ONC+ Developer’s Guide
Contents

Preface xv

Part I Introduction

1. Introduction to ONC+ Technologies 3
   Brief Description of ONC+ Technologies 4
      TI-RPC 4
      XDR 4
      NFS 4
      NIS+ 5

Part II Remote Procedure Call

2. Introduction to TI-RPC 9
   What Is TI-RPC 9
   TI-RPC Issues 10
      Parameter Passing 11
      Binding 11
      Transport Protocol 11
      Call Semantics 11
      Data Representation 12
   Program, Version, and Procedure Numbers 12
   Overview of Interface Routines 12
      Simplified Interface Routines 13
      Standard Interface Routines 13
   Network Selection 16
Network Types/Transport Selection 50
Command Line Define Statements 50
Server Response to Broadcast Calls 51
Port Monitor Support 51
Time-out Changes 52
Client Authentication 53
Dispatch Tables 53
Debugging Applications 55

4. The Programmer's Interface to RPC 57
RPC Is Multithread Safe 57
Simplified Interface 57
Client 58
Server 60
Hand-Coded Registration Routine 61
Passing Arbitrary Data Types 62
Standard Interfaces 65
Top Level Interface 66
Intermediate Level Interface 70
Expert Level Interface 73
Bottom Level Interface 77
Server Caching 78
Low-Level Data Structures 79
Testing Programs Using Low-level Raw RPC 81
Advanced RPC Programming Techniques 84
poll on the Server Side 84
Broadcast RPC 86
Batching 88
Authentication 91
<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>Type Definitions</td>
<td>217</td>
</tr>
<tr>
<td>Declarations</td>
<td>217</td>
</tr>
<tr>
<td>Simple Declarations</td>
<td>218</td>
</tr>
<tr>
<td>Fixed-Length Array Declarations</td>
<td>218</td>
</tr>
<tr>
<td>Variable-Length Array Declarations</td>
<td>218</td>
</tr>
<tr>
<td>Pointer Declarations</td>
<td>219</td>
</tr>
<tr>
<td>Structures</td>
<td>219</td>
</tr>
<tr>
<td>Unions</td>
<td>220</td>
</tr>
<tr>
<td>Programs</td>
<td>220</td>
</tr>
<tr>
<td>Special Cases</td>
<td>221</td>
</tr>
<tr>
<td>rpcbind Protocol</td>
<td>223</td>
</tr>
<tr>
<td>rpcbind Operation</td>
<td>228</td>
</tr>
<tr>
<td>Version 4 rpcbind</td>
<td>230</td>
</tr>
<tr>
<td>Bibliography</td>
<td>231</td>
</tr>
</tbody>
</table>

C. **XDR Protocol Specification** 233

<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>XDR Protocol Introduction</td>
<td>233</td>
</tr>
<tr>
<td>Graphic Box Notation</td>
<td>234</td>
</tr>
<tr>
<td>Basic Block Size</td>
<td>234</td>
</tr>
<tr>
<td>XDR Data Type Declarations</td>
<td>234</td>
</tr>
<tr>
<td>Signed Integer</td>
<td>235</td>
</tr>
<tr>
<td>Unsigned Integer</td>
<td>235</td>
</tr>
<tr>
<td>Enumerations</td>
<td>236</td>
</tr>
<tr>
<td>Booleans</td>
<td>237</td>
</tr>
<tr>
<td>Hyper Integer and Unsigned Hyper Integer</td>
<td>237</td>
</tr>
<tr>
<td>Floating Point</td>
<td>238</td>
</tr>
<tr>
<td>Quadruple-Precision Floating Point</td>
<td>239</td>
</tr>
<tr>
<td>Fixed-Length Opaque Data</td>
<td>240</td>
</tr>
<tr>
<td>Variable-Length Opaque Data</td>
<td>240</td>
</tr>
</tbody>
</table>
Counted Byte Strings 241
Fixed-Length Array 242
Variable-Length Array 243
Structure 244
Discriminated Union 244
Void 245
Constant 245
Typedef 246
Optional-Data 246
The XDR Language Specification 247
Notational Conventions 247
Lexical Notes 247
Syntax Notes 249
XDR Data Description 250
RPC Language Reference 251
Enumerations 252
Constants 252
Type Definitions 252
Declarations 253
Simple Declarations 253
Fixed-Length Array Declarations 253
Variable-Length Array Declarations 254
Pointer Declarations 254
Structures 254
Unions 255
Programs 256
Special Cases 257

D. Live RPC Code Examples 259

x ONC+ Developer's Guide ♦ August, 1997
/etc/saf/pmtag/_pmtab 307
/etc/saf/pmtag/svctag 307
/etc/saf/pmtag/_pid 307
/etc/saf/pmtag/_pmpipe 307
/var/saf/_log 307
/var/saf/pmtag 307

Glossary 309

Index 311
The ONC+ Developer's Guide describes the programming interfaces to remote procedure call (RPC) and NIS+, a network name service, which belong to the ONC+ distributed services developed at SunSoft, Inc.

In this guide, the terms SunOS and Solaris are used interchangeably because the interfaces described in this manual are common to both. Solaris 2.5 is SunSoft's distributed computing operating environment. It is comprised of SunOS release 5.5 with the ONC+ technologies, OpenWindows, ToolTalk, DeskSet, and OPEN LOOK as well as other utilities.

All utilities, their options, and library functions in this manual reflect the current Solaris system software developed by SunSoft Inc. If you are using a previous version of Solaris system software, some utilities and library functions may function differently.

Who Should Use This Book

The guide assists you in converting an existing single-computer application to a networked, distributed application, or developing and implementing distributed applications.

Use of this guide assumes basic competence in programming, a working familiarity with the C programming language, and a working familiarity with theUNIX operating system. Previous experience in network programming is helpful, but is not required to use this manual.
How This Book Is Organized

Part One—Introduction
Chapter 1, gives a high-level introduction to the ONC+ distributed computing platform and services.

Part Two—Remote Procedure Call (RPC)
Chapter 2, introduces TI-RPC.
Chapter 3, describes how the \texttt{rpcgen} tool generates client and server stubs.
Chapter 4, describes the use of RPC in the programming environment.

Part Three—NIS+ Applications Programming Interface
Chapter 5, describes the NIS+ applications programming interface.

Appendixes
Appendix A, describes XDR and how it is used in data formatting and type conversion.
Appendix B, describes the protocol of RPC usage, both syntax and limitations.
Appendix C, describes the XDR protocol and language.
Appendix D, contains complete functional listings of some of the code included in the document as examples.
Appendix E, describes the portmap utility and its function. This appendix is included in this document to aid migrating applications written to run on earlier releases of SunOS.
Appendix F, describes the process of writing a port monitor application under the SAF and is included as a reference for applications development.
Related Books

The following on-line System AnswerBook® products cover related network programming topics:

- Solaris 2.5 Reference Manual AnswerBook
- Solaris 2.6 Software Developer AnswerBook Vol 1

For information on SunSoft’s NFS® distributed computing file system, see the following sources:


The following third-party books and articles are excellent sources on network programming topics:

What Typographic Changes and Symbols Mean

The following table describes the typographic changes used in this book.

**TABLE P–1**  Typographic Conventions

<table>
<thead>
<tr>
<th>Typeface or Symbol</th>
<th>Meaning</th>
<th>Example</th>
</tr>
</thead>
<tbody>
<tr>
<td>AaBbCc123</td>
<td>The names of commands, files, and directories; on-screen computer output</td>
<td>Edit your <code>.login</code> file.</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Use <code>ls -a</code> to list all files.</td>
</tr>
<tr>
<td></td>
<td></td>
<td><code>machine_name%</code> <em>You have mail.</em></td>
</tr>
<tr>
<td>AaBbCc123</td>
<td>What you type, contrasted with on-screen computer output</td>
<td><code>machine_name% su</code></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Password:</td>
</tr>
<tr>
<td>AaBbCc123</td>
<td>Command-line placeholder: replace with a real name or value</td>
<td>To delete a file, type <code>rm filename</code>.</td>
</tr>
<tr>
<td>AaBbCc123</td>
<td>Book titles, new words or terms, or words to be emphasized</td>
<td><code>Read Chapter 6 in User's Guide. These are called class options.</code></td>
</tr>
<tr>
<td></td>
<td></td>
<td><code>You must be root to do this.</code></td>
</tr>
</tbody>
</table>

Shell Prompts in Command Examples

The following table shows the default system prompt and superuser prompt for the C shell, Bourne shell, and Korn shell.
<table>
<thead>
<tr>
<th>Shell</th>
<th>Prompt</th>
</tr>
</thead>
<tbody>
<tr>
<td>C shell prompt</td>
<td>machine_name%</td>
</tr>
<tr>
<td>C shell superuser prompt</td>
<td>machine_name#</td>
</tr>
<tr>
<td>Bourne shell and Korn shell prompt</td>
<td>$</td>
</tr>
<tr>
<td>Bourne shell and Korn shell superuser</td>
<td>#</td>
</tr>
</tbody>
</table>
Part 1 is an introduction to the ONC+ services.

- Chapter 1
Introduction to ONC+ Technologies

This chapter is a short introduction to ONC+ technologies, Sun’s open systems distributed computing environment. The ONC+ technologies are the core services available to developers who implement distributed applications in a heterogeneous distributed computing environment. ONC+ technologies also includes tools to administer client/server networks.

Figure 1–1 shows an integrated view of how client-server applications are built on top of ONC+ technologies, and how they sit on top of the low-level networking protocols:

![Figure 1–1 ONC+ Distributed Computing Platform](image-url)
Brief Description of ONC+Technologies

ONC+ technologies are comprised of a family of technologies, services, and tools. It is backward compatible and interoperates with the installed base of ONC services. The main components are described. This guide covers the technologies that require the use of programming facilities.

TI-RPC

Transport-independent remote procedure call (TI-RPC) was developed by Sun and AT&T as part of the UNIX System V Release 4 (SVR4). It makes RPC applications transport-independent by allowing a single binary version of a distributed program to run on multiple transports. Previously, with transport-specific RPC, the transport was bound at compile time so that applications could not use other transports unless the program was rebuilt. With TI-RPC, applications can use new transports if the system administrator updates the network configuration file and restarts the program. Thus, no changes are required to the binary application.

XDR

External data representation (XDR) is an architecture-independent specification for representing data. It resolves the differences in data byte ordering, data type size, representation, and alignment between different architectures. Applications that use XDR may exchange data across heterogeneous hardware systems.

NFS

NFS is Sun’s distributed computing file system that provides transparent access to remote filesystems on heterogeneous networks. In this way, users can share files among PCs, workstations, mainframes, and supercomputers. As long as they are connected to the same network, the files appear as though they are on the user’s desktop. The NFS environment features Kerberos authentication, multithreading, the network lock manager, and the automounter.
NFS does not have programming facilities, so it is not covered in this guide. However, the specification for NFS V.3 is available through anonymous ftp. See “Related Books” on page xvii for more information.

**NIS+**

NIS+ is the enterprise naming service in Solaris. It provides a scalable and secure information base for host names, network addresses, and user names. It is designed to make administration of large, multivendor client/server networks easier by being the central point for adding, removing, and relocating network resources. Changes made to the NIS+ information base are automatically and immediately propagated to replica servers across the network; this ensures that system uptime and performance is preserved. Security is integral to NIS+. Unauthorized users and programs are prevented from reading, changing, or destroying naming service information.
PART II    Remote Procedure Call

Part 2 covers RPC topics
- Chapter 2
- Chapter 3
- Chapter 4
Introduction to TI-RPC

This chapter provides an overview of TI-RPC, also known as Sun RPC. The information presented is most useful to someone new to RPC. (See also Glossary for the definition of the terms used in this guide.)

- “What Is TI-RPC” on page 9
- “TI-RPC Issues” on page 10
- “Overview of Interface Routines” on page 12
- “Network Selection” on page 16
- “Transport Selection” on page 17
- “Address Lookup Services” on page 19

What Is TI-RPC

TI-RPC is a powerful technique for constructing distributed, client-server based applications. It is based on extending the notion of conventional, or local procedure calling, so that the called procedure need not exist in the same address space as the calling procedure. The two processes may be on the same system, or they may be on different systems with a network connecting them.

By using RPC, programmers of distributed applications avoid the details of the interface with the network. The transport independence of RPC isolates the application from the physical and logical elements of the data communications mechanism and allows the application to use a variety of transports.
Figure 2–1  How RPC Works

An RPC is analogous to a function call. Like a function call, when an RPC is made, the calling arguments are passed to the remote procedure and the caller waits for a response to be returned from the remote procedure.

Figure 2–1 shows the flow of activity that takes place during an RPC call between two networked systems. The client makes a procedure call that sends a request to the server and waits. The thread is blocked from processing until either a reply is received, or it times out. When the request arrives, the server calls a dispatch routine that performs the requested service, and sends the reply to the client. After the RPC call is completed, the client program continues.

RPC specifically supports network applications. TI- RPC runs on available networking mechanisms such as TCP/IP. Other RPC standards are OSF DCE (based on Apollo’s NCS system), Xerox Courier, and Netwise.

TI-RPC Issues

There are a number of issues that help to characterize a particular RPC implementation.
Parameter Passing
TI-RPC allows a single parameter to be passed from client to server. If more than one parameter is required, the components can be combined into a structure which is counted as a single element. Information passed from server to client is passed as the function’s return value. It is not possible to pass information back from server to client through the parameter list.

Binding
The client must know how to contact the service it is interested in. The two aspects are finding out which host the server is on, and then connecting to the actual server process. On each host, a service called rpcbind manages RPC services. TI-RPC uses the available host-naming services, such as the hosts file, NIS+, and DNS, to locate a host.

Transport Protocol
The transport protocol specifies how the call message and the reply message are transmitted between client and server. TS-RPC used TCP and UDP as transport protocols, but the current version of TI-RPC is transport independent; that is, it works with any transport protocol supported by Solaris 2.x.

Call Semantics
Call semantics has to do with what the client can assume about the execution of the remote procedure, in particular, how many times the procedure was executed. This is important in dealing with error conditions. The three alternatives are exactly once, at most once, and at least once. ONC+ provides at least once semantics. Procedures called remotely are idempotent: they should return the same result each time they are called, even if it is several times.
Data Representation

Data representation describes the format used for parameters and results as they are passed between processes. For RPC to function on a variety of system architectures requires a standard data representation. TI-RPC uses external data representation (XDR). XDR is a machine-independent data description and encoding protocol. Using XDR, RPC can handle arbitrary data structures, regardless of different hosts’ byte orders or structure layout conventions. For a detailed discussion of XDR, see Appendix C, and Appendix A.”

Program, Version, and Procedure Numbers

A remote procedure is uniquely identified by the triple:

\[(\text{program number}, \text{version number}, \text{procedure number})\]

The program number identifies a group of related remote procedures, each of which has a unique procedure number.

A program may consist of one or more versions. Each version consists of a collection of procedures which are available to be called remotely. Version numbers enable multiple versions of an RPC protocol to be available simultaneously.

Each version contains a a number of procedures that can be called remotely. Each procedure has a procedure number.

“Program and Procedure Numbers” on page 196, lists the range of values and their significance and tells you how to have a program number assigned to your RPC program. A list of mappings of RPC service name to program number is available in the rpc network database, /etc/rpc.

Overview of Interface Routines

RPC has multiple levels of application interface to its services. These levels provide different degrees of control balanced with different amounts of interface code to implement. In order of increasing control and complexity. This section gives a summary of the routines available at each level.
Simplified Interface Routines

The simplified interfaces are used to make remote procedure calls to routines on other machines, and specify only the type of transport to use. The routines at this level are used for most applications. Descriptions and code samples can be found in the section, “Simplified Interface” on page 57.

**TABLE 2-1**  RPC Routines—Simplified Level

<table>
<thead>
<tr>
<th>Routine</th>
<th>Function</th>
</tr>
</thead>
<tbody>
<tr>
<td>rpc_reg</td>
<td>Registers a procedure as an RPC program on all transports of the specified type.</td>
</tr>
<tr>
<td>rpc_call</td>
<td>Remote calls the specified procedure on the specified remote host.</td>
</tr>
<tr>
<td>rpc_broadcast</td>
<td>Broadcasts a call message across all transports of the specified type.</td>
</tr>
</tbody>
</table>

Standard Interface Routines

The standard interfaces are divided into top level, intermediate level, expert level, and bottom level. These interfaces give a developer much greater control over communication parameters such as the transport being used, how long to wait before responding to errors and retransmitting requests, and so on.

Top Level Routines

At the top level, the interface is still simple, but the program has to create a client handle before making a call or create a server handle before receiving calls. If you want the application to run on all transports, use this interface. Use of these routines and code samples can be found in “Top Level Interface” on page 66.
TABLE 2–2  RPC Routines—Top Level

<table>
<thead>
<tr>
<th>Routine</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>clnt_create</td>
<td>Generic client creation. The program tells clnt_create where the server is located and the type of transport to use.</td>
</tr>
<tr>
<td>clnt_create_timed</td>
<td>Similar to clnt_create but lets the programmer specify the maximum time allowed for each type of transport tried during the creation attempt.</td>
</tr>
<tr>
<td>svc_create</td>
<td>Creates server handles for all transports of the specified type. The program tells svc_create which dispatch function to use.</td>
</tr>
<tr>
<td>clnt_call()</td>
<td>Client calls a procedure to send a request to the server.</td>
</tr>
</tbody>
</table>

Intermediate Level Routines

The intermediate level interface of RPC lets you control details. Programs written at these lower levels are more complicated but run more efficiently. The intermediate level enables you to specify the transport to use. Use of these routines and code samples can be found in “Intermediate Level Interface” on page 70.

TABLE 2–3  RPC Routines—Intermediate Level

<table>
<thead>
<tr>
<th>Routine</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>clnt_tp_create</td>
<td>Creates a client handle for the specified transport.</td>
</tr>
<tr>
<td>clnt_tp_create_timed</td>
<td>Similar to clnt_tp_create but lets the programmer specify the maximum time allowed.</td>
</tr>
<tr>
<td>svc_tp_create</td>
<td>Creates a server handle for the specified transport.</td>
</tr>
<tr>
<td>clnt_call()</td>
<td>Client calls a procedure to send a request to the server.</td>
</tr>
</tbody>
</table>
Expert Level Routines

The expert level contains a larger set of routines with which to specify transport-related parameters. Use of these routines and code samples can be found in “Expert Level Interface” on page 73.

<table>
<thead>
<tr>
<th>Routine</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>clnt_tli_create</td>
<td>Creates a client handle for the specified transport.</td>
</tr>
<tr>
<td>svc_tli_create</td>
<td>Creates a server handle for the specified transport.</td>
</tr>
<tr>
<td>rpcb_set</td>
<td>Calls rpcbind to set a map between an RPC service and a network address.</td>
</tr>
<tr>
<td>rpcb_unset</td>
<td>Deletes a mapping set by rpcb_set.</td>
</tr>
<tr>
<td>rpcb_getaddr</td>
<td>Calls rpcbind to get the transport addresses of specified RPC services.</td>
</tr>
<tr>
<td>svc_reg</td>
<td>Associates the specified program and version number pair with the specified dispatch routine.</td>
</tr>
<tr>
<td>svc_unreg</td>
<td>Deletes an association set by svc_reg.</td>
</tr>
<tr>
<td>clnt_call()</td>
<td>Client calls a procedure to send a request to the server.</td>
</tr>
</tbody>
</table>

Bottom Level Routines

The bottom level contains routines used for full control of transport options. “Bottom Level Interface” on page 77, describes these routines.
TABLE 2–5  RPC Routines—Bottom Level

<table>
<thead>
<tr>
<th>Routine</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>clnt_dg_create</td>
<td>Creates an RPC client handle for the specified remote program, using a connectionless transport.</td>
</tr>
<tr>
<td>svc_dg_create</td>
<td>Creates an RPC server handle, using a connectionless transport.</td>
</tr>
<tr>
<td>clnt_vc_create</td>
<td>Creates an RPC client handle for the specified remote program, using a connection-oriented transport.</td>
</tr>
<tr>
<td>svc_vc_create</td>
<td>Creates an RPC server handle, using a connection-oriented transport.</td>
</tr>
<tr>
<td>clnt_call()</td>
<td>Client calls a procedure to send a request to the server.</td>
</tr>
</tbody>
</table>

Network Selection

You can write programs to run on a specific transport or transport type, or to operate on a system- or user-chosen transport. There are two mechanisms for network selection, the /etc/netconfig database and the environmental variable NETPATH. These mechanisms allow a fine degree of control over network selection: a user can specify a preferred transport, and if it can, an application uses it. If the specified transport is inappropriate, the application automatically tries others with the right characteristics.

/etc/netconfig lists the transports available to the host and identifies them by type. NETPATH is optional and allows a user to specify a transport or selection of transports from the list in /etc/netconfig. By setting the NETPATH, the user specifies the order in which the application tries the available transports. If NETPATH is not set, the system defaults to all visible transports specified in /etc/netconfig, in the order they appear in that file.

For more details on network selection, refer to the Transport Interfaces Programming Guide or see the getnetconfig(3N) and netconfig(4) manpages.

RPC divides selectable transports into the following types:
### TABLE 2-6  nettype Parameters

<table>
<thead>
<tr>
<th>Value</th>
<th>Meaning</th>
</tr>
</thead>
<tbody>
<tr>
<td>NULL</td>
<td>Same as selecting netpath.</td>
</tr>
<tr>
<td>visible</td>
<td>Uses the transports chosen with the visible flag ('v') set in their</td>
</tr>
<tr>
<td></td>
<td>/etc/netconfig entries.</td>
</tr>
<tr>
<td>circuit_v</td>
<td>Same as visible, but restricted to connection-oriented</td>
</tr>
<tr>
<td></td>
<td>transports. Transports are selected in the order listed in /etc/</td>
</tr>
<tr>
<td></td>
<td>netconfig.</td>
</tr>
<tr>
<td>datagram_v</td>
<td>Same as visible, but restricted to connectionless transports.</td>
</tr>
<tr>
<td>circuit_n</td>
<td>Uses the connection-oriented transports chosen in the order defined in</td>
</tr>
<tr>
<td></td>
<td>NETPATH.</td>
</tr>
<tr>
<td>datagram_n</td>
<td>Uses the connectionless transports chosen in the order defined in</td>
</tr>
<tr>
<td></td>
<td>NETPATH.</td>
</tr>
<tr>
<td>udp</td>
<td>Specifies Internet user datagram protocol (UDP).</td>
</tr>
<tr>
<td>tcp</td>
<td>Specifies Internet transport control protocol (TCP).</td>
</tr>
</tbody>
</table>

---

**Transport Selection**

RPC services are supported on both circuit-oriented and datagram transports. The selection of the transport depends on the requirements of the application.

A datagram transport is the transport of choice if the application has all of the following characteristics:

- Calls to the procedures do not change the state of the procedure or of associated data.
- The size of both the arguments and results is smaller than the transport packet size.
- The server is required to handle hundreds of clients. A datagram server does not keep any state data on clients, so it can potentially handle many clients. A circuit-oriented server keeps state data on each open client connection, so the number of clients is limited by the host resources.

---

*Introduction to TI-RPC  17*
A circuit-oriented transport is the transport of choice if the application has any of the following characteristics:

- The application can tolerate or amortize the higher cost of connection setup compared to datagram transports.
- Calls to the procedures can change the state of the procedure or of associated data.
- The size of either the arguments or the results exceed the maximum size of a datagram packet.

Name-to-Address Translation

Each transport has an associated set of routines that translate between universal network addresses (string representations of transport addresses) and the local address representation. These universal addresses are passed around within the RPC system (for example, between rpcbind and a client). A run-time linkable library that contains the name-to-address translation routines is associated with each transport. Table 2–7 shows the main translation routines.

For more details on these routines, see the netdir(3N) manpage and the Transport Interfaces Programming Guide. Note that the netconfig structure in each case provides the context for name-to-address translations.

<table>
<thead>
<tr>
<th>Table 2–7 Name-to-Address Translation Routines</th>
</tr>
</thead>
<tbody>
<tr>
<td>netdir_getbyname</td>
</tr>
<tr>
<td>uaddr2taddr</td>
</tr>
<tr>
<td>taddr2uaddr</td>
</tr>
</tbody>
</table>
Address Lookup Services

Transport services do not provide address-lookup services. They provide only message transfer across a network. A client program needs a way to obtain the address of its server program. In earlier system releases this service was performed by portmap. rpcbind replaces the portmap utility.

RPC makes no assumption about the structure of a network address. It deals with universal addresses specified only as null-terminated strings of ASCII characters. RPC translates universal addresses into local transport addresses by using routines specific to the transport. For more details on these routines, see the netdir(3N) and rpcbind(3N) manpages.

rpcbind provides the operations:

- Add a registration
- Delete a registration
- Get address of a specified program number, version number, and transport
- Get the complete registration list
- Perform a remote call for a client
- Return the time

Registering Addresses

rpcbind maps RPC services to their addresses, so its address must be known. The name-to-address translation routines must reserve a known address for each type of transport used. For example, in the Internet domain, rpcbind has port number 111 on both TCP and UDP. When rpcbind is started, it registers its location on each of the transports supported by the host. rpcbind is the only RPC service that must have a known address.

For each supported transport, rpcbind registers the addresses of RPC services and makes the addresses available to clients. A service makes its address available to clients by registering the address with the rpcbind daemon. The address of the service is then available to rpcinfo(1M) and to programs using library routines named in the rpcbind(3N) manpage. No client or server can assume the network address of an RPC service.

Client and server programs and client and server hosts are usually distinct but they need not be. A server program can also be a client program. When one server calls another rpcbind server it makes the call as a client.

To find a remote program’s address, a client sends an RPC message to a host's rpcbind daemon. If the service is on the host, the daemon returns the
address in an RPC reply message. The client program can then send RPC messages to the server’s address. (A client program can minimize its calls to rpcbind by storing the network addresses of recently called remote programs.)

The RPCBPROC_CALLIT procedure of rpcbind lets a client make a remote procedure call without knowing the address of the server. The client passes the target procedure’s program number, version number, procedure number, and calling arguments in an RPC call message. rpcbind looks up the target procedure’s address in the address map and sends an RPC call message, including the arguments received from the client, to the target procedure.

When the target procedure returns results, RPCBPROC_CALLIT passes them to the client program. It also returns the target procedure’s universal address so that the client can later call it directly.

The RPC library provides an interface to all rpcbind procedures. Some of the RPC library procedures also call rpcbind automatically for client and server programs. For details, see Appendix B.

**Reporting RPC Information**

rpcinfo is a utility that reports current RPC information registered with rpcbind. rpcinfo (with either rpcbind or the portmap utility) reports the universal addresses and the transports for all registered RPC services on a specified host. It can call a specific version of a specific program on a specific host and report whether a response is received. It can also delete registrations. For details, see the rpcinfo(1M) manpage.
This chapter introduces the `rpcgen` tool and provides a tutorial with code examples and usage of the available compile-time flags. See Glossary for the definition of the terms used in this chapter.

- “SunOS 5.x Features” on page 22
- “An rpcgen Tutorial” on page 23
- “Compile-Time Flags” on page 36
- “rpcgen Programming Techniques” on page 49

What is `rpcgen`?

The `rpcgen` tool generates remote program interface modules. It compiles source code written in the RPC Language. RPC Language is similar in syntax and structure to C. `rpcgen` produces one or more C language source modules, which are then compiled by a C compiler.

The default output of `rpcgen` is:

- A header file of definitions common to the server and the client
- A set of XDR routines that translate each data type defined in the header file
- A stub program for the server
- A stub program for the client

`rpcgen` can optionally generate:

- Various transports
- A time-out for servers
Server stubs that are MT safe
Server stubs that are not main programs
C-style arguments passing ANSI C-compliant code
An RPC dispatch table that checks authorizations and invokes service routines

rpcgen significantly reduces the development time that would otherwise be spent developing low-level routines. Handwritten routines link easily with the rpcgen output. (For a discussion of RPC programming without rpcgen, see Chapter 4.)

SunOS 5.x Features

This section lists the features found in the SunOS 5.x rpcgen code generator that are not found in the SunOS 4.x version.

Template Generation

rpcgen generates client-side, server-side, and makefile templates. See “Client and Server Templates” on page 37 for the list of options.

C-style Mode

rpcgen has two compilation modes, C-style and default. C-style mode lets arguments be passed by value, instead of as pointers to a structure. It also supports passing multiple arguments. The default mode is the same as in previous releases. See “C-style Mode” on page 38 for the example code for both modes.

Multithread-Safe Code

rpcgen can now generate MT-safe code for use in a threaded environment. By default, the code generated by rpcgen is not MT-safe. See “MT-Safe Code” on page 41 for the description and example code.
Multithread Auto Mode

`rpcgen` can generate MT-safe server stubs that operate in the MT Auto mode. See “MT Auto Mode” on page 47 for the definition and example code.

Library Selection

`rpcgen` can use library calls for either TS-RPC or TI-RPC. See “TI-RPC or TS-RPC Library Selection” on page 48.

ANSI C -compliant Code

The output generated by `rpcgen` conforms to ANSI C standards. The code can also be used in the SPARCompiler™ C++ 3.0 environment. See “ ANSI C-compliant Code” on page 48.

An `rpcgen` Tutorial

`rpcgen` provides programmers a simple and direct way to write distributed applications. Server procedures may be written in any language that observes procedure-calling conventions. They are linked with the server stub produced by `rpcgen` to form an executable server program. Client procedures are written and linked in the same way. This section presents some basic `rpcgen` programming examples. Refer also to the `rpcgen`(1) manpage.

Converting Local Procedures to Remote Procedures

Assume that an application runs on a single computer and you want to convert it to run in a “distributed” manner on a network. This example shows the stepwise conversion of this program that writes a message to the system console. Code Example 3–1 shows the original program.

**CODE EXAMPLE 3–1** Single Process Version of printmsg.c

```c
/* printmsg.c: print a message on the console */
#include <stdio.h>
```
main(argc, argv)
int argc;
char *argv[];
{
    char *message;
    if (argc != 2) {
        fprintf(stderr, "usage: %s <message>\n", argv[0]);
        exit(1);
    }
    message = argv[1];
    if (!printmessage(message)) {
        fprintf(stderr, "%s: couldn’t print your message\n", argv[0]);
        exit(1);
    }
    printf("Message Delivered!\n");
    exit(0);
}

/* Print a message to the console.
 * Return a boolean indicating whether
 * the message was actually printed. */

printmessage(msg)
char *msg;
{
    FILE *f;
    f = fopen("/dev/console", "w");
    if (f == (FILE *)NULL) {
        return (0);
    }
    fprintf(f, "%s\n", msg);
    fclose(f);
    return(1);
}

For local use on a single machine, this program could be compiled and executed as follows:

$ cc printmsg.c -o printmsg
$ printmsg "Hello, there."
Message delivered!
$
If the printmessage function is turned into a remote procedure, it can be called from anywhere in the network. rpcgen makes it easy to do this.

First, determine the data types of all procedure-calling arguments and the result argument. The calling argument of printmessage is a string, and the result is an integer. We can write a protocol specification in RPC language that describes the remote version of printmessage. The RPC language source code for such a specification is:

```c
/* msg.x: Remote msg printing protocol */
program MESSAGEPROG {
    version PRINTMESSAGEVERS {
        int PRINTMESSAGE(string) = 1;
    } = 1;
    } = 0x20000001;
}
```

Remote procedures are always declared as part of remote programs. The code above declares an entire remote program that contains the single procedure PRINTMESSAGE. In this example, the PRINTMESSAGE procedure is declared to be procedure 1, in version 1 of the remote program MESSAGEPROG, with the program number 0x20000001. (See Appendix B for guidance on choosing program numbers.) Version numbers are incremented when functionality is changed in the remote program. Existing procedures can be changed or new ones can be added. More than one version of a remote program can be defined and a version can have more than one procedure defined.

Note that the program and procedure names are declared with all capital letters. This is not required, but is a good convention to follow.

Note also that the argument type is string and not char * as it would be in C. This is because a char * in C is ambiguous. char usually means an array of characters, but it could also represent a pointer to a single character. In RPC language, a null-terminated array of char is called a string.

There are just two more programs to write:

- The remote procedure itself
- The main client program that calls it

Code Example 3-2 is a remote procedure that implements the PRINTMESSAGE procedure in Code Example 3-1.

**CODE EXAMPLE 3-2**  RPC Version of printmsg.c

```c
/* 
 * msg_proc.c: implementation of the 
 * remote procedure "printmessage"
 */
#include <stdio.h>
#include "msg.h"  /* msg.h generated by rpcgen */

(Continued)
```
```c
int *
printmessage_1(msg, req)
    char **msg;
    struct svc_req *req; /* details of call */
{ static int result; /* must be static! */
    FILE *f;

    f = fopen("/dev/console", "w");
    if (f == (FILE *)NULL) {
        result = 0;
        return (&result);
    }
    fprintf(f, "%s\n", *msg);
    fclose(f);
    result = 1;
    return (&result);
}
```

Note that the declaration of the remote procedure `printmessage_1` differs from that of the local procedure `printmessage` in four ways:

1. It takes a pointer to the character array instead of the pointer itself. This is true of all remote procedures when the `−N` option is not used: They always take pointers to their arguments rather than the arguments themselves. Without the `−N` option, remote procedures are always called with a single argument. If more than one argument is required the arguments must be passed in a `struct`.

2. It is called with two arguments. The second argument contains information on the context of an invocation: the program, version, and procedure numbers, raw and canonical credentials, and an `SVCXPRT` structure pointer (the `SVCXPRT` structure contains transport information). This information is made available in case the invoked procedure requires it to perform the request.

3. It returns a pointer to an integer instead of the integer itself. This is also true of remote procedures when the `−N` option is not used: They return pointers to the result. The result should be declared `static unless` the `−M` (multithread) or `−A` (Auto mode) options are used. Ordinarily, if the result is declared local to the remote procedure, references to it by the server stub are invalid after the remote procedure returns. In the case of `−M` and `−A` options, a pointer to the result is passed as a third argument to the procedure, so the result is not declared in the procedure.
4. An _1 is appended to its name. In general, all remote procedures calls generated by rpcgen are named as follows: the procedure name in the program definition (here PRINTMESSAGE) is converted to all lowercase letters, an underbar (_) is appended to it, and the version number (here 1) is appended. This naming scheme allows multiple versions of the same procedure.

Code Example 3–3 shows the main client program that calls the remote procedure.

**CODE EXAMPLE 3–3**  
Client Program to Call printmsg.c

```c
/*
* rprintmsg.c: remote version
* of "printmsg.c"
*/
#include <stdio.h>
#include "msg.h" /* msg.h generated by rpcgen */

main(argc, argv)
int argc;
char *argv[];
{
    CLIENT *clnt;
    int *result;
    char *server;
    char *message;
    if (argc != 3) {
        fprintf(stderr, "usage: %s host
message\n", argv[0]);
        exit(1);
    }
    server = argv[1];
    message = argv[2];
    /*
    * Create client "handle" used for
    * calling MESSAGEPROG on the server
    * designated on the command line.
    */
    clnt = clnt_create(server, MESSAGEPROG,
PRINTMESSAGEVERS,
    "visible");
    if (clnt == (CLIENT *)NULL) {
        /*
        * Couldn’t establish connection
        * with server.
        * Print error message and die.
        */
        clnt_pcreateerror(server);
        exit(1);
    }
    /*
(Continued)
```
* Call the remote procedure
* "printmessage" on the server
*/
result = printmessage_1(&message, clnt);
if (result == (int *)NULL) {
  /*
   * An error occurred while calling
   * the server.
   * Print error message and die.
   */
  clnt_perror(clnt, server);
  exit(1);
}
/* Okay, we successfully called
 * the remote procedure.
*/
if (*result == 0) {
  /*
   * Server was unable to print
   * our message.
   * Print error message and die.
   */
  fprintf(stderr,
"%s: could not print your message\n",argv[0]);
  exit(1);
}
/* The message got printed on the
 * server's console
*/
printf("Message delivered to %s\n", server);
clnt_destroy( clnt );
exit(0);
}

Note the following about Code Example 3–3:

1. First, a client handle is created by the RPC library routine clnt_create. This client handle is passed to the stub routine that calls the remote procedure. (The client handle can be created in other ways as well. See Chapter 4 for details.) If no more calls are to be made using the client handle, destroy it with a call to clnt_destroy to conserve system resources.

2. The last parameter to clnt_create is visible, which specifies that any transport noted as visible in /etc/netconfig can be used. For further information on this, see the /etc/netconfig file and its description in Transport Interfaces Programming Guide.
3. The remote procedure `printmessage_1` is called exactly the same way as it is declared in `msg_proc.c`, except for the inserted client handle as the second argument. It also returns a pointer to the result instead of the result.

4. The remote procedure call can fail in two ways. The RPC mechanism can fail or there can be an error in the execution of the remote procedure. In the former case, the remote procedure `printmessage_1` returns a NULL. In the latter case, the error reporting is application dependent. Here, the error is returned through `*result`.

Here are the compile commands for the printmsg example:

```bash
$ rpcgen msg.x
$ cc rprintmsg.c msg_clnt.c -o rprintmsg -lnsl
$ cc msg_proc.c msg_svc.c -o msg_server -lnsl
```

First, `rpcgen` was used to generate the header files (`msg.h`), client stub (`msg_clnt.c`), and server stub (`msg_svc.c`). Then, two programs are compiled: the client program `rprintmsg` and the server program `msg_server`. The C object files must be linked with the library `libnsl`, which contains all of the networking functions, including those for RPC and XDR.

In this example, no XDR routines were generated because the application uses only the basic types that are included in `libnsl`.

Here is what `rpcgen` did with the input file `msg.x`:

1. It created a header file called `msg.h` that contained `#define` statements for `MESSAGEPROG`, `MESSAGEVERS`, and `PRINTMESSAGE` for use in the other modules. This file must be included by both the client and server modules.
2. It created the client stub routines in the `msg_clnt.c` file. Here there is only one, the `printmessage_1` routine, that was called from the `rprintmsg` client program. If the name of an `rpcgen` input file is `FOO.x`, the client stub’s output file is called `FOO_clnt.c`.
3. It created the server program in `msg_svc.c` that calls `printmessage_1` from `msg_proc.c`. The rule for naming the server output file is similar to that of the client: for an input file called `FOO.x`, the output server file is named `FOO_svc.c`.

Once created, the server program is installed on a remote machine and run. (If the machines are homogeneous, the server binary can just be copied. If they are not, the server source files must be copied to and compiled on the remote machine.) For this example, the remote machine is called `remote` and the local machine is called `local`. The server is started from the shell on the remote system:

```
remote$ msg_server
```
Server processes generated with `rpcgen` always run in the background. It is not necessary to follow the server’s invocation with an ampersand (&). Servers generated by `rpcgen` can also be invoked by port monitors like `listen` and `inetd`, instead of from the command line.

Thereafter, a user on `local` can print a message on the console of machine `remote` as follows:

```
local$ rprintmsg remote "Hello, there."
```

Using `rprintmsg`, a user can print a message on any system console (including the `local` console) when the server `msg_server` is running on the target system.

**Passing Complex Data Structures**

“Converting Local Procedures to Remote Procedures” on page 23 shows how to generate client and server RPC code. `rpcgen` can also be used to generate XDR routines (the routines that convert local data structures into XDR format and vice versa).

Code Example 3–4 presents a complete RPC service: a remote directory listing service, built using `rpcgen` both to generate stub routines and to generate the XDR routines.

**CODE EXAMPLE 3–4 RPC Protocol Description File: dir.x**

```c
/*
 * dir.x: Remote directory listing protocol
 *
 * This example demonstrates the functions of rpcgen.
 */

const MAXNAMELEN = 255; /* max length of directory entry */
typedef string nametype<MAXNAMELEN>; /* directory entry */
typedef struct nodenode *namelist; /* link in the listing */

/* A node in the directory listing */
struct nodenode {
    nametype name; /* name of directory entry */
    namelist next; /* next entry */
};

/*
 * The result of a READDIR operation
 *
 * a truly portable application would use
 * an agreed upon list of error codes
 * rather than (as this sample program

(Continued)
* does) rely upon passing UNIX errno’s
* back.
*
* In this example: The union is used
* here to discriminate between successful
* and unsuccessful remote calls.
*/

union readdir_res switch (int errno) {
  case 0:
    namelist list; /* no error: return directory listing */
    default:
    void; /* error occurred: nothing else to return */
  }

/* The directory program definition */
program DIRPROG { 
  version DIRVERS { 
    readdir_res 
    READDIR(nametype) = 1;
  } = 1;
  } = 0x20000076;

You can redefine types (like readdir_res in the example above) using the
struct, union, and enum RPC language keywords. These keywords are not
used in later declarations of variables of those types. For example, if you
define a union, foo, you declare using only foo, and not union foo.

rpcgen compiles RPC unions into C structures. Do not declare C unions
using the union keyword.

Running rpcgen on dir.x generates four output files: (1) the header file, (2)
the client stub, (3) the server skeleton, and (4) the XDR routines in the file
dir_xdr.c. This last file contains the XDR routines to convert declared data
types from the host platform representation into XDR format, and vice versa.

For each RPCL data type used in the .x file, rpcgen assumes that libnsl
contains a routine whose name is the name of the data type, prepended by
the XDR routine header xdr_ (for example, xdr_int). If a data type is
defined in the .x file, rpcgen generates the required xdr_ routine. If there is
no data type definition in the .x source file (for example, msg.x), then no
_xdr.c file is generated.

You can write a .x source file that uses a data type not supported by libnsl,
and deliberately omit defining the type (in the .x file). In doing so, you must
provide the xdr_ routine. This is a way to provide your own customized
xdr_ routines. See Chapter 4,“ for more details on passing arbitrary data
types. The server-side of the READDIR procedure is shown in Code Example 3–5.

**CODE EXAMPLE 3–5  Server dir_proc.c Example**

```c
/*
* dir_proc.c: remote readdir
* implementation
*/
#include <dirent.h>
#include "dir.h" /* Created by rpcgen */
extern int errno;
extern char *malloc();
extern char *strdup();

readdir_res *
readdir_1(dirname, req)
nametype *dirname;
struct svc_req *req;
{
  DIR *dirp;
  struct dirent *d;
  namelist nl;
  namelist *nlp;
  static readdir_res res; /* must be static! */
/* Open directory */
dirp = opendir(*dirname);
if (dirp == (DIR *)NULL) {
  res.errno = errno;
  return (&res);
}
/* Free previous result */
xdr_free(xdr_readdir_res, &res);
/*
* Collect directory entries.
* Memory allocated here is free by
* xdr_free the next time readdir_1
* is called
* /
  nlp = &res.readdir_res_u.list;
  while (d = readdir(dirp)) {
    nl = *nlp = (namenode *)
      malloc(sizeof(namenode));
    if (nl == (namenode *) NULL) {
      res.errno = EAGAIN;
      closedir(dirp);
      return(&res);
    }
    nl->name = strdup(d->d_name);
    nlp = &nl->next;
  }
  *nlp = (namelist)NULL;
/* Return the result */
(Continued)
```
res.errno = 0;
closedir(dirp);
return (&res);
}

Code Example 3–6 shows the client-side implementation of the READDIR procedure.

**CODE EXAMPLE 3–6** Client-side Implementation of rls.c

```c
/*
 * rls.c: Remote directory listing client
 */
#include <stdio.h>
#include "dir.h" /* generated by rpcgen */
extern int errno;

int main(argc, argv)
int argc;
char *argv[];
{
    CLIENT *clnt;
    char *server;
    char *dir;
    readdir_res *result;
    namelist nl;
    if (argc != 3) {
        fprintf(stderr, "usage: %s host
directory\n", argv[0]);
        exit(1);
    }
    server = argv[1];
    dir = argv[2];
    /* Create client "handle" used for
     * calling MESSAGEPROG on the server
     * designated on the command line.
     */
    clnt = clnt_create(server, DIRPROG, 
                      DIRVERS, "tcp");
    if (clnt == (CLIENT*)NULL) {
        clnt_pcreateerror(server);
        exit(1);
    }
```
result = readdir_1(dir, clnt);
if (result == (readdir_res *)NULL) {
    clnt_perror(clnt, server);
    exit(1);
}
/* Okay, we successfully called
 ** the remote procedure.
 */
if (result->errno != 0) {
    /* Remote system error. Print
     * error message and die.
     */
    errno = result->errno;
    perror(dir);
    exit(1);
}
/* Successfully got a directory listing.
 * Print it.
 */
for (nl = result->readdir_res_u.list;
    nl != NULL;
    nl = nl->next) {
    printf("%s\n", nl->name);
}
clnt_destroy(clnt);
exit(0);}

As in other examples, execution is on systems named local and remote.
The files are compiled and run as follows:

remote$ rpcgen dir.x
remote$ cc -c dir_xdr.c
remote$ cc rls.c dir_clnt.c dir_xdr.o -o rls -lnsl
remote$ cc dir_svc.c dir_proc.c dir_xdr.o -o dir_svc -lnsl
remote$ dir_svc

When you install rls on system local, you can list the contents of
/usr/share/lib on system remote as follows:

local$ rls remote /usr/share/lib
ascii
eqnchar
greek
kbd
marg8	abclr
tabs
tabs4
local$ rpcgen generated client code does not release the memory allocated for the results of the RPC call. Call xdr_free to release the memory when you are finished with it. It is similar to calling the free routine, except that you pass the XDR routine for the result. In this example, after printing the list, xdr_free(xdr_readdir_res, result); was called.

Note - Use xdr_free to release memory allocated by malloc. Failure to use xdr_free to release memory results in memory leaks.

Preprocessing Directives

rpcgen supports C and other preprocessing features. C preprocessing is performed on rpcgen input files before they are compiled. All standard C preprocessing directives are allowed in the .x source files. Depending on the type of output file being generated, five symbols are defined by rpcgen.

rpcgen provides an additional preprocessing feature: any line that begins with a percent sign (%) is passed directly to the output file, with no action on the line's content. Caution is required because rpcgen does not always place the lines where you intend. Check the output source file and, if needed, edit it.

<table>
<thead>
<tr>
<th>Symbol</th>
<th>Use</th>
</tr>
</thead>
<tbody>
<tr>
<td>RPC_HDR</td>
<td>Header file output</td>
</tr>
<tr>
<td>RPC_XDR</td>
<td>XDR routine output</td>
</tr>
<tr>
<td>RPC_SVC</td>
<td>Server stub output</td>
</tr>
<tr>
<td>RPC_CLNT</td>
<td>Client stub output</td>
</tr>
<tr>
<td>RPC_TBL</td>
<td>Index table output</td>
</tr>
</tbody>
</table>

TABLE 3-1 rpcgen Preprocessing Directives

Code Example 3-7 is a simple rpcgen example. Note the use of rpcgen's pre-processing features.
CODE EXAMPLE 3–7  Time Protocol rpcgen Source

/*
 * time.x: Remote time protocol
 */
program TIMEPROG {
    version TIMEVERS {
        unsigned int TIMEGET() = 1;
    } = 1;
    () = 0x20000044;
}

#ifdef RPC_SVC
%int *
%timeget_1() %{
    static int thetime;
    % thetime = time(0);
    % return (&thetime);
%
#endif

cpp Directive

rpcgen supports C preprocessing features. rpcgen defaults to use
/usr/ccs/lib/cpp as the C preprocessor. If that fails, rpcgen tries to use
/lib/cpp. You may specify a library containing a different cpp to rpcgen
with the −Y flag.

For example, if /usr/local/bin/cpp exists, you can specify it to rpcgen
as follows:

rpcgen −Y /usr/local/bin test.x

Compile-Time Flags

This section describes the rpcgen options available at compile time. The
following table summarizes the options which are discussed in this section.
### TABLE 3-2  \texttt{rpcgen} Compile-time Flags

<table>
<thead>
<tr>
<th>Option</th>
<th>Flag</th>
<th>Comments</th>
</tr>
</thead>
<tbody>
<tr>
<td>Templates</td>
<td>−a, −Sc, −Ss, −Sm</td>
<td>See Table 3-3</td>
</tr>
<tr>
<td>C-style</td>
<td>−N</td>
<td>Also called Newstyle mode</td>
</tr>
<tr>
<td>ANSI C</td>
<td>−C</td>
<td>Often used with the -N option</td>
</tr>
<tr>
<td>MT-Safe code</td>
<td>−M</td>
<td>For use in multithreaded environments</td>
</tr>
<tr>
<td>MT Auto mode</td>
<td>−A</td>
<td>-A also turns on -M option</td>
</tr>
<tr>
<td>TS-RPC library</td>
<td>−b</td>
<td>TI-RPC library is default</td>
</tr>
<tr>
<td>xdr_inline count</td>
<td>−i</td>
<td>Uses 5 packed elements as default, but other number may be specified</td>
</tr>
</tbody>
</table>

### Client and Server Templates

\texttt{rpcgen} generates sample code for the client and server sides. Use these options to generate the desired templates.

### TABLE 3-3  \texttt{rpcgen} Template Selection Flags

<table>
<thead>
<tr>
<th>Flag</th>
<th>Function</th>
</tr>
</thead>
<tbody>
<tr>
<td>−a</td>
<td>Generate all template files</td>
</tr>
<tr>
<td>−Sc</td>
<td>Generate client-side template</td>
</tr>
<tr>
<td>−Ss</td>
<td>Generate server-side template</td>
</tr>
<tr>
<td>−Sm</td>
<td>Generate makefile template</td>
</tr>
</tbody>
</table>

The files can be used as guides or by filling in the missing parts. These files are in addition to the stubs generated.
A C-style mode server template is generated from the add.x source by the command:

```bash
rpcgen -N -Ss -o add_server_template.c add.x
```

The result is stored in the file `add_server_template.c`. A C-style mode, client template for the same add.x source is generated with the command line:

```bash
rpcgen -N -Sc -o add_client_template.c add.x
```

The result is stored in the file `add_client_template.c`. A make file template for the same add.x source is generated with the command line:

```bash
rpcgen -N -Sm -o mkfile_template add.x
```

The result is stored in the file `mkfile_template`. It can be used to compile the client and the server. If the `-a` flag is used as follows:

```bash
rpcgen -N -a add.x
```

rpcgen generates all three template files. The client template goes into `add_client.c`, the server template to `add_server.c`, and the makefile template to `makefile.a`. If any of these files already exists, rpcgen displays an error message and exits.

*Note* - When you generate template files, give them new names to avoid the files being overwritten the next time rpcgen is executed.

### C-style Mode

Also called Newstyle mode, The `-N` flag causes rpcgen to produce code in which arguments are passed by value and multiple arguments are passed without a `struct`. These changes allow RPC code that is more like C and other high-level languages. For compatibility with existing programs and make files, the previous (standard) mode of argument passing is the default. The following examples demonstrate the new feature. The source modules for both modes, C-style and default, are given in Code Example 3-8 and in Code Example 3-9 respectively.

**CODE EXAMPLE 3-8**  C-style Mode Version of add.x

```c
/*
 * This program contains a procedure
 * to add 2 numbers. It demonstrates
 * the C-style mode argument passing.
 */
```
* Note that add() has 2 arguments.
*/
program ADDPROG { /* program number */
    version ADDVER { /* version number */
        int add(int, int) = 1; /* procedure */
    } = 1;
} = 0x20000199;

CODE EXAMPLE 3-9  Default Mode Version of add.x

/*
 * This program contains a procedure
 * to add 2 numbers. It demonstrates
 * the "default" mode argument passing.
 * In this mode rpcgen can process
 * only one argument.
 */
struct add_arg {
    int first;
    int second;
};
program ADDPROG { /* program number */
    version ADDVER { /* version number */
        int add (add_arg) = 1; /* procedure */
    } = 1;
} = 0x20000199;

The next four figures show the resulting client-side templates.

CODE EXAMPLE 3-10  C-style Mode Client Stub for add.x

/*
 * The C-style client side main
 * routine calls the add() function
 * on the remote rpc server
 */
#include <stdio.h>
#include "add.h"
main(argc, argv)
int argc;
char *argv[];
{
    CLIENT *clnt;
    int *result,x,y;

    if(argc != 4) {
        printf("usage: \ts host num1
               num2\n" argv[0]);
        exit(1);
    }
    /* create client handle -
     * bind to server
     */
    clnt = clnt_create(argv[1], ADDPROG,
} (Continued)
ADDVER, "udp");
if (clnt == NULL) {
    clnt_pcreateerror(argv[1]);
    exit(1);
}  
int x = atoi(argv[2]);
int y = atoi(argv[3]);
/*
 * invoke remote procedure: Note that
 * multiple arguments can be passed to
 * add_1() instead of a pointer
 */
    int result = add_1(x, y, clnt);
    if (result == (int*) NULL) {
        clnt_perror(clnt, "call failed:");
        exit(1);
    } else {
        printf("Success: %d + %d = %d\n", x, y, *result);
    }
exit(0);
}

Code Example 3–11 shows how the default mode code differs from C-style mode code.

**CODE EXAMPLE 3–11  Default Mode Client**

```c
arg.first = atoi(argv[2]);
arg.second = atoi(argv[3]);
/*
 * invoke remote procedure -- note
 * that a pointer to the argument has
 * to be passed to the client stub
 */
    result = add_1(&arg, clnt);
```

The server-side procedure in C-style mode is shown in Code Example 3–12.

**CODE EXAMPLE 3–12  C-style Mode Server**

```c
#include "add.h"

int *
add_1(arg1, arg2, rqstp)
    int arg1;
    int arg2;
    struct svc_req *rqstp;
```
The server side procedure in default mode is shown in Code Example 3–13.

**CODE EXAMPLE 3–13  Default Mode Server Stub**

```c
#include "add.h"

int *add_1(argp, rqstp)
    add_arg *argp;
    struct svc_req *rqstp;
{ 
    static int result;
    result = argp->first + argp->second;
    return(&result);
}
```

**MT-Safe Code**

By default, the code generated by `rpcgen` is not MT safe. It uses unprotected global variables and returns results in the form of static variables. The `-M` flag generates MT-safe code which can be used in a multithreaded environment. This can be used with the C-style flag, the ANSI C flag, or both.

An example of an MT-safe program with this interface follows. The `rpcgen` protocol file is `msg.x`, shown in Code Example 3–14.

**CODE EXAMPLE 3–14   MT-Safe Program: msg.**

```c
program MESSAGEPROG {
version PRINTMESSAGE {
    int PRINTMESSAGE (string) = 1;
} = 1;
} = 0x4001;
```

A string is passed to the remote procedure, which prints it and returns the length of the string to the client. The MT-Safe stubs are generated with:

```bash
% rpcgen -M msg.x
```

A possible client-side code that could be used with this is shown in Code Example 3–15.
#include "msg.h"

void
messageprog_1(host)
char *host;
{
CLIENT *clnt;
enum clnt_stat retval_1;
int result_1;
char * printmessage_1_arg;

clnt = clnt_create(host, MESSAGEPROG,
PRINTMESSAGE,
"netpath");
if (clnt == (CLIENT *) NULL) {
clnt_pcreateerror(host);
exit(1);
}
printmessage_1_arg =
(char *) malloc(256);
strcpy(printmessage_1_arg, "Hello World");

retval_1 = printmessage_1(&printmessage_1_arg,
&result_1,clnt);
if (retval_1 != RPC_SUCCESS) {
clnt_perror(clnt, "call failed");
}
printf("result = %d\n", result_1);
clnt_destroy(clnt);
}

main(argc, argv)
int argc;
char *argv[];
{
char *host;

if (argc < 2) {
printf("usage: %s server_host\n", argv[0]);
exit(1);
}
host = argv[1];
messageprog_1(host);
}

Note that a pointer to both the arguments and the results needs to be passed in to the rpcgen-generated code. This is to preserve reentrancy. The value returned by the stub function indicates whether this call is a success or a failure. The stub returns RPC_SUCCESS if the call is successful. Compare the
MT-safe client stub (generated with the -M option) and the not MT-safe client stub shown in Code Example 3–16. The client stub that is not MT safe uses a static to store returned results and can use only one thread at a time.

**CODE EXAMPLE 3–16**  Client Stub (Not MT Safe)

```c
int *
printmessage_1(argp, clnt)
    char **argp;
    CLIENT *clnt;
{
    static int clnt_res;
    memset((char *)&clnt_res, 0,
           sizeof (clnt_res));
    if (clnt_call(clnt, PRINTMESSAGE,
                 (xdrproc_t) xdr_wrapstring,
                 (caddr_t) argp,
                 (xdrproc_t) xdr_int, (caddr_t) &clnt_res,
                 TIMEOUT) != RPC_SUCCESS) {
        return (NULL);
    }
    return (&clnt_res);
}
```

The server side code is shown in Code Example 3–17.

**Note** - When compiling a server that uses MT-safe mode, you must link in the threads library. To do this, specify the `-lthread` option in the compile command.

**CODE EXAMPLE 3–17**  MT-Safe Server Stub

```c
#include "msg.h"
#include <syslog.h>

bool_t
printmessage_1_svc(argp, result, rqstp)
    char **argp;
    int *result;
    struct svc_req *rqstp;
{
    int retval;
    if (*argp == NULL) {
        syslog(LOG_INFO, "argp is NULL\n");
        *result = 0;
    } else {
        syslog("argp is \%s\n", *argp);
        *result = strlen (*argp);
    }
}
```

(Continued)
The server side code should not use statics to store returned results. A pointer to the result is passed in and this should be used to pass the result back to the calling routine. A return value of 1 indicates success to the calling routine, while 0 indicates a failure.

In addition, the code generated by `rpcgen` also generates a call to a routine to free any memory that may have been allocated when the procedure was called. To prevent memory leaks, any memory allocated in the service routine needs to be freed in this routine. `messageprog_1_freeresult` frees the memory.

Normally, `xdr_free` frees any allocated memory for you (in this case, no memory was allocated, so no freeing needs to take place).

As an example of the use of the `-M` flag with the C-style and ANSI C flag, consider the following file, `add.x`, shown in Code Example 3–18.

**CODE EXAMPLE 3–18**  MT-Safe Program: add.x

```c
program ADDPROG {
    version ADDVER {
        int add(int, int) = 1;
    } = 1;
} = 199;
```

This program adds two numbers and returns its result to the client. `rpcgen` is invoked on it, with the following command: `% rpcgen -N -M -C add.x`

The multithreaded client code to call this is shown in Code Example 3–19.
CODE EXAMPLE 3–19  MT-Safe Client: add.x

```c
/*
 * This client-side main routine
 * starts up a number of threads,
 * each of which calls the server
 * concurrently.
 */

#include "add.h"

CLIENT *clnt;
#define NUMCLIENTS 5
struct argrec {
    int arg1;
    int arg2;
};

/* Keeps count of number of
 * threads running */
int numrunning;
mutex_t numrun_lock;
cond_t condnum;

void
addprog(struct argrec *args)
{
    enum clnt_stat retval;
    int result;
    /* call server code */
    retval = add_1(args->arg1, args->arg2,
                    &result, clnt);
    if (retval != RPC_SUCCESS) {
        clnt_perror(clnt, "call failed");
    } else
        printf("thread #%x call succeeded,
                result = %d\n", thr_getself(),
                result);
    /* decrement the number of running
     * threads */
    mutex_lock(&numrun_lock);
    numrunning--;
    cond_signal(&condnum);
    mutex_unlock(&numrun_lock);
    thr_exit(NULL);
}

main(int argc, char *argv[])
{
    char *host;
    struct argrec args[NUMCLIENTS];
    int i;
    thread_t mt;
    int ret;
}
```

(Continued)
if (argc < 2) {
    printf("usage: %s server_host\n", argv[0]);
    exit(1);
}
host = argv[1];
clnt = clnt_create(host, ADDPROG, ADDVER, "netpath");
if (clnt == (CLIENT *) NULL) {
    clnt_pcreateerror(host);
    exit(1);
};
mutex_init(&numrun_lock, USYNC_THREAD, NULL);
cond_init(&condnum, USYNC_THREAD, NULL);
umrunning = 0;
/* Start up separate threads */
for (i = 0; i < NUMCLIENTS; i++) {
    args[i].arg1 = i;
    args[i].arg2 = i + 1;
    ret = thr_create(NULL, NULL, addprog,
        (char *) &args[i],
        THR_NEW_LWP, &mt);
    if (ret == 0)
        numrunning++;
}
mutex_lock(&numrun_lock);
/* are any threads still running ? */
while (numrunning != 0)
    cond_wait(&condnum, &numrun_lock);
mutex_unlock(&numrun_lock);
clnt_destroy(clnt);
}
The server-side procedure is shown in Code Example 3–20.

**Note** - When compiling a server that uses MT-safe mode, you must link in the threads library. To do this, specify the `-lthread` option in the compile command.

**CODE EXAMPLE 3–20**  MT-Safe Server: add.x

```c
add_1_svc(int arg1, int arg2,
        int *result, struct svc_req *rqstp)
```
bool_t retval;
/* Compute result */
*result = arg1 + arg2;
retval = 1;
return (retval);
}

/* Routine for freeing memory that may
* be allocated in the server procedure */
int
addprog_1_freeresult(SVCXPRT *transp,
xdrproc_t xdr_result,
caddr_t result)
{
(void) xdr_free(xdr_result, result);
}

MT Auto Mode

MT Auto mode enables RPC servers to automatically use Solaris threads to
process client requests concurrently. Use the -A option to generate RPC code
in MT Auto mode. The -A option also has the effect of turning on the -m
option, so -m does not need to be explicitly specified. The -m option is
necessary because any code generated has to be multithread safe.

Further discussion on multithreaded RPC begins on “Multithreaded RPC
Programming” on page 105; see also “MT Auto Mode” on page 112.

Here is an example of an Auto mode program generated by rpcgen. The
rpcgen protocol file time.x is shown in Code Example 3–21. A string is
passed to the remote procedure, which prints it and returns the length of the
string to the client. The MT-safe stubs are generated with:

CODE EXAMPLE 3–21  MT Auto Mode: time.x

program TIMEPROG {
    version TIMEVERS {
        unsigned int TIMEGET(void) = 1;
        void TIMESET(unsigned) = 2;
    } = 1;
} = 0x20000044;

% rpcgen -A time.x

Note - When the -A option is used, the generated server code will contain
instructions for enabling MT Auto mode for the server.

When compiling a server that uses MT Auto mode, you must link in the
threads library. To do this, specify the -lthread option in the compile
command.
TI-RPC or TS-RPC Library Selection

In older SunOS releases, rpcgen created stubs that used the socket functions. With the current SunOS release, you can use either the transport-independent RPC (TI-RPC) or the transport-specific socket (TS-RPC) routines. This provides backward compatibility with previous releases. The default uses the TI-RPC interfaces. The −b flag tells rpcgen to create TS-RPC variant source code as its output.

ANSI C-compliant Code

rpcgen can also produce output that is compatible with ANSI C or SPARCompiler C++ 3.0. This feature is selected with the −C compile flag and is most often used with the −N flag, described in “C-style Mode” on page 38. The add.x example of the server template is generated by the command:

```
rpcgen -N -C -Ss -o add_server_template.c add.x
```

It is important to note that on the C++ 3.0 server, remote procedure names require an _svc suffix. In the following example, the add.x template and the −C compile flag produce the client side add_1 and the server stub add_1_svc.

CODE EXAMPLE 3–22  rpcgen ANSI C Server Template

```c
/*
 * This is a template. Use it to
 * develop your own functions.
 */
#include <c_varieties.h>
#include "add.h"

int *
add_1_svc(int arg1, int arg2,
            struct svc_req *rqstp)
{
    static int result;
    /* insert server code here */
    return(&result);
}
```
This output conforms to the syntax requirements and structure of ANSI C. The header files that are generated when this option is invoked can be used with ANSI C or with C++.

xdr_inline Count

rpcgen tries to generate more efficient code by using xdr_inline when possible (see xdr_admin(3) manpage). When a structure contains elements that xdr_inline can be used on (for example integer, long, bool), the relevant portion of the structure is packed with xdr_inline. A default of five or more packed elements in sequence causes in-line code to be generated. This default can be changed with the −i flag. For example:

```
rpcgen −i 3 test.x
```

causes rpcgen to start generating in-line code after three qualifying elements are found in sequence. The example:

```
rpcgen −i 0 test.x
```

prevents any in-line code from being generated.

In most situations, there is no reason to use the −i flag. The _xdr.c stub is the only file affected by this feature.

rpcgen Programming Techniques

This section suggests some common RPC and rpcgen programming techniques. Each topic is covered in its own subsection.

- Network Type
  rpcgen can produce server code for specific transport types.

- Define Statements
  C-preprocessing symbols can be defined on rpcgen command lines.

- Broadcast Calls
  Servers need not send error replies to broadcast calls.

- Debugging Applications
  Debug as normal function calls, then change to a distributed application.

- Port Monitor Support
Port monitors can “listen” on behalf of RPC servers.

- Dispatch Tables
  Programs can access server dispatch tables.

- Time-out Changes
  Client default time-out periods can be changed.

- Authentication
  Clients may authenticate themselves to servers; interested servers can examine client authentication information.

**Network Types/Transport Selection**

`rpcgen` takes optional arguments that allow a programmer to specify desired network types or specific network identifiers. (For details of network selection, see *Transport Interfaces Programming Guide*).

The `-s` flag creates a server that responds to requests on the specified type of transport. For example, the invocation

```
rpcgen -s datagram_n prot.x
```

writes a server to standard output that responds to any of the connectionless transports specified in the `NETPATH` environment variable (or in `/etc/netconfig`, if `NETPATH` is not defined). A command line can contain multiple `-s` flags and their network types.

Similarly, the `-n` flag creates a server that responds only to requests from the transport specified by a single network identifier.

---

**Caution** - Be careful using servers created by `rpcgen` with the `-n` flag. Network identifiers are host specific, so the resulting server may not run as expected on other hosts.

**Command Line Define Statements**

You can define C-preprocessing symbols and assign values to them from the command line. Command line define statements can, for example, be used to generate conditional debugging code when the `DEBUG` symbol is defined. For example:

```
$ rpcgen -DDEBUG proto.x
```
Server Response to Broadcast Calls

When a procedure has been called through broadcast RPC and cannot provide a useful response, the server should send no reply to the client. This reduces network traffic. To prevent the server from replying, a remote procedure can return NULL as its result. The server code generated by rpcgen detects this and sends no reply.

Code Example 3–23 is a procedure that replies only if it is an NFS server.

**CODE EXAMPLE 3–23**  
NFS Server Response to Broadcast Calls

```c
void * reply_if_nfsserver()  
{  
char notnull; /* only here so we can use its address */  
  if( access( "/etc/dfs/sharetab",  
         F_OK ) < 0 ) {  
    /* prevent RPC from replying */  
    return( (void *) NULL );  
  }  
  /* assign notnull a non-null value */  
  /* so RPC will send a reply */  
  return( (void *) &notnull );  
}
```

A procedure must return a non-NULL pointer when it wants RPC library routines to send a reply.

In Code Example 3–23, if the procedure reply_if_nfsserver is defined to return non void values, the return value (&notnull) should point to a static variable.

Port Monitor Support

Port monitors such as inetd and listen can monitor network addresses for specified RPC services. When a request arrives for a particular service, the port monitor spawns a server process. After the call has been serviced, the server can exit. This technique conserves system resources. The main server function generated by rpcgen allows invocation by inetd. See “Using inetd” on page 100 “for details.

It may be useful for services to wait for a specified interval after satisfying a service request, in case another request follows. If there is no call in the specified time, the server exits, and some port monitors, like inetd, continue to monitor for the server. If a later request for the service occurs, the port
monitor gives the request to a waiting server process (if any), rather than spawning a new process.

**Note** - When monitoring for a server, some port monitors, like `listen`, always spawn a new process in response to a service request. If a server is used with such a monitor, it should exit immediately on completion.

By default, services created using `rpcgen` wait for 120 seconds after servicing a request before exiting. The programmer can change the interval with the `-K` flag. In this example,

```
$ rpcgen -K 20 proto.x
```

the server waits for 20 seconds before exiting. To create a server that exits immediately, zero value can be used for the interval period:

```
$ rpcgen -K 0 proto.x.
```

To create a server that never exits, the value is `−K 1`.

### Time-out Changes

After sending a request to the server, a client program waits for a default period (25 seconds) to receive a reply. This time-out may be changed using the `clnt_control` routine. See “Standard Interfaces” on page 65, for additional uses of the `clnt_control` routine. See also the `rpc(3N)` manpage. When considering time-out periods, be sure to allow the minimum amount of time required for “round-trip” communications over the network. Code Example 3–24 illustrates the use of `clnt_control`.

**CODE EXAMPLE 3-24**  clnt_control Routine

```c
struct timeval tv;
CLIENT *clnt;
clnt = clnt_create( "somehost", SOMEPROG,
                    SOMEVERS, "visible" );
if (clnt == (CLIENT *)NULL)
    exit(1);
tv.tv_sec = 60; /* change time-out to
                 * 60 seconds
                 */
tv.tv_usec = 0;
clnt_control(clnt, CLSET_TIMEOUT, &tv);
```
Client Authentication

The client create routines do not have any facilities for client authentication. Some clients may have to authenticate themselves to the server.

The following example illustrates one of the least secure authentication methods in common use. See “Authentication” on page 91, for information on the more secure DES authentication technique.

**CODE EXAMPLE 3–25**  AUTH_SYS Authentication Program

```c
CLIENT *clnt;
clnt = clnt_create( "somehost", SOMEPROG,
                   SOMEVERS, "visible" );
if (clnt != (CLIENT *)NULL) {
   /* To set AUTH_SYS style authentication */
   clnt->cl_auth = authsys_createdefault();
}
```

Authentication information is important to servers that have to achieve some level of security. This extra information is supplied to the server as a second argument.

Code Example 3–26 is a server that checks client authentication data. It is modified from printmessage_1 in “An rpcgen Tutorial” on page 23 and only allows superusers to print a message to the console.

**CODE EXAMPLE 3–26**  printmsg_1 for Superuser

```c
int *
printmessage_1(msg, req)
   char **msg;
   struct svc_req *req;
{
   static int result; /* Must be static */
   FILE *f;
   struct authsys_parms *aup;

   aup = (struct authsys_parms *)req->rq_clntcred;
   if (aup->aup_uid != 0) {
      result = 0;
      return (&result)
   }
   /* Same code as before. */
}
```

Dispatch Tables

It is sometimes useful for programs to have access to dispatch tables used by the RPC package. For example, the server dispatch routine may check authorization and then invoke the service routine; or a client library may deal with the details of storage management and XDR data conversion.
When invoked with the −T option, rpcgen generates RPC dispatch tables for each program defined in the protocol description file, proto.x, in the file proto_tbl.i. The suffix i stands for “index.” rpcgen may be invoked with the −t option to build only the header file. rpcgen cannot be invoked in C-style mode (−N) with either the −T or −t flag.

Each entry in the dispatch table is a struct rpcgen_table, defined in the header file proto.h as follows:

```c
struct rpcgen_table {  
    char *(proc);  
    xdrproc_t xdr_arg;  
    unsigned len_arg;  
    xdrproc_t xdr_res;  
    xdrproc_t len_res  
};
```

where:

- **proc** is a pointer to the service routine
- **xdr_arg** is a pointer to the input (argument) xdr routine
- **len_arg** is the length in bytes of the input argument
- **xdr_res** is a pointer to the output (result) xdr routine
- **len_res** is the length in bytes of the output result

The table, named dirprog_1_table for the dir.x example, is indexed by procedure number. The variable dirprog_1_nproc contains the number of entries in the table.

An example of how to locate a procedure in the dispatch tables is shown by the routine find_proc:

```c
struct rpcgen_table *  
find_proc(proc)  
    u_long proc;  
{  
    if (proc >= dirprog_1_nproc)  
        /* error */  
    else  
        return (&dirprog_1_table[proc]);  
}
```

Each entry in the dispatch table contains a pointer to the corresponding service routine. However, that service routine is usually not defined in the client code. To avoid generating unresolved external references, and to require only one source file for the dispatch table, the rpcgen service routine initializer is RPCGEN_ACTION(proc_ver).

This way, the same dispatch table can be included in both the client and the server. Use the following define statement when compiling the client:

```c
#define RPCGEN_ACTION(routine) 0
```
And use the following define when writing the server:

```
#define RPCGEN_ACTION(routine)routine
```

### Debugging Applications

You can simplify the testing and debugging process. First test the client program and the server procedure in a single process by linking them with each other rather than with the client and server skeletons. Comment out calls to the client create RPC library routines (see the `rpc_clnt_create(3N)` manpage) and the authentication routines. Do not link with libnsl.

Link the procedures from previous example by:

```
cc rls.c dir_clnt.c dir_proc.c -o rls
```

With the RPC and XDR functions commented out, the procedure calls execute as ordinary local function calls, and the program is debugged with a local debugger such as `dbxtool`. When the program works, the client program is linked to the client skeleton produced by `rpcgen` and the server procedures are linked to the server skeleton produced by `rpcgen`.

You can also use the Raw RPC mode to test the XDR routines. See “Testing Programs Using Low-level Raw RPC” on page 81 for details.

There are two kinds of errors that can happen in an RPC call. The first kind of error is caused by a problem with the mechanism of the remote procedure calls. Examples of these are (1) the procedure is not available, (2) the remote server is not responding, and (3) the remote server is unable to decode the arguments. In Code Example 3–26, an RPC error happens if `result` is `NULL`. The reason for the failure can be displayed by using `clnt_perror`, or an error string can be returned through `clnt_sperror`.

The second type of error is caused by the server itself. In Code Example 3–26, an error can be returned by `opendir`. The handling of these errors is application specific and is the responsibility of the programmer.

Note that the mechanism illustrated by the paragraphs above does not function with the `−C` option because of the `_svc` suffix added to the server-side routines.
This chapter addresses the C interface to RPC and describes how to write network applications using RPC. For a complete specification of the routines in the RPC library, see the \texttt{rpc(3N)} and related man pages.

- “Simplified Interface” on page 57
- “Standard Interfaces” on page 65
- “Testing Programs Using Low-level Raw RPC” on page 81
- “Advanced RPC Programming Techniques” on page 84
- “Multithreaded RPC Programming” on page 105
- “MT Auto Mode” on page 112
- “MT User Mode” on page 116
- “Porting From TS-RPC to TI-RPC” on page 128

RPC Is Multithread Safe

The client and server interfaces described in this chapter are multithread safe, except where noted (such as raw mode). This means that applications that contain RPC function calls can be used freely in a multithreaded application.

Simplified Interface

The simplified interface is the easiest level to use because it does not require the use of any other RPC routines. It also limits control of the underlying...
communications mechanisms. Program development at this level can be rapid, and is directly supported by the rpcgen compiler. For most applications, rpcgen and its facilities are sufficient.

Some RPC services are not available as C functions, but they are available as RPC programs. The simplified interface library routines provide direct access to the RPC facilities for programs that do not require fine levels of control. Routines such as rusers are in the RPC services library librpcsvc. Code Example 4–1 is a program that displays the number of users on a remote host. It calls the RPC library routine, rusers.

**CODE EXAMPLE 4–1  rusers Program**

```c
#include <rpc/rpc.h>
#include <rpcsvc/rusers.h>
#include <stdio.h>

/*
 * a program that calls the
 * rusers() service
 */

main(argc, argv)
int argc;
char **argv;
{
    int num;

    if (argc != 2) {
        fprintf(stderr, "usage: %s hostname\n", argv[0]);
        exit(1);
    }

    if ((num = rnusers(argv[1])) < 0) {
        fprintf(stderr, "error: rusers\n");
        exit(1);
    }

    fprintf(stderr, "%d users on %s\n", num, argv[1] );
    exit(0);
}
```

Compile the program in Code Example 4–1 with:

```bash
cc program.c -lrpcsvc -lnsl
```

**Client**

There is just one function on the client side of the simplified interface: rpc_call. It has nine parameters:
int 0 or error code
rpc_call (  
    char *host /* Name of server host */  
    u_long prognum /* Server program number */  
    u_long versnum /* Server version number */  
    xdrproc_t inproc /* XDR filter to encode arg */  
    char *in /* Pointer to argument */  
    xdr_proc_t outproc /* Filter to decode result */  
    char *out /* Address to store result */  
    char *nettype /* For transport selection */  
);  

This function calls the procedure specified by prognum, versnum, and procnun on the host. The argument to be passed to the remote procedure is pointed to by the \textit{in} parameter, and \textit{inproc} is the XDR filter to encode this argument. The \textit{out} parameter is an address where the result from the remote procedure is to be placed. \textit{outproc} is an XDR filter which will decode the result and place it at this address.

The client blocks on \texttt{rpc\_call} until it receives a reply from the server. If the server accepts, it returns RPC\_SUCCESS with the value of zero. It will return a non-zero value if the call was unsuccessful. This value can be cast to the type \texttt{clnt\_stat}, an enumerated type defined in the RPC include files and interpreted by the \texttt{clnt\_sperrno} function. This function returns a pointer to a standard RPC error message corresponding to the error code.

In the example, all “visible” transports listed in /\texttt{/etc/netconfig} are tried. Adjusting the number of retries requires use of the lower levels of the RPC library.

Multiple arguments and results are handled by collecting them in structures.

The example in Code Example 4–1, changed to use the simplified interface, looks like Code Example 4–2.

\textbf{CODE EXAMPLE 4–2} rusers Program Using Simplified Interface

```
#include <stdio.h>
#include <utmp.h>
#include <rpc/rpc.h>
#include <rpcsvc/rusers.h>

/* a program that calls the RUSERSPROG * RPC program */

main(argc, argv)
int argc;
char **argv;
{
  unsigned long nusers;
  enum clnt_stat cs;
```

(Continued)
if (argc != 2) {
    fprintf(stderr, "usage: rusers hostname\n");
    exit(1);
}

if( cs = rpc_call(argv[1], RUSERSPROG, RUSERSVERS, RUSERSPROC_NUM, xdr_void, (char *)0, xdr_u_long, (char *)&nusers, "visible") != RPC_SUCCESS ) {
    clnt_perrno(cs);
    exit(1);
}

fprintf(stderr, "%d users on %s\n", nusers, argv[1] );
exit(0);
}

Since data types may be represented differently on different machines, rpc_call needs both the type of, and a pointer to, the RPC argument (similarly for the result). For RUSERSPROC_NUM, the return value is an unsigned long, so the first return parameter of rpc_call is xdr_u_long (which is for an unsigned long) and the second is &nusers (which points to unsigned long storage). Because RUSERSPROC_NUM has no argument, the XDR encoding function of rpc_call is xdr_void and its argument is NULL.

Server

The server program using the simplified interface is very straightforward. It simply calls rpc_reg to register the procedure to be called, and then it calls svc_run, the RPC library’s remote procedure dispatcher, to wait for requests to come in.

rpc_reg has the following arguments:

```
rpc_reg (  
    u_long prognum /* Server program number */  
    u_long versnum /* Server version number */  
    u_long procnum /* server procedure number */  
    char *procname /* Name of remote function */  
    xdrproc_t inproc /* Filter to encode arg */  
    xdrproc_t outproc /* Filter to decode result */  
    char *nettype /* For transport selection */  
);  
```
svc_run invokes service procedures in response to RPC call messages. The
dispatcher in rpc_reg takes care of decoding remote procedure arguments
and encoding results, using the XDR filters specified when the remote
procedure was registered. Some notes about the server program:

- Most RPC applications follow the naming convention of appending a _1 to
  the function name. The sequence _n is added to the procedure names to
  indicate the version number n of the service.
- The argument and result are passed as addresses. This is true for all
  functions that are called remotely. If you pass NULL as a result of a
  function, then no reply is sent to the client. It is assumed that there is no
  reply to send.
- The result must exist in static data space because its value is accessed after
  the actual procedure has exited. The RPC library function that builds the
  RPC reply message accesses the result and sends the value back to the
  client.
- Only a single argument is allowed. If there are multiple elements of data,
  they should be wrapped inside a structure which can then be passed as a
  single entity.
- The procedure is registered for each transport of the specified type. If the
  type parameter is (char *)NULL, the procedure is registered for all
  transports specified in NETPATH.

Hand-Coded Registration Routine

You can sometimes implement faster or more compact code than can
rpcgen. rpcgen handles the generic code-generation cases. The following
program is an example of a hand-coded registration routine. It registers a
single procedure and enters svc_run to service requests. Hand-Coded
Registration Server

```c
#include <stdio.h>
#include <rpc/rpc.h>
#include <rpcsvc/rusers.h>
void *rusers();

main()
{
    if(rpc_reg(RUSERSPROG, RUSERSVERS,
        RUSERSPROC_NUM, rusers,
        xdr_void, xdr_u_long,
        "visible") == -1) {
        fprintf(stderr, "Couldn’t Register\n");
        exit(1);
    }
    svc_run(); /* Never returns */

(Continued)
```
fprintf(stderr, "Error: svc_run returned!\n");
exit(1);
}

rpc_reg can be called as many times as is needed to register different programs, versions, and procedures.

Passing Arbitrary Data Types

Data types passed to and received from remote procedures can be any of a set of predefined types, or can be programmer-defined types. RPC handles arbitrary data structures, regardless of different machines’ byte orders or structure layout conventions, by always converting them to a standard transfer format called external data representation (XDR) before sending them over the transport. The conversion from a machine representation to XDR is called serializing, and the reverse process is called deserializing.

The translator arguments of rpc_call and rpc_reg can specify an XDR primitive procedure, like xdr_u_long, or a programmer-supplied routine that processes a complete argument structure. Argument processing routines must take only two arguments: a pointer to the result and a pointer to the XDR handle.

<table>
<thead>
<tr>
<th>TABLE 4–1</th>
<th>XDR Primitive Type Routines</th>
</tr>
</thead>
<tbody>
<tr>
<td>xdr_int</td>
<td>xdr_netobj</td>
</tr>
<tr>
<td>xdr_long</td>
<td>xdr_float</td>
</tr>
<tr>
<td>xdr_short</td>
<td>xdr_double</td>
</tr>
<tr>
<td>xdr_char</td>
<td>xdr_quadruple</td>
</tr>
</tbody>
</table>
The nonprimitive xdr_string, which takes more than two parameters, is called from xdr_wrapstring.

For an example of a programmer-supplied routine, the structure:

```c
struct simple {
    int a;
    short b;
} simple;
```

contains the calling arguments of a procedure. The XDR routine xdr_simple translates the argument structure as shown in Code Example 4–3.

**CODE EXAMPLE 4–3  xdr_simple Routine**

```c
#include <rpc/rpc.h>
#include "simple.h"

bool_t
xdr_simple(xdrsp, simplep)
XDR *xdrsp;
struct simple *simplep;
{
    if (!xdr_int(xdrsp, &simplep->a))
        return (FALSE);
    if (!xdr_short(xdrsp, &simplep->b))
        return (FALSE);
    return (TRUE);
}
```

An equivalent routine can be generated automatically by rpcgen.

An XDR routine returns nonzero (a C TRUE) if it completes successfully, and zero otherwise. A complete description of XDR is provided in Appendix C.

**TABLE 4–2**

Prefabricated Routines

<table>
<thead>
<tr>
<th>xdr_array</th>
<th>xdr_bytes</th>
<th>xdr_reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>xdr_vector</td>
<td>xdr_union</td>
<td>xdr_pointer</td>
</tr>
<tr>
<td>xdr_string</td>
<td>xdrOpaque</td>
<td></td>
</tr>
</tbody>
</table>

For example, to send a variable-sized array of integers, it is packaged in a structure containing the array and its length:
struct varintarr {
    int *data;
    int arrlnth;
} arr;

Translate the array with xdr_varintarr, as shown in Code Example 4–4.

**CODE EXAMPLE 4–4**  xdr_varintarr Syntax Use

```c
bool_t
xdr_varintarr(xdrsp, arrp)
XDR *xdrsp;
struct varintarr *arrp;
{
    return(xdr_array(xdrsp, (caddr_t)&arrp->data,
                     (u_int *)&arrp->arrlnth, MAXLEN,
                     sizeof(int), xdr_int));
}
```

The arguments of xdr_array are the XDR handle, a pointer to the array, a pointer to the size of the array, the maximum array size, the size of each array element, and a pointer to the XDR routine to translate each array element. If the size of the array is known in advance, use xdr_vector, as shown in Code Example 4–5.

**CODE EXAMPLE 4–5**  xdr_vector Syntax Use

```c
int intarr[SIZE];

bool_t
xdr_intarr(xdrsp, intarr)
XDR *xdrsp;
int intarr[];
{
    return (xdr_vector(xdrsp, intarr, SIZE,
                        sizeof(int),
                        xdr_int));
}
```

XDR converts quantities to 4-byte multiples when serializing. For arrays of characters, each character occupies 32 bits. xdr_bytes packs characters. It has four parameters similar to the first four parameters of xdr_array.

Null-terminated strings are translated by xdr_string. It is like xdr_bytes with no length parameter. On serializing it gets the string length from strlen, and on deserializing it creates a null-terminated string.

Code Example 4–6 calls the built-in functions xdr_string and xdr_reference, which translates pointers to pass a string, and struct simple from the previous examples.
CODE EXAMPLE 4–6  xdr_reference Syntax Use

struct finalexample {
    char *string;
    struct simple *simplep;
} finalexample;

bool_t
xdr_finalexample(xdrsp, finalp)
    XDR *xdrsp;
    struct finalexample *finalp;
{
    if (!xdr_string(xdrsp, &finalp->string,
                    MAXSTRLEN))
        return (FALSE);
    if (!xdr_reference( xdrsp, &finalp->simplep,
                        sizeof(struct simple), xdr_simple))
        return (FALSE);
    return (TRUE);
}

Note that xdr_simple could have been called here instead of xdr_reference.

Standard Interfaces

Interfaces to standard levels of the RPC package provide increasing control over RPC communications. Programs that use this control are more complex. Effective programming at these lower levels requires more knowledge of computer network fundamentals. The top, intermediate, expert, and bottom levels are part of the standard interfaces.

This section shows how to control RPC details by using lower levels of the RPC library. For example, you can select the transport protocol, which can be done at the simplified interface level only through the NETPATH variable. You should be familiar with the TLI in order to use these routines.

The routines shown in Table 4–3 cannot be used through the simplified interface because they require a transport handle. For example, there is no way to allocate and free memory while serializing or deserializing with XDR routines at the simplified interface.
TABLE 4–3  XDR Routines Requiring a Transport Handle

<table>
<thead>
<tr>
<th>Do Not Use With Simplified Interface</th>
</tr>
</thead>
<tbody>
<tr>
<td>clnt_call</td>
</tr>
<tr>
<td>clnt_perrno</td>
</tr>
<tr>
<td>svc_destroy</td>
</tr>
</tbody>
</table>

Top Level Interface

At the top level, the application can specify the type of transport to use but not the specific transport. This level differs from the simplified interface in that the application creates its own transport handles, in both the client and server.

Client

Assume the header file in Code Example 4–7.

**CODE EXAMPLE 4–7  time_prot.h Header File**

```c
/* time_prot.h */
#include <rpc/rpc.h>
#include <rpc/types.h>

struct timev {
  int second;
  int minute;
  int hour;
};
typedef struct timev timev;
bool_t xdr_timev();
#define TIME_PROG ((u_long)0x40000001)
#define TIME_VERS ((u_long)1)
#define TIME_GET ((u_long)1)
```

Code Example 4–8 shows the client side of a trivial date service using top-level service routines. The transport type is specified as an invocation argument of the program.
#include <stdio.h>
#include "time_prot.h"

#define TOTAL (30)
/*
 * Caller of trivial date service
 * usage: calltime hostname
 */

main(argc, argv)
int argc;
char *argv[];
{
struct timeval time_out;
CLIENT *client;
enum clnt_stat stat;
struct timev timev;
char *nettype;

if (argc != 2 && argc != 3) {
    fprintf(stderr,"usage:%s host[nettype]\n",argv[0]);
    exit(1);
}
if (argc == 2)
    nettype = "netpath"; /* Default */
else
    nettype = argv[2];
client = clnt_create(argv[1], TIME_PROG,
    TIME_VERS, nettype);
if (client == (CLIENT *) NULL) {
    clnt_pcreateerror(''Couldn't create client'');
    exit(1);
}
time_out.tv_sec = TOTAL;
time_out.tv_usec = 0;
stat = clnt_call( client, TIME_GET,
    xdr_void, (caddr_t)NULL,
    xdr_timev, (caddr_t)&timev,
    time_out);
if (stat != RPC_SUCCESS) {
    clnt_perror(client, "Call failed");
    exit(1);
}
fprintf(stderr,"%s: %02d:%02d:%02d GMT\n",nettype,timev.hour, timev.minute,
timev.second);
(void) clnt_destroy(client);
exit(0);
}
If `nettype` is not specified in the invocation of the program, the string `netpath` is substituted. When RPC libraries routines encounter this string, the value of the NETPATH environment variable governs transport selection.

If the client handle cannot be created, display the reason for the failure with `clnt_pcreateerror`, or get the error status by reading the contents of the global variable `rpc_createerr`.

After the client handle is created, `clnt_call` is used to make the remote call. Its arguments are the remote procedure number, an XDR filter for the input argument, the argument pointer, an XDR filter for the result, the result pointer, and the time-out period of the call. The program has no arguments, so `xdr_void` is specified. Clean up by calling `clnt_destroy`.

In the above example, if the programmer wished to bound the time allowed for client handle creation to thirty seconds, the call to `clnt_create` should be replaced with a call to `clnt_create_timed` as shown in the following code segment:

```c
struct timeval timeout;
timeout.tv_sec = 30; /* 30 seconds */
timeout.tv_usec = 0;
client = clnt_create_timed(argv[1],
    TIME_PROG, TIME_VERS, nettype,
    &timeout);
```

Code Example 4-9 shows a top-level implementation of a server for the trivial date service.

**CODE EXAMPLE 4–9   Server for Trivial Date Service**

```c
#include <stdio.h>
#include <rpc/rpc.h>
#include "time_prot.h"

static void time_prog();

main(argc,argv)
    int argc;
    char *argv[];
{
    int transeum;
    char *nettype;

    if (argc > 2) {
        fprintf(stderr, "usage: %s [nettype]\n",
            argv[0]);
        exit(1);
    }
    if (argc == 2) {
        ...
    }
    ...
```

(Continued)
nettype = argv[1];
else
    nettype = "netpath"; /* Default */
transpnum =
svc_create(time_prog,TIME_PROG,TIME_VERS,nettype);
if (transpnum == 0) {
    fprintf(stderr,"'\%s: cannot create %s service.'\n",
        argv[0], nettype);
    exit(1);
}
svc_run();
}

/* The server dispatch function */
static void
time_prog(rqstp, transp)
    struct svc_req *rqstp;
    SVCXPRT *transp;
{
    struct timev rslt;
    time_t thetime;
    switch(rqstp->rq_proc) {
    case NULLPROC:
        svc_sendreply(transp, xdr_void, NULL);
        return;
    case TIME_GET:
        break;
    default:
        svcerr_noproc(transp);
        return;
    }
    thetime = time({time_t *} 0);
    rslt.second = thetime % 60;
    thetime /= 60;
    rslt.minute = thetime % 60;
    thetime /= 60;
    rslt.hour = thetime % 24;
    if (!svc_sendreply(transp, xdr_timev, &rslt)) {
        svcerr_systemerr(transp);
    }
}

svc_create returns the number of transports on which it created server handles. time_prog is the service function called by svc_run when a request specifies its program and version numbers. The server returns the results to the client through svc_sendreply.
When rpcgen is used to generate the dispatch function, svc_sendreply is called after the procedure returns, so rslt (in this example) must be declared static in the actual procedure. svc_sendreply is called from inside the dispatch function, so rslt is not declared static.

In this example, the remote procedure takes no arguments. When arguments must be passed, the calls:

```
svc_getargs( SVCXPRT_handle, XDR_filter, argument_pointer);
svc_freeargs( SVCXPRT_handle, XDR_filter argument_pointer);
```

fetch, deserialize (XDR decode), and free the arguments.

**Intermediate Level Interface**

At the intermediate level, the application directly chooses the transport to use.

**Client**

Code Example 4–10 shows the client side of the time service from “Top Level Interface” on page 66, written at the intermediate level of RPC. In this example, the user must name the transport over which the call is made on the command line.

**CODE EXAMPLE 4–10**  Client for Time Service, Intermediate Level

```c
#include <stdio.h>
#include <rpc/rpc.h>
#include <netconfig.h> /* For netconfig structure */
#include "time_prot.h"

#define TOTAL (30)

main(argc,argv)
int argc;
char *argv[];
{
    CLIENT *client;
    struct netconfig *nconf;
    char *netid;
    /* Declarations from previous example */

    if (argc != 3) {
        fprintf(stderr, "usage: %s host netid\n", argv[0]);
        exit(1);
    }

(Continued)
In this example, the netconfig structure is obtained by a call to
getnetconfig (netid). (See the getnetconfig(3N) man page and
Transport Interfaces Programming Guide for more details.) At this level, the
program explicitly selects the network.

In the above example, if the programmer wished to bound the time allowed
for client handle creation to thirty seconds, the call to clnt_tp_create
should be replaced with a call to clnt_tp_create_timed as shown in the
following code segment:

```
struct timeval timeout;
timeout.tv_sec = 30; /* 30 seconds */
timeout.tv_usec = 0;
client = clnt_tp_create_timed(argv[1],
    TIME_PROG, TIME_VERS, nconf,
    &timeout);
```

Server

Code Example 4–11 shows the corresponding server. The command line that
starts the service must specify the transport over which the service is
provided.
/*  * This program supplies Greenwich mean  *  * time to the client that invokes it.  *  * The call format is: server netid  */
#include <stdio.h>
#include <rpc/rpc.h>
#include <netconfig.h> /* For netconfig structure */
#include "time_prot.h"

static void time_prog();

main(argc, argv)
int argc;
char *argv[];
{
SVCXPRT *transp;
struct netconfig *nconf;

if (argc != 2) {
fprintf(stderr, "usage: %s netid\n", argv[0]);
exit(1);
}
if ((nconf = getnetconfigent(argv[1])) == (struct netconfig *) NULL) {
fprintf(stderr, "Could not find info on %s\n", argv[1]);
exit(1);
}
transp = svc_tp_create(time_prog, TIME_PROG,
TIME_VERS, nconf);
if (transp == (SVCXPRT *) NULL) {
fprintf(stderr, "%s: cannot create
%s service\n", argv[0], argv[1]);
exit(1)
}
freenetconfigent(nconf);
svc_run();
}

static
void time_prog(rqstp, transp)
struct svc_req *rqstp;
SVCXPRT *transp;
{
/* Code identical to Top Level version */
Expert Level Interface

At the expert level, network selection is done the same as at the intermediate level. The only difference is in the increased level of control that the application has over the details of the CLIENT and SVCXPRT handles. These examples illustrate this control, which is exercised using the clnt_tli_create and svc_tli_create routines. For more information on TLI, see Transport Interfaces Programming Guide.

Client

Code Example 4-12 shows a version of clntudp_create (the client creation routine for UDP transport) using clnt_tli_create. The example shows how to do network selection based on the family of the transport you choose. clnt_tli_create is used to create a client handle and to:

- Pass an open TLI file descriptor, which may or may not be bound
- Pass the server's address to the client
- Specify the send and receive buffer size

CODE EXAMPLE 4-12  Client for RPC Lower Level

```c
#include <stdio.h>
#include <rpc/rpc.h>
#include <netconfig.h>
#include <netinet/in.h>
/*
 * In earlier implementations of RPC,
 * only TCP/IP and UDP/IP were supported.
 * This version of clntudp_create()
 * is based on TLI/Streams.
 */
CLIENT *
c1ntudp_create(raddr, prog, vers, wait, sockp)
    struct sockaddr_in *raddr; /* Remote address */
    u_long prog; /* Program number */
    u_long vers; /* Version number */
    struct timeval wait; /* Time to wait */
    int *sockp; /* fd pointer */
{
    CLIENT *cl; /* Client handle */
    int madefd = FALSE; /* Is fd opened here */
    int fd = *sockp; /* TLI fd */
    struct t_bind *tbind; /* bind address */
    struct netconfig *nconf; /* netconfig structure */
    void *handlep;

    if ((handlep = setnetconfig()) == (void *) NULL) {
        /* Error starting network configuration */
        rpc_createerr.cf_stat = RPC_UNKNOWNPROTO;
```

(Continued)
return((CLIENT *) NULL);
*/
/*
* Try all the transports until it gets one that is
* connectionless, family is INET, and preferred name is UDP
*/
while (nconf = getnetconfig( handlep)) {
    if ((nconf->nc_semantics == NC_TPI_CLTS) &&
        (strcmp( nconf->nc_protofmly, NC_INET ) == 0) &&
        (strcmp( nconf->nc_proto, NC_UDP ) == 0))
        break;
}
if (nconf == (struct netconfig *) NULL)
    rpc_createerr.cf_stat = RPC_UNKNOWNPROTO;
    goto err;
}  
if (fd == RPC_ANYFD) {
    fd = t_open(nconf->nc_device, O_RDWR, &tinfo);
    if (fd == -1) {
        rpc_createerr.cf_stat = RPC_SYSTEMERROR;
        goto err;
    }
    if (raddr->sin_port == 0) { /* remote addr unknown */
        u_short sport;
/*
* rpcb_getport() is a user-provided routine that calls
* rpcb_getaddr and translates the netbuf address to port
* number in host byte order.
*/
        sport = rpcb_getport(raddr, prog, vers, nconf);
        if (sport == 0) {
            rpc_createerr.cf_stat = RPC_PROGUNAVAIL;
            goto err;
        }
        raddr->sin_port = htons(sport);
    }
/* Transform sockaddr_in to netbuf */
    thbind = (struct t_bind *) t_alloc(fd, T_BIND, T_ADDR);
    if (thbind == (struct t_bind *) NULL)
        rpc_createerr.cf_stat = RPC_SYSTEMERROR;
        goto err;
}  
if (t_bind->addr.maxlen < sizeof( struct sockaddr_in))
    goto err;
(void) memcpy( thbind->addr.buf, (char *)raddr,
    sizeof(struct sockaddr_in));
thbind->addr.len = sizeof(struct sockaddr_in);
/* Bind fd */
    if (t_bind( fd, NULL, NULL) == -1) {
        rpc_createerr.ct_stat = RPC_TLIERROR;
        goto err;
    }

(Continued)
cl = clnt_tli_create(fd, nconf, &tbind->addr, prog, vers,
               tinfo.tsdu, tinfo.tsdu);
/* Close the netconfig file */
(void) endnetconfig( handlep);
(void) t_free((char *) tbind, T_BIND);
if (cl) {
    *sockp = fd;
    if (madefd == TRUE) {
        /* fd should be closed while destroying the handle */
        (void)clnt_control(cl, CLSET_FD_CLOSE, (char *) NULL);
    }
    /* Set the retry time */
    (void) clnt_control(1, CLSET_RETRY_TIMEOUT,
                        (char *) &wait);
    return(cl);
}
err:
if (madefd == TRUE)
    (void) t_close(fd);
(void) endnetconfig(handlep);
return((CLIENT *) NULL);

The network is selected using setnetconfig, getnetconfig, and endnetconfig.

Note - endnetconfig is not called until after the call to clnt_tli_create, near the end of the example.

clnudp_create can be passed an open TLI fd. If passed none (fd == RPC_ANYFD), it opens its own using the netconfig structure for UDP to find the name of the device to pass to t_open.

If the remote address is not known (raddr->sin_port == 0), it is obtained from the remote rpcbind.

After the client handle has been created, you can modify it using calls to clnt_control. The RPC library closes the file descriptor when destroying the handle (as it does with a call to clnt_destroy when it opens the fd itself) and sets the retry time-out period.

Server

Code Example 4–13 shows the server side of Code Example 4–12. It is called svcudp_create. The server side uses svc_tli_create.
svc_tli_create is used when the application needs a fine degree of control, particularly to:

- Pass an open file descriptor to the application.
- Pass the user's bind address.
- Set the send and receive buffer sizes. The fd argument may be unbound when passed in. If it is, then it is bound to a given address, and the address is stored in a handle. If the bind address is set to NULL and the fd is initially unbound, it will be bound to any suitable address.

Use rpcb_set to register the service with rpcbind.

CODE EXAMPLE 4–13  Server for RPC Lower Level

```c
#include <stdio.h>
#include <rpc/rpc.h>
#include <netconfig.h>
#include <netinet/in.h>

SVCXPRT *
svcudp_create(fd)
register int fd;
{
  struct netconfig *nconf;
  SVCXPRT *svc;
  int madefd = FALSE;
  int port;
  void *handlep;
  struct tu_info tinfo;

  /* If no transports available */
  if ((handlep = setnetconfig()) == (void *) NULL) {
    nc_perror("server");
    return((SVCXPRT *) NULL);
  }

  /* Try all the transports until it gets one which is
   * connectionless, family is INET and, name is UDP *
   */
  while (nconf = getnetconfig( handlep)) {
    if ((nconf->nc_semantics == NC_TPI_CLTS) &&
        (strcmp( nconf->nc_protofmly, NC_INET) == 0 ) &&
        (strcmp( nconf->nc_proto, NC_UDP) == 0 ))
      break;
  }

  if (nconf == (struct netconfig *) NULL) {
    endnetconfig(handlep);
    return((SVCXPRT *) NULL);
  }

  if (fd == RPC_ANYFD) {
    fd = t_open(nconf->nc_device, O_RDWR, &tinfo);
    if (fd == -1) {
      (void) endnetconfig();
      return((SVCXPRT *) NULL);
    }

    /* Bind to the port on the device */
    port = nconf->nc_port;
    if (port == 0) {
      port = t_getport(handlep);
      if (port < 0) {
        (void) endnetconfig();
        return((SVCXPRT *) NULL);
      }
    }

    /* Create a bind address */
    char *bind = t_sockaddr(handlep);
    int bindaddrlen;
    void *bindhandlep;
    if ((bindaddrlen = t_sockextend(handlep)) < 0) {
      (void) endnetconfig();
      return((SVCXPRT *) NULL);
    }
    if ((bindhandlep = bindsockaddr(bind, bindaddrlen)) == (void *) NULL) {
      (void) endnetconfig();
      return((SVCXPRT *) NULL);
    }

    /* Set the bind address */
    setbindaddress(handlep);

    /* Bind the transport */
    bindsock(handlep, bindhandlep, port);
    madefd = TRUE;
  }

  /* Create the transport */
  svc = svcudp_create(fd);
  if (svc == (SVCXPRT *) NULL) {
    (void) endnetconfig();
    return((SVCXPRT *) NULL);
  }

  return(svc);
}
```

(Continued)
The network selection here is accomplished similar to `clntudp_create`. The file descriptor is not bound explicitly to a transport address because `svc_tli_create` does that.

`svcudp_create` can use an open `fd`. It will open one itself using the selected `netconfig` structure, if none is provided.

**Bottom Level Interface**

The bottom-level interface to RPC lets the application control all options. `clnt_tli_create` and the other expert-level RPC interface routines are based on these routines. You rarely use these low-level routines.

Bottom-level routines create internal data structures, buffer management, RPC headers, and so on. Callers of these routines, like the expert level routine `clnt_tli_create`, must initialize the `cl_netid` and `cl_tp` fields in the client handle. For a created handle, `cl_netid` is the network identifier (for example `udp`) of the transport and `cl_tp` is the device name of that transport (for example `/dev/udp`). The routines `clnt_dg_create` and `clnt_vc_create` set the `clnt_ops` and `cl_private` fields.

**Client**

Code Example 4–14 shows calls to `clnt_vc_create` and `clnt_dg_create`. 

```c
} else
    t_getinfo(fd, &tinfo);
    svc = svc_tli_create(fd, nconf, (struct t_bind *) NULL,
                        tinfo.tsdu, tinfo.tsdu);
(void) endnetconfig(handlep);
if (svc == (SVCXPRT *) NULL) {
    if (madefd)
        (void) t_close(fd);
    return((SVCXPRT *)NULL);}
return (svc);

The network selection here is accomplished similar to `clntudp_create`. The file descriptor is not bound explicitly to a transport address because `svc_tli_create` does that.

`svcudp_create` can use an open `fd`. It will open one itself using the selected `netconfig` structure, if none is provided.

**Bottom Level Interface**

The bottom-level interface to RPC lets the application control all options. `clnt_tli_create` and the other expert-level RPC interface routines are based on these routines. You rarely use these low-level routines.

Bottom-level routines create internal data structures, buffer management, RPC headers, and so on. Callers of these routines, like the expert level routine `clnt_tli_create`, must initialize the `clnetid` and `cltp` fields in the client handle. For a created handle, `cl_netid` is the network identifier (for example `udp`) of the transport and `cl_tp` is the device name of that transport (for example `/dev/udp`). The routines `clnt_dg_create` and `clnt_vc_create` set the `clnt_ops` and `cl_private` fields.

**Client**

Code Example 4–14 shows calls to `clnt_vc_create` and `clnt_dg_create`. 

```c
} else
    t_getinfo(fd, &tinfo);
    svc = svc_tli_create(fd, nconf, (struct t_bind *) NULL,
                        tinfo.tsdu, tinfo.tsdu);
(void) endnetconfig(handlep);
if (svc == (SVCXPRT *) NULL) {
    if (madefd)
        (void) t_close(fd);
    return((SVCXPRT *)NULL);}
return (svc);
```
CODE EXAMPLE 4–14  Client for Bottom Level

```c
/*
 * variables are:
 * cl: CLIENT *
 * tinfo: struct t_info returned from either t_open or t_getinfo
 * svcaddr: struct netbuf *
 */
switch(tinfo.servtype) {
  case T_COTS:
  case T_COTS_ORD:
    cl = clnt_vc_create(fd, svcaddr, 
                      prog, vers, sendsz, recvsz);
    break;
  case T_CLTS:
    cl = clnt_dg_create(fd, svcaddr, 
                      prog, vers, sendsz, recvsz);
    break;
  default:
    goto err;
}
```

These routines require that the file descriptor is open and bound. `svcaddr` is the address of the server.

Server

Code Example 4–15 is an example of creating a bottom-level server.

CODE EXAMPLE 4–15  Server for Bottom Level

```c
/*
 * variables are:
 * xprt: SVCXPRT *
 */
switch(tinfo.servtype) {
  case T_COTS_ORD:
  case T_COTS:
    xprt = svc_vc_create(fd, sendsz, recvsz);
    break;
  case T_CLTS:
    xprt = svc_dg_create(fd, sendsz, recvsz);
    break;
  default:
    goto err;
}
```

Server Caching

`svc_dg_enablecache` initiates service caching for datagram transports. Caching should be used only in cases where a server procedure is a "once
only" kind of operation, because executing a cached server procedure multiple times will yield different results.

```c
svc_dg_enablecache(xprt, cache_size)
SVCXprt *xprt;
unsigned long cache_size;
```

This function allocates a duplicate request cache for the service endpoint `xprt`, large enough to hold `cache_size` entries. A duplicate request cache is needed if the service contains procedures with varying results. Once enabled, there is no way to disable caching.

**Low-Level Data Structures**

The following are for reference only. The implementation may change.

First is the client RPC handle, defined in `<rpc/clnt.h>`. Low-level implementations must provide and initialize one handle per connection, as shown in Code Example 4–16.

**CODE EXAMPLE 4–16 RPC Client Handle Structure**

```c
typedef struct {
  AUTH *cl_auth; /* authenticator */
  struct clnt_ops {
    enum clnt_stat (*cl_call)(); /* call remote procedure */
    void (*cl_abort)(); /* abort a call */
    void (*cl_geterr)(); /* get specific error code */
    bool_t (*cl_freeres)(); /* frees results */
    void (*cl_destroy)(); /* destroy this structure */
    bool_t (*cl_control)(); /* the ioctl() of rpc */
  } *cl_ops;
  caddr_t cl_private; /* private stuff */
  char *cl_netid; /* network token */
  char *cl_tp; /* device name */
} CLIENT;
```

The first field of the client-side handle is an authentication structure, defined in `<rpc/auth.h>`. By default, it is set to `AUTH_NONE`. A client program must initialize `cl_auth` appropriately, as shown in Code Example 4–17.

**CODE EXAMPLE 4–17 Client Authentication Handle**

```c
typedef struct {
  struct opaque_auth ah_cred;
  struct opaque_auth ah_verf;
  union des_block ah_key;
  struct auth_ops {
    void (*ah_nextverf)();
    int (*ah_marshall)(); /* nextverf & serialize */
    int (*ah_validate)(); /* validate varifier */
    int (*ah_refresh)(); /* refresh credentials */
    void (*ah_destroy)(); /* destroy this structure */
  }
}...
```
In the AUTH structure, ah_cred contains the caller's credentials, and ah_verf contains the data to verify the credentials. See “Authentication” on page 91 for details.

Code Example 4–18 shows the server transport handle.

**CODE EXAMPLE 4–18  Server Transport Handle**

```c
typedef struct {
    int xp_fd;
    #define xp_sock xp_fd
    u_short xp_port; /* associated port number. Obsoleted */
    struct xp_ops {
        bool_t (*xp_recv)(); /* receive incoming requests */
        enum xprt_stat (*xp_stat)(); /* get transport status */
        bool_t (*xp_getargs)(); /* get arguments */
        bool_t (*xp_reply)(); /* send reply */
        bool_t (*xp_freeargs)(); /* free mem alloc for args */
        void (*xp_destroy)(); /* destroy this struct */
    } *xp_ops;
    int xp_addrlen; /* length of remote addr. Obsolete */
    char *xp_tp; /* transport provider device name */
    char *xp_netid; /* network token */
    struct netbuf xp_ltaddr; /* local transport address */
    struct netbuf xp_rtaddr; /* remote transport address */
    char xp_raddr[16]; /* remote address. Now obsoleted */
    caddr_t xp_p1; /* private: for use by svc ops */
    caddr_t xp_p2; /* private: for use by svc ops */
    caddr_t xp_p3; /* private: for use by svc lib */
} SVCXPRT;
```

Table 4–4 shows the fields for the server transport handle.

**TABLE 4–4  RPC Server Transport Handle Fields**

<table>
<thead>
<tr>
<th>Field</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>xp_fd</td>
<td>The file descriptor associated with the handle. Two or more server handles can share the same file descriptor.</td>
</tr>
<tr>
<td>xp_netid</td>
<td>The network identifier (e.g. udp) of the transport on which the handle is created and xp_fp is the device name associated with that transport.</td>
</tr>
<tr>
<td>xp_ltaddr</td>
<td>The server's own bind address.</td>
</tr>
</tbody>
</table>
TABLE 4–4  RPC Server Transport Handle Fields  (continued)

<table>
<thead>
<tr>
<th>Field</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>xp_rtaddr</td>
<td>The address of the remote caller (and so may change from call to call).</td>
</tr>
<tr>
<td>xp_netid</td>
<td>Initialized by svc_tli_create and other expert-level routines.</td>
</tr>
<tr>
<td>xp_tp</td>
<td></td>
</tr>
<tr>
<td>xp_ltaddr</td>
<td></td>
</tr>
</tbody>
</table>

The rest of the fields are initialized by the bottom-level server routines svc_dg_create and svc_vc_create.

For connection-oriented endpoints, the fields in Table 4–5 are not valid until a connection has been requested and accepted for the server:

TABLE 4–5  RPC Connection-Oriented Endpoints

<table>
<thead>
<tr>
<th>Fields Not Valid Until Connection Is Accepted</th>
</tr>
</thead>
<tbody>
<tr>
<td>xp_fd</td>
</tr>
<tr>
<td>xp_p2</td>
</tr>
<tr>
<td>xp_ltaddr</td>
</tr>
</tbody>
</table>

Testing Programs Using Low-level Raw RPC

There are two pseudo-RPC interface routines that bypass all the network software. The routines shown in Code Example on page 82, clnt_raw_create and svc_raw_create, do not use any real transport.

Note - Do not use raw mode on production systems. Raw mode is intended as a debugging aid only. Raw mode is not MT safe.

Code Example on page 82 is compiled and linked using the following Makefile:
CODE EXAMPLE 4–19  Simple Program Using Raw RPC

/*
 * A simple program to increment a number by 1
 */

#include <stdio.h>
#include <rpc/rpc.h>
#include <rpc/raw.h>
define prognum 0x40000001
define versnum 1
define INCR 1

struct timeval TIMEOUT = {0, 0};
static void server();

main (argc, argv)
int argc;
char **argv;
{
CLIENT *cl;
SVCXPRTR *svc;
int num = 0, ans;
int flag;

if (argc == 2) 
    num = atoi(argv[1]);
svc = svc_raw_create();
if (svc == (SVCXPRT *) NULL) {
    fprintf(stderr, "Could not create server handle
");
exit(1);
}
flag = svc_reg( svc, prognum, versnum, server,
    (struct netconfig *) NULL );
if (flag == 0) {
    fprintf(stderr, "Error: svc_reg failed.
");
exit(1);
}
cl = clnt_raw_create( prognum, versnum );
if (cl == (CLIENT *) NULL) {
    clnt_pcreateerror("Error: clnt_raw_create");
exit(1);
}
if (clnt_call(cl, INCR, xdr_int, (caddr_t) &num, xdr_int,
    (caddr_t) &ans, TIMEOUT)
!= RPC_SUCCESS) {
    clnt_perror(cl, "Error: client_call with raw");
exit(1);
}
printf("Client: number returned \d\n", ans);

(Continued)
exit(0);
}

static void
server(rqstp, transp)
struct svc_req *rqstp;
SVCXPRT *transp;
{
  int num;

  fprintf(stderr, "Entering server procedure.\n");

  switch(rqstp->rq_proc) {
    case NULLPROC:
      if (svc_sendreply( transp, xdr_void,
        (caddr_t) NULL) == FALSE) {
        fprintf(stderr, "error in null proc\n");
        exit(1);
      }
      return;
    case INCR:
      break;
    default:
      svcerr_noproc(transp);
      return;
  }

  if (!svc_getargs( transp, xdr_int, &num)) {
    svcerr_decode(transp);
    return;
  }

  fprintf(stderr, "Server procedure: about to increment.\n");
  num++;
  if (svc_sendreply( transp, xdr_int, &num) == FALSE) {
    fprintf(stderr, "error in sending answer\n");
    exit (1);
  }
  fprintf(stderr, "Leaving server procedure.\n");
}

Note the following points in Code Example on page 82:

- The server must be created before the client.
- svc_raw_create has no parameters.
- The server is not registered with rpcbind. The last parameter to svc_reg is (struct netconfig *) NULL, which means that it will not be registered with rpcbind.
- svc_run is not called.
- All the RPC calls occur within the same thread of control.
The server-dispatch routine is the same as for normal RPC servers.

Advanced RPC Programming Techniques

This section addresses areas of occasional interest to developers using the lower level interfaces of the RPC package. The topics are:

- **poll on the server**— how a server can call the dispatcher directly if calling `svc_run` is not feasible
- Broadcast RPC — how to use the broadcast mechanisms
- Batching — how to improve performance by batching a series of calls
- Authentication — what methods are available in this release
- Port monitors — how to interface with the `inetd` and `listener` port monitors
- Multiple program versions — how to service multiple program versions

**poll on the Server Side**

This section applies only to servers running RPC in single-threaded (default) mode.

A process that services RPC requests and performs some other activity cannot always call `svc_run`. If the other activity periodically updates a data structure, the process can set a `SIGALRM` signal before calling `svc_run`. This allows the signal handler to process the data structure and return control to `svc_run` when done.

A process can bypass `svc_run` and access the dispatcher directly with the `svc_getreqset` call. Given the file descriptors of the transport endpoints associated with the programs being waited on, the process can have its own `poll` that waits on both the RPC file descriptors and its own descriptors.

Code Example 4-20 shows `svc_run`. `svc_pollset` is an array of `pollfd` structures that is derived, through a call to `__rpc_select_to_poll`, from `svc_fdset`. The array can change every time any RPC library routine is called, because descriptors are constantly being opened and closed. `svc_getreq_poll` is called when `poll` determines that an RPC request has arrived on some RPC file descriptors.
The functions __rpc_dtbsize and __rpc_select_to_poll are not part of the SVID, but they are available in the libnsl library. The descriptions of these functions are included here so that you may create versions of these functions for non-Solaris implementations.

```c
int __rpc_select_to_poll(int fdmax, fd_set *fdset,
                        struct pollfd *pollset)
```

Given an `fd_set` pointer and the number of bits to check in it, this function initializes the supplied `pollfd` array for RPC’s use. RPC polls only for input events. The number of `pollfd` slots that were initialized is returned.

```c
int __rpc_dtbsize()
```

This function calls the `getrlimit` function to determine the maximum value that the system may assign to a newly created file descriptor. The result is cached for efficiency.

For more information on the SVID routines in this section, see the `rpc_svc_calls(3N)` and the `poll(2)` man pages.

**CODE EXAMPLE 4–20** svc_run() and poll()

```c
void
svc_run()
{
    int nfds;
    int dtbsize = __rpc_dtbsize();
    int i;
    struct pollfd svc_pollset[fd_setsize];

    for (i = 0; i < nfds;)
        /*
        * Check whether there is any server fd on which we may have
        * to wait.
        */
        nfds = __rpc_select_to_poll(dtbsize, &svc_fdset,
                                    svc_pollset);
    if (nfds == 0)
        break; /* None waiting, hence quit */

    switch (i = poll(svc_pollset, nfds, -1)) {
    case -1:
        /*
        * We ignore all errors, continuing with the assumption
        * that it was set by the signal handlers (or any
        * other outside event) and not caused by poll().
        */
    case 0:
        continue;
    default:
        svc_getreq_poll(svc_pollset, i);
    
```
Broadcast RPC

When an RPC broadcast is issued, a message is sent to all rpcbind daemons on a network. An rpcbind daemon with which the requested service is registered forwards the request to the server. The main differences between broadcast RPC and normal RPC calls are:

- Normal RPC expects one answer; broadcast RPC expects many answers (one or more answer from each responding machine).
- Broadcast RPC works only on connectionless protocols that support broadcasting, such as UDP.
- With broadcast RPC, all unsuccessful responses are filtered out; so, if there is a version mismatch between the broadcaster and a remote service, the broadcaster never hears from the service.
- Only datagram services registered with rpcbind are accessible through broadcast RPC; service addresses may vary from one host to another, so rpc_broadcast sends messages to rpcbind’s network address.
- The size of broadcast requests is limited by the maximum transfer unit (MTU) of the local network; the MTU for Ethernet is 1500 bytes.

Code Example 4–21 shows how rpc_broadcast is used and describes its arguments.

**CODE EXAMPLE 4–21**  RPC Broadcast

```c
/*
 * bcast.c: example of RPC broadcasting use.
 */
#include <stdio.h>
#include <rpc/rpc.h>
main(argc, argv)
int argc;
char *argv[];
{
    enum clnt_stat rpc_stat;
    int argc;
    char *argv[];
    { 
        enum clint_stat rpc_stat;
        u_long prognum, vers;
        struct rpcent *re;

        if(argc != 3) {
            fprintf(stderr, "usage : %s RPC_PROG VERSION\n", argv[0]);
            exit(1);
        }
    }
(Continued)
```

86  ONC+ Developer’s Guide ♦ August, 1997
if (isdigit((argv[1])))
    prognum = atoi(argv[1]);
else {
    re = getrpcbyname(argv[1]);
    if (! re) {
        fprintf(stderr, "Unknown RPC service \%s\n", argv[1]);
        exit(1);
    }
    prognum = re->r_number;
}
    vers = atoi(argv[2]);
    rpc_stat = rpc_broadcast(prognum, vers, NULLPROC, xdr_void,
    (char *)NULL, xdr_void, (char *)NULL, bcast_proc,
    NULL);
    if ((rpc_stat != RPC_SUCCESS) && (rpc_stat != RPC_TIMEDOUT)) {
        fprintf(stderr, "broadcast failed: \%s\n",
        clnt_sperrno(rpc_stat));
        exit(1);
    }
    exit(0);
}

The function in Code Example 4–22 collects replies to the broadcast. Normal operation is to collect either the first reply or all replies. bcast_proc displays the IP address of the server that has responded. Since the function returns FALSE it will continue to collect responses, and the RPC client code will continue to resend the broadcast until it times out.

CODE EXAMPLE 4–22 Collect Broadcast Replies

bool_t
bcast_proc(res, t_addr, nconf)
void *res; /* Nothing comes back */
struct t_bind *t_addr; /* Who sent us the reply */
struct netconfig *nconf;
{
    register struct hostent *hp;
    char *naddr;
    naddr = taddr2naddr(nconf, &taddr->addr);
    if (naddr == (char *) NULL) {
        fprintf(stderr,"Responded: unknown\n");
    } else {
        fprintf(stderr,"Responded: \%s\n", naddr);
        free(naddr);
    }
    return(FALSE);
}
If done is TRUE, then broadcasting stops, and rpc_broadcast returns successfully. Otherwise, the routine waits for another response. The request is rebroadcast after a few seconds of waiting. If no responses come back, the routine returns with RPC_TIMEDOUT.

Batching

RPC is designed so that clients send a call message and wait for servers to reply to the call. This implies that a client is blocked while the server processes the call. This is inefficient when the client does not need each message acknowledged.

RPC batching lets clients process asynchronously. RPC messages can be placed in a pipeline of calls to a server. Batching requires that:

- The server does not respond to any intermediate message.
- The pipeline of calls is transported on a reliable transport, such as TCP.
- The result’s XDR routine in the calls must be NULL.
- The RPC call’s time-out must be zero.

Because the server does not respond to each call, the client can send new calls in parallel with the server processing previous calls. The transport can buffer many call messages and send them to the server in one write system call. This decreases interprocess communication overhead and the total time of a series of calls. The client should end with a nonbatched call to flush the pipeline.

shows the unbatched version of the client. It scans the character array, buf, for delimited strings and sends each string to the server.

**CODE EXAMPLE 4–23 Unbatched Client**

```c
#include <stdio.h>
#include <rpc/rpc.h>
#include "windows.h"

main(argc, argv)
int argc;
char **argv;
{
    struct timeval total_timeout;
    register CLIENT *client;
    enum clnt_stat clnt_stat;
    char buf[1000], *s = buf;

    if ((client = clnt_create( argv[1], WINDOWPROG, WINDOWVERS,
        "circuit_v") == (CLIENT *) NULL) {
```

(Continued)
clnt_pcreateerror("clnt_create");
exit(1);}

total_timeout.tv_sec = 20;
total_timeout.tv_usec = 0;
while (scanf( "%s", s ) != EOF) {
    if (clnt_call(client, RENDERSTRING, xdr_wrapstring, &s,
                 xdr_void, (caddr_t) NULL, total_timeout) != RPC_SUCCESS) {
        clnt_perror(client, "rpc");
        exit(1);
    }
}
clnt_destroy( client );
exit(0);

shows the batched version of the client. It does not wait after each string is
sent to the server. It waits only for an ending response from the server.

CODE EXAMPLE 4–24  Batched Client

#include <stdio.h>
#include <rpc/rpc.h>
#include "windows.h"

main(argc, argv)
    int argc;
    char **argv;
{
struct timeval total_timeout;
    register CLIENT *client;
enum clnt_stat clnt_stat;
    char buf[1000], *s = buf;

    if ((client = clnt_create( argv[1], WINDOWPROG, WINDOWVERS,
                               "circuit_v") == (CLIENT *) NULL) {
        clnt_pcreateerror("clnt_create");
        exit(1);
    }
    timerclear(&total_timeout);
    while (scanf("%s", s) != EOF)
        clnt_call(client, RENDERSTRING_BATCHED, xdr_wrapstring,
                  xdr_void, (caddr_t) NULL, total_timeout);
    /* Now flush the pipeline */
total_timeout.tv_sec = 20;
clnt_stat = clnt_call(client, NULLPROC, xdr_void, 
(caddr_t) NULL, xdr_void, (caddr_t) NULL, 
total_timeout);
if (clnt_stat != RPC_SUCCESS) {
    clnt_perror(client, "rpc");
    exit(1);
}
clint_destroy(client);
exit(0);

Code Example 4–25 shows the dispatch portion of the batched server. Because the server sends no message, the clients are not notified of failures.

CODE EXAMPLE 4–25  Batched Server

```c
#include <stdio.h>
#include <rpc/rpc.h>
#include "windows.h"

void
windowdispatch(rqstp, transp)
struct svc_req *rqstp;
SVCXPRT *transp;
{
    char *s = NULL;

    switch(rqstp->rq_proc) {
        case NULLPROC:
            if (!svc_sendreply( transp, xdr_void, NULL))
                fprintf(stderr, "can't reply to RPC call\n");
            return;
        case RENDERSTRING:
            if (!svc_getargs( transp, xdr_wrapstring, &s)) {
                fprintf(stderr, "can't decode arguments\n");
                /* Tell caller an error occurred */
                svcerr_decode(transp);
                break;
            }
            /* Code here to render the string s */
            if (!svc_sendreply( transp, xdr_void, (caddr_t) NULL))
                fprintf( stderr, "can't reply to RPC call\n");
            break;
        case RENDERSTRING_BATCHED:
            if (!svc_getargs(transp, xdr_wrapstring, &s)) {
                fprintf(stderr, "can't decode arguments\n");
                /* Code here to render the string s */
                if (!svc_sendreply( transp, xdr_void, (caddr_t) NULL))
                    fprintf(stderr, "can't reply to RPC call\n");
                break;
            }
            /* Code here to render the string s */
            if (!svc_sendreply( transp, xdr_void, (caddr_t) NULL))
                fprintf(stderr, "can't reply to RPC call\n");
            break;
    }
}
```
/* Be silent in the face of protocol errors */
break;

/* Code here to render string s, but send no reply! */
break;
default:
    svcerr_noproc(transp);
    return;

/* Now free string allocated while decoding arguments */
svc_freeargs(transp, xdr_wrapstring, &s);
}

Batching Performance

To illustrate the benefits of batching, the examples in , , and Code Example 4–25 were completed to render the lines in a 25144-line file. The rendering service simply throws the lines away. The batched version of the application was four times as fast as the unbatched version.

Authentication

In all of the preceding examples in this chapter, the caller has not identified itself to the server, and the server has not required identification of the caller. Some network services, such as a network file system, require caller identification. Refer to System Administration Guide, to implement any of the authentication methods described in this section.

Just as different transports can be used when creating RPC clients and servers, different “flavors” of authentication can be associated with RPC clients. The authentication subsystem of RPC is open ended. So, many flavors of authentication can be supported. The authentication protocols are further defined in Appendix B.”

Sun RPC currently supports the authentication flavors shown in .
<table>
<thead>
<tr>
<th>Authentication Method</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>AUTH_NONE</td>
<td>Default. No authentication performed</td>
</tr>
<tr>
<td>AUTH_SYS</td>
<td>An authentication flavor based on UNIX operating system, process permissions authentication</td>
</tr>
<tr>
<td>AUTH_SHORT</td>
<td>An alternate flavor of AUTH_SYS used by some servers for efficiency. Client programs using AUTH_SYS authentication can receive AUTH_SHORT response verifiers from some servers. See Appendix B for details</td>
</tr>
<tr>
<td>AUTH_DES</td>
<td>An authentication flavor based on DES encryption techniques</td>
</tr>
<tr>
<td>AUTH_KERB</td>
<td>Version 4 Kerberos authentication based on DES framework</td>
</tr>
</tbody>
</table>

When a caller creates a new RPC client handle as in:
```c
clnt = clnt_create(host, prognum, versnum, nettype);
```

the appropriate client-creation routine sets the associated authentication handle to:
```c
clnt->cl_auth = authnone_create();
```

If you create a new instance of authentication, you must destroy it with `auth_destroy(clnt->cl_auth)`. This should be done to conserve memory.

On the server side, the RPC package passes the service-dispatch routine a request that has an arbitrary authentication style associated with it. The request handle passed to a service-dispatch routine contains the structure `rq_cred`. It is opaque, except for one field: the flavor of the authentication credentials.

```c
/*
 * Authentication data
 */
struct opaque_auth {
    enum_t oa_flavor; /* style of credentials */
    caddr_t oa_base;  /* address of more auth stuff */
    u_int  oa_length; /* not to exceed MAX_AUTH_BYTES */
};
```

The RPC package guarantees the following to the service-dispatch routine:

- The `rq_cred` field in the `svc_req` structure is well formed. So, you can check `rq_cred.oa_flavor` to get the flavor of authentication. You can
also check the other fields of \texttt{rq\_cred} if the flavor is not supported by RPC.

- The \texttt{rq\_clntcred} field that is passed to service procedures is either \texttt{NULL} or points to a well-formed structure that corresponds to a supported flavor of authentication credential. There is no authentication data for the \texttt{AUTH\_NONE} flavor. \texttt{rq\_clntcred} can be cast only as a pointer to an \texttt{authsys_parms}, \texttt{short\_hand\_verf}, \texttt{authkerb\_cred}, or \texttt{authdes\_cred} structure.

**AUTH\_SYS Authentication**

The client can use \texttt{AUTH\_SYS} (called \texttt{AUTH\_UNIX} in previous releases) style authentication by setting \texttt{clnt\textgreater\textgreater cl\_auth} after creating the RPC client handle:

\begin{verbatim}
crnt\textgreater\textgreater cl\_auth = authsys\_create\_default();
\end{verbatim}

This causes each RPC call associated with \texttt{clnt} to carry with it the following credentials-authentication structure shown in .

**CODE EXAMPLE 4–26**  
**AUTH\_SYS Credential Structure**

```c
/*
 * AUTH\_SYS flavor credentials.
 */
struct authsys_parms {
    u_long aup_time;  /* credentials creation time */
    char *aup_machname;  /* client’s host name */
    uid_t aup_uid;  /* client’s effective uid */
    gid_t aup_gid;  /* client’s current group id */
    u_int aup_len;  /* element length of aup\_gids*/
    gid_t *aup_gids;  /* array of groups user is in */
}:
```

\texttt{rpc\_broadcast} defaults to \texttt{AUTH\_SYS} authentication.

Code Example 4–27 shows a server, with procedure \texttt{RUSERPROC\_1}, that returns the number of users on the network. As an example of authentication, it checks \texttt{AUTH\_SYS} credentials and does not service requests from callers whose \texttt{uid} is 16.

**CODE EXAMPLE 4–27**  
**Authentication Server**

```c
nuser(rqstp, transp)
    struct svc\_req *rqstp;
    SVCXPRT *transp;
{
    struct authsys\_parms *sys\_cred;
    uid\_t uid;
    unsigned long nusers;

    /* NULLPROC should never be authenticated */
```
if (rqstp->rq_proc == NULLPROC) {
    if (!svc_sendreply( transp, xdr_void, (caddr_t) NULL))
        fprintf(stderr, "can’t reply to RPC call\n");
    return;
}

/* now get the uid */
switch(rqstp->rq_cred.oa_flavor) {
    case AUTH_SYS:
        sys_cred = (struct authsys_parms *) rqstp->rq_clntcred;
        uid = sys_cred->aup_uid;
        break;
    default:
        svcerr_weakauth(transp);
        return;
}

switch(rqstp->rq_proc) {
    case RUSERSPROC_1:
        /* make sure caller is allowed to call this proc */
        if (uid == 16) {
            svcerr_systemerr(transp);
            return;
        }
        /*
         * Code here to compute the number of users and assign it
         * to the variable nusers
         */
        if (!svc_sendreply( transp, xdr_u_long, &nusers))
            fprintf(stderr, "can’t reply to RPC call\n");
        return;
    default:
        svcerr_noproc(transp);
        return;
}
}

Note the following:

- The authentication parameters associated with the NULLPROC (procedure number zero) are usually not checked.

- The server calls svcerr_weakauth if the authentication parameter’s flavor is too weak; there is no way to get the list of authentication flavors the server requires.

- The service protocol should return status for access denied; in Code Example 4–27, the protocol calls the service primitive svcerr_systemerr, instead.

The last point underscores the relation between the RPC authentication package and the services: RPC deals only with authentication and not with an individual service’s access control. The services themselves must establish access-control policies and reflect these policies as return statuses in their protocols.
AUTH_DES Authentication

Use AUTH_DES authentication for programs that require more security than
AUTH_SYS provides. AUTH_SYS authentication is easy to defeat. For
example, instead of using authsys_create_default, a program can call
authsys_create and change the RPC authentication handle to give itself
any desired user ID and hostname.

AUTH_DES authentication requires that keyserv daemons are running on
both the server and client hosts. The NIS or NIS+ naming service must also
be running. Users on these hosts need public/private key pairs assigned by
the network administrator in the publickey database. They must also have
decrypted their secret keys with the keylogin command (normally done by
login unless the login password and secure-RPC password differ).

To use AUTH_DES authentication, a client must set its authentication handle
appropriately. For example:

cl->cl_auth = authdes_seccreate(servername, 60, server,
(char *)NULL);

The first argument is the network name or “netname” of the owner of the
server process. Server processes are usually root processes, and you can get
their netnames with the following call:

char servername[MAXNETNAMELEN];
host2netname(servername, server, (char *)NULL);

servername points to the receiving string and server is the name of the host
the server process is running on. If the server process was run by a non-root
user, use the call user2netname as follows:

char servername[MAXNETNAMELEN];
user2netname(servername, serveruid(), (char *)NULL);

serveruid is the user id of the server process. The last argument of both
functions is the name of the domain that contains the server. NULL means
“use the local domain name.”

The second argument of authdes_seccreate is the lifetime (known also as
the window) of the client’s credential, here, 60 seconds. A credential will
expire 60 seconds after the client makes an RPC call. If a program tries to
reuse the credential, the server RPC subsystem recognizes that it has expired
and does not service the request carrying the expired credential. If any
program tries to reuse a credential within its lifetime, it is rejected, because
the server RPC subsystem saves credentials it has seen in the near past and
does not serve duplicates.

The third argument of authdes_seccreate is the name of the timehost
used to synchronize clocks. AUTH_DES authentication requires that server
and client agree on the time. The example specifies to synchronize with the
server. A (char *) NULL says not to synchronize. Do this only when you are sure that the client and server are already synchronized.

The fourth argument of authdes_seccreate points to a DES encryption key to encrypt time stamps and data. If this argument is (char *) NULL, as it is in this example, a random key is chosen. The ah_key field of the authentication handle contains the key.

The server side is simpler than the client. Code Example 4–28 shows the server in Code Example 4–27 changed to use AUTH_DES.

**CODE EXAMPLE 4–28  AUTH_DES Server**

```c
#include <rpc/rpc.h>
...
...
server(rqstp, transp)
struct svc_req *rqstp;
SVCXPRT *transp;
{
    struct authdes_cred *des_cred;
    uid_t uid;
    gid_t gid;
    int gidlen;
    gid_t gidlist[10];

    /* NULLPROC should never be authenticated */
    if (rqstp->rq_proc == NULLPROC) {
        /* same as before */
    }
    /* now get the uid */
    switch(rqstp->rq_cred.oa_flavor) {
    case AUTH_DES:
        des_cred = (struct authdes_cred *) rqstp->rq_clntcred;
        if (! netname2user( des_cred->adc_fullname.name, &uid,
                       &gid, &gidlen, gidlist)) {
            fprintf(stderr, "unknown user: %s
",
                       des_cred->adc_fullname.name);
            svcerr_systemerr(transp);
            return;
        }
        break;
    default:
        svcerr_weakauth(transp);
        return;
    }
    /* The rest is the same as before */
```

Note the routine netname2user converts a network name (or “netname” of a user) to a local system ID. It also supplies group IDs (not used in this example).
AUTH_KERB Authentication

SunOS 5.x includes support for most client-side features of Kerberos 4.0, except klogin. AUTH_KERB is conceptually similar to AUTH_DES; the essential difference is that DES passes a network name and DES-encrypted session key, while Kerberos passes the encrypted service ticket. The other factors that affect implementation and interoperability are given in the following subsections.

For more information, see the kerberos(3N) man page and the Steiner-Neuman-Shiller paper\(^1\) on the MIT Project Athena implementation of Kerberos. You may access MIT documentation through the FTP directory /pub/kerberos/doc on athena-dist.mit.edu, or through Mosaic, using the document URL, ftp://athena-dist.mit.edu/pub/kerberos/doc.

Time Synchronization

Kerberos uses the concept of a time window in which its credentials are valid. It does not place restrictions on the clocks of the client or server. The client is required to determine the time bias between itself and the server and compensate for the difference by adjusting the window time specified to the server. Specifically, the window is passed as an argument to authkerb_seccreate; the window does not change. If a timehost is specified as an argument, the client side gets the time from the timehost and alters its timestamp by the difference in time. Various methods of time synchronization are available. See the kerberos_rpc(3N) man page for more information.

Well-Known Names

Kerberos users are identified by a primary name, instance, and realm. The RPC authentication code ignores the realm and instance, while the Kerberos library code does not. The assumption is that user names are the same between client and server. This enables a server to translate a primary name into user identification information. Two forms of well-known names are used (omitting the realm):

- root\(\_\)host represents a privileged user on client host.
- user\_ignored represents the user whose user name is user. The instance is ignored.

Encryption

Kerberos uses cipher block chaining (CBC) mode when sending a full name credential (one that includes the ticket and window), and electronic code book (ECB) mode otherwise. CBC and ECB are two methods of DES encryption. See the des_crypt(3) man page for more information. The session key is used as the initial input vector for CBC mode. The notation

\[ \text{xdr\_type(object)} \]

means that XDR is used on object as a type. The length in the next code section is the size, in bytes of the credential or verifier, rounded up to 4-byte units. The full name credential and verifier are obtained as follows:

\begin{align*}
\text{xdr\_long(timestamp.seconds)} \\
\text{xdr\_long(timestamp.useconds)} \\
\text{xdr\_long(window)} \\
\text{xdr\_long(window - 1)}
\end{align*}

After encryption with CBC with input vector equal to the session key, the output is two DES cipher blocks:

\begin{align*}
\text{CB0} \\
\text{CB1.low} \\
\text{CB1.high}
\end{align*}

The credential is:

\begin{align*}
\text{xdr\_long(AUTH\_KERB)} \\
\text{xdr\_long(length)} \\
\text{xdr\_enum(AN\_FULLNAME)} \\
\text{xdr\_bytes(ticket)} \\
\text{xdr\_opaque(CB1.high)}
\end{align*}

The verifier is:

\begin{align*}
\text{xdr\_long(AUTH\_KERB)} \\
\text{xdr\_long(length)} \\
\text{xdr\_opaque(CB0)} \\
\text{xdr\_opaque(CB1.low)}
\end{align*}

The nickname exchange yields:

\begin{align*}
\text{xdr\_long(timestamp.seconds)} \\
\text{xdr\_long(timestamp.useconds)}
\end{align*}

The nickname is encrypted with ECB to obtain ECB0, and the credential is:

\begin{align*}
\text{xdr\_long(AUTH\_KERB)} \\
\text{xdr\_long(length)} \\
\text{xdr\_enum(AN\_NICKNAME)} \\
\text{xdr\_opaque(akc\_nickname)}
\end{align*}
The verifier is:

\begin{verbatim}
xdr_long(AUTH_KERB)
xdr_long(length)
xdr_opaque(ECB0)
xdr_opaque(0)
\end{verbatim}

**Using Port Monitors**

RPC servers can be started by port monitors such as `inetd` and `listen`. Port monitors listen for requests and spawn servers in response. The forked server process is passed file descriptor 0 on which the request has been accepted. For `inetd`, when the server is done, it may exit immediately or wait a given interval for another service request.

For `listen`, servers should exit immediately after replying because `listen` always spawns a new process. The following function call creates a `SVCXPRT` handle to be used by the services started by port monitors:

\begin{verbatim}
transp = svc_tli_create(0, nconf, (struct t_bind *)NULL, 0, 0)
\end{verbatim}

`nconf` is the `netconfig` structure of the transport from which the request is received.

Because the port monitors have already registered the service with `rpcbind`, there is no need for the service to register with `rpcbind`. But it must call `svc_reg` to register the service procedure:

\begin{verbatim}
svc_reg(transp, PROGNUM, VERSNUM, dispatch,(struct netconfig *)NULL)
\end{verbatim}

The `netconfig` structure here is NULL to prevent `svc_reg` from registering the service with `rpcbind`.

---

**Note** - Study `rpcgen`-generated server stubs to see the sequence in which these routines are called.

For connection-oriented transports, the following routine provides a lower level interface:

\begin{verbatim}
transp = svc_fd_create(0, recvsize, sendsize);
\end{verbatim}

A 0 file descriptor is the first argument. You can set the value of `recvsize` and `sendsize` to any appropriate buffer size. A 0 for either argument causes a system default size to be chosen. Application servers that do not do any listening of their own use `svc_fd_create`. 

The Programmer's Interface to RPC 99
Using **inetd**

Entries in `/etc/inet/inetd.conf` have different formats for socket-based, TLI-based, and RPC services. The format of `inetd.conf` entries for RPC services is:

```
rpc_prog/vers endpoint_type rpc/proto flags user pathname args
```

where:

**TABLE 4-7** RPC inetd Services

<table>
<thead>
<tr>
<th>rpc_prog/vers</th>
<th>The name of an RPC program followed by a <code>/</code> and the version number or a range of version numbers.</th>
</tr>
</thead>
<tbody>
<tr>
<td>endpoint_type</td>
<td>One of <code>dgram</code> (for connectionless sockets), <code>stream</code> (for connection mode sockets), or <code>tli</code> (for TLI endpoints).</td>
</tr>
<tr>
<td>proto</td>
<td>May be <code>*</code> (for all supported transports), a nettype, a netid, or a comma separated list of nettype and netid.</td>
</tr>
<tr>
<td>flags</td>
<td>Either <code>wait</code> or <code>nowait</code>.</td>
</tr>
<tr>
<td>user</td>
<td>Must exist in the effective passwd database.</td>
</tr>
<tr>
<td>athname</td>
<td>Full path name of the server daemon.</td>
</tr>
<tr>
<td>args</td>
<td>Arguments to be passed to the daemon on invocation.</td>
</tr>
</tbody>
</table>

For example:

```
rquotad/1 tli rpc/udp wait root /usr/lib/nfs/rquotad rquotad
```

For more information, see the `inetd.conf(4)` man page.

**Using the Listener**

Use `pmadm` to add RPC services:

```
pmadm -a -p pm_tag -s svctag -i id -v vers \  
-m 'nlsadmin -c command -D -R prog:vers'
```

The arguments are: `-a` means to add a service, `-p pm_tag` specifies a tag associated with the port monitor providing access to the service, `-s svctag` is the server’s identifying code, `-i id` is the `/etc/passwd` user name assigned
to service svctag, \(-v\) ver is the version number for the port monitor’s data base file, and \(-m\) specifies the nlsadmin command to invoke the service. nlsadmin can have additional arguments. For example, to add version 1 of a remote program server named rusersd, a pmadm command is:

```
# pmadm -a -p tcp -s rusers -i root -v 4 \n   -m 'nlsadmin -c /usr/sbin/rpc.ruserd -D -R 100002:1'
```

The command is given root permissions, installed in version 4 of the listener data base file, and is made available over TCP transports. Because of the complexity of the arguments and options to pmadm, use a command script or the menu system to add RPC services. To use the menu system, enter sysadm ports and choose the \(-p\) port_services option.

After adding a service, the listener must be re-initialized before the service is available. To do this, stop and restart the listener, as follows (note that rpcbind must be running):

```
# sacadm -k -p pmtag
# sacadm -s -p pmtag
```

For more information, such as how to set up the listener process, see the listen(1M), pmadm(1M), sacadm(1M) and sysadm(1M) man pages and the TCP/IP and Data Communications Administration Guide.

### Multiple Server Versions

By convention, the first version number of a program, PROG, is named PROGVERS_ORIG and the most recent version is named PROGVERS. Program version numbers must be assigned consecutively. Leaving a gap in the program version sequence can cause the search algorithm to not find a matching program version number that is defined.

Version numbers should never be changed by anyone other than the owner of a program. Adding a version number to a program that you do not own can cause severe problems when the owner increments the version number. Sun registers version numbers and answers questions about them (rpc@Sun.com).

Suppose a new version of the ruser program returns an unsigned short rather than a long. If you name this version RUSERSVERS_SHORT, a server that wants to support both versions would do a double register. The same server handle is used for both registrations.

#### CODE EXAMPLE 4–29 Server Handle for Two Versions of Single Routine

```c
if (!svc_reg(transp, RUSERSPROG, RUSERSVERS_ORIG, nuser, nconf))
{
    fprintf(stderr, "can’t register RUSER service\n");
}
```

The Programmer’s Interface to RPC 101
Both versions can be performed by a single procedure.

**CODE EXAMPLE 4–30**  Procedure for Two Versions of Single Routine

```c
void
nuser(rqstp, transp)
    struct svc_req *rqstp;
    SVCXPRT *transp;
{
    unsigned long nusers;
    unsigned short nusers2;
    switch(rqstp->rq_proc) {
    case NULLPROC:
        if (! svc_sendreply( transp, xdr_void, 0))
            fprintf(stderr, "can’t reply to RPC call\n");
        return;
    case RUSERSPROC_NUM:
        /*
         * Code here to compute the number of users
         * and assign it to the variable nusers
         */
        switch(rqstp->rq_vers) {
        case RUSERSVERS_ORIG:
            if (! svc_sendreply( transp, xdr_u_long, &nusers))
                fprintf(stderr, "can’t reply to RPC call\n");
            break;
        case RUSERSVERS_SHORT:
            nusers2 = nusers;
            if (! svc_sendreply( transp, xdr_u_short, &nusers2))
                fprintf(stderr, "can’t reply to RPC call\n");
            break;
        default:
            svcerr_noproc(transp);
            return;
        }
    }
    return;
}
```

**Multiple Client Versions**

Since different hosts may run different versions of RPC servers, a client should be capable of accommodating the variations. For example, one server may run the old version of RUSERSPROG(RUSERSVERS_ORIG) while another server runs the newer version (RUSERSVERS_SHORT).

If the version on a server does not match the version number in the client creation call, clnt_call fails with an RPCPROGVERSMISMATCH error. You
can get the version numbers supported by a server and then create a client handle with the appropriate version number. Use either the routine in Code Example 4–31, or clnt_create_vers. See the rpc(3N) man page for more details.

**CODE EXAMPLE 4–31   RPC Versions on Client Side**

```c
main()
{
enum clnt_stat status;
u_short num_s;
u_int num_l;
struct rpc_err rpcerr;
int maxvers, minvers;
CLIENT *clnt;

clnt = clnt_create("remote", RUSERSPROG, RUSERSVERS_SHORT,
    "datagram_v");
if (clnt == (CLIENT *) NULL) {
    clnt_pcreateerror("unable to create client handle");
    exit(1);
}
to.tv_sec = 10; /* set the time outs */
to.tv_usec = 0;
status = clnt_call(clnt, RUSERSPROC_NUM, xdr_void,
    (caddr_t) NULL, xdr_u_short,
    (caddr_t)&num_s, to);
if (status == RPC_SUCCESS) { /* Found latest version number */
    printf("num = %d\n", num_s);
    exit(0);
}
if (status != RPC_PROGVERSMISMATCH) { /* Some other error */
    clnt_perror(clnt, "rusers");
    exit(1);
}
/* This version not supported */
clnt_geterr(clnt, &rpcerr);
maxvers = rpcerr.re_vers.high; /* highest version supported */
minvers = rpcerr.re_vers.low; /* lowest version supported */
if (RUSERSVERS_SHORT < minvers || RUSERSVERS_SHORT > maxvers)
    { /* doesn't meet minimum standards */
    clnt_perror(clnt, "version mismatch");
    exit(1);
}
(void) clnt_control(clnt, CLSET_VERSION, RUSERSVERS_ORIG);
status = clnt_call(clnt, RUSERSPROC_NUM, xdr_void,
    (caddr_t) NULL, xdr_u_long, (caddr_t)&num_l, to);
if (status == RPC_SUCCESS)
    /* We found a version number we recognize */
    printf("num = %d\n", num_l);
else {
    clnt_perror(clnt, "rusers");
    exit(1);
}
}(Continued)
```
Using Transient RPC Program Numbers

Occasionally, it is useful for an application to use RPC program numbers that are generated dynamically. This could be used for implementing callback procedures, for example. In the callback example, a client program typically registers an RPC service using a dynamically generated, or transient, RPC program number and passes this on to a server along with a request. The server will then call back the client program using the transient RPC program number in order to supply the results. Such a mechanism may be necessary if processing the client’s request will take a huge amount of time and the client cannot block (assuming it is single-threaded); in this case, the server will acknowledge the client’s request, and call back later with the results. Another use of callbacks is to generate periodic reports from a server; the client makes an RPC call to start the reporting, and the server periodically calls back the client with reports using the transient RPC program number supplied by the client program.

Dynamically generated, or transient, RPC program numbers are in the transient range, 0x40000000 - 0x5fffffff. The following routine creates a service based on a transient RPC program for a given transport type. The service handle and the transient rpc program number are returned. The caller supplies the service dispatch routine, the version, and the transport type.

**CODE EXAMPLE 4–32** Transient RPC Program—Server Side

```c
SVCXPRT *register_transient_prog(dispatch, program, version, netid)
void (*dispatch)(); /* service dispatch routine */
u_long *program; /* returned transient RPC number */
u_long version; /* program version */
char *netid; /* transport id */
{
    SVCXPRT *transp;
    struct netconfig *nconf;
    u_long progrnum;
    if (!(nconf = getnetconfigent(netid)) == (struct netconfig *)NULL)
        return ((SVCXPRT *)NULL);
    if ((transp = svc_tli_create(RPC_ANYFD, nconf,
        (struct t_bind *)NULL, 0, 0)) == (SVCXPRT *)NULL) {
        freemessage((struct SVCXPRT *)NULL);
        return ((SVCXPRT *)NULL);
    }
    *program = progrnum;
    *version = version;
    *dispatch = dispatch;
    return (transp);
}
```
Multithreaded RPC Programming

This manual does not cover basic topics and code examples for the Solaris implementation of multithread programming. Instead, refer to the Multithreaded Programming Guide for background on the following topics:

- Thread creation
- Scheduling
- Synchronization
- Signals
- Process resources
- Light-weight processes (lwp)
- Concurrency
- Data locking strategies

TI-RPC supports multithreaded RPC servers in Solaris 2.4 and higher. The difference between a multithreaded server and a single-threaded server is that a multithreaded server uses threading technology to process incoming client requests concurrently. Multithreaded servers can have higher performance and availability compared with single-threaded servers.

The section “MT Server Overview” on page 110 is a good place to start reading about the interfaces available in this release.
MT Client Overview

In a multithread client program, a thread can be created to issue each RPC request. When multiple threads share the same client handle, only one thread at a time will be able to make an RPC request. All other threads will wait until the outstanding request is complete. On the other hand, when multiple threads make RPC requests using different client handles, the requests are carried out concurrently. Figure 4–1 illustrates a possible timing of a multithreaded client implementation consisting of two client threads using different client handles.

Code Example 4–33 shows the client side implementation of a multithreaded rstat program. The client program creates a thread for each host. Each thread creates its own client handle and makes various RPC calls to the given host. Because the client threads are using different handles to make the RPC calls, they can carry out the RPC calls concurrently.

Figure 4–1 Two Client Threads Using Different Client Handles (Real time)
Note - You must link in the thread library when writing any RPC multi-threaded-safe application. The thread library must be the last named library on the link line. To do this, specify the −lthread option in the compile command.

Compile the program in Code Example 4–33 by typing:

$ cc rstat.c -lnsl -lthread

CODE EXAMPLE 4–33  Client for MT rstat

```c
/* @(#)rstat.c 2.3 93/11/30 4.0 RPCSRC */
/*
 * Simple program that prints the status of a remote host,
 * in a format similar to that used by the 'w' command.
 */

#include <thread.h> /* thread interfaces defined */
#include <synch.h> /* mutual exclusion locks defined */
#include <stdio.h>
#include <sys/param.h>
#include <rpc/rpc.h>
#include <rpcsvc/rstat.h>
#include <errno.h>

mutex_t tty; /* control of tty for printf's */
cond_t cv_finish;
int count = 0;

main(argc, argv)
int argc;
char **argv;
{
    int i;
    thread_t tid;
    void *do_rstat();

    if (argc < 2) {
        fprintf(stderr, "usage: %s \"host\" [...]\n", argv[0]);
        exit(1);
    } else
        fprintf(stderr, "tid: %d\n", tid);

    mutex_lock(&tty);
    for (i = 1; i < argc; i++) {
        if (thr_create(NULL, 0, do_rstat, argv[i], 0, &tid) < 0) {
            fprintf(stderr, "thr_create failed: %d\n", i);
            exit(1);
        } else
            fprintf(stderr, "tid: %d\n", tid);
    }

    while (count < argc-1) {
        // (Continued)
```
printf('%argc = %d, count = %d\n', argc-1, count);
cond_wait(&cv_finish, &tty);
}
exit(0);
}
bool_t rstatproc_stats();

void *
do_rstat(host)
char *host;
{
CLIENT *rstat_clnt;
statstime host_stat;
bool_t rval;
struct tm *tmp_time;
struct tm host_time;
struct tm host_uptime;
char days_buf[16];
char hours_buf[16];
mutex_lock(&tty);
printf('%s: starting\n', host);
mutex_unlock(&tty);
/* client handle to rstat */
rstat_clnt = clnt_create(host, RSTATPROG, RSTATVERS_TIME,
"udp");
if (rstat_clnt == NULL) {
mutex_lock(&tty); /* get control of tty */
cint_pcreateerror(host);
count++;
cond_signal(&cv_finish);
mutex_unlock(&tty); /* release control of tty */
thr_exit(0);
}
rval = rstatproc_stats(NULL, &host_stat, rstat_clnt);
if (!rval) {
mutex_lock(&tty); /* get control of tty */
cint_perror(rstat_clnt, host);
count++;
cond_signal(&cv_finish);
mutex_unlock(&tty); /* release control of tty */
thr_exit(0);
}
tmp_time = localtime_r(&host_stat.curtime.tv_sec, &host_time);

host_stat.curtime.tv_sec = host_stat.boottime.tv_sec;

tmp_time = gmtime_r(&host_stat.curtime.tv_sec, &host_uptime);

if (host_uptime.tm_yday != 0)
    sprintf(days_buf, "%d day%s", host_uptime.tm_yday,
            (host_uptime.tm_yday > 1) ? "s" : "");
else
    days_buf[0] = '\0';

if (host_uptime.tm_hour != 0)
    sprintf(hours_buf, "%2d:%02d", host_uptime.tm_hour, host_uptime.tm_min);
else if (host_uptime.tm_min != 0)
    sprintf(hours_buf, "%2d mins", host_uptime.tm_min);
else
    hours_buf[0] = '\0';

mutex_lock(&tty); /* get control of tty */
printf("%s: \n", host);
printf(" %2d:%02d%cm up %s%s load average: %.2f %.2f %.2f\n",
       (host_time.tm_hour > 12) ? host_time.tm_hour - 12
       : host_time.tm_hour,
       (host_time.tm_hour >= 12) ? 'p' : 'a',
       days_buf,
       hours_buf,
       (double)host_stat.avenrun[0]/FSCALE,
       (double)host_stat.avenrun[1]/FSCALE,
       (double)host_stat.avenrun[2]/FSCALE);
count++;
cond_signal(&cv_finish);
mutex_unlock(&tty); /* release control of tty */
clnt_destroy(rstat_clnt);
sleep(10);
thr_exit(0);
}

/* Client side implementation of MT rstat program */
/* Default timeout can be changed using clnt_control() */
static struct timeval TIMEOUT = { 25, 0 };

bool_t
rstatproc_stats(argp, clnt_resp, clnt)
void *argp;
statstime *clnt_resp;
CLIENT *clnt;
{
    memset((char *)clnt_resp, 0, sizeof (statstime));
    if (clnt_call(clnt, RSTATPROC_STATS,
        (xdrproc_t) xdr_void, (caddr_t) argp,
        (xdrproc_t) xdr_statstime, (caddr_t) clnt_resp,
        TIMEOUT) != RPC_SUCCESS) {
        return (FALSE);
    }
    return (TRUE);
}

MT Server Overview

Prior to Solaris 2.4, RPC servers were single threaded. That is, they process client requests sequentially, as the requests come in. For example, if two requests come in, and the first takes 30 seconds to process, and the second takes only 1 second to process, the client that made the second request will still have to wait for the first request to complete before it receives a response. This is not desirable, especially in a multiprocessor server environment, where each CPU could be processing a different request simultaneously; or in a situation where one request is waiting for I/O to complete, other requests could be processed by the server.

Solaris 2.4 and higher provides facilities in the RPC library for service developers to create multithreaded servers that deliver better performance to end users. Two modes of server multithreading are supported in TI-RPC: the Automatic MT mode and the User MT mode.

In the Auto mode, the server automatically creates a new thread for every incoming client request. This thread processes the request, sends a response, and exits. In the User mode, the service developer decides how to create and manage threads for concurrently processing the incoming client requests. The Auto mode is much easier to use than the User mode, but the User mode offers more flexibility for service developers with special requirements.
Note - You must link in the thread library when writing RPC multithreaded-safe applications. The thread library must be the last named library on the link line. To do this, specify the -lthread option in the compile command.

The two calls that support server multithreading are rpc_control and svc_done. The rpc_control call is used to set the MT mode, either Auto or User mode. If the server uses Auto mode, it does not need to invoke svc_done at all. In User mode, svc_done must be invoked after each client request is processed, so that the server can reclaim the resources from processing the request. In addition, multithreaded RPC servers must call on svc_run. Note that svc_getreqpoll and svc_getreqset are unsafe in MT applications.

Note - If the server program does not invoke any of the MT interface calls, it remains in single-threaded mode, which is the default mode.

You are required to make RPC server procedures multithreaded safe regardless of which mode the server is using. Usually, this means that all static and global variables need to be protected with mutex locks. Mutual exclusion and other synchronization APIs are defined in synch.h. See the condition(3T), rwlock(3T), and mutex(3T) man pages for a list of the various synchronization interfaces.

Figure 4–2 illustrates a possible timing of a server implemented in one of the MT modes of operation.
Sharing the Service Transport Handle

The service transport handle, SVCXPRT, contains a single data area for decoding arguments and encoding results. Therefore, in the default, single-threaded mode, this structure cannot be freely shared between threads that call functions that perform these operations. However, when a server is operating in the MT Auto or User modes, a copy of this structure is passed to the service dispatch procedure in order to enable concurrent request processing. Under these circumstances, some routines which would otherwise be unsafe, become safe. Unless otherwise noted, the server interfaces are generally MT safe. See the rpc_svc_calls(3N) man page for more details on safety for server-side interfaces.

MT Auto Mode

In the Automatic mode, the RPC library creates and manages threads. The service developer invokes a new interface call, rpc_control, to put the server into MT Auto mode before invoking the svc_run call. In this mode, the programmer needs only to ensure that service procedures are MT safe. rpc_control allows applications to set and modify global RPC attributes. At present, it supports only server-side operations. Table 4–8 shows the
rpc_control operations defined for Auto mode. See also the rpc_control(3N) man page for additional information.

**TABLE 4-8  rpc_control() Library Routines**

<table>
<thead>
<tr>
<th>Routine</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>RPC_SVC_MTMODE_SET</td>
<td>Set multithread mode</td>
</tr>
<tr>
<td>RPC_SVC_MTMODE_GET</td>
<td>Get multithread mode</td>
</tr>
<tr>
<td>RPC_SVC_THRMAX_SET</td>
<td>Set Maximum number of threads</td>
</tr>
<tr>
<td>RPC_SVC_THRMAX_GET</td>
<td>Get Maximum number of threads</td>
</tr>
<tr>
<td>RPC_SVC_THRTOTAL_GET</td>
<td>Total number of threads currently active</td>
</tr>
<tr>
<td>RPC_SVC_THRCREATES_GET</td>
<td>Cumulative total number of threads created by the RPC library</td>
</tr>
<tr>
<td>RPC_SVC_THRERRORS_GET</td>
<td>Number of thr_create errors within RPC library</td>
</tr>
</tbody>
</table>

**Note** - All of the get operations in Table 4-8, except RPC_SVC_MTMODE_GET, apply only to the Auto MT mode. If used in MT User mode or the single-threaded default mode, the results of the operations may be undefined.

By default, the maximum number of threads that the RPC server library creates at any time is 16. If a server needs to process more than 16 client requests concurrently, the maximum number of threads must be set to the desired number. This parameter may be set at any time by the server, and it allows the service developer to put an upper bound on the thread resources consumed by the server. An example RPC program written in MT Auto mode. In this case, the maximum number of threads is set at 20.

MT performance is enhanced if the function svc_getargs is called by every procedure other than NULLPROCS, even if there are no arguments (xdr_void may be used in this case). This is true for both the MT Auto and MT User modes. For more information on this call, see the rpc_svc_calls(3N) man page.

illustrates the server in MT Auto mode.
**Note** - You must link in the thread library when writing RPC multithreaded-safe applications. The thread library must be the last named library on the link line. To do this, specify the `-lthread` option in the compile command.

Compile the program in by typing:

```
$ cc time_svc.c -lnsl -lthread
```

**CODE EXAMPLE 4–34**  
Server for MT Auto Mode

```c
#include <stdio.h>
#include <rpc/rpc.h>
#include <synch.h>
#include <thread.h>
#include "time_prot.h"

void time_prog();

main(argc, argv)
int argc;
char *argv[];
{
    int transpnum;
    char *nettype;
    int mode = RPC_SVC_MT_AUTO;
    int max = 20; /* Set maximum number of threads to 20 */

    if (argc > 2) {
        fprintf(stderr, "usage: %s [nettype]\n", argv[0]);
        exit(1);
    }

    if (argc == 2)
        nettype = argv[1];
    else
        nettype = "netpath";

    if (!rpc_control(RPC_SVC_MTMODE_SET, &mode)) {
        printf("RPC_SVC_MTMODE_SET: failed\n");
        exit(1);
    }

    if (!rpc_control(RPC_SVC_THRMAX_SET, &max)) {
        printf("RPC_SVC_THRMAX_SET: failed\n");
        exit(1);
    }

    transpnum = svc_create( time_prog, TIME_PROG, TIME_VERS, nettype);

    if (transpnum == 0) {
        fprintf(stderr, "%s: cannot create %s service.\n", argv[0], nettype);
        exit(1);
    }
```

(Continued)
/*
* The server dispatch function.
* The RPC server library creates a thread which executes
* the server dispatcher routine time_prog(). After which
* the RPC library will take care of destroying the thread.
*/

static void
time_prog(rqstp, transp)
    struct svc_req *rqstp;
    SVCXPRT *transp;
{
    switch (rqstp->rq_proc) {
    case NULLPROC:
        svc_sendreply(transp, xdr_void, NULL);
        return;
    case TIME_GET:
        dotime(transp);
        break;
    default:
        svcerr_noproc(transp);
        return;
    }
}

dotime(transp)
SVCXPRT *transp;
{

    struct timev rslt;
    time_t thetime;

    thetime = time((time_t *)0);
    rslt.second = thetime % 60;
    thetime /= 60;
    rslt.minute = thetime % 60;
    thetime /= 60;
    rslt.hour = thetime % 24;
    if (!svc_sendreply(transp, xdr_timev, (caddr_t) &rslt)) {
        svcerr_systemerr(transp);
    }
}

Code Example 4–35 shows the time_prot.h header file for the server.
MT User Mode

In MT User mode, the RPC library will not create any threads. This mode works, in principle, like the single-threaded, or default mode. The only difference is that it passes copies of data structures (such as the transport service handle to the service dispatch routine) to be MT safe.

The RPC server developer takes the responsibility for creating and managing threads through the thread library. In the dispatch routine, the service developer can assign the task of procedure execution to newly created or existing threads. The thr_create API is used to create threads having various attributes. All thread library interfaces are defined in thread.h. See the pthread_create(3T) man page for more details.

There is a lot of flexibility available to the service developer in this mode. Threads can now have different stack sizes based on service requirements. Threads may be bound. Different procedures may be executed by threads with different characteristics. The service developer may choose to run some services single threaded. The service developer may choose to do special thread-specific signal processing.

As in the Auto mode, the rpc_control library call is used to turn on User mode. Note that the rpc_control operations shown in Table 4–8 (except for RPC_SVC_MTMODE_GET) apply only to MT Auto mode. If used in MT User mode or the single-threaded default mode, the results of the operations may be undefined.
Freeing Library Resources in User Mode

In the MT User mode, service procedures must invoke svc_done before
returning. svc_done frees resources allocated to service a client request
directed to the specified service transport handle. This function is invoked
after a client request has been serviced, or after an error or abnormal
condition that prevents a reply from being sent. After svc_done is invoked,
the service transport handle should not be referenced by the service
procedure. shows a server in MT User mode.

Note - svc_done must only be called within MT User mode. For more
information on this call, see the rpc_svc_calls(3N)man page.

CODE EXAMPLE 4–36  MT User Mode: rpc_test.h

```c
#define SVC2_PROG 0x30000002
#define SVC2_VERS ((u_long) 1)
#define SVC2_PROC_ADD ((u_long) 1)
#define SVC2_PROC_MULT ((u_long) 2)

struct intpair {
    u_short a;
    u_short b;
};

typedef struct intpair intpair;

struct svc2_add_args {
    long argument;
    SVCXPRT *transp;
};

struct svc2_mult_args {
    intpair mult_argument;
    SVCXPRT *transp;
};

extern bool_t xdr_intpair();

#define NTHREADS_CONST 500
```

Code Example 4–37 is the client for MT User mode.

CODE EXAMPLE 4–37  Client for MT User Mode

```c
#define __REENTRANT
#include <stdio.h>
#include <rpc/rpc.h>
#include <sys/uio.h>
#include <netconfig.h>
#include <netdb.h>

(Continued)
```
#include <rpc/nettype.h>
#include <thread.h>
#include "rpc_test.h"

void *doclient();
int NTHREADS;
struct thread_info {
    thread_t client_id;
    int client_status;
};
struct thread_info save_thread[NTHREADS_CONST];

main(argc, argv)
    int argc;
    char *argv[];
{
    int index, ret;
    int thread_status;
    thread_t departedid, client_id;
    char *hosts;
    if (argc < 3) {
        printf("Usage: do_operation \[n\] host\n\n"),
        printf("\twhere n is the number of threads
\n"),
        exit(1);
    } else
    if (argc == 3) {
        NTHREADS = NTHREADS_CONST;
        hosts = argv[1]; /* live_host */
    } else {
        NTHREADS = atoi(argv[1]);
        hosts = argv[2];
    }
    for (index = 0; index < NTHREADS; index++)
        if (ret = thr_create(NULL, NULL, doclient,
            (void *) hosts, THR_BOUND, &client_id)) {
            printf("thr_create failed: return value \%d", ret);
            printf(" for \%dth thread\n", index);
            exit(1);
        }
    save_thread[index].client_id = client_id;
    for (index = 0; index < NTHREADS; index++)
        if (thr_join(save_thread[index].client_id, &departedid,
            (void *) &thread_status)) {
            printf("thr_join failed for thread \%d\n",
                save_thread[index].client_id);
            exit(1);
        }
    save_thread[index].client_status = thread_status;
}

void *doclient(host)
    char *host;
{
(Continued)
struct timeval tout;
enum clnt_stat test;
long result = 0;
u_short mult_result = 0;
long add_arg;
long EXP_RSLT;
intpair pair;
CLIENT *clnt;
if ((clnt = clnt_create(host, SVC2_PROG, SVC2_VERS, "udp"
==NULL)) {
    clnt_pcreateerror("clnt_create error: ");
    thr_exit((void *) -1);
}
tout.tv_sec = 25;
tout.tv_usec = 0;
memset((char *) &result, 0, sizeof (result));
memset((char *) &mult_result, 0, sizeof (mult_result));
if (thr_self() % 2){
    EXP_RSLT = thr_self() + 1;
    add_arg = thr_self();
    test = clnt_call(clnt, SVC2_PROC_ADD, (xdrproc_t) xdr_long,
        (caddr_t) &add_arg, (xdrproc_t) xdr_long, (caddr_t) &result,
        tout);
} else {
    pair.a = (u_short) thr_self();
    pair.b = (u_short) 1;
    EXP_RSLT = (long) pair.a * pair.b;
    test = clnt_call(clnt, SVC2_PROC_MULT, (xdrproc_t)
        xdr_intpair,
        (caddr_t) &pair, (xdrproc_t) xdr_u_short,
        (caddr_t) &mult_result, tout);
    result = (long) mult_result;
} else {
    printf("THREAD: %d clnt_call hav
        thr_exit((void *) -1);
    );
    thr_exit((void *) 0);
}

Code Example 4–38 shows the server side in MT User mode. MT performance
is enhanced if the function svc_getargs is called by every procedure other
than NULLPROC, even if there are no arguments (xdr_void may be used in this
case). This is true for both the MT Auto and MT User modes. For more
information on this call, see the rpc_svc_calls(3N)man page.
**Note** - You must link in the thread library when writing RPC multithreaded-safe applications. The thread library must be the last named library on the link line. To do this, specify the `-lthread` option in the compile command.

**CODE EXAMPLE 4–38  Server for MT User Mode**

```c
#define _REENTRANT
#include <stdio.h>
#include <rpc/rpc.h>
#include <sys/types.h>
#include <netinet/in.h>
#include <sys/uio.h>
#include <signal.h>
#include <thread.h>
#include "operations.h"

SVCXPRT *xprt;
void add_mult_prog();
void *svc2_add_worker();
void *svc2_mult_worker();
main(argc, argv)
    int argc;
    char **argv;
    { int transpnum;
        char *nettype;
        int mode = RPC_SVC_MT_USER;
        if(rpc_control(RPC_SVC_MTMODE_SET,&mode) == FALSE){
            printf(" rpc_control is failed to set AUTO mode\n");
            exit(0);
        }
        if (argc > 2) {
            fprintf(stderr, "usage: %s [nettype]\n", argv[0]);
            exit(1);
        }
        if (argc == 2)
            nettype = argv[1];
        else
            nettype = "netpath";
        transpnum = svc_create(add_mult_prog, SVC2_PROG,
                               SVC2_VERS, nettype);
        if (transpnum == 0) {
            fprintf(stderr, "%s: cannot create %s service.\n", argv[0],
                    nettype);
            exit(1);
        }
        svc_run();
    }
    void add_mult_prog (rqstp, transp)
        struct svc_req *rqstp,
       SVCXPRT *transp;
        {
```
long argument;
unsigned short mult_arg();
intpair mult_argument;
bool_t (*xdr_argument)();
struct svc2_mult_args *sw_mult_data;
struct svc2_add_args *sw_add_data;
int ret;
thread_t worker_id;

switch ((long) rqstp->rq_proc){
  case NULLPROC:
    svc_sendreply(transp, xdr_void, (char *) 0);
    svc_done(transp);
    break;
  case SVC2_PROC_ADD:
    xdr_argument = xdr_long;
    (void) memset((char *) &argument, 0, sizeof (argument));
    if (!svc_getargs(transp, xdr_argument,
                     (char *) &argument)) {
      printf("problem with getargs\n");
      svcerr_decode(transp);
      exit(1);
    }
    sw_add_data = (struct svc2_add_args *)
      malloc(sizeof (struct svc2_add_args));
    sw_add_data->transp = transp;
    sw_add_data->argument = argument;
    if (ret = thr_create(NULL, THR_MIN_STACK + 16 * 1024,
                         svc2_add_worker, (void *) sw_add_data, THR_DETACHED, 
                         printf("SERVER: thr_create failed:");
                         printf(" return value %d", ret);
                         printf(" for add thread\n");
                         exit(1);
                     })
      break;
  case SVC2_PROC_MULT:
    xdr_argument = xdr_intpair;
    (void) memset((char *) &mult_argument, 0, sizeof (mult_argument));
    if (!svc_getargs(transp, xdr_argument,
                     (char *) &mult_argument)) {
      printf("problem with getargs\n");
      svcerr_decode(transp);
      exit(1);
    }
    sw_mult_data = (struct svc2_mult_args *)
      malloc(sizeof (struct svc2_mult_args));
    sw_mult_data->transp = transp;
    sw_mult_data->mult_argument.a = mult_argument.a;
    sw_mult_data->mult_argument.b = mult_argument.b;
    if (ret = thr_create(NULL, THR_MIN_STACK + 16 * 1024,
                         svc2_mult_worker, (void *) sw_mult_data, THR_DETACHED, 
                         &worker_id)) {
      printf("SERVER: thr_create failed:");
      (Continued)
printf("return value %d", ret);
printf("for multiply thread\n");
exit(1);
break;
default:
    svcerr_noproc(transp);
svc_done(transp);
break;
}
}
void *svc2_add_worker(add_arg)
struct svc2_add_args *add_arg;
{ long *result;
    bool_t (*xdr_result)();
    xdr_result = xdr_long;
    result = (long *) malloc(sizeof (long));
    *result = add_one(add_arg->argument);
    if (!svc_sendreply(add_arg->transp, xdr_result,
        (caddr_t) result)){
        printf("sendreply failed\n");
        svcerr_systemerr(add_arg->transp);
        svc_done(add_arg->transp);
        thr_exit((void *) -1);
    }
    svc_done(add_arg->transp);
    thr_exit((void *) 0);
}
void *svc2_mult_worker(m_arg)
struct svc2_mult_args *m_arg;
{ u_short *result;
    bool_t (*xdr_result)();
    xdr_result = xdr_u_short;
    result = (u_short *) malloc(sizeof (u_short));
    *result = mult_arg(&m_arg->mult_argument);
    if (!svc_sendreply(m_arg->transp, xdr_result,
        (caddr_t) result)){
        printf("sendreply failed\n");
        svcerr_systemerr(m_arg->transp);
        svc_done(m_arg->transp);
        thr_exit((void *) -1);
    }
    svc_done(m_arg->transp);
    thr_exit((void *) 0);
    u_short mult_arg(pair)
    intpair *pair;
    { u_short result;
        result = pair->a * pair->b;
        return (result); }
long add_one(arg)
    long arg;
    {
        return (++arg);
    }

Connection-Oriented Transports

Code Example 4–39 copies a file from one host to another. The RPC send call reads standard input and sends the data to the server receive, which writes the data to standard output. This also illustrates an XDR procedure that behaves differently on serialization and on deserialization. A connection-oriented transport is used.

CODE EXAMPLE 4–39  Remote Copy (Two-Way XDR Routine)

    /*
    * The xdr routine:
    * on decode, read wire, write to fp
    * on encode, read fp, write to wire
    */
    #include <stdio.h>
    #include <rpc/rpc.h>

    bool_t
    xdr_rcp(xdrs, fp)
    XDR *xdrs;
    FILE *fp;
    {
        unsigned long size;
        char buf[BUFSIZ], *p;

        if (xdrs->x_op == XDR_FREE) /* nothing to free */
            return(TRUE);
        while (TRUE) {
            if (xdrs->x_op == XDR_ENCODE) {
                if ((size = fread(buf, sizeof(char), BUFSIZ, fp))
                    == 0 && ferror(fp)) {
                    fprintf(stderr, "can’t fread
");
                } else {
                    if (xdrs->x_op == XDR_FREE) {
                        /* free wire */
                        free((char *)buf);
                    } else {
                        if (xdrs->x_op == XDR_DECODE) {
                            /* decode, read wire, write to fp */
                            if (fwrite(buf, size, 1, fp))
                                return(TRUE);
                        } else {
                            /* encode, read fp, write to wire */
                            if (fwrite(xhr, size, 1, fp))
                                return(TRUE);
                        }
                    }
                }
            } else {
                return(FALSE);
            }
        }
    }
In Code Example 4–40 and Code Example 4–41, the serializing and deserializing are done only by the xdr_rcp routine shown in Code Example 4–39.

**CODE EXAMPLE 4–40  Remote Copy Client Routines**

```c
/* The sender routines */
#include <stdio.h>
#include <netdb.h>
#include <rpc/rpc.h>
#include <sys/socket.h>
#include <sys/time.h>
#include "rcp.h"

main(argc, argv)
    int argc;
    char **argv;
{
    int xdr_rcp();
    if (argc != 2) {
        fprintf(stderr, "usage: %s servername\n", argv[0]);
        exit(1);
    }
    if( callcots( argv[1], RCPPROG, RCPPROC, RCPVERS, xdr_rcp,
        stdin, 
        xdr_void, 0 ) != 0 )
        exit(1);

    return(FALSE);
} else
    return(TRUE);
} p = buf;
if (! xdr_bytes( xdrs, &p, &size, BUFSIZ))
    return(0);
if (size == 0)
    return(1);
if (xdrs->x_op == XDR_DECODE) {
    if (fwrite( buf, sizeof(char), size, fp) != size) {
        fprintf(stderr, "can't fwrite\n");
        return(FALSE);
    } else
        return(TRUE);
}
```

(Continued)
exit(0);
}
callcots(host, prognum, procnum, versnum, inproc, in, outproc, out)
    char *host, *in, *out;
    xdrproc_t inproc, outproc;
{
    enum clnt_stat clnt_stat;
    register CLIENT *client;
    struct timeval total_timeout;

    if ((client = clnt_create( host, prognum, versnum,
        "circuit_v")
        == (CLIENT *) NULL)) {
        clnt_pcreateerror("clnt_create");
        return(-1);
    }
    total_timeout.tv_sec = 20;
    total_timeout.tv_usec = 0;
    clnt_stat = clnt_call(client, procnum, inproc, in, outproc,
        out,
            total_timeout);
    clnt_destroy(client);
    if (clnt_stat != RPC_SUCCESS)
        clnt_perror("callcots");
    return((int)clnt_stat);
}

The receiving routines are defined in Code Example 4–41. Note that in the server, xdr_rcp did all the work automatically.

**CODE EXAMPLE 4–41** Remote Copy Server Routines

```c
/*
 * The receiving routines
 */
#include <stdio.h>
#include <rpc/rpc.h
#include "rcp.h"

main()
{
    void rcp_service();
    if (svc_create(rcp_service,RCPPROG,RCPVERS,"circuit_v") == 0) {
        fprintf(stderr, "svc_create: errpr\n");
    }
}
```

(Continued)
exit(1);
}
svc_run();    /* never returns */
fprintf(stderr, "svc_run should never return\n");
}

void
rcp_service(rqstp, transp)
    register struct svc_req *rqstp;
    register SVCXPRT *transp;
{
    switch(rqstp->rq_proc) {
      case NULLPROC:
        if (svc_sendreply(transp, xdr_void, (caddr_t) NULL) == FALSE)
            fprintf(stderr, "err: rcp_service");
        return;
      case RCPPROC:
        if (!svc_getargs( transp, xdr_rcp, stdout)) {
            svcerr_decode(transp);
            return();
        }
        if(!svc_sendreply(transp, xdr_void, (caddr_t) NULL)) {
            fprintf(stderr, "can’t reply\n");
            return();
        }
        return();
      default:
        svcerr_noproc(transp);
        return();
    }
}

Memory Allocation With XDR

XDR routines normally serialize and deserialze data. XDR routines often automatically allocate memory and free automatically allocated memory. The convention is to use a NULL pointer to an array or structure to indicate that an XDR function must allocate memory when deserializing. The next example, xdr_chararr1, processes a fixed array of bytes with length SIZE and cannot allocate memory if needed:
If space has already been allocated in `chararr`, it can be called from a server like this:

```c
char chararr[SIZE];
svc_getargs(transp, xdr_chararr1, chararr);
```

Any structure through which data is passed to XDR or RPC routines must be allocated so that its base address is at an architecture-dependent boundary. An XDR routine that does the allocation must be written so that it can:

- Allocate memory when a caller requests
- Return the pointer to any memory it allocates

In the following example, the second argument is a NULL pointer, meaning that memory should be allocated to hold the data being deserialized.

```c
xdr_chararr2(xdrsp, chararrp)
XDR *xdrsp;
char **chararrp;
{
  int len;
  len = SIZE;
  return (xdr_bytes(xdrsp, chararrp, &len, SIZE));
}
```

The corresponding RPC call is:

```c
char *arrptr;
arrptr = NULL;
svc_getargs(transp, xdr_chararr2, &arrptr);
/*
 * Use the result here
 */
svc_freeargs(transp, xdr_chararr2, &arrptr);
```

After use, the character array should be freed through `svc_freeargs`. `svc_freeargs` does nothing if passed a NULL pointer as its second argument.
To summarize:

- An XDR routine normally serializes, deserializes, and frees memory.
- `svc_getargs` calls the XDR routine to deserialize.
- `svc_freeargs` calls the XDR routine to free memory.

Porting From TS-RPC to TI-RPC

The transport-independent RPC (TI-RPC) routines allow the developer stratified levels of access to the transport layer. The highest-level routines provide complete abstraction from the transport and provide true transport-independence. Lower levels provide access levels similar to the TI-RPC of previous releases.

This section is an informal guide to porting transport-specific RPC (TS-RPC) applications to TI-RPC. Table 4–9 shows the differences between selected routines and their counterparts. For information on porting issues concerning sockets and transport layer interface (TLI), see the *Transport Interfaces Programming Guide*.

Porting an Application

An application based on either TCP or UDP can run in binary-compatibility mode. For some applications you only recompile and relink all source files. This may be true of applications that use simple RPC calls and use no socket or TCP or UDP specifics.

Some editing and new code may be needed if an application depends on socket semantics or features specific to TCP or UDP. Examples use the format of host addresses or rely on the Berkeley UNIX concept of privileged ports.

Applications that are dependent on the internals of the library or the socket implementation, or depend on specific transport addressing probably require more effort to port and may require substantial modification.

Benefits of Porting

Some of the benefits of porting are:

- Applications transport independence means they operate over more transports than before.
- Use of new interfaces make your application more efficient.
- Binary compatibility is less efficient than native mode.
- Old interfaces could be removed from future releases.

Porting Issues

libnsl Library
libc no longer includes networking functions. libnsl must be explicitly specified at compile time to link the network services routines.

Old Interfaces
Many old interfaces are supported in the libnsl library, but they work only with TCP or UDP transports. To take advantage of new transports, you must use the new interfaces.

Name-to-Address Mapping
Transport independence requires opaque addressing. This has implications for applications that interpret addresses.

Differences Between TI-RPC and TS-RPC
The major differences between transport-independent RPC and transport-specific RPC are illustrated in Table 4–9. Also see section “Comparison Examples” on page 133 for code examples comparing TS-RPC with TI-RPC.

<table>
<thead>
<tr>
<th>Category</th>
<th>TI-RPC</th>
<th>TS-RPC</th>
</tr>
</thead>
<tbody>
<tr>
<td>Default Transport Selection</td>
<td>TI-RPC uses the TLI interface.</td>
<td>TS-RPC uses the socket interface.</td>
</tr>
<tr>
<td>RPC Address Binding</td>
<td>TI-RPC uses rpcbind for service binding. rpcbind keeps address in universal address format.</td>
<td>TS-RPC uses portmap for service binding.</td>
</tr>
</tbody>
</table>
### TABLE 4–9  Differences Between TI-RPC and TS-RPC  (continued)

<table>
<thead>
<tr>
<th>Category</th>
<th>TI-RPC</th>
<th>TS- RPC</th>
</tr>
</thead>
<tbody>
<tr>
<td>Transport Information</td>
<td>Transport information is kept in a local file, /etc/netconfig. Any transport identified in netconfig is accessible.</td>
<td>Only TCP and UDP transports are supported.</td>
</tr>
<tr>
<td>Loopback Transports</td>
<td>rpcbind service requires a secure loopback transport for server registration</td>
<td>TS-RPC services do not require a loopback transport.</td>
</tr>
<tr>
<td>Host Name Resolution</td>
<td>The order of host name resolution in TI-RPC depends on the order of dynamic libraries identified by entries in /etc/netconfig.</td>
<td>Host name resolution is done by name services. The order is set by the state of the hosts database.</td>
</tr>
<tr>
<td>File Descriptors</td>
<td>Descriptors are assumed to be TLI endpoints.</td>
<td>Descriptors are assumed to be sockets.</td>
</tr>
<tr>
<td>rpcgen</td>
<td>The TI-RPC rpcgen tool adds support for multiple arguments, pass-by values, sample client files, and sample server files.</td>
<td>rpcgen in SunOS 4.1 and previous releases do not support the features listed for TI-RPC rpcgen.</td>
</tr>
<tr>
<td>Libraries</td>
<td>TI-RPC requires that applications be linked to the libnl library.</td>
<td>All TS-RPC functionality is provided in libc.</td>
</tr>
<tr>
<td>MT Support</td>
<td>Multithreaded RPC clients and servers are supported.</td>
<td>Multithreaded RPC is not supported.</td>
</tr>
</tbody>
</table>

### Function Compatibility Lists

The RPC library functions are listed in this section and grouped into functional areas. Each section includes lists of functions that are unchanged, have added functionality, and are new relative to previous releases.

**Note** - Functions marked with an asterisk are retained for ease of porting and may be not be supported in future releases of Solaris.
Creating Client Handles

The following functions are unchanged from the previous release and available in the current SunOS release:

clnt_destroy
clnt_pcreateerror
clnt_raw_create
clnt_spcreateerror
clnt_tcp_create
clntudp_bufcreate
clntudp_create
clntraw_create
clntraw_create
clnt_create
clntraw_create_timed
clntraw_create_vers
clntraw_dg_create
clntraw_raw_create
clntraw_tli_create
clntraw_tp_create
clntraw_tp_create_timed
clntraw_vc_create

Creating and Destroying Services

The following functions are unchanged from the previous releases and available in the current SunOS release:

svc_destroy
cvcfd_create
svc_raw_create
svc_tp_create
svcudp_create
svc_udp_bufcreate
svc_create
svc_dp_create
csvcfd_create
csvc_raw_create
csvc_tli_create
csvc_tp_create
csvc_vc_create

Registering and Unregistering Services

The following functions are unchanged from the previous releases and available in the current SunOS release:

*registerrpc
*srv_register
*srv_unregister
xprt_register
xprt_unregister
rpc_reg
svc_reg
svc_unregister
SunOS 4.x Compatibility Calls

The following functions are unchanged from previous releases and available in the current SunOS release:

*callrpc
clnt_call
svc_getcaller - works only with IP-based transports
rpc_call
svc_getrpccaller

Broadcasting

The following call has the same functionality as in previous releases, although it is supported for backward compatibility only:

*clnt_broadcast

clnt_broadcast can broadcast only to the portmap service. It does not support rpcbind.

The following function that broadcasts to both portmap and rpcbind is also available in the current release of SunOS:

rpc_broadcast

Address Management Functions

The TI-RPC library functions interface with either portmap or rpcbind. Since the services of the programs differ, there are two sets of functions, one for each service.

The following functions work with portmap:

pmap_set
pmap_unset
pmap_getport
pmap_getmaps
pmap_rmtcall

The following functions work with rpcbind:

rpcb_set
rpcb_unset
rpcb_getaddr
rpcb_getmaps
rpcb_rmtcall
Authentication Functions

The following calls have the same functionality as in previous releases. They are supported for backward compatibility only:

- authdes_create
- authunix_create
- authunix_create_default
- authdes_seccreate
- authsys_create
- authsys_create_default

Other Functions

rpcbind provides a time service (primarily for use by secure RPC client-server time synchronization), available through the rpcb_gettime function. pmap_getport and rpcb_getaddr can be used to get the port number of a registered service. rpcb_getaddr communicates with any server running version 2, 3, or 4 of rpcbind. pmap_getport can only communicate with version 2.

Comparison Examples

The changes in client creation from TS-RPC to TI-RPC are illustrated in Code Example 4–42 and Code Example 4–43. Each example

- Creates a UDP descriptor.
- Contacts the remote host’s RPC binding process to get the services address.
- Binds the remote service’s address to the descriptor.
- Creates the client handle and set its time out.

CODE EXAMPLE 4–42 Client Creation in TS-RPC

```c
struct hostent *h;
struct sockaddr_in sin;
int sock = RPC_ANYSOCK;
u_short port;
struct timeval wait;

if ((h = gethostbyname( "host" )) == (struct hostent *) NULL)
{
    syslog(LOG_ERR, "gethostbyname failed");
    exit(1);
}

sin.sin_addr.s_addr = *(u_long *) h->h_addr;
if ((port = pmap_getport(&sin, PROGRAM, VERSION, "udp")) == 0) {
    syslog (LOG_ERR, "pmap_getport failed");
    exit(1);
} else
    sin.sin_port = htons(port);
```
wait.tv_sec = 25;
wait.tv_usec = 0;
clnudp_create(sin, PROGRAM, VERSION, wait, &sock);

The TI-RPC version assumes that the UDP transport has the netid udp. A netid is not necessarily a well-known name.

**CODE EXAMPLE 4–43**  Client Creation in TI-RPC

```c
struct netconfig *nconf;
struct netconfig *getnetconfig();
struct timeval wait;

nconf = getnetconfig("udp");
if (nconf == (struct netconfig *) NULL) {
    syslog(LOG_ERR, "getnetconfig for udp failed");
    exit(1);
}
fd = t_open(nconf->nc_device, O_RDWR, (struct t_info *)NULL);
if (fd == -1) {
    syslog(LOG_ERR, "t_open failed");
    exit(1);
}
tbind = (struct t_bind *) t_alloc(fd, T_BIND, T_ADDR);
if (tbind == (struct t_bind *) NULL) {
    syslog(LOG_ERR, "t_bind failed");
    exit(1);
}
if (rpcb_getaddr( PROGRAM, VERSION, nconf, &tbind->addr, "host")
    == FALSE) {
    syslog(LOG_ERR, "rpcb_getaddr failed");
    exit(1);
}
cl = clnt_tli_create(fd, nconf, &tbind->addr, PROGRAM, VERSION,
                      0, 0);
(void) t_free((char *) tbind, T_BIND);
if (cl == (CLIENT *) NULL) {
    syslog(LOG_ERR, "clnt_tli_create failed");
    exit(1);
}
wait.tv_sec = 25;
wait.tv_usec = 0;
cln_control(cl, CLSET_TIMEOUT, (char *) &wait);
```

Code Example 4–44 and Code Example 4–45 show the differences between broadcast in TS-RPC and TI-RPC. The older clnt_broadcast is similar to the newer rpc_broadcast. The primary difference is in the collectnames function: deletes duplicate addresses and displays the names of hosts that reply to the broadcast.

**CODE EXAMPLE 4–44**  Broadcast in TS-RPC

```c
statstime sw;
extern int collectnames();
```
Code Example 4–45 shows the Broadcast for TI-RPC:

CODE EXAMPLE 4–45 Broadcast in TI-RPC

```c
statstime sw;
extern int collectnames();

cpycast(RSTATPROG, RSTATVERS_TIME, RSTATPROC_STATS,
     xdr_void, NULL, xdr_statstime, &sw, collectnames);

... collectnames(resultsp, raddrp)
     char *resultsp;
     struct sockaddr_in *raddrp;
{
     u_long addr;
     struct entry *entryp, *lim;
     struct hostent *hp;
     extern int curentry;
     /* weed out duplicates */
     addr = raddrp->sin_addr.s_addr;
     lim = entry + curentry;
     for (entryp = entry; entryp < lim; entryp++)
       if (addr == entryp->addr)
         return (0);
     ...
   /* print the host’s name (if possible) or address */
   hp = gethostbyaddr(raddrp->sin_addr.s_addr, sizeof(u_long),
                      AF_INET);
   if( hp == (struct hostent *) NULL)
     printf("0x%x", addr);
   else
     printf("%s", hp->h_name);
}
```

```c

cpycast(RSTATPROG, RSTATVERS_TIME, RSTATPROC_STATS,
     xdr_void, NULL, xdr_statstime, &sw, collectnames, (char *)
     0);

... collectnames(resultsp, taddr, nconf)
     char *resultsp;
     struct t_bind *taddr;
     struct netconfig *nconf;
{
     struct entry *entryp, *lim;
     struct netconfig *nconf;
     extern int curentry;
     extern int netbufeq();
     /* weed out duplicates */
     lim = entry + curentry;
     for (entryp = entry; entryp < lim; entryp++)
       if (netbufeq( &taddr->addr, entryp->addr))
         return (0);
     ...
   /* print the host’s name (if possible) or address */
```
if (netdir_getbyaddr( nconf, &hs, &taddr->addr ) == ND_OK)
    printf("%s", hs->h_hostservs->h_host);
else {
    char *uaddr = taddr2uaddr(nconf, &taddr->addr);
    if (uaddr) {
        printf("%s\n", uaddr);
        (void) free(uaddr);
    } else
        printf("unknown");
}

netbufeq(a, b)
struct netbuf *a, *b;
{
    return(a->len == b->len && memcmp( a->buf, b->buf, a->len));
}
Part 3 discusses the NIS+ API.

- Chapter 5
NIS+ Programming Guide

This chapter presents the fundamental principles of the NIS+ applications programming interface and a detailed sample program. The NIS+ API is for programmers who need to build applications for Solaris-based networks. It provides the essential features for supporting enterprise-wide applications.

- “NIS+ Overview” on page 139
- “NIS+ API” on page 143
- “NIS+ Sample Program” on page 148

The NIS+ network name service addresses the requirements of client/server networks ranging in size from 10 clients supported by a few servers on a simple local area network to 10,000 multi-vendor clients supported by 20 to 100 specialized servers located in sites throughout the world and connected by several public networks.

NIS+ Overview

Domains

NIS+ supports hierarchical domains, illustrated as a simple case in Figure 5–1
A NIS+ domain is a set of data describing the workstations, users, and network services in a portion of an organization. NIS+ domains can be administered independently of each other. This allows NIS+ to be used in a range of networks, from small to very large.

**Servers**

Each domain is supported by a set of servers. The principal server is called the master server, and the backup servers are called replicas. Both master and replica servers run NIS+ server software. The master server stores the original tables, and the backup servers store copies.

NIS+ accepts incremental updates to the replicas. Changes are first made on the master server. Then they are automatically propagated to the replica servers and are soon available to the entire namespace.

**Tables**

NIS+ stores information in tables instead of maps or zone files. NIS+ provides 16 types of predefined, or system, tables, shown in Figure 5–2.
Each table stores a different type of information. For instance, the Hosts table stores host name/Internet address pairs, and the Password table stores information about users of the network.

NIS+ tables have two major improvements over NIS maps. First, a NIS+ table can be accessed by any column, not just the first column (sometimes referred to as the “key”). This eliminates the need for duplicate maps, such as the `hosts.byname` and `hosts.byaddr` maps of NIS. Second, access to the information in NIS+ tables can be controlled at three levels of granularity: the table level, the entry level, and the column level.

### NIS+ Security

The NIS+ security model provides both authorization and authentication mechanisms. First, every object in the namespace specifies the type of operation it accepts and from whom. This is authorization. Second, NIS+ attempts to authenticate every requestor accessing the namespace. Once it identifies the originator of the request, it determines whether the object has authorized that particular operation for that particular principal. Based on its authentication and the object’s authorization, NIS+ carries out or denies the access request.

### Name Service Switch

NIS+ works in conjunction with a separate facility called the Name Service Switch. The Name Service Switch, sometimes referred to as “the Switch,” lets Solaris 2.x-based workstations obtain their information from more than one network information service; specifically, from local, or `/etc` files, from NIS maps, from DNS zone files, or from NIS+ tables. The Switch not only offers a choice of sources, but allows a workstation to specify different sources for different types of information. The name service is configured through the file `/etc/nsswitch.conf`.

### NIS+ Administration Commands

NIS+ provides a full set of commands for administering a namespace. Table 5–1 summarizes them.
<table>
<thead>
<tr>
<th>Command</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>nischgrp</td>
<td>Changes the group owner of a NIS+ object.</td>
</tr>
<tr>
<td>nischmod</td>
<td>Changes an object’s access rights.</td>
</tr>
<tr>
<td>nischown</td>
<td>Changes the owner of a NIS+ object.</td>
</tr>
<tr>
<td>nisgrpadm</td>
<td>Creates or destroys a NIS+ group, or displays a list of its members. Also adds members to a group, removes them, or tests them for membership in the group.</td>
</tr>
<tr>
<td>niscat</td>
<td>Displays the contents of NIS+ tables.</td>
</tr>
<tr>
<td>nisgrep</td>
<td>Searches for entries in a NIS+ table.</td>
</tr>
<tr>
<td>nisls</td>
<td>Lists the contents of a NIS+ directory.</td>
</tr>
<tr>
<td>nismatch</td>
<td>Searches for entries in a NIS+ table.</td>
</tr>
<tr>
<td>nisaddent</td>
<td>Adds information from /etc files or NIS maps into NIS+ tables.</td>
</tr>
<tr>
<td>nistbladm</td>
<td>Creates or deletes NIS+ tables, and adds, modifies or deletes entries in a NIS+ table.</td>
</tr>
<tr>
<td>nisaddcred</td>
<td>Creates credentials for NIS+ principals and stores them in the Cred table.</td>
</tr>
<tr>
<td>nispasswd</td>
<td>Changes password information stored in the NIS+ Passwd table.</td>
</tr>
<tr>
<td>nisupdkeys</td>
<td>Updates the public keys stored in a NIS+ object.</td>
</tr>
<tr>
<td>nisinit</td>
<td>Initializes a NIS+ client or server.</td>
</tr>
<tr>
<td>nismkdir</td>
<td>Creates a NIS+ directory and specifies its master and replica servers.</td>
</tr>
<tr>
<td>nisrmdir</td>
<td>Removes NIS+ directories and replicas from the namespace.</td>
</tr>
<tr>
<td>nissetup</td>
<td>Creates org_dir and groups_dir directories and a complete set of (unpopulated) NIS+ tables for a NIS+ domain.</td>
</tr>
<tr>
<td>rpc.nisd</td>
<td>The NIS+ server process.</td>
</tr>
<tr>
<td>nis_cachemgr</td>
<td>Starts the NIS+ Cache Manager on a NIS+ client.</td>
</tr>
</tbody>
</table>
TABLE 5-1  NIS+ Namespace Administration Commands  (continued)

<table>
<thead>
<tr>
<th>Command</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>nischt tl</td>
<td>Changes a NIS+ object’s time-to-live value.</td>
</tr>
<tr>
<td>nisdefaults</td>
<td>Lists a NIS+ object’s default values: domain name, group name, workstation name, NIS+ principal name, access rights, directory search path, and time-to-live.</td>
</tr>
<tr>
<td>nisln</td>
<td>Creates a symbolic link between two NIS+ objects.</td>
</tr>
<tr>
<td>nisrm</td>
<td>Removes NIS+ objects (except directories) from the namespace.</td>
</tr>
<tr>
<td>nisshowcache</td>
<td>Lists the contents of the NIS+ shared cache maintained by the NIS+ Cache Manager.</td>
</tr>
</tbody>
</table>

NIS+ API

The NIS+ application programming interface (API) is a group of functions that can be called by an application to access and modify NIS+ objects. The NIS+ API has 54 functions that fall into nine categories:

- Object Manipulation Functions (nis_names)
- Table Access Functions (nis_tables)
- Local Name Functions (nis_local_names)
- Group Manipulation Functions (nis_groups)
- Server Related Functions (nis_server)
- Database Access Functions (nis_db)
- Error Message Display Functions (nis_error)
- Transaction Log Functions (nis_admin)
- Miscellaneous Functions (nis_subr)

The functions in each category are summarized in Table 5-2. The category names match the names by which they are grouped in the NIS+ manpages.
<table>
<thead>
<tr>
<th>Function</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>nis_names</td>
<td>Locate and Manipulate Objects</td>
</tr>
<tr>
<td>nis_lookup</td>
<td>Returns a copy of an NIS+ object. Can follow links. Though it cannot search for an entry object, if a link points to one, it can return an entry object.</td>
</tr>
<tr>
<td>nis_add</td>
<td>Adds an NIS+ object to the namespace.</td>
</tr>
<tr>
<td>nis_remove</td>
<td>Removes an NIS+ object in the namespace.</td>
</tr>
<tr>
<td>nis_modify</td>
<td>Modifies an NIS+ object in the namespace.</td>
</tr>
<tr>
<td>nis_tables</td>
<td>Search and Update Tables.</td>
</tr>
<tr>
<td>nis_list</td>
<td>Searches a table in the NIS+ namespace and returns entry objects that match the search criteria. Can follow links and search paths from one table to another.</td>
</tr>
<tr>
<td>nis_add_entry</td>
<td>Adds an entry object to an NIS+ table. Can be instructed to either fail or overwrite if the entry object already exists. Can return a copy of the resulting object if the operation was successful.</td>
</tr>
<tr>
<td>nis_freeresult</td>
<td>Frees all memory associated with a nis_result structure.</td>
</tr>
<tr>
<td>nis_remove_entry</td>
<td>Removes one or more entry objects from an NIS+ table. Can identify the object to be removed by using search criteria or by pointing to a cached copy of the object. If using search criteria, can remove all objects that match the search criteria; therefore, with the proper search criteria, can remove all entries in a table. Can return a copy of the resulting object if the operation was successful.</td>
</tr>
<tr>
<td>nis_modify_entry</td>
<td>Modifies one or more entry objects in an NIS+ table. Can identify the object to be modified by using search criteria or by pointing to a cached copy of the object.</td>
</tr>
<tr>
<td>nis_first_entry</td>
<td>Returns a copy of the first entry object in an NIS+ table.</td>
</tr>
<tr>
<td>Function</td>
<td>Description</td>
</tr>
<tr>
<td>--------------------------</td>
<td>---------------------------------------------------------------------------------------------------------------------------------------------</td>
</tr>
<tr>
<td>nis_next_entry</td>
<td>Returns a copy of the “next” entry object in an NIS+ table. Because a table can be updated and entries removed or modified between calls to this function, the order of entries returned may not match the actual order of entries in the table.</td>
</tr>
<tr>
<td>nis_local_names</td>
<td>Get Default Names for the Current Process</td>
</tr>
<tr>
<td>nis_local_directory</td>
<td>Returns the name of the workstation’s NIS+ domain.</td>
</tr>
<tr>
<td>nis_local_host</td>
<td>Returns the fully-qualified name of the workstation. A fully qualified name has the form <code>&lt;host-name&gt;.&lt;domain-name&gt;</code>.</td>
</tr>
<tr>
<td>nis_local_group</td>
<td>Returns the name of the current NIS+ group, which is specified by the environment variable NIS_GROUP.</td>
</tr>
<tr>
<td>nis_local_principal</td>
<td>Returns the name of the NIS+ principal whose UID is associated with the calling process.</td>
</tr>
<tr>
<td>nis_getnames</td>
<td>Returns a list of possible expansions to a particular name.</td>
</tr>
<tr>
<td>nis_freenames</td>
<td>Frees the memory containing the list generated by nis_getnames.</td>
</tr>
<tr>
<td>nis_groups</td>
<td>Group Manipulation and Authorization</td>
</tr>
<tr>
<td>nis_ismember</td>
<td>Test whether a principal is a member of a group.</td>
</tr>
<tr>
<td>nis_addmember</td>
<td>Adds a member to a group. The member can be a principal, a group, or a domain.</td>
</tr>
<tr>
<td>nis_removemember</td>
<td>Deletes a member from a group.</td>
</tr>
<tr>
<td>nis_creategroup</td>
<td>Create a group object.</td>
</tr>
<tr>
<td>nis_destroygroup</td>
<td>Delete a group object.</td>
</tr>
<tr>
<td>nis_verifygroup</td>
<td>Tests whether a group object exists.</td>
</tr>
<tr>
<td>nis_print_group_entry</td>
<td>Lists the principals that are members of a group object.</td>
</tr>
</tbody>
</table>
TABLE 5–2  NIS+ API Functions  (continued)

<table>
<thead>
<tr>
<th>Function</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>nis_server</td>
<td>Various services for NIS+ applications.</td>
</tr>
<tr>
<td>nis_mkdir</td>
<td>Creates the databases to support service for a named directory on a specified host.</td>
</tr>
<tr>
<td>nis_rmdir</td>
<td>Removes the directory from a host.</td>
</tr>
<tr>
<td>nis_servstate</td>
<td>Sets and reads state variables of NIS+ servers and flushes internal cashes.</td>
</tr>
<tr>
<td>nis_stats</td>
<td>Retrieves statistics about a server’s performance.</td>
</tr>
<tr>
<td>nis_getservlist</td>
<td>Returns a list of servers that support a particular domain.</td>
</tr>
<tr>
<td>nis_freeservlist</td>
<td>Frees the list of servers returned by nis_getservlist.</td>
</tr>
<tr>
<td>nis_freetags</td>
<td>Frees the memory associated with the results of nis_servstate and nis_stats.</td>
</tr>
<tr>
<td>nis_db</td>
<td>Interface Between the NIS+ Server and the Database. Not To Be Used By a NIS+ Client.</td>
</tr>
<tr>
<td>db_first_entry</td>
<td>Returns a copy of the first entry of the specified table.</td>
</tr>
<tr>
<td>db_next_entry</td>
<td>Returns a copy of the entry succeeding the specified entry.</td>
</tr>
<tr>
<td>db_reset_next_entry</td>
<td>Terminates a first/next entry sequence.</td>
</tr>
<tr>
<td>db_list_entries</td>
<td>Returns copies of entries that meet specified attributes.</td>
</tr>
<tr>
<td>db_remove_entry</td>
<td>Removes all entries that meet specified attributes.</td>
</tr>
<tr>
<td>db_add_entry</td>
<td>Replaces an entry in a table identified by specified attributes with a copy of the specified object, or adds the object to the table.</td>
</tr>
<tr>
<td>db_checkpoint</td>
<td>Reorganizes the contents of a table to make access to the table more efficient.</td>
</tr>
<tr>
<td>Function</td>
<td>Description</td>
</tr>
<tr>
<td>--------------</td>
<td>-----------------------------------------------------------------------------</td>
</tr>
<tr>
<td>db_standby</td>
<td>Advises the database manager to release resources.</td>
</tr>
<tr>
<td>nis_error</td>
<td>Functions that supply descriptive strings equivalent to NIS+ status values</td>
</tr>
<tr>
<td>nis_sperrno</td>
<td>Returns a pointer to the appropriate string constant.</td>
</tr>
<tr>
<td>nis_perror</td>
<td>Displays the appropriate string constant on standard output.</td>
</tr>
<tr>
<td>nis_lerror</td>
<td>Sends the appropriate string constant to syslog</td>
</tr>
<tr>
<td>nis_sperror</td>
<td>Returns a pointer to a statically allocated string to be used or to be copied with strdup.</td>
</tr>
<tr>
<td>nis_admin</td>
<td>Transaction logging functions used by servers</td>
</tr>
<tr>
<td>nis_ping</td>
<td>Used by the master server of a directory to time stamp it. This forces replicas of the directory to be updated.</td>
</tr>
<tr>
<td>nis_checkpoint</td>
<td>Forces logged data to be stored in the table on disk.</td>
</tr>
<tr>
<td>nis_subr</td>
<td>Functions To Help Operate on NIS+ Names and Objects.</td>
</tr>
<tr>
<td>nis_leaf_of</td>
<td>Returns the first label in an NIS+ name. The returned name does not have a trailing dot.</td>
</tr>
<tr>
<td>nis_name_of</td>
<td>Removes all domain-related labels and returns only the unique object portion of the name. The name passed to the function must be either in the local domain or in one of its child domains, or the function returns NULL.</td>
</tr>
<tr>
<td>nis_domain_of</td>
<td>Returns the name of the domain in which an object resides. The returned name ends in a dot.</td>
</tr>
<tr>
<td>nis_dir_cmp</td>
<td>Compares any two NIS+ names. The comparison ignores case and states whether the names are the same, descendants of each other, or not related.</td>
</tr>
</tbody>
</table>
TABLE 5–2   NIS+ API Functions  (continued)

<table>
<thead>
<tr>
<th>Function</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>nis_clone_object</td>
<td>Creates an exact duplicate of an NIS+ object.</td>
</tr>
<tr>
<td>nis_destroy_object</td>
<td>Destroys an object created by nis_clone_object.</td>
</tr>
<tr>
<td>nis_print_object</td>
<td>Prints the contents of an NIS+ object structure to stdout.</td>
</tr>
</tbody>
</table>

NIS+ Sample Program

This program performs the following tasks:

- Determines the local principal and local domain
- Looks up the local directory object
- Creates a directory called foo under the local domain
- Creates the groups_dir and org_dir directories under domain foo
- Creates a group object admins.foo
- Adds the local principal to the admins group
- Creates a table under org_dir.foo
- Adds two entries to the org_dir.foo table
- Retrieves and displays the new membership list of the admins group
- Lists the namespace under the foo domain using callbacks
- Lists the contents of the table created using callbacks
- Cleans up all the objects that were created by removing the following:
  - the local principal from the admins group
  - the admins group
  - the entries in the table followed by the table
  - the groups_dir and org_dir directory objects
  - the foo directory object

The example program is not a typical application. In a normal situation the directories and tables would be created or removed through the command line interface, and applications would manipulate NIS+ entry objects.
Unsupported Macros

The sample program uses unsupported macros that are defined in the file <rpcsvc/nis.h>. These are not public APIs and can change or disappear in the future. They are used for illustration purposes only and if you choose to use them, you do so at your own risk. The macros used are:

- NIS_RES_OBJECT
- ENTRY_VAL
- DEFAULT_RIGHTS

Functions Used in the Example

The use of the following NIS+ C API functions is illustrated through this example:

nis_add nis_add_entry nis_addmember
nis_creategroup nis_destroygroup nis_domain_of
nis_freeresult nis_leaf_of nis_list
nis_local_directory nis_local_principal nis_lookup
nis_mkdir nis_perror nis_remove
nis_remove_entry nis_removemember

Program Compilation

The program shown in Code Example 5–1 assumes that the NIS+ principal running this application has permission to create directory objects in the local domain. The program is compiled:

yourhost% cc -o example.c example -lnsl

It is invoked:

yourhost% example [dir]

where dir is the NIS+ directory in which the program creates all the NIS+ objects. Specifying no directory argument causes the objects to be created in the parent directory of the local domain. Note that for the call to nis_lookup, a space and the name of the local domain are appended to the string that names the directory. The argument is the name of the NIS+ directory in which to create the NIS+ objects. The principal running this program should have create permission in the directory.
#include <stdio.h>
#include <string.h>
#include <stdlib.h>
#include <rpcsvc/nis.h>

#define MAX_MESG_SIZE 512
#define BUFFER_SIZE 64
#define TABLE_TYPE "test_data"

main(argc, argv)
int argc;
char *argv[];
{
    char *saved_grp, *saved_name, *saved_owner;
    char dir_name[NIS_MAXNAMELEN];
    char local_domain[NIS_MAXNAMELEN];
    char local_princip [NIS_MAXNAMELEN];
    char org_dir_name [NIS_MAXNAMELEN];
    char grp_name [NIS_MAXNAMELEN];
    char grp_dir_name [NIS_MAXNAMELEN];
    char table_name [NIS_MAXNAMELEN];
    nis_object *dirobj, entdata;
    nis_result *pres;
    u_int saved_num_servers;
    int err;

    if (argc == 2)
        sprintf (local_domain, "%s.", argv[1]);
    else
        strcpy (local_domain, "");

    strcat (local_domain, (char *) nis_local_directory());

    strcpy (local_princip, (char *) nis_local_principal());

    /*
     * Lookup the directory object for the local domain for two
     * reasons:
     * 1.To get a template of a nis_object.
     * 2.To reuse some of the information contained in the
     * directory
     * object returned. We could have declared a static
     * nis_object, but
     * since we need to change very little, it is easier to make
     * changes and not initialize the nis_object structure.
     */
    pres = nis_lookup (local_domain, 0);
    if (pres->status != NIS_SUCCESS) {
        nis_perror (pres->status, "unable to lookup local
                directory");
        exit (1);
    }
}
/* re-use most of the fields in the parent directory object -
save
* pointers to the fields that are being changed so that we can
* free the original object and avoid dangling pointer
* references.
*/
dirobj = NIS_RES_OBJECT (pres);
saved_name = dirobj->DI_data.do_name;
saved_owner = dirobj->zo_owner;
saved_grp = dirobj->zo_group;

/* set the new name, group, owner and new access rights for the
* foo domain.
*/
sprintf (dir_name, "%s.%s", "foo", local_domain);
sprintf (grp_name, "%s.%s", "admins", dir_name);
dirobj->DI_data.do_name = dir_name;
dirobj->zo_group = grp_name;
dirobj->zo_owner = local_princip;

/* Access rights in NIS+ are stored in a u_long with the
* order bytes reserved for the "nobody" category, the next
* eight bytes reserved for the owner, followed by group and world.
* In this example we are giving access to the directory based
* on the
* "----rmcdrmcd----" access right pattern.
*/
dirobj->zo_access = ((NIS_READ_ACC + NIS_MODIFY_ACC
+ NIS_CREATE_ACC + NIS_DESTROY_ACC) << 16)
| ((NIS_READ_ACC + NIS_MODIFY_ACC
+ NIS_CREATE_ACC + NIS_DESTROY_ACC) << 8);

/* Save the number of servers the parent directory object had
so that we can restore this value before calling
nis_freeresult()
* later and avoid memory leaks.
*/
saved_num_servers = dirobj->DI_data.do_servers.do_servers_len;

/* We want only one server to serve this directory */
dirobj->DI_data.do_servers.do_servers_len = 1;
dir_create (dir_name, dirobj);

/* create the groups_dir and org_dir directories under foo. */
sprintf (grp_dir_name, "groups_dir.%s", dir_name);
dirobj->DI_data.do_name = grp_dir_name;
dir_create (grp_dir_name, dirobj);

sprintf (org_dir_name, "org_dir.%s", dir_name);
dirobj->DI_data.do_name = org_dir_name;
dir_create (org_dir_name, dirobj);

grp_create (grp_name);

printf ("Adding principal %s to group %s ... \n", 
    local_princip, grp_name);
err = nis_addmember (local_princip, grp_name);

if (err != NIS_SUCCESS) {
    nis perror (err, 
        "unable to add local principal to group.");
    exit (1);
}

sprintf (table_name, "test_table.org_dir.%s", dir_name);
tbl_create (dirobj, table_name);

/*
 * Now create NIS+ entry objects in the table that was just
 * created
 *
 */
stuff_table (table_name);

/* Display what we stuffed */
list_objs(dir_name, table_name, grp_name);

/* Clean out everything we created. */
cleanup (local_princip, grp_name, table_name, dir_name, 
dirobj);

/*
 * Restore the saved pointers from the original pres structure
 * so that we can free up the associated memory and have no
 * memory leaks.
 *
 */
dirobj->DI_data.do_name = saved_name;
dirobj->zo_group = saved_grp;
dirobj->zo_owner = saved_owner;
dirobj->DI_data.do_servers.do_servers_len = 
    saved_num_servers; 
(void) nis_freeresult (pres);
}
Code Example 5–2 shows the routine is called by main to create directory objects.

**CODE EXAMPLE 5–2  NIS+ Routine to Create Directory Objects**

```c
void
dir_create (dir_name, dirobj)
nis_name dir_name;
nis_object *dirobj;
{
    nis_result *cres;
    nis_error err;

    printf ("\n Adding Directory %s to namespace ... \n", dir_name);
    cres = nis_add (dir_name, dirobj);
    if (cres->status != NIS_SUCCESS) {
        nis_perror (cres->status, "unable to add directory foo.");
        exit (1);
    }

    (void) nis_freeresult (cres);
    /*
     * NOTE: you need to do a nis_mkdir to create the table to
     * store the contents of the directory you are creating.
     */
    err = nis_mkdir (dir_name, dirobj->DI_data.do_servers.do_servers_val);
    if (err != NIS_SUCCESS) {
        (void) nis_remove (dir_name, 0);
        nis_perror (err, "unable to create table for directory object foo.");
        exit (1);
    }
}
```

This routine is called by main to create the group object. Since nis_creategroup works only on group objects, the “groups_dir” literal is not needed in the group name.

**CODE EXAMPLE 5–3  NIS+ Routine to Create Group Objects**

```c
void
grp_create (grp_name)
```
The routine shown in Code Example 5–3 is called by main to create a table object laid out as shown in Table 5–3.

**TABLE 5–3** NIS+ Table Objects

<table>
<thead>
<tr>
<th>Column1</th>
<th>Column2</th>
</tr>
</thead>
<tbody>
<tr>
<td>Name: id</td>
<td>name</td>
</tr>
<tr>
<td>Attributes:</td>
<td>Searchable, Text</td>
</tr>
<tr>
<td>Access Rights</td>
<td>---rmcdr---r---</td>
</tr>
</tbody>
</table>

The **TA_SEARCHABLE** constant indicates to the service that the column is searchable. Only **TEXT** (the default) columns are searchable. **TA_CASE** indicates to the service that the column value is to be treated in a case-insensitive manner during searches.

**CODE EXAMPLE 5–4** NIS+ Routine to Create Table Objects

```c
#define TABLE_MAXCOLS 2
#define TABLE_COLSEP ':'
#define TABLE_PATH 0

void
tbl_create (dirobj, table_name)
nis_object *dirobj; /* need to use some of the fields */
nis_name table_name;
{
    nis_result  *cres;
    static nis_object  tblobj;
    static table_col  tbl_cols[TABLE_MAXCOLS] = {
        {'Id', TA_SEARCHABLE | TA_CASE, DEFAULT_RIGHTS},
        {'Name', TA_SEARCHABLE | TA_CASE, DEFAULT_RIGHTS}
    };

    tblobj.zo_owner = dirobj->zo_owner;
```
The routine shown in Code Example 5–5 is called by main() to add entry objects to the table object. Two entries are added to the table object. Note that the column width in both entries is set to include the NULL character for a string terminator.

**CODE EXAMPLE 5–5  NIS+ Routine to Add Objects to Table**

```c
#define MAXENTRIES 2
void
stuff_table(table_name)
nis_name table_name;
{
  int i;
  nis_object entdata;
  nis_result *cres;
  static entry_col ent_col_data[MAXENTRIES][TABLE_MAXCOLS] = {
    {0, 2, "1", 0, 5, "John"},
    {0, 2, "2", 0, 5, "Mary"}
  };

  printf ("\n Adding entries to table ... \n");
  /*
   * Look up the table object first since the entries being added
   * should have the same owner, group owner and access rights as
   * the table they go in.
   */

  tblobj.zo_group = dirobj->zo_group;
tblobj.zo_access = DEFAULT_RIGHTS; /* macro defined in nis.h */
tblobj.zo_data.zo_type = TABLE_OBJ; /* enumerated type in nis.h */
tblobj.TA_data.ta_type = TABLE_TYPE;
tblobj.TA_data.ta_maxcol = TABLE_MAXCOLS;
tblobj.TA_data.ta_sep = TABLE_COLSEP;
tblobj.TA_data.ta_path = TABLE_PATH;
tblobj.TA_data.ta_cols.ta_cols_len = tblobj.TA_data.ta_maxcol; /* ALWAYS */
tblobj.TA_data.ta_cols.ta_cols_val = tbl_cols;

  /*
   * Use a fully qualified table name i.e. the "org_dir" literal
   * should be embedded in the table name. This is necessary because
   * nis_add * operates on all types of NIS+ objects and needs the full path
   * name
   * if a table is created.
   */
  printf ("\n Creating table %s ... \n", table_name);
cres = nis_add (table_name, &tblobj);
  if (cres->status != NIS_SUCCESS) {
    nis_perror (cres->status, "unable to add table.");
    exit (1);
  }
  (void) nis_freeresult (cres);
}
```

The routine shown in Code Example 5–5 is called by main() to add entry objects to the table object. Two entries are added to the table object. Note that the column width in both entries is set to include the NULL character for a string terminator.
The routine shown in Code Example 5–6 is the print function for the
nis_list call. When list_objs calls nis_list, a pointer to
print_info is one of the calling arguments. Each time the service calls this
function, it prints the contents of the entry object. The return value indicates
to the library to call with the next entry from the table.

**CODE EXAMPLE 5–6  NIS+ Routine for nis_list Call**

```c
int
print_info (name, entry, cbdata)
    nis_name name; /* Unused */
    nis_object *entry; /* The NIS+ entry object */
    void *cbdata; /* Unused */
    {  
        static u_int firsttime = 1;
        entry_col *tmp; /* only to make source more readable */
        u_int i, terminal;

        if (firsttime) {  
            printf ("Id.			Name\n");
            printf ("---			----\n");
            firsttime = 0;
        }
        for (i = 0; i < entry->EN_data.en_cols.en_cols_len; ++i) {  
            tmp = &entry->EN_data.en_cols.en_cols_val[i];
```
The routine shown in Code Example 5–7 is called by main to list the contents of the group, table and directory objects. The routine demonstrates the use of callbacks also. It retrieves and displays the membership of the group. The group membership list is not stored as the contents of the object. So, it is queried through the nis_lookup instead of the nis_list call. You must use the “groups_dir” form of the group name since nis_lookup works on all types of NIS+ objects.

CODE EXAMPLE 5–7  NIS+ Routine to List Objects

```c
void
list_objs(dir_name, table_name, grp_name)
    nis_name dir_name, table_name, grp_name;
{
    group_obj *tmp; /* only to make source more readable */
    u_int i;
    char grp_obj_name [NIS_MAXNAMELEN];
    nis_result *cres;
    char index_name [BUFFER_SIZE];

    sprintf (grp_obj_name, "%s.groups_dir.%s",
        nis_leaf_of (grp_name), nis_domain_of (grp_name));

    cres = nis_lookup(grp_obj_name, 0);
    if (cres->status != NIS_SUCCESS) {
        nis_perror (cres->status, "Unable to lookup group object.");
        exit(1);
    }

    tmp = &(NIS_RES_OBJECT(cres)->GR_data);
    for (i = 0; i < tmp->gr_members.gr_members_len; ++i)
        printf ("%s\n", tmp->gr_members.gr_members_val[i]);
    (void) nis_freeresult (cres);

    printf ("\nContents of Directory %s are: \n", dir_name);
    cres = nis_list (dir_name, 0, 0, 0);
    if (cres->status != NIS_SUCCESS) {
        nis_perror (cres->status,
```
"Unable to list Contents of Directory foo.");
    exit(1);
}
for (i = 0; i < NIS_RES_NUMOBJ(cres); ++i)
    printf("\t%s\n", NIS_RES_OBJECT(cres)[i].zo_name);
(void) nis_freeresult (cres);

/*
 * List the contents of the table we created using the callback form
 * of nis_list().
 */
printf ("\n Contents of Table %s are: \n", table_name);
    cres = nis_list (table_name, 0, print_info, 0);
    if(cres->status != NIS_CBRESULTS && cres->status !=
    NIS_NOTFOUND)
        nis_perror (cres->status,
            "Listing entries using callback failed");
        exit(1);
    }
    (void) nis_freeresult (cres);

/*
 * List only one entry from the table we created. We will use indexed names to do this retrieval.
 */
printf("\n Entry corresponding to id 1 is:\n");
    /*
 * The name of the column is usually extracted from the table object, which would have to be retrieved first.
 */
    sprintf(index_name, 
            
   [Id=1],%s", table_name);
    cres = nis_list (index_name, 0, print_info, 0);
    if(cres->status != NIS_CBRESULTS && cres->status !=
    NIS_NOTFOUND)
        nis_perror (cres->status,
            "Listing entry using indexed names and callback failed");
        exit(1);
    }
    (void) nis_freeresult (cres);

The routine in Code Example 5-8 is called by cleanup to remove a directory object from the namespace. It also informs the servers serving the directory about this deletion. Notice that the memory containing result structure, pointed to by cres, must be freed after the result has been tested.

CODE EXAMPLE 5-8  NIS+ Routine to Remove Directory Objects

void
dir_remove(dir_name, srv_list, numservers)
nis_name dir_name;
nis_server *srv_list;

u_int numservers;
{
    nis_result *cres;
    nis_error err;
    u_int i;

    printf("\nRemoving $s directory object from namespace ... \n", dir_name);
    cres = nis_remove (dir_name, 0);
    if (cres->status != NIS_SUCCESS) {
        nis_perror (cres->status, "unable to remove directory");
        exit (1);
    }
    (void) nis_freeresult (cres);

    for (i = 0; i < numservers; ++i) {
        err = nis_rmdir (dir_name, &srv_list[i]);
        if (err != NIS_SUCCESS) {
            nis_perror (err, "unable to remove server from directory");
            exit (1);
        }
    }
}

This routine, Code Example 5–9, is called by main to delete all the objects that were created in this example. Note the use of the REM_MULTIPLE flag in the call to nis_remove_entry. All entries must be deleted from a table before the table itself can be deleted.

**CODE EXAMPLE 5–9  NIS+ Routine to Remove All Objects**

```c
void
cleanup(local_princip, grp_name, table_name, dir_name, dirobj)
    nis_name local_princip, grp_name, table_name, dir_name;
    nis_object *dirobj;
{
    char grp_dir_name [NIS_MAXNAMELEN];
    char org_dir_name [NIS_MAXNAMELEN];
    nis_error err;
    nis_result *cres;

    sprintf(grp_dir_name, "%s.%s", "groups_dir", dir_name);
    sprintf(org_dir_name, "%s.%s", "org_dir", dir_name);

    printf("\n\nStarting to Clean up ... \n");
    printf("\nRemoving principal $s from group $s \n", local_princip, grp_name);
    err = nis_removemember (local_princip, grp_name);

    if (err != NIS_SUCCESS) {
        nis_perror (err, "unable to delete local principal from group.");
        exit (1);
    }
}
```
/*
 * Delete the admins group. We do not use the "groups_dir" form
 * of the group name since this API is applicable to groups
 * only.
 * It automatically embeds the groups_dir literal in the name
 * of
 * the group.
 */
printf("\nRemoving %s group from namespace ... \n",
grp_name);
err = nis_destroygroup (grp_name);
if (err != NIS_SUCCESS) {
nis_perror (err, "unable to delete group.");
exit (1);
}

printf("\n Deleting all entries from table %s ... \n",
table_name);
cres = nis_remove_entry(table_name, 0, REM_MULTIPLE);
switch (cres->status) {
case NIS_SUCCESS:
case NIS_NOTFOUND:
break;
default:
nis_perror(cres->status, "Could not delete entries from
 table");
exit (1);
}
(void) nis_freeresult (cres);
printf("\n Deleting table %s itself ... \n", table_name);
cres = nis_remove(table_name, 0);
if (cres->status != NIS_SUCCESS) {
nis_perror(cres->status, "Could not delete table.");
exit (1);
}
(void) nis_freeresult (cres);

/* delete the groups_dir, org_dir and foo directory objects. */
dir_remove (grp_dir_name,
dirobj->DI_data.do_servers.do_servers_val,
dirobj->DI_data.do_servers.do_servers_len);
dir_remove (org_dir_name,
dirobj->DI_data.do_servers.do_servers_val,
dirobj->DI_data.do_servers.do_servers_len);
dir_remove (dir_name, dirobj->
>DI_data.do_servers.do_servers_val,
dirobj->DI_data.do_servers.do_servers_len);
}

Running the program displays on the screen, as shown in Figure 5-3.
myhost% **domainname**
sun.com
myhost% ./*sample
Adding Directory foo.sun.com. to namespace ...
Adding Directory groups_dir.foo.sun.com. to namespace ...
Adding Directory org_dir.foo.sun.com. to namespace ...
Adding admins.foo.sun.com. group to namespace ...
Adding principal myhost.sun.com. to group admins.foo.sun.com. ...
Creating table test_table.org_dir.foo.sun.com. ...
Adding entries to table ...
Group admins.foo.sun.com. membership is:
        myhost.sun.com.
Contents of Directory foo.sun.com. are:
groups_dir
 org_dir

Contents of Table test_table.org_dir.foo.sun.com. are:
<table>
<thead>
<tr>
<th>Id.</th>
<th>Name</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>John</td>
</tr>
<tr>
<td>2</td>
<td>Mary</td>
</tr>
</tbody>
</table>

Entry corresponding to id 1 is:
        1 John

Starting to Clean up ...

Removing admins.foo.sun.com. group from namespace ...
Deleting all entries from table test_table.org_dir.foo.sun.com. ...
Deleting table test_table.org_dir.foo.sun.com. itself ...
Removing groups_dir.foo.sun.com. directory object from namespace ...
Removing org_dir.foo.sun.com. directory object from namespace ...
Removing foo.sun.com. directory object from namespace ...

myhost%

**Figure 5–3  NIS+ Program Execution**

As a debugging aid, the same operations are performed by the following command sequence. The first command:

% **niscat** -o 'domainname'

displays the name of the master server. Substitute the master server name where the variable **master** appears below.

% **nismkdir** -m master foo."domainname".

# Create the org_dir.foo subdirectory with the specified master
% **nismkdir** -m master org_dir.foo."domainname".
# Create the groups_dir.foo subdirectory with the specified master
% **nismkdir** -m master groups_dir.foo."domainname".
# Create the "admins" group
% **nisgrpadm** -c admins.foo."domainname".
# Add yourself as a member of this group
% **nisgrpadm** -a admins.foo."domainname". "nisdefaults" -p
# Create a test_table with two columns : Id and Name

NIS+ Programming Guide 161
% nistbladm -c test_data id=SI Name=SI \ test_table.org_dir.foo.'domainname'
# Add one entry to that table.
% nistbladm -a id=1 Name=John test_table.org_dir.foo.'domainname'.
# Add another entry to that table.
% nistbladm -a id=2 Name=Mary test_table.org_dir.foo.'domainname'.

# List the members of the group admins
% nisgrpadm -l admins.foo.'domainname'.
# List the contents of the foo directory
% nisls foo.'domainname'.
# List the contents of the test_table along with its header
% niscat -h test_table.org_dir.foo.'domainname'.

# Get the entry from the test_table where id = 1
% nismatch id=1 test_table.org_dir.foo.'domainname'.

# Delete all we created.
# First, delete yourself from the admins group
% nisgrpadm -r admins.foo.'domainname'. 'nisdefaults -p'
# Delete the admins group
% nisgrpadm -d admins.foo.'domainname'.
# Delete all the entries from the test_table
% nistbladm -r '[]',test_table.org_dir.foo.'domainname'./
# Delete the test_table itself.
% nistbladm -d test_table.org_dir.foo.'domainname'.
# Delete all three directories that we created
% nisrmdir groups_dir.foo.'domainname'.
% nisrmdir org_dir.foo.'domainname'.
% nisrmdir foo.'domainname'.
XDR Technical Note

This appendix is a technical note on SunSoft’s implementation of the external data representation (XDR) standard, a set of library routines that enable C programmers to describe arbitrary data structures in a machine-independent fashion.

What is XDR

XDR is the backbone of SunSoft’s Remote Procedure Call package, in the sense that data for RPCs are transmitted using this standard. XDR library routines should be used to transmit data accessed (read or written) by more than one type of machine.

XDR works across different languages, operating systems, and machine architectures. Most users (particularly RPC users) only need the information in the sections on Number Filters, Floating Point Filters, and Enumeration Filters. Programmers wanting to implement RPC and XDR on new machines will be interested in this technical note and the protocol specification.

rpcgen can be used to write XDR routines even in cases where no RPC calls are being made.

C programs that use XDR routines must include the file <rpc/xdr.h>, which contains all the necessary interfaces to the XDR system. Since the library libnsl.a contains all the XDR routines, compile as follows:

example% cc program.c

In many environments several criteria must be observed to accomplish porting. It is not always easy to see the ramifications of a small programmatic change, but they can often have far reaching implications.
Consider the examples of a program to read/write a line of text, shown in Code Example A–0 and Code Example A–0.

**CODE EXAMPLE A–1**  Writer Example (initial)

```c
#include <stdio.h>

main() /* writer.c */
{
    long i;
    for (i = 0; i < 8; i++) {
        if (fwrite((char *) &i, sizeof(i), 1, stdout) != 1) {
            fprintf(stderr, "failed!
");
            exit(1);
        }
    }
    exit(0);
}
```

**CODE EXAMPLE A–2**  Reader Example (initial)

```c
#include <stdio.h>

main() /* reader.c */
{
    long i, j;
    for (j = 0; j < 8; j++) {
        if (fread((char *) &i, sizeof(i), 1, stdin) != 1) {
            fprintf(stderr, "failed!
");
            exit(1);
        }
        printf("%ld ", i);
    }
    printf("\n");
    exit(0);
}
```

The two programs appear to be portable, because (a) they pass lint checking, and (b) they exhibit the same behavior when executed on two different hardware architectures, a SPARC and a VAX.

Piping the output of the writer program to the reader program gives identical results on SPARC or VAX.

```
sun% writer | reader
0 1 2 3 4 5 6 7
sun%
vax% writer | reader
0 1 2 3 4 5 6 7
vax%
```
With the advent of local area networks and 4.2BSD came the concept of “network pipes”—a process produces data on one machine, and a second process consumes data on another machine. A network pipe can be constructed with `writer` and `reader`. Here are the results if the first produces data on a SPARC, and the second consumes data on a VAX.

```
sun% writer | rsh vax reader
0 16777216 33554432 50331648 67108864 83886080 100663296
117440512
sun%
```

Identical results can be obtained by executing `writer` on the VAX and `reader` on the SPARC. These results occur because the byte ordering of long integers differs between the VAX and the SPARC, even though word size is the same. Note that 16777216 is $2^{24}$—when four bytes are reversed, the 1 is placed in the 24th bit.

Whenever data is shared by two or more machine types, there is a need for portable data. Programs can be made data-portable by replacing the `read` and `write` calls with calls to an XDR library routine, `xdr_long`, a filter that knows the standard representation of a long integer in its external form. The revised versions of `writer` are shown in Code Example A–0.

**CODE EXAMPLE A–3**  Writer Example (XDR modified)

```
#include <stdio.h>
#include <rpc/rpc.h> /* xdr is a sub-library of rpc */

main() /* writer.c */
{
    XDR xdrs;
    long i;

    xdrstdio_create(&xdrs, stdout, XDR_ENCODE);
    for (i = 0; i < 8; i++) {
        if (!xdr_long(&xdrs, &i)) {
            fprintf(stderr, "failed!
");
            exit(1);
        }
    }
    exit(0);
}
```

Code Example A–0 shows the revised versions of `reader`.

**CODE EXAMPLE A–4**  Reader Example (XDR modified)
The new programs were executed on a SPARC, on a VAX, and from a SPARC to a VAX; the results are shown below.

```
sun% writer | reader
0 1 2 3 4 5 6 7
sun%
vax% writer | reader
0 1 2 3 4 5 6 7
vax%
sun% writer | rsh vax reader
0 1 2 3 4 5 6 7
sun%
```

**Note** - Integers are just the tip of the portable-data iceberg. Arbitrary data structures present portability problems, particularly with respect to alignment and pointers. Alignment on word boundaries may cause the size of a structure to vary from machine to machine. And pointers, which are very convenient to use, have no meaning outside the machine where they are defined.

---

**A Canonical Standard**

XDR’s approach to standardizing data representations is canonical. That is, XDR defines a single byte order, a single floating-point representation (IEEE), and so on. Any program running on any machine can use XDR to create
portable data by translating its local representation to the XDR standard representations. Similarly, any program running on any machine can read portable data by translating the XDR standard representations to its local equivalents. The single standard completely decouples programs that create or send portable data from those that use or receive portable data. The advent of a new machine or a new language has no effect upon the community of existing portable-data creators and users. A new machine joins this community by being “taught” how to convert the standard representations and its local representations; the local representations of other machines are irrelevant. Conversely, to existing programs running on other machines, the local representations of the new machine are also irrelevant; such programs can immediately read portable data produced by the new machine because such data conforms to the canonical standards that they already understand.

There are strong precedents for XDR’s canonical approach. For example, TCP/IP, UDP/IP, XNS, Ethernet, and, indeed, all protocols below layer five of the ISO model, are canonical protocols. The advantage of any canonical approach is simplicity; in the case of XDR, a single set of conversion routines is written once and is never touched again. The canonical approach has a disadvantage, but it is unimportant in real-world data transfer applications. Suppose two Little-Endian machines are transferring integers according to the XDR standard. The sending machine converts the integers from Little-Endian byte order to XDR (Big-Endian) byte order; the receiving machine performs the reverse conversion. Because both machines observe the same byte order, their conversions are unnecessary. The point, however, is not necessity, but cost as compared to the alternative.

The time spent converting to and from a canonical representation is insignificant, especially in distributed applications. Most of the time required to prepare a data structure for transfer is not spent in conversion but in traversing the elements of the data structure. To transmit a tree, for example, each leaf must be visited and each element in a leaf record must be copied to a buffer and aligned there; storage for the leaf may have to be de-allocated as well. Similarly, to receive a tree, storage must be allocated for each leaf, data must be moved from the buffer to the leaf and properly aligned, and pointers must be constructed to link the leaves together. Every machine pays the cost of traversing and copying data structures whether or not conversion is required. In distributed applications, communications overhead—the time required to move the data down through the sender’s protocol layers, across the network and up through the receiver’s protocol layers—dwarfs conversion overhead.
The XDR Library

The XDR library not only solves data portability problems, it also allows you to write and read arbitrary C constructs in a consistent, specified, well-documented manner. Thus, it can make sense to use the library even when the data is not shared among machines on a network.

The XDR library has filter routines for strings (null-terminated arrays of bytes), structures, unions, and arrays, to name a few. Using more primitive routines, you can write your own specific XDR routines to describe arbitrary data structures, including elements of arrays, arms of unions, or objects pointed at from other structures. The structures themselves may contain arrays of arbitrary elements, or pointers to other structures.

Look closely at the two programs. There is a family of XDR stream creation routines in which each member treats the stream of bits differently. In the example, data is manipulated using standard I/O routines, so you use xdrstdio_create. The parameters to XDR stream creation routines vary according to their function. In the example, xdrstdio_create takes a pointer to an XDR structure that it initializes, a pointer to a FILE that the input or output is performed on, and the operation. The operation may be XDR_ENCODE for serializing in the writer program, or XDR_DECODE for deserializing in the reader program.

Note - RPC users never need to create XDR streams; the RPC system itself creates these streams, which are then passed to the users.

The xdr_long primitive is characteristic of most XDR library primitives and all client XDR routines. First, the routine returns FALSE (0) if it fails, and TRUE (1) if it succeeds. Second, for each data type, xxx, there is an associated XDR routine of the form:

```c
xdr_xxx(xdrs, xp)
    XDR *xdrs;
    xxx *xp;
    }
```

In this case, xxx is long, and the corresponding XDR routine is a primitive, xdr_long. The client could also define an arbitrary structure xxx in which case the client would also supply the routine xdr_xxx, describing each field by calling XDR routines of the appropriate type. In all cases the first parameter, xdrs can be treated as an opaque handle, and passed to the primitive routines.

XDR routines are direction independent; that is, the same routines are called to serialize or deserialize data. This feature is critical to software engineering of portable data. The idea is to call the same routine for either
operation—this almost guarantees that serialized data can also be deserialized. One routine is used by both producer and consumer of networked data. This is implemented by always passing the address of an object rather than the object itself—only in the case of deserialization is the object modified. This feature is not shown in our trivial example, but its value becomes obvious when nontrivial data structures are passed among machines. If needed, the user can obtain the direction of the XDR operation. For details, see the section, “Operation Directions” on page 183.

A slightly more complicated example follows. Assume that a person’s gross assets and liabilities are to be exchanged among processes. Also assume that these values are important enough to warrant their own data type:

```c
struct gnumbers {
    long g_assets;
    long g_liabilities;
};
```

The corresponding XDR routine describing this structure is:

```c
bool_t /* TRUE is success, FALSE is failure */
xdr_gnumbers(xdrs, gp)
    XDR *xdrs;
    struct gnumbers *gp;
{
    if (xdr_long(xdrs, &gp->g_assets) && 
        xdr_long(xdrs, &gp->g_liabilities)) 
        return(TRUE);
    return(FALSE);
}
```

Note that the parameter `xdrs` is never inspected or modified; it is only passed on to the subcomponent routines. It is imperative to inspect the return value of each XDR routine call, and to give up immediately and return `FALSE` if the subroutine fails.

This example also shows that the type `bool_t` is declared as an integer whose only values are `TRUE` (1) and `FALSE` (0). This document uses the following definitions:

```c
#define bool_t int
#define TRUE 1
#define FALSE 0
```

Keeping these conventions in mind, `xdr_gnumbers` can be rewritten as follows:

```c
xdr_gnumbers(xdrs, gp)
    XDR *xdrs;
    struct gnumbers *gp;
{
    return(xdr_long(xdrs, &gp->g_assets) && 
        xdr_long(xdrs, &gp->g_liabilities));
}
```

This document uses both coding styles.
XDR Library Primitives

This section gives a synopsis of each XDR primitive. It starts with memory allocation and the basic data types, then moves on to constructed data types. Finally, XDR utilities are discussed. The interface to these primitives and utilities is defined in the include file <rpc/xdr.h>, automatically included by <rpc/rpc.h>.

Memory Requirements for XDR Routines

When using XDR routines, there is sometimes a need to pre-allocate memory (or to determine memory requirements). In these instances where the developer needs to control the allocation and de-allocation of memory for XDR conversion routines to use there is a routine, xdr_sizeof, that is used to return the number of bytes needed to encode and decode data using one of the XDR filter functions (func). xdr_sizeof’s output does not include RPC headers or record markers and they must be added in to get a complete accounting of the memory required. xdr_sizeof returns a zero on error.

xdr_sizeof(xdrproc_t func, void *data)

xdr_sizeof is specifically useful the allocation of memory in applications that use XDR outside of the RPC environment; to select between transport protocols; and at the lower levels of RPC to perform client and server creation functions.

Code Example A–0 and Code Example A–0 illustrate two uses of xdr_sizeof.

CODE EXAMPLE A–5 xdr_sizeof Example #1

```c
#include <rpc/rpc.h>

/*
 * This function takes as input a CLIENT handle, an XDR function and
 * a pointer to data to be XDR'd. It returns TRUE if the amount of
 * data to be XDR'd may be sent using the transport associated
 * with the CLIENT handle, and false otherwise.
 */
bool_t
cansend(cl, xdrfunc, xdrdata)
    CLIENT *cl;
    xdrproc_t xdrfunc;
```

(Continued)
void *xdrdata;
{
int fd;
struct t_info tinfo;

if (clnt_control(cl, CLGET_FD, &fd) == -1) {
    /* handle clnt_control() error */
    return (FALSE);
}

if (t_getinfo(fd, &tinfo) == -1) {
    /* handle t_getinfo() error */
    return (FALSE);
} else {
    if (tinfo.servtype == T_CLTS) {
        /* This is a connectionless transport. Use xdr_sizeof() * 
        * to compute the size of this request to see whether it   
        * is too large for this transport.                    
        */
        switch(tinfo.tsdu) {
        case 0: /* no concept of TSDUs */          
            return (FALSE);
            break;
        case -2: /* can't send normal data */      
            return (FALSE);
            break;
        case -1: /* no limit on TSDU size */       
            return (TRUE);
            break;
        default:
            if (tinfo.tsdu < xdr_sizeof(xdrfunc, xdrdata))
                return (FALSE);
            else
                return (TRUE);
        }
    } else {                                    
        return (TRUE);
    }
}

Code Example A-0 is the second xdr_sizeof example.

CODE EXAMPLE A-6  xdr_sizeof Example #2

#include <sys/statvfs.h>
#include <sys/sysmacros.h>

/*
 * This function takes as input a file name, an XDR function, and a
 * pointer to data to be XDR'd. It returns TRUE if the filesystem
(Continued)
Number Filters

The XDR library provides primitives to translate between numbers and their corresponding external representations. Primitives cover the set of numbers in the types:

\[[\text{signed, unsigned}] \times \text{[short, int, long]}\]

Specifically, the eight primitives are:

```c
bool_t xdr_char(xdrs, op)
    XDR *xdrs;
    char *cp;

bool_t xdr_u_char(xdrs, ucp)
    XDR *xdrs;
    unsigned char *ucp;

bool_t xdr_int(xdrs, ip)
    XDR *xdrs;
    int *ip;

bool_t xdr_u_int(xdrs, up)
    XDR *xdrs;
    unsigned *up;

bool_t xdr_long(xdrs, lip)
    XDR *xdrs;
    long *lip;

bool_t xdr_u_long(xdrs, lup)
    XDR *xdrs;
```

* on which the file resides has room for the additional amount of data to be XDR’d. Note that since the information `statvfs(2)` returns about the filesystem is in blocks you must convert the `*` value returned by `xdr(sizeof)` from bytes to disk blocks.

```c
bool_t canwrite(file, xdrfunc, xdrdata)
    char *file;
    xdrproc_t xdrfunc;
    void *xdrdata;
{
    struct statvfs s;

    if (statvfs(file, &s) == -1) {
        /* handle statvfs() error */
        return (FALSE);
    }

    if (s.f_bavail >= btod(xdr(sizeof(xdrfunc, xdrdata))))
        return (TRUE);
    else
        return (FALSE);
}
```
The first parameter, \textit{xdrs}, is an XDR stream handle. The second parameter is the address of the number that provides data to the stream or receives data from it. All routines return \textit{TRUE} if they complete successfully, and \textit{FALSE} otherwise.

**Floating Point Filters**

The XDR library also provides primitive routines for C floating point types:

```c
bool_t xdr_float(xdrs, fp)
XDR *xdrs;
float *fp;
```

The first parameter, \textit{xdrs} is an XDR stream handle. The second parameter is the address of the floating point number that provides data to the stream or receives data from it. Both routines return \textit{TRUE} if they complete successfully, and \textit{FALSE} otherwise.

\textbf{Note} - Since the numbers are represented in IEEE floating point, routines may fail when decoding a valid IEEE representation into a machine-specific representation, or vice versa.

**Enumeration Filters**

The XDR library provides a primitive for generic enumerations. The primitive assumes that a C \textit{enum} has the same representation inside the machine as a C integer. The Boolean type is an important instance of the \textit{enum}. The external representation of a Boolean is always \textit{TRUE} (1) or \textit{FALSE} (0).

```c
#define bool_t int
#define FALSE 0
#define TRUE 1
#define enum_t int
bool_t xdr_enum(xdrs, ep)
```
The second parameters $ep$ and $bp$ are addresses of the associated type that provides data to, or receives data from, the stream $xdrs$.

**No-Data Routine**

Occasionally, an XDR routine must be supplied to the RPC system, even when no data is passed or required. The library provides such a routine:

```c
bool_t xdr_void(); /* always returns TRUE */
```

**Constructed Data Type Filters**

Constructed or compound data type primitives require more parameters and perform more complicated functions than the primitives discussed previously. This section includes primitives for strings, arrays, unions, and pointers to structures.

Constructed data type primitives may use memory management. In many cases, memory is allocated when deserializing data with XDR_DECODE. Therefore, the XDR package must provide means to de-allocate memory. This is done by an XDR operation, XDR_FREE. To review, the three XDR directional operations are XDR_ENCODE, XDR_DECODE, and XDR_FREE.

**Strings**

In the C language, a string is defined as a sequence of bytes terminated by a null byte, which is not considered when calculating string length. However, when a string is passed or manipulated, a pointer to it is employed. Therefore, the XDR library defines a string to be a char *, and not a sequence of characters. The external representation of a string is drastically different from its internal representation.

Externally strings are represented as sequences of ASCII characters, while internally they are represented with character pointers. Conversion between the two representations is accomplished with the routine `xdr_string`:

```c
bool_t xdr_string(xdrs, sp, maxlength)
XDR *xdrs;
char **sp;
_u_int maxlength;
```

The first parameter $xdrs$ is the XDR stream handle. The second parameter $sp$ is a pointer to a string (type char **). The third parameter $maxlength$
specifies the maximum number of bytes allowed during encoding or
decoding. Its value is usually specified by a protocol. For example, a protocol
specification may say that a file name may be no longer than 255 characters.
The routine returns FALSE if the number of characters exceeds maxlength,
and TRUE if it doesn’t.

The behavior of xdr_string is similar to the behavior of other routines
discussed in this section. The direction XDR_ENCODE is easiest to understand.
The parameter sp points to a string of a certain length; if the string does not
exceed maxlength, the bytes are serialized.

The effect of deserializing a string is subtle. First the length of the incoming
string is determined; it must not exceed maxlength. Next sp is dereferenced; if
the value is NULL, a string of the appropriate length is allocated and *sp is set
to this string. If the original value of *sp is nonnull, the XDR package
assumes that a target area has been allocated, which can hold strings no
longer than maxlength. In either case, the string is decoded into the target
area. The routine then appends a null character to the string.

In the XDR_FREE operation the string is obtained by dereferencing sp. If the
string is not NULL, it is freed and *sp is set to NULL. In this operation
xdr_string ignores the maxlength parameter.

Note that you can use XDR on an empty string (""") but not on a NULL string.

Byte Arrays

Often variable-length arrays of bytes are preferable to strings. Byte arrays
differ from strings in the following three ways: (1) the length of the array (the
byte count) is explicitly located in an unsigned integer, (2) the byte sequence
is not terminated by a null character, and (3) the external representation of the
bytes is the same as their internal representation. The primitive xdr_bytes
converts between the internal and external representations of byte arrays:

```c
bool_t xdr_bytes(xdrs, bpp, lp, maxlength)
XDR *xdrs;
char **bpp;
unsigned int *lp;
unsigned int maxlength;
```

The usage of the first, second, and fourth parameters is identical to the first,
second and third parameters of xdr_string respectively. The length of the
byte area is obtained by dereferencing lp when serializing; *lp is set to the
byte length when deserializing.
Arrays

The XDR library package provides a primitive for handling arrays of arbitrary elements. The xdr_bytes routine treats a subset of generic arrays, in which the size of array elements is known to be 1, and the external description of each element is built-in. The generic array primitive, xdr_array requires parameters identical to those of xdr_bytes plus two more: the size of array elements, and an XDR routine to handle each of the elements. This routine is called to encode or decode each element of the array.

```c
bool_t xdr_array(xdrs, ap, lp, maxlength, elementsize, xdr_element)
XDR *xdrs;
char **ap;
uint *lp;
uint maxlength;
uint elementsize;
bool_t (*xdr_element)();
```

The parameter `ap` is the address of the pointer to the array. If `*ap` is NULL when the array is being deserialized, XDR allocates an array of the appropriate size and sets `*ap` to that array. The element count of the array is obtained from `*lp` when the array is serialized; `*lp` is set to the array length when the array is deserialized. The parameter `maxlength` is the maximum number of elements that the array is allowed to have; `elementsiz` is the byte size of each element of the array (the C function `sizeof` can be used to obtain this value). The `xdr_element` routine is called to serialize, deserialize, or free each element of the array.

Before defining more constructed data types, it is appropriate to present three examples.

Array Example 1

A user on a networked machine can be identified by (a) the machine name, such as krypton; (b) the user’s UID; see the geteuid man page; and (c) the group numbers to which the user belongs; see the getgroups man page. A structure with this information and its associated XDR routine could be coded as in Code Example A–0.

```c
struct netuser {
    char *nu_machinename;
    int nu_uid;
    uint nu_glen;
    int *nu_gids;
};
#define NLEN 255 /* machine names < 256 chars */
#define NGRPS 20 /* user can’t be in > 20 groups */
```

```c
bool_t
xdr_array(xdrs, ap, lp, maxlength, elementsize, xdr_element)  
    XDR *xdrs;
    char **ap;
    u_int *lp;
    u_int maxlength;
    u_int elementsize;
    bool_t (*xdr_element)();

    The parameter ap is the address of the pointer to the array. If *ap is NULL when the array is being deserialized, XDR allocates an array of the appropriate size and sets *ap to that array. The element count of the array is obtained from *lp when the array is serialized; *lp is set to the array length when the array is deserialized. The parameter maxlength is the maximum number of elements that the array is allowed to have; elementsiz is the byte size of each element of the array (the C function sizeof can be used to obtain this value). The xdr_element routine is called to serialize, deserialize, or free each element of the array.

Before defining more constructed data types, it is appropriate to present three examples.

Array Example 1

A user on a networked machine can be identified by (a) the machine name, such as krypton; (b) the user’s UID; see the geteuid man page; and (c) the group numbers to which the user belongs; see the getgroups man page. A structure with this information and its associated XDR routine could be coded as in Code Example A–0.

CODE EXAMPLE A–7  Array Example #1

```c
struct netuser {
    char *nu_machinename;
    int nu_uid;
    u_int nu_glen;
    int *nu_gids;
};
#define NLEN 255 /* machine names < 256 chars */
#define NGRPS 20 /* user can’t be in > 20 groups */
```

```c
bool_t
xdr_array(xdrs, ap, lp, maxlength, elementsize, xdr_element)  
    XDR *xdrs;
    char **ap;
    u_int *lp;
    u_int maxlength;
    u_int elementsize;
    bool_t (*xdr_element)();

    The parameter ap is the address of the pointer to the array. If *ap is NULL when the array is being deserialized, XDR allocates an array of the appropriate size and sets *ap to that array. The element count of the array is obtained from *lp when the array is serialized; *lp is set to the array length when the array is deserialized. The parameter maxlength is the maximum number of elements that the array is allowed to have; elementsiz is the byte size of each element of the array (the C function sizeof can be used to obtain this value). The xdr_element routine is called to serialize, deserialize, or free each element of the array.

Before defining more constructed data types, it is appropriate to present three examples.

Array Example 1

A user on a networked machine can be identified by (a) the machine name, such as krypton; (b) the user’s UID; see the geteuid man page; and (c) the group numbers to which the user belongs; see the getgroups man page. A structure with this information and its associated XDR routine could be coded as in Code Example A–0.

```c
struct netuser {
    char *nu_machinename;
    int nu_uid;
    u_int nu_glen;
    int *nu_gids;
};
#define NLEN 255 /* machine names < 256 chars */
#define NGRPS 20 /* user can’t be in > 20 groups */
```
Array Example 2

A party of network users could be implemented as an array of netuser structure. The declaration and its associated XDR routines are as shown in Code Example A–0.

**CODE EXAMPLE A–8**  Array Example #2

```c
struct party {
    u_int p_len;
    struct netuser *p_nusers;
};
#define PLEN 500 /* max number of users in a party */
bool_t xdr_party(xdrs, pp)
    XDR *xdrs;
    struct party *pp;
{ return(xdr_array(xdrs, &pp->p_nusers, &pp->p_len, PLEN,
        sizeof (struct netuser), xdr_netuser));
}
```

Array Example 3

The well-known parameters to main, argc and argv can be combined into a structure. An array of these structures can make up a history of commands. The declarations and XDR routines might look like Code Example A–0.

**CODE EXAMPLE A–9**  Array Example #3

```c
struct cmd {
    u_int c_argc;
    char **c_argv;
};
#define ALEN 1000 /* args cannot be > 1000 chars */
#define NARGC 100 /* commands cannot have > 100 args */
struct history {
    u_int h_len;
    struct cmd *h_cmds;
};
#define NCMDS 75 /* history is no more than 75 commands */
bool_t xdr_wrapstring(xdrs, sp)
```
XDR *xdrs;
char **sp;
{
    return(xdr_string(xdrs, sp, ALEN));
}

bool_t
xdr_cmd(xdrs, cp)
    XDR *xdrs;
    struct cmd *cp;
{
    return(xdr_array(xdrs, &cp->c_argv, &cp->c_argc, NARGC,
        sizeof (char *), xdr_wrapstring));
}

bool_t
xdr_history(xdrs, hp)
    XDR *xdrs;
    struct history *hp;
{
    return(xdr_array(xdrs, &hp->h_cmds, &hp->h_len, NCMDS,
        sizeof (struct cmd), xdr_cmd));
}

The most confusing part of this example is that the routine xdr_wrapstring is needed to package the xdr_string routine, because the implementation of xdr_array passes only two parameters to the array element description routine; xdr_wrapstring supplies the third parameter to xdr_string.

By now the recursive nature of the XDR library should be obvious. Let’s continue with more constructed data types.

Opaque Data

In some protocols, handles are passed from a server to client. The client passes the handle back to the server at some later time. Handles are never inspected by clients; they are obtained and submitted. That is to say, handles are opaque. The xdr_opaque primitive is used for describing fixed sized, opaque bytes.

bool_t
xdr_opaque(xdrs, p, len)
    XDR *xdrs;
    char *p;
    u_int len;

The parameter p is the location of the bytes; len is the number of bytes in the opaque object. By definition, the actual data contained in the opaque object are not machine portable.
In SunOS/SVR4 there is another routine for manipulating opaque data. This routine, xdr_netobj sends counted opaque data, much like xdr_opaque. Code Example A-0 illustrates the syntax of xdr_netobj.

**CODE EXAMPLE A-10  xdr_netobj Routine**

```c
struct netobj {
    u_int n_len;
    char *n_bytes;
};
typedef struct netobj netobj;

bool_t xdr_netobj(xdrs, np)
    XDR *xdrs;
    struct netobj *np;
    {
        return(TRUE);
    }
```

The xdr_netobj routine is a filter primitive that translates between variable length opaque data and its external representation. The parameter np is the address of the netobj structure containing both a length and a pointer to the opaque data. The length may be no more than MAX_NETOBJ_SZ bytes. This routine returns TRUE if it succeeds, FALSE otherwise.

**Fixed-Length Arrays**

The XDR library provides a primitive, xdr_vector, for fixed-length arrays, shown in Code Example A-0.

**CODE EXAMPLE A-11  xdr_vector Routine**

```c
#define NLEN 255 /* machine names must be < 256 chars */
#define NGRPS 20 /* user belongs to exactly 20 groups */

struct netuser {
    char *nu_machinename;
    int nu_uid;
    int nu_gids[NGRPS];
};

bool_t xdr_netuser(xdrs, nup)
    XDR *xdrs;
    struct netuser *nup;
    {
        int i;
        if (!xdr_string(xdrs, &nup->nu_machinename, NLEN))
            return(FALSE);
        if (!xdr_int(xdrs, &nup->nu_uid))
            return(FALSE);
        if (!xdr_vector(xdrs, nup->nu_gids, NGRPS, sizeof(int),
            xdr_int))
            return(FALSE);
        return(TRUE);
    }
```
Discriminated Unions

The XDR library supports discriminated unions. A discriminated union is a C union and an \texttt{enum\_t} value that selects an “arm” of the union.

\begin{verbatim}
struct xdr_discrim {
    enum\_t value;
    bool\_t (*proc)();
};

bool\_t xdr_union(xdrs, dscmp, unp, arms, defaultarm)
    XDR *xdrs;
    enum\_t *dscmp;
    char *unp;
    struct xdr_discrim *arms;
    bool\_t (*defaultarm)(); /* may equal NULL */

    First the routine translates the discriminant of the union located at \*dscmp. The discriminant is always an \texttt{enum\_t}. Next the union located at \*unp is translated. The parameter \texttt{arms} is a pointer to an array of \texttt{xdr\_discrim} structures. Each structure contains an ordered pair of \texttt{[value,proc]}. If the union’s discriminant is equal to the associated \texttt{value}, then the \texttt{proc} is called to translate the union. The end of the \texttt{xdr\_discrim} structure array is denoted by a routine of value \texttt{NULL} (0). If the discriminant is not found in the \texttt{arms} array, then the \texttt{defaultarm} procedure is called if it is nonnull; otherwise the routine returns \texttt{FALSE}.
\end{verbatim}

Discriminated Union Example

Suppose the type of a union may be integer, character pointer (a string), or a \texttt{gnumbers} structure. Also, assume the union and its current type are declared in a structure. The declaration is:

\begin{verbatim}
enum utype {INTEGER=1, STRING=2, GNUMBERS=3};
struct u\_tag {
    enum utype utype; /* the union’s discriminant */
    union {
        int ival;
        char *pval;
        struct gnumbers gn;
    } uval;
};
\end{verbatim}

Code Example A–0 constructs and XDR procedure (de)serialize the discriminated union.

\textbf{CODE EXAMPLE A–12} XDR Discriminated Union
struct xdr_discrim u_tag_arms[4] = {  
  {INTEGER, xdr_int},  
  {GNUMBERS, xdr_gnumbers}  
  {STRING, xdr_wrapstring},  
  {__dontcare__, NULL}  
  /* always terminate arms with a NULL xdr_proc */  
}

bool_t  
xdr_u_tag(xdrs, utp)  
XDR *xdrs;  
struct u_tag *utp;  
{  
  return(xdr_union(xdrs, &utp->utype, &utp->uval,  
                   u_tag_arms, NULL));  
}

The routine xdr_gnumbers was presented above in the XDR Library section. xdr_wrapstring was presented in example C. The default arm parameter to xdr_union (the last parameter) is NULL in this example. Therefore the value of the union’s discriminant may legally take on only values listed in the u_tag_arms array. This example also demonstrates that the elements of the arm’s array do not need to be sorted.

It is worth pointing out that the values of the discriminant may be sparse, though in this example they are not. It is always good practice to assign explicitly integer values to each element of the discriminant’s type. This practice both documents the external representation of the discriminant and guarantees that different C compilers emit identical discriminant values.

Exercise
Implement xdr_union using the other primitives in this section.

Pointers
In C it is often convenient to put pointers to another structure within a structure. The xdr_reference primitive makes it easy to serialize, deserialize, and free these referenced structures.

bool_t  
xdr_reference(xdrs, pp, size, proc)  
XDR *xdrs;  
char **pp;  
u_int ssize;  
bool_t (*proc)();  

Parameter pp is the address of the pointer to the structure; parameter ssize is the size in bytes of the structure (use the C function sizeof to obtain this.
value); and proc is the XDR routine that describes the structure. When decoding data, storage is allocated if *pp is NULL.

There is no need for a primitive xdr_struct to describe structures within structures, because pointers are always sufficient.

Exercise

Implement xdr_reference using xdr_array.

Warning  •  xdr_reference and xdr_array are NOT interchangeable external representations of data.

Pointer Example

Suppose there is a structure containing a person’s name and a pointer to a gnumbers structure containing the person’s gross assets and liabilities. The construct is:

```c
struct pgn {
    char *name;
    struct gnumbers *gnp;
};
```

The corresponding XDR routine for this structure is:

```c
bool_t xdr_pgn(xdrs, pp)
    XDR *xdrs;
    struct pgn *pp;
{
    return(xdr_string(xdrs, &pp->name, NLEN) &&
        xdr_reference(xdrs, &pp->gnp, sizeof(struct gnumbers),
        xdr_gnumbers));
}
```

Pointer Semantics

In many applications, C programmers attach double meaning to the values of a pointer. Typically the value NULL (or zero) means data is not needed, yet some application-specific interpretation applies. In essence, the C programmer is encoding a discriminated union efficiently by overloading the interpretation of the value of a pointer. For instance, in example E a NULL pointer value for gnp could indicate that the person’s assets and liabilities are unknown. That is, the pointer value encodes two things: whether or not the data is known; and if it is known, where it is located in memory. Linked lists are an extreme example of the use of application-specific pointer interpretation.
The primitive `xdr_reference` cannot and does not attach any special meaning to a null-value pointer during serialization. That is, passing an address of a pointer whose value is NULL to `xdr_reference` when serializing data will most likely cause a memory fault and, on the UNIX system, a core dump.

`xdr_pointer` correctly handles NULL pointers.

**Nonfilter Primitives**

XDR streams can be manipulated with the primitives discussed in this section.

```c
u_int xdr_getpos(xdrs)  
   XDR *xdrs;

bool_t xdr_setpos(xdrs, pos)  
   XDR *xdrs;  
   u_int pos;

xdr_destroy(xdrs)  
   XDR *xdrs;
```

The routine `xdr_getpos` returns an unsigned integer that describes the current position in the data stream. Warning: In some XDR streams, the value returned by `xdr_getpos` is meaningless; the routine returns a -1 in this case (though -1 should be a legitimate value).

The routine `xdr_setpos` sets a stream position to `pos`. Warning: In some XDR streams, setting a position is impossible; in such cases, `xdr_setpos` will return `FALSE`. This routine will also fail if the requested position is out-of-bounds. The definition of bounds varies from stream to stream.

The `xdr_destroy` primitive destroys the XDR stream. Usage of the stream after calling this routine is undefined.

**Operation Directions**

At times you may want to optimize XDR routines by taking advantage of the direction of the operation—`XDR_ENCODE`, `XDR_DECODE` or `XDR_FREE`. The value `xdrs->x_op` always contains the direction of the XDR operation. An example in “Linked Lists” on page 188 demonstrates the usefulness of the `xdrs->x_op` field.
Stream Access

An XDR stream is obtained by calling the appropriate creation routine. These creation routines take arguments that are tailored to the specific properties of the stream. Streams currently exist for (de)serialization of data to or from standard I/O FILE streams, record streams, and UNIX files, and memory.

Standard I/O Streams

XDR streams can be interfaced to standard I/O using the xdrstdio_create() routine:

```c
#include <stdio.h>
#include <rpc/rpc.h> /* xdr is part of rpc */

void
xdrstdio_create(xdrs, fp, x_op)
    XDR *xdrs;
    FILE *fp;
    enum xdr_op x_op;

The routine xdrstdio_create initializes an XDR stream pointed to by xdrs. The XDR stream interfaces to the standard I/O library. Parameter fp is an open file, and x_op is an XDR direction.

Memory Streams

Memory streams allow the streaming of data into or out of a specified area of memory:

```c
#include <rpc/rpc.h>

void
xdrmem_create(xdrs, addr, len, x_op)
    XDR *xdrs;
    char *addr;
    u_int len;
    enum xdr_op x_op;

The routine xdrmem_create initializes an XDR stream in local memory. The memory is pointed to by parameter addr; parameter len is the length in bytes of the memory. The parameters xdrs and x_op are identical to the corresponding parameters of xdrstdio_create. Currently, the datagram implementation of RPC uses xdrmem_create. Complete call or result messages are built in memory before calling the t_sndndata TLI routine.
Record (TCP/IP) Streams

A record stream is an XDR stream built on top of a record marking standard that is built on top of the UNIX file or 4.2 BSD connection interface.

```c
#include <rpc/rpc.h> /* xdr is part of rpc */
xdrrec_create(xdrs, sendsize, recvszie, iohandle, readproc, writeproc)
    XDR *xdrs;
    u_int sendsize, recvszie;
    char *iohandle;
    int (*readproc)(), (*writeproc)();
```

The routine `xdrrec_create` provides an XDR stream interface that allows for a bidirectional, arbitrarily long sequence of records. The contents of the records are meant to be data in XDR form. The stream’s primary use is for interfacing RPC to TCP connections. However, it can be used to stream data into or out of normal UNIX files.

The parameter `xdrs` is similar to the corresponding parameter described above. The stream does its own data buffering similar to that of standard I/O. The parameters `sendsize` and `recvszie` determine the size in bytes of the output and input buffers, respectively; if their values are zero (0), then predetermined defaults are used. When a buffer needs to be filled or flushed, the routine `readproc` or `writeproc` is called, respectively. The usage and behavior of these routines are similar to the UNIX system calls `read` and `write`. However, the first parameter to each of these routines is the opaque parameter `iohandle`. The other two parameters (and `nbytes`) and the results (byte count) are identical to the system routines. If `xxx` is `readproc` or `writeproc`, then it has the following form:

```c
/* returns the actual number of bytes transferred. -1 is an error */
int xxx(iohandle, buf, len)
    char *iohandle;
    char *buf;
    int nbytes;
```

The XDR stream provides means for delimiting records in the byte stream. Abstract data types needed to implement the XDR stream mechanism are discussed in "XDR Stream Implementation " on page 186. The protocol RPC uses to delimit XDR stream records is discussed in "Record-Marking Standard" on page 202.

The primitives that are specific to record streams are as follows:

```c
bool_t
xdrrec_endofrecord(xdrs, flushnow)
    XDR *xdrs;
    bool_t flushnow;

bool_t
xdrrec_skiprecord(xdrs)
```
The routine xdrrec_endofrecord causes the current outgoing data to be marked as a record. If the parameter flushnow is TRUE, then the stream's writeproc will be called; otherwise, writeproc will be called when the output buffer has been filled.

The routine xdrrec_skiprecord causes an input stream's position to be moved past the current record boundary and onto the beginning of the next record in the stream.

If there is no more data in the stream's input buffer, then the routine xdrrec_eof returns TRUE. That is not to say that there is no more data in the underlying file descriptor.

XDR Stream Implementation

This section provides the abstract data types needed to implement new instances of XDR streams.

The XDR Object

The structure in Code Example A-0 defines the interface to an XDR stream.

CODE EXAMPLE A-13  XDR Stream Interface Example

```c
enum xdr_op {XDR_ENCODE=0, XDR_DECODE=1, XDR_FREE=2};

typedef struct {
  enum xdr_op x_op;
  struct xdr_ops {
    bool_t (*x_getlong)(); /* get long from stream */
    bool_t (*x_putlong)(); /* put long to stream */
    bool_t (*x_getbytes)(); /* get bytes from stream */
    bool_t (*x_putbytes)(); /* put bytes to stream */
    u_int (*x_getpostn)(); /* return stream offset */
    bool_t (*x_setpostn)(); /* reposition offset */
    caddr_t (*x_inline)(); /* ptr to buffered data */
    VOID (*x_destroy)(); /* free private area */
  } *x_ops;
} *xdsrs;

bool_t
xdrrec_eof(xdsrs)
  XDR *xdsrs;

(Continued)
caddr_t x_public; /* users’ data */
caddr_t x_private; /* pointer to private data */
caddr_t x_base; /* private for position info */
int x_handy; /* extra private word */
} XDR;

The `x_op` field is the current operation being performed on the stream. This field is important to the XDR primitives, but should not affect a stream’s implementation. That is, a stream’s implementation should not depend on this value. The fields `x_private, x_base,` and `x_handy` are private to the particular stream’s implementation. The field `x_public` is for the XDR client and should never be used by the XDR stream implementations or the XDR primitives. `x_getposn, x_setposn,` and `x_destroy` are macros for accessing operations. The operation `x_inline` has two parameters: an `XDR *, and an unsigned integer, which is a byte count. The routine returns a pointer to a piece of the stream’s internal buffer. The caller can then use the buffer segment for any purpose. From the stream’s point of view, the bytes in the buffer segment have been consumed. The routine may return NULL if it cannot return a buffer segment of the requested size. (The `x_inline` routine is used to squeeze cycles, and the resulting buffer is not data portable. Users are cautioned against using this feature.)

The operations `x_getbytes and x_putbytes` blindly get and put sequences of bytes from or to the underlying stream; they return TRUE if they are successful, and FALSE otherwise. The routines have identical parameters (replace `xxx`):

```c
bool_t
xxxbytes(xdrs, buf, bytecount)
XDR *xdrs;
char *buf;
_u_int bytecount;
```

The operations `x_getlong` and `x_putlong` receive and put long numbers from and to the data stream. It is the responsibility of these routines to translate the numbers between the machine representation and the (standard) external representation. The UNIX primitives `htonl` and `ntohl` can be helpful in accomplishing this. The higher-level XDR implementation assumes that signed and unsigned long integers contain the same number of bits, and that nonnegative integers have the same bit representations as unsigned integers. The routines return TRUE if they succeed, and FALSE otherwise.

They have identical parameters:

```c
bool_t
xxxlong(xdrs, lp)
XDR *xdrs;
long *lp;
```
Implementors of new XDR streams must make an XDR structure (with new operation routines) available to clients, using some kind of create routine.

Advanced Topics

This section describes techniques for passing data structures that are not covered in the preceding sections. Such structures include linked lists (of arbitrary lengths). Unlike the simpler examples covered in the earlier sections, the following examples are written using both the XDR C library routines and the XDR data description language. Appendix C, describes this language in detail.

Linked Lists

The “Pointer Example” on page 182 presented a C data structure and its associated XDR routines for an individual’s gross assets and liabilities. Code Example A–0 uses a linked list to duplicate the pointer example.

CODE EXAMPLE A–14  Linked List

```c
struct gnumbers {
    long g_assets;
    long g_liabilities;
};

bool_t xdr_gnumbers(xdrs, gp)
XDR *xdrs;
struct gnumbers *gp;
{
    return(xdr_long(xdrs, &(gp->g_assets) &&
        xdr_long(xdrs, &(gp->g_liabilities))));
}
```

Now assume that you want to implement a linked list of such information. A data structure could be constructed as follows:

```c
struct gnumbers_node {
    struct gnumbers gn_numbers;
    struct gnumbers_node *gn_next;
};
typedef struct gnumbers_node *gnumbers_list;
```

The head of the linked list can be thought of as the data object; that is, the head is not merely a convenient shorthand for a structure. Similarly the gn_next field is used to indicate whether the object has terminated. Unfortunately, if the object continues, the gn_next field is also the address
of where it continues. The link addresses carry no useful information when the object is serialized.

The XDR data description of this linked list is described by the recursive declaration of `gnumbers_list`:

```c
struct gnumbers {
    int g_assets;
    int g_liabilities;
};
struct gnumbers_node {
    gnumbers gn_numbers;
    gnumbers_node *gn_next;
};
```

In this description, the Boolean indicates whether there is more data following it. If the Boolean is FALSE, it is the last data field of the structure. If it is TRUE, it is followed by a `gnumbers` structure and (recursively) by a `gnumbers_list`. Note that the C declaration has no Boolean explicitly declared in it (though the `gn_next` field implicitly carries the information), while the XDR data description has no pointer explicitly declared in it.

Hints for writing the XDR routines for a `gnumbers_list` follow easily from the XDR description above. Note how the primitive `xdr_pointer` is used to implement the XDR union above.

**CODE EXAMPLE A–15  xdr_pointer**

```c
bool_t
xdr_gnumbers_node(xdrs, gn)
    XDR *xdrs;
    gnumbers_node *gn;
{
    return(xdr_gnumbers(xdrs, &gn->gn_numbers) &&
           xdr_gnumbers_list(xdrs, &gn->gn_next));
}

bool_t
xdr_gnumbers_list(xdrs, gnp)
    XDR *xdrs;
    gnumbers_list *gnp;
{
    return(xdr_pointer(xdrs, gnp, sizeof(struct gnumbers_node),
                       xdr_gnumbers_node));
    xdr_pointer
}
```

The unfortunate side effect of using XDR on a list with these routines is that the C stack grows linearly with respect to the number of nodes in the list. This is due to the recursion. Code Example A–0 collapses the above two mutually recursive routines into a single, nonrecursive one.

**CODE EXAMPLE A–16  Nonrecursive Stack in XDR**
bool_t
wdr_gnumbers_list(xdrs, gnp)
XDR *xdrs;
gnumbers_list *gnp;
{
    bool_t more_data;
    gnumbers_list *nextp;

    for(;;) {
        more_data = (*gnp != NULL);
        if (!xdr_bool(xdrs, &more_data))
            return(FALSE);
        if (! more_data)
            break;
        if (xdrs->x_op == XDR_FREE)
            nextp = &(*gnp)->gn_next;
        if (!xdr_reference(xdrs, gnp,
            sizeof(struct gnumbers_node), xdr_gnumbers))
            return(FALSE);
        gnp = (xdrs->x_op == XDR_FREE) ? nextp : &(*gnp)->gn_next;
    }'*gnp = NULL;
    return(TRUE);
}

The first task is to find out whether there is more data, so that this Boolean
information can be serialized. Notice that this statement is unnecessary in the
XDR_DECODE case, since the value of more_data is not known until you
deserialize it in the next statement.

The next statement implements XDR on the more_data field of the XDR
union. Then if there is truly no more data, you set this last pointer to NULL
to indicate the end of the list, and return TRUE because you are done. Note
that setting the pointer to NULL is only important in the XDR_DECODE case,
since it is already NULL in the XDR_ENCODE and XDR_FREE cases.

Next, if the direction is XDR_FREE, the value of nextp is set to indicate the
location of the next pointer in the list. We do this now because you need to
dereference gnp to find the location of the next item in the list, and after the
next statement, the storage pointed to by gnp will be freed up and no be
longer valid. We can’t do this for all directions though, because in the
XDR_DECODE direction the value of gnp won’t be set until the next statement.

Next, you XDR the data in the node using the primitive xdr_reference.
xdr_reference is like xdr_pointer which you used before, but it does
not send over the Boolean indicating whether there is more data. We use it
instead of xdr_pointer because you have already used XDR on this
information yourself. Notice that the XDR routine passed is not the same
type as an element in the list. The routine passed is xdr_gnumbers, but
each element in the list is actually of type gnumbers_node. You don’t pass
xdr_gnumbers_node because it is recursive. Instead use xdr_gnumbers
which uses XDR on all of the nonrecursive part. Note that this trick works
only if the `gn_numbers` field is the first item in each element, so that their addresses are identical when passed to `xdr_reference`.

Finally, you update `gnp` to point to the next item in the list. If the direction is `XDR_FREE`, you set it to the previously saved value; otherwise you can dereference `gnp` to get the proper value. Though harder to understand than the recursive version, this nonrecursive routine will run more efficiently since much of the procedure call overhead has been removed. Most lists are small though (in the hundreds of items or less) and the recursive version should be sufficient for them.
RPC Protocol and Language Specification

This appendix specifies a message protocol used in implementing the RPC package. The message protocol is specified with the XDR language. The companion appendix to this one is Appendix C.

- “Protocol Overview” on page 193
- “Program and Procedure Numbers” on page 196
- “Authentication Protocols” on page 203
- “The RPC Language Specification” on page 214

Protocol Overview

The RPC protocol provides for the following:

- Unique specification of a procedure to be called.
- Provisions for matching response messages to request messages.
- Provisions for authenticating the caller to service and vice-versa. In addition, the RPC package provides features that detect the following:
  - RPC protocol mismatches
  - Remote program protocol version mismatches
  - Protocol errors (such as incorrect specification of a procedure’s parameters)
  - Reasons why remote authentication failed

Consider a network file service composed of two programs. One program may deal with high-level applications such as file system access control and
The other may deal with low-level file I/O and have procedures like "read" and "write." A client machine of the network file service would call the procedures associated with the two programs of the service on behalf of some user on the client machine. In the client-server model a remote procedure call is used to call the service.

The RPC Model

The RPC model is similar to the local procedure call model. In the local case, the caller places arguments to a procedure in some well-specified location. The caller then transfers control to the procedure, and eventually regains control. At that point, the results of the procedure are extracted from a well-specified location, and the caller continues execution.

The RPC is similar, in that one thread of control logically winds through two processes. One is the caller’s process, the other is a server’s process. Conceptually, the caller process sends a call message to the server process and waits (blocks) for a reply message. The call message contains the procedure’s parameters, among other things. The reply message contains the procedure’s results, among other things. Once the reply message is received, the results of the procedure are extracted, and the caller’s execution is resumed.

On the server side, a process is dormant awaiting the arrival of a call message. When one arrives, the server process extracts the procedure’s parameters, computes the results, sends a reply message, and then awaits the next call message.

Note that in this description only one of the two processes is active at any given time. However, this need not be the case. The RPC protocol makes no restrictions on the concurrency model implemented. For example, an implementation may choose to have RPC calls be asynchronous, so that the client may do useful work while waiting for the reply from the server. Another possibility is to have the server create a task to process an incoming request, so that the server can be free to receive other requests.

Transports and Semantics

The RPC protocol is independent of transport protocols. That is, RPC disregards how a message is passed from one process to another. The protocol deals only with specification and interpretation of messages.

RPC does not attempt to ensure transport reliability. Therefore, you must supply the application with information about the type of transport protocol used under RPC. If you tell the RPC service it is running on top of a reliable transport such as TCP, most of the work is already done for it. On the other hand, if RPC is running on top of an unreliable transport such as UDP, the
service must devise its own retransmission and time-out policy. RPC does not provide this service.

Because of transport independence, the RPC protocol does not attach specific semantics to the remote procedures or their execution. Semantics can be inferred from (but should be explicitly specified by) the underlying transport protocol. For example, consider RPC running on top of an unreliable transport. If an application retransmits RPC messages after short time-outs, the only thing it can infer if it receives no reply is that the procedure was executed zero or more times. If it does receive a reply, it can infer that the procedure was executed at least once.

A server may choose to remember previously granted requests from a client and not regrant them to insure some degree of execute-at-most-once semantics. A server can do this by taking advantage of the transaction ID that is packaged with every RPC request. The main use of this transaction ID is by the RPC client for matching replies to requests. However, a client application may choose to reuse its previous transaction ID when retransmitting a request. The server application, checking this fact, may choose to remember this ID after granting a request and not regrant requests with the same ID. The server is not allowed to examine this ID in any other way except as a test for equality.

On the other hand, if using a reliable transport such as TCP, the application can infer from a reply message that the procedure was executed exactly once, but if it receives no reply message, it cannot assume the remote procedure was not executed. Note that even if a connection-oriented protocol like TCP is used, an application still needs time-outs and reconnection to handle server crashes.

### Binding and Rendezvous Independence

The act of binding a client to a service is not part of the remote procedure call specification. This important and necessary function is left up to some higher-level software. (The software may use RPC itself; see “rpcbind Protocol” on page 223.)

Implementers should think of the RPC protocol as the jump-subroutine instruction (JSR) of a network; the loader (binder) makes JSR useful, and the loader itself uses JSR to accomplish its task. Likewise, the network makes RPC useful, enabling RPC to accomplish this task.

The RPC protocol provides the fields necessary for a client to identify itself to a service and vice-versa. Security and access control mechanisms can be built on top of the message authentication. Several different authentication protocols can be supported. A field in the RPC header specifies the protocol being used. More information on authentication protocols can be found in the section “Record-Marking Standard” on page 202.
Program and Procedure Numbers

The RPC call message has three unsigned fields that uniquely identify the procedure to be called:

- remote program number
- remote program version number
- remote procedure number

Program numbers are administered by a central authority, as described in “Program Number Registration” on page 198.

The first implementation of a program will most likely have version number 1. Because most new protocols evolve into better, stable, and mature protocols, a version field of the call message identifies the version of the protocol the caller is using. Version numbers make “speaking” old and new protocols through the same server process possible.

The procedure number identifies the procedure to be called. These numbers are documented in the individual program’s protocol specification. For example, a file service’s protocol specification may state that its procedure number 5 is “read” and procedure number 12 is “write.”

Just as remote program protocols may change over several versions, the RPC message protocol itself may change. Therefore, the call message also has in it the RPC version number, which is always equal to 2 for the version of RPC described here.

The reply message to a request message has enough information to distinguish the following error conditions:

- The remote implementation of RPC does not “speak” protocol version 2. The lowest and highest supported RPC version numbers are returned.
- The remote program is not available on the remote system.
- The remote program does not support the requested version number. The lowest and highest supported remote program version numbers are returned.
- The requested procedure number does not exist. (This is usually a caller-side protocol or programming error.)
- The parameters to the remote procedure appear to be garbage from the server’s point of view. (Again, this is usually caused by a disagreement about the protocol between client and service.)

Provisions for authentication of caller to service and vice versa are provided as a part of the RPC protocol. The call message has two authentication fields, the credentials and verifier. The reply message has one authentication field,
the response verifier. The RPC protocol specification defines all three fields to be the following opaque type:

```c
enum auth_flavor {
    AUTH_NONE = 0,
    AUTH_SYS = 1,
    AUTH_SHORT = 2,
    AUTH_DES = 3,
    AUTH_KERB = 4
    /* and more to be defined */
};
struct opaque_auth {
    enum auth_flavor; /* style of credentials */
    caddr_t oa_base; /* address of more auth stuff */
    u_int   oa_length; /* not to exceed MAX_AUTH_BYTES */
};
```

An opaque_auth structure is an auth_flavor enumeration followed by bytes that are opaque to the RPC protocol implementation.

The interpretation and semantics of the data contained within the authentication fields are specified by individual, independent authentication protocol specifications. (See “Record-Marking Standard” on page 202 for definitions of the various authentication protocols.)

If authentication parameters are rejected, the response message contains information stating why they are rejected.

Program Number Assignment

Program numbers are distributed in groups of 0x20000000, as shown in Table B-1.

<table>
<thead>
<tr>
<th>Program Numbers</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>00000000 - 1fffffff</td>
<td>Defined by Sun</td>
</tr>
<tr>
<td>20000000 - 3fffffff</td>
<td>Defined by user</td>
</tr>
<tr>
<td>40000000 - 5fffffff</td>
<td>Transient (Reserved for customer-written applications)</td>
</tr>
<tr>
<td>60000000 - 7fffffff</td>
<td>Reserved</td>
</tr>
</tbody>
</table>
TABLE B–1  RPC Program Number Assignment  (continued)

<table>
<thead>
<tr>
<th>Program Numbers</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>80000000 - 9fffffff</td>
<td>Reserved</td>
</tr>
<tr>
<td>a0000000 - bfffffff</td>
<td>Reserved</td>
</tr>
<tr>
<td>c0000000 - dfffffff</td>
<td>Reserved</td>
</tr>
<tr>
<td>e0000000 - ffffffff</td>
<td>Reserved</td>
</tr>
</tbody>
</table>

Sun Microsystems administers the first group of numbers, which should be identical for all customers. If a customer develops an application that might be of general interest, that application should be given an assigned number in the first range.

The second group of numbers is reserved for specific customer applications. This range is intended primarily for debugging new programs.

The third group is reserved for applications that generate program numbers dynamically.

The final groups are reserved for future use, and should not be used.

Program Number Registration

To register a protocol specification, send a request by email to rpc@sun.com, or write to: RPC Administrator Sun Microsystems 2550 Garcia Avenue Mountain View, CA 94043

Please include a compilable rpcgen .x file describing your protocol. You will be given a unique program number in return.

The RPC program numbers and protocol specifications of standard RPC services can be found in the include files in /usr/include/rpcsrv. These services, however, constitute only a small subset of those that have been registered.
Other Uses of the RPC Protocol

The intended use of this protocol is for calling remote procedures. That is, each call message is matched with a response message. However, the protocol itself is a message-passing protocol with which other (non-RPC) protocols can be implemented. Some of the non-RPC protocols supported by the RPC package are batching and broadcasting.

Batching

Batching allows a client to send an arbitrarily large sequence of call messages to a server; batching typically uses reliable byte stream protocols (like TCP) for its transport. In batching, the client never waits for a reply from the server, and the server does not send replies to batch requests. A sequence of batch calls is usually finished by a non-batch RPC call to flush the pipeline (with positive acknowledgment). For more information, see “Batching” on page 88.

Broadcast RPC

In broadcast RPC, the client sends a broadcast packet to the network and waits for numerous replies. Broadcast RPC uses connectionless, packet-based protocols (like UDP) as its transports. Servers that support broadcast protocols only respond when the request is successfully processed, and are silent in the face of errors. Broadcast RPC uses the rpcbind service to achieve its semantics. See “Broadcast RPC” on page 86 and “rpcbind Protocol” on page 223 for further information.

The RPC Message Protocol

This section defines the RPC message protocol in the XDR data description language. The message is defined in a top-down style, as shown in Code Example B–0.

CODE EXAMPLE B–1  RPC Message Protocol

```c
enum msg_type {
    CALL = 0,
    REPLY = 1
};

/*
 * A reply to a call message can take on two forms: The message was

(Continued)```
* either accepted or rejected.
*/
enum reply_stat {
    MSG_ACCEPTED = 0,
    MSG_DENIED = 1
};

/*
 * Given that a call message was accepted, the following is the
 * status of an attempt to call a remote procedure.
 */
enum accept_stat {
    SUCCESS = 0, /* RPC executed successfully */
    PROG_UNAVAIL = 1, /* remote service hasn’t exported prog */
    PROC_MISMATCH = 2, /* remote service can’t support versn # */
    PROC_UNAVAIL = 3, /* program can’t support proc */
    GARBAGE_ARGS = 4 /* procedure can’t decode params */
};

/*
 * Reasons a call message was rejected:
 */
enum reject_stat {
    RPC_MISMATCH = 0, /* RPC version number != 2 */
    AUTH_ERROR = 1 /* remote can’t authenticate caller */
};

/*
 * Why authentication failed:
 */
enum auth_stat {
    AUTH_BADCRED = 1, /* bad credentials */
    AUTH_REJECTEDCRED = 2, /* clnt must do new session */
    AUTH_BADVERF = 3, /* bad verifier */
    AUTH_REJECTEDVERF = 4, /* verif expired or replayed */
    AUTH_TOOWEAK = 5 /* rejected for security */
};

/*
 * The RPC message:
 * All messages start with a transaction identifier, xid, followed
 * by a two-armed discriminated union. The union’s discriminant is
 * a msg_type which switches to one of the two types of the
 * message.
 * The xid of a REPLY message always matches that of the
 * initiating CALL message. NB: The xid field is only used for
 * clients matching reply messages with call messages or for servers
 * detecting retransmissions; the service side cannot treat this id as
 * any type of sequence number.
 */
struct rpc_msg {
    unsigned int xid;
    union switch (msg_type mtype) {
        case CALL:
            /*...
            */
    }
call_body cbody;
case REPLY:
    reply_body rbody;
} body;
}

/*
* Body of an RPC request call:
* In version 2 of the RPC protocol specification, rpcvers must be
* equal to 2. The fields prog, vers, and proc specify the remote
* program, its version number, and the procedure within the
* remote program to be called. After these fields are two
* authentication parameters: cred (authentication credentials) and
* verf (authentication verifier). The two authentication parameters
* are followed by the parameters to the remote procedure, which are
* specified by the specific program protocol.
*/
struct call_body {
    unsigned int rpcvers; /* must be equal to two (2) */
    unsigned int prog;
    unsigned int vers;
    unsigned int proc;
    opaque_auth cred;
    opaque_auth verf;
    /* procedure specific parameters start here */
};

/*
* Body of a reply to an RPC request:
* The call message was either accepted or rejected.
*/
union reply_body switch (reply_stat stat) {
    case MSG_ACCEPTED:
        accepted_reply areply;
    case MSG_DENIED:
        rejected_reply rreply;
} reply;

/*
* Reply to an RPC request that was accepted by the server: there
* could be an error even though the request was accepted. The
* first field is an authentication verifier that the server
* generates in order to validate itself to the caller. It is
* followed by a union whose discriminant is an enum accept_stat.
* The SUCCESS arm of the union is protocol specific.
* The PROG_UNAVAIL, PROC_UNAVAIL, and GARBAGE_ARGP arms of
* the union are void. The PROG_MISMATCH arm specifies the lowest
* and highest version numbers of the remote program supported by
* the server.
*/
struct accepted_reply {
    opaque_auth verf;
    union switch (accept_stat stat) {
        /* (Continued) */
case SUCCESS:
   opaque results[0];
   /* procedure-specific results start here */
case PROG_MISMATCH:
   struct {
      unsigned int low;
      unsigned int high;
   } mismatch_info;
default:
   /*
   * Void. Cases include PROG_UNAVAIL, PROC_UNAVAIL, and
   * GARBAGE_ARGS.
   */
   void;
   } reply_data;
};

/*
 * Reply to an RPC request that was rejected by the server:
 * The request can be rejected for two reasons: either the server
 * is not running a compatible version of the RPC protocol
 * (RPC_MISMATCH), or the server refuses to authenticate the
 * caller AUTH_ERROR). In case of an RPC version mismatch,
 * the server returns the lowest and highest supported RPC
 * version numbers. In case of refused authentication, failure
 * status is returned.
 */
union rejected_reply switch (reject_stat stat) {
   case RPC_MISMATCH:
      struct {
         unsigned int low;
         unsigned int high;
      } mismatch_info;
      case AUTH_ERROR:
      auth_stat stat;
};

Record-Marking Standard

When RPC messages are passed on top of a byte stream transport (like TCP),
it is necessary, or at least desirable, to delimit one message from another to
detect and possibly recover from user protocol errors. This is called record
marking (RM). One RPC message fits into one RM record.

A record is composed of one or more record fragments. A record fragment is
a four-byte header followed by 0 to \((2^{31}) - 1\) bytes of fragment data.
The bytes encode an unsigned binary number; as with XDR integers, the byte
order is the network byte order.

The header encodes two values:
- A Boolean that specifies whether the fragment is the last fragment of the record (bit value 1 implies the fragment is the last fragment).

- A 31-bit unsigned binary value that is the length in bytes of the fragment’s data. The Boolean value is the highest-order bit of the header; the length is the 31 low-order bits. (This record specification is not in XDR standard form).

## Authentication Protocols

Authentication parameters are opaque but open-ended to the rest of the RPC protocol. This section defines some flavors of authentication that have already been implemented. Other sites are free to invent new authentication types, with the same rules of flavor number assignment for program number assignment. SunSoft maintains and administers a range of authentication flavors. To have authentication numbers (like RPC program numbers) allocated (or registered to them), contact the Sun RPC Administrator, as described in “Program Number Registration” on page 198.

**AUTH_NONE**

Calls are often made where the caller does not authenticate itself and the server disregards who the caller is. In these cases, the flavor value (the “discriminant” of the opaque_auth “union”) of the RPC message’s credentials, verifier, and response verifier is AUTH_NONE. The body length is zero when AUTH_NONE authentication flavor is used.

**AUTH_SYS**

This is the same as the authentication flavor previously known as AUTH_UNIX. The caller of a remote procedure may wish to identify itself using traditional UNIX process permissions authentication. The flavor of the opaque_auth of such an RPC call message is AUTH_SYS. The bytes of the body encode the following structure:

```c
struct auth_sysparms {
    unsigned int stamp;
    string machinename<255>;
    uid_t uid;
    gid_t gid;
    gid_t gids<10>;
};
```
- *stamp* is an arbitrary ID that the caller machine may generate.
- *machinename* is the name of the caller’s machine.
- *uid* is the caller’s effective user ID.
- *gid* is the caller’s effective group ID.
- *gids* is a counted array of groups in which the caller is a member.

The *flavor* of the verifier accompanying the credentials should be **AUTH_NONE**.

### The AUTH_SHORT Verifier

When using **AUTH_SYS** authentication, the *flavor* of the response verifier received in the reply message from the server may be **AUTH_NONE** or **AUTH_SHORT**.

If **AUTH_SHORT**, the bytes of the response verifier’s string encode a *short_hand_verf* structure. This opaque structure may now be passed to the server instead of the original **AUTH_SYS** credentials.

The server keeps a cache that maps shorthand opaque structures (passed back by way of an **AUTH_SHORT** style response verifier) to the original credentials of the caller. The caller can save network bandwidth and server cpu cycles by using the new credentials.

The server may flush the shorthand opaque structure at any time. If this happens, the remote procedure call message will be rejected owing to an authentication error. The reason for the failure will be **AUTH_REJECTEDCRED**.
At this point, the caller may wish to try the original AUTH_SYS style of credentials. See Figure B–1.

![Authentication Process Map](image)

**Figure B–1** Authentication Process Map

**AUTH_DES Authentication**

AUTH_SYS authentication has the following problems:

1. Caller identification cannot be guaranteed to be unique if machines with differing operating systems are on the same network.
2. There is no verifier, so credentials can easily be faked. AUTH_DES authentication attempts to fix these two problems.

The first problem is handled by addressing the caller by a simple string of characters instead of by an operating system specific integer. This string of characters is known as the netname or network name of the caller. The server should not interpret the caller’s name in any way other than as the identity...
of the caller. Thus, netnames should be unique for every caller in the naming domain.

It is up to each operating system’s implementation of AUTH_DES authentication to generate netnames for its users that ensure this uniqueness when they call remote servers. Operating systems already distinguish users local to their systems. It is usually a simple matter to extend this mechanism to the network. For example, a user with a user ID of 515 might be assigned the following netname: ‘‘UNIX.515@sun.com’’. This netname contains three items that serve to ensure it is unique. Going backward, there is only one naming domain called sun.com in the Internet. Within this domain, there is only one UNIX user with user ID 515. However, there may be another user on another operating system, for example VMS, within the same naming domain who, by coincidence, happens to have the same user ID. To ensure that these two users can be distinguished you add the operating system name. So one user is ‘‘UNIX.515@sun.com’’ and the other is ‘‘VMS.515@sun.com’’.

The first field is actually a naming method rather than an operating system name. It just happens that there is almost a one-to-one correspondence between naming methods and operating systems. If the world could agree on a naming standard, the first field could be a name from that standard, instead of an operating system name.

**AUTH_DES Authentication Verifiers**

Unlike AUTH_SYS authentication, AUTH_DES authentication does have a verifier so the server can validate the client’s credential (and vice versa). The contents of this verifier are primarily an encrypted timestamp. The server can decrypt this timestamp, and if it is close to its current real time, then the client must have encrypted it correctly. The only way the client could encrypt it correctly is to know the conversation key of the RPC session. If the client knows the conversation key, it must be the real client.

The conversation key is a DES [5] key that the client generates and notifies the server of in its first RPC call. The conversation key is encrypted using a public key scheme in this first transaction. The particular public key scheme used in AUTH_DES authentication is Diffie-Hellman [3] with 192-bit keys. The details of this encryption method are described later.

The client and the server need the same notion of the current time for this to work. If network time synchronization cannot be guaranteed, then client can synchronize with the server before beginning the conversation. rpcbind provides a procedure, RPCBPROC_GETTIME, which may be used to obtain the current time.

A server can determine if a client timestamp is valid. For any transaction after the first, the server checks for two things:
The timestamp is greater than the one previously seen from the same client.

The timestamp has not expired. A timestamp is expired if the server’s time is later than the sum of the client’s timestamp plus what is known as the client’s window. The window is a number the client passes (encrypted) to the server in its first transaction. The window can be thought of as a lifetime for the credential.

For the first transaction, the server checks that the timestamp has not expired. As an added check, the client sends an encrypted item in the first transaction known as the window verifier which must be equal to the window minus 1, or the server will reject the credential.

The client must check the verifier returned from the server to be sure it is legitimate. The server sends back to the client the encrypted timestamp it received from the client, minus one second. If the client gets anything other than this, it will reject it.

Nicknames and Clock Synchronization

After the first transaction, the server’s AUTH_DES authentication subsystem returns in its verifier to the client an integer nickname that the client may use in its further transactions instead of passing its netname, encrypted DES key and window every time. The nickname is most likely an index into a table on the server that stores for each client its netname, decrypted DES key and window. It should however be treated an opaque data by the client.

Though originally synchronized, client and server clocks can get out of sync. If this happens, the client RPC subsystem most likely will receive an RPC_AUTHERROR at which point it should resynchronize.

A client may still get the RPC_AUTHERROR error even though it is synchronized with the server. The reason is that the server’s nickname table is a limited size, and it may flush entries whenever it wants. The client should resend its original credential and the server will give it a new nickname. If a server crashes, the entire nickname table will be flushed, and all clients will have to resend their original credentials.

DES Authentication Protocol (in XDR language)

CODE EXAMPLE B–2  AUTH_DES Authentication Protocol
There are two kinds of credentials: one in which the client uses its full network name, and one in which it uses its "nickname" (just an unsigned integer) given to it by the server. The client must use its full name in its first transaction with the server, in which the server will return to the client its nickname. The client may use its nickname in all further transactions with the server. There is no requirement to use the nickname, but it is wise to use it for performance reasons.

```c
enum authdes_namekind {
    ADN_FULLNAME = 0,
    ADN_NICKNAME = 1
};
```

A 64-bit block of encrypted DES data

```c
typedef opaque des_block[8];
```

Maximum length of a network user’s name

```c
const MAXNETNAMELEN = 255;
```

A fullname contains the network name of the client, an encrypted conversation key and the window. The window is actually a lifetime for the credential. If the time indicated in the verifier timestamp plus the window has passed, then the server should expire the request and not grant it. To insure that requests are not replayed, the server should insist that timestamps be greater than the previous one seen, unless it is the first transaction. In the first transaction, the server checks instead that the window verifier is one less than the window.

```c
struct authdes_fullname {
    string name<MAXNETNAMELEN>; /* name of client */
    des_block key; /* PK encrypted conversation key */
    unsigned int window; /* encrypted window */
}; /* NOTE: PK means "public key" */
```

A credential is either a fullname or a nickname

```c
union authdes_credswitch(authdes_namekind adc_namekind) {
    case ADN_FULLNAME:
        authdes_fullname adc_fullname;
    case ADN_NICKNAME:
        unsigned int adc_nickname;
};
```

(Continued)
* A timestamp encodes the time since midnight, January 1, 1970.
* /
struct timestamp {
    unsigned int seconds;  /* seconds */
    unsigned int useconds; /* and microseconds */
};

/* Verifier: client variety
 */
struct authdes_verf_clnt {
    timestamp adv_timestamp; /* encrypted timestamp */
    unsigned int adv_winverf; /* encrypted window verifier */
};

/* Verifier: server variety
 * The server returns (encrypted) the same timestamp the client gave
 * it minus one second. It also tells the client its nickname to be
 * used in future transactions (unencrypted).
 */
struct authdes_verf_svr {
    timestamp adv_timeverf; /* encrypted verifier */
    unsigned int adv_nickname; /* new nickname for clnt */
};

**Diffie-Hellman Encryption**

In this scheme, there are two constants, PROOT and HEXMODULUS. The particular values chosen for these for the DES authentication protocol are:

```plaintext
const PROOT = 3;
const HEXMODULUS = /* hex */
  "d4a0ba0250b6fd2ec626e7ef637df76c716e22d0944b88b";
```

The way this scheme works is best explained by an example. Suppose there are two people “A” and “B” who want to send encrypted messages to each other. A and B each generate a random secret key that they do not disclose to anyone. Let these keys be represented as SK(A) and SK(B). They also publish in a public directory their public keys. These keys are computed as follows:

```plaintext
PK(A) = (PROOT ** SK(A)) mod HEXMODULUS
PK(B) = (PROOT ** SK(B)) mod HEXMODULUS
```

The ** notation is used here to represent exponentiation.

Now, both A and B can arrive at the common key between them, represented here as CK(A,B), without disclosing their secret keys.

A computes:
CK(A, B) = (PK(B) ** SK(A)) mod HEXMODULUS

while B computes:
CK(A, B) = (PK(A) ** SK(B)) mod HEXMODULUS

These two can be shown to be equivalent:
(PK(B)**SK(A)) mod HEXMODULUS =
(PK(A)**SK(B)) mod HEXMODULUS. Drop the mod HEXMODULUS parts
and assume modulo arithmetic to simplify the process:

PK(B) ** SK(A) = PK(A) ** SK(B)

Then replace PK(B) by what B computed earlier and likewise for PK(A).

((PROOT ** SK(B)) ** SK(A) = (PROOT ** SK(A)) ** SK(B)

which leads to:
PROOT ** (SK(A) * SK(B)) = PROOT ** (SK(A) * SK(B))

This common key CK(A,B) is not used to encrypt the timestamps used in
the protocol. It is used only to encrypt a conversation key that is then used
to encrypt the timestamps. The reason for doing this is to use the common
key as little as possible, for fear that it could be broken. Breaking the
conversation key is a far less serious offense, because conversations are
comparatively short-lived.

The conversation key is encrypted using 56-bit DES keys, yet the common
key is 192 bits. To reduce the number of bits, 56 bits are selected from the
common key as follows. The middle-most 8 bytes are selected from the
common key, and then parity is added to the lower order bit of each byte,
producing a 56-bit key with 8 bits of parity.

**AUTH_KERB Authentication**

To avoid compiling Kerberos code into the operating system kernel, the
kernel used in the SunOS 5.x implementation of AUTH_KERB uses a proxy
RPC daemon called kerbd. The daemon exports three procedures. Refer to
the kerbd(1M) manpage for more details.

1. KGETKCREDD is used by the server-side RPC to check the authenticator
   presented by the client.
2. KSETKCREDD returns the encrypted ticket and DES session key, given a
   primary name, instance, and realm.
3. KGETUCREDD is UNIX-specific. It returns the user’s ID, the group ID, and
   groups list, assuming that the primary name is mapped to a user name
   known to the server.

The best way to describe how Kerberos works is to use an example based on
a service currently implementing Kerberos: the network file system (NFS).
The NFS service on server s is assumed to have the well-known principal
name nfs. A privileged user on client $c$ is assumed to have the primary name root and an instance $c$. Note that (unlike AUTH_DES) when the user's ticket-granting ticket has expired, \texttt{kinit} must be reinvoked. NFS service for Kerberos mounts will fail until a new ticket-granting ticket is obtained.

NFS Mount Example

This section follows an NFS mount request from start to finish using AUTH_KERB. Since mount requests are executed as root in SunOS, the user’s identity is root.$c$. Client $c$ makes a MOUNTPROC_MOUNT request to the server $s$ to obtain the file handle for the directory to be mounted. The client mount program makes an NFS mount system call, handing the client kernel the file handle, mount flavor, time synchronization address, and the server’s well-known name, nfs.$s$. Next the client kernel contacts the server at the time synchronization host to obtain the client-server time bias.

The client kernel makes the following RPC calls: (1) \texttt{KSETKCRED} to the local kerbd to obtain the ticket and session key, (2) \texttt{NFSPROC_GETATTR} to the server’s NFS service, using the full name credential and verifier. The server receives the calls and makes the \texttt{KGETKCRED} call to its local kerbd to check the client’s ticket.

The server’s kerbd and the Kerberos library decrypt the ticket and return, among other data, the principal name and DES session key. The server checks that the ticket is still valid, uses the session key to decrypt the DES-encrypted portions of the credential and verifier, and checks that the verifier is valid.

The possible Kerberos authentication errors returned at this time are:

- \texttt{AUTH_BADCRED} is returned if the verifier is invalid (the decrypted $win$ in the credential and $win +1$ in the verifier do not match), or the timestamp is not within the window range
- \texttt{AUTH_REJECTEDCRED} is returned if a replay is detected
- \texttt{AUTH_BADVERF} is returned if the verifier is garbled

If no errors are received, the server caches the client’s identity and allocates a nickname (small integer) to be returned in the NFS reply. The server then checks if the client is in the same realm as the server. If it is, the server calls \texttt{KGETUCRED} to its local kerbd to translate the principal’s primary name into UNIX credentials. If it is not translatable, the user is marked anonymous. The server checks these credentials against the file system’s export information.

There are three cases to consider:

1. If the \texttt{KGETUCRED} call fails and anonymous requests are allowed, the UNIX credentials of the anonymous user are assigned.
2. If the \texttt{KGETUCRED} call fails and anonymous requests are not allowed, the NFS call fails with the \texttt{AUTH_TOO_WEAK}. 

RPC Protocol and Language Specification 211
3. If the KGETUCRED call succeeds, the credentials are assigned, and normal protection checking follows, including checking for root permission. Next the server sends an NFS reply, including the nickname and server’s verifier. The client receives the reply, decrypts and validates the verifier, and stores the nickname for future calls. The client makes a second NFS call to the server, and the calls to the server described earlier are repeated. The client kernel makes an NFSPROC_STATVFS call to the server’s NFS service, using the nickname credential and verifier described previously. The server receives the call and validates the nickname. If it is out of range, the error AUTH_BADCREDS is returned. The server uses the session key just obtained to decrypt the DES-encrypted portions of the verifier and validates the verifier.

The possible Kerberos authentication errors returned at this time are:
- AUTH_REJECTEDVERF is returned if the timestamp is invalid, a replay is detected, or if the timestamp is not within the window range
- AUTH_TIMEEXPIRE is returned if the service ticket is expired

If no errors are received, the server uses the nickname to retrieve the caller’s UNIX credentials. Then it checks these credentials against the file system’s export information, and sends an NFS reply that includes the nickname and the server’s verifier. The client receives the reply, decrypts and validates the verifier, and stores the nickname for future calls. Last, the client’s NFS mount system call returns, and the request is finished.

KERB Authentication Protocol (in XDR Language)

Code Example B–0 (AUTH_KERB) has many similarities to the one for AUTH_DES, shown in Code Example B–0. Note the differences.

```c
#define AUTH_KERB 4
/
* There are two kinds of credentials: one in which the client
* sends the (previously encrypted) Kerberos ticket, and one in
* which it uses its ‘’nickname’’ (just an unsigned integer)
* given to it by the server. The client must use its full name
* in its first transaction with the server, in which the server
* will return to the client its nickname. The client may use
* its nickname in all further transactions with the server
* (until the ticket expires). There is no requirement to use
* the nickname, but it is wise to use it for performance reasons.
*/
enum authkerb_namekind {
    AKN_FULLNAME = 0,
    AKN_NICKNAME = 1
};
```

(Continued)
/*
 * A fullname contains the encrypted service ticket and the
 * window. The window is actually a lifetime
 * for the credential. If the time indicated in the verifier
 * timestamp plus the window has passed, then the server should
 * expire the request and not grant it. To insure that requests
 * are not replayed, the server should insist that timestamps be
 * greater than the previous one seen, unless it is the first
 * transaction. In the first transaction, the server checks
 * instead that the window verifier is one less than the window.
 */

struct authkerb_fullname {
    KTEXT_ST ticket; /* Kerberos service ticket */
    unsigned long window; /* encrypted window */
};

/*
 * A credential is either a fullname or a nickname
 */
union authkerb_credswitch(authkerb_namekind akc_namekind){
    case AKN_FULLNAME:
        authkerb_fullname akc_fullname;
    case AKN_NICKNAME:
        unsigned long akc_nickname;
};

/*
 * A timestamp encodes the time since midnight, January 1, 1970.
 */
struct timestamp {
    unsigned long seconds; /* seconds */
    unsigned long useconds; /* and microseconds */
};

/*
 * Verifier: client variety
 */
struct authkerb_verf_clnt {
    timestamp akv_timestamp; /* encrypted timestamp */
    unsigned long akv_winverf; /* encrypted window verifier */
};

/*
 * Verifier: server variety
 * The server returns (encrypted) the same timestamp the client
gave it minus one second. It also tells the client its
 * nickname to be used in future transactions (unencrypted).
 */
struct authkerb_verf_svr {

The RPC Language Specification

Just as there was a need to describe the XDR data types in a formal language, there is also need to describe the procedures that operate on these XDR data types in a formal language as well. The RPC Language, an extension to the XDR language, serves this purpose. The following example is used to describe the essence of the language.

An Example Service Described in the RPC Language

Code Example B–0 shows the specification of a simple ping program.

CODE EXAMPLE B–4 ping Service Using RPC Language

```c
/*
 * Simple ping program
 */
program PING_PROG {
    version PING_VERS_PINGBACK {
        void
        PINGPROC_NULL(void) = 0;
        /*
        * ping the caller, return the round-trip time
        * in milliseconds. Return a minus one (-1) if
        * operation times-out
        */
        int
        PINGPROC_PINGBACK(void) = 1;
        /* void - above is an argument to the call */
    } = 2;
    /*
```

(Continued)
The first version described is PING_VERS_PINGBACK with two procedures, PINGPROC_NULL and PINGPROC_PINGBACK.

PINGPROC_NULL takes no arguments and returns no results, but it is useful for such things as computing round-trip times from the client to the server and back again. By convention, procedure 0 of any RPC program should have the same semantics, and never require authentication.

The second procedure returns the amount of time (in microseconds) that the operation used.

The next version, PING_VERS_ORIG, is the original version of the protocol and it does not contain PINGPROC_PINGBACK procedure. It is useful for compatibility with old client programs, and as this program matures it may be dropped from the protocol entirely.

**RPCL Syntax**

The RPC language (RPCL) is similar to C. This section describes the syntax of the RPC language, showing a few examples along the way. It also shows how RPC and XDR type definitions get compiled into C type definitions in the output header file.

An RPC language file consists of a series of definitions.

```
definition-list:
  definition;
  definition; definition-list
```

It recognizes six types of definitions.

```
definition:
  enum-definition
  const-definition
  typedef-definition
  struct-definition
  union-definition
  program-definition
```
Definitions are not the same as declarations. No space is allocated by a definition – only the type definition of a single or series of data elements. This means that variables still must be declared.

The RPC language is identical to the XDR language, except for the added definitions described in Table B–2.

<table>
<thead>
<tr>
<th>Term</th>
<th>Definition</th>
</tr>
</thead>
<tbody>
<tr>
<td>program-definition</td>
<td>program program-ident {version-list} = value</td>
</tr>
<tr>
<td>version-list</td>
<td>version;</td>
</tr>
<tr>
<td></td>
<td>version; version-list</td>
</tr>
<tr>
<td>version</td>
<td>version version-ident {procedure-list} = value</td>
</tr>
<tr>
<td>procedure-list</td>
<td>procedure;</td>
</tr>
<tr>
<td></td>
<td>procedure; procedure-list</td>
</tr>
<tr>
<td>procedure</td>
<td>type-ident procedure-ident (type-ident) = value</td>
</tr>
</tbody>
</table>

- The following keywords are added and cannot be used as identifiers: program version.
- Neither version name nor a version number can occur more than once within the scope of a program definition.
- Neither a procedure name nor a procedure number can occur more than once within the scope of a version definition.
- Program identifiers are in the same name space as constant and type identifiers.
- Only unsigned constants can be assigned to programs, versions, and procedures.

**Enumerations**

RPC/XDR enumerations have the same syntax as C enumerations.
enum-definition:
   "enum" enum-ident "{" 
   enum-value-list 
   "}" 
enum-value-list:
   enum-value 
   enum-value "," enum-value-list 
enum-value:
   enum-value-ident 
   enum-value-ident "=" value 

Here is an example of an XDR enum and the C enum to which it gets compiled.

```
enum colortype { 
    RED = 0, 
    GREEN = 1, 
    BLUE = 2 
};
```

typedef enum colortype colortype;

Constants

XDR symbolic constants may be used wherever an integer constant is used. For example, in array size specifications:

```
const-definition:
   const const-ident = integer 
```

The following example defines a constant, DOZEN as equal to 12:

```
const DOZEN = 12; --> #define DOZEN 12
```

Type Definitions

XDR typedefs have the same syntax as C typedefs.

```
typedef-definition: 
   typedef declaration
```

This example defines an fname_type used for declaring file name strings that have a maximum length of 255 characters.

```
typedef string fname_type<255>; --> typedef char *fname_type;
```

Declarations

In XDR, there are four kinds of declarations. These declarations must be a part of a struct or a typedef; they cannot stand alone:

```
declaration: 
   simple-declaration 
   fixed-array-declaration 
   variable-array-declaration 
   pointer-declaration
```
Simple Declarations

Simple declarations are just like simple C declarations:

\[ \text{simple-declaration:} \quad \text{type-ident variable-ident} \]

Example:

\[
\text{colortype color; \text{--> colortype color;}}
\]

Fixed-Length Array Declarations

Fixed-length array declarations are just like C array declarations:

\[ \text{fixed-array-declaration:} \quad \text{type-ident variable-ident [value]} \]

Example:

\[
\text{colortype palette[8]; \text{--> colortype palette[8];}}
\]

Many programmers confuse variable declarations with type declarations. It is important to note that rpcgen does not support variable declarations. This example is a program that will not compile:

```
int data[10];
program P {
    version V {
        int PROC(data) = 1;
    } = 1;
} = 0x200000;
```

The example above will not compile because of the variable declaration:

```
int data[10]
```

Instead, use:

```
typedef int data[10];
```

or

```
struct data {int dummy [10]};
```

Variable-Length Array Declarations

Variable-length array declarations have no explicit syntax in C. The XDR language does have a syntax, using angle brackets:

\[ \text{variable-array-declaration:} \quad \text{type-ident variable-ident <value> < >} \]

The maximum size is specified between the angle brackets. The size may be omitted, indicating that the array may be of any size:

```
int heights<12>; /* at most 12 items */
int widths<>; /* any number of items */
```
Because variable-length arrays have no explicit syntax in C, these declarations are compiled into struct declarations. For example, the heights declaration compiled into the following struct:

```c
struct {
    u_int heights_len; /* # of items in array */
    int *heights_val; /* pointer to array */
} heights;
```

The number of items in the array is stored in the _len component and the pointer to the array is stored in the _val component. The first part of each component name is the same as the name of the declared XDR variable (heights).

**Pointer Declarations**

Pointer declarations are made in XDR exactly as they are in C. Address pointers are not really sent over the network; instead, XDR pointers are useful for sending recursive data types such as lists and trees. The type is called “optional-data,” not “pointer,” in XDR language:

```c
pointer-declaration:
    type-ident *variable-ident
```

Example:

```c
listitem *next; --> listitem *next;
```

**Structures**

An RPC/XDR struct is declared almost exactly like its C counterpart. It looks like the following:

```c
struct-definition:
    struct struct-ident "{"
    declaration-list
    "}"* declaration-list:
    declaration ";"
    declaration ";" declaration-list
```

The following XDR structure is an example of a two-dimensional coordinate and the C structure that it compiles into:

```c
struct coord {
    int x;
    int y;
};
```

```c
typedef struct coord coord;
```
Unions

XDR unions are discriminated unions, and do not look like C unions – they are more similar to Pascal variant records:

union-definition:
    "union" union-ident "switch" "("simple declaration")" "{" 
    case-list 
"}" 

case-list: 
    "case" value ":" declaration ";" 
    case-list 
    "default" ":" declaration ";" 

The following is an example of a type returned as the result of a “read data” operation: If there is no error, return a block of data; otherwise, don’t return anything.

union read_result switch (int errno) {
    case 0:
        opaque data[1024];
    default:
        void;
};

It compiles into the following:

struct read_result {
    int errno;
    union {
        char data[1024];
    } read_result_u;
};

typedef struct read_result read_result;

Notice that the union component of the output struct has the same name as the type name, except for the trailing _u.

Programs

RPC programs are declared using the following syntax:

program-definition:
    "program" program-ident "{" 
    version-list 
"}" 
version-list:
    version ";" 
version ";" version-list
version:
 "version" version-ident "[
   procedure-list
  ]" "=" value;
procedure-list:
 procedure ";
 procedure ";" procedure-list
procedure:
  type-ident procedure-ident "(" type-ident ")" "=" value;

When the −N option is specified, rpcgen also recognizes the following syntax:

procedure:
  type-ident procedure-ident "(" type-ident-list ")" "=" value;
  type-ident-list:
    type-ident
    type-ident "," type-ident-list

For example:

    /*
    * time.x: Get or set the time. Time is represented as seconds
    * since 0:00, January 1, 1970.
    */
    program TIMEPROG {
      version TIMEVERS {
        unsigned int TIMEGET(void) = 1;
        void TIMESET(unsigned) = 2;
      } = 1;
    } = 0x20000044;

Note that the void argument type means that no argument is passed.

This file compiles into these #define statements in the output header file:

#define TIMEPROG 0x20000044
#define TIMEVERS 1
#define TIMEGET 1
#define TIMESET 2

Special Cases
There are several exceptions to the RPC language rules.

C-style Mode
In the new features section we talked about the features of the C-style mode of rpcgen. These features have implications with regard to the passing of void arguments. No arguments need be passed if their value is void.
Booleans

C has no built-in boolean type. However, the RPC library uses a boolean type called bool_t that is either TRUE or FALSE. Parameters declared as type bool in XDR language are compiled into bool_t in the output header file.

Example:

bool married; --> bool_t married;

Strings

The C language has no built-in string type, but instead uses the null-terminated char * convention. In C, strings are usually treated as null-terminated single-dimensional arrays.

In XDR language, strings are declared using the string keyword, and compiled into type char * in the output header file. The maximum size contained in the angle brackets specifies the maximum number of characters allowed in the strings (not counting the NULL character). The maximum size may be omitted, indicating a string of arbitrary length.

Examples:

string name<32>; --> char *name;
string longname<>; --> char *longname;

Note - NULL strings cannot be passed; however, a zero-length string (that is, just the terminator or NULL byte) can be passed.

Opaque Data

Opaque data is used in XDR to describe untyped data, that is, sequences of arbitrary bytes. It may be declared either as a fixed length or variable length array. Examples:

opaque diskblock[512]; --> char diskblock[512];
opaque filedatal<1024>; --> struct {
    u_int filedata_len;
    char *filedata_val;
} filedatal;

Voids

In a void declaration, the variable is not named. The declaration is just void and nothing else. Void declarations can only occur in two places: union definitions and program definitions (as the argument or result of a remote procedure, for example no arguments are passed.)
**rpcbind Protocol**

*rpcbind* maps RPC program and version numbers to universal addresses, thus making dynamic binding of remote programs possible.

*rpcbind* is bound to a well-known address of each supported transport, and other programs register their dynamically allocated transport addresses with it. *rpcbind* then makes those addresses publicly available. Universal addresses are string representations of the transport-dependent address. They are defined by the addressing authority of the given transport.

*rpcbind* also aids in broadcast RPC. RPC programs will have different addresses on different machines, so there is no way to broadcast directly to all these programs. *rpcbind*, however, has a well-known address. So, to broadcast to a given program, the client actually sends its message to the *rpcbind* process on the machine it chooses to reach. *rpcbind* picks up the broadcast and calls the local service specified by the client. When *rpcbind* gets a reply from the local service, it passes the reply on to the client.

**CODE EXAMPLE B–5  rpcbind Protocol Specification (in RPC Language)**

```c
/*
 * rpcb_prot.x
 * RPCBIND protocol in rpc language
 */

/*
 A mapping of (program, version, network ID) to universal address
 */
struct rpcb {
    unsigned long r_prog; /* program number */
    unsigned long r_vers; /* version number */
    string r_netid<>;       /* network id */
    string r_addr<>;        /* universal address */
    string r_owner<>;       /* owner of this service */
};

/* A list of mappings */
struct rpcblist {
    rpcb rpcb_map;
    struct rpcblist *rpcb_next;
};

/* Arguments of remote calls */
struct rpcb_rmtcallargs {
    unsigned long prog;  /* program number */
    unsigned long vers;  /* version number */
    unsigned long proc;  /* procedure number */
    opaque args<>;        /* argument */
};

/* Results of the remote call */
struct rpcb_rmtcallres {
    string addr<>;        /* remote universal address */
    opaque results<>;     /* result */
};
```

(Continued)
struct rpcb_entry {
    string r_maddr<>;  /* merged address of service */
    string r_nc_netid<>;  /* netid field */
    unsigned long r_nc_semantics; /* semantics of transport */
    string r_nc_protofmly<>;  /* protocol family */
    string r_nc_proto<>;  /* protocol name */
};

/* A list of addresses supported by a service. */
struct rpcb_entry_list {
    rpcb_entry rpcb_entry_map;
    struct rpcb_entry_list *rpcb_entry_next;
};

typedef rpcb_entry_list *rpcb_entry_list_ptr;

/* rpcbind statistics */
const rpcb_highproc_2 = RPCBPROC_CALLIT;
const rpcb_highproc_3 = RPCBPROC_TADDR2UADDR;
const rpcb_highproc_4 = RPCBPROC_GETSTAT;
const RPCBSTAT_HIGHPROC = 13; /* # of procs in rpcbind V4 plus one */
const RPCBVERS_STAT = 3; /* provide only for rpcbind V2, V3 and V4 */
const RPCBVERS_4_STAT = 2;
const RPCBVERS_3_STAT = 1;
const RPCBVERS_2_STAT = 0;

/* Link list of all the stats about getport and getaddr */
struct rpcbs_addrlist {
    unsigned long prog;
    unsigned long vers;
    int success;
    int failure;
    string netid<>;
    struct rpcbs_addrlist *next;
};

/* Link list of all the stats about rmtcall */
struct rpcbs_rmtcalllist {
    unsigned long prog;
    unsigned long vers;
    unsigned long proc;
typedef int rpcbs_proc[RPCBSTAT_HIGHPROC];
typedef rpcbs_addrlist *rpcbs_addrlist_ptr;
typedef rpcbs_rmtcalllist *rpcbs_rmtcalllist_ptr;

struct rpcb_stat {
  rpcbs_proc info;
  int setinfo;
  int unsetinfo;
  rpcbs_addrlist_ptr addrinfo;
  rpcbs_rmtcalllist_ptr rmtinfo;
};

/*
 * One rpcb_stat structure is returned for each version of rpcbind
 * being monitored.
 */
typedef rpcb_stat rpcb_stat_byvers[RPCBVERS_STAT];

/* rpcbind procedures */
program RPCBPROG {
  version RPCBVERS {
    void
    RPCBPROC_NULL(void) = 0;
    /*
     * Registers the tuple [r_prog, rVers, r_addr, r_owner,
     * r_netid]. The rpcbind server accepts requests for this
     * procedure on only the loopback transport for security
     * reasons. Returns TRUE if successful, FALSE on failure.
     */
    bool
    RPCBPROC_SET(rpcb) = 1;
    /*
     * Unregisters the tuple [r_prog, rVers, r_owner, r_netid].
     * If vers is zero, all versions are unregistered. The rpcbind
     * server accepts requests for this procedure on only the
     * loopback transport for security reasons. Returns TRUE if
     * successful, FALSE on failure.
     */
    bool
    RPCBPROC_UNSET(rpcb) = 2;
    /*
     * Returns the universal address where the triple [r_prog,
* r_vers, r_netid] is registered. If r_addr specified,
  * return a universal address merged on r_addr. Ignores
  * r_owner. Returns FALSE on failure.

*/
string
RPCBPROC_GETADDR(rpcb) = 3;

/* Returns a list of all mappings. */
rpcblist
RPCBPROC_DUMP(void) = 4;

/*
 * Calls the procedure on the remote machine. If it is not
 * registered, this procedure IS quiet; that is, it DOES NOT
 * return error information.
 */
rpcb_rmtcallres
RPCBPROC_CALLIT(rpcb_rmtcallargs) = 5;

/*
 * Returns the time on the rpcbind server’s system.
 */
unsigned int
RPCBPROC_GETTIME(void) = 6;

struct netbuf
RPCBPROC_UADDR2TADDR(string) = 7;

string
RPCBPROC_TADDR2UADDR(struct netbuf) = 8;

}) = 3;
version RPCBVERS4 {
bool
RPCBPROC_SET(rpcb) = 1;

bool
RPCBPROC_UNSET(rpcb) = 2;

string
RPCBPROC_GETADDR(rpcb) = 3;

rpcblist_ptr
RPCBPROC_DUMP(void) = 4;

/*
 * NOTE: RPCBPROC_BCAST has the same functionality as CALLIT;
*/
* the new name is intended to indicate that this procedure  
* should be used for broadcast RPC, and RPCBPROC_INDIRECT  
* should be used for indirect calls.  
*/

rpcb_rmtcallres
RPCBPROC_BCAST(rpcb_rmtcallargs) = RPCBPROC_CALLIT;

unsigned int
RPCBPROC_GETTIME(void) = 6;

struct netbuf
RPCBPROC_UADDR2TADDR(string) = 7;

string
RPCBPROC_TADDR2UADDR(struct netbuf) = 8;

/*
* Same as RPCBPROC_GETADDR except that if the given version
* number is not available, the address is not returned.
*/

string
RPCBPROC_GETVERSADDR(rpcb) = 9;

/*
* Calls the procedure on the remote machine. If it is not
* registered, this procedure IS NOT quiet; that is, it DOES
* return error information.
*/

rpcb_rmtcallres
RPCBPROC_INDIRECT(rpcb_rmtcallargs) = 10;

/*
* Same as RPCBPROC_GETADDR except that it returns a list of
* addresses registered for the combination (prog, vers).
*/

rpcb_entry_list_ptr
RPCBPROC_GETADDRLIST(rpcb) = 11;

/*
* Returns statistics about the rpcreg server’s activity.
*/

rpcb_stat_byvers
RPCBPROC_GETSTAT(void) = 12;
**rpcbind Operation**

* rpcbind is contacted by way of an assigned address specific to the transport being used. For TCP/IP and UDP/IP, for example, it is port number 111. Each transport has such an assigned well known address. The following is a description of each of the procedures supported by rpcbind.

**RPCBPROC_NULL**

This procedure does no work. By convention, procedure zero of any program takes no parameters and returns no results.

**RPCBPROC_SET**

When a program first becomes available on a machine, it registers itself with the rpcbind program running on the same machine. The program passes its program number `prog`; version number `vers`; network identifier `netid`; and the universal address `uaddr`; on which it awaits service requests.

The procedure returns a Boolean response with the value `TRUE` if the procedure successfully established the mapping and `FALSE` otherwise. The procedure refuses to establish a mapping if one already exists for the ordered set `(prog, vers, netid)`.

Note that neither `netid` nor `uaddr` can be `NULL`, and that `netid` should be a valid network identifier on the machine making the call.

**RPCBPROC_UNSET**

When a program becomes unavailable, it should unregister itself with the rpcbind program on the same machine.

The parameters and results have meanings identical to those of `RPCBPROC_SET`. The mapping of the `(prog, vers, netid)` tuple with `uaddr` is deleted.

If `netid` is `NULL`, all mappings specified by the ordered set `(prog, vers, *)` and the corresponding universal addresses are deleted. Only the owner of the service or the super-user is allowed to unset a service.
RPCBPROC_GETADDR

Given a program number `prog`, version number `vers`, and network identifier `netid`, this procedure returns the universal address on which the program is awaiting call requests.

The `netid` field of the argument is ignored and the `netid` is inferred from the `netid` of the transport on which the request came in.

RPCBPROC_DUMP

This procedure lists all entries in rpcbind’s database.

The procedure takes no parameters and returns a list of program, version, `netid`, and universal addresses. Call this procedure using a stream rather than a datagram transport to avoid the return of a large amount of data.

RPCBPROC_CALLIT

This procedure allows a caller to call another remote procedure on the same machine without knowing the remote procedure’s universal address. It is intended for supporting broadcasts to arbitrary remote programs via rpcbind’s universal address.

The parameters `prog`, `vers`, `proc`, and the `args_ptr` are the program number, version number, procedure number, and parameters of the remote procedure.

**Note** - This procedure sends a response only if the procedure was successfully executed, and is silent (no response) otherwise.

The procedure returns the remote program’s universal address, and the results of the remote procedure.

RPCBPROC_GETTIME

This procedure returns the local time on its own machine in seconds since midnight of January 1, 1970.

RPCBPROC_UADDR2TADDR

This procedure converts universal addresses to transport (netbuf) addresses. RPCBPROC_UADDR2TADDR is equivalent to uaddr2taddr. See the netdir(3N) manpage. Only processes that cannot link to the name-to-address library modules should use RPCBPROC_UADDR2TADDR.
RPCBPROC_TADDR2UADDR

This procedure converts transport (netbuf) addresses to universal addresses. RPCBPROC_TADDR2UADDR is equivalent to taddr2uaddr. See the netdir(3N) manpage. Only processes that can not link to the name-to-address library modules should use RPCBPROC_TADDR2UADDR.

Version 4 rpcbind

Version 4 of the rpcbind protocol includes all of the above procedures, and adds several others.

RPCBPROC_BCAST

This procedure is identical to the version 3 RPCBPROC_CALLIT procedure. The new name indicates that the procedure should be used for broadcast RPCs only. RPCBPROC_INDIRECT, defined in the following text, should be used for indirect RPC calls.

RPCBPROC_GETVERSADDR

This procedure is similar to RPCBPROC_GETADDR. The difference is the r_vers field of the rpcb structure can be used to specify the version of interest. If that version is not registered, no address is returned.

RPCBPROC_INDIRECT

This procedure is similar to RPCBPROC_CALLIT. Instead of being silent about errors (such as the program not being registered on the system), this procedure returns an indication of the error. This procedure should not be used for broadcast RPC. It is intended to be used with indirect RPC calls only.

RPCBPROC_GETADDR LIST

This procedure returns a list of addresses for the given rpcb entry. The client may be able use the results to determine alternate transports that it can use to communicate with the server.
RPCBPROC_GETSTAT

This procedure returns statistics on the activity of the rpcbind server. The information lists the number and kind of requests the server has received.

**Note** - All procedures except RPCBPROC_SET and RPCBPROC_UNSET can be called by clients running on a machine other than a machine on which rpcbind is running. rpcbind accepts only RPCPROC_SET and RPCPROC_UNSET requests on the loopback transport.

---

**Bibliography**

For further information on the technologies and architectures discussed in this appendix, reference the following resources.


XDR Protocol Specification

This appendix contains the XDR Protocol Language Specification.

- “XDR Protocol Introduction” on page 233
- “XDR Data Type Declarations” on page 234
- “The XDR Language Specification” on page 247

XDR Protocol Introduction

External data representation (XDR) is a standard for the description and encoding of data. The XDR protocol is useful for transferring data between different computer architectures and has been used to communicate data between such diverse machines as the Sun® Workstation®, VAX®, IBM® PC, and Cray. XDR fits into the ISO reference model’s presentation layer (layer 6) and is roughly analogous in purpose to X.409, ISO Abstract Syntax Notation. The major difference between the two is that XDR uses implicit typing, while X.409 uses explicit typing.

XDR uses a language to describe data formats and only can be used to describe data; it is not a programming language. This language makes it possible to describe intricate data formats in a concise manner. The XDR language is similar to the C language. Protocols such as RPC and the NFS use XDR to describe the format of their data.

The XDR standard assumes that bytes (or octets) are portable and that a byte is defined to be 8 bits of data.
Graphic Box Notation

This appendix uses graphic box notation for illustration and comparison. In most illustrations, each box depicts a byte. The representation of all items requires a multiple of four bytes (or 32 bits) of data. The bytes are numbered 0 through \( n-1 \). The bytes are read or written to some byte stream such that byte \( m \) always precedes byte \( m+1 \). The \( n \) bytes are followed by enough (0 to 3) residual zero bytes, \( r \), to make the total byte count a multiple of four. Ellipses (\( \ldots \)) between boxes show zero or more additional bytes where required. For example:

```
<table>
<thead>
<tr>
<th>byte 0</th>
<th>byte 1</th>
<th>byte n-1</th>
<th>0</th>
<th>0</th>
</tr>
</thead>
<tbody>
<tr>
<td>n bytes</td>
<td></td>
<td></td>
<td>r bytes</td>
<td></td>
</tr>
<tr>
<td>n+r (where (n+r) mod 4 = 0)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
```

Basic Block Size

Choosing the XDR block size requires a trade off. Choosing a small size such as two makes the encoded data small, but causes alignment problems for machines that are not aligned on these boundaries. A large size such as eight means the data will be aligned on virtually every machine, but causes the encoded data to grow too large. Four was chosen as a compromise. Four is big enough to support most architectures efficiently, except for rare machines such as the eight-byte aligned Cray.

This is not to say that the computers cannot utilize standard XDR, just that they do so at a greater overhead per data item than 4-byte (32-bit) architectures. Four is also small enough to keep the encoded data restricted to a reasonable size.

The same data should encode into an equivalent result on all machines, so that encoded data can be compared or checksummed. So, variable length data must be padded with trailing zeros.

XDR Data Type Declarations

Each of the sections that follow:

- Describe a data type defined in the XDR standard
- Show how that data type is declared in the language
- Include a graphic illustration of the encoding
For each data type in the language we show a general paradigm declaration. Note that angle brackets (< and >) denote variable length sequences of data and square brackets ([ and ]) denote fixed-length sequences of data. $n$, $m$ and $r$ denote integers. For the full language specification, refer to “The XDR Language Specification” on page 247.

For some data types, specific examples are included. A more extensive example is given in the section, “XDR Data Description” on page 250.

### Signed Integer

**Description**

An XDR signed integer is a 32-bit datum that encodes an integer in the range $[-2147483648, 2147483647]$. The integer is represented in two’s complement notation; the most and least significant bytes are 0 and 3, respectively.

**Declaration**

Integers are declared:
```
int identifier;
```

**Encoding**

Integer

<table>
<thead>
<tr>
<th>(MSB)</th>
<th>(LSB)</th>
</tr>
</thead>
<tbody>
<tr>
<td>byte 0</td>
<td>byte 1</td>
</tr>
</tbody>
</table>

32 bits

### Unsigned Integer

**Description**

An XDR unsigned integer is a 32-bit datum that encodes a nonnegative integer in the range $[0, 4294967295]$. The integer is represented by an unsigned binary number whose most and least significant bytes are 0 and 3, respectively.
Declaration
An unsigned integer is declared as follows:

\[
\text{unsigned int identifier;}
\]

Encoding
Unsigned Integer

<table>
<thead>
<tr>
<th>(MSB)</th>
<th>(LSB)</th>
</tr>
</thead>
<tbody>
<tr>
<td>byte 0</td>
<td>byte 1</td>
</tr>
</tbody>
</table>

| 32 bits |

Enumerations

Description
Enumerations have the same representation as signed integers and are handy for describing subsets of the integers.

Declaration
Enumerated data is declared as follows:

\[
\text{enum \{name-identifier = constant, \ldots \} identifier;}
\]

For example, an enumerated type could represent the three colors red, yellow, and blue as follows:

\[
\text{enum \{RED = 2, YELLOW = 3, BLUE = 5\} colors;}
\]

It is an error to assign to an \textit{enum} an integer that has not been assigned in the \textit{enum} declaration.

Encoding
See “Signed Integer” on page 235.
Booleans

Description
Booleans are important enough and occur frequently enough to warrant their own explicit type in the standard. Booleans are integers of value 0 or 1.

Declaration
Booleans are declared as follows:
```c
bool identifier;
```
This is equivalent to:
```c
enum {FALSE = 0, TRUE = 1} identifier;
```

Encoding
See “Signed Integer” on page 235.

Hyper Integer and Unsigned Hyper Integer

Description
The standard defines 64-bit (8-byte) numbers called hyper int and unsigned hyper int whose representations are the obvious extensions of integer and unsigned integer, defined above. They are represented in two’s complement notation; the most and least significant bytes are 0 and 7, respectively.

Declaration
Hyper integers are declared as follows:
```c
hyper int identifier;
unsigned hyper int identifier;
```

Encoding
Hyper Integer

<table>
<thead>
<tr>
<th>(MSB)</th>
<th>(LSB)</th>
</tr>
</thead>
<tbody>
<tr>
<td>byte 0</td>
<td>byte 1</td>
</tr>
</tbody>
</table>

64 bits
Floating Point

Description
The standard defines the floating-point data type `float` (32-bits or 4-bytes). The encoding used is the IEEE standard for normalized single-precision floating-point numbers [1]. The following three fields describe the single-precision floating-point number:

S: The sign of the number. Values 0 and 1 represent positive and negative, respectively. One bit.

E: The exponent of the number, base 2. There are eight bits in this field. The exponent is biased by 127.

F: The fractional part of the number’s mantissa, base 2. There are 23 bits are in this field.

Therefore, the floating-point number is described by:

\((-1)^S \times 2^{(E-Bias)} \times 1.F\)

Declaration
Single-precision floating-point data is declared as follows:

```c
float identifier;
```

Double-precision floating-point data is declared as follows:

```c
double identifier;
```

Encoding
Double-Precision Floating Point

Just as the most and least significant bytes of an integer are 0 and 3, the most and least significant bits of a double-precision floating-point number are 0 and 63. The beginning bit (and most significant bit) offsets of S, E, and F are 0, 1, and 12, respectively.

These offsets refer to the logical positions of the bits, not to their physical locations (which vary from medium to medium).

The IEEE specifications should be consulted about the encoding for signed zero, signed infinity (overflow), and de-normalized numbers (underflow) [1].
According to IEEE specifications, the NaN (not a number) is system dependent and should not be used externally.

**Quadruple-Precision Floating Point**

**Description**

The standard defines the encoding for the quadruple-precision floating-point data type `quadruple` (128 bits or 16 bytes). The encoding used is the IEEE standard for normalized quadruple-precision floating-point numbers [1]. The standard encodes the following three fields, which describe the quadruple-precision floating-point number:

- **s**: The sign of the number. Values 0 and 1 represent positive and negative, respectively. One bit.
- **e**: The exponent of the number, base 2. There are 15 bits in this field. The exponent is biased by 16383.
- **f**: The fractional part of the number’s mantissa, base 2. There are 111 bits in this field.

Therefore, the floating-point number is described by:

\((-1)^s \times 2^{(e-Bias)} \times 1.f\)

**Declaration**

```c
quadruple identifier;
```

**Encoding**

![Quadruple-Precision Floating Point Encoding Diagram]

Just as the most and least significant bytes of an integer are 0 and 3, the most and least significant bits of a quadruple-precision floating-point number are 0 and 127. The beginning bit (and most significant bit) offsets of S, E, and F are 0, 1, and 16, respectively. These offsets refer to the logical positions of the bits, not to their physical locations (which vary from medium to medium).

The IEEE specifications should be consulted about the encoding for signed zero, signed infinity (overflow), and de-normalized numbers (underflow) [1].
According to IEEE specifications, the NaN (not a number) is system
dependent and should not be used externally.

Fixed-Length Opaque Data

Description
At times, fixed-length uninterpreted data needs to be passed among
machines. This data is called opaque.

Declaration
Opaque data is declared as follows:

\[
\text{opaque identifier}[n];
\]

where the constant \(n\) is the (static) number of bytes necessary to contain the
opaque data. The \(n\) bytes are followed by enough (0 to 3) residual zero bytes,
\(r\), to make the total byte count of the opaque object a multiple of four.

Encoding
The \(n\) bytes are followed by enough (0 to 3) residual zero bytes, \(r\), to make
the total byte count of the opaque object a multiple of four.

Fixed-Length Opaque

```
+---+---+---+---+---+---+---+---+
| 0 | 1 | ... | 0 | 0 |
+---+---+---+---+---+---+---+---+
```

Variable-Length Opaque Data

Description
The standard also provides for variable-length (counted) opaque data,
declared as a sequence of \(n\) (numbered 0 through \(n-1\)) arbitrary bytes to be
the number \(n\) encoded as an unsigned integer (as described subsequently),
and followed by the \(n\) bytes of the sequence.
Byte $b$ of the sequence always precedes byte $b+1$ of the sequence, and byte 0 of the sequence always follows the sequence’s length (count). The $n$ bytes are followed by enough (0 to 3) residual zero bytes, $r$, to make the total byte count a multiple of four.

**Declaration**

Variable-length opaque data is declared in the following way:

opaque identifier<$m$>;

or

opaque identifier<>; ;

The constant $m$ denotes an upper bound of the number of bytes that the sequence may contain. If $m$ is not specified, as in the second declaration, it is assumed to be $(2^{*32}) - 1$, the maximum length. For example, a filing protocol may state that the maximum data transfer size is 8192 bytes, as follows:

opaque filedata<8192>;

**Encoding**

**Variable-Length Opaque**

```
  length n   n bytes    r bytes
  n+r (where (n+r) mod4 = 0)
```

It is an error to encode a length greater than the maximum described in the specification.

**Counted Byte Strings**

**Description**

The standard defines a string of $n$ (numbered 0 through $n-1$) ASCII bytes to be the number $n$ encoded as an unsigned integer (as described previously), and followed by the $n$ bytes of the string. Byte $b$ of the string always precedes byte $b+1$ of the string, and byte 0 of the string always follows the string’s length. The $n$ bytes are followed by enough (0 to 3) residual zero bytes, $r$, to make the total byte count a multiple of four.
Declaration

Counted byte strings are declared as follows:

```c
string object<m>;
```
or

```c
string object<>;
```

The constant `m` denotes an upper bound of the number of bytes that a string may contain. If `m` is not specified, as in the second declaration, it is assumed to be `(2**32) - 1`, the maximum length. The constant `m` would normally be found in a protocol specification. For example, a filing protocol may state that a file name can be no longer than 255 bytes, as follows:

```c
string filename<255>;
```

Encoding

String

```
<table>
<thead>
<tr>
<th>1</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>...</th>
</tr>
</thead>
<tbody>
<tr>
<td>length n</td>
<td>byte 0</td>
<td>byte 1</td>
<td>byte n-1</td>
<td>0</td>
</tr>
</tbody>
</table>
```

It is an error to encode a length greater than the maximum described in the specification.

Fixed-Length Array

Fixed-length arrays of elements numbered 0 through `n-1` are encoded by individually encoding the elements of the array in their natural order, 0 through `n-1`. Each element's size is a multiple of four bytes. Though all elements are of the same type, the elements may have different sizes. For example, in a fixed-length array of strings, all elements are of type `string`, yet each element will vary in its length.

Declaration

Declarations for fixed-length arrays of homogenous elements are in the following form:

```c
type-name identifier[n];
```

Encoding

```
Fixed-Length Array
```
Variable-Length Array

Description
Counted arrays allow variable-length arrays to be encoded as homogeneous elements: the element count \( n \) (an unsigned integer) is followed by each array element, starting with element 0 and progressing through element \( n-1 \).

Declaration
The declaration for variable-length arrays follows this form:

\[
\text{type-name identifier}^{<m>};
\]

or

\[
\text{type-name identifier}^{<>};
\]

The constant \( m \) specifies the maximum acceptable element count of an array. If \( m \) is not specified, it is assumed to be \( (2^{32}) - 1 \).

Encoding
Counted Array

It is an error to encode a length greater than the maximum described in the specification.
Structure

Description
The components of the structure are encoded in the order of their declaration in the structure. Each component’s size is a multiple of four bytes, though the components may be different sizes.

Declaration
Structures are declared as follows:

```c
struct {
    component-declaration-A;
    component-declaration-B;
    ...
} identifier;
```

Encoding
Structure

<table>
<thead>
<tr>
<th>component A</th>
<th>component B</th>
</tr>
</thead>
</table>

Discriminated Union

Description
A discriminated union is a type composed of a discriminant followed by a type selected from a set of prearranged types according to the value of the discriminant. The type of discriminant is either `int`, `unsigned int`, or an enumerated type, such as `bool`. The component types are called arms of the union, and are preceded by the value of the discriminant that implies their encoding.

Declaration
Discriminated unions are declared as follows:

```c
union switch (discriminant-declaration) {
    case discriminant-value-A:
        arm-declaration-A;
    case discriminant-value-B:
        arm-declaration-B;
    ...
}
```
Each `case` keyword is followed by a legal value of the discriminant. The
default arm is optional. If it is not specified, then a valid encoding of the
union cannot take on unspecified discriminant values. The size of the
implied arm is always a multiple of four bytes.

The discriminated union is encoded as its discriminant followed by the
encoding of the implied arm.

### Encoding

**Discriminated Union**

<p>| | | | | | | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

- `discriminant` (4 bytes)
- `implied arm`

### Void

**Description**

An XDR `void` is a 0-byte quantity. Voids are useful for describing operations
that take no data as input or no data as output. They are also useful in
unions, where some arms may contain data and others do not.

**Declaration**

The declaration is simply as follows:

```plaintext
void;
```

### Constant

**Description**

`const` is used to define a symbolic name for a constant; it does not declare
any data. The symbolic constant may be used anywhere a regular constant
may be used.

The following example defines a symbolic constant `DOZEN`, equal to 12.
const DOZEN = 12;

Declaration

The declaration of a constant follows this form:
const name-identifier = n;

Typedef

typedef does not declare any data either, but serves to define new identifiers for declaring data. The syntax is:
typedef declaration;

The new type name is actually the variable name in the declaration part of the typedef. The following example defines a new type called eggbox using an existing type called egg and the symbolic constant DOZEN:
typedef egg eggbox[DOZEN];

Variables declared using the new type name have the same type as the new type name would have in the typedef, if it were considered a variable. For example, the following two declarations are equivalent in declaring the variable fresheggs:
eggbox fresheggs;
egg fresheggs[DOZEN];

When a typedef involves a struct, enum, or union definition, there is another (preferred) syntax that may be used to define the same type. In general, a typedef of the following form:
typedef <<struct, union, or enum definition>> identifier;

may be converted to the alternative form by removing the typedef part and placing the identifier after the struct, enum, or union keyword, instead of at the end. For example, here are the two ways to define the type bool:
typedef enum { /* using typedef */
    FALSE = 0,
    TRUE = 1
} bool;
enum bool { /* preferred alternative */
    FALSE = 0,
    TRUE = 1
};

This syntax is preferred because one does not have to go to the end of a declaration to learn the name of the new type.

Optional-Data

Optional-data is one kind of union that occurs so frequently that it is given a special syntax of its own for declaring it. It is declared as follows:
type-name *identifier;

This is equivalent to the following union:

```c
union switch (bool opted) {
  case TRUE:
    type-name element;
  case FALSE:
    void;
} identifier;
```

It is also equivalent to the following variable-length array declaration, since the Boolean `opted` can be interpreted as the length of the array:

```c
type-name identifier<1>;
```

Optional-data is useful for describing recursive data-structures, such as linked lists and trees.

---

The XDR Language Specification

Notational Conventions

This specification uses a modified Backus-Naur Form notation for describing the XDR language. Here is a brief description of the notation:

1. The characters |, (,), [], and * are special.
2. Terminal symbols are strings of any characters embedded in quotes (").
3. Nonterminal symbols are strings of nonspecial italic characters.
4. Alternative items are separated by a vertical bar (|).
5. Optional items are enclosed in brackets.
6. Items are grouped together by enclosing them in parentheses.
7. A * following an item means 0 or more occurrences of the item.

For example, consider the following pattern:

```
"a " "very" (", " " very")* [" cold " "and"] " rainy "
   ("day" | "night")
```

An infinite number of strings match this pattern. A few of them are:

- a very rainy day
- a very, very rainy day
- a very cold and rainy day
- a very, very, very cold and rainy night

Lexical Notes

1. Comments begin with /* and end with */.
2. White space serves to separate items and is otherwise ignored.
3. An identifier is a letter followed by an optional sequence of letters, digits, or underbars (_). The case of identifiers is not ignored.
4. A constant is a sequence of one or more decimal digits, optionally preceded by a minus-sign (-).

**CODE EXAMPLE C–1**  XDR Specification

<table>
<thead>
<tr>
<th>Syntax Information</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>declaration:</strong></td>
</tr>
<tr>
<td>type-specifier identifier</td>
</tr>
<tr>
<td>type-specifier identifier &quot;[&quot; value &quot;]&quot;</td>
</tr>
<tr>
<td>type-specifier identifier &quot;&lt;&quot; [ value ] &quot;&gt;&quot;</td>
</tr>
<tr>
<td>&quot;opaque&quot; identifier &quot;[&quot; value &quot;]&quot;</td>
</tr>
<tr>
<td>&quot;opaque&quot; identifier &quot;&lt;&quot; [ value ] &quot;&gt;&quot;</td>
</tr>
<tr>
<td>&quot;string&quot; identifier &quot;&lt;&quot; [ value ] &quot;&gt;&quot;</td>
</tr>
<tr>
<td>type-specifier &quot;***&quot; identifier</td>
</tr>
<tr>
<td>&quot;void&quot;</td>
</tr>
<tr>
<td><strong>value:</strong></td>
</tr>
<tr>
<td>constant</td>
</tr>
<tr>
<td>identifier</td>
</tr>
<tr>
<td><strong>type-specifier:</strong></td>
</tr>
<tr>
<td>[ &quot;unsigned&quot; ] &quot;int&quot;</td>
</tr>
<tr>
<td>[ &quot;unsigned&quot; ] &quot;hyper&quot;</td>
</tr>
<tr>
<td>&quot;float&quot;</td>
</tr>
<tr>
<td>&quot;double&quot;</td>
</tr>
<tr>
<td>&quot;quadruple&quot;</td>
</tr>
<tr>
<td>&quot;bool&quot;</td>
</tr>
<tr>
<td>enum-type-spec</td>
</tr>
<tr>
<td>struct-type-spec</td>
</tr>
<tr>
<td>union-type-spec</td>
</tr>
<tr>
<td>identifier</td>
</tr>
<tr>
<td><strong>enum-type-spec:</strong></td>
</tr>
<tr>
<td>&quot;enum&quot; enum-body</td>
</tr>
<tr>
<td><strong>enum-body:</strong></td>
</tr>
<tr>
<td>{</td>
</tr>
<tr>
<td>( identifier &quot;=&quot; value )</td>
</tr>
<tr>
<td>( &quot;,&quot; identifier &quot;=&quot; value )*</td>
</tr>
<tr>
<td>}</td>
</tr>
<tr>
<td><strong>struct-type-spec:</strong></td>
</tr>
<tr>
<td>&quot;struct&quot; struct-body</td>
</tr>
<tr>
<td><strong>struct-body:</strong></td>
</tr>
<tr>
<td>{</td>
</tr>
<tr>
<td>( declaration &quot;;&quot; )</td>
</tr>
<tr>
<td>( declaration &quot;;&quot; )*</td>
</tr>
<tr>
<td>}</td>
</tr>
<tr>
<td><strong>union-type-spec:</strong></td>
</tr>
</tbody>
</table>
"union" union-body

union-body:
"switch" "(" declaration ")" ";"
{ "case" value ":" declaration ";" }
{ "case" value ":" declaration ";" }
{ "default" ":" declaration ";" }
"

constant-def:
"const" identifier "=" constant ";"

type-def:
"typedef" declaration ";"
| "enum" identifier enum-body ";"
| "struct" identifier struct-body ";"
| "union" identifier union-body ";"

definition:
type-def
| constant-def

specification:
definition *

Syntax Notes

1. The following are keywords and cannot be used as identifiers:

<table>
<thead>
<tr>
<th>TABLE C-1</th>
<th>XDR Keywords</th>
</tr>
</thead>
<tbody>
<tr>
<td>bool</td>
<td>const</td>
</tr>
<tr>
<td>cas</td>
<td>default</td>
</tr>
<tr>
<td>cha</td>
<td>double</td>
</tr>
</tbody>
</table>

2. Only unsigned constants may be used as size specifications for arrays. If an identifier is used, it must have been declared previously as an unsigned constant in a const definition.

2. Constant and type identifiers within the scope of a specification are in the same name space and must be declared uniquely within this scope.
3. Similarly, variable names must be unique within the scope of `struct` and `union` declarations. Nested `struct` and `union` declarations create new scopes.

4. The discriminant of a `union` must be of a type that evaluates to an integer. That is, `int`, `unsigned int`, `bool`, an `enum` type, or any `typedef` that evaluates to one of these. Also, the case values must be legal discriminant values. Finally, a case value may not be specified more than once within the scope of a `union` declaration.

**XDR Data Description**

Here is a short XDR data description of a file data structure, which might be used to transfer files from one machine to another.

**CODE EXAMPLE C-2  XDR File Data Structure**

```c
const MAXUSERNAME = 32; /* max length of a user name */
const MAXFILELEN = 65535; /* max length of a file */
const MAXNAMELEN = 255; /* max length of a file name */

/* Types of files: */
enum filekind {
    TEXT = 0, /* ascii data */
    DATA = 1, /* raw data */
    EXEC = 2 /* executable */
};

/* File information, per kind of file: */
union filetype switch (filekind kind) {
    case TEXT: void; /* no extra information */
    case DATA:
        string creator<MAXNAMELEN>; /* data creator */
    case EXEC:
        string interpretor<MAXNAMELEN>; /* proginterp*/
};

/* A complete file: */
struct file {
    string filename<MAXNAMELEN>; /* name of file */
    filetype type; /* info about file */
    string owner<MAXUSERNAME>; /* owner of file */
    opaque data<MAXFILELEN>; /* file data */
};
```

Suppose now that there is a user named `john` who wants to store his LISP program `sillyprog` that contains just the data "quit." His file would be encoded as follows:
**TABLE C-2  XDR Data Description Example**

<table>
<thead>
<tr>
<th>Offset</th>
<th>Hex Bytes</th>
<th>ASCII</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>00 00 00 09</td>
<td>....</td>
<td>Length of filename = 9</td>
</tr>
<tr>
<td>4</td>
<td>73 69 6c 6c</td>
<td>sill</td>
<td>Filename characters</td>
</tr>
<tr>
<td>8</td>
<td>79 70 72 6f</td>
<td>ypro</td>
<td>... and more characters ...</td>
</tr>
<tr>
<td>12</td>
<td>67 00 00 00</td>
<td>g...</td>
<td>.. and 3 zero-bytes of fill</td>
</tr>
<tr>
<td>16</td>
<td>00 00 00 02</td>
<td>....</td>
<td>Filekind is EXEC = 2</td>
</tr>
<tr>
<td>20</td>
<td>00 00 00 04</td>
<td>....</td>
<td>Length of interpretor = 4</td>
</tr>
<tr>
<td>24</td>
<td>6c 69 73 70</td>
<td>lisp</td>
<td>Interpretor characters</td>
</tr>
<tr>
<td>28</td>
<td>00 00 00 04</td>
<td>....</td>
<td>Length of owner = 4</td>
</tr>
<tr>
<td>32</td>
<td>6a 6f 68 6e</td>
<td>john</td>
<td>Owner characters</td>
</tr>
<tr>
<td>36</td>
<td>00 00 00 06</td>
<td>....</td>
<td>Length of file data = 6</td>
</tr>
<tr>
<td>40</td>
<td>28 71 75 69</td>
<td>(qu</td>
<td>File data bytes ...</td>
</tr>
<tr>
<td>44</td>
<td>74 29 00 00</td>
<td>t)...</td>
<td>... and 2 zero-bytes of fill</td>
</tr>
</tbody>
</table>

**RPC Language Reference**

The RPC language is an extension of the XDR language. The sole extension is the addition of the program and version types.

For a description of the RPC extensions to the XDR language, see Appendix B.

The RPC language is similar to C. This section describes the syntax of the RPC language, showing a few examples along the way. It also shows how RPC and XDR type definitions get compiled into C type definitions in the output header file.

An RPC language file consists of a series of definitions.

```
definition-list:
  definition;
```
It recognizes six types of definitions.

Enumerations

RPC/XDR enumerations have similar syntax as C enumerations.

```
enum-definition:
    "enum" enum-ident "{" 
    enum-value-list 
    "}"

enum-value-list:
    enum-value 
    enum-value"," enum-value-list

definition:

typedef-definition:

struct-definition

union-definition

program-definition

Definitions are not the same as declarations. No space is allocated by a
definition – only the type definition of a single or series of data elements.
This means that variables still must be declared.

Constants

XDR symbolic constants may be used wherever an integer constant is used.
For example, in array size specifications:

```
const-definition:
    const const-ident = integer
```

The following example defines a constant, DOZEN as equal to 12:

```
const DOZEN = 12; --> #define DOZEN 12
```

Type Definitions

XDR typedefs have the same syntax as C typedefs.
typedef declaration

This example defines an `fname_type` used for declaring file name strings that have a maximum length of 255 characters.

```c
typedef string fname_type<255>;  -->  typedef char *fname_type;
```

Declarations

In XDR, there are four kinds of declarations. These declarations must be a part of a `struct` or a `typedef`; they cannot stand alone:

```c
declaration:
  simple-declaration
  fixed-array-declaration
  variable-array-declaration
  pointer-declaration
```

Simple Declarations

Simple declarations are just like simple C declarations:

```c
simple-declaration:
  type-ident variable-ident
```

Example:

```c
colortype color;  -->  colortype color;
```

Fixed-Length Array Declarations

Fixed-length array declarations are just like C array declarations:

```c
fixed-array-declaration:
  type-ident variable-ident [value]
```

Example:

```c
colortype palette[8];  -->  colortype palette[8];
```

Many programmers confuse variable declarations with type declarations. It is important to note that rpcgen does not support variable declarations. This example is a program that will not compile:

```c
int data[10];
program P {
  version V {
    int PROC(data) = 1;
  } = 1;
} = 0x200000;
```

The example above will not compile because of the variable declaration:

```c
int data[10]
```

Instead, use:

```c
typedef int data[10];
```
or

```c
struct data {int dummy [10]};
```

## Variable-Length Array Declarations

Variable-length array declarations have no explicit syntax in C. The XDR language does have a syntax, using angle brackets:

```c
type-ident variable-ident <value>
type-ident variable-ident < >
```

The maximum size is specified between the angle brackets. The size may be omitted, indicating that the array may be of any size:

```c
int heights<12>; /* at most 12 items */
int widths<>; /* any number of items */
```

Because variable-length arrays have no explicit syntax in C, these declarations are compiled into struct declarations. For example, the heights declaration compiled into the following struct:

```c
struct {
    u_int heights_len; /* # of items in array */
    int *heights_val; /* pointer to array */
} heights;
```

The number of items in the array is stored in the _len component and the pointer to the array is stored in the _val component. The first part of each component name is the same as the name of the declared XDR variable (heights).

## Pointer Declarations

Pointer declarations are made in XDR exactly as they are in C. Address pointers are not really sent over the network; instead, XDR pointers are useful for sending recursive data types such as lists and trees. The type is called “optional-data,” not “pointer,” in XDR language:

```c
type-ident *variable-ident
```

Example:

```c
listitem *next; --> listitem *next;
```

## Structures

An RPC/XDR struct is declared almost exactly like its C counterpart. It looks like the following:

```c
struct-definition:
    struct struct-ident "{*"
    declaration-list
```
The following XDR structure is an example of a two-dimensional coordinate and the C structure that it compiles into:

```
struct coord {
    int x;
    int y;
};
```

```
typedef struct coord coord;
```

The output is identical to the input, except for the added `typedef` at the end of the output. This enables one to use `coord` instead of `struct coord` when declaring items.

### Unions

XDR unions are discriminated unions, and do not look like C unions – they are more similar to Pascal variant records:

```
union-definition:
    "union" union-ident "switch" "("simple declaration")" "{" case-list "}"
```

```
case-list:
    "case" value ":" declaration ";"
    "case" value ":" declaration ";" case-list
    "default" ":" declaration ";"
```

The following is an example of a type returned as the result of a “read data” operation: If there is no error, return a block of data; otherwise, don’t return anything.

```
union read_result switch (int errno) {
    case 0:
        opaque data[1024];
    default:
        void;
};
```

It compiles into the following:

```
struct read_result {
    int errno;
    union {
        char data[1024];
    } read_result_u;
};
typedef struct read_result read_result;
```
Notice that the union component of the output struct has the same name as the type name, except for the trailing \_u.

### Programs

RPC programs are declared using the following syntax:

```
program-definition:
  "program" program-ident "{" 
  version-list 
  "}" "=" value;
version-list:
  version 
  version 
version:
  "version" version-ident "{" 
  procedure-list 
  "}" "=" value;
procedure-list:
  procedure 
  procedure 
procedure:
  type-ident procedure-ident "{" type-ident 
  procedure-list 
  "}" "=" value;
```

When the \(-N\) option is specified, rpcgen also recognizes the following syntax:

```
procedure:
  type-ident procedure-ident "{" type-ident-list 
  procedure-list 
  "}" "=" value;
type-ident-list:
  type-ident 
  type-ident 
type-ident:
  type-ident 
  type-ident 
```

For example:

```c
/*
* time.x: Get or set the time. Time is represented as seconds
* since 0:00, January 1, 1970.
*/
program TIMEPROG {
  version TIMEVERS {
    unsigned int TIMEGET(void) = 1;
    void TIMESET(unsigned) = 2;
  } = 1;
} = 0x20000044;
```

Note that the \texttt{void} argument type means that no argument is passed.

This file compiles into these \#define statements in the output header file:

```c
#define TIMEPROG 0x20000044
#define TIMEVERS 1
#define TIMEGET 1
#define TIMESET 2
```
Special Cases
There are several exceptions to the RPC language rules.

C-style Mode
In the new features section we talked about the features of the C-style mode of rpcgen. These features have implications with regard to the passing of void arguments. No arguments need be passed if their value is void.

Booleans
C has no built-in boolean type. However, the RPC library uses a boolean type called bool_t that is either TRUE or FALSE. Parameters declared as type bool in XDR language are compiled into bool_t in the output header file.

Example:
bool married; --> bool_t married;

Strings
The C language has no built-in string type, but instead uses the null-terminated char * convention. In C, strings are usually treated as null-terminated single-dimensional arrays.

In XDR language, strings are declared using the string keyword, and compiled into type char * in the output header file. The maximum size contained in the angle brackets specifies the maximum number of characters allowed in the strings (not counting the NULL character). The maximum size may be omitted, indicating a string of arbitrary length.

Examples:
string name<32>; --> char *name;
string longname<>; --> char *longname;

Note - NULL strings cannot be passed; however, a zero-length string (that is, just the terminator or NULL byte) can be passed.

Opaque Data
Opaque data is used in XDR to describe untyped data, that is, sequences of arbitrary bytes. It may be declared either as a fixed length or variable length array. Examples:

opaque diskblock[512]; --> char diskblock[512];
opaque filedata<1024>; --> struct {

Voids

In a void declaration, the variable is not named. The declaration is just void and nothing else. Void declarations can only occur in two places: union definitions and program definitions (as the argument or result of a remote procedure, for example no arguments are passed.)
Live RPC Code Examples

This appendix contains copies of the complete live code modules used in the rpcgen and RPC chapters of this book. They are compilable as they are written and will run (unless otherwise noted to be pseudo-code or the like). They are provided for informational purposes only. SunSoft assumes no liability from their use.

Directory Listing Program and Support Routines (rpcgen)

CODE EXAMPLE D–1   rpcgen Program: dir.x

/*
 * dir.x: Remote directory listing
 * protocol
 * This source module is a rpcgen source module
 * used to demonstrate the functions of the rpcgen
 * tool.
 * It is compiled with the rpcgen -h -T switches to
 * generate both the header (.h) file and the
 * accompanying data structures.
 */
const MAXNAMELEN = 255; /*maxlengthofadirectoryentry*/
typedef string nametype<MAXNAMELEN>; /* directory entry */
typedef struct namenode *namelist; /*linkinthelisting*/

(Continued)
struct namenode {
    nametype name; /* name of directory entry */
    namelist next; /* next entry */
};

union readdir_res switch (int errno) {
    case 0:
        namelist list; /* no error: return directory listing */
    default:
        void; /* error occurred: nothing else to return */
};

union readdir_res switch (int errno) {
    case 0:
        namelist list; /* no error: return directory listing */
    default:
        void; /* error occurred: nothing else to return */
};

program DIRPROG {
    version DIRVERS {
        readdir_res
        READDIR(nametype) = 1;
    }
    } = 1;
} = 0x20000076;

CODE EXAMPLE D–2 Remote dir_proc.c

/*
 * dir_proc.c: remote readdir implementation
 */
#include <rpc/rpc.h> /* Always needed */
#include <dirent.h>
#include "dir.h" /* Created by rpcgen */

extern int errno;
extern char *malloc();
extern char *_strdup();

(Continued)
/* ARGSUSED1 */
readdir_res *
readdir_1(dirname, req)
  nametype *dirname;
  struct svc_req *req;
{
  DIR *dirp;
  struct dirent *d;
  namelist nl;
  namelist *nlp;
  static readdir_res res; /* must be static! */

  /* Open directory */
  dirp = opendir(*dirname);
  if (dirp == (DIR *)NULL) {
    res.errno = errno;
    return (&res);
  }
  /* Free previous result */
  xdr_free(xdr_readdir_res, &res);
  /* Collect directory entries. Memory allocated here is freed by
   * xdr_free the next time readdir_1 is called.
   */
  nlp = &res.readdir_res_u.list;
  while (d = readdir(dirp)) {
    nl = *nlp = (namenode *) malloc(sizeof(namenode));
    if (nl == (namenode *) NULL) {
      res.errno = EAGAIN;
      closedir(dirp);
      return (&res);
    }
    nl->name = strdup(d->d_name);
    nlp = &nl->next;
  }
  *nlp = (namelist)NULL;
  /* Return the result */
  res.errno = 0;
  closedir(dirp);
  return (&res);
}
/*
 * rls.c: Remote directory listing client
 */

#include <stdio.h>
#include <rpc/rpc.h> /* always need this */
#include "dir.h" /* generated by rpcgen */

extern int errno;

main(argc, argv)
    int argc;
    char *argv[];
{
    CLIENT *cl;
    char *server;
    char *dir;
    readdir_res *result;
    namelist n;

    if (argc != 3) {
        fprintf(stderr, "usage: %s host directory\n", argv[0]);
        exit(1);
    }
    server = argv[1];
    dir = argv[2];
    cl = clnt_create(server, DIRPROG, DIRVERS, "visible");
    if (cl == (CLIENT *)NULL) {
        clnt_pcreateerror(server);
        exit(1);
    }
    result = readdir_1(&dir, cl);
    if (result == (readdir_res *)NULL) {
        clnt_perror(cl, server);
        exit(1);
    }
    /* Okay, we successfully called the remote procedure. */
    if (result->errno != 0) {
        /*
         * A remote system error occurred. Print error message and
         * die.
         */
        if (result->errno < sys_nerr)
            fprintf(stderr, "%s : %s\n", dir,
                    sys_enlist[result->errno]);
        errno = result->errno;
        perror(dir);
    }
}
exit(1);
}

/* Successfully got a directory listing. Print it out. */
for(nl = result->readdir_res_u.list; nl != NULL; nl = nl->next) {
    printf("%s\n", nl->name);
}
exit(0);

Time Server Program (rpcgen)

CODE EXAMPLE D–4  rpcgen Program: time.x

/*
 * time.x: Remote time protocol
 */
program TIMEPROG {
    version TIMEVERS {
        unsigned int TIMEGET(void) = 1;
    } = 1;
} = 0x20000044;

#ifdef RPC_SVC
%int *
%timeget_1()
{%
 static int thetime;
 %
 % thetime = time(0);
 % return (sthetime);
 %}
#endif
Add Two Numbers Program (\texttt{rpcgen})

\textbf{CODE EXAMPLE D–5} \begin{verbatim}
/* This program contains a procedure to add 2 numbers to demonstrate
 * some of the features of the new \texttt{rpcgen}. Note that \texttt{add()} takes 2
 * arguments in this case. */

program ADDPROG { /* program number */
    version ADDVER { /* version number */
        int add ( int, int ) /* procedure */
        = 1;
    }
}
\end{verbatim}

Spray Packets Program (\texttt{rpcgen})

Refer to the notes section on the \texttt{spray(1M)} man page for information about using this tool.

\textbf{CODE EXAMPLE D–6} \begin{verbatim}
/*
 * Copyright (c) 1987, 1991 by Sun Microsystems, Inc.
 */

/* from spray.x */

#ifndef RPC_HDR
勃,pather "(#)spray.h 1.2 91/09/17 SMI"
#endif

/*
 * Spray a server with packets
 * Useful for testing flakiness of network interfaces
 */

const SPRAYMAX = 8845; /* max amount can spray */
\end{verbatim}

(Continued)
struct spraytimeval {
  unsigned int sec;
  unsigned int usec;
};

struct spraycumul {
  unsigned int counter;
  spraytimeval clock;
};

typedef opaque sprayarr<SPRAYMAX>;

program SPRAYPROG {
  version SPRAYVERS {
    void SPRAYPROC_SPRAY(sprayarr) = 1;
    spraycumul SPRAYPROC_GET(void) = 2;
    void SPRAYPROC_CLEAR(void) = 3;
  } = 1;
} = 100012;
Print Message Program With Remote Version

CODE EXAMPLE D–7 printmesg.c

```c
/* printmsg.c: print a message on the console */
#include <stdio.h>

main(argc, argv)
  int argc;
  char *argv[];
{
  char *message;
  if (argc != 2) {
      fprintf(stderr, "usage: %s <message>\n", argv[0]);
      exit(1);
  }
  message = argv[1];
  if( !printmessage(message) ) {
      fprintf(stderr, "%s: couldn't print your message\n", argv[0]);
      exit(1);
  }
  printf("Message Delivered!\n");
  exit(0);
}

/* Print a message to the console. */

/*
 * Return a boolean indicating whether the message was actually printed.
 */
printmessage(msg)
  char *msg;
{
  FILE *f;
  if = fopen("/dev/console","w");
  if (f == (FILE *)NULL)
      return (0);
  fprintf(f,"\ksen", msg);
  fclose(f);
  return (1);
}
```

CODE EXAMPLE D–8 Remote Version of printmesg.c
main(argc, argv)
    int argc;
    char *argv[];
{
    CLIENT *cl;
    int *result;
    char *server;
    char *message;
    extern int sys_nerr;
    extern char *sys_errlist[];

    if (argc != 3) {
        fprintf(stderr, "usage: %s host messagen", argv[0]);
        exit(1);
    }
    /* Save values of command line arguments */
    server = argv[1];
    message = argv[2];
    /* Create client"handle" used for calling
    * MESSAGEPROG on the server
    * designated on the command line.
    */
    cl = clnt_create(server, MESSAGEPROG, PRINTMESSAGEVERS,
        "visible");
    if (cl == (CLIENT *)NULL) {
        /* Couldn’t establish connection with server.
        * Print error message and die.
        */
        clnt_pcreateerror(server);
        exit(1);
    }
    /* Call the remote procedure "printmessage" on the server */
    result = printmessage_1(&message, cl);
    if (result == (int *)NULL) {
        /* An error occurred while calling the server.
        * Print error message and die.
        */
        clnt_perror(cl, server);
        exit(1);
    }
    /* Okay, we successfully called the remote procedure. */
    if (*result == 0) {
        /* Server was unable to print our message.
        * Print error message and die.
        */
        fprintf(stderr, "%s"
/* The message got printed on the server's console */
printf("Message delivered to %s!\n", server);
exit(0);
}

CODE EXAMPLE D–9  rpcgen Program: msg.x

/* msg.x: Remote message printing protocol */
program MESSAGEPROG {
  version MESSAGEVERS {
    int PRINTMESSAGE(string) = 1;
  } = 1;
} = 0x20000001;

CODE EXAMPLE D–10  mesg_proc.c

/* msg_proc.c: implementation of the remote
 * procedure "printmessage"
 */
#include <stdio.h>
#include <rpc/rpc.h> /* always needed */
#include "msg.h" /* msg.h generated by rpcgen */

/*ARGSUSED1*/
int printmessage_1(msg, req)
  char **msg;
  struct svc_req *req;
{
  static int result; /* must be static! */
  FILE *f;

  f = fopen("/dev/console", "w");
  if (f == (FILE *)NULL) {
    result = 0;
    return (&result);
  }
  fprintf(f, "%sn", *msg);
  fclose(f);
  result = 1;
  return (&result);
}
Batched Code Example

CODE EXAMPLE D–11  Batched Client Program

```c
#include <stdio.h>
#include <rpc/rpc.h>
#include "windows.h"

main(argc, argv)
int argc;
char **argv;
{
    struct timeval total_timeout;
    register CLIENT *client;
    enum clnt_stat clnt_stat;
    char buff[1000], *s = buff;

    if ((client = clnt_create(argv[1], WINDOWPROG, WINDOWVERS,
                               "CIRCUIT_V") == (CLIENT *) NULL) {
        clnt_pcreateerror("clnt_create");
        exit(1);
    }

    timerclear(&total_timeout);
    while (scanf("%s", s) != EOF) {
        clnt_call(client, RENDERSTRING_BATCHED, xdr_wrapstring,
                   &s, xdr_void, (caddr_t) NULL, total_timeout);
    }

    /* Now flush the pipeline */
    total_timeout.tv_sec = 20;
    clnt_stat = clnt_call(client, NULLPROC, xdr_void,
                          (caddr_t) NULL, xdr_void, (caddr_t) NULL,
                          total_timeout);
    if (clnt_stat != RPC_SUCCESS) {
        clnt_perror(client, "rpc");
        exit(1);
    }
    clnt_destroy(client);
    exit(0);
}
```

CODE EXAMPLE D–12  Batched Server Program

```c
#include <stdio.h>
#include <rpc/rpc.h>
#include "windows.h"

void windowdispatch();

main()
```
int num;

num = svc_create(windowdispatch, WINDOWPROG, WINDOWVERS,
 "CIRCUIT_V");
if (num == 0) {
    fprintf(stderr, "can't create an RPC server\n");
    exit(1);
}
svc_run(); /* Never returns */
    fprintf(stderr, "should never reach this point\n");

void
windowdispatch(rqstp, transp)
    struct svc_req *rqstp;
    SVCXPRT *transp;
{
    char *s = NULL;

    switch (rqstp->rq_proc) {
        case NULLPROC:
            if (!svc_sendreply(transp, xdr_void, 0))
                fprintf(stderr, "can't reply to RPC call\n");
            return;
        case RENDERSTRING:
            if (!svc_getargs(transp, xdr_wrapstring, &s)) {
                fprintf(stderr, "can't decode arguments\n");
                /* Tell caller an error occurred */
                svcerr_decode(transp);
            break;
            }/* Code here to render the string s */
            if (!svc_sendreply(transp, xdr_void, (caddr_t) NULL))
                fprintf(stderr, "can't reply to RPC call\n");
            break;
        case RENDERSTRING_BATCHED:
            if (!svc_getargs(transp, xdr_wrapstring, &s)) {
                fprintf(stderr, "can't decode arguments\n");
                /* Be silent in the face of protocol errors */
                break;
            }/* Code here to render string s, but send no reply! */
            break;
        default:
            svcerr_noproc(transp);
            return;
    }/* Now free string allocated while decoding arguments */
    svc_freeargs(transp, xdr_wrapstring, &s);
}
Non-Batched Example

This example is included for reference only. It is a version of the batched client string rendering service, Code Example D–0, written in as a non-batched program.

**CODE EXAMPLE D–13  Unbatched Version of Batched Client**

```c
#include <stdio.h>
#include <rpc/rpc.h>
#include "windows.h"

main(argc, argv)
int argc;  
char **argv;
{
    struct timeval total_timeout;
    register CLIENT *client;
    enum clnt_stat clnt_stat;
    char buf[1000], *s = buf;

    if ((client = clnt_create(argv[1], WINDOWPROG, WINDOWVERS, "CIRCUIT_V")) == (CLIENT *) NULL) {
        clnt_pcreateerror("clnt_create");
        exit(1);
    }
    total_timeout.tv_sec = 20;
    total_timeout.tv_usec = 0;
    while (scanf("%s", s) != EOF) {
        if(clnt_call(client, RENDERSTRING, xdr_wrapstring, &s, xdr_void, (caddr_t) NULL, total_timeout) != RPC_SUCCESS) {
            clnt_perror(client, "rpc");
            exit(1);
        }
    }
    clnt_destroy(client);
    exit(0);
}
```
The portmap Utility

The `rpcbind` utility replaces the `portmap` utility available in previous releases of SunOS. This appendix is included to help you understand the history of port and network address resolution using the `portmap` utility.

SunOS 4.x RPC-based services use `portmap` as a system registration service. It manages a table of correspondences between ports (logical communications channels) and the services registered at them. It provides a standard way for a client to look up the TCP/IP or UDP/IP port number of an RPC program supported by the server.

System Registration Overview

For client programs to find distributed services on a network, they need a way to look up the network addresses of server programs. Network transport (protocol) services do not provide this function. Their task is to provide process-to-process message transfer across a network (that is, a message is sent to a transport-specific network address). A network address is a logical communications channel—by listening on a specific network address, a process receives messages from the network.

The way a process waits on a network address varies from one operating system to the next, but all provide mechanisms by which a process can synchronize its activity with arriving messages. Messages are not sent across networks to receiving processes, but rather to the network address at which receiving processes pick them up. Network addresses are valuable because they allow message receivers to be specified in a way that is independent of the conventions of the receiving operating system. TI-RPC, being transport independent, makes no assumptions about the structure of a network address. It uses a universal addresses. This universal address is specified as a
null-terminated string of characters. Such universal addresses are translated into local transport addresses by routines specific to each transport provider.

The *rpcbind* protocol defines a network service that provides a standard way for clients to look up the network address of any remote program supported by a server. Because it can be implemented on any transport, it provides a single solution to a general problem that works for all clients, all servers, and all networks.

**portmap Protocol**

The *portmap* program maps RPC program and version numbers to transport-specific port numbers. This program makes dynamic binding of remote programs possible.

![Figure E–1  Typical Portmap Sequence (For TCP/IP Only)](image)

Figure E–1 shows the process:
1. The server registers with *portmap*
2. The client gets the server’s port from *portmap*
3. The client calls the server.

The range of reserved port numbers is small and the number of potential remote programs is very large. By running only the port mapper on a well known port, the port numbers of other remote programs can be ascertained by querying the port mapper. In Figure E–1, a, 111, b, and c represent port numbers, where 111 is the assigned portmapper port number.
The port mapper also aids in broadcast RPC. A given RPC program usually has different port number bindings on different machines, so there is no way to directly broadcast to all of these programs. The port mapper, however, does have a fixed port number. So, to broadcast to a given program, the client actually sends its message to the port mapper located at the broadcast address. Each port mapper that receives the broadcast then calls the local service specified by the client. When portmap gets the reply from the local service, it returns the reply to the client. The portmap protocol specification is shown in Code Example E–0.)

**CODE EXAMPLE E–1**  portmap Protocol Specification (in RPC Language)

```c
const PMAP_PORT = 111;  /* portmapper port number */
/
* A mapping of (program, version, protocol) to port number */
struct pmap {
    unsigned int prog;
    unsigned int vers;
    unsigned int prot;
    unsigned int port;
};
/*
* Supported values for the "prot" field */
const IPPROTO_TCP = 6;  /* protocol number for TCP/IP */
const IPPROTO_UDP = 17;  /* protocol number for UDP/IP */
/*
* A list of mappings */
struct pmaplist {
    pmap map;
    pmaplist *next;
};
/*
* Arguments to callit */
struct call_args {
    unsigned int prog;
    unsigned int vers;
    unsigned int proc;
    opaque args<>;
};
/*
* Results of callit */
struct call_result {
    unsigned int port;
    opaque res<>;
};
/*
* Port mapper procedures */
program PMAP_PROG {
(Continued)
```
portmap Operation

portmap currently supports two protocols (UDP/IP and TCP/IP). portmap is contacted by talking to it on assigned port number 111 (SUNRPC (5)) on either of these protocols. The following is a description of each of the portmapper procedures.

PMAPPROC_NULL

This procedure does no work. By convention, procedure zero of any protocol takes no parameters and returns no results.

PMAPPROC_SET

When a program first becomes available on a machine, it registers itself with the local port map program. The program passes its program number prog, version number vers, transport protocol number prot, and the port port on which it receives service requests. The procedure refuses to establish a mapping if one already exists for the specified port and it is bound. If the mapping exists and the port is not bound, portmap unregisters the port and
performs the requested mapping. The procedure returns TRUE if the procedure successfully established the mapping and FALSE otherwise. See also the pmap_set function in the rpc_soc(3N) manpage.

PMAPPROC_UNSET
When a program becomes unavailable, it should unregister itself with the port mapper program on the same machine. The parameters and results have meanings identical to those of PMAPPROC_SET. The protocol and port number fields of the argument are ignored. See also the pmap_unset function in the rpc_soc(3N) manpage.

PMAPPROC_GETPORT
Given a program number prog, version number vers, and transport protocol number prot, this procedure returns the port number on which the program is awaiting call requests. A port value of zeros means the program has not been registered. The port field of the argument is ignored. See also the pmap_getport function in the rpc_soc(3N) manpage.

PMAPPROC_DUMP
This procedure enumerates all entries in the port mapper’s database. The procedure takes no parameters and returns a list of program, version, protocol, and port values. See also the pmap_getmaps function in the rpc_soc(3N) manpage.

PMAPPROC_CALLIT
This procedure allows a caller to call another remote procedure on the same machine without knowing the remote procedure’s port number. It is intended for supporting broadcasts to arbitrary remote programs via the well-known port mapper’s port. The parameters prog, vers, proc, and the bytes of args are the program number, version number, procedure number, and parameters of the remote procedure. See also the pmap_rmtcall function in the rpc_soc(3N) manpage.

This procedure only sends a response if the procedure was successfully executed and is silent (no response) otherwise. It also returns the remote program’s port number, and the bytes of results are the results of the remote procedure.
The port mapper communicates with the remote program using UDP/IP only.

Bibliography

Writing a Port Monitor With the Service Access Facility (SAF)

This appendix gives a brief description of the functions a port monitor must perform to run under the service access facility (SAF) and service access controller (SAC).

- “What Is the SAF” on page 279
- “What Is the SAC” on page 280
- “SAF Files” on page 283
- “The SAC/Port Monitor Interface” on page 284
- “The Port Monitor Administrative Interface” on page 287
- “Configuration Files and Scripts” on page 293
- “Sample Port Monitor Code” on page 298
- “Logic Diagram and Directory Structure” on page 304

What Is the SAF

The SAF generalizes the procedures for service access so that login access on the local system and network access to local services are managed in similar ways. Under the SAF, systems may access services using a variety of port monitors, including ttymon, the listener, and port monitors written expressly for a user’s application.

The manner in which a port monitor observes and manages access ports is specific to the port monitor and not to any component of the SAF. Users may therefore extend their systems by developing and installing their own port
monitors. One of the important features of the SAF is that it can be extended in this way by users.

Relative to the SAF, a service is a process that is started. There are no restrictions on the functions a service may provide.

The SAF consists of a controlling process, the service access controller (SAC), and two administrative levels corresponding to two levels in the supporting directory structure. The top administrative level is concerned with port monitor administration, the lower level with service administration.

From an administrative point of view, the SAF consists of the following components:

- The SAC
- A per-system configuration script
- The SAC administrative file
- The SAC administrative command sacadm
- Port monitors
- Optional per-port monitor configuration scripts
- An administrative file (for each port monitor)
- The administrative command pmadm
- Optional per-service configuration scripts

What Is the SAC

The SAC is the SAF’s controlling process. The SAC is started by init by means of an entry in /etc/inittab. Its function is to maintain the port monitors on the system in the state specified by the system administrator.

The administrative command sacadm is used to tell the SAC to change the state of a port monitor. sacadm can also be used to add or remove a port monitor from SAC supervision and to list information about port monitors known to the SAC.

The SAC’s administrative file contains a unique tag for each port monitor known to the SAC and the path name of the command used to start each port monitor.

The SAC performs three main functions:

- Customizes its own environment
- Starts the appropriate port monitors
- Polls its port monitors and initiates recovery procedures when necessary
Basic Port Monitor Functions

A port monitor is a process that is responsible for monitoring a set of homogeneous, incoming ports on a machine. A port monitor’s major purpose is to detect incoming service requests and to dispatch them appropriately.

A port is an externally seen access point on a system. A port may be an address on a network (TSAP or PSAP), a hardwired terminal line, an incoming phone line, etc. The definition of what constitutes a port is strictly a function of the port monitor itself.

A port monitor performs certain basic functions. Some of these are required to conform to the SAF; others may be specified by the requirements and design of the port monitor itself.

Port monitors have two main functions:

- Managing ports
- Monitoring ports for indications of activity

Port Management

The first function of a port monitor is to manage a port. The actual details of how a port is managed are defined by the person who defines the port monitor. A port monitor is not restricted to handling a single port; it may handle multiple ports simultaneously.

Note - Some examples of port management are setting the line speed on incoming phone connections, binding an appropriate network address, reinitializing the port when the service terminates, outputting a prompt, etc.

Activity Monitoring

The second function of a port monitor is to monitor the port or ports for which it is responsible for indications of activity. Two types of activity may be detected.

1. The first is an indication to the port monitor to take some port monitor-specific action. Pressing the break key to indicate that the line speed should be cycled is an example of a port monitor activity. Not all port monitors need to recognize and respond to the same indications. The indication used to attract the attention of the port monitor is defined by the person who defines the port monitor.

2. The second is an incoming service request. When a service request is received, a port monitor must be able to determine which service is being
requested from the port on which the request is received. Note that the same service may be available on more than one port.

Other Port Monitor Functions
This section briefly describes other port monitor functions.

Restricting Access to the System
A port monitor must be able to restrict access to the system without disturbing services that are still running. In order to do this, a port monitor must maintain two internal states: enabled and disabled. The port monitor starts in the state indicated by the ISTATE environment variable provided by the sac. See “The SAC/Port Monitor Interface” on page 284.

Enabling or disabling a port monitor affects all ports for which the port monitor is responsible. If a port monitor is responsible for a single port, only that port will be affected. If a port monitor is responsible for multiple ports, the entire collection of ports will be affected.

Enabling or disabling a port monitor is a dynamic operation: It causes the port monitor to change its internal state. The effect does not persist across new invocations of the port monitor.

Enabling or disabling an individual port, however, is a static operation: It causes a change to an administrative file. The effect of this change will persist across new invocations of the port monitor.

Creating utmp Entries
Port monitors are responsible for creating utmp entries with the type field set to USER_PROCESS for services they start, if this action has been specified (that is, if −fu was specified in the pmadm line that added the service). These utmp entries may in turn be modified by the service. When the service terminates, the utmp entry must be set to DEAD_PROCESS.

Port Monitor Process IDs and Lock Files
When a port monitor starts, it writes its process id into a file named _pid in the current directory and places an advisory lock on the file.

Changing the Service Environment: Running doconfig
Before invoking the service designated in the port monitor administrative file, _pmtab, a port monitor must arrange for the per-service configuration
script to be run (if one exists by calling the library function doconfig. Because the per-service configuration script may specify the execution of restricted commands, as well as for other security reasons, port monitors are invoked with root permissions. The details of how services are invoked are specified by the person who defines the port monitor.

**Terminating a Port Monitor**

A port monitor must terminate itself gracefully on receipt of the signal `SIGTERM`. The termination sequence is the following:

1. The port monitor enters the stopping state; no further service requests are accepted.
2. Any attempt to re-enable the port monitor will be ignored.
3. The port monitor yields control of all ports for which it is responsible. It must be possible for a new instantiation of the port monitor to start correctly while a previous instantiation is stopping.
4. The advisory lock on the process id file is released. Once this lock is released, the contents of the process id file are undefined and a new invocation of the port monitor may be started.

**SAF Files**

This section briefly covers the files used by the SAF.

**The Port Monitor Administrative File**

A port monitor's current directory contains an administrative file named `_pmtab`. `_pmtab` is maintained by the `pmadm` command in conjunction with a port monitor-specific administrative command.

**Note** - The port monitor administrative command for a listen port monitor is `nlsadmin`; the port monitor administrative command for `ttymon` is `ttyadm`. Any port monitor written by a user must be provided with an administrative command specific to that port monitor to perform similar functions.
Per-Service Configuration Files

A port monitor’s current directory also contains the per-service configuration scripts, if they exist. The names of the per-service configuration scripts correspond to the service tags in the _pmtab file.

Private Port Monitor Files

A port monitor may create private files in the directory /var/saf/tag, where tag is the name of the port monitor. Examples of private files are log files or temporary files.

The SAC/Port Monitor Interface

The sac creates two environment variables for each port monitor it starts:

1. PMTAG
2. ISTATE

This variable is set to a unique port monitor tag by the sac. The port monitor uses this tag to identify itself in response to sac messages. ISTATE is used to indicate to the port monitor what its initial internal state should be. ISTATE is set to “enabled” or “disabled” to indicate that the port monitor is to start in the enabled or disabled state respectively. The sac performs a periodic sanity poll of the port monitors.

The sac communicates with port monitors through FIFOs. A port monitor should open _pmpipe, in the current directory, to receive messages from the sac and ../_sacpipe to send return messages to the sac.

Message Formats

This section describes the messages that may be sent from the sac to a port monitor (sac messages), and from a port monitor to the sac (port monitor messages). These messages are sent through FIFOs and are in the form of C structures. See Code Example F-0.

sac Messages

The format of messages from the sac is defined by the structure sacmsg:
The `sac` may send four types of messages to port monitors. The type of message is indicated by setting the `sc_type` field of the `sacmsg` structure to one of the following:

- `SC_STATUS` status request
- `SC_ENABLE` enable message
- `SC_DISABLE` disable message
- `SC_READDB` message indicating that the port monitor's `_pmtab` file should be read

`sc_size` indicates the size of the optional data part of the message. See “Message Classes” on page 286. For Solaris, `sc_size` should always be set to 0.

A port monitor must respond to every message sent by the `sac`.

### Port Monitor Messages

The format of messages from a port monitor to the `sac` is defined by the structure `pmmsg`:

```c
struct pmmsg {
    char pm_type;          /* type of message */
    unsigned char pm_state; /* current state of port monitor */
    char pm_maxclass;      /* maximum message class this port
                            monitor understands */
    char pm_tag[PMTAGSIZE + 1]; /* port monitor’s tag */
    int pm_size;           /* size of optional data portion */
};
```

Port monitors may send two types of messages to the `sac`. The type of message is indicated by setting the `pm_type` field of the `pmmsg` structure to one of the following:

- `PM_STATUS` state information
- `PM_UNKNOWN` negative acknowledgment

For both types of messages, the `pm_tag` field is set to the port monitor’s tag and the `pm_state` field is set to the port monitor’s current state. Valid states are:

- `PM_STARTING` starting
- `PM_ENABLED` enabled
- `PM_DISABLED` disabled
- `PM_STOPPING` stopping

The current state reflects any changes caused by the last message from the `sac`.

The status message is the normal return message. The negative acknowledgment should be sent only when the message received is not understood.

`pm_size` indicates the size of the optional data part of the message. `pm_maxclass` is used to specify a message class. Both are discussed under "Message Classes" on page 286.
"Message Classes" on page 286. In Solaris, always set pm_maxclass to 1 and sc_size to 0.

Port monitors may never initiate messages; they may only respond to messages that they receive.

Message Classes

The concept of message class has been included to accommodate possible SAF extensions. The messages described above are all class 1 messages. None of these messages contains a variable data portion; all pertinent information is contained in the message header.

If new messages are added to the protocol, they will be defined as new message classes (for example, class 2). The first message the sac sends to a port monitor will always be a class 1 message. Since all port monitors, by definition, understand class 1 messages, the first message the sac sends is guaranteed to be understood. In its response to the sac, the port monitor sets the pm_maxclass field to the maximum message class number for that port monitor. The sac will not send messages to a port monitor from a class with a larger number than the value of pm_maxclass. Requests that require messages of a higher class than the port monitor can understand will fail. For Solaris, always set pm_maxclass to 1.

Note - For any given port monitor, messages of class pm_maxclass and messages of all classes with values lower than pm_maxclass are valid. Thus, if the pm_maxclass field is set to 3, the port monitor understands messages of classes 1, 2, and 3. Port monitors may not generate messages; they may only respond to messages. A port monitor’s response must be of the same class as the originating message.

Since only the sac can generate messages, this protocol will function even if the port monitor is capable of dealing with messages of a higher class than the sac can generate.

pm_size (an element of the pmmsg structure) and sc_size (an element of the sacmsg structure) indicate the size of the optional data part of the message. The format of this part of the message is undefined. Its definition is inherent in the type of message. For Solaris, always set both sc_size and pm_size to 0.
The Port Monitor Administrative Interface

This section discusses the administrative files available under the SAC.

The SAC Administrative File _sactab

The service access controller’s administrative file contains information about all the port monitors for which the SAC is responsible. This file exists on the delivered system. Initially, it is empty except for a single comment line that contains the version number of the SAC. Port monitors are added to the system by making entries in the SAC’s administrative file. These entries should be made using the administrative command sacadm with a -a option. sacadm is also used to remove entries from the SAC’s administrative file.

Each entry in the SAC’s administrative file contains the following information, shown in Table F–1.

<table>
<thead>
<tr>
<th>Fields</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>PMTAG</td>
<td>A unique tag that identifies a particular port monitor. The system administrator is responsible for naming a port monitor. This tag is then used by the SAC to identify the port monitor for all administrative purposes. PMTAG may consist of up to 14 alphanumeric characters.</td>
</tr>
<tr>
<td>PMTYPE</td>
<td>The type of the port monitor. In addition to its unique tag, each port monitor has a type designator. The type designator identifies a group of port monitors that are different invocations of the same entity. ttymon and listen are examples of valid port monitor types. The type designator is used to facilitate the administration of groups of related port monitors. Without a type designator, the system administrator has no way of knowing which port monitor tags correspond to port monitors of the same type. PMTYPE may consist of up to 14 alphanumeric characters.</td>
</tr>
<tr>
<td>FLGS</td>
<td>The flags that are currently defined are: -d When started, do not enable the port monitor. -x Do not start the port monitor. If no flag is specified, the default action is taken. By default a port monitor is started and enabled.</td>
</tr>
</tbody>
</table>
### TABLE F-1  Service Access Controller _sactab File  (continued)

<table>
<thead>
<tr>
<th>Fields</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>RCNT</td>
<td>The number of times a port monitor may fail before being placed in a failed state. Once a port monitor enters the failed state, the SAC will not try to restart it. If a count is not specified when the entry is created, this field is set to 0. A restart count of 0 indicates that the port monitor is not to be restarted when it fails.</td>
</tr>
<tr>
<td>COMMAND</td>
<td>A string representing the command that will start the port monitor. The first component of the string, the command itself, must be a full path name.</td>
</tr>
</tbody>
</table>

### The Port Monitor Administrative File _pmtab

Each port monitor will have two directories for its exclusive use. The current directory will contain files defined by the SAF (_pmtab, _pid) and the per-service configuration scripts, if they exist. The directory `/var/saf/pmtag`, where pmtag is the tag of the port monitor, is available for the port monitor’s private files.

Each port monitor has its own administrative file. The pmadm command should be used to add, remove, or modify service entries in this file. Each time a change is made using pmadm, the corresponding port monitor rereads its administrative file. Each entry in a port monitor’s administrative file defines how the port monitor treats a specific port and what service is to be invoked on that port.

Some fields must be present for all types of port monitors. Each entry must include a service tag to identify the service uniquely and an identity to be assigned to the service when it is started (for example, root).

**Note** - The combination of a service tag and a port monitor tag uniquely define an instance of a service. The same service tag may be used to identify a service under a different port monitor. The record must also contain port monitor specific data (for example, for a ttymon port monitor, this will include the prompt string which is meaningful to ttymon). Each type of port monitor must provide a command that takes the necessary port monitor-specific data as arguments and outputs these data in a form suitable for storage in the file. The ttyadm command does this for ttymon and nlsadmin does it for listen. For a user-defined port monitor, a similar administrative command must also be supplied.

Each service entry in the port monitor administrative file must have the following format and contain the information listed below:
SVCTAG is a unique tag that identifies a service. This tag is unique only for
the port monitor through which the service is available. Other port monitors
may offer the same or other services with the same tag. A service requires
both a port monitor tag and a service tag to identify it uniquely.

SVCTAG may consist of up to 14 alphanumeric characters. The service entries
are defined in Table F–2.

### TABLE F–2  SVCTAG Service Entries

<table>
<thead>
<tr>
<th>Service Entries</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>FLGS</td>
<td>Flags with the following meanings may currently be included in this field:</td>
</tr>
<tr>
<td></td>
<td>−x Do not enable this port. By default the port is enabled.</td>
</tr>
<tr>
<td></td>
<td>−u Create a <code>utmp</code> entry for this service. By default no <code>utmp</code> entry is created for the service.</td>
</tr>
<tr>
<td>ID</td>
<td>The identity under which the service is to be started. The identity has the form of a login name as it appears in <code>/etc/passwd</code>.</td>
</tr>
<tr>
<td>PMSPECIFIC</td>
<td>Examples of port monitor information are addresses, the name of a process to execute, or the name of a STREAMS pipe to pass a connection through. This information will vary to meet the needs of each different type of port monitor.</td>
</tr>
<tr>
<td>COMMENT</td>
<td>A comment associated with the service entry.</td>
</tr>
</tbody>
</table>

**Note** - Port monitors may ignore the −u flag if creating a `utmp` entry for the service is not appropriate to the manner in which the service is to be invoked. Some services may not start properly unless `utmp` entries have been created for them (for example, `login`).

Each port monitor administrative file must contain one special comment of the form:

```
# VERSION=value
```

where `value` is an integer that represents the port monitor’s version number. The version number defines the format of the port monitor administrative file. This comment line is created automatically when a port monitor is added to the system. It appears on a line by itself, before the service entries.
The SAC Administrative Command *sacadm*

*sacadm* is the administrative command for the upper level of the *SAF* hierarchy, that is, for port monitor administration (see the *sacadm*(1M) manpage). Under the *SAF*, port monitors are administered by using the *sacadm* command to make changes in the *SAC*'s administrative file. *sacadm* performs the following functions:

- Prints requested port monitor information from the *SAC* administrative file
- Adds or removes a port monitor
- Enables or disables a port monitor
- Starts or stops a port monitor
- Installs or replaces a per-system configuration script
- Installs or replaces a per-port monitor configuration script
- Asks the *SAC* to reread its administrative file

The Port Monitor Administrative Command *pmadm*

*pmadm* is the administrative command for the lower level of the *SAF* hierarchy, that is, for service administration (see the *pmadm*(1M) manpage). A port may have only one service associated with it although the same service may be available through more than one port. *pmadm* performs the following functions:

- Prints service status information from the port monitor’s administrative file
- Adds or removes a service
- Enables or disables a service
- Installs or replaces a per-service configuration script

Note that in order to identify an instance of a service uniquely, the *pmadm* command must identify both the service (−s) and the port monitor or port monitors through which the service is available (−p or −t).

Monitor-Specific Administrative Command

In the previous section, two pieces of information included in the _pmtab_ file were described: the port monitor’s version number and the port monitor part of the service entries in the port monitor’s _pmtab_ file. When a new port monitor is added, the version number must be known so that the
_pmtab file can be correctly initialized. When a new service is added, the port monitor part of the _pmtab entry must be formatted correctly.

Each port monitor must have an administrative command to perform these two tasks. The person who defines the port monitor must also define such an administrative command and its input options. When the command is invoked with these options, the information required for the port monitor part of the service entry must be correctly formatted for inclusion in the port monitor’s _pmtab file and must be written to the standard output. To request the version number the command must be invoked with a −V option; when it is invoked in this way, the port monitor’s current version number must be written to the standard output.

If the command fails for any reason during the execution of either of these tasks, no data should be written to standard output.

The Port Monitor/Service Interface

The interface between a port monitor and a service is determined solely by the service. Two mechanisms for invoking a service are presented here as examples.

New Service Invocations

The first interface is for services that are started anew with each request. This interface requires the port monitor to first fork a child process. The child will eventually become the designated service by performing an exec. Before the exec happens, the port monitor may take some port monitor-specific action; however, one action that must occur is the interpretation of the per-service configuration script, if one is present. This is done by calling the library routine doconfig.

Standing Service Invocations

The second interface is for invocations of services that are actively running. To use this interface, a service must have one end of a stream pipe open and be prepared to receive connections through it.

Port Monitor Requirements

To implement a port monitor, several generic requirements must be met. This section summarizes these requirements. In addition to the port monitor itself, an administrative command must be supplied.
Initial Environment
When a port monitor is started, it expects an initial execution environment in which:
- It has no file descriptors open
- It cannot be a process group leader
- It has an entry in /etc/utmp of type LOGIN_PROCESS
- An environment variable, ISTATE, is set to “enabled” or “disabled” to indicate the port monitor’s correct initial state
- An environment variable, PMTAG, is set to the port monitor’s assigned tag
- The directory that contains the port monitor’s administrative files is its current directory
- The port monitor is able to create private files in the directory /var/saf/tag, where tag is the port monitor’s tag
- The port monitor is running with user id 0 (root)

Important Files
Relative to its current directory, the following key files exist for a port monitor.

<table>
<thead>
<tr>
<th>File</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>_config</td>
<td>The port monitor’s configuration script. The port monitor configuration script is run by the SAC. The SAC is started by init as a result of an entry in /etc/inittab that calls sac.</td>
</tr>
<tr>
<td>_pid</td>
<td>The file into which the port monitor writes its process id.</td>
</tr>
<tr>
<td>_pmtab</td>
<td>The port monitor’s administrative file. This file contains information about the ports and services for which the port monitor is responsible.</td>
</tr>
<tr>
<td>_pmpipe</td>
<td>The FIFO through which the port monitor will receive messages from sac.</td>
</tr>
<tr>
<td>svctag</td>
<td>The per-service configuration script for the service with the tag svctag.</td>
</tr>
<tr>
<td>../_sacpipe</td>
<td>The FIFO through which the port monitor will send messages to sac.</td>
</tr>
</tbody>
</table>
Port Monitor Responsibilities

A port monitor is responsible for performing the following tasks in addition to its port monitor function:

- Write its process id into the file _pid and place an advisory lock on the file
- Terminate gracefully on receipt of the signal SIGTERM.
- Follow the protocol for message exchange with sac

A port monitor must perform the following tasks during service invocation:

- Create a utmp entry if the requested service has the “−u” flag set in _pmtab

**Note** - Port monitors may ignore this flag if creating a utmp entry for the service does not make sense because of the manner in which the service is to be invoked. On the other hand, some services may not start properly unless utmp entries have been created for them.

- Interpret the per-service configuration script for the requested service, if it exists, by calling the doconfig library routine

Configuration Files and Scripts

Interpreting Configuration Scripts With doconfig

The library routine doconfig, defined in libnsl.so, interprets the configuration scripts contained in the files /etc/saf/_sysconfig (the per-system configuration file), and /etc/saf/pmtag/_config (per-port monitor configuration files); and in /etc/saf/pmtag/svctag (per-service configuration files). Its syntax is:

```c
#include <sac.h>
int doconfig (int fd, char *script, long rflag);
```

script is the name of the configuration script; fd is a file descriptor that designates the stream to which stream manipulation operations are to be applied; rflag is a bitmask that indicates the mode in which script is to be interpreted. rflag may take two values, NORUN and NOASSIGN, which may be or’d. If rflag is zero, all commands in the configuration script are eligible to be interpreted. If rflag has the NOASSIGN bit set, the assign command is considered illegal and will generate an error return. If rflag has the NORUN bit
set, the run and runwait commands are considered illegal and will generate error returns.

If a command in the script fails, the interpretation of the script ceases at that point and a positive integer is returned; this number indicates which line in the script failed. If a system error occurs, a value of -1 is returned.

If a script fails, the process whose environment was being established should not be started.

In the example, doconfig is used to interpret a per-service configuration script.

```c
... if ((i = doconfig (fd, svctag, 0)) != 0){
    error ("doconfig failed online %d of script %s", i, svctag);
}
```

The Per-System Configuration File

The per-system configuration file, `/etc/saf/_sysconfig`, is delivered empty. It may be used to customize the environment for all services on the system by writing a command script in the interpreted language described in this chapter and on the doconfig(3N) manpage. When the SAC is started, it calls the doconfig function to interpret the per-system configuration script. The SAC is started when the system enters multiuser mode.

Per-Port Monitor Configuration Files

Per-port monitor configuration scripts (`/etc/saf/pmtag/_config`) are optional. They allow the user to customize the environment for any given port monitor and for the services that are available through the ports for which that port monitor is responsible. Per-port monitor configuration scripts are written in the same language used for per-system configuration scripts.

The per-port monitor configuration script is interpreted when the port monitor is started. The port monitor is started by the SAC after the SAC has itself been started and after it has run its own configuration script, `/etc/saf/_sysconfig`

The per-port monitor configuration script may override defaults provided by the per-system configuration script.
Per-Service Configuration Files

Per-service configuration files allow the user to customize the environment for a specific service. For example, a service may require special privileges that are not available to the general user. Using the language described in the doconfig(3N) manpage, you can write a script that will grant or limit such special privileges to a particular service offered through a particular port monitor.

The per-service configuration may override defaults provided by higher-level configuration scripts. For example, the per-service configuration script may specify a set of STREAMS modules other than the default set.

The Configuration Language

The language in which configuration scripts are written consists of a sequence of commands, each of which is interpreted separately. The following reserved keywords are defined: assign, push, pop, runwait, and run. The comment character is #. Blank lines are not significant. No line in a command script may exceed 1024 characters.

assign variable=value

Used to define environment variables. variable is the name of the environment variable and value is the value to be assigned to it. The value assigned must be a string constant; no form of parameter substitution is available. value may be quoted. The quoting rules are those used by the shell for defining environment variables. assign will fail if space cannot be allocated for the new variable or if any part of the specification is invalid.

push module1[, module2, module3, ...]

Used to push STREAMS modules onto the stream designated by fd. See the doconfig(3N) manpage. module1 is the name of the first module to be pushed, module2 is the name of the second module to be pushed, and so on. The command will fail if any of the named modules cannot be pushed. If a module cannot be pushed, the subsequent modules on the same command line will be ignored and modules that have already been pushed will be popped.

pop [module]

Used to pop STREAMS modules off the designated stream. If pop is invoked with no arguments, the top module on the stream is popped. If an argument is given, modules will be popped one at a time until the named module is at
the top of the stream. If the named module is not on the designated stream, the stream is left as it was and the command fails. If \textit{module} is the special keyword \texttt{ALL}, then all modules on the stream will be popped. Note that only modules above the topmost driver are affected.

\texttt{runwait command}

The \texttt{runwait} command runs a command and waits for it to complete. \texttt{command} is the path name of the command to be run. The command is run with \texttt{/bin/sh -c} prepended to it; shell scripts may thus be executed from configuration scripts. The \texttt{runwait} command will fail if \texttt{command} cannot be found or cannot be executed, or if \texttt{command} exits with a nonzero status.

\texttt{run command}

The \texttt{run} command is identical to \texttt{runwait} except that it does not wait for \texttt{command} to complete. \texttt{command} is the path name of the command to be run. \texttt{run} will not fail unless it is unable to create a child process to execute the command.

Although they are syntactically indistinguishable, some of the commands available to \texttt{run} and \texttt{runwait} are interpreter built-in commands. Interpreter built-ins are used when it is necessary to alter the state of a process within the context of that process. The \texttt{doconfig} interpreter built-in commands are similar to the shell special commands and, like these, they do not spawn another process for execution. See the \texttt{sh(1)manpage}. The initial set of built-in commands is:

\texttt{cd ulimit umask}

\textbf{Printing, Installing, and Replacing Configuration Scripts}

This section describes the form of the \texttt{SAC} and port monitor administrative commands used to install the three types of configuration scripts. Per-system and per-port monitor configuration scripts are administered using the \texttt{sacadm} command. Per-service configuration scripts are administered using the \texttt{pmadm} command.

\textbf{Per-System Configuration Scripts}

\texttt{sacadm \textasciitilde G [ \textasciitilde z script ]}
The −G option is used to print or replace the per-system configuration script. The −G option by itself prints the per-system configuration script. The −G option in combination with a −z option replaces /etc/saf/_sysconfig with the contents of the file script. Other combinations of options with a −G option are invalid.

Sample Per-System Configuration Script

The _sysconfig file in the example sets the time zone variable, TZ.

assign TZ=EST5EDT # set TZ
runwait echo SAC is starting > /dev/console

Per-Port Monitor Configuration Scripts

sacadm −g −p pmtag [ −z script ]

The −g option is used to print, install, or replace the per-port monitor configuration script. A −g option requires a −p option. The −g option with only a −p option prints the per-port monitor configuration script for port monitor pmtag. The −g option with a −p option and a −z option installs the file script as the per-port monitor configuration script for port monitor pmtag, or, if /etc/saf/pmtag/_config exists, it replaces _config with the contents of script. Other combinations of options with −g are invalid.

Sample Per-Port Monitor Configuration Script

In the hypothetical _config file in the figure, the command /usr/bin/daemon is assumed to start a daemon process that builds and holds together a STREAMS multiplexor. By installing this configuration script, the command can be executed just before starting the port monitor that requires it.

# build a STREAMS multiplexor
run /usr/bin/daemon
runwait echo $PMTAG is starting > /dev/console

Per-Service Configuration Scripts

pmadm −g −p pmtag −s svctag [ −z script ]
pmadm −g −s svctag −t type −z script

Per-service configuration scripts are interpreted by the port monitor before the service is invoked.
Note - The SAC interprets both its own configuration file, \_sysconfig, and the port monitor configuration files. Only the per-service configuration files are interpreted by the port monitors.

The \texttt{-g} option is used to print, install, or replace a per-service configuration script. The \texttt{-g} option with a \texttt{-p} option and a \texttt{-s} option prints the per-service configuration script for service \texttt{svctag} available through port monitor \texttt{pmtag}. The \texttt{-g} option with a \texttt{-p} option, a \texttt{-s} option, and a \texttt{-z} option installs the per-service configuration script contained in the file \texttt{script} as the per-service configuration script for service \texttt{svctag} available through port monitor \texttt{pmtag}. The \texttt{-g} option with a \texttt{-s} option, a \texttt{-t} option, and a \texttt{-z} option installs the file \texttt{script} as the per-service configuration script for service \texttt{svctag} available through any port monitor of type \texttt{type}. Other combinations of options with \texttt{-g} are invalid.

Sample Per-Service Configuration Script

The following per-service configuration script does two things: It specifies the maximum file size for files created by a process by setting the process’s \texttt{ulimit} to \texttt{4096}. It also specifies the protection mask to be applied to files created by the process by setting \texttt{umask} to \texttt{077}.

\begin{verbatim}
runwait ulimit 4096
runwait umask 077
\end{verbatim}

Sample Port Monitor Code

Code Example F-0 shows an example of a “null” port monitor that simply responds to messages from the SAC.

CODE EXAMPLE F-1 Sample Port Monitor

\begin{verbatim}
#include <stdlib.h>
#include <stdio.h>
#include <unistd.h>
#include <fcntl.h>
#include <signal.h>
#include <sac.h>

char Scratch[BUFSIZ]; /* scratch buffer */
\end{verbatim}
(Continued)
main(argc, argv)
    int argc;
    char *argv[];
    {
        char *istate;
        strcpy(Tag, getenv("PMTAG"));
        /* open up a log file in port monitor’s private directory */
        sprintf(Scratch, "/var/saf/%s/log", Tag);
        Fp = fopen(Scratch, "a+");
        if (Fp == (FILE *)NULL)
            exit(1);
        log(Fp, "starting");
        /* retrieve initial state (either "enabled" or "disabled") and set 
         * State accordingly */
        istate = getenv("ISTATE");
        sprintf(Scratch, "ISTATE is %s", istate);
        log(Fp, Scratch);
        if (!strcmp(istate, "enabled"))
            State = PM_ENABLED;
        else if (!strcmp(istate, "disabled"))
            State = PM_DISABLED;
        else {
            log(Fp, "invalid initial state");
            exit(1);
        }
        sprintf(Scratch, "PMTAG is %s", Tag);
        log(Fp, Scratch);
        /* set up pid file and lock it to indicate that we are active */
        Tfp = fopen("_pid", "w");
        if (Tfp == (FILE *)NULL) {
            log(Fp, "couldn’t open pid file");
            exit(1);
        }
        if (lockf(fileno(Tfp), F_TEST, 0) < 0) {
            log(Fp, "pid file already locked");
            exit(1);
        }
        fprintf(Tfp, "%d", getpid());
        fflush(Tfp);
        log(Fp, "locking file");
if (lockf(fileno(Tfp), F_LOCK, 0) < 0) {
    log(Fp, "lock failed");
    exit(1);
}

/*
 * handle poll messages from the sac ... this function never returns
 */
handlepoll();
pause();
fclose(Tfp);
fclose(Fp);
}

handlepoll()
{
    int pfd; /* file descriptor for incoming pipe */
    int sfd; /* file descriptor for outgoing pipe */
    struct sacmsg sacmsg; /* incoming message */
    struct pmmsg pmmsg; /* outgoing message */
    /*
    * open pipe for incoming messages from the sac
    */
    pfd = open("_pmpipe", O_RDONLY|O_NONBLOCK);
    if (pfd < 0) {
        log(Fp, "_pmpipe open failed");
        exit(1);
    }
    /*
    * open pipe for outgoing messages to the sac
    */
    sfd = open("../_sacpipe", O_WRONLY);
    if (sfd < 0) {
        log(Fp, "_sacpipe open failed");
        exit(1);
    }
    /*
    * start to build a return message; we only support class 1 messages
    */
    strcpy(pmmsg.pm_tag, Tag);
    pmmsg.pm_size = 0;
    pmmsg.pm_maxclass = 1;
    /*
    * keep responding to messages from the sac
    */
    for (;;) {
        if (read(pfd, &sacmsg, sizeof(sacmsg)) != sizeof(sacmsg)) {
            log(Fp, "_pmpipe read failed");
            exit(1);
        }
        /*
        * determine the message type and respond appropriately
        */
switch (sacmsg.sc_type) {
    case SC_STATUS:
        log(Fp, "Got SC_STATUS message");
        pmmsg.pm_type = PM_STATUS;
        pmmsg.pm_state = State;
        break;
    case SC_ENABLE:
        /*note internal state change below*/
        log(Fp, "Got SC_ENABLE message");
        pmmsg.pm_type = PM_STATUS;
        State = PM_ENABLED;
        pmmsg.pm_state = State;
        break;
    case SC_DISABLE:
        /*note internal state change below*/
        log(Fp, "Got SC_DISABLE message");
        pmmsg.pm_type = PM_STATUS;
        State = PM_DISABLED;
        pmmsg.pm_state = State;
        break;
    case SC_READDB:
        /*
         * if this were a fully functional port monitor it
         * would read _pmtab here and take appropriate action
         */
        log(Fp, "Got SC_READDB message");
        pmmsg.pm_type = PM_STATUS;
        pmmsg.pm_state = State;
        break;
    default:
        sprintf(Scratch, "Got unknown message <%d>",
                 sacmsg.sc_type);
        log(Fp, Scratch);
        pmmsg.pm_type = PM_UNKNOWN;
        pmmsg.pm_state = State;
        break;
}

/*
 * send back a response to the poll
 * indicating current state
 */
if (write(sfd, &pmmsg, sizeof(pmmsg)) != sizeof(pmmsg))
    log(Fp, "sanity response failed");
}

/*
 * general logging function
 */
log(fp, msg)
FILE *fp;
char *msg;
Code Example F–0 shows the sac.h header file.

**CODE EXAMPLE F–2  sac.h Header File**

```c
/* length in bytes of a utmp id */
#define IDLEN 4
/* wild character for utmp ids */
#define SC_WILDC 0xff
/* max len in bytes for port monitor tag */
#define PMTAGSIZE 14
/*
 * values for rflag in doconfig()
 */
/* don’t allow assign operations */
#define NOASSIGN 0x1
/* don’t allow run or runwait operations */
#define NORUN 0x2
/*
 * message to SAC (header only). This header is forever fixed. The
 * size field (pm_size) defines the size of the data portion of
 * the
 * message, which follows the header. The form of this optional
 * data
 * portion is defined strictly by the message type (pm_type).
 */
struct pmmsg {
    char pm_type;        /* type of message */
    unchar pm_state;     /* current state of pm */
    char pm_maxclass;    /* max message class this port
                          monitor
                          understands */
    char pm_tag[PMTAGSIZE + 1]; /* pm’s tag */
    int pm_size;         /* size of opt data portion */
};
/*
 * pm_type values
 */
#define PM_STATUS 1  /* status response */
#define PM_UNKNOWN 2 /* unknown message was received */
/*
 * pm_state values
 */
/*
 * Class 1 responses
 */
```

(Continued)
# define PM_STARTING 1 /* monitor in starting state */
# define PM_ENABLED 2 /* monitor in enabled state */
# define PM_DISABLED 3 /* monitor in disabled state */
# define PM_STOPPING 4 /* monitor in stopping state */
/*
 * message to port monitor
 */
struct sacramsg {
    int sc_size; /* size of optional data portion */
    char sc_type; /* type of message */
};
/*
 * sc_type values
 * These represent commands that the SAC sends to a port monitor.
 * These commands are divided into "classes" for extensibility.
 * Each subsequent "class" is a superset of the previous "classes" plus
 * the new commands defined within that "class". The header for all
 * commands is identical; however, a command may be defined such that
 * an optional data portion may be sent in addition to the header.
 * The format of this optional data piece is self-defining based on
 * the command. Important note: the first message sent by the SAC
 * will always be a class 1 message. The port monitor response
 * indicates the maximum class that it is able to understand. Another note
 * is that port monitors should only respond to a message with an
 * equivalent class response (i.e. a class 1 command causes a
 * class 1 response).
 */
/*
 * Class 1 commands (currently, there are only class 1 commands)
 */
# define SC_STATUS 1 /* status request */
# define SC_ENABLE 2 /* enable request */
# define SC_DISABLE 3 /* disable request */
# define SC_READDB 4 /* read pmtab request */
/*
 * 'errno' values for Saferrno, note that Saferrno is used by both
 * pmadm and sacadm and these values are shared between them
 */
# define E_BADARGS 1 /* bad args/ill-formed cmd line */
# define E_NOPRIV 2 /* user not priv for operation */
# define E_SAFERR 3 /* generic SAF error */
# define E_SYSERR 4 /* system error */
# define E_NOEXIST 5 /* invalid specification */
# define E_DUP 6 /* entry already exists */
# define E_PMRUN 7 /* port monitor is running */
Logic Diagram and Directory Structure

Figure F–2 is a logical diagram of the SAF. It illustrates how a single service access controller may spawn a number of port monitors on a per-system
basis. This means that several monitors may be running concurrently, providing for the simultaneous operation of several different protocols.

![SAF Logical Framework](image)

*Figure F–1*  SAF Logical Framework
"/etc/saf/_sysconfig" on page 306 is the corresponding directory structure diagram. Following the diagram is a description of the files and directories.

Figure F–2  SAF Directory Structure

/etc/saf/_sysconfig
The per-system configuration script.

/etc/saf/_sactab
The SAC’s administrative file. Contains information about the port monitors for which the SAC is responsible.

/etc/saf/pmtag
The home directory for port monitor pmtag.

/etc/saf/pmtag/_config
The per-port monitor configuration script for port monitor pmtag.
/etc/saf/pmtag/_pmtab
Port monitor pmtag’s administrative file. Contains information about the services for which pmtag is responsible.

/etc/saf/pmtag/svctag
The file in which the per-service configuration script for service svctag (available through port monitor pmtag) is placed.

/etc/saf/pmtag/_pid
The file in which a port monitor writes its process id in the current directory and places an advisory lock on the file.

/etc/saf/pmtag/_pmpipe
The file in which the port monitor receives messages from the sac and ../_sacpipe and sends return messages to the sac.

/var/saf/_log
The SAC’s log file.

/var/saf/pmtag
The directory for files created by port monitor pmtag, for example its log file.
Glossary

RPC Programming Terms

The following terms define the RPC concepts used throughout this manual.

**client**
A program or system that uses the services of a remote program or system.

**client handle**
A client process data structure that represents the binding of the client to a particular server's RPC program.

**connection-oriented transport**
See stream transport.

**connectionless transport**
See datagram transport.

**datagram transport**
Datagram transports have less overhead than connection-oriented transports but are considered less reliable and data transmissions are limited by buffer size.

**deserialize**
Convert data from XDR format to a machine-specific representation.

**handle**
An abstraction used by the service libraries to refer to a file or a file-like object such as a socket.

**host**
A computer (mainframe, mini, server, workstation, or personal computer) connected to a network.

**MT hot**
An interface is multithreaded hot if the library or call automatically creates threads.

**MT safe**
An interface is multithreaded safe if it can be called in a threaded environment. An MT-safe interface may be invoked concurrently for multiple threads.

**network client**
Usually client. A process that makes remote procedure calls to services.
<table>
<thead>
<tr>
<th>Term</th>
<th>Definition</th>
</tr>
</thead>
<tbody>
<tr>
<td>network server</td>
<td>Usually server. A process that performs a network service. A server may support more than one version of a remote program to be forward compatible with changing protocols.</td>
</tr>
<tr>
<td>network service</td>
<td>A collection of one or more remote service programs.</td>
</tr>
<tr>
<td>ping</td>
<td>The ping service is used to verify activity on a remote system.</td>
</tr>
<tr>
<td>remote program</td>
<td>A program that implements one or more remote procedures.</td>
</tr>
<tr>
<td>RPC Language (RPCL)</td>
<td>A C-like programming language translated by the rpcgen compiler. RPCL is a superset of XDR Language.</td>
</tr>
<tr>
<td>RPC library</td>
<td>libnsl, specified to the link editor at compile time. Also known as the RPC package.</td>
</tr>
<tr>
<td>RPC protocol</td>
<td>The message-passing protocol that is the basis of the RPC package.</td>
</tr>
<tr>
<td>RPC/XDR</td>
<td>See RPC Language.</td>
</tr>
<tr>
<td>serialize</td>
<td>Converting data from a machine representation to XDR format.</td>
</tr>
<tr>
<td>server</td>
<td>A process that provides remote service to clients.</td>
</tr>
<tr>
<td>stream transport</td>
<td>Stream transport is considered reliable. It supports byte-stream deliveries of unlimited data size.</td>
</tr>
<tr>
<td>transport</td>
<td>The fourth layer of the Open Systems Interconnection (OSI) Reference Model.</td>
</tr>
<tr>
<td>transport handle</td>
<td>An abstraction used by the RPC libraries to refer to the transport’s data structures.</td>
</tr>
<tr>
<td>TI-RPC</td>
<td>Transport-independent RPC. The version of RPC supported in SunOS 5.x.</td>
</tr>
<tr>
<td>TS-RPC</td>
<td>Transport-specific RPC. The version of RPC supported in SunOS 4.x. TS-RPC is also supported in SunOS 5.x.</td>
</tr>
<tr>
<td>universal address</td>
<td>A machine-independent representation of a transport address.</td>
</tr>
<tr>
<td>virtual circuit</td>
<td>See streamtransport.</td>
</tr>
<tr>
<td>transport</td>
<td></td>
</tr>
<tr>
<td>XDR Language</td>
<td>A data description language and data representation protocol.</td>
</tr>
</tbody>
</table>

% preprocessing directive, 35
& server invocation and, 30
_1 suffix, 61

A
access control
    authentication vs., 94
port monitors and, 282
add.x source file, 38, 40, 44, 47, 48
adding
    address registrations, 19
NIS+ database entries, 146
NIS+ group members, 145
NIS+ objects to namespace, 144
NIS+ table entry objects, 144, 155, 156
port monitor services, 288
port monitors, 280
two numbers, 264
ADDPROG program, 264
addresses
    information reporting for, 20
lookup services, 15, 19, 20
management functions, 132
name-to-address translation
        routines, 18, 19, 129
network, 273, 274
overview, 273, 274
passing arguments as, 26, 27, 61
passing server’s address to client, 73
passing user’s bind address, 76
transport (netbuf), 19, 229, 230
universal, 18, 228–230, 274, 310
unregistering
    current vs. previous release, 131
portmap routine, 277
rpcbind routine, 19, 225, 228
rpcinfo routine, 20
addresses, see portmap routine
    mapping RPC services to,
        registering,
    ah_cred field, 80
    ah_key field, 96
    ah_verf field, 80
allocating memory, see memory\x0d
ampersand (& server invocation and, 30
ANSI C standards
    rpcgen tool and, 23, 37, 48
application programming interface (API)
    NIS+, 143, 148
applications
    porting from TS-RPC to TI-RPC, 128
arguments (remote procedures)
        overview, 11
    passing arbitrary data types, 62, 65
    passing by address, 26, 27, 61
    passing by value, 38, 41
    passing open TLI file descriptors, 73, 76
    passing server’s address to client, 73
    passing user’s bind address, 76
void, 221
arrays, xv
    converting to XDR format, 64, 126, 128
declarations
  RPC language, 218, 219
  XDR language, 242, 243, 249
  XDR code examples, 175, 178–180
 assign configuration-script keyword, 295
 asynchronous mode, 84, 86
 AUTH_BADCRED error, 211, 212
 AUTH_BADVERF error, 211
 AUTH_DES authentication, 95, 96, 205, 210
   common key, 209, 210
   conversation key, 206, 208, 210
   credentials, 95, 208
   described, 92, 95, 96, 205, 206
   Diffie-Hellman encryption, 96, 209, 210
   errors, 207
   handle, 95, 96
   -Hellman encryption, 206
   nicknames, 207–209
   protocol in XDR language, 207, 210
   server, 96
   time synchronization, 96, 207
   verifiers, 206, 207, 209
 auth_destroy routine, 92
 AUTH_KERB authentication, 97, 99, 210, 214
   credentials, 97, 98, 211, 213
   described, 92, 97, 99, 210
   encryption, 98, 99
   errors, 211, 212
   NFS and, 211, 212
   nicknames, 98, 211–213
   protocol in XDR language, 212, 214
   time synchronization, 97, 213
   verifiers, 98, 99, 212–214
 AUTH_NONE authentication, 92, 93, 203
 AUTH_REJECTEDCRED error, 205, 211
 AUTH_REJECTEDVERF error, 212
 AUTH_SHORT authentication, 92
 AUTH_SHORT verifier, 204, 205
 AUTH_SYS authentication, 92–94, 203, 205
 AUTH_SYS) authentication, 93
 AUTH_TIMEEXPIRE error, 212
 AUTH_TOOWEAK error, 212

AUTH_UNIX (AUTH_SYS)
  authentication, 92, 94, 203, 205
 authdes_create routine, 133
 authdes_seccreate routine, 95, 133
 authentication, 91, 99, 203, 214
   access control vs., 94
   allocating authentication
     numbers, 203
   AUTH_DES, 92, 95, 96, 205, 210
   AUTH_KERB, 92, 97, 99, 210, 214
   AUTH_NONE, 92, 93, 203
   AUTH_SHORT, 92, 204, 205
   AUTH_SYS (AUTH_UNIX), 92–94, 203, 205
 credentials
   AUTH_DES, 95, 208
   AUTH_KERB, 97, 98, 211, 213
   described, 197
   window (lifetime) of, 95, 97, 207, 208, 213
 current vs. previous release, 133
 destroying an, 92
 errors
   AUTH_DES, 207
   AUTH_KERB, 211, 212
   AUTH_SYS, 94, 205
 handles, 79, 95, 96
 low-level data structures and, 79
 methods supported, 92, 104
 nicknames
   AUTH_DES, 207–209
   AUTH_KERB, 98, 211–213
 NIS+, 141
 overview, 91, 92, 203
 registering authentication
 numbers, 203
 RPC protocol and, 197
 rpcgen tool and, 50, 53
 servers, 92–94, 96
 service-dispatch routine and, 92, 93
time synchronization
  AUTH_DES authentication, 96, 206, 207
  AUTH_KERB authentication, 97, 213
verifiers
  AUTH_DES, 206, 207
  AUTH_KERB, 98, 99, 212–214
  AUTH_SYS, 204, 205
described, 197
authkerb_seccreate routine, 97
authorization, xv
  NIS+, 141
authsys_create routine, 95, 133
authsys_create_default routine, 95, 133
authunix_create routine, 133
authunix_create_default routine, 133
Automatic MT mode, see MT Auto
mode.

B
batching, 88, 91, 199, 269, 271
bcast.c program, 86, 87
bcast_proc routine, 87, 88
binding, xv
  dynamic, 274
  TI-RPC and, 11, 195
booleans
  RPC language, 222
  XDR language, 237
bottom level interface routines (RPC), 15, 77, 78
broadcast RPC, 86, 88
current vs. previous release, 132
overview, 86, 199
portmap routine and, 275, 277
routines for, 13, 86, 88
rpcbind routine and, 223, 230
server response to, 49, 51, 87, 88
TS-RPC vs. TI-RPC, 134, 137
buffer size
  specifying send and receive, 73, 76
byte arrays, xv
  XDR, 175

C
rpcgen tool and
  ANSI C compliance, 23, 37, 48
  C-style mode, 22, 37, 38, 41, 221
  preprocessing directives, 35, 36, 49, 50
  SPARCompiler C++ 3.0 compatibility, 48
XDR routines and, 163
C-style mode
rpcgen tool, 22, 37, 38, 41, 221
caching
  NIS+, 142
  server, 79
call semantics
  TI-RPC, 11, 195
callback procedures
  NIS+, 157, 158
transient RPC program numbers
  and, 104
  uses for, 104
callrpc routine, 132
cbc command, 296
changing
  NIS+ objects in namespace, 144
  NIS+ table entry objects, 144
  port monitor configuration
    scripts, 296, 298
  port monitor services, 288
  version numbers of programs, 101
cipher block chaining (CBC) mode, 98
circuit-oriented transports
  when to use, 18
circuit_n transport type, 17
circuit_v transport type, 17
cl_auth field, 79, 92
cl_netid field, 77
cl_private field, 77
cl_tp field, 77
classes of messages, 286
client authentication, see client authentication
client handles
creating

bottom level interface, 16, 77
current vs. previous release, 131
expert level interface, 15, 73, 75
intermediate level interface, 14,
70, 71
top level interface, 13, 14, 28, 29,
66, 68
defined, 309

destroying

current vs. previous release, 131
expert level interface, 75
top level interface, 28, 68

low-level data structures for, 79, 80

cient programs

remote copy, 124, 125
rpcgen tool and
 ANSI C-compliance, 48
 complex data structure
 passing, 33, 35
debugging, 55
directory listing service, 33, 35,
 263
message printing code
 example, 27, 30
MT-safety, 22, 44, 46
overview, 23, 27, 30
simplified interface, 58, 60

cient stub routines

rpcgen tool and, 21, 28, 29
 C-style mode, 39, 40
 MT Auto mode, 47
 MT-safe, 41, 43
 MT-unsafe, 43
 preprocessor directive, 35

cient templates

rpcgen tool, 22, 37, 39, 40
cient time-out periods
 creating timed clients, 14, 68, 71
rpcgen tool and, 50, 52

cients

batched, 88, 90, 269
defined, 309
multiple versions, 102, 104

multithreaded

overview, 106, 110
safety, 41, 43, 44, 46, 57, 309
User mode, 116, 117, 119
NIS+, 142
transaction IDs and, 195
TS-RPC vs. TI-RPC, 133, 134

_clnt.c suffix, 29
clt_broadcast routine, 132, 134, 135
clt_call routine
 current vs. previous release, 132
described, 14, 15
RPCPROGVERSISMATCH
 error, 103
top level interface and, 68
clt_control routine
 current vs. previous release, 131
described, 52, 75
 using, 52
clt_create routine
 code example, 27, 28
current vs. previous release, 131
described, 14
clt_create_timed routine
 current vs. previous release, 131
described, 14
 using, 68
clt_create_vers routine, 103, 131
clt_destroy routine
 current vs. previous release, 131
described, 28, 68, 75
clt_dg_create routine
 current vs. previous release, 131
described, 16, 77
 using, 77
clt_ops field, 77
clt_pcreateerror routine

current vs. previous release, 131
described, 68
clt_pererror routine, 55
clt_raw_create routine, 81, 84, 131
clt_spcreateerror routine, 131
clt_sperror routine, 55
clt_tli_create routine
 current vs. previous release, 131
described, 15, 73, 77
using, 73, 75
clnf tp_create routine, 14, 131
clnf tp_create_timed routine
current vs. previous release, 131
described, 14
using, 71
clnf vc_create routine
current vs. previous release, 131
described, 16, 77
using, 77
clnfraw_create routine, 131
clnf tcp_create routine, 131
clnfudp_bufcreate routine, 131
clnfudp_create routine, 73, 75, 131
clock synchronization, see time
conl synchronization
comments
XDR language, 247
compatibility
library functions, current vs. previous release, 130, 133
compilation
NIS+, 149
cpp directive
rpcgen tool and, 36
converting
addresses, 18, 19, 129
from XDR format, 62, 64, 70, 126, 127, 169, 309
local procedures to remote procedures, 23, 30
to XDR format, 30, 35, 62, 65, 126, 128, 165, 166, 168, 169, 310
copying
NIS+ database entries, 146
NIS+ objects, 148
NIS+ table entry objects, 144
remote, 123, 126
counted byte strings, see string declarations
cpp directive
cppgen tool and, 36
creases
server, 195, 207
creating
NIS+ databases, 146
NIS+ directory objects, 153
NIS+ group objects, 145, 153, 154
NIS+ table objects, 154, 156
utmp entries, 282, 289, 293
creating, see client handles, creating client handles,
creating, see server handles, creating server handles,
credentials
AUTH的信任, 95, 208
AUTH_KERB, 97, 98, 211, 213
described, 197
window (lifetime) of, 95, 97, 207, 208, 213
connection-oriented endpoints, 81
connection-oriented transports, see client handle creation for, 16
defined, 309
nettype parameters for, 17
port monitors and, 99
remote copy code example, 123, 126
server handle creation for, 16
connectionless transports, see client handle creation for, 16
defined, 309
nettype parameters for, 17
server handle creation for, 16
UDP (user datagram protocol), see
constants
RPC language, 217
XDR language, 245, 248, 250
constructed data type filters
XDR, 64, 174
conversation key
AUTH_DES authentication, 206, 208, 210
crashes
server, 195, 207
cpp directive
cppgen tool and, 36
creating
NIS+ databases, 146
NIS+ directory objects, 153
NIS+ group objects, 145, 153, 154
NIS+ table objects, 154, 156
utmp entries, 282, 289, 293
creating, see client handles, creating client handles,
creating, see server handles, creating server handles,
credentials
AUTH_DES, 95, 208
AUTH_KERB, 97, 98, 211, 213
described, 197
window (lifetime) of, 95, 97, 207, 208, 213
Index-315
daemons
   kerbd, 210, 211
   rpcbind, 19, 86
data representation, xv
   TI-RPC, 12
data structures, xv
   converting to XDR format, 30, 35, 62,
   65, 126, 128, 166, 168, 169
   low-level, 79
   MT safe, 116
   packing with xdr_inline, 37, 49
   recursive, 188, 191, 247
   rpcgen tool and, 30, 35
structure declarations, xv
data types
   passing arbitrary, 62, 65
database access functions (NIS+), 143, 146
datagram transports, xv
   broadcast RPC and, 86
defined, 309
   nettype parameters, 17, 24
   when to use, 17
datagram_n transport type, 17
datagram_v transport type, 17
date service
   intermediate level client for, 70, 71
   intermediate level server for, 71
top level client for trivial, 66, 68
top level server for, 68, 70
db_add_entry function, 146
db_checkpoint function, 146
db_first_entry function, 146
db_list_entries function, 146
db_next_entry function, 146
db_remove_entry function, 146
db_reset_next_entry function, 146
db_standby function, 147
dbxtool routine, 55
deallocating memory, , see memory,
releasing \x0d
debugging
   raw mode and, 81, 84
   rpcgen tool and, 49, 50, 55
declarations
   RPC language, 217, 222
XDR language, 234, 247
defaults
   maximum number of threads, 113
   single-threaded mode, 111
   defaults, see rpcgen tool, defaults\x0d
define statements, command line, rpcgen
tool, 49, 50
definitions
   RPC language, 215, 216, 251
deleting, xv
   address registrations, 131
   associations, 15
   mappings, 15
   NIS+ database entries, 146
   NIS+ directory from host, 146
   NIS+ group members, 145
   NIS+ group objects, 145, 159, 163
   NIS+ objects from namespace, 144,
   158, 159
   NIS+ table entry objects, 144, 159, 163
   port monitor services, 288
   port monitors, 280
memory, releasing, xv
deleting, see unregistering\x0d
   address registrations,
   DES encryption, 96, 98, 99, 206, 209, 210
deserializing, xv, 62, 64, 70, 126, 128, 169,
   309
destroying, xv
   client authentication handles, 92
   client handles, 28, 68, 75, 131
   NIS+ objects, 148
   server handles, 131
   XDR streams, 183
diffie-hellman encryption, 96, 206, 209,
   210
   dir.x program, 30, 32, 259, 260
   dir_proc.c routine, 32, 33, 259, 261
dir_remove function, 160
directories
   remote directory listing service, 30,
   35, 259, 263
   SAF (service access facility), 307, 306
disabling port monitors, xv, 282, 284
discriminated unions
declarations
  RPC language, 220
  XDR language, 220, 244, 250
XDR code samples, 180, 181
dispatch tables
  rpcgen tool, 50, 53, 55
doconfig function, 283, 291, 293, 294
domains (NIS+)
  administration commands, 142
  functions, 145, 147
  overview, 139, 140
dynamic binding, 274
dynamic program numbers, 104, 105, 197

E
ECB (electronic code book) mode, 98
electronic code book (ECB) mode, 98
enabling
  port monitors, 282, 284
  server caching, 79
encryption
  AUTH_DES authentication
    (Diffie-Hellman), 96, 206, 209, 210
  AUTH_KERB authentication, 98, 99
endnetconfig routine, 75
endpoints
  connection-oriented, 81
enumeration filters
  XDR primitives, 173, 174
enumerations
  RPC language, 31, 216, 217
  XDR language, 236
errors
authentication
  AUTH_DES, 207
  AUTH_KERB, 211, 212
  AUTH_SYS, 94, 205
client handle creation, 68
multiple client version, 103
NIS+ error message display
  functions, 143, 147
RPC, 29, 55, 196
  thr_create, 113
/etc/inet/inetd.conf file, 100
/etc/netconfig database, 16, 50, 130
/etc/rpc database, 12
/etc/saf/_pid file, 282, 292, 307, 299
/etc/saf/ directory, 306
/etc/saf/_config file, 292–294, 297
/etc/saf/_pmpipe file, 284, 292, 307
/etc/saf/_pmtab file, 283, 288, 289, 292, 307
/etc/saf/_svctag file, 288, 289, 292, 293, 307
/etc/saf/_sactab files, 287
/etc/saf/_sactab files, 306
/etc/saf/_sysconfig file, 293, 294, 297, 306
expert level interface routines (RPC), 15, 73, 77
  client, 73, 75
  overview, 15, 73
  server, 75, 77
external data representation, , see XDR (external data representation)

F
file data structure
  XDR language, 250
file descriptors, passing open TLI, 73, 76
file system, , see NFS (network file system)
filters (XDR)
  arrays, 175, 178–180
  constructed (compound) data
    type, 64, 174
  enumeration, 173, 174
  floating point, 62, 173
  number, 62, 63, 172, 173
  opaque data, 178, 179
  strings, 64, 174, 175
  unions, 180, 181
fixed-length arrays, xv
declarations
  RPC language, 218
  XDR language, 242
XDR code sample, 179, 180
fixed-length opaque data
  XDR language, 240

Index-317
flags, see rpcgen tool, flags
floating point
  XDR language, 238, 240
floating point filters
  XDR primitives, 62, 173
free routine, 35
freeing resources, xv
freenetconfig routine, 71

G
getnetconfig routine, 75
getnetconfig routine, 71
glossary, 309, 310
groups (NIS+)
  administration commands, 142, 143
  manipulation functions, 143, 145
  sample programs, 153, 154, 159, 163

H
hand-coded registration routine, 61
handles, xv
  authentication, 79, 95, 96
defined, 309
transport handles, xv
header files
  rpcgen tool and, 29, 35
hosts
  defined, 309
hyper integers
  XDR language, 237

I
.i suffix, 54
I/O streams
  XDR, 184
idempotent
  defined, 11
identifiers
  XDR language, 248
identifying, xv
  port monitor services, 288
  remote procedures, 12, 195, 196, 198
naming, xv
index table
rpcgen tool and, 35
indirect RPC, 227, 230
inetcda port monitor
  RPC services, 100, 104
  rpcgen tool and, 30, 51, 52
  using, 99, 100
inetcda.conf file, 100
information reporting
  addresses, 20
  NIS+, 145
  remote host status, 106, 110
  RPC, 20
  rpcbind server, 231
  server callbacks, 104
installing port monitor configuration
  scripts, 296, 298
integers, xv
  XDR language, 165, 166
  integers, XDR language, 235, 237
interfaces, see RPC (remote procedure call), interface routines
intermediate level interface routines (RPC), 14, 70
Internet protocols, see TCP (transport control protocol)
ISTATE environment variable, 282, 284, 292

K
kerbd daemon, 210, 211
Kerberos authentication, see AUTH_KERB
  authentication
keywords
  RPC language, 31
  XDR language, 249
KGETKCRED procedure, 210, 211
KGETUCRED procedure, 210–212
KSETKCRED procedure, 210, 211

L
lib library, 34
libc library, 129, 130
libnsl library, 29, 31, 129, 130
libraries
  lib, 34
  libc, 129, 130
  libnsl, 29, 31, 129, 130
  librpcsvc, 58
  lthread, 107
  RPC functions, 130, 133
  rpcgen tool and
    libnsl, 29, 31, 129, 130
    selecting TI-RPC or TS-RPC
    library, 23, 37, 48
  XDR, xv, 168, 169
  librpcsvc library, 58
  lifetime of credentials, , see window of credentials
limits
  broadcast request size, 86
  maximum number of threads, 113
linked lists
  XDR, 188, 191, 247
listen port monitor
  administrative command for, 283
  rpcgen tool and, 30, 51, 52
  using, 99–101
listing, xv
  NIS+ objects, 148, 157, 158
  NIS+ principals, 145
  NIS+ servers, 146
  NIS+ table objects, 144, 156, 158
  portmap mappings, 277
  remote directory listing service, 30, 35, 259, 263
  rpcbind addresses, 230
  rpcbind mappings, 12, 226, 229
live code examples, 259, 271
  adding two numbers program, 264
  batched code, 269, 271
  directory listing program, 259, 263
  print message program, 266, 268
  spray packets program, 264, 265
  time server program, 263
loading, , see binding
local procedures
  converting to remote procedures, 23, 30
locks
  mutex, multithreaded mode and, 111
  port monitor IDs and lock files, 282, 292, 293, 307
log functions
  NIS+ transaction, 143, 147
  _log file
    SAC (service access controller), 307
low-level data structures, 79
lthread library, 107
M
  main server function, 51
makefile templates
  rpcgen tool, 22, 37
  mapping, xv, 15
  rpcbind routine, xv
master servers
  NIS+, 140, 142, 161, 162
maximums
  broadcast request size, 86
  number of threads, 113
memory
  allocating with XDR, 126, 128
  releasing
    clnt_destroy routine, 28, 68, 75
    free routine, 35
    messageprog_1_freeresult routine, 44
    NIS+, 144–146
    simplified interface and, 65
    svc_done routine, 117
    svc_freeargs routine, 127
    XDR_FREE operation, 174
    xdr_free routine, 35, 44
    XDR primitive requirements for, 170, 172, 174
memory streams
  XDR, 184
memory, releasing, xv
mesg_proc.c routine, 268
message classes, 286
message interface (SAF), 284, 286, 292, 307, 298, 304
messageprog_1_freeresult routine, 44
modifying, , see changing
msg.h header file, 29
msg.x program, 41
msg_clnt.c routine, 29
msg_svc.c program, 29
msg_svc.c routine, 29
MT Auto mode, 110, 112, 116
  code examples, 113, 116
described, 110, 112
rpcgen tool and, 23, 37, 47
service transport handle and, 112
MT hot
defined, 309
MT RPC programming, see multithreaded RPC
programming
MT User mode, xv, 110, 112, 116, 123
MT-safe code, xv
  clients, 22, 41, 43, 44, 46, 57, 309
defined, 309
rpcgen tool and, 22, 37, 41, 47
servers, 22, 43, 44, 46, 47, 57, 111, 309
multiple client versions, 102, 104
multiple server versions, 101, 102
multithreaded Auto mode, see MT Auto
mode
multithreaded hot
defined, 309
multithreaded RPC programming, xv, 105, 123
  clients
    overview, 106, 110
    safety, 22, 41, 43, 44, 46, 57, 309
    User mode, 116, 117, 119
library, 107
maximum number of threads, 113
overview, 105
performance enhancement, 113, 119
rpcgen tool and, 22, 37, 41, 47
servers
  Auto mode, 23, 37, 47, 110, 112, 116
  overview, 105, 110, 112
  safety, 22, 43, 44, 46, 47, 57, 111, 309
timing diagram, 111
unsafe routines, 111
User mode, 110, 112, 116, 119, 123
multithreaded User mode, 110, 112, 116, 123
multithreaded-safe code, see MT-safe
code
mutex locks
  multithreaded mode and, 111

N
Name Service Switch, 141
name-to-address translation, 18, 19, 129
naming, xv
  client stub programs by rpcgen, 29
  netnames, 95, 206
  programs by version number, 101
  remote procedure calls by rpcgen, 27
  server programs by rpcgen, 29
  standard for, 206
template files for rpcgen, 38
naming service, see NIS+ (Network
  Information Services
  Plus)
NIS+ (Network Information Services
  Plus), xv
netbuf addresses, 229, 230
netconfig database, 16, 50, 130
netnames, 95, 206
NETPATH environment variable, 16, 50, 68
nettype parameters, 17, 24
network addresses, see addresses
  network file system, see NFS (network
  file system)
Network Information Services Plus, see
  NIS+ (Network
  Information Services
  Plus)
network names, 95, 206
network pipes, 165
network selection
  RPC, 16
  rpcgen tool, 49, 50
network services
defined, 310
Newstyle (C-style) mode
  rpcgen tool, 22, 37, 38, 41
NFS (network file system)
defined, 310
Kerberos authentication and, 211, 212
NFSPROC_GETATTR procedure, 211
NFSPROC_STATVFS procedure, 212
nicknames
  AUTH_DES, 207–209
  AUTH_KERB, 98, 99, 211–213
NIS+
  servers
    sample program, 161, 162
NIS+ (Network Information Services Plus), 139, 163
application programming interface
  (API), 143, 148
cache administration commands, 142
client administration commands, 142
compilation, 149
database access functions, 143, 146
domains
  administration commands, 142
  functions, 145, 147
  overview, 139, 140
time synchronization, 147
event message display functions, 143, 147
groups
  administration commands, 142, 143
  manipulation functions, 143, 145
  sample programs, 153, 154, 159, 163
namespace administration
  commands, 142, 143
objects
  administration commands, 142, 143
  manipulation functions, 143, 147
  sample programs, 149, 163
overview, 5, 139, 143
sample program, 148, 163
security, 141
  administration commands, 142
servers
  administration commands, 142
  functions, 143, 146
  overview, 140
tables
  access functions, 143, 144
  administration commands, 142
  overview, 140, 141
  sample programs, 154, 156
time synchronization, 147
transaction log functions, 143, 147
unsupported macros, 148
nis_add function, 144, 149, 153, 155
nis_add_entry function, 144, 149, 156
nis_addmember function, 145, 149, 152
nis_admin functions, 147
nis_cachemgr command, 142
nis_checkpoint function, 147
nis_clone_object function, 148
nis_creategroup function, 145, 149, 153, 154
nis_db functions, 146
nis_destru_object function, 148
nis_destroygroup function, 145, 149, 159, 160
nis_dir_cmp function, 147
nis_domain_of function, 147, 149
nis_error functions, 147
nis_first_entry function, 144
nis_freenames function, 145
nis_freeresult function, 144, 149
nis_freeservlist function, 146
nis_freevags function, 146
nis_getnames function, 145
nis_getservlist function, 146
nis_groups functions, 145
Index-321
nis_ismember function, 145
nis_leaf_of function, 147, 149, 157
nis_terror function, 147
nis_list function, 144, 149, 156, 158
nis_local_directory function, 145, 149, 150
nis_local_group function, 145
nis_local_host function, 145
nis_local_names functions, 145
nis_local_principal function, 145, 149, 150
nis_lookup function, 144, 149, 151, 156, 157
nis_mkdir function, 146, 149, 153
nis_modify function, 144
nis_modify_entry function, 144
nis_name_of function, 147
nis_next_entry function, 145
nis_pererror function, 147, 149
nis_ping function, 147
nis_print_group_entry function, 145
nis_print_object function, 148
nis_remove function, 144, 149, 153, 158–160
nis_remove_entry function, 144, 149, 159, 160
nis_removemember function, 145, 149, 159
nis_rmdir function, 146
nis_server functions, 146
nis_servstate function, 146
nis_spermo function, 147
nis_serror function, 147
nis_stats function, 146
nis_subr functions, 147, 148
nis_tables functions, 144
nis_verifygroup function, 145
nisaddcred command, 142
nisaddent command, 142
nischgrp command, 142
nischmod command, 142
nischown command, 142
nischt1l command, 143
nisdefaults command, 143
nisgrep command, 142
nisgpradm command, 142
nissinit command, 142
nisln command, 143
nisls command, 142
nismatch command, 142
nismkdir command, 142
nispasswd command, 142
nisrm command, 143
nisrmdir command, 142
nissetup command, 142
nistbladm command, 142
nisupdkeys command, 142
nlsadmin command, 283
no-data routine
XDR, 174
NULL arguments, 61
NULL pointers, 182
NULL strings, 222
NULL transport type, 17
number filters, XDR, xv, 62, 63, 172, 173
number of users
on a network, 93, 94
on a remote host, 58
numbers, xv
adding two, 264
version numbers, xv

O
objects (NIS+)
administration commands, 142, 143
manipulation functions, 143, 147
sample programs, 149, 163
ONC+ overview, 4, 5
opaque data
declarations
RPC language, 222
XDR language, 240, 241
XDR code examples, 178, 179
open TLI file descriptors
passing, 73, 76
optional-data unions
XDR language, 246

P
parameters, , see arguments\x0d
passing parameters, , see arguments\x0d
percent sign (%), preprocessing directive, 35
_pid file, 282, 292, 307, 299
ping program, 214, 215, 310
pipes
  network, 165
    _pmpipe file, 284, 292, 307
    _sacpipe file, 284, 292, 307
pm_maxclass field, 286
pm_size field, 286
pmadm command, 100, 101, 283, 288, 290, 297, 298
pmap_getmaps routine, 132
pmap_getport routine, 132, 133
pmap_rmtcall routine, 132
pmap_unset routine, 132
pmmsg structure, 285
_pmpipe file, 284, 292, 307
_pmtab file, 283, 288, 292, 307
/p directory, 284, 288, 307
/pmtag directory, 306
PMTAG environment variable, 284, 292
pointers
  remote procedures, 26, 27
  RPC language, 219
  XDR code examples, 181, 183
poll routine, 84, 86
pop configuration-script keyword, 295
port monitors
  activity monitoring, 281
  adding, 280
  adding services, 288
administrative commands
  monitor-specific command, 291
  pmadm, 100, 101, 283, 288, 290, 297, 298
  sacadm, 101, 280, 287, 290, 296, 297
administrative files
  _pmtab, 283, 288, 292, 307
  _sactab, 287, 306
administrative interface, 287, 293
changing port monitor services, 288
configuration scripts, 293, 298
  installing, 296, 298
  language for writing, 295, 296
  per-port monitor, 292–294, 297
  per-service, 283, 284, 291–293, 295, 297, 298
  per-system, 293, 294, 296, 297, 306
  printing, 296, 298
  replacing, 296, 298
deleting services, 288
disabling, 282, 284
enabling, 282, 284
files
  key, 292
  per-port monitor configuration, 292–294, 297
  per-service configuration, 283, 284, 291–293, 307, 297, 298
  per-system configuration, 293, 294, 296, 297, 306
  private, 284, 288
  process ID, 282, 292, 293, 307
functions, 281, 283, 293
home directory for, 306
identifying services, 288
management function, 281
message interface, 284, 286, 292, 307, 298, 304
_pmpipe file, 284, 292, 307
private files, 284, 288
process IDs and lock files, 282, 292, 293, 307
removing, 280
requirements for implementing, 291, 292
restricting access to system, 282

Index-323
rpcgen tool and, 30, 50–52
sample code, 298, 304
service interface, 291
terminating, 283, 293
types of, 287
using, 99, 101
utmp entry creation, 282, 289, 293
version numbers, 289, 291
port monitors, see SAF (service access facility)
writing with service access facility,
port numbers, xv
getting for registered services, 133, 273, 278
portmap routine, 276
rpcbind routine, 19, 228
TCP/IP protocol, 19, 228, 276
UDP/IP protocol, 19, 228, 276
porting data, see XDR (external data representation)
porting TS-RPC to TI-RPC, 128, 137
applications, 128
benefits, 128
code comparison examples, 133, 137
differences between TI-RPC and
TS-RPC, 129, 133, 137
function compatibility lists, 130, 133
libc library and, 129
libnsl library and, 129
name-to-address mapping and, 129
old interfaces and, 129
portmap routine, 273, 278
address management functions, 132
broadcast RPC and, 275, 277
operation of, 276, 278
overview, 273, 274
port number, 276
protocol specification for, 274, 276
replacement of, 19, 273
TS-RPC and, 130
ports
defined, 281
preprocessing directives
rpcgen tool, 35, 36, 49, 50
printing, xv
message to system console, 23, 30, 266, 268
port monitor configuration
scripts, 296, 298
printmsg.c program
remote version, 25, 30, 266, 268
single process version, 23, 24, 266
procedure numbers, xv
described, 12, 196
error conditions, 196
procedure-lists
RPC language, 216
procedures
registering as RPC programs, 13, 61
RPC language, 216
program declarations
RPC language, 220, 221
program numbers, xv, 196, 198
assigning, 197, 198
described, 12, 196
error conditions, 196
registering, 198
transient (dynamically assigned), 104, 105, 198
program numbers, see portmap routine
mapping,
program-definitions
RPC language, 216
PROGVERS program name, 101
PROGVERS_ORIG program name, 101
protocols, xv
AUTH_DES, 207, 210
rpcbind protocol specification, 223, 228
specifying in RPC language, 25
XDR language, xv
push configuration-script keyword, 295

Q
quadruple-precision floating point
XDR language, 239, 240

R
raw RPC
testing programs using low-level, 81, 84
READDIR procedure, 30, 35, 259, 263
record streams
XDR, 184, 186, 202, 203
record-marking standard, 202, 203
recursive data structures, 188, 191, 247
registering
authentication numbers, 203
current vs. previous release, 131
hand-coded registration routine, 61
procedures as RPC programs, 13, 60, 61
program numbers, 198
program version numbers, 101
registering, see portmap routine
addresses,
registerrpc routine, 132
releasing memory, see memory, releasing
remote copy routine, 123, 126
remote directory listing service, 30, 35
remote procedure call, see RPC (remote procedure call)\x0d
remote procedures
converting local procedures to, 23, 30
identifying, 12, 195, 196, 198
remote time protocol, 35, 47
removing, see deleting
rendezvousing
TI-RPC and, 195
replacing, see changing
replica servers
NIS+, 140, 142
reporting, see information reporting\x0d
rlsc routine, 35, 263
RPC (remote procedure call), xv
address lookup services, 15, 19, 20
address reporting, 20
address translation, 18, 19, 129
asynchronous mode, 84, 86
batching, 88, 91, 199, 269, 271
described, 9, 10, 194
described, 9, 10, 194
described, 9, 10, 194
errors, 29, 55, 196
failure of, 29
identifying remote procedures, 12, 195, 196, 198
indirect, 227, 230
information reporting, 20, 104
interface routines, 12, 16, 57, 78
bottom level, 15, 77, 78
caching servers, 79
expert level, 15, 73, 77
intermediate level, 14, 70
low-level data structures, 79
overview, 12, 16, 57, 58, 65
simplified, 13, 58, 65
standard, 13, 16, 65
top level, 13, 28, 29, 66, 70
multiple client versions, 102, 104
multiple server versions, 101, 102
name-to-address translation, 18, 19, 129
network selection, 16
poll routine, 84, 86
port monitor usage, 99, 101
raw, testing programs using
low-level, 81
record-marking standard, 202, 203
standards, 10, 202, 203
transient RPC program numbers, 104, 105, 197
transport selection, 17
transport types, 17, 24
rpcgen tool, xv
RPC (remote procedure call), see
broadcast RPC\x0d
broadcast,
RPC (remote procedure call), see portmap
routine
address registration,
RPC (remote procedure call), see TI-RPC
(transport-independent remote procedure call),
interface routines\x0d
interface routines,
RPC (remote procedure call), see TI-RPC
(transport-independent remote procedure
call)protocol\x0d
protocol,
RPC language
reference, 251, 258
RPC language (RPCL), xv, 214, 222
arrays, 218, 219
booleans, 222
C vs., 21
C-style mode and, 221
declarations, 217, 219
definitions, 215, 216, 251
discriminated unions, 31, 220
enumerations, 31, 216, 217
example protocol described in, 25
example service described in, 214, 215
fixed-length arrays, 218
keywords, 31
opaque data, 222
overview, 251, 310
pointers, 219
portmap protocol specification in, 275,
276
program declarations, 220, 221
simple declarations, 218
special cases, 221, 222
specification for, 214, 222
strings, 25, 222
structures, 31, 219
syntax, 215, 216, 251
type definitions, 217
unions, 31, 220
variable-length arrays, 218
voids, 222
XDR language vs., 214, 216, 251
rpc.nisd command, 142
RPC/XDR, , see RPC language
(RPCL)\0d
RPC_AUTHERROR error, 207
rpc_broadcast routine
  current vs. previous release, 132, 134
described, 13
  source code, 135, 137
  using, 86, 88
rpc_call routine
  current vs. previous release, 132
described, 13, 58
  using, 58, 60
RPC_CLNT preprocessing directive, 35
rpc_control routine, 111, 112, 116
rpc_createerr global variable, 68
__rpc_dtbsize function, 85
RPC_HDR preprocessing directive, 35
rpc_reg routine
  current vs. previous release, 132
described, 13
  using, 60, 62
__rpc_select_to_poll function, 84, 85
RPC_SVC preprocessing directive, 35
RPC_SVC_MTMODE_GET library
  routine, 113, 116
RPC_SVC_MTMODE_SET library
  routine, 113
RPC_SVC_THRCREATES_GET library
  routine, 113
RPC_SVC_THRERRORS_GET library
  routine, 113
RPC_SVC_THRMAX_GET library
  routine, 113
RPC_SVC_THRMAX_SET library
  routine, 113
RPC_SVC_THRTOTAL_GET library
  routine, 113
RPC_TBL preprocessing directive, 35
RPC_XDR preprocessing directive, 35
rpcb_getaddr routine, 15, 133
rpcb_getmaps routine, 133
rpcb_gettime routine, 133
rpcb_rmtcall routine, 133
rpcb_set routine
  current vs. previous release, 133
described, 15, 76
rpcb_unset routine, 15, 133
rpcbind daemons
  broadcast RPC and, 86
  registering addresses with, 19
rpcbind routine, xv
  address management functions, 132
  broadcast RPC and, 223, 230
calling, 15
described, 11, 19, 20, 129, 133, 206, 223
  listing mappings, 12, 226, 229
operation of, 228, 231
port number, 19, 228
portmap routine replaced by, 19, 273
protocol specification for, 223, 228
time service, 133, 206, 229
version 4, 230, 231
RPCBPROC_BCAST procedure, 227, 230
RPCBPROC_CALLIT procedure, 20, 226,
229, 230
RPCBPROC_DUMP procedure, 226, 229
RPCBPROC_GETADDR procedure, 226,
229
RPCBPROC_GETADDRLIST
procedure, 227, 230
RPCBPROC_GETSTAT procedure, 227,
231
RPCBPROC_GETTIME procedure, 206,
226, 229
RPCBPROC_GETVERSADDR
procedure, 227, 230
RPCBPROC_INDIRECT procedure, 227,
230
RPCBPROC_NULL procedure, 225, 228
RPCBPROC_SET procedure, 225, 228, 231
RPCBPROC_TADDR2UADDR
procedure, 226, 227, 230
RPCBPROC_UADDR2TADDR
procedure, 226, 227, 229
RPCBPROC_UNSET procedure, 225, 228,
231
rpcgen tool, 21, 55
add two numbers program, 264
advantages, 22
arguments, 26, 27, 38, 41, 61, 62, 65,
221
authentication and, 50, 53, 95, 99
batched code example, 269, 271
broadcast call server response, 49, 51
C and
ANSI C compliance, 23, 37, 48
C-style mode, 22, 37, 38, 41, 221
preprocessing directives, 35, 36,
49, 50
SPARCCompiler C++ 3.0
compatibility, 48
compilation modes, 22, 38, 41
complex data structure passing, 30, 35
converting local procedures to remote
procedures, 23, 30
cpp directive, 36
debugging, 49, 50, 55
defaults
argument passing mode, 38, 40
C preprocessor, 36
client time-out period, 52
compilation mode, 22
library selection, 48
MT-safety, 22, 41
output, 21
server exit interval, 52
define statements on command
line, 49, 50
described, 21, 22
directory listing program, 30, 35, 259,
263
dispatch tables, 50, 53, 55
failure of remote procedure calls, 29
flags
-A (MT Auto mode), 37, 47
-a (templates), 37
-b (TS-RPC library), 37, 48
-i (xdr_inline count), 37, 49
listed, 36
-M (MT-safe code), 37, 41
-N (C-style mode), 37, 38
-Sc (templates), 37
-Sm (templates), 37
-Ss (templates), 37
hand-coding vs., 61
libraries
libnsl, 29, 31, 129, 130
selecting TI-RPC or TS-RPC
library, 23, 37, 48
MT (multithread) Auto mode, 23, 37,
47, 112, 116
MT (multithread)-safe code, 22, 37,
41, 47
naming remote procedure calls, 27
network types/transport selection, 49,
50
Index-327
Newstyle (C-style) mode, 22, 37, 38, 41
optional output, 21
pointers, 26, 27
port monitor support, 30, 49, 51, 52
preprocessing directives, 35, 36, 49, 50
print message program, 23, 30, 266, 268
programming techniques, 49, 55
socket functions and, 48
spray packets program, 264, 265
SunOS 5.x features, 22, 23
templates, 22, 37, 39, 40
TI-RPC vs. TS-RPC, 129
TI-RPC vs. TS-RPC library
selection, 23, 37, 48
time server program, 35, 47, 263
time-out changes, 50, 52
tutorial, 23, 37
variable declarations and, 218
XDR routine generation, 30, 35, 163
xdr_inline count, 37, 49
rpcgen tool, see client handles
handles and,
rpcgen tool, see client handles\x0d
client handles and,
rpcgen tool, see client programs\x0d
client programs and,
rpcgen tool, see client stub routines
stub routines and,
rpcgen tool, see client stub routines\x0d
client stub routines and,
rpcgen tool, see server handles\x0d
server handles and,
rpcgen tool, see server programs\x0d
server programs and,
rpcgen tool, see server stub routines
server stub routines and,
rpcinfo routine, 19, 20
RPCPROGVERSMISMATCH error, 103
/rpcsvc directory, 198
rprintmsg routine, 30
rq_clntcred field, 93
rq_cred field, 92, 93
rstat program
multithreaded, 106, 110

run configuration-script keyword, 296
runwait configuration-script keyword, 296
rusers routine, 61
rusersDefault Para Font routine, 58

S
SAC (service access controller)
described, 280, 283
key files, 287, 292, 306
log file, 307
message interface, 284, 286, 292, 307, 298, 304
go header file, 302, 304
sacadm command, 101, 280, 287, 290, 296, 297
_sacpipe file, 284, 292, 307
_sactab file, 287, 306
starting, 292, 294
_sacpipe file, 302, 304
sacadm command, 101, 280, 287, 290, 296, 297
_sacpipe file, 284, 292, 307
_sactab file, 287, 306
SAF (service access facility), xv, 279, 307
administrative interface, 287, 293, 296, 298
key files, 292
monitor-specific command, 291
pmadm command, 100, 101, 283, 288, 290, 297, 298
_pmcmd file, 283, 288, 289, 292, 307
port monitor implementation
requirements, 291, 292
port monitor responsibilities, 293
sacadm command, 101, 280, 287, 290, 296, 297
_sactab file, 287, 306
service interface, 291
configuration scripts, 293, 298
  installing, 296, 298
  language for writing, 295, 296
  per-port monitor, 292–294, 297
  per-service, 283, 284, 291–293, 295, 297, 298
  per-system, 293, 294, 296, 297, 306
  printing, 296, 298
  replacing, 296, 298
directory structure, 307, 306
files used by, 283, 287–289, 292
logic diagram, 305
message interface, 284, 286, 292, 307, 298, 304
overview, 279, 283
port monitor functions and, 281, 283, 293
SAC (service access controller)
  and, 280, 283, 284, 286
sample code, 298, 304
terminating port monitors, 283, 293
sc_size field, 285, 286
searching NIS+ tables, 144
security, xv
  NIS+, 141
authentication, xv
semantics
  TI-RPC call, 11, 195
  serializing, xv, 30, 35, 62, 65, 126, 128, 165, 166, 168, 169, 310
server handles, xv
  creating
    bottom level interface, 16, 78
    current vs. previous release, 131
    expert level interface, 15, 75, 77
    intermediate level interface, 14, 71
    top level interface, 13, 14, 68, 70
destroying, current vs. previous release, 131
  low-level data structures for, 80
server programs
  remote copy, 125, 126
transaction IDs and, 195
service access controller, see SAC (service access controller)

service access facility, see SAF (service access facility)

service transport handle (SVCXPRT), 99, 112

service-distribution routine

   authentication and, 92, 93

setnetconfig routine, 75

signed integers

   XDR language, 235

simple declarations

   RPC language, 218

simplified interface routines (RPC), 13, 58, 65

   client, 58, 60

   hand-coded registration routine, 61

   MT safety of, 57

   overview, 13, 58

   server, 60

   XDR conversion, 62, 65

single-threaded mode

   as default, 111

   poll routine and, 84, 86

socket functions, see TS-RPC

   (transport-specific remote procedure call)

SPARC Compiler C++ 3.0

rpcgen tool and, 48

spray.x (spray packets) program, 264, 265

standard interface routines (RPC), 13, 16, 65

   bottom level routines, 15, 77, 78

   expert level routines, 15, 73, 77

   intermediate level routines, 14, 70

   low-level data structures, 79

   MT safety of, 57

   server caching, 79

   top level routines, 13, 28, 29, 66, 70

standards

   ANSI C standard, rpcgen tool and, 23, 37, 48

   naming standard, 206

   record-marking standard, 202, 203

   RPC, 10, 202, 203

   XDR canonical standard, 167

   statistics, see information reporting

   status reporting, see information reporting

   stream transports, xv

   defined, 310

STREAMS modules

   port monitor configuration

   and, 295–297

stream, see XDR (external data representation), streams

string declarations

   RPC language, 25, 222

   XDR language, 241, 242

string representation, xv

   XDR routines, 64, 174, 175

structure declarations, xv

   RPC language, 31, 219

   XDR language, 244, 250

stub routines, see client stub routines

Sun RPC, see TI-RPC

   (transport-independent remote procedure call)

SunOS 5.x

   rpcgen tool features, 22, 23

   _svc suffix, 48

   _svc.c suffix, 29

   svc_create routine

   current vs. previous release, 131

   described, 14, 69

   svc_destroy routine, 131

   svc_dg_create routine

   current vs. previous release, 131

   described, 16, 81

   using, 78

   svc_dg_enablecache routine, 79

   svc_done routine, 111, 117

   svc_fd_create routine, 99, 131

   svc_freeargs routine, 127

   svc_getargs routine

   described, 128

   MT performance and, 113, 119

   svc_getcaller routine, 132
svc_getreqpoll routine, 84, 111
svc_getreqset routine, 84, 111
svc_getrccaller routine, 132
svc_pollset routine, 84
svc_raw_create routine, 81, 84, 131
svc_reg routine
  current vs. previous release, 132
described, 15
port monitors and, 99
svc_register routine, 132
svc_run routine
  bypassing, 84
described, 60, 69
multithreaded RPC servers and, 111
poll routine and, 84–86
svc_sendreply routine, 69
svc_tli_create routine
  current vs. previous release, 131
described, 15, 75
using, 75, 77, 99
svc_tp_create routine, 14, 131
svc_udp_bufcreate routine, 131
svc_unreg routine, 15, 132
svc_unregister routine, 132
svc_vc_create routine
  current vs. previous release, 131
described, 16, 81
using, 78
svcerr_systemerr routine, 94
svcfd_create routine, 131
svctag file, 288, 289, 292, 293, 307
svcudp_create routine, 131
SVCXPRT service transport handle, 99, 112
svrerr_weakauth routine, 94
synchronization, see time
  synchronization
syntax
  RPC language, 215, 216, 251
  XDR language, 249, 250
 Sysconfig file, 293, 294, 297, 306

T
  tables (NIS+)
    access functions, 143, 144
    administration commands, 142

overview, 140, 141
sample programs, 154, 156
/tag directory, see /pmtag directory
TCP (transport control protocol), xv
  nettype parameter for, 17
porting TCP applications from
    TS-RPC to TI-RPC, 128
  portmap port number for, 276
  portmap sequence for, 274
  RPC protocol and, 195
  rpcbind port number for, 19, 228
  server crashes and, 195
tcp transport type, 17
TCP/IP protocol, see TCP (transport
  control protocol)
TCP/IP streams
  XDR, 184, 186, 202, 203
templates
  rpcgen tool, 22, 37, 39, 40
terminating port monitors, 283, 293
testing
  NIS+ groups, 145
  programs using low-level raw
    RPC, 81, 84
thr_create routine, 116
thread library
  thread, 107
thread.h file, 116
threads, see MT Auto mode
TI-RPC (transport-independent remote
  procedure call), xv
  address lookup services, 15, 19, 20
  address reporting, 20
  address translation, 18, 19, 129
call semantics, 11
data representation, 12
described, 4, 9, 10
  identifying remote procedures, 12,
    195, 196, 198
  information reporting, 20, 104

Index-331
interface routines, 12, 16, 57, 78
bottom level, 15, 77, 78
caching servers, 79
expert level, 15, 73, 77
intermediate level, 14, 70
low-level data structures, 79
overview, 12, 16, 57, 58, 65
simplified, 13, 58, 65
standard, 13, 16, 65
top level, 13, 28, 29, 66, 70
library selection, rpcgen tool and, 23, 37, 48
name-to-address translation, 18, 19, 129
network selection, 16
protocol, xv, 193, 203
authentication and, 197
binding and rendezvous
independence and, 11, 195
identifying procedures, 12, 195, 196, 198
in XDR language, 199, 202
overview, 193, 195
record-marking standard, 202, 203
transport protocols and semantics
and, 11, 194
version number, 196
raw, testing programs using
low-level, 81, 84
transient RPC program numbers, 104, 105, 198
transport selection, 17
transport types, 17, 24
rpcgen tool, xv
TI-RPC (transport-independent remote
procedure call), see
arguments
parameter passing,
TI-RPC (transport-independent remote
procedure call), see
porting TS-RPC to
TI-RPC
TS-RPC and,
TI-RPC (transport-independent remote
procedure call), see
rpcbind routine
address registration,
time
obtaining current, 133, 206, 229
ping program, 214, 215
time server program, 35, 47, 263
time service
intermediate level client for, 70, 71
intermediate level server for, 71
rpcbind routine, 133, 206, 229
top level client for, 66, 68
top level server for, 68, 70
time synchronization
AUTH_DES authentication, 96, 207
AUTH_KERB authentication, 97, 213
NIS+, 147
time-out periods
rpcgen tool and, 50, 52
time.x program, 35, 47, 263
time_prog routine, 69
timed client creation
intermediate level interface, 14, 71
top level interface, 14, 68
timestamps, see time
synchronization
TLI (transport-level interface), see expert
level interface
routines
TLI file descriptors
passing open, 73, 76
top level interface routines (RPC), 13, 28,
29, 66, 70
client, 28, 29, 66, 68
overview, 13, 66
server, 68, 70
transaction IDs, 12, 195
transaction log functions (NIS+), 143, 147
transient program numbers, 104, 105, 197
translating, see converting
transport
defined, 310
transport addresses (netbuf addresses), 229, 230
transport handles
  defined, 310
server, 80
SVCXPRRT service, 99, 112
XDR routines requiring, 66, 69
transport protocols, xv
  RPC protocol and, 194
UDP (user datagram protocol), xv
transport selection
  RPC, 17
  rpcgen tool, 49, 50
transport types, xv
  described, 17, 24
  interfaces and, 66
  rpcgen tool and, 49, 50
connectionless transports, xv
transport-independent remote procedure call, , see TI-RPC
  (transport-independent remote procedure call)\x0d
transport-level interface (TLI) routines, , see expert level interface routines\x0d
transport-level interface file descriptors passing open, 73, 76
transport-specific remote procedure call, , see TS-RPC
  (transport-specific remote procedure call)\x0d
trees, 247
TS-RPC (transport-specific remote procedure call)
  library selection, rpcgen tool and, 37, 48
TS-RPC (transport-specific remote procedure call), see porting TS-RPC to TI-RPC\x0d
porting to TI-RPC,
ttyadm command, 283
ttymon port monitor, 283
tutorials
  rpcgen tool, 23, 36
type definitions
  RPC language, 217
  XDR language, 246, 249
U
UDP (user datagram protocol), xv
  broadcast RPC and, 86
  client creation routines for, 73, 75, 77
  nettype parameter for, 17
  porting UDP applications from
    TS-RPC to TI-RPC, 128
  portmap port number for, 276
  RCP protocol and, 195
  rpcbind port number for, 19, 228
  server creation routines for, 75, 77, 78
  udp transport type, 17
UDP/IP protocol, , see UDP (user datagram protocol)\x0d
ulimit command, 296
umask command, 296
unions
  declarations
    RPC language, 31, 220
    XDR language, 244, 246, 250
    XDR code samples, 180, 181
  universal addresses, 18, 228–230, 274, 310
  unregistering, xv
    current vs. previous release, 131
    portmap routine, 277
    rpcbind routine, 19, 225, 228
    rpcinfo routine, 20
  unsigned hyper integers
    XDR language, 237
  unsigned integers
    XDR language, 235
user datagram protocol, , see UDP (user datagram protocol)\x0d
User MT mode, 110, 112, 116, 123
user’s bind address
  passing, 76
users
  number of
    on a network, 93, 94
    on a remote host, 58
/usr/include/rpcsvc directory, 198
/usr/share/lib directory, 34
utmp entries
Index-333
creating, 282, 289, 293

V
/var/saf/_log file, 307
/var/saf/ directory, 284, 288, 307
variable declarations, 218
variable-length array declarations, xv
  RPC language, 218
  XDR language, 243
variable-length opaque data, xv
  XDR language, 240, 241
verifiers
  AUTH_DES, 206, 207, 209
  AUTH_KERB, 98, 99, 212–214
  AUTH_SYS, 204, 205
described, 197
version numbers, xv
  assigning, 101
  changing, 101
described, 12, 101, 196
  error conditions, 196
message protocol, 196
  multiple client versions, 102, 104
  multiple server versions, 101, 102
port monitors, 289, 291
described, 101
versions, xv
version numbers, see portmap routine
version-lists
  RPC language, 216
versions, xv
library functions, compatibility of
  current vs. previous
release, 130, 133
RPC language, 216
visible transport type, 17
void arguments, 221
void declarations
  RPC language, 222
  XDR language, 245

W
window of credentials

AUTH_DES authentication, 95, 208
AUTH_KERB authentication, 97, 213
defined, 207
window verifiers, 207

X
x_base field, 187
x_destroy macro, 187
x_getbytes routine, 187
x_getlong routine, 187, 188
x_getpostn macro, 187
x_handy field, 187
x_inline routine, 187
x_op field, 187
x_private field, 187
x_public field, 187
x_putchar routine, 187
x_putchar routine, 187, 188
x_setpostn macro, 187
.x suffix, 31
XDR (external data representation), xv,
  163, 191
  block size, 234
canonical standard, 167
converting from (deserializing), 62,
  64, 70, 126, 127, 169, 309
converting to (serializing), 30, 35, 62,
  65, 126, 128, 165, 166, 168,
  169, 310
cost of conversion, 167
described, 4, 163, 166, 233
direction determination for
  operations, 183
direction independence of
  routines, 169
described, 4, 163, 166, 233
file data structure in, 250
graphic box notation, 234
library, xv, 168, 169
linked lists, 188, 191, 247
memory allocation with, 126, 128
optimizing routines, 183
primitive routines, 62, 65, 168, 170, 186
arrays, 175, 178–180
byte arrays, 175
described (compound) data
type filters, 64, 174
discriminated unions, 180, 181
enumeration filters, 173, 174
fixed-length arrays, 179, 180
floating point filters, 62, 173
memory requirements, 170, 172, 174
no-data routine, 174
nonfilter, 183
number filters, 62, 63, 172, 173
opaque data, 178, 179
overview, 170
pointers, 181, 183
strings, 64, 174, 175
unions, 180, 181
remote copy (two-way XDR)
routine, 123, 126
rpcgen tool and, 30, 35, 163
streams
accessing, 184
creation by RPC system, 168
implementing new instances
of, 186, 188
interface to, 186, 188
memory, 184
nonfilter primitives for, 183
record (TCP/IP), 184, 186, 202, 203
standard I/O, 184
transport handles and, 66, 69
XDR (external data representation), see
XDR language
protocol,
XDR language, 234, 250
arrays, 242, 243, 249
AUTH_DES authentication protocol
in, 207, 210
booleans, 237
comments, 248
constants, 245, 248, 249
counted byte strings, 241, 242
declarations, 234, 247
discriminated unions, 244, 246, 250
enumerations, 236
file data structure described in, 250
fixed-length arrays, 242
fixed-length opaque data, 240
floating point, 238, 240
hyper integers, 237
identifiers, 248
integers, 165, 166, 235, 237
keywords, 249
opaque data, 240, 241
optional-data unions, 246
overview, 233, 234, 310
quadruple-precision floating
point, 239, 240
RPC language vs., 214, 216, 251
RPC message protocol in, 199, 202
signed integers, 235
specification for, 247, 250
strings, 241, 242
structures, 244, 250
syntax, 249, 250
type definitions, 246, 249
unions, 244, 246, 250
unsigned hyper integers, 237
unsigned integers, 235
variable-length arrays, 243
variable-length opaque data, 240, 241
voids, 245
xdr_prefix, 31
xdr_array routine, 176, 178, 182
xdr_bool routine, 62
xdr_bytes routine, 64, 175
xdr_char routine, 62, 173
xdr_chararr routines, 126, 127
xdr_cnd routine, 178
XDR_DECODE operation, 174
xdr_destroy routine, 183
xdr_double routine, 62, 173
xdr_element routine, 176
XDR_ENCODE operation, 174
xdr_enum routine, 62
xdr_float routine, 62, 173
XDR_FREE operation, 174

Index-335
xdr_free routine, 35, 44
xdr_getpos routine, 183
xdr_gnumbers routine, 169, 188, 191
xdr_history routine, 178
xdr_inline count, 37, 49
xdr_int routine, 62, 173
xdr_long routine, 62, 165, 166, 168
xdr_netobj routine, 62, 179
xdrOpaque routine, 178
xdr_party routine, 177
xdr_pgn routine, 182
xdr_pointer routine, 183, 189, 191
xdr_quadruple routine, 62
xdr_rcp routine, 123, 126, 134
xdr_reference routine, 64, 181, 183, 191
xdr_setpos routine, 183
xdr_short routine, 62
xdr_simple routine, 63, 65
xdr_sizeof routine, 170, 172
xdr_string routine, 63, 64, 174, 178
xdr_type(object) notation, 98
xdr_u_char routine, 62, 173
xdr_u_int routine, 62, 173
xdr_u_long routine, 62
xdr_u_short routine, 62
xdr_union routine, 180, 181
xdr_varintarr routine, 64
xdr_vector routine, 64, 179, 180
xdr_void routine, 62, 174
xdr_wrapstring routine, 62, 178
xdrmem_create routine, 184
xdrrec_create routine, 185
xdrrec_endofrecord routine, 186
xdrrec_eof routine, 186
xdrrec_skiprecord routine, 186
xdrs-x_op field, 183
xdrstdio_create routine, 168, 184
xp_fd field, 80
xp_ltaddr field, 80
xp_netid field, 80
xp_rtaddr field, 81
xp_tp field, 81
xprt_register routine, 132
xprt_unregister routine, 132