



Sun Cluster データサービス開発ガイド (Solaris OS 版)

Sun Microsystems, Inc.
4150 Network Circle
Santa Clara, CA 95054
U.S.A.

Part No: 819-0185-10
2004 年 9 月, Revision A

Copyright 2004 Sun Microsystems, Inc. 4150 Network Circle, Santa Clara, CA 95054 U.S.A. All rights reserved.

本製品およびそれに関連する文書は著作権法により保護されており、その使用、複製、頒布および逆コンパイルを制限するライセンスのもとにおいて頒布されます。サン・マイクロシステムズ株式会社による事前の許可なく、本製品および関連する文書のいかなる部分も、いかなる方法によっても複製することが禁じられます。

本製品の一部は、カリフォルニア大学からライセンスされている Berkeley BSD システムに基づいていることがあります。UNIX は、X/Open Company, Ltd. が独占的にライセンスしている米国ならびに他の国における登録商標です。フォント技術を含む第三者のソフトウェアは、著作権により保護されており、提供者からライセンスを受けているものです。

Federal Acquisitions: Commercial Software—Government Users Subject to Standard License Terms and Conditions.

本製品に含まれる HG-MinchoL、HG-MinchoL-Sun、HG-PMinchoL-Sun、HG-GothicB、HG-GothicB-Sun、および HG-PGothicB-Sun は、株式会社リコーがリコービマジクス株式会社からライセンス供与されたタイプフェイスマスタをもとに作成されたものです。HeiseiMin-W3H は、株式会社リコーが財団法人日本規格協会からライセンス供与されたタイプフェイスマスタをもとに作成されたものです。フォントとして無断複製することは禁止されています。

Sun、Sun Microsystems、docs.sun.com、AnswerBook、AnswerBook2、NetBeans、Sun StorEdge、Sun Cluster、SunPlex は、米国およびその他の国における米国 Sun Microsystems, Inc. (以下、米国 Sun Microsystems 社とします) の商標もしくは登録商標です。

サンのロゴマークおよび Solaris は、米国 Sun Microsystems 社の登録商標です。

すべての SPARC 商標は、米国 SPARC International, Inc. のライセンスを受けて使用している同社の米国およびその他の国における商標または登録商標です。SPARC 商標が付いた製品は、米国 Sun Microsystems 社が開発したアーキテクチャに基づくものです。

OPENLOOK、OpenBoot、JLE は、サン・マイクロシステムズ株式会社の登録商標です。

Wnn は、京都大学、株式会社アステック、オムロン株式会社で共同開発されたソフトウェアです。

Wnn6 は、オムロン株式会社、オムロンソフトウェア株式会社で共同開発されたソフトウェアです。© Copyright OMRON Co., Ltd. 1995-2000. All Rights Reserved. © Copyright OMRON SOFTWARE Co., Ltd. 1995-2002 All Rights Reserved.

「ATOK」は、株式会社ジャストシステムの登録商標です。

「ATOK Server/ATOK12」は、株式会社ジャストシステムの著作物であり、「ATOK Server/ATOK12」にかかる著作権その他の権利は、株式会社ジャストシステムおよび各権利者に帰属します。

本製品に含まれる郵便番号辞書 (7 桁/5 桁) は郵政事業庁が公開したデータを元に制作された物です (一部データの加工を行なっています)。

本製品に含まれるフェイスマーク辞書は、株式会社ビレッジセンターの許諾のもと、同社が発行する『インターネット・パソコン通信フェイスマークガイド '98』に添付のものを使用しています。© 1997 ビレッジセンター

Unicode は、Unicode, Inc. の商標です。

本書で参照されている製品やサービスに関しては、該当する会社または組織に直接お問い合わせください。

OPEN LOOK および Sun™ Graphical User Interface は、米国 Sun Microsystems 社が自社のユーザおよびライセンス実施権者向けに開発しました。米国 Sun Microsystems 社は、コンピュータ産業用のビジュアルまたはグラフィカル・ユーザインタフェースの概念の研究開発における米国 Xerox 社の先駆者としての成果を認めるものです。米国 Sun Microsystems 社は米国 Xerox 社から Xerox Graphical User Interface の非独占的ライセンスを取得しており、このライセンスは米国 Sun Microsystems 社のライセンス実施権者にも適用されます。

DiComboBox ウィジェットと DtSpinBox ウィジェットのプログラムおよびドキュメントは、Interleaf, Inc. から提供されたものです。(© 1993 Interleaf, Inc.)

本書は、「現状のまま」をベースとして提供され、商品性、特定目的への適合性または第三者の権利の非侵害の黙示の保証を含みそれに限定されない、明示的であるか黙示的であるかを問わない、なんらの保証も行われぬものとします。

本製品が、外国為替および外国貿易管理法 (外為法) に定められる戦略物資等 (貨物または役務) に該当する場合、本製品を輸出または日本国外へ持ち出す際には、サン・マイクロシステムズ株式会社の事前の書面による承諾を得ることのほか、外為法および関連法規に基づく輸出手続き、また場合によっては、米国商務省または米国所轄官庁の許可を得ることが必要です。

原典: Sun Cluster Data Services Developer's Guide for Solaris OS

Part No: 817-6555-10

Revision A



041202@10536



目次

はじめに	11
1	リソース管理の概要 17
	Sun Cluster アプリケーション環境 17
	RGM モデル 19
	リソースタイプ 19
	リソース 20
	リソースグループ 20
	リソースグループマネージャ 21
	コールバックメソッド 21
	プログラミングインタフェース 22
	RMAPI 22
	データサービス開発ライブラリ (DSDL) 23
	SunPlex Agent Builder 23
	Resource Group Manager の管理インタフェース 24
	SunPlex Manager 24
	管理コマンド 24
2	データサービスの開発 25
	アプリケーションの適合性の分析 25
	使用するインタフェースの決定 27
	データサービス作成用開発環境の設定 28
	▼ 開発環境の設定方法 29
	データサービスをクラスタに転送する方法 30
	リソースとリソースタイププロパティの設定 30
	リソースタイププロパティの宣言 31

リソースプロパティの宣言	34
拡張プロパティの宣言	37
コールバックメソッドの実装	39
リソースとリソースグループのプロパティ情報へのアクセス	39
メソッドの呼び出し回数への非依存性	39
汎用データサービス	40
アプリケーションの制御	40
リソースの起動と停止	40
Init、Fini、Boot の各メソッド	43
リソースの監視	44
メッセージログのリソースへの追加	45
プロセス管理の提供	45
リソースへの管理サポートの提供	46
フェイルオーバーリソースの実装	46
スケーラブルリソースの実装	47
スケーラブルサービスの妥当性検査	50
データサービスの作成と検証	50
キープアライブの使用法	51
HA データサービスの検証	51
リソース間の依存関係の調節	52
3 リソースタイプの更新	55
概要	55
リソースタイプ登録ファイル	56
リソースタイプ名	56
ディレクティブ	57
RTR ファイル内の RT_Version の変更	57
以前のバージョンの Sun Cluster のリソースタイプ名	58
Type_version リソースプロパティ	58
リソースを別のバージョンへ移行	59
リソースタイプのアップグレードとダウングレード	60
▼ リソースタイプをアップグレードする方法	60
▼ 古いバージョンのリソースタイプにダウングレードする方法	61
デフォルトのプロパティ値	62
リソースタイプ開発者の文書	63
リソースタイプ名とリソースタイプモニターの実装	63
アプリケーションのアップグレード	64
リソースタイプのアップグレード例	64

リソースタイプパッケージのインストール要件	68
RTR ファイルの変更前に認識しておくべき事項	68
モニターコードの変更	69
メソッドコードの変更	69
4 Resource Management API リファレンス	71
RMAPI アクセスメソッド	72
RMAPI シェルコマンド	72
C 関数	73
RMAPI コールバックメソッド	77
メソッドの引数	77
終了コード	78
制御および初期化コールバックメソッド	78
管理サポートメソッド	79
ネットワーク関連コールバックメソッド	80
モニター制御コールバックメソッド	80
5 サンプルデータサービス	83
サンプルデータサービスの概要	83
リソースタイプ登録ファイルの定義	84
RTR ファイルの概要	84
サンプル RTR ファイルのリソースタイププロパティ	85
サンプル RTR ファイルのリソースプロパティ	86
すべてのメソッドに共通な機能の提供	90
コマンドインタプリタの指定およびパスのエクスポート	90
PMF_TAG と SYSLOG_TAG 変数の宣言	90
関数の引数の構文解析	91
エラーメッセージの生成	93
プロパティ情報の取得	93
データサービスの制御	94
Start メソッド	94
Stop メソッド	97
障害モニターの定義	100
検証プログラム	101
Monitor_start メソッド	106
Monitor_stop メソッド	107
Monitor_check メソッド	108

プロパティ更新の処理	109
Validate メソッド	109
Update メソッド	114
6 データサービス開発ライブラリ (DSDL) 117	
DSDL の概要	117
構成プロパティの管理	118
データサービスの起動と停止	118
障害モニターの実装	119
ネットワークアドレス情報へのアクセス	120
実装したリソースタイプのデバッグ	120
高可用性ローカルファイルシステムの有効化	121
7 リソースタイプの設計。 123	
RTR ファイル	124
Validate メソッド	124
Start メソッド	126
Stop メソッド	127
Monitor_start メソッド	128
Monitor_stop メソッド	129
Monitor_check メソッド	129
Update メソッド	129
Init、Fini、Boot の各メソッド	130
障害モニターデーモンの設計	131
8 サンプル DSDL リソースタイプの実装 135	
X Font Server について	135
X Font Server の構成ファイル	136
TCP ポート番号	136
命名規約	136
SUNW.xfnts の RTR ファイル	137
scds_initialize() 関数	138
xfnts_start メソッド	138
起動前のサービスの検証	138
サービスの起動	139
svc_start() からの復帰	140
xfnts_stop メソッド	142

xfnts_monitor_start メソッド	144
xfnts_monitor_stop メソッド	145
xfnts_monitor_check メソッド	146
SUNW.xfnts 障害モニター	147
xfnts_probe のメインループ	147
svc_probe() 関数	149
障害モニターのアクションの決定	151
xfnts_validate メソッド	152
xfnts_update メソッド	154

9 SunPlex Agent Builder 157

Agent Builder の概要	157
Agent Builder の使用にあたって	158
複数の独立したプロセスツリーを持つリソースタイプの作成	158
Agent Builder の使用	159
アプリケーションの分析	159
Agent Builder のインストールと構成	160
Agent Builder 画面	161
Agent Builder の起動	161
Agent Builder のナビゲーション	162
作成画面の使用	166
構成画面の使用	168
Agent Builder の Korn シェルベース \$hostnames 変数の使用	171
プロパティ変数	171
完成した作業内容の再利用	173
▼ コマンド行バージョンの Agent Builder を使用する方法	175
ディレクトリ構造	175
Agent Builder の出力	176
ソースファイルとバイナリファイル	176
ユーティリティースクリプトとマニュアルページ	178
サポートファイル	179
パッケージディレクトリ	179
rtconfig ファイル	180
Agent Builder の Cluster Agent モジュール	180
▼ Cluster Agent モジュールをインストールし設定する方法	180
▼ Cluster Agent モジュールを起動する方法	181
Cluster Agent モジュールの使用	183
Cluster Agent モジュールと Agent Builder の違い	184

- 10 汎用データサービス 187
 - GDS の概要 187
 - コンパイル済みリソースタイプ 188
 - GDS を使用することの利点と欠点 188
 - GDS を使用するサービスの作成方法 188
 - GDS によるイベントのロギング 189
 - 必須の GDS プロパティ 190
 - 任意の GDS プロパティ 190
 - SunPlex Agent Builder を使って GDS ベースのサービスを作成 194
 - スクリプトの作成と構成 194
 - SunPlex Agent Builder の出力 198
 - 標準的な Sun Cluster 管理コマンドを使って GDS ベースのサービスを作成 199
 - ▼ Sun Cluster 管理コマンドを使って GDS ベースの高可用性サービスを作成する方法 199
 - ▼ Sun Cluster 管理コマンドを使って GDS ベースのスケラブルサービスを作成する方法 200
 - SunPlex Agent Builder のコマンド行インタフェース 201
 - ▼ コマンド行バージョンの Agent Builder を使用して GDS ベースのサービスを作成する方法 201

- 11 データサービス開発ライブラリのリファレンス 203
 - DSDL 関数 203
 - 汎用関数 203
 - プロパティ関数 205
 - ネットワークリソースアクセス関数 205
 - TCP 接続を使用する障害監視 206
 - PMF 関数 206
 - 障害監視関数 207
 - ユーティリティ関数 207

- 12 CRNP 209
 - CRNP の概要 209
 - CRNP プロトコルの概要 210
 - CRNP が使用するメッセージのタイプ 212
 - クライアントをサーバーに登録する方法 213
 - 管理者によるサーバー設定の前提 213
 - サーバーによるクライアントの識別方法 213
 - クライアントとサーバー間での SC_CALLBACK_REG メッセージの受け渡し方法 213

クライアントに対するサーバーの応答方法	215
SC_REPLY メッセージの内容	216
クライアントによるエラー状況の処理	216
サーバーがクライアントにイベントを配信する方法	217
イベント配信の保証	218
SC_EVENT メッセージの内容	218
CRNP によるクライアントとサーバーの認証	221
CRNP を使用する Java アプリケーションの作成	221
▼ 環境の設定	222
▼ 作業の開始	223
▼ コマンド行引数の解析	224
▼ イベント受信スレッドの定義	225
▼ コールバックの登録と登録解除	226
▼ XML の生成	227
▼ 登録メッセージと登録解除メッセージの作成	230
▼ XML パーサーの設定	232
▼ 登録応答の解析	233
▼ コールバックイベントの解析	235
▼ アプリケーションの実行	238
A 標準プロパティ	239
リソースタイププロパティ	239
リソースのプロパティ	245
リソースグループのプロパティ	257
リソースプロパティの属性	263
B データサービスのコード例	267
リソースタイプ登録ファイルのリスト	267
start メソッド	270
stop メソッド	273
gettime ユーティリティ	275
PROBE プログラム	276
Monitor_start メソッド	282
Monitor_stop メソッド	284
Monitor_check メソッド	285
Validate メソッド	287
Update メソッド	291

- C サンプル DSDL リソースタイプのコード例 293**
 - xfnts.c 293
 - xfnts_monitor_check メソッド 305
 - xfnts_monitor_start メソッド 306
 - xfnts_monitor_stop メソッド 307
 - xfnts_probe メソッド 308
 - xfnts_start メソッド 311
 - xfnts_stop メソッド 312
 - xfnts_update メソッド 313
 - xfnts_validate メソッドのコードリスト 315

- D 有効な RGM 名と値 317**
 - 有効な RGM 名 317
 - 命名規則 (リソースタイプ名を除く) 317
 - リソースタイプ名の形式 318
 - RGM の値 319

- E 非クラスタ対応のアプリケーションの要件 321**
 - 多重ホストデータ 321
 - 多重ホストデータを配置するためのシンボリックリンクの使用 322
 - ホスト名 323
 - 多重ホームホスト 323
 - INADDR_ANY へのバインドと特定の IP アドレスへのバインド 324
 - クライアントの再試行 325

- F CRNP のドキュメントタイプ定義 327**
 - SC_CALLBACK_REG XML DTD 327
 - NVPAIR XML DTD 329
 - SC_REPLY XML DTD 330
 - SC_EVENT XML DTD 331

- G CrnpClient.java アプリケーション 333**
 - CrnpClient.java のコンテンツ 333

索引 355

はじめに

このマニュアルでは、RMAPI (Resource Management (リソース管理) API) を使用して、SPARC® と x86 ベースシステムの両方で Sun™ Cluster データサービスを開発する方法について説明します。

注 - このマニュアルでは、「x86」という用語は、Intel 32 ビット系列のマイクロプロセッサチップ、および AMD が提供する互換マイクロプロセッサチップを意味します。

注 - Sun Cluster ソフトウェアは、SPARC と x86 の 2 つのプラットフォーム上で稼動します。このマニュアル内の情報は、章、節、注、箇条書き項目、図、表、または例などで特に明記されていない限り両方に適用されます。

対象読者

このマニュアルは、Sun のソフトウェアとハードウェアについて豊富な知識を持っている経験のある開発者を対象にしています。このマニュアルの情報は、Solaris™ オペレーティングシステムの知識があることを前提としています。

このマニュアルの構成

このマニュアルは、次の章と付録で構成されています。

- 第 1 章では、データサービスを開発するのに必要な概念について説明します。
- 第 2 章では、データサービスの開発に関する詳細な情報を説明します。
- 第 3 章では、リソースタイプの開発者がリソースタイプの更新や移行を行うために必要な情報を提供します。
- 第 4 章では、Resource Management API (RMAPI) を構成するアクセス関数とコールバックメソッドに関する情報を説明します。
- 第 5 章では、`in.named()` アプリケーション用の Sun Cluster データサービスの例を示します。
- 第 6 章では、Data Services Development Library (DSDL) を形成するアプリケーションプログラミングインタフェースの概要を説明します。
- 第 7 章では、リソースタイプの設計と実装における DSDL の代表的な使用例について説明します。
- 第 8 章では、DSDL により実装されるリソースタイプの例を説明します。
- 第 9 章では、SunPlex™ Agent Builder について説明します。
- 第 10 章では、一般的なデータサービスの作成方法について説明します。
- 第 11 章では、DSDL API 関数について説明します。
- 第 12 章では、Cluster Reconfiguration Notification Protocol (CRNP) について説明します。CRNP を使用することで、フェイルオーバー用のアプリケーションや拡張性のあるアプリケーションを「クラスタ対応」として設定できます。
- 付録 A では、標準リソースタイプ、リソースグループ、およびリソースプロパティについて説明します。
- 付録 B では、データサービスの例について、それぞれのメソッドの完全なコードを示します。
- 付録 C では、`SUNW.xfnts()` リソースタイプにおける各メソッドの完全なコードを示します。
- 付録 D では、Resource Group Manager (RGM) の名前と値についての文字の要件を説明します。
- 付録 E では、クラスタに対応していない、通常アプリケーションを高可用性に適用させる要件を説明します。
- 付録 F では、CRNP のドキュメントタイプ定義を説明します。
- 付録 G では、第 12 章で説明されている `CrnpClient.java` の完全なアプリケーションを示します。

関連マニュアル

関連する Sun Cluster トピックについての情報は、以下の表に示すマニュアルを参照してください。すべての Sun Cluster マニュアルは、<http://docs.sun.com> で参照できます。

トピック	マニュアル
概要	『Sun Cluster の概要 (Solaris OS 版)』
概念	『Sun Cluster の概念 (Solaris OS 版)』
ハードウェアの設計と管理	『Sun Cluster 3.x Hardware Administration Manual for Solaris OS』 各ハードウェア管理ガイド
ソフトウェアのインストール	『Sun Cluster ソフトウェアのインストール (Solaris OS 版)』
データサービスのインストールと管理	『Sun Cluster データサービスの計画と管理 (Solaris OS 版)』 各データサービスガイド
データサービスの開発	『Sun Cluster データサービス開発ガイド (Solaris OS 版)』
システム管理	『Sun Cluster のシステム管理 (Solaris OS 版)』
エラーメッセージ	『Sun Cluster Error Messages Guide for Solaris OS』
コマンドと関数のリファレンス	『Sun Cluster Reference Manual for Solaris OS』

Sun Cluster のマニュアルの完全なリストについては、お使いの Sun Cluster ソフトウェアのリリースノートをお使いの <http://docs.sun.com> で参照してください。

Sun Cluster マニュアルの完全なリストについては、ご使用の Sun Cluster リリースのリリースノートをお参照してください。リリースノートは、<http://docs.sun.com> にあります。

問い合わせについて

Sun Cluster のインストールまたは使用で問題が発生した場合は、ご購入先に連絡し、次の情報をお伝えください。

- 名前と電子メールアドレス (利用している場合)

- 会社名、住所、および電話番号
- システムのモデルとシリアル番号
- オペレーティングシステムのバージョン番号 (例: Solaris 10)
- Sun Cluster のバージョン番号 (例: Sun Cluster 3.1)

ご購入先に知らせる、システム上の各ノードについての情報を収集するには、次のコマンドを使用します。

コマンド	機能
<code>prtconf -v</code>	システムメモリのサイズと周辺デバイス情報を表示します
<code>psrinfo -v</code>	プロセッサの情報を表示する
<code>showrev -p</code>	インストールされているパッチを報告する
<code>SPARC: prtdiag -v</code>	システム診断情報を表示する
<code>/usr/cluster/bin/scinstall -pv</code>	Sun Cluster のリリースとパッケージバージョン情報を表示する

上記の情報にあわせて、`/var/adm/messages` ファイルの内容もご購入先にお知らせください。

Sun のオンラインマニュアル

`docs.sun.com` では、Sun が提供しているオンラインマニュアルを参照することができます。マニュアルのタイトルや特定の主題などをキーワードとして、検索を行うこともできます。URL は、`http://docs.sun.com` です。

表記上の規則

このマニュアルでは、次のような字体や記号を特別な意味を持つものとして使用します。

表 P-1 表記上の規則

字体または記号	意味	例
AaBbCc123	コマンド名、ファイル名、ディレクトリ名、画面上のコンピュータ出力、コード例を示します。	.login ファイルを編集します。 ls -a を使用してすべてのファイルを表示します。 system%
AaBbCc123	ユーザーが入力する文字を、画面上のコンピュータ出力と区別して示します。	system% su password:
AaBbCc123	変数を示します。実際に使用する特定の名前または値で置き換えます。	ファイルを削除するには、rm <i>filename</i> と入力します。
『 』	参照する書名を示します。	『コードマネージャ・ユーザーズガイド』を参照してください。
「 」	参照する章、節、ボタンやメニュー名、強調する単語を示します。	第 5 章「衝突の回避」を参照してください。 この操作ができるのは、「スーパーユーザー」だけです。
\	枠で囲まれたコード例で、テキストがページ行幅を超える場合に、継続を示します。	sun% grep `^#define \ XV_VERSION_STRING`

コード例は次のように表示されます。

■ C シェル

```
machine_name% command y|n [filename]
```

■ C シェルのスーパーユーザー

```
machine_name# command y|n [filename]
```

■ Bourne シェルおよび Korn シェル

```
$ command y|n [filename]
```

■ Bourne シェルおよび Korn シェルのスーパーユーザー

```
# command y|n [filename]
```

[] は省略可能な項目を示します。上記の例は、*filename* は省略してもよいことを示しています。

| は区切り文字 (セパレータ) です。この文字で分割されている引数のうち 1 つだけを指定します。

キーボードのキー名は英文で、頭文字を大文字で示します (例: Shift キーを押します)。ただし、キーボードによっては Enter キーが Return キーの動作をします。

ダッシュ (-) は 2 つのキーを同時に押すことを示します。たとえば、Ctrl-D は Control キーを押したまま D キーを押すことを意味します。

第 1 章

リソース管理の概要

このマニュアルでは、Oracle®、Sun Java™ System Web Server (以前の Sun™ ONE Web Server)、DNS などのソフトウェアアプリケーション用のリソースタイプを作成するためのガイドラインを説明します。したがって、このマニュアルはリソースタイプの開発者を対象としています。

この章では、データサービスを開発するために理解しておく必要がある概念について説明します。この章の内容は、次のとおりです。

- 17 ページの「Sun Cluster アプリケーション環境」
- 19 ページの「RGM モデル」
- 21 ページの「リソースグループマネージャ」
- 21 ページの「コールバックメソッド」
- 22 ページの「プログラミングインタフェース」
- 24 ページの「Resource Group Manager の管理インタフェース」

注 - このマニュアルでは、「リソースタイプ」と「データサービス」という用語を同じ意味で使用しています。また、このマニュアルではほとんど使用されることはありませんが、「エージェント」という用語も「リソースタイプ」や「データサービス」と同じ意味で使用されます。

Sun Cluster アプリケーション環境

Sun Cluster システムを使用すると、アプリケーションを高度な可用性とスケーラビリティを備えたリソースとして実行および管理できます。RGM (Resource Group Manager) というクラスタ機能は、高可用性とスケーラビリティを実現するための機構を提供します。この機能を利用するためのプログラミングインタフェースを形成する要素は、次のとおりです。

- ユーザーが作成するコールバックメソッド。RGM は、このコールバックメソッドを利用してクラスタ上のアプリケーションを制御します。

- Resource Management API (RMAPI)。コールバックメソッドの作成に使用する低レベルの API コマンドおよび API 関数です。これらの API は、libscha.so ライブラリに実装されます。
- クラスタ上のプロセスを監視および再起動するプロセス管理機能。
- データサービス開発ライブラリ (Data Service Development Library: DSDL)。低レベル API およびプロセス管理機能をより高レベルでカプセル化し、コールバックメソッドの作成を支援するいくつかの機能を追加するライブラリ関数です。これらの関数は、libdsdev.so ライブラリに実装されます。

次の図は、これらの要素の相互関係を示しています。

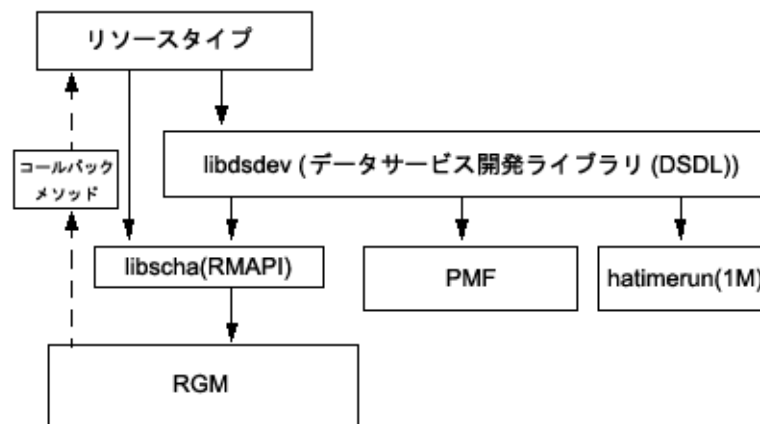


図 1-1 プログラミングアーキテクチャ

Sun Cluster パッケージには、データサービスの作成プロセスを自動化する SunPlex™ Agent Builder というツールが含まれています (第 9 章 を参照)。Agent Builder はデータサービスのコードを C 言語または Korn シェル (ksh) のどちらでも生成できます。前者の場合は DSDL 関数、後者の場合は低レベルの API コマンドを使ってコールバックメソッドを作成します。

RGM は各クラスタ上でデーモンとして動作して、事前構成したポリシーに従って、選択したノード上のリソースを自動的に起動および停止します。リソースの高可用性を実現するために、RGM は、ノードが異常終了または再起動すると、影響を受けるノード上でリソースを停止し、別のノード上でリソースを起動します。また、リソースに固有のモニター (監視機能) を起動および停止することによって、障害のあるリソースを検出し、別のノードに再配置したり、さまざまな視点からリソース性能を監視します。

RGM はフェイルオーバーリソースとスケラブルリソースの両方をサポートします。フェイルオーバーリソースとは、同時に 1 つのノード上だけでオンラインになることができるリソースのことです。スケラブルリソースとは、同時に複数のノード上でオンラインになることができるリソースのことです。

RGM モデル

ここでは、基本的な用語をいくつか紹介し、RGM とそれに関連するインタフェースについて詳細に説明します。

RGM は、「リソースタイプ」、「リソース」、「リソースグループ」という 3 種類の相互に関連するオブジェクトを処理します。これらのオブジェクトを紹介するために、次のような例を使用します。

開発者は、既存の Oracle DBMS アプリケーションを高可用性にするためのリソースタイプ `ha-oracle` を実装します。エンドユーザーは、マーケティング、エンジニアリング、および財務ごとに異なるデータベースを定義し、それぞれのリソースタイプを `ha-oracle` にします。クラスタ管理者は、上記リソースを異なるリソースグループに分類することによって、異なるノード上で実行したり、個別にフェイルオーバーできるようにします。開発者は、もう 1 つのリソースタイプ `ha-calendar` を作成し、Oracle データベースを必要とする高可用性のカレンダーサーバーを実装します。クラスタ管理者は、財務カレンダーリソースと財務データベースリソースを同じリソースグループに分類することによって、両方のリソースを同じノード上で実行したり、一緒にフェイルオーバーできるようにします。

リソースタイプ

リソースタイプは、クラスタ上で実行されるソフトウェアアプリケーション、アプリケーションをクラスタリソースとして管理するために RGM がコールバックメソッドとして使用する制御プログラムおよびクラスタの静的な構成の一部を形成するプロパティセットからなります。RGM は、リソースタイププロパティを使って特定のタイプのリソースを管理します。

注 - リソースタイプは、ソフトウェアアプリケーションだけでなく、その他のシステムリソース (ネットワークアドレスなど) も表します。

リソースタイプの開発者は、リソースタイププロパティを指定し、その値をリソースタイプ登録 (RTR) ファイルに設定します。RTR ファイルの形式は、30 ページの「リソースとリソースタイププロパティの設定」と `rt_reg(4)` のマニュアルページの記述に従い、明確に定義されています。リソースタイプ登録ファイルの例については、84 ページの「リソースタイプ登録ファイルの定義」を参照してください。

239 ページの「リソースタイププロパティ」に、リソースタイププロパティのリストを示します。

クラスタ管理者は、リソースタイプの実装と実際のアプリケーションをクラスタにインストールし、登録します。さらに登録手順で、リソースタイプ登録ファイルの情報をクラスタ構成に入力します。データサービスの登録手順については、『*Sun Cluster データサービスの計画と管理 (Solaris OS 版)*』を参照してください。

リソース

リソースは、そのリソースタイプからプロパティと値を継承します。さらに、開発者は、リソースタイプ登録ファイルでリソースプロパティを宣言できます。[245 ページ](#)の「[リソースのプロパティ](#)」にリソースプロパティのリストが記載されています。

クラスタ管理者は、リソースタイプ登録 (RTR) ファイルにプロパティを指定することによって、特定のプロパティの値を変更できます。たとえば、プロパティ定義に値の許容範囲を指定しておきます。これにより、プロパティが調節可能なときに、作成時、常時、不可などを指定できます。このような許容範囲内であれば、クラスタ管理者は管理コマンドでプロパティを変更できます。

クラスタ管理者は、同じタイプのリソースを多数作成して、各リソースに独自の名前とプロパティ値のセットを持たせることができます。これによって、実際のアプリケーションの複数のインスタンスをクラスタ上で実行できます。このとき、各インスタンスにはクラスタ内で一意の名前が必要です。

リソースグループ

各リソースはリソースグループに構成する必要があります。RGM は、同じグループのすべてのリソースを同じノード上でオンラインかオフラインにします。このとき、グループ内の個々のリソースに対してコールバックメソッドを呼び出します。

リソースグループがオンラインになっているノードを主ノードと呼びます。リソースグループは、その主ノードによってマスター (制御) されます。各リソースグループは、クラスタ管理者が設定した独自の `Nodelist` プロパティを持っており、これによってリソースグループの潜在的な主ノードを識別します。

リソースグループはプロパティセットも持っています。このようなプロパティには、クラスタ管理者が設定できる構成プロパティや、RGM が設定してリソースグループのアクティブな状態を反映する動的プロパティがあります。

RGM は、2 種類のリソースグループ、フェイルオーバー (failover) とスケーラブル (scalable) を定義します。フェイルオーバーリソースグループは、同時に 1 つのノード上だけでオンラインになることができます。一方、スケーラブルリソースグループは、同時に複数のノード上でオンラインになることができます。RGM は、各種類のリソースグループを作成するためのプロパティセットを提供します。このようなプロパティの詳細については、[30 ページ](#)の「[データサービスをクラスタに転送する方法](#)」と [39 ページ](#)の「[コールバックメソッドの実装](#)」を参照してください。

[257 ページ](#)の「[リソースグループのプロパティ](#)」にリソースグループプロパティのリストが記載されています。

リソースグループマネージャ

Resource Group Manager (RGM) は `rgmd` デモンとして実装され、クラスタの各メンバー (ノード) 上で動作します。 `rgmd` プロセスはすべて互いに通信し、単一のクラスタ規模の機能として動作します。

RGM は、次の機能をサポートします。

- ノードが起動またはクラッシュしたとき、管理対象のすべてのリソースグループを適切なマスター上で自動的にオンラインにし、その可用性を維持します。
- 特定のリソースが異常終了した場合、そのモニタープログラムはリソースグループを同じマスター上で再起動するか、新しいマスターに切り替えるかを要求できません。
- クラスタ管理者は管理コマンドを発行して、次のいずれかのアクションを要求できます。
 - リソースグループをマスターする権利の変更
 - リソースグループ内の特定のリソースの有効化または無効化
 - リソース、リソースグループ、リソースタイプの作成、削除、変更

RGM は、構成を変更するとき、そのアクションをクラスタのすべてのメンバー (ノード) 間で調整します。このような動作を「再構成」と呼びます。RGM は、個々のリソースの状態を変更をするため、各リソース上でリソースタイプに固有のコールバックメソッドを呼び出します。

コールバックメソッド

Sun Cluster フレームワークは、コールバックメソッドを使用して、データサービスと RGM 間の通信を実現します。また、コールバックメソッド (引数と戻り値を含む) のセットと、RGM が個々のメソッドを呼び出す環境を定義します。

データサービスを作成するには、個々のコールバックメソッドのセットをコーディングし、個々のメソッドを RGM から呼び出し可能な制御プログラムとして実装します。つまり、データサービスは、単一の実行可能コードではなく、多数の実行可能なスクリプト (`ksh`) またはバイナリ (C 言語) から構成されており、それぞれを RGM から直接呼び出すことができます。

コールバックメソッドを RGM に登録するには、リソースタイプ登録 (RTR) ファイルを使用します。RTR ファイルには、データサービスとして実装した各メソッドのプログラムを指定します。システム管理者がデータサービスをクラスタに登録すると、RGM は RTR ファイルにあるさまざまな情報の中からコールバックプログラムの識別情報を読み取ります。

リソースタイプの必須コールバックメソッドは、起動メソッド (Start または Prenet_start) と停止メソッド (Stop または Postnet_stop) だけです。

コールバックメソッドは、次のようなカテゴリに分類できます。

- 制御および初期化メソッド
 - Start と Stop は、オンラインまたはオフラインにするグループ内のリソースを起動または停止します。
 - Init、Fini、Boot は、リソース上で初期化と終了コードを実行します。
- 管理サポートメソッド
 - Validate は、管理アクションによって設定されるプロパティを確認します。
 - Update は、オンラインリソースのプロパティ設定を更新します。
- ネットワーク関連メソッド
 - Prenet_start と Postnet_stop は、同じリソースグループ内のネットワークアドレスが「起動」に構成される前、または「停止」に構成された後に、特別な起動アクションまたは停止アクションを実行します。
- モニター制御メソッド
 - Monitor_start と Monitor_stop は、リソースのモニターを起動または停止します。
 - Monitor_check は、リソースグループがノードに移動される前に、ノードの信頼性を査定します。

コールバックメソッドの詳細については第 4 章と `rt_callbacks(1HA)` のマニュアルページを参照してください。コールバックメソッドの使用例については、第 5 章および第 8 章を参照してください。

プログラミングインタフェース

リソース管理アーキテクチャは、データサービス用のコードを作成するため、低レベルのベース API、ベース API 上のものより高いレベルのライブラリを提供します。さらに、いくつかの基本的な入力情報をもとにデータサービスを自動的に生成するツール、SunPlex Agent Builder を提供します。

RMAPI

データサービスは、RMAPI (Resource Management API) の低レベルルーチンを使って、システム内のリソース、リソースタイプ、リソースグループの情報にアクセスします。ローカルの再起動やフェイルオーバーの要求、リソースの状態の設定も行います。これらの関数にアクセスするには、`libscha.so` ライブラリを使用します。

RMAPI は、これらのコールバックメソッドを、シェルコマンドまたは C 関数の形で提供できます。RMAPI ルーチンの詳細については、`scha_calls(3HA)` のマニュアルページと第 4 章を参照してください。データサービスコールバックメソッドにおける RMAPI の使用例については、第 5 章を参照してください。

データサービス開発ライブラリ (DSDL)

データサービス開発ライブラリ (Data Service Development Library: DSDL) は、RMAPI 上に構築されており、RGM のメソッドコールバックモデルを基盤にして上位レベルの統合フレームワークを提供します。DSDL は、次のようなさまざまなデータサービス開発向けの機能を提供します。

- `libscha.so`—低レベルのリソース管理 API
- PMF—プロセスとその子孫の監視、停止したプロセスの再起動などを実行するプロセス管理機能 (`pmfadm(1M)` および `rpc.pmf(1M)` のマニュアルページを参照)
- `hatimerun`—タイムアウトを適用してプログラムを実行するための機能 (`hatimerun(1M)` のマニュアルページを参照)

DSDL は、大多数のアプリケーションに対して、データサービスの構築に必要なほとんどまたはすべての機能を提供します。DSDL は、低レベルの API の代わりになるのではなく、低レベルの API をカプセル化および拡張するためのものであることに注意してください。実際、多くの DSDL 関数は `libscha.so` 関数を呼び出します。`libscha.so` 関数を直接呼び出すこともできますが、DSDL を使用することにより、データサービスの大部分を作成できます。DSDL 関数は `libdsdev.so` ライブラリとして実装されています。

DSDL の詳細については、第 6 章と `scha_calls(3HA)` のマニュアルページを参照してください。

SunPlex Agent Builder

Agent Builder は、データサービスの作成を自動化するツールです。このツールでは、ターゲットアプリケーションと作成するデータサービスについての基本的な情報を入力します。Agent Builder は、ソースコードと実行コード (C または Korn シェル)、カスタマイズされた RTR ファイル、Solaris™ パッケージを利用してデータサービスを作成します。

大多数のアプリケーションでは、Agent Builder を使用することにより、わずかなコードを手作業で変更するだけで完全なデータサービスを生成できます。追加プロパティの妥当性検査を必要とするような、より要件の厳しいアプリケーションには、Agent Builder では対応できないこともあります。しかし、このような場合でもコードの大部分を生成できるので、手作業によるコーディングは残りの部分だけで済みます。Agent Builder を使用すれば、少なくとも独自の Solaris パッケージを生成することができます。

Resource Group Manager の管理インタフェース

Sun Cluster はクラスタを管理するために、グラフィカルユーザーインタフェースとコマンドセットの両方を提供します。

SunPlex Manager

SunPlex Manager は、次の作業を実行できる Web ベースのツールです。

- クラスタのインストール
- クラスタの管理
- リソースおよびリソースグループの作成と構成
- Sun Cluster ソフトウェアを使ったデータサービスの構成

SunPlex Manager のインストール方法、SunPlex Manager によるクラスタソフトウェアのインストール方法については、『*Sun Cluster* ソフトウェアのインストール (Solaris OS 版)』を参照してください。管理作業については、SunPlex Manager のオンラインヘルプを参照してください。

管理コマンド

RGM オブジェクトの管理用 Sun Cluster コマンドは、`scrgadm(1M)`、`scswitch(1M)`、および `scstat(1M) -g` の 3 つです。

`scrgadm` コマンドでは、RGM が使用するリソースタイプ、リソースグループ、リソースオブジェクトの表示、作成、構成が可能です。このコマンドはクラスタの管理インタフェースの一部であり、この章の残りで説明しているアプリケーションインタフェースとは異なったプログラミングコンテキストで使用されます。しかし、このコマンドを使って、API が動作するクラスタ構成を構築することもできます。管理インタフェースを理解すると、アプリケーションインタフェースも理解しやすくなります。このコマンドで実行できる管理作業の詳細については、`scrgadm(1M)` のマニュアルページを参照してください。

`scswitch` コマンドでは、指定のノード上のリソースグループのオンラインとオフラインの切り替えや、リソースまたはそのモニターの有効と無効の切り替えが可能です。`scswitch` コマンドで実行できる管理作業の詳細については、`scswitch(1M)` のマニュアルページを参照してください。

`scstat -g` コマンドでは、すべてのリソースグループおよびリソースの現在の動的な状態を表示できます。

第 2 章

データサービスの開発

この章では、データサービスを開発するための詳細な方法について説明します。

この章の内容は次のとおりです。

- 25 ページの「アプリケーションの適合性の分析」
- 27 ページの「使用するインタフェースの決定」
- 28 ページの「データサービス作成用開発環境の設定」
- 30 ページの「リソースとリソースタイププロパティの設定」
- 39 ページの「コールバックメソッドの実装」
- 40 ページの「汎用データサービス」
- 40 ページの「アプリケーションの制御」
- 44 ページの「リソースの監視」
- 45 ページの「メッセージログのリソースへの追加」
- 45 ページの「プロセス管理の提供」
- 46 ページの「リソースへの管理サポートの提供」
- 46 ページの「フェイルオーバーリソースの実装」
- 47 ページの「スケーラブルリソースの実装」
- 50 ページの「データサービスの作成と検証」

アプリケーションの適合性の分析

データサービスを作成するための最初の手順では、ターゲットアプリケーションが高可用性またはスケーラビリティを備えるための要件を満たしているかどうかを判定します。すべての要件を満たしていない場合は、要件を満たすようにアプリケーションのソースコードを変更します。

次に、アプリケーションが高可用性またはスケーラビリティを備えるための要件を要約します。要件に関する詳細情報を確認したい場合や、アプリケーションのソースコードを変更する必要がある場合は、[付録 B](#) を参照してください。

注 - スケーラブルサービスを実現するためには、次に示す高可用性の要件をすべて満たした上で、いくつかの追加要件も満たしている必要があります。

- Sun Cluster 環境では、ネットワーク対応 (クライアントサーバーモデル) とネットワーク非対応 (クライアントレス) のアプリケーションはどちらも、高可用性またはスケーラビリティを備えることが可能です。ただし、タイムシェアリング環境では、アプリケーションはサーバー上で動作し、telnet または rlogin 経由でアクセスされるため、Sun Cluster の可用性を強化することはできません。
- アプリケーションはクラッシュに対する耐障害性 (クラッシュトレラント) を備えていなければなりません。つまり、ノードが予期せぬ停止状態になった後、アプリケーションは再起動時に必要なディスクデータを復元できなければなりません。さらに、クラッシュ後の復元時間にも制限が課せられます。ディスクを復元し、アプリケーションを再起動できる能力は、データの整合性に関わる問題であるため、クラッシュトレラントであることは、アプリケーションが高可用性を備えるための前提条件となります。データサービスは接続を復元できる必要はありません。
- アプリケーションは、自身が動作するノードの物理的なホスト名に依存してはなりません。詳細については、[323 ページの「ホスト名」](#)を参照してください。
- アプリケーションは、複数の IP アドレスが構成されている環境で正しく動作する必要があります。たとえば、ノードが複数のパブリックネットワーク上に存在する多重ホームホスト環境や、単一のハードウェアインタフェース上に複数の論理インタフェースが構成されているノードが存在する環境で正しく動作しなければなりません。
- 高可用性を備えるには、アプリケーションデータはクラスタファイルシステム内に格納されている必要があります。[321 ページの「多重ホストデータ」](#)を参照してください。

アプリケーションがデータの格納先を示すのに固定されたパス名を使用している場合、アプリケーションのソースコードを変更しなくても、そのパスをクラスタファイルシステム内の場所を指すシンボリックリンクに変更できる場合があります。詳細については、[322 ページの「多重ホストデータを配置するためのシンボリックリンクの使用」](#)を参照してください。
- アプリケーションのバイナリとライブラリは、ローカルの各ノードまたはクラスタファイルシステムのどちらにも格納できます。クラスタファイルシステム上に格納する利点は、1 箇所にインストールするだけで済む点です。欠点としては、アプリケーションが RGM の制御下で動作している間はバイナリファイルが使用中になるので、ローリングアップグレードの問題が生じることが挙げられます。
- 初回の照会がタイムアウトした場合、クライアントは自動的に照会を再試行できる必要があります。アプリケーションとプロトコルがすでに単一サーバーのクラッシュと再起動に対応できている場合、関連するリソースグループのフェイルオーバーまたはスイッチオーバーにも対応する必要があります。詳細については、[325 ページの「クライアントの再試行」](#)を参照してください。
- アプリケーションは、クラスタファイルシステム内で UNIX[®] ドメインソケットまたは名前付きパイプを使用してはなりません。

さらに、スケーラブルサービスは、次の要件も満たしている必要があります。

- アプリケーションは、複数のインスタンスを実行でき、すべてのインスタンスがクラスタファイルシステム内の同じアプリケーションデータを処理できる必要があります。
- アプリケーションは、複数のノードからの同時アクセスに対してデータの整合性を保証する必要があります。
- アプリケーションは、クラスタファイルシステムのように、広域的に使用可能な機構を備えたロック機能を実装している必要があります。

スケーラブルサービスの場合、アプリケーションの特性により負荷均衡ポリシーが決定されます。たとえば、負荷均衡ポリシー LB_WEIGHTED は、任意のインスタンスがクライアントの要求に応答できるポリシーですが、クライアント接続にサーバー上のメモリー内キャッシュを使用するアプリケーションには適用されません。この場合、特定のクライアントのトラフィックをアプリケーションの1つのインスタンスに制限する負荷均衡ポリシーを指定する必要があります。負荷均衡ポリシー LB_STICKY と LB_STICKY_WILD は、クライアントからのすべての要求を同じアプリケーションインスタンスに繰り返して送信します。この場合、アプリケーションはメモリー内キャッシュを使用できます。異なるクライアントから複数の要求が送信された場合、RGM はサービスの複数のインスタンスに要求を分配します。スケーラブルデータサービスに対応した負荷均衡ポリシーを設定する方法については、[46 ページ](#)の「フェイルオーバーリソースの実装」を参照してください。

使用するインタフェースの決定

Sun Cluster 開発者サポートパッケージ (SUNWscdev) は、データサービスメソッドのコーディング用に 2 種類のインタフェースセットを提供します。

- Resource Management API (RMAPI) - 低レベルのルーチンセット (libscha.so ライブラリとして実装されている)
- データサービス開発ライブラリ (Data Services Development Library: DSDL) - RMAPI の機能をカプセル化し、いくつかの追加機能を提供する、より高いレベルの関数セット (libdsdev.so ライブラリとして実装されている)

Sun Cluster 開発者サポートパッケージには、データサービスの作成を自動化するツールである SunPlex Agent Builder も含まれています。

次に、データサービスを開発する際の推奨手順を示します。

1. C 言語または Korn シェルのどちらでコーディングするかを決定します。DSDL は C 言語用のインタフェースしか提供しないため、Korn シェルでコーディングする場合は使用できません。
2. Agent Builder を実行します。必要な情報を入力するだけで、ソースコード、実行可能コード、RTR ファイル、パッケージを含むデータサービスを生成できます。
3. 生成されたデータサービスをカスタマイズする必要がある場合は、生成されたソースファイルに DSDL コードを追加できます。Agent Builder は、ソースファイル内において独自のコードを追加できる場所にコメント文を埋め込みます。

4. ターゲットアプリケーションをサポートするために、さらにコードをカスタマイズする必要がある場合は、既存のソースコードに RMAPI 関数を追加できます。

実際には、データサービスを作成する方法はいくつもあります。たとえば、Agent Builder によって生成されたコード内の特定の場所に独自のコードを追加する代わりに、生成されたメソッドの 1 つや生成された監視プログラムを DSDL や RMAPI 関数を使って最初から作成したプログラムで完全に置き換えることができます。しかし、使用方法に関わらず、ほとんどの場合は Agent Builder を使って開発作業を開始することをお勧めします。次に、その理由を示します。

- Agent Builder が生成するコードは本質的に汎用であり、多数のデータサービスでテストされています。
- Agent Builder は、RTR ファイル、make ファイル、リソースのパッケージなど、データサービス用のサポートファイルを作成します。データサービスのコードをまったく使用しない場合でも、このようなサポートファイルを使用することによってかなりの作業を省略できます。
- 生成されたコードは変更できます。

注 - RMAPI は C 言語用の関数セットとスクリプト用のコマンドセットを提供しますが、DSDL は C 言語用の関数インタフェースしか提供しません。DSDL は ksh コマンドを提供しないので、Agent Builder で Korn shell (ksh) 出力を指定した場合、生成されるソースコードは RMAPI を呼び出します。

データサービス作成用開発環境の設定

データサービスの開発を始める前に、Sun Cluster 開発パッケージ (SUNWscdev) をインストールして、Sun Cluster のヘッダーファイルやライブラリファイルにアクセスできるようにする必要があります。このパッケージがすでにすべてのクラスタノード上にインストールされている場合でも、通常は、クラスタノード上にはない独立した開発マシン、すなわちクラスタノード以外の開発マシンで開発を行います。このような場合、pkgadd を使って、開発マシンに SUNWscdev パッケージをインストールする必要があります。

コードをコンパイルおよびリンクするとき、ヘッダーファイルとライブラリファイルを識別するオプションを設定する必要があります。

注 - 互換モードでコンパイルされた C++ コードと標準モードでコンパイルされた C++ コードを Solaris オペレーティングシステム製品や Sun Cluster 製品で併用することはできません。したがって、Sun Cluster で使用する C++ ベースのデータサービスを作成する場合は、次の点に注意してください。

- Sun Cluster 3.0 以前のバージョンで使用する場合は、互換モードでコンパイルする必要があります。
 - Sun Cluster 3.1 以降のバージョンで使用する場合は、標準モードでコンパイルする必要があります。
-

(クラスタノード以外の) 開発マシンで開発が終了すると、完成したデータサービスをクラスタに転送して、実行および検証できます。

注 - Solaris オペレーティングシステムの Solaris 5.8 またはそれ以降の Developer または Entire Distribution ソフトウェアグループを使用してください。

この節では、次の手順を使用します。

- Sun Cluster 開発パッケージ (SUNWscdev) をインストールして、適切なコンパイラオプションとリンカーオプションを設定します。
- データサービスをクラスタに転送します。

▼ 開発環境の設定方法

SUNWscdev パッケージをインストールして、コンパイラオプションとリンカーオプションをデータサービス開発用に設定する方法について説明します。

1. スーパーユーザーになるか、あるいは同等の役割を持つ役割を宣言し、使用したい **CD-ROM** ディレクトリに移動します。

```
# cd CD-ROM_directory
```

2. **SUNWscdev** パッケージを現在のディレクトリにインストールします。

```
# pkgadd -d . SUNWscdev
```

3. **Makefile** に、データサービスのコードが使用する **include** ファイルとライブラリファイルを示すコンパイラオプションとリンカーオプションを指定します。

-I オプションは、Sun Cluster のヘッダーファイルを指定します。-L オプションは、開発システム上にあるコンパイル時ライブラリの検索パスを指定します。-R オプションはクラスタの実行時リンカーの検索パスを指定します。

```
# サンプルデータサービスの Makefile  
...
```

```
-I /usr/cluster/include  
  
-L /usr/cluster/lib  
  
-R /usr/cluster/lib  
...
```

データサービスをクラスタに転送する方法

開発マシン上でデータサービスの開発が完了した場合、クラスタに転送して検証する必要があります。この転送を行うときは、エラーが発生する可能性を減らすために、データサービスのコードと RTR ファイルを一緒にパッケージに保管して、その後、クラスタのすべてのノード上でパッケージをインストールすることを推奨します。

注 - データサービスをインストールするときは、pkgadd を使用するかどうかに関わらず、すべてのクラスタノード上にデータサービスをインストールする必要があります。Agent Builder は、RTR ファイルとデータサービスのコードを自動的にパッケージ化します。

リソースとリソースタイププロパティの設定

Sun Cluster は、データサービスの静的な構成を定義するためのリソースタイププロパティおよびリソースプロパティのセットを提供します。リソースタイププロパティでは、リソースのタイプ、そのバージョン、API のバージョンと同時に、各コールバックメソッドへのパスも指定できます。239 ページの「リソースタイププロパティ」に、すべてのリソースタイププロパティのリストを示します。

リソースプロパティ (Failover mode, Thorough probe interval など) やメソッドタイムアウトも、リソースの静的な構成を定義します。動的なリソースプロパティ (Resource state や Status など) は、管理対象のリソースの活動状況を反映します。リソースプロパティについては、245 ページの「リソースのプロパティ」を参照してください。

リソースタイプおよびリソースプロパティは、データサービスの重要な要素であるリソースタイプ登録 (RTR) ファイルで宣言します。RTR ファイルは、クラスタ管理者が Sun Cluster でデータサービスを登録するとき、データサービスの初期構成を定義します。

Agent Builder が宣言するプロパティセットはどのようなデータサービスにとっても有用かつ必須です。したがって、独自のデータサービス用の RTR ファイルを生成するときは、Agent Builder を使用することを推奨します。たとえば、ある種のプロパ

ティ (Resource_type など) が RTR ファイルで宣言されていない場合、データサービスの登録は失敗します。必須ではなくても、その他のプロパティも RTR ファイルで宣言されていなければ、システム管理者はそれらのプロパティを利用することはできません。いくつかのプロパティは RTR ファイルで宣言されていなくても使用することができますが、これは RGM がそのプロパティを定義し、そのデフォルト値を提供しているためです。このような複雑さを回避するためにも、Agent Builder を使用して、適切な RTR ファイルを生成するようにしてください。必要であれば、Agent Builder で生成した RTR ファイルを編集して、特定の値を変更することもできます。

以降では、Agent Builder で作成した RTR ファイルの例を示します。

リソースタイププロパティの宣言

クラスタ管理者は、RTR ファイルで宣言されているリソースタイププロパティを構成することはできません。このようなリソースタイププロパティは、リソースタイプの恒久的な構成の一部を形成します。

注 - Installed_nodes というリソースタイププロパティは、システム管理者が構成できます。事実、Installed_nodes はシステム管理者が構成できる唯一のリソースタイププロパティであり、RTR ファイルでは宣言できません。

次に、リソースタイプ宣言の構文を示します。

```
property_name = value;
```

注 - RGM はプロパティ名の大文字と小文字を区別します。Sun が提供する RTR ファイルのプロパティに対する命名規則では、名前の最初の文字が大文字で、残りが小文字です (メソッド名は例外)。メソッド名は、プロパティ属性と同様にすべて大文字です。

次に、サンプルのデータサービス (smpl) 用の RTR ファイルにおけるリソースタイプ宣言を示します。

```
# Sun Cluster Data Services Builder template version 1.0
# Registration information and resources for smpl
#
#NOTE: Keywords are case insensitive, i.e., you can use
#any capitalization style you prefer.
#
Resource_type = "smpl";
Vendor_id = SUNW;
RT_description = "Sample Service on Sun Cluster";

RT_version = "1.0";
```

```

API_version = 2;
Failover = TRUE;

Init_nodes = RG_PRIMARYES;

RT_basedir=/opt/SUNWsmpl/bin;

Start          =   smpl_svc_start;
Stop           =   smpl_svc_stop;

Validate       =   smpl_validate;
Update         =   smpl_update;

Monitor_start  =   smpl_monitor_start;
Monitor_stop   =   smpl_monitor_stop;
Monitor_check  =   smpl_monitor_check;

```

ヒント – RTR ファイルの最初のエントリには、`Resource_type` プロパティを宣言する必要があります。宣言しないと、リソースタイプの登録は失敗します。

リソースタイプ宣言の最初のセットは、次のようなリソースタイプについての基本的な情報を提供します。

`Resource_type` および `Vendor_id`

リソースタイプの名前を提供します。リソースタイプ名は `Resource_type` プロパティ (この例では「`smpl`」) 単独で指定できます。 `Vendor_id` を接頭辞として使用し、リソースタイプ (この例では「`SUNW.smpl`」) との区切りにドット (.) を入力することもできます。 `Vendor_id` を使用する場合、リソースタイプを定義する企業の略号にします。リソースタイプ名はクラスタ内で一意である必要があります。

注 – 便宜上、リソースタイプ名

(`Resource_type` と `Vendor_id`) はパッケージ名として使用されます。パッケージ名は9文字に制限されているので、これら2つのプロパティの文字数の合計も9文字以内に制限することをお勧めします。RGM は9文字の制限を適用しません。一方、`Agent Builder` はリソースタイプ名からパッケージ名を系統だてて生成します。つまり、`Agent Builder` は9文字の制限を適用しません。

RT_version	サンプルデータサービスのバージョンです。
API_version	API のバージョンです。たとえば、「API_version =2」は、データサービスが Sun Cluster バージョン 3.0 の管理下で動作していることを示します。
Failover = TRUE	このデータサービスが、複数のノード上で同時にオンラインにできるリソースグループ上では実行できないサービス、すなわちフェイルオーバーデータサービスであることを示します。詳細については、30 ページの「データサービスをクラスタに転送する方法」を参照してください。
Start、Stop、Validate など	RGM によって呼び出されるコールバックメソッドプログラムのパスを提供します。これらのパスは、RT_basedir で指定されたディレクトリからの相対パスです。
リソースタイプ宣言の残りのセットは、次のような構成情報を提供します。	
Init_nodes = RG_PRIMARYES	RGM が、データサービスをマスターできるノード上でのみ Init、Boot、Fini、Validate の各メソッドを呼び出すように指定します。RG_PRIMARYES で指定されたノードは、データサービスがインストールされているすべてのノードのサブセットです。この値に RT_INSTALLED_NODES を設定した場合、RGM は、データサービスがインストールされているすべてのノード上で上記メソッドを呼び出します。
RT_basedir	コールバックメソッドパスのような完全な相対パスとして、/opt/SUNWsample/bin をポイントします。
Start、Stop、Validate など	RGM によって呼び出されるコールバックメソッドプログラムのパスを提供します。これらのパスは、RT_basedir で指定されたディレクトリからの相対パスです。

リソースプロパティの宣言

リソースタイププロパティと同様に、リソースプロパティも RTR ファイルで宣言します。便宜上、リソースプロパティ宣言は RTR ファイルのリソースタイププロパティ宣言の後に行います。リソース宣言の構文では、一連の属性と値のペアを記述して、全体を中括弧で囲みます。

```
{
    Attribute = Value;
    Attribute = Value;
    .
    .
    Attribute = Value;
}
```

Sun Cluster が提供するリソースプロパティ (システム定義プロパティ) の場合、特定の属性は RTR ファイルで変更できます。たとえば、Sun Cluster はコールバックメソッドごとにメソッドタイムアウトプロパティを定義して、そのデフォルト値を提供します。RTR ファイルを使用すると、異なるデフォルト値を指定できます。

Sun Cluster が提供するプロパティ属性を使用することにより、RTR ファイル内に新しいリソースプロパティ (拡張プロパティ) を定義することもできます。263 ページの「リソースプロパティの属性」に、リソースプロパティを変更および定義するための属性を示します。拡張プロパティ宣言は RTR ファイルのシステム定義プロパティ宣言の後に行います。

システム定義リソースプロパティの最初のセットでは、コールバックメソッドのタイムアウト値を指定します。

...

```
# リソースプロパティ宣言は中括弧で囲まれたリストであり、
#リソースタイププロパティ宣言の後で宣言する。
#プロパティ名宣言は、リソースプロパティエントリの左中括弧の
# 直後にある最初の属性でなければならない。
#
# メソッドタイムアウト用の最小値とデフォルト値を設定する。
{
    PROPERTY = Start_timeout;
    MIN=60;
    DEFAULT=300;
}

{
    PROPERTY = Stop_timeout;
    MIN=60;
    DEFAULT=300;
}

{
    PROPERTY = Validate_timeout;
    MIN=60;
```

```

        DEFAULT=300;
    }
    {
        PROPERTY = Update_timeout;
        MIN=60;
        DEFAULT=300;
    }
    {
        PROPERTY = Monitor_Start_timeout;
        MIN=60;
        DEFAULT=300;
    }
    {
        PROPERTY = Monitor_Stop_timeout;
        MIN=60;
        DEFAULT=300;
    }
    {
        PROPERTY = Monitor_Check_timeout;
        MIN=60;
        DEFAULT=300;
    }
}

```

プロパティ名 (PROPERTY = *value*) は、各リソースプロパティ宣言における最初の属性でなければなりません。リソースプロパティは、RTR ファイルのプロパティ属性で定義された範囲内で構成することができます。たとえば、各メソッドタイムアウト用のデフォルト値は 300 秒です。システム管理者はこの値を変更できますが、指定できる最小値は (MIN 属性で指定されているように) 60 秒です。263 ページの「リソースプロパティの属性」にリソースプロパティ属性のリストを示します。

リソースプロパティの次のセットは、データサービスにおいて特定の目的に使用されるプロパティを定義します。

```

{
    PROPERTY = Failover_mode;
    DEFAULT=SOFT;
    TUNABLE = ANYTIME;
}
{
    PROPERTY = Thorough_Probe_Interval;
    MIN=1;
    MAX=3600;
    DEFAULT=60;
    TUNABLE = ANYTIME;
}
# ある期限内に再試行する回数。この回数を超えると、
# 当該ノード上ではアプリケーションが起動できないと判断される。
{
    PROPERTY = Retry_Count;
    MAX=10;
    DEFAULT=2;
    TUNABLE = ANYTIME;
}

```

```

# Retry_Interval に 60 の倍数を設定する。
# この値は秒から分に変換され、切り上げられる。
# たとえば、50 秒は 1 分に変更される。このプロパティを使用して、
# 再試行回数 (Retry_Count) を指定する。
{
    PROPERTY = Retry_Interval;
    MAX=3600;
    DEFAULT=300;
    TUNABLE = ANYTIME;
}

{
    PROPERTY = Network_resources_used;
    TUNABLE = WHEN_DISABLED;
    DEFAULT = "";
}

{
    PROPERTY = Scalable;
    DEFAULT = FALSE;
    TUNABLE = AT_CREATION;
}

{
    PROPERTY = Load_balancing_policy;
    DEFAULT = LB_WEIGHTED;
    TUNABLE = AT_CREATION;
}

{
    PROPERTY = Load_balancing_weights;
    DEFAULT = "";
    TUNABLE = ANYTIME;
}

{
    PROPERTY = Port_list;
    TUNABLE = AT_CREATION;
    DEFAULT = ;
}

```

上記のリソースプロパティ宣言では、システム管理者が値を設定し、制限を設けることができる TUNABLE 属性が追加されています。AT_CREATION は、システム管理者が値を指定できるのはリソースの作成時だけであり、後で変更できないことを示します。

上記のプロパティのほとんどは、特に理由がないかぎり、Agent Builder が生成するデフォルト値を使用しても問題ありません。こうしたプロパティのあとには、次のような情報が続きます (詳細については、[245 ページの「リソースのプロパティ」](#)または `r_properties(5)` のマニュアルページを参照してください)。

Failover mode

Start または Stop メソッドの失敗時、RGM がリソースグループを再配置するか、ノードを停止するかを指定します。

`Thorough_probe_interval`、`Retry_count`、`Retry_interval`
障害モニターで使用します。障害モニターが適切に機能していない場合、システム管理者はいつでも調整できます。

`Network_resources_used`
データサービスで使用される論理ホスト名または共有アドレスリソースのリスト。このプロパティは、Agent Builder によって宣言されるので、システム管理者はデータサービスを構成するとき必要に応じてリソースのリストを指定できます。

`Scalable`
FALSE に設定した場合、このリソースはクラスタネットワーキング (共有アドレス) 機能を使用しません。この設定は、リソースタイプ `Failover` プロパティに TRUE を設定して、フェイルオーバーサービスを指定するのと同じです。このプロパティの詳しい使用方法については、30 ページの「データサービスをクラスタに転送する方法」および 39 ページの「コールバックメソッドの実装」を参照してください。

`Load_balancing_policy`、`Load_balancing_weights`
これらのプロパティは自動的に宣言されますが、フェイルオーバーリソースタイプでは使用されません。

`Port_list`
サーバーが待機するポートのリストです。このプロパティは Agent Builder によって宣言されるので、システム管理者はデータサービスを構成するときポートのリストを指定できます。

拡張プロパティの宣言

次に、RTR ファイルの最後の例として、拡張プロパティを示します。

```
# 拡張プロパティ
#
# クラスタ管理者は、このプロパティに値を設定して、アプリケーション
# が使用する構成ファイルが格納されているディレクトリを指定する
# 必要がある。このアプリケーション (smp1) の場合、PXFS 上に
# ある構成ファイル (通常は named.conf) のパスを指定する。

{
    PROPERTY = Confdir_list;
    EXTENSION;
    STRINGARRAY;
    TUNABLE = AT_CREATION;
    DESCRIPTION = "The Configuration Directory Path(s)";
}

# 次の 2 つのプロパティは、障害モニターの再起動を制御する。
{
    PROPERTY = Monitor_retry_count;
    EXTENSION;
    INT;
    DEFAULT = 4;
}
```

```

    TUNABLE = ANYTIME;
    DESCRIPTION = "Number of PMF restarts allowed for fault monitor.";
}
{
    PROPERTY = Monitor_retry_interval;
    EXTENSION;
    INT;
    DEFAULT = 2;
    TUNABLE = ANYTIME;
    DESCRIPTION = "Time window (minutes) for fault monitor restarts.";
}
# 検証用のタイムアウト値 (秒)。
{
    PROPERTY = Probe_timeout;
    EXTENSION;
    INT;
    DEFAULT = 120;
    TUNABLE = ANYTIME;
    DESCRIPTION = "Time out value for the probe (seconds)";
}
# PMF 用の子プロセス監視レベル (pmfadm の -C オプション)。
# デフォルトの -1 は、pmfadm -C オプションを使用しないこと
# を示す。
# 0 より大きな値は、目的の子プロセス監視レベルを示す。
{
    PROPERTY = Child_mon_level;
    EXTENSION;
    INT;
    DEFAULT = -1;
    TUNABLE = ANYTIME;
    DESCRIPTION = "Child monitoring level for PMF";
}
# ユーザー追加コード-- BEGIN VVVVVVVVVVVV
# ユーザー追加コード-- END ^^^^^^^^^^^^^^

```

次に示すように、Agent Builder は、ほとんどのデータサービスにとって有用な拡張プロパティを作成します。

Confdir_list

アプリケーション構成ディレクトリへのパスを指定します。このプロパティは多くのアプリケーションにとって有用な情報です。データサービスを構成するときに、システム管理者はこのディレクトリの場所を指定できます。

Monitor_retry_count、Monitor_retry_interval、Probe_timeout
サーバーデーモンではなく、障害モニター自体の再起動を制御します。

Child_mon_level

PMF による監視レベルを設定します。詳細については、pmfadm(1M) のマニュアルページを参照してください。

「ユーザー追加コード」というコメント文で囲まれた部分に、追加の拡張プロパティを作成できます。

コールバックメソッドの実装

この節では、コールバックメソッドの実装に関する一般的な情報について説明します。

リソースとリソースグループのプロパティ情報へのアクセス

一般に、コールバックメソッドはリソースのプロパティにアクセスする必要があります。RMAPI は、リソースのシステム定義プロパティと拡張プロパティにアクセスするために、コールバックメソッドで使用できるシェルコマンドと C 関数の両方を提供します。詳細については、`scha_resource_get(1HA)` と `scha_resource_get(3HA)` のマニュアルページを参照してください。

DSDL は、システム定義プロパティにアクセスするための C 関数セット (プロパティごとに 1 つ) と、拡張プロパティにアクセスするための関数を提供します。詳細については、`scds_property_functions(3HA)` と `scds_get_ext_property(3HA)` のマニュアルページを参照してください。

`Status` と `Status_msg` の設定を除き、リソースプロパティを設定する API 関数が存在しないため、プロパティ機構を使用して、データサービスの動的な状態情報を格納することはできません。したがって、動的な状態情報は、広域ファイルに格納するようにします。

注 - クラスタ管理者は、`scrgadm` コマンド、グラフィカル管理コマンド、またはグラフィカル管理インタフェースを使用して、ある種のリソースプロパティを設定することができます。ただし、`scrgadm` はクラスタの再構築時、すなわち RGM がメソッドを呼び出した時点でエラー終了するため、コールバックメソッドから呼び出さないようにします。

メソッドの呼び出し回数への非依存性

一般に、RGM は、同じリソース上で同じメソッドを (同じ引数で) 何回も連続して呼び出すことはありません。ただし、`Start` メソッドが失敗した場合には、リソースが起動していなくても RGM はそのリソース上で `Stop` メソッドを呼び出すことができます。同様に、リソースデーモンが自発的に停止している場合でも、RGM はそのリソース上で `Stop` メソッドを呼び出すことができます。`Monitor_start` メソッドと `Monitor_stop` メソッドにも、同じことが当てはまります。

このような理由のため、`Stop` メソッドと `Monitor_stop` メソッドは呼び出し回数に依存しないように組み込む必要があります。同じリソース上で、同じパラメータを指定して `Stop` メソッドまたは `Monitor_stop` メソッドを何回連続して呼び出しても、1 回だけ呼び出したときと同じ結果になります。

呼び出し回数に依存しないということは、リソースまたはモニターがすでに停止し、動作していなくても、Stop メソッドと Monitor_stop メソッドが 0 (成功) を返す必要があるということも意味します。

注 - Init、Fini、Boot、Update の各メソッドも呼び出し回数に依存しない必要があります。Start メソッドは呼び出し回数に依存してもかまいません。

汎用データサービス

汎用データサービス (GDS) は、単純なアプリケーションを Sun Cluster の Resource Group Manager フレームワークに組み込むことにより、スケーラビリティと高可用性を実現する機構です。この機構では、アプリケーションを可用性の高いものにしたリ、スケーラブルなものにするために通常必要になるエージェントのコーディングは必要ありません。

GDS モデルは、コンパイル済みのリソースタイプ SUNQ.gds により、RGM フレームワークとやりとりします。

詳細については、[第 10 章](#)を参照してください。

アプリケーションの制御

RGM は、ノードがクラスタに結合されるとき、またはクラスタから切り離されるとき、コールバックメソッドを使って実際のリソース (アプリケーション) を制御できます。

リソースの起動と停止

リソースタイプを実装するには、少なくとも、Start メソッドと Stop メソッドが必要です。RGM は、リソースタイプのメソッドプログラムを適切なノード上で適切な回数だけ呼び出して、リソースグループをオフラインまたはオンラインにします。たとえば、クラスタノードのクラッシュ後、そのノードがマスターしているリソースグループを新しいノードに移動します。Start メソッドを実装して、RGM が正常に動作しているホストノード上で各リソースを再起動できるようにする必要があります。

Start メソッドは、ローカルノード上でリソースが起動し、使用可能な状態になるまで終了してはいけません。初期化に時間がかかるリソースタイプでは、その Start メソッドに、十分な長さのタイムアウト値を設定する必要があります。Start_timeout プロパティのデフォルト値と最小値は、リソースタイプ登録ファイルで設定します。

Stop メソッドは、RGM がリソースをオフラインにする状況に合わせて実装する必要があります。たとえば、リソースグループがノード 1 上でオフラインになり、ノード 2 上でもう一度オンラインになると仮定します。リソースグループをオフラインにしている間、RGM は、そのリソースグループ内のリソース上で Stop メソッドを呼び出して、ノード 1 上のすべての活動を停止しようとしています。ノード 1 上ですべてのリソースの Stop メソッドが完了したなら、RGM は、ノード 2 上でそのリソースグループを再度オンラインにします。

Stop メソッドは、ローカルノード上でリソースがすべての活動を完全に停止し完全にシャットダウンするまで終了してはいけません。Stop メソッドは、ローカルノード上でリソース関連のすべてのプロセスを終了することでもっとも安全に実装できます。シャットダウンに時間がかかるリソースタイプでは、その Stop メソッドに十分な長さのタイムアウト値を設定する必要があります。Stop_timeout プロパティはリソースタイプ登録ファイルで設定します。

Stop メソッドが失敗またはタイムアウトすると、リソースグループはエラー状態になり、システム管理者の介入が必要となります。この状態を回避するには、Stop および Monitor_stop メソッドがすべてのエラー状態から回復するようにする必要があります。理想的には、これらのメソッドは 0 (成功) のエラー状態で終了し、ローカルノード上でリソースとそのモニターのすべての活動を正常に停止するべきです。

Start および Stop メソッドを使用するかどうかの決定

この節では、Start メソッドと Stop メソッドを使用するか、または、Prenet_start メソッドと Postnet_stop メソッドを使用するかを決定するときのいくつかの注意事項について説明します。どちらのメソッドを使用するのが適切かを決定するには、クライアントおよびデータサービスのクライアントサーバー型ネットワークプロトコルについて十分に理解している必要があります。

ネットワークアドレスリソースを使用するサービスでは、論理ホスト名のアドレス構成から始まる順番で、起動手順または停止手順を実行する必要があります。コールバックメソッドの Prenet_start と Postnet_stop を使用してリソースタイプを実装すると、同じリソースグループ内のネットワークアドレスが「起動」に構成される前、または「停止」に構成されたあとに、特別な起動アクションまたは停止アクションを行います。

RGM は、データサービスの Prenet_start メソッドを呼び出す前に、ネットワークアドレスを取り付ける (plumb、ただし起動には構成しない) メソッドを呼び出します。RGM は、データサービスの Postnet_stop メソッドを呼び出したあとに、ネットワークアドレスを取り外す (unplumb) メソッドを呼び出します。RGM がリソースグループをオンラインにするときは、次のような順番になります。

1. ネットワークアドレスを取り付けます。

2. データサービスの `Prenet_start` メソッドを呼び出します (存在する場合)。
3. ネットワークアドレスを「起動」に構成します。
4. データサービスの `Start` メソッドを呼び出します (存在する場合)。

RGM がリソースグループをオフラインにするときは、逆の順番になります。

1. データサービスの `Stop` メソッドを呼び出します (存在する場合)。
2. ネットワークアドレスを「停止」に構成します。
3. データサービスの `Postnet_stop` メソッドを呼び出します (存在する場合)。
4. ネットワークアドレスを取り外します。

`Start`、`Stop`、`Prenet_start`、`Postnet_stop` のうち、どのメソッドを使用するかを決定するには、まずサーバー側を考えます。データサービスアプリケーションリソースとネットワークアドレスリソースの両方を持つリソースグループをオンラインにするとき、RGM は、データサービスリソースの `Start` メソッドを呼び出す前に、ネットワークアドレスを「起動」に構成するメソッドを呼び出します。したがって、データサービスを起動するときにネットワークアドレスが「起動」に構成されている必要がある場合は、`Start` メソッドを使用してデータサービスを起動します。

同様に、データサービスアプリケーションリソースとネットワークアドレスリソースの両方を持つリソースグループをオフラインにするとき、RGM は、データサービスリソースの `Stop` メソッドを呼び出したあとに、ネットワークアドレスを「停止」に構成するメソッドを呼び出します。したがって、データサービスを停止するときにネットワークアドレスが「起動」に構成されている必要がある場合は、`Stop` メソッドを使用してデータサービスを停止します。

たとえば、データサービスを起動または停止するときに、データサービスの管理ユーティリティまたはライブラリを呼び出す必要がある場合もあります。また、クライアントサーバー型ネットワークインタフェースを使用して管理を実行するような管理ユーティリティまたはライブラリを持っているデータサービスもあります。つまり、管理ユーティリティがサーバーデーモンを呼び出すので、管理ユーティリティまたはライブラリを使用するためには、ネットワークアドレスが「起動」に構成されている必要があります。このような場合は、`Start` メソッドと `Stop` メソッドを使用します。

データサービスが起動および停止するときにネットワークアドレスが「停止」に構成されている必要がある場合は、`Prenet_start` メソッドと `Postnet_stop` メソッドを使用してデータサービスを起動および停止します。クラスタ再構成 (SCHA_GIVEOVER 引数を指定した `scha_control()` または `scswitch` によるスイッチオーバー) のあとネットワークアドレスとデータサービスのどちらが最初にオンラインになるかによってクライアントソフトウェアの応答が異なるかどうかを考えます。たとえば、クライアントの実装が最小限の再試行を行うだけで、データサービスのポートが利用できないと判断すると、すぐにあきらめる場合もあります。

データサービスを起動するときにネットワークアドレスが「起動」に構成されている必要がない場合、ネットワークインタフェースが「起動」に構成される前に、データサービスを起動します。すると、ネットワークアドレスが「起動」に構成されるとすぐに、データサービスはクライアントの要求に応答できます。したがって、クライアントが再試行を停止する可能性も減ります。このような場合は、`Start` ではなく、`Prenet_start` メソッドを使用してデータサービスを起動します。

Postnet_stop メソッドを使用した場合、ネットワークアドレスが「停止」に構成されている時点では、データサービスリソースは「起動」のままです。Postnet_stop メソッドを呼び出すのは、ネットワークアドレスが「停止」に構成されたあとだけです。結果として、データサービスの TCP または UDP のサービスポート (つまり、その RPC プログラム番号) は、常に、ネットワーク上のクライアントから利用できます。ただし、ネットワークアドレスが応答しない場合を除きます。

注 - クラスタに RPC サービスをインストールする場合、サービスはプログラム番号 100141、10014、および 100248 を使用できません。これらの番号は、Sun Cluster デーモン rgmd_receptionist、fed、および pmfd 用に予約されています。これらのプログラム番号を使用する RPC サービスをインストールした場合は、別のプログラム番号を使用するように変更する必要があります。

Start メソッドと Stop メソッドを使用するか、Prenet_start メソッドと Postnet_stop メソッドを使用するか、または両方を使用するかを決定するには、サーバーとクライアントの要件と動作を考慮に入れる必要があります。

Init、Fini、Boot の各メソッド

RGM は、3 つの任意のメソッド Init、Fini、Boot を使用し、リソース上で初期化と終了コードを実行できます。リソースを管理下に置くとき (リソースが属しているリソースグループを管理していない状態から管理している状態に切り替えるとき、またはすでに管理されているリソースグループでリソースを作成するとき)、RGM は Init メソッドを呼び出して、1 回だけリソースの初期化を実行します。

リソースを管理下から外すとき (リソースが属しているリソースグループを管理していない状態に切り替えるとき、またはすでに管理されているリソースグループからリソースを削除するとき)、RGM は Fini を呼び出して、リソースをクリーンアップします。クリーンアップは呼び出し回数に依存しない必要があります。つまり、すでにクリーンアップが行われている場合、Fini は 0 (成功) で終了する必要があります。

RGM は、新たにクラスタに結合したノード、すなわち起動または再起動したノード上で、Boot メソッドを呼び出します。

Boot メソッドは、通常、Init と同じ初期化を実行します。この初期化は呼び出し回数に依存しない必要があります。つまり、ローカルノード上ですでにリソースが初期化されている場合、Boot と Init は 0 (成功) で終了する必要があります。

リソースの監視

通常、モニターは、リソース上で定期的に障害検証を実行し、検証したリソースが正しく動作しているかどうかを検出するように実装します。障害検証が失敗した場合、モニターはローカルで再起動するか、RMAPI 関数 `scha_control()` または DSDL 関数 `scds_fm_action()` を呼び出して、影響を受けるリソースグループのフェイルオーバーを要求できます。

また、リソースの性能を監視して、性能を調節または報告できます。可能であれば、リソースタイプに固有な障害モニターを作成することを推奨します。このような障害モニターを作成しなくても、リソースタイプは Sun Cluster により基本的なクラスタの監視が行われます。Sun Cluster は、ホストハードウェアの障害、ホストのオペレーティングシステムの全体的な障害、およびパブリックネットワーク上で通信できるホストの障害を検出します。

RGM がリソースモニターを直接呼び出すことはありませんが、RGM は自動的にリソース用のモニターを起動する準備を整えます。リソースをオフラインにするとき、RGM は、リソース自体を停止する前に、`Monitor_stop` メソッドを呼び出して、ローカルノード上でリソースのモニターを停止します。リソースをオンラインにするとき、RGM は、リソース自体を起動した後に、`Monitor_start` メソッドを呼び出します。

RMAPI 関数 `scha_control()` と `scha_control()` を呼び出す DSDL 関数 `scds_fm_action()` を使用すると、リソースモニターは異なるノードへのリソースグループのフェイルオーバーを要求できます。`Monitor_check` が定義されている場合、`scha_control` は健全性検査の一環としてこの関数を呼び出して、リソースが属するリソースグループをマスターするのに要求されたノードが十分信頼できるかどうかを判断します。`Monitor_check` が「このノードは信頼できない」と報告した場合、あるいはメソッドがタイムアウトした場合、RGM はフェイルオーバー要求に適する別のノードを探します。すべてのノードで `Monitor_check` が失敗した場合、フェイルオーバーは取り消されます。

リソースモニターは、モニターから見たリソースの状態を反映するように `Status` と `Status_msg` プロパティを設定します。これらのプロパティを設定するには、RMAPI 関数 `scha_resource_setstatus()` または `scha_resource_setstatus` コマンド、あるいは DSDL 関数 `scds_fm_action()` を使用します。

注 - `Status` と `Status_msg` はリソースモニターに固有の使用方法ですが、これらのプロパティは任意のプログラムで設定できます。

RMAPI による障害モニターの実装例については、100 ページの「障害モニターの定義」を参照してください。DSDL による障害モニターの実装例については、147 ページの「SUNW.xfnts 障害モニター」を参照してください。Sun が提供するデータサービスに組み込まれている障害モニターについては、『Sun Cluster データサービスの計画と管理 (Solaris OS 版)』を参照してください。

メッセージログのリソースへの追加

状態メッセージをほかのクラスタメッセージと同じログファイルに記録する場合は、`scha_cluster_getlogfacility()` 関数を使用して、クラスタメッセージを記録するために使用されている機能番号を取得します。

この機能番号を通常の Solaris `syslog()` 関数で使用して、状態メッセージをクラスタログに書き込みます。`scha_cluster_get()` 汎用インタフェースからでも、クラスタログ機能情報にアクセスできます。

プロセス管理の提供

リソースモニターとリソース制御コールバックを実装するために、プロセス管理機能が RMAPI および DSDL に提供されています。RMAPI は次の機能を定義します。これらのコマンドとプログラムの詳細については、各マニュアルページを参照してください。

プロセス監視機能: `pmfadm` および `rpc.pmfd`

プロセス監視機能 (Process Monitor Facility: PMF) は、プロセスとその子孫を監視し、プロセスが終了したときに再起動する手段を提供します。この機能は、監視するプロセスを起動および制御する `pmfadm` コマンドと、`rpc.pmfd` デーモンからなります。

`halockrun`

ファイルロックを保持しながら子プログラムを実行するためのプログラムです。このコマンドはシェルスクリプトで使用すると便利です。

`hatimerun`

タイムアウト制御下で子プログラムを実行するためのプログラムです。このコマンドはシェルスクリプトで使用すると便利です。

DSDL は、`hatimerun` 機能を実装するために `scds_hatimerun` 関数を提供します。

DSDL は、PMF 機能を実装するための関数セット (`scds_pmf_*`) を提供します。DSDL の PMF 機能の概要と、個々の関数のリストについては、[206 ページの「PMF 関数」](#) を参照してください。

リソースへの管理サポートの提供

リソース上での管理アクションには、リソースプロパティの設定と変更があります。このような管理アクションを行うために、API は `Validate` と `Update` というコールバックメソッドを定義しています。

リソースの作成時や、管理アクションによるリソースまたはリソースグループのプロパティの更新時、RGM は、オプションの `Validate` メソッドを呼び出します。RGM は、リソースとそのリソースグループのプロパティ値を `Validate` メソッドに渡します。RGM は、リソースタイプの `Init_nodes` プロパティが示す複数のクラスタノード上で `Validate` を呼び出します。`Init_nodes` の詳細については、[239 ページ](#) の「リソースタイププロパティ」か、`rt_properties(5)` のマニュアルページを参照してください。RGM は、作成または更新が行われる前に `Validate` を呼び出します。任意のノード上でメソッドから失敗の終了コードが戻ってくると、作成または更新は取り消されます。

RGM が `Validate` メソッドを呼び出すのは、管理アクションがリソースまたはグループのプロパティを変更したときだけです。RGM がプロパティを設定したときや、モニターがリソースプロパティ `Status` と `Status_msg` を設定したときではありません。

RGM は、オプションの `Update` メソッドを呼び出して、プロパティが変更されたことを実行中のリソースに通知します。RGM は、管理アクションがリソースまたはそのリソースグループのプロパティの設定に成功したあとに、`Update` を呼び出します。RGM は、リソースがオンラインであるノード上で、このメソッドを呼び出します。このメソッドは、API アクセス関数を使用して、アクティブなリソースに影響する可能性があるプロパティ値を読み取り、その値に従って、実行中のリソースを調節できます。

フェイルオーバーリソースの実装

フェイルオーバーリソースグループには、ネットワークアドレス (組み込みリソースタイプである論理ホスト名や共有アドレスなど) やフェイルオーバーリソース (フェイルオーバーデータサービス用のデータサービスアプリケーションリソースなど) があります。データサービスがフェイルオーバーするかスイッチオーバーされると、ネットワークアドレスリソースは関連するデータサービスリソースと共にクラスタノード間を移動します。RGM は、フェイルオーバーリソースの実装をサポートするプロパティをいくつか提供します。

ブール型リソースタイププロパティ `Failover` を `TRUE` に設定し、同時に複数のノード上でオンラインになることができるリソースグループだけで構成されるようにリソースを制限します。このプロパティのデフォルト値は `FALSE` です。したがって、フェイルオーバーリソースを実現するためには、RTR ファイルで `TRUE` として宣言する必要があります。

`Scalable` リソースプロパティは、リソースがクラスタ共有アドレス機能を使用するかどうかを決定します。フェイルオーバーリソースの場合、フェイルオーバーリソースは共有アドレスを使用しないので、`Scalable` には `FALSE` を設定します。

`RG_mode` リソースグループプロパティを使用すると、クラスタ管理者はリソースグループがフェイルオーバーまたはスケーラブルのどちらであるかを識別できます。`RG_mode` が `FAILOVER` の場合、RGM はリソースグループの `Maximum primaries` プロパティを 1 に設定して、リソースグループが単一のノードでマスターされるように制限します。RGM は、`Failover` プロパティが `TRUE` であるリソースを、`RG_mode` が `SCALABLE` であるリソースグループで作成することを禁止します。

`Implicit_network_dependencies` リソースグループプロパティは、リソースグループ内におけるネットワークアドレスリソース (論理ホスト名や共有アドレス) への非ネットワークアドレスリソースの暗黙で強力な依存関係を、RGM が強制することを指定します。これは、リソースグループ内のネットワークアドレスが「起動」に構成されるまで、リソースグループ内の非ネットワークアドレス (データサービス) リソースが、自分の `Start` メソッドを呼び出さないことを意味します。この `Implicit_network_dependencies` プロパティのデフォルト値は `TRUE` です。

スケーラブルリソースの実装

スケーラブルリソースは、同時に複数のノード上でオンラインになることができます。スケーラブルリソースには、Sun Cluster HA for Sun Java System Web Server (以前の Sun Cluster HA for Sun ONE Web Server) や Sun Cluster HA for Apache などのデータサービスがあります。

RGM は、スケーラブルリソースの実装をサポートするプロパティをいくつか提供します。

ブール型リソースタイププロパティの `Failover` を `FALSE` に設定し、一度に複数のノードでオンラインにできるリソースグループ内でリソースが構成されるようにします。

`Scalable` リソースプロパティは、リソースがクラスタ共有アドレス機能を使用するかどうかを決定します。スケーラブルサービスは共有アドレスリソースを使用するので (スケーラブルサービスの複数のインスタンスが単一のサービスであるかのようにクライアントに見せるため)、`Scalable` には `TRUE` を設定します。

`RG_mode` プロパティを使用すると、クラスタ管理者はリソースグループがフェイルオーバーまたはスケーラブルのどちらであるかを識別できます。`RG_mode` が `SCALABLE` の場合、RGM は `Maximum primaries` が 1 より大きな値を持つこと、つ

まり同時に複数のノードがグループをマスターすることを許可します。RGM は、Failover プロパティが FALSE であるリソースが、RG_mode が SCALABLE であるリソースグループ内でインスタンス化されることを許可します。

クラスタ管理者は、スケーラブルサービスリソースが属するためのスケーラブルリソースグループを作成します。また、スケーラブルリソースが依存する共有アドレスリソースが属するためのフェイルオーバーリソースグループも別に作成します。

クラスタ管理者は、RG_dependencies リソースグループプロパティを使用して、あるノード上でリソースグループをオンラインまたはオフラインにする順番を指定します。スケーラブルリソースとそれらが依存する共有アドレスリソースは異なるリソースグループに属するので、この順番はスケーラブルサービスにとって重要です。スケーラブルデータサービスが起動する前に、そのネットワークアドレス (共有アドレス) リソースが構成されていることが必要です。したがって、クラスタ管理者は、スケーラブルサービスが属するリソースグループの RG_dependencies プロパティを設定して、共有アドレスリソースが属するリソースグループを組み込む必要があります。

リソースの RTR ファイルでスケーラブルプロパティを宣言した場合、RGM はそのリソースに対して、次のようなスケーラブルプロパティのセットを自動的に作成します。

Network_resources_used	このリソースによって使用される共有アドレスリソースです。このプロパティのデフォルト値は空の文字列です。したがって、クラスタ管理者は、リソースを作成するときに、スケーラブルサービスが使用する実際の共有アドレスのリストを指定する必要があります。scsetup コマンドと SunPlex Manager は、スケーラブルサービスに必要なリソースとグループを自動的に設定する機能を提供します。
Load_balancing_policy	リソースの負荷均衡ポリシーを指定します。このポリシーは RTR ファイルに明示的に設定しても、デフォルトの LB_WEIGHTED を使用してもかまいません。どちらの場合でも、クラスタ管理者はリソースを作成するときに値を変更できます (RTR ファイルで Load_balancing_policy の Tunable を NONE または FALSE に設定していない場合)。有効な値は次のとおりです。 LB_WEIGHTED Load_balancing_weights プロパティに設定されている重みにより、さまざまなノードに負荷が分散されます。 LB_STICKY スケーラブルサービスの指定のクライアント (クライアントの IP アドレスで識別される) は、常に同じクラスタノードに送信されます。

LB_STICKY_WILD

指定のクライアント (クライアントの IP アドレスで識別される) はワイルドカードスティッキーサービスの IP アドレスに接続され、送信時に使用されるポート番号とは無関係に、常に同じクラスタノードに送信されます。

Load_balancing_policy、LB_STICKY、LB_STICKY_WILD を持つスケーラブルなサービスの場合、サービスがオンラインの状態

Load_balancing_weights を変更すると、既存のクライアントとの関連がリセットされることがあります。リセットされると、(同じクラスタ内にある) 今までサービスを行っていたノードとは別のノードが、後続のクライアント要求を処理します。

同様に、サービスの新しいインスタンスをクラスタ上で開始すると、既存のクライアントとの関連がリセットされることがあります。

Load_balancing_weights

個々のノードへ負荷を送信することを指定します。形式は *weight@node, weight@node* です。weight は、node に分散される負荷の相対的な割り当てを示す整数です。ノードに分散される負荷の割合は、このノードのウェイトをアクティブなインスタンスのすべてのウェイトの合計で割った値になります。たとえば、1@1, 3@2 は、ノード 1 に負荷の 1/4 が割り当てられ、ノード 2 に負荷の 3/4 が割り当てられることを意味します。

Port_list

サーバーが待機するポートです。このプロパティのデフォルト値は空の文字列です。ポートのリストは RTR ファイルに指定できます。このファイルで指定しない場合、クラスタ管理者は、リソースを作成するときに、実際のポートのリストを提供する必要があります。

データサービスは、管理者がスケーラブルまたはフェイルオーバーのどちらにでも構成できるように作成できます。このためには、データサービスの RTR ファイルにおいて、Failover リソースタイププロパティと Scalable リソースプロパティの両方を FALSE に宣言します。Scalable プロパティは作成時に調整できるように指定します。

Failover プロパティが FALSE の場合、リソースはスケーラブルリソースグループに構成できます。管理者はリソースを作成するときに Scalable を TRUE に変更する、すなわちスケーラブルサービスを作成することによって、共有アドレスを有効にできます。

一方、Failover が FALSE の場合でも、管理者はリソースをフェイルオーバーリソースグループに構成して、フェイルオーバーサービスを実装できます。この場合、Scalable の値 (FALSE) は変更しません。このような偶然性に対処するために、Scalable プロパティの validate メソッドで妥当性を検査する必要があります。Scalable が FALSE の場合、リソースがフェイルオーバーリソースグループに構成されていることを確認します。

スケーラブルリソースの詳細については、『Sun Cluster の概念 (Solaris OS 版)』を参照してください。

スケーラブルサービスの妥当性検査

Scalable プロパティが TRUE であるリソースが作成または更新されるたびに、RGM は、さまざまなリソースプロパティの妥当性を検査します。プロパティが正しく構成されていない場合、RGM は作成または更新を拒否します。RGM は次の検査を行います。

- Network_resources_used プロパティは、空の文字列であってはならず、既存の共有アドレスリソースの名前を含む必要があります。スケーラブルリソースを含むリソースグループの NodeList にあるすべてのノードは、指定した共有アドレスリソースの 1 つである NetIfList プロパティまたは AuxNodeList プロパティに存在する必要があります。
- スケーラブルリソースを含むリソースグループの RG_dependencies プロパティは、スケーラブルリソースの Network_resources_used プロパティに存在する、すべての共有アドレスリソースのリソースグループを含む必要があります。
- Port_list プロパティは、空の文字列であってはならず、ポートとプロトコル (tcp または udp) のペアのリストを含む必要があります。たとえば、

```
Port_list=80/tcp,40/udp  
となります。
```

データサービスの作成と検証

この節では、データサービスを作成および検証する方法について説明します。

キープアライブの使用法

サーバー側で TCP キープアライブを有効にしておく、サーバーはダウン時の (または、ネットワークで分割された) クライアントのリソースを浪費しません。長時間稼働するようなサーバーでこのようリソースがクリーンアップされない場合、浪費されたリソースが無制限に大きくなり、最終的にはクライアントに障害が発生して再起動します。

クライアントサーバー通信が TCP ストリームを使用する場合、クライアントとサーバーは両方とも TCP キープアライブ機構を有効にしなければなりません。これは、非高可用性の単一サーバーの場合でも適用されます。

ほかにも、キープアライブ機構を持っている接続指向のプロトコルは存在します。

クライアント側で TCP キープアライブを有効にしておく、ある物理ホストから別の物理ホストに論理ホストがフェイルオーバーまたはスイッチオーバーしたとき、接続の切断がクライアントに通知されます。このようなネットワークアドレスリソースの転送 (フェイルオーバーやスイッチオーバー) が発生すると、TCP 接続が切断されます。しかし、クライアント側で TCP キープアライブを有効にしておかなければ、接続が休止したとき、必ずしも接続の切断はクライアントに通知されません。

たとえば、クライアントが、実行に時間がかかる要求に対するサーバーからの応答を待っており、また、クライアントの要求メッセージがすでにサーバーに到着しており、TCP 層で認識されているものと想定します。この状況では、クライアントの TCP モジュールは要求を再転送し続ける必要はないので、クライアントアプリケーションはブロックされて、要求に対する応答を待ちます。

TCP キープアライブ機構は必ずしもあらゆる限界状況に対応できるわけではなく、クライアントアプリケーションは、可能であれば、TCP キープアライブ機構に加えて、独自の定期的なキープアライブをアプリケーションレベルで実行する必要があります。アプリケーションレベルのキープアライブ機構を使用するには、通常、クライアントサーバー型プロトコルが NULL 操作、または、少なくとも効率的な読み取り専用操作 (状態操作など) をサポートする必要があります。

HA データサービスの検証

この節では、高可用性環境におけるデータサービスの実装を検証する方法について説明します。この検証は一例であり、完全ではないことに注意してください。実際に稼働させるマシンに影響を与えないように、検証時は、検証用の Sun Cluster 構成にアクセスする必要があります。

リソースグループが物理ホスト間で移動するような場合を想定して、HA データサービスが適切に動作するかどうかを検証します。たとえば、システムがクラッシュした場合や、scswitch コマンドを使用した場合です。また、このような場合にクライアントマシンがサービスを受け続けられるかどうかを検証します。

メソッドの呼び出し回数への非依存性を検証します。たとえば、各メソッドを一時的に、元のメソッドを2回以上呼び出す短いシェルスクリプトに変更します。

リソース間の依存関係の調節

あるクライアントサーバーのデータサービスが、クライアントからの要求を満たすために、別のクライアントサーバーのデータサービスに要求を行うことがあります。このように、データサービス A が自分のサービスを提供するために、データサービス B にそのサービスを提供してもらう場合、データサービス A はデータサービス B に依存していると言えます。この要件を満たすために、Sun Cluster では、リソースグループ内でリソースの依存関係を構築できます。依存関係は、Sun Cluster がデータサービスを起動および停止する順番に影響します。詳細については、scrgadm(1M) のマニュアルページを参照してください。

あるリソースタイプのリソースが別のリソースタイプのリソースに依存する場合、データサービス開発者は、リソースとリソースグループを適切に構成するようにユーザーに指示するか、これらを正しく構成するスクリプトまたはツールを提供する必要があります。依存するリソースを依存されるリソースと同じノード上で実行する必要がある場合、両方のリソースを同じリソースグループ内で構成する必要があります。

明示的なリソースの依存関係を使用するか、このような依存関係を省略して、HA データサービス独自のコードで別のデータサービスの可用性をポーリングするかを決定します。依存するリソースと依存されるリソースが異なるノード上で動作できる場合は、これらのリソースを異なるリソースグループ内で構成します。この場合、グループ間にはリソースの依存関係を構築できないため、ポーリングが必要です。

データサービスによっては、データを自分自身で直接格納せず、別のバックエンドデータサービスに依頼して、すべてのデータを格納してもらうものもあります。このようなデータサービスは、すべての読み取り要求と更新要求をバックエンドデータサービスへの呼び出しに変換します。たとえば、すべてのデータを SQL データベース (Oracle など) に格納するようなクライアントサーバー型のアポイントメントカレンダーサービスの場合、このサービスは独自のクライアントサーバー型ネットワークプロトコルを持っています。たとえば、RPC 仕様言語 (ONC RPC など) を使用するプロトコルを定義している場合があります。

Sun Cluster 環境では、HA-ORACLE を使用してバックエンド Oracle データベースを高可用性にできます。つまり、アポイントメントカレンダーデーモンを起動および停止する簡単なメソッドを作成できます。エンドユーザーは Sun Cluster でアポイントメントカレンダーのリソースタイプを登録できます。

アポイントメントカレンダーアプリケーションが Oracle データベースと同じノード上で動作する必要がある場合、エンドユーザーは、HA-ORACLE リソースと同じリソースグループ内でアポイントメントカレンダーリソースを構築して、アポイントメントカレンダーリソースを HA-ORACLE リソースに依存するようにします。この依存関係を指定するには、scrgadm の Resource_dependencies プロパティを使用します。

アポイントメントカレンダーリソースが HA-ORACLE リソースとは別のノード上で動作できる場合、エンドユーザーはこれらのリソースを2つの異なるリソースグループ内で構成します。カレンダーリソースグループのリソースグループ依存関係を、Oracle リソースグループ上で構築することもできます。しかし、リソースグループ依存関係が有効になるのは、両方のリソースグループが同時に同じノード上で起動または停止されたときだけです。したがって、カレンダーデータサービスデーモンは、起動後、Oracle データベースが利用可能になるまで、ポーリングして待機します。この場合、通常、カレンダーリソースタイプの Start メソッドは単に成功を戻すだけです。これは、Start メソッドが無限にブロックされると、そのリソースグループがビジー状態になり、それ以降、リソースグループで状態の変化 (編集、フェイルオーバー、スイッチオーバーなど) が行われなくなるためです。しかし、カレンダーリソースの Start メソッドがタイムアウトまたは非ゼロで終了すると、Oracle データベースが利用できない間、リソースグループが複数のノード間でやりとりを無限に繰り返す可能性があります。

第3章

リソースタイプの更新

この章では、リソースタイプの開発者がリソースタイプの更新や移行を行うために必要な情報を提供します。

- 55 ページの「概要」
- 56 ページの「リソースタイプ登録ファイル」
- 58 ページの「Type_version リソースプロパティ」
- 59 ページの「リソースを別のバージョンへ移行」
- 60 ページの「リソースタイプのアップグレードとダウングレード」
- 62 ページの「デフォルトのプロパティ値」
- 63 ページの「リソースタイプ開発者の文書」
- 63 ページの「リソースタイプ名とリソースタイプモニターの実装」
- 64 ページの「アプリケーションのアップグレード」
- 64 ページの「リソースタイプのアップグレード例」
- 68 ページの「リソースタイプパッケージのインストール要件」

概要

システム管理者は、既存のリソースタイプの新しいバージョンをインストールおよび登録できなければなりません。これは、リソースを削除したり作成し直したりすることなく、複数のバージョンのリソースタイプを登録したり、既存のリソースを新しいバージョンのリソースタイプに移行したりする必要があるからです。リソース開発者は、リソースタイプのアップグレードや移行の要件を把握しておく必要があります。

アップグレードを念頭に置いたリソースタイプの開発を「アップグレード対応」と呼びます。

新しいバージョンのリソースタイプは、次の点で前のバージョンとは異なっている可能性があります。

- リソースタイププロパティの属性

- 標準プロパティ、拡張プロパティを含む宣言済みリソースプロパティ
- リソースプロパティの属性 (default、min、max、arraymin、arraymax) または tunable 属性
- 宣言済みメソッド
- メソッドやモニターの実装

リソースタイプ開発者は、既存のリソースを新しいバージョンへ移行するタイミングを次の Tunable 属性のオプションによって特定します。制約の小さいものから順に、次のようなオプションがあります。

- いつでもよい (ANYTIME)
- リソースが監視されていないとき (WHEN_UNMONITORED)
- リソースがオフラインのとき (WHEN_OFFLINE)
- リソースが無効のとき (WHEN_DISABLED)
- リソースグループが管理されていないとき (WHEN_UNMANAGED)
- 作成時 (AT_CREATION)

各オプションについては、58 ページの「Type_version リソースプロパティ」を参照してください。

注 - この章では、アップグレード手順の記述で常に scrgadm コマンドを使用します。これは、管理者が scrgadm コマンドしか使用できないということではありません。GUI、scsetup コマンドなども使用できます。

リソースタイプ登録ファイル

リソースタイプ名

リソースタイプ名は、RTR ファイルに指定されたプロパティ、Vendor_id、Resource_type、RT_version の3つで構成されます。scrgadm コマンドは、ピリオドとコロンを区切り文字として使用し、リソースタイプ名を作成します。

```
vendor_id.resource_type:rt_version
```

Vendor_id 接頭辞では、同一名でベンダーの異なる2つの登録ファイルを識別できません。RT_version では、同じベースリソースタイプの複数のバージョン(アップグレード)を識別できます。重複を防ぐため、Vendor_id には、リソースタイプの作成元の会社のストックシンボルを使用することをお勧めします。

RT_version の文字列に、空白文字や、タブ、スラッシュ (/)、バックスラッシュ (\)、アスタリスク (*)、疑問符 (?)、コンマ (,)、セミコロン (;)、左角かっこ ([)、右角かっこ (]) が含まれていると、リソースタイプの登録は失敗します。

RT_Version プロパティは、Sun Cluster 3.0 まではオプションでしたが、Sun Cluster 3.1 以降では必須です。

完全名は、次のコマンドで取得できます。

```
scha_resource_get -O Type -R resource_name -G resource_group_name
```

Sun Cluster 3.1 以前に登録されたリソースタイプ名では、引き続き次の構文を使用します。

```
vendor_id.resource_type
```

ディレクティブ

アップグレード対応リソースタイプの RTR ファイルには、次の形式の #`$upgrade` ディレクティブが必要です。ほかにディレクティブがある場合は、このディレクティブの後ろに続きます。

```
#$upgrade_from version tunability
```

`upgrade_from` ディレクティブは、文字列 #`$upgrade_from`、RT_Version、リソースの Tunable 属性の制約で構成されます。アップグレードを行う前のリソースタイプにバージョンがない場合、RT_Version は、以下の例の最後の行のように空文字列として指定されます。

```
#$upgrade_from "1.1" when_offline
#$upgrade_from "1.2" when_offline
#$upgrade_from "1.3" when_offline
#$upgrade_from "2.0" when_unmonitored
#$upgrade_from "2.1" anytime
#$upgrade_from "" when_unmanaged
```

システム管理者がリソース `Type_version` を変更しようとする、RGM によってこれらの制約が課されます。現在のリソースタイプのバージョンがリストに表示されない場合、Tunable 属性は `WHEN_UNMANAGED` になります。

これらのディレクティブは、RTR ファイル内のリソースタイププロパティ宣言とリソース宣言セクションの間に指定されます。rt_reg(4) を参照してください。

RTR ファイル内の RT_Version の変更

RTR ファイルの内容が変更されたときは、必ず RTR ファイル内の RT_Version 文字列を変更します。このプロパティの値は、どちらのリソースタイプが新しく、どちらが古いかをはっきりと示す必要があります。RTR ファイルに変更が加えられていなければ、RT_Version 文字列を変更する必要はありません。

以前のバージョンの Sun Cluster のリソースタイプ名

Sun Cluster 3.0 のリソースタイプ名には、バージョン接尾辞がありません。

```
vendor_id.resource_name
```

元々 Sun Cluster 3.0 で登録されたリソースタイプは、クラスタリングソフトウェアを Sun Cluster 3.1 以上にアップグレードしたあとも、この形式の名前を持ちます。RTR ファイルを Sun Cluster 3.1 またはそれ以上のソフトウェアで登録した場合でも、RTR ファイル内に #`$upgrade` ディレクティブの指定がないリソースタイプは、バージョン接尾辞のない Sun Cluster 3.0 の形式の名前を付与されます。

Sun Cluster 3.0 では、#`$upgrade` ディレクティブや #`$upgrade_from` ディレクティブを使った RTR ファイルの登録は可能ですが、既存のリソースの新しいリソースタイプへの移行はサポートされません。

Type_version リソースプロパティ

標準リソースプロパティ `Type_version` は、リソースタイプの `RT_Version` プロパティを格納します。このプロパティは、RTR ファイル内には指定されません。システム管理者は、次のコマンドを使って、`Type_version` プロパティを編集します。

```
scrgadm -c -j resource -y Type_version=new_version
```

このプロパティの `Tunable` 属性は、次の項目によって決まります。

- 現在のリソースタイプのバージョン
- RTR ファイル内の #`$upgrade_from` ディレクティブ

#`$upgrade_from` ディレクティブの値は次のとおりです。

`ANYTIME`

リソースをいつアップグレードしてもよい場合。リソースを完全にオンラインにできます。

`WHEN_UNMONITORED`

新しいリソースタイプのバージョンの `Update`、`Stop`、`Monitor_check`、および `Postnet_stop` メソッドが古いリソースタイプのバージョンの起動メソッド (`Pre-net_stop` および `Start`) と互換することがわかっている場合と、新しいリソースタイプのバージョンの `Finis` メッセージが古いバージョンの `Init` メソッドと互換することがわかっている場合。アップグレード前にリソース監視プログラムを停止するだけですみます。

`WHEN_OFFLINE`

新しいリソースタイプのバージョンの `Update`、`Stop`、`Monitor_check`、または `Postnet_stop` メソッドと古いリソースタイプのバージョンの起動メソッド

(Prenet_stop および Start) に互換性がないが、これらのメソッドと古いリソースタイプのバージョンの Init メソッドには互換性があることがわかっている場合、タイプのアップグレード時にリソースはオフラインにする必要があります。

WHEN_DISABLED

WHEN_OFFLINE と同じです。ただし、より厳しい条件で、リソースが無効化されている場合になります。

WHEN_UNMANAGED

新しいリソースタイプのバージョンの Fini メソッドが古いバージョンの Init メソッドと互換しないことがわかっている場合、リソースのアップグレード前に既存のリソースグループを管理されていない状態にする必要があります。

AT_CREATION

新しいリソースタイプのバージョンにアップグレードできないリソースの場合、新しいバージョンの新しいリソースしか作成できません。

Tunable 属性が AT_CREATION の場合、リソースタイプ開発者は既存のリソースを新しいタイプに移行することを禁止できます。この場合、システム管理者は、リソースを削除して作成し直す必要があります。これは、リソースのバージョンが作成時にだけ設定されるという宣言と同じことです。

リソースを別のバージョンへ移行

システム管理者がリソースの Type_version プロパティを編集すると、既存のリソースは新しいリソースタイプバージョンを取得します。これは、その他のリソースプロパティの編集時と同じ規則に従います。ただし、一部の情報が現在のバージョンではなく新しいリソースタイプから取得されるという点を除きます。

- min、max、arraymin、arraymax、デフォルト、Tunable 属性など、すべてのプロパティのリソースプロパティ属性は、新しいリソースタイプのバージョンから取得されます。
- Type_version プロパティに適用される Tunable 属性は、RTR ファイル内の #Upgrade_from ディレクティブと、既存のリソースのリソースタイプの RT_version プロパティから取得されます。Tunable 属性は、property_attributes(5) のマニュアルページの説明とは異なります。
- 新しいリソースタイプのバージョンの Validate メソッドが適用されます。このため、プロパティ属性は、新しいリソースタイプで有効です。既存のリソースプロパティ属性が新しいリソースタイプのバージョンの妥当性検査の条件を満たしていない場合、システム管理者は、scrgadm コマンドに、こうしたプロパティの有効な値を指定する必要があります。この手続きは、新しいリソースタイプのバージョンが、以前のバージョンでは宣言されていなかったプロパティを使用し、デフォルト値がない場合に行われます。既存のリソースが、新しいリソースタイプのバージョンでは無効なプロパティ値を持っている場合にも、この手続きが行われず。

- 古いバージョンのリソースタイプで宣言されたリソースプロパティが新しいバージョンでは宣言されないことがあります。こうしたリソースを新しいバージョンに移行すると、リソースからプロパティが削除されます。

注 - Validate メソッドは、リソースの新しい `Type_version` (Validate コマンド行に渡される) だけでなく現在の `Type_version` も照会できます (`scha_resource_get` を使用)。このため、サポートされないバージョンのアップグレードを Validate によって禁止できます。

リソースタイプのアップグレードとダウングレード

リソースタイプのアップグレードおよび移行の詳細については、『*Sun Cluster データサービスの計画と管理 (Solaris OS 版)*』の「リソースタイプのアップグレード」を参照してください。

▼ リソースタイプをアップグレードする方法

1. 新しいリソースタイプのアップグレード手順の説明を読み、リソースタイプの変更とリソースの **Tunable** 属性の制約を確認します。
2. すべてのクラスタノードに、リソースタイプアップグレードパッケージをインストールします。

新しいリソースタイプパッケージのインストールは、段階的に行うことをお勧めします。ノードが非クラスタモードで起動している間に `pkgadd` を実行します。

クラスタモードのノードに新しいリソースタイプパッケージをインストールする場合は、条件に合わせて異なった方法を使用します。

- リソースタイプパッケージのインストールによってメソッドコードが変更されず、モニターだけが更新される場合は、インストール中にそのタイプのすべてのリソース上で監視を停止する必要があります。
 - リソースタイプパッケージのインストールによってメソッドとモニターの両方のコードが変更されない場合は、ディスク上に新しい RTR ファイルが追加されるだけなので、インストール中にリソース上での監視を停止する必要はありません。
3. アップグレードの RTR ファイルを参照し、`scrgadm` または同等のコマンドを使って新しいリソースタイプのバージョンを登録します。
RGM により、次の形式の新しいリソースタイプが作成されます。

```
vendor_id.resource_type:version
```

- リソースタイプのアップグレードがノードのサブセットにだけインストールされる場合は、新しいリソースタイプの **Installed_nodes** プロパティに実際のインストール先ノードを設定する必要があります。

リソースが新しいタイプ (新しく作成された、または更新された) を取得するとき、RGM は、リソースグループ `nodelist` がリソースタイプの `Installed_nodes` リストのサブセットであることを要求します。

```
scrgadm -c -t resource_type -h installed_node_list
```

- 以前にタイプがアップグレードされ、これからそのタイプに移行する各リソースに対して、**scswitch** を呼び出し、そのリソース自体またはそのリソースグループを、アップグレード手順に示されている適切な状態に変更します。
- 以前にタイプがアップグレードされ、これからそのタイプに移行する各リソースを編集し、**Type_version** プロパティに新しいバージョンを設定します。

```
scrgadm -c -j resource -y Type_version=new_version
```

必要に応じて、同じコマンドを使って、同じリソースのその他のプロパティに適切な値を設定します。

- 手順 5 で呼び出したコマンドを逆方向に実行すると、リソースまたはリソースグループの以前の状態を回復できます。

▼ 古いバージョンのリソースタイプにダウングレードする方法

リソースをダウングレードして古いバージョンのリソースタイプにすることが可能です。古いリソースタイプバージョンにダウングレードする場合は、新しいリソースタイプバージョンにアップグレードする場合よりも条件が厳しくなります。まず、リソースグループの管理を解除する必要があります。アップグレードが可能なリソースタイプバージョンにしかダウングレードできないということにも注意してください。アップグレード可能なバージョンは `scrgadm!`) `-p` コマンドを使用して確認できます。アップグレード可能なバージョンの場合、接尾辞 `:version` が表示されます。

- ダウングレードしたいリソースを含んでいるリソースグループをオフラインに切り替えます。

```
scswitch -F -g resource_group
```

- ダウングレードするリソースと、そのリソースグループ内のすべてのリソースを無効にします。

```
scswitch -n -j resource_to_downgrade
```

```
scswitch -n -j resource1
```

```
scswitch -n -j resource2
```

```
scswitch -n -j resource3
```

```
...
```

注 - リソースの無効化は、依存性の高いもの (アプリケーションリソース) から開始し、もっとも依存性の低いもの (ネットワークアドレスリソース) で終了するように行ってください。

3. リソースグループを非管理状態に切り替えます。

```
scswitch -u -g resource_group
```

4. ダウングレード後のリソースタイプバージョンとする古いリソースバージョンがクラスタ内にまだ登録されているかどうか確認します。

- 登録されている場合は、次の手順に進みます。
- 登録されていない場合は、希望する旧バージョンを登録し直します。

```
scrgadm -a -t resource_type_name
```

5. 希望する旧バージョンを **Type_version** に指定し、リソースをダウングレードします。

```
scrgadm -c -j resource_to_downgrade -y Type_version=old_version
```

必要に応じて、同じコマンドを使って、同じリソースのその他のプロパティに適切な値を設定します。

6. ダウングレードしたリソースを含んでいるリソースグループを管理状態にし、すべてのリソースを有効にしたあと、このグループをオンラインに切り替えます。

```
scswitch -Z -g resource_group
```

デフォルトのプロパティ値

RGM は、システム管理者によって明示的に設定されなかったためリソースの CCR (クラスタ構成リポジトリ) に格納されていないデフォルトのプロパティなど、すべてのリソースを格納します。RGM は、CCR からリソースが読み込まれるとき、欠落したリソースプロパティのデフォルト値をそのリソースタイプから取得します。リソースタイプにデフォルト値が定義されていない場合は、システム定義のデフォルト値を使用します。アップグレードされたリソースタイプは、この方法により、新しいプロパティや既存のプロパティの新しいデフォルト値を定義することができます。

リソースプロパティが変更された場合、RGM は編集コマンドで指定されたプロパティを CCR に格納します。

リソースタイプのアップグレードされたバージョンが、デフォルトのプロパティの新しいデフォルト値を宣言する場合、新しいデフォルト値は既存のリソースから継承されます。この手続きは、プロパティが作成時または無効化されている場合だ

け「Tunable」と宣言されている場合でも変わりません。新しいデフォルトのアプリケーションで、Stop、Postnet_stop、Finiなどのメソッドが失敗する場合、リソースタイプの実装者は、アップグレード時のリソースの状態を適切に制限する必要があります。具体的には、Type_version プロパティの Tunable 属性を制限します。

新しいリソースタイプのバージョンの Validate メソッドでは、既存のプロパティ属性が適切であるかどうかを検査できます。既存のプロパティ属性が適切でない場合、システム管理者は、Type_version プロパティを編集するコマンドを使って、新しいリソースタイプのバージョンにリソースをアップグレードできるように、プロパティの属性値を変更できます。

注 – Sun Cluster 3.0 で作成されたリソースは、新しいバージョンに移行しても、リソースタイプからデフォルトのプロパティ属性を継承しません。これは、これらのリソースのデフォルトのプロパティ値が CCR に格納されているからです。

リソースタイプ開発者の文書

リソースタイプ開発者は、新しいリソースに関する文書を提供する必要があります。必須記載事項は次のとおりです。

- プロパティの追加、変更、削除の説明
- プロパティを新しい要件に準拠させた方法の説明
- リソースの Tunable 属性の制約
- 新しいデフォルトプロパティ属性
- Type_version プロパティの編集に使用するコマンドを使って、既存のリソースプロパティを適切な値に変更し、リソースを新しいリソースタイプのバージョンにアップグレードできることの通知 (システム管理者向けの情報)

リソースタイプ名とリソースタイプモニターの実装

Sun Cluster 3.0 でもアップグレード対応のリソースタイプを登録することはできますが、これらの名前はバージョン接尾辞なしで CCR に記録されています。Sun Cluster 3.0 と Sun Cluster 3.1 (および以降のリリース) の両方で正常に実行するには、このリソースタイプのモニターが両方の命名規則を処理できなければなりません。

```
vendor_id.resource_name:version
vendor_id.resource_name
```

モニターコードは、次のコマンドと同等のコマンドを実行することにより、適切な名前を判断できます。

```
scha_resourcetype_get -O RT_VERSION -T VEND.myrt
scha_resourcetype_get -O RT_VERSION -T VEND.myrt:vers
```

次に、`vers` と出力値を比較します。同じリソースタイプの同じバージョンを別の名前で2回以上登録することはできません。したがって、`vers` の特定の値に対して成功するのは、上記のコマンドのいずれか一方のみとなります。

アプリケーションのアップグレード

アプリケーションコードのアップグレードは、いくつかの共通点はあるものの、エージェントコードのアップグレードとは大きく異なっています。アプリケーションのアップグレードでは、リソースタイプのアップグレードが必要な場合とそうでない場合があります。

リソースタイプのアップグレード例

以下では、リソースタイプのインストールとアップグレードの例をいくつか紹介します。Tunable 属性とパッケージ情報は、リソースタイプ実装の変更に基づいて選択されています。Tunable 属性は、リソースを新しいリソースタイプに移行するときに適用されます。

すべての例は、次の条件に従うものとします。

- リソースタイプは Solaris パッケージで配布されています。pkgadd(1M) および pkgrm(1M) のマニュアルページを参照してください。
- リソースタイプの以前のバージョンは1つしかないので、新しい RTR ファイル内には #supgrade_from ディレクティブが1つしかありません。
- RGM がディスクから削除されたメソッドを呼び出すことができる場合、インストールによってメソッドが削除または上書きされることはありません。
- 特に注記がない場合、新しいメソッドと古いメソッドには互換性があります。
- リソースおよびリソースグループは、インストールまたは移行前に正しい scswitch(1M) コマンド (または同等のコマンド) により適切な状態に設定されます。次に示すのは、リソースグループを管理対象外に設定する例です。


```
scswitch -M -n -j resource
scswitch -n -j resource
scswitch -F -g resource_group
scswitch -u -g resource_group
```

- リソースタイプの登録は次のコマンドで行います。

```
scrgadm -a -t resource_type -f path_to_RTR_file
```

- リソースの移行は次のコマンドで行います。

```
scrgadm -c -j resource -y Type_version=version \
-y property=value \
-x property=value ...
```

- リソースおよびリソースグループは、移行後に適切な `scswitch (1M)` コマンド (または同等のコマンド) により以前の状態に戻されます。

```
scswitch -M -e -j resource
scswitch -e -j resource
scswitch -o -g resource_group
scswitch -Z -g resource_group
```

場合によっては、リソースタイプ開発者は、例で使用されているものより制限の厳しい `Tunable` 属性の値を指定する必要があります。`Tunable` 属性の値は、リソースタイプ実装の変更内容によって決定します。リソースタイプ開発者は、例で使用されている `Solaris` パッケージ以外のパッケージスキーマを選択することもできます。

表 3-1 リソースタイプのアップグレード例

変更の種類	Tunable 属性	パッケージ処理	手続き
プロパティの変更は、RTR ファイルのみに加えられました。	ANYTIME	新しい RTR ファイルを配布するだけです。	新しい RTR ファイルの <code>pkgadd</code> をすべてのノード上で実行します。 新しいリソースタイプを登録します。 リソースを移行します。
メソッドが更新されました。	ANYTIME	更新されたメソッドを古いメソッドとは異なったパスに配置します。	更新されたメソッドの <code>pkgadd</code> をすべてのノード上で実行します。 新しいリソースタイプを登録します。 リソースを移行します。

表 3-1 リソースタイプのアップグレード例 (続き)

変更の種類	Tunable 属性	パッケージ処理	手続き
新しいモニタープログラムです。	WHEN_UNMONITORED	以前のバージョンのモニターを上書きするだけです。	監視を無効化します。 新しい監視プログラムの pkgadd をすべてのノード上で実行します。 新しいリソースタイプを登録します。 リソースを移行します。 監視を有効化します。
メソッドが更新されました。新しい Update/ Stop メソッドと古い Start メソッドには互換性がありません。	WHEN_OFFLINE	更新されたメソッドを古いメソッドとは異なったパスに配置します。	更新されたメソッドの pkgadd をすべてのノード上で実行します。 新しいリソースタイプを登録します。 リソースをオフラインにします。 リソースを移行します。 リソースをオンラインにします。
メソッドが更新され、RTR ファイルに新しいプロパティが追加されました。新しいメソッドには新しいプロパティが必要です。リソースの所属リソースグループをオンラインのまま保持しながらリソースがオンラインになることを防ぐ必要があります。このためには、リソースグループをノード上でオフラインの状態からオンラインの状態に変更します。	WHEN_DISABLED	以前のバージョンのメソッドを上書きします。	リソースを無効にします。 各ノード上で次の処理を行います。 ■ クラスタからノードを削除 ■ 更新されたメソッドの pkgrm/pkgadd の実行 ■ クラスタにノードを戻す 新しいリソースタイプを登録します。 リソースを移行します。 リソースを有効にします。

表 3-1 リソースタイプのアップグレード例 (続き)

変更の種類	Tunable 属性	パッケージ処理	手続き
メソッドが更新され、RTR ファイルに新しいプロパティが追加されました。新しいメソッドは新しいプロパティを必要としません。	ANYTIME	以前のバージョンのメソッドを上書きします。	各ノード上で次の処理を行います。 <ul style="list-style-type: none"> ■ クラスタからノードを削除 ■ 更新されたメソッドの pkgrm/pkgadd の実行 ■ クラスタにノードを戻す この手続きの間に、RGM は新しいメソッドを呼び出します。これは、まだ移行が完了しておらず、新しいプロパティが構成されていない場合でも変わりません。新しいメソッドは、新しいプロパティなしで正常に機能しなければなりません。 新しいリソースタイプを登録します。 リソースを移行します。
メソッドが更新されました。新しい Fini メソッドと古い Init メソッドには互換性がありません。	WHEN_UNMANAGED	更新されたメソッドを古いメソッドとは異なったパスに配置します。	リソースの所属リソースグループを管理対象外にします。 更新されたメソッドの pkgadd をすべてのノード上で実行します。 リソースタイプを登録します。 リソースを移行します。 リソースの所属リソースグループを管理対象にします。
メソッドが更新されました。RTR ファイルは変更されていません。	該当しません。RTR ファイルは変更されていません。	以前のバージョンのメソッドを上書きします。	各ノード上で次の処理を行います。 <ul style="list-style-type: none"> ■ クラスタからノードを削除 ■ 更新されたメソッドの pkgadd を実行 ■ クラスタにノードを戻す RTR ファイルが変更されなかったため、リソースを登録または移行する必要はありません。

リソースタイプパッケージのインストール要件

新しいリソースタイプをインストールするときは、次の2つの要件が満たされていなければなりません。

- リソースタイプが登録されている場合、ディスク上の RTR ファイルにアクセスできなければなりません。
- 新しいタイプのリソースを作成した場合、新しいタイプの宣言済みメソッドのパス名および監視プログラムがディスク上に存在し、実行できなければなりません。リソースが使用されている間は、以前のメソッドおよび監視プログラムを定位置に確保しておく必要があります。

最適のパッケージを決定するため、リソースタイプの実装者は、次のことを考慮する必要があります。

- RTR ファイルが変更されたか
- プロパティのデフォルト値または `tunable` 属性が変更されたか
- プロパティの `min` または `max` 値が変更されたか
- アップグレードによってプロパティが追加されたか、または削除されたか
- メソッドコードが変更されたか
- モニターコードが変更されたか
- 新しいメソッドまたはモニターコードが以前のバージョンのものと互換するか

RTR ファイルの変更前に認識しておくべき事項

リソースタイプをアップグレードしても、新しいメソッドまたはモニターコードが呼び出されない場合があります。たとえば、リソースプロパティのデフォルト値または `tunable` 属性の値だけが変更される場合があります。メソッドコードは変更されないため、読み取り可能な RTR ファイルの有効なパス名を指定するだけでアップグレードをインストールできます。

古いリソースタイプを登録し直す必要がない場合、新しい RTR ファイルによって以前のバージョンのものが上書きされます。それ以外の場合、新しい RTR ファイルは新しいパス名で配置できます。

アップグレードによってプロパティのデフォルト値または `tunable` 属性が変更される場合、移行時に新しいバージョンの `validate` メソッドで既存のプロパティ属性が新しいリソースタイプでも有効かどうかを検査できます。アップグレードによってプロパティの `min`、`max` または `type` 属性が変更される場合、移行時に `scrgadm` コマンドでこれらの制約が自動的に検査されます。

アップグレード文書には、新しいデフォルトのプロパティ値をすべて記載する必要があります。この文書では、システム管理者に、`Type_version` コマンドを編集するコマンドを使って、新しいリソースタイプのバージョンにプロパティをアップグレードできるように、既存のリソースプロパティの値を変更できることを通知します。

アップグレードによってプロパティが追加または削除される場合、コールバックメソッドまたはモニターコードの一部を変更しなければならないことがあります。

モニターコードの変更

更新後のリソースタイプでモニターコードだけが変更される場合、パッケージのインストールによってモニターのバイナリが上書きされます。文書には、システム管理者向けの情報として、新しいパッケージをインストールする前に監視を一時停止する指示を記述します。

メソッドコードの変更

更新後のリソースタイプでメソッドコードだけが変更される場合、新しいメソッドコードが以前のバージョンのものと互換するかどうかを確認する必要があります。これにより、新しいメソッドコードを新しいパス名で格納するか、古いメソッドを上書きするかが決定します。

新しい Stop、Postnet_stop および Fini メソッドが宣言されていて、古いバージョンの Start、Prenet_stop または Init メソッドによって初期化または起動されたリソースに適用できる場合、新しいメソッドで古いメソッドを上書きできます。

新しいメソッドコードが以前のバージョンのものと互換しない場合、アップグレードされたリソースタイプに移行する前に、古いバージョンのメソッドを使ってリソースを停止または構成を解除する必要があります。新しいメソッドが古いメソッドを上書きする場合、リソースタイプをアップグレードする前に、そのタイプのすべてのリソースを停止 (場合によっては、さらに管理対象外に設定) しなければならないことがあります。新しいメソッドが古いものと別の場所に格納されていて、両方に同時にアクセスできる場合は、後方互換性がなくても、新しいリソースタイプのバージョンをインストールし、リソースを1つずつアップグレードすることができます。

新しいメソッドに後方互換性がある場合でも、その他のリソースが古いメソッドを使用し続けている間は、リソースを1つずつアップグレードして、新しいメソッドを使用できるようにする必要があります。また、新しいメソッドは、古いメソッドを上書きすることがないように別のディレクトリに格納する必要があります。

個々のリソースタイプのバージョンのメソッドを別々のディレクトリに格納すると、新しいバージョンで問題が発生したとき、元のリソースタイプのバージョンに切り替える手続きが簡単であるという点で有利です。

引き続きサポートされている以前のバージョンをすべてパッケージに含めるという方法もあります。この方法では、古いメソッドのパスを上書きまたは削除することなく、新しいパッケージのバージョンで古いバージョンを置き換えることができます。サポートされる以前のバージョンの数は、リソースタイプ開発者が決定します。

注 - メソッド、または現在クラスタ内にあるノード上の `pkgrm/ pkgadd` メソッドの上書きはお勧めしません。RGM がディスク上のアクセス不能なメソッドを呼び出そうとすると、予測できない結果を招くことがあります。実行中のメソッドのバイナリの削除または置き換えでも、予測できない結果を招く可能性があります。

第 4 章

Resource Management API リファレンス

この章では、Resource Management API (RMAPI) を構成するアクセス関数やコールバックメソッドに関する情報を提供します。ここでは、各関数やメソッドについて簡単に説明します。詳細については、RMAPI のそれぞれのマニュアルページを参照してください。

この章の内容は、次のとおりです。

- 72 ページの「RMAPI アクセスメソッド」 - シェルスクリプトコマンドと C 関数
 - `scha_resource_get(1HA)`、`scha_resource_close(3HA)`、`scha_resource_get(3HA)`、`scha_resource_open(3HA)`
 - `scha_resource_setstatus(1HA)`、`scha_resource_setstatus(3HA)`
 - `scha_resourcetype_get(1HA)`、`scha_resourcetype_close(3HA)`、`scha_resourcetype_get(3HA)`、`scha_resourcetype_open(3HA)`
 - `scha_resourcegroup_get(1HA)`、`scha_resourcegroup_get(3HA)`、`scha_resourcegroup_close(3HA)`、`scha_resourcegroup_open(3HA)`
 - `scha_control(1HA)`、`scha_control(3HA)`
 - `scha_cluster_get(1HA)`、`scha_cluster_close(3HA)`、`scha_cluster_get(3HA)`、`scha_cluster_open(3HA)`
 - `scha_cluster_getlogfacility(3HA)`
 - `scha_cluster_getnodename(3HA)`
 - `scha_streerror(3HA)`
- 77 ページの「RMAPI コールバックメソッド」 - `rt_callbacks(1HA)` のマニュアルページで説明されている内容
 - `Start`
 - `Stop`
 - `Init`
 - `Fini`
 - `Boot`
 - `Prenet_start`

- Postnet_stop
- Monitor_start
- Monitor_stop
- Monitor_check
- Update
- ValidateS

RMAPI アクセスメソッド

API は、リソース、リソースタイプ、リソースグループのプロパティ、および他のクラスタ情報にアクセスするための関数を提供します。これらの関数はシェルコマンドと C 関数の両方の形で提供されるため、リソースタイプの開発者はシェルスクリプトまたは C プログラムのどちらでも制御プログラムを実装できます。

RMAPI シェルコマンド

シェルコマンドは、クラスタの RGM によって制御されるサービスを表すリソースタイプのコールバックメソッドを、シェルスクリプトで実装するときに使用します。このコマンドを使用すると、次のことを行えます。

- リソース、リソースタイプ、リソースグループ、クラスタについての情報にアクセスする。
- モニターと併用し、リソースの Status プロパティと Status_msg プロパティを設定する。
- リソースグループの再起動と再配置を要求する。

注 - この節では、シェルコマンドについて簡単に説明します。詳細については、個々のマニュアルページのセクション 1HA を参照してください。特に注記しないかぎり、各コマンドと同じ名前前のマニュアルページがあります。

RMAPI リソースコマンド

以下のコマンドを使用すると、リソースについての情報にアクセスしたり、リソースの Status プロパティや Status_msg プロパティを設定できます。

scha_resource_get

RGM の制御下のリソースまたはリソースタイプに関する情報にアクセスできます。このコマンドは、scha_resource_get () 関数と同じ情報を提供します。

scha_resource_setstatus

RGM の制御下のリソースの Status および Status_msg プロパティを設定します。このコマンドはリソースのモニターによって使用され、モニターから見たリ

ソースの状態を反映します。このコマンドは、C 関数 `scha_resource_setstatus()` と同じ機能を提供します。

注 - `scha_resource_setstatus()` はリソースモニター専用の関数ですが、任意のプログラムから呼び出すことができます。

リソースタイプコマンド

このコマンドは、RGM に登録されているリソースタイプについての情報にアクセスします。

`scha_resourcetype_get`
このコマンドは、C 関数 `scha_resourcegroup_get()` と同じ機能を提供します。

リソースグループコマンド

以下のコマンドを使用すると、リソースグループについての情報にアクセスしたり、リソースグループを再起動できます。

`scha_resourcegroup_get`
RGM の制御下のリソースグループに関する情報にアクセスできます。このコマンドは、C 関数 `scha_resourcegroup_get()` と同じ機能を提供します。

`scha_control`
RGM の制御下のリソースグループの再起動、または別のノードへの再配置を要求します。このコマンドは、C 関数 `scha_control()` と同じ機能を提供します。

クラスタコマンド

このコマンドは、クラスタについての情報(ノード名、ノード ID、ノードの状態、クラスタ名、リソースグループなど)にアクセスします。

`scha_cluster_get`
このコマンドは、C 関数 `scha_cluster_get()` と同じ情報を提供します。

C 関数

C 関数は、クラスタの RGM によって制御されるサービスを表すリソースタイプのコールバックメソッドを、C プログラムで実装するときに使用します。この関数を使用すると、次のことを行えます。

- リソース、リソースタイプ、リソースグループ、クラスタについての情報にアクセスする。

- モニターと併用し、リソースの `Status` プロパティと `Status_msg` プロパティを設定する。
- リソースグループの再起動と再配置を要求する。
- エラーコードを適切なエラーメッセージに変換する。

注 - この節では、C 関数について簡単に説明します。詳細については、各関数の (3HA) マニュアルページを参照してください。特に注記しないかぎり、各関数と同じ名前のマニュアルページがあります。C 関数の出力関数および戻りコードについては、`scha_calls` (3HA) のマニュアルページを参照してください。

リソース関数

以下の関数は、RGM に管理されているリソースについての情報にアクセスします。モニターから見たリソースの状態を表します。

`scha_resource_open()`、`scha_resource_get()`、`scha_resource_close()`
これらの関数は、RGM によって管理されるリソースの情報にアクセスします。
`scha_resource_open()` 関数は、リソースへのアクセスを初期化し、
`scha_resource_get()` のハンドルを戻します。`scha_resource_get()` 関数は、リソースの情報にアクセスします。`scha_resource_close()` 関数は、ハンドルを無効にし、`scha_resource_get()` の戻り値に割り当てられているメモリーを解放します。

`scha_resource_open()` 関数がリソースのハンドルを戻したあとに、クラスタの再構成や管理アクションによって、リソースが変更されることがあります。この場合、`scha_resource_get()` 関数がハンドルを通じて獲得した情報は正しくない可能性があります。リソースでクラスタの再構成や管理アクションが行われた場合、RGM は `scha_err_seqid` エラーコードを `scha_resourcegroup_get()` 関数に戻し、リソースが変更されたことを示します。このメッセージは致命的なエラーメッセージではないため、関数は正常に終了します。したがって、このメッセージは無視してもかまいません。また、現在のハンドルを閉じて新しいハンドルを開き、リソースの情報にアクセスし直してもかまいません。

これら 3 つの関数は 1 つのマニュアルページ内で説明しています。このマニュアルページには、個々の関数名 `scha_resource_open`(3HA)、`scha_resource_get`(3HA)、`scha_resource_close`(3HA) でアクセスできます。

`scha_resource_setstatus()`
RGM の制御下のリソースの `Status` および `Status_msg` プロパティを設定します。この関数はリソースのモニターによって使用され、モニターから見たリソースの状態を反映します。

注 - `scha_resource_setstatus()` はリソースモニター専用の関数ですが、任意のプログラムから呼び出すことができます。

リソースタイプ関数

これらの関数は、RGMに登録されているリソースタイプについての情報にアクセスします。

`scha_resourcetype_open()`、`scha_resourcetype_get()`、`scha_resourcetype_close()`

`scha_resourcetype_open()` 関数は、リソースへのアクセスを初期化し、`scha_resourcetype_get()` のハンドルを戻します。
`scha_resourcetype_get()` 関数はリソースタイプの情報にアクセスします。
`scha_resourcetype_close()` 関数は、ハンドルを無効にし、`scha_resourcetype_get()` の戻り値に割り当てられているメモリーを解放します。

`scha_resourcetype_open()` 関数がリソースタイプのハンドルを戻したあとに、クラスタの再構成や管理アクションによって、リソースタイプが変更されることがあります。この場合、`scha_resourcetype_get()` 関数がハンドルを通じて獲得した情報は正しくない可能性があります。リソースタイプ上でクラスタの再構成や管理アクションが行われた場合、RGMは `scha_err_seqid` エラーコードを `scha_resourcetype_get()` 関数に戻し、リソースタイプが変更されたことを示します。このメッセージは致命的なエラーメッセージではないため、関数は正常に終了します。したがって、このメッセージは無視してもかまいません。また、現在のハンドルを閉じて新しいハンドルを開き、リソースタイプの情報にアクセスし直してもかまいません。

これら3つの関数は1つのマニュアルページ内で説明しています。このマニュアルページには、個々の関数名 `scha_resourcetype_open(3HA)`、`scha_resourcetype_get(3HA)`、`scha_resourcetype_close(3HA)` からアクセスできます。

リソースグループ関数

以下の関数を使用すると、リソースグループについての情報にアクセスしたり、リソースグループを再起動できます。

`scha_resourcegroup_open(3HA)`、`scha_resourcegroup_get(3HA)`、`scha_resourcegroup_close(3HA)`

これらの関数は、RGMによって管理されるリソースグループの情報にアクセスします。`scha_resourcegroup_open()` 関数は、リソースグループへのアクセスを初期化し、`scha_resourcegroup_get()` のハンドルを戻します。
`scha_resourcegroup_get()` 関数は、リソースグループの情報にアクセスします。`scha_resourcegroup_close()` 関数は、ハンドルを無効にし、`scha_resourcegroup_get()` の戻り値に割り当てられているメモリーを解放します。

`scha_resourcegroup_open()` 関数がリソースグループのハンドルを戻したあとに、クラスタの再構成や管理アクションによって、リソースグループが変更されることがあります。この場合、`scha_resourcegroup_get()` 関数がハンドルを通じて獲得した情報は正しくない可能性があります。リソースグループ上でクラスタ

の再構成や管理アクションが行われた場合、RGM は `scha_err_seqid` エラーコードを `scha_resourcegroup_get()` 関数に戻し、リソースグループが変更されたことを示します。このメッセージは致命的なエラーメッセージではないため、関数は正常に終了します。したがって、このメッセージは無視してもかまいません。また、現在のハンドルを閉じて新しいハンドルを開き、リソースグループの情報にアクセスし直してもかまいません。

`scha_control(3HA)`

RGM の制御下のリソースグループの再起動、または別のノードへの再配置を要求します。

クラスタ関数

以下の関数は、クラスタについての情報にアクセスし、その情報を戻します。

`scha_cluster_open(3HA)`、`scha_cluster_get(3HA)`、

`scha_cluster_close(3HA)`

これらの関数は、ノード名、ID、状態、クラスタ名、リソースグループなど、クラスタに関する情報にアクセスします。

`scha_cluster_open()` 関数がクラスタのハンドルを戻したあとに、再構成や管理アクションによって、クラスタが変更されることがあります。この場合、`scha_cluster_get()` 関数がハンドルを通じて獲得した情報は正しくない可能性があります。クラスタ上でクラスタの再構成や管理アクションが行われた場合、RGM は `scha_err_seqid` エラーコードを `scha_cluster_get()` 関数に戻し、クラスタが変更されたことを示します。このメッセージは致命的なエラーメッセージではないため、関数は正常に終了します。したがって、このメッセージは無視してもかまいません。また、現在のハンドルを閉じて新しいハンドルを開き、クラスタの情報にアクセスし直してもかまいません。

`scha_cluster_getlogfacility(3HA)`

クラスタログとして使用されるシステムログ機能の数を戻します。戻された番号を Solaris の `syslog()` 関数で使用すると、イベントと状態メッセージをクラスタログに記録できます。

`scha_cluster_getnodename(3HA)`

関数が呼び出されたクラスタノードの名前を戻します。

ユーティリティ関数

この関数は、エラーコードをエラーメッセージに変換します。

`scha_strerror(3HA)`

`scha_` 関数によって戻されるエラーコードを適切なエラーメッセージに変換します。この関数を `logger` と共に使用すると、メッセージをシステムログ (`syslog`) に記録できます。

RMAPI コールバックメソッド

コールバックメソッドは、リソースタイプを実装するための API が提供する重要な要素です。コールバックメソッドを使用すると、RGM は、クラスタのメンバーシップが変更されたとき (ノードが起動またはクラッシュしたとき) にクラスタ内のリソースを制御できます。

注 - クライアントプログラムがクラスタシステム上の HA サービスを制御するため、コールバックメソッドはルートのアクセス権を持つ RGM によって実行されます。したがって、このようなコールバックメソッドをインストールおよび管理するときは、ファイルの所有権とアクセス権を制限します。特に、このようなコールバックメソッドには、特権付き所有者 (bin や root など) を割り当てます。

さらに、このようなコールバックメソッドは、書き込み可能にはなりません。この節では、コールバックメソッドの引数と終了コードについて説明し、次のカテゴリのコールバックメソッドについて説明します。

- 制御および初期化メソッド
- 管理サポートメソッド
- ネットワーク関連メソッド
- モニター制御メソッド

注 - この節では、メソッドが呼び出されるタイミングやよそうされるリソースへの影響など、コールバックメソッドについて簡単に説明します。詳細については `rt_callbacks (1HA)` のマニュアルページを参照してください。

メソッドの引数

RGM はコールバックメソッドを呼び出すとき、次のような引数を使用します。

```
method -R resource-name -T type-name -G group-name
```

method は、`start` や `stop` などのコールバックメソッドとして登録されているプログラムのパス名です。リソースタイプのコールバックメソッドは、それらの登録ファイルで宣言します。

コールバックメソッドの引数はすべて、フラグ付きの値として渡されます。-R はリソースインスタンスの名前を示し、-T はリソースのタイプを示し、-G はリソースが構成されているグループを示します。このような引数をアクセス関数で使用すると、リソースについての情報を取得できます。

`validate` メソッドを呼び出すときは、追加の引数 (リソースのプロパティ値と呼び出されるリソースグループ) を使用します。

詳細については、`scha_calls(3HA)` のマニュアルページを参照してください。

終了コード

終了コードは、すべてのコールバックメソッドで共通で、メソッドの呼び出しによるリソースの状態への影響を示すように定義されています。これらの終了コードについては、`scha_calls(3HA)` のマニュアルページを参照してください。終了コードには、以下のものがあります。

- 0 (ゼロ) – メソッドは成功しました。
- ゼロ以外の任意の値 – メソッドは失敗しました。

RGM は、コールバックメソッドの実行の異常終了 (タイムアウトやコアダンプ) も処理します。

メソッドは、各ノード上で `syslog` を使用して障害情報を出力するように実装する必要があります。 `stdout` や `stderr` に書き込まれる出力は、ローカルノードのコンソール上には表示されますが、それをユーザーが確認するかどうかは保証できないためです。

制御および初期化コールバックメソッド

制御および初期化コールバックメソッドは、主に、リソースを起動および停止します。その他にも、リソース上で初期化と終了コードを実行します。

Start

この必須メソッドは、リソースを含むリソースグループがクラスタノード上でオンラインになったとき、そのクラスタノード上で呼び出されます。このメソッドは、そのノード上でリソースを起動します。

Start メソッドは、ローカルノード上でリソースが起動し、使用可能な状態になるまで終了してはなりません。したがって、Start メソッドは終了する前にリソースをポーリングし、リソースが起動しているかどうかを判断する必要があります。さらに、このメソッドには、十分な長さのタイムアウト値を設定する必要があります。たとえば、あるリソース (データベースデーモンなど) が起動するのに時間がかかる場合、そのメソッドには十分な長さのタイムアウト値を設定する必要があります。

RGM が Start メソッドの失敗に応答する方法は、`Failover_mode` プロパティの設定によって異なります。

リソースの Start メソッドのタイムアウト値を設定するには、リソースタイプ登録ファイルの `START_TIMEOUT` プロパティを使用します。

Stop

この必須メソッドは、リソースを含むリソースグループがクラスタノード上でオフラインになったとき、そのクラスタノード上で呼び出されます。このメソッドは、リソースを (アクティブであれば) 停止します。

Stop メソッドは、ローカルノード上でリソースがすべての活動を完全に停止し、すべてのファイル記述子を終了するまで終了してはなりません。そうしないと、RGM が (実際にはアクティブであるのに) リソースが停止したと判断するため、データが破壊されることがあります。データの破壊を防ぐために最も安全な方法は、ローカルノード上でリソースに関連するすべてのプロセスを停止することです。

Stop メソッドは終了する前にリソースをポーリングし、リソースが停止しているかどうかを判断する必要があります。さらに、このメソッドには、十分な長さのタイムアウト値を設定する必要があります。たとえば、あるリソース (データベースデーモンなど) が停止するのに時間がかかる場合、そのメソッドには十分長めのタイムアウト値を設定する必要があります。

RGM が Stop メソッドの失敗に応答する方法は、Failover_mode プロパティの設定によって異なります (245 ページの「リソースのプロパティ」を参照)。

リソースの Stop メソッドのタイムアウト値を設定するには、リソースタイプ登録ファイルの STOP_TIMEOUT プロパティを使用します。

Init

このオプションメソッドは、リソースを管理下に置くとき (リソースが属しているリソースグループを管理していない状態から管理している状態に切り替えるとき、またはすでに管理されているリソースグループでリソースを作成するとき)、1 回だけ呼び出され、リソースの初期化を実行します。このメソッドは、Init_nodes リソースプロパティが示すノード上で呼び出されます。

Fini

このオプションメソッドは、リソースを管理下から外すとき (リソースが属しているリソースグループを管理していない状態に切り替えるとき、またはすでに管理されているリソースグループからリソースを削除するとき) に呼び出され、リソースをクリーンアップします。このメソッドは、Init_nodes リソースプロパティが示すノード上で呼び出されます。

Boot

このオプションメソッドは、Init と同様に、リソースの所属リソースグループが RGM の管理下に置かれたあと、クラスタを結合するノード上のリソースを初期化します。このメソッドは、Init_nodes リソースプロパティが示すノード上で呼び出されます。Boot メソッドは、起動または再起動の結果とし、ノードがクラスタに結合または再結合したときに呼び出されます。

注 - Init、Fini、Boot メソッドが失敗すると、syslog() 関数がエラーメッセージを生成しますが、それ以外は RGM のリソース管理に影響しません。

管理サポートメソッド

リソース上での管理アクションには、リソースプロパティの設定と変更があります。Validate と Update コールバックメソッドを使用してリソースタイプを実装すると、このような管理アクションを行うことができます。

Validate

この任意メソッドは、リソースが作成される時、および、リソースまたはリソースグループ (リソースを含む) のプロパティが管理アクションによって更新されるときに呼び出されます。このメソッドは、リソースタイプの `Init_nodes` プロパティが示す複数のクラスタノード上で呼び出されます。`validate` は、作成または更新が行われる前に呼び出されます。任意のノード上でメソッドから失敗の終了コードが戻ると、作成または更新は取り消されます。

`Validate` メソッドが呼び出されるのは、管理アクションによってリソースまたはリソースグループのプロパティが変更されたときのみです。`RGM` がプロパティを設定したときや、モニターがリソースプロパティ `Status` と `Status_msg` を設定したときではありません。

Update

このオプションメソッドを呼び出して、プロパティが変更されたことを実行中のリソースに通知することができます。`Update` は、管理アクションがリソースまたはリソースグループのプロパティの設定に成功したあとに呼び出されます。このメソッドは、リソースがオンラインであるノード上で呼び出されます。このメソッドは、API アクセス関数を使用し、アクティブなリソースに影響する可能性があるプロパティ値を読み取り、その値に従って実行中のリソースを調節します。

`Update` メソッドが失敗すると、`syslog()` 関数がエラーメッセージを生成しますが、それ以外は `RGM` のリソース管理に影響しません。

ネットワーク関連コールバックメソッド

ネットワークアドレスリソースを使用するサービスでは、ネットワークアドレス構成から始まる順番で、起動手順または停止手順を実行する必要があります。任意コールバックメソッドの `Prenet_start` と `Postnet_stop` を使用してリソースタイプを実装すると、関連するネットワークアドレスが「起動」に構成される前、または、「停止」に構成されたあとに、特別な起動アクションまたはシャットダウンアクションを行うことができます。

`Prenet_start`

このオプションメソッドを呼び出して、同じリソースグループ内のネットワークアドレスが構成される前に特殊な起動アクションを実行することができます。

`Postnet_stop`

このオプションメソッドを呼び出して、同じリソースグループ内のネットワークアドレスの構成後に特殊な終了アクションを実行することができます。

モニター制御コールバックメソッド

リソースタイプは、オプションとして、リソースの性能を監視したり、その状態を報告したり、リソースの障害に対処するようなプログラムを含むようにも実装できます。`Monitor_start`、`Monitor_stop`、`Monitor_check` メソッドは、リソースタイプへのリソースモニターの実装をサポートします。

Monitor_start

このオプションメソッドを呼び出して、リソースの起動後にリソースの監視を開始することができます。

Monitor_stop

この任意メソッドは、リソースが停止する前に呼び出され、リソースのモニターを停止します。

Monitor_check

このオプションメソッドを呼び出して、リソースグループを新しいノードに配置する前に、そのノードの信頼性を査定することができます。Monitor_check メソッドは、並行して実行中のその他のメソッドと競合しない方法で実装する必要があります。

第 5 章

サンプルデータサービス

この章では、`in.named` アプリケーションを Sun Cluster データサービスとして稼働する HA-DNS について説明します。`in.named` デーモンは Solaris におけるドメインネームサービス (DNS) の実装です。サンプルのデータサービスでは、Resource Management API を使用して、アプリケーションの高可用性を実現する方法を示します。

RMAPI は、シェルスクリプトと C プログラムの両方のインタフェースをサポートします。この章のサンプルアプリケーションはシェルスクリプトインタフェースで作成されています。

この章の内容は、次のとおりです。

- 83 ページの「サンプルデータサービスの概要」
- 84 ページの「リソースタイプ登録ファイルの定義」
- 90 ページの「すべてのメソッドに共通な機能の提供」
- 94 ページの「データサービスの制御」
- 100 ページの「障害モニターの定義」
- 109 ページの「プロパティ更新の処理」

サンプルデータサービスの概要

サンプルのデータサービスはクラスタのイベント (管理アクション、アプリケーションの異常終了、ノードの異常終了など) に応じて、DNS アプリケーションの起動、停止、再起動や、クラスタノード間での DNS アプリケーションの切り替えを行います。

アプリケーションの再起動は、プロセス監視機能 (PMF) によって管理されます。アプリケーションの障害が再試行最大期間または再試行最大回数を超えると、障害モニターはアプリケーションリソースを含むリソースグループを別のノードにフェイルオーバーします。

サンプルのデータサービスは、PROBE メソッドという形で障害監視機能を提供します。PROBE メソッドは、nslookup コマンドを使用し、アプリケーションが正常な状態であることを保証します。DNS サービスのハングを検出すると、PROBE メソッドは、DNS アプリケーションをローカルで再起動することによって、この状況を修正しようとします。この方法で状況が改善されず、サービスの問題が繰り返し検出される場合、PROBE メソッドは、サービスをクラスタ内の別のノードにフェイルオーバーしようとします。

サンプルのアプリケーションには、具体的に、次のような機能が含まれています。

- リソースタイプ登録ファイル - データサービスの静的なプロパティを定義します。
- Start コールバックメソッド - HA-DNS データサービスを含むリソースグループがオンラインになるときに RGM によって呼び出され、in.named デーモンを起動します。
- Stop コールバックメソッド - HA-DNS データサービスを含むリソースグループがオフラインになるときに RGM によって呼び出され、in.named デーモンを停止します。
- 障害モニター - DNS サーバーが動作しているかどうかを確認することによって、サービスの信頼性を検査します。障害モニターはユーザー定義の PROBE メソッドによって実装され、Monitor_start と Monitor_stop コールバックメソッドによって起動および停止されます。
- Validate コールバックメソッド - RGM によって呼び出され、サービスの構成ディレクトリがアクセス可能であるかどうかを検査します。
- Update コールバックメソッド - システム管理者がリソースプロパティの値を変更したときに RGM によって呼び出され、障害モニターを再起動します。

リソースタイプ登録ファイルの定義

この例で使用するサンプルのリソースタイプ登録 (RTR) ファイルは、DNS リソースタイプの静的な構成を定義します。このタイプのリソースは、RTR ファイルで定義されているプロパティを継承します。

RTR ファイル内の情報は、クラスタ管理者が HA-DNS データサービスを登録したときに RGM によって読み取られます。

RTR ファイルの概要

RTR ファイルの形式は明確に定義されています。リソースタイププログラム、システム定義リソースプロパティ、拡張プロパティという順番で並んでいます。詳細は、rt_reg(4) のマニュアルページと30 ページの「リソースとリソースタイププロパティの設定」を参照してください。

この節では、サンプルの RTR ファイルの特定のプロパティについて説明します。この節で扱うリストは、サンプルの RTR ファイルの一部だけです。サンプルの RTR ファイルの完全なリストについては、267 ページの「リソースタイプ登録ファイルのリスト」を参照してください。

サンプル RTR ファイルのリソースタイププロパティ

次のリストに示すように、サンプルの RTR ファイルはコメントから始まり、そのあとに、HA-DNS 構成を定義するリソースタイププロパティが続きます。

```
#
# Copyright (c) 1998-2004 by Sun Microsystems, Inc.
# All rights reserved.
#
# Registration information for Domain Name Service (DNS)
#

#pragma ident    "@(#)SUNW.sample  1.1  00/05/24 SMI"

RESOURCE_TYPE = "sample";
VENDOR_ID = SUNW;
RT_DESCRIPTION = "Domain Name Service on Sun Cluster";

RT_VERSION = "1.0";
API_VERSION = 2;
FAILOVER = TRUE;

RT_BASEDIR=/opt/SUNWsample/bin;
PKGLIST = SUNWsample;

START      =  dns_svc_start;
STOP       =  dns_svc_stop;

VALIDATE   =  dns_validate;
UPDATE     =  dns_update;

MONITOR_START =  dns_monitor_start;
MONITOR_STOP  =  dns_monitor_stop;
MONITOR_CHECK =  dns_monitor_check;
```

ヒント – RTR ファイルの最初のエントリには、Resource_type プロパティを宣言する必要があります。宣言しないと、リソースタイプの登録は失敗します。

注 - RGM はプロパティ名の大文字と小文字を区別します。Sun が提供する RTR ファイルのプロパティに対する命名規則では、名前の最初の文字が大文字で、残りが小文字です (メソッド名は例外)。メソッド名は、プロパティ属性と同様にすべて大文字です。

次に、これらのプロパティについての情報を説明します。

- リソースタイプ名は、Resource_type プロパティだけで指定できます (例: sample)。あるいは、Vendor_id + "." + Resource_type という形式でも指定できます (例: SUNW.sample)。
Vendor_id を使用する場合、リソースタイプを定義する企業の略号にします。リソースタイプ名はクラスタ内で一意である必要があります。
- RT_version プロパティは、ベンダーによって指定されたサンプルのデータサービスのバージョンを識別します。
- API_version プロパティは Sun Cluster のバージョンを識別します。たとえば、「API_version =2」は、データサービスが Sun Cluster バージョン 3.0 の管理下で動作していることを示します。
- Failover = TRUE の場合、同時に複数のノード上でオンラインになることができるリソースグループでは、データサービスが動作できないことを示します。
- RT_basedir は相対パス (コールバックメソッドのパスなど) を補完するためのディレクトリパスで、/opt/SUNWsample/bin を指します。
- Start、Stop、Validate などは、RGM によって呼び出される個々のコールバックメソッドプログラムへのパスを提供します。これらのパスは、RT_basedir で指定されたディレクトリからの相対パスです。
- Pkglist は、SUNWsample をサンプルのデータサービスのインストールを含むパッケージとして識別します。

この RTR ファイルに指定されていないリソースタイププロパティ (Single_instance、Init_nodes、Installed_nodes など) は、デフォルト値を取得します。リソースタイププロパティの完全なリストとそのデフォルト値については、239 ページの「リソースタイププロパティ」を参照してください。

クラスタ管理者は、RTR ファイルのリソースタイププロパティに指定されている値を変更できません。

サンプル RTR ファイルのリソースプロパティ

慣習上、RTR ファイルでは、リソースプロパティをリソースタイププロパティのあとに宣言します。リソースプロパティには、Sun Cluster が提供するシステム定義プロパティと、データサービス開発者が定義する拡張プロパティが含まれます。どちらのタイプの場合でも、Sun Cluster が提供するプロパティ属性の数 (最小、最大、デフォルト値など) を指定できます。

RTR ファイルのシステム定義プロパティ

次のリストは、サンプル RTR ファイルのシステム定義プロパティを示しています。

```
#リソースタイプ宣言のあとに、中括弧に囲まれたリソースプロパティ宣言の
# リストが続く。プロパティ名宣言は、各エントリの左中括弧の直後にある
# 最初の属性である必要がある。
# <method>_timeout プロパティの値は、RGM がメソッドの呼び出しが
# 失敗したと結論するまでの時間（秒）を設定する。
# すべてのメソッドタイムアウトの MIN 値は 60 秒に設定されている。
# これは、管理者が短すぎる時間を設定するのを防ぐためである。短すぎる
# 時間を設定すると、スイッチオーバーやフェイルオーバーの性能が上がらず、
# さらには予期せぬ RGM アクションが発生する可能性がある（誤った
# フェイルオーバー、ノードの再起動、リソースグループの
# ERROR_STOP_FAILED 状態への移行など、管理者の介入を必要とする
# RGM アクション）。
# メソッドタイムアウトに短すぎる時間を設定すると、データサービス全体の
# 可用性が低下する。
{
    PROPERTY = Start_timeout;
    MIN=60;
    DEFAULT=300;
}

{
    PROPERTY = Stop_timeout;
    MIN=60;
    DEFAULT=300;
}

{
    PROPERTY = Validate_timeout;
    MIN=60;
    DEFAULT=300;
}

{
    PROPERTY = Update_timeout;
    MIN=60;
    DEFAULT=300;
}

{
    PROPERTY = Monitor_Start_timeout;
    MIN=60;
    DEFAULT=300;
}

{
    PROPERTY = Monitor_Stop_timeout;
    MIN=60;
    DEFAULT=300;
}

{
    PROPERTY = Thorough_Probe_Interval;
    MIN=1;
    MAX=3600;
    DEFAULT=60;
    TUNABLE = ANYTIME;
}
```

```

}

# 当該ノード上でアプリケーションを正常に起動できないと結論するまでに
# 指定された時間内 (Retry_Interval) に行う再試行回数
{
    PROPERTY = Retry_Count;
    MIN=0;
    MAX=10;
    DEFAULT=2;
    TUNABLE = ANYTIME;
}

# Retry_Interval には 60 の倍数を指定する。これは、この値は秒から分に変換され、
# 端数が切り上げられるためである。
# たとえば、50 (秒) は 1 分に変換される。このプロパティ値は再試行回数
# (Retry_Count) のタイミングを指定する。
{
    PROPERTY = Retry_Interval;
    MIN=60;
    MAX=3600;
    DEFAULT=300;
    TUNABLE = ANYTIME;
}

{
    PROPERTY = Network_resources_used;
    TUNABLE = AT_CREATION;
    DEFAULT = "";
}

```

Sun Cluster はシステム定義プロパティを提供しますが、リソースプロパティ属性を使用すると、異なるデフォルト値を設定できます。リソースプロパティに適用するために利用できる属性の完全なリストについては、[263 ページの「リソースプロパティの属性」](#)を参照してください。

サンプルの RTR ファイル内のシステム定義リソースプロパティについては、次の点に注意してください。

- Sun Cluster は、すべてのタイムアウトに最小値 (1 秒) とデフォルト値 (3600 秒) を提供します。サンプルの RTR ファイルは、最小値をそのまま (60 秒) にし、デフォルト値を 300 秒に変更しています。クラスタ管理者は、このデフォルト値を使用することも、タイムアウト値を変更することもできます (たとえば、60 秒以上)。Sun Cluster は正当な最大値を持っていません。
- Thorough_Probe_Interval、Retry_count、Retry_interval プロパティの TUNABLE 属性は ANYTIME に設定されています。この設定は、データサービスが動作中でも、クラスタ管理者がこれらのプロパティの値を変更できることを意味します。上記のプロパティは、サンプルのデータサービスによって実装される障害モニターによって使用されます。サンプルのデータサービスは、管理アクションによってさまざまなリソースが変更されたときに障害モニターを停止および再起動するように、Update を実装します。詳細は、[114 ページの「Update メソッド」](#)を参照してください。
- リソースプロパティは次のように分類されます。

- 必須—クラスタ管理者はリソースを作成するときに必ず値を指定する必要があります。
- 任意—クラスタ管理者が値を指定しない場合、システムがデフォルト値を提供します。
- 条件付き—RTR ファイルで宣言されている場合だけ、RGM はプロパティを作成します。

サンプルのデータサービスの障害モニターは、`Thorough_probe_interval`、`Retry_count`、`Retry_interval`、`Network_resources_used` という条件付きプロパティを使用しているため、開発者はこれらのプロパティを RTR ファイルで宣言する必要があります。プロパティの分類方法については、`r_properties` (5) のマニュアルページまたは 245 ページの「リソースのプロパティ」を参照してください。

RTR ファイルの拡張プロパティ

次に、RTR ファイルの最後の例として、拡張プロパティを示します。

```
# 拡張プロパティ
# クラスタ管理者は、このプロパティの値として、アプリケーションによって
# 使用される構成ファイルが格納されているディレクトリのパスを指定する。
# このアプリケーション (DNS) は、PXFS 上の DNS 構成ファイルのパス (通常
# named.conf) のパスを指定する。
{
    PROPERTY = Confdir;
    EXTENSION;
    STRING;
    TUNABLE = AT_CREATION;
    DESCRIPTION = "The Configuration Directory Path";
}

# 検証の失敗が宣言されるまでのタイムアウト値 (秒)。
{
    PROPERTY = Probe_timeout;
    EXTENSION;
    INT;
    DEFAULT = 120;
    TUNABLE = ANYTIME;
    DESCRIPTION = "Time out value for the probe (seconds)";
}
```

サンプルの RTR ファイルは 2 つの拡張プロパティ、`Confdir` と `Probe_timeout` を定義します。`Confdir` は、DNS 構成ディレクトリへのパスを指定します。このディレクトリには、DNS が正常に動作するために必要な `in.named` ファイルが格納されています。サンプルのデータサービスの `Start` と `Validate` メソッドはこのプロパティを使用し、DNS を起動する前に、構成ディレクトリと `in.named` ファイルがアクセス可能であるかどうかを確認します。

データサービスが構成されるとき、`Validate` メソッドは、新しいディレクトリがアクセス可能であるかどうかを確認します。

サンプルのデータサービスの PROBE メソッドは、Sun Cluster コールバックメソッドではなく、ユーザー定義メソッドです。したがって、Sun Cluster はこの Probe_timeout プロパティを提供しません。開発者はこの拡張プロパティを RTR ファイルに定義し、クラスタ管理者が Probe_timeout の値を構成できるようにする必要があります。

すべてのメソッドに共通な機能の提供

この節では、サンプルのデータサービスのすべてのコールバックメソッドで使用される次のような機能について説明します。

- 90 ページの「コマンドインタプリタの指定およびパスのエクスポート」。
- 90 ページの「PMF_TAG と SYSLOG_TAG 変数の宣言」。
- 91 ページの「関数の引数の構文解析」。
- 93 ページの「エラーメッセージの生成」。
- 93 ページの「プロパティ情報の取得」。

コマンドインタプリタの指定およびパスのエクスポート

シェルスクリプトの最初の行は、コマンドインタプリタを指定します。サンプルのデータサービスの各メソッドスクリプトは、次に示すように、コマンドインタプリタを指定します。

```
#!/bin/ksh
```

サンプルアプリケーション内のすべてのメソッドスクリプトは、Sun Cluster のバイナリとライブラリへのパスをエクスポートします。ユーザーの PATH 設定には依存しません。

```
#####  
# MAIN  
#####  
  
export PATH=/bin:/usr/bin:/usr/cluster/bin:/usr/sbin:/usr/proc/bin:$PATH
```

PMF_TAG と SYSLOG_TAG 変数の宣言

すべてのメソッドスクリプト (Validate を除く) は、リソース名を渡し、pmfadm を使用してデータサービスまたはモニターのいずれかを起動 (または停止) します。各スクリプトは変数 PMF_TAG を定義し、pmfadm に渡すことによって、データサービスまたはモニターを識別できます。

同様に、各メソッドスクリプトは、logger コマンドを使用してメッセージをシステムログに記録します。各スクリプトは変数 SYSLOG_TAG を定義し、-t オプションで logger に渡すことによって、メッセージが記録されるリソースのリソースタイプ、リソースグループ、リソース名を識別できます。

すべてのメソッドは、次に示す例と同じ方法で SYSLOG_TAG を定義します。dns_probe、dns_svc_start、dns_svc_stop、および dns_monitor_check の各メソッドは、次のように PMF_TAG を定義します。なお、pmfadm と logger は dns_svc_stop のものを使用しています。

```
#####
# MAIN
#####

PMF_TAG=$RESOURCE_NAME.named

SYSLOG_TAG=$RESOURCETYPE_NAME,$RESOURCEGROUP_NAME,$RESOURCE_NAME

# データサービスに SIGTERM シグナルを送り、タイムアウト値の 80%
# が経過するまで待機する。
pmfadm -s $PMF_TAG.named -w $SMOOTH_TIMEOUT TERM
if [ $? -ne 0 ]; then
    logger -p ${SYSLOG_FACILITY}.info \
        -t [${SYSLOG_TAG}] \
        "${ARGV0} Failed to stop HA-DNS with SIGTERM; Retry with \
        SIGKILL"
```

dns_monitor_stop、dns_monitor_stop、dns_update の各メソッドは、次のように PMF_TAG を定義します (なお、pmfadm は dns_monitor_stop メソッドのものを使用しています)。

```
#####
# MAIN
#####

PMF_TAG=$RESOURCE_NAME.monitor
SYSLOG_TAG=$RESOURCETYPE_NAME,$RESOURCEGROUP_NAME,$RESOURCE_NAME
...

# in.named が実行中であるかどうかを確認し、実行中であれば強制終了する。
if pmfadm -q $PMF_TAG.monitor; then
    pmfadm -s $PMF_TAG.monitor KILL
```

関数の引数の構文解析

RGM は、次に示すように、すべてのコールバックメソッド (Validate を除く) を呼び出します。

```
method_name -R resource_name -T resource_type_name -G resource_group_name
```

`method_name` は、コールバックメソッドを実装するプログラムのパス名です。データサービスは、各メソッドのパス名を RTR ファイルに指定します。このようなパス名は、RTR ファイルの `RT_basedir` プロパティに指定されたディレクトリからのパスになります。たとえば、サンプルのデータサービスの RTR ファイルでは、ベースディレクトリとメソッド名は次のように指定されます。

```
RT_BASEDIR=/opt/SUNWsample/bin;
Start = dns_svc_start;
Stop = dns_svc_stop;
...
```

コールバックメソッドの引数はすべて、フラグ付きの値として渡されます。-R はリソースインスタンスの名前を示し、-T はリソースのタイプを示し、-G はリソースが構成されているグループを示します。コールバックメソッドの詳細については、`rt_callbacks (1HA)` のマニュアルページを参照してください。

注 - `Validate` メソッドを呼び出すときは、追加の引数 (リソースのプロパティ値と呼び出されるリソースグループ) を使用します。詳細は、[109 ページの「プロパティ更新の処理」](#) を参照してください。

各コールバック、メソッドには、渡された引数を構文解析する関数が必要です。すべてのコールバックメソッドには同じ引数が渡されるので、データサービスは、アプリケーション内のすべてのコールバックメソッドで使用される単一の構文解析関数を提供します。

次に、サンプルのアプリケーションのコールバックメソッドで使用される `parse_args ()` 関数を示します。

```
#####
# プログラム引数の解析。
#
function parse_args # [args ...]
{
    typeset opt

    while getopts 'R:G:T:' opt
    do
        case "$opt" in
            R)
                # DNS リソース名。
                RESOURCE_NAME=$OPTARG
                ;;
            G)
                # リソースが構成されたリソースグループ名。
                RESOURCEGROUP_NAME=$OPTARG
                ;;
            T)
                # リソースタイプ名。
                RESOURCETYPE_NAME=$OPTARG
                ;;
        esac
    done
}
```

```

*)
  logger -p ${SYSLOG_FACILITY}.err \
  -t [$RESOURCE_NAME,$RESOURCEGROUP_NAME,$RESOURCE_NAME] \
  "ERROR: Option $OPTARG unknown"
  exit 1
  ;;
esac
done
}

```

注 - サンプルのアプリケーションの PROBE メソッドはユーザー定義メソッドですが、Sun Cluster コールバックメソッドと同じ引数で呼び出されます。したがって、このメソッドには、他のコールバックメソッドと同じ構文解析関数が含まれています。

構文解析関数は、次に示すように、MAIN の中で呼び出されます。

```
parse_args "$@"
```

エラーメッセージの生成

エンドユーザーに対してエラーメッセージを出力するには、syslog 機能をメソッドに使用することを推奨します。サンプルのデータサービスのすべてのコールバックメソッドは、次に示すように、scha_cluster_get () 関数を使用し、クラスタログ用に使用されている syslog 機能番号を取得します。

```
SYSLOG_FACILITY='scha_cluster_get -O SYSLOG_FACILITY'
```

この値はシェル変数 SYSLOG_FACILITY に格納されます。logger コマンドの機能として使用すると、エラーメッセージをクラスタログに記録できます。たとえば、サンプルのデータサービスの Start メソッドは、次に示すように、SYSLOG_FACILITY を取得し、データサービスが起動したことを示すメッセージを記録します。

```

SYSLOG_FACILITY='scha_cluster_get -O SYSLOG_FACILITY'
...
if [ $? -eq 0 ]; then
  logger -p ${SYSLOG_FACILITY}.err \
  -t [$SYSLOG_TAG] \
  "${ARGV0} HA-DNS successfully started"
fi

```

詳細については、scha_cluster_get (1HA) のマニュアルページを参照してください。

プロパティ情報の取得

ほとんどのコールバックメソッドは、データサービスのリソースとリソースタイプのプロパティについての情報を取得する必要があります。このために、API は scha_resource_get () 関数を提供しています。

リソースプロパティには2種類(システム定義プロパティと拡張プロパティ)あります。システム定義プロパティは事前に定義されており、拡張プロパティはデータサービス開発者が RTR ファイルに定義します。

`scha_resource_get()` を使用してシステム定義プロパティの値を取得するときは、`-O` パラメータでプロパティの名前を指定します。このコマンドは、プロパティの値だけを返します。たとえば、サンプルのデータサービスの `Monitor_start` メソッドは検証プログラムを特定し、起動できるようにしておく必要があります。検証プログラムはデータサービスのベースディレクトリ (`RT_basedir` プロパティが指す位置) 内に存在します。したがって、`Monitor_start` メソッドは、次に示すように、`RT_basedir` の値を取得し、その値を `RT_basedir` 変数に格納します。

```
RT_BASEDIR=`scha_resource_get -O RT_BASEDIR -R $RESOURCE_NAME -G \  
$RESOURCEGROUP_NAME`
```

拡張プロパティの場合、データサービス開発者は、これが拡張プロパティであることを示す `-O` パラメータを指定し、最後のパラメータとしてプロパティの名前を指定する必要があります。拡張プロパティの場合、このコマンドは、プロパティのタイプと値の両方を返します。たとえば、サンプルのデータサービスの検証プログラムは、次に示すように、`probe_timeout` 拡張プロパティのタイプと値を取得し、次に `awk` コマンドを使用して値だけを `PROBE_TIMEOUT` シェル変数に格納します。

```
probe_timeout_info=`scha_resource_get -O Extension -R $RESOURCE_NAME \  
-G $RESOURCEGROUP_NAME Probe_timeout`  
PROBE_TIMEOUT=`echo $probe_timeout_info | awk '{print $2}'`
```

データサービスの制御

データサービスは、クラスタ上でアプリケーションデーモンを起動するために `Start` メソッドまたは `Prenet_start` メソッドを提供し、クラスタ上でアプリケーションデーモンを停止するために `Stop` メソッドまたは `Postnet_stop` メソッドを提供する必要があります。サンプルのデータサービスは、`Start` メソッドと `Stop` メソッドを実装します。代わりに `Prenet_start` メソッドと `Postnet_stop` メソッドを使用する場合は、41 ページの「[Start および Stop メソッドを使用するかどうかの決定](#)」を参照してください。

Start メソッド

データサービスリソースを含むリソースグループがクラスタノード上でオンラインになるとき、あるいは、リソースグループがすでにオンラインになっていて、そのリソースが有効なとき、RGM はそのノード上で `Start` メソッドを呼び出します。サンプルのアプリケーションでは、`Start` メソッドはそのノード上で `in.named` (DNS) デーモンを起動します。

この節では、サンプルのアプリケーションの Start メソッドの重要な部分だけを説明します。parse_args() 関数や syslog 機能番号を取得する方法など、すべてのコールバックメソッドに共通な機能については説明しません。このような機能については、90 ページの「すべてのメソッドに共通な機能の提供」を参照してください。

Start メソッドの完全なリストについては、270 ページの「Start メソッド」を参照してください。

Start の概要

DNS を起動する前に、サンプルのデータサービスの Start メソッドは、構成ディレクトリと構成ファイル (named.conf) がアクセス可能で利用可能であるかどうかを確認します。DNS が正常に動作するためには、named.conf の情報が重要です。

このコールバックメソッドは、PMF (pmfadm) を使って DNS デーモン (in.named) を起動します。DNS がクラッシュしたり、起動に失敗したりすると、このメソッドは、一定の期間に一定の回数だけ DNS の起動を再試行します。再試行の回数と期間は、データサービスの RTR ファイル内のプロパティで指定されます。

構成の確認

DNS が動作するためには、構成ディレクトリ内の named.conf ファイルからの情報が必要です。したがって、Start メソッドは、DNS を起動しようとする前にいくつかの妥当性検査を実行し、ディレクトリやファイルがアクセス可能であるかどうかを確認します。

Confdir 拡張プロパティは、構成ディレクトリへのパスを提供します。プロパティ自身は RTR ファイルに定義されています。しかし、実際の位置は、クラスタ管理者がデータサービスを構成するときに指定します。

サンプルのデータサービスでは、Start メソッドは scha_resource_get() 関数を使用して構成ディレクトリの位置を取得します。

注 - Confdir は拡張プロパティであるため、scha_resource_get() はタイプと値の両方を返します。したがって、awk コマンドで値だけを取得し、シェル変数 CONFIG_DIR に格納します。

```
# クラスタ管理者がリソースの追加時に設定した Confdir の値を検索。
config_info=`scha_resource_get -O Extension -R $RESOURCE_NAME \
-G $RESOURCEGROUP_NAME Confdir`

# scha_resource_get は、拡張プロパティの値とともにタイプを返す。
# 拡張プロパティの値だけを取得。
CONFIG_DIR=`echo $config_info | awk '{print $2}'`
```

次に、Start メソッドは CONFIG_DIR の値を使用し、ディレクトリがアクセス可能であるかどうかを確認します。アクセス可能ではない場合、Start メソッドはエラーメッセージを記録し、エラー状態で終了します。97 ページの「Start の終了状態」を参照してください。

```
# CONFIG_DIR がアクセス可能かどうかを確認。
if [ ! -d $CONFIG_DIR ]; then
  logger -p ${SYSLOG_FACILITY}.err \
    -t [SYSLOG_TAG] \
    "${ARGV0} Directory $CONFIG_DIR is missing or not mounted"
  exit 1
fi
```

アプリケーションデーモンを起動する前に、このメソッドは最終検査を実行し、named.conf ファイルが存在するかどうかを確認します。存在しない場合、Start メソッドはエラーメッセージを記録し、エラー状態で終了します。

```
# データファイルに相対パス名が含まれている場合は $CONFIG_DIR
# ディレクトリに移動。
cd $CONFIG_DIR

# named.conf ファイルが $CONFIG_DIR ディレクトリに存在することを確認。
if [ ! -s named.conf ]; then
  logger -p ${SYSLOG_FACILITY}.err \
    -t [SYSLOG_TAG] \
    "${ARGV0} File $CONFIG_DIR/named.conf is missing or empty"
  exit 1
fi
```

アプリケーションの起動

このメソッドは、プロセス監視機能 (pmfadm) を使用してアプリケーションを起動します。pmfadm コマンドを使用すると、アプリケーションを再起動するときの期間と回数を指定できます。このため、RTR ファイルには2つのプロパティ Retry_count と Retry_interval があります。Retry_count は、アプリケーションを再起動する回数を指定し、Retry_interval は、アプリケーションを再起動する期間を指定します。

Start メソッドは、scha_resource_get () 関数を使用して Retry_count と Retry_interval の値を取得し、これらの値をシェル変数に格納します。次に、-n オプションと -t オプションを使用し、これらの値を pmfadm に渡します。

```
# RTR ファイルから再試行回数の値を取得する。
RETRY_CNT=`scha_resource_get -O Retry_Count -R $RESOURCE_NAME \
-G $RESOURCEGROUP_NAME`
# RTR ファイルから次の再試行までの時間 (秒数) を取得。この値は pmfadm
# に渡されるため分数に変換される。変換時に端数が切り上げられる点に注意。
# たとえば 50 秒は 1 分に変換される。
((RETRY_INTRVAL=`scha_resource_get -O Retry_Interval -R $RESOURCE_NAME \
-G $RESOURCEGROUP_NAME` / 60))

# PMF 制御下で in.named デーモンを起動する。RETRY_INTERVAL の期間、
```



```

# $RETRY_COUNT の回数だけクラッシュおよび再起動できる。
# それ以上の回数クラッシュした場合、PMF はそれ以上再試行しない。
# タグで登録済みのプロセスがある場合、PMF はプロセスが
# すでに実行中であるという警告メッセージを送信する。
pmfadm -c $PMF_TAAG -n $RETRY_CNT -t $RETRY_INTRVAL \
    /usr/sbin/in.named -c named.conf

# HA-DNS が起動していることを示すメッセージを記録。
if [ $? -eq 0 ]; then
    logger -p ${SYSLOG_FACILITY}.err \
        -t [$SYSLOG_TAG] \
        "${ARGV0} HA-DNS successfully started"
fi
exit 0

```

Start の終了状態

Start メソッドは、実際のアプリケーションが本当に動作して実行可能になるまで、成功状態で終了してはなりません。特に、ほかのデータサービスが依存している場合は注意する必要があります。正常に終了したかどうかを検証するための1つの方法は、Start メソッドが終了する前に、アプリケーションが動作しているかどうかを確認することです。複雑なアプリケーション(データベースなど)の場合、RTR ファイルの `Start_timeout` プロパティに十分高い値を設定することによって、アプリケーションが初期化され、クラッシュ回復を実行できる時間を提供します。

注 - サンプルのデータサービスのアプリケーションリソース DNS は直ちに起動するため、サンプルのデータサービスは、成功状態である前に、ポーリングでアプリケーションが動作していることを確認していません。

このメソッドが DNS の起動に失敗し、失敗状態で終了すると、RGM は `Failover_mode` プロパティを検査し、どのように対処するかを決定します。サンプルのデータサービスは明示的に `Failover_mode` プロパティを設定していないため、このプロパティはデフォルト値 `NONE` が設定されています(ただし、クラスタ管理者がデフォルトを変更して異なる値を指定していないと仮定します)。したがって、RGM は、データサービスの状態を設定するだけで、ほかのアクションは行いません。同じノード上で再起動したり、別のノードにフェイルオーバーしたりするには、ユーザーの介入が必要です。

Stop メソッド

HA-DNS リソースを含むリソースグループがクラスタノード上でオフラインになるとき、あるいはリソースグループがオンラインでリソースが無効なとき、Stop メソッドが呼び出されます。このメソッドは、そのノード上で `in.named` (DNS) デモンを停止します。

この節では、サンプルのアプリケーションの Stop メソッドの重要な部分だけを説明します。parse_args() 関数や syslog 機能番号を取得する方法など、すべてのコールバックメソッドに共通な機能については説明しません。このような機能については、90 ページの「すべてのメソッドに共通な機能の提供」を参照してください。

Stop メソッドの完全なリストについては、273 ページの「Stop メソッド」を参照してください。

Stop の概要

データサービスを停止するときは、考慮すべきことが 2 点あります。1 点は、停止処理を正しい順序で行うことです。これを実現する最良の方法は、pmfadm 経由で SIGTERM シグナルを送信することです。

もう 1 点は、データサービスが本当に停止していることを保証することによって、データベースが Stop_failed 状態にならないようにすることです。これを実現する最良の方法は、pmfadm 経由で SIGKILL シグナルを送信することです。

サンプルのデータサービスの Stop メソッドは、このような点を考慮しています。まず、SIGTERM シグナルを送信し、このシグナルがデータサービスの停止に失敗した場合は、SIGKILL シグナルを送信します。

DNS を停止しようとする前に、この Stop メソッドは、プロセスが実際に動作しているかどうかを確認します。プロセスが動作している場合には、PMF (pmfadm) を使ってこれを停止します。

この Stop メソッドは何回か呼びだしてもその動作が変わらないことが保証されます。RGM は、Start の呼び出しでデータサービスを起動せずに、Stop メソッドを 2 回呼び出すことはありません。しかし、RGM は、リソースが起動されていなくても、あるいは、リソースが自発的に停止している場合でも、Stop メソッドをそのリソース上で呼び出すことができます。つまり、DNS がすでに動作していない場合でも、この Stop メソッドは成功状態です。

アプリケーションの停止

Stop メソッドは、データサービスを停止するために 2 段階の方法を提供します。pmfadm 経由で SIGTERM シグナルを使用する規則正しい方法と、SIGKILL シグナルを使用する強制的な方法です。Stop メソッドは、Stop メソッドが戻るまでの時間を示す stop_timeout 値を取得します。次に、Stop メソッドはこの時間の 80% を規則正しい方法に割り当て、15% を強制的な方法に割り当てます (5% は予約済み)。次の例を参照してください。

```
STOP_TIMEOUT=`scha_resource_get -O STOP_TIMEOUT -R $RESOURCE_NAME  
\`  
  
-G $RESOURCEGROUP_NAME  
( (SMOOTH_TIMEOUT=$STOP_TIMEOUT * 80/100) )  
  
( (HARD_TIMEOUT=$STOP_TIMEOUT * 15/100) )
```

Stop メソッドは `pmfadm -q` を使用し、DNS デーモンが動作しているかどうかを確認します。動作している場合、Stop メソッドはまず `pmfadm -s` を使用して TERM シグナルを送信し、DNS プロセスを終了します。このシグナルを送信してからタイムアウト値の 80% が経過してもプロセスが終了しない場合、Stop メソッドは SIGKILL シグナルを送信します。このシグナルを送信してからタイムアウト値の 15% が経過してもプロセスが終了しない場合、STOP メソッドはエラーメッセージを記録し、エラー状態で終了します。

`pmfadm` がプロセスを終了した場合、STOP メソッドはプロセスが停止したことを示すメッセージを記録し、成功状態で終了します。

DNS プロセスが動作していない場合、STOP メソッドは DNS プロセスが動作していないことを示すメッセージを記録しますが、成功状態で終了します。次のコード例に、Stop メソッドがどのように `pmfadm` を使用して DNS プロセスを停止するかを示します。

```
# in.named が実行中であるかどうかを確認し、実行中であれば強制終了する。
if pmfadm -q $PMF_TAG; then
# データサービスに SIGTERM シグナルを送り、タイムアウト値の 80%
# が経過するまで待機する。
pmfadm -s $RESOURCE_NAME.named -w $SMOOTH_TIMEOUT TERM
if [ $? -ne 0 ]; then
    logger -p ${SYSLOG_FACILITY}.err \
        -t [$RESOURCE_TYPE_NAME,$RESOURCE_GROUP_NAME,$RESOURCE_NAME] \
        "${ARGV0} Failed to stop HA-DNS with SIGTERM; Retry with \
        SIGKILL"

# SIGTERM シグナルでデータサービスが停止しないので、今度は SIGKILL を
# 使って、合計タイムアウト値の残りの 15% が経過するまで待機する。
pmfadm -s $PMF_TAG -w $HARD_TIMEOUT KILL
if [ $? -ne 0 ]; then
    logger -p ${SYSLOG_FACILITY}.err \
        -t [$SYSLOG_TAG]
        "${ARGV0} Failed to stop HA-DNS; Exiting UNSUCCESSFUL"

    exit 1
fi
fi
else
# この時点でデータサービスは実行されていない。メッセージを記録して
# 成功状態で終了する。

logger -p ${SYSLOG_FACILITY}.err \
    -t [$SYSLOG_TAG] \
    "HA-DNS is not started"

# HA-DNS が実行中でなくても、データサービスリソースを STOP_FAILED
# 状態にするのを避けるため成功状態で終了する。

exit 0

fi

# DNS の停止に成功。メッセージを記録して成功状態で終了する。
```

```
logger -p ${SYSLOG_FACILITY}.err \  
-t [$RESOURCE_TYPE_NAME,$RESOURCEGROUP_NAME,$RESOURCE_NAME]  
\  
"HA-DNS successfully stopped"  
exit 0
```

Stop の終了状態

Stop メソッドは、実際のアプリケーションが本当に停止するまで、成功状態で終了してはなりません。特に、ほかのデータサービスが依存している場合は注意する必要があります。そうしなければ、データが破壊される可能性があります。

複雑なアプリケーション (データベースなど) の場合、RTR ファイルの Stop_timeout プロパティに十分高い値を設定することによって、アプリケーションが停止中にクリーンアップできる時間を提供します。

このメソッドが DNS の停止に失敗し、失敗状態で終了すると、RGM は Failover_mode プロパティを検査し、どのように対処するかを決定します。サンプルのデータサービスは明示的に Failover_mode プロパティを設定していないため、このプロパティはデフォルト値 NONE が設定されています (ただし、クラスタ管理者がデフォルトを変更して異なる値を指定していないと仮定します)。したがって、RGM は、データサービスの状態を Stop_failed に設定するだけで、ほかのアクションは行いません。アプリケーションを強制的に停止し、Stop_failed 状態をクリアするには、ユーザーの介入が必要です。

障害モニターの定義

サンプルのアプリケーションは、DNS リソース (in.named) の信頼性を監視する基本的な障害モニターを実装します。障害モニターは、次の要素から構成されます。

- dns_probe - nslookup を使用し、サンプルのデータサービスの制御下にある DNS リソースが動作しているかどうかを確認するユーザー定義プログラム。DNS が動作していない場合、このメソッドは DNS をローカルで再起動しようとします。あるいは、再起動の再試行回数によっては、RGM がデータサービスを別のノードに再配置することを要求します。
- dns_monitor_start - dns_probe を起動するコールバックメソッド。監視が有効である場合、RGM は、サンプルのデータサービスがオンラインになった後、自動的に dns_monitor_start を呼び出します。
- dns_monitor_stop - dns_probe を停止するコールバックメソッド。RGM は、サンプルのデータサービスがオフラインになる前に、自動的に dns_monitor_stop を呼び出します。
- dns_monitor_check - PROBE プログラムがデータサービスを新しいノードにフェイルオーバーするとき、Validate メソッドを呼び出し、構成ディレクトリが利用可能であるかどうかを確認するコールバックメソッド。

検証プログラム

dns_probe プログラムは、サンプルのデータサービスの管理下にある DNS リソースが動作しているかどうかを確認する、連続して動作するプロセスを実行します。dns_probe は、サンプルのデータサービスがオンラインになったあと、RGM によって自動的に呼び出される dns_monitor_start メソッドによって起動されます。データサービスは、サンプルのデータサービスがオフラインになる前、RGM によって呼び出される dns_monitor_stop メソッドによって停止されます。

この節では、サンプルのアプリケーションの PROBE メソッドの重要な部分だけを説明します。parse_args() 関数や syslog 機能番号を取得する方法など、すべてのコールバックメソッドに共通な機能については説明しません。このような機能については、90 ページの「すべてのメソッドに共通な機能の提供」を参照してください。

PROBE メソッドの完全なリストについては、276 ページの「PROBE プログラム」を参照してください。

検証プログラムの概要

検証プログラムは無限ループで動作します。検証プログラムは、nslookup を使用し、適切な DNS リソースが動作しているかどうかを確認します。DNS が動作している場合、検証プログラムは一定の期間 (Thorough_probe_interval システム定義プロパティに設定されている期間) だけ休眠し、その後、再び検証を行います。DNS が動作していない場合、検証プログラムは DNS をローカルで再起動しようとするか、再起動の再試行回数によっては、RGM がデータサービスを別のノードに再配置することを要求します。

プロパティ値の取得

このプログラムには、次のプロパティ値が必要です。

- Thorough_probe_interval - 検証プログラムが休眠する期間を設定します。
- Probe_timeout - nslookup コマンドが検証を行う期間 (タイムアウト値) を設定します。
- Network_resources_used - DNS が動作するサーバーを設定します。
- Retry_count と Retry_interval - 再起動を行う回数と期間を設定します。
- RT_basedir - PROBE プログラムと gettime ユーティリティが格納されているディレクトリを取得します。

scha_resource_get() 関数は、次に示すように、上記プロパティの値を取得し、シェル変数に格納します。

```
PROBE_INTERVAL=`scha_resource_get -O THOROUGH_PROBE_INTERVAL \  
-R $RESOURCE_NAME -G $RESOURCEGROUP_NAME`
```

```
probe_timeout_info=`scha_resource_get -O Extension -R $RESOURCE_NAME
```

```

\  

-G $RESOURCEGROUP_NAME Probe_timeout\  

PROBE_TIMEOUT='echo $probe_timeout_info | awk '{print $2}''  

\  

DNS_HOST='scha_resource_get -O NETWORK_RESOURCES_USED -R $RESOURCE_NAME\  

\  

-G $RESOURCEGROUP_NAME\  

\  

RETRY_COUNT='scha_resource_get -O RETRY_COUNT -R $RESOURCE_NAME\  

-G\  

  $RESOURCEGROUP_NAME\  

\  

RETRY_INTERVAL='scha_resource_get -O RETRY_INTERVAL -R $RESOURCE_NAME\  

-G\  

  $RESOURCEGROUP_NAME\  

\  

RT_BASEDIR='scha_resource_get -O RT_BASEDIR -R $RESOURCE_NAME -G\  

  $RESOURCEGROUP_NAME'

```

注 - システム定義プロパティ (Thorough_probe_interval など) の場合、
 scha_resource_get () は値だけを返します。拡張プロパティ (Probe_timeout など) の場合、
 scha_resource_get () はタイプと値を返します。値だけを取得するには
 awk コマンドを使用します。

サービスの信頼性の検査

検査プログラム自身は、nslookup コマンドの while による無限ループです。
 while ループの前に、nslookup の応答を保管する一時ファイルを設定します。
 probefail 変数と retries 変数は 0 に初期化されます。

```

# nslookup の応答用の一時ファイルを設定。DNSPROBEFILE=/tmp/.$RESOURCE_NAME.probe
probfail=0
retries=0

```

while ループ自身は、次の作業を行います。

- 検査プログラム用の休眠期間を設定します。
- hatimerun を使用し、nslookup に Probe_timeout の値とターゲットホストを渡して起動します。
- nslookup の戻りコード (成功または失敗) に基づいて、probfail 変数を設定します。
- probefail が 1 (失敗) に設定された場合、nslookup への応答がサンプルのデータサービスから来ており、他の DNS サーバーから来ているのではないことを確認します。

次に、while ループコードを示します。

```

while :
do
  # 検査が実行される時間は THOROUGH_PROBE_INTERVAL プロパティ

```

```

# に指定されている。したがって、THOROUGH_PROBE_INTERVAL の間
# 検証が sleep するように設定する。
sleep $PROBE_INTERVAL

# DNS がサービスを提供している IP アドレス上で nslookup コマンドを実行する。
hatimerun -t $PROBE_TIMEOUT /usr/sbin/nslookup $DNS_HOST $DNS_HOST \
> $DNSPROBEFILE 2>&1

    retcode=$?
    if [ $retcode -ne 0 ]; then
        probefail=1
    fi

# nslookup への応答が/etc/resolv.conf ファイルに指定されている
# その他のネームサーバーではなく HA-DNS サーバーから返されて
# いることを確認する。
if [ $probefail -eq 0 ]; then
# nslookup 照会に応答したサーバーの名前を取得する。
SERVER=`awk ' $1!="Server:" { print $2 }' \
$DNSPROBEFILE | awk -F. ' { print $1 } ' `
if [ -z "$SERVER" ]; then
    probefail=1
else
    if [ $SERVER != $DNS_HOST ]; then
        probefail=1
    fi
fi
fi
fi

```

再起動とフェイルオーバーの評価

probefail 変数が 0 (成功) 以外である場合、nslookup コマンドがタイムアウトしたか、あるいは、サンプルのサービスの DNS 以外のサーバーから応答が来ていることを示します。どちらの場合でも、DNS サーバーは期待どおりに機能していないので、障害モニターは `decide_restart_or_failover()` 関数を呼び出し、データサービスをローカルで起動するか、RGM がデータサービスを別のノードに再配置することを要求するかを決定します。*probefail* 変数が 0 の場合、検証が成功したことを示すメッセージが生成されます。

```

if [ $probefail -ne 0 ]; then
    decide_restart_or_failover
else
    logger -p ${SYSLOG_FACILITY}.err \
-t [${SYSLOG_TAG}] \
"${ARGV0} Probe for resource HA-DNS successful"
fi

```

`decide_restart_or_failover()` 関数は、再試行最大期間 (`Retry_interval`) と再試行最大回数 (`Retry_count`) を使用し、DNS をローカルで再起動するか、RGM がデータサービスを別のノードに再配置することを要求するかを決定します。この関数は、次のような条件付きコードを実装します。コードリストについては、276 ページの「PROBE プログラム」にある `decide_restart_or_failover()` を参照してください。

- 最初の障害である場合、データサービスをローカルで再起動します。エラーメッセージを記録し、retries 変数の再試行カウンタをインクリメントします。
- 最初の障害ではなく、再試行時間が再試行最大期間を過ぎている場合、データサービスをローカルで再起動します。エラーメッセージを記録し、再試行カウンタをリセットし、再試行時間をリセットします。
- 再試行時間が再試行最大期間を過ぎておらず、再試行カウンタが再試行最大回数を超えている場合、別のノードにフェイルオーバーします。フェイルオーバーが失敗すると、エラーメッセージを記録し、検証プログラムを状態 1 (失敗) で終了します。
- 再試行時間が再試行最大期間を過ぎておらず、再試行カウンタが再試行最大回数を超えていない場合、データサービスをローカルで再起動します。エラーメッセージを記録し、retries 変数の再試行カウンタをインクリメントします。

期限 (再試行最大期間) 内に再起動の回数 (再試行カウンタ) が制限 (再試行最大回数) に到達した場合、この関数は、RGM がデータサービスを別のノードに再配置することを要求します。再起動の回数が制限に到達していない場合、あるいは、再試行最大期間を過ぎていて、再試行カウンタをリセットする場合、この関数は DNS を同じノード上で再起動しようとします。この関数については、次の点に注意してください。

- gettimeofday ユーティリティを使用すると、再起動間の時間を追跡できます。このユーティリティは C プログラムで、(RT_basedir) ディレクトリ内にあります。
- Retry_count と Retry_interval のシステム定義リソースプロパティは、再起動を行う回数と期間を決定します。RTR ファイルのデフォルト値は、Retry_count が 2 回、Retry_interval が 5 分 (300 秒) です。クラスタ管理者はこのデフォルトを変更できます。
- restart_service() 関数は、同じノード上でデータサービスの再起動を試行する場合に呼び出されます。この関数の詳細については、104 ページの「データサービスの再起動」を参照してください。
- API 関数 scha_control() は、GIVEOVER オプションを指定すると、サンプルデータサービスを含むリソースグループをオフラインにし、別のノード上でオンラインにし直します。

データサービスの再起動

restart_service() 関数は、decide_restart_or_failover() によって呼び出され、同じノード上でデータサービスの再起動を試行します。この関数は次の作業を行います。

- データサービスが PMF 下にまだ登録されているかどうかを調べます。サービスが登録されている場合、この関数は次の作業を行います。
 - データサービスの Stop メソッド名と Stop_timeout 値を取得します。
 - hatimerun を使用してデータサービスの Stop メソッドを起動し、Stop_timeout 値を渡します。

- (データサービスが正常に停止した場合) データサービスの Start メソッド名と Start_timeout 値を取得します。
- hatimerun を使用してデータサービスの Start メソッドを起動し、Start_timeout 値を渡します。
- データサービスが PMF 下に登録されていない場合は、データサービスが PMF 下で許可されている再試行最大回数を超過していることを示しています。したがって、GIVEOVER オプションを指定して scha_control() 関数を呼び出し、データサービスを別のノードにフェイルオーバーします。

```
function restart_service
{
    # データサービスを再起動するには、まずデータサービス自体が PMF に
    # 登録されているかどうかを確認する。

    pmfadm -q $PMF_TAG
    if [[ $? -eq 0 ]]; then
        # データサービスの TAG が PMF に登録されている場合、データ
        # サービスを停止し、再起動する。
        # 当該リソースの Stop メソッド名と
        # STOP_TIMEOUT 値を取得する。
        STOP_TIMEOUT=`scha_resource_get -O STOP_TIMEOUT \
            -R $RESOURCE_NAME -G $RESOURCEGROUP_NAME`
        STOP_METHOD=`scha_resource_get -O STOP \
            -R $RESOURCE_NAME -G $RESOURCEGROUP_NAME`
        hatimerun -t $STOP_TIMEOUT $RT_BASEDIR/$STOP_METHOD \
            -R $RESOURCE_NAME -G $RESOURCEGROUP_NAME \
            -T $RESOURCE_TYPE_NAME

        if [[ $? -ne 0 ]]; then
            logger-p ${SYSLOG_FACILITY}.err -t [[SYSLOG_TAG] \
                "${ARGV0} Stop method failed."
            return 1
        fi

        # 当該リソースの Start メソッド名と START_TIMEOUT 値を取得する。
        START_TIMEOUT=`scha_resource_get -O START_TIMEOUT \
            -R $RESOURCE_NAME -G $RESOURCEGROUP_NAME`
        START_METHOD=`scha_resource_get -O START \
            -R $RESOURCE_NAME -G $RESOURCEGROUP_NAME`
        hatimerun -t $START_TIMEOUT $RT_BASEDIR/$START_METHOD \
            -R $RESOURCE_NAME -G $RESOURCEGROUP_NAME \
            -T $RESOURCE_TYPE_NAME

        if [[ $? -ne 0 ]]; then
            logger-p ${SYSLOG_FACILITY}.err -t [[SYSLOG_TAG] \
                "${ARGV0} Start method failed."
            return 1
        fi
    fi

    else
        # データサービスの TAG が PMF に登録されていない場合、
```

```

# データサービスが PMF で許可されている最大再試行回数
# を超過している。したがって、データサービスを再起動しては
# ならない。代わりに、同じクラスタ内の別のノードへの
# フェイルオーバーを試みる。
scha_control -O GIVEOVER -G $RESOURCEGROUP_NAME \
              -R $RESOURCE_NAME
fi

return 0
}

```

検証プログラムの終了状態

ローカルでの再起動が失敗したり、別のノードへのフェイルオーバーが失敗したりすると、サンプルのデータサービスの PROBE プログラムは失敗で終了し、Failover attempt failed (フェイルオーバーは失敗しました)" というエラーメッセージを記録します。

Monitor_start メソッド

サンプルのデータサービスがオンラインになったあと、RGM は Monitor_start メソッドを呼び出し、dns_probe メソッドを起動します。

この節では、サンプルアプリケーションの Monitor_start メソッドの重要な部分だけを説明します。parse_args() 関数や syslog 機能番号を取得する方法など、すべてのコールバックメソッドに共通な機能については説明しません。このような機能については、90 ページの「すべてのメソッドに共通な機能の提供」を参照してください。

Monitor_start メソッドの完全なリストについては、282 ページの「Monitor_start メソッド」を参照してください。

Monitor_start の概要

このメソッドは PMF (pmfadm) を使って検証プログラムを起動します。

検証プログラムの起動

Monitor_start メソッドは、RT_basedir プロパティの値を取得し、PROBE プログラムへの完全パス名を構築します。このメソッドは、pmfadm の無限再試行オプション (-n -1, -t -1) を使用して検証プログラムを起動します。つまり、検証プログラムの起動に失敗しても、PMF メソッドは検証プログラムを無限に再起動します。

```

# リソースの RT_BASEDIR プロパティを取得し、検証プログラムが存在する
# 場所を確認する。
RT_BASEDIR=`scha_resource_get -O RT_BASEDIR -R $RESOURCE_NAME -G \

```

```

$RESOURCEGROUP_NAME`

# PMF の制御下でデータサービスの検証を開始する。無限再試行オプションを使って
# 検証プログラムを起動する。リソースの名前、タイプ、グループを検証
# プログラムに渡す。
pmfadm -c $RESOURCE_NAME.monitor -n -1 -t -1 \
  $RT_BASEDIR/dns_probe -R $RESOURCE_NAME -G $RESOURCEGROUP_NAME \
  -T $RESOURCECETYPE_NAME

```

Monitor_stop メソッド

サンプルのデータサービスがオフラインになるとき、RGM は Monitor_stop メソッドを呼び出し、dns_probe の実行を停止します。

この節では、サンプルアプリケーションの Monitor_stop メソッドの重要な部分だけを説明します。parse_args() 関数や syslog 機能番号を取得する方法など、すべてのコールバックメソッドに共通な機能については説明しません。このような機能については、90 ページの「すべてのメソッドに共通な機能の提供」を参照してください。

Monitor_stop メソッドの完全なリストについては、284 ページの「Monitor_stop メソッド」を参照してください。

Monitor_stop の概要

このメソッドは、PMF (pmfadm) を使用して検証プログラムが動作しているかどうかを判断し、動作している場合は検証プログラムを停止します。

検証プログラムの停止

Monitor_stop メソッドは、pmfadm -q を使用して検証プログラムが動作しているかどうかを判断し、動作している場合は pmfadm -s を使用して検証プログラムを停止します。検証プログラムがすでに停止している場合でも、このメソッドは成功状態です。これによって、メソッドが呼び出し回数に依存しないことが保証されます。

```

# 検証プログラムが動作しているかどうかを判断し、動作している場合は停止する。
if pmfadm -q $PMF_TAG; then
  pmfadm -s $PMF_TAG KILL
  if [ $? -ne 0 ]; then
    logger -p ${SYSLOG_FACILITY}.err \
      -t [${SYSLOG_TAG}] \
      "${ARGV0} Could not stop monitor for resource " \
      $RESOURCE_NAME
    exit 1
  else
    # 検証プログラムの停止に成功。メッセージを記録する。

```

```

logger -p ${SYSLOG_FACILITY}.err \
-t [${SYSLOG_TAG}] \
"${ARGV0} Monitor for resource " $RESOURCE_NAME \
" successfully stopped"
fi
fi
exit 0

```



注意 - 検証プログラムを停止するときは、必ず、pmfadm で KILL シグナルを使用するようにしてください。絶対に、マスク可能なシグナル (TERM など) は使用しないでください。そうしないと、Monitor_stop メソッドが無限にハングし、結果としてタイムアウトする可能性があります。この問題の原因は、PROBE メソッドがデータサービスを再起動またはフェイルオーバーする必要があるときに、scha_control () を呼び出すところにあります。scha_control () がデータサービスをオフラインにするプロセスの一部として Monitor_stop メソッドを呼び出したときに、Monitor_stop メソッドがマスク可能なシグナルを使用していると、Monitor_stop メソッドは scha_control () が終了するのを待ち、scha_control () は Monitor_stop メソッドが終了するのを待つため、結果として両方がハングします。

Monitor_stop の終了状態

PROBE メソッドを停止できない場合、Monitor_stop メソッドはエラーメッセージを記録します。RGM は、主ノード上でサンプルのデータサービスを MONITOR_FAILED 状態にするため、そのノードに障害が発生することがあります。

Monitor_stop メソッドは、検証プログラムが停止するまで終了してはなりません。

Monitor_check メソッド

PROBE メソッドが、データサービスを含むリソースグループを新しいノードにフェイルオーバーしようとするとき、RGM は Monitor_check メソッドを呼び出します。

この節では、サンプルアプリケーションの Monitor_check メソッドの重要な部分だけを説明します。parse_args () 関数や syslog 機能番号を取得する方法など、すべてのコールバックメソッドに共通な機能については説明しません。このような機能については、90 ページの「すべてのメソッドに共通な機能の提供」を参照してください。

Monitor_check メソッドの完全なリストについては、285 ページの「Monitor_check メソッド」を参照してください。

Monitor_check メソッドは、並行して実行中のその他のメソッドと競合しない方法で実装する必要があります。

Monitor_check メソッドは Validate メソッドを呼び出し、新しいノード上で DNS 構成ディレクトリが利用可能かどうかを確認します。Confdir 拡張プロパティが DNS 構成ディレクトリを指します。したがって、Monitor_check は Validate メソッドのパスと名前、および Confdir の値を取得します。Monitor_check は、次のように、この値を Validate に渡します。

```
# リソースタイプの RT_BASEDIR プロパティから Validate メソッドの
# 完全パスを取得する。
RT_BASEDIR=`scha_resource_get -O RT_BASEDIR -R $RESOURCE_NAME \
  -G $RESOURCEGROUP_NAME`

# 当該リソースの Validate メソッド名を取得する。
VALIDATE_METHOD=`scha_resource_get -O VALIDATE \
  -R $RESOURCE_NAME -G $RESOURCEGROUP_NAME`

# データサービスを起動するための Confdir プロパティの値を取得する。
# 入力されたリソース名とリソースグループを使用し、リソースを追加
# するときに設定した Confdir の値を取得する。
config_info=`scha_resource_get -O Extension -R $RESOURCE_NAME \
  -G $RESOURCEGROUP_NAME Confdir`

# scha_resource_get は、Confdir 拡張プロパティの値とともにタイプも戻す。
# awk を使用し、Confdir 拡張プロパティの値だけを取得する。
CONFIG_DIR=`echo $config_info | awk '{print $2}'`

# Validate メソッドを呼び出し、データサービスを新しいノードに
# フェイルオーバーできるかどうかを確認する。
$RT_BASEDIR/$VALIDATE_METHOD -R $RESOURCE_NAME -G $RESOURCEGROUP_NAME \
  -T $RESOURCETYPE_NAME -x Confdir=$CONFIG_DIR
```

ノードがデータサービスのホストとして最適であるかどうかをサンプルアプリケーションが確認する方法については、109 ページの「Validate メソッド」を参照してください。

プロパティ更新の処理

サンプルのデータサービスは、クラスタ管理者によるプロパティの更新を処理するために、Validate メソッドと Update メソッドを実装します。

Validate メソッド

リソースが作成されたとき、および、リソースまたは(リソースを含む)リソースグループのプロパティが管理アクションによって更新されるとき、RGM は Validate メソッドを呼び出します。RGM は、作成または更新が行われる前に、Validate メソッドを呼び出します。任意のノード上でメソッドから失敗の終了コードが戻ると、作成または更新は取り消されます。

RGM が `Validate` メソッドを呼び出すのは、管理アクションがリソースまたはグループのプロパティを変更したときだけです。RGM がプロパティを設定したときや、モニターがリソースプロパティ `Status` と `Status_msg` を設定したときではありません。

注 - `PROBE` メソッドがデータサービスを新しいノードにフェイルオーバーする場合、`Monitor_check` メソッドも明示的に `Validate` メソッドを呼び出します。

Validate の概要

RGM は、他のメソッドに渡す追加の引数 (更新されるプロパティと値を含む) を指定して、`Validate` メソッドを呼び出します。したがって、サンプルのデータサービスの `Validate` メソッドは、追加の引数を処理する別の `parse_args()` 関数を実装する必要があります。

サンプルのデータサービスの `Validate` メソッドは、単一のプロパティである `Confdir` 拡張プロパティを確認します。このプロパティは、DNS が正常に動作するために重要な DNS 構成ディレクトリを指します。

注 - DNS が動作している間、構成ディレクトリは変更できないため、`Confdir` プロパティは `RTR` ファイルで `TUNABLE = AT_CREATION` と宣言します。したがって、`Validate` メソッドが呼び出されるのは、更新の結果として `Confdir` プロパティを確認するためではなく、データサービスリソースが作成されているときだけです。

RGM が `Validate` メソッドに渡すプロパティの中に `Confdir` が存在する場合、`parse_args()` 関数はその値を取得および保存します。次に、`Validate` メソッドは、`Confdir` の新しい値が指すディレクトリがアクセス可能かどうか、および、`named.conf` ファイルがそのディレクトリ内に存在し、データを持っているかどうかを確認します。

`parse_args()` 関数が、RGM から渡されたコマンド行引数から `Confdir` の値を取得できない場合でも、`Validate` メソッドは `Confdir` プロパティの妥当性を検査しようとします。まず、`Validate` メソッドは `scha_resource_get()` 関数を使用し、静的な構成から `Confdir` の値を取得します。次に、同じ検査を実行し、構成ディレクトリがアクセス可能かどうか、および、空でない `named.conf` ファイルがそのディレクトリ内に存在するかどうかを確認します。

`Validate` メソッドが失敗で終了した場合、`Confdir` だけでなく、すべてのプロパティの更新または作成が失敗します。

Validate メソッドの構文解析関数

RGM は、ほかのコールバックメソッドとは異なるパラメータを Validate メソッドに渡します。したがって、Validate メソッドには、ほかのメソッドとは異なる引数を構文解析する関数が必要です。Validate メソッドやその他のコールバックメソッドに渡される引数の詳細については、`rt_callbacks(1HA)` のマニュアルページを参照してください。次に、Validate メソッドの `parse_args()` 関数を示します。

```
#####
# Validate 引数の構文解析。
#
function parse_args # [args...]
{
    typeset opt
    while getopts 'cur:x:g:R:T:G:' opt
    do
        case "$opt" in
            R)
                # DNS リソース名。
                RESOURCE_NAME=$OPTARG
                ;;
            G)
                # リソースが構成されたリソースグループ名。
                RESOURCEGROUP_NAME=$OPTARG
                ;;
            T)
                # リソースタイプ名。
                RESOURCETYPE_NAME=$OPTARG
                ;;
            r)
                # メソッドはシステム定義プロパティ
                # にアクセスしていない。したがって、
                # このフラグは動作なし。
                ;;
            g)
                # メソッドはリソースグループプロパティに
                # アクセスしていない。したがって、
                # このフラグは動作なし。
                ;;
            c)
                # Validate メソッドがリソースの作成中に
                # 呼び出されていることを示す。したがって、
                # このフラグは動作なし。
                ;;
            u)
                # リソースがすでに存在しているときは、
                # プロパティの更新を示す。Confdir
                # プロパティを更新する場合、Confdir
                # がコマンド行引数に現れる。現れない場合、メソッドは
                # scha_resource_get を使用して
                # Confdir を探す必要がある。
                UPDATE_PROPERTY=1
        esac
    done
}
```

```

;;
x)
# 拡張プロパティのリスト。プロパティ
# と値のペア。区切り文字は「=」
PROPERTY='echo $OPTARG | awk -F= '{print $1}''
VAL='echo $OPTARG | awk -F= '{print $2}''
# Confdir 拡張プロパティがコマンド行
# 上に存在する場合、その値を記録する。
if [ $PROPERTY == "Confdir" ]; then
    CONFDIR=$VAL
    CONFDIR_FOUND=1
fi
;;
*)
logger -p ${SYSLOG_FACILITY}.err \
-t [SYSLOG_TAG] \
"ERROR: Option $OPTARG unknown"
exit 1
;;
esac
done
}

```

ほかのメソッドの `parse_args()` 関数と同様に、この関数は、リソース名を取得するためのフラグ (R)、リソースグループ名を取得するためのフラグ (G)、RGM から渡されるリソースタイプを取得するためのフラグ (T) を提供します。

このメソッドはリソースが更新されるときに拡張プロパティの妥当性を検査するために呼び出されるため、`r` フラグ (システム定義プロパティを示す)、`g` フラグ (リソースグループプロパティを示す)、`c` フラグ (リソースの作成中に妥当性の検査が行われていることを示す) は無視されます。

`u` フラグは、`UPDATE_PROPERTY` シェル変数の値を 1 (TRUE) に設定します。`x` フラグは、更新されているプロパティの名前と値を取得します。更新されているプロパティの中に `Confdir` が存在する場合、その値が `CONFDIR` シェル変数に格納され、`CONFDIR_FOUND` 変数が 1 (TRUE) に設定されます。

Confdir の妥当性検査

`Validate` メソッドはまず、その `MAIN` 関数において、`CONFDIR` 変数を空の文字列に設定し、`UPDATE_PROPERTY` と `CONFDIR_FOUND` を 0 に設定します。

```

CONFDIR=""
UPDATE_PROPERTY=0
CONFDIR_FOUND=0

```

次に、`Validate` メソッドは `parse_args()` 関数を呼び出し、RGM から渡された引数を構文解析します。

```

parse_args "$@"

```


次に、Validate は、Validate メソッドがプロパティの更新の結果として呼び出されているかどうか、および、Confdir 拡張プロパティがコマンド行上に存在するかどうかを検査します。次に、Validate メソッドは、Confdir プロパティが値を持っているかどうかを確認します。値を持っていない場合、Validate メソッドはエラーメッセージを記録し、失敗状態で終了します。

```
if ( ( ( $UPDATE_PROPERTY == 1 ) ) && ( ( CONFDIR_FOUND == 0 ) ) ); then
    config_info='scha_resource_get -O Extension -R $RESOURCE_NAME \
    -G $RESOURCEGROUP_NAME Confdir'
    CONFDIR='echo $config_info | awk '{print $2}'`
fi

# Confdir プロパティが値を持っているかどうかを確認する。持っていな
# い場合、状態 1 (失敗) で終了する。
if [ ! -z $CONFDIR ]; then
    logger -p ${SYSLOG_FACILITY}.err \
        "${ARGV0} Validate method for resource "$RESOURCE_NAME " failed"
    exit 1
fi
```

注 - 上記コードにおいて、Validate メソッドが更新の結果として呼び出されているのか (\$UPDATE_PROPERTY == 1)、および、プロパティがコマンド行上に存在しないのか (CONFDIR_FOUND == 0) を検査し、両者が TRUE である場合に、scha_resource_get () 関数を使用して Confdir の既存の値を取得するところに注目してください。Confdir がコマンド行上に存在する (CONFDIR_FOUND == 1) 場合、CONFDIR の値は、scha_resource_get () 関数からではなく、parse_args () 関数から取得されます。

次に、Validate メソッドは CONFDIR の値を使用し、ディレクトリがアクセス可能であるかどうかを確認します。アクセス可能ではない場合、Validate メソッドはエラーメッセージを記録し、エラー状態で終了します。

```
# $CONFDIR がアクセス可能であるかどうかを検査する。
if [ ! -d $CONFDIR ]; then
    logger -p ${SYSLOG_FACILITY}.err \
        -t [${SYSLOG_TAG}] \
        "${ARGV0} Directory $CONFDIR missing or not mounted"
    exit 1
fi
```

Confdir プロパティの更新の妥当性を検査する前に、Validate メソッドは最終検査を実行し、named.conf ファイルが存在するかどうかを確認します。存在しない場合、Validate メソッドはエラーメッセージを記録し、エラー状態で終了します。

```
# named.conf ファイルが Confdir ディレクトリ内
# に存在するかどうかを検査する。
if [ ! -s $CONFDIR/named.conf ]; then
    logger -p ${SYSLOG_FACILITY}.err \
        -t [${SYSLOG_TAG}] \
        "${ARGV0} File $CONFDIR/named.conf is missing or empty"
    exit 1
fi
```

最終検査を通過した場合、Validate メソッドは、成功を示すメッセージを記録し、成功状態で終了します。

```
# Validate メソッドが成功したことを示すメッセージを記録する。
logger -p ${SYSLOG_FACILITY}.err \
  -t [SYSLOG_TAG] \
  "${ARGV0} Validate method for resource "$RESOURCE_NAME \
  " completed successfully"

exit 0
```

Validate の終了状態

Validate メソッドが成功 (0) で終了すると、新しい値を持つ Confdir が作成されます。Validate メソッドが失敗 (1) で終了すると、Confdir を含むすべてのプロパティが作成されず、理由を示すメッセージがクラスタ管理者に送信されます。

Update メソッド

リソースのプロパティが変更されたとき、RGM は Update メソッドを呼び出し、動作中のリソースにその旨を通知します。RGM は、管理アクションがリソースまたはそのリソースグループのプロパティの設定に成功したあとに、Update を呼び出します。このメソッドは、リソースがオンラインであるノード上で呼び出されます。

Update の概要

Update メソッドはプロパティを更新しません。プロパティの更新は RGM が行います。その代わりに、動作中のプロセスに更新が発生したことを通知します。サンプルのデータサービスでは、プロパティの更新によって影響を受けるプロセスは障害モニターだけです。したがって、Update メソッドは、障害モニターを停止および再起動します。

Update メソッドは、障害モニターが動作していることを確認してから、pmfadm で障害モニターを強制終了する必要があります。UPDATE メソッドは、障害モニターを実装する検証プログラムの位置を取得し、その後、もう一度 pmfadm で障害モニターを再起動します。

Update による障害モニターの停止

Update メソッドは、pmfadm -q を使用し、障害モニターが動作していることを確認します。動作している場合、pmfadm -s TERM で障害モニターを強制終了します。障害モニターが正常に終了した場合、その影響を示すメッセージが管理ユーザーに送信されます。障害モニターが停止できない場合、Update メソッドは、エラーメッセージを管理ユーザーに送信し、失敗状態で終了します。

```

if pmfadm -q $RESOURCE_NAME.monitor; then

# すでに動作している障害モニターを強制終了する。
pmfadm -s $PMF_TAG TERM
    if [ $? -ne 0 ]; then
        logger -p ${SYSLOG_FACILITY}.err \
            -t [${SYSLOG_TAG}] \
                "${ARGV0} Could not stop the monitor"
        exit 1
    else
# DNS の停止に成功。メッセージを記録する。
        logger -p ${SYSLOG_FACILITY}.err \
            -t [${RESOURCE_TYPE_NAME}, ${RESOURCE_GROUP_NAME}, ${RESOURCE_NAME}] \
                "Monitor for HA-DNS successfully stopped"
    fi

```

障害モニターの再起動

障害モニターを再起動するために、Update メソッドは検証プログラムを実装するスクリプトの位置を見つける必要があります。検証プログラムはデータサービスのベースディレクトリ (RT_basedir プロパティが指すディレクトリ) 内にあります。Update は、次に示すように、RT_basedir の値を取得し、RT_BASEDIR 変数に格納します。

```

RT_BASEDIR=`scha_resource_get -O RT_BASEDIR -R $RESOURCE_NAME -G \
$RESOURCEGROUP_NAME`

```

次に、Update は、RT_BASEDIR の値を pmfadm で使用し、dns_probe プログラムを再起動します。検証プログラムを再起動できた場合、Update メソッドはその影響を示すメッセージを管理ユーザーに送信し、成功状態で終了します。pmfadm が検証プログラムを再起動できない場合、Update メソッドはエラーメッセージを記録し、失敗状態で終了します。

Update の終了状態

Update メソッドが失敗すると、リソースが“update failed” (更新失敗) の状態になります。この状態は RGM のリソース管理に影響しません。しかし、syslog 機能を通じて、管理ツールへの更新アクションが失敗したことを示します。

第 6 章

データサービス開発ライブラリ (DSDL)

この章では、データサービス開発ライブラリ (Data Service Development Library: DSDL) を構成するアプリケーションプログラミングインタフェースの概要について説明します。DSDL は `libdsdev.so` ライブラリとして実装されており、Sun Cluster パッケージに含まれています。

この章の内容は次のとおりです。

- 117 ページの「DSDL の概要」
- 118 ページの「構成プロパティの管理」
- 118 ページの「データサービスの起動と停止」
- 119 ページの「障害モニターの実装」
- 120 ページの「ネットワークアドレス情報へのアクセス」
- 120 ページの「実装したリソースタイプのデバッグ」

DSDL の概要

DSDL API は、RMAPI の最上位の階層を形成します。DSDL API は RMAPI の代わりになるものではなく、RMAPI をカプセル化および拡張するためのものであることに注意してください。DSDL は、特定の Sun Cluster 統合問題に対する事前定義されたソリューションを提供することによって、データサービスの開発を簡素化します。その結果、アプリケーションに本来求められている高可用性とスケーラビリティの実現に、より多くの開発時間を割くことが可能になります。また、アプリケーションの起動、シャットダウン、および監視機能を Sun Cluster に統合する際に、多くの時間を費やすこともありません。

構成プロパティの管理

すべてのコールバックメソッドは構成プロパティにアクセスする必要があります。DSDL は、以下により、プロパティへのアクセスを容易にします。

- 環境の初期化
- プロパティ値を簡単に取得できる関数セットの提供

`scds_initialize` 関数 (各コールバックメソッドの開始時に呼び出す必要がある) は、次の処理を行います。

- RGM がコールバックメソッドに渡すコマンド行引数 (`argc` と `argv[]`) を検査および処理します。そのため、コマンド行解析関数を作成する必要はありません。
- 他の DSDL 関数が使用できるように内部データ構造を設定します。たとえば、DSDL で提供されている関数によって RGM から取得されたプロパティ値はこのデータ構造に格納されます。同様に、コマンド行から入力された値 (RGM から取得された値よりも優先される) もこのデータ構造に格納されます。

注 - `Validate` メソッドの場合、`scds_initialize` はコマンド行で渡されたプロパティ値を解析します。そのため、`Validate` 用の解析関数を作成する必要はありません。

また、`scds_initialize` 関数はロギング環境を初期化して、障害モニターの検証設定の妥当性を検査します。

DSDL は、リソース、リソースタイプ、リソースグループのプロパティ、および、よく使用される拡張プロパティを取得するための関数セットを提供します。これらの関数は、次のような規則に従って、プロパティへのアクセスを標準化しています。

- 各関数は、`scds_initialize` から戻されるハンドル引数だけを取ります。
- 各関数が特定のプロパティに対応します。関数の戻り値のタイプは取得するプロパティ値のタイプに一致します。
- 値は `scds_initialize` によってあらかじめ算出されているため、関数はエラーを戻しません。新しい値がコマンド行で渡された場合を除き、関数は RGM から値を取得します。

データサービスの起動と停止

`Start` メソッドは、クラスタノード上でデータサービスを起動するために必要なアクションを実行します。通常、このようなアクションには、リソースプロパティの取得、アプリケーション固有の実行可能ファイルおよび構成ファイルの格納先の特定、および適切なコマンド行引数を用いたアプリケーションの起動が含まれます。

`scds_initialize` 関数はリソース構成を取得します。Start メソッドはプロパティ用の DSDL 関数を使用して、アプリケーションを起動するのに必要な構成ディレクトリや構成ファイルを識別するための特定のプロパティ (`Confdir_list` など) の値を取得します。

Start メソッドは、`scds_pmf_start` を呼び出して、プロセス監視機能 (PMF) の制御下でアプリケーションを起動します。PMF を使用すると、プロセスに適用する監視レベルを指定したり、異常終了したプロセスを再起動したりできます。DSDL で実装する Start メソッドの例については、138 ページの「`xfnts_start` メソッド」を参照してください。

Stop メソッドは呼び出し回数に依存しないように実装されていなければなりません。つまり、アプリケーションが動作していないときにノード上で呼び出された場合でも、正常終了する必要があります。Stop メソッドが失敗した場合、停止するリソースが `STOP_FAILED` 状態に設定され、クラスタの再起動を招いてしまう可能性があります。

リソースが `STOP_FAILED` 状態になるのを防止するために、Stop メソッドはあらゆる手段を構じてリソースを停止する必要があります。`scds_pmf_stop` 関数は、段階的にリソースを停止しようとします。まず、`SIGTERM` シグナルを使用してリソースを停止しようとします。これに失敗した場合は、`SIGKILL` シグナルを使用します。詳細は、`scds_pmf_stop` (3HA) のマニュアルページを参照してください。

障害モニターの実装

DSDL は、事前に定義されたモデルを提供することによって、障害モニターを実装する際の煩雑さをほとんど取り除きます。リソースがノード上で起動すると、`Monitor_start` メソッドは PMF の制御下で障害モニターを起動します。リソースがノード上で動作している間、障害モニターは無限ループを実行します。次に、DSDL 障害モニターのロジックの概要を示します。

- `scds_fm_sleep` 関数は `Thorough_probe_interval` プロパティを使用して、検証を行う期間を決定します。この期間中に PMF がアプリケーションプロセスの失敗を決定した場合、リソースは再起動されます。
- 検証機能自身は、障害の重要度を示す値を戻します。この値の範囲は、0 (障害なし) から 100 (致命的な障害) までです。
- 検証機能が戻した値は、`scds_action` 関数に送信されます。`scds_action` 関数は、`Retry_interval` プロパティの期間中に、障害の履歴を累積します。
- `scds_action` 関数は、次に示すような、障害が発生した場合の処置を決定します。
 - 累積した障害が 100 より少ない場合は、何もしません。
 - 累積した障害が 100 に到達した場合 (完全な障害)、データサービスを再起動します。`Retry_interval` を超えた場合、障害の履歴をリセットします。

- `Retry_interval` で指定された期間中に、再起動の回数が `Retry_count` プロパティを上回った場合、データサービスをフェイルオーバーします。

ネットワークアドレス情報へのアクセス

DSDL は、リソースおよびリソースグループのネットワークアドレス情報を戻す関数を提供します。たとえば、`scds_get_netaddr_list` は、リソースが使用するネットワークアドレスリソースを取得して、障害モニターがアプリケーションを検証できるようにします。

また、DSDL は TCP ベースの監視を行う関数セットも提供します。通常、このような関数はサービスとの間に単純なソケット接続を確立し、サービスのデータを読み書きした後で、サービスとの接続を切断します。検証の結果を DSDL の `scds_fm_action` 関数に送信し、次に実行すべき処理を決定できます。

TCP ベースの障害監視の例については、152 ページの「`xfnts_validate` メソッド」を参照してください。

実装したリソースタイプのデバッグ

DSDL は、データサービスをデバッグするときに役立つ組み込み機能を提供します。

DSDL の `scds_syslog_debug()` ユーティリティは、実装したリソースタイプにデバッグ文を追加するための基本的なフレームワークを提供します。デバッグレベル (1 から 9 までの数字) は、各クラスタノード上のリソースタイプごとに動的に設定できます。ファイル `/var/cluster/rgm/rt/rtname/loglevel` は、1 から 9 までの整数だけが含まれているファイルであり、すべてのリソースタイプコールバックメソッドはこのファイルを読み取ります。DSDL の `scds_initialize()` ルーチンはこのファイルを読み取って、内部デバッグレベルを指定されたレベルに設定します。デフォルトのデバッグレベルは 0 であり、この場合、データサービスはデバッグメッセージを記録しません。

`scds_syslog_debug()` 関数は、`LOG_DEBUG` の優先順位において、`scha_cluster_getlogfacility()` 関数から戻された機能を使用します。このようなデバッグメッセージは `/etc/syslog.conf()` で構成できます。

`scds_syslog` ユーティリティを使用すると、いくつかのデバッグメッセージをリソースタイプの通常の動作 (おそらくは `LOG_INFO` 優先順位) における情報メッセージとして使用することができます。第 8 章のサンプル DSDL アプリケーションでは、`scds_syslog_debug` と `scds_syslog` 関数が多用されています。

高可用性ローカルファイルシステムの有効化

HAStoragePlus リソースタイプを使用すると、ローカルファイルシステムを Sun Cluster 環境内で高可用性にすることができます。このためには、ローカルファイルシステムのパーティションを広域ディスクグループ内に配置しなければなりません。また、アフィニティスイッチオーバーを有効にし、Sun Cluster 環境をフェイルオーバー用に構成する必要もあります。これによって、多重ホストディスクに直接接続された任意のホストから、多重ホストディスク上の任意のファイルシステムにアクセスできるようになります。入出力が多いデータサービスでは、高可用性のローカルファイルシステムを使用することを強く推奨します。HAStoragePlus リソースタイプの構成については、『Sun Cluster データサービスの計画と管理 (Solaris OS 版)』の「HA ローカルファイルシステムの有効化」を参照してください。

第 7 章

リソースタイプ的设计。

この章では、リソースタイプの设计や実装で DSDL を通常どのように使用するかについて説明します。また、リソース構成を検証したり、リソースの开始、停止、および監視を行なったりするためのリソースタイプの设计についても説明します。そして、最後に、リソースタイプのコールバックメソッドを DSDL を使って導入する方法を説明します。

詳細は、`rt_callbacks(1HA)` のマニュアルページを参照してください。

これらの作業を行うには、リソースのプロパティ設定値にアクセスできなければなりません。DSDL ユーティリティー `scds_initialize()` を使用すると、統一された方法でリソースプロパティにアクセスできます。この機能は、各コールバックメソッドの始めの部分で呼び出す必要があります。このユーティリティー関数は、クラスタフレームワークからリソースのすべてのプロパティを取り出します。これによって、これらのプロパティは、`scds_getname()` 関数群からアクセスできるようになります。

この章の内容は次のとおりです。

- 124 ページの「RTR ファイル」
- 124 ページの「Validate メソッド」
- 126 ページの「Start メソッド」
- 127 ページの「Stop メソッド」
- 128 ページの「Monitor_start メソッド」
- 129 ページの「Monitor_stop メソッド」
- 129 ページの「Monitor_check メソッド」
- 129 ページの「Update メソッド」
- 130 ページの「Init、Fini、Boot の各メソッド」
- 131 ページの「障害モニターデーモンの设计」

RTR ファイル

RTR (Resource Type Registration、リソースタイプ登録) ファイルは、リソースタイプの重要コンポーネントです。Sun Cluster は、リソースタイプの詳細な情報をこのファイルから取得します。この詳細情報には、この実装に必要なプロパティや、それらのデータタイプやデフォルト値、リソースタイプの実装に必要なコールバックメソッドのファイルシステムパス、システム定義プロパティのさまざまな設定値などがあります。

ほとんどのリソースタイプ実装では、DSDL に添付されるサンプル RTR ファイルだけで十分なはずですが。必要な作業は、基本的な要素 (リソースタイプ名、リソースタイプのコールバックメソッドのパス名など) の編集だけです。リソースタイプを実装する際に新しいプロパティが必要な場合は、そのプロパティをリソースタイプ実装のリソースタイプ登録 (RTR) ファイルに拡張プロパティとして宣言します。新しいプロパティには、DSDL の `scds_get_ext_property()` ユーティリティを使ってアクセスできます。

Validate メソッド

リソースタイプ実装の `validate` メソッドは、次のどちらかの条件のときに RGM によって呼び出されます。

- このリソースタイプの新規リソースが作成されつつある。
- このリソースまたはリソースグループのプロパティが更新されつつある。

この2つの操作は、リソースの `validate` メソッドに渡されるコマンド行オプション `-c` (作成) と `-u` (更新) によって区別されます。

`validate` メソッドは、リソースタイププロパティ `INIT_NODES` の値で定義されるノード群の各ノードに対して呼び出されます。たとえば、`INIT_NODES` に `RG_PRIMARYES` が設定されている場合、`validate` は、そのリソースのリソースグループを収容できる (その主ノードになりうる) 各ノードに対して呼び出されます。`INIT_NODES` が `RT_INSTALLED_NODES` に設定されている場合、`validate` は、リソースタイプソフトウェアがインストールされている各ノード (通常は、クラスタのすべてのノード) に対して呼び出されず、`INIT_NODES` のデフォルト値は `RG_PRIMARYES` です (`rt_reg(4)` のマニュアルページを参照)。`validate` メソッドが呼び出される時点では、RGM はまだリソースを作成していません (作成コールバックの場合)。あるいは、更新するプロパティの更新値をまだ適用していません (更新コールバックの場合)。リソースタイプ実装の `validate` コールバックメソッドの目的は、リソースの新しい設定値 (リソースに対して指定された新しいプロパティ設定値) がそのリソースタイプにとって有効であるかどうかを検査することにあります。

注 - HASToragePlus によって管理されるローカルファイルシステムを使用している場合は、`scds_hasp_check` を使って HASToragePlus リソースの状態を検査します。この情報は、そのリソースが依存するすべての SUNW.HASToragePlus (5) リソースの状態 (オンラインかどうか) から、そのリソースに定義されたシステムプロパティ `Resource_dependencies` または `Resource_dependencies_weak` を使って取得されます。`scds_hasp_check` 呼び出しから返される状態コードの完全なリストについては、`scds_hasp_check(3HA)` のマニュアルページを参照してください。

DSDL 関数 `scds_initialize()` は、リソースの作成や更新をそれぞれ次のように処理します。

- リソースの作成では、コマンド行から渡された新しいリソースプロパティを解析します。これによって、リソースタイプの開発者は、リソースプロパティの新しい値を、そのリソースがすでにシステムに作成されているかのように使用できます。
- リソースやリソースグループの更新では、管理者によって更新されようとしているプロパティの新しい値をコマンド行から読み込み、残りのプロパティ (値が更新されないもの) をリソース管理 API を使って Sun Cluster から読み込みます。ただし、リソースタイプの開発者は、DSDL を使用する限り、このような初期作業を行う必要はありません。さらに、開発者は、リソースのすべてのプロパティが使用可能であるものとして、リソースの検証を行うことができます。

次の図に示す `svc_validate()` は、リソースプロパティの検証を行う関数です。この関数は、`scds_get_name()` 関数群を使って、検証しようとするプロパティを検査します。リソースの設定が有効ならこの関数から戻りコード 0 が返されるとすると、リソースタイプの `validate` メソッドは、次のコード部分のようになります。

```
int
main(int argc, char *argv[])
{
    scds_handle_t handle;
    int rc;

    if (scds_initialize(&handle, argc, argv) != SCHA_ERR_NOERR) {
        return (1); /* 初期エラー */
    }
    rc = svc_validate(handle);
    scds_close(&handle);
    return (rc);
}
```

さらに、検証関数は、リソースの設定が有効でない場合は、その理由を記録する必要があります。`svc_validate()` 関数の例 (詳細は省略) は、次のようになります (実際の検証ルーチンについては、次の章を参照)。

```
int
svc_validate(scds_handle_t handle)
{
    scha_str_array_t *confdirs;
    struct stat      statbuf;
```

```

confdirs = scds_get_confdir_list(handle);
if (stat(confdirs->str_array[0], &statbuf) == -1) {
return (1); /* 無効なリソースプロパティ設定 */
}
return (0); /* 有効な設定 */
}

```

このように、リソースタイプの開発者は、`svc_validate()` 関数を使用することだけに集中できます。リソースタイプ実装の典型的な例としては、`app.conf` というアプリケーション構成ファイルを `Confdir_list` プロパティの下に置く処理があります。この処理は、`Confdir_list` プロパティから取り出した適切なパス名に対して `stat()` システム呼び出しを実行することによって実装できます。

Start メソッド

リソースタイプ実装の `Start` コールバックメソッドは、特定のクラスタノードのリソースを開始するときに RGM によって呼び出されます。リソースグループ名とリソース名、およびリソースタイプ名はコマンド行から渡されます。`Start` メソッドは、クラスタノードでデータサービスリソースを開始するために必要なアクションを行います。通常、このようなアクションには、リソースプロパティの取得や、アプリケーション固有の実行可能ファイルと構成ファイル (または、どちらか) の場所の特定、適切なコマンド行引数を使用したアプリケーションの起動などがあります。

DSDL では、リソース構成ファイルが `scds_initialize()` ユーティリティによってすでに取得されています。アプリケーションの起動アクションは、`svc_start()` 関数に指定できます。さらに、アプリケーションが実際に起動されたかどうかを確認するために、`svc_wait()` 関数を呼び出すことができます。`Start` メソッドのコード (詳細は省略) は、次のようになります。

```

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

    if (scds_initialize(&handle, argc, argv) != SCHA_ERR_NOERR) {
return (1); /* 初期化エラー */
}
    if (svc_validate(handle) != 0) {
return (1); /* 無効な設定 */
}
    if (svc_start(handle) != 0) {
return (1); /* 起動に失敗 */
}
    return (svc_wait(handle));
}

```

この起動メソッドの実装では、`svc_validate()` を呼び出してリソース構成を検証します。検証結果が正しくない場合は、リソース構成とアプリケーション構成が一致していないか、このクラスタノードのシステムに関して何らかの問題があることを示

しています。たとえば、リソースに必要な広域ファイルシステムが現在このクラスタノードで使用できない可能性などが考えられます。その場合には、このクラスタノードでこのリソースを起動しても意味がないので、RGM を使って別のノードのリソースを起動すべきです。ただし、この場合、`svc_validate()` は十分に限定的であるものとします。その場合、このルーチンは、アプリケーションが必要とするリソースがあるかどうかをそのクラスタノードだけで検査します。そうでないと、このリソースはすべてのクラスタノードで起動に失敗し、`START_FAILED` の状態になる可能性があります。この状態については、`scswitch(1M)` のマニュアルページおよび『*Sun Cluster* データサービスの計画と管理 (Solaris OS 版)』を参照してください。

`svc_start()` 関数は、このノードでリソースの起動に成功した場合は戻りコード 0 を、問題を検出した場合は 0 以外の戻りコードをそれぞれ返す必要があります。この関数から 0 以外の値が返されると、RGM は、このリソースを別のクラスタノードで起動しようと試みます。

DSDL を最大限に活用するには、`svc_start()` 関数で `scds_pmf_start()` ユーティリティを使って、アプリケーションを PMF (プロセス管理機能) のもとで起動できます。このユーティリティは、PMF の障害コールバックアクション機能 (`pmfadm(1M)` の `-a` アクションフラグを参照) を使って、プロセス障害の検出を実装します。

Stop メソッド

リソースタイプ実装の Stop コールバックメソッドは、特定のクラスタノードでアプリケーションを停止するときに RGM によって呼び出されます。Stop メソッドのコールバックが有効であるためには、次の条件が必要です。

- Stop メソッドは結果に依存しない命令 (*idempotent*) でなければなりません。つまり、Stop メソッドは、そのノードで Start メソッドが正常に終了していても、RGM から呼び出されることがあります。したがって、Stop メソッドは、そのクラスタノードでアプリケーションが動作していない場合でも (したがって、特別な処理が必要ない場合でも)、正常に (終了コード 0 で) 終了しなければなりません。
- リソースタイプの Stop メソッドが特定のクラスタノードで失敗に終わると (戻りコードが 0 以外だと)、そのリソースタイプは `STOP_FAILED` の状態になります。この場合、リソースの `Failover_mode` 設定によっては、このクラスタノードが RGM によって強制的に再起動されることがあります。したがって、Stop メソッドの設計時には、アプリケーションを停止する手段をメソッドに組み込んでおくことが重要です。たとえば、アプリケーションが停止しない場合は、`SIGKILL` などを使って、アプリケーションを強制的かつ即時に停止する必要があります。さらに、この処理は一定の時間内に行われなければなりません。Stop_timeout で設定した時間が経過すると、停止が失敗したものとみなされ、リソースは `STOP_FAILED` の状態になるからです。

ほとんどのアプリケーションには、DSDL ユーティリティー `scds_pmf_stop()` で十分なはずです。このユーティリティーは、まず、アプリケーションが PMF の `scds_pmf_start()` で起動されたものとみなして、アプリケーションを SIGTERM で「静かに」停止しようとしています。これで停止しない場合は、プロセスに対して SIGKILL を適用します。このユーティリティーの詳細については、206 ページの「PMF 関数」を参照してください。

アプリケーションを停止するそのアプリケーション固有の関数を `svc_stop()` とし、これまで使用してきたコードモデルに従うとするなら、Stop メソッドは、次のように実装できます。`svc_stop()` の実装で `scds_pmf_stop()` が使用されているかどうかは、ここでは関係ありません。それが使用されているかどうかは、アプリケーションが PMF のもとで Start メソッドによって起動されているかどうかによって依存します。

```
if (scds_initialize(&handle, argc, argv) != SCHA_ERR_NOERR)
{
    return (1);    /* 初期化エラー */
}
return (svc_stop(handle));
```

Stop メソッドの実装では、`svc_validate()` メソッドは使用されません。システムに問題があったとしても、Stop メソッドは、このノードでこのアプリケーションを停止すべきだからです。

Monitor_start メソッド

RGM は、Monitor_start メソッドを呼び出して、リソースに対する障害モニターを起動します。障害モニターは、このリソースによって管理されているアプリケーションの状態を監視します。リソースタイプの実装では、通常、障害モニターはバックグラウンドで動作する独立したデーモンとして実装されます。このデーモンの起動には、適切な引数をもつ Monitor_start コールバックメソッドが使用されません。

モニターデーモン自体は障害が発生しやすいため (たとえば、モニターは、アプリケーションを、監視されない状態にしたまま停止することがある)、モニターデーモンは、PMF を使って起動すべきです。DSDL ユーティリティー `scds_pmf_start()` には、障害モニターを起動する機能が組み込まれています。このユーティリティーは、モニターデーモンプログラムの相対パス名 (リソースタイプコールバックメソッド実装の場所を表す `RT_basedir` との相対パス) を使用します。さらに、ユーティリティーは、DSDL によって管理される `Monitor_retry_interval` と `Monitor_retry_count` 拡張プロパティを使って、デーモンが際限なく再起動されるのを防止します。モニターデーモンのコマンド行構文には、コールバックメソッドに対して定義されたコマンド行構文と同じものが使用されます (`-R resource -G resource_group -T resource_type`)。ただし、モニターデーモンが RGM から直接呼び出されることはありません。このユーティリティーでは、モニターデーモン実装自体が `scds_initialize()` ユーティリティーを使って独自の環境を設定できます。したがって、主な作業は、モニターデーモン自体を設計することです。

Monitor_stop メソッド

RGM は、Monitor_stop メソッドを使って、Monitor_start メソッドで起動された障害モニターデーモンを停止します。このコールバックメソッドの失敗は、Stop メソッドの失敗とまったく同じように処理されます。したがって、Monitor_stop メソッドは、Stop メソッドと同じように強固なものでなければなりません。

障害モニターデーモンを `scds_pmf_start()` ユーティリティを使って起動したら、`scds_pmf_stop()` ユーティリティで停止する必要があります。

Monitor_check メソッド

クラスタノードが特定のリソースを支配できるかどうかを確認するために (つまり、そのリソースによって管理されるアプリケーションがそのノードで正常に動作するかどうかを確認するために)、そのリソースの Monitor_check コールバックメソッドがそのリソースのノードで呼び出されます。)。通常、この状況では、アプリケーションに必要なすべてのシステムリソースが本当にクラスタノードで使用可能かどうかを確認されます。124 ページの「Validate メソッド」で述べたように、開発者が使用する `svc_validate()` 関数では、少なくともこの確認が行われなければなりません。

リソースタイプ実装によって管理されているアプリケーションによっては、Monitor_check メソッドでその他の作業を行うことがあります。Monitor_check メソッドは、並行して実行中のその他のメソッドと競合しない方法で実装する必要があります。DSDL を使用する場合には、リソースプロパティに対するアプリケーション固有の検証を行なうために作成された `svc_validate()` 関数を Monitor_check メソッドで活用することをお勧めします。

Update メソッド

RGM は、リソースタイプ実装の Update メソッドを呼び出して、システム管理者が行なったすべての変更をアクティブリソースの構成に適用します。Update メソッドは、そのリソースがオンラインになっているすべてのノードに対して呼び出されます。

リソースの構成に対して行われた変更は、リソースタイプ実装にとって必ず有効なものです。RGM は、リソースタイプの Update メソッドを呼び出す前に Validate メソッドを呼び出すからです。Validate メソッドは、リソースやリソースグループの

プロパティが変更される前に呼び出されます。したがって、Validate メソッドは新しい変更を拒否できます。変更が適用されると、Update メソッドが呼び出され、新しい設定値がアクティブ (オンライン) リソースに通知されます。

リソースタイプの開発者は、どのプロパティを動的に変更できるようにするかを慎重に決定し、RTR ファイルでこれらのプロパティに TUNABLE = ANYTIME を設定する必要があります。通常、障害モニターデーモンによって使用される、リソースタイプ実装のプロパティは、すべて動的に変更できるように設定できます。ただし、Update メソッドの実装が少なくともモニターデーモンを再起動できなければなりません。

このようなプロパティの候補には次のものがあります。

- Thorough_Probe_Interval
- Retry_Count
- Retry_Interval
- Monitor_retry_count
- Monitor_retry_interval
- Probe_timeout

これらのプロパティは、障害モニターデーモンがサービスの状態をどのようにチェックするかや、このチェックをどのような頻度で行うか、エラーをどのような履歴間隔で追跡管理するか、あるいは、PMF がどのような再起動しきい値を設定するかに影響を及ぼします。DSDL には、これらのプロパティの更新を行なうための `scds_pmf_restart()` ユーティリティが備わっています。

リソースプロパティを動的に更新できなければならないがプロパティの変更によって動作中のアプリケーションに影響が及ぶ可能性があるという場合は、適切なアクションを行なうことによって、プロパティに対する変更が動作中のアプリケーションインスタンスに正しく適用されるようにしなければなりません。DSDL には、現在のところ、この問題をサポートする機能はありません。変更されたプロパティをコマンド行から Update に渡すことはできません (Validate に渡すことは可能)。

Init、Fini、Boot の各メソッド

これらのメソッドは、「1 度だけのアクション」を行うためのものです (リソース管理 API 仕様の定義を参照)。DSDL のサンプル実装には、これらのメソッドの使い方は示されていません。しかし、これらのメソッドを使用する必要がある場合には、DSDL のすべての機能をこれらのメソッドでも使用できます。通常、「1 度だけのアクション」を使用するリソースタイプ実装では、Init メソッドと Boot メソッドはまったく同じように機能します。Fini メソッドは、一般に、Init メソッドや Boot メソッドのアクションを「取り消す」ためのアクションに使用されます。

障害モニターデーモンの設計

DSDL を使用したリソースタイプ実装の障害モニターデーモンには、通常、次の役割があります。

- 管理されているアプリケーションの状態を定期的に監視します。モニターデーモンのこの役割はアプリケーションに大きく依存します。したがって、リソースタイプによって大幅に異なることがあります。DSDL には、TCP に基づく簡単なサービスの状態を検査するいくつかのユーティリティ関数が組み込まれています。HTTP、NNTP、IMAP、POP3 など、ASCII ベースのプロトコルを実装するアプリケーションは、これらのユーティリティを使って実装できます。
- アプリケーションによって検出された問題をリソースプロパティ `Retry_interval` や `Retry_count` を使って追跡します。さらに、アプリケーションが異常停止した場合には、PMF アクションスクリプトを使ってサービスを再起動すべきかどうかや、アプリケーションの障害が頻繁に発生するためにフェイルオーバーを考慮すべきかどうかを判断します。DSDL では、この機構の使用を助けるユーティリティとして `scds_fm_action()` と `scds_fm_sleep()` が提供されます。
- アプリケーションを再起動するか、リソースを含むリソースグループのフェイルオーバーを試みるなど、適切なアクションを実行します。DSDL ユーティリティ `scds_fm_action()` には、このようなアルゴリズムが使用されています。そのために、このアルゴリズムは、過去の `Retry_interval` で指定した秒数の間に起った検証障害の累積を計算します。
- リソースの状態を更新します。これによって、`scstat` コマンドやクラスタ管理 GUI からアプリケーションの状態を知ることができます。

DSDL ユーティリティの設計では、障害モニターデーモンの主要ループは次のようになっています。

DSDL を使って実装された障害モニターでは、

- アプリケーションプロセスの異常停止は、`scds_fm_sleep()` によって比較的迅速に検出されます。これは、PMF によるプロセス停止の通知が非同期に行われるためです。これは、障害モニターが時々ループから復帰してサービスの状態を検査し、アプリケーションの停止を検出する方法と対比的です。DSDL を使用した障害モニターでは障害検出時間が大幅に短縮されるため、サービスの可用性が向上します。
- RGM が `scha_control(3HA)` API によるサービスのフェイルオーバーを拒否すると、`scds_fm_action()` は、現在の障害履歴を「リセット」(消去)します。このようにするのは、障害履歴が `Retry_count` の値をすでに超えているからです。さらに、モニターデーモンは、次のサイクルでスリープから復帰した後に、デーモンの状態検査を正常に完了できないと、`scha_control()` を再び呼び出そうとするはずですが、しかし、前回のサイクルで呼び出しが拒否され状況が依然として残っていれば、この呼び出しは今回も拒否されるはずですが。履歴がリセットされていれば、障害モニターは、少なくとも、次のサイクルでアプリケーションの再起

動などによってその状況を内部的に訂正しようとしています。

- 再起動が失敗に終わった場合、`scds_fm_action()` は、アプリケーション障害履歴をリセットしません。これは、状況が訂正されなければ、`scha_control()` が間もなく呼び出される可能性が高いからです。
- ユーティリティー `scds_fm_action()` は、障害履歴に従って、`SCHA_RSSTATUS_OK`、`SCHA_RSSTATUS_DEGRADED`、`SCHA_RSSTATUS_FAULTED` のどれかをリソースステータスに設定します。これによって、ステータスをクラスタシステム管理から使用できるようになります。

ほとんどの場合、アプリケーション固有の状態検査アクションは、スタンドアロンの別個のユーティリティー (たとえば、`svc_probe()`) として実装してから、この汎用的なメインループに統合できます。

```
for (;;) {

    /* 正常な検証と検証の間の thorough_probe_interval
     * だけスリープする。 */
    (void) scds_fm_sleep(scds_handle,
        scds_get_rs_thorough_probe_interval(scds_handle));

    /* 使用するすべての ipaddress を検証する。
     * 次の各要素を繰り返し検証する。
     * 1. 使用するすべてのネットリソース
     * 2. 特定のリソースのすべての ipaddresses
     * 検証する ipaddress ごとに
     * 障害履歴を計算する。 */
    probe_result = 0;
    /* すべてのリソースを繰り返し調べて、
     * svc_probe() の呼び出しに使用する各 IP アドレスを取得する。 */
    for (ip = 0; ip < netaddr->num_netaddrs; ip++) {
        /* 状態を検証する必要があるホスト名とポート
         * を取得する。
         */
        hostname = netaddr->netaddrs[ip].hostname;
        port = netaddr->netaddrs[ip].port_proto.port;
        /*
         * HA-XFS は、1 つのポートしかサポートしないため
         * ポート配列の最初のエン트리から
         * ポート値を取得する。
         */
        ht1 = gethrtime(); /* Latch probe start time */
        probe_result = svc_probe(scds_handle,

            hostname, port, timeout);
        /*
         * サービス検証履歴を更新し、
         * 必要に応じてアクションを実行する。
         * 検証終了時刻を保存する。
         */
        ht2 = gethrtime();
        /* ミリ秒に変換する。 */
        dt = (ulong_t)((ht2 - ht1) / 1e6);
```

```
/*
 * 障害履歴を計算し、
 * 必要に応じてアクションを実行する。
 */
(void) scds_fm_action(scds_handle,
probe_result, (long)dt);
} /* 各ネットワークリソース */
} /* 検証を続ける */
```


第 8 章

サンプル DSDL リソースタイプの実装

この章では、DSDL で実装したサンプルのリソースタイプ `SUNW.xfnts` について説明します。データサービスは C 言語で作成されています。使用するアプリケーションは TCP/IP ベースのサービスである X Font Server です。

この章の内容は、次のとおりです。

- 135 ページの「X Font Server について」
- 137 ページの「`SUNW.xfnts` の RTR ファイル」
- 138 ページの「`scds_initialize()` 関数」
- 138 ページの「`xfnts_start` メソッド」
- 142 ページの「`xfnts_stop` メソッド」
- 144 ページの「`xfnts_monitor_start` メソッド」
- 145 ページの「`xfnts_monitor_stop` メソッド」
- 146 ページの「`xfnts_monitor_check` メソッド」
- 147 ページの「`SUNW.xfnts` 障害モニター」
- 152 ページの「`xfnts_validate` メソッド」

X Font Server について

X Font Server は、フォントファイルをクライアントに提供する、簡単な TCP/IP ベースのサービスです。クライアントはサーバーに接続してフォントセットを要求します。サーバーはフォントファイルをディスクから読み取って、クライアントにサービスを提供します。X Font Server デーモンはサーバーバイナリである `/usr/openwin/bin/xfns` から構成されます。このデーモンは通常、`inetd` から起動されますが、このサンプルでは、`/etc/inetd.conf` ファイル内の適切なエントリが (たとえば、`fsadmin -d` コマンドによって) 無効にされているものと想定します。したがって、デーモンは Sun Cluster だけの制御下にあります。

X Font Server の構成ファイル

デフォルトでは、X Font Server はその構成情報をファイル `/usr/openwin/lib/X11/fontserver.cfg` から読み取ります。このファイルのカタログエントリには、デーモンがサービスを提供できるフォントディレクトリのリストが入っています。クラスタ管理者は広域ファイルシステム上のフォントディレクトリの格納先を指定できます。こうすることによって、システム上でフォントデータベースのコピーを1つだけ保持すれば済むので、Sun Cluster 上の X Font Server の使用を最適化できます。広域ファイルシステム上のフォントディレクトリの格納先を指定するには、`fontserver.cfg` を編集して、フォントディレクトリの新しいパスを反映させる必要があります。

構成を簡単にするために、管理者は構成ファイル自身も広域ファイルシステム上に配置できます。`xfst` デーモンはデフォルトの格納先 (このファイルの組み込み場所) を変更するためのコマンド行引数を提供します。`SUNW.xfnts` リソースタイプは、次のコマンドを使用して、Sun Cluster の制御下でデーモンを起動します。

```
/usr/openwin/bin/xfst -config <location_of_cfg_file>/fontserver.cfg \  
-port <portnumber>
```

`SUNW.xfnts` v リソースタイプの実装では、`Confdir_list` プロパティを使用して、`fontserver.cfg` 構成ファイルの格納場所を管理できます。

TCP ポート番号

`xfst` サーバーデーモンが通信する TCP ポートの番号は、一般に「fs」ポート (通常、`/etc/services` ファイルで 7100 と定義されている) です。ただし、`xfst` コマンド行で `-port` オプションを指定することにより、システム管理者はデフォルトの設定を変更できます。`SUNW.xfnts` リソースタイプの `Port_list` プロパティを使用すると、デフォルト値を設定したり、`xfst` コマンド行で `-port` オプションを指定できるようになります。RTR ファイルにおいて、このプロパティのデフォルト値を `7100/tcp` と定義します。`SUNW.xfnts` の `Start` メソッドで、`Port_list` を `xfst` コマンド行の `-port` オプションに渡します。このようにすると、このリソースタイプのユーザーはポート番号を指定する必要がなくなります。つまり、デフォルトのポートが `7100/tcp` になります。ただし、リソースタイプを構成するときに、`Port_list` プロパティに異なる値を指定することにより、別のポートを指定することも可能です。

命名規約

次の命名規則を覚えておけば、サンプルコード内で関数とメソッドを簡単に識別できます。

- RMAPI 関数の名前は、`scha_` で始まります。

- DSDL 関数の名前は、scds_ で始まります。
- コールバックメソッドの名前は、xfnts_ で始まります。
- ユーザー定義関数の名前は、svc_ で始まります。

SUNW.xfnts の RTR ファイル

この節では、SUNW.xfnts の RTR ファイルで宣言されている、いくつかの重要なプロパティについて説明します。各プロパティの目的については説明しません。プロパティの詳細については、30 ページの「リソースとリソースタイププロパティの設定」を参照してください。

次に示すように、Confdir_list 拡張プロパティは構成ディレクトリ (または、ディレクトリのリスト) を指定します。

```
{  
    PROPERTY = Confdir_list;  
    EXTENSION;  
    STRINGARRAY;  
    TUNABLE = AT_CREATION;  
    DESCRIPTION = "The Configuration Directory Path(s)";  
}
```

Confdir_list プロパティには、デフォルト値は設定されていません。クラスタ管理者はリソースを作成するときに、構成 ディレクトリを指定する必要があります。

「TUNABLE = AT_CREATION」が指定されているので、作成時以降、この値を変更することはできません。

次に示すように、Port_list プロパティは、サーバーデーモンがリッスンするポートを指定します。

```
{  
    PROPERTY = Port_list;  
    DEFAULT = 7100/tcp;  
    TUNABLE = AT_CREATION;  
}
```

このプロパティにはデフォルト値が設定されているため、クラスタ管理者はリソースを作成するときに、新しい値を指定するか、デフォルト値を使用するかを選択します。「TUNABLE = AT_CREATION」が指定されているので、作成時以降、この値を変更することはできません。

scds_initialize() 関数

DSDL では、各コールバックメソッドがメソッドの開始時に `scds_initialize` (3HA) 関数を呼び出す必要があります。この関数は次の作業を行います。

- フレームワークがデータサービスメソッドに渡すコマンド行引数 (`argc` と `argv`) を検査および処理します。そのため、データサービスメソッドは、コマンド行引数について追加の処理を実行する必要はありません。
- 他の DSDL 関数が使用できるように内部データ構造を設定します。
- ロギング環境を初期化します。
- 障害モニターの検証設定の妥当性を検査します。

`scds_close()` 関数を使用すると、`scds_initialize()` が割り当てたリソースを再利用できます。

xfnts_start メソッド

データサービスリソースを含むリソースグループがオンラインになったとき、あるいは、リソースが有効になったとき、RGM はそのクラスタノード上で `start` メソッドを呼び出します。サンプルの `SUNW.xfnts` リソースタイプでは、`xfnts_start` メソッドが当該ノード上で `xfns` デーモンを起動します。

`xfnts_start` メソッドは `scds_pmf_start()` を呼び出して、PMF の制御下でデーモンを起動します。PMF は、自動障害通知、再起動機能、および障害モニターとの統合を提供します。

注 - `xfnts_start` は、`scds_initialize()` を最初に呼び出します。これによって、いくつかのハウスキーピング関数が実行されます。詳細については、[138 ページの「scds_initialize\(\) 関数」](#)と、`scds_initialize(3HA)` のマニュアルページを参照してください。

起動前のサービスの検証

次に示すように、`xfnts_start` メソッドは X Font Server を起動する前に `svc_validate()` を呼び出して、`xfns` デーモンをサポートするための適切な構成が存在していることを確認します。詳細については、[152 ページの「xfnts_validate メソッド」](#)を参照してください。

```

rc = svc_validate(scds_handle);
if (rc != 0) {
    scds_syslog(LOG_ERR,
        "Failed to validate configuration.");
    return (rc);
}

```

サービスの起動

xfnts_start メソッドは、xfnts.c で定義されている svc_start() メソッドを呼び出して、xfs デーモンを起動します。ここでは、svc_start() について説明します。

以下に、xfs デーモンを起動するためのコマンドを示します。

```
xfs -config config_directory/fontserver.cfg -port port_number
```

Confdir_list 拡張プロパティには config_directory を指定します。一方、Port_list システムプロパティには port_number を指定します。クラスタ管理者はデータサービスを構成するときに、これらのプロパティの値を指定します。

次に示すように、xfnts_start メソッドはこれらのプロパティを文字列配列として宣言し、scds_get_ext_confdir_list() と scds_get_port_list() 関数 (scds_property_functions(3HA)) で説明されている) を使用して、管理者が設定した値を取得します。

```

scha_str_array_t *confdirs;
scds_port_list_t  *portlist;
scha_err_t  err;

/* confdir_list プロパティから構成ディレクトリを取得する。*/
confdirs = scds_get_ext_confdir_list(scds_handle);

(void) sprintf(xfnts_conf, "%s/fontserver.cfg", confdirs->str_array[0]);

/* Port_list プロパティから XFS が使用するポートを取得する。*/
err = scds_get_port_list(scds_handle, &portlist);
if (err != SCHA_ERR_NOERR) {
    scds_syslog(LOG_ERR,
        "Could not access property Port_list.");
    return (1);
}

```

confdirs 変数は配列の最初の要素 (0) を指していることに注意してください。

次に示すように、xfnts_start メソッドは sprintf を使用して、xfs 用のコマンド行を形成します。

```

/* xfs デーモンを起動するコマンドを構築する。*/
(void) sprintf(cmd,
    "/usr/openwin/bin/xfs -config %s -port %d 2>/dev/null",

```

```
xfnts_conf, portlist->ports[0].port);
```

出力が dev/null にリダイレクトされていることに注意してください。こうすることによって、デーモンが生成するメッセージが抑制されます。

次に示すように、xfnts_start メソッドは xfs コマンド行を scds_pmf_start() に渡して、PMF の制御下でデータサービスを起動します。

```
scds_syslog(LOG_INFO, "Issuing a start request.");
err = scds_pmf_start(scds_handle, SCDS_PMF_TYPE_SVC,
                    SCDS_PMF_SINGLE_INSTANCE, cmd, -1);

if (err == SCHA_ERR_NOERR) {
    scds_syslog(LOG_INFO,
                "Start command completed successfully.");
} else {
    scds_syslog(LOG_ERR,
                "Failed to start HA-XFS ");
}
}
```

scds_pmf_start() を呼び出すときは、次のことに注意してください。

- SCDS_PMF_TYPE_SVC パラメータには、データサービスアプリケーションとして起動するプログラムを指定します。このメソッドは他のタイプのアプリケーション (障害モニターなど) も起動できます。
- SCDS_PMF_SINGLE_INSTANCE パラメータには、これが単一インスタンスのソースであることを指定します。
- cmd パラメータは、すでに生成されているコマンド行を示します。
- 最後のパラメータである -1 には、子プロセスの監視レベルを指定します。-1 は、PMF がすべての子プロセスを親プロセスと同様に監視することを示します。

次に示すように、svc_pmf_start() は portlist 構造体に割り当てられているメモリーを解放してから戻ります。

```
scds_free_port_list(portlist);
return (err);
```

svc_start() からの復帰

svc_start() が正常終了したときでも、使用するアプリケーションが起動に失敗した可能性があります。そのため、svc_start() はアプリケーションを検証して、アプリケーションが動作していることを確認してから、正常終了のメッセージを戻す必要があります。このとき、アプリケーションがただちに利用できない理由として、アプリケーションの起動にはある程度時間がかかるということを考慮しておく必要があります。次に示すように、svc_start() メソッドは xfnts.c で定義されている svc_wait() を呼び出して、アプリケーションが動作していることを確認します。

```
/* サービスが完全に起動するまで待つ。 */
scds_syslog_debug(DBG_LEVEL_HIGH,
```

```

        "Calling svc_wait to verify that service has started.");

rc = svc_wait(scds_handle);

scds_syslog_debug(DBG_LEVEL_HIGH,
    "Returned from svc_wait");

if (rc == 0) {
    scds_syslog(LOG_INFO, "Successfully started the service.");
} else {
    scds_syslog(LOG_ERR, "Failed to start the service.");
}

```

次に示すように、`svc_wait()` 関数は `scds_get_netaddr_list(3HA)` を呼び出して、アプリケーションを検証するのに必要なネットワークアドレスリソースを取得します。

```

/* 検証に使用するネットワークリソースを取得する。*/
if (scds_get_netaddr_list(scds_handle, &netaddr)) {
    scds_syslog(LOG_ERR,
        "No network address resources found in resource group.");
    return (1);
}

/* ネットワークリソースが存在しない場合は、エラーを戻す。*/
if (netaddr == NULL || netaddr->num_netaddrs == 0) {
    scds_syslog(LOG_ERR,
        "No network address resource in resource group.");
    return (1);
}

```

次に示すように、`svc_wait()` は `start_timeout` と `stop_timeout` 値を取得します。

```

svc_start_timeout = scds_get_rs_start_timeout(scds_handle)
probe_timeout = scds_get_ext_probe_timeout(scds_handle)

```

サーバーの起動に時間がかかることを考慮して、`svc_wait()` は `scds_svc_wait()` を呼び出して、`start_timeout` 値の 3% であるタイムアウト値を渡します。次に、`svc_wait()` は `svc_probe()` を呼び出して、アプリケーションが起動していることを確認します。`svc_probe()` メソッドは指定されたポート上でサーバーとの単純ソケット接続を確立します。ポートへの接続が失敗した場合、`svc_probe()` は値 100 を戻して、致命的な障害であることを示します。ポートとの接続は確立したが、切断に失敗した場合、`svc_probe()` は値 50 を戻します。

`svc_probe()` が完全にまたは部分的に失敗した場合、`svc_wait()` は `scds_svc_wait()` をタイムアウト値 5 で呼び出します。`scds_svc_wait()` メソッドは、検証の周期を 5 秒ごとに制限します。また、このメソッドはサービスを起動しようとした回数も数えます。この回数が `Retry_interval` プロパティで指定された期間内にリソースの `Retry_count` プロパティの値を超えた場合、`scds_svc_wait()` メソッドは失敗します。この場合、`svc_start()` 関数も失敗します。

```

#define    SVC_CONNECT_TIMEOUT_PCT    95
#define    SVC_WAIT_PCT                3
    if (scds_svc_wait(scds_handle, (svc_start_timeout * SVC_WAIT_PCT)/100)
        != SCHA_ERR_NOERR) {

        scds_syslog(LOG_ERR, "Service failed to start.");
        return (1);
    }

do {
    /*
     * ネットワークリソースの IP アドレスと portname 上で
     * データサービスを検証する。
     */
    rc = svc_probe(scds_handle,
        netaddr->netaddrs[0].hostname,
        netaddr->netaddrs[0].port_proto.port, probe_timeout);
    if (rc == SCHA_ERR_NOERR) {
        /* 成功。リソースを解放して終了。*/
        scds_free_netaddr_list(netaddr);
        return (0);
    }

    /* サービスが何度も失敗する場合は、scds_svc_wait() を呼び出す。
    if (scds_svc_wait(scds_handle, SVC_WAIT_TIME)
        != SCHA_ERR_NOERR) {
        scds_syslog(LOG_ERR, "Service failed to start.");
        return (1);
    }

/* RGM のタイムアウトを待つプログラムを終了する。*/
} while (1);

```

注 - xfnts_start メソッドは終了する前に scds_close() を呼び出して、scds_initialize() が割り当てたりソースを再利用します。詳細については、[138 ページの「scds_initialize\(\) 関数」](#)と [scds_close\(3HA\)](#) のマニュアルページを参照してください。

xfnts_stop メソッド

xfnts_start メソッドは scds_pmf_start() を使用して PMF の制御下でサービスを起動するため、xfnts_stop は scds_pmf_stop() を使用してサービスを中止します。

注 - `xfnts_stop` は `scds_initialize()` を最初に呼び出します。これによって、いくつかのハウスキーピング関数が実行されます。詳細については、`scds_initialize(3HA)` のマニュアルページを参照してください。

次に示すように、`xfnts_stop` メソッドは、`xfnts.c` で定義されている `svc_stop()` メソッドを呼び出します。

```
scds_syslog(LOG_ERR, "Issuing a stop request.");
err = scds_pmf_stop(scds_handle,
    SCDS_PMF_TYPE_SVC, SCDS_PMF_SINGLE_INSTANCE, SIGTERM,
    scds_get_rs_stop_timeout(scds_handle));

if (err != SCHA_ERR_NOERR) {
    scds_syslog(LOG_ERR,
        "Failed to stop HA-XFS.");
    return (1);
}

scds_syslog(LOG_INFO,
    "Successfully stopped HA-XFS.");
return (SCHA_ERR_NOERR); /* 正常に停止。 */
```

`svc_stop()` から `scds_pmf_stop()` 関数を呼び出すときは、次のことに注意してください。

- `SCDS_PMF_TYPE_SVC` パラメータには、データサービスアプリケーションとして停止するプログラムを指定します。このメソッドは他のタイプのアプリケーション(障害モニターなど)も停止できます。
- `SCDS_PMF_SINGLE_INSTANCE` パラメータには、シグナルを指定します。
- `SIGTERM` パラメータには、リソースインスタンスを停止するのに使用するシグナルを指定します。このシグナルでインスタンスを停止できなかった場合、`scds_pmf_stop()` は `SIGKILL` を送信してインスタンスを停止しようとします。このシグナルでもインスタンスを停止できなかった場合、タイムアウトエラーで戻ります。詳細については、`scds_pmf_stop(3HA)` のマニュアルページを参照してください。
- タイムアウト値は、リソースの `Stop_timeout` プロパティの値を示します。

注 - `xfnts_stop` メソッドは終了する前に `scds_close()` を呼び出して、`scds_initialize()` が割り当てたリソースを再利用します。詳細については、[138 ページの「`scds_initialize\(\)` 関数](#)」と `scds_close(3HA)` のマニュアルページを参照してください。

xfnts_monitor_start メソッド

リソースがノード上で起動したあと、RGMはそのノード上で Monitor_start メソッドを呼び出して障害モニターを起動します。xfnts_monitor_start メソッドは scds_pmf_start() を使用して PMF の制御下でモニターデーモンを起動します。

注 - xfnts_monitor_start は、scds_initialize() を最初に呼び出します。これによって、いくつかのハウスキーピング関数が実行されます。詳細については、138 ページの「scds_initialize() 関数」と、scds_initialize(3HA) のマニュアルページを参照してください。

次に示すように、xfnts_monitor_start メソッドは、xfnts.c に定義されている mon_start メソッドを呼び出します。

```
scds_syslog_debug(DBG_LEVEL_HIGH,
    "Calling Monitor_start method for resource <%s>.",
    scds_get_resource_name(scds_handle));

/* scds_pmf_start を呼び出し、検証の名前を渡す。 */
err = scds_pmf_start(scds_handle, SCDS_PMF_TYPE_MON,
    SCDS_PMF_SINGLE_INSTANCE, "xfnts_probe", 0);

if (err != SCHA_ERR_NOERR) {
    scds_syslog(LOG_ERR,
        "Failed to start fault monitor.");
    return (1);
}

scds_syslog(LOG_INFO,
    "Started the fault monitor.");

return (SCHA_ERR_NOERR); /* モニターを正常に起動。 */
}
```

svc_mon_start() から scds_pmf_start() 関数を呼び出すときは、次のことに注意してください。

- SCDS_PMF_TYPE_MON パラメータには、障害モニターとして起動するプログラムを指定します。このメソッドは他のタイプのアプリケーション (データサービスなど) も起動できます。
- SCDS_PMF_SINGLE_INSTANCE パラメータには、これが単一インスタンスのリソースであることを指定します。
- xfnts_probe パラメータには、起動するモニターデーモンを指定します。このモニターデーモンは、他のコールバックプログラムと同じディレクトリに存在するものと想定されます。

- 最後のパラメータである 0 は、子プロセスの監視レベルを指定します。この場合、モニターデーモンだけを監視することを示します。

注 - `xfnts_monitor_start` メソッドは終了する前に `scds_close()` を呼び出して、`scds_initialize()` が割り当てたリソースを再利用します。詳細については、138 ページの「`scds_initialize()` 関数」と `scds_close(3HA)` のマニュアルページを参照してください。

xfnts_monitor_stop メソッド

`xfnts_monitor_start` メソッドは `scds_pmf_start()` を使用して PMF の制御下でモニターデーモンを起動するので、`xfnts_monitor_stop` は `scds_pmf_stop()` を使用してモニターデーモンを停止します。

注 - `xfnts_monitor_stop` は、`scds_initialize()` を最初に呼び出します。これによって、いくつかのハウスキーピング関数が実行されます。詳細については、138 ページの「`scds_initialize()` 関数」と `scds_initialize(3HA)` のマニュアルページを参照してください。

次に示すように、`xfnts_monitor_stop()` メソッドは `xfnts.c` で定義されている `mon_stop` メソッドを呼び出します。

```
scds_syslog_debug(DBG_LEVEL_HIGH,
    "Calling scds_pmf_stop method");

err = scds_pmf_stop(scds_handle, SCDS_PMF_TYPE_MON,
    SCDS_PMF_SINGLE_INSTANCE, SIGKILL,
    scds_get_rs_monitor_stop_timeout(scds_handle));

if (err != SCHA_ERR_NOERR) {
    scds_syslog(LOG_ERR,
        "Failed to stop fault monitor.");
    return (1);
}

scds_syslog(LOG_INFO,
    "Stopped the fault monitor.");

return (SCHA_ERR_NOERR); /* モニターを正常に停止。 */
}
```

`svc_mon_stop()` から `scds_pmf_stop()` 関数を呼び出すときは、次のことに注意してください。

- SCDS_PMF_TYPE_MON パラメータには、障害モニターとして停止するプログラムを指定します。このメソッドは他のタイプのアプリケーション(データサービスなど)も停止できます。
- SCDS_PMF_SINGLE_INSTANCE パラメータには、これが単一インスタンスのリソースであることを指定します。
- SIGKILL パラメータには、リソースインスタンスを停止するのに使用するシグナルを指定します。このシグナルでインスタンスを停止できなかった場合、scds_pmf_stop() はタイムアウトエラーで戻ります。詳細については、scds_pmf_stop(3HA) のマニュアルページを参照してください。
- タイムアウト値は、リソースの Monitor_stop_timeout プロパティの値を示します。

注 - xfnts_monitor_stop メソッドは終了する前に scds_close() を呼び出して、scds_initialize() が割り当てたリソースを再利用します。詳細については、138 ページの「scds_initialize() 関数」と scds_close(3HA) のマニュアルページを参照してください。

xfnts_monitor_check メソッド

障害モニターがリソースが属するリソースグループを別のノードにフェイルオーバーしようとするたびに、RGM は Monitor_check メソッドを呼び出します。xfnts_monitor_check メソッドは svc_validate() メソッドを呼び出して xfs デーモンをサポートするための適切な構成が存在していることを確認します。詳細については、152 ページの「xfnts_validate メソッド」を参照してください。次に、xfnts_monitor_check のコードを示します。

```

/* RGM から渡された引数を処理し、syslog を初期化する。*/
if (scds_initialize(&scds_handle, argc, argv) != SCHA_ERR_NOERR)
{
    scds_syslog(LOG_ERR, "Failed to initialize the handle.");
    return (1);
}

rc = svc_validate(scds_handle);
scds_syslog_debug(DBG_LEVEL_HIGH,
    "monitor_check method "
    "was called and returned <%d>.", rc);

/* scds_initialize が割り当てたすべてのメモリーを解放する。*/
scds_close(&scds_handle);

/* モニター検査の一環として実行した検証メソッドの結果を戻す。*/
return (rc);

```

```
}
```

SUNW.xfnts 障害モニター

リソースがノード上で起動したあと、RGM は、PROBE メソッドを直接呼び出すのではなく、Monitor_start メソッドを呼び出してモニターを起動します。

xfnts_monitor_start メソッドは PMF の制御下で障害モニターを起動します。

xfnts_monitor_stop メソッドは障害モニターを停止します。

SUNW.xfnts 障害モニターは、次の処理を実行します。

- 単純な TCP ベースのサービス (xfs など) を検査するために特別に設計されたユーティリティを使用して、定期的に xfs サーバーデーモンの状態を監視します。
- (Retry_count と Retry_interval プロパティを使用して) ある期間内にアプリケーションが遭遇した問題を追跡し、アプリケーションが完全に失敗した場合に、データサービスを再起動するか、フェイルオーバーするかどうかを決定します。scds_fm_action() と scds_fm_sleep() 関数は、この追跡および決定機構の組み込みサポートを提供します。
- scds_fm_action() を使用して、フェイルオーバーまたは再起動の決定を実装します。
- リソースの状態を更新して、管理ツールや GUI で利用できるようにします。

xfnts_probe のメインループ

xfnts_probe メソッドは ループを実行します。ループを実行する前に、xfnts_probe は次の処理を行います。

- 次に示すように、xfnts リソース用のネットワークアドレスリソースを取得します。

```
/* 当該リソース用に利用できる IP アドレスを取得する。*/  
if (scds_get_netaddr_list(scds_handle, &netaddr)) {  
    scds_syslog(LOG_ERR,  
        "No network address resource in resource group.");  
    scds_close(&scds_handle);  
    return (1);  
}  
  
/* ネットワークリソースが存在しない場合、エラーを戻す。*/  
if (netaddr == NULL || netaddr->num_netaddrs == 0) {  
    scds_syslog(LOG_ERR,  
        "No network address resource in resource group.");  
    return (1);  
}
```

- `scds_fm_sleep()` を呼び出し、タイムアウト値として `Thorough_probe_interval` の値を渡します。検証を実行する間、検証機能は `Thorough_probe_interval` で指定された期間、休止状態になります。

```
timeout = scds_get_ext_probe_timeout(scds_handle);

for (;;) {
    /*
     * 連続する検証の間、thorough_probe_interval で指定された期間、
     * 休眠状態になる。
     */
    (void) scds_fm_sleep(scds_handle,
        scds_get_rs_thorough_probe_interval(scds_handle));
```

`xfnts_probe` メソッドは次のようなループを実装します。

```
for (ip = 0; ip < netaddr->num_netaddrs; ip++) {
    /*
     * 状態を監視するホスト名とポートを取得する。
     */
    hostname = netaddr->netaddrs[ip].hostname;
    port = netaddr->netaddrs[ip].port_proto.port;
    /*
     * HA-XFS がサポートするポートは 1 つだけなので、
     * ポート値はポートの配列の最初の
     * エントリから取得する。
     */
    ht1 = gethrtime(); /* 検証開始時間を取得する。 */
    scds_syslog(LOG_INFO, "Probing the service on port: %d.", port);

    probe_result =
        svc_probe(scds_handle, hostname, port, timeout);

    /*
     * サービス検証履歴を更新し、
     * 必要に応じて、アクションを行う。
     * 検証終了時間を取得する。
     */
    ht2 = gethrtime();

    /* ミリ秒に変換する。 */
    dt = (ulong_t)(ht2 - ht1) / 1e6);

    /*
     * 障害の履歴を計算し、必要に応じて
     * アクションを実行する。
     */
    (void) scds_fm_action(scds_handle,
        probe_result, (long)dt);
} /* ネットワークリソースごとに */
} /* 検証を永続的に繰り返す。 */
```

svc_probe() 関数は検証ロジックを実装します。svc_probe() からの戻り値は scds_fm_action() に渡されます。そして scds_fm_action() は、アプリケーションを再起動するか、リソースグループをフェイルオーバーするか、あるいは何もしないかを決定します。

svc_probe() 関数

svc_probe() 関数は、scds_fm_tcp_connect() を呼び出すことによって、指定されたポートとの単純ソケット接続を確立します。接続に失敗した場合、svc_probe() は 100 の値を戻して、致命的な障害であることを示します。接続には成功したが、切断に失敗した場合、svc_probe() は 50 の値を戻して、部分的な障害であることを示します。接続と切断の両方に成功した場合、svc_probe() は 0 の値を戻して、成功したことを示します。

次に、svc_probe() のコードを示します。

```
int svc_probe(scds_handle_t scds_handle,
char *hostname, int port, int timeout)
{
    int rc;
    hrtime_t t1, t2;
    int sock;
    char testcmd[2048];
    int time_used, time_remaining;
    time_t connect_timeout;

    /*
     * probe the data service by doing a socket connection to the port *
     * specified in the port_list property to the host that is
     * serving the XFS data service. If the XFS service which is configured
     * to listen on the specified port, replies to the connection, then
     * the probe is successful. Else we will wait for a time period set
     * in probe_timeout property before concluding that the probe failed.
     */

    /*
     * Use the SVC_CONNECT_TIMEOUT_PCT percentage of timeout
     * to connect to the port
     */
    connect_timeout = (SVC_CONNECT_TIMEOUT_PCT * timeout)/100;
    t1 = (hrtime_t)(gethrtime()/1E9);

    /*
     * the probe makes a connection to the specified hostname and port.
     * The connection is timed for 95% of the actual probe_timeout.
     */
    rc = scds_fm_tcp_connect(scds_handle, &sock, hostname, port,
        connect_timeout);
    if (rc) {
        scds_syslog(LOG_ERR,
```

```

        "Failed to connect to port <%d> of resource <%s>.",
        port, scds_get_resource_name(scds_handle));
/* this is a complete failure */
return (SCDS_PROBE_COMPLETE_FAILURE);
}

t2 = (hrtime_t)(gethrtime()/1E9);

/*
 * Compute the actual time it took to connect. This should be less than
 * or equal to connect_timeout, the time allocated to connect.
 * If the connect uses all the time that is allocated for it,
 * then the remaining value from the probe_timeout that is passed to
 * this function will be used as disconnect timeout. Otherwise, the
 * the remaining time from the connect call will also be added to
 * the disconnect timeout.
 *
 */

time_used = (int)(t2 - t1);

/*
 * Use the remaining time(timeout - time_took_to_connect) to disconnect
 */

time_remaining = timeout - (int)time_used;

/*
 * If all the time is used up, use a small hardcoded timeout
 * to still try to disconnect. This will avoid the fd leak.
 */
if (time_remaining <= 0) {
    scds_syslog_debug(DBG_LEVEL_LOW,
        "svc_probe used entire timeout of "
        "%d seconds during connect operation and exceeded the "
        "timeout by %d seconds. Attempting disconnect with timeout"
        " %d ",
        connect_timeout,
        abs(time_used),
        SVC_DISCONNECT_TIMEOUT_SECONDS);

    time_remaining = SVC_DISCONNECT_TIMEOUT_SECONDS;
}

/*
 * Return partial failure in case of disconnection failure.
 * Reason: The connect call is successful, which means
 * the application is alive. A disconnection failure
 * could happen due to a hung application or heavy load.
 * If it is the later case, don't declare the application
 * as dead by returning complete failure. Instead, declare
 * it as partial failure. If this situation persists, the
 * disconnect call will fail again and the application will be
 * restarted.
 */

```

```

rc = scds_fm_tcp_disconnect(scds_handle, sock, time_remaining);
if (rc != SCHA_ERR_NOERR) {
    scds_syslog(LOG_ERR,
        "Failed to disconnect to port %d of resource %s.",
        port, scds_get_resource_name(scds_handle));
    /* this is a partial failure */
    return (SCDS_PROBE_COMPLETE_FAILURE/2);
}

t2 = (hrtime_t)(gethrtime()/1E9);
time_used = (int)(t2 - t1);
time_remaining = timeout - time_used;

/*
 * If there is no time left, don't do the full test with
 * fsinfo. Return SCDS_PROBE_COMPLETE_FAILURE/2
 * instead. This will make sure that if this timeout
 * persists, server will be restarted.
 */
if (time_remaining <= 0) {
    scds_syslog(LOG_ERR, "Probe timed out.");
    return (SCDS_PROBE_COMPLETE_FAILURE/2);
}

/*
 * The connection and disconnection to port is successful,
 * Run the fsinfo command to perform a full check of
 * server health.
 * Redirect stdout, otherwise the output from fsinfo
 * ends up on the console.
 */
(void) sprintf(testcmd,
    "/usr/openwin/bin/fsinfo -server %s:%d > /dev/null",
    hostname, port);
scds_syslog_debug(DBG_LEVEL_HIGH,
    "Checking the server status with %s.", testcmd);
if (scds_timerun(scds_handle, testcmd, time_remaining,
    SIGKILL, &rc) != SCHA_ERR_NOERR || rc != 0) {

    scds_syslog(LOG_ERR,
        "Failed to check server status with command <%s>",
        testcmd);
    return (SCDS_PROBE_COMPLETE_FAILURE/2);
}
return (0);
}
}

```

svc_probe() は終了時に、成功 (0)、部分的な障害 (50)、または致命的な障害 (100) を戻します。xfnts_probe メソッドはこの値を scds_fm_action() に渡します。

障害モニターのアクションの決定

xfnts_probe メソッドは scds_fm_action() を呼び出して、行うべきアクションを決定します。scds_fm_action() のロジックは次のとおりです。

- `Retry_interval` プロパティで指定された期間中に、障害の履歴を累積します。
- 累積した障害が 100 に到達した場合 (致命的な障害)、データサービスを再起動します。`Retry_interval` を超えた場合、障害の履歴をリセットします。
- `Retry_interval` で指定された期間中に、再起動の回数が `Retry_count` プロパティを上回った場合、データサービスをフェイルオーバーします。

たとえば、検証機能が xfs サーバーに正常に接続したが、切断に失敗したものと想定します。これは、サーバーは動作しているが、ハングしていたり、一時的に過負荷状態になっている可能性を示しています。切断に失敗すると、`scds_fm_action()` に部分的な障害 (50) が送信されます。この値は、データサービスを再起動するしきい値を下回っていますが、値は障害の履歴に記録されます。

次の検証でもサーバーが切断に失敗した場合、`scds_fm_action()` が保持している障害の履歴に値 50 が再度追加されます。累積した障害の履歴が 100 になるので、`scds_fm_action()` はデータサービスを再起動します。

xfnts_validate メソッド

リソースが作成されたとき、および、リソースまたは (リソースを含む) リソースグループのプロパティが管理アクションによって更新される時、RGM は `Validate` メソッドを呼び出します。RGM は、作成または更新が行われる前に、`Validate` メソッドを呼び出します。任意のノード上でメソッドから失敗の終了コードが戻ると、作成または更新は取り消されます。

RGM が `Validate` メソッドを呼び出すのは、管理アクションがリソースまたはグループのプロパティを変更したときだけです。RGM がプロパティを設定したときや、モニターがリソースプロパティ `Status` と `Status_msg` を設定したときではありません。

注 - `PROBE` メソッドがデータサービスを新しいノードにフェイルオーバーする場合、`Monitor_check` メソッドも明示的に `Validate` メソッドを呼び出します。

RGM は、他のメソッドに渡す追加の引数 (更新されるプロパティと値を含む) を指定して、`Validate` メソッドを呼び出します。`xfnts_validate` の開始時に実行される `scds_initialize()` の呼び出しにより、RGM が `xfnts_validate` に渡したすべての引数が解析され、その情報が `scds_handle` パラメータに格納されます。この情報は、`xfnts_validate` が呼び出すサブルーチンによって使用されます。

`xfnts_validate` メソッドは `svc_validate()` を呼び出して、次のことを検証します。

- `Confdir_list` プロパティがリソース用に設定されており、単一のディレクトリが定義されているかどうか。


```

scha_str_array_t *confdirs;
confdirs = scds_get_ext_confdir_list(scds_handle);

/* confdir_list 拡張プロパティが存在しない場合、エラーを戻す。*/
if (confdirs == NULL || confdirs->array_cnt != 1) {
    scds_syslog(LOG_ERR,
        "Property Confdir_list is not set properly.");
    return (1); /* 検証に失敗 */
}

```

- Confdir_list で指定されたディレクトリに fontserver.cfg ファイルが存在しているかどうか。

```
(void) sprintf(xfnts_conf, "%s/fontserver.cfg", confdirs->str_array[0]);
```

```

if (stat(xfnts_conf, &statbuf) != 0) {
    /*
     * errno.h プロトタイプには void 引数がないので、
     * lint エラーが抑制される。
     */
    scds_syslog(LOG_ERR,
        "Failed to access file <%s> : <%s>",
        xfnts_conf, strerror(errno)); /*lint !e746 */
    return (1);
}

```

- サーバーデーモンバイナリがクラスタノード上でアクセスできるかどうか。

```

if (stat("/usr/openwin/bin/xfns", &statbuf) != 0) {
    scds_syslog(LOG_ERR,
        "Cannot access XFS binary : <%s> ", strerror(errno));
    return (1);
}

```

- Port_list プロパティが単一のポートを指定しているかどうか。

```

scha_port_list_t *portlist;
err = scds_get_port_list(scds_handle, &portlist);
if (err != SCHA_ERR_NOERR) {
    scds_syslog(LOG_ERR,
        "Could not access property Port_list: %s.",
        scds_error_string(err));
    return (1); /* 検証に失敗 */
}

#ifdef TEST
if (portlist->num_ports != 1) {
    scds_syslog(LOG_ERR,
        "Property Port_list must have only one value.");
    scds_free_port_list(portlist);
    return (1); /* 検証に失敗 */
}
#endif

```

- データサービスが属するリソースグループに、少なくとも1つのネットワークアドレスリソースが属しているかどうか。

```

scds_net_resource_list_t *snrlp;
if ((err = scds_get_rs_hostnames(scds_handle, &snrlp))
    != SCHA_ERR_NOERR) {
    scds_syslog(LOG_ERR,
        "No network address resource in resource group: %s.",
        scds_error_string(err));
    return (1); /* 検証に失敗 */
}

/* ネットワークアドレスリソースが存在しない場合エラーを戻す。*/
if (snrlp == NULL || snrlp->num_netresources == 0) {
    scds_syslog(LOG_ERR,
        "No network address resource in resource group.");
    rc = 1;
    goto finished;
}

```

次に示すように、`svc_validate()` は戻る前に、割り当てられているすべてのリソースを解放します。

```

finished:
    scds_free_net_list(snrlp);
    scds_free_port_list(portlist);

    return (rc); /* 検証結果を戻す。*/

```

注 - `xfnts_validate` メソッドは終了する前に `scds_close()` を呼び出して、`scds_initialize()` が割り当てたリソースを再利用します。詳細については、[138 ページの「scds_initialize\(\) 関数」](#)と `scds_close(3HA)` のマニュアルページを参照してください。

xfnts_update メソッド

プロパティが変更された場合、RGM は `Update` メソッドを呼び出して、そのことを動作中のリソースに通知します。`xfnts` データサービスにおいて変更可能なプロパティは、障害モニターに関連したもののだけです。したがって、プロパティが更新されたとき、`xfnts_update` メソッドは `scds_pmf_restart_fm()` を呼び出して、障害モニターを再起動します。

- * 障害モニターがすでに動作していることを検査し、動作している場合、
- * 障害モニターを停止および再起動する。`scds_pmf_restart_fm()` への
- * 2 番目のパラメータは、再起動する必要がある障害モニターの名

```

* インスタンスを一意に識別する。
*/

scds_syslog(LOG_INFO, "Restarting the fault monitor.");
result = scds_pmf_restart_fm(scds_handle, 0);
if (result != SCHA_ERR_NOERR) {
    scds_syslog(LOG_ERR,
        "Failed to restart fault monitor.");
    /* scds_initialize が割り当てたすべてのメモリーを解放する。*
    scds_close(&scds_handle);
    return (1);
}

scds_syslog(LOG_INFO,
    "Completed successfully.");

```

注 - `scds_pmf_restart_fm()` への 2 番目のパラメータは、複数のインスタンスが存在する場合に、再起動する障害モニターのインスタンスを一意に識別します。この例の場合、値 0 は障害モニターのインスタンスが 1 つしか存在しないことを示します。

第 9 章

SunPlex Agent Builder

この章では、Resource Group Manager (RGM) の管理下で動作するリソースタイプ (データサービス) の作成を自動化するツール、SunPlex Agent Builder と Agent Builder 用の Cluster Agent モジュールについて説明します。リソースタイプとは、アプリケーションが RGM の制御下にあるクラスタ環境で動作できるようにするアプリケーションのラッパーのことです。

次にこの章の内容を示します。

- 157 ページの「Agent Builder の概要」
- 158 ページの「Agent Builder の使用にあたって」
- 159 ページの「Agent Builder の使用」
- 175 ページの「ディレクトリ構造」
- 176 ページの「Agent Builder の出力」
- 180 ページの「Agent Builder の Cluster Agent モジュール」

Agent Builder の概要

Agent Builder は、アプリケーションや作成したいリソースタイプの種類に関する情報を入力するための画面ベースのインタフェースを提供します。

注 - グラフィカルユーザーインタフェース版の Agent Builder を使用できない場合は、コマンド行から Agent Builder を使用できます。詳細は、175 ページの「コマンド行バージョンの Agent Builder を使用する方法」を参照してください。

Agent Builder は、入力された情報に基づいて、次のソフトウェアを生成します。

- フェイルオーバーまたはスケーラブルリソースタイプ向けの C、Korn シェル (ksh) または汎用データサービス (GDS) ソールファイル群。これらのソースファイルは、ネットワーク対応 (クライアントサーバーモデル) アプリケーションの場合

も、非ネットワーク対応(クライアントなし)アプリケーションの場合も、リソースタイプの各メソッドコールバックに対応しています。

- C シェルまたは Korn シェルのソースコードを生成する場合は、カスタマイズされたリソースタイプ登録 (Resource Type Registration: RTR) ファイル。
- リソースタイプのインスタンス (リソース) を起動、停止、および削除するためのカスタマイズされたユーティリティスクリプト。また、これらのファイルの使用方法を説明するカスタマイズされたマニュアルページ。
- C のソースを生成する場合はバイナリを含む Solaris パッケージとユーティリティスクリプト。C または Korn シェルのソースコードを生成する場合は RTR ファイルを含む Solaris パッケージとユーティリティスクリプト。

Agent Builder は、クライアントとの通信にネットワークを使用するネットワーク対応アプリケーションと、非ネットワーク対応のスタンドアロンアプリケーションをサポートします。Agent Builder を使って、プロセス監視機能 (PMF) によって個別に監視および再起動される複数の独立したプロセスツリーを持つアプリケーション用のリソースタイプを生成できます。158 ページの「[複数の独立したプロセスツリーを持つリソースタイプの作成](#)」を参照してください。

Agent Builder の使用にあたって

以下の各節では、Agent Builder を使用する前に知っておく必要がある事項について説明します。

複数の独立したプロセスツリーを持つリソースタイプの作成

Agent Builder は、複数の独立したプロセスツリーを持つアプリケーション用のリソースタイプを作成できます。プロセスツリーが独立しているということは、PMF が各プロセスツリーを個別に監視および起動することを意味しています。PMF は独自のタグを使用して各プロセスツリーを起動します。

注 - Agent Builder を使って、複数の独立したプロセスツリーをもつリソースタイプを作成できますが、そのためには、生成されるソースコードとして C か GDS を指定する必要があります。Agent Builder を使って、このようなリソースタイプを Korn シェル用に作成することはできません。Korn シェル用にこれらのリソースタイプを作成するには、手動でそれらのコードを作成する必要があります。

複数の独立したプロセスツリーを持つベースアプリケーションの場合、1つのコマンド行だけでアプリケーションを起動することはできません。代わりに、アプリケーションの各プロセスツリーを起動するコマンドへの完全パスを行ごとに記述したテキストファイルを作成します。このファイルには空白行を含めることはできません。そして、このファイルへのパスを構成画面の「起動コマンド」テキストフィールドに指定します。

このファイルに実行権が設定されていないことを確認する必要があります。これにより、Agent Builder は、このファイルが複数のプロセスツリーを起動するためのものであり、単に複数のコマンドが記述されている実行可能スクリプトではないことを認識できます。このテキストファイルに実行権を与えた場合、リソースはクラスタ上で問題なく動作するよう見えますが、すべてのコマンドが1つの PMF タグ下で起動されるため、PMF は各プロセスツリーを個別に監視および再起動することはできません。

Agent Builder の使用

この節では、Agent Builder の使用方法と、Agent Builder を使用する前に実行する作業について説明します。リソースタイプコードを生成したあとで Agent Builder を活用する方法についても説明します。

この章では、以下の内容について説明します。

- 159 ページの「アプリケーションの分析」
- 160 ページの「Agent Builder のインストールと構成」
- 161 ページの「Agent Builder 画面」
- 161 ページの「Agent Builder の起動」
- 162 ページの「Agent Builder のナビゲーション」
- 166 ページの「作成画面の使用」
- 168 ページの「構成画面の使用」
- 171 ページの「Agent Builder の Korn シェルベース `$hostnames` 変数の使用」
- 171 ページの「プロパティ変数」
- 173 ページの「既存のリソースタイプからクローンを作成する方法」
- 174 ページの「生成されたソースコードの編集」
- 175 ページの「コマンド行バージョンの Agent Builder を使用する方法」

アプリケーションの分析

Agent Builder を使用する前に、アプリケーションが高可用性またはスケーラビリティを備えるための要件を満たしているかどうかを判定します。この分析はアプリケーションの実行時特性だけに基づくものなので、Agent Builder はこの分析を行うことができません。詳細については、25 ページの「アプリケーションの適合性の分析」を参照してください。

Agent Builder は必ずしもアプリケーション用の完全なリソースタイプを作成できるわけではありませんが、ほとんどの場合、少なくとも部分的なソリューションを提供します。たとえば、より複雑なアプリケーションでは、Agent Builder がデフォルトで生成しないコード (プロパティの妥当性検査を追加するコードや Agent Builder がエクスポートしないパラメータを調節するコード) を追加しなければならない場合があります。このような場合、生成されたコードまたは RTR ファイルを修正する必要があります。Agent Builder は、このような柔軟性を提供するように設計されています。

Agent Builder は、ソースファイル内において独自のリソースタイプコードを追加できる場所にコメント文を埋め込みます。ソースコードを修正した後、Agent Builder が生成した Makefile を使用すれば、ソースコードを再コンパイルし、リソースタイプパッケージを生成し直すことができます。

Agent Builder が生成したリソースタイプコードを使用せずに、リソースタイプコードを完全に作成し直す場合でも、Agent Builder が生成した Makefile やディレクトリ構造を使用すれば、独自のリソースタイプ用の Solaris パッケージを作成できます。

Agent Builder のインストールと構成

Agent Builder を個別にインストールする必要はありません。Agent Builder は、Sun Cluster ソフトウェアの標準インストールの際にデフォルトでインストールされる SUNWscdev パッケージに含まれています。詳細は『Sun Cluster ソフトウェアのインストール (Solaris OS 版)』を参照してください。

Agent Builder を使用する前に、次のことを確認してください。

- Java 実行時環境が \$PATH 変数に含まれているかどうか。Agent Builder は Java (Java Development Kit バージョン 1.3.1 以降) に依存しています。Java が \$PATH に含まれていないと、scdsbuilder からエラーメッセージが返され、表示されません。
- Solaris 8 以降の「Developer System Support」ソフトウェアグループがインストールされていること。
- cc コンパイラが \$PATH 変数に含まれているかどうか。Agent Builder は \$PATH 変数内で最初に現れる cc を使用して、リソースタイプの C バイナリコードを生成するコンパイラを識別します。cc が \$PATH に存在しない場合、Agent Builder は C コードを生成するオプションを無効にします。詳細は166 ページの「作成画面の使用」を参照してください。

注 – Agent Builder では、標準の `cc` コンパイラ以外のコンパイラも使用できます。このためには、`$PATH` において、`cc` から別のコンパイラ (`gcc` など) にシンボリックリンクを作成します。もう 1 つの方法は、Makefile におけるコンパイラ指定を変更して (現在は、`CC=cc`)、別のコンパイラへの完全パスを指定します。たとえば、エージェントが生成する Makefile において、`CC=cc` を `CC=pathname/gcc` に変更します。この場合、エージェントを直接実行することはできません。代わりに、`make` や `make pkg` コマンドを使用して、データサービスコードとパッケージを生成する必要があります。

Agent Builder 画面

Agent Builder は 2 つのステップ (画面) からなるウィザードです。Agent Builder では、次の 2 つの画面を使用して、新しいリソースタイプを作成します。

1. 「作成」この画面では、作成するリソースタイプの基本的な情報 (その名前や生成するファイルの作業ディレクトリなど) を指定します。作業ディレクトリは、リソースタイプテンプレートの作成や構成に使用する場所です。作成するリソースの種類 (スケーラブルまたはフェイルオーバー)、ベースアプリケーションがネットワーク対応かどうか (つまり、ネットワークを使用してクライアントと通信するかどうか)、生成するコードのタイプ (C、Korn シェル (ksh) または GDS) も指定できます。GDS の詳細は、第 10 章を参照してください。この画面の情報はすべて必須です。入力を終えたら「作成」を選択してその出力を生成します。その後「構成」画面を表示できます。
2. 「構成」この画面には完全なコマンド行を指定します。通常は、任意の UNIX[®] シェルに渡すことによって基本アプリケーションを起動できるものを指定します。オプションとして、アプリケーションを停止するコマンドや検証するコマンドも提供できます。これらのコマンドを指定しないと、生成される出力は、信号を送信してアプリケーションを停止し、デフォルトの検証メカニズムを提供します。検証コマンドの説明は、168 ページの「構成画面の使用」を参照してください。また、この画面では、上記の各コマンドのタイムアウト値も変更できます。

Agent Builder の起動

注 – グラフィカルユーザーインターフェイス版の Agent Builder を使用できない場合は、コマンド行から Agent Builder を使用できます。詳細は、175 ページの「コマンド行バージョンの Agent Builder を使用する方法」を参照してください。

注 - 既存のリソースタイプの作業ディレクトリから起動する場合、Agent Builder は、作成および構成画面を既存のリソースタイプの値に初期化します。

Agent Builder は次のコマンドで起動します。

```
% /usr/cluster/bin/scdsbuilder
```

「作成」画面が表示されます。

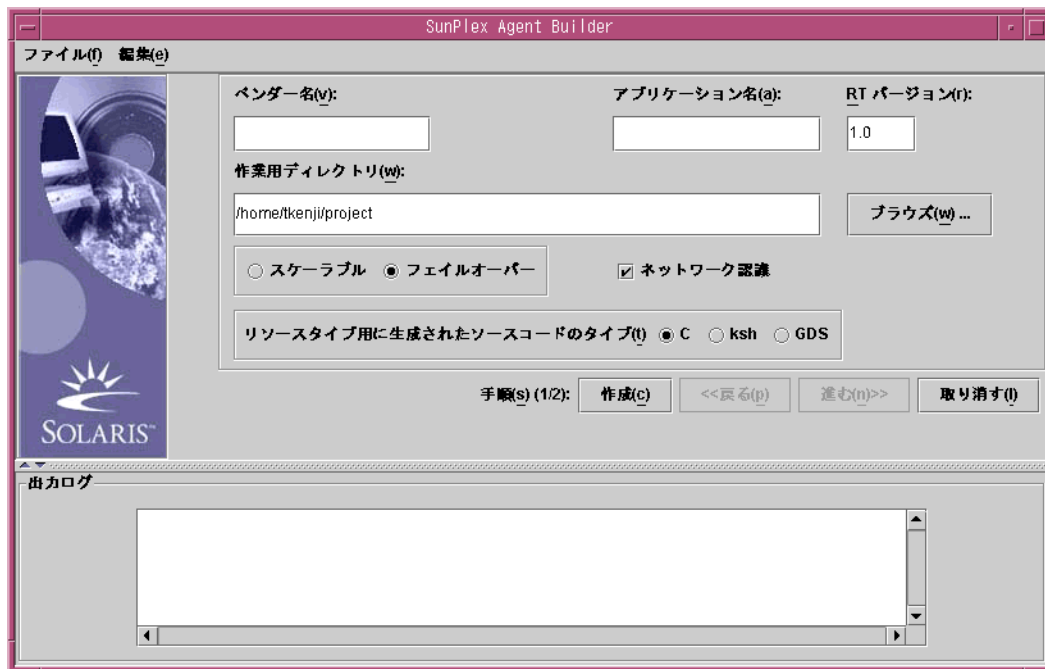


図 9-1 作成画面

Agent Builder のナビゲーション

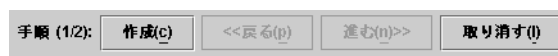
「作成」や「構成」画面の情報は、次のようにして入力します。

- フィールドに情報を入力。
- ディレクトリ構造をブラウズして、ファイルまたはディレクトリを選択。
- ラジオボタンの中から1つだけ選択 (たとえば、「スケラブル」または「フェイルオーバー」)。

- ボックスのオンまたはオフ。たとえば、「ネットワーク認識」ボックスをクリックすると、ベースアプリケーションがネットワーク対応であることを指定します。クリックしなければ、アプリケーションが非ネットワーク対応であることを指定します。

各画面の下にあるオプションを使用すると、作業を完了したり、次の画面に進んだり、以前の画面に戻ったり、Agent Builder を終了したりできます。Agent Builder は、状況に応じてこれらのオプションを強調表示したり、グレイ表示したりします。

たとえば、「作成」画面において、必要なフィールドに入力し、希望のオプションにチェックマークを付けてから、画面の下にある「作成」ボタンをクリックします。この時点で、以前の画面は存在しないので、「戻る」ボタンはグレイ表示されます。また、この作業が完成するまで次の手順には進めないで、「進む」ボタンもグレイ表示されます。



Agent Builder は、画面の下にある出力ログ領域に進捗メッセージを表示します。作業が終了したとき、Agent Builder は成功メッセージまたは警告メッセージを出力ログに表示します。「進む」が強調表示されます。あるいは、これが最後の画面の場合、「キャンセル」だけが強調表示されます。

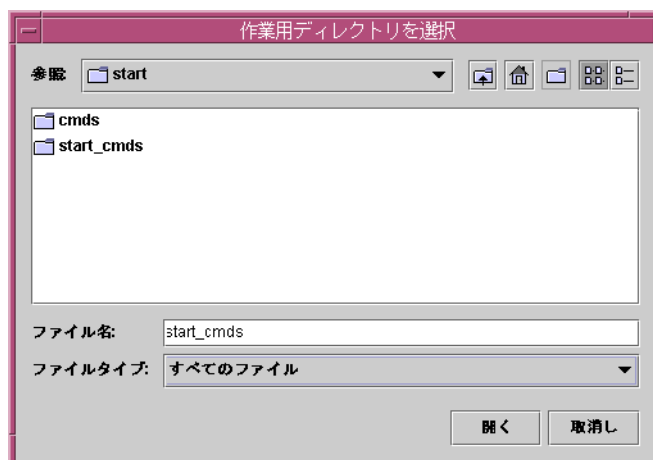
「キャンセル」ボタンを押すと、いつでも Agent Builder を終了できます。

検索

Agent Builder のフィールドの中には、情報を直接入力することも、「ブラウズ」ボタンをクリックしてディレクトリ構造をブラウズし、ファイルまたはディレクトリを選択することも可能なフィールドがあります。



「ブラウズ」をクリックすると、このような画面が表示されます。



フォルダをダブルクリックすると、フォルダが開きます。カーソルをファイルに移動すると、ファイルの名前が「ファイル名」ボックスに表示されます。必要なファイルを見つけ、そこにカーソルを移動したら、「選択」をクリックします。

注 - ディレクトリをブラウズする場合は、必要なディレクトリにカーソルを移動し、「開く」をクリックします。サブディレクトリがない場合、Agent Builder はブラウズウィンドウを閉じて、カーソルで指示されたディレクトリ名を適切なフィールドに表示します。サブディレクトリがある場合、「閉じる」をクリックすると、ブラウズウィンドウが閉じて、以前の画面に戻ります。Agent Builder は、カーソルで指示されたディレクトリ名を適切なフィールドに表示します。

画面の右上隅にあるアイコンは、次のような処理を行います。



ディレクトリツリーの1つ上のレベルに移動します。



ホームフォルダに戻ります。



現在選択しているフォルダの下に新しいフォルダを作成します。



ビューを切り替えます。将来のために予約されています。

メニュー

Agent Builder には、「ファイル」と「編集」プルダウンメニューがあります。

「ファイル」メニュー

「ファイル」メニューには、次の2つのオプションがあります。

- 「リソースタイプをロード」 既存のリソースタイプをロードします。Agent Builder が提供するブラウザ画面を使用して、既存のリソースタイプ用の作業ディレクトリを選択します。Agent Builder を起動したディレクトリにリソースタイプが存在する場合、Agent Builder は自動的にそのリソースタイプをロードします。「リソースタイプをロード」コマンドを使用すると、任意のディレクトリから Agent Builder を起動した後、既存のリソースタイプを選択して、新しいリソースタイプを作成するためのテンプレートとして使用できます。詳細は、[173 ページの「既存のリソースタイプからクローンを作成する方法」](#)を参照してください。
- 「終了」 Agent Builder を終了します。「作成」または「構成」画面で「キャンセル」をクリックして Agent Builder を終了することもできます。

「編集」メニュー

「編集」メニューには、次のオプションがあります。

- 「出力ログをクリア」。出力ログの情報を消去します。「作成」または「構成」を選択するたびに、Agent Builder は状態メッセージを出力ログに追加します。対話形式で繰り返してソースコードを修正し、Agent Builder で出力を生成し直しているときに、出力の生成ごとに状態メッセージを記録したい場合は、出力ログを使用するたびにログファイルの内容を保存および消去できます。
- 「出力ログを保存」。ログ出力をファイルに保存します。Agent Builder が提供するブラウザ画面を使用すると、ディレクトリを選択して、ファイル名を指定できます。

作成画面の使用

作成画面

リソースタイプを作成する最初の段階では、Agent Builder を起動したときに表示される作成画面に必要な情報を入力します。すると、次の画面が表示されます。

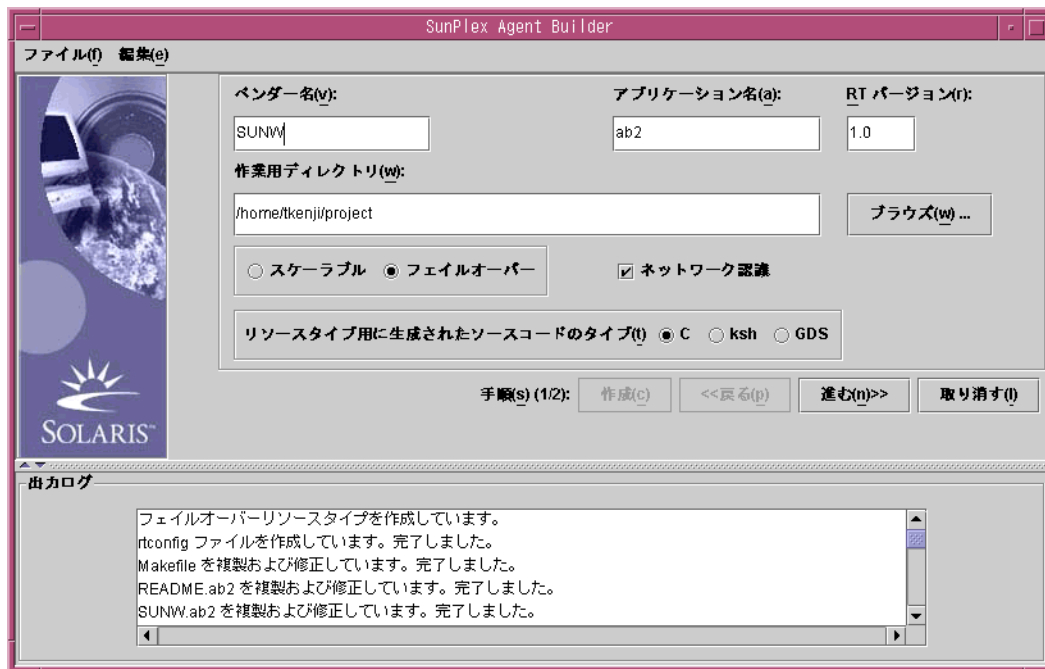


図 9-2 作成画面

作成画面には、次のフィールド、ラジオボタン、およびチェックボックスがあります。

- 「ベンダー名」。リソースタイプのベンダーの名前。通常、ベンダーの略号を指定します。ベンダーを一意に識別する名前であれば、どのような名前でも有効です。英数字文字だけを使用します。
- 「アプリケーション名」。リソースタイプの名前。英数字文字だけを使用します。

注 - ベンダー名とアプリケーション名の両方で、リソースタイプの完全な名前が形成されます。完全な名前は 9 文字を超えてはなりません。

- 「RT バージョン」。生成されるリソースタイプのバージョン。同一のベースリソースタイプのバージョン (またはアップグレード) が複数登録されている場合は、「RT バージョン」でそれらを区別します。
「RT バージョン」フィールドで次の文字を使用することはできません。空白文字、タブ、スラッシュ (/)、逆スラッシュ (\)、アスタリスク (*)、疑問符 (?)、コンマ (,)、セミコロン (;)、左角かっこ ((), 右角かっこ ())。
- 「作業ディレクトリ」Agent Builder は、このディレクトリの下に、ターゲットリソースタイプ用のすべてのファイルを格納するディレクトリ構造を作成します。1 つの作業ディレクトリには 1 つのリソースタイプしか作成できません。Agent Builder は、このフィールドを Agent Builder が起動されたディレクトリのパスで初期化します。ただし、別のディレクトリ名を入力したり、「ブラウズ」を使用して異なるディレクトリを指定することもできます。
Agent Builder は、作業ディレクトリの下にリソースタイプ名を持つサブディレクトリを作成します。たとえば、ベンダー名が SUNW で、アプリケーション名が ftp である場合、Agent Builder はこのサブディレクトリに SUNWftp という名前を付けます。
Agent Builder は、ターゲットリソースタイプのすべてのディレクトリとファイルをこのサブディレクトリの下に置きます。詳細は、175 ページの「ディレクトリ構造」を参照してください。
- 「スケラブル」または「フェイルオーバー」。ターゲットのリソースタイプがフェイルオーバーなのかスケラブルなのかを指定します。
- 「ネットワーク認識」。ベースアプリケーションがネットワーク対応かどうかを指定します。つまり、アプリケーションがネットワークを使用してクライアントと通信するかどうかを指定します。ネットワーク対応を指定する場合は、「ネットワーク認識」チェックボックスを選択します。非ネットワーク対応を指定する場合は選択しません。
- 「C, ksh」。生成されるソースコードの言語を指定します。この 2 つのオプションを同時に指定することはできませんが、Agent Builder では、Korn シェルで生成されたコードでリソースタイプを作成してから、同じ情報を再使用して、C で生成されたコードを作成することができます。詳細は、173 ページの「既存のリソースタイプからクローンを作成する方法」を参照してください。
- 「GDS」。このサービスが汎用データサービスであることを示します。汎用データサービスの詳しい作成および構成方法については、第 10 章を参照してください。

注 - cc コンパイラが \$PATH に含まれていないと、Agent Builder は「C」ラジオボタンをグレイ表示し、「ksh」ラジオボタンを選択可能にします。異なるコンパイラを指定する方法については、160 ページの「Agent Builder のインストールと構成」の最後にある注を参照してください。

必要な情報を入力した後、「Create」をクリックします。画面の一番下にある「出力ログ」ウィンドウには、Agent Builder が実行するアクションが表示されます。「編集」メニューの「Save Output Log」を使用すれば、出力ログ内の情報を保存できます。

これが終わると、Agent Builder は、成功メッセージか警告メッセージを表示します。

- Agent Builder がこのステップを正常に終了できなかった場合は、出力ログの詳細な情報を調べてください。
- Agent Builder が成功メッセージを表示した場合は、「進む」をクリックして「構成」画面を表示します。この画面でリソースタイプの生成を完結することができます。

注 – 完全なリソースタイプを生成するには、2 段階の作業が必要ですが、最初の段階 (つまり、作成) が完了した後に Agent Builder を終了しても、入力した情報や Agent Builder で作成した内容が失われることはありません。詳細は、173 ページの「完成した作業内容の再利用」を参照してください。

構成画面の使用

構成画面

Agent Builder がリソースタイプの作成を終了した後に、「作成」画面の「次へ」を選択すると、以下の図に示す「構成」画面が表示されます。リソースタイプの作成が完了していなければ、構成画面にはアクセスできません。

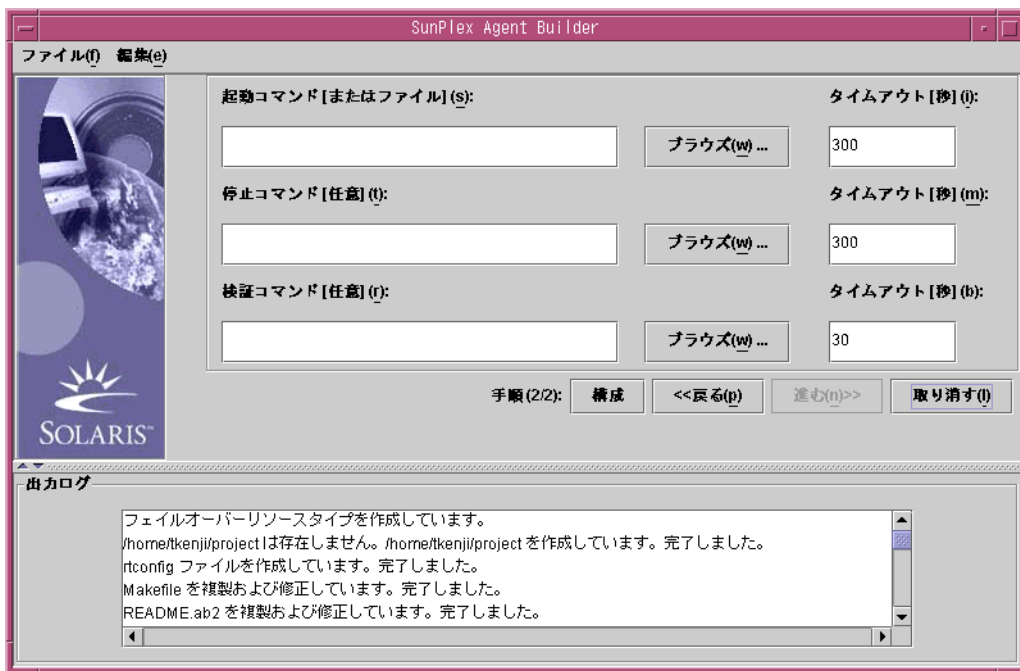


図 9-3 構成画面

構成画面には、次のフィールドがあります。

- 「起動コマンド」。ベースアプリケーションを起動するために任意の UNIX シェルに渡すことができる完全なコマンド行。これには、起動コマンドを指定する必要があります。このフィールドにコマンドを入力するか、「ブラウズ」を使用して、アプリケーションを起動するコマンドが記述されているファイルを指定します。

完全なコマンド行には、アプリケーションを起動するのに必要なすべての要素が含まれていなければなりません。たとえば、ホスト名、ポート番号、構成ファイルへのパスなどです。あるいは、プロパティ変数を指定することもできます。この変数については、171 ページの「プロパティ変数」を参照してください。Korn シェルベースのアプリケーションにコマンド行からホスト名を指定する必要がある場合は、Agent Builder が定義する \$hostnames 変数を使用できます。詳細は、171 ページの「Agent Builder の Korn シェルベース \$hostnames 変数の使用」を参照してください。

コマンドは二重引用符 (") で囲んではなりません。

注 - ベースアプリケーションが複数の独立したプロセスツリーを持ち、各プロセスツリーが Process Monitor Facility (PMF) の制御下で独自のタグによって起動される場合、単一のコマンドは指定できません。代わりに、各プロセスツリーを起動するための個々のコマンドを記述したテキストファイルを作成し、そのファイルへのパスを「起動コマンド」テキストフィールドに指定する必要があります。このファイルが適切に機能するために必要な特性については、158 ページの「複数の独立したプロセスツリーを持つリソースタイプの作成」を参照してください。

- 「停止コマンド」。ベースアプリケーションを停止するために任意の UNIX シェルに渡すことができる完全なコマンド行。このフィールドにコマンドを入力するか、「ブラウザ」を使用して、アプリケーションを停止するコマンドが記述されているファイルを指定します。あるいは、プロパティ変数を指定することもできます。この変数については、171 ページの「プロパティ変数」を参照してください。Korn シェルベースのアプリケーションにコマンド行からホスト名を指定する必要がある場合は、Agent Builder が定義する \$hostnames 変数を使用できます。詳細は、171 ページの「Agent Builder の Korn シェルベース \$hostnames 変数の使用」を参照してください。

このコマンドは省略可能です。停止コマンドを指定しない場合、生成されるコードは、次に示すように、Stop メソッドでシグナルを使用して、アプリケーションを停止します。

- Stop メソッドは SIGTERM を送信してアプリケーションを停止しようとします。そして、アプリケーション用のタイムアウト値の 80% だけ待機して、停止しない場合は終了します。
- SIGTERM シグナルが失敗した場合、Stop メソッドは SIGKILL を送信して、アプリケーションを停止しようとします。そして、アプリケーション用のタイムアウト値の 15% だけ待機して、停止しない場合は終了します。
- SIGKILL が失敗した場合、Stop メソッドは異常終了します。タイムアウト値の残りの 5% はオーバーヘッドとみなされます。



注意 - 停止コマンドは、アプリケーションが完全に停止するまで戻らないことに注意してください。

- 「検証コマンド」。定期的に行われ、アプリケーションの状態を検査して、0 (正常) から 100 (致命的な障害) の範囲の終了状態に戻すコマンド。このコマンドは省略可能です。このフィールドにコマンドの完全パスを入力するか、「ブラウザ」を使用して、アプリケーションを検証するコマンドが記述されているファイルを指定します。

通常は、単にベースアプリケーションのクライアントを指定します。検証コマンドを指定しない場合、生成されるコードは、リソースが使用するポートへの接続と切断を試みます。接続と切断に成功すれば、アプリケーションの状態が正常であると判断します。あるいは、プロパティ変数を指定することもできます。この変数については、171 ページの「プロパティ変数」を参照してください。Korn シェルベー

スのアプリケーションに検証コマンド行からホスト名を指定する必要がある場合は、Agent Builder が定義する `$hostnames` 変数を使用できます。詳細は、171 ページの「Agent Builder の Korn シェルベース `$hostnames` 変数の使用」を参照してください。

- 「Timeout.」各コマンドのタイムアウト値 (秒数)。新しい値を指定するか、Agent Builder が提供するデフォルト値を受け入れます。起動コマンドと停止コマンドのデフォルト値は 300 秒です。検証コマンドのデフォルト値は 30 秒です。

Agent Builder の Korn シェルベース `$hostnames` 変数の使用

多くのアプリケーション (特にネットワーク対応のアプリケーション) では、アプリケーションが待機するホスト名や、顧客が求めるサービスをコマンド行からアプリケーションに渡す必要があります。多くの場合、ホスト名は、構成画面において、ターゲットリソースタイプの起動、停止、および検証コマンドに指定する必要があります。つまり、ホスト名はリソースがクラスタで実行されるときに決められ、Agent Builder がリソースタイプコードを生成する時点に決めることはできません。

この問題を解決するために、Agent Builder は `$hostnames` 変数を提供します。この変数を使用すると、起動、停止、および検証コマンドのコマンド行にホスト名を指定できます。

注 - `$hostnames` 変数は、Korn シェルベースのサービスでのみサポートされます。つまり、C ベースや GDS ベースのサービスではサポートされません。

`$hostnames` 変数を指定する方法は、実際のホスト名を指定する方法と同じです。たとえば、次のようになります。

```
% /opt/network_aware/echo_server -p port_no -l $hostnames
```

ターゲットリソースタイプのリソースがあるクラスタ上で動作するとき、(リソースの `Network_resources_used` リソースプロパティで) そのリソースに構成されている `LogicalHostname` または `SharedAddress` ホスト名が `$hostnames` 変数の値に置き換えられます。

`Network_resources_used` プロパティに複数のホスト名を構成している場合、すべてのホスト名をコンマで区切って `$hostnames` 変数に指定します。

プロパティ変数

プロパティ変数を使用すれば、Sun Cluster のリソース、リソースタイプ、リソースグループの一部のプロパティの値を RGM フレームワークから取り出すことができます。Agent Builder は、起動、検証、停止のスクリプトをスキャンし、プロパティ変数があれば、スクリプトを起動する前にこれらの変数を対応する値で置き換えます。

注 - プロパティ変数は、Korn シェルベースのサービスではサポートされません。

プロパティ変数のリスト

以下のリストは、スクリプトで使用できるプロパティ変数を示しています。Sun Cluster のリソース、リソースタイプ、リソースグループのプロパティについては、[付録 A](#) を参照してください。

以下は、リソースプロパティ変数のリストです。

- HOSTNAMES
- RS_CHEAP_PROBE_INTERVAL
- RS_MONITOR_START_TIMEOUT
- RS_MONITOR_STOP_TIMEOUT
- RS_NAME
- RS_NUM_RESTARTS
- RS_RESOURCE_DEPENDENCIES
- RS_RESOURCE_DEPENDENCIES_WEAK
- RS_RETRY_COUNT
- RS_RETRY_INTERVAL
- RS_SCALABLE
- RS_START_TIMEOUT
- RS_STOP_TIMEOUT
- RS_THOROUGH_PROBE_INTERVAL
- SCHA_STATUS

以下は、リソースタイププロパティ変数のリストです。

- RT_API_VERSION
- RT_BASEDIR
- RT_FAILOVER
- RT_INSTALLED_NODES
- RT_NAME
- RT_RT_VERSION
- RT_SINGLE_INSTANCE

以下は、リソースグループプロパティ変数のリストです。

- RG_DESIRED_PRIMARYES
- RG_GLOBAL_RESOURCES_USED
- RG_IMPLICIT_NETWORK_DEPENDENCIES
- RG_MAXIMUM_PRIMARYES
- RG_NAME
- RG_NODELIST
- RG_NUM_RESTARTS
- RG_PATHPREFIX
- RG_PINGPONG_INTERVAL

- `RG_RESOURCE_LIST`

プロパティ変数の構文

プロパティ変数を指定する場合は、プロパティ名の前にパーセント符号 (%) を指定します。次はその例です。

```
# /opt/network_aware/echo_server -t %RS_STOP_TIMEOUT -n %RG_NODELIST
```

上の例の場合、Agent Builder はこれらのプロパティ変数を解釈し、たとえば、次の値を使って `echo_server` スクリプトを起動します。

```
# /opt/network_aware/echo_server -t 300 -n phys-node-1,phys-node-2,phys-node-3
```

Agent Builder によるプロパティ変数の置き換え

Agent Builder では、プロパティ変数のタイプは次のように解釈されます。

- 整数は、その実際の値 (たとえば、300) で置き換えられます。
- ブール値は、文字列 `TRUE` か `FALSE` で置き換えられます。
- 文字列は、実際の文字列 (たとえば、`phys-node-1`) で置き換えられます。
- 文字列リストの場合は、リストが、コンマで区切られた実際の値で置き換えられます (たとえば、`phys-node-1,phys-node-2,phys-node-3`)。
- 整数リストの場合は、リストが、コマンドで区切られた実際の整数 (たとえば、`1,2,3`) で置き換えられます。
- 列挙タイプは、その値 (文字列形式) で置き換えられます。

完成した作業内容の再利用

Agent Builder を使用すると、次に示すように、完成した作業内容を再利用できます。

- Agent Builder で作成した既存のリソースタイプのクローンを作成できます。
- Agent Builder が生成したソースコードを編集して、そのコードを再コンパイルすれば、新しいパッケージを作成できます。

▼ 既存のリソースタイプからクローンを作成する方法

Agent Builder で作成した既存のリソースタイプのクローンを作成するには、次の手順に従います。

1. 次のどちらかの方法で既存のリソースタイプを **Agent Builder** にロードします。
 - Agent Builder で作成された既存のリソースタイプの作業ディレクトリ (`rtconfig` ファイルが含まれている) から Agent Builder を起動します。Agent Builder がこのリソースタイプの値を「作成」や「構成」画面にロードします。

- 「ファイル」プルダウンメニューの「Load Resource Type」オプションを使用します。
2. 作成画面で作業ディレクトリを変更します。
「ブラウズ」を使ってディレクトリを選択する必要があります。新しいディレクトリ名を入力するだけでは不十分です。ディレクトリを選択したあと、Agent Builder は「作成」ボタンを有効に戻します。
 3. 必要な変更を行います。
この手順は、リソースタイプ用に生成されたコードのタイプを変更するときに使用できます。たとえば、最初は Korn シェルバージョンのリソースタイプを作成していたが、後で C バージョンのリソースタイプが必要になった場合などです。この場合、既存の Korn シェルバージョンのリソースタイプをロードし、出力用の言語を C に変更してから Agent Builder で C バージョンのリソースタイプを構築します。
 4. リソースタイプのクローンを作成します。
「作成」をクリックして、リソースタイプを作成します。「次へ」をクリックして「構成」画面を表示します。「構成」をクリックしてリソースタイプを構成し、次に「キャンセル」をクリックして終了します。

生成されたソースコードの編集

リソースタイプを作成するプロセスを簡単にするために、Agent Builder は入力数を制限しています。必然的に、生成されるリソースタイプの範囲も制限されます。したがって、より複雑な機能、たとえば、追加のプロパティの妥当性をチェックしたり、Agent Builder がエクスポートしないパラメータを調整したりする機能を追加するには、生成されたソースコードまたは RTR ファイルを修正する必要があります。

ソースファイルは `install_directory/rt_name/src` ディレクトリに置かれます。Agent Builder は、ソースコード内においてコードを追加できる場所にコメント文を埋め込みます。このようなコメントの形式は次のとおりです (C コードの場合)。

```
/* User added code -- BEGIN vvvvvvvvvvvvvvvvvv */
/* User added code -- END   ^^^^^^^^^^^^^^^^^^^^ */
```

注 - コメントは Korn シェルソースコードのものと同じですが、Korn シェルソースコードの場合は、ポンド記号 (#) がコメントの始めを表します。

たとえば、`rt_name.h` は、異なるプログラムが使用するすべてのユーティリティールーチンを宣言します。宣言リストの最後はコメント文になっており、ここでは自分のコードに追加したいルーチンを宣言できます。

`install_directory/rt_name/src` ディレクトリには、適切なターゲットとともに、`makefile` も生成されます。`make` コマンドを使用すると、ソースコードを再コンパイルできます。また、`make pkg` コマンドを使用すると、リソースタイプパッケージを生成し直すことができます。

RTR ファイルは `install_directory/rt_name/etc` ディレクトリに置かれます。RTR ファイルは、普通のテキストエディタで編集できます。RTR ファイルについては 30 ページの「リソースとリソースタイププロパティの設定」を、プロパティについては 付録 A をそれぞれ参照してください。

▼ コマンド行バージョンの Agent Builder を使用する 方法

コマンド行バージョンの Agent Builder でも、グラフィカルユーザーインターフェースと同様の基本手順を使用します。ただし、グラフィカルユーザーインターフェースでは情報を入力しましたが、コマンド行インターフェースでは `scdscreate` や `scdsconfig` コマンドにパラメータを渡します。詳細は、`scdscreate(1HA)` と `scdsconfig(scdsconfig)` のマニュアルページを参照してください。

コマンド行バージョンの Agent Builder の使用方法は次のとおりです。

1. アプリケーションに高可用性またはスケーラビリティを持たせるため、**scdscreate** を使って **Sun Cluster** リソースタイプテンプレートを作成します。
2. **scdsconfig** を使って、**scdscreate** で作成したリソースタイプテンプレートを構成します。
プロパティ変数を指定できます。プロパティ変数については、171 ページの「プロパティ変数」を参照してください。
3. 作業ディレクトリの **pkg** サブディレクトリに移動します。
4. **pkgadd** コマンドを実行して、**scdscreate** で作成したパッケージをインストールします。
5. 必要に応じて、生成されたソースコードを編集します。
6. 起動スクリプトを実行します。

ディレクトリ構造

Agent Builder は、ターゲットリソースタイプ用に生成するすべてのファイルを格納するためのディレクトリ構造を作成します。「作成」画面で作業ディレクトリを指定します。開発するリソースタイプごとに異なるインストールディレクトリを指定する必要があります。Agent Builder は、作業ディレクトリの下に、「作成」画面で入力されたベンダー名とリソースタイプ名を連結した名前を持つサブディレクトリを作成します。たとえば、SUNW というベンダー名を指定し、ftp というリソースタイプを作成した場合、Agent Builder は SUNWftp というディレクトリを作業ディレクトリの下に作成します。

Agent Builder は、このサブディレクトリの下に、次のようなディレクトリを作成し、各ディレクトリにファイルを配置します。

ディレクトリ名	目次
bin	C 出力の場合、ソースファイルからコンパイルしたバイナリファイルが格納されます。Korn シェル出力の場合、src ディレクトリと同じファイルが格納されます。
etc	RTR ファイルが格納されます。Agent Builder は、ベンダー名とアプリケーション名をピリオド区切り (.) で結合して RTR ファイル名を作成します。たとえば、ベンダー名が SUNW で、リソースタイプ名が ftp である場合、RTR ファイル名は SUNW.ftp となります。
man	start、stop および remove ユーティリティースクリプト用にカスタマイズされたマニュアルページが格納されます。たとえば、startftp(1M)、stopftp(1M)、および removeftp(1M) が格納されます。 これらのマニュアルページを見る場合は、man -M オプションでこのパスを指定します。次に例を示します。 <pre># man -M install_directory/SUNWftp/man removeftp</pre>
pkg	最終的なパッケージが格納されます。
src	Agent Builder によって生成されたソースファイルが格納されます。
util	Agent Builder によって生成された start、stop、および remove ユーティリティースクリプトが格納されます。178 ページの「ユーティリティースクリプトとマニュアルページ」を参照してください。Agent Builder は、これらのスクリプト名にアプリケーション名を追加します。たとえば、startftp、stopftp、および removeftp のようになります。

Agent Builder の出力

この節では、Agent Builder の出力について説明します。

ソースファイルとバイナリファイル

リソースグループと、最終的にはクラスタ上のリソースを管理する Resource Group Manager (RGM) は、コールバックモデル上で動作します。つまり、特定のイベント (ノードの障害など) が発生したとき、RGM は、当該ノード上で動作しているリソースごとにリソースタイプのメソッドを呼び出します。たとえば、RGM は Stop メソッドを呼び出して、当該ノード上で動作しているリソースを停止します。次に、Stop メソッドを呼び出して、異なるノード上でリソースを起動します。このモデルについては、19 ページの「RGM モデル」と 21 ページの「コールバックメソッド」のほかに、rt_callbacks(1HA) のマニュアルページを参照してください。

このモデルをサポートするために Agent Builder は、コールバックメソッドの役目をする 8 つの実行可能 C プログラムまたは Korn シェルスクリプトを *install_directory/rt_name/bin* ディレクトリに生成します。

注 - 厳密には、障害モニターを実装する *rt_name_probe* プログラムはコールバックプログラムではありません。RGM は *rt_name_probe* を直接呼び出すのではなく、*rt_name_monitor_start* および *rt_name_monitor_stop* を呼び出します。そして、これらのメソッドが *rt_name_probe* を呼び出すことによって、障害モニターが起動および停止されます。

Agent Builder が生成する 8 つのメソッドは次のとおりです。

- *rt_name_monitor_check*
- *rt_name_monitor_start*
- *rt_name_monitor_stop*
- *rt_name_probe*
- *rt_name_svc_start*
- *rt_name_svc_stop*
- *rt_name_update*
- *rt_name_validate*

各メソッドに固有な情報については、*rt_callbacks(1HA)* のマニュアルページを参照してください。

Agent Builder は、*install_directory/rt_name/src* ディレクトリ (C 出力の場合) に、次のファイルを生成します。

- ヘッダーファイル (*rt_name.h*)
- すべてのメソッドに共通なコードが記述されているソースファイル (*rt_name.c*)
- 共通なコード用のオブジェクトファイル (*rt_name.o*)
- 各メソッド用のソースファイル (*.c)
- 各メソッド用のオブジェクトファイル (*.o)

Agent Builder は *rt_name.o* ファイルを各メソッドの *.o* ファイルにリンクして、実行可能ファイルを *install_directory/rt_name/bin* ディレクトリに作成します。

Korn シェル出力の場合、*install_directory/rt_name/bin* ディレクトリと *install_directory/rt_name/src* ディレクトリの内容は同じです。それぞれのディレクトリには、7 つのコールバックメソッドと Probe メソッドに対応する 8 つの実行可能スクリプトが含まれています。

注 - Korn シェル出力には、2 つのコンパイル済みユーティリティプログラム (*gettime* と *gethostnames*) が含まれています。これらのプログラムは、特定のコールバックメソッドが時間の取得や、検証を行う際に必要です。

ソースコードを編集して、make コマンドを実行すると、コードを再コンパイルできます。さらに、再コンパイル後、make pkg コマンドを実行すると、新しいパッケージを生成できます。ソースコードの修正をサポートするために、Agent Builder はソースコード中の適切な場所 (コードを追加できる場所) にコメント文を埋め込みます。174 ページの「生成されたソースコードの編集」を参照してください。

ユーティリティースクリプトとマニュアルページ

リソースタイプを生成し、そのパッケージをクラスタにインストール後、リソースタイプのインスタンス (リソース) をクラスタ上で実行する必要があります。一般に、リソースを実行するには、管理コマンドまたは SunPlex Manager を使用します。Agent Builder は、ターゲットリソースタイプのリソースを起動するためのカスタマイズされたユーティリティースクリプトに加え、リソースを停止および削除するスクリプトも生成します。これら 3 つのスクリプトは *install_directory/rt_name/util* ディレクトリに格納されており、次のような処理を行います。

- 起動スクリプトリソースタイプを登録し、必要なリソースグループとリソースを作成します。また、アプリケーションがネットワーク上のクライアントと通信するためのネットワークアドレスリソース (LogicalHostname または SharedAddress) も作成します。
- 停止スクリプトリソースを停止し、無効にします。
- 削除スクリプト起動スクリプトによる作業を取り消します。つまり、このスクリプトは、リソース、リソースグループ、ターゲットリソースタイプを停止し、システムから削除します。

注 - これらのスクリプトは内部的な規則を使用して、リソースとリソースグループの名前付けを行います。そのため、削除スクリプトを使用できるリソースは、対応する起動スクリプトで起動されたリソースだけです。

Agent Builder は、スクリプト名にアプリケーション名を追加することにより、スクリプトの名前付けを行います。たとえば、アプリケーション名が ftp の場合、各スクリプトは startftp、stopftp、および removeftp になります。

Agent Builder は、各ユーティリティースクリプト用のマニュアルページを *install_directory/rt_name/man/man1m* ディレクトリに格納します。これらのマニュアルページにはスクリプトに渡す必要があるパラメータについての説明が記載されているので、各スクリプトを起動する前に、これらのマニュアルページをお読みください。

これらのマニュアルページを表示するには、man コマンドに -M オプションを付けて、上記のマニュアルページが格納されているディレクトリへのパスを指定する必要があります。たとえば、ベンダーが SUNW で、アプリケーション名が ftp である場合、startftp(1M) のマニュアルページを表示するには、次のコマンドを使用します。

```
% man -M install_directory/SUNWftp/man startftp
```

クラスタ管理者は、マニュアルページユーティリティスクリプトも利用できます。Agent Builder が生成したパッケージをクラスタ上にインストールすると、ユーティリティスクリプト用のマニュアルページは、`/opt/rt_name/man` ディレクトリに格納されます。たとえば、`startftp(1M)` のマニュアルページを表示するには、次のコマンドを使用します。

```
% man -M /opt/SUNWftp/man startftp
```

サポートファイル

Agent Builder はサポートファイル (`pkginfo`、`postinstall`、`postremove`、`preremove` など) を `install_directory/rt_name/etc` ディレクトリに格納します。このディレクトリには、Resource Type Registration (RTR) ファイルも格納されます。RTR ファイルは、ターゲットリソースタイプが利用できるリソースとリソースタイププロパティを宣言して、リソースをクラスタに登録するときにプロパティ値を初期化します。詳細は、30 ページの「リソースとリソースタイププロパティの設定」を参照してください。RTR ファイルの名前は、ベンダー名とリソースタイプ名をピリオドで区切って連結したものです (たとえば、`SUNW.ftp`)。

RTR ファイルは、ソースコードを再コンパイルしなくても、標準のテキストエディタで編集および修正できます。ただし、`make pkg` コマンドを使用して、パッケージを再構築する必要があります。

パッケージディレクトリ

`install_directory/rt_name/pkg` ディレクトリには、Solaris パッケージが格納されます。パッケージの名前は、ベンダー名とアプリケーション名を連結したものです (たとえば、`SUNWftp`)。 `install_directory/rt_name/src` ディレクトリ内の `Makefile` は、新しいパッケージを作成するのに役立ちます。たとえば、ソースファイルを修正し、コードを再コンパイルした場合、あるいは、パッケージユーティリティスクリプトを修正した場合、`make pkg` コマンドを使用して新しいパッケージを作成します。

パッケージをクラスタから削除する場合、同時に複数のノードから `pkgrm` コマンドを実行しようとする、`pkgrm` コマンドが失敗する可能性があります。この問題を解決するには、次の 2 つの方法があります。

- クラスタの 1 つのノードで `remove rt_name` スクリプトを実行してから、任意のノードで `pkgrm` を実行します。
- クラスタの 1 つのノードで `pkgrm` を実行して、必要なすべてのクリーンアップを行なった後、(必要であれば同時に) 残りのノードで `pkgrm` を実行します。

同時に複数のノードから `pkgrm` を実行しようとして失敗した場合は、再度いずれかのノードで `pkgrm` を実行した後、残りのノードで `pkgrm` を実行します。

rtconfig ファイル

C または Korn シェルソースコードを作業ディレクトリ内に生成する場合、Agent Builder は構成ファイル `rtconfig` を生成します。このファイルには、作成および構成画面でユーザーが入力した情報が含まれています。既存のリソースタイプの作業ディレクトリから Agent Builder を起動すると (または「ファイル」プルダウンメニューの「リソースタイプのロード」を選択して既存のリソースタイプをロードすると)、Agent Builder は、`rtconfig` ファイルを読み取り、リソースタイプの作成時に指定された情報を「作成」や「構成」画面に表示します。この機能は、既存のリソースタイプのクローンを作成したい場合に便利です。詳細は、[173 ページの「既存のリソースタイプからクローンを作成する方法」](#)を参照してください。

Agent Builder の Cluster Agent モジュール

Agent Builder の Cluster Agent モジュールは、NetBeans™ モジュールです。この Cluster Agent モジュールを使用することで、Sun Java Studio (以前の Sun ONE Studio) 製品のユーザーは統合された開発環境で Sun Cluster ソフトウェアのリソースタイプ (データサービス) を作成できます。この Cluster Agent モジュールは、ユーザーが作成するリソースタイプの種類を説明する画面ベースのインタフェースを提供します。

注 - Sun Java Studio 製品の設定、インストール、使用についての詳細は、Sun Java Studio マニュアルに記載されています。

▼ Cluster Agent モジュールをインストールし設定する方法

Cluster Agent モジュールは、Sun Cluster ソフトウェアのインストール時にインストールされます。Sun Cluster のインストールツールは、Cluster Agent モジュールファイルを `/usr/cluster/lib/scdsbuilder` の `scdsbuilder.jar` に配置します。Sun Java Studio ソフトウェアで Cluster Agent モジュールを使用するには、このファイルに対してシンボリックリンクを作成する必要があります。

注 – Cluster Agent モジュールを実行する予定のシステムには、Sun Cluster 製品、Sun Java Studio 製品、および Java 1.4 がすでにインストールされ、使用可能な状況でなければなりません。

1. ユーザー全員が **Cluster Agent** モジュールを使用できるようにするか、あるいは自分だけが使用できるようにします。

- ユーザー全員に使用を認めるには、スーパーユーザーになるか、あるいは同等の役割を宣言し、汎用的なモジュールディレクトリにシンボリックリンクを作成します。

```
# cd /opt/s1studio/ee/modules
# ln -s /usr/cluster/lib/scdsbuilder/scdsbuilder.jar
```

注 – Sun Java Studio ソフトウェアを /opt/s1studio/ee 以外のディレクトリにすでにインストールしてある場合は、このディレクトリパスを、使用したパスに読み替えてください。

- 自分だけが使用できるようにするには、自分の `modules` サブディレクトリにシンボリックリンクを作成します。

```
% cd ~your-home-dir/ffjuser40ee/modules
% ln -s /usr/cluster/lib/scdsbuilder/scdsbuilder.jar
```

2. Sun Java Studio ソフトウェアを停止し、再起動します。

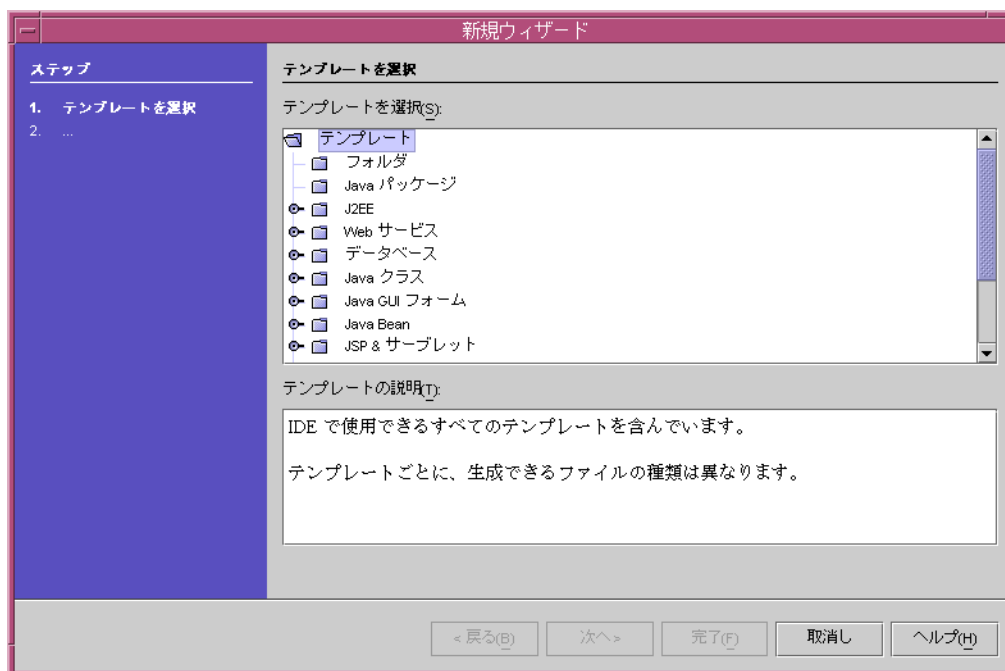
▼ Cluster Agent モジュールを起動する方法

次に、Sun Java Studio ソフトウェアから Cluster Agent モジュールを起動する手順を示します。

1. Sun Java Studio の「ファイル」メニューから「新規」を選択するか、あるいはツールバーの「新規」アイコンをクリックします。



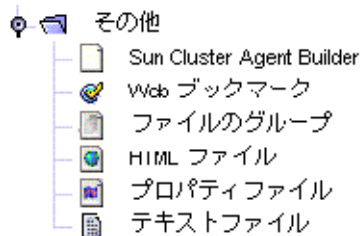
「新規ウィザード」画面が表示されます。



2. 「テンプレートを選択」ウィンドウで、必要に応じて下方向へスクロールし、その他フォルダの横に表示されている鍵マークをクリックします。

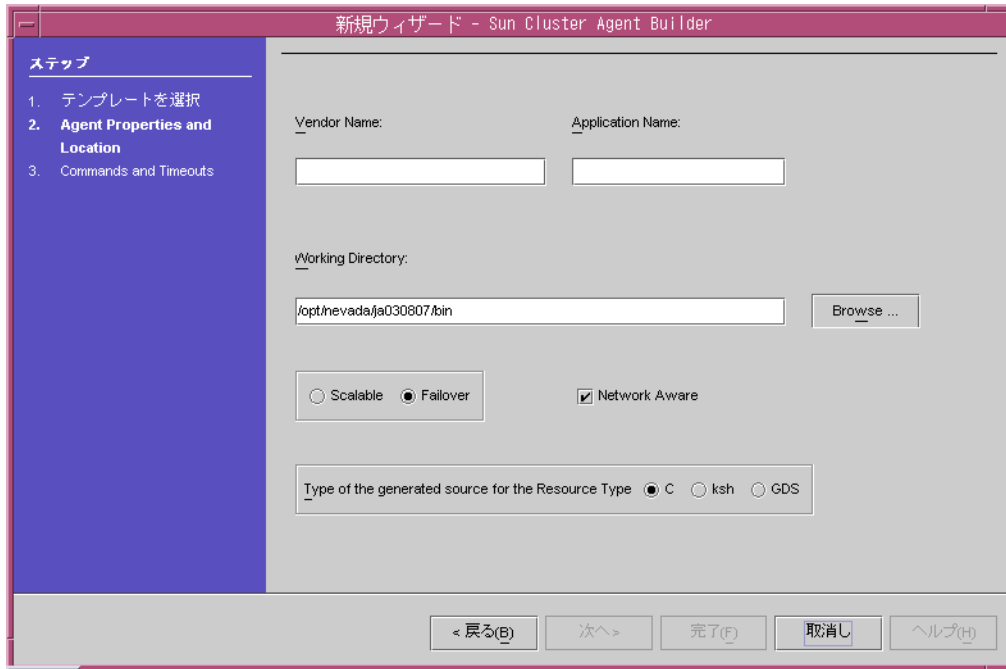


「その他」フォルダが開きます。



3. 「その他」フォルダから **Sun Cluster Agent Builder** を選択し、「次へ」をクリックします。

Sun Java Studio 起動のための Cluster Agent モジュール最初の「新規ウィザード - Sun Cluster Agent Builder」画面が表示されます。



Cluster Agent モジュールの使用

Cluster Agent モジュールは、Agent Builder ソフトウェアと同様に使用できます。インターフェースは英語版の Agent Builder ソフトウェアと全く同じです。たとえば次の図では、英語版 Agent Builder ソフトウェアの「Create」画面と Cluster Agent モジュールの最初の「新規ウィザード - Sun Cluster Agent Builder」画面には同じフィールドと選択肢が存在することがわかります。

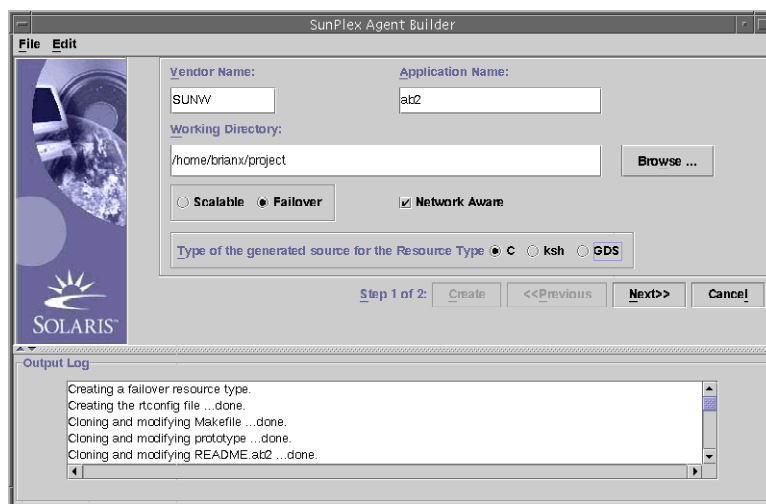


図 9-4 英語版 Agent Builder ソフトウェアの作成画面

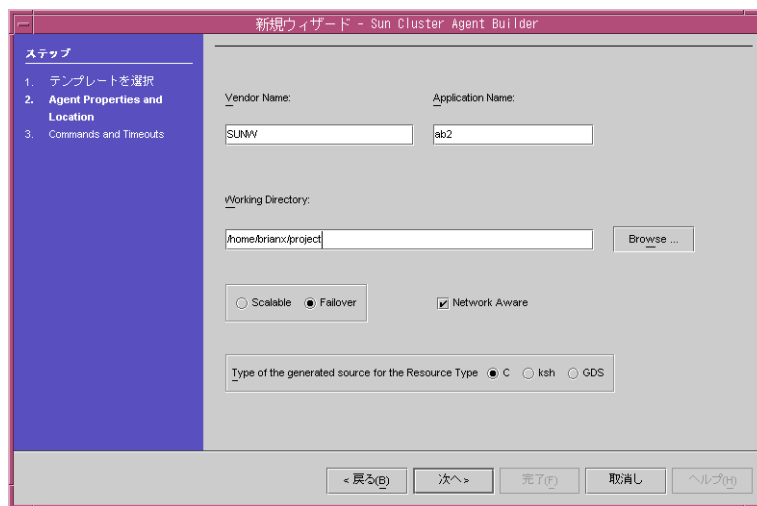


図 9-5 Cluster Agent モジュールの「新規ウィザード - Sun Cluster Agent Builder」画面

Cluster Agent モジュールと Agent Builder の違い

Cluster Agent モジュールと Agent Builder は似ていますが、小さな違いがいくつかあります。

- Cluster Agent モジュールでは、2 つ目の「新規ウィザード - Sun Cluster Agent Builder」画面で「完了」をクリックした時点でリソースタイプの作成と構成が完了します。最初の「新規ウィザード - Sun Cluster Agent Builder」画面で「次へ」をクリックした時点ではリソースタイプは作成されません。

英語版 Agent Builder では、「Create」画面で「Create」をクリックした時点でリソースタイプがただちに作成され、「構成」画面で「構成」をクリックする時点で構成が行われます。

- 英語版 Agent Builder の「Output Log」ウィンドウに表示される情報は、Sun Java Studio 製品では別の出カウィンドウで表示されます。

第 10 章

汎用データサービス

この章では、汎用データサービス (GDS) の概要を述べてから、GDS を使用するサービスの作成方法について説明します。このサービスの作成には、SunPlex Agent Builder か標準的な Sun Cluster 管理コマンドを使用します。

この章の内容は、次のとおりです。

- 187 ページの「GDS の概要」
- 194 ページの「SunPlex Agent Builder を使って GDS ベースのサービスを作成」
- 199 ページの「標準的な Sun Cluster 管理コマンドを使って GDS ベースのサービスを作成」
- 201 ページの「SunPlex Agent Builder のコマンド行インタフェース」

GDS の概要

GDS とは、簡単なネットワーク対応や非ネットワーク対応のアプリケーションを高可用性にしたり、スケーラブルにしたりするためのメカニズムです。そのためには、これらのアプリケーションを Sun Cluster Resource Group Management (RGM) フレームワークに組み込む必要があります。この機構では、アプリケーションを高可用性にしたり、スケーラブルにしたりするために通常必要になるエージェントのコーディングは必要ありません。

GDS は、あらかじめコンパイルされた単一のデータサービスです。このアプローチでは、コールバックメソッド (`rt_callbacks(1HA)`) の実装やリソースタイプ登録ファイル (`rt_reg(4)`) など、コンパイル済みのデータサービスやそのコンポーネントを変更することはできません。

コンパイル済みリソースタイプ

汎用データサービスのリソースタイプ `SUNW.gds` は、`SUNWscgds` パッケージに含まれています。このパッケージは、クラスタのインストール時に `scinstall` ユーティリティによってインストールされます (`scinstall (1M)` のマニュアルページを参照)。`SUNWscgds` パッケージには次のファイルが格納されています。

```
# pkgchk -v SUNWscgds

/opt/SUNWscgds
/opt/SUNWscgds/bin
/opt/SUNWscgds/bin/gds_monitor_check
/opt/SUNWscgds/bin/gds_monitor_start
/opt/SUNWscgds/bin/gds_monitor_stop
/opt/SUNWscgds/bin/gds_probe
/opt/SUNWscgds/bin/gds_svc_start
/opt/SUNWscgds/bin/gds_svc_stop
/opt/SUNWscgds/bin/gds_update
/opt/SUNWscgds/bin/gds_validate
/opt/SUNWscgds/etc
/opt/SUNWscgds/etc/SUNW.gds
```

GDS を使用することの利点と欠点

GDS には、`SunPlex Agent Builder` が生成するソースコードモデル (`scdscreate (1HA)` のマニュアルページを参照) や標準的な `Sun Cluster` 管理コマンドを使用するのに比べ、次の利点があります。

- GDS は使い易いデータサービスです。
- GDS とそのメソッドはコンパイル済みであるため、変更できません。
- `SunPlex Agent Builder` を使って、アプリケーション用のスクリプトを生成できます。これらのスクリプトは、複数のクラスタ間で再利用できる `Solaris` パッケージに含まれています。

GDS には多くの利点がありますが、GDS 機構の使用が適さない場合もあります。次のような場合は、GDS を使用すべきではありません。

- コンパイル済みリソースタイプを使用する場合よりも高度な制御が必要な場合。たとえば拡張プロパティを追加する場合や、デフォルト値を変更する場合など
- 特別な機能を追加するためにソースコードを変更する必要がある場合

GDS を使用するサービスの作成方法

GDS を使用するサービスの作成方法は 2 通りあります。

- `SunPlex Agent Builder` を使用
- 標準的な `Sun Cluster` 管理コマンドを使用

GDS と SunPlex Agent Builder

SunPlex Agent Builder を使用します。生成するソースコードのタイプとして GDS を選択します。特定のアプリケーションのリソースを設定するスクリプト群を生成するためにユーザーの入力が必要です。

GDS と標準的な Sun Cluster 管理コマンド

この方法では、SUNWscgds に含まれているコンパイル済みデータサービスコードを使用します。システム管理者は、標準的な Sun Cluster 管理コマンドを使ってリソースの作成と構成を行う必要があります。詳細は、scrgadm(1M) と scswitch(1M) のマニュアルページを参照してください。

GDS ベースのサービスを作成する方法の選択

199 ページの「Sun Cluster 管理コマンドを使って GDS ベースの高可用性サービスを作成する方法」や 200 ページの「Sun Cluster 管理コマンドを使って GDS ベースのスケラブルサービスを作成する方法」の手順からわかるように、適切な scrgadm や scswitch コマンドを実行するためには、かなりの分量の入力を行う必要があります。

GDS と SunPlex Agent Builder を使用する方法では、この処理が簡単になります。この方法では、生成されるスクリプトがユーザーに代わって scrgadm と scswitch コマンドを出力するからです。

GDS によるイベントのロギング

GDS から、GDS を起動するスクリプトに渡される有益な情報をロギングできます。この情報は、起動、検証、停止の各メソッドやプロパティ変数などからなります。この情報を使ってスクリプトの問題やエラーを診断したり、この情報を他の目的に利用したりできます。

GDS でロギングすべきメッセージのレベル (つまり、タイプ) の指定には、Log_level プロパティを使用します (詳細は 193 ページの「Log_level プロパティ」を参照)。プロパティの値には NONE、INFO、または ERR を指定できます。

GDS ログファイル

次の 2 つの GDS ログファイルが、`/var/cluster/logs/DS/resource_group_name/resource_name` ディレクトリに置かれます。

- `start_stop_log.txt` には、リソース起動メソッドや停止メソッドによってロギングされるメッセージが書き込まれます。

- probe_log.txt には、リソースモニターによってロギングされるメッセージが書き込まれます。

start_stop_log.txt には、次のようなメッセージが書き込まれます。

```
10/20/2004 12:38:05 phys-node-1 START-INFO> Start succeeded. [/home/brianx/sc/start_cmd]
10/20/2004 12:42:11 phys-node-1 STOP-INFO> Successfully stopped the application
```

probe_log.txt には、次のようなメッセージが書き込まれます。

```
10/20/2004 12:38:15 phys-node-1 PROBE-INFO> The GDS monitor (gds_probe) has been started
10/20/2004 12:39:15 phys-node-1 PROBE-INFO> The probe result is 0
10/20/2004 12:40:15 phys-node-1 PROBE-INFO> The probe result is 0
10/20/2004 12:41:15 phys-node-1 PROBE-INFO> The probe result is 0
```

必須の GDS プロパティ

アプリケーションが非ネットワーク対応の場合は、Start_command 拡張プロパティと Port_list プロパティを両方とも指定する必要があります。アプリケーションがネットワーク対応の場合は、Port_list プロパティだけが必須です。

Start_command 拡張プロパティ

Start_command 拡張プロパティに指定される起動コマンドは、アプリケーションの起動を行います。このコマンドは、引数を備えた UNIX コマンドでなければなりません。コマンドは、アプリケーションを起動するシェルに直接渡されます。

Port_list プロパティ

Port_list プロパティは、アプリケーションが待機するポート群を指定したものです。Port_list プロパティは、SunPlex Agent Builder によって生成される start スクリプトか、scrgadm コマンド (標準的な Sun Cluster 管理コマンドを使用する場合) に指定されていなければなりません。

任意の GDS プロパティ

次のプロパティは任意の GDS プロパティです。

- Network_resources_used
- Stop_command (拡張プロパティ)
- Probe_command (拡張プロパティ)
- Start_timeout
- Stop_timeout
- Probe_timeout (拡張プロパティ)

- Child_mon_level (標準的な管理コマンドだけで使用される拡張プロパティ)
- Failover_enabled (拡張プロパティ)
- Stop_signal (拡張プロパティ)
- Log_level (拡張プロパティ)

Network_resources_used プロパティ

このプロパティのデフォルト値は Null です。アプリケーションが1つ以上の特定のアドレスにバインドする必要がある場合は、このプロパティを指定する必要があります。このプロパティを省略するか、このプロパティが Null の場合、アプリケーションはすべてのアドレスに対して待機するものとみなされます。

GDS リソースを作成する場合は、LogicalHostname か SharedAddress リソースがあらかじめ構成されていなければなりません。LogicalHostname または SharedAddress リソースの構成方法については、『Sun Cluster データサービスの計画と管理 (Solaris OS 版)』を参照してください。

値を指定する場合は、1つまたは複数のリソース名を指定します。個々のリソース名には、1つ以上の LogicalHostname か 1つ以上の SharedAddress を指定できます。詳細は、r_properties(5)のマニュアルページを参照してください。

Stop_command プロパティ

stop コマンドは、アプリケーションを停止し、アプリケーションが完全に停止した後で終了します。このコマンドは、アプリケーションを停止するシェルに直接渡される完全な UNIX コマンドでなければなりません。

Stop_command 拡張プロパティが指定されていると、GDS 停止メソッドは、停止タイムアウトの 80% を指定して停止コマンドを起動します。さらに、GDS 停止メソッドは、停止コマンドの起動結果がどうであれ、停止タイムアウトの 15% を指定して SIGKILL を送信します。タイムアウトの残り 5% は、処理のオーバーヘッドのために使用されます。

stop コマンドが省略されていると、GDS は、Stop_signal に指定されたシグナルを使ってアプリケーションを停止します。

Probe_command プロパティ

probe コマンドは、特定のアプリケーションの状態を周期的にチェックします。このコマンドは、引数を備えた UNIX コマンドでなければなりません。コマンドは、アプリケーションを検証するシェルに直接渡されます。アプリケーションの状態が正常であれば、検証コマンドは終了ステータスとして 0 を返します。

検証コマンドの終了ステータスは、アプリケーションの障害の重大度を判断するために使用されます。終了ステータス (検証ステータス) は、0 (正常) から 100 (全面的な障害) までの整数でなければなりません。ただし、検証ステータスは、特殊な値として 201 をとることがあります。この場合、アプリケーションは、

Failover_enabled が FALSE に設定されている場合を除き、直ちにフェイルオーバーされます。GDS プローブアルゴリズム (scds_fm_action(3HA) のマニュアルページを参照) は、この probe ステータスを使って、アプリケーションをローカルに再起動するか別のノードにフェイルオーバーするかを決定します。終了ステータス 201 なら、アプリケーションは直ちにフェイルオーバーされます。

検証コマンドを省略すると、GDS はそれ自身の簡単な検証を行います。つまり、Network_resources_used プロパティや scds_get_netaddr_list の出力から得られる一連の IP アドレスを使ってアプリケーションに接続します (scds_get_netaddr_list(3HA) のマニュアルページを参照)。この検証では、接続に成功すると、接続を直ちに切り離します。接続と切り離しが両方とも正常なら、アプリケーションは正常に動作しているものとみなされます。

注 - GDS 提供の検証は、全機能を備えたアプリケーション固有の検証を代替するものではありません。

Start_timeout プロパティ

このプロパティでは、起動コマンドの起動タイムアウトを指定します。詳細は、190 ページの「Start_command 拡張プロパティ」を参照してください。Start_timeout のデフォルトは 300 秒です。

Stop_timeout プロパティ

このプロパティでは、停止コマンドの停止タイムアウトを指定します。詳細は、191 ページの「Stop_command プロパティ」を参照してください。Stop_timeout のデフォルトは 300 秒です。

Probe_timeout プロパティ

このプロパティでは、検証コマンドのタイムアウト値を指定します。詳細は、191 ページの「Probe_command プロパティ」を参照してください。Probe_timeout のデフォルトは 30 秒です。

Child_mon_level プロパティ

注 - 標準的な Sun Cluster 管理コマンドを使用する場合は、このオプションを使用できます。SunPlex Agent Builder を使用する場合は、このオプションを使用できません。

このプロパティを使用すれば、Process Monitor Facility (PMF) を通して監視するプロセスを制御できます。このプロパティは、フォークされた子プロセスをどのようなレベルで監視するかを表します。このプロパティは、pmfadm コマンドの `-c` 引数と同等の働きをします。詳細は、pmfadm(1M) のマニュアルページを参照してください。

このプロパティを省略するか、このプロパティにデフォルト値の `-1` を指定することは、pmfadm コマンドで `--c` オプションを省略するのと同じ効果があります。つまり、すべての子プロセスとその子孫プロセスが監視されます。

Failover_enabled プロパティ

ブール値のこの拡張プロパティでは、リソースのフェイルオーバー動作を制御します。この拡張プロパティに `true` を設定すると、アプリケーションは、再起動回数が `retry_interval` 秒間に `retry_count` を超えるとフェイルオーバーされます。

このプロパティに `false` を設定すると、再起動回数が `retry_interval` 秒間に `retry_count` を超えてもアプリケーションの再起動やフェイルオーバーは行なわれません。

このプロパティにより、アプリケーションリソースによるリソースグループのフェイルオーバーを防ぐことができます。このプロパティのデフォルトの値は `true` です。

Stop_signal プロパティ

GDS は、整数値のこの拡張プロパティを使って、PMF によるアプリケーションの停止に使用するシグナルを判別します。指定可能な整数値については、`signal` (3HEAD) のマニュアルページを参照してください。デフォルト値は `15` (SIGTERM) です。

Log_level プロパティ

このプロパティでは、GDS でロギングする診断メッセージのレベル (タイプ) を指定します。このプロパティには、`NONE`、`INFO`、または `ERR` を指定できます。`NONE` を指定すると、診断メッセージはロギングされません。`INFO` を指定すると、情報メッセージだけがロギングされます。`ERR` を指定すると、エラーメッセージだけがロギングされます。デフォルトでは、診断メッセージはロギングされません (`NONE`)。

SunPlex Agent Builder を使って GDS ベースのサービスを作成

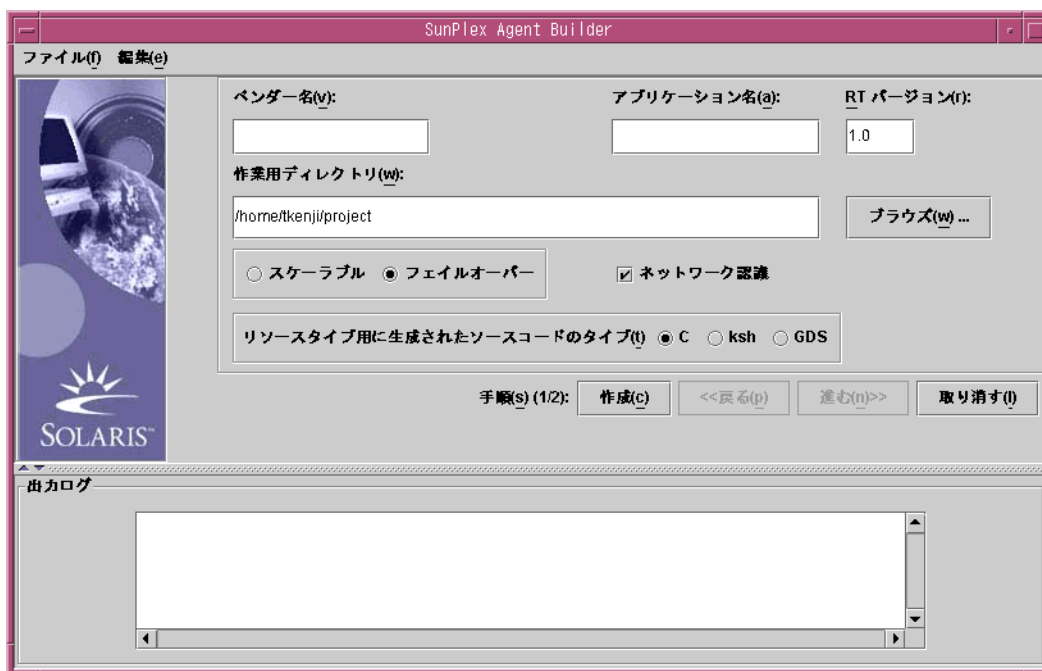
SunPlex Agent Builder を使って、GDS を使用するサービスを作成できます。SunPlex Agent Builder の詳細については、[第 9 章](#)を参照してください。

スクリプトの作成と構成

▼ SunPlex Agent Builder を起動し、スクリプトを作成する方法

1. スーパーユーザーになるか、同等の役割になります。
2. **SunPlex Agent Builder** を起動します。

```
# /usr/cluster/bin/scdsbuilder
```
3. 「**SunPlex Agent Builder Create**」画面が表示されます。



4. 「ベンダー名」に値を入力します。
5. 「アプリケーション名」に値を入力します。

注 - 「ベンダー名」と「アプリケーション名」の合計は9文字以内でなければなりません。この組み合わせは、スクリプトのパッケージ名として使用されます。

6. 作業ディレクトリに移動します。
このパスを入力する代わりに、「ブラウズ」プルダウンメニューを使ってディレクトリを選択することもできます。
7. データサービスがスケラブルなのかフェイルオーバーなのかを選択します。
GDSを作成するときには「ネットワーク認識」がデフォルトですので、これを選択する必要はありません。
8. 「GDS」を選択します。
9. (省略可能) 「RT バージョン」フィールドにデフォルト値が表示されるので、これを変更します。

注 - 「RT バージョン」フィールドで次の文字を使用することはできません。空白文字、タブ、スラッシュ (/)、逆スラッシュ (\)、アスタリスク (*)、疑問符 (?)、コンマ (,), セミコロン (;)、左角かっこ ((), 右角かっこ ()).

10. 「Create」をクリックしてください。

Agent Builder がスクリプトを作成します。サービスの作成結果が「Output Log」ウィンドウに表示されます。



「Create」がグレイ表示されます。これで、スクリプトの構成を始めることができます。

11. 「進む」をクリックする。
「構成」画面が表示されます。

▼ スクリプトを構成する方法

スクリプトの作成が終わったら、新しいサービスを構成する必要があります。

1. 起動コマンドの場所を指定するか、「ブラウズ」をクリックして起動コマンドの場所を選択します。

プロパティ変数を指定できます。プロパティ変数については、171 ページの「プロパティ変数」を参照してください。

2. (省略可能) 停止コマンドを指定するか、「ブラウズ」をクリックして停止コマンドを選択します。

プロパティ変数を指定できます。プロパティ変数については、171 ページの「プロパティ変数」を参照してください。

3. (省略可能) 検証コマンドを指定するか、「ブラウズ」をクリックして検証コマンドを選択します。

プロパティ変数を指定できます。プロパティ変数については、171 ページの「プロパティ変数」を参照してください。

4. (省略可能) 起動、停止、検証コマンドのタイムアウト値を指定します。

5. 「構成」をクリックします。

Agent Builder が起動し、スクリプトを構成します。

注 - Agent Builder は、ベンダー名とアプリケーション名を連結してパッケージ名を作成します。

スクリプトのパッケージが作成され、次のディレクトリに置かれます。

```
working-dir/vendor_nameapplication/pkg
```

たとえば、/export/wdir/NETapp/pkg のようになります。

6. スーパーユーザーで、完成したパッケージをクラスタのすべてのノードにインストールします。

```
# cd /export/wdir/NETapp/pkg
# pkgadd -d . NETapp
```

7. **pkgadd** によって以下のファイルがインストールされます。

```
/opt/NETapp
/opt/NETapp/README.app
/opt/NETapp/man
/opt/NETapp/man/man1m
/opt/NETapp/man/man1m/removeapp.1m
/opt/NETapp/man/man1m/startapp.1m
/opt/NETapp/man/man1m/stopapp.1m
/opt/NETapp/man/man1m/app_config.1m
/opt/NETapp/util
/opt/NETapp/util/removeapp
/opt/NETapp/util/startapp
/opt/NETapp/util/stopapp
/opt/NETapp/util/app_config
```

注 - マニュアルページとスクリプト名は、上で入力したアプリケーション名の前にスクリプト名を付けたものです。たとえば、startapp のようになります。

マニュアルページを表示するには、マニュアルページへのパスを指定する必要があります。たとえば、startapp(1M) のマニュアルページを表示する場合は、次のように指定します。

```
# man -M /opt/NETapp/man startapp
```

8. クラスタのいずれかのノードでリソースを構成し、アプリケーションを起動します。

```
# /opt/NETapp/util/startapp -h logicalhostname -p port_and_protocol_list
```

startapp スクリプトの引数は、リソースのタイプがフェイルオーバーかスケラブルかで異なります。カスタマイズしたマニュアルページを検査するか、startapp スクリプトを引数なしで実行して引数リストを表示してください。

```
# /opt/NETapp/util/startapp
```

The resource name of LogicalHostname or SharedAddress must be specified. For failover services:

```
Usage: startapp -h logicalhostname
        -p port_and_protocol_list
        [-n ipmpgroup_adapter_list]
```

For scalable services:

```
Usage: startapp -h shared_address_name
        -p port_and_protocol_list
        [-l load_balancing_policy]
        [-n ipmpgroup/adapter_list]
        [-w load_balancing_weights]
```

SunPlex Agent Builder の出力

SunPlex Agent Builder は 3 つの起動スクリプトと、パッケージ作成時の入力内容に基づく構成ファイルを生成します。構成ファイルには、リソースグループとリソースタイプの名前が指定されます。

作成されるスクリプトは次のとおりです。

- 起動スクリプト:リソースを構成し、RGM 制御のもとにあるアプリケーションを起動します。
- 停止スクリプト:アプリケーションを停止し、リソースやリソースグループを停止します。
- 削除スクリプト:起動スクリプトによって作成されたリソースやリソースグループを削除します。

これらのスクリプトのインタフェースや動作は、SunPlex Agent Builder によって非 GDS ベースのエージェント用に作成されるユーティリティスクリプトのものと同じです。これらのスクリプトは、Solaris パッケージに含まれており、複数のクラスタで再利用できます。

構成ファイルをカスタマイズすれば、リソースグループの名前など、一般には `scrgadm` コマンドへの入力として指定されるパラメータを独自に設定できます。スクリプトをカスタマイズしないと、SunPlex Agent Builder が `scrgadm` のパラメータに対しデフォルト値を設定します。

標準的な Sun Cluster 管理コマンドを使って GDS ベースのサービスを作成

この節では、GDS のパラメータをどのように入力するかについて説明します。GDS の使用や管理には、既存の Sun Cluster 管理コマンド (`scrgadm` や `scswitch` など) を使用します。

スクリプトに適切な機能が含まれている場合は、この節で述べる低位レベルの管理コマンドを入力する必要はありません。ただし、GDS ベースのリソースをより細かく制御したい場合は、低位レベルの管理コマンドを使用できます。これらのコマンドはスクリプトによって実行されます。

▼ Sun Cluster 管理コマンドを使って GDS ベースの高可用性サービスを作成する方法

1. スーパーユーザーになるか、同等の役割になります。

2. リソースタイプ `SUNW.gds` を登録します。

```
# scrgadm -a -t SUNW.gds
```

3. **LogicalHostname** リソースとフェイルオーバーサービス自体を含むリソースグループを作成します。

```
# scrgadm -a -g haapp_rg
```

4. **LogicalHostname** リソースのリソースを作成します。

```
# scrgadm -a -L -g haapp_rs -l hhead
```

5. フェイルオーバーサービス自体のリソースを作成します。

```
# scrgadm -a -j haapp_rs -g haapp_rg -t SUNW.gds \  
-y Scalable=false -y Start_timeout=120 \  
-y Stop_timeout=120 -x Probe_timeout=120 \  
-
```

```

-y Port_list="2222/tcp" \
-x Start_command="/export/ha/appctl/start" \
-x Stop_command="/export/ha/appctl/stop" \
-x Probe_command="/export/app/bin/probe" \
-x Child_mon_level=0 -y Network_resources_used=hhead \
-x Failover_enabled=true -x Stop_signal=9

```

- リソースグループ **haapp_rg** をオンラインにします。

```
# scswitch -Z -g haapp_rg
```

▼ Sun Cluster 管理コマンドを使って GDS ベースのスケラブルサービスを作成する方法

- スーパーユーザーになるか、同等の役割になります。
- リソースタイプ **SUNW.gds** を登録します。

```
# scrgadm -a -t SUNW.gds
```

- SharedAddress** リソースのリソースグループを作成します。

```
# scrgadm -a -g sa_rg
```

- SharedAddress** リソースを **sa_rg** に作成します。

```
# scrgadm -a -S -g sa_rg -l hhead
```

- スケラブルサービスのリソースグループを作成します。

```
# scrgadm -a -g app_rg -y Maximum primaries=2 \
-y Desired primaries=2 -y RG_dependencies=sa_rg
```

- スケラブルサービス自体のリソースグループを作成します。

```
# scrgadm -a -j app_rs -g app_rg -t SUNW.gds \
-y Scalable=true -y Start_timeout=120 \
-y Stop_timeout=120 -x Probe_timeout=120 \
-y Port_list="2222/tcp" \
-x Start_command="/export/app/bin/start" \
-x Stop_command="/export/app/bin/stop" \
-x Probe_command="/export/app/bin/probe" \
-x Child_mon_level=0 -y Network_resource_used=hhead \
-x Failover_enabled=true -x Stop_signal=9
```

- ネットワークリソースを含むリソースグループをオンラインにします。

```
# scswitch -Z -g sa_rg
```

- リソースグループ **app_rg** をオンラインにします。

```
# scswitch -Z -g app_rg
```

SunPlex Agent Builder のコマンド行インタフェース

SunPlex Agent Builder のコマンド行インタフェースでも、グラフィカルユーザーインタフェースで使えるのと同じ機能を使用できます。コマンド行インタフェースは `scdscreate` と `scdsconfig` コマンドからなります。詳細は、`scdscreate(1HA)` と `scdsconfig(1HA)` のマニュアルページを参照してください。

▼ コマンド行バージョンの Agent Builder を使用して GDS ベースのサービスを作成する方法

この節では、194 ページの「SunPlex Agent Builder を使って GDS ベースのサービスを作成」と同じ手順を、コマンド行インタフェースを使ってどのように行うかについて説明します。

1. スーパーユーザーになるか、同等の役割になります。

2. サービスを作成します。

ファイルオーバーサービスの場合:

```
# scdscreate -g -V NET -T app -d /export/wdir
```

スケーラブルサービスの場合:

```
# scdscreate -g -s -V NET -T app -d /export/wdir
```

注 - `-d` パラメータは任意です。このパラメータを省略すると、現在のディレクトリが作業ディレクトリになります。

3. サービスを構成します。

```
# scdsconfig -s "/export/app/bin/start" -t "/export/app/bin/stop" \  
-m "/export/app/bin/probe" -d /export/wdir
```

プロパティ変数を指定できます。プロパティ変数については、171 ページの「プロパティ変数」を参照してください。

注 - `start` コマンドだけが必須で、他のパラメータはすべて任意です。

4. 完成したパッケージをクラスタのすべてのノードにインストールします。

```
# cd /export/wdir/NETapp/pkg  
# pkgadd -d . NETapp
```

5. **pkgadd** によって以下のファイルがインストールされます。

```
/opt/NETapp  
/opt/NETapp/README.app  
/opt/NETapp/man  
/opt/NETapp/man/man1m  
/opt/NETapp/man/man1m/removeapp.1m  
/opt/NETapp/man/man1m/startapp.1m  
/opt/NETapp/man/man1m/stopapp.1m  
/opt/NETapp/man/man1m/app_config.1m  
/opt/NETapp/util  
/opt/NETapp/util/removeapp  
/opt/NETapp/util/startapp  
/opt/NETapp/util/stopapp  
/opt/NETapp/util/app_config
```

注 - マニュアルページとスクリプト名は、上で入力したアプリケーション名の前にスクリプト名を付けたものです。たとえば、`startapp` のようになります。

マニュアルページを表示するには、マニュアルページへのパスを指定する必要があります。たとえば、`startapp(1M)` のマニュアルページを表示する場合は、次のように指定します。

```
# man -M /opt/NETapp/man startapp
```

6. クラスタのいずれかのノードでリソースを構成し、アプリケーションを起動します。

```
# /opt/NETapp/util/startapp -h logicalhostname -p port_and_protocol_list
```

`startapp` スクリプトの引数は、リソースのタイプがフェイルオーバーかスケラブルかで異なります。カスタマイズしたマニュアルページを検査するか、`startapp` スクリプトを引数なしで実行して引数リストを表示してください。

```
# /opt/NETapp/util/startapp  
The resource name of LogicalHostname or SharedAddress must be  
specified.  
For failover services:  
Usage: startapp -h logicalhostname  
        -p port_and_protocol_list  
        [-n ipmpgroup/adapter_list]  
For scalable services:  
Usage: startapp -h shared_address_name  
        -p port_and_protocol_list  
        [-l load_balancing_policy]  
        [-n ipmpgroup/adapter_list]  
        [-w load_balancing_weights]
```

第 11 章

データサービス開発ライブラリのリファレンス

この章では、データサービス開発ライブラリ (Data Service Development Library: DSDL) の API 関数の一覧を示し、概要を述べます。個々の DSDL 関数の詳細については、そのマニュアルページ (3HA) を参照してください。DSDL は C 言語用のインタフェースだけを定義します。スクリプトベースの DSDL インタフェースはありません。

DSDL には、次のタイプの関数があります。

- 203 ページの「汎用関数」
- 205 ページの「プロパティ関数」
- 205 ページの「ネットワークリソースアクセス関数」
- 206 ページの「PMF 関数」
- 207 ページの「障害監視関数」
- 207 ページの「ユーティリティ関数」

DSDL 関数

この節では、DSDL 関数の各カテゴリを簡単に説明します。DSDL 関数を定義するリファレンスについては、個々のマニュアルページ (3HA) を参照してください。

汎用関数

このカテゴリの関数は、さまざまな機能を提供します。これらの関数では、次の操作を行うことができます。

- DSDL 環境を初期化します。
- リソース、リソースタイプ、およびリソースグループの名前、ならびに、拡張プロパティの値を取得します。

- リソースグループをフェイルオーバーおよび再起動し、リソースを再起動します。
- エラー文字列をエラーメッセージに変換します。
- タイムアウトを適用してコマンドを実行します。

次の関数は、呼び出しメソッドを初期化します。

- `scds_initialize(3HA)` – リソースを割り当て、DSDL 環境を初期化します。
- `scds_close(3HA)` – `scds_initialize` によって割り当てられた空きリソース

次の関数は、リソース、リソースタイプ、リソースグループ、および拡張プロパティについての情報を取得します。

- `scds_get_resource_name(3HA)` – 呼び出しメソッド用のリソースの名前を取得します。
- `scds_get_resource_type_name(3HA)` – 呼び出しメソッド用のリソースタイプの名前を取得します。
- `scds_get_resource_group_name(3HA)` – 呼び出しメソッド用のリソースグループの名前を取得します。
- `scds_get_ext_property(3HA)` – 指定された拡張プロパティの値を取得します。
- `scds_free_ext_property(3HA)` – `scds_get_ext_property` で割り当てられたメモリを解放します。

次の関数は、リソースが使用している SUNW.HAStoragePlus リソースについての状態情報を取得します。

- `scds_hasp_check(3HA)` – リソースによって使用される SUNW.HAStoragePlus (5) リソースの状態の情報を取得します。当該リソース用に定義されている `Resource_dependencies` または `Resource_dependencies_weak` のシステム属性を使用することによって、当該リソースが依存しているすべての SUNW.HAStoragePlus リソース状態 (オンラインであるか、オンラインでないか) についての情報が得られます。

次の関数は、リソースまたはリソースグループをフェイルオーバーまたは再起動します。

- `scds_failover_rg(3HA)` – リソースグループをフェイルオーバーします。
- `scds_restart_rg(3HA)` – リソースグループを再起動します。
- `scds_restart_resource(3HA)` – リソースを再起動します。

次の関数は、タイムアウトを適用してコマンドを実行し、エラーコードをエラーメッセージに変換します。

- `scds_timerun(3HA)` – タイムアウト値のもとでコマンドを実行します。
- `scds_error_string(3HA)` – エラーコードをエラー文字列に変換します。

プロパティ関数

このカテゴリの関数は、関連するリソース、リソースグループ、およびリソースタイプ (よく使用される一部の拡張プロパティも含む) に固有なプロパティにアクセスするのに有用な API を提供します。DSDL は、`scds_initialize` を使用して、コマンド行引数を解析します。`scds_initialize(3HA)` 関数は、関連するリソース、リソースグループ、およびリソースタイプの様々なプロパティをキャッシュに入れます。

これらの関数の説明は `scds_property_functions(3HA)` にあります。次にさまざまなサポート関数について説明します。

- `scds_get_rs_property_name`
- `scds_get_rg_property_name`
- `scds_get_rt_property_name`
- `scds_get_ext_property_name`

ネットワークリソースアクセス関数

このカテゴリの関数は、リソースおよびリソースグループが使用するネットワークリソースを、取得、出力、および解放します。ここで説明する `scds_get_*` 関数は、RMAPI 関数を使用して `Network_resources_used` や `Port_list` などのプロパティを照会しなくても、ネットワークリソースを取得できる便利な方法を提供します。`scds_print_name()` 関数は、`scds_get_name()` 関数から戻されたデータ構造から値を出力します。`scds_free_name()` 関数は、`scds_get_name()` 関数が割り当てたメモリーを解放します。

ホスト名を扱う関数には次のものがあります。

- `scds_get_rg_hostnames(3HA)` – リソースグループ内のネットワークリソースによって使用されているホスト名のリストを取得します。
- `scds_get_rs_hostnames(3HA)` – リソースによって使用されているホスト名のリストを取得します。
- `scds_print_net_list(3HA)` – `scds_get_rg_hostnames()` または `scds_get_rs_hostnames()` から返されたホスト名リストの内容を印刷します。
- `scds_free_net_list(3HA)` – `scds_get_rg_hostnames()` または `scds_get_rs_hostnames()` によって割り当てられたメモリーを解放します。

ポートリストを扱う関数には次のものがあります。

- `scds_get_port_list(3HA)` – リソースによって使用されているポート/プロトコル・ペアのリストを取得します。
- `scds_print_port_list(3HA)` – `scds_get_port_list()` によって返されたポート/プロトコル・ペアのリストの内容を印刷します。
- `scds_free_port_list(3HA)` – `scds_get_port_list()` によって割り当てられたメモリーを解放します。

ネットワークアドレスを扱う関数には次のものがあります。

- `scds_get_netaddr_list` (3HA) – リソースによって使用されているネットワークアドレスのリストを取得します。
- `scds_print_netaddr_list` (3HA) – `scds_get_netaddr_list` によって返されたネットワークアドレスリストの内容を印刷します。
- `scds_free_netaddr_list` (3HA) – `scds_get_netaddr_list` によって割り当てられたメモリを解放します。

TCP 接続を使用する障害監視

このカテゴリの関数は、TCP ベースの監視を行います。通常、障害モニターはこれらの関数を使用して、サービスとの単純ソケット接続を確立し、サービスのデータを読み書きしてサービスの状態を確認した後、サービスとの接続を切断します。

次にさまざまなサポート関数について説明します。

- `scds_fm_tcp_connect` (3HA) – IPv4 アドレッシングだけを使用するプロセスへの TCP 接続を確立します。
- `scds_fm_net_connect` (3HA) – IPv4 か IPv6 アドレッシングのどちらかを使用するプロセスへの TCP 接続を確立します。
- `scds_fm_tcp_read` (3HA) – TCP 接続を使って、監視されているプロセスからデータを読み取ります。
- `scds_fm_tcp_write` (3HA) – TCP 接続を使って、監視されているプロセスにデータを書き込みます。
- `scds_simple_probe` (3HA) – プロセスへの TCP 接続を確立し、停止することによってプロセスを検証します。この関数は IPv4 アドレスだけを扱います。
- `scds_simple_net_probe` (3HA) – プロセスへの TCP 接続を確立し、停止することによってプロセスを検証します。この関数は IPv4 と IPv6 アドレスを扱います。
- `scds_fm_tcp_disconnect` (3HA) – IPv4 アドレッシングを使用し、かつ監視されているプロセスへの接続を停止します。
- `scds_fm_net_disconnect` (3HA) – IPv4 か IPv6 アドレッシングのどちらかを使用し、かつ監視されているプロセスへの接続を停止します。

PMF 関数

このカテゴリの関数は、PMF 機能をカプセル化します。PMF 経由の監視における DSDL モデルは、`pmfadm` (1M) に対して、暗黙のタグ値を作成および使用します。また、PMF 機能は、`Restart_interval`、`Retry_count`、および `action_script` 用の暗黙値も使用します (`pmfadm` の `-t`、`-n`、および `-a` オプション)。最も重要な点は、DSDL が、PMF によって検出されたプロセス停止履歴を、障害モニターによって検出されたアプリケーション障害履歴に結びつけ、再起動またはフェイルオーバーのどちらを行うかを決定することです。

次にさまざまなサポート関数について説明します。

- `scds_pmf_get_status(3HA)` – 指定するインスタンスが PMF 制御のもとで監視されているかどうかを判別します。
- `scds_pmf_restart_fm(3HA)` – PMF を使って障害モニターを再起動します。
- `scds_pmf_signal(3HA)` – PMF 制御のもとで動作するプロセスツリーに、指定するシグナルを送信します。
- `scds_pmf_start(3HA)` – 指定するプログラム (障害モニターを含む) を PMF 制御のもとで実行します。
- `scds_pmf_stop(3HA)` – PMF 制御のもとで動作しているプロセスを停止します。
- `scds_pmf_stop_monitoring(3HA)` – PMF 制御のもとで動作しているプロセスの監視を停止します。

障害監視関数

このカテゴリの関数は、障害履歴を保持し、その履歴を `Retry_count` および `Retry_interval` プロパティと関連付けて評価することにより、障害監視の事前定義モデルを提供します。

次にさまざまなサポート関数について説明します。

- `scds_fm_sleep(3HA)` – 障害モニター制御ソケットに関するメッセージを待ちます。
- `scds_fm_action(3HA)` – 検証の完了後にアクションをとります。
- `scds_fm_print_probes(3HA)` – 検証状態の情報をシステムログに書き込みます。

ユーティリティ関数

このカテゴリの関数は、メッセージやデバッグ用メッセージをシステムログに書き込みます。次にさまざまなサポート関数について説明します。

- `scds_syslog(3HA)` – メッセージをシステムログに書き込みます。
- `scds_syslog_debug(3HA)` – デバッグメッセージをシステムログに書き込みます。

第 12 章

CRNP

この章では、Cluster Reconfiguration Notification Protocol (CRNP) について説明します。CRNP を使用することで、フェイルオーバー用のアプリケーションや拡張性のあるアプリケーションを「クラスタ対応」として設定できます。具体的には、Sun Cluster 再構成イベントにアプリケーションを登録し、それらのイベントの後続の非同期通知を受け取ることができます。イベント通知の受信登録が可能なのは、クラスタの内部で動作するデータサービスと、クラスタの外部で動作するアプリケーションです。イベントは、クラスタ内のメンバーシップに変化があった場合と、リソースグループまたはリソースの状態に変化があった場合に生成されます。

注 - SUNW.Event のリソース型実装は、高可用性を備えた Sun Cluster の CRNP サービスを提供します。SUNW.Event のリソース型実装については、SUNW.Event (5) のマニュアルページを参照してください。

- 209 ページの「CRNP の概要」
- 212 ページの「CRNP が使用するメッセージのタイプ」
- 213 ページの「クライアントをサーバーに登録する方法」
- 215 ページの「クライアントに対するサーバーの応答方法」
- 217 ページの「サーバーがクライアントにイベントを配信する方法」
- 221 ページの「CRNP によるクライアントとサーバーの認証」
- 221 ページの「CRNP を使用する Java アプリケーションの作成」

CRNP の概要

CRNP は、クラスタ再構成イベントの生成、クラスタへの配信、それらのイベントを要求しているクライアントへの送信を行うメカニズムとデーモンを提供します。

クライアントとの通信を行うのは、cl_apid デーモンです。クラスタ再構成イベントの生成は、Sun Cluster Resource Group Manager (RGM) によって行われます。これらのデーモンは、syseventd (1M) を使用して各ローカルノードにイベントを転送します。cl_apid デーモンは、TCP/IP 上で XML (Extensible Markup Language) を使用して要求クライアントとの通信を行います。

次の図は、CRNP コンポーネント間のイベントの流れを簡単に示したものです。この図では、一方のクライアントはクラスタノード 2 で動作し、他方のクライアントはクラスタに属していないコンピュータ上で動作しています。

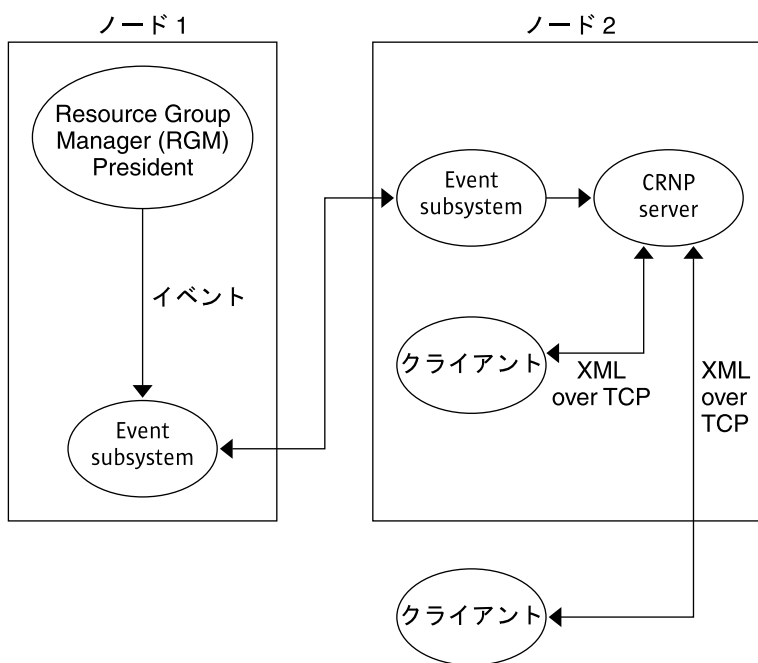


図 12-1 CRNP の動作

CRNP プロトコルの概要

CRNP は、標準の 7 層 OSI (Open System Interconnect) プロトコルスタックにおけるアプリケーション層、プレゼンテーション層、およびセッション層を定義します。トランスポート層は TCP でなければならず、ネットワーク層は IP でなければなりません。CRNP は、データリンク層および物理層とは無関係です。CRNP 内で交換されるアプリケーション層メッセージはすべて、XML 1.0 をベースとしたものです。

CRNP プロトコルのセマンティクス

クライアントは、サーバーへ登録メッセージ (SC_CALLBACK_RG) を送信することによって通信を開始します。この登録メッセージは、通知を受信したいイベントタイプと、イベントの配信先として使用できるポートを指定するものです。登録用接続のソース IP と指定ポートから、コールバックアドレスが構成されます。

クライアントが配信を希望しているイベントがクラスタ内で生成されるたびに、サーバーはこのコールバックアドレス (IP とポート) を持つクライアントと通信を行い、イベント (SC_EVENT) をクライアントに配信します。サーバーには、そのクラスタ内で稼動している高可用マシンが使用されます。サーバーは、クラスタの再起動後も維持されるストレージにクライアントの登録情報を格納します。

登録解除を行う場合、クライアントはサーバーに登録メッセージ (REMOVE_CLIENT_MESSAGE) が入った SC_CALLBACK_RG) を送信します。サーバーから SC_REPLY_MESSAGE) を受け取ったあとで、クライアントは接続を閉じることができます。

次の図は、クライアントとサーバー間の通信の流れを示します。

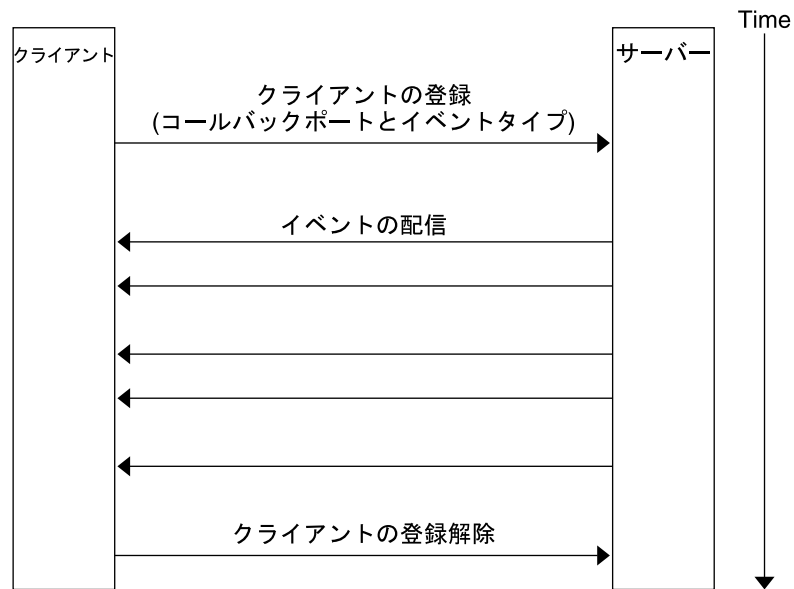


図 12-2 クライアントとサーバー間の通信の流れ

CRNP が使用するメッセージのタイプ

CRNP は、3 種類のメッセージを使用します。これらはすべて、次の表に示すように XML ベースのメッセージです。これらのメッセージタイプの内容と使用法の詳細は、この章で後述します。

メッセージのタイプ	説明
SC_CALLBACK_REG	<p>このメッセージには、4 つのフォーム、ADD_CLIENT、REMOVE_CLIENT、ADD_EVENTS、および REMOVE_EVENTS を指定できます。これらの各フォームには、次の情報が含まれます。</p> <ul style="list-style-type: none">■ プロトコルバージョン■ ASCII 形式で示されたコールバックポート (バイナリ形式ではない) <p>ADD_CLIENT、ADD_EVENTS、REMOVE_EVENTS の各フォームには制約のないイベントタイプリストも含まれ、それぞれに次の情報が含まれます。</p> <ul style="list-style-type: none">■ イベントクラス■ イベントサブクラス (省略可能)■ 名前と値がペアになったリスト (省略可能) <p>イベントクラスとイベントサブクラスにより一意の「イベントタイプ」が定義されます。SC_CALLBACK_REG のクラスを生成する DTD (ドキュメントタイプ定義) は、SC_CALLBACK_REG です。この DTD の詳細は、付録 F を参照してください。</p>
SC_EVENT	<p>このメッセージには次の情報が含まれます。</p> <ul style="list-style-type: none">■ プロトコルバージョン■ イベントクラス■ イベントサブクラス■ ベンダー■ パブリッシャー■ 名前と値のペアリスト (名前と値をペアにした 0 個以上のデータ構造)<ul style="list-style-type: none">■ 名前 (文字列)■ 値 (文字列または文字列配列) <p>SC_EVENT 内の値はタイプとしては分類されていません。SC_EVENT のクラスを生成する DTD (ドキュメントタイプ定義) は SC_EVENT です。この DTD の詳細は、付録 F を参照してください。</p>
SC_REPLY	<p>このメッセージには次の情報が含まれます。</p> <ul style="list-style-type: none">■ プロトコルバージョン■ エラーコード■ エラーメッセージ: <p>SC_REPLY のクラスを生成する DTD (ドキュメントタイプ定義) は SC_REPLY です。この DTD の詳細は、付録 F を参照してください。</p>

クライアントをサーバーに登録する方法

この節では、サーバーの設定、クライアントの識別、アプリケーション層とセッション層での情報送信、エラー状況などについて説明します。

管理者によるサーバー設定の前提

システム管理者は、汎用 IP アドレス (クラスタ内の特定のマシン専用でない IP アドレス) とポート番号を使用してサーバーを構成し、クライアントとなるマシンにこのネットワークアドレスを公開する必要があります。CRNP では、クライアントがこのサーバー名をどのように取得するかは定義されていません。管理者は、ネーミングサービスを使用することも (この場合、クライアントは動的にサーバーのネットワークアドレスを検出できる)、ネットワーク名を構成ファイルに追加してクライアントに読み取らせることもできます。サーバーは、クラスタ内でフェイルオーバーリソースタイプとして動作します。

サーバーによるクライアントの識別方法

各クライアントは、そのコールバックアドレス (IP アドレスとポート番号) で識別されます。ポートは `SC_CALLBACK_REG` メッセージで指定され、IP アドレスは登録用の TCP 接続から取得されます。CRNP は、同じコールバックアドレスを持つ後続の `SC_CALLBACK_REG` メッセージは同じクライアントから送信されたと想定します。これは、メッセージの送信元であるソースポートが異なる場合でも同様です。

クライアントとサーバー間での `SC_CALLBACK_REG` メッセージの受け渡し方法

クライアントは、サーバーの IP アドレスとポート番号に対して TCP 接続を開くことによって登録を開始します。TCP 接続が確立され書き込みの用意ができたところで、クライアントはその登録メッセージを送信する必要があります。この登録メッセージは正しい書式の `SC_CALLBACK_REG` メッセージでなければならず、メッセージの前後に余分なバイトを含めることはできません。

バイトがすべてストリームに書き込まれたあと、クライアントはサーバーから応答を受け取ることができるように接続をオープン状態に保つ必要があります。クライアントが不正な書式のメッセージを送信した場合、サーバーはそのクライアントを登録せず、クライアントに対してエラー応答を送信します。サーバーが応答を送信する前にクライアントがソケット接続を閉じた場合、サーバーはそのクライアントを正常なクライアントとして登録します。

クライアントは、いつでもサーバーと通信を行うことができます。サーバーと通信を行うごとに、クライアントは `SC_CALLBACK_REG` メッセージを送信する必要があります。書式が不正なメッセージ、順不同のメッセージ、無効なメッセージなどを受け取った場合、サーバーはクライアントにエラー応答を送信します。

クライアントは、それ自身が `ADD_CLIENT` メッセージを送信するまでは `ADD_EVENTS`、`REMOVE_EVENTS`、`REMOVE_CLIENT` メッセージを送信できません。また、`ADD_CLIENT` メッセージを送信しないかぎり `REMOVE_CLIENT` メッセージを送信できません。

クライアントが `ADD_CLIENT` メッセージを送信したが、そのクライアントがすでに登録されていたという場合は、サーバーがこのメッセージを黙認することがあります。このような場合、サーバーは報告なしに古いクライアント登録を削除し、2 つめの `ADD_CLIENT` メッセージに指定された新しいクライアント登録に置き換えます。

通常、クライアントはその起動時に `ADD_CLIENT` メッセージを送信することによって、サーバーに一度だけ登録を行います。また、登録の解除もサーバーに `REMOVE_CLIENT` メッセージを送信して一度だけ行います。しかし、CRNP はクライアントが必要に応じてイベントタイプリストを動的に変更できるだけの柔軟性を備えています。

SC_CALLBACK_REG メッセージの概念

`ADD_CLIENT`、`ADD_EVENTS`、および `REMOVE_EVENTS` メッセージには、それぞれイベントリストが含まれます。次の表は、CRNP が受け付けるイベントタイプを、必要となる名前と値のペアと共に示して説明しています。

クライアントが以下の作業のどちらか一方を行うと、

- まだ登録が行われていないイベントタイプを 1 つ以上指定する `REMOVE_EVENTS` メッセージを送信する
- 同じイベントタイプを 2 度登録する

サーバーはクライアントに通知することなくこれらのメッセージを無視します。

クラスとサブクラス	名前と値のペア	説明
<code>EC_Cluster</code>	必須: なし	クラスタメンバーシップの変更 (ノードの停止または結合) に関連するあらゆるイベントに登録する
<code>ESC_cluster_membership</code>	(省略可能) なし	
<code>EC_Cluster</code>	次の条件で 1 つ必要:	<code>rg_name</code> (リソースグループ名) のあらゆる状態変更イベントに登録する
<code>ESC_cluster_rg_state</code>	<code>rg_name</code>	
	値のタイプ: 文字列型 (省略可能) なし	

クラスとサブクラス	名前と値のペア	説明
EC_Cluster ESC_cluster_r_state	次の条件で1つ必要: r_name 値のタイプ: 文字列型 (省略可能) なし	r_name (リソース名) のあらゆる状態変更イベントに登録する
EC_Cluster なし	必須: なし (省略可能) なし	あらゆる Sun Cluster イベントに登録する

クライアントに対するサーバーの応答方法

登録を処理したあと、サーバーは SC_REPLY メッセージを送信します。この送信は、登録要求を行なったクライアントがオープン状態に維持している TCP 接続に対して行われます。このあとサーバーはこの接続を閉じます。クライアントは、サーバーから SC_REPLY メッセージを受信するまで TCP 接続をオープン状態に保つ必要があります。

クライアントは次のような作業を行います。

1. サーバーに対して TCP 接続を開きます。
2. 接続が「writeable (書き込み可能)」になるまで待機します。
3. SC_CALLBACK_REG メッセージ (このメッセージには ADD_CLIENT メッセージが入っている) を送信します。
4. SC_REPLY メッセージの到着を待機します。
5. SC_REPLY メッセージを受け取ります。
6. サーバーが接続を閉じたことを示すインジケータを受信します (ソケットから 0 バイトを読み取る)。
7. 接続を閉じます。

その後クライアントは以下の作業を行います。

1. サーバーに対して TCP 接続を開きます。
2. 接続が「writeable (書き込み可能)」になるまで待機します。
3. SC_CALLBACK_REG メッセージ (このメッセージには REMOVE_CLIENT メッセージが入っている) を送信します。
4. SC_REPLY メッセージの到着を待機します。
5. SC_REPLY メッセージを受け取ります。

6. サーバーが接続を閉じたことを示すインジケータを受信します (ソケットから 0 バイトを読み取る)。
7. 接続を閉じます。

クライアントから SC_CALLBACK_REG メッセージを受け取るたびに、サーバーは同じ接続に SC_REPLY メッセージを送信します。このメッセージは、処理が正常に完了したか失敗したかを示すものです。SC_REPLY メッセージの XML ドキュメントタイプ定義とこのメッセージ内で示されるエラーメッセージについては、[330 ページ](#)の「SC_REPLY XML DTD」を参照してください。

SC_REPLY メッセージの内容

SC_REPLY メッセージは、処理が正常に完了したか失敗したかを示します。このメッセージには、CRNP プロトコルメッセージのバージョン、ステータスコード、およびステータスコードの詳細を説明したステータスメッセージが含まれます。次の表は、ステータスコードの値を説明しています。

ステータスコード	説明
OK	メッセージは正常に処理されました。
RETRY	一時的なエラーのためにクライアントの登録はサーバーに拒否されました。クライアントは別のパラメータを使用して登録をもう一度試す必要があります。
LOW_RESOURCE	クラスタのリソースが少ないため、クライアントはあとでもう一度試すか、システム管理者にクラスタのリソースを増やしてもらう必要があります。
SYSTEM_ERROR	重大な問題が発生しました。クラスタのシステム管理者に連絡してください。
FAIL	承認の失敗などの問題が発生し、登録が失敗しました。
MALFORMED	XML 要求の形式が正しくないため解析が失敗しました。
INVALID	XML 要求が無効です (XML 仕様を満たしていない)。
VERSION_TOO_HIGH	メッセージのバージョンが高すぎて、メッセージを正常に処理できませんでした。
VERSION_TOO_LOW	メッセージのバージョンが低すぎて、メッセージを正常に処理できませんでした。

クライアントによるエラー状況の処理

通常、SC_CALLBACK_REG メッセージを送信するクライアントは登録の成功または失敗を知らせる応答を受け取ります。

しかし、クライアントが登録を試みる際にサーバーからの `SC_REPLY` メッセージの送信を妨げるエラーが発生することがあります。この場合、エラーが発生する前に登録が正常に完了することも、登録が失敗することも、あるいは登録処理が行われないまま終了することもあります。

サーバーはクラスタ上でフェイルオーバー (高可用) サーバーとして機能するため、このエラーがサービスの終了を意味するわけではありません。実際、サーバーは新しく登録されたクライアントに対してすぐにイベント送信を開始できます。

これらの状況を修復するには、クライアントは次の 2 つの作業を行う必要があります。

- `SC_REPLY` メッセージを待機している登録用接続にアプリケーションレベルのタイムアウトを強制します (このあと、登録を再試行する必要があります)。
- イベントコールバックの登録を行う前に、イベント配信用のコールバック IP アドレスとポート番号で待機を開始します。クライアントは、登録確認メッセージとイベント配信を同時に待機することになります。確認メッセージを受信する前にイベントを受信し始めた場合は、クライアントはそのまま登録接続を閉じる必要があります。

サーバーがクライアントにイベントを配信する方法

クラスタ内でイベントが生成されると、CRNP サーバーはそのタイプのイベントを要求したすべてのクライアントにイベントの配信を行います。この配信では、クライアントのコールバックアドレスに `SC_EVENT` メッセージが送信されます。各イベントの配信は、新たな TCP 接続で行われます。

クライアントが `ADD_CLIENT` メッセージまたは `ADD_EVENT` メッセージが入った `SC_CALLBACK_REG` メッセージを通してイベントタイプの配信登録を行うと、サーバーはただちにクライアントに対してそのタイプの最新イベントを送信します。続いてクライアントは、後続のイベントを送信するシステムの現在の状態を検出できます。

クライアントに対して TCP 接続を開始する際に、サーバーはその接続に `SC_EVENT` メッセージを 1 つだけ送信します。続いてサーバーは全二重通信を閉じます。

クライアントは次のような作業を行います。

1. サーバーが TCP 接続を開始するのを待機します。
2. サーバーからの着信接続を受け入れます。
3. `SC_EVENT` メッセージの到着を待機します。
4. `SC_EVENT` メッセージを読み取ります。

5. サーバーが接続を閉じたことを示すインジケータを受信します (ソケットから 0 バイトを読み取る)。
6. 接続を閉じます。

すべてのクライアントが登録を終了した時点で、それらのクライアントはイベント配信のための着信接続を受け入れるために常にコールバックアドレス (IP アドレスとポート番号) で待機する必要があります。

クライアントとの通信に失敗してイベントを配信できなかった場合、サーバーはユーザーが設定してある回数と周期に従ってイベントの配信を繰り返し試みます。それらの試行がすべて失敗に終わった場合、そのクライアントはサーバーのクライアントリストから削除されます。イベントをそれ以上受け取るためには、クライアントは `ADD_CLIENT` メッセージが入った `SC_CALLBACK_REG` メッセージを別途送信して登録をもう一度行う必要があります。

イベント配信の保証

クラスタ内では、クライアントごとに配信順序を守るという方法で、トータル的にイベント生成を順序付けます。たとえば、クラスタ内でイベント A の生成後イベント B が生成された場合、クライアント X はイベント A を受け取ってからイベント B を受け取ります。しかし、全クライアントに対するイベント配信の全体的な順序付けは保持されません。つまり、クライアント Y はクライアント X がイベント A を受け取る前にイベント A と B の両方を受け取る可能性があります。この方法では、低速のクライアントのために全クライアントへの配信が停滞するということはありません。

サーバーが配信するイベントはすべて (サブクラス用の最初のイベントとサーバーエラーのあとに発生するイベントを除く)、クラスタが生成する実際のイベントに応答して発生します。ただし、クラスタで生成されるイベントを見逃すようなエラーが発生する場合は、サーバーは各イベントタイプの現在のシステム状態を示すイベントをそれらのイベントタイプごとに生成します。各イベントは、そのイベントタイプの配信登録を行なったクライアントに送信されます。

イベント配信は、「1 回以上」というセマンティクスに従って行われます。つまり、サーバーは 1 台のクライアントに対して同じイベントを複数回送信できます。この許可は、サーバーが一時的に停止して復帰した際に、クライアントが最新の情報を受け取ったかどうかをサーバーが判断できないという場合に不可欠なものです。

SC_EVENT メッセージの内容

`SC_EVENT` メッセージには、クラスタ内で生成されて `SC_EVENT XML` メッセージ形式に合うように変換された実際のメッセージが入っています。次の表は、CRNP が配信するイベントタイプ (名前と値のペア、パブリッシャー、ベンダーなど) を説明したものです。

クラスとサブクラス	パブリッシャーとベンダー	名前と値のペア	記入欄
EC_Cluster	パブリッシャー: rgm	名前: node_list	state_list の配列要素は、node_list の配列要素と同期をとるように配置されます。つまり、node_list 配列内で最初に出現しているノードの状態は、state_list 配列の先頭に示されます。
ESC_cluster_membership	ベンダー: SUNW	値のタイプ: 文字配列 名前: state_list 値のタイプ: 文字配列	

state_list には、ASCII 形式の数字だけが入っています。各数字は、クラスタにおけるそのノードの現在のインカーネーション番号を示します。この番号が前のメッセージで受信した番号と同じである場合、ノードとクラスタの関係は変化していません (離脱、結合、または再結合が行われていない)。インカーネーション番号が -1 の場合、ノードはクラスタのメンバーではありません。インカーネーション番号が負の値以外の数字である場合、ノードはクラスタのメンバーです。

ev_ で始まるほかの名前や、それらの名前に関連した値が存在する場合がありますが、クライアントによる使用を意図したものではありません。

クラスとサブクラス	パブリッシャーとベンダー	名前と値のペア	記入欄
EC_Cluster ESC_cluster_rg_state	パブリッシャー: rgm ベンダー: SUNW	名前: rg_name 値のタイプ: 文字列型 名前: node_list 値のタイプ: 文字配列 名前: state_list 値のタイプ: 文字配列	state_list の配列要素は、node_list の配列要素と同期をとるように配置されます。つまり、node_list 配列内で最初に出現しているノードの状態は、state_list 配列の先頭に示されます。 state_list には、リソースグループの状態を示す文字列が入っています。有効な値は、scha_cmds (1HA) コマンドで取得できる値です。 ev_ で始まるほかの名前や、それらの名前に関連した値が存在する場合がありますが、クライアントによる使用を意図したものではありません。
EC_Cluster ESC_cluster_r_state	パブリッシャー: rgm ベンダー: SUNW	次の条件で3つ必要: 名前: r_name 値のタイプ: 文字列型 名前: node_list 値のタイプ: 文字配列 名前: state_list 値のタイプ: 文字配列	state_list の配列要素は、node_list の配列要素と同期をとるように配置されます。つまり、node_list 配列内で最初に出現しているノードの状態は、state_list 配列の先頭に示されます。 state_list には、リソースの状態を示す文字列が入っています。有効な値は、scha_cmds (1HA) コマンドで取得できる値です。 ev_ で始まるほかの名前や、それらの名前に関連した値が存在する場合がありますが、クライアントによる使用を意図したものではありません。

CRNP によるクライアントとサーバーの認証

サーバーは、TCP ラッパーを使用してクライアントの認証を行います。この場合、登録メッセージのソース IP アドレス (これはイベントの配信先であるコールバック IP アドレスとしても使用される) がサーバー側の「許可されたユーザー」リストに含まれていなければなりません。ソース IP アドレスと登録メッセージが「拒否されたクライアント」リストに存在してはなりません。ソース IP アドレスと登録メッセージがリスト中に存在しない場合、サーバーは要求を拒否し、クライアントに対してエラー応答を返します。

サーバーが `SC_CALLBACK_REG ADD_CLIENT` メッセージを受け取る場合、そのクライアントの後続の `SC_CALLBACK_REG` メッセージには最初のメッセージ内のものと同じソース IP アドレスが含まれていなければなりません。この条件を満たさない `SC_CALLBACK_REG` を受信した場合、CRNP サーバーは次のどちらかを選択します。

- 要求を無視し、クライアントにエラー応答を送信する
- その要求が新しいクライアントからのものであると想定する (`SC_CALLBACK_REG` メッセージの内容にもとづいて判断)

このセキュリティメカニズムは、正規クライアントの登録の解除を試みるサービス拒否攻撃の防止に役立ちます。

クライアントも、同様のサーバー認証を行う必要があります。クライアントは、それ自体が使用した登録 IP アドレスおよびポート番号と同じソース IP アドレスおよびポート番号を持つサーバーからのイベント配信を受け入れるだけです。

CRNP サービスのクライアントはクラスタを保護するファイアウォール内に配置されるのが一般的なため、CRNP にセキュリティメカニズムは提供されていません。

CRNP を使用する Java アプリケーションの作成

以下の例は、CRNP を使用する `CrnpClient` というシンプルな Java アプリケーションを作成する方法を示しています。このアプリケーションでは、クラスタ上の CRNP サーバーへのイベントコールバックの登録、イベントコールバックの待機、イベントの処理 (内容の出力) を行い、終了前にイベントコールバック要求の登録解除を行います。

この例を参照する場合は、以下の点に注意してください。

- このアプリケーション例は、JAXP (XML 処理用の Java API) による XML 生成と解析を行います。この例は JAXP の使用方法を説明したものではありません。JAXP の詳細は、<http://java.sun.com/xml/jaxp/index.html> で説明しています。
- この例は、付録 G に示されている完全なアプリケーションコードを断片的に示したものです。この章の例は個々の概念を効果的に示すことをねらっており、付録 G に示されている完全なアプリケーションコードと多少異なります。
- また、簡潔に示すため、この章の例ではコード例からコメントを除いてあります。付録 G に示されている完全なアプリケーションコードにはコメントが含まれています。
- この例に示しているアプリケーションは終了するだけでほとんどのエラー状況に対応できるものですが、ユーザーが実際に使用するアプリケーションではエラーを徹底的に処理する必要があります。

▼ 環境の設定

まず、環境の設定を行う必要があります。

1. JAXP と、正しいバージョンの Java コンパイラおよび Virtual Machine をダウンロードし、インストールを行います。

作業手順は、<http://java.sun.com/xml/jaxp/index.html> に示されています。

注 - この例は、バージョン 1.3.1 以降の Java を必要とします。

2. コンパイラが JAXP クラスを見つけることができるように、コンパイルのコマンド行に必ず **classpath** を指定する必要があります。ソースファイルが置かれているディレクトリから、次のように入力します。

```
% javac -classpath JAXP_ROOT/dom.jar:JAXP_ROOTjaxp-api. \
jar:JAXP_ROOTsax.jar:JAXP_ROOTxalan.jar:JAXP_ROOT/xercesImpl \
.jar:JAXP_ROOT/xsltc.jar -sourcepath . SOURCE_FILENAME.java
```

上記コマンドの JAXP_ROOT には、JAXP jar ファイルが置かれているディレクトリの絶対パスまたは相対パスを指定してください。SOURCE_FILENAME には、Java ソースファイルの名前を指定してください。

3. アプリケーションの実行時に、アプリケーションが適切な JAXP クラスファイルを読み込むことができるように **classpath** を指定します (**classpath** の最初のパスは現在のディレクトリ)。

```
java -cp .:JAXP_ROOT/dom.jar:JAXP_ROOTjaxp-api. \
jar:JAXP_ROOTsax.jar:JAXP_ROOTxalan.jar:JAXP_ROOT/xercesImpl \
.jar:JAXP_ROOT/xsltc.jar SOURCE_FILENAME ARGUMENTS
```

以上で環境の構成が終了し、アプリケーションの開発を行える状況となります。

▼ 作業の開始

サンプルアプリケーションのこの段階では、コマンド行引数を解析して `CrnpClient` オブジェクトの構築を行うメインメソッドを使用し、`CrnpClient` という基本的なクラスを作成します。このオブジェクトは、コマンド行引数をこのクラスに渡し、ユーザーがアプリケーションを終了するのを待って `CrnpClient` で `shutdown` を呼び出し、その後終了します。

`CrnpClient` クラスのコンストラクタは、以下の作業を実行する必要があります。

- オブジェクトを処理する XML を設定する
- イベントコールバックを待機するスレッドを作成する
- CRNP サーバーと通信し、イベントコールバックを受け取る登録をする
- 上記のロジックを実装する **Java** コードを作成します。

次の例は、`CrnpClient` クラスのスケルトンコードを示しています。コンストラクタ内で参照される 4 つのヘルパーメソッドと停止メソッドの実装はあとで示します。ここでは、ユーザーが必要とするパッケージをすべてインポートするコードを示しています。

```
import javax.xml.parsers.*;
import javax.xml.transform.*;
import javax.xml.transform.dom.*;
import javax.xml.transform.stream.*;
import org.xml.sax.*;
import org.xml.sax.helpers.*;
import org.w3c.dom.*;

import java.net.*;
import java.io.*;
import java.util.*;

class CrnpClient
{
    public static void main(String []args)
    {
        InetAddress regIp = null;
        int regPort = 0, localPort = 0;

        try {
            regIp = InetAddress.getByName(args[0]);
            regPort = (new Integer(args[1])).intValue();
            localPort = (new Integer(args[2])).intValue();
        } catch (UnknownHostException e) {
            System.out.println(e);
            System.exit(1);
        }

        CrnpClient client = new CrnpClient(regIp, regPort, localPort,
            args);
        System.out.println("Hit return to terminate demo...");
        try {
```

```

        System.in.read();
    } catch (IOException e) {
        System.out.println(e.toString());
    }
    client.shutdown();
    System.exit(0);
}

public CrnpClient(InetAddress regIpIn, int regPortIn, int localPortIn,
    String []clArgs)
{
    try {
        regIp = regIpIn;
        regPort = regPortIn;
        localPort = localPortIn;
        regs = clArgs;

        setupXmlProcessing();
        createEvtRecepThr();
        registerCallbacks();

    } catch (Exception e) {
        System.out.println(e.toString());
        System.exit(1);
    }
}

public void shutdown()
{
    try {
        unregister();
    } catch (Exception e) {
        System.out.println(e);
        System.exit(1);
    }
}

private InetAddress regIp;
private int regPort;
private EventReceptionThread evtThr;
private String regs[];

public int localPort;
public DocumentBuilderFactory dbf;
}

```

メンバー変数についての詳細は後述します。

▼ コマンド行引数の解析

- コマンド行引数の解析方法については、付録 G 内のコードを参照してください。

▼ イベント受信スレッドの定義

イベント受信はコード内で個別のスレッドで行われるようにする必要があります。これは、イベントスレッドがイベントコールバックを待機している間アプリケーションが継続してほかの作業を行えるようにするためです。

注 - XML の設定については後述します。

1. コード内で、**ServerSocket** を作成してソケットにイベントが到着するのを待機する **EventReceptionThread** という **Thread** サブクラスを定義します。
サンプルコードのこの部分では、イベントの読み取りもイベントの処理も行われません。イベントの読み取りと処理については後述します。
EventReceptionThread は、ワイルドカード IP アドレス上に **ServerSocket** を作成します。**EventReceptionThread** は、**CrnpClient** オブジェクトにイベントを送信して処理できるように、**CrnpClient** オブジェクトに対する参照も維持します。

```
class EventReceptionThread extends Thread
{
    public EventReceptionThread(CrnpClient clientIn) throws IOException
    {
        client = clientIn;
        listeningSock = new ServerSocket(client.localPort, 50,
            InetAddress.getLocalHost());
    }

    public void run()
    {
        try {
            DocumentBuilder db = client.dbf.newDocumentBuilder();
            db.setErrorHandler(new DefaultHandler());

            while(true) {
                Socket sock = listeningSock.accept();
                // ソケットストリームからイベントを作成し、処理する。
                sock.close();
            }
            // 到達不能

        } catch (Exception e) {
            System.out.println(e);
            System.exit(1);
        }
    }

    /* プライベートメンバー変数 */
    private ServerSocket listeningSock;
    private CrnpClient client;
}
```

2. 以上で、**EventReceptionThread** クラスがどのように動作するか確認ができませんでした。次は、**createEvtRecepThr** オブジェクトを構築します。

```
private void createEvtRecepThr() throws Exception
{
    evtThr = new EventReceptionThread(this);
    evtThr.start();
}
```

▼ コールバックの登録と登録解除

登録は以下の作業によって行います。

- 登録用の IP アドレスとポートに対して基本的な TCP ソケットを開く
- XML 登録メッセージを作成する
- ソケット上で XML 登録メッセージを送信する
- ソケットから XML 応答メッセージを読み取る
- ソケットを閉じる

1. 上記のロジックを実装する **Java** コードを作成します。

以下の例は、**CrnpClient** クラスの **registerCallbacks** メソッド (**CrnpClient** コンストラクタによって呼び出される) の実装を示しています。**createRegistrationString()** と **readRegistrationReply()** の呼び出しの詳細は後述します。

regIp と **regPort** は、コンストラクタによって設定されるオブジェクトメンバーです。

```
private void registerCallbacks() throws Exception
{
    Socket sock = new Socket(regIp, regPort);
    String xmlStr = createRegistrationString();
    PrintStream ps = new
        PrintStream(sock.getOutputStream());
    ps.print(xmlStr);
    readRegistrationReply(sock.getInputStream());
    sock.close();
}
```

2. **unregister** メソッドを実装します。このメソッドは、**CrnpClient** の **shutdown** メソッドによって呼び出されます。**createUnregistrationString** の実装の詳細は後述します。

```
private void unregister() throws Exception
{
    Socket sock = new Socket(regIp, regPort);
    String xmlStr = createUnregistrationString();
    PrintStream ps = new PrintStream(sock.getOutputStream());
    ps.print(xmlStr);
    readRegistrationReply(sock.getInputStream());
    sock.close();
}
```

▼ XML の生成

以上で、アプリケーション構造の設定と、通信用のコードの作成が終了しました。次は、XML の生成と解析を行うコードを作成します。初めに、SC_CALLBACK_REG XML 登録メッセージを生成するコードを作成します。

SC_CALLBACK_REG メッセージは、登録のタイプ (ADD_CLIENT、REMOVE_CLIENT、ADD_EVENTS、または REMOVE_EVENTS)、コールバックポート、および要求するイベントの一覧から構成されます。各イベントはクラスとサブクラスから構成され、名前と値のペアリストが続きます。

この例のこの段階では、登録タイプ、コールバックポート、および登録イベントの一覧を格納する CallbackReg クラスを作成します。このクラスは、それ自体を SC_CALLBACK_REG XML メッセージにシリアル化することもできます。

このクラスには、クラスメンバーから SC_CALLBACK_REG XML メッセージ文字列を作成する convertToXml という興味深いメソッドがあります。このメソッドを使用したコードの詳細は、<http://java.sun.com/xml/jaxp/index.html> の JAXP ドキュメントに記載されています。

次に、Event クラスの実装を示します。CallbackReg クラスは、イベントを1つ保存してそのイベントを XML Element に変換できる Event クラスを使用します。

1. 上記のロジックを実装する Java コードを作成します。

```
class CallbackReg
{
    public static final int ADD_CLIENT = 0;
    public static final int ADD_EVENTS = 1;
    public static final int REMOVE_EVENTS = 2;
    public static final int REMOVE_CLIENT = 3;

    public CallbackReg()
    {
        port = null;
        regType = null;
        regEvents = new Vector();
    }

    public void setPort(String portIn)
    {
        port = portIn;
    }

    public void setRegType(int regTypeIn)
    {
        switch (regTypeIn) {
            case ADD_CLIENT:
                regType = "ADD_CLIENT";
                break;
            case ADD_EVENTS:
                regType = "ADD_EVENTS";
                break;
        }
    }
}
```

```

        case REMOVE_CLIENT:
            regType = "REMOVE_CLIENT";
            break;
        case REMOVE_EVENTS:
            regType = "REMOVE_EVENTS";
            break;
        default:
            System.out.println("Error, invalid regType " +
                regTypeIn);
            regType = "ADD_CLIENT";
            break;
    }
}

public void addRegEvent(Event regEvent)
{
    regEvents.add(regEvent);
}

public String convertToXml()
{
    Document document = null;
    DocumentBuilderFactory factory =
        DocumentBuilderFactory.newInstance();
    try {
        DocumentBuilder builder = factory.newDocumentBuilder();
        document = builder.newDocument();
    } catch (ParserConfigurationException pce) {
        // 指定されたオプションを持つパーサーを構築できない。
        pce.printStackTrace();
        System.exit(1);
    }

    // root 要素を作成する。
    Element root = (Element) document.createElement(
        "SC_CALLBACK_REG");

    // 属性を追加する。
    root.setAttribute("VERSION", "1.0");
    root.setAttribute("PORT", port);
    root.setAttribute("regType", regType);

    // イベントを追加する。
    for (int i = 0; i < regEvents.size(); i++) {
        Event tempEvent = (Event)
            (regEvents.elementAt(i));
        root.appendChild(tempEvent.createXmlElement(
            document));
    }
    document.appendChild(root);

    // 全体を文字列に変換する。
    DOMSource domSource = new DOMSource(document);
    StringWriter strWrite = new StringWriter();
    StreamResult streamResult = new StreamResult(strWrite);

```

```

TransformerFactory tf = TransformerFactory.newInstance();
try {
    Transformer transformer = tf.newTransformer();
    transformer.transform(domSource, streamResult);
} catch (TransformerException e) {
    System.out.println(e.toString());
    return ("");
}
return (strWrite.toString());
}

private String port;
private String regType;
private Vector regEvents;
}

```

2. Event クラスと NVPair クラスを実装します。

CallbackReg クラスは、NVPair クラスを使用する Event クラスを使用します。

```

class Event
{
    public Event()
    {
        regClass = regSubclass = null;
        nvpairs = new Vector();
    }

    public void setClass(String classIn)
    {
        regClass = classIn;
    }

    public void setSubclass(String subclassIn)
    {
        regSubclass = subclassIn;
    }

    public void addNvpair(NVPair nvpair)
    {
        nvpairs.add(nvpair);
    }

    public Element createXmlElement(Document doc)
    {
        Element event = (Element)
            doc.createElement("SC_EVENT_REG");
        event.setAttribute("CLASS", regClass);
        if (regSubclass != null) {
            event.setAttribute("SUBCLASS", regSubclass);
        }
        for (int i = 0; i < nvpairs.size(); i++) {
            NVPair tempNv = (NVPair)
                (nvpairs.elementAt(i));

```

```

        event.appendChild(tempNv.createXmlElement(
            doc));
    }
    return (event);
}

private String regClass, regSubclass;
private Vector nvpairs;
}

class NVPair
{
    public NVPair()
    {
        name = value = null;
    }

    public void setName(String nameIn)
    {
        name = nameIn;
    }

    public void setValue(String valueIn)
    {
        value = valueIn;
    }

    public Element createXmlElement(Document doc)
    {
        Element nvpair = (Element)
            doc.createElement("NVPAIR");
        Element eName = doc.createElement("NAME");
        Node nameData = doc.createCDATASection(name);
        eName.appendChild(nameData);
        nvpair.appendChild(eName);
        Element eValue = doc.createElement("VALUE");
        Node valueData = doc.createCDATASection(value);
        eValue.appendChild(valueData);
        nvpair.appendChild(eValue);

        return (nvpair);
    }

    private String name, value;
}

```

▼ 登録メッセージと登録解除メッセージの作成

XML メッセージを生成するヘルパークラスの作成が終了したところで、次は `createRegistrationString` メソッドを実装します。このメソッドは、`registerCallbacks` メソッド (詳細は 226 ページの「コールバックの登録と登録解除」) によって呼び出されます。

createRegistrationString は、CallbackReg オブジェクトを構築し、その登録タイプとポートを設定します。続いて、createRegistrationString は、createAllEvent、createMembershipEvent、createRgEvent、および createREvent ヘルパーメソッドを使用して各種のイベントを構築します。各イベントは、CallbackReg オブジェクトが作成されたあとでこのオブジェクトに追加されます。最後に、createRegistrationString は CallbackReg オブジェクト上で convertToXml メソッドを呼び出し、String 形式の XML メッセージを取得します。

regs メンバー変数は、ユーザーがアプリケーションに指定するコマンド行引数を格納します。5つ目以降の引数は、アプリケーションが登録を行うイベントを指定します。4つ目の引数は登録のタイプを指定しますが、この例では無視されています。付録 G に挙げられている完全なコードでは、この4つ目の引数の使用方法も示されています。

1. 上記のロジックを実装する Java コードを作成します。

```
private String createRegistrationString() throws Exception
{
    CallbackReg cbReg = new CallbackReg();
    cbReg.setPort("" + localPort);

    cbReg.setRegType(CallbackReg.ADD_CLIENT);

    // イベントを追加する
    for (int i = 4; i < regs.length; i++) {
        if (regs[i].equals("M")) {
            cbReg.addRegEvent(
                createMembershipEvent());
        } else if (regs[i].equals("A")) {
            cbReg.addRegEvent(
                createAllEvent());
        } else if (regs[i].substring(0,2).equals("RG")) {
            cbReg.addRegEvent(createRgEvent(
                regs[i].substring(3)));
        } else if (regs[i].substring(0,1).equals("R")) {
            cbReg.addRegEvent(createREvent(
                regs[i].substring(2)));
        }
    }

    String xmlStr = cbReg.convertToXml();
    return (xmlStr);
}

private Event createAllEvent()
{
    Event allEvent = new Event();
    allEvent.setClass("EC_Cluster");
    return (allEvent);
}

private Event createMembershipEvent()
```

```

    {
        Event membershipEvent = new Event();
        membershipEvent.setClass("EC_Cluster");
        membershipEvent.setSubclass("ESC_cluster_membership");
        return (membershipEvent);
    }

private Event createRgEvent(String rgname)
{
    Event rgStateEvent = new Event();
    rgStateEvent.setClass("EC_Cluster");
    rgStateEvent.setSubclass("ESC_cluster_rg_state");

    NVPair rgNvpair = new NVPair();
    rgNvpair.setName("rg_name");
    rgNvpair.setValue(rgname);
    rgStateEvent.addNvpair(rgNvpair);

    return (rgStateEvent);
}

private Event createREvent(String rname)
{
    Event rStateEvent = new Event();
    rStateEvent.setClass("EC_Cluster");
    rStateEvent.setSubclass("ESC_cluster_r_state");

    NVPair rNvpair = new NVPair();
    rNvpair.setName("r_name");
    rNvpair.setValue(rname);
    rStateEvent.addNvpair(rNvpair);

    return (rStateEvent);
}

```

2. 登録解除文字列を作成します。

イベントを指定する必要がない分、登録解除文字列の作成は登録文字列の作成よりも簡単です。

```

private String createUnregistrationString() throws Exception
{
    CallbackReg cbReg = new CallbackReg();
    cbReg.setPort("" + localPort);
    cbReg.setRegType(CallbackReg.REMOVE_CLIENT);
    String xmlStr = cbReg.convertToXml();
    return (xmlStr);
}

```

▼ XML パーサーの設定

以上で、アプリケーションの通信用コードと XML 生成コードの生成が終わります。最後のステップとして、登録応答とイベントコールバックの解析と処理を行います。CrnpClient コンストラクタは setupXmlProcessing メソッドを呼び出します。

このメソッドは、DocumentBuilderFactory オブジェクトを作成し、そのオブジェクトに各種の解析プロパティを設定します。このメソッドの詳細は、<http://java.sun.com/xml/jaxp/index.html> の JAXP ドキュメントに記載されています。

- 上記のロジックを実装する **Java** コードを作成します。

```
private void setupXmlProcessing() throws Exception
{
    dbf = DocumentBuilderFactory.newInstance();

    // 検証を行う必要はない。
    dbf.setValidating(false);
    dbf.setExpandEntityReferences(false);

    // コメントと空白文字は無視したい。
    dbf.setIgnoringComments(true);
    dbf.setIgnoringElementContentWhitespace(true);

    // CDATA セクションを TEXT ノードに結合する。
    dbf.setCoalescing(true);
}
```

▼ 登録応答の解析

登録メッセージまたは登録解除メッセージに回答して CRNP サーバーが送信する SC_REPLY XML メッセージを解析するには、RegReply ヘルパークラスが必要です。このクラスは、XML ドキュメントから構築できます。このクラスは、ステータスコードとステータスメッセージのアクセッサを提供します。サーバーからの XML ストリームを解析するには、新しい XML ドキュメントを作成してそのドキュメントの解析メソッドを使用する必要があります(このメソッドの詳細は <http://java.sun.com/xml/jaxp/index.html> の JAXP ドキュメントを参照)。

1. 上記のロジックを実装する **Java** コードを作成します。

readRegistrationReply メソッドは、新しい RegReply クラスを使用します。

```
private void readRegistrationReply(InputStream stream) throws Exception
{
    // ドキュメントビルダーを作成する。
    DocumentBuilder db = dbf.newDocumentBuilder();
    db.setErrorHandler(new DefaultHandler());

    //入力ファイルを解析する。
    Document doc = db.parse(stream);

    RegReply reply = new RegReply(doc);
    reply.print(System.out);
}
```

2. **RegReply** クラスを実装します。

retrieveValues メソッドは XML ドキュメント内の DOM ツリーを回り、ステータスコードとステータスメッセージを抽出します。詳細は、<http://java.sun.com/xml/jaxp/index.html> の JAXP ドキュメントに記載されています。

```
class RegReply
{
    public RegReply(Document doc)
    {
        retrieveValues(doc);
    }

    public String getStatusCode()
    {
        return (statusCode);
    }

    public String getStatusMsg()
    {
        return (statusMsg);
    }

    public void print(PrintStream out)
    {
        out.println(statusCode + ": " +
            (statusMsg != null ? statusMsg : ""));
    }

    private void retrieveValues(Document doc)
    {
        Node n;
        NodeList nl;
        String nodeName;

        // SC_REPLY 要素を見つける。
        nl = doc.getElementsByTagName("SC_REPLY");
        if (nl.getLength() != 1) {
            System.out.println("Error in parsing: can't find "
                + "SC_REPLY node.");
            return;
        }

        n = nl.item(0);

        // statusCode 属性の値を取得する。
        statusCode = ((Element)n).getAttribute("STATUS_CODE");

        // SC_STATUS_MSG 要素を検出する。
        nl = ((Element)n).getElementsByTagName("SC_STATUS_MSG");
        if (nl.getLength() != 1) {
            System.out.println("Error in parsing: can't find "
                + "SC_STATUS_MSG node.");
            return;
        }
        // TEXT セクションを取得する (存在する場合)。
        n = nl.item(0).getFirstChild();
    }
}
```

```

        if (n == null || n.getNodeType() != Node.TEXT_NODE) {
            // 1 つも存在しなくてもエラーではないため、そのまま戻る。
            return;
        }

        // 値を取得する。
        statusMsg = n.getNodeValue();
    }

private String statusCode;
private String statusMsg;
}

```

▼ コールバックイベントの解析

最後のステップは、実際のコールバックイベントの解析と処理です。この作業をスムーズに行うため、[227 ページの「XML の生成」](#)で作成した Event クラスを変更します。このクラスを使用して XML ドキュメントから Event を構築し、XML Element を作成できます。この変更は、XML ドキュメントを受け付ける別のコンストラクタ、retrieveValues メソッド、2 つの新たなメンバー変数 (vendor と publisher)、全フィールドのアクセッサメソッド、および出力メソッドを必要とします。

1. 上記のロジックを実装する **Java** コードを作成します。

このコードは、[233 ページの「登録応答の解析」](#)で説明している RegReply クラスのコードに似ていることに注目してください。

```

public Event(Document doc)
{
    nvpairs = new Vector();
    retrieveValues(doc);
}

public void print(PrintStream out)
{
    out.println("\tCLASS=" + regClass);
    out.println("\tSUBCLASS=" + regSubclass);
    out.println("\tVENDOR=" + vendor);
    out.println("\tPUBLISHER=" + publisher);
    for (int i = 0; i < nvpairs.size(); i++) {
        NVPair tempNv = (NVPair)
            (nvpairs.elementAt(i));
        out.print("\t\t");
        tempNv.print(out);
    }
}

private void retrieveValues(Document doc)
{
    Node n;
    NodeList nl;
    String nodeName;
}

```

```

// SC_EVENT 要素を検出する。
nl = doc.getElementsByTagName("SC_EVENT");
if (nl.getLength() != 1) {
    System.out.println("Error in parsing: can't find "
        + "SC_EVENT node.");
    return;
}

n = nl.item(0);

//
// CLASS、SUBCLASS、VENDOR、および PUBLISHER
// 属性の値を取得する。
//
regClass = ((Element)n).getAttribute("CLASS");
regSubclass = ((Element)n).getAttribute("SUBCLASS");
publisher = ((Element)n).getAttribute("PUBLISHER");
vendor = ((Element)n).getAttribute("VENDOR");

// すべての nv ペアを取得する。
for (Node child = n.getFirstChild(); child != null;
    child = child.getNextSibling())
{
    nvpairs.add(new NVPair((Element)child));
}

}

public String getRegClass()
{
    return (regClass);
}

public String getSubclass()
{
    return (regSubclass);
}

public String getVendor()
{
    return (vendor);
}

public String getPublisher()
{
    return (publisher);
}

public Vector getNvpairs()
{
    return (nvpairs);
}

private String vendor, publisher;

```

2. XML 解析をサポートする、**NVPair** クラスのコンストラクタとメソッドを別途実装します。

手順1でEventクラスに変更を加えたため、NVPairクラスにも類似した変更を加える必要があります。

```
public NVPair(Element elem)
{
    retrieveValues(elem);
}
public void print(PrintStream out)
{
    out.println("NAME=" + name + " VALUE=" + value);
}
private void retrieveValues(Element elem)
{
    Node n;
    NodeList nl;
    String nodeName;

    // NAME 要素を検出する。
    nl = elem.getElementsByTagName("NAME");
    if (nl.getLength() != 1) {
        System.out.println("Error in parsing: can't find "
            + "NAME node.");
        return;
    }
    // TEXT セクションを取得する。
    n = nl.item(0).getFirstChild();
    if (n == null || n.getNodeType() != Node.TEXT_NODE) {
        System.out.println("Error in parsing: can't find "
            + "TEXT section.");
        return;
    }

    // 値を取得する。
    name = n.getNodeValue();

    // ここで要素を取得する。
    nl = elem.getElementsByTagName("VALUE");
    if (nl.getLength() != 1) {
        System.out.println("Error in parsing: can't find "
            + "VALUE node.");
        return;
    }
    // TEXT セクションを取得する。
    n = nl.item(0).getFirstChild();
    if (n == null || n.getNodeType() != Node.TEXT_NODE) {
        System.out.println("Error in parsing: can't find "
            + "TEXT section.");
        return;
    }

    // 値を取得する。
    value = n.getNodeValue();
}

public String getName()
```

```

        {
            return (name);
        }

        public String getValue()
        {
            return (value);
        }
    }

```

3. **EventReceptionThread** でイベントコールバックを待機する **while** ループを実装します (**EventReceptionThread** については [225 ページの「イベント受信スレッドの定義」](#) を参照)。

```

while(true) {
    Socket sock = listeningSock.accept();
    Document doc = db.parse(sock.getInputStream());
    Event event = new Event(doc);
    client.processEvent(event);
    sock.close();
}

```

▼ アプリケーションの実行

- アプリケーションを実行します。

```
# java CrnpClient crnpHost crnpPort localPort ...
```

完全な CrnpClient アプリケーションコードは、[付録 G](#) に示されています。

標準プロパティ

この付録では、標準リソースタイプ、リソースグループ、リソースプロパティについて説明します。また、システム定義プロパティの変更および拡張プロパティの作成に使用するリソースプロパティ属性についても説明します。

この付録は、次の節で構成されています。

- 239 ページの「リソースタイププロパティ」
- 245 ページの「リソースのプロパティ」
- 257 ページの「リソースグループのプロパティ」
- 263 ページの「リソースプロパティの属性」

リソースタイププロパティ

以下に、Sun Cluster によって定義されるリソースタイププロパティを示します。プロパティ値は、「カテゴリ」で次のように分類されます。

- 必須 – プロパティはリソースタイプ登録 (RTR) ファイルに明示的な値を必要とします。そうでない場合、プロパティが属するオブジェクトは作成できません。空白文字または空の文字列を値として指定することはできません。
- 条件付 — このプロパティが存在するためには、RTR ファイル内で宣言する必要があります。宣言がない場合、RGM はこのプロパティを作成しません。したがって、このプロパティを管理ユーティリティから利用することはできません。空白文字または空の文字列を値として指定できます。プロパティが RTR ファイル内で宣言されており、値が指定されていない場合には、RGM はデフォルト値を使用します。
- 条件付/明示 — このプロパティが存在するためには、明示的に値を指定し、RTR ファイル内で宣言する必要があります。宣言がない場合、RGM はこのプロパティを作成しません。したがって、このプロパティを管理ユーティリティから利用することはできません。空白文字または空の文字列を値として指定することはでき

ません。

- 任意 — プロパティを RTR ファイル内で宣言できます。プロパティが RTR ファイル内で宣言されていない場合は、RGM がこれを作成し、デフォルト値を与えます。プロパティが RTR ファイル内で宣言されており、値が指定されていない場合は、RGM は、プロパティが RTR ファイル内で宣言されないときのデフォルト値と同じ値を使用します。

リソースタイププロパティを管理ユーティリティで更新することはできません。ただし、RTR ファイル内に宣言できない `Installed_nodes` と `RT_system` は除きます。これらのプロパティは、管理者が設定する必要があります。

以下にプロパティ名とその説明を示します。

API_version (整数型)

このリソースタイプの実装が使用するリソース管理 API のバージョン。

Sun Cluster の各バージョンによってサポートされる最新の API_version は次のとおりです。

3.1 以前	2
3.1 10/03	3
3.1 4/04	4
3.1 9/04	5

RTR ファイルで API_version に 2 より大きい値を宣言すると、そのリソースタイプは、その値より古いバージョンをサポートする Sun Cluster にはインストールされません。たとえば、あるリソースタイプに API_version=5 を宣言すると、このリソースタイプは、3.1 9/04 より前にリリースされた Sun Cluster のバージョンにはインストールされません。

カテゴリ: 任意
デフォルト: 2
調整: No

Boot (文字列型)

任意のコールバックメソッド。RGM がノード上で呼び出すプログラムのパスを指定します。このプログラムは、このリソース型が管理対象になっているとき、クラスタの結合または再結合を行います。このメソッドは、このタイプのリソースに対して、Init メソッドと同様の初期化アクションを行う必要があります。

カテゴリ: 条件付き / 明示
デフォルト: なし
調整: No

Failover (プール型)

TRUE の場合、複数のノード上で同時にオンラインにできるグループ内にこの型のリソースを構成することはできません。

カテゴリ: 任意
デフォルト: FALSE
調整: No

Finis (文字列型)

任意のコールバックメソッド。この型のリソースを RGM 管理の対象外にすると RGM によって呼び出されるプログラムのパスです。

カテゴリ: 条件付き / 明示
デフォルト: なし
調整: No

Init (文字列型)

任意のコールバックメソッド。この型のリソースを RGM 管理対象にすると RGM によって呼び出されるプログラムのパスです。

カテゴリ: 条件付き / 明示
デフォルト: なし
調整: No

Init_nodes (列挙型)

指定できる値は、RG primaries (リソースをマスターできるノードのみ) または RT_installed_nodes (このリソース型がインストールされる全てのノード) のいずれかです。RGM が Init、Finis、Boot、Validate メソッドをコールするノードを示します。

カテゴリ: 任意
デフォルト: RG primaries
調整: No

Installed_nodes (文字配列型)

リソースタイプの実行が許可されるクラスタノード名のリスト。このプロパティは RGM によって自動的に作成されます。クラスタ管理者は値を設定できます。RTR ファイル内には宣言できません。

カテゴリ: クラスタ管理者による構成が可能です。
デフォルト: すべてのクラスタノード
調整: Yes

Is_logical_hostname (ブール型)

TRUEは、このリソース型が、フェイルオーバーインターネットプロトコル (IP) アドレスを管理するLogicalHostname リソース型のいずれかのバージョンであることを示します。

カテゴリ: 照会のみ
デフォルト: デフォルトなし

調整: No

Is_shared_address (ブール型)

TRUE は、このリソース型が、フェイルオーバーインターネットプロトコル (IP) アドレスを管理する共有アドレスリソース型のいずれかのバージョンであることを示します。

カテゴリ: 照会のみ

デフォルト: デフォルトなし

調整: No

Monitor_check (文字列型)

任意のコールバックメソッド。障害モニターの要求によってこのリソース型のフェイルオーバーを実行する前に、RGM によって呼び出されるプログラムのパスです。

カテゴリ: 条件付き / 明示

デフォルト: デフォルトなし

調整: No

Monitor_start (文字列型)

任意のコールバックメソッド。この型のリソースの障害モニターを起動するために RGM によって呼び出されるプログラムのパスです。

カテゴリ: 条件付き / 明示

デフォルト: デフォルトなし

調整: No

Monitor_stop (文字列型)

Monitor_start が設定されている場合の、必須のコールバックメソッド。この型のリソースの障害モニターを停止するために RGM によって呼び出されるプログラムのパスです。

カテゴリ: 条件付き / 明示

デフォルト: デフォルトなし

調整: No

Pkglist (文字配列型)

リソース型のインストールに含まれている任意のパッケージリストです。

カテゴリ: 条件付き / 明示

デフォルト: デフォルトなし

調整: No

Postnet_stop (文字列型)

任意のコールバックメソッド。この型のリソースが、あるネットワークアドレスリソースに依存している場合、そのネットワークアドレスリソースの Stop メソッドの呼び出し後に RGM によって呼び出されるプログラムのパスです。ネットワークインタフェースが構成され、停止状態にされた場合、このメソッドは Stop アクションを行う必要があります。

カテゴリ: 条件付き / 明示

デフォルト: デフォルトなし

調整: No

Prenet_start (文字列型)

任意のコールバックメソッド。この型のリソースがネットワークアドレスリソースに依存している場合、このネットワークアドレスリソースの Start メソッドの呼び出し前に RGM によって呼び出されるプログラムのパスです。ネットワークインタフェースが構成される前に必要な Start アクションを行う必要があります。

カテゴリ: 条件付き / 明示

デフォルト: デフォルトなし

調整: No

Resource_type (文字列型)

リソースタイプの名前。現在登録されているリソースタイプ名を表示するには、次のコマンドを使用します。

```
scrgadm -p
```

Sun Cluster 3.1 以降のリリースでは、リソースタイプ名にバージョンが含まれます (必須)。

```
vendor_id.resource_type:version
```

リソース型の名前は、RTR ファイル内に指定された 3 つのプロパティ *Vendor_id*、*Resource_type*、*RT_version* で構成されます。scrgadm コマンドがピリオド (.) と コロン (:) をプロパティの間に挿入します。リソースタイプ名の最後の部分、*RT_version* には、*RT_version* プロパティと同じ値が入ります。重複を防ぐため、*Vendor_id* には、リソースタイプの作成元の会社のストックシンボルを使用することをお勧めします。Sun Cluster 3.1 以前に登録されたリソースタイプ名では、引き続き次の構文を使用します。

```
vendor_id.resource_type
```

カテゴリ: 必須

デフォルト: 空の文字列

調整: No

RT_basedir (文字列型)

コールバックメソッドの相対パスのを補完するディレクトリパスです。このパスは、リソースタイプパッケージのインストール場所に設定します。このパスには、

スラッシュ (/) で開始する完全なパスを指定する必要があります。すべてのメソッドパス名が絶対パスの場合は、指定しなくてもかまいません。

カテゴリ: 必須 (絶対パスでないメソッドパスがある場合)

デフォルト: デフォルトなし

調整: No

RT_description (文字列型)

リソース型の簡単な説明です。

カテゴリ: 条件付き

デフォルト: 空の文字列

調整: No

RT_system (ブール型)

あるリソースタイプに対してこのプロパティが TRUE に設定されている場合、このリソースタイプに対して許されている scrgadm(1M) の操作は制限されます。

RT_system の値が TRUE に設定されているリソースタイプをシステムリソースタイプといいます。RT_system の現在の値が何であれ、このプロパティ自体の編集は制限されていません。

カテゴリ: 任意

デフォルト: FALSE

調整: Yes

RT_version (文字列型)

Sun Cluster 3.1 以降では、このリソースタイプの実装の必須バージョン文字列。

RT_version は、完全なリソースタイプ名の末尾の部分です。RT_version プロパティの設定は、Sun Cluster 3.0 では任意でしたが、Sun Cluster 3.1 以降のリリースでは必須です。

カテゴリ: (任意) 明示または必須

デフォルト: デフォルトなし

調整: No

Single_instance (ブール型)

TRUE は、この型のリソースがクラスタ内に 1 つだけ存在できることを示します。

RGM は、同時に 1 つのこのリソースタイプだけに、クラスタ全体に渡っての実行を許可します。

カテゴリ: 任意

デフォルト: FALSE

調整: No

Start (文字列型)

コールバックメソッド。この型のリソースを起動するために RGM によって呼び出されるプログラムのパスです。

カテゴリ: RTR ファイルで Prenet_start メソッドが宣言されていないか
ぎり必須

デフォルト: デフォルトなし

調整: No

Stop (文字列型)

コールバックメソッド。この型のリソースを停止するために RGM によって呼び出されるプログラムのパスです。

カテゴリ: RTR ファイルで Postnet_stop メソッドが宣言されていないか
ぎり必須

デフォルト: デフォルトなし

調整: No

Update (文字列型)

任意のコールバックメソッド。この型の実行中のリソースのプロパティが変更されたとき RGM によって呼び出されるプログラムのパスです。

カテゴリ: 条件付き / 明示

デフォルト: デフォルトなし

調整: No

Validate (文字列型)

任意のコールバックメソッド。この型のリソースのプロパティ値を検査するために呼び出されるプログラムのパスです。

カテゴリ: 条件付き / 明示

デフォルト: デフォルトなし

調整: No

Vendor_ID (文字列型)

Resource_type を参照してください。

カテゴリ: 条件付き

デフォルト: デフォルトなし

調整: No

リソースのプロパティ

この節では、Sun Cluster によって定義されているリソースプロパティについて説明します。プロパティ値は、「カテゴリ」で次のように分類されます。

- 必須 — 管理者は、管理ユーティリティでリソースを作成するときに、必ず値を指定する必要があります。
- 任意 — 管理者がリソースグループの作成時に値を指定しない場合、システムがデフォルト値を提供します。
- 条件付き — プロパティが RTR ファイルで宣言されている場合にのみ、RGM がプロパティを作成します。宣言されていない場合プロパティは存在せず、システム管理者はこれを利用できません。RTR ファイルで宣言されている条件付きのプロパティは、デフォルト値が RTR ファイル内で指定されているかどうかによって、必須または任意になります。詳細については、各条件付きプロパティの説明を参照してください。
- 照会のみ — 管理ツールから直接設定できません。

Tunable 属性では、リソースプロパティを更新できるかどうかや、いつ更新できるかを指定します。

NONE または FALSE	しない
TRUE または ANYTIME	任意の時点 (Anytime)
AT_CREATION	リソースをクラスタに追加するとき
WHEN_DISABLED	リソースが無効なとき

以下にプロパティ名とその説明を示します。

Affinity_timeout (整数型)

リソース内のサービスのクライアント IP アドレスからの接続は、この時間 (秒数) 内に同じサーバーノードに送信されます。

このプロパティは、Load_balancing_policy が Lb_sticky または Lb_sticky_wild の場合にかぎり有効です。さらに、Weak_affinity を FALSE (デフォルト値) に設定されている必要があります。

このプロパティは、スケーラブルサービス専用です。

カテゴリ: 任意
 デフォルト: デフォルトなし
 調整: ANYTIME

Cheap_probe_interval (整数型)

リソースの即時障害検証の呼び出しの間隔 (秒数)。このプロパティは RGM によって作成されます。RTR ファイルに宣言されている場合にかぎり、管理者は使用を許可されます。

RTR ファイル内でデフォルト値が指定されている場合、このプロパティは任意です。リソース型ファイル内に Tunable 属性が指定されていない場合、このプロパティの Tunable 値は WHEN_DISABLED になります。

このプロパティが RTR ファイルに宣言され、かつ Default 属性が指定されていない場合、このプロパティは必須です。

カテゴリ: 条件付き
デフォルト: デフォルトなし
調整: WHEN_DISABLED

拡張プロパティ

そのリソースのタイプの RTR ファイルで宣言される拡張プロパティ。リソースタイプの実装によって、これらのプロパティを定義します。拡張プロパティに設定可能な各属性については、263 ページの「リソースプロパティの属性」を参照してください。

カテゴリ: 条件付き
デフォルト: デフォルトなし
調整: 特定のプロパティに依存

Failover_mode (列挙型)

NONE、SOFT、HARD は、起動メソッド (Prenet_start または Start) が失敗した場合、フェイルオーバー動作だけに影響を与えます。リソースがいったん正常に起動すると、NONE、SOFT、HARD は、その後にリソースモニターが scha_control (1HA) や scha_control (3HA) を使って行うリソース再起動やギブオーバー動作に対して何の影響も及ぼしません。NONE (デフォルト) は、メソッドの失敗時に RGM がリソース状態を設定し、ユーザーの介入を待つことを意味します。SOFT は、Start メソッドが失敗した場合、RGM がリソースのグループを別のノードに再配置することを意味します。Stop または Monitor_stop メソッドが失敗した場合、RGM は、リソースを Stop_failed 状態に、リソースグループを Error_stop_failed 状態にそれぞれ設定します。そして、ユーザーの介入を待ちます。Stop または Monitor_stop が失敗した場合の NONE と SOFT は同じ意味です。HARD は、Start メソッドが失敗した場合、RGM がグループを再配置することを意味します。Stop または Monitor_stop メソッドが失敗した場合、RGM はクラスターノードを放棄することによってリソースを停止します。HARD、NONE、SOFT は、Start または Prenet_start メソッドが失敗した場合、フェイルオーバー動作に影響を与えます。

NONE、SOFT、HARD とは異なり、RESTART_ONLY と LOG_ONLY は、すべてのフェイルオーバー動作に影響を与えます。フェイルオーバー動作には、モニター起動 (scha_control) によるリソースやリソースグループの再起動や、リソースモニター (scha_control) によって開始されるギブオーバーなどがあります。RESTART_ONLY は、モニターが scha_control を使ってリソースを再起動することはできますが、その後に scha_control を使ってリソースグループの再起動やギブオーバーをしようとするとう失敗することを意味します。RGM では、Retry_interval の間に Retry_count 回数だけ再起動を試行できます。試行回数が Retry_count を超えると、それ以後、リソース再起動は許可されません。Failover_mode に LOG_ONLY を設定すると、リソース再起動は許可されませ

ん。Failover_mode に LOG_ONLY を設定するのは、Failover_mode に RESTART_ONLY を設定し、Retry_count にゼロを設定するのと同じことです。起動メソッドが失敗した場合、RESTART_ONLY と LOG_ONLY は NONE と同じことです。つまり、フェイルオーバーは起こりません。リソースは start_failed 状態になります。

カテゴリ: 任意
デフォルト: デフォルトなし
調整: ANYTIME

Load_balancing_policy (文字列型)

使用する負荷均衡ポリシーを定義する文字列。このプロパティは、スケーラブルサービス専用です。RTR ファイルに Scalable プロパティが宣言されている場合、RGM は自動的にこのプロパティを作成します。Load_balancing_policy には次の値を設定できます。

Lb_weighted (デフォルト)。Load_balancing_weights プロパティで設定されているウエイトに従って、さまざまなノードに負荷が分散されます。

Lb_sticky。スケーラブルサービスの指定のクライアント (クライアントの IP アドレスで識別される) は、常に同じクラスタノードに送信されます。

Lb_sticky_wild。ワイルドスティッキーサービスの IP アドレスに接続する Lb_sticky_wild で指定されたクライアントの IP アドレスは、IP アドレスが到着するポート番号とは無関係に、常に同じクラスタノードに送られます。

カテゴリ: 条件付き/任意
デフォルト: Lb_weighted
調整: AT_CREATION

Load_balancing_weights (文字配列型)

このプロパティは、スケーラブルサービス専用です。RTR ファイルに Scalable プロパティが宣言されている場合、RGM は自動的にこのプロパティを作成します。形式は、「weight@node,weight@node」になります。ここで、weight は、指定したノード (node) に対する負荷分散の相対的な割り当てを示す整数になります。ノードに分散される負荷の割合は、すべてのウエイトの合計でこのノードのウエイトを割った値になります。たとえば、1@1,3@2 は、ノード 1 が負荷の 1/4 を受け取り、ノード 2 が 3/4 を受け取ることを指定します。空の文字列 ("") は、負荷を均等に分散することを意味します (デフォルト)。明示的にウエイトを割り当てられていないノードのウエイトは、デフォルトで 1 になります。

Tunable 属性がリソースタイプファイルに指定されていない場合は、プロパティの Tunable 値は ANYTIME (任意の時点) になります。このプロパティを変更すると、新しい接続時にのみ分散が変更されます。

カテゴリ: 条件付き/任意
デフォルト: 空の文字列 ("")
調整: ANYTIME

`method_timeout` (この「タイプ」のコールバックメソッドごとの) (整数型)
RGM がメソッドの呼び出しに失敗したと判断するまでの時間 (秒)。

カテゴリ: 条件付き/任意

デフォルト: RTR ファイルにメソッド自体が宣言されている場合は 3600 (1 時間)

調整: ANYTIME

`Monitored_switch` (列挙型)

クラスタ管理者が管理ユーティリティを使用してモニターを有効または無効にすると、RGM によって Enabled または Disabled に設定されます。Disabled に設定されると、再び有効に設定されるまで、モニターは Start メソッドを呼び出しません。リソースが、モニターのコールバックメソッドを持っていない場合は、このプロパティは存在しません。

カテゴリ: 照会のみ

デフォルト: デフォルトなし

調整: しない

`Network_resources_used` (文字配列型)

リソースが使用する論理ホスト名または共有アドレスネットワークリソースのリスト。スケーラブルサービスの場合、このプロパティは別のリソースグループに存在する共有アドレスリソースを参照する必要があります。フェイルオーバーサービスの場合、このプロパティは同じリソースグループに存在する論理ホスト名または共有アドレスを参照します。RTR ファイルに Scalable プロパティが宣言されている場合、RGM は自動的にこのプロパティを作成します。Scalable が RTR ファイルで宣言されていない場合、`Network_resources_used` は RTR ファイルで明示的に宣言されていない限り使用できません。

Tunable 属性がリソースタイプファイルに指定されていない場合は、プロパティの Tunable 値は、AT_CREATION (作成時) になります。

注 - CRNP を使用する場合のこのプロパティの設定方法については、SUNW.Event (5) のマニュアルページを参照してください。

カテゴリ: 条件付き/必須

デフォルト: デフォルトなし

調整: AT_CREATION

`Num_resource_restarts` (クラスタノードごとの) (整数型)

このプロパティは RGM によって設定されるため、ユーザーが直接設定することはできません。RGM は、`scha_control`、`Resource_restart`、または `Resource_is_restarted` 呼び出しが、過去 n 秒間にこのノードでこのリソースに対して行われた回数をこのプロパティに設定します。 n は、リソースの `Retry_interval` プロパティの値です。リソース再起動カウンタは、このリソ

スによって `scha_control` ギブオーバーが実行されるたびに、その成否とは関係なく RGM によってゼロにリセットされます。

リソース型が `Retry_interval` プロパティを宣言していない場合、この型のリソースに `Num_resource_restarts` プロパティを使用できません。

カテゴリ: 照会のみ
デフォルト: デフォルトなし
調整: No

Num_rg_restarts on each cluster node (整数型)

このプロパティは RGM によって設定されるため、ユーザーが直接設定することはできません。RGM は、`scha_control Restart` 呼び出しが、過去 n 秒間にこのノードでこのリソースによって、そのリソースを包含するリソースグループに対して行われた回数をこのプロパティに設定します。 n はこのリソースの `Retry_interval` プロパティの値です。リソース型が `Retry_interval` プロパティを宣言していない場合、この型のリソースに `Num_resource_restarts` プロパティを使用できません。

カテゴリ: 説明を参照
デフォルト: デフォルトなし
調整: No

On_off_switch (列挙型)

クラスタ管理者が管理ユーティリティを使用してリソースを有効または無効にすると、RGM によって `Enabled` または `Disabled` に設定されます。無効に設定されたリソースはオフラインにされます。再び有効にされるまでコールバックが呼び出されることはありません。

カテゴリ: 照会のみ
デフォルト: デフォルトなし
調整: しない

Port_list (文字配列型)

サーバーが待機するポートの番号リストです。ポート番号には、スラッシュ (/) と、そのポートで使用されるプロトコルが付加されます (たとえば、`Port_list=80/tcp` や `Port_list=80/tcp6,40/udp6` など)。プロトコルには、次のものを指定できます。

- `tcp` (TCP IPv4)
- `tcp6` (TCP IPv6)
- `udp` (UDP IPv4)
- `udp6` (UDP IPv6) `Scalable` プロパティが RTR ファイルで宣言されている場合、RGM は自動的に `Port_list` を作成します。それ以外の場合、このプロパティは RTR ファイルで明示的に宣言されていないかぎり使用できません。

Apache 用にこのプロパティを設定する方法は、『*Sun Cluster Data Service for Apache* ガイド (Solaris OS 版)』を参照してください。

カテゴリ: 条件付き/必須
デフォルト: デフォルトなし
調整: AT_CREATION

R_description (文字列)
リソースの簡単な説明。

カテゴリ: 任意
デフォルト: 空の文字列
調整: ANYTIME

Resource_dependencies (文字配列型)

このリソースが強く依存している、同じグループまたは異なるグループに属するリソースのリスト。このリソースを起動するためには、リストのすべてのリソースがオンラインになっていなければなりません。このリソースと、リストのリソースの1つが同時に起動されると、RGM は、リストのリソースが始動してからこのリソースを起動します。このリソースの Resource_dependencies リストにあるリソースが始動しない場合は、このリソースもオフラインのままです。リストのリソースが始動しない原因は、リストのリソースのリソースグループがオフラインであるか、リストのリソースが start_failed 状態にあるためである可能性があります。このリソースが、異なるリソースグループの始動しないリソースに依存していることが原因でオフラインのままになっている場合、このリソースのグループは Pending_online_blocked 状態になります。

このリソースが、リストのリソースと同時にオフラインにされる場合は、このリソースが停止されてから、リストのほかのリソースが停止されます。ただし、このリソースがオンラインのままであったり、停止に失敗した場合でも、異なるリソースグループに属するリストのリソースは停止されます。このリソースが先に無効にしなければ、リスト内のリソースは無効にできません。

同じリソースグループ内では、デフォルトとして、アプリケーションリソースがネットワークアドレスリソースに対して暗黙的に強いリソース依存性を持っています。詳細は、257 ページの「リソースグループのプロパティ」の Implicit_network_dependencies を参照してください。

同じリソースグループ内では、依存性の順序に従って Prenet_start メソッドが Start メソッドより先に実行されます。同様に Postnet_stop メソッドは Stop メソッドより後に実行されます。異なるリソースグループ内では、依存されるリソースが Prenet_start と Start を終了してから、依存するリソースが Prenet_start を実行します。同様に、依存するリソースが Stop と Postnet_stop を終了してから、依存されるリソースが Stop を実行します。

カテゴリ: 任意
デフォルト: 空のリスト
調整: ANYTIME

Resource_dependencies_restart (文字配列型)

このリソースが再起動に関して依存している、同じグループまたは異なるグループに属するリソースのリスト。

このプロパティの機能は、次の点を除き Resource_dependencies と同じです。つまり、このプロパティでは、再起動依存リストのリソースが再起動されると、このリソースも再起動されます。このリソースの再起動は、リストのリソースがオンラインに復帰した後で行われます。

カテゴリ: 任意

デフォルト: 空のリスト

調整: ANYTIME

Resource_dependencies_weak (文字配列型)

このリソースによる依存性が低い、同じグループまたは異なるグループに属するリソースのリスト。低い依存性によってメソッドの呼び出し順序が決まります。

RGM は、リストのリソースの start メソッドを呼び出してからこのリソースの start メソッドを呼び出します。さらに、RGM は、このリソースの stop メソッドを呼び出してからリストのリソースの stop メソッドを呼び出します。リスト内のリソースが始動に失敗したりオフラインのままであっても、リソースは起動されます。

このリソースと、Resource_dependencies_weak リストのリソースが同時に起動されると、RGM は、リストのリソースが始動してからこのリソースを起動します。リストのリソースが始動しない場合でも (たとえば、リストのリソースのリソースグループがオフラインのままであったり、リストのリソースが start_failed 状態である場合でも)、このリソースは起動されます。このリソースのリソースグループは、このリソースの Resource_dependencies_weak リストのリソースが始動する間、一時的に pending_online_blocked 状態になることがあります。リストのすべてのリソースが始動するか始動に失敗すると、このリソースが起動され、そのグループが再び pending_online 状態になります。

このリソースが、リストのリソースと同時にオフラインにされる場合は、このリソースが停止されてから、リストのほかのリソースが停止されます。このリソースがオンラインのままであったり、停止に失敗した場合でも、リストのリソースは停止されます。リストのリソースを無効にするためには、このリソースをまず無効にする必要があります。

同じリソースグループ内では、依存性の順序に従って prenet_start メソッドが start メソッドより先に実行されます。同様に postnet_stop メソッドは stop メソッドより後に実行されます。異なるリソースグループ (複数) 内では、依存されるリソースが prenet_start と start を終了してから、依存するリソースが prenet_start を実行します。同様に、依存するリソースが stop と postnet_stop を終了してから、依存されるリソースが stop を実行します。

カテゴリ: 任意

デフォルト: 空のリスト

調整: ANYTIME

Resource_name (文字列)

リソースインスタンスの名前です。この名前はクラスタ構成内で一意にする必要があります。リソースが作成されたあとで変更はできません。

カテゴリ: 必須
デフォルト: デフォルトなし
調整: しない

Resource_project_name (文字列)

リソースに関連付けられた Solaris プロジェクト名。このプロパティは、CPU の共有、クラスタデータサービスのリソースプールといった Solaris のリソース管理機能に適用できます。RGM は、リソースをオンラインにすると、このプロジェクト名を持つ関連プロセスを起動します。このプロパティが指定されていない場合、プロジェクト名は、リソースが属しているリソースグループの `RG_project_name` プロパティから取得されます。`rg_properties(5)` を参照してください。どちらのプロパティも指定されなかった場合、RGM は事前定義済みのプロジェクト名 `default` を使用します。プロジェクトデータベース内に存在するプロジェクト名を指定する必要があります。また、`root` ユーザーは、このプロジェクトのメンバーとして構成されている必要があります。このプロパティは、Solaris 9 以降のバージョンでサポートされます。

注 - このプロパティへの変更を有効にするためには、リソースを起動し直す必要があります。

カテゴリ: 任意
デフォルト: Null
調整: ANYTIME

Resource_state (各クラスタノードの) (列挙型)

RGM が判断した各クラスタノード上のリソースの状態。この状態には、`Online`、`Offline`、`Start_failed`、`Stop_failed`、`Monitor_failed`、`Online_not_monitored`、`Starting`、`Stopping` があります。

ユーザーはこのプロパティを構成できません。

カテゴリ: 照会のみ
デフォルト: デフォルトなし
調整: しない

Retry_count (整数型)

起動に失敗したリソースをモニターが再起動する回数です。このプロパティは RGM によって作成されます。RTR ファイルに宣言されている場合にかぎり、管理者は使用を許可されます。デフォルト値が RTR ファイルに指定されている場合、`Retry_count` の指定は任意です。

リソース型ファイル内に Tunable 属性が指定されていない場合、このプロパティの Tunable 値は WHEN_DISABLED になります。

このプロパティが RTR ファイルに宣言され、かつ Default 属性が指定されていない場合、このプロパティは必須です。

カテゴリ: 条件付き
デフォルト: デフォルトなし
調整: WHEN_DISABLED

Retry_interval (整数型)

失敗したリソースを再起動するまでの秒数。リソースモニターは、このプロパティと Retry_count を組み合わせて使用します。このプロパティは RGM によって作成されます。RTR ファイルに宣言されている場合にかぎり、管理者は使用を許可されます。デフォルト値が RTR ファイルに指定されている場合、Retry_interval の指定は任意です。

リソース型ファイル内に Tunable 属性が指定されていない場合、このプロパティの Tunable 値は WHEN_DISABLED になります。

このプロパティが RTR ファイルに宣言され、かつ Default 属性が指定されていない場合、このプロパティは必須です。

カテゴリ: 条件付き
デフォルト: デフォルトなし
調整: WHEN_DISABLED

Scalable (ブール型)

リソースがスケーラブルであるかどうか、つまり、リソースが Sun Cluster の ネットワーキング負荷分散機能を使用するかどうかを表します。

このプロパティが RTR ファイルで宣言されている場合は、そのタイプのリソースに対して、RGM は、次のスケーラブルサービスプロパティを自動的に作成します。Affinity_timeout、Load_balancing_policy、Load_balancing_weights、Network_resources_used、Port_list、UDP_affinity、Weak_affinity。これらのプロパティは、RTR ファイル内で明示的に宣言されない限り、デフォルト値を持ちます。RTR ファイルに Scalable が宣言されている場合、このプロパティのデフォルトは TRUE です。

RTR ファイルにこのプロパティが宣言されている場合、AT_CREATION 以外の Tunable 属性の割り当ては許可されません。

RTR ファイルにこのプロパティが宣言されていない場合、このリソースはスケーラブルではないため、このプロパティを調整することはできません。RGM は、スケーラブルサービスプロパティをいっさい設定しません。ただし、RTR ファイルに明示的に Network_resources_used および Port_list プロパティを宣言できます。これらのプロパティは、スケーラブルサービスだけでなく、非スケーラブルサービスでも有用です。

このリソースプロパティと Failover リソースタイププロパティの組み合わせ、およびその説明は次の通りです。

このリソースプロパティと Failover リソースタイププロパティを併用する場合には、`r_properties` (5) のマニュアルページを参照してください。

カテゴリ: 任意
デフォルト: デフォルトなし
調整: AT_CREATION

Status (各クラスターノードの) (列挙型)

リソースモニターが `scha_resource_setstatus` (1HA) か `scha_resource_setstatus` (3HA) を使って設定します。指定可能な値は、OK、degraded、faulted、unknown、および offline です。リソースがオンラインまたはオフラインにされると、RGM は Status 値を自動的に設定します。ただし、Status 値がリソースのモニターやメソッドによって設定される場合を除きます。

カテゴリ: 照会のみ
デフォルト: デフォルトなし
調整: しない

Status_msg (各クラスターノードの) (文字列型)

リソースモニターによって、Status プロパティと同時に設定されます。リソースがオンラインまたはオフラインにされると、RGM は自動的にこのプロパティを空文字列でリセットします。ただし、このプロパティがリソースのメソッドによって設定される場合を除きます。

カテゴリ: 照会のみ
デフォルト: デフォルトなし
調整: しない

Thorough_probe_interval (整数型)

高オーバーヘッドのリソース障害検証の呼び出し間隔 (秒)。このプロパティは RGM によって作成されます。RTR ファイルに宣言されている場合にかぎり、管理者は使用を許可されます。デフォルト値が RTR ファイルに指定されている場合、Thorough_probe_interval の指定は任意です。

リソース型ファイル内に Tunable 属性が指定されていない場合、このプロパティの Tunable 値は WHEN_DISABLED になります。

RTR ファイルのプロパティ宣言内に Default 属性が指定されていない場合、このプロパティは必須です。

カテゴリ: 条件付き
デフォルト: デフォルトなし
調整: WHEN_DISABLED

Type (文字列型)

このリソースがインスタントであるリソースタイプ。

カテゴリ: 必須
デフォルト: デフォルトなし
調整: しない

Type_version (文字列型)

現在このリソースに関連付けられているリソース型のバージョンを指定します。このプロパティは RTR ファイル内に宣言できません。したがって、RGM によって自動的に作成されます。このプロパティの値は、リソースタイプの RT_version プロパティと等しくなります。リソースの作成時、Type_version プロパティはリソースタイプ名の接尾辞として表示されるだけで、明示的には指定されません。リソースを編集する際に Type_version を新しい値に変更できます。

このプロパティの調整については、次の情報から判断されます。

- 現在のリソースタイプのバージョン
- RTR ファイル内の #`$upgrade_from` ディレクティブ

カテゴリ: 説明を参照
デフォルト: デフォルトなし
調整: 説明を参照

UDP_affinity (ブール型)

`true` の場合、指定のクライアントからの UDP トラフィックはすべて現在クライアントの TCP トラフィックを処理しているサーバーノードに送信されます。

このプロパティは、Load_balancing_policy が Lb_sticky または Lb_sticky_wild の場合にかぎり有効です。さらに、Weak_affinity を FALSE (デフォルト値) に設定されている必要があります。

このプロパティは、スケーラブルサービス専用です。

カテゴリ: 任意
デフォルト: デフォルトなし
調整: WHEN_DISABLED

Weak_affinity (ブール型)

`true` の場合、弱い形式のクライアントアフィニティが有効になります。弱い形式のクライアントアフィニティが有効になっている場合、特定のクライアントからの接続は、次の場合を除き、同じサーバーノードに送信されます。

- 障害モニターの再起動や、リソースのフェイルオーバーまたはスイッチオーバー、障害後のノードとクラスタの再結合などが行われたことに対応してサーバーリスナーが起動する。
- 管理アクションによってスケーラブルリソースの Load_balancing_weights が変更される。

弱いアフィニティはメモリーの消費とプロセッササイクルの点で、デフォルトの形式よりもオーバーヘッドを低く抑えられます。

このプロパティは、Load_balancing_policy が Lb_sticky または Lb_sticky_wild の場合にかぎり有効です。

このプロパティは、スケーラブルサービス専用です。

カテゴリ: 任意

デフォルト: デフォルトなし

調整: WHEN_DISABLED

リソースグループのプロパティ

以下に、Sun Cluster によって定義されるリソースグループプロパティを示します。プロパティ値は、「カテゴリ」で次のように分類されます。

- 必須 — 管理者は、管理ユーティリティでリソースグループを作成するときに、必ず値を指定する必要があります。
- 任意 — 管理者がリソースグループの作成時に値を指定しない場合、システムがデフォルト値を提供します。
- 照会のみ — 管理ツールから直接設定できません。

各説明では、初期設定後に、そのプロパティが更新可能 (Yes) か更新不可能 (No) かを示します。

以下にプロパティ名とその説明を示します。

Auto_start_on_new_cluster (ブール型)

このプロパティを使用すると、新しいクラスタを形成するとき、リソースグループの自動起動を無効にすることができます。

TRUE の場合、クラスタが再起動するとき、Resource Group Manager はリソースグループを自動的に起動して、Desired primaries を有効にしようと試みます。FALSE に設定されている場合、クラスタのすべてのノードが同時に再起動したとき、Resource Group Manager はリソースグループを自動的に起動しません。

カテゴリ: 任意

デフォルト: TRUE

調整: Yes

Desired primaries (整数型)

グループが同時にオンラインになることができるノードの数。

RG_mode プロパティが Failover の場合、このプロパティの値を 1 より大きく設定することはできません。RG_mode プロパティが Scalable の場合は、1 より大きな値を設定できます。

カテゴリ: 任意

デフォルト: 1

調整: Yes

Failback (ブール型)

クラスタのメンバーシップが変更されたとき、グループがオンラインになっているノードセットを再計算するかどうかを示すブール値です。再計算により、RGM は優先度の低いノードをオフラインにし、優先度の高いノードをオンラインにすることができます。

カテゴリ: 任意

デフォルト: FALSE

調整: Yes

Global_resources_used (文字配列型)

クラスタファイルシステムがこのリソースグループ内のリソースによって使用されるかどうかを指定します。管理者はアスタリスク (*) か空文字列 ("") を指定できます。すべてのグローバルリソースを指定するときはアスタリスク、グローバルリソースを一切指定しない場合は空文字列を指定します。

カテゴリ: 任意

デフォルト: すべてのグローバルリソース

調整: Yes

Implicit_network_dependencies (ブール型)

TRUE の場合、RGM は、グループ内のネットワークアドレスリソースで非ネットワークアドレスリソースに対する強い依存を強制します。ネットワークアドレスリソースには、論理ホスト名と共有アドレスリソース型があります。

スケラブルリソースグループの場合、ネットワークアドレスリソースを含んでいないため、このプロパティは効果がありません。

カテゴリ: 任意

デフォルト: TRUE

調整: Yes

Maximum primaries (整数型)

グループを同時にオンラインにできるノードの最大数です。

RG_mode プロパティが Failover の場合、このプロパティの値を 1 より大きく設定することはできません。RG_mode プロパティが Scalable の場合は、1 より大きな値を設定できます。

カテゴリ: 任意

デフォルト: 1

調整: Yes

Nodelist (文字配列型)

優先順位に従ってグループをオンラインにできるクラスタノードのリスト。これらのノードは、リソースグループの潜在的な主ノードまたはマスターです。

カテゴリ: 任意

デフォルト: すべてのクラスタノードのリスト

調整: Yes

Pathprefix (文字配列型)

リソースグループ内のリソースが重要な管理ファイルを書き込むことができるクラスタファイルシステム内のディレクトリ。一部のリソースの必須プロパティです。Pathprefix の値はリソースグループごとに固有の値を指定します。

カテゴリ: 任意

デフォルト: 空の文字列

調整: Yes

Pingpong_interval (整数型)

負数ではない整数値 (秒)。RGM は、この値を使って、リソースグループをどこでオンラインにするかを決めます。このプロパティが必要になることがある条件には、次のものがあります。

- 再構成が行われる場合
- `scha_control -O GIVEOVER` コマンドが実行されるか、`scha_control()` 関数が `SCHA GIVEOVER` 引数で実行される場合。再構成が行われる場合、リソースグループが特定のノードで `Pingpong_interval` 秒内に 2 回以上オンラインにならないことがあると、そのノードは、このリソースグループのホストとしての資格がないものと見なされ、RGM は別のマスターを探します。リソースグループがオンラインにならない理由は、リソースの `Start` または `Prenet_start` メソッドが非ゼロで終わったか、タイムアウトになったためです。

リソースの `scha_control` コマンドまたは関数の呼び出しによって、`Pingpong_interval` で指定した秒数内に特定のノード上でリソースグループがオフラインになった場合、別のノードから生じる後続の `scha_control()` 呼び出しの結果、そのノードはリソースグループのホストとして不適切だと判断されます。

カテゴリ: 任意

デフォルト: 3600 (1 時間)

調整: Yes

Resource_list (文字配列型)

グループに含まれるリソースのリスト。管理者はこのプロパティを直接設定しません。このプロパティは、管理者がリソースグループにリソースを追加したり、リソースを削除したときに、RGM によって更新されます。

カテゴリ: 照会のみ
デフォルト: デフォルトなし
調整: No

RG_affinities (文字列型)

RGM は、リソースグループを、別のあるリソースグループの現行マスターであるノードに求める (ポジティブアフィニティ) か、リソースグループを、あるリソースグループの現行マスターではないノードに求めようとしています (ネガティブアフィニティ)。

RG_affinities には次の文字列を設定できます。

- ++ (強いポジティブアフィニティ)
- + (弱いポジティブアフィニティ)
- - (弱いネガティブアフィニティ)
- -- (強いネガティブアフィニティ)
- +++ (フェイルオーバーの権限を委譲された強いポジティブアフィニティ) たとえば、RG_affinities=+RG2,--RG3 は、このリソースグループが RG2 に対して弱いポジティブアフィニティを、RG3 に対して強いネガティブアフィニティをもつことを表しています。

RG_affinities の使い方については、『Sun Cluster データサービスの計画と管理 (Solaris OS 版)』の「データサービスリソースの管理」を参照してください。

カテゴリ: 任意
デフォルト: 空の文字列
調整: Yes

RG_dependencies (文字配列型)

リソースグループのリスト (任意)。このリストは、同じノードでほかのグループをどのような順序でオンラインまたはオフラインにしたいかを表しています。すべての強い RG_affinities (ポジティブおよびネガティブ) と RG_dependencies の関係図式の中に循環が含まれてはなりません。

たとえば、リソースグループ RG1 の RG_dependencies リストにリソースグループ RG2 が含まれているとします。つまり、RG1 には RG2 に対するリソースグループの依存性があるものとします。次の各項は、このリソースグループの依存性がどのような結果をもたらすかを説明したものです。

- ノードがクラスタに結合されると、そのノードでは、RG2 のすべてのリソースに対する Boot メソッドが終わってから、RG1 のリソースに対する Boot メソッドが実行されます。
- RG1 と RG2 が共に同じノードで同時に Pending_online 状態にある場合は、RG2 のすべてのリソースに対する起動メソッド (Prenet_start または Start) が終了してから、RG1 のリソースに対する起動メソッドが実行されます。
- RG1 と RG2 が共に同じノードで同時に Pending_offline 状態にある場合は、RG2 のすべてのリソースに対する停止メソッド (Stop または Postnet_stop) が終了してから、RG1 のリソースに対する停止メソッドが実行されます。

行されます。

- RG1 または RG2 の主ノードをスイッチする場合、それによって RG1 がいずれかのノードでオンラインに、RG2 がすべてのノードでオフラインになる場合は、このスイッチは失敗します。詳細は、scswitch(1M) と scsetup(1M) のマニュアルページを参照してください。
- RG2 に対する Desired primaries がゼロに設定されている場合は、RG1 に対する Desired primaries プロパティをゼロより大きい値に設定することはできません。
- RG2 に対する Auto_start_on_new_cluster が FALSE に設定されている場合は、RG1 に対する Auto_start_on_new_cluster プロパティを TRUE に設定することはできません。

カテゴリ: 任意

デフォルト: 空のリスト

調整: Yes

RG_description (文字列型)

リソースグループの簡単な説明。

カテゴリ: 任意

デフォルト: 空の文字列

調整: Yes

RG_is_frozen (ブール型)

あるリソースグループが依存している大域デバイスをスイッチオーバーするかどうかを表します。このプロパティが TRUE に設定されている場合、大域デバイスはスイッチオーバーされます。このプロパティが FALSE に設定されている場合、大域デバイスはスイッチオーバーされません。リソースグループが大域デバイスに依存するかどうかは、Global_resources_used プロパティの設定によります。

RG_is_frozen プロパティをユーザーが直接設定することはありません。

RG_is_frozen プロパティは、大域デバイスのステータスが変わったときに、RGM によって更新されます。

カテゴリ: 任意

デフォルト: デフォルトなし

調整: No

RG_mode (列挙型)

リソースグループがフェイルオーバーグループかスケラブルグループかを指定します。値が Failover の場合、RGM はグループの Maximum primaries プロパティの値を 1 に設定し、リソースグループのマスターを単一のノードに制限します。

このプロパティの値が Scalable の場合、RGM は Maximum primaries プロパティに 1 より大きい値を設定することを許可します。その結果、グループを複数のノードで同時にマスターできます。Failover プロパティの値が TRUE のリソースを、RG_mode の値が Scalable のリソースグループに追加することはできません。

Maximum primaries が 1 の場合、デフォルトは Failover です。
Maximum primaries が 1 より大きい場合、デフォルトは Scalable です。

カテゴリ: 任意
デフォルト: Maximum primaries の値によります。
調整: No

RG_name (文字列型)

リソースグループの名前。この名前は、クラスタ内で一意である必要があります。

カテゴリ: 必須
デフォルト: デフォルトなし
調整: No

RG_project_name (文字列型)

リソースグループに関連付けられた Solaris プロジェクト名。このプロパティは、CPU の共有、クラスタデータサービスのリソースプールといった Solaris のリソース管理機能に適用できます。RGM は、リソースグループをオンラインにするとき、Resource_project_name プロパティセットを持たないリソースに対して、このプロジェクトで関連付けられたプロセスを起動します。プロジェクトデータベース内に存在するプロジェクト名を指定する必要があります。また、root ユーザーは、このプロジェクトのメンバーとして構成されている必要があります。

このプロパティは、Solaris 9 以降のバージョンでサポートされます。

注 - このプロパティへの変更を有効にするためには、リソースを起動し直す必要があります。

カテゴリ: 任意
デフォルト: 文字列 "default"
調整: ANYTIME

各クラスタノード上の RG_state (列挙型)

RGM によって Unmanaged、Online、Offline、Pending_online、Pending_offline、Pending_online_blocked、Error_stop_failed、Online_faulted、または Pending_online_blocked に設定されます。これは、そのグループが各クラスタノードでどのような状態にあるかを表します。

ユーザーはこのプロパティを構成できません。ただし、`scswitch(1M)` を呼び出すか、これと同等の `scsetup(1M)` または `SunPlex Manager` コマンドを使用すれば、このプロパティをユーザーが間接的に設定することができます。

カテゴリ: 照会のみ

デフォルト: デフォルトなし

調整: No

RG_system (ブール型)

リソースグループの `RG_system` プロパティの値が `TRUE` の場合、そのリソースグループとそのリソースグループ内のリソースに関する特定の操作が制限されます。この制限は、重要なリソースグループやリソースを間違えて変更または削除してしまうことを防ぐためにあります。このプロパティによって影響を受けるのは、`scrgadm(1M)` と `scswitch(1M)` コマンドだけです。`scha_control(1HA)` と `scha_control(3HA)` の操作には影響を与えません。

リソースグループ (またはリソースグループ内のリソース) の制限操作を実行する前には、まず、リソースグループの `RG_system` プロパティを `FALSE` に設定する必要があります。クラスタサービスをサポートするリソースグループ (または、リソースグループ内のリソース) を変更または削除するときには注意してください。

`RG_system` の値が `TRUE` に設定されているリソースグループをシステムリソースグループといいます。 `RG_system` の現在の値が何であれ、このプロパティ自体の編集は制限されていません。制限については、`rg_properties(5)` のマニュアルページを参照してください。

カテゴリ: 任意

デフォルト: FALSE

調整: Yes

リソースプロパティの属性

次の情報は、システム定義のプロパティの変更または拡張プロパティの作成に使用できるリソースプロパティ属性を示したものです。



注意 - `boolean`、`enum`、`int` タイプのデフォルト値に、`Null` または空の文字列 ("") は指定できません。

以下にプロパティ名とその説明を示します。

Array_maxsize
stringarray タイプの場合、設定できる配列要素の最大数。

Array_minsize
stringarray タイプの場合、設定できる配列要素の最小数。

Default
プロパティのデフォルト値を示します。

Description
プロパティを簡潔に記述した注記 (文字列)。RTR ファイル内でシステム定義プロパティに対する Description 属性を設定することはできません。

Enumlist
enum タイプの場合、プロパティに設定できる文字列値のセット。

Extension
リソースタイプの実装によって定義された拡張プロパティが RTR ファイルのエントリで宣言されていることを示します。拡張プロパティが使用されていない場合、そのエントリはシステム定義プロパティです。

Max
int タイプの場合、プロパティに設定できる最大値。

Maxlength
string および stringarray タイプの場合、設定できる文字列の最大。

Min
int タイプの場合、プロパティに設定できる最小値。

Minlength
string および stringarray タイプの場合、設定できる文字列の最小長。

Property
リソースプロパティの名前。

Tunable
クラスタ管理者がリソースのプロパティ値をいつ設定できるかを示します。管理者にプロパティの設定を許可しない場合は、NONE または FALSE に設定します。管理者にプロパティの調整を許可する属性値は、次のとおりです。TRUE または ANYTIME (任意の時点)、AT_CREATION (リソースの作成時のみ)、WHEN_DISABLED (リソースがオフラインのとき)。「監視をいつ無効にするか」や「いつオフラインにするか」など、ほかの調整条件を設定する場合は、この属性に ANYTIME を設定し、Validate メソッドを使ってリソースの状態を検証します。

デフォルトは、個々の標準的なリソースプロパティによって異なります (次節を参照)。拡張プロパティの調整に関するデフォルト設定は、RTR ファイルに別の指定がある場合を除き、TRUE (ANYTIME) です。

プロパティのタイプ

指定可能な型は、string、boolean、int、enum、stringarray です。RTR ファイル内で、システム定義プロパティの型の属性を設定することはできません。タイプは、RTR ファイルのエントリに登録できる、指定可能なプロパティ値とタイ

プ固有の属性を決定します。enumタイプは、文字列値のセットです。

付録 B

データサービスのコード例

この付録では、データサービスの各メソッドの完全なコード例を示します。また、リソースタイプ登録 (RTR) ファイルの内容も示します。

この付録に含まれるコードリストは、次のとおりです。

- 267 ページの「リソースタイプ登録ファイルのリスト」
- 270 ページの「Start メソッド」
- 273 ページの「Stop メソッド」
- 275 ページの「gettime ユーティリティ」
- 276 ページの「PROBE プログラム」
- 282 ページの「Monitor_start メソッド」
- 284 ページの「Monitor_stop メソッド」
- 285 ページの「Monitor_check メソッド」
- 287 ページの「Validate メソッド」
- 291 ページの「Update メソッド」

リソースタイプ登録ファイルのリスト

リソースタイプ登録 (RTR) ファイルには、クラスタ管理者がデータサービスを登録するとき、データサービスの初期構成を定義するリソースとリソースタイプのプロパティ宣言が含まれています。

例 B-1 SUNW.Sample RTR ファイル

```
#
# Copyright (c) 1998-2004 by Sun Microsystems, Inc.
# All rights reserved.
#
# ドメインネームサービス (DNS) の登録情報
#
```

例 B-1 SUNW.Sample RTR ファイル (続き)

```
#pragma ident    "@(#)SUNW.sample  1.1  00/05/24 SMI"

RESOURCE_TYPE = "sample";
VENDOR_ID = SUNW;
RT_DESCRIPTION = "Domain Name Service on Sun Cluster";

RT_VERSION = "1.0";
API_VERSION = 2;
FAILOVER = TRUE;

RT_BASEDIR=/opt/SUNWsample/bin;
PKGLIST = SUNWsample;

START          = dns_svc_start;
STOP           = dns_svc_stop;

VALIDATE       = dns_validate;
UPDATE        = dns_update;

MONITOR_START  = dns_monitor_start;
MONITOR_STOP   = dns_monitor_stop;
MONITOR_CHECK  = dns_monitor_check;

# リソースタイプ宣言のあとに、中括弧に囲まれたリソースプロパティ宣言
# のリストが続く。プロパティ名宣言は、各エントリの左中括弧の直後にある
# 最初の属性である必要がある。
#
# <method>_timeout プロパティは、RGM がメソッド呼び出しが失敗
# したという結論を下すまでの時間 (秒) を設定する。
#
# すべてのメソッドタイムアウトの MIN 値は、60 秒に設定されている。
# これは、管理者が短すぎる時間を設定することを防ぐためである。短すぎる
# 時間を設定するとスイッチオーバーやフェイルオーバーの性能が上がらず、
# さらには、予期せぬ RGM アクションが発生する可能性がある (間違った
# フェイルオーバー、ノードの再起動、リソースグループの
# ERROR_STOP_FAILED 状態への移行、オペレータの介入の必要性など)。
# メソッドタイムアウトに短すぎる時間を設定すると、データサービス全体
# の可用性を下げることになる。
#
{
    PROPERTY = Start_timeout;
    MIN=60;
    DEFAULT=300;
}

{
    PROPERTY = Stop_timeout;
    MIN=60;
    DEFAULT=300;
}

{
    PROPERTY = Validate_timeout;
```

例 B-1 SUNW.Sample RTR ファイル (続き)

```
        MIN=60;
        DEFAULT=300;
    }
    {
        PROPERTY = Update_timeout;
        MIN=60;
        DEFAULT=300;
    }
    {
        PROPERTY = Monitor_Start_timeout;
        MIN=60;
        DEFAULT=300;
    }
    {
        PROPERTY = Monitor_Stop_timeout;
        MIN=60;
        DEFAULT=300;
    }
    {
        PROPERTY = Thorough_Probe_Interval;
        MIN=1;
        MAX=3600;
        DEFAULT=60;
        TUNABLE = ANYTIME;
    }
}

# 当該ノード上でアプリケーションを正常に起動できないと結論を下すまで
# に、ある期間 (Retry_Interval) に行う再試行の回数。
{
    PROPERTY = Retry_Count;
    MIN=0;
    MAX=10;
    DEFAULT=2;
    TUNABLE = ANYTIME;
}

# Retry_Interval には 60 の倍数を設定する。これは、秒から分に変換さ
# れ、端数が切り上げられるためである。たとえば、50 (秒) という値を
# 指定すると、1 分に変換される。
# このプロパティは再試行数 (Retry_Count) のタイミングを決定する。
{
    PROPERTY = Retry_Interval;
    MIN=60;
    MAX=3600;
    DEFAULT=300;
    TUNABLE = ANYTIME;
}

{
    PROPERTY = Network_resources_used;
    TUNABLE = AT_CREATION;
    DEFAULT = "";
}
}
```

例 B-1 SUNW.Sample RTR ファイル (続き)

```
#
# 拡張プロパティ
#
# クラスタ管理者はこのプロパティの値を設定して、アプリケーションが使用
# する構成ファイルが入っているディレクトリを示す必要がある。このアプリ
# ケーションの場合、DNS は PXFS (通常は named.conf) 上の DNS 構成ファ
# イルのパスを指定する。
{
    PROPERTY = Confdir;
    EXTENSION;
    STRING;
    TUNABLE = AT_CREATION;
    DESCRIPTION = "The Configuration Directory Path";
}

# 検証が失敗したと宣言するまでのタイムアウト値 (秒)
{
    PROPERTY = Probe_timeout;
    EXTENSION;
    INT;
    DEFAULT = 30;
    TUNABLE = ANYTIME;
    DESCRIPTION = "Time out value for the probe (seconds)";
}
```

Start メソッド

データサービスリソースを含むリソースグループがオンラインになったとき、あるいは、リソースが有効になったとき、RGM はそのクラスタノード上で Start メソッドを呼び出します。サンプルのアプリケーションでは、Start メソッドはそのノード上で in.named (DNS) デーモンを起動します。

例 B-2 dns_svc_start メソッド

```
#!/bin/ksh
#
# HA-DNS の Start メソッド
#
# このメソッドは PMF の制御下でデータサービスを起動する。DNS の
# in.named プロセスを起動する前に、いくつかの妥当性検査を実行する。
# データサービスの PMF タグは $RESOURCE_NAME.named である。
# PMF は、指定された回数 (Retry_count) だけ、サービスを起動しようとす
# る。そして、指定された期間 (Retry_interval) 内で試行回数がこの値を超えた
# 場合、PMF はサービスの起動に失敗したことを報告する。
# Retry_count と Retry_interval は両方とも RTR ファイルに設定されて
```

例 B-2 dns_svc_start メソッド (続き)

いるリソースプロパティである。

```
#pragma ident    "@(#)dns_svc_start    1.1    00/05/24 SMI"
```

```
#####
```

```
# プログラム引数を構文解析する。
```

```
#
```

```
function parse_args # [args ...]
```

```
{
```

```
    typeset opt
```

```
    while getopts 'R:G:T:' opt
```

```
    do
```

```
        case "$opt" in
```

```
        R)
```

```
            # DNS リソース名。
```

```
            RESOURCE_NAME=$OPTARG
```

```
            ;;
```

```
        G)
```

```
            # リソースが構成されているリソース
```

```
            # グループの名前。
```

```
            RESOURCEGROUP_NAME=$OPTARG
```

```
            ;;
```

```
        T)
```

```
            # リソースタイプ名。
```

```
            RESOURCETYPE_NAME=$OPTARG
```

```
            ;;
```

```
        *)
```

```
            logger -p ${SYSLOG_FACILITY}.err \
```

```
            -t [${RESOURCETYPE_NAME},${RESOURCEGROUP_NAME},${RESOURCE_NAME}] \
```

```
            "ERROR: Option $OPTARG unknown"
```

```
            exit 1
```

```
            ;;
```

```
        esac
```

```
    done
```

```
}
```

```
#####
```

```
# MAIN
```

```
#
```

```
#####
```

```
export PATH=/bin:/usr/bin:/usr/cluster/bin:/usr/sbin:/usr/proc/bin:$PATH
```

```
# メッセージの記録に使用する syslog 機能番号を取得する。
```

例 B-2 dns_svc_start メソッド (続き)

```
SYSLOG_FACILITY=`scha_cluster_get -O SYSLOG_FACILITY`

# このメソッドに渡された引数を構文解析する。
parse_args "$@"

PMF_TAG=$RESOURCE_NAME.named
SYSLOG_TAG=$RESOURCE_TYPE_NAME,$RESOURCEGROUP_NAME,$RESOURCE_NAME

# DNS を起動するため、リソースの Confdir プロパティの値を取得する。
# 入力されたリソース名とリソースグループを使用して、リソースを
# 追加するときにクラスタ管理者が設定した Confdir の値を見つける。
config_info=scha_resource_get -O Extension -R $RESOURCE_NAME \
-G $RESOURCEGROUP_NAME Confdir`
# scha_resource_get は拡張プロパティの「タイプ」と「値」を戻す。
# 拡張プロパティの値だけを取得する。
CONFIG_DIR=`echo $config_info | awk '{print $2}'`

# Check if $CONFIG_DIR がアクセス可能であるかどうかを検査する。
if [ ! -d $CONFIG_DIR ]; then
    logger -p ${SYSLOG_FACILITY}.err -t [$SYSLOG_TAG] \
        "${ARGV0} Directory $CONFIG_DIR missing or not mounted"
    exit 1
fi

# データファイルへの相対パスが存在する場合、$CONFIG_DIR
# ディレクトリに移動する。
cd $CONFIG_DIR

# named.conf ファイルが $CONFIG_DIR ディレクトリ内に存在するか
# を検査する。
if [ ! -s named.conf ]; then
    logger -p ${SYSLOG_FACILITY}.err -t [$SYSLOG_TAG] \
        "${ARGV0} File $CONFIG_DIR/named.conf is missing or empty"
    exit 1
fi

# RTR ファイルから Retry_count の値を取得する。
RETRY_CNT=`scha_resource_get -O Retry_Count -R $RESOURCE_NAME \
-G $RESOURCEGROUP_NAME`

# RTR ファイルから Retry_interval の値を取得する。この値の単位は秒
# であり、pmfadm に渡すときは分に変換する必要がある。変換時、端数は
# 切り上げられるので注意すること。たとえば、50 秒は 1 分に切り上げられる。
((RETRY_INTRVAL = `scha_resource_get -O Retry_Interval -R $RESOURCE_NAME \
-G $RESOURCEGROUP_NAME` 60))

# PMF の制御下で in.named デーモンを起動する。$RETRY_INTERVAL の期
# 間、$RETRY_COUNT の回数だけ、クラッシュおよび再起動できる。どちら
# かの値以上クラッシュした場合、PMF は再起動を中止する。
# というタグですすでにプロセスが登録されて
# いる場合、PMF はすでにプロセスが動作していることを示す警告メッセ
# ージを送信する。
echo "Retry interval is "$RETRY_INTRVAL
```


例 B-2 dns_svc_start メソッド (続き)

```
pmfadm -c $PMF_TAG.named -n $RETRY_CNT -t $RETRY_INTRVAL \  
  /usr/sbin/in.named -c named.conf  
  
# HA-DNS が起動していることを示すメッセージを記録する。  
if [ $? -eq 0 ]; then  
    logger -p ${SYSLOG_FACILITY}.info -t [${SYSLOG_TAG}] \  
        "${ARGV0} HA-DNS successfully started"  
fi  
exit 0
```

Stop メソッド

HA-DNS リソースを含むリソースグループがクラスタノード上でオフラインになるとき、あるいは、HA-DNS リソースが無効になるとき、RGM は Stop メソッドを呼び出します。このメソッドは、そのノード上で in.named (DNS) デーモンを停止します。

例 B-3 dns_svc_stop メソッド

```
#!/bin/ksh  
#  
# HA-DNS の Stop メソッド  
#  
# このメソッドは、PMF を使用するデータサービスを停止する。サービス  
# が動作していない場合、このメソッドは状態 0 で終了する。その他の値  
# は戻さない。リソースは STOP_FAILED 状態になる。  
  
#pragma ident  "@(#)dns_svc_stop 1.1 00/05/24 SMI"  
  
#####  
# プログラム引数を構文解析する。  
#  
function parse_args # [args ...]  
{  
    typeset opt  
  
    while getopts `R:G:T:` opt  
    do  
        case "$opt" in  
            R)  
                # DNS リソースの名前。  
                RESOURCE_NAME=$OPTARG  
                ;;  
            G)  
                # リソースが構成されているリソース  
                # グループの名前。  
                ;;  
        esac  
    done  
}
```

例 B-3 dns_svc_stop メソッド (続き)

```

        RESOURCEGROUP_NAME=$OPTARG
        ;;
T)
    # リソースタイプの名前。
    RESOURCETYPE_NAME=$OPTARG
    ;;
*)
    logger -p ${SYSLOG_FACILITY}.err \
    -t [$RESOURCECETYPE_NAME,$RESOURCEGROUP_NAME,$RESOURCE_NAME] \
    "ERROR: Option $OPTARG unknown"
    exit 1
    ;;
esac

done

}

#####
# MAIN
#
#####

export PATH=/bin:/usr/bin:/usr/cluster/bin:/usr/sbin:/usr/proc/bin:$PATH

# メッセージの記録に使用する syslog 機能番号を取得する。
SYSLOG_FACILITY=`scha_cluster_get -O SYSLOG_FACILITY`

# このメソッドに渡された引数を構文解析する。
parse_args "$@"

PMF_TAG=$RESOURCE_NAME.named
SYSLOG_TAG=$RESOURCECETYPE_NAME,$RESOURCEGROUP_NAME,$RESOURCE_NAME

# RTR ファイルから Stop_timeout 値を取得する。
STOP_TIMEOUT=`scha_resource_get -O STOP_TIMEOUT -R $RESOURCE_NAME -G \
$RESOURCEGROUP_NAME`

# PMF 経由で SIGTERM シグナルを使用する規則正しい方法でデータサービ
# スを停止しようとする。SIGTERM がデータサービスを停止できるまで、
# Stop_timeout 値の 80% だけ待つ。停止できない場合、SIGKILL を送信
# して、データサービスを停止しようとする。SIGKILL がデータサービ
# スを停止できるまで、Stop_timeout 値の 15% だけ待つ。停止できない場合、
# メソッドは何か異常があったと判断し、0 以外の状態で終了する。
# Stop_timeout の残りの 5% は他の目的のために予約されている。
((SMOOTH_TIMEOUT=$STOP_TIMEOUT * 80/100))

((HARD_TIMEOUT=$STOP_TIMEOUT * 15/100))

# in.named が動作しているかどうかを調べて、動作していれば停止する。
if pmfadm -q $PMF_TAG.named; then
    # シグナルをデータサービスに送信して、合計タイムアウト値
    # の 80% だけ待つ。

```

例 B-3 dns_svc_stop メソッド (続き)

```
pmfadm -s $PMF_TAG.named -w $SMOOTH_TIMEOUT TERM
if [ $? -ne 0 ]; then
    logger -p ${SYSLOG_FACILITY}.info -t [SYSLOG_TAG] \
        "${ARGV0} Failed to stop HA-DNS with SIGTERM; Retry with \
        SIGKILL"

    # SIGTERM シグナルでデータサービスが停止しないので、今度は
    # SIGKILL を使用して、合計タイムアウト値の 15% だけ待つ。

    pmfadm -s $PMF_TAG.named -w $HARD_TIMEOUT KILL
    if [ $? -ne 0 ]; then
        logger -p ${SYSLOG_FACILITY}.err -t [SYSLOG_TAG] \
            "${ARGV0} Failed to stop HA-DNS; Exiting UNSUCCESSFUL"

        exit 1
    fi
fi
else
    # この時点でデータサービスは動作していない。メッセージを記録して、
    # 成功で終了する。

    logger -p ${SYSLOG_FACILITY}.info -t [SYSLOG_TAG] \
        "HA-DNS is not started"

    # HA-DNS が動作していない場合でも、成功で終了し、データサービス
    # リソースが STOP_FAILED 状態にならないようにする。

    exit 0
fi

# DNS の停止に成功。メッセージを記録して、成功で終了する。
logger -p ${SYSLOG_FACILITY}.info -t [SYSLOG_TAG] \
    "HA-DNS successfully stopped"
exit 0
```

gettime ユーティリティ

gettime ユーティリティは、検証の再起動間の経過時間を PROBE プログラムが追跡するための C プログラムです。このプログラムは、コンパイル後、コールバックメソッドと同じディレクトリ (RT_basedir プロパティが指すディレクトリ) に格納する必要があります。

例 B-4 gettime.c ユーティリティプログラム

```
#
# このユーティリティプログラムは、データサービスの検証メソッドによ
# って使用され、既知の参照ポイント (基準点) からの経過時間 (秒) を
```

例 B-4 gettimeofday.c ユーティリティプログラム (続き)

```
# 追跡する。このプログラムは、コンパイル後、データサービスのコール
# バックメソッドと同じディレクトリ (RT_basedir) に格納しておくこと。

#pragma ident    "@(#)gettimeofday.c    1.1    00/05/24 SMI"

#include <stdio.h>
#include <sys/types.h>
#include <time.h>

main()
{
    printf("%d\n", time(0));
    exit(0);
}
```

PROBE プログラム

PROBE プログラムは、nslookup(1M) コマンドを使用して、データサービスの可用性を検査します。このプログラムは、Monitor_start コールバックメソッドによって起動され、Monitor_stop コールバックメソッドによって停止されます。

例 B-5 dns_probe プログラム

```
#!/bin/ksh
#pragma ident    "@(#)dns_probe    1.1    00/04/19 SMI"
#
# HA-DNS の Probe メソッド。
#
# このプログラムは、nslookup を使用して、データサービスの可用性を検査
# する。nslookup は DNS サーバーに照会することによって、DNS
# サーバー自身を探す。サーバーが応答しない場合、あるいは、別のサー
# バーが照会に応答した場合、probe メソッドはデータサービスまたはク
# ラスタ内の別のノードになんらかの問題が発生したという結論を下す。
# 検証は、RTR ファイルの THOROUGH_PROBE_INTERVAL で設定さ
# れた間隔で行われる。

#pragma ident    "@(#)dns_probe    1.1    00/05/24 SMI"

#####
# プログラム引数を構文解析する。
function parse_args # [args ...]
{
    typeset opt

    while getopts 'R:G:T:' opt
    do
        case "$opt" in

```

例 B-5 dns_probe プログラム (続き)

```

R)
    # DNS リソースの名前。
    RESOURCE_NAME=$OPTARG
    ;;
G)
    # リソースが構成されているリソース
    # グループの名前。

    RESOURCEGROUP_NAME=$OPTARG
    ;;
T)
    # リソースタイプの名前。
    RESOURCETYPE_NAME=$OPTARG
    ;;
*)
    logger -p ${SYSLOG_FACILITY}.err \
    -t [$RESOURCE_NAME,$RESOURCEGROUP_NAME,$RESOURCETYPE_NAME] \
    "ERROR: Option $OPTARG unknown"
    exit 1
    ;;
esac
done
}

#####
# restart_service ()
#
# この関数は、まずデータサービスの stop メソッドを呼び出し、
# 次に start メソッドを呼び出すことによって、データサービスを再起動
# しようとする。データサービスがすでに起動しておらず、
# データサービスのタグが PMF に登録されていない場合、
# この関数はデータサービスをクラスタ内の
# 別のノードにフェイルオーバーする。
#
function restart_service
{
    # データサービスを再起動するには、まず、データサービス自身が
    # PMF 下に登録されているかどうかを確認する。

    pmfadm -q $PMF_TAG
    if [[ $? -eq 0 ]]; then
        # データサービスの TAG が PMF に登録されている場合、
        # データサービスを停止し、起動し直す。
        # 当該リソースの stop メソッド名と STOP_TIMEOUT 値を取得する。

        STOP_TIMEOUT=`scha_resource_get -O STOP_TIMEOUT \
            -R $RESOURCE_NAME -G $RESOURCEGROUP_NAME`
        STOP_METHOD=`scha_resource_get -O STOP \
            -R $RESOURCE_NAME -G $RESOURCEGROUP_NAME`
        hatimerun -t $STOP_TIMEOUT $RT_BASEDIR/$STOP_METHOD \
            -R $RESOURCE_NAME -G $RESOURCEGROUP_NAME \
            -T $RESOURCETYPE_NAME
    fi
}

```

例 B-5 dns_probe プログラム (続き)

```

        if [[ $? -ne 0 ]]; then
            logger-p ${SYSLOG_FACILITY}.err -t [${SYSLOG_TAG}] \
                "${ARGV0} Stop method failed."
            return 1
        fi

        # 当該リソースの start メソッド名と START_TIMEOUT 値を取得する。
        START_TIMEOUT=`scha_resource_get -O START_TIMEOUT \
            -R $RESOURCE_NAME -G $RESOURCEGROUP_NAME`
        START_METHOD=`scha_resource_get -O START \
            -R $RESOURCE_NAME -G $RESOURCEGROUP_NAME`
        hatimerun -t $START_TIMEOUT $RT_BASEDIR/$START_METHOD \
            -R $RESOURCE_NAME -G $RESOURCEGROUP_NAME \
            -T $RESOURCECETYPE_NAME

        if [[ $? -ne 0 ]]; then
            logger-p ${SYSLOG_FACILITY}.err -t [${SYSLOG_TAG}] \
                "${ARGV0} Start method failed."
            return 1
        fi

    else
        # データサービスの TAG が PMF に登録されていない場合、
        # データサービスが PMF 下で許可されている再試行最大回数を
        # 超えていることを示す。したがって、データサービスを再起動
        # してはならない。その代わりに、同じクラスタ内にある別のノード
        # にフェイルオーバーを試みる。

        scha_control -O GIVEOVER -G $RESOURCEGROUP_NAME \
            -R $RESOURCE_NAME
    fi

    return 0
}

#####
# decide_restart_or_failover ()
#
# この関数は、検証が失敗したときに行うべきアクション、つまり、デー
# タサービスをローカルで再起動するか、クラスタ内の別のノードに
# フェイルオーバーするかを決定する。
#
function decide_restart_or_failover
{
    # 最初の再起動の試行であるかどうかを検査する。
    if [ $retries -eq 0 ]; then
        # 最初の失敗である。
        # 最初の試行の時刻を記録する。

        start_time=`$RT_BASEDIR/gettimè`
        retries=`expr $retries + 1`
        # 最初の失敗であるので、データサービスを

```

例 B-5 dns_probe プログラム (続き)

```
# 再起動しようと試行する。

restart_service
if [ $? -ne 0 ]; then
    logger -p ${SYSLOG_FACILITY}.err -t [${SYSLOG_TAG}] \
        "${ARGV0} Failed to restart data service."
    exit 1
fi
else
# 最初の失敗ではない。
current_time=`$RT_BASEDIR/gettimè
time_diff=`expr $current_time - $start_time`
if [ $time_diff -ge $RETRY_INTERVAL ]; then
# この失敗は再試行最大期間後に発生した。
# したがって、再試行カウンタをリセットし、
# 再試行時間をリセットし、さらに再試行する。

retries=1
start_time=$current_time
# 前回の失敗が Retry_interval よりも以前に発生しているので、
# データサービスを再起動しようと試行する。

restart_service
if [ $? -ne 0 ]; then
    logger -p ${SYSLOG_FACILITY}.err \
        -t [${SYSLOG_TAG}
            "${ARGV0} Failed to restart HA-DNS."
    exit 1
fi
elif [ $retries -ge $RETRY_COUNT ]; then
# 再試行最大期間内であり、再試行カウンタは満了
# している。したがって、フェイルオーバーする。

retries=0
scha_control -O GIVEOVER -G $RESOURCEGROUP_NAME \
    -R $RESOURCE_NAME
if [ $? -ne 0 ]; then
    logger -p ${SYSLOG_FACILITY}.err -t [${SYSLOG_TAG}] \
        "${ARGV0} Failover attempt failed."
    exit 1
fi
else
# 再試行最大期間内であり、再試行カウンタは満了
# していない。したがって、さらに再試行する。

retries=`expr $retries + 1`
restart_service
if [ $? -ne 0 ]; then
    logger -p ${SYSLOG_FACILITY}.err -t [${SYSLOG_TAG}] \
        "${ARGV0} Failed to restart HA-DNS."
    exit 1
fi
fi
fi
```

例 B-5 dns_probe プログラム (続き)

```
fi
}

#####
# MAIN
#####

export PATH=/bin:/usr/bin:/usr/cluster/bin:/usr/sbin:/usr/proc/bin:$PATH

# メッセージの記録に使用する syslog 機能番号を取得する。
SYSLOG_FACILITY=`scha_resource_get -O SYSLOG_FACILITY`

# このメソッドに渡された引数を構文解析する。
parse_args "$@"

PMF_TAG=$RESOURCE_NAME.named
SYSLOG_TAG=$RESOURCE_TYPE_NAME,$RESOURCEGROUP_NAME,$RESOURCE_NAME

# 証が行われる間隔はシステム定義プロパティ THOROUGH_PROBE_INTERVAL
# に設定されている。scha_resource_get でこのプロパティの値を取得する。
PROBE_INTERVAL=scha_resource_get -O THOROUGH_PROBE_INTERVAL \
-R $RESOURCE_NAME -G $RESOURCEGROUP_NAME

# 検証用のタイムアウト値を取得する。この値は RTR ファイルの
# PROBE_TIMEOUT 拡張プロパティに設定されている。nslookup のデフォル
# トのタイムアウトは 1.5 分。
probe_timeout_info=`scha_resource_get -O Extension -R $RESOURCE_NAME \
-G $RESOURCEGROUP_NAME Probe_timeout`
PROBE_TIMEOUT=`echo $probe_timeout_info | awk '{print $2}'`

# リソースの NETWORK_RESOURCES_USED プロパティの値を取得して、
# DNS がサービスを提供するサーバーを見つける。
DNS_HOST=`scha_resource_get -O NETWORK_RESOURCES_USED -R $RESOURCE_NAME \
-G $RESOURCEGROUP_NAME

# システム定義プロパティ Retry_count から再試行最大回数を取得する。
RETRY_COUNT=`scha_resource_get -O RETRY_COUNT -R $RESOURCE_NAME \
-G $RESOURCEGROUP_NAME

# システム定義プロパティ Retry_interval から再試行最大期間を取得する。
Retry_interval
RETRY_INTERVAL=scha_resource_get -O RETRY_INTERVAL -R $RESOURCE_NAME \
-G $RESOURCEGROUP_NAME

# リソースタイプの RT_basedir プロパティから gettime ユーティリティの
# 完全パスを取得する。
RT_BASEDIR=scha_resource_get -O RT_BASEDIR -R $RESOURCE_NAME \
-G $RESOURCEGROUP_NAME

# 検証は無限ループで動作し、nslookup コマンドを実行し続ける。
# nslookup 応答用の一時ファイルを設定する。
DNSPROBEFILE=/tmp/.$RESOURCE_NAME.probe
probefail=0
```


例 B-5 dns_probe プログラム (続き)

```
retries=0

while :
do
# 検証が動作すべき期間は <THOROUGH_PROBE_INTERVAL> プロパティに指
# 定されている。したがって、THOROUGH_PROBE_INTERVAL の間、検証
# プログラムが休眠するように設定する。
    sleep $PROBE_INTERVAL

# DNS がサービスを提供している IP アドレス上で nslookup コマンド
# を実行する。

    hatimerun -t $PROBE_TIMEOUT /usr/sbin/nslookup $DNS_HOST $DNS_HOST \
        > $DNSPROBEFILE 2>&1

    retcode=$?
        if [ retcode -ne 0 ]; then
            probefail=1
        fi

# nslookup への応答が HA-DNS サーバーから来ており、
# /etc/resolv.conf ファイル内に指定されているほかのネームサーバー
# から来ていないことを確認する。

    if [ $probfail -eq 0 ]; then
        # nslookup 照会に回答するサーバーの名前を取得する。
        SERVER=`awk ' $1=="Server:" {print $2 }' \
            $DNSPROBEFILE | awk -F. ' { print $1 } ' `
        if [ -z "$SERVER" ];
        then
            probefail=1
        else
            if [ $SERVER != $DNS_HOST ]; then
                probefail=1
            fi
        fi
    fi

# probefail 変数が 0 以外である場合、nslookup コマンドがタイム
# アウトしたか、あるいは、別のサーバー (/etc/resolv.conf ファイ
# ルに指定されている) から照会への応答が来ていることを示す。
# どちらの場合でも、DNS サーバーは応答していないので、
# このメソッドは decide_restart_or_failover を呼び出して、
# データサービスをローカルで起動するか、あるいは、別のノードに
# フェイルオーバーするかを評価する。

    if [ $probfail -ne 0 ]; then
        decide_restart_or_failover
    else
        logger -p ${SYSLOG_FACILITY}.info -t [${SYSLOG_TAG}] \
            "${ARGV0} Probe for resource HA-DNS successful"
    fi
fi
```

例 B-5 dns_probe プログラム (続き)

done

Monitor_start メソッド

このメソッドは、データサービスの PROBE プログラムを起動します。

例 B-6 dns_monitor_start メソッド

```
#!/bin/ksh
#
# HA-DNS の Monitor_start メソッド
#
# このメソッドは、PMF の制御下でデータサービスのモニター (検証) を
# 起動する。モニターは一定の間隔でデータサービスを検証するプロセス
# で、問題が発生すると、データサービスを同じノード上で再起動するか、
# クラスタ内の別のノードにフェイルオーバーする。モニター用の PMF
# タグは $RESOURCE_NAME.monitor。

#pragma ident "@(#)dns_monitor_start 1.1 00/05/24 SMI"

#####
# プログラム引数を構文解析する。
#
function parse_args # [args ...]
{
    typeset opt

    while getopts `R:G:T:` opt
    do
        case "$opt" in
            R)
                # DNS リソースの名前。
                RESOURCE_NAME=$OPTARG
                ;;
            G)
                # リソースが構成されているリソース
                # グループの名前。

                RESOURCEGROUP_NAME=$OPTARG
                ;;
            T)
                # リソースタイプの名前。
                RESOURCETYPE_NAME=$OPTARG
                ;;
            *)
                logger -p ${SYSLOG_FACILITY}.err \
                    -t [${RESOURCETYPE_NAME}, ${RESOURCEGROUP_NAME}, ${RESOURCE_NAME}] \

```

例 B-6 dns_monitor_start メソッド (続き)

```
                "ERROR: Option $OPTARG unknown"
                exit 1
                ;;
        esac
done

}

#####
# MAIN
#
#####

export PATH=/bin:/usr/bin:/usr/cluster/bin:/usr/sbin:/usr/proc/bin:$PATH

# メッセージの記録に使用する syslog 機能を取得する。
SYSLOG_FACILITY=`scha_cluster_get -O SYSLOG_FACILITY`

# のメソッドに渡された引数を構文解析する。
parse_args "$@"

PMF_TAG=$RESOURCE_NAME.monitor
SYSLOG_TAG=$RESOURCE_TYPE_NAME,$RESOURCEGROUP_NAME,$RESOURCE_NAME

# データサービスの RT_BASEDIR プロパティを取得することによって、検
# 証メソッドが存在する場所を見つける。
RT_BASEDIR=`scha_resource_get -O RT_BASEDIR -R $RESOURCE_NAME \
-G $RESOURCEGROUP_NAME`

# PMF の制御下でデータサービスの検証を開始する。無限再試行オプション
# を使用して検証メソッドを起動する。リソースの名前、タイプ、および
# グループを検証メソッドに渡す。
pmfadm -c $PMF_TAG.monitor -n -1 -t -1 \
    $RT_BASEDIR/dns_probe -R $RESOURCE_NAME -G $RESOURCEGROUP_NAME \
    -T $RESOURCE_TYPE_NAME

# HA-DNS のモニターが起動されたことを示すメッセージを記録する。
if [ $? -eq 0 ]; then
    logger -p ${SYSLOG_FACILITY}.info -t [${SYSLOG_TAG}] \
        "${ARGV0} Monitor for HA-DNS successfully started"
fi
exit 0
```

Monitor_stop メソッド

このメソッドは、データサービスの PROBE プログラムを停止します。

例 B-7 dns_monitor_stop メソッド

```
#!/bin/ksh
# HA-DNS の Monitor_stop メソッド
# PMF を使用して動作しているモニターを停止する。

#pragma ident    "@(#)dns_monitor_stop    1.1    00/05/24 SMI"

#####
# プログラム引数を構文解析する。
#
function parse_args # [args ...]
{
    typeset opt

    while getopts `R:G:T:` opt
    do
        case "$opt" in
            R)
                # DNS リソースの名前。
                RESOURCE_NAME=$OPTARG
                ;;
            G)
                # リソースが構成されているリソース
                # グループの名前。

                RESOURCEGROUP_NAME=$OPTARG
                ;;
            T)
                # リソースタイプの名前。
                RESOURCETYPE_NAME=$OPTARG
                ;;
            *)
                logger -p ${SYSLOG_FACILITY}.err \
                -t [${RESOURCETYPE_NAME},${RESOURCEGROUP_NAME},${RESOURCE_NAME}] \
                "ERROR: Option $OPTARG unknown"
                exit 1
                ;;
        esac
    done
}

#####
# MAIN
#
#####

export PATH=/bin:/usr/bin:/usr/cluster/bin:/usr/sbin:/usr/proc/bin:$PATH
```

例 B-7 dns_monitor_stop メソッド (続き)

```
# メッセージの記録に使用する syslog 機能を取得する。
SYSLOG_FACILITY=`scha_cluster_get -O SYSLOG_FACILITY`

# このメソッドに渡された引数を構文解析する。
parse_args "$@"

PMF_TAG=$RESOURCE_NAME.monitor
SYSLOG_TAG=$RESOURCETYPE_NAME,$RESOURCEGROUP_NAME,$RESOURCE_NAME

# モニターが動作しているかどうかを調べて、動作していれば停止する。
if pmfadm -q $PMF_TAG.monitor; then
    pmfadm -s $PMF_TAG.monitor KILL
    if [ $? -ne 0 ]; then
        logger -p ${SYSLOG_FACILITY}.err -t [SYSLOG_TAG] \
            "${ARGV0} Could not stop monitor for resource " \
            $RESOURCE_NAME
        exit 1
    else
        # モニターは正常に停止している。メッセージを記録する。
        logger -p ${SYSLOG_FACILITY}.info -t [SYSLOG_TAG] \
            "${ARGV0} Monitor for resource " $RESOURCE_NAME \
            " successfully stopped"
    fi
fi
exit 0
```

Monitor_check メソッド

このメソッドは、Confdir プロパティが示すディレクトリの存在を確認します。PROBE メソッドがデータサービスを新しいノードにフェイルオーバーするとき、RGM は Monitor_check を呼び出し、また、潜在マスターであるノードを検査します。

例 B-8 dns_monitor_check メソッド

```
#!/bin/ksh#
# DNS の Monitor_check メソッド
#
# 障害モニターがデータサービスを新しいノードにフェイルオーバー
# するとき、RGM はこのメソッドを呼び出す。Monitor_check は
# Validate メソッドを呼び出して、新しいノード上で構成ディレク
# トリおよびファイルが利用できるかどうかを確認する。
#pragma ident "@(#)dns_monitor_check 1.1 00/05/24 SMI"

#####
# プログラム引数を構文解析する。
```

例 B-8 dns_monitor_check メソッド (続き)

```
function parse_args # [args ...]
{
    typeset opt

    while getopts 'R:G:T:' opt
    do
        case "$opt" in

            R)
                # DNS リソースの名前。
                RESOURCE_NAME=$OPTARG
                ;;

            G)
                # リソースが構成されているリソース
                # グループの名前。

                RESOURCEGROUP_NAME=$OPTARG
                ;;

            T)
                # リソースタイプの名前。
                RESOURCETYPE_NAME=$OPTARG
                ;;

            *)
                logger -p ${SYSLOG_FACILITY}.err \
                -t [$RESOURCECETYPE_NAME,$RESOURCEGROUP_NAME,$RESOURCE_NAME] \
                "ERROR: Option $OPTARG unknown"
                exit 1
                ;;
            esac
        done
    }

#####
# MAIN
#####

export PATH=/bin:/usr/bin:/usr/cluster/bin:/usr/sbin:/usr/proc/bin:$PATH

# メッセージの記録に使用する syslog 機能を取得する。
SYSLOG_FACILITY=`scha_cluster_get -O SYSLOG_FACILITY`

# このメソッドに渡された引数を構文解析する。
parse_args "$@"

PMF_TAG=$RESOURCE_NAME.named
SYSLOG_TAG=$RESOURCECETYPE_NAME,$RESOURCEGROUP_NAME,$RESOURCE_NAME

# リソースタイプの RT_BASEDIR プロパティから validate メソッドの
# 完全パスを取得する。
```

例 B-8 dns_monitor_check メソッド (続き)

```
RT_BASEDIR=`scha_resource_get -O RT_BASEDIR -R $RESOURCE_NAME \
-G $RESOURCEGROUP_NAME`

# 当該リソースの Validate メソッド名を取得する。
VALIDATE_METHOD=`scha_resource_get -O VALIDATE -R $RESOURCE_NAME \
-G $RESOURCEGROUP_NAME`

# データサービスを起動するための Confdir プロパティの値を取得する。
# 入力されたリソース名とリソースグループを使用して、リソースを
# 追加するときに設定した Confdir の値を取得する。
config_info=`scha_resource_get -O Extension -R $RESOURCE_NAME \
-G $RESOURCEGROUP_NAME Confdir`

# scha_resource_get は、拡張プロパティの値とともにタイプも戻す。
# awk を使用して、拡張プロパティの値だけを取得する。
CONFIG_DIR=`echo $config_info | awk '{print $2}'`

# Validate メソッドを呼び出して、データサービスを新しいノードに
# フェイルオーバーできるかどうかを確認する。
$RT_BASEDIR/$VALIDATE_METHOD -R $RESOURCE_NAME -G $RESOURCEGROUP_NAME \
-T $RESOURCETYPE_NAME -x Confdir=$CONFIG_DIR

# モニター検査が成功したことを示すメッセージを記録する。
if [ $? -eq 0 ]; then
    logger -p ${SYSLOG_FACILITY}.info -t [SYSLOG_TAG] \
        "${ARGV0} Monitor check for DNS successful."
    exit 0
else
    logger -p ${SYSLOG_FACILITY}.err -t [SYSLOG_TAG] \
        "${ARGV0} Monitor check for DNS not successful."
    exit 1
fi
```

Validate メソッド

このメソッドは、Confdir プロパティが示すディレクトリの存在を確認します。RGM がこのメソッドを呼び出すのは、クラスタ管理者がデータサービスを作成したときと、データサービスのプロパティを更新したときです。障害モニターがデータサービスを新しいノードにフェイルオーバーしたときは、Monitor_check メソッドは常にこのメソッドを呼び出します。

例 B-9 dns_validate メソッド

```
#!/bin/ksh
# # HA-DNS の Validate メソッド
# このメソッドは、リソースの Confdir プロパティを妥当性検査する。
```

例 B-9 dns_validate メソッド (続き)

```
# Validate メソッドが呼び出されるのは、リソースが作成されたときと、リソース
# プロパティが更新されたときの 2 つである。リソースが作成されたとき、
# Validate メソッドは -c フラグで呼び出され、すべてのシステム定義プロ
# パティと拡張プロパティがコマンド行引数として渡される。リソースプロ
# パティが更新されたとき、Validate メソッドは -u フラグで呼び出され、
# 更新されるプロパティのプロパティ/値のペアだけがコマンド行引数とし
# て渡される。
#
# 例: リソースが作成されたとき、コマンド行引数は次のようになる。
#
# dns_validate -c -R <...> -G <...> -T <...> -r <sysdef-prop=value>...
#       -x <extension-prop=value>.... -g <resourcegroup-prop=value>....
#
# 例: リソースプロパティが更新されたとき、コマンド行引数は次のようになる。
#
# dns_validate -u -R <...> -G <...> -T <...> -r <sys-prop_being_updated=value>
# または
# dns_validate -u -R <...> -G <...> -T <...> -x <extn-prop_being_updated=value>

#pragma ident    "@(#)dns_validate    1.1    00/05/24 SMI"

#####
# プログラム引数を構文解析する。
#
function parse_args # [args ...]
{
    typeset opt

    while getopts `cur:x:g:R:T:G:` opt
    do
        case "$opt" in
            R)
                # DNS リソースの名前。
                RESOURCE_NAME=$OPTARG
                ;;
            G)
                # リソースが構成されているリソース
                # グループの名前。
                RESOURCEGROUP_NAME=$OPTARG
                ;;
            T)
                # リソースタイプの名前。
                RESOURCETYPE_NAME=$OPTARG
                ;;
            r)
                #メソッドはシステム定義プロパティにアクセスして
                # いない。したがって、このフラグは動作なし。
                ;;
            g)
                # メソッドはリソースグループプロパティにアクセスして
                # いない。したがって、このフラグは動作なし。
                ;;
            c)
                ;;
        esac
    done
}
```


例 B-9 dns_validate メソッド (続き)

```

        # Validate メソッドがリソースの作成中に呼び出されてい
# ることを示す。したがって、このフラグは動作なし。
        ;;
    u)
        # リソースがすでに存在しているときは、プロパティの更新
# を示す。Confdir プロパティを更新する場合、Confdir
# がコマンド行引数に現れるはずである。現れない場合、
# メソッドは scha_resource_get を使用して
# Confdir を探す必要がある。
        UPDATE_PROPERTY=1
        ;;
    x)
        # 拡張プロパティのリスト。プロパティと値のペア。
# 区切り文字は「=」。

        PROPERTY=`echo $OPTARG | awk -F= '{print $1}'`
        VAL=`echo $OPTARG | awk -F= '{print $2}'`

        # Confdir 拡張プロパティがコマンド行上に存在する場合、
# その値を記録する。
        if [ $PROPERTY == "Confdir" ];
        then
            CONFDIR=$VAL
            CONFDIR_FOUND=1
        fi
        ;;
    *)
        logger -p ${SYSLOG_FACILITY}.err \
        -t [${SYSLOG_TAG}] \
        "ERROR: Option $OPTARG unknown"
        exit 1
        ;;
esac

done
}

#####
# MAIN
#
#####

export PATH=/bin:/usr/bin:/usr/cluster/bin:/usr/sbin:/usr/proc/bin:$PATH

# メッセージの記録に使用する syslog 機能を取得する。
SYSLOG_FACILITY=`scha_cluster_get -O SYSLOG_FACILITY`

# CONFDIR の値を NULL に設定する。この後、このメソッドは Confdir
# プロパティの値を、コマンド行から取得するか、scha_resource_get を
# 使って取得する。
CONFDIR=""
UPDATE_PROPERTY=0
CONFDIR_FOUND=0

```

例 B-9 dns_validate メソッド (続き)

```
# このメソッドに渡された引数を構文解析する。
parse_args "$@"

# プロパティの更新の結果として呼び出されている場合、Validate メソッドはコマンド行から Confdir 拡張プロパティの値を取得する。そうでない場合、scha_resource_get を使用して Confdir の値を取得する。
if ( (( $UPDATE_PROPERTY == 1 )) && (( CONFDIR_FOUND == 0 )) ); then
    config_info=scha_resource_get -O Extension -R $RESOURCE_NAME \
        -G $RESOURCEGROUP_NAME Confdir`
    CONFDIR=`echo $config_info | awk '{print $2}'`
fi

# Confdir プロパティが値を持っているかどうかを確認する。持っていない場合、状態 1 (失敗) で終了する。
if [[ -z $CONFDIR ]]; then
    logger -p ${SYSLOG_FACILITY}.err \
        "${ARGV0} Validate method for resource "$RESOURCE_NAME " failed"
    exit 1
fi

# 実際の Confdir プロパティ値の妥当性検査はここから始まる。
# $CONFDIR がアクセス可能であるかどうかを検査する。
if [ ! -d $CONFDIR ]; then
    logger -p ${SYSLOG_FACILITY}.err -t [${SYSLOG_TAG}] \
        "${ARGV0} Directory $CONFDIR missing or not mounted"
    exit 1
fi

# named.conf ファイルが Confdir ディレクトリ内に存在するかどうかを
# 検査する。
if [ ! -s $CONFDIR/named.conf ]; then
    logger -p ${SYSLOG_FACILITY}.err -t [${SYSLOG_TAG}] \
        "${ARGV0} File $CONFDIR/named.conf is missing or empty"
    exit 1
fi

# Validate メソッドが成功したことを示すメッセージを記録する。
logger -p ${SYSLOG_FACILITY}.info -t [${SYSLOG_TAG}] \
    "${ARGV0} Validate method for resource "$RESOURCE_NAME \
    " completed successfully"

exit 0
```

Update メソッド

プロパティが変更された場合、RGM は Update メソッドを呼び出して、そのことを動作中のリソースに通知します。

例 B-10 dns_update メソッド

```
#!/bin/ksh
# HA-DNS の Update メソッド
# 実際のプロパティの更新は RGM が行う。更新の影響を受けるのは障害モ
# ニターだけである。したがって、このメソッドは障害モニターを再起動
# する必要がある。
#pragma ident "@(#)dns_update 1.1 00/05/24 SMI"

#####
# プログラム引数を構文解析する。
#
function parse_args # [args ...]
{
    typeset opt

    while getopts `R:G:T:` opt
    do
        case "$opt" in
            R)
                # DNS リソースの名前。
                RESOURCE_NAME=$OPTARG
                ;;
            G)
                # リソースが構成されているリソース
                # グループの名前。

                RESOURCEGROUP_NAME=$OPTARG
                ;;
            T)
                # リソースタイプの名前。
                RESOURCETYPE_NAME=$OPTARG
                ;;
            *)
                logger -p ${SYSLOG_FACILITY}.err \
                -t [${RESOURCETYPE_NAME},${RESOURCEGROUP_NAME},${RESOURCE_NAME}] \
                "ERROR: Option $OPTARG unknown"
                exit 1
                ;;
        esac
    done
}
#####
# MAIN
#####
export PATH=/bin:/usr/bin:/usr/cluster/bin:/usr/sbin:/usr/proc/bin:$PATH
```

例 B-10 dns_update メソッド (続き)

```
# メッセージの記録に使用する syslog 機能を取得する。
SYSLOG_FACILITY=`scha_cluster_get -O SYSLOG_FACILITY`

# このメソッドに渡された引数を構文解析する。
parse_args "$@"

PMF_TAG=$RESOURCE_NAME.monitor
SYSLOG_TAG=$RESOURCE_TYPE_NAME,$RESOURCEGROUP_NAME,$RESOURCE_NAME

# リソースの RT_BASEDIR プロパティを取得することによって、
# 検証メソッドが存在する場所を見つける。
RT_BASEDIR=`scha_resource_get -O RT_BASEDIR -R $RESOURCE_NAME \
-G $RESOURCEGROUP_NAME`

# Update メソッドが呼び出されると、RGM は更新されるプロパティの値を
# 更新する。このメソッドは、障害モニター (検証メソッド) が動作し
# ているかどうかを検査し、動作している場合は強制終了し、再起動
# する必要がある。
if pmfadm -q $PMF_TAG.monitor; then

# すでに動作している障害モニターを強制終了する。
    pmfadm -s $PMF_TAG.monitor TERM
    if [ $? -ne 0 ]; then
        logger -p ${SYSLOG_FACILITY}.err -t [${SYSLOG_TAG}] \
            "${ARGV0} Could not stop the monitor"
        exit 1
    else
        # DNS の停止に成功。メッセージを記録する。
        logger -p ${SYSLOG_FACILITY}.info -t [${SYSLOG_TAG}] \
            "Monitor for HA-DNS successfully stopped"
    fi
# モニターを再起動する。
    pmfadm -c $PMF_TAG.monitor -n -1 -t -1 $RT_BASEDIR/dns_probe \
        -R $RESOURCE_NAME -G $RESOURCEGROUP_NAME -T $RESOURCE_TYPE_NAME
    if [ $? -ne 0 ]; then
        logger -p ${SYSLOG_FACILITY}.err -t [${SYSLOG_TAG}] \
            "${ARGV0} Could not restart monitor for HA-DNS "
        exit 1
    else
        logger -p ${SYSLOG_FACILITY}.info -t [${SYSLOG_TAG}] \
            "Monitor for HA-DNS successfully restarted"
    fi
fi
exit 0
```

付録 C

サンプル DSDL リソースタイプのコード例

この付録では、SUNW.xfnts リソースタイプの各メソッドの完全なコード例を示します。また、コールバックメソッドが呼び出すサブルーチンのコードを含む、xfnts.c のコード例を示します。この付録に含まれるコードリストは、次のとおりです。

- 293 ページの「xfnts.c」
- 305 ページの「xfnts_monitor_check メソッド」
- 306 ページの「xfnts_monitor_start メソッド」
- 307 ページの「xfnts_monitor_stop メソッド」
- 308 ページの「xfnts_probe メソッド」
- 311 ページの「xfnts_start メソッド」
- 312 ページの「xfnts_stop メソッド」
- 313 ページの「xfnts_update メソッド」
- 315 ページの「xfnts_validate メソッドのコードリスト」

xfnts.c

このファイルは、SUNW.xfnts メソッドが呼び出すサブルーチンを実装します。

例 C-1 xfnts.c

```
/*
 * Copyright (c) 1998-2004 by Sun Microsystems, Inc.
 * All rights reserved.
 *
 * xfnts.c - HA-XFS 用の一般的なユーティリティー
 *
 * This utility has the methods for performing the validation, starting and
 * stopping the data service and the fault monitor. It also contains the method
 * to probe the health of the data service. The probe just returns either
 * success or failure. Action is taken based on this returned value in the
 * method found in the file xfnts_probe.c
```

例 C-1 xfnts.c (続き)

```
*
*/

#pragma ident "@(#)xfnts.c 1.47 01/01/18 SMI"

#include <stdio.h>
#include <stdlib.h>
#include <strings.h>
#include <unistd.h>
#include <sys/types.h>
#include <sys/stat.h>
#include <sys/socket.h>
#include <sys/wait.h>
#include <netinet/in.h>
#include <scha.h>
#include <rgm/libdsdev.h>
#include <errno.h>
#include "xfnts.h"

/*
 * The initial timeout allowed for the HAXFS data service to
 * be fully up and running. We will wait for 3 % (SVC_WAIT_PCT)
 * of the start_timeout time before probing the service.
 */
#define SVC_WAIT_PCT 3

/*
 * We need to use 95% of probe_timeout to connect to the port and the
 * remaining time is used to disconnect from port in the svc_probe function.
 */
#define SVC_CONNECT_TIMEOUT_PCT 95

/*
 * SVC_WAIT_TIME is used only during starting in svc_wait().
 * In svc_wait() we need to be sure that the service is up
 * before returning, thus we need to call svc_probe() to
 * monitor the service. SVC_WAIT_TIME is the time between
 * such probes.
 */
#define SVC_WAIT_TIME 5

/*
 * This value will be used as disconnect timeout, if there is no
 * time left from the probe_timeout.
 */
#define SVC_DISCONNECT_TIMEOUT_SECONDS 2

/*
 * svc_validate():
 */
```

例 C-1 xfnts.c (続き)

```
* Do HA-XFS specific validation of the resource configuration.
*
* svc_validate will check for the following
* 1. Confdir_list extension property
* 2. fontserver.cfg file
* 3. xfs binary
* 4. port_list property
* 5. network resources
* 6. other extension properties
*
* If any of the above validation fails then, Return > 0 otherwise return 0 for
* success
*/

int
svc_validate(scds_handle_t scds_handle)
{
    char    xfnts_conf[SCDS_ARRAY_SIZE];
    scha_str_array_t *confdirs;
    scds_net_resource_list_t *snrlp;
    int rc;
    struct stat statbuf;
    scds_port_list_t *portlist;
    scha_err_t err;

    /*
     * Get the configuration directory for the XFS dataservice from the
     * confdir_list extension property.
     */
    confdirs = scds_get_ext_confdir_list(scds_handle);

    /* Return an error if there is no confdir_list extension property */
    if (confdirs == NULL || confdirs->array_cnt != 1) {
        scds_syslog(LOG_ERR,
            "Property Confdir_list is not set properly.");
        return (1); /* Validation failure */
    }

    /*
     * Construct the path to the configuration file from the extension
     * property confdir_list. Since HA-XFS has only one configuration
     * we will need to use the first entry of the confdir_list property.
     */
    (void) sprintf(xfnts_conf, "%s/fontserver.cfg", confdirs->str_array[0]);

    /*
     * Check to see if the HA-XFS configuration file is in the right place.
     * Try to access the HA-XFS configuration file and make sure the
     * permissions are set properly
     */
    if (stat(xfnts_conf, &statbuf) != 0) {
        /*
         * suppress lint error because errno.h prototype

```

例 C-1 xfnts.c (続き)

```
    * is missing void arg
    */
    scds_syslog(LOG_ERR,
        "Failed to access file <%s> : <%s>",
        xfnts_conf, strerror(errno)); /*lint !e746 */
    return (1);
}

/*
 * Make sure that xfs binary exists and that the permissions
 * are correct. The XFS binary are assumed to be on the local
 * File system and not on the Global File System
 */
if (stat("/usr/openwin/bin/xfs", &statbuf) != 0) {
    scds_syslog(LOG_ERR,
        "Cannot access XFS binary : <%s> ", strerror(errno));
    return (1);
}

/* HA-XFS will have only port */
err = scds_get_port_list(scds_handle, &portlist);
if (err != SCHA_ERR_NOERR) {
    scds_syslog(LOG_ERR,
        "Could not access property Port_list: %s.",
        scds_error_string(err));
    return (1); /* Validation Failure */
}

#ifdef TEST
    if (portlist->num_ports != 1) {
        scds_syslog(LOG_ERR,
            "Property Port_list must have only one value.");
        scds_free_port_list(portlist);
        return (1); /* Validation Failure */
    }
#endif

/*
 * Return an error if there is an error when trying to get the
 * available network address resources for this resource
 */
if ((err = scds_get_rs_hostnames(scds_handle, &snrlp))
    != SCHA_ERR_NOERR) {
    scds_syslog(LOG_ERR,
        "No network address resource in resource group: %s.",
        scds_error_string(err));
    return (1); /* Validation Failure */
}

/* Return an error if there are no network address resources */
if (snrlp == NULL || snrlp->num_netresources == 0) {
    scds_syslog(LOG_ERR,
        "No network address resource in resource group.");
}
```


例 C-1 xfnts.c (続き)

```
        rc = 1;
        goto finished;
    }

    /* Check to make sure other important extension props are set */
    if (scds_get_ext_monitor_retry_count(scds_handle) <= 0)
    {
        scds_syslog(LOG_ERR,
            "Property Monitor_retry_count is not set.");
        rc = 1; /* Validation Failure */
        goto finished;
    }
    if (scds_get_ext_monitor_retry_interval(scds_handle) <= 0) {
        scds_syslog(LOG_ERR,
            "Property Monitor_retry_interval is not set.");
        rc = 1; /* Validation Failure */
        goto finished;
    }

    /* All validation checks were successful */
    scds_syslog(LOG_INFO, "Successful validation.");
    rc = 0;

finished:
    scds_free_net_list(snrlp);
    scds_free_port_list(portlist);

    return (rc); /* return result of validation */
}

/*
 * svc_start():
 *
 * Start up the X font server
 * Return 0 on success, > 0 on failures.
 *
 * The XFS service will be started by running the command
 * /usr/openwin/bin/xfns -config <fontserver.cfg file> -port <port to listen>
 * XFS will be started under PMF. XFS will be started as a single instance
 * service. The PMF tag for the data service will be of the form
 * <resourcegroupname,resourceinstance,instance_number.svc>. In case of XFS, since
 * there will be only one instance the instance_number in the tag will be 0.
 */

int
svc_start(scds_handle_t scds_handle)
{
    char    xfnts_conf[SCDS_ARRAY_SIZE];
    char    cmd[SCDS_ARRAY_SIZE];
    scha_str_array_t *confdirs;
    scds_port_list_t    *portlist;
    scha_err_t    err;
```

例 C-1 xfnts.c (続き)

```
/* get the configuration directory from the confdir_list property */
confdirs = scds_get_ext_confdir_list(scds_handle);

(void) sprintf(xfnts_conf, "%s/fontserver.cfg", confdirs->str_array[0]);

/* obtain the port to be used by XFS from the Port_list property */
err = scds_get_port_list(scds_handle, &portlist);
if (err != SCHA_ERR_NOERR) {
    scds_syslog(LOG_ERR,
        "Could not access property Port_list.");
    return (1);
}

/*
 * Construct the command to start HA-XFS.
 * NOTE: XFS daemon prints the following message while stopping the XFS
 * "/usr/openwin/bin/xfns notice: terminating"
 * In order to suppress the daemon message,
 * the output is redirected to /dev/null.
 */
(void) sprintf(cmd,
    "/usr/openwin/bin/xfns -config %s -port %d 2>/dev/null",
    xfnts_conf, portlist->ports[0].port);

/*
 * Start HA-XFS under PMF. Note that HA-XFS is started as a single
 * instance service. The last argument to the scds_pmf_start function
 * denotes the level of children to be monitored. A value of -1 for
 * this parameter means that all the children along with the original
 * process are to be monitored.
 */
scds_syslog(LOG_INFO, "Issuing a start request.");
err = scds_pmf_start(scds_handle, SCDS_PMF_TYPE_SVC,
    SCDS_PMF_SINGLE_INSTANCE, cmd, -1);

if (err == SCHA_ERR_NOERR) {
    scds_syslog(LOG_INFO,
        "Start command completed successfully.");
} else {
    scds_syslog(LOG_ERR,
        "Failed to start HA-XFS ");
}

scds_free_port_list(portlist);
return (err); /* return Success/failure status */
}

/*
 * svc_stop():
 *
 * Stop the XFS server
 * Return 0 on success, > 0 on failures.
 */
```

例 C-1 xfnts.c (続き)

```
*
* svc_stop will stop the server by calling the toolkit function:
* scds_pmf_stop.
*/
int
svc_stop(scds_handle_t scds_handle)
{
    scha_err_t    err;

    /*
     * The timeout value for the stop method to succeed is set in the
     * Stop_Timeout (system defined) property
     */
    scds_syslog(LOG_ERR, "Issuing a stop request.");
    err = scds_pmf_stop(scds_handle,
        SCDS_PMF_TYPE_SVC, SCDS_PMF_SINGLE_INSTANCE, SIGTERM,
        scds_get_rs_stop_timeout(scds_handle));

    if (err != SCHA_ERR_NOERR) {
        scds_syslog(LOG_ERR,
            "Failed to stop HA-XFS.");
        return (1);
    }

    scds_syslog(LOG_INFO,
        "Successfully stopped HA-XFS.");
    return (SCHA_ERR_NOERR); /* Successfully stopped */
}

/*
 * svc_wait():
 *
 * wait for the data service to start up fully and make sure it is running
 * healthy
 */
int
svc_wait(scds_handle_t scds_handle)
{
    int rc, svc_start_timeout, probe_timeout;
    scds_netaddr_list_t *netaddr;

    /* obtain the network resource to use for probing */
    if (scds_get_netaddr_list(scds_handle, &netaddr) {
        scds_syslog(LOG_ERR,
            "No network address resources found in resource group.");
        return (1);
    }

    /* Return an error if there are no network resources */
    if (netaddr == NULL || netaddr->num_netaddrs == 0) {
        scds_syslog(LOG_ERR,
            "No network address resource in resource group.");
    }
}
```

例 C-1 xfnts.c (続き)

```
    return (1);
}

/*
 * Get the Start method timeout, port number on which to probe,
 * the Probe timeout value
 */
svc_start_timeout = scds_get_rs_start_timeout(scds_handle);
probe_timeout = scds_get_ext_probe_timeout(scds_handle);

/*
 * sleep for SVC_WAIT_PCT percentage of start_timeout time
 * before actually probing the dataservice. This is to allow
 * the dataservice to be fully up in order to reply to the
 * probe. NOTE: the value for SVC_WAIT_PCT could be different
 * for different data services.
 * Instead of calling sleep(),
 * call scds_svc_wait() so that if service fails too
 * many times, we give up and return early.
 */
if (scds_svc_wait(scds_handle, (svc_start_timeout * SVC_WAIT_PCT)/100)
    != SCHA_ERR_NOERR) {

    scds_syslog(LOG_ERR, "Service failed to start.");
    return (1);
}

do {
    /*
     * probe the data service on the IP address of the
     * network resource and the portname
     */
    rc = svc_probe(scds_handle,
        netaddr->netaddrs[0].hostname,
        netaddr->netaddrs[0].port_proto.port, probe_timeout);
    if (rc == SCHA_ERR_NOERR) {
        /* Success. Free up resources and return */
        scds_free_netaddr_list(netaddr);
        return (0);
    }

    /*
     * Dataservice is still trying to come up. Sleep for a while
     * before probing again. Instead of calling sleep(),
     * call scds_svc_wait() so that if service fails too
     * many times, we give up and return early.
     */
    if (scds_svc_wait(scds_handle, SVC_WAIT_TIME)
        != SCHA_ERR_NOERR) {
        scds_syslog(LOG_ERR, "Service failed to start.");
        return (1);
    }
}
```

例 C-1 xfnts.c (続き)

```
/* We rely on RGM to timeout and terminate the program */
} while (1);

}

/*
 * This function starts the fault monitor for a HA-XFS resource.
 * This is done by starting the probe under PMF. The PMF tag
 * is derived as <RG-name,RS-name,instance_number.mon>. The restart option
 * of PMF is used but not the "infinite restart". Instead
 * interval/retry_time is obtained from the RTR file.
 */

int
mon_start(scds_handle_t scds_handle)
{
    scha_err_t    err;

    scds_syslog_debug(DBG_LEVEL_HIGH,
        "Calling MONITOR_START method for resource <%s>.",
        scds_get_resource_name(scds_handle));

    /*
     * The probe xfnts_probe is assumed to be available in the same
     * subdirectory where the other callback methods for the RT are
     * installed. The last parameter to scds_pmf_start denotes the
     * child monitor level. Since we are starting the probe under PMF
     * we need to monitor the probe process only and hence we are using
     * a value of 0.
     */
    err = scds_pmf_start(scds_handle, SCDS_PMF_TYPE_MON,
        SCDS_PMF_SINGLE_INSTANCE, "xfnts_probe", 0);

    if (err != SCHA_ERR_NOERR) {
        scds_syslog(LOG_ERR,
            "Failed to start fault monitor.");
        return (1);
    }

    scds_syslog(LOG_INFO,
        "Started the fault monitor.");

    return (SCHA_ERR_NOERR); /* Successfully started Monitor */
}

/*
 * This function stops the fault monitor for a HA-XFS resource.
 * This is done via PMF. The PMF tag for the fault monitor is
 * constructed based on <RG-name_RS-name,instance_number.mon>.
 */

int
```

例 C-1 xfnts.c (続き)

```
mon_stop(scds_handle_t scds_handle)
{
    scha_err_t    err;

    scds_syslog_debug(DBG_LEVEL_HIGH,
        "Calling scds_pmf_stop method");

    err = scds_pmf_stop(scds_handle, SCDS_PMF_TYPE_MON,
        SCDS_PMF_SINGLE_INSTANCE, SIGKILL,
        scds_get_rs_monitor_stop_timeout(scds_handle));

    if (err != SCHA_ERR_NOERR) {
        scds_syslog(LOG_ERR,
            "Failed to stop fault monitor.");
        return (1);
    }

    scds_syslog(LOG_INFO,
        "Stopped the fault monitor.");

    return (SCHA_ERR_NOERR); /* Successfully stopped monitor */
}

/*
 * svc_probe(): Do data service specific probing. Return a float value
 * between 0 (success) and 100(complete failure).
 *
 * The probe does a simple socket connection to the XFS server on the specified
 * port which is configured as the resource extension property (Port_list) and
 * pings the dataservice. If the probe fails to connect to the port, we return
 * a value of 100 indicating that there is a total failure. If the connection
 * goes through and the disconnect to the port fails, then a value of 50 is
 * returned indicating a partial failure.
 */
int
svc_probe(scds_handle_t scds_handle, char *hostname, int port, int
timeout)
{
    int rc;
    hrtime_t  t1, t2;
    int sock;
    char testcmd[2048];
    int time_used, time_remaining;
    time_t connect_timeout;

    /*
     * probe the dataservice by doing a socket connection to the port
     * specified in the port_list property to the host that is
     * serving the XFS dataservice. If the XFS service which is configured
     * to listen on the specified port, replies to the connection, then
     * the probe is successful. Else we will wait for a time period set

```

例 C-1 xfnts.c (続き)

```
* in probe_timeout property before concluding that the probe failed.
*/

/*
 * Use the SVC_CONNECT_TIMEOUT_PCT percentage of timeout
 * to connect to the port
 */
connect_timeout = (SVC_CONNECT_TIMEOUT_PCT * timeout)/100;
t1 = (hrtime_t)(gethrtime()/1E9);

/*
 * the probe makes a connection to the specified hostname and port.
 * The connection is timed for 95% of the actual probe_timeout.
 */
rc = scds_fm_tcp_connect(scds_handle, &sock, hostname, port,
    connect_timeout);
if (rc) {
    scds_syslog(LOG_ERR,
        "Failed to connect to port <%d> of resource <%s>.",
        port, scds_get_resource_name(scds_handle));
    /* this is a complete failure */
    return (SCDS_PROBE_COMPLETE_FAILURE);
}

t2 = (hrtime_t)(gethrtime()/1E9);

/*
 * Compute the actual time it took to connect. This should be less than
 * or equal to connect_timeout, the time allocated to connect.
 * If the connect uses all the time that is allocated for it,
 * then the remaining value from the probe_timeout that is passed to
 * this function will be used as disconnect timeout. Otherwise, the
 * the remaining time from the connect call will also be added to
 * the disconnect timeout.
 *
 */

time_used = (int)(t2 - t1);

/*
 * Use the remaining time(timeout - time_took_to_connect) to disconnect
 */

time_remaining = timeout - (int)time_used;

/*
 * If all the time is used up, use a small hardcoded timeout
 * to still try to disconnect. This will avoid the fd leak.
 */
if (time_remaining <= 0) {
    scds_syslog_debug(DBG_LEVEL_LOW,
        "svc_probe used entire timeout of "
        "%d seconds during connect operation and exceeded the "
```

例 C-1 xfnts.c (続き)

```
        "timeout by %d seconds. Attempting disconnect with timeout"
        " %d ",
        connect_timeout,
        abs(time_used),
        SVC_DISCONNECT_TIMEOUT_SECONDS);

    time_remaining = SVC_DISCONNECT_TIMEOUT_SECONDS;
}

/*
 * Return partial failure in case of disconnection failure.
 * Reason: The connect call is successful, which means
 * the application is alive. A disconnection failure
 * could happen due to a hung application or heavy load.
 * If it is the later case, don't declare the application
 * as dead by returning complete failure. Instead, declare
 * it as partial failure. If this situation persists, the
 * disconnect call will fail again and the application will be
 * restarted.
 */
rc = scds_fm_tcp_disconnect(scds_handle, sock, time_remaining);
if (rc != SCHA_ERR_NOERR) {
    scds_syslog(LOG_ERR,
        "Failed to disconnect to port %d of resource %s.",
        port, scds_get_resource_name(scds_handle));
    /* this is a partial failure */
    return (SCDS_PROBE_COMPLETE_FAILURE/2);
}

t2 = (hrtime_t)(gethrtime()/1E9);
time_used = (int)(t2 - t1);
time_remaining = timeout - time_used;

/*
 * If there is no time left, don't do the full test with
 * fsinfo. Return SCDS_PROBE_COMPLETE_FAILURE/2
 * instead. This will make sure that if this timeout
 * persists, server will be restarted.
 */
if (time_remaining <= 0) {
    scds_syslog(LOG_ERR, "Probe timed out.");
    return (SCDS_PROBE_COMPLETE_FAILURE/2);
}

/*
 * The connection and disconnection to port is successful,
 * Run the fsinfo command to perform a full check of
 * server health.
 * Redirect stdout, otherwise the output from fsinfo
 * ends up on the console.
 */
(void) sprintf(testcmd,
    "/usr/openwin/bin/fsinfo -server %s:%d > /dev/null",
```


例 C-1 xfnts.c (続き)

```
        hostname, port);
scds_syslog_debug(DBG_LEVEL_HIGH,
    "Checking the server status with %s.", testcmd);
if (scds_timerun(scds_handle, testcmd, time_remaining,
    SIGKILL, &rc) != SCHA_ERR_NOERR || rc != 0) {

    scds_syslog(LOG_ERR,
        "Failed to check server status with command <%s>",
        testcmd);
    return (SCDS_PROBE_COMPLETE_FAILURE/2);
}
return (0);
}
```

xfnts_monitor_check メソッド

このメソッドは、基本的なリソースタイプ構成が有効であることを確認します。

例 C-2 xfnts_monitor_check.c

```
/*
 * Copyright (c) 1998-2004 by Sun Microsystems, Inc.
 * All rights reserved.
 *
 * xfnts_monitor_start.c - HA-XFS のモニター起動メソッド
 */

#pragma ident "@(#)xfnts_monitor_check.c 1.11 01/01/18
SMI"

#include <rgm/libdsdev.h>
#include "xfnts.h"

/*
 * このメソッドは、HA-XFS リソース用の障害モニターを起動する。
 * そのためには、検証機能を PMF の制御下で起動する。PMF タグの形式は
 * <RG-name, RS-name.mon> である。PMF の再起動オプションを
 * 使用するが、無限に再起動しない。その代わりに、
 * interval/retry_time を RTR ファイルから取得する。
 */

int
main(int argc, char *argv[])
{
    scds_handle_t    scds_handle;
    int    rc;
```

例 C-2 xfnts_monitor_check.c (続き)

```
/* RGM から渡された引数を処理して、syslog を初期化する。*/
if (scds_initialize(&scds_handle, argc, argv) != SCHA_ERR_NOERR)
{
    scds_syslog(LOG_ERR, "Failed to initialize the handle.");
    return (1);
}

rc = svc_validate(scds_handle);
scds_syslog_debug(DBG_LEVEL_HIGH,
    "monitor_check method "
    "was called and returned <%d>.", rc);

/* scds_initialize が割り当てたすべてのメモリーを解放する。*/
scds_close(&scds_handle);

/* monitor_start メソッドの結果を戻す。*/
return (rc);
}
```

xfnts_monitor_start メソッド

このメソッドは、xfnts_probe メソッドを起動します。

例 C-3 xfnts_monitor_start.c

```
/*
 * Copyright (c) 1998-2004 by Sun Microsystems, Inc.
 * All rights reserved.
 *
 * xfnts_monitor_stop.c - HA-XFS のモニター停止メソッド
 */

#pragma ident "@(#)xfnts_monitor_start.c 1.10 01/01/18
SMI"

#include <rgm/libdsdev.h>
#include "xfnts.h"

/*
 * このメソッドは、HA-XFS リソース用の障害モニターを停止する。
 * この処理は PMF 経由で行われる。障害モニター用の
 * PMF タグの形式は<RG-name_RS-name.mon> である。
 */

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

例 C-3 xfnts_monitor_start.c (続き)

```
scds_handle_t    scds_handle;
int    rc;

/* RGM から渡された引数を処理して、syslog を初期化する。*/
if (scds_initialize(&scds_handle, argc, argv) != SCHA_ERR_NOERR)
{
    scds_syslog(LOG_ERR, "Failed to initialize the handle.");
    return (1);
}

rc = mon_start(scds_handle);

/* scds_initialize が割り当てたすべてのメモリーを解放する。*/
scds_close(&scds_handle);

/* monitor_stop メソッドの結果を戻す。*/
return (rc);
}
```

xfnts_monitor_stop メソッド

このメソッドは、xfnts_probe メソッドを停止します。

例 C-4 xfnts_monitor_stop.c

```
/*
 * Copyright (c) 1998-2004 by Sun Microsystems, Inc.
 * All rights reserved.
 *
 * xfnts_monitor_stop.c - HA-XFS のモニター停止メソッド
 */

#pragma ident "@(#)xfnts_monitor_stop.c 1.9 01/01/18 SMI"

#include <rgm/libdsdev.h>
#include "xfnts.h"

/*
 * このメソッドは、HA-XFS リソース用の障害モニターを停止する。
 * この処理は PMF 経由で行われる。障害モニター用の
 * PMF タグの形式は <RG-name_RS-name.mon> である。
 */

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

例 C-4 xfnts_monitor_stop.c (続き)

```
scds_handle_t    scds_handle;
int              rc;

/* RGM から渡された引数を処理して、syslog を初期化する。*/
if (scds_initialize(&scds_handle, argc, argv) != SCHA_ERR_NOERR)
{
    scds_syslog(LOG_ERR, "Failed to initialize the handle.");
    return (1);
}
rc = mon_stop(scds_handle);

/* scds_initialize が割り当てたすべてのメモリーを解放する。*/
scds_close(&scds_handle);

/* Return the result of monitor stop method */
return (rc);
}
```

xfnts_probe メソッド

xfnts_probe メソッドは、アプリケーションの可用性を検査して、データサービスをフェイルオーバーするか、再起動するかを決定します。xfnts_probe メソッドは、xfnts_monitor_start コールバックメソッドによって起動され、xfnts_monitor_stop コールバックメソッドによって停止されます。

例 C-5 xfnts_probe.c+

```
/*
 * Copyright (c) 1998-2004 by Sun Microsystems, Inc.
 * All rights reserved.
 *
 * xfnts_probe.c - HA-XFS の検査
 */

#pragma ident "@(#)xfnts_probe.c 1.26 01/01/18 SMI"

#include <stdio.h>
#include <stdlib.h>
#include <strings.h>
#include <unistd.h>
#include <signal.h>
#include <sys/time.h>
#include <sys/socket.h>
#include <strings.h>
#include <rgm/libdsdev.h>
#include "xfnts.h"
```

例 C-5 xfnts_probe.c+ (続き)

```
/*
 * main():
 * sleep() を実行して、PMF アクションスクリプトが sleep() に割り込むのを
 * 待機する無限ループ。sleep() への割り込みが発生すると、HA-XFS 用の
 * 起動メソッドを呼び出して、再起動する。
 */

int
main(int argc, char *argv[])
{
    int          timeout;
    int          port, ip, probe_result;
    scds_handle_t scds_handle;

    hrttime_t    ht1, ht2;
    unsigned long dt;

    scds_netaddr_list_t *netaddr;
    char *hostname;

    if (scds_initialize(&scds_handle, argc, argv) != SCHA_ERR_NOERR)
    {
        scds_syslog(LOG_ERR, "Failed to initialize the handle.");
        return (1);
    }

    /* 当該リソースに利用できる IP アドレスを取得する。*/
    if (scds_get_netaddr_list(scds_handle, &netaddr)) {
        scds_syslog(LOG_ERR,
            "No network address resource in resource group.");
        scds_close(&scds_handle);
        return (1);
    }

    /* ネットワークリソースが存在しない場合、エラーを戻す。*/
    if (netaddr == NULL || netaddr->num_netaddrs == 0) {
        scds_syslog(LOG_ERR,
            "No network address resource in resource group.");
        return (1);
    }

    /*
     * X プロパティからタイムアウト値を設定する。つまり、
     * 当該リソース用に構成されたすべてのネットワークリソース間で
     * タイムアウト値を分割するのではなく、検証を行うたびに、
     * 各ネットワークリソースに設定されているタイムアウト値を
     * 取得することを意味する。
     */
    timeout = scds_get_ext_probe_timeout(scds_handle);
}
```

例 C-5 xfnets_probe.c+ (続き)

```
for (;;) {  
  
    /*  
    * 連続する検証の間、thorough_probe_interval  
    * の期間、スリープ状態になる。  
  
    */  
    (void) scds_fm_sleep(scds_handle,  
        scds_get_rs_thorough_probe_interval(scds_handle));  
  
    /*  
    * 使用するすべての IP アドレスを検証する。  
    * 以下をループで検証する。  
    * 1. 使用するすべてのネットワークリソース  
    * 2. 指定されたリソースのすべての IP アドレス  
    * 検証する IP アドレスごとに、障害履歴を計算する。  
    */  
    probe_result = 0;  
    /*  
    * すべてのリソースを繰り返し検証して、svc_probe() の  
    * 呼び出しに使用する各 IP アドレスを取得する。  
    */  
    for (ip = 0; ip < netaddr->num_netaddrs; ip++) {  
        /*  
        * 状態を監視するホスト名とポートを取得する。  
        */  
        hostname = netaddr->netaddrs[ip].hostname;  
        port = netaddr->netaddrs[ip].port_proto.port;  
        /*  
        * HA-XFS がサポートするポートは 1 つだけなので、  
        * ポート値はポートの配列の最初の  
        * エントリから取得する。  
  
        */  
        ht1 = gethrtime(); /* Latch probe start time */  
        scds_syslog(LOG_INFO, "Probing the service on "  
            "port: %d.", port);  
  
        probe_result =  
            svc_probe(scds_handle, hostname, port, timeout);  
  
        /*  
        * サービス検証履歴を更新し、  
        * 必要に応じて、アクションを実行する。  
        * 検証終了時間を取得する。  
  
        */  
        ht2 = gethrtime();  
  
        /* ミリ秒に変換する。*/  
        dt = (ulong_t)((ht2 - ht1) / 1e6);  
  
        /*
```

例 C-5 xfnts_probe.c+ (続き)

```
        * 障害の履歴を計算し、
        * 必要に応じて、アクションを実行する。
        */
        (void) scds_fm_action(scds_handle,
            probe_result, (long)dt);
    } /* ネットワークリソースごと*/
} /* 検証を永続的に繰り返す。*/
}
```

xfnts_start メソッド

データサービスリソースを含むリソースグループがオンラインになったとき、あるいは、リソースが有効になったとき、RGMはそのクラスタノード上で start メソッドを呼び出します。xfnts_start メソッドはそのノード上で xfs デーモンを起動します。

例 C-6 xfnts_start.c

```
/*
 * Copyright (c) 1998-2004 by Sun Microsystems, Inc.
 * All rights reserved.
 *
 * xfnts_svc_start.c - HA-XFS の起動メソッド
 */

#pragma ident "@(#)xfnts_svc_start.c 1.13 01/01/18 SMI"

#include <rgm/libdsdev.h>
#include "xfnts.h"

/*
 * HA-XFS 用の起動メソッド。リソース設定に対していくつかの
 * 健全性検査を行なったあと、アクションスクリプトを使用して HA-XFS を
 * PMF の制御下で起動する。
 */

int
main(int argc, char *argv[])
{
    scds_handle_t    scds_handle;
    int rc;

    /*
     * RGM から渡された引数を処理して、syslog を初期化する。
     */

    if (scds_initialize(&scds_handle, argc, argv) != SCHA_ERR_NOERR)
```

例 C-6 xfnets_start.c (続き)

```
{
    scds_syslog(LOG_ERR, "Failed to initialize the handle.");
    return (1);
}

/* 構成の妥当性を検査する。エラーがあれば戻る。*/
rc = svc_validate(scds_handle);
if (rc != 0) {
    scds_syslog(LOG_ERR,
        "Failed to validate configuration.");
    return (rc);
}

/* データサービスを起動する。失敗した場合、エラーで戻る。*/
rc = svc_start(scds_handle);
if (rc != 0) {
    goto finished;
}

/* サービスが完全に起動するまで待つ。*/
scds_syslog_debug(DBG_LEVEL_HIGH,
    "Calling svc_wait to verify that service has started.");

rc = svc_wait(scds_handle);

scds_syslog_debug(DBG_LEVEL_HIGH,
    "Returned from svc_wait");

if (rc == 0) {
    scds_syslog(LOG_INFO, "Successfully started the service.");
} else {
    scds_syslog(LOG_ERR, "Failed to start the service.");
}

finished:
/* 割り当てられた環境リソースを解放する。*/
scds_close(&scds_handle);

return (rc);
}
```

xfnets_stop メソッド

HA-XFS リソースを含むリソースグループがクラスタのノード上でオフラインになったとき、あるいは、リソースが無効になったとき、RGM はそのクラスタノード上で stop メソッドを呼び出します。xfnets_stop メソッドはそのノード上で xfs デモンを停止します。

例 C-7 xfnts_stop.c

```
/*
 * Copyright (c) 1998-2004 by Sun Microsystems, Inc.
 * All rights reserved.
 *
 * xfnts_svc_stop.c - HA-XFS の停止メソッド
 */

#pragma ident "@(#)xfnts_svc_stop.c 1.10 01/01/18 SMI"

#include <rgm/libdsdev.h>
#include "xfnts.h"

/*
 * PMF を使用して HA-XFS プロセスを停止する。
 */

int
main(int argc, char *argv[])
{
    scds_handle_t    scds_handle;
    int              rc;

    /* RGM から渡された引数を処理して、syslog を初期化する。*/
    if (scds_initialize(&scds_handle, argc, argv) != SCHA_ERR_NOERR)
    {
        scds_syslog(LOG_ERR, "Failed to initialize the handle.");
        return (1);
    }

    rc = svc_stop(scds_handle);

    /* scds_initialize が割り当てたすべてのメモリーを解放する。*/
    scds_close(&scds_handle);

    /* svc_stop メソッドの結果を戻す。*/
    return (rc);
}
```

xfnts_update メソッド

リソースのプロパティが変更されたとき、RGM は Update メソッドを呼び出し、動作中のリソースに通知します。RGM は、管理アクションがリソースまたはそのリソースグループのプロパティの設定に成功したあとに、Update を呼び出します。

例 C-8 xfnts_update.c

```
#pragma ident "@(#)xfnts_update.c 1.10 01/01/18 SMI"

/*
```

例 C-8 xfnts_update.c (続き)

```
* Copyright (c) 1998-2004 by Sun Microsystems, Inc.
* All rights reserved.
*
* xfnts_update.c - HA-XFS の更新メソッド
*/

#include <stdio.h>
#include <stdlib.h>
#include <signal.h>
#include <rgm/libdsdev.h>

/*
 * リソースのプロパティが更新された可能性がある。
 * このような更新可能なプロパティはすべて障害モニターに関連するもので
 * あるため、障害モニターを再起動する必要がある。
 */

int
main(int argc, char *argv[])
{
    scds_handle_t    scds_handle;
    scha_err_t      result;

    /* RGM から渡された引数を処理して、syslog を初期化する。*/
    if (scds_initialize(&scds_handle, argc, argv) != SCHA_ERR_NOERR)
    {
        scds_syslog(LOG_ERR, "Failed to initialize the handle.");
        return (1);
    }

    /*
     * 障害モニターがすでに動作していることを検査し、
     * 動作している場合、障害モニターを停止および再起動する。
     * scds_pmf_restart_fm() への 2 番目のパラメータは、再起動する
     * 必要がある障害モニターのインスタンスを一意に識別する。
     */

    scds_syslog(LOG_INFO, "Restarting the fault monitor.");
    result = scds_pmf_restart_fm(scds_handle, 0);
    if (result != SCHA_ERR_NOERR) {
        scds_syslog(LOG_ERR,
            "Failed to restart fault monitor.");
        /* scds_initialize が割り当てたすべてのメモリーを解放する。*/
        scds_close(&scds_handle);
        return (1);
    }

    scds_syslog(LOG_INFO,
        "Completed successfully.");

    /* scds_initialize が割り当てたすべてのメモリーを解放する。*/
    scds_close(&scds_handle);
}
```

例 C-8 xfnts_update.c (続き)

```
    return (0);
}
```

xfnts_validate メソッドのコードリスト

xfnts_validate メソッドは、Confdir_list プロパティが示すディレクトリの存在を確認します。RGM がこのメソッドを呼び出すのは、クラスタ管理者がデータサービスを作成したときと、データサービスのプロパティを更新したときです。障害モニターがデータサービスを新しいノードにフェイルオーバーしたときは、Monitor_check メソッドは常にこのメソッドを呼び出します。

例 C-9 xfnts_validate.c

```
/*
 * Copyright (c) 1998-2004 by Sun Microsystems, Inc.
 * All rights reserved.
 *
 * xfnts_validate.c - HA-XFS の検証メソッド
 */

#pragma ident "@(#)xfnts_validate.c 1.9 01/01/18 SMI"

#include <rgm/libdsdev.h>
#include "xfnts.h"

/*
 * プロパティが正しく設定されていることを確認する。
 */

int
main(int argc, char *argv[])
{
    scds_handle_t    scds_handle;
    int    rc;

    /* RGM から渡された引数を処理して、syslog を初期化する。*/
    if (scds_initialize(&scds_handle, argc, argv) != SCHA_ERR_NOERR)
    {
        scds_syslog(LOG_ERR, "Failed to initialize the handle.");
        return (1);
    }
    rc = svc_validate(scds_handle);

    /* scds_initialize が割り当てたすべてのメモリーを解放する。*/
```

例 C-9 xfnts_validate.c (続き)

```
scds_close(&scds_handle);  
  
/* 検証メソッドの結果を戻す。*/  
return (rc);  
  
}
```

付録 D

有効な RGM 名と値

この付録では、Resource Group Manager (RGM) の名前と値に指定できる文字の条件について説明します。

有効な RGM 名

RGM 名は、次のカテゴリに分類されます。

- リソースグループ名
- リソースタイプ名
- リソース名
- プロパティ名
- 列挙型リテラル名

命名規則 (リソースタイプ名を除く)

リソースタイプ名を除き、他の名前はすべて次の規則に従う必要があります。

- ASCII であること。
- 先頭は必ず文字にする。
- 名前に使用できる文字は、英字の大文字と小文字、数字、ハイフン (-)、下線 (_)。
- 255 文字を超えないこと。

リソースタイプ名の形式

リソースタイプの完全な名前形式は、リソースタイプが何かによって異なります。

- リソースタイプのリソースタイプ登録 (RTR) ファイルに `#$upgrade` ディレクティブが含まれている場合、形式は次のようになります。

```
vendor-id.base-rt-name:version
```

- リソースタイプの RTR ファイルに `#$upgrade` ディレクティブが含まれていない場合、形式は次のようになります。

```
vendor-id.base-rt-name
```

`vendor-id` と `base-rt-name` の間をピリオドで、`base-rt-name` と `version` の間をコロンでそれぞれ区切ります。

これらの変数の意味は次のとおりです。

<code>vendor-id</code>	ベンダー ID 接頭辞を指定します。ベンダー ID 接頭辞は、RTR ファイル内の <code>Vendor_id</code> リソースタイププロパティの値です。
<code>base-rt-name</code>	ベースリソースタイプ名を指定します。ベースリソースタイプ名は、RTR ファイル内の <code>Resource_type</code> リソースタイププロパティの値です。
<code>version</code>	バージョン接尾辞を指定します。バージョン接尾辞は、RTR ファイル内の <code>RT_version</code> リソースタイププロパティの値です。RTR ファイルに <code>#\$upgrade</code> ディレクティブが含まれている場合、バージョン接尾辞は完全なリソースタイプ名の一部に過ぎません。 <code>#\$upgrade</code> ディレクティブは Sun Cluster のリリース 3.1 で導入されました。

注 - ベースリソースタイプ名のバージョンが 1 つだけ登録されている場合は、`scrgadm(1M)` コマンドに完全な名前を指定する必要はありません。ベンダー ID 接頭辞かバージョン番号接尾辞、またはその両方を省略できます。

リソースタイププロパティについては、239 ページの「リソースタイププロパティ」を参照してください。

例 D-1 リソースタイプの完全な名前 (`#$upgrade` ディレクティブが指定されている場合)

これは完全なリソースタイプ名の例です。このリソースタイプのプロパティが RTR ファイルに次のように設定されているものとします。

- `Vendor_id=SUNW`
- `Resource_type=sample`
- `RT_version=2.0`

RTR ファイルによって定義される完全なリソースタイプ名は次のようになります。

```
SUNW.sample:2.0
```

例 D-2 リソースタイプの完全な名前 (#\$upgrade ディレクティブが指定されていない場合)

これは完全なリソースタイプ名の例です。このリソースタイプのプロパティが RTR ファイルに次のように設定されているものとします。

- Vendor_id=SUNW
- Resource_type=nfs

RTR ファイルによって定義される完全なリソースタイプ名は次のようになります。

SUNW.nfs

RGM の値

RGM の値は、プロパティ値と記述値という 2 つのカテゴリに分類されます。両カテゴリの規則は同じです。

- 値は ASCII であること。
- 値の最大長は 4M - 1 バイト (つまり、4,194,303 バイト) であること。
- NULL、
 - Null
 - 復帰改行
 - コンマ
 - セミコロン

非クラスタ対応のアプリケーションの要件

通常、非クラスタ対応のアプリケーションの高可用性 (HA) を実現するには、特定の要件を満たす必要があります。そのための要件の一覧が、25 ページの「アプリケーションの適合性の分析」に示されています。この付録では、それらの要件のうち、特定のものについて詳細に説明します。

アプリケーションの高可用性を実現するには、そのリソースをリソースグループで構成します。アプリケーションのデータは、高可用性の広域ファイルシステムに格納されます。したがって、1つのサーバーが異常終了しても、正常に動作しているサーバーがデータにアクセスできます。『Sun Cluster の概念 (Solaris OS 版)』のクラスタファイルシステムに関する情報を参照してください。

ネットワーク上のクライアントがネットワークにアクセスする場合、論理ネットワーク IP アドレスは、データサービスリソースと同じリソースグループにある論理ホスト名リソースで構成されます。データサービスリソースとネットワークアドレスリソースは共にフェイルオーバーします。この場合、データサービスのネットワーククライアントは新しいホスト上のデータサービスリソースにアクセスします。

多重ホストデータ

高可用性の広域ファイルシステムのデバイスは多重ホスト化されているため、ある物理ホストがクラッシュしても、正常に動作している物理ホストの1つがデバイスにアクセスできます。アプリケーションの高可用性を実現するには、そのデータが高可用性であること、つまり、そのデータが広域 HA ファイルシステムに格納されていることが必要です。

広域ファイルシステムは、独立したものであるように作成されたデバイスグループにマウントされます。ユーザーは、あるデバイスグループをマウントされた広域ファイルシステムとして使用し、別のディスクグループをデータサービス (HA Oracle など) で使用する raw デバイスとして使用することもできます。

アプリケーションは、データファイルの位置を示すコマンド行スイッチまたは構成ファイルを持っていることがあります。アプリケーションが固定されたパス名を使用する場合は、アプリケーションのコードを変更せずに、このパス名を広域ファイルシステム内のファイルの位置を指すシンボリックリンクに変更できます。シンボリックリンクの詳しい使用方法については、[322 ページの「多重ホストデータを配置するためのシンボリックリンクの使用」](#)を参照してください。

最悪の場合は、実際のデータの位置を示すような何らかの機構を使用するように、アプリケーションのソースコードを変更する必要があります。この作業は、コマンド行スイッチを追加することにより行うことができます。

Sun Cluster は、ボリューム管理ソフトウェアに構成されている UNIX® UFS ファイルシステムと HA の raw デバイスの使用をサポートします。インストールおよび構成するとき、システム管理者はどのディスクリソースを UFS ファイルシステムまたは raw デバイス用に使用するかを指定する必要があります。通常、raw デバイスを使用するのは、データベースサーバーとマルチメディアサーバーだけです。

多重ホストデータを配置するためのシンボリックリンクの使用

アプリケーションの中には、そのデータファイルへのパス名が固定されており、しかも、固定されたパス名を変更する機構がないものがあります。このような場合に、シンボリックリンクを使用すればアプリケーションのコードを変更せずに、済ませられる場合もあります。

たとえば、アプリケーションがそのデータファイルに固定されたパス名 `/etc/mydatafile` を指定すると仮定します。このパスは、論理ホストのファイルシステムの 1 つにあるファイルを示す値を持つシンボリックリンクに変更できます。たとえば、`/global/phys-schost-2/mydatafile` へのシンボリックリンクに変更できます。

ただし、データファイルの名前を内容とともに変更するアプリケーション (または、その管理手順) の場合には、シンボリックリンクをこのように使用すると問題が生じる可能性があります。たとえば、アプリケーションが更新を実行するとき、まず、新しい一時ファイル `/etc/mydatafile.new` を作成すると仮定します。次に、このデータベースは `rename (2)` システムコール (または `mv (1)` プログラム) を使用し、この一時ファイルの名前を実際のファイルの名前に変更します。一時ファイルを作成し、その名前を実際のファイルの名前に変更することにより、データサービスは、そのデータファイルの内容が常に適切であるようにします。

この時、`rename (2)` の操作はシンボリックリンクを破壊します。このため、`/etc/mydatafile` という名前は通常ファイルとなり、クラスタの広域ファイルシステムの中ではなく、`/etc` ディレクトリと同じファイルシステムの中に存在することになります。`/etc` ファイルシステムは各ホスト専用であるため、フェイルオーバーまたはスイッチオーバー後はデータが利用できなくなります。

このような状況の根本的な問題は、既存のアプリケーションがシンボリックリンクに気付かない、つまり、シンボリックリンクを考慮するように作成されていないことにあります。シンボリックリンクを使用し、データアクセスを論理ホストのファイルシ

システムにリダイレクトするには、アプリケーション実装がシンボリックリンクを消去しないように動作する必要があります。したがって、シンボリックリンクは、論理ホストのファイルシステムへのデータの配置に関する問題をすべて解決できるわけではありません。

ホスト名

データサービス開発者は、データサービスが動作しているサーバーのホスト名を、データサービスが知る必要があるかどうかを判断する必要があります。知る必要があると判断した場合は、物理ホストではなく、論理ホストのホスト名 (つまり、アプリケーションリソースと同じリソースグループ内にある論理ホスト名リソース内に構成されているホスト名) を使用するようにデータサービスを変更する必要があります。

データサービスのクライアントサーバープロトコルでは、サーバーが自分のホスト名をクライアントへのメッセージの一部としてクライアントに戻すことがあります。このようなプロトコルでは、クライアントは戻されたホスト名をサーバーに接続するときのホスト名として使用できます。戻されたホスト名をフェイルオーバーやスイッチオーバーが発生した後も使用できるようにするには、物理ホストではなく、リソースグループの論理ホスト名を使用する必要があります。物理ホスト名を使用している場合は、論理ホスト名をクライアントに戻すようにデータサービスのコードを変更する必要があります。

多重ホームホスト

多重ホームホストとは、複数のパブリックネットワーク上にあるホストのことです。このようなホストは複数 (つまり、ネットワークごとに1つ) のホスト名/IP アドレスのペアを持ちます。Sun Cluster は、1つのホストが複数のネットワーク上に存在できるように設計されています。1つのホストが単一のネットワーク上に存在することも可能ですが、このような場合は「多重ホームホスト」とは呼びません。物理ホスト名が複数のホスト名/IP アドレスのペアを持つように、各リソースグループも複数 (つまり、パブリックネットワークごとに1つ) のホスト名/IP アドレスのペアを持ちます。Sun Cluster がリソースグループをある物理ホストから別の物理ホストに移動するとき、そのリソースグループに対するホスト名 / IP アドレスのペアもすべて移動します。

リソースグループに対するホスト名/IP アドレスのペアは、リソースグループに含まれる論理ホスト名リソースとして構成されます。このようなネットワークアドレスリソースは、システム管理者がリソースグループを作成および構成するときに指定します。Sun Cluster データサービス API は、このようなホスト名 / IP アドレスのペアを照会する機能を持っています。

Solaris オペレーティングシステム用に書かれているほとんどの市販のデータサービスデーモンは、多重ホームホストを適切に処理できます。ネットワーク通信を行うとき、多くのデータサービスは Solaris のワイルドカードアドレス `INADDR_ANY` にバインドします。すると、`INADDR_ANY` は、すべてのネットワークインタフェースのすべての IP アドレスを自動的に処理します。`INADDR_ANY` は、現在マシンに構成されているすべての IP アドレスに効率的にバインドします。一般的に、`INADDR_ANY` を使用するデータサービスデーモンは、変更しなくても、Sun Cluster 論理ネットワークアドレスを処理できます。

INADDR_ANY へのバインドと特定の IP アドレスへのバインド

Sun Cluster の論理ネットワークアドレスの概念では、多重ホーム化されていない環境でも、マシンは複数の IP アドレスを持つことができます。つまり、独自の物理ホストの IP アドレスを 1 つだけ持ち、さらに、現在マスターしているネットワークアドレス (論理ホスト名) リソースごとに 1 つの IP アドレスを持ちます。ネットワークアドレスリソースのマスターになるとき、マシンは動的に追加の IP アドレスを獲得します。ネットワークアドレスリソースのマスターを終了するとき、マシンは動的に IP アドレスを放棄します。

データサービスの中には、`INADDR_ANY` にバインドしていると、Sun Cluster 環境で適切に動作しないものもあります。このようなデータサービスは、リソースグループのマスターになるとき、またマスターをやめるときに、バインドしている IP アドレスのセットを動的に変更する必要があります。このようなデータサービスが再バインドする方法の 1 つが、起動メソッドと停止メソッドを使用し、データサービスのデーモンを強制終了および再起動するという方法です。

`Network_resources_used` リソースプロパティを使用すると、エンドユーザーは、アプリケーションリソースをバインドすべきネットワークアドレスリソースを構成できます。この機能が必要なリソースタイプの場合、そのリソースタイプの RTR ファイルで `Network_resources_used` プロパティを宣言する必要があります。

リソースグループをオンラインまたはオフラインにするとき、Sun Cluster は、データサービスリソースメソッドを呼び出す順番に従って、ネットワークアドレスの取り付け (`plumb`)、取り外し (`unplumb`)、「起動」または「停止」への構成を行います。詳細は、41 ページの「[Start および Stop メソッドを使用するかどうかの決定](#)」を参照してください。

データサービスは、`Stop` メソッドが戻るまでに、リソースグループのネットワークアドレスの使用を終了している必要があります。同様に、データサービスは、`Start` メソッドが戻るまでに、リソースグループのネットワークアドレスの使用を開始している必要があります。

データサービスが、個々の IP アドレスではなく、INADDR_ANY にバインドする場合、データサービスリソースメソッドが呼び出される順番とネットワークアドレスメソッドが呼び出される順番には重要な関係があります。

データサービスの停止メソッドと起動メソッドでデータサービスのデーモンを終了および再起動する場合、データサービスは適切な時間にネットワークアドレスの使用を停止および開始します。

クライアントの再試行

ネットワーククライアントから見ると、フェイルオーバーやスイッチオーバーは、論理ホストに障害が発生し、高速再起動しているように見えます。したがって、クライアントアプリケーションとクライアントサーバプロトコルは、このような場合に何回か再試行するように構成されていることが理想的です。すでに、単一サーバの障害と高速再起動を処理するように構成されているアプリケーションとプロトコルは、上記のような場合も、リソースグループのテイクオーバーやスイッチオーバーとして処理します。無限に再試行するようなアプリケーションもあります。また、何回も再試行していることをユーザーに通知し、さらに継続するかどうかをユーザーにたずねるような、より洗練されたアプリケーションもあります。

付録 F

CRNP のドキュメントタイプ定義

この付録では、Cluster Reconfiguration Notification Protocol (CRNP) のドキュメントタイプ定義 (DTD) を示します。

SC_CALLBACK_REG XML DTD

注 - SC_CALLBACK_REG と SC_EVENT の両方で使用する NVPAIR データ構造は、1 度だけ定義します。

```
<!-- SC_CALLBACK_REG XML format specification
      Copyright 2001-2004 Sun Microsystems, Inc. All rights reserved.
      Use is subject to license terms.
```

Intended Use:

A client of the Cluster Reconfiguration Notification Protocol should use this xml format to register initially with the service, to subsequently register for more events, to subsequently remove registration of some events, or to remove itself from the service entirely.

A client is uniquely identified by its callback IP and port. The port is defined in the SC_CALLBACK_REG element, and the IP is taken as the source IP of the registration connection. The final attribute of the root SC_CALLBACK_REG element is either an ADD_CLIENT, ADD_EVENTS, REMOVE_CLIENT, or REMOVE_EVENTS, depending on which form of the message the client is using.

The SC_CALLBACK_REG contains 0 or more SC_EVENT_REG sub-elements.

One SC_EVENT_REG is the specification for one event type. A client may specify only the CLASS (an attribute of the SC_EVENT_REG element), or may specify a SUBCLASS (an optional attribute) for further granularity. Also, the SC_EVENT_REG has as subelements 0 or more

NVPAIRS, which can be used to further specify the event.

Thus, the client can specify events to whatever granularity it wants. Note that a client cannot both register for and unregister for events in the same message. However a client can subscribe to the service and sign up for events in the same message.

Note on versioning: the VERSION attribute of each root element is marked "fixed", which means that all message adhering to these DTDs must have the version value specified. If a new version of the protocol is created, the revised DTDs will have a new value for this "fixed" VERSION attribute, such that all message adhering to the new version must have the new version number.

->

<!-- SC_CALLBACK_REG definition

The root element of the XML document is a registration message. A registration message consists of the callback port and the protocol version as attributes, and either an ADD_CLIENT, ADD_EVENTS, REMOVE_CLIENT, or REMOVE_EVENTS attribute, specifying the registration type. The ADD_CLIENT, ADD_EVENTS, and REMOVE_EVENTS types should have one or more SC_EVENT_REG subelements. The REMOVE_CLIENT should not specify an SC_EVENT_REG subelement.

ATTRIBUTES:

| | |
|----------|---|
| VERSION | The CRNP protocol version of the message. |
| PORT | The callback port. |
| REG_TYPE | The type of registration. One of:
ADD_CLIENT, ADD_EVENTS, REMOVE_CLIENT, REMOVE_EVENTS |

CONTENTS:

SUBELEMENTS: SC_EVENT_REG (0 or more)

->

<!ELEMENT SC_CALLBACK_REG (SC_EVENT_REG*)>

<!-- ATTLIST SC_CALLBACK_REG

| | | |
|----------|---|-----------|
| VERSION | NMTOKEN | #FIXED |
| PORT | NMTOKEN | #REQUIRED |
| REG_TYPE | (ADD_CLIENT ADD_EVENTS REMOVE_CLIENT REMOVE_EVENTS) | #REQUIRED |

>

<!-- SC_EVENT_REG definition

The SC_EVENT_REG defines an event for which the client is either registering or unregistering interest in receiving event notifications. The registration can be for any level of granularity, from only event class down to specific name/value pairs that must be present. Thus, the only required attribute is the CLASS. The SUBCLASS attribute, and the NVPAIRS sub-elements are optional, for higher granularity.

Registrations that specify name/value pairs are registering interest in notification of messages from the class/subclass specified with ALL name/value pairs present. Unregistrations that specify name/value pairs are unregistering interest in notifications that have EXACTLY those name/value pairs in granularity previously specified. Unregistrations that do not specify name/value pairs unregister interest in ALL event notifications of the specified class/subclass.

ATTRIBUTES:

| | |
|--------|--|
| CLASS: | The event class for which this element is registering or unregistering interest. |
|--------|--|


```

                SUBCLASS:      The subclass of the event (optional).

    CONTENTS:
        SUBELEMENTS: 0 or more NVPAIRs.
->

<!ELEMENT SC_EVENT_REG (NVPAIR*)>
<!-- ATTLIST SC_EVENT_REG
        CLASS          CDATA          #REQUIRED
        SUBCLASS       CDATA          #IMPLIED
-->

```

NVPAIR XML DTD

```

<!-- NVPAIR XML format specification

```

```

    Copyright 2001-2004 Sun Microsystems, Inc. All rights reserved.
    Use is subject to license terms.

```

```

    Intended Use:

```

```

        An nvpair element is meant to be used in an SC_EVENT or SC_CALLBACK_REG
        element.

```

```

->

```

```

<!-- NVPAIR definition

```

```

    The NVPAIR is a name/value pair to represent arbitrary name/value combinations.
    It is intended to be a direct, generic, translation of the Solaris nvpair_t
    structure used by the sysevent framework. However, there is no type information
    associated with the name or the value (they are both arbitrary text) in this xml
    element.

```

```

    The NVPAIR consists simply of one NAME element and one or more VALUE elements.
    One VALUE element represents a scalar value, while multiple represent an array
    VALUE.

```

```

    ATTRIBUTES:

```

```

    CONTENTS:

```

```

        SUBELEMENTS: NAME(1), VALUE(1 or more)

```

```

->

```

```

<!ELEMENT NVPAIR (NAME,VALUE+)>

```

```

<!-- NAME definition

```

```

    The NAME is simply an arbitrary length string.

```

```

    ATTRIBUTES:

```

```

    CONTENTS:
        Arbitrary text data. Should be wrapped with <![CDATA[...]]> to prevent XML
        parsing inside.
->
<!ELEMENT NAME (#PCDATA)>

<!-- VALUE definition
    The VALUE is simply an arbitrary length string.

    ATTRIBUTES:

    CONTENTS:
        Arbitrary text data. Should be wrapped with <![CDATA[...]]> to prevent XML
        parsing inside.
->
<!ELEMENT VALUE (#PCDATA)>

```

SC_REPLY XML DTD

```

<!-- SC_REPLY XML format specification

    Copyright 2001-2004 Sun Microsystems, Inc. All rights reserved.
    Use is subject to license terms.
->

<!-- SC_REPLY definition

    The root element of the XML document represents a reply to a message. The reply
    contains a status code and a status message.

    ATTRIBUTES:
        VERSION:          The CRNP protocol version of the message.
        STATUS_CODE:     The return code for the message. One of the
                        following: OK, RETRY, LOW_RESOURCES, SYSTEM_ERROR, FAIL,
                        MALFORMED, INVALID_XML, VERSION_TOO_HIGH, or
                        VERSION_TOO_LOW.

    CONTENTS:
        SUBELEMENTS: SC_STATUS_MSG(1)
->

<!ELEMENT SC_REPLY (SC_STATUS_MSG)>
<!-- ATTLLIST SC_REPLY
    VERSION          NMTOKEN          #FIXED "1.0"
    STATUS_CODE      OK|RETRY|LOW_RESOURCE|SYSTEM_ERROR|FAIL|MALFORMED|INVALID,\
                    VERSION_TOO_HIGH, VERSION_TOO_LOW) #REQUIRED
>

```

```

<!-- SC_STATUS_MSG definition
    The SC_STATUS_MSG is simply an arbitrary text string elaborating on the status
    code.  Should be wrapped with <![CDATA[...]]> to prevent XML parsing inside.

    ATTRIBUTES:

    CONTENTS:
        Arbitrary string.
->

<!ELEMENT SC_STATUS_MSG (#PCDATA)>

```

SC_EVENT XML DTD

注 - SC_CALLBACK_REG と SC_EVENT の両方で使用する NVPAIR データ構造は、1 度だけ定義します。

```

<!-- SC_EVENT XML format specification

    Copyright 2001-2004 Sun Microsystems, Inc. All rights reserved.
    Use is subject to license terms.

    The root element of the XML document is intended to be a direct, generic,
    translation of the Solaris syseventd message format.  It has attributes to
    represent the class, subclass, vendor, and publisher, and contains any number of
    NVPAIR elements.

    ATTRIBUTES:
        VERSION:          The CRNP protocol version of the message.
        CLASS:            The sysevent class of the event
        SUBCLASS:        The subclass of the event
        VENDOR:          The vendor associated with the event
        PUBLISHER:       The publisher of the event

    CONTENTS:
        SUBELEMENTS: NVPAIR (0 or more)
->

<!ELEMENT SC_EVENT (NVPAIR*)>
<!ATTLIST SC_EVENT
    VERSION          NMTOKEN          #FIXED "1.0"
    CLASS            CDATA            #REQUIRED
    SUBCLASS        CDATA            #REQUIRED
    VENDOR           CDATA            #REQUIRED
    PUBLISHER       CDATA            #REQUIRED
>

```


付録 G

CrnpClient.java アプリケーション

この付録では、CrnpClient.java アプリケーションの完全なコードを示します (詳細は、第 12 章を参照)。

CrnpClient.java のコンテンツ

```
/*
 * CrnpClient.java
 * =====
 *
 * 解析についての注意:
 *
 * このプログラムは、Sun Java Architecture for XML Processing (JAXP) API を
 * 使用しています。API ドキュメントや利用についての情報は、
 * http://java.sun.com/xml/jaxp/index.html を参照して下さい。
 *
 * このプログラムは、Java 1.3.1 以降を対象に作成されています。
 *
 * プログラムの概要:
 *
 * このプログラムのメインスレッドは、CrnpClient オブジェクトを作成し、
 * ユーザーがデモを終了するのを待機し、CrnpClient オブジェクトで
 * shutdown を呼び出し、最後にプログラムを終了します。
 *
 * CrnpClient コンストラクタは、EventReceptionThread オブジェクトを作成し、
 * (コマンド行で指定されたホストとポートを使用して) CRNP サーバーに対して
 * 接続を開き、(コマンド行の指定にもとづいて) 登録メッセージを作成し、登録
 * メッセージを送信し、応答の読み取りと解析を行います。
 *
 * EventReceptionThread は、このプログラムが動作するマシンのホスト名と
 * コマンド行に指定されるポートにバインドされる待機ソケットを作成します。
 * EventReceptionThread は、イベントコールバックの着信を待機し、受信した
 * ソケットストリームから XML ドキュメントを構築し、これを CrnpClient*
```

```

* オブジェクトに返して処理を行わせます。
*
* CrnpClient 内の shutdown メソッドは、単に登録解除用の
* (REMOVE_CLIENT) SC_CALLBACK_REG メッセージを crnp サーバーへ送信
* するだけです。
*
* エラー処理についての注意: 説明を簡潔にするため、このプログラムはほとんどの
* エラーに対して単に終了するだけですが、実際のアプリケーションではさまざまな
* 方法でエラー処理がなされます (適宜再試行するなど)。
*/

// JAXP パッケージ
import javax.xml.parsers.*;
import javax.xml.transform.*;
import javax.xml.transform.dom.*;
import javax.xml.transform.stream.*;
import org.xml.sax.*;
import org.xml.sax.helpers.*;
import org.w3c.dom.*;

// 標準パッケージ
import java.net.*;
import java.io.*;
import java.util.*;

/*
* クラス CrnpClient
* -----
* 上記のファイルヘッダーコメントを参照。
*/
class CrnpClient
{
    /*
    * main
    * ----
    * 実行のエントリーポイント main は、コマンド行引数の数を
    * 検証し、すべての作業を行う CrnpClient インスタンスを
    * 作成する。
    */
    public static void main(String []args)
    {
        InetAddress regIp = null;
        int regPort = 0, localPort = 0;

        /* コマンド行引数の数を検証する */
        if (args.length < 4) {
            System.out.println(
                "Usage: java CrnpClient crnpHost crnpPort "
                + "localPort (-ac | -ae | -re) "
                + "[(M | A | RG=name | R=name) [...] ]");
            System.exit(1);
        }

        /*

```

```

    * コマンド行には crnp サーバーの IP とポート、
    * 待機するローカルポート、登録タイプを示す
    * 引数が入る。
    */
    try {
        regIp = InetAddress.getByName(args[0]);
        regPort = (new Integer(args[1])).intValue();
        localPort = (new Integer(args[2])).intValue();
    } catch (UnknownHostException e) {
        System.out.println(e);
        System.exit(1);
    }

    // CrnpClient を作成する。
    CrnpClient client = new CrnpClient(regIp, regPort, localPort,
        args);

    // ユーザーがプログラムを終了したくなるまで待機する。
    System.out.println("Hit return to terminate demo...");

    // ユーザーが何か入力するまで読み取りはブロックする。
    try {
        System.in.read();
    } catch (IOException e) {
        System.out.println(e.toString());
    }

    // クライアントを停止する。
    client.shutdown();
    System.exit(0);
}

/*
 * =====
 * public メソッド
 * =====
 */

/*
 * CrnpClient コンストラクタ
 * -----
 * crnp サーバーとの通信方法を知るためにコマンド行引数を解析
 * し、イベント受信スレッドを作成し、このスレッドの実行を開始
 * し、XML DocumentBuilderFactory オブジェクトを作成し、
 * 最後に crnp サーバーにコールバックの登録を行う。
 */
public CrnpClient(InetAddress regIpIn, int regPortIn, int localPortIn,
    String []clArgs)
{
    try {
        regIp = regIpIn;
        regPort = regPortIn;
        localPort = localPortIn;
        regs = clArgs;
    }
}

```

```

    /*
     * xml 処理用のドキュメントビルダー
     * ファクトリを設定する。
     */
    setupXmlProcessing();

    /*
     * ServerSocket を作成してこれをローカル IP と
     * ポートにバインドする EventReception を作成する。
     */
    createEvtRecepThr();

    /*
     * crnp サーバーに登録する。
     */
    registerCallbacks();

} catch (Exception e) {
    System.out.println(e.toString());
    System.exit(1);
}
}

/*
 * processEvent
 * -----
 * CrnpClient (イベントコールバックを受信する際に
 * EventReceptionThread によって使用される) にコールバックする。
 */
public void processEvent(Event event)
{
    /*
     * ここでは、説明の都合上、単純にイベントを System.out
     * へ出力。実際のアプリケーションでは、通常、イベント
     * をなんらかの方法で使用する。
     */
    event.print(System.out);
}

/*
 * shutdown
 * -----
 * CRNP サーバーに対する登録を解除する。
 */
public void shutdown()
{
    try {
        /* サーバーに登録解除メッセージを送信する */
        unregister();
    } catch (Exception e) {
        System.out.println(e);
        System.exit(1);
    }
}
}

```



```

/*
 * =====
 * private ヘルパーメソッド
 * =====
 */

/*
 * setupXmlProcessing
 * -----
 * xml 応答と xml イベントを解析するためにドキュメント
 * ビルダーファクトリを作成する。
 */
private void setupXmlProcessing() throws Exception
{
    dbf = DocumentBuilderFactory.newInstance();

    // わざわざ検証する必要はない。
    dbf.setValidating(false);
    dbf.setExpandEntityReferences(false);

    // コメントと空白文字は無視したい。
    dbf.setIgnoringComments(true);
    dbf.setIgnoringElementContentWhitespace(true);

    // CDATA セクションを TEXT ノードに結合する。
    dbf.setCoalescing(true);
}

/*
 * createEvtRecepThr
 * -----
 * 新しい EventReceptionThread オブジェクトを作成し、待機
 * ソケットがバインドされる IP とポートを保存し、スレッドの
 * 実行を開始する。
 */
private void createEvtRecepThr() throws Exception
{
    /* スレッドオブジェクトを作成する */
    evtThr = new EventReceptionThread(this);

    /*
     * イベント配信コールバックを待機し始めるために
     * スレッドの実行を開始する。
     */
    evtThr.start();
}

/*
 * registerCallbacks
 * -----
 * crnp サーバーに対するソケット接続を作成し、
 * イベント登録メッセージを送信する。

```

```

*/
private void registerCallbacks() throws Exception
{
    System.out.println("About to register");

    /*
     * crnp サーバーの登録 IP / ポートに接続されたソケット
     * を作成し、登録情報を送信する。
     */
    Socket sock = new Socket(regIp, regPort);
    String xmlStr = createRegistrationString();
    PrintStream ps = new PrintStream(sock.getOutputStream());
    ps.print(xmlStr);

    /*
     * 応答を読み取る。
     */
    readRegistrationReply(sock.getInputStream());

    /*
     * ソケット接続を閉じる。
     */
    sock.close();
}

/*
 * unregister
 * -----
 * registerCallbacks の場合と同様に、crnp サーバーに対する
 * ソケット接続を作成し、登録解除メッセージを送信し、
 * サーバーからの応答を待機し、ソケットを閉じる。
 */
private void unregister() throws Exception
{
    System.out.println("About to unregister");

    /*
     * crnp サーバーの登録 IP / ポートに接続された
     * ソケットを作成し、登録解除情報を送信する。
     */
    Socket sock = new Socket(regIp, regPort);
    String xmlStr = createUnregistrationString();
    PrintStream ps = new PrintStream(sock.getOutputStream());
    ps.print(xmlStr);

    /*
     * 応答を読み取る。
     */
    readRegistrationReply(sock.getInputStream());

    /*
     * ソケット接続を閉じる。
     */
    sock.close();
}

```

```

/*
 * createRegistrationString
 * -----
 * このプログラムのコマンド行引数にもとづいて CallbackReg
 * オブジェクトを作成し、CallbackReg オブジェクトから XML
 * 文字列を取得する。
 */
private String createRegistrationString() throws Exception
{
    /*
     * 実際の CallbackReg クラスを作成し、ポートを設定する。
     */
    CallbackReg cbReg = new CallbackReg();
    cbReg.setPort("" + localPort);

    // 登録タイプを設定する。
    if (regs[3].equals("-ac")) {
        cbReg.setRegType(CallbackReg.ADD_CLIENT);
    } else if (regs[3].equals("-ae")) {
        cbReg.setRegType(CallbackReg.ADD_EVENTS);
    } else if (regs[3].equals("-re")) {
        cbReg.setRegType(CallbackReg.REMOVE_EVENTS);
    } else {
        System.out.println("Invalid reg type: " + regs[3]);
        System.exit(1);
    }

    // イベントを追加する。
    for (int i = 4; i < regs.length; i++) {
        if (regs[i].equals("M")) {
            cbReg.addRegEvent(
                createMembershipEvent());
        } else if (regs[i].equals("A")) {
            cbReg.addRegEvent(
                createAllEvent());
        } else if (regs[i].substring(0,2).equals("RG")) {
            cbReg.addRegEvent(createRgEvent(
                regs[i].substring(3)));
        } else if (regs[i].substring(0,1).equals("R")) {
            cbReg.addRegEvent(createREvent(
                regs[i].substring(2)));
        }
    }

    String xmlStr = cbReg.convertToXml();
    System.out.println(xmlStr);
    return (xmlStr);
}

/*
 * createAllEvent
 * -----
 * クラス EC_Cluster を使用して (サブクラスは使用しない)
 * XML 登録イベントを作成する。

```

```

*/
private Event createAllEvent()
{
    Event allEvent = new Event();
    allEvent.setClass("EC_Cluster");
    return (allEvent);
}

/*
 * createMembershipEvent
 * -----
 * クラス EC_Cluster、サブクラス ESC_cluster_memberhip を
 * 使用して XML 登録イベントを作成する。
 */
private Event createMembershipEvent()
{
    Event membershipEvent = new Event();
    membershipEvent.setClass("EC_Cluster");
    membershipEvent.setSubclass("ESC_cluster_membership");
    return (membershipEvent);
}

/*
 * createRgEvent
 * -----
 * クラス EC_Cluster、サブクラス ESC_cluster_rg_state、
 * および "rg_name" nvpair (入力パラメータにもとづく) を
 * 1 つ使用して XML 登録イベントを作成する。
 */
private Event createRgEvent(String rgname)
{
    /*
     * rgname リソースグループ用の
     * リソースグループ状態変更イベントを作成する。
     * このイベントタイプには、どのリソースグループに興味
     * があるのかを示すため、名前/ 値ペア (nvpair) を指定
     * する。
     */
    /*
     * イベントオブジェクトを作成し、クラスとサブクラスを設定する。
     */
    Event rgStateEvent = new Event();
    rgStateEvent.setClass("EC_Cluster");
    rgStateEvent.setSubclass("ESC_cluster_rg_state");

    /*
     * nvpair オブジェクトを作成し、これをこのイベントに追加する。
     */
    NVPair rgNvpair = new NVPair();
    rgNvpair.setName("rg_name");
    rgNvpair.setValue(rgname);
    rgStateEvent.addNvpair(rgNvpair);

    return (rgStateEvent);
}

```

```

/*
 * createREvent
 * -----
 * クラス EC_Cluster、サブクラス ESC_cluster_r_state、
 * および "r_name" nvpair (入力パラメータにもとづく) を
 * 1 つ使用して XML 登録イベントを作成する。
 */
private Event createREvent(String rname)
{
    /*
     * rgname リソース用のリソース
     * 状態変更イベントを作成する。
     * このイベントタイプには、どのリソースグループに興味
     * があるかを示すため、名前/ 値ペア (nvpair) を指定
     * する。
     */
    Event rStateEvent = new Event();
    rStateEvent.setClass("EC_Cluster");
    rStateEvent.setSubclass("ESC_cluster_r_state");

    NVPair rNvpair = new NVPair();
    rNvpair.setName("r_name");
    rNvpair.setValue(rname);
    rStateEvent.addNvpair(rNvpair);

    return (rStateEvent);
}

/*
 * createUnregistrationString
 * -----
 * REMOVE_CLIENT CallbackReg オブジェクトを作成し、
 * CallbackReg オブジェクトから XML 文字列を取得する。
 */
private String createUnregistrationString() throws Exception
{
    /*
     * CallbackReg オブジェクトを作成する。
     */
    CallbackReg cbReg = new CallbackReg();
    cbReg.setPort("" + localPort);
    cbReg.setRegType(CallbackReg.REMOVE_CLIENT);

    /*
     * 登録を OutputStream に整形化する。
     */
    String xmlStr = cbReg.convertToXml();

    // デバッグのために文字列を出力する。
    System.out.println(xmlStr);
    return (xmlStr);
}

```

```

/*
 * readRegistrationReply
 * -----
 * xml を解析してドキュメントにし、このドキュメントから
 * RegReply オブジェクトを構築し、RegReply オブジェクトを
 * 出力する。実際のアプリケーションでは、通常、RegReply
 * オブジェクトの status_code にもとづいて処理をする。
 */
private void readRegistrationReply(InputStream stream)
    throws Exception
{
    // ドキュメントビルダーを作成する。
    DocumentBuilder db = dbf.newDocumentBuilder();

    //
    // 解析前に ErrorHandler を設定する。
    // ここではデフォルトハンドラを使用。
    //
    db.setErrorHandler(new DefaultHandler());

    // 入力ファイルを解析する。
    Document doc = db.parse(stream);

    RegReply reply = new RegReply(doc);
    reply.print(System.out);
}

/* private 指定のメンバー変数 */
private InetAddress regIp;
private int regPort;
private EventReceptionThread evtThr;
private String regs[];

/* public 指定のメンバー変数 */
public int localPort;
public DocumentBuilderFactory dbf;
}

/*
 * クラス EventReceptionThread
 * -----
 * 上記のファイルヘッダーコメントを参照。
 */
class EventReceptionThread extends Thread
{
    /*
     * EventReceptionThread コンストラクタ
     * -----
     * ローカルホスト名とワイルドカードポートにバインドされる
     * 新しい ServerSocket を作成する。
     */
    public EventReceptionThread(CrnpClient clientIn) throws IOException
    {
        /*
         * イベントの取得時に再度呼び返すことができるように、

```

```

    * クライアントに対する参照を保持する。
    */
    client = clientIn;

    /*
    * バインドする IP を指定する。これは、ローカル
    * ホスト IP である。このマシンに複数のパブリック
    * インタフェースが構成されている場合は、
    * InetAddress.getLocalHost によって検出される
    * ものをどれでも使用する。
    *
    */
    listeningSock = new ServerSocket(client.localPort, 50,
        InetAddress.getLocalHost());
    System.out.println(listeningSock);
}

/*
* run
* ---
* Thread.Start メソッドによって呼び出される。
*
* ServerSocket で着信接続を待機し、永続的にループする。
*
* 各着信接続が受け入れられる際に xml ストリームから
* Event オブジェクトが作成される。続いてこのオブジェクト
* が CrnpClient オブジェクトに返されて処理される。
*/
public void run()
{
    /*
    * 永続的にループする。
    */
    try {
        //
        // CrnpClient 内のドキュメントビルダーファクトリを
        // 使用してドキュメントビルダーを作成する。
        //
        DocumentBuilder db = client.dbf.newDocumentBuilder();

        //
        // 解析前に ErrorHandler を設定する。
        // ここではデフォルトハンドラを使用。
        //
        db.setErrorHandler(new DefaultHandler());

        while(true) {
            /* サーバーからのコールバックを待機 */
            Socket sock = listeningSock.accept();

            // 入力ファイルを解析する。
            Document doc = db.parse(sock.getInputStream());

            Event event = new Event(doc);
            client.processEvent(event);
        }
    }
}

```

```

        /* ソケットを閉じる */
        sock.close();
    }
    // 到達不能

    } catch (Exception e) {
        System.out.println(e);
        System.exit(1);
    }
}

/* private 指定のメンバー変数 */
private ServerSocket listeningSock;
private CrnpClient client;
}

/*
 * クラス NVPair
 * -----
 * このクラスは名前/ 値ペア (両方とも文字列) を格納する。
 * このクラスは、そのメンバーから NVPAIR XML メッセージを構築し、
 * NVPAIR XML 要素を解析してそのメンバーにすることができる。
 *
 * NVPAIR の形式仕様では複数の値が許可されているが、ここでは
 * 単純に値は 1 つだけという前提を下す。
 */
class NVPair
{
    /*
     * 2 つのコンストラクタ: 最初のコンストラクタは空の NVPair を
     * 作成する。2 つ目は NVPAIR XML 要素から NVPair を作成する。
     */
    public NVPair()
    {
        name = value = null;
    }

    public NVPair(Element elem)
    {
        retrieveValues(elem);
    }

    /*
     * Public 指定のセッター。
     */
    public void setName(String nameIn)
    {
        name = nameIn;
    }

    public void setValue(String valueIn)
    {
        value = valueIn;
    }
}

```



```

/*
 * 1 行で名前と値を出力する。
 */
public void print(PrintStream out)
{
    out.println("NAME=" + name + " VALUE=" + value);
}

/*
 * createXmlElement
 * -----
 * メンバー変数から NVPAIR XML 要素を作成する。
 * この要素を作成できるように、ドキュメントをパラメータとして
 * 受け付ける。
 */
public Element createXmlElement(Document doc)
{
    // 要素を作成する。
    Element nvpair = (Element)
        doc.createElement("NVPAIR");
    //
    // 名前を追加する。実際の名前は別の
    // CDATA セクションであることに注意。
    //
    Element eName = doc.createElement("NAME");
    Node nameData = doc.createCDATASection(name);
    eName.appendChild(nameData);
    nvpair.appendChild(eName);
    //
    // 値を追加する。実際の値は別の
    // CDATA セクションであることに注意。
    //
    Element eValue = doc.createElement("VALUE");
    Node valueData = doc.createCDATASection(value);
    eValue.appendChild(valueData);
    nvpair.appendChild(eValue);

    return (nvpair);
}

/*
 * retrieveValues
 * -----
 * XML 要素を解析して名前と値を取得する。
 */
private void retrieveValues(Element elem)
{
    Node n;
    NodeList nl;

    //
    // NAME 要素を検出する。
    //
    nl = elem.getElementsByTagName("NAME");

```

```

    if (nl.getLength() != 1) {
        System.out.println("Error in parsing: can't find "
            + "NAME node.");
        return;
    }

    //
    // TEXT セクションを取得する。
    //
    n = nl.item(0).getFirstChild();
    if (n == null || n.getNodeType() != Node.TEXT_NODE) {
        System.out.println("Error in parsing: can't find "
            + "TEXT section.");
        return;
    }

    // 値を取得する。
    name = n.getNodeValue();

    //
    // ここで値要素を取得する。
    //
    nl = elem.getElementsByTagName("VALUE");
    if (nl.getLength() != 1) {
        System.out.println("Error in parsing: can't find "
            + "VALUE node.");
        return;
    }

    //
    // TEXT セクションを取得する。
    //
    n = nl.item(0).getFirstChild();
    if (n == null || n.getNodeType() != Node.TEXT_NODE) {
        System.out.println("Error in parsing: can't find "
            + "TEXT section.");
        return;
    }

    // 値を取得する。
    value = n.getNodeValue();
}

/*
 * Public 指定のアクセッサ
 */
public String getName()
{
    return (name);
}

public String getValue()
{
    return (value);
}

```

```

    }

    // Private 指定のメンバー変数
    private String name, valu;
}

/*
 * クラス Event
 * -----
 * このクラスは、クラス、サブクラス、ベンダー、パブリッシャー、名前/
 * 値ペアのリストから成るイベントを格納する。このクラスでは、そのメンバー
 * から SC_EVENT_REG XML 要素を作成し、この要素を解析してそのメンバーに
 * することができる。次の非対称性に注意: SC_EVENT 要素を解析するが、
 * 作成するのは SC_EVENT_REG 要素である。これは、SC_EVENT_REG 要素が
 * 登録メッセージ (これは作成の必要がある) 内で使用され、SC_EVENT 要素
 * がイベント配信 (これは解析の必要がある) 内で使用されるためである。
 * 違いは、SC_EVENT_REG 要素にはベンダーとパブリッシャーがないことだけ
 * である。
 */
class Event
{
    /*
     * 2 つのコンストラクタ: 最初のコンストラクタは空のイベントを
     * 作成し、2 つ目は SC_EVENT XML ドキュメントからイベントを
     * 作成する。
     */
    public Event()
    {
        regClass = regSubclass = null;
        nvpairs = new Vector();
    }

    public Event(Document doc)
    {
        nvpairs = new Vector();

        //
        // デバッグで使用できるようにドキュメントを文字列に
        // 変換して出力する。
        //
        DOMSource domSource = new DOMSource(doc);
        StringWriter strWrite = new StringWriter();
        StreamResult streamResult = new StreamResult(strWrite);
        TransformerFactory tf = TransformerFactory.newInstance();
        try {
            Transformer transformer = tf.newTransformer();
            transformer.transform(domSource, streamResult);
        } catch (TransformerException e) {
            System.out.println(e.toString());
            return;
        }
        System.out.println(strWrite.toString());
    }
}

```

```

        // 実際に解析する。
        retrieveValues(doc);
    }

    /*
     * Public 指定のセッター。
     */
    public void setClass(String classIn)
    {
        regClass = classIn;
    }

    public void setSubclass(String subclassIn)
    {
        regSubclass = subclassIn;
    }

    public void addNvpair(NVPair nvpair)
    {
        nvpairs.add(nvpair);
    }

    /*
     * createXmlElement
     * -----
     * メンバー変数から SC_EVENT_REG XML 要素を作成する。
     * この要素を作成できるように、ドキュメントをパラメータとして
     * 受け付ける。NVPair createXmlElement 機能を使用。
     */
    public Element createXmlElement(Document doc)
    {
        Element event = (Element)
            doc.createElement("SC_EVENT_REG");
        event.setAttribute("CLASS", regClass);
        if (regSubclass != null) {
            event.setAttribute("SUBCLASS", regSubclass);
        }
        for (int i = 0; i < nvpairs.size(); i++) {
            NVPair tempNv = (NVPair)
                (nvpairs.elementAt(i));
            event.appendChild(tempNv.createXmlElement(
                doc));
        }
        return (event);
    }

    /*
     * メンバー変数を複数行に出力する。
     */
    public void print(PrintStream out)
    {
        out.println("\tCLASS=" + regClass);
        out.println("\tSUBCLASS=" + regSubclass);
        out.println("\tVENDOR=" + vendor);
    }

```

```

        out.println("\tPUBLISHER=" + publisher);
        for (int i = 0; i < nvpairs.size(); i++) {
            NVPair tempNv = (NVPair)
                (nvpairs.elementAt(i));
            out.print("\t\t");
            tempNv.print(out);
        }
    }

    /*
    * retrieveValues
    * -----
    * XML ドキュメントを解析し、クラス、サブクラス、ベンダー、
    * パブリッシャー、および nvpair を取得する。
    */
    private void retrieveValues(Document doc)
    {
        Node n;
        NodeList nl;

        //
        // SC_EVENT 要素を検出する。
        //
        nl = doc.getElementsByTagName("SC_EVENT");
        if (nl.getLength() != 1) {
            System.out.println("Error in parsing: can't find "
                + "SC_EVENT node.");
            return;
        }

        n = nl.item(0);

        //
        // CLASS, SUBCLASS, VENDOR, および PUBLISHER
        // 属性の値を取得する。
        //
        regClass = ((Element)n).getAttribute("CLASS");
        regSubclass = ((Element)n).getAttribute("SUBCLASS");
        publisher = ((Element)n).getAttribute("PUBLISHER");
        vendor = ((Element)n).getAttribute("VENDOR");

        //
        // すべての nv ペアを取得する。
        //
        for (Node child = n.getFirstChild(); child != null;
            child = child.getNextSibling())
        {
            nvpairs.add(new NVPair((Element)child));
        }
    }

    /*
    * public 指定のアクセッサメソッド。
    */
    public String getRegClass()

```

```

    {
        return (regClass);
    }

    public String getSubclass()
    {
        return (regSubclass);
    }

    public String getVendor()
    {
        return (vendor);
    }

    public String getPublisher()
    {
        return (publisher);
    }

    public Vector getNvpairs()
    {
        return (nvpairs);
    }

    // Private 指定のメンバー変数
    private String regClass, regSubclass;
    private Vector nvpairs;
    private String vendor, publisher;
}

/*
 * クラス CallbackReg
 * -----
 * このクラスは、ポートと登録タイプ（どちらも文字列）、およびイベントリストを
 * 格納する。このクラスは、そのメンバーから SC_CALLBACK_REG XML メッセージ
 * を作成できる。
 *
 * SC_CALLBACK_REG メッセージを解析する必要があるのは CRNP サーバー
 * だけであるため、このクラスで SC_CALLBACK_REG メッセージを解析でき
 * なくてもよい。
 */
class CallbackReg
{
    // setRegType メソッドに便利な定義
    public static final int ADD_CLIENT = 0;
    public static final int ADD_EVENTS = 1;
    public static final int REMOVE_EVENTS = 2;
    public static final int REMOVE_CLIENT = 3;

    public CallbackReg()
    {
        port = null;
        regType = null;
        regEvents = new Vector();
    }
}

```

```

}

/*
 * Public 指定のセッター。
 */
public void setPort(String portIn)
{
    port = portIn;
}

public void setRegType(int regTypeIn)
{
    switch (regTypeIn) {
        case ADD_CLIENT:
            regType = "ADD_CLIENT";
            break;
        case ADD_EVENTS:
            regType = "ADD_EVENTS";
            break;
        case REMOVE_CLIENT:
            regType = "REMOVE_CLIENT";
            break;
        case REMOVE_EVENTS:
            regType = "REMOVE_EVENTS";
            break;
        default:
            System.out.println("Error, invalid regType " +
                regTypeIn);
            regType = "ADD_CLIENT";
            break;
    }
}

public void addRegEvent(Event regEvent)
{
    regEvents.add(regEvent);
}

/*
 * convertToXml
 * -----
 * メンバー変数から SC_CALLBACK_REG XML ドキュメントを構築する。
 * Event createElement 機能を使用。
 */
public String convertToXml()
{
    Document document = null;
    DocumentBuilderFactory factory =
        DocumentBuilderFactory.newInstance();
    try {
        DocumentBuilder builder = factory.newDocumentBuilder();
        document = builder.newDocument();
    } catch (ParserConfigurationException pce) {
        // 指定したオプションを持つパーサーを構築できない。
    }
}

```

```

        pce.printStackTrace();
        System.exit(1);
    }
    Element root = (Element) document.createElement(
        "SC_CALLBACK_REG");
    root.setAttribute("VERSION", "1.0");
    root.setAttribute("PORT", port);
    root.setAttribute("REG_TYPE", regType);
    for (int i = 0; i < regEvents.size(); i++) {
        Event tempEvent = (Event)
            (regEvents.elementAt(i));
        root.appendChild(tempEvent.createXmlElement(
            document));
    }
    document.appendChild(root);

    //
    // ここでドキュメントを文字列に変換する。
    //
    DOMSource domSource = new DOMSource(document);
    StringWriter strWrite = new StringWriter();
    StreamResult streamResult = new StreamResult(strWrite);
    TransformerFactory tf = TransformerFactory.newInstance();
    try {
        Transformer transformer = tf.newTransformer();
        transformer.transform(domSource, streamResult);
    } catch (TransformerException e) {
        System.out.println(e.toString());
        return ("");
    }
    return (strWrite.toString());
}

// private 指定のメンバー変数
private String port;
private String regType;
private Vector regEvents;
}

/*
 * クラス RegReply
 * -----
 * このクラスは、status_code と status_msg (どちらも文字列) を格納する。
 * このクラスは、SC_REPLY XML 要素を解析し、そのメンバーにできる。
 */
class RegReply
{
    /*
     * 1 つのコンストラクタが XML ドキュメントを受け入れて解析を行う。
     */
    public RegReply(Document doc)
    {
        //
        // ここでドキュメントを文字列に変換する。
    }
}

```



```

//
DOMSource domSource = new DOMSource(doc);
StringWriter strWrite = new StringWriter();
StreamResult streamResult = new StreamResult(strWrite);
TransformerFactory tf = TransformerFactory.newInstance();
try {
    Transformer transformer = tf.newTransformer();
    transformer.transform(domSource, streamResult);
} catch (TransformerException e) {
    System.out.println(e.toString());
    return;
}
System.out.println(strWrite.toString());

retrieveValues(doc);
}

/*
 * Public 指定のアクセッサ
 */
public String getStatusCode()
{
    return (statusCode);
}

public String getStatusMsg()
{
    return (statusMsg);
}

/*
 * 1 行で情報を出力する。
 */
public void print(PrintStream out)
{
    out.println(statusCode + ": " +
        (statusMsg != null ? statusMsg : ""));
}

/*
 * retrieveValues
 * -----
 * XML ドキュメントを解析し、statusCode と statusMsg を取得する。
 */
private void retrieveValues(Document doc)
{
    Node n;
    NodeList nl;

    //
    // SC_REPLY 要素を検出する。
    //
    nl = doc.getElementsByTagName("SC_REPLY");
    if (nl.getLength() != 1) {
        System.out.println("Error in parsing: can't find "

```

```

        + "SC_REPLY node.");
    return;
}

n = nl.item(0);

// STATUS_CODE 属性の値を取得する。
statusCode = ((Element)n).getAttribute("STATUS_CODE");

//
// SC_STATUS_MSG 要素を検出する。
//
nl = ((Element)n).getElementsByTagName("SC_STATUS_MSG");
if (nl.getLength() != 1) {
    System.out.println("Error in parsing: can't find "
        + "SC_STATUS_MSG node.");
    return;
}

//
// TEXT セクションが存在する場合は、それを取得する。
//
n = nl.item(0).getFirstChild();
if (n == null || n.getNodeType() != Node.TEXT_NODE) {
    // 存在しなくてもエラーではないため、
    // このまま戻る。
    return;
}

// 値を取得する。
statusMsg = n.getNodeValue();
}

// private 指定のメンバー変数。
private String statusCode;
private String statusMsg;
}

```

索引

数字・記号

- #\$upgrade_from directive, 58
- #\$upgrade_from ディレクティブ, 57
 - ANYTIME, 58
 - AT_CREATION, 59
 - Tunable 属性の値, 58
 - WHEN_DISABLED, 59
 - WHEN_OFFLINE, 58
 - WHEN_UNMANAGED, 59
 - WHEN_UNMONITORED, 58
- #\$upgrade ディレクティブ, 318

A

- Affinity_timeout, リソースプロパティ, 246
- Agent Builder
 - Cluster Agent モジュール, 180
 - 違い, 184
 - rtconfig ファイル, 180
 - アプリケーションの分析, 159
 - インストール, 160
 - 画面の構成, 168
 - 完成した作業内容の再利用, 173
 - 既存のリソースタイプのクローン作成, 173
 - 起動, 162, 194
 - 構成, 160
 - コマンド行バージョン, 175
 - 作成画面, 166
 - サポートファイル, 179
 - 出力, 198
 - 使用, 159

Agent Builder (続き)

- スクリプト, 178
- 生成されたソースコードの編集, 174
- 説明, 18, 23
- ソースファイル, 176
- で GDS を作成, 189
- ディレクトリ構造, 175
- ナビゲーション, 162
 - 「ファイル」メニュー, 165
 - 「ブラウズ」, 163
 - 「編集」メニュー, 165
 - メニュー, 165
- のコマンド行バージョンを使って GDS ベースのサービスを作成, 201
- バイナリファイル, 176
- パッケージディレクトリ, 179
- マニュアルページ, 178
- を使って、GDS を使用するサービスを作成, 194
- Agent Builder のインストール, 160
- ANYTIME, #\$upgrade_from ディレクティブ, 58
- API_version, リソースタイププロパティ, 240
- API、Resource Management, 「RMAPI」を参照
- Array_maxsize, リソースプロパティ属性, 264
- Array_minsize, リソースプロパティ属性, 264
- arraymax, リソースタイプの移行, 56
- arraymin, リソースタイプの移行, 56

AT_CREATION, #`$upgrade_from` ディレク
タイプ, 59
Auto_start_on_new_cluster, リソースグ
ループプロパティ, 257

B

Boot, リソースタイププロパティ, 240
Boot メソッド、使用, 43
Bootメソッド、使用, 79

C

CCR (クラスタ構成リポジトリ), 62
Cheap_probe_interval, リソースプロパ
ティ, 246
Cluster Reconfiguration Notification Protocol,
「CRNP」を参照
Cluster Agent モジュール
Agent Builder との違い, 184
インストール, 180
起動, 181
使用, 183
設定, 180
説明, 180
CRNP
Java アプリケーションの例, 221
SC_CALLBACK_REG メッセージ, 213
エラー状況, 216
機能, 210
クライアント, 213
クライアント識別プロセス, 213
クライアントとサーバーの登録, 213
サーバー, 213
サーバーイベントの配信, 217
サーバーの応答, 215
説明, 209
通信, 211
認証, 221
プロトコル, 210
プロトコルの意味論, 211
メッセージのタイプ, 212
CRNP クライアントとサーバーの登録, 213
C プログラムの関数, RMAPI, 73

D

Default, リソースプロパティ属性, 264
Description, リソースプロパティ属性, 264
Desired primaries, リソースグループプロ
パティ, 257
DSDL
libdsdev.so, 18
PMF (Process Monitor Facility) 関数, 206
概要, 18
関数, 203
高可用性ローカルファイルシステムの有効
化, 121
コンポーネント, 23
実装される, 18
障害モニター, 206
障害モニター関数, 207
障害モニターの実装, 119
説明, 117, 118
データサービスの起動, 118
データサービスの停止, 118
ネットワークアドレスのアクセス, 120
ネットワークリソースアクセス関数, 205
汎用関数, 203
プロパティ関数, 205
ユーティリティー関数, 207
リソースタイプ実装のサンプル
scds_initialize() 関数, 138
SUNW.xfnts 障害モニター, 147
SUNW.xfnts の RTR ファイル, 137
svc_probe() 関数, 149
svc_start() からの復帰, 140
TCP ポート番号, 136
X Font Server, 135
xfnts_monitor_check メソッド, 146
xfnts_monitor_start メソッド, 144
xfnts_monitor_stop メソッド, 145
xfnts_probe のメインループ, 147
xfnts_start メソッド, 138
xfnts_stop メソッド, 142
xfnts_update メソッド, 154
xfnts_validate メソッド, 152
サービスの起動, 139
サービスの検証, 138
障害モニターのアクションの決定, 151
リソースタイプ実装の例
X Font Server の構成ファイル, 136
リソースタイプのデバッグ, 120

DSDLによる高可用性ローカルファイルシステムの有効化, 121
DSDLによるデータサービスの起動, 118
DSDLによるデータサービスの停止, 118
DSDLによるリソースタイプのデバッグ, 120

E

Enumlist, リソースプロパティ属性, 264
Extension, リソースプロパティ属性, 264

F

Failback, リソースグループプロパティ, 258
Failover, リソースタイププロパティ, 240
Failover_mode, リソースプロパティ, 247
Fini, リソースタイププロパティ, 241
Fini メソッド、使用, 43,79

G

GDS

Agent Builder のコマンド行バージョンを使ってサービスを作成, 201
Child_mon_level プロパティ, 193
Failover_enabled プロパティ, 193
Log_level プロパティ, 193
Network_resources_used プロパティ, 191
Port_list プロパティ, 190
Probe_command プロパティ, 191
Probe_timeout プロパティ, 192
Start_command 拡張プロパティ, 190
Start_timeout プロパティ, 192
Stop_command プロパティ, 191
Stop_signal プロパティ, 193
Stop_timeout プロパティ, 192
Sun Cluster 管理コマンドで使用, 189
SunPlex Agent Builder で使用, 189
SunPlex Agent Builder を使って、GDS を使用するサービスを作成, 194
SUNW.gds リソースタイプ, 188
コマンドを使って、GDS を使用するサービスを作成, 199
使用する場合, 188

GDS (続き)

使用する利点, 188
使用方法, 188
説明, 187
定義, 40
必須のプロパティ, 190
Global_resources_used, リソースグループプロパティ, 258

H

halockrun, 説明, 45
hatimerun, 説明, 45
HA データサービス, 検証, 51

I

Implicit_network_dependencies, リソースグループプロパティ, 258
Init, リソースタイププロパティ, 241
Init_nodes, リソースタイププロパティ, 241
Init メソッド、使用, 43,79
Installed_nodes, リソースタイププロパティ, 241
Is_logical_hostname, リソースタイププロパティ, 241
Is_shared_address, リソースタイププロパティ, 242

J

Java, CRNP を使用するアプリケーションの例, 221

L

libdsdev.so, DSDL, 18
libscha.so, RMAPI, 18
Load_balancing_policy, リソースプロパティ, 248
Load_balancing_weights, リソースプロパティ, 248

M

max, リソースタイプの移行, 56
Max, リソースプロパティ属性, 264
Maximum primaries, リソースグループプロパティ, 258
Maxlength, リソースプロパティ属性, 264
method timeout, リソースプロパティ, 248
min, リソースタイプの移行, 56
Min, リソースプロパティ属性, 264
Minlength, リソースプロパティ属性, 264
Monitor_check, リソースタイププロパティ, 242
Monitor_check メソッド
 互換性, 58
 使用, 81
Monitor_start, リソースタイププロパティ, 242
Monitor_start メソッド, 使用, 81
Monitor_stop, リソースタイププロパティ, 242
Monitor_stop メソッド, 使用, 81
Monitored_switch, リソースプロパティ, 249

N

Network_resources_used, リソースプロパティ, 249
Nodelist, リソースグループプロパティ, 259
Num_resource_restarts, リソースプロパティ, 249
Num_rg_restarts, リソースプロパティ, 250

O

On_off_switch, リソースプロパティ, 250

P

Pathprefix, リソースグループプロパティ, 259
Pingpong_interval, リソースグループプロパティ, 259
Pkglist, リソースタイププロパティ, 242

PMF

関数、DSDL, 206
目的, 45
Port_list, リソースプロパティ, 250
Postnet_start メソッド, 使用, 80
Postnet_stop
 互換性, 58
 リソースタイププロパティ, 242
Prenet_start, リソースタイププロパティ, 243
Prenet_start メソッド, 使用, 80
Property, リソースプロパティ属性, 264

R

R_description, リソースプロパティ, 251
Resource_dependencies, リソースプロパティ, 251
Resource_dependencies_restart, リソースプロパティ, 251
Resource_dependencies_weak, リソースプロパティ, 252
Resource Group Manager, 「RGM」を参照
Resource_list, リソースグループプロパティ, 259
Resource Management API, 「RMAPI」を参照
Resource_name, リソースプロパティ, 252
Resource_project_name, リソースプロパティ, 253
Resource_state, リソースプロパティ, 253
Resource_type, 移行, 56
Resource_type, リソースタイププロパティ, 243
Retry_count, リソースプロパティ, 253
Retry_interval, リソースプロパティ, 254
RG_affinities, リソースグループプロパティ, 260
RG_dependencies, リソースグループプロパティ, 260
RG_description, リソースグループプロパティ, 261
RG_is_frozen, リソースグループプロパティ, 261
RG_mode, リソースグループプロパティ, 261
RG_name, リソースグループプロパティ, 262
RG_project_name, リソースグループプロパティ, 262

RG_state, リソースグループプロパティ, 262
RG_system, リソースグループプロパティ, 263

RGM

「リソースグループマネージャ」を参照

説明, 21

目的, 18

リソースグループの処理, 19

リソースタイプの処理, 19

リソースの処理, 19

RMAPI, 18

C プログラムの関数, 73

libscha.so, 18

クラスタ関数, 76

クラスタコマンド, 73

コールバックメソッド, 77

コンポーネント, 22

シェルコマンド, 72

実装場所, 18

終了コード, 78

メソッドの引数, 77

ユーティリティ関数, 76

リソース関数, 74

リソースグループ関数, 75

リソースグループコマンド, 73

リソースコマンド, 72

リソースタイプ関数, 75

リソースタイプコマンド, 73

RT_basedir, リソースタイププロパティ, 243

RT_description, リソースタイププロパティ, 244

RT_system, リソースタイププロパティ, 244

RT_version, 移行, 56

RT_Version

変更しないとき, 57

変更するとき, 57

目的, 57

RT_version, リソースタイププロパティ, 244

rtconfig ファイル, 180

RTR

説明, 21

ファイル

SUNW.xfnts, 137

移行, 56

説明, 124

変更, 68

S

Scalable, リソースプロパティ, 254

scds_initialize() 関数, 138

Single_instance, リソースタイププロパティ, 244

Start, リソースタイププロパティ, 244

Startメソッド、使用, 41, 78

Status, リソースプロパティ, 255

Status_msg, リソースプロパティ, 255

Stop, リソースタイププロパティ, 245

Stopメソッド

互換性, 58

使用, 41, 78

Sun Cluster

GDS による使用, 188

コマンド, 24

SunPlex Agent Builder

「Agent Builder」を参照

起動, 194

出力, 198

で GDS を作成, 189

のコマンド行バージョンを使って GDS ベースのサービスを作成, 201

を使って、GDS を使用するサービスを作成, 194

SunPlex Manager, 説明, 24

SUNW.xfnts

RTR ファイル, 137

障害モニター, 147

svc_probe() 関数, 149

T

TCP 接続, DSDL 障害モニターによる, 206

Thorough_probe_interval, リソースプロパティ, 255

Tunable, リソースプロパティ属性, 264

Tunable 属性のオプション, 56

ANYTIME, 58

AT_CREATION, 59

WHEN_DISABLED, 59

WHEN_OFFLINE, 58

WHEN_UNMANAGED, 59

WHEN_UNMONITORED, 58

Tunable 属性の制約, 文書の要件, 63

Type, リソースプロパティ, 255

Type_version, リソースプロパティ, 256

Type_version リソースプロパティ, 58
Tunable 属性, 58
編集, 58

U

UDP_affinity, リソースプロパティ, 256
Update, リソースタイププロパティ, 245
Update メソッド
互換性, 58
使用, 46, 80

V

Validate, リソースタイププロパティ, 245
Validate メソッド
アップグレード, 60
アップグレードのためにプロパティ値を検
査, 63
使用, 46, 80
Vendor_id
移行, 56
識別, 56
Vendor_ID, リソースタイププロパティ, 245

W

Weak_affinity, リソースプロパティ, 256
WHEN_DISABLED, #supgrade_from ディレク
ティブ, 59
WHEN_OFFLINE, #supgrade_from ディレク
ティブ, 58
WHEN_UNMANAGED, #supgrade_from ディレ
クティブ, 59
WHEN_UNMONITORED, #supgrade_from
ディレクティブ, 58

X

X Font Server
構成ファイル, 136
定義, 135
xfnts_monitor_check, 146
xfnts_monitor_start, 144

xfnts_monitor_stop, 145
xfnts_start, 138
xfnts_stop, 142
xfnts_update, 154
xfnts_validate, 152
xfs サーバー, ポート番号, 136

あ 値

デフォルトのプロパティ, 62
リソースグループマネージャ, 319
アップグレード
デフォルトのプロパティ値, 62
文書の要件, 63
リソースタイプの例, 64
アップグレード対応, 定義済み, 55

い

依存関係, リソース間の調節, 52
インストール要件, リソースタイプパッケ
ージ, 68
インタフェース, コマンド行, 24

え

エラー状況, CRNP, 216

お

オプション, Tunable 属性, 56

か

拡張, リソースプロパティ, 247
拡張プロパティ, 宣言, 37
画面
構成, 168
作成, 166
画面の構成, Agent Builder, 168
関数
DSDL, 203

関数 (続き)

- DSDL PMF (Process Monitor Facility), 206
- DSDL 障害モニター, 207
- DSDL ネットワークリソースアクセス, 205
- DSDL プロパティ, 205
- DSDL ユーティリティー, 207
- RMAPI C プログラム, 73
- RMAPI クラスタ, 76
- RMAPI ユーティリティー, 76
- RMAPI リソース, 74
- RMAPI リソースグループ, 75
- RMAPI リソースタイプ, 75
- scds_initialize(), 138
- svc_probe(), 149
- 汎用 DSDL, 203
- 完成した作業内容の再利用, Agent Builder, 173
- 完全修飾名, 取得方法, 57
- 完全修飾名の取得, 57
- 管理コマンド, を使って, GDS を使用するサービスを作成, 199

き

- キープアライブ, 使用, 51
- 記述値, 規則, 319
- 規則
 - 記述値, 319
 - プロパティ値, 319
 - プロパティ名, 317
 - リソースグループ名, 317
 - リソース名, 317
 - 列挙型リテラル名, 317
- 既存のリソースタイプのクローン作成, Agent Builder, 173

く

- クライアント, CRNP, 213
- クラスタ関数, RMAPI, 76
- クラスタ構成リポジトリ, 62
- クラスタコマンド, RMAPI, 73

け

- 検証
 - HA データサービス, 51
 - データサービス, 50
- 検証チェック, スケーラブルサービス, 50

こ

- 構成, Agent Builder, 160
- 構文
 - 記述値, 319
 - プロパティ値, 319
 - プロパティ名, 317
 - リソースグループ名, 317
 - リソースタイプ名, 318-319
 - リソース名, 317
 - 列挙型リテラル名, 317

コード

- RMAPI の終了, 78
- メソッドの変更, 69
- モニターの変更, 69

コールバックメソッド

- Monitor_check, 81
- Monitor_start, 81
- Monitor_stop, 81
- Postnet_start, 80
- Prenet_start, 80
- RMAPI, 77
- Update, 80
- Validate, 80
- 概要, 17
- 使用, 46
- 初期化, 78
- 制御, 78
- 説明, 21

コマンド

- halockrun, 45
- hatimerun, 45
- RMAPI リソースタイプ, 73
- Sun Cluster, 24
- で GDS を作成, 189
- を使って, GDS を使用するサービスを作成, 199

コマンド行

- Agent Builder, 175
- コマンド行上のコマンド, 24

コンポーネント, RMAPI, 22

さ

サーバー

CRNP, 213

X Font

構成ファイル, 136

定義, 135

xfns

ポート番号, 136

作成画面, Agent Builder, 166

サポートファイル, Agent Builder, 179

サンプル, データサービス, 83

サンプル DSDL

scds_initialize() 関数, 138

SUNW.xfnts 障害モニター, 147

SUNW.xfnts の RTR ファイル, 137

svc_probe() 関数, 149

svc_start() からの復帰, 140

TCP ポート番号, 136

X Font Server, 135

xfnts_monitor_check メソッド, 146

xfnts_monitor_start メソッド, 144

xfnts_monitor_stop メソッド, 145

xfnts_probe メインループ, 147

xfnts_start メソッド, 138

xfnts_stop メソッド, 142

xfnts_update メソッド, 154

xfnts_validate メソッド, 152

X Font Server の構成ファイル, 136

サービスの起動, 139

サービスの検証, 138

障害モニターのアクションの決定, 151

サンプルデータサービス

Monitor_check メソッド, 108

Monitor_start メソッド, 106

Monitor_stop メソッド, 107

RTR ファイル, 85

RTR ファイルの拡張プロパティ, 89

RTR ファイルのサンプルプロパティ, 86

Start メソッド, 94

Stop メソッド, 97

Update メソッド, 114

Validate メソッド, 109

エラーメッセージの生成, 93

共通の機能, 90

検証プログラム, 101

障害モニターの定義, 100

プロパティ更新の処理, 109

プロパティ情報の取得, 93

し

シェルコマンド, RMAPI, 72

実装

DSDL 障害モニター, 119

RMAPI, 18

リソースタイプ名, 63

リソースタイプモニター, 63

終了コード, RMAPI, 78

主ノード, 20

障害モニター

SUNW.xfnts, 147

関数, DSDL, 207

デーモン

設計, 131

す

スクリプト

Agent Builder, 178

構成, 196

作成, 194

スケーラブルサービス, 検証, 50

スケーラブルリソース, 実装, 47

せ

生成された Agent Builder ソースコードの編集, 174

そ

ソースコード, 生成された Agent Builder の編集, 173

ソースファイル, Agent Builder, 176

属性, リソースプロパティ, 263

た

タイプ, リソースプロパティ属性, 264

ち

チェック, スケーラブルサービスの検証, 50

て

ディレクティブ

- #\$upgrade, 318
 - #\$upgrade_from, 57, 58
 - RTR ファイル内の配置, 57
 - Tunable 属性の制約, 57
 - デフォルトの Tunable 属性, 57
- ディレクトリ, Agent Builder, 179
- ディレクトリ構造, Agent Builder, 175
- データサービス
- HA の検証, 51
 - 開発環境の設定, 28
 - 検証, 50
 - 検証のためにクラスタに転送する, 30
 - 作成, 50
 - インタフェースの決定, 27
 - 適合性の分析, 25
 - サンプル, 83
 - Monitor_check メソッド, 108
 - Monitor_start メソッド, 106
 - Monitor_stop メソッド, 107
 - RTR ファイル, 85
 - RTR ファイルの拡張プロパティ, 89
 - RTR ファイルのリソースプロパティ, 86
 - Start メソッド, 94
 - Stop メソッド, 97
 - Update メソッド, 114
 - Validate メソッド, 109
 - エラーメッセージの生成, 93
 - 共通の機能, 90
 - 検証プログラム, 101
 - 障害モニターの定義, 100
 - データサービスの制御, 94
 - プロパティ更新の処理, 109
 - プロパティ情報の取得, 93
- データサービス開発ライブラリ, 「DSDL」を参照
- データサービスサンプル, データサービスの制御, 94
- データサービスの作成, 50
- デーモン, 障害モニターの設定, 131
- デフォルトのプロパティ値
- Sun Cluster 3.0, 63
 - アップグレード, 62
 - アップグレード用の新しい値, 62
 - クラスタ構成リポジトリ, 62
 - 継承されるとき, 63

な

ナビゲーション, Agent Builder の, 162

ね

- ネットワークアドレスのアクセス, DSDL による, 120
- ネットワークリソースアクセス関数, DSDL, 205

は

- バイナリファイル, Agent Builder, 176
- パッケージディレクトリ, Agent Builder, 179
- 汎用データサービス
 - 「GDS」を参照

ひ

引数, RMAPI メソッド, 77

ふ

ファイル

- Agent Builder におけるサポート, 179
- Agent Builder のソース, 176
- Agent Builder のバイナリ, 176
- rtconfig, 180
- フェイルオーバーリソース, 実装, 46
- 複数の登録バージョンの識別, RT_version, 56
- 「ブラウズ」, Agent Builder, 163
- プログラミングアーキテクチャ, 18
- プロセス監視機能, 「PMF」を参照
- プロセス管理, 45
- プロセス管理機能, 概要, 18
- プロトコル, CRNP, 210
- プロパティ
 - Child_mon_level, 193
 - Failover_enabled, 193
 - GDS、必須, 193
 - Log_level, 193
 - Network_resources_used, 191
 - Port_list, 190
 - Probe_command, 191

プロパティ (続き)

- Probe_timeout, 192
- Start_command 拡張, 190
- Start_timeout, 192
- Stop_command, 191
- Stop_signal, 193
- Stop_timeout, 192
- 拡張プロパティの宣言, 37
- リソース, 245
- リソースグループ, 257
- リソースタイプ, 239
- リソースタイプの設定, 30
- リソースタイプの宣言, 31
- リソースの設定, 30, 46
- リソースの宣言, 34
- リソースの変更, 46
- プロパティ関数, DSDL, 18
- プロパティ属性, リソース, 263
- プロパティ値
 - 規則, 319
 - デフォルト, 62
- プロパティ変数, 171
 - そのタイプを Agent Builder がどのように置き換えるか, 173
 - の構文, 173
 - のリスト, 172
 - リソースグループのリスト, 172
 - リソースタイプのリスト, 172
 - リソースのリスト, 172
- プロパティ名, 規則, 317
- 文書の要件
 - Tunable 属性の制約, 63
 - アップグレードの, 63

へ

変数

- Agent Builder はプロパティのタイプをどのように置き換えるか, 173
 - プロパティ, 171
 - プロパティの構文, 173
 - プロパティのリスト, 172
 - リソースグループプロパティのリスト, 172
 - リソースタイププロパティのリスト, 172
 - リソースプロパティのリスト, 172
- ベンダーの識別, Vendor_id, 56

ま

- マスター, 説明, 20
- マニュアルページ, Agent Builder, 178

め

メソッド

- Boot, 43, 79, 130
- Fini, 43, 79, 130
- Init, 43, 79, 130
- Monitor_check, 81, 129
- Monitor_check コールバック, 81
- Monitor_start, 81, 128
- Monitor_start コールバック, 81
- Monitor_stop, 81, 129
- Monitor_stop コールバック, 81
- Postnet_start, 80
- Postnet_start コールバック, 80
- Prenet_start, 80
- Prenet_start コールバック, 80
- Start, 41, 78, 126
- Stop, 41, 78, 127
- Update, 46, 80, 129
- Update コールバック, 80
- Validate, 46, 80, 124
- Validate コールバック, 80
- xfnts_monitor_check, 146
- xfnts_monitor_start, 144
- xfnts_monitor_stop, 145
- xfnts_start, 138
- xfnts_stop, 142
- xfnts_update, 154
- xfnts_validate, 152
- コールバック, 46
 - 初期化, 78
 - 制御, 78
- 呼び出し回数への非依存性, 39
- メソッドコード, 変更, 69
- メソッドの引数, RMAPL, 77
- メッセージ, SC_CALLBACK_REG CRNP, 213
- メッセージログ, リソースへの追加, 45
- メニュー
 - Agent Builder, 165
 - Agent Builder の「ファイル」, 165
 - Agent Builder の「編集」, 165

も

モニターコード, 変更, 69

ゆ

有効な名前, リソースグループマネージャ, 317-319

ユーティリティ関数, DSDL, 207

ユーティリティ関数, RMAPI, 76

よ

呼び出し回数への非依存性, メソッド, 39

り

リソース

間の依存関係の調節, 52

監視, 44

起動, 40

スケラブルリソースの実装, 47

説明, 20

停止, 40

フェイルオーバーの実装, 46

別のバージョンへ移行, 59

メッセージログの追加, 45

リソース関数, RMAPI, 74

リソースグループ

スケラブル, 20

説明, 20

フェイルオーバー, 20

プロパティ, 20

リソースグループ関数, RMAPI, 75

リソースグループコマンド, RMAPI, 73

リソースグループプロパティ, 257

Auto_start_on_new_cluster, 257

Desired primaries, 257

Failback, 258

Global_resources_used, 258

Implicit_network_dependencies, 258

Maximum primaries, 258

Nodelist, 259

Pathprefix, 259

Pingpong_interval, 259

Resource_list, 259

リソースグループプロパティ (続き)

RG_affinities, 260

RG_dependencies, 260

RG_description, 261

RG_is_frozen, 261

RG_mode, 261

RG_name, 262

RG_project_name, 262

RG_state, 262

RG_system, 263

についての情報へのアクセス, 39

リソースグループマネージャ

値, 319

有効な名前, 317-319

リソースグループ名, 規則, 317

リソースコマンド, RMAPI, 72

リソースタイプ

DSDLによるデバッグ, 120

アップグレード, 60

移行の要件, 55

説明, 19

複数のバージョン, 55

リソースタイプ関数, RMAPI, 75

リソースタイプコマンド, RMAPI, 73

リソースタイプ登録, 「RTR」を参照

リソースタイプのアップグレード, 例, 64

リソースタイプの移行, 55

リソースタイプパッケージ, インストール要件, 68

リソースタイププロパティ, 239

API_version, 240

Boot, 240

Failover, 240

Fini, 241

Init, 241

Init_nodes, 241

Installed_nodes, 241

Is_logical_hostname, 241

Is_shared_address, 242

Monitor_check, 242

Monitor_start, 242

Monitor_stop, 242

Pkglist, 242

Postnet_stop, 242

Prenet_start, 243

Resource_type, 243

RT_basedir, 243

RT_description, 244

リソースタイププロパティ (続き)

- RT_system, 244
- RT_version, 244
- Single_instance, 244
- Start, 244
- Stop, 245
- Update, 245
- Validate, 245
- Vendor_ID, 245
- 設定, 30
- 宣言, 31

リソースタイプ名

- Sun Cluster 3.0, 58
- 規則, 318-319
- 実装, 63
- 制限, 56
- バージョン接尾辞, 56
- バージョン接尾辞なし, 58
- 例, 318, 319

リソースタイプモニター, 実装, 63

リソースの依存関係, 調節, 52

リソースプロパティ, 245

- Affinity_timeout, 246
- Cheap_probe_interval, 246
- Failover_mode, 247
- Load_balancing_policy, 248
- Load_balancing_weights, 248
- method_timeout, 248
- Monitored_switch, 249
- Network_resources_used, 249
- Num_resource_restarts, 249
- Num_rg_restarts, 250
- On_off_switch, 250
- Port_list, 250
- R_description, 251
- Resource_dependencies, 251
- Resource_dependencies_restart, 251
- Resource_dependencies_weak, 252
- Resource_name, 252
- Resource_project_name, 253
- Resource_state, 253
- Retry_count, 253
- Retry_interval, 254
- Scalable, 254
- Status, 255
- Status_msg, 255
- Thorough_probe_interval, 255
- Type, 255

リソースプロパティ (続き)

- Type_version, 256
- UDP_affinity, 256
- Weak_affinity, 256
- 拡張, 247
- 設定, 30, 46
- 宣言, 34
- についての情報へのアクセス, 39
- 変更, 46

リソースプロパティ属性, 263

- Array_maxsize, 264
- Array_minsize, 264
- Default, 264
- Description, 264
- Enumlist, 264
- Extension, 264
- Max, 264
- Maxlength, 264
- Min, 264
- Minlength, 264
- Property, 264
- Tunable, 264
- タイプ, 264

リソース名, 規則, 317

れ

例

- CRNP を使用する Java アプリケーション, 221

- リソースタイプのアップグレード, 64

- 列挙型リテラル名, 規則, 317

ろ

- ログ, リソースへの追加, 45