

# **Oracle® Tuxedo**

Oracle Tuxedo/Oracle Exalogic Users Guide

12c Release 1 (12.1.1)

June 2012

**ORACLE®**

Copyright © 1996, 2012, Oracle and/or its affiliates. All rights reserved.

This software and related documentation are provided under a license agreement containing restrictions on use and disclosure and are protected by intellectual property laws. Except as expressly permitted in your license agreement or allowed by law, you may not use, copy, reproduce, translate, broadcast, modify, license, transmit, distribute, exhibit, perform, publish, or display any part, in any form, or by any means. Reverse engineering, disassembly, or decompilation of this software, unless required by law for interoperability, is prohibited.

The information contained herein is subject to change without notice and is not warranted to be error-free. If you find any errors, please report them to us in writing.

If this is software or related documentation that is delivered to the U.S. Government or anyone licensing it on behalf of the U.S. Government, the following notice is applicable:

U.S. GOVERNMENT END USERS: Oracle programs, including any operating system, integrated software, any programs installed on the hardware, and/or documentation, delivered to U.S. Government end users are "commercial computer software" pursuant to the applicable Federal Acquisition Regulation and agency-specific supplemental regulations. As such, use, duplication, disclosure, modification, and adaptation of the programs, including any operating system, integrated software, any programs installed on the hardware, and/or documentation, shall be subject to license terms and license restrictions applicable to the programs. No other rights are granted to the U.S. Government.

This software or hardware is developed for general use in a variety of information management applications. It is not developed or intended for use in any inherently dangerous applications, including applications that may create a risk of personal injury. If you use this software or hardware in dangerous applications, then you shall be responsible to take all appropriate fail-safe, backup, redundancy, and other measures to ensure its safe use. Oracle Corporation and its affiliates disclaim any liability for any damages caused by use of this software or hardware in dangerous applications.

Oracle and Java are registered trademarks of Oracle and/or its affiliates. Other names may be trademarks of their respective owners.

This software or hardware and documentation may provide access to or information on content, products, and services from third parties. Oracle Corporation and its affiliates are not responsible for and expressly disclaim all warranties of any kind with respect to third-party content, products, and services. Oracle Corporation and its affiliates will not be responsible for any loss, costs, or damages incurred due to your access to or use of third-party content, products, or services.

# Oracle Tuxedo/Oracle Exalogic Users Guide

|  |     |
|--|-----|
| Overview . . . . .   | -1  |
| Scope . . . . .  | -2  |
| Prerequisites . . . . .  | -2  |
| Terminology . . . . .  | -2  |
| Feature Summary . . . . .                                      | -3  |
| Direct Cross Node Communication Leveraging RDMA . . . . .      | -3  |
| Self-Tuning Lock Mechanism . . . . .                           | -4  |
| Oracle Tuxedo SDP Support . . . . .                            | -4  |
| Use of Shared Memory for Inter Process Communication . . . . . | -4  |
| Read-Only Optimization for XA . . . . .                        | -4  |
| Shared Applications Staging . . . . .                          | -4  |
| Oracle Tuxedo installation . . . . .                           | -5  |
| Configuration for Exalogic . . . . .                           | -5  |
| Network . . . . .  | -5  |
| Storage . . . . .  | -6  |
| Choosing Oracle Tuxedo Home . . . . .                          | -6  |
| Start installation . . . . .                                   | -6  |
| Console-Mode Installation . . . . .                            | -6  |
| Oracle Tuxedo Configuration . . . . .                          | -8  |
| Direct Cross Node Communication Leveraging RDMA . . . . .      | -9  |
| UBBCONFIG File . . . . .                                       | -9  |
| Self-Tuning Lock Mechanism . . . . .                           | -12 |
| Oracle Tuxedo SDP Support . . . . .                            | -14 |
| MP . . . . .   | -14 |
| GWTDOMAIN . . . . .  | -16 |

|  |     |
|--|-----|
| WSL .....  | -18 |
| /WS client .....   | -19 |
| JSL .....  | -19 |
| WTC .....  | -20 |
| Use of Shared Memory for Inter Process Communication ..... | -20 |
| Read-only Optimization for XA .....                        | -21 |
| Database .....   | -21 |
| Oracle Tuxedo File Transfer Default Directory Name .....   | -22 |
| Choosing APPDIR .....                                      | -22 |
| limits.conf .....  | -24 |
| EXALOGIC_MSGQ_CACHE_SIZE .....                             | -25 |
| Running Oracle Tuxedo .....                                | -25 |
| Start/Stop tux_msgq_monitor .....                          | -25 |
| Assistant Tools .....                                      | -26 |
| Shell scripts for start/stop Oracle Tuxedo .....           | -26 |
| Tool for Computing the Shared Memory Size .....            | -27 |
| Running Oracle Tuxedo on OVM .....                         | -30 |
| SALT and TMA on Exalogic .....                             | -30 |
| Migration .....  | -31 |

# Oracle Tuxedo/Oracle Exalogic Users Guide

This chapter contains the following topics:

- [Overview](#)
- [Oracle Tuxedo installation](#)
- [Oracle Tuxedo Configuration](#)
- [Running Oracle Tuxedo](#)
- [Running Oracle Tuxedo on OVM](#)
- [SALT and TMA on Exalogic](#)
- [Migration](#)

## Overview

This section contains the following topics

- [Scope](#)
- [Prerequisites](#)
- [Terminology](#)
- [Feature Summary](#)

## Scope

This document will introduce all special features for Exalogic platform, and describe the instruction on how to configure and run Oracle Tuxedo with these features.

## Prerequisites

Before install Oracle Tuxedo, ensure set Oracle Exalogic machine properly that include storage and networking. For more information, see [Oracle Exalogic Machine Owner's Guide](#).

## Terminology

- **Oracle Home:** An Oracle Home directory serves as the central support directory for all Oracle products installed on the target system.
- **Compute Node:** A physical machine in an Exalogic rack that is meant for running an Oracle Fusion Middleware instance at any given time.
- **Shared Storage:** Shared storage refers to the Sun ZFS Storage 7320 appliance that is accessible by all compute nodes in the Oracle Exalogic Machine. All compute nodes in the Exalogic machine can access this storage appliance simultaneously for both read and write operations.
- **UBBCONFIG:** Each Oracle Tuxedo domain is controlled by a configuration file in which installation-dependent parameters are defined. The text version of the configuration file is referred to as UBBCONFIG. The UBBCONFIG file for a Oracle Tuxedo domain contains all the information necessary to boot the application.
- **TUXCONFIG:** The TUXCONFIG file is a binary version of the UBBCONFIG file. It is created by running the `tmlodcf(1)` command, which parses UBBCONFIG and loads the binary TUXCONFIG file to the location referenced by the `TUXCONFIG` environment variable.
- **XA:** The eXtended Architecture (XA) is the standard-defined interface between the Transaction Manager and the Resource Manager. XA allows program control of Resource Managers that are involved in distributed transactions
- **InfiniBand:** It is a switched fabric communications link used in high-performance computing and enterprise data centers. Its features include high throughput, low latency, quality of service and failover, and it is designed to be scalable.
- **IB:** InfiniBand.

- IBA: InfiniBand Architecture
- EMSQ: Exalogic Message Queue based on InfiniBand technology.
- Msgq\_daemon: A daemon process that manages EMSQ resources.
- IPoIB: IP over InfiniBand.
- EOIB: Ethernet over InfiniBand.
- SDP: Sockets Direct Protocol

## Feature Summary

[Table 1](#) lists the Exalogic supported Oracle Tuxedo features.

**Table 1 Exalogic Supported Oracle Tuxedo Features**

| Feature Name   | Oracle Tuxedo Version   |
|--|---|
| Direct Cross Node Communication Leveraging RDMA      | Oracle Tuxedo 11gR1 (11.1.1.3.0),<br>Oracle Tuxedo 12cR1 (12.1.1) |
| Self-tuning Lock Mechanism                           | Oracle Tuxedo 11gR1 (11.1.1.3.0),<br>Oracle Tuxedo 12cR1 (12.1.1) |
| Oracle Tuxedo SDP Support                            | Oracle Tuxedo 11gR1 (11.1.1.3.0),<br>Oracle Tuxedo 12cR1 (12.1.1) |
| Use of Shared Memory for Inter Process Communication | Oracle Tuxedo 12cR1 (12.1.1)                                      |
| Read-only Optimization for XA                        | Oracle Tuxedo 12cR1 (12.1.1)                                      |
| Shared Applications Staging                          | Oracle Tuxedo 12cR1 (12.1.1)                                      |

### Direct Cross Node Communication Leveraging RDMA

In previous releases, messages between local client and remote server must go through bridge. For example, first the message will be sent to local bridge through IPC queue, next the local bridge sends it to remote bridge through network, then the remote bridge sends the message to server's IPC queue, finally the server retrieves the message from its IPC queue, so the Bridge will become a bottleneck under high concurrency. In this release, if Direct Cross Node

Communication Leveraging RDMA is enabled in the TUXCONFIG file, the local client and remote server can skip Bridge, and transfer message directly.

## **Self-Tuning Lock Mechanism**

In previous Oracle Tuxedo Release, the tuxedo process will try to lock the bulletin board at user level for SPINCOUNT times before blocking on a UNIX semaphore. Typically it can get high performance if a proper SPINCOUNT is specified to make the process getting the lock at user level. Different load in an Oracle Tuxedo application may need different proper SPINCOUNT that the user must find it himself. In this release if Self-tuning Lock Mechanism is enabled, the Oracle Tuxedo will find the proper SPINCOUNT automatically under kinds of load at run time.

## **Oracle Tuxedo SDP Support**

This feature enables Oracle Tuxedo component those are using BSD Socket APIs can leverage the advantages of SDP network protocol provided by Exalogic (such as high bandwidth, low latency as well as reduced CPU involvement).

## **Use of Shared Memory for Inter Process Communication**

Oracle Tuxedo 12c significantly enhances performance of Tuxedo applications on Exalogic with use of shared memory queues instead of IPC Message Queues for inter process communication on the same Tuxedo node. With the use of shared memory queues, the sender and receiver processes can exchange pre-allocated messages in shared memory, thus eliminating the need to copy messages several times before message reaches its intended target and resulting in much better throughput and lower latency.

## **Read-Only Optimization for XA**

Optimized Distributed Transaction processing within and across Tuxedo domains for read-only transaction, including global transaction across Tuxedo domain and WTC( in WLS 12.1.1 - Contact Oracle Support for a patch, or higher release of WLS). One of the typical scenarios is every branch of the global transaction access the same Oracle Database instance.

## **Shared Applications Staging**

With Oracle Tuxedo 12c, one can share application directory (APPDIR) among many compute nodes of the storage appliance on an Exalogic system, making it easier to manage application deployment.



# Oracle Tuxedo installation

This section contains the following topics:

- [Configuration for Exalogic](#)
- [Choosing Oracle Tuxedo Home](#)
- [Start installation](#)

## Configuration for Exalogic

Before the Oracle Tuxedo installation, you should understand the current state of the Exalogic environment.

It is assumed that you have completed all tasks described in the Oracle Exalogic Machine Owner's Guide, which discusses your data center site preparation, Oracle Exalogic machine commissioning, initial networking configuration including IP address assignments, and initial setup of the Sun ZFS Storage 7320 appliance.

This section contains the following topics:

- Network
- Storage

## Network

Run the Oracle `OneCommand` tool to complete the following tasks (as described in "Initial Configuration of an Exalogic Machine Using Oracle OneCommand" in the [Oracle Exalogic Machine Owner's Guide](#)).

- Configuration of IP addresses for all Exalogic compute nodes and the Sun ZFS Storage 7320 appliance.
- Configuration of InfiniBand gateway switches.
- Configuration of the Cisco Ethernet management switch.
- Setup and verification of the default IP over InfiniBand (IPoIB) link spanning all compute nodes.
- Setup and verification of the default Ethernet over InfiniBand (EoIB) link for connectivity with components of the topology running on Ethernet.

- Configuration of the default InfiniBand partition that covers all of the compute nodes in Exalogic Machine.

## Storage

The initial configuration of the Sun ZFS Storage 7320 appliance on your Oracle Exalogic machine is completed at the time of manufacturing. For more information about default shares (Exported File Systems), see the “Default Storage Configuration” section in the [Oracle Exalogic Machine Owner's Guide](#).

After completing this initial configuration, you can proceed to create custom shares as needed.

## Choosing Oracle Tuxedo Home

We recommend that you can install the Oracle Tuxedo product binaries in one of the shares on Sun ZFS Storage 7320 appliance locations, so you can run Oracle Tuxedo on any Exalogic nodes by one binary copy.

**Notes:** The share, which is a shared file system, must be accessible by all compute nodes. You can create a local user account for each node, and ensure it has the same uid and gid (to avoid permission access issues), or create NIS accounts for users.

Oracle Tuxedo must be installed in a different directory if you want to develop the Oracle Tuxedo plug-in interface with different implementation for each Oracle Tuxedo.

## Start installation

The installation procedure is the same as previous Oracle Tuxedo releases. For more information, see [Installing the Oracle Tuxedo System](#).

## Console-Mode Installation

The following is a sample for Console-Mode installation on Exalogic, assume we will install Oracle Tuxedo on the node named slce04cn01, and the shared directory `"/home/oracle"` is available.

1. Log in to slce04cn01 as the Oracle Tuxedo administrator.

Go to the directory where you downloaded the installer and invoke the installation procedure by entering the following command:

```
prompt>sh./tuxedo111130_64_Linux_01_x86.bin-i console
```

The Choose Locale screen is displayed.

2. In the Choose Locale screen, enter 1, which is associated with English.  
The Introduction screen is displayed.
3. In the Introduction screen, press <ENTER> to continue.  
The Choose Install Set screen is displayed.
4. In the Choose Install Set screen, enter 1, which is associated with Full Install.  
The Choose Oracle Home screen is displayed.
5. In the Choose Oracle Home screen, enter 1, which is associated with Create new Oracle Home.  
The Specify a new Oracle Home directory screen is prompted.
6. Enter your Oracle Home directory. Be sure to use the full pathname when specifying the new Oracle Home directory.  
  
The Oracle Home should be on the shared file system on the Sun Storage 7000 Unified Storage System, and can be accessible by all compute nodes in the Oracle Exalogic machine.  
  
For example, enter `/home/oracle` as the Tuxedo home directory.  
The Choose Product Directory screen is displayed.
7. In the Choose Product Directory screen, enter 2, which is associated with Use Current Selection.  
The Install Samples (Y/N) is prompted.
8. Enter Y to install the samples.  
The Pre- Installation Summary screen is displayed.
9. In the Pre- Installation Summary screen, press <ENTER> to continue.  
The Ready To Install screen is displayed.
10. In the Ready To Install screen, press <ENTER> to install.  
The Installing screen is displayed.
11. In the Installing screen, no user input is required.  
When it finishes, the Configure tlisten Service screen is displayed.

12. In the Configure tlisten Service screen, enter a tlisten password of your choice. Your password must be a string of alphanumeric characters in clear-text format that is no more than 80 characters in length. Then Verify your password.

**Note:** A common password is required for all the nodes in an Oracle Tuxedo application (domain) to communicate successfully. For this reason, you must use the same password when you install Oracle Tuxedo on more than one compute nodes for a single application.

The SSL Installation Choice screen is displayed.

13. In the SSL Installation Choice screen, you can choose to enter 1, which is associated with YES (This is not mandatory for the installation).

The Enter Your LDAP Settings for SSL Support screen is displayed.

14. In the Enter Your LDAP Settings for SSL Support screen, input your LDAP Service Name, LDAP PortID, LDAP BaseObject and LDAP Filter File Location (this is not mandatory for the installation).

The Installation Complete screen is displayed.

15. In the Installation Complete screen, press <ENTER> to exit the installer.

## Oracle Tuxedo Configuration

This section introduces the basic Oracle Tuxedo feature configuration on Exalogic. For more information, see the [Oracle Tuxedo 12c Release 1 \(12.1.1\) Release Notes](#) and [Setting Up an Oracle Tuxedo Application](#).

- [Direct Cross Node Communication Leveraging RDMA](#)
- [Self-Tuning Lock Mechanism](#)
- [Oracle Tuxedo SDP Support](#)
- [Use of Shared Memory for Intra-node Communication](#)
- [Read-only Optimization for XA](#)
- [Database](#)
- [Oracle Tuxedo File Transfer Default Directory Name](#)
- [Choosing APPDIR](#)

- [limits.conf](#)
- [EXALOGIC\\_MSGQ\\_CACHE\\_SIZE](#)

## Direct Cross Node Communication Leveraging RDMA

Using Direct Cross Node Communication Leveraging RDMA requires UBBCONFIG file configuration.

### UBBCONFIG File

Direct Cross Node Communication Leveraging RDMA is only supported under MP mode. To enable this feature, you must specify both `EXALOGIC` and `RDMA` in `OPTIONS`, otherwise the message goes through the Bridge.

There is one attribute for Direct Cross Node Communication Leveraging RDMA in the `*RESOURCE` section.

- `EXALOGIC_SHARED_PATH`

The directory name of Oracle Tuxedo file transfer. The function of `EXALOGIC_SHARED_PATH` here is the same as that of environment variable `EXALOGIC_SHARED_PATH`; however, at Tuxedo runtime, such environment variable has higher priority. `EXALOGIC_SHARED_PATH` must be a shared directory with read/write permissions for all Tuxedo nodes and can be specified in `*RESOURCE` section only if RDMA is enabled.

There are five attributes for Direct Cross Node Communication Leveraging RDMA in the `*MACHINES` section.

- `RDMAEAEMONIP`

The IP address where the `Msgq_daemon` is bound. It must be configured, and must be an IPoIB address (not an Ethernet based IP address). You should configure one `Msgq_daemon` for one logic machine.

- `RDMAEAEMONPORT`

The port number where `Msgq_daemon` listens on. It must be configured.

- `RDMAQSIZE`

The EMSQ queue size. The default value is 65536 bytes if not defined in the UBBCONFIG file.

- `RDMAQENTRIES`

The EMSQ queue entry number, that is the maximum number of messages allowed in this queue. The default value is 64 if not defined in the UBBCONFIG file.

- EXALOGIC\_MSGQ\_CACHE\_SIZE

The entry number for Oracle Tuxedo EMSQ cache. The function of EXALOGIC\_MSGQ\_CACHE\_SIZE here is the same as that of the environment variable EXALOGIC\_MSGQ\_CACHE\_SIZE; however, the environment variable has the higher priority. The value must be between 32 and 2048 inclusive. EXALOGIC\_MSGQ\_CACHE\_SIZE can be specified in \*MACHINE only when RDMA is enabled.

After enable RDMA option in the \*RESOURCES section. Attribute “TYPE” of \*MACHINES section cannot be set, since by default, any machines in MP mode should be Exalogic machine (with the same type) to support RDMA feature.

You can also get/change the configuration via TM\_MIB. For more information, see [File Formats, Data Descriptions, MIBs, and System Processes Reference](#).

[Listing 1](#) shows a example UBBCONFIG File with Direct Cross Node Communication Leveraging RDMA enabled.

### Listing 1 UBBCONFIG File Example with Direct Cross Node Communication Leveraging RDMA Enabled

---

```
*RESOURCES
IPCKEY          87654

MASTER         site1,site2
MAXACCESSERS   40
MAXSERVERS     40
MAXSERVICES    40
MODEL          MP
OPTIONS        LAN,EXALOGIC, RDMA
LDBAL          Y

*MACHINES
slce04cn01     LMID=site1
```

```

APPDIR="/home/oracle/tuxedo12.1.1.0/samples/atmi/simpapp"
TUXCONFIG="/home/oracle/tuxedo12.1.1.0/samples/atmi/simpapp/
tuxconfig"
TUXDIR="/home/oracle/tuxedo12.1.1.0"
UID=601
GID=601
RDMADAEMONIP="192.168.10.1"
RDMADAEMONPORT=9800
RDMAQSIZE=65536
RDMAQENTRIES=64

```

```

slce04cn02    LMID=site2

```

```

APPDIR="/home/oracle/tuxedo12.1.1.0/samples/atmi/simpapp/slave"
TUXCONFIG="/home/oracle/tuxedo12.1.1.0/samples/atmi/simpapp/
slave/tuxconfig"
TUXDIR="/home/oracle/tuxedo12.1.1.0"
UID=601
GID=601
RDMADAEMONIP="192.168.10.2"
RDMADAEMONPORT=9800
RDMAQSIZE=65536
RDMAQENTRIES=64

```

```

*GROUPS

```

```

GROUP1

```

```

LMID=site1    GRPNO=1    OPENINFO=NONE

```

```

GROUP2

```

```
LMID=site2      GRPNO=2      OPENINFO=NONE

*NETWORK

site1  NADDR="//slce04cn01:5432"
       NLSADDR="//slce04cn01:5442"

site1  NADDR="//slce04cn02:5432"
       NLSADDR="//slce04cn02:5442"

*SERVERS

DEFAULT:

        CLOPT="-A"

simpserv      SRVGRP=GROUP2  SRVID=3

*SERVICES

TOUPPER
```

---

## Self-Tuning Lock Mechanism

To enable Self-tuning Lock Mechanism, the options `EXALOGIC` and `SPINTUNING` must be specified in the `UBBCONFIG` file `*RESOURCE` section. Whenever the feature enabled, `SPINCOUNT` in `*MACHINE` cannot be set.

Two other optional attributes are added in `*MACHINES` section:

- `SPINTUNING_FACTOR`: Controls the tuning target. More CPUs may be used with a larger value under a heavy load; higher performance may be gained.
- `SPINTUNING_MINIDLECPU`: Specifies the CPU idle time.

**Note:** If not specified, the default values for these attributes are used.



For more information, see `UBBCONFIG(5)` and `UBBCONFIG(5) Additional Information, Example 2 Self-Tuning Lock Mechanism Configuration`, in [File Formats, Data Descriptions, MIBs, and System Processes Reference](#).

You can also set the configuration via `TM_MIB`. For more information, see `TM_MIB(5)` in [File Formats, Data Descriptions, MIBs, and System Processes Reference](#).

[Listing 2](#) shows a `UBBCONFIG` file example with Self-Tuning Lock Mechanism enabled.

## Listing 2 UBBCONFIG File Example with Self-Tuning Lock Mechanism Enabled

---

```
*RESOURCES

IPCKEY                123456
DOMAINID              simpapp
MASTER                ALLEN
MAXACCESSERS 10
MAXSERVERS            5
MAXSERVICES           10
MODEL                 SHM
LDBAL                 N
OPTIONS               EXALOGIC, SPINTUNING

*MACHINES

ALLENHOST             LMID="ALLEN"
APPDIR="/home/allen/Workspace/Tuxedo10/simpdir"
TUXCONFIG="/home/allen/Workspace/Tuxedo10/simpdir/tuxconfig"
TUXDIR="/home/allen/Software/OraHome/tuxedo11gR1"
SPINTUNING_FACTOR=1000
SPINTUNING_MINIDLECPU=20

*GROUPS

GROUP1
```

```
LMID=ALLEN GRPNO=1 OPENINFO=NONE

*SERVERS
DEFAULT:
CLOPT="-A"
simplserv  SRVGRP=GROUP1  SRVID=1

*SERVICES
TOUPPER
```

---

## Oracle Tuxedo SDP Support

To enable Oracle Tuxedo SDP Support, you must specify `EXALOGIC` for `OPTIONS` in `* RESOURCE` section, and set the relevant configuration in `UBBCONFIG` file or `DMCONFIG` file.

You can also get/change the configuration via `TM_MIB`. For more information, see [File Formats, Data Descriptions, MIBs, and System Processes Reference](#).

This section covers the following configurations:

- [MP](#)
- [GWTDOMAIN](#)
- [WSL](#)
- [/WS client](#)
- [JSL](#)
- [WTC](#)

### MP

According to the requirements, `MP` should work inside `IB` clusters, i.e., both master and slave machines are inside `IB` cluster, so only consider `SDP` and `IPoIB` are used inside `IB` cluster, and in bootstrap phase, `tmboot`, `tlisten`, `bsbridge` and `bridge` are using `Socket API` to communicate with each other.

- Configure SDP in MP deployment model.

Prefixed “sdp:” to the network address, and the network address must be an IPoIB address.

[Listing 3](#) shows a UBBCONFIG file SDP example.

### Listing 3 UBBCONFIG File SDP Configuration Example

---

```
*NETWORK

SITE1
    NADDR="sdp://IB_IP: 9003"
    NLSADDR="sdp://IB_IP: 3050"

SITE2
    NADDR="sdp://IB_IP: 9003"
    NLSADDR="sdp://IB_IP: 3050"
```

To start tlisten:

```
tlisten -d /dev/tcp -l sdp://IB_IP: 3050
```

---

- Configure IPoIB in MP deployment model

Using the IPoIB address as the network address.

[Listing 4](#) shows a UBBCONFIG file IPoIB configuration example.

### Listing 4 UBBCONFIG File IPoIB Configuration Example

---

```
*NETWORK

SITE1
    NADDR="//IB_IP: 9003"
    NLSADDR="//IB_IP: 3050"

SITE2
    NADDR="//IB_IP: 9003"
    NLSADDR="//IB_IP: 3050"
```

To start tlisten:

```
tlisten -d /dev/tcp -l //IB_IP: 3050
```

## GWTDOMAIN

If the node running GWTDOMAIN has multiple network interfaces (multi-homed) with multiple IP addresses, it's better to use explicit IP address when configuring GWTDOMAIN in DMCONFIG file instead of host name. Typically, every Exalogic node has at least two types of network interface, i.e., IB interface and Ethernet interface, in order to facilitate to demonstrate how to configure GWTDOMAIN, just presume IB interface is bound to IP address IB\_IP, and Ethernet interface with IP address ETH\_IP.

Functionally, GWTDOMAIN acts as both server and client in role, as server, it will listen on a configured IP address and port number in the DMCONFIG file to accept connection request from other GWTDOMAIN, as client, it will initiate connection request to other GWTDOMAIN by policy configured in the DMCONFIG file.

- Configure GWTDOMAIN to listen on SDP.

Prefixed “sdp:” to the network address, and the network address must be an IPoIB address, as shown in [Listing 5](#).

---

### Listing 5 GWTDOMAIN Listening on SDPD DMCONFIG File Configuration Example

---

```
*DM_LOCAL
SCLCU03          GWGRP=DOMGRP
                  TYPE=TDOMAIN
*DM_TDOMAIN
                  SCLCU03 NWADDR="sdp://IB_IP: 27610"
```

---

- Configure GWTDOMAIN to connect using SDP

Prefixed “sdp:” to the network address, and the network address must be an IPoIB address as shown in [Listing 6](#).

**Listing 6 GWTDOMAIN Connection Using SDP DMCONFIG File Configuration Example**

---

```

*DM_LOCAL

    SCLCU03      GWGRP=DOMGRP

                  TYPE=TDOMAIN

*DM_REMOTE

SCLCU05 TYPE=TDOMAIN

                  DOMAINID= "EXALOGIC_SCLCU05"

*DM_TDOMAIN

                  SCLCU05 NWADDR="sdp://IB_IP: 27610"

```

---

- Configure GWTDOMAIN to listen on IPoIB.

Using the IPoIB address as the network address, the following is the DMCONFIG file configuration example for IPoIB:

**Listing 7 GWTDOMAIN Listen on IPoIB DMCONFIG File Configuration Example**

---

```

*DM_LOCAL

SCLCU03          GWGRP=DOMGRP

                  TYPE=TDOMAIN

*DM_TDOMAIN

SCLCU03 NWADDR="//IB_IP: 27610"

```

---

- Configure GWTDOMAIN to connect using IPoIB.

Using the IPoIB address as the network address, the following is the DMCONFIG file configuration example for IPoIB:

### Listing 8 GWTDOMAIN Connection Using IPoIB DMCONFIG File Configuration Example

---

```
*DM_LOCAL
SCLCU03          GWGRP=DOMGRP
                  TYPE=TDOMAIN

*DM_REMOTE
SCLCU05 TYPE=TDOMAIN
                  DOMAINID= "EXALOGIC_SCLCU05"

*DM_TDOMAIN
SCLCU05 NWADDR="//IB_IP: 27610"
```

---

## WSL

- Configure WSL listen on SDP

Prefixed “sdp:” to the network address, and the network address must be an IPoIB address as shown in [Listing 9](#).

### Listing 9 WSL Listening on SDP DMCONFIG File Configuration Example

---

```
*SERVERS
DEFAULT:          CLOPT="-A"
WSL               SRVGRP=WSGRP SRVID=1001
                  CLOPT="-A -- -nsdp://IB_IP: 11101 -m1 -M10 -x1"
```

---

- To configure WSL listen on IPoIB

Using the IPoIB address as the network address, as shown in [Listing 10](#).

**Listing 10 WSL Listening on IpoIB DMCONFIG File Configuration Example**

---

```

*SERVERS

DEFAULT:      CLOPT="-A"
WSL           SRVGRP=WSGRP SRVID=1001

              CLOPT="-A -- -n //IB_IP: 11101 -m1 -M10 -x1"

```

---

**/WS client**

- Configure /WS client to use SDP

Prefixed “sdp:” to network address for WSNADDR, and the network address must be an IPoIB address, for example:

```
export WSNADDR=sdp://IB_IP:1001
```

- Configure /WS client to use IPoIB

Using the IPoIB address as the network address for WSNADDR, for example:

```
export WSNADDR://IB_IP:1001
```

**JSL**

- Configure JSL listen on SDP

Prefixed “sdp:” to the network address, and the network address must be an IPoIB address as shown in [Listing 11](#).

**Listing 11 JSL Listening on SDP UBBCONFIG File Configuration Example**

---

```

*SERVERS

DEFAULT:      CLOPT="-A"
JSL           SRVGRP=WSGRP SRVID=1001

              CLOPT="-A -- -nsdp: //IB_IP: 11101 -m1 -M10 -x1"

```

---

- To configure JSL listen on IPoIB

Using the IPoIB address as the network address, the following is the UBBCONFIG file configuration example for JSL with IPoIB enabled:

### Listing 12 JSL Listening on IpoIB UBBCONFIG File Configuration Example

---

```

*SERVERS

DEFAULT:          CLOPT="-A "
JSL               SRVGRP=WSGRP SRVID=1001
                  CLOPT="-A -- -n //IB_IP: 11101 -m1 -M10 -x1"
    
```

---

## WTC

To enable SDP connection between WTC and Oracle Tuxedo, do the following steps:

1. Specify the NWAddr of the WTC service Local/Remote Access Points as follows:

```
sdp://IB_IP:port
```

It is the same as the GWTDOMAIN NWADDR configuration in the DMCONFIG file.

2. Add additional Java Option “-Djava.net.preferIPv4Stack=true” to the java command line to start up WLS server.

**Notes:** If the WTC access point has SSL enabled, after configuring for the SDP, the SSL configuration is ignored.

Only Weblogic Server 12c (12.1.1) and higher can connect to Oracle Tuxedo via SDP. For more information, see [Enable IPv4 for SDP transport](#), [NWAddr attribute for WTC local Tuxedo Domain configuration](#), and [NWAddr attribute for WTC remote Tuxedo Domain configuration](#).

## Use of Shared Memory for Inter Process Communication

To enable this feature, the options EXALOGIC and SHMQ must be specified in the \*RESOURCE section of UBBCONFIG file.

Another optional attribute is provided in \*RESOURCE section:

- SHMQMAXMEM *numeric\_value*



which specifies the maximum shared memory size used for message buffers. To use **SHMQMAXMEM**, the option **SHMQ** must be specified.

The range of *numeric\_value* is from 1 to 96000 inclusive. If **SHMQ** is specified while either **SHMQMAXMEM** is not configured or its value is too small, a recommended minimum value will be used.

Run `tmloadcf -c` to get recommended minimum value. For more information, refer to [tmloadcf\(1\)](#).

## Read-only Optimization for XA

There is a new option in **\*RESOURCES** section of **UBBCONFIG** for this Read-only Optimization for XA feature.

```
OPTIONS { [...|RDONLY1PC], * }
```

It indicates that the global transaction owned by the local TDomain can support Read-only Optimization for XA. This option can be specified only if the option **EXALOGIC** is already specified; besides, if **NO\_XA** is specified, **RDONLY1PC** cannot be specified at the same time.

### Listing 13 Configuration Example

---

```
*RESOURCES
    OPTIONS LAN, EXALOGIC, RDONLY1PC
```

---

You can also get/change the configuration via **TM\_MIB**. For more information, see [File Formats, Data Descriptions, MIBs, and System Processes Reference](#).

## Database

The Oracle Tuxedo system uses the X/Open XA interface for communicating with the various resource managers. The XA Standard is widely supported in all the major database vendor products.

You can use SDP (Sockets Direct Protocol) for Oracle Database invocations. There is no special requirement for Oracle Tuxedo application.

Please configure the database to support InfiniBand, as described in [Configuring SDP Protocol Support for Infiniband Network Communication to the Database Server](#) in the Oracle Database Net Services Administrators Guide.

**Note:** The following SDP parameters affect performance when sending large data blocks to the database.

- `recv_poll`

The receiving poll time before going to sleep waiting for data. You can improve the write performance by decreasing the `recv_poll` value. It is recommended to set it to zero.

- `sdp_zcopy_thresh`

The threshold for zero copy sending data. If the messages are longer than the `sdp_zcopy_thresh` value, it causes the user buffer to be pinned and the data sent directly from the original buffer. This results in decreased CPU use. The default value is 64K; it is recommended to set it to zero.

For example, you can set the two parameters in “`/etc/modprobe.conf`” on the server node as follows:

```
options ib_sdp sdp_zcopy_thresh=0 recv_poll=0
```

## Oracle Tuxedo File Transfer Default Directory Name

Before starting Oracle Tuxedo, ensure that there is a shared directory for all Exalogic nodes when Direct Cross Node Communication Leveraging RDMA is enabled. Make sure that access permissions are properly set.

The default name is `/u01/common/patches/tuxttmpfile`, you can also set your own directory using the `EXALOGIC_SHARED_PATH` environment variable. It is used for Oracle Tuxedo file transfer. When the EMSQ is full, or the message size exceeds the queue size, Oracle Tuxedo puts this message into a temporary file under the `/u01/common/patches/tuxttmpfile` directory, and sends notification directly to the remote process queue. The remote process can then get the file as long as it receives the notification.

## Choosing APPDIR

You can deploy your Oracle Tuxedo application to a shared directory on Exalogic in MP environment (named Shared Applications Staging) with the requirements that both EXALOGIC

option and MP mode are set and enabled. Before booting the Oracle Tuxedo application, ensure the following parameters are set correctly in the UBBCONFIG file:

- TUXCONFIG

The TUXCONFIG must be different for each node.

- TLOGDEVICE

The TLOGDEVICE must be different for each node.

- ULOGPFX

Set different path for ULOGPFX if you want to have a separate ULOG.

- Access Permission for shared APPDIR

Users from different Exalogic nodes must have the same *uid* and *gid* of OS.

Besides above, each node had better use distinctive `TMIFRSVR repository_file`, standard output/error file, `AUDITLOG` file, and `ALOGPFX` to have a clear logging system. All applications should be set distinctive names to use the Shared Applications Staging feature better.

[Listing 14](#) shows a UBBCONFIG file shared APPDIR example.

#### Listing 14 UBBCONFIG File Shared APPDIR

---

```
...
*MACHINES
slce04cn01 LMID=site1

    APPDIR="/home/oracle/tuxapp"
    TUXCONFIG="/home/oracle/tuxapp/tuxconfig_cn01"
    TUXDIR="/home/oracle/tuxedo11gR1"

    TLOGDEVICE="/home/oracle/tuxapp/TLOG1"
ULOGPFX="/ home/oracle/tuxapp /ULOG_cn01"
    RDMADAEMONIP="192.168.10.1"
    RDMADAEMONPORT=9800
    RDMAQSIZE=1048576
    RDMAQENTRIES=1024

slce04cn02 LMID=site2
```

```
APPDIR=" home/oracle/tuxapp"  
TUXCONFIG=" home/oracle/tuxapp/tuxconfig_cn02"  
TUXDIR="/home/oracle/tuxedollgR1"  
  
TLOGDEVICE="/home/oracle/tuxapp/TLOG2"  
ULOGPFX="/home/oracle/tuxapp /ULOG_cn02"  
RDMADAEMONIP="192.168.10.2"  
RDMADAEMONPORT=9800  
RDMAQSIZE=1048576  
RDMAQENTRIES=1024
```

...

---

If the SECUTIRY is set in UBBCONFIG file while neither EXALOGIC option nor MP mode is set, the Shared APPDIR cannot be used. In that case, you must use a different APPDIR and have a copy for each node.

## limits.conf

Set the memlock in `/etc/security/limits.conf` properly. Use the following formula to get the minimum value:

$[\text{Msgq\_daemon shared memory size}] * 2 + \text{MAXACCESSERS} * 14\ 000\ \text{kb}$

`Msgq_daemon` shared memory size: The size of shared memory allocated by `Msgq_daemon`. For more information, see [Start/Stop tux\\_msgq\\_monitor](#)

`MAXACCESSERS`: An attribute in the UBBCONFIG file.

For example:

`Msgq_daemon` shared memory size:  $200 * 1024 * 1024$  bytes

`MAXACCESSERS`: 100

The minimum memlock value is 1853030400 bytes.

In `/etc/security/limits.conf`, set memlock las follows:

- \* hard memlock 1853030
- \* soft memlock 1853030

## EXALOGIC\_MSGQ\_CACHE\_SIZE

Each Oracle Tuxedo thread has an EMSQ runtime cache; the default entry number is 32. You can change it between 32 and 2048 using the environment `EXALOGIC_MSGQ_CACHE_SIZE` variable before the Oracle Tuxedo application starts. Increasing the number can improve Oracle Tuxedo performance, but `Msgq_daemon` consumes more shared memory.

Performance can be improved using this environment variable when one process sends messages to many queues. Note the following examples:

- If there are 40 remote Oracle Tuxedo servers providing the same service and clients call the service 100 times, setting `EXALOGIC_MSGQ_CACHE_SIZE` to a value equal to or greater than 40 on the client improves performance.
- If there are 50 WSHs, and each WSH receives response messages from the same remote server, setting `EXALOGIC_MSGQ_CACHE_SIZE` to a value equal to or greater than 50 on the server environment improves performance.

## Running Oracle Tuxedo

There is a difference running Oracle Tuxedo on a non-Exalogic platform if Direct Cross Node Communication Leveraging RDMA is enabled. The `tux_msgq_monitor` must be started before booting an Oracle Tuxedo application. This section includes the following topics:

- [Start/Stop `tux\_msgq\_monitor`](#)
- [Assistant Tools](#)
- [Oracle Tuxedo File Transfer Default Directory Name](#)

### Start/Stop `tux_msgq_monitor`

- Start `tux_msgq_monitor`

`tux_msgq_monitor` functions as the `Msgq_daemon` default monitor. It is responsible for starting `Msgq_daemon` and checking its running status. When `Msgq_daemon` terminates abnormally, it restarts using `tux_msgq_monitor`.

`tux_msgq_monitor` should be started before booting the Oracle Tuxedo application. One monitor can only serve one Oracle Tuxedo application on one logical machine.

Before starting `tux_msgq_monitor`, ensure the environment variables `TUXCONFIG`, `LD_LIBRARY_PATH` and `TUXDIR` are properly set. Run `tux_msgq_monitor` as follows:

```
tux_msgq_monitor -i <IPoIB address> -d <port number> -M <shared memory size> -K <shared memory key to be used by Msgq_daemon>.
```

[Table 2](#) lists `tux_msgq_monitor` options.

**Table 2 tux\_msgq\_monitor: Options**

| Options | Description  |
|---------|--|
| -i      | IPoIB address, it must be identical to the value of RDMADAEMONIP in UBBCONFIG file.  |
| -d      | Port number, it must be identical to the value of RDMADAEMONPORT in UBBCONFIG file.  |
| -M      | Shared memory size allocated by <code>Msgq_daemon</code> , you can get it by “ <code>tmloadcf -c</code> ”. For more information, see <a href="#">Tool for Computing the Shared Memory Size</a> . |
| -K      | Shared memory key used by <code>Msgq_daemon</code> .   |

When `tux_msgq_monitor` is running, there will be two active processes named as `tux_msgq_monitor` and `Msgq_daemon`.

After starting `tux_msgq_monitor` successfully, you can boot the Oracle Tuxedo application.

- **Shutdown `tux_msgq_monitor`**

Execute the following command:

```
kill pid -- tux_msgq_monitor's pid
```

**Note:** If you execute the `kill` command using `-9`, do the following steps to clear all the resources.

```
kill pid-- Msgq_daemon pid
```

```
ipcrm -m shmid -- the shared memory created by Msgq_daemon.
```

## Assistant Tools

### Shell scripts for start/stop Oracle Tuxedo

There are some shell scripts that simplify the startup/shutdown procedure. Using these tools, you can only run one command to start/stop both `tux_msgq_monitor` and an Oracle Tuxedo

application. Before running these commands, ensure the environment variables `TUXCONFIG`, `LD_LIBRARY_PATH` and `APPDIR` are set properly.

For example, on the master node you can start/stop Oracle Tuxedo as follows:

- `tmboot.sh -i daemon_ip -d daemon_port -M shm_size -K shm_key [-l nlsaddr]`

This script starts up `tux_msgq_monitor`, executes `tmboot` to start the Oracle Tuxedo application, and starts `tlisten` if option "-l" specified.

- `tmshut.sh`

Stops both the Oracle Tuxedo application and `tux_msgq_monitor`.

On slave node, there are two shell scripts:

- `tlisten_start.sh -l nlsaddr -i daemon_ip -d daemon_port -M shm_size -K shm_key`

This script starts the `tux_msgq_monitor` and `tlisten`.

- `tlisten_stop.sh`:

This script terminates `tlisten` and `tux_msgq_monitor`:

**Note:** In MP mode, for startup, you should run the commands in the following sequence:

1. `tmlisten_start.sh` on all the slave nodes.
2. `tmboot.sh` on the master node

For shutdown, you should run commands in the following sequence:

1. `tmshut.sh` on the master node
2. `tmlisten_stop.sh` on all the slave nodes.

## Tool for Computing the Shared Memory Size

When `Msgq_daemon` starts, it allocates a shared memory. You can get the memory size by using `tmloadcf -c ubb`, as shown in [Listing 15](#).

### Listing 15 UBBCONFIG File \*MACHINES Section

---

```
*MACHINES
ex03  LMID=sitel
...
RDMADAEMONIP="192.168.10.1"
RDMADAEMONPORT=9800
```

```
RDMAQSIZE=100000
RDMAQENTRIES=100
MAXACCESSERS=100
...
ex03_1 LMID=site2
...
RDMAEAEMONIP="192.168.10.2"
RDMAEAEMONPORT=9800
RDMAQENTRIES=1000
MAXACCESSERS=200
...
ex04 LMID=site3
...
RDMAEAEMONIP="192.168.10.3"
RDMAEAEMONPORT=9800
RDMAQSIZE=100000
RDMAQENTRIES=100
MAXACCESSERS=200
MAXSERVERS=100
...
ex04_1 LMID=site4
...
RDMAEAEMONIP="192.168.10.4"
RDMAEAEMONPORT=9800
RDMAQSIZE=1000000
RDMAQENTRIES=1000
MAXACCESSERS=100
...
```

---

Run command `tmloadcf -c ubb`, get the output shown in [Listing 16](#)

---

#### Listing 16 `tmloadcf -c ubb` Output Example

---

```
Ipc sizing (minimum /T values only) ...
Fixed Minimums Per Node
```



```
SHMMIN: 1
SHMALL: 1
SEMMAP: SEMMNI
```

## Variable Minimums Per Node

| SEMUME, | A      | SHMMAX  |        |        |        |        |        |        |
|---------|--------|---------|--------|--------|--------|--------|--------|--------|
|         |        | SEMMNU, | *      |        |        |        |        | *      |
| Node    | SEMMNS | SEMMSL  | SEMMSL | SEMMNI | MSGMNI | MSGMAP | SHMSEG | RCDMSZ |
| ----    | -----  | -----   | -----  | -----  | -----  | -----  | -----  | -----  |
| ex03    | 126    | 15      | 120    | A + 2  | 26     | 52     | 1178K  | 220M   |
| ex04    | 221    | 28      | 220    | A + 1  | 26     | 52     | 1340K  | 340M   |
| ex04_1  | 121    | 15      | 120    | A + 1  | 26     | 52     | 1178K  | 1300M  |
| ex03_1  | 221    | 28      | 220    | A + 1  | 25     | 50     | 1340K  | 2500M  |

RCDMSZ increases linearly when any of following items configured in UBB increases:

- MAXACCESSERS.
- RDMAQSIZE.
- The number of logic machines in the UBBCONFIG file.

After getting the RCDMSZ, you can adjust the actual size according to the following runtime factors:

- Frequency of start and exit of processes.

For performance reasons, shared memory resource are not immediately freed after the queue is destroyed. The more frequently processes start/stop, the more frequently queues are created/destroyed, the more resource s are occupied by stale queues. At most, you should allocate more shared memory using the following formula:

$$96 * \text{MAXSERVERS} * \text{RDMAQSIZE}$$

MAXSERVERS: A UBBCONFIG file attribute.

RDMAQSIZE: A UBBCONFIG file attribute.

If the `EXALOGIC_MSGQ_CACHE_SIZE` environment variable is set, at most  $3 * EXALOGIC\_MSGQ\_CACHE\_SIZE * MAXSERVERS * RDMAQSIZE$  more bytes of shared memory should be allocated.

For example, if `MAXSERVERS` is 100, `RDMAQSIZE` is 100000 and processes on machine `ex04` start/stop frequently, you should configure the shared memory size with at most 1300M (340M+960M) instead of 340M.

- The actual number of machines in the domain

The number of machines in the `UBBCONFIG` file is considered in calculating `RCDMSZ`. If new machines are added using `MIB`, you should increase the size linearly.

- The entry number of queue cache

In some scenarios, `Msgq_daemon` needs to allocate more shared memory when increasing queue cache entry number using `EXALOGIC_MSGQ_CACHE_SIZE`. For example, if Oracle Tuxedo process `start/stop` or `tpinit/tpterm` frequently. Then you should consider adjusting the initial `Msgq_daemon` shared memory size, otherwise the message transfer may fail.

The default value is 32 for each Oracle Tuxedo thread, and `tmloadcf -c` always assumes the value of 32. If you change the default value, adjust the `Msgq_daemon` shared memory size using the following method:

```
MEMSZ * CACHENUM / 32
```

`MEMSZ`: The current shared memory size calculated by above method.

`CACHENUM`: The maximum entry queue cache number for each Oracle Tuxedo thread.

**Note:** You can adjust the size using this method only when setting `EXALOGIC_MSGQ_CACHE_SIZE` at the machine level.

## Running Oracle Tuxedo on OVM

There are no special requirements for Oracle Tuxedo running on OVM.

## SALT and TMA on Exalogic

Before installing `SALT 11gR1 11.1.1.2.0` and `TMA 11gR1 11.1.1.2.0` on Exalogic, you must run `preinstall` first, as follows.

```
preinstall <Oracle Home>
```

`<Oracle Home>`: The directory you specified when Oracle Tuxedo installation.

After the installation, run postinstall, for example:

```
postinstall <Oracle Home >
```

<Oracle Home>: The directory you specified when Oracle Tuxedo installation.

## Migration

There are no special requirements if you do not use the Exalogic features. For more information, see [Oracle Tuxedo Interoperability Guide](#).

**Note:** Oracle Tuxedo 12c (12.1.1) only support hot upgrading from Tuxedo 11gR1 (11.1.1.3.0) if any of the Exalogic features are enabled.

