

Configuring Enterprise Manager for High Availability

*An Oracle White Paper
April 2004*

Configuring Enterprise Manager for High Availability

Introduction	3
Architecture Overview	3
Installation and Configuration for High Availability	5
Management Agent.....	5
Configure the agent to automatically start on boot and restart on failure.....	5
Configuring Restart for the Management Agent.....	5
Configure the connection between Management Agents and the Management Service	6
Install the Management Agent software on redundant storage.....	8
Configure all out-of-band notifications	9
Management Service.....	9
Configure SLB to abstract the underlying Management Service host names for easier reconnect after failure	9
Management Service installation should be done to non-clustered servers	10
Configure Management Service to use client side Oracle Net load balancing for Failover and Load Balancing.....	11
Install the Management Service software on redundant storage.....	11
Management Repository	11
Install into an existing RAC repository	11
Consider (physical) data guard for redundancy	12
Configuration within Grid Control	12
Console warnings, alerts and notifications.....	13
Configure additional error reporting mechanisms.....	13
Component Backup.....	13
Troubleshooting.....	13
Upload Delay for Monitoring Data.....	14
Notification Delay of Target State change	14
Conclusion.....	14
Appendix 1 Testing Architecture.....	15
Appendix 2 Agent configuration in Active/Passive configurations.....	16

Configuring Enterprise Manager for High Availability

INTRODUCTION

Every day, more Oracle customers deploy systems that are considered critical to their business. These systems often have strict availability requirements and maintenance windows. Downtime is often measured in minutes and maintenance windows are short. Oracle has addressed this business need with the rollout of the 'Unbreakable' database and blueprints for highly available systems such as the Maximum Availability Architecture. Now with the release of Oracle Enterprise Manager 10g Grid Control, Oracle has increased the manageability of highly available systems. This also increases the availability requirements for the manageability infrastructure.

The purpose of this paper is to describe a highly available deployment of Grid Control. After reviewing reading this paper the reader should understand the steps needed to configure each component for high availability. The reader will also be introduced to the strengths and limitations of the current solution and will have an understanding of how to recover from outages of each tier.

ARCHITECTURE OVERVIEW

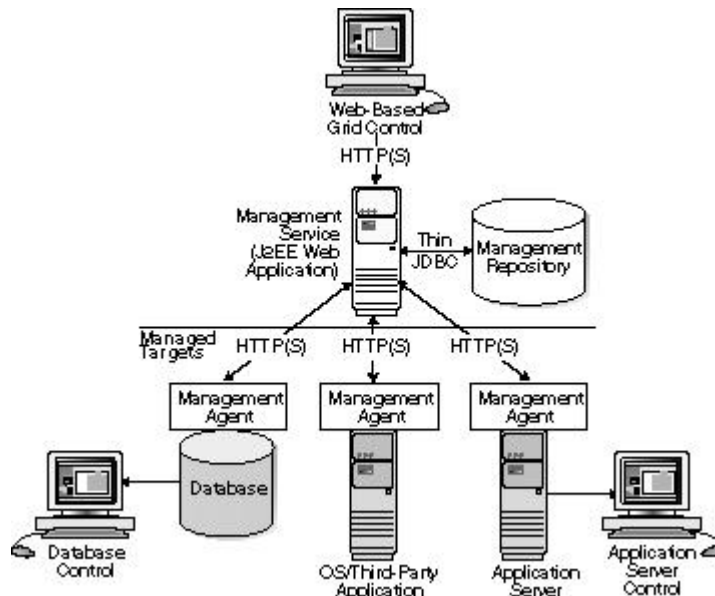
The architecture for a highly available Grid Control deployment is based on two key concepts, redundancy and component monitoring. Each component of Grid Control can be configured to apply both these concepts.

The components of Grid Control discussed in this paper include:

- The Management Agent - A process that is deployed on each monitored host, is responsible for monitoring all targets on the host, for communicating that information to the middle-tier Management Service, and for managing and maintaining the host and its targets.
- The Management Service - A J2EE Web application that renders the user interface for the Grid Control Console, works with all Management Agents to process monitoring and jobs information, and uses the Management Repository as its data store.
- The Management Repository - The schema in an Oracle Database that contain all available information about administrators, targets, and applications managed within Enterprise Manager.

The Management Agent uploads collected monitoring data to a Management Service. The Management Service in turn loads the data into the Management Repository. The Management Repository represents the persistent historic view of collected information that is presented to clients via a web user interface.

Changes in a target state either in an availability state change or detection of a notification dependent upon a metric threshold being crossed results in a notification being sent. The Management Agent detects this change and is



responsible for forwarding the information to the Management Service that in turn, records the state change in the Repository. Any registered users requesting notification have messages posted via registered notification methods by the Management Service and the console display updated

Figure 1: Overview of Enterprise Manage Architecture Components

For more information about the Grid Control architecture, see the Oracle Enterprise Manager 10g documentation:

- Oracle Enterprise Manager Grid Control Installation and Basic Configuration
- Oracle Enterprise Manager Concepts
- Oracle Enterprise Manager Advanced Configuration

The Oracle Enterprise Manager 10g documentation is available at the following location on the Oracle Technology Network (OTN):

<http://otn.oracle.com/documentation/oem.html>

Details on using Enterprise Manager to configure high availability features such as RMAN and Data Guard can be found in the Oracle documentation and in the Enterprise Manager Grid Control on-line help

INSTALLATION AND CONFIGURATION FOR HIGH AVAILABILITY

The following sections document best practices for installation and configuration of each Grid Control component.

Management Agent

Enterprise Manager uses a software process called the Oracle Management Agent to monitor a target. The Management Agent is a system daemon that consists of two processes, a process that provides monitoring, alerting and job system capabilities as well as a watchdog process that is responsible for insuring the Management Agent is up and available.

The data that is collected by the Management Agent is stored temporarily on the monitored host in files. Once the Management Agent deems it necessary to upload the information to the Grid Control system, it contacts the Management Service to establish a connection and uploads the data.

The Management Service accepts the data from the Management Agent, stores the information as files local to the Management Service and acknowledges receipt of the information to the Management Agent. Depending on the volume of work the Management Service is performing, a period of time may elapse before the Management Service loads the data into the Management Repository.

Notifications of alerts, warnings and target state changes do not follow this delayed model. When the Management Agent uploads the information, the Management Service commits the data immediately to the Management Repository before acknowledgement is returned to the Management Agent.

The Management Agent and its watchdog are started thru the command `‘$ORACLE_HOME/bin/emctl start agent.’`

Configure the agent to automatically start on boot and restart on failure

The Management Agent is started manually. It is important that the Management Agent be automatically started when the host is booted to insure monitoring of critical resources on the administered host. To that end, use any and all operating system mechanisms to automatically start the agent. For example, on UNIX systems this is done by placing an entry in the UNIX `/etc/init.d` that calls the agent on `boot`. or by the setting the Windows service to start automatically.

Configuring Restart for the Management Agent

Once the agent is started, the watchdog process monitors the Management Agent and attempts to restart it in the event of a failure. The behavior of the watchdog is controlled by environment variables set before the agent process starts. The

variables that control this behavior are listed below. All testing for this paper was done with the default settings.

- **EM_MAX_RETRIES** – This is the maximum number of times the watchdog will attempt to restart the agent within the **EM_RETRY_WINDOW**. The default in this release is to attempt restart of the Management Agent 3 times.
- **EM_RETRY_WINDOW** - This is the time interval in seconds that is used together with the **EM_MAX_RETRIES** environmental variable to determine whether the agent is to be restarted. The default in this release is 600 seconds.

The watchdog will not restart the agent if the watchdog detects that the agent has required restart more than **EM_MAX_RETRIES** within the **EM_RETRY_WINDOW** time period,

Configure the connection between Management Agents and the Management Service

Management Agents do not maintain a persistent connection to the Management Service. When a Management Agent needs to upload collected monitoring data or an urgent target state change, the Management Agent establishes a connection to the Management Service. If the connection is not possible, such as in the case of a network failure or a host failure, the Management Agent retains the data and re-attempts to send the information later.

There are two possible connection configurations between a Management Agent and its Management Service. Each of the alternatives defined below have strengths and drawbacks.

Alternative 1: Use a persistent connection from a Server Load Balance for Management Agent Communication to the Management Service

Server Load Balancers (SLBs) such as the F5 Networks Big-IP® provide logical service abstractions for network clients. Clients establish connections to the virtual service exposed by the SLB. The SLB routes the request to any one of a number of available servers that provide the requested service. The service chosen by an SLB as the destination is dependent upon the virtual service definition. One such criterion is whether a service is capable of accepting connections.

The Grid Control Management Service is a network service that can be fronted by a SLB to address the need for resiliency.

To accomplish the goal of having a highly available Management Service that the Management Agents can use for data upload, configure a virtual pool that consists of the hosts and the services that the hosts provide. In the case of the Management Services pool, the hostname and agent upload port would be specified. To insure a highly available Management Service, you should have two or more Management Services defined within the virtual pool.

Declare the pool such that any new connection between a Management Agent and the virtual pool member is persistent. This relationship is maintained and the Management Agent will upload to that the Management Service till the persistence period has elapsed or the Management Service is deemed to be inaccessible.

In order to have the Management Agent now upload information thru the SLB virtual service, modify the 'REPOSITORY_URL' setting in the `emd.properties` file located in the `$AGENT_HOME1/sysman/config` directory. The hostname and port specified must be that of the SLB Virtual Service.

This configuration will allow the Management Agent to make a permanent connection to a Management Service unless for some reason that management service becomes unavailable. In that event, the load balancer will connect the agent to any surviving Management Service, based on the policies configured in the load balancer (e.g. Round Robin or Least Loaded)

There is some risk within this proposed configuration. The Grid Control Data architecture requires that add collected monitoring data be loaded in chronological order. The Management Service, upon receipt of data, stores it temporarily in as a local file and acknowledges receipt to the Management Agent. The Management Service loads the data in a background thread in chronological order.

Should the data not get uploaded and the Management Agent fail over and attempt the upload of collected data to another Management Service, that new Management Service may upload data for data points more recent than was previously sent to the other host. The older data as a result may never get loaded and simply discarded. This may result in potential holes in historic data.

To minimize this potential, carefully monitor the amount of data pending load to the Management Repository via the Management System tab within the Grid Control console. Monitor the transaction backlogs and maintain enough Management Service processing capacity to reduce the potential for data loss

Alternative 2: Manually load balance connections of Management Agents to Management Services

Management Servers not only exist as the receivers of upload information from Management Agents, but also serve User interface in the form of HTML pages to clients as well as perform background processing tasks such as notification delivery and dispatch of jobs.

It is due to these additional tasks that are performed that the manual assignment of Management Agents to Management Services must be carefully managed and

¹ The AGENT_HOME will differ depending on the type of host. If the node is a part of a cluster managed with third party vendor cluster software, the Oracle Installation Software will append the name of the host to the end of the string specified for the Agent home directory for each node in the cluster. Otherwise, the AGENT_HOME will be the proper ORACLE_HOME path where the agent was installed.

balanced. Improper distribution of load from Management Agents to Management Services may result in perceived:

- sluggish user interface response
- delays in delivering notification messages
- delays in dispatching jobs
- backlog in monitoring information being uploaded to the Management Repository.

Management Agents are initially assigned to a Management Service as part of the installation process. This association is recorded on the Management Agent alone as the property 'REPOSITORY_URL' in the `$AGENT_HOME/sysman/config/emd.properties` file. This property defines the single Management Service that the Management Agent is to upload information to.

To keep the workload evenly distributed, the administrator must be aware of how many agents are configured for each Management Service and balance the Management Agents accordingly. The list of Management Agents and the Management Services that are uploaded is shown in the Management System tab of the Grid Control Console.

Maintenance of a single connection for the life of a Management Agent avoids the problem of management data potentially being received out of order and causing a loss of data as stated in the previous section. However, in the event of a Management Service outage, Management Agents will be unable to upload any information or state changes (although traffic can still be routed to agents from surviving Management Services). This means that telemetry and any alerts from the Management Agent will not be processed until the Management Service is restarted or recovered.

Connect each Management Agent in a RAC environment through to different Management Services. This provides a measure of redundancy in the event of a Management Service outage. If an outage occurs, other Management Agents monitoring nodes in the cluster will still be able to report their instance status to the Grid Control Console. This gives an administrator the ability to continue to monitor the cluster in question, while providing a quick diagnostic clue to the root cause of the problem.

Install the Management Agent software on redundant storage

The Management Agent persists its intermediate state and collected information using local files in the `$AGENT_HOME/$HOSTNAME/sysman/emd` sub tree under the Agent home directory.

In the event that these files are lost or corrupted before being uploaded to the repository, a loss of monitoring data and any pending alerts not yet uploaded to the Management Repository will occur.

At a minimum, configure these sub-directories on striped redundant or mirrored storage. Availability would be further enhanced by placing the entire \$AGENT_HOME on redundant storage. The agent home directory is shown by entering the command 'emctl getemhome' and the command line or from the Management Services/Agent tab in Grid Control

Configure all out-of-band notifications

The Enterprise Manager Grid Control deployment is configured out of the box such that connection failures between the Management Service and the Management Agent are detected. This is thru a process of heartbeats that the Management Agent performs against the Management Service and if the Management Service determines it has not heard back from the Agent, it pings it.

This condition does not however correct for the condition where the Management Agents are up and available however there are no Management Services to upload to or more importantly process notifications. For just such a situation, the Management Agent has the capability of sending an emergency notification when it is still up but has lost contact with the Management Service.

This provides another mechanism to alert the administrator of a Management Service failure. For more information, see the section 'Configuration with Grid Control.'

In the emd.properties file located in the \$AGENT_HOME/sysman/config directory, modify the property values for emd_email_address and emd_email_gateway to reflect a valid email address in your system. The parameter emd_from_email_address should also be modified to reflect the name of the system sending the alert for faster root cause identification.

In addition, any custom notification script can be executed by the Agent in the event of a failure to communicate with the Management Service. This script can be set to execute by modifying the 'emdFailureScript' entry in the Agent emd.properties file

Management Service

The Management Service element of the Enterprise Manager Grid Control product acts both as the receiver of information from Management Agents as well as serves out the User Interface in the form of HTML pages. It does this by maintaining a connection to the configurations database repository and responding to requests over HTTP.

Configure SLB to abstract the underlying Management Service host names for easier reconnect after failure

A hardware server load balancer (SLB) such as F5 Networks Big-IP® can be used as the front end to abstract the number and location of Management Services and appear as a single service. Under that abstraction, the SLB parcels the work to any

number of Management Service processes that it has in its 'virtual pool.'. For any Grid Control installation with an availability requirement there should be a minimum of two Management Service processes installed. Coupled with a Server Load Balancer (SLB), this provides a method for constant communication to the Grid Control Console in the event of the failure of a Management Service.

Key to the liveliness of the system is active detection of Management Service failures. This is done in a variety of means however the F5 Networks Big-IP® SLB can be configured to monitor the underlying Management Service. This is accomplished by configuring a 'monitor' on the SLB. The monitor definition indicates the HTTP request that is to be sent to a Management Service, the expected result in the event of success and the frequency of evaluation.

During testing, the Big-IP® was configured to check the state of the Management Service every 5 seconds. On three successive failures, the SLB would mark the component as unavailable and no longer route requests to it. The monitor was configured to send the string 'GET /em/upload' over HTTP and expect to get the response 'Http XML File receiver'.

To complete the abstraction of Management Services via the SLB changes are required within the Apache configuration files located in the Management Service install home. The most straightforward way to accomplish this abstraction is to edit the Apache Configuration files (`httpd.conf` or in the event the Management Services are configured for SSL, `ssl.conf`) and declare for the specific `VirtualHost` a `ServerName`.

The `ServerName` specified must match the SLB virtual service that has been configured. In the event of a Management Service failure when the end user tries to reconnect, they will be automatically redirected to the next available Management Service.

Beware not to mask the `ServiceName` returned when the `VirtualHost` for Management Agent uploads is used ². Setting it this way will allow for redirection back through the SLB and still allow it to determine which agents are uploading through each OMS.

Management Service installation should be done to non-clustered servers

During testing, we determined that Management Service processes in this release cannot be installed on any machines running under a cluster, whether it is CRS or vendor cluster software. Install Management Services to single nodes and use the method described above for failover and availability.

² The Management Service listens on an alternate set of ports for Management Agent uploads. The port is typically 4889 for HTTP and 4444 when the Management Agents are secured.

Configure Management Service to use client side Oracle Net load balancing for Failover and Load Balancing

When you use a RAC cluster, a standby system, or both to provide high availability for the Management Repository, the Management Service can be configured to use an Oracle Net connect string that will take advantage of redundancy in the repository. Correctly configured, the Management service process will continue to process data from Agents even during a database node outage.

In the `$OMS_HOME/sysman/config` directory, modify the `emdRepConnectDescriptor` entry in the `emoms.properties` file to point to the appropriate repository instances. The following example shows a connect string required to support a 2-node RAC configuration. Note the backslash (\) before each equal sign (=) sign.

```
oracle.sysman.eml.mntr.emdRepConnectDescriptor=
(DESCRIPTION\=(ADDRESS_LIST\=(FAILOVER\=ON)
(ADDRESS\=(PROTOCOL\=TCP)(HOST\=haem1.us.oracle.com)
(PORT\=1521)))(ADDRESS\=(PROTOCOL\=TCP)
(HOST\=haem2.us.oracle.com)(PORT\=1521)))
(CONNECT_DATA\=(SERVICE_NAME\=em10))
```

Install the Management Service software on redundant storage

The Management Service contains results of the intermediate collected data before it is loaded into the repository. The `$OMS_HOME/sysman/recv` directory contains these files and is typically empty when the Management Service is able to load data as quickly as it is received. Once the files are received by the Management Service, the Management Agent considers them committed and therefore removes its local copy. In the event that these files are lost before being uploaded to the repository, data loss will occur. At a minimum, configure these sub-directories on striped redundant or mirrored storage.

Similar to the Management Agent directories, availability would be further enhanced by placing the entire Management Service software tree on redundant storage. This can also be determined at the command line using the `'emctl getemhome'` or by using the Management Services Tab in Grid Control

Management Repository

The Management Repository is the central location for all historical data managed by Grid Control. Redundancy at this tier is provided by standard database features and best practices.

Install into an existing RAC repository

The Grid Control installation process does not directly support installation into a RAC repository. The recommended installation method is to install the 9.2 database software first and create a RAC database. When this is complete, install

the Enterprise Manager software, selecting the ‘Enterprise Manager 10g Grid Control Using an Existing Database’ installation option.

The installation does not transparently support the installation of the Enterprise Manager 10g Grid Control into a RAC database. Specify the SID of one of the cluster instances when prompted for during the installation. After the installation of the Enterprise Manager 10g Grid Control Management Service, you should modify the connection string the Management Service uses to take advantage of client failover in the event of a RAC host outage (refer to section, ‘Configure Management Service to use client side Oracle Net load balancing for Failover and Load Balancing’)

The installation process also does not allow modification of the size of the required Enterprise Manager tablespaces (although it does allow for specification of the name and location of data files that are to be used by the Enterprise Manager 10g Grid Control schema). The default sizes for the initial data file extents depend on using the AUTOEXTEND feature and as such are insufficient for a production installation. This is particularly problematic where storage for the RAC is on a raw device.

If the RAC database being used for the repository is configured with raw devices there are two options for increasing the size of the repository. You can create multiple raw partitions, with the first one equal to the default size of the tablespace as defined by the installation process. Alternatively, you can create the tablespace using the default size, create a dummy object that will increase the size of the tablespace to the end of the raw partition, then drop that object. Regardless, if raw devices are used, disable the default space management for these objects, which is to auto-extend.

Consider (physical) data guard for redundancy

Clients who require greater uptime or an off-site copy of the repository can use Oracle Data Guard in conjunction with Grid Control. This alternative can be used regardless of whether or not you are using a RAC database. Currently, only the use of physical data guard is supported.

A Data Guard instance must be created manually using the steps documented in the Data Guard documentation.

CONFIGURATION WITHIN GRID CONTROL

Grid Control comes preconfigured with a series of default rules to monitor many common targets. These rules can be extended to monitor the Grid Control infrastructure as well as the other targets on your network to meet specific monitoring needs.

Console warnings, alerts and notifications

The following list is a set of recommendations that extend the default monitoring performed by Enterprise Manager. Use the Notification Rules link on the Preferences page to adjust the default rules provided on the Configuration/Rules page:

- Ensure the 'Agent Unreachable' rule is set to alert on all 'agent unreachable' and agent clear errors.
- Ensure the 'Repository Operations Availability' rule is set to notify on any unreachable problems with the OMS or repository nodes. Also modify this rule to alert on the 'Targets Not Providing Data' condition and any database alerts that are detected against the database serving as the Enterprise Manager 10g Grid Repository.

Modify the 'Agent Upload Problems' Rule to alert when the 'Management Service' status has hit a warning or clear threshold

Configure additional error reporting mechanisms

Enterprise Manager provides error reporting mechanisms through email notifications, PL/SQL packages, and SNMP alerts. Configure these mechanisms based on the infrastructure of the production site. If using email for notifications, configure the notification rule through the Grid Control Console to notify administrators using multiple SMTP servers if they are available. This can be done by modifying the default email server setting under Setup/Notifications Methods

Component Backup

Backup procedures for the database are well established standards. Configure backup for the repository using the RMAN interface provided in the Grid Control Console. Refer to the RMAN documentation or the Maximum Availability architecture document for detailed implementation instructions.

In addition to the repository, the Management Service and Management Agent should also have regular backups. Backups should be performed after any configuration change. Best practices for backing up these tiers are documented in the 'Enterprise manager Backup and Recovery' paper.

Troubleshooting

In the event of a problem with Grid Control, the starting point for any diagnostic effort is the console itself. The Management System tab provides access to an overview of all Management Service operations and current alerts. Other pages summarize the health of Management Service processes and logged errors. These pages are useful for determining the causes of any performance problems as the summary page shows at a historical view of the amount of files waiting to be loaded to the Management repository and the amount of work waiting to be completed by agents.

Upload Delay for Monitoring Data

When assessing the health and availability of targets through the grid control console, information is slow to appear in the GUI, especially after a Management Service outage. The state of a target in the console may be delayed after a state change on the monitored host. Use the Management System page to gauge backlog for pending files to be processed.

Notification Delay of Target State change

The model used by the agent to assess the state of health for any particular monitored target is poll based. Management Agents immediately post a notification to the Management Service as soon as a change in state is detected. This infers that there is some potential delay for the agent to actually detect a change in state.

CONCLUSION

The architecture for a highly available Enterprise Manager deployment is based on two key concepts, redundancy and component monitoring. With the release of Enterprise Manager 10g Grid Control, Enterprise Manager is more capable than ever of being the single backbone console to monitor the entire data center. Implementation of the recommendations listed above this paper will greatly increase the availability of the Grid Control Console for continuous use in management of high available operations. As with any other highly available applications requiring high availability, best practices and testing both before and during production are the keys to success.

APPENDIX 1 TESTING ARCHITECTURE

Repository

2 x Sun Ultra 250 clustered with 2 processors and 2 GB memory

Solaris 2.8, patch level 18

Oracle 9.2.0.4 no additional patches

Veritas 3.5

Grid Control

Enterprise Manager 10.1.0.2 for Agents and Management Services

2 Management Services on Sun Ultra 250 with 2GB of memory

Testing environment included 20 hosts with 160 targets

Management service redundancy configured using a F5 Networks Big-IP
540 ® Load balancer

APPENDIX 2 AGENT CONFIGURATION IN ACTIVE/PASSIVE CONFIGURATIONS

An active/passive install describes a particular type of clustered environment where one node of a two node cluster is running an application and is configured using the cluster software to move the application to the second node in the event of a failure.. This is done by using the concept of a floating IP address and global storage. This configuration is done using cluster ware provided by third party vendors such as Sun or HP. Refer to vendor specific documentation for implementation details.

During the install of the Management Agent, the discovery process will find the floating IP address that was configured.. Upon node failure, the Management Agent moves with the application to the surviving node. A management agent will follow the application during a failover and be able to report on the application targets regardless of the node



Configuring Enterprise Manager 10g for High Availability

April 2004

Author: James Viscusi

Contributing Authors: Nestor Dutko

Oracle Corporation

World Headquarters

500 Oracle Parkway

Redwood Shores, CA 94065

U.S.A.

Worldwide Inquiries:

Phone: +1.650.506.7000

Fax: +1.650.506.7200

www.oracle.com

Copyright © 2004, Oracle. All rights reserved.

This document is provided for information purposes only and the contents hereof are subject to change without notice.

This document is not warranted to be error-free, nor subject to any other warranties or conditions, whether expressed orally or implied in law, including implied warranties and conditions of merchantability or fitness for a particular purpose. We specifically disclaim any liability with respect to this document and no contractual obligations are formed either directly or indirectly by this document. This document may not be reproduced or transmitted in any form or by any means, electronic or mechanical, for any purpose, without our prior written permission.

Oracle is a registered trademark of Oracle Corporation and/or its affiliates. Other names may be trademarks of their respective owners.