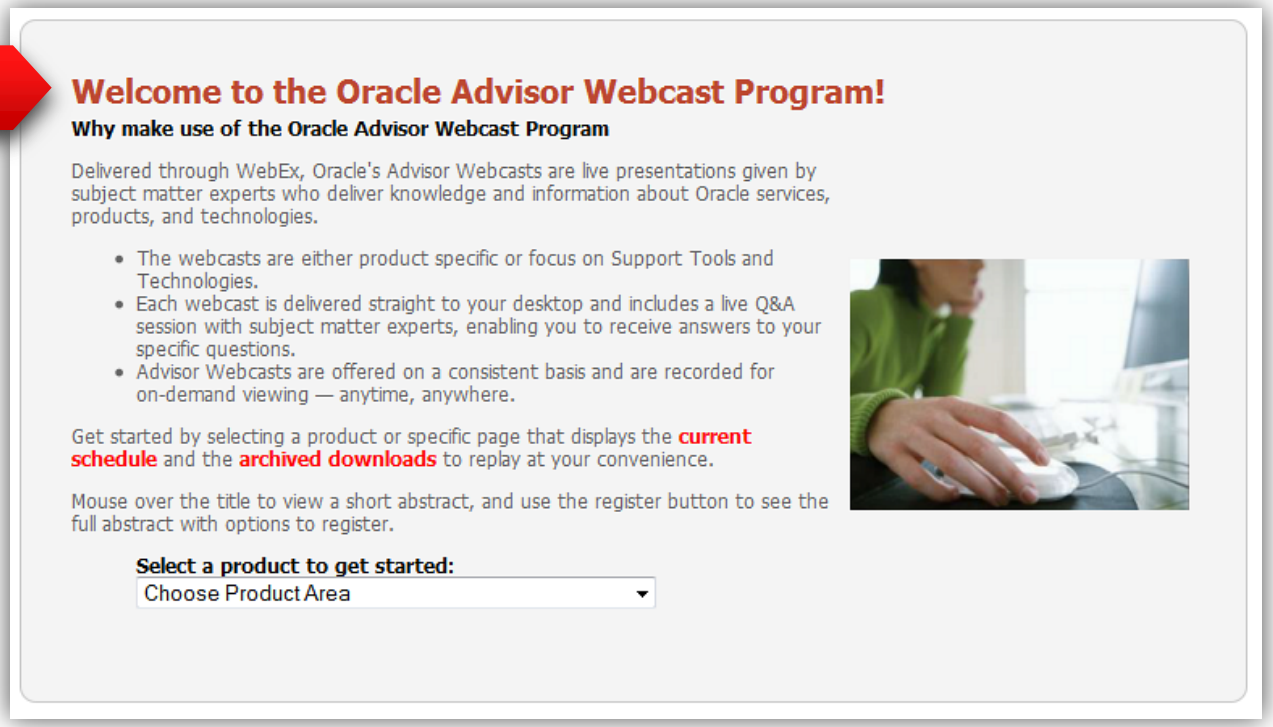
Upcoming: Exadata计算节点的系统还原

Oracle Advisor Webcast Program

- Current schedule
- Archived recordings
- Doc ID 740966.1



Upcoming: Exadata计算节点的系统还原

Copyright © 2014, Oracle and/or its affiliates. All rights reserved.

议程

- 演讲部分 – 大约45分钟
- Q&A部分–大约15分钟

  – 网络与会者请使用Q&A面板提问
  – 电话与会者请使用Q&A面板或者电话提问

Upcoming: Exadata计算节点的系统还原

Copyright © 2014, Oracle and/or its affiliates. All rights reserved.

注意 – 音频的选择

您可以

- 在电脑上收听音频广播
- 或者加入电话会议（通过电话接入号码）

Upcoming: Exadata计算节点的系统还原

Copyright © 2014, Oracle and/or its affiliates. All rights reserved.

流媒体 – 音频广播

- 仅限于收听而不能提问
- 好处: 不需要拨号接入
- 如果有问题怎么办?
  把您的问题输入到Q&A 面板上
- 如果您需要通过语音直接发问，就请您接入我们的电话会议系统。
- 电话会议的接入细节将在下一页介绍。

Upcoming: Exadata计算节点的系统还原

Copyright © 2014, Oracle and/or its affiliates. All rights reserved.

注意 – 电话会议的接入信息

电话会议的接入细节:

1. 会议 ID: 2104416
2. 中国北方地区免费接入号码: 1080 074 413 29
3. 中国南方地区免费接入号码: 1080 044 111 82
4. 中国台湾地区免费接入号码: 0080 104 4259
5. 各国的免费接入号码在 Note 1148600.1 可以查到

您可以在任何时候通过 WebEx 菜单来查看这些信息:

Communicate > Teleconference > Join Teleconference

Upcoming: Exadata计算节点的系统还原

Copyright © 2014, Oracle and/or its affiliates. All rights reserved.

ORACLE

Oracle Advisor Webcast

Exadata计算节点的系统还原

Peng Wang
Senior Technical Support Engineer

ORACLE SUPPORT

议程

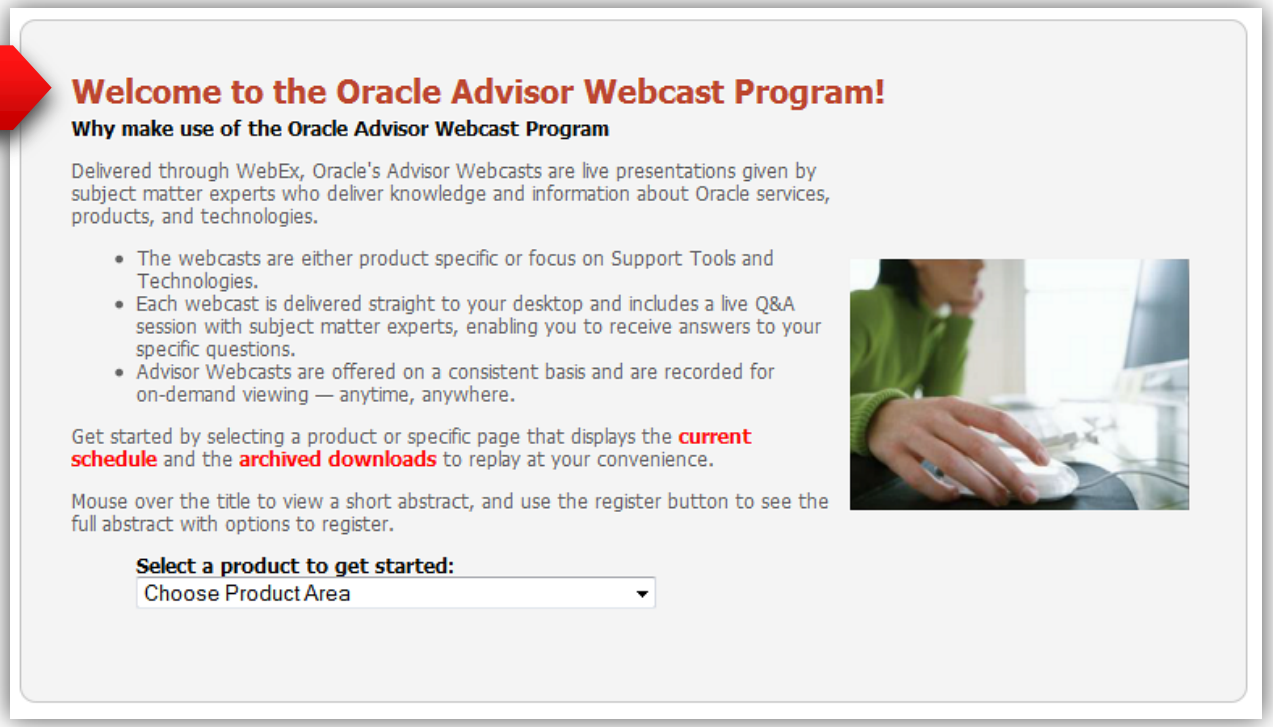
- 演讲部分 – 大约45分钟
- Q&A部分–大约15分钟

  – 网络与会者请使用Q&A面板提问
  – 电话与会者请使用Q&A面板或者电话提问

Copyright © 2014, Oracle and/or its affiliates. All rights reserved.

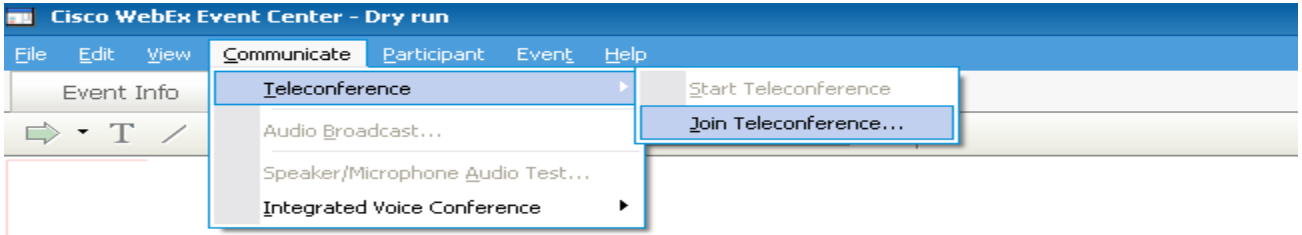
Oracle Advisor Webcast Program

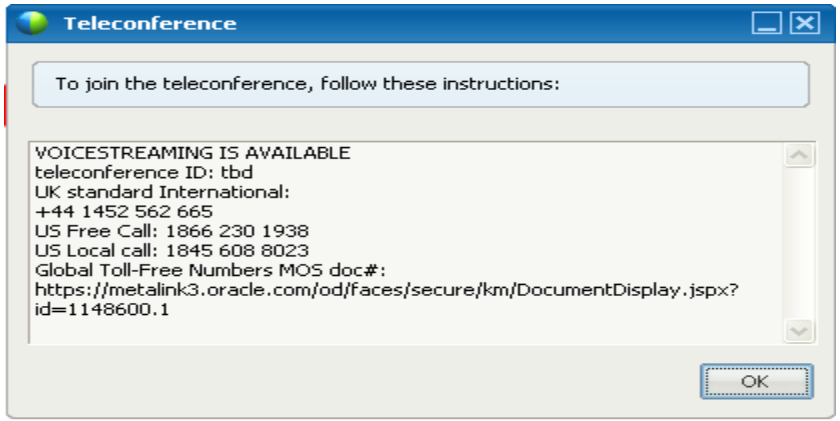
- Current schedule
- Archived recordings
- Doc ID 740966.1



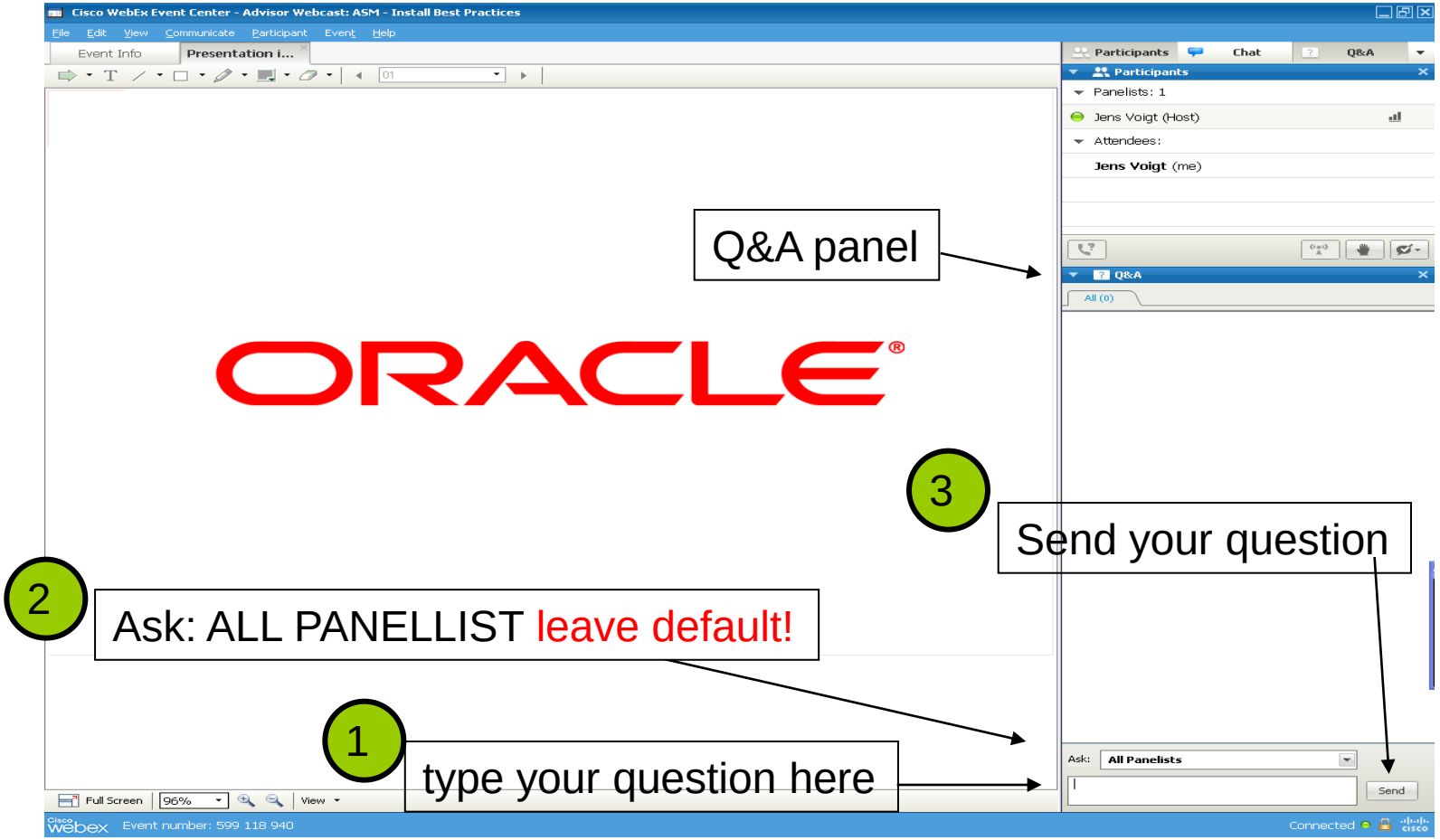
Copyright © 2014, Oracle and/or its affiliates. All rights reserved.

AUDIO INFO – Join Teleconference

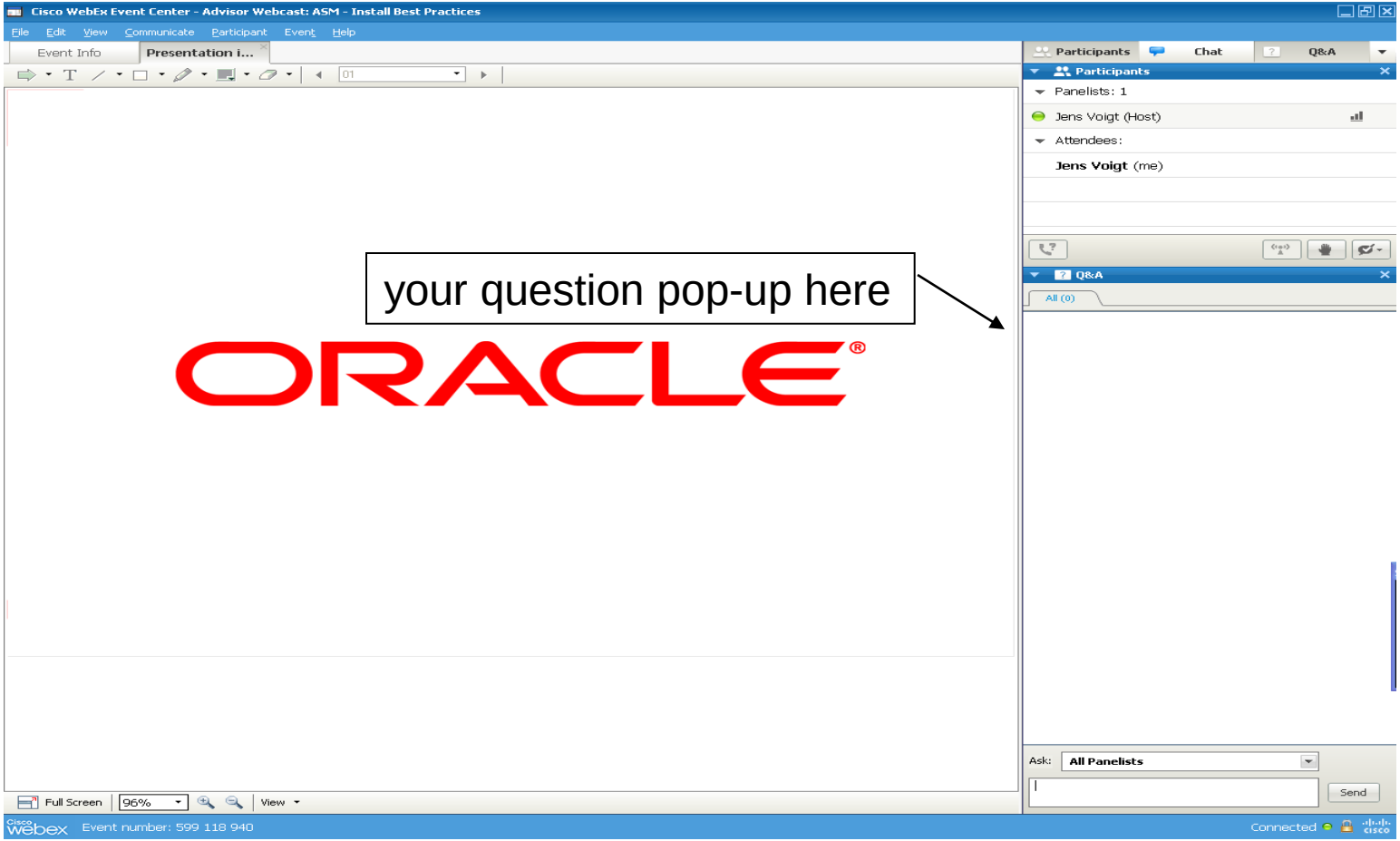




Copyright © 2014, Oracle and/or its affiliates. All rights reserved.



Copyright © 2014, Oracle and/or its affiliates. All rights reserved.



Copyright © 2014, Oracle and/or its affiliates. All rights reserved.

Safe Harbor Statement

The following is intended to outline our general product direction. It is intended for information purposes only, and may not be incorporated into any contract. It is not a commitment to deliver any material, code, or functionality, and should not be relied upon in making purchasing decision. The development, release, and timing of any features or functionality described for Oracle's products remains at the sole discretion of Oracle.

Copyright © 2014, Oracle and/or its affiliates. All rights reserved.

ORACLE

Oracle Advisor Webcast

Exadata计算节点的系统还原

Peng Wang
Senior Technical Support Engineer

ORACLE SUPPORT

Objectives

- Know when to do BMR
- Perform a Bare Metal Restore
- How to avoid need of a BMR

Copyright © 2014, Oracle and/or its affiliates. All rights reserved.

Agenda

Bare Metal Restore of a Compute Node.

- Determine when a Bare Metal Procedure (BMR) is necessary
- Prerequisites for Bare Metal Procedure
- Overview of Bare Metal Procedure
- Recorded Demo of Bare Metal Procedure
- Backup Options that help to prevent the need for a BMR

Copyright © 2014, Oracle and/or its affiliates. All rights reserved.

Agenda

What is not covered?

- Hardware diagnosis of failed node.
- Restoring custom scripts, cron jobs, 3rd party monitoring software, etc
- The restoring of any site specific OS customizations or provisioning
- Rare cases where there are no surviving database nodes - Reimaging and OneCommand is needed

Copyright © 2014, Oracle and/or its affiliates. All rights reserved.

Determine When a Bare Metal Restore is needed

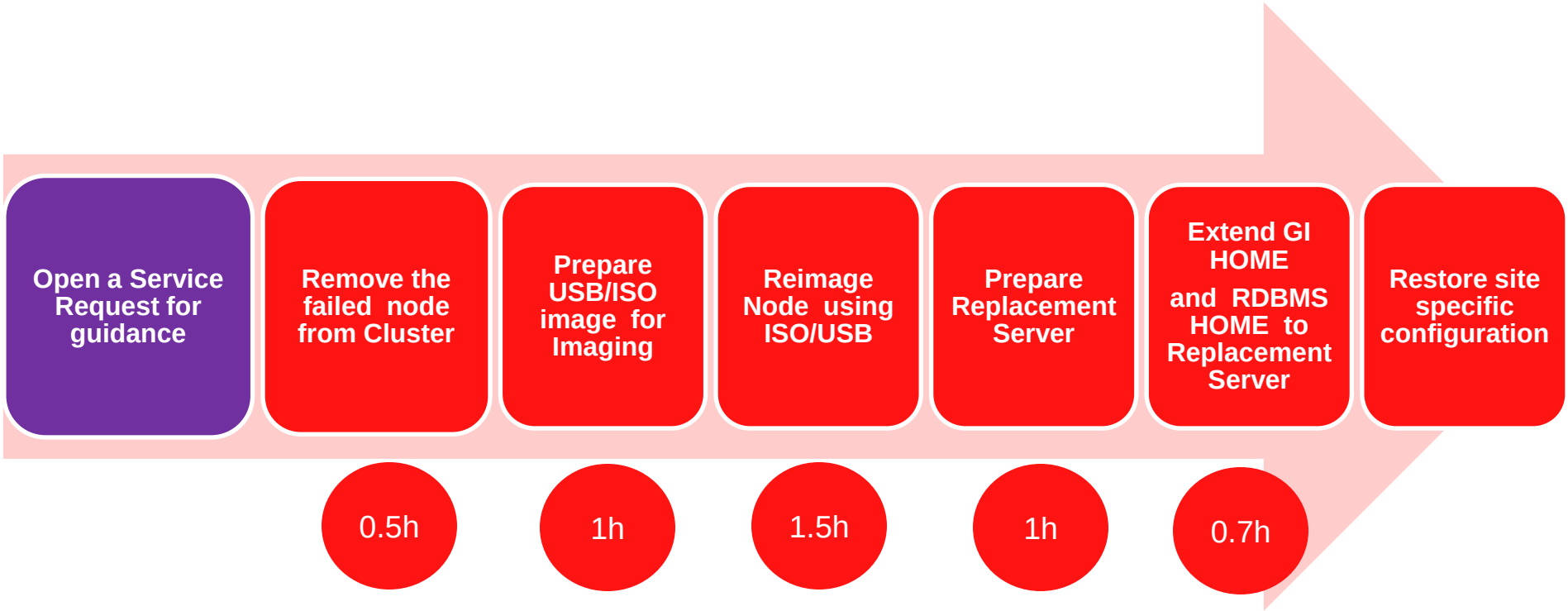
A Bare Metal Restore procedure is needed when:

- Node can not be brought online due to root file system corruption
- Typically as a result of hardware failures or filesystem corruption
- Can also be due to user error (LVM resizing, rm –rf /, mkfs, etc)
- When a node is physically replaced or reprovisioned to a different rack

Copyright © 2014, Oracle and/or its affiliates. All rights reserved.

Overview Bare Metal Restore Procedure

The high level steps are:



Copyright © 2014, Oracle and/or its affiliates. All rights reserved.

Open A Service request for guidance

- Have Service request open to obtain assistance
- A member of Engineered Systems Support Team will provide guidance
- Engage your DBA and Sys Admin
- The customer is responsible for driving the BMR

Copyright © 2014, Oracle and/or its affiliates. All rights reserved.

Information about our demo environment

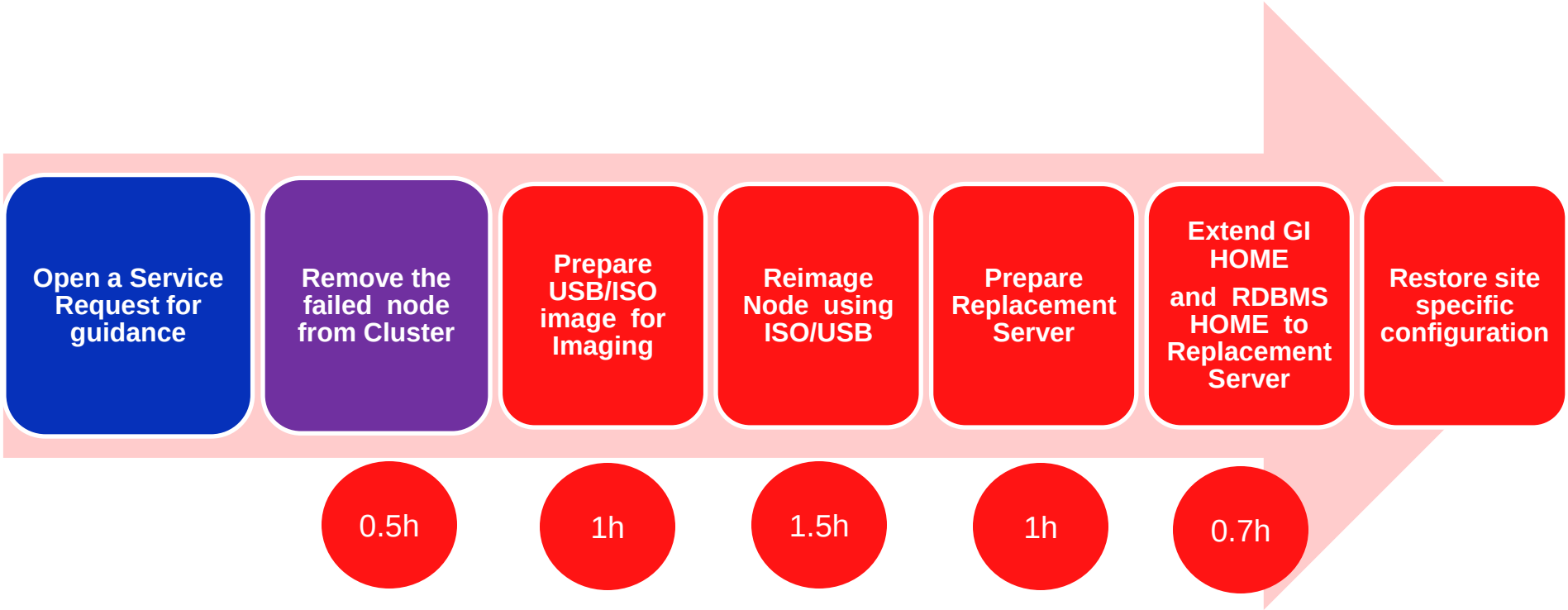
To keep in mind while following the demos and sample commands:

Quarter rack – 2 database nodes
1st node: dmorldb05 – surviving node, also referenced as root@surviving
2nd node: dmorldb06 - failed node, also referenced as root@reimaged
Grid Infrastructure (GI) Home: /u01/app/11.2.0.3/grid
RDBMS Home: /u01/app/oracle/product/11.2.0.3/db_home1
oracle user is the software owner of both GI and RDBMS homes
If the command prompt is prefixed with “#”, the command is run as root
If the command prompt is prefixed with “$”, command is run as oracle

Copyright © 2014, Oracle and/or its affiliates. All rights reserved.

Overview Bare Metal Restore Procedure

The high level steps are:



Copyright © 2014, Oracle and/or its affiliates. All rights reserved.

Remove the failed node from Cluster.

Perform on only one of the surviving nodes

- Remove failed Node from Clusterware Configuration
- Update Oracle Inventory to remove failed node from node list
- Verify that the node is completely removed
- Remove SSH host identification key of failed node from all surviving nodes for both oracle and root users. dcli can also be used

```
ssh-keygen -R <hostname of failed node>
```


Copyright © 2014, Oracle and/or its affiliates. All rights reserved.

Remove the failed Node

Perform on a single surviving node – GRID HOME

- Disable Listener

```
[oracle@surviving]$ srvctl disable listener -n dmorldb06
```

- Make sure the Failed Node is unpinned

```
[oracle@surviving]$ olsnodes -s -t
```

- Determine name of VIP Resources for the failed node and stop it

```
[root@surviving ~]# srvctl status nodeapps
VIP dmorl05-vip is enabled
VIP dmorl05-vip is running on node: dmorldb05
VIP dmorl06-vip is enabled
VIP dmorl06-vip is running on node: dmorldb05
[root@surviving ~]# srvctl stop vip -i dmorldb06-vip
```


Copyright © 2014, Oracle and/or its affiliates. All rights reserved.

Remove the failed Node

Perform on only one of the surviving nodes – RDBMS Home

- Update the Node List in Inventory (RDBMS_HOME)

```
oracle@surviving]$ ./runInstaller -updateNodeList  \
ORACLE_HOME=/u01/app/oracle/product/11.2.0.3/db_home1  "CLUSTER_NODES=dmorldb05”
```

➢*(CLUSTER_NODES contains the comma separated list of surviving nodes)*

Copyright © 2014, Oracle and/or its affiliates. All rights reserved.

Remove the failed Node

Perform on a single surviving node GRID HOME

- Delete the Node From Cluster Registry

```
[root@surviving]# crsctl delete node -n dmorldb06
CRS-4661: Node dmorldb06 successfully deleted.
```

- Update the Node List for Grid Home in central and local inventories:

```
[oracle@surviving]$ ./runInstaller -updateNodeList ORACLE_HOME=/u01/app/11.2.0.3/grid
"CLUSTER_NODES=dmorldb05“ CRS=TURE
```

(CLUSTER_NODES contains the comma separated list of surviving nodes)

Copyright © 2014, Oracle and/or its affiliates. All rights reserved.

Remove the failed Node -Verify

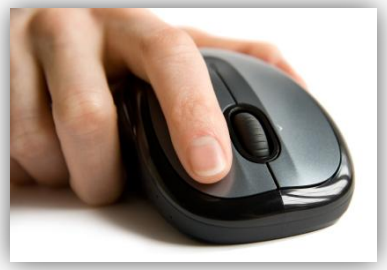
Perform on a single surviving node

- Verify the node deletion is successful

```
[oracle@surviving]$ cluvfy stage -post nodedel -n dmorldb06 –verbose
Performing post-checks for node removal
Checking CRS integrity...
The Oracle clusterware is healthy on node "dmorldb05“
CRS integrity check passed
Result:
Node removal check passed
Post-check for node removal was successful.
```


Copyright © 2014, Oracle and/or its affiliates. All rights reserved.

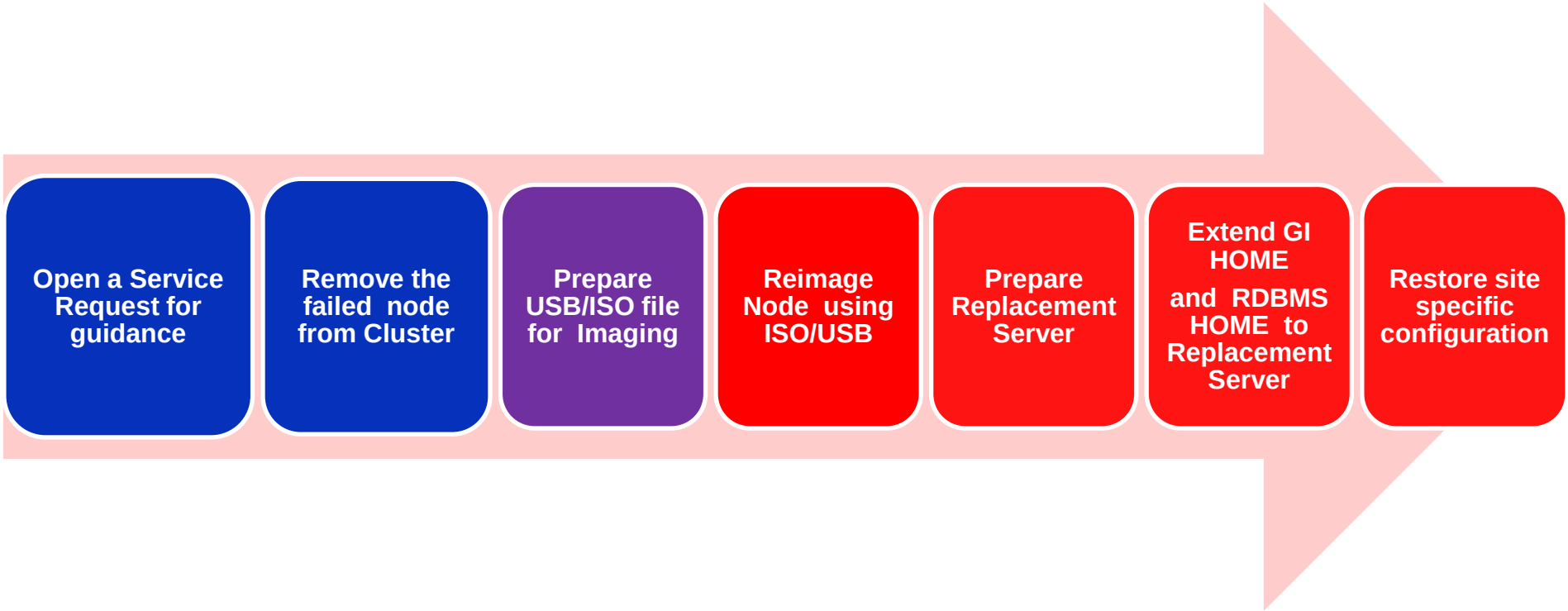
Demo – Removing node



Copyright © 2014, Oracle and/or its affiliates. All rights reserved.

Overview Bare Metal Restore Procedure

The high level steps are:



Copyright © 2014, Oracle and/or its affiliates. All rights reserved.

Prepare USB/ISO for Reimaging

- USB or ISO ?
- Download the Software
- Prepare USB/ISO

Copyright © 2014, Oracle and/or its affiliates. All rights reserved.

Prepare USB/ISO for Reimaging

USB or ISO Considerations

- ISO Image for attaching to ILOM console as virtual CD-ROM drive.
- For ISO method, the workstation used for reimaging must have good uplink bandwidth to the database node, as ISO is ~1.5 GB in size
- ILOM CD-ROM virtualization requires 32-bit JRE/JDK
- USB flash/thumb drive must be physically attached to the DB node
- **Access to video console (via ILOM Remote Console, KVM - if applicable - or from data center via notebook connected to the Cisco management switch or via DB node's VGA port) is required. Access to ILOM's serial /SP/console is not a substitute. – See slide notes for further info.**

Copyright © 2014, Oracle and/or its affiliates. All rights reserved.

Prepare USB/ISO for Reimaging

Download the Software

- Identify current Exadata software version from a Surviving node.
- Run # *imageinfo* to identify the current version
- For current versions older than 11.2.3.1.0, it may be required to reimage to the initial factory image. Run # imagehistory to determine the initial factory image version.
- Download the software from edelivery.oracle.com
  - For "*Select a Product Pack*", select *Oracle Database*
  - For "*Platform*", select *Linux x86-64*
  - Select the appropriate 11.2.X version of the *Oracle Exadata Storage Software Media Pack*
  - Download the appropriate version of *Oracle Database Machine Database Host*

Copyright © 2014, Oracle and/or its affiliates. All rights reserved.

Prepare USB/ISO for Reimaging

Stage the software and create the ISO or USB image

- The downloaded zip archive must be staged on an EL or RHEL Linux machine that has grub 0.97 and tar with bzip2 support. Running *grub –version* will show the grub version
- All actions must be run as root
- Unpack the software:

-unzip the V<xxxx>.zip file and the computeImageMaker<version>.tar.zip file
-untar the computeImageMaker<version>.tar file by running:
  tar -pxvf computeImageMaker<version>.tar
cd to dl360 directory

- Review the **README_FOR_FACTORY.txt** file for additional reimaging methods beyond those covered here: PXE, NFS

Copyright © 2014, Oracle and/or its affiliates. All rights reserved.

Prepare USB/ISO for Reimaging

Steps to create the ISO image

- Run as root:

```
# ./makeImageMedia.sh -dualboot no <chosen_filename>.iso
```

- Once the ISO file is generated, copy it to the workstation which will be used for the reimaging and where you will run the ILOM Remote Console from

Copyright © 2014, Oracle and/or its affiliates. All rights reserved.

Prepare USB/ISO for Reimaging

Steps to create the USB image

- Insert the USB thumb drive, then determine the assigned block device name by running lsscsi or by reviewing the entries added to /var/log/messages after insertion
- Run as root:

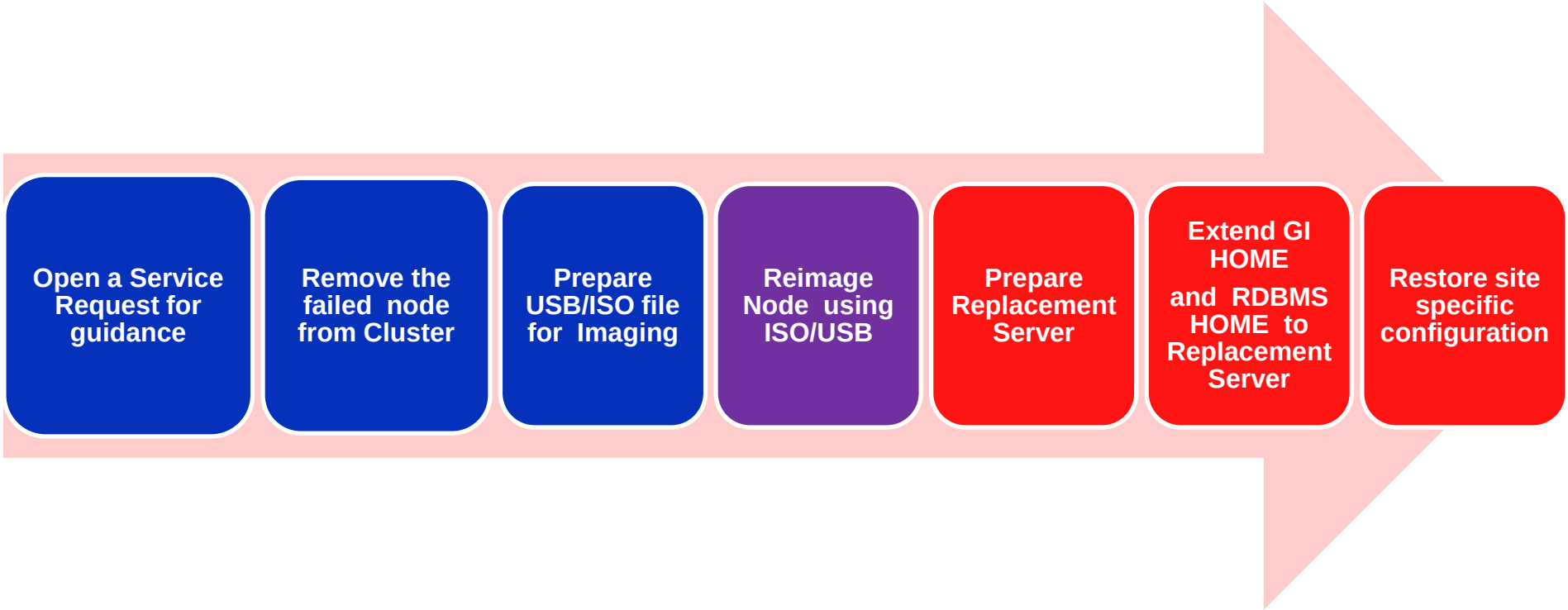
```
# ./makeImageMedia.sh -dualboot no
```

- Once the thumb drive is ready, remove it and connect it to DB Node

Copyright © 2014, Oracle and/or its affiliates. All rights reserved.

Overview Bare Metal Restore Procedure

The high level steps are as below



Copyright © 2014, Oracle and/or its affiliates. All rights reserved.

Reimage the Server

Attach USB to Node or ISO to ILOM console

- If using ISO method:
  - Launch the ILOM Remote Console and attach the CD-ROM ISO image
  - Set the Next Boot Device as CD-ROM, then power cycle the Node
- If using the USB method:
  - Power cycle the Node, then enter the one-time boot menu (Press **F8** during BIOS POST) and select the USB drive as boot device
- Once imaging media has booted, follow the prompts

Copyright © 2014, Oracle and/or its affiliates. All rights reserved.

Demo – Reimage the node

Copyright © 2014, Oracle and/or its affiliates. All rights reserved.

Re-IP the node

The ipconf at first boot

- After reimage is complete, once node boots after power on, it will prompt for IPs, hostnames, DNS, NTP, etc
- Use the /opt/oracle.cellos/cell.conf of a surviving node in same rack
- Make sure to apply the appropriate offset to hostnames and IPs
- If manual changes to interfaces were made post deployment, consult sysadmin and supply the current/accurate information to ipconf, not the obsolete one in the cell.conf
- See slide notes for steps to redirect ipconf to serial console for copy/paste support

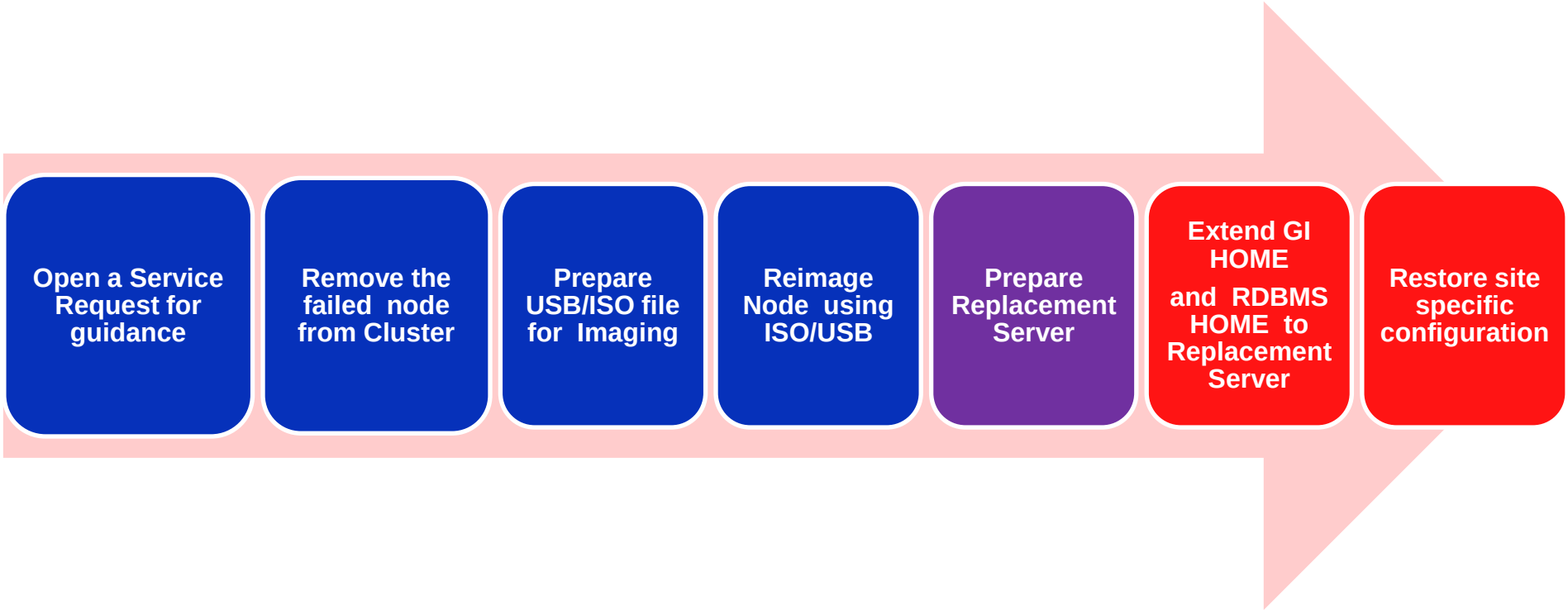
Copyright © 2014, Oracle and/or its affiliates. All rights reserved.

Demo – Re-IP the node

Copyright © 2014, Oracle and/or its affiliates. All rights reserved.

Overview Bare Metal Restore Procedure

The high level steps are as below



Copyright © 2014, Oracle and/or its affiliates. All rights reserved.

Prepare the Replacement Server:

Copy Configuration files :

- COPY /PREPARE CONFIG files…
  - Copy the **/etc/security/limits.conf** file
  - Merge the contents of **/etc/hosts**
  - Copy the ***/etc/oracle/cell/network-config/cellinit.ora*** file and update the IP address to reflect the IP address of the bondib0 interface on the replacement database server
  - Copy the **/etc/oracle/cell/network-config/cellip.ora** file. The content of the cellip.ora file should be the same on all database servers.

Copyright © 2014, Oracle and/or its affiliates. All rights reserved.

Prepare the Replacement Server:

Prepare OS Groups and Users

- CREATE OS USERs and GROUPs

  Add the group (or groups) for the Oracle software owner (typically, the owner is oracle):

- On the surviving node, obtain the current group information:

```
[root@surviving]# id oracle
uid=1000(oracle) gid=1001(oinstall)  groups=1001(oinstall),1002(dba)
```

- On the replacement node, use the groupadd command to add the groups:

```
[root@replacement#  groupadd -g 1001 oinstall
[root@replacement]#  groupadd –g 1002 dba
```

- On the replacement node, use the useradd command to add the users:

```
[root@replacement]#  useradd -u 1000 -g 1001 –G 1001,1002 -m -d /home/oracle -s /bin/bash
oracle
```

- On the replacement node, set passwords for oracle user
- Set the same ownership on the cellip.ora and cellinit.ora files as on surviving node

Copyright © 2014, Oracle and/or its affiliates. All rights reserved.

Prepare the Replacement Server:

SSH equivalence is for oracle user is needed before Extending Clusterware

- Set SSH equivalence. On reimaged node:

➢ i. Switch to the oracle account:

```
[root@replacement]# su - oracle
```

➢ ii. Copy/create the group file listing the nodes in the Oracle Cluster: /opt/oracle.SupportTools/dbs_group

➢ iii. Run (do not supply any password when prompted):

```
[root@replacement]$ ssh-keygen -t rsa
[root@replacement]$ ssh-keygen -t dsa
```

➢ iii. Propagate the generated public key to remote nodes by running:

```
dcli -l oracle -k -g /opt/oracle.SupportTools/dbs_group
```

- supply password for oracle user for other nodes when prompted

Copyright © 2014, Oracle and/or its affiliates. All rights reserved.

Prepare the Replacement Server:

SSH equivalence is for oracle user is needed before Extending Clusterware

- Set SSH equivalence. On **each** surviving node:

➢ i. switch to the oracle account:

```
[root@surviving]# su - oracle
```

➢ ii. Run:

```
[root@surviving]# dcli -k -c dmorldb06 -l oracle
```

- supply password for oracle user set on reimaged node when prompted

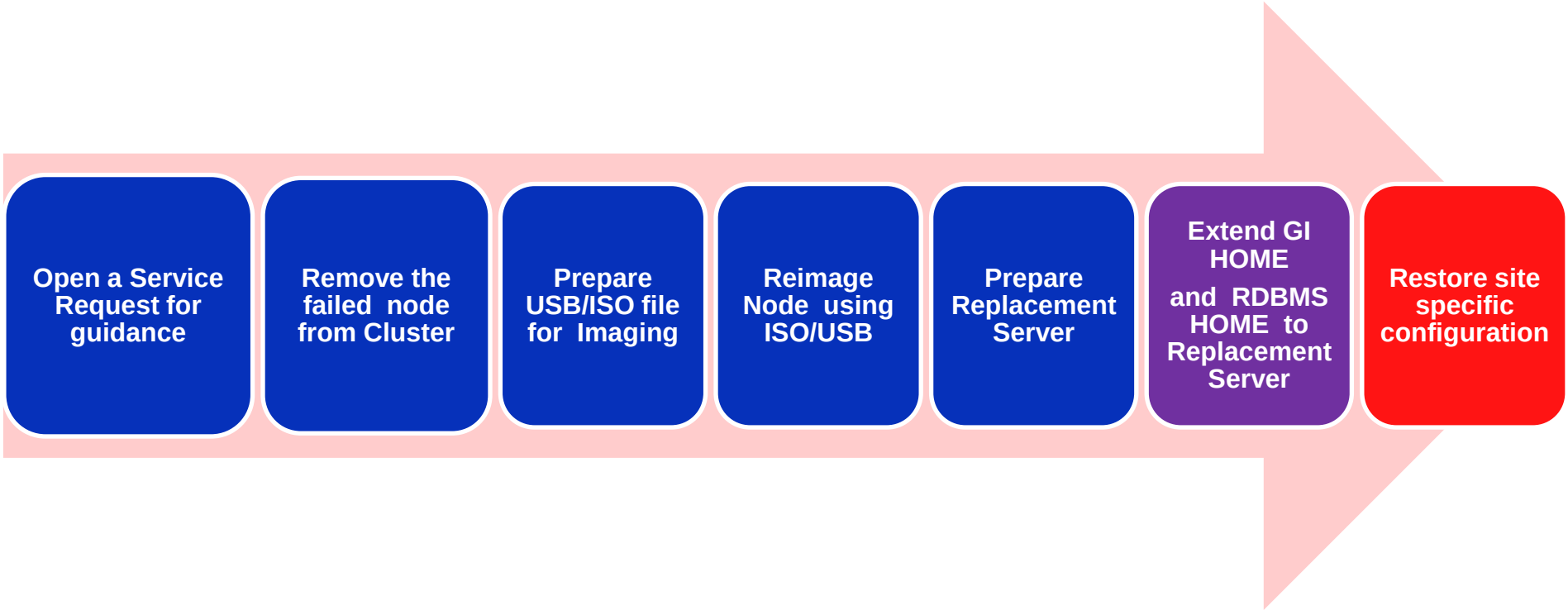
Copyright © 2014, Oracle and/or its affiliates. All rights reserved.

Demo – Prepare reimaged node

Copyright © 2014, Oracle and/or its affiliates. All rights reserved.

Overview Bare Metal Restore Procedure

The high level steps are as below



Copyright © 2014, Oracle and/or its affiliates. All rights reserved.

Add Replacement Server to Clusterware

Clone Oracle Grid Infra HOME (prechecks)

- Verify the hardware and operating system installations with the Cluster Verification Utility (CVU):

```
[oracle@surviving]$ cluvfy stage -post hwos -n dmorldb05,dmorldb06 –verbose
```

- Verify peer compatibility:

```
[oracle@surviving]$ cluvfy comp peer -refnode dmorldb05 -n dmorldb06 -orainv oinstall -osdba dba \
 | grep -B 3 -A 2 mismatched
```

- Perform requisite checks for node addition:

```
[oracle@surviving]$ cluvfy stage -pre nodeadd -n dmorldb06 -fixup –fixupdir /home/oracle/fixupdir
```


Copyright © 2014, Oracle and/or its affiliates. All rights reserved.

Add Replacement Server to Clusterware

Clone Oracle Grid Infra HOME

- Run addNode.sh from Surviving node:

```
[oracle@surviving]$ cd /u01/app/11.2.0/grid/oui/bin/
[oracle@surviving]$ ./addNode.sh -silent "CLUSTER_NEW_NODES={dmorldb06}"
"CLUSTER_NEW_VIRTUAL_HOSTNAMES={dmorl06-vip}“
```

- Run the orainstRoot.sh and root.sh scripts for the reimaged node when prompted by addNode.sh
- Re-enable the listener:

```
[root@replacement]# /u01/app/11.2.0/grid/bin/srvctl enable listener -l LISTENER -n dmorldb06
[root@replacement]# /u01/app/11.2.0/grid/bin/srvctl start listener -l LISTENER  -n dmorldb06
```


Copyright © 2014, Oracle and/or its affiliates. All rights reserved.

Add Replacement Server to Clusterware

Clone RDBMS HOME

- Add the replacement database server into the cluster

addNode.sh from Surviving node

```
[oracle@surviving]$ cd /u01/app/oracle/product/11.2.0/dbhome_1/oui/bin/
[oracle@surviving]$ ./addNode.sh -silent "CLUSTER_NEW_NODES={dmorldb06}
```

- Run script as root on reimaged node when prompted

```
[root@replacement]# /u01/app/oracle/product/11.2.0/dbhome_1/root.sh
```

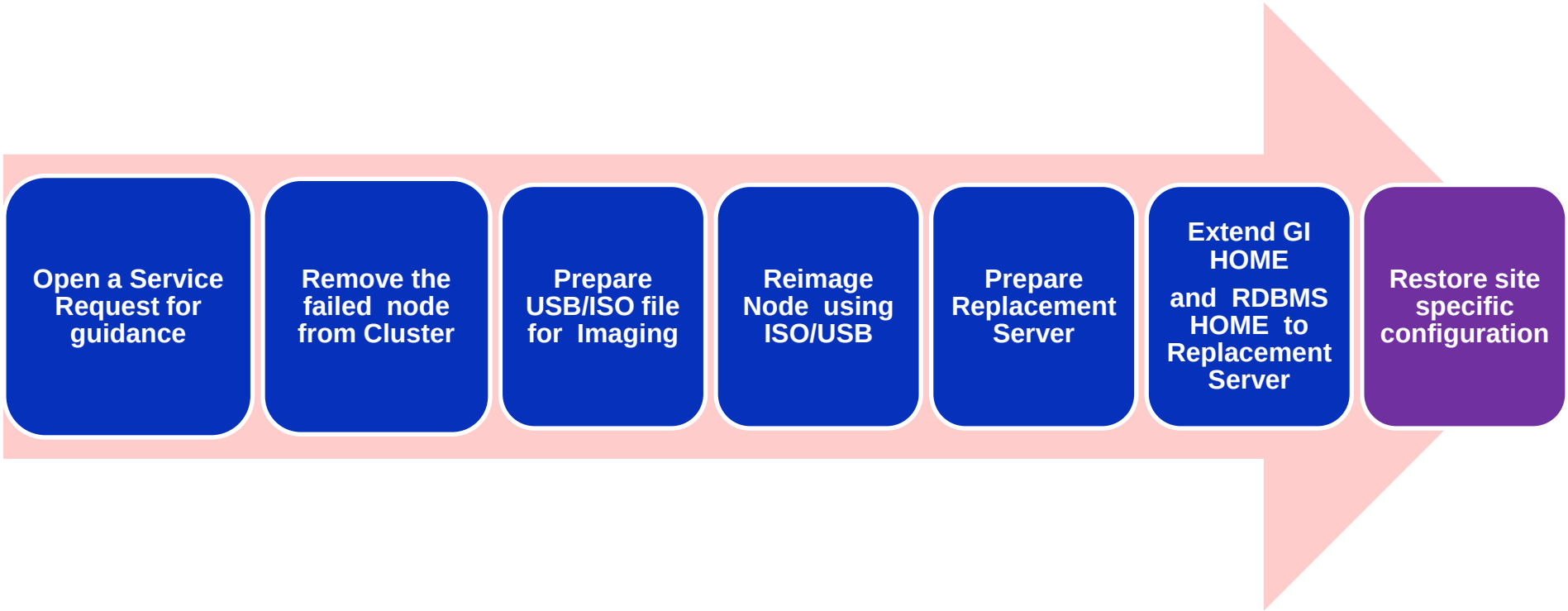
- Refer to the follow note

How to Add Node/Instance or Remove Node/Instance in 10gR2, 11gR1 and 11gR2 Oracle Clusterware and RAC (Doc ID 1332451.1)

Copyright © 2014, Oracle and/or its affiliates. All rights reserved.

Overview Bare Metal Restore Procedure

The high level steps are as below



Copyright © 2014, Oracle and/or its affiliates. All rights reserved.

Tip

- Backup each node using steps in Exadata Owner's Guide section 7.125 to avoid BMR

Copyright © 2014, Oracle and/or its affiliates. All rights reserved.

Summary

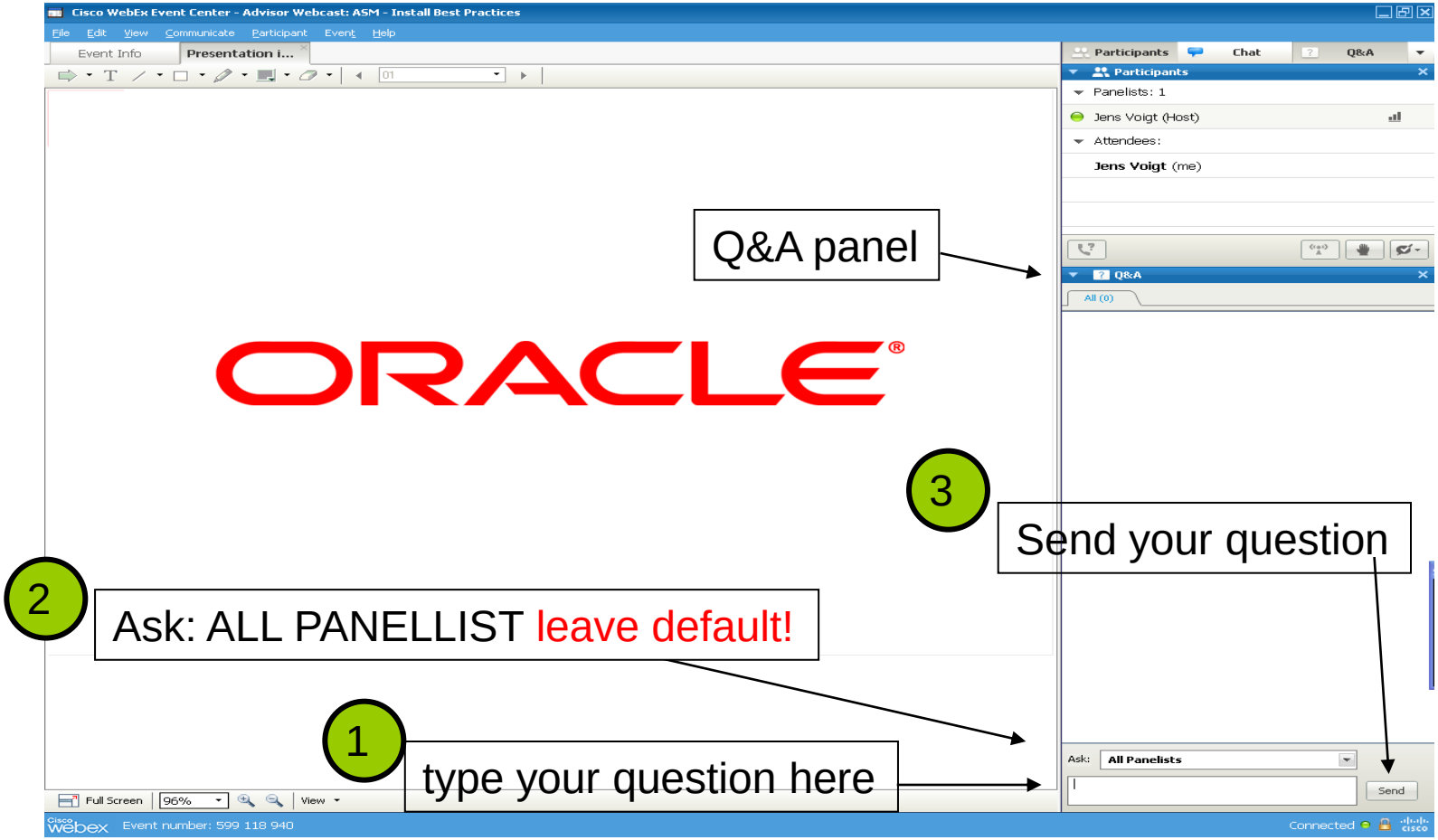
What we covered today?

- BMR is a complex procedure
- Preparation will depend on site specifics
- Sysadmin, DBA and Oracle Support involved

Copyright © 2014, Oracle and/or its affiliates. All rights reserved.

Questions & Answers

Copyright © 2014, Oracle and/or its affiliates. All rights reserved.



Copyright © 2014, Oracle and/or its affiliates. All rights reserved.

Further Info & Help

- Generic Advisor Webcast Note (Doc ID 740966.1)
- Oracle Engineered Systems Advisor Webcast Schedule and Archive recordings(Doc ID 1456202.1)
- DB Newsletter (Doc ID 1284265.1)
- MOS Community 中文社区=> **Exadata**

Copyright © 2014, Oracle and/or its affiliates. All rights reserved.

Oracle Advisor Webcast Program

- Current schedule
- Archived recordings
- Doc ID 740966.1

Copyright © 2014, Oracle and/or its affiliates. All rights reserved.

Oracle Engineered Systems Advisor Webcast Schedule and Archive recordings (Doc ID 1456202.1)

To Bottom

Modified: 26-Dec-2013 Type: ANNOUNCEMENT Status: PUBLISHED [EXTERNAL]

(1)

GET PROACTIVE

Get Proactive Product Page Index ▶

Stay Informed

Advisor Webcasts ▶

Proactive Best Practices ▶

Learn My Oracle Support ▶

Oracle Database Product News

Exadata

Proactive Social Media

Advisor Webcasts for Oracle Engineered Systems

Why attend the Oracle Advisor Webcast Program

The Oracle Advisor Webcast Program brings interactive expertise straight to your desktop using Oracle Web Conferencing technology, at no cost. This technology brings you and Oracle experts together to access information about support services, products, technologies, best practices and more.

Archives

Tabs to view ar… …mn Headers to sort …t.

Schedule

Archives

Current Schedule | Archived 2014 | Archived 2013 | Archived 2012 | Archived 2011

Product Family	Title	Date & Time	Register
数据库	高性能Exadata最佳实践 - Mandarin only	January 15, 2014, 15:00 Beijing	REGISTER
Exadata General	Oracle SuperCluster Best Practices	February 14, 2014, 08:00 PT	REGISTER

For more information regarding the advisor webcast program, please contact Oracle.

Copyright © 2014, Oracle and/or its affiliates. All rights reserved.

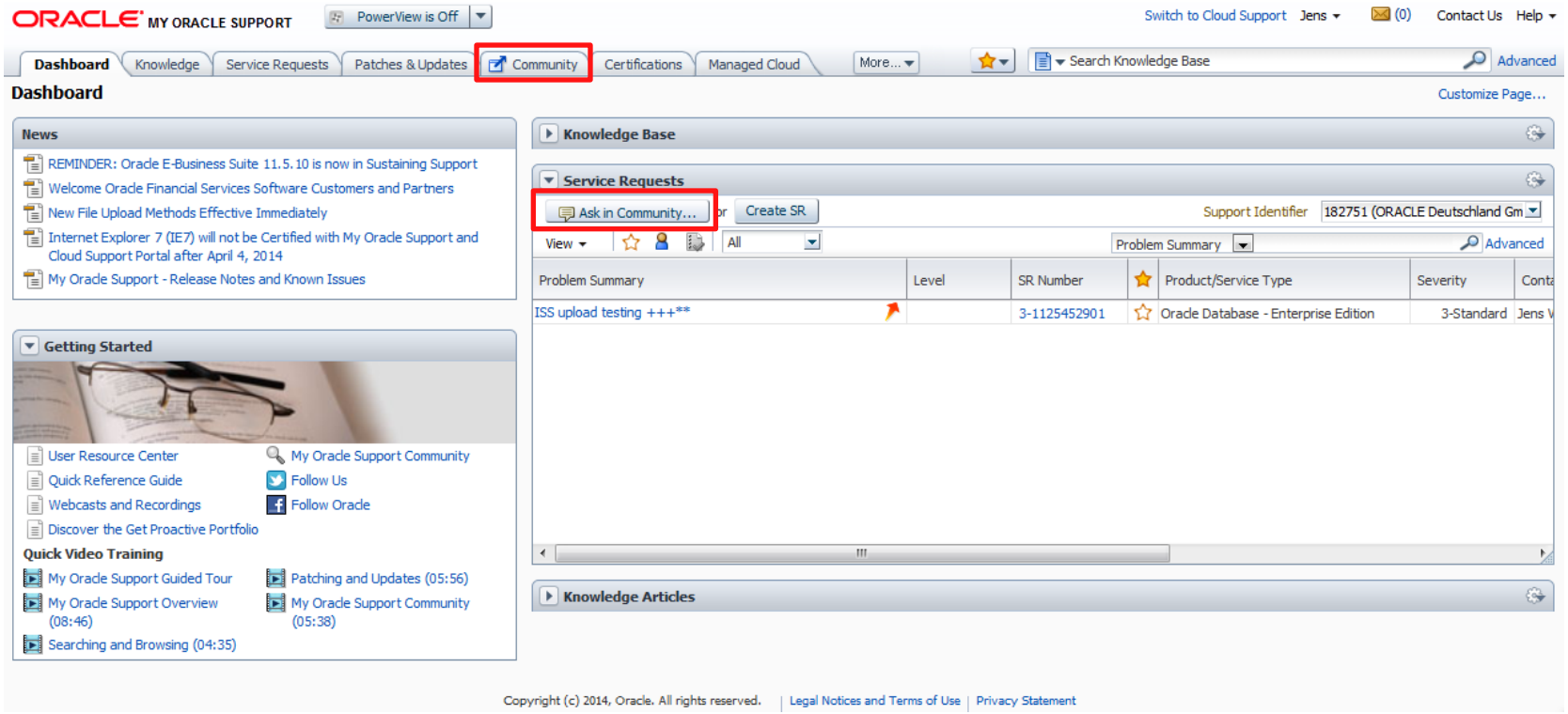
Copyright © 2014, Oracle and/or its affiliates. All rights reserved.

Learn More

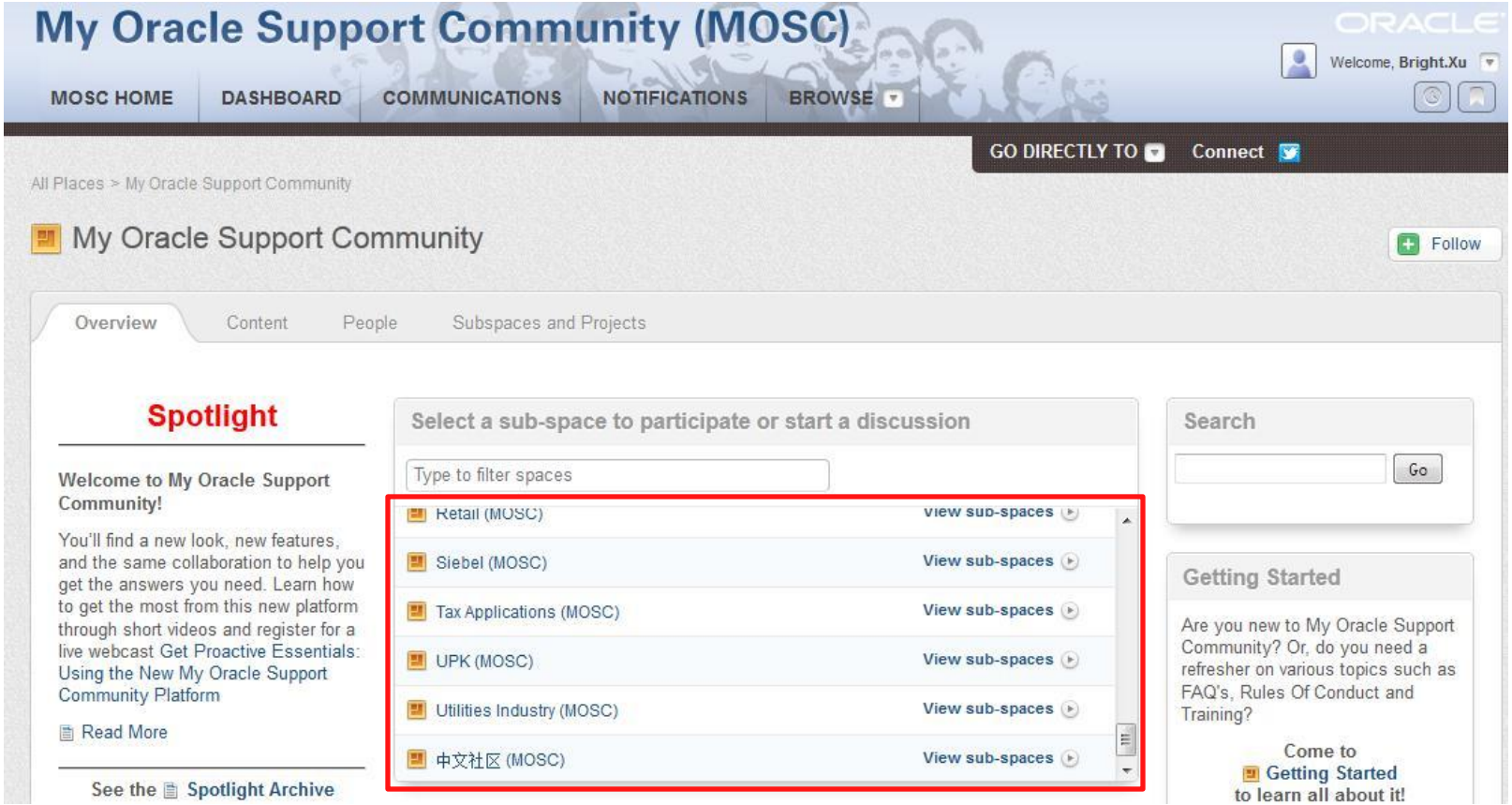
Available References and Resources to Get Proactive

- About Oracle Support Best Practices
  www.oracle.com/goto/proactivesupport
- Get Proactive in My Oracle Support
  https://support. oracle.com | Doc ID: 432.1
- Get Proactive Blog
  https://blogs.oracle.com/getproactive/
- Ask the Get Proactive Team
  get-proactive_ww@oracle.com

Copyright © 2014, Oracle and/or its affiliates. All rights reserved.



Copyright © 2014, Oracle and/or its affiliates. All rights reserved.



Copyright © 2014, Oracle and/or its affiliates. All rights reserved.

THANK YOU

Copyright © 2014, Oracle and/or its affiliates. All rights reserved.

Hardware and Software

ORACLE

Engineered to Work Together