

管信融合 RAC 安装操作手册

文件控制

文档分类

文档名称	
版本号码	
版本日期	
作者	

版本控制

版本	日期	备注
V1.0		第一版

修改记录

日期	作者	版本	修改记录
2015-5-26	杜诚文	V1.0	第一次创建

拷贝

拷贝No.	姓名	单位

审阅

职务	姓名	职位

批准

职务	签名	日期

目 录

1 安装环境规划.....	3
1.1 硬件环境.....	3
1.2 主机配置.....	3
1.3 数据库软件.....	3
1.4 安装目录规划.....	3
1.5 存储规划.....	3
1.6 网络规划.....	4
1.7 密码规划.....	4
2 主机硬件环境检查.....	5
2.1 查看主机配置.....	5
2.2 检查主机物理内存.....	7
2.3 检查主机交互空间.....	8
2.4 查看/tmp 的空间.....	8
2.5 查看磁盘空间.....	8
2.6 查看网络信息.....	8
2.7 查看操作系统架构.....	10
3 操作系统配置.....	12
3.1 操作系统版本要求.....	12
3.2 操作系统软件包.....	12
3.3 操作系统补丁包.....	13
3.4 调整操作系统参数.....	15
3.5 调整网络参数.....	16
3.6 调整异步 IO.....	18
3.7 创建 GRID,ORACLE 安装用户和组.....	18
3.8 配置 grid,oracle 用户 SSH 互信设置.....	22
4 配置共享存储.....	25
4.1 存储方式与文件类型对应关系.....	25
4.2 配置 ASM 使用的共享裸磁盘设备.....	25
5 安装 GRID 软件.....	30
5.1 解压缩软件.....	30
5.2 运行 runInstaller 脚本.....	30
5.3 完成 GI 安装.....	41
5.4 执行 GI 安装后校验.....	41
6 配置 ASM 磁盘组.....	43
7 安装配置 RAC 数据库.....	45
7.1 安装数据库软件.....	45
7.2 创建数据库.....	49
8 数据库补丁安装.....	56
8.1 Database 11.2.0.4 PSU.....	56
8.2 USLA heap fix.....	56
9 RAC 数据库日常管理命令.....	57
9.1 ORACLE RAC 自动启动.....	57
9.2 查看当前 RAC 服务状态.....	57
9.3 启动和停止 RAC 数据库.....	58
9.4 启动和停止 GI.....	58
10 参考文档.....	59
10.1 IBM 官方文档.....	59
10.2 ORACLE 官方文档.....	59
10.3 MOS 参考文档.....	59

1 安装环境规划

1.1 硬件环境

操作系统	AIX 6.1 TL6
主机	IBM P590
存储	DS8870

1.2 主机配置

CPU 数量	8
内存容量	16G
网卡数量	3
RAC 节点数	2

1.3 数据库软件

集群软件版本	Oracle Grid Infrastructure 11.2.0.4.0 AIX5L
数据库版本	Oracle Database 11.2.0.4.0 Enterprise Edition AIX5L
数据库补丁	
其他补丁	

1.4 安装目录规划

ORACLE_BASE	/oracle/app/oracle
ORACLE_HOME	/oracle/app/oracle/11.2/db_1
GRID_HOME	/oracle/app/grid/11.2
RAC 节点数	2

1.5 存储规划

LUN 编号	挂载磁盘	容量 (GB)	磁盘别名	用途
1	hdisk2	186		未使用
2	hdisk3	100		db01 oracle 安装目录
3	hdisk4	100		db02 oracle 安装目录
4	hdisk5	100	ASMDISK_GRID	OCR & VOTE
5	hdisk6	100	ASMDISK_DATA1	ASM 数据磁盘
6	hdisk7	100	ASMDISK_DATA2	ASM 数据磁盘
7	hdisk8	100	ASMDISK_DATA3	ASM 数据磁盘
8	hdisk9	100	ASMDISK_DATA4	ASM 数据磁盘
9	hdisk10	100	ASMDISK_DATA5	ASM 数据磁盘
10	hdisk11	100	ASMDISK_DATA6	ASM 数据磁盘
11	hdisk12	100	ASMDISK_DATA7	ASM 数据磁盘

12	hdisk13	100	ASMDISK_DATA8	ASM 数据磁盘
----	---------	-----	---------------	----------

1.6 网络规划

每台主机需要有两个网络接口，一个作为公有网卡，一个作为私有网卡，如下表：

主机名	IP 地址	网卡	说明
db01	10.217.229.144	en1	公有网络地址
db01-vip	10.217.229.145	en1	虚拟网络地址
db01-priv	192.168.0.132	en0	私有网络地址
db02	10.217.229.146	en1	公有网络地址
db02-vip	10.217.229.147	en1	虚拟网络地址
db02-priv	192.168.0.134	en0	私有网络地址
db0-scan	10.217.229.148	en1	服务地址
默认网关	10.217.229.129	公有网络必须配置，且各主机一致	

1.7 密码规划

以下密码为安装系统后的初始密码，数据库安装完成后，需进行更改

1.7.1 操作系统密码

主机	用户名	临时密码
xgbusidb1	root	root
xgbusidb1	grid	Qwer4321
xgbusidb1	oracle	Qwer4321
xgbusidb2	root	root
xgbusidb2	grid	Qwer4321
xgbusidb2	oracle	Qwer4321

1.7.2 数据库密码

数据库名	主机名	用户名	临时密码
gxoradb	xgbusidb1	sys/system	Tlh_150202
	xgbusidb2	sys/system	Tlh_150202
	主机名	用户名	临时密码
ASM	xgbusidb1	sys/system	Tlh_150202
	xgbusidb2	sys/system	Tlh_150202

2 主机硬件环境检查

2.1 查看主机配置

查看当前主机配置信息，包含 CPU、内存、网卡、磁盘等信息。

2.1.1 prtconf

```
# prtconf
System Model: IBM,9119-590
Machine Serial Number: 83DA24D
Processor Type: PowerPC_POWER5
Processor Implementation Mode: POWER 5
Processor Version: PV_5_3
Number Of Processors: 4
Processor Clock Speed: 2102 MHz
CPU Type: 64-bit
Kernel Type: 64-bit
LPAR Info: 10 OAP1_8_
Memory Size: 16384 MB
Good Memory Size: 16384 MB
Platform Firmware level: SF240_418
Firmware Version: IBM,SF240_418
Console Login: enable
Auto Restart: true
Full Core: false

Network Information
  Host Name: localhost
  IP Address: 10.218.64.83
  Sub Netmask: 255.255.255.192
  Gateway: 10.218.64.65
  Name Server:
  Domain Name:

Paging Space Information
  Total Paging Space: 16384MB
  Percent Used: 1%

Volume Groups Information
=====
rootvg:
PV_NAME      PV STATE      TOTAL PPs   FREE PPs   FREE DISTRIBUTION
hdisk0       active        546         330        109..03..00..109..109
=====

INSTALLED RESOURCE LIST

The following resources are installed on the machine.
+/- = Added or deleted from Resource List.
*   = Diagnostic support not available.

Model Architecture: chrp
Model Implementation: Multiple Processor, PCI bus

+ sys0                                     System Object
+ sysplanar0                             System Planar
* vio0                                   Virtual I/O Bus
* vsa0      U9119.590.83DA24D-V10-C0      LPAR Virtual Serial Adapter
* vty0      U9119.590.83DA24D-V10-C0-L0    Asynchronous Terminal
* pci3      U5791.001.9920Z7D-P2           PCI Bus
* pci8      U5791.001.9920Z7D-P2           PCI Bus
+ ent2      U5791.001.9920Z7D-P2-C06-T1    2-Port 10/100/1000 Base-TX PCI-X
Adapter (14108902)
```

+ ent3	U5791. 001. 9920Z7D-P2-C06-T2	2-Port 10/100/1000 Base-TX PCI-X
Adapter (14108902)		
* pci9	U5791. 001. 9920Z7D-P2	PCI Bus
+ fcs2	U5791. 001. 9920Z7D-P2-C07-T1	FC Adapter
* fcnet2	U5791. 001. 9920Z7D-P2-C07-T1	Fibre Channel Network Protocol
Device		
+ fscsi2	U5791. 001. 9920Z7D-P2-C07-T1	FC SCSI I/O Controller Protocol
Device		
* hdisk2	U5791. 001. 9920Z7D-P2-C07-T1-W5005076305031788-L4010400200000000	IBM MPIO FC 2107
* hdisk3	U5791. 001. 9920Z7D-P2-C07-T1-W5005076305031788-L4010400300000000	IBM MPIO FC 2107
* hdisk4	U5791. 001. 9920Z7D-P2-C07-T1-W5005076305031788-L4010400400000000	IBM MPIO FC 2107
* hdisk5	U5791. 001. 9920Z7D-P2-C07-T1-W5005076305031788-L4010400500000000	IBM MPIO FC 2107
* hdisk6	U5791. 001. 9920Z7D-P2-C07-T1-W5005076305031788-L4010400600000000	IBM MPIO FC 2107
* hdisk7	U5791. 001. 9920Z7D-P2-C07-T1-W5005076305031788-L4010400700000000	IBM MPIO FC 2107
* hdisk8	U5791. 001. 9920Z7D-P2-C07-T1-W5005076305031788-L4010400800000000	IBM MPIO FC 2107
* hdisk9	U5791. 001. 9920Z7D-P2-C07-T1-W5005076305031788-L4010400900000000	IBM MPIO FC 2107
* hdisk10	U5791. 001. 9920Z7D-P2-C07-T1-W5005076305031788-L4011400200000000	IBM MPIO FC 2107
* hdisk11	U5791. 001. 9920Z7D-P2-C07-T1-W5005076305031788-L4011400300000000	IBM MPIO FC 2107
* hdisk12	U5791. 001. 9920Z7D-P2-C07-T1-W5005076305031788-L4011400400000000	IBM MPIO FC 2107
* hdisk13	U5791. 001. 9920Z7D-P2-C07-T1-W5005076305031788-L4011400500000000	IBM MPIO FC 2107
* hdisk14	U5791. 001. 9920Z7D-P2-C07-T1-W5005076305031788-L4011400600000000	IBM MPIO FC 2107
* hdisk15	U5791. 001. 9920Z7D-P2-C07-T1-W5005076305031788-L4011400700000000	IBM MPIO FC 2107
* hdisk16	U5791. 001. 9920Z7D-P2-C07-T1-W5005076305031788-L4011400800000000	IBM MPIO FC 2107
* sfwcomm2	U5791. 001. 9920Z7D-P2-C07-T1-W0-L0	Fiber Channel Storage Framework
Comm		
+ fcs3	U5791. 001. 9920Z7D-P2-C07-T2	FC Adapter
+ fscsi3	U5791. 001. 9920Z7D-P2-C07-T2	FC SCSI I/O Controller Protocol
Device		
* sfwcomm3	U5791. 001. 9920Z7D-P2-C07-T2-W0-L0	Fiber Channel Storage Framework
Comm		
* fcnet3	U5791. 001. 9920Z7D-P2-C07-T2	Fibre Channel Network Protocol
Device		
* pci2	U5791. 001. 9920Z7D-P2	PCI Bus
* pci6	U5791. 001. 9920Z7D-P2	PCI Bus
+ fcs0	U5791. 001. 9920Z7D-P2-C08-T1	FC Adapter
* fcnet0	U5791. 001. 9920Z7D-P2-C08-T1	Fibre Channel Network Protocol
Device		
+ fscsi0	U5791. 001. 9920Z7D-P2-C08-T1	FC SCSI I/O Controller Protocol
Device		
* sfwcomm0	U5791. 001. 9920Z7D-P2-C08-T1-W0-L0	Fiber Channel Storage Framework
Comm		
+ fcs1	U5791. 001. 9920Z7D-P2-C08-T2	FC Adapter
+ fscsi1	U5791. 001. 9920Z7D-P2-C08-T2	FC SCSI I/O Controller Protocol
Device		
* sfwcomm1	U5791. 001. 9920Z7D-P2-C08-T2-W0-L0	Fiber Channel Storage Framework
Comm		
* fcnet1	U5791. 001. 9920Z7D-P2-C08-T2	Fibre Channel Network Protocol
Device		
* pci7	U5791. 001. 9920Z7D-P2	PCI Bus
+ ent0	U5791. 001. 9920Z7D-P2-C09-T1	2-Port 10/100/1000 Base-TX PCI-X
Adapter (14108902)		
+ ent1	U5791. 001. 9920Z7D-P2-C09-T2	2-Port 10/100/1000 Base-TX PCI-X
Adapter (14108902)		
* pci1	U5791. 001. 99B0FBT-P2	PCI Bus
* pci5	U5791. 001. 99B0FBT-P2	PCI Bus
+ scsi0	U5791. 001. 99B0FBT-P2-T5	Wide/Ultra-3 SCSI I/O Controller
+ hdisk0	U5791. 001. 99B0FBT-P2-T5-L8-L0	16 Bit LVD SCSI Disk Drive (146800 MB)
+ hdisk1	U5791. 001. 99B0FBT-P2-T5-L9-L0	16 Bit LVD SCSI Disk Drive (146800 MB)
+ ses0	U5791. 001. 99B0FBT-P2-T5-L15-L0	SCSI Enclosure Services Device
* pci0	U5791. 001. 99B0FBT-P1	PCI Bus
* pci4	U5791. 001. 99B0FBT-P1	PCI Bus
+ sisscsia0	U5791. 001. 99B0FBT-P1-C10	PCI-XDDR Dual Channel Ultra320 SCSI
Adapter		
+ scsi1	U5791. 001. 99B0FBT-P1-C10-T1	PCI-X Dual Channel Ultra320 SCSI
Adapter bus		
+ rmt0	U5791. 001. 99B0FBT-P1-C10-T1-L0-L0	LVD SCSI 4mm Tape Drive (36000 MB)
+ cd0	U5791. 001. 99B0FBT-P1-C10-T1-L5-L0	16 Bit LVD SCSI DVD-RAM Drive (4700 MB)
+ scsi2	U5791. 001. 99B0FBT-P1-C10-T2	PCI-X Dual Channel Ultra320 SCSI

Adapter bus	
+ L2cache0	L2 Cache
+ mem0	Memory
+ proc0	Processor
+ proc2	Processor
+ proc4	Processor
+ proc6	Processor

2.1.2 lparstat -i

lparstat -i

```

Node Name                : localhost
Partition Name           : OAP1_8_
Partition Number         : 10
Type                     : Dedicated-SMT
Mode                     : Capped
Entitled Capacity        : 4.00
Partition Group-ID       : 32778
Shared Pool ID           : -
Online Virtual CPUs      : 4
Maximum Virtual CPUs     : 8
Minimum Virtual CPUs     : 2
Online Memory            : 16384 MB
Maximum Memory           : 32768 MB
Minimum Memory           : 4096 MB
Variable Capacity Weight : -
Minimum Capacity         : 2.00
Maximum Capacity         : 8.00
Capacity Increment       : 1.00
Maximum Physical CPUs in system : 64
Active Physical CPUs in system : 32
Active CPUs in Pool      : -
Shared Physical CPUs in system : -
Maximum Capacity of Pool : -
Entitled Capacity of Pool : -
Unallocated Capacity     : -
Physical CPU Percentage   : 100.00%
Unallocated Weight       : -
Memory Mode              : Dedicated
Total I/O Memory Entitlement : -
Variable Memory Capacity Weight : -
Memory Pool ID           : -
Physical Memory in the Pool : -
Hypervisor Page Size     : -
Unallocated Variable Memory Capacity Weight : -
Unallocated I/O Memory entitlement : -
Memory Group ID of LPAR  : -
Desired Virtual CPUs     : 4
Desired Memory           : 16384 MB
Desired Variable Capacity Weight : -
Desired Capacity        : 4.00
Target Memory Expansion Factor : -
Target Memory Expansion Size : -
Power Saving Mode        : -

```

2.2 检查主机物理内存

GI 安装需要所需内存最小为 2.5 GB。

DB 安装需要所需内存最小为 1.5 GB。

```
# /usr/sbin/lsattr -E -l sys0 -a realmem
realmem 16777216 Amount of usable physical memory in Kbytes False
```

2.3 检查主机交互空间

交换空间要求设置一般为内存的 2 倍，但不超过 32GB，磁盘须有足够的空间。

```
# /usr/sbin/lspv -a
```

Page Space	Physical Volume	Volume Group	Size	%Used	Active	Auto	Type	Chksum
hd6	hdisk0	rootvg	16384MB	1	yes	yes	lv	0

2.4 查看/tmp 的空间

查看/tmp 空间大小，/tmp 可用空间不能小于 1GB。

```
# df -g /tmp
```

Filesystem	GB blocks	Free	%Used	Iused	%Iused	Mounted on
/dev/hd3	10.00	9.99	1%	87	1%	/tmp

如果 tmp 空间不足，可以通过以下方式，指定新的 temp 目录：

```
export TEMP=/new_tmp
```

```
export TMPDIR=/new_tmp
```

2.5 查看磁盘空间

GI 软件安装需要的最小空间为 13GB

DB 软件安装需要的最小空间为 7.5GB

```
# df -g
```

Filesystem	GB blocks	Free	%Used	Iused	%Iused	Mounted on
/dev/hd4	5.50	5.24	5%	10544	1%	/
/dev/hd2	5.25	3.01	43%	50933	7%	/usr
/dev/hd9var	5.50	5.11	8%	8455	1%	/var
/dev/hd3	5.50	3.72	33%	1871	1%	/tmp
/dev/hd1	5.50	5.50	1%	120	1%	/home
/dev/hd11admin	5.25	5.25	1%	5	1%	/admin
/proc	-	-	-	-	-	/proc
/dev/hd10opt	5.75	5.51	5%	8755	1%	/opt
/dev/livedump	0.25	0.25	1%	4	1%	/var/adm/ras/livedump
/dev/fs1v01	99.00	70.83	29%	63831	1%	/oracle

2.6 查看网络信息

2.6.1 查看主机名配置

```
# more /etc/hosts
```

```

127.0.0.1      localhost      # loopback (lo0) name/address

10.217.229.149 xgbusidbl
10.217.229.150 xgbusidbl-vip
192.168.12.149 xgbusidbl-priv

10.217.229.151 xgbusidb2
10.217.229.152 xgbusidb2-vip
192.168.12.151 xgbusidb2-priv

10.217.229.153 xgbusidb-scan

```

2.6.2 查看网卡信息

`/usr/sbin/ifconfig -a`

```

en0:
flags=5e080863, c0<UP, BROADCAST, NOTRAILERS, RUNNING, SIMPLEX, MULTICAST, GROUPRT, 64BIT, CHECKSUM_OFFLOA
D(ACTIVE), PSEG, LARGESEND, CHAIN>
    inet 10.217.229.151 netmask 0xffffffe0 broadcast 10.217.229.159
    tcp_sendspace 131072 tcp_recvspace 65536 rfc1323 0
en2:
flags=5e080863, c0<UP, BROADCAST, NOTRAILERS, RUNNING, SIMPLEX, MULTICAST, GROUPRT, 64BIT, CHECKSUM_OFFLOA
D(ACTIVE), PSEG, LARGESEND, CHAIN>
    inet 10.10.0.68 netmask 0xfffff00 broadcast 10.10.0.255
    tcp_sendspace 131072 tcp_recvspace 65536 rfc1323 0
en3:
flags=5e080863, c0<UP, BROADCAST, NOTRAILERS, RUNNING, SIMPLEX, MULTICAST, GROUPRT, 64BIT, CHECKSUM_OFFLOA
D(ACTIVE), PSEG, LARGESEND, CHAIN>
    inet 192.168.12.151 netmask 0xfffff00 broadcast 192.168.12.255
    tcp_sendspace 131072 tcp_recvspace 65536 rfc1323 0
lo0:
flags=e08084b, c0<UP, BROADCAST, LOOPBACK, RUNNING, SIMPLEX, MULTICAST, GROUPRT, 64BIT, LARGESEND, CHAIN>
    inet 127.0.0.1 netmask 0xff000000 broadcast 127.255.255.255
    inet6 ::1%1/0
    tcp_sendspace 131072 tcp_recvspace 131072 rfc1323 1

```

网卡信息获取脚本:

```

for i in en0 en2
do
    echo $i

    for attribut in netaddr netmask broadcast state
    do
        lsattr -El $i -a $attribut
    done
done

```

```

en0
netaddr 10.218.64.83 Internet Address True
netmask 255.255.255.192 Subnet Mask True
broadcast Broadcast Address True
state up Current Interface Status True

en2
netaddr 10.10.0.58 Internet Address True
netmask 255.255.255.0 Subnet Mask True
broadcast Broadcast Address True
state up Current Interface Status True

```

2.6.3 查看路由信息

netstat -rn

Routing tables							
Destination	Gateway	Flags	Refs	Use	If	Exp	Groups
Route Tree for Protocol Family 2 (Internet):							
default	10.217.229.129	UG	3	5623	en0	-	-
10.10.0.0	10.10.0.68	UHSb	0	0	en2	-	- =>
10.10/24	10.10.0.68	U	1	70354	en2	-	-
10.10.0.68	127.0.0.1	UGHS	0	230	lo0	-	-
10.10.0.255	10.10.0.68	UHSb	0	4	en2	-	-
10.217.229.128	10.217.229.151	UHSb	0	0	en0	-	- =>
10.217.229.128/27	10.217.229.151	U	1	2	en0	-	-
10.217.229.151	127.0.0.1	UGHS	0	80	lo0	-	-
10.217.229.159	10.217.229.151	UHSb	0	0	en0	-	-
127/8	127.0.0.1	U	17	595890	lo0	-	-
192.168.12.0	192.168.12.151	UHSb	0	0	en3	-	- =>
192.168.12/24	192.168.12.151	U	0	36	en3	-	-
192.168.12.151	127.0.0.1	UGHS	0	83	lo0	-	-
192.168.12.255	192.168.12.151	UHSb	0	0	en3	-	-
Route Tree for Protocol Family 24 (Internet v6):							
::1%1	::1%1	UH	1	174	lo0	-	-



说明

- 每台主机的网卡用途与主机名类型必须匹配(如 eth0 统一对应公有网卡)。
- 公有网络必须支持 TCP/IP 协议。
- 公有 IP，虚拟 IP，SCAN IP 地址在同一网段。
- 安装前，默认网关地址必须能够正常访问(通过 route 命令查看)。
- 私有网络必须支持 UDP 协议。
- 私有网络必须与公有网络网段隔离。
- 私有网络必须通过交换机连接，不能使用心跳线方式。
- 私有 IP 地址范围为：10.*.* 或者 192.168.*.*。
- 如果使用 DNS，可以配置 3 个静态的 IP 作为 SCAN 地址。
- SCAN 地址名称中不能有“_”，可以使用“-”。
- SCAN 地址对应的主机名长度不能超过 16 个字符。
- 通过主机名解析的方式，只能配置一个 SCAN IP。
- 不启用 GNS 的情况下，公有 IP 和虚拟 IP 必须为静态地址。
- 从 11.2.0.2 开始，oracle 提供网络冗余特性，不需要再进行操作系统配置。

2.7 查看操作系统架构

```
# /usr/bin/getconf HARDWARE_BITMODE
64
```

```
# bootinfo -k
```

```
64
```

```
# oslevel -s
```


3 操作系统配置

3.1 操作系统版本要求

AIX 7.1 TL 0 SP1 (" 7100-00-01") 64-bit kernel
AIX 6.1 TL 04 SP1 ("6100-02-01"), 64-bit kernel
AIX 5L V5.3 TL 09 SP1 ("5300-09-01"), 64 bit kernel or later

```
# oslevel -r  
6100-06
```

3.2 操作系统软件包

3.2.1 AIX 6.1

以下软件包是 AIX 6.1 所必需安装的软件包。

```
bos.adt.base  
bos.adt.lib  
bos.adt.libm  
bos.perf.libperfstat 6.1.2.1 or later  
bos.perf.perfstat  
bos.perf.proctools  
rsct.basic.rte  
rsct.compat.clients.rte  
xlc.aix61.rte.10.1.0.0 or later  
xlc.rte:10.1.0.0 or later  
  
# lsllpp -l |grep -E \  
'bos.adt.base|bos.adt.lib|bos.adt.libm|bos.perf.libperfstat  
|bos.perf.perfstat|bos.perf.proctools|rsct.basic.rte|rsct.compat.clie  
nts.rte|xlc.aix61.rte |xlc.rte'
```

检测 AIX 软件包:

```
# lsllpp -l | grep bos.adt.base  
bos.adt.base          6.1.6.1  COMMITTED  Base Application Development  
bos.adt.base          6.1.6.1  COMMITTED  Base Application Development  
# lsllpp -l | grep bos.adt.lib  
bos.adt.lib           6.1.2.0  COMMITTED  Base Application Development  
bos.adt.libm          6.1.6.0  COMMITTED  Base Application Development  
bos.adt.libmio        6.1.6.0  COMMITTED  Modular IO Library  
# lsllpp -l | grep bos.perf.libperfstat  
bos.perf.libperfstat  6.1.6.1  COMMITTED  Performance Statistics Library  
bos.perf.libperfstat  6.1.6.1  COMMITTED  Performance Statistics Library  
# lsllpp -l | grep bos.perf.perfstat  
bos.perf.perfstat     6.1.6.0  COMMITTED  Performance Statistics  
bos.perf.perfstat     6.1.6.0  COMMITTED  Performance Statistics  
# lsllpp -l | grep bos.perf.proctools  
bos.perf.proctools    6.1.6.0  COMMITTED  Proc Filesystem Tools  
# lsllpp -l | grep xlc.aix61  
xlc.aix61.rte         11.1.0.1  COMMITTED  XL C/C++ Runtime for AIX 6.1  
# lsllpp -l | grep xlc.rte  
xlc.rte               11.1.0.1  COMMITTED  XL C/C++ Runtime
```

3.2.2 AIX 7.1

以下软件包是 AIX 7.1 所必需安装的软件包。

```
bos.adt.base
bos.adt.lib
bos.adt.libm
bos.perf.libperfstat
bos.perf.perfstat
bos.perf.proctools
xlc.rte.11.1.0.2 or later
```

3.2.3 HACMP (不可以安装)

安装 11G RAC 时，不得安装其他第三方集群软件。因此必须卸载 HACMP。如果有检测到 HACMP 集群，联系操作系统管理员卸载该集群。

1. 检查集群配置是否存在

```
# cat /etc/objrepos/HACMPnode
```

2. 检查集群是否运行

```
# odmget HACMPcluster
```

3. 检查其他集群信息

```
# cat /etc/objrepos/HACMPadapter
# cat /etc/objrepos/HACMPcluster
```

3.3 操作系统补丁包

3.3.1 AIX 6.1

以下补丁是 AIX 6.1 所需的补丁。

```
IZ41855
IZ51456
IZ52319
IZ97457
IZ89165
```

These 6.1 fixes are already present in the following TL levels:

```
AIX 6.1 TL-02 SP-04 and later
AIX 6.1 TL-03 SP-02 and later
AIX 6.1 TL-04
```

```
# instfix -i -k "IZ41855 IZ51456 IZ52319 IZ97457 IZ89165"
```

当前系统版本为 AIX 6.1 TL-06， 不需要安装补丁。

3.3.2 AIX 7.1

以下补丁是 AIX 7.1 所需的补丁。

```
IZ87216
IZ87564
IZ89165
IZ97035
```

AIX 7.1 TL01 SP01
IV09541

AIX7.1 TL01
IV23859
IV21116

AIX71 TL02
IV21878

```
# /usr/sbin/instfix -ik "IZ87216"
# /usr/sbin/instfix -ik "IZ87564"
# /usr/sbin/instfix -ik "IZ89165"
# /usr/sbin/instfix -ik "IZ97035"
# /usr/sbin/instfix -ik "IV09541"
# /usr/sbin/instfix -ik "IV23859"

# /usr/sbin/instfix -ik "IV21116"
  There was no data for IV21116 in the fix database.
# /usr/sbin/instfix -ik "IV21878"
  All filesets for IV21878 were found.
```

3.3.3 其他建议补丁

1. Paging space growth leads to node failure/eviction:

64K paging taking place when available system RAM exists, the fix will avoid unexpected paging space growth and node failure. Below is a matrix of APAR for various TL/SP level

6100 TL5	6100-05	IZ71603
6100 TL4 SP4	6100-04-04-1014	IZ71191
6100 TL3 SP4	6100-03-04-1014	IZ72031
6100 TL2 SP7	6100-02-07-1014	IZ71850
6100 TL1 SP8	6100-01-08-1014	IZ71987
5300 TL12	5300-12	IZ71460
5300 TL11 SP4	5300-11-04-1015	IZ73687
5300 TL10 SP4	5300-10-04-1015	IZ73754
5300 TL9 SP7	5300-09-07-1015	IZ73864
5300 TL8 SP10	5300-08-10-1015	IZ67445

2. gc block lost or IPC send timeout or instance eviction

VIOS Server will not forward traffic from its VIO Clients to the external network, interrupts do not reach the trunk adapter, the fix will avoid SEA/VIO client hang. Below is a matrix of APAR for various TL/SP level

7100 TL0 SP3	7100-00-03-1115	IZ97035
6100 TL6 SP5	6100-06-05-1115	IZ96155
6100 TL5 SP6	6100-05-06-1119	IZ97457
6100 TL4 SP10	6100-04-10-1119	IZ97605
5300 TL12 SP4	5300-12-04-1119	IZ98126
5300 TL11 SP7	5300-11-07-1119	IZ98424

3. Other kernel hang fix

- * IZ91983 lockl performance issue, hang
- * IV04047: shlap64 unable to process Oracle request leading to kernel hang

4. Excessive CPU usage in LPAR in shared processor mode

If LPAR is in shared processor mode, without the following fix, LPAR may see excessive CPU usage:

APARS for WAITPROC IDLE LOOPING CONSUMES CPU:

```
IV01111 AIX 6.1 TL05 if before SP08 (fixed in SP08)
IV06197 AIX 6.1 TL06 if before SP07 (fixed in SP07)
IV10172 AIX 6.1 TL07 if before SP02 (fixed in SP02)
IV09133 AIX 7.1 TL00 if before SP05 (fixed in SP05)
IV10484 AIX 7.1 TL01 if before SP02 (fixed in SP02)
```

This problem can effect POWER7 systems running any level of Ax720 firmware prior to Ax720_101. But it is recommended to update to the latest available firmware. If required, AIX and Firmware fixes can be obtained from IBM Support Fix Central:
<http://www-933.ibm.com/support/fixcentral/main/System+p/AIX>

```
5. Crash in netinfo_unixdomnlist while running netstat
6100 TL6 SP6 6100-06-06-1140 IZ97166
6100 TL5 SP7 6100-05-07-1140 IZ97353
6100 TL4 SP11 6100-04-11-1140 IV00634
```

3.4 调整操作系统参数

`smit chgsys`

3.4.1 修改参数

```
vmo -p -o minperm%=3;
vmo -p -o maxperm%=90;
vmo -p -o maxclient%=90;
vmo -p -o lru_file_repage=0;
vmo -p -o strict_maxperm=0;
vmo -p -o strict_maxclient=1;
vmo -p -o page_steal_method=1;
chdev -l sys0 -a 'minpout=4096 maxpout=8193';
chdev -l sys0 -a maxuproc=16384
chdev -l sys0 -a ncargs='128'
vmo -p -o maxpin%=80
vmo -p -o v_pinshm=1
ioo -p -o aio_maxreqs=131072

schedo -o vpm_xvcpus=-1
```

3.4.2 参数说明

maxperm	用于缓存非计算性分页的最大内存量
minperm	非计算性分页的最低内存量
maxclient	缓存非计算性客户端分页的最大内存量的限制
strict_maxperm	AIX 内核允许非计算性分页的最大内存量 <code>maxperm%</code> 限制的范围
strict_maxclient	对 <code>maxclient%</code> 进行硬性限制, AIX 内核不允许非计算性的客户端文件缓存超出 <code>maxclient</code> 限制的范围
lru_file_repage	AIX 内核始终选择非计算性的分页进行操作
page_steal_method	如果设置 <code>page_steal_method=0</code> , 在寻找空闲页面时 LRUD 不得不扫描它们控制的所有内存页面, 尽管很可能只释放持久页面。 如果设置 <code>page_steal_method=1</code> , LRUD 会改用基于列表的页面管理方案。这意味着 LRUD 把内存划分为一个持久页面列表和一个工作存储页面列表。
maxpout	当线程从挂起状态唤起时指定最少计划页数

minpout	当线程被挂起前指定能被计划页数
	maxpout 和 minpout 可以帮助防止对文件执行顺序写入操作的线程占用过多的系统资源
maxuproc	最大用户进程数
ncargs	命令行的最大长度
maxpin	表示支持的最大的可 pin 住内存的比例
v_pinsh	表示 aix 将支持 pin 住内存, 默认是 0.
lgpg_size	large page 大小
aio_maxreqs	系统最大允许的 AIO 请求数量。5L 和 6L 默认为 64K, 7L 默认为 128K。
vpm_xvcpus	虚拟处理器管理功能, -1 表示禁用虚拟处理器管理功能。

3.4.3 参数检查

```
vmo -L minperm%
vmo -L maxperm%
vmo -L maxclient%
vmo -L lru_file_repage
vmo -L strict_maxperm
vmo -L strict_maxclient
vmo -a | grep page_steal_method
lsattr -E -l sys0 -a maxuproc
lsattr -E -l sys0 -a ncargs
vmo -L maxpin%
vmo -L v_pinshm
vmo -a | grep lgpg_size
vmo -a | grep lgpg_regions
ioo -o aio_maxreqs
schedo -o vpm_xvcpus
```

3.5 调整网络参数

3.5.1 修改网卡参数

```
no -p -o tcp_sendspace=1048576
no -p -o tcp_recvspace=1048576
no -p -o udp_sendspace=1048576
no -p -o udp_recvspace=1048576
no -o rfc1323=1
no -p -o sb_max=4194304
no -r -o ipqmaxlen=512
no -p -o tcp_ephemeral_low=9000
no -p -o tcp_ephemeral_high=65500
no -p -o udp_ephemeral_low=9000
no -p -o udp_ephemeral_high=65500
```

3.5.1.1 判断是否系统是否为兼容模式

```
# /usr/sbin/lsattr -E -l sys0 -a pre520tune
```

```
pre520tune disable Pre-520 tuning compatibility mode True
```

属于非兼容模式。

3.5.1.2 非兼容模式参数修改

```
# /usr/sbin/no -r -o ipqmaxlen=512 #该参数需要重启
```

其他参数:

```
# /usr/sbin/no -p -o parameter=value
```

3.5.2 修改网络配置文件

1. 修改配置文件/etc/resolv.conf 后,添加

```
vi /etc/resolv.conf
```

```
options timeout:1
```

```
options retry:1
```

2. 检查配置文件/etc/netsvc.conf

```
cat /etc/netsvc.conf|grep hosts
```

检查是否存在该行

```
hosts = local,bind
```

3.5.3 修改 ntp 配置

```
# ntpq -p
```

调整 NTP 使用 slew 模式:

```
# vi /etc/rc.tcpip
```

```
# Start up Network Time Protocol (NTP) daemon
```

```
#start xntpd "$src_running"
```

修改为:

```
start /usr/sbin/xntpd "$src_running" "-x"
```

重启 ntp 服务或重启主机:

```
# stopsrc -s xntpd
```

```
# startsrc -s xntpd -a "-x"
```

3.5.4 参数说明

tcp_recvspace	接收 TCP 数据的系统缺省套接字缓冲区大小。
tcp_sendspace	发送数据的系统缺省套接字缓冲区大小。
udp_sendspace	发送 UDP 数据的系统缺省套接字缓冲区大小。
udp_recvspace	接收 UDP 数据的系统缺省套接字缓冲区大小。Must be < sb_max
rfc1323	启用由 RFC 1323 (TCP 扩展以得到高性能) 指定的窗口定标和时间图标。
sb_max	确定 socket 使用缓存的最大值。
ipqmaxlen	指定接收包的数目,这些包可以列在 IP 协议输入队列中。
tcp_ephemeral_low	指定为 TCP 短暂端口分配的最小端口号。
tcp_ephemeral_high	指定为 TCP 短暂端口分配的最大端口号。
udp_ephemeral_low	指定为 UDP 短暂端口分配的最小端口号。
udp_ephemeral_high	指定为 UDP 短暂端口分配的最大端口号。

3.5.5 参数检查

```
# vi net_check.sh
```

```
no -a |grep tcp_recvspace
no -a |grep tcp_sendspace
no -a |grep udp_recvspace
no -a |grep udp_sendspace
no -a |grep 'rfc1323'
no -a |grep 'sb_max'
no -a |grep 'ipqmaxlen'
no -a |grep tcp_ephemeral_low
no -a |grep tcp_ephemeral_high
no -a |grep udp_ephemeral_low
no -a |grep udp_ephemeral_high
```

3.6 调整异步 IO

3.6.1 参数修改

```
ioo -p -o aio_maxreqs=131072
```

3.6.2 参数说明

```
# ioo -a|grep server
        aio_maxservers = 80
        aio_minservers = 40
    aio_server_inactivity = 300
    posix_aio_maxservers = 30
    posix_aio_minservers = 3
    posix_aio_server_inactivity = 300
```

注意：

1. Starting with AIX 6.1, minservers and maxservers are replaced by the aio_minservers and aio_maxservers parameters respectively.
2. Adjust the initial value of maxservers to 10 times the number of disks divided by the number of CPUs that are to be used concurrently but no more than 80.

3.7 创建 GRID,ORACLE 安装用户和组

在创建 grid,oracle 安装用户和组时，需要保证所有节点的用户和组的 ID 一致。

3.7.1 检查是否已经存在 Oracle Inventory

```
# more /etc/oraInst.loc
```

如果文件没有返回信息，表明没有安装数据库软件或安装不规范。

如果文件已经存在，并且返回的结果类似：

```
inventory_loc=/oracle/app/oracle/oraInventory
inst_group=oinstall
```

表明已经存在 oracle 软件安装信息。

3.7.2 创建 grid,oracle 安装组

```
mkgroup -'A' id='1000' oinstall
mkgroup -'A' id='1001' dba
mkgroup -'A' id='1002' oper
mkgroup -'A' id='1005' asmadmin
mkgroup -'A' id='1006' asmdba
mkgroup -'A' id='1007' asmoper
```

3.7.3 创建 grid,oracle 安装用户

➤ 检查 oracle 用户是否存在

```
# id oracle
# id grid
```

如果命令没有返回信息，表明用户不存在，

➤ 新建用户，用户默认路径为 /home/用户名

```
# mkuser id='1000' pgrp='oinstall' groups='asmadmin,asmdba,asmoper,dba'
grid
# mkuser id='1001' pgrp='oinstall' groups='asmdba,dba,oper' oracle
```

AIX 支持 mkuser 和 useradd 两种命令方式：

```
# useradd -u 1000 -g oinstall -G asmadmin,asmdba,asmoper,dba grid
# useradd -u 1001 -g oinstall -G asmdba,dba,oper oracle
```

➤ 删除用户：

```
# rmuser grid
# userdel grid
```

➤ 修改用户：

```
# usermod -g oinstall -G asmdba,dba oracle
```

3.7.4 创建 nobody 安装用户

```
# id nobody
# /bin/mkuser id='1002' nobody
```

3.7.5 修改 grid,oracle 用户口令

修改口令后，需要使用用户进行登录，重置用户口令。

```
# passwd grid
# passwd oracle
```

3.7.6 为用户添加 CAP_NUMA_ATTACH 等能力

```
# chuser capabilities=CAP_NUMA_ATTACH,CAP_BYPASS_RAC_VMM,CAP_PROPAGATE
grid
```

```
# chuser capabilities=CAP_NUMA_ATTACH,CAP_BYPASS_RAC_VMM,CAP_PROPAGATE
oracle
```

```
# chuser capabilities=CAP_NUMA_ATTACH,CAP_BYPASS_RAC_VMM,CAP_PROPAGATE
root
```

■ 查看用户属性

```
# lsuser grid
```

```
# lsuser oracle
# lsuser root
```

3.7.7 配置 shell 限制

所有节点，调整 root、grid、oracle 用户的 shell 限制。

```
# vi /etc/security/limits
```

```
default:
    fsize = -1
    core = 2097151
    cpu = -1
    data = -1
    rss = -1
    stack = -1
    nofiles = -1
```

统一修改用户的 limit。

通过 ulimit 命令进行检查

```
$ ulimit -a
time(seconds)          unlimited
file(blocks)           unlimited
data(kbytes)           unlimited
stack(kbytes)          4194304
memory(kbytes)         unlimited
coredump(blocks)       0
nofiles(descriptors)   unlimited
threads(per process)   unlimited
processes(per user)    unlimited
```

3.7.7.1 参数说明

fsize:文件大小限制,以块为单位。
core:内核转储(dump)大小,512 字节为单位。
cpu:用户进程可使用 CPU 限制,以秒为单位。
data:数据段大小限制,以字节为单位。
rss:用户进程常驻内存限制,以字节为单位。
stack:用户进程堆栈大小限制,以字节为单位。
nofiles:每个进程打开文件限制

3.7.7.2 注意事项

尽量提前修改该配置,否则 ftp 上传文件都会失败。

3.7.8 创建 grid,oracle 软件安装目录

■ GRID 软件目录

```
# export ORACLE_BASE=/oracle/app/oracle
# export ORACLE_HOME=/oracle/app/11.2.0/grid

# mkdir -p $ORACLE_BASE
# mkdir -p $ORACLE_HOME
# chmod -R 775 /oracle/
# chown -R grid:oinstall /oracle
```

■ DB 软件目录

```
# export ORACLE_BASE=/oracle/app/oracle
# export ORACLE_HOME=/oracle/app/oracle/product/11.2.0/db_1

# mkdir -p $ORACLE_HOME
# chown -R oracle:oinstall $ORACLE_HOME
# chmod -R 775 $ORACLE_HOME
```

3.7.9 配置 grid 用户环境变量

```
$ vi /home/grid/.profile
```

```
set -o vi
export umask=022
export LANG=C

export ORACLE_BASE=/oracle/app/oracle
export GRID_HOME=/oracle/app/11.2.0/grid
export ORACLE_SID=[根据实例进行设置]
export NLS_LANG=AMERICAN_AMERICA.ZHS16GBK
export NLS_DATE_FORMAT='YYYY-MM-DD HH24:MI:SS'

export LD_LIBRARY_PATH=$ORACLE_HOME/lib:/lib:/usr/lib
export PATH=$GRID_HOME/bin:$GRID_HOME/OPatch:/sbin:$PATH

export PS1="\`hostname`:\`$PWD> ";

if [ -t 0 ]; then
    stty intr ^C
fi
```

3.7.10 配置 oracle 用户环境变量

```
$ vi /home/oracle/.profile
```

```
set -o vi
export umask=022
export LANG=C

export ORACLE_BASE=/oracle/app/oracle
export GRID_HOME=/oracle/app/11.2.0/grid
export ORACLE_HOME=/oracle/app/oracle/product/11.2.0/db_1
export ORACLE_SID=[根据实例进行设置]
export ORACLE_UNQNAME=[根据数据库进行设置]
export NLS_LANG=american_america.zhs16gbk
export NLS_DATE_FORMAT='YYYY-MM-DD HH24:MI:SS'

export LD_LIBRARY_PATH=$ORACLE_HOME/lib:/lib:/usr/lib
export
PATH=$ORACLE_HOME/bin:$ORACLE_HOME/OPatch:$GRID_HOME/bin:/sbin:$PATH

export PS1="\`hostname`:\`$PWD> ";

if [ -t 0 ]; then
    stty intr ^C
fi
```

3.7.11 屏蔽 stty 造成的安装错误

GRID 安装过程中, 系统的隐含文件(如: .bashrc 或 .cshrc)如果包含 stty 命令, 会引起安装错误。

要避免该问题，必须在每个 `grid,oracle` 安装所有者的用户主目录中修改这些文件以取消所有 `STDERR` 输出安装错误。

■ Bourne, Bash, or Korn shell:

```
if [ -t 0 ]; then
    stty intr ^C
fi
```

■ C shell:

```
test -t 0
if ($status == 0) then
    stty intr ^C
endif
```

3.8 配置 `grid,oracle` 用户 SSH 互信设置

3.8.1 检查 `ssh` 进程运行状态

```
$ ps -ef | grep sshd
```

注意 bug 5033250:

运行 `NetCA` 和 `DBCA`，需要 `/usr/local/bin` 目录下有 `scp` 和 `ssh` 命令，如果 `/usr/local/bin` 目录下不存在这 2 个命令，需要创建符号链接文件。

```
# mkdir -p /usr/local/bin
# ln -s /etc/ssh /usr/local/etc
# ln -s /usr/bin /usr/local/bin
# ln -s /usr/bin/ksh /bin/bash
# ln -s /usr/bin/ssh-keygen /usr/local/bin/ssh-keygen
```

为避免连接自动关闭，可以修改 `/etc/ssh/sshd_config` 配置文件，添加

```
# vi /etc/ssh/sshd_config
LoginGraceTime 0
```

VIPCA 配置需要

```
# ln -s /usr/sbin/lsattr /etc/lsattr
```

3.8.2 所有节点配置 SSH

```
$ mkdir ~/.ssh
$ chmod 700 ~/.ssh
```

```
$ /usr/bin/ssh-keygen -t rsa    #生成 RSA key
$ /usr/bin/ssh-keygen -t dsa    #生成 DSA key
```

3.8.3 将 key 文件加入 `authorized key` 文件（在一个节点操作）

3.8.3.1 生成 `authorized_keys` 文件

```
$ touch ~/.ssh/authorized_keys
$ cd ~/.ssh
$ ls
```

```
$ chmod 600 ~/.ssh/authorized_keys
```

3.8.3.2 所有节点 key 信息加入 authorized_keys 文件

将所有节点的 id_rsa.pub 和 id_dsa.pub 信息加入到 authorized_keys 文件中，为 oracle 用户提供免口令的能力。

```
$ ssh node1 cat /home/grid/.ssh/id_rsa.pub >> authorized_keys
$ ssh node1 cat /home/grid/.ssh/id_dsa.pub >> authorized_keys

$ ssh node2 cat /home/grid/.ssh/id_rsa.pub >> authorized_keys
$ ssh node2 cat /home/grid/.ssh/id_dsa.pub >> authorized_keys
```

3.8.3.3 复制 authorized_keys 文件到其他节点

将 authorized_keys 文件复制到集群中的其他节点

```
$ scp authorized_keys node2:/home/grid/.ssh/
```

3.8.4 测试 SSH 配置

```
$ more /etc/hosts | grep -E -v '^#|^$|127.0.0.1|vip|:' | awk '{print "ssh"
" $2 " date;"}' > ping.sh
```

```
$ sh ping.sh
$ sh ping.sh
```

注意：

执行两次 ping.sh 命令，第一次需要输入口令时，输入 yes 后回车。

```
The authenticity of host 'node1 (140.87.152.153)' can't be established.
RSA key fingerprint is 7z:ez:e7:f6:f4:f2:4f:8f:9z:79:85:62:20:90:92:z9.
Are you sure you want to continue connecting (yes/no)?
```

如果第二次仍然需要输入口令，SSH 配置存在问题。

3.8.5 所有节点激活 SSH 互信（认证不同的情况下）

如果 SSH 配置正确，仍然无法免密码进行操作，可以使用下面的方法，在安装程序节点进行操作。

```
$ exec /usr/bin/ssh-agent $SHELL
$ /usr/bin/ssh-add
```

3.8.6 设置 x11 转发为 NO

启用 x11 转发，会造成安装失败，因此 oracle 用户需要屏蔽该特性

```
$ vi ~oracle/.ssh/config
Host *
    ForwardX11 no
```


3.8.7 设置图形环境

```
$ export DISPLAY=REMOTEIP:0.0
```

REMOTEIP 为远程客户端主机的 IP 地址。

4 配置共享存储

4.1 存储方式与文件类型对应关系

Storage Option	File Types Supported			
	OCR & Voting Disk	Oracle Software	Database	Recovery
ASM	Yes	No	Yes	Yes
GPFS	Yes	Yes	Yes	Yes
Local storage	No	Yes	No	No
Raw LVM by HACMP	No	No	No	No



说明

- OCR & voting disks 要求至少 256 MB。
- 11G OCR & voting disks 只支持 ASM 和 GPFS 方式。
- 标准版 RAC 安装，只支持 ASM 存放数据文件和闪回文件。

4.2 配置 ASM 使用的共享裸磁盘设备

4.2.1 查看可用磁盘

```
# /usr/sbin/lssdev -cc disk
```

```
hdisk0 Available 05-08-00-8,0 16 Bit LVD SCSI Disk Drive
hdisk1 Available 05-08-00-9,0 16 Bit LVD SCSI Disk Drive
hdisk2 Available 09-08-02 IBM MPIO FC 2107
hdisk3 Available 09-08-02 IBM MPIO FC 2107
hdisk4 Available 09-08-02 IBM MPIO FC 2107
hdisk5 Available 09-08-02 IBM MPIO FC 2107
hdisk6 Available 09-08-02 IBM MPIO FC 2107
hdisk7 Available 09-08-02 IBM MPIO FC 2107
hdisk8 Available 09-08-02 IBM MPIO FC 2107
hdisk9 Available 09-08-02 IBM MPIO FC 2107
hdisk10 Available 09-08-02 IBM MPIO FC 2107
hdisk11 Available 09-08-02 IBM MPIO FC 2107
hdisk12 Available 09-08-02 IBM MPIO FC 2107
hdisk13 Available 09-08-02 IBM MPIO FC 2107
hdisk14 Available 09-08-02 IBM MPIO FC 2107
hdisk15 Available 09-08-02 IBM MPIO FC 2107
hdisk16 Available 09-08-02 IBM MPIO FC 2107
```

如果有未被列出的磁盘，通过下面命令进行查找

```
# /usr/sbin/cfgmgr
```

从上面的返回结果可知，hdisk2-hdisk16 是来至存储的磁盘设备。

4.2.2 查看所有节点的磁盘设备是否一致

```
# /usr/sbin/lspv | grep -i none
```

hdisk1	none	None
hdisk2	none	None
hdisk3	none	None
hdisk4	none	None
hdisk5	none	None
hdisk6	none	None
hdisk7	none	None
hdisk8	none	None
hdisk9	none	None
hdisk10	none	None
hdisk11	none	None
hdisk12	none	None
hdisk13	none	None
hdisk14	none	None
hdisk15	none	None
hdisk16	none	None

如果磁盘没有 **PVID** 属性，通过下面命令进行分配，方便在各节点之间进行信息对比：
/usr/sbin/chdev -l hdiskn -a pv=yes

给磁盘分配 **PVID**

```
# /usr/sbin/lspv | grep -i none | awk '{print "/usr/sbin/chdev -l \"$1\" -a pv=yes"}'
```

```
/usr/sbin/chdev -l hdisk2 -a pv=yes
/usr/sbin/chdev -l hdisk3 -a pv=yes
/usr/sbin/chdev -l hdisk4 -a pv=yes
/usr/sbin/chdev -l hdisk5 -a pv=yes
/usr/sbin/chdev -l hdisk6 -a pv=yes
/usr/sbin/chdev -l hdisk7 -a pv=yes
/usr/sbin/chdev -l hdisk8 -a pv=yes
/usr/sbin/chdev -l hdisk9 -a pv=yes
/usr/sbin/chdev -l hdisk10 -a pv=yes
/usr/sbin/chdev -l hdisk11 -a pv=yes
/usr/sbin/chdev -l hdisk12 -a pv=yes
/usr/sbin/chdev -l hdisk13 -a pv=yes
/usr/sbin/chdev -l hdisk14 -a pv=yes
/usr/sbin/chdev -l hdisk15 -a pv=yes
/usr/sbin/chdev -l hdisk16 -a pv=yes
```

查看磁盘 **PVID** 信息

```
# /usr/sbin/lspv
```

hdisk0	00cda24d4c6ce504	rootvg	active
hdisk1	none	None	
hdisk2	00cda24d8eccfc23	None	
hdisk3	00cda24d8eccfc86	None	
hdisk4	00cda24d8eccfcde	None	
hdisk5	00cda24d8eccfd36	None	
hdisk6	00cda24d8eccfd8e	None	
hdisk7	00cda24d8eccfde6	None	
hdisk8	00cda24d8eccfe3e	None	
hdisk9	00cda24d8eccfe96	None	
hdisk10	00cda24d8eccfeef	None	
hdisk11	00cda24d8eccff46	None	
hdisk12	00cda24d8eccff9e	None	
hdisk13	00cda24d8eccfff6	None	
hdisk14	00cda24d8ecd004e	None	
hdisk15	00cda24d8ecd00a6	None	
hdisk16	00cda24d8ecd00fe	None	

查看其他节点对于 **PVID** 的磁盘设备名称

```
# /usr/sbin/lspv | grep -i "0009005fb9c23648"
```

```
# lspcfg -vl hdisk17
```

4.2.3 设置磁盘并发访问的 ODM 属性

Disk Type	Attribute	Value
EMC, HDS, IBM DS4000, and non-MPIO capable	reserve_lock	no
ESS, EMC, HDS, CLARiion, DS8000, DS6000	reserve_policy	no_reserve

4.2.3.1 查看磁盘属性

```
# /usr/sbin/lssattr -E -l hdiskn | grep reserve
```

如果属性与要求不一致，通过下面命令进行修改

■ EMC 存储

```
# /usr/sbin/chdev -l hdiskn -a reserve_lock=no
```

■ IBM 存储

```
# /usr/sbin/chdev -l hdiskn -a reserve_policy=no_reserve
```

■ HDS 存储

```
# /usr/sbin/chdev -l hdiskn -a dlmsvlevel=no_reserve
```

本次使用 IBM 的 DS8870 存储：

```
# /usr/sbin/lspv | awk '{print "/usr/sbin/chdev -l \"$1\" -a reserve_policy=no_reserve"}'
```

```
/usr/sbin/chdev -l hdisk2 -a reserve_policy=no_reserve
/usr/sbin/chdev -l hdisk3 -a reserve_policy=no_reserve
/usr/sbin/chdev -l hdisk4 -a reserve_policy=no_reserve
/usr/sbin/chdev -l hdisk5 -a reserve_policy=no_reserve
/usr/sbin/chdev -l hdisk6 -a reserve_policy=no_reserve
/usr/sbin/chdev -l hdisk7 -a reserve_policy=no_reserve
/usr/sbin/chdev -l hdisk8 -a reserve_policy=no_reserve
/usr/sbin/chdev -l hdisk9 -a reserve_policy=no_reserve
/usr/sbin/chdev -l hdisk10 -a reserve_policy=no_reserve
/usr/sbin/chdev -l hdisk11 -a reserve_policy=no_reserve
/usr/sbin/chdev -l hdisk12 -a reserve_policy=no_reserve
/usr/sbin/chdev -l hdisk13 -a reserve_policy=no_reserve
/usr/sbin/chdev -l hdisk14 -a reserve_policy=no_reserve
/usr/sbin/chdev -l hdisk15 -a reserve_policy=no_reserve
/usr/sbin/chdev -l hdisk16 -a reserve_policy=no_reserve
```

4.2.4 检查所有磁盘属性

执行 disk_info.sh 脚本，进行信息校验。

```
for i in 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16
do
    lsattr -El hdisk$i | grep reserve | awk '{print $1,$2 }'| read rp1 rp2
    lsattr -El hdisk$i | grep pvid | awk '{print $1,$2 }'| read pv1 pv2
    bootinfo -s hdisk$i | awk '{print $1 }'| read sz1
    lsattr -El hdisk$i | grep lun_id | awk '{print $1,$2 }'| read li1 li2
    if [ "$li1" != "" ]
    then
        echo hdisk$i' -> '$li1' = '$li2' / '$rp1' = '$rp2' / '$pv1' = '$pv2'
        / 'size' = '$sz1
    fi
done
echo
for i in 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16
do
    lscfg -v1 hdisk$i
done
```

```
ls -la /dev/hdisk*
ls -la /dev/rdisk*
ls -la /dev/ASM*
```

4.2.5 清除需要使用的磁盘的 PVID

```
# /usr/sbin/lspv | grep -i none | awk '{print "/usr/sbin/chdev -l \"$1\" -a pv=clear"}'
```

```
/usr/sbin/chdev -l hdisk2 -a pv=clear
/usr/sbin/chdev -l hdisk3 -a pv=clear
/usr/sbin/chdev -l hdisk4 -a pv=clear
/usr/sbin/chdev -l hdisk5 -a pv=clear
/usr/sbin/chdev -l hdisk6 -a pv=clear
/usr/sbin/chdev -l hdisk8 -a pv=clear
/usr/sbin/chdev -l hdisk9 -a pv=clear
/usr/sbin/chdev -l hdisk10 -a pv=clear
/usr/sbin/chdev -l hdisk11 -a pv=clear
/usr/sbin/chdev -l hdisk12 -a pv=clear
/usr/sbin/chdev -l hdisk13 -a pv=clear
/usr/sbin/chdev -l hdisk14 -a pv=clear
/usr/sbin/chdev -l hdisk16 -a pv=clear
```

4.2.6 修改设备的权限

去除本地磁盘和作为挂载点的磁盘。需要过滤的磁盘包含：hdisk0、hdisk1、hdisk7、hdisk15。

4.2.6.1 修改磁盘权限

```
# ls -la /dev/rhdisk* | awk '{print "chown grid:asmadmin \"$10\"}'
chown grid:asmadmin /dev/rhdisk10
chown grid:asmadmin /dev/rhdisk11
chown grid:asmadmin /dev/rhdisk12
chown grid:asmadmin /dev/rhdisk13
chown grid:asmadmin /dev/rhdisk14
chown grid:asmadmin /dev/rhdisk16
chown grid:asmadmin /dev/rhdisk2
chown grid:asmadmin /dev/rhdisk3
chown grid:asmadmin /dev/rhdisk4
chown grid:asmadmin /dev/rhdisk5
chown grid:asmadmin /dev/rhdisk6
chown grid:asmadmin /dev/rhdisk8
chown grid:asmadmin /dev/rhdisk9
```

4.2.6.2 修改磁盘属组

```
# ls -la /dev/rhdisk* | awk '{print "chmod 660 \"$10\"}'
chmod 660 /dev/rhdisk10
chmod 660 /dev/rhdisk11
chmod 660 /dev/rhdisk12
chmod 660 /dev/rhdisk13
chmod 660 /dev/rhdisk14
chmod 660 /dev/rhdisk16
chmod 660 /dev/rhdisk2
chmod 660 /dev/rhdisk3
chmod 660 /dev/rhdisk4
chmod 660 /dev/rhdisk5
chmod 660 /dev/rhdisk6
chmod 660 /dev/rhdisk8
chmod 660 /dev/rhdisk9
```

4.2.6.3 查询结果

```
# ls -la /dev/rhdisk*
```

4.2.7 参考文档

Placement of Voting disk and OCR Files in Oracle RAC 10g and 11gR1
(文档 ID 293819.1)

5 安装 GRID 软件

5.1 解压缩软件

AIX 下默认没有 zip 软件，可以用 jdk 自带的 jar 命令进行解压：

```
$ cd /oracle/software
$ /usr/java6/bin/jar -xvf p10404530_112030_AIX64-5L_3of7.zip
$ /usr/java6/bin/jar -xvf p10404530_112030_AIX64-5L_1of7.zip
$ /usr/java6/bin/jar -xvf p10404530_112030_AIX64-5L_2of7.zip
```

5.2 运行 runInstaller 脚本

在图形界面命令行中用 oracle 用户执行命令启动安装向导：

```
$ export DISPLAY={客户端 IP 地址}:0.0
$ cd /oracle/software/grid
$ ./runInstaller
```

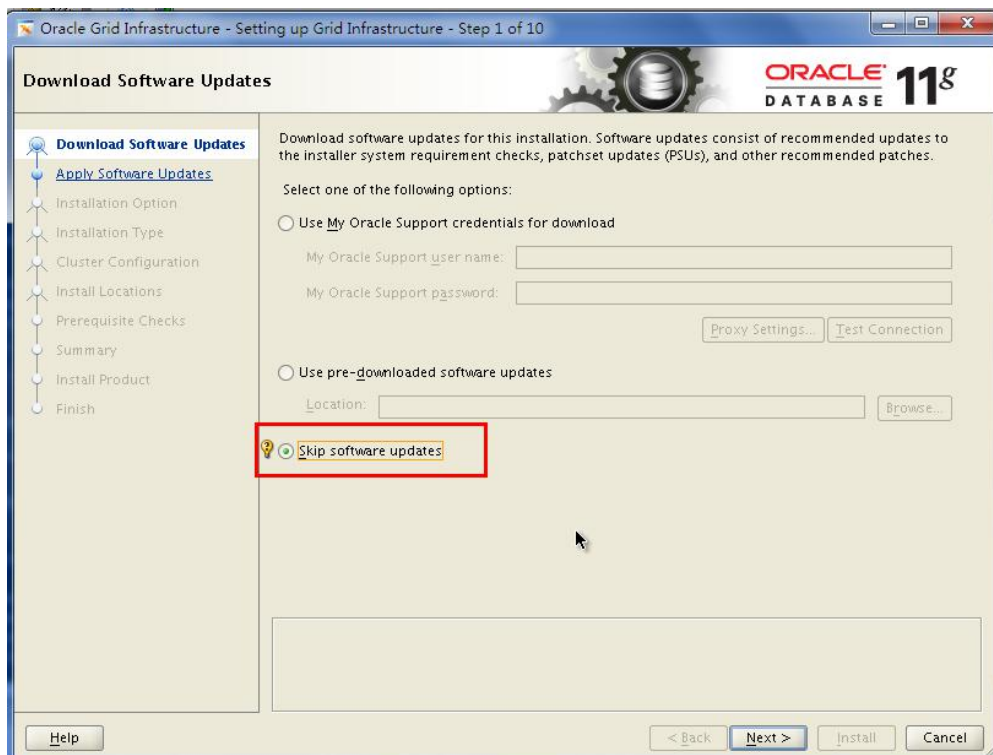
```
$ ./runInstaller
*****
Your platform requires the root user to perform certain pre-installation
OS preparation. The root user should run the shell script 'rootpre.sh' before
you proceed with Oracle installation. rootpre.sh can be found at the top level
of the CD or the stage area.

Answer 'y' if root has run 'rootpre.sh' so you can proceed with Oracle
installation.
Answer 'n' to abort installation and then ask root to run 'rootpre.sh'.
*****
Has 'rootpre.sh' been run by root on all nodes? [y/n] (n)
```

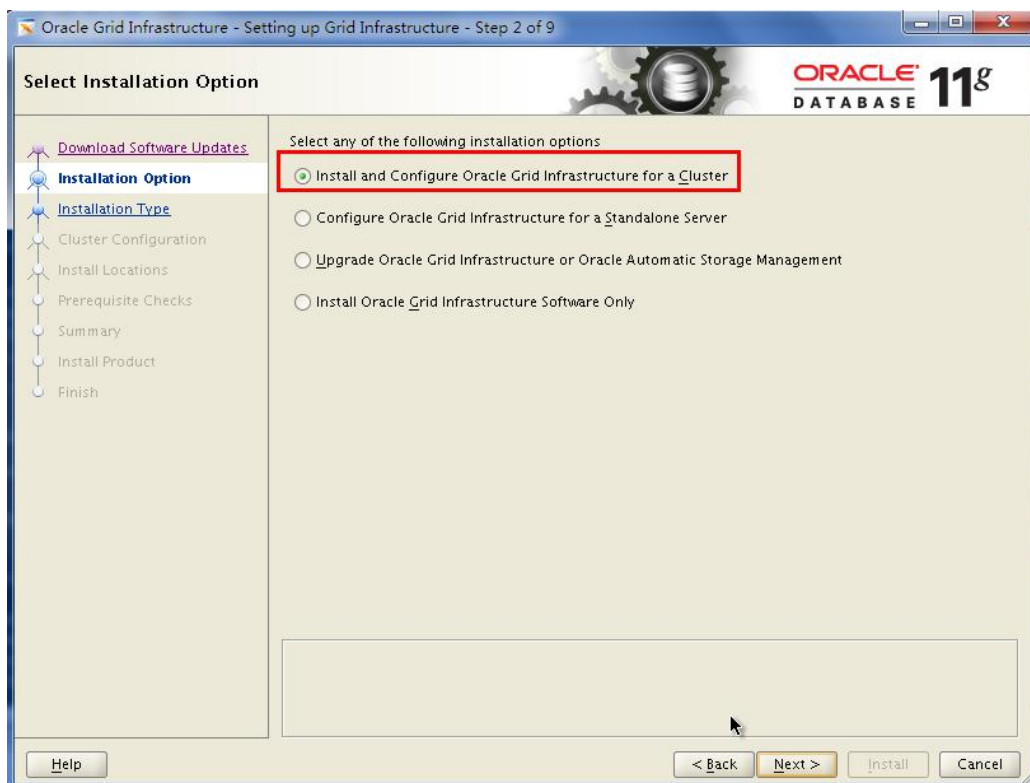
根据提示，使用 root 用户，运行 rootpre.sh 脚本。进行安装前检测。

执行过 rootpre.sh 脚本后，输入 y 回车，出现安装界面。

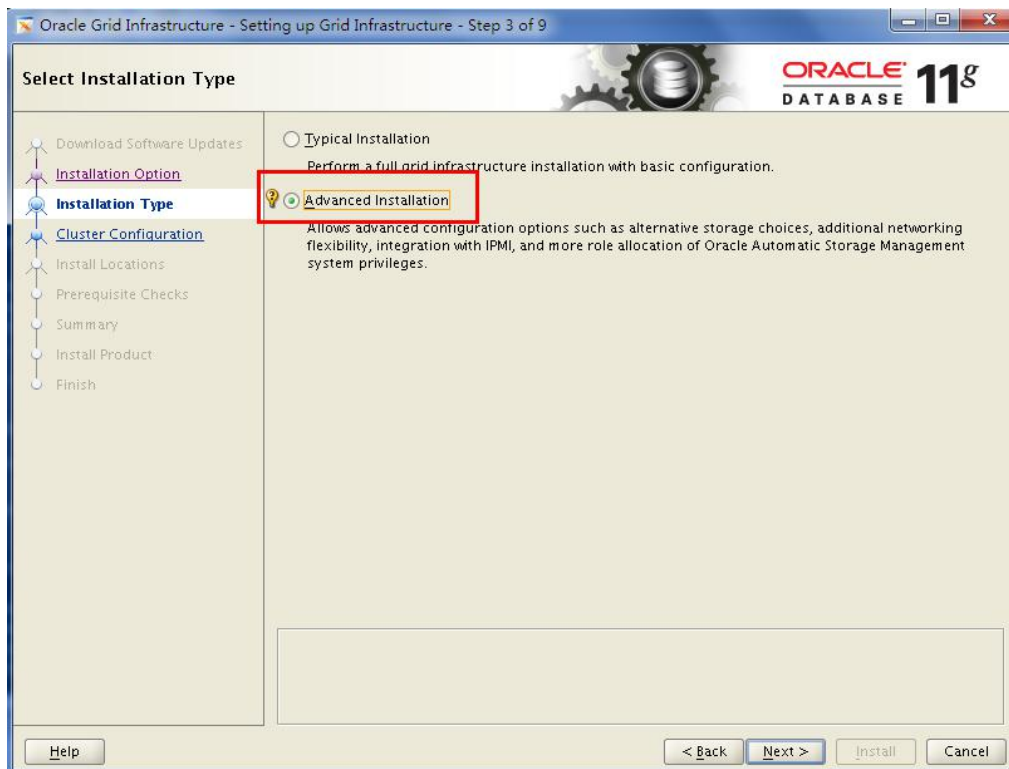




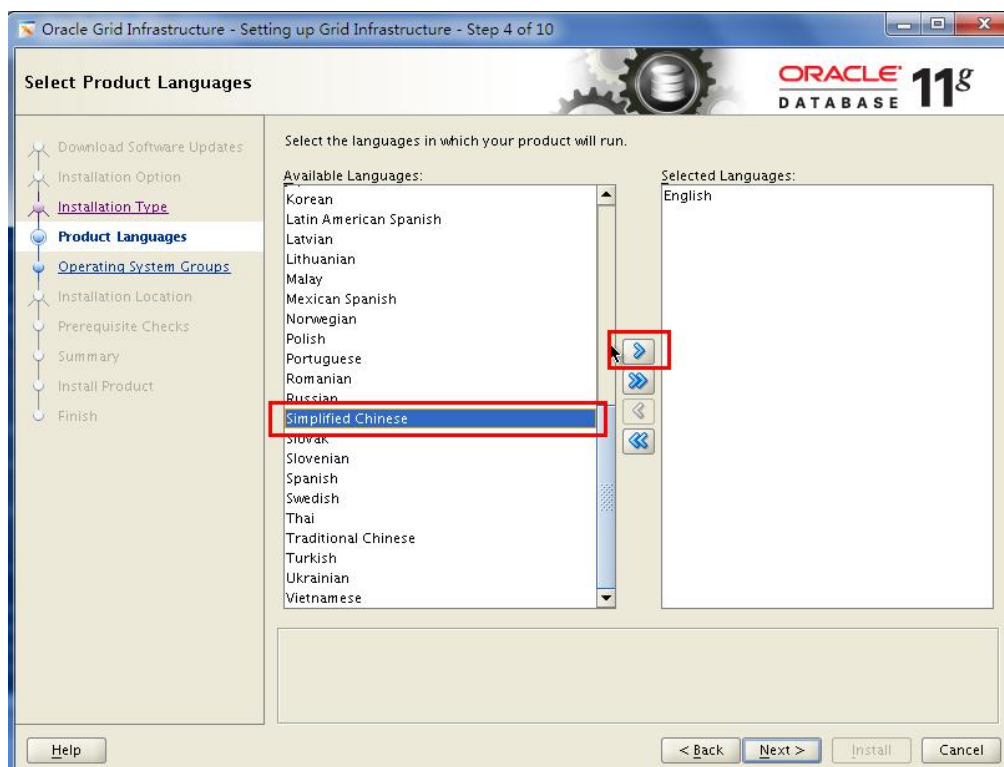
选择跳过软件更新。点击 Next，进入下一步



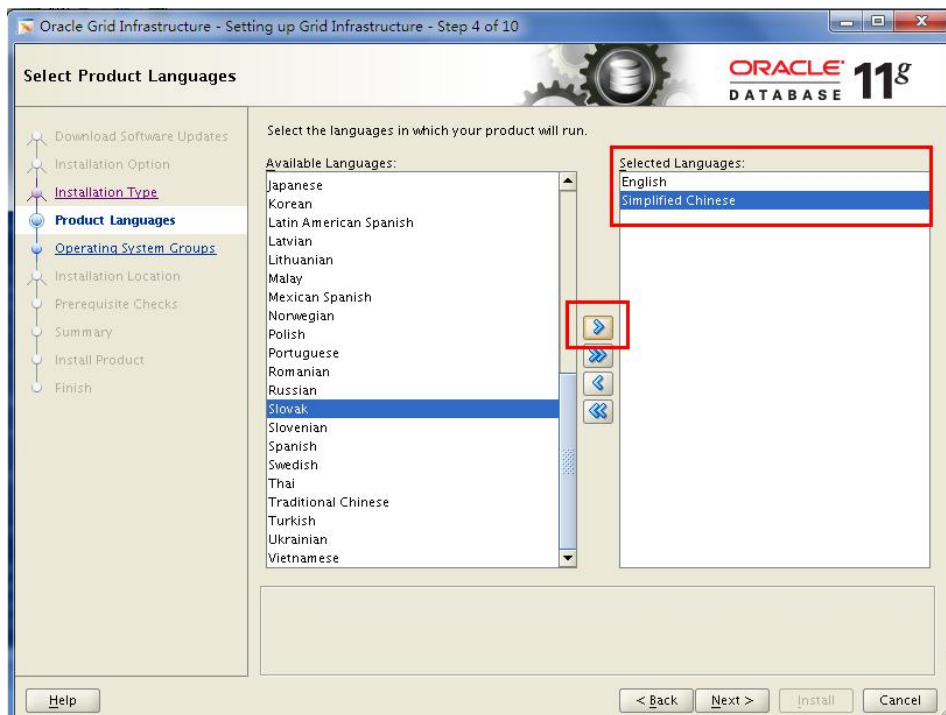
选择集群安装方式，点击 Next，进入下一步



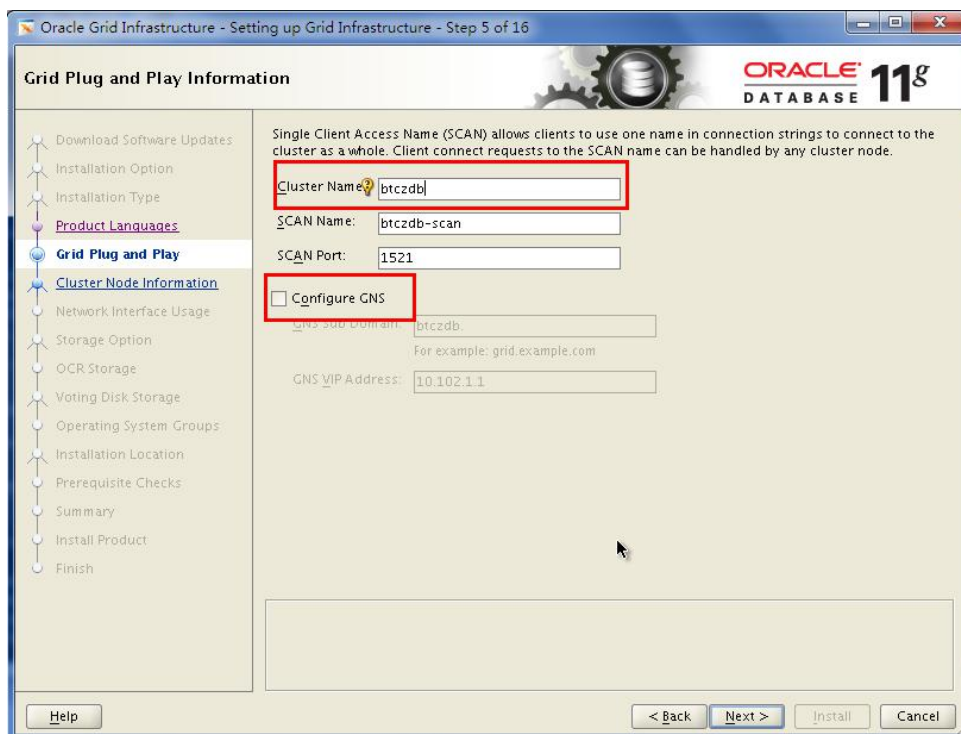
选择高级安装方式，点击 **Next**，进入下一步



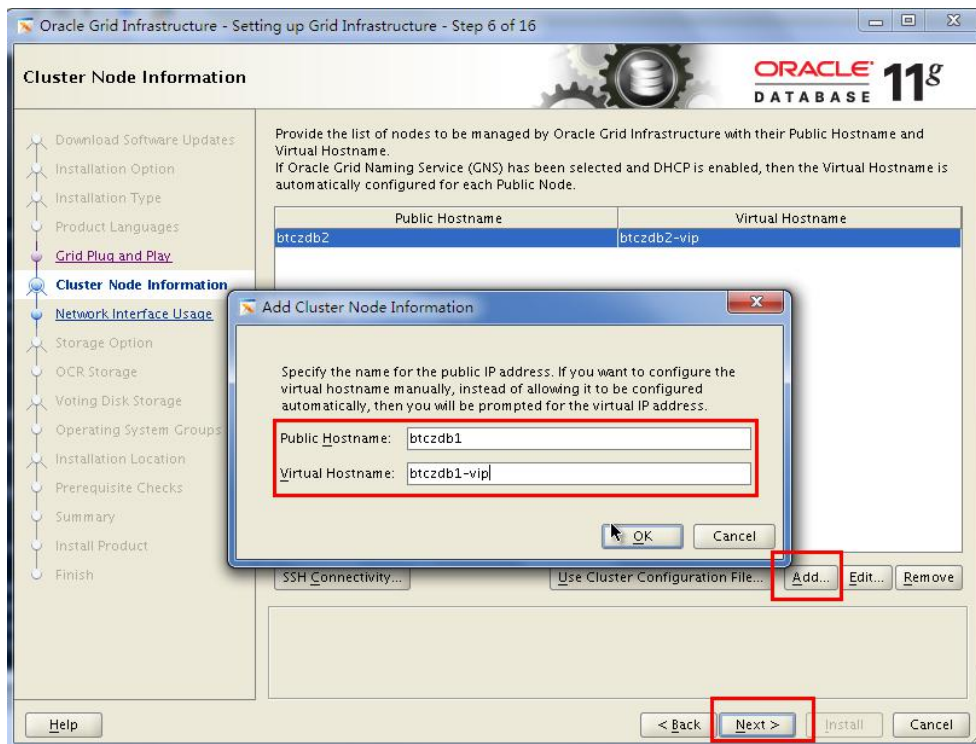
添加简体中文语言



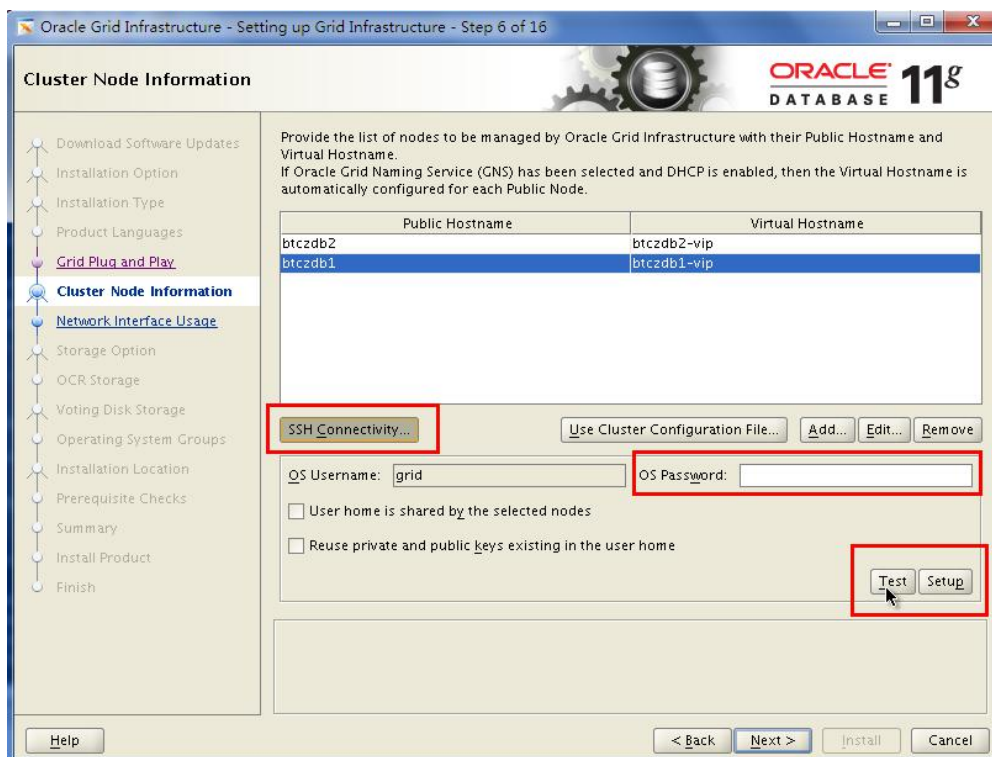
点击 Next，进入下一步



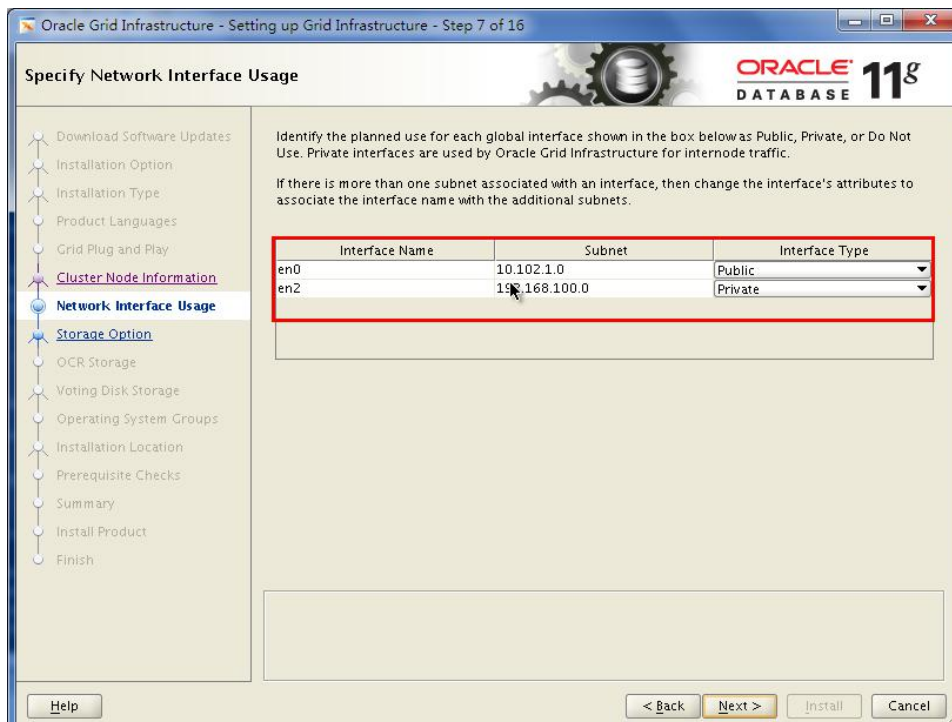
集群名称不能超过 15 位，可以包含“-”。取消 GNS，点击 Next，进入下一步



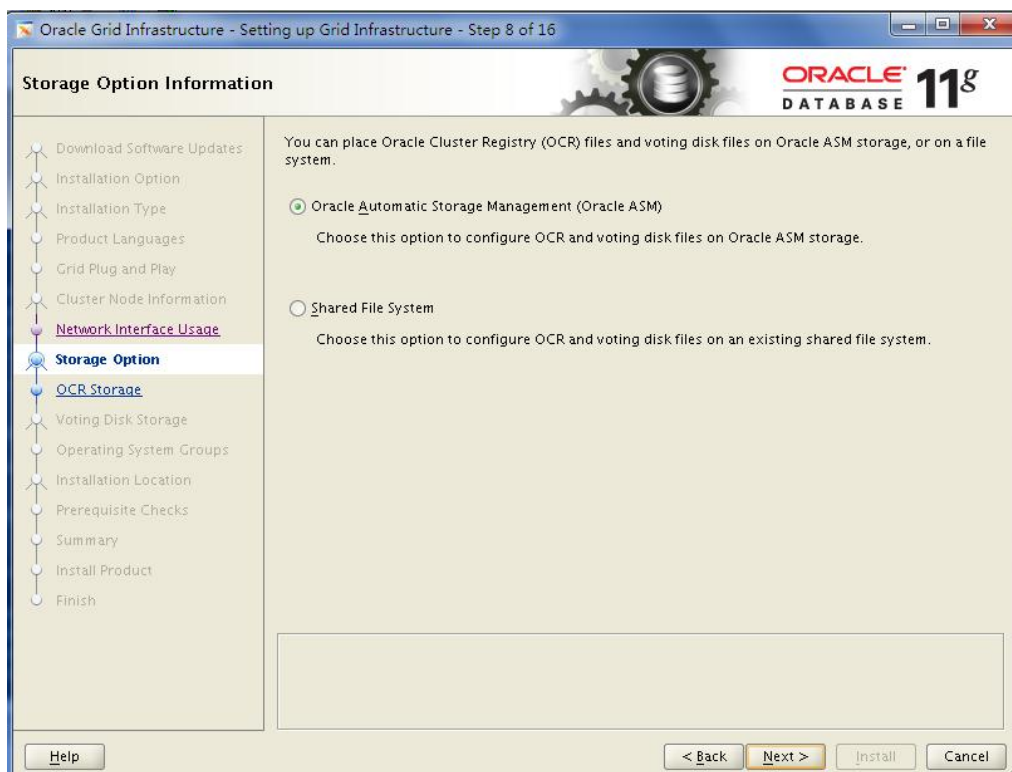
点击 Add 按钮，添加新节点的主机名和虚拟主机名，点击 OK，完成主机名的添加



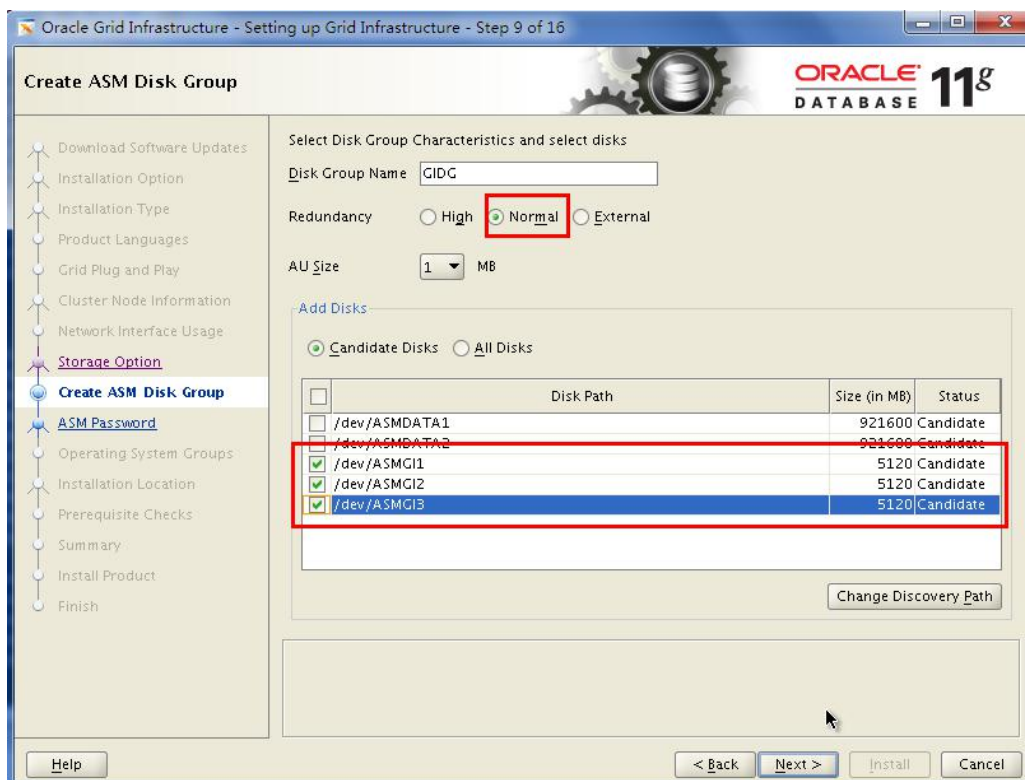
输入 grid 用户的口令，点击 Setup 按钮，自动完成 SSH 互信设置，点击 Next，自动完成 SSH 校验，校验通过后，进入下一步



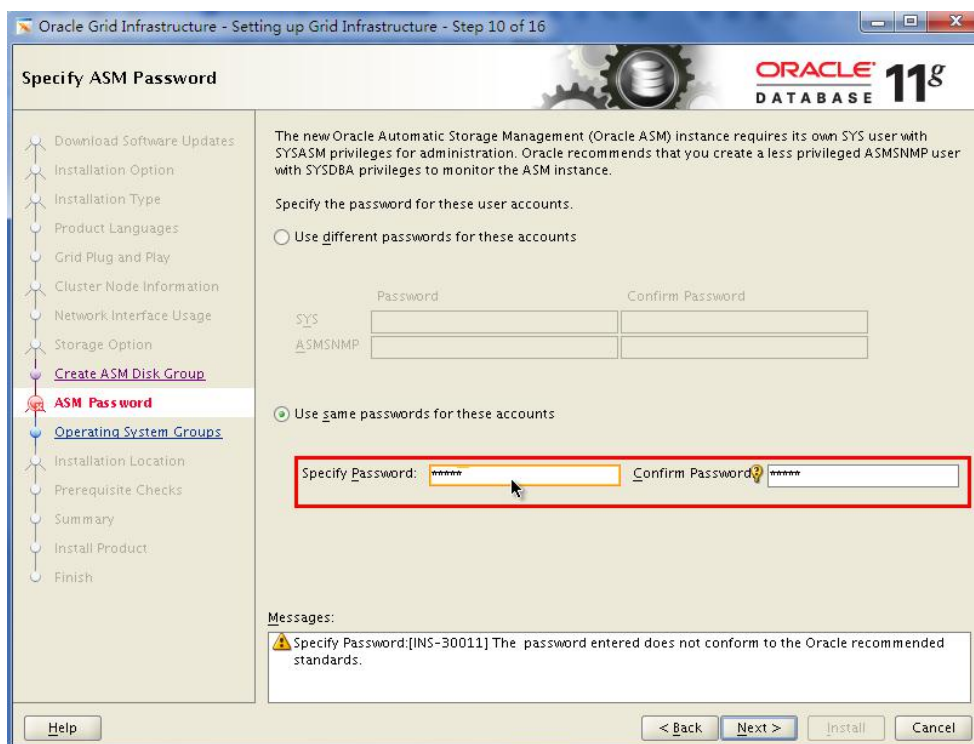
选择对应的网卡做为不同的接口类型，点击 Next，进入下一步



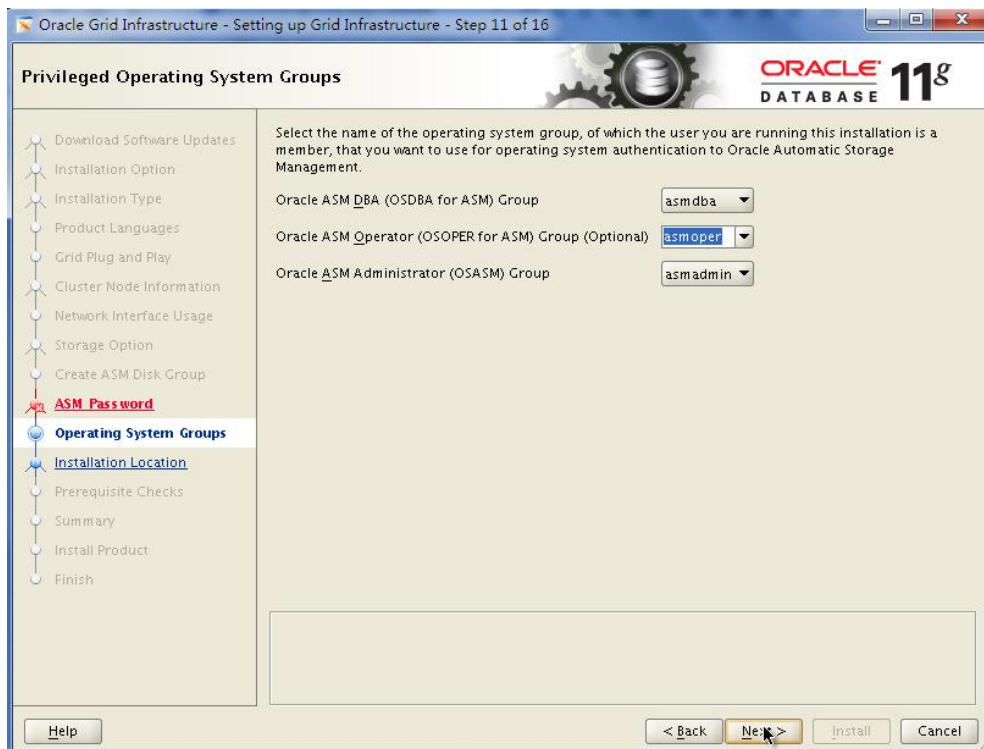
选择存储类型为 Oracle ASM，点击 Next，进入下一步



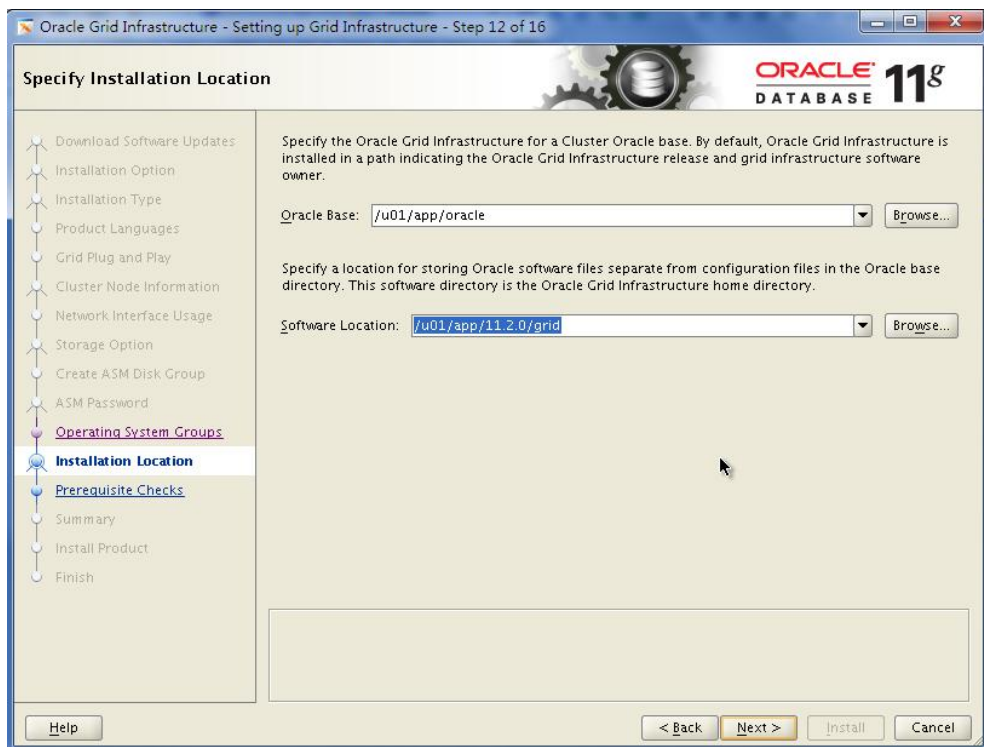
新建磁盘组，名为 GIDG，用于存储表决磁盘和 OCR 文件，使用普通冗余，需要 3 个裸设备文件，大小为 5G。点击 Next，进入下一步



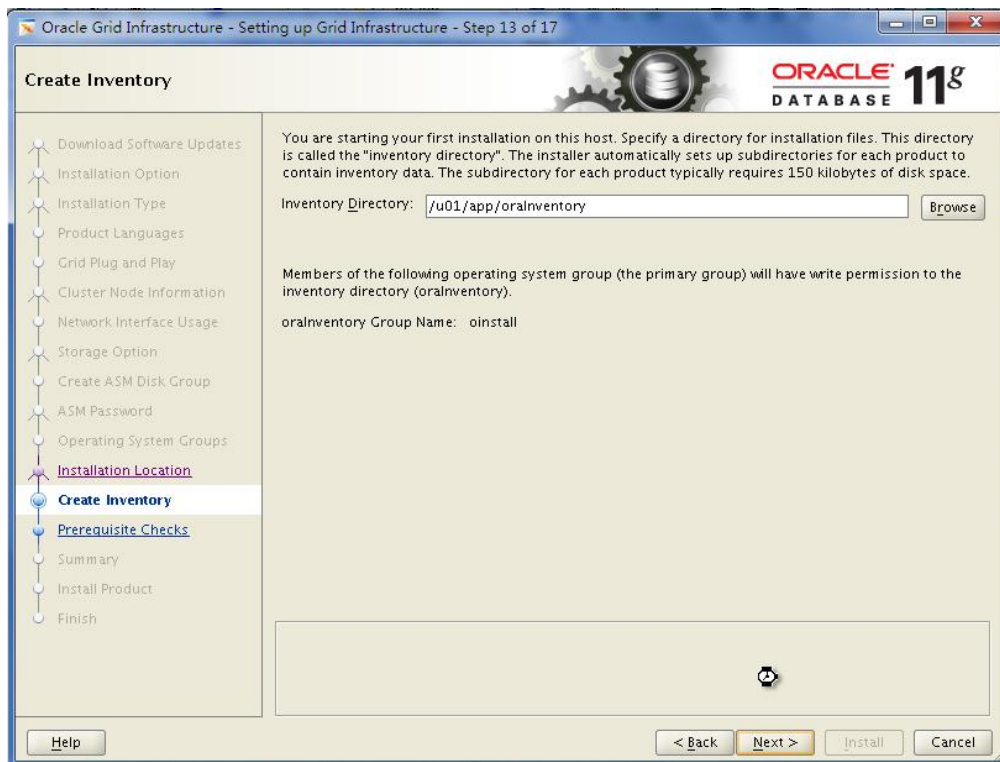
指定 SYS 用户的临时口令为 Admin1234，忽略警告信息，点击 Next，进入下一步



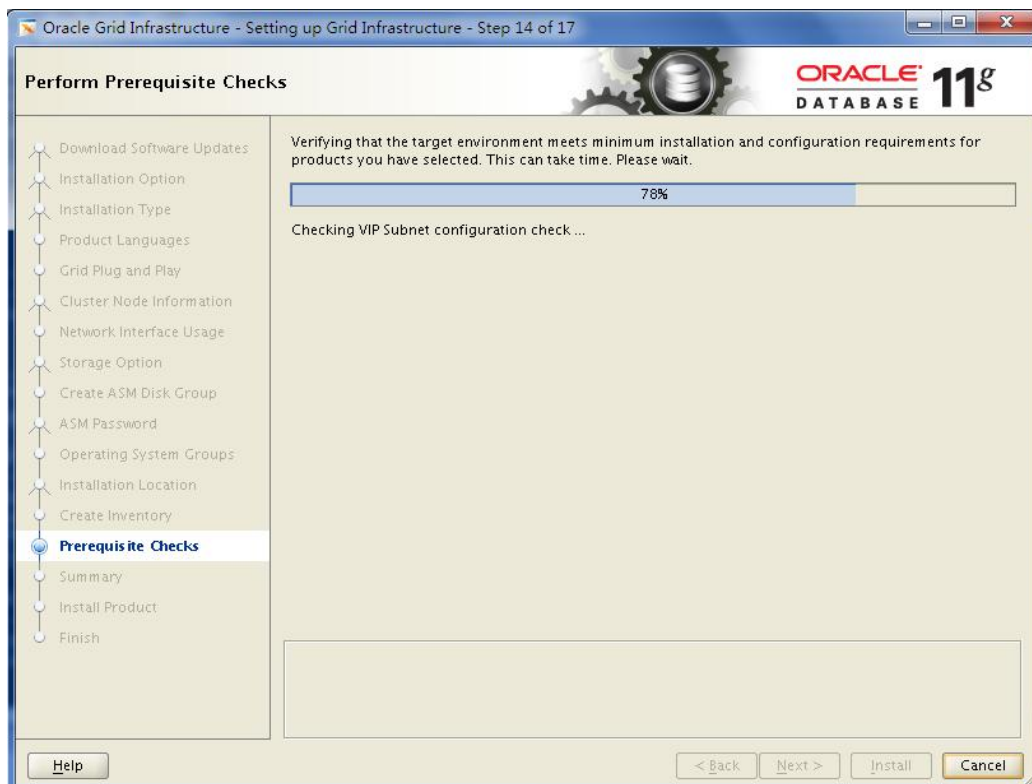
确认安装信息的用户属组，点击 Next，进入下一步



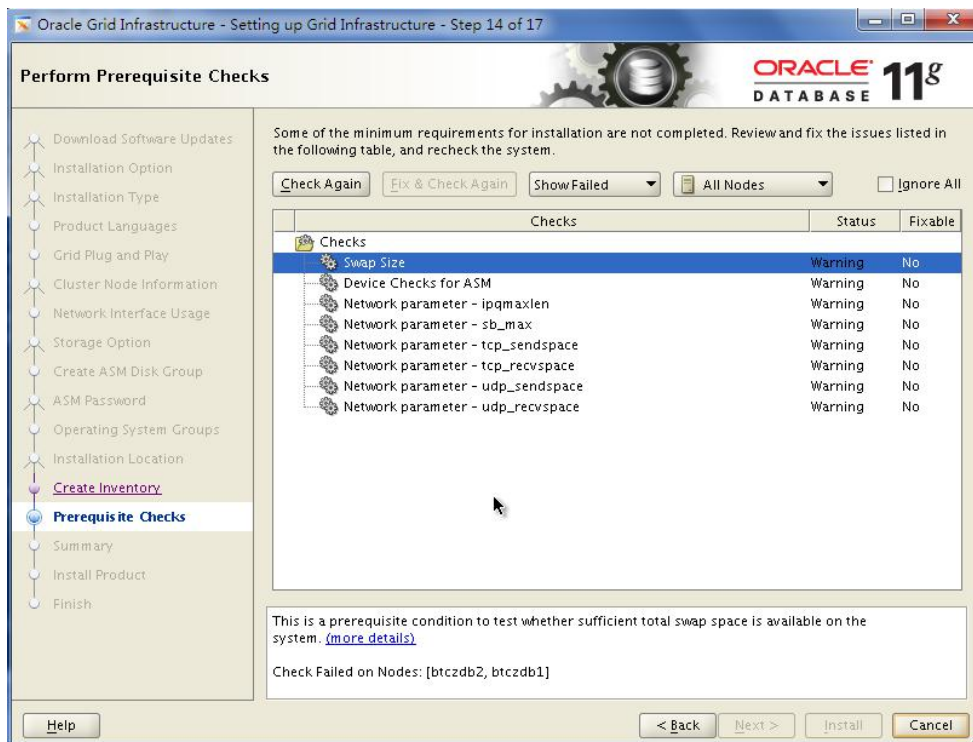
确认安装目录结构信息，点击 Next，进入下一步



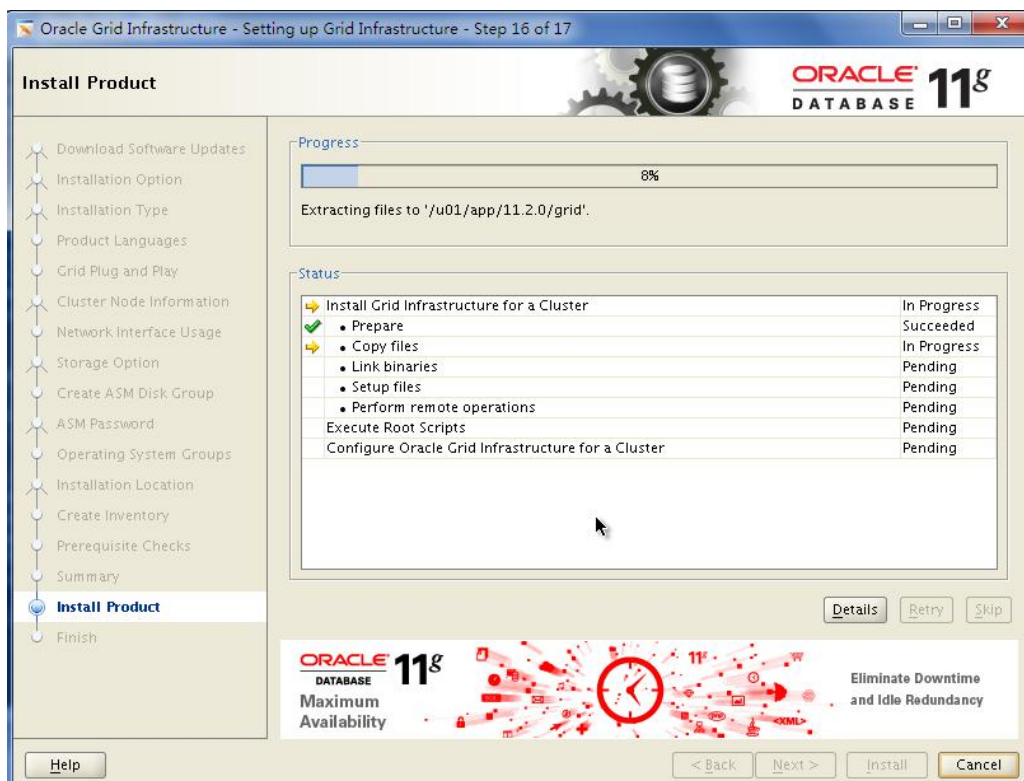
确认 Oracle Inventory 目录信息，点击 Next，进入下一步



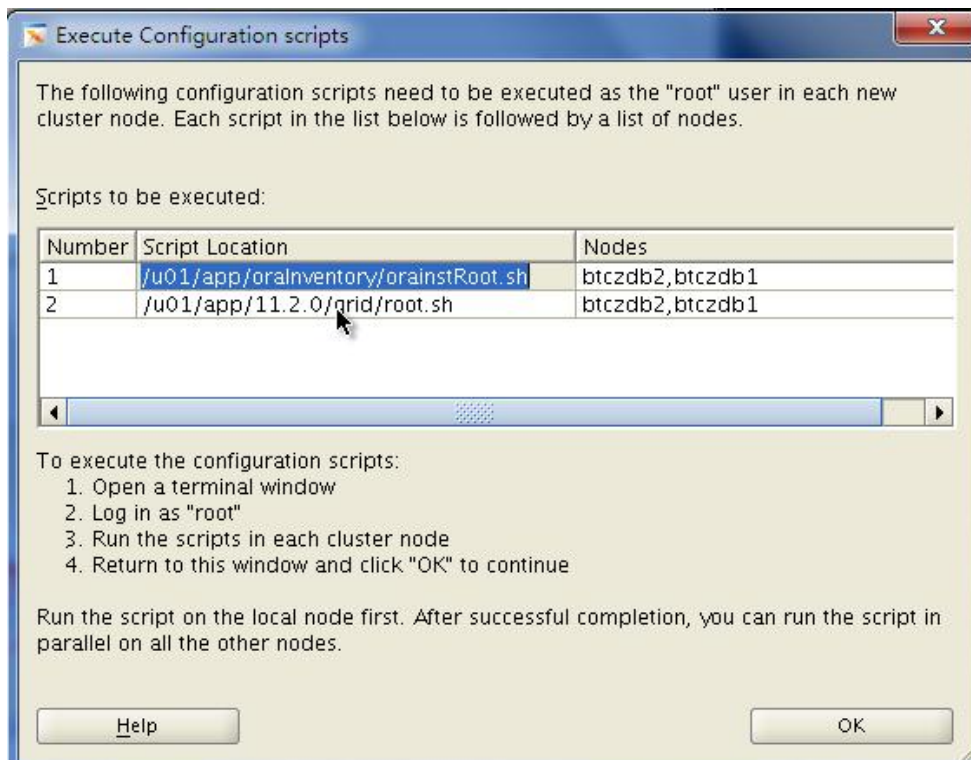
进行安装环境信息检测，



针对提示的告警信息，进行修复，修复完成后，点击 Next，进入下一步



进入 GI 安装过程



根据提示，运行 orainstRoot.sh 和 root.sh 脚本，进行 GI 安装配置

orainstRoot.sh 脚本

```
# /oracle/app/oraInventory/orainstRoot.sh
```

Changing permissions of /oracle/app/oraInventory.

Adding read,write permissions for group.

Removing read,write,execute permissions for world.

Changing groupname of /oracle/app/oraInventory to oinstall.

The execution of the script is complete.

root.sh 脚本

```
# /oracle/app/11.2.0/grid/root.sh
```

在 2 个节点，根据提示，分别运行脚本

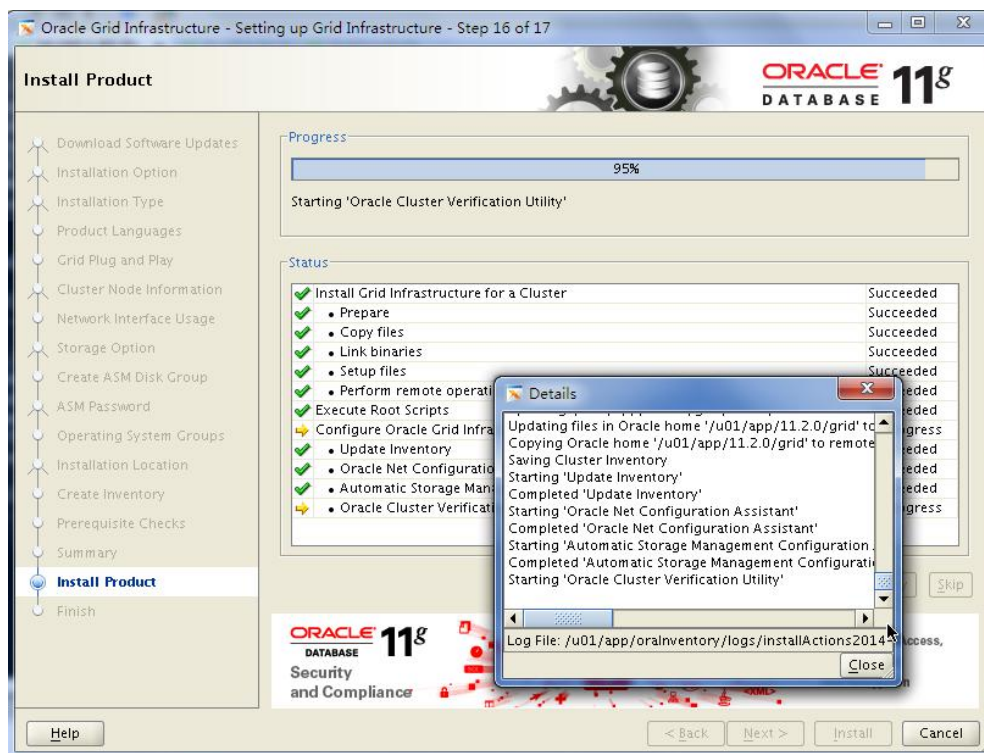
```

CRS-4266: Voting file(s) successfully replaced
##  STATE      File Universal Id      File Name Disk group
---  -
1.  ONLINE     94d2406553c84fe2bf00b844d32407b6  (/dev/ASMG11) [GIDG]
2.  ONLINE     5eff446a4ce34f88bf87a9bdca1cea1c  (/dev/ASMG12) [GIDG]
3.  ONLINE     513181e9e2934fbbef43757d678d9c30  (/dev/ASMG13) [GIDG]
Located 3 voting disk(s).
CRS-2672: Attempting to start 'ora.asm' on 'btczdb2'
CRS-2676: Start of 'ora.asm' on 'btczdb2' succeeded
CRS-2672: Attempting to start 'ora.GIDG.dg' on 'btczdb2'
CRS-2676: Start of 'ora.GIDG.dg' on 'btczdb2' succeeded
CRS-2672: Attempting to start 'ora.registry.acfs' on 'btczdb2'
CRS-2676: Start of 'ora.registry.acfs' on 'btczdb2' succeeded
Configure Oracle Grid Infrastructure for a Cluster ... succeeded

```

在最后一个节点运行完 root.sh 后，点击 OK 按钮。

5.3 完成 GI 安装



5.4 执行 GI 安装后校验

5.4.1 crs_stat

```
$ crs_stat -t
```

5.4.2 olsnodes

```
$ olsnodes
```

5.4.3 crsctl

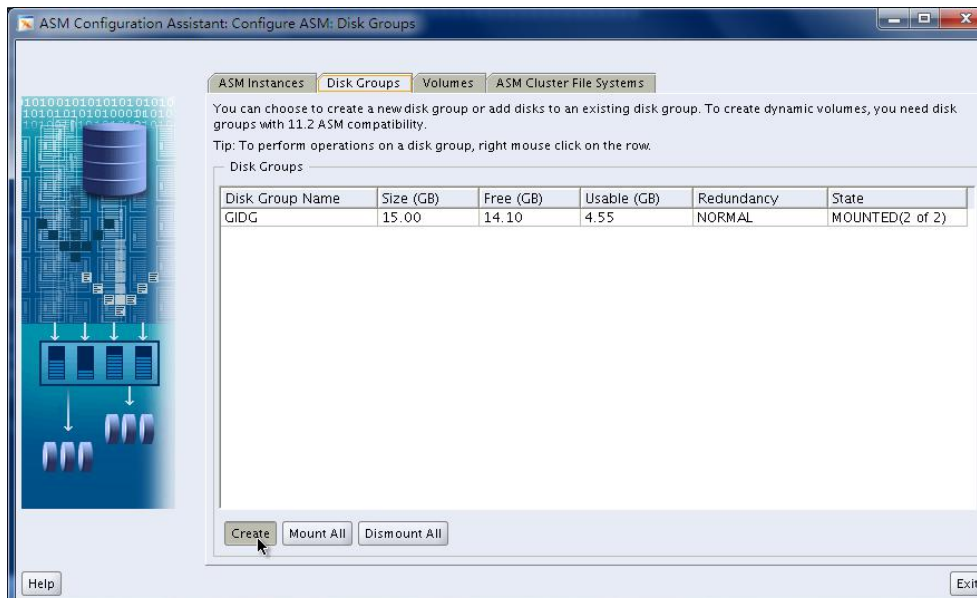
```
$ crsctl check crs
```

5.4.4 GI 版本

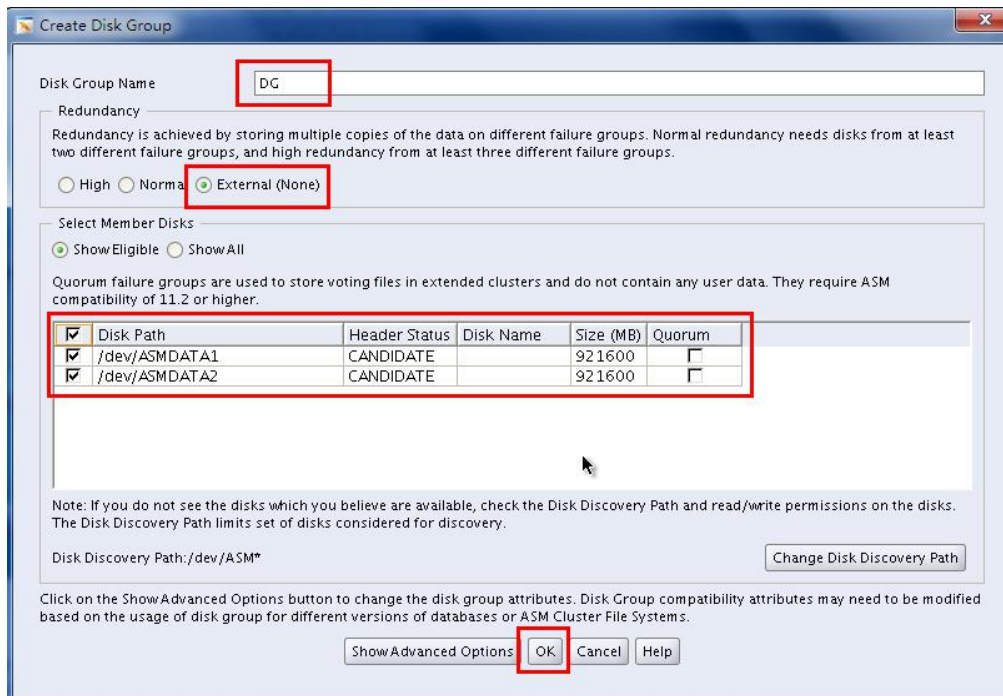
```
$ crsctl query crs activeversion
```

6 配置 ASM 磁盘组

在 grid 用户下，执行 asmca 命令，启动 ASM 配置界面
\$ asmca



点击 Create 按钮，添加新的磁盘组



输入磁盘组名称，选择磁盘组中的磁盘，点击 OK 按钮，完成磁盘组的新建

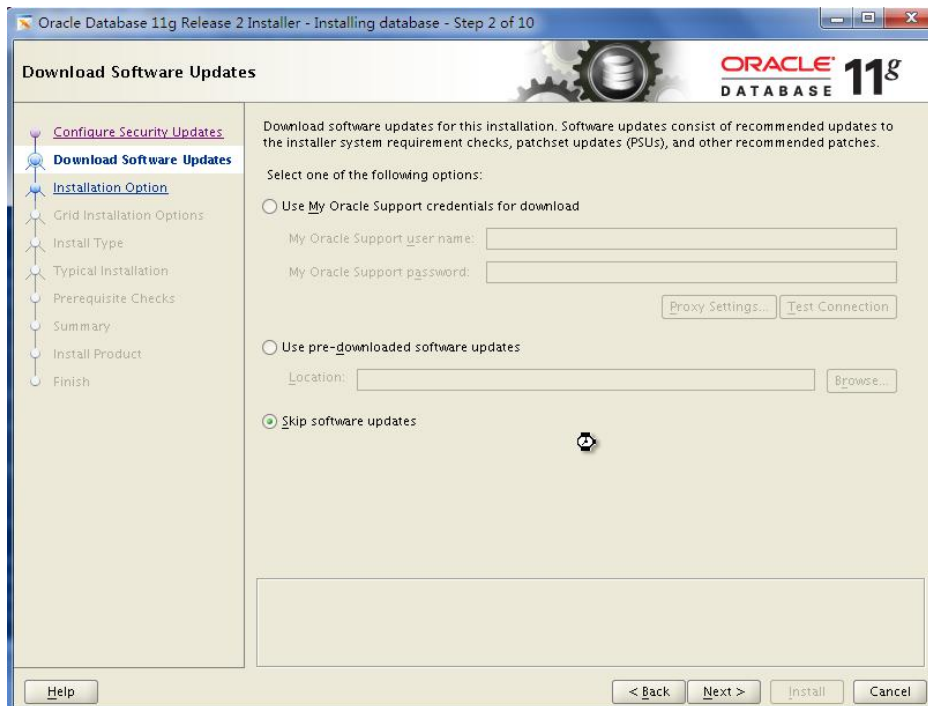


7 安装配置 RAC 数据库

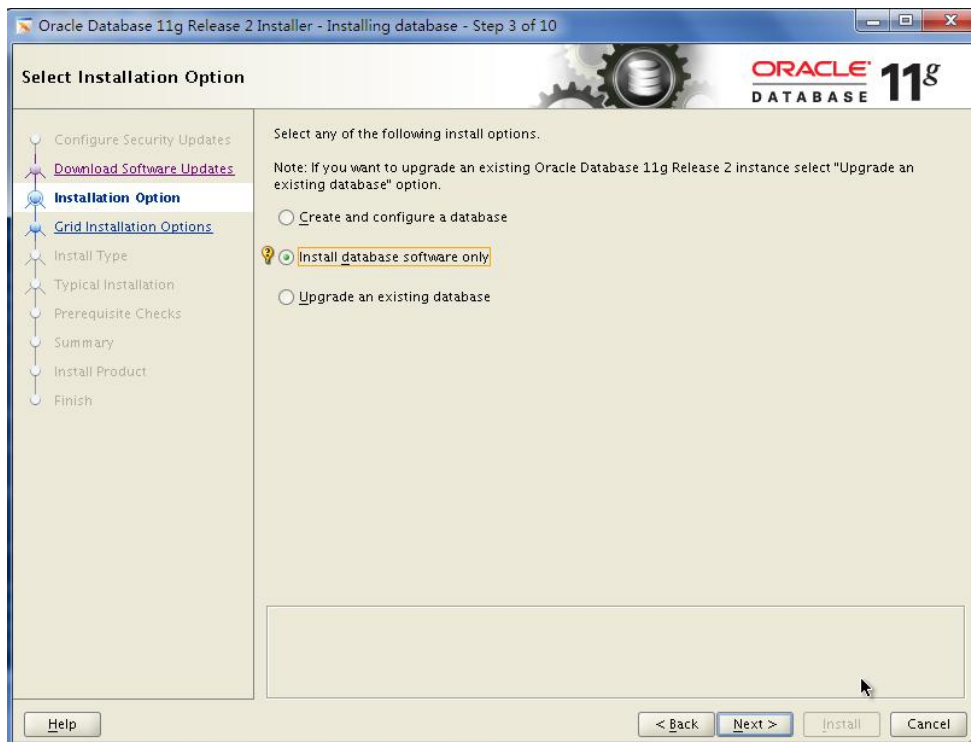
7.1 安装数据库软件

进入到数据库所在目录，执行安装程序

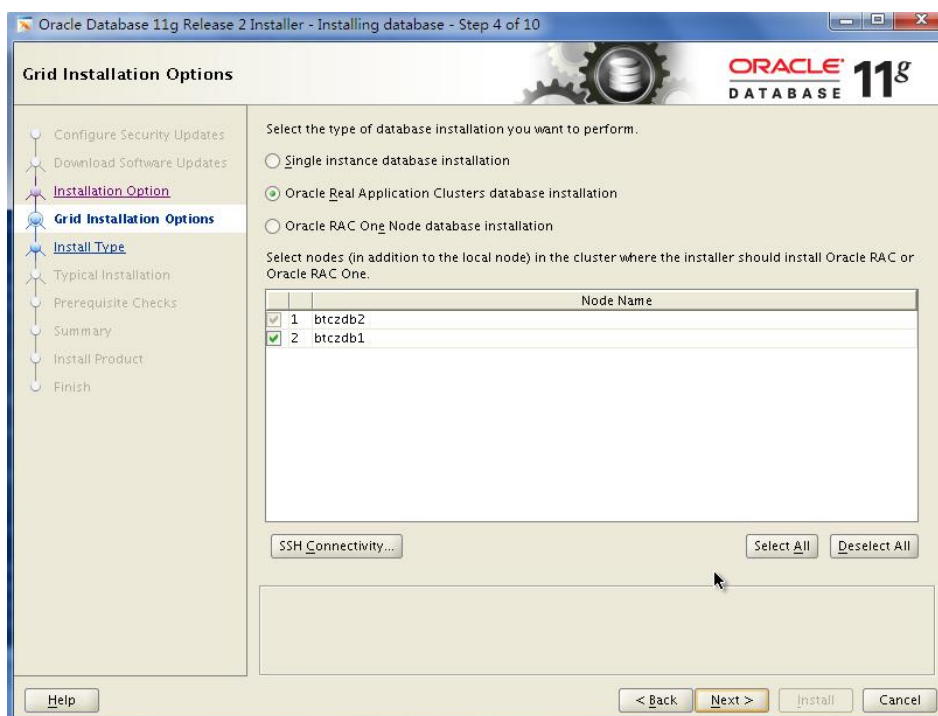
```
$ cd /oracle/software/database  
$ ./runInstaller
```



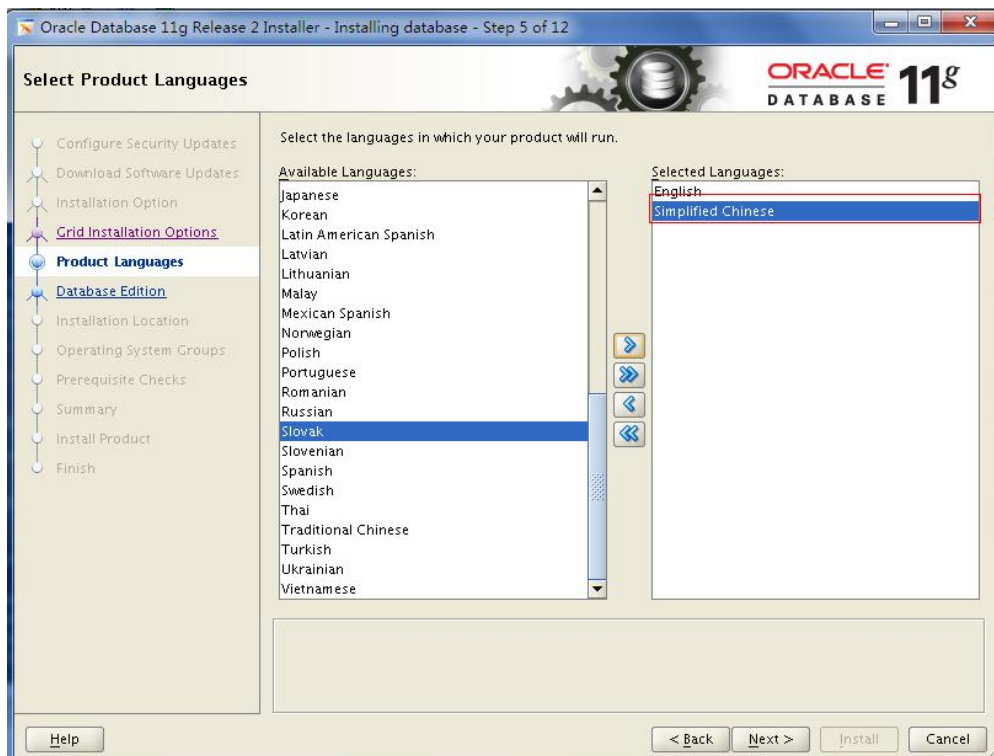
选择跳过软件更新，点击 **Next**，进入下一步



选择只安装数据库软件，点击 **Next**，进入下一步



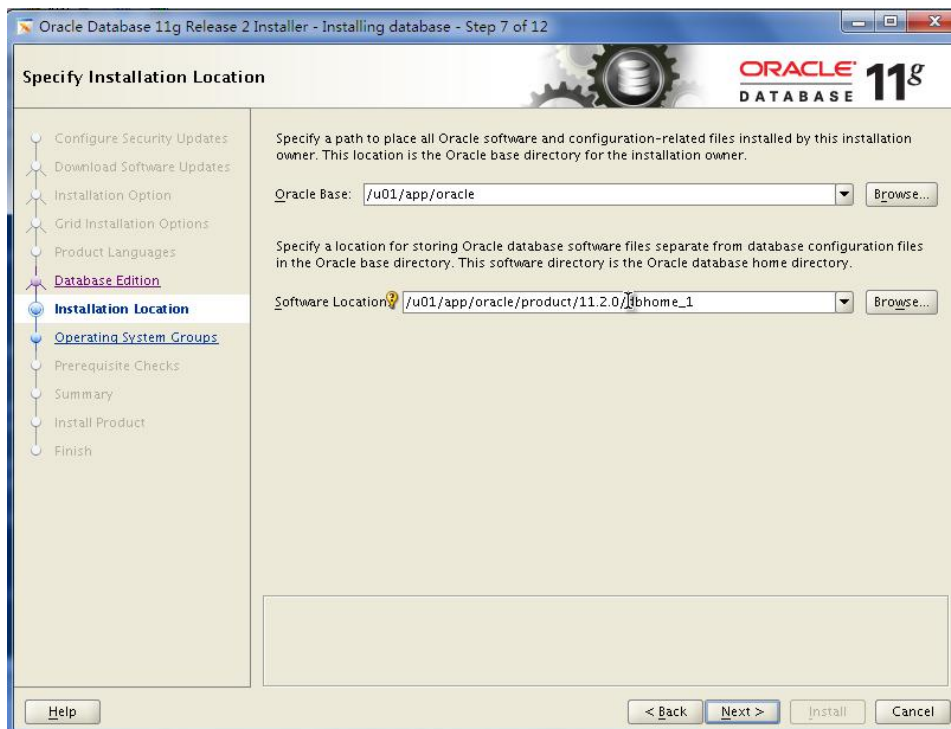
选择数据库安装节点，配置 **ssh** 认证，点击 **Next**，进入下一步



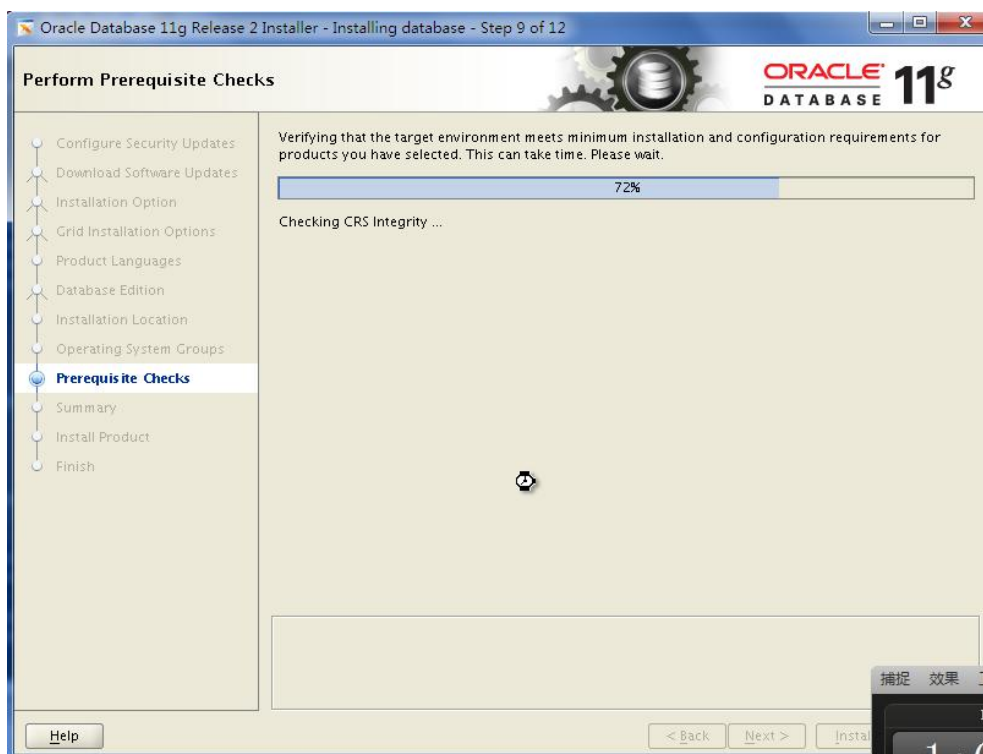
添加中文语言，点击 **Next**，进入下一步



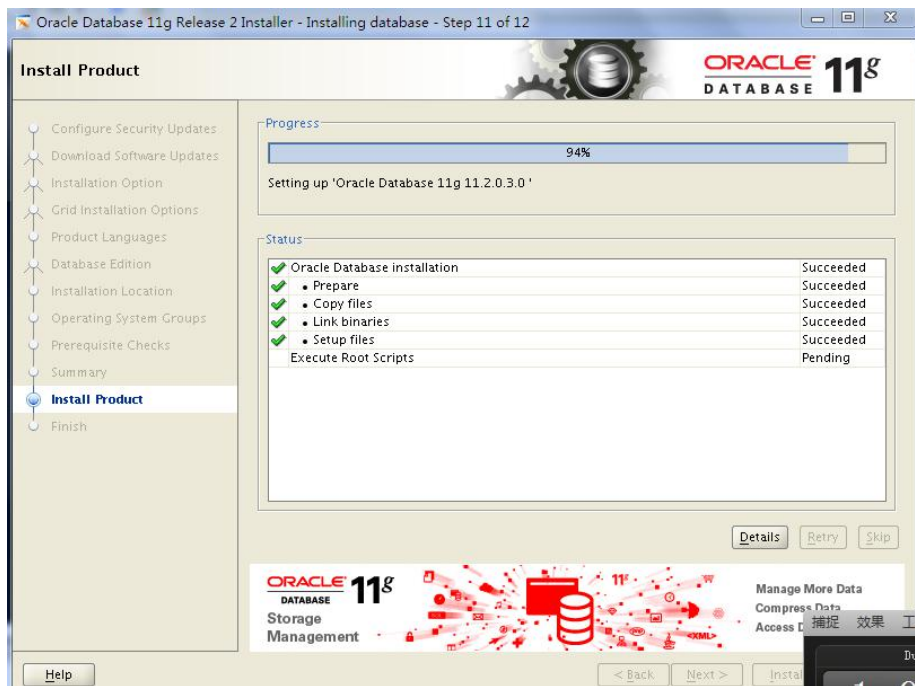
选择安装企业版，点击 **Next**，进入下一步



选择数据库软件安装目录，点击 **Next**，进入下一步



执行安装信息检查

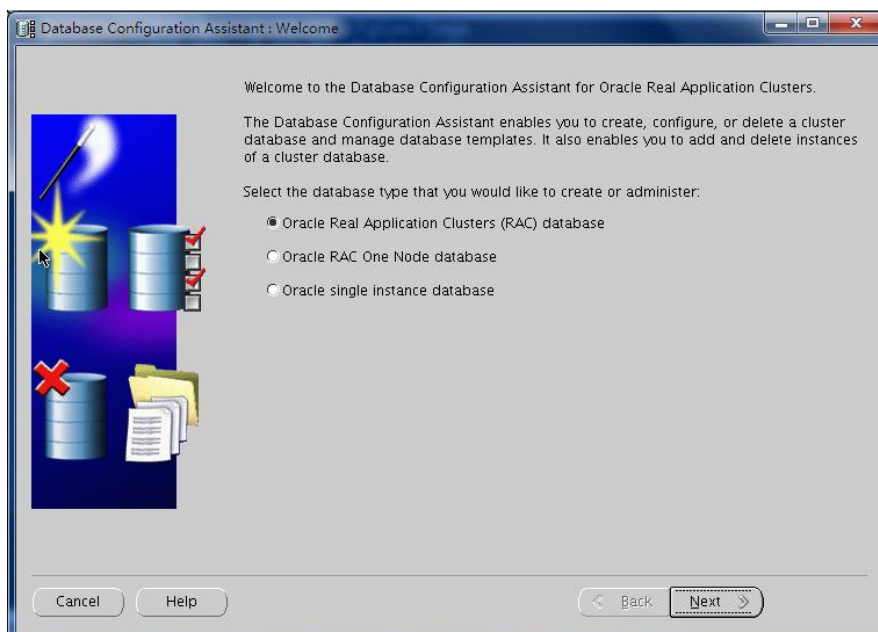


完成数据库软件安装。

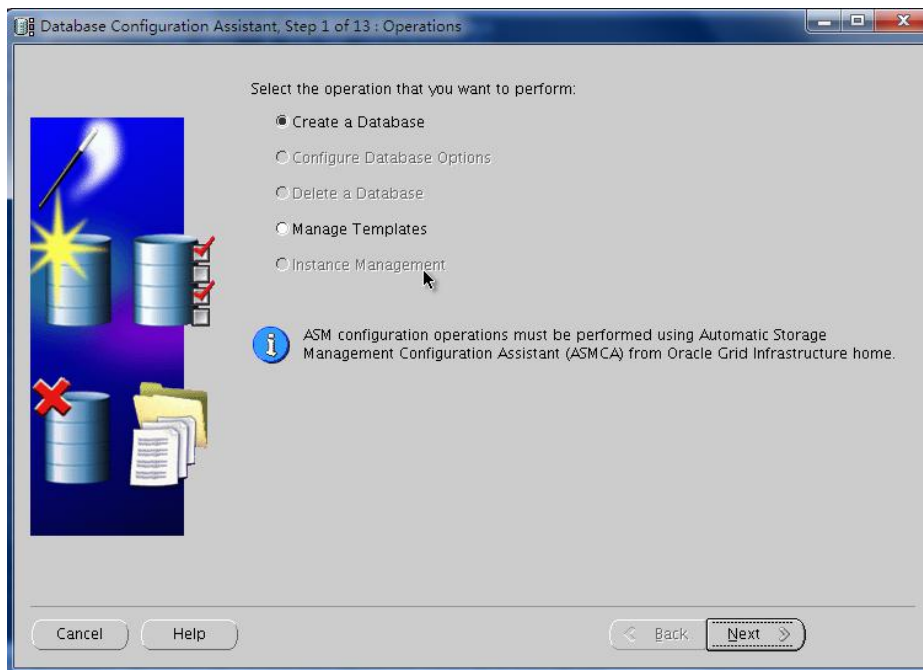
7.2 创建数据库

使用 oracle 用户，运行 dbca 命令，启动数据库创建程序

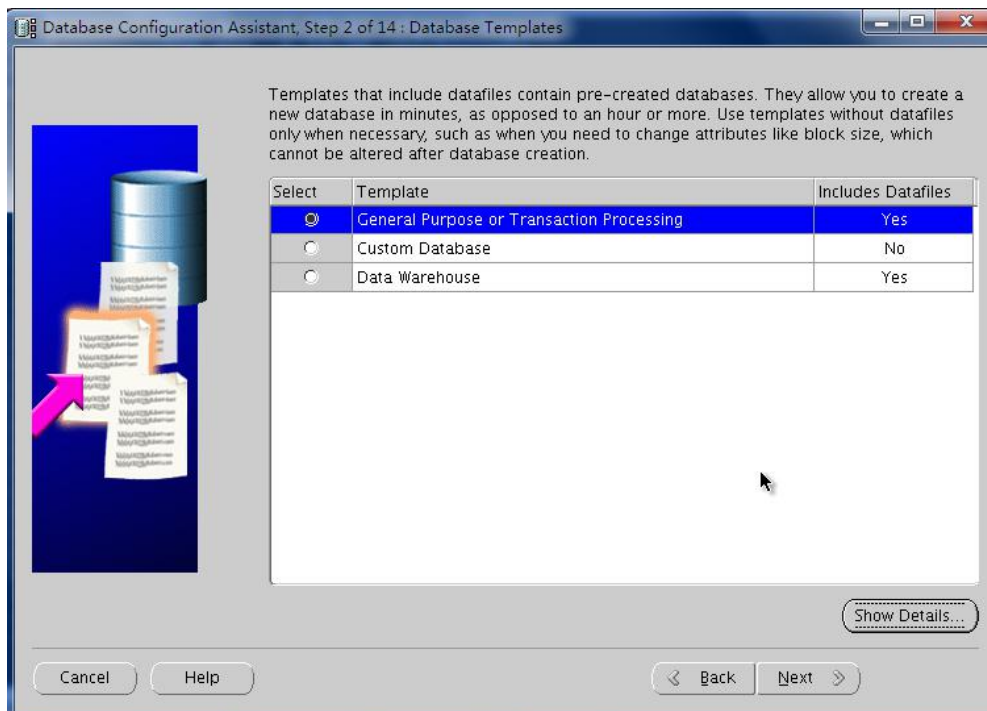
```
$ dbca
```



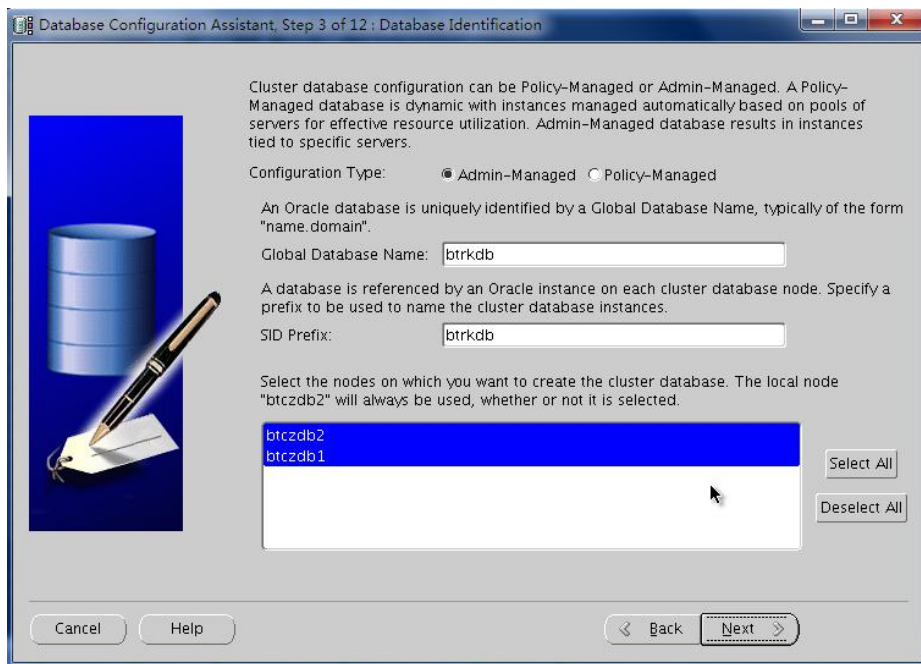
选择集群数据库，点击 Next，进入下一步



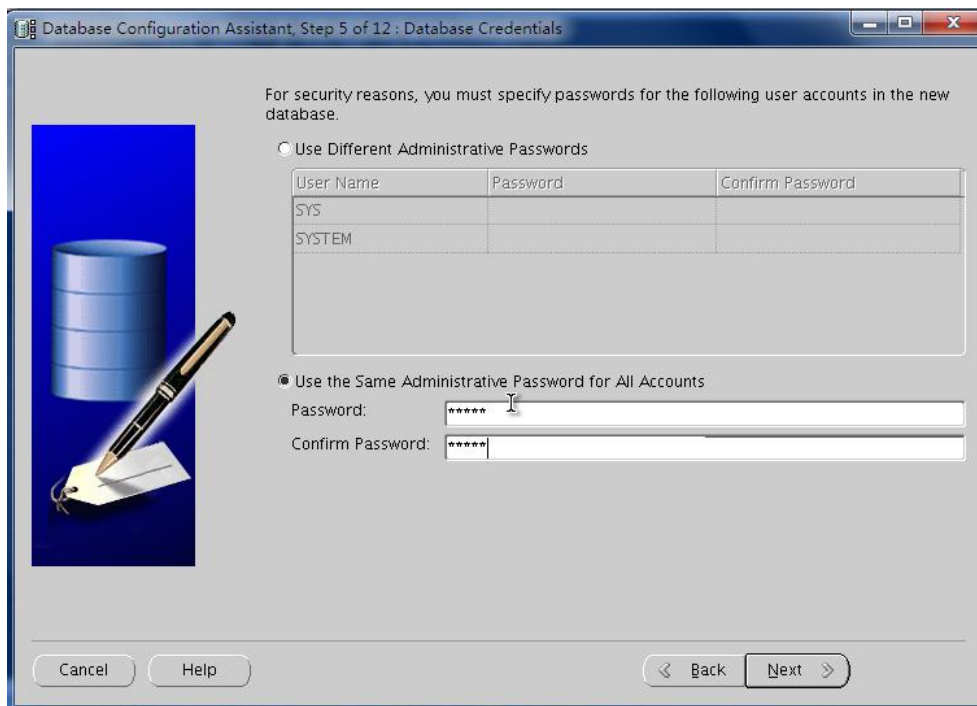
选择创建数据库，点击 Next，进入下一步



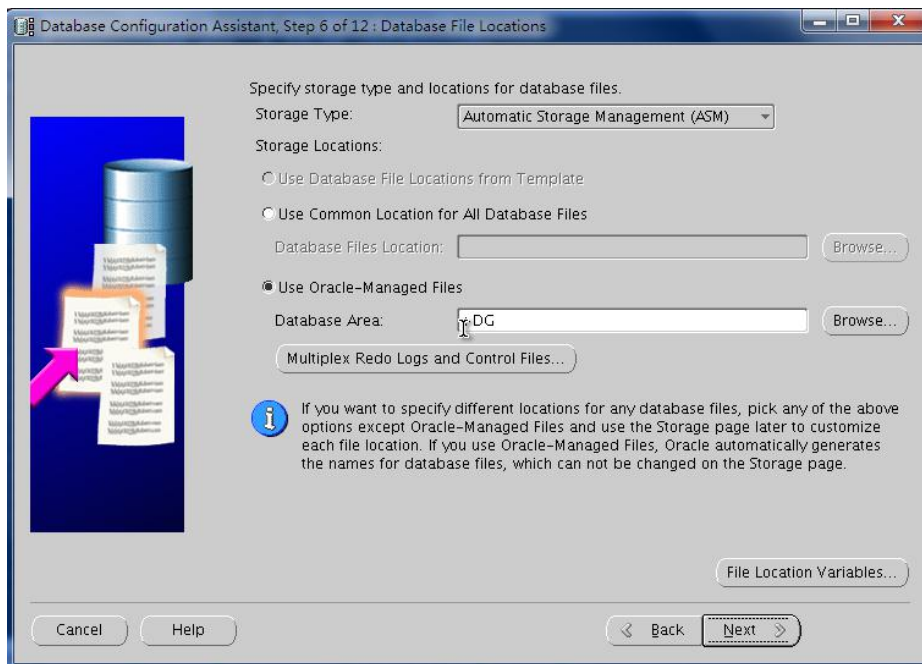
选择创建通用型数据库，点击 Next，进入下一步



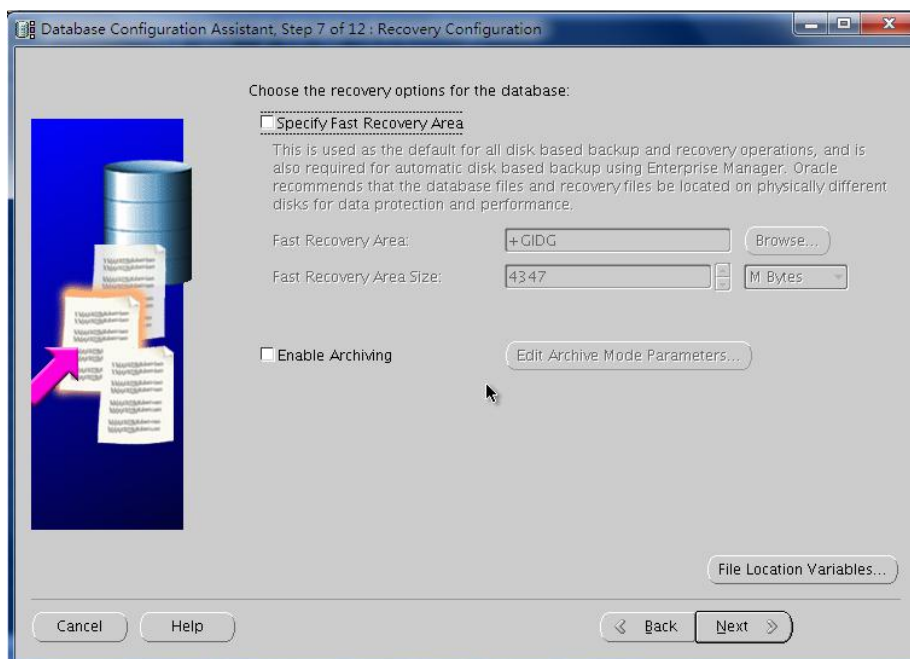
输入数据库名称: btrkdb, 选择所有节点, 点击 Next, 进入下一步



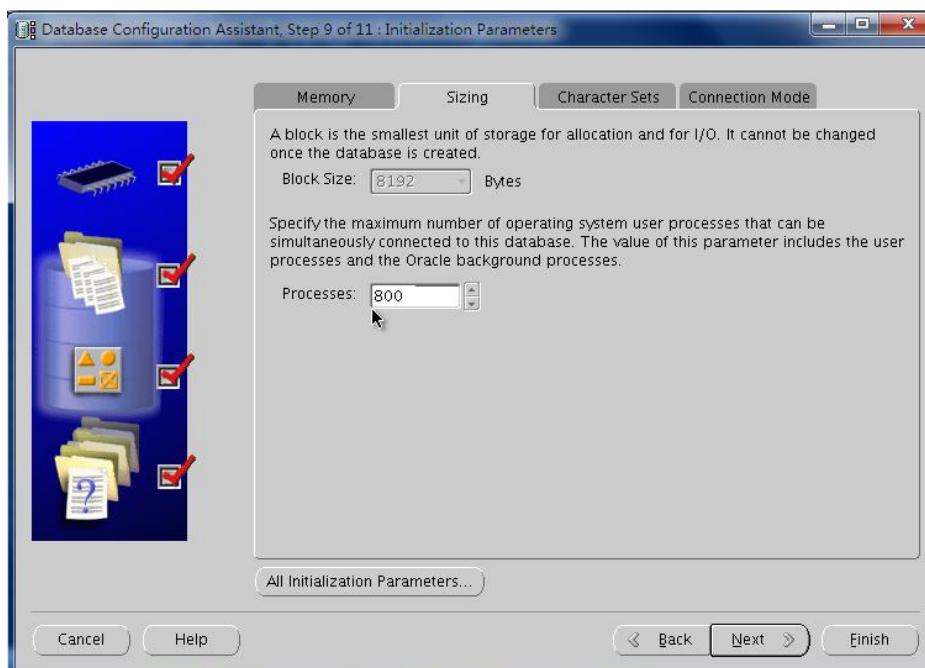
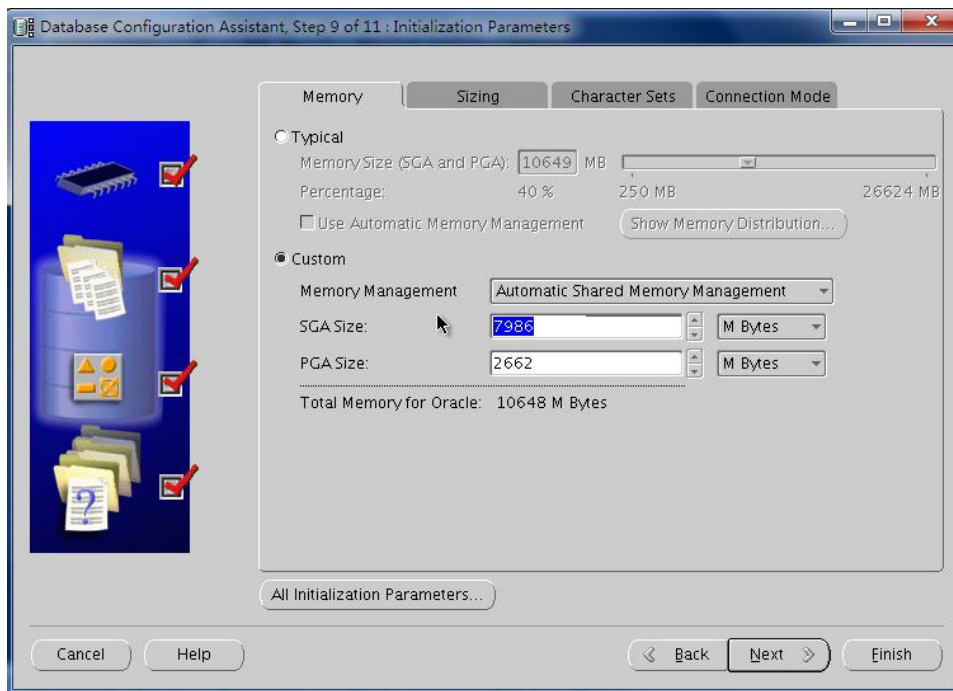
设置数据库用户默认口令为 admin, 点击 Next, 进入下一步

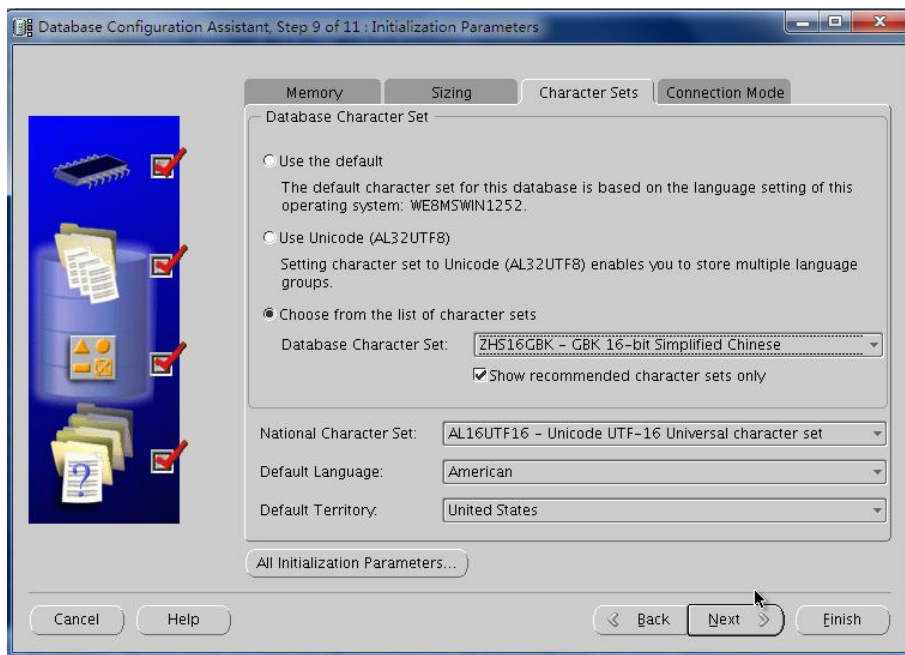


选择数据文件存放磁盘组，点击 Next，进入下一步

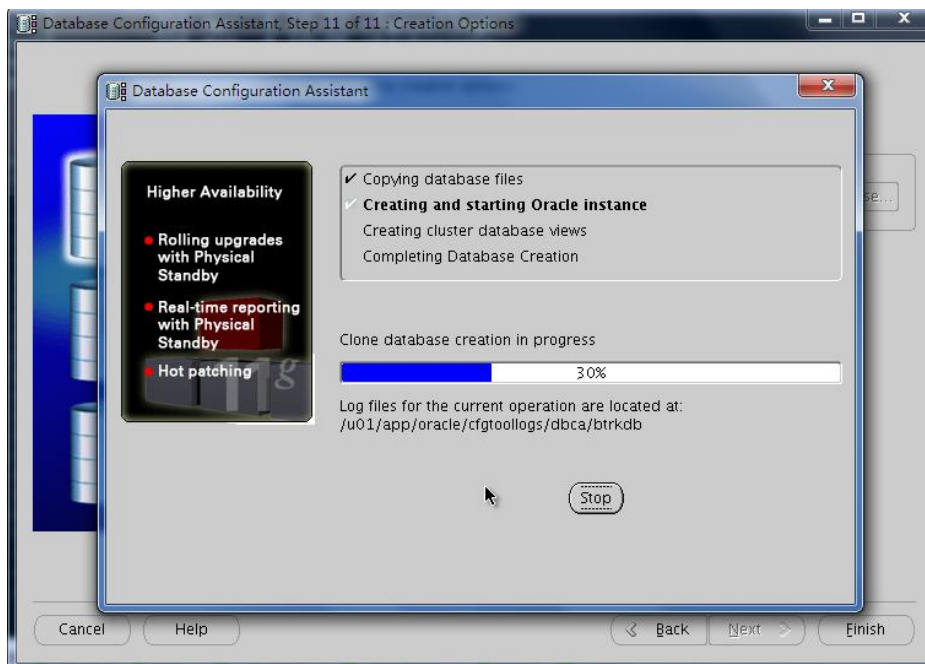


为提高系统性能，取消闪回配置，在完成数据迁移之前，不打开归档，点击 Next，进入下一步

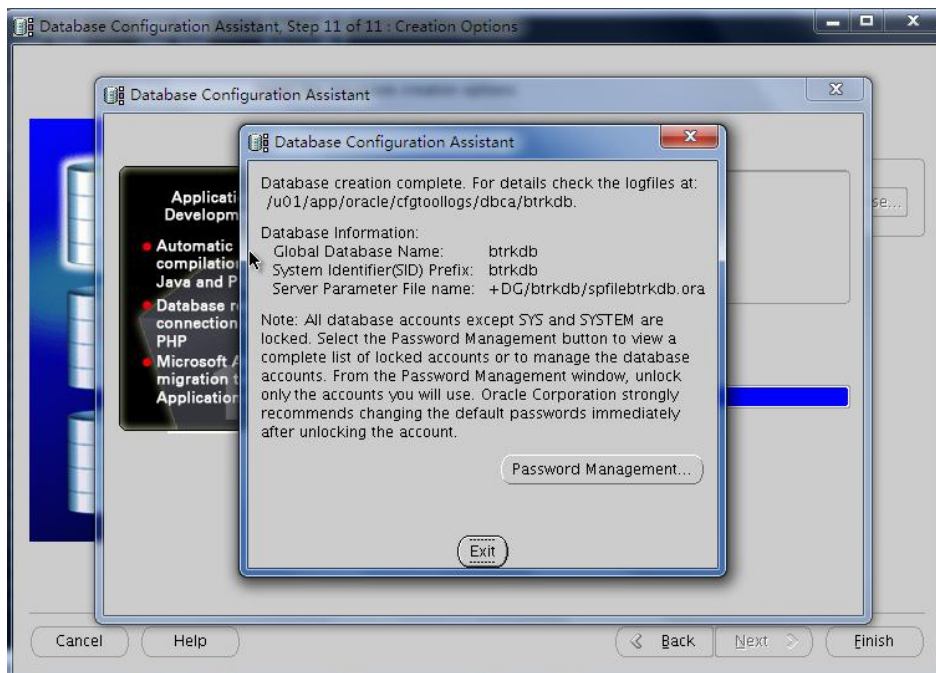




修改内存参数配置，调整进程数，修改数据库字符集，修改完成后，点击 Next，进入下一步



进行数据库创建操作。



完成数据库创建。

8 数据库补丁安装

8.1 Database 11.2.0.4 PSU

Master Note for Database Proactive Patch Program (Doc ID 756671.1)

8.2 USLA heap fix

For AIX 6.1 TL07 SP02/AIX 7.1 TL01 SP02 or later, apply patch 13443029
For AIX 6.1 TL07 or AIX 7.1 TL01, install AIX 6.1 TL-07 APAR IV09580, AIX 7.1 TL-01 APAR IV09541, and apply patch 13443029
For other AIX level, apply patch 10190759, this will disable Oracle's online patching mechanism
New connection can be slow to establish without fix for bug 13494030 which is fixed in 11.2.0.4

9 RAC 数据库日常管理命令

9.1 ORACLE RAC 自动启动

正常情况下每次节点重新启动的时候，RAC 的所有服务都会自动启动。

```
$ crsctl check crs
```

```
CRS-4638: Oracle High Availability Services is online
```

```
CRS-4537: Cluster Ready Services is online
```

```
CRS-4529: Cluster Synchronization Services is online
```

```
CRS-4533: Event Manager is online
```

系统状态为 healthy，表明集群运行正常。

9.2 查看当前 RAC 服务状态

在服务器中用 root 用户执行命令：

```
# . /home/oracle/.profile
```

```
# crs_stat -t
```

```
$ crs_stat -t
```

Name	Type	Target	State	Host
ora.DG.dg	ora....up.type	ONLINE	ONLINE	db01
ora.GIDG.dg	ora....up.type	ONLINE	ONLINE	db01
ora....ER.lsnr	ora....er.type	ONLINE	ONLINE	db01
ora....N1.lsnr	ora....er.type	ONLINE	ONLINE	db01
ora.asm	ora.asm.type	ONLINE	ONLINE	db01
ora....SM2.asm	application	ONLINE	ONLINE	db01
ora....B1.lsnr	application	ONLINE	ONLINE	db01
ora....db1.gsd	application	OFFLINE	OFFLINE	
ora....db1.ons	application	ONLINE	ONLINE	db01
ora....db1.vip	ora....t1.type	ONLINE	ONLINE	db01
ora....SM1.asm	application	ONLINE	ONLINE	db02
ora....B2.lsnr	application	ONLINE	ONLINE	db02
ora....db2.gsd	application	OFFLINE	OFFLINE	
ora....db2.ons	application	ONLINE	ONLINE	db02
ora....db2.vip	ora....t1.type	ONLINE	ONLINE	db02
ora.btrkdb.db	ora....se.type	ONLINE	ONLINE	db01
ora.cvu	ora.cvu.type	ONLINE	ONLINE	db01
ora.gsd	ora.gsd.type	OFFLINE	OFFLINE	
ora....network	ora....rk.type	ONLINE	ONLINE	db01
ora.oc4j	ora.oc4j.type	ONLINE	ONLINE	db01
ora.ons	ora.ons.type	ONLINE	ONLINE	db01
ora....ry.acfs	ora....fs.type	ONLINE	ONLINE	db01
ora.scan1.vip	ora....ip.type	ONLINE	ONLINE	db01

如果所有服务都是 online 的，说明 RAC 数据库状态正常。

9.3 启动和停止 RAC 数据库

9.3.1 停止数据库

在服务器中用 oracle 用户执行命令：

停止数据库

```
[oracle@db01 ~]$ srvctl stop database -d db0
```

停止 nodeapps

```
[oracle@db01 ~]$ srvctl stop nodeapps -n db01  
[oracle@db01 ~]$ srvctl stop nodeapps -n db02
```

9.3.2 启动数据库

在服务器中用 oracle 用户执行命令：

启动 nodeapps

```
[oracle@db01 ~]$ srvctl start nodeapps -n db01  
[oracle@db01 ~]$ srvctl start nodeapps -n db02
```

启动数据库

```
[oracle@db01 ~]$ srvctl start database -d db0
```

9.4 启动和停止 GI

9.4.1 停止 GI:

```
[root@db01 ~]# . /home/oracle/.profile  
[root@db01 ~]# crsctl stop crs
```

9.4.2 启动 GI:

```
[root@db01 ~]# . /home/oracle/.profile  
[root@db01 ~]# crsctl start crs
```

10 参考文档

10.1 IBM 官方文档

10.2 ORACLE 官方文档

Oracle® Database Oracle Clusterware and Oracle Real Application Clusters Installation Guide 10g Release 2 (10.2) for AIX

10.3 MOS 参考文档

10.3.1 文档 ID 169706.1

AIX: Top Things to DO NOW to Stabilize 11gR2 GI/RAC Cluster (Doc ID 1427855.1)

10.3.2 文档 ID 169706.1

Oracle Database (RDBMS) on Unix AIX, HP-UX, Linux, Mac OS X, Solaris, Tru64 Unix Operating Systems Installation and Configuration Requirements Quick Reference (8.0.5 to 11.2) (文档 ID 169706.1)

10.3.3 文档 ID 1307544.1

Certification Information for Oracle Database on IBM AIX on Power systems (文档 ID 1307544.1)

■ 12.1

OS versions and minimum levels:

- 6.1 - minimum TL7 SP3 ("6100-07-03-1207") or later, 64-bit kernel. Oracle recommends TL9 or later.
- 7.1 - minimum TL1 SP3 ("7100-01-03-1207") or later, 64-bit kernel. Oracle recommends TL3 or later.

■ 11.2

OS versions and minimum levels:

- 5.3
- 6.1
- 7.1 - the following APARs are required:
 - ✧ IZ87216
 - ✧ IZ87564
 - ✧ IZ89165

■ 10.2

OS versions and minimum levels

- 5.2
- 5.3
- 6.1

10.3.4 文档 ID 282036.1

Minimum Software Versions and Patches Required to Support Oracle Products on IBM Power Systems (文档 ID 282036.1)

■ AIX 6.1\7.1

IV04047: SHLAP64 UNABLE TO PROCESS ORACLE REQUEST LEADING TO KERNEL HANG
IV16603: SYSTEM CRASH DUE TO FREED SOCKET WHEN SOCKETPAIR() CALL USED

IV16250: MULTICAST UDP PACKETS NOT DELIVERED TO ALL LISTENERS IN WPAR ENV
IV20595 The first retransmit packet may be sent after 64 seconds, (Doc Number=5795)

IV09942: SYSTEM CRASH IN NETINFO_UNIXDOMNLIST

iv08797 SYSTEM CRASH

IZ75919: KRLOCK SERIALIATION ISSUE

IZ76433: KRLOCK SERIALIATION ISSUE APPLIES TO AIX 6100-06

IZ97166: CRASH IN NETINFO_UNIXDOMNLIST WHILE RUNNING NETSTAT APPLIES TO AIX 6100-06

IY83611: AIX CRASHES WHEN USING 16GB PAGES WITH ORACLE

■ AIX 5.3

IY68989: WRITE TO MMAPPED SPACE HANGS 05/09/16 PTF PECHANGE

IY84780: KERNEL MEMORY GARBAGE COLLECTOR FAILS TO FREE LISTS

IY83611: AIX CRASHES WHEN USING 16GB PAGES WITH ORACLE

■ AIX 5.2

IY63133: PERFORMANCE DEGREDATION WITH MANY CPUS AND VOLUME GROUPS

10.3.5 文档 ID 1156586.1

Master Note For Oracle Database Server Installation (文档 ID 1156586.1)

10.3.6 文档 ID 811293.1

RAC and Oracle Clusterware Best Practices and Starter Kit (AIX) (文档 ID 811293.1)

RAC 和 Oracle Clusterware 最佳实践和初学者指南 (AIX) (文档 ID 1526555.1)

■ Tune virtual memory parameters. IBM recommended numbers are:

vmo -p -o maxperm%=90;

vmo -p -o minperm%=3;

vmo -p -o maxclient%=90;

vmo -p -o strict maxperm=0;

vmo -p -o strict maxclient=1;

vmo -p -o lru_file_repage=0;

vmo -r -o page_steal_method=1; (need to reboot to take into effect)

vmo -p -o strict_maxclient=1

vmo -p -o strict_maxperm=0;

■ On AIX 5.3, apply APAR IY84780 to fix a known kernel issue with per-cpu freelists.

■ Set AIXTHREAD_SCOPE=S in the environment: export AIXTHREAD_SCOPE=S for improved performance (default of S on AIX 6.1 and above).

- Ensure that the GI and ORACLE owner account has the CAP_NUMA_ATTACH, CAP_BYPASS_RAC_VMM, and CAP_PROPAGATE capabilities.
#/usr/bin/lsuser -a capabilities grid
#/usr/bin/chuser
capabilities=CAP_NUMA_ATTACH,CAP_BYPASS_RAC_VMM,CAP_PROPAGATE grid

- Do not assign PVIDs (Physical volume IDs) to disks or volumes that are being used for ASM Diskgroups.

- Ensure that the network tuning parameters are set in accordance with the following to ensure optimal interconnect performance:

```
tcp_recvspace = 65536
tcp_sendspace = 65536
udp_sendspace = ((DB_BLOCK_SIZE * DB_FILE_MULTIBLOCK_READ_COUNT) + 4 KB)
but no lower than 65536
udp_recvspace = 655360 (Minimum recommended value is 10x udp_sendspace,
parameter value must be less than sb_max)
rfc1323 = 1
sb_max = 4194304
ipqmaxlen = 512
```