Understanding the In-Depth Technology of Oracle Clusterware and Case Studies to Fix the Clusterware Issues

Amit Das

Principal Database Engineering Architect

Salesforce Database Engineering Team

salesforce

Introducing our talented team members

David Green– Our visionary and inspiring Leader

Chris Kifuthu– Our Very Expert and talented Engineer

Jeff Iskin – Our Master Automation Expert

George Yao– Our Delivery leader, maintain high quality of Service

Saravanan Jayavelu– Our newest addition exadata expert engineer

Bohan Chen – The Project Delivery expert.

Who Am I (past)?

- 3+ years in PayPal/Ebay Inc

- Was Engineering lead at Ebay Inc for the followings

  - Oracle Cluterware, Oracle RAC, Optimization of RAC Reconfiguration, ASM, Exadata, POC for New Oracle Technology
    POC of new Storage, Gold Image build for Patching and New Installation, Patching automation, which reduced 4+ hours downtime to less than 6 minutes downtime Cluster build automation, which has reduced from 2 weeks to 1 day.

- Frequent presenter at

  - Oracle Open World
  - IOUG
  - NoCOUG
  - Delivered lots of training to internal Engineers

Who Am I (past)?

- Before joining Ebay Inc: More than 18 years in IT

- 11 years in Oracle RAC Development team.

- Contributed more than 50 Oracle products
  Architect of Oracle Autopatching tool
  Developed Onecommad to simplify RAC/GRID installation
  Helped more than 20 big customers to implement and successes in Oracle RAC.

- Tech Lead on Global Exadata Project for the following big accounts
  Apple : World first Exadata Production Go-live, while at Oracle. Commonwealth Bank Australia : (On their largest charge back system)

  Rakuten Japan : Largest Online Retail System
  Canadian Imperial Bank : Internal Analytics

Agenda

- Architectural Overview

- Grid Infrastructure Processes

- Case studies for Multiple CRS issues

- Q&A

Grid Infrastructure Overview

- GI introduced in Oracle 9.2 with a special release

- Oracle Clusterware version 10.1 - major release with increased functionality

- ASM Home was outside of Clusterware on 10.1

- ASM Home integrated with Oracle Clustware from 11gR1

- 11gR2 is the most stable and integrated version of Oracle Clusterware

Grid Infrastructure Overview

- Grid Infrastructure consists of:
    - Oracle Cluster Ready Services (CRS)
    - Oracle Automatic Storage Management (ASM)
- The Grid Home contains the software for both products
- CRS can also be Standalone for ASM and/or Oracle Restart.
- CRS can run by itself or in combination with other vendor clusterware
- Grid Home and RDBMS home must be installed in different locations
    - The installer locks the Grid Home path by setting root permissions.

salesforce

Grid Infrastructure Overview

- CRS requires shared Oracle Cluster Registry (OCR) and Voting files
  - Must be in ASM or CFS (raw not supported for install)
  - OCR backed up every 4 hours automatically to $GIHOME/cdata
  - OCR backup kept 4, 8,12 hours, 1 day, 1 week
  - Restored with ocrconfig
  - Voting file backed up into OCR at each change.
  - Voting file restored with crsctl

Grid Infrastructure Overview

- Network requirements for CRS:
  - One high speed, low latency, redundant private network for inter node communications
  - Should be a separate physical network
  - VLANS are supported with restrictions
  - Used for :-
    - Clusterware messaging
    - RDBMS messaging and block transfer
    - ASM messaging

Grid Infrastructure Overview

- Network requirements for CRS:
  - Standard set up Public Network
    - One Public IP and VIP per node in DNS
    - One Scan name set up in DNS.
  - Or let Grid Naming Service (GNS) set up Public Network
    - One Public IP per node (recommended)
    - One GNS VIP per cluster
    - DHCP allocation of hostnames

Grid Infrastructure Overview

- Single Client Access Name (SCAN)
    - Single name for clients to access Oracle Databases running in a cluster
    - Cluster alias for databases in the cluster
    - Provides load balancing and failover for client connections to database
    - Cluster topology changes do not require client configuration changes
    - Allows clients to use the EZConnect client and the simple JDBC thin URL for transparent access to any database running in the cluster
        - **sqlplus system/manager@sales1-scan:1521/oltp**
        - **jdbc:oracle:thin:@sales1-scan:1521/oltp**

salesforce

Grid Infrastructure Overview

- SCAN in the Cluster
  - Each SCAN IP has a SCAN listener dispersed across the cluster.
  - [oracle@mynode] **srvctl config scan_listener**

    SCAN Listener LISTENER_SCAN1 exists. Port: TCP:1521
    SCAN Listener LISTENER_SCAN2 exists. Port: TCP:1521
    SCAN Listener LISTENER_SCAN3 exists. Port: TCP:1521
  - [oracle@mynode] **srvctl config scan**

    SCAN name: sales1-scan, Network: 1/133.22.67.0/255.255.255.0/
    SCAN VIP name: scan1, IP: /sales1-scan.example.com/133.22.67.192
    SCAN VIP name: scan2, IP: /sales1-scan.example.com/133.22.67.193
    SCAN VIP name: scan3, IP: /sales1-scan.example.com/133.22.67.194

Grid Infrastructure Overview

- Only one set of Clusterware daemons can run on each node
- The CRS stack all spawns from Oracle HA Services Daemon (ohasd)
- On Unix ohasd runs out of inittab with respawn
- A node can be evicted when deemed unhealthy
  - May require node reboot
  - Will require CRS stack restart (rebootless restart)
- CRS provides Cluster Time Synchronization services
  - Always runs; If ntpd configured, runs in "observer" mode
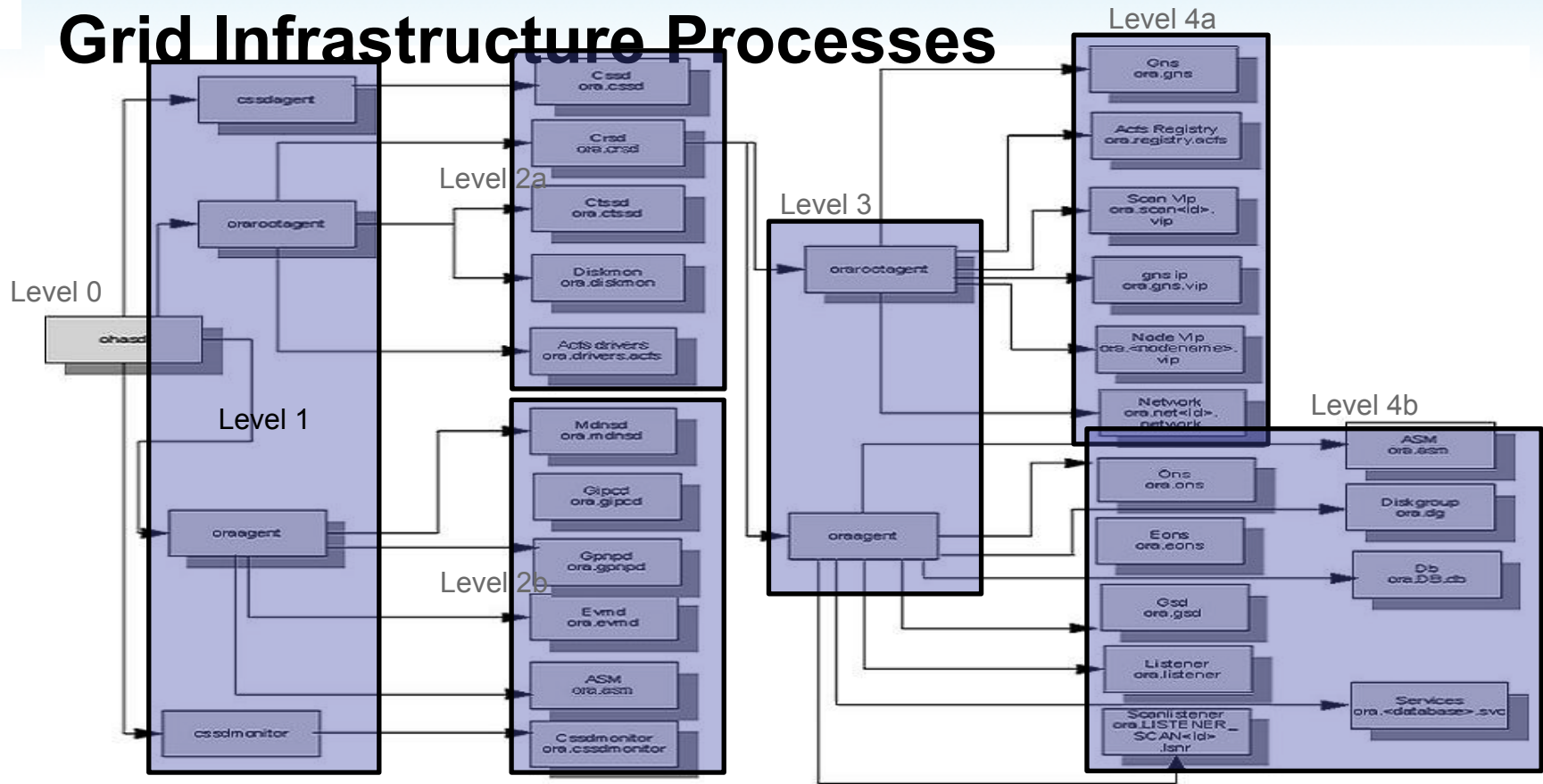
Grid Infrastructure Overview

- Nodes only lease a node number
  - Not guaranteed for stack to always start with same node number
  - Only way to influence numbering is at first install/upgrade, and then ensure nodes remain fairly active (almost true)
  - Pre 11.2 databases cannot handle leased node numbers
    - Pin node numbers – only allows pinning to current leased number

- CRS stack should be started/stopped on boot/shutdown by init or
  - crsctl start/stop crs for local clusterware stack
  - crsctl start/stop cluster for all nodes (ohasd must be running)

Grid Infrastructure processes

11.2 Agents changed everything!

- Multi-threaded Daemons

- Manage multiple resources and types

- Implements entry points for multiple resource types
  - Start, stop check, clean, fail

- oraagent, orarootagent, application agent, script agent, cssdagent

- Single process started from init on Unix (ohasd)

- Diagram in next page shows all core resources

salesforce

Grid Infrastructure Processes

Grid Infrastructure Processes: Init Scripts

- /etc/init.d/ohasd (location O/S dependent)
  - RC script with "start" and "stop" actions
  - Initiates Oracle Clusterware autostart
  - Control file coordinates with CRSCTL
- /etc/init.d/init.ohasd (location O/S dependent)
  - OHASD Framework Script runs from init/upstart
  - Control file coordinates with CRSCTL

Linux

```
h1:3:respawn:/etc/init.d/init.ohasd run >/dev/null 2>&1 </dev/null
```

Grid Infrastructure Processes
Startup Sequence 11gR2

- Level 1: OHASD Spawns:
  - cssdagent - Agent responsible for spawning CSSD
  - orarootagent - Agent responsible for managing all root owned ohasd resources
  - oraagent - Agent responsible for managing all oracle owned ohasd resources
  - cssdmonitor - Monitors CSSD and node health (along with the cssdagent)

Grid Infrastructure Processes

Startup Sequence 11gR2.

- Level 2a: OHASD oraagent spawns:
  - Mdnsd – Multicast DNS daemon
  - GIPCD – Grid IPC Daemon
  - GpnpD – Grid Plug and Play Daemon
  - EVMD – Event Monitor Daemon

salesforce

Grid Infrastructure Processes
Startup Sequence 11gR2.

- Level 2b: OHASD rootagent spawns:

  - ➢ CTSSD - Cluster Time Synchronization Services Daemon

  - ➢ Diskmon (Exadata)

  - ➢ ASM – ASM instance started here as may be required by CRSD

  - ➢ ACFS (ASM Cluster File System) Drivers

  - ➢ CRSD - Primary daemon responsible for managing cluster resources

Grid Infrastructure Processes

Startup Sequence 11gR2.

- Level 2b: OHASD rootagent spawns:
  - ➢ CTSSD - Cluster Time Synchronization Services Daemon
  - ➢ Diskmon (Exadata)
  - ➢ ASM – ASM instance started here as may be required by CRSD
  - ➢ ACFS (ASM Cluster File System) Drivers
  - ➢ CRSD - Primary daemon responsible for managing cluster resources

Grid Infrastructure Processes

Startup Sequence 11gR2.

- Level 3: CRSD spawns:
  - orarootagent - Agent responsible for managing all root owned crsd resources
  - oraagent - Agent responsible for managing all nonroot owned crsd resources. One is spawned for every user that has CRS ressources to manage

Grid Infrastructure Processes

Startup Sequence 11gR2.

- Level 4: CRSD oraagent spawns:
  - ASM Resouce - ASM Instance(s) resource (proxy resource)
  - Diskgroup - Used for managing/monitoring ASM diskgroups.
  - DB Resource - Used for monitoring and managing the DB and instances
  - SCAN Listener - Listener for single client access name, listening on SCAN VIP
  - Listener - Node listener listening on the Node VIP
  - Services - Used for monitoring and managing services
  - ONS - Oracle Notification Service
  - eONS - Enhanced Oracle Notification Service (pre 11.2.0.2)
  - GSD - For 9i backward compatibility
  - GNS (optional) - Grid Naming Service - Performs name resolution

Grid Infrastructure Processes

ohasd managed resources

Resource Name	Agent Name	Owner
ora.gipcd	oraagent	crs user
ora.gpnpd	oraagent	crs user
ora.mdnsd	oraagent	crs user
ora.cssd	cssdagent	root
ora.cssdmonitor	cssdmonitor	root
ora.diskmon	orarootagent	root
ora.ctssd	orarootagent	root
ora.evmd	oraagent	crs user
ora.crsd	orarootagent	root
ora.asm	oraagent	crs user
ora.driver.acfs	orarootagent	root

Cluster Ready Services Daemon

➢ Primarily responsible for maintaining the availability of application resources

➢ It is responsible for starting and stopping these resources

➢ Relocating them when required to another node in the event of failure

➢ Maintaining the resource profiles in the OCR (Oracle Cluster Registry).

➢ It is responsible for overseeing the caching of the OCR for faster access

➢ Responsible for backing up the OCR.

➢ Log file is GI_HOME/log/<node>/crsd/crsd.log

- Rotation policy 10MB
- Retention policy 10 logs

Grid Infrastructure Processes

CRSD's oraagent manages

➢ all database, instance, service and diskgroup resources node listeners,SCAN listeners, and ONS

➢ If the Grid Infrastructure owner is different from the RDBMS home owner then you would have 2 oraagents each running as one of the installation owners.

➢ The database, and service resources would be managed by the RDBMS home owner and other resources by the Grid Infrastructure home owner.

➢ Log file is GI_HOME/log/<node>/agent/crsd/oraagent_<user>/ oraagent_<user>.log

Grid Infrastructure Processes

CRSD orarootagent

> CRSD's rootagent manages
>> GNS and it's VIP
>> Node VIP
>> SCAN VIP
>> network resources.

> Log file is GI_HOME/log/<node>/agent/crsd/ orarootagent_root/oraagent_root.log

salesforce

Grid Infrastructure Processes

Multicast Domain Name Service Daemon (mDNS(d))

➤ Used by Grid Plug and Play to locate profiles in the cluster, as well as by GNS to perform name resolution.

➤ The mDNS process is a background process on Linux and UNIX and on Windows.

➤ Uses multicast for cache updates on service advertisement arrival/departure.

➤ Advertises/serves on all found node interfaces.

➤ Log is GI_HOME/log/<node>/mdnsd/mdnsd.log

Grid Infrastructure Processes

Grid Plug 'n' play daemon (gpnp(d))

- ➢ Provides access to the Grid Plug and Play profile
- ➢ Coordinates updates to the profile from clients among the nodes of the cluster
- ➢ Ensures all nodes have the most recent profile
- ➢ Registers with mdns to advertise profile availability
- ➢ Log is GI_HOME/log/<node>/gpnpd/gpnpd.log

Grid Infrastructure Processes

cssd agent and monitor

- Same functionality in both agent and monitor
- Functionality of several pre-11.2 daemons consolidated in both
  - OPROCD – system hang
  - OMON – oracle clusterware monitor
  - VMON – vendor clusterware monitor
- Run realtime with locked down memory, like CSSD
- Provides enhanced stability and diagnosability
- Logs are
  - GI_HOME/log/<node>/agent/oracssdagent_root/oracssdagent_root.log
  - GI_HOME/log/<node>/agent/oracssdmonitor_root/oracssdmonitor_root.log

salesforce

Grid Infrastructure Processes

cssd agent and monitor

- ➢ oprocd

  - ❖ The basic objective of both OPROCD and OMON was to ensure that the perceptions of other nodes was correct

    - ● If CSSD failed, other nodes assumed that the node would fail within a certain amount of time and OMON ensured that it would

    - ● If the node hung for long enough, other nodes would assume that it was gone and OPROCD would ensure that it was gone

  - ❖ The goal of the change is to do this more accurately and avoid false terminations
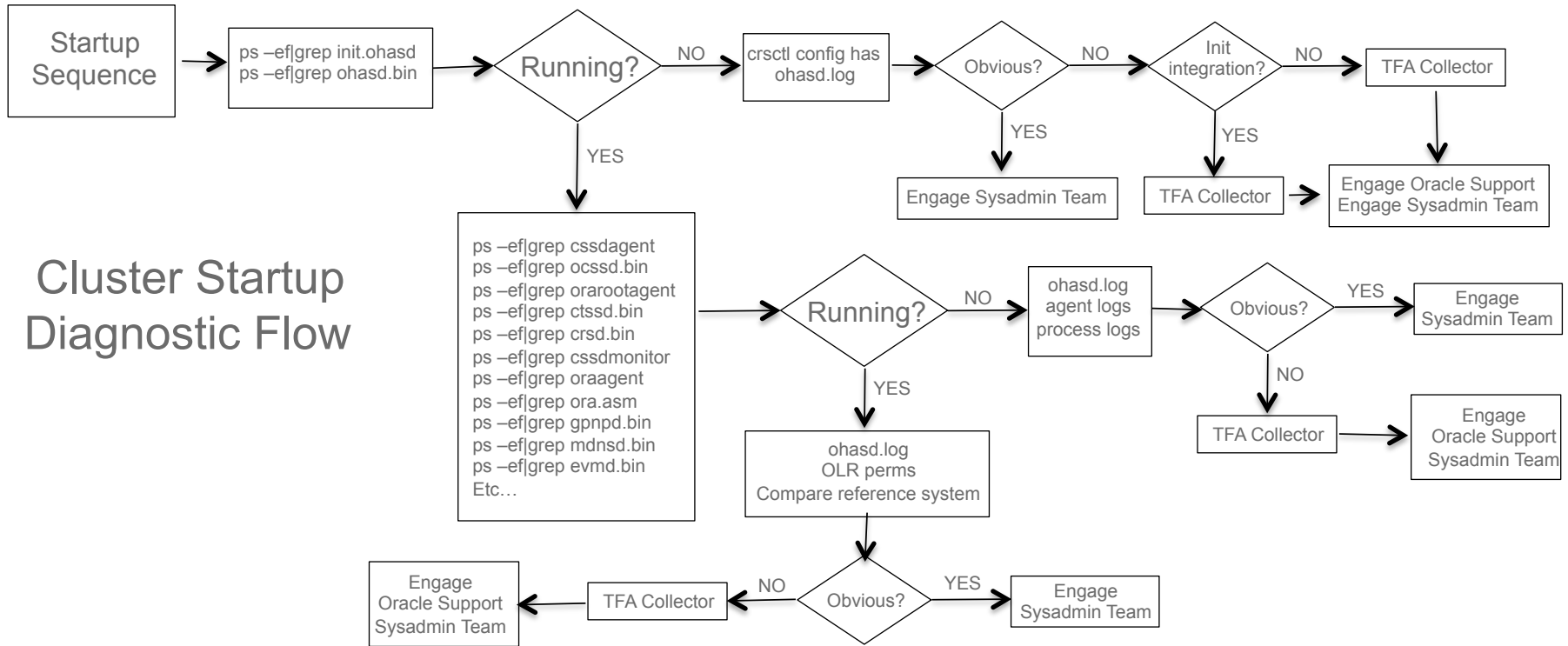
Grid Infrastructure Processes

Cluster Time Synchronisation Services daemon

❑Provides time management in a cluster for Oracle.

➢Observer mode when Vendor time synchronisation s/w is found

● Logs time difference to the CRS alert log

➢Active mode when no Vendor time sync s/w is found

Troubleshooting Scenarios

Cluster Startup Problem Triage (11.2+)

Startup Sequence → ps –ef|grep init.ohasd / ps –ef|grep ohasd.bin → Running?

- NO → crsctl config has ohasd.log → Obvious?
  - YES → Engage Sysadmin Team
  - NO → Init integration?
    - YES → TFA Collector → Engage Oracle Support / Engage Sysadmin Team
    - NO → TFA Collector → Engage Oracle Support / Engage Sysadmin Team
- YES → (Cluster Startup Diagnostic Flow)

Cluster Startup Diagnostic Flow

ps –ef|grep cssdagent
ps –ef|grep ocssd.bin
ps –ef|grep orarootagent
ps –ef|grep ctssd.bin
ps –ef|grep crsd.bin
ps –ef|grep cssdmonitor
ps –ef|grep oraagent
ps –ef|grep ora.asm
ps –ef|grep gpnpd.bin
ps –ef|grep mdnsd.bin
ps –ef|grep evmd.bin
Etc…

→ Running?

- NO → ohasd.log / agent logs / process logs → Obvious?
  - YES → Engage Sysadmin Team
  - NO → TFA Collector → Engage Oracle Support / Sysadmin Team
- YES → ohasd.log / OLR perms / Compare reference system → Obvious?
  - YES → Engage Sysadmin Team
  - NO → TFA Collector → Engage Oracle Support / Sysadmin Team

salesforce

Troubleshooting Scenarios

## Cluster Startup Problem Triage

<?xml version="1.0" encoding="UTF-8"?>

<gpnp:GPnP-Profile Version="1.0" xmlns="http://www.grid-pnp.org/2005/11/gpnp-profile" xmlns:gpnp="http://www.grid-pnp.org/2005/11/gpnp-profile" xmlns:orcl="http://www.oracle.com/gpnp/2005/11/gpnp-profile" xmlns:xsi="http://www.w3.org/2001/XMLSchema-instance" xsi:schemaLocation="http://www.grid-pnp.org/2005/11/gpnp-profile gpnp-profile.xsd" **ProfileSequence="6" ClusterUId="b1eec1fcdd355f2bbf7910ce9cc4a228" ClusterName="staij-cluster"** PALocation="">

<gpnp:Network-Profile><gpnp:HostNetwork id="gen" HostName="*">
<gpnp:Network id="net1" IP="140.87.152.0" Adapter="eth0" Use="public"/>
<gpnp:Network id="net2" IP="140.87.148.0" Adapter="eth1" Use="cluster_interconnect"/>
</gpnp:HostNetworkcss"></gpnp:Network-Profile>

<orcl:CSS-Profile id=" DiscoveryString="+asm" LeaseDuration="400"/>

<orcl:ASM-Profile id="asm" DiscoveryString="" SPFile="+SYSTEM/staij-cluster/asmparameterfile/registry.
253.693925293"/>

<ds:Signature xmlns:ds="http://www.w3.org/2000/09/xmldsig#"><ds:SignedInfo><ds:CanonicalizationMethod Algorithm="http://www.w3.org/2001/10/xml-exc-c14n#"/><ds:SignatureMethod Algorithm="http://www.w3.org/2001/10/xml-exc-c14n#"> <InclusiveNamespaces xmlns="http://www.w3.org/2001/10/xml-exc-c14n#" PrefixList="gpnp orcl xsi"/></ds:Transform></ds:Transforms><ds:DigestMethod Algorithm="http://www.w3.org/2000/09/xmldsig#sha1"/><ds:DigestValue>x1H9LWjyNyMn6BsOykHhMvxnP8U=</ds:DigestValue></ds:Reference></ds:SignedInfo><ds:SignatureValue>N+20jG4=</ds:SignatureValue>**</ds:Signature>**

</gpnp:GPnP-Profile>

Troubleshooting Scenarios

- Agent return codes
  - Check entry must return one of the following return codes:
    - ONLINE
    - UNPLANNED_OFFLINE: Target=online, may be recovered failed over
    - PLANNED_OFFLINE:
    - UNKNOWN: Cannot determine, if previously online, partial then monitor
    - PARTIAL: Some of a resources services are available. Instance up but not open.
    - FAILED: Requires clean action

Installation Diagnostics and Troubleshooting

Install/Upgrade Scenario Process Flow
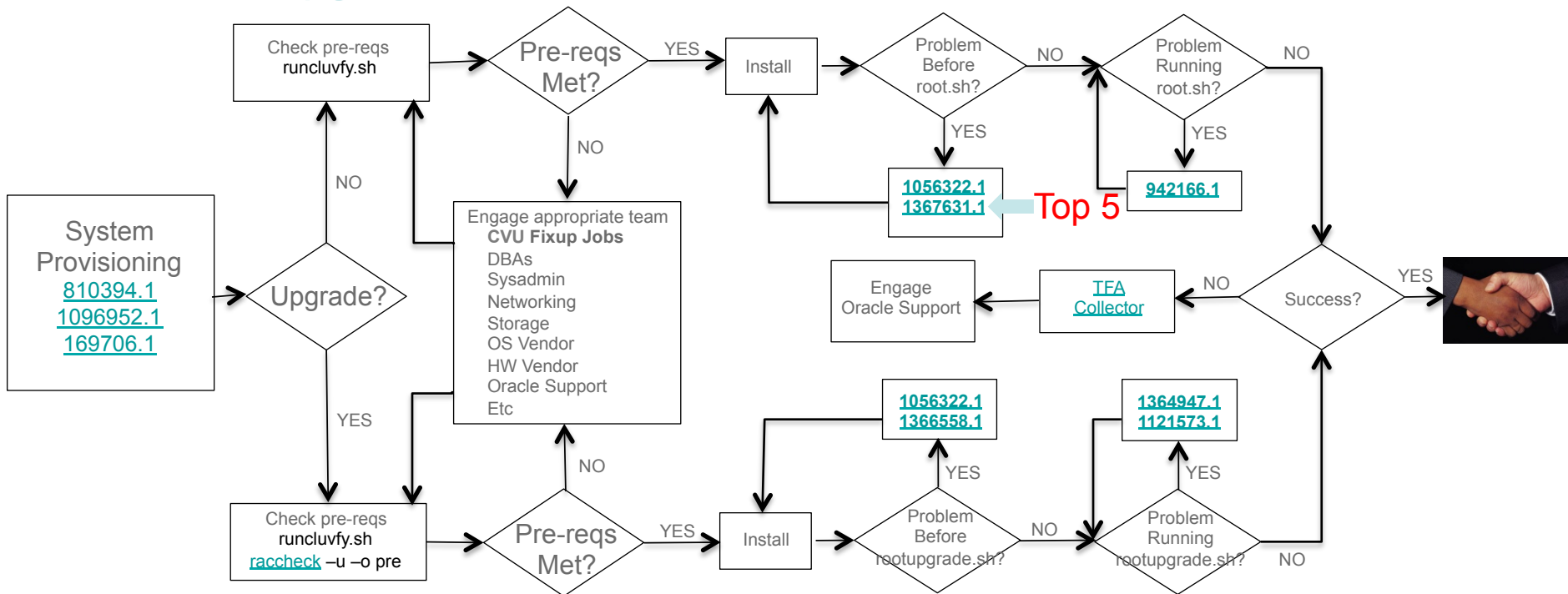
Installation Diagnostics and Troubleshooting
Top 5 CRS/Grid Infrastructure Install issues

- # 1: root.sh fails to startup 11.2 GI stack due to known defects
  - Symptom
    - GI install failure when running root.sh
  - Cause
    - Known issues for which fixes already exist
  - Solution
    - In-flight application of most recent PSU
      - Proceed with install up to step requiring running root.sh
      - Before running root.sh script apply PSU
      - In general you'll want the latest PSUs anyway but this step may help avoid problems
    - For upgrades run ./raccheck –u –o pre prior to beginning
      - Checks for pre-req patches

Installation Diagnostics and Troubleshooting

Top 5 CRS/Grid Infrastructure Install issues

(Case Study#2)

Cluster Startup Problem Triage, after upgrade Upgrade the cluster from 11.2.0.3 to 11.2.0.4 on Exadata

– CSSD is UP, ASM is UP, but CRSD is not starting

[ohasd(46497)]CRS-2765:Resource 'ora.crsd' has failed on server 'db01'.

[crsd(66199)]CRS-1013:The OCR location in an ASM disk group is inaccessible. Details in /oracle/db/grid/11204/log/exa01/crsd/crsd.log.

[crsd(66199)]CRS-0804:Cluster Ready Service aborted due to Oracle Cluster Registry error [PROC-26: Error while accessing the physical storage ORA-15077: could not locate ASM instance serving a required diskgroup]. Details at (:CRSD00111:) in /oracle/db/grid/11204/log/exa01/crsd/crsd.log.

[ohasd(46497)]CRS-2765:Resource 'ora.crsd' has failed on server 'db01'.

Installation Diagnostics and Troubleshooting

Top 5 CRS/Grid Infrastructure Install issues

(Case Study#2) Cluster Startup Problem Triage

```
<?xml version="1.0" encoding="UTF-8"?>
<gpnp:GPnP-Profile Version="1.0" xmlns="http://www.grid-pnp.org/2005/11/gpnp-profile" xmlns:gpnp="http://www.grid- pnp.org/2005/11/gpnp-profile" xmlns:orcl="http://www.oracle.com/gpnp/2005/11/gpnp-profile" xmlns:xsi="http://www.w3.org/2001/XMLSchema-instance" xsi:schemaLocation="http://www.grid- pnp.org/2005/11/gpnp-profile gpnp-profile.xsd" ProfileSequence="6" ClusterUId="b1eec1fcdd355f2bbf7910ce9cc4a228" ClusterName="lali-cluster" PALocation="">
<gpnp:Network-Profile><gpnp:HostNetwork id="gen" HostName="*"> <gpnp:Network id="net1" IP="180.87.152.0" Adapter="eth0" Use="public"/> </gpnp:HostNetworkcss"></gpnp:Network-Profile>
<orcl:CSS-Profile id=" DiscoveryString="+asm" LeaseDuration="400"/>
<orcl:ASM-Profile id="asm" DiscoveryString="" SPFile="+SYSTEM/staij- cluster/asmparameterfile/registry.253.693925293"/>
<ds:Signature xmlns:ds="http://www.w3.org/2000/09/xmldsig#"><ds:SignedInfo><ds:CanonicalizationMethod Algorithm="http://www.w3.org/2001/10/xml-exc-c14n#"/><ds:SignatureMethod Algorithm="http://www.w3.org/2001/10/xml-exc-c14n#"> <InclusiveNamespaces xmlns="http://www.w3.org/2001/10/xml-exc-c14n#" PrefixList="gpnp orcl xsi"/></ds:Transform></ds:Transforms><ds:DigestMethod Algorithm="http://www.w3.org/2000/09/xmldsig#sha1"/><ds:DigestValue>x1H9LWjyNyMn6BsOykHhMvxnP8U=</d s:DigestValue></ds:Reference></ds:SignedInfo><ds:SignatureValue>N+20jG4=</ds:SignatureValue></ds:Signat ure>
</gpnp:GPnP-Profile>
```

Installation Diagnostics and Troubleshooting

Top 5 CRS/Grid Infrastructure Install issues

(Case Study#2) Cluster Startup Problem Triage
Most important configuration file : $GI_HOME/crs/install/crsconfig_params
vi $GI_HOME/crs/install/crsconfig_params

CRS_NODEVIPS="node1-vip/255.255.255.0/bondeth1,node2-vip/
255.255.255.0/bondeth1" NODELIST=node1,node2
NETWORKS="bondeth1"/172.30.6.0:public,"bondib1"/
192.168.5.0:cluster_interconnect ASM_DISK_GROUP=DBFS_FRA
ASM_DISCOVERY_STRING=o/*/* ASM_DISKS=o/192.168.10.11...........
ASM_REDUNDANCY=EXTERNAL

Installation Diagnostics and Troubleshooting

Top 5 CRS/Grid Infrastructure Install issues

(Case Study#2) Cluster Startup Problem Triage
Deconfig the node

$GI_HOME/crs/install/rootcrs.pl –deconfig –force

Reconfig the node (will rebuild correct gpnp profile)

$GI_HOME/root.sh

Validate the upgrade

crsctl query crs activeversion

Lists the Oracle Clusterware operating version

crsctl query crs releaseversion

Lists the Oracle Clusterware release version

Installation Diagnostics and Troubleshooting

Top 5 CRS/Grid Infrastructure Install issues

(Case Study#2) Cluster Startup Problem Triage

```
<?xml version="1.0" encoding="UTF-8"?>
<gpnp:GPnP-Profile Version="1.0" xmlns="http://www.grid-pnp.org/2005/11/gpnp-profile" xmlns:gpnp="http://www.grid- pnp.org/2005/11/gpnp-profile" xmlns:orcl="http://www.oracle.com/gpnp/2005/11/gpnp-profile" xmlns:xsi="http://www.w3.org/2001/XMLSchema-instance" xsi:schemaLocation="http://www.grid- pnp.org/2005/11/gpnp-profile gpnp-profile.xsd" ProfileSequence="6" ClusterUId="b1eec1fcdd355f2bbf7910ce9cc4a228" ClusterName="lali-cluster" PALocation="">
<gpnp:Network-Profile><gpnp:HostNetwork id="gen" HostName="*">
<gpnp:Network id="net1" IP="180.87.152.0" Adapter="eth0" Use="public"/>
<gpnp:Network id="net2" IP="180.87.148.0" Adapter="eth1" Use="cluster_interconnect"/> </gpnp:HostNetworkcss"></gpnp:Network-Profile>
<orcl:CSS-Profile id=" DiscoveryString="+asm" LeaseDuration="400"/>
<orcl:ASM-Profile id="asm" DiscoveryString="" SPFile="+SYSTEM/staij- cluster/asmparameterfile/registry.253.693925293"/>
<ds:Signature xmlns:ds="http://www.w3.org/2000/09/xmldsig#"><ds:SignedInfo><ds:CanonicalizationMethod Algorithm="http://www.w3.org/2001/10/xml-exc-c14n#"/><ds:SignatureMethod Algorithm="http://www.w3.org/2001/10/xml-exc-c14n#"> <InclusiveNamespaces xmlns="http://www.w3.org/2001/10/xml-exc-c14n#" PrefixList="gpnp orcl xsi"/></ds:Transform></ds:Transforms><ds:DigestMethod Algorithm="http://www.w3.org/2000/09/xmldsig#sha1"/><ds:DigestValue>x1H9LWjyNyMn6BsOykHhMvxnP8U=</d s:DigestValue></ds:Reference></ds:SignedInfo><ds:SignatureValue>N+20jG4=</ds:SignatureValue></ds:Signat ure>
```

Installation Diagnostics and Troubleshooting

Top 5 CRS/Grid Infrastructure Install issues

(Case Study#3) Node Eviction Triage Due to time drift

- Cluster Time Synchronisation Services daemon
  - Provides time management in a cluster for Oracle.

- Observer mode when Vendor time synchronisation s/w is found
  - Logs time difference to the CRS alert log

- Active mode when no Vendor time sync s/w is found
- Check the alert log of Clusterware for the time drift message

Installation Diagnostics and Troubleshooting

Top 5 CRS/Grid Infrastructure Install issues

(Case Study#4) Node Eviction Triage

- Node dies with following error

  [cssd(22589)]CRS-1611:Network communication with node node2 (2) missing for 50% of timeout interval. Removal of this node from cluster in 14.535 seconds

  [cssd(22589)]CRS-1611:Network communication with node node2 (2) missing for 75% of timeout interval. Removal of this node from cluster in 7.505 seconds

  [cssd(22589)]CRS-1610:Network communication with node kwhq2 (2) missing for 90% of timeout interval. Removal of this node from cluster in 2.520 seconds

Installation Diagnostics and Troubleshooting

Top 5 CRS/Grid Infrastructure Install issues

Troubleshooting Scenarios (Case Study#4) Node Eviction Triage

Network HB Setting

– Default setting 30 sec for non-Exadata.
– In Exadata, it has set to 60 sec.
– crsctl get css misscount

CRS-4678: Successful get misscount 60 for Cluster Synchronization Services.

Installation Diagnostics and Troubleshooting

(Case Study#4a) ASM Eviction Triage

Using Infiniband network for CRS/ASM/DB
    Lost network communication between ASM instances.
    It is split brain situation for ASM (may be with DB)
    ASM instance died with ORA-29740

• Why split brain noticed in ASM not in Oracle Clusterware, though the CRS and ASM are using same network?

Installation Diagnostics and Troubleshooting

Troubleshooting Scenarios (Case Study#4a) ASM Eviction Triage

- RCA for ASM crash
  - Oracle Clusterware use TCP/IP over IB.
  - ASM is using RDS over IB
- HANG observed due to a bug in RDS layer.
  - Not affected the Oracle Clusterware as it is using TCP
  - ASM is the victim here
  - DB did not crash
- Before DB instance eviction, ASM evicted faster.
- DB was using UDP due to misconfiguration.

salesforce

Installation Diagnostics and Troubleshooting

Top 5 CRS/Grid Infrastructure Install issues

(Case Study#4b) DB Instance Eviction Triage

- Using Infiniband network for CRS/ASM/DB
- Lost network communication between DB instances.
    - It is split brain situation for DB only
    - DB instance died with ORA-29740

• Why split brain noticed in DB neither in Oracle Clusterware, nor in ASM, though the CRS/ASM/DB are using same network?

salesforce

Installation Diagnostics and Troubleshooting

Troubleshooting Scenarios (Case Study#4b) ASM Eviction Triage

RCA for DB Instance crash
- Oracle Clusterware use TCP/IP over IB.
- ASM is using UDP over IB
- DB is using RDS over IB

HANG observed due to a bug in RDS layer.
- Not affected the Oracle Clusterware as it is using TCP
- Not affected to the ASM as it is using UDP
- DB crashed, as it is using RDS

Installation Diagnostics and Troubleshooting

Top 5 CRS/Grid Infrastructure Install issues

CRS/Grid Infrastructure Install issues (Case Study#5)
- # 5: Installation fails because network requirements aren't met
- Symptom
  - o Clusterware startup problems
  - o Individual clusterware component startup problems
- Cause
  - o Improper network configuration for public and/or private network
- Solution
  - o Prior to installation Run cluvfy
- How to Validate Network and Name Resolution Setup for the Clusterware and RAC (Doc ID 1054902.1)
- Grid Infrastructure Startup During Patching, Install or Upgrade May Fail Due to Multicasting Requirement (Doc ID 1212703.1)

Installation Diagnostics and Troubleshooting

Top 5 CRS/Grid Infrastructure Install issues

CRS/Grid Infrastructure Install issues (Case Study#6)

HAIP bug is blocking us to use HAIP in Solaris (11.2.0.2)

- Symptom
  - o Node will be evicted due to ONE Private network failure
- Cause
  - o It is a combination of IPMP setup vs CRS
- Solution
  - o Disable HAIP, but how?

crsctl stat res res ora.cluster_interconnect.haip –init –p

crsctl modify res ora.cluster_interconnect.haip -attr

ENABLED=0 –init

- Run the same on all node
  - Stop CRS
  - Start CRS

salesforce

Installation Diagnostics and Troubleshooting

Top 5 CRS/Grid Infrastructure Install issues

CRS/Grid add node issues (Case Study#7)

ASM binary on new node compiled with different IPC compare to existing node in the cluster.

- Symptom
    - CRS start will fail, as ASM instance startup failed
- Cause

    IPC of all node must be same
    Error will be like "ORA-15160: rolling migration internal fatal error in module SKGXP,valNorm:not-native"
- Syntax to validate IPC is "skgxpinfo"