

RAC Startup sequence

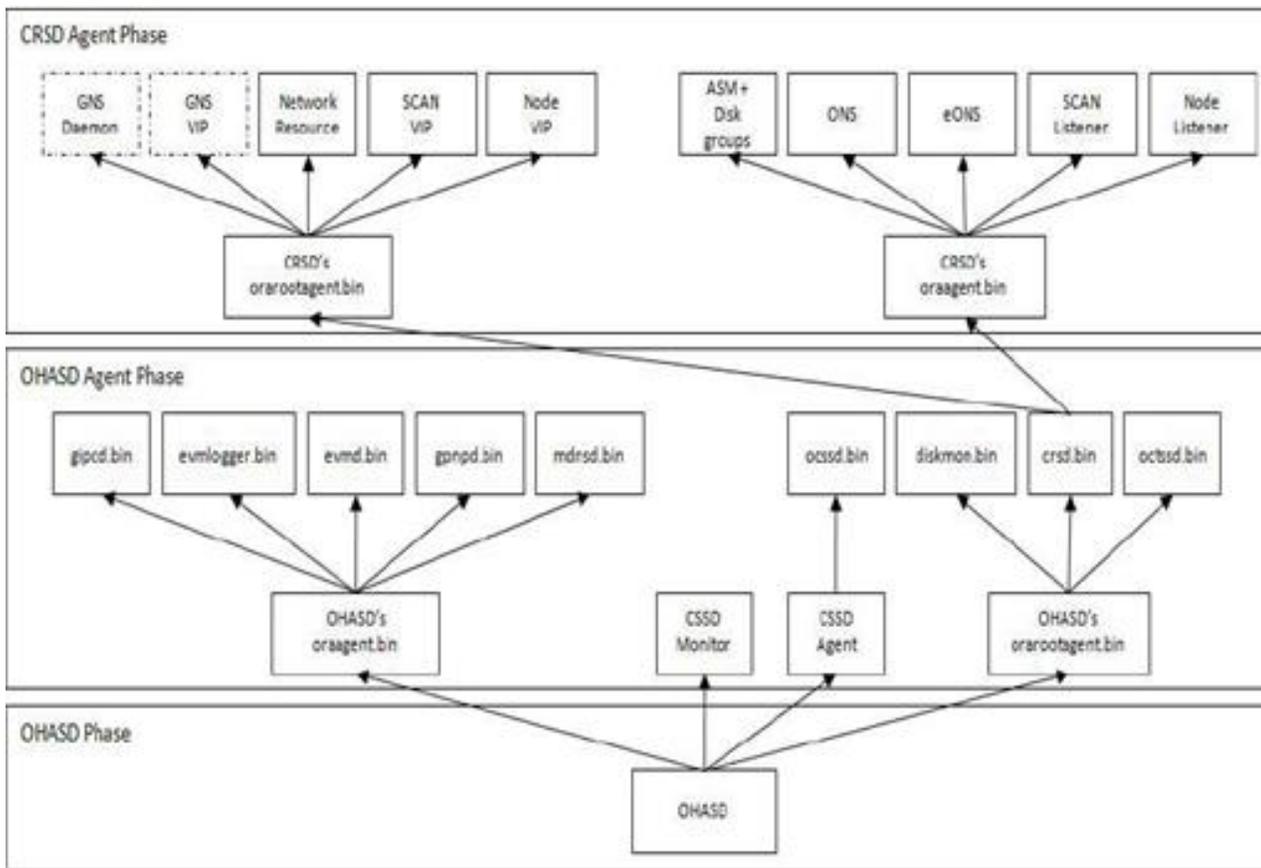
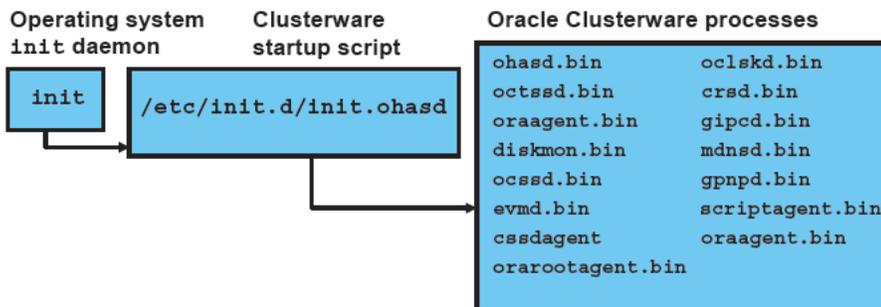
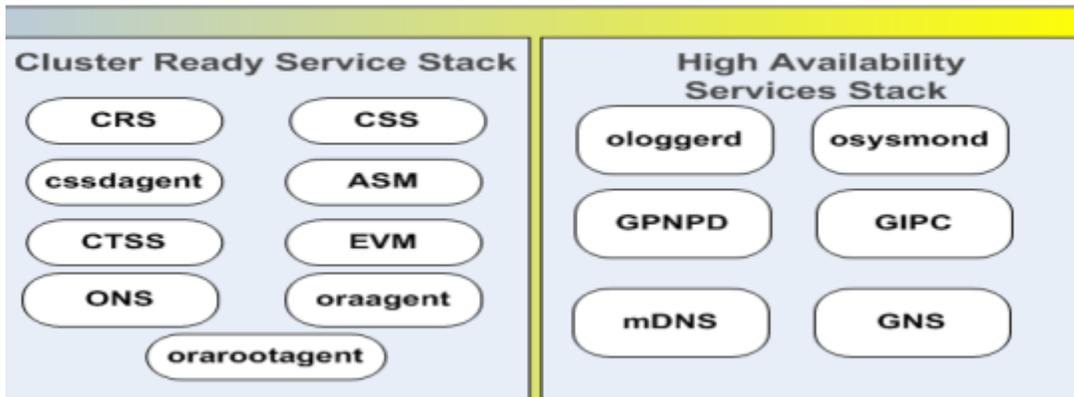


Figure 8-1. The Grid Infrastructure software stack and process dependencies



Oracle 11g R2 Clusterware Stack



Starting in Oracle 11g Real Application Clusters (RAC), we see a new process called Oracle High Availability Services or OHAS for short.

OHASD Phase:-

OHASD (Oracle High Availability Server Daemon) starts Firsts and it will start

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).

OHASD Agent Phase:-

Level 2: OHASD oraagent spawns:

- MDNSD(ora.mdnsd) - Used for DNS lookup
- GIPCD(ora.gipcd) - Used for inter-process and inter-node communication
- GPMPD(ora.gmpd) - Grid Plug & Play Profile Daemon
- EVMD(ora.evmd) - Event Monitor Daemon
- ASM(ora.asm) - Resource for monitoring ASM instances

- **OHASD Agent** starts and in turn this will start

gipcd	Grid interprocess communication daemon, used for monitoring cluster interconnect
mdnsd	Multicast DNS service It resolves DNS requests on behalf of GNS
gns	The Grid Naming Service (GNS), a gateway between DNS and mdnsd, resolves DNS requests
gpnpd	Grid Plug and Play Daemon, Basically a profile similar like OCR contents stored in XML format in \$GI_HOME/gpnp/profiles/<peer> etc., this is where used by OCSSD also to read the ASM disk locations to start up with out having ASM to be up, moreover this also provides the plug and play profile where this can be distributed across nodes to cluster
evmd/ evmlogger	Evm service will be provided by evmd daemon, which is a information about events happening in cluster, stop node, start node, start instance etc.

- **cssdagent** (cluster synchronization service agent), in turn starts

ocssd	Cluster synchronization service daemon which manages node membership in the cluster
-------	---

If cssd found that ocssd is down, it will reboot the node to protect the [data integrity](#).

- **cssdmonitor** (cluster synchronization service monitor), replaces oprocd and provides I/O fencing
- **OHASD orarootagent** starts and in turn starts

Level 2: OHASD rootagent spawns:

- CSD (ora.cssd) - Cluster Synchronization Services
- CRSD(ora.crsd) - Primary daemon responsible for managing cluster resources.
- CTSSD(ora.ctssd) - Cluster Time Synchronization Services Daemon
- Diskmon(ora.diskmon)
- ACFS (ASM Cluster File System) Drivers

crsd.bin	Cluster ready services, which manages high availability of cluster resources , like stopping , starting, failing over etc.
diskmon.bin	disk monitor (diskdaemon monitor) provides I/O fencing for exadata storage
octssd.bin	Cluster synchronization time services , provides Network time protocol services but manages its own rather depending on OS

CRSD Agent Phase:- crsd.bin starts two more agents

Level 3: CRSD spawns:

- orarootagent - Agent responsible for managing all root owned crsd resources.
- oraagent - Agent responsible for managing all oracle owned crsd resources.

- **CRSD orarootagent(Oracle root agent)** starts and in turn this will start

Level 4: CRSD rootagent spawns:

- Network resource - To monitor the public network
- SCAN VIP(s) - Single Client Access Name Virtual IPs
- Node VIPs - One per node
- ACFS Registry - For mounting ASM Cluster File System
- GNS VIP (optional) - VIP for GNS

gns	Grid interprocess communication daemon, used for monitoring cluster interconnect
gns vip	Multicast DNS service It resolves DNS requests on behalf of GNS
Network	Monitor the additional networks to provide HAIP to cluster interconnects
Scan vip	Monitor the scan vip, if found fail or unreachable failed to other node
Node vip	Monitor the node vip, if found fail or unreachable failed to other node

- **crsd oraagent(Oracle Agent)** starts and in turn it will start (the same functionality in 11gr1 and 10g managed by racgmain and racgimon background process) which is now managed by crs Oracle agent itself.

Level 4: CRSD oraagent spawns:

- ASM Resource - ASM Instance(s) 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
- GSD - For 9i backward compatibility
- GNS (optional) - Grid Naming Service - Performs name resolution

ASM & disk groups	Start & monitor local asm instance
ONS	FAN feature, provides notification to interested client
eONS	FAN feature, provides notification to interested client
SCAN Listener	Start & Monitor scan listener
Node Listener	Start & monitor the node listener (rdbms?)

How does 11gr2 CRS start if OCR is in ASM

Startup Sequence

init -> OHAS →(OLR) -> CSSD/GPnPD (access ASMDISK/LUN and read Voting Disks) -> CRSD (open DISKGROUP and read OCR).

Beginning with the version 11g Release 2, the ASM spfile is stored automatically in the first disk group created during Grid Infrastructure installation. Since voting disk/OCR are stored on ASM, ASM needs to be started on the node. To startup ASM, its Spfile is needed. But SPFILE is again located on ASM diskgroup only. How does Clusterware resolve this issue?

- 1) When a node of an Oracle Clusterware cluster restarts, OHASD is started by platform-specific means. OHASD accesses OLR (Oracle Local Registry) stored on the local file system to get the data needed to complete OHASD initialization.
- 2) OHASD brings up GPnPD and CSSD. CSSD accesses the GPnPD Profile stored on the local file system which contains the following vital bootstrap data;
 - a. **ASM_DISKSTRING** parameter (if specified) to locate the disks on which ASM disks are configured.
 - b. **ASM SPFILE location** : Name of the diskgroup containing ASM spfile
 - c. **Location of Voting Files**: ASM
- 3) CSSD scans the headers of all ASM disks (as indicated in ASM_DISKSTRING in GPnPD profile) to identify the disk containing the voting file. Using the pointers in ASM disk headers, the Voting Files locations on ASM Disks are accessed by CSSD and CSSD is able to complete initialization and start or join an existing cluster.
- 4) To read the ASM spfile during the ASM instance startup, it is not necessary to open the disk group. All information necessary to access the data is stored in the device's header. **OHASD reads the header of the ASM disk containing ASM SPfile** (as read from GPnPD profile) **and using the pointers in disk header, contents of ASM spfile are read.** Thereafter, **ASM instance is started.**
- 5) With an ASM instance operating and its Diskgroups mounted, access to Clusterware's OCR is available to CRSD.

6) OHASD starts CRSD with access to the OCR in an ASM Diskgroup.

7) Clusterware completes initialization and brings up other services under its control.

Note: CRS starts after ASM is up and the diskgroup for OCR and voting disks are mounted.

1. CSSD and CTSSD are up before ASM
2. Voting disks used by CSSD are discovered by reading the header of the disks using kfed utility, not through ASM.
3. Then ASM instance is started up and the diskgroup for OCR and voting disks are mounted.
4. CRS is now started by reading the OCR file inside ASM Diskgroup.