Evictions - What, Why, How?

Aman Sharma

OTNYatra-2016
Who Am I?

- Aman Sharma
- About 12+ years using Oracle Database
- **Oracle ACE**
- Frequent Contributor to OTN Database forum (*Aman....*)
- Oracle Certified
- Sun Certified

[@amansharma81]

[http://blog.aristadba.com]
Agenda

• **RAC Overview**
• **Understanding Node Eviction**
• **Instance Eviction**
• **Q/A**
Oracle RAC Overview

• Software based clustering using Grid Infrastructure software
• Cluster nodes contain only database and ASM instances
• Homogenous configuration
• Dedicated access to the shared storage for the cluster nodes
• Applications/users connect via nodes outsides the cluster
• Reflects Point-to-Point model
Oracle RAC Architecture - Simplified

- SGA
- Clusterware
- CFS
- SGA
- Clusterware
- Voting Disk(s)

Aman Sharma-OTNYatra(2016)
Why Evictions Happen in Clusterware

• If the nodes are not able to see each other, every node(group) *thinks* that it is the only survivor

• This causes *Split-Brain* scenario

• To avoid this phenomenon, culprit nodes and/or instances are evicted
Concept of Master Node in RAC

Three types

• OCR Master node
• Resource Master node
• Cluster Master node
Cluster Master Node

- Clusterware Master node
- Generally is the first node that joins the cluster is considered the master node
- In multi-node RAC evictions
  - Largest node group survives
  - If nodes are even in the groups, subgroup having the CSS master node would survive
- Can be checked by looking into the OCSSD logs/CRSD logs
Eviction Types in Oracle CW & RAC

- **Node Eviction**
  - Mechanism in which a node(s) is evicted (rebooted) by other cluster nodes from the cluster
  - Happens if any critical issue is detected at Cluster level
  - Is done to ensure that the cluster won’t enter into Split-Brain scenario

- **Instance Eviction**
  - Mechanism in which an instance(s) is evicted by group of instances
  - Happens if any critical issue is detected at the RAC instance/database level
  - Is done to ensure that the entire database won’t be required to get crashed due to issue in one or few instances
Clusterware - When All is Well!

Node-1

SGA

I can see Node 2 😊

Clusterware

NW Heartbeat

CFS

SGA

I can see Node 1 😊

Clusterware

Node-2

VD Heartbeat

Voting Disk(s)

Aman Sharma-OTNYatra(2016)
Clusterware - When Things Go Wrong!

Node-1

SGA

I can see Node 2

Clusterware

Node-2

I can’t see Node 2

Clusterware

VD Heartbeat

CFS

NW Heartbeat

Voting Disk(s)

Aman Sharma-OTNYatra(2016)
• Clusterware uses both the heartbeats to know whether the nodes are “alive” or not
• Both the heartbeats are checked by CSSD constantly to avoid Split-Brain scenario
• Any node(s) failing to respond within a pre-defined threshold value for either of the two heartbeats would be evicted
Network & Disk Heartbeat

- If Disk HB is completed within it’s threshold AND NW HB is successful —
  - **Node remains in the cluster**

- If NW Heartbeat is missed then
  - **Node is evicted**
    - Disk HB doesn’t matter!

- If Disk HB isn’t complete in the defined threshold
  - **Node is evicted**
Node Evictions at Clusterware Level

• Node(s) failing to respond is evicted by other nodes
• Initiated by the CRS (Cluster Read Service)
• Reasons for Node evictions
  – Network Heartbeat Missed
  – Voting Disk Heartbeat Missed
  – CSSDAgent/CSSMonitor being hung
  – RDBMS Instance being hung and leading to node eviction

Aman Sharma-OTNYatra(2016)
What is The MissCount

• Maintained by CSS component of the Clusterware
• The threshold time in which the response is requested before the node eviction commences
• CSSD’s multithreaded architecture maintains the check for both Disk & Network Heartbeats
• Two types
  – CSS Misscount a.k.a Network Heartbeat
  – Disk Misscount a.k.a Disk Heartbeat
Node Eviction - NW HeartBeat Missed

- Node-1
  - SGA
  - Clusterware
  - I can see Node 2 😞

- Node-2
  - SGA
  - Clusterware
  - I can’t see Node 2 😞

- CFS

- NW Heartbeat

- VD Heartbeat

- Voting Disk(s)

Aman Sharma-OTNYatra(2016)
CSS Misscount

- CSS Timeout=network latency in a node-node communication (in Seconds)
- Can be checked through
  \textdollar crsctl get css misscount
- Default CSS Misscount=30 seconds
- For Vendor (non-Oracle) clusterware=600 seconds
- Default value can be changed (MOS#284752.1)
NW HeartBeat Miss - Reasons & Actions

• **Reasons**
  – No communication between the nodes (Private Network)
  – Mostly would be due to faulty network Hardware

• **Actions to be taken**
  – Check OCSSD Log & Alert Log of the evicted node
  – Confirm that the interconnect network is working fine
  – Confirm that there isn’t any packet drops in the private network
Node Eviction - Voting Disk Heartbeat Missed

I can’t see Node 2 😞

I can’t see Node 1 😞

Aman Sharma-OTNYatra(2016)
Cluster nodes ping constantly into the Voting disk to mark their presence.

Disk heartbeat has an internal timeout - Disk Timeout (DTO); *in Seconds*

Disk heartbeat i/o timeout interval is directly related to the misscount parameter setting.

Default value = 200 seconds

Default value can be changed (MOS#284752.1)
**What Is Inside a Voting Disk**

- Shared device accessible to all the nodes; used to achieve **Quorum**
- **Voting Disk Contains**
  - Information about the cluster nodes (Static) i.e.
    - Node(s) which are part of the cluster
    - Node(s) leaving or joining the cluster
  - Node Heartbeat Ping (Dynamic)
- OCSSD process(s) of each node write into the Voting Disk & also read the heartbeat writes of other nodes
- To stay in the cluster, each node must access >50% of the Voting disks
- Writes are done in the unit of an OS block size (OS dependent)
- Voting disk count must be in an ODD number
- \#Voting Disk=F*(2+1) where F=Voting Disks expected to be failed
- Maximum Voting Disks can be **15**
- From 11.2, can also be stored on ASM
(Voting) Disk Heartbeat Timeout & Reboot Time

• Disk Heartbeat is internally calculated
  – Value is different across different releases of Oracle
  – Disk Heartbeat Timeout=Maximum time allowed for Voting Disk IO to be completed
    $crsctl\ get\ css\ disktimeout(200\ seconds)$

• Reboot Time
  – Time allowed for a node to do a reboot after being evicted
    $crsctl\ get\ css\ reboottime(3\ seconds)$
Possible Reasons for Disk Timeout

• Mostly it would be due to some sort of HW related issue
• Bad cable/disk throughput causing more latency in doing the IO over the Voting Disk
• Wrong setting for permissions
• High CPU workload that may cause delay for CSSD to ping into the voting disk
• Useful to check OS, SAN, and storage logs
Processes Responsible For Node Reboots

Clusterware

CSSD
CSSDAGENT/CSSMonitor
OCLSKD
OPROCD
Processes Responsible For Node Reboots

- **CSSD—Cluster Synchronization Service Daemon**
  - Startup sequence:
    - `init -> init.ohasd -> ohasd -> ohasd.bin -> cssdagent -> ocssd -> ocssd.bin`
  - Runs as Oracle/Grid user
  - Manages cluster node membership via Network & Disk Heartbeat
  - Failure results in node restart

- **CSSDAgent**
  - Startup sequence:
    - `INIT -> init.ohasd -> ohasd -> ohasd.bin -> cssdagent`
  - Introduced in 11.2
  - Runs with elevated priority as ROOT user
  - Monitors the cluster and provides I/O fencing (used to be handled by OPROCD)
  - Stops, starts, checks the status of CSSD
  - Failure results in node restart
Processes Responsible For Node Reboots

• **CSSDMonitor**
  - Startup sequence:
    - INIT --> init.ohasd --> ohasd --> ohasd.bin --> cssdmonitor
  - Introduced in 11.2
  - Runs with elevated priority as ROOT user
  - Monitors node & OCSSD process hangs
  - Monitors Vendor Clusterware

• **OCLSKD** *(Oracle CW Kill Daemon)*
  - New from 11.1.0.6
  - Handles RDBMS Instance Hang-up issues causing Instance Eviction
  - Also can nodes eviction (Member Kill Escalation)

• **OPROCD** *(Oracle Process Monitor Daemon)*
  - Used to provide IO Fencing (versions <11g)
  - Runs as Root, spawned in init.CSSD
  - Used to detect HW and driver freeze on the node
  - From 11g onwards, monitors the system state of the nodes
  - Implemented on Unix if not using Vendor Clusterware
  - From 10.2.0.4, implemented on Linux
  - Works along with Hangcheck Timer
  - Process failure/halt result in Node eviction
Troubleshooting OCSSD Evictions

- Network failure or latency between nodes. It would take 30 consecutive missed check-ins to cause a node eviction.
- Problems writing to or reading from the CSS voting disk. If the node cannot perform a disk heartbeat to the majority of its voting files, then the node will be evicted.
An unexpected failure or hang of the OCSSD process,

Can be caused due to some software bug
Troubleshooting Node Eviction Due to CSSDAgent Hung

- CSSDAgent & CSSMonitor run in elevated priority
- Both process can cause a Node eviction if
  - There is an OS scheduler problem
  - There is some CPU starvation due to excessive workload or undersized CPU resource being available
  - A thread(s) within the CSS daemon hung
  - An OS I/O call is blocked due to some hardware issue
- Can be also caused by some Oracle bug

```
[grid@oranode01 oranode01]$ tailf /u01/app/11.2.0/grid/log/oranode01/alerthost01.log
[ohasd(12412)]CRS-8013:reboot advisory message text: Rebooting after limit 28100 exceeded; disk timeout 28100, network timeout 27880, last heartbeat from CSSD at epoch seconds 352436647.013, 34280 milliseconds ago based on invariant clock
  Node 2 is rebooted and network connection with it breaks value of 294678040
[ohasd(12412)]CRS-8013:reboot advisory message text: Rebooting after limit 28100 exceeded; disk timeout 28100, network timeout 27880, last heartbeat from CSSD at epoch seconds 352436647.013, 34280 milliseconds ago based on invariant clock
  Node 2 is rebooted and network connection with it breaks value of 294678040
[ohasd(12412)]CRS-8013:reboot advisory message text: Rebooting after limit 28100 exceeded; disk timeout 28100, network timeout 27880, last heartbeat from CSSD at epoch seconds 352436647.013, 34280 milliseconds ago based on invariant clock
  Node 2 is rebooted and network connection with it breaks value of 294678040
[cssd(14493)]CRS-1612:Network communication with node oranode02 (2) missing for 50% of timeout interval. Removal of this node from cluster in 14.330 seconds
```
Node Eviction Due to Member Kill Escalation

- Each RAC Instances pings into the control files of each other
- Mechanism is known as IMR (Instance Membership Recovery)
- If for some reason, RDBMS instance becomes hung, available instance(s) would ask the hung instance to kill itself and reboot the node
- If before the CSS misscount threshold is crossed, instance kill gets escalated to node kill

LMS0 (ospid: 31771) has detected no messaging activity from instance 2
LMS0 (ospid: 31771) issues an IMR to resolve the situation
Please check LMS0 trace file for more detail.
Fri Nov 20 11:22:14 2015
Remote instance kill is issued with system inc 30
Remote instance kill map (size 1) : 2
LMON received an instance eviction notification from instance 1
The instance eviction reason is 0x20000000

Database Alert Log output snippet

[cssd(14493)]CRS-1607: Node oranode02 is being evicted in cluster incarnation 267943929; details at (:CSSNM00007:) in

Clusterware Alert Log output snippet
Node Evictions-Which Process Is the Culprit?

- From 11.2 onwards, in the Cluster Alert Log, the process responsible would be mentioned.
- This facilitates to find the root cause more quickly.
- If there is no message in the evicted node’s CW Alert Log, check lastgap of the local node:
  - `/etc/oracle/lastgasp/cssagent_<node>.lg`
- CW alert log on the other nodes

Network Hearbeat Missed:

```
[ohasd(11243)] CRS-8013: reboot advisory message text: Rebooting after limit 28500 exceeded; disk timeout 27630, network timeout 28500, last heartbeat from CSSD at epoch seconds 1241543005.340, 4294967295 milliseconds ago based on invariant clock value of 93235653
```
Reboot Less Node Eviction

• Prior releases of Oracle CW would prevent a Split-Brain by asking for the culprit node immediately.
• It would result in the node reboot along with all the applications on it - Oracle & Non-Oracle.
• From 11.2.0.2 onwards, mechanism is changed to safeguard especially non-Oracle applications from being forcefully restarted.
Reboot Less Node Eviction (contd....)

- With Reboot Less Reboot, when a node is found to be the culprit and has to be evicted
  - CW will restart only the offending process(s)
  - Entire node reboot doesn’t happen
  - May result in the stop of all Oracle related processes to be stopped

- On the culprit node
  - Clusterware would be stopped
  - OHASD daemon would keep on trying to start CRS daemon & stack
  - A restart of the CRS would be eventually initiated
Node Reboot in Reboot-less Node Eviction

• Even with Reboot-less eviction, node reboot may happen

• Node will still reboot if
  – All Oracle related processes can’t be stopped
  – IO related processes can’t be stopped for some reason e.g. hung due to some hardware issue
Alert Log Output....
[cssd(2876)]CRS-1649: An I/O error occurred for voting file: ORCL:ASMDISK22; details at (:CSSNM00059:) ...
[cssd(2876)]CRS-1606: The number of voting files available, 0, is less than the minimum number of voting files required, 1, resulting in CSSD termination to ensure data integrity;
[cssd(2876)]CRS-1656: The CSS daemon is terminating due to a fatal error; [cssd(2876)]CRS-1652: **Starting clean up of CRSD resources.** 2015-11-11 12:40:30.795 ...
[cssd(2876)]CRS-1654: **Clean up of CRSD resources finished successfully.**

[root@oracle-node02 ~]# crsctl check crs
CRS-4638: **Oracle High Availability Services is online**
CRS-4535: Cannot communicate with Cluster Ready Services
CRS-4530: Communications failure contacting Cluster Synchronization Services daemon
CRS-4534: Cannot communicate with Event Manager
Understanding Instance Eviction

• Unlikely node eviction (reboot), one or few instances of the cluster database would be evicted

• An instance is evicted if
  – Instance’s processes are not responding to the other instances or Control File e.g. LMON doesn’t ping to the control file in every 3 seconds
  – Communication link is down
  – Instance is termed to be inactive (no respond from the instance to anything)
Instance Eviction-Possible Causes

• Network issues
• OS Load issues
• Memory or CPU starvation issues
• Instance eviction may lead to Node Eviction(Member Kill escalation)
  – LMON may request CSS to remove an instance from the cluster
  – If NW timeout occurs before it happens, node itself is evicted(11g onwards)
Reasons for Instance Evictions

• For instance evictions-errors reported are
  – **ORA-29740, 29770, 481, 484**
• For ORA-29740, check the REASONS(1,2,3)
  – Reason 0 = No Reconfiguration-> No Eviction
  – Reason 1 = Node Monitor triggered the reconfiguration(DBA issued the reconfiguration of the nodes)
  – Reason 2 = An instance death is detected (no HB to the Control File by LMON
  – Reason 3 = Communication Failure(NW issue or High Load)
  – Reason 4 = Reconfiguration after suspend
• Files to be checked
  – CW Alert Log (Evicted & Surviving Node)
  – LMON Trace File (Evicted & Surviving Node)
Troubleshooting Node Evictions - Where To look?

Useful Files to look into

• Clusterware alert log in <GRID_HOME>/log/<nodename>
• The cssdagent log(s) in <GRID_HOME>/log/<nodename>/agent/ohasd/oracssdagent_root
• The cssdmonitor log(s) in <GRID_HOME>/log/<nodename>/agent/ohasd/oracssdmonitor_root
• The ocssd log(s) in <GRID_HOME>/log/<nodename>/cssd
• The lastgasp log(s) in /etc/oracle/lastgasp or /var/opt/oracle/lastgasp
• The OCLSKD log in <GRID_HOME>/log/<nodename>client/oclskd.log
• Cluster Health Monitor & OS Watcher data
• Message Files
  – Linux: /var/log/messages
  – Sun: /var/adm/messages
  – HP-UX: /var/adm/syslog/syslog.log
  – IBM: /bin/errpt -a > messages.out
Node Eviction-Tools That Can Help in RCA!

• OS Watcher (MOS# 301137.1)
• Cluster Health Monitor (MOS# 736752.1)
• TFA Collector- Tool for Enhanced Diagnostic Gathering (MOS# 1513912.1)
• ORACHk (MOS# 1268927.2)
Few Recommendations

• Ensure proper HW resources are available
• For database, use AMM & Resource Manager for proper memory management
• Run ORACHK & implement it’s given suggestions
• Install & use OSWatcher
• Use faster network, especially between the nodes(Private Network)
• Private interconnect card should be very fast-at least 1Gigabit
• Don’t use crossover cables but use a Switch instead
• Ensure to use NTP(Network Time Protocol)
Thank You!

@amansharma81

Q & A!

amansharma@aristadba.com