\*\*Split-brain syndrome\*\* in Oracle RAC (Real Application Clusters) refers to a situation where multiple instances (nodes) in the cluster lose communication with each other, but continue to access the shared storage independently, each believing they are the only active instance. This can lead to data corruption because each instance might modify the same data blocks without coordinating with the others.

**Causes of Split-brain in RAC**

**Split-brain can occur due to**:

1. \*\*Network Partitioning (Interconnect Failure)\*\*: When the private interconnect between the RAC nodes fails, the nodes are no longer able to communicate. Each node may assume the other node(s) are down and attempt to take control of the database resources.

2. \*\*Clusterware Failure\*\*: If Oracle Clusterware fails or is misconfigured, it might not properly detect that other nodes are still operational.

3. \*\*Disk I/O Issues\*\*: If nodes cannot access shared storage due to storage network issues but think they can, a split-brain scenario might arise.

**Impact of Split-brain Syndrome**

In split-brain, multiple instances might modify the same data blocks in the shared storage, causing:

**Data Corruption**: Without coordination between the nodes, inconsistent changes to data blocks can occur.

**Lost Transactions**: Some transactions processed by one node may not be visible to other nodes.

**System Instability**: The database can become unstable, and the instances might crash due to data corruption or inconsistencies.

**How Oracle RAC Handles Split-brain**

Oracle RAC is designed to prevent split-brain by using a \*\*voting disk\*\* and \*\*fencing mechanisms\*\*. If the interconnect between nodes fails, Oracle Clusterware and the voting disk will ensure that only one subset of nodes can continue to access the shared database, while the other nodes are "fenced" off to prevent them from accessing shared resources. The fencing mechanism is critical in resolving split-brain situations by terminating the connection of any node that has lost communication.

**Voting Disk and Quorum**

Voting Disk: Each RAC cluster is configured with one or more voting disks. These disks are used to maintain the quorum among the nodes, which ensures that only a majority of nodes can access the database.

Quorum: When a communication failure occurs, the nodes that maintain a majority of the voting disks form the quorum and are allowed to continue functioning. The minority nodes are fenced and shut down.

**Node Eviction**

Oracle RAC uses a \*\*node eviction mechanism\*\* to prevent split-brain. The clusterware will forcefully shut down any nodes that lose quorum. This ensures that only one set of instances remains active and can modify the shared data.

**Example of Split-brain Syndrome**

Consider a two-node Oracle RAC cluster (Node A and Node B):

1. \*\*Normal Operation\*\*: Both nodes (Node A and Node B) are communicating through the private interconnect, and they both access the shared storage (datafiles) and coordinate their changes using the Global Cache Service (GCS).

2. \*\*Interconnect Failure\*\*: The private network connection between Node A and Node B fails due to a network issue.

- \*\*Node A's Perspective\*\*: Node A assumes that Node B has failed and continues processing transactions.

- \*\*Node B's Perspective\*\*: Node B assumes that Node A has failed and also continues processing transactions.

3. \*\*Split-brain Scenario\*\*: Both Node A and Node B now believe they are the only active node and start modifying the shared database independently. For instance:

- Node A updates a row in Table 1 and writes it to disk.

- Node B simultaneously updates the same row in Table 1 and writes a different value to disk.

4. \*\*Data Corruption\*\*: Since there is no communication between the two nodes, both are unaware of the changes made by the other, leading to potential corruption of the database when these conflicting changes are written to the shared storage.

**Preventing Split-brain with Fencing and Voting Disk**

In the above scenario, Oracle Clusterware and the voting disk mechanism would prevent both Node A and Node B from continuing their operations independently.

1. \*\***Fencing Action**\*\*: Upon detecting a loss of communication between Node A and Node B, Oracle Clusterware will initiate a fencing mechanism.

- If Node A holds the majority of voting disks (or maintains access to the voting disk), it will continue operating.

- Node B, having lost access to the voting disk or quorum, will be evicted (shutdown), ensuring only Node A continues working.

2. \*\***Quorum Resolution**\*\*: Once communication is restored, Node B can rejoin the cluster after restarting, and the Global Cache Service will synchronize the data across nodes.

**How to Monitor Split-brain in Oracle RAC**

1. \*\***Cluster Logs**\*\*: Oracle Clusterware logs (in the `CRS` logs) can provide details about node evictions or split-brain situations.

Example log location:

/u01/app/grid/diag/crs/hostname/crs/trace/alert.log

2. \*\***Alert Logs**\*\*: The database alert logs for each instance will show any errors or issues related to split-brain.

3. \*\***Voting Disk Status**\*\*:

You can check the voting disk status to ensure it is functioning correctly:

crsctl query css votedisk

Example of Node Eviction in Log

The log file may show messages like the following during node eviction:

Node eviction completed. Reason: Rebooting due to split-brain. Voting disk quorum lost.

### **Conclusion**

Oracle RAC's split-brain syndrome occurs when communication between nodes in the cluster is lost, and multiple instances try to take control of the shared storage independently, which can lead to data corruption. Oracle's fencing mechanism, voting disk, and node eviction processes are designed to detect and prevent split-brain scenarios by ensuring that only one set of nodes can access the shared storage at a time.