Oracle Data Guard – Fast Start Failover understood!

Dr. Martin Wunderli
http://www.trivadis.com
Principal Consultant Partner
Trivadis Facts & Figures

- 12 locations
  - D: Dusseldorf, Frankfurt, Freiburg, Hamburg, Munich, Stuttgart
  - A: Vienna
  - CH: Baden, Basle, Bern, Lausanne, Zurich

- Consolidated income
  CHF 85 million / EUR 53 million

- Over 470 employees

- Over 450 clients

- Over 1'400 projects per year

- Over 110 Service Level Agreements

- About 4'000 training participants per year
FSFO understood!

- Data Guard Concepts & History
- The startup issue
- Fast Start Failover
- Conclusion

Data is always part of the game.
Oracle Standby Databases and Data Guard – Overview

Log Transport:
Either via asynchronous or synchronous

Primary Site
- Primary Database
- Online Log Files
- Local Archiving

Standby Site
- Standby Database
- Archived Log Files

Standby Log Files
- Either Physical Standby (redo apply)
- or Logical Standby (sql apply)
- Either Realtime or via archived Redo Logs
Standby Databases: A short history

- **Oracle 7.3:** Creating and mounting a standby database

- **Oracle 8i:** Automated archived redo log transport and application, TAF, open read-only of standby database

- **Oracle 9i:** Data Guard and Data Guard Broker with switchover, close log gap, delayed redo application, GUI and no-data-loss setups (sync transport)

- **Oracle 10g:** Simplified syntax, RAC support, partial failover cluster support, reuse of old primary as new standby database, automatic standby activation
Why Data Guard (and not e.g. a Failover Cluster)?

- In case of a disaster protection setup (data must be mirrored between at least two locations), bandwidth usage is smaller: Even high transaction systems typically need only approx. 70 MBit/s bandwidth.

- No extra software layer and license needed (if you already licensed Oracle Enterprise Edition on primary and standby server…)

- No file system or instance recovery of database needed after crash of primary server (standby is up to date in case of No-Data-Loss setup and 10gR2)
Why a Failover Cluster (and not Data Guard)?

- File system based mirroring is needed because of non-database files
- IP address failover is needed for e.g. an application server
- DBA knowledge is not available

- In case that instance recovery time and bandwidth between locations is also crucial, a combination of Failover Cluster and Data Guard between the same machines may be necessary
Failover Cluster vs. Data Guard

- Remember…

- Fast Start Failover for Data Guard is **not** a failover cluster with
  - Two connections between nodes (network and disk) where the loss of one connection results in node shutdown
  - A single location of data files (from the point of view of Oracle RDBMS)

- These two points have positive and negative impact
  - Nodes stay longer up and in their role in case of partial inter node connection loss
  - Automatic failover may not be possible after partial inter node connection loss
FSFO understood!

- Data Guard Concepts & History
- The startup issue
- Fast Start Failover
- Conclusion
Physical Standby: Startup Behavior 10g versus 9i

- `startup nomount` (Primary 9i)
- `alter database mount;`
- `alter database open;` # Primary
- `recover managed standby database;` # Standby
- Data Guard Broker 10g takes over and opens primary database if...
- Data Guard Broker 9i takes over
Physical Standby – Startup Issue

- What is the biggest problem for data consistency in a cluster?

  - Split Brain!

- What is the biggest problem for data consistency in a Data Guard environment?

  - More than one primary!

- How can this happen?
  Primary startup (Hardware fixed etc.) after standby activation
Primary Startup after Failover: Network connected

- **OK:** STARTUP MOUNT of former primary database
  - Data Guard Broker takes over and handles startup process
  - The Broker knows about the failover and the resulting change of the primary database
  - The former primary database is not started

  ```
  DGMGRL> show configuration;
  Error: ORA-16795: database resource guard detects that database re-creation is required
  Configuration details cannot be determined by DGMGRL
  ```

- **BAD:** STARTUP of former primary database
  - Results in two primary databases since sqlplus does not know of the Data Guard Broker configuration
Primary Startup after Failover: Network interrupted

- **BAD**: STARTUP of former primary database
  - Results in two primary databases since sqlplus does not know of the Data Guard Broker configuration

- **BAD**: STARTUP MOUNT of former primary database
  - Data Guard Broker tries to verify the Data Guard configuration
  - After 5 unsuccessful requests, Data Guard Broker opens the former primary database 😞
Startup: Variants

1. Only mount Primary and Standby Database during system boot
2. Manual database startup after system boot
3. Adapt TNSNAMES or LDAP server so that old Primary is not found anymore. But local jobs…
4. Is there a better solution? Yes, see Fast Start Failover!
FSFO understood!

- Data Guard Concepts & History
- The startup issue
- Fast Start Failover
- Conclusion

Data is always part of the game.
Fast-Start Failover

- Main criticism of Oracle standby databases: too much manual interaction

1. Manual interaction is required for a failover
   - Need some administrative checks before to validate the status of the standby database, e.g. if all redo are applied
   - More downtime 😞

2. Manual interaction to recreate a new standby database
   - No HA until the setup of the new standby is finished

3. Manual interaction is needed for startup if two primaries have to be avoided at all cost
   - Fast Start Failover addresses all three problems!
Fast Start Failover: Concept

1. Observed Data Guard environment

2. Fast-Start-Failover (automatic)

3. Reinstallte (automatic)
When is a Fast-Start Failover triggered?

- **Primary site failure**
  - Server crash or server shutdown (without database shutdown)

- **Primary database failure**
  - Instance failure (last running instance if RAC)
  - Shutdown abort (but not with normal or immediate)
  - Data file is taken offline

- **Network failure (special case)**
  - Documentation of when and when not automatic activation will happen is quite large. Read and test carefully. We will show one case.
Network Failure (1)

Select fs_failover_status, fs_failover_observer_present
from v$database; -- on primary site

<table>
<thead>
<tr>
<th>FS_FAILOVER_STATUS</th>
<th>FS_FAILOVER_OBSERVER_PRESENT</th>
</tr>
</thead>
<tbody>
<tr>
<td>SYNCHRONIZED</td>
<td>NO</td>
</tr>
</tbody>
</table>

observer

log transport
Network Failure (2)

Select fs_failover_status,fs_failover_observer_present
from v$database; -- on primary site

FS_FAILOVER_STATUS    FS_FAILOVER_OBSERVER_PRESENT
---------------------- ----------------------------
STALLED               NO
Network Failure (3)

Select `fs_failover_status,fs_failover Observer_PRESENT` from `v$database`; -- on new primary site

<table>
<thead>
<tr>
<th>FS_FAILOVER_STATUS</th>
<th>FS_FAILOVER_OBSERVER_PRESENT</th>
</tr>
</thead>
<tbody>
<tr>
<td>REINSTATE REQUIRED</td>
<td>YES</td>
</tr>
</tbody>
</table>

---

Observer

Database STALLED

New primary DB
Network Failure (4)

Select fs_failover_status, fs_failover_observer_present
from v$database; -- on primary site and standby site

<table>
<thead>
<tr>
<th>FS_FAILOVER_STATUS</th>
<th>FS_FAILOVER_OBSERVER_PRESENT</th>
</tr>
</thead>
<tbody>
<tr>
<td>SYNCHRONIZED</td>
<td>YES</td>
</tr>
</tbody>
</table>
Observer location?

Main computing center

1. Observer

primary DB

log transport

2. Observer

standby DB

Standby computing center

3. Observer

4. Observer

5. Observer
Observer location …

- Best is three locations:
  - One for primary database
  - One for standby database
  - One for observer

- In many real life situations (no three locations…)
  - Observer on primary site will be the best choice if avoiding 'false' activations is most important
  - Observer on standby site will be the best choice if protection from computation center loss is most important
Compromise to minimize false activates
Observer – Installation Requirements

- Observer machine with Oracle Net configuration
- Special entry in Data Guard Broker configuration which requires…
  - MaxAvailability protection mode for Primary Database
    - but: special startup behavior
    - but: primary stalls in certain situations
  - Flashback database activated on Primary and Standby Database
Observer - Data Guard additional Configuration

- Not much to configure

```
edit database 'PHYS_LUCERNE'
    set property FastStartFailedoverTarget = 'PHYS_TOKYO';
edit database 'PHYS_TOKYO'
    set property FastStartFailedoverTarget = 'PHYS_LUCERNE';
edit configuration
    set property FastStartFailedoverThreshold = 15;
enable fast_start failover;
```

- Fast-Start Failover is a feature of Oracle Data Guard, and cannot run without a Data Guard Broker configuration!
FSFO – Does it work? (1)

- Usually it works 😊

- An interrupt (network, server crash etc.) during reinstate often results in problems
  - FSFO configuration hangs
    - The reinstating instance will not continue
    - The observer cannot be stopped (with stop observer)
  - How to solve the problem

```
dgmgrl
connect system@<new_primary>
disable fast_start failover force;
reinstall database '<old_primary>';
enable fast_start failover;
```

- If the "disable fast_start failover force" also hangs, kill/start the observer and restart the new primary instance
FSFO – Does it work? (2)

- In rare cases, the whole broker configuration is corrupted
  - Remove the configuration
    - On both nodes / instances
      ```sql
      sql> shutdown immediate
      cd $ORACLE_BASE/admin/DG1/pfile/
      mv dr1DG1.dat dr1DG1.dat.bck
      mv dr2DG1.dat dr2DG1.dat.bck
      sql> startup mount
      ```
  - Recreate the configuration (good to have scripts 😊)
    ```
    dgmgr1> create configuration 'DG1'
    dgmgr1> add database
    dgmgr1> edit database / edit configuration
    dgmgr1> enable configuration;
    dgmgr1> enable fast_start failover;
    ```
FSFO understood!

- Data Guard Concepts & History
- The startup issue
- Fast Start Failover
- Conclusion

Data is always part of the game.
Fast Start Failover understood: Core Messages

- FSFO addresses three major problems of 9i Data Guard
- Observer location is not easy to decide
- Things can become corrupt: Be prepared to recreate the Data Guard configuration

Data is always part of the game.