Seamless Application Failover with Oracle Data Guard

Ashish Ray
Senior Director of Product Management
Database High Availability

Mike Smith
Principal Member of Technical Staff
Maximum Availability Architecture

Martin McGeough
Database Technical Architect
VocaLink
The following is intended to outline our general product direction. It is intended for information purposes only, and may not be incorporated into any contract. It is not a commitment to deliver any material, code, or functionality, and should not be relied upon in making purchasing decisions. The development, release, and timing of any features or functionality described for Oracle’s products remains at the sole discretion of Oracle.
Oracle OpenWorld
Latin America 2010

December 7–9, 2010
Oracle Products Available Online

Oracle Store

Buy Oracle license and support online today at oracle.com/store
Agenda

• Overview
  – Oracle Maximum Availability Architecture (MAA)
  – Application failover types
  – Oracle Data Guard, Switchover / Failover
  – New application failover features in Oracle Database 11g Release 2

• Seamless Application Failover with Data Guard
  – Design principles and requirements
  – Best practices: JDBC and OCI
  – Considerations for Oracle Applications & WebLogic
  – Considerations for previous releases

• VocaLink Case Study
  – VocaLink: business summary
  – Description of VocaLink MAA configuration
  – Failover configuration & business benefits
Oracle Maximum Availability Architecture
Low cost, Integrated, Fully Active, High ROI

**Production Site**

- **RAC**
  - Scalability
  - Server HA

- **Flashback**
  - Human error correction

- **ASM**
  - Volume Management

- **RMAN & Fast Recovery Area**
  - On-disk backups

**Active Replica**

- **Active Data Guard**
  - Data Protection, DR
  - Query Offload

- **GoldenGate**
  - Active-active
  - Heterogeneous

- **Oracle Secure Backup**
  - Backup to tape / cloud

Edition-based Redefinition, Online Redefinition, Data Guard, GoldenGate
- Minimal downtime maintenance, upgrades, migrations

Ref. [http://www.oracle.com/goto/maa](http://www.oracle.com/goto/maa)
Application Failover

Principles

• Definition
  – In the event of a planned / unplanned outage, how efficiently can user sessions be directed to a secondary site / database, with minimal interruption

• Types
  – Full site failover
  – Partial site failover / Database failover

• Focus of this presentation
  – Application failover when database failover is facilitated by Data Guard
Full Site Failover

Manual, for Disaster Recovery

Production Site A

Web Servers

App Servers

Database Servers

Traffic Manager

DR Site B

Web Servers

App Servers

Database Servers

Traffic Manager

DR Site A

Web Servers

App Servers

Database Servers

Production Site B

DR Failover
Partial Site / Database Failover
Automated, for Local High Availability
Oracle Data Guard Overview

Best Data Protection

- Data availability and data protection for the Oracle Database
- Up to thirty standby databases in a single configuration
- Standby database used for queries, reports, test, or backups
Data Guard Role Transition

Switchover & Failover

- Switchover
  - Planned role reversal, never any data loss
  - No database re-instantiation required
  - Used for database upgrades, tech refresh, data center moves, etc.
  - Manually invoked via Grid Control, DGMGRL, or SQL*Plus

- Failover
  - Unplanned failure of primary
  - Flashback Database used to reinstate original primary
  - Manually invoked via Grid Control, DGMGRL, or SQL*Plus
  - May also be done automatically: Fast-Start Failover
Data Guard Fast-Start Failover
Reduce Unplanned Downtime

- Automatic database failover, upon …
  - Database down
  - Designated health-check conditions
  - Request of an application

- Supported with
  - Maximum Availability (10.2)
  - Maximum Performance (11.1)
Overview

- Single name for clients to access an Oracle Database running in a cluster
- Configured during the installation of the Grid Infrastructure
- Resolves to 3 IP addresses in the cluster, each associated with a SCAN Listener

```
SALES=(DESCRIPTION=
  (ADDRESS=(PROTOCOL=tcp)(HOST=Austin-scan)(PORT=1521))
  (CONNECT_DATA=(SERVICE_NAME=OrderEntry)))
```

Load Balancing

- Client load balancing: Oracle Net randomly connects clients to one of the 3 SCAN VIPs and connects to that node’s SCAN Listener
- Server load balancing: SCAN Listener checks for least loaded instance for the requested service, re-directs connection to the local listener on that node

Benefits

- No need to change client connection configuration if cluster changes
- Single-name method much more manageable for large clusters – avoids need to list a series of IP address strings in `tnsnames.ora`
Oracle Database 11.2 Enhancements
Role Specific Database Services

• Overview & Benefit
  – Services can be automatically active in specific database roles on databases configured with Oracle Clusterware or Oracle Restart
  – Data Guard Broker interacts with Oracle Clusterware or Oracle Restart to ensure appropriate services are active after a role transition: no need to write triggers on system startup events

• Example
  – Service “OrderEntry” active in primary database role on the primary database Austin
  – Service “OrderReport” active in physical standby database role on the standby database Houston

```
srvctl add service -d Austin -s OrderEntry -l PRIMARY -q TRUE -e SELECT
                   -m BASIC -w 10 -z 150
srvctl add service -d Houston -s OrderReport -l PHYSICAL_STANDBY -q TRUE -e SELECT
                   -m BASIC -w 10 -z 150
```

Note:

1. For failover / switchover, Service “OrderEntry” should also be configured on Houston in primary role and “OrderReport” be configured on Austin in physical standby role

2. The “OrderReport” service must be explicitly started with `SRVCTL START SERVICE` and stopped with `SRVCTL STOP SERVICE` on primary database to ensure its information is propagated via redo to physical standby
Overview

- Following a Data Guard Failover (manual or Fast-Start Failover), Data Guard Broker now automatically publishes a FAN (Fast Application Notification) event to clean up connections to the failed primary database.
- Upon receipt of that event, FAN client subscribers can automatically reconnect to the service started on the new primary database.
- Requires databases configured with Oracle Clusterware or Oracle Restart.
- Supports publishing FAN events with either ONS (Oracle Notification Services – for JDBC clients) or AQ (Advanced Queueing – for OCI or ODP.NET clients), for failover to either physical or logical standby database.

Benefit

- Integrated Broker / FAN automation removes the need of an external program to publish these FAN events, and manual triggers to invoke that external program.
Overview

- Oracle Maximum Availability Architecture (MAA)
- Application failover types
- Oracle Data Guard, Switchover / Failover
- New application failover features in Oracle Database 11g Release 2

Seamless Application Failover with Data Guard

- Design principles and requirements
- Best practices: JDBC and OCI
- Considerations for Oracle Applications & WebLogic
- Considerations for previous releases

VocaLink Case Study

- VocaLink: business summary
- Description of VocaLink MAA configuration
- Failover configuration & business benefits
Seamless Application Failover
Overall Design Principles

• Focus of this presentation
  – Seamless application redirect during database failover by Data Guard
  – Ideal for maintaining app uptime during local outages
    ▪ Examples: disk failure, data corruption, hardware fault, software hang, etc.,
      much more frequent than site-disasters

• Seamless application failover: three main components
  – Restarting services on the new primary
  – Notifying applications to terminate existing connections
  – Enabling applications to reconnect efficiently

• Required Oracle technologies:
  – Oracle Clusterware
  – Oracle Restart for single instance
  – Oracle Data Guard, including Data Guard Broker
Partial Site Failover – Steady State

App Server Farm

Primary Connection
- JDBC: Subscribe to ONS
- Connected to OrderEntry Service

2-node RAC Primary
- JDBC: ONS Daemon
- Service: OrderEntry

Data Guard Redo Transport

Data Guard Broker

HOUSTON Data Center

Standby Connection
- JDBC: Subscribe to ONS

2-node RAC Standby
- JDBC: ONS Daemon
- Service: OrderEntry

AUSTIN Data Center
Partial Site Failover – Database Failover

AUSTIN Data Center

2-node RAC Primary

1

2-node RAC Standby

2

Primary Connection
- JDBC: Timeout ensues
- No longer connected to OrderEntry Service

Standby Connection
- JDBC: Subscribe to ONS

3

Failover Started by Data Guard Broker

HOUSTON Data Center
Partial Site Failover – New Primary

- Standby Connection
  - JDBC: Subscribe to ONS

- AUSTIN Data Center
- HOUSTON Data Center

- 2-node RAC Primary

- Old Primary Existing Connections
  - JDBC: Timeout continues
  - No longer connected to OrderEntry Service

- New Connections Directed to New Primary
  - JDBC: Subscribe to ONS
  - Connected to OrderEntry Service

- Service: OrderEntry auto-started on new primary

- App Server Farm

- Failover Completed by Data Guard Broker

- New Connections Directed to New Primary
  - JDBC: Subscribe to ONS
  - Connected to OrderEntry Service

- Service: OrderEntry auto-started on new primary
Partial Site Failover – Notification

New Connections to New Primary

Old Primary Existing Connections
- JDBC: Timeout continues
- No longer connected to OrderEntry Service

Broker Notification to Terminate Existing Connections
- JDBC: FAN event sent to ONS subscribers

JDBC: ONS Daemon
Service: OrderEntry

JDBC: ONS Daemon
Service: OrderEntry

Data Guard Broker

App Server Farm

2-node RAC Primary

AUSTIN Data Center

HOUSTON Data Center
Partial Site Failover – Application Redirect

App Server Farm

- 2-node RAC Primary

AUSTIN Data Center

- Apps no longer connected to old primary

2-node RAC Primary

Data Guard Broker

HOUSTON Data Center

- JDBC: Subscribe to ONS
- Connected to OrderEntry Service

All Connections Directed to New Primary

JDBC: ONS Daemon
Service: OrderEntry
Partial Site Failover – Steady State

App Server Farm

Primary Connection
- OCI: Send info for Client IP table
- Connected to OrderEntry Service

2-node RAC Primary

OCI: Client IP Info table
Service: OrderEntry

Data Guard Redo Transport

Data Guard Broker

Standby Connection
- OCI: No info exchanged

2-node RAC Standby

AUSTIN Data Center

HOUSTON Data Center

OCI

ORACLE
**Partial Site Failover – Database Failover**

1. **2-node RAC Primary**
   - **OCI:** Client IP Info table
   - **Service:** OrderEntry

2. **Primary Connection**
   - **OCI:** Timeout ensues
   - **No longer connected to OrderEntry Service**

3. **Failover Started by Data Guard Broker**

4. **Standby Connection**
   - **OCI:** No info exchanged

**AUSTIN Data Center**

**HOUSTON Data Center**
Partial Site Failover – New Primary

Old Primary Existing Connections
- OCI: Timeout continues
- No longer connected to OrderEntry Service

New Connections Directed to New Primary
- OCI: Send info for Client IP table
- Connected to OrderEntry Service

Failover Completed by Data Guard Broker

Service: OrderEntry auto-started on new primary

OCI: Client IP Info table

AUSTIN Data Center

HOUSTON Data Center
Partial Site Failover – Notification

New Connections to New Primary

Old Primary Existing Connections
- OCI: Timeout continues
- No longer connected to OrderEntry Service

Broker Notification to Terminate Existing Connections
- OCI: HA AQ event posted for entries in Client IP info table

OCI: Client IP Info table
Service: OrderEntry

OCI: Client IP Info table
Service: OrderEntry

Data Guard Broker

2-node RAC Primary

AUSTIN Data Center

HOUSTON Data Center
Partial Site Failover – Application Redirect

App Server Farm

AUSTIN Data Center

2-node RAC Primary

Data Guard Broker

HOUSTON Data Center

2-node RAC Primary

OCI: Client IP Info table
Service: OrderEntry

All Connections Directed to New Primary
- OCI: Send info for Client IP table
- Connected to OrderEntry Service

Apps no longer connected to old primary
Application Failover

JDBC / OCI Requirements

- Application notification is done via Fast Application Notification (FAN)
- Application configuration for FAN includes
  - JDBC
    - Install Oracle JDBC driver
    - Enable Fast Connection Failover (FCF)
    - Use Implicit Connection Cache
  - OCI
    - Enable OCI_EVENTS and OCI_THREADED mode
    - Link application with threads library
    - Note additional requirements for ODP.Net and OLEDB
Configuration Example – JDBC App
Create Role Based Services

- Services only start up on appropriate database role
- Must be created on both primary and standby clusters

Primary cluster:
```
srvctl add service -d Austin
   -s OrderEntry -r ssa1,ssa2 -l PRIMARY
   -q FALSE -e NONE -m NONE -w 0 -z 0
```

Standby cluster:
```
srvctl add service -d Houston
   -s OrderEntry -r ssb1,ssb2 -l PRIMARY
   -q FALSE -e NONE -m NONE -w 0 -z 0
```

*Do not enable TAF or OCI HA Event (aq_ha_notifications)*
Services that are to be active on standby must be also created on primary in anticipation of switchover / failover

Primary cluster:
```
srvctl add service -d Austin -s OrderReport -r ssa1,ssa2 -l PHYSICAL_STANDBY -q FALSE -e NONE -m NONE -w 0 -z 0
```

Standby cluster:
```
srvctl add service -d Houston -s OrderReport -r ssb1,ssb2 -l PHYSICAL_STANDBY -q FALSE -e NONE -m NONE -w 0 -z 0
```

- For OrderReport service, must run `SRVCTL START SERVICE` and `SRVCTL STOP SERVICE` on the primary: that’s how standby knows about this service through redo transmission
Configure JDBC clients to use a connect descriptor that includes an address list that includes the SCAN address for each site

```
"jdbc:oracle:thin:@" +
"(DESCRIPTION_LIST=" +
  "(LOAD_BALANCE=off)" +
  "(FAILOVER=on)" +
  "(DESCRIPTION=" +
    "(ADDRESS_LIST=" +
      "(LOAD_BALANCE=on)" +
      "(ADDRESS=(PROTOCOL=TCP)(HOST=Austin-scan)(PORT=1521)))" +
      "(CONNECT_DATA=(SERVICE_NAME=OrderEntry)))" +
    "(DESCRIPTION=" +
      "(ADDRESS_LIST=" +
        "(LOAD_BALANCE=on)" +
        "(ADDRESS=(PROTOCOL=TCP)(HOST=Houston-scan)(PORT=1521)))" +
        "(CONNECT_DATA=(SERVICE_NAME=OrderEntry)))" ];
```
After connection made using the JDBC URL:

- Oracle Net contacts DNS and resolves primary SCAN to a total of 3 IP addresses
- Randomly picks one of the 3 IP address and attempts to make a connection
- If the connection to primary site is unsuccessful, it then contacts DNS and resolves standby SCAN to 3 addresses
- It then randomly picks up one of the IP addresses and tries to connect
Configuration Example – JDBC App
Configure for Fast Reconnect

- Setting the `TCP_CONNTIMEOUT_STR` property enables the JDBC client to quickly traverse an address_list

```java
Properties prop = new Properties();
prop.put(oracle.net.ns.SQLnetDef.TCP_CONNTIMEOUT_T_STR, ""+5000); // 5000ms
pds.setConnectionProperties(prop);
```
Enable Fast Connection Failover (FCF) to receive FAN events

Configure Application to connect to all primary and standby ONS daemons

```java
pds.setFastConnectionFailoverEnabled(true);
pds.setONSConfiguration("nodes=hasun05:6200,hasun06:6200,hasun07:6200,hasun08:6200");
```
Considerations for OCI Services

- Service created on both primary and standby clusters
- Must enable TAF and OCI HA Event (aq_ha_notifications)
- Configure DELAY and RETRY parameters

Primary cluster:
```
srvctl add service -d Austin -s OrderEntry -r ssa1,ssa2 -l PRIMARY -q TRUE -e SESSION -m BASIC -w 10 -z 150
```

Standby cluster:
```
srvctl add service -d Houston -s OrderReport -r ssb1,ssb2 -l PRIMARY -q TRUE -e SESSION -m BASIC -w 10 -z 150
```
Considerations for OCI

Enabling FAN

- Enable FAN for OCI clients by initializing the environment with the OCI_EVENTS parameter
  \[\text{OCIEnvCreate}(\ldots\text{OCI_EVENTS}\ldots)\]

- Link the OCI client applications with thread library \texttt{libthread} or \texttt{libpthread}

- Configure application to check for events and register a callback whenever events are received
Considerations for OCI
Oracle Net Alias

• Oracle Net alias should specify both the primary and standby SCAN hostnames

SALES=
 (DESCRIPTION_LIST=
  (LOAD_BALANCE=off) (FAILOVER=on)
  (DESCRIPTION=
   (LOAD_BALANCE=on) (CONNECT_TIMEOUT=10) (RETRY_COUNT=3)
    (ADDRESS_LIST=
     (ADDRESS=(PROTOCOL=TCP)(HOST=Austin-scan)(PORT=1521)))
    (CONNECT_DATA=(SERVICE_NAME=OrderEntry)))
  (DESCRIPTION=
   (LOAD_BALANCE=on) (CONNECT_TIMEOUT=10) (RETRY_COUNT=3)
    (ADDRESS_LIST=
     (ADDRESS=(PROTOCOL=TCP)(HOST=Houston-scan)(PORT=1521)))
    (CONNECT_DATA=(SERVICE_NAME=OrderEntry))))
Considerations for OCI

New Oracle Net Parameters

• Three new parameters in Oracle Database 11g Release 2 used in previous example
  – `CONNECT_TIMEOUT` controls the overall time to connect to the service
  – `TRANSPORT_CONNECT_TIMEOUT` is the amount of time for the TCP connection to complete
    • `CONNECT_TIMEOUT` set to a value slightly greater than `TRANSPORT_CONNECT_TIMEOUT`
  – `RETRY_COUNT` parameter specifies the number of times an address list is traversed before the connection attempt is terminated
Considerations for OCI

New Oracle Net Parameters – Example

• Be careful with optimal values, e.g. if Austin server/clusterware is down:

  SALES=
  (DESCRIPTION LIST=
    (LOAD_BALANCE=off) (FAILOVER=on)
    (DESCRIPTION=
      (LOAD_BALANCE=on) (CONNECT_TIMEOUT=5) (RETRY_COUNT=2)
      (ADDRESS LIST=
        (ADDRESS=(PROTOCOL=TCP) (HOST=Austin-scan) (PORT=1521))
        (CONNECT_DATA=(SERVICE_NAME=OrderEntry)))
    (DESCRIPTION=
      (LOAD_BALANCE=on) (CONNECT_TIMEOUT=5) (RETRY_COUNT=2)
      (ADDRESS LIST=
        (ADDRESS=(PROTOCOL=TCP) (HOST=Austin-scan) (PORT=1521))
        (CONNECT_DATA=(SERVICE_NAME=OrderEntry))))

• A new connection spends $5 \times 3 = 15$ seconds to iterate through 3 Austin-SCAN VIPs
• This is retried 2 times: additional $2 \times 15 = 30$ seconds
• So connection fails over to Houston after $15 + 30 = 45$ seconds
Considerations for Oracle Applications

- PeopleSoft supports seamless client failover
  - Simply follow OCI configuration steps
  - PeopleSoft PeopleTools version 8.50.09 and higher supports FAN
  - Consult MAA whitepaper: “Reducing PeopleSoft Downtime Using a Local Standby Database”,
Considerations for Oracle Applications

• E-Business Suite and Siebel
  – Create role based services
  – Follow Oracle Net alias and JDBC URL examples
  – On the mid tier tune the OS TCP timeout for faster disconnects
Considerations for Oracle WebLogic

– For Current Release
  • Use Multi Pool Data source
    – Multi Pool Data Source contains data sources for both primary and secondary sites
    – Requires Database Service available at primary site only
    – Configure connect timeout at each data source level
    – Does not use FAN: uses a polling mechanism instead
  – Upcoming enhancement
    • Grid Link
      – Integrates with usage of FAN
      – Can consume Data Guard specific FAN events
      – Can integrate with best practices for SCAN addresses with Oracle Database 11.2

Considerations for Data Guard Switchover

**JDBC and OCI**

- You get client failover for switchover free if you have followed the steps (*well….sorta*)

- **Physical standby**
  - Clients disconnected as primary is converted to a standby
  - Clients go through TAF retry logic (OCI) or application retry logic (JDBC)
  - Clients connected to the standby disconnected as it is converted to primary
  - Once both databases come up in new roles, services start and clients reconnect

- **Logical standby**
  - Services are stopped automatically if Data Guard Broker switchover
  - Manually disconnect connections to both primary and standby
  - Perform switchover
  - Once both databases come up in new roles, services start and clients reconnect
Efficient Routing of New Connections

JDBC and OCI

- The example Oracle Net Alias / JDBC URL discussed previously connects to SCAN address as an ordered list
  - Disable `LOAD_BALANCE` at the `DESCRIPTION_LIST` level
    
    `SALES=`
    `(DESCRIPTION_LIST=
    (LOAD_BALANCE=off) (FAILOVER=on)
    (DESCRIPTION=
    )`  
    - First `DESCRIPTION` is always processed first
    - Second `DESCRIPTION` is only attempted once all attempts in first `DESCRIPTION` fail
    - This is the optimal method if failover / switchovers are rare

- If role transitions occur frequently consider randomizing connections between `DESCRIPTION`
  - Enable `LOAD_BALANCE` at the `DESCRIPTION_LIST` level
    
    `SALES=`
    `(DESCRIPTION_LIST=
    (LOAD_BALANCE=on) (FAILOVER=on)
    (DESCRIPTION=`
Efficient Routing of New Connections
Active Data Guard Configuration Example

Traffic Manager

Web Servers

App Servers

Database Servers

Austin
Production Database
OrderEntry Service

Houston
Active Data Guard Standby
OrderReport Service
Efficient Routing of New Connections

Active Data Guard: JDBC and OCI

• For environments that include Active Data Guard disable `LOAD_BALANCE` at the `DESCRIPTION_LIST` level.

• Oracle Net alias / JDBC URL for the primary application should list the primary SCAN hostname in the first `DESCRIPTION`.

• Oracle Net alias / JDBC URL for the read only connections going to Active Data Guard list the standby SCAN hostname in the first `DESCRIPTION`. 
Efficient Routing of New Connections
Active Data Guard TNSNAMES Example

SALES=
(DESCRIPTION_LIST=
 (LOAD_BALANCE=off) (FAILOVER=on)
 (DESCRIPTION=
  (LOAD_BALANCE=on) (CONNECT_TIMEOUT=10) (RETRY_COUNT=3)
   (ADDRESS_LIST=
    (ADDRESS=(PROTOCOL=TCP) (HOST=Austin-scan) (PORT=1521)))
   (CONNECT_DATA=(SERVICE_NAME=OrderEntry)))
 (DESCRIPTION=
  (LOAD_BALANCE=on) (CONNECT_TIMEOUT=10) (RETRY_COUNT=3)
   (ADDRESS_LIST=
    (ADDRESS=(PROTOCOL=TCP) (HOST=Houston-scan) (PORT=1521)))
   (CONNECT_DATA=(SERVICE_NAME=OrderEntry))))

Primary Database Connection
Austin: Primary Database, listed earlier
OrderEntry: Read/Write Service

REPORTS=
(DESCRIPTION_LIST=
 (LOAD_BALANCE=off) (FAILOVER=on)
 (DESCRIPTION=
  (LOAD_BALANCE=on) (CONNECT_TIMEOUT=10) (RETRY_COUNT=3)
   (ADDRESS_LIST=
    (ADDRESS=(PROTOCOL=TCP) (HOST=Houston-scan) (PORT=1521)))
   (CONNECT_DATA=(SERVICE_NAME=OrderReport)))
 (DESCRIPTION=
  (LOAD_BALANCE=on) (CONNECT_TIMEOUT=10) (RETRY_COUNT=3)
   (ADDRESS_LIST=
    (ADDRESS=(PROTOCOL=TCP) (HOST=Austin-scan) (PORT=1521)))
   (CONNECT_DATA=(SERVICE_NAME=OrderReport))))

Active Standby Database Connection
Houston: Active Standby Database, listed earlier
OrderReport: Read-only Service
Considerations for Previous Releases

**JDBC and OCI**

- Oracle Database 11g Release 2 greatly simplified client failover setup and configuration over previous releases
- Earlier releases need some manual configuration for services and FAN events
  - Configure Oracle Net alias to include all node VIP names instead of SCAN names
  - Create triggers that manage application services so that application services are started for the correct database role
  - Configure wrapper script and configuration file for the ONS Publisher
  - Create trigger based on the `DB_ROLE_CHANGE` system event to execute the ONS Publisher wrapper script
Considerations for Previous Releases

Other Noteworthy Items

• **RETRY_COUNT** : new in Oracle Database 11.2
  – Previous releases may need to specifically code additional retries for new connection attempts

• **SQLNET.OUTBOUND_CONNECT_TIMEOUT** in Oracle Database 11.1 and Oracle Database 10.2 can only be set in the sqlnet.ora
  – That means all Oracle Net aliases inherit that one value
Agenda

• Overview
  – Oracle Maximum Availability Architecture (MAA)
  – Application failover types
  – Oracle Data Guard, Switchover / Failover
  – New application failover features in Oracle Database 11g Release 2

• Seamless Application Failover with Data Guard
  – Design principles and requirements
  – Best practices: JDBC and OCI
  – Considerations for Oracle Applications & WebLogic
  – Considerations for previous releases

• VocaLink Case Study
  – VocaLink: business summary
  – Description of VocaLink MAA configuration
  – Failover configuration & business benefits
VocaLink and BGC Application Failover

Martin McGeough
Database Technical Architect
### VocaLink: at the heart of the transaction
A specialist provider of payment transaction services

| Our history | Driving automated payments for more than 40 years
|            | From domestic supplier to large-scale international provider of modern payment services |
| Our scale | We securely process over 9 billion payments a year, including 15% of all European bank-to-bank payments |
|            | On a peak day the payment platform processes over 90 million transactions and its switching technology powers the world's busiest network of over 60,000 ATMs |
| Our customers | The world’s top banks, their corporate customers and Government departments |
| Our services | Real-Time Payments | Euro Services | Sterling clearing services | Connectivity | Cards and ATM services |
| 2009 awards | Best payment system deployment (Faster Payments Service) |
|            | Best outsourcing partnership (BGC) |
|            | Overall winner (Faster Payments Service) |
VocaLink’s history with Oracle

• In 2007 VocaLink initiated a joint programme with BGC – the Swedish clearing house - to renew the BGC Service – the goal was to replace the heritage mainframe technology with a highly scalable, highly available and modular infrastructure that would reduce costs while simultaneously improving performance.

• Main Architecture requirements
  – High Availability – No single points of failure
  – Disaster Recovery – remote site with zero data loss
  – Site Failover – Site failover SLA is 15 minutes
  – High Throughput – Process payments within very tight SLA’s
  – Manageability – Ability to manage independent payments services
VocaLink’s history with Oracle contd.

• Technology
  – WebLogic Server was chosen to provide the Application Server software
  – Sun/Oracle was chosen to provide the hardware.
  – Oracle Database was chosen to provide the Database software using RAC (Real Application Cluster) for a solution that provided high availability and performance and Data Guard + Data Guard Broker provided an easily managed DR solution.

• Oracle’s MAA framework promotes and supports VocaLink’s Architecture requirements by providing the tools and a tested and supported set of configurations.
BGC Architecture

**Primary Site**

**Application Tier – WLS 10.3**

Connection Pools using loadbalance multipools

**Database Tier – 11gR1**

High-speed interconnect

Database instance 1

Database instance 2

RCVR

**Standby Site**

**Application Tier – WLS 10.3**

Connection Pools using loadbalance multipools

**Database Tier – 11gR1**

High-speed interconnect

Database instance 1

Database instance 2

LGWR (Sync)

RCVR

Datafiles

Redo Logs

Flashback Logs

Standby Redo Logs

Datafiles

Redo Logs

Flashback Logs

Standby Redo Logs
BGC Architecture – Explained

• The Standby site has two DWDM links the longest is 90Km and the network latency has been measured at just under 3ms, the Data Guard protection level is set to Max Availability – this allows the primary to keep processing even through loss of the standby servers or the inter-site link.
BGC Architecture - Explained

- At present the Active Data Guard standby is not used for anything other than ensuring the Standby application has a warm cache of Reference Data to enable VocaLink to meet the 15 minute SLA, however for a future release of the BGC service it is being considered for providing read-only reporting application. This is reliant on 11gR2 and the automated monitoring of the latency between the primary and standby databases providing a suitable framework.
BGC Architecture - Explained

• Each site is mutually exclusive at present therefore the WLS tiers are only configured to connect to the local database servers. However with the use of the load-balanced multipools configurations has reduced the amount of manual interventions required after the old primary databases are restarted after a Switchover or other failure scenarios i.e. an instance crash or a physical server crash.
VocaLink – BGC Service Site Failover

• Control of the Failover/Switchover process are controlled by a job scheduler (Tivoli Workload Scheduler) due to the external dependencies above the Application and Database tiers – this controls the Data Guard Broker and the application restart processes.

• A failover has been timed at 12 mins whilst a switchover has been timed at just under 15mins (14:45).

• The failover includes network, proxy, application and database tiers. The database failover takes about 5 minutes and the application tier takes about a further 3 minutes. The rest of the time was taken by the network and some ancillary systems switching sites.
VocaLink – BGC Service Site Failover

• During a Site Failover of the BGC Service VocaLink use the following Oracle technologies to meet the strict 15 minutes SLA - Data Guard, Data Guard Broker, Flashback and Active Data Guard.

• The Data Guard configuration allows for the data to be replicated to VocaLink’s standby site without data loss.

• Data Guard Broker is used to simplify the process of performing a Failover/Switchover – one command to control the whole process.
VocaLink – BGC Service Site Failover

• Active Data Guard is implemented along with a custom caching application to keep the reference data up to date on the standby site – allowing the application to restart very quickly after a failover is initiated.

• Flashback is implemented to allow us to quickly rebuild a failed primary database.

• The WLS multi pool configuration along with Active Data Guard allows the standby site to be maintained and kept up to date with the minimal interaction from the system administrators.
Key HA Sessions, Labs, & Demos by Oracle Development

Monday, 20 Sep – Moscone South *
3:30p Extreme Consolidation with RAC One Node, Rm 308
4:00p Edition-Based Redefinition, Hotel Nikko, Monterey I / II
5:00p Five Key HA Innovations, Rm 103
5:00p GoldenGate Strategy & Roadmap, Moscone West, Rm 3020

Tuesday, 21 Sep – Moscone South *
11:00a App Failover with Data Guard, Rm 300
12:30p Oracle Data Centers & Oracle Secure Backup, Rm 300
12:30p Oracle Net Services: Best Practices, Rm 302
2:00p ASM Cluster File System, Rm 308
2:00p Exadata: OLTP, Warehousing, Consolidation, Rm 103
3:30p Deep Dive into OLTP Table Compression, Rm 104
3:30p MAA for E-Business Suite R12.1, Moscone West, Rm 2020
5:00p Instant DR by Deploying on Amazon Cloud, Rm 300

Wednesday, 22 Sep – Moscone South *
11:30a RMAN Best Practices, Rm 103
11:30a Database & Exadata Smart Flash Cache, Rm 307
11:30a Configure Oracle Grid Infrastructure, Rm 308
1:00p Top HA Best Practices, Rm 103
1:00p Exadata Backup/Recovery Best Practices, Rm 103
4:45p GoldenGate Architecture, Hotel Nikko, Peninsula

Thursday, 23 Sep – Moscone South *
10:30a Active Data Guard Under the Hood, Rm 103
1:30p Minimal Downtime Upgrades, Rm 306
3:00p DR for Database Machine, Rm 103

Demos Moscone West DEMOGrounds
Mon & Tue 9:45a - 5:30p; Wed 9:00a - 4:00p
Maximum Availability Architecture (MAA)
Oracle Active Data Guard
Oracle Secure Backup
Oracle Recovery Manager & Flashback
Oracle GoldenGate
Oracle Real Application Clusters
Oracle Automatic Storage Management

Hands-on Labs Marriott Marquis, Salon 10 / 11
Monday, Sep 20, 12:30 pm - 1:30 pm Oracle Active Data Guard
Tuesday, Sep 21, 5:00 pm - 6:00 pm Oracle Active Data Guard

* All session rooms are at Moscone South unless otherwise noted
* After Oracle OpenWorld, visit
http://www.oracle.com/goto/availability

ORACLE