Volkswagen AG

Deploying Oracle Streams and Data Guard
High Availability and Disaster Recovery on a Global Basis

"The combination of Oracle Streams and Data Guard saves Volkswagen between $100,000 and $200,000/hour in downtime costs when failures occur."

Holger Urban
Volkswagen AG

Volkswagen AG

Corporate Profile

- Largest carmaker in Europe
- Over 5.7 million vehicles delivered in 2006
- Eight different brands
- 325,000 employees.
- 50 production plants located in Europe, Americas, Asia, and Africa
- http://www.volkswagenag.com

Oracle Technology Used

- Oracle Database (10.2.0.2)
- Oracle Streams
- Oracle Data Guard
- Oracle Recovery Manager (RMAN)

OVERVIEW

The Volkswagen Group with its headquarters in Wolfsburg Germany is one of the world's leading automobile manufacturers and the largest carmaker in Europe. The Group is made up of eight brands from six European countries: Volkswagen, Audi, Bentley, Bugatti, Lamborghini, SEAT, Škoda and Volkswagen Commercial Vehicles. The Group operates 44 production plants in twelve European countries and a further six countries in the Americas, Asia and Africa.

The Product Data Management (PDM) department at Volkswagen supports engineering teams operating at multiple locations spread across four continents. From an information technology perspective, this high degree of geographic separation is completely transparent with regard to information access, management, and high availability. A central repository of engineering metadata serves as the hub of Volkswagen's information technology architecture. Spoke databases are synchronized with the hub to provide remote engineering organizations local access to a replica of the central repository.

Volkswagen uses Oracle Streams [1] bi-directional replication to implement their hub and spoke architecture. The hub database serves three functions. It consolidates engineering data into a single global repository. It replicates key tables to remote locations so that engineering teams have fast, reliable, local access to engineering data. And it provides remote locations continuous access to data in the event that their local spoke database becomes unavailable for any reason.

The hub database is protected against failures and data corruptions by a Data Guard [2] standby database. In the event the hub should fail, a Data Guard failover enables the standby database to assume the production role of the hub and maintain synchronization with spoke databases, insuring that there is no single point of failure that can impact the availability of the entire system.

This combination of Oracle Streams and Oracle Data Guard provide Volkswagen a level of high availability (HA) and data protection to prevent unplanned downtime, which can cost between $100,000 and $200,000 per hour, per site.

This case study details the use of Oracle Streams and Oracle Data Guard in the high availability architecture deployed by the Volkswagen Group.

VOLKSWAGEN REQUREMENTS

As a multinational automobile manufacturer with distributed engineering teams located in facilities on four different continents, Volkswagen faced a considerable challenge to make the geographic separation between teams transparent in terms of information access and availability. Remote engineering teams require continuous access to an up-to-date global repository of engineering data. A solution was required that could:

1. Consolidate the activities of distributed engineering teams located at seven geographically remote locations into a single global repository for engineering data.
2. Provide remote locations fast, local access to the global repository.
3. Efficiently utilize network bandwidth by providing flexible configuration options to minimize the volume of data transmitted over the WAN.
4. Be easy to manage, and be easy to provide services should new remote locations be added at a later date.
5. Be resilient to failures at any level – providing for both high availability and disaster recovery.
6. Be resilient to change – application upgrades, database upgrades, and database maintenance with little or no downtime.
7. Be flexible – to accommodate different platforms, operating systems, and HA architectures implemented by remote locations and various subsidiaries within the Volkswagen Group.

Volkswagen determined that the combination of Oracle Streams replication and Oracle Data Guard, both included features of the Oracle Database, was the best solution for this set of requirements.

ORACLE STREAMS REPLICATION

Oracle Streams captures DML and DDL changes made to database objects and replicates those changes to one or more destination databases. The destination databases allow DML and DDL changes to the same database objects, and these changes can also be propagated to other databases in the environment should the user desire. Streams can also be configured to propagate changes between one or more databases bi-directionally. The tables for which data is shared do not need to be identical copies at all databases; both the structure and the contents of these tables can differ at different databases. Please refer to Oracle Streams Concepts and Administration [3] for a complete discussion of capabilities.

Using Streams for High Availability (HA)

Streams can be used to implement a distributed HA architecture that includes multiple independent databases, each synchronized with the other such that should one of the databases fail, applications can be redirected to a surviving database and continue processing. The simplest HA configuration deployed in this fashion would include two databases synchronized using Streams bi-directional replication. This approach has several benefits:

- All resources are fully utilized. All databases in the Streams configuration are open for read-write transactions.
- There is a high degree of autonomy between databases in a Streams configuration. Hardware architecture, OS, and Oracle version can be different. Database structure and table structures can be different.
- Replication is very configurable. Replicate only the data that must be replicated. Transformations and other customizations are possible. Data may even be replicated to non-Oracle targets.

High Availability & Disaster Recovery in a Hub and Spoke Configuration

A Hub database becomes a very important node in a distributed environment. Should it fail, all communications flowing through the Hub will fail. This concern is addressed using the same techniques used to make any Oracle database resilient to failures. Oracle RAC can be used to protect against instance and node failures. Whether or not Oracle RAC is utilized, all Hub databases should have an accompanying Data Guard physical standby database to protect from disasters and data errors. Volkswagen has deployed Data Guard Redo Apply (physical standby database) to protect its Hub database. Details are provided in the Data Guard section below.

The flexibility of Streams allows this model to be extended to additional databases connected in an n-way configuration, each capable of being a failover target for any database in the configuration. This model works best when there is a reliable, low latency connection between sites. In Volkswagen's environment, the high network latency that is a byproduct of the geographic separation between locations combined with concerns for managing replication between a large number of databases led to the selection of a variant of an n-way configuration called a "hub and spoke" architecture.

A hub and spoke configuration funnels changes from many "Spoke" databases into a central "Hub" database. The hub database manages the replication of changes originating from each spoke back to the other spokes in the configuration. New source or destination databases are simply connected to the hub database, rather than needing to establish connections to every other database. The hub database provides a single point of control that greatly simplifies management of the configuration.

IMPLEMENTING A HUB AND SPOKE CONFIGURATION

Users who are approaching Streams for the first time are encouraged to utilize the Streams configuration tool in Enterprise Manager or the single procedure provided in the `DBMS_STREAMS_ADM` package. Both approaches will quickly configure all of the Streams components for a replication environment with two databases. Additional destination databases can then be added to create a hub and spoke configuration.

Volkswagen chose to implement by directly interfacing with Streams API's for maximum flexibility as documented in the Oracle Streams Replication Administrators Guide [4] for setting up a hub and spoke configuration. An overview of Volkswagen's implementation is provided in the sections below.

In a hub and spoke configuration, changes are captured, propagated, and applied in the following way:

- The primary database (hub) captures local changes to the shared data and propagates these changes to all secondary databases (spokes), where these changes are applied at each secondary database locally.
- Each spoke database captures local changes to the shared data and propagates these changes to the hub database only, where these changes are applied at the hub database locally.
- The hub database applies changes from each spoke database locally. Next, these changes are captured at the hub database and propagated to all spoke databases, except for the one at which the change originated. This avoids a problem referred to as "change cycling", where changes are sent back to the database from which they originated. Each spoke database applies the changes from the other databases locally, after they have gone through the hub database.

System Configuration

Hub - Central Repository

- 16-processor Sun, 32GB memory, Solaris 10 64-bit
- 1 TB database
- Oracle Database 10g Release 2 (10.2.0.2)
- Oracle Streams bi-directional replication
- Data Guard Redo Apply – physical standby database

Spokes - Remote Databases

- 6 remote databases in different countries on 4 continents
- Various hardware and O.S. architectures
- Network links with widely varying bandwidth and RTT latencies

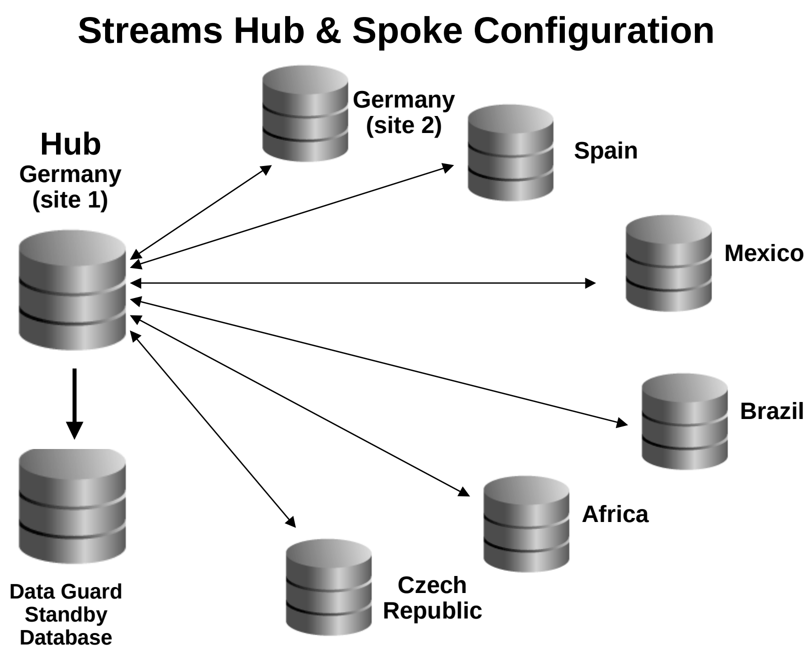

Figure 1

Hub Configuration Details

A 1 TB Oracle Database (10.2.0.2) serves as the central repository of metadata for engineering documents pertaining to automotive construction across Volkswagen product lines. This central repository is the hub database, residing on a 16-processer SUN server with 32GB of memory running 64-bit Solaris 10 operating system.

There are three steps to the Streams replication process – Capture, Propagation, and Apply, used by the hub to maintain synchronization with the spokes. Because Volkswagen's hub and spoke configuration uses bi-directional replication, these same steps occur for data changes originating from the spoke as well. While the basics of replication are described below, please note that the Streams documentation includes many ways to customize how and where these processes are executed in order to address a wide range of replication requirements.

- **Capture:** A Streams capture process creates one or more logical change records (LCRs) and queues them to a Capture Queue. An LCR is a message with a specific format that describes a database change. A capture process reformats changes captured from the redo log into LCRs. If the change was a data manipulation language (DML) operation, then each LCR encapsulates a row change resulting from the DML operation to a shared table at the source database. If the change was a data definition language (DDL) operation, then an LCR encapsulates the DDL change that was made to a shared database object at a source database.
- **Propagation:** Streams propagation propagates the staged LCR to another queue residing in the destination database where apply will occur.
- **Apply:** Once the LCR has reached the destination database, a Streams apply process consumes the change by applying the LCR to the shared database object.

The hub database is configured with a single Streams Capture Queue. A separate Propagation and Apply process is defined for each remote spoke. Volkswagen has defined capture by table in order to minimize the volume of data that needs to be transmitted over the Wide Area Network (WAN). The hub database has over 200 tables in the base schema from which only one third must be replicated to maintain the required level of synchronization. For example, local configuration tables and local journaling tables are excluded from the replication process. For simplicity, Volkswagen is able to globally define Streams Propagation and Apply properties such that capture and apply queues are handled in the same way throughout the hub and spoke configuration.

Efficient Network Utilization

Streams provides complete control over the replication process reducing the amount of data Volkswagen transmits over the WAN by 80%. The ability to restrict replication to specific tables, DML, and DDL – makes it possible for a Hub and Spoke configuration to function on high latency networks spanning thousands of miles.

The ability to perform granular replication of a subset of the database is an important attribute of the configuration given the limited network bandwidth and high network latency between the hub and spokes, some of which are located as far as 6,000 miles away. In contrast, complete replication of all redo data generated by the hub database would result in 5x the data volume compared to the more granular replication enabled by Oracle Streams. Without the advanced features of Streams, Volkswagen would find it prohibitively expensive to replicate the entire contents of the hub database over such distances and latencies given the volume of data that would be required.

Volkswagen uses Streams "Tags" to control how DDL and DML are captured, propagated, and applied by the hub and spoke databases. Every redo entry in a redo log has a tag associated with it. By default, when a user or application generates redo entries, the value of the tag is NULL for each redo entry, and a NULL tag consumes no space. Configuring how tag values are interpreted provides fine-grained control over replication. For example, a tag can be used to determine whether an LCR contains a change that originated in the local database or at a different database. This is how Volkswagen avoids "change cycling" – a circumstance where the Hub would send an LCR back to the spoke database where it originated. Propagation will only occur to spoke databases that did not originate the change.

When operational:

- The hub receives and applies LCR's propagated from each spoke (using the separate Apply process defined for each spoke).
- These changes and other local changes made at the Hub to tables defined for replication by Streams are in-turn captured by Streams (via the single capture process defined).
- Streams propagates the changes back to the spokes (using the separate Propagation process defined for each spoke).

Spoke Configuration Details

There are six spoke databases located in different countries scattered across four continents. Since each spoke database is only concerned with communicating with the hub, each is configured with a single Capture Queue and a single Apply Queue. Each spoke inherits the same Streams Propagation and Apply definitions that have been defined globally for Volkswagen's configuration.

Local sites have utilized the flexibility of Streams in the following ways:

- The remote locations can utilize platforms that are the most convenient for them to support since there is no requirement for common hardware architecture or operating system.
- Remote locations have the flexibility to implement their own local HA architecture. One of the remote locations has chosen to implement an additional level of "local" HA to minimize the frequency of redirecting local users to the remote hub by using Oracle RAC to protect against server failure. The site has configured its spoke using an Oracle RAC [5] database running on two servers. In the event a server fails, the surviving server assumes the workload of the failed system and local processing is unaffected. If a catastrophic event causes the last surviving server to fail, users are then directed to the remote hub.
- A second remote site has utilized Streams to more easily maintain their local test systems. This site uses the downstream capture feature of Streams to synchronize a local replica of their spoke database for test purposes. Downstream capture offloads the overhead of capture, propagation and apply to the test database, avoiding overhead on the spoke database. Streams has proven to be a faster and more flexible mechanism for maintaining a test system than standard cloning procedures. Streams can selectively synchronize tables on the test system at the desired interval to make it current with the production database while keeping previous test data and other aspects of the test database intact.

Automatic Conflict Detection and Resolution

Conflicts can occur in a bi-directional hub and spoke configuration due to concurrent data manipulation language (DML) operations on the same data at multiple databases. A conflict is a mismatch between the old values in an LCR and the expected data in a table. Typically, a conflict results when the same row in the source database and destination database is changed at approximately the same time. The Streams Apply process automatically detects conflicts caused by row LCRs. Streams offers prebuilt handlers enabling a conflict resolution system to be defined that resolves conflicts in accordance with specific business rules. If a conflict cannot be resolved automatically, or if a handler procedure raises an error, then all messages in the transaction that raised the error are saved in the error queue for later analysis and possible re-execution.

Automatic Conflict Detection

Any bi-directional replication solution must provide facilities to deal with conflicts that are caused when the same row in both a source and destination database are changed at approximately the same time. The Streams apply process automatically detects when such conflicts occur, and provides built-in handlers that administrators can use to resolve them automatically.

Because the Streams Apply process cannot detect DDL conflicts, it is wise to designate a single site as the primary administrative location for DDL changes. At Volkswagen, by default, DDL changes are not replicated. This allows each site to perform local DDL, such as adding an index to a table, as needed. Volkswagen administrators can choose to selectively enable replication of DDL from the hub location. This allows changes resulting from application upgrades to be automatically propagated and applied to all spokes.

The nature of the workload in Volkswagen's configuration minimizes the likelihood that DML conflicts will occur. While engineering teams frequently require read-access to data that originates from other locations, they typically only modify data for projects that are owned locally. This minimizes the chance of row conflicts caused by simultaneous updates at multiple sites. If such conflicts do occur, Volkswagen utilizes the prebuilt `OVERWRITE` update conflict handler, one of several prebuilt handlers provided by Streams. When a conflict occurs, the `OVERWRITE` handler will replace the current value at the destination database with the new value in the LCR from the source database.

WORKLOAD & WIDE AREA NETWORK

Database workload for the 1 TB central metadata repository is moderate. Peak volumes occur during application upgrades and other database maintenance when updates to the hub database can generate up to 1MB/second of redo data. Because only a subset of tables are replicated to each spoke database, the volume transmitted over the WAN to each of the remote locations is less than .2MB/second. Volkswagen maintains a minimum WAN bandwidth of 1Mbps between the hub and each spoke database to insure that Streams replication is provided sufficient bandwidth to keep all spoke databases synchronized with the hub.

MONITORING STREAMS REPLICATION

The hub and spoke architecture dramatically simplifies the monitoring of Volkswagen's Streams configuration. This is accomplished by creating a "heartbeat table", enabling the administrator to determine the status of the complete configuration in a single view. Inside the heartbeat table, each row represents information about a location (hub and spokes), using ID, Database Name (or Global Name) and a timestamp. The hub and spoke databases periodically update the table with timestamps for the most recent data replicated to each site. The table also includes other attributes for apply, propagation, capture, queuing, and apply error information. The Streams Replication Best Practices section of the Oracle Streams Replication Administrators Guide [4] contains details on creating a heartbeat table as well as other recommended administrative practices in a Streams environment.

STANDBY DATABASE FOR THE HUB DATABASE

Data Guard Redo Apply (physical standby database) is used to protect against failures and data corruptions on the hub database. The standby database is located in a separate facility 5 kilometers away from the hub database. The standby database runs on a 12-processer SUN server with 24GB of memory. Data Guard is configured using Maximum Performance protection mode.

Data Guard Configuration

Standby Database for Hub

- 12-processor Sun, 24GB memory, Solaris 10 64-bit
- Data Guard Maximum Performance Protection Mode – ARCH redo transport services

The Data Guard physical standby database is initially created from a backup of the hub. Once created, Data Guard maintains synchronization by shipping redo data generated by the hub database and applying the redo data to the standby database. The proximity, high bandwidth, and low latency of the network link between the hub and standby database can easily accommodate the 1MB/second peak volume required to maintain a complete replica of the hub. The standby database is an exact copy of the hub in every respect; even to the point where it is possible to off-load online backups to the standby database and eliminate the overhead of taking backups on the hub system [6].

Volkswagen is able to use a Data Guard switchover to move production from the hub to the standby database to perform planned hardware and operating system maintenance on the hub with minimal downtime. In the event that the hub should fail unexpectedly, a Data Guard failover is used to transition the standby database to the production role of the hub [7].

The Data Guard standby database is a critical element of the hub and spoke configuration. Without a functioning hub database, the activities of the remote engineering organizations are not consolidated into a single global repository. If the hub is not available, spoke databases will quickly contain stale data and remote sites will be subject to extended downtime should a second failure occur that impacts the spoke database. Data Guard effectively isolates failures that impact the hub database and insures that there is an identical replica of the hub database ready to assume processing in the event the hub should fail.

Additional benefit can be realized due to the fact that the Data Guard standby apply process does not consume the full resources of the standby server while in standby role. Thus the standby server also hosts other Oracle databases used for testing and quality assurance. These non-critical activities are deferred in the event a failover or switchover promotes the standby database to the primary role, insuring all system resources are available to support production.

BACKUP AND RECOVERY

Oracle Recovery Manager (RMAN) [8] is used to implement a common backup strategy across locations. In the event a spoke database experiences a failure, local applications are redirected to the hub database while the error is corrected. If the spoke database needs to be recreated, this accomplished using the most current RMAN backup and resynchronizing with the hub database.

CONCLUSION

Using Oracle technologies, Volkswagen has engineers located in Germany, Mexico, Brazil, Africa, Spain and the Czech Republic that all have access to the latest information on engineering documents pertaining to automotive construction across Volkswagen product lines. Oracle Streams is used to simultaneously consolidate updates that can occur at any one of seven remote facilities into a central metadata repository, while updating local replicas at each remote facility with updates originating from other facilities.

Volkswagen uses the flexibility of Streams replication to accommodate the limited bandwidth and high network latency given geographic separation between sites that extends up to 6,000 miles. This same flexibility is used to configure a Streams hub and spoke configuration such that management can continue to scale as new remote locations are added to the configuration. The Streams hub provides high availability for remote locations should a local database fail – applications are simply redirected to the hub until the problems at a local level are resolved.

A standby database maintained by Oracle Data Guard insures that a hub database is always available to keep spoke databases synchronized and to continue to provide remote locations with high availability in the event of a simultaneous failure of a hub and spoke database.

Volkswagen has used included features of the Oracle Database (Streams, Data Guard, RMAN) to successfully implement a high availability architecture that fully utilizes all computing resources, provides the necessary level of autonomy to remote locations, and is able to maintain a high level of service across a WAN that spans four continents.

REFERENCES

1. Oracle Streams
   http://www.oracle.com/technology/products/dataint/index.html
2. Oracle Data Guard
   http://www.oracle.com/technology/deploy/availability/htdocs/DataGuardOverview.html
3. Oracle Streams Concepts and Administration
   http://download-west.oracle.com/docs/cd/B19306_01/server.102/b14229/strms_over.htm
4. Oracle Streams Replication Administrators Guide
   http://download-west.oracle.com/docs/cd/B19306_01/server.102/b14228/toc.htm
5. Oracle RAC
   http://www.oracle.com/technology/products/database/clustering/index.html
6. Using Recovery Manager (RMAN) with Oracle Data Guard
   http://www.oracle.com/technology/deploy/availability/pdf/RMAN_DataGuard_10g_wp.pdf
7. MAA Best Practices: Data Guard Switchover and Failover
   http://www.oracle.com/technology/deploy/availability/pdf/MAA_WP_10gR2_SwitchoverFailoverBestPractices.pdf
8. Oracle Recovery Manager - RMAN
   http://www.oracle.com/technology/deploy/availability/htdocs/rman_overview.htm

The Volkswagen Group and Oracle Streams
July 2007
Authors: Volker Kuhr, Maria Pratt, Joseph Meeks

Oracle Corporation
World Headquarters
500 Oracle Parkway
Redwood Shores, CA 94065
U.S.A.

Worldwide Inquiries:
Phone: +1.650.506.7000
Fax: +1.650.506.7200
oracle.com