Business Comparison: Data Replication Solutions

Choosing the Best Option for Data Replication in Oracle Environments: Evaluating and Comparing Dell and Oracle Solutions

Contents

- What to Look for in a Data Replication Solution .................................................. 2
- Consider Your Options: Dell and Oracle ............................................................... 4
- Comparing Dell SharePlex With Oracle Data Replication Solutions ...................... 6
- Conclusion ........................................................................................................... 10

Data replication has always been a vital requirement for IT organizations looking to make their systems more agile, flexible and collaborative. But significant changes in data replication applications, technologies and best practices are forcing IT leaders to rethink their strategies for how to design, deploy and manage data replication solutions.

For instance, take today’s reality of multiple-database environments. Long gone is the era of the so-called “Oracle shop,” even though Oracle remains the number-one database management system in the industry. In the past two years, Oracle’s market share slipped below the 50% mark: research firm IDC said Oracle 2012 market share was 45%, followed by Microsoft with 20% and IBM with 18%. That means that it’s typical to have enterprise database infrastructure with Oracle and one or more additional databases — especially with the popularity of open source databases like Hadoop and others.
Companies are looking for ways to deal with the demands of the proliferation of data and big data. Solutions now must provide cost-effective and flexible support for the movement of Oracle data to their business intelligence platforms. And, solutions must be able to support the requirement for multi-location databases that have full data accuracy and integrity, and are updated in lockstep in real time in order to ensure that every user sees and accesses the same data at exactly the same time.

Doing so now — especially in the face of smaller and less-specialized in-house staffs and tighter budget constraints — means that IT organizations need to look for new solutions that offer increased flexibility, optimized performance and improved cost efficiency for the total solution. These solutions must offer greater functionality that is designed for tight integration and easier, faster deployment. Functionality such as reporting, disaster recovery, data comparison, load balancing, migration, real-time change data capture, data transformation, ETL, data warehousing and monitoring/management provide maximum value.

What to Look for in a Data Replication Solution

IT decision-makers looking for data replication solutions for maximum business continuity, advanced analytics, data warehousing, business intelligence and other applications should take time to create a functionality requirements framework for data replication solutions. Having a well-considered set of “must-have” features and functions will go a long way toward helping make a decision on specific products and suppliers to meet their needs. At the same time, it’s essential to think of this decision as part of a long-term strategy that can affordably evolve and scale as business requirements and technical infrastructure evolve over time, so think about “buying ahead” of your current requirements.

• Integrating Oracle data. Even though Oracle is still king of the hill when it comes to relational DBMS installations, non-Oracle databases have gained significant momentum in recent years. Oracle data may need to get to or integrate with tried-and-true competitive databases such as SQL Server, Sybase, DB2, MySQL, Netezza and Teradata. These databases have dramatically improved their functionality and appeal to data-intensive enterprises, including next-generation data platforms such as Hadoop, MongoDB and others that have taken root in big data applications. This means that your data replication tools must have the flexibility to be able to cost-effectively and quickly support multiple-database environments and the new technologies introduced to the market.

• Real-time copies of production data without impacting OLTP performance and availability. As more and more organizations conduct operations around the clock, the idea of interrupting business-critical transactions to and from production databases in order to replicate data to secondary databases or remote locations is unacceptable. Today’s data replication solutions must be continuously copied from one database server to another without pre-determined interruptions, including replication of any data changes on the source database to the target database. This ensures that data on all databases are always aligned and that the correct data is always available. Ideally, your solution should employ a continuous streaming process outside the database instance in order to minimize the impact on performance and network capacity.
• **Data accuracy.** In order to ensure that data targets are fully accurate representations of source databases, look for solutions that employ compare-and-repair functionality to unearth anomalies between target and source for fast, full remediation. Be sure that your solution includes these tools in its native implementation, rather than being “bolted on” post-deployment. Peer-to-peer monitoring and conflict detection and resolution are key requirements for hyper-accurate data replication.

• **Support for key Oracle advanced security features.** With the lion’s share of enterprise and mid-sized organization database installations, Oracle compatibility is essential from a security standpoint. You’ll want support for such features as Transparent Data Encryption (TDE) and Hybrid Columnar Compression (HCC).

• **Transaction modeling.** Database administrators need statistical analysis of database transactions on specific tables during specified timeframes in order to understand table relationships so they can configure optimal throughput for replication. This will ensure transaction consistency between the source and target database as well as identify the need for additional objects to be considered for replication.

• **Support for high availability and disaster recovery.** Ideal data replication solutions are built upon flexible architectures that enable both directional and unidirectional data replication for maximum availability, as well as fast automated failover to secondary databases in the event of an unplanned service interruption. The disaster recovery system should be validated to ensure access to the secondary system with minimal service interruption. The ideal solution should allow for testing of the disaster recovery process to occur periodically.

• **Consider your deployment strategies.** Few, if any, organizations can afford having curtailed access to production databases while replication solutions are being deployed. Be sure you can maintain productivity by ensuring ongoing access to production databases while setting up the new environment and adding to existing environments.

• **Look for solutions with flexible design for reporting.** Rather than having to replicate an entire database to generate reports, there are solutions that allow you to replicate only the tables required for reporting. This saves you considerable cost in data storage, as well as allowing for faster and more efficient reporting.

• **Latency is the enemy of transaction-intensive databases, and should be minimized in every way possible.** Latency used to be thought of as only a storage hardware issue, but no longer. Data replication solutions have to be able to move Oracle data with very low latency in order to avoid processing bottlenecks and ensure real-time access to data. Real-time replicas can be deployed on a different database running on low-cost hardware.
Consider Your Options: Dell and Oracle

While there are a number of data replication tools on the market, the most widely adopted tools for data replication of Oracle-to-Oracle and Oracle-to-other environments come from either Oracle itself or Dell.

Oracle remains the leading database company in the world, even in the face of major competitive inroads made by Microsoft, IBM and others in recent years. Research firm IDC said Oracle held a 45% market share in 2012, followed by Microsoft at 20% and IBM with 18%. Oracle has offered a variety of both in-house-developed replication tools and products from third-party vendors that were acquired to round out Oracle’s portfolio. Oracle’s primary solutions in this area are GoldenGate, Data Guard/Active Data Guard and Streams.

Oracle GoldenGate

GoldenGate 11g Release 2 is Oracle's latest version of the product it assimilated in 2009 when Oracle bought GoldenGate. This product is described by Oracle as a “high-performance software application for real-time transaction change data capture, transformation and delivery, offering log-based bidirectional data replication.”

GoldenGate 11g Release 2’s feature set includes:

- Integrated Capture to support Oracle Database and Oracle Exadata compression types.
- Expanded support for global deployments.
- Increased extensibility by taking defined actions based on event records in the transaction log or Trail file.
- Intelligent Conflict Detection and Resolution for Active-Active and Multi-Master implementations.
- Expanded heterogeneity with IBM DB2, PostgreSQL, SQL Server, Sybase, MySQL, Teradata and IBM z/OS.
- Certified near-zero downtime upgrade for Oracle Communications Billing and Revenue Management applications.

The new release is said to capture and deliver change data in real time to production data stores, reporting systems, data warehouses and other OLTP databases with minimal impact. Its three-component architecture — Capture, Trail Files and Delivery — is reportedly designed to let each component perform its task independently of the others, which helps to speed replication and provide the highest data integrity.

Oracle Active Data Guard

Active Data Guard is a paid add-on solution to Oracle’s free Data Guard disaster recovery software available only for Oracle Database Standard Edition. Active Data Guard is marketed by Oracle as a way to “transform your disaster recovery system into a reliable, productive asset used to offload read-only workloads from your production database.” It is designed to be a data loss protection and prevention solution, working in concert with a data replication solution like GoldenGate, rather than performing as a data replication platform itself.
Active Data Guard performs as a standby database, providing an exact copy of the primary database; it delivers open, read-only access, while continuously applying changes sent from the primary database. As an active standby database, it helps offload queries, reporting and incremental backups from the primary database, helping the production database’s performance, responsiveness and availability.

The product’s architecture is designed to allow standby databases to be isolated from failures that may occur at the primary database. Changes from the primary are sent directly from memory in order to shield the standby database from possible errors and anomalies that may take place at the primary. Corruption checks are performed in order to insure that data is in exact alignment before data is copied to a standby database.

In the event of an outage affecting availability of the production database, Active Data Guard is transitioned to the target standby database function to prevent downtime and data loss. That transition can be done either manually or automatically.

**Oracle Streams**

Streams is a long-standing, legacy tool used to help Oracle users move information within and between databases. It has been offered only on Oracle’s Enterprise Edition systems since 2002, and has largely been replaced as a data replication solution for new deployments by Oracle GoldenGate after that product was integrated into the Oracle mix following the 2009 acquisition of GoldenGate. Oracle has indicated that its strategic focus for data replication going forward is now on GoldenGate and that Streams will receive significantly less focus as a data replication platform.

In its November 2012 “GoldenGate Statement of Direction” designed to help explain Oracle’s strategic plans for data integration and data replication products, Oracle noted the following about Streams:

“Oracle Streams is a built-in feature of the Oracle Database that allows information sharing among multiple Oracle databases. With Oracle Streams, transactional changes and events are captured, propagated and applied within an Oracle database or between Oracle databases, creating a flexible replication solution for a homogeneous all-Oracle environment.

“Oracle Streams continues to be supported but will not be enhanced to support any new database features in future releases. This means that Streams will continue to support exactly what it supports today (11.2) in future releases. It also means that Streams will not support any new database features, data types or capabilities. Customers who want to take advantage of the new data types and features will need to license Oracle GoldenGate.”

**Dell SharePlex**

SharePlex is a strategic part of Dell’s software portfolio, which was significantly expanded by the company’s acquisition of Quest Software in 2012. SharePlex is a product that was developed by Quest and continuously upgraded and enhanced over a number of years; SharePlex 8.0, the most recent version of the product line, was introduced in May 2013 and offers significant new and expanded features over previous versions.
Most significantly, expanding from its core Oracle-to-Oracle replication environment dramatically enhanced SharePlex’s capabilities. SharePlex 8.0 now supports near-term-time data integration to a Java Message Service (JMS) for facilitating data delivery to a wide variety of popular non-Oracle databases such as SQL Server, DB2, Sybase, Netezza and Teradata, as well as newer, unstructured databases that are gaining momentum in advanced analytics and big data environments, such as Hadoop and Greenplum.

SharePlex’s data integration feature replaces traditional data extraction techniques with its proprietary, industry-leading change data capture (CDC) methodology, making data integration far more efficient and cost effective by its ability to isolate only data changed in the data sources to be moved to the transform and load components of data aggregation.

The product also was enhanced with new support for high-availability requirements in Oracle environments, including TDE and HCC security features, to enable extended Oracle compression on Exadata platforms.

Another element of SharePlex 8.0 is its ability to ensure low latency in access to Oracle data, even in heavy transaction-centric environments. The maintenance of a real-time replica of the production database allows for operational reporting and other types of queries to be offloaded from the primary database to a secondary database, often running on lower-cost hardware.

Dell also said SharePlex 8.0 supports a wide range of Oracle application environments, such as Oracle E-business Suite, PeopleSoft, Cerner Millennium, SAP ERP, JD Edwards, Siebel and custom applications. SharePlex is designed to optimize operational efficiency during migrations, consolidations and restructuring by mitigating risk and downtime during those events. SharePlex allows DBAs to maintain business productivity in the form of continuous availability and tightly managed risk while setting up a new system, since users can continue to analyze, move, store and process data during migration.

Finally, SharePlex is integrated into Dell’s new ZeroIMPACT migration services toolkit for SAP solutions. The ZeroIMPACT migration suite for SAP is designed to help customers transition their legacy Oracle environments to a more modernized approach by keeping production systems operational during the migration. SharePlex’s real-time streaming process outside the database limits the impact on database performance, and uses log-based replication to endure data integrity and accuracy without affecting or changing the production application.

**Comparing Dell SharePlex With Oracle Data Replication Solutions**

Both Dell and Oracle offer IT organizations and database administrators strong choices when it comes to data replication solutions. The two companies’ respective solutions have been in the market for a number of years and have a history of consistent upgrades and enhancements to keep up with customers’ changing requirements.

There are, however, some important differences between Dell SharePlex 8.0 and the latest versions of Oracle’s relevant products — GoldenGate, Active Data Guard and Streams. Buyers should give strong consideration to these differences when making decisions on next-generation
data replication platforms, because those choices can have significant impact on nearly every aspect of an organization’s database activities for a number of years. Additionally, buyers should note the economic differences between the two approaches. For instance, Oracle GoldenGate typically requires additional charges for add-on tools or more functionality, whereas SharePlex provides full functionality in a single price tag. This results in significant advantages for SharePlex when it comes to both affordability and return on investment.

**Dell SharePlex 8.0 vs. Oracle GoldenGate**

Both products offer near-zero downtime during data replication processes, as well as real-time continuous availability and limited impact on performance due to overhead. But there are several reasons why SharePlex may be a better choice:

- One of the perceived challenges in doing a full GoldenGate implementation is the need for some IT organizations to purchase and integrate add-on components (such as Active Data Guard for operational reporting or Veridata for data comparison) in order to achieve full functionality. SharePlex, by comparison, has designed all data replication features into the core product, making it easier and faster to fully deploy the solution and get maximum productivity out of the solution.

- SharePlex has been designed to be set up using two quick steps: create the configuration file and then activate it. There is no need to set up procedures and turn on pieces manually, since it is designed for maximum automation to create and sync all queues and processes. Additional built-in features include reconcile, flush and compare-and-repair. These features help to ensure that the source and target are fully aligned and easily identify and rectify any anomalies without having to re-instantiate or go through the painful recovery process of the replica database. GoldenGate has the ability to do some of these features, but typically not in the baseline package. Upgrading to some of these features — where GoldenGate can even support them — requires an additional cost to the standard licensing fee. The lack of automated validation as part of the core function set is said to be a challenge for GoldenGate users because it requires either the purchase of Veridata as an additional cost or the dedication of DBA personnel to manually validate and resolve the accuracy of the replicated data on the secondary database.

- Scalability is another area of obvious importance to IT organizations, and Oracle often requires users to upgrade to Oracle Real Application Clusters (RAC) or Oracle Enterprise Edition to meet greater workloads. By comparison, SharePlex can help organizations easily and flexibly scale — at no additional cost — through such approaches as load balancing between multiple databases without overhead and costs of RAC. SharePlex also provides the option of running on low-cost hard. Eliminating the need to upgrade to Oracle RAC helps reduce I/O latency, since RAC cluster software has a tendency to increase wait time.

- In general, GoldenGate requires more dedicated DBA staff to handle such activities as deployment and maintenance, when compared with SharePlex. This is often due to SharePlex’s integrated and highly automated processes that house all data replication and data integration function under one software suite, as compared with GoldenGate which often is combined with other modules or other specialized Oracle programs for peak functionality.
• SharePlex offers in-flight data integrity to allow for immediate detection of a data discrepancy. All open transactions are checked immediately and handled in the appropriate manner. By comparison, GoldenGate transactions are “stuck” in trail files on the source database, which could result in lost transactions should a service interruption occur, or it could cause network bottlenecks during high activity transmissions.

• Finally, SharePlex’s design flexibility, simplicity of implementation and fully automated functionality allows customers to purchase and implement the product far less expensively than a similar GoldenGate implementation would cost. SharePlex is offered in two licensing models — one for Oracle Enterprise Edition and a lower price point for Standard Edition — which can save customers money if they don’t need all the functionality of the Enterprise Edition version or upgrade their database infrastructure, as is often the case with GoldenGate. In general, SharePlex’s costs are about one-half that of a similar GoldenGate implementation, considering initial licensing costs, purchasing of add-on modules to extend GoldenGate’s functionality and the need to dedicate more DBAs to implementation, training and data validation. SharePlex also is available in shorter-duration licenses for customers that only need a data replication solution for a short-term project, instead of being required to purchase the standard one-year license as is the case with GoldenGate. SharePlex’s new JMS capability introduced in version 8 provides an affordable solution to facilitate the integration of data with other database platforms. Users can apply the change data held in the JMS to the other database platform without having to pay for a SharePlex license on the target platform.

SharePlex 8.0 vs. Oracle Active Data Guard

Active Data Guard extends the disaster discovery capabilities of Data Guard, which is bundled into Oracle databases at no extra charge. Active Data Guard’s capabilities typically come with a significant additional price tag — usually around $10,000 per CPU. For this extra cost, users can have both their primary and secondary instances up and running, although the secondary systems are for read-only requirements, not for writing data. Queries are allowed on the secondary system using Active Data Guard, but not updates, deletes or additions. This makes Active Data Guard a viable solution for reporting applications, but not for full-scale data replication.

An important challenge in using Active Data Guard is that reporting applications that use Data Guard need to have all redo and archive log files sent to the target system, adding to cost and making it a challenge for network bandwidth. Since Active Data Guard is a read-only database, most reporting applications leave trail information inside the database requiring write access. As such, reporting applications have to be redesigned so they can run with Active Data Guard.

Active Data Guard users are confronted with the same Oracle restrictions as the core Data Guard solution, such as the requirement that the source and target systems must be running on the same system configuration, including platform, operating system and database version. As a result, customers have to enact a scheduled outage for system component upgrades, resulting in downtime. Additionally, the entire database must be replicated, users cannot pick and choose objects to be replicated and the target instance can’t be customized for reporting with Active Data Guard, resulting in sub-optimal reporting performance because adding indices is not allowed on the target instance. Finally, some customers may find that they need to consolidate data
from multiple source systems when building a reporting instance; users of Active Data Guard will discover that they won’t be able to do this due to the restriction of allowing only a one-to-one relationship.

SharePlex’s ability to replicate between different operating systems and databases help circumvent Oracle restrictions faced by Active Data Guard. SharePlex’s flexible design helps overcome the above-detailed restrictions and provide customers with maximum availability by avoiding planned outages for upgrades.

SharePlex also supports a wide variety of use cases with no additional costs in the baseline licensing fee, which is considerably lower than a similarly configured Oracle solution combining GoldenGate and Active Data Guard. SharePlex also is a superior solution for heterogeneous environments (mixed operating systems and/or different databases) since SharePlex runs on the server side and doesn’t query the database side.

Unlike Active Data Guard, SharePlex also does not require having the redo and archive files be sent to the target, allowing for immediate replication to secondary systems. This not only improves availability and responsiveness, but it reduces network traffic and eliminates performance bottlenecks. If there is an unplanned outage, data is automatically and immediately applied to the target system.

**SharePlex 8.0 vs. Oracle Streams**

As Oracle has stated, its Streams product is being de-emphasized from a development standpoint, with customers being encouraged to migrate to GoldenGate for data replication requirements. This means that Streams users essentially have three choices: Migrate to GoldenGate, stay with Streams for as long as Oracle maintains support, or transition to a new platform like SharePlex.

For Streams users, there are a number of important benefits in switching to SharePlex instead of remaining on Streams, which is clearly in “wind-down” mode over the next several years. These focus primarily upon such areas as optimizing full availability in production environments and maximizing continuous performance.

- Streams is embedded within the Oracle production database, which increases risk in the event of an unplanned service interruption. It also can utilize precious in-database resources and thus impact performance of production transactions. SharePlex, by contrast, operates externally from the production database, essentially isolating replicated data from the primary database in case a problem should arise.

- SharePlex achieves higher reliability than Streams because it provides fault-tolerant management of in-transit data with a file-based queuing architecture that automatically manages and recovers from interruptions. Streams, however, maintains queues within the database, which impacts database processing performance and potentially impacts availability of the production instance and the Oracle services shared by Streams with other transactions.

- Interoperability can be a challenge with Streams since it has dependencies upon specific Oracle versions, making upgradeability a big challenge. SharePlex has no dependencies on operating systems, database versions or Oracle versions.
• Out-of-sync issues are not easily addressed with Streams. Locating and fixing the anomalies is left up to manual processing of the DBA. SharePlex comes with a Compare-and-Repair feature that locates and fixes the anomalies, without having to re-instantiate or go through the painful recovery process of the replica database.

• Latency is a bigger problem with Streams because it begins moving data to the replica instance only after the data has been committed. In contrast, SharePlex extracts transaction data directly from redo logs before the transaction has been committed, then transmits the data directly to the target server. This cuts latency and facilitates near-real-time data on the target far more reliably and efficiently than Streams.

**Conclusion**

Data replication has become an extremely important function for IT organizations and database administrators to enable in this era of advanced analytics, business intelligence and big data. The need for optimized performance in transaction-intensive applications, as well as to ensure the highest levels of availability and data integrity, puts pressure on organizations to use data replication tools that combine sophisticated performance techniques and flexible design with low overhead and easier, less expensive deployment.

As the database vendor with the largest installed base, Oracle has necessarily developed a number of different data replication and data integration solutions over the past decade. Since some of these solutions were developed in-house and others were part of key acquisitions, Oracle has had to work hard to find ways to integrate the functionality of the different solutions in lockstep with continued technology advancements on their core Oracle relational database management solutions.

The biggest challenges Oracle faces with its current lineup of solutions are its lack of a single, comprehensive, integrated feature set for all relevant data replication/data integration requirements, and the significant cost and complexity in deploying its solutions, including its flagship GoldenGate product.

Dell SharePlex, by comparison, has been designed from the start to be a fully integrated, highly automated, flexible, and cost-efficient approach to data replication and near-real-time data integration for Oracle environments. Now, with SharePlex 8.0, that product is well positioned for the reality of today’s Oracle to Oracle and mixed database environments, offering sustained peak performance, high availability, full Oracle compatibility, facilitating support for non-Oracle environments and secure operation in the event of an unplanned service interruption.