An Oracle White Paper
Updated October 2009

Real-Time Access to Real-Time Information: Oracle GoldenGate

ORACLE

Executive Overview 1
Introduction 1
Understanding the Requirements for Transactional Data 3
Transactional Data, Databases, and Atomicity, Consistency, Isolation, and Durability 3
Transactional Data Management and Oracle GoldenGate 4
One Platform, Many Solutions 5
Continuous Availability and Disaster Tolerance 5
Real-Time Data Integration 10
Implement and Expand 14
Architecture Overview 16
Oracle GoldenGate Capture 16
Oracle GoldenGate Trail Files 17
Oracle GoldenGate Delivery 19
Oracle GoldenGate Manager 21
Additional Oracle GoldenGate Features 21
Flexible Topology Support and Bidirectional Configurations 21
Event-Based Processing 22
Robust Data Security 23
Conflict Detection and Resolution 23
Dynamic Rollback 23
Customized Oracle GoldenGate Processing 24
Initial Data Loads 24

Associated Products 25
Management Pack for Oracle GoldenGate 25
Oracle GoldenGate Veridata 27
Conclusion 28
Appendix: Supported Actions for Conflict Avoidance, Detection, and Resolution 29

Executive Overview

Business applications have become increasingly critical for transaction processing. As a result, end users must access, analyze, act on, integrate, store, and verify transactional data faster than ever—often in real time—and without system interruption or downtime. This demand for low-latency data is compounded by the exponential growth in transactional data volumes and an increasingly heterogeneous enterprise IT environment, creating a need for data integration and replication solutions that are easy to implement and have little to no impact on business-critical applications.

Oracle GoldenGate is used by major Fortune 500 companies and other large enterprises on more than 4,000 solution implementations worldwide to support mission-critical systems for data availability and integration. Written for business project owners, key stakeholders, and the entire IT organization, this white paper provides a detailed look at Oracle GoldenGate and its underlying technology architecture.

Introduction

Most business activities—whether with customers, partners, employees, or third parties—are conducted through transactions. Common transactions include making purchases, returning items, placing orders for shipping, managing patient records within hospital networks, paying bills, updating account and customer information, and making or changing reservations.

Today, across nearly all industries, an overwhelming majority of business transactions are executed digitally. As businesses expand their use of online applications, transactional data—and its importance—are growing at an exponential rate. To succeed in this real-time environment, companies must eliminate interruptions in data availability,

sluggish application performance, and stale data. There are four aspects to this challenge.

- **Availability.** Business-critical applications and underlying data must be accessible at or near 24/7/365 without service interruption or performance degradation.
- **Reduced latency.** Data must remain fresh. As it ages, data becomes less relevant and less valuable—day-old data is often insufficient in today's competitive landscape.
- **Heterogeneity and IT flexibility.** Integration and replication solutions must have the flexibility to be easily modified and distributed across diverse IT systems.
- **Transaction integrity.** Data completeness and accuracy must be ensured as it is moved between systems.

In short, companies need a platform that allows business applications to benefit from continuous access to real-time information, without compromising performance and data integrity or demanding significant resources to deploy and manage.

Oracle GoldenGate empowers organizations to capture, route, transform, and deliver transactional data between heterogeneous databases in real time with minimal overhead. Using unique real-time, log-based CDC and replication technology, Oracle provides continuous availability, disaster tolerance, and real-time data integration solutions that enable the management and movement of transactional data across the enterprise. Oracle GoldenGate is designed for low-impact implementation, operation, modification, and extension to support the evolving needs of enterprise information.

Understanding the Requirements for Transactional Data

Transactional data—data related to business transactions such as customer orders, payments, shipments, and service requests—and business applications that rely on processing, reporting, and analyzing that transactional data are critical for day-to-day operations. Businesses need to ensure that transactions, hence transactional data, are properly stored, accessed, and updated to avoid errors and delays that can impact customer satisfaction, operational costs, and revenues.

Transactional data must be fully understood in order to evaluate solutions for managing it effectively. Each transaction creates a number of data changes in the underlying database. For example, the online transfer of funds from a checking account to a savings account can be seen by the end user as a single transaction, but it requires multiple operations to database tables. In the context of databases, a *transaction* is a collection of related, logical operations on the data, namely inserts, updates, or delete operations.

The processing of a transaction within the database involves a number of steps to ensure its integrity, and it is not successfully completed from the application's perspective until all the underlying tasks are performed. Furthermore, any aborted or failed transactions must be rolled back so that the true history is preserved. In the example above, if the checking account was debited, but the savings account was never credited, the transaction should not be committed.

Once transactions are completed and committed at the database, that new data must be shared with other systems for a wide variety of reasons—such as integration into data warehouses to support reporting needs, or for making copies of data via replication to backup systems. As those transactions are moved from the original source database to other target systems, it is crucial that transactional integrity is maintained. Otherwise, the true business activity and history can be lost or misrepresented at the target system.

Transactional Data, Databases, and Atomicity, Consistency, Isolation, and Durability

To preserve the integrity of the database, there are four key features for processing transactions that occur within the database: Atomicity, Consistency, Isolation, and Durability (ACID). These properties guarantee that transactions are processed reliably within the database.

- **Atomicity.** Either all or none of the tasks of a transaction are performed.
- **Consistency.** A transaction can't break the rules or integrity constraints of the database; if it attempts to do so, it is aborted.
- **Isolation.** The operations in a transaction must appear isolated from all other operations in the application. No operation outside the transaction can ever see the data while it is in an intermediate state.

- **Durability.** The transaction will persist and not be undone after it is committed. A transaction can only be deemed committed after it is safely written into a database transaction log.

Oracle GoldenGate only captures and moves committed database transactions that meet the ACID guidelines. The application carefully ensures the integrity of data as it is moved from the source database and applied to any number of target databases or messaging systems.

Transactional Data Management and Oracle GoldenGate

Oracle GoldenGate offers a real-time, log-based change data capture (CDC) and replication software platform to meet the needs of today's transaction-driven applications. The software provides capture, routing, transformation, and delivery of transactional data across heterogeneous environments in real time. Using this technology, customers can achieve continuous availability for critical systems and real-time data integration for fast, easy access to current and accurate data. See Table 1.

TABLE 1. ORACLE GOLDENGATE KEY FEATURES AND DIFFERENTIATORS

FEATURE	DETAIL
Real-time data feeds	Provides continuous capture and delivery of data from sources to targets with end-to-end subsecond latency. Operates at high performance with low overhead even at high volumes.
Heterogeneity	Captures and delivers data between a variety of relational, open systems/open source, and legacy databases on all major platforms.
Transactional integrity	Maintains the reliability and accuracy of transactional data as it is moved between systems by enforcing ACID properties and referential integrity.

Since 1995, Oracle GoldenGate has been time tested and stress tested by transaction-intensive applications across countless industries, including banking, financial services, healthcare, telecommunications, cable, media, retail, e-business, travel, hospitality, energy, utilities, and the public sector. More than 400 customers representing more than 4,000 solution deployments worldwide have put their trust in Oracle to enable the movement and management of their critical, rapidly changing transactional data.



Figure 1. Oracle GoldenGate provides real-time access to real-time information through a comprehensive view of operational systems.

One Platform, Many Solutions

In addition to the Oracle GoldenGate core platform for real-time data movement, Oracle provides the Management Pack for Oracle GoldenGate—a visual management and monitoring solution for Oracle GoldenGate deployments—as well as Oracle GoldenGate Veridata, which allows high-speed, high-volume comparison between two in-use databases.

Together, these products enable companies to easily and successfully implement a variety of solutions for improving the availability, accessibility, performance, and integration of critical data across the enterprise. There are two primary business needs addressed by these solutions:

- Continuous availability and disaster tolerance
- Real-time data integration

Continuous Availability and Disaster Tolerance

Businesses are facing increased pressure to meet end-users' expectations for 24/7 system availability and to surpass competitors' service levels. This means the IT organization must implement solutions that keep key business applications operating with an acceptable level of continuity, even under a range of potentially damaging or interrupting conditions.

Typically, companies might first consider traditional disaster recovery products and practices. But for databases that support applications requiring high-availability targets, such as fast recovery time objectives and strict recovery point objectives, traditional disaster recovery technologies alone are not likely to be sufficient, because with these solutions database recovery is typically required, which extends the failover time.

Continuous availability and disaster tolerance of critical data systems is more than "after-the-fact" data recovery. Oracle GoldenGate allows customers to avoid experiencing downtime or the data loss effects of a disaster. In addition, it improves system availability when faced with not only unplanned outages but also planned outages and performance-related issues that impact the end-users' perception of availability—regardless whether the system is down or available. Oracle GoldenGate can be implemented to support each of these states of availability. See Figure 2.

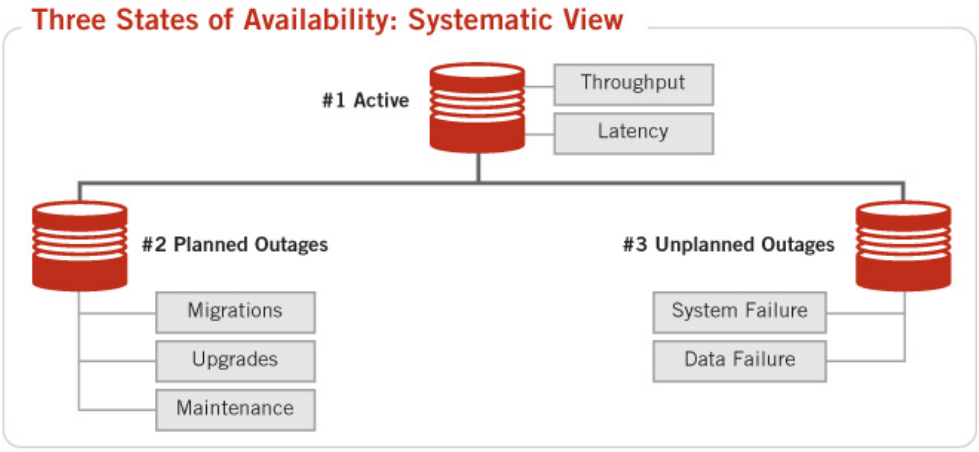


Figure 2. Continuous availability solutions must address planned and unplanned outages and support access for critical applications.

State #1: Active but with Performance Challenges

If the application or database is up but experiencing performance degradation that is noticeably affecting throughput and response times, one of the most effective approaches is to offload or distribute the processing activity.

Deploying Oracle GoldenGate for Data Distribution and Synchronization

Oracle GoldenGate's capabilities for real-time, heterogeneous, multidirectional data movement enable companies to reliably deploy distributed systems to drive global operations and improve operational efficiency. Oracle GoldenGate distributes data in real time across Oracle and non-Oracle environments and enables access to consistent data across geographically dispersed data centers. In addition to handling large and growing data volumes, this configuration allows continuous availability for business operations if one data center is not available.

Deploying Oracle GoldenGate for Active-Active Databases

Deploying Oracle GoldenGate for active-active databases allows for the highest availability and performance on critical systems, because it enables load balancing by continuously and multidirectionally moving data between two or more active and heterogeneous databases, as shown in Figure 3. The solution also offers strong capabilities for conflict detection and resolution that can be required in multimaster configurations.

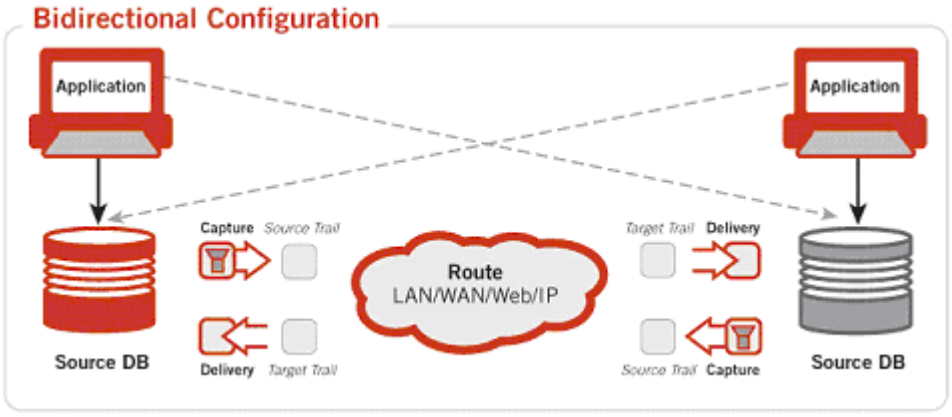


Figure 3. Oracle GoldenGate enables multidirectional data replication to support a multimaster, active-active application configuration.

Deploying Oracle GoldenGate for Query Offloading

High volume, read-only queries that run on online transaction processing (OLTP) systems—such as product searches by customers—can cause significant overhead on the production environment, leading to high CPU utilization and performance degradation. In many cases, to address the issue, organizations upgrade their existing expensive legacy systems with additional investments to be able to provide acceptable performance to end users. To improve transaction processing performance and reduce costs, Oracle GoldenGate creates a real-time replica of the production environment on lower-cost platforms and allows read-only queries to be pointed to this environment. With this architecture, the production environment is freed up from supporting expensive read-only activity and can support transaction processing activity with better performance and a longer lifetime.

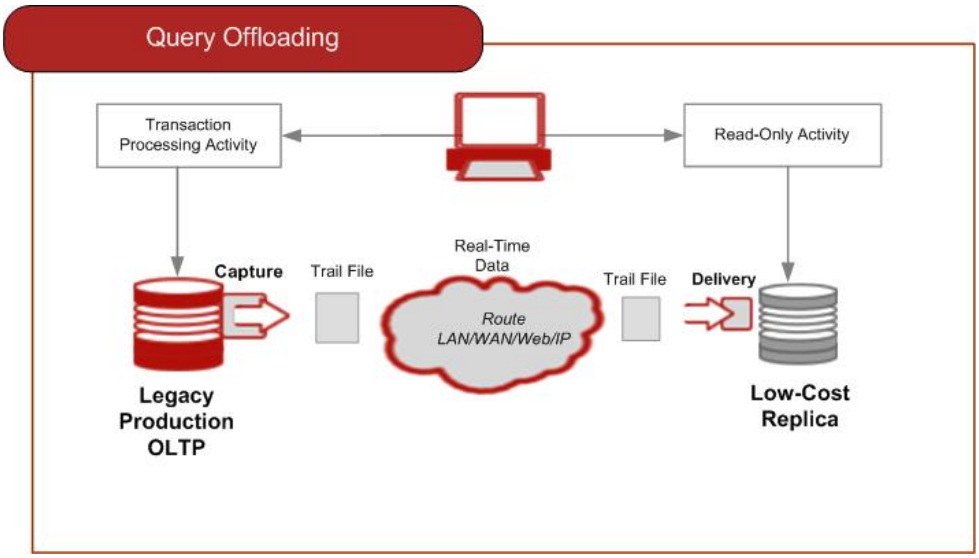


Figure 4. With Oracle GoldenGate, you can offload queries from production systems to reduce impact and optimize reporting performance.

State #2: Planned Outages

IT groups occasionally need to schedule planned outages to modify hardware or database software, upgrade applications or databases, apply software patches, or migrate to a different computing architecture. Oracle GoldenGate eliminates the downtime traditionally associated with these planned outages.

Deploying Oracle GoldenGate for Zero-Downtime Operations

Through real-time, bidirectional data movement and synchronization between old and new systems, Oracle GoldenGate enables automated switchover from the existing system to the new system, when the new system is ready—without ever denying access to the application. The solution can support heterogeneous environments for cross-platform upgrades and database migrations from non-Oracle databases to Oracle databases. In addition, the solution provides robust failback contingencies by keeping the old and new environments in sync in real time. By simultaneously using Oracle GoldenGate Veridata, users can identify and report data discrepancies across systems before switchover, without impacting the production environment.

Oracle GoldenGate offers a unique solution for upgrades from Oracle 8*i*, Oracle 9*i*, and Oracle 10*g* to Oracle 11*g*, where the database downtime is reduced to mere seconds or minutes during the upgrade. In addition, the same Oracle GoldenGate product can be used to upgrade to Oracle's Siebel Customer Relationship Management (CRM) 8.x from Siebel CRM 6 or Siebel CRM 7 without any downtime and with options to perform phased user migrations. Through an additional product—Oracle GoldenGate Application Adapters—Oracle provides prebuilt upgrade and downgrade transformation logic for Oracle's Siebel CRM solutions.

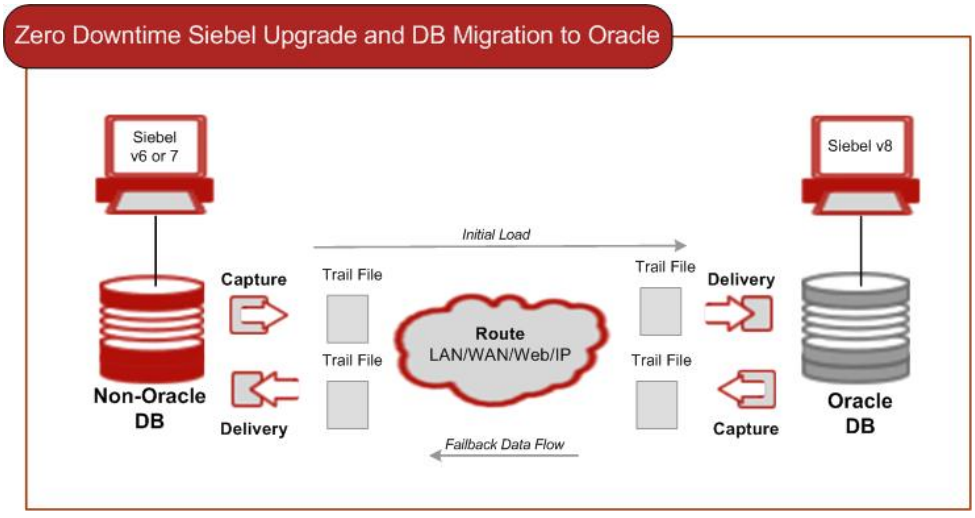


Figure 5. Oracle GoldenGate enables organizations to upgrade and migrate mission-critical applications with no downtime.

State #3: Unplanned Outages

Unplanned outages may be caused by system or data failures at the site or system level. For mission-critical systems, businesses must confidently fail over to a backup as quickly as possible, and easily revert back to normal operating conditions once the primary system is back online.

Deploying Oracle GoldenGate for Disaster Recovery and Data Protection

When configured for disaster recovery and data protection, Oracle GoldenGate provides a continuous availability software implementation that significantly improves recovery time for business-critical systems. Oracle GoldenGate's disaster recovery and data protection configuration delivers up-to-the-second data to the backup system and enables immediate and automatic switchover to the new system if an outage occurs. It also immediately initiates real-time data capture from the standby database to update the primary system, once it is online, with any new data processed by the standby system. In addition to automating switchover in case of system outages, Oracle GoldenGate has capabilities to detect issues with low system performance and can alert responsible teams.

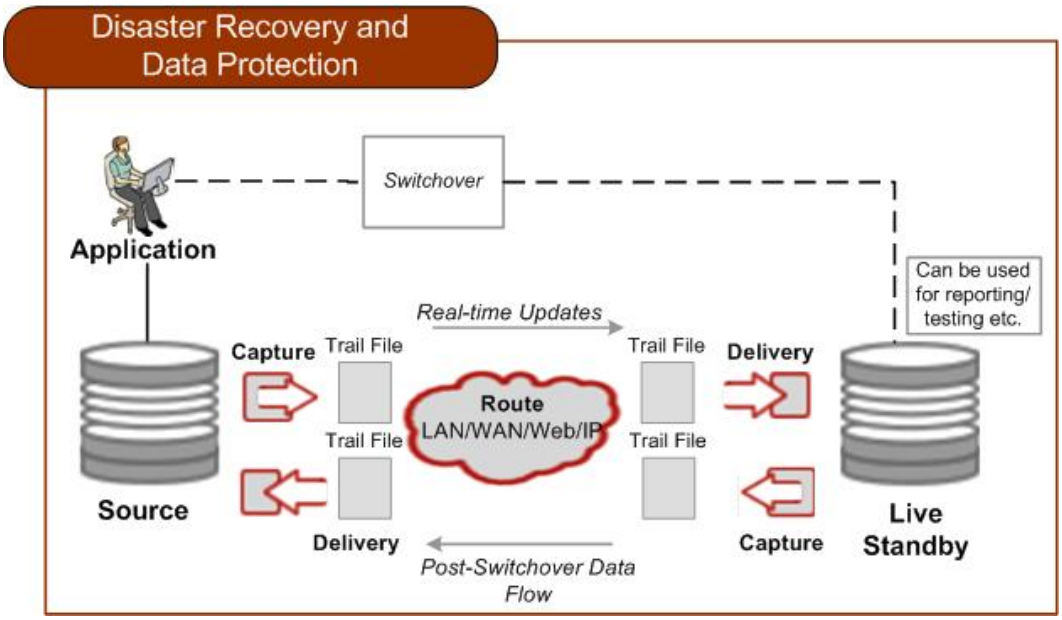


Figure 6. Oracle GoldenGate provides a robust solution for failover in the event of an unplanned outage.

Real-Time Data Integration

Another key market trend is the low-latency integration of data from various business applications. As data latency increases, the information becomes less consistent with reality and its organizational value diminishes. Some business operations require business insights that use the most current data to enable employees to take action with completely reliable and accurate information. To improve operational efficiency and effectiveness, companies need to rely on business intelligence (BI) that uses timely operational information as well as historical context.

For data integration needs, companies have generally relied upon moving bulk data periodically to their analytical systems. The overhead of running bulk data extract processes impacts OLTP systems' performance significantly and typically requires downtime; therefore, this batch data extract is processed typically in the night when less transaction volumes are expected. For many mission-critical systems that need to support users at or near 24/7 uptime, these batch windows can be a major issue, because they have to limit the time the systems are unavailable. As data volumes increase, IT teams might not be able to complete the extraction of data they need to move within the allotted time window. The same problem occurs if they need to restart the batch process for any reason; the allotted time window might not be adequate.

To overcome this dilemma, companies should avoid focusing on custom scripts or piecemeal bulk data handling solutions. Instead, they should adopt more-comprehensive data integration approaches that combine both bulk data movement and transformation with real-time data integration to enable highly available mission-critical systems and real-time BI.

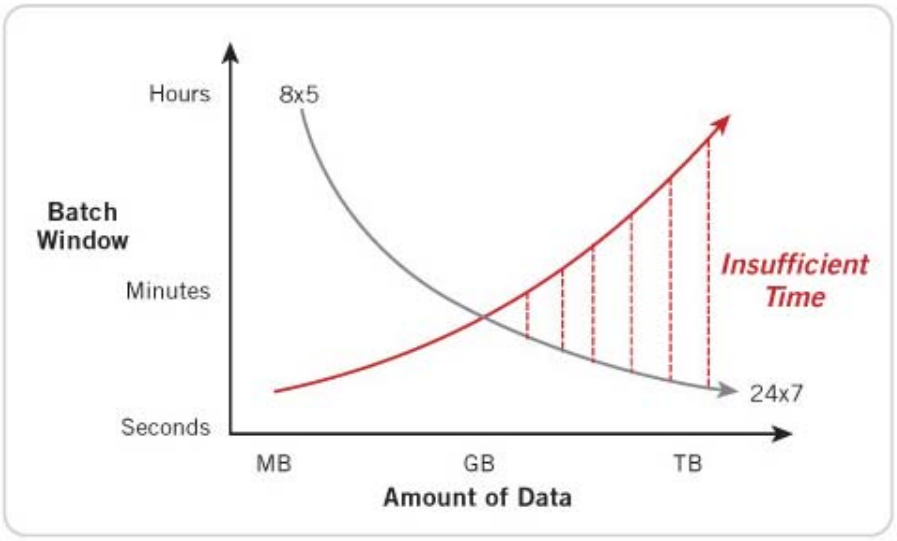


Figure 7. Transactional data volume has exceeded the point in which batch processing of data can be completed within the desired time window. Real-time changed data movement eliminates these risks and operates without business interruption.

Augmenting extract, transform, and load (ETL) systems with a real-time, log-based change data capture (CDC) solution enables IT teams to meet the requirements of mission-critical systems. Through a log-based CDC approach, organizations can source data from OLTP systems without impacting performance and feed the ETL system with a continuous stream of real-time data. This method not only decreases data latency for BI systems, but it also eliminates the reliance on batch-processing windows, while allowing continuous operations for business-critical systems.

Design for Real Time

Oracle GoldenGate enables the continuous, real-time capture, routing, transformation, and delivery of transactional data across heterogeneous environments. As new or updated data is committed at the source system, it is continuously captured and applied to one or more target systems with subsecond latency. Only changed data is moved, so a lower burden is placed on the infrastructure.

Oracle GoldenGate offers several key advantages:

- Continuous, real-time data movement with subsecond latency
- Negligible impact and overhead on source and target systems
- No requirement for a middle-tier server
- Integration with Oracle Data Integrator Enterprise Edition for complex transformations
- No downtime for batch processing
- Complete data recoverability in case of outages or failures
- Read-consistent data movement while maintaining referential integrity
- Ability to apply transformations and mappings within the target database
- Ability to use the same product in different topologies for different solutions such as continuous availability and zero-downtime upgrades and migrations

Customers worldwide are increasingly implementing Oracle GoldenGate for real-time data integration. Among the popular solutions from Oracle in this area are

- Real-time BI
- Operational reporting
- Operational data integration

Deploying Oracle GoldenGate for Real-Time Business Intelligence

With its real-time, log-based CDC and delivery capabilities, Oracle GoldenGate complements Oracle Data Integrator Enterprise Edition. By combining these features with Oracle Data Integrator's data quality and transformation capabilities and its ELT architecture, customers can

achieve a low-impact, cost-effective, real-time data integration solution for their enterprise data warehouse implementations. In real-time BI and data warehousing use cases, Oracle GoldenGate brings changed data to the staging tables of the data warehouse in real time, at which point, Oracle Data Integrat**or** performs high-speed optimized transformations.

Oracle GoldenGate allows companies to implement real-time BI solutions, empowering a much larger number of front-line users across the enterprise to make better-informed operational decisions and create a sustainable competitive advantage.

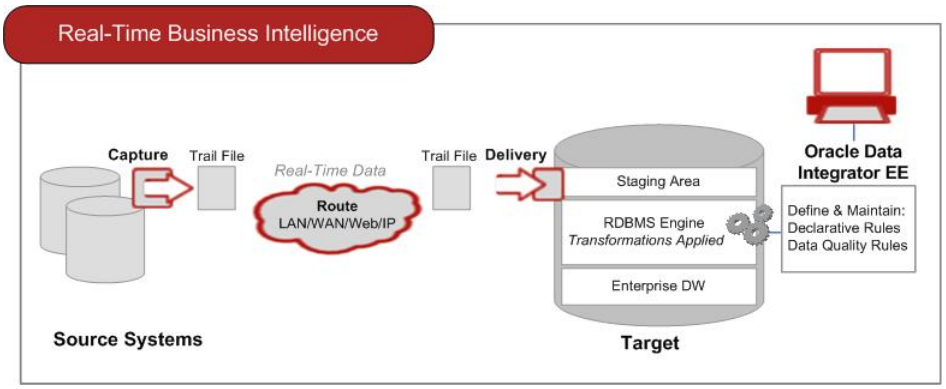


Figure 8. Oracle GoldenGate can rapidly populate data warehouses with real-time data without impacting system performance.

Deploying Oracle GoldenGate for Operational Reporting

Many operational reporting activities ideally seek to use the latest data available, and that often means running reports against production databases, which degrades performance. With Oracle GoldenGate, a secondary, cost-effective system can be deployed to serve the purposes of real-time, operational reporting, freeing up the resources of the critical source systems.

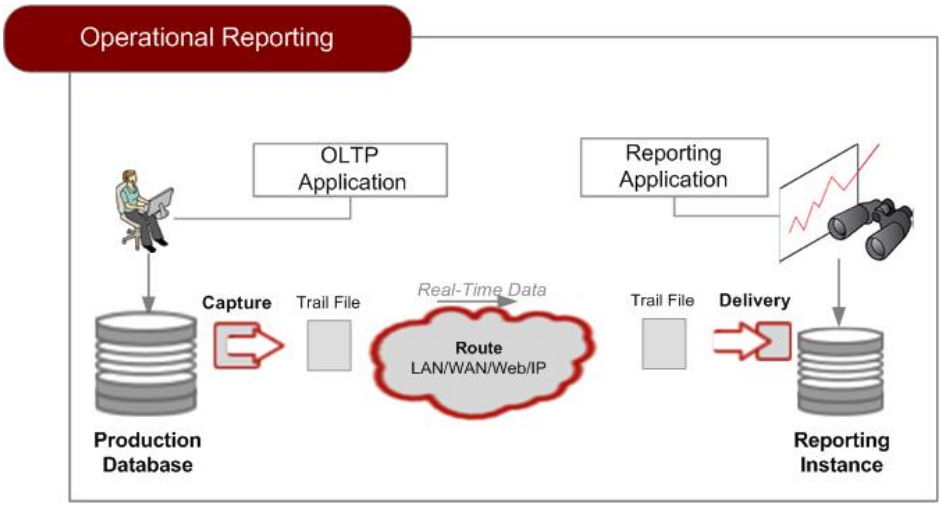


Figure 9. By continuously moving source data to a reporting database, users can access real-time information with no production impact.

Oracle GoldenGate can complement other ETL products with real-time CDC capabilities to help customers realize better value from their existing investments. The solution reads changed data from database transaction logs, eliminates batch windows, and causes negligible impact on source systems while maintaining transaction integrity. Oracle GoldenGate persists changed data outside of the databases, which means the solution enables complete recoverability from failures in the case of an outage.

Oracle GoldenGate offers three different methods to augment ETL solutions:

- **Staging tables.** Oracle GoldenGate moves real-time transactional data from multiple production databases into staging tables within the data warehouse, at which point the ETL system pulls the data from the staging tables and loads the user tables after making any required transformations. This method is the recommended method for integration with Oracle Data Integrator Enterprise Edition. With Oracle Data Integrator's ELT architecture, there is no need for a middle-tier transformation server. This method optimizes performance and scalability and lowers overall solution costs.
- **Flat files.** Oracle GoldenGate writes the changed data into a flat file that is stored on a server of the customer's choosing. The ETL system reads from the flat file, performs the necessary transformations, and loads the user tables of the data warehouse. Oracle GoldenGate has the ability to provide the data in a variety of formats to create the optimal feeding mechanism for existing ETL technology.
- **Messaging systems.** Oracle GoldenGate can publish changed data in real time from source OLTP systems to messaging systems. From there, ETL systems can receive the changed data in XML format in real time.

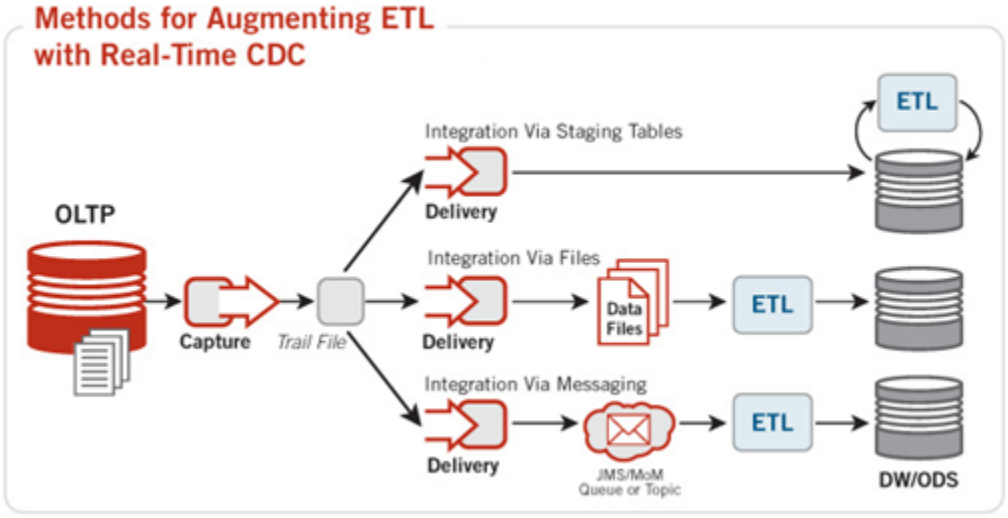


Figure 10. Oracle GoldenGate can leverage—and enhance—existing investments to optimize real-time enterprise insight.

Deploying Oracle GoldenGate for Operational Data Integration

Oracle GoldenGate offers a solution for distributing timely data between transaction processing systems. Configuring Oracle GoldenGate for operational data integration provides a nonintrusive, fast-to-deploy method for moving data in real time, with minimal impact, using an architecture that is substantially easier to customize and maintain as application requirements change. By deploying Oracle GoldenGate Application Adapters for Java Message Service (JMS), in conjunction with Oracle GoldenGate, organizations can publish changed data to JMS-based systems from heterogeneous transaction processing or analytical systems to support service-oriented architecture (SOA) and enable event-driven architectures (EDAs). This enhanced version of the product can integrate with Oracle SOA Suite with its ability to publish data to Oracle WebLogic Server, IBM WebSphere MQ, TIBCO Enterprise Message Service, Red Hat JBoss, Apache ActiveMQ, and others.

Implement and Expand

Fortune 1000 companies and other large business entities leverage the Oracle GoldenGate product and configuration solutions to manage critical initiatives in a range of business areas including continuous availability, disaster tolerance, and real-time data integration. Oracle's technology is extremely versatile in its ability to meet a wide variety of today's business requirements. Leveraging the application's modular architecture and design, these organizations quickly scale their investment in Oracle GoldenGate across the enterprise as return on investment is achieved.

Table 2 highlights how Oracle GoldenGate deployments have progressed for three such customers in their pursuit of a robust, real-time data infrastructure.

TABLE 2. CUSTOMER SUCCESS: EXPANDING THE ORACLE GOLDENGATE FOOTPRINT ACROSS THE ENTERPRISE

PHASE	OVERSTOCK.COM	BANK OF AMERICA	MONTEFIORE MEDICAL CENTER
1	**Zero-downtime operations.** Minimized downtime during an enterprisewide, cross-platform migration of its OLTP systems from Oracle 9*i* to Oracle 10*g.*	**Data distribution.** Data generated by its 18,000+ ATM network is backed up in real time to multiple data centers for immediate failover.	**Data integration.** New and changed data from its clinical information system application is immediately fed to the data warehouse to enhance reporting and business intelligence.
2	**Data integration for real-time business intelligence.** Real-time data feeds from OLTP systems into the enterprise data warehouse to enable a "single view of the customer" and improve reporting capabilities.	**Zero-downtime operations.** ATM business processes are kept online during large upgrade of production servers.	**Continuous availability.** Patient data systems are backed up to a live standby system for disaster recovery support.
3			**Data verification.** Oracle GoldenGate Veridata runs frequent discrepancy checks on data shared between production and secondary databases.
	"With a clear, accurate understanding of the customer, we can give them the right information at the right time, which directly impacts satisfaction levels and essentially revenue and growth."	"We evaluated other products; however, Oracle GoldenGate offered us benefits that would also enable us to meet our longer-term goals."	"Accessing the data in real time using Oracle GoldenGate, we can immediately see if we're profitable and if our business processes are working."
	—Sam Peterson CIO Overstock.com	—Michele Schwappach VP of ATM/Debit Applications Bank of America	—Carl Baylis Assistant Director EHIT/Montefiore Medical Center

Architecture Overview

The Oracle GoldenGate architecture consists of decoupled modules that can be combined across the enterprise to provide maximum flexibility, availability, and performance. This architecture facilitates the movement of transactional data in four simple, yet powerful steps.

- **Capture.** Oracle GoldenGate captures changed data operations committed in the database transaction logs in a nonintrusive, high-performance, low-overhead implementation.
- **Route.** Oracle GoldenGate can use a variety of transport protocols, and it can compress and encrypt changed data prior to routing. Transactional data can be delivered via Open Database Connectivity–compliant databases or through a specialized adapter to a JMS message queue or topic.
- **Transform.** At any point prior to applying the data to the target system, Oracle GoldenGate can be used to execute a number of built-in functions, such as filtering and transformations.
- **Apply.** Oracle GoldenGate applies the changed transactional data to one or more database targets with only subsecond latency, preserving transactional integrity.

Oracle GoldenGate consists of four distinct modules and components:

- Oracle GoldenGate Capture
- Oracle GoldenGate Trail Files
- Oracle GoldenGate Delivery
- Oracle GoldenGate Manager

Oracle GoldenGate Capture

The Capture module grabs committed transactions resulting from insert, update, and delete operations executed against a database, and routes them for distribution.

High-Speed, Low-Impact Data Capture

The Capture module does not require any changes to be made to the source database or the application it supports. To maintain optimal performance, the Capture module employs a range of CDC techniques against the source database. For instance, in databases that include transaction logs, such as Oracle databases; Microsoft SQL Server; IBM DB2 (LUW) and z/OS; Sybase; Ingres; and those running on HP NonStop/Enscribe, SQL/MP, and SQL/MX, changes are captured through direct access to native database transaction logs (redo logs, if applicable).

Transaction logs contain all changes made to the database and are automatically maintained by the database application independently of Oracle GoldenGate. Consequently, no additional tables are required to run the Capture module, and overhead is greatly reduced as compared with

trigger-based capture techniques. Customers report only single-digit percentage overhead when running the Capture module on the source database. The Capture module can automatically adjust its transaction memory based on the size and number of the transactions it is capturing, which optimizes memory usage, allowing even lower overhead on the source systems.

Table, Row, and Column Selectivity

When not all changed data from the source needs to be replicated to the target system—such as for real-time reporting purposes—the Capture module allows users to filter tables and rows based on user-defined criteria and ignores the entries in the transaction log that don't meet the end-user's needs. Users can optionally select and apply transformation rules to specific columns via built-in Oracle GoldenGate functions, user-supplied code, stored procedures, or Oracle Data Integrator Enterprise Edition.

Efficient Network Use and Large Data Volumes

The Capture module can route transactions over WANs and LANs as well as the internet, and it can reduce network bandwidth requirements in a number of ways. Typically, the amount of data transmitted is only a fraction of the data that is generated by the database and stored in transaction logs. Because only committed transactions are propagated, intermediate activities and rolled-back operations are not transferred. Traffic is optimized by bundling individual records into larger, more-efficient packets and avoiding record-at-a-time bottlenecks. Several levels of data compression are available to further reduce the amount of network bandwidth required for transmission. Depending on datatypes, data compression can reduce byte transfer by 75 percent or more.

For scenarios requiring very large changed data volumes, users can deploy multiple Capture modules to minimize the lag between source and target systems.

Checkpoints for Reliable Data Delivery

Oracle GoldenGate creates a checkpoint at the last changed transaction whenever a commit boundary is encountered. This enables the delivery of all committed records to the target, even in the event of a restart or cluster failover. Checkpoints store the current position as processed by both the Capture and Delivery modules. Following a network or system outage, Oracle GoldenGate restarts from the last good checkpoint.

Oracle GoldenGate Trail Files

Trail Files—Oracle GoldenGate's unique queuing mechanism—contain the most recent changed data in a transportable, platform-independent format called the Oracle GoldenGate Universal Data Format, and can be converted to XML and other popular formats for consumption by different applications. The Capture module can create unique as well as overlapping sets of data

in each Trail File. Based on the requirements of the implementation, users can store Trail Files on the target system, the source system, or both. Trail Files can be delivered to alternative queue types and application interfaces.

Flexible, Decoupled Architecture

A decoupled architecture addresses numerous problems inherent in tightly coupled alternatives. Process-to-process coupling creates a dependency between data capture and delivery. For example, if delivery is slower than capture, capture activities must be held up. In the event of an unplanned outage, decoupling ensures that the nonimpacted system continues to operate.

Tightly coupled or process-to-process implementations can impose scalability challenges. A great deal of interprocess checkpointing needs to occur to ensure no data is lost, thereby creating many more messages and still more overhead. Network outages lasting more than a few minutes can also cause excessive resource consumption, because outstanding transactions need to be queued in memory and eventually swapped to disk. Neither the physical nor the virtual memory activities are persistent; therefore if the process fails, data inconsistencies—or even loss—ensues.

By staging data in Trail Files, Oracle GoldenGate decouples the datasource and target for heterogeneous support. Unlike architectures that implement a tight process-to-process coupling, this decoupled architecture allows each module to perform its tasks independently.

Oracle GoldenGate also provides flexibility in the choice of hardware, operating system, and databases for sources and targets. For maximum flexibility and ease of use, customers can use different versions of Capture, Delivery, and Trail Files in the same implementation.

Data Pumps

Depending on the configuration and environment, it might be preferable to create Trail Files on the source system and use a separate Oracle GoldenGate feature, called a Data Pump, to continuously push—or "pump"—the Trail Files from the source system to the target system(s). This configuration enhances the fault tolerance and reliability of the overall Oracle GoldenGate environment. In the event of a network failure (between the source and the target systems), Oracle GoldenGate can continue to capture transactions because the data can be queued up locally in the Trail Files on the source, enhancing the recoverability in case of database failures.

The Data Pump feature is strongly recommended if data needs to be distributed to multiple targets from the same source (one to many). Whereas the Capture module can focus solely on capturing transactions, individual Data Pumps can be set up to distribute the data to those targets, increasing the efficiency of the overall environment. Fault tolerance is also greatly increased in such a configuration because any failure associated with one target has no impact on the source capture or delivery to other targets—transactions will continue to be captured, routed, and applied to the other targets even when one of them is down. Data Pumps can also be used to route data through an intermediate system, even if that system doesn't have a database installed.

Archival and Audit Capabilities

Trail Files can create an archive of purged information from the source database by transforming delete and update records into inserts in a different location. For auditing and compliance purposes, Oracle GoldenGate can also maintain a separate history table to track each update to individual records as they change.

Oracle GoldenGate Delivery

The Delivery module takes any changed transactional data that has been placed in a Trail File and immediately applies it to the target database. Through the use of Oracle GoldenGate Application Adapters, Oracle GoldenGate also has the capability to publish changed data to a messaging system in XML or other formats, as well as provide data in flat files for third-party products, such as an ETL system.

Data Integrity and Transaction Consistency

The Delivery module applies each database change in the same order as it was committed in the source database to provide data and referential integrity. In addition, it applies changes within the same transaction context as they were on the source system for consistency on the target.

Column Mapping and Transformation

As with Capture, users can configure the Delivery module via user-defined criteria to not only specify target tables but also individual rows and columns. By default, the Delivery module populates any target table column with data from a source table column if the two columns share the same name, and this is also true of like-named tables. However, you can easily configure Oracle GoldenGate to move data from a single table into multiple target tables or vice versa. This can be used to normalize or denormalize data in a data warehouse or OLTP environment.

Users can also define explicit mapping and transformation rules, ranging from simple column assignments to more-complex transformations for which Oracle GoldenGate provides a suite of date, math, string, and utility functions. The module also supports the use of stored database procedures and functions and enables implicit mapping and explicit rules to be combined. If additional transformations, data quality, aggregation, and other functionality are required, Oracle GoldenGate integrates with Oracle Data Integrator Enterprise Edition to support end-to-end data integration.

Optimized High-Speed, High-Volume Data Delivery

The Delivery module provides a variety of techniques to optimize the posting of changed transactions to the target database. Oracle GoldenGate's posting processes, where possible, run local to the target database, maximizing throughput by avoiding network limitations. The Delivery module also minimizes disk I/O while preserving original transaction properties. In

addition, where possible, updates are executed via native database interfaces rather than through middleware, and internal caches are used to ensure fast execution of repetitive statements.

Multiple Delivery modules can be deployed to minimize lag time in the event of high data volumes during peak processing times or seasonality. This capture-route-transform-apply process runs continuously, so that the most recent transactions committed at the source are immediately moved and delivered to the target.

Deferred Apply

For maximum flexibility, the Delivery module can apply data immediately or at a deferred time chosen by the user, without losing transactional integrity. This allows an additional layer of data protection when needed and keeps the secondary system at a consistent state behind the primary system. In this configuration, Oracle GoldenGate routes the changed data to the Trail File on the target server but does not deliver it to the target database until all captured changes have been delivered to the target Trail File. All changed data in the target Trail File can then be immediately applied to the target database, to bring it to a consistent state relative to the source, whenever the user chooses.

Java Message Service Delivery

In addition to databases, Oracle GoldenGate can also publish changed data to JMS queues and topics by using Oracle GoldenGate Application Adapters. After capturing from source database transaction logs, Oracle GoldenGate converts captured records into JMS text and map messages (name-value pairs), and it formats text in any way, including XML. Changed data can be published as transactions with preserved integrity or as individual database operations such as inserts or deletes. As mentioned earlier, this allows Oracle GoldenGate to provide improved support for SOA and enable event-driven architectures

Flat File Delivery

Using Oracle GoldenGate Application Adapters, Oracle GoldenGate can publish changed data in the form of flat files to integrate with third-party data management products such as ETL. For those ETL systems that perform faster reading files than scanning staging tables, this method minimizes storage resources and system maintenance. It also enables the user to decrease the data latency by configuring the frequency of microbatches. Oracle GoldenGate has the ability to provide the data in a variety of formats, including delimited text files and binary files, to create the optimal feeding mechanism.

Oracle GoldenGate Manager

To give users control over Oracle GoldenGate processes, the Manager module provides a command-line interface to perform a variety of administrative, housekeeping, and reporting activities, including

- Setting parameters to configure and fine-tune Oracle GoldenGate processes
- Starting, stopping, and monitoring the Capture and Delivery modules
- Critical, informational event, and threshold reporting
- Resource management
- Trail File management

The Manager module executes requests on demand as well as unattended. For example, it can be used to restart Oracle GoldenGate components as well as monitor latency. The module also automatically recycles Trail File data when no longer needed, providing insurance against inadvertent disk-full conditions and offering an alternative to error-prone manual housekeeping procedures.

For enhanced management of Oracle GoldenGate processes and solutions, customers should consider adding the Management Pack for Oracle GoldenGate.

Additional Oracle GoldenGate Features

In addition to its core modules and functionality, Oracle GoldenGate provides the following features and benefits to support a comprehensive real-time information environment.

Flexible Topology Support and Bidirectional Configurations

As a result of its decoupled modular design, Oracle GoldenGate easily supports a wide variety of topologies. These include one-to-one, one-to-many, many-to-one, and many-to-many—for both unidirectional and bidirectional configurations.

For unlimited scalability, cascading topologies can be created to eliminate any potential bottlenecks. By staging specific sets of database changes on the source or target system, different requirements can be met through a single pass on the datasource. Each set of staged data can contain unique or overlapping sets of data.

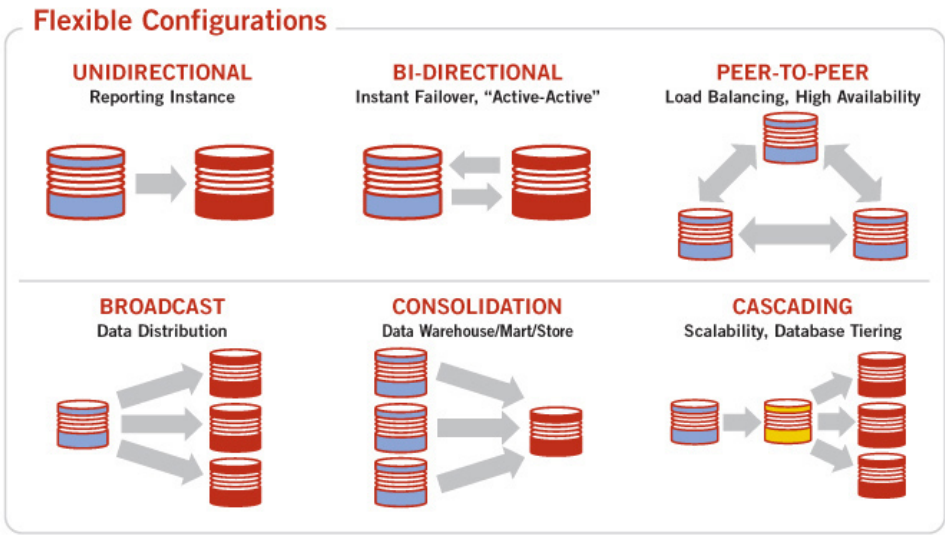


Figure 11. Oracle GoldenGate supports numerous data propagation solutions to support real-time visibility across the enterprise.

As mentioned in the Delivery section, Oracle GoldenGate can also integrate with JMS-based messaging systems for increased flexibility in distributing real-time data. This capability, along with the flat file delivery feature, allows Oracle GoldenGate to provide different architectures to augment existing investments.

Event-Based Processing

Using an event marker infrastructure, Oracle GoldenGate processes can act upon a specific database operation captured and stored in the Trail Files. For example, when a transaction that meets user-defined criteria is captured, Oracle GoldenGate can automatically start or stop Capture or Delivery processes, perform certain transformations, or provide statistics. This capability significantly improves Oracle GoldenGate's flexibility and manageability in a variety of use cases including

- Automatic switchover to the secondary system during planned outages
- Better monitoring over source systems' performance and automated switchover to the standby system in case of an outage with the primary system
- Automatic switchover from initial load to changed data movement
- Automatic synchronization of any type of batch processing taking place on both the source and target databases for database consistency
- Automatic stoppage of the Delivery module to allow end-of-day reporting
- Finding, tracking, and reporting on transactions that are of interest including the ones that do not have primary keys or transaction record numbers

Robust Data Security

Oracle GoldenGate provides 128-bit encryption to offer secure, confidential data transmissions. The encryption algorithm can accommodate a variable-length key, from 32 bits to 128 bits, making it suitable for both domestic and international use.

Conflict Detection and Resolution

Conflict detection and resolution are key prerequisites of active-active or multimaster database configurations. When both systems are processing transactions and the activity is shared across multiple systems, detecting and addressing data conflicts becomes an essential requirement.

Oracle GoldenGate provides a wide variety of options for avoiding, detecting, and resolving conflicts. These options can be implemented globally, on an object-by-object basis, based on data values and filters, or through event-driven criteria including database error messages. See the Appendix for a complete list of supported actions

Dynamic Rollback

Oracle GoldenGate provides a Dynamic Rollback feature that eliminates the need for full restore, helps maintain large test databases, and enables point-in-time and selective data recovery.

Eliminate the Need for Full Restore

Oracle GoldenGate can perform selective back-out (reverse) processing on enterprise databases, eliminating the need for full restore operations that, for large databases, usually require several hours or more to complete. The Dynamic Rollback feature captures and uses "before" and "after" images to undo database changes for user-defined tables, records, and time periods. This makes it ideal when data becomes corrupted or erroneously deleted.

Maintain Large Test Databases

Dynamic Rollback is also very effective for maintaining large test databases. Administrators can restore a test database to its original state before a test run, enabling test cycles to occur more quickly against a predictable baseline of data. Since Dynamic Rollback can undo all changes, reverting the database takes a fraction of the time compared to a full restore.

Point-in-Time and Selective Data Recovery

Using the Capture module to retrieve the database changes that have been committed, Dynamic Rollback can be applied to reverse operations to a specific point in time. The data is then analyzed and prepared for rollback by inverting the order of the database operations retrieved. This guarantees that records with the same key will be properly applied when done so in reverse order. The before and after image indicator in each record is modified, delete operations are

changed to inserts, and inserts are changed to deletes. Then, the begin and end transaction indicators are reversed to delimit each transaction. The Delivery module is invoked to apply the before images back to the database. Before executing, users can review the changes to be applied.

Customized Oracle GoldenGate Processing

Oracle GoldenGate functionality can be customized to meet specific needs.

User Exits

User exits are custom routines that can be called at different points during processing. With user exits, customers can respond to database events when they occur without altering production programs. For example, users can perform arithmetic operations, implement archival functions, gather statistics, or use exit routines as an alternative to native data transformation functions.

Stored Procedures and Queries

Stored procedures and queries can be called from Oracle GoldenGate to perform custom operations in the database's native procedural language. They can execute a query and they can retrieve output parameters for input into Oracle GoldenGate's filtering and mapping functions.

Macros

Macros offer a way to easily reuse parameters, commands, and conversion functions. They enhance productivity by enabling users to implement multiple uses of a statement, consolidate multiple commands, and invoke other macros.

Initial Data Loads

Oracle GoldenGate can be used for real-time CDC as well as for the initial loading to instantiate a database. Oracle GoldenGate's initial load capabilities provide continuous uptime and allow the application to be introduced into the data environment quickly and effectively.

Oracle GoldenGate allows you to carry out an initial load across multiple systems nonintrusively and without downtime, in three steps.

- **Start Capture.** Places any changed data after its start in a Trail File. The Capture module acquires source data in arrays instead of rows for improved performance.
- **Snapshot Load.** Takes a snapshot from the source and loads it directly into the target.
- **Start Delivery.** Applies the data previously placed by in the Trail File to the target to "catch up" the database until both are fully synchronized. Oracle GoldenGate can be configured to automatically switch to CDC after the initial load is completed.

Associated Products

There are two primary products that augment Oracle GoldenGate to enhance your real-time information platform:

- **Management Pack for Oracle GoldenGate.** A tool for visually deploying and managing Oracle GoldenGate processes across the enterprise.
- **Oracle GoldenGate Veridata.** A data comparison utility that quickly compares data between two online databases and reports any discrepancies (can run as a standalone product).

Management Pack for Oracle GoldenGate

Management Pack for Oracle GoldenGate is a centralized, server-based graphical enterprise application that offers an intuitive way to define, configure, manage, and report Oracle GoldenGate processes. It leverages the management services of the core Oracle GoldenGate platform to help users reduce the deployment time for their continuous availability and real-time data integration configurations.

Management Pack for Oracle GoldenGate minimizes human error and enables the rapid production of multiserver solutions from a central location to improve productivity and responsiveness, lower costs, and increase return on investment. The application centralizes configuration, management, monitoring, and reporting for Capture and Delivery processes running across multiple hosts, and provides real-time feedback on configuration and status changes—problems are instantly highlighted for corrective action.

Management Pack for Oracle GoldenGate does not impose any additional load on other deployed Oracle GoldenGate components. Further, it can be extended to work with third-party infrastructure management solutions to integrate into existing enterprise environments. Management Pack for Oracle GoldenGate includes a server component, a rich client, a Web client, and an administration console.

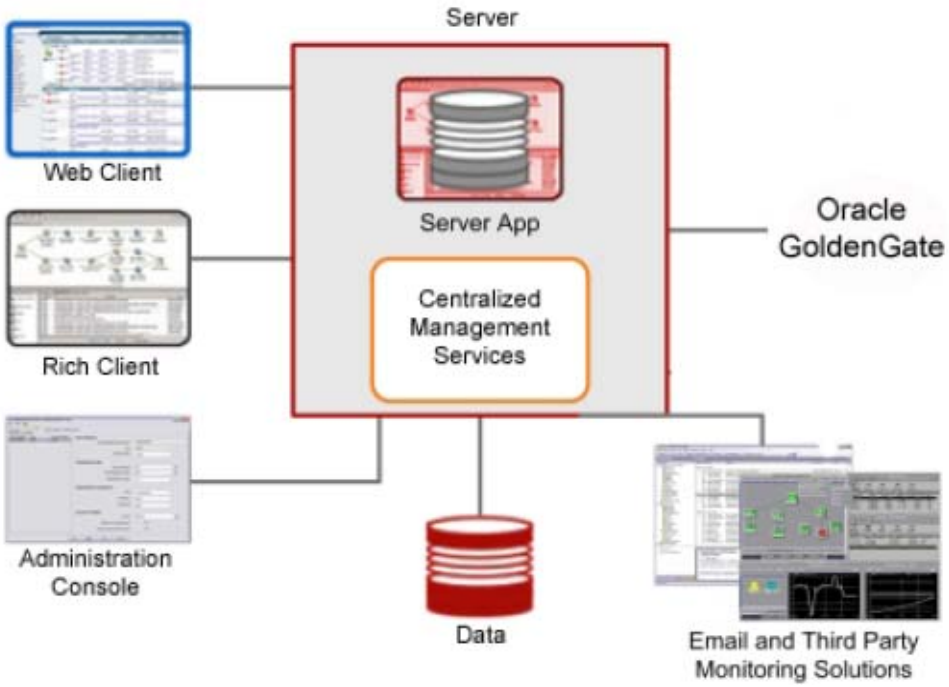


Figure 12. Management Pack for Oracle GoldenGate provides a visual interface to optimize Capture and Delivery management.

Residing on a central server, the Management Pack for Oracle GoldenGate server component communicates with deployed Oracle GoldenGate processes and third-party monitoring systems. Running on top of a Java application server, the server application

- Interacts with the rich client and Web client
- Provides a centralized repository for security settings, event logs, lag times, records processing, and other statistics about the Capture and Delivery modules
- Establishes and manages connections with Oracle GoldenGate processes
- Triggers e-mail notification alerts
- Enables integration with other monitoring solutions through prebuilt and custom adapters
- Supports remote command line sessions

Management Pack for Oracle GoldenGate's rich client interface is a Java-based client application that provides customized network views; easy, visual management of Oracle GoldenGate processes and tasks on any number of hosts; consolidated event logs; and e-mail and custom alert notifications. Management Pack for Oracle GoldenGate also provides a browser-based interface that offers some of the features of the rich client without the need to install additional software.

Oracle GoldenGate Veridata

Oracle GoldenGate Veridata is a high-speed, low-impact data comparison solution that identifies and reports data discrepancies between two databases, without interrupting those systems or the business processes they support. A standalone product, Oracle GoldenGate Veridata does not depend on the presence of Oracle GoldenGate's core components.

Data discrepancies result from a wide variety of causes. User errors, application errors, and infrastructure problems can all lead to out-of-sync conditions. Some of the causes are unintentional, such as input errors, unintended use, application bugs, system failures, disk corruption, and network outages. Other causes include malicious activity, such as security breaches or internal vandalism. In either case, inaccurate data leads to erroneous or ill-informed decision-making; failed service-level agreements; and, ultimately, financial and legal risk exposure.

Oracle GoldenGate Veridata reduces the amount of time and resources required to compare data. It eliminates the introduction of human errors and allows users to instantly identify and address potential problems. Oracle GoldenGate Veridata offers an intuitive, Web-based graphical user interface as well as command-line capabilities to allow for scheduled comparisons.

Through an intuitive Web interface, users of Oracle GoldenGate Veridata can

- Operate against live databases to compare data sets at very high speeds, but with almost no impact on the infrastructure
- Compare only the data that changed since the initial comparison, enabling high performance
- Select pertinent tables or data fields on the source and target databases to compare—not an all-or-nothing approach
- Run multiple comparison jobs in parallel
- Compare heterogeneous databases, including Oracle, Teradata, Enscribe, and SQL/MP, on different operating systems

To initiate the comparison, users simply perform a few simple clicks in Oracle GoldenGate Veridata's graphical interface or enter a command through its command-line interface. When Oracle GoldenGate Veridata starts, its agents on each datasource start to read the data from their respective databases. This allows the verification of data to be unattended.

Oracle GoldenGate Veridata agents use patent-pending technology to create row signatures that are not only very compact in size, but are also guaranteed to be unique. As the data on the source and target databases continue to change, Oracle GoldenGate Veridata server compares these row signatures and reports on persistent differences, as well as in-flight data that continues to change during the comparison. Users have the flexibility to determine how to handle in-flight data based on their own business requirements.

Conclusion

To remain competitive, businesses need to be able to access transactional data across disparate systems without interruption, and then rapidly and accurately convert it into information that is readily available across the enterprise. This enables them to streamline their business processes, which helps to optimize customer service and capture market opportunities.

However, managing transactional data and ensuring its availability, accessibility, and accuracy is not a trivial task. Oracle is committed to providing the highest-quality real-time CDC and replication software solutions that fully address the ever-evolving needs of mission-critical transactional systems. Oracle GoldenGate and supporting products are used to enable a range of IT solutions for hundreds of companies today, spanning business needs for continuous availability and real-time data integration. The architecture of Oracle GoldenGate allows for rapid deployment and enables Oracle to deliver real-time access to real-time information.

Appendix: Supported Actions for Conflict Avoidance, Detection, and Resolution

TABLE A1. SUPPORTED ACTIONS FOR CONFLICT AVOIDANCE

ACTION	DESCRIPTION
Map specific columns from specific sources	Where appropriate, different datasources can supply different segments of the target row.
Wait on an external event	Conflicts can sometimes be avoided by delaying the delivery of data until an external event occurs. Rules can be invoked at multiple levels so that operations coincide correctly with dependent actions at the target database.

TABLE A2. SUPPORTED ACTIONS FOR CONFLICT DETECTION

ACTION	DESCRIPTION
Identify conflicts in “dissimilar” data	Oracle GoldenGate facilitates mapping, transformation, combination, and normalization between datasources and targets. In these types of scenarios, a more-sophisticated detection rule can be created to correctly identify conflict.
Application-specific detection options	Application-specific rules can determine whether a conflict has occurred and specify what action is to take place in either case. Rules can be established by examining the current column values in the row, the column values of the operation, the source row values before the operation, database lookups, stored procedure results, and other information.
Invoke database stored procedures	Conflict detection rules can be created in stored procedures via familiar languages such as PL/SQL and Transact-SQL.
Invoke user exits	Custom code can be invoked to detect conflicts.

TABLE A3. SUPPORTED ACTIONS FOR CONFLICT RESOLUTION

ACTION	DESCRIPTION
Decide based upon the time stamp	Using the commit time stamps, conflicts can be resolved, where the latest or the first committed transaction can overrule the others.
Decide based upon the datasource	Conflicts can be resolved also based on the source of the transaction. For example, if one site "owns" the data more than another, this type of strategy might be appropriate.
Apply "delta changes"	In some cases, it might be appropriate to apply delta changes for specific columns. This method does not require any changes to the structure of the objects. It does require that both the before and after images of the delta change column be included in the Trail File.
Retry the operation	The operation can be retried at periodic intervals until the conflict is resolved externally. A maximum number of retries can also be specified.
Terminate processing	The conflict causes synchronization to terminate, pending external action. Synchronization can then proceed after external resolution. Processing can also be terminated after a specified number of conflict errors are encountered.
Log the operation into an "exceptions table"	Conflicting operations can be inserted into one or more exceptions tables. Logged data can include column values, type of operation, time stamp, and reason for conflict. Exceptions tables can subsequently be processed manually or via a custom program to resolve conflicts.

User exits can be invoked to resolve conflicts with more-complex rules—when existing methods mentioned above are not adequate.



Real-Time Access to Real-Time Information:
Oracle GoldenGate
Updated October 2009

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



Oracle is committed to developing practices and products that help protect the environment

Copyright © 2008, 2009, Oracle and/or its affiliates. All rights reserved. This document is provided for information purposes only and the contents hereof are subject to change without notice. This document is not warranted to be error-free, nor subject to any other warranties or conditions, whether expressed orally or implied in law, including implied warranties and conditions of merchantability or fitness for a particular purpose. We specifically disclaim any liability with respect to this document and no contractual obligations are formed either directly or indirectly by this document. This document may not be reproduced or transmitted in any form or by any means, electronic or mechanical, for any purpose, without our prior written permission.

Oracle is a registered trademark of Oracle Corporation and/or its affiliates. Other names may be trademarks of their respective owners.

0109