



Real-Time Operational Reporting for E-Business Suite via GoldenGate Replication to an Operational Data Store

An Oracle Technical White Paper
May 2012

ORACLE

Abstract	1
Oracle's Operational Reporting Problem in Service Contracts	2
Getting Real (Time)	5
How GoldenGate Works	6
Solution Architecture and Implementation	6
Results	11
Conclusion	15

Abstract

At Oracle, real-time operational reporting was a pipe dream for too many areas of its internal operations. The volume of data and velocity of change forced over 2,000 potential users across Sales, Operations, and Finance to make do with stale data or un-performant reports that were missing critical dimensions or attributes required to drive better decisions.

Nowhere was this pain more apparent than in Oracle's \$20 billion service contracts business. The ten year old contracts reporting solution running within Oracle's global single instance of E-Business Suite had reached a breaking point with the near-tripling of contract line volume brought on the acquisition of Sun Microsystems.

In 2011, Oracle implemented a new operational reporting system using GoldenGate real-time data replication to an operational data store. GoldenGate though proven for Siebel customers, had less frequently been used to replicate data from E-Business Suite. However, other methods, such as snap mirror copying of production OLTP systems to dedicated reporting instances, were unacceptable. They failed to reduce the data set and as a result required very expensive and ongoing SQL optimization across too many reports. Therefore this somewhat novel use of GoldenGate appeared to offer the best approach to real-time reporting on a reduced OLTP data set.

This paper describes the successful deployment of GoldenGate real-time data replication against Oracle's internal E-Business Suite and ancillary systems. In addition, the paper includes recommendations for how to architect and configure the solution, as well as before-and-after quantitative findings.

Although service contracts was among the first subject areas enabled, the use of this real-time reporting solution is quickly expanding to other Oracle business areas. It is our recommendation that all Oracle E-Business Suite customers with sizable deployments and a need for performant real-time operational reporting consider adopting the technologies and solution design patterns included herein.

Oracle's Operational Reporting Problem in Service Contracts

Service contracts represent over \$20 billion of revenue per year for Oracle, with well over 100,000 renewals and new contracts to be booked every quarter. Not surprisingly, there is a real and urgent need for real-time visibility into this business; a need that had increasingly gone unsatisfied. Sales Operations need to know which contract renewals have already been booked, which are in process, and which still need to be pursued with customers. Management needs rapid insight into key operational metrics, such as the bookings to date and which contracts have been cancelled and why.

Yet Oracle's ever-increasing volume of data and velocity of change made real-time service contracts reporting difficult to achieve on legacy systems. The service contracts business generates up to 800,000 changed rows of data in Oracle's global single instance (GSI) of E-Business Suite every day. Daily volume spikes as quarter end approaches – exactly when real-time visibility matters most. Increasingly stale data had begun to deprive Oracle of a good day-to-day picture of a major part of its business.

When OKI Was No Longer OK

Prior to 2010, service contracts operational reporting came from Oracle Contracts Intelligence (OKI); an older Daily Business Intelligence-based product bundled with the Oracle Contracts E-Business Suite application. OKI was deployed within Oracle's global single instance of E-Business Suite, and aggregated contracts reporting data directly from Oracle Service Contracts on a fixed refresh schedule. In effect, OKI was making very large data-aggregating queries directly against Oracle's global single instance. Oracle Discover then ran queries on OKI data to produce service contracts reports. Since OKI sat within Oracle's E-Business Suite instance, this burdened Oracle's single global instance with a second set of large and complex queries. Periodic extracts were also performed from OKI to Oracle's Global Corporate Warehouse, to produce analytical reports. This analytical reporting is outside the scope of this paper. Figure 1 shows the architecture of Oracle's legacy service contracts reporting solution.

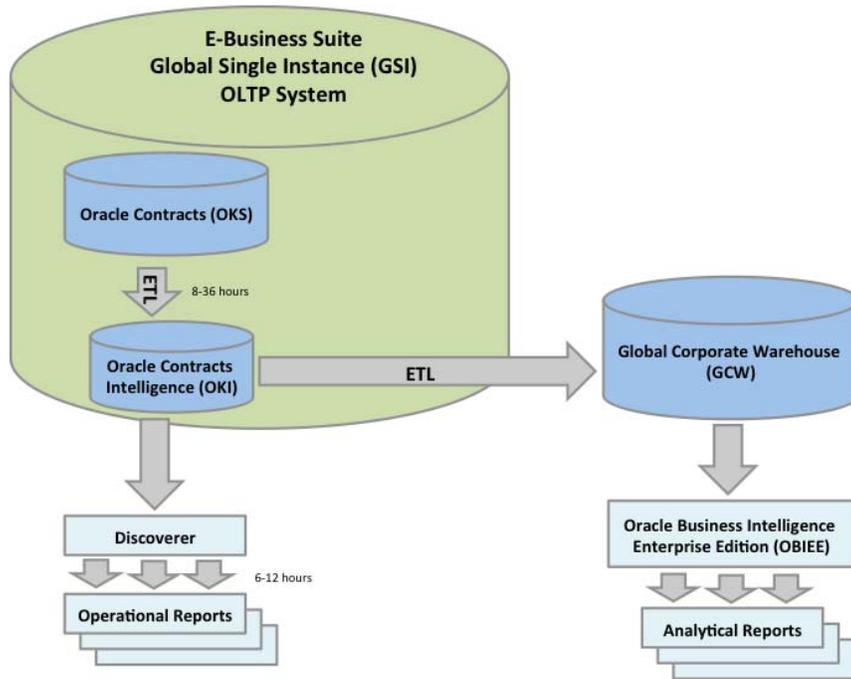


Figure 1. Functional Architecture of Oracle's Legacy Service Contracts Reporting Solution

This reporting solution had handled Oracle's volume when it was placed into service more than 10 years earlier. However as Oracle's service contracts volume grew, the OKI data grew staler and report runs lengthened. With the increase in contract line volume brought on by the acquisition of Sun Microsystems, the system neared a breaking point. .

Days-Old Data

Since it was placed into service in 2002, OKI had ran an incremental ETL process to aggregate service contracts reporting data twice a day, starting at 4a.m. and 4p.m. Pacific Time. Each incremental ETL had a typical complete time of three hours, and a longer-running complete refresh was performed weekly. From 2003 through 2009, completion time lengthened incrementally with the organic increase in Oracle's service contracts volume, but still completed comfortably in a twice-daily cycle. This meant that the data in OKI was never more than 12 hours old.

In July 2010 however, following the acquisition of Sun, Oracle merged Sun's legacy contracts data into Oracle's core systems. This increased Oracle's service contracts installed base to more than 600,000 active contracts, and contract lines from 14 million to 47 million. The size of daily updates made to contracts data increased dramatically.

Under ballooning volume, OKI could no longer complete twice-daily refreshes and had to be switched to once-daily aggregation. This should have meant that the data in OKI was never more than 24 hours old. However, to limit adverse performance impact on the ERP system, incremental ETL loads were limited to the number of transactions that could load in 12 hours. Transactions that failed to make this

cut fell to the next day's ETL, and delays cascaded until the OKI could receive its complete refresh on the weekend. Data lags in the run-up to the end of fiscal Q3 (March 2010), sometimes exceeded 48 hours.

This meant during Q3 OKI was at times more than 48 hours old before report runs even began. Only substantial code changes to the ETL process prevented the team from having to increase lag times even further. Figure 2 shows the sharp increase staleness of OKI data.

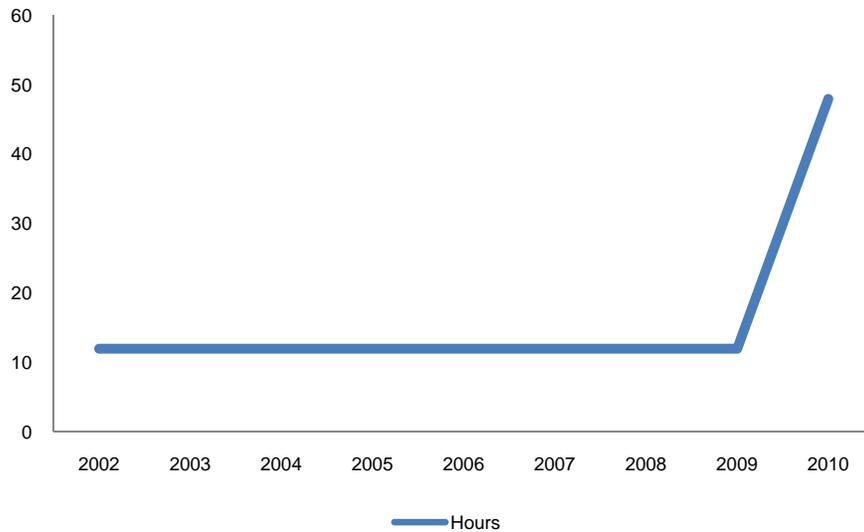


Figure 2. Maximum Staleness of OKI Service Contracts Data Year-Over-Year, Excluding Report Run Time

A hardware upgrade in fiscal Q4 helped to stabilize the system, but did not significantly wind back the data staleness problem.

Slow Report Runs

Even after data aggregation into OKI was completed, increasingly long report runs introduced additional delays. Most queries were run against huge volumes of data, since OKI stored several years worth of data. Therefore even the most efficient queries took significant time to complete. Queries that aggregated data further exacerbated the performance problem. The run times for all key OKI reports were so long that they had to be exempted from GSP's eight-hour kill script in order to complete at all.

At the system's worst point in early 2011, the very large Contract Base Detail report took more than eight hours to run. This eight hours of report run time stacked on top of up to 48 hours of ETL refresh cycle delay. A new transaction could therefore take up to 56 hours to show up in the Contract Base Detail report. Figure 3 below shows a breakdown of refresh delay and report run times for six key service contracts reports. Not reflected in the typical run times is the fact that larger reports would frequently exceed 12 hours of run time and therefore be terminated before completion.

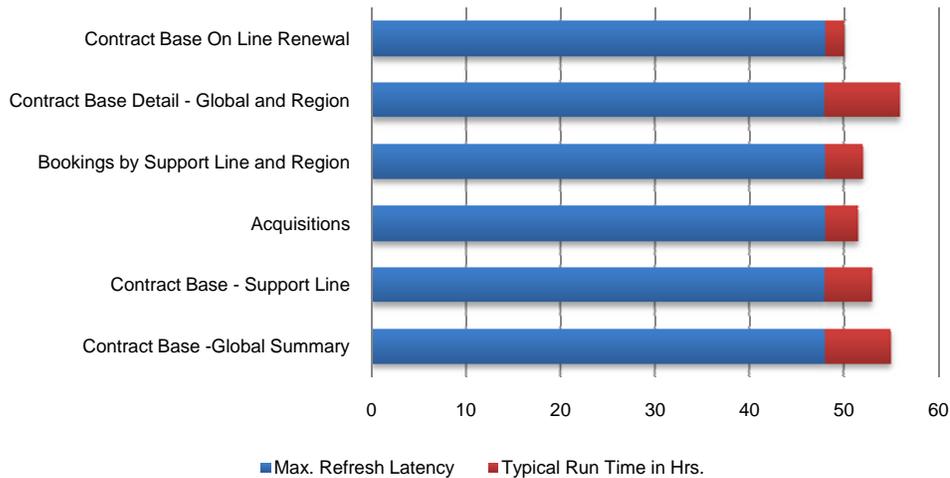


Figure 3. Maximum Combined Delay from Transaction to Report Completion for Key Service Contracts Reports

Stale data and long report runs forced the service contracts business to drive forward while looking in the rear view mirror. Operations and tactical decisions were handicapped by outdated information. It was not possible to provide an accurate daily bookings summary to executive staff because the report would be obsolete before it could be delivered. It was clear to both IT and business users that service contracts reporting required a radical overhaul. The question was, which approach would provide real-time insight on massive volumes of rapidly changing data?

Getting Real (Time)

Like many of its customers, Oracle considered a number of solutions to its operational reporting problem.

Table 1. Approaches Considered and Rejected By Oracle for Service Contracts Operational Reporting

OPTION	APPROACH	ADVANTAGES	REASON REJECTED
Snap Mirror-Copying	Mirror-copy production OLTP systems to dedicated reporting instance	Removes performance burden of report runs and ETL from main ERP instance	- "Snap" copying would take more than 8 hours - Fails to reduce dataset for reporting, requiring SQL optimization for every report.
Data Warehouse	Move service contracts operational reporting from ERP to Oracle's Global Corporate Warehouse	Removes performance burden of report runs and "TL" portion of ETL from main ERP instance Reduces dataset for reporting	- Daily/Weekly ETL refresh cycle of data warehouse still leaves data 24 hours stale prior to report run

Both of the approaches considered above could work under lower data volumes and/or less stringent requirements for data freshness. For example, Oracle's Global Corporate Warehouse is well-suited to analytical reporting where historical data is more important and up-to-the-minute data freshness less

so. However both of the approaches above had fatal flaws for the operational reporting requirements of Oracle's service contracts business.

Ultimately, Oracle IT concluded that no batch replication method could replicate the required volume of data frequently enough and report on it quickly enough to provide anything approaching real-time reporting. The solution needed to replicate transactions immediately, and run reports on them in minutes instead of hours. To do this, the system had to replicate only changes as they occurred in the ERP system, and even then only the relevant subset of those changes. It was this requirement that ultimately led the team to consider GoldenGate.

While GoldenGate real-time data-integration software had previously been proven for Oracle's Siebel customers, it had rarely if ever used to replicate data from Oracle E-Business Suite. However the real-time replication provided by GoldenGate appeared to offer the best method to provide real-time reporting on a reduced OLTP data set.

How GoldenGate Works

GoldenGate is real-time data-integration software that essentially replicates only changed data. Rather than doing batch ETL from the source database, GoldenGate monitors database redo (transaction) logs for change activity. It then replicates only these changed transactions into a target data store. To ensure reliable replication, GoldenGate also includes some built-in mechanisms for detecting errors and flagging exceptions.

Since not all transactions may be relevant to a particular use, GoldenGate can be configured to only replicate certain kinds of changed transactions. The resulting replicated data set can therefore be much smaller than the source database.

GoldenGate has been used extensively by Oracle customers against CRM data sources such as Siebel. A smaller number of customers had previously deployed GoldenGate against Oracle E-Business Suite. However GoldenGate technology appeared to be a good match for Oracle's operational reporting requirements. A successful deployment of GoldenGate against Oracle's large Global Single Instance (GSI) of E-Business Suite would not only solve the service contracts reporting problem, but also be quickly extensible to other subject areas.

Solution Architecture and Implementation

The final functional architecture of the solution is shown in Figure 4. This new architecture uses GoldenGate to integrate data into a new database dubbed the Operational Data Store (ODS). Reports are then drawn from the ODS using a dedicated instance of Oracle Business Intelligence Enterprise Edition (OBI EE).

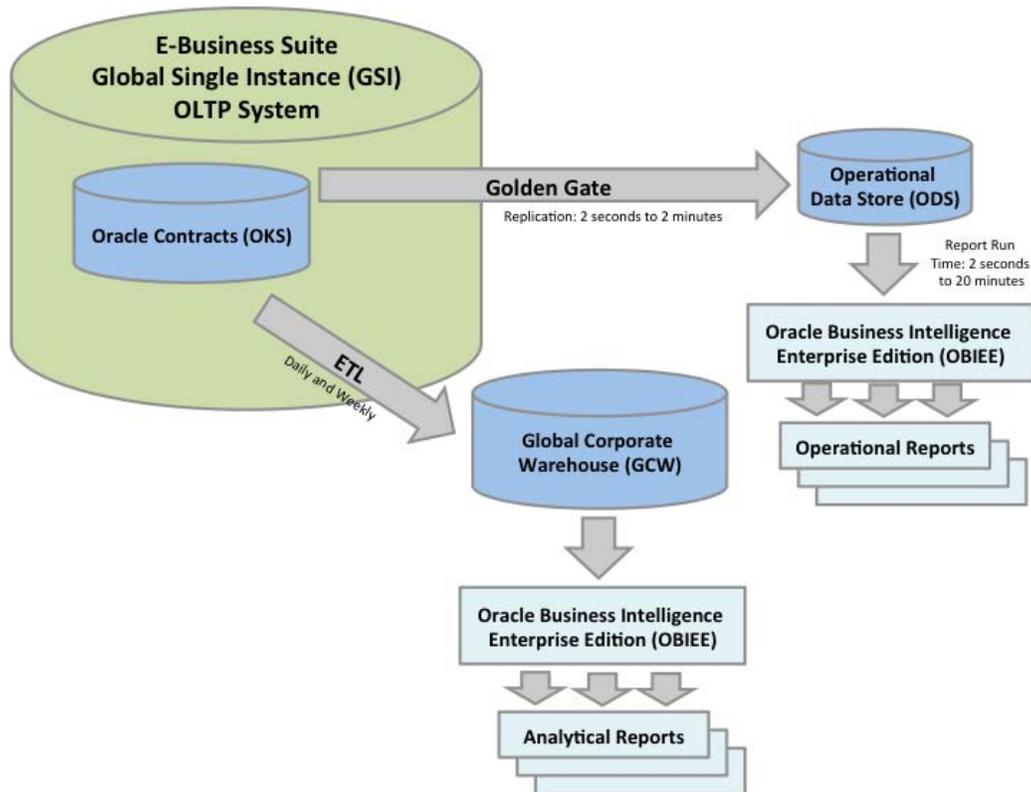


Figure 4. Functional Architecture of Oracle's Service Contracts Reporting Using GoldenGate and OBI EE

Operational Data Store

The Operational Data Store was designed to contain only the subset of source system data required for operational reporting. Oracle IT made two key architectural choices in defining the ODS: first to leave out historical data and second to duplicate the source schema as much as possible.

Since operational reporting is primarily concerned with what's new and what's changed, fresh data and fast report runs are paramount. Oracle's vast store of historical data – more than 10 years of service contracts data even before the Sun acquisition – is invaluable for analytical reporting. For operational reporting, this huge data store was an impediment.

Therefore the team determined that ODS would keep two quarters of data, and any data beyond this horizon would automatically be purged. Reports on data older than this would fall outside of our definition of operational reporting, and would be run against the Global Customer Warehouse rather than the operational reporting solution.

The second key architectural decision was to essentially replicate the schema of the source system, albeit with much less data. This simplified and improved performance of replication. Direct schema replication did impose a cost in the run performance of certain types of reports. However, because the ODS would contain a greatly reduced data set – typically about 2% of the GSI source data – the team was confident that they could draw performant reports directly from the replicated tables.

In practice, the ODS schema is very similar, though not identical, to the source schema. Replicated tables comprise about 90% of ODS data. They directly reflect the corresponding tables in Oracle's E-Business Suite instance, and are fed real-time by GoldenGate. Most operational reports were developed directly on these replicated tables. The few summary reports needed from the ODS – such as a report of the number of service contracts cancelled in the last month – are run against a small number of fact tables that pre-aggregate this summary data.

GoldenGate

As previously mentioned, GoldenGate replicates a defined set (subset) of new and changed transactions from the source system into the ODS. Conceptually, the main components in GoldenGate's decoupled architecture are Capture, Transmit, and Apply. Deploying GoldenGate replication involves setting up three main processes called Extract, Pump and Replicate. Extract pulls the changed data from the GSI redo log into transitional "trail" files. Setting up the Extract includes defining the subset of changed data that should be replicated. Pump processes the extracted "trail" files and transfers trail file contents over to the ODS environment in real-time. Finally, Replicate takes the "trail" file contents and loads them into the ODS database –reconstructing in effect the relevant subset of new and changed transactions. Setup includes an instantiation phase to create a one-time first copy of the source tables at the ODS. This can be achieved using GoldenGate's Initial Load method or other mechanisms such as a physical backup.

To iron out any kinks in this replication process, Oracle IT completed setup of these processes and began replicating data to the ODS three months ahead of go live. This allowed the team to fine tune replication settings and patch bugs before the functional go live for business users. Purge routines were also developed during this time to delete data older than two quarters. These purge routines became part of the system team's monthly operational schedule.

Although Oracle IT was not the first customer to deploy GoldenGate against a very large-scale E-Business Suite environment, we were among the early adopters of this use case; and were therefore very conservative in our timeline. Having ironed out many of the difficulties a customer might encounter, we believe a shorter period of advance replication will be sufficient for most deployments. That said, Oracle IT strongly recommends as a best practice separating the technical go live from the business go live. An early technical go live enables the project team to work out issues with data, systems and processes before users encounter them. Beginning replication at least several weeks prior to the business go live is one example of this approach.

OBI EE

The operational reporting solution uses a dedicated instance of OBI EE to deliver operational reports on the ODS data. The report footprint includes: quote requesting, quoting, ordering, shipping, planning, order holds, service contracts, new product introductions, and Spares management.

To create this expanded report footprint, functional "superusers" were identified from the various business areas. These superusers, typically process owners for their functional areas, helped design data sets and built reports based on templates provided by IT.

To enforce data governance policies, row level data security was implemented based on Oracle's Virtual Private Database technology. This security restricts user access by a combination of product and geography, as well as by role. For example, a sales representative from North America responsible for Exadata is only eligible to view orders in that region for that product.

Monitoring Tools and Tracking Reports

Since real-time speed was the overarching requirement of this solution, Oracle IT developed several tools to measure replication, database and reporting performance in real time.

The Heartbeat tool injects a transaction into a heartbeat table in the OLTP system, and reports how quickly it appears in the ODS. The team initially pointed GoldenGate to read archive logs from Oracle's disaster recovery instance, to mitigate the risk of overloading the OLTP. However the Heartbeat report showed that this configuration produced an average replication lag of about 20 minutes, as shown in Figure 5.

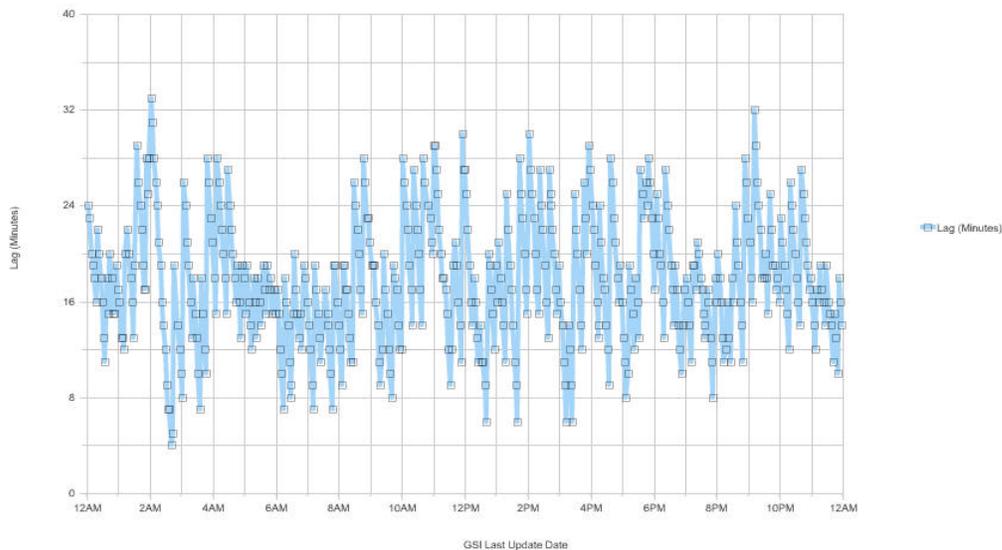


Figure 5. Heartbeat Report Log Graph Using GoldenGate with Archive Logs

Based on this data, the team re-pointed the extract process directly to the OLTP redo logs. This change dramatically improved replication speed. Transactions appeared in the ODS seconds after they occurred in the OLTP, as shown in Figure 6. Oracle IT now recommends this approach as a best practice for customers implementing GoldenGate.

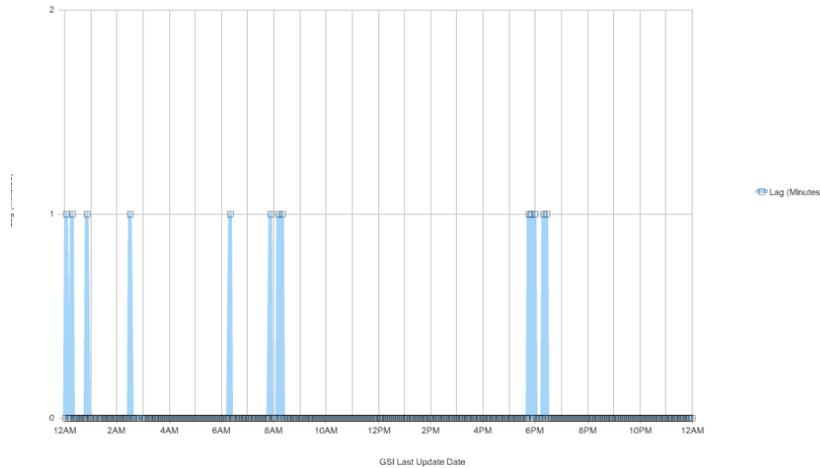


Figure 6. Representative Heartbeat Report Log after Implementing GoldenGate with Redo Logs

The Heartbeat tool also provides interactive reporting on transaction lags. This allows users to monitor lags for specific transactions at specific times. A sample of this report is shown in Figure 7.

GSI Transaction Last Updated Date	GSI Transaction Capture Date	ODS Transaction Last Replicated Date	Lag (Days Hours Minutes Seconds)
12/9/2011 2:37:46 PM	12/9/2011 2:37:46 PM	12/9/2011 2:37:48 PM	0 Days 0 Hours 0 Minutes 2 Seconds

ODS Current Lag Information - Replicate Process 1 : 0 Days 0 Hours 0 Minutes 2 Seconds

Figure 7. Lag Report Sample

Finally, the Lag Alert tool notifies appropriate personnel by email whenever Heartbeat lags exceed a certain threshold. This threshold was initially set at two minutes, but in the final configuration was reduced to two seconds. A sample of these alerts is shown in Figure 8.

ODSAP_Lag_Alert				
REPLICATE	LAST_UPDATE_TS	AUDIT_TS	SYSDATE	LAG_IN_SECONDS
Replicate # 1	2/9/2012 12:00:00 AM	2012-02-08 23:59:58	2/9/2012 12:00:03 AM	5
Replicate # 2	2/9/2012 12:00:01 AM	2012-02-08 23:59:58	2/9/2012 12:00:03 AM	5

Figure 8. Lag Alert Message Sample

In addition to the Heartbeat tool, several tracking reports were developed. The Daily Alert report monitors the daily average and standard deviation of elapsed report times. Automated Workload Reports (AWR), a database utility, provides key technical information about the load on the database – enabling the team to identify and weed out long running queries. As a fail-safe, a kill script was put in place to kill any query running longer than 30 minutes. Tracking reports were also developed to identify and show report performance for the most used reports and heaviest-consuming users, in preparation for the expansion of the ODS to a larger user base and more business subject areas.

As a complement to the monitoring tools and reports discussed above, the team developed load utilities to simulate production-like traffic on our development and test instances. While no load utility can produce all of the corner cases that occur in actual usage, these utilities did help us test the scalability of our system and assess the performance impact of any changes.

Challenges and Lessons Learned

Oracle IT encountered some lags in replication performance that were attributable to issues in the GoldenGate product that surfaced in this novel use case. These performance issues were addressed by the development team and are part of the current GA product release. . Some other challenges were not due to product bugs, but have been or will be solved by product enhancements in GoldenGate. One example is the handling of compressed tables. Tables that were compressed on the source OLTP needed to be uncompressed in order for GoldenGate to be able to replicate them. Decompression of tables would require downtime on the OLTP instance. Since our deployment, Oracle GoldenGate 11.2.1, which support Compression, has been released. Customers using the compression feature of the Oracle database on tables they wish to replicate should implement this GoldenGate version.

Oracle IT also had to put into place strict process controls to ensure that any changes to the OLTP data model were also made to the ODS before data replication was restarted. GoldenGate now provides automatic replication of data definition language (DDL) changes, which automates synchronization of such data model changes. Oracle IT plans to deploy this feature in the near future, and once deployed, process controls will be relaxed.

Results

Using GoldenGate on E-Business Suite to drive Oracle's service contracts operational reporting solution has exceeded expectations. Data freshness is now measured in seconds, and even very large reports now run in seconds or minutes rather than hours. These improvements in have already driven a significant increase in usage, and the ODS solution is now being expanded from service contracts into other areas of Oracle's business.

Data Freshness Improved 1000X

Since the functional go live of the GoldenGate ODS solution, data freshness has improved literally by orders of magnitude. This is the case not only in comparison with the performance of OKI at its most overloaded in 2010, but also even in comparison with OKI running twice-daily refreshes circa 2002-2009. Table 2 below displays differences in data freshness between OKI and ODS at 2012 data volumes.

Table 2. Differences in Data Freshness Between OKI and the GoldenGate/ODS Solution

DATA FRESHNESS IN SECONDS	OKI	ODS	DECREASE	% DECREASE	% IMPROVEMENT
Worst	172,800.00	78.00	172,722.00	99.95486%	221,438%
Best	86,400.00	1.00	86,399.00	99.99884%	8,639,900%
Average	129,600.00	11.88	129,588.13	99.99084%	1,091,268%

It should be noted that replication speed of GoldenGate is sensitive to hardware I/O bandwidth. Oracle IT initially encountered Input/Output (I/O) bottlenecks in ODS performance because of the older hardware on which the system was deployed. To reduce the cost of the project, the team initially implemented the system on a Sun X4150 with NetApp storage. This was not an ideal platform because real-time operational reporting and replication are both sensitive to I/O bandwidth between the server and storage systems.

In February of 2012, the team upgraded the hardware for the ODS to an Exadata X2-8 engineered server, which the ODS shares with Oracle’s Global Corporate Warehouse. This system includes two 40 GBPS Infiniband switches and a 5.3 terabyte flash cache, and therefore removed the I/O bottleneck and provided further performance gains. Except where otherwise noted, the performance numbers cited for the ODS are for this final configuration.

While the great majority of performance gains are due to the structural difference between ODS and OKI, this upgrade did provide further performance gains. Most notably, it eliminated occasional spikes in replication lag time which had previously degraded data freshness from seconds to as long as 12 minutes. Current replication lags range from one second to no more than 120 seconds.

Report Run Performance Improved 60X

Report run performance has also shown order-of-magnitude improvements using the ODS, particularly for the six large and complex key service contracts reports. The longest-running of these, Contract Base Detail, currently has shown more than 23X improvement, running in 20 minutes instead of eight hours. This substantial improvement is actually the lowest performance improvement of all service contracts reports. The other five key service contracts reports range from 84X to 149X improvement – with a mean improvement for all service contracts reports of 64X. Figure 9 contrasts report run time for the legacy OKI solution with the same reports run from the ODS on Exadata hardware.

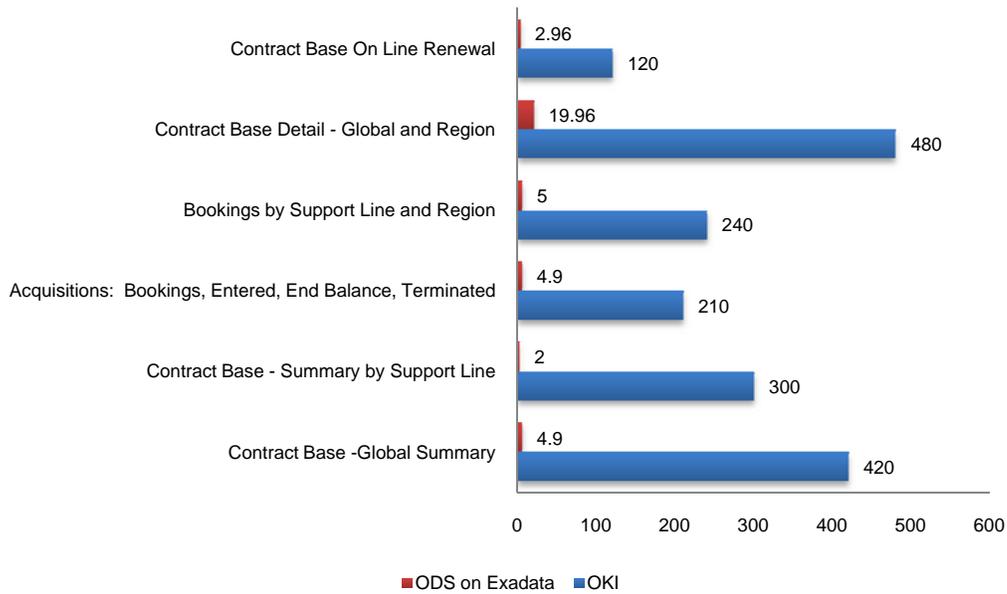


Figure 9. Representative Runtimes for Key Service Contracts Reports in Minutes

Because this improvement in runtime combines with improvements in data freshness, reports such as Daily Bookings Summary, which would previously have been outdated on arrival, can now be provided to management.

Usage Expansion

Not surprisingly, user adoption and system usage has increased substantially since real-time service contracts reporting went live. Figure 10 shows the rise in service contracts report queries and distinct users through March of 2012.

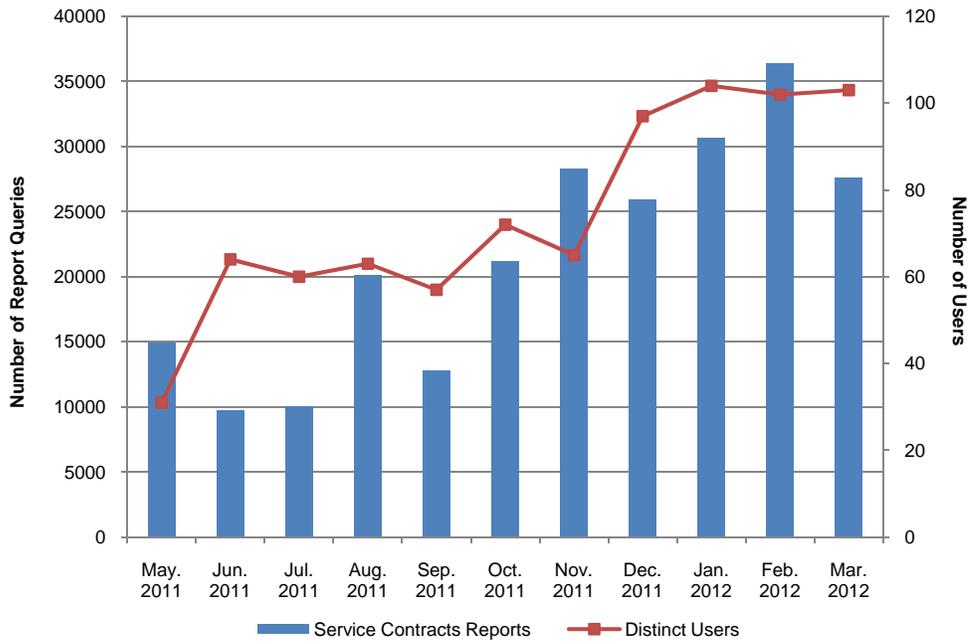


Figure 10. ODS Service Contracts Report Queries and Distinct Users

In addition, the success of the ODS solution has led to its rapid expansion into other business areas. Report content is either in use or in development for quoting, usage tracking, and especially multiple areas of supply chain management. Figure 11 below shows the increase in report runs and distinct users across all subject areas.

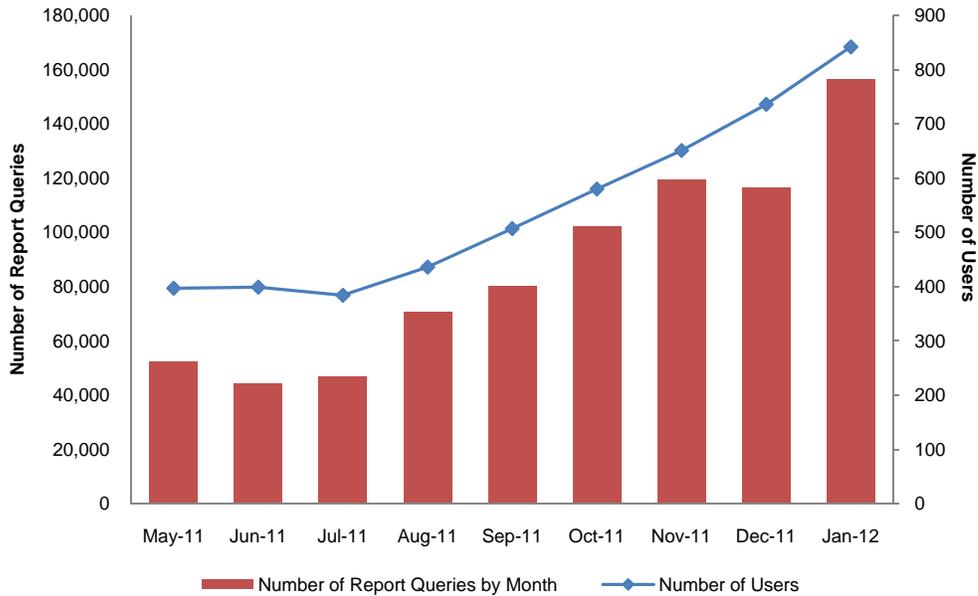


Figure 11. ODS All Report Queries and Distinct Users

Across all subject areas, the ODS currently runs an average of more than 5000 report queries per day, with an average run time under 30 seconds.

Conclusion

Real-time replication to an operational data store via GoldenGate has proven highly successful or Oracle’s service contracts reporting – with gains in both data freshness and report performance in the thousands of percents.

Oracle IT believes based on our experience that moving operational reporting to a real-time ODS can substantially improve operations and operational decision-making in the business. In addition, the ODS also eliminates a source of strain on Oracle’s Global Single Instance by removing large data aggregations and multi-hour report runs from within GSI.

Oracle IT now recommends this approach for any E-Business Suite customers facing similar requirements for real-time operational reporting.



Oracle's Internal Real-Time Operational Reporting for E-Business Suite via GoldenGate Replication to an Operational Data Store

May 2012

Authors: Vinay Deshmukh, Senior Director; Rajesh Panda, Group Manager; Mark Field, Vice President; Jeffrey Pease, Vice President; Dave Stephens, Group Vice President; Alok Pareek, Vice President

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 © 2012, 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 and Java are registered trademarks of Oracle and/or its affiliates. Other names may be trademarks of their respective owners.

Intel and Intel Xeon are trademarks or registered trademarks of Intel Corporation. All SPARC trademarks are used under license and are trademarks or registered trademarks of SPARC International, Inc. AMD, Opteron, the AMD logo, and the AMD Opteron logo are trademarks or registered trademarks of Advanced Micro Devices. UNIX is a registered trademark licensed through X/Open Company, Ltd. 0611

Hardware and Software, Engineered to Work Together