Five New Data Integration Requirements

And How to Meet Them with Oracle Data Integration
Disclaimer

The following is intended to outline our general product direction. It is intended for information purposes only, and may not be incorporated into any contract. It is not a commitment to deliver any material, code, or functionality, and should not be relied upon in making purchasing decisions. The development, release, and timing of any features or functionality described for Oracle’s products remains at the sole discretion of Oracle.
Executive Overview ........................................................................................................... 2
Five New Requirements in Data Integration .............................................................. 2
Oracle’s Data Integration Strategy ............................................................................ 3
Contending with Big Data Requirements .................................................................. 5
Any Source, Any Target ......................................................................................... 8
Data Integration On-Premise and in the Cloud ......................................................... 9
Meeting “Zero Downtime” Requirements ............................................................... 11
Real-Time Analytics for Trusted Information ....................................................... 13
Next-Generation Analytics with Oracle Engineered Systems ............................. 15
Ensuring Data Quality for Trusted Information ................................................... 18
Conclusion ............................................................................................................ 20
Executive Overview

Today’s decision-makers need timely, accurate data to create value for the enterprise. In order to gain complete insight about the business, they must be able to access all relevant enterprise data, regardless of the source or originating format, and integrate it across social, mobile, and cloud technologies.

As the business world responds to these challenges and opportunities, many enterprises are turning to data integration technology to obtain pervasive, continuous access to trusted data. In addition to assisting with modern initiatives such as big data, real-time analytics, and cloud computing, data integration tools help organizations with fundamental tasks related to building, deploying, and managing data warehouses, data marts, decision support systems, data hubs, and many other critical projects. Data integration is also essential to many strategic IT initiatives including master data management (MDM), modernization, and service-oriented architecture (SOA).

Customers need data integration technology to support the volume, velocity, and variety of big data as well as to ensure that their analytic initiatives are able to support a growing number of operational users with access to up-to-date, trusted data. Data integration also plays a leading role in data migration, consolidation, and modernization initiatives by structuring the interchange of information and improving system performance.

This white paper describes Oracle’s mature, well-crafted strategy for meeting these integration requirements. Oracle Data Integration delivers pervasive and continuous access to timely and trusted data across heterogeneous systems. It includes a broad family of products designed to deliver maximum performance with low cost of ownership, ease of use, and reliability. Its comprehensive capabilities support the fundamental requirements of the enterprise including real-time and bulk data movement, data transformation, bi-directional replication, data services, data federation, and data quality for customer and product domains.

Five New Requirements in Data Integration

As businesses strive to compete more effectively, they increasingly rely on data as a critical asset. Fresh, accurate data allows them to make better decisions and develop a more complete understanding of the competitive and operational aspects of the enterprise. Data volumes are growing exponentially, yet the speed of business requires organizations to process this data faster to gain accurate and actionable
information to create better results for the business, such as cost savings via operational efficiencies and higher revenue via an improved customer experience.

The onset of big data processing and cloud computing has made it much more difficult to bring disparate data silos together. Obtaining accurate insights from large volumes of data gets more complicated as data sources and IT environments change. Customers, partners and employees need to be able to access information on a 24/7 basis, with high performance and inherent security.

Clearly the requirements for data integration technologies have evolved significantly. Organizations must be able to integrate any data from any source, stored on premise or in the cloud, with maximum performance, so that they can provide 24/7 access to timely and trusted information.

Figure 1. Five new requirements for data integration

Oracle’s Data Integration Strategy

The Oracle data integration strategy fulfills these fundamental requirements:

- Pervasive access to any data of any source
- On premises or in the cloud
- Continuous access to information with maximum availability and performance
- Using timely data in decision making
- Ensuring high data quality for trusted information

Oracle executes on this strategy with a portfolio of integrated data integration products:

Oracle Data Integrator (ODI) offers unique Extract Load and Transform (ELT) technology that improves performance and reduces data integration costs—even across heterogeneous systems. Hot-pluggable Knowledge Modules provide out-of-the-box modularity, flexibility, and extensibility. Unlike conventional Extract Transform and Load (ETL) tools, ODI delivers the productivity of a declarative
design approach and the benefits of an active integration platform for seamless batch and real-time integration.

**Oracle GoldenGate** offers real-time, log-based change data capture, routing, transformation, and delivery between heterogeneous systems with minimal overhead. The software offers bidirectional data movement and enables critical systems to support 24/7 operations as well as gain access to real-time data for better business insight. A typical environment includes capture processes, trail files, and delivery processes, each of which can run on most of the popular operating systems and databases.

**Oracle Enterprise Data Quality** enables organizations to measure, improve, and manage the quality of data from any domain, including customer and product data, and minimize exposure to risk. Oracle’s data quality products provide powerful data profiling, cleansing, matching, and monitoring capabilities, unparalleled ease of use, and integration with Oracle Master Data Management.

**Oracle Data Service Integrator** provides real-time data access and data federation, enabling companies to quickly develop and manage federated data services for accessing single views of disparate information. Oracle Data Service Integrator is the only data federation technology that supports the creation of bidirectional (read and write) data services from multiple data sources. In addition, Oracle Data Service Integrator offers the breakthrough capability of eliminating coding by graphically modeling both simple and complex updates to heterogeneous data sources.

**Oracle Data Integration 12c**

The new Oracle Data Integrator 12c and Oracle GoldenGate 12c product releases advance Oracle’s leadership in the market even further by bringing extreme performance and future-ready technologies in simplified and integrated solutions.

**Oracle Data Integration 12c**

*Delivering Real-Time Integration for Big Data and Cloud*

![Diagram](image)

- **Future Ready**
- **Extreme Performance**
- **Fast Time to Value**

Figure 02. Oracle offers a complete set of solutions for modern data integration requirements.

Both Oracle GoldenGate 12c and Oracle Data Integrator 12c introduced new features that deliver higher IT productivity and extreme performance, and have deeper integration that enables end-to-end real-time data integration in a single platform. In addition, Oracle’s data integration products
Five New Requirements for Data Integration

have tighter integration with other Oracle technologies, including Oracle Database, Oracle Coherence, and Oracle Applications, enabling faster time to value and streamlined solutions for customers. This flexible and robust offering helps organizations easily keep pace with new data-oriented technology trends, including cloud computing, big data analytics, and real-time replication.

The three pillars of benefits of Oracle Data Integration 12c are higher IT productivity, future-ready, flexible solutions that deliver extreme performance for today’s and future processing needs.

- **Higher IT productivity.** Oracle GoldenGate 12c creates an abstraction layer where deployment details do not need to be specified and the product automatically optimally tunes the delivery processes, saving time while improving performance. Oracle Data Integrator 12c’s framework already contains such an abstraction layer, further adding to productivity gains. The new flow-based declarative user interface brings superior developer productivity, ease of use, and ultimately fast time to market for end users. Ability to seamlessly reuse mapping logic speeds development significantly as well.

- **Future-ready, flexible technology.** Featuring architecture designed to work with heterogeneous systems and deep integration with the Oracle technology stack, these products make it very easy to adjust applications to changing business requirements. Especially the tight integration with Oracle Database, Oracle Fusion Middleware, and Oracle’s application offerings, which means that as these products evolve and provide more advanced features, Oracle’s data integration solutions will be able to support them right away. For instance, Oracle provides a utility to help customers move to Data Integrator from the terminal release of Oracle Warehouse Builder to take advantage of the newest Data Integrator features.

- **Extreme performance.** When it comes to reliable performance, the new releases further separate Oracle from the rest of the market. As an example, Oracle GoldenGate 12c’s Integrated Delivery feature enables higher throughput via a special application programming interface into Oracle Database, while Oracle Integrator 12c brings parallelism that significantly increases its performance as well. Recent benchmarks show, at a minimum, the new feature delivers two times faster performance, while customers note up to five times higher performance.

Please access and review our free resources to learn more about Oracle Data Integrator 12c and Oracle GoldenGate 12c.

The following sections describe how Oracle’s data integration products support today’s and future data integration requirements.

Contending with Big Data Requirements

Due in part to the adoption of cloud and social computing models, enterprises are seeing a huge growth in the amount of raw data collected on a daily basis—commonly referred to as big data. Big data is very complex. There are multiple technologies, formats and sources. In order to get the most from big data, businesses must acquire, organize, and analyze the data deluge. Structured and unstructured data must coexist and be used in conjunction with each other in order to gain maximum
insight. In addition to ERP, CRM and other enterprise information systems, organizations must contend with data from sensors, conversations, e-commerce websites, social networks, and many other sources.

Big data presents huge opportunities in healthcare, manufacturing, government, retail, and many other industries. For example, the combination of social media data, weblogs, and structured data in a data warehouse or CRM system gives retailers a more complete view of not only what their customers purchased, but also what they might purchase.

Big data changes the question from “How do you look at your data? to “How do you look at the data that is relevant to you?” It’s a subtle semantic difference, but a huge leap in information-management best practices. With the right data integration technology, companies can tap the data that is relevant to them in a variety of forms and formats.

Figure 3. Big data presents information in a variety of forms and formats.

Solving Big Data Challenges with DI Technology

Clearly, big data presents a new set of IT challenges. However, it is also an opportunity to utilize unstructured information in a way that can provide fresh insight and new value. Big data imposes three basic requirements on data integration tools:

- The need to process high volumes of data at high velocity in any format, supporting a range of heterogeneous sources
- The need for unified design-time environments to help correlate big data with other enterprise data
- The need to integrate or “pre-bundle” big data technologies with relational technologies to streamline operations

To make effective use of big data, forward-looking organizations use data integration technology to achieve a common, consistent way to manage customer and partner information across the
information lifecycle. Data integration technologies help them to maintain tight linkages among all types of data, so they can achieve a complete portrait of the business.

In addition to Oracle Big Data Appliance for acquiring and organizing big data, Oracle offers Oracle Big Data Connectors that enable an integrated data set for analysis. Big Data Connectors include:

- Oracle Loader for Hadoop, a MapReduce utility that optimizes data loading from Hadoop to Oracle Database
- Oracle Direct Connect for Hadoop Distributed File System (HDFS), a high-speed connector for accessing HDFS data directly from an Oracle Database
- Oracle R Connector for Hadoop, which enables interactive access to Hadoop data from R
- Oracle Data Integrator Application Adapter for Hadoop, which allows organizations to build Hadoop metadata within Oracle Data Integrator, load data into Hadoop, transform data within Hadoop, and load data easily and directly into Oracle Database utilizing Oracle Loader for Hadoop

Oracle Data Integrator has the necessary capabilities for integrating structured, semi-structured, and unstructured data. It provides a unified design environment for developing MapReduce transformations, for transforming structured data, and for developing Hive-based transformations, all in a single utility. By providing prebuilt integration with Big Data technologies, such as HDFS and Hive, businesses are able to leverage additional sources of data previously too large and unwieldy to gain benefits from. Companies can run big data transformations that integrate directly with Oracle Loader for Hadoop to merge data into and out of Hadoop environments.

Oracle Data Integrator 12c delivers integration for Big Data technologies leveraging a metadata driven process while enabling data integration resources to easily manage how Big Data is extracted, loaded, and transformed.

Figure 4. Oracle Data Integrator and Big Data Connectors handle all types of loading requirements.
Any Source, Any Target

In addition to handling all types of data, data integration tools need to support different sources and targets. Oracle Data Integration product family provides the level of heterogeneity that organizations require. Both Oracle Data Integrator and Oracle GoldenGate support target and source database software from different vendors and operating systems. In addition to capturing both compressed and non-compressed data from Oracle Database, Oracle GoldenGate supports non-Oracle databases such as IBM DB2 for iSeries, PostgreSQL, Teradata, Microsoft SQL Server, Sybase, and MySQL. GoldenGate can also capture, map and transform data from databases using multibyte/Unicode character sets.

Oracle Data Integrator also offers complete heterogeneity with its flexible Knowledge Module components. It provides heterogeneous support for non-Oracle data sources and data warehouse appliances. While Oracle Data Integrator is certified with Oracle technologies and applications it is optimized for mixed targets, sources and applications.

The Heterogeneity of Oracle Data Integration

- IBM DB2 UDB
- IBM DB2 z Series
- IBM DB2 i Series
- Enscribe
- SQL/MP
- SQL/MX
- MySQL
- Sybase ASE
- Ingres
- JMS
- Teradata
- Nelezza
- Hadoop
- Sybase IQ
- TimesTen
- PostgreSQL
- Greenplum
- HP Neoview
- SAS
- SalesForce
- SAP ERP & BW
- Generic SQL
- Hypersonic SQL
- Microsoft SQL
- Oracle Database
- Oracle Exadata
- Oracle Big Data Appliance
- Enterprise Data Quality
- Oracle E-Business Suite
- JD Edwards Enterprise One
- Oracle Enterprise Service Bus
- Oracle Hyperion
- Oracle OLAP
- Oracle PeopleSoft
- Oracle Siebel CRM
- Oracle Communications BRM
- Microsoft Access
- Oracle BI EE
- Oracle BI Apps
- Linux
- Oracle Solaris
- Windows 2000, 2003, XP
- HP NonStop
- HP-UX
- IBM AIX
- zLinux
- ...

Figure 5. Oracle Data Integrator and GoldenGate work with all major database and OS environments.

Leveraging Mainframe Data Across the Enterprise

Heterogeneity provides the flexibility to move data anywhere. Oracle Data Integrator and Oracle GoldenGate support multiple operating environments, permitting customers to leverage many types of data. One common scenario is to move mainframe data to other OLTP systems and create a reporting database to offload MIPS-intensive queries. For example, GoldenGate captures changed data non-invasively from IBM DB2 on iSeries and zSeries and delivers it to Oracle Database in real time. This capability permits users to access up-to-date information without adding overhead to the mainframe environment.

GoldenGate’s query offloading solution improves the performance of the production systems and decreases costs by limiting MIPS usage and delaying or avoiding mainframe upgrades. It can replicate all or a portion of data and even manipulate the data within any of these processes for both heterogeneous environments and also different database schemas.
Data Integration On-Premise and in the Cloud

Most enterprises have spent years avoiding the data “silos” that inhibit productivity. Do cloud applications represent a new silo? As organizations and individual departments adopt cloud applications, in many cases they create point-to-point connections that bypass well-established integration principles. Lack of centralized monitoring and management causes extra work for system administrators, especially as cloud data and functions are shared with on-premise information systems.

Data integration solutions play a vital role in migrating data simply, efficiently, and reliably to the cloud. They are essential to platform-as-a-service (PaaS) implementations because they support cloud deployments with data-layer application integration between on-premise and cloud environments.
Consolidating into Cloud

Oracle Data Integrator and Oracle GoldenGate can connect on-premise enterprise systems to a private cloud by moving data in bulk or as real-time transactions across geographies. GoldenGate's real-time and bidirectional data replication capabilities between heterogeneous systems enable companies to consolidate data without interrupting business operations. Oracle GoldenGate 12c is optimized for Oracle Database 12c and supports its new multitenant architecture, making it a perfect solution for consolidation without impacting business operations. A failback option allows the IT team to test the target environment as long as necessary, minimizing risk. In addition, Oracle Data Integrator offers data services for service-oriented architecture (SOA) and business process management (BPM) solutions that enable agile application delivery and shared data services for cloud deployments.

Integrating Public Cloud Environments with On Premises Systems

Some of the key requirements in keeping the data fresh in hybrid environments include—minimal latency, reliability and security. Data must remain fresh. As data ages it becomes less relevant and less valuable—day-old data is often insufficient in today's competitive landscape. Reliability must be guaranteed despite system or connectivity issues that can occur between the cloud and on-premises instances. Oracle GoldenGate offers real-time data replication between on-premises databases and databases hosted on public cloud environments to enable reliable, secure and timely data integration.

To learn more please review the knowledge article (ID - 1588484.1) 'Replicating between Cloud and On-Premises using Oracle GoldenGate'.

Oracle Cloud enables an organization to move data through Oracle SOA Suite using REST APIs to Oracle Messaging Cloud Service—a new service that enables applications deployed in Oracle Cloud to securely and reliably communicate over Java Messaging Service. Oracle Data Integrator also supports a knowledge module for Salesforce.com, now available on AppExchange. Oracle also offers data adapters for Workday, Amazon, and other cloud providers. Oracle Cloud implements these same technologies in addition to best-of-breed database, middleware, security, business intelligence applications, and engineered systems.
Meeting “Zero Downtime” Requirements

As business processes move online, and operate globally 24/7, the tolerance for any type of system outage is approaching zero. This is particularly true for cloud applications with strict availability requirements. Mission-critical systems and many cloud service providers can’t afford to take down their information systems. This means they need to minimize downtime for upgrades, migrations, software patches, and other maintenance work, which often leads to postponing upgrade or modernization projects. Some data integration products focus on unplanned outages with disaster recovery oriented solutions. However 24/7 operations must also accommodate planned outages, which is why many organizations now seek solutions that address all types of maintenance, upgrade, and migration scenarios as well.

Continuous Access

With its bidirectional and heterogeneous transactional data replication capabilities, Oracle GoldenGate eliminates not only unplanned downtime but also planned downtime with support for flexible yet complete solutions. Oracle GoldenGate enables organizations to complete “rolling upgrades” and “rolling migrations” with little or no application downtime, enabling continuous operations and access to information.

GoldenGate’s bi-directional replication capabilities enable reverse synchronization from the new system to the old systems after the switchover, enabling users to failback to the old system if unexpected issues arise. It can also allow two systems to run in parallel during the migration process until the company decides to retire the old system. This active-active database set up enables phased migrations and minimizes risks for mission-critical systems. Oracle GoldenGate Veridata offers data comparison and validation between heterogeneous systems before switchover to the new environment in the case of upgrades, or for ongoing testing of the standby database. Oracle GoldenGate Veridata is a high-speed data-comparison solution that identifies and reports on data discrepancies between heterogeneous databases without interrupting ongoing business processes.

Oracle GoldenGate’s bidirectional data replication capabilities support heterogeneous databases and operating systems and enable Active-Active database replication among geographically distinct platforms and data centers. This solution also addresses unplanned downtime with complete seamless transitions and also improves system performance by distributing the load to the servers.

With its flexible and robust solutions, Oracle GoldenGate is a key component of the Oracle Database Maximum Availability Architecture. The product also provides high availability solutions for non-Oracle databases such as MySQL, HP NonStop, DB2, SQL Server, and Sybase ASE.

Maximum Availability Architecture and GoldenGate

The Oracle Database Maximum Availability Architecture (MAA) leverages several industry-leading high availability (HA) technologies, extensive validation performed by the Oracle MAA development team, and the accumulated experience of customers who have successfully deployed business-critical
Five New Requirements for Data Integration

Oracle GoldenGate has a key role within Oracle Database MAA and extends the availability of Oracle Database in the following ways:

- Zero downtime upgrades and migrations
- Online platform and application upgrades
- Bi-directional and multi-master replication for ‘active-active’ scenarios
- An alternative to physical replication for site protection
- Flexible planned maintenance and heterogeneous migrations

With the new 12c release Oracle GoldenGate offers integration with Oracle Data Guard FSFO to provide automated and transparent failover of Oracle GoldenGate components with the failover/switchover of the primary database. This allows the replication to continue without any manual intervention. With the new release GoldenGate’s Downstream Capture can also utilize the Data Guard redo log transport mechanism to process transactions off-source. When redo logs are shipped to a remote machine where GoldenGate Capture process is installed, the solution eliminates the risk of data loss completely and meets very strict data recovery objectives.

Bank of Valletta Enjoys Continuous Availability

Bank of Valletta, a leading financial services provider in Malta, uses Oracle Database 11g as the foundation for its critical customer-facing and back-office systems. The bank deployed Oracle GoldenGate to enable rapid recovery in the event of an outage at the production site, where Bank of Valetta runs its core business applications. The technology enables the bank to restore customer access to all services within 20 minutes of a failure. Oracle GoldenGate and Oracle Tuxedo are also embedded in third-party applications to ensure continuous uptime and high performance for the bank’s online customer services. The dynamic scalability of Oracle GoldenGate enables customers to withdraw cash, make card purchases, carry out internet banking, and shop online with good response time, even during peak times. Bank of Valletta achieved greater than 99% service availability for the
applications running on Oracle Database as well as for the third-party applications such as BASE24 payment engine, into which Oracle GoldenGate is embedded. The bank also achieved greater than 99 percent service availability for its business intelligence platform.

BT Maximizes Uptime during Migration

British Telecom (BT) deployed Oracle GoldenGate to safely migrate its business critical system without user impact, reducing downtime by 90 percent. BT Group plc is one of the world’s leading communications solutions and services providers, with operations in 170 countries. BT’s network inventory planning and service provisioning application handles 40,000 transactions per minute and is key to providing end-to-end services to customers. This application stores the entire organization’s network inventory information, such as circuits, routers, switches, and locations, and is a single source of truth for other interfacing components, such as network event monitoring systems, business intelligence systems, and data warehouses. Because this application is so crucial to BT’s customer delivery environment, any system outages could severely impact customer-facing services and troubleshooting capabilities. BT maintained system availability for 40,000 transactions per minute and 2,000 concurrent sessions during the migration of HP-Tru64 to Oracle Linux. With Oracle GoldenGate, the company avoided more than 20 hours of estimated system downtime and completed the migration in less than 2 hours.

Real-Time Analytics for Trusted Information

Timely data is essential for operational decision-making. As the latency increases, the data becomes less relevant and the ability to act on it decreases, potentially causing a company to improve operational efficiencies or to add more revenue. Real-time data integration ensures that the data used in decision-making is as current as possible—in some cases, that the analytic information is virtually in sync with operational information.

Traditional BI/data warehousing solutions give users an excellent view of past events and entities—often by accessing historical data. These solutions typically monitor slowly moving trends, such as product defects or sales performance. Users gain insight into these activities by querying a data warehouse, which is updated periodically.

Historical analysis is important, but it mainly targets strategic, long-term decision-making by analyzing trends, and it doesn’t let people respond to operational events happening in the organization. As data latency increases, its relevance to operational events diminishes. BI solutions for day-to-day operations need to leverage the most current data to enable employees to take action right away. Data integration solutions that leverage heterogeneous change data capture (CDC) technology can establish real-time links to production data sources rather than—or in addition to—performing bulk updates. These solutions work efficiently by moving only the changed data to the decision support system.
Enabling Real-Time Data for Analytics

Oracle’s strategy for real-time analytics incorporates the real-time data replication capabilities of GoldenGate in conjunction with other technologies including transformation (operational and analytical data), data federation, and the ability to apply data quality rules as part of the integration process. With the latest release of Oracle Data Integrator 12c, customers can deploy Oracle Data Integrator with Oracle GoldenGate easier, and achieve fast time to market for real-time data integration solutions. Oracle Data Integrator works with Oracle Enterprise Data Quality to ensure that the data in your analytics environment is current, consistent and of high quality. Oracle ties these integration solutions into its analytics tools, performance management tools, data discovery tools and in-memory analytics.

Oracle data integration tools are often used with Oracle’s business intelligence products to feed real-time reports, dashboards, and analytic applications. The new release of Oracle BI Applications 11g, release 11.1.1.7.1, uses Oracle Data Integrator to integrate data end-to-end across the full BI Applications architecture. Oracle Data Integrator supports capabilities such as data-lineage which helps business users identify report-to-source capabilities. Leveraging the power of ODI, the new release of BI Applications enables customers to increase IT efficiency and reduce costs with a comprehensive data integration platform that covers all data integration requirements – including big data, application integration, as well as BI / data warehousing. In addition, Oracle GoldenGate’s real-time data integration offering is an optional capability for Oracle BI Applications users to allow analysis with timely data non-intrusively. GoldenGate log-based change data capture technology does not impact application performance or require application modifications. Oracle GoldenGate option is only available with the version 11.1.1.7.1 of Oracle BI Applications which includes Oracle Data Integrator for extract, load and transformation.
MegaFon Centralizes Real-Time Billing Data

MegaFon is the only Russian telecommunications provider with a network that spans all of Russia, as well as the Republics of Abkhazia, South Ossetia, and Tajikistan. The company was the first in the country to launch 3G commercial operations. It is now Russia’s top provider of mobile internet solutions, and is ranked second highest for the number of active communications subscribers. Building on its growth, MegaFon recently acquired Synterra, a Russian mobile carrier. Following the acquisition, the acquired company’s billing information was in eight separate regional billing systems across Russia. To sustain growth in the multiple fixed and wireless segments, MegaFon needed to ensure accurate, trusted, and timely data for all corporate departments. The company deployed Oracle GoldenGate to extract billions of monthly transactions from eight regional billing systems. The data was integrated and centralized onto Oracle Database 11g Enterprise Edition and distributed to business-critical subsystems for revenue, fraud, and security analysis.

Oracle GoldenGate also enables MegaFon to create sophisticated, targeted marketing campaigns based on enhanced mobile subscriber information, maximizing market penetration and facilitating continued growth. With real-time access to key performance indicators, field offices and data centers can better meet their strategic goals. In addition, the solution enables more business-specific analysis of billing data—such as revenue assurance, customer analytics, and fraud detection—to accommodate revenue growth, customer insight, and security requirements.

Next-Generation Analytics with Oracle Engineered Systems

Oracle Data Integration solutions are tightly integrated with all Oracle Engineered Systems, including Oracle Exadata, Oracle Exalytics, and Oracle Big Data Appliance.

Oracle Big Data Appliance provides an integrated platform for big data based on an engineered system that has been optimized for acquiring, organizing, and loading unstructured data into Oracle Database.
This engineered system combines optimized hardware components with new software solutions to deliver the most complete big data solution on the market.

The Oracle Big Data Appliance is pre-configured with hardware and software to optimize any big data environment. It comes loaded with NoSQL, Cloudera Hadoop, Oracle R, Oracle Loader, Oracle Data Integrator, and the HDFS Connector. Oracle Big Data Connectors can stream data into an Oracle Exadata Database Machine at 15 TB/hour.

As described in the earlier section on big data, Oracle Data Integrator offers a unified design environment for developing MapReduce transformations, as well as transformations for structured data. It supports Hive-based transformations for running big data transformations, and integrates directly with Oracle Loader for Hadoop to load data into and out of Hadoop environments.

Data Integration for Oracle Exadata

As organizations contend with escalating demands for greater quantities of information, Oracle Exadata makes their database workloads faster, easier, and less expensive. This popular database machine enables superior business decision-making by allowing analysts to consider more possibilities and do deeper analysis more quickly. Normally data must be extracted from an online transaction processing (OLTP) system and loaded into a separate data warehouse for reporting and query processing, consuming valuable resources and requiring the IT staff to maintain two separate database systems. Oracle Exadata combines these two types of database activities into one consolidated system so organizations can run their business intelligence operations against live OLTP data.

Oracle provides a data integration solution that is optimized for this industry-leading engineered system. It meets the needs of BI, data warehousing, and big data, characterized by the following industry-leading advancements:

- Optimized bulk loads that leverage the performance of Exadata’s Infiniband network components and Smart Storage infrastructure.
- No need for time-consuming data transfers to and from a separate ETL server.
- Transformations that use Exadata-optimized SQL.
- Native batch and real-time integration capability.
- Zero downtime migrations and consolidation

Data Integration for Oracle Exalytics

Oracle Exalytics In-Memory Machine is the industry’s first high-speed engineered system featuring in-memory business intelligence (BI) software and hardware to deliver extreme performance for analytic and performance management applications. With Oracle Exalytics, organizations can deliver custom and packaged analytic and performance management applications that run faster, are easier to use, and support more users than ever before. With the engineered system organizations can complement their dashboard and reporting projects with scenario modeling, planning, and forecasting all running in a single environment.
An open solution, Oracle Exalytics is suited for use in heterogeneous IT environments and can access and analyze data from Oracle or third-party relational, OLAP, or other data sources. Oracle Data Integration technology supports real-time and batch integration of data from IBM DB2, IBM Netezza, Microsoft SQL Server and Analysis Services, SAP Business Information Warehouse (BW), Sybase Adaptive Server Enterprise (ASE), and Teradata Warehouse, among others, in any combination.

Data Warehousing using E-LT Architecture

The most efficient data integration tools use an E-LT architecture, where transformations take place in the target database. These tools do not require a middle-tier server to execute the transformation logic, like traditional ETL tools do, which means they require less hardware, administration, and energy costs to run. E-LT solutions offer IT architects a choice as to where transformations are performed—either on the target or on the source destination—allowing for greater flexibility, improved scalability, and greater performance. E-LT approaches can also reduce IT infrastructure costs. Because the target system has a powerful database engine and transformation engine available, it makes sense to utilize this resource. The goal is to move data as quickly as possible into that target data warehouse system and then use the capabilities of that engine to do the transformations.

E-LT scenarios are often a key component of real-time data warehouse scenarios. For example, if it takes eight hours to run an ETL process, an organization is introducing unnecessary latency into its operational reporting system. Instead, with Oracle Data Integration products, the latency can be reduced to a few minutes. Oracle GoldenGate reads online transaction logs as transactions commit in the source system and captures the data without impacting the source system. It “trickle feeds” changed data on the source system with very low overhead and latency. Oracle Data Integrator is integrated with Oracle GoldenGate and can use GoldenGate’s non-invasive log-based change data capture capabilities before the transformation step within the target database.

Turkcell Consolidates IT Environment with E-LT Technology

Increasing the efficiency of network operations that impact its data warehouse was a significant challenge for Turkcell, Turkey’s largest GSM carrier, with more than 34 million subscribers. The company has continuously increased the variety of its services based on mobile, audio, and data communication while increasing quality levels, boosting its total number of subscribers. By deploying Oracle Exadata, Turkcell decreased the number of server racks in its data center from 11 to one. It went from 250 TB of uncompressed data to 25 TB of compressed data and achieved 10-fold performance increases in query response times. Also, by moving from a traditional ETL architecture to an E-LT architecture with Oracle Data Integrator, Turkcell improved its ETL processing performance by 500 percent. With the new E-LT architecture the company has achieved 167 percent ROI after nine months and performance improvement as high as 20 times over previous statistics. The combined solution helped the company decrease data preparation time for data mining from 27 days to three days. Turkcell can better optimize its network operations by gathering information from the whole network infrastructure using Oracle Data Integrator.
Ensuring Data Quality for Trusted Information

Oracle keeps data quality as a central focus of integration projects. Oracle Enterprise Data Quality ensures that companies have trusted information in their data warehouse or BI system to improve business insight and reduce the risk of making decisions based on faulty information. Data quality is also an important element in governance, risk, and compliance initiatives. As shown in Figure 11, enterprise data is changing constantly, which makes it easy to inadvertently use outdated information. Bad data can have real, tangible risks because it cascades throughout the enterprise and can quickly pervade numerous systems, compromising all types of business initiatives. Oracle’s data quality technology identifies and corrects flaws in the data by automating data profiling, cleansing, matching, enrichment, and monitoring activities.

Oracle Enterprise Data Quality Profile and Audit
Oracle Enterprise Data Quality Parsing and Standardization
Oracle Enterprise Data Quality Match and Merge
Oracle Enterprise Data Quality Address Verification

Figure 12. Enterprise data is changing fast in all categories.¹

Oracle Enterprise Data Quality is an end-to-end enterprise data quality solution that enables customers to better understand, profile and analyze their customer data and product data. Recognized for its broad capabilities and ease of use for both business and IT users, Oracle Enterprise Data Quality is designed to work with Oracle Master Data Management and Oracle Data Integration, as well as create data quality services that can be accessed by any system in the enterprise. Oracle Enterprise Data Quality helps customers make sure their data is trusted, cleansed, and standardized and ‘fit for use’ in any application, data warehouse, or BI report.

Continuously monitoring data health is critical. In addition to industry leading profiling and cleansing capabilities, Oracle Enterprise Data Quality offers ongoing governance including case management and remediation capabilities as well as easy-to-understand executive dashboards for auditing, profiling, parsing, and standardizing process results. As shown in figure 13, it enables end users to see data quality trends over time and allows continuous monitoring and quick discovery of issues before they impact the business.

Oracle Enterprise Data Quality products include:

Five New Requirements for Data Integration

- Oracle Enterprise Data Quality Product Data Parsing and Standardization
- Oracle Enterprise Data Quality Product Data Match and Merge

Oracle Enterprise Data Quality executive dashboards display the results of ongoing profiling and audit processes.

BSI Improves Accuracy, Boosts Sales with Single Customer View

BSI (British Standards Institution) is the United Kingdom’s National Standards Body and the originator of many of the world’s most commonly used standards. The company works with more than 64,000 clients in 150 countries. BSI used Oracle’s Enterprise Data Quality products to create a single, accurate, complete record of each of these clients in just one month.

The project began with a simple recognition: BSI wanted to optimize customer insight by creating master customer records that captured each client’s profile, purchasing history, relationships, and other attributes in a single view. The goal was to eliminate inaccurate, incomplete, nonstandard, multi-format, and duplicate customer and transactional data from the customer database, which was growing three percent to four percent each year. The need for standardization also extended to the publications, training documents, tools, and services that BSI sold online. The company needed to ensure consistent coding, description, and pricing formats in its electronic catalogue. Always conscious of costs, BSI wanted to complete this master data management (MDM) project using its existing staff resources, despite a 20% year-over-year increase in data volume.

BSI chose Oracle Enterprise Data Quality, a suite of automated cleansing, matching, and profiling solutions, for its intuitive functionality, adaptability, and value. IT professionals used the software to aggregate more than 5 million disparate data records, held in multiple databases and formats, into a set
of “golden” customer records that span a four-year period and include transaction histories of all products and services. The results have been impressive:

- BSI used the Oracle technology to deliver a single customer view to 2,000 workers in 60 countries, while enforcing best practices in data governance and management within the organization.
- BSI eliminated inaccurate, incomplete, nonstandard, multi-format, and duplicate customer and transactional data for a client database that is growing 3% to 4% annually.
- Better marketing campaigns have improved sales by boosting cross- and up-selling opportunities and, in turn, providing a more complete customer experience.
- BSI now has standardized product categorizations and consistent coding and descriptions to accommodate a 20% expansion in data volume each year, without adding staff.

As a result of this project, the accuracy of BSI’s customer and corporate data has improved to nearly 100%, and the company can refresh information four times faster.

Conclusion

Businesses are demanding more from their data as they struggle to deal with the deluge of information that’s presented to them. To differentiate themselves from competitors and improve business results, they need to turn their data into a strategic asset. Achieving this goal not only means more pervasive BI and analytics, but also more powerful and flexible data integration technologies that can capture real-time data, transform it into a usable form, and make it available both on-premise and in the cloud. In today’s 24/7 business environment, uninterrupted access to critical information is a key requirement.

Oracle provides all the technologies needed to deliver pervasive, continuous access to trusted data. Oracle Data Integration 12c enables businesses to leverage any data, any source, at any latency—on-
premise or in the cloud—with zero downtime and maximum performance. Businesses can trust their data’s quality and timeliness when making strategic and operational decisions. Oracle Data Integration 12c products are designed to simplify deployment, increase productivity, and deliver higher performance while providing organizations with solutions that can easily keep pace with new data-oriented technology trends, such as cloud computing, big data analytics, and real-time data replication. No other vendor offers such a complete set of data integration capabilities that are optimized for performance, reliability, productivity, and scale.
Five New Requirements for Data Integration

January 2013

Authors: David Baum, Irem Radzik, Dain Hansen, Brad Adelberg

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

Copyright © 2013, 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 of The Open Group. 0113

Hardware and Software, Engineered to Work Together