Executive Summary

For companies to navigate turbulent business conditions and add value to their products, customer relationships, and business partnerships, they need timely, accurate, and relevant information. The more responsive and agile a company is, the better it can respond to globalization, mergers and acquisitions, regulatory compliance, fierce competition, tight operating budgets, high customer expectations, and demands for ever-faster product delivery. To speed interactions between systems and achieve the agility they need to respond effectively to powerful market forces, companies must be able to integrate data enterprise-wide in real and near-real time.

The Real-Time Enterprise Enables Business Agility

For decades, organizations have supported their operations with a host of IT systems, many of them specific to division or departmental needs. As market conditions and business processes become more complex, though, these systems can no longer operate in siloes. The company’s success depends on their ability to interact smoothly and in a timely way.

The Real-Time Enterprise ensures that smooth, timely interaction by adopting business and IT solutions and best practices that always give critical business processes access to accurate, trusted, available information about products, customers, and partners. This allows the Real-Time Enterprise to react faster, better, and more cost-effectively to evolving customer demands, changing market conditions, and emerging competitive threats.

The Information Latency Continuum

Different business processes have different latency requirements for information. Analytical data integration for projects such as business intelligence reports based on historical data has latencies ranging from days to weeks. Operational data integration, on the other hand, demands the delivery of information within hours, minutes, or seconds. Examples of IT projects that require operational data integration include real-time data warehouses, operational data hubs, data synchronization and data replication projects, and data services or data virtualization as part of a service-oriented architecture strategy.

The success of data integration projects typically depends on the ability to meet service level agreements related to data latency, data completeness, and data accuracy. Following are questions to consider when planning a data integration project:
• Is the data integration project primarily for analytical or operational business purposes?
• Does the data integration project require data delivered frequently (e.g., real-time or near-real-time) or infrequently (e.g., weekly batch windows)?
• Does the data integration project require point-of-entry data cleansing to ensure quality and avoid propagating bad data to downstream applications?
• Does the data integration project require access to a variety of data sources, including external data, and advanced data transformations to ensure completeness?
• Does the data integration project require movement of large data volumes or small datasets between applications?
• Does the data integration project require ready access to the most current data available in operational systems?

Figure 1. The Information Latency Continuum describes the range of information latency requirements for analytical and operational business processes.
Data Integration and the Real-Time Enterprise

Timely, trusted information is essential to sell products, service customers, and source from partners in today’s global markets. To use this information in critical systems and processes, organizations must be able to access any data source at any latency and of any size, integrate the data, cleanse it of duplicates or errors, and deliver it to any consumer channel in real time or near-real time. However, as the volume, velocity, and variety of data continues to grow, integrating, cleansing, enriching, and delivering it across an equally expanding number and variety of channels such as portals, web services, and applications is becoming increasingly complex.

Traditional approaches to data integration, such as enterprise application integration or hand-coding, are time-consuming, costly, and difficult to build and maintain. Furthermore, they fail to address issues of data quality and data governance. The Real-Time Enterprise needs an alternative that supports all of these typical data integration projects and initiatives:

Real-time Data Warehousing and Operational Reporting

A real-time data warehouse is typically a mix of batch data integration processing and real-time integration processing. Real-time integration processing must generally meet two key requirements. First, it must be able to deal with very large amounts of data movement. When an operational data store (ODS) is loaded, data is consolidated and aggregated in the data warehouse (DWH) to provide a consistent forecast and history of business operations. Second, real-time integration processing must take action based on information analyzed in real time or near-real time.

For large volumes of data, loading an ODS once a night, along with complex data aggregation to load the data warehouse, may not be enough. It is therefore much better to smooth the ODS loading throughout the day. Loading an ODS in real time also provides a better recovery mechanism in case of failure, since in traditional large data warehouses, if an error occurs during the ODS or DWH load, it is difficult to reload during the batch windows.

“Using Informatica’s Pushdown Optimization with Teradata allowed us to combine wireline and wireless information into an enterprise data warehouse, meeting the SLA established by our users. The daily load time for the enterprise data warehouse was decreased by 50 percent, from twelve hours to six. The ultimate result will be a unified view of our customers. Adding Informatica’s change data capture (CDC) for ADABAS helped us to deliver, within a very tight timeframe, customer information to the fraud agency. These accomplishments would not have been possible without the significant performance capabilities offered by the Informatica and Teradata platforms.”

- Operations and Architecture Manager, Largest Telecommunications Company in Brazil
Data Replication

Data replication physically copies data between systems. For example, a data replication project may copy relational database (RDBMS) tables from a source system to another RDBMS of the same or different type.

Data replication must be able to move data continuously from one system to another as transactions occur in the source system. It must also be able to empower business decision-making by providing quick access to the freshest data available in operational systems without impeding the performance of those systems. This calls for database-independent, heterogeneous, real-time transaction replication software that is highly scalable, reliable, and easy to use without disrupting operational performance. It must provide access to the most current data regardless of the complexity and diversity of the IT landscape. Most of all, it must work in concert with the core platform to meet all the data integration, data movement, data quality, and data currency needs of the organization, such as operational reporting or operational synchronization.

Data Synchronization

Data synchronization projects are a type of data copy in which the data copied may not be easily accessible and may require some transformation, conversion, and additional information to fit the target system. Data profiling prior to data movement is a key step in minimizing inconsistencies in the data while streamlining development efforts. Most projects transfer data either one way or bidirectionally between one or more applications.

An enterprise-grade data synchronization solution can enable your IT organization to achieve accurate and consistent data across operational and transactional systems. With such a solution, your IT organization can quickly build business logic that handles the most complex data synchronization projects and reuse the same logic across projects in batch, near-real time, and real time. This results in high performance, greater accuracy and consistency, and a lower total cost of ownership. Operational Intelligence

“Right-time data synchronization and cleansing is a key component of our migration strategy to a responsive and agile architecture. Informatica delivers right-time synchronization of trade and account data between our legacy mainframe systems and newer transactional systems to support the high performance requirements of one of the busiest options exchanges in the world. Informatica enabled us to implement the project in half the time than alternative methods.

- First Vice President, Derivatives Trading Systems Development
  Back Office Enterprise Architecture Development, Major Stock Exchange
Operational Intelligence

It is becoming increasingly important to respond to business events as they arise, or anticipate certain events before it is too late to react. Capturing changes in a number of environments as they occur enables your IT organization to deliver up-to-the-minute data to the business. However, traditional EAI solutions only capture selected portions of data published by predicted transactions. In many cases, data changes are not predictable, resulting in data irregularities and inconsistencies. Moreover, traditional data integration only captures data from source systems, cleanses and transforms it, and delivers it to a store. Any analysis for operational decision-making usually occurs offline with the aid of business intelligence tools.

By adding complex event processing and data integration, it is possible to create an event-driven architecture that automates the entire process of capturing, correlating, and analyzing data in real time. Organizations can then respond rapidly and intelligently to emerging opportunities and threats the data reveals.

Real-Time Data Quality

Data quality—ensuring that data is accurate, consistent, and complete—is critical to operational efficiency and agility. However, when data enters a database through multiple automated and manual entry points, the organization cannot be sure of its quality. That situation leads to a need for repeated data cleansing across databases, at the point of entry, and every time new data is acquired or integrated.

The alternative is to build data normalization, validation and data quality rules once and apply them automatically in real time whenever data is created, acquired, or integrated. This ensures that data meets the standards of the business, reduces the amount of effort required to continually address data quality, and drives down the cost and effort of operational development efforts. Typically, this real-time information is used in conjunction with master data management platforms to deliver consolidated and reliable business-critical data to critical applications.

“Informatica gives Zyme the ability to maintain our competitive advantage by helping us deliver timely and accurate data to our customers. Using Informatica, Zyme is on-boarding EDI feeds faster than before—freeing up development resources to address other important tasks.”

- Senior Vice President, Operations, Zyme Solutions
Multi-Enterprise Data Integration

As enterprise infrastructures and business processes move to hybrid models, data becomes increasingly location-independent and resides inside and outside the firewall, including in business partner, channel partner or supplier systems, in SaaS applications, or on devices. As a result, enterprises must streamline how, when, and where they share data externally. A more centralized and flexible approach to connecting, managing, monitoring, and transforming data with external partners eliminates operational silos. As a result, the Real-Time Enterprise can onboard new external data sources faster, manage a broad range of business partner technical capabilities, standardize the onboarding process and shift it to non-development resources, and reuse automated rules to minimize redundant integration development and reduce maintenance costs.

Operational Data Hub

A typical application-to-application data integration scenario involves numerous files and large sets of data or small payloads moving between applications in real time or in batch. This scenario often uses a constantly changing array of complex, proprietary hand-coding and point-to-point interfaces that do not support all the formats of files or data, have no central point of control, and provide minimal visibility into the movement of information. Maintenance is a nightmare, there are multiple points of failure, and the lack of control over data quality as it propagates across systems leads to inconsistent business processes.

The alternative is a data integration hub, an architectural pattern that supports internal and external information exchange among applications. Ideally, the data integration hub leverages a sophisticated data integration platform and can support self-service, auditing, data quality, and the ability to integrate data at any latency.

Message-Driven Data Integration

Today’s CIOs and enterprise architects face an exponential growth in event data, as well as the need to distribute and interact with that data more frequently in real time to accelerate business operations. Customers also demand more proactive alerts to issues, real-time awareness of situations, and insight into orders and other business objects. Finally, business users want more fine-grained tracking of events within their operations and among business partners to be more proactive in the face of more complex and interconnected processes.

To achieve all these goals, organizations need modern peer-to-peer messaging middleware that will support event-driven architectures for many kinds of applications, including—but not limited to—those identified as involving big data. Not to be confused with point-to-point messaging, peer-to-peer messaging leverages a distributed architecture similar to the “shared nothing” architectures used across the Web and in many of today’s massively scalable database systems for OLTP and data warehousing.
Informatica offers solutions that meet all of these needs, providing a highly available and scalable enterprise data integration platform for accessing, discovering, cleansing, and integrating data from virtually any business system, in any format, and delivering that data throughout the enterprise at any speed. With Informatica, enterprises can integrate and provision operational data in real time with high-performance data integration software. Informatica supports all types of business processes that require real-time or near-real-time agility, including cross-selling by customer service, point-of-sale fraud detection, real-time insurance policy quotations, SEPA and AML regulatory compliance, personalized best offers for marketing, supply chain optimization, retail out-of-stock replenishment, straight-through trade processing and reconciliation, 24x7 global operations, and many more.

With Informatica, you can support the entire data integration lifecycle for all your analytical and operational IT projects. Informatica has helped thousands of companies successfully overcome the challenges of global uncertainty to improve operational efficiency, better serve customers, improve communications with trading partners, and minimize exposure to risk. Informatica provides products and services to cost-effectively access and deliver timely and trusted data in support of both analytical and operational business processes, enabling the Real-Time Enterprise.
About Informatica

Informatica Corporation (Nasdaq:INFA) is the world’s number one independent provider of data integration software. Organizations around the world rely on Informatica to realize their information potential and drive top business imperatives. Informatica Vibe, the industry’s first and only embeddable virtual data machine (VDM), powers the unique “Map Once. Deploy Anywhere.” capabilities of the Informatica Platform. Worldwide, over 5,500 enterprises depend on Informatica to fully leverage their information assets from devices to mobile to social to big data residing on-premise, in the Cloud and across social networks. For more information, call +1 650-385-5000 (1-800-653-3871 in the U.S.), or visit www.informatica.com.