Complex Event Processing

Relevant, Timely and Trustworthy Decisions for End-Users

Table of Contents

Introduction

Organizations are challenged to intelligently respond to opportunities and threats represented by the ever-increasing velocity, volume, and variety of informational events—occurring real-time, both inside and outside the corporate walls. Situational awareness, cost containment, risk mitigation, and growth: these operational imperatives require timely detection of and response to key business events. Organizational success and survival, compounded by economic, political, social, and competitive pressures, mandates decisiveness, responsiveness, and accountability.

For an organization to be able to effectively meet these challenges, they must adopt solutions that can be leveraged across the enterprise—empowering all employees. Organizations must be able to detect events as they occur instead of hours, weeks, or even months later. Users must be provided intelligent and contextual information in real time in order to make timely decisions. Users need flexibility to determine what events are monitored, how events are analyzed, and how responses should be activated. And, all of these capabilities must minimize the impact on existing systems and processes.

Events such as new sales orders, a series of credit card transactions, a customer service request, several changes in geographic position, or an updated investigative report represent the foundation of new and changing data throughout an organization. These events must be continually monitored to identify opportunities and threats. Individual events, if not already detected by existing business systems, can be quickly harnessed using traditional Event Stream Processing (ESP) techniques. However, complex events, comprised of related individual events, traditionally go undetected by both users and existing systems. These events may have internal and external informational characteristics, and may occur at different times spanning geographic locations.

Below are two real-world examples of normally undetected complex events, along with potential responses:

Example 1 – Vessel Tracking

Scenario

A cargo vessel departing the Middle East will enter a U.S. port in 24 hours carrying a passenger whose name is on a watch list and transporting a questionable cargo item.

Complex Events

Identifying the appearance of a high-risk vessel based on data derived from the vessel's position, travel history, cargo, and passenger manifest.

Response

Alert appropriate watch officers who may request interdiction of the vessel prior to its entry into port so that the crew, passengers, and cargo can be properly inspected.

Example 2 – Transaction Monitoring for Charge Card Programs

Scenario

Three separate credit card transactions of $2,100, $2,000, and $1,700 by the same individual occur across a 24-hour period with the same vendor.

Complex Events

Identify the transactions as waste, fraud, or abuse based on the total transactions exceeding pre-set spending limits, occurring within a 24-hour period, and conducted with the same vendor.

Response

Alert the appropriate manager of the potential issue while providing supporting information in order to discuss the questionable transactions with the employee.

Today, complex events like the examples above are not typically detected in a timely manner. The high-risk vessel may proceed directly to port, exposing citizens or infrastructure to potential danger. The corporate card program manager would be forced to manually identify the potential fraud, waste, or abuse—often weeks or months after the occurrence.

Intelligently monitoring these events across windows of time, over geographic boundaries, and across stove-piped systems or domains demands user-friendly complex event processing (CEP) capabilities.

What's Driving Complex Event Processing?

What's driving CEP, or as some may call it, operational intelligence, real-time business intelligence, or business event processing?

The Information Explosion Continues

In recent years, the quantity, quality and access to information have increased exponentially. These increases have largely been driven by the emergence of efficient systems that record, track and monitor events, as well as communications mediums such as the Internet and wireless technologies that deliver this information to end users, not only in business and government, but also to consumers.

Events, whether system-generated or human-generated, are pervasive and ever increasing in their volume, velocity and variability. In the commercial world, robust software packages track business events such as orders, shipments, payments and other transactional data for record keeping and compliance purposes as well as for improved decision making. In the government world, similar systems are used for these purposes as well as for critical and time-sensitive decision making regarding national security. In the consumer world, online and mobile applications deliver event information concerning news, entertainment, transportation, and safety.

Moving Beyond BI, SOA, and EDA

While this explosion of event-driven information and simplified information availability has provided significant benefits, the sheer amount of information has created issues with managing, manipulating, and analyzing such fragmented data, often paralyzing the decision making process, forcing erroneous assumptions and limiting the utility of the information. As a result, several tools have been developed to accomplish the task of connecting disparate systems, analyzing and processing complex information, and delivering this information to decision makers. Data warehouse and business intelligence (BI) systems support decision making by increasing the ability to drill down and produce effective historical reports.

The adoption of Service-Oriented Architectures (SOA) coupled with Event-Driven Architectures (EDA) promise faster implementation time of solutions while breaking apart silos of information assets. While SOA and EDA solutions are used by IT professionals to enable better connectivity, BI software is dominated by tools that query and move massive amounts of data, often populating pre-built reports to provide historical analysis of particular aspects of a business or situation. While creating reports serves certain business requirements, traditional BI systems are impeded by latency and an inability to perform on-the-fly correlations and persistent, around the clock, analysis. These limitations render such tools ineffective in many situations that require dynamic analysis and decision making in real time based on information from disparate internal and external data sources.

The Need for User-Oriented Persistent Analytic Tools

The task of analyzing, processing and delivering actionable information is further complicated by the disparate needs of users in various industries, business units, and markets, as well as the varying need for more updated or real-time information, and the fact that crucial decisions in business and government are based upon multiple events that often occur independently of one another. Consequently, it is extremely difficult to identify, in real time, related events that signify emerging opportunities and threats. Furthermore, the complexity of analyzing such events grows exponentially with added events. For example, a large withdrawal from a bank account is not notable in isolation, but when combined with one or more other factors (a similarly large deposit made immediately prior, the lack of such activity in the past or the name on the account being the same as a fugitive recently in the news), it becomes highly suspicious and must be addressed—ideally within minutes or seconds as opposed to weeks or months. The need for flexible tools that can assess such events from a multitude of sources, process them and deliver the actionable results to decision makers has become apparent across many commercial and government applications.

THE AGILE ORGANIZATION MUST BE ABLE TO ACCESS, ANALYZE, AND ACT ON INFORMATION BY PROVIDING:

- Real-time persistent analytics

- End-user control

- Actionable intelligence

- Enterprise-wide collaboration

- Flexible technology framework

As these needs and the lack of robust solutions to address them have become more recognized, complex event processing (CEP) has emerged as a meaningful category of software with broad applicability across commercial and government users.

CEP software identifies correlations and establishes connections among events, then initiates triggers or responses when a trend, cross-source correlation, or other event emerges. CEP allows organizations to seize opportunities and mitigate risks in real time. New and changing data across an enterprise, and often external to an enterprise, must be continually monitored to identify opportunities and threats. Complex events comprised of contextually-related individual events have traditionally gone undetected by both users, due to information overload, and by existing systems, due to a lack of real-time cross-platform processing capability. These events may have internal and external informational components and may occur at different times. Software that can make sense of the flow of events from various data sources, within time windows and geospatial boundaries, while allowing non-technical end users to personally define what is important to them is critical to achieving real-time intelligence, as well as to obtaining tactical and operational superiority.

Most CEP solutions available today have been custom-built for specific uses, and typically require significant involvement of the IT organization and developers to implement, deploy and customize. As such, while most CEP solutions add significant value beyond classic business intelligence and other tools, these solutions are limited by a lack of flexibility and ease of use for the ultimate end users—those outside the IT department.

Core Attributes of Complex Event Processing

Over the last ten years, Informatica has worked with its customers to meet unique event processing needs for end users. During this time, our experience has led us to conclude that an effective CEP solution must possess several key attributes to meet the needs of an organization—beyond the technical experts of the IT organization. These attributes include the following.

1. Rules must automatically detect complex business events in near-real time, from any information source, without compromising security or data integrity.

2. Processing of live data events must be correlated and conditionally analyzed with many information sources including internal and external data feeds, data warehouses, or other live data sources and sensors—the CEP solution must be able to receive and retrieve data.

3. Prior to a rule triggering an alert or initiating a process, there must exist the ability to perform secondary analysis or correlation in order to extend knowledge about detected events and to discover additional facts to ensure a complete picture of event details—especially when deployed to address high-risk situations.

4. Responses to detected events must be automatic, targeted, specific, and contextually appropriate. This includes delivering information-rich alerts to authorized subscribers, automatically triggering business processes, and launching existing analytic tools.

5. Rule sets used for ingesting events, detection and correlation, secondary analytics, and response generation must be constructed and maintained using a graphical interface. This interface must be usable by business users and domain experts, and not require IT personnel for changes and additions to business rules. In addition, end-user management facilities must be available to allow users to share rules or define personalized domain-specific rules.

A robust CEP engine must serve as the core for detecting and responding to events, and it must be capable of using detection and correlation rule sets in real time as they are defined by users.

As events are asynchronous in nature, the CEP engine must operate through the use of server-based processes generated from rule set definitions. Additionally, CEP capabilities must be deployed without changes to existing systems and processes, without duplication of sensitive data, and without compromising information security.

Key Features of Enterprise CEP Solutions—For Users

Complex event processing must address an organization's true need for speed and agility in decision making. Widespread deployment of CEP capabilities is only possible if the solution is focused on automating and extending intelligence gathering and decision making for end users (analysts, watch officers, managers, and operators), not just programmers. Traditionally, employees tasked with identifying events to forecast impending threats or opportunities must deal with reporting latency, inaccessible information, or turnaround time from IT staff to implement custom solutions that are often outdated when, or if, delivered. The consequences of missed opportunities or threats are often determined in seconds or minutes. The need for user-driven real-time persistent analysis of data is apparent and, as a result, specifies that the following features be offered.

User-Driven

Non-technical end users must have the ability to configure rules on-the-fly, specific to their mission needs. Dynamic rule creation allows users across organizations to quickly adapt during critical situations when presented with new information that redefines what events constitute opportunities and threats. User-defined rules must leverage temporal analysis capabilities (e.g., alert me when X and Y occur more than three times within a time window), geospatial analytics, event feeds defined by other users, a diverse response library, and other external data services. The CEP software must make it possible for an end user to configure how his or her own rules should interact with any number of data sources, analytic routines, and response capabilities.

Flexibility

The CEP software must not merely exist as a point solution. Instead, it must be implemented (and simultaneously used, as appropriate) across a variety of organizations, each having unique sets of requirements. Whether it is providing complex correlation of high-interest vessel sensor data, or monitoring a bank's customer-facing external systems for signs of fraud, the software must be deployable within days or weeks, integrated with existing data sources and networked application services—internal or external.

Automated Data Enrichment and Knowledge Extension

Knowledge of detected events must be extended in real time through the correlation of known parameters such as people, geographic locations, email addresses, time, and other multi-source event data. This capability enables discovery of additional events, resulting in contextual, information-rich alerts, and allows actions to be automatically initiated from a more informed position.

Intelligent Response Management

CEP software must enable the ability to detect complex events, but also the ability to initiate complex responses based on detected conditions. For example, a user may define a response to send particular event attributes to a personal GIS tool and only a simple notification via Instant Message or a VoIP call in the event of a high-priority event.

Collaboration and Knowledge Preservation

End users must have the ability to share personal rules (and thus scenarios comprised of multiple rules), data source definitions, analytics and watchlists, allowing for the deployment of user-driven CEP systems that grow with an organization while fostering information sharing and the preservation of best practices and specific domain knowledge.

Robust Developer Framework

The CEP software must also provide extensibility by including well-documented application programming interfaces (APIs) that allow other user interfaces and applications to leverage the CEP engine's capabilities across IT environments and platforms without redevelopment.

Case Studies

Informatica's innovative approach to CEP focuses on automating and extending real-time intelligence processing and response for the end user. Informatica's software provides a solid foundation for user-defined event detection, data correlation, and enterprise alerting. The software is used throughout the Federal Government's national security community to perform critical processes such as fusion of multiple data sources, rule creation for tactical decision making, and homeland security-oriented threat identification.

Informatica's software addresses numerous commercial business requirements, such as Anti-Money Laundering (AML) automation, anti-phishing investigation automation, corporate payment card policy compliance, employee fraud monitoring, logistics optimization, and geospatial tracking. In all cases, Informatica provides end users with the power of continuous all-source analysis, heightened situational awareness and ultimately situational dominance when time-critical action is required.

The following case studies outline specific use cases of Informatica's CEP products Informatica RulePoint® and Informatica Real-Time Alert Manager™. Due to the extremely flexible and easy-to-use nature of Informatica's products, only minimal configuration and training is required to deliver results in the diverse uses described below.

Case Study 1 – Battlespace Situational Awareness / Command and Control

Challenge

Battlespace commanders, watch officers and analysts must be continually aware of battlefield information from many different sources to properly assess threats and target opportunities. With the massive amount of information and changes to keep track of, this process can be time-consuming. Worse yet, threats are easily overlooked or missed within an environment that demands real-time decisions and reactions. Relevant or related changes in battlefield information includes an overwhelming number of events and data sources including friendly and enemy troop movement, weapons activity, weather events and unattended CBRNE (Chemical, Biological, Radiological/Nuclear, Explosive) sensors. Due to the sheer volume of data, and both the velocity and volatility of changes to the data, these sources are extremely difficult to continuously track, process, and assess. For example, it is impossible for a person to manually detect and respond to a cross-correlation scenario when particular enemy troop movement has occurred within a particular geographic area more than three times, within a 24-hour period, and within 10 kilometers of friendly forces, or to manually detect and respond to a CBRNE event in conjunction with a forthcoming weather event. The only way to solve this in a timely and accurate manner is through automation.

Solution

Informatica RulePoint provides a comprehensive way to automatically detect actionable changes in battlefield information and receive instant intelligence. As event data flows into the software from multiple systems and sources, user-defined rules and analysis logic are applied on-the-fly. When relevant threats or target opportunities are detected using both geospatial and temporal analysis, RulePoint generates contextually enriched alerts and disseminates them to the appropriate people as actionable intelligence.

Results

The detection of battlefield threats or target opportunities is performed automatically with RulePoint. Based on these threats or opportunities, commanders receive personalized information-rich alerts that contain geospatial visualization, contextual action links to initiate a response to the threat, and access to supporting information that substantiates the detection and response scenario. This expeditious recognition of threats and opportunities decreases the time required to have clear and complete situational awareness—resulting in a more rapid response that can save lives and resources, as well as in rapid enemy target identification and response coordination. The volatility of these situations is exacerbated by the complexity of correlations across data sources and the time requirements necessary to initiate a successful response. User-driven CEP using RulePoint represents the only way to achieve the objectives. When world events and situations are changing every second, RulePoint provides the only means of instantly adapting to the intelligence and response requirements of commanders, watch officers, and analysts.

Case Study 2 – Transaction Monitoring for Corporate Card Payment Programs

Challenge

Businesses with corporate payment cards must be able to mitigate waste, fraud, and abuse associated with employee spending activity in real time, while also being able to increase the type and size of transactions settled through payment card programs. Manual reporting and ad hoc data analysis techniques are used to mine the massive amounts of data to identify transactions that are out of compliance with comprehensive and often-changing corporate policy. These methods are very time-consuming and labor-intensive—making the traditional management of charge card programs inefficient and error-prone.

Solution

Informatica's enterprise CEP software, RulePoint, gives corporate charge card providers a comprehensive way for their customers to dynamically detect non-compliant charge card use by employees. Through a web-based, self-service model that enables program managers to set threshold criteria, RulePoint allows end users to automate the analytic logic derived from the labor-intensive and error-prone manual compliance checks. RulePoint receives streams of transaction data from diverse systems and continuously, in real time, applies the manager's analysis logic on the data. This is in contrast to previous solutions that ran reports against a data warehouse on a weekly or monthly basis. Upon identifying non-compliant transactions, RulePoint generates alerts and sends them to the appropriate person for remediation.

After reviewing many options on the market, including those offered by the card associations, Informatica's client found RulePoint to be unique in its delivery of a broad, cost-effective solution to support its corporate payment card customers—each having similar yet unique policies and thresholds for employee card usage.

Results

Analysis is instant, comprehensive and precise with RulePoint handling persistent identification of non-compliant transactions. Alerts of non-compliance are delivered instantly to program managers, functional managers, and even directly to employees for immediate action. No longer must the organization wait a month or longer to read and digest a report of a wasteful or fraudulent employee before taking action. Transaction analysis logic can be adapted quickly and easily through web-based forms, without involving technical support, thus enabling companies to capitalize on the ability to drive savings across ordering, invoicing and payment. Leveraging Informatica's solution, one large, Fortune 500 customer expects a 20% reduction in fraudulent and wasteful credit card activities while forecasting $5 million in annual savings in overall processing costs.

Case Study 3 – Maritime Domain Awareness

Challenge

To facilitate timely, accurate decisions about threats within ports, harbors, canals, and across the open sea, analysts, watch officers and commanders need instant, comprehensive and precise actionable geospatial intelligence and situational awareness. With too much information to sift through and so many changes happening within the maritime domain, the task of collecting, fusing, analyzing, displaying and disseminating operationally relevant information becomes impossible for humans alone. Manual analytic efforts must be supplemented with automated analysis logic and correlated all-source data to achieve timely and persistent awareness across the ever-changing maritime domain.

Solution

RulePoint automatically detects geospatial maritime threats and efficiently invokes appropriate responses, alerts, and GIS tool updates. Through an easy-to-use web-based interface, RulePoint allows users to capture and automate analysis logic previously performed manually. RulePoint receives data, information and intelligence from legacy systems, databases, sensors, and other emerging technologies that track vessels, cargo, crew and passenger manifests, maritime geospatial areas of interest, ports, the environment, and maritime critical infrastructure, all of which in isolation or various combinations may represent threats and important activities.

RulePoint then correlates this data using the user-defined analysis logic as soon as the data is received. For example, an analyst creates a geospatial rule to alert her when a particular unknown ship comes within a certain distance of any number of ships already of interest. Other rules monitor ships carrying certain types of cargo, specific passengers, or suspicious movement over time. RulePoint persistently watches for these maritime threat scenarios—and when a threat is detected, it automatically generates appropriate alerts and actions containing

relevant information and intelligence, ultimately sending those alerts to the appropriate decision makers for remediation. Informatica customers have derived significant instant value by sending alerts into geospatial visualization tools instead of viewing all of their data at once, leveraging RulePoint to make sure that the only "dots and tracks" that the analyst has to look at are relevant to operational maritime tracking requirements.

Results

With RulePoint, decision makers receive instant intelligence and precise information, dramatically increasing the accuracy and completeness of information sharing, situational awareness and collaborative planning. For example, before RulePoint, many Informatica customers were able to display geospatial information such as the position of a ship on their visualization tools. Now they are able to see through the thousands of unknown blinking dots to the correlated, actionable geospatial events that require a response interdiction. Informatica not only delivers maritime situational awareness—but even more so, maritime information superiority.

Complex Event Processing and Informatica

For nearly a decade, Informatica has been focused on its mission of delivering easy-to-use enterprise software that continuously delivers timely and relevant intelligence for seizing hidden opportunities and eliminating imminent and unforeseen risks. Our products continue to evolve with emphasis on eliminating the time required to detect and respond to threats and opportunities by automating the analysis of data and delivering instant intelligence to an individual or user.

Informatica's fourth-generation CEP products allow analysts, managers, operators, investigators, and watch officers to monitor data in real time across diverse sources and cross-reference, filter and correlate multiple events over time to identify key threats and opportunities. Decision makers are provided with context-appropriate and prioritized alerts via multiple channels such as email, IM, and web browser, and automatic responses to data events are initiated as appropriate.

Informatica RulePoint

Informatica RulePoint is an enterprise-class, user-driven CEP software product that rapidly integrates into an organization's existing IT infrastructure to deliver real-time operational BI to key decision makers. RulePoint uses pre-packaged adapters (configurable interfaces that link to data sources) to receive information from relevant streams of data flowing from different data sources—even batch files. Each piece of incoming information is referred to as an event. These events are published into RulePoint and grouped into familiar subscription categories, referred to as topics. User-defined rules then reference one or more topics, apply conditional or advanced analysis routines to the event data derived from the topic, and ultimately activate other services to coordinate responses to detected events.

Most notably, RulePoint enables users to create rules themselves, ensuring that an organization's responsiveness to changing conditions is not hindered by the traditional IT software development cycle. RulePoint offers three web browser-based modes for creating rules: Wizard, Template and Advanced Mode. Wizard Mode offers a three-step point-and-click selection of topics, evaluation conditions and responses. Template Mode allows the most novice user to fill in only select parameters of pre-defined rules. Finally, Advanced Mode allows for natural language-like rule creation, with auto-complete assistance that guides the user through rule creation by making contextual syntax suggestions along with each keystroke.

RulePoint's rule creation syntax provides in-stream event-enrichment capabilities that can be reused and natively leveraged across rules. RulePoint enriches event data by executing on-the-fly database queries or web service calls and then performs further processing of the results, along with the original event data. The additional data can be used for correlation or to provide supporting information to the resulting response message.

Users write rules to respond to correlations within and across voluminous, complex streams of data and heterogeneous information sources. These correlations are nearly impossible for human beings to detect in real time, which has resulted in traditional reporting and historical trend analysis for post-mortem identification of threats and opportunities that have already passed. RulePoint detects and delivers intelligence in real time, enabling employees within organizations to respond to important events as they are occurring—to literally capture opportunities and stop threats as they occur.

Unlike other CEP and event stream processing (ESP) vendors whose tools are designed as development platforms for IT staff and programmers, RulePoint was designed as an application for true end users. Informatica developed RulePoint to meet the needs of customers in environments where real-time situational awareness is critical to the success of life-or-death missions, and as such, the RulePoint architecture more than meets the needs of enterprises. RulePoint differs from pure database-oriented solutions because it can access diverse, non-database-oriented data sources (such as RSS, web services, file systems, email systems, etc.), as well as receive real-time events from event producing sources through its web service API, TCP socket listeners, or inbound instant messages.
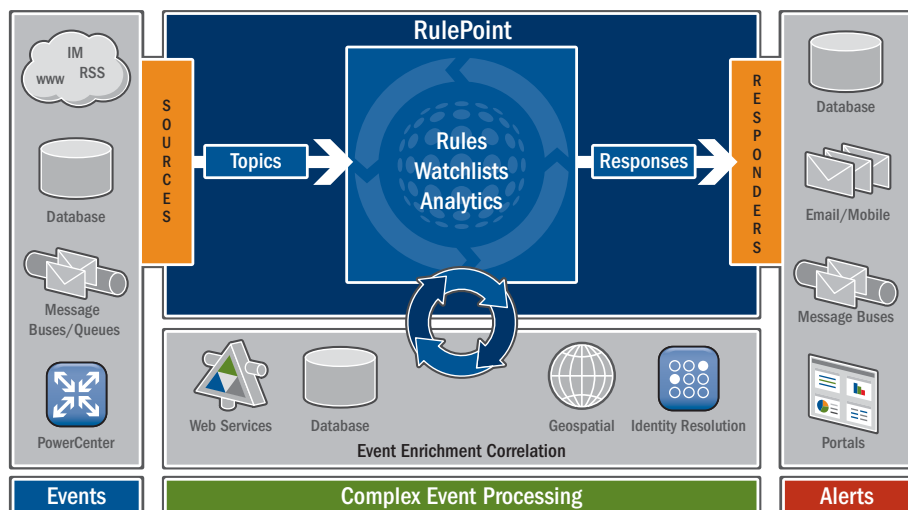

Figure 1 – RulePoint Architecture

Informatica Real-Time Alert Manager

Informatica Real-Time Alert Manager is a web-based application for managing alerts received from RulePoint. Real-Time Alert Manager provides a persistent communications channel for receiving and responding to notifications through point-and-click actions. Based on user-defined rules in RulePoint, Real-Time Alert Manager allows end users to easily sort, manage, share and prioritize alerts into dynamic channels, while limiting excessive messages. Executives and knowledge workers receive an overwhelming amount of information as part of their day-to-day responsibilities; thus, the need for a centralized view of actionable intelligence derived from real-time operations is more critical than ever. Real-Time Alert Manager allows users across organizations to monitor and act upon numerous alerts in a timely, persistent manner. Real-Time Alert Manager empowers the real-time organization by notifying appropriate individuals about key incidents or emerging issues, displaying aggregated data to enrich alerts and automating the prioritization, classification, and routing of alerts.

Summary

The sheer volume, variety, and velocity of events occurring across organizations and the globe is staggering. Organizations either fall behind or adapt and embrace technologies to operate in real time.

Informatica's software provides robust capabilities required to deliver effective complex event processing, while providing ease of use and flexibility that enables applicability of CEP far beyond traditional applications. Our products are ideally suited for environments where the ability to provide real-time event detection, personalized alerts, and automated responses determine mission success or failure. Due to the product's broad applicability and flexible design, it is easily configured to meet the specific requirements of clients in very different industries.

Learn More

Learn more about the Informatica Platform. Visit us at www.informatica.com or call +1 650-385-5000 (1-800-653-3871 in the U.S.).

About Informatica

Informatica Corporation (NASDAQ: INFA) is the world's number one independent leader in data integration software. The Informatica Platform provides organizations with a comprehensive, unified, open, and economical approach to lower IT costs and gain competitive advantage from their information assets. Nearly 4,000 enterprises worldwide rely on Informatica to access, integrate, and trust their information assets held in the traditional enterprise and in the Internet cloud. Visit www.informatica.com.