ORACLE®

FUSION MIDDLEWARE

An Oracle White Paper
March 28th, 2013

Big Data Gets Real-time:
Oracle Fast Data

ORACLE®

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.

What is Fast Data?

To Today, both IT and business users alike are facing business scenarios where they need better information to differentiate, innovate, and radically transform their business. In many cases, that transformation is being enabled by a move to "Big Data." Organizations are increasingly collecting vast quantities of real-time data from a variety of sources, from online social media data to highly-granular transactional data to data from embedded sensors. Once collected, users or businesses are mining the data for meaningful patterns that can be used to drive business decisions or actions.

Big Data uses specialized technologies (like Hadoop and NoSQL) to process vast amounts of information in bulk. But most of the focus on Big Data so far has been on situations where the data being managed is basically fixed—it's already been collected and stored in a Big Data database

This is where *fast data* comes in. Fast data is a term that was coined recently by Ovum's Tony Baer who said, "Fast Data, the velocity side of Big Data, is not new, but technology price/performance trends are making Fast Data applications more widely available". [1]

"Fast data is a complimentary approach to big data" writes Hasan Rizvi, "for managing large quantities of "in-flight" data that helps organizations get a jump on those business-critical decisions. Fast data is the continuous access and processing of events and data in real-time for the purposes of gaining instant awareness and instant action. Fast data can leverage big data sources, but it also adds a real-time component of being able to take action on events and information before they even enter a Big Data system."[2]

---

[1] "What is Fast Data", Tony Baer, Ovum, Nov 2012

[2] Fast Data Gets A Jump On Big Data, Forbes, Hasan Rizvi, March 1st, 2013

Fueling the Demand for Velocity Solutions

There are a couple of reasons why the appeal of fast data has grown. As increased volumes of data become commonplace across many industry sectors, the value of applying a fast data strategy as an end-to-end solution has become more apparent. For instance, the Big Data Market is $5 billion growing at 58% CAGR over next 5 years to reach $50 billion [3]. This clearly an indication that volumes, varieties and consequently the velocity of data is on the rise.

Secondly, customer touch points are increasing. As the membership to online social communities continues to skyrocket, so does the demand for instantaneous responsiveness as well as improved customer experiences. Fast data means getting the most current information to customers, customer service representatives, and business analysts in a timely manner.

Finally, another factor which has fueled the demand for velocity solutions is the internet of things. There are more devices than ever before. In fact we're seeing an explosion in the number of devices over the next 5 years by 22X. [4] There's now more connectivity of all of this information that we can take advantage of in real-time. As more and more devices come on board we're also seeing an uptake in solutions that require *device to datacenter* (D2D); fast data is an important element in connecting streams of data together from device to the datacenter. This data that exists beyond the companies four walls can also be tapped – and when it is tapped it's important to keep only the most current information. For example, sensors within a city center can help first responders react faster in a storm or a critical situation.

Are You Running your Business Fast Enough?

While speeds and feeds do not usually justify a business case alone, in this case speed *is* the business case. By capturing data faster, being able to move data faster, means analyzing it and acting on it faster. Today more and more cases within IT are demanding the elimination of latency as part of solving a business problem. Data loses value at a faster rate and

---

[3] http://wikibon.org/wiki/v/Big_Data_Market_Size_and_Vendor_Revenues

[4] CISCO VNI Mobile 2012

consequently the value you get from your data diminishes when you don't tap into it soon enough. The following are four key drivers that we're seeing when it comes to fast data solutions:

New Services	Better Customer Experience	Improved Efficiency	Higher Quality In Operations

- **To build new services** –organizations need to have direct insights into their data. For example: building a location based offer requires a collection of real-time information, geo-spatial technologies, as well as marketing data.

- **To improve customer experience** –companies need instant access to customer information, claims transactions, support information, social media metrics

- **To improve efficiencies** – This could be hardware offloading costs, or improved asset utilization by faster data processing.

- **To develop higher quality in operations** –one needs to look at using operational data for an advantage. For example, by collecting events fast and eliminating latencies for reporting companies can shorten their supply chain cycles while reducing gaps in key business processes.

Fast Data across Industries

Financial Services Industry

One of the most traditional styles of fast data is using event correlation to contextualize available financial data. This can inform trading behavior, specifically algorithmic trading, by identifying opportunities or threats that indicate traders (or automatic trading systems) should buy or sell. Algorithmic trading is a growing trend in competitive financial markets. Today, a wide variety of financial applications use event processing technologies, including profit, loss, and risk management systems, order and liquidity analysis, quantitative trading and signal generation systems, and others. The growing trend in this industry is to expand not only the use of event processing technologies but also real-time data integration and analytics as well.

Logistics/Travel

Fast data solutions help monitor all airline's operational events, from passenger check-in, baggage handling (bag on conveyor, bag off conveyor, bag loaded on ULD, ULD loaded on plane, etc.), and flight operations (flight leaves gate, flight lands, etc.). By applying a fast data approach, an airline can better ensure that the right bags get loaded correctly, ensure that flights leave the gate and take off on time, and ultimately manage airline operations in an intelligent, automated way.

Public Sector

Fast Data solutions help public sector organizations respond faster to incidents as they occur in city centers across a range of scenarios: informational, criminal, and terrorist incidents. By being able to correlate and filter a vast number of events (which rapidly decay in value over

time) city centers can support comprehensive real-time data collection and analysis for real time intelligent traffic management, forensic analysis, vehicle journey history, traffic hotspots, crime density.

Telco

Many Telco organizations are faced by the challenges of managing their resources effectively, optimizing capital expenditure (CAPEX) on network infrastructure while lowering or maintaining operational expenditure (OPEX). This requires the ability to gain insights as they occur to understand allocation of network resources based on traffic and application requirements, network usage patterns. Ultimately, fast data can help gain real-insights based on hard data as it comes in, not just approximate trending.

Energy

Energy organizations that have implemented smart grid or smart metering may be utilizing a combination of devices and high volume of events which are delivered over a vast network; these events need to be correlated and processed in real-time to make important real-time decisions. For example, would the Northeast blackout actually have occurred if the right systems had detected the alarms on time? What would have been a manageable local blackout cascaded into widespread distress on the electric grid.[5]

Manufacturing

Capture events from business applications, manufacturing systems, and other process devices and proving real-time dashboards and actionable analysis. Control and monitor operations, analyze real-time metrics that for immediate results to take corrective action before a failure occurs. Recognizing the need for maintenance can save in repair costs and costly delays.

Customer Experience / Retail

Customer service centers are using Fast Data for click-stream analysis and customer experience management. Fast Data can factor real-time information about millions of events (clicks or other interactions) per second into business intelligence and other decision-support applications. These "recommendation applications" help agents provide personalized service

---

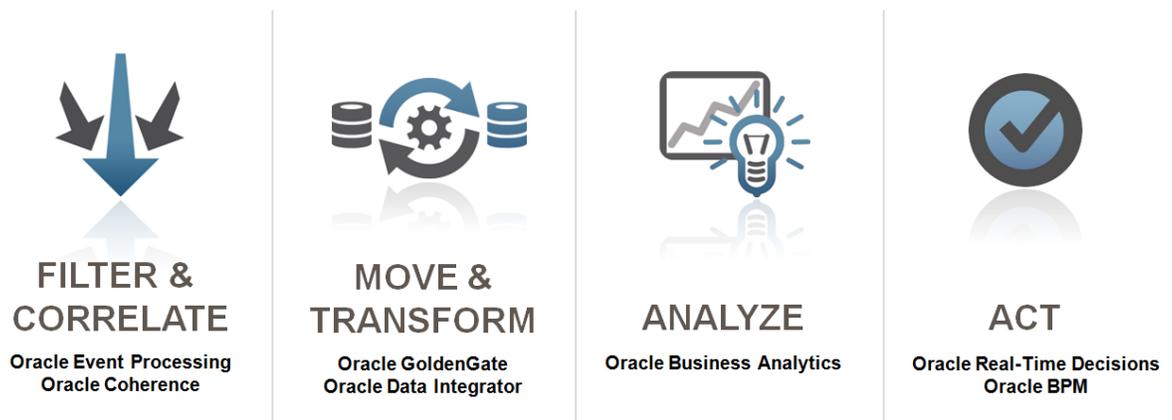[5] http://en.wikipedia.org/wiki/Northeast_blackout_of_2003

based on each customer's experience. Click-stream analysis can help target offers and even initiative customer service interactions. For example "it looks like you're having difficulty deciding which camera to purchase… would you like to chat online with a technical representative?").

Healthcare

Another example of Fast Data in practice is in the healthcare industry. One example, is the HyReminder system, developed by the Worcester Polytechnic Institute and UMass Medical School, continually tracks healthcare workers for hygiene compliance (e.g. sanitizing hands and wearing masks), reminding them to perform hygiene when appropriate to prevent the spread of infectious disease. Each worker wears an RFID badge that displays a green (safe), yellow (warning) or red (violation) light, depending on what behavior the RFID chip has observed.

Oracle's Solution for Fast Data

Oracle's Fast Data solutions offer multiple technologies that work hand-in-hand to create value out of high-velocity, high-volume data. They are designed to optimize the efficiency, scale for processing high volume events and transactions.

FILTER & CORRELATE
Oracle Event Processing
Oracle Coherence

MOVE & TRANSFORM
Oracle GoldenGate
Oracle Data Integrator

ANALYZE
Oracle Business Analytics

ACT
Oracle Real-Time Decisions
Oracle BPM

- **Filter and Correlate**. With Oracle Event Processing you can use predefined rules to filter and correlate data through big data sources. In addition, this solution has the advantage that it can run in-memory to optimize performance, scale by integrating with Oracle Coherence.

- **Move and Transform.** With Oracle Data Integrator and Oracle GoldenGate you can capture data (structured or unstructured) and immediately move information where it is needed—and in the right format—to best support decision making.

- **Analyze.** Oracle Business Analytics enable you to discover what's possible for real-time analysis.

- **Act.** Oracle Real-time Decisions together with Oracle BPM, helps support both automated decision-making as well as more complex, human-based interactions, such as business process management.

All of these components run on a rich data tier environment which supports both: Big Data Hadoop/NoSQL, as well as SQL, and scales elastically on standards based Java platforms. Only Oracle has a complete and integrated solution with the best-of-breed components when it comes to processing events, data, and analytics in real time. And, only Oracle can run these fast data components on diverse deployment architectures, including device gateways, big data environments, or engineered systems.

Oracle Event Processing and Oracle Coherence

Oracle Event Processing is a complete solution for building applications to filter, correlate and process events in real-time so that downstream applications, service oriented architectures and event-driven architectures are driven by true, real-time intelligence. Oracle Event Processing provides organizations with a complete "top-down" solution for designing, defining, development and implementing event processing applications that not only meet business requirements but perform to the highest levels of enterprise expectation.

One of the first set of requirements in a fast data architecture is processing against live dynamic data at the rate of the fastest event stream. This requires one to apply known queries against unknown, incoming data. Oracle Event Processing is built on industry-standards including ANSI SQL, Java, Spring DM and OSGI to provide an open architecture for sourcing, processing, and publishing complex events throughout the enterprise.

Oracle Event Processing is a complete solution for building applications to filter, correlate, and process events in real time. It ensures that Ensure downstream applications, and service-oriented and event-driven architectures are driven by true, real-time intelligence. It helps

8

eliminate, consolidate, correlate, and filter data before it clutters the data warehouse and the Hadoop clusters. It also helps analyze in real-time massive data streams, as they get inserted into the big data repositories. The OEP Platform supports a variety of data "cartridges" including those for Hadoop and NoSQL databases and allow for the "filtering and fresh data" scenario often required for analysis in Hadoop implementations.

Oracle Event Processing has both a visual development environment as well as standard Java-based tooling. This ensures that your IT team can be developing event-driven applications without the hurdle of specialized training or unique skill-set investment.

At its core, Oracle Event Processing is a Java container implemented with a lightweight, modular architecture based OSGi™ based on Spring Dynamic Modules. It provides a complex event processing (CEP) engine and a standards based language "Continuous Query Language" (CQL), (SQL for data streams) with extreme performance and deterministic garbage collection. This results in event processing times almost as fast as the event stream.
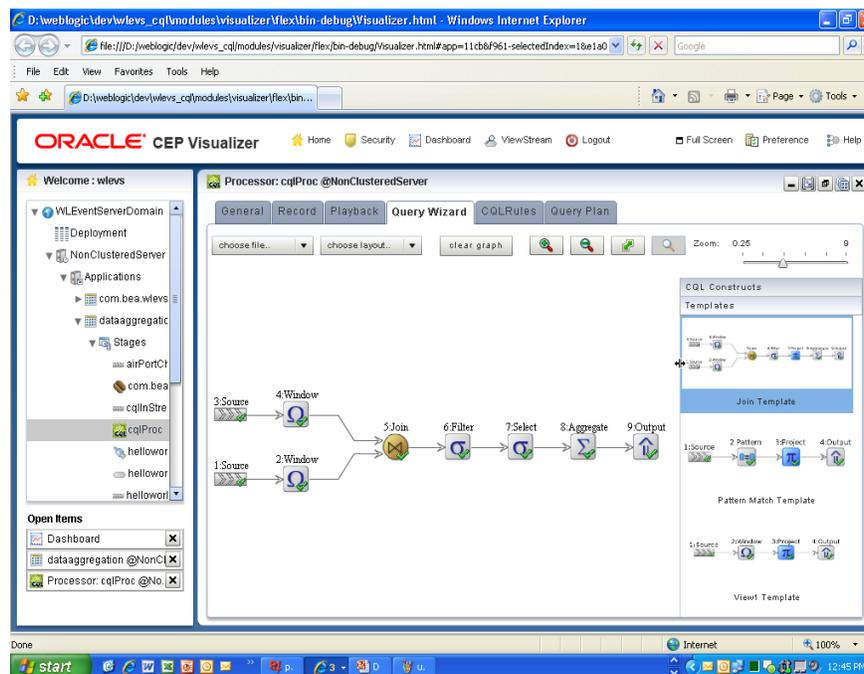
In addition, Oracle Event Processing can easily manage externalized queries which may be injected in runtime by trained business analysts. The queries could be parameterized, allowing business analysts to tweak the parameters through a web interface. Using the Oracle Event Processing Visualizer it is possible to replay or reprocess events while the built in Oracle

9

Coherence data grid serves as a cache for look ups, enrichment and high availability. This allows for the easy creation of event queries requiring close proximity to large data sets that maybe static in comparison to event streams.

Oracle Data Integration

Oracle Data Integration delivers a complete solution for integrating data 'fast' which means moving data continuously, pervasively, and accurately across Oracle technologies as well as 3rd party systems and applications. It combines all the elements of data integration—real-time and bulk data movement, data synchronization, big data transformation, data quality, and data services into a unified platform. Oracle Data Integration is designed to deliver maximum performance with lower cost of ownership, ease of use, and reliability.

Fast Data and Oracle's Real-time Data Integration

The essence of fast data is in ensuring that 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.

Oracle's strategy for real-time analytics incorporates the real-time data replication capabilities of Oracle 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. 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. Only Oracle provides complete integration from "report to source," permitting organizations to trace the data lifecycle from the source of heterogeneous data systems all the way to the business dashboards and reports.

Oracle Business Analytics

Oracle Business Analytics provide a full range of analytics that unlocks the value from various data scenarios including Fast Data - to get answers at the speed of thought and empower business users to take action based on facts with confidence. Acting on data in real-time – several Oracle products enable solutions that make fast data actionable.
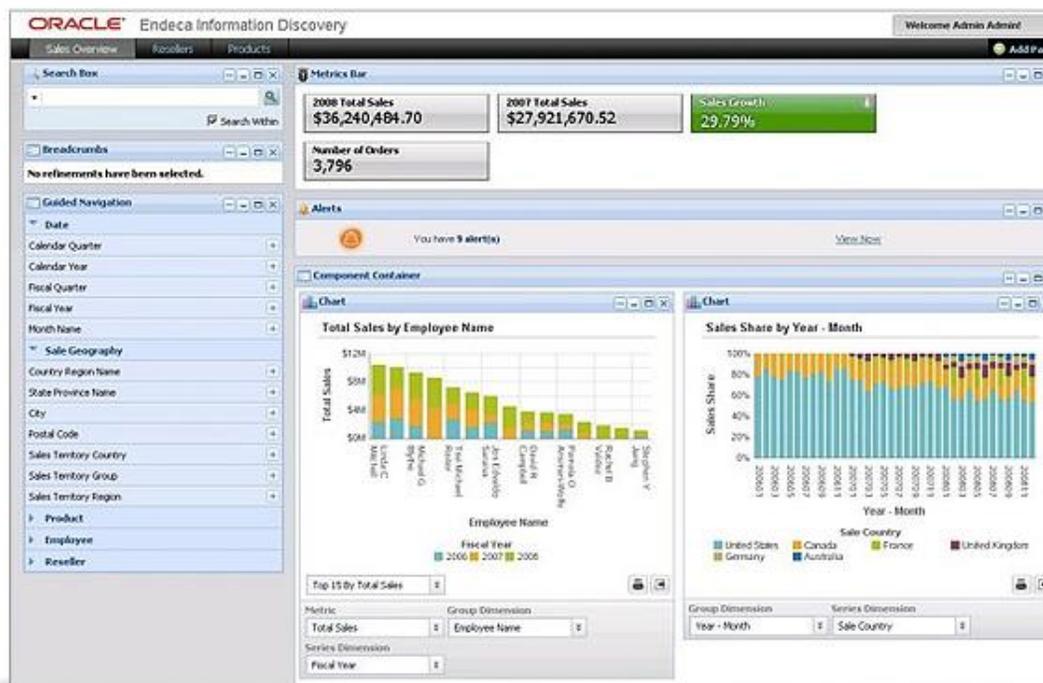
Oracle Real Time Decisions

Oracle Real Time Decisions (RTD) provides a real-time path for action that enables response time in milliseconds: fast enough to optimize interaction with customers and prospects. RTD can use analytics run outside of the real-time path, but accessed in real time, to enhance the decision making process. Oracle Advanced Analytics (OAA) builds predictive models and generates scores that guide RTD decisions. RTD combines the data stream with OAA predictive models and scores, business rules and performance goals to deliver the Next Best

11

Action to the customer or prospect. RTD enhances Oracle Business Process Management (BPM) by enabling real-time reports and optimization of business processes.

For the highest velocity streams, Oracle Event Processing can be used to filter the stream and identify potential actionable events that can be passed onto Oracle Real-Time Decisions. Oracle NoSQL Database can be used to store recent events for fast retrieval by OEP, enabling accurate identification of more complex events (e.g. one GPS signal gives location; the last 10 combined give direction and speed as well as location).

Oracle Business Intelligence Foundation

Oracle provides a rich portfolio of analytics from discovery of diverse and unstructured data to reporting Key Performance Indicators to planning and forecasting using what if scenario modeling. Oracle Business Intelligence Foundation integrates the capabilities of multiple BI tools into a single, unified technology platform that supports the full spectrum of enterprise requirements for business analytics including reports, dashboards, scorecards, scenario planning, and forecasting models, as well as ad hoc queries and analysis.

Oracle Endeca Information Discovery

One of the key challenges of analyzing Big Data is the variety of information, especially unstructured data. Users frequently don't know what they're looking for or how to find it. They want the freedom to explore and discover new insights and meaningful conclusions. Oracle Endeca Information Discovery helps users explore and analyze diverse and unstructured data, to find better ways to listen to customers, improve the relationship between companies and their employees, get to the real root causes behind issues, or create new products and services through a better understanding of the market.

Oracle Fast Data on Engineered Systems

Oracle's engineered systems have a lot to offer to Oracle's Fast Data Platform – they of course run fast data faster, but in addition they improve time to value, lower costs and simplify operations. Oracle Engineered Systems are optimized to achieve enterprise performance levels that are unmatched in the industry.

They enable a faster time to production by implementing pre-engineered and pre-assembled hardware and software bundles. And finally Oracle's single-vendor stack simplifies and reduces the costs associated with purchasing, deploying, and supporting IT environments

while minimizing the business risk associated with such projects. Many of the solutions in the Fast Data arsenal, including Oracle Event Processing, Oracle Data Integration, and Oracle Business Analytics are highly tuned and optimized for Oracle's engineered systems. Let's look at a few examples.

Oracle Event Processing with Oracle Coherence on Oracle Exalogic

Oracle Event Processing on Oracle Exalogic delivers even higher throughputs and lower latencies that are not possible on commodity hardware. Scenarios that require an Oracle Coherence cache (where in-memory caching is required) and database access will see an even more substantial benefit in terms of performance, throughput, latency, and scalability on Exalogic. For example, running in Oracle labs, one million events were processed per second; the average event latency for the full processing path on the server was 32 microseconds with 99.99 percent of the events processed in less than 2 milliseconds.

Oracle Data Integration and Oracle Exadata Database Machine

The Oracle Exadata Database Machine delivers unbeatable performance and scalability for all your database applications. However companies can only leverage their investment if the optimal data loading strategy is in place. Integrating data efficiently and at high velocity is what fast data is all about. Oracle's data integration product line, including Oracle Data Integrator Enterprise Edition, Oracle GoldenGate, and Oracle Data Quality, offers the best-of-breed solution for providing clean, consistent, continuous access to your data. The solution provides extreme data integration performance with maximum availability. Oracle's data integration solution is fully certified with and optimized for the Oracle Exadata Database Machine. Oracle's data integration solutions are optimized to run on, interoperate with and integrate to the Oracle Exadata Database Machine. In independently verified trials, Oracle's data integration solutions speed the loads of the Oracle Exadata Database Machine by 500% while providing continuous access to business critical information across heterogeneous sources.

13

Oracle Business Analytics and Oracle Exalytics

As analytic applications become more sophisticated and calculation-intensive, the use of mobile BI expands, user adoption increases, and data volumes explode making the need for speed and efficiency more important than ever. In-memory technology can dramatically accelerate analytic performance. Oracle Exalytics In-Memory Machine is the industry's first engineered system for analytics that combines analytic software, in-memory database software, and hardware all engineered and optimized to work together to deliver extreme performance. As a result, users can visually navigate and drill into information at the speed of thought, without limits on the complexity of their questions or the volume of the underlying data. Oracle Exalytics drives a new class of smarter and more powerful analytic applications – from discovery to planning - that simply weren't possible using conventional BI software and generic hardware configurations.

With an end-to-end family of analytic solutions optimized on Engineered Systems, Oracle helps organizations thrive by enabling them to gain insight into every aspect of their business and market, predict and plan ahead, and act with confidence.

Summary

While the term Fast data might be new; it's a set of mature ideas and technologies around the rapid processing of events and high volumes of data for improved analytical insights. It's recently been impacted by the explosion in volume and variety of data [big data]; the ubiquitous nature of devices and internet of things, and ultimately the heightened expectations of customers. No longer do we see fast data in only financial stock exchanges or limited to Hadoop based deployments; but it's becoming more and more pervasive across industries as companies are recognizing the need to run their businesses in real-time.

15

Only Oracle provides a complete and best-of-breed platform for fast data to handle increasing volume, velocity & variety of events and data to generate value for customers looking to implement real time analytics use cases. The components of this platform are engineered to work together. In addition, each component is based on open standards and interfaces that allow customers to leverage their existing enterprise standard components in areas where investments may have already been made. With the use of Oracle Engineered Systems, Oracle is the only vendor today that can add even more increased value by tapping your high velocity data with unmatched performance, reduce business risk and deliver innovative new services.

For more information on Oracle's complete solution for Fast Data go to:
www.oracle.com/fastdata

ORACLE®

Big Data Gets Real-time
February 2013
Authors: Dain Hansen

Contributors: Irem Radzik, Amit Zavery, Ben Bhargava, Shiva Bhajekar, Rashmi Vittal

Oracle Corporation
World Headquarters
500 Oracle Parkway
Redwood Shores, CA 94065
U.S.A.

Worldwide Inquiries:
Phone: +1.650.506.7000
Fax: +1.650.506.7200

oracle.com

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

Hardware and Software, Engineered to Work Together