A HURWITZ WHITE PAPER

Winning Mobile and Web Customers with Fast Data

Marcia Kaufman,
COO and Principal Analyst

Daniel Kirsch,
Research Analyst

HURWITZ & ASSOCIATES
Insight to Action

Sponsored by VMWare

Winning Mobile and Web Customers with Fast Data

Business leaders are finding that they need to dramatically transform the way their organizations engage with customers to remain competitive. Fast Data – the data architecture that supports real-time decision-making – is at the heart of this transformation. Customer's expectations have evolved. Now, customers expect to have the information needed to make a purchasing decision at the time and place that is right for them. In addition, customers expect customized and instantaneous responses from their suppliers. Businesses increasingly understand that customers expect to be able to interact though web and mobile devices.

Fast Data has emerged as a requirement to provide companies with the responsiveness and agility needed to increase the level of partner and customer engagement.

Therefore, smart businesses are creating innovative mobile and web applications that delight their customers and build loyalty. But this transformation doesn't happen in a vacuum. Leaders have to rethink their approach to data management. The tables have been turned. Now, the customer controls much of the engagement process. For example, customers increasingly want to choose which device and platform they will use to communicate with their service providers. In addition, customers expect the engagement experience to be seamless and personalized based on their requirements and their relationship to their providers.

To meet this changing engagement model requires that the right information be available at the right time. Fast Data has emerged as a requirement to provide companies with the responsiveness and agility needed to increase the level of partner and customer engagement. This paper will discuss how organizations are using Fast Data to provide customers with the right product or service offering at the right time. We will also provide an example of a Fast Data solution based on VMware's SQLFire database, which is part of the vFabric cloud application platform.

Fast Data does not live in isolation. In many situations companies are dealing with massive amounts of data – Big Data - that has to be delivered at unparalleled speed. Therefore, these companies need to manage Big and Fast Data. If a company only focuses on Big Data, they may not be able to provide that data at the right time to customers. For additional information on Fast Data in combination with Big Data, please refer to the Hurwitz & Associates companion White Paper, "Data Management in Transition: Where are We Heading with Big and Fast Data?"

Defining Fast Data Architecture

Fast Data is a data architecture that supports storage, access, and analysis of data at speeds that allow for real-time decision-making. It is the lynchpin core technology required to support a variety of emerging mobile and web applications. Fast Data implementations require a framework that moves data both physically and logically closer to users and applications in order to eliminate data latency problems that are inherent with traditional databases.

HURWITZ & ASSOCIATES
Insight to Action

Why is this transition to Fast Data needed? Historically, large amounts of transactional data have been collected and stored in relational databases. While these traditional systems of record hold a company's most prized customer and product information, the data can't always be processed quickly enough to be immediately actionable. Fast Data allows companies to leverage this treasure house of information with the right velocity and within the right timeframe.

For example, imagine if a retail store employee could determine when a customer who purchased one product would be 80 percent likely to buy a related product. If that offer is made at the point of sale, the business has the opportunity to extend and enhance the monetary value of the transaction. Alternatively imagine the impact on a hospital's effectiveness if patient data from tests could be combined with data collected from real-time monitoring systems in order to detect changes in a patient's medical status. If a physician can immediately detect an important change in a patient's health, lives can be saved

Analyzing customer behavior patterns in real time, allows businesses to present customized user experiences, leading to greater loyalty and increased sales opportunities.

Customer Demand Drives the Fast Data Evolution

Customers are increasingly demanding faster and more customized business experiences. Until recently, customers were willing to purchase products and services from one-size fits all websites that had static, generic offerings. Companies like Google, Amazon and Facebook have helped to change customer expectations by demonstrating that highly customized and fast user experiences are possible. As a result, customers expect vendors to recognize and anticipate their preferences, and compete for their business with innovative and highly customized offerings. The bar for delivering fast and personalized responses has been raised even higher as customer move to mobile devices.

Businesses are making use of a range of advanced technologies, including cloud computing, collaboration tools, and advanced analytics to help meet these new demands. Combining data elements with these tools enables businesses to implement targeted and personalized offerings based on customer behaviors and requirements. Analyzing customer behavior patterns in real time, allows businesses to present customized user experiences, leading to greater loyalty and increased sales opportunities. Retailers, for example, need up to date positioning and browsing data to make personalized offers to customers. Creating new business strategies based on deepening the level of customer engagement requires a major transformation in both the speed and performance of the underlying technology infrastructure. Data delays of just a few seconds can disengage users, making it nearly impossible to deliver the customized, fast experience that customers expect.

Companies are adopting Fast Data technologies to reduce latencies in customer and product data and support new customer focused business strategies. Table 1 provides several examples of Fast Data implementations.

HURWITZ & ASSOCIATES
Insight to Action

Table 1. The Role of Fast Data

Business Initiative	The Role of Fast Data
Personalized Web Offers	Web commerce sites offers discounts based on a customer's new purchases and historical data. Fast Data is used to immediately bring a customized offer to the customer at the point of sale.
Web Pricing	Travel sites need to continually change bundled pricing based on the number of services a customer buys and the availability of inventory. Fast Data allows the website to continually monitor multiple vendor offerings and inventory. The website is able to select the most suitable options for the customer.
Mobile Location Services	Mobile applications are able to give users customized offerings based on their real-time location. Fast Data enables the application to deliver deals that are based on the user's exact location.
Mobile Gaming	Mobile gaming applications allow users to compete with friends. Fast Data is used to improve the dynamic quality of the game experience for users.
Infrastructure Monitoring	Utility companies keep track of energy usage and potential outages based on monitoring data from thousands of sensors streaming into a back-end database. Fast Data supports a utility company's ability to monitor this infrastructure in real-time. As a result, the company can prevent outages and identify problems well before they interfere with energy delivery.
Fraud Monitoring and Prevention	Financial services firms want to compare new transactions against historical patterns of fraud. Fast Data allows the company to immediately analyze credit transactions and only authorize legitimate purchases.

The Path to Fast Data

Fast Data is becoming part of a larger data lifecycle. In the real world, companies will continue to leverage their traditional transactional systems of record. This data is typically centered on customer transactions and is not likely to change dramatically. However, there increasingly needs to be a way to combine system of record data with real-time or near real-time data. Management must be able to leverage this data in new ways to change the customer engagement model. Combining transactional data with real-time data about immediate customer actions will enable businesses to more readily interact with customers at the point of engagement. Successful execution of these emerging engagement models depends on a businesses ability to leverage a database structure with lightening fast response times.

Combining transactional data with real-time data about immediate customer actions will enable businesses to more readily interact with customers at the point of engagement.

HURWITZ & ASSOCIATES
Insight to Action

This requirement for increased speed and performance initially became apparent in the early e-commerce era when web developers were struggling to improve responsiveness. During this time, vendors conducted extensive research to determine how sensitive customers were to slow and poorly performing shopping sites. Research showed that users would leave a website and abandon a shopping cart if the application took too long to load. Although website responsiveness has improved, user expectations for speed and the overall quality of the experience has continued to escalate. Just a few years ago a 2-3 second delay may have been acceptable. Today, users are dissatisfied with delays of even a tenth of a second. In fact, a Google research study showed that very small variations in speed could make the difference between the success and failure of a search engine. [1] When users have a choice they are more likely to choose the fast and better performing service. In other words, speed matters to customers.

Achieving the right speed for your customers depends on how quickly your applications can access, process, and analyze the relevant data. Managing data at the right speed is challenging when a large number of data sources must be relied on to make a highly personalized offer. The data from these distributed and varied sources needs to be ingested quickly so analysis can be done in time to provide a recommendation or solution at the point of engagement. In addition, delivering the right level of performance depends on how quickly your application can scale to support a large numbers of users.

Executing a strategy to enable fast, customer centric experiences, requires a new approach to data management infrastructure. This approach needs to eliminate the data bottlenecks that occur when systems are tasked with executing thousands of simultaneous requests.

Fast Data Infrastructure Requirements

Executing a strategy to enable fast, customer centric experiences, requires a new approach to data management infrastructure. This approach needs to eliminate the data bottlenecks that occur when systems are tasked with executing thousands of simultaneous requests. Traditional relational databases must constantly read and write to disk, creating multiple copies during a transaction. The end result is that data latency increases as message requests begin to stack up while the system is trying to read or write to the physical disk. There are four key elements that are needed to support a Fast Data Infrastructure: in-memory database structures, support for WAN replication over a distributed network, scalability, and support for real-time analytics.

In-memory data

One of the key elements of this new approach is to store more of the relevant data in memory. This approach contrasts with traditional database management systems that store data on disk. Data that is stored in-memory is closer to the application eliminating performance delays that occur when an application needs to retrieve data from storage or write it to the disk. As a result, in-memory databases can increase processing speeds by 10 to 100 times over traditional relational databases. One of the benefits of the in-memory model is that it

1. J. Brutlag. Speed matters for Google web search, June 2009. http://services.google.com/fh/files/blogs/google_delayexp.pdf

HURWITZ & ASSOCIATES
Insight to Action

reduces the need to do multiple data transfers and replications. This is because information is directly fed from memory to the requesting application.

Wide Area Network (WAN)

A Wide Area Network links disparate locations together and may cover a very large geographic area. Multiple servers (or nodes) can be located at different locations within the WAN allowing relevant data to be brought closer to individual users. This approach enables continuous availability of the data and results in near real-time performance. For example, transaction speed is improved for all users when application users in North America access information from a datacenter in the United States, while those in Europe are served data from a datacenter in Germany.

When working with an in-memory system, data is not necessarily persisted to disk, so it is critical that there is a process in place to ensure reliability in case of service disruptions. High availability is maintained by implementing a WAN to connect multiple nodes. Automating and streamlining processes for data replication and persistence to disk ensures that data is not lost if a node fails. Distributing data has two main benefits. First, it ensures that the data is able to serve applications even in the event of a major disaster to one datacenter. Second, by distributing the data across multiple locations, latency is decreased because applications will use the data closest to the user. Implementing a WAN helps to maintain high availability ensuring that in-memory data can run across multiple datacenters or multiple clouds.

High availability is maintained by implementing a WAN to connect multiple nodes. Automating and streamlining processes for data replication and persistence to disk ensures that data is not lost if a node fails.

Scalability

Well-designed, scalable databases can support the addition of servers (or nodes) onto the existing infrastructure. Balancing between the additional nodes and the previous ones should be handled automatically and not require changes to the database schema or application code. This type of linear scaling is known as scaling out (horizontal scaling) rather than scaling up (vertical scaling) because an organization isn't purchasing new, bigger systems, but is instead increasing their capacity through the addition of more nodes.

Real-time analytics support

Big Data environments such as Hadoop and Greenplum are complimentary technologies to Fast Data. Big Data platforms provide companies with the capability to manage and analyze large volumes of data. Fast Data provides companies with the opportunity to quickly process data. When companies use these technologies in combination, they can analyze large volumes of complex data and process results in the right timeframe to make attractive offers to customers. For example, a retailer can use Big and Fast Data to predict customer responses to various offerings and make a customized discount offer to a customer at the point of sale. To achieve this level of performance requires parallel processing techniques such as Hadoop and MapReduce. Hadoop provides the ability to manage and analyze highly distributed data, while MapReduce collects results to execute on a task.

HURWITZ & ASSOCIATES
Insight to Action

VMware vFabric SQLFire

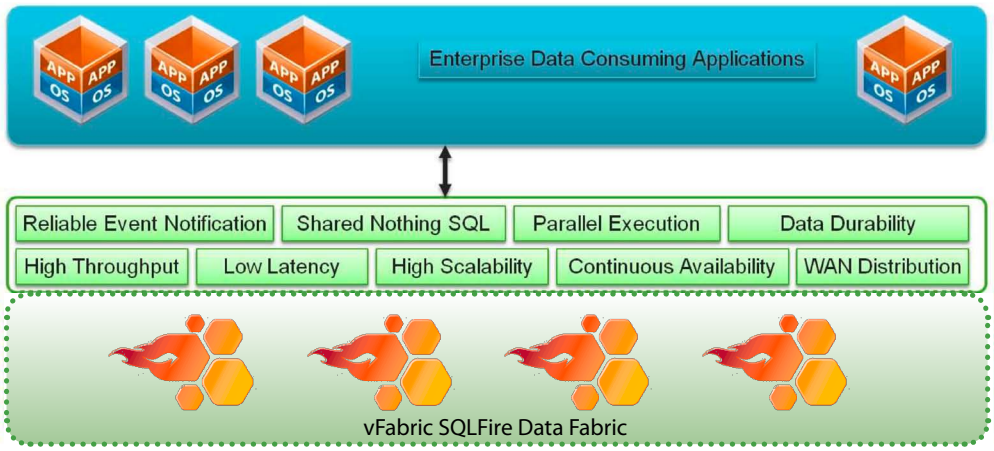
vFabric SQLFire is VMware's approach to meet the demands of customers, who are seeking to create fast, customized experiences for their application users. SQLFire is an in-memory distributed SQL database designed to support high-speed applications. SQLFire is one of the foundational data services within the vFabric Cloud Application Platform. VMware's vFabric is a middleware suite for developing data intensive applications. It is built on the Spring framework – an application development framework for enterprise Java. vFabric leverages the server architecture of VMware vSphere and provides services such as application management, application monitoring, and Java application runtime support. By supporting standards such as SQL, JDBC (for Java applications), and ADO.NET (for Microsoft .Net applications), SQLFire can integrate into an existing enterprise infrastructure and connect with existing applications that use relational databases.

vFabric SQLFire is VMware's approach to meet the demands of customers, who are seeking to create fast, customized experiences for their application users.

SQLFire Deployment Models

Many organizations use vFabric SQLFire as a primary database for new applications. Typically SQLFire would be used to help eliminate potential data bottlenecks in new mobile and web environments. As shown in Figure 1, SQLFire acts as a full data management system.

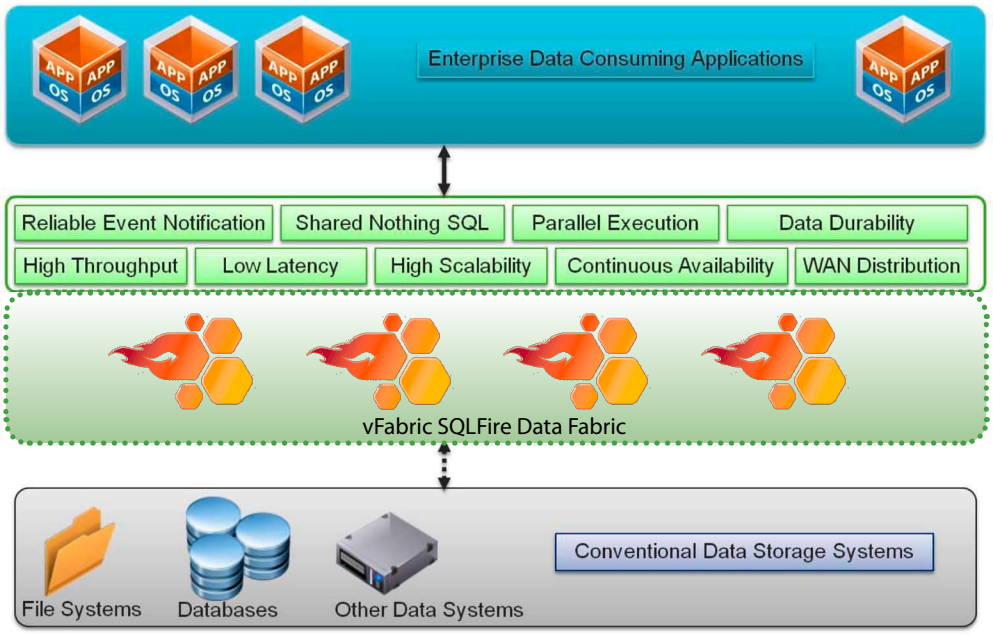
Figure 1: SQLFire as the primary database



Another common deployment option for SQLFire is to integrate it with an existing traditional database or an analytics engine. In this model, SQLFire acts as a caching layer to enhance performance (see Figure 2). The combination of SQLFire with traditional technologies provides a high-speed system for data ingestion along with the required deep analytics.

The combination of SQLFire with traditional technologies provides a high-speed system for data ingestion along with the required deep analytics.

Figure 2: SQLFire as a data accelerator with a traditional database as the system of record



SQLFire Core Characteristics

vFabric SQLFire incorporates the key elements needed to support a Fast Data infrastructure. SQLFire is designed for speed. It uses asynchronous messaging to make the process of writing to disk highly efficient and provide both speed and persistence. Six core characteristics of SQLFire are described below.

Wide Area Network (WAN) Clustering

SQLFire WAN gateways can be implemented to support a globally distributed data layer. For example, replicated data can be housed in Europe, North America and Asia. These SQLFire WAN clusters enable a very large number of nodes to act as one logical system and provide a consistent view of data around the world. As a result of this capability, data is brought closer to users and is protected from service disruptions at one site. Additionally, WAN clustering allows systems to expand and contract dynamically to support the horizontal scalability requirements of online and mobile applications.

Multisite deployments enabling disaster avoidance and recovery as well as consistent data

SQLFire takes an active-active approach to replicating data across geographies. This provides significant advantages for backup and disaster recovery and ultimately disaster avoidance compared to the active-passive approach. The active-passive method means that an offline (passive) database is used to mirror the active (live) database. If there is a disaster, you will need to bring the passive backup database online. Under this scenario you may lose messages as a result of the disaster. With SQLFire messages may be delayed, but they will not be lost.

SQLFire multisite deployments allow nodes to be hot-swapped for disaster recovery because all of the nodes in the cluster are active. Therefore, if a datacenter in Europe fails, an alternative datacenter in North America can be accessed. As the European datacenter gets back online, WAN connectivity allows it to easily and rapidly resynchronize with the North American datacenter. Disaster avoidance can also be achieved with a WAN deployment. Although one datacenter may go down, applications will seek data from an alternative, replicated datacenter and users will be unaffected. Multisite deployments also allow the data to be physically closer to users, which allows for faster access and consistent views of the data no matter where a user is accessing it.

Shared-nothing architecture

Shared-nothing architecture means that each node in the SQLFire environment is logically independent. As a result, an application will be able to function even if one of the nodes fails. Without this protection a single node failure could lead to additional node failures bringing down an application. Drivers that are shipped with SQLFire allow the system to automatically reconnect to a working node in the case of a node failure.

Scale-out infrastructure

SQLFire includes elastic scaling and auto-balancing capabilities. When resources approach capacity, new nodes can easily be added to the environment as needed without disrupting applications. When a new node is added, data is replicated to the node and workloads are automatically shifted. SQLFire supports commodity servers which helps keep hardware costs under control when scaling. In addition, SQLFire deployments can be implemented entirely in the cloud, and when additional infrastructure is required, more resources can be purchased from a cloud vendor.

Tunable data replication

Although, SQLFire typically places related data on the same node to increase both speed and availability, some adjustments to the data management process may be needed to deliver the fastest results. The standard approach is to replicate data tables on each node. This delivers fast speed and also makes sure that the data is still available if one node is down. This approach is best if the tables are relatively small, frequently accessed, and not likely to change often, such as product data. Alternatively, for very large data sets that change rapidly

SQLFire includes elastic scaling and auto-balancing capabilities. When resources approach capacity, new nodes can easily be added to the environment as needed without disrupting applications.

HURWITZ & ASSOCIATES
Insight to Action

(i.e. stock prices, inventory data, or mobile game data) you would partition a table across multiple nodes to increase speed and also create a redundant copy to ensure reliability.

Choice of in-memory or disk-based storage

SQLFire provides two options for data storage. Data can be stored completely in-memory along with redundant in-memory copies or data can be persisted to disk.. As a result of this flexibility, organizations can ensure that all processing is performed in-memory to reduce the latency problems with disk-based systems and then have data written to disk after processing.

With SQLFire the transaction rate increased from 50-60 transactions per second to 600 transactions per second.

Case Study: How SQLFire Speeds Transaction Performance

A telecommunications company was losing market share because of its poor performance with existing customers. One of the major sources of customer dissatisfaction could be traced to the inflexibility and slow speed of its customer-facing systems. Responses to customer queries were often delivered several days late if at all. In addition, the company was falling behind its competitors due to its inability to offer attractive offers to prospects in a timely manner. While the company had plenty of historical data about customers it was not able to use this data quickly enough to provide adequate and timely responses to customers. With the existing system, the company could only process about 50-60 transactions per second. When the system was originally designed, speed was not a priority because it was intended for batch processing. But as the business changed the database now had to perform new functions that it was ill suited for. The large number of dependencies between the applications and the database made it very difficult to improve the transaction speed without starting from scratch with a new system. But since the existing application still served a valuable purpose, IT executives could not justify the expense to rewrite it from scratch. The company needed a solution that would enhance performance leveraging the existing backend application.

The company used SQLFire as a front-end data management system to improve data responsiveness for new web and mobile applications. With SQLFire the transaction rate increased from 50-60 transactions per second to 600 transactions per second. To implement SQLFire, the company extracted the data from the application through the use of VMware tools. This extracted data was placed in a container that was attached to SQLFire. With the improved transaction performance, the company was now able to more effectively respond to customer issues. In addition, with SQLFire's speed, the company could deliver personalized offers that are now helping them win new business.

HURWITZ & ASSOCIATES
Insight to Action

Conclusion

In the face of increasingly demanding customers, companies must be able to respond to opportunities and threats at the right speed and at the right time. In addition, applications must be robust enough to support a highly distributed user base, a variety of device types, and wide swings in usage. This, in turn, requires companies to develop a new vision for data management that positions them for the era of real time data. While companies need the ability to leverage traditional data management platforms, they also need the flexibility to leverage alternative platforms that support Fast Data.

What is the benefit to companies that implement a Fast Data infrastructure? An in-memory database designed for speed and low latency will give companies the scalability and flexibility needed to support new business initiatives. By combining an in-memory database with Big Data environments like Greenplum, companies can begin to solve a whole new class of problems. The ability to analyze customer-buying patterns in real time can lead to personalized offers and deeper more positive customer experiences. Ultimately, a Fast Data infrastructure helps businesses anticipate their customers' needs and create a sophisticated and responsive customer experience.

An in-memory database designed for speed and low latency will give companies the scalability and flexibility needed to support new business initiatives.

HURWITZ & ASSOCIATES
Insight to Action

About Hurwitz & Associates

Hurwitz & Associates is a consulting, market research and analyst firm that focuses on how technology solutions solve real world business problems. The firm's research concentrates on disruptive technologies, such as Cloud Computing, Information Management including Big and Fast Data, Service Oriented Architecture and Web 2.0, Service Management, wireless computing, security and Social and Collaborative Computing. We help our customers understand how these technologies are reshaping the market and how they can apply them to meet business objectives. The team provides direct customer research, competitive analysis, actionable strategic advice, thought leadership, white papers and speeches. Additional information on Hurwitz & Associates can be found at www.hurwitz.com.

175 Highland Avenue, 3rd Floor • Needham, MA 02494 • Tel: 617-597-1724
www.hurwitz.com