Cisco, Fusion-io, and Oracle Deliver Extreme Performance to Oracle NoSQL Database Big Data Applications

Solution Brief
September 2012

Highlights

Extreme Performance

- Cisco Unified Computing System™ (Cisco UCS®), Fusion ioDrive2, and Oracle NoSQL Database deliver 1.2 million operations per second to big data applications.

Large-Volume Management of Big Data

- Oracle NoSQL Database delivers predictable latency to support real-time big data applications, with transactional semantics for easy data access, high availability, and scalable throughput.

High-Performance Intelligent Infrastructure

- Cisco UCS delivers high performance and radically simplifies deploying big data applications with its intelligent infrastructure.

Real-Time Response and Millisecond Latency

- Fusion ioDrive2 delivers the millisecond latency big data applications need to maintain real-time response when processing tens of terabytes of data.

The superior performance of the Cisco Unified Computing System™ (Cisco UCS®) and Fusion ioDrive2 matches the scalability, reliability, and simplicity of the Oracle NoSQL Database to deliver 1.2 million Oracle NoSQL operations per second.

The combination of the Oracle NoSQL Database, Cisco UCS, and Fusion ioDrive2 provides a compelling and cost-effective solution for big data environments. Accelerated performance was validated when Cisco, Fusion-io, and Oracle partnered to test the solution's capability to support real-time applications. Testing results show that the system delivered over 1.2 million operations per second and achieved an average latency of 0.88 milliseconds for read operations and 4 milliseconds for update operations. This extreme performance enables businesses to quickly collect hundreds of terabytes of data and analyze it to glean insights for strategic, competitive advantage.

Big Data Requirements

"Big data" is an informal term that encompasses the collection, storage, and analysis of massive amounts of information, including data from web logs, sensors, tweets, blogs, user reviews, Short Message Service (SMS) messages, and data-intensive Internet applications. The data is characterized by high volumes of extremely large data sets (hundreds or thousands of terabytes in size) and typically has no structure. Data must be delivered at high velocity, on the order of hundreds of thousands of operations per second, so it can be consumed in near-real time. By analyzing all available information, companies can gain insight as to how the business is performing against tactical and strategic goals.

Deliver Answers Fast

Delivering answers quickly under fluctuating workloads is a critical requirement for big data processing. For example, consider a big data solution that manages user profiles for e-commerce websites. The capability to look up a specific user's

profile with extremely low latency is critical to maintaining user satisfaction. In addition, e-commerce activity can fluctuate significantly over time and experience peaks in demand, causing data access needs to be highly variable. A comprehensive solution must be able to gracefully handle such changes in demand without requiring a restart. This requirement necessitates around-the-clock delivery of required throughput with predictable low latency under widely varying workload conditions.

Handle Growing Data Volumes

Enterprises continue to struggle with the demands of a connected economy, with dramatic increases in Internet applications generating massive data growth. More important, extracting valuable information from repositories in real time is essential to support business applications. Together, these stresses have created a surge of interest in big data solutions. With more enterprises adopting big data processing for its business benefits, open source and commercial vendors are providing a variety of solutions aimed at addressing specific big data processing needs.

The Problem with Conventional Technology

The need for performance is not limited to certain applications or market segments. Nearly every company can benefit from faster access to the right information. For example, retail e-commerce applications must respond to user requests within one or two seconds to help ensure high user retention. Similarly, healthcare applications must capture and monitor data from sensors monitoring hundreds of thousands of patients while processing and responding to critical medical events reliably and predictably without data loss. These applications are evidence that the capability to provide services in real time requires high throughput, predictable latency, high availability, and dynamic scalability from underlying infrastructure.

High Throughput

A common technique for helping ensure high throughput and low latency is to store all information in memory. Yet the massive size and unpredictable quantity of data volumes makes in-memory solutions expensive choices for big data processing. Typically, big data solutions store the vast majority of information on disk, and use memory for caching the most frequently accessed data subsets. As a result, system performance often is limited by the speed of data storage and retrieval. In particular, the number of input and output operations per second (IOPS) delivered by a disk can dictate the performance characteristics of the entire solution.

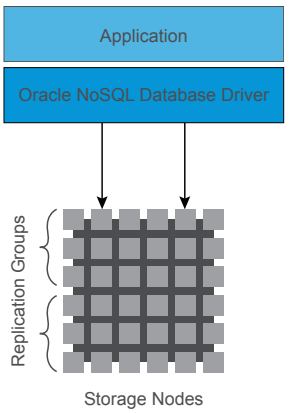

Figure 1. Oracle NoSQL Database Groups Multiple Storage Nodes into Replication Groups for Redundancy

Random Access

Modern spinning disks deliver fast sequential access and poor sustained random performance. Because the storage requirements of big data applications far exceed the capacity of a single disk, high-performance solutions use multiple disks per server to increase I/O bandwidth. Although this approach may be adequate for smaller data sets, the processing of large data sets requires external arrays, with escalating costs associated with the hardware and maintenance required for dynamic system scalability and low latency.

Innovative Business Solutions

In the second half of 2011, Cisco and Oracle joined forces to provide innovative solutions for this growing problem. These industry giants delivered complementary technologies (software from Oracle and hardware from Cisco) to manage and process big data to increase business value. In 2012, Cisco and Oracle extended the partnership to include Fusion-io in an engineering effort to address big data issues in the data center. Working in unison, Cisco, Oracle, and Fusion-io products provide exceptional performance and significant advantages over other open source and proprietary solutions.

Oracle NoSQL Database : Large-Volume Management of Big Data

The Oracle NoSQL (Not SQL or Not Only SQL) Database (Figure 1) is a

distributed, simple key-and-value database for large-volume, latency-sensitive applications or web services that use big data. A shared-nothing architecture with predictable latency support real-time big data applications, delivering transactional semantics for easy data access, high availability, and scalable throughput. Although it is possible to process big data using traditional SQL-based products and solutions, Oracle NoSQL Database provides a more cost-effective, higher-performing, and horizontally scalable alternative. Furthermore, the software complements and can be integrated with traditional SQL-based solutions to provide comprehensive business advantages to the enterprise. Because big data solutions can have many components, commercial-class solutions such as Oracle NoSQL Database are preferred over open source solutions for business-critical applications.

Cisco UCS: High Performance Intelligent Infrastructure

Cisco UCS offers a unified computing, networking, storage access, and management platform that radically simplifies the deployment of big data solutions (Figure 2). Its low-latency, high-bandwidth, 10 Gigabit Ethernet unified fabric and servers with intelligent Intel® Xeon® processors bring industry-leading performance to big data applications. The system is designed with no single point of failure to help ensure data reliability and availability when and where it is needed. This approach is a radical simplification compared to traditional systems, helping streamline data center operations while reducing floor space, power, and cooling requirements, to lower total cost of ownership (TCO).

Fusion ioDrive2: Real-Time Response and Millisecond Latency

The Fusion ioDrive2 solution delivers the millisecond latency big data applications need to maintain real-time response when processing hundreds of terabytes of data—a load that in-memory databases cannot handle. At the same time, Fusion ioDrive2 products provide persistent storage, with capacities ranging from hundreds of gigabytes to multiple terabytes. Extreme I/O performance keeps CPUs busy, a task that disk arrays cannot achieve without many racks of storage and high-bandwidth network infrastructure. In addition, powerful features eliminate reliability concerns. Adaptive Flashback, an intelligent, self-healing feature, provides complete fault tolerance and enables self-repair after a single or multichip failure without disrupting business continuity.

Extreme Big Data Performance

For the test, 15 Cisco UCS C240 M3 Rack Servers were used to run the Oracle NoSQL Database and 15 additional clients ran the Yahoo! Cloud Serving Benchmark (YCSB). A workload with 95 percent read operations and 5 percent update operations was used to measure latency and throughput. Each workload executed ten million operations using a two billion record store that amounted to two terabytes of data. Latency and throughput results, measured by the YCSB clients during

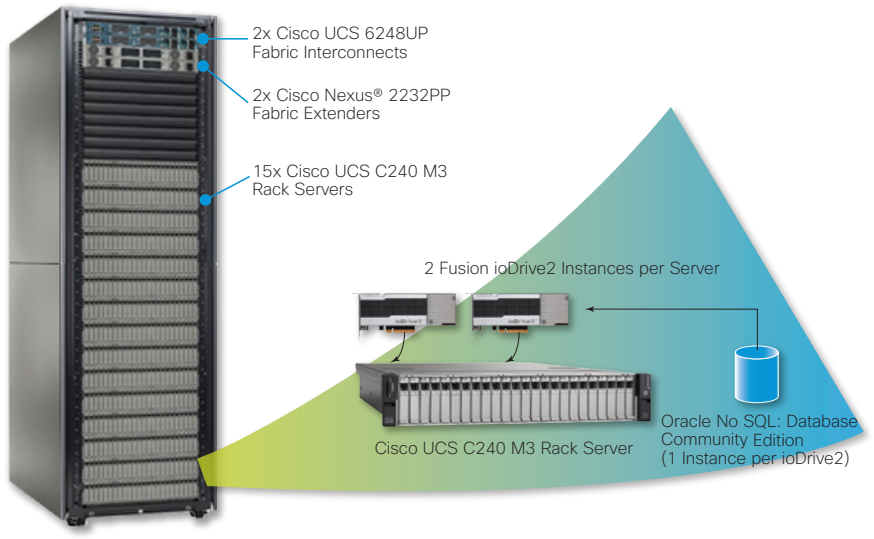

Figure 2. Cisco, Oracle, and Fusion-io Big Data Solution

Table 1. Oracle NoSQL Database Performance

Performance Characteristic	Number of Shards	2	4	8	10
Throughput	Insert throughput (operations per second)	56,634	99,322	186,674	226,277
	Mixed (95% read and 5% update) throughput (operations per second)	302,152	558,569	1,028,868	1,244,550
Latency	Average insert latency (milliseconds)	3.17	3.62	3.85	3.97
	Mixed workload read latency (milliseconds)	0.76	0.79	0.85	0.88
	Mixed workload update latency (milliseconds)	3.08	3.82	4.29	4.47
Scalability	Insert workload scalability ratio	1.0	0.88	0.82	0.8
	Mixed workload scalability ratio	1.0	0.92	0.85	0.82

the tests, are summarized in Table 1. Across 10 shards, the system achieved near-linear scalability for both insert and mixed-workload scenarios.

Read operations require random I/O (seeks) on conventional disks. With Fusion ioDrive2 products, random I/O and sequential I/O speeds are nearly identical. Consequently, the solution offers increased performance compared to conventional hard disk drives and scales nearly linearly while delivering significantly lower latency.

A Winning Combination

The combination of Cisco UCS, Fusion ioDrive2, and Oracle NoSQL Database delivers dramatic performance capabilities to big data applications. The solution delivers fast access to information that enables organizations to glean important insights for guiding business decisions. The fact is, commonly used disk drives cannot deliver the performance or latency needed within a reasonable cost structure. By deploying Cisco, Fusion-io, and Oracle together, companies can use this enterprise-class, high-performance, low-latency, reliable, and scalable big data solution to increase customer and user loyalty and trust, gain competitive advantage, and lower operating costs.

For More Information

- For more information about Cisco big data solutions, please visit http://www.cisco.com/go/bigdata.
- For more information about Cisco UCS, please visit http://www.cisco.com/go/ucs.
- For more information about Fusion ioDrive2, please visit http://www.fusionio.com/products/iodrive2/.
- For more information about Oracle NoSQL Database, please visit http://www.oracle.com/us/products/database/nosql/overview/index.html.

Americas Headquarters
Cisco Systems, Inc.
San Jose, CA

Asia Pacific Headquarters
Cisco Systems (USA) Pte. Ltd.
Singapore

Europe Headquarters
Cisco Systems International BV Amsterdam,
The Netherlands

Cisco has more than 200 offices worldwide. Addresses, phone numbers, and fax numbers are listed on the Cisco Website at **www.cisco.com/go/offices**.

LE-33009-00 09/12