Terracotta and Intel: Breaking Down Barriers to In-memory Big Data Management

Big data solutions have changed the data analysis landscape, and enterprises are increasingly harnessing the power of big data. In-memory data management platforms—a natural progression of the big data revolution—are pushing data management and analytics into the realm of real-time. Use cases that were difficult or impossible five years ago—such as real-time fraud detection and digital ad marketplaces—are now a reality.

Enterprises increasingly need to take immediate action on big data intelligence, which is driving changes to the traditional enterprise architecture. In particular, these enterprises are keeping data instantly available in high-speed machine memory, rather than locking it away in slow, disk-bound databases. Terracotta BigMemory Max*, combined with servers equipped with the Intel® Xeon® processor E7 v2 family, can give enterprises high performance, highly available in-memory data access and management, and predictable speeds at any scale.

This white paper provides an overview of Terracotta BigMemory Max and the benefits of running BigMemory Max on servers powered by the Intel Xeon processor E7 v2 family. The paper describes how Terracotta BigMemory Max and Intel technologies can work together to enable real-time analysis of large data sets.

What's Holding Back Real-time Big Data Analysis?

Since the age of mainframes, enterprises have relied on analyzing structured data stored in databases for making business decisions. Analysis was often slow and complex, and infrastructure performance wasn't up to the task of providing real-time analysis of high-velocity data. While solutions such as Apache Hadoop* have opened up new avenues to capture, store, and analyze data, enterprises that have embraced big data solutions are now grappling with the challenge of managing massive amounts of hard-disk-based data across large server clusters, and the delays that can be caused by latency.

For example, a traditional Apache Hadoop cluster is a distributed network of servers, or nodes, running the Apache Hadoop software. A component of Apache Hadoop is the Hadoop Distributed File System* (HDFS), which distributes and mirrors data across the cluster nodes. Each node stores data on traditional storage devices, such as hard disks or solid-state drives (SSDs). Apache Hadoop optimizes analysis tasks such that the analysis takes place as close as possible to the source data, which can help reduce latency. But while hard disk speeds have improved over the years, hard disk performance still pales in comparison to DRAM or even SSD speeds. Even with an aggressively optimized Apache Hadoop cluster, real-time analysis can be difficult given the performance restraints of traditional storage.

Real-time analysis can also suffer due to the methods many enterprises use to load data into an Apache Hadoop cluster. Enterprises typically rely on traditional relational databases to store information. This data is then batch-loaded into Apache

Table of Contents

Hadoop clusters, often outside of normal business hours so as to avoid impacting the performance of the database. Enterprises that employ this data load model are limited to analyzing data that can be hours or days old, which diminishes the ability to act on quickly changing trends.

Terracotta BigMemory Max* Enables Real-time Big Data Analysis

A combination of factors—including advances in CPU and networking technologies, and the creation of advanced data gathering and analysis tools such as Apache Hadoop and the MapReduce framework—changed the analysis landscape forever. Now, enterprises are pushing the envelope even further with in-memory data management and analysis.

Terracotta provides a high-performance alternative to disk-based storage with Terracotta BigMemory Max. BigMemory Max overcomes the weaknesses of traditional batch-loaded data analysis by storing data in the server's machine memory. Enterprises can use this data storage space to store terabytes of data in machine memory, which is orders of magnitude faster than loading data onto disk-based storage. Storing data in memory provides low, predictable latency and eliminates delays associated with disk access. These features make data available instantly for analysis, with BigMemory Max automatically optimizing memory use with changing data patterns.

BigMemory Max can use as much memory as is physically available on the server, and provides the option of building arrays of interconnected servers to provide an even larger memory storage footprint. With BigMemory Max, enterprises can reduce the amount of time required for complex analysis from hours to minutes, or even seconds, which opens up unprecedented opportunities to quickly turn high-velocity transactional data into actionable business decisions.

Terracotta is already a recognized leader of in-memory data management solutions. CERN, the European Organization for Nuclear Research, is one of the world's leading particle physics research centers. The organization provides scientists from around the world with the ability to conduct high-energy physics experiments using equipment such as the Large Hadron Collider (LHC) particle accelerator. CERN uses a variety of control systems to monitor critical equipment, including the Technical Infrastructure Monitoring (TIM) platform, which monitors approximately 120,000 sensors, and Diagnostics and Monitoring (DIAMON), which monitors some of the equipment related to the particle accelerators.

TIM and DIAMON require a high level of performance and availability. CERN engineers used Terracotta BigMemory Max to build a stable, redundant system that can process incoming data in less than one second. As a component of the CERN Control and Monitoring Platform (C2MON), Terracotta BigMemory Max holds current status data from TIM and DIAMON in main memory, and provides high availability by supporting automatic failover between mirrored servers. The combination of in-memory data storage and high availability improves overall system response times, increases throughput, and helps C2MON achieve 99.99984 percent availability for the TIM platform.[1]

The advantages of BigMemory Max also extend to the financial services industry. A Fortune 500* online payments processor relied upon a platform that could only handle 50 fraud detection rules, and was limited to 2,000 fraud detection transactions per second. Latency was also inconsistent. As a result, the company failed to meet service-level agreements (SLAs) as often as ten percent of the time, costing an estimated $10 million a year in fines and fraudulent charges. To help reduce SLA violations and fraud and improve customer satisfaction, the company built a high performance system with BigMemory Max that can hold over 100 times the number of fraud detection rules in memory, boost transaction throughput by over 30 times to 65,000 transactions per second, and allow the company to approve or flag transactions within 650 milliseconds. The company estimates it will save tens of millions of dollars by avoiding missed SLAs fines and fraudulent charges.[2]

How BigMemory Max Works

Traditional servers typically rely on the venerable concept of hierarchal storage. Operating systems and applications, which need to run as fast as possible, load from slower disk-based storage into a server's high-speed RAM. On the other hand, bulk data—such as documents, media files, or raw data captured from transactional systems—is stored on slower, often less expensive hard-disk-based storage, and only loaded into RAM when required by an application. Loading data from disk into RAM is time-consuming, and can significantly slow analysis tasks.

Now, servers can utilize hundreds of gigabytes, or even terabytes, of cost-effective RAM to store data. Operating systems and applications often only require gigabytes or tens of gigabytes of memory, which can leave large amounts of RAM unused. BigMemory Max configures unused memory as storage space that applications can use to store, retrieve, and analyze bulk data.

By combining multiple servers into arrays, BigMemory Max can create large pools of high-speed memory that applications can use to store, retrieve, and analyze extremely large datasets. In a single-server environment, the BigMemory Max client manages data in the server's local RAM. In a distributed environment, the BigMemory Max client manages data within the server's RAM and also maintains a network connection between the members of the BigMemory Max server array, moving data as needed among the member servers' RAM.

BigMemory Max also protects data and can deliver 99.999% uptime by mirroring the data across multiple servers. RAM is volatile and does not persist data when power is removed. If a server outage occurs, any data in that server's RAM is lost. BigMemory Max pairs two servers in a mirrored configuration where each server keeps a complete copy of the data. If either server experiences a power failure or some other fault, the data is not lost.

How Applications Access BigMemory Max

Applications can access BigMemory Max storage using several methods. A common method is the industry-standard Java* Ehcache API that provides common methods for storing and retrieving data, in addition to advanced search and analysis capabilities. Other common APIs, such as those found in C# and C++ libraries, can also access and manage data in BigMemory Max.

BigMemory Max provides an option for storing data as Java objects, which simplifies application programming and reduces the amount of data transformation overhead that relational databases require. Software engineers can use common protocols such as MOM, HTTP, REST, and SOAP to access BigMemory Max.

Terracotta and Intel: Powering the New Generation of Data Analysis

Built on the Intel 22 nanometer architecture, the Intel Xeon processor E7 v2 family provides key scalability, performance, and reliability enhancements that increase the performance and flexibility of BigMemory Max. These features can help enterprises reduce the complexity and total cost of ownership of their analytical engines while enjoying the benefits of higher performance and uptime.

Scalability

Servers equipped with Intel Xeon processor E7 v2 family provide higher memory capacities than previous generation Intel Xeon processors.

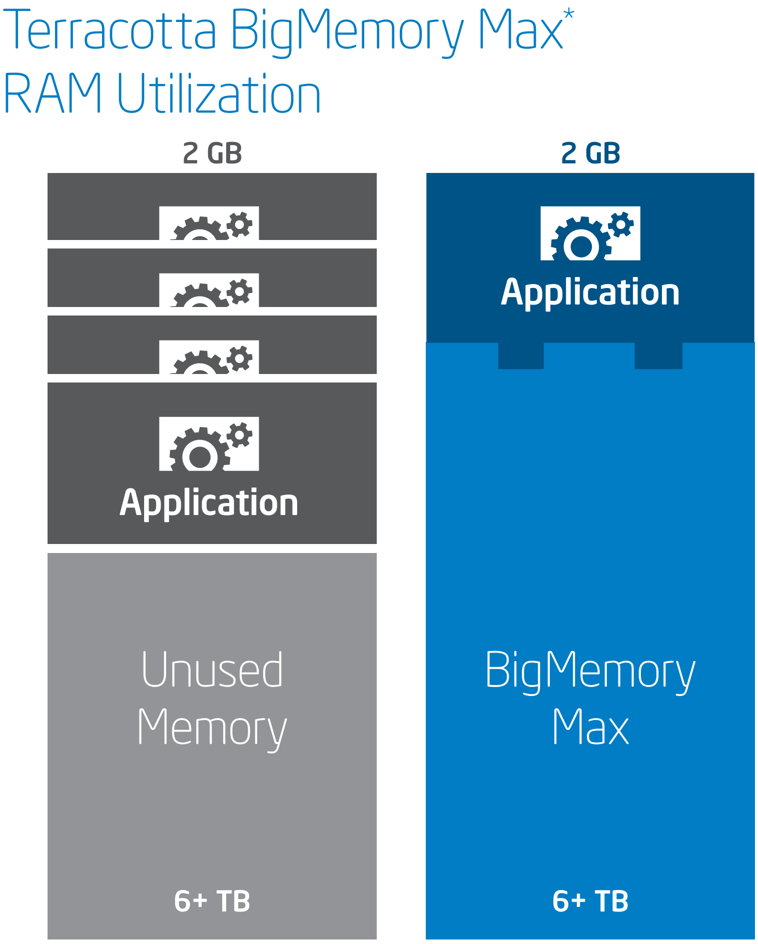

Figure 1. Terracotta BigMemory Max* creates a data store from unused RAM, which is accessible by applications in microseconds

Four-socket configurations can accommodate up to 6 terabytes of RAM and 60 cores, while eight-socket configurations can provide up to 12 terabytes of RAM. More memory provides greater analytical flexibility by allowing applications such as Terracotta BigMemory Max to store and access larger in-memory datasets. In a clustered environment, Terracotta BigMemory Max can scale out in-memory access to petabytes of data.

Performance

Built on Java, Terracotta BigMemory Max can run as a standalone application using only the Java Development Kit (JDK), or on many leading Java application stacks, including IBM WebSphere*, Oracle WebLogic*, and Apache Tomcat*. Intel works closely with Java Virtual Machine (JVM) vendors to increase the performance of their JVMs on Intel hardware. Applications such as BigMemory Max benefit from JVM optimization, which takes advantage of the latest Intel microarchitecture enhancements.

The Intel Xeon processor E7 v2 family also provides a number of key performance enhancements over the previous generation of the Intel Xeon processor E7 family:

- Utilizes a 22 nm process technology, which allows higher operating speeds while reducing power consumption and heat
- Allows for up to 15 cores and 30 logical processors per socket—a 50 percent increase over the previous generation Intel Xeon processor E7 family

Reliability

The Intel Xeon processor E7 v2 family provides enterprises with world-class reliability and uptime. Previous generations of Intel Xeon processors introduced reliability, availability, and serviceability (RAS) features that give servers RISC-like capabilities. The Intel Xeon processor E7 v2 family continues to

Terracotta BigMemory Max* Server Array Configuration

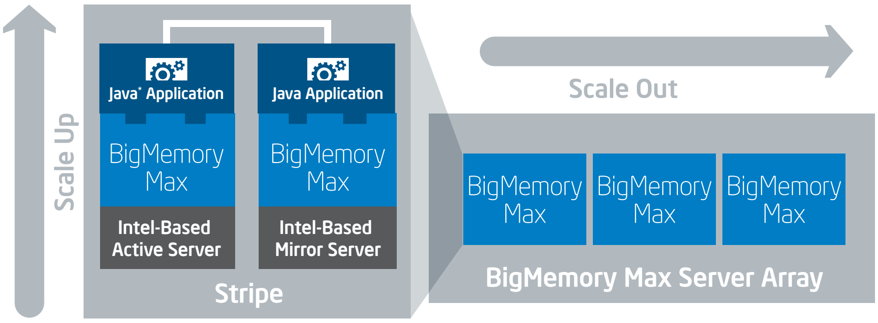

Figure 2. Terracotta BigMemory Max* can pool RAM from multiple servers in an array configuration, and provides high availability and resiliency by mirroring the data across two or more servers in an array configuration

build upon these features to provide even greater uptime and data integrity.

Designed for systems with 99.999% uptime requirements, the Intel Xeon processor E7 v2 family provides continuous self-monitoring and self-healing capabilities that rival those of RISC-based systems.[3]

Some of these features include:

- Machine Check Architecture (MCA) Recovery, which lets CPUs and operating systems compartmentalize errors that could crash the server, such as unrecoverable memory errors.
- MCA input/output (I/O), which provides information on uncorrectable I/O errors to the operating system. The operating system or monitoring tools can then use this information to determine the cause of system errors and enable preventive maintenance.
- MCA Recovery Execution Path, which handles uncorrectable data errors passed to the CPU. This feature enables operating systems and applications to assist in recovering from errors that cannot be corrected at the hardware level.
- Enhanced Machine Check Architecture (eMCA) Gen 1, which provides enhanced logging information to the operating system and applications, which can use this information to better diagnose errors and proactively predict failures.
- Peripheral Component Interconnect Express (PCIe) Live Error Recovery (LER),[4] which provides recovery from and containment of PCIe errors.

Lower Operating Costs

In recent benchmark tests using common Apache Hadoop workloads, servers equipped with the Intel Xeon processor E7 v2 family outperformed servers equipped with the previous generation Intel Xeon processor E5 family by up to 3.5 times.[5] With the Intel Xeon processor E7 v2 family, enterprises can achieve the same or greater benefits by scaling their analytics clusters up instead of out by running fewer, higher-powered servers. An infrastructure that follows the scale-up model can reduce complexity and lower power, management, and cooling costs, all while still providing the computational benefits of a larger cluster running on lower-powered servers. These traits can result in a lower overall total cost of ownership (TCO).

The Performance Advantage of Intel Technologies

One of the most important aspects of high-performance applications is the ability to behave predictably under load at any scale. Minor performance fluctuations on individual servers can become magnified as multiple servers are pooled into larger server arrays.

In recent tests, Intel and Terracotta engineers created an environment to determine the performance and latency predictability of servers equipped with the Intel Xeon processor E7 v2 family and BigMemory Max. The benchmark measured performance with large, in-memory data sets by measuring the total number of transactions per second and transaction latency, which is the round trip time in milliseconds to complete each transaction.

The test configuration consisted of a single server running BigMemory Max 4, and was equipped with four of the Intel Xeon processor E7-4890 v2 and 6 terabytes of RAM. The tests were repeated against multiple BigMemory Max data store sizes of 2 TB, 4 TB, and 5.5 terabytes.[6] Seven client machines were used to produce the workload, and were each equipped with two of the Intel Xeon processor E5-2697 v2. The server and client machines communicated over a 1 gigabit Ethernet network.

The client machines used the Ehcache API to read from and write to the BigMemory Max data store on the server. A client machine bulk loaded data into the data store up to its maximum capacity of 5.5 terabytes. Once the data load completed, each of the seven clients executed up to 128 threads concurrently for a maximum of 896 threads, with each thread representing a transaction between the client and server. To simulate real-world conditions, the tests consisted of a ratio of 90 percent reads from the data store to 10 percent writes to the data store. Each transaction carried an average payload of 2 kilobytes stored as a binary array.

The results in Figure 3 demonstrate how BigMemory Max provides predictable throughput under load as the number of concurrent threads increases. As the number of threads increased, the transactions per second (TPS) increased in a predictable manner. At 224 concurrent threads, the server produced 46,506 TPS, and increased to 63,492 TPS at 896 concurrent threads. The CPU utilization peaked at only 15 percent with seven clients running a total of 896 concurrent threads.

Figure 4 illustrates the scalability of BigMemory Max. The server memory capacity and BigMemory Max data store was increased incrementally to 2 TB, 4 TB, and 5.5 TB. The performance tests were run with each memory configuration. The volume of data used in each test increased in relation to the size of the BigMemory

Terracotta BigMemory Max* Throughput

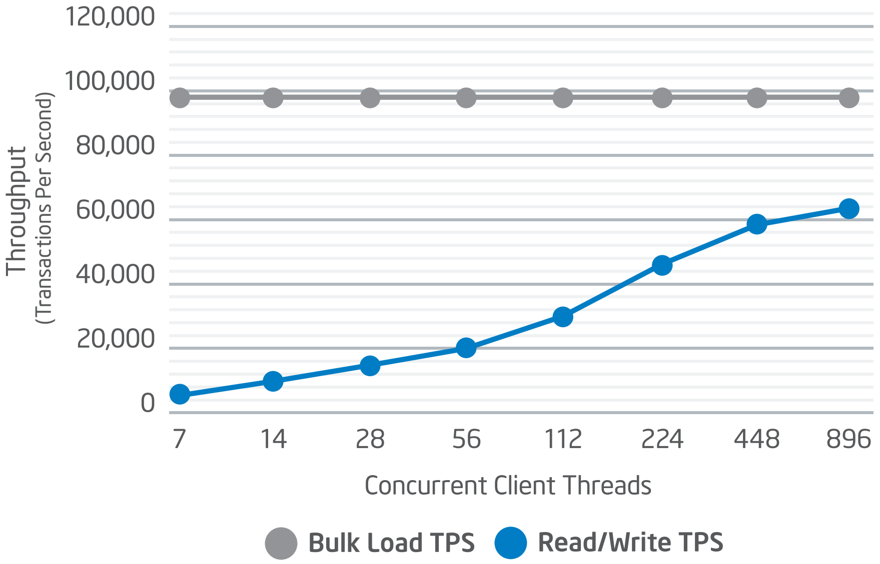

Figure 3: Terracotta BigMemory Max* provides predictable throughput as the number of concurrent client threads increases[7]

Terracotta BigMemory Max* Performance Scaling

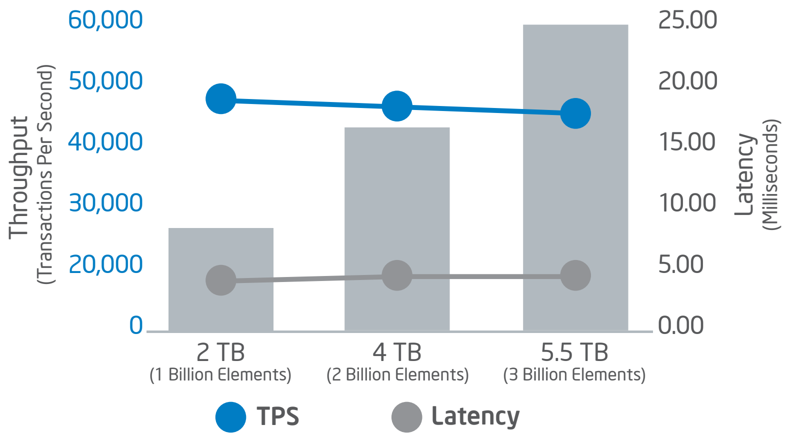

Figure 4. Terracotta BigMemory Max* provides predicable latency and throughput as data volume increases[7]

Max data store from 1 billion, 2 billion, and 3 billion elements, with each element consisting of an average of 2 kilobytes.

The results demonstrate that even as the volume of data increased at each each memory configuration point, latency and throughput remained stable. The latency remained fairly constant at 4.2 to 4.3 milliseconds, while the TPS remained stable between 45,230 and 46,915.7

The combination of BigMemory Max running on the Intel Xeon processor E7 v2 platform easily handled the load generated by the seven client machines producing a maximum of 350 concurrent threads over a 1 gigabit Ethernet network, even when the server memory was scaled to 6 TB. The CPU utilization peaked at 15 percent, which means the server was not being utilized to its maximum capacity. The limiting factor in these tests was the 1 gigabit network and the number of client machines. In a production environment using 10 gigabit Ethernet to connect clients and servers, a single server would be capable of producing higher TPS and handling a higher number of connections. The capacity and throughput can be driven even higher by combining multiple servers into a Terracotta Server Array, as BigMemory demonstrates linear scale out of TPS at a constant low latency rate.

Systems architects can use this predictability to harness the power of BigMemory Max and the high performance and large memory capabilities of the Intel Xeon processor E7 v2 family to scale their in-memory data management systems as their needs require. As demand increases, enterprises can simply add additional compute resources to quickly scale their systems. Architects can also use the performance predictability of BigMemory Max and the Intel Xeon processor E7 v2 family to design server arrays with specific performance and latency goals.

Conclusion

By harnessing the power of Terracotta BigMemory Max and servers equipped with the Intel Xeon processor E7 v2 family, enterprises can build a foundation that enables data analysis use cases that were, up until recently, difficult or impossible to implement. Intel and Terracotta engineers demonstrated that the combination of BigMemory Max and the Intel Xeon processor E7 v2 family can utilize up to 6 TB of memory on a single server and deliver consistent, predictable throughput and latency.

In-memory data management and analysis is the next step in big data evolution. BigMemory Max running on servers equipped with the Intel Xeon processor E7 v2 family provides a path for enterprises to scale up and out, which opens up new avenues to reap the benefits of real-time big data analysis.

To learn more how Intel can accelerate your big data solutions, visit www.intel.com/bigdata.

For more information about Terracotta BigMemory Max, visit www.terracotta.org/products/bigmemorymax.

[1] For more information, please visit https://www.softwareag.com/corporate/images/SAG_Terracotta_CERN_RS_Jan14_WEB_tcm16-116352.pdf.

[2] For more information, please visit http://terracotta.org/resources/success-stories/financial-services/better-faster-fraud-detection-drives-higher-profits.

[3] No computer system can provide absolute reliability, availability, or serviceability. Requires the Intel® Xeon® processor E7-8800/4800/2800 v2 product families or an Intel® Itanium® 9500 series-based system (or follow-on generations of either). Built-in reliability features available on select Intel® processors may require additional software, hardware, services and/or an internet connection. Results may vary depending upon configuration. Consult your system manufacturer for more details.

[4] PCIe Live Error Recovery implementations are specific to an OEM's implementation. Consult your OEM for specific recovery capabilities.

[5] For more information see the paper "Accelerate Big Data Analysis with Intel® Technologies" available at www.intel.com.

[6] With the server hardware maximum of 6 TB of RAM, the BigMemory Max data store was configured for a maximum of 5.5 TB instead of 6 TB in order to leave enough RAM available for the operating system. 6 TB of memory achieved using Micron* 64 GB load-reduced DIMMs (LR-DIMMs) built with Micron's latest 4Gb quad-die, stacked DRAM components. For more information, see http://www.micron.com/products/dram-modules.

[7] Based on Intel Corp. internal testing. The test environment consisted of a single server and seven client machines. The server was configured with four of the Intel® Xeon® processor E7-4890 v2 product family running at 2.2 Ghz, 96 x 64 GB DDR3-1333 LRDIMM (6 TB total), 2 x 10 gigabit Intel® Ethernet Controller X540-AT2, 2 x Seagate 300 GB 15K RPM SAS hard drives in a RAID 0 configuration, and an LSI SAS2308 RAID controller. The operating system was Red Hat Enterprise Linux* 6.4, and the software used was Terracotta BigMemory Max* 4.0.5, Java* 1.7.0_25, Java SE Runtime Environment (build 1.7.0_25-b15), and Java HotSpot* 64-Bit Server VM (build 23.25-b01, mixed mode). The client machines were configured with two of the Intel Xeon processor E5-2697 product family running at 2.7 GHz, 8 x 4 GB 1333 MHz memory modules, a 1 TB 7200 RPM SATA 6 GB/s 2.5" hard drive, and an Intel® Ethernet Controller I350 gigabit network interface controller. The operating system was Red Hat Enterprise Linux 6.4, and the software used was Ehcache* 2.8.0, Java 1.7.0_25, Java SE Runtime Environment (build 1.7.0_25-b15), and Java HotSpot 64-Bit Server VM (build 23.25-b01, mixed mode). The clients used Ehcache 2.8.0 to read from and write to the BigMemory Max. Two benchmark tests were run against the configuration. The first test measured compute scaling where the number of threads on each client was increased from 1 to 128, with a fixed data size of 500 GB. The second test demonstrated data scaling where the client thread count was fixed at 50, while the data volume was increased to fill the capacity of each memory configuration.

Printed in USA 0114/TA/PRW/PDF Please Recycle 330106-001US