White Paper

A Revolution in Data Management: Oracle Exadata X8M

Sponsored by: Oracle

Carl W. Olofson
April 2020

IN THIS WHITE PAPER

This white paper reviews the significant improvements to Oracle Exadata Database Machine through the advent of its latest version: X8M. The white paper describes the advanced technologies incorporated in the system and their impact on both the transaction throughput and query performance of Oracle Database running on Oracle Exadata X8M. It also discusses the significance of this new version of Exadata for business users.

SITUATION OVERVIEW

Oracle Exadata Background

Oracle initially introduced Exadata in October 2008 as an integrated compute and storage hardware system designed specifically to support Oracle Database with capabilities made possible by a custom configuration of industry-standard infrastructure enhanced by complementary software. This first version was built using hardware made by HP Inc. It included compute nodes performing complex SQL query and transaction processing and storage nodes with internal hard disks storing the data. These transaction processing and storage nodes would respond to simple SQL queries to retrieve and update the data. The two layers communicated through a dedicated InfiniBand private network. Using dedicated computers with internal storage as the storage resource enabled Oracle to make the communication between compute and storage much simpler and more efficient than has been the case with conventional do-it-yourself (DIY) database platforms built using standalone general-purpose servers and storage. This is because the Exadata storage nodes only return data relevant to a query and not whole blocks of data from files, as is conventionally done. The private network in Exadata also enabled its database servers and storage servers to communicate much faster than its DIY equivalents because it eliminated the competing network traffic that commonly slows database performance.

Over the years, Oracle has greatly improved Exadata, first by switching from the hardware made by HP Inc. to hardware produced through the newly acquired Sun Microsystems, and then by improving the efficiency of communication between the compute and storage layers. Oracle also built in the same in-memory columnar hybrid compressed format, with vector processing (optimizing data entry into the processor to exploit the speed of the processor cache combined with single instruction multiple data [SIMD] execution), that is featured in the Oracle Database In-Memory Option. Finally, Oracle added the storage option of all flash and moving data using the InfiniBand remote direct memory access (RDMA) capability, coupled with NVMe modules front ending the flash, for still greater performance.

Improvements in Exadata X8M

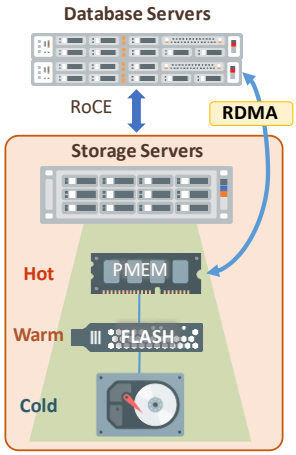
The latest member of the Exadata portfolio includes all of the machine learning and artificial intelligence (AI) capabilities of Exadata X8 and features a new high-performance and cost-effective memory and interconnect architecture. Exadata X8M is one of the first scalable database infrastructure platforms in the world to utilize Intel Optane DC persistent memory modules (PMEMs). X8M uses PMEM as another data tier, with data exchange rates similar to direct random-access memory (DRAM) but with four times the capacity per module and the persistence capability. It also speeds the system by accepting transaction log records at a much greater rate than was possible with flash.

Oracle implemented PMEM in Application Direct Mode, together with RDMA over a 100Gbps Ethernet interconnect, or RDMA over Converged Ethernet (RoCE). This combination enables many database read, write, and logging activities to bypass the entire operating system (OS), I/O operations, and network software stacks, eliminating expensive CPU interrupts and context switches and reducing SQL read latency by over 10 times, from 250µs to less than 19µs. In contrast to PMEM implementations that require customers to change their applications to take advantage of the PMEM capabilities, the Oracle Database team made all the required changes within Exadata X8M and Oracle Database. Users of Oracle Database 19c can take advantage of the higher throughput and lower latency available with Exadata X8M immediately, without making any application changes.

For those who like to segregate Exadata workloads using virtual machines (VMs), Exadata X8M now supports the Linux Kernel Virtual Machine (KVM), making those VM environments more compatible with third-party software that interacts with VM functionality. This architecture is illustrated in Figure 1.

FIGURE 1

Oracle Exadata X8M Architecture



Scale-out design

- 2-14 Database Servers per rack
- 3-18 Storage Servers per rack
- Up to 18 racks per system

Shared Persistent Memory (PMEM) performance, redundancy, and security

- Aggregate performance of all PMEM available to all database servers
- Automatic PMEM mirroring across storage servers transparently protects from PMEM failures
- No direct access to PMEM from DB hosts, only databases can access

RDMA over Converged Ethernet (RoCE) latency and throughput enhancements

- Enables use of low-latency PMEM by bypassing the OS, I/O, and network stacks
- Zero packet-loss networking
- Prioritization of critical messages

Source: Oracle Corp., 2020

RDMA over Converged Ethernet

Most database management system (DBMS) implementations involve a compute layer and a storage layer that communicate over a conventional Ethernet network. This means that they must exchange

packets and conform to Ethernet packet switching protocols, which add considerable overhead to their communications. These systems are often from different suppliers and are not optimized to work together or to run Oracle Database. In contrast, even the earliest Exadata machines were much more efficient: co-engineered with the Oracle Database, with a private InfiniBand network connecting the database servers and the storage servers — all specialized for Oracle Database data access and processing.

Oracle Exadata also uses RDMA to make communications between server layers simpler still. Instead of using a message protocol and sending packets to be processed by each other's CPUs, RDMA allows each system to read from and write to each other's memory directly without involving the CPU. This eliminates the message switching and CPU processing overhead. The result is much faster and cleaner processing. Oracle Exadata X8M takes this a step further by using RoCE instead of InfiniBand for its internal networking. The use of RoCE gives Oracle Database a significant performance boost due to both its high throughput and low latency, enabling the Exadata X8M components to communicate across a faster 100Gbps network.

Intel Optane DC Persistent Memory

The most important aspect of the new Exadata X8M machine is its use of Intel Optane DC persistent memory (PMEM). This delivers several important benefits. Each PMEM fits in a standard memory expansion slot, and it can be accessed as ordinary memory, in the usual manner, or as managed persistent memory using App Direct Mode.

On the transaction side, there is a significant boost in throughput. Most memory-optimized database systems use flash drives to store transaction log records. A transaction cannot be considered fully committed until the log record has been successfully written. Traditionally, log-write completions are dependent on storage I/O latency, which makes it the single biggest drag on transaction performance. Using PMEM on Exadata X8M for log-write commits instead of flash enables Oracle Database to process transactions at a much higher rate: 2.5 times faster and 10 times lower latency than with the previous generation Exadata.

Oracle Exadata X8M uses PMEM as an active data tier that is as much as 15 times faster than flash storage. It also offers four times as much memory as a conventional DRAM module. Because Oracle can keep some active data in PMEM instead of DRAM with no noticeable performance impact, restarting the system can be much faster, since the in-memory data cache does not need to be reloaded.

When used in App Direct Mode, as the application connects to the module to send data, it gets an encryption key (all data put in PMEM is encrypted). If the system is shut down and restarted for any reason, presenting this key to the module enables the application to gain access to the data. The key is not lost when the system stops running, and it is secured during the shutdown.

With this state-of-the-art technology, Exadata X8M delivers much higher Oracle Database performance for both transactional and analytical workloads than does Exadata X8. According to Oracle, Exadata X8M's sub-19µs latency when running Oracle Database 19c, and its I/O rate of over 16 million IOPS for SQL 8Kb reads, combine to deliver substantial application-level improvements for OLTP workloads. In addition, Oracle reports that Exadata X8M features analytic scan throughput of 560GBps and over 1TBps analytic scans with columnar data in flash, enabling complicated real-time analytics. Exadata X8M is currently only available on premises.

Linux Kernel Virtual Machine

Previously, Oracle had offered the ability to segregate workloads using Oracle Virtual Machine (OVM), but users were limited in terms of their use of third-party software in this context. With Exadata X8M, Oracle now supports the Linux Kernel Virtual Machine, enabling a broader range of software to be used within the Exadata system.

Benefits

Although the improvements to Exadata X8M are technical in nature, the benefits to users can be clearly understood in business terms.

Accelerated Transaction Rates

Oracle reports that Exadata X8M delivers 2.5 times faster I/O and 10 times lower latency than Exadata X8 when running Oracle Database 19c. With this level of OLTP performance, enterprises are able to process more business transactions on Exadata X8M without needing to increase the system's size. This is important given the accelerating pace of business and the demands on transaction processing being made not only by traditional ERP and sales applications but also by new use cases such as network intrusion detection, high-frequency stock trading, IoT data processing, real-time fraud detection, and the variety of customer-facing applications on the web and mobile devices.

For example, for IoT workflows, the power of PMEM enables a much larger cache and accelerates transactions by using PMEM to capture log records, enabling Exadata X8M to ingest hundreds of millions of IoT data points per second. Exadata X8M can also accelerate queries to obtain greater IoT data insights.

An eighth rack Exadata X8M, which delivers over 1 million IOPS of SQL 8Kb reads, starts at $180,000, bringing a new level of performance, scale, and security, making these new use cases available to a broader customer base as well.

Shorter Time to Insights

Oracle puts Exadata X8M's analytic scan throughput per rack at 560GBps, which increases to over 1TBps per rack with columnar data in flash. Because of this, complex queries are more quickly and easily processed on Exadata X8M. The result is that both business analysts and data scientists can crunch the numbers without long wait times or degrading other processing on the same system.

More Power to Guest VMs

Users seeking to create guest VMs without complicating the system processing and third-party software will work without modification, thanks to the use of KVM in the Exadata X8M system. This enables them to use standard technology to maximize database isolation, security, and resource management in multitenant environments.

Greater Availability and Consolidation for Lower Management Cost

In Exadata X8M, if any storage server is taken offline or shut down for maintenance, the PMEM components retain their current data, enabling much faster restart times. Exadata X8M is also simpler to configure and manage than DIY infrastructure, and it can leverage its scalable PMEM capability, with up to 21TB per rack, to create greater memory sizes and increase Oracle Database consolidation efficiency while requiring no application changes. As such, Oracle Exadata X8M enables enterprises to replace large, multivendor DIY infrastructures with fewer and smaller Exadata systems that

dramatically reduce complexity as well as datacenter footprint and power requirements. In addition, Oracle Database components, such as Automatic Indexing (a technology from Oracle Autonomous Database) and AI-based root cause analysis, are exclusively available on Oracle Exadata and use its built-in machine learning capabilities to reduce time spent on tuning databases for critical business applications and troubleshooting performance issues.

For example, with Automatic Indexing, Oracle found that an Exadata system running Oracle Database19c created 6,000 application indexes for NetSuite software in only 24 hours. These indexes replaced 9,000 indexes created manually by NetSuite DBAs over a 15-year period and delivered superior application-level performance. This demonstrates the business value of Exadata X8M's machine learning capabilities, which enable companies to focus on more strategic projects while the system and Oracle Database tune themselves.

More Secure

In contrast to DIY infrastructures, Exadata X8M is a single integrated system that includes servers, storage, networking, and operating system. These capabilities are managed and patched together, so customers can rapidly update their systems security without time- and resource-intensive testing. Customers can also thereby avoid applying patch sets from multiple component vendors. As such, Exadata X8M significantly reduces surface exposure to cyberattacks and eases timely patching with regular single bundle implementations. In addition, Oracle Database encrypts data end to end, so it is protected at rest, in motion, and even while stored in data protection and archive environments.

FUTURE OUTLOOK

It is likely that other memory-optimized database systems will take advantage of the benefits of PMEM in the coming years, but the bespoke design elements of Exadata, including its Smart Scan query offload capability and the very precisely designed use of RoCE for inter-server data transfer, serving the exact requirements of the Oracle Database, will continue to offer a differentiating advantage for Exadata in terms of power and performance. The benefits of Exadata X8M and its successors will apply to users using both on-premises and cloud deployments.

CHALLENGES/OPPORTUNITIES

The database market, at the high end, is a marathon, not a sprint. As users calculate the benefit-to-cost ratio of database systems, some with less demanding requirements will, no doubt, choose less potent on-premises solutions or move to the cloud.

For those who need a scalable blend of transaction and analytic processing with open standard capability at the high end, or who just want an efficient platform onto which to consolidate their Oracle Databases to improve security and operations, Oracle Exadata X8M sets a high bar at present, but one that will continue to be challenged by competitors seeking to win more business in the high-end database management market.

CONCLUSION

Oracle Exadata X8M is not only a major step forward for Oracle Exadata and its users but represents the future of database management systems. That future includes much faster transaction processing

than was possible before, combined with faster queries thanks to larger in-memory data collections, and very rapid database recovery from any outage, planned or unplanned, particularly on the analytic side.

Enterprises that are considering running Oracle Database on generic cloud infrastructure or DIY configurations need to be aware that such environments do not deliver anywhere near the high throughput, low latency, and simplicity offered by Oracle Exadata. With Oracle Exadata X8M, those advantages are taken to another level.

IDC research indicates that the move is on to bring database applications to the cloud. Those evaluating Oracle Database on generic cloud infrastructure need to ensure that those database servers can deliver the performance and resiliency they need. They also need to ensure that the database servers are close enough to their middleware and application platforms, so speed-of-light delays don't impact end-user performance and that their total cost of ownership is in line with their expectations. If generic infrastructure cannot meet those requirements, and the enterprise still wants to move to the cloud with the benefits of Oracle Exadata X8M, then it should consider some combination of Oracle's high-performance Autonomous Database, Exadata Cloud Service, and Exadata Cloud at Customer offerings as alternatives.

About IDC

International Data Corporation (IDC) is the premier global provider of market intelligence, advisory services, and events for the information technology, telecommunications and consumer technology markets. IDC helps IT professionals, business executives, and the investment community make fact-based decisions on technology purchases and business strategy. More than 1,100 IDC analysts provide global, regional, and local expertise on technology and industry opportunities and trends in over 110 countries worldwide. For 50 years, IDC has provided strategic insights to help our clients achieve their key business objectives. IDC is a subsidiary of IDG, the world's leading technology media, research, and events company.

Global Headquarters

5 Speen Street
Framingham, MA 01701
USA
508.872.8200
Twitter: @IDC
idc-community.com
www.idc.com