

White Paper

Oracle Exadata X8M Leaves Competitors Behind in Running Oracle Database

Exadata X8M Boasts In-memory Performance, RDMA Fabric, and Shared Storage for an Unmatched Oracle Database Platform

By Mark Peters, Principal Analyst & Practice Director

September 2019

This ESG White Paper was commissioned by Oracle and is distributed under license from ESG.



Contents

Introduction and Executive Summary: The End to DIY Database Architecture Headaches	3
The Continuous Need for More Database Speed	3
Introducing the Exadata X8M	4
RDMA over Converged Ethernet (RoCE)	5
Persistent Memory	6
Virtualization with KVM	6
The Bigger Truth.....	6

Introduction and Executive Summary: The End to DIY Database Architecture Headaches

Oracle Database is a powerful piece of software; it is foundational to many businesses and to all our lives. It's glib but true that "operating at the speed of business" has never been more important or more difficult because today's required speed of business is preferably instantaneous, and delivering that blistering speed invariably requires more steps, such as calculations, look-ups, and I/O, than even a few years back.

Oracle Database sits bang in the middle of all this, as a mission-critical component of myriad organizations' abilities to compete and to deliver. Oracle has been delivering multiple integrated platforms to run its Database for a while now; its flagship Exadata just celebrated a decade in the market and (especially the latest X8 version) is an extremely powerful and operationally sophisticated "Engineered System." But there's still a sizeable group of Oracle Database users who are always stretching for more; they are making large investments in terms of money, time, and energy and generating more than a few headaches to build their own nodes or pods to drive their Oracle Database ever faster. But now Oracle has added the Exadata X8M to its lineup, meaning that these organizations' need to build DIY Database platforms, and suffer the associated headaches, is gone.

The specifications and capabilities of the X8M are truly a step change and category-shaping. The serious answer as to why it is called X8M is that the "M" stands for "memory." The far more flippant interpretation could be that Oracle chose it because X8 "OMG" would be a tad too colloquial for its restrained corporate tastes (though no less true). With latencies (of 19 microseconds) that *start* at 5 times better than the most popular competitor, it is genuinely hard to view most alternatives as legitimate competition in this market-segment.

As with many of the best advances in IT, the essence of the X8M is not hard to explain; it can seem counterintuitive, but often the more evolutionary and smaller steps can take reams of explanation because they are based on the delicate and arcane application of small technology increments. Conversely, X8M succeeds by unifying the potential of multiple technology heavyweights, most notably an RDMA fabric and persistent memory, optimized via new virtualization software. Let's dig a little deeper.

The Continuous Need for More Database Speed

The speed at which an organization can validate a transaction or react to a request has an immediate impact on business success and customer satisfaction. And with customer experience being so crucial (in a time where organizations are digitally transforming to put the customers' needs first), it's no surprise that the performance and efficiency of the entire infrastructure stack is critical.

- **Throughput and latency** – the ability to handle more complex workloads at scale. Whether it is measured as more transactions/sec or processing queries in microseconds instead of milliseconds, organizations seek the latest and greatest technology to offset performance burdens or concerns.
- **Performance from an operational standpoint** – improving the efficiency of common workflows and processes to ensure that the high levels of performance are maintained. Relevant metrics here include: How quickly can staff react to a new request? Maintain uptime? Ensure secure access to data? Reduce risk exposure and maintain compliance through reliable and timely upgrades?

Whether it is a retailer looking to improve customer experience by completing orders faster, or a high-frequency trading house looking to reduce latency between transactions, the faster the underlying infrastructure can respond to requests, the more value it delivers to the business. This need for speed is supported by ESG research:¹

¹ Source: ESG Research Survey, *Enterprise Database Trends*, January 2017.



35% of organizations cite meeting database performance requirements as a top challenge faced by their organization when it comes to their current database environment and supporting infrastructure.

For organizations considering a new database and/or underlying infrastructure to support it, **performance is the most important attribute.**

With in-memory databases growing in popularity, the idea of leveraging persistent memory closer to the application is more appealing than ever before. While misconceptions and concerns remain about the best ways to leverage new in-memory technology, and whether applications will require modifications to utilize and benefit from it, one thing remains true: If the performance gains give the organization a competitive edge, it will look for any way to get there, and vendors that can offer a shortcut to affordably, credibly, and easily get there will come out on top.

When those sorts of advantages are *not* available from the vendor community, some users will invest heavily (in time, money, and IP) to try to build their own. But running to your local car-parts store to try to replicate a Bugatti or an AMG is unlikely to be successful at all; unless you are Lewis Hamilton, you won't have a legion of Formula 1 AMG engineers behind you. Similarly, individual users won't have the R&D budget, source code, or legion of engineers that Oracle does. Most users would simply end up over-spent, missing SLAs, unsatisfied, and with a literal, virtual, and system headache. Exadata X8M can thus be seen as Oracle delivering the Database super-car that's needed for the racetrack of contemporary business. In this instance, buy beats build, both operationally and financially.

Introducing the Exadata X8M

Before launching into the advanced capabilities of the X8M, it's worth quickly reviewing the foundational elements of the Exadata family, which still apply. Oracle's own vision for Exadata is a good summary.

ESG's take is that Exadata's value proposition is easy to convey: It is a choice of optimized platforms on which to run Oracle Database. While Oracle Database can run on myriad other systems, the optimal experience is running it on Exadata. It is an Oracle "Engineered System" that is tuned to be best for Oracle Database and integrates unique value-add features that are unavailable from other vendors. As such, *it is not general purpose, nor simply fit-for-purpose*: It is purpose-built. Over its decade in the market, Exadata has grown as the—identical everywhere—foundation for on-premises, Oracle Cloud, and Oracle Autonomous Database operations. Oracle states that Exadata is run by 77% of the Fortune Global 100 (25% of which have also adopted Exadata Cloud Service).²

Exadata systems provide an ideal Database platform that combines scale-out, Database-optimized compute, networking, and storage *hardware* (to deliver both performance and economy) with smart *software* in the form of specialized algorithms that can vastly improve all aspects of business-critical OLTP applications.

The Exadata X8M stays true to those roots, but adds impressive new capabilities:

² Paragraph modified from ESG White Paper, *Oracle Databases Achieve Unprecedented Speed Running on the New Exadata Database Machine X8*, published June 2019.

- Exadata X8M looks to provide organizations with an easy to understand approach to gaining fast in-memory performance and ultra-low latency in a cost-effective way through the use of new RDMA fabric, persistent memory, and new virtualization software.
- Compared with the X8 and its InfiniBand fabric, the X8M will offer 100Gb RDMA over Converged Ethernet (RoCE) as the internal fabric to deliver latency of under 19 microseconds (10X improvement over the X8).
- With 1.5 TB of persistent memory (PMEM) per storage server and up to 21.5 TB of PMEM per standard full rack, organizations can achieve up to 16 million OLTP 8K read IOPS, 2.5X the X8.
- Log writes gain a significant improvement to reliably complete more transactions faster and at scale.
- By utilizing Linux KVM, organizations gain a cost-effective virtualization layer than enables better consolidation without sacrificing performance.

There's a carefully curated abundance of brute force, sophistication, and intelligence in the X8M. The following are highlights of the three main new technology areas.

RDMA over Converged Ethernet (RoCE)

Using RoCE3 is all about speed and performance combined with risk reduction and reliability. RDMA provides extreme performance, enabling access to data on a remote computer without the overhead of the OS or CPU. The network card directly reads/writes memory with no extra copying or buffering and at very low latency. RoCE on Exadata X8M supports all Exadata RDMA optimizations and the benefits are significant.⁴

However, it's important to note the additional benefits on top of performance that can be gained by leveraging RoCE; for example, organizations can benefit from automated "class of service" abilities, as well as "messaging prioritization," which ensures that the higher priority messages that require low latency are not impacted by high-throughput workloads. So, if the cluster is running a batch request or doing a backup (high-throughput workloads), committing a transaction or ensuring the cluster is up and running via heartbeat checks are not slowed or impacted.

Looking away from just the performance aspects, Exadata RoCE can help mitigate risks. For instance, Exadata RoCE utilizes a feature called Priority-based Flow Control, which can detect the speed of incoming packets. If the packets are coming too fast or unable to be serviced, the feature tells the sender to pause until the packet buffer has caught up. With Explicit Congestion Notification (ECN), fast packet senders can be detected and forced to slow down to ensure no packets are dropped. Exadata also has a unique ability to leverage RoCE for instant failure detection; traditionally, detecting a failure has been determined by a long timeout, making it hard to detect a slow response versus a server failure, but now Exadata uses RDMA and frequent heartbeat messages between nodes to detect possible failures at sub-second speeds (plus, RDMA uses hardware, so remote ports respond even if the software is slow⁵).

³ RDMA over Converged Ethernet (RoCE) is a protocol that runs InfiniBand RDMA software on top of Ethernet.

⁴ For the more technical reader, these include (Oracle-provided list):

1. High throughput and low-CPU usage for large data transfers.
2. Unique Direct-to-Wire Protocol to deliver 3x faster inter-node OLTP cluster messaging.
3. Unique Smart Fusion Block Transfer that eliminates log write on inter-node block move.
4. Unique RDMA protocol to coordinate transactions between nodes.
5. Ultra low-latency I/Os to persistent memory in storage servers.
6. Ultra low-latency writes of database logs to persistent memory in storage servers.
7. Ability to unify cluster switching and client switching operations and management to save costs.

⁵ How it works: Four RDMA reads are sent to the server suspected of being down across all combinations of source and target ports. If all four RDMA reads fail, the server is evicted from cluster.

Persistent Memory

Oracle has incorporated Intel Optane DC persistent memory (PMEM) in the Exadata X8M. Unlike traditional DRAM, the persistent memory enables data to survive power failures, read like DRAM, and persist writes through the utilization of intelligent software. The PMEM accelerator sits in front of the flash, adding another tier of capacity and performance. With the Oracle Database using RDMA instead of regular I/O protocols to read/write to PMEM, it can achieve **2.5x higher IOPS throughput compared with the recently introduced X8 Exadata, and 10x better latency.**

Also noteworthy: PMEM automatically tiers and shares data across databases, helping to increase the effective capacity by 10x; for log writes, the PMEM accelerator can address the criticality of low latency to improve transaction commit times; and, by bypassing the software and network stack, a one-way RDMA writes to PMEM across multiple storage servers, which can show log write performance improvements of up to 8x.

Virtualization with KVM

Exadata X8M is standardized on KVM-based virtualization. While this is mainly due to RoCE and persistent memory only being supported on KVM, it turns out to be advantageous; it delivers 2x more guest VM memory (1.5TB/server), 50% more guest VMs per server, and faster network latency.

The Bigger Truth

Yet again Oracle has raised the performance bar in running its eponymous Database. Exadata X8M takes the DNA of the successful and proven Exadata and applies a jaw-dropping set of enhancements to an extent that can only be described as game-changing. Users can enjoy the ultra-low latency of persistent memory combined with the IOPS of flash at pretty much the cost of disk and adding 100Gb RoCE means all this horsepower can be applied to drive higher net throughput for operations that need it. This, for example, can mean improving OLTP workloads with 10x lower I/O latency, 2.5x more IOPS, and super-fast log writes. It enables improved mixed workload performance for both OLTP and data warehouses, and it frees CPU loading on the storage server for workloads associated with data warehousing, such as smart scans. As stated earlier, Exadata X8M can be seen as *the Database super-car that's needed for the racetrack of contemporary business*. And, viewed thus as today's ultimate Database Engineered System, it obviates the need for a competitive analysis section in this paper...as to attempt one would be disingenuous and unfair to both Oracle and those that are a lap or two behind (whether that's the other system vendors, or self-assembly efforts).

Yet all the available power does not complete the story. Three remaining factors are vital to understanding the practical user value of X8M, beyond the amazing "specs." First, it is implemented in a way that mitigates and precludes risk; secondly, Oracle has done all the heavy lifting because all the benefits listed above can be achieved without users needing to change their applications; and last but not least, X8M means that users no longer need to stay up all night trying to build their own super-Database platforms. Now, there is a SKU for that.

All trademark names are property of their respective companies. Information contained in this publication has been obtained by sources The Enterprise Strategy Group (ESG) considers to be reliable but is not warranted by ESG. This publication may contain opinions of ESG, which are subject to change from time to time. This publication is copyrighted by The Enterprise Strategy Group, Inc. Any reproduction or redistribution of this publication, in whole or in part, whether in hard-copy format, electronically, or otherwise to persons not authorized to receive it, without the express consent of The Enterprise Strategy Group, Inc., is in violation of U.S. copyright law and will be subject to an action for civil damages and, if applicable, criminal prosecution. Should you have any questions, please contact ESG Client Relations at 508.482.0188.



Enterprise Strategy Group is an IT analyst, research, validation, and strategy firm that provides actionable insight and intelligence to the global IT community.

© 2019 by The Enterprise Strategy Group, Inc. All Rights Reserved.

