Author: Marc Staimer

Oracle Exadata X8M:
Game Over for DIY Database Systems

Oracle Exadata X8M: Game Over for DIY Database Systems

Executive Summary

There has been a tech battle that's raged since the beginning of the electronic age. It's between the vendors that deliver complete integrated, engineered, implemented, and managed turnkey systems and the do-it-yourselfer (DIY) commodity components approach. DIY implementers believe they can match or beat any turnkey systems functionality and performance at a much lower cost if they build it themselves with the same off the shelf components.

Back when stereo systems first became available to the general public in 1970s, many DIYs built their own from components. This repeated itself when the PCs became standardized on Intel chips, off-the-shelf HDDs, networking, and Microsoft DOS back in the 1970s and 1980s. It repeated again in the late 1990s and early 2000s as servers became standardized on similar architectures, Microsoft Windows and open source Linux. Building their own was as easy as buying the commodity components, software, etc., and putting it all together. All it took was knowhow, time, and sweat equity. As the Internet proliferated and services such as Google's YouTube™ sprang up, how-to videos made DIY even easier.

Then why is it that few if any DIYs build their own stereo systems, PCs, or servers anymore? Even the hyperscalers merely design the products they use and have someone else manufacture them. So, what happened? The turnkey products kept becoming more sophisticated with designs, capabilities, and functionality not duplicatable by the DIY system. The DIY implementers discover they can't match the turnkey system functionality, innovation, and many attractive features. Mass production delivers cost points they can't get close to equivalence.

At some point it dawns on the DIY that it takes them more time to procure, design, build, maintain, and support their roll your own creation than the turnkey system. They also realize that they can't keep up with turnkey vendor's innovations and functionality. And ultimately, they recognize it's costing them a lot more than they anticipated and more than the turnkey system.

This battle is now playing to its conclusion with database systems. IT managers/architects have constructed their own DIY database system for years. That's primarily because the database software suppliers rarely provided turnkey database systems. These systems were few and far between and quite expensive. The IT managers were generally forced to be the DIYers. They loaded database software onto servers, virtual machines (VM), containers, and clusters. They bought fast internal direct attached storage, shared SAN, or shared NAS storage over standardized networks such as fibre channel or Ethernet. They manually tuned their systems to get maximum performance.

This started to change in the 2000s. Database vendors started delivering turnkey database systems in mass. These database systems were engineered to deliver optimal performance. Every type of database could be acquired as a turnkey database system or as software licensed only for the DIY system. The IT manager's previous DIY experience convinced them that they could get equal or better performance at much lower costs than any turnkey database system. The common wisdom stated that this was especially true if they could leverage the latest server, storage, and network interconnect technology.

Oracle has continuously demonstrated the fallacy of that common wisdom with its Exadata Database Machine, the Oracle Database turnkey system. Each release of the Oracle Exadata system over the past 10 years has widened the gap between Exadata and DIY systems. The gap has continually grown in performance, performance density (IOPS and/or throughput per rack unit or RU), automation, security, database resilience, scalability, and most of all total cost of ownership (TCO).

The release of Exadata X8 in June 2019 made the gaps so wide in every aspect that it did not matter what server (generic x86, IBM Power Systems, HP-UX Integrity servers, etc.), storage (Dell EMC or Pure all flash array, NVMe-oF), interconnect, or storage media the DIY utilized, they could not come close to matching Exadata X8. And the DIY costs were outrageously higher than Exadata X8. It was enough to completely demoralize the most ardent DIYer. Now that Oracle has released the latest advanced Exadata X8M, the gap is completely unreachable for DIY systems.

Game Over.

Let's put that in perspective. If the pros in the IT department are attempting to put together or just considering putting together an Oracle Database DIY system, or any other database DIY system…***STOP***! Just stop. Take a breath. Read this document. Doing a DIY database system today is a huge waste of effort, personnel, and especially time and money. The Exadata X8M will deliver more performance and automation, in a much simpler package, for a significantly lower total cost. It ***annihilates*** every database system or DIY database system. This document shows why.

Table of Contents

Why Oracle Exadata X8M is Game Over for DIY Database Systems

The essential difference between Oracle Exadata X8M and DIY Oracle Database systems is the software + hardware co-engineering. That's not simply a marketecture claim, it's a real architectural difference. Oracle significantly modified the Oracle Database software, Oracle Linux operating system[1] in the database servers and storage servers, and Oracle storage server software to make them work synergistically. That synergy optimizes performance, performance density, automation, AI machine learning, database resilience, security, and total cost of ownership (TCO). Those software modifications are not available when the Oracle Database is running on any hardware other than Exadata. Exadata now has 70+ unique X8M capabilities not found in any other hardware running the Oracle Database. Those capabilities are what makes Exadata X8M untouchable by any DIY component assembly.

Proving this assertion only requires comparing Exadata X8M to DIY systems. Comparisons require articulating the underlying assumptions. These **assumptions** include:

- DIY[2] system utilizes the same Oracle Database as Exadata X8M
    - Without the 70+ unique Exadata X8M features available to DIY systems
- Exadata X8M is standard off the shelf hardware
- DIY hardware is commodity-off-the-shelf (COTS)
    - Servers – X86 or IBM POWER based servers
    - NICs
    - Switches
    - Storage
        - Utilizing latest purported fastest commercially available shared storage
        - Dell EMC PowerMax storage
- There are no specialized hardware ASICs or FPGAs in either DIY systems or Exadata X8M

Look at how they compare.

Performance

Oracle Exadata X8M	DIY
Unique Oracle Exadata software designed to specifically enhance and accelerate performance:	No unique software to enhance or accelerate performance. Zero. DIY systems require that the hardware to provide the performance.
• **Persistent memory (PMEM) tiering (Intel Optane)** Reads are like DRAM reads. But writes required sophisticated new software to effectively use PMEM as memory instead of storage. The new software calls special instructions to flush out the database server CPU cache to the PMEM and to complete or back out write sequences interrupted by a crash. ○ **Much faster non-volatile storage tier in front of flash** ○ **Up to 27TB of PMEM per rack in storage servers** ○ **Not much slower than DRAM**	• **Persistent memory (PMEM)** No off-the-shelf software enabling "App Direct Mode" for the Oracle Database. This means any use of PMEM can only be used in "Memory Mode". Memory Mode uses PMEM as a lower cost slightly slower memory behind DRAM that acts as a L4 cache. All the data in this configuration is considered volatile. PMEM persistent capabilities require App Direct Mode which the Oracle Database does not support. ○ **No PMEM capabilities**
• **PMEM data accelerator** Database servers utilize RDMA over 100G Converged Ethernet (RoCE) to read PMEM directly from multiple	• **PMEM SSDs** In shared storage, the preferred storage for Oracle Database and RAC, PMEM is implemented using slightly faster NVMe

---

[1] Oracle moved to its own version of Linux years ago because RedHat Enterprise Linux could not keep up with the performance Oracle was demanding.

[2] There are other databases in the market making the assertion they're faster than the Oracle Database. Those assertions are primarily apples to oranges comparisons. They are never like-to-like comparisons vs. Exadata. Nor do these databases compare functionally to the Oracle Database. Take time series databases as an example. Time series databases are very fast for time-based data and that's it. The time series database vendors are merely hyping the performance of their one-trick pony time series database not comparing it. The Oracle Database also has built-in very fast time series databases, but it also uniquely provides relational, data warehousing, object, key value, document, spatial, graphical, JSON, and more, in a single multi-tenant multi-database license. There is no comparison. This is similar for other types of databases as well. Doing a realistic comparison to DIY systems requires utilizing the same database.

storage servers, thus bypassing IO software, interrupts, and context switches.

- o **Reduces latency by 10X vs Exadata X8**
- o **IO latencies ≤ 19µs 8K SQL database reads**
- o **Increases effective cache capacity by 10X as well**

- **PMEM commit accelerator**
Database servers use RDMA over 100 Gb/sec Converged Ethernet (RoCE) to write redo to persistent memory directly in multiple storage servers instead of IO, thus bypassing the storage controllers, IO software, interrupts, and context switches. Faster log writes provide much faster transaction commit times.

    - o **Up to 8X faster log writes than Exadata X8**

- **RoCE Class of Service (CoS) High Priority Networking**
Allows packets to be sent on multiple classes of service, each with separate network buffers for independent processing. Exadata X8M uniquely chooses the most optimal CoS for each database message. Messages requiring low latency (cluster heartbeat, transaction commit, cache fusion) aren't hindered or slowed by high throughput messages (backup, reporting, batch.)

    - o **Minimally 3X faster inter-nodal messaging**
    - o **Much faster OLTP**

- **RoCE Instant RAC Node Failure Detection**
Exadata X8M uniquely uses frequent heartbeat messages between nodes to detect possible server failure. Rather than relying on lengthy timeouts to detect server failure, Exadata X8M uses a series of RDMA reads to rapidly identify server node failures.

    - o **Very fast recoveries from Oracle Database RAC node failures - ≤ 3 secs vs. minutes**

- **Just a few other Unique Exadata Oracle Database Performance Accelerators**

    Computational Storage Servers. Offloads CPU-intensive workloads (SQL, XML, JSON, decrypt, RMAN backup filter or BCT, fast file creation, in-database analytics, and AI-ML), from Database servers to Storage servers.

    Automatic IO reduction from flash caching that delivers NVMe flash performance at HDD capacities and cost; and storage indexes eliminating IO not relevant to a particular query.

    Analytics Optimized Columnar Format. Hybrid Columnar Compression (HCC) speeds up analytics by 10-15X because the Oracle Database analyzes HCC data without requiring rehydration.

    Storage In-Memory Analytics. Delivers in-memory performance in shared storage. Exadata Storage servers automatically transform table data into in-memory database columnar formats within Exadata Flash cache. Exadata flash throughput approaches DRAM throughput enabling much faster vector processing for storage server queries. New with Exadata Smart Software 19.3.0, it decrypts only the columns needed for a query and now offloads Oracle Database in-memory aggregation algorithms to Exadata, further accelerating the Oracle Database.

    Hybrid Columnar Compression (HCC). Exadata HCC automatically compresses Oracle Database data up to 10-15X In the Exadata Storage servers. More importantly, the Oracle Database can transform and cache HCC data automatically into SIMD compatible formats (no rehydration) accelerating.

SSDs. Latency and IOPS are marginally better than NVMe flash, but at a much higher cost.

- o **Much higher storage costs**
- o **No noticeable improvement in performance**

- **RoCE**
None of the Oracle Database software that leverages RoCE is available on DIY systems.

    - o **Higher costs**
    - o **But no noticeable performance improvement**

- **RAC Node Failure Detection**
Tied to very lengthy TCP/IP timeouts.

    - o **Reduced performance from failures**
    - o **Multiple orders of magnitude longer failure detection**
    - o **Long time for RAC node reconfigurations**

- **High Performance NVMe-oF Storage**
Low latency storage from major storage vendors such as Dell EMC claim up to 15 million 4K IOPS in 2 racks, or 7.5 million IOPS per rack for their latest and greatest PowerMax. That drops minimally in half for 8K Oracle SQL IOPS to 3.75 million IOPS per rack. The NVMe-oF does reduce latency between the database server and storage.

    - o **Storage does not offload Oracle Database**
    - o **Latency is minimally 120 µs server - storage**
    - o **Latency does not include Oracle Database processing**
    - o **Storage is exceedingly expensive**

- **Built-in Storage Deduplication and Compression**
Generally designed to work effectively with unstructured data. Does not work very well with Oracle Database data.

    - o **Typically, only getting a 2-3X data reduction**
    - o **Data must be rehydrated to be read by Oracle Database**

- **DIY Systems have no commercial database accelerators**
None available off-the-shelf. It would require significant programming, testing, documentation, ongoing QA, patching, etc., and still could not come close to what comes standard with Exadata.

    - o **Definitively, not worth the effort, time, and not cost effective**

Advantage: *Oracle Exadata X8M*

Performance Results Proof

Performance as reported by the manufacturers and adjusted to be apples to apples.

Oracle Exadata X8M	DIY
Max 8K SQL Read IOPS per rack: up to ~ 16 million • Annihilates DIY with at least **4.27 X** more IOPS per rack	**Max 8K SQL Read IOPS per rack: up to ~ 3.75 million** • Generous estimate based on industry's purported fastest commercially available shared storage , the Dell EMC PowerMax
Max throughput per rack: min 350GB/s to 560 GB/s *Depending on configuration* • Outstanding for batch analytics, AI-ML, and big data • Devastates DIY competition with minimally ~ **3.2 X** > throughput	**Max throughput per rack: up to ~ 175 GB/S** • Generous estimate based on industry's purported fastest commercially available shared storage, the Dell EMC PowerMax • Does not take into consideration database server throughput limitations
8K SQL database read latency: ~ ≤ 19µs ○ Obliterates DIY with at least **5.26 X** lower latency	**8K SQL database read latency: ~ ≤ 100µs** • Generous estimate based on industry's purported fastest commercially available shared storage, the Dell EMC PowerMax • Does not take into consideration database server additional latency

Advantage: ***Oracle Exadata X8M***

Performance Density

Performance density matters significantly in the modern data center. Data center real estate is precious and expensive. Rack units (RU) affect costs including power, cooling, allocated overhead, cables, conduits, transceivers, etc. More performance per RU reduces the amount of hardware needed to meet database performance requirements. Less hardware equates to a lower Oracle Database license cost and overall much lower TCO. Performance density matters.

Oracle Exadata X8M	DIY

Max 8K SQL Read IOPS density: up to ~ 381 K per RU
 	
• Crushes DIY with ≥ **4.27 X** more IOPS per RU
	Max 8K SQL Read IOPS density: up to ~ 89 K per RU
	
	• Depends on the number of database servers
	  • A full Dell EMC PowerMax rack is 40 RU by itself
	  • Should be a minimum of 4 powerful database servers for RAC
	  • Each DB server is at least 1 RU and more likely 2 RU

Max throughput density: min 8.33 GB/s to 13.33 GB/s per RU
 	
• Demolishes DIY systems with up to ≥ **3.2 X** > throughput
	Max throughput density: up to ~ 4.17 GB/S per RU
	
	• Depends on the number of database servers
	  • A full Dell EMC PowerMax rack is 40 RU by itself
	  • Should be a minimum of 4 powerful database servers for RAC
	  • Each DB server is at least 1 RU and more likely 2 RU

Advantage: ***Oracle Exadata X8M***

Automation

Automation is all about reducing and eliminating time consuming database administrator (DBA) manual labor-intensive tasks. Tasks such as indexing, tuning, troubleshooting, and more. As these tasks are automated, the DBA can focus on the applications, new projects, and time to market. Less time spent on time wasting tasks means more productivity.
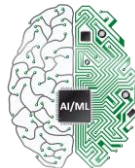
Oracle Exadata X8M	DIY
• **Auto Indexing** Automatically delivers extremely efficient and usable indexing via built-in AI-ML, adjusting quickly to changing workloads. Reinforced learning empowers Exadata to learn from its own actions. Candidate indexes are validated before implementation. And it does all of this in a small fraction of the time and much more efficiently versus the most experienced DBA. • **Auto Tuning** Automatically adjusts the Exadata system to deliver the best possible database performance for all functions based on priority. • **Auto Self-Diagnosing and Repair** Automatically detects and recovers from sick or failed servers, storage, switches, or links. Exadata additionally provides real-time hang management, anomaly detection, maintenance slot identification, and bug identification. • **Auto Query Optimization** Automatically optimizes each query in real-time. Ensures the most critical queries are never hindered by less critical operations. • **Auto Patching** All patches including security patches, for all databases, software, and hardware components are combined in a single patch each quarter. Implementations are non-disruptive. • **Auto Statistics** Automatically gathers statistics in real-time as DML executes.	• **Nominal Automation** Only found in Dell EMC PowerMax storage in that it will place data on different storage tiers based on policy and history. However, that's it. Indexing takes many DBA hours. And the indexes have to be refined, tuned, and re-indexed manually by the DBA. Tuning is an ongoing exercise consuming much of the DBAs time. Troubleshooting is frequently a lengthy process and can be quite frustrating when determining where the root cause of the problem is in the multi-vendor DIY hardware environment Multi-vendor finger pointing adds to that frustration. Patches alone require on average more than 8700% more manual procedures than Oracle Exadata X8M. And many of them are operationally disruptive requiring scheduled downtime.

Advantage: ***Oracle Exadata X8M***

AI Machine Learning (AI-ML)

AI machine learning has emerged as the next evolution in analytics providing greater value across industries. Additional analytical value runs from 30% to 128% according to the McKinsey Global Institute Analysis. Oracle Exadata X8M comes with AI-ML built into the systems. That AI-ML does not require learning Python. It leverages SQL commands and algorithms production proven in the Oracle Autonomous Database.

Oracle Exadata X8M	DIY
• **Built-in AI-ML** Enables utilization of SQL commands, built-in schemas, sample data, and a drag and drop interface to implement sophisticated AI-ML analytics on customer databases.	• **Built-in AI-ML** Available as long as they are utilizing Oracle Database 19c Enterprise Edition or better.

Advantage: ***Equal***

Database Resilience

The best ability is always availability. There is no DIY equivalent.

Oracle Exadata X8M	DIY
• **AL4 fault tolerance** Only database system to achieve IDC's AL4 fault tolerance. AL4 fault tolerance makes faults and failures unnoticeable to users because of zero data loss. Their databases keep running as if nothing happened. Only the DBA is made aware of the faults or failures. AL4 fault tolerance is automatically configurable in the Oracle Exadata X8M when implemented in an Oracle Maximum Availability Architecture (MAA) Gold or higher environment. • **Automates Backup, Recovery, and DR** Automatically and intuitively leverages RMAN; ability to recover databases, PDBs, or CDBs to any point-in-time; Flashback Database time enabling views or returns of database objects past states without point-in-time media recoveries. Oracle Zero Data Loss Recovery Appliance reduces RMAN backup time by up to 95%; and Oracle Data Guard synchronization of one or more database copies.	• **AL4 fault tolerance is extremely difficult to achieve** Technically possible. But, it requires extensive time, planning, implementation, ongoing tuning, updating, documenting, quality assurance testing, and ongoing effort. It's a manual process. None of the Oracle backup, recovery, and DR tools are automatically implemented. • **Lack of Backup, Recovery, and DR Automation** None of the advanced Oracle exceptional backup, recovery, and DR tools are automatically implemented. This makes achieving Oracle MAA a non-trivial task.

Advantage: ***Oracle Exadata X8M***

Security

Security has become a top priority for CIOs. Malware attacks have been increasing at an exponential rate, especially ransomware. Personal identifiable information (PII) protection laws and regulations have made security more important than ever with harsh fines and stock reducing data breach publicity.

Oracle Exadata X8M	DIY
• **Storage server decrypt offload** All database cybersecurity measures demand processing, memory, and I/O resources. decrypt is resource intensive reducing database performance by as much as 20%. Exadata leverages its built-in computational storage servers to offload those decrypt functions ensuring nominal impact to the Oracle Database performance at approximately 2% or less[3]. Encryption occurs after the data is compressed so the encryption cost is decreased by the compression ratio. Exadata is able to query fully encrypted and compressed databases with minimal overhead at a rate of hundreds of gigabytes of (original) user data per second. This is something no DIY database system can do. • **Unique Exadata Engineered Built-in Security** **Secure Boot**. Ensures system UEFI firmware only allows execution of cryptographically signed boot loaders the system recognizes as trustworthy. Each server reboot makes sure every executed component is verified, preventing malware from hiding embedded code in the boot chain.	• **Only the Security Inherent in the Oracle Database 19c** Oracle Database 19c Enterprise Edition automated database security includes core database security for users, roles, authentication, etc.; Network Communication Encryption; Transparent Data Encryption (TDE); Column-Level Encryption; Oracle Key Vault; Database Auditing; Oracle Audit Vault; Oracle Database Firewall; Oracle Database Vault; Label Security; Real Application Security; Virtual Private Database; and Data Masking and Subsetting. More detailed Oracle security information can be found here: Oracle Database Security Guide. However, these capabilities are also part of Exadata X8M and are not unique. And those unique capabilities in Exadata X8M are not available in any DIY database system.

---

[3] Some vendors assert self-encrypting drives (SED) eliminates encrypt/decrypt overhead more than Oracle Exadata Transparent Data Encryption (TDE) offload. For data encrypt/decrypt at rest, it is equivalent. However, SEDs do not encrypt data in-flight or data backed up – required for most PII laws and regulations. That still requires TDE and its overhead.

Stored Data Encryption. Storage flash SSDs and HDDs encrypts all data as it enters the devices. **Secure Erase.** Leverage stored data encryption when an Exadata is repurposed or decommissioned instantly erasing all the data by changing the encryption keys and deleting the original keys. **Minimized Oracle Linux Distribution.** Installs and enables just the RPMs required to run the Oracle Database. Strengthens system security by reducing potential malware attack vectors avoiding default Linus vulnerabilities. **Oracle Linux ksplice.** Enables security updates to be applied to the OS while staying online.	

Advantage: ***Oracle Exadata X8M***

Total Cost of Ownership (TCO)

There is an enduring myth that DIY systems will be less expensive than Oracle Exadata. This perception is tied to the belief that commodity-off-the-shelf (COTS) hardware will be less costly than the Exadata system and the Oracle Database license costs will be the same. Both are completely untrue. These perceptions are based on false underlying assumptions. The first is that price and cost are the same. They're not. The unique Exadata software that accelerates the Oracle Database by as much as an order of magnitude means less hardware is required to achieve the same database objectives. HCC means reduced storage hardware is required to achieve the same database objective. Reduced database server and storage hardware equates into reduced supporting infrastructure hardware too: fewer NICs, cables, transceivers, switches, less rack space, etc. The Oracle Database is licensed by cores. Exadata always requires fewer database server cores than any DIY system because it can offload to the Exadata Storage servers. Oracle Database licensing does not count the Exadata Storage servers in the license subscription. Fewer database server cores equals reduced database subscription licensing. Less total hardware equals reduced hardware capital expenditures – CapEx, and operating expenditures (power, cooling, rack space, allocated overhead, personnel) – OpEx.

That's just the cost side of the ledger. What about the revenue side? Exadata X8M's much accelerated Oracle Database means data gets analyzed faster, producing usable actionable information sooner. Much of that information may result in quicker revenues or higher productivity and reduced organizational or supply chain costs. Faster time-to market means capturing revenue that would never have been captured. That too must be taken into consideration.

Oracle Exadata X8M	DIY
• **Database server hardware** Less	• **Database server hardware** More
• **Storage hardware** Less	• **Storage hardware** More
• **Supporting infrastructure** Less	• **Supporting infrastructure** More
• **Oracle Database subscription license** Less	• **Oracle Database subscription license** More
• **Time spent: implement, ops, mgmt, troubleshooting** Less	• **Time spent: implement, ops, mgmt, troubleshooting** More
• **CapEx** Less	• **CapEx** More
• **OpEx** Less	• **OpEx** More

Humbling Advantage: ***Oracle Exadata X8M***

Conclusion

Oracle Exadata X8M's performance, performance density, automation, AI-ML, database resilience, security, TCO, and time-to-market advantages makes it impossible for any DIY database systems or any other database system to match …period. DIY Oracle Database systems waste time, effort, and treasure upfront and ongoing. There is nothing to be gained and much to be lost. Few IT pros build their own DIY servers, storage, or laptops today for the same reasons. This is why the Oracle Exadata X8M's "**unfair advantage**" has made it crystal clear that it's "**Game Over**" for DIY database systems.

The choice is simple:

1. Spend a lot more effort, energy, time (especially time), CapEx, and OpEx for a less capable database system.
2. Or get a lot more database system for a whole lot less effort, energy, time, CapEx, and OpEx.

You decide.

For More Information on Oracle Exadata X8M

Go to: Oracle Exadata