ORACLE

Oracle Exadata

—

A guide for decision makers

October 2021

The Oracle Exadata Database Machine[1] (Exadata) is a computing and storage platform specialized for running Oracle databases.

The goal of Exadata[2] is higher performance and availability at lower cost by optimizing and integrating hardware and software at all levels and moving database algorithms and intelligence into storage and networking, bypassing the traditional layers of general-purpose servers.[3]

Exadata is a combined hardware and software platform that includes scale-out compute servers, scale-out intelligent storage servers, ultra-fast networking, persistent memory (PMem), NVMe flash, and specialized Exadata Software[4] in a wide range of shapes and price points. Exadata Storage features high-performance servers to store data and run Exadata Software for data-intensive database processing directly in the shared storage tier.

Exadata debuted[5] in 2008 as the first offering in Oracle's family of Engineered Systems[6] for use in corporate data centers deployed as "private clouds". In October 2015, Exadata became available in the Oracle Cloud as a subscription service, known as the Exadata Cloud Service.[7]

Oracle Databases deployed in the Exadata Cloud Service[8] are 100% compatible with databases deployed on Exadata on-premises, which enables customers to transition to the Oracle Cloud with zero application changes. Oracle manages this service, including hardware, network, Linux software and Exadata software, while customers retain complete control over their databases.

In early 2017, a third Exadata deployment choice became available. Exadata Cloud@Customer[9] is Exadata Cloud Service technology deployed on-premises (behind the customer's firewall) and managed by Oracle Cloud experts. Like the Exadata Cloud Service, Exadata Cloud@Customer is owned and managed by Oracle, and acquired through a pay-as-you-go subscription. The Oracle Cloud@Customer[10] program brings all the benefits of the Oracle public cloud while addressing potential network latency, security, and regulatory concerns.

In 2018 Oracle introduced Oracle Autonomous Database[64] – a cloud-based self-driving, self-securing, self-protecting database that provides mission critical availability and security while reducing management costs. Oracle Autonomous Database is available for Exadata Cloud Service and Cloud@Customer deployments.

In 2019, the release of Exadata X8M substantially enhanced Exadata's performance through the addition of two major technical breakthroughs - persistent memory (PMem) and RDMA (Remote Direct Memory Access) over Converged Ethernet (RoCE)[65]. Oracle Exadata X8M uses RDMA directly from the database to access persistent memory in smart storage servers, bypassing the entire OS, IO, and network software stacks. This enables dramatically lower latency and higher throughput.

Overwhelming Industry Adoption

The success of Exadata is evidenced by its adoption by the largest and most influential organizations in almost all industries. In Financial Services, 9 of the top 10 retail banks, insurers and credit card issuers, and 7 of the top 10 investment banks use Exadata. All 10 of the largest Health Care providers, and 9 of the top 10 Retailers. Exadata has been adopted by 87% of the Fortune Global 100; nearly one-third of those companies deploy Exadata Cloud Service in the Oracle Cloud[11].

ORACLE

In 2021, the release of Exadata X9M saw further performance increase with the inclusion of PCI Express version 4.0, providing an increase in network and storage throughput for the RoCE network fabric and NVMe Flash. Exadata Software version 21.2 includes new features that not only enhance the investment customers have already made in Exadata, but also include features that complement the X9M hardware such as enhanced decompression and decryption algorithms. The result of enhancements in both hardware and software in X9M results in a 70% and 87% improvement in OLTP and Analytics operations, respectively.

Exadata Use Cases

Exadata is designed to optimally run any Oracle Database workload or combination of workloads, such as OLTP applications running simultaneously with analytics processing. The platform is frequently used to consolidate many databases that were previously running on dedicated database servers. Exadata's scale-out architecture is naturally suited to running in the Oracle Cloud, where computing requirements can dynamically grow and shrink.

Historically, specialized database computing platforms were designed for a particular workload, such as data warehousing, and were poor or unusable for other workloads, such as OLTP. Exadata has optimizations for all database workloads, implemented such that mixed workloads share system resources fairly. Resource management features allow for prioritized allocation of system resources, such as always favoring workloads servicing interactive users over reporting and batch, even if they are accessing the same data.

Long running requests, typical of data warehouses, reports, batch jobs and analytics, run many times faster compared to a conventional, non-Exadata database server[12]. Customer references often cite performance gains of 10x or greater. Analytics workloads can also use the Oracle Database In-Memory[13] option on Exadata for additional acceleration, and in-memory databases on Exadata have been extended to take advantage of Flash memory, whose capacity is many times larger than the capacity of DRAM. Exadata's Hybrid Columnar Compression[4] feature is intended to reduce the storage consumption of data warehouses and archival data as well as increase performance by reducing the amount of I/O.

Transactional (OLTP) workloads on Exadata benefit from the incorporation of persistent memory and Flash memory into Exadata's storage hierarchy, and the automatic "tiering" of data into DRAM, persistent memory, Flash or disk storage. Special algorithms optimize persistent memory and Flash for response time sensitive database operations such as log writes. For the most demanding OLTP, all-Flash storage eliminates the latency of disk media completely.

Exadata Design Concepts

To better understand the design of Exadata it helps to compare it with a traditional database computing platform, assembled from separate hardware and software components operating independently.

Traditional computing platforms are general-purpose

The hardware components that make up a typical database computing platform are a compute server connected over a network to a storage array. The database software runs on the compute server and sends or receives data to and from the storage array over the network. The hardware components use standard software protocols to communicate with each other. This separation via standard interfaces is what allows a general-purpose computing platform to run a wide variety of workloads, software and hardware from different vendors. All the application logic and the processing of data are performed on the compute server, to which all the data must be sent. This approach enables using a computing platform for a wide range of software applications, though it will not be optimized for any particular application.

ORACLE

Oracle Database is the focus of Exadata

The goal of Exadata was to create a complete stack of software and hardware tailored to the Oracle Database, that moved processing to its optimal location. Because Exadata is only processing Oracle Database requests it can take advantage of that focus in all the software layers. The hardware design includes technologies such as very fast RoCE Ethernet networking, Flash memory and persistent memory, integrated into the architecture to yield the most advantages to Oracle Database applications. Given the importance of data storage to databases, Oracle placed particular focus on optimizing that aspect of the Exadata platform.

Oracle wanted a storage layer for Exadata that could easily scale out and parallelize Oracle Database requests. The addition of Flash memory and (later) persistent memory to Exadata Storage Servers also opened up a range of possibilities for optimizing performance in the storage layer. For example, as the performance and capacity of Flash storage increased at a rapid rate, the network became a bottleneck for traditional database platforms, whereas Exadata's offloading of database processing into Exadata Storage Servers avoided that problem. The subsequent addition of persistent memory in the Exadata storage layer exposed the limitations of traditional platforms even more acutely.

Adding database intelligence to storage

At the time Exadata was conceived, Oracle had several decades of experience developing database software, and was well aware of the limitations and performance bottlenecks imposed by traditional computing platforms[16]. To fulfill the Exadata mission, Oracle needed a storage layer that could easily scale out and parallelize Oracle Database requests. Oracle also recognized the opportunity for storage to cooperate in the processing of database requests beyond just storing and shipping data. For example, rather than send an entire database table across the network to the compute server to find a small number of records, such data filtering could be done in storage so that only the resulting records need be sent across the network.

In summary, Oracle recognized the need for a powerful server that could run intelligent database software and act as a storage array, with a modular design that could easily grow in capacity and performance as the database grew. Building a "database-aware" storage server that could cooperate with compute servers in the execution of database requests became a compelling undertaking, enabled by focusing Exadata on what is best for the Oracle Database.

The database-aware Exadata Storage Server[17][18], invented by Oracle to replace the traditional storage array, is the foundation of Exadata.

Optimizing across the full stack

To maximize the effectiveness of Exadata, Oracle had to control the software and hardware components of the platform, so that coordinated improvements could be tightly integrated and made anywhere at any time.

Oracle already had a broad portfolio of software products when Exadata was conceived, covering most of the software layers that are required to run a database platform, such as the Oracle Linux operating system, storage management software, monitoring and administrative tools, virtual machine software, and, of course, Oracle Database and options software.

The initial 2008 release of Exadata (V1) was a joint development between Oracle (software) and Hewlett-Packard (hardware). The second generation (V2) of Exadata switched to hardware from Sun Microsystems, and shortly thereafter Oracle acquired Sun Microsystems, and thus gained ownership of the main hardware components of Exadata.

Owning the main hardware components of Exadata completed Oracle's ability to develop an entire computing platform optimized around the Oracle Database. A further benefit for customers was the ability to support the entire

ORACLE

Exadata platform from one vendor; something impossible with a traditional computing platform of hardware and software components from multiple vendors.

Exadata Smart Software

With the initial versions proving the value of the Exadata concept, Oracle embarked on a regular release schedule of new software capabilities closely aligned with the release and incorporation of leading-edge hardware components.

Using knowledge of Oracle Database formats and algorithms, plus an understanding of the application workload, these "smart" software enhancements, running in the Oracle Database and all software and firmware in the platform, had the potential to significantly outperform the capabilities of a traditional database platform. The ability to run Oracle Database routines in the Exadata Storage Server software was a major enabler of this advantage.

Software enhancements specific to Exadata achieved better performance in some areas due to the integration of hardware and software. For example, when an OLTP application commits a transaction to the database software, that request is viewed by Exadata as a critical operation and prioritized accordingly within the network and in the storage servers. The commit request will jump ahead of less urgent messages on the network and in the I/O queue[15].

Another example is the use of Flash memory for caching data being accessed by analytics applications. Because a columnar data format is more effective for analytics workloads, when Exadata moves the row data from disk to Flash, it automatically reformats the data into a columnar format[16].

These examples illustrate how the Oracle Database understands the intent of the application and cascades this understanding to the network and storage software, which then behaves accordingly.

Since Exadata's debut in 2008, there have often been at least two significant Exadata software releases per year, delivering dozens of "smart" software enhancements. Most of these enhancements are based on a smaller number of essential technical foundations, as explained below.

Technical foundations

Offloading to Storage - refers to the execution of data-intensive database operations within the Exadata Storage Servers, such as data scans, table joins and filtering of rows and columns. Sending just the description of the operation, and getting back the results, substantially reduces the network traffic between the compute and storage servers. This avoids the network bottleneck of traditional architectures where data-intensive operations require shipping large amounts of data between compute servers and storage. Offloading is possible because Exadata storage is built on standard servers, capable of running database functions in coordination with the compute server, simultaneous with storage I/O. Over time, more database functions and more data types have been offloaded. In addition, "reverse offloading" will push an operation back to the compute servers if Exadata storage is too busy.

Storage Indexes - enable the avoidance of I/O by tracking column values within small regions of storage. Storage indexes are automatically maintained and kept in memory on Exadata Storage Servers. If the Storage Index indicates that an I/O to a region will not find a match, that I/O is avoided, which yields a significant performance benefit. Initially Storage Indexes tracked value ranges within a small number of columns. Over time, more columns and more sophisticated value tracking have been added, so that additional classes of I/O operations can be avoided. Storage Indexes are persisted to local storage on the Exadata Storage Servers to improve performance consistency after a server has been restarted.

Flash and PMem Caching - delivers the low latency (fast response) of Flash and persistent memory, while preserving the lower cost of disk for storing large databases, for the best I/O performance at the lowest cost. In general, a small percentage of a database is active at any one time. For instance, if just the active data is held in Flash, the I/O performance would be equal to all-Flash storage, at a much lower cost. Exadata monitors the current workloads and

ORACLE

keeps the most active data in Flash or persistent memory, in the optimal format. For example, Exadata knows when an I/O is part of a database backup, and not an indication of an active data block, whereas traditional platforms view any I/O as a "hot" block. Flash caching will also reformat rows into columnar format in Flash for data being accessed for analytics. Initially, Flash caching was only used for reading data, then extended to log writes and all other write I/O. Later, Flash was used as an extension of Oracle's Database In-Memory columnar data store, for significantly larger in-memory databases than DRAM capacity alone would permit. PMem caching debuted with Exadata X8M, adding an even faster cache in storage and substantially improving IOPS (I/Os per second) and latency.

Hybrid Columnar Compression (HCC) - reduces the amount of storage consumed by infrequently updated data, such as data warehouses (which can grow to enormous sizes). Conventional data compression algorithms yield between 2x and 4x compression, whereas HCC averages between 10x and 15x compression due to the greater compressibility of columnar formats. Such a large reduction in the amount of I/O can also substantially improve performance. Initially, HCC tables did not support row-level locking, limiting their use with OLTP applications. In 2016, support for row-level locking was added to HCC on Exadata, improving the performance of mixed workloads with HCC data. HCC's hybrid format enables Exadata to avoids the performance pitfalls of columnar-only databases.

Resource Management - allocates Exadata system resources, such as CPU, I/O and network bandwidth, to databases, applications or users based on priorities. When consolidating many databases on Exadata, Resource Management ensures the appropriate quality of service. I/O Resource Management debuted in V1 of Exadata. Network Resource Management was added in Exadata X4.

In-Memory Databases - offer exceptional performance for Analytics workloads, leveraging DRAM on compute servers, a complement to Exadata's emphasis on storage and networking. Oracle Database In-Memory became available in 2014 on Exadata, leveraging its fast InfiniBand network for in-memory Fault Tolerance. To support larger in-memory databases, Exadata Storage Servers implement in-memory routines and in-memory data formats in Exadata Flash, as an extension of the same in-memory processing that occurs on compute servers.

Smart software enhancements

A more detailed listing of software enhancements is below, grouped by their value to analytics or OLTP workloads, and their impact on database availability and security. Similar enhancements cannot be duplicated on other platforms because they require software and API modifications and integration across database software, operating systems, networking and storage.

Refer to the Exadata documentation[17] and data sheet[1] for descriptions of these features.

For ANALYTICS	For OLTP
Automatically parallelize and offload data scans to storage	Database-aware NVMe PCIe flash interface
Filter rows in storage based on 'where' clause	Exadata Smart Flash Logging
Filter rows in storage based on columns selected	Write-back Flash Cache
Filter rows in storage based on join with another table	I/O Resource Management by database, user, or workload to ensure QoS
JSON and XML offload	Network Resource Management
Hybrid Columnar Compression	Exafusion Direct-to-Wire Protocol for inter-node data transfer
Storage Index data skipping to avoid I/Os	EXAchk full-stack validation
Automatic transformation to In-Memory columnar format in Flash Cache	Active AWR includes storage stats for end-to -end monitoring
Temp I/O to Flash Cache for large analytic queries and loads	Cell-to-cell rebalance preserving flash cache

ORACLE

Offload index fast full scans	In-Memory commit cache
Offload scans on encrypted data, with FIPS compliance	Memory optimized OLTP and IoT lookups
Storage offload for LOBs and CLOBs	Persistent memory data accelerator (PMem Cache)
Storage offload for min/max operations	Persistent memory commit accelerator (PMem Log)
Data Mining offload	Database-specific control of PMem usage
All ports active InfiniBand messaging	Cell-to-cell rebalance preserves PMem Cache
Reverse offload to DB servers if storage CPUs are busy	Smart Flash Log write-back
Just in time smart columnar decryption	
Smart aggregation with columnar cache	
Fast In-Memory columnar cache creation	

For AVAILABILITY	For SECURITY
Instant detection of node or cell failure	Fast, hardware based (AES) encryption/decryption
Sub-second failover of I/O on stuck disk or flash	Full-stack security scanning
Offload incremental backups to storage servers	Minimal Linux distribution (reduced attack surface)
Hardware Assisted Resilient Data validation (H.A.R.D.)	Database and ASM-scoped security
Instant datafile creation	Fast, secure eraser of disk and Flash (crypto erase)
Prioritized rebalance of critical files	InfiniBand partitioning
Automatic hard disk scrub and repair	Audit record of all logons and configuration changes
Power-cycle failed drives to eliminate false drive failures	Firewall-protected Exadata Storage Servers secure network access
Avoid reading disks predicted to fail	IPV6 support
Cell software transparent restart	Restricting the binaries used to boot the system with Secure Boot
Flash and disk life cycle management alert	Integrated Lights Out Management (ILOM) of database and storage servers
Prevent shutdown if mirror server is down	Comprehensive monitoring and auditing at the server, network, database, and storage layers
Automatic Software Updates on an entire fleet of Exadata systems with a single operation	Advanced Intrusion Detection Environment (AIDE)
Hot pluggable Flash cards	FIPS 140-2 certification
Automatic monitoring of CPU, network and memory using Machine Learning	PCI-DSS compliance
Exadata I/O latency capping	Least privilege: processes run with the lowest privileges needed to perform a task
Exadata Storage Server I/O timeout threshold	Highly restrictive password enforcement
Automatic identification of underperforming disks	Secure computing filter to restrict system calls
Automatic rebalance on disks predicted to fail	Isolation policies to secure network traffic, databases and storage
Exadata storage snapshots	Secure RDMA fabric isolation
Cell-to-cell rebalance to preserve Flash Cache population	Multi-pass secure erase of disks and flash
Prefetch OLTP data into secondary mirror Flash Cache	Securing Storage Server Software processes with memory protection keys
Online Linux patching (Ksplice)	

ORACLE

Database Software

Exadata compute servers run the Oracle Linux 7 operating system and Oracle Database 11g Release 2 Enterprise Edition through Oracle Database 19c (Oracle's Long Term Support release), and now Oracle Database 21c Enterprise Edition.

Oracle Database is also known as the Oracle Converged Database – a term used describe the multi-data model storage and processing support Oracle has built into Oracle Database over the past several decades. Exadata provides optimizations for many of the data types and processing methods in the Oracle Converged Database, such as JSON and XML Offload, and Data Mining Scoring on the Exadata Storage Servers, that enhance performance of applications that use both relational and non-relational data types.

Exadata system resources can be optionally virtualized using the Xen-based Oracle VM (prior to X8M) or KVM-based Oracle VM (X8M and later). All Oracle Database options, such as Real Application Clusters, Multitenant, Database In-Memory, Advanced Compression, Advanced Security, Partitioning, Active Data Guard and others are optionally available with Exadata. Applications that are certified to a supported version of the Oracle Database are automatically compatible with Exadata. No additional modifications or certifications are required[18].

The same database software that runs on Exadata on-premises will run in the Exadata Cloud Service and Exadata Cloud@Customer. In addition, on-premises software licenses are eligible for the BYOL[19] (Bring Your Own License) transfer into the Oracle Cloud or Cloud@Customer.

Networking

Exadata X8M and later provides high-speed networks for internal and external connectivity. A 100 gigabits per second (100 Gbit/s) active/active RDMA over Converged Ethernet (RoCE) fabric is used for internal connectivity between compute and storage servers, and 25 and 10 Gbit/s Ethernet ports are included for data center connectivity. The RoCE fabric is also used as the cluster interconnect between compute servers. Previous generations of Exadata featured 40 gigabits per second (40 Gbit/s) InfiniBand network for internal connectivity. Because Oracle controls the entire hardware and software architecture of Exadata, introducing newer technologies, such as the RoCE fabric can be achieved transparently without encumbering customers with integration and testing that comes with such change.

Exadata uses a custom-designed, database-oriented protocol over the Ethernet fabric to achieve higher performance. It makes extensive use of Remote Direct Memory Access to improve efficiency by avoiding operating system overhead and extra copies when moving data between servers. Exadata also implements a direct-to-wire protocol[20] that allows the database to communicate directly to the RoCE hardware.

Exadata also takes advantage of RoCE Class of Service in its Network Resource Management[15] feature to prioritize important traffic across the network. In this feature the Oracle Database software tags network messages that require low latency, such as transaction commits, lock messages and I/O operations issued by interactive users, enabling them to overtake messages issued by less critical high-throughput workloads such as reporting and batch. The result is analogous to how an emergency vehicle with its siren on is able to move more quickly through heavy traffic - high-priority network messages are moved to the front of the server, network switch, and storage queues, bypassing lower-priority messages and resulting in shorter and more predictable response times.

Management Software

Oracle Enterprise Manager Cloud Control (EM) manages Oracle software and hardware, including the Exadata Database Machine. EM integrates with the built-in Exadata management tooling, as well as with customers' existing systems management and helpdesk tools. The Exadata plug-in for EM provides an integrated view of compute servers, storage servers, switches, and topology. It also provides discovery, monitoring and alerting capability for Exadata systems management.

ORACLE

The Exadata Cloud Service and Exadata Cloud@Customer platforms are managed by Oracle, while customers control and manage the software and databases running on the compute servers. Exadata Cloud platforms include a web-based provisioning wizard through which customers can quickly provision their chosen Exadata system and subsequently their database instances. Customers also perform familiar database administration and OS administration tasks aided by cloud automation for backup, patching, and upgrades.

Hardware

Exadata is available in two models: one based on two-socket compute servers and the other based on eight-socket compute servers. The two models differ only in the hardware used for the compute servers. The networking, storage servers and software are the same in both models.

Exadata Database Servers

The most recent Oracle Exadata Database Machine is the X9M[21] generation, introduced in September 2021.

The X9M-2 compute servers feature a small form factor, 1 RU (Rack Unit) in height. They employ 2-socket Intel Xeon processors; each socket with 32 compute cores for 64 total cores per compute server. Memory starts at 512 gigabytes (GB) and can be expanded to 2 terabytes (TB).

The Exadata Database Machine base configuration has 2 compute servers and 3 storage servers, referred to as a Quarter Rack. The same hardware is also available in an Eighth Rack configuration with half of the processing and half of the storage capacity. As the database workload and/or data size increases, additional compute and storage servers may be added to increase the volume of work performed in parallel, using Exadata's Elastic Configuration.

The X8M-8 compute server (continued from the X8M-8 Exadata Database Machine) uses eight-socket compute servers that consume 5 RU in height, have higher memory capacity at 3 to 6 TB per server, and contains 192 compute cores. This allows large database workloads to easily scale-up within a compute server while still supporting Exadata's scale-out expandability across multiple servers. The larger memory capacity of the X8M-8 also favors In-Memory Databases and very large OLTP, consolidation, and data warehouse workloads. Like the X9M-2, the Exadata X9M-8 base configuration has 2 compute servers and 3 storage servers, however consumes a half rack of space. Additional compute and storage servers may be added until the rack is full.

Exadata Storage Servers

There are three choices for Exadata storage servers: Extreme Flash (EF), High Capacity (HC) and Extended (XT).

The X9M-2 Extreme Flash Storage Server is all-flash storage containing 8 PCIe flash drives for a total of 51.2 TB of raw storage capacity and 1.5 TB of persistent memory.

The X9M-2 High Capacity Storage Server contains 12 disks of 18 TB each, for a total of 216 TB of raw storage capacity, plus 25.6 TB of PCIe flash and 1.5 TB of persistent memory to cache active data blocks. Exadata's Persistent Memory Data Accelerator, Persistent Memory Commit Accelerator, Smart Flash Cache, Smart Flash Log, Columnar Flash Cache and Write Back Flash Cache features automatically tier data on persistent memory and flash.

The X9M-2 Extended Storage Server contains 12 disks, 18 TB each, for a total of 216 TB of raw storage capacity.

In addition to adding storage servers into an Exadata Database Machine base configuration, storage servers may also be acquired with or added to Exadata Storage Expansion[22] racks.

ORACLE

Storage performance specifications for a full rack Exadata configuration are as follows:

EXADATA STORAGE SERVER	SCAN RATE	READ IOPS	WRITE IOPS
X9M-2 Extreme Flash	1,050 GB/sec	27,600,000*	8,596,000
X9M-2 High Capacity	630 GB/sec	22,400,000	8,596,000

* Based on an elastic configuration of 10 compute and 12 storage servers. All other rates are based on a standard full rack configuration of 8 compute and 14 storage servers[23]

Note: IOPS = 8K I/O Operations per second from SQL

With the introduction of Exadata X8, Oracle added Exadata Storage Server Extended (XT) for low-cost storage of infrequently accessed data. The X9M-2 XT storage server contains the same disk capacity as the X9M-2 High Capacity Storage Server, but does not contain Flash storage or PMem. It also can be configured without the Exadata Storage Server Software license.

Memory-level performance with shared storage

Architects of traditional computing platforms have always had to cope with technology change affecting the design of their systems. The goal is to eliminate bottlenecks so that the storage output moves through the network and is processed by compute servers without any slowdown. Solving an imbalance generally involves adding faster or more network connections or compute servers. This was before the advent of ultra-fast PCIe Flash memory and the NVMe Flash interface, and before persistent memory.

Flash memory started to become mainstream in corporate computing around 2010, used as a cache in front of hard disks or as a replacement for disks entirely. Every year after that, Flash capacity and performance increased significantly. In 2017, leading-edge Flash performance crossed a threshold, where the most advanced networks were unable to match the performance of Flash, and became a substantial bottleneck. As an example, a popular all-Flash storage system with 480 Flash cards is rated at only 75 GB/s of data throughput, whereas without a network bottleneck, that many Flash cards should produce over 2,600 GB/s of data throughput. Offloading to storage in Exadata bypasses this network bottleneck by filtering out unneeded data in storage before sending the remaining data across the network. The addition of persistent memory, which is faster than Flash, increased the value of Exadata offloading even further. Note that while plugging Flash directly into a compute server removes the network bottleneck, it also removes the ability to share storage across multiple compute servers. Exadata's approach does not suffer from this limitation.

ORACLE

Hardware specifications

EXADATA GENERATION (2-SOCKET)*	V1	V2	X2-2	X3-2	X4-2	X5-2	X6-2	X7-2	X8-2	X8M-2	X9M-2
Date Introduced	Sep 2008	Sep 2009	Sep 2010	Sep 2012	Nov 2013	Jan 2015	Apr 2016	Oct 2017	Apr 2019	Sep-2019	Sep 2021
Last Ship Date	Oct 2009	Oct 2010	Sep 2012	Feb 2014	Mar 2015	Jul-2016	Nov 2017	Jun 2019	still shipping	still shipping	still shipping
Specifications	N/A	Data Sheet [24]	Data Sheet [25]	Data Sheet [26]	Data Sheet [27]	Data Sheet [28]	Data Sheet [29]	Data Sheet [30]	Data Sheet [31]	Data Sheet [66]	Data Sheet [1]
Operating System	Linux	Linux	Linux	Linux	Linux	Linux	Linux	Linux	Linux	Linux	Linux
Disk Storage (raw TB)	168	336	504	504	672	1344	1344	1680	2352	2352	3024
Flash Cache (raw TB)	N/A	5.3	5.3	22.4	44.8	89.6	179.2	358.4	358.4	358.4	358.4
Persistent Memory (TB)	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	21	21
Extreme Flash (raw TB)	N/A	N/A	N/A	N/A	N/A	179.2	358.4	716.8	716.8	716.8	716.8
Compute Cores	64	64	96	128	192	288	352	384	384	384	512
Max Memory (GB)	256	576	1,152	2,048	4,096	6,144	12,288	12,288	12,288	12,288	16,384
RDMA Network Fabric (Gb/s)	20	40	40	40	40	40	40	40	40	100	100
Ethernet (Gb/sec)	8	24	184	400	400	400	400	800	800	800	800

* Per 1 full rack with 8 compute servers and 14 storage servers

ORACLE

EXADATA GENERATION (8-SOCKET)**	X2-8	X3-8	X4-8	X5-8	X6-8	X7-8	X8-8	X8M-8	X9M-8
Date Introduced	Sep 2010	Sep 2012	Jul 2014	Nov 2015	Apr 2016	Oct 2017	Apr 2019	Sep- 2019	Sep 2021
Last Ship Date	Nov 2012	Dec 2014	Oct 2015	Mar 2016	Nov 2017	Jun 2019	still shipping	still shipping	still shipping
Specifications	Data Sheet [32]	Data Sheet [33]	Data Sheet [34]	Data Sheet [35]	Data Sheet [36]	Data Sheet [37]	Data Sheet [38]	Data Sheet [39]	Data Sheet [67]
Operating System	Linux	Linux	Linux	Linux	Linux	Linux	Linux	Linux	Linux
Disk Storage (raw TB)	504	504	672	1344	1344	1680	2352	2352	3024
Flash Cache (raw TB)	5.3	22.4	89.6	89.6	179.2	358.4	358.4	358.4	358.4
Persistent Memory (TB)	N/A	N/A	N/A	N/A	N/A	N/A	N/A	21	21
Extreme Flash (raw TB)	N/A	N/A	179.2	179.2	358.4	716.8	716.8	716.8	716.8
Compute Cores	96	160	240	288	288	384	384	384	384
Max Memory (TB)	4	4	12	12	12	12	12	12	12
RDMA Network Fabric (Gb/s)	40	40	40	40	40	40	40	100	100
Ethernet (Gb/sec)	176	176	180	180	180	540	540	540	540

** Per 1 full rack with 2 compute servers and 14 storage servers

ORACLE

Elastic Configurations

Prior to the X5-2 generation, Exadata systems were only available in fixed-size configurations of Eighth, Quarter, Half and Full Rack sizes. The X5-2 Exadata release in January 2015 introduced elastic configurations. An elastic configuration has a customer-specified combination of database servers and storage servers, allowing individual storage or compute servers to be added to a base configuration until the physical rack is full.

The ratio of compute to storage servers can vary, depending on the characteristics of the intended workload. For example, an Exadata system optimized for in-memory database processing would add many compute servers, each with maximum memory. Conversely, an Exadata system optimized for a large data warehouse could add many High-Capacity storage servers. Elastic configurations may also be used to scale out earlier generation Exadata systems using the latest compatible servers. For example, X9M-2 Exadata Database and Storage Servers can be added to an existing X8M-2 Exadata Database Machine enabling customers to grow capacity as required. Exadata Database Machines have always been able to span multiple racks using the built-in internal network fabric connections extending Exadata's scale out design beyond a single physical rack.

Exadata Evolution

Oracle releases a new generation of Exadata every twelve to eighteen months. At each release, Oracle refreshes most hardware components to the latest CPU processors, and memory, disk, flash and networking technologies. These hardware refreshes result in performance increases with every release. Exadata software innovations, delivered with each generation and periodically in between, consistently enhance performance, availability, security, management and workload consolidation. In October 2015, features to support the Oracle Cloud were introduced.

The evolution of Exadata is best understood through the innovations introduced in each generation.

Exadata V1, released in 2008, focused on accelerating data warehousing by delivering the full throughput of storage to the database. Exadata achieved this by moving database filtering operations into storage, instead of sending all data to the compute servers and filtering it there. Oracle refers to this capability as Exadata Smart Scan.[40][41] Exadata V1 also supported a consolidation feature for allocating I/O bandwidth between databases or workloads, called IORM (I/O Resource Manager).[42] Exadata V1 was available in Full Rack or Half Rack sizes, and the choice of High Performance or High Capacity storage servers.

Exadata V2[24][43][44], released in 2009, added a Quarter Rack configuration and support for OLTP workloads via Flash storage and database-aware Flash Caching.[45]

Exadata V2 also introduced Hybrid Columnar Compression to reduce the amount of storage consumed by large Data Warehouse tables.

Storage Indexes[46] in Exadata V2 increased performance by eliminating the need to read entire regions of storage, based on the storage server's knowledge of the values contained in the region.

Exadata X2-2[25], the third generation, was released in 2010 and a second model of Exadata, **Exadata X2-8**[32], was introduced. The X2-8 and subsequent "8 socket" Exadata models feature Intel processors targeted at large memory, scale-up workloads. The use of Flash storage beyond caching began in this release with a Smart Flash Logging[47][48] feature. Support for 10 Gigabit per second (Gb/sec) Ethernet connectivity was also added.

Exadata X2-2 also encouraged data security through encryption with the incorporation of hardware decryption[49] , which largely eliminated the performance overhead of software decryption.

A Storage Expansion Rack[50] based on Exadata X2-2 was added in 2011 to accommodate large, fast-growing data warehouses and archival databases. All subsequent 2-socket Exadata generations have included a new Storage Expansion Rack.

ORACLE

Exadata X3-2[51] and **X3-8** were released in 2012, including a new Eighth Rack X3-2 entry-level configuration. Flash storage capacity quadrupled and OLTP write throughput reportedly increased by 20x via the Write-Back Flash Cache feature.

A number of availability enhancements were added, bypassing slow or failed storage media[52], reducing the duration of storage server brownouts and simplifying replacement of failed disks.

Exadata X4-2 was released in 2013. Flash capacity doubled and Flash compression was added, effectively doubling capacity again. Network Resource Management[15] was introduced, automatically prioritizing critical messages. InfiniBand bandwidth doubled with support for active/active connections.

Exadata X4-8 released in 2014, adding Capacity on Demand[53] licensing, I/O latency capping and timeout thresholds.

Exadata X5-2 and **X5-8** were released in 2015 with a major set of enhancements. Flash and disk capacity doubled. Elastic configurations[54] were introduced to enable expansion one server at a time. Virtualization was added as an option to Exadata along with Trusted Partitions[55] for flexible licensing within a virtual machine. Database snapshots[56] on Exadata storage enabled efficient development and testing. Oracle Database In-Memory on Exadata included Fault Tolerant[57][58] redundancy. The High Performance Exadata storage servers were replaced with all-Flash (Extreme Performance) storage servers and Exadata became the first major vendor to adopt the NVMe Flash interface. Columnar Flash cache was introduced to automatically reformat analytics data into columnar format in Flash. IPv6 support was completed. Exadata Cloud Service[59] was launched on the Oracle Cloud.

Exadata X6-2 and **X6-8** were released in 2016. Flash capacity doubled. Exafusion Direct-to-Wire protocol[60] reduced messaging overhead in a cluster and Smart Fusion Block Transfer[61] eliminated log write delays for OLTP applications in a cluster. Exadata Cloud@Customer debuted, enabling Oracle Cloud benefits within corporate data centers.

Exadata X7-2 and **X7-8** were released in 2017[62]. Flash capacity doubled. Flash cards became hot-pluggable for online replacement. 10 Terabyte (TB) disk drives debuted along with 25 Gb/sec Ethernet connectivity. Oracle Database In-Memory processing was extended into Flash storage, and storage server DRAM was utilized for faster OLTP.

Exadata X8-2 and **X8-8** were released in April 2019. Exadata Storage Server Extended (XT) was introduced for low-cost storage of infrequently accessed data. 14 Terabyte (TB) disk drives debuted along with 60% more compute cores in Exadata storage servers. Machine Learning algorithms were added that automatically monitor CPU, network and memory to detect anomalies such as stuck processes, memory leaks and flaky networks, and to automatically create (Auto index), rebuild or drop indexes. Optimizer statistics are also gathered in real-time as DML executes. For enhanced security, Advanced Intrusion Detection Environment (AIDE) was added to detect and alert when unknown changes to system software are made.

Exadata X8M-2[66] and **X8M-8** were released in September, 2019. Substantial performance increases[63] resulted from the addition of Intel Optane DC Persistent Memory in Exadata Storage Servers, and a new 100 Gbit/s internal network fabric based on RoCE (RDMA over Converged Ethernet), replacing the previous InfiniBand fabric. These changes increased Read I/O throughput by 2.5x and lowered I/O latency 10x. In addition. a new KVM based virtual machine replaced the previous Xen based one, doubling the amount of memory available to a guest VM.

Exadata X9M-2[1] and **X9M-8**[67] were released in September 2021 and includes the latest generation of Intel Optane Persistent Memory and PCIe Gen 4 which led to significant performance gains over the previous generation. OLTP Read I/O throughput increased a further 1.6x and the 1TB/s Smart Scan threshold was crossed within a single rack.

ORACLE

References

1.   Data Sheet: Oracle Exadata X9M-2 (PDF).
2.   Loaiza, J., et al. Oracle Exadata: World's Fastest Database Machine (video), August 2011.
3.   Umamageswaran, K. and Goindi, G. Oracle Exadata: Architecture and Internals Technical Deep Dive (PDF), October 2018.
4.   Oracle Exadata System Software User's Guide 19.3.0.
5.   Exadata Turns 10. siliconangle.com, June 2019
6.   Hollis, C., Chuck's Blog: Grown-up IT for Grown-Up Applications, September 2015.
7.   Oracle Database Exadata Cloud Service Technical Data Sheet (PDF).
8.   Spendolini, Brian, Oracle Database Exadata Cloud Service: A Beginner's Guide. McGraw-Hill, November 2018.
9.   Oracle Gen 2 Exadata Cloud@Customer (PDF).
10. Oracle Cloud@Customer Data Sheet and Resources.
11. Exadata Customer Success Stories.
12. Gartner Peer Insights: Oracle Exadata Database Machine. Gartner peer insights.
13. Data Sheet: Oracle Database In-Memory (PDF). Oracle Technology Network.
14. A Technical Overview of the Oracle Exadata Database Machine and Exadata Storage Server(PDF).
15. Loaiza, J. Guaranteeing Fast Response for OLTP Messages (video), July 2016.
16. Rivenes, A. Columnar Formats in Exadata Flash Cache. Oracle blogs, October 2017.
17. Oracle Exadata Database Machine Documentation Release 19.3.
18. Oracle Exadata Statement of Direction (PDF), October 2019.
19. Oracle Cloud Pricing.
20. Loaiza, Juan. Direct Access from Database to Network for Much Faster OLTP, July 2016.
21. Loaiza, Juan. Introducing Oracle Exadata X8M, September 2019.
22. Data Sheet: Oracle Exadata Storage Expansion Rack X8-2 (PDF).
23. Oracle Exadata Configuration Assistant (OECA).
24. Data Sheet: Sun Oracle Database Machine (PDF).
25. Data Sheet: Oracle Exadata Database Machine X2-2 (PDF).
26. Data Sheet: Oracle Exadata Database Machine X3-2 (PDF).
27. Data Sheet: Oracle Exadata Database Machine X4-2 (PDF).
28. Data Sheet: Oracle Exadata Database Machine X5-2 (PDF).
29. Data Sheet: Oracle Exadata Database Machine X6-2 (PDF).
30. Data Sheet: Oracle Exadata Database Machine X7-2 (PDF).
31. Data Sheet: Oracle Exadata Database Machine X8-2 (PDF).
32. Data Sheet: Oracle Exadata Database Machine X2-8 (PDF).
33. Data Sheet: Oracle Exadata Database Machine X3-8 (PDF).
34. Data Sheet: Oracle Exadata Database Machine X4-8 (PDF).
35. Data Sheet: Oracle Exadata Database Machine X5-8 (PDF).
36. Data Sheet: Oracle Exadata Database Machine X6-8 (PDF).
37. Data Sheet: Oracle Exadata Database Machine X7-8 (PDF).
38. Data Sheet: Oracle Exadata Database Machine X8-8 (PDF).
39. Data Sheet: Oracle Exadata Database Machine X8M-8 (PDF).
40. Loaiza, J. Exadata Smart Scan: Moving Queries to Data is Required to Realize Flash Performance (video), July 2018.
41. Osborne, Kerry. Understanding Exadata Offloading (PDF). Kerry Osborne's Oracle Blog, August 2011.
42. Fusek and Spiller. Oracle Exadata System Software User's Guide, (20.1.0).
43. Data Sheet: Sun Oracle Exadata Storage Server (PDF).
44. Oracle and Sun Unveil Exadata Version 2 (video), Sept 2009
45. Subramaniam, M. Exadata Smart Flash Cache Features and the Oracle Exadata Database Machine (PDF), December 2013.
46. Loaiza, J. Storage Indexes: By Understanding Data, Exadata Can Radically Reduce I/O (video), July 2016.
47. Loaiza, J. Smart Logging Enables Fast OLTP on Flash (video), July 2016.
48. Fusek, P. Oracle Exadata Smart Flash Log (video), November 2011.
49. Nguyen, K. Securing Databases with Intel Advanced Encryption Standard New Instructions (AES-NI). Intel, December 2012.
50. Various authors (2017). Data Sheet: Oracle Exadata Storage Expansion Rack X7-2 (PDF), 2017.
51. Gorbachev, A. A Brief History of Exadata Time (PDF). NoCOUG Journal, May 2013.
52. Loaiza, J. Instant Server Failure Detection (video), July 2016.
53. Stern, J. Exadata Database Machine Licensing Information User's Guide, 18c. pp. 5–6., 2018
54. Stern, J. Exadata Database Machine System Overview. pp. 14–17, 2018.
55. Goindi, G. Controlling Software Costs on Exadata. Oracle Blogs, February 2015.
56. Shah, M. Storage Efficient Database Copy Options with Exadata (PDF). Oracle Technology Network, March 2015.
57. Loaiza, J. Duplicating In-Memory Data Assures Availability (video), July 2016.
58. Chavan, S. Exadata High Availability: In-Memory Fault Tolerance (video), June 2019.
59. Data Sheet: Oracle Database Exadata Cloud Service (PDF).
60. Loaiza, J. Direct Access From Database to Network for Much Faster OLTP (video), July 2016.
61. Loaiza, J. Non-Blocking Transfers of OLTP Data on Clusters (video), July 2016.

ORACLE

62.   Loaiza, J. Oracle's Next Generation Exadata Database Machine X7, 2017
63.   Brown, V. Oracle Unleashes World's Fastest Database Machine, September 2019.
64.   Oracle Autonomous Database.
65.   Loaiza, J. New Exadata X8M PMem and RoCE Capabilities and Benefits, September 2019.
66.   Data Sheet: Oracle Exadata Database Machine X8M-2 (PDF).
67.   Data Sheet: Oracle Exadata Database Machine X9M-8 (PDF).

All references by Oracle unless otherwise noted.

Principal author: Oracle

Redwood Shores, Calif., September 2021

Connect with us

Call +**1.800.ORACLE1** or visit **oracle.com**. Outside North America, find your local office at: **oracle.com/contact**.

  blogs.oracle.com   facebook.com/oracle   twitter.com/oracle