ORACLE

Oracle Exadata Storage Expansion X9M-2

The Oracle Exadata Database Machine is engineered to deliver higher performance and capacity at lower costs than other platforms, for all types of modern database workloads including Online Transaction Processing (OLTP), Data Warehousing (DW), In-Memory Analytics, Internet of Things (IoT), financial, gaming and compliance data management, as well as efficient consolidation of mixed workloads. The Oracle Exadata Storage Expansion Rack is engineered to be the simplest, fastest and most robust way to add additional storage capacity to an Exadata Database Machine. A natural extension of the Exadata Database Machine, the Exadata Storage Expansion Rack can be used to satisfy the requirements of the largest mission critical databases.

Oracle Exadata Storage Expansion X9M-2

Engineered for Fast and Reliable Deployment

The Exadata Database Machine is the most cost-efficient and highest performance platform for running Oracle Databases. Exadata is easy to deploy even for the most mission-critical systems, as the database servers, storage servers and network are pre-configured, pre-tuned and pre-tested by Oracle experts. Extensive end-to-end testing and validation ensures all components including database software, operating system, hypervisor, drivers, firmware, work seamlessly together and that there are no performance bottlenecks or single points of failure. The Exadata Storage Expansion Rack is built using the same secure, highly available and scale-out architecture as the Exadata Database Machine enabling storage to be added when it is needed, rather than having to size for peak capacity requirements.

Extreme Performance and Capacity

The Exadata Storage Expansion X9M rack enables you to grow the Exadata storage capacity and bandwidth of any RoCE based Exadata Database Machine. It is designed for database deployments that require very large amounts of data including: historical or archive data, backups, documents, images, XML, LOBs, etc. The expansion rack is extremely simple to configure as there are no LUNs or mount points to set up. Storage is configured and added to a database online with a few simple commands.

Key Features

- Grow the storage capacity of RDMA over Coverged Ethernet (RoCE) Oracle Exadata Database Machines

- 100 Gb/sec RoCE Network

- Includes from 4 to 19 Oracle Exadata Storage Servers

- Mirrored usable capacity of up to 1,663TB per rack before compression

- Up to 608 CPU cores dedicated to SQL processing in storage

- Up to 28.5 TB of Persistent Memory Acceleration per rack

- Complete redundancy for high availability

ORACLE

Extreme System Scalability and Growth with Elastic Configurations

The Exadata Storage Expansion X9M offers you more flexibility than ever to grow. Exadata Storage Expansion X9M can be configured and purchased as small as a Quarter Rack with four storage servers, additional storage servers can be added one at a time or as many as needed up to a maximum of 19. With the flexibility of adding between 4 and 19 storage servers, there is a configuration that fits any application. In addition to upgrading from a small to large Exadata Storage Expansion X9M, Oracle continues to use a building-block approach to connect the Exadata Storage Expansion X9M to the Exadata Database Machine X9M using the integrated RoCE Network fabric to easily scale the system to any size. Exadata Storage Expansion X9M can be coupled to Exadata Database Machine X9M Elastic Rack systems in almost any combination. Up to 12 Exadata Database Machine X9M racks and Exadata Storage Expansion X9M racks can be easily connected via RoCE cables and internal switches. An Exadata X9M-2 traditional full rack configuration[1], along with 11 Exadata Storage Expansion X9M racks[2], has a raw disk capacity of 47 Petabytes (48,168 TB) and 7,136 CPU cores dedicated to SQL processing. Larger configurations can be built with additional RoCE switches.

As new Exadata Storage Expansion X9M racks are connected to a RoCE based Exadata Database Machine, the storage capacity and performance of the system grows. The system can be run in single system image mode or logically partitioned for consolidation of multiple databases. Scaling out is easy with Exadata Database Machine and Exadata Storage Expansion Racks. Automatic Storage Management (ASM) dynamically and automatically balances the data across Exadata Storage Servers, online, evenly spreading the I/O load across the racks, fully utilizing all the hardware and easily integrating the expansion rack into the configuration. The I/O Resource Manager can also be used to apportion I/O bandwidth to different databases and users of the system to deliver on business service level targets.

Extreme Performance by Offloading Data Intensive Processing

As data volumes grow exponentially, conventional storage arrays struggle to quickly transfer data from disk and flash to database servers at a rate that keeps the CPUs busy. Modern servers with many CPUs can consume data at many tens to hundreds of gigabytes a second. This is far faster than conventional architectures that use storage arrays can deliver data through their storage heads and the storage network.

Key Benefits

- Pre-configured, pre-tested system optimized for all database applications

- Uncompressed I/O bandwidth of up to 1,425 GB/second per rack from SQL

- Engineered scale-out storage architecture pre-configured to easily expand system capacity and performance, online

- Simple upgrade to meet the needs of any size application

- Over 13 Petabytes of user data can be stored in a rack using the included Hybrid Columnar Compression

- Scale the configuration by connecting up to 12 Exadata Database Machines and Exadata Storage Expansion Racks without external switches. Larger configurations can be built with additional RoCE switches

---

[1] Traditional full rack configuration with 8x Exadata X9M-2 Database Servers and 14x Exadata X9M-2 High Capacity Storage Servers

[2] Fully configured Exadata Storage Expansion Rack with 19x Exadata X9M-2 High Capacity Storage Servers

ORACLE

The scale-out architecture of the Exadata Database Machine X9M not only provides high performance and scalability, it also includes a unique technology that offloads data intensive SQL operations into the Oracle Exadata Storage Servers. By pushing SQL processing to the Exadata Storage Servers, data filtering and processing occurs immediately and in parallel across all storage servers as data is read from disk, flash and persistent memory. Only the rows and columns that are directly relevant to a query are sent to the database servers.

For example, if a query is executed to identify the customers who placed sales orders over $1000 in the month of March, an Exadata system will: offload the scanning of the table to the Exadata storage; filter out all sales orders that are less than $1000; filter out sales orders not in March; and extract just the relevant customer names. The result is that the data transferred to the database servers is reduced by orders of magnitude. This greatly accelerates query execution, eliminates bottlenecks, and significantly reduces the CPU usage of the database servers.

Each Exadata Storage Server X9M-2 has two Xeon® x86 processors that are used for database offload. A single Exadata Storage Expansion X9M rack with 19 Storage Servers has a total of 608 processor cores that can be used to offload the database servers. The CPUs in Exadata Storage Servers do not replace database CPUs. Instead they accelerate data intensive workloads in a manner analogous to how graphics cards accelerate image intensive workloads.

Optimizing Storage Use and I/O Through Compression

The Exadata Storage Server provides a unique compression capability called **Hybrid Columnar Compression (HCC) that enables dramatic reductions in storage for large databases**. Hybrid Columnar Compression technology is an innovative method of organizing data within a database table that uses a combination of both row and columnar methods for storing data. This hybrid approach achieves the compression benefits of columnar storage, while avoiding the performance shortfalls of a pure columnar format.

With Hybrid Columnar Compression, Exadata enables the highest levels of data compression possible with Oracle Database and provides substantial cost-savings and performance improvements due to reduced I/O, especially for analytic workloads. Storage savings is data-dependent and often ranges from 5x to 20x. Average storage savings is an industry-leading 10x. On conventional systems, enabling high data compression has the drawback of reducing performance as it adds the load of decompression to the CPU. Exadata Database Machine is able to offload decompression to processors in Exadata storage, and in addition there is reduced I/O need because of the high compression achieved, most analytic workloads run faster using Hybrid Columnar Compression than without it.

Two modes of Hybrid Columnar Compression are available. **Warehouse compression** mode is suitable for read-intensive workloads and provides large storage savings and enhanced analytic performance. **Archive compression** mode provides the highest degree of compression and is targeted at data that is

ORACLE

seldom accessed but still must be kept online. In addition, this data can now be seamlessly stored on the XT storage server for further cost reduction.

On OLTP systems, Hybrid Columnar Compression can be used to compress older, less active data while newer, more active, and update-intensive data can be compressed using Advanced Row Compression. Oracle Database Release 18c and above provides the ability to change the type of compression used by individual table partitions online (even if there are global indexes on the table), to ensure seamless tiering across different compression types as data ages and becomes less active.

For data analytics, Exadata implements a unique algorithm to accelerate reporting and analytical queries, called **Exadata Columnar Flash Cache**. Columnar Flash Caching implements a dual format architecture in Exadata flash by automatically transforming frequently scanned Hybrid Columnar Compressed data into a pure columnar format as it is loaded into the flash cache. Smart scans on pure columnar data in flash run faster because they read only the selected columns, reducing flash I/Os and storage server CPU consumption. This accelerates reporting and analytic queries while maintaining excellent performance for OLTP-style single row lookups.

Groundbreaking RDMA based Network Fabric

The Exadata X9M release uses the same ultra-fast cloud scale networking fabric that was introduced in Exadata X8M, RDMA over Converged Ethernet (RoCE). RDMA (Remote Direct Memory Access) allows one computer to directly access data from another without Operating System or CPU involvement, for high bandwidth and low latency. The network card directly reads/writes memory with no extra copying or buffering and with very low latency. RDMA is an integral part of the Exadata high-performance architecture, and has been tuned and enhanced over the past decade, underpinning several Exadata-only technologies such as Exafusion Direct-to-Wire Protocol and Smart Fusion Block Transfer. As the RoCE API infrastructure is identical to InfiniBand's, **all existing Exadata performance features are available on RoCE.**

The Exadata X9M release implements a dual port PCIe Gen 4 network interface card capable of 2x 100Gb/sec active-active RoCE network for a total throughput of 200Gb/sec. This makes the world's fastest database machine even faster. Real world database workloads running on Exadata X9M, deployed with the new **shared persistent memory accelerator**, beating the previous benchmark of 16M Read IOPS, set by Exadata X8M, with **27.6 Million Read OLTP Read IOPS (8K IOs)[3].**

Shared Persistent Memory Acceleration

Exadata X8M introduced persistent memory (PMEM) data and commit accelerators in front of flash cache, enabling orders of magnitude lower latency accessing remotely stored data. Persistent memory is a modern silicon technology, adding a distinct storage tier of performance, capacity, and price

---

[3] Elastic configuration with 10x Exadata X9M-2 Database Servers and 12x Exadata X9M-2 Extreme Flash Storage Servers

ORACLE

between DRAM and Flash. As the persistent memory is physically present on the memory bus of the storage server, reads perform at memory speed, much faster than flash. Writes are persistent, surviving power cycles, unlike DRAM. By utilizing RDMA to access persistent memory remotely, **Exadata Persistent Memory Data and Commit Accelerators bypass the network and I/O stack, eliminating expensive CPU interrupts and context switches, reducing latency by 10x**, from 200μs to less than 19μs. Smart Exadata System Software also ensures data is mirrored across storage servers, which provides additional fault-tolerance. Exadata's unique end-to-end integration between Oracle Database and Exadata Storage Servers automatically caches the hottest data blocks efficiently between the database buffer cache, persistent memory, and flash cache. Adding persistent memory to the storage tier means the aggregate performance of this new cache tier can be dynamically used by any database on any server. This is a significant advantage over general-purpose storage architectures, which preclude sharing storage resources across database instances.

Another smart Exadata System Software feature boosts log write performance. Log write latency is critical for OLTP performance, a faster log write means faster commit times. Inversely, any slowdown of log writes can cause the database to stall. With the RoCE based Exadata, **Exadata Persistent Commit Accelerator** automatically enables the database to issue a one-way RDMA log write to persistent memory. RDMA and persistent memory technologies allow the log write to occur without acknowledgement, and smart software distributes the write across multiple servers for resilience. This leads to a performance increase for log write operations.

Security and management of this tier are also automated. Persistent memory is configured automatically at installation time, with no user interaction required. Hardware monitoring is pre-configured. Persistent memory is only accessible to databases using database access controls, ensuring end to end security of data. Deploying persistent memory in Exadata X9M is so simple, it's transparent.

Figure 1. Intel® Optane™ Persistent Memory modules

Extreme Flash Storage Server: Record-breaking I/O Performance

Exadata **Extreme Flash (EF) Storage Server** is the foundation of a database-optimized all-flash Exadata Database Machine. Each EF Storage Server contains eight 6.4 TB Flash Accelerator F640v3 NVMe PCI Flash drives, offering 51.2 TB raw flash capacity per EF Storage Server, with an expected endurance of 8 years or more for typical database workloads. Exadata delivers ultra-high performance by placing these flash devices directly on the high speed PCIe4 interface rather than behind slow disk controllers. Exadata X9M includes shared persistent

ORACLE

memory as an acceleration tier. Twelve 128 GB Intel® Optane™ Persistent Memory modules in front of flash to boost performance even more.

Exadata X9M uses a combination of scale-out storage, RDMA over Converged Ethernet networking, database offload, persistent memory accelerator and PCIe Flash to deliver extremely high performance from memory and flash. When added to an Exadata Database Machine X9M-2, a full rack of Exadata Storage Expansion X9M with 19 Extreme Flash storage servers, using the persistent memory accelerator, can add up to **43.7 Million random 8K database read and 11.6 Million random 8K flash write I/O operations per second (IOPS).**

For data warehouse environments that require the highest performance, Exadata X9M Extreme Flash storage servers are capable of **scanning up to 75GB/s per server** and increasing scan throughput by up to **1.39TB/s**[45]**.**

These are real-world end-to-end performance figures measured running SQL workloads with standard 8K database I/O sizes inside a single rack Exadata system. Exadata's performance on real Oracle Database workloads is orders of magnitude faster than traditional storage array architectures, and is also much faster than current all-flash storage arrays, whose architecture bottlenecks flash throughput.

Figure 2. Flash Accelerator PCIe Card

High Capacity Storage Server: Tiered Disk Flash And Persistent Memory Deliver Cost Of Disk With Shared Memory Performance

The second Exadata storage option is the **High Capacity (HC) Storage Server**. This server includes twelve 18 TB SAS disk drives with 216 TB total raw disk capacity. It also has four Flash Accelerator F640v3 NVMe PCIe cards with a total raw capacity of 25.6 TB of flash memory. Exadata X9M adds the shared persistent memory acceleration tier, twelve 128 GB Intel® Optane™ Persistent Memory modules in front of flash to boost performance even more. Deployed using smart software, **Exadata Smart PMEM Cache**, only the hottest database blocks are automatically cached in this new tier. Accessible over RDMA direct from the database delivers the highest I/O rates at an extremely low latency.

Flash in the HC Storage Server can be used directly as flash disks, but is almost always configured as a flash cache (**Exadata Smart Flash Cache**) in front of disk storage behind the PMEM Cache to deliver the best performance. Exadata Smart Flash Cache is used in-sync with PMEM Cache to automatically cache frequently accessed data while keeping infrequently accessed data on disk, delivering the high I/O rates and fast response times of flash with the large capacity and low

---

[4] Smart Scan rate measured at 1425 GB/sec

[5] Fully configured Exadata Storage Expansion Rack with 19x Exadata X9M-2 High Capacity Storage Servers

ORACLE

cost of disk. Exadata uniquely understands database workloads and **knows when to avoid caching data that will negatively affect overall performance**. For example, if large write I/Os caused by backups or large table scans are likely to disrupt higher priority OLTP or scan operations, those large I/Os will bypass the flash cache and go straight to disk. Otherwise, Exadata System Software will utilize additional spare flash capacity and I/O bandwidth to optimize performance by caching these I/Os. In addition to automatic caching, administrators can optionally provide SQL directives to ensure that specific tables, indexes, or partitions are preferentially retained in the flash cache.

It is common for hit rates in the Exadata Smart Flash Cache to be over 95%, or even 99% in real-world database workloads, yielding an effective flash capacity many times larger than the physical flash. For example, a traditional full rack[6], often has an effective flash capacity close to the usable disk capacity of 900 TB

The Exadata Smart Flash Cache also caches database block writes using Exadata Write Back Flash Cache technology. Write caching eliminates disk bottlenecks in large scale OLTP and batch workloads. The flash write capacity of a single full rack Exadata Storage Expansion X9M-2[7] adds up to **11.6 Million 8K flash write I/O operations per second (IOPS)** when connected to an existing Exadata Database Machine using the RoCE Fabric. The Exadata write cache is transparent, persistent, and fully redundant, with performance comparable to dozens of enterprise disk arrays with thousands of disk drives.

The automatic data tiering between RAM, persistent memory, flash and disk in Exadata provides tremendous advantages over other flash-based solutions. Many storage vendors have developed flash-only arrays to achieve higher performance than traditional arrays. These flash-only arrays deliver better performance but cannot match the cost advantages of Exadata's smart tiering of data between disk and flash, as the overall size of data that can benefit from flash is limited to the size of expensive flash. And these flash arrays are unable to benefit from Exadata's unique database-aware storage optimization technologies. Generic data deduplication provided by some flash arrays is effective for Virtual Desktop Infrastructure environments, but not for databases.

Exadata not only delivers much more capacity than generic all-flash arrays, it also delivers better performance. Flash-only storage arrays cannot match the throughput of Exadata's integrated and optimized architecture with 100 Gb/sec RDMA over Converged Ethernet based scale-out network, fast PCI Flash, offload of data intensive operations to storage, and algorithms throughout that are specifically optimized for databases.

Extended Capacity Storage Server: Much Lower Cost Exadata Storage for Low Use Data

A third Exadata storage option is the **Extended (XT) Storage Server**. Each Exadata XT Storage Server includes twelve 18 TB SAS disk drives with 216 TB total raw disk capacity.

---

[6] Traditional full rack with 8x Exadata X9M-2 Database Servers and 14x Exadata X9M-2 High Capacity Storage Servers

[7] Fully configured Exadata Storage Expansion Rack with 19x Exadata X9M-2 High Capacity Storage Servers

ORACLE

This storage option extends the operational and management benefits of Exadata to rarely accessed data that must be kept online. Exadata's Extended (XT) Storage Server is:

- Efficient – The XT server offers the same high capacity as the HC Storage server, including Hybrid Columnar Compression
- Simple – The XT server adds capacity to Exadata while remaining transparent to applications, transparent to SQL, and retains the same operational model
- Secure – The XT server enables customers to extend to low-use data the same Exadata security model and encryption used for online data
- Fast and Scalable – Unlike other low-use data storage solutions, the XT server is integrated to the Exadata fabric, for fast access and easy scale-out
- Compatible – The XT server is just another flavor of Exadata Storage server – you can just add XT servers to any Exadata rack
- Inexpensive –Exadata Storage Software licenses are optional.

With Exadata Extended (XT) Storage Server, enterprises can meet their long-term data retention compliance requirements with the same trusted and continually validated Exadata solution, avoiding the operational risks and costs of managing information lifecycle across multiple platforms.

The ability to combine Extreme Flash, High Capacity and Extended Storage within an Exadata configuration allows customers to define a true Information Lifecycle Management policy. As data ages, it can be moved between the three tiers of storage to ensure data is stored on the right medium for their usage and retention requirements. Coupled with Automatic Data Optimization, part of Oracle Advanced Compression, customers can define policies to automate this movement as well as moving between compression levels.

Extreme Backup & Recover Performance

A key strength of the Exadata Storage Expansion rack is when used as a destination for Exadata Database Machine backups. A full database backup can be created at up to 48 TB/hour when backing up uncompressed data that is being written to mirrored disk in an Exadata Storage Expansion rack. Taking incremental backups this increases the effective backup rate to **terabytes per hour**. When combining incremental backups and Hybrid Columnar Compression further increases the effective backup rate to **petabytes per hour.**

A disk backup on an Exadata Storage Expansion rack is usable directly without loss of performance and without having to do a restore. This is a unique backup capability only available when backing up to an Exadata Storage Expansion. It is by far the fastest and simplest way to backup and recover your Oracle Exadata Database Machine.

Mission Critical High Availability

The Exadata Database Machine is engineered to provide the highest levels of availability. **All types of failures are protected against** from simple failures such as disk, server, or network, to complex site failures and human errors. Each Exadata Database Machine has **completely redundant hardware**, including redundant networking, redundant Power Distribution Units (PDU), redundant

ORACLE

power supplies, and redundant database and storage servers. Oracle RAC protects against database server failure. Oracle ASM provides data mirroring to protect against disk or storage server failures. Oracle RMAN provides extremely fast and efficient backups to disk or tape. Oracle's Flashback technology allows backing out user errors at the database, table or even row level. Using Oracle Data Guard, a second Exadata Database Machine can be deployed in a Maximum Availability Architecture (MAA) configuration to transparently maintain a real-time copy of the database at a remote site and provide full protection against primary database failures and site disasters.

Exadata in an MAA configuration is recognized by the analyst firm IDC as a system that **delivers at least 5-nines availability** and is categorized in the IDC AL4 fault-tolerant market segment[8].

The Exadata principle of deep hardware and software integration is also evident in the many ways Exadata uniquely assures high availability across several different failure conditions. One such unique capability is **Instant Failure Detection**. On non-Exadata platforms, detecting a server failure requires waiting for a long timeout, leading to extended application brownouts. RoCE based Exadata implements a unique RDMA-based **sub-second node death detection**, leading to virtual elimination of application brownout conditions.

Disk and flash devices occasionally exhibit very long latency I/O operations due to internal recovery of failed sectors, internal firmware reboots, or wear leveling. These long I/O operations can cause stalls in mission critical OLTP databases. With Exadata's **unique I/O Latency Capping**, Oracle Exadata System software automatically redirects read I/O operations to an ASM-mirrored copy of the data when the latency of a read I/O is much longer than expected. Similarly, it automatically redirects high latency write I/O operations to a healthy flash device, eliminating outliers during write operations. Exadata System Software uses machine learning techniques to predict components susceptible to failure and takes proactive action to gracefully take such components out of service. If disks do fail, ASM performs a rebalance operation for the data that was resident on the disk, while applications continue to access the database with no interruption. Exadata allows hot swapping of disks, fans, power supplies, and PCIe Flash cards to avoid downtime. Exadata System software takes rebalance one step further by preserving the flash cache population and storage indexes when moving data between storage servers to maintain consistent application performance. On rare occasions when there are outliers within the networking subsystem, Exadata redirects the I/O issued by the database server to another storage server.

Exadata automates monitoring of CPU, memory, input/output subsystems, file system, and network. This automation combines machine learning techniques with the deep lessons learned from thousands of mission critical real-world deployments. For example, Exadata can detect that anomalous use of system resources that negatively impact database performance and automatically

8 Worldwide AL4 Server Market Shares, 2019: Fault-Tolerant Systems Become Digital Transformation Platforms, IDC, Paul Maguranis Peter Rutten, July 2020

ORACLE

identifies the process responsible and issues an alert – all without any manual intervention.

As a result of its industry leading availability, the Exadata Database Machine has been deployed by leading companies for their most critical applications including interbank fund transfers, online securities trading, real-time call tracking, and web-based retailing, and many more. Exadata's Mission Critical availability capabilities are not restricted to OLTP workloads; they also apply to warehousing and analytics.

Highest Level of Service

Oracle offers a complete set of support services for the Exadata family of products including: 24x7 hardware service, system monitoring, software installation and configuration among other standard and custom offerings.

Of particular value is **Oracle Platinum Services** that is available exclusively for Oracle's Engineered Systems. Platinum Services provides fault monitoring, faster response times, and expedited escalation to development. With Platinum Services, Oracle support engineers perform software maintenance and patching remotely. Platinum Services provides a higher level of support than has ever been available before for all software and hardware within an Engineered System including the Oracle Database. Platinum Services is provided at no extra charge to Exadata customers.

IT Agility

Exadata is a complete system for running databases including storage, servers, and internal networks. Management of a traditional database system is typically spread across the management teams of each of the components such as the database team, the storage team, and the system administration team. In contrast, an **Exadata system is typically managed by a unified Database Machine Administration (DMA) team**. Database Machine Administrators have full control of all resources in the Exadata Database Machine including storage resources. New database deployments and configuration changes can be implemented by the Database Machine Administrators without coordination across different component management teams that are often overloaded and have differing priorities. Database Machine Administrators can focus on application and business specific enhancements rather than coordinating across component teams, or tuning and triaging of low level configuration issues.

Dramatically Lower Costs

Due to the extreme performance, high storage capacity, and unique compression capabilities delivered by the Exadata Database Machine, workloads that would require very large traditional hardware systems can be run on much smaller Exadata systems. In sizing exercises it is typical to see a 2-4x reduction in Exadata system size compared to a traditional system.

ORACLE

Exadata provides a huge RAM, flash, and disk footprint for large data sets. Raw disk storage on a full rack Exadata Storage Expansion[9] can reach 3.9 Petabytes while raw flash storage can be up to 972 TB. In addition, Hybrid Columnar Compression often expands storage and memory capacity 10X. By intelligently moving active data across disk, flash, and memory tiers, Exadata simultaneously delivers the highest performance and the lowest cost.

Exadata has the **unique** ability to consolidate many databases supporting multiple workloads in a single cloud platform. High-end OLTP, analytics, batch, reporting, and backups can all run simultaneously within and across databases with extreme performance. **The extreme performance and capacity of Exadata enables very large numbers of databases and workloads to be consolidated on Exadata**. Consolidating databases on Exadata reduces system hardware cost, software cost, and greatly reduces ongoing operations cost.

The uniformity of Exadata Database Machine configurations results in large cost savings. **Exadata standardizes not just technologies, but also integration, testing, security, hardening, tuning, and support**. Customers deploy Exadata systems much faster and with a lot less labor than traditional systems. Low level tuning, integration, and maintenance is reduced or eliminated. All Exadata users run a configuration that is identical to thousands of other users, and is identical to Oracle's internal configurations, it is far less likely that issues will be encountered, and issue resolution is quicker and simpler reducing both operations cost and downtime cost. When issues do occur, customers have to deal with one supplier – Oracle, as the entire system – hardware, firmware, operating system, hypervisor, and database layers are all owned and supported by Oracle. The "one-hand-to-shake" support model enables faster problem resolution times and reduces downtime further increasing benefits.

Exadata Business Benefits

Beyond the operational benefits of extreme performance, availability, security, and deployment flexibilities across on-premises and cloud, Exadata also directly benefits the business.

Exadata accelerates time to market for new business applications since the time needed for system configuration, tuning, and testing is largely eliminated. Deployment times are reduced from months to days, and the risk of unexpected system level issues after go-live is greatly reduced. When a new application is deployed, it is common for unanticipated application usage patterns to create performance issues. Exadata's huge I/O, network, and compute throughput can absorb spikes created by unanticipated workloads without slowing response times of mission critical workloads. Overall Exadata speeds application deployment and reduces risk, allowing businesses to innovate faster.

Exadata's extreme performance, large memory, and flash capacity enhances employee productivity and customer satisfaction by greatly improving user response times. **Users spend more time doing useful work, and less time waiting** for the system to respond.

---

[9] Fully configured Exadata Storage Expansion Rack with 19x Exadata X9M-2 High Capacity Storage Servers

ORACLE

Exadata's extreme performance does not just improve business efficiency, it also **enables business users to make smarter decisions, discover growth opportunities, and reduce costs**. Users can analyze data in real-time, explore different possibilities, and perform rapid iteration to find better solutions. Exadata enables:

- Real-time business data analysis
- Faster financial closes
- Better planning and budgeting
- More effective and faster projections

Conclusion

Exadata delivers a fully integrated database platform with the latest hardware technologies and **unique** software to deliver extreme performance, availability, and security. This coupled with cost savings, ease of management, and enhanced supportability result in greater business agility and efficiency. Given what can be achieved with Exadata, it is no surprise it is the new global standard for running Oracle Databases –on-premises, or in the cloud.

ORACLE

Exadata Storage Expansion X9M Key Capacity and Performance Metrics

METRIC	QUARTER RACK HIGH CAPACITY	QUARTER RACK EXTREME FLASH	SINGLE SERVER HIGH CAPACITY (HC)	SINGLE SERVER EXTREME FLASH (EF)	SINGLE SERVER EXTENDED (XT)	MAXIMUM CONFIG HIGH CAPACITY	MAXIMUM CONFIG EXTREME FLASH
Number of Storage Servers	4	4	1	1	1	19	19
Flash Metrics							
Maximum SQL flash bandwidth [2]	180 GB/s	300 GB/s	45 GB/s	75 GB/s	0 GB/s	855 GB/s	1425 GB/s
Maximum SQL flash read IOPS [3]	9,200,000	9,200,000	2,300,000	2,300,000	0	43,700,000	43,700,000
Maximum SQL flash write IOPS [4]	2,456,000	2,456,000	614,000	614,000	0	11,666,000	11,666,000
PCI Flash capacity (raw)	102.4 TB	204.8 TB	25.6 TB	51.2 TB	0.0 TB	486.4 TB	972.8 TB
Disk Metrics							
Maximum SQL disk bandwidth [2]	7.2 GB/s	N/A	1.8 GB/s	N/A	N/A	34 GB/s	N/A
Maximum SQL disk IOPS [3]	10,000	N/A	2,600	N/A	N/A	49,000	N/A
Data capacity (raw) [5]	864 TB	204 TB	216 TB	51 TB	216 TB	4104 TB	973 TB
Combined Metrics							
Data Capacity (usable) Normal Redundancy [6]	327.1 TB	75 TB	82 TB	19 TB	82 TB	1664 TB	382.5 TB
Data Capacity (usable) High Redundancy [6]	256.6 TB	59.0 TB	64.1 TB	14.7 TB	64.1 TB	1218.7 TB	280.2 TB
Maximum data load rate [7]	10 TB/hour	10 TB/hour	2.5 TB/hour	2.5 TB/hour	2.5 TB/hour	47.5 TB/hour	47.5 TB/hour

Actual system performance varies by application.

[1] EF = Extreme Flash; HC = High Capacity

[2] Bandwidth is peak physical scan bandwidth achieved running SQL, assuming no database compression. Effective user data bandwidth is higher when database compression is used.

[3] Based on 8K IO requests running SQL. Note that the IO size greatly affects Flash IOPS. Others quote IOPS based on smaller IOs and are not relevant for databases.

[4] Based on 8K IO requests running SQL. Flash write I/Os measured at the storage servers after ASM mirroring, which usually issues multiple storage IOs to maintain redundancy.

[5] Raw capacity is measured in standard disk drive terminology with 1 GB = 1 billion bytes. Usable capacity is measured using normal powers of 2 space terminology with 1 TB = 1024 * 1024 * 1024 * 1024 bytes.

[6] Actual space available for a database after mirroring (ASM normal redundancy) while also providing adequate space (one disk on Quarter and Half Racks and two disks on a Full Rack) to reestablish the mirroring protection after a disk failure in the normal redundancy case.

[7] Effective Flash Capacity is larger than the physical flash capacity and takes into account the high flash hit ratios due to Exadata's intelligent flash caching algorithms, and the size of the underlying disk storage. It is the size of the data files that can often be stored in Exadata and be accessed at the speed of flash memory.

[8] Load rates are typically limited by database server CPU, not IO. Rates vary based on load method, indexes, data types, compression, and partitioning.

ORACLE

Exadata Storage Expansion X9M Hardware

QUARTER RACK

4 x Exadata Storage Server X9M-2:
 	
    • 128 CPU cores for SQL processing
    • 48x 128 GB Persistent Memory
    • 48x 18 TB High Capacity Drive and 16x 6.4 TB NVMe PCI Flash Cards for HC Quarter Rack, or
    • 32x 6.4 TB NVMe PCI Flash Drives for EF Quarter Rack
    • 3x 36-port 100 Gb/s RoCE switches

Additional Hardware Components:
 	
    • 42U Rack
    • Ethernet switch for administrative connectivity
    • 2x Redundant Power Distribution Units (PDUs)

Included Spare Parts Kit Contains:
 	
    • 1x 6.4 TB NVMe PCI Flash Card and 1 x 18 TB High Capacity disk, or
    • 1x 6.4 TB NVMe PCI Flash drive

Exadata Storage Expansion X9M Connectivity and Upgrades

CONNECTION TO EXADATA DATABASE MACHINE	UPGRADABILITY

Connect any combination of up to 12 Exadata Database Machine X9M racks or Exadata Storage Expansion X9M Racks via the RoCE Network Fabric.
 	
Larger Configurations can be built with external RoCE switches. Connected racks can be X9M or X8M generation hardware
	After the initial quarter rack, additional HC, EF, XT or combination of HC, EF and XT storage servers can be added one at a time or as many as needed up to a maximum configuration (19 storage servers).
	
	Hardware Components included with the upgrade:
	
	    • RoCE and Ethernet cables and adapters to connect all the components
	    • 12 x 18 TB High Capacity Drives, 4 x 6.4 TB NVMe PCI Flash Cards, and 12 x 128 GB Persistent Memory Modules for each additional HC storage server, or
	    • 8 x 6.4 TB NVMe PCI Flash Cards and 12 x 128 GB Persistent Memory Modules for each additional EF storage server, or
	    • 12 x 18 TB High Capacity Drives for each additional XT storage

Upgrade Support Services:
• Hardware Installation and Software Configuration

ORACLE

Exadata Storage Expansion X9M Environmental Specifications

METRIC	QUARTER RACK	MAXIMUM CONFIGURATION	SINGLE SERVER
Height	78.74 in. (2000 mm)		3.42 in. (86.9 mm)
Width	23.66 in. (601 mm)		17.52 in. (445.0 mm)
Depth	47.13 in. (1197 mm)		29.88 in. (759.0 mm)
Acoustic Noise (operating)	9.2 B	9.5 B	8.0 B
Environmentals With High Capacity Disks			
Weight	919.7 lb (417.2 kg)	2069.2 lb (938.6 kg)	76.7 lb (34.8 kg)
Maximum Power Usage	20.1 kW (20.5 kVA)	17.2 kW (17.6 kVA)	0.8 kW (0.9 kVA)
Typical Power Usage [1]	3.2 kW (3.3 kVA)	12.1 kW (12.3 kVA)	0.6 kW (0.6 kVA)
Cooling at Maximum Usage	15,590 BTU/hour	58,839 BTU/hour	2,883 BTU/hour
	16,448 kJ/hour	62,075 kJ/hour	3,042 kJ/hour
Cooling at Typical Usage	10,913 BTU/hour	41,187 BTU/hour	2,018 BTU/hour
	11,513 kJ/hour	43,453 kJ/hour	2,129 kJ/hour
Airflow at Maximum Usage [2]	722 CFM	2724 CFM	133 CFM
Airflow at Typical Usage [2]	505 CFM	1907 CFM	93 CFM
Environmentals With Extreme Flash Disks			
Weight	855.3 lb (388.0 kg)	1763.3 lb (799.8 kg)	60.6 lb (27.5 kg)
Maximum Power Usage	4.6 kW (4.7 kVA)	17.4 kW (17.7 kVA)	0.9 kW (0.9 kVA)
Typical Power Usage [1]	3.2 kW (3.3 kVA)	12.2 kW (12.4 kVA)	0.6 kW (0.6 kVA)
Cooling at Maximum Usage	15,686 BTU/hour	59,293 BTU/hour	2,907 BTU/hour
	16,548 kJ/hour	62,554 kJ/hour	3,067 kJ/hour
Cooling at Typical Usage	10,980 BTU/hour	41,505 BTU/hour	2,035 BTU/hour
	11,584 kJ/hour	43,788 kJ/hour	2,147 kJ/hour
Airflow at Maximum Usage [2]	726 CFM	2745 CFM	135 CFM
Airflow at Typical Usage [2]	508 CFM	1922 CFM	94 CFM

Operating temperature/humidity: 5 ºC to 32 ºC (41 ºF to 89.6 ºF), 10% to 90% relative humidity, non-condensing
Altitude Operating: Up to 3,048 m, max. ambient temperature is de-rated by 1° C per 300 m above 900 m
[1] Typical power usage varies by application load.
[2] Airflow must be front-to-back.

ORACLE

Exadata Storage Expansion X9M Regulations and Certifications

Regulations 1,2,3	Product Safety:	UL/CSA 60950-1, EN 60950-1, IEC 60950-1 CB Scheme with all country differences
		UL/CSA 62368-1, EN 62368-1, IEC 62368-1 CB Scheme with all country differences
	EMC	
	Emissions:	FCC CFR 47 Part 15, ICES-003, EN55032, KN32, EN61000-3-11, EN61000-3-12
	Immunity:	EN55024, KN35
Certifications [2,3]		North America (NRTL), CE (European Union), International CB Scheme, HSE Exemption (India), BSMI (Taiwan), CCC (PRC), EAC (EAEU including Russia), KC (Korea), RCM (Australia), VCCI (Japan), UKCA (United Kingdom)
European Union Directives [3]		2014/35/EU Low Voltage Directive. 2014/30/EU EMC Directive, 2011/65/EU RoHS Directive, 2012/19/EU WEEE Directive

[1] All standards and certifications referenced are to the latest official version. For additional detail, please contact your sales representative.
[2] Other country regulations/certifications may apply.
[3] In some cases, as applicable, regulatory and certification compliance were obtained for the shelf-level systems only.

Exadata Storage Expansion X9M Support Services

Components:

- Hardware Warranty: 1 year with a 4 hour web/phone response during normal business hours (Mon-Fri 8AM-5PM), with 2 business day on-site response/Parts Exchange
- Oracle Premier Support for Systems includes Oracle Linux and Solaris support and 24x7 with 2 hour on-site hardware service response (subject to proximity to service center)
- Oracle Premier Support for Operating Systems
- Oracle Customer Data and Device Retention
- System Installation Services
- Software Configuration Services
- Oracle Platinum Services
- Business Critical Service for Systems
- Oracle Exadata Start-Up Pack
- System Upgrade Support Services including hardware installation and software configuration
- Oracle Auto Service Request (ASR)

Key Features and Functionality

Exadata and Database Software Features - Analytics

- Unique Automatic Parallelization and Offload of Data Scans to storage
- Unique Filtering of Rows in Storage based on 'where' clause
- Unique Filtering of Rows in Storage based on columns selected
- Unique Storage Offload of JSON and XML Analytic Queries
- Unique Filtering of rows in Storage based on Join with other Table
- Unique Hybrid Columnar Compression
- Unique Storage Index Data Skipping
- Unique I/O Resource Management by User, Query, Service, DB, etc.
- Unique Automatic Transformation to Columnar Format in Flash Cache
- Unique Smart Flash Caching for Table Scans
- Unique Storage Offload of Index Fast Full Scans
- Unique Storage Offloads of Scans on Encrypted Data, with FIPS compliance
- Unique Storage Offload for LOBs and CLOBs
- Unique Storage Offload for min/max operations
- Unique Data Mining Offload to Storage

ORACLE

- Unique Reverse Offload to DB servers if Storage CPUs are Busy
- Unique Automatic Data Columnarization
- Unique Automatic Conversion of Data to In-Memory Formats when Loading into Flash Cache

Exadata and Database Software Features - OLTP

- Unique Persistent Memory Data Accelerator
- Unique Persistent Memory Commit Accelerator
- Unique Database Aware PCI Flash
- Unique Exadata Smart Flash Caching
- Unique Exadata Smart Flash Logging
- Unique Smart Write-back Flash Cache
- Unique I/O Prioritization by cluster, workload, DB or user to ensure QOS
- Unique Exafusion Direct-to-Wire Protocol
- Unique Database Intelligent Network Resource Management
- Unique Exachk full-stack validation
- Unique Full-stack security scanning
- Unique Database scoped security
- Unique Cell-to-Cell Rebalance preserving Flash Cache and Storage Index
- Unique Full-Stack Secure Erase
- Unique Instant Data File Creation
- Unique Smart Fusion Block Transfer
- Unique Control of Flash Cache Size per Database
- Unique In-Memory OLTP Acceleration
- Unique Undo-Block Remote RDMA Read
- Unique Support for 4000 Pluggable Databases per Container Database with Multitenant Option

Exadata and Database Software Features - High Availability

- Unique Instant Detection of Node or Cell Failure
- Unique In-Memory Fault Tolerance
- Unique Sub-second Failover of I/O on stuck disk or Flash
- Unique Offload backups to storage servers
- Unique Exadata Data Validation (extended H.A.R.D.)
- Unique Prioritize Recovery of Critical Database Files
- Unique Automatic Repair of Corrupt Disk Data By Reading Other Storage Servers
- Unique Avoidance of Read I/Os on Predictive failed disks
- Unique Confinement and power cycle of temporarily poor performing drives
- Unique Shutdown Prevention If Mirror Storage Server is Down
- Unique Detection and Disabling of Unreliable Network Links
- Unique Preservation of Storage Index on Rebalance

Manageability Features

- Oracle Embedded Integrated Lights Out Manager (ILOM) with upgrade pre-staging optimizations
- Oracle Enterprise Manager Exadata Plug-in
- Unique Active AWR includes storage stats for end to end monitoring
- IPv6 Support for Ethernet Connections
- Capacity on Demand
- Cell software transparent restart
- Flash and disk life cycle management alert
- Automatic Disk Scrub and Repair
- Trusted Partitions for Oracle Linux Virtualization
- Automated VLAN Creation

ORACLE

- Oracle Exadata Deployment Assistant
- Separate Management Switch and Connectivity
- Exacli command line management from remote servers
- Cellcli command line management of Storage Servers
- DCLI distributed command line automation tool
- Automatic Service Request and Patch Manager (patchmgr) support for:
  - database servers,
  - storage servers
  - power distribution units, and
  - Cisco RoCE and management switches

Oracle Database Software (available separately):

- For database servers: Oracle Database 11g Release 2 Enterprise Edition, Oracle Database 12c Enterprise Edition Release 1 and 2, Oracle Database 18c Enterprise Edition, Oracle Database 19c Enterprise Edition, and Oracle Database 21c Enterprise Edition. Oracle Database Options such as Oracle Real Application Clusters, Oracle Partitioning, Oracle Multitenant, Oracle Active Data Guard. See the release specific documentation for feature support.
- For storage servers: Oracle Exadata System Software. Licenses are transferable from one system to another, or to a new system

Oracle Software (included):

- For database servers: Oracle Linux 7 Update 9 with the Unbreakable Enterprise Kernel 5. Zero-loss Zero-copy Datagram Protocol (ZDP) RoCEv2 protocol used to communicate between the Exadata Storage Servers and the Oracle Database which is based on the Reliable Datagram Sockets (RDS) OpenFabrics Enterprise Distribution (OFED)

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

ORACLE