WHITEPAPER / SEPTEMBER 12, 2017 / QL835520

Comparing IBM Power Systems and Oracle Exadata Database Machine for Transaction Processing

Sponsored by: IBM

Market Situation

Server architectures continue to evolve to better address the growing size and diversity of data workloads. While analytics, Big Data, and cloud computing topics tend to dominate today's headlines, transaction processing infrastructures continue to be the driving force serving core business operations. At the same time, in-memory software advances and processor innovations have prompted a new generation of computational paradigms and associated platform architectures.

This paper examines the fundamental technology differences between two distinct server platforms in support of online transaction processing (OLTP) workloads: the Oracle Exadata Database Machine, an engineered system that essentially consists of a SPARC or Intel x86-based compute and storage server bundle running Oracle Linux and Oracle Database; and the POWER8-based IBM Power Systems platform, running the AIX operating system and Oracle Database.

Comparisons include estimated three-year costs for use of Oracle Exadata Database Machine and IBM Power Systems with Oracle Database as platforms for facilitating enterprise OLTP applications. Cost estimates draw upon the experiences of organizations that have deployed the Oracle Database in support of enterprise OLTP applications on these platforms.

A broad survey of 181 organizations representing a variety of industries was used to construct 10 installation profiles in manufacturing, healthcare, distribution, agribusiness, finance, engineering and construction, media, and IT services companies. Companies ranged from 50 to 17,500 employees with revenues up to $6 billion.

Costs for use of IBM Power Systems compared to Oracle Exadata Database Machine as the underlying infrastructure for running Oracle Database averaged 40 percent less (Figure 1). Hardware employed for comparisons includes Exadata X6-2, Exadata X6-8, Power S814, Power S824, and/or Power E850 systems. Power Systems storage options include IBM FlashSystem and Storwize families. Configurations on both

Exadata and Power Systems use Oracle Database 12*c* as the foundation for applications such as Oracle E-Business Suite, PeopleSoft, and/or Financial solutions. Cost differences between the two platforms are driven by factors such as server architecture and software licensing policies.

The Exadata Database Machine, designed specifically for Oracle Database, employs a split architecture optimized for analytics workloads with clustered storage and database nodes connected by an InfiniBand network. The Power Systems platform, designed with flexibility and scalability in mind, includes standard server models equipped with industry leading processors that integrate with high performance flash storage to provide database, web, and application serving capabilities. Organizations also benefit from robust reliability, availability, serviceability (RAS), and security features offered by Power Systems with PowerVM and AIX in addition to highly granular workload management and innovative autonomic features.

Oracle Exadata Database Machine Technology

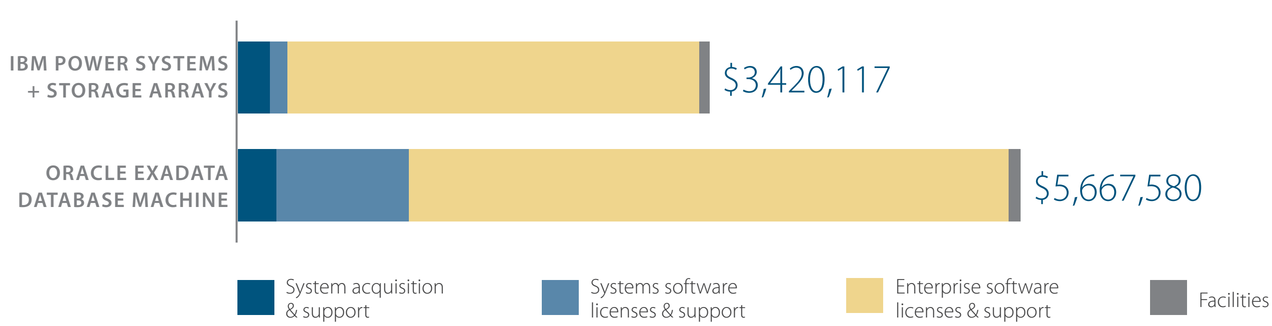
Oracle engineered the Exadata Database Machine using Intel x86 processor and hardware technology modified to address high-volume, scan-intensive, online analytical processing (OLAP) operations. The original SAGE (Storage Appliance for Grid Environments) engineered system from Oracle was launched and co-branded as the HP Oracle Database Machine in 2008. The system was reintroduced in 2009 as the Sun Oracle Database Machine, just prior to Oracle acquiring Sun Microsystems.

In April 2016, the latest generation of x86-based Exadata machines became available, offering the two-socket X6-2 model containing Intel Xeon E5 v4 processors and the eight-socket X6-8 model featuring Intel Xeon E7 v3 processors. A SPARC-based Exadata system, SL6, has been available since February 2017. The two-socket SL6 model contains SPARC M7-based database servers and Intel Xeon E5 v4 storage servers and is offered at a comparable price point to the x86-based models.

Oracle's entry into the hardware sector with its Exadata engineered system involved significant design modification of the conventional server hardware model.

The Exadata Database Machine hardware design splits and distributes the Oracle Database and Real Application Cluster (RAC) architecture between two separate subsystems, each with its own compute

FIGURE 1: Average Three-year IT Costs by Platform for Transactional Processing



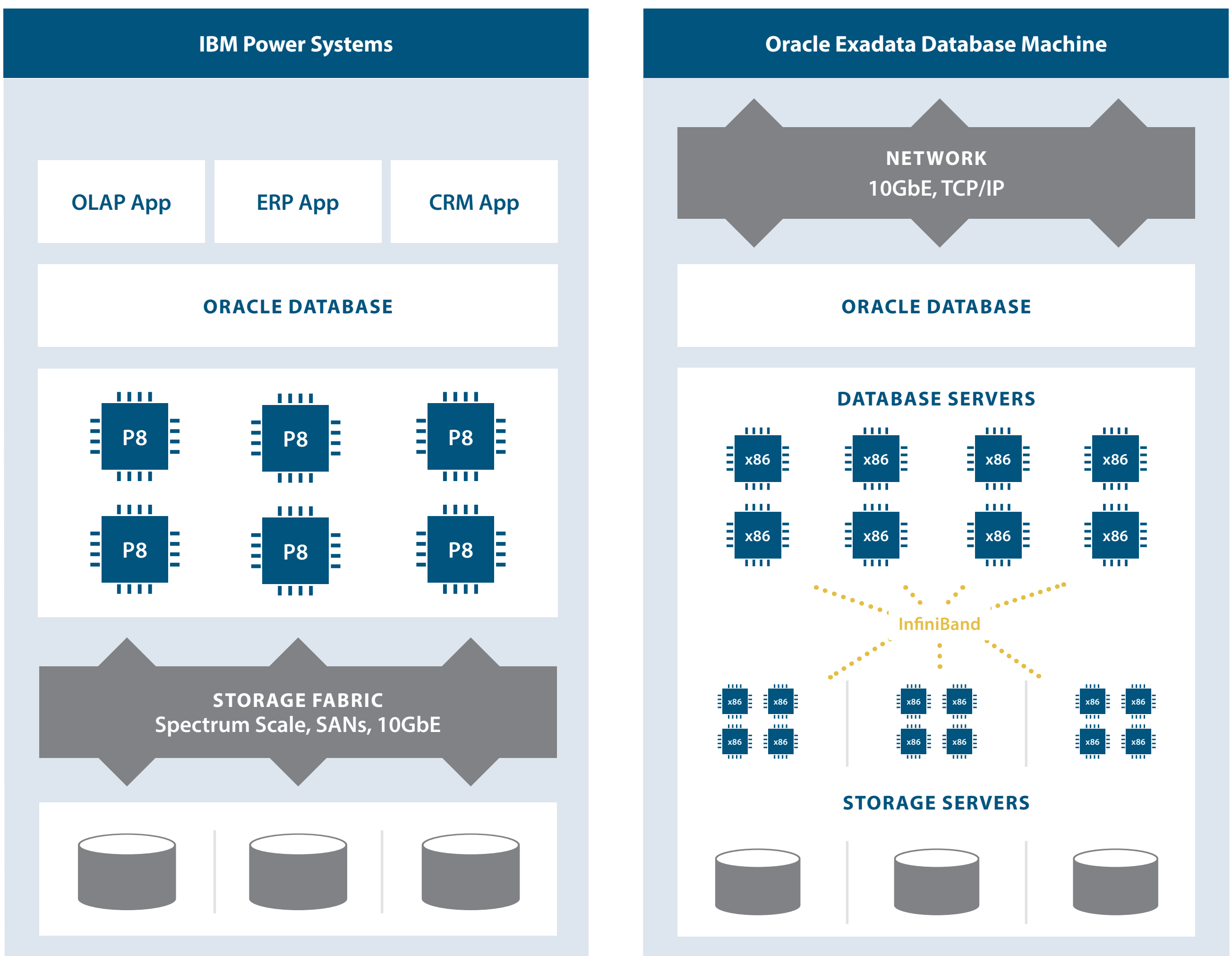
SOURCE: Quark + Lepton (August 2017)

nodes and supporting memory infrastructure. The Storage Server subsystem takes on input/output (I/O) processing while the Oracle Database and RAC implementations reside in the Database Server subsystem. Server software manages the I/O processing communications overhead that connects these subsystems. The InfiniBand fabric moves these loads at high speeds, but the communication complexity required to minimize data volumes being read or written can be easily underestimated.

Figure 2 illustrates the hardware basis for the Oracle Database split and paired-server hosting design unique to Exadata.

Exadata Storage Servers include a collection of specialized technologies created over the years to assist with I/O processing communication challenges. Specifically, Hybrid Columnar Compression (HCC), Smart Scan, Storage Indexes, Smart Flash Cache, and Write Back Flash Cache are storage server-based features that together play a critical role in working to minimize traffic over the InfiniBand fabric. These technologies contribute to OLAP optimization, but Exadata's storage subsystem remains essentially a "black box" otherwise, as only Oracle Database takes advantage of these features. In contrast, Power

FIGURE 2: IBM Power Systems and Oracle Exadata Database Machine Designs



SOURCE: Quark + Lepton (August 2017)

Systems with IBM storage tend to be more flexible, equipped with advanced storage subsystem features that benefit various workloads.

The Exadata Database Machine runs Oracle Linux, which is a full Linux distribution based on Red Hat Enterprise Linux, that has been modified to accommodate specific Oracle software and hardware requirements. Oracle Linux is available free and open source; however, a support contract with Oracle is required for related implementation, performance, and operating issues. Oracle also offers maintenance services for Exadata Database and Storage Servers as well as support for Oracle Database and Exadata Storage Server software.

Exadata has maintained its original bifurcated Database Server software/Storage Server software design from the X2 through the current X6 generation. The hybridization of its data storage format and processing, supporting a unique I/O read/write hand-off from Database Server compute nodes to Storage Server compute nodes, remains at the heart of the system's design.

This specialized hybridization undermines the practicality of hosting large enterprise database-dependent applications on the Exadata platform directly. Oracle does not support applications on Exadata, and organizations are encouraged to purchase Oracle's Exalogic Elastic Cloud or other Oracle solutions at significant additional cost if they wish to deploy new applications. The overall result of this hybrid architecture does not help the processing speeds of OLTP tasks, and may only marginally improve mixed workload performance.

IBM Power Systems Technology

IBM Power Systems originated as a general-purpose UNIX server, and has been refined and enhanced since the introduction of the industry's first multi-core processor, POWER4. Within each new generation of Power Systems, IBM expertise in processor manufacturing, supercomputer and mainframe designs, and operating system and compiler optimization has focused on providing broad support for mixed workloads with industry-leading, high-performance computing speeds and throughput. Power Systems architecture and flexible component-level interconnectivity is illustrated in Figure 2.

The POWER8 processor has been designed to handle a variety of high-performance computing workloads such as artificial intelligence (AI), in-memory OLTP, and analytics. POWER8 processors, for example, support larger numbers of threads per core (by a factor of four), greater memory bandwidth, a higher clock speed, and more cache (compared to the E7-8895 v3 processor and the E5-2699 v4 processor used in the X6-8 and X6-2 models, respectively). In the POWER8 design, L1 through L3 cache are on-chip, while up to 128 megabytes (MB) of L4 cache is added off-chip. Intel offers no equivalent to POWER8's L4 cache.

Dedication to the improvement of Power Systems OLTP workload management and performance has resulted in the close integration of PowerVM with AIX. In a mixed workload environment, the amount of work a server can perform over time depends not only on processor, memory, and I/O throughput, but also on the mechanisms that allocate and reallocate system resources as demands change.

The evolution of Power Systems capabilities has been continuously aligned, through close collaboration, with more than a quarter century of Oracle Database updates and major releases. Power Systems, equipped with the tightly integrated AIX operating system and PowerVM, can simultaneously and efficiently host databases and applications in development, test, and/or production environments, with robust and protected hardware isolation.

Design Differences

The different design strategies implemented by Oracle and IBM in these two Oracle Database server platforms are explored below within the context of performance, virtualization, storage, security, and RAS. These differences are summarized in Table 1.

PERFORMANCE

When evaluating design differences between Exadata and Power Systems from a database-hosting performance perspective, it is important to focus on the distinct characteristics of OLAP and OLTP data processing. *OLAP workloads* often involve complex queries targeting a large data set, requiring heavy read I/O and aggregation of data. *OLTP workloads* are characterized by a large volume of frequent, much smaller I/O actions. Performance for OLAP workloads tend to focus on improving throughput, whereas fast response times are desirable for OLTP workloads.

Power Systems are flexible and integrate easily into an existing IT environment. Designed with performance-boosting database, hardware, and storage innovations, Power Systems are intended for use as a single hosting platform for a broad mix of enterprise software. To address the specific needs of OLTP workloads, POWER8 processors provide up to 12 cores per socket with clock speeds of over 4 GHz.

For analytics workloads, the increased number of POWER8 cores supported by higher I/O bandwidths provide industry-leading performance levels for OLAP workloads as well. POWER8-based Power Systems servers can deliver I/O bandwidth of up to 192 gigabytes per second (GB/s).

Additional POWER8 features that reinforce mixed workload capabilities include the following:

- **Intelligent Threading** allows workloads to be executed using from one to eight threads per core. The system can automatically determine how many threads per core to use for optimum performance, or system administrators may select the number of threads to be employed.
- **Intelligent Cache** allows systems to dynamically vary cache utilization to minimize overall cache latency as workload characteristics change. The system may automatically determine the appropriate level of cache for specific workloads.

In contrast, each Exadata rack combines compute servers and storage servers over an InfiniBand network. This unconventional split architecture aims to minimize data transfer through I/O processing done by storage servers using Smart Scan software.

TABLE 1: Principal Differences Between Oracle Exadata Database Machine and IBM Power Systems with Storage Arrays

	ORACLE EXADATA DATABASE MACHINE	IBM POWER SYSTEMS & STORAGE ARRAYS
Design	Specialized hardware & software appliance designed for high-volume, scan-intensive analytical workloads requiring high levels of read I/O performance & throughput.	General-purpose UNIX server & storage array designs configurable & optimized for simultaneously hosting multiple types of workloads & applications. Industry-leading benchmarks for OLTP & OLAP performance.
Virtualization & workload sharing	Oracle VM (OVM), based on Xen, is a software-based virtualization with limited isolation & workload prioritization capabilities, increasing overhead. Licensing forbids live migration of VMs to other OVM servers. Requires additional servers for applications.	Advanced virtualization server architecture supports multiple hardware- & software-based partition types. PowerVM with Live Partition Mobility avoids downtime for planned maintenance. VMs allow databases & applications to be hosted on the same physical servers.
Configurability	Capacity on Demand available for database server cores, but not storage server cores. Must use Exadata Storage Servers. Supports Oracle 11*g* R2 or 12*c* on Oracle Linux.	Server & IBM or 3rd-party storage arrays separately configurable with high levels of granularity. Various operating systems & system software available. Supports older Oracle Database versions with earlier AIX releases.
Availability optimization	Standard Intel hardware reliability, availability & serviceability (RAS) features present. Double or triple disk mirroring is recommended, but increases costs. Supports Oracle Database, Data Guard, Active Data Guard, & Maximum Availability Architecture. Potential single points of failure for some configurations due to heavy reliance on InfiniBand.	Advanced hardware RAS & operating system-level availability optimization features enabled. Servers & storage arrays support multiple RAID levels. Supports same Oracle availability solutions as Exadata. IBM & third-party solutions for high availability (HA) clustering, failover, & recovery enabled.
Backup & restore procedures	DBAs must use Oracle Recovery Manager (RMAN) software running on database server nodes to do backups. Backups, disaster recovery, & other storage management very limited & typically require more specialized infrastructure than that commonly found in existing data centers. Backups may also require external Oracle ZFS Storage Appliances, especially over larger geographic areas, increasing costs.	Supports storage-level database backup & extended SAN features, including snapshots (FlashCopy), remote mirroring, storage tiering (e.g., Easy Tier), LAN-free backups, & thin provisioning. Power Systems easily integrated into existing Exadata environment.
Security	Uses Intel Advanced Encryption Standard (AES) New Instruction (NI) cryptographic acceleration. Oracle Advanced Security leverages Transparent Data Encryption & Data Redaction policies to meet FIPS Level 2 security.	Leverages PowerVM bare metal isolation & KMIP compliant self-encrypting storage. Can integrate with IBM PCIe Cryptographic Coprocessors to achieve the highest cryptographic security available commercially, FIPS Level 4.

SOURCE: Quark + Lepton (August 2017)

By adding compute power to these customized storage servers, Oracle has employed unique hardware constraints to deliver an approach to increasing performance in data warehouse and analytical workloads. I/O processing by the Smart Scan maximizes efficiency when processing large data sets. However, this hybrid structure does not deliver these performance gains when processing OLTP workloads. OLTP operations do not involve reading and writing large tables or aggregates. OLTP queries typically involve INSERT, UPDATE, and DELETE transactions for individual records. As a result, storage cores remain largely dormant when facilitating transaction-intensive database activities.

Exadata X6-2 database servers utilize two 22-core Intel Xeon E5 v4 processors, up to a total of eight two-socket servers with 352 cores for compute nodes in a full rack. Exadata X6-8 racks contain up to four eight-socket servers with 18-core Intel Xeon E7 v3 chips, up to a total 576 cores per rack.

Exadata X6-2 racks also contain up to 14 two-socket storage servers with 10-core Intel Xeon E5 v4 chips, up to a total of 280 cores. For transaction processing, the compute capabilities of up to 280 cores per rack are dramatically underutilized. Additionally, the processors in both X6-2 and X6-8 database servers support up to two threads per core. In contrast, POWER8 cores support up to eight threads per core, boosting the number of instructions that can be executed concurrently. In support of the multithreading, it should be noted that the sharing of cache among threads can affect system memory performance. POWER8 processors contain a larger caches than the processors used in Exadata database servers. A larger cache contributes to better performance, as multithreading can increase the potential for cache misses. There is evidence suggesting that a three times performance decrease may occur when the working set exceeds the L3 cache size. For mixed workloads, a larger cache is beneficial for OLTP tasks.

Table 2 provides detailed specifications for Power Systems' POWER8 processors as well as Intel x86 processors used in Oracle Exadata Database Servers.

Since 2009, Oracle has placed a strong marketing emphasis on another performance feature, Exadata Smart Flash Cache (ESFC). However, performance gains achieved from using ESFC vary depending on workload and configuration.

Adding data to ESFC is typically faster than writing to disk, but it is significantly slower than latency associated with writing to memory. Exadata Storage Server, by default, stores only small I/Os in ESFC. Data requested by table scans will not be stored in ESFC unless manually configured to do so. ESFC is also a tertiary cache, used only when requested data cannot be found in the local buffer cache or the RAC global cache. Efficient local buffer caches or global cache will reduce any benefits gained by ESFC. Typical workloads, on all but exceptionally large enterprise databases, do not realize significant improvement in performance from the use of ESFC.

Oracle's primary performance focus on data warehousing and analytics is also evident in the implementation of HCC. This technology utilizes both row and columnar formats for data storage. Columnar compression techniques can result in higher storage savings and improved tables-scan performance. Using HCC, column values for an appropriate set of rows are compressed and stored in compression units. This compression function synergizes with Smart Scan to improve query performance involving large sequential data sets. HCC's benefits are maximized for historic data that does not require updating. Making changes to HCC data requires intensive compressing and decompressing sequences and may significantly increase overhead.

Oracle's Advanced Compression, a feature of Oracle Database, can be leveraged by both Power Systems and Exadata. Differences in database operations in Power Systems and Exadata for both OLAP and OLTP workloads are shown in Figures 3 and 4.

VIRTUALIZATION

Virtualization and consolidation techniques have allowed businesses to increase server efficiency and utilization. Although both Oracle VM (OVM) and IBM PowerVM are promoted as server consolidation solutions, the strategies used by each virtualization technology are fundamentally different. When consolidating workloads on Exadata, additional servers are required for application and Web serving. In contrast, Power Systems with PowerVM allow for the consolidation of both databases and applications on the same platform.

In such areas as virtualization, multitasking, and workload management, POWER8-based systems running AIX are significantly differentiated from x86-based Exadata equivalents running Oracle Linux. For example, POWER8-based systems running AIX can support higher workload densities, over a greater number of partitions, and with greater efficiency, than can the Exadata system. PowerVM bare metal architecture also supports various operating systems, such as popular enterprise Linux distributions, providing unmatched flexibility to meet the needs of diverse workloads.

TABLE 2: Processor Specifications for Latest Generation Power Systems and Oracle Exadata Database Machine

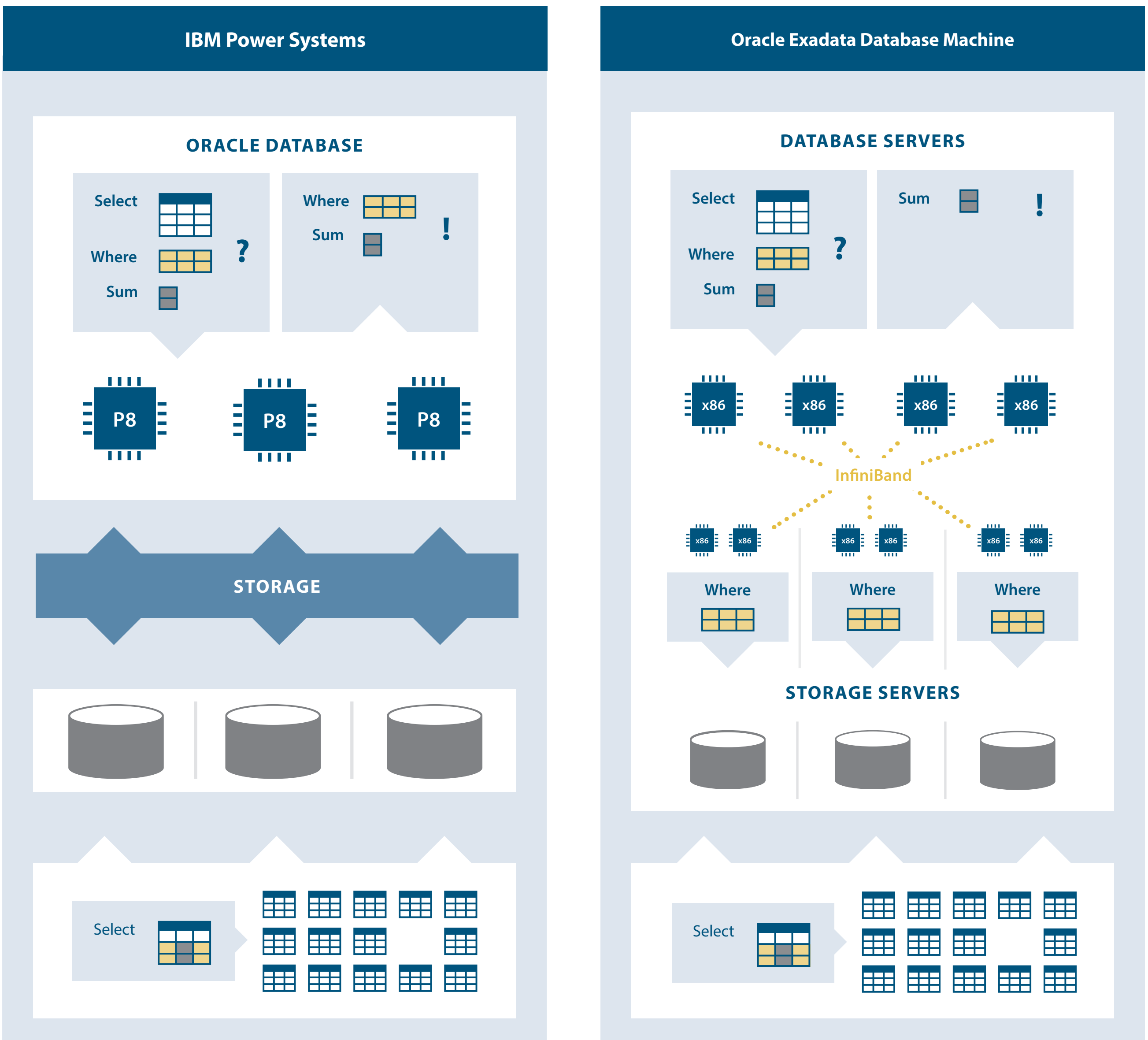
FEATURES	POWER SYSTEMS	EXADATA X6-8	EXADATA X6-2
Processor	*POWER8*	*Intel Xeon E7-8895 v3*	*Intel Xeon E5-2699 v4*
Cores/chip	12	18	22
Max threads/core	8	2	2
Total threads	96	36	44
Max clock rate*	4.2 GHz	3.5 GHz	3.6 GHz
Max L-3 cache	96 MB	45 MB	55 MB
Max L-4 cache	128 MB	N/A	N/A
Max memory bandwidth	230 GB/s	102 GB/s	76.8 GB/s
Max sockets	16	8	2
Process technology	22 nm	22 nm	14 nm

*Refers to max turbo frequency for Intel Xeon CPUs
SOURCE: Quark + Lepton (August 2017), Intel ARK, IBM POWER8 Specifications

For virtualization, Exadata employs OVM, based on open source Xen hypervisor technology. On Exadata systems, allocation of system resources is primarily handled by the Resource Manager, a standard feature of Oracle Database 11*g* and 12*c*, along with the Exadata-specific I/O Resource Manager (IORM), which manages access to I/O resources provided by storage servers based on user-defined application priorities and service-level targets.

Because OVM is licensed on a per-CPU basis, CPU pinning is sometimes required to bind virtual CPUs to a physical one. CPU pinning is not controlled or accessed in the Database Resource Manager or the IORM management tool and requires manual modification of virtual machine (VM) configuration files.

FIGURE 3: Oracle Database OLAP Workloads on IBM Power Systems and Oracle Exadata Database Machine

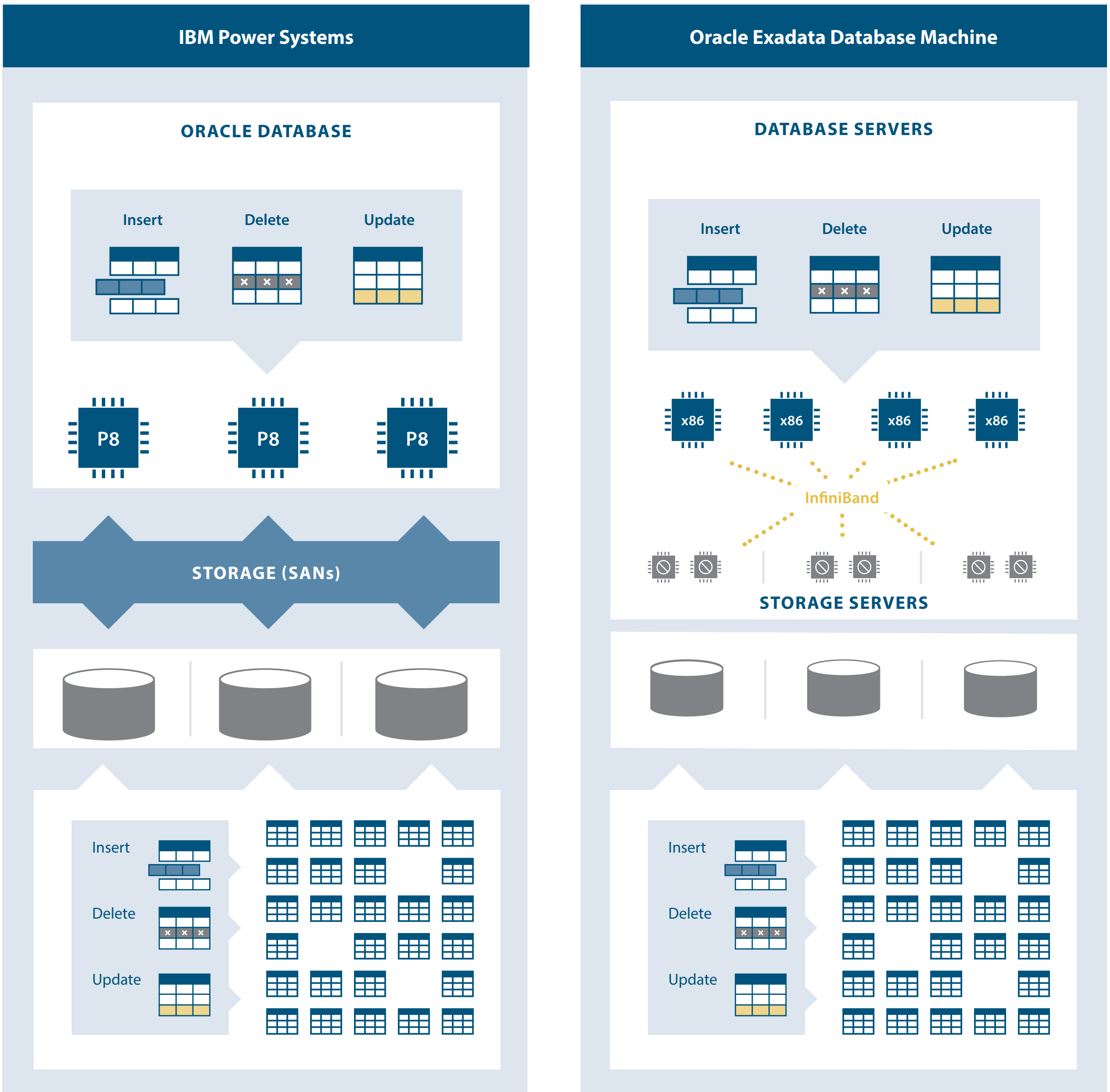


SOURCE: Quark + Lepton (August 2017)

Increasing VM resource utilization efficiency requires manually modifying hypervisor rules that assign each VM's virtual CPUs near its memory to minimize non-uniform memory access (NUMA) latency.

Inefficient partitioning creates the potential for lower levels of capacity utilization. However, the extent to which this occurs in practice depends on mechanisms that allocate system resources between, and monitor and control workload execution processes across partitions. If these mechanisms are ineffective, a high proportion of system capacity may be idle over time.

FIGURE 4: Oracle Database OLTP Workloads on IBM Power Systems and Oracle Exadata Database Machine



SOURCE: Quark + Lepton (August 2017)

Administrators will, moreover, tend to operate systems well below capacity to avoid risks of bottlenecks. These effects explain why use of x86 hypervisors falls short of their architectural potential in many organizations.

According to Oracle's own disclaimers, Exadata VMs are not recommended for virtualization of heavyweight applications, such as its own E-business Suite.

In contrast, PowerVM's unique firmware based hardware virtualization minimizes overhead. PowerVM allows for more processing power, a much higher level of security and overall enhanced performance. PowerVM's industry-leading virtualization capabilities include I/O isolation through Virtual I/O Servers (VIOS), a feature not offered by Oracle VM or other virtualization technologies. PowerVM bare metal implementation is ideal when organizations are restricted by compliance measures, typically requiring strict data security and privacy controls.

Through firmware-implemented isolation, bare metal servers give customers complete control of a VM's physical resources. Power Systems with PowerVM support up to 1,000 VMs per server using logical partitions (LPARs) for dividing system resources. Dedicated LPARs use dedicated static system resources and can be installed and operated separately. Dynamic LPARs (DLPARs) allow administrators to add and delete resources such as processors, memory, and I/O components from live partitions. In addition, the Live Partition Mobility feature of PowerVM also allows for live migration of LPARs from one system to another for organizations that meet the licensing requirements for this feature.

To achieve even better workload balance, IBM's Micro-Partitioning, originally a mainframe innovation, allows for the partitioning of processing resources up to a granularity of 1/100th of a core (with a minimum of 1/20th of a core per partition).

STORAGE

Power Systems support a wide selection of storage solutions, including disk and flash arrays or hybrid storage deployments, from IBM as well as other vendors. External storage capacity may be dynamically allocated and reallocated between systems. IBM SAN Volume Controller (SVC) controlled storage, such as IBM Storwize or IBM XIV systems, can be deployed with IBM Real-Time Compression to gain up to 80 percent in storage capacity savings. Thin provisioning can be used with SVC to optimize storage efficiency by flexibly allocating available storage capacity between users and applications.

For high performance storage, IBM introduced the FlashSystem 840 with new RAS features in 2014. This was the first release of reengineered technology IBM obtained in its acquisition of Texas Memory Systems (TMS) in 2012. The IBM FlashSystem 900, introduced in February 2015, is designed for Tier 0 high-end performance storage. The FlashSystem 900 can be configured to deliver RAID 5 redundant disk array striping with up to 57 terabytes (TBs) per array, ensuring its proven ability to deliver 99.999 percent availability.

IBM FlashSystem family is a popular flash storage solution for Power System that harnesses the high speed and performance of IBM MicroLatency Modules. Using IBM FlashCore, a proprietary, hardware-only controller, the FlashSystem family reduces the overhead generated by software data management produced by other types of flash storage.

The FlashSystem 900 is an easy-to-deploy entry model designed to accelerate business-critical applications. The high capacity FlashSystem V9000 model, suitable for large enterprises, is capable of petabyte scale storage in a single array. The latest releases in the FlashSystem line, the A9000 and its more scalable brother, A9000R, are cloud-optimized offerings that can be easily integrated with most existing infrastructures.

For exceptionally I/O-intensive workloads, Power Systems also support most new types of all-flash arrays being offered by a growing number of vendors, including IBM itself.

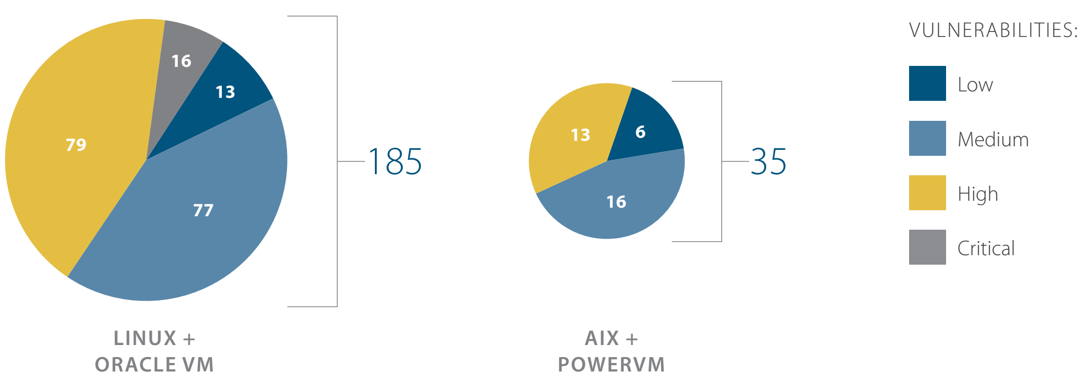
Oracle's promotion of its latest X6 Extreme Flash (EF) models highlights the increase in I/O performance as the result of using all flash drives in storage servers. Exadata also offers High Capacity (HC) models featuring hybrid storage servers. While Exadata flash storage does show an improvement in input/output operations per second (IOPS) rates over HDD devices, it remains comparable to conventional flash storage in access latency.

Although Exadata touts impressive claims of I/O throughput and scalability, they only apply to large full-rack systems without mirroring. IOPS rates for smaller Exadata systems, such as a quarter rack, are only a fraction of the rates published for a full rack. The throughput rates published by Oracle also do not take into account the effects of mirroring. When using Oracle's recommended ASM high-redundancy mirroring, effective I/O throughput drops to one third of published rates. Typical workloads will not take advantage of the high throughput Oracle claims Exadata is capable of.

SECURITY

Contributing to the reliability of Power Systems are the robust security features included in the hardware and AIX operating system. From January 2014 through June 2017, IBM AIX received 35 security advisories, as reported by the National Vulnerability Database from the National Institute of Standards and Technology (NIST). Power VM has never received a security advisory.

FIGURE 5: Comparative Vulnerability Data: January 2014 – June 2017



SOURCE: Quark + Lepton (August 2017), NIST Computer Security Division, National Vulnerability Database, CVSS Metrics Version 3

In contrast, Oracle Linux and Oracle VM received 157 and 28 vulnerability advisories respectively in the same time period. For detailed vulnerability statistics comparing AIX with PowerVM and Oracle Linux with Oracle VM, refer to Figure 5.

The AIX operating system provides security using an Encrypted File System (EFS) capability, which encrypts and decrypts all files in the file system transparently for non-RAC environments. Once enabled, EFS utilizes user credentials to effect encryption protection for all user-associated files. IBM storage solutions, such as the DS8000, FlashSystem, or XIV families, enable data-at-rest encryption through use of embedded data encryption engines. The drive-level encryption capability protects storage media directly, and reduces risk without adding overhead to database servers.

PowerVM logical partitions add an additional layer of protection through separation and isolation of system resources in virtualized environments. PowerSC can also be deployed to enable security and compliance automation in virtual environments.

In addition to the security features provided by the AIX operating system, IBM POWER8 processors are capable of leveraging hardware acceleration for high-performance, in-core cryptography using the Advanced Encryption Standard instruction set. IBM offers additional security and compliance capabilities through the PowerSC solution for AIX or Linux environments running on Power Systems with PowerVM.

For Exadata, Oracle uses a combination of isolation and access control policies to secure the system. OVM offers software based virtualization, but software partitioned environments are less isolated than hardware based virtualization, offering less protection. Fine grain database isolation can be implemented using Oracle Database Vault, Oracle Virtual Private Database, and Oracle Label Security solutions. Oracle Advanced Security, certified Level 2 through FIPS 140-2, data encryption capabilities and isolation policies can also be deployed.

RELIABILITY, AVAILABILITY, AND SERVICEABILITY

POWER8-based Power Systems benefit from hardware- and microcode-based RAS features derived from earlier models as well as from IBM mainframes. These RAS features include unique capabilities not found in Intel-based designs. Power Systems exploit the greater speed, expanded memory management, and processor allocation capabilities of the POWER8 processor.

The close integration of Power Systems, PowerVM, and AIX provide users with the means to manipulate a wider range of RAS variables, including subsystems, threads, processors, cache, main memory and I/O, multiple types of partitioning, multiple threads, and dedicated and pooled processors, and to do so with higher levels of granularity and flexibility than can be managed with the Exadata platform.

Power Systems may realize significantly higher levels of capacity utilization over time than less well-optimized platforms. Experiences with enterprise resource planning (ERP) systems, for example, have shown that, over periods of months to years, Power Systems may execute workload volumes up to 40 percent larger than indicated by point-in-time performance measurements. Such measurements have been used to compare OLTP capacity across competitive systems, including the Exadata system.

Another feature, Active Memory Expansion (AME), enables system-managed compression and decompression of data in memory. This in-memory data compression is transparent to applications and users. AME is achieved through the compression of in-memory data, increasing the amount of data that can be placed into memory. Compression rates of up to 50 percent are supported; i.e., allowing the effective maximum memory capacity to be up to 100 percent larger than true physical memory.

According ITIC's 2016 Global Server Hardware, Server OS Reliability Report, 61 percent of IBM Power Systems had 99.999 percent uptime, which equates to 5.26 minutes of downtime per server/per year. Only 40 percent of Oracle hardware systems met the "five nines" percent uptime standard.

The same study found that Power Systems server administrators spend less time rebooting operating systems, including planned reboots needed to add or reconfigure system resources. In fact, 77 percent of IBM AIX administrators responded that they rarely or never reboot the server.

RAS features included in IBM POWER8-based servers running AIX are shown in Table 3.

Cost Comparison

This detailed cost section presents total cost of ownership estimates derived from survey results, and summarizes key differentiators responsible for driving cost variations between these two database platforms. As the Exadata X6 series is relatively new, many users still employ older Exadata models. Configurations for Exadata and Power Systems in this report used the latest-generation models and were selected based on estimated comparative OLTP performance.

Overall results from cost comparisons are summarized in Figure 1. Even allowing for aggressive Oracle discounting and use of Exadata Capacity on Demand (CoD), costs for use of Power Systems for all solutions averaged 40 percent less than for Exadata systems.

Costs for both systems include system acquisition and maintenance, licenses, and support for systems and Oracle enterprise software, and facilities costs. Systems software includes Oracle Database 12*c*, operating systems, clustering, virtualization, diagnostics, and tuning.

Key drivers in the higher costs for using Exadata are system software and enterprise software licensing (Table 4). System software license and support costs average 87 percent less for use of Power Systems compared to Oracle Exadata, and enterprise software license and support costs average 31 percent less for use of Power Systems compared to Oracle Exadata.

Comparisons are between: latest-generation Exadata X6-2 eighth-, quarter-, and half-rack models with Oracle Linux, Oracle VM, and Exadata EF or HC storage servers; and IBM POWER8-based Power Systems S814, S824, and E850 server models with the AIX operating system, PowerVM virtualization, and IBM FlashSystem or Storwize storage.

Comparisons are for production systems only. Most organizations also employed separate systems for test, development, quality assurance, other non-production functions, and for backup. All calculations

TABLE 3: Examples of POWER8 RAS Technologies

RAS FEATURE	DESCRIPTION
Chipkill	Technology capable of detecting & correcting single-bit as well as 2-, 3- & 4-bit errors in memory devices, including cache & memory interfaces. Employs RAID-like striping of data across memory to provide redundancy & enable reinstatement of original data in the event of a failure. Significantly more reliable than conventional error correction code (ECC) technology. Can survive double memory failures in 8-DRAM groups.
First Failure Data Capture (FFDC)	Employs 1,000+ embedded sensors that identify errors in any system component. Root causes of errors are determined without the need to recreate problems or run tracing or diagnostics programs.
Dynamic Processor Deallocation	Detects error patterns indicating potential processor failures, & stops & deallocates untrustworthy processors.
Dynamic Processor Sparing	Allows alternates to be automatically activated as replacements for failed processors.
Code Error Recovery: Processor Instruction Retry Alternate Processor Recovery Partition Core Contained Checkstop	If an instruction fails to execute, the system automatically retries the operation. If the failure persists, the operation is repeated on a different processor. If this does not succeed, the failed processor is taken out of service (checkstopped). Only LPARs supported by the failed processor are affected.
Partition Availability Priority	In the event of a processor failure, reallocates LPAR-based workloads based on assigned priorities.
Hot-Swappable Disk/SSD Bays	Drive replacement can be performed without interrupting data writing on adjacent drives.
Processor Instruction Retry	PowerVM can work with the processor hardware to migrate a workload running on a failing processor to a spare or alternate processor.
Redundant Hot-Plug Power/Cooling	Redundant components allow upgrading or replacement without requiring the shutting down of the host hardware.
Node Add, Node Repair, Memory Upgrade	Hardware can be added, repaired, or upgraded while the system is powered on.
Dual VIOS	Redundant VIOS partitions, each consisting of a virtual Ethernet adapter, either of which can be used as the control channel between the two adapters, should one fail.

SOURCE: Quark + Lepton (August 2017)

except those for energy consumption are based on industry standard discounted prices as reported by users. Energy costs are based on system specifications and national averages.

Personnel costs for database administrators (DBAs) and application specialists were similar, and are not included. There was also no substantive difference in full-time equivalent (FTE) staffing for server and storage administration tasks. However, Exadata DBAs tend to require additional skills that DBAs working with less specialized systems do not need.

COST VARIATIONS

Although results were generally consistent for all types of solutions, there were a number of cost variations between system sizes and applications.

As Figure 6 depicts, three-year costs for use of Power Systems averaged 54 percent less for installations with fewer than 250 users, 59 percent less for 250 to 499 users, 71 percent less for 500 to 999 users, and 34 percent less for 1,000+ users.

Sample configurations were based on an aggregation of customer survey results and publicly available company information. Systems sizing was based on estimated number of active users and company workloads. Licensing costs were based on per core, per processor or fixed fee licensing metrics. Enterprise software licensing was based on industry standards for company workloads.

Three-year Power Systems costs for deployment of Oracle E-Business Suite averaged 37 percent less than for use of Exadata, while Oracle PeopleSoft, Financial and Retail solutions averaged 51 percent, 53 percent, and 42 percent less respectively.

PACKAGING AND DEPLOYMENT

Cost differences also reflect the packaging and deployment options for Oracle Database on Oracle Exadata and Power Systems.

Oracle Exadata X6-2 eighth-, quarter-, half-, and full-rack models are configured with 44, 88, 176, 352 Intel database server cores, respectively. Exadata X6-8 half-and full-rack models contain 288 and 576

TABLE 4: Average Three-year Costs for Transaction Processing Systems

COST CATEGORY	IBM POWER SYSTEMS + DISK ARRAYS	ORACLE EXADATA DATABASE MACHINE
System acquisition & support	233,733	274,791
Systems software licenses & support	125,436	969,097
Enterprise software licenses & support	3,039,282	4,379,213
Facilities	21,666	44,479
THREE-YEAR TOTAL ($)	**3,409,284**	**5,642,085**

SOURCE: Quark + Lepton (August 2017)

database cores, respectively. CoD can be used to deactivate certain numbers of Exadata database cores as a means of reducing total cost of ownership (TCO). However, CoD cannot be used with Oracle VM.

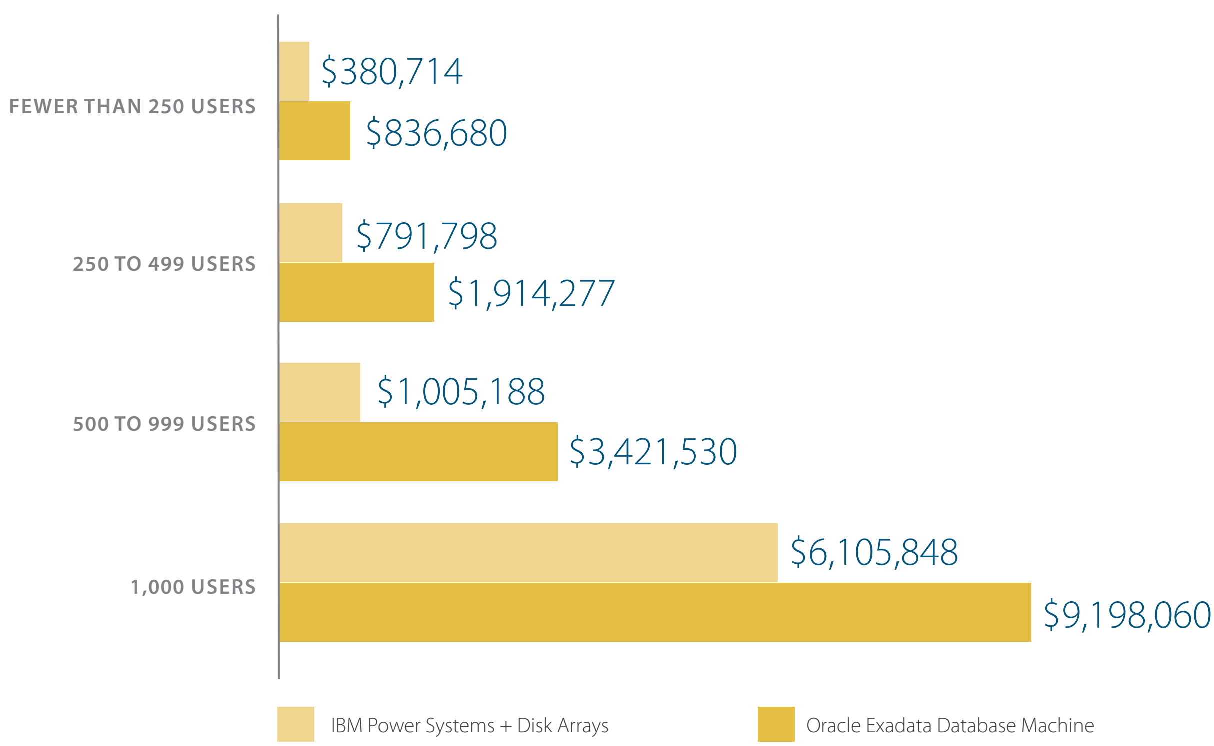
The lack of CoD for Exadata storage servers, coupled with mirroring recommendations, exacerbates factors increasing licensing costs. Oracle recommends the triple-mirroring high redundancy setting of Oracle Automatic Storage Management (ASM) to ensure optimal high availability. If Data Guard is enabled, ASM normal redundancy (double mirroring) can be used. When using either double or triple mirroring, organizations are required to license the full physical capacity of their Exadata storage. This requirement significantly drives up costs based on the usable mirrored capacity, although the used storage capacity is effectively a fraction of the total storage capacity.

Storage software licensing prices for Exadata storage servers are listed as $10,000 per disk and $20,000 per flash drive. A full X6 rack can contain up to 14 storage servers with either 8 NVMe flash drives (EF) or 12 disks and 4 flash drives. Storage licensing costs, before discounts, for a full rack can reach $2.24 to $2.8 million for a single Exadata Database Machine.

In contrast, Power Systems servers offer more granular configurations of 4-, 6-, 8-, 12-, 16-, 24-, 32-, 48-, and 192-cores. Besides database workloads, Power Systems servers can be application-serving as well.

Power Systems compatibility with a wide range of storage solutions also affects pricing. Storage hardware and software can be supplied by IBM or other vendors and tailored to customers' needs. An organization's existing storage solutions may also be used.

FIGURE 6: Average Three-Year Costs for Use of IBM Power Systems and Oracle Exadata Database Machine by Numbers of Active Users



SOURCE: Quark + Lepton (August 2017)

Higher Exadata X6 core counts, compared to Power Systems, as well as previous generation Exadata systems, results in higher system and enterprise software licensing costs. Power Systems can also support additional operating systems, such as enterprise Linux distributions from Red Hat, SUSE, and Canonical as well as system management and tuning tools from IBM.

Conclusions

The commodity hardware, hybridized software, and customized infrastructure of the Oracle Exadata Database Machine has been shown to negatively impact data center TCO costs, integration considerations, and performance cost-effectiveness when it is used principally to run OLTP workloads. Three year costs for deployment of Oracle Database averaged 40 percent less on IBM Power Systems than on Oracle Exadata.

As the Exadata Database Machine architecture continues to require specialized skills for its successful deployment, administration, and maintenance, the costs and risks associated with its uniqueness must be carefully assessed. In terms of capability, adaptability, and future impacts of selecting this platform—specifically for OLTP-intensive operations—serious implications apply.

Architecture and technology differences in database design, virtualization, and storage contribute to cost and performance variances between the two platforms. Although Exadata touts impressive I/O throughput performance, the values are effectively only a fraction of the advertised rates once system redundancy is considered. Storage server processing, although effective for OLAP workloads, does not benefit OLTP workloads.

Strict licensing schemes for Oracle Exadata coupled with high processor counts for database and storage servers drive up costs as well. Pre-integrated storage, operating system, and engineered system software also limit customers' flexibility and ability to tailor their IT infrastructure to their needs.

The wide support IBM Power Systems offer for various operating systems, middleware, applications, and storage solutions provide organizations with great flexibility, while maximizing the compatibility of existing IT investments with future ones. IBM AIX and Power Systems' long history of robust RAS and security features give users the confidence that their investments will deliver protected value, regardless of possible adjustments that may be necessary in optimizing their operational plans.

Index

LIST OF FIGURES

LIST OF TABLES

LIST OF REFERENCES

Oracle Advanced Compression with Oracle Database 12*c* Release 2 found at www.oracle.com/technetwork/database/options/compression/advanced-compression-wp-12c-1896128.pdf

Oracle Exadata and OVM Best Practice Overview found at www.oracle.com/technetwork/database/availability/exadata-ovm-2795225.pdf

Intel Xeon Processor Specifications found at ark.intel.com/

National Institute of Standards and Technology found at nvd.nist.gov/

ITIC, 2016 Global Server Hardware, Server OS Reliability Report found at www-01.ibm.com/common/ssi/cgi-bin/ssialias?htmlfid=ZSL03380USEN

Quark + Lepton™

CORPORATE OFFICE
Boulder, Colorado USA

www.quarkandlepton.com
info@quarkandlepton.com

IBM sponsored this publication, however, the information and conclusions contained in this publication do not necessarily represent the positions of IBM or other referenced sources.