Oracle Exadata: What's New And What's Coming

Gurmeet Goindi

Exadata Product Management

ORACLE

Safe Harbor Statement

The following is intended to outline our general product direction. It is intended for information purposes only, and may not be incorporated into any contract. It is not a commitment to deliver any material, code, or functionality, and should not be relied upon in making purchasing decisions. The development, release, and timing of any features or functionality described for Oracle's products remains at the sole discretion of Oracle.

Exadata Vision

Dramatically Better Platform for All Database Workloads

- Ideal Database Hardware - Scale-out, database optimized compute, networking, and storage for fastest performance and lowest costs
- Smart System Software – specialized algorithms vastly improve all aspects of database processing: **OLTP, Analytics, Consolidation**
- Full-Stack Integration – Database-to-disk optimization, automation, testing, patching, and support to reduce operational costs

Identical On-Premises and Oracle Public Cloud

Exadata Cloud Service

Proven at Thousands of Critical Deployments since 2008

Half OLTP - Half Analytics - Many Mixed

- **Petabyte Warehouses**
- **Online Financial Trading**
- **Business Applications**
  - SAP, Oracle, Siebel, PSFT, …
- **Massive DB Consolidation**
- **Public SaaS Clouds**
  - Oracle Fusion Apps, Sales*f*orce, SAS, …

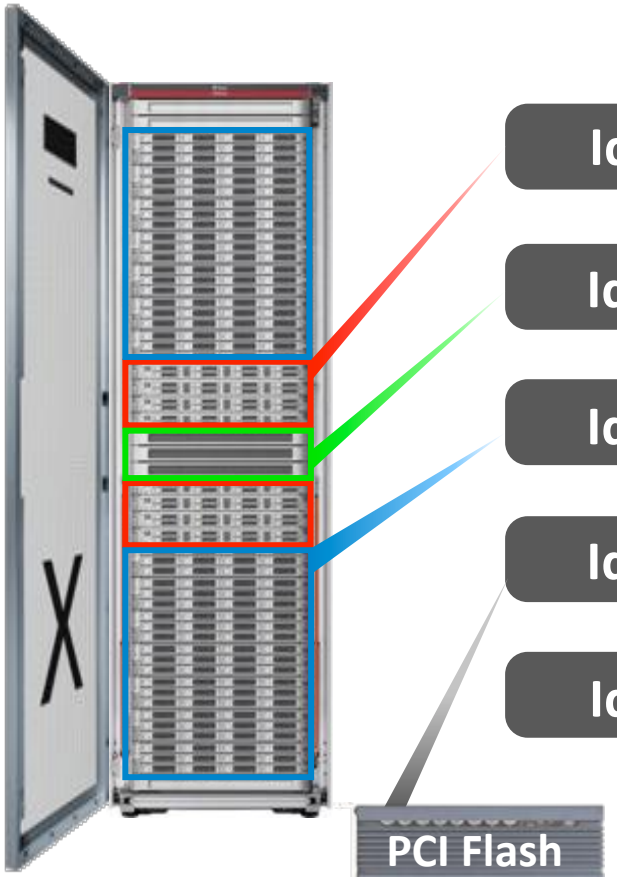
4 OF THE TOP 5

BANKS, TELCOS, RETAILERS RUN EXADATA

- Ideal Database Hardware
- Smart System Software
- Full-Stack Integration

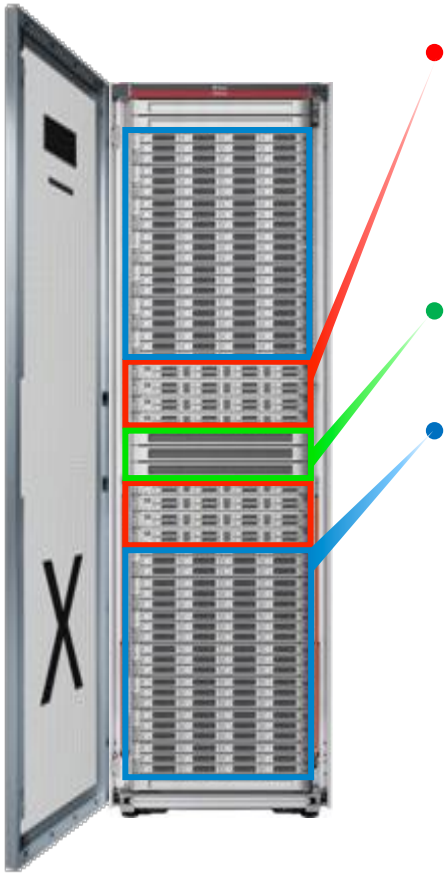
Ideal Hardware Architecture for Database

Most Advanced - Highest Performance - Always Available - Starts Small, Scales Huge



Ideal Server Architecture	Scale-Out with Fastest CPUs
Ideal Network Architecture	Unified Ultra-fast InfiniBand
Ideal Storage Architecture	Scale-Out Servers + Database Offload
Ideal Flash Architecture	Ultra-fast NVMe PCIe Flash
Ideal Capacity Architecture	Tier PCIe Flash & Huge 8TB SAS3 Disks

Exadata X6-2 Hardware Details (new in red)



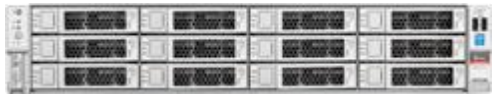
- **Scale-Out 2-Socket Database Servers**
  - Fastest Intel 22-core Broadwell: E5-2699 v4, 25% faster
  - DDR4 DRAM frequency increased 13%
- **Ultra-Fast Unified InfiniBand Internal Fabric**
- **Scale-Out Intelligent 2-Socket Storage Servers**
  - Intel 10-core Broadwell CPUs (25% faster)
    - Offload database processing
  - 8TB Helium Disk Drives (2X Larger since Oct)
  - 2X larger and 2X faster 3D V-NAND NVMe Flash cards

- **High Capacity Rack - 1.3 PB Disk, 180 TB PCIe Flash**
- **Extreme Flash Rack - 360 TB PCIe Flash**

Database Server

44 CPU cores

High Capacity (HC) Storage



12.8 TB PCI Flash
96 TB disk
20 CPU cores

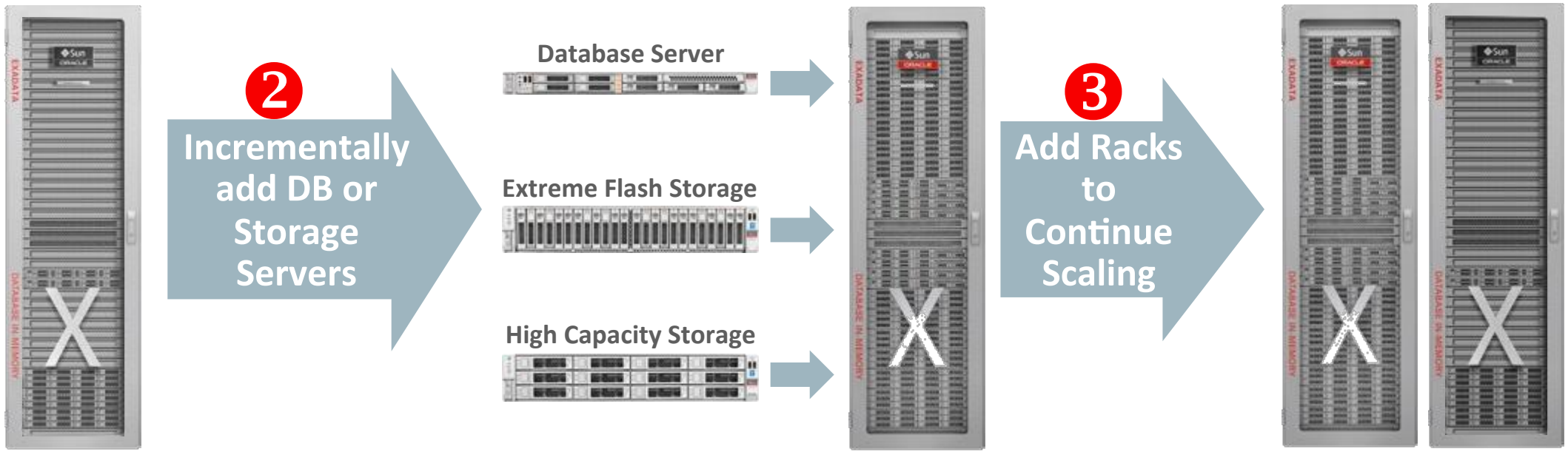
Extreme Flash (EF) Storage

25.6 TB PCI Flash
20 CPU cores

ORACLE

Elastic Configurations Incrementally Scale Servers

Achieve any Level of Performance with Minimum Hardware



- **Enable Database CPU cores as needed with Capacity on Demand**
- **Expand older Exadata machines with new X6-2 servers**

Exadata X6 Delivers Breakthrough DB I/O Performance

301 GB/sec Analytic Throughput

5.6 Million 8K OLTP Read IOPS

5.2 Million 8K OLTP Write IOPS

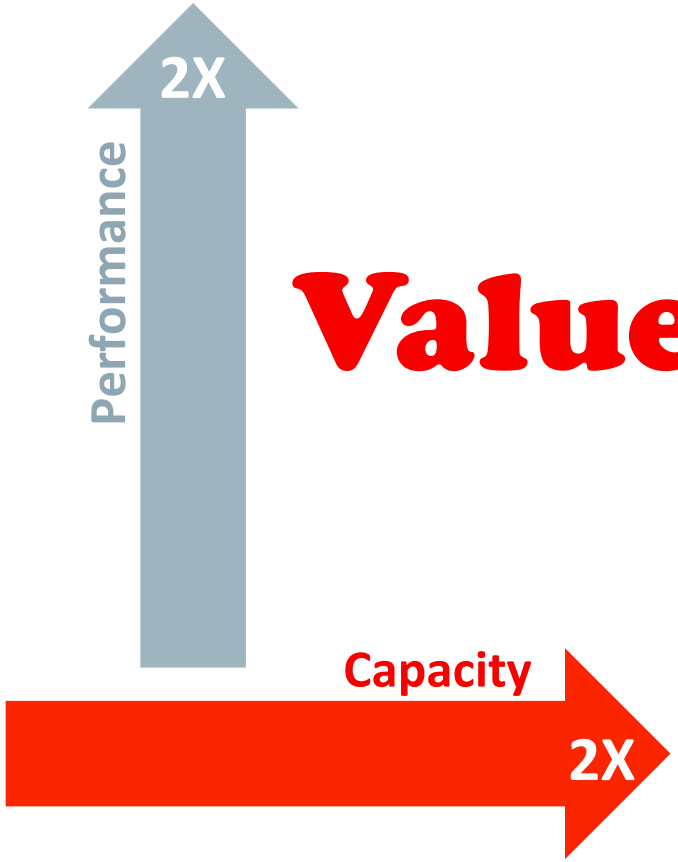
250 us I/O latency at 2.4 Million IOPS

(Scales higher as racks are added)

Performance of 1 Exadata Rack with 10 DB servers and 12 Extreme Flash storage servers

ORACLE

Exadata X6 Pricing Unchanged to Greatly Improve Value



- Huge benefits over X5 for **same price**
  - **2X disk capacity, 2X flash capacity, 2X faster flash**
  - **25% faster X6-2 Broadwell CPUs**
  - Faster DRAM
- Minimum DB licenses unchanged (COD)
  - 16 cores for Eighth Rack, 28 cores for Quarter Rack
- New X6-8 Exadata updated to X6 storage
- **Recovery Appliance** (ZDLRA) updated to X6
- Oracle **SuperCluster** updated to X6 storage

- Ideal Database Hardware
- Smart System Software
- Full-Stack Integration

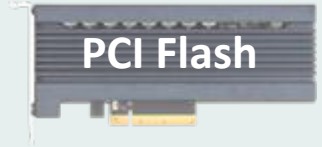
Smart System Software Highlights

Smart Analytics

- Move **queries to storage**, not storage to queries
- Automatically **offload and parallelize** queries across all storage servers
- **100X** faster analytics

Smart Storage

- **Hybrid Columnar Compression** reduces space usage by **10X**
- Database-aware **Flash Caching** gives speed of flash with capacity of disk



Smart OLTP

- **Special InfiniBand protocol** enables highest speed, lowest latency OLTP
- Ultra-fast transactions using DB optimized **flash logging** algorithms
- **Fault-tolerant In-Memory DB** by mirroring memory across servers

Smart Consolidation

- **Workload prioritization** from CPU to network to storage ensures QoS
- **4X** more Databases in same hardware

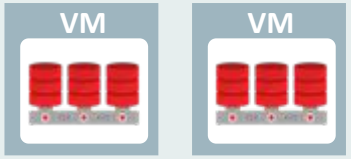
Smart System Software Introduced in 2015

Smart Analytics

- **5X** faster scans by converting data to **Columnar** format in Flash Cache
- **3X** faster **JSON/XML** by offloading to storage servers

Smart Consolidation

- Zero overhead **VMs**
- **Snapshots** for test/dev
- Set flash cache min size per DB to ensure QoS
- InfiniBand partitioning
- IPv6 for Ethernet



Smart OLTP

- **3X** faster OLTP messaging using **direct DB to InfiniBand** access
- **Instant** detection of node failure
- **Sub-second** capping of I/O latency by rerouting I/Os to faster storage

Smart Licensing

- **Capacity-on-Demand** reduces license cost by disabling unneeded cores
- **Trusted Partitions** limit license scope of specialized options

Exadata Implements Much More Smart System Software

Smart Analytics

- Storage Index data skipping
- Storage offload for min/max operations
- Data mining storage offload
- Storage offload for LOBs and CLOBs
- Auto flash caching for table scans
- Reverse offload to DB servers
- Offload index fast full scans
- Offloads scans on encrypted data, with FIPS compliance

Smart OLTP

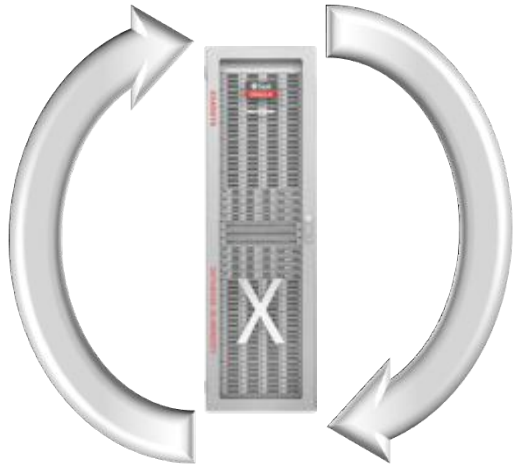
- Active AWR end to end monitoring
- Smart write-back flash cache
- Cell-to-cell rebalance preserving Flash Cache and Storage Indexes
- Database scoped security
- Full-stack security scanning
- Exachk full-stack validation
- Instant data file creation
- Active bonding of InfiniBand
- Disabling of unreliable network links
- Critical DB messages always jump to head of queue for ultra-fast latency
- I/O issued by interactive users and important workloads is prioritized

Smart Availability

- Prioritize rebalance of critical files
- Elimination of false drive failures
- Flash and disk life cycle mgmt alert
- Avoid reading predictive failed disks
- Cell software transparent restart
- I/O hang hardening
- Prevent shutdown if mirror server is down
- Confinement of temporarily poor performing drives
- Data integrity validation (HARD)
- Automatic disk scrub and repair

Super Fast Software Updates

- 2.5X speed up in Storage Server Software Update
  - Parallel firmware upgrades across components such as hard disks, flash, ILOM/BIOS, InfiniBand card
  - Reduced reboots for Software updates
- Imaging speedup
  - X5 High Capacity cell reimaged in less than 20 minutes

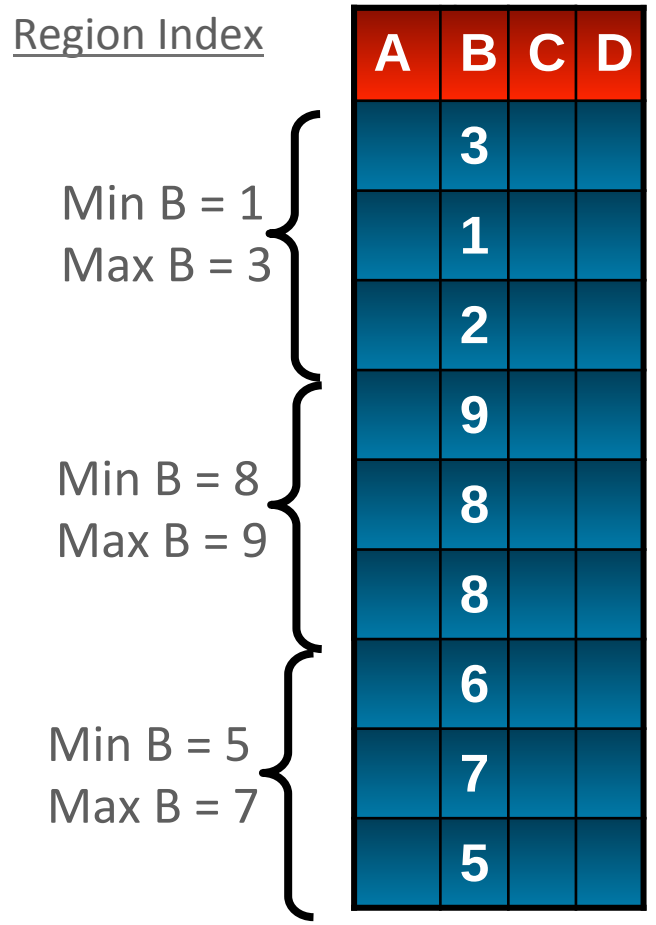


High Redundancy on Quarter and Eighth Racks

- Problem: On an eighth or a quarter rack the Voting Disks were on a Normal Redundancy disk group and thus susceptible to concurrent failures
- Solution: Create quorum disks on database servers in addition to those on storage servers
- Best Practice: Use HIGH redundancy for DATA diskgroup and place voting disk in HIGH redundancy diskgroup
- Oracle Exadata Deployment Assistant automatically creates quorum disks
- Quorum Disk Manager Utility creates and manage quorum disks
- Minimum Grid Infrastructure Software version required:
  - Oracle Database 12c Release 1 (12.1) release 12.1.0.2.160119 with these patches: 22722476 and 22682752

Storage Index Preservation across Rebalance

- In event of a disk failure data needs to be rebalanced out to disks on other cells
- Previously, storage indexes created for the regions on the failed disk were lost and recreated on the next scan
- Storage index entries will be moved along with data to the new disk during cell to cell offloaded rebalance
- Maintains application performance during rebalance
- Minimum Grid Infrastructure software version:
  - Oracle Database 12c Release 1 (12.1) release 12.1.0.2.160119 with patch 22682752



- Ideal Database Hardware
- Smart System Software
- Full-Stack Integration

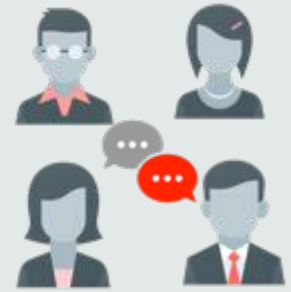
Exadata Full-Stack Integration Reduces Operations Costs

Unique Full-Stack Integration

- All layers pre-**configured**, pre-**tuned**, pre-**debugged**
  - DB, OS, drivers, firmware, network, servers, storage

Unique Full-Stack Reliability

- All users run identical full stack
- You get:
  - Bank tested full-stack **HA**
  - Telco tested full-stack **scaling**
  - Government tested full **security**

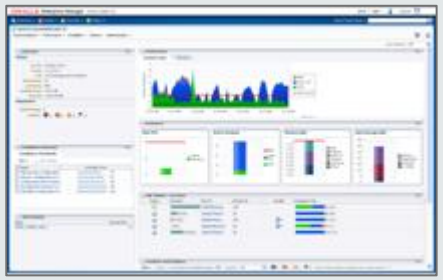


Unique Full-Stack Support

- **One support team** expert in and accountable for full stack
- Oracle performs **free full-stack updates and 24/7 monitoring**

Unique Full-Stack Management

- **Full-stack management tool**
- Drill down from DB to storage and up from storage to DB



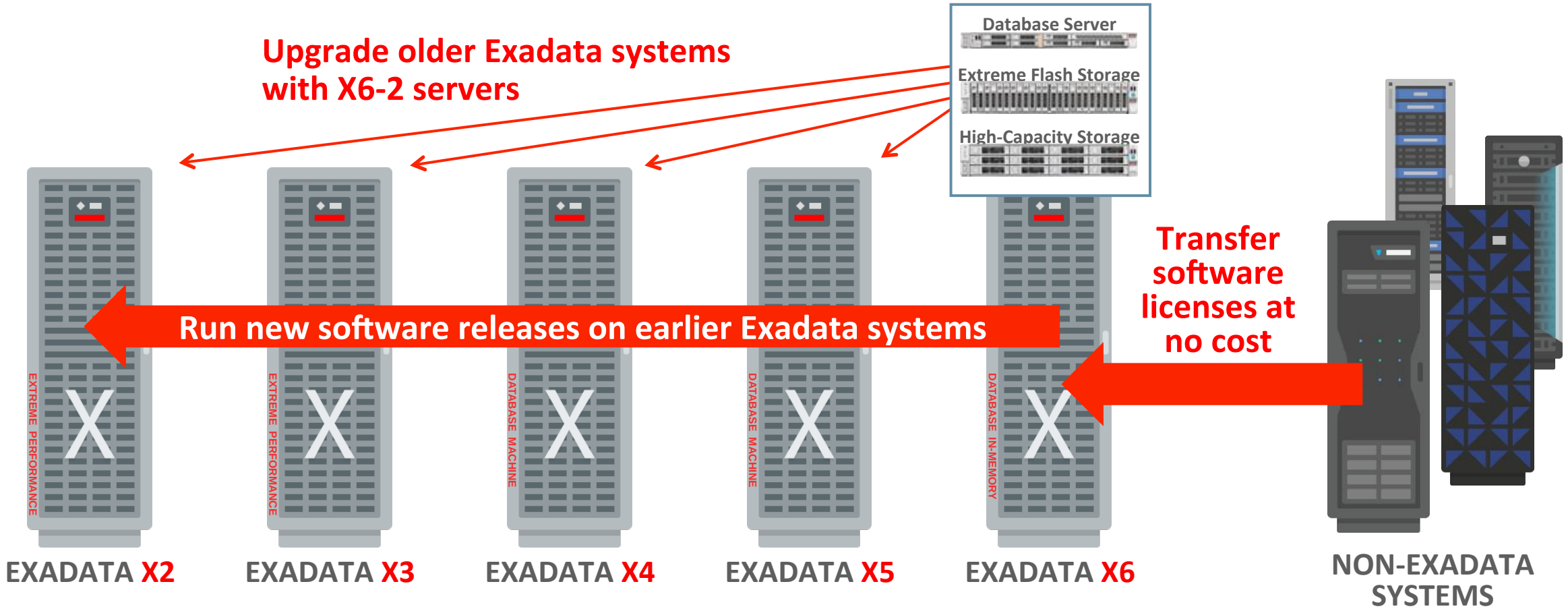
Exadata Hardware Generational Advances*

	V1	V2	X2	X3	X4	X5	X6	
								
	Sept 2008	Sept 2009	Sept 2010	Sept 2012	Nov 2013	Dec 2014	Apr 2016	
	Xeon E5430 Q4, 2007 Intel GA	Xeon E5540 Q1, 2009 Intel GA	Xeon X5670 Q1, 2010 Intel GA	Xeon E5-2690 Q1, 2012 Intel GA	Xeon E5-2697v2 Q3, 2013 Intel GA	Xeon E5-2699 v3 Q3, 2014 Intel GA	Xeon E5-2699 v4 Q1, 2016 Intel GA	**X6 Growth Factor**
Storage (TB)	168	336	504	504	672	1344	1344	8 X
Flash (TB)	0	5.3	5.3	22.4	44.8	89.6	179.2	32 X
CPU (Cores)	64	64	96	128	192	288	352	5.5 X
Memory (GB)	256	576	1152	2048	4096	6144	6144	24 X
Ethernet (Gb/s)	8	24	184	400	400	400	400	50 X

** Assumes typical full rack configuration of 8 database servers and 14 storage servers*

Exadata Investment Protection

Expand Existing Systems with X6-2 Servers - Run New Software on Older Hardware



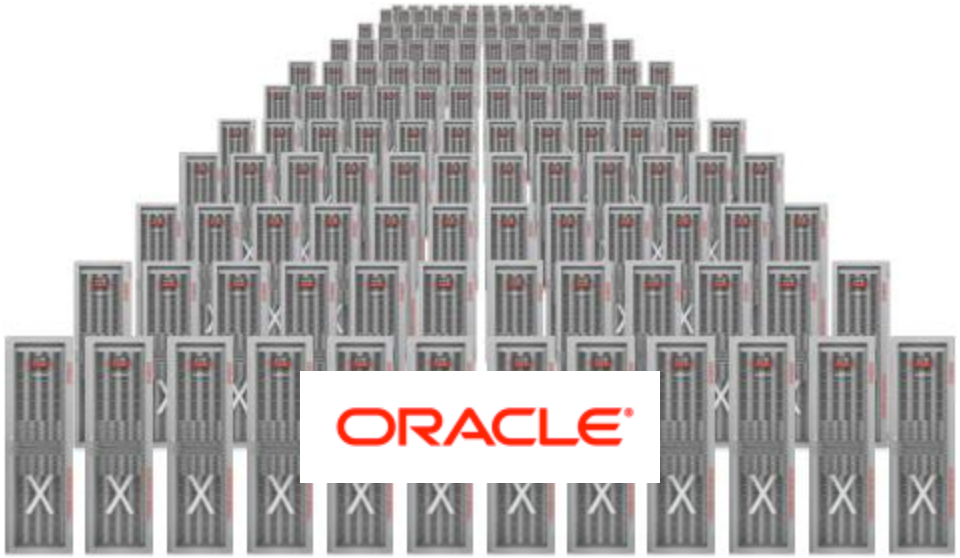
The Exadata "Community Effect"

Exadata Takes Standardization to the Next Level

- Standard Technologies
- Standard Configurations
- Standard Integration
- Standard Tuning
- Standard Support

The **Scalability of Telecom**,
The **Availability of Banking**,
The **Security of Government**

Oracle Public Cloud
Oracle Development & Support
1,000's of Customers and Partners



Server Vendor A
Storage Vendor B
Network Vendor C
Database Vendor D
OS Vendor E
VM Vendor F

The New Global Standard **vs** **A Company Standard**

Exadata Advantages Increase Every Year

Dramatically Better Platform for All Database Workloads

- **Exadata Cloud Service**
- In-Memory Columnar in Flash
- Smart Fusion Block Transfer
- In-Memory Fault Tolerance
- Direct-to-wire Protocol
- JSON and XML offload
- Instant failure detection
- Network Resource Management
- Multitenant Aware Resource Mgmt
- Prioritized File Recovery
- 3D V-NAND Flash
- Software-in-Silicon
- IO Priorities
- Data Mining Offload
- Offload Decrypt on Scans
- Tiered Disk/ Flash

Smart Software

- Database Aware Flash Cache
- Storage Indexes
- Columnar Compression
- PCIe NVMe Flash
- Unified InfiniBand
- Smart Scan
- InfiniBand Scale-Out

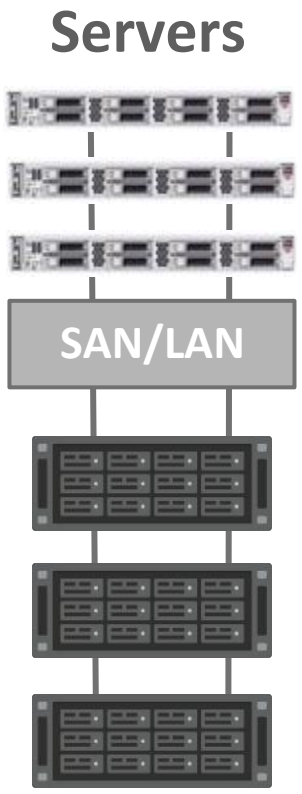
Smart Hardware

- DB Processors in Storage
- Scale-Out Storage
- Scale-Out Servers

Exadata Industry Breakthroughs

- Exadata Cloud Service
- Shared Flash at Memory Speed
- Real-Time, All-the-Time OLTP

Shared Storage Has Many Advantages over Local Storage



- Much better **space utilization**
- Much better **security, management, reliability**
- Enables DB **consolidation**, DB **high availability**, RAC **scale-out**
- **Shares storage performance**
  - Aggregate performance of shared storage can be dynamically used by any server that needs it

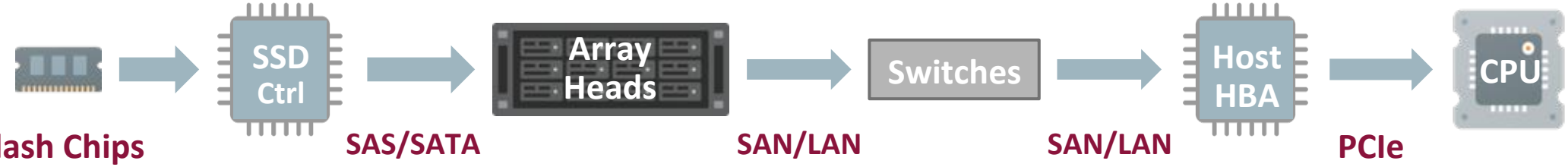
Flash Performance is Wasted by Shared Storage Arrays

SAN Link = 40Gb
5 GB/sec

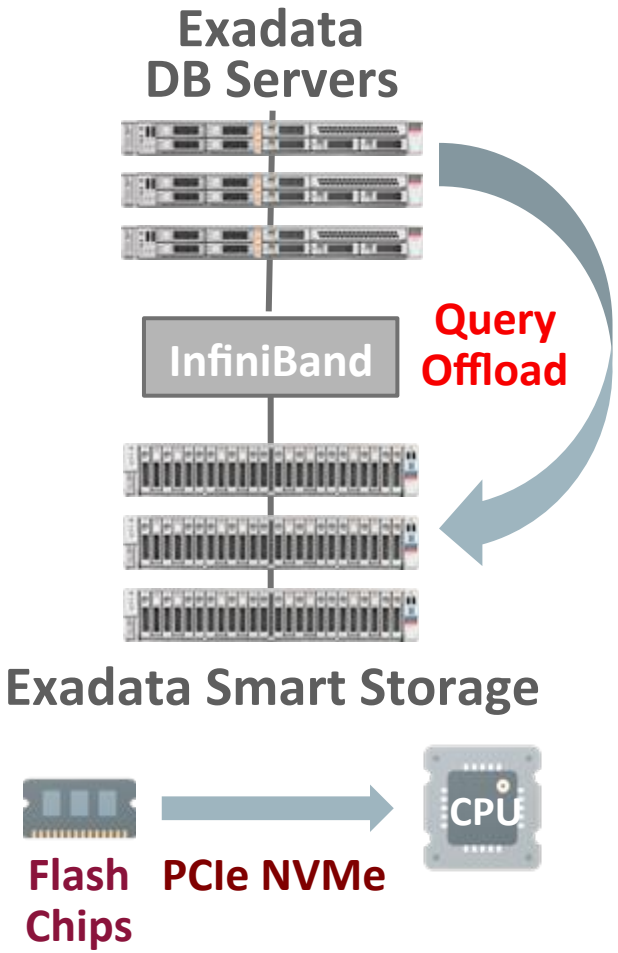
Latest PCIe Flash
5.4 GB/sec

- **New improvements in flash performance are causing 100X bottlenecks across shared storage stack**
  - Speed of one flash card is now similar to fastest SAN or LAN link
  - Throughput of a few flash cards is too fast to transfer to servers

All-Flash Storage Array IO Path: many steps, each adds latency and creates bottlenecks



Exadata Achieves Memory Performance with Shared Flash



- **Exadata X6 delivers 300GB/sec flash bandwidth to any server**
  - Approaches 800GB/sec aggregate DRAM bandwidth of DB servers
- **Must move compute to data to achieve full flash potential**
  - Requires owning full stack, can't be solved in storage alone
- **Fundamentally, storage arrays can share flash capacity but not flash performance**
  - Even with next gen scale-out, PCIe networks, or NVMe over fabric
- **Shared storage with memory-level bandwidth is a paradigm change in the industry**
  - Get near DRAM throughput, with the capacity of shared flash

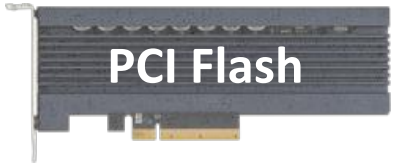
Exadata Industry Breakthroughs

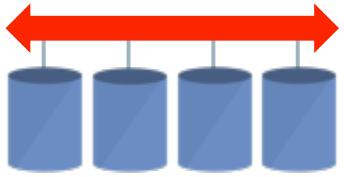
- Exadata Cloud Service
- Shared Flash at Memory Speed
- Real-Time, All-the-Time OLTP

Users Demand Instant Fulfillment of Any Information Need

- **Patience is an extinct virtue**
- **Customers** demand instant product information and immediate action
  - Prospects lose interest and move-on in seconds
- **Employees** demand instant results from internal applications
- **Instant fulfillment requires OLTP systems to deliver results in** real-time, all-the-time
  - Scale OLTP to any volume
  - Deliver instant results all-the-time
  - Even when consolidating databases to reduce costs

Exadata Uniquely Eliminates Bottlenecks to Scaling OLTP

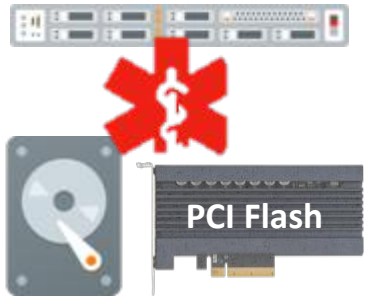




- **Exadata eliminates the traditional OLTP bottleneck: Random I/O**
  - Unique **scale-out storage** with ultra-fast PCIe flash, ultra-fast NVMe protocol, ultra-fast InfiniBand, and ultra-fast iDB protocol delivers:
    - Over **5 Million** DB reads or writes per rack; **¼ millisecond** latency
  - Unique **Smart Flash Logging** eliminates OLTP slowdowns and bottlenecks due to logging
- **Exadata eliminates Inter-node coordination bottlenecks for OLTP**
  - Unique **Direct-to-Wire Protocol** gives **3x** faster inter-node OLTP messaging
  - Unique **Smart Fusion Block Transfer** eliminates the need to write the log file when moving blocks between nodes
- **Exadata delivers instant analytic results directly on OLTP data**
  - Unique **query offload** to storage, shared PCIe flash, InfiniBand, and In-Memory Database scaling

Exadata Uniquely Delivers OLTP in Real-Time, All-the-Time

- **Exadata eliminates mixed-workload interference on OLTP operations**
  - Unique **OLTP message bypass** of prior queued messages at host, switches, and storage
  - Unique **prioritization of OLTP I/Os** from DB to disk bypassing I/Os from batch/reports/ analytics
  - Unique **storage offload** of RMAN backup
  - Unique **prioritization of operations** against important database vs less important

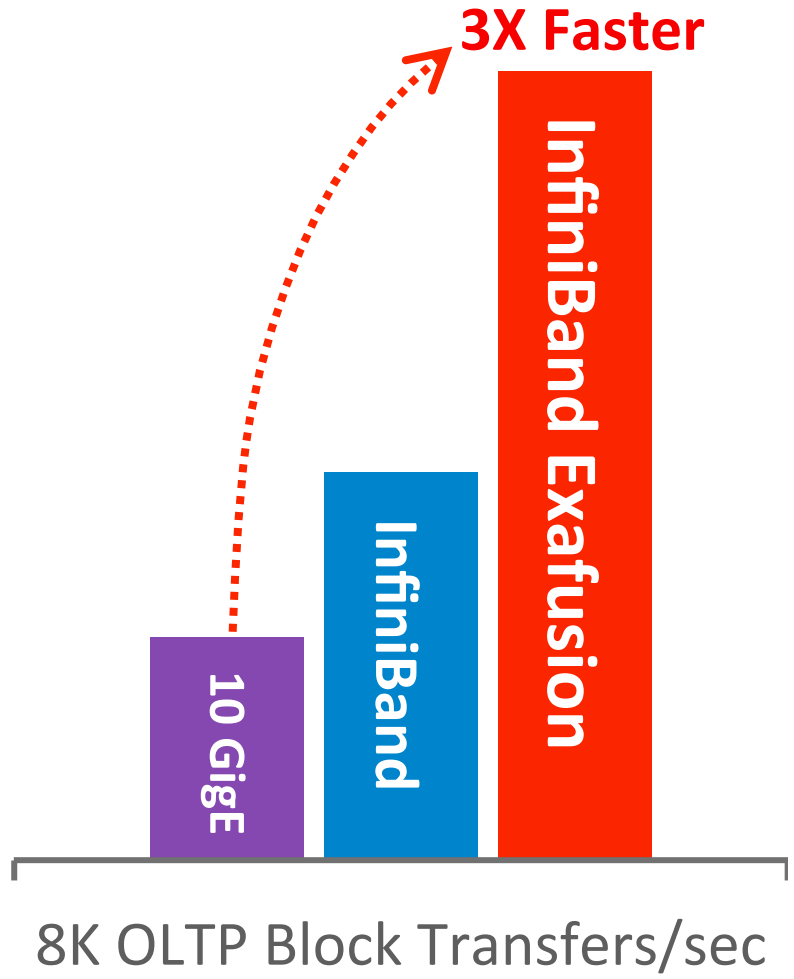


- **Exadata eliminates OLTP stalls from failed or sick components**
  - Unique **full-stack integration** for fastest recovery of failed server, storage or switch/link
  - Unique **detection** of server failures without a long timeout avoids system hangs
  - Unique **sub-second redirection** of I/Os around sick devices avoids database hangs
  - Unique **In-Memory Database Fault Tolerance**

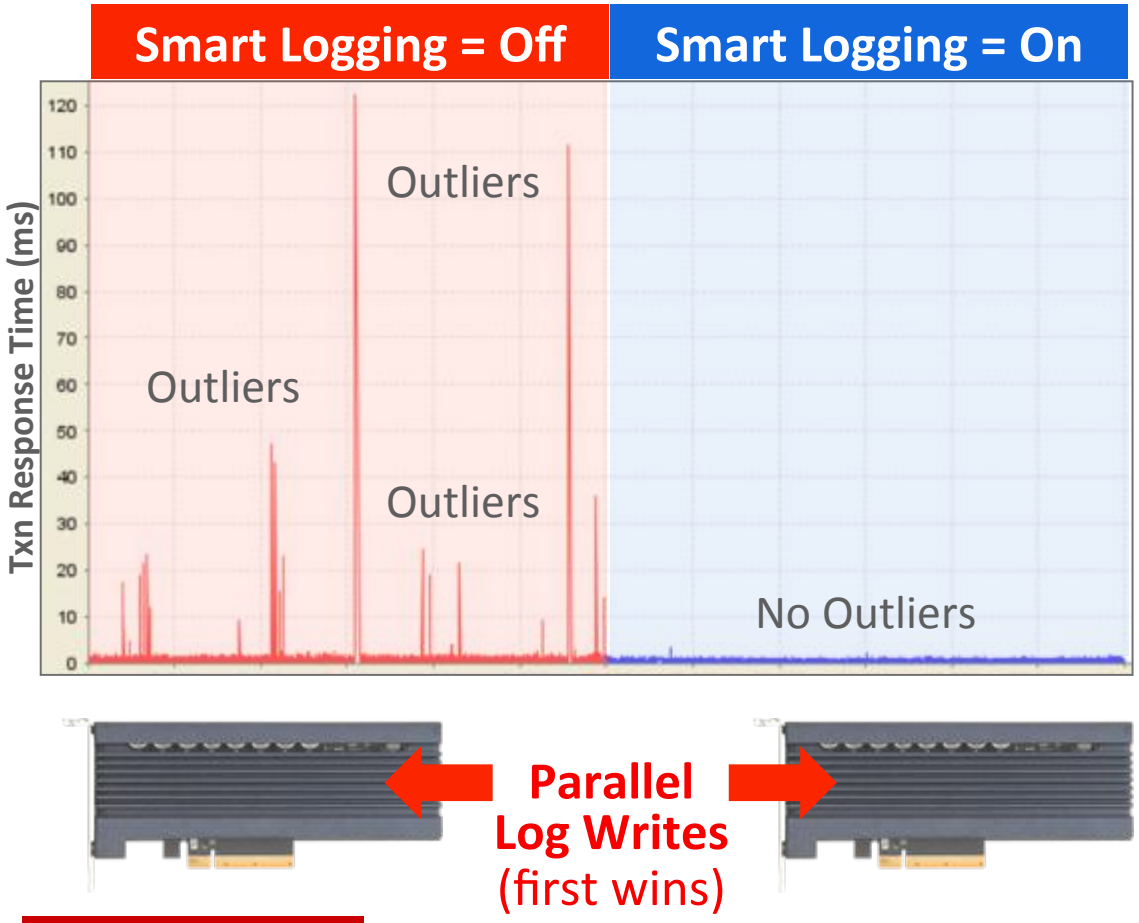
Exafusion Direct-to-Wire OLTP Protocol

World's First Database to InfiniBand Protocol

- **InfiniBand has great throughput**
  - But overheard of calling through OS on every message limits small message rate
- **Exafusion re-implements RAC Cache Fusion**
- **Database directly calls InfiniBand hardware**
  - Bypasses networking software stack, interrupts, scheduling



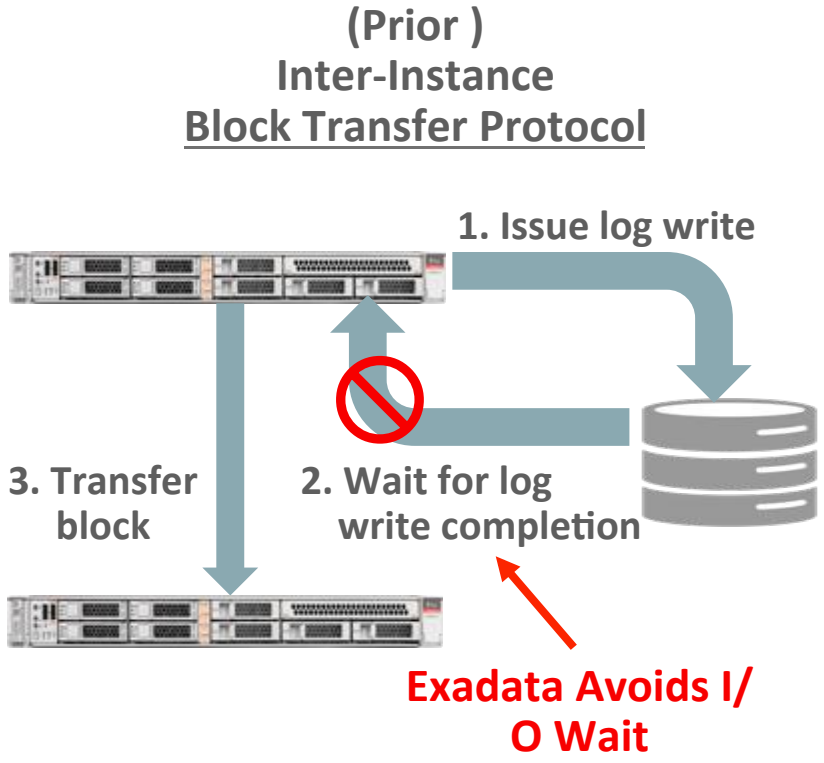
Exadata Smart Flash Logging



- **Flash delivers great I/O throughput for OLTP**
- But **Flash has 100X spikes in response time**
- Response spikes are not a problem for most DB I/Os, but are **big issue for OLTP log writes**
  - Transaction commits must occur in order, so a slow log write stalls all transactions
- Smart Flash Logging uses flash as a parallel write cache to disk controller cache or other flash
- Write that completes first wins (disk or flash)
- Shown to greatly improve OLTP throughput and response times in real workloads

Smart Fusion Block Transfer

- **OLTP workloads can have hot blocks that are frequently updated (e.g. right-growing index)**
  - Log file must be written before transferring a hot block between RAC instances so the block can be recovered
  - Adds latency and reduces throughput
- **On Exadata, Oracle does not wait for the log write**
  - Exadata ensures the log write completes before changes to block on another instance commit, guaranteeing durability
  - Wait for Log I/O during transfer of hot blocks is eliminated
  - Up to **40% throughput** and **33% response time improvement** in some heavily contended OLTP workloads



Integrated Cloud

Applications & Platform Services


ORACLE®