ORACLE

Exadata Hardware Overview

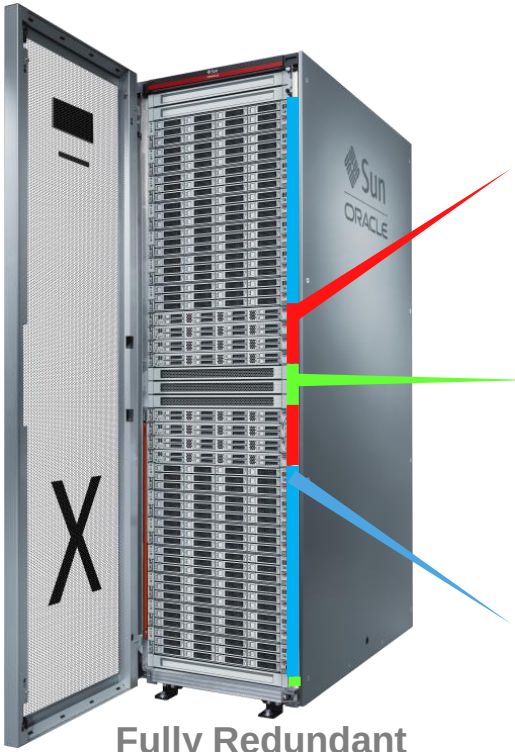
Exadata Database Machine

- **The ultimate platform for all database workloads**
  - OLTP, Warehousing, Database as a Service
- **Most advanced <u>hardware</u>**
  - Fully scale-out servers and intelligent storage with unified InfiniBand connectivity and PCI flash
- **Most advanced <u>software</u>**
  - Database optimized compute, storage, and networking algorithms dramatically improve performance and cost
- **Standardized, optimized, hardened end-to-end**

Exadata Hardware

Exadata Architecture

Complete | Optimized | Standardized | Hardened Database Platform



Fully Redundant

Standard Database Servers

- 8x 2-socket servers ➔ 192 cores, 2TB DRAM
  or
- 2x 8-socket servers ➔ 160 cores, 4TB DRAM

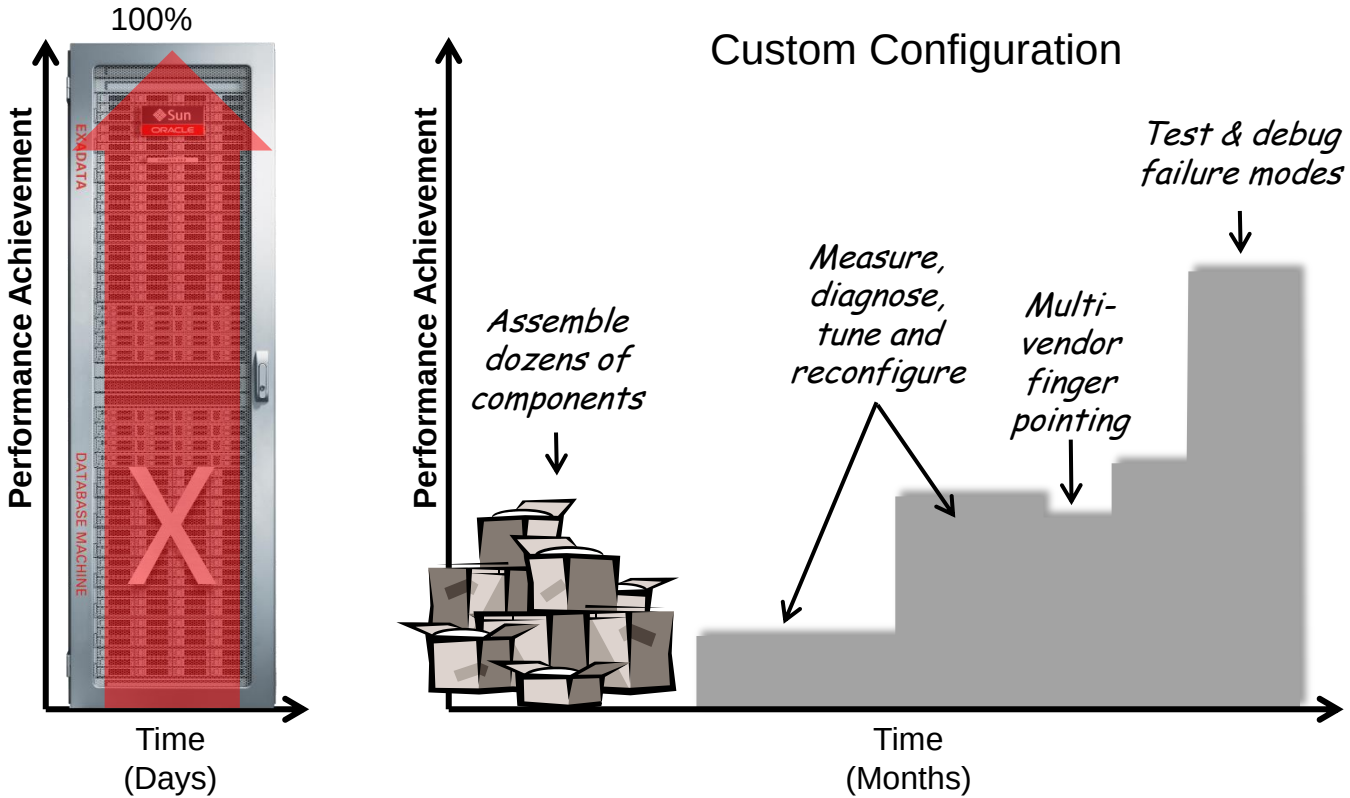
Unified Ultra-Fast Network

- 40 Gb InfiniBand internal connectivity ➔ all ports active
- 10 Gb or 1 Gb Ethernet data center connectivity

Scale-out Intelligent Storage Servers

- 14x 2-socket servers ➔ 168 cores in storage
- 168 SAS disk drives ➔ 672 TB HC or 200 TB HP
- 56 Flash PCI cards ➔ 44 TB Flash + compression

Pre-built and Optimized Out-of-the-Box



Standardized and Simple to Deploy

Deploy in Days, Not Months

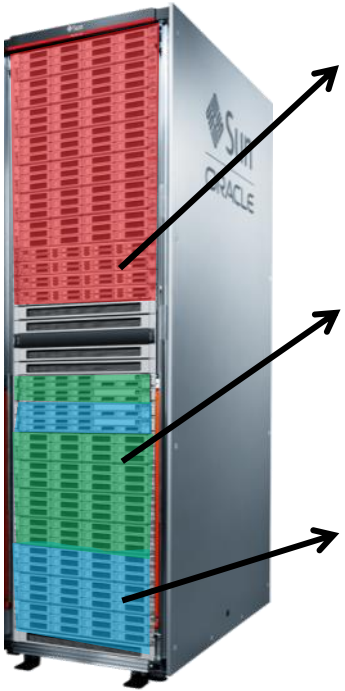
- **All Database Machines are the same**
  - Delivered ready-to-run
  - Tested
  - Highly supportable
  - No unique configuration issues
  - Identical to configuration used by Oracle Engineering
- **Runs existing OLTP and DW applications**
  - Full 30 years of Oracle DB capabilities
  - No Exadata certification required
- **Leverages Oracle ecosystem**
  - Skills, knowledge base, people and partners

Scalable from Eighth-Rack to Multi-Rack

Field Upgradeable - Supports Multiple Generations of Hardware

Seamless Upgrades and Expansions

Upgrade Example



X4-2
Half to Full Upgrade in 2014

X3-2
Qtr to Half Upgrade in 2013

X2-2
Initial Quarter Rack deployed in 2011

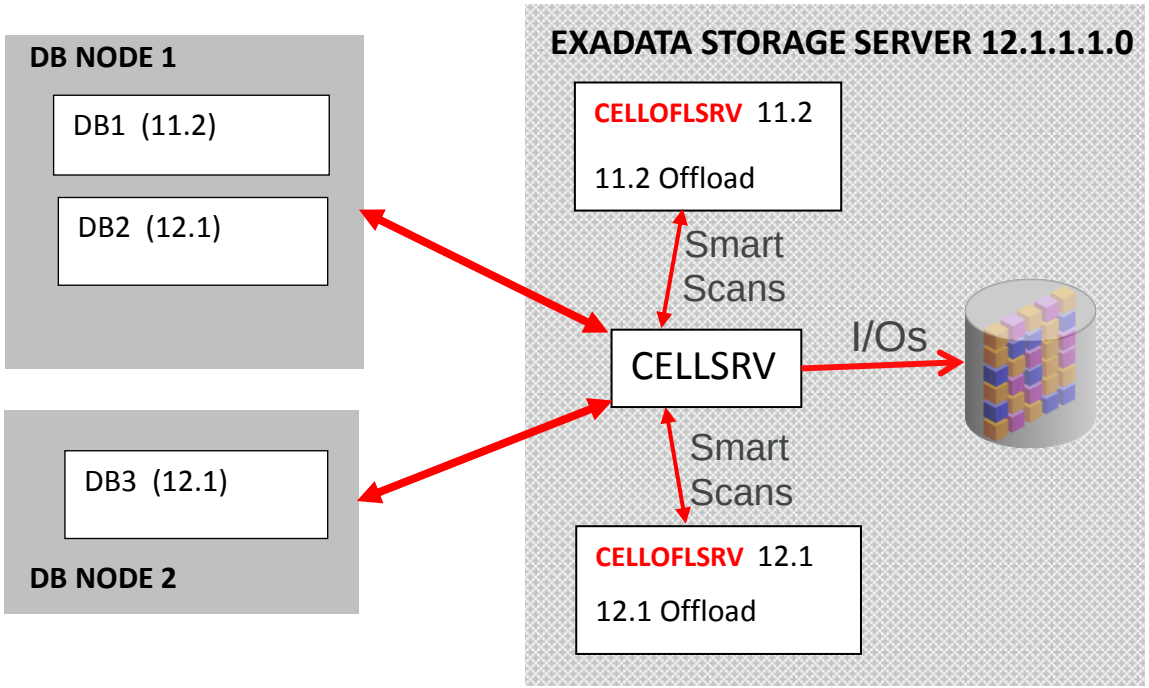
- **A single Database Machine can have servers from different generations**
- **Databases and Clusters can span across multiple hardware generations**
- **New software runs on older hardware**

Hardware Generational Advances

Storage (TB)	168	336	504	504	672	**4X**
Flash (TB)	0	5.3	5.3	22.4	44.8	**8X**
CPU (Cores)	64	64	96	128	192	**3X**
Memory (GB)	256	576	1152	2048	4096	**16X**
Connectivity (Gb/s)	8	24	184	400	400	**50X**

ORACLE

Support for Multiple DB versions on Exadata



- Exadata Systems can run a mix of 11.2 and 12.1 databases
- Different offload server process for each DB version
  - Matches DB libraries

X4-2 Database Machine IO Performance from SQL

		X4-2 Full Rack	X4-2 Half Rack	X4-2 Quarter	X4-2 Eighth
Flash Cache SQL Bandwidth[1,3]	High Cap Disk	100 GB/s	50 GB/s	21.5 GB/s	10.7 GB/s
	High Perf Disk	100 GB/s	50 GB/s	21.5 GB/s	10.7 GB/s
Flash SQL IOPS[2,3]	8K Reads	2,660,000	1,330,000	570,000	285,000
	8K Writes	1,960,000	980,000	420,000	210,000
Disk SQL Bandwidth[1,3]	High Cap Disk	20 GB/s	10 GB/s	4.5 G/s	2.25 GB/s
	High Perf Disk	24 GB/s	12 GB/s	5.2 GB/s	2.6 GB/s
Disk SQL IOPS	High Cap Disk	32,000	16,000	7,000	3,500
	High Perf Disk	50,000	25,000	10,800	5,400
Data Load Rate[4]		20 TB/hr	10 TB/hr	5 TB/hr	2.5 TB/hr

1 - Bandwidth is peak physical scan bandwidth achieved running SQL, assuming no compression. Effective data bandwidth will be much higher when compression is factored in.
2 - IOPS – Based on read IO requests of size 8K running SQL, typically with sub-millisecond latencies. Note that the IO size greatly effects flash IOPS. Others quote IOPS based on smaller IOs that are not relevant for databases and measure IOs using low level tools instead of SQL.
3- Actual Performance varies by application.
4 –Load rates are typically limited by database server CPU, not IO. Rates vary based on load method, indexes, data types, compression, and partitioning

X4-2 Storage Expansion IO Performance from SQL

		X4-2 Full Rack	X4-2 Half Rack	X4-2 Quarter	X4-2 Single Cell
Flash Cache SQL Bandwidth[1,3]	High Cap Disk	130 GB/s	65 GB/s	29 GB/s	7.25 GB/s
	High Perf Disk	130 GB/s	65 GB/s	29 GB/s	7.25 GB/s
Flash SQL IOPS[2,3]	8K Reads	3,420,000	1,710,000	760,000	190,000
	8K Writes	2,520,000	1,260,000	560,000	140,000
Disk SQL Bandwidth[1,3]	High Cap Disk	26 GB/s	13 GB/s	6 GB/s	1.5 GB/s
	High Perf Disk	30 GB/s	15 GB/s	7 GB/s	1.75 GB/s
Disk SQL IOPS	High Cap Disk	42,000	21,000	9,500	2,400
	High Perf Disk	64,000	32,000	14,400	3,600

1 - Bandwidth is peak physical scan bandwidth achieved running SQL, assuming no compression. Effective data bandwidth will be much higher when compression is factored in.
2 - IOPS – Based on read IO requests of size 8K running SQL, typically with sub-millisecond latencies. Note that the IO size greatly effects flash IOPS. Others quote IOPS based on smaller IOs that are not relevant for databases and measure IOs using low level tools instead of SQL. Exadata Flash read IOPS are so high they are typically limited by database server CPU, not IO. This is especially true for expansion racks.
3- Actual Performance varies by application.

Exadata Database Machine X4-2 Full Rack

Pre-Configured for Extreme Performance

- 8 Xeon-based Dual-processor Database Servers
  - 192 cores (24 per server)
  - 2048 GB memory expandable to 4096 GB (256 GB per server expandable to 512 GB)
  - 10 Gig E-connectivity to Data Center
    - 40 x 10Gb E-ports (5 per server)
- 44.8 TB High Speed Flash
- 14 Exadata Storage Servers X4-2
  - All with High Performance 12 x 1.2 TB SAS disks

  OR
  - All with High Capacity 12 x 4 TB SAS disks
- 2 Sun Datacenter InfiniBand Switch 36
  - 36-port Managed QDR (40Gb/s) switch
- 1 “Admin” Ethernet switch
- Redundant Power Distributions Units (PDUs)

Add more racks for additional scalability

Exadata X4-2 Memory Expansion Kit

- Optional kit for X4-2 customers who consolidate large numbers of databases and require more memory
  - Changes memory configuration of the X4-2 database server from 256 GB (16 x 16 GB) to 512 GB (16 x 32 GB)

- Available for X4-2 Full, Half, Quarter, and Eighth Racks

Exadata Storage Expansion X4-2 Full Rack

For additional storage – backups, historical data, unstructured data

- 18 Exadata Storage Servers X4-2
  - All with High Performance 12 x 1.2 TB SAS disks

  OR
  - All with High Capacity 12 x 4 TB SAS disks
- 57.6 TB High Speed Flash
- 3 Sun Datacenter InfiniBand Switch 36
  - 36-port Managed QDR (40Gb/s) switch
- 1 “Admin” Ethernet switch
- Redundant Power Distributions Units (PDUs)

Add storage capacity to Database Machine online

X4-2 Database Server

New 12-core “IvyBridge” CPUs, Faster InfiniBand Card, Larger Disks

Processors	**2 Twelve-Core Intel® Xeon® E5-2697 v2 Processors (2.7GHz)**
Memory	256 GB (16 x 16GB) – Expandable to 512GB (16 X 32GB) via memory kits
Local Disks	4 x **600GB** 10K RPM SAS Disks (Hot-Swappable)
Disk Controller	Disk Controller HBA with 512MB Cache **– Battery Online Replaceable**
Network	2 x InfiniBand 4X QDR (40Gb/s) Ports (**PCIe 3.0**) – **Both Ports Active** 4 x 1GbE/10GbE Base-T Ethernet Ports 2 x 10GbE Ethernet SFP+ Ports (1 Dual-port 10GbE PCIe 2.0 network card based on the Intel 82599 10GbE Controller technology)
Remote Management	1 Ethernet port (ILOM)
Power Supplies	Redundant Hot-Swappable power supplies and fans

X4-2 Storage Server

6-core IvyBridge CPUs, Larger disks, Larger Flash Cards, Flash Compression

Processors	**2 Six-Core Intel® Xeon® E5-2630 v2 Processors (2.6 GHz) - Faster clock**
Memory	**96 GB** (4 x 8GB + 4 x 16GB) **- More memory needed to manage larger flash**
Disks	12 x **1.2 TB 10K RPM** High Performance SAS (hot-swap) – **2.5" disk size** OR 12 x **4 TB** 7.2K RPM High Capacity SAS (hot-swap) – 3.5" disk size
Flash	4 x **800 GB** Sun Flash Accelerator **F80** PCIe Cards – **Hardware Compression**
Disk Controller	Disk Controller HBA with 512MB Cache **- Battery Online Replaceable**
Network	2 InfiniBand 4X QDR (40Gb/s) Ports (**PCIe 3.0**) – **Both Ports Active** Embedded Gigabit Ethernet Ports for management connectivity
Remote Management	1 Ethernet port (ILOM)
Power Supplies	Redundant Hot-Swappable power supplies and fans

Exadata Database Machine Hardware Summary

		X4-2 Full	X4-2 Half	X4-2 Quarter	X4-2 Eighth
Database Servers		8	4	2	2
Database Grid Cores		192	96	48	24
Database Grid Memory (GB)		2048 (max 4096)	1024 (max 2048)	512 (max 1024)	512 (max 1024)
InfiniBand switches		2	2	2	2
Ethernet switch		1	1	1	1
Exadata Storage Servers		14	7	3	3
Storage Grid CPU Cores		168	84	36	18
Raw Flash Capacity		44.8 TB	22.4 TB	9.6 TB	4.8 TB
Raw Storage Capacity	**High Perf**	200 TB	100 TB	43.2 TB	21.6 TB
	High Cap	672 TB	336 TB	144 TB	72 TB
Usable mirrored capacity	**High Perf**	90 TB	45 TB	19 TB	9 TB
	High Cap	300 TB	150 TB	63 TB	30 TB
Usable Triple mirrored capacity	**High Perf**	60 TB	30 TB	13 TB	6.3 TB
	High Cap	200 TB	100 TB	43 TB	21.5 TB

Exadata Storage Expansion Rack Summary

		X4-2 Full	X4-2 Half	X4-2 Quarter	Single Cell
InfiniBand switches		3	3	2	-
Ethernet switch		1	1	1	-
Exadata Storage Servers		18	9	4	1
Storage Grid CPU Cores		216	108	48	12
Raw Flash Capacity		57.6 TB	28.8 TB	12.8 TB	3.2 TB
Raw Storage Capacity	**High Perf**	258 TB	129 TB	57 TB	14.4 TB
	High Cap	864 TB	432 TB	192 TB	48 TB
Usable mirrored capacity	**High Perf**	116 TB	58 TB	25 TB	6 TB
	High Cap	387 TB	194 TB	85 TB	20 TB
Usable Triple mirrored capacity	**High Perf**	78 TB	39 TB	17 TB	4 .25 TB
	High Cap	260 TB	130TB	58 TB	14.5 TB

Database Machine Capacity (Uncompressed)

		X4-2 Full	X4-2 Half	X4-2 Quarter	X4-2 Eighth
Raw Flash Capacity[1,4]		44.8 TB	22.4 TB	9.6 TB	4.8 TB
Effective Flash Capacity Due to Caching[4]		440 TB	220TB	96TB	48TB
Raw Disk Capacity[1]	High Cap Disk	672 TB	336 TB	144 TB	72 TB
	High Perf Disk	200 TB	100 TB	43.2 TB	21.6 TB
Usable Mirrored Capacity [2,3]	High Cap Disk	300 TB	150 TB	63 TB	30 TB
	High Perf Disk	90 TB	45 TB	19 TB	9 TB
Usable Triple Mirrored Capacity[2,3]	High Cap Disk	200 TB	100 TB	43 TB	21.5 TB
	High Perf Disk	60 TB	30 TB	13 TB	6.3 TB

1 - Raw Disk Capacity defined using standard disk drive terminology of 1 TB = 1000 * 1000 * 1000 * 1000 bytes.
2 - Capacity calculated using normal space terminology of 1 TB = 1024 * 1024 * 1024 * 1024 bytes.
3 - Actual space available for a database after mirroring (ASM normal or high redundancy). For the ASM normal redundancy case, assume one disk (Quarter and Half) or two disks (Full Rack) of free space to automatically remirror after disk failures.
4 – 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 data files that often can be stored in Exadata and be accessed at the speed of flash memory.

Exadata Storage Expansion Rack Summary

		X4-2 Full	X4-2 Half	X4-2 Quarter	Single Cell
InfiniBand switches		3	3	2	-
Ethernet switch		1	1	1	-
Exadata Storage Servers		18	9	4	1
Storage Grid CPU Cores		216	108	48	12
Raw Flash Capacity		57.6 TB	28.8 TB	12.8 TB	3.2 TB
Raw Storage Capacity	**High Perf**	258 TB	129 TB	57 TB	14.4 TB
	High Cap	864 TB	432 TB	192 TB	48 TB
Usable mirrored capacity	**High Perf**	116 TB	58 TB	25 TB	6 TB
	High Cap	387 TB	194 TB	85 TB	20 TB
Usable Triple mirrored capacity	**High Perf**	78 TB	39 TB	17 TB	4 .25 TB
	High Cap	260 TB	130TB	58 TB	14.5 TB

Operating System Updates in 11.2.3.3.0

- Oracle Linux distribution updated to 5.9
  - Including all recent security updates

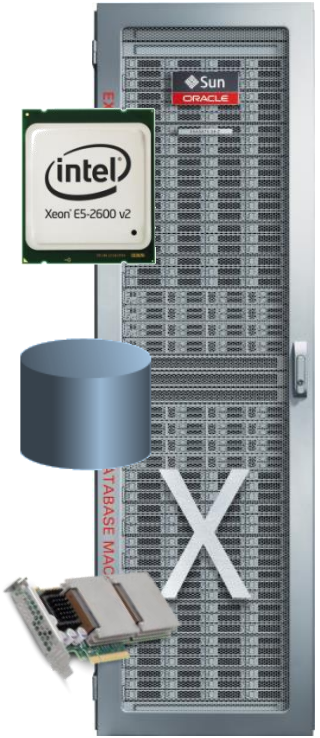
- Linux kernel updated to UEK2 kernel (2.6.39-400)
  - Same kernel on database servers (2-socket and 8-socket) and the storage servers

- Oracle Solaris updated to S11 Update 1 SRU 9

Exadata Innovations

Unique Software Optimizes Database Processing

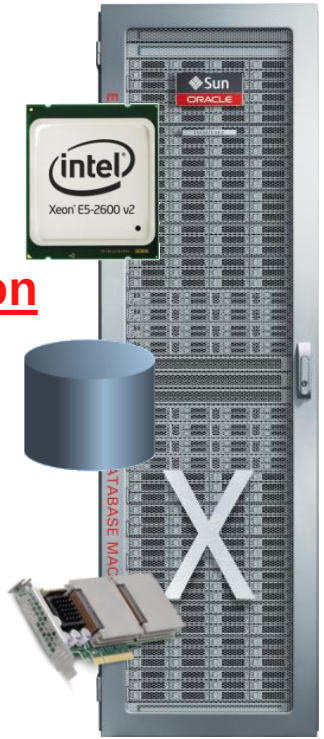
- **Query offload in storage**
  - Data intensive query operations offloaded to storage CPUs
  - 100 GB/sec SQL data throughput
  - Storage Index data skipping
- **Database storage compression**
  - Hybrid Columnar for 10x DB size reduction and faster analytics
- **Database optimized PCI Flash**
  - Smart caching of database data
  - 2.66 Million Database IOs/sec
  - Smart Flash log speeds transactions



- **Database optimized QoS**
  - End-to-end prioritization from application to DB and storage
- **Database optimized availability**
  - Fastest recovery of failed database, server, storage or switch
  - Fastest backup. Incremental offload
  - Exachk top-to-bottom validation of hardware, software, settings
- **Database optimized messaging**
  - SQL optimized InfiniBand protocol for high throughput low latency SQL

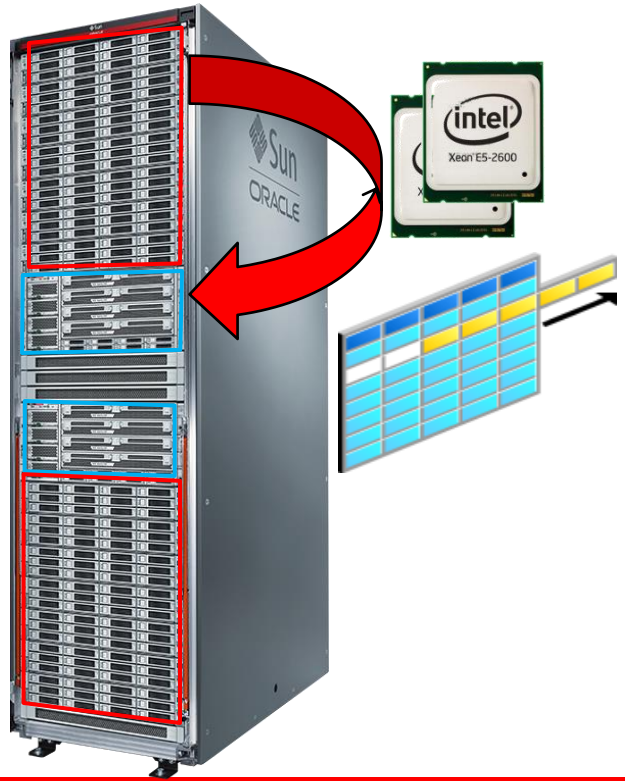
Innovation Continues: Recent Enhancements

- **Query offload in storage**
  - Offload searches on LOBs (12*c*)
  - Offload joins for non-parallel queries (11.2.0.4)
- **Database optimized compression**
  - Hybrid Columnar enhanced for OLTP
  - and for Spatial and Text data (12*c*)
- **Database optimized PCI Flash**
  - Ultra high speed flash compression (X3 & X4) at multi-million IOs/sec
  - Automatic caching for table scans
  - Faster file initialization



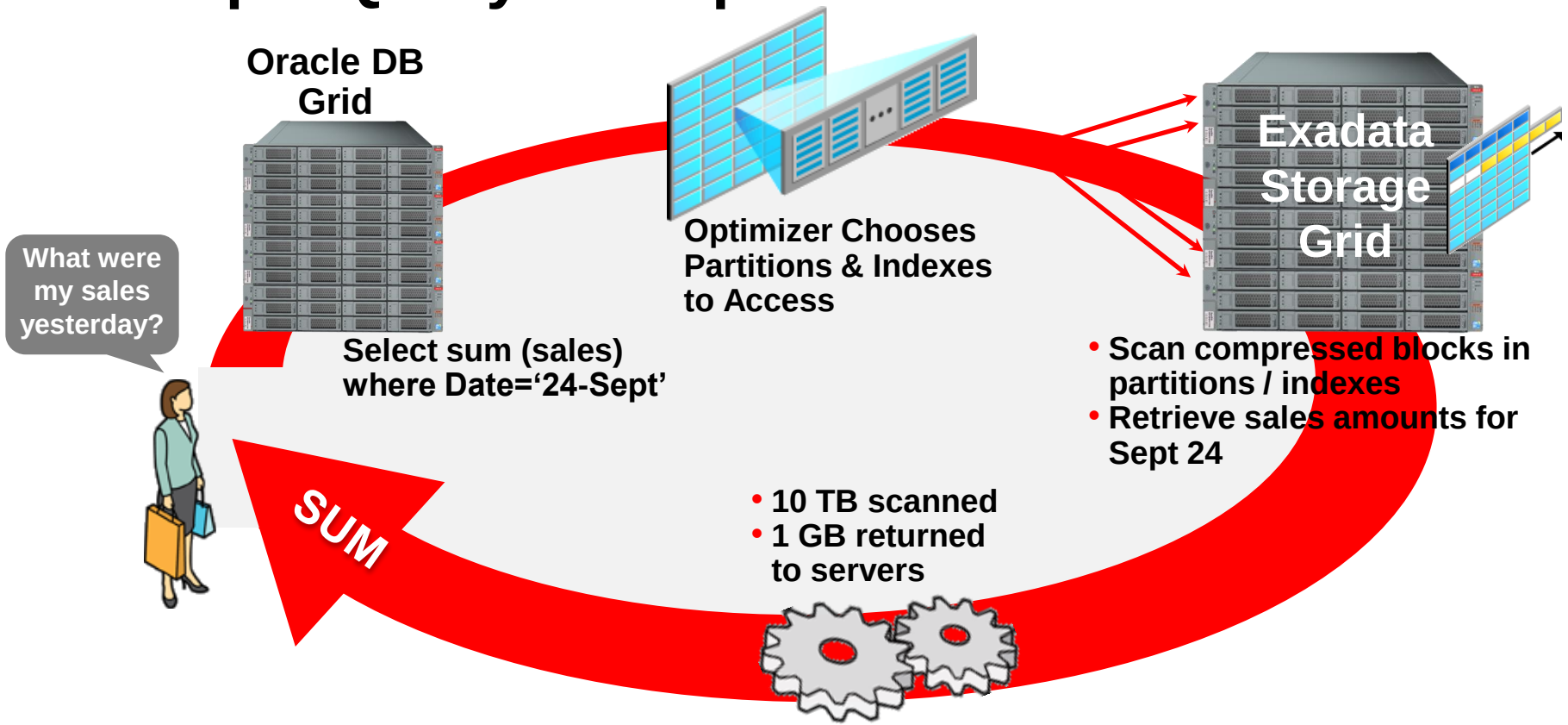
- **Database optimized QoS**
  - Prioritization of CPU and IO by multitenant pluggable database (12*c*)
- **Database optimized availability**
  - Prioritize recovery of critical DB files (11.2.0.4)
- **Database optimized messaging**
  - End-to-End prioritization of critical database messages (11.2.0.4), including log writes and RAC

Exadata Intelligent Storage Grid

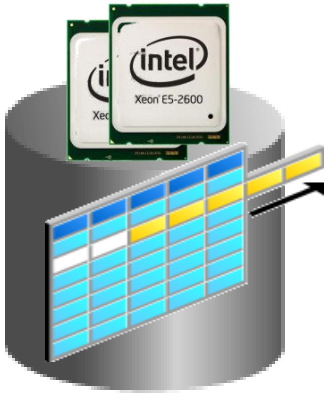


- **Data Intensive processing runs in Exadata Storage Grid**
  - Filter rows and columns as data streams from disks (168 Intel Cores)
- Example: How much product X sold last quarter
  - Exadata Storage Reads 10TB from disk
  - Exadata Storage Filters rows by Product & Date
  - Sends 100GB of matching data to DB Servers
- Scale-out storage parallelizes execution and removes bottlenecks

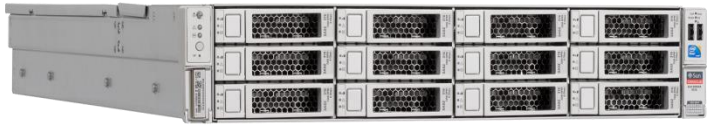
Simple Query Example



Exadata Intelligent Storage

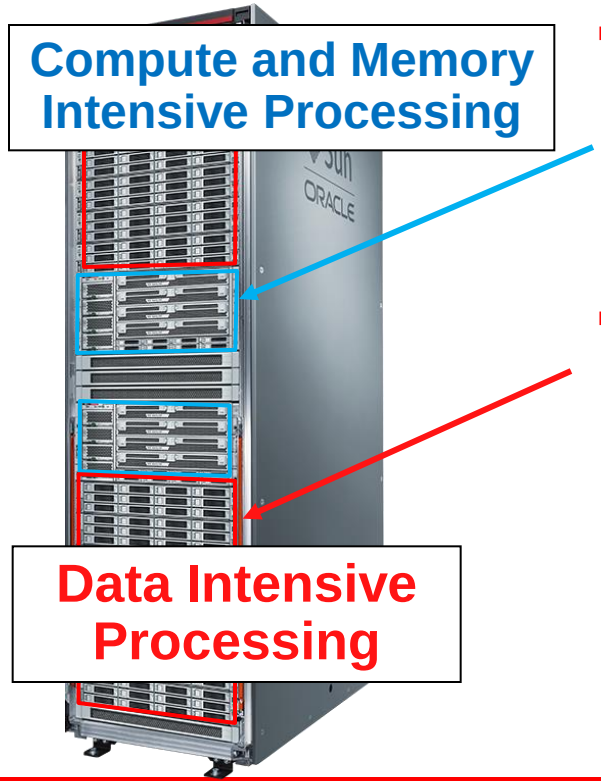


Exadata Intelligent Storage Grid



- Exadata storage servers also run more complex operations in storage
  - **Join filtering**
  - **Incremental backup filtering**
  - **I/O prioritization**
  - **Storage Indexing**
  - **Database level security**
  - **Offloaded scans on encrypted data**
  - **Data Mining Model Scoring**
- 10x reduction in data sent to DB servers is common

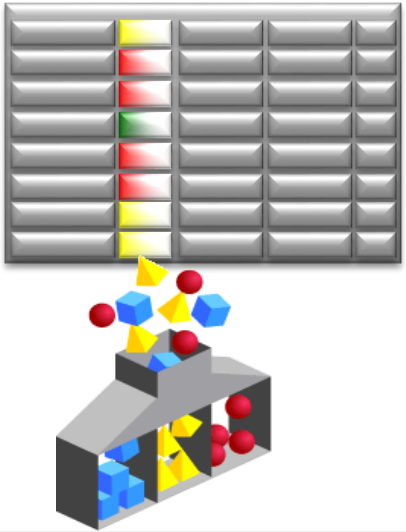
Exadata is Smart Storage



- **Database Servers**
  - Perform complex database processing such as joins, aggregation, etc.
- **Exadata Storage Servers**
  - Storage Server is smart storage, not a DB node
  - Search tables and indexes filtering out data that is not relevant to a query
  - Cells serve data to multiple databases **enabling OLTP and consolidation**
  - Simplicity, and robustness of storage appliance

Exadata Hybrid Columnar Compression

Highest Capacity, Lowest Cost



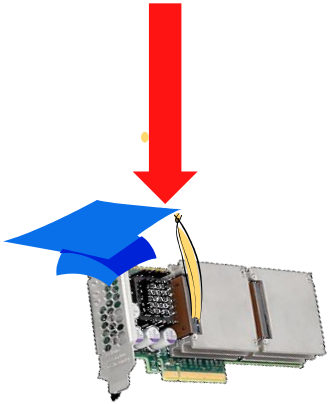
- Data is organized and compressed by column
  - Dramatically better compression
- Speed Optimized **Query Mode** for Data Warehousing
  - **10X compression typical**
  - **Runs faster because of Exadata offload!**
- Space Optimized **Archival Mode** for infrequently accessed data
  - **15X to 50X compression typical**

Faster and Simpler
Backup, DR, Caching, Reorg, Clone

Benefits Multiply

Exadata Smart Flash Cache

I/Os

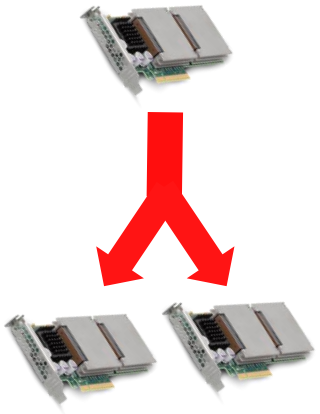


2.66 Million 8K Read
1.96 Million 8K Write
IOPS from SQL

- Caches Read and Write I/Os in PCI flash
- Transparently accelerates read and write intensive workloads
  - Up to 2.66 million 8K read IOPS from SQL
  - Up to 1.96 million 8K write IOPS from SQL
- Persistent write cache speeds database recovery
- Exadata Flash Cache is much more effective than flash tiering architectures used by others
  - Caches current hot data, not yesterday's
  - Caches data in granules 8x to 16x smaller than tiering
    - Greatly improves the effectiveness of flash

Exadata Flash Cache Compression

New

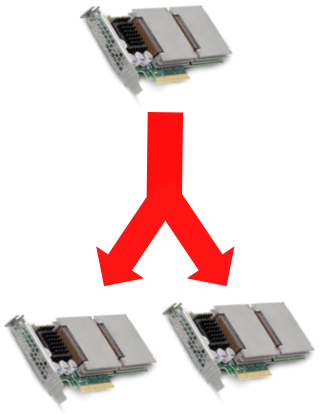


- Exadata uses compression to expand smart flash cache
  - Data automatically compressed as it is written to flash cache
  - Automatically decompressed when it is read out of flash cache
  - Up to 2X more data fits in smart flash cache, so flash hit rates will improve and performance will improve for large data sets
- Flash cache compress/decompress implemented in hardware
  - Performance is same as uncompressed – millions of I/Os per second
    - ZERO performance overhead
  - Supported on X3 or X4 storage servers (requires F40 or F80 cards)
- Note – Flash cache compression does not change DB format, so does not improve backups, buffer cache, network, scan rates, etc.

Exadata Flash Cache Compression

New

- As always, compression benefits vary based on data

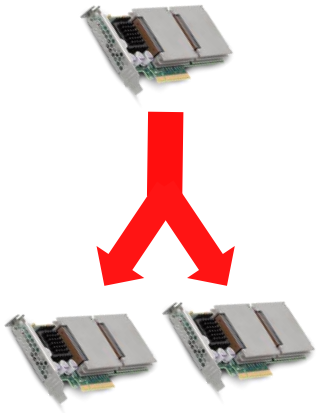


Data Type	Typical Compression
Uncompressed Tables	1.3X to 4X
OLTP Compressed Tables	1.2X to 2X
Indexes	1.3X to 4X
Oracle E-biz uncompressed DB	3x to 5x
HCC Compressed Tables or Compressed LOBs	Minimal

Many OLTP Databases will see 2x Flash Increase

- X4 with flash cache compression stores up to 80TB of data in flash
  - Up to 4X more than X3 (depending on compressibility of data)

Flash Cache Compression Commands



- Trivial to implement, no management
- Enable using simple cell command:
  - **`On X4 machines: alter cell flashCompression=TRUE`**
  - **`On X3 machines also run:`**
    - **`alter cell FlashCacheCompX3Support= TRUE`**
- Amount of data cached in Exadata Smart Flash Cache grows and shrinks dynamically and automatically based on data compressibility
- Monitor Flash Cache Compression using cell metric FC_BY_USED
  - Reported flash cache size will increase to up to double physical flash size

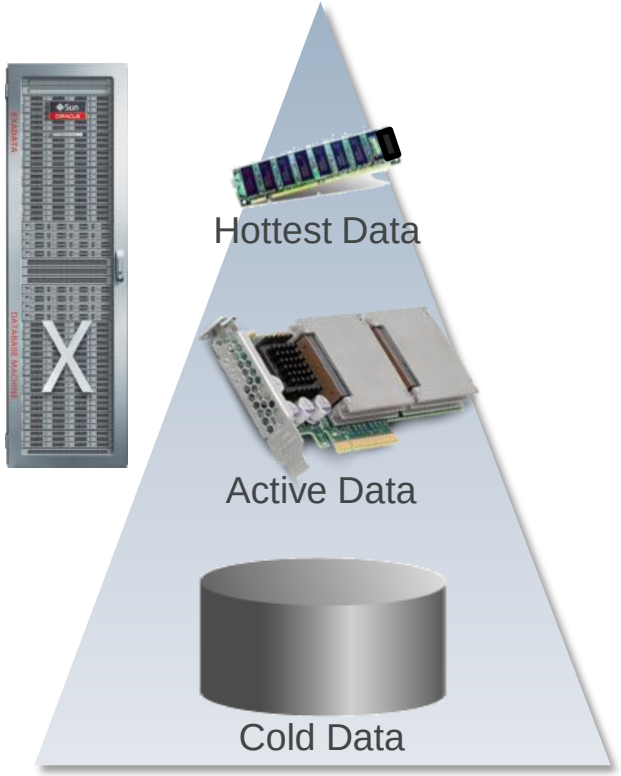
Exadata Smart Flash Table Caching

New

- Smarter flash caching for large table scans
  - Exadata software understands database table and partition scans and automatically caches then when it makes sense
  - Avoids thrashing flash cache when tables are too big or scanned infrequently or scanned by maintenance jobs
  - If scanned table is larger than flash, then subset of table is cached
  - No need to manually “KEEP” tables that are only scanned

Scan

Business Benefits of X4 Database Machine



- Sub-millisecond latency
  - Interactive OLTP with millions of users
- Sub-second analytics
  - Real-time decision making, instant reports
- 10x faster parallel jobs
  - Quarter close, payroll, supply planning, field inventory, pricing, route planning, sub-ledger accounting

Exadata Management

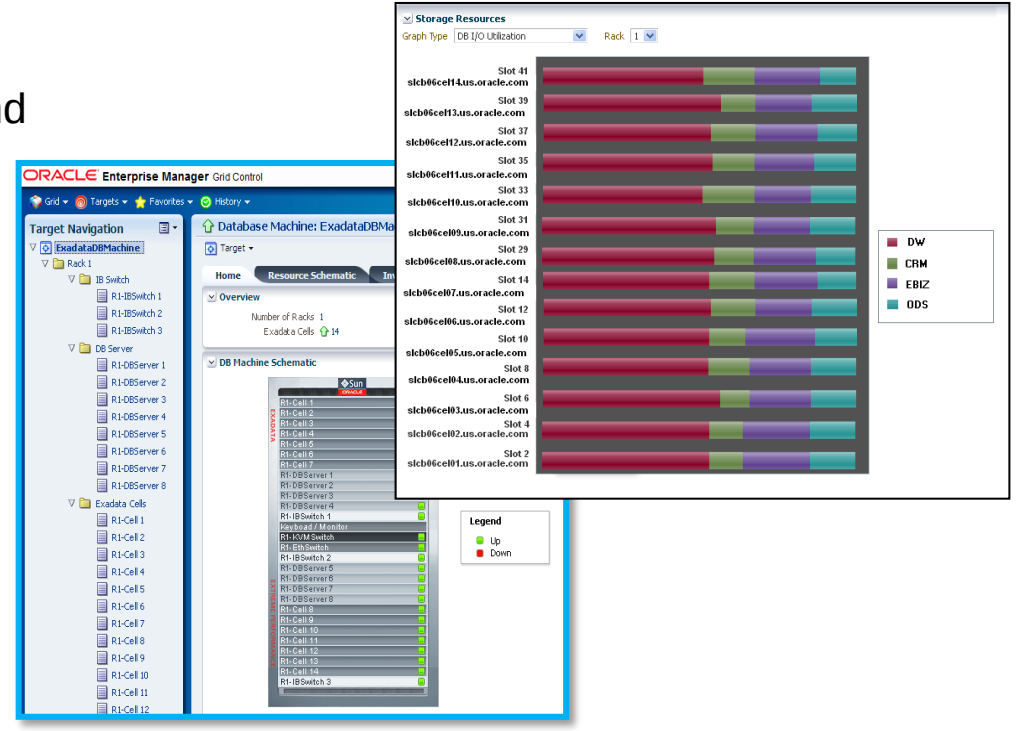
Exadata Storage Management & Administration

- Enterprise Manager
  - Manage & administer Database and ASM
  - Monitor the Exadata Database Machine Hardware
- Auto Service Request (ASR)
  - File SRs automatically for common hardware faults
- Comprehensive CLI
  - Local Exadata Storage cell management
  - Distributed shell utility to execute CLI across multiple cells
- Embedded Integrated Lights Out Manager (ILOM)
  - Remote management and administration of hardware

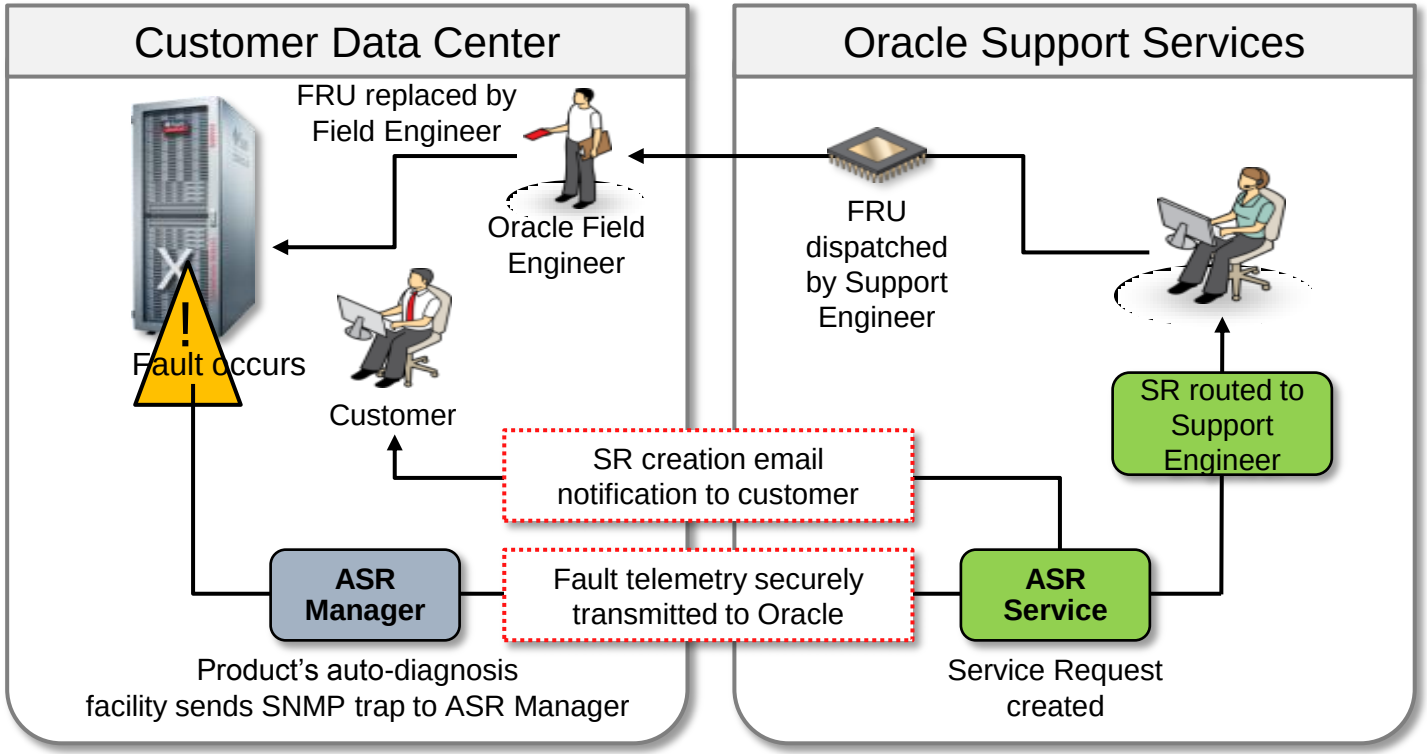
Enterprise Manager 12c

Integrated H/W + S/W management for Exadata

- **Hardware view**
  - Schematic of cells, compute nodes and switches
  - Hardware components alerts
- **Software/system view**
  - Performance, availability, usage by databases, services, clusters
  - Software alerts db, cluster, ASM
  - Topology view of DB systems/clusters
- **Configuration view**
  - Version summary of all components along with patch recommendations



Automated Service Request (ASR)



Comprehensive Fault Coverage

- CPU
- Disk controllers
- Disks
- Flash Cards
- Flash modules
- InfiniBand
- Cards
- Memory
- System Board
- Power supplies
- Fans

Platinum Support for Exadata

ORACLE EXADATA ORACLE DATABASE

ORACLE PREMIER SUPPORT

Complete. Integrated. Proactive.

- **24/7 support**
- **Specialized Engineered Systems Support Team**
- **2-hour onsite response** to hardware issues[1]
- **New Updates and Upgrades for Database, Server, Storage, and OS software**
- **My Oracle Support** proactive support portal
- **"Phone home"** automated service requests (ASR)

ORACLE PLATINUM SERVICES

High Availability. No Additional Cost.

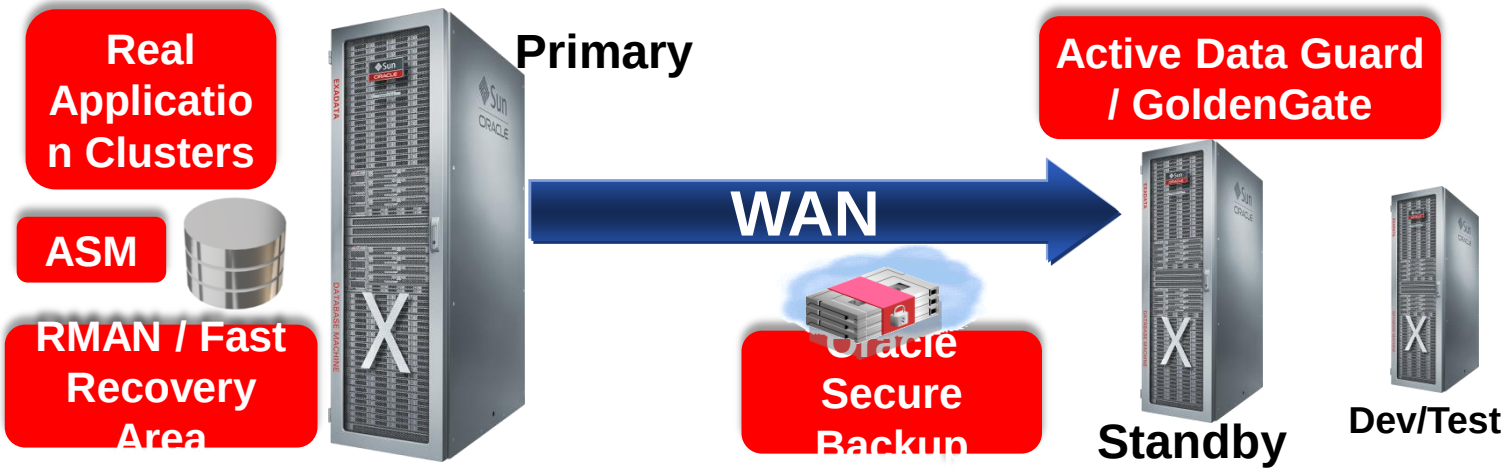
- **Better support for the complete Oracle stack**
  - Includes higher support levels for Database software
- **Proactive remote monitoring for faults**
- **Industry leading service level response times:**
  - **5 Minute** Fault Notification
  - **15 Minute** Restoration or Escalation to Development
  - **30 Minute** Joint Debugging with Development
- **Oracle Engineers perform quarterly patching and updates**

Available for certified configurations on Exadata

[1] *Covered system must be within an Oracle two-hour service area to receive two-hour response as a standard service.*

ORACLE

Exadata Maximum Availability Architecture



Comprehensive protection from failures

Server – Storage – Network – Site – Corruptions

Active Disaster Recovery: Real-time standby open for query offload

Correction from human errors: database, table, row, transaction

Online indexing and table redefinition

Online patching and upgrades

InfiniBand Network

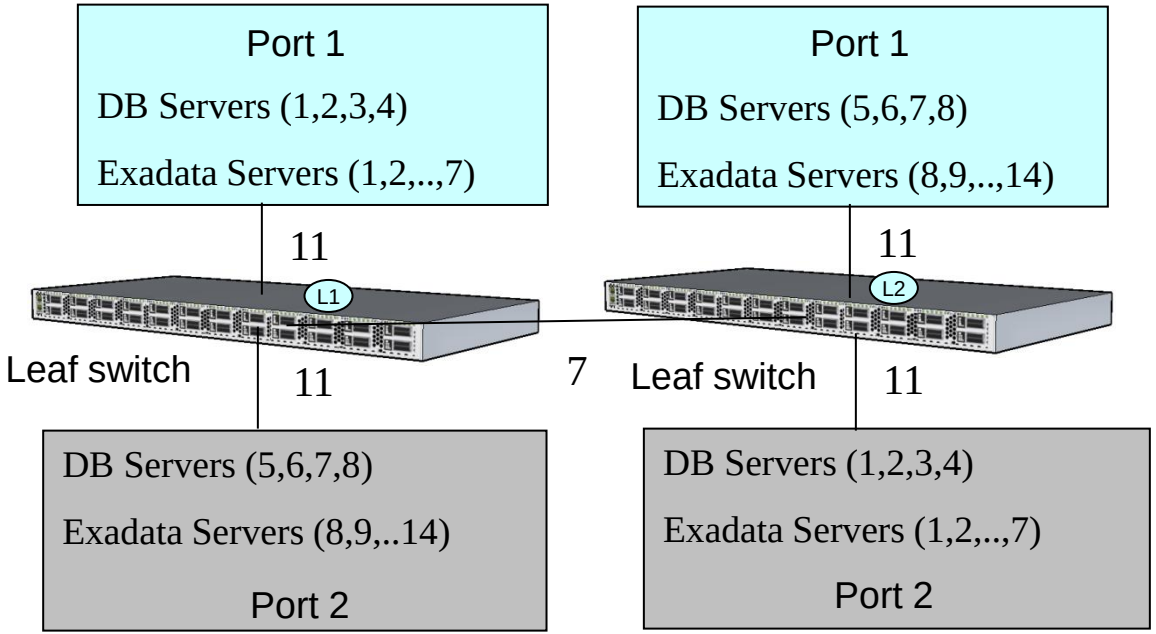
InfiniBand Network

- Unified InfiniBand Network
  - Storage Network
  - RAC Interconnect
  - External Connectivity (optional)
- High Performance, Low Latency Network
  - 80 Gb/s bandwidth per link (40 Gb/s each direction)
  - SAN-like Efficiency (Zero copy, buffer reservation)
  - Simple manageability like IP network
- Protocols
  - Zero-copy Zero-loss Datagram Protocol (ZDP RDSv3)
    - Low CPU overhead (Transfer 3 GB/s with 2% CPU usage)
  - Internet Protocol over InfiniBand (IPoIB)
    - Looks like normal Ethernet to host software (tcp/ip, udp, http, ssh,…)

InfiniBand Network

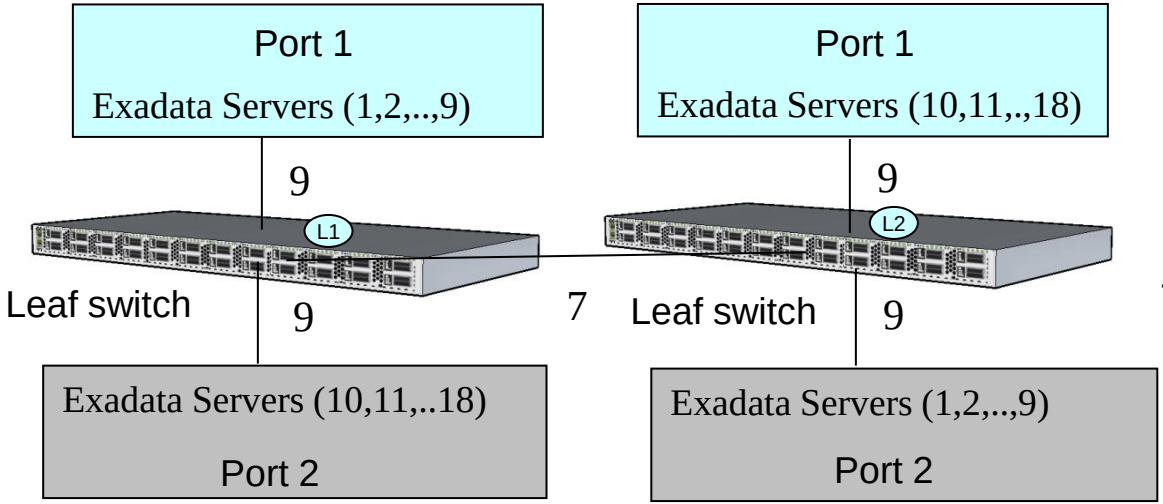
- Uses Sun Datacenter 36-port Managed QDR (40Gb/s) InfiniBand switches
  - Runs subnet manager and automatically discovers network topology
  - Only one subnet manager active at a time
  - 2 “leaf” switches to connect individual server IB ports
  - 1 “spine” switch needed in each rack to scale out to additional Racks
- Database Server and Exadata Servers
  - Each server has Dual-port QDR (40Gb/s) IB HCA (X3-8 database server has 4 dual-port QDR Express modules)
  - Active-Active Bonding – Assign two IP addresses per dual-port HCA for X4-2 database and X4-2 storage servers
    - X3-8 Database servers continue with Active-Passive as it is still PCI 2.0
  - Connect one port from the HCA to one leaf switch and the other port to the second leaf switch for redundancy
  - Connections pre-wired in the Factory

X4-2 Full Rack – Server to Leaf Switch



- Each Server HCA connects to both leaf switches for redundancy
- Full Bandwidth even if switch fails
- Connections pre-wired at factory
- Half, Quarter, and Eighth rack follow same cabling methodology except they have fewer servers

Storage Expansion Rack– Server to Leaf Switch



- Each Server HCA connects to both leaf switches for redundancy
- Active & Passive ports balanced between leaf switches
- Full Bandwidth even if switch fails
- Connections pre-wired at factory
- Half and Quarter Expansion rack follow same cabling methodology except they have fewer servers

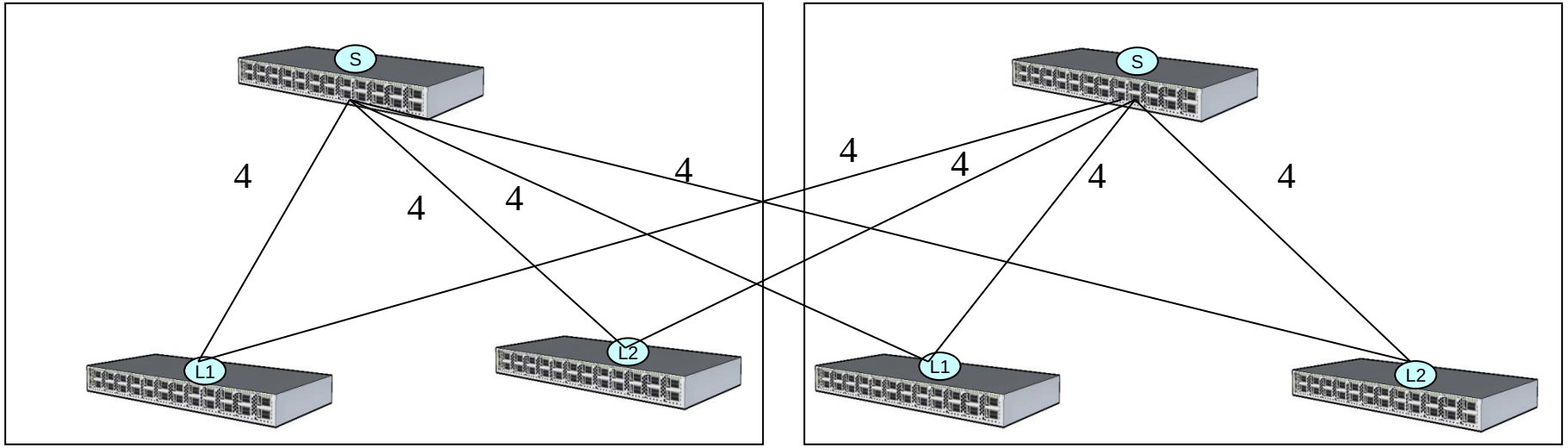
Multi-rack cabling

- Need to purchase IB spine switches (Expansion Switch Kit) when inter-connecting X4-2 and X3-8 racks
  - Except for quarter rack in most cases
- Need to purchase IB cables as spares kits no longer contains IB cables
- Same multi-rack cabling methodology as before
  - Same fat-tree topology to connect multiple racks
  - Can now connect more than 2 Qtr racks by simply purchasing the spine switches for the Quarter racks

Multi-rack Cabling – Fat Tree Topology

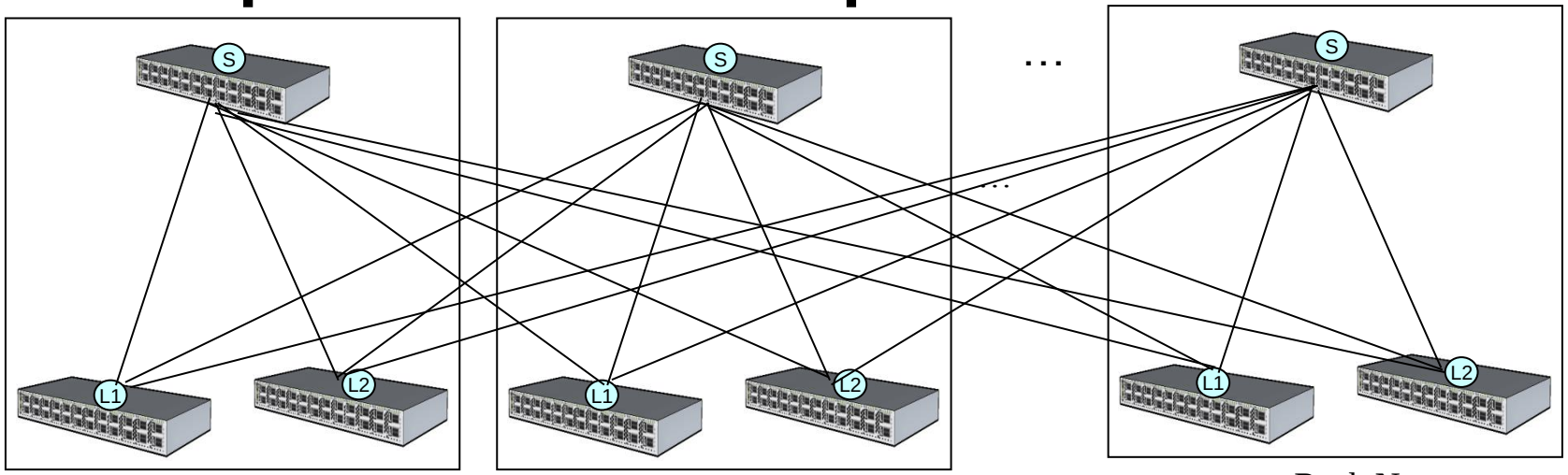
- Single InfiniBand Network
- Switch to a “Fat Tree” Topology
  - Valid up to 18 Racks with no external switches
  - Can go to 36 racks with 2 external switches
  - Every “leaf” switch inter-connected with every “spine” switch
  - “Leaf” switches not connected with other “leaf” switches
  - “Spine” switches not connected with other “spine” switches
  - Database and Exadata Server cabling unchanged.
  - Inter-rack cabling done at installation time
- Refer to the Database Machine Owner’s Guide for exact port mappings

Two Rack Case – Fat Tree Topology



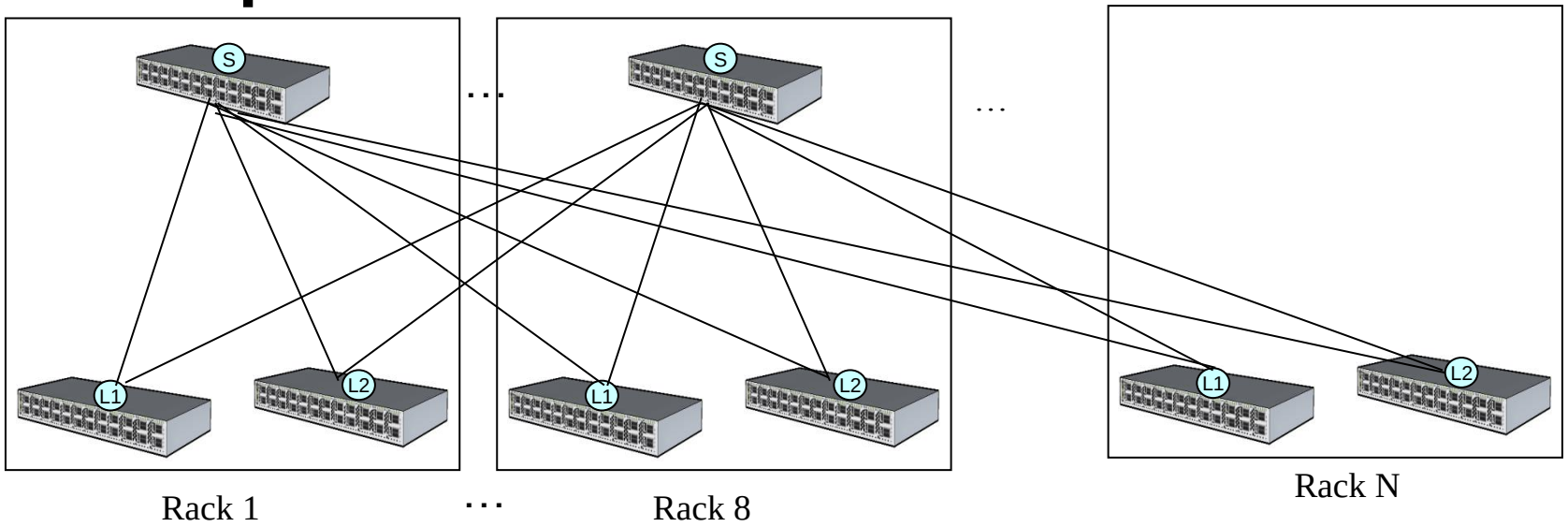
- Install Spine Switch in each rack, if no spine switch exists
- Remove inter-switch links between leaf switches (L1,L2) in each rack
- Connect each leaf switch to both the spine switches (4 links per pair)

Multiple Rack Case – Up to 8 Racks



- Install Spine Switch in each rack, if no spine switch exists
- Remove inter-switch links between leaf switches (L1,L2) in each rack
- Distribute 8 links from every “leaf” switch to every “spine” switch
  - Example – 4 rack case – 2 links from each leaf to each spine

Multiple Rack Case – 9 to 18 Racks

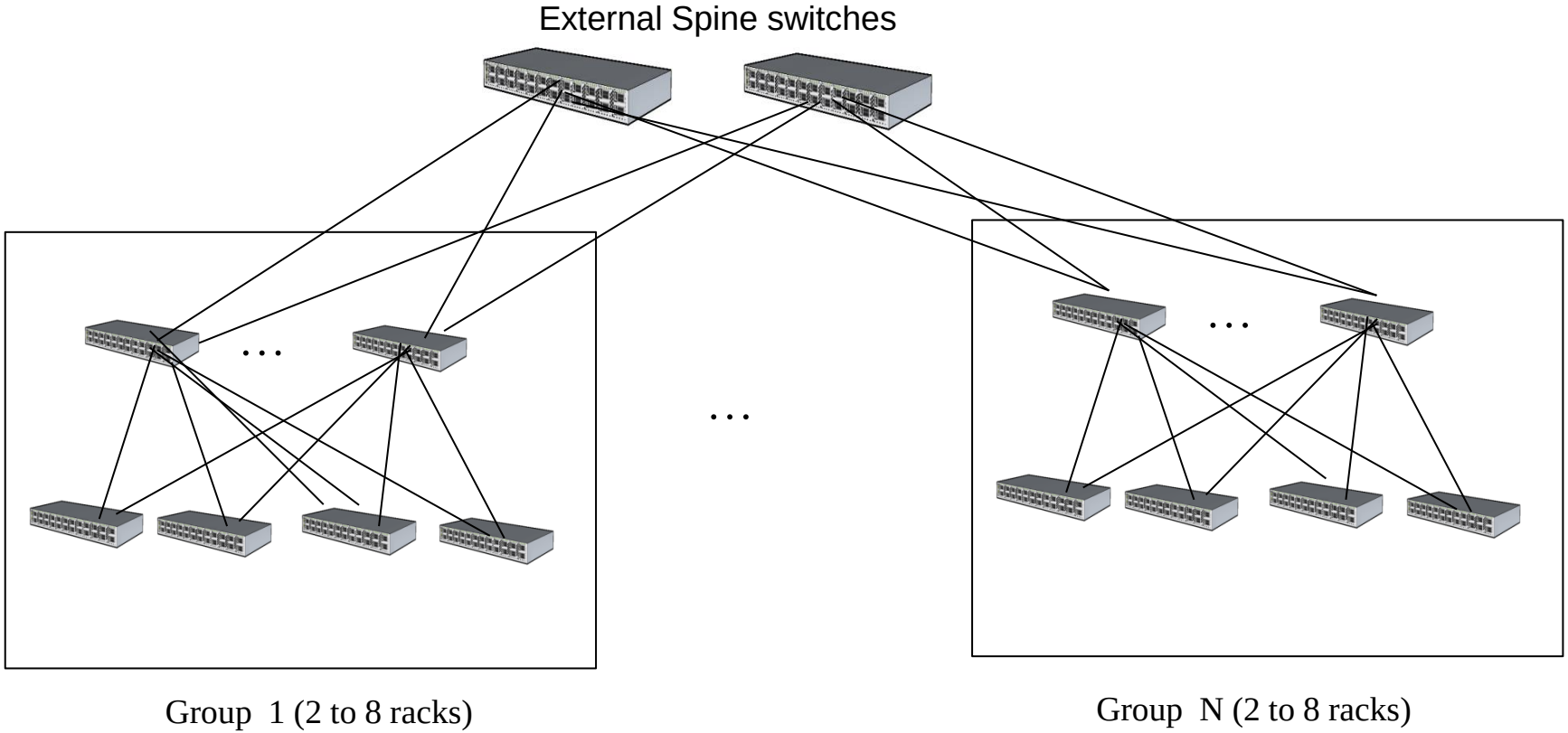


- Install Spine Switch in 1st eight racks, if no spine switch exists
- Remove inter-switch links between leaf switches (L1,L2) in each rack
- Distribute 8 links from every “leaf” switch to “spine” switch in 1st 8 racks

Scaling Out to 36 Racks

- Single InfiniBand Network
- Require 2 “external spine” 36-port InfiniBand switches
  - Add another level to the Fat Tree topology
- Divide Racks into groups each having 2-8 racks
  - Group related racks together – for example, racks that run a common database or are part of a common cluster
  - Inter-connect racks within each group using the two level Fat Tree topology
- Inter-connect the groups via the two “external spine” switches
  - Connect each “internal spine” switch in each rack to the two “external spine” switches
- Bandwidth from/to each group is 80 Gb/s * number of racks in the group
- Can be done online

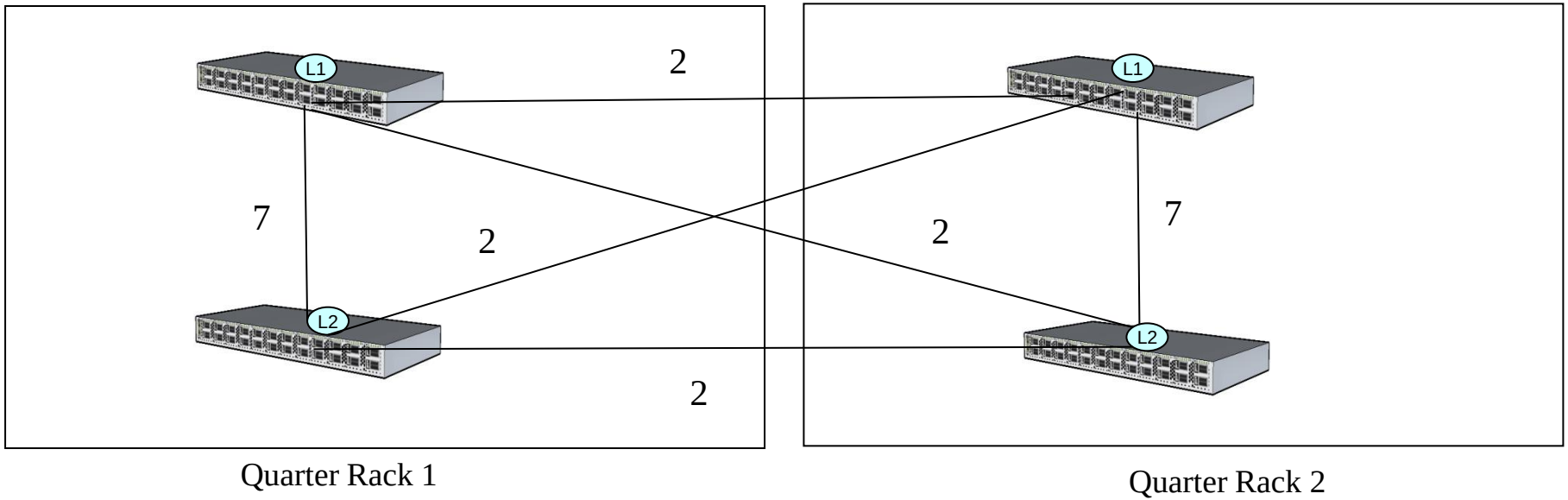
Multiple Rack Case – 9 to 36 Racks



Inter-connecting Quarter Racks

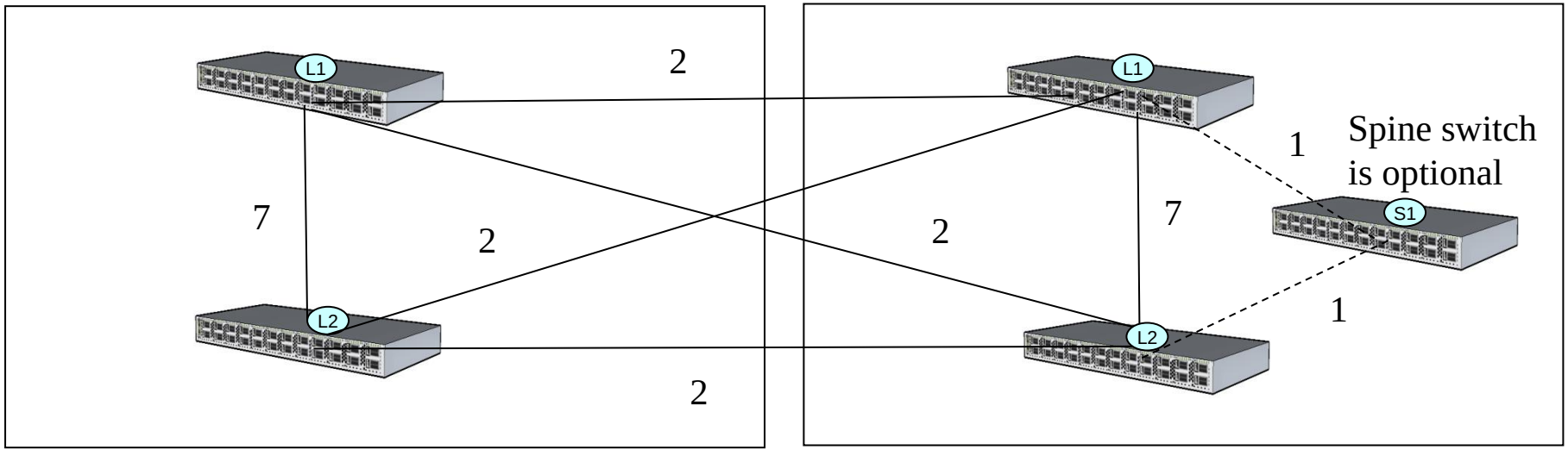
- Allow limited situations where a Quarter Rack can be inter-connected with other rack without requiring a spine switch in the Quarter rack
- Case 1 - Connect two Quarter Racks
  - The pre-requisite is that at least 4 ports reserved for external connectivity are open on each leaf switch
- Case 2 – Connect one Quarter Rack with one other Half or Full Rack
  - The pre-requisite is that at least 4 ports reserved for external connectivity are open on each leaf switch
- Case 3 – Connect one Quarter Rack to two or more racks (all of which are either Exadata Half or Full Racks)

Case 1 – Two Quarter Racks



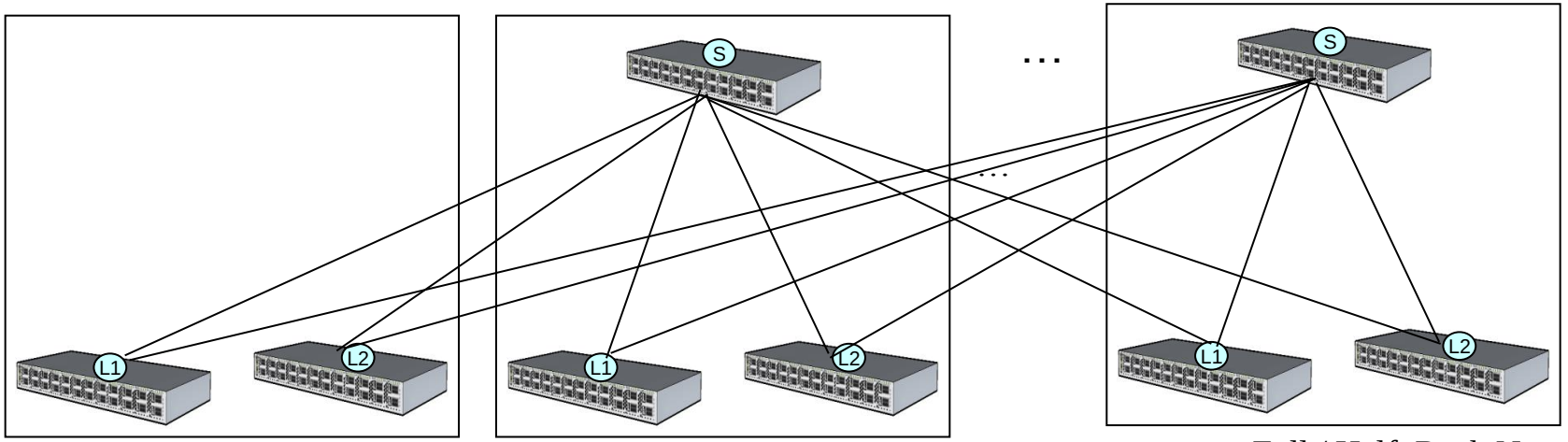
- Inter-connect the leaf switches between racks with 2 links each
  - Use the ports reserved for external connectivity

Case 2 –Quarter Rack with One Half or Full Rack



- Inter-connect the leaf switches between racks with 2 links each
  - Use the ports reserved for external connectivity

Case 3 – Quarter Rack with Two or more Racks



- Remove the 7 inter-switch links between leaf switches in the Quarter Rack
- Connect each leaf switch of the Quarter Rack to each spine switch of the other Full / Half Racks with 2 links each (if there are less than 4 racks) or 1 link each (if there are more than 4 racks)

InfiniBand Network – External Connectivity

- Six ports in each of the two leaf switches per Exadata rack available for external connectivity
  - Ports 5B,6A,6B,7A,7B, 12A per leaf switch (total 12 per rack)
  - Connecting to media servers for Tape backup
  - Data Loading
  - Client / Application Access
  - For HA connections, connect 1 port to one Leaf switch and another to the second Leaf switch

- Important – Never connect servers directly to spine switch

Expansion Switch Kit

- Purpose
  - Spine switch needed to inter-connect multiple rack.
  - One per rack being inter-connected.
  - Needed only for systems that did not ship with a spine switch (V2 Half rack, All X4-2 racks, X3-8 Full rack with X4-2 storage servers)

- Contains
  - Sun Datacenter 36-port Switch to be used as a spine switch in a rack

InfiniBand Network Addresses

	X4-2			X3-8	Storage Expansion X4-2		
	Full	Half	Quarter / Eighth	Full	Full	Half	Quarter
Ib0 and ib1 for Database Servers	16	8	4	8	-	-	-
Ib0 and ib1 for Exadata Storage Servers	28	14	6	28	36	18	8
Total	44	22	10	36	36	18	8

Ethernet Network

“Admin” Ethernet Network

Database Server (ILOM)

Database Server (“mgmt” port)

Exadata Storage Server (ILOM)

Exadata Storage Server (“mgmt” port)

IB Switch (Mgmt port)

Cisco switch

1

To Data Center Network

- “Admin” Ethernet switch pre-cabled connections
  - Connected to hardware mgmt ports
    - ILOM port on the Exadata Storage Servers and Database Servers
  - Connected to software mgmt ports
    - eth0 on Exadata storage server and eth0 on the Database Server
    - For software admin, root access, ..
- Connected to InfiniBand Switch management ports
- User can connect network ports in the PDUs directly to the data center network for monitoring

External Connectivity – Ethernet

- Management network
  - 1 port from “Admin” Ethernet switch
  - 2 from the PDUs
- Database / Client / Application Access
  - Need to connect a minimum of 1 GbE or 1 10GbE port per Database Server
  - X4-2 Database Server
    - 3 x 1GbE / 10 GbE Base-T Ethernet ports available
    - 2 x 10GbE Ethernet SFP+ ports available
  - X3-8 Database Server
    - 7 x 1GbE Ethernet ports available
    - 8 x 10GbE Ethernet SFP+ ports available

Additional Ethernet Network Requirements

- Additional Ethernet IP addresses needed if multiple clusters are deployed on the Database Machine Full Rack, Half Rack, or Quarter Rack
  - 3 IP addresses per additional cluster for SCAN addresses
- Static IP addressing only (no DHCP)
- All Ethernet IP addresses must be registered in DNS

Ethernet Network Addresses

	X4-2				X3-8	Storage Expansion X4-2		
	Full	Half	Qtr	Eighth	Full	Full	Half	Qtr
ILOM for Database Servers	8	4	2	2	2	-	-	-
ILOM for Exadata Cells	14	7	3	3	14	18	9	4
Eth0 (mgmt) for DB Servers	8	4	2	2	2	-	-	
Eth0 (mgmt) for Exadata Cells	14	7	3	3	14	18	9	4
Mgmt port for IB switches	2	2	2	2	3	3	3	2
IP address for Ethernet Switch	1	1	1	1	1	1	1	1
IP address for PDUs	2	2	2	2	2	2	2	2
Mgmt Subnet Total	**49**	**27**	**15**	**15**	**38**	**42**	**24**	**13**
Client Access Database Servers	8	4	2	2	2	-	-	
VIPs for Database Servers	8	4	2	2	2	-	-	-
SCAN Addresses (per Cluster)	3	3	3	3	3	-	-	-
Client Access Total	**19**	**11**	**7**	**7**	**7**	**-**	**-**	**-**
Total	**68**	**38**	**22**	**22**	**45**	**42**	**24**	**13**

Exadata Database Machine

- **The ultimate platform for all database workloads**
  - OLTP, Warehousing, Database as a Service
- **Most advanced <u>hardware</u>**
  - Fully scale-out servers and intelligent storage with unified InfiniBand connectivity and PCI flash
- **Most advanced <u>software</u>**
  - Database optimized compute, storage, and networking algorithms dramatically improve performance and cost
- **Standardized, optimized, hardened end-to-end**

Hardware and Software

ORACLE

Engineered to Work Together


ORACLE®

Exadata Database Machine X4-2 Half Rack

Pre-Configured for Extreme Performance

- 4 Xeon-based Dual-processor Database Servers
  - 96 cores (24 per server)
  - 1024 GB memory expandable to 2048 GB (256 GB per server expandable to 512 GB)
  - 10 Gig E-connectivity to Data Center
    - 40 x 10Gb E-ports (5 per server)
- 22.4 TB High Speed Flash
- 7 Exadata Storage Servers X4-2
  - All with High Performance 12 x 1.2 TB SAS disks

  OR
  - All with High Capacity 12 x 4 TB SAS disks
- 2 Sun Datacenter InfiniBand Switch 36
  - 36-port Managed QDR (40Gb/s) switch
- 1 “Admin” Ethernet switch
- Redundant Power Distributions Units (PDUs)

Upgradable to Full Rack

Exadata Database Machine X4-2 Quarter Rack

Pre-Configured for Extreme Performance

- 2 Xeon-based Dual-processor Database Servers
  - 48 cores (24 per server)
  - 512 GB memory expandable to 1024GB (256 GB per server expandable to 512 GB)
  - 10 Gig E-connectivity to Data Center
    - 10 x 10Gb E-ports (5 per server)
- 9.6 TB High Speed Flash
- 3 Exadata Storage Servers X4-2
  - All with High Performance 12 x 1.2 TB SAS disks

  OR
  - All with High Capacity 12 x 4TB SAS disks
- 2 Sun Datacenter InfiniBand Switch 36
  - 36-port Managed QDR (40Gb/s) switch
- 1 “Admin” Ethernet switch
- Redundant Power Distributions Units (PDUs)

Upgradable to Half Rack

Exadata Database Machine X4-2 Eighth Rack

Entry Level Configuration

- 2 Xeon-based Dual-processor Database Servers
  - 24 cores (12 per server)
  - 512 GB memory expandable to 1024GB (256 GB per server expandable to 512 GB)
  - 10 Gig E-connectivity to Data Center
    - 10 x 10Gb E-ports (5 per server)
- 4.8 TB High Speed Flash
- 3 Exadata Storage Servers X4-2
  - All with High Performance 6 x 1.2 TB SAS disks

  OR
  - All with High Capacity 6 x 4TB SAS disks
- 2 Sun Datacenter InfiniBand Switch 36
  - 36-port Managed QDR (40Gb/s) switch
- 1 “Admin” Ethernet switch
- Redundant Power Distributions Units (PDUs)

Upgradable via software to Quarter Rack

Exadata Storage Expansion X4-2 Half Rack

For additional storage – backups, historical data, unstructured data

- 9 Exadata Storage Servers X4-2
  - All with High Performance 12 x 1.2 TB SAS disks

  OR

  - All with High Capacity 12 x4TB SAS disks
- 28.8 TB High Speed Flash
- 3 Sun Datacenter InfiniBand Switch 36
  - 36-port Managed QDR (40Gb/s) switch
- 1 “Admin” Ethernet switch
- Redundant Power Distributions Units (PDUs)

Upgradable to Full Rack

Exadata Storage Expansion X4-2 Quarter Rack

For additional storage – backups, historical data, unstructured data

- 4 Exadata Storage Servers X4-2
  - All with High Performance 12 x 61.2 TB SAS disks

  OR
  - All with High Capacity 12 x 4 TB SAS disks
- 12.8 TB High Speed Flash
- 2 Sun Datacenter InfiniBand Switch 36
  - 36-port Managed QDR (40Gb/s) switch
- 1 “Admin” Ethernet switch
- Redundant Power Distributions Units (PDUs)

Upgradable to Quarter Rack