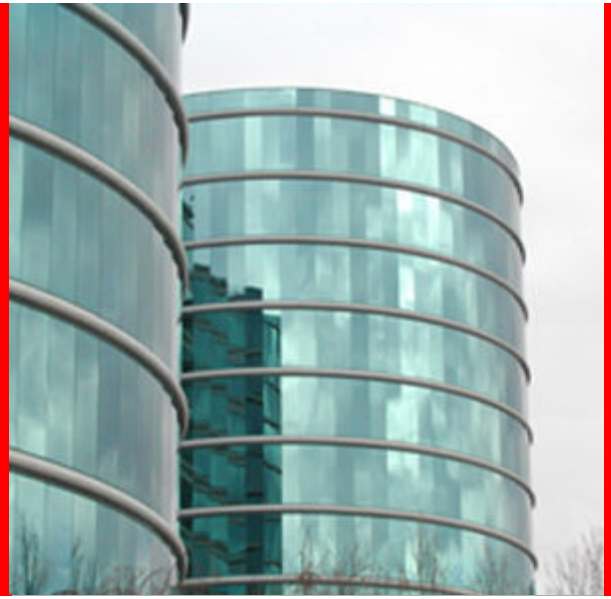


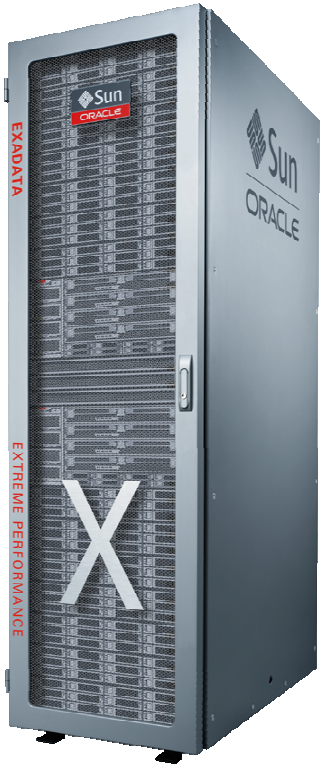
ORACLE®

SGSS Tech Day
Oracle Exadata Database Machine Overview

Copyright © 2010, Oracle Corporation and/or its affiliates

Exadata Database Machine

Best Platform to Run the Oracle Database

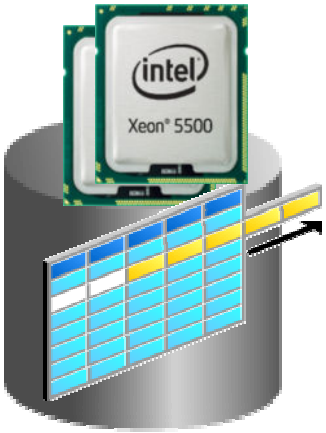


- Best Machine for **Data Warehousing**
- Best Machine for **OLTP**
- Best Machine for **Database Consolidation**

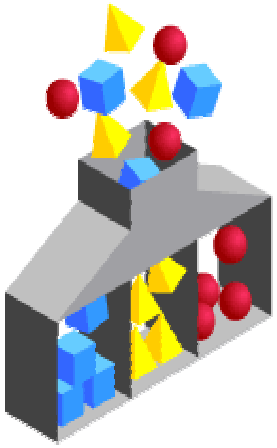
- **Unique** Architecture Makes it
  - Fastest, Lowest Cost

Copyright © 2010, Oracle Corporation and/or its affiliates

Keys to Speed and Cost Advantage



Exadata Intelligent Storage Grid



Exadata Hybrid Columnar Compression

Exadata Smart Flash Cache

© 2010 Oracle Corporation

Exadata in the Marketplace

Rapid adoption in all geographies and industries

Copyright © 2010, Oracle Corporation and/or its affiliates

Exadata Hardware Architecture

Scaleable Grid of industry standard servers for Compute and Storage

- Eliminates long-standing tradeoff between Scalability, Availability, Cost

Database Grid

- 8 Dual-processor x64 database servers

OR

- 2 Eight-processor x64 database servers

InfiniBand Network

- Redundant 40Gb/s switches
- Unified server & storage network

Intelligent Storage Grid

- 14 High-performance low-cost storage servers

- 100 TB **High Speed** disk, or 336 TB **High Capacity** disk
- **5.3 TB PCI Flash**
- Data mirrored across storage servers

Copyright © 2010, Oracle Corporation and/or its affiliates

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 config 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, partners

ORACLE

Copyright © 2010, Oracle Corporation and/or its affiliates

Exadata Storage Server Building Block

- **Hardware by Sun**
- **Software by Oracle**

- High-performance storage server built from industry standard components
- 12 disks - 600 GB 15000 RPM High Performance SAS or 2TB 7200 RPM High Capacity SAS
- 2 Six-Core Intel Xeon Processors (L5640)
- Dual ported 40 Gb/sec InfiniBand
- 4 x 96 GB Flash Cards
- Intelligent Exadata Storage Server Software

Copyright © 2010, Oracle Corporation and/or its affiliates

New - Exadata Database Machine X2-8 Full Rack

Extreme Performance for Consolidation, Large OLTP and DW

- 2 x64 Eight-processor Database servers (Sun Fire 4800)
  - High Core, High Memory Database Servers
  - 128 CPU cores (64 per server)
  - 2 TB (1 TB per server)
  - 10 GigE connectivity to Data Center
    - 16 x 10GbE ports (8 per server)
- 14 Exadata Storage Servers X2-2
  - All with High Performance 600GB SAS disks

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

Add more racks for additional scalability

ORACLE

Copyright © 2010 Oracle Corporation and/or its affiliates

Exadata Database Machine X2-2 Full Rack

Pre-Configured for Extreme Performance

- 8 x64 Dual-procesor Database Servers (Sun Fire X4170 M2)
  - 96 cores (12 per server)
  - 768 GB memory (96GB per server)
  - 10 GigE connectivity to Data Center
    - 16 x 10GbE ports (2 per server)
- 14 Exadata Storage Servers X2-2
  - All with High Performance 600GB SAS disks

  OR
  - All with High Capacity 2 TB SAS disks
- 3 Sun Datacenter InfiniBand Switch 36
  - 36-port Managed QDR (40Gb/s) switch
- 1 “Admin” Cisco Ethernet switch
- Keyboard, Video, Mouse (KVM) hardware
- Redundant Power Distributions Units (PDUs)

Add more racks for additional scalability

ORACLE

Copyright © 2010 Oracle Corporation and/or its affiliates

Start Small and Grow

Field Upgradeable

Quarter Rack **Half Rack** **Full Rack**

Balanced Incremental Scaling for OLTP and DW

Copyright © 2010, Oracle Corporation and/or its affiliates

Scale to 8 Racks by Just Adding Cables

Full Bandwidth and Redundancy

Copyright © 2010, Oracle Corporation and/or its affiliates

Complete Family Of Database Machines

For OLTP, Data Warehousing & Consolidated Workloads

Oracle Exadata X2-2

- Quarter, Half, Full and Multi-Racks

Oracle Exadata X2-8

- Full and Multi-Racks

Copyright © 2010, Oracle Corporation and/or its affiliates

Database Server Operating System Choices

- Two Operating System Choices on the database servers
  - Oracle Linux
  - Solaris 11 Express (x86) – Coming Soon
- Customers will choose their preferred Database Server OS at installation time
- Exadata Storage Servers will continue to be Oracle Linux

ORACLE

Copyright © 2010, Oracle Corporation and/or its affiliates

Exadata Product Capacity

		X2-8 Full Rack	X2-2 Full Rack	X2-2 Half Rack	X2-2 Quarter Rack
Raw Disk[1]	High Perf Disk	100 TB	100 TB	50 TB	21 TB
	High Cap Disk	336 TB	336 TB	168 TB	72 TB
Raw Flash[1]		5.3 TB	5.3 TB	2.6 TB	1.1 TB
User Data[2] (assuming no compression)	High Perf Disk	28 TB	28 TB	14 TB	6 TB
	High Cap Disk	100 TB	100 TB	50 TB	21 TB

1 – Raw capacity calculated using 1 GB = 1000 x 1000 x 1000 bytes and 1 TB = 1000 x 1000 x 1000 x 1000 bytes.

2 - User Data: Actual space for end-user data, computed after single mirroring (ASM normal redundancy) and after allowing space for database structures such as temp, logs, undo, and indexes. Actual user data capacity varies by application. User Data capacity calculated using 1 TB = 1024 * 1024 * 1024 * 1024 bytes.

ORACLE

Copyright © 2010, Oracle Corporation and/or its affiliates

Exadata Product Performance

		X2-8 Full Rack	X2-2 Full Rack	X2-2 Half Rack	X2-2 Quarter Rack
Raw Disk Data Bandwidth[1,4]	High Perf Disk	25 GB/s	25 GB/s	12.5 GB/s	5.4 GB/s
	High Cap Disk	14 GB/s	14 GB/s	7 GB/s	3 GB/s
Raw Flash Data Bandwidth[1,4]		50 GB/s	50 GB/s	25 GB/s	11 GB/s
Disk IOPS[3,4]	High Perf Disk	50,000	50,000	25,000	10,800
	High Cap Disk	25,000	25,000	12,500	5,400
Flash IOPS[3,4]		1,000,000	1,000,000	500,000	225,000
Data Load Rate[4]		5 TB/hr	5 TB/hr	2.5 TB/hr	1 TB/hr

1 – Bandwidth is peak physical disk scan bandwidth, assuming no compression.

2 - Max User Data Bandwidth assumes scanned data is compressed by factor of 10 and is on Flash.

3 – IOPs – Based on IO requests of size 8K

4 - Actual performance will vary by application.

ORACLE

Copyright © 2010, Oracle Corporation and/or its affiliates

Exadata Storage Software Unique Features

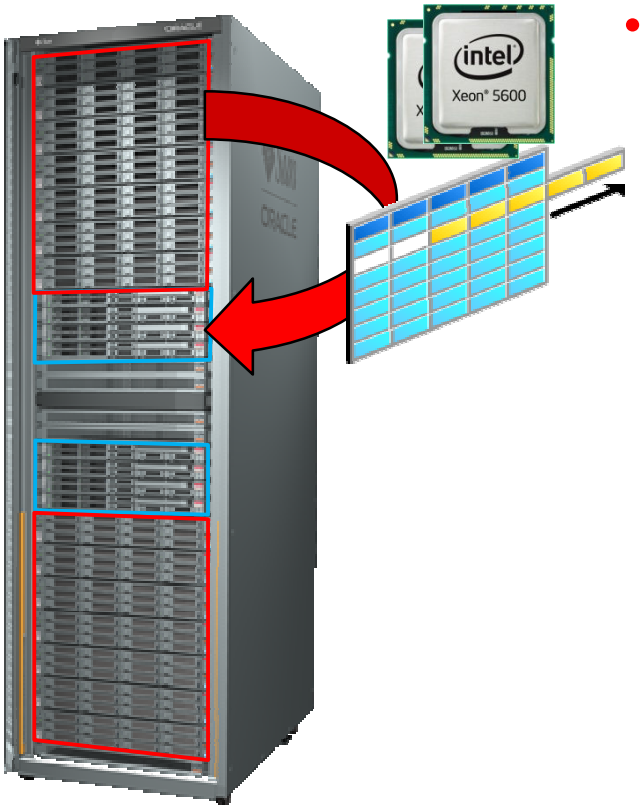
- Exadata Smart Scans
  - 10X or greater reduction in data sent to database servers
- Exadata Storage Indexes
  - Eliminate unnecessary I/Os
- Hybrid Columnar Compression
  - Efficient compression increases effective storage capacity and increases user data scan bandwidths by a factor of up to 10X
- Exadata Smart Flash Cache
  - Breaks random I/O bottleneck by increasing IOPs by up to 20X
  - Doubles user data scan bandwidths
- I/O Resource Manager (IORM)
  - Enables storage grid by prioritizing I/Os to ensure predictable performance
- Quality of Service (QoS)
  - Actively meet and maintain SLAs
  - Memory Guard to protect existing current transactions from memory-based failures

ORACLE

Copyright © 2010, Oracle Corporation and/or its affiliates

Exadata Intelligent Storage Grid

Most Scalable Data Processing

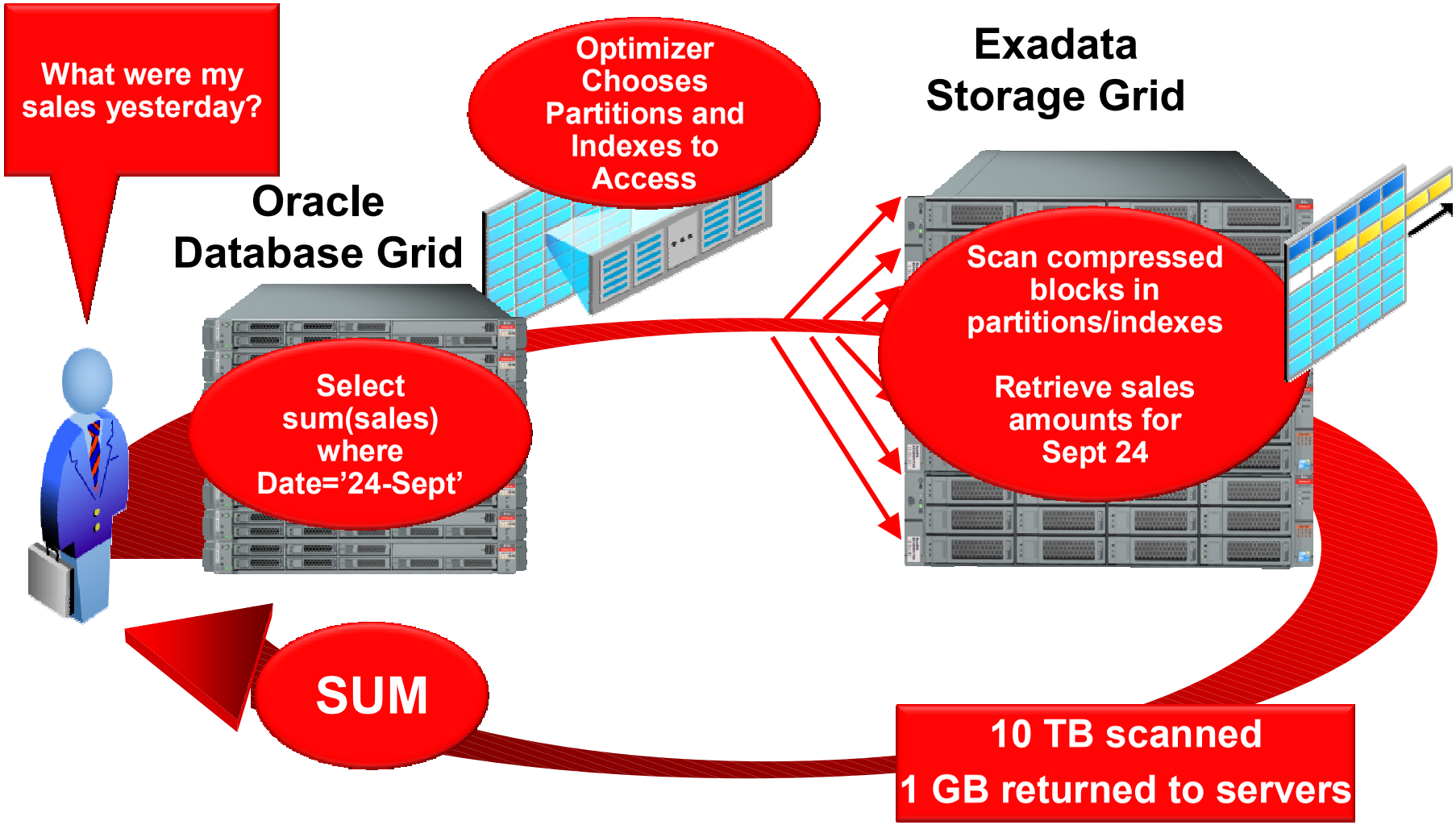


- **Data Intensive processing runs in Exadata Storage Grid**
  - Filter rows and columns as data streams from disks (168 Intel Cores)
  - Example: How much product X sold last quarter
  - Exadata Storage Reads 10TB from disk
  - Exadata Storage Filters rows by Product & Date
  - Sends 100GB of matching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

ORACLE

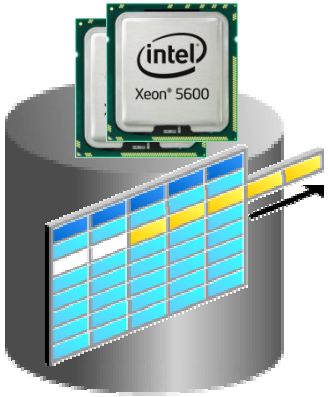
Copyright © 2010, Oracle Corporation and/or its affiliates

Simple Query Example



Copyright © 2010, Oracle Corporation and/or its affiliates

Exadata Intelligent Storage

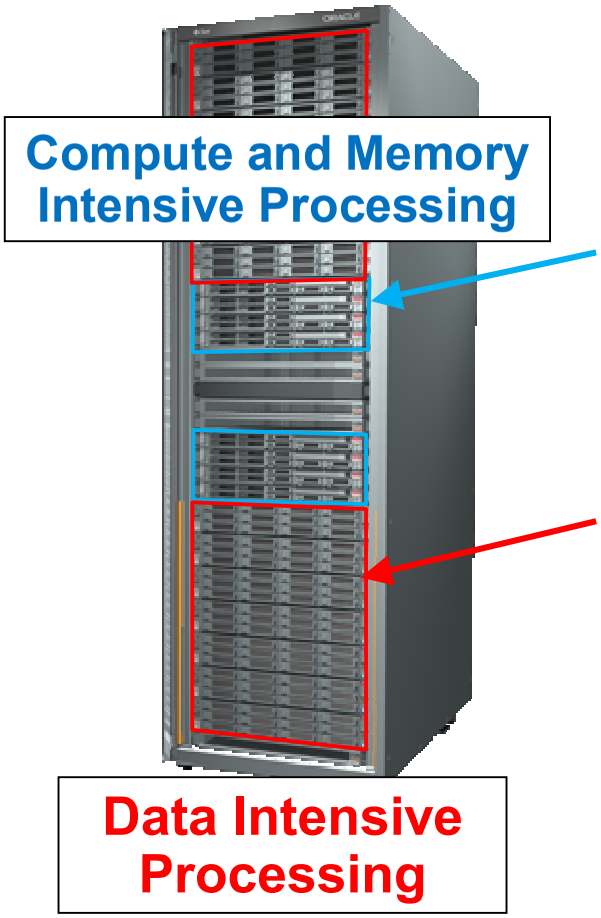


Exadata Intelligent Storage Grid

- 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

Copyright © 2010, Oracle Corporation and/or its affiliates

Exadata is Smart Storage



- Storage Server is smart storage, not a DB node
  - Storage remains an independent tier
- **Database Servers**
  - Perform complex database processing such as joins, aggregation, etc.
- **Exadata Storage Servers**
  - 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

Copyright © 2010, Oracle Corporation and/or its affiliates

Exadata Storage Index

Transparent I/O Elimination with No Overhead

Table **Index**

A	B	C	D
	1		
	3		
	5		
	5		
	8		
	3		

Min B = 1
Max B =5

Min B = 3
Max B =8

- Exadata Storage Indexes maintain summary information about table data in memory
  - Store MIN and MAX values of columns
  - Typically one index entry for every MB of disk
- Eliminates disk I/Os if MIN and MAX can never match "where" clause of a query
- Completely **automatic and transparent**

Select * from Table where B<2 - Only first set of rows can match

ORACLE

Copyright © 2010, Oracle Corporation and/or its affiliates

Storage Index with Partitions Example

Orders Table

Order#	Order_Date Partitioning Column	Ship_Date	Item
1	2007	2007	
2	2008	2008	
3	2009	2009	

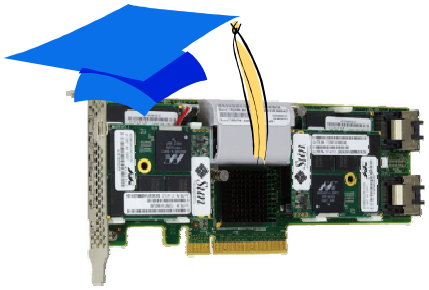
- Queries on Ship_Date do not benefit from Order_Date partitioning
  - However Ship_date and Order# are highly correlated with Order_Date
  - e.g. Ship dates are usually near Order_Dates and are never less
- Storage index provides partition pruning like performance for queries on Ship_Date and Order#
  - Takes advantage of ordering created by partitioning or sorted loading

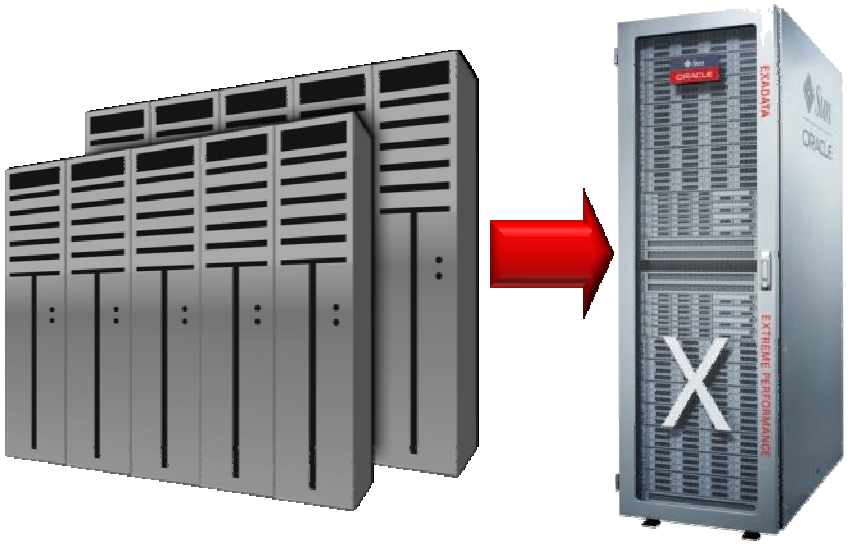
ORACLE

Copyright © 2010, Oracle Corporation and/or its affiliates

Exadata Smart Flash Cache

Extreme Performance OLTP & DW





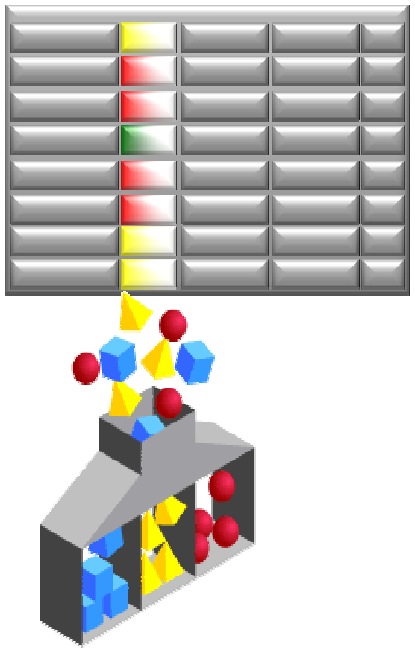
5X More I/Os than 1000 Disk Enterprise Storage Array

- Exadata has **5 TB** of flash
  - **56 Flash PCI cards avoid disk controller bottlenecks**
- **Intelligently manages flash**
  - Smart Flash Cache holds hot data
    - Avoids large scan wipe-outs of cache
  - **Gives speed of flash, cost of disk**
- Exadata flash cache achieves:
  - Over **1 million IO/sec from SQL** (8K)
  - Sub-millisecond response times

Copyright © 2010, Oracle Corporation and/or its affiliates

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

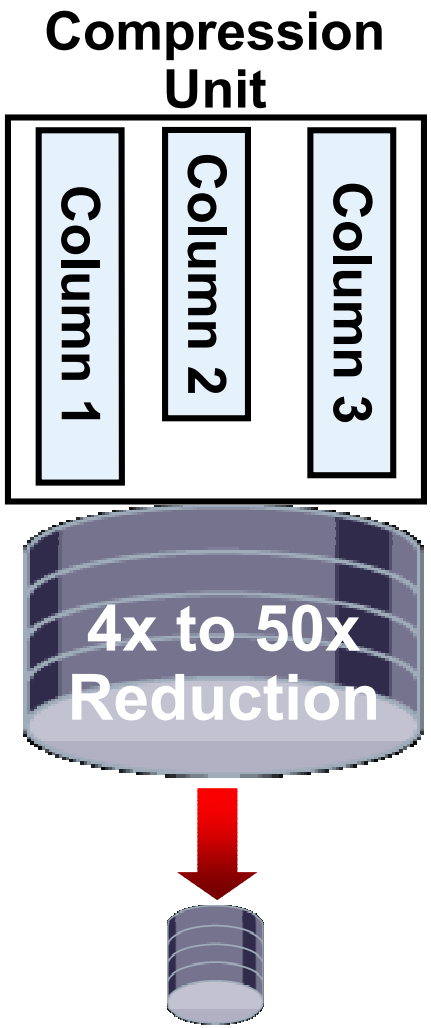
Benefits Multiply

ORACLE

Copyright © 2010, Oracle Corporation and/or its affiliates

Exadata Hybrid Columnar Compression

How it works

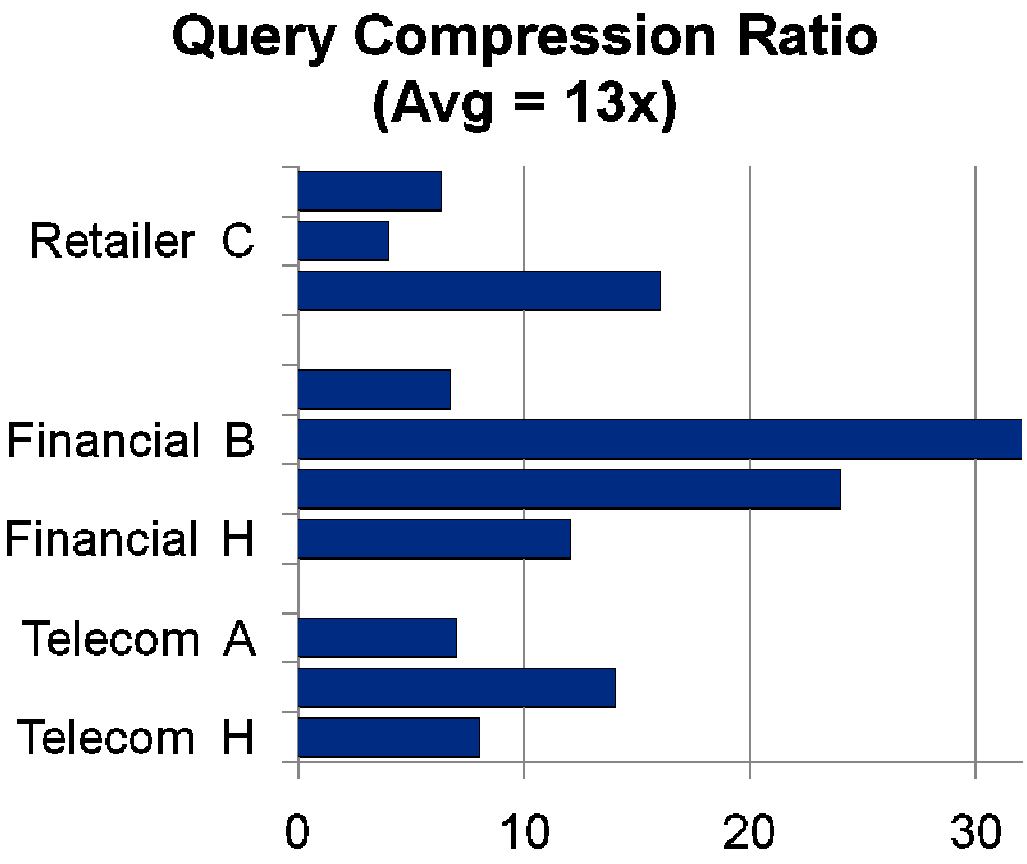


- Tables are organized into sets of a few thousand rows
  - Compression Units (CUs)
- Within CU, data is organized by column, then compressed
  - Column organization brings similar values close together, enhancing compression
- Useful for data that is bulk loaded and queried
  - Update activity is light
- Compared to best conventional algorithms – Gzip, Bzip2
  - Typically 2X the compression, 10X the performance
- Exadata servers offload filtering, projection, etc. for scans on compressed data
  - Indexed accesses return compressed blocks to database so buffer cache benefits from compression

ORACLE

Copyright © 2010, Oracle Corporation and/or its affiliates

Compression Ratio of Real-World Data



- Compression Ratio varies by customer and table
- Trials were run on largest table at 10 ultra large companies
  - Average revenue > $60 BB
- Average Query Compression ratio was 13x
  - On top of Oracle's already highly efficient format

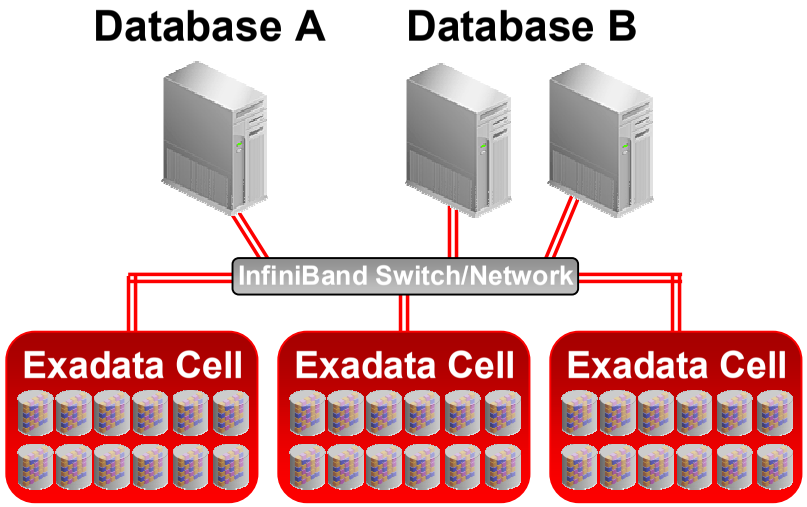
ORACLE

Copyright © 2010, Oracle Corporation and/or its affiliates

Exadata I/O Resource Management

Mixed Workloads and Multi-Database Environment

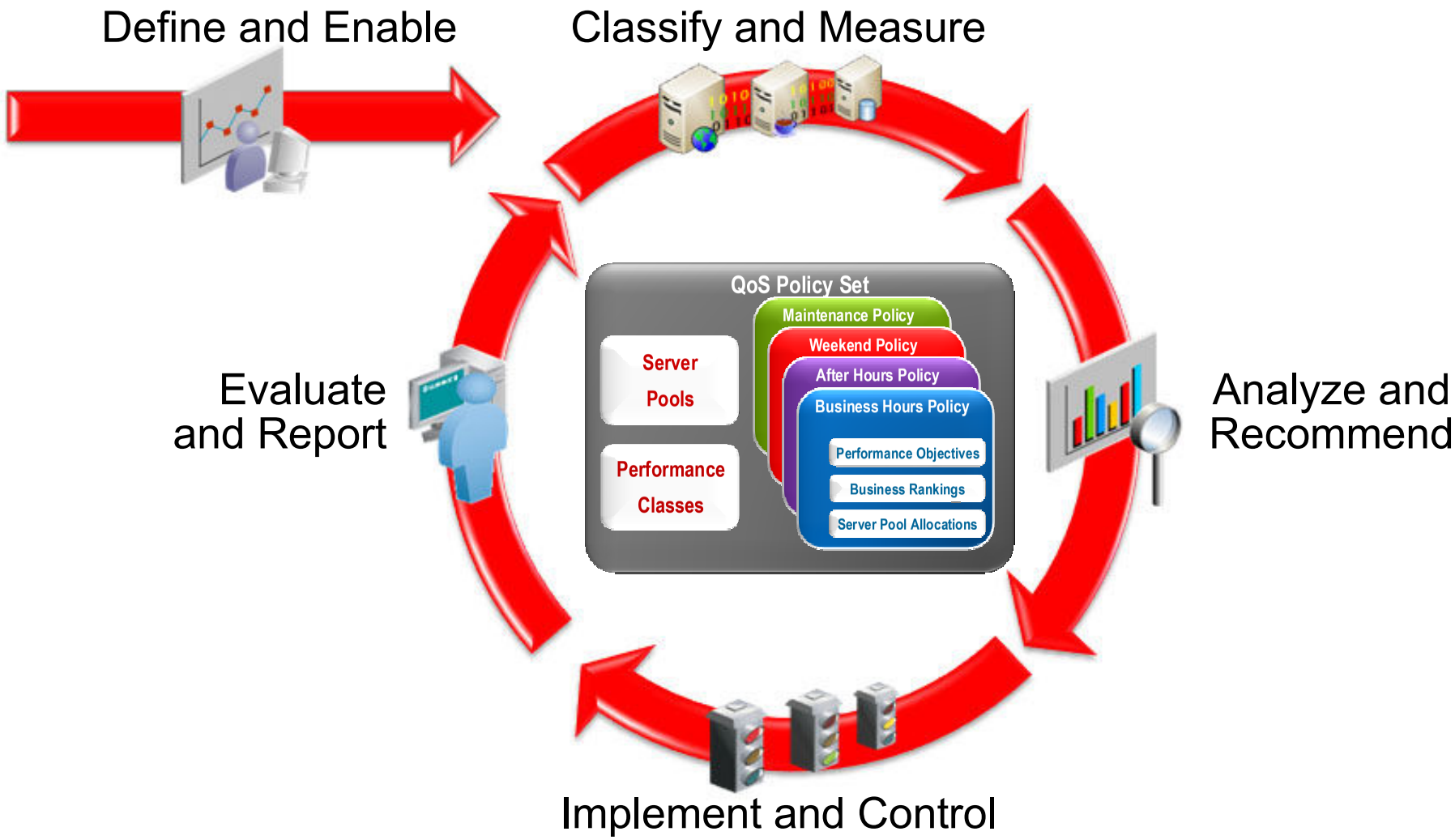
- Ensure different databases are allocated the correct relative amount of I/O bandwidth
  - Database A: 33% I/O resources
  - Database B: 67% I/O resources
- Ensure different users and tasks within a database are allocated the correct relative amount of I/O bandwidth
  - Database A:
    - Reporting: 60% of I/O resources
    - ETL: 40% of I/O resources
  - Database B:
    - Interactive: 30% of I/O resources
    - Batch: 70% of I/O resources



Copyright © 2010, Oracle Corporation and/or its affiliates

Quality of Service Management

The Oracle Approach in Action



Copyright © 2010, Oracle Corporation and/or its affiliates

Integrated, Task-Based Management

New EM Quality of Service Management Tool

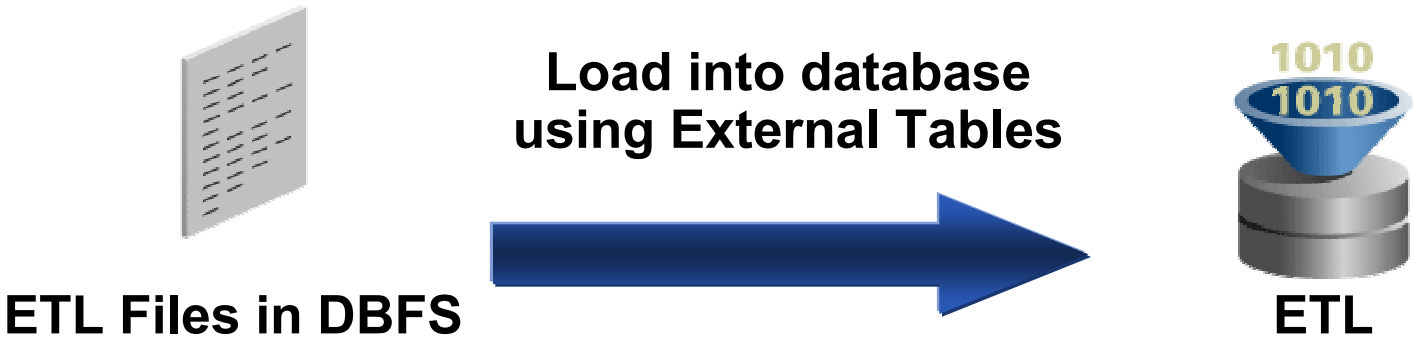
- Define and Enable Policies
  - ✓ Policy Definition Wizard
- Monitor Application Performance
  - ✓ Bottleneck Dashboard and Alerts
- Reallocate for Resource Problems
  - ✓ Expert-system Recommendations
- Audit Application Performance
  - ✓ Historical Metrics Views

ORACLE

Copyright © 2010, Oracle Corporation and/or its affiliates

DBFS - Scalable Shared File System

- Database Machine comes with DBFS shared Linux file system
  - Shared storage for ETL staging, scripts, reports and other application files
- Files stored as SecureFile LOBs in database tables stored in Exadata
  - Protected like any DB data – mirroring, DataGuard, Flashback, etc.
- 5 to 7 GB/sec file system I/O throughput



More File Throughput than High-End NAS Filer

Copyright © 2010, Oracle Corporation and/or its affiliates

Exadata Storage Management & Administration

- Enterprise Manager
  - Manage & administer Database and ASM
  - Monitor the Exadata Database Machine Hardware
    - Plug-ins to monitor the Exadata Storage Cells and other components in the Database Machine
- 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

ORACLE

Copyright © 2010, Oracle Corporation and/or its affiliates

First Secure Database Machine

- Moves decryption from software to hardware
  - Over 5x faster
- Near zero overhead for fully encrypted database
- Queries decrypt data at hundreds of Gigabytes/second

Copyright © 2010, Oracle Corporation and/or its affiliates

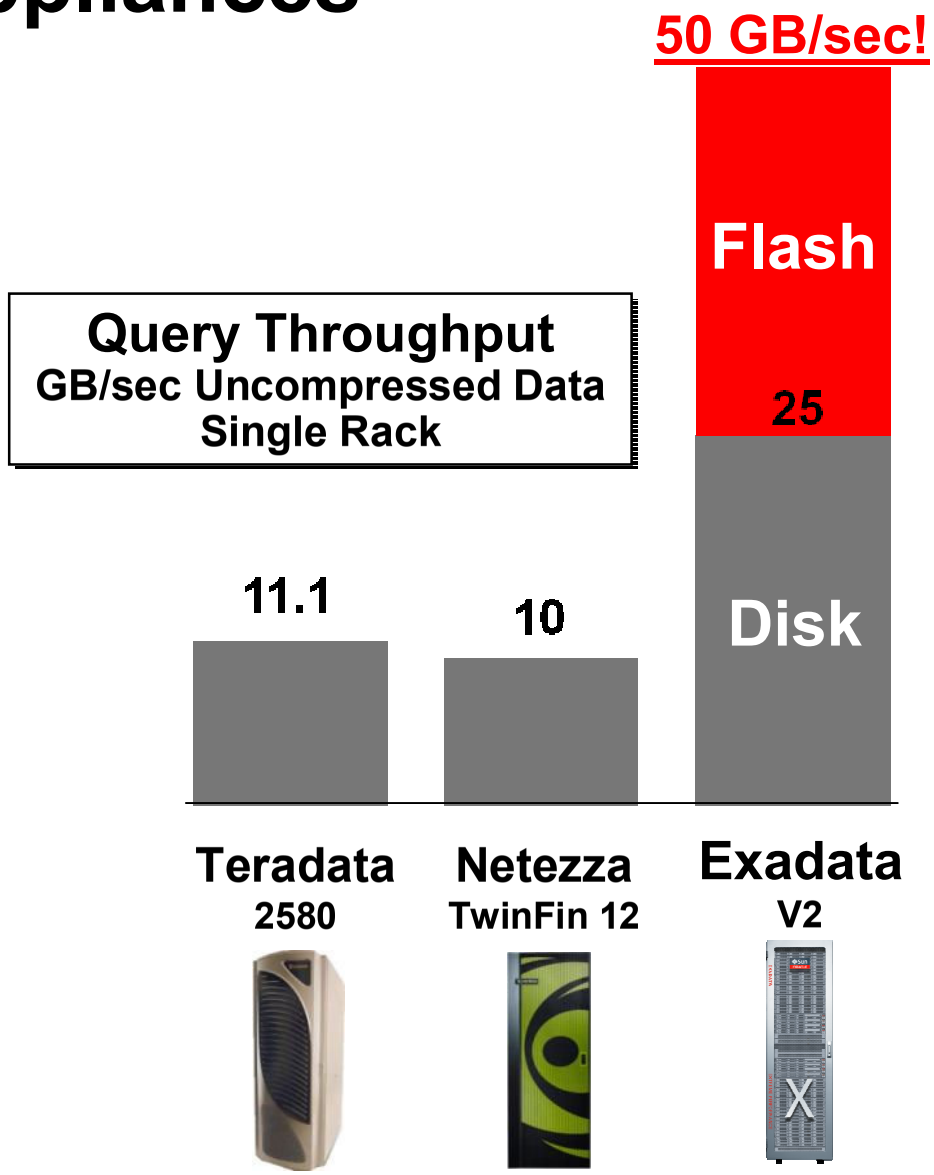
Best Machine for Database Consolidation

- Exadata serves as farm/cloud for databases
  - Large memory enables many databases to be consolidated
  - Extreme performance for complex workloads that mix OLTP, DW, batch, reporting
  - I/O and CPU resource management isolates workloads

Copyright © 2010, Oracle Corporation and/or its affiliates

Faster than DW Appliances

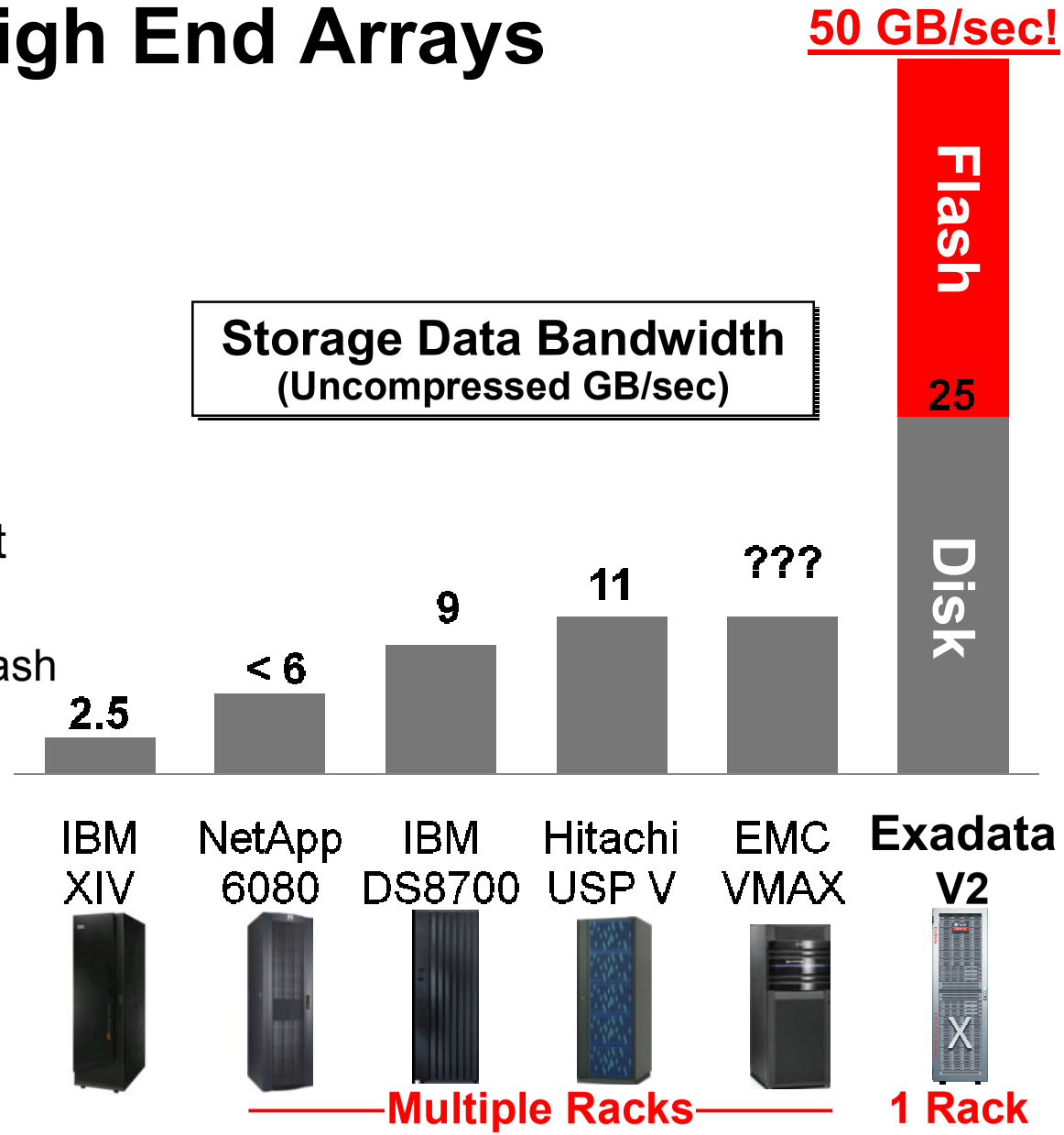
- Faster Throughput from Disk
- Much Faster with Flash
- **50 TB of data fits in Flash**
  - Using 10x Query Compression
- Effective Query Throughput on compressed data is even higher
  - **Hundreds of GB/sec**



Copyright © 2010, Oracle Corporation and/or its affiliates

Faster than High End Arrays

- Exadata is fastest
  - and scales with more racks
- Storage Arrays already can't deliver disk bandwidth
  - **No** extra bandwidth from Flash
  - No CPU offload
  - No Columnar Compression
  - No InfiniBand



Copyright © 2010, Oracle Corporation and/or its affiliates

More Data Capacity

Exadata V2
10x Compression

Netezza TwinFin
2x to 4x Compression

EMC VMAX
3x Oracle Compression

Teradata 2580
1.3x Compression

Systems with Equal User Data

All with Largest Disks, Best Compression

ORACLE

Copyright © 2010, Oracle Corporation and/or its affiliates

Exadata is Faster and 6X Lower Cost than IBM's Best - Power 795

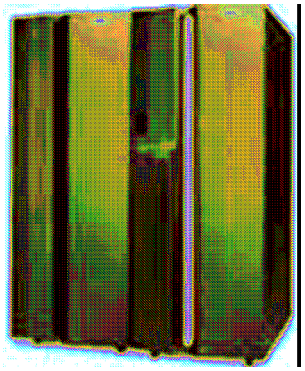
2 Exadata X2-8

$**3**,000,000

- More **CPU Cores**
- More I/O Performance
- Same Storage Capacity
  - Excluding Exadata compression!
- HA Configuration

IBM P795 + 4 DS8700s with Flash

$**18**,860,000

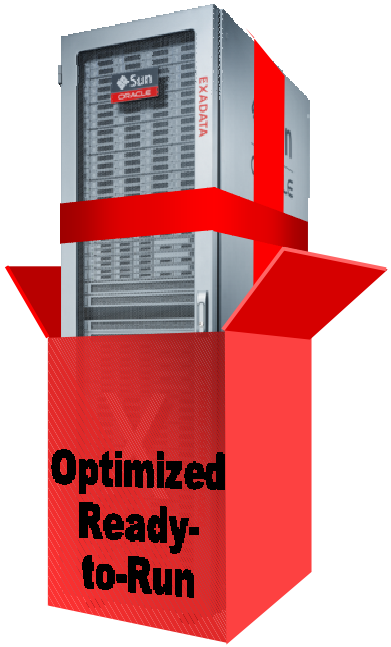


Comparison of Hardware Prices

Copyright © 2010, Oracle Corporation and/or its affiliates

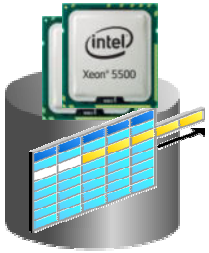
Exadata Database Machine

Ideal Database Platform

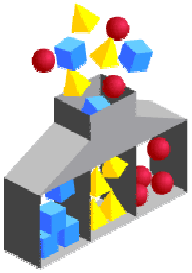


- **Best for Data Warehousing**
- **Best for OLTP**
- **Best for Database Consolidation**

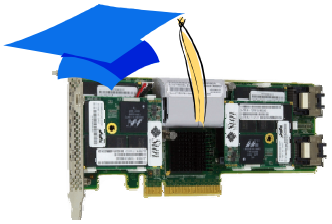
<u>Fastest, Lowest Cost</u>



Intelligent Storage Grid



Hybrid Columnar Compression



Smart Flash Cache

Hardware Decryption

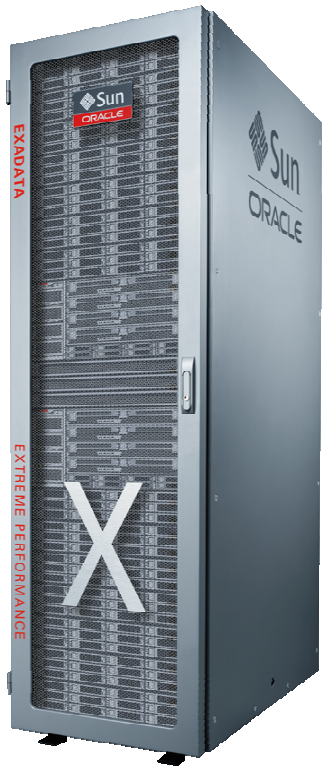
Business answers in seconds, not hours

ORACLE

Copyright © 2010, Oracle Corporation and/or its affiliates

Exadata Value Proposition

- ✓Large Volume
- ✓Structured and Unstructured Data
- ✓Mixed Workload
- ✓Scale on Demand
- ✓Roadmap for growth
- ✓Price-Performance
- ✓Differentiate with competition



Fastest, Lowest Cost

Business answers in seconds, not hours

Copyright © 2010, Oracle Corporation and/or its affiliates

Copyright © 2010, Oracle Corporation and/or its affiliates