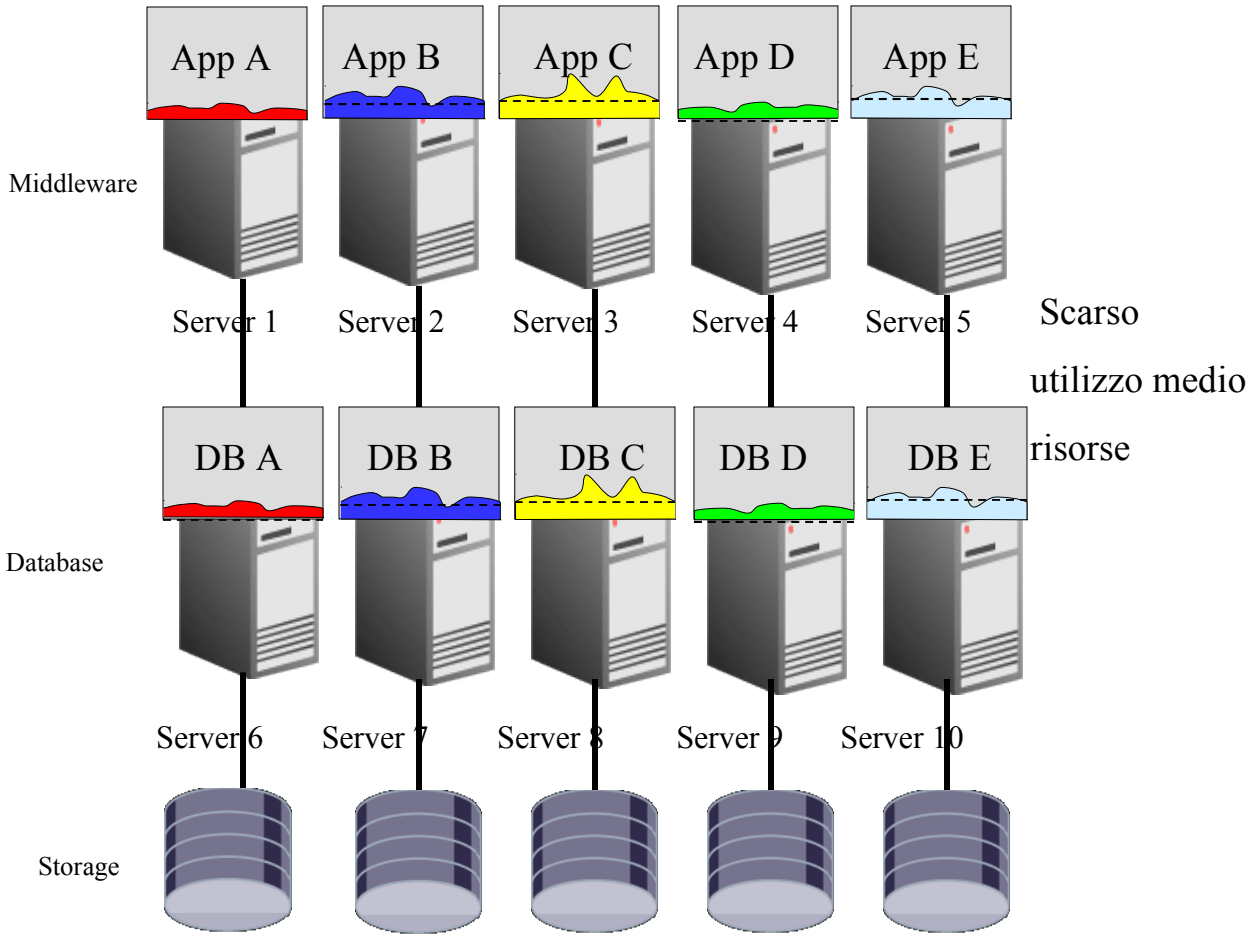
ORACLE

ORACLE THIRTY YEARS 30

Risparmiare consolidando con Sun Oracle Database Machine

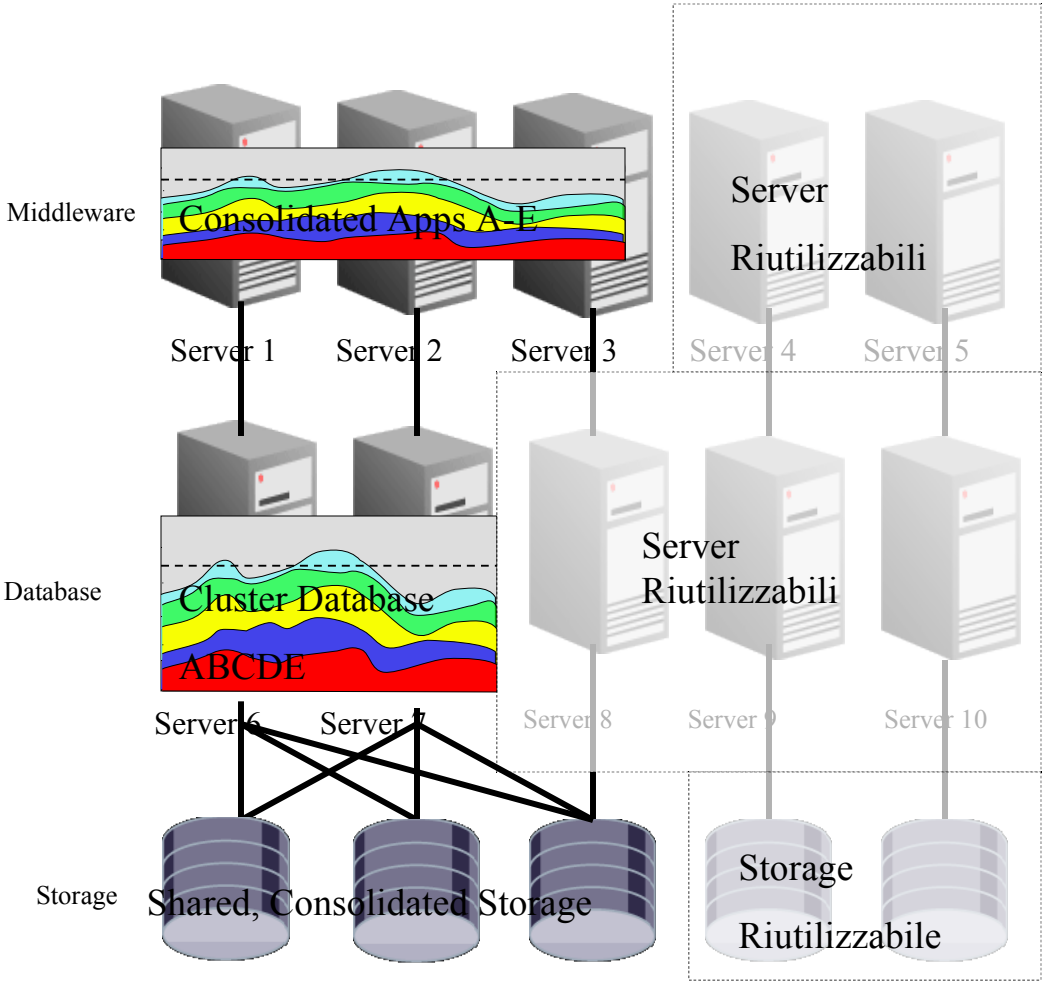
Alessandro Bracchini
Sales Consulting – Oracle Italia

Infrastruttura tradizionale: costosa ed inefficiente



- Architettura a “Silos” lascia sottoutilizzate per diverso tempo le risorse HW disponibili
- Alta Disponibilita’ ottenuta raddoppiando Server – Alti costi
- Non ecologica: richiede piu’ elettricita’, spazio, condizionamento, …

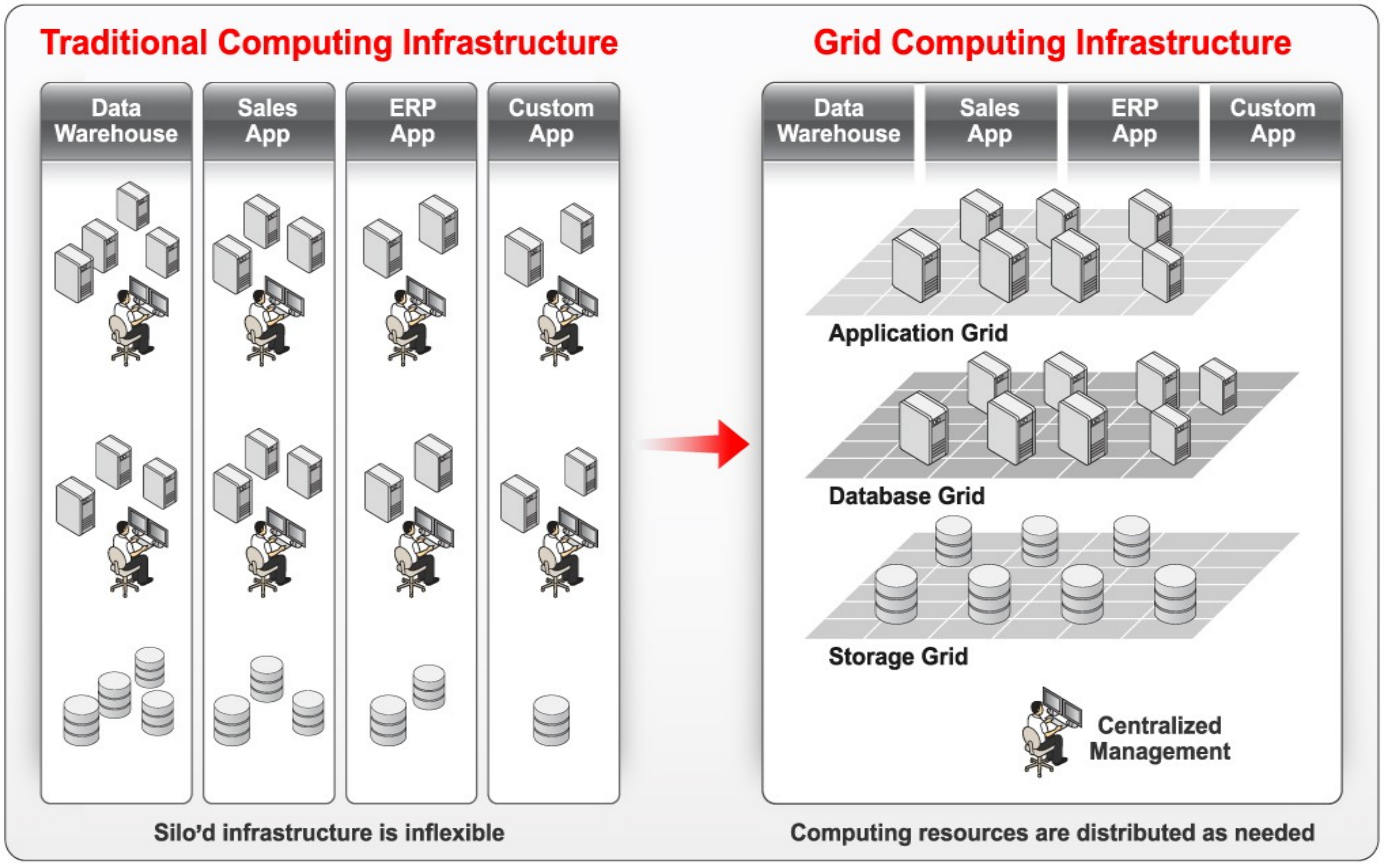
Consolidamento con il Grid Computing



- E' possibile far coesistere su risorse condivise virtuali carichi di lavoro complementari
- Si ottimizza il livello di utilizzo medio dei Server risparmiando sugli acquisti HW
- L'architettura e' intrinsecamente affidabile
- Maggiore efficienza = Green Computing

Grid Computing

Consolidamento e Virtualizzazione risorse IT



- Dimensionati per picchi di carico
- Rigida
- Complessa da gestire

- Gruppi di risorse condivise
- Ridistribuzione risorse se necessario
- Gestione centralizzata

Best Machine for Consolidating Databases

- Più database su uno stesso server
- Un unico storage per tutti i database
- Architettura RAC per l'alta affidabilità e le prestazioni
- Singoli DB possono usare l'*instance caging*

Sun Oracle Database Machine

- Ottimizzata per il Data Warehousing
  - Risposte veloci per qualunque tipo di query su grandi moli di dati di dettaglio
    - Elaborazione delle query direttamente nello storage con una banda di 21 GB/sec per rack, arrivando a 50 GB/sec con le momorie Flash
- Ottimizzata per l'Oltp ed il Warehousing operazionale
  - Query brevi e circostanziate e di media durata tipiche dei sistemi operazionali
    - In-memory parallel execution
    - Indicizzazione automatica avanzata ed utilizzo trasparente delle memorie Flash
- Ottimizzata per il caricamento veloce di file
  - "Multi-version read consistency" e capacità di caricamento di 5TB/hr

© 2009 Oracle Corporation

Sun Oracle Database Machine

Extreme Performance

Oracle Database Server Grid

- Millions of transactions per minute
- Tens of millions of queries per minute
- Billions of rows per minute

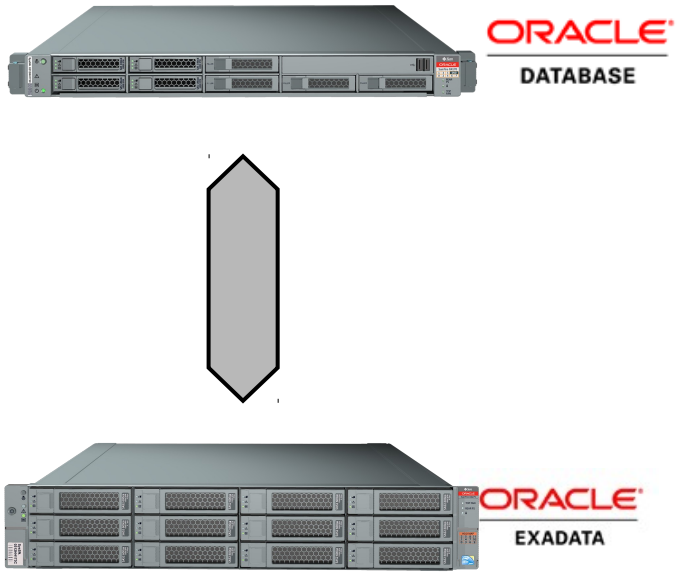
InfiniBand Network

- 880 Gb/sec aggregate throughput

Exadata Storage Server Grid

- 21 GB/sec disk bandwidth
- 50 GB/sec flash bandwidth
- 1 million I/Os per second

L’architettura di base di Exadata



- Database Server
  - Potenza di calcolo e disponibilità di memoria per processare tutte le operazioni sui dati
  - Tutte le operazioni eseguite in parallelo
- Connessioni Infiniband
- Exadata Storage Server
  - Esecuzione dell’I/O intensivo nello storage con selezione intelligente dei soli dati necessari alla query

I Database Server e gli Exadata storage lavorano congiuntamente per alle query SQL
La cella Exadata è uno storage intelligente, non un nodo di Database

© 2009 Oracle Corporation

Sun Oracle Database Machine

Get on the Grid Faster

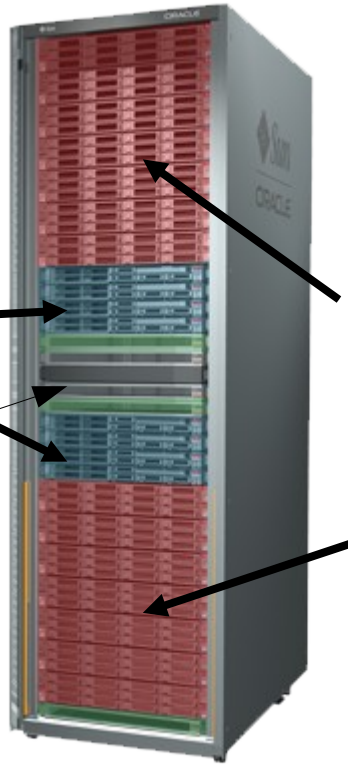
- Highest performance, lowest cost, fault tolerant, scalable on demand
- Database Machine is an engineered, optimized, standardized, and tested grid for Oracle database – with intelligent storage

Oracle Database Server Grid

- 8 compute servers
- 2 Intel quad-core Xeons each (64 cores)
- 576 GB DRAM

InfiniBand Network

- 40 Gb/sec unified server and storage network
- Fault Tolerant



Exadata Storage Server Grid

- 14 storage servers
- 100 TB raw SAS disk storage

or

336 TB raw SATA disk storage

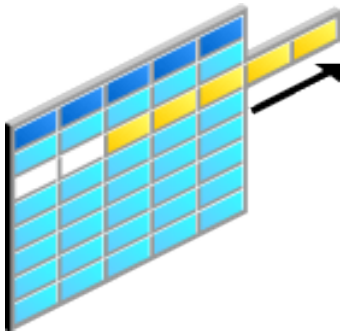
- **5TB flash storage!**
- Offload queries into storage

Exadata Software Features

- Exadata Smart Scans
  - 10X or greater reduction in data sent to database servers
- OLTP Compression & Hybrid Columnar Compression (HCC)
  - Efficient compression increases effective storage capacity and increases user data scan bandwidths by a factor of 10X
- Exadata Smart Flash Cache
  - Breaks random I/O bottleneck by increasing IOPs by 20X
  - Doubles user data scan bandwidths
- I/O Resource Manager (IORM)
  - Enables storage grid by prioritizing I/Os to ensure predictable performance
- Inter-leaved Grid Disks
  - Enables storage grid that allows multiple applications to place frequently accessed data on faster portions of the disk

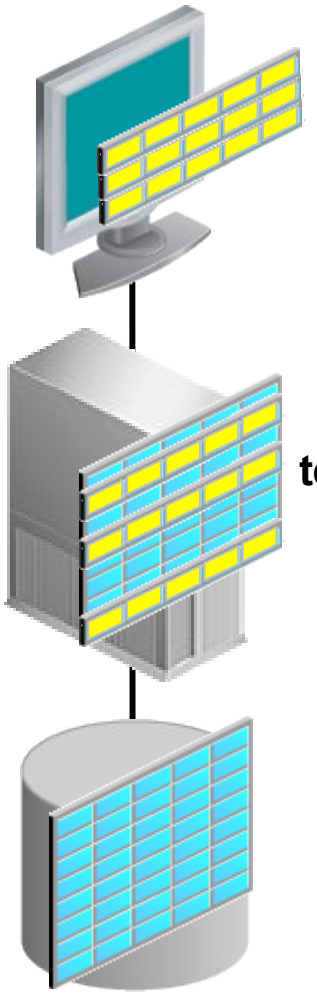
Smart Scan Offload Processing

- Exadata Storage Servers implement smart scans to greatly reduce the data that needs to be processed by database hosts
  - Offload predicate evaluation
  - Only return relevant rows and columns to host
  - Join filtering



- Data reduction is usually very large
  - 10x data reduction is common
- Completely transparent
  - Even if a cell or disk fails during a query
- Smart Scan Example:
  - Telco wants to identify customers that spend more than $200 on a single phone call
  - The information about these premium customers occupies 2MB in a 1 terabyte table

Traditional Scan Processing



❶
```
SELECT
customer_name
  FROM calls
WHERE amount >
200;
```

❷ **Table Extents Identified**

❸ **I/Os Issued**

❹ **I/Os Executed: 1 terabyte of data returned to hosts**

❺ **DB Host reduces terabyte of data to 1000 customer names that are returned to client**

❻ **Rows Returned**

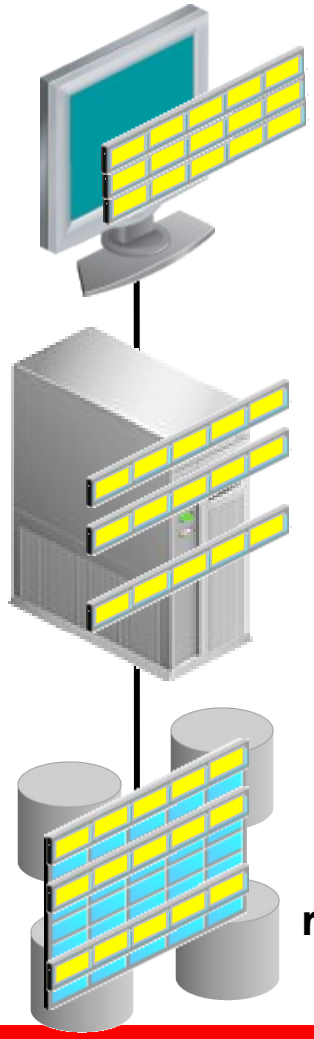
- With traditional storage, all database intelligence resides in the database hosts
- Very large percentage of data returned from storage is discarded by database servers
- Discarded data consumes valuable resources, and impacts the performance of other workloads

Exadata Smart Scan Processing

❶
```
SELECT
customer_name
  FROM calls
WHERE amount >
  200;
```

❷ **Smart Scan Constructed And Sent To Cells**

❸ **Smart Scan identifies rows and columns within terabyte table that match request**



❹ **2MB of data returned to server**

❺ **Consolidated Result Set Built From All Cells**

❻ **Rows Returned**

- Only the relevant columns
  customer_name
  and required rows
  where amount>200
  are returned to hosts
- CPU consumed by predicate evaluation is offloaded to Exadata
- Moving scan processing off the database host frees host CPU cycles and eliminates massive amounts of unproductive messaging
  - Returns the needle, not the entire hay stack

Exadata Software Features

- Exadata Smart Scans
  - 10X or greater reduction in data sent to database servers
- Exadata Smart Flash Cache
  - Breaks random I/O bottleneck by increasing IOPs by 20X
  - Doubles user data scan bandwidths
- OLTP Compression & Hybrid Columnar Compression (HCC)
  - Efficient compression increases effective storage capacity and increases user data scan bandwidths by a factor of 10X
- I/O Resource Manager (IORM)
  - Enables storage grid by prioritizing I/Os to ensure predictable performance
- Inter-leaved Grid Disks
  - Enables storage grid that allows multiple applications to place frequently accessed data on faster portions of the disk

Exadata Smart Flash Cache

- Caches Hot Data Transparently in the 4 Flash Cards (96GB each)
- Use PCI Express based Flash Cards for greater throughput and IOPs and avoid disk controller limitations
- Smart Caching
  - Smarter than basic LRU algorithm
  - Knows when to skip caching objects to avoid polluting or flushing the cache
- Allows applications to explicitly optimize caching

Oracle is the First Flash Optimized Database

Exadata Software Features

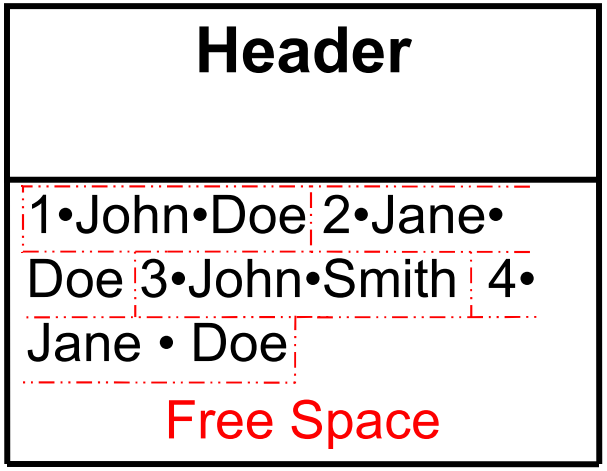
- Exadata Smart Scans
  - 10X or greater reduction in data sent to database servers
- Exadata Smart Flash Cache
  - Breaks random I/O bottleneck by increasing IOPs by 20X
  - Doubles user data scan bandwidths
- OLTP Compression & Hybrid Columnar Compression (HCC)
  - Efficient compression increases effective storage capacity and increases user data scan bandwidths by a factor of 10X
- I/O Resource Manager (IORM)
  - Enables storage grid by prioritizing I/Os to ensure predictable performance
- Inter-leaved Grid Disks
  - Enables storage grid that allows multiple applications to place frequently accessed data on faster portions of the disk

OLTP Table Compression

Employee Table

ID	FIRST_NAME	LAST_NAME
1	John	Doe
2	Jane	Doe
3	John	Smith
4	Jane	Doe

Initially Uncompressed Block




```
INSERT INTO EMPLOYEE
    VALUES (5, 'Jack', 'Smith');
COMMIT;
```

OLTP Table Compression

Employee Table

ID	FIRST_NAME	LAST_NAME
1	John	Doe
2	Jane	Doe
3	John	Smith
4	Jane	Doe
5	Jack	[illegible]

Compressed Block

Header

John=❶|Doe=❶|Jane=❷|Smith=❸

1•❶•❶ 2•❷•❶ 3•❶•❸ 4 • ❷ • ❶ 5•Jack•❸

Free Space

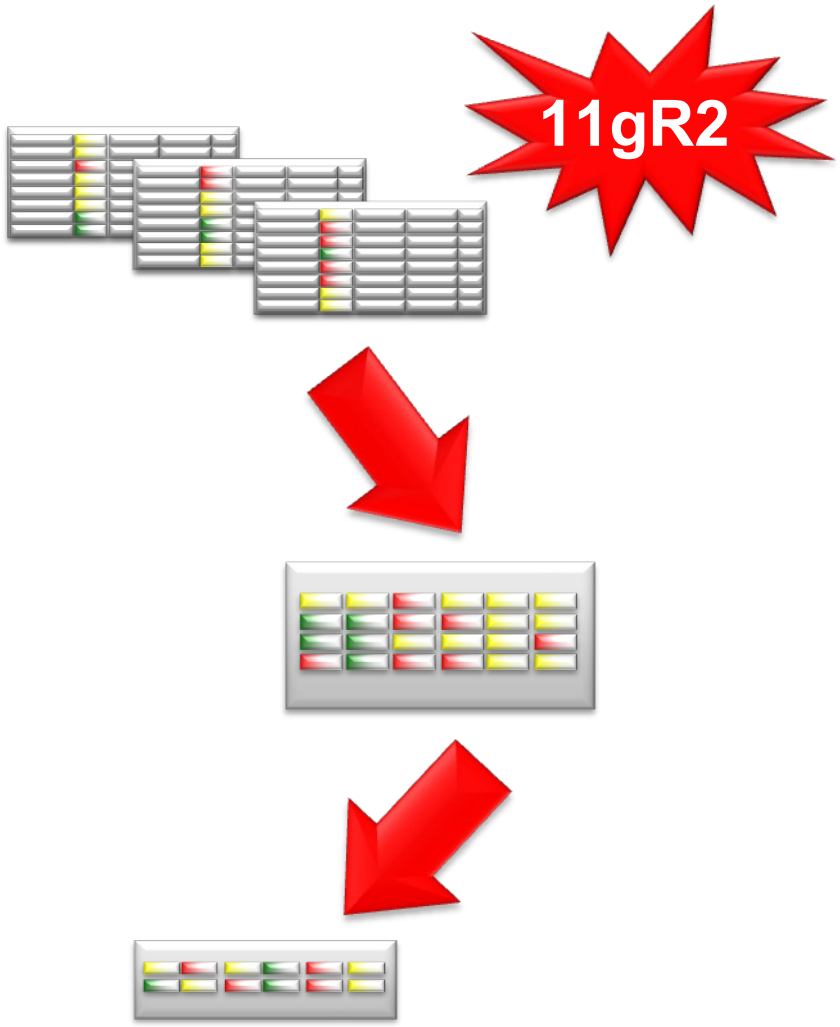
Local Symbol Table

Exadata Hybrid Columnar Compression

- Data is grouped by column and then compressed

- **Query Mode** for data warehousing
  - Optimized for speed
  - **10X compression** typical
  - Scans improve proportionally

- **Archival Mode** for infrequently accessed data
  - Optimized to reduce space
  - 15X compression is typical
  - Up to 50X for some data



Exadata Software Features

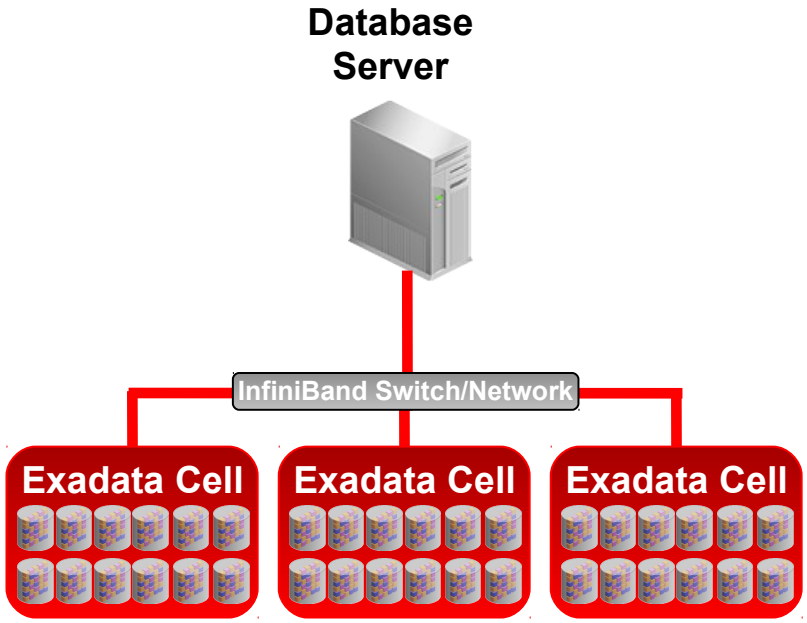
- Exadata Smart Scans
  - 10X or greater reduction in data sent to database servers
- OLTP Compression & Hybrid Columnar Compression (HCC)
  - Efficient compression increases effective storage capacity and increases user data scan bandwidths by a factor of 10X
- Exadata Smart Flash Cache
  - Breaks random I/O bottleneck by increasing IOPs by 20X
  - Doubles user data scan bandwidths
- I/O Resource Manager (IORM)
  - Enables storage grid by prioritizing I/Os to ensure predictable performance
- Inter-leaved Grid Disks
  - Enables storage grid that allows multiple applications to place frequently accessed data on faster portions of the disk

Exadata I/O Resource Management

Mixed Workload Environments

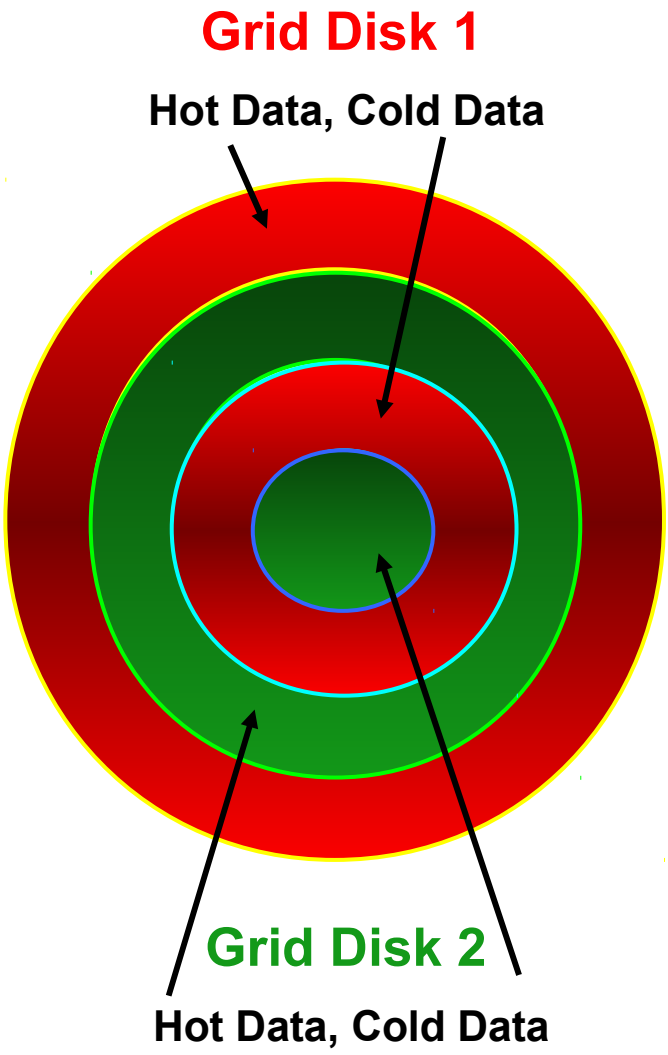
- With traditional storage,creating and managing shared storage is hampered by the inability to balance the work between users on the same database or on multiple databases sharing the storage subsystem
  - Hardware isolation is the approach to ensure separation
- **Exadata I/O resource management ensures different users and tasks within a database are allocated the correct relative amount of I/O resources**

- For example:
  - Interactive: 50% of I/O resources
  - Reporting: 30% of I/O resources
  - ETL: 20% of I/O resources



Interleaved Grid Disks

- Grid disks are optionally split and interleaved to place frequently accessed data in all grid disks on higher performing outer tracks
- All applications benefit from higher performance outer tracks of disks



Sun Oracle Database Machine

Get on the Grid Faster

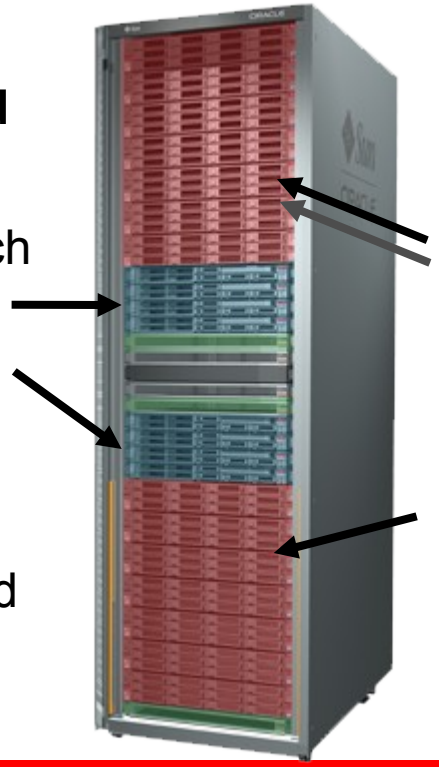
- Highest performance, lowest cost, fault tolerant, scalable on demand
- Database Machine is an engineered, optimized, standardized, and tested grid for Oracle database – with intelligent storage

Oracle Database Server Grid

- 8 compute servers
- 2 Intel quad-core Xeons each (64 cores)
- 576 GB DRAM

InfiniBand Network

- 40 Gb/sec unified server and storage network
- Fault Tolerant



Exadata Storage Server Grid

- 14 storage servers
- 100 TB raw SAS disk storage

or

336 TB raw SATA disk storage

- **5TB flash storage!**
- Offload queries into storage

Sun Oracle Database Machine

Get on the Grid Faster

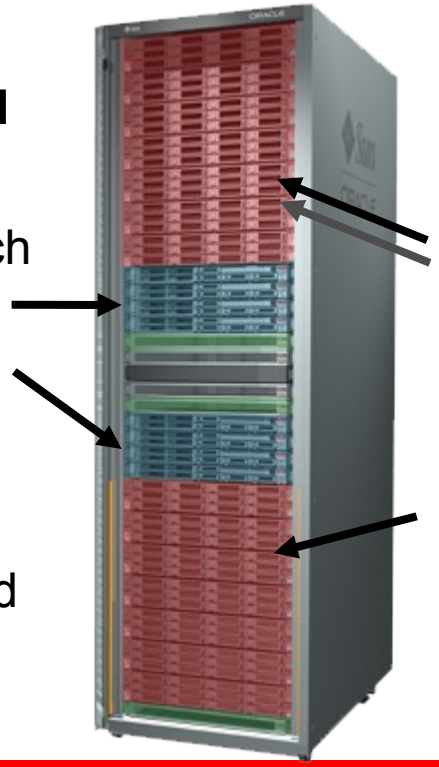
- Highest performance, lowest cost, fault tolerant, scalable on demand
- Database Machine is an engineered, optimized, standardized, and tested grid for Oracle database – with intelligent storage

Oracle Database Server Grid

- 8 compute servers
- 2 Intel quad-core Xeons each (64 cores)
- 576 GB DRAM

InfiniBand Network

- 40 Gb/sec unified server and storage network
- Fault Tolerant



Exadata Storage Server Grid

- 14 storage servers
- 100 TB raw SAS disk storage

or

336 TB raw SATA disk storage

- **5TB flash storage!**
- Offload queries into storage

Exadata Product Performance

		Quarter Rack	Half Rack	Full Rack
Raw Disk Data Bandwidth[1,4]	SAS	4.5 GB/s	10.5 GB/s	21 GB/s
	SATA	2.5 GB/s	6 GB/s	12 GB/s
Raw Flash Data Bandwidth[1,4]		11 GB/s	25 GB/s	50 GB/s
Max User Data Bandwidth[2,4] (10x compression & Flash)		110 GB/s	250 GB/s	500 GB/s
Disk IOPS[3,4]	SAS	10,800	25,000	50,000
	SATA	4,300	10,000	20,000
Flash IOPS[3,4]		225,000	500,000	1,000,000
Data Load Rate[4]		1 TB/hr	2.5 TB/hr	5 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.

Start Small and Grow

Quarter Rack

Half Rack

Full Rack

Sun Oracle Database Machine

Product Family

	Quarter Rack	Half Rack	Full Rack	2-8 Full Racks
Database Servers	2	4	8	16-64
Exadata Storage Servers	3	7	14	28-112
User Data (uncompressed)	6 TB	14 TB	28 TB	56 – 224 TB
Flash cache	1.1 TB	2.6 TB	5 TB	10 – 50 TB
I/O Throughput (disks)	4.5 GB/sec	10.5 GB/sec	21 GB/sec	42 - 168 GB/sec
I/O Throughput (flash)	11 GB/sec	25 GB/sec	50 GB/sec	100 - 400 GB/sec
I/O per Second (IOPS)	225,000	500,000	1,000,000	2M – 8M
Racks	1	1	1	2-8

Scale Performance and Capacity

- **Scalable**
  - Scales to 8 rack database machine by just adding wires
    - More with external InfiniBand switches
  - Scales to hundreds of storage servers
    - Multi-petabyte databases
- **Redundant and Fault Tolerant**
  - Failure of any component is tolerated
  - Data is mirrored across storage servers

Drastically Simplified Deployments

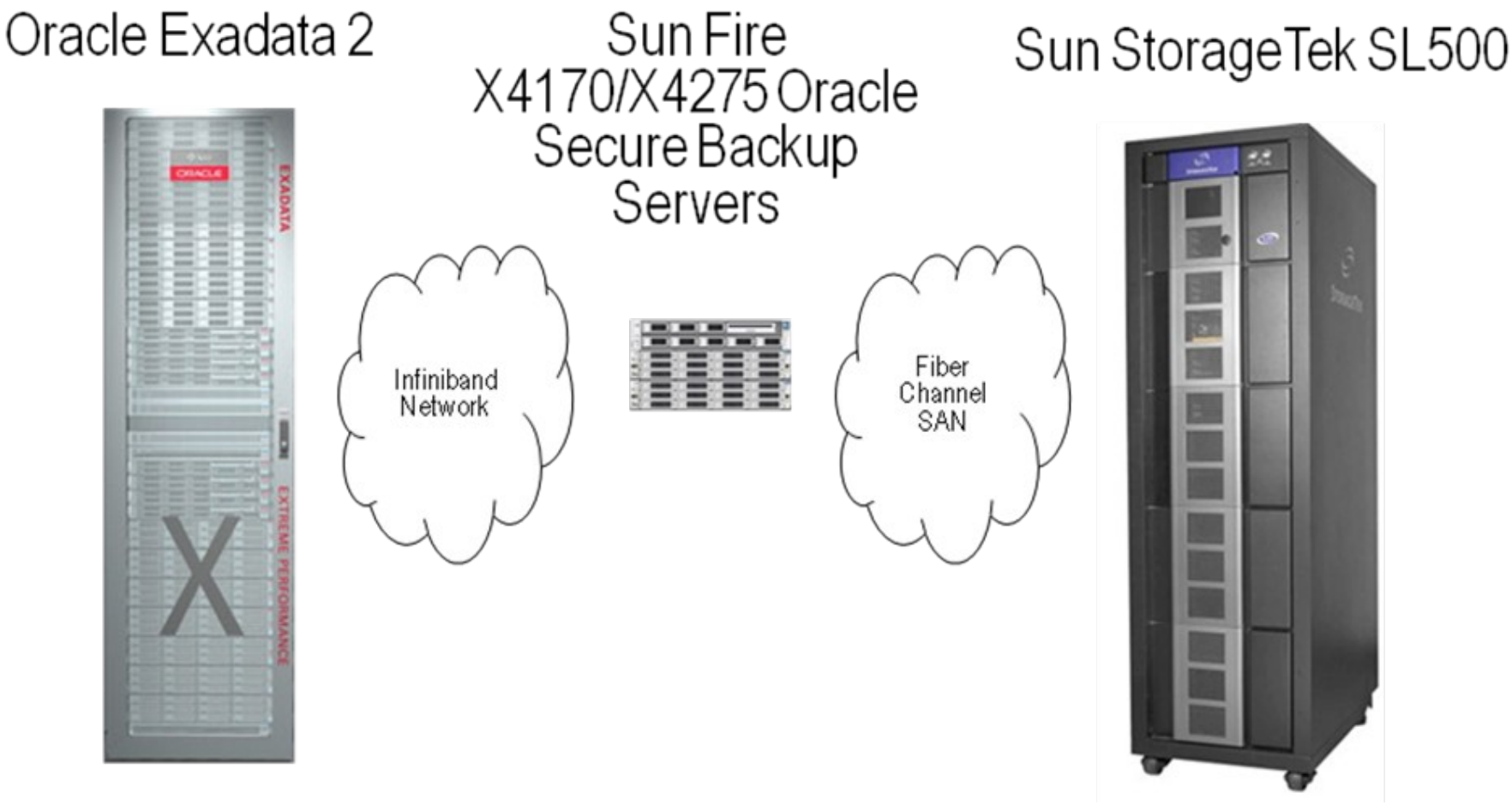
Months to Days

- Database Machine eliminates the complexity of deploying database systems
  - Months of configuration, troubleshooting, tuning
- Database Machine is ready on day one
  - Pre-built, tested, standard, supportable configuration
  - **Runs existing applications unchanged**
- Extreme performance out of the box

Data Protection Solutions

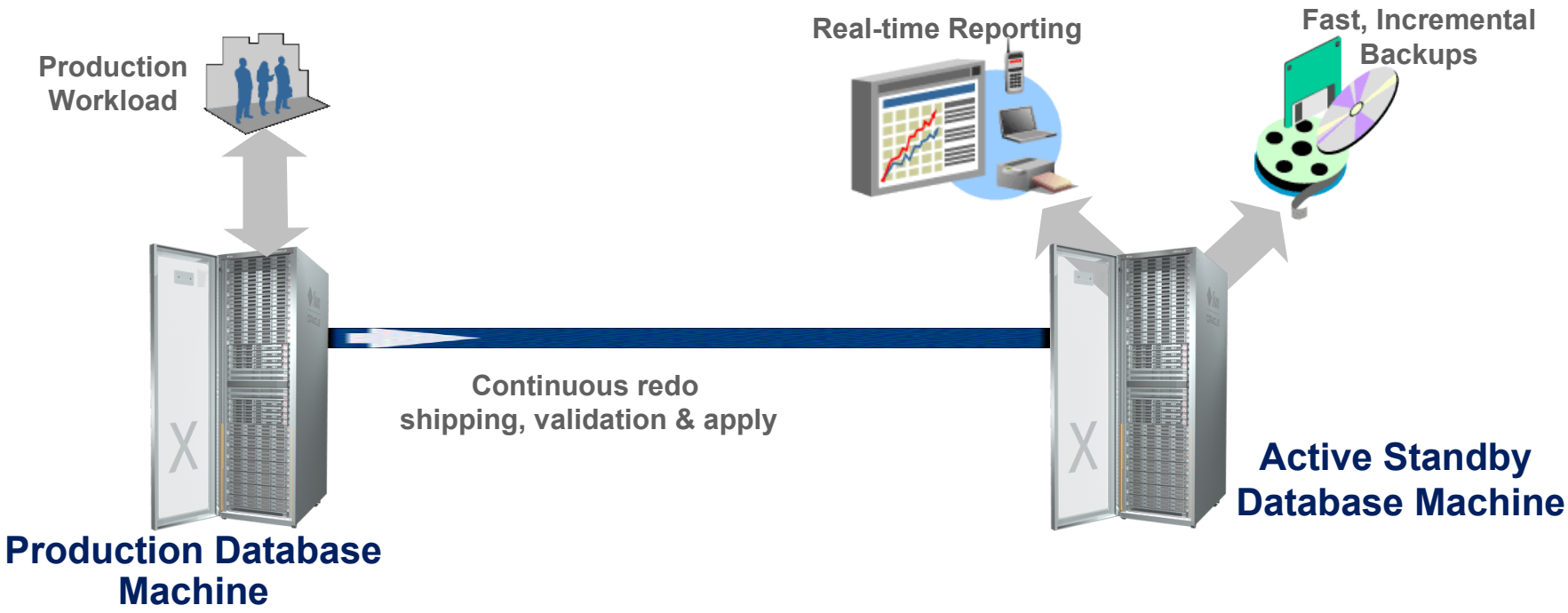
- All single points of failure eliminated by the Exadata Storage architecture
- Hardware Assisted Resilient Data (HARD) built in to Exadata Storage
  - Prevent data corruption before it happens
- Data Guard provides disaster protection and data corruption protection
  - Automatically maintains one or more copies of the database
- Flashback provides human error protection
  - Snapshot-like capabilities to rewind database to before error
- Recovery Manager (RMAN) provides backup to disk
  - Archiving and corruption protection
  - Can be used with Oracle Secure Backup (OSB) or third party tape backup software
- These work just as they do for traditional non-Exadata storage
  - Users and database administrator use familiar tools

Example Tape backup architecture



Sun Oracle Database Machine

Disaster Recovery with Data Guard



- **Both sides benefit from Exadata features!**
  - **Offload query processing, Fast Interconnect, Scale-out Storage**
  - **Use offloaded incremental backups on a physical standby**
  - **Fast File Creation**

Oracle Enterprise Manager (SYSMAN) - Data Guard - Microsoft Internet Explorer

File Edit View Favorites Tools Help

ORACLE Enterprise Manager

Setup Preferences Help Logout

Home | Targets | Configuration | Alerts | Jobs | Management System

Hosts | Databases | Application Servers | Web Applications | Groups | All Targets

Cluster: drcluster > Cluster Database: San Francisco > Data Guard

Data Guard

Page Refreshed August 1, 2003 5:27:25 PM EDT

Overview

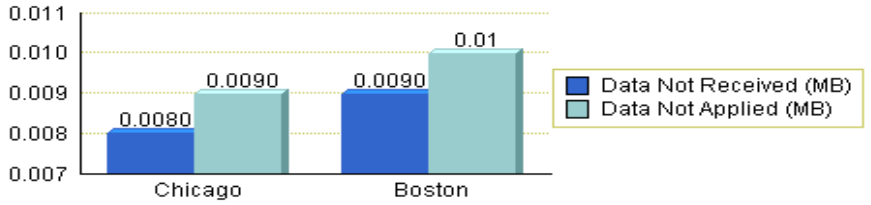
Overall Status ✔ **Normal**
Protection Mode Maximum Performance

Primary Cluster Database

Name San Francisco
Cluster drcluster
Status ✔ Normal
Current Log Multiple Threads
Related Link Edit

Standby Progress Summary

This chart shows the amount of data that each standby has not yet received and applied.



Standby Databases

Add Standby Database | Edit | Remove | Switchover | Failover

Select	Name	Host	Status	Role	Last Received Log	Last Applied Log
◉	Chicago	drlab3	✔ Normal	Physical Standby	Multiple Threads	Multiple Threads
○	Boston	drlab4	✔ Normal	Logical Standby	Multiple Threads	Multiple Threads

Performance

Performance Overview
Log File Details

Additional Administration

Verify
Remove Data Guard Configuration

Home | **Targets** | Configuration | Alerts | Jobs | Management System | Setup | Preferences | Help | Logout

Copyright © 1996, 2003, Oracle. All rights reserved.
About Oracle Enterprise Manager

ORACLE

Complete, Open, Integrated Security

Monitoring

Configuration Management

Audit Vault

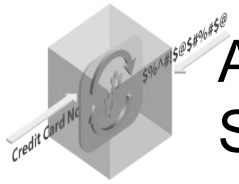
Total Recall

Access Control

Database Vault

Label Security

Encryption and Masking



Advanced Security

Secure Backup

Data Masking

Exadata Storage Management & Administration

- Enterprise Manager
  - Manage & administer Database and ASM
- Exadata Storage Plug-in
  - Enterprise Manager Grid Control Plug-in to monitor & manage Exadata Storage Cells
- Sun Embedded Integrated Lights Out Manager (ILOM)
  - Remote management and administration of hardware

Best Machine for Consolidating Databases

- The Database Machine runs any combination of workloads with extreme performance
  - Warehouse oriented **bulk data processing**
  - OLTP oriented **random updates**
  - Real time BI against transactional data
- ASM spreads data from all databases evenly across storage
  - **Exadata I/O resource manager** prioritizes I/Os to ensure predictable performance
- Large databases **span nodes** using RAC
- Multiple small databases run on a single node
  - Instance caging ensures predictable performance

SOFTWARE. HARDWARE. COMPLETE.