ORACLE®

Oracle Maximum Availability Architecture with Exadata Database Machine

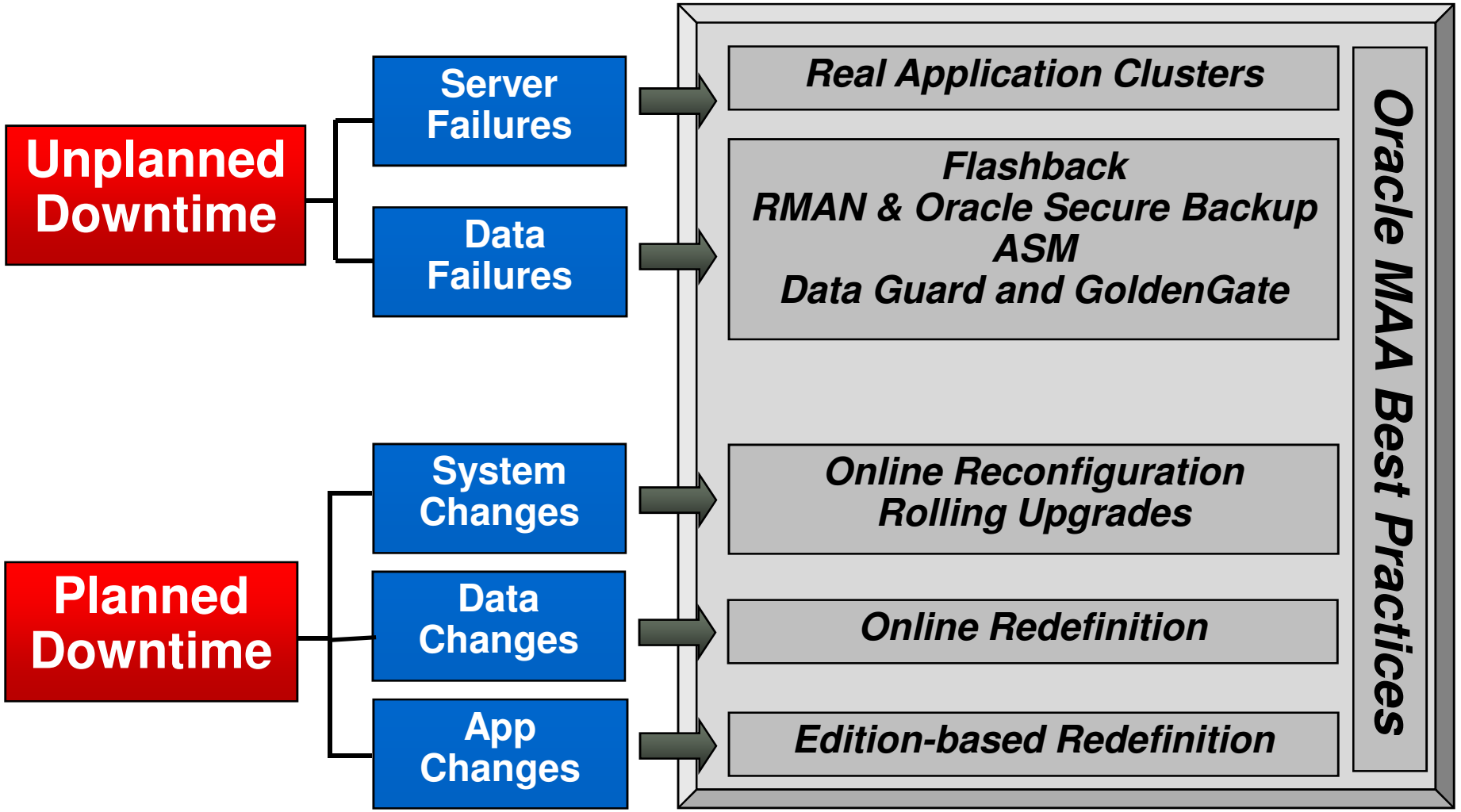
Morana Kobal Butković
Principal Sales Consultant
Oracle Hrvatska

MAA is Oracle's Availability Blueprint

- Oracle's MAA is a best practices blueprint for:
  - High Availability (HA)
  - Disaster Recovery (DR)
  - Backup and Recovery (BR)
- Validated in our labs and proven with customers
- Operational and Configuration Best Practices for
  - Fault tolerance and fast repair
  - Online changes and rolling upgrades
  - Lowest downtime and highest availability
- Exadata Database Machine is MAA's Reference Platform

http://www.oracle.com/goto/maa

Oracle's Database HA Solution Set



Exadata Architecture

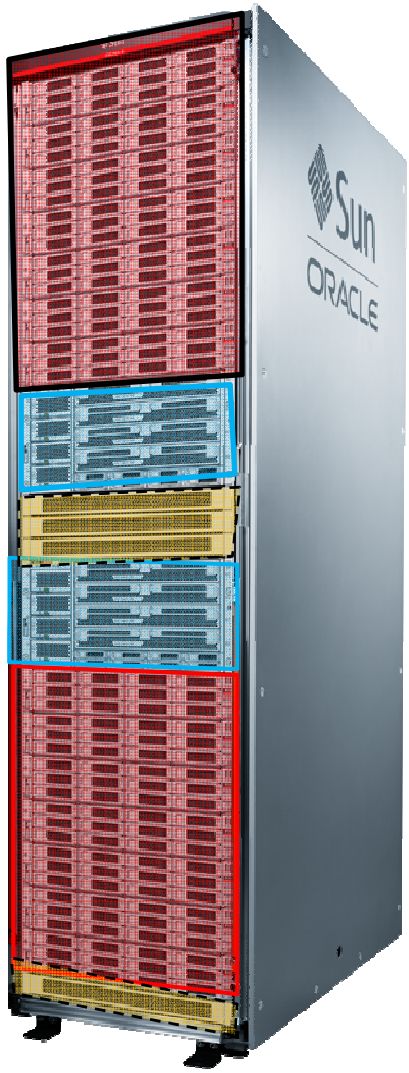
Database Grid

- 8/4/2 Database Servers
- 2 x 6-core Intel Xeon processors per server
- 96GB DRAM per server
- Oracle Database 11.2
- Oracle Linux or Solaris

InfiniBand Network

- 3 x 36-port 40Gb/s switches
- Unified server & storage network

1 or 10 Gb Ethernet to Data Center



Intelligent Storage Grid

- 14/7/3 Storage Servers

- 12 x 600GB High Performance or
  12 x 2TB High Capacity disks
- 2 x 6-core Intel Xeon processors per server
- 4 x 96GB Flash PCIe per server

Intelligent Exadata Storage Server Software

Exadata Out-of-the-Box Availability

Delivered pre-optimized, pre-configured, validated configuration defaults for MAA

Database Grid

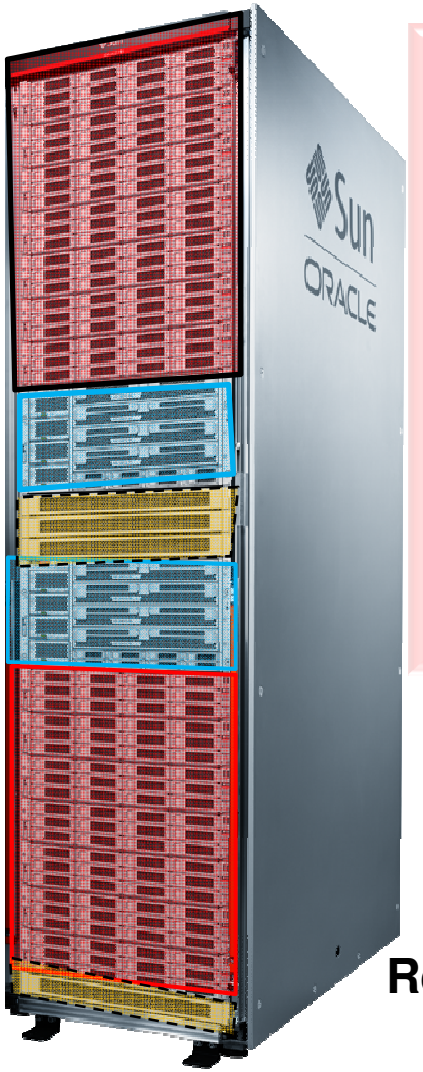
- Multiple Database Servers in Real Application Cluster
- Protection from server failures

InfiniBand Network

- Multiple IB switches
- Dual ported IB Host Channel Adapters
- Redundant GigE and IB links

1 or 10 Gb Ethernet

- Redundant ports



Intelligent Storage Grid

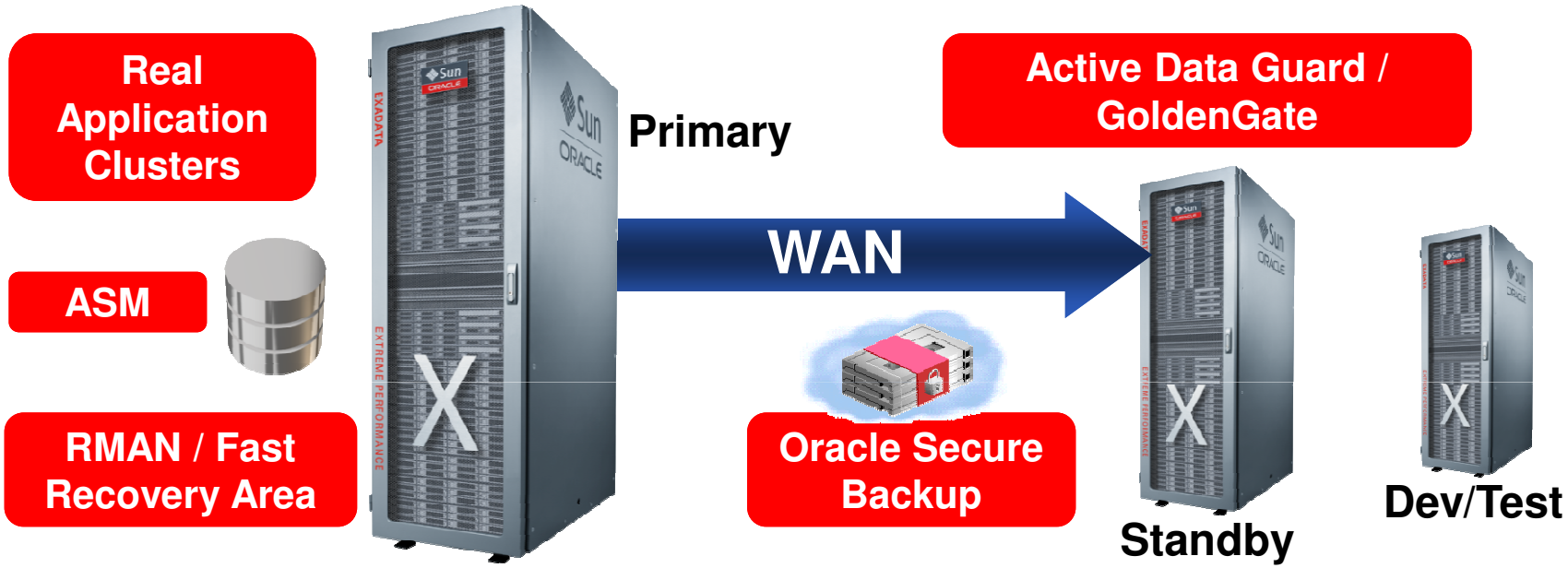
- Redundant Storage Servers

- Fault tolerance for both cell and disk failures using ASM redundancy
- Exadata storage built-in corruption checks

Redundant PDUs

Best Practices for Availability

Maximum Availability Architecture for Exadata



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

ASM Disk Group Configuration

Recommended ASM disk group configuration

- Disk Groups
  - DATA on the OUTER sections of all disks and cells
  - RECO on the OUTER-INNER sections of all disks and cells
  - DBFS_DG on the INNER sections of all disk and cells

- Key advantages
  - All databases use the same disk groups
  - All operations have access to full IO bandwidth if required
  - IO Resource Manager can set priority of IO operations

ASM Disk Group Configuration

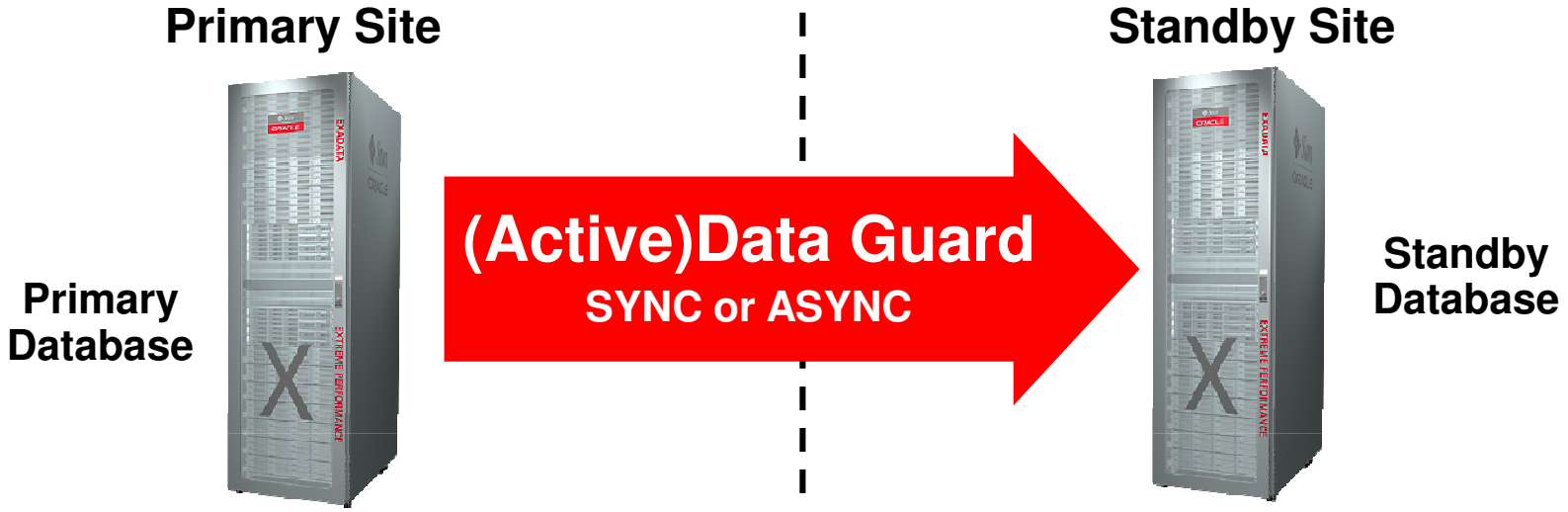
High Redundancy Disk Group Recommendation

- High Redundancy Benefits
  - Prevents loss of cluster and disk group due to dual storage failures
  - Tolerates storage failure during Exadata planned maintenance
  - Prevents data loss with storage failure and latent defects
- Recommend at least one High Redundancy disk group
  - If DATA is HIGH, application remains available
  - If RECO is HIGH, database can be restored with zero data loss
  - Pick the disk group configuration option during deployment
  - Half Rack Database Machine is minimum

Disaster Recovery Options

Oracle Data Guard

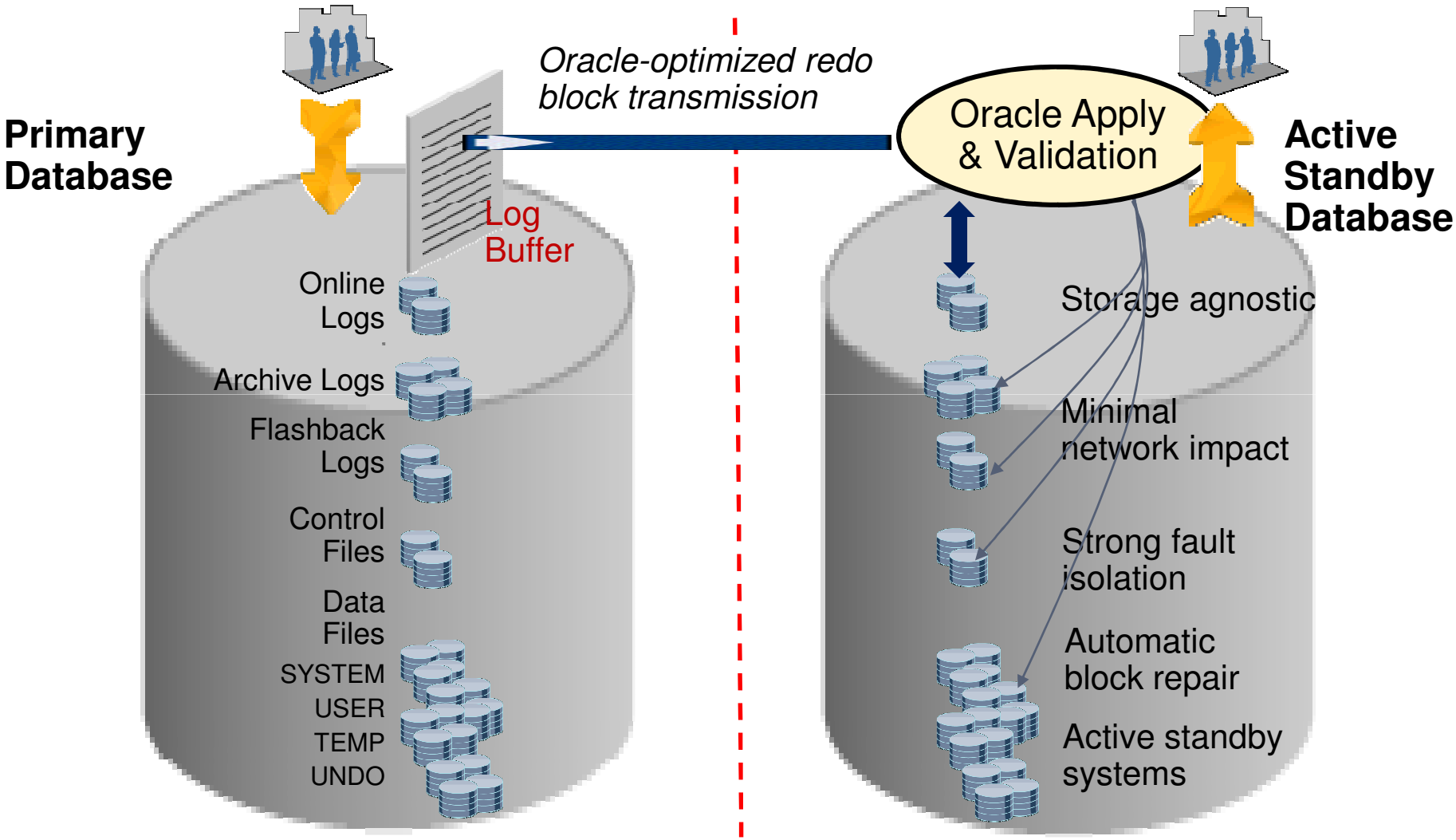
Data Protection and Availability for Oracle Database



- Data protection – continuous synchronization and data validation
- Data availability - automatic failover to standby database
- Minimize planned downtime - database rolling upgrades (transient logical), Standby-First Patch Apply >=11.2.0.1
- Reduction in primary database workload
  - Backup operations
  - Running large reports on active data off the primary

Oracle Data Guard

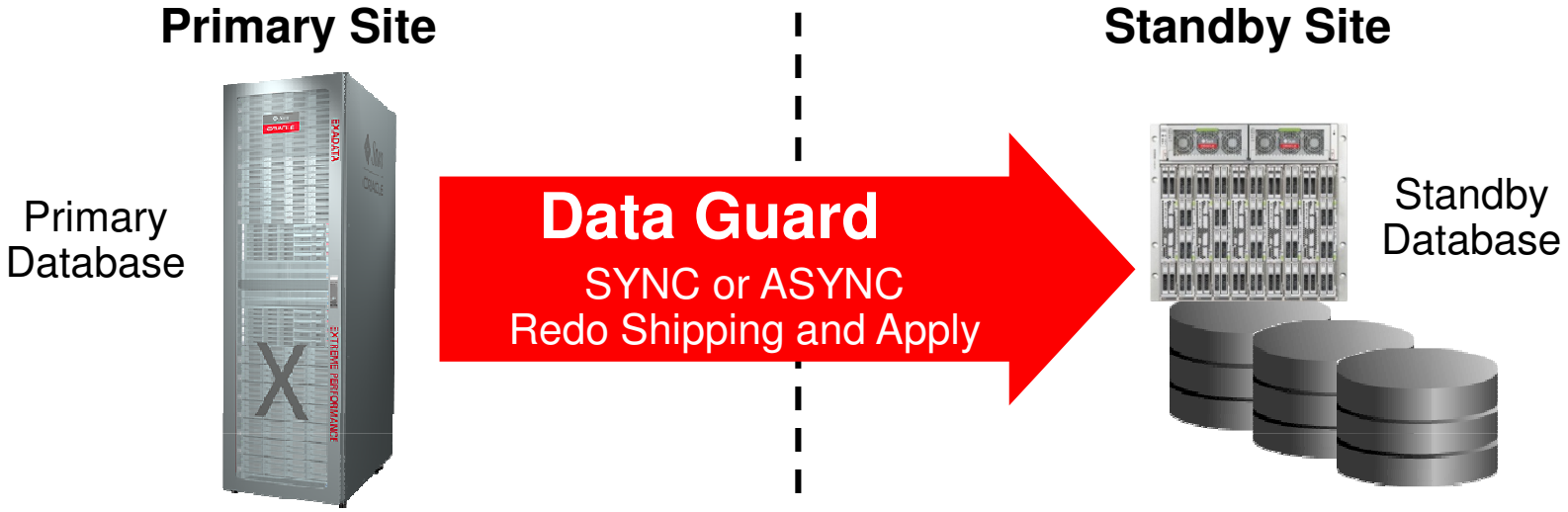
Optimized Usage of Network Resources



Also: standby database protected from primary database block corruptions

Non-Exadata Server and Storage

Watch Out - Exadata Hybrid Columnar Compression



- Upon failover - Exadata Hybrid Columnar Compression (EHCC) tables will need to be uncompressed impacting RTO
- Requires more storage (10x-15x) for Stdby system with performance cost
- Active Data Guard cannot be used to read EHCC tables on non-Exadata Stdby
- Post failover role reversal – Primary Exa will not be able to leverage EHCC

MOS Note 413484.1 OS must match from site to site --- Example: (Primary) OL → (Remote) OL

Oracle Golden Gate for Exadata

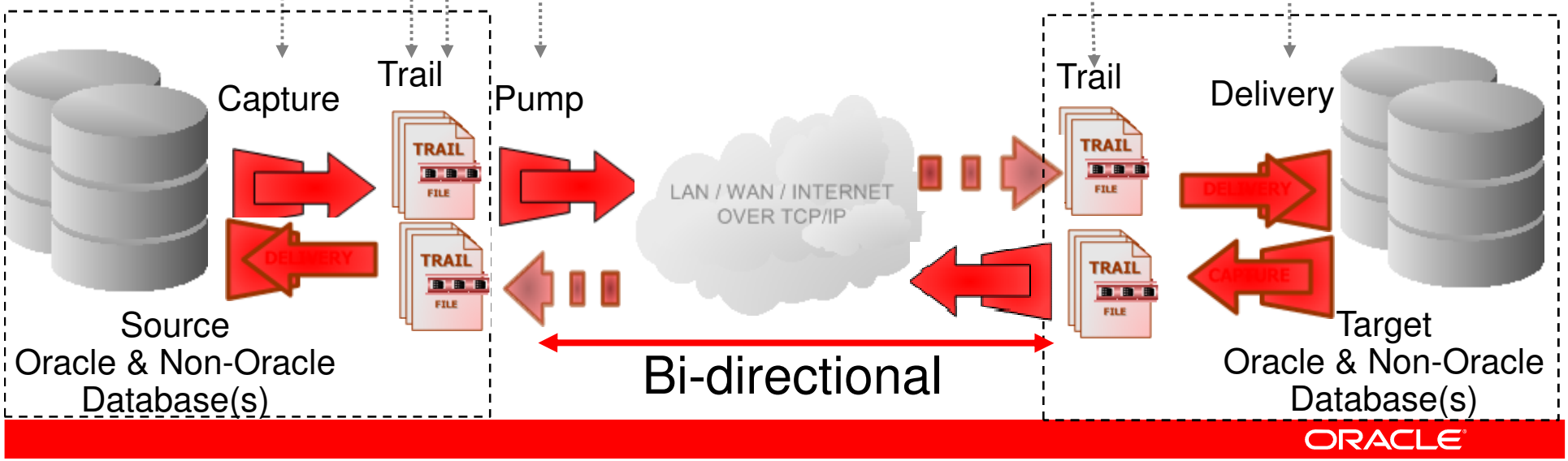
Capture: committed transactions are captured (and can be filtered) as they occur by reading the transaction logs.

Trail: stages and queues data for routing.

Pump: distributes data for routing to target(s).

Route: data is compressed, encrypted for routing to target(s).

Delivery: applies data with transaction integrity, transforming the data as required.



Backup & Recovery Options

Best Practices for Database Backups

Fastest

21TB/hour

27 TB/hour

Up to 50 TB/hour for incremental

Flexible

6-8 TB/hour

Snapshots/Clones

Economical

5-8 TB/hour

(Dependent on # of tape drives)

Disk Based Backup & Recovery

RMAN Backup to Exadata Storage

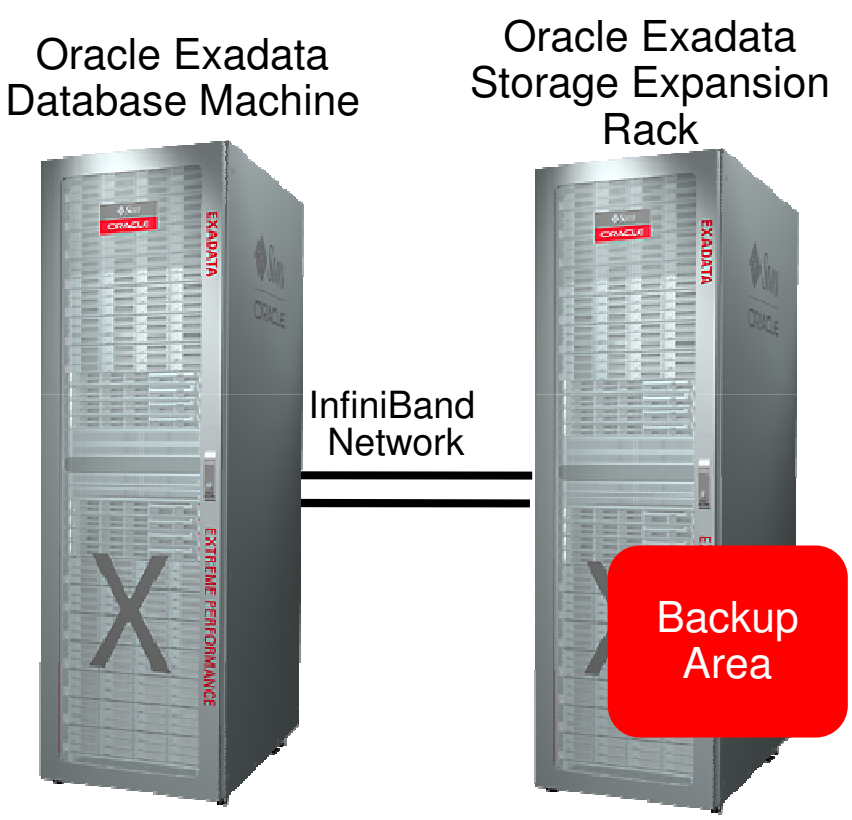
Oracle Exadata Database Machine

- **Oracle Suggested Backup Strategy**
  - Use RMAN **Incrementally Updated** backups
  - Image Copy stored in Fast Recovery Area
  - Nightly Incremental Backups created in Fast Recovery Area
  - Changed blocks are tracked using **block change tracking** bitmap file
  - Backs up only changed blocks – Exadata offload
  - Incremental Backups merged into Image Copies on a 24 hour delay basis
- Fastest backup and recovery
  - RTO in range of “minutes” - switch to copy capability

http://www.oracle.com/technetwork/database/features/availability/maa-tech-wp-sundbm-backup-11202-183503.pdf

Disk Based Backup & Recovery

Using additional Storage Cells



- Simplest, fastest and most robust way to add additional storage capacity
- Uses same 2TB high-capacity SAS drives and Smart Flash Cache as Database Machine
- Available in three different sizes; up to eight racks can be connected without additional IB switches
- Key Benefits
  - High performance
  - Simple unified management

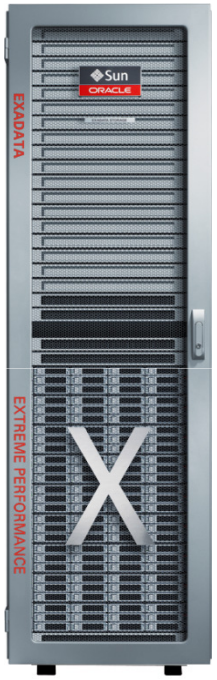
Exadata Storage Expansion Racks

Expand Database Machine Storage Capacity Online

Quarter Rack

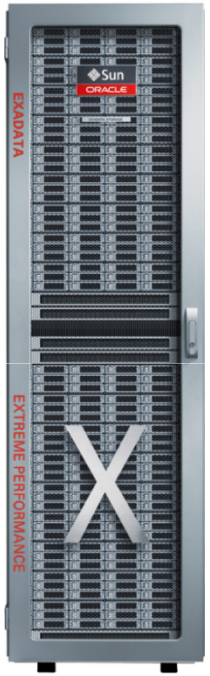
96 TB Disk
1.5 TB Flash
4 Storage Servers 48 CPU cores

Half Rack



216 TB Disk
3.4 TB Flash
9 Storage Servers
108 CPU cores

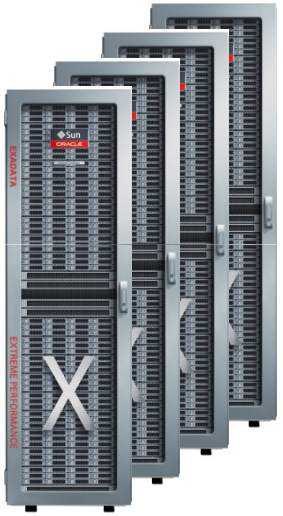
Full Rack



432 TB Disk
6.75 TB Flash
18 Storage Servers
216 CPU cores

Multi Rack

8+ Racks



InfiniBand Connected

Database Backups, Historical Data, Files, Images, XML

Disk Based Backup & Recovery

- MAA WP: Backup and Recovery Performance and Best Practices for Exadata Cell and Oracle Exadata Database Machine, August 2011

http://www.oracle.com/technetwork/database/features/availability/maa-tech-wp-sundbm-backup-11202-183503.pdf

- Disk backup and restore testing was performed with image copy formats using a fast recovery area located on Exadata storage and using varying degrees of RMAN parallelism
- <5% CPU was used

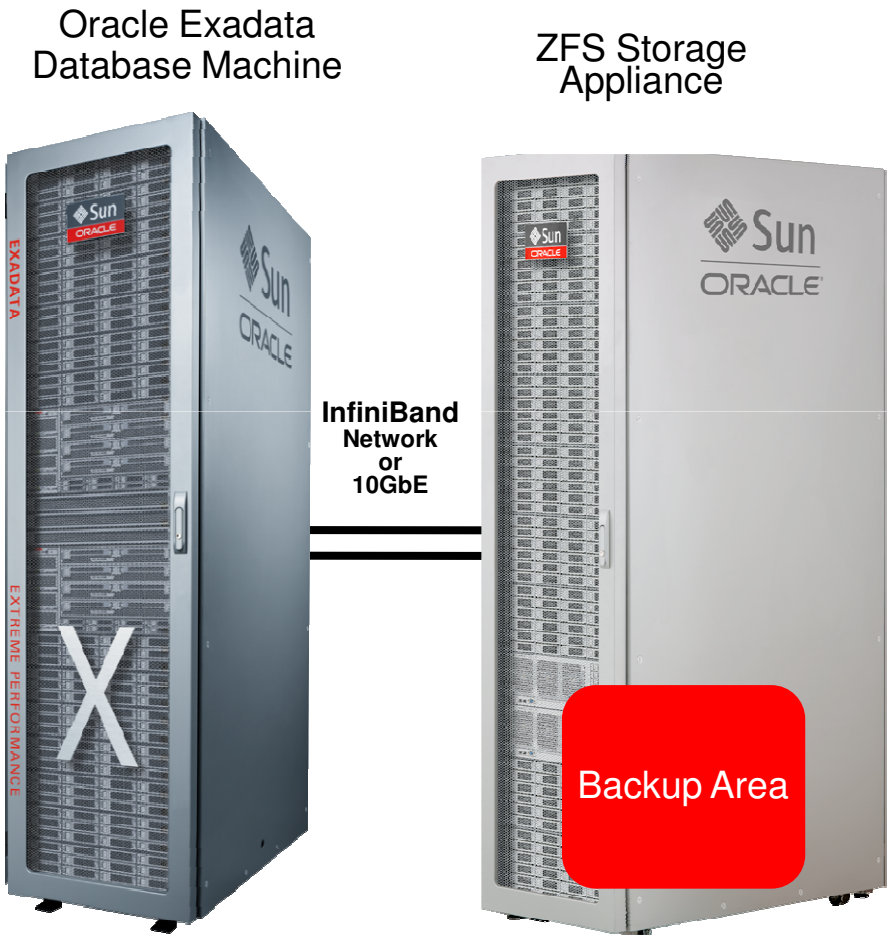
Disk Based Backup & Recovery

Performance

FULL DATABASE BACKUP TO DISK USING IMAGE COPIES[1]			
Instances and Channels	**Quarter Rack**	**Half Rack**	**Full Rack**
X2-2 (11.2.0.2) All instances 2 RMAN channels per instance	4 TB/hour	8-9 TB/hour	17-18 TB/hour
FULL DATABASE INCREMENTAL BACKUP TO DISK (10% CHANGE) USED DAILY			
2 instances, 2 RMAN channels per instance	Measured effective backup rate 20 to 50 TB/hr, depending on workload		
FULL DATABASE RESTORE FROM DISK			
X2-2 (11.2.0.2) All instances 2 RMAN channels per instance	4 TB/hour	7 TB/hour	14TB/hour (DATA is HIGH redundancy disk group) 17 TB/hour (DATA is NORMAL redundancy disk group)

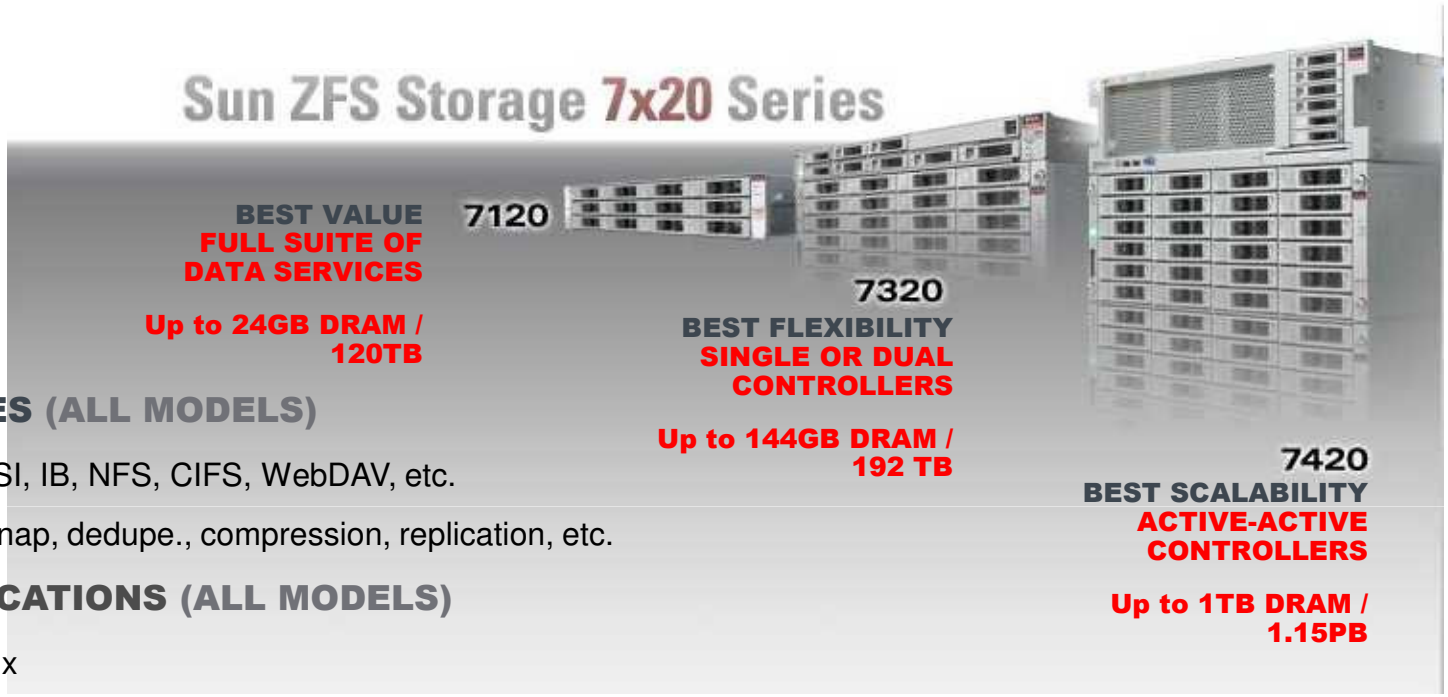
Backup to Sun ZFS Storage Appliance

Flexible



- All Oracle solution - tested, validated & supported together
- Simple Architecture
  - Directly connected to IB, can use 1 GbE, 10 GigE
  - No changes to standard RMAN
  - No media server if tape not used
- ZFS works alongside RMAN to ensure against data corruption
  - End-to-end checksumming and data integrity prevent corrupt files when restoring old backups
- Data services available
  - Compression
  - Snapshot s
  - Augment Oracle DR utilizing replicated copies of RMAN backups

ZFS Storage Appliances



STANDARD FEATURES (ALL MODELS)

All Data Protocols: FC, iSCSI, IB, NFS, CIFS, WebDAV, etc.

Advanced Data Services: Snap, dedupe., compression, replication, etc.

CLIENTS AND APPLICATIONS (ALL MODELS)

Oracle Solaris • Oracle Linux

Oracle Database, Middleware, and Applications

Oracle VM • VMware • Windows

More than 50 business applications supported

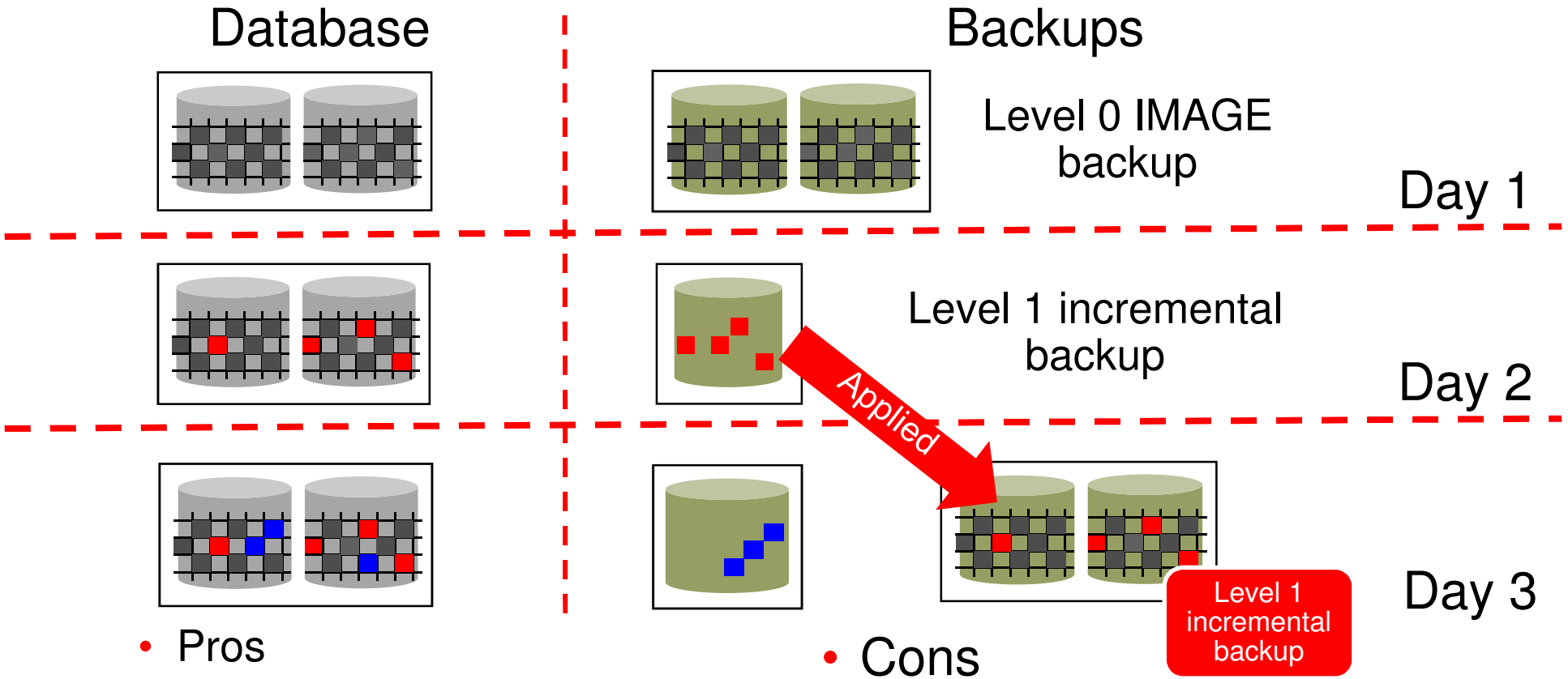
NEW BENEFITS

Best Density and Scale: Industry-leading density, scale up to 1PB for Consolidation

Flash Everywhere and More Of It: Industry-leading flash capacity for Application Performance

Doubled the Processing Power: Performance to drive enterprise Data Protection

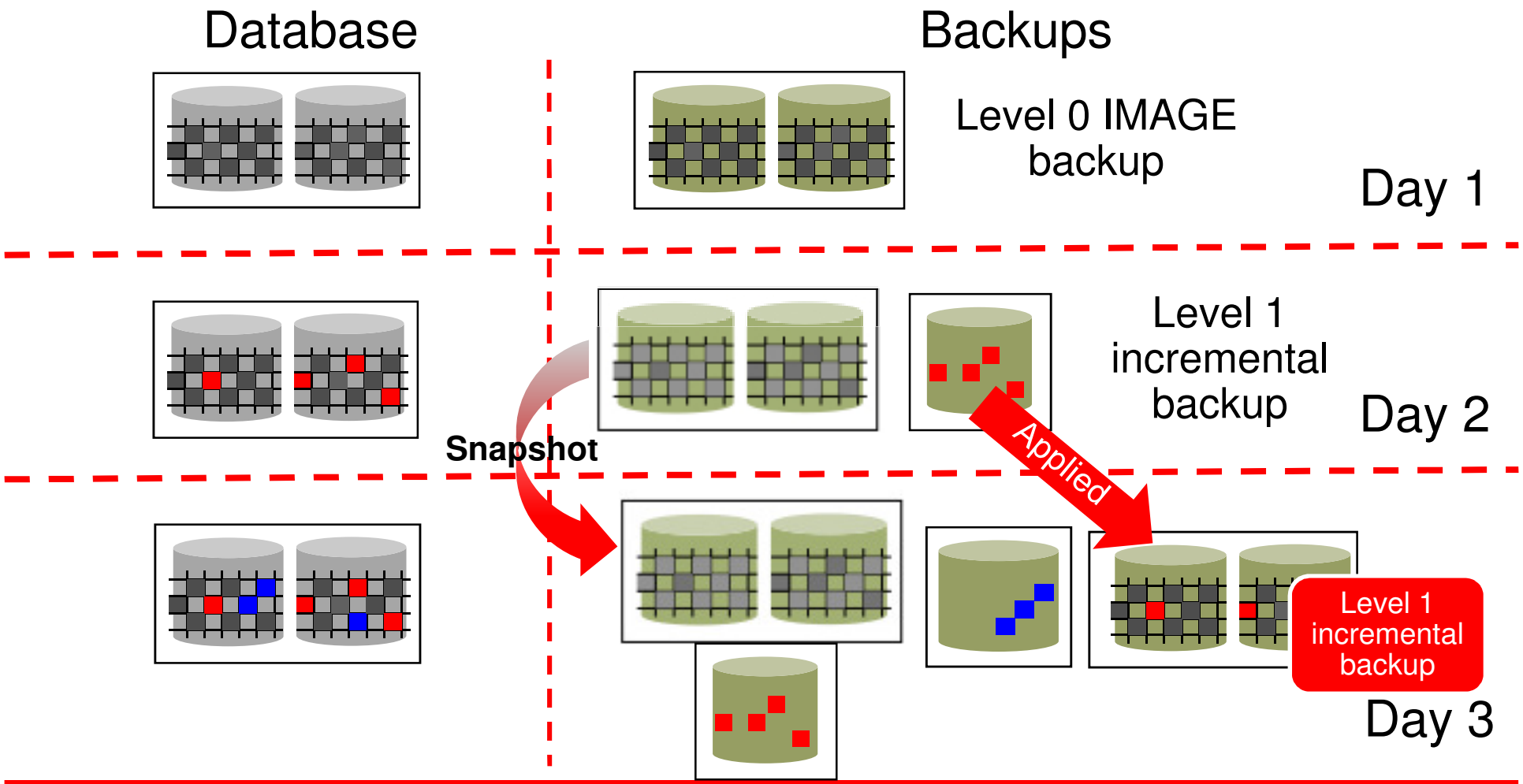
Oracle Suggested Backup Strategy



- Pros
  - Faster recovery from up-to-date image copy
  - Less data to backup = shortened backup windows
- Cons
  - Can not recover data older than 24 hours
  - Image copy can not be compressed by RMAN

Oracle Suggested Backup Strategy

Leveraging the ZFS Storage Appliance's Snapshot



Backup to ZFS Storage Appliance

Performance

- Backup throughput
  - **8.9 TB/hour image copy**
- Restore throughput
  - **6.9 TB/hour image copy**
- Backup and restore operations can be automatically parallelized across all database nodes and Sun ZFS Storage Appliance channels and controllers
- ORACLE RMAN BACKUP AND RESTORE THROUGHPUT FOR A SUN ZFS STORAGE APPLIANCE WITH 2 HEADS, 4 TRAYS, 2 POOLS, AND 16 SHARES
- 16 RMAN channels

Oracle White Paper, September 2011

Protecting Oracle Exadata with the Sun ZFS Storage Appliance: Configuration Best Practices

http://www.oracle.com/technetwork/articles/systems-hardware-architecture/exadata-7000-367640.pdf

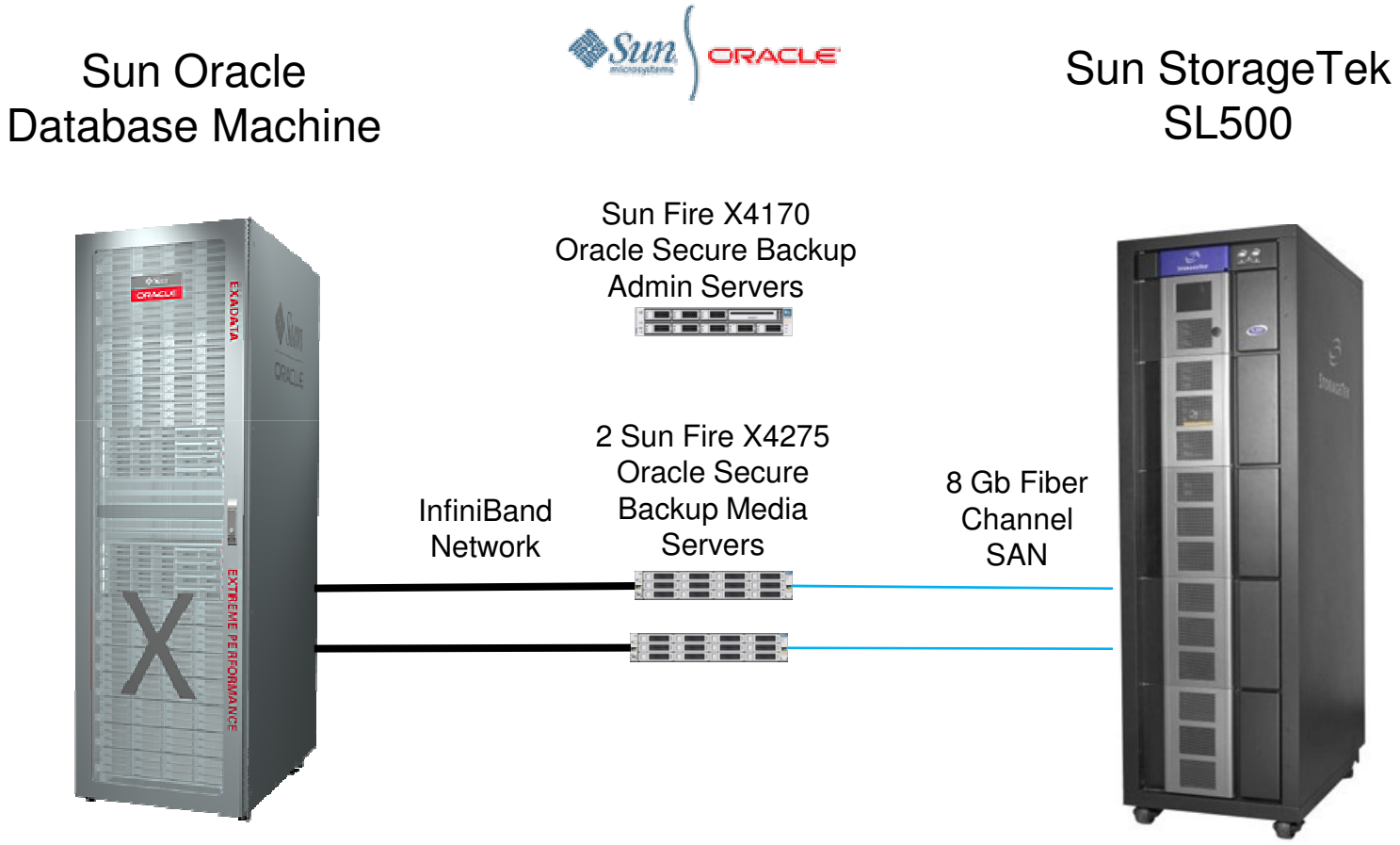
Tape Based Backup & Recovery

- MAA WP: Backup and Recovery Performance and Best Practices for Exadata Cell and Oracle Exadata Database Machine

http://www.oracle.com/technetwork/database/features/availability/maa-tech-wp-sundbm-backup-11202-183503.pdf

- Testing was performed using Oracle Secure Backup Release 10.3, with 2 media servers attached to a Sun StorageTek SL500 tape library and 14 LTO-4 tape drives via SAN
- Any tape backup product that integrates with RMAN is automatically supported
- Backup and restore rates to tape were limited by the aggregate tape transfer rates

MAA Validated Architecture



Tape Based Backup & Recovery

Performance

FULL DATABASE BACKUP TO TAPE			
Instances and Tape Drives	**Quarter Rack**	**Half Rack**	**Full Rack**
All instances, 14 tape drives, 1 RMAN channel per tape drive	2509 MB/sec or 8.6 TB/hr or 179 MB/sec per tape drive	2509 MB/sec or 8.6 TB/hr or 179 MB/sec per tape drive	2509 MB/sec or 8.6 TB/hr or 179 MB/sec per tape drive
FULL DATABASE INCREMENTAL BACKUP TO TAPE (10% CHANGE)			
All instances, 14 tape drives, 1 RMAN channel per tape drive	Measured effective backup rate 10 to 70 TB/hr		
FULL DATABASE RESTORE FROM TAPE			
All instances, 14 tape drives, 1 RMAN channel per tape drive[2]	1800 MB/sec or 6.1 TB/hr or 128 MB/sec per tape drive	2271 MB/sec or 7.8 TB/hr or 162 MB/sec per tape drive	2271 MB/sec or 7.8 TB/hr or 162 MB/sec per tape drive

- <1 CPU core used per instance
- Allocate 1 RMAN channel per tape drive, use all instances

Tape Based Backup & Recovery

- Benefits
  - Fault Isolation from Exadata Storage
  - Maximizes Exadata Database Machine capacity and bandwidth
  - Move backup off-site easily
  - Keep multiple copies of backups in a cost effective manner
- Trade-Offs
  - Disk-based solutions have better recovery times for data and logical corruptions and certain tablespace point in time recovery scenarios
  - No differential incremental backups are available
- Top limiting factors
  - Number and type of tape drives
  - External network (IB, 1 GbE or 10GbE) linking the tape library to Exadata

Backup & Recovery: Complete Story