ORACLE

Exadata V2 - Oracle Exadata Database Machine

Robert Pastijn
Platform Technology Services (PTS) – Product Development

Exadata V2 Goals

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

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

Agenda

- **Hardware Architecture**
- Key Technologies
- Consolidation & Protection
- Migration

Exadata Hardware Architecture

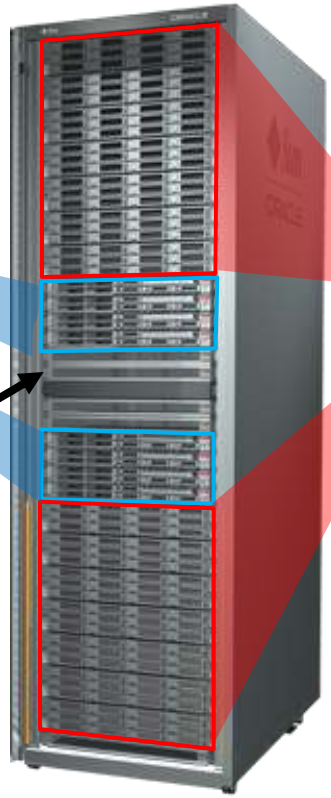
Scaleable Grid of industry standard servers for **Compute and Storage**

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

Database Grid

- 8 compute servers (1U)

- 64 Intel cores



InfiniBand Network

- Redundant 40Gb/s switches
- **Unified** server & storage net

Storage Grid

- 14 storage servers (2U)

- **112 Intel cores in storage**
- 100 TB SAS disk, or 336 TB SATA disk
- **5 TB PCI Flash**
- Data mirrored across storage servers

Start Small and Grow

Field Upgradeable

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

Balanced Incremental Scaling for OLTP and DW

Scales to 8 Racks by Just Adding Cables

Full Bandwidth and Redundancy

Standardized and Simple to Deploy

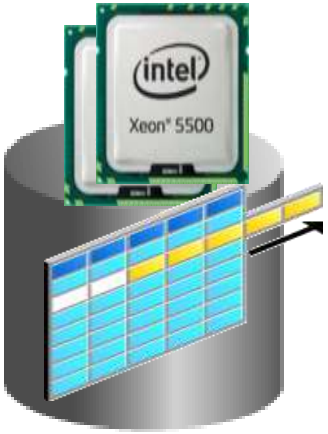
- All Database Machines are the same
  - Delivered Tested and Ready-to-Run
  - Highly Optimized
  - Highly Supportable
  - No unique configuration issues
  - **Identical to config used by Oracle Engineering**
- **Runs existing OLTP and DW applications**
  - Full 30 years of Oracle DB capabilities
  -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

Deploy in Days, Not Months

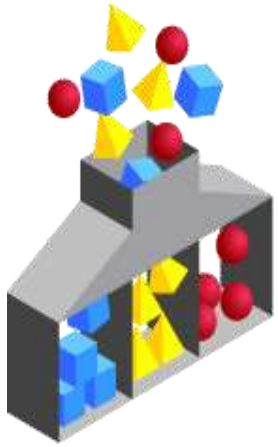
Agenda

- Hardware Architecture
- **Key Technologies**
- Consolidation & Protection
- Migration

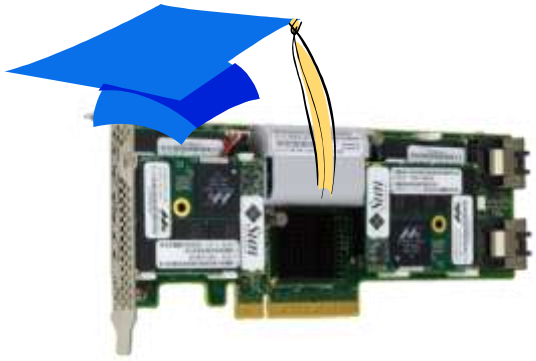
Keys to Speed and Cost Advantage



Exadata Intelligent Storage Grid



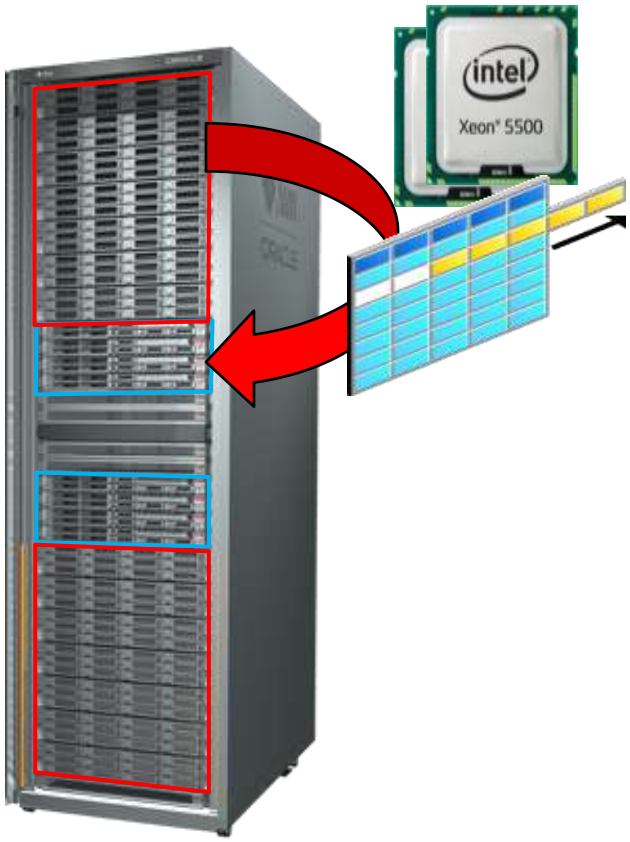
Exadata Hybrid Columnar Compression



Exadata Smart Flash Cache

Exadata Intelligent Storage Grid

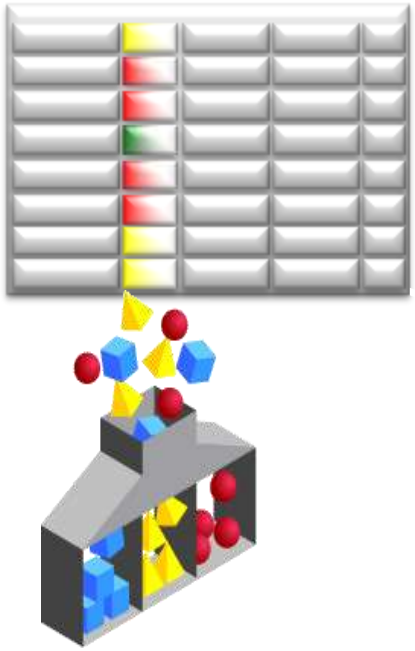
Most Scalable Data Processing



- **Data Intensive processing runs in Exadata Storage Grid**
  - Filter rows and columns as data streams from disks (112 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

Exadata Hybrid Columnar Compression

Highest Capacity, Lowest Cost



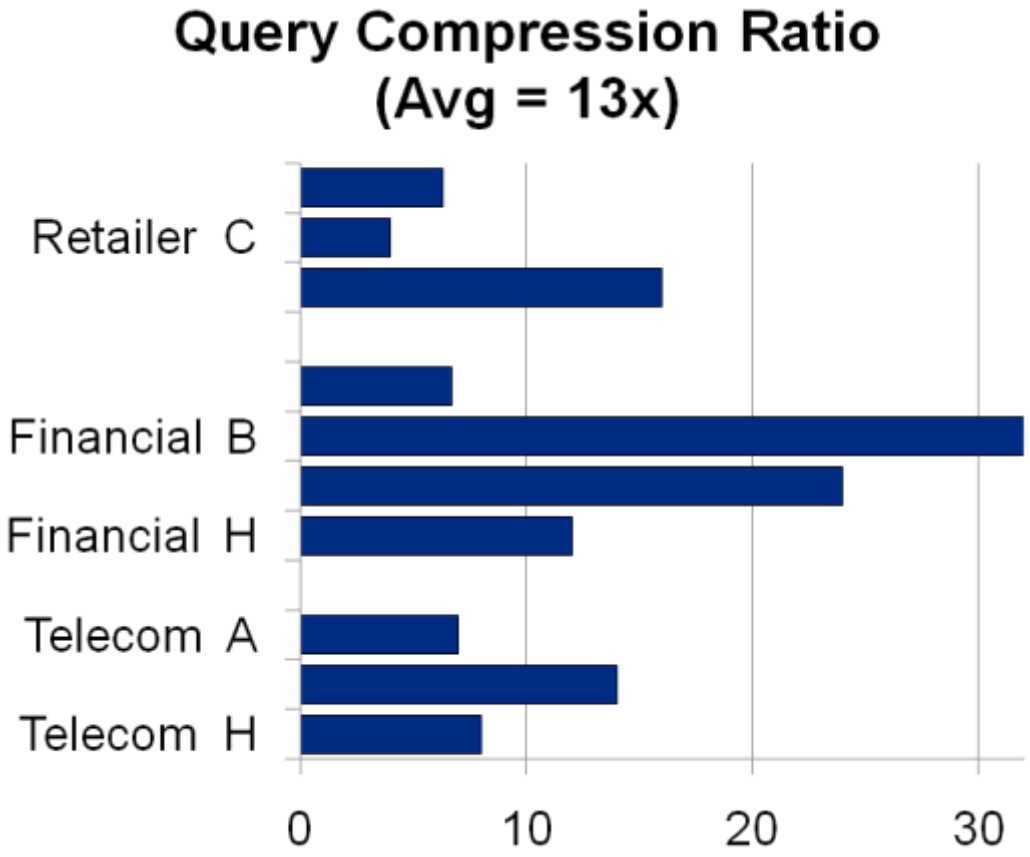
- Data is organized and compressed by column
  - Dramatically better compression
- Speed Optimized **Query Mode** for Data Warehousing
  - **10X compression typical**
  - **Runs faster because of Exadata offload!**
- Space Optimized **Archival Mode** for infrequently accessed data
  - **15X to 50X compression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

Faster and Simpler

Backup, DR, Caching, Reorg, Clone

Benefits Multiply

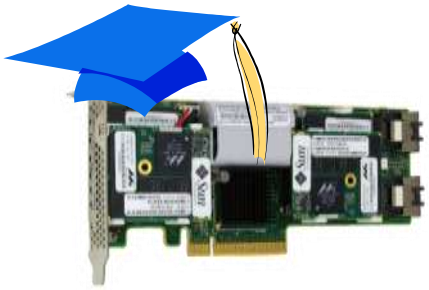
Compression Ratio of Real-World Data

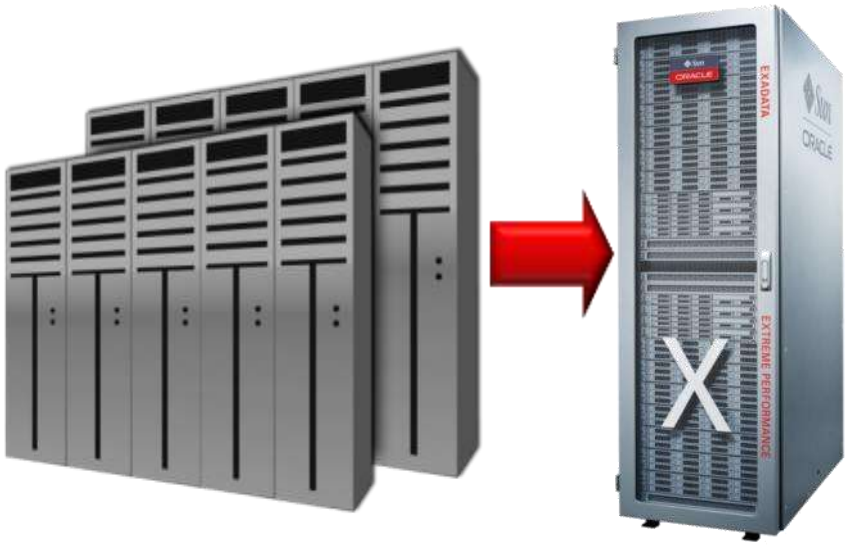


- Compression Ratio varies by customer and table
- Trials were run on largest table at 10 ultra large companies
  - Average revenue > $60 BB
- Average Query Compression ratio was 13x
  - On top of Oracle's already highly efficient format

Exadata Smart Flash Cache

Extreme Performance OLTP



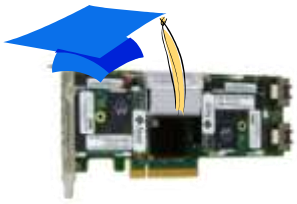


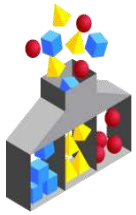
5X More I/Os than 1000 Disks in an Enterprise Storage Array

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

Exadata Flash Warehousing

Fastest Query Throughput

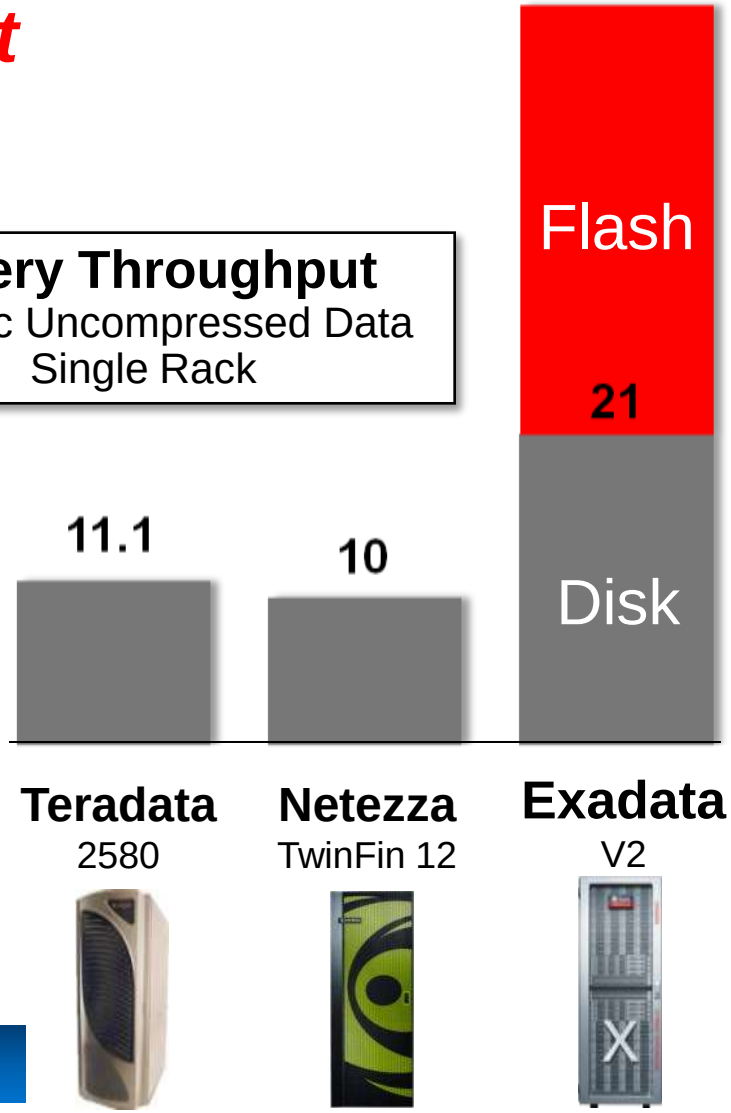




- **50 TB of data fits in Flash**
  - Using 10x Query Compression
- Effective Query Throughput on compressed data is even higher
  - **Hundreds of GB/sec**
- **Easily keep recent data in flash, older data on disk**

Business answers in seconds, not hours

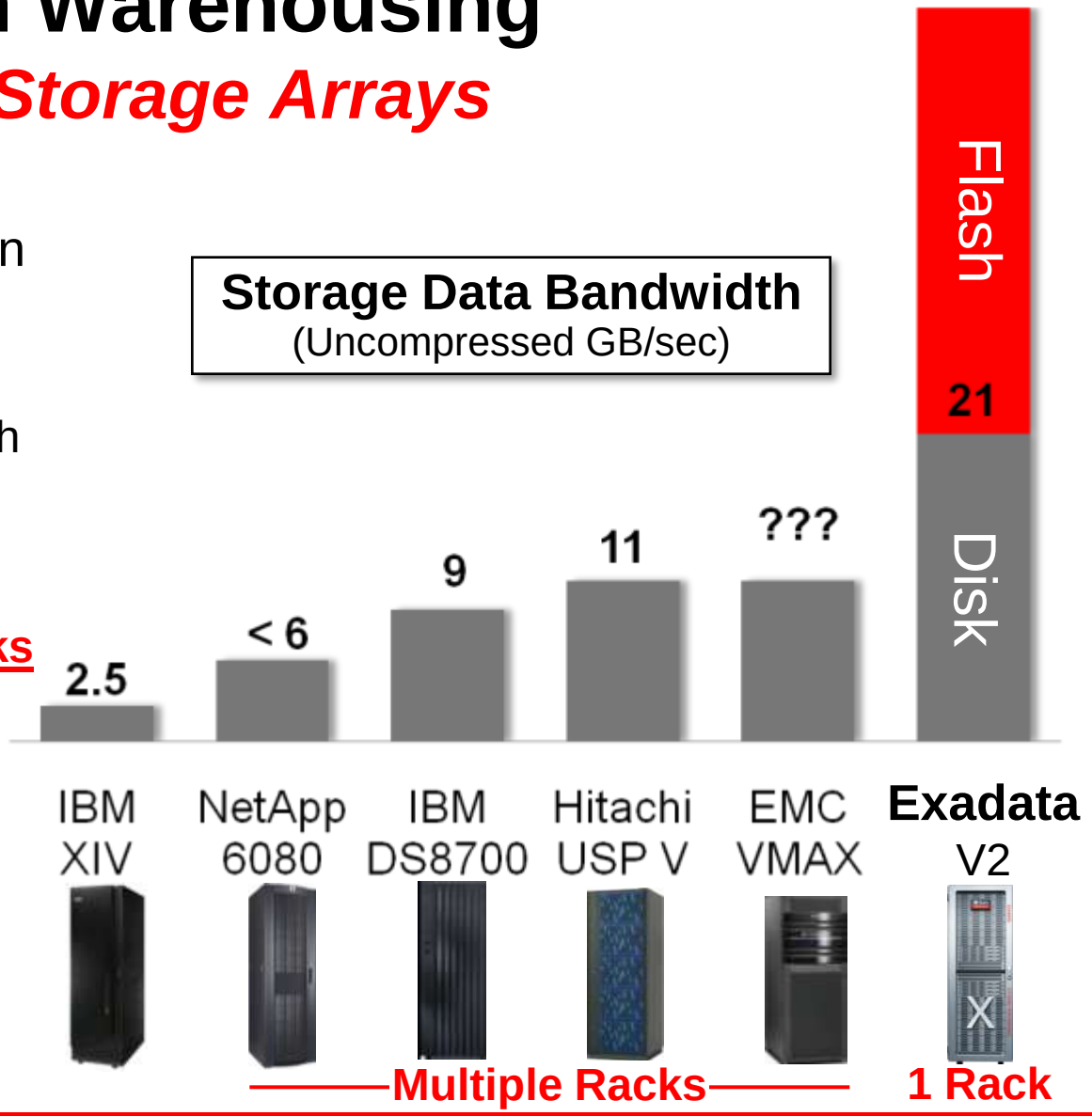
Query Throughput
GB/sec Uncompressed Data
Single Rack



Exadata Flash Warehousing

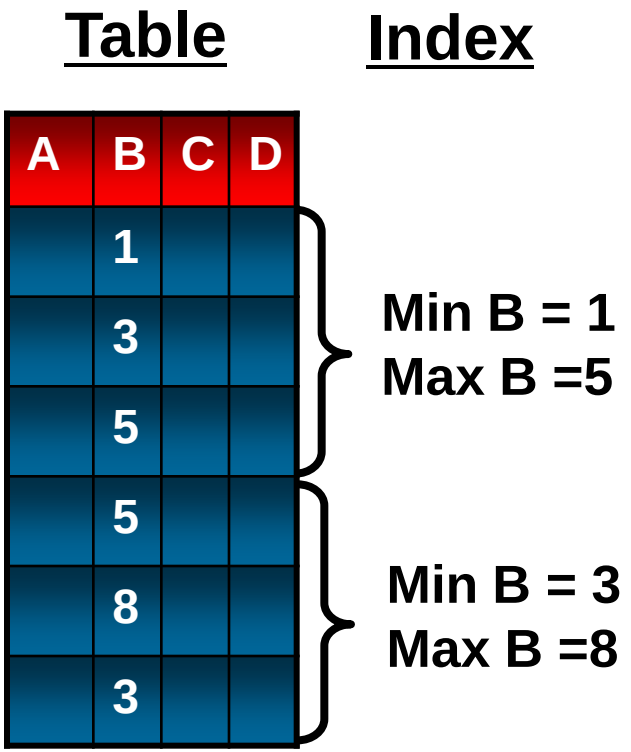
Comparison to Storage Arrays

- Storage Arrays bottleneck on back-end connectivity and controller performance
  - Flash provides no bandwidth increase
- **Exadata is fastest**
  - **and scales with more racks**
- Arrays don't scale and:
  - No CPU offload
  - No Columnar Compression
  - No InfiniBand
  - Expensive



Exadata Storage Index

Transparent I/O Elimination with No Overhead



- 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

Storage Index with Partitions Example

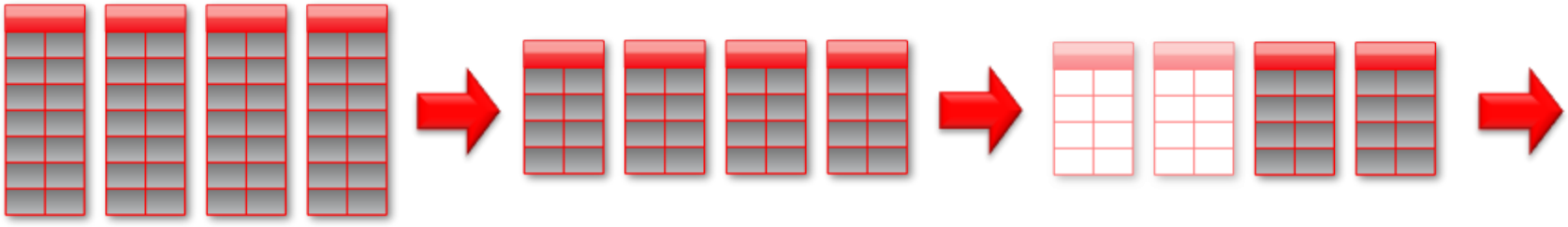
Orders Table

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

Benefits Multiply

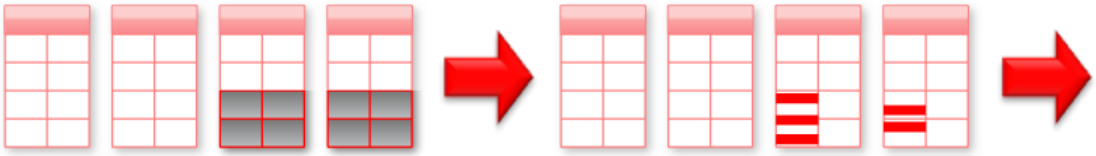
Example



10 TB of user data
Normally 10 TB of IO

1 TB
with 10x compression

100 GB
with partition pruning



50 GB
with Storage Indexes

Scan 50GB in **Flash**
10 GB returned after
Exadata Filtering

Seconds
On Exadata

Data is 10x Smaller, Scan is 2000x faster

Agenda

- Hardware Architecture
- Key Technologies
- **Consolidation & Protection**
- Migration

Best Machine for Database Consolidation

Same Day Database Deployments on Exadata Farm/Cloud

- Consolidation is key to reducing costs
  - Administration, hardware, software, data center
- Many databases can be consolidated on Exadata
  - Multiple small databases within a node
  - Large databases can span nodes using RAC
  - Exadata serves as farm/cloud for databases
- Exadata **delivers extreme performance for complex workloads that mix OLTP and DW**
  - Complex OLTP with batch and reporting
  - Complex Warehousing with thousands of users
  - Multiple databases running different applications

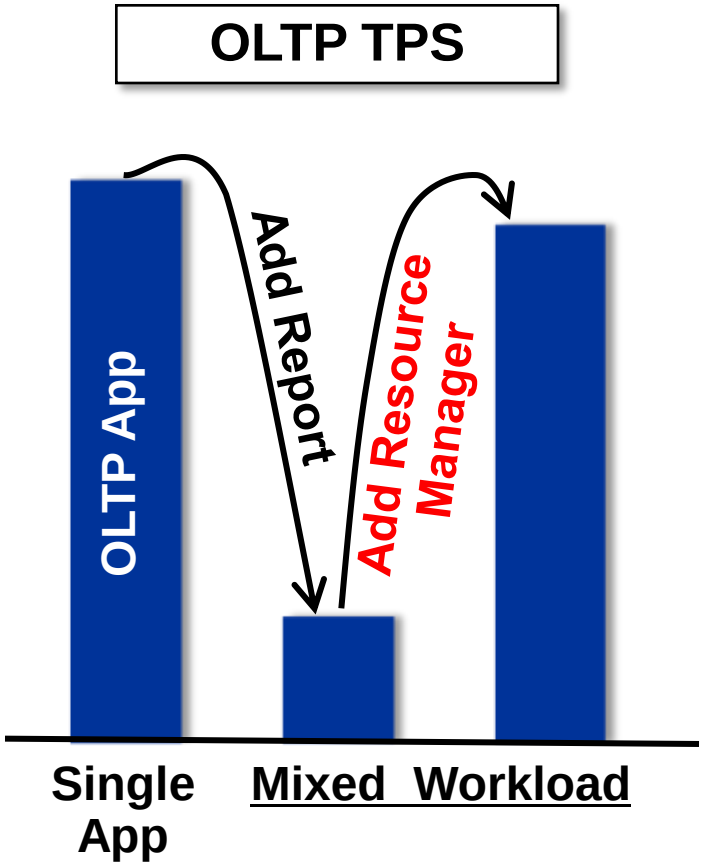
Consolidate Database Storage

- Exadata and ASM allow all storage servers to be shared across databases
- Shared Configuration
  - Advanced ASM data striping spreads **every database across all storage servers**
  - Eliminates hot-spots and captive unused space
  - Full storage grid performance available to all databases
  - Database or cluster level storage security
- Predictable Performance
  - **Exadata I/O resource manager** prioritizes I/Os to ensure predictable performance
    - At user, job, application, or database level
  - No need for isolated storage islands

Consolidate Database Servers

- Many databases can run on Database Machine servers
- Shared Configuration
  - Applications connect to a database **service** that runs on one or more database servers
    - Services can **grow, shrink, & move** dynamically
  - Large databases can **span nodes** using RAC
  - Multiple small databases can run on a single node
- Predictable performance
  - **Instance caging** provides predictable CPU resources when multiple databases run **on the same node**
    - Restricts a database to subset of processors

Performance Protection



- Resource Manager reserves and prioritizes I/O and CPU across and within databases
- **Unique ability to prioritize by dynamic attributes**
  - Workload, Job, User
- In addition to static
  - Database, Application

Failure Protection

RAC

ASM

Secure Backup

Flashback

Online Redefinition

WAN

Active Data Guard

GoldenGate Replication

- Redundant Hardware
  - Servers, Storage, Network
- Database Level HA
  - Tolerate failures and changes

- **Real-Time Active Replica**

Agenda

- Hardware Architecture
- Key Technologies
- Consolidation & Protection
- **Migration**

Migration is quite simple

- Target is a standard 11g release 2 database
  - Which uses ASM
  - Happens to be stored on Exadata Cells
- All DBA knowledge can be (re-)used
  - Make sure your DBA's have up to date (11gR2) knowledge
    - Besides Exadata knowledge
  - New versions might have new options (Data Pump etc)
- All best practices for 11g release 2 are valid for DBM
  - Check MOS note 785351.1
    - 11g release 2 Upgrade Companion

Non-Oracle to Exadata

- Source system is non-Oracle
  - Check default options for migration to Oracle
  - http://otn.oracle.com/migration
    - Oracle Migration Workbench
    - SQL*Developer
    - Flat file loading
    - Third Parties
- Some partners are very specialized
  - Example: Teradata to Oracle migrations
    - Inspirer / SQLWays Migration or Oracle

Oracle to Oracle

- Your target is always the same
  - Oracle 11g release 2 on Linux 64 bit (with RAC ?)
  - ASM based storage
  - Little Endian storage format (Intel x86)
- Sources are usually not on 11g Release 2
  - Application needs to be validated against 11g Release 2
  - Application needs to be certified to use ASM
  - Database needs to be upgraded
- Sources are usually on non-Linux OS
  - Source is usually on Big Endian (Risc, Itanium)
- Databases are build without latest features and options
  - Block size, Result Cache, Active Data Guard, Real Application Clusters, Exadata Columnar Compression etc etc etc

Prepare your migration

- Get the application certified
  - Upgrade your test database to 11g Release 2
  - Test your application until you are satisfied
  - Oracle VM is your friend
    - If you do not have a test environment
    - If you do not have a RAC enabled test environment
    - Easy to replay steps
- Think about more than the application
  - How about backup and recovery
  - How is the database loaded with data after the migration
    - Batch jobs
    - Reporting tools

Perform Capacity Planning

- Understand the Capacity of the current environment
  - Usable Database Size
  - IOPS (I/Os per second)
  - MBPS (Megabytes per second)
  - Use AWR to get statistics
- Size the Exadata Storage Cells
  - Size for both - Performance and Capacity
- Size for worst case
  - Ensure that the post failure IO Capacity is sufficient to tolerant failures
  - Enough Capacity after failures to have the same redundancy
  - Enough performance after failures (MBPS and IOPS)

Two migration options

- Physical Migration
  - Move entire table spaces and data files
  - Do not look at the data inside
  - Easiest way to migrate (if it is possible)
- Logical Migration
  - Get the data out
    - Data pump
    - Import/Export
    - Dump to files
    - Golden Gate / Streams
  - More difficult to setup
  - More flexible

Physical Migration

- If source is on Linux and 11.x
  - Data Guard Physical Standby
  - Transportable Database
  - Restore backup to Exadata
- If source is on >10.2
  - Cross-Platform Transportable tablespaces
    - Possible need of an Endian Change (using RMAN) on intermediate storage
- If source is <10.2
  - Upgrade first to 10.2
  - Or choose for Logical migration

Physical Migration

- Upside
  - No worries that all data is migrated
  - Easy to setup and execute (just a few files)
- Downside
  - Copy more data than needed
  - No best practices to apply
  - Data in target is as good or as bad as in source
  - No changes possible to setup
    - AU size of ASM (4Mb)
    - Large extends (8M)
    - Compression
    - Partitioning

Logical Migration

- Get the data out of the source system
  - Data Pump
  - Import / Export
  - CTAS over database link
  - Dump to text files and External Tables
- Synchronize the data
  - Golden Gate
  - Streams

Logical Migration

- Upside
  - Less data to be moved
    - No empty parts of the data file and table spaces
    - Only data, not the indexes
  - Works on versions of v7 and up
  - Lowest downtime option (Golden Gate)
  - Changes possible, apply best practices
  - Loading can be very fast (5+TB/hour)

- Downside
  - Check if all data is there
  - Takes longer to setup
  - Needs to be monitored

How long does it take ?

- Bottleneck is the network
  - How to get it to the Database Machine
- Gigabit Ethernet
  - Maximum 125MB/sec = 0.4TB/hour
  - Remember that you have 8 nodes, use them
  - Bundling Ethernet is possible (we have 3 per system)
- Fiber channel
  - No hardware changes allowed on the Database Machine
- Infiniband
  - Supply your old environment with Infiniband
  - Between 1.5 and 2.5GB/sec per link (5TB/hour)
    - We have 8 database nodes
    - Think about the number and speed of source nodes

What is my downtime ?

- How much time do you have ?
  - SLA's, Maintenance Windows
  - Fallback scenario
  - Weigh speed versus best practices
- Least downtime
  - 11g on Windows / Linux to Database Machine
    - Physical Standby
  - Other situations
    - Golden Gate (where supported)
    - Oracle Streams
- The less the downtime, the more the preparation !

Best Practices

- Take time to decide which approach you will take
  - Take a look at all options
  - Decide the best for your environment
  - Implement as much recommendations as possible
    - 4MB ASM AU size
    - Large extents (8MB)
    - Table compression
    - Partitioning scheme
    - 8k or 16k block size

Best Practices

- Get a step by step script
  - Make sure all area's of the migration are covered
  - All steps documented and fool-proof
    - There might be some Oracle bugs you need to find
  - Time every step to make sure it fits in the time window
    - "Your mileage may vary" compared to theoretical values
- Think about a fall-back strategy
  - What happens in case the unforeseen happens ?

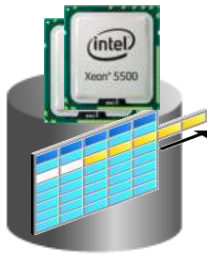
TEST, TEST, TEST !

Conclusion - Exadata V2

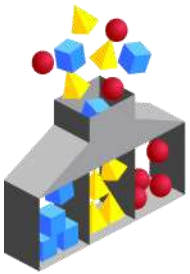
Ideal Database Platform

- **Best for Data Warehousing**
- **Best for OLTP**
- **Best for Database Consolidation**
- **Migration is not Rocket Science**

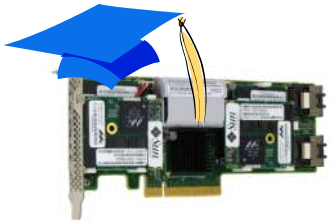
Fastest, Lowest Cost



Intelligent Storage Grid



Hybrid Columnar Compression



Smart Flash Cache

Exadata in real life

- Who gets to administer the system ?
  - Oracle DBA ?
  - Storage admin ?
  - System admin ?
- How do we backup the environment ?
  - Existing backup environment ? Is it fast enough ?
- How do we architect the system ?
  - S.A.M.E. ? High Availability ? Disaster Recovery ?
- How about test and development ?

Exadata Proof-of-Concepts

- Take your 10 worst performing queries..
  - Or your batch job
  - Or your Real Application Testing recording
  - Or ….
- And run it in a POC Lab
  - Oracle Solution Center (former ETC) Reading (UK)
  - VX Company (Baarn) or Atos Origin (Groningen)
- All situations Exadata was faster by default
  - 11.2 software, Exadata options, good hardware
  - Most of them were faster (5-1000x)

Exadata ordering

- Sales is happy to assist you !
- Some preparation needs to be done before ordering..
  - Does the system fit into the datacenter ?
  - Can we get the machine in the datacenter ?
  - Enough cooling, power ?
  - Which connectors are needed ?
  - *Oracle (Sun) engineer can assist with this..*
- Some preparation needs to be done before delivery..
  - IP ranges
  - Physical network connections
  - *Oracle engineer can assist with this..*

Exadata delivery and install

- Order time is usually 6-8 weeks
  - Systems are assembled in-factory and pre-tested
  - Shipped fully assembled
  - Needs 1-2 days for acclimatizing
- Physically connect the system is easy (~ 1 day)
  - Power connection hookup
  - Network preparation (no hookup yet !)
- Installation of Exadata and Database software (~2-3 days)
  - Storage is pre-imaged but might need the latest patches
  - Host config and DB software (including patches) install is automatic
    - Setup and install using special tool (OneCommand)

Exadata first usage

- During install (by Oracle) initial database is created
  - Storage is configured for Data and Recovery (according to survey)
  - Initial DB is created using special Exadata template
- After the physical setup and software install the system is ready
  - Start using the initial database (production setup)
  - Create other databases according to need
- Changing storage and DB setup is very easy

Exadata to production

- Migration is easy, but slow and error-prone
  - Human nature is one of the bottlenecks
  - Existing network speed (connection old-new system) is bottleneck
- Take enough time to test application on new system
  - NL customer did not do enough testing and ran into a (generic 11.1) bug
  - Next migration he doubled his effort on application and had no issues
- Think about fall-back scenario is usually forgotten or not useable
  - NL customer did setup fallback but throughput of network of old system could not keep up
  - After a week running 4-5 days behind

Exadata in day-to-day usage

- Databases running smooth
- Normal Oracle database issues
  - Tablespaces growing
  - Patches need to be applied
  - Performance needs to be monitored
- If something goes wrong, log an SR
  - PLEASE use the Exadata SR number
  - For hardware issues, usually a phone call is also needed
  - Automated Service Request system can be implemented for hardware
  - Oracle engineer will replace all hardware on-site (24/7 2 hour response)
    - (if premier support was purchased)

Exadata performance and tuning

- Usually application runs between 2x and 50x faster without tuning
  - What happens if you drop the indexes ?
  - Tuning your environment like any other new system !
  - New to RAC ? Should have tested !
  - Oracle Enterprise Manager is your friend !
    - Free Exadata and OS monitoring *(if support has been purchased)*
    - Various tools to investigate and optimize *(needs EM packs at additional $)*
- Create tablespace in 30 seconds instead of 15 minutes
  - (25GB tablespace on a customer system with and without Exadata)
- Incremental backups without usual overhead

Resources

- **Oracle.com:**
  http://www.oracle.com/exadata
- **Oracle Exadata Technology Portal on OTN:**
  http://www.oracle.com/technology/products/bi/db/exadata
- **Oracle Exadata white papers:**
  http://www.oracle.com/technology/products/bi/db/exadata/pdf/exadata-technical-whitepaper.pdf

  http://www.oracle.com/technology/products/bi/db/exadata/pdf/migration-to-exadata-whitepaper.pdf

Resources continued…

- Oracle Exadata Migration white paper: www.oracle.com/technology/products/bi/db/exadata/pdf/migration-to-exadata-whitepaper.pdf
- MyOracle Support (Metalink) Note: 762540.1: "Consolidated Reference List for Migration Upgrades"
- MyOracle Support (Metalink) Note: 785351.1:
- "11g Release 2 Upgrade Companion"


QUESTIONS
&
ANSWERS


ORACLE®