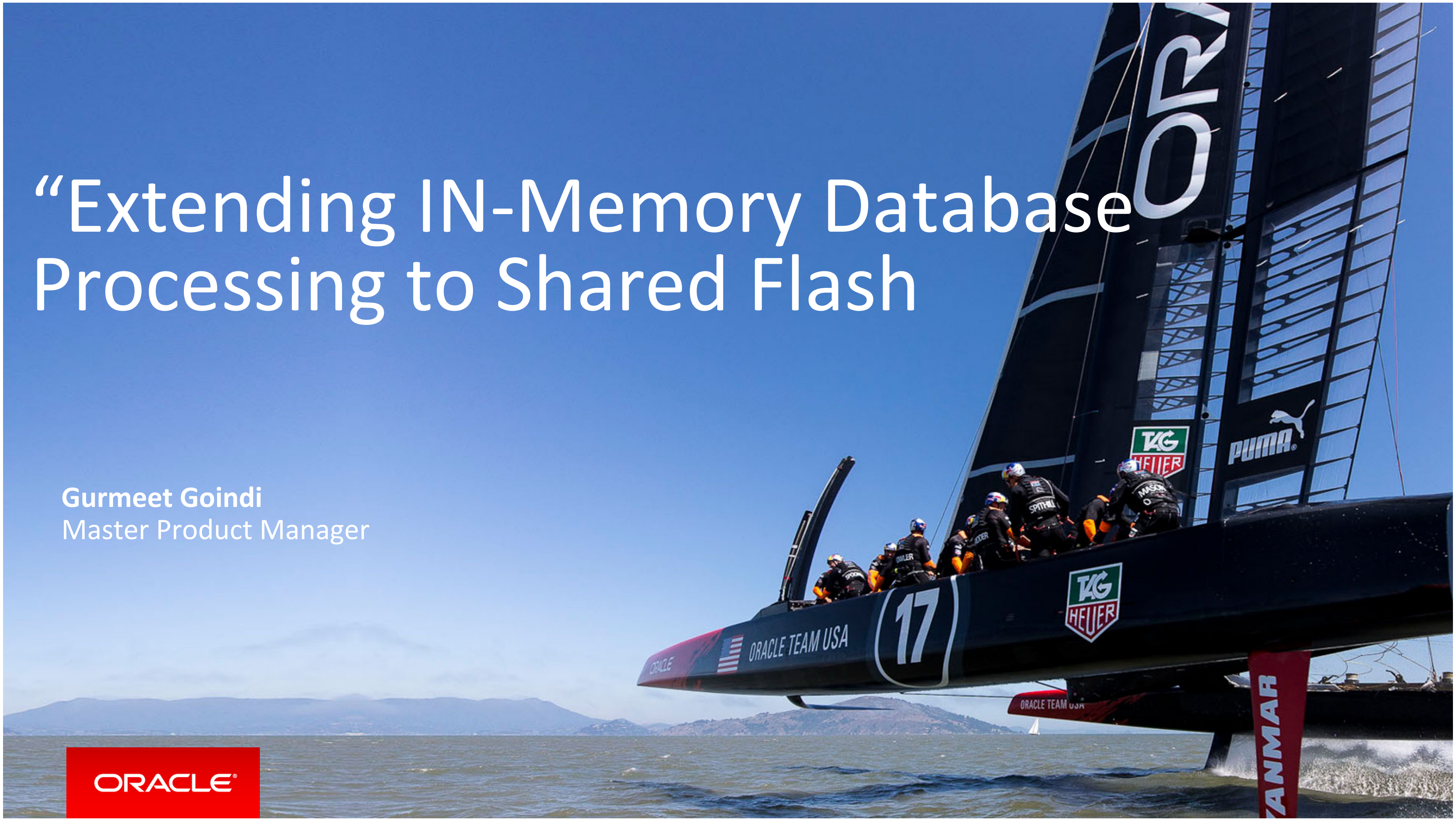
“Extending IN-Memory Database Processing to Shared Flash

Gurmeet Goindi
Master Product Manager



Safe Harbor Statement

The following is intended to outline our general product direction. It is intended for information purposes only, and may not be incorporated into any contract. It is not a commitment to deliver any material, code, or functionality, and should not be relied upon in making purchasing decisions. The development, release, and timing of any features or functionality described for Oracle's products remains at the sole discretion of Oracle.

Exadata Database Machine

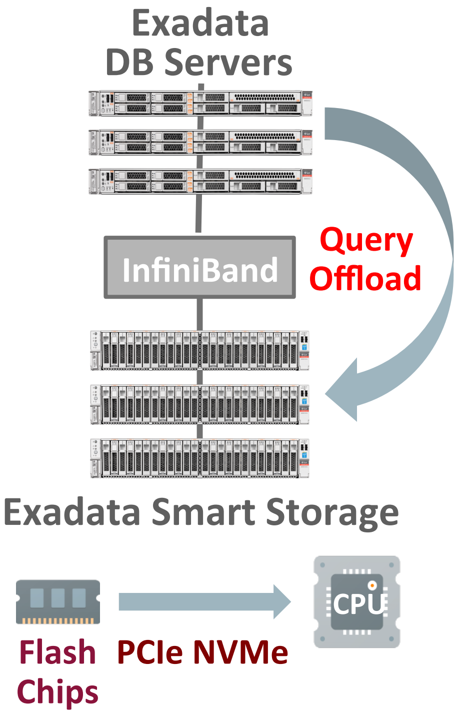
Performance, Availability and Security

Best Platform for Oracle Databases on-premises and in the Cloud

Enabled by:

- Single-vendor accountability
- Exclusive focus on databases
- Deep h/w and s/w integration
- Revolutionary approach to storage

Exadata Achieves Memory Performance with Shared Flash



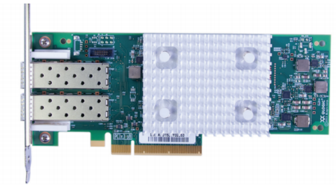
- Exadata X6 delivers 300GB/sec flash bandwidth to any server
  - Approaches 800GB/sec aggregate DRAM bandwidth of DB servers
- Must move compute to data to achieve full flash potential
  - Requires owning full stack, can't be solved in storage alone
- Fundamentally, storage arrays can share flash capacity but not flash performance
  - Even with next gen scale-out, PCIe networks, or NVMe over fabric
  - e.g. new EMC DSSD has 3-6 times lower throughput than Exadata X6
- Shared storage with memory-level bandwidth is a paradigm change in the industry
  - Get near DRAM throughput, with the capacity of shared flash

NVMe PCI-e Flash Disrupts the Storage Array Model

New improvements are causing 100X bottlenecks across shared storage stack

Latest PCIe Flash

5.4 GB/sec



SAN Link = 40Gb

5 GB/sec

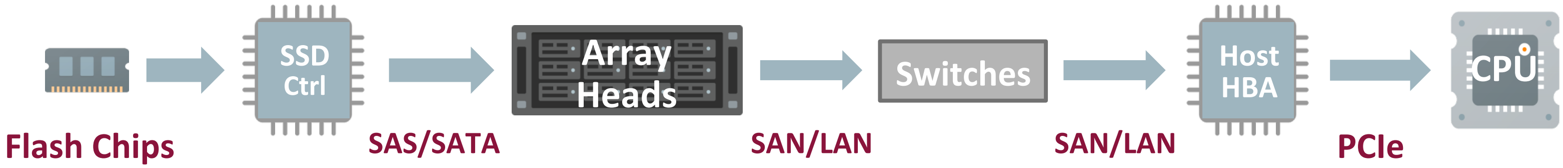
Less than 1 Flash card

Leading All Flash Array

24 GB/sec

Less than 5 Flash card

All-Flash Storage Array IO Path: many steps, each adds latency and creates bottlenecks

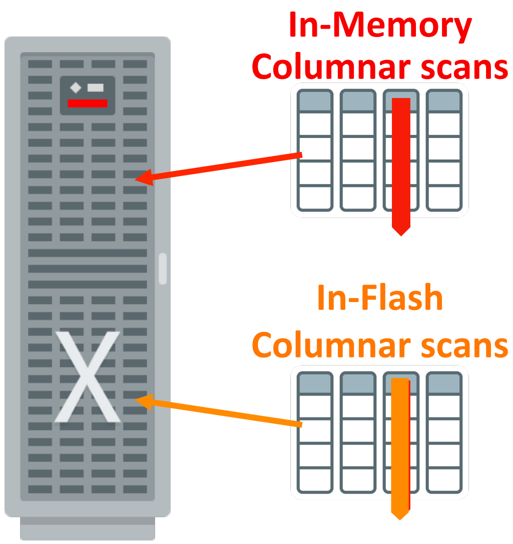


ORACLE

Redesigning Scan Offload for Memory Throughput

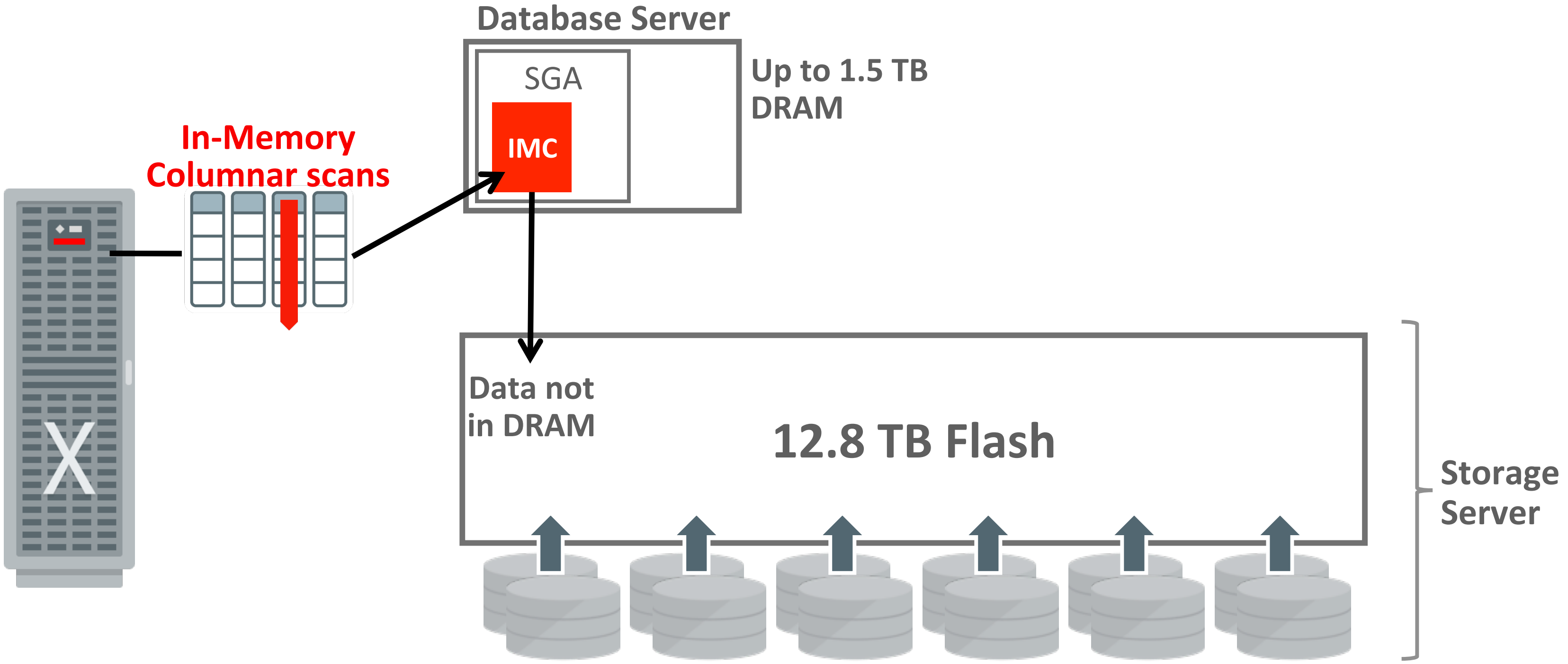
NEW IN 12.2

- With Exadata Flash throughput approaching memory throughput, SQL bottleneck moves from I/O to CPU
- Exadata will automatically transform table data into In-Memory DB columnar formats in Exadata flash cache
  - Dual format architecture extended from DRAM to flash
- Enables fast vector processing for storage server queries
  - Smart Scan results sent to DB using In-Memory Columnar format to reduce DB CPU usage
- Uniquely optimizes next generation flash as memory



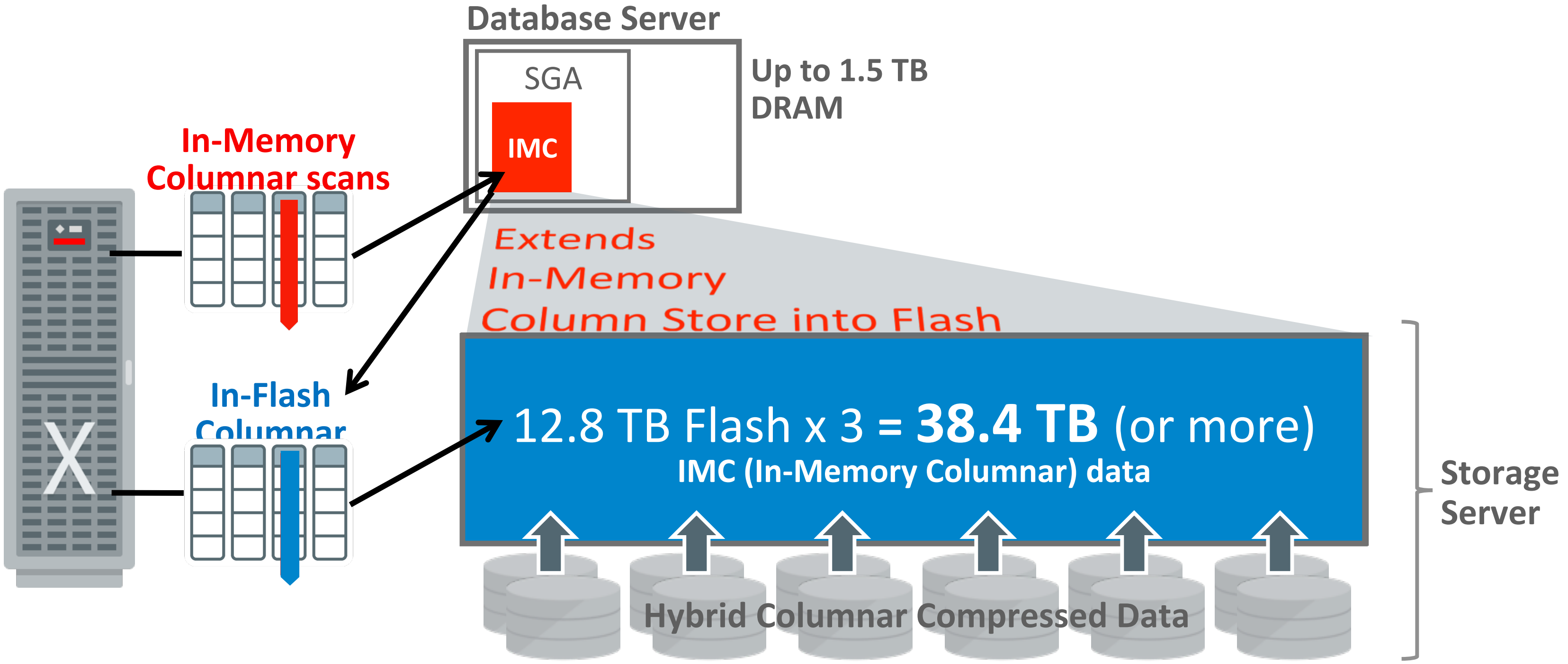
In-Memory Columnar Formats in DRAM (pre 12.2.1.1.0)

Super-Fast Scans from Memory, but All Queries Complete



In-Memory Columnar Formats in Flash Cache (12.2.1.1.0)

3 - 4x Overall Analytics Performance Improvement



Smart Analytics: Join and Aggregation Smart Scan

NEW
DB 12.2

- Extend In-Memory Aggregation technique into storage (vector joins and vector aggregation)
- Find Sales per country

```
SELECT /*+ VECTOR_TRANSFORM  */ country_id,
sum(amount_sold) amount_sold
FROM customers, sales
WHERE customers.cust_id = sales.cust_id
GROUP BY customers.country_id
ORDER BY customers.country_id;
```

- Storage cells scanning sales fact table return tuples

  `{country_id, sum_amount_sold }`

- Join and Aggregation offloaded to the storage server

Smart Analytics: More Smart Scan Enhancements

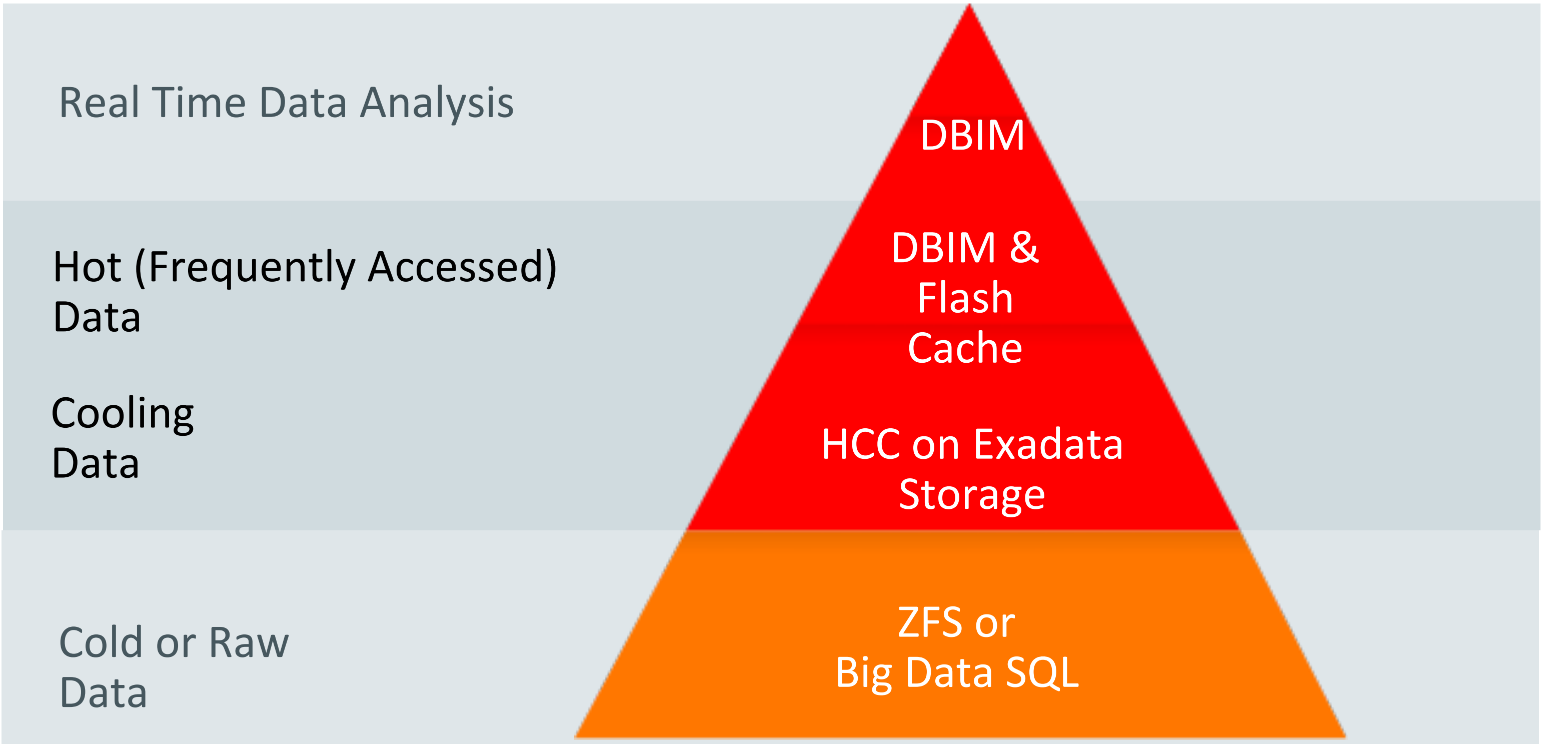
NEW
DBIM 12.2

- Smart Scan enhancements for XML and JSON
  - JSON_EXISTS, JSON_VALUE, JSON_QUERY,

  "IS JSON" and "IS NOT JSON"
  - XML: XMLExists, XMLCast(XMLQuery())
- Significant speedup in JSON analytic workloads

```
select count(*)
 from pictures
 where json_value(photo, ‘$.tag’)
 like ‘%spain%’;
```

Data Tiering



Real Time Data Analysis

DBIM

Hot (Frequently Accessed) Data

DBIM & Flash Cache

Cooling Data

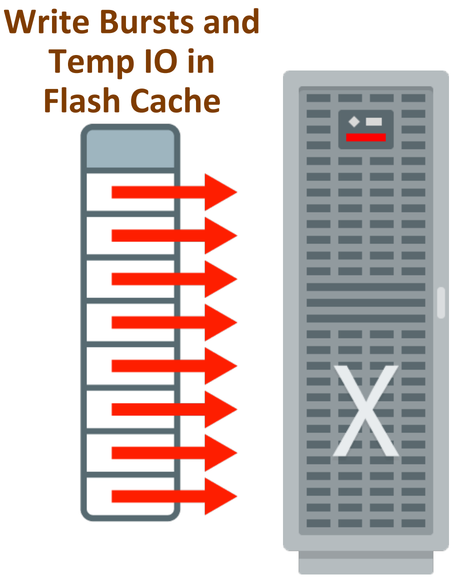
HCC on Exadata Storage

Cold or Raw Data

ZFS or Big Data SQL

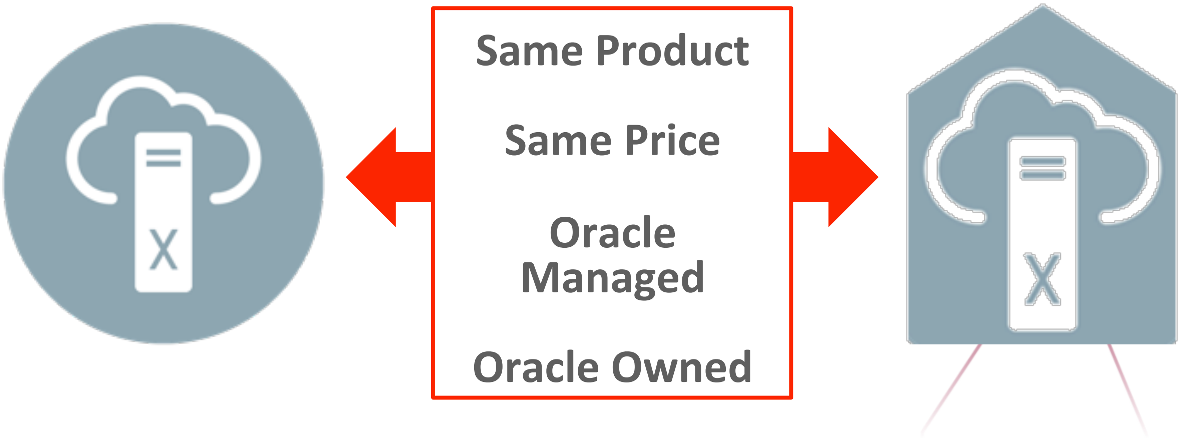
Smart Analytics: Smart Write Bursts and Temp IO in Flash Cache

- Write throughput of four flash cards has become greater than the write throughput of 12-disks
- When database write throughput exceeds throughput of disks, Smart Flash Cache intelligently caches writes
- When queries write a lot of temp IO, Smart Flash Cache intelligently caches temp IO
  - Writes to flash for temp spill reduces elapsed time
  - Reads from flash for temp reduces elapsed time further
- Smart Flash Cache prioritizes OLTP data and does not remove hot OLTP lines from the cache
- Smart flash wear management for large writes
- Supports Database 11.2.0.4, 12.1.0.2 and 12.2.0.1



Accelerates Large Joins and Sorts and Large Data Loads

Exadata Cloud – Your Way



Exadata Cloud Service

Exadata Cloud Machine

Same Product

Same Price

Oracle Managed

Oracle Owned

Oracle Public Cloud

Cloud at Customer

Exadata Customer Case Studies

NTT docomo : MoBills（Mobile Billing System）

Benefits

	Faster Billing Processing	Maximum Availability	Reduced Operational Cost	Reduced Introduction Cost	Data Center Cost Savings
"MoBills is a very important position as a mission-critical system to promote efforts toward the realization of “+d”. Oracle Exadata is running very stable as a expected performance. We will continue to use the "Oracle Exadata" and we would like to establish a further advantage for our business." - Shimamura, Manager, Information System Department, NTT docomo	10X speedup 3 million SQL /sec	Local & Remote Standby 24 7	50%	25% $	90% Space Reduction

Business Objectives

- Real-Time Billing Platform for 66 million customer
- Dramatically improve performance and availability
- Reduce cost and complexity

Solution

- Oracle Exadata : 30 racks
- Oracle MAA (RAC / Active Data Guard - Local & Remote Standby database)

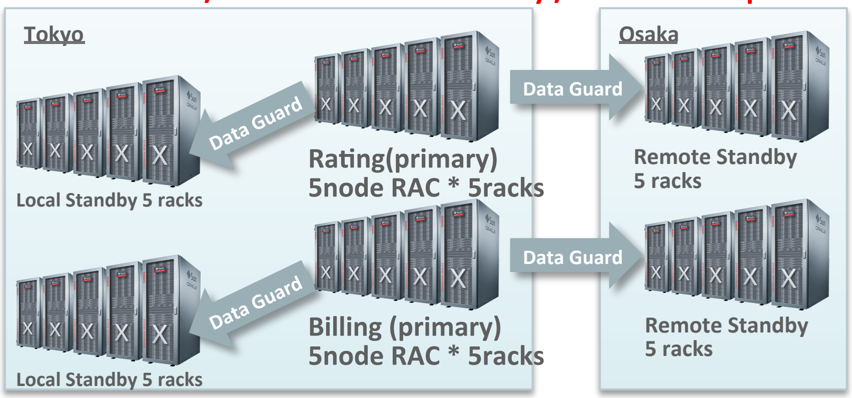
Pre-Exadata

- **Real-Time Billing Processing**
  - **High-end SMP Server**

  + High-end Storage : 350 racks
  - **Storage Mirror Backup**
- **Storage Mirror Replication**
- **Oracle 9i Database Release 2**

Exadata MAA

30 racks / Local & Remote Standby / RMAN backup



DCM Holdings : System Consolidation of 3 companies

Benefits	Faster Batch Processing Reduced Introduction Cost		High Consolidation Ratio Improved Manageability	Simplified Support
Realized the Database consolidation and integration due to the high performance provided by Oracle Exadata. And, Oracle Database 12c Multitenant Architecture also achieved high consolidation ratio while maintaining the independence of each group companies. *Platinum Service could improve the service level, Oracle Full-stack products could provide One-Stop Support.*	Standardization 2X speedup	40% Off	Multitenant Architecture 6DBs Consolidation	Oracle Full Stack Non Stop Support

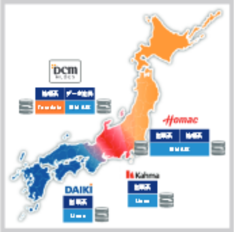
Business Objectives

- $10 billion Sales, Faster M&A
- High Consolidation ratio and improve service level
- Reduce operational cost

Solution

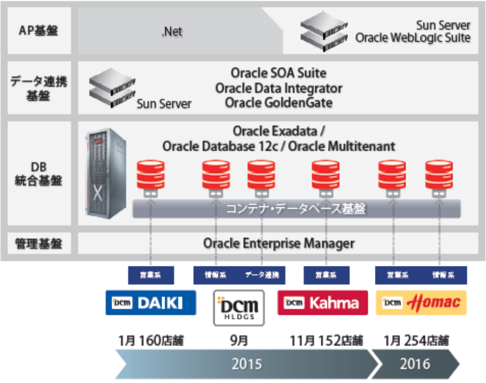
- Oracle Database 12c Multitenant on Exadata
- Oracle Full Stack (Middleware and Server products)

Pre-Exadata



Integration / Consolidation

Oracle Multitenant on Exadata



- **Consolidation and Integration 3 group companies (Homac, Kahma and DAIKI) of system infrastructure**
- **Replaced from IBM p Servers**
- **Teradata Migration to Exadata**
- **Oracle Database 12c Multitenant**
- **Platinum Service**
- **Zero down time System Migration by using GoldenGate**

Sprint: Call Data Record - Data Warehouse

Benefits

Faster Queries	Faster Reports	Storage Savings	Maximum Availability	Data Center Cost Savings
> 10x	24 X	6 x	No unplanned downtime	3:1 Consolidation
15 billion transactions/day	7 days to 7 hours	+ removed 150 TB Indexes		

> *"We reduced the queries from 30 seconds down to sub-second response time. Quick information, quick queries give Customer Care the ability to do their job better and meet the customer's needs."*
>
> - **Richard Ewald, Senior Technical Architect, Data Warehousing**

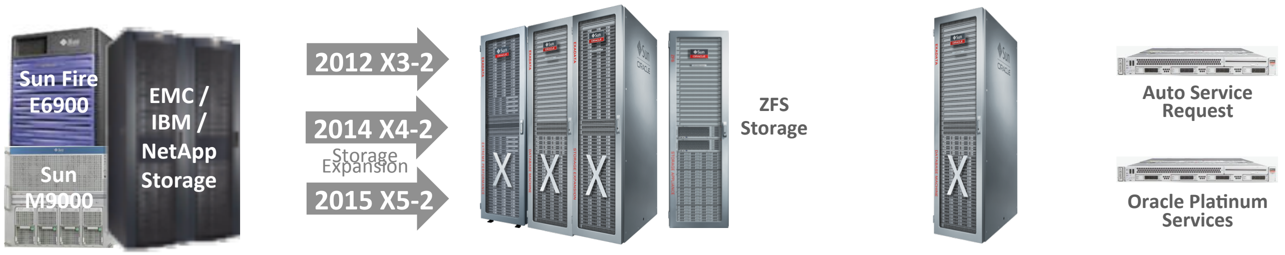
Business Objectives

- Improve performance
- Improve sustainability
- Improve availability and maintainability

Solution

- Full Rack (Prod), Half Rack (Dev/Test); ZFS
- Storage Expansion
- Half Rack (Prod)

Pre-Exadata | **Production** | **Dev/Test**



- **4 x Sun Fire E6900, 1 x M9000**
- **Mixed Storage**
- **Multiple backup systems**
- **90 Day CDR DW 1.15 PB**
- **Oracle DB 11gR2**

- **Exadata X3-2 Full Rack**
- **HCC: 950 TB to 150 TB**
- **ZFS Storage Appliance (Backup)**
- **Exadata Storage Expansion**
- **Exadata X5-2 Half Rack**

- **Exadata X3-2 Half Rack**

Pulte Group: Multitenant Consolidation

PulteGroup™

Benefits	Business Impact	Faster Applications	Lower Admin & Support Costs	Cost Savings
"Exadata delivered tremendous improvements in productivity. Users no longer have to wait for data. Data sharing is now real time." **- Brian Pawlik, IS Manager, Pulte Homes**	40% Productivity ↑ Monthly Close 2 Days Faster	2x -15x Faster	40% Reduction	40% CapEx

Business Objectives

- Scalability
- Supportability
- Sustainability

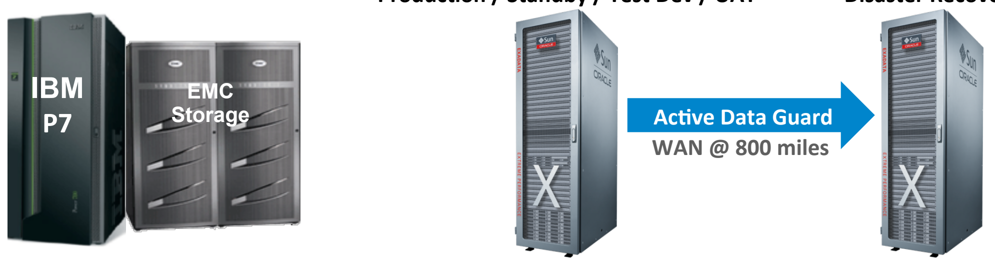
Solution

- quarter rack & eighth rack

Pre-Exadata

Exadata Quarter Rack
Production / Standby / Test Dev / UAT

Exadata Eighth Rack
Disaster Recovery



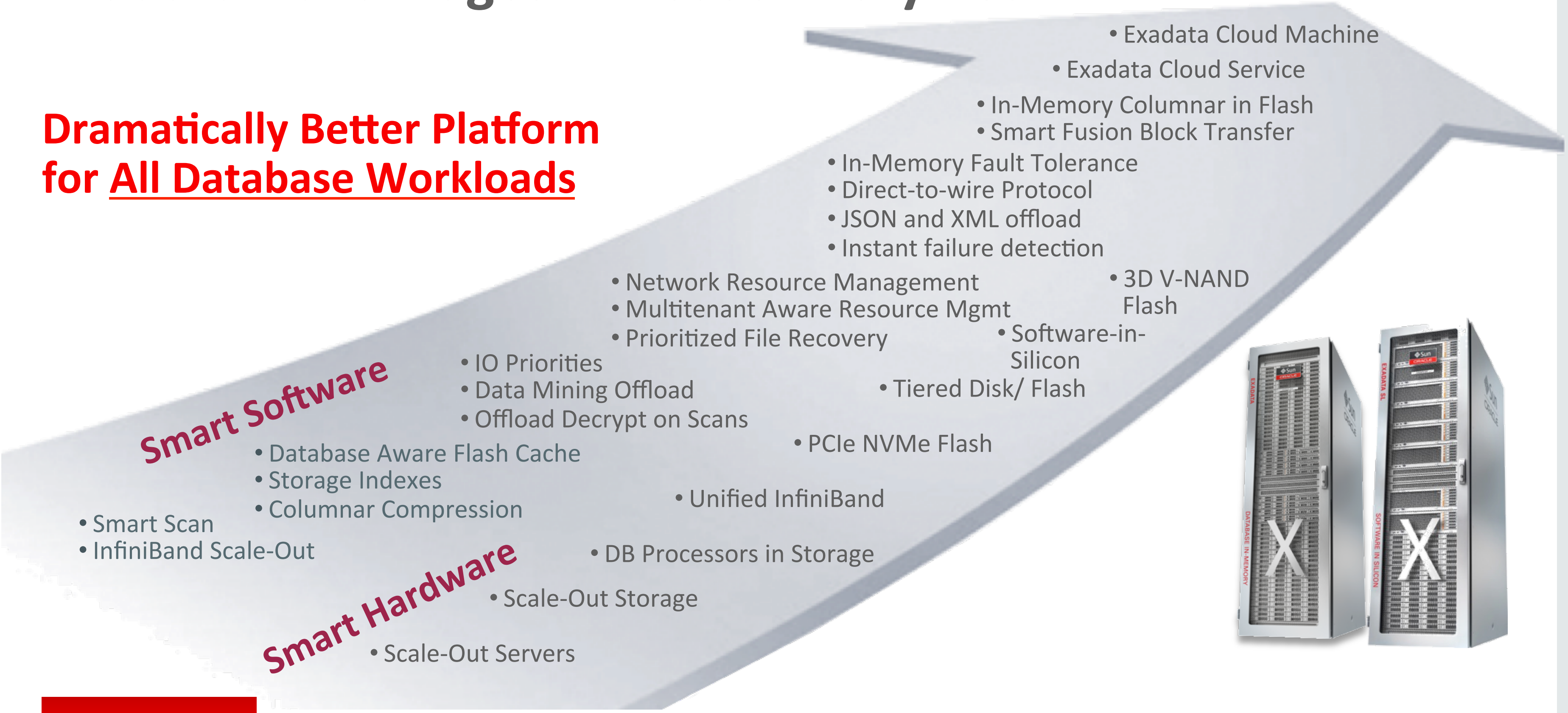
- **IBM P7**
- **EMC storage arrays**

- **Infor Lawson S3 ERP; Rebate Tracking**
- **Consolidate 35 DBs: 4 CDBs, 35 PDBs**
- **Production, Local Standby and QA**
- **Primary databases: > 5 TB**

ORACLE®

Exadata Advantages Increase Every Year

Dramatically Better Platform for All Database Workloads



- Exadata Cloud Machine
- Exadata Cloud Service
- In-Memory Columnar in Flash
- Smart Fusion Block Transfer
- In-Memory Fault Tolerance
- Direct-to-wire Protocol
- JSON and XML offload
- Instant failure detection
- 3D V-NAND Flash
- Network Resource Management
- Multitenant Aware Resource Mgmt
- Prioritized File Recovery
- Software-in-Silicon
- IO Priorities
- Data Mining Offload
- Offload Decrypt on Scans
- Tiered Disk/ Flash
- PCIe NVMe Flash
- Database Aware Flash Cache
- Storage Indexes
- Columnar Compression
- Unified InfiniBand
- Smart Scan
- InfiniBand Scale-Out
- DB Processors in Storage
- Scale-Out Storage
- Scale-Out Servers

Integrated Cloud

Applications & Platform Services

ORACLE®