Oracle Exadata: Achieving Memory-Level Performance with NVME Flash

Gurmeet Goindi
Master Product Manager

ORACLE®

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
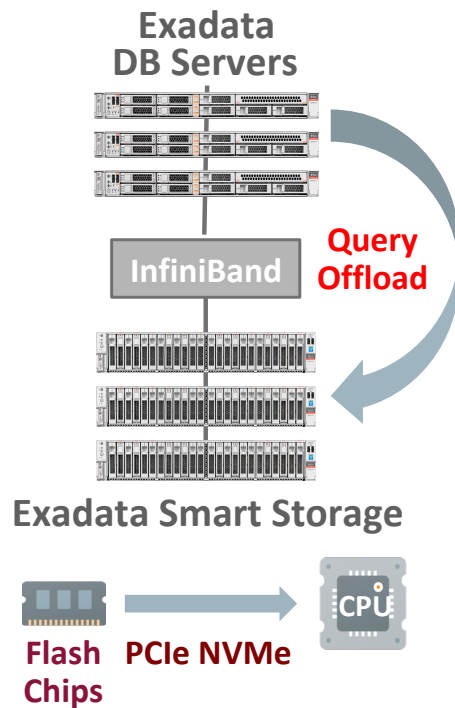
Performance, Availability and Security
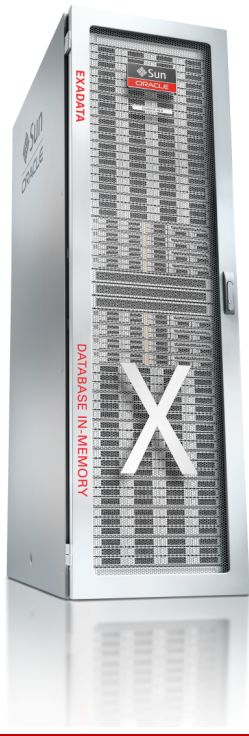
Best Platform for Oracle Databases on-premises and in the Cloud

Enabled by:

- Single-vendor accountability

- Exclusive focus on databases

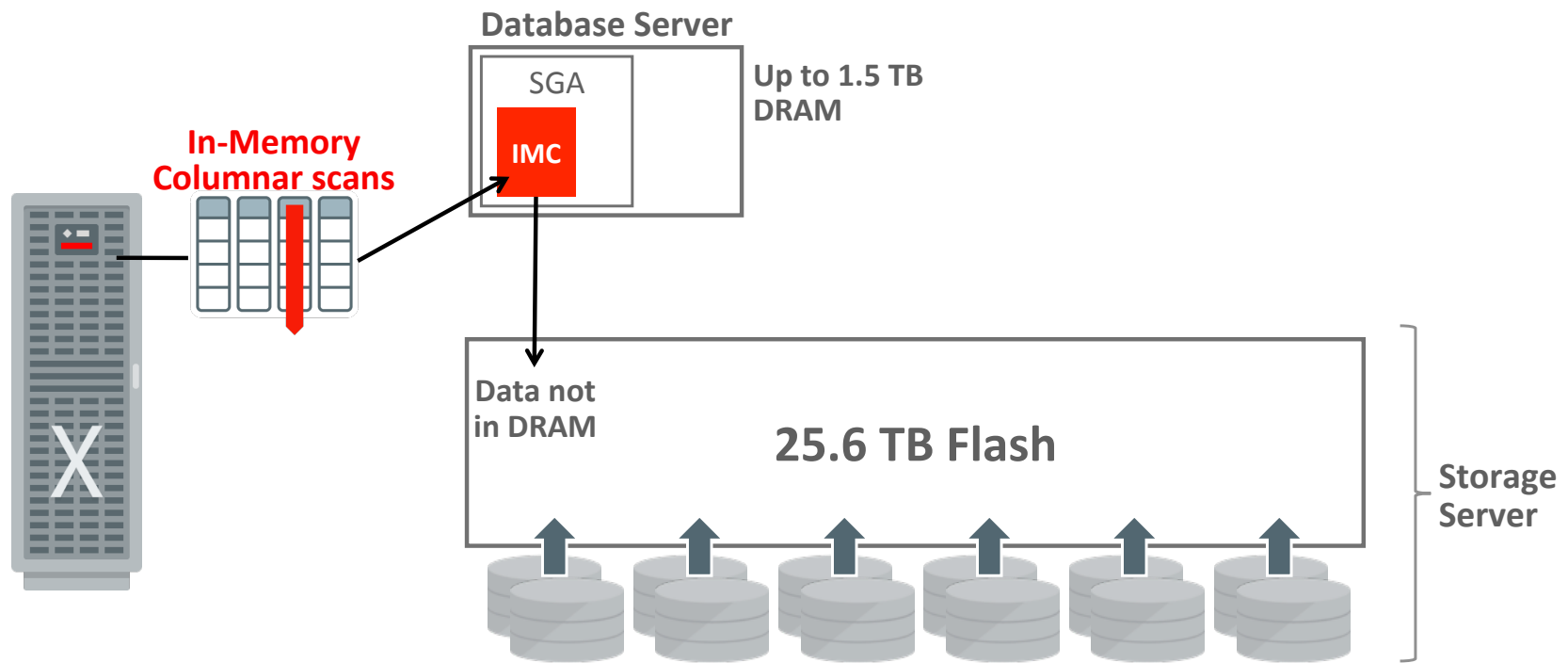- Deep h/w and s/w integration

- Revolutionary approach to storage

Exadata Achieves Memory Performance with Shared Flash

**Exadata
DB Servers**

InfiniBand

**Query
Offload**

Exadata Smart Storage

**Flash
Chips** **PCIe NVMe** CPU

- Exadata X7 delivers **350GB/sec flash bandwidth** to <u>any</u> 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 <u>capacity</u> but not flash <u>performance</u>
  - Even with next gen scale-out, PCIe networks, or NVMe over fabric
- Shared storage with memory-level bandwidth is a paradigm change in the industry
  - Get near DRAM throughput, with the capacity of shared flash

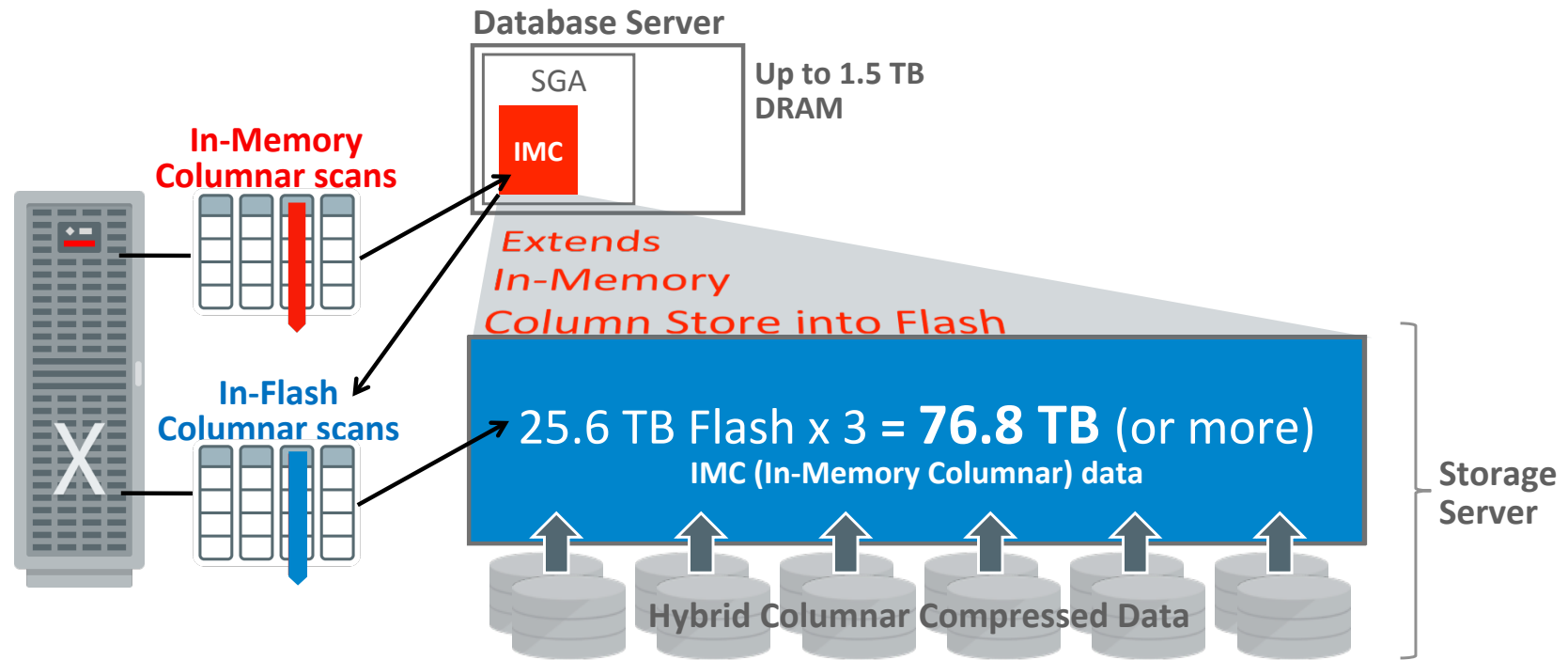# In-Memory Columnar Formats in DRAM
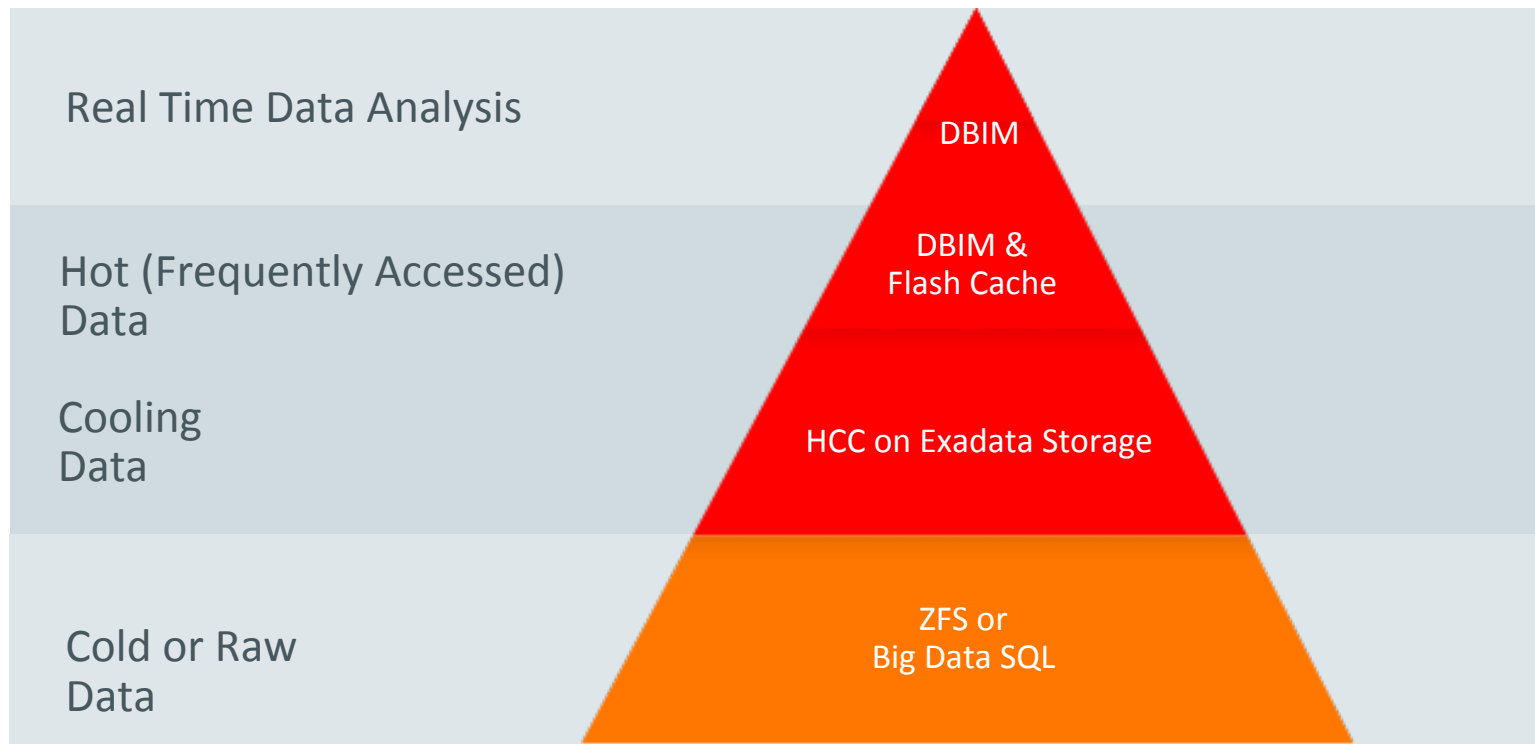## Super-Fast Scans from Memory, but All Queries Complete

Database Server

SGA

IMC

Up to 1.5 TB DRAM

In-Memory Columnar scans

Data not in DRAM

25.6 TB Flash

Storage Server

In-Memory Columnar Formats in Flash Cache (12.2.1.1.0)

3 - 4x Overall Analytics Performance Improvement

Database Server

SGA

IMC

Up to 1.5 TB DRAM

In-Memory Columnar scans

In-Flash Columnar scans

Extends In-Memory Column Store into Flash

25.6 TB Flash x 3 = **76.8 TB** (or more)

IMC (In-Memory Columnar) data

Hybrid Columnar Compressed Data

Storage Server

ORACLE®

Data Tiering

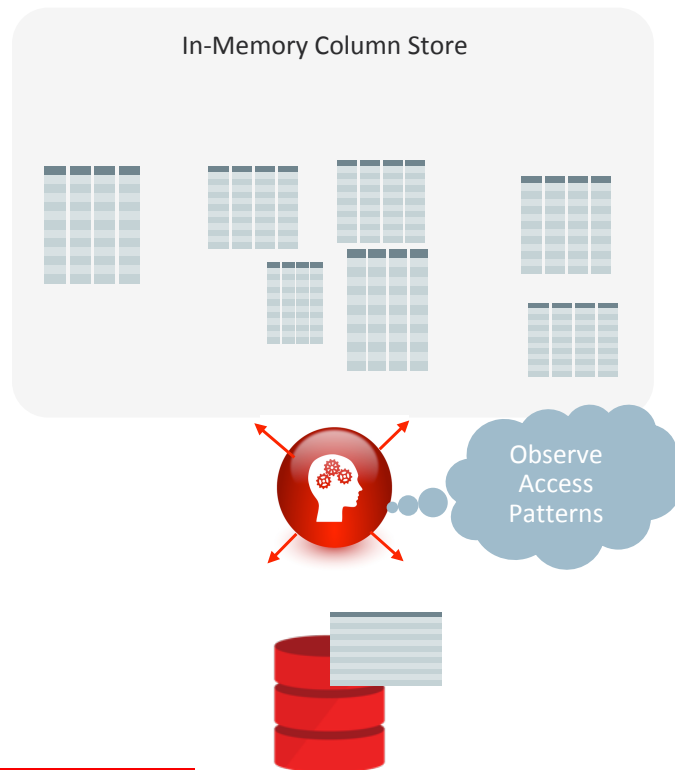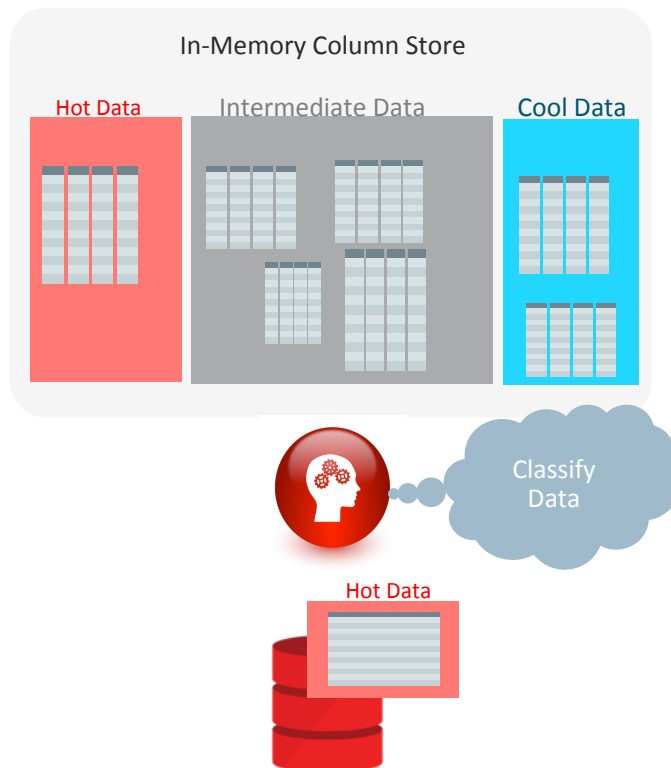|	
Real Time Data Analysis	DBIM
Hot (Frequently Accessed) Data	DBIM & Flash Cache
Cooling Data	HCC on Exadata Storage
Cold or Raw Data	ZFS or Big Data SQL

Introducing Automatic In-Memory

NEW IN 18ᶜ

- Eliminates trial and error regarding in-memory area contents
- Constant background action:
  - Classifies data as hot, intermediate or cold
  - Hotter in-memory tables automatically populated
  - Colder in-memory tables automatically removed
  - Intelligent algorithm takes into account space-benefit tradeoffs
- Controlled by new parameter **inmemory_automatic_level**
- Useful for autonomous cloud services since no user intervention required

In-Memory Column Store

Observe Access Patterns

ORACLE®

Introducing Automatic In-Memory

NEW IN 18c

In-Memory Column Store

Hot Data · Intermediate Data · Cool Data

Classify Data

Hot Data

- Eliminates trial and error regarding in-memory area contents
- Constant background action:
  – Classifies data as hot, intermediate or cold
  – Hotter in-memory tables automatically populated
  – Colder in-memory tables automatically removed
  – Intelligent algorithm takes into account space-benefit tradeoffs
- Controlled by new parameter **inmemory_automatic_level**
- Useful for autonomous cloud services since no user intervention required
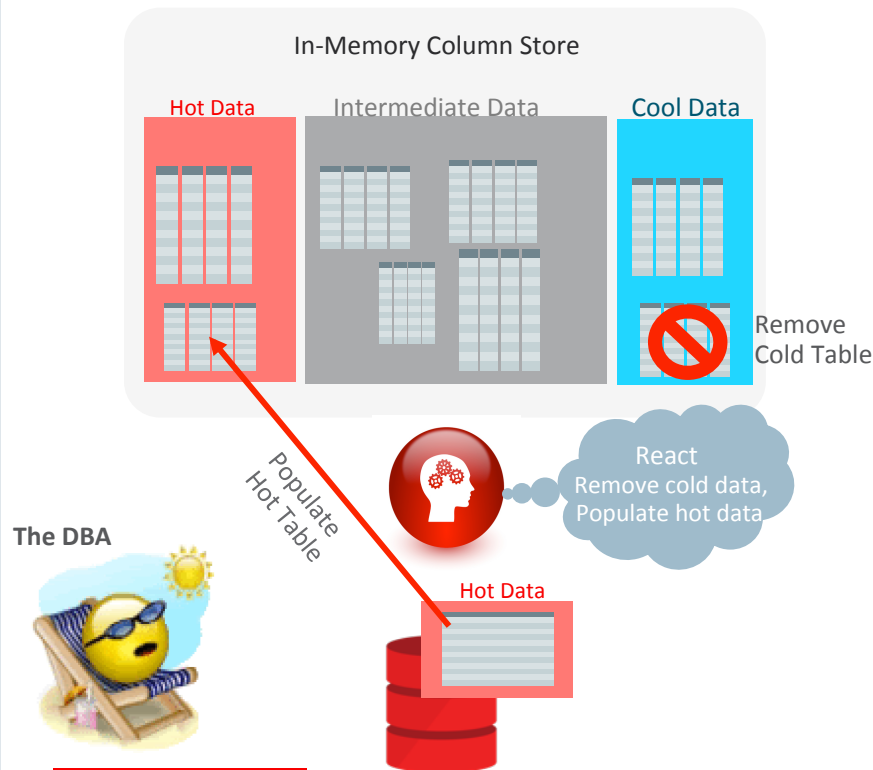
Introducing Automatic In-Memory

In-Memory Column Store

Hot Data Intermediate Data Cool Data

Remove Cold Table

Populate Hot Table

React
Remove cold data,
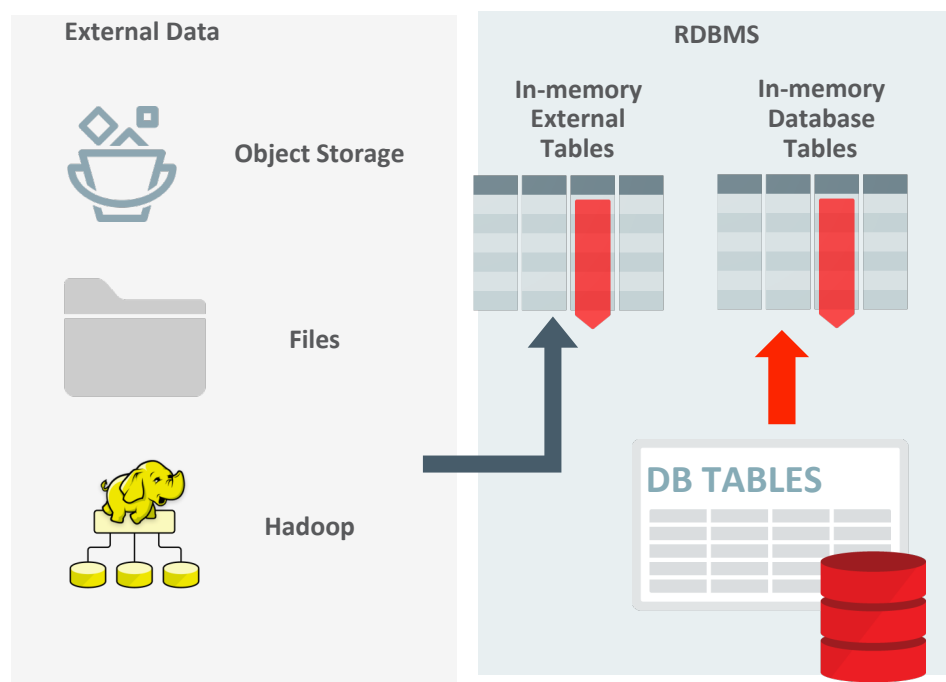Populate hot data

The DBA

Hot Data

- Eliminates trial and error regarding in-memory area contents
- Constant background action:
  - Classifies data as hot, intermediate or cold
  - Hotter in-memory tables automatically populated
  - Colder in-memory tables automatically removed
  - Intelligent algorithm takes into account space-benefit tradeoffs
- Controlled by new parameter **inmemory_automatic_level**
- Useful for autonomous cloud services since no user intervention required

ORACLE®

In-Memory For External Tables
Fast Analytics on External Data

External Data

Object Storage

Files

Hadoop

RDBMS

In-memory External Tables

In-memory Database Tables

DB TABLES

- External Tables allow transparent SQL on external data

- In-Memory External Tables: **100x faster** analytics on external data

- All In-Memory Optimizations apply
  - Vector processing, JSON expressions extend transparently to external data

- Simple to enable via CREATE / ALTER:

```
create table EXT1(…)
organization external(…)inmemory

alter table EXT2 inmemory
```
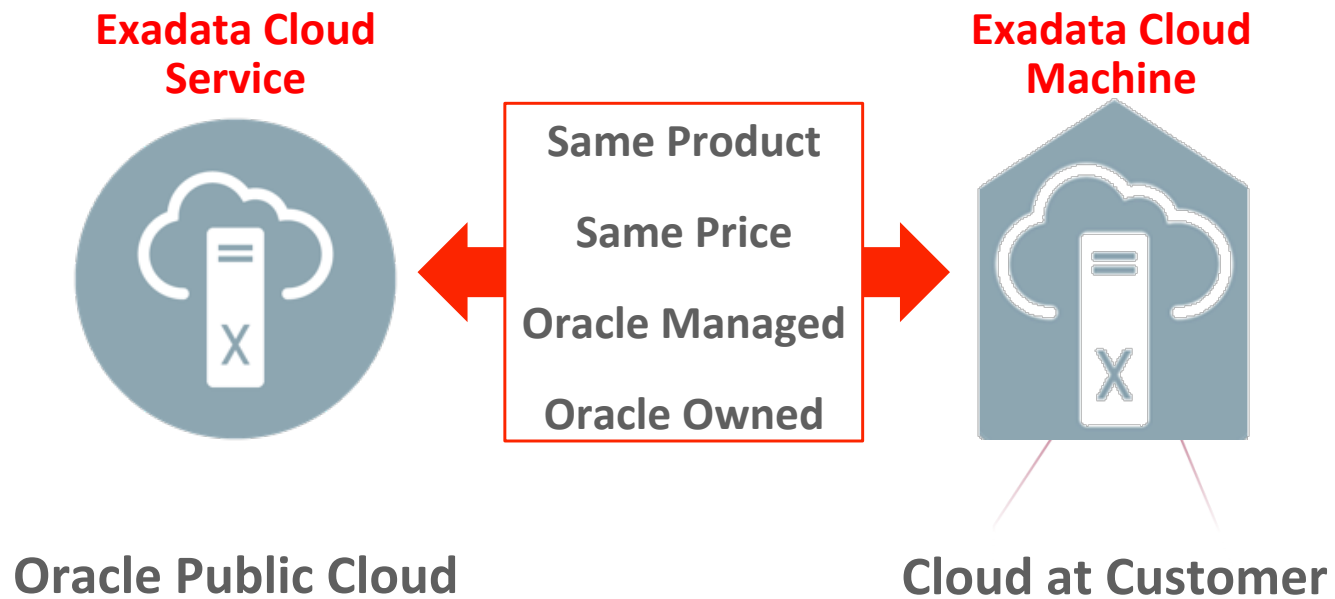
Note: HIVE/HDFS support targeted for 19c

NEW IN
18c

Preview: Non-volatile Memory Tier in Exadata Storage

Compute Server

Storage Server

RDMA

NVRAM

Hot

Warm

Cold
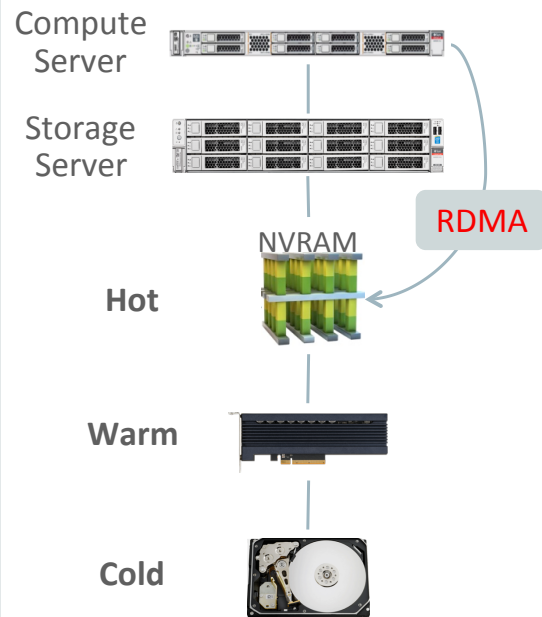
- **Exadata Storage Servers will add a non-volatile memory (NVRAM) cache in front of Flash memory**
  - Similar to current Flash cache in front of disk
  - RDMA direct access to NVRAM gives **20x lower latency** than Flash

- **NVRAM used as a cache effectively increases its capacity 10x**

- **Expensive NVRAM shared across servers for lower cost**

- **NVRAM mirrored across storage servers for fault-tolerance**

ORACLE®

Exadata Cloud – Your Way

Exadata Cloud Service

Same Product

Same Price

Oracle Managed

Oracle Owned

Exadata Cloud Machine

Oracle Public Cloud

Cloud at Customer

Exadata Customer Case Studies

ORACLE®

Industry Examples of Heavy Ingest Workloads

SK telecom

- Korea's number one mobile operator

- 65 billion transactions per day

- 18TB of data per day

- All data processing occurs on Oracle Database running on Exadata

U.S. Department of Homeland Security

- One of world's largest law enforcement orgs

- ~3 billion transactions per day

- ~32 billion queries per day

- Database is over 1PB

- Deployed on Oracle Database on Exadata

NYSE

- World's largest stock exchange

- ~1000 million database transactions per day
  - 180,000 messages/sec

- ~ 15 TB of data per day

- All data captured and processed in an Oracle Database on Exadata

Heavy Transactional Workloads with Oracle Exadata

GARMIN

- Garmin Connect Mobile
- 4 million active users
- 6 Billion miles of user activity a day
- All user data & geospatial data is store in an Oracle Database on Exadata

RWE
The energy to lead

- Leading electricity and gas providers in Europe
- Ingests and processes 2.4 Billion smart meter reads a day
- System runs on Oracle Database on Exadata

Canon

- Leading camera and printer manufacture
- Remote monitoring of over 1 million multifunction printers from 100 countries
- System runs on Oracle Database on Exadata

Exadata Database Machine

Performance, Availability and Security
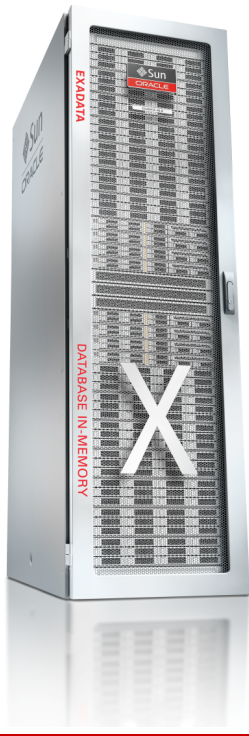
Best Platform for Oracle Databases on-premises and in the Cloud

Delivers:

- Memory-Level Performance

- Automatic Data Tiering

- 5 Nines Availability

Exadata Advantages Increase Every Year

Dramatically Better Platform for <u>All Database Workloads</u>

Smart Software

Smart Hardware

- 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

- Network Resource Management
- Multitenant Aware Resource Mgmt
- Prioritized File Recovery
- 3D V-NAND Flash
- IO Priorities
- Software-in-Silicon
- Data Mining Offload
- Tiered Disk/ Flash
- Offload Decrypt on Scans
- Database Aware Flash Cache
- PCIe NVMe Flash
- Storage Indexes
- Columnar Compression
- Smart Scan
- Unified InfiniBand
- InfiniBand Scale-Out
- DB Processors in Storage
- Scale-Out Storage
- Scale-Out Servers

Integrated Cloud
Applications & Platform Services

19

ORACLE®