Oracle Exadata: Achieving Memory-Level Performance with NVME 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 X7 delivers **350GB/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
• **Shared storage with memory-level bandwidth** is a paradigm change in the industry
  – Get near DRAM throughput, with the capacity of shared flash

---

Exadata DB Servers

Exadata Smart Storage

Flash Chips

PCIe NVMe

InfiniBand

Query Offload

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

In-Memory Columnar scans

Database Server

SGA

Up to 1.5 TB DRAM

IMC

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

In-Memory Columnar scans

In-Flash Columnar scans

Database Server

SGA

Up to 1.5 TB DRAM

Extends In-Memory Column Store into Flash

25.6 TB Flash x 3 = 76.8 TB (or more)
IMC (In-Memory Columnar) data

Storage Server

Hybrid Columnar Compressed Data
Introducing **Automatic In-Memory**

- 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**

- 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**

- 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
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
Preview: Non-volatile Memory Tier in Exadata Storage

- 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
Exadata Cloud – Your Way

Oracle Public Cloud

Exadata Cloud Service

Same Product
Same Price
Oracle Managed
Oracle Owned

Exadata Cloud Machine

Cloud at Customer
Exadata Customer Case Studies
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

- **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

- **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

<table>
<thead>
<tr>
<th>Company</th>
<th>Description</th>
</tr>
</thead>
</table>
| **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**   | • 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 All Database Workloads
Integrated Cloud
Applications & Platform Services