Oracle In-Memory Processing

Tirthankar Lahiri
Vice President
Data and In-Memory Technologies
Oracle Database

Class-Leading In-Memory Technology
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.
Database Platform of the **Future** – **Complete and Integrated**

**Services**
- OLTP, IoT
- Any Application
- Any Data

**Architecture**
- In-Memory Database
- Fast, Low Cost, Secure
- Scalable, Available

**Systems**
- Optimized Compute
- Smart Storage
- Fastest Networking

**Cloud**
- Autonomous, Elastic
- Wherever You Want

**Data Science, AI, ML**

**Spatial, Graph, Text, Media**

**Web, Micro Services**

**Real Time Analytics**

**Big Data**

**REST**

Copyright © 2018, Oracle and/or its affiliates. All rights reserved. |
In-Memory Database Technology Facts
The Time is NOW

• Next Generation Enterprises must be real-time
  – In-Memory is essential for real-time processing
  – In-Memory is essential in next-gen database platforms

• Current Memory sizes enable In-Memory today
  – Oracle X7-2 server: Up to 1.5TB of DRAM
  – Not using in-memory is like using Windows 98 in 2018
In-Memory Across Tiers in the Oracle Database Stack

**Application-Tier**

- **TimesTen In-Memory Database**
  - Latency Critical custom OLTP applications
  - Microsecond response time
  - Standalone or Cache for Oracle Database

**Database-Tier**

- **Database In-Memory**
  - Dual Format In-Memory Database
  - Billions of Rows/sec analytic data processing
  - 2-3x Faster Mixed Workloads

**Storage-Tier**

- **In-Memory on Exadata Storage**
  - In-memory column format on Exadata Flash Cache
  - 5-10x faster smart scan in storage
  - 15x increase in total columnar capacity
Oracle TimesTen In-Memory Database

Relational Database
- Pure in-memory
- ACID compliant
- Standard SQL
- Entire database in DRAM

Persistent and Recoverable
- Database and Transaction logs persisted on local disk or flash storage

Extremely Fast
- Microseconds response time
- Very high throughput

Highly Available
- Active-Standby and multi-master replication
- Very high performance parallel replication
Application-Tier Database Cache for Oracle Database

- Cache subset of Oracle Database tables in TimesTen for better response time
  - With full persistence to local storage
- Read-write caching
  - Transaction execution and persistence in TimesTen
- Read-only caching
  - Transactions executed in Oracle Database
- HA and fault tolerance in the application-tier
New Release: TimesTen Scaleout In-Memory Database
Proven TimesTen technology with Scale-Out for High Performance, and Multi-Copy for High Availability

- For High-Velocity **Extreme OLTP** applications
  - IOT, trading, mobile, click stream, billing, orders, fraud, etc.

- Performance-Oriented Design
  - Pure In-Memory, Full SQL, Full ACID Transactions
  - Fault-Tolerant Scale-Out
    - All nodes active for read/writes
    - Multiple data copies for HA
  - Sophisticated and Parallel SQL for reporting and batch

- Extremely easy to install and operate
  - **Less than 15 mins** to install, deploy, and run application
Scale-Out Shared Nothing In-Memory Database

Single-Image DB with High Availability and Elasticity

- Appears as a single DB to applications
  - **Not** a sharded database
- Adding and removing DB nodes (elements)
  - Data automatically redistributed
  - Workload automatically uses new nodes
- Built-in HA via fully-active node copies
  - Node copies automatically kept in sync
- Highly compatible with Oracle (subset)
  - Data types, SQL & PLSQL, Oracle Call Interface
YCSB Workload A (50% Read 50% Update)

**Workload:**
- YCSB version 0.15.0
- 1KB record (100-byte x 10 Fields)
- 100M records / Replica Set
- Uniform Distribution

**Configuration**
- Oracle Bare Metal Cloud
- 32 x DenseIO.52 hosts
- Oracle Linux x86 64-bit
- Intel Platinum 8167M @2GHz
- 25G Ethernet
- NVMe disk storage
YCSB Workload B (95% Read 5% Update)

Workload:
- YCSB version 0.15.0
- 1KB record
  (100-byte x 10 Fields)
- 100M records / Replica Set
- Uniform Distribution

Configuration
- Oracle Bare Metal Cloud
- 32 x DenseIO.52 hosts
- Oracle Linux x86 64-bit
- Intel Platinum 8167M @2GHz
- 25G Ethernet
- NVMe disk storage
YCSB Workload C (100% Read)

Workload:
- YCSB version 0.15.0
- 1KB record
  (100-byte x 10 Fields)
- 100M records / Replica Set
- Uniform Distribution

Configuration
- Oracle Bare Metal Cloud
- 32 x DenseIO.52 hosts
- Oracle Linux x86 64-bit
- Intel Platinum 8167M @2GHz
- 25G Ethernet
- NVMe disk storage
Linear and Massive Transaction Scalability

SQL Transactions/sec

**Workload**
- TPTBM telco workload
- Shipped with TimesTen for 10+ years
- 80% read / 20% write workload

**Configuration**
- Oracle Bare Metal Cloud
- 64 x BM.HighIO1.36
- Oracle Linux x86 64-bit
- E5-2699 v3 @ 2.30GHz
- 10G Ethernet
- NVMe Disk storage
Linear and Massive Query Scalability

1.2 Billion SQL Selects/sec

Workload
- TPTBM telco workload
- Shipped with TimesTen for 10+ years
- 100% read workload

Configuration
- Oracle Bare Metal Cloud
- 64 x BM.HighIO1.36
- Oracle Linux x86 64-bit
- E5-2699 v3 @ 2.30GHz
- 10G Ethernet
- NVMe Disk storage
In-Memory Across Tiers

**Application-Tier**

- TimesTen In-Memory Database
  - Latency Critical custom OLTP applications
  - Microsecond response time
  - Standalone or Cache for Oracle Database

**Database-Tier**

- Database In-Memory
  - Dual Format In-Memory Database
  - Billions of Rows/sec analytic data processing
  - 2-3x Faster Mixed Workloads

**Storage-Tier**

- In-Memory on Exadata Storage
  - In-memory column format on Exadata Flash Cache
  - 5-10x faster smart scan in storage
  - 15x increase in total columnar capacity
Row Format Databases vs. Column Format Databases

**Row**
- **Transactions** run faster on row format
  - Example: Insert or query a sales order
  - Fast processing for few rows, many columns

**Column**
- **Analytics** run faster on column format
  - Example: Report on sales totals by region
  - Fast accessing few columns, many rows

Until Now Must Choose One Format and Suffer Tradeoffs
What is a Controversy?

“A discussion marked especially by the expression of opposing views”

Merriam Webster
Decades Long Controversy in Database Systems

“A discussion marked especially by the expression of opposing views”

*Merriam Webster*
Oracle Database In-Memory: Dual Format Architecture

- BOTH row and column formats for same table
- Simultaneously active and consistent
- OLTP uses existing row format
- Analytics uses new In-Memory Column format
In-Memory Columnar Format

- Pure in-memory column format
- Cheap to maintain – no logging or IO
- Allows efficient OLTP
- No changes to disk format
  - Transparent to Applications
- Can be enabled for any subset of database
  - Partitions of tables
  - Tables
  - Tablespaces
  - Full database if needed
Vector Processing: Additional Advantage of Column Format

- Each CPU core scans only required columns
- SIMD vector instructions used to process multiple values in each instruction
  - E.g. Intel AVX instructions
- Billions of rows/sec scan rate per CPU core
  - Row format is millions/sec

Example: Find all sales in state of CA

> 100x Faster
Joining and Combining Data Also Dramatically Faster

Example: Find total sales in outlet stores

- Converts joins of data in multiple tables into fast filtered column scans
- Joins tables 10x faster
Generates Reports Instantly

**Example:** Report sales of footwear in outlet stores

- Dynamically creates in-memory report outline
- Then report outline filled-in during fast fact scan
- Reports run **10x** faster
  - Without predefined cubes
Scale-Out In-Memory Database to Any Size

• Scale-Out across servers to grow memory and CPUs

• In-Memory *queries parallelized* across servers to access local column data

• **Direct-to-wire** InfiniBand protocol speeds messaging on Engineered Systems
### In-Memory Across Tiers

#### Application-Tier

- TimesTen In-Memory Database
  - Latency Critical custom OLTP applications
  - **Microsecond** response time
  - Standalone or Cache for Oracle Database

#### Database-Tier

- **Database In-Memory**
  - Dual Format In-Memory Database
  - **Billions of Rows/sec** analytic data processing
  - **2-3x** Faster Mixed Workloads

#### Storage-Tier

- **In-Memory on Exadata Storage**
  - In-memory column format on Exadata Flash Cache
  - **5-10x** faster smart scan in storage
  - **15x** increase in total columnar capacity

---

<table>
<thead>
<tr>
<th>Application-Tier</th>
<th>Database-Tier</th>
<th>Storage-Tier</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>SALES</strong></td>
<td><strong>SALES</strong></td>
<td><strong>SALES</strong></td>
</tr>
<tr>
<td>Row Format</td>
<td>Column Format</td>
<td></td>
</tr>
</tbody>
</table>

---

[Image: Oracle logo and sales database example]
Exadata: The Ultimate Database Platform

• **Ideal Database Hardware** - Scale-out, database optimized compute, networking, and storage

• **Smart Database Software** – Specialized algorithms for **Analytics, OLTP, and Consolidation**

• **Full-Stack Automation** – Automation and optimization of configuration, updates, performance, resource management
Exadata: Transparent Storage Tiering
In-Memory Speed + Flash Bandwidth + Disk Capacity

- Size not limited by memory
- Data transparently accessed across tiers
- Each tier has specialized algorithms & compression
- Simultaneously Achieve:
  - **Speed** of DRAM
  - **I/Os** of Flash
  - **Cost** of Disk
Exadata Achieves Memory Performance with Shared Flash

- Exadata X7 delivers **400GB/sec flash bandwidth**
  - Approaches 800GB/sec aggregate DRAM bandwidth of DB servers
- Must move compute to data to achieve full flash potential
  - Push query processing into storage tier
  - Requires owning full stack, can’t be solved in storage or DB alone
- Flash storage arrays scale flash **capacity** but not **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
In-Memory Columnar Format in Columnar Flash Cache

• In-Memory format in Smart Columnar Flash
  – Enables SAME in-memory optimizations on data in Exadata flash as available on Exadata DB compute nodes DRAM

• Extends in-memory from DB compute nodes to Storage nodes
  – **15x** Columnar Capacity (**100s** of TB possible on full rack X7)
  – New in-memory format – offloaded queries **10x faster**

• Does not require any user intervention
  – Powers the new Autonomous Data Warehouse
Summary: In-Memory is a Cross-Tier Technology

Application-Tier

- **TimesTen In-Memory Database**
  - Latency Critical custom OLTP applications
  - **Microsecond** response time
  - Standalone or Cache for Oracle Database

Database-Tier

- **Database In-Memory**
  - Dual Format In-Memory Database
  - **Billions of Rows/sec** analytic data processing
  - **2-3x** Faster Mixed Workloads

Storage-Tier

- **In-Memory on Exadata Storage**
  - In-memory column format on Exadata Flash Cache
  - **5-10x** faster smart scan in storage
  - **15x** increase in total columnar capacity
Oracle In-Memory Databases Scored Highest by Forrester on both Current Offering and Strategy


The Forrester Wave™ is copyrighted by Forrester Research, Inc. Forrester and Forrester Wave™ are trademarks of Forrester Research, Inc. The Forrester Wave™ is a graphical representation of Forrester’s call on a market and is plotted using a detailed spreadsheet with exposed scores, weightings, and comments. Forrester does not endorse any vendor, product, or service depicted in the Forrester Wave. Information is based on best available resources. Opinions reflect judgment at the time and are subject to change.
In-Memory Enables the Real Time Enterprise

- **Data Driven** – Rapidly make decisions based on real-time data
- **Agile** – Respond quickly to change
- **Efficient** – Continually improve processes and profitability

Real-Time Enterprise
Integrated Cloud
Applications & Platform Services