Oracle SPARC Systems

Michael Coleman
Chief Technologist, Systems, Asia Pacific Division

SKA Colloquium, Auckland New Zealand
February 2017
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.
Agenda

1. SPARC Roadmap and Overview
2. Software in Silicon
3. Performance
4. Analytics
5. SuperCluster
Oracle SPARC/Solaris Platform Roadmap

- **Solaris 11.1**
  - M5: +6x Throughput, +1.5x Thread Strength
  - T5: +2.5x Throughput, +1.2x Thread Strength

- **Solaris 11.2**
  - M6: +2x Throughput, >1x Thread Strength

- **Solaris 11.3**
  - M7: +2x Throughput, +1.5x Thread Strength
  - S7: Fully Integrated Software in Silicon V1

- **SPARC IaaS**
  - Oracle Public Cloud
  - Dedicated Compute

- **@Customer**
  - Dedicated Metered & Non-Metered

- **Solaris 11.next**
  - Cloud Deployment & Integration Enhancements

- **SPARC next**
  - 1.4x Throughput
  - 1.5x Thread Strength
  - Software in Silicon V2

- **SPARC next+**
  - Increased Cache
  - Increased Bandwidth
  - Software in Silicon V3

As of January 13, 2017

Copyright © 2017, Oracle and/or its affiliates. All rights reserved.
Consistent Execution
SPARC @ Oracle
7 Processors in 6 Years

<table>
<thead>
<tr>
<th>Year</th>
<th>SPARC</th>
<th>2nd/3rd/4th Gen</th>
<th>Cores</th>
<th>L2/L3 Cache</th>
<th>Frequency</th>
</tr>
</thead>
<tbody>
<tr>
<td>2010</td>
<td>T3</td>
<td>16 x 2nd Gen</td>
<td>16</td>
<td>6MB L2, 8MB L3</td>
<td>1.65 GHz</td>
</tr>
<tr>
<td>2011</td>
<td>T4</td>
<td>8 x 3rd Gen</td>
<td>16</td>
<td>4MB L3</td>
<td>3.0 GHz</td>
</tr>
<tr>
<td>2013</td>
<td>T5</td>
<td>16 x 3rd Gen</td>
<td>16</td>
<td>8MB L3</td>
<td>3.6 GHz</td>
</tr>
<tr>
<td>2013</td>
<td>M5</td>
<td>6 x 3rd Gen</td>
<td>12</td>
<td>48MB L3</td>
<td>3.6 GHz</td>
</tr>
<tr>
<td>2013</td>
<td>M6</td>
<td>12 x 3rd Gen</td>
<td>16</td>
<td>48MB L3</td>
<td>4.13 GHz</td>
</tr>
<tr>
<td>2015</td>
<td>M7</td>
<td>32 x 4th Gen</td>
<td>32</td>
<td>64MB L3</td>
<td>4.13 GHz</td>
</tr>
<tr>
<td>2016</td>
<td>S7</td>
<td>8 x 4th Gen</td>
<td>8</td>
<td>16MB L3</td>
<td>4.27 GHz</td>
</tr>
</tbody>
</table>

Including Software in Silicon:
- Silicon Secured Memory
- Encryption Acceleration
- Data Analytics Accelerators
- More...

Copyright © 2015, Oracle and/or its affiliates. All rights reserved.
Software in Silicon
Software in Silicon: Next Step in Hardware Evolution


Mid-1990’s
Larger Memory Support, Greater Accuracy

Mid-2000’s
Multi-core, Multi-threaded Computing

Open Systems
Virtualization

Today
Software in Silicon: Software Functions on Chip
SPARC M7: Software In Silicon Features

Security in Silicon:
Silicon Secured Memory
Cryptography Acceleration

SQL in Silicon:
Database In Memory Accelerator Engines

Capacity in Silicon:
Decompression Engines
Silicon Secured Memory: Always-On Intrusion Protection

Breakthrough security and reliability for applications

- Unique hardware-based memory protection
- Stops malicious programs from accessing other application memory. Ex: HeartBleed, Venom
- Can be always on: hardware approach has negligible performance impact
- Easily activated for existing applications
- Extremely efficient for software development
Silicon Secured Memory: Always-On Intrusion Protection

Breakthrough security and reliability in hardware

• Silicon Secured Memory implements fine grained memory protection in hardware
  – Hidden “color” bits added to pointers (key), and content (lock)

• Pointer color (key) must match content color or program is aborted
  – Set on memory allocation, changed on memory free

• Helps prevent access off end of structure, stale pointer access, malicious attacks, etc. plus improves developer productivity
SPARC M7: Designed for Security

The most complete set of encryption standards

• 15 crypto algorithms
• 25 user level crypto instructions
• 32 Crypto Accelerators per Processor
• To accelerate:
  – Asymmetric (Public Key Encryption)
  – Symmetric Key (Bulk Encryption)
  – Message Digest (Hash Functions)
Oracle Solaris Stops Malware Before It Gets In

Immutable Systems and Virtual Machines
– Stop attackers from establishing a foothold
– Prevent administrator mistakes
– Update, even though it’s un-writable by users and applications

Tamper-Evident Software
– From firmware all the way to applications
– Install only known, trusted software
– If not signed, won’t install
– Verified Boot
SQL In Silicon: In-Memory Query Acceleration

- Dedicated SQL accelerators built on chip
  - Independently process streams of compressed data placed in system memory
  - Like adding 32 additional specialized cores to chip
  - Up to 170 Billion rows per second!

- Frees processor cores to run other applications, such as OLTP

- Decompresses data simultaneously to processing SQL functions
  - Like adding 64 additional specialized cores
M7 Query Accelerator Engine

• 32 In-silicon offload engines
• Cores/Threads operate synchronous or asynchronous to offload engines
• User level synchronization through shared memory
• High performance at low power
• 3x more memory bandwidth than x86
World’s Fastest Microprocessor
Balanced Design
SPARC M7: Setting 20 World Records in Performance

#1 SPECint_rate2006: 1,200 peak
#1 SPECfp_rate2006: 832 peak
#1 SPECjEnterprise2010: 25,093.06 EjOPs
#1 SAP-SD 2 processor: 30,800 users

And more...
Real-Time Enterprise: Simultaneous OLTP & In-memory

SPARC M7: Faster analytics, faster OLTP, and better response time

Analytics: SPARC M7 8.2x faster per chip
OLTP: SPARC M7 2.9x faster per chip

One SPARC M7

OLTP Performance
Throughput: 338k Transactions per sec
Response Time: 11ms

Analytics
Queries per min: 267

OLTP Performance
Throughput: 236k Transactions per sec
Response Time: 12ms

Two x86 E5 v3

Analytics
Queries per min: 65
Benchmark Disclosure Statement

- **Benchmark 2015, Oracle &/or its affiliates. All rights reserved. Oracle & Java are registered trademarks of Oracle &/or its affiliates. Other names may be trademarks of their respective owners.**

- **SPEC and the benchmark name SPECjEnterprise are registered trademarks of the Standard Performance Evaluation Corporation. Results from www.spec.org as of 10/25/2015.**
  - SPARC T7-1, 25,818.85 SPECjEnterprise2010 EJOPS (unsecure); SPARC T7-1, 25,093.06 SPECjEnterprise2010 EJOPS (secure); Oracle Server X5-2, 21,504.30 SPECjEnterprise2010 EJOPS (unsecure); IBM Power S824, 22,543.34 SPECjEnterprise2010 EJOPS (unsecure); IBM x3650 M5, 19,282.14 SPECjEnterprise2010 EJOPS (unsecure).

- **SPEC and the benchmark name SPECvirt_sc are registered trademarks of the Standard Performance Evaluation Corporation. Results from www.spec.org as of 10/25/2015.**
  - SPECvirt_sc2013 3026@168 VMs; HP DL580 Gen9, SPECvirt_sc2013 3020@168 VMs; Lenovo x850 X6; SPECvirt_sc2013 2655@147 VMs; Huawei FusionServer RH2288H V3, SPECvirt_sc2013 1616@95 VMs; HP ProLiant DL360 Gen9, SPECvirt_sc2013 1614@95 VMs; IBM Power S824, SPECvirt_sc2013 1371@79 VMs.

- **SPEC and the benchmark names SPECfp and SPECint are registered trademarks of the Standard Performance Evaluation Corporation. Results as of October 25, 2015 from www.spec.org and this report.**

- **SPEC and the benchmark name SPECrpoj are registered trademarks of Standard Performance Evaluation Corporation (SPEC). Results from http://www.spec.org as of 11/13/2015 and IBM announcement.**

- **SPEC and the benchmark name SPEC OMP are registered trademarks of the Standard Performance Evaluation Corporation. Results as of October 25, 2015 from www.spec.org and this report.**

- **Two-tier SAP Sales and Distribution (SD) standard application benchmarks, SAP Enhancement Package 5 for SAP ERP 6.0 as of 10/23/15: SPARC T7-2 (2 processors, 64 cores, 512 threads): 30,800 SAP SD users, 2 x 4.13 GHz SPARC M7, 1 TB memory, Oracle Database 12c, Oracle Solaris 11, Cert# 2015050. IBM Power System S824 (4 processors, 24 cores, 192 threads): 21,212 SAP SD users, 4 x 3.52 GHz POWER8, 512 GB memory, DB2 10.5, AIX 7, Cert#201401. Dell PowerEdge R730 (2 processors, 36 cores, 72 threads) 16,500 SAP SD users, 2 x 2.3 GHz Intel Xeon Processor E5-2699 v3 256 GB memory, SAP ASE 16, RHEL 7, Cert#2014033. HP ProLiant DL380 Gen9 (2 processors, 36 cores, 72 threads) 16,101 SAP SD users, 2 x 2.3 GHz Intel Xeon Processor E5-2699 v3 256 GB memory, SAP ASE 16, RHEL 6.5, Cert#2014032. SAP, R/3, reg TM of SAP AG in Germany and other countries. More info www.sap.com/benchmark.

- **Additional Info: http://blogs.oracle.com/bestperf**
Analytics
Legacy Approach Cannot Sustain Modern Analytics

Complexity and Latency

Latency and Delayed Insight

Data Warehouse/OLAP (Column Format)

Copy DB to Run Cost Management

BI Server (Analytics & Reports)

ETL
OLTP and Data Warehousing Consolidation

With SPARC you can efficiently run analytics and OLTP simultaneously
Solution: Extremely Efficient Analytics in Silicon

Real Time Analytics on up to the Moment Data

- BOTH row and column formats for same table
- Analytics & reporting use new in-memory column format
- Specialized analytics accelerators on-chip, with maximum bandwidth
SQL in Silicon: Data Analytics Accelerator (DAX)

- Stream Processor for data parallel operations
- DSP-like pipe for efficient filtering operations, typical of first phase of any query
- Cache sparing design with more complex processing in general purpose cores
# The Power of DAX for Analytics and Machine Learning

## Community Driven Innovation

<table>
<thead>
<tr>
<th>Application</th>
<th>DAX Advantage</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Oracle Database</strong></td>
<td>Up to 10x faster</td>
</tr>
<tr>
<td><strong>Apache Spark SQL</strong></td>
<td>Up to 21x faster</td>
</tr>
<tr>
<td><strong>Java Streams</strong></td>
<td>3x to 21x faster</td>
</tr>
<tr>
<td><strong>Active Pivot</strong></td>
<td>6x to 8x faster</td>
</tr>
<tr>
<td><strong>Machine Learning</strong></td>
<td>4x to 12x faster</td>
</tr>
</tbody>
</table>

Open access: [http://SWiSdev.oracle.com/DAX](http://SWiSdev.oracle.com/DAX)
Accelerating Apache Spark In-Memory Columnar
Open DAX with Apache Spark SQL Over 20x Faster Than Latest 2-chip x86

• Apache Spark runs on a Java Virtual Machine
• SPARC DAX offloads critical functions
  – SQL: filtering, dictionary encoding, 1 & 2 predicate, join processing, additional compression, etc.
  – 1-chip SPARC M7 DAX: 12.9 Billion rows/sec 2-predicate scan
  – 20X faster than 2 of the latest 20-core x86
SuperCluster
SuperCluster M7: Hardware Architecture

- **ZS3 Mixed-use Storage**
  - 160 TB (raw) storage for Virtual Machine and system data

- **QDR InfiniBand Unified Ultra-fast Network**
  - 40Gb/s QDR InfiniBand IO backplane

- **M7 Servers for Databases & Applications**
  - 1 or 2 M7 Chassis per system (Elastic Configurations)
  - 2 Physical Domains per M7 chassis, 1 - 4 processors ea.
  - Up to 8TB RAM per rack

- **Exadata Storage Servers for Oracle Database**
  - From 3 to 11 per configuration (Flex. Config.)
  - High Capacity (96TB raw disk and 12.8TB of raw flash ea.)
  - Extreme Flash (25.6TB raw flash ea.)
**Smart Analytics: In-Memory Formats in Columnar Flash Cache**

- In-Memory formats used in Smart Columnar Flash Cache
- Enables vector processing on storage server during smart scans
  - Multiple column values evaluated in single instruction
- Faster decompression speed than Hybrid Columnar Compression
- Enables dictionary lookup and avoids processing unnecessary rows
- Smart Scan results sent back to database in In-Memory Columnar format
  - Reduces Database node CPU utilization
- In-memory performance seamlessly extended from DB node DRAM memory to 10x larger capacity flash in storage
  - Even bigger differentiation against all-flash arrays and other in-memory databases
- Supports Database 12.1.0.2 and Database 12.2.0.1
Exadata Storage Server In-Memory Fault Tolerance

- Similar to storage mirroring
- Duplicate in-memory columns on another node
  - Enabled per table/partition
  - Application transparent
- Downtime eliminated by using duplicate after failure