Hot Interconnects 2015

Commercial Computing Trends and its Impact on Interconnect

Rick Hetherington
Vice President
SPARC Architecture and Performance

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.

Keynote Agenda

Who am I and why am I here

Server market trend : impact on interconnect

On-premise to Cloud Computing : impact on interconnect

Oracle Engineered Systems : dependence on interconnect

Sonoma Launch and how it summarizes this talk

My background:

More than 35 years in the Industry

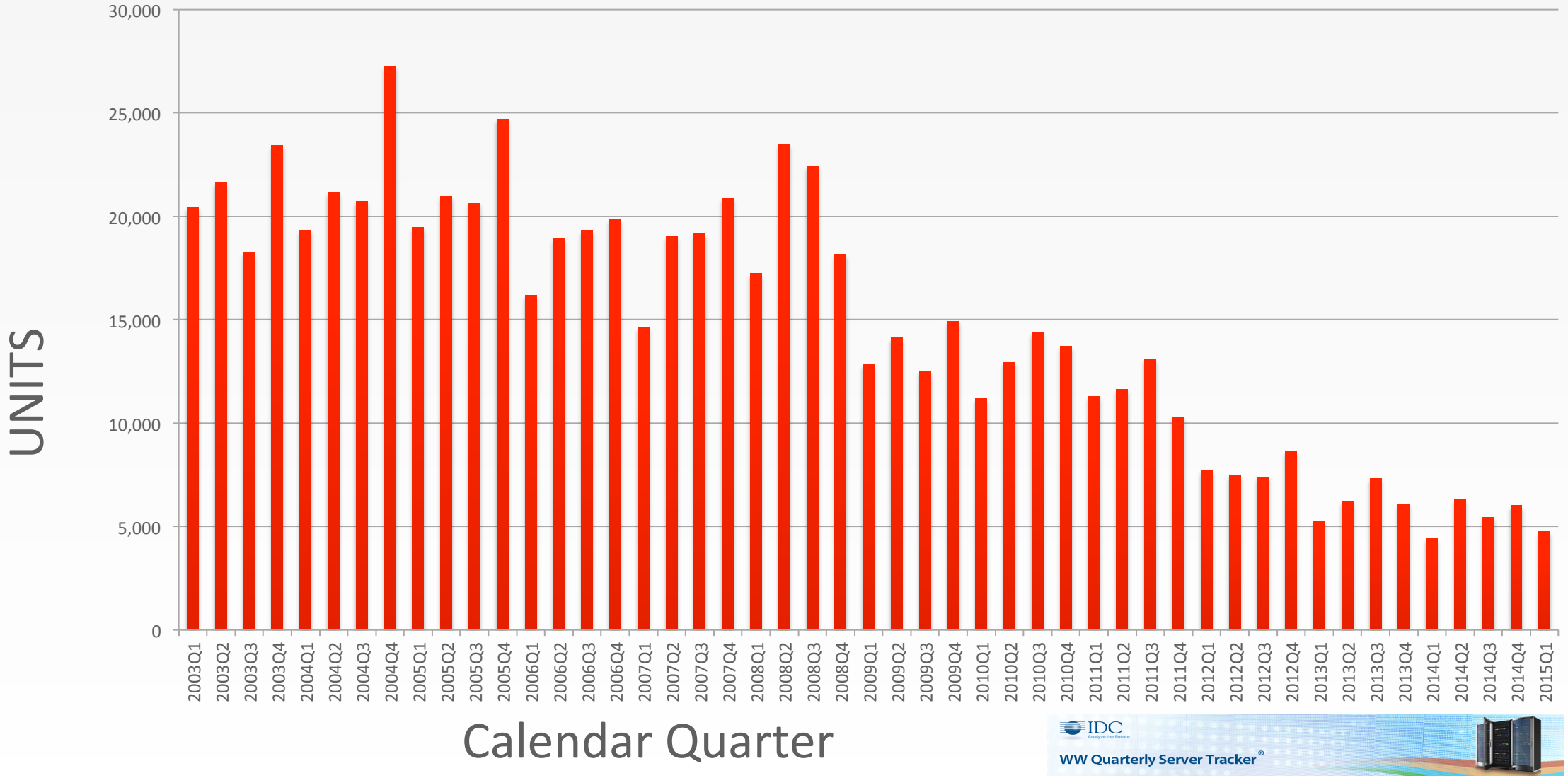
Sun/Oracle since 1996 as SPARC Architect

‘81-’96 with Digital Doing Vaxes and Alphas

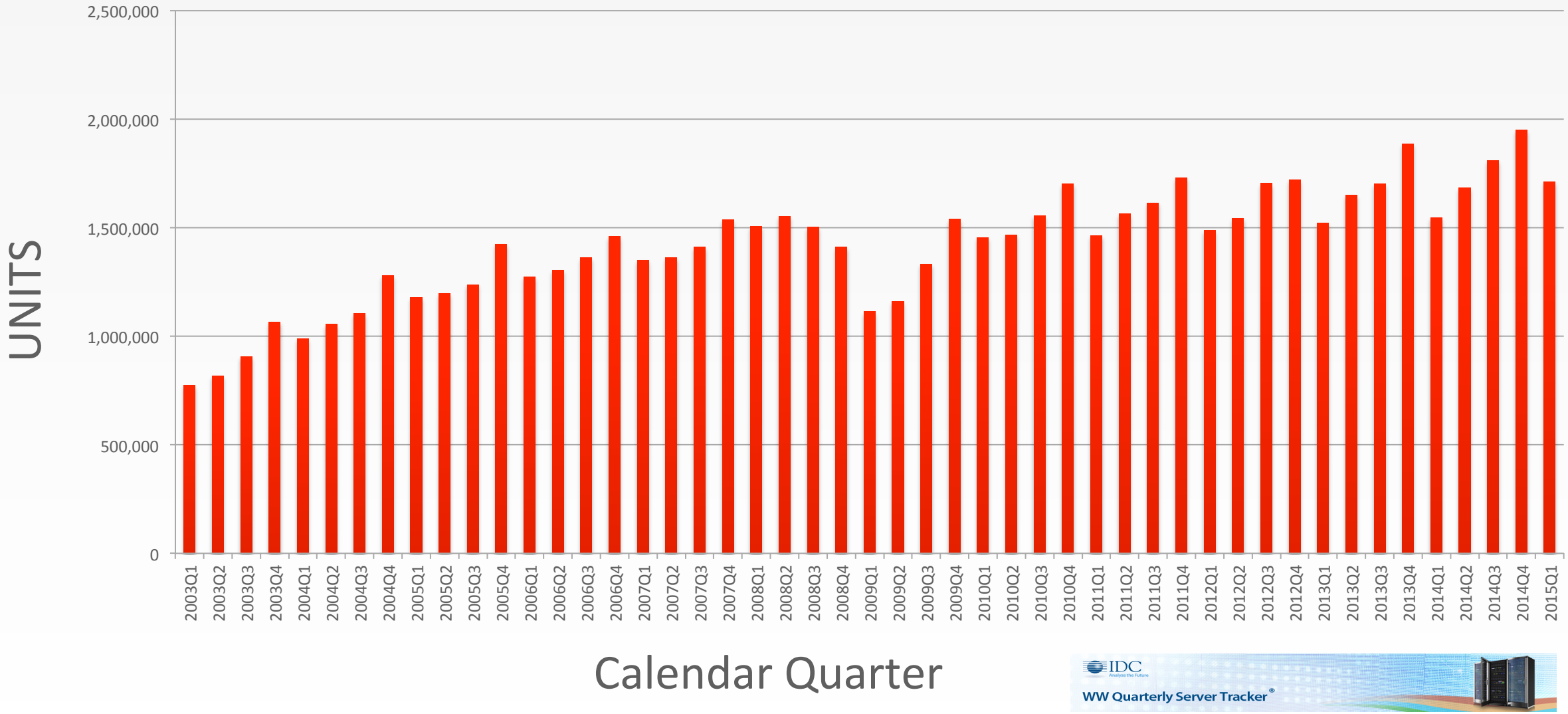
Alpha EV6 System Architect – first DDR SRAM + System Interface

Niagara Architect brought SERDES technology into SUN

8-64 Socket Server Market Trends



2 Socket Server Market Trends



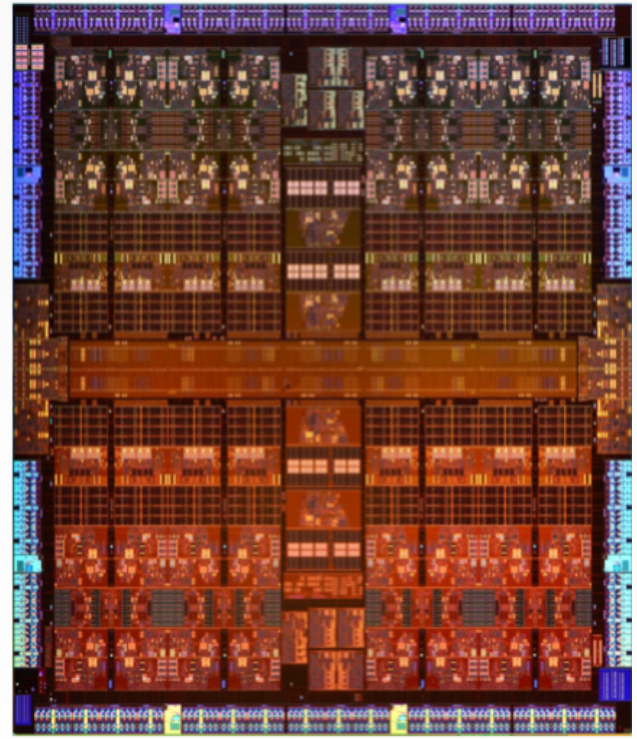
What is contributing to this trend?

Capability, Cost, Reliability, Power, Footprint, Applications, Cloud

M9000 - 64

=

M7 Processor



What is the impact on interconnect?

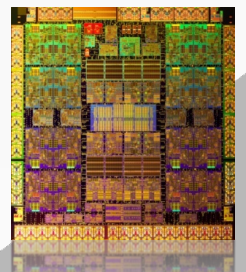
- Higher throughput (core count) concentrated in fewer processors
  - Large Dies but still a finite and limited pin count
- Critical interconnect is on die - higher levels of integration
- Memory Bandwidth is a limiter
  - 5 – 8GBs (delivered) per threaded-core on commercial workloads
  - Latency vs Bandwidth vs Capacity battle is never ending
- Extremely Fast Coherence Links for linear scaling
- IO
  - Fewer Adapters at much much higher frequency interconnect
  - Storage moving closer to the processor with NVMe
  - Storage moving next to the processor with persistent memory devices

SPARC @ Oracle

7 Processors in 6 Years

Including Software in Silicon

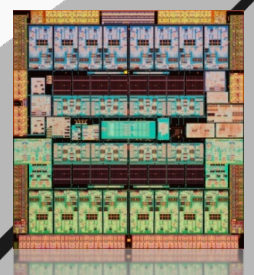
- App Data Integrity
- DB Query Acceleration
- Inline Decompression
- More…



2010

SPARC T3

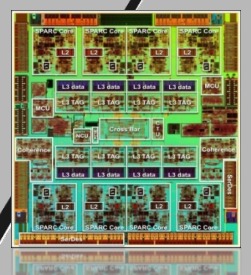
FBDIMM – DDR2-800
10GE
PCI-E 2.0 5.0Gbs
Coherence Links
9.8GbS



2011

SPARC T4

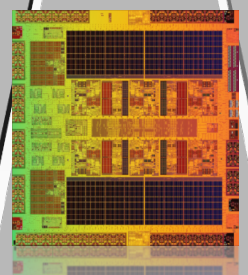
RDIMM DDR3 – 1066
10GE
PCI-E Rev 2.0 5.0Gbs
Coherence Links
9.8Gbs



2013

SPARC T5

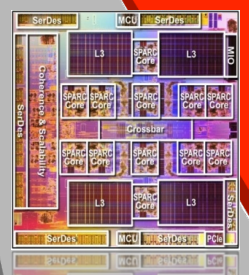
RDIMM-DDR3-1066
Host Links 12.8 Gbs
PCI-E Rev 3.0 8.0Gbs
Coherence Links
12.8 Gbs



2013

SPARC M5

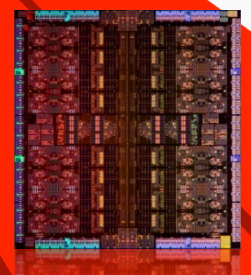
RDIMM-DDR3-1066
Host Links 12.8 Gbs
PCI-E Rev 3.0 8.0Gbs
Coherence Links
12.8 Gbs



2013

SPARC M6

RDIMM-DDR3-1066
Host Links 12.8 Gbs
PCI-E Rev 3.0 8.0Gbs
Coherence Links
12.8 Gbs

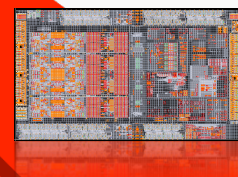


2015

SPARC M7

RDIMM-DDR4-2133
Host Links 16 Gbs
I/O Links 16 Gbs
Coherence Links
16 Gbs

Coming Soon



2016

Sonoma

RDIMM-DDR4-2400
PCI-E Rev 3.0 8 Gbs
Coherence Links
16 Gbs
IB at FDR

Coming in 2016

Oracle Cloud

- Data as a Service
- Software as a Service
- Platform as a Service
- Infrastructure as a Service

Interconnect in the Oracle Public Cloud

120K+

VMs in 19 Data Centers

#2

SaaS Provider in the world

70M+

Users on the Oracle Cloud Every Day

31B+

Transactions on the Oracle Cloud Every Day

Today:
- Server Edge: 10G
- Network Core: 40G

Trending:
- Server Edge: 25G-40Gbe
- Network Core: 100Gbe

Cloud trends pushing higher interconnect speeds?

- More capable processors with much high memory capacity
  - SPARC M7 - 32 cores and 256 threads
  - SPARC M7 – 52 bit Physical Address - 16 DDR4 DIMM Slots
- Higher Concentration of Virtual Machines (VM)/socket
  - Today less than 10 VMs/Socket
  - SPARC M7 can easily support 30 – 100 VMs
- Higher I/O Requirements
  - Smaller Servers with High Concentration of VMs requires much higher pin BW
  - IaaS driven by cost,fewer ports and cables to higher speed switches and routers

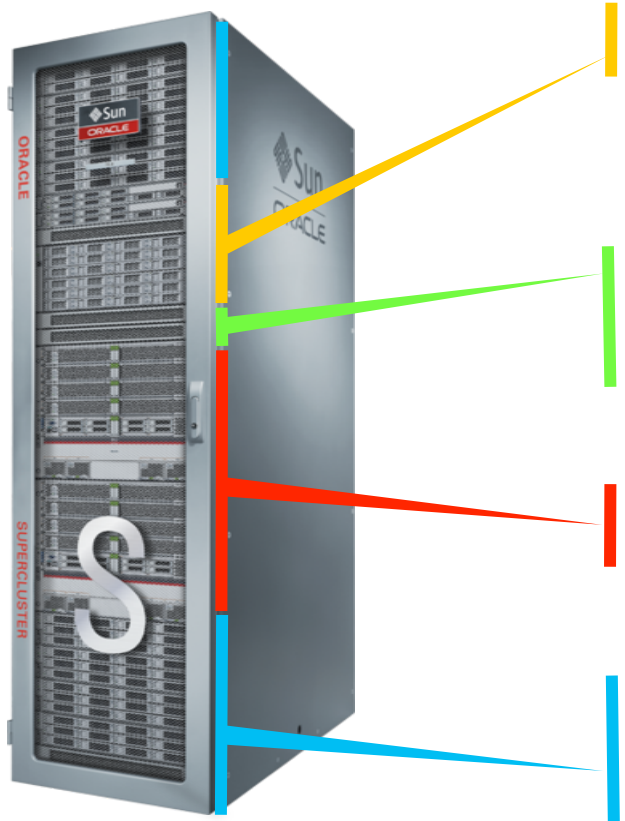
Oracle SuperCluster T5-8 and M6

SPARC Based Engineered Systems

ORACLE

Oracle SuperCluster T5-8 and M6-32 Architecture

Complete | Optimized | Standardized



Fully Redundant
Active Components

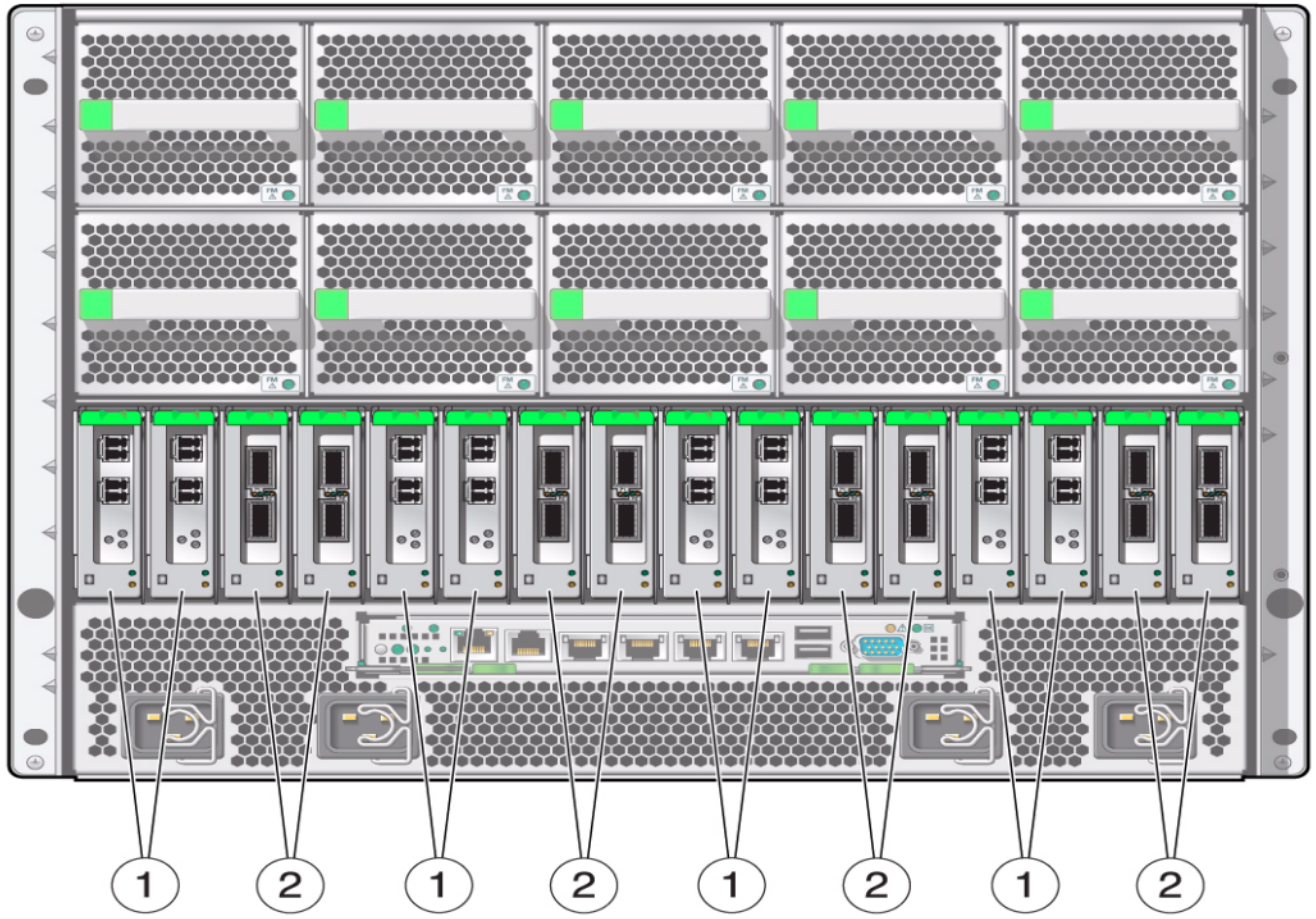
- **Integrated Enterprise NAS Storage**
  - System storage (system images, logs, test/dev databases, backup)
- **Unified Ultra-Fast Network**
  - InfiniBand internal I/O backplane
  - Ethernet data center connectivity
- **Database & Application Servers**
  - **T5-8:** 16 CPU (16 cores), 4TB RAM
  - **M6-32:** 32 CPU (12 cores), 32TB RAM
- **Exadata Storage Servers**
  - Optimized for Oracle Database
  - Intelligent scale-out storage grid

Fully Redundant
Active Components

ORACLE

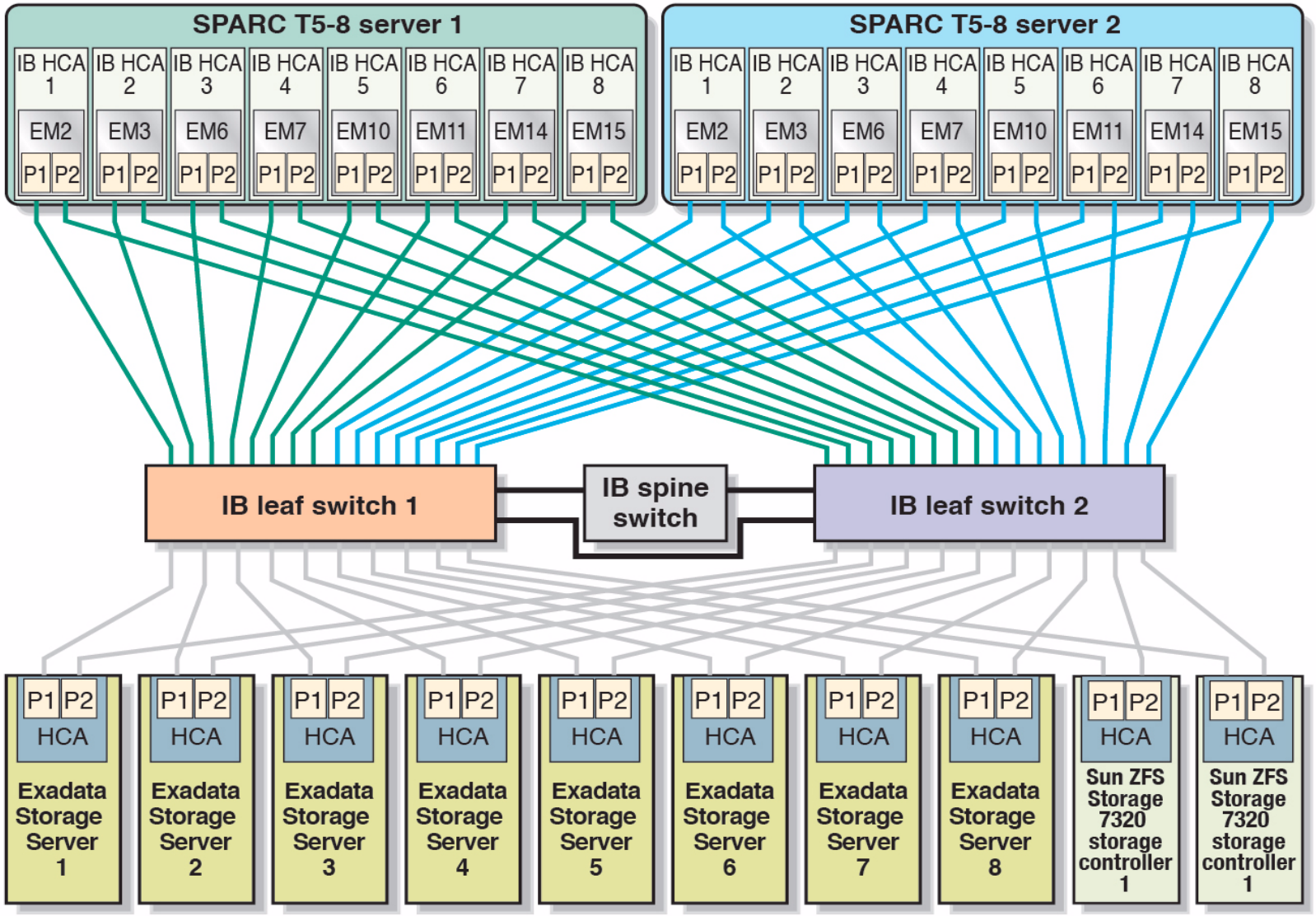
Interconnect of the SPARC T5-8 Server

2 T5-8 Systems in a Full Rack Interconnected with Infiniband



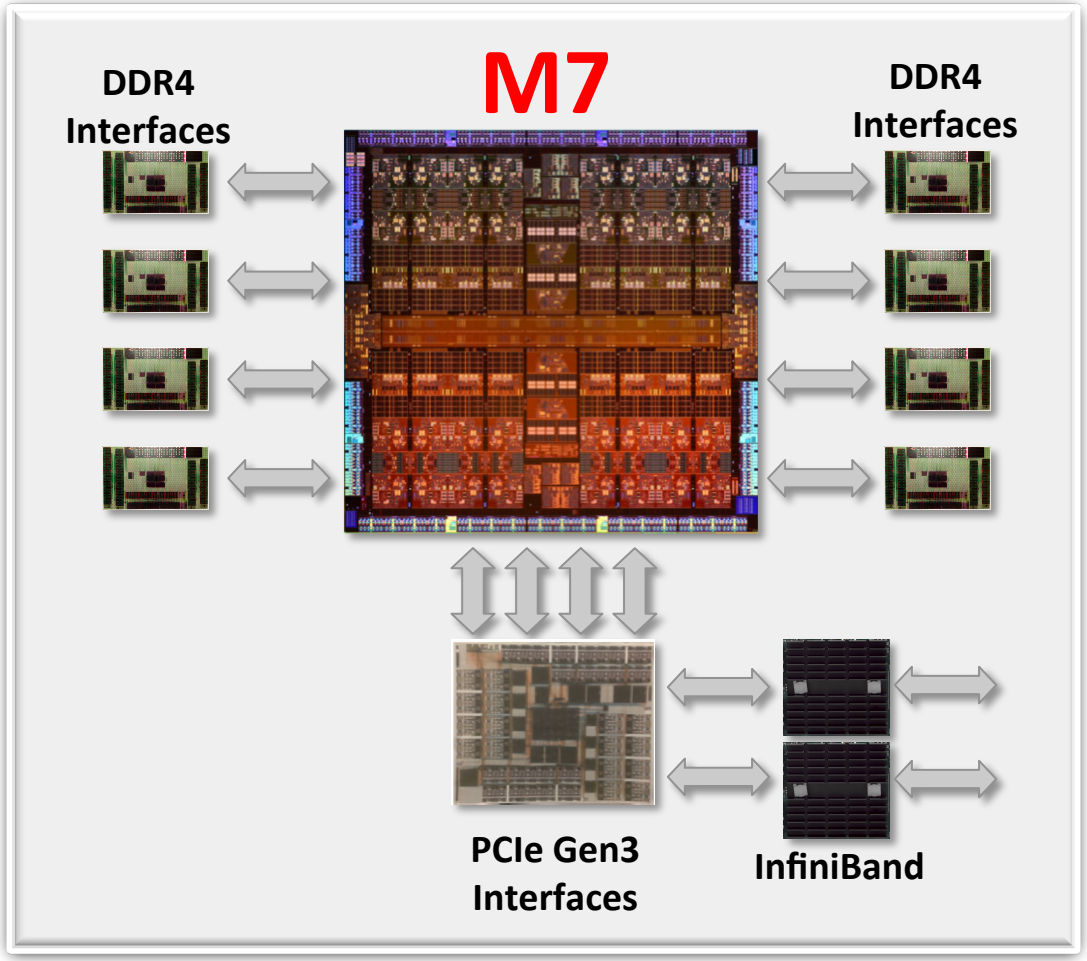
- (1) Dual-port 10 GbE network interface cards, for connection to the 10 GbE client access network
- (2) Dual-port InfiniBand host channel adapters, for connection to the InfiniBand network

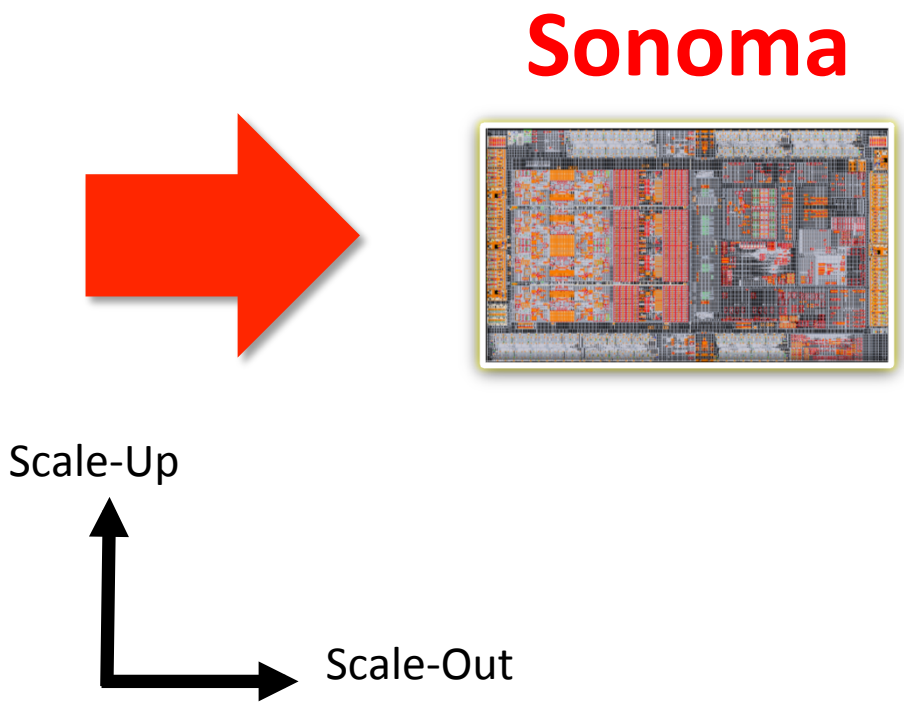
InfiniBand Network on Full Rack



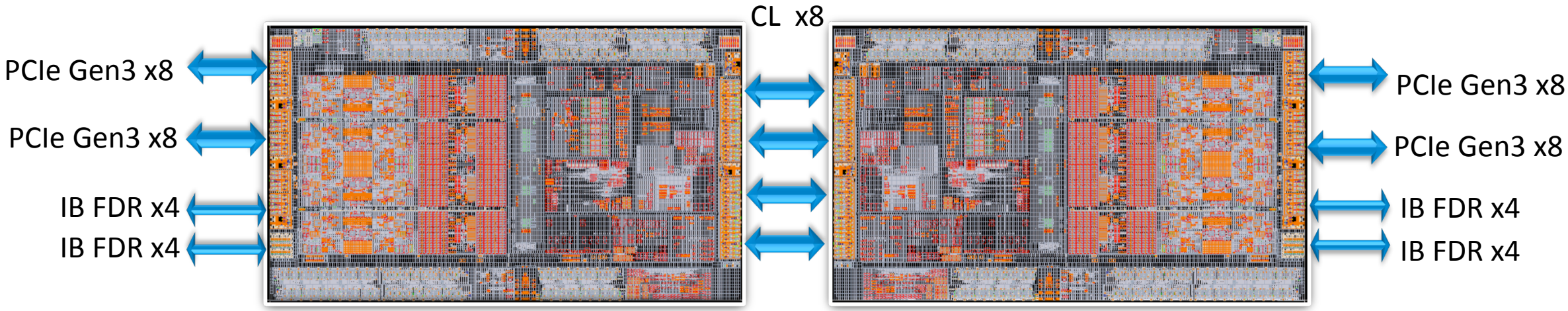
Sonoma

Fully Integrated to Lower Latency, Power, and Cost for Scale-Out





Connectivity Optimized for Scale-Out



- 2 InfiniBand links @ FDR (56Gbps)
  - Low latency scale-out networking interconnect for DB and clusters
  - 28 GB/s Bidirectional Bandwidth
- 2 PCIe links @ Gen3 (64Gbps)
  - 32 GB/s Bidirectional Bandwidth
- 4 Scale-Up Coherence links @ 16Gbps (128Gbps)
  - 128 GB/s bidirectional bandwidth
  - Auto frame retry, auto link retrain, and single lane failover

Sonoma: The Perfect Choice for Scale-Out

Cost

High system integration: networking, memory, fabric

Mainstream volume process technology

Mainstream TDP

Hardware offloads

Convergence

Direct attached memory

Integrated PCIe

Integrated InfiniBand

Lower latency, higher bandwidth

Cloud

Real-time application security

Excellent throughput

Software in Silicon

Optimized for Oracle software

ORACLE

Commercial Computing is converging on:

Two Socket Scale-out Topology

Processors with Many Cores and Many Threads per Core

Enterprise Processors have BW needs met with proprietary interconnect

Cloud Processors require efficiency in cost, power, packaging, virtualization

Smaller Capable Systems with standardized ports require extreme pin BW

ORACLE

Hardware and Software
Engineered to Work Together