Ethernet Alliance
Technology Exploration Forum 2014
"The Rate Debate"

Server Diversity:
Why One Speed Is No Longer Enough
A hastily prepared and poorly organized dissertation on server networking!

David Chalupsky, Intel
October 16, 2014

This presentation has been developed within the Ethernet Alliance, and is intended to educate and promote the exchange of information. Opinions expressed during this presentation are the views of the presenters, and should not be considered the views or positions of the Ethernet Alliance or Intel Corporation.

www.ethernetalliance.org

Abstract

- Server connectivity is often used as a leading use case driving a new Ethernet standards project. The question is often asked, "What speed does a server need?" While great efficiencies could be realized by the industry by adopting a single network interface for the market majority, servers today are defying a simple description and instead diversifying in many aspects. The result is a proliferation in the Ethernet ports speeds and media being defined and adopted. This presentation aims to describe server market variables such as hardware varieties, software applications and benchmarks, and design & distribution ecosystem.

Motivation

- Server connectivity has been highlighted as the leading use case for many new Ethernet standards development projects
  - 40 Gb/s in HSSG, 2007
  - 100G Gb/s 4x25G Cu, 2010
  - Next Generation (40G) BASE-T CFI, 2012
  - 25 Gb/s CFI, 2014

That leaves me with questions:
- Are we that bad at predicting the future?
  - ...or is there just more than one answer?
- What's the driving use case for the next rate?
  - 40G serial? 50G? 200G?

CFIs with Server Use Case
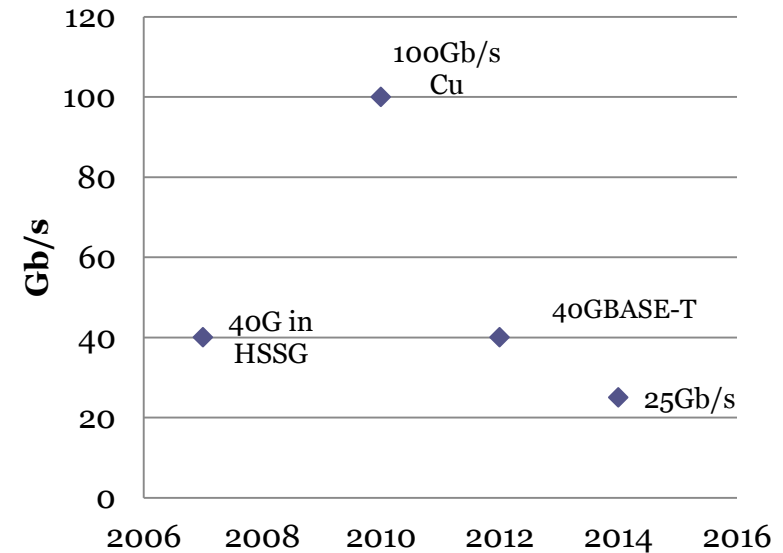
What speed does a Server need?
...are you expecting one answer?

10/16/2014

The Perfect Balance

- If we could identify the perfect balance of server hardware, applications and Network I/O, we could predict the next optimal rate for Ethernet

- This seemingly simple task is elusive because of the great diversity in server
  - Hardware
  - Applications and workloads
  - End-user priorities

www.ethernetalliance.org

Server Hardware

- **Definition**: "IDC defines a server system as a multiuser computing device that accesses and delivers services via a network… " *IDC's Worldwide Server Tracker Taxonomy, 2014*

- Major Components
  - CPUs (Central Processing Unit)
  - Local Memory (DRAM)
  - Storage (hard disk v. solid state)
  - I/O
    - expansion interfaces, like PCI Express
    - Network

- Server-like hardware is also used for other system types
  - Storage, security appliance, load balance, ADC, Communications
  - Server Class Ethernet controllers used in all areas and hard to track

- Diversity starts here…

# A Plethora of CPUs for Datacenters

- Multiple architectures
  - x86, EPIC, Power, Sparc, HP-RISC, ARM, other RISC & CISC...
- Core counts from 1 to dozens
  - Both Big Cores & Little Cores!
- Power dissipation from a few watts to a few hundred watts
- Memory capacity from GB to TB
- List price from 2-digits to 4-digits per socket ($US)

Just the x86 server CPU line is confusing

- 7-8 tiers
- 100's of SKUs
- Users have the choice to pay for the performance they need

Intel's Data Center Group (DCG) Public Roadmap

www.ethernetalliance.org
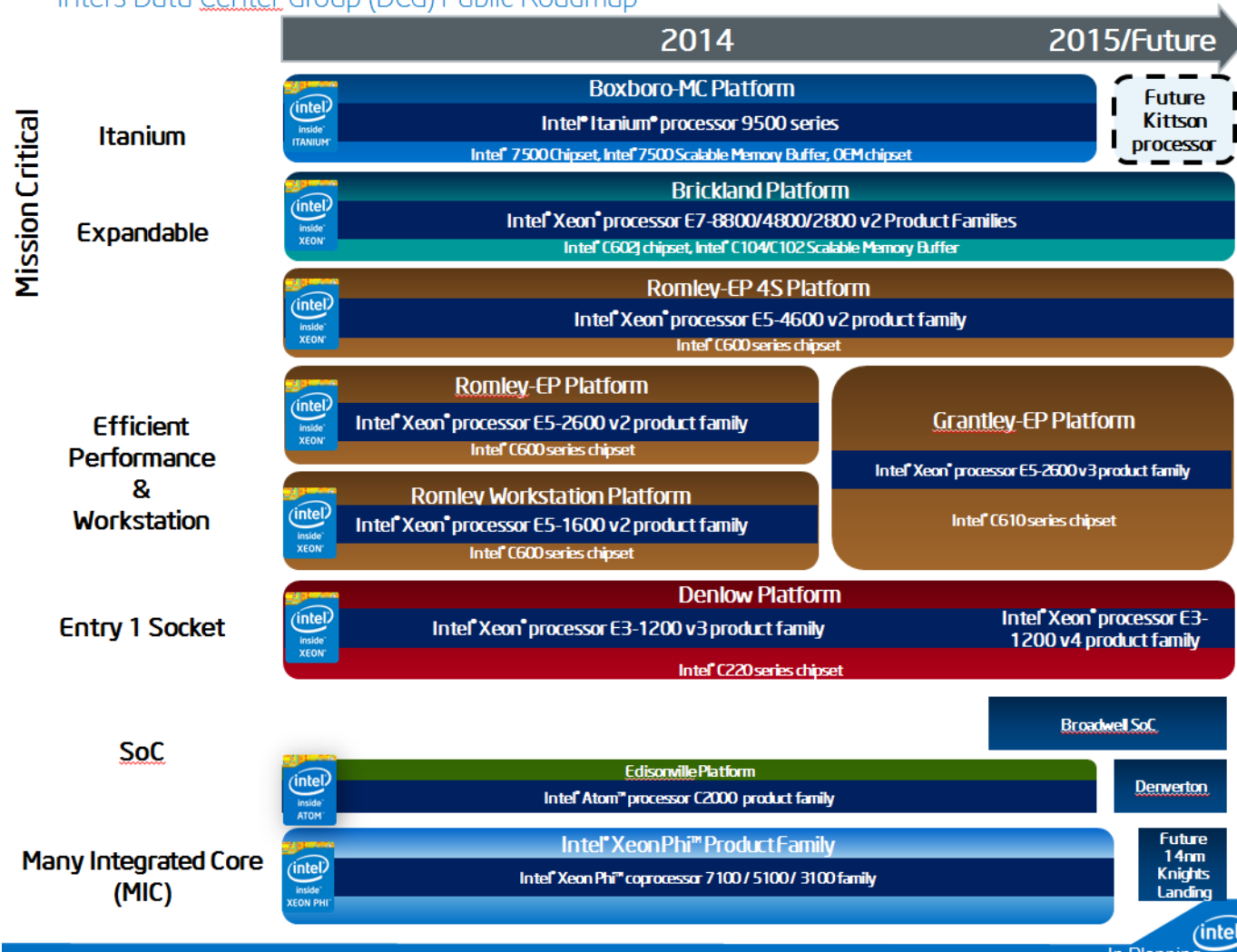
Storage & Memory

Drives

- *Warehouse Scale Computer* (WSC) originally built with desktop-class HDDs to save cost
- Industry transition from HDD to SSD
    - Significant reduction in power dissipation
    - Massive increase in access speed.

- *"Flash's performance is so high that it becomes a challenge to use it effectively in distributed storage systems since it demands much higher bandwidth from the WSC fabric"*
    - *Louis Barroso, Urs Holzle, Jimmy Clidaras - Google*

- Memory
    - Faster interface: DDR4 shipping now
    - More capacity: ~1/2TB mainstream, ~12TB high end
    - System memory easily expanded via socketed DIMMs

What the Analysts Track

- Multiple CPU architectures
- 4 major form factor types
  - Blade, Rack optimized, Density Optimized, Tower
- 6 operating system categories
- 23 price bands (3 to 7 digits)
- 18 workload categories

...more

- *IDC Server Tracker*

- *Are you still expecting one answer?*

www.ethernetalliance.org

Hardware Summary

- We want a balanced, optimized system

- There are so many choices in
- CPU
- Memory
- And Storage...
- The choices in Ethernet seem limited in comparison

www.ethernetalliance.org

Server Workloads

- "In the Future…
  - Systems will be workload optimized
  - Infrastructure will be software defined
  - Analytics will be pervasive"
- Intel Developer Forum keynote, 2014

Key Metrics

- Parallelizability
- Amount of Inter-Process Communication (IPC)
- Size of the data set

Scale Out vs. Scale Up

- Scale out: adding servers in parallel
  - Usually means low cost servers
- Scale up: bigger servers
  - more cores sharing more memory

- "Scaling up is far more widespread than scaling out across all workloads." IDC Server Workloads Forecast, 2012

- "Clusters of low-end servers are the preferred building blocks for WSCs (warehouse scale computers) today"
  - *The Datacenter as a Computer, An Introduction to the Design of Warehouse-Scale Machines*, Barroso, Clidaras, and Hölzle, Google, 2013

- Why can't we agree?
  - Parallelization, IPC, and size of data set...

The Relationship

- Many workloads are parallelizable
  - Some "embarrassingly parallel" ibid

- IPC is faster in memory than over the LAN
  - Nanoseconds v. microseconds
- Workloads vary in the amount of IPC required
  - High IPC, want the database in one machine (scale up)
  - Low IPC, scale out vs. up doesn't matter much

- If the database is huge, no choice but to scale out
  - Indexing the entire internet: a really big database

- A scale out application with high IPC can benefit from faster LAN
  - It makes the whole cluster look like one machine… a warehouse scale computer

In-Memory Computing: Game Changing

In-Memory **148x** Faster

Disk Based

Test results based on an internal 10 TB database Proof of Performance and Scalability (POPS) benchmark with IBM® DB2® 10.1 running on a 4-socket Intel® Xeon® processor E7-4870 with 1,024 GB of memory, and DB2 10.5 with BLU Acceleration running on a 4-socket Intel Xeon processor E7-4890 with 1,024 GB of memory

From Intel Xeon E7 v2 launch, February 2014
Memory capacity increase form 4TB to 12TB, increasing the utility of scale up

10/16/2014

Big Data Analytics

- Big Databases
- Highly compute intensive
- Myriad new applications emerging for scale up

- My favorite example: Genome sequencing
  - From 3 to >250 sequences computed per hour
  - We can sequence your healthy cells...
  - Sequence your cancer cells...
  - And engineer antibodies that target the cancer
  - Ari E. Berman, PhD, Director of Government Services and Principal Investigator, BioTeam, Inc.

  - We are finally doing something useful with computers!

Rapidly Changing Workloads

- "Key pieces of Google's services have release cycles on the order of a couple of weeks… This environment… makes it hard for a system designer to extract useful benchmarks even from established applications."
  - *The Datacenter as a Computer, An Introduction to the Design of Warehouse-Scale Machines*, Barroso, Clidaras, and Hölzle, Google, 2013

- The workloads will change in nature over the life of the datacenter

- That makes it harder to find one answer…

Summary #2

- Server hardware is very diverse...
- Server workloads are very diverse
- The data rate options of Ethernet seem... kinda limited in comparison

# End User Environments

Cloud vs. Enterprise

Tiny definition:
- Public Cloud: computing *is* the business
- Enterprise IT: computing supports the business

- "…expect technology bifurcation driven by the contrasting needs of enterprise datacenters and cloud service providers to continue."
  - IDC server workload forecast 2012

Server Deployment in Cloud vs. Enterprise

Figure 1: Server Shipments

Source: Dell'Oro Controller and Adapter Report, July 2014
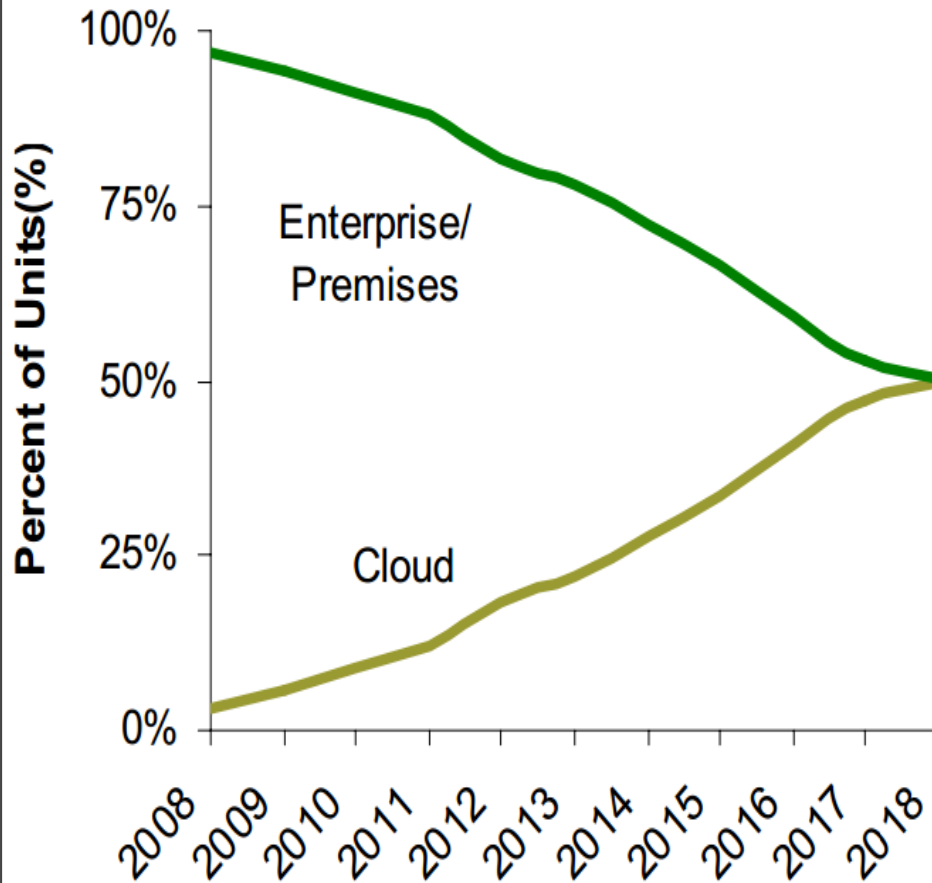
Cloud: ~1/4 of server units 2014

- Growing fast
- Mix expected to stabilize
- A small number of entities comprise the majority of this market
- Most, but not all, hyperscale datacenters

Enterprise:

- Largely follow TIA-942, and deploy a variety of access switch placements including MoR and EoR.
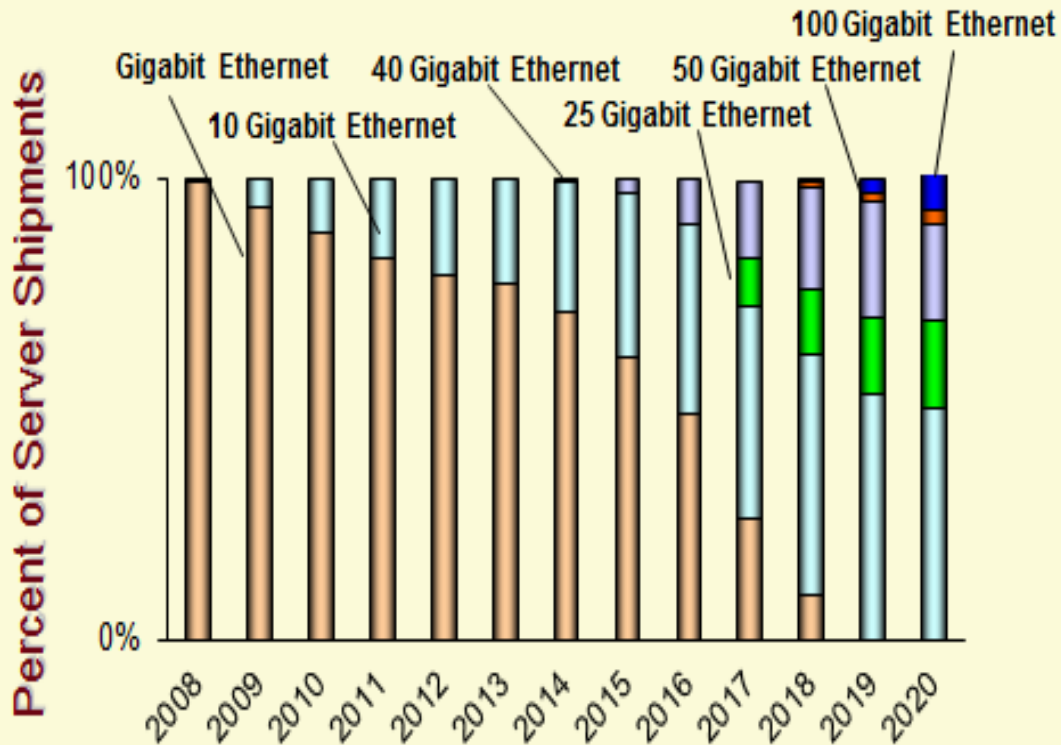
Speed Migration on Cloud Servers

**Speed Migration on Cloud Servers
(Included in Dell'Oro Group's Server Report)**
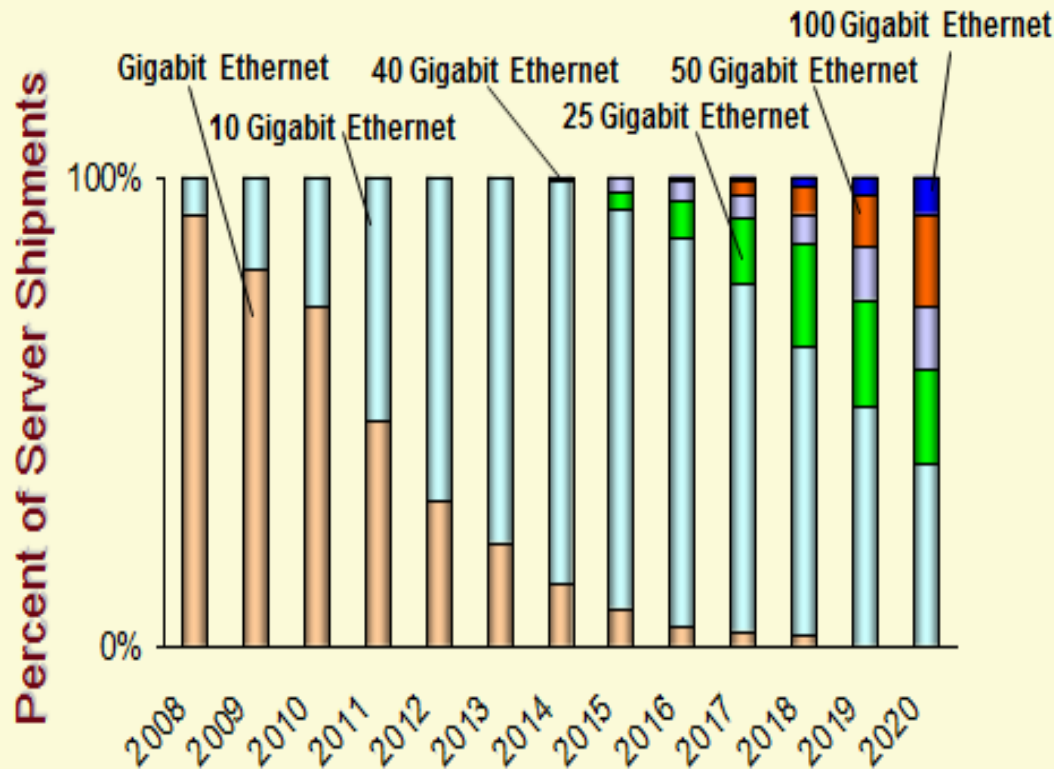
- Transition to 10G nearly complete
- Early adoption of 40G, but 25G/50G expected to overtake
- Standardization not imperative
  - Engineered systems

Speed Migration on Premise Servers

- 1000BASE-T still majority of ports
- 10G growing
- 40G still significant future
- IT & datacenters are a cost center
- Investment depends upon economy
- Standardization important
  - 25G appears once standard
  - multi-vendor, off the shelf hardware

What is Cloud Scale?

Microsoft Online Services:

- over 200 cloud services
- to 1B customers,
- 15M businesses,
- 90 markets.
- Over 1 million servers in service
- Kushagra Vaid, GM Server Engineering, Microsoft Cloud & Enterprise Division, 2014

The Enterprise

Very Rough Guess: how many Servers are owned by medium & large companies in the U.S.?

Org Size	Average # Servers[1]	company count[2]	total servers
Medium Organization (100-1000)	193	30004	5,790,772
Large Organizations (>1000)	720	8910	6,415,200
			12,205,972

- Of course, this misses the 5.6 million U.S. firms <100 employees… they might have a few servers…

- "The number of servers medium and large enterprises have might be sobering for networking vendors that like to talk about how large their solutions can scale: the majority of enterprises have ~100 servers and need only about 6 ToRs with 48 ports of 10GE."[1]

- 1 - Infonetics Data Center Strategies North American Enterprise Survey, September 24, 2014
- 2 – United States Census Bureau, 2011

10/16/2014

www.ethernetalliance.org

In Summary

- Thee is a huge diversity in server hardware
- There is a huge diversity and rapid change in server workloads
- There are large differences in end user behavior
- The choices for Ethernet data rate seem limited in comparison

- We cannot expect that one Ethernet speed will serve all
- 1G to 100G Ethernet will coexist for years to come

Server Diversity:
Why One Speed Is No Longer Enough

I hope you found the answer!!!!

David Chalupsky
Intel Corp.
October 2014