



Interconnect Challenges in a Many Core Compute Environment

Jerry Bautista, PhD
Gen Mgr, New Business Initiatives
Intel, Tech and Manuf Grp

**Interconnect Challenges in a Many Core
Compute Environment**
*Jerry Bautista, Director, Technology
Management, Microprocessor Programming
and Research Lab, Intel*

Agenda

Microprocessor – general trends

Implications

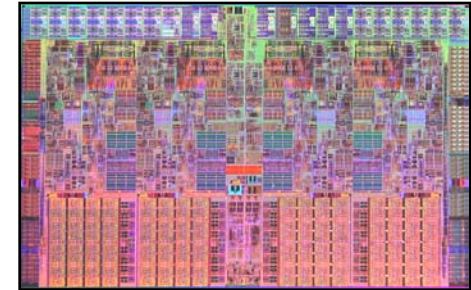
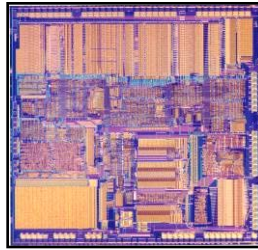
Tradeoffs

Summary

Microprocessor Evolution

2009

1985



Intel386

Nehalem

Transistor Count

280 Thousand

731 Million

Frequency

16 MHz

> 3.6 GHz

Core Count

1

4

Cache Size

None

8 MB

I/O Peak BW

64 MB/s

50 GB/s

Adaptive Circuits

None

Sleep Mode

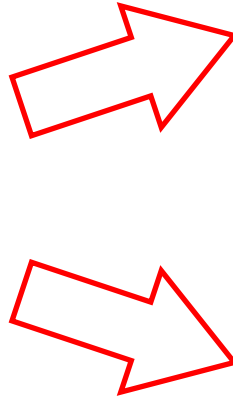
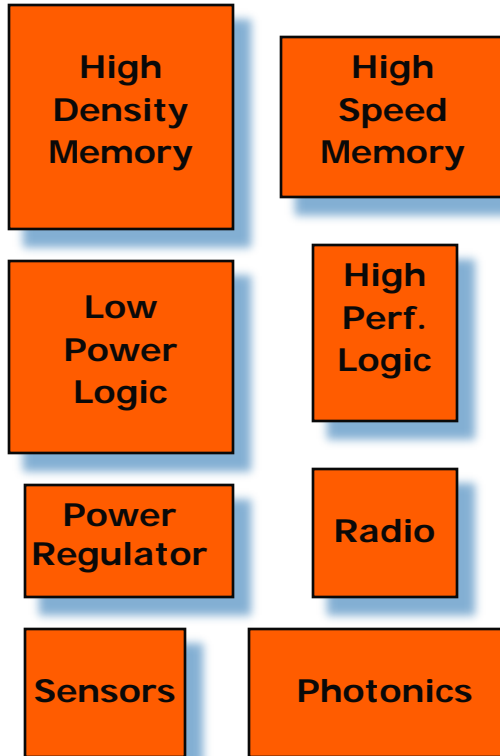
Turbo Mode

Power Gating

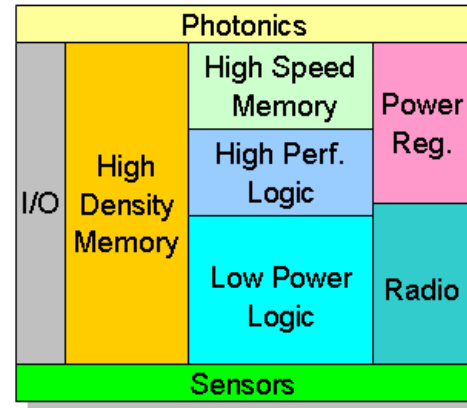
Adaptive Frequency Clocking

System Integration

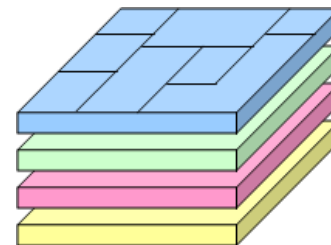
Discrete



2-D Integration (SoC)



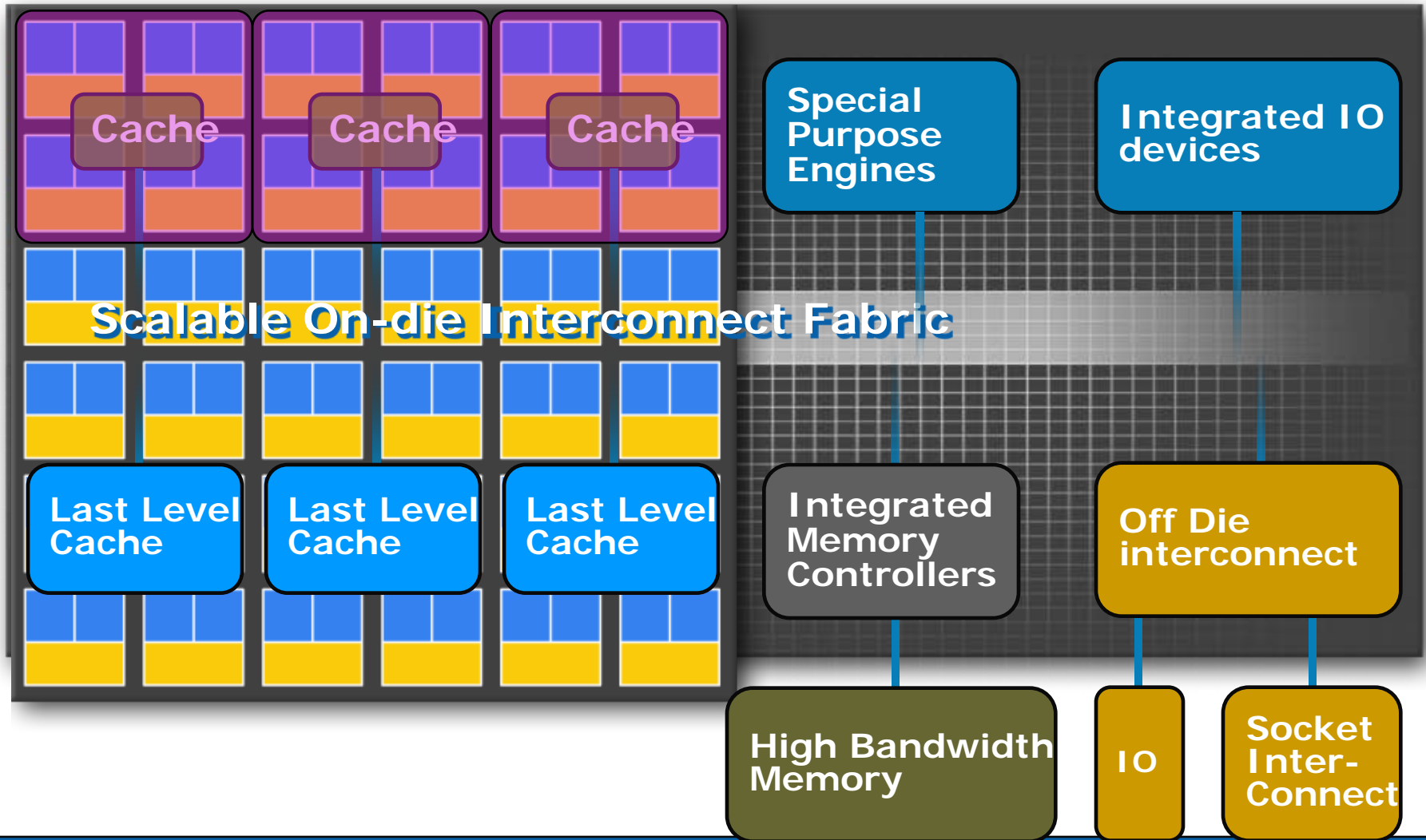
3-D Integration



Logic
Memory
Power Reg.
Radio
Sensors
Photonics

System integration needed for performance, power, form factor
Challenge is to integrate wider range of heterogeneous elements

A Tera-scale Platform Vision



Multi-core today becomes many core tomorrow

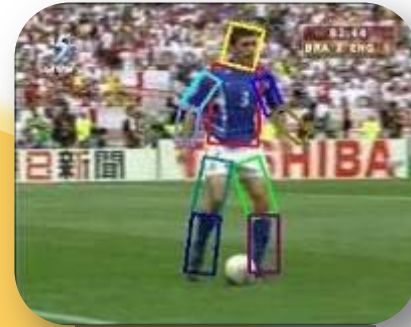


Visual Computing – 3D and more



**Ray-traced
Graphics**

*Looks real
Acts real*

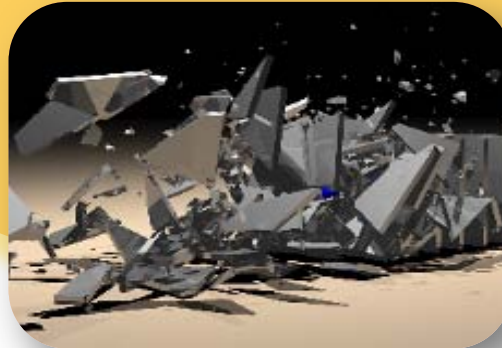


Video Search

...Feels real



Expressive Faces



**Physics-based
Animation**



Computer Vision

Acts Real = Physics and Behavioral Intelligence
Looks Real = Life-like Images

Platform Performance Demands

SERVERS: 10x More Work

MMORPGS

WoW

2500

VWs

Second Life

160

CLIENTS: 3x CPU, 20x GPU

APPLICATION

% CPU
UTILIZATION

% GPU
UTILIZATION

2D Websites

20

0-1

Google Earth

50

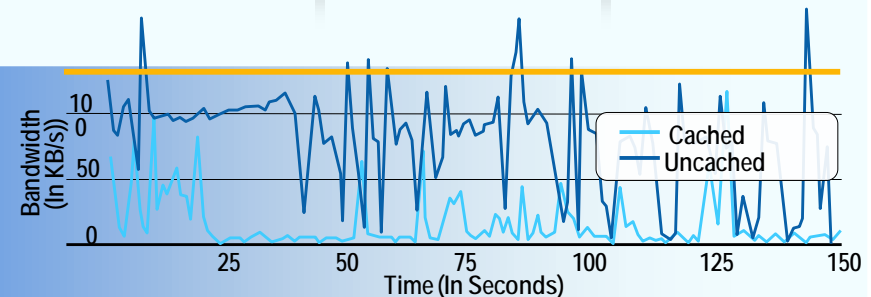
10-15

Second Life

70

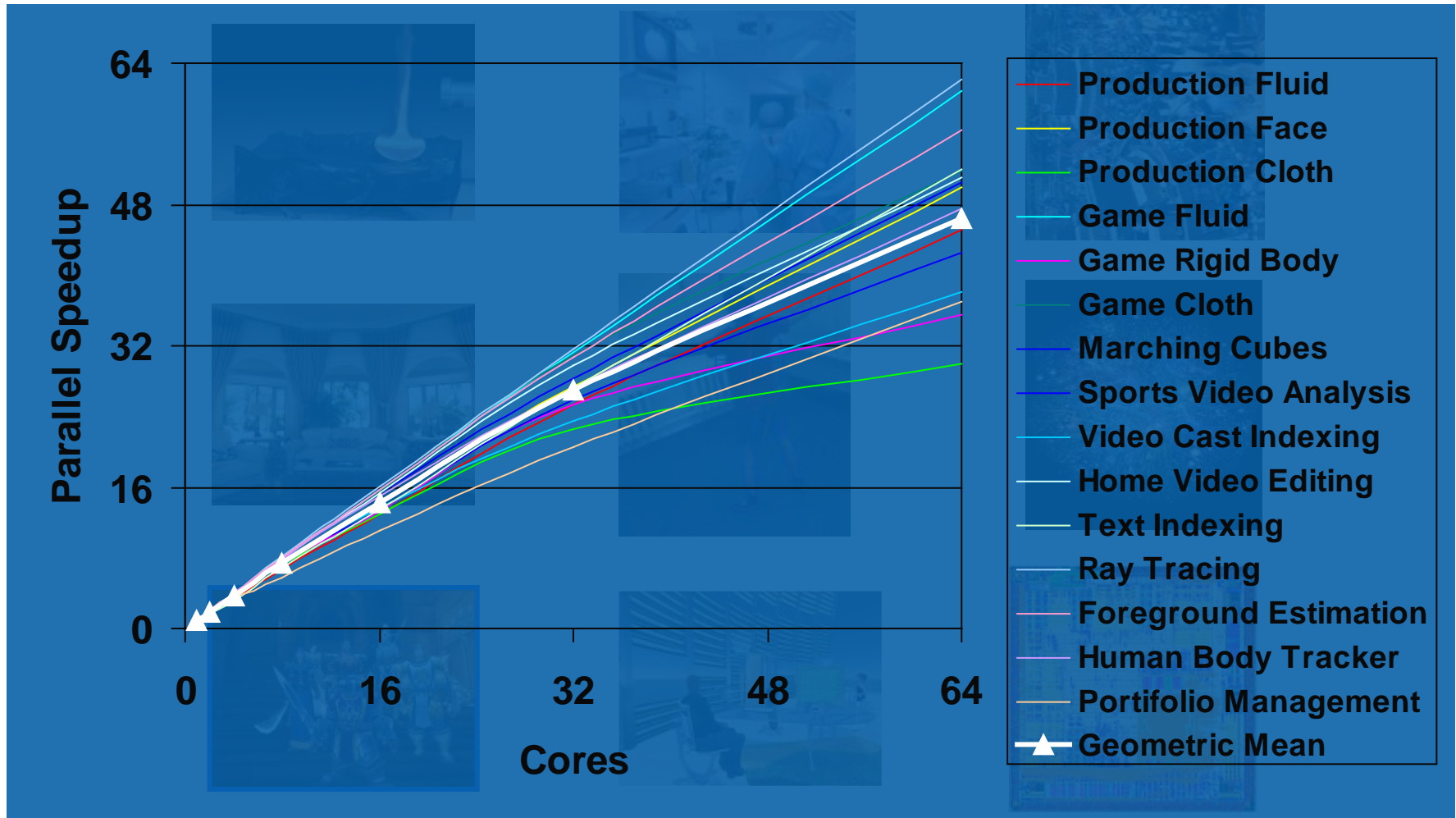
35-75

NETWORK: 100x Bandwidth
Maximum Bandwidth Limited by
Server to Client



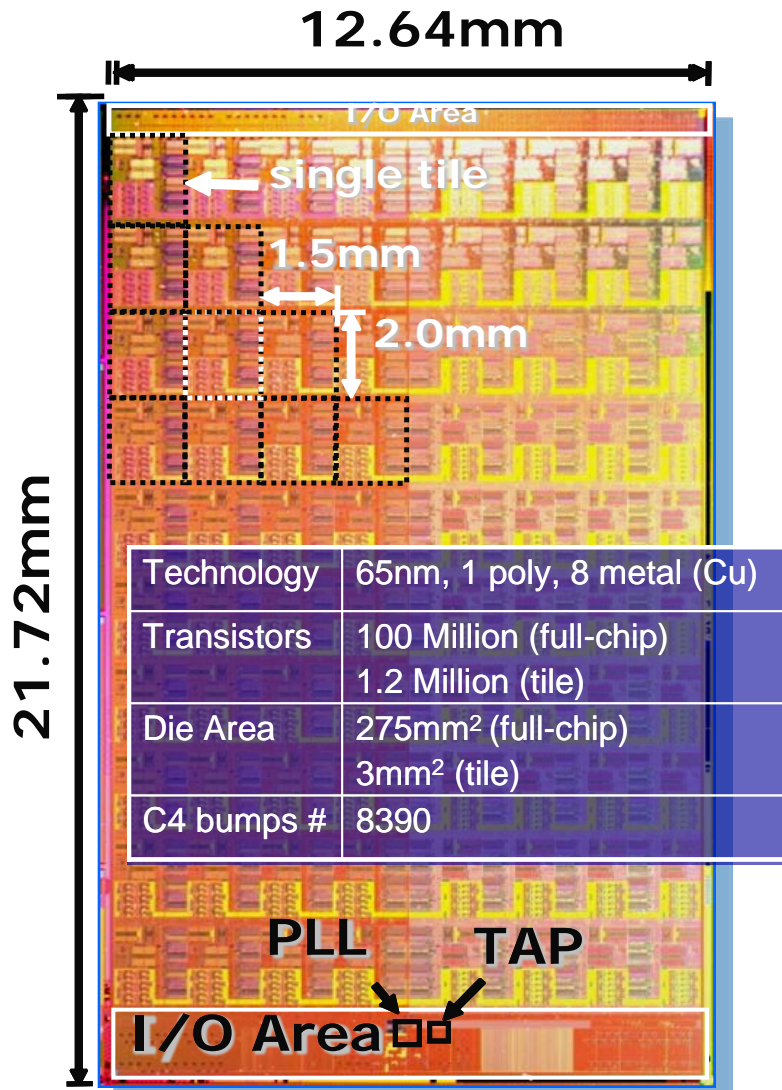
Sources: WoW data (source www.warcraftrealms.com), Second Life data (source CTO-CTO meeting and www.secondlife.com), and Intel measurements.

Research shows Applications Scale Well



Many applications can be “parallelized” across a large number of core with almost 1:1 speed-up

Teraflops Research Processor



Goals:

Deliver Tera-scale performance

- Single precision TFLOP at desktop power
- Frequency target 5GHz
- Bi-section B/W order of Terabits/s
- Link bandwidth in hundreds of GB/s

Prototype two key technologies

- On-die interconnect fabric
- 3D stacked memory

Develop a scalable design methodology

- Tiled design approach
- Mesochronous clocking
- Power-aware capability

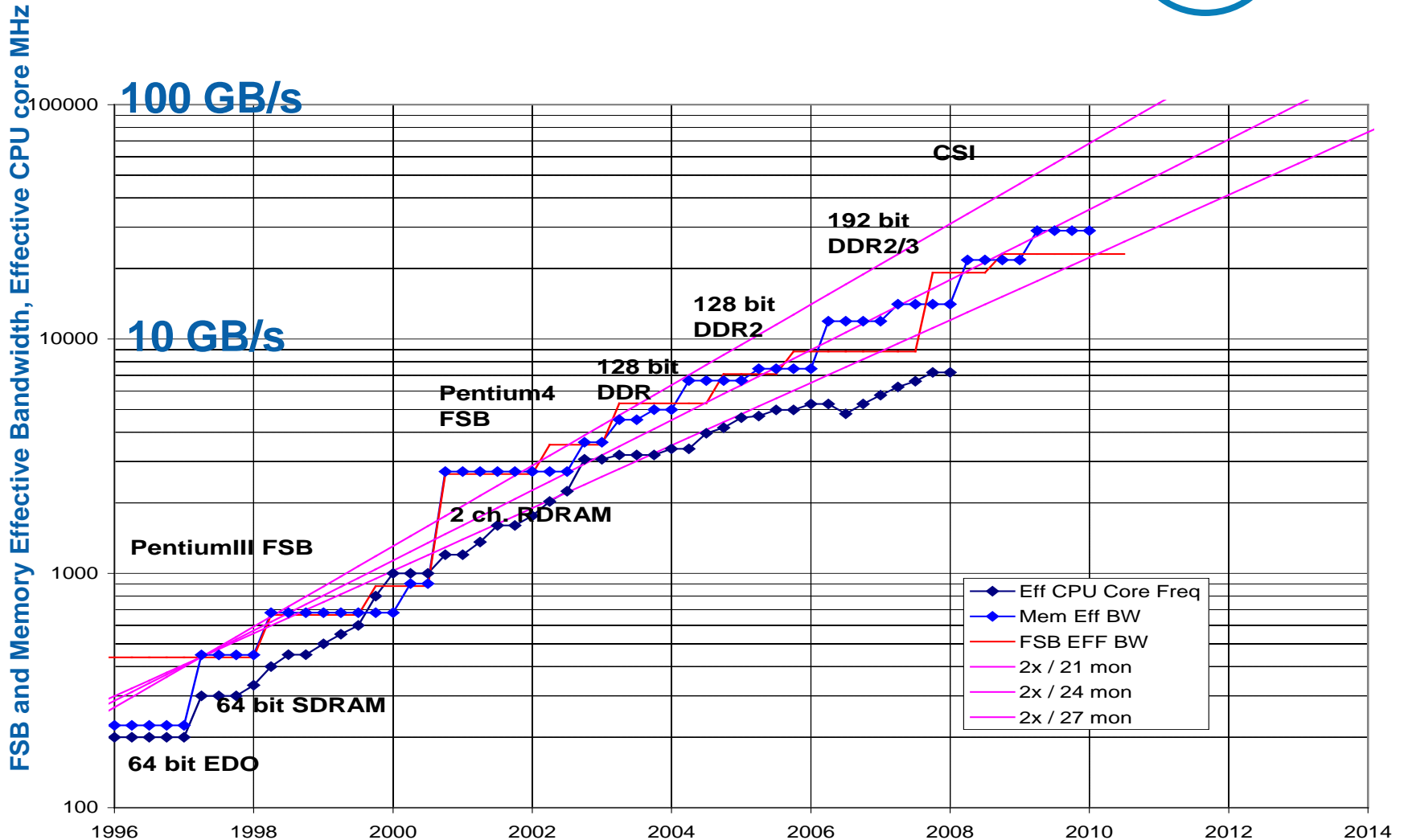
1 TeraFLOP in < 60 W envelope

100 Million Transistors ● 80 Tiles ● 275mm²

Desktop Platform Bandwidth Roadmap

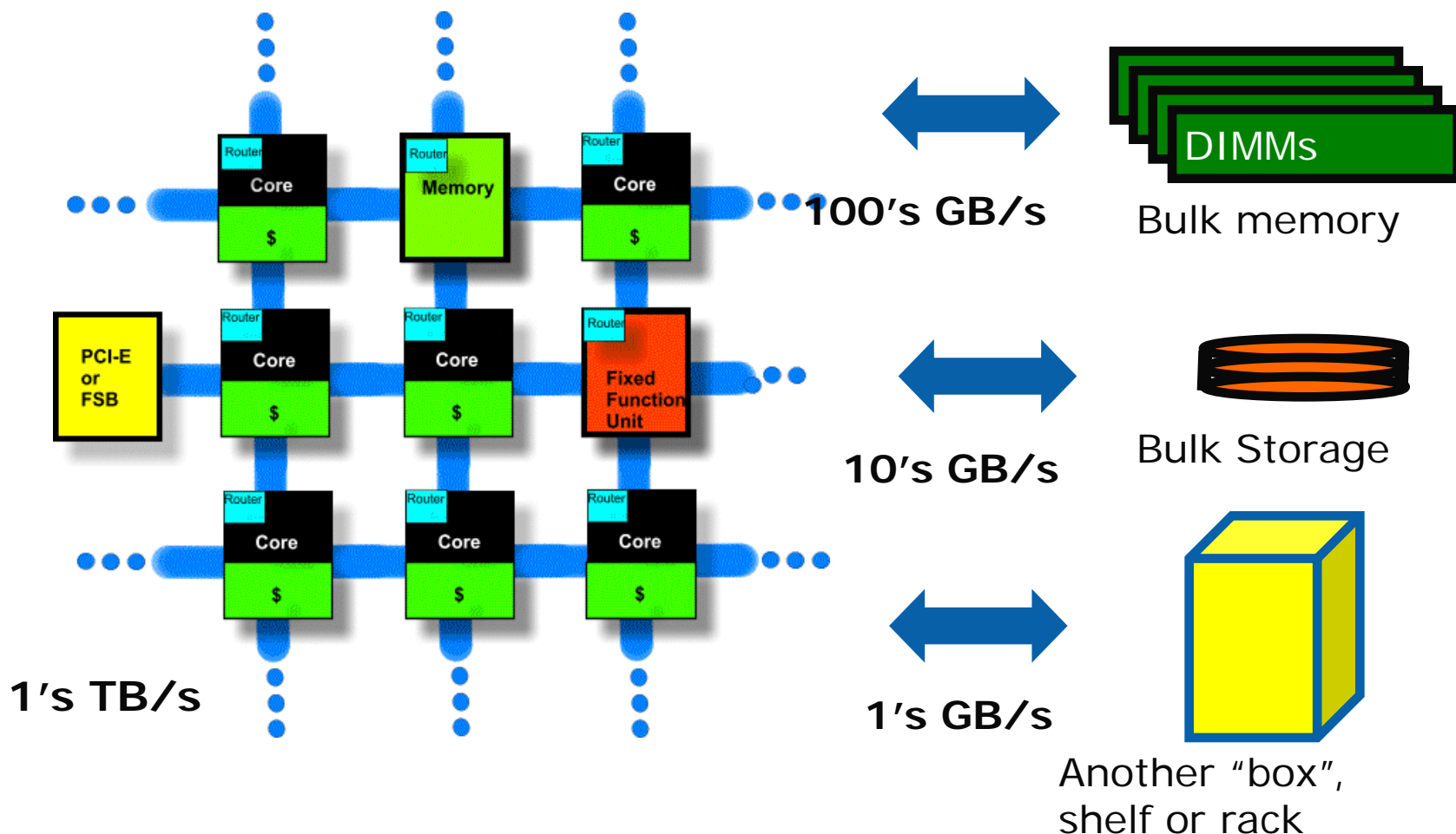
- driven by 2x perf/2 yrs

Tera-scale
Computing
100's GB/s



Staying "on-trend" puts memory BW target at 25-40 GB/s for the 2009-2011 time frame - Tera-scale implies a 10x increase beyond that!

Future CPU's – Tera-Scale Performance



Tera FLOPS implies tera bytes/sec bandwidth

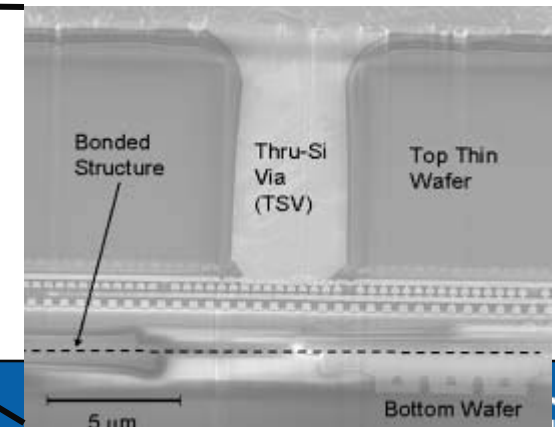
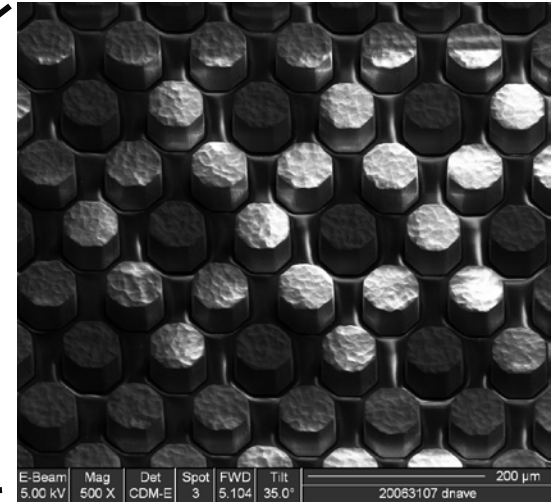
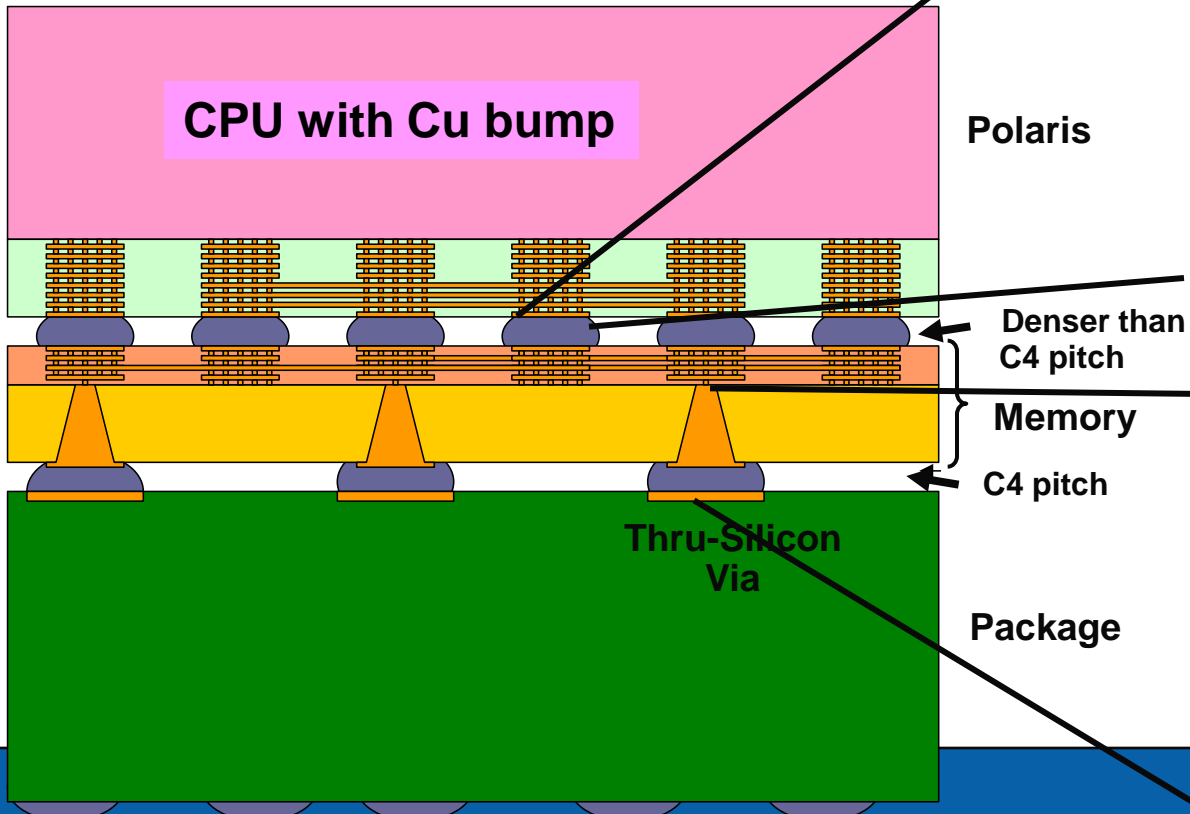


I/O Link Attributes

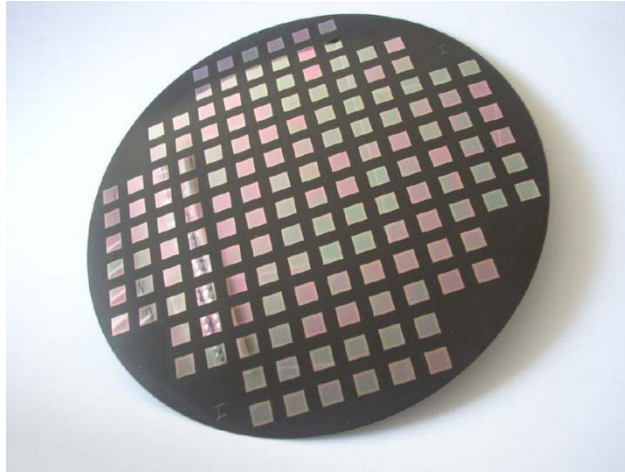
Application Attributes	Box-Box	Package-Package	Chip-Chip	On-Die
Distance	Long (~10m)	Medium (~10's cm)	Medium (~1cm)	Short (~1mm)
Speed	Low (10Gbps)	Medium (10-40Gb/s)	High (40Gbps)	High (40+Gbps)
Power	High (10's mW/Gbps)	Medium (<5mW/Gbps)	Low (<1mW/Gbps)	Low (<1mW/Gbps)
Temperature	Low (<70C)	Medium (up to ~80C)	High (up to ~100C)	High (up to ~100C)
Form Factor	Large (cm's)	Medium (~1cm)	Small (a few mm's)	Small (a few mm's)
Physical Layer	Optical or Electrical Std	Hybrid and/or Optical Electrical	3D stack or Die array	On-die metal layers

Stacking Memory w/ CPU

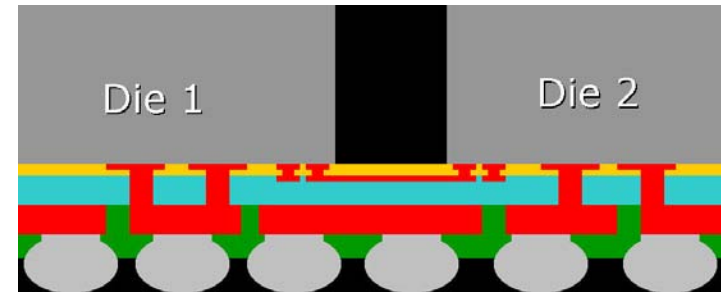
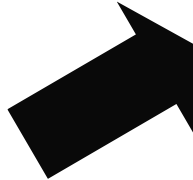
256 KB SRAM per core
4X C4 bump density
8490 thru-silicon vias



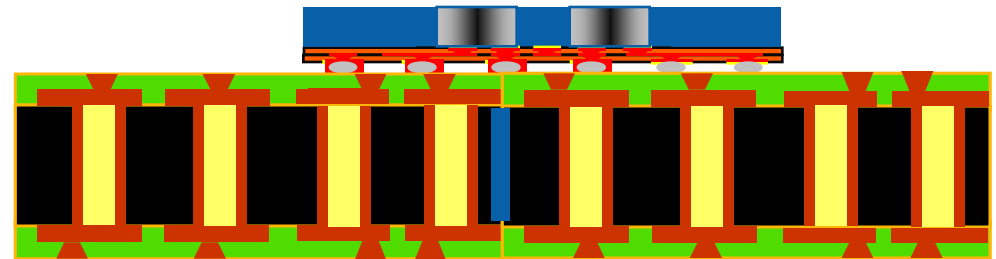
BBUL-W Approach



Dies reconstituted into molded reconfigured wafer*



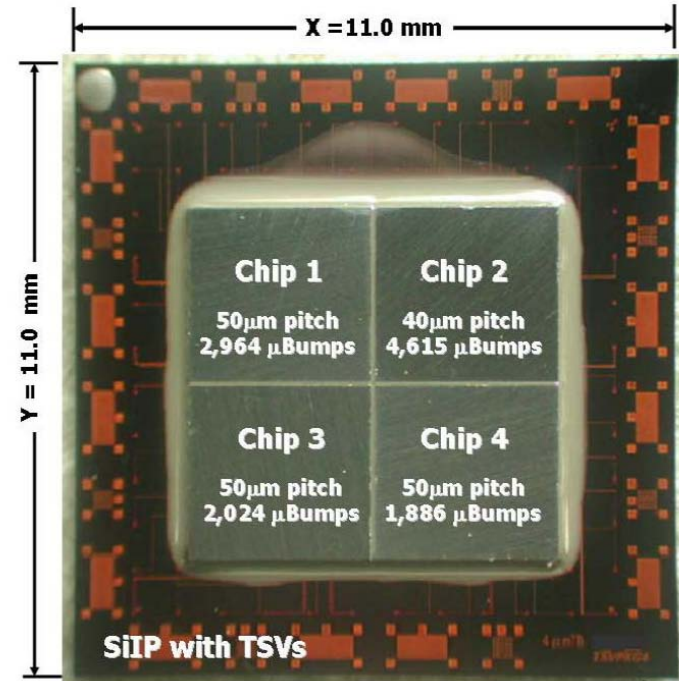
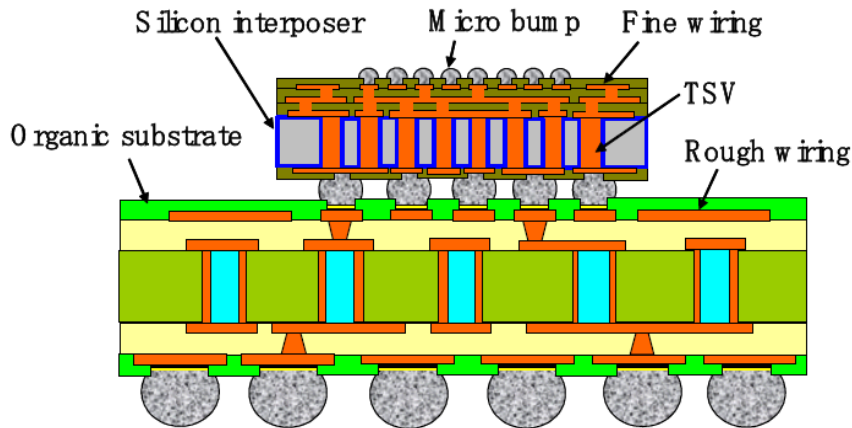
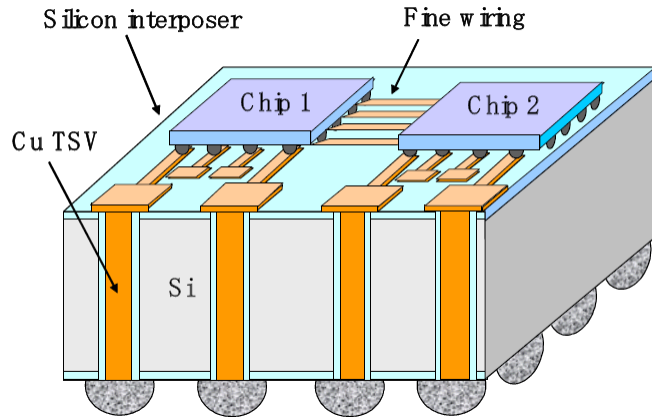
Wafer thin film processes used to interconnect dies



Singulated BBUL-W package connected to interposer

*Ref.:
M. Brunnbauer *et al.*, "An embedded device technology based on a molded reconfigured wafer," in *Proc. ECTC*, pp. 547, May 2006.

Silicon Interposer



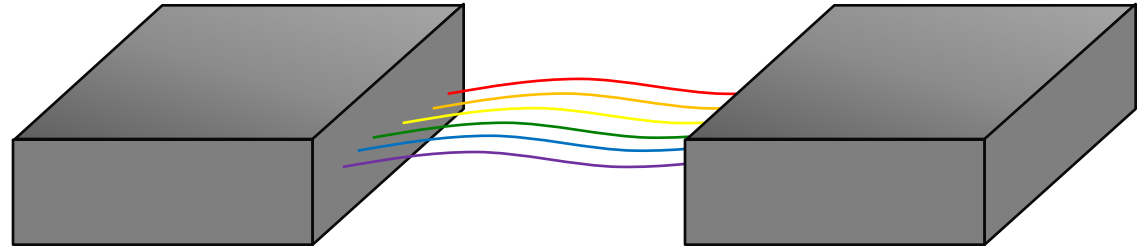
Refs.:

M. Sunohara *et al.*, "Silicon interposer with TSVs (through silicon vias) and fine multilayer wiring," in *Proc. ECTC*, pp. 847, May 2008.

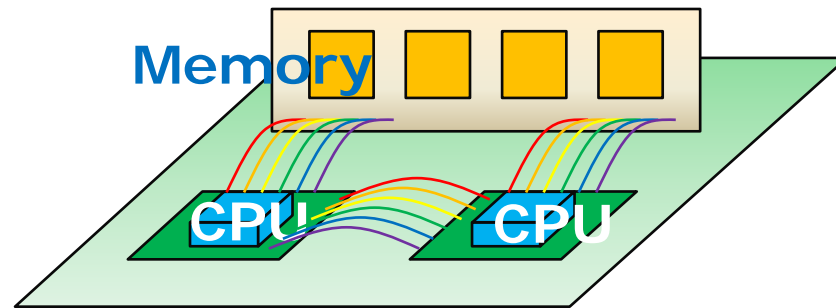
K. Kumagai *et al.*, "A silicon interposer BGA package with Cu-filled TSV and multi-layer Cu-plating interconnect," in *Proc. ECTC*, pp. 571, May 2008.

Optical Interconnect on CPU Platform

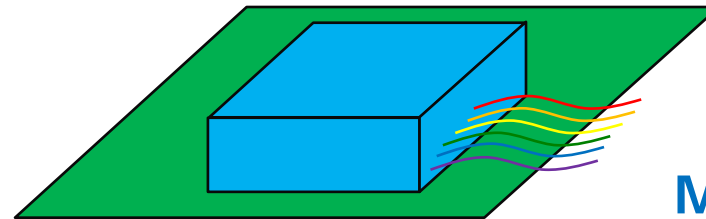
Box-Box



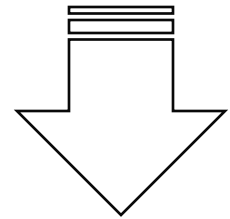
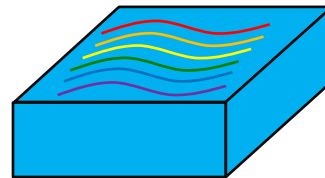
Package-Package



Chip-Chip



On-Die/On-Package



More Difficult
Longer Term

Optical Infiniband*



Product Parameters:

- Copper cable replacement
- 4 channels in each direction at up to 5Gb/s per channel for DDR product (10Gb/s per channel for QDR product)

Product description:

- Two 4 channel transceivers coupled with 12 channel 62.5 or 50um fiber
- Plastic lens for optical coupling
- 1x4 array VCSEL and PIN
- 4 channel tx and rx IC designed by Intel, and fabbed by TSMC
- dies are attached directly onto a PCB and wire bonding from chip to chip and chip to pcb

*No longer an Intel product...

Summary

- Parallelism in Many-Core platforms is key to delivering continued performance and fine-grained power management
- Future applications stress the entire interconnect stack, from on-die to the broader network.
- New multi-chip packaging approaches are required to address more aggressive integration.
- Optical to electrical transition shrinking...now approaching sub 1 m, but highly dependent upon specific application/topology.
- Entire HW/SW stack is impacted by the move to parallelism requiring broad ecosystem enablement