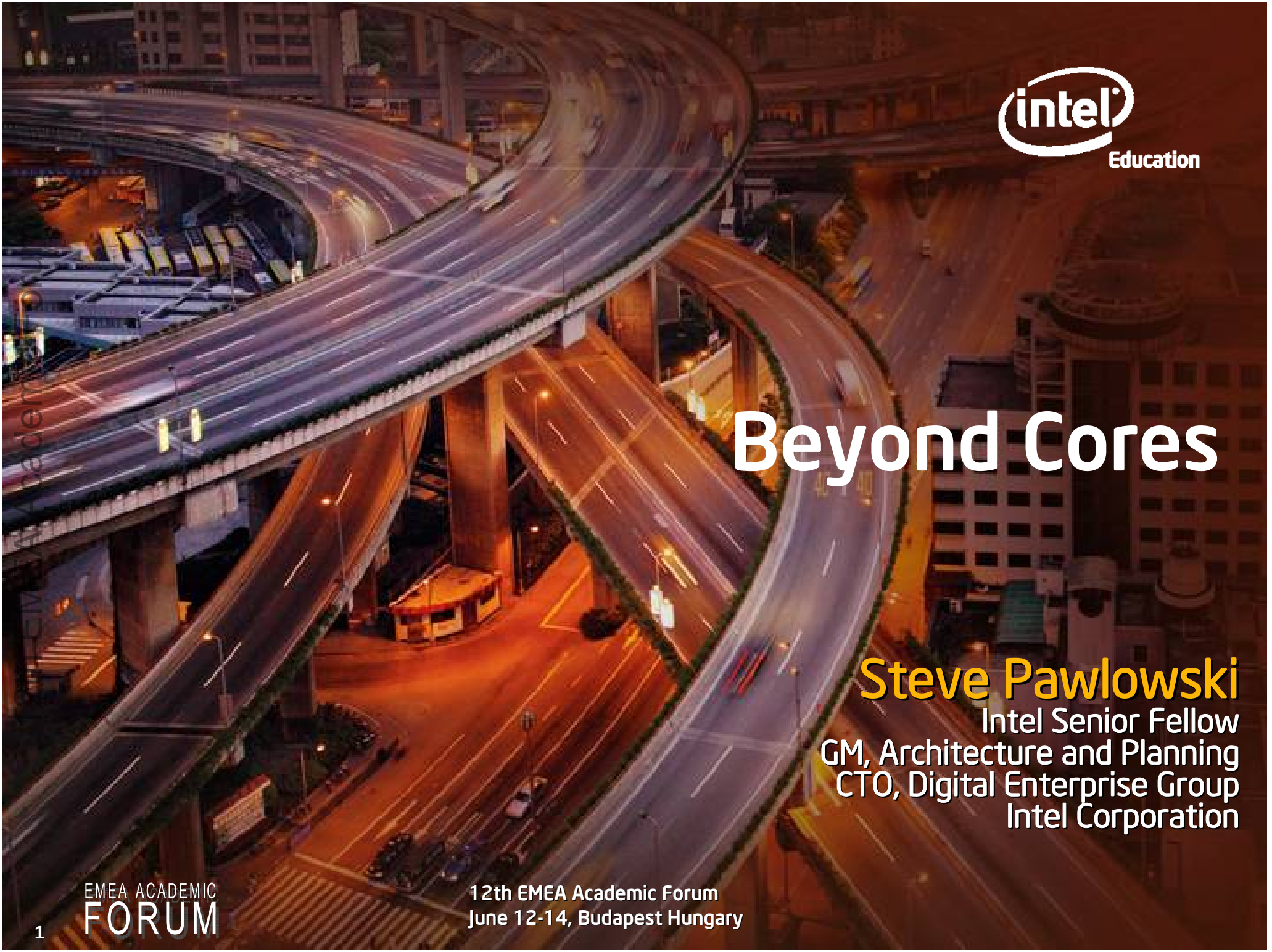

intel
Education
Beyond Cores
Steve Pawlowski
Intel Senior Fellow
GM, Architecture and Planning
CTO, Digital Enterprise Group
Intel Corporation
EMEA ACADEMIC
FORUM
12th EMEA Academic Forum
June 12-14, Budapest Hungary

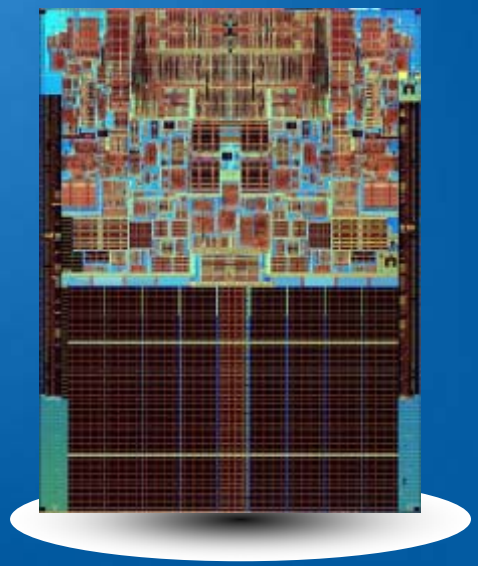
Legal Disclaimer

- INFORMATION IN THIS DOCUMENT IS PROVIDED IN CONNECTION WITH INTEL® PRODUCTS. EXCEPT AS PROVIDED IN INTEL'S TERMS AND CONDITIONS OF SALE FOR SUCH PRODUCTS, INTEL ASSUMES NO LIABILITY WHATSOEVER, AND INTEL DISCLAIMS ANY EXPRESS OR IMPLIED WARRANTY RELATING TO SALE AND/OR USE OF INTEL PRODUCTS, INCLUDING LIABILITY OR WARRANTIES RELATING TO FITNESS FOR A PARTICULAR PURPOSE, MERCHANTABILITY, OR INFRINGEMENT OF ANY PATENT, COPYRIGHT, OR OTHER INTELLECTUAL PROPERTY RIGHT.
- Intel products are not intended for use in medical, life saving, life sustaining, critical control or safety systems, or in nuclear facility applications.
- Intel Corporation may have patents or pending patent applications, trademarks, copyrights, or other intellectual property rights that relate to the presented subject matter. The furnishing of documents and other materials and information does not provide any license, express or implied, by estoppel or otherwise, to any such patents, trademarks, copyrights, or other intellectual property rights.
- Intel may make changes to specifications, product descriptions, and plans at any time, without notice.
- The processors, the chipsets, or the ICH components may contain design defects or errors known as errata, which may cause the product to deviate from published specifications. Current characterized errata are available upon request.
- All dates provided are subject to change without notice. All dates specified are target dates, are provided for planning
purposes only and are subject to change.

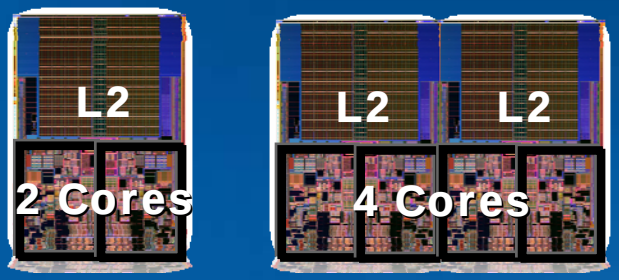
- Intel and are trademarks or registered trademarks of Intel Corporation or its subsidiaries in the United States and other countries.
- * Other names and brands may be claimed as the property of others.
-

A Snapshot of Today

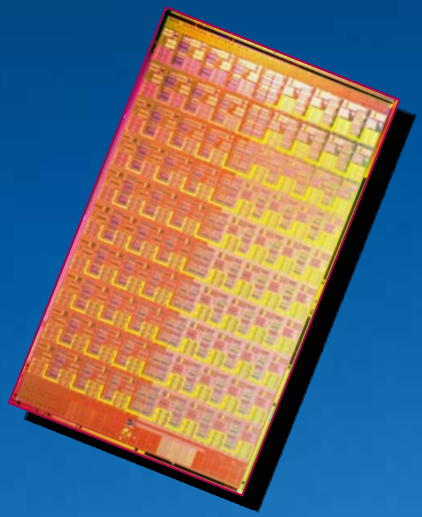
- Moore's Law continues its momentum
- Power consumption is becoming a major concern
- Multiple cores on-die has become the standard to deliver performance at reasonable power



Intel® Core™ 2 Duo
(Merom, 65nm Process)



Intel® Core™ 2 Duo
(Penryn, 45nm Process)



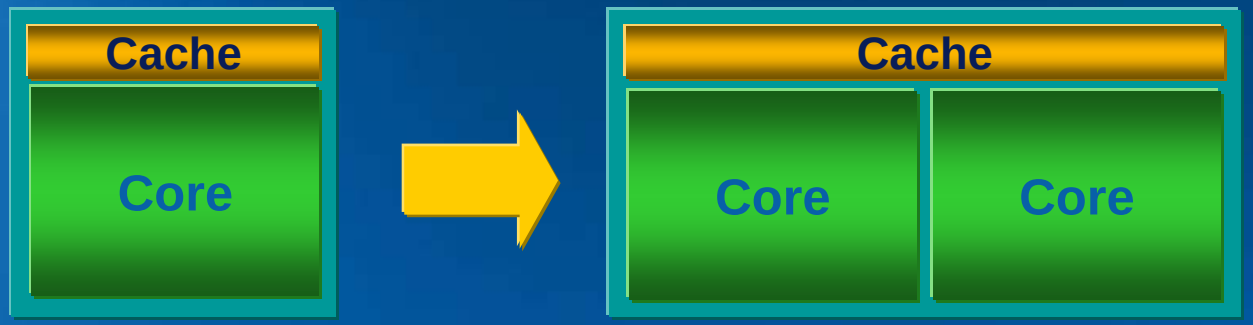
Intel Research Chip: 80 core Polaris

Why Multi/Many Cores?

Performance within the Power Envelope

Rule of thumb

Voltage	Frequency	Power	Performance
1%	1%	3%	0.66%

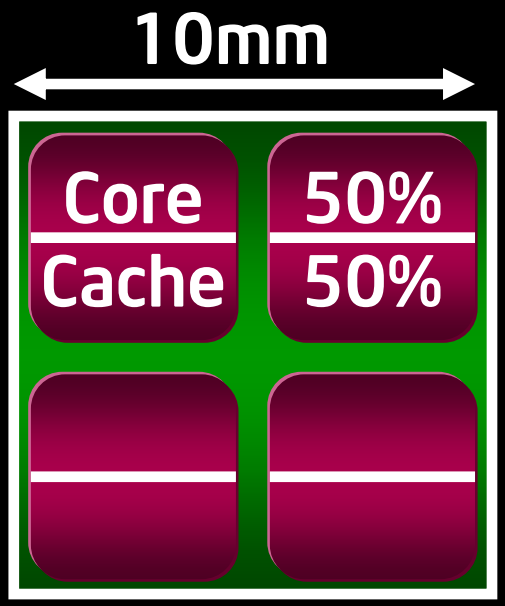


Voltage = 1
Freq = 1
Power = 1
Perf = 1

Voltage = -20%
Freq = -20%
Power = 1
Perf = ~1.7

Put Moore's Law into Great Use

10s and 100s of Cores - Not a Dream



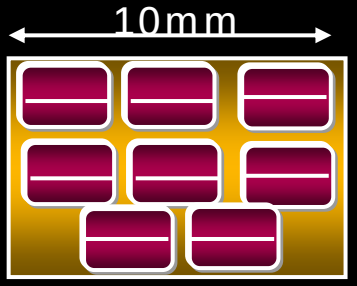
65nm, 4 Cores

1V, 3GHz
10mm die, 5mm each core
Core Logic: 6MT, Cache: 44MT
Total transistors: 200M

Per core power reduction is based on:

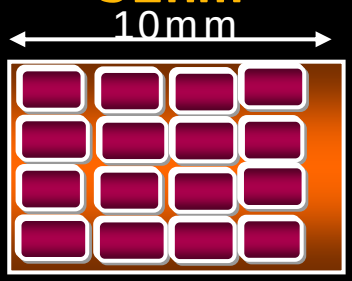
- Capacitance, voltage and frequency scaling.
- Assuming voltage and frequency scaling will slow down

45nm



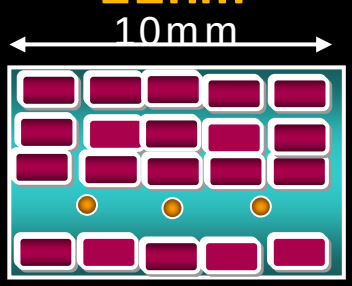
8 Cores, 1V, 3GHz
3.5mm each core
Total: 400MT

32nm



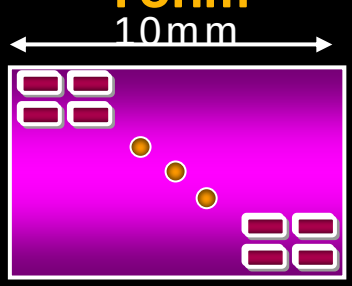
16 Cores, 1V, 3GHz
2.5mm each core
Total: 800MT

22nm



32 Cores, 1V, 3GHz
1.8mm each core
Total: 1.6BT

16nm

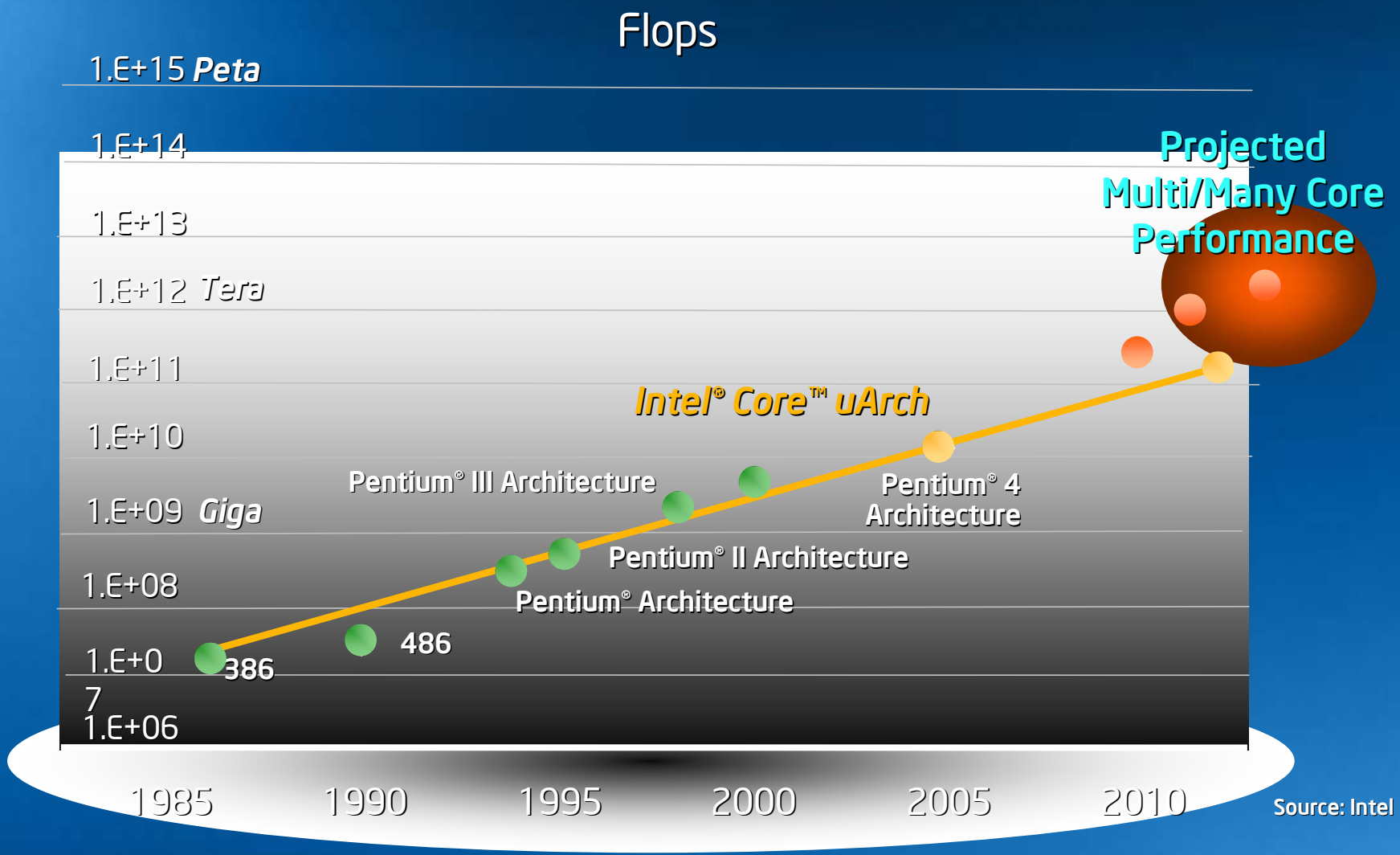


64 Cores, 1V, 3GHz
1.3mm each core
Total: 3.2BT

Note: the above pictures don't represent any current or future Intel products

Assume: Voltage and Frequency constant

Where Are We Heading with Many Core?



Source: Intel

12th EMEA Academic Forum

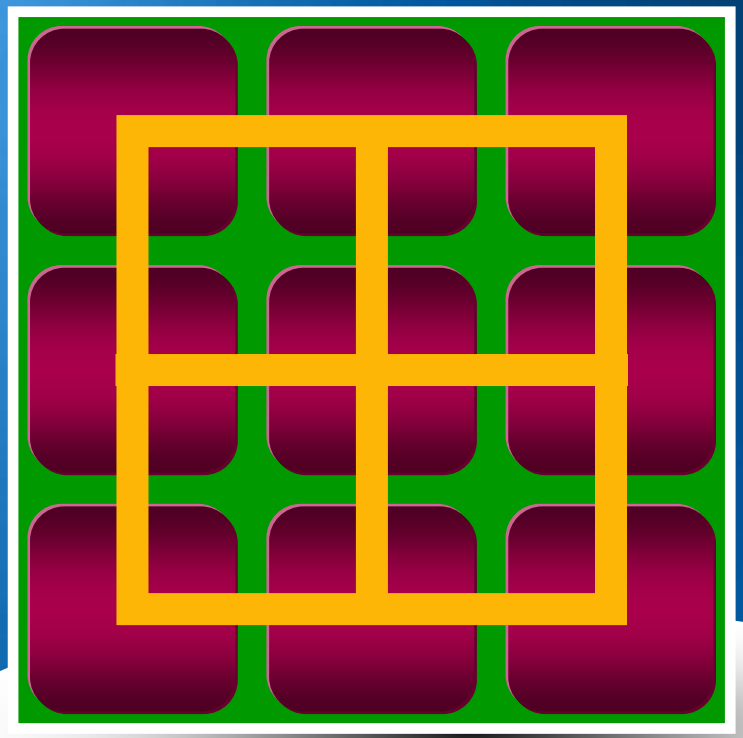
Technology Vision: Addressing the Pain Point

Interconnections for Parallelized Platforms: Cores, Memory and I/O

for Bandwidth, Capacity and Power

How Do We Connect the Cores?

Shared Bus for Future Many Core Chips?



Issues:

Slow: one core at a time <300MHz

Limited scalability

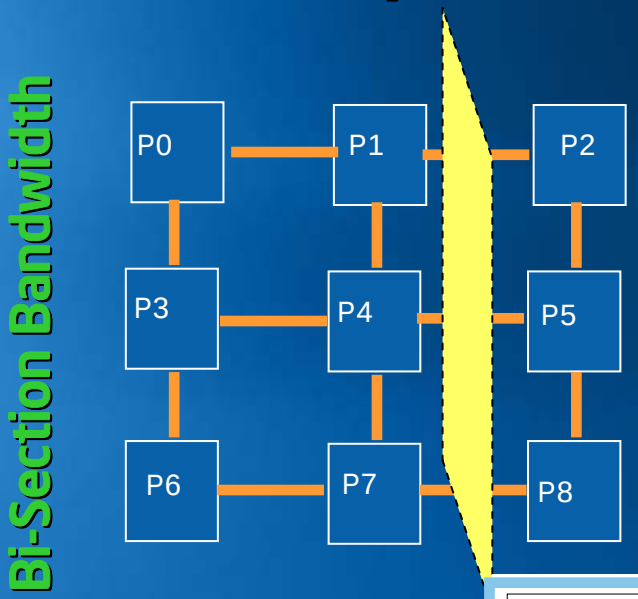
Benefits:

Power?

Simpler cache coherency

Traditional Bus is Not the Best Interconnect Option

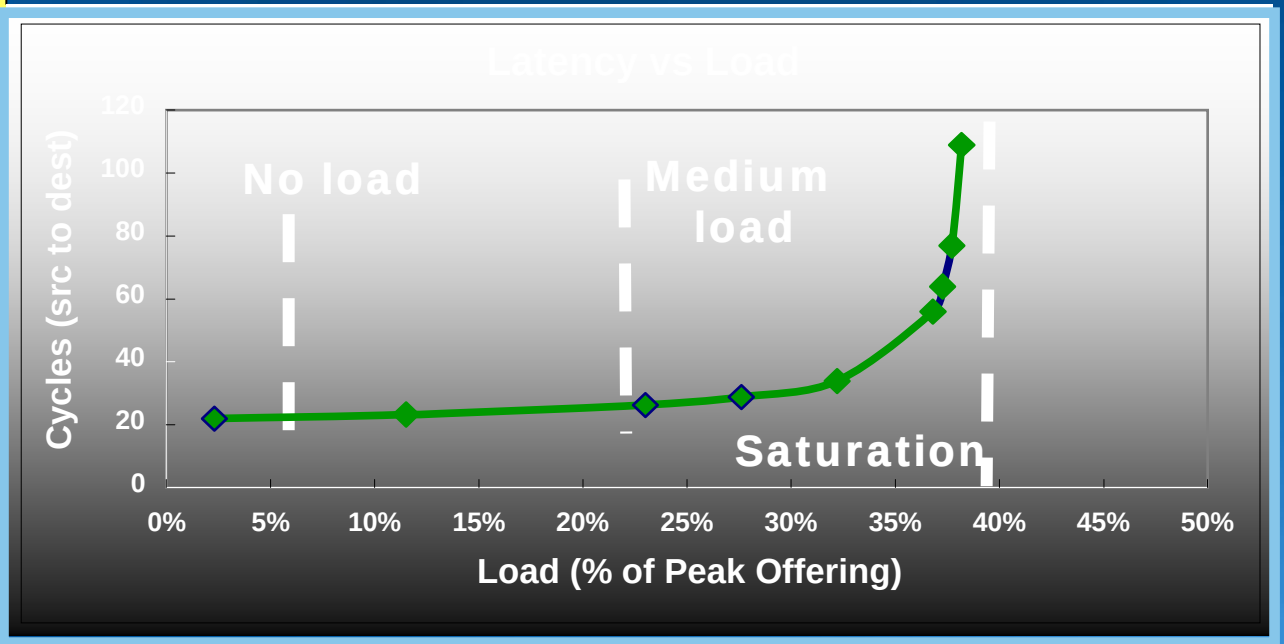
Intra-chip Interconnect Performance



Bandwidth Demand is increasing 2X per generation

- Will hit Terabytes per second soon
- Implies link bandwidth in hundreds of GB/s

- The load shown is expressed as fraction of peak (injection) capacity
- Load may also be expressed as % of network (or 2Xbisection) capacity
- Example: if network capacity (2X bisection) = 50% of peak (uniform random traffic on 4x4 mesh), then, saturation at 40% of peak implies 80% of capacity



Note: This graph is to illustrate the effect of saturation. The absolute numbers are meaningless.

Associated Costs

Power

- Interconnect fabric can consume up to 36% of chip power!
- Increasing fabric bandwidth increases power
- Need dynamic on-demand power management techniques

Area

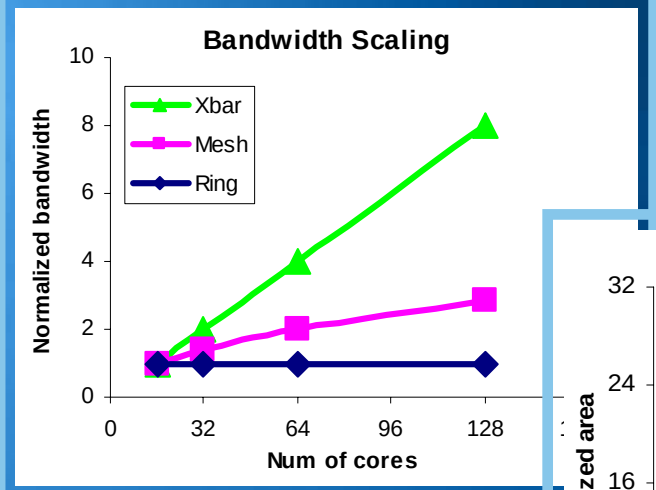
- Fabric area can be more than 20% of core area!
- Trading off compute area for higher bandwidth fabric is not desirable

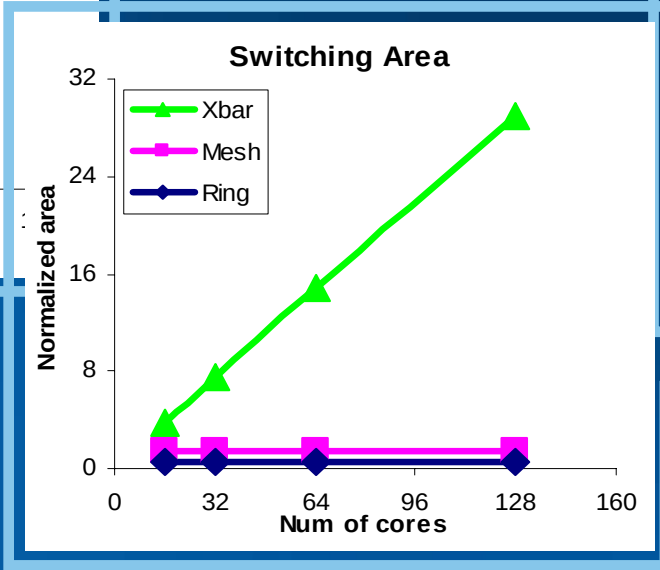
Design complexity

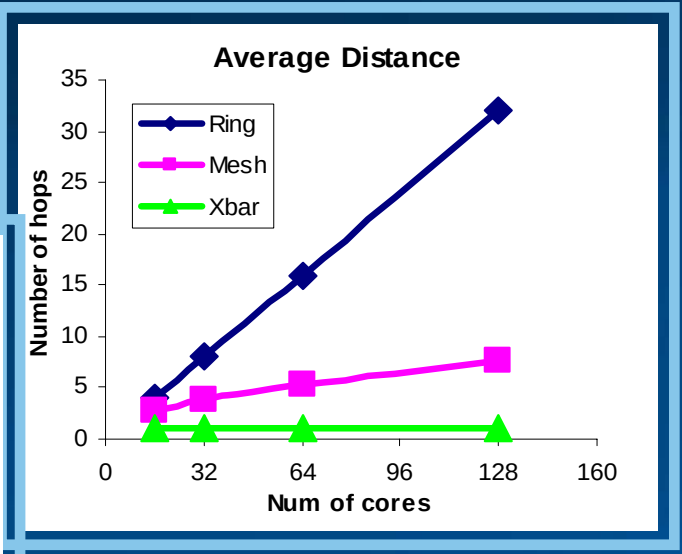
- Weighing architectural properties vs. design difficulty

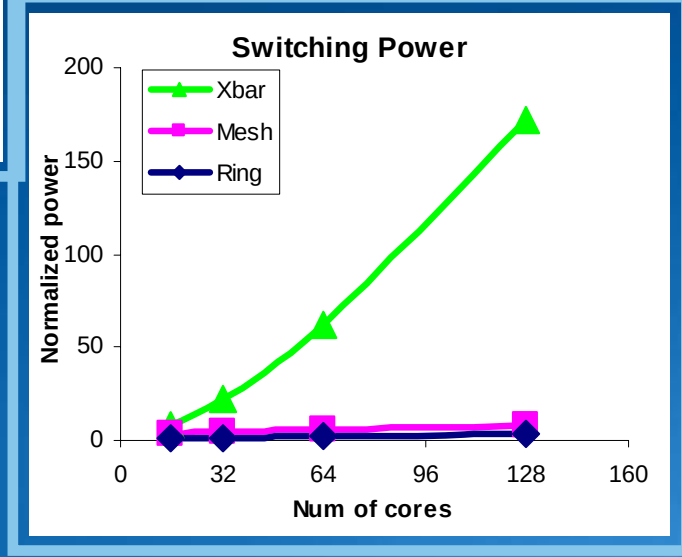
Ref: Wang et al MICRO 36, 2003

Intra-chip Interconnect Requires Topology Tradeoffs





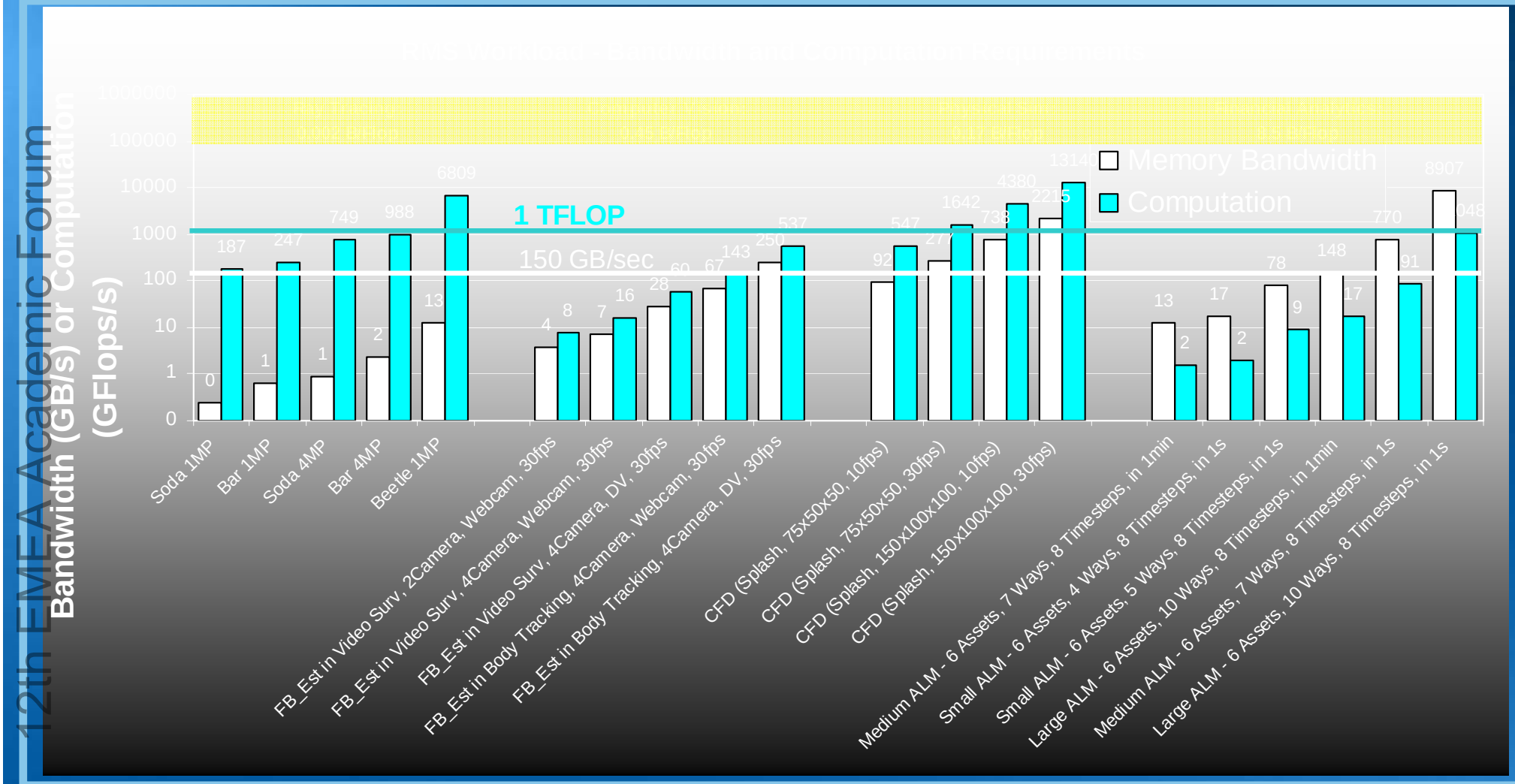




- Crossbar BW and hop count scale the best
- Crossbar area and power scale the worst
- Need techniques to improve bandwidth and latency scaling for the Mesh and Ring

Source: Intel

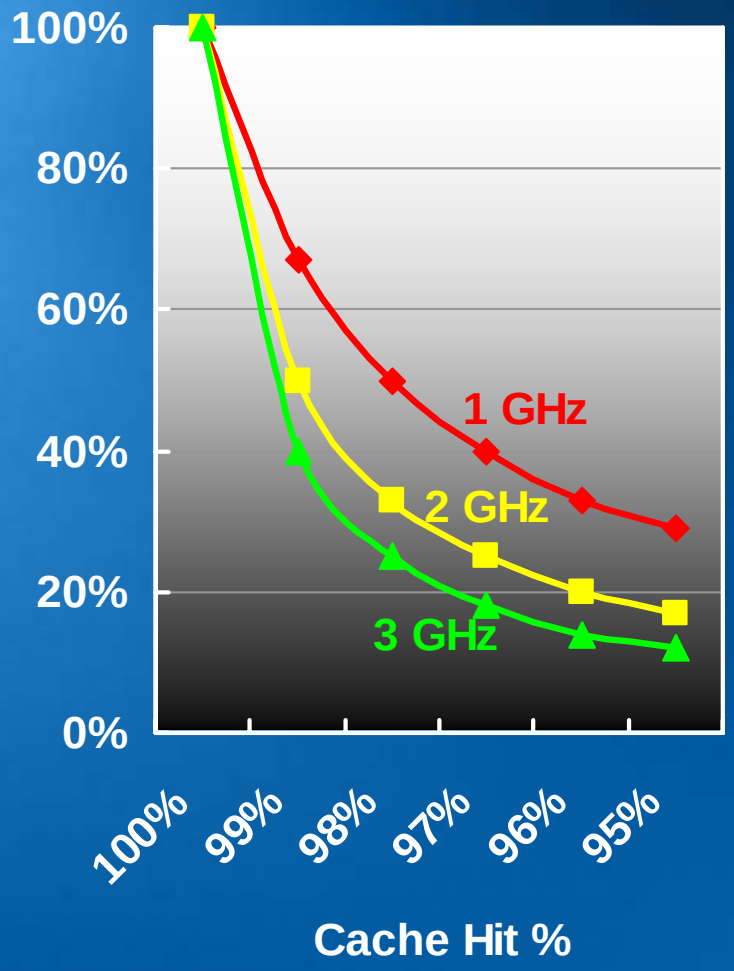
How Do We Feed the Machine?



Source: Intel Labs

Memory Bandwidth and Processor Performance Need to Keep Pace

Reduce Memory Stall Penalty through Multi-Threading



Thermals & Power Delivery designed for full HW utilization

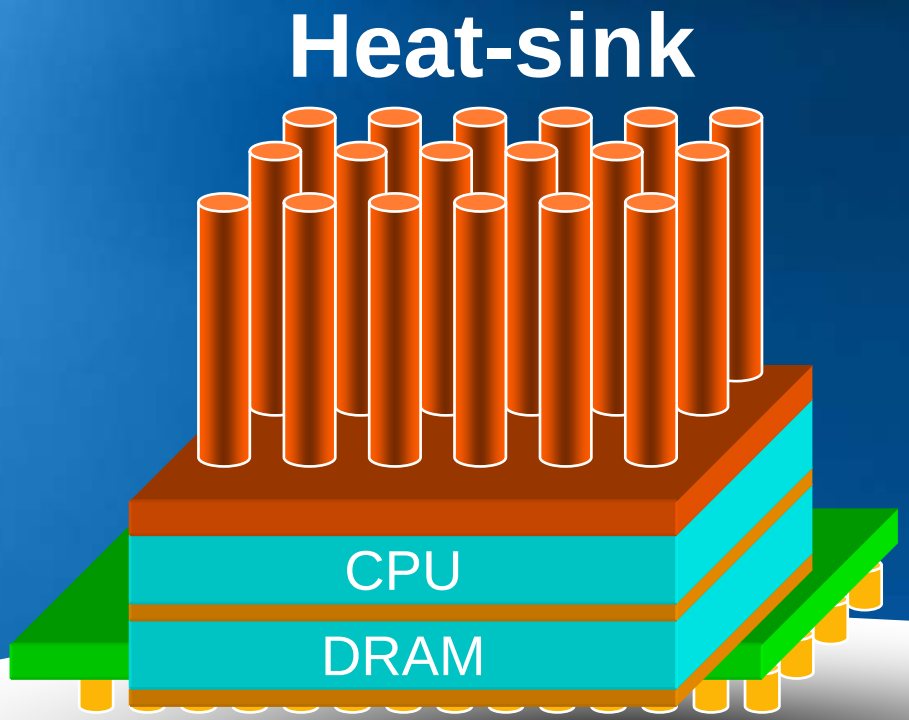
Single Thread (ST)

Multi-Threading (MT)

Multi-Threading Increases Performance and Reduce Power

Increase Memory Bandwidth through 3D Die Stacking

High Performance by Bringing Memory Closer to the Cores



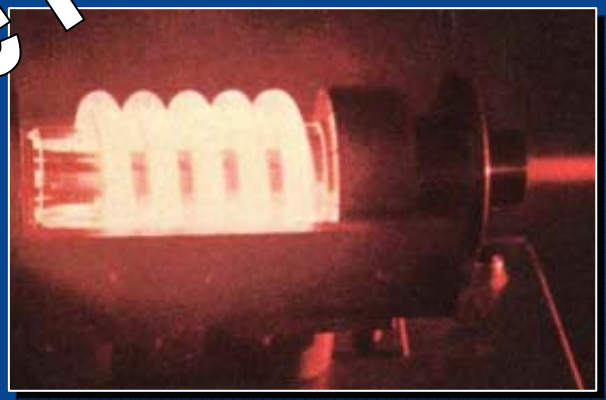
- Power and IO signals go through DRAM to CPU
- Thin DRAM die
- Through DRAM vias

DRAM, Voltage Regulators, and High Voltage I/O All on the 3D integrated die

How about Laser for Interconnect?

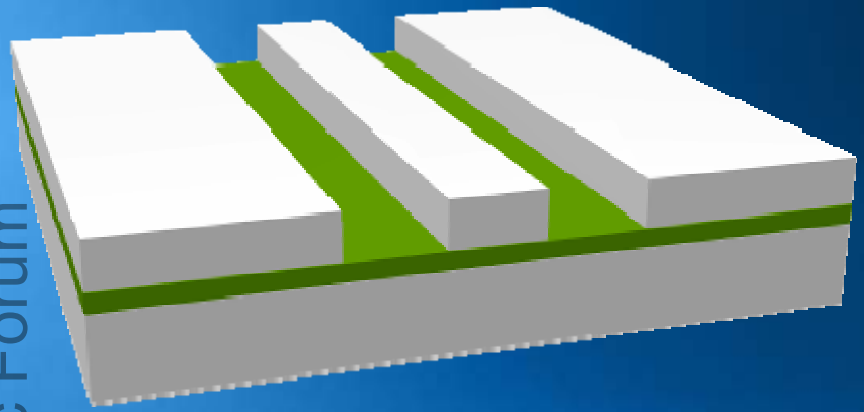
Developed by Maiman, this ruby laser used a flash bulb as an optical pump

The First Laser



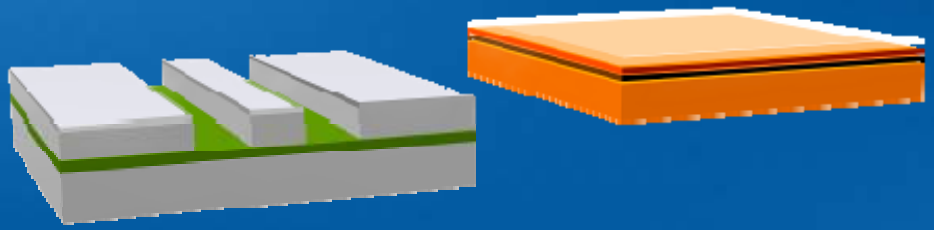
Published in *Nature*, August 6, 1960

Today's Silicon Photonics - Hybrid Laser

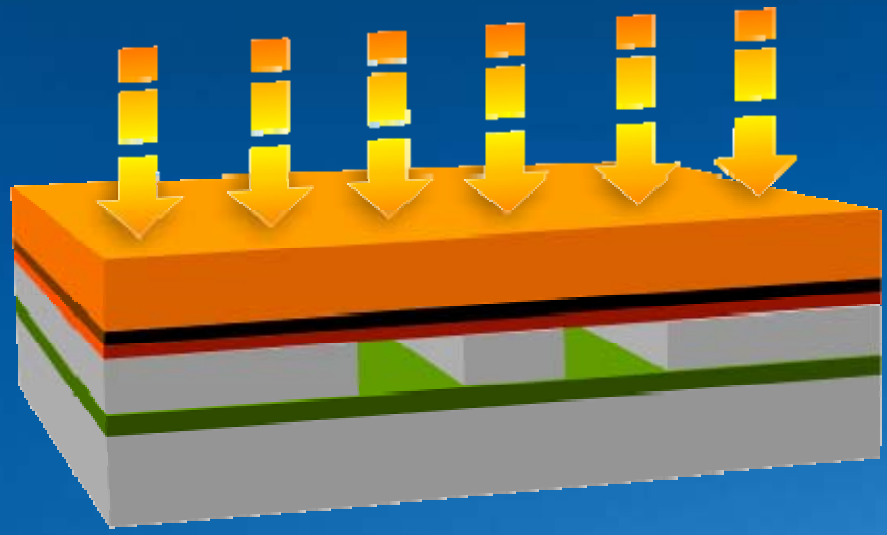


1) A waveguide is etched in silicon

2) The Indium phosphide is processed to make it a good light emitter

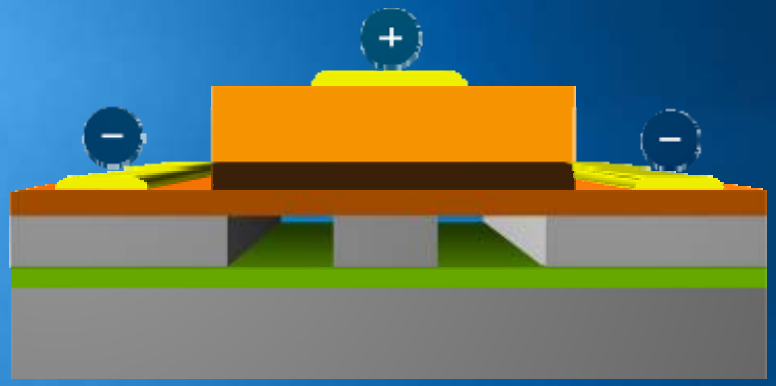


3) Both materials are exposed to the oxygen plasma to form the "glass-glue"

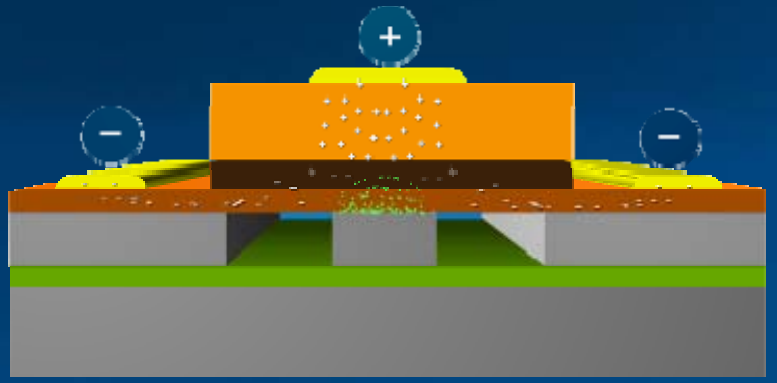


4) The two materials are bonded together under low heat

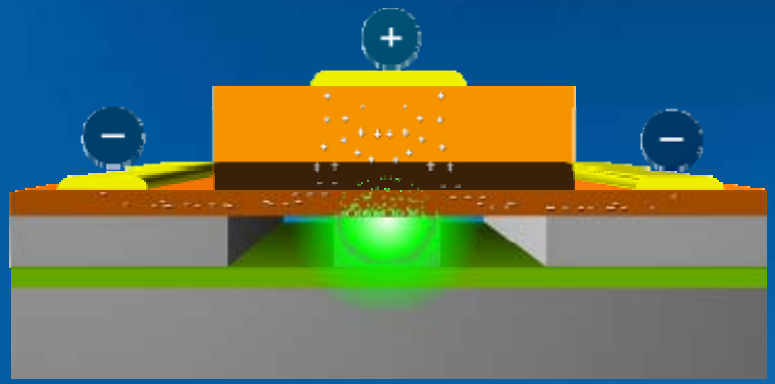
Process Animation



5) The Indium phosphide is etched and electrical contacts are added



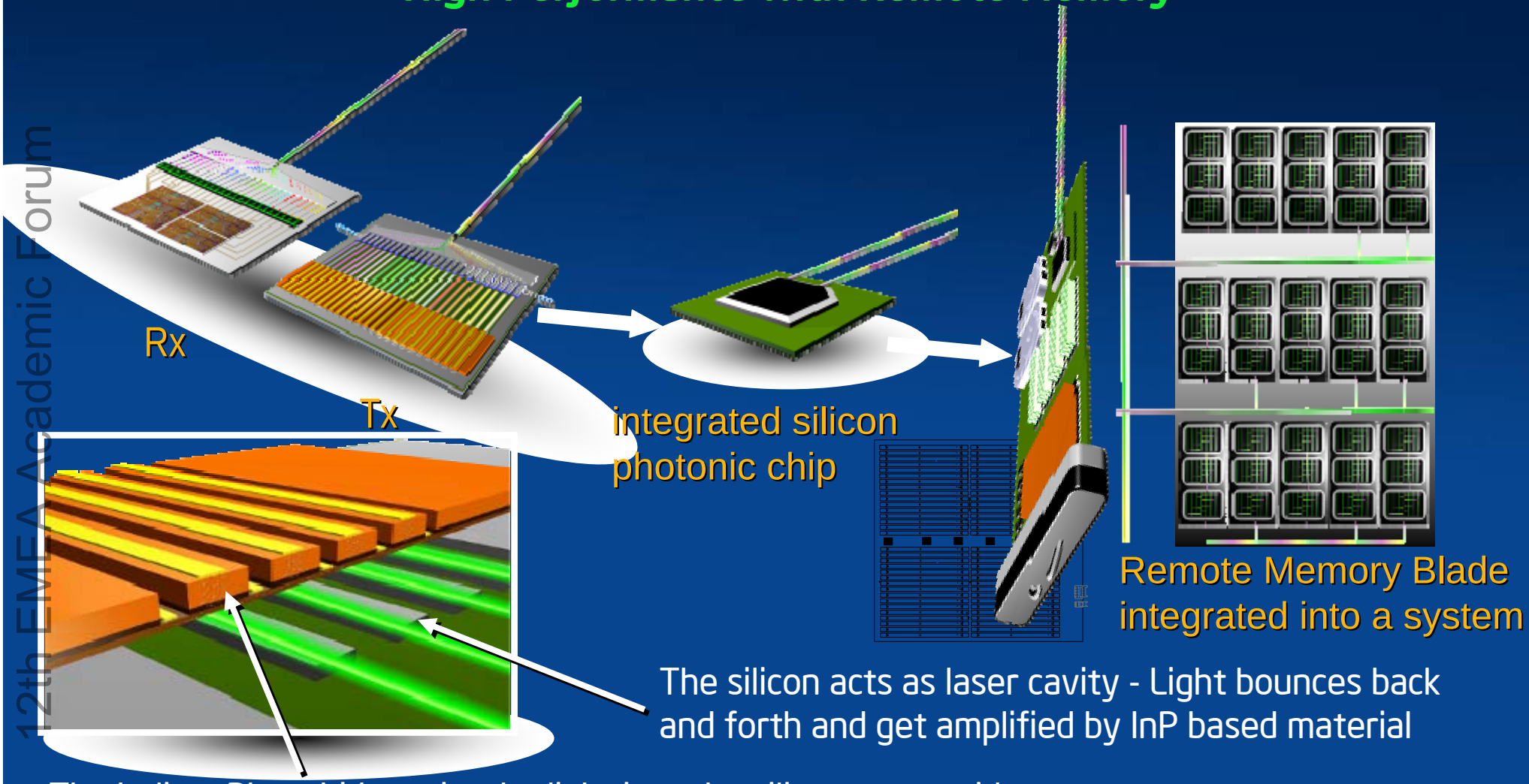
6) Photons are emitted from the Indium Phosphide when a voltage is applied



7) The light is coupled into the silicon waveguide which forms the laser cavity. Laser light emanates from the device.

Photonics For Memory Bandwidth and Capacity

High Performance with Remote Memory



The silicon acts as laser cavity - Light bounces back and forth and get amplified by InP based material

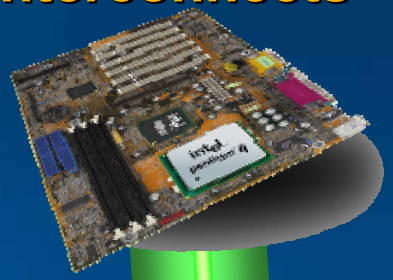
The Indium Phosphide emits the light into the silicon waveguide

Integrated Terabit Per Second (Tb/s) Optical Link on a Single Chip

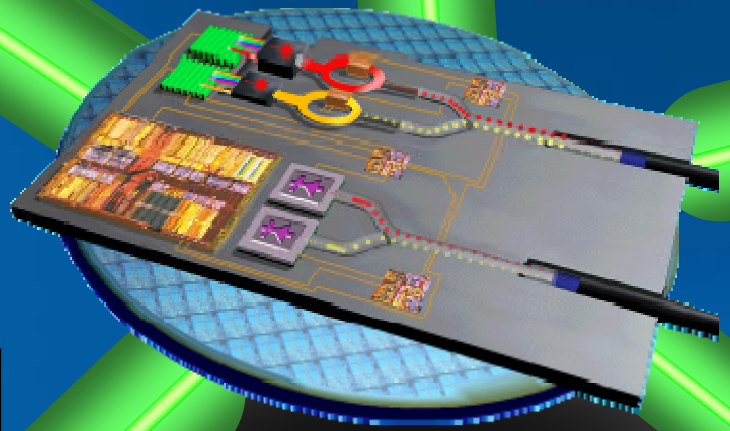
Silicon Photonics Future I/O Vision

HPC and Data Center Fabrics

Chip-to-Chip Interconnects



Backplane and Display Interconnects

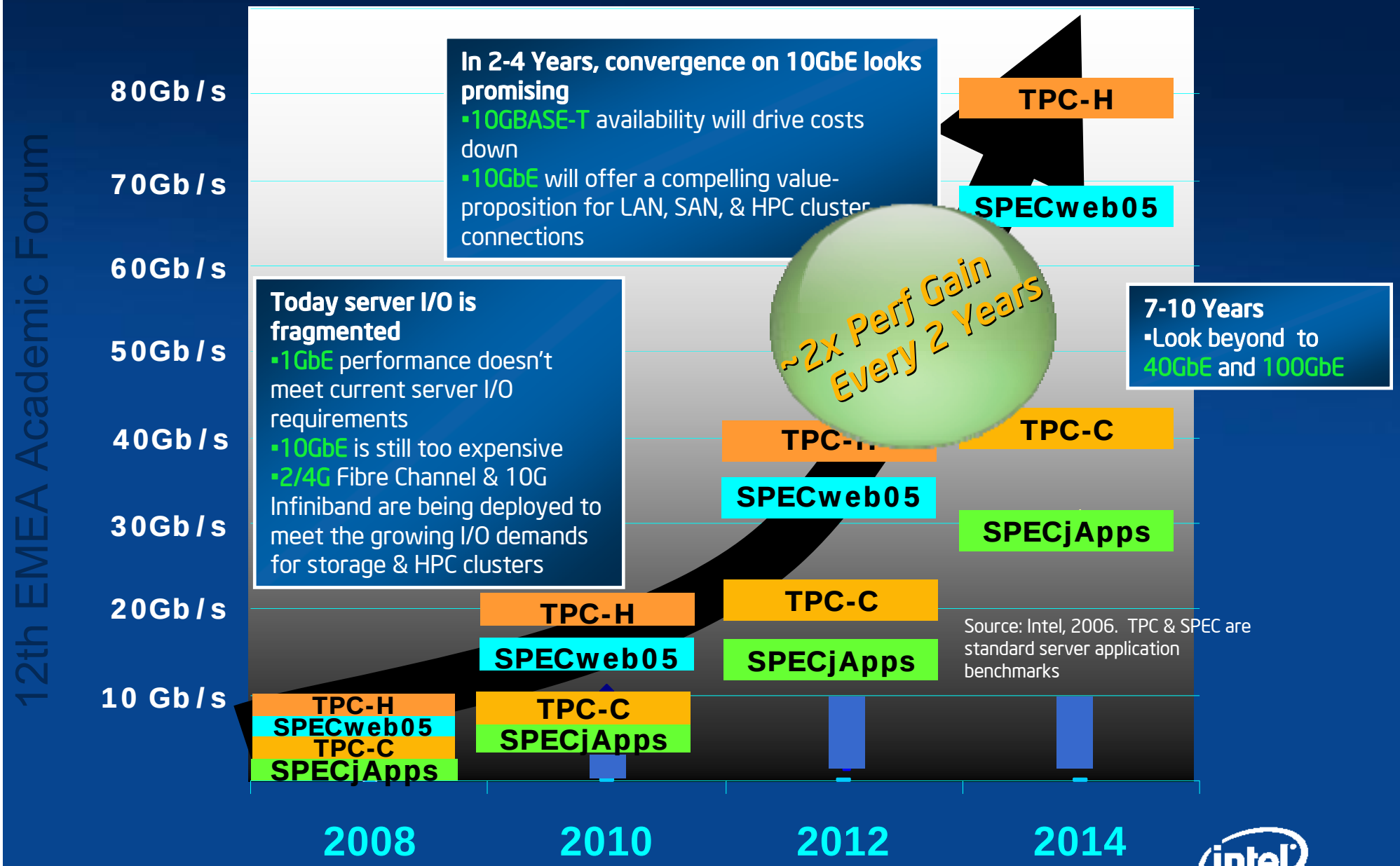


Chemical Analysis

Medical Lasers

Boost Performance with I/O

Increase Ethernet Bandwidth



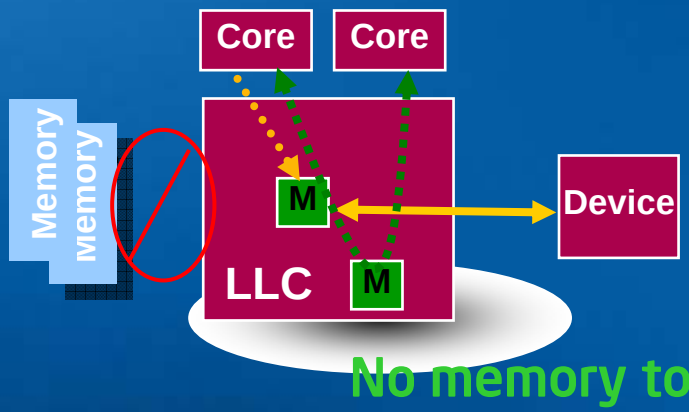
Reducing CPU and Memory Usage with Caching Hints

Problem:

- I/O related data movement & data access limits performance
  - CPU makes multiple "forced" trips to memory during the I/O processing data flow
- High CPU utilization due to "compulsory" cache misses

Impact:

- High system interconnect bandwidth consumption
- Throughput reduction per core

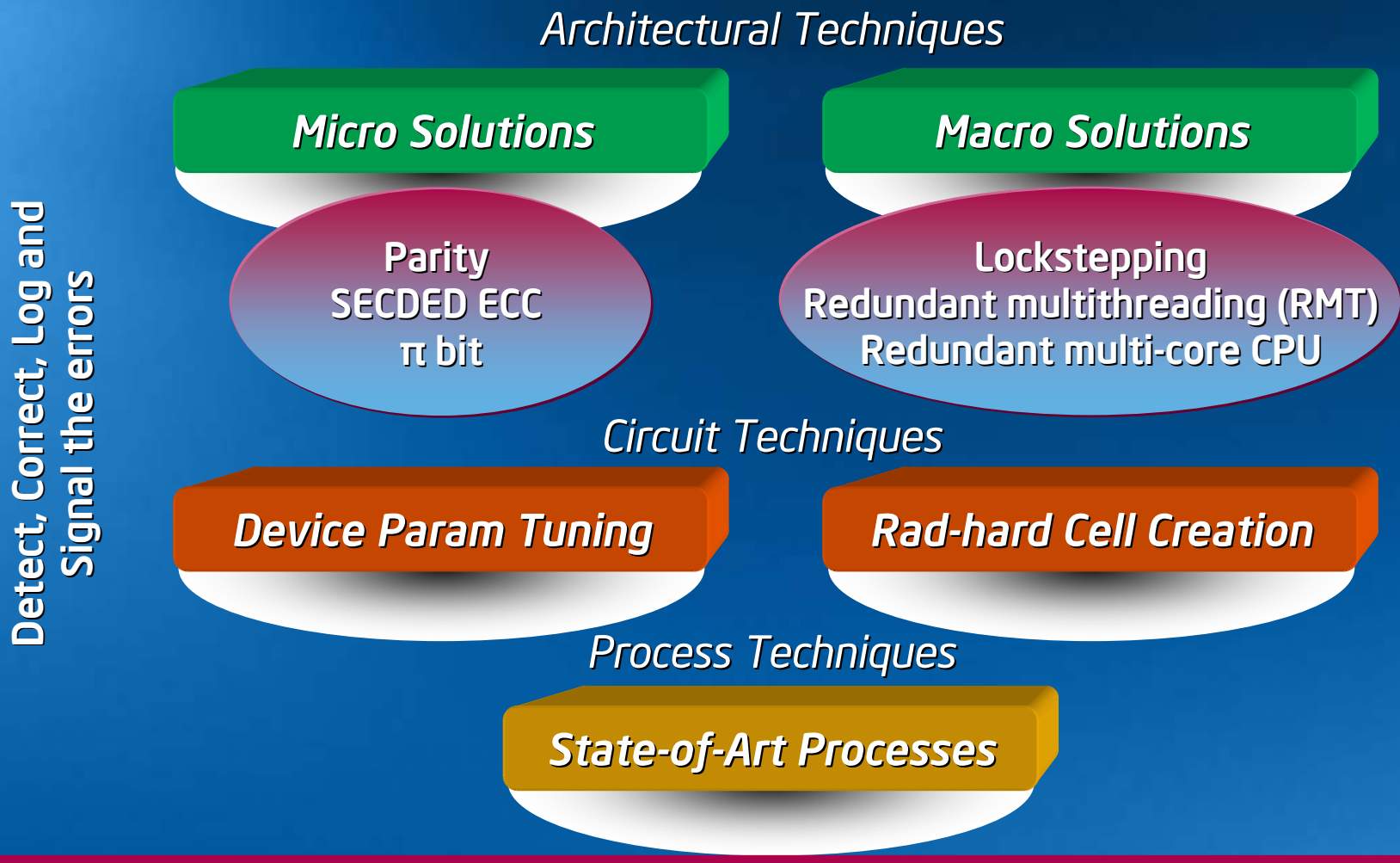


No memory touch

Time on wire at 10Gb/s @ 1518B		1230ns
Per frame processing time at 2000 inst/frame, CPI =1, 3.5Ghz core		575ns
Measured WDC (Bensley) Copy Performance (1460B)		
Source	Destination	
L2	L2	183ns
Memory	L2	730ns
Memory	Memory	1460ns
Single L2 cache miss		107ns (idle)

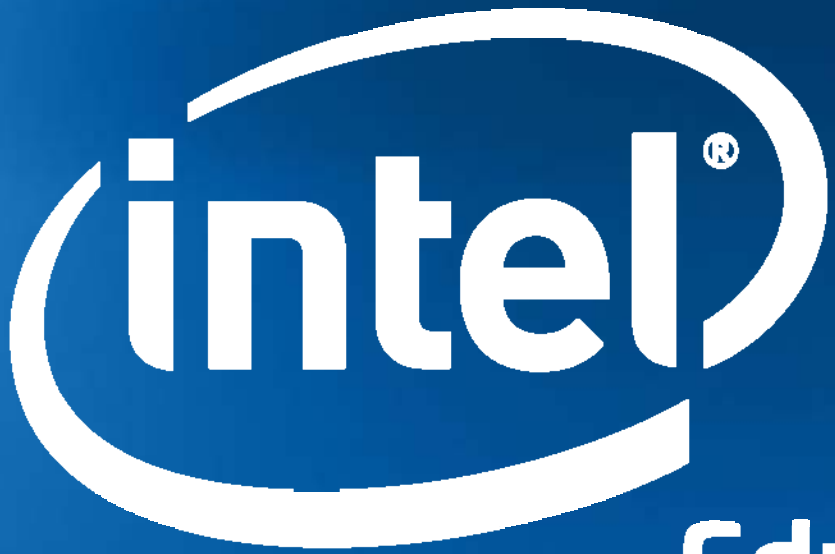
Cannot afford to copy from memory

Reliable Systems With Unreliable Components



Future Reliability Faces Big Challenges – Solution is in the Platform

Questions?



Education