The OptIPuter – Toward a Terabit LAN

Talk at
The ON*VECTOR Terabit LAN Workshop
Hosted by Calit2
University of California, San Diego
January 29 , 2005

Dr. Larry Smarr
Director, California Institute for Telecommunications and Information Technology
Harry E. Gruber Professor,
Dept. of Computer Science and Engineering
Jacobs School of Engineering, UCSD

The OptIPuter Philosophy

Bandwidth is getting cheaper faster than storage.
Storage is getting cheaper faster than computing.
Exponentials are crossing.

"A global economy designed
to waste transistors, power, and silicon area
-and conserve bandwidth above all-
is breaking apart and reorganizing itself
to waste bandwidth
and conserve power, silicon area, and transistors."
George Gilder Telecosm (2000)

OptIPuter

it²

Terabits to the Desktop by 2010

- Simplified User View
- Terabit Fiber Connection To The Desktop
- Integrated Photonics And Electronics
- Single Fiber Dense-WDM
- Packets And Flows
- Encryption
- …

“ Ethernet ”	
1990	10 Mb
1995	100 Mb
1998	1 Gb
2002	10 Gb
2006	100 Gb
2008	1 Tb
2010	10 Tb

OptIPuter

Source: Steven Squires, Chief Scientist HP

it²

The Future of Devices

- **Trends**
  - Processor < Memory Channel
  - Electron < Photon Channel
  - Flux < High Flux
- **Limits**
  - Processor Memory Interface
  - Photon Electron Interface
  - Photon Processing
- **Futures**
  - Integrated Processor Memory Flux
  - Integrated Photonic Electronics
  - Integrated Photonic Logic

Source: Stephen Squires, Hewlett-Packard

OptIPuter is Prototyping The PC of 2010

- Terabits to the Desktop…
- 100 Megapixels Display
  - 55-Panel
- 1/3 Terabit/sec I/O
  - 30 x 10GE interfaces
  - Linked to OptIPuter
- 1/4 TeraFLOP
  - Driven by 30 Node Cluster of 64 bit Dual Opterons
- 1/8 TB RAM
- 60 TB Disk



OptIPuter

Source: Jason Leigh, Tom DeFanti, EVL@UIC
OptIPuter Co-PIs

it²

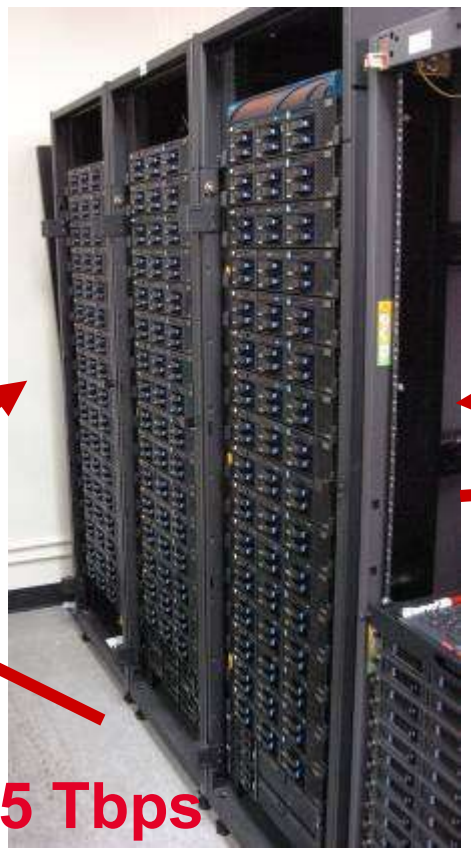
UCSD Campus LambdaStore Architecture —If we Used 10Gb NICs—A Terabit LAN!

SIO Ocean Supercomputer



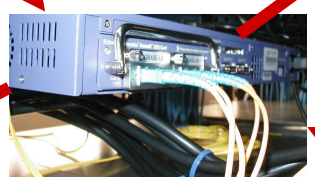
1.3 Tbps

IBM Storage Cluster



Streaming Microscope







Extreme switch with 2 ten gigabit uplinks

0.5 Tbps





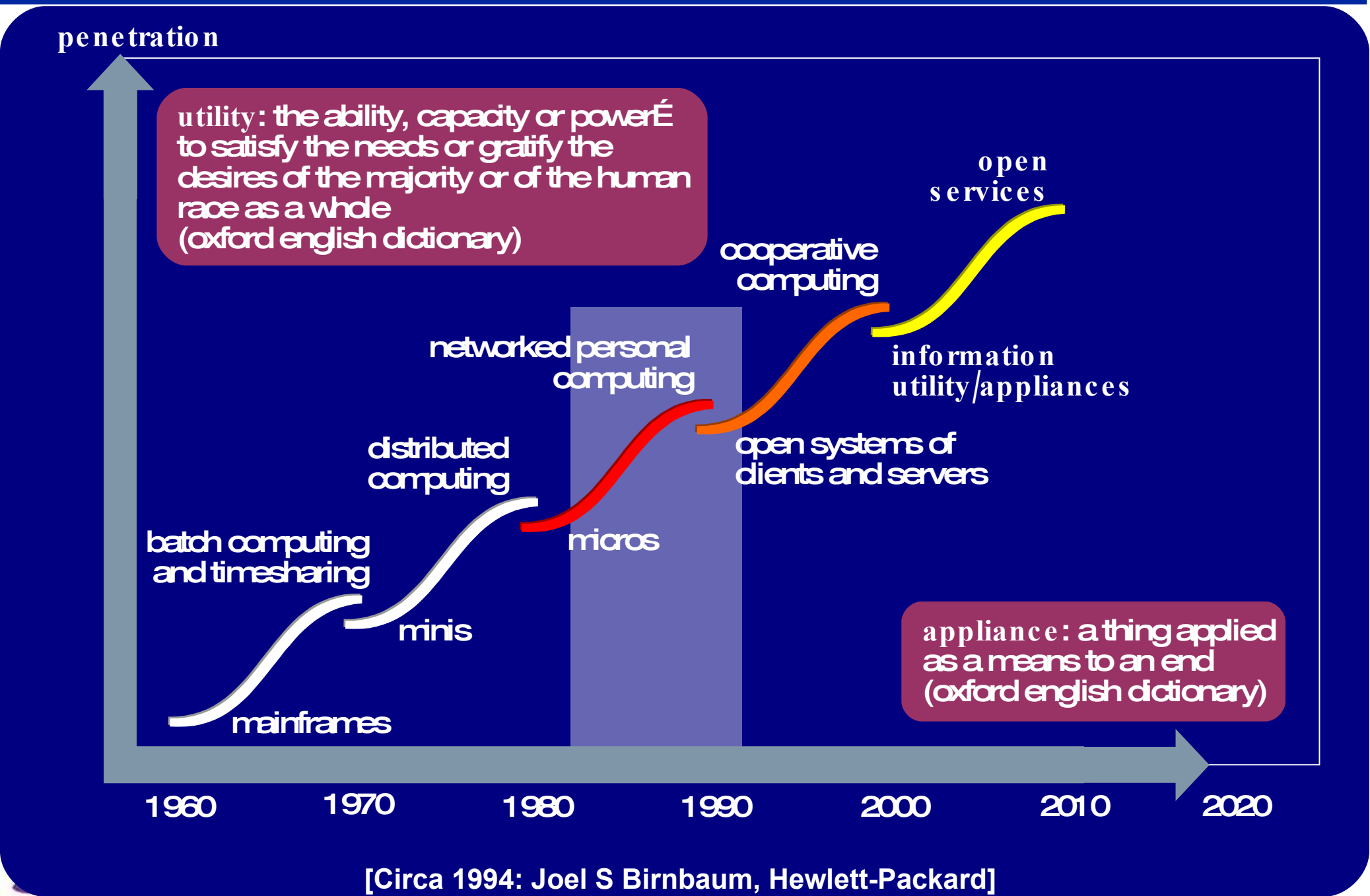


0.1 Tbps

OptIPuter

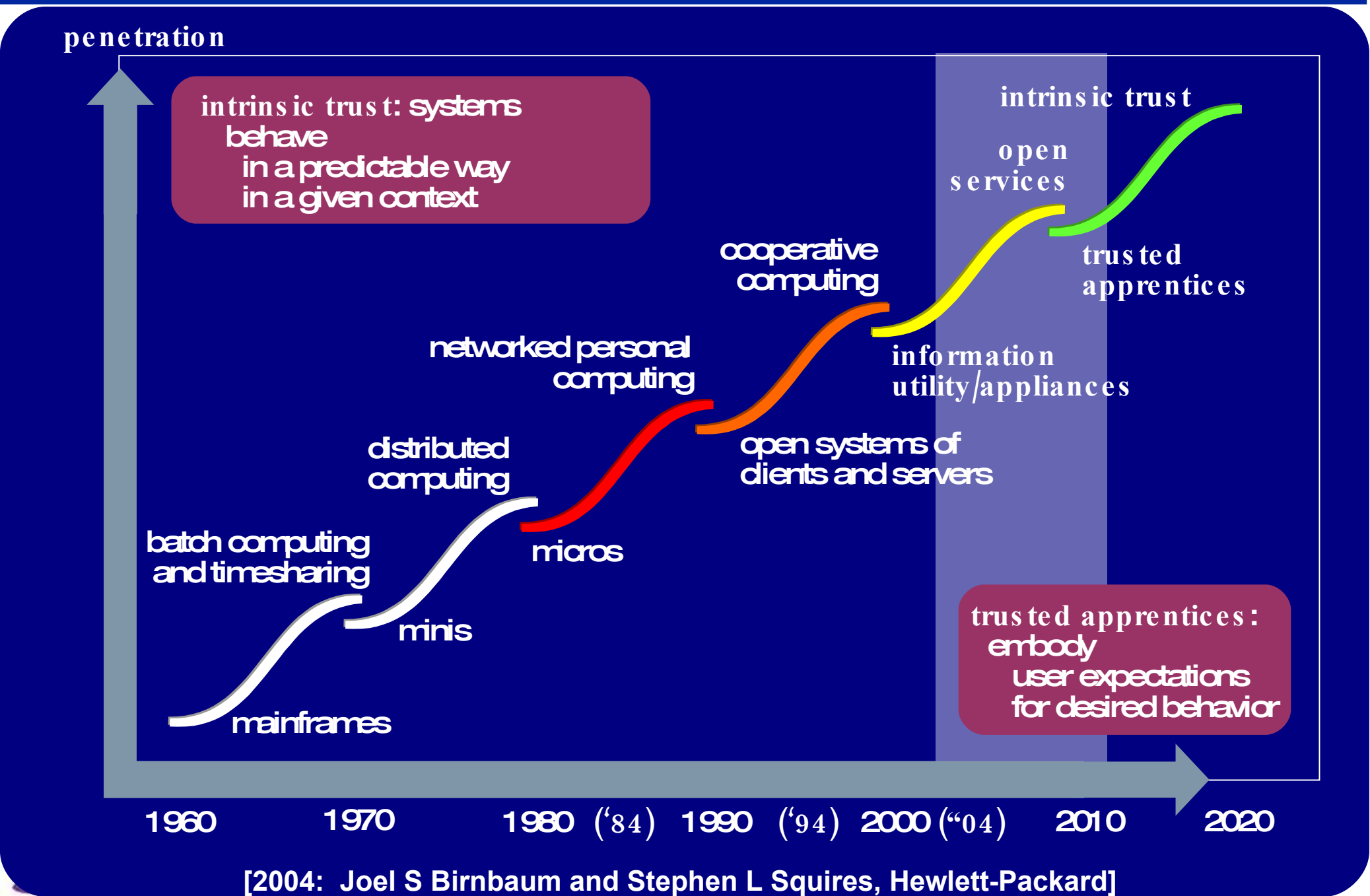
it²

Toward Pervasive Information Systems



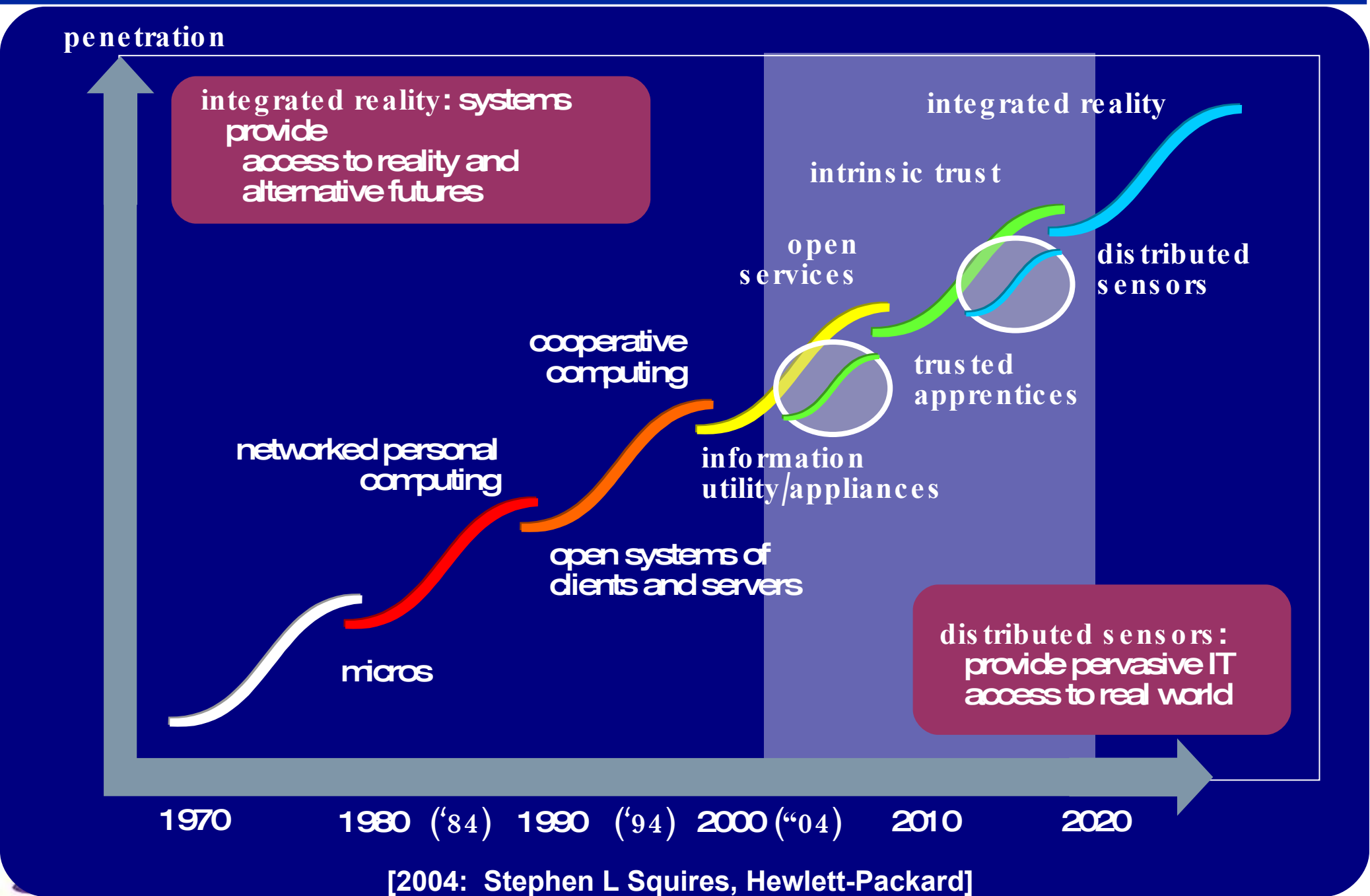
[Circa 1994: Joel S Birnbaum, Hewlett-Packard]

Toward Critical And Pervasive Trusted Information Systems



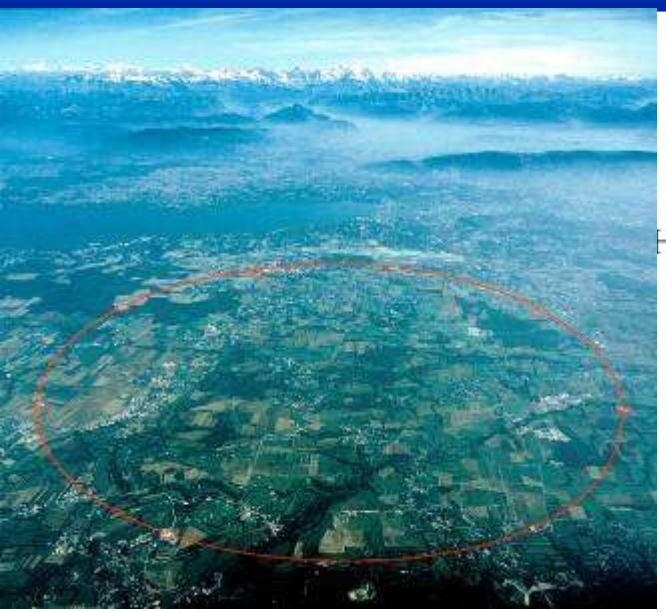
[2004: Joel S Birnbaum and Stephen L Squires, Hewlett-Packard]

Beyond *Critical and Pervasive* Toward *Integrated Reality*



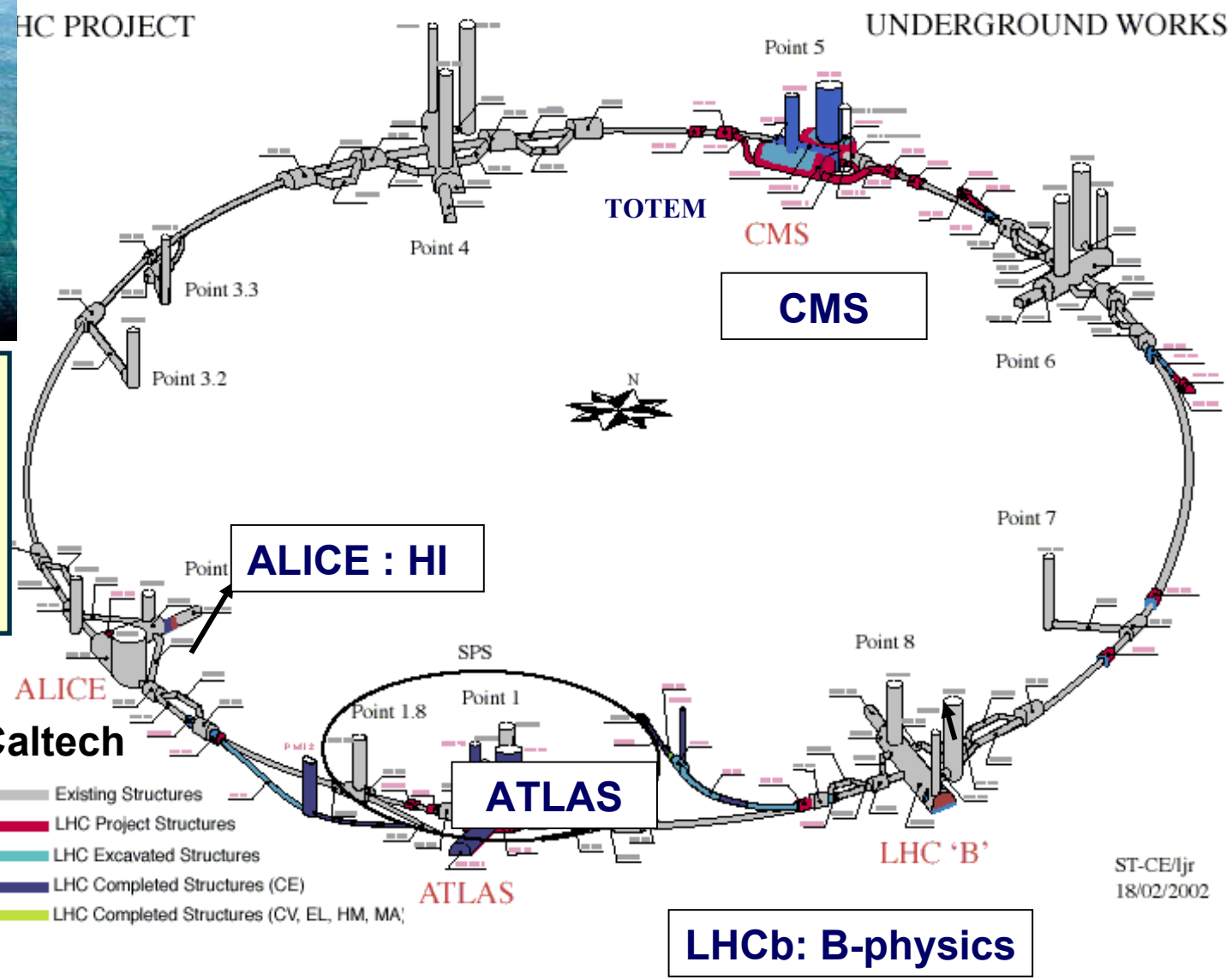
[2004: Stephen L Squires, Hewlett-Packard]

Large Hadron Collider (LHC)
e-Science Driving Global Cyberinfrastructure



- pp $\sqrt{s}$ =14 TeV L=10^{34} cm^{-2} s^{-1}
- 27 km Tunnel in Switzerland & France

First Beams:
April 2007
Physics Runs:
from Summer 2007



Source: Harvey Newman, Caltech

OptIPuter

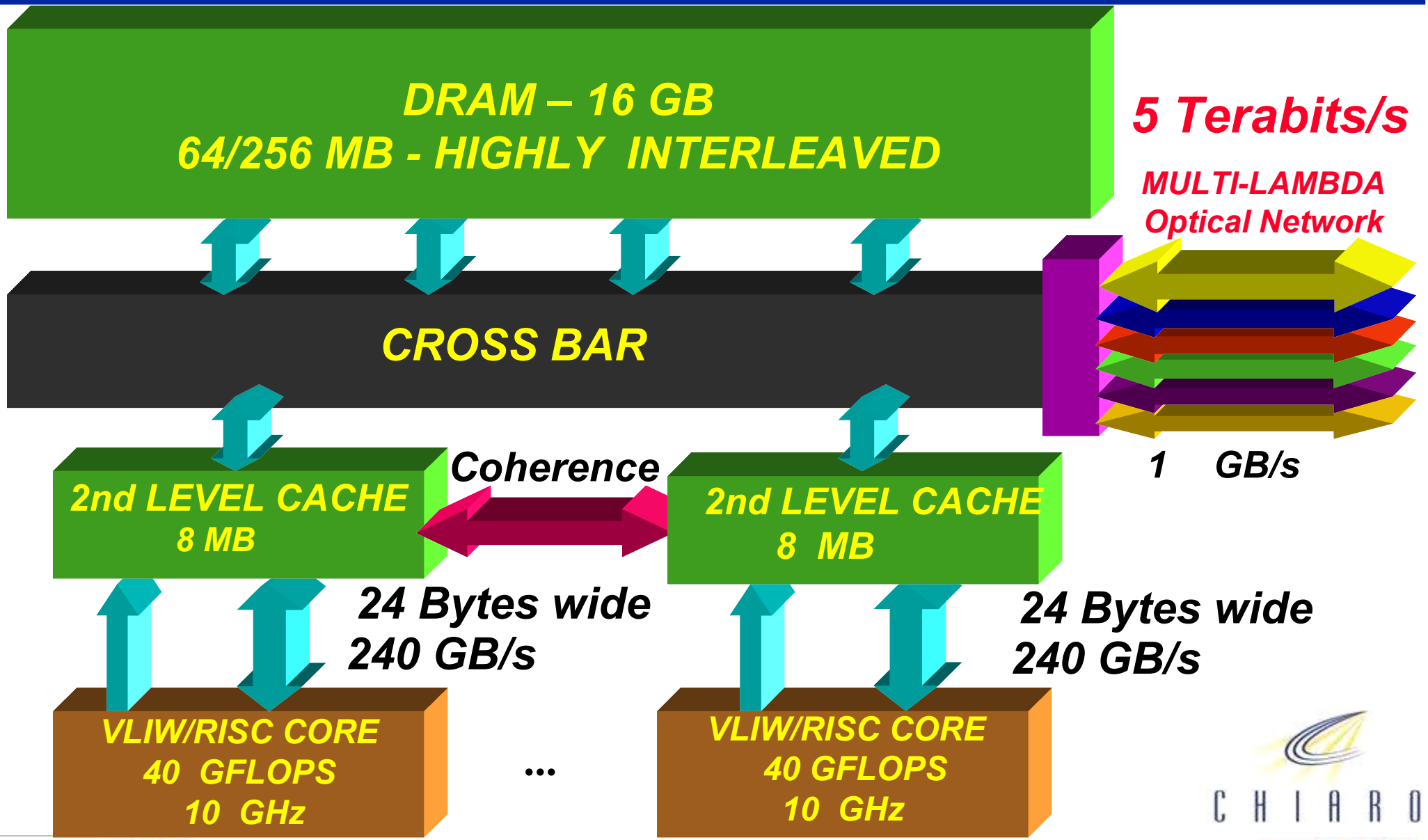
High Energy and Nuclear Physics
A Terabit/s WAN by 2010

Year	Production	Experimental	Remarks
2001	0.155	0.622-2.5	SONET/SDH
2002	0.622	2.5	SONET/SDH DWDM; GigE Integ.
2003	2.5	10	DWDM; 1 + 10 GigE Integration
2005	10	2-4 X 10	λ Switch; λ Provisioning
2007	2-4 X 10	~10 X 10; 40 Gbps	1st Gen. λ Grids
2009	~10 X 10 or 1-2 X 40	~5 X 40 or ~20-50 X 10	40 Gbps λ Switching
2011	~5 X 40 or ~20 X 10	~25 X 40 or ~100 X 10	2nd Gen λ Grids Terabit Networks
2013	~Terabit	~MultiTbps	~Fill One Fiber

Source: Harvey Newman, Caltech

Continuing the Trend: ~1000 Times Bandwidth Growth Per Decade; We are Rapidly Learning to Use Multi-Gbps Networks Dynamically

OptIPuter Inspiration--Node of a 2009 PetaFLOPS Supercomputer



CHIARO

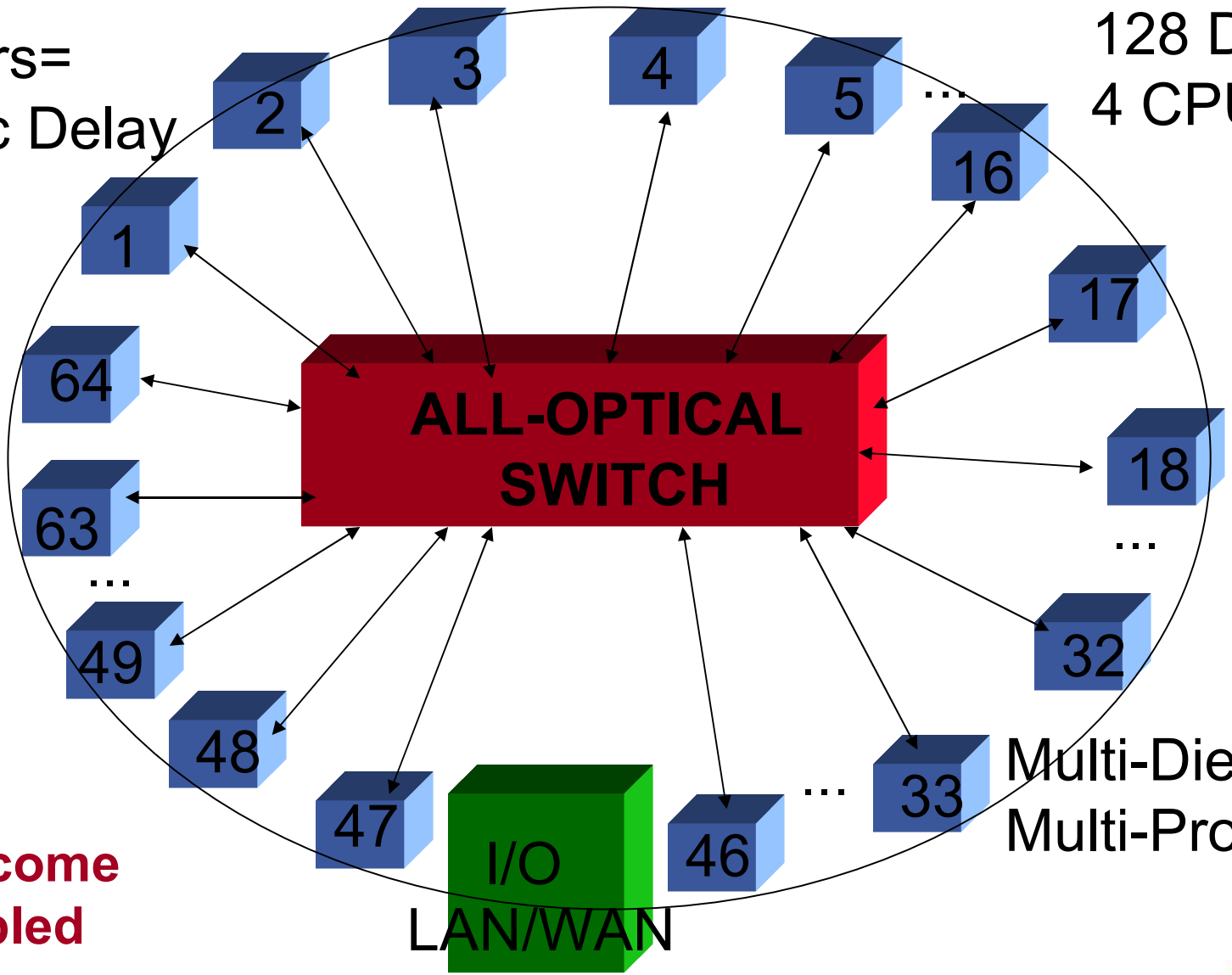
OptIPuter

Updated From Steve Wallach, Supercomputing 2000 Keynote

Global Architecture of a 2009 COTS PetaFLOPS System

1 meters=
50 nanosec Delay

128 Die/Box
4 CPU/Die



Multi-Die
Multi-Processor

CHIARO

Systems Become GRID Enabled

OptIPuter

Source: Steve Wallach, Supercomputing 2000 Keynote

it²

Convergence of Networking Fabrics

- **Today's Computer Room**
  - Router For External Communications (WAN)
  - Ethernet Switch For Internal Networking (LAN)
  - Fibre Channel For Internal Networked Storage (SAN)

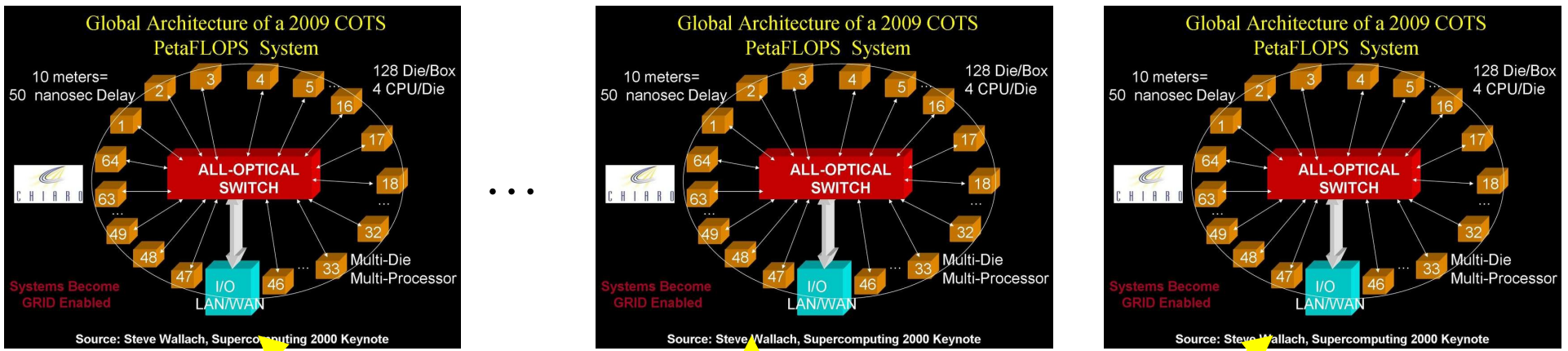
- **Tomorrow's Grid Room**
  - A Unified Architecture Of LAN/WAN/SAN Switching
  - More Cost Effective
    - One Network Element vs. Many
  - One Sphere of Scalability
  - ALL Resources are GRID Enabled
    - Layer 3 Switching and Addressing Throughout

OptIPuter

Source: Steve Wallach, Chiaro Networks

it²

Grid ExaFlop in 2009



One Ring to rule them all. One Ring to find them.
One Ring to bring them all and in the Darkness bind them.

World-Wide SONET RING

Source: Steve Wallach, Chiaro Networks