“End-to-end Optical Fiber Cyberinfrastructure for Data-Intensive Research: Implications for Your Campus”

Featured Speaker EDUCAUSE 2010

Anaheim Convention Center

Anaheim, CA

October 13, 2010

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

Follow me on Twitter: lsmarr

Abstract

Most campuses today only provide shared Internet connectivity to the end user's labs, in spite of the existence of national-scale optical fiber networking, capable of multiple wavelengths of 10Gbps dedicated bandwidth. This "last mile gap" requires campus CIOs to plan for installing a more ubiquitous fiber infrastructure on campus and rethinking the centralization of storage and computing. Such a set of high-bandwidth campus "on-ramps" will also be required if remote clouds are to be useful for storing gigabyte to terabyte size data objects, which are routinely produced by modern scientific instruments. I will review experiments at UCSD which give a preview of how to build a 21st century data-intensive research campus.

The Data Intensive Era Requires High Performance Cyberinfrastructure

- **Growth of Digital Data is Exponential**
  - **"Data Tsunami"**
- **Driven by Advances in Digital Detectors, Networking, and Storage Technologies**
- **Shared Internet Optimized for Megabyte-Size Objects**
- **Need New Cyberinfrastructure for Gigabyte Objects**
- **Making Sense of it All is the New Imperative**
  - **Data Analysis Workflows**
  - **Data Mining**
  - **Visual Analytics**
  - **Multiple-database Queries**
  - **Data-driven Applications**



SDSC

Source: SDSC

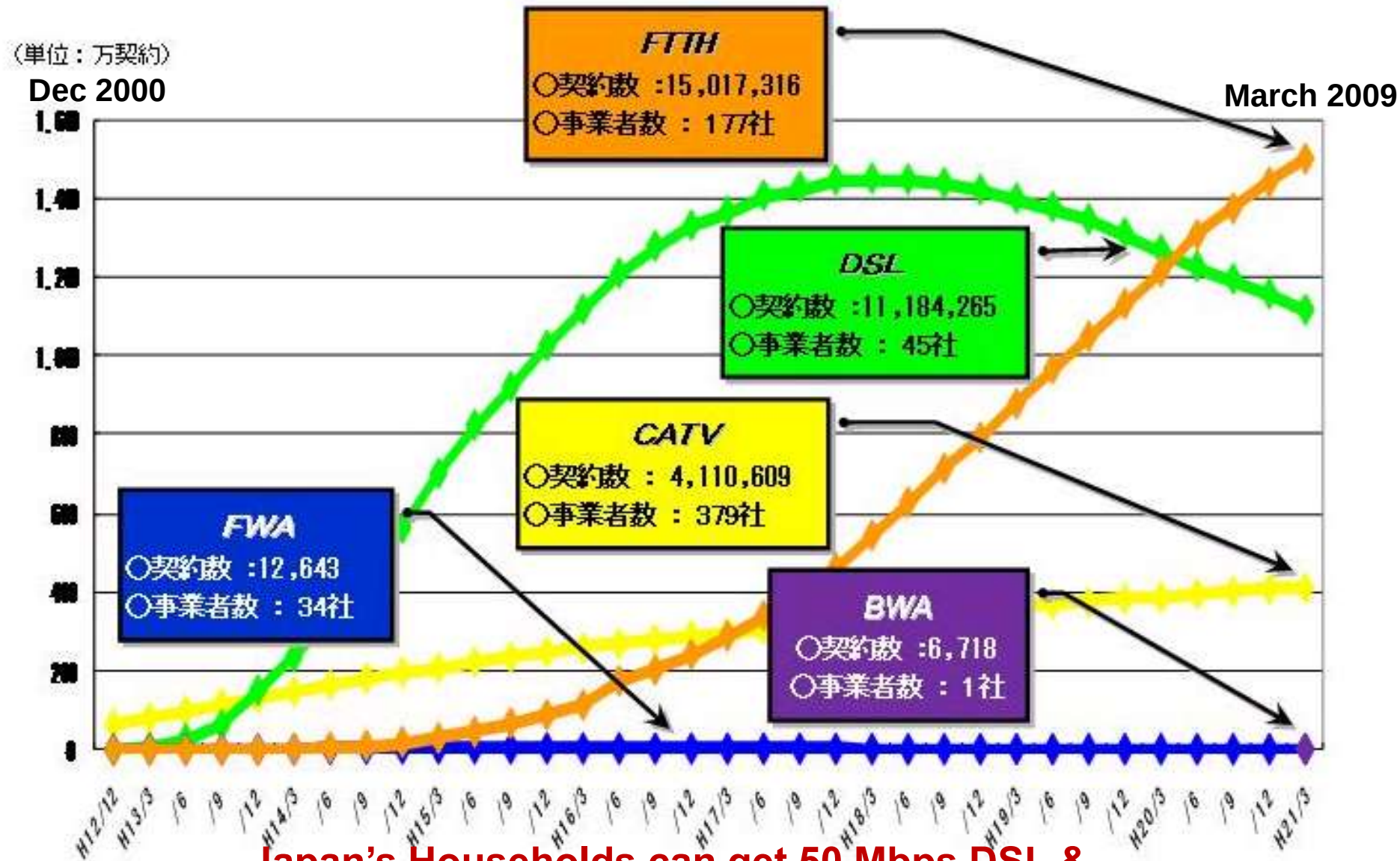
it²

What Are the Components of High Performance Cyberinfrastructure?

- High Performance Optical Networks
- Data-Intensive Visualization and Analysis
- End-to-End Wide Area CI
- Data-Intensive Research CI

High Performance Optical Networks

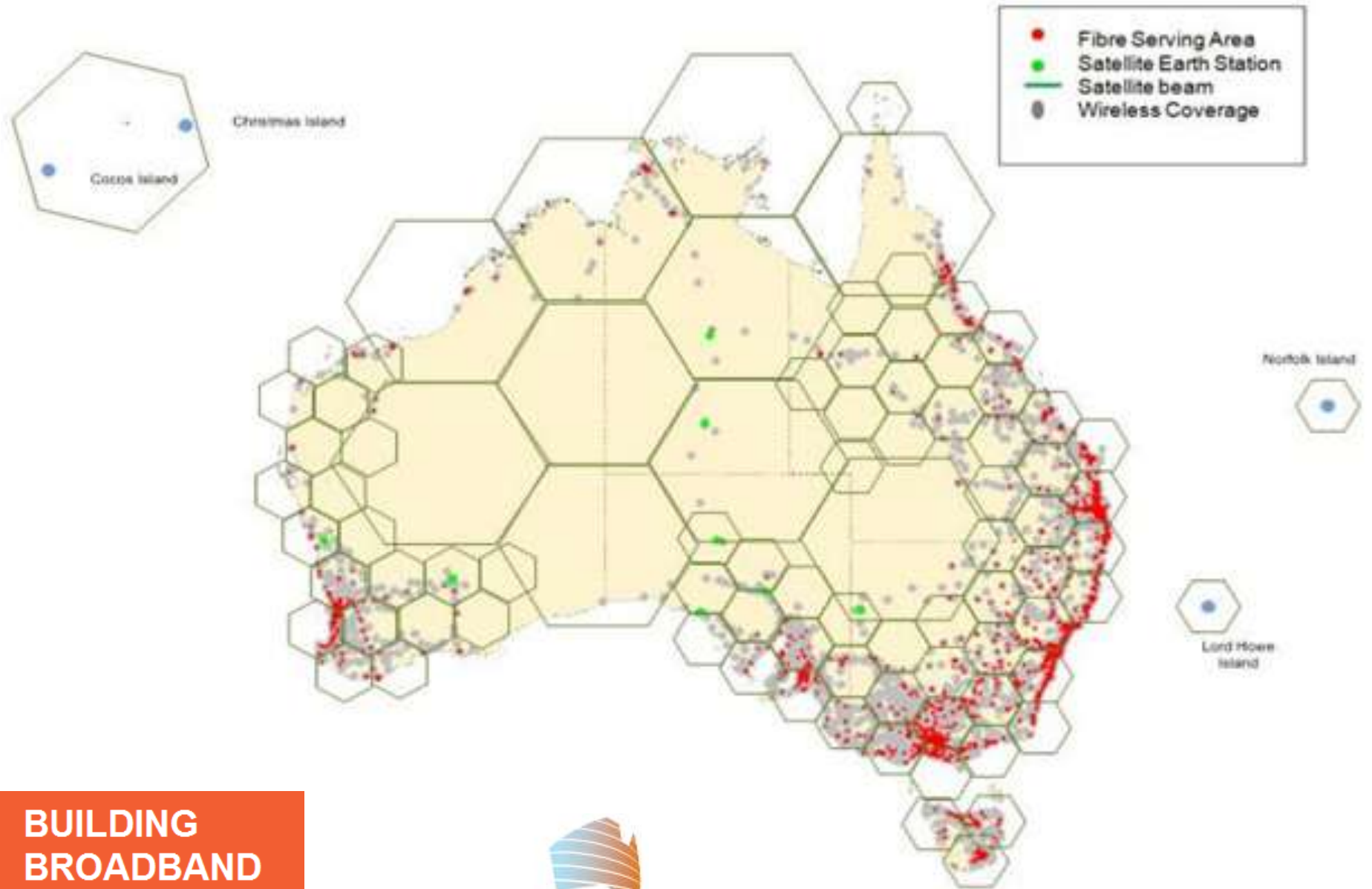
In Japan, FTTH Has Become the Dominant Broadband--Subscribers to "Slow" 40 Mbps ADSL Are Decreasing!



Japan's Households can get 50 Mbps DSL &
100Mbps to1Gbps FTTH Services with Competitive Prices

Source: Japan's Ministry of Internal Affairs and Communications
http://tilgin.wordpress.com/2009/12/17/japan-the-land-of-fiber/

Australia—The Broadband Nation:
Universal Coverage with Fiber, Wireless, Satellite



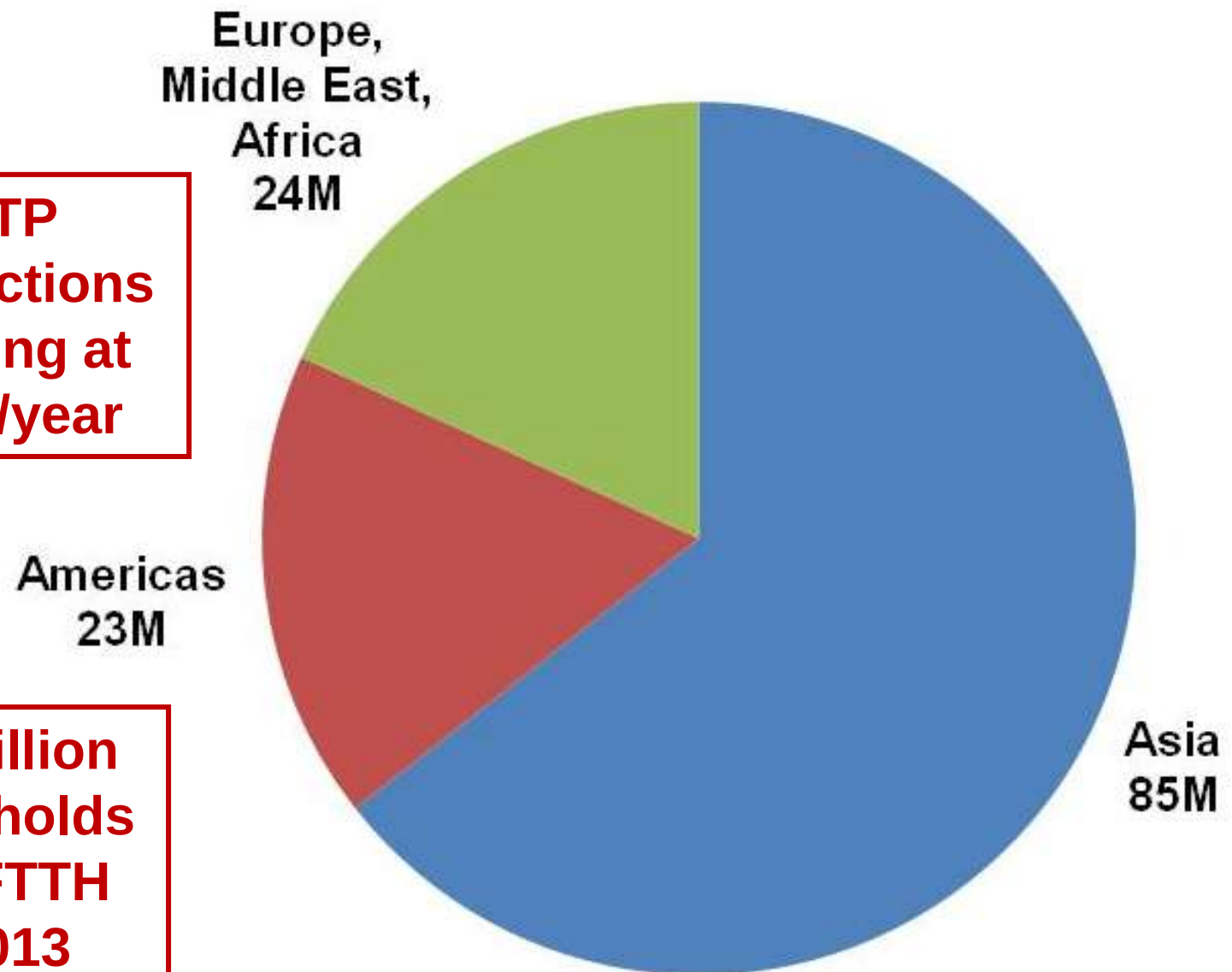
BUILDING BROADBAND FOR ALL AUSTRALIANS

NBNCo Limited

www.nbnco.com.au

Globally Fiber to the Premise is Growing Rapidly, Mostly in Asia

FTTP Connections Growing at ~30%/year

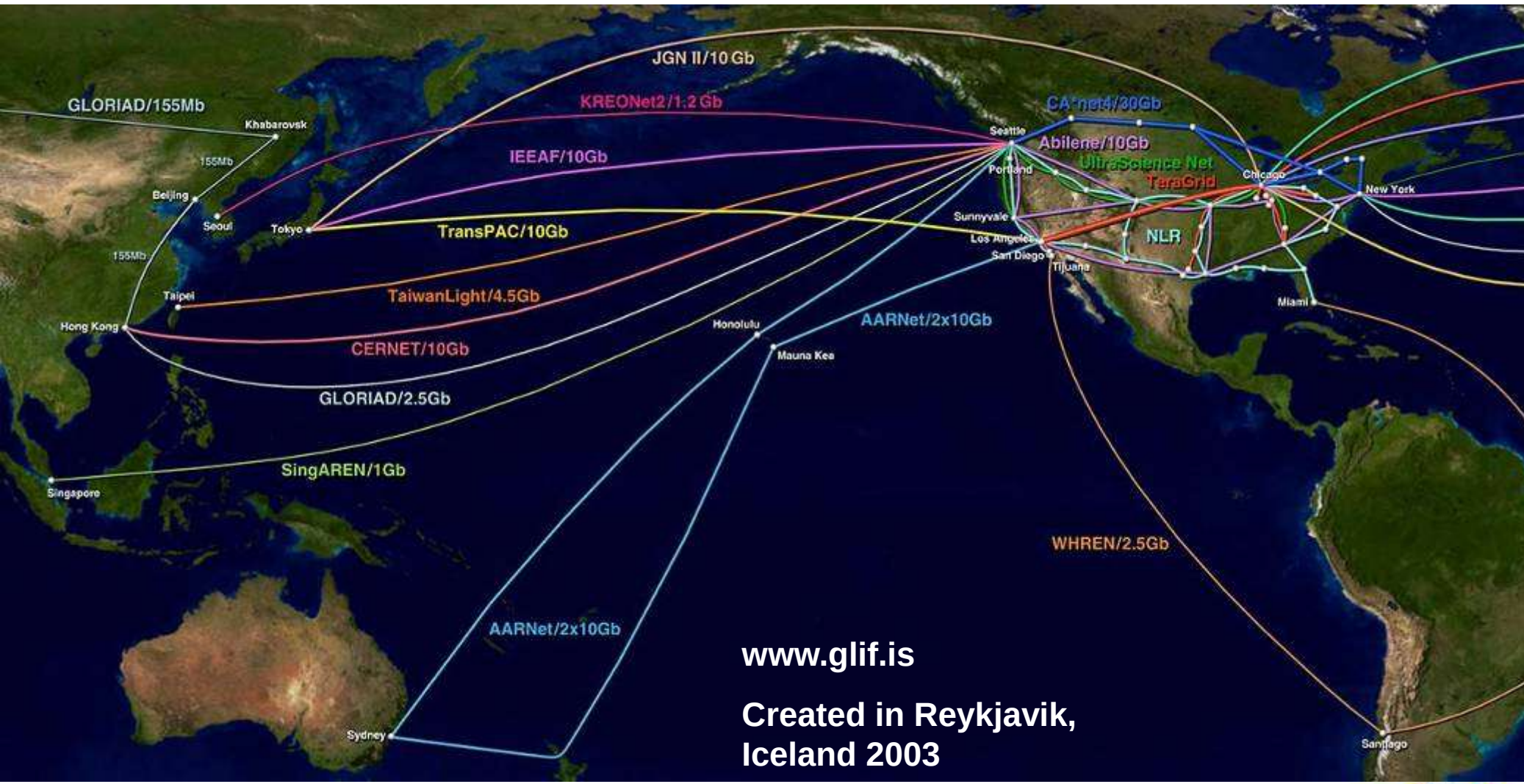


130 Million Households with FTTH in 2013

Source: Heavy Reading (www.heavyreading.com), the market research division of Light Reading (www.lightreading.com).

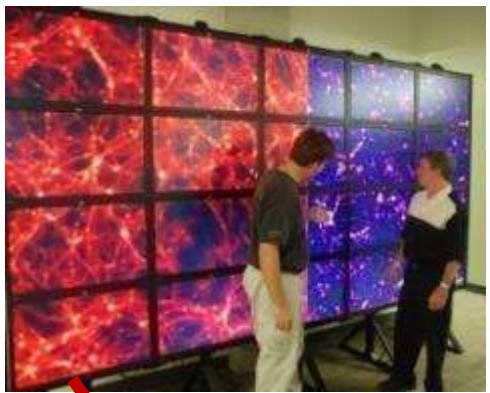
The Global Lambda Integrated Facility--
Creating a Planetary-Scale High Bandwidth Collaboratory

Research Innovation Labs Linked by 10G GLIF



Visualization courtesy of
Bob Patterson, NCSA.

Academic Research "OptIPlatform" Cyberinfrastructure: A 10Gbps "End-to-End" Lightpath Cloud



End User OptIPortal



Instruments



HD/4k Telepresence



HD/4k Video Cams

HPC



10G Lightpaths

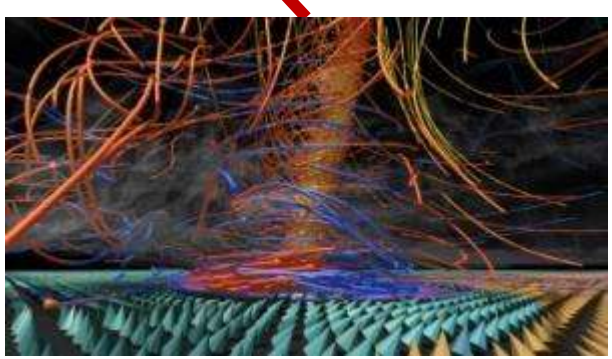


TeraGrid

Campus Optical Switch



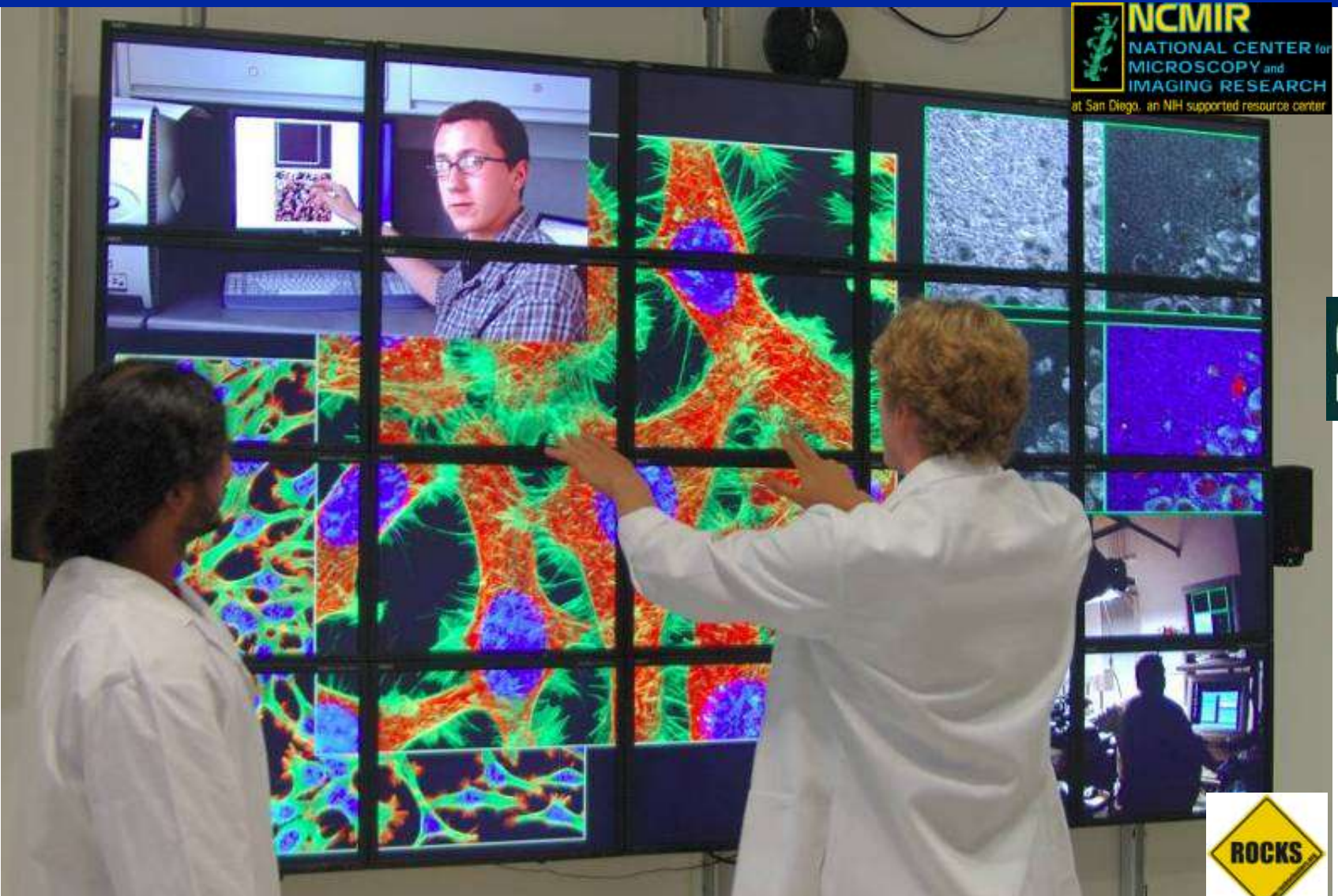
Data Repositories & Clusters



HD/4k Video Images

Data-Intensive Visualization and Analysis

The OptIPuter Project: Creating High Resolution Portals Over Dedicated Optical Channels to Global Science Data



Scalable Adaptive Graphics Environment (SAGE)

evl electronic visualization laboratory

NSF

Picture Source: Mark Ellisman, David Lee, Jason Leigh

Calit2 (UCSD, UCI), SDSC, and UIC Leads—Larry Smarr PI

Univ. Partners: NCSA, USC, SDSU, NW, TA&M, UvA, SARA, KISTI, AIST

Industry: IBM, Sun, Telcordia, Chiaro, Calient, Glimmerglass, Lucent

On-Line Resources
Help You Build Your Own OptIPortal

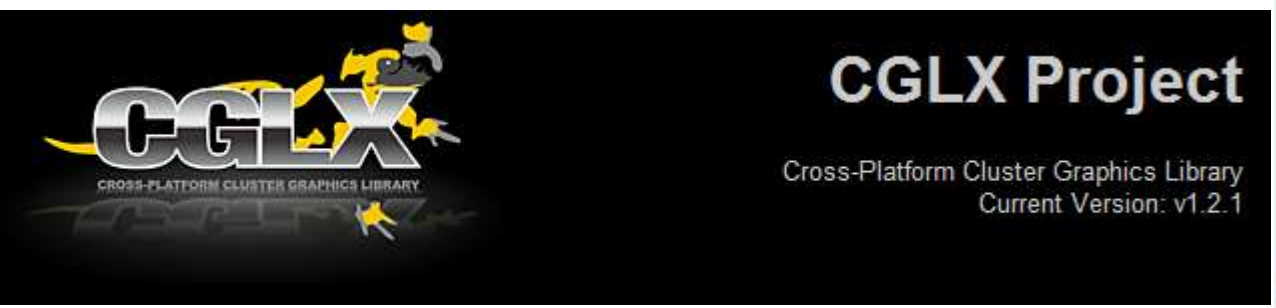
www.optiputer.net

http://wiki.optiputer.net/optiportal



www.evl.uic.edu/cavern/sage/





http://vis.ucsd.edu/~cglx/



OptIPortals Are Built
From Commodity PC Clusters and LCDs
To Create a 10Gbps Scalable Termination Device

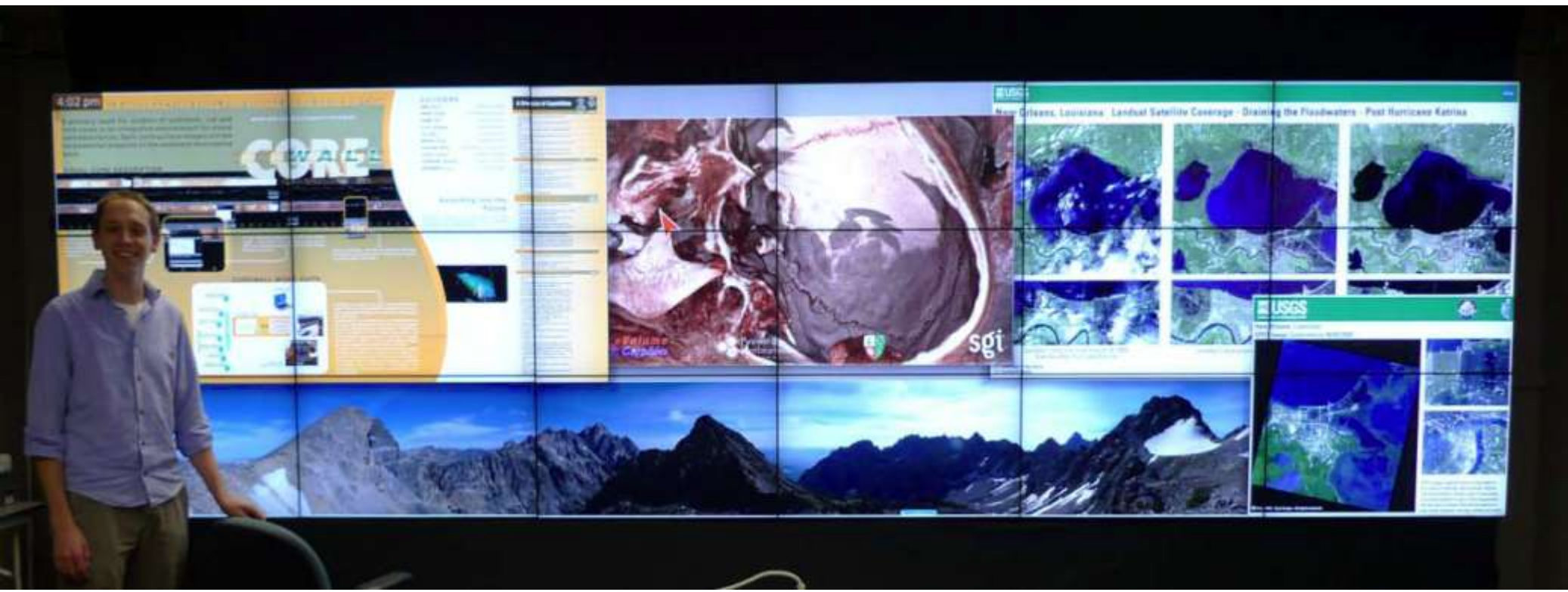
1/3 Billion Pixel OptIPortal Used to Study NASA Earth Satellite Images of October 2007 Wildfires



Source: Falko Kuester, Calit2@UCSD

Nearly Seamless AESOP OptIPortal

46” NEC Ultra-Narrow Bezel 720p LCD Monitors



Source: Tom DeFanti, Calit2@UCSD;

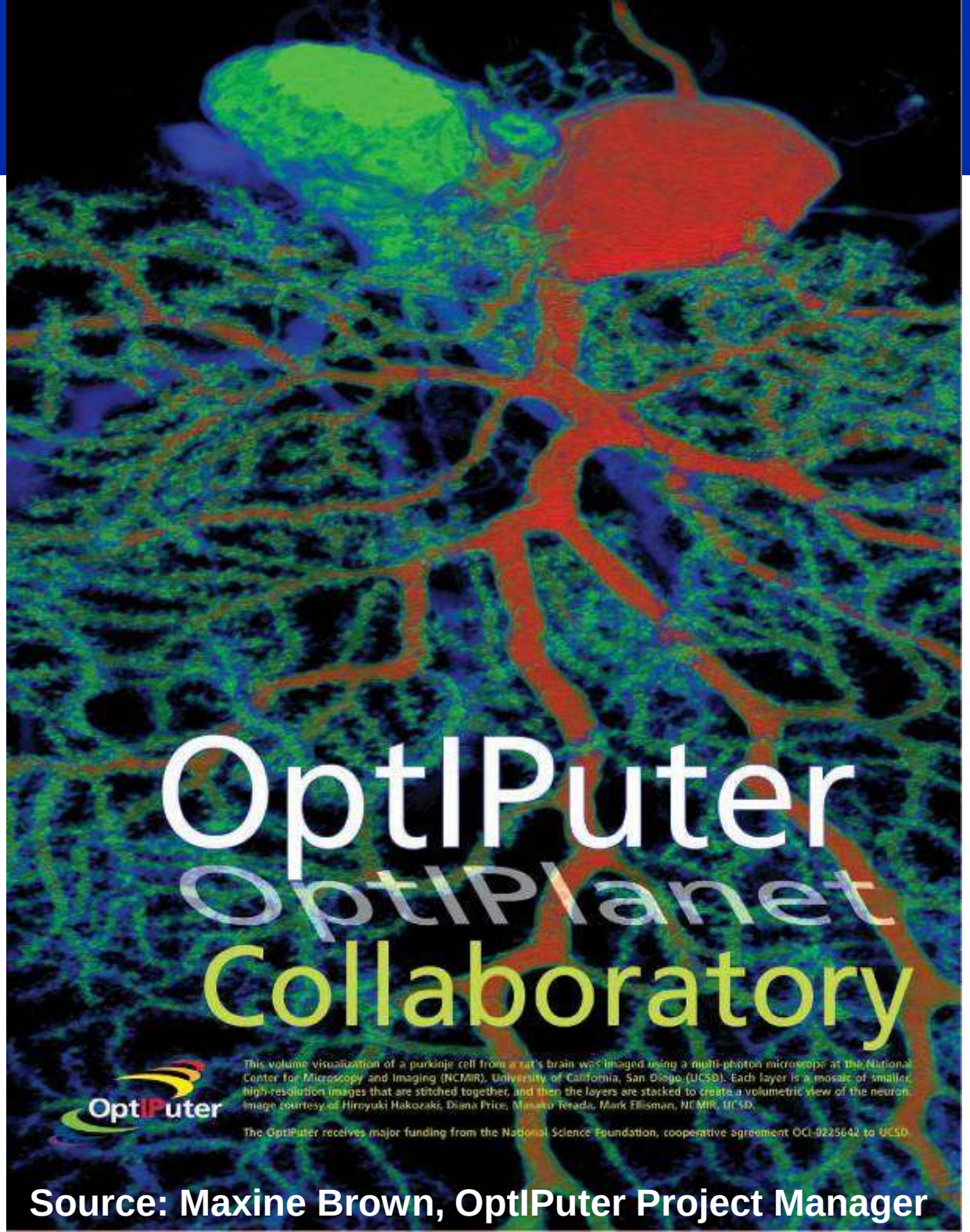
3D Stereo Head Tracked OptIPortal: NexCAVE



Array of JVC HDTV 3D LCD Screens
KAUST NexCAVE = 22.5MPixels

www.calit2.net/newsroom/article.php?id=1584

Source: Tom DeFanti, Calit2@UCSD



Source: Maxine Brown, OptIPuter Project Manager

Green Initiative:

Can Optical Fiber Replace Airline Travel for Continuing Collaborations ?

Multi-User Global Workspace: San Diego, Chicago, Saudi Arabia



Source: Tom DeFanti, KAUST Project, Calit2

CineGrid 4K Remote Microscopy
USC to Calit2



Richard Weinberg, USC

First Tri-Continental Premier of a Streamed 4K Feature Film With Global HD Discussion



CRCA

CISCO
C-wave

CineGrid

4K Transmission Over 10Gbps--
4 HD Projections from One 4K Projector

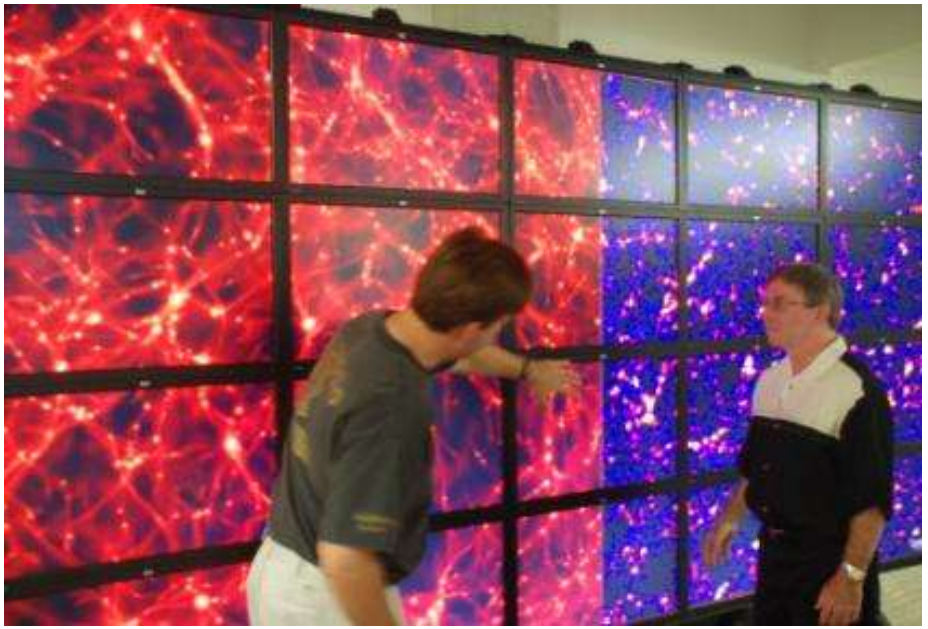
it²

End-to-end WAN
HPCI

Project StarGate Goals: Combining Supercomputers and Supernetworks

- Create an "End-to-End" 10Gbps Workflow
- Explore Use of OptIPortals as Petascale Supercomputer "Scalable Workstations"
- Exploit Dynamic 10Gbps Circuits on ESnet
- Connect Hardware Resources at ORNL, ANL, SDSC
- Show that Data Need Not be Trapped by the Network "Event Horizon"

OptIPortal@SDSC



Rick Wagner Mike Norman

Source: Michael Norman, SDSC, UCSD

- ANL * Calit2 * LBNL * NICS * ORNL * SDSC

Using Supernetworks to Couple End User's OptIPortal to Remote Supercomputers and Visualization Servers

Source: Mike Norman, SDSC

Argonne NL

DOE Eureka
100 Dual Quad Core Xeon Servers
200 NVIDIA Quadro FX GPUs in 50 Quadro Plex S4 1U enclosures
3.2 TB RAM



SDSC

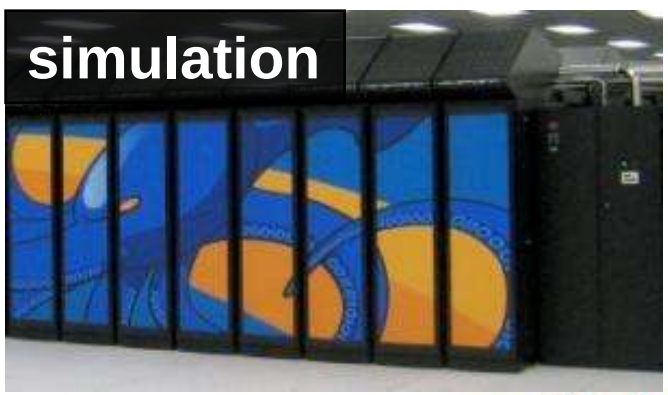


ESnet
10 Gb/s fiber optic network

NICS ORNL

Calit2/SDSC OptIPortal1
20 30" (2560 x 1600 pixel) LCD panels
10 NVIDIA Quadro FX 4600 graphics cards > 80 megapixels
10 Gb/s network throughout

NSF TeraGrid Kraken
Cray XT5
8,256 Compute Nodes
99,072 Compute Cores
129 TB RAM



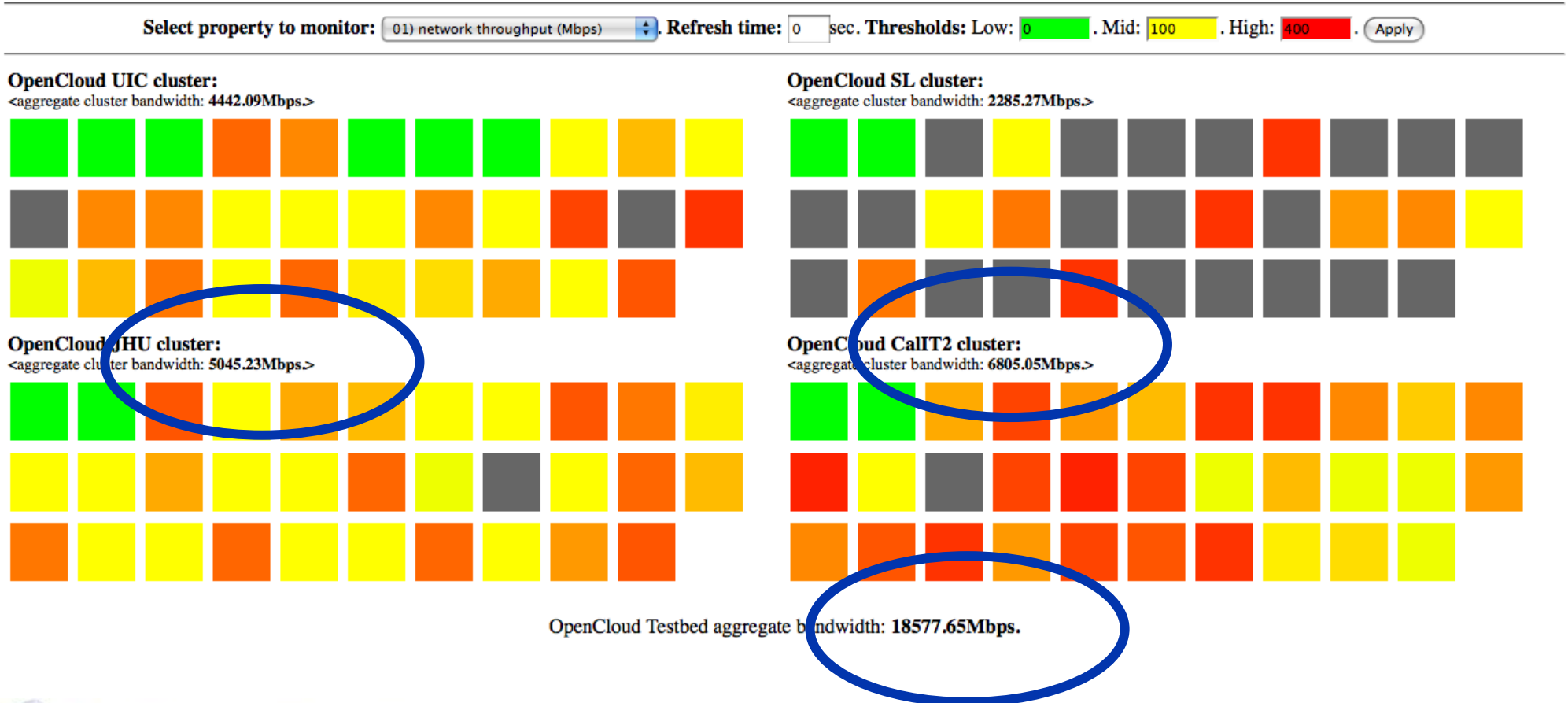
SDSC

***ANL * Calit2 * LBNL * NICS * ORNL * SDSC**

it²

Wavelengths and the Appropriate Cloud Middleware Make Wide Area Clouds Practical

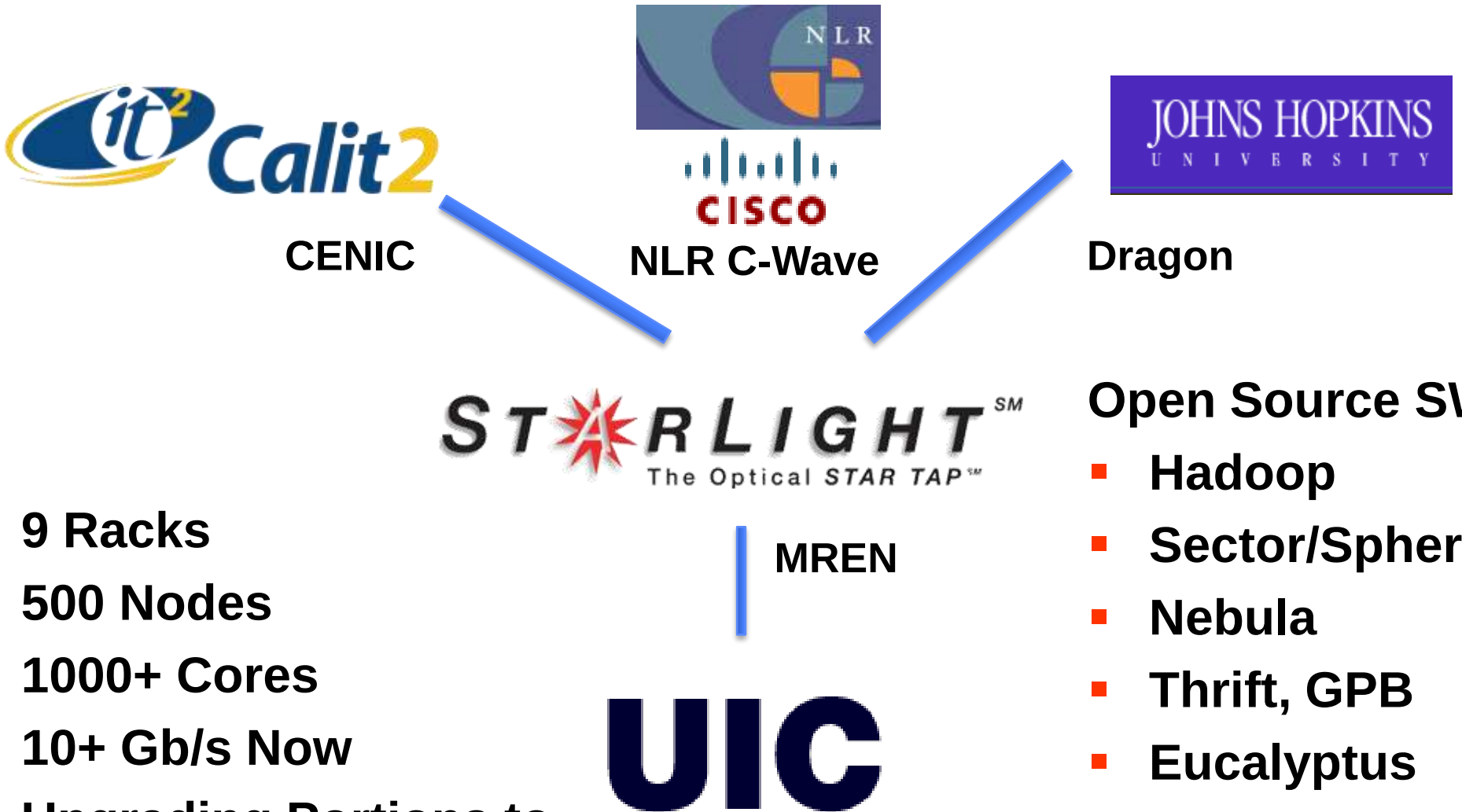
Terasort on Open Cloud Testbed
Sorting 10 Billion Records (1.2 TB) at 4 Sites (120 Nodes)
Sustaining >5 Gbps--Only 5% Distance Penalty



Open Cloud Consortium

it²

Open Cloud OptIPuter Testbed--Manage and Compute Large Datasets Over 10Gbps Lambdas



- 9 Racks
- 500 Nodes
- 1000+ Cores
- 10+ Gb/s Now
- Upgrading Portions to 100 Gb/s in 2010/2011

Open Source SW

- Hadoop
- Sector/Sphere
- Nebula
- Thrift, GPB
- Eucalyptus
- Benchmarks

Open Cloud Consortium

Source: Robert Grossman, UChicago

Sector Won the SC 08 and SC 09 Bandwidth Challenge



2009: Sector/Sphere Sustained Over 100 Gbps Cloud Computation Across 4 Geographically Distributed Data Centers

2008: Sector/Sphere Used for a Variety of Scientific Computing Applications on Open Cloud Testbed.



Open Cloud Consortium

Source: Robert Grossman, UChicago

it²

California and Washington Universities Are Testing a 10Gbps Connected Commercial Data Cloud

- **Amazon Experiment for Big Data**
  - **Only Available Through CENIC & Pacific NW GigaPOP**
    - **Private 10Gbps Peering Paths**
  - **Includes Amazon EC2 Computing & S3 Storage Services**
- **Early Experiments Underway**
  - **Robert Grossman, Open Cloud Consortium**
  - **Phil Papadopoulos, Calit2/SDSC Rocks**

amazon web services™

cenic

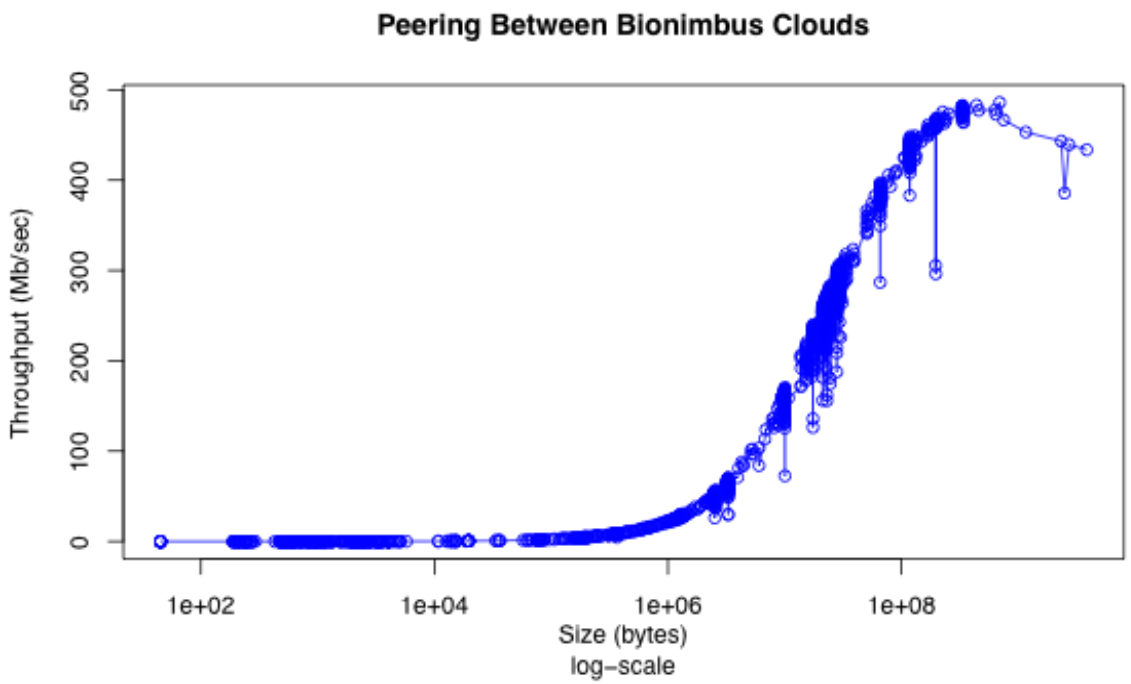
PACIFIC NORTHWEST GIGAPOP

it²

Hybrid Cloud Computing with modENCODE Data

- Computations in Bionimbus Can Span the Community Cloud & the Amazon Public Cloud to Form a Hybrid Cloud
- Sector was used to Support the Data Transfer between Two Virtual Machines
  - One VM was at UIC and One VM was an Amazon EC2 Instance
- Graph Illustrates How the Throughput between Two Virtual Machines in a Wide Area Cloud Depends upon the File Size







Biological data (Bionimbus)

Source: Robert Grossman, UChicago

Moving into the Clouds: Rocks and EC2

- **We Can Build Physical Hosting Clusters & Multiple, Isolated Virtual Clusters:**
  - **Can I Use Rocks to Author "Images" Compatible with EC2? (We Use Xen, They Use Xen)**
  - **Can I Automatically Integrate EC2 Virtual Machines into My Local Cluster (Cluster Extension)**
    - **Submit Locally**
    - **My Own Private + Public Cloud**
- **What This Will Mean**
  - **All your Existing Software Runs Seamlessly Among Local and Remote Nodes**
  - **User Home Directories are Mounted**
  - **Queue Systems Work**
  - **Unmodified MPI Works**

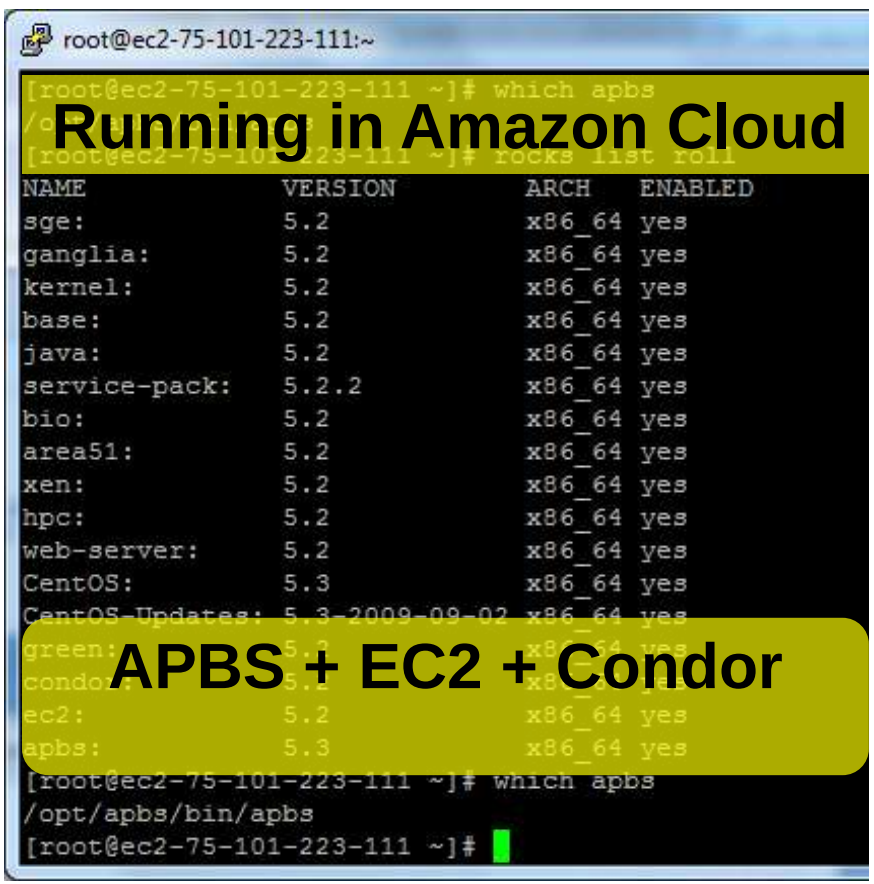
SDSC

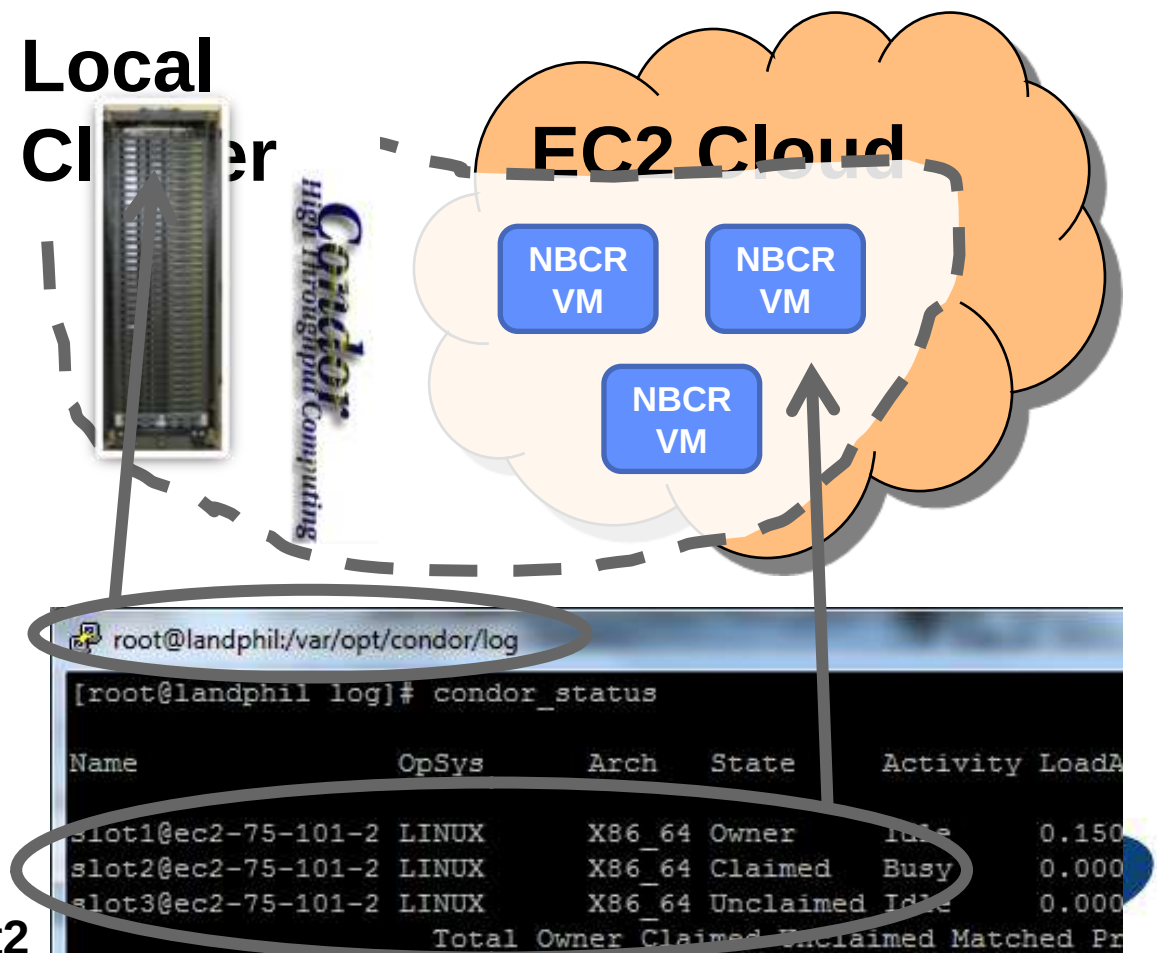
Source: Phil Papadopoulos, SDSC/Calit2

it²

Proof of Concept Using Condor and Amazon EC2 Adaptive Poisson-Boltzmann Solver (APBS)

- **APBS Rocks Roll (NBCR) + EC2 Roll + Condor Roll = Amazon VM**
- **Cluster extension into Amazon using Condor**





Source: Phil Papadopoulos, SDSC/Calit2

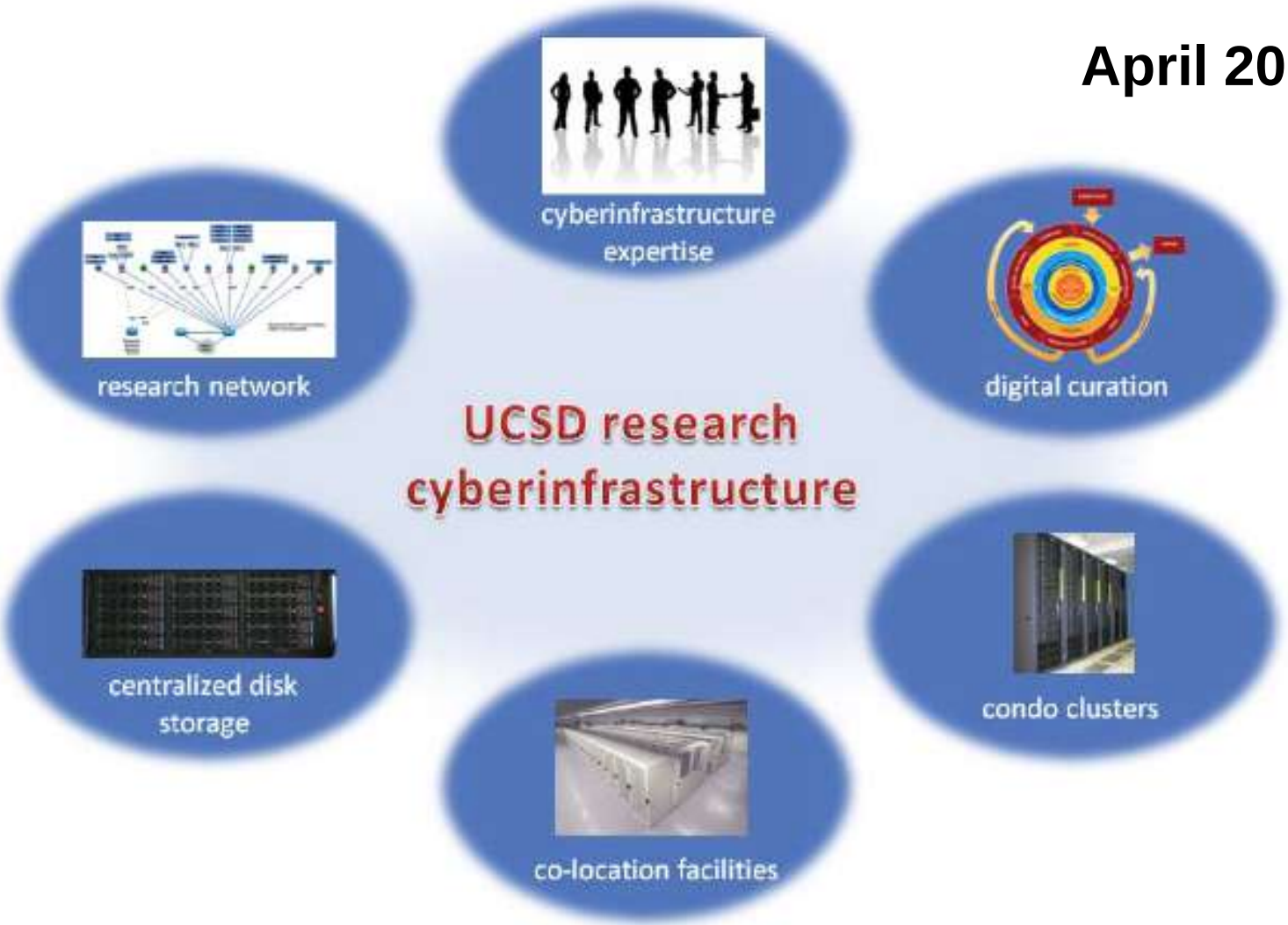
Data-Intensive Research Campus CI

"Blueprint for the Digital University"--Report of the UCSD Research Cyberinfrastructure Design Team

- **Focus on Data-Intensive Cyberinfrastructure**

April 2009

No Data Bottlenecks --Design for Gigabit/s Data Flows



http://research.ucsd.edu/documents/rcidt/RCIDTReportFinal2009.pdf

SDSC

it²

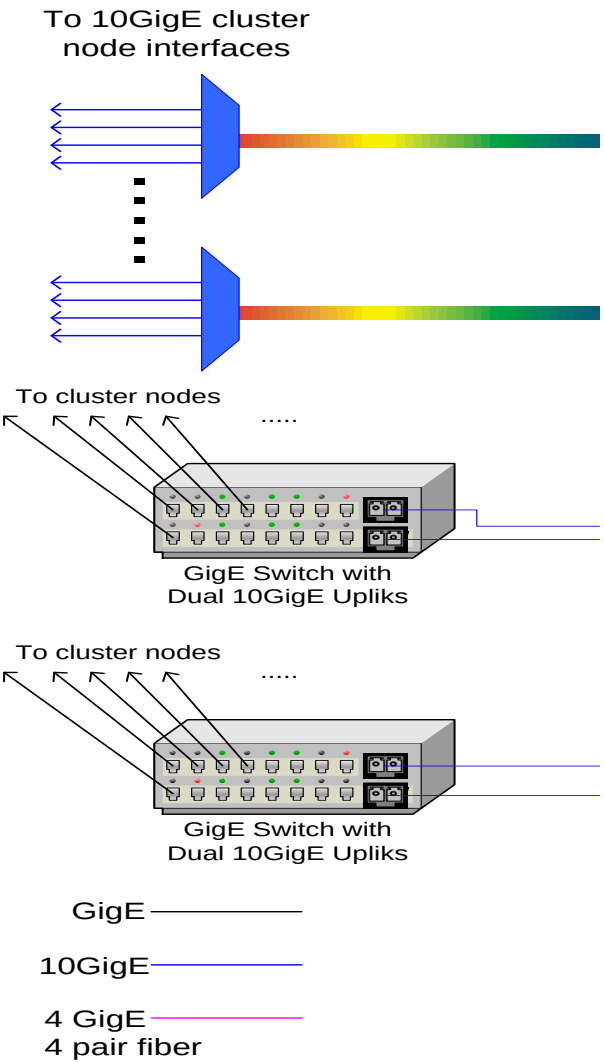
Broad Campus Input to Build the Plan and Support for the Plan

- **Campus Survey of CI Needs-April 2008**
  - 45 Responses (Individuals, Groups, Centers, Depts)
  - #1 Need was Data Management
    - 80% Data Backup
    - 70% Store Large Quantities of Data
    - 64% Long Term Data Preservation
    - 50% Ability to Move and Share Data
- **Vice Chancellor of Research Took the Lead**
- **Case Studies Developed from Leading Researchers**
- **Broad Research CI Design Team**
  - Chaired by Mike Norman and Phil Papadopoulos
  - Faculty and Staff:
    - Engineering, Oceans, Physics, Bio, Chem, Medicine, Theatre
    - SDSC, Calit2, Libraries, Campus Computing and Telecom

SDSC

it²

Current UCSD Optical Core: Bridging End-Users to CENIC L1, L2, L3 Services



Enpoints:

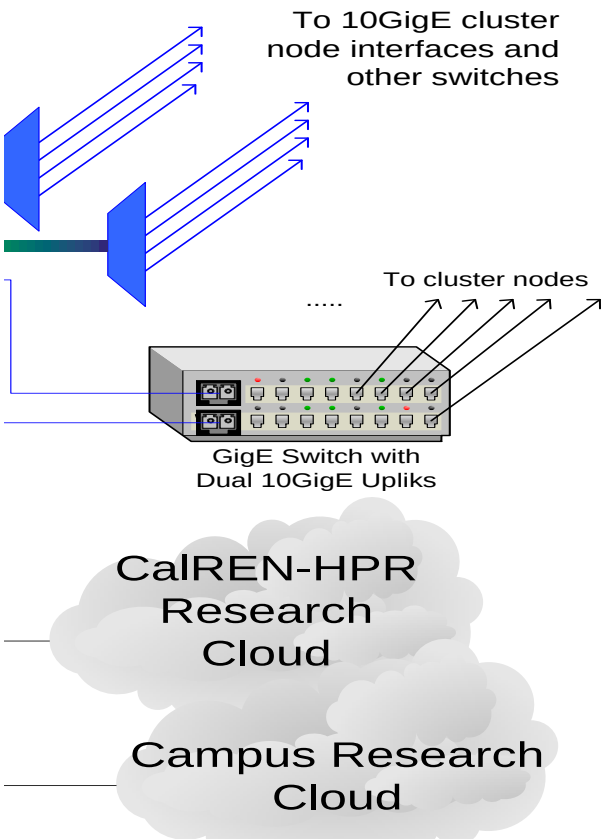
>= 60 endpoints at 10 GigE

>= 32 Packet switched

>= 32 Switched wavelengths

>= 300 Connected endpoints

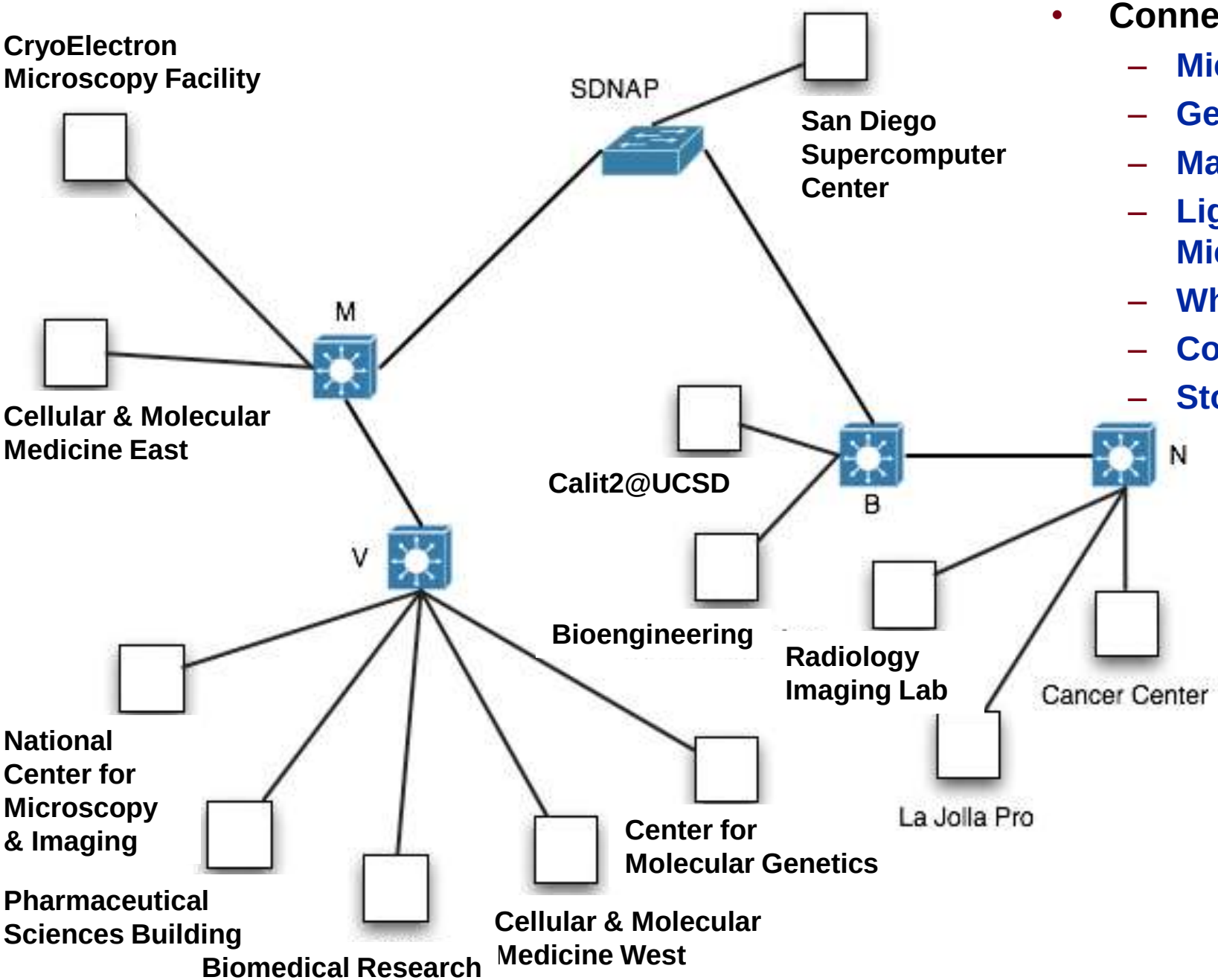
Approximately 0.5 TBit/s Arrive at the "Optical" Center of Campus. Switching is a Hybrid of: Packet, Lambda, Circuit -- OOO and Packet Switches



Source: Phil Papadopoulos, SDSC/Calit2 (Quartzite PI, OptIPuter co-PI)

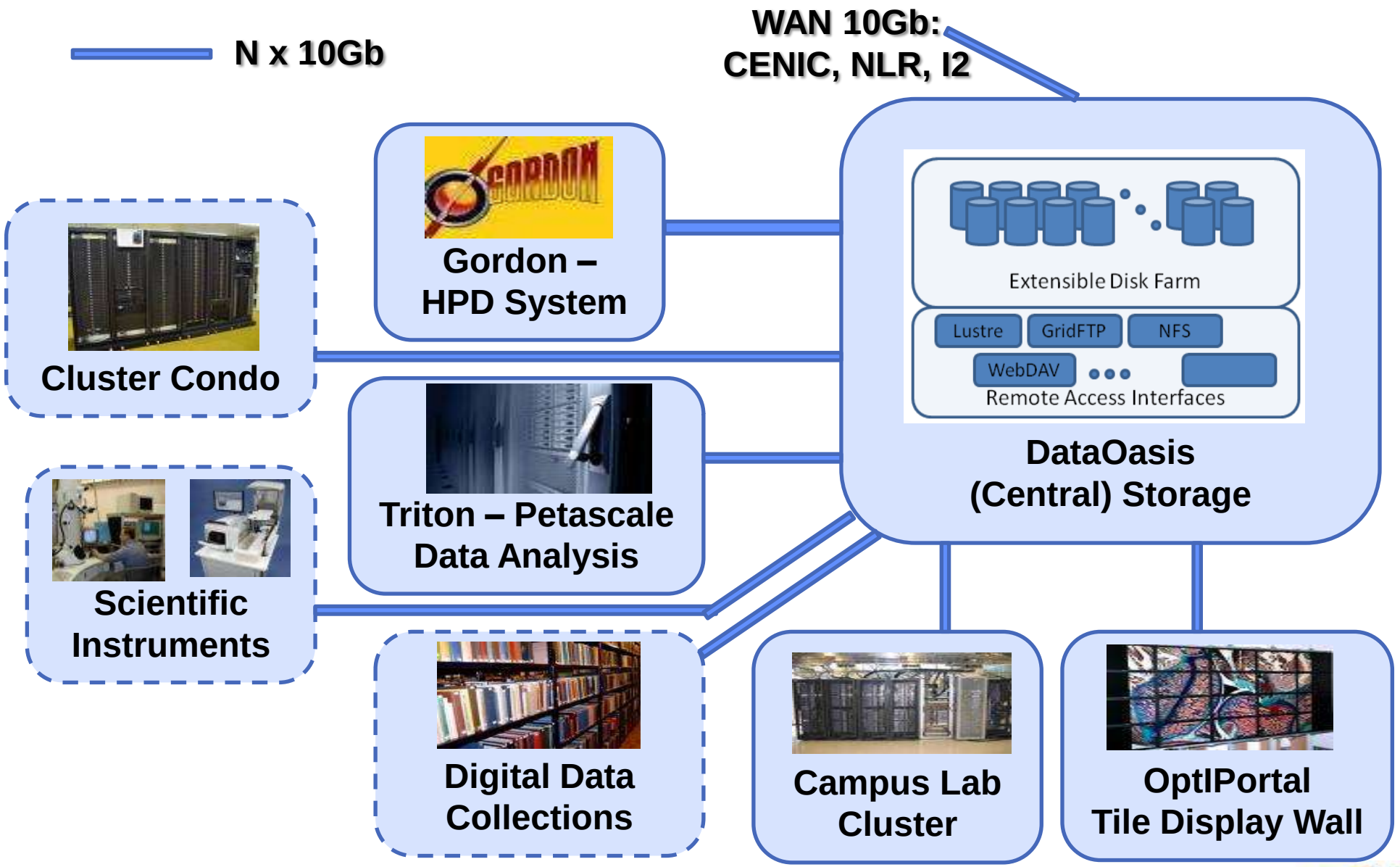
Quartzite Network MRI #CNS-0421555; OptIPuter #ANI-0225642

UCSD Planned Optical Networked Biomedical Researchers and Instruments



- **Connects at 10 Gbps :**
  - **Microarrays**
  - **Genome Sequencers**
  - **Mass Spectrometry**
  - **Light and Electron Microscopes**
  - **Whole Body Imagers**
  - **Computing**
  - **Storage**

UCSD Campus Investment in Fiber Enables Consolidation of Energy Efficient Computing & Storage



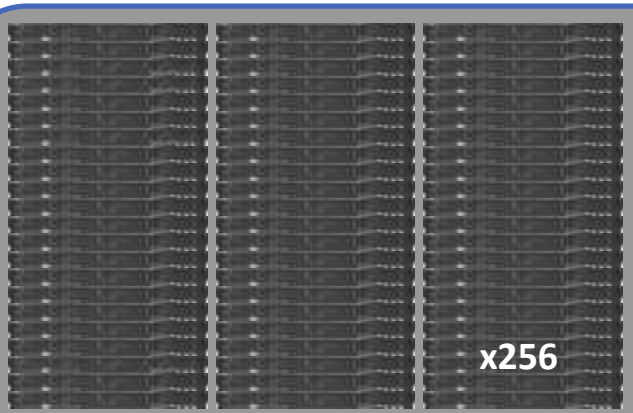
Source: Philip Papadopoulos, SDSC/Calit2

Moving to a Shared Campus Data Storage and Analysis Resource: Triton Resource @ SDSC

Triton Resource

Large Memory PSDAF
- 256/512 GB/sys
- 9TB Total
- 128 GB/sec
- ~ 9 TF





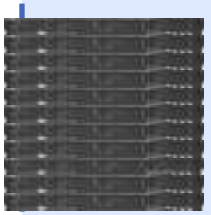
Shared Resource Cluster
- 24 GB/Node
- 6TB Total
- 256 GB/sec
- ~ 20 TF



Large Scale Storage
- 2 PB
- 40 – 80 GB/sec
- 3000 – 6000 disks
- Phase 0: 1/3 TB, 8GB/s



UCSD Research Labs







Campus Research Network

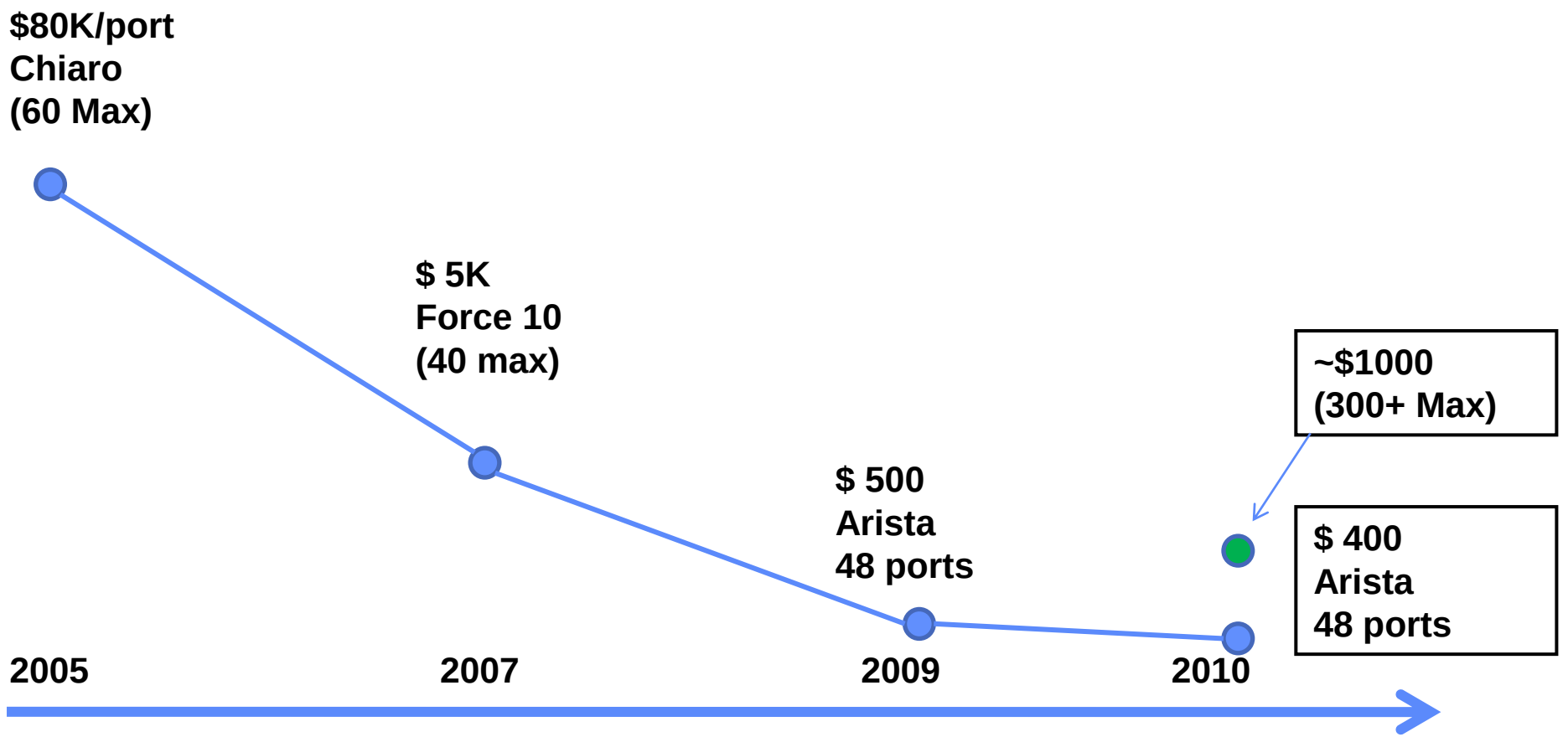
SDSC

Source: Philip Papadopoulos, SDSC/Calit2

it²

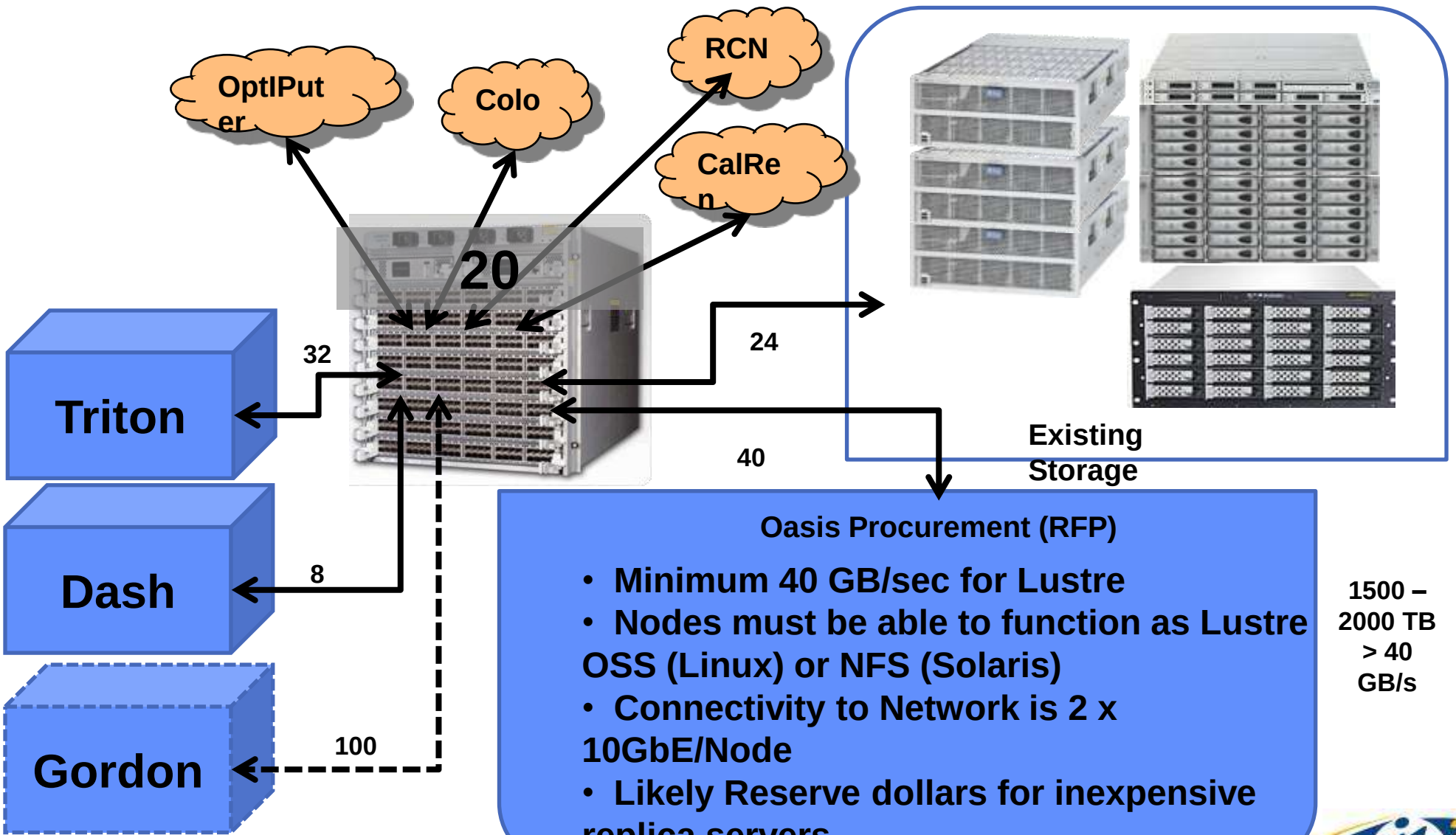
Rapid Evolution of 10GbE Port Prices Makes Campus-Scale 10Gbps CI Affordable

- **Port Pricing is Falling**
- **Density is Rising – Dramatically**
- **Cost of 10GbE Approaching Cluster HPC Interconnects**



Source: Philip Papadopoulos, SDSC/Calit2

10G Switched Data Analysis Resource: Data Oasis (RFP Underway)



Source: Philip Papadopoulos, SDSC/Calit2

High Performance Computing (HPC) vs. High Performance Data (HPD)

Attribute	HPC	HPD
Key HW metric	Peak FLOPS	Peak IOPS
Architectural features	Many small-memory multicore nodes	Fewer large-memory vSMP nodes
Typical application	Numerical simulation	Database query Data mining
Concurrency	High concurrency	Low concurrency or serial
Data structures	Data easily partitioned e.g. grid	Data not easily partitioned e.g. graph
Typical disk I/O patterns	Large block sequential	Small block random
Typical usage mode	Batch process	Interactive

SDSC

Source: Mike Norman, SDSC

it²

What is Gordon?

- **Data-Intensive Supercomputer Based on SSD Flash Memory and Virtual Shared Memory SW**
  - ***Emphasizes MEM and IOPS over FLOPS***
- **System Designed to Accelerate Access to Massive Data Bases being Generated in all Fields of Science, Engineering, Medicine, and Social Science**
- **The NSF's Most Recent Track 2 Award to the San Diego Supercomputer Center (SDSC)**
- **Coming Summer 2011**



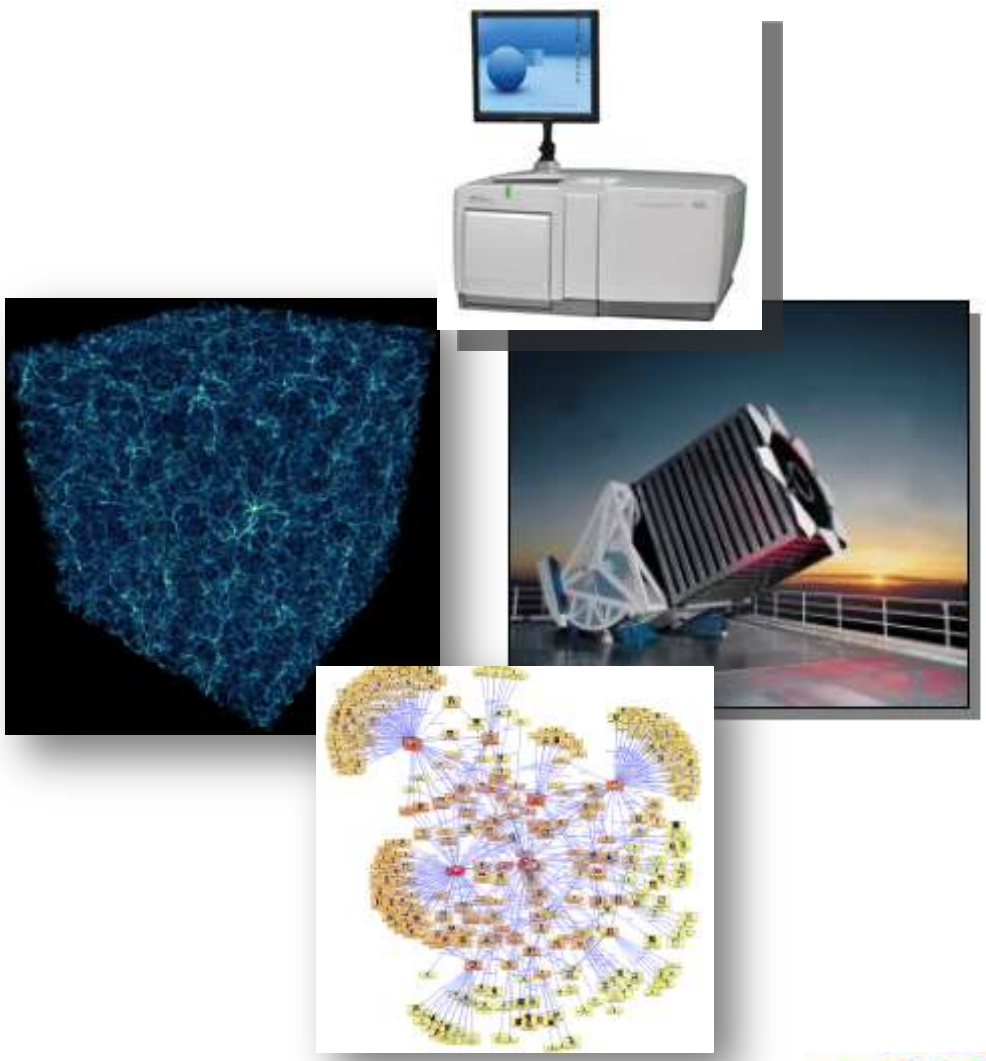
Source: Mike Norman, SDSC

SDSC

it²

Data Mining Applications will Benefit from Gordon

- ***De Novo* Genome Assembly from Sequencer Reads & Analysis of Galaxies from Cosmological Simulations & Observations**
  - *Will Benefit from Large Shared Memory*
- **Federations of Databases & Interaction Network Analysis for Drug Discovery, Social Science, Biology, Epidemiology, Etc.**
  - *Will Benefit from Low Latency I/O from Flash*



SDSC

Source: Mike Norman, SDSC

it²

Grand Challenges in Data-Intensive Sciences

October 26-28, 2010

San Diego Supercomputer Center , UC San Diego

Confirmed conference *topics* and speakers :

- ***Needs and Opportunities in Observational Astronomy*** - **Alex Szalay, JHU**
- ***Transient Sky Surveys*** – **Peter Nugent, LBNL**
- ***Large Data-Intensive Graph Problems*** – **John Gilbert, UCSB**
- ***Algorithms for Massive Data Sets*** – **Michael Mahoney, Stanford U.**
- ***Needs and Opportunities in Seismic Modeling and Earthquake Preparedness*** - **Tom Jordan, USC**
- ***Needs and Opportunities in Fluid Dynamics Modeling and Flow Field Data Analysis*** – **Parviz Moin, Stanford U.**
- ***Needs and Emerging Opportunities in Neuroscience*** – **Mark Ellisman, UCSD**
- ***Data-Driven Science in the Globally Networked World*** – **Larry Smarr, UCSD**

SDSC

it²

You Can Download This Presentation at lsmarr.calit2.net



ABOUT LARRY SMARR

Larry Smarr is the founding Director of the California Institute for Telecommunications and Information Technology (Calit2), a UC San Diego/UC Irvine partnership, and holds the Harry E. Gruber professorship in Computer Science and Engineering (CSE) at UCSD's Jacobs School. At Calit2, Smarr has

WHAT I'M READING

Complete List of Larry's Recommendations

Homo Britannicus: The Incredible Story of Human Life in Britain (2006)

American College and University

RECENT TWITTER POSTS

Twitter? Click To Follow Me

Ismarr: SDSC Founding Director Sid Karin and I are making a few comments on the history of the NSF Supercomputing Centers at