UC Capabilities Supporting High-Performance Collaboration and Data-Intensive Sciences

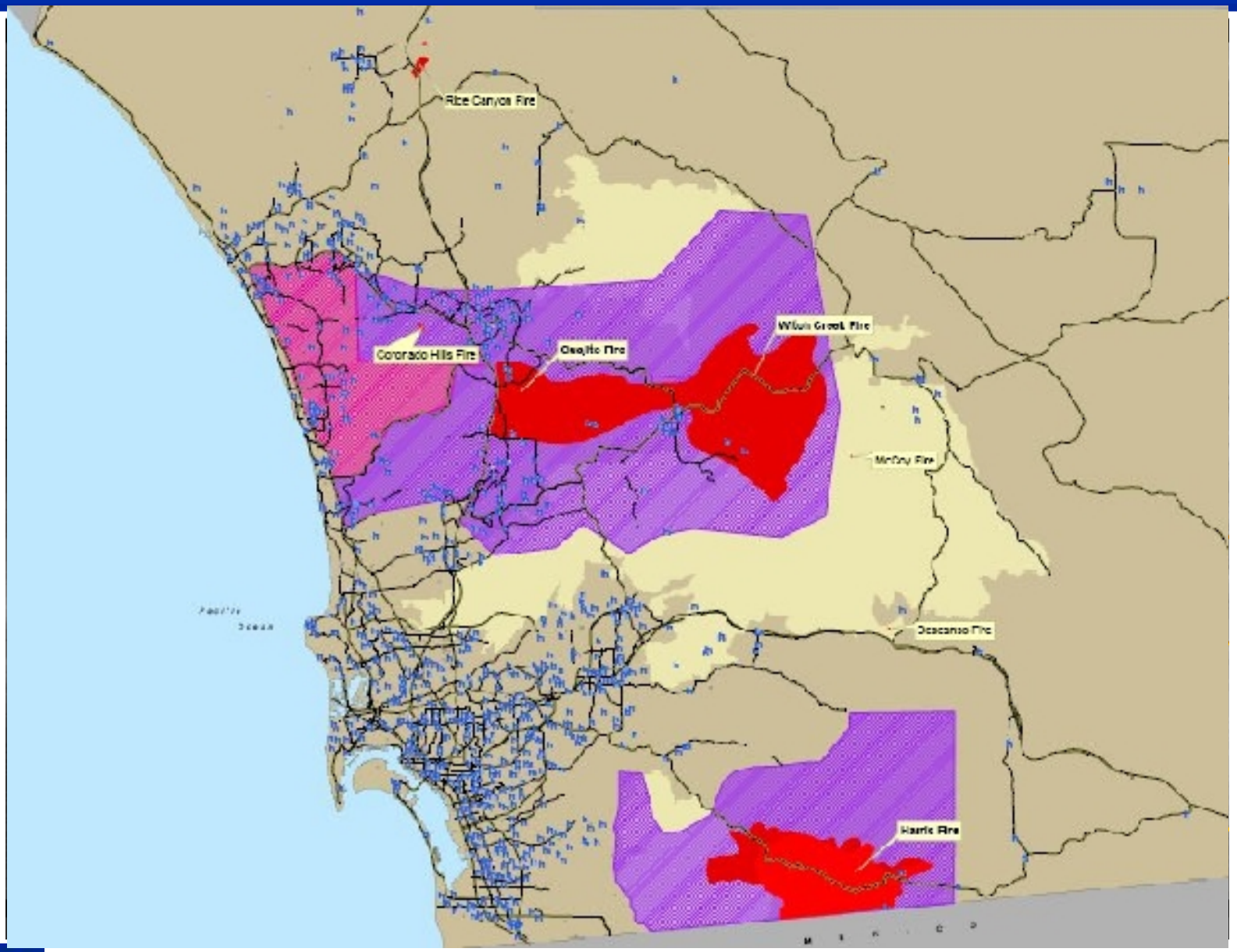
University of California Council of Research
University of California, Irvine
October 22, 2007

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

UCSD
UCIrvine

it²

Southern California Visible Satellite
250,000 Evacuated



UCSD
UCIrvine

NASA Aqua MODIS Image
2pm Monday PDT



UCSD
UCIrvine

Data Intensive e-Science Instruments Will Require SuperNetworks



UCSD
UCIrvine

it²

Access to CERN Large Hadron Collider
A Terabit/s WAN by 2013!

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
2005-6	10-20	2-4 X 10	λ Switch; λ Provisioning
2007-8	30-40	~100 or 2 X 40 Gbps	1st Gen. λ Grids
2009-10	60-80	~5 X 40 or ~2 X 100	40 or 100 Gbps λ Switching
2011-12	~5 X 40 or ~2 X 100	~25 X 40 or ~10 X 100	2nd Gen λ Grids Terabit Networks
2013-14	~Terabit	~MultiTbps	~Fill One Fiber

Source: Harvey Newman, Caltech

UCSD
UCIrvine

The Unrelenting Exponential Growth of Data Requires an Exponential Growth in Bandwidth

- **"Each LHC experiment foresees a recorded raw data rate of 1 to several PetaBytes/year"**
  - **Dr. Harvey Neuman (Cal Tech), Professor of Physics**

- **"The VLA facility is now able to generate 700 Gbps of astronomical data and the Extended VLA will reach 3.2 Terabits per second by 2009."**
  - **Dr. Steven Durand, National Radio Astronomy Observatory, E-VLBI Workshop, MIT Haystack Observatory., Sep 2006**

- **"The Global Information Grid will need to store and access exabytes of data on a realtime basis by 2010"**
  - **Dr. Henry Dardy (DOD), Optical Fiber Conference, Los Angeles, CA USA, Mar 2006**

- **"US Bancorp backs up 100 TB financial data every night – now."**
  - **David Grabski (VP Information Tech. US Bancorp), Qwest High Performance Networking Summit, Denver, CO. USA, June 2006.**

UCSD
UCIrvine

Source: Jerry Sobieski MAX / University of Maryland

it²

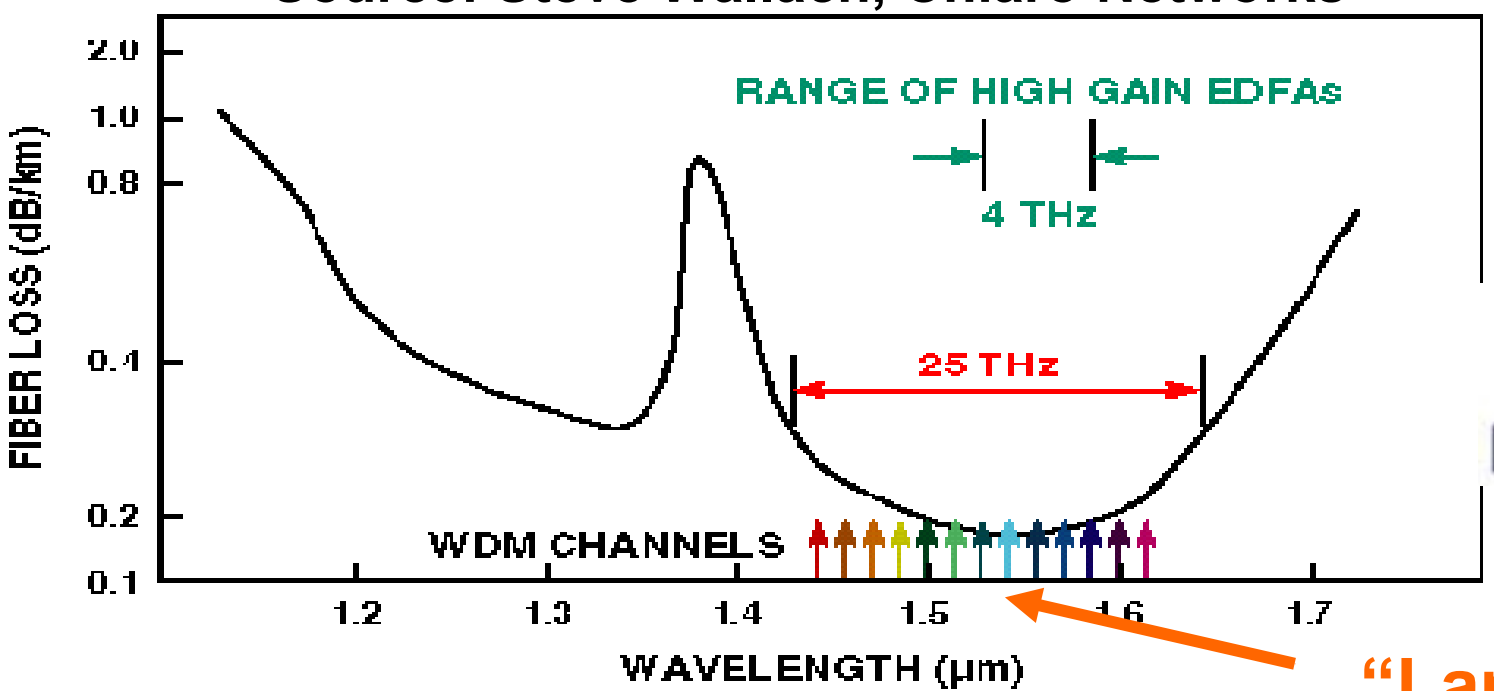
Dedicated Optical Channels Makes High Performance Cyberinfrastructure Possible

WAVELENGTH DIVISION MULTIPLEXING (WDM)

10 Gbps per User ~ 200x Shared Internet Throughput

- EXPLOITS
  - ENORMOUS BANDWIDTH OF SILICA FIBER
  - HIGH-GAIN WIDEBAND OPTICAL AMPLIFIERS

Source: Steve Wallach, Chiaro Networks



"Lambdas"

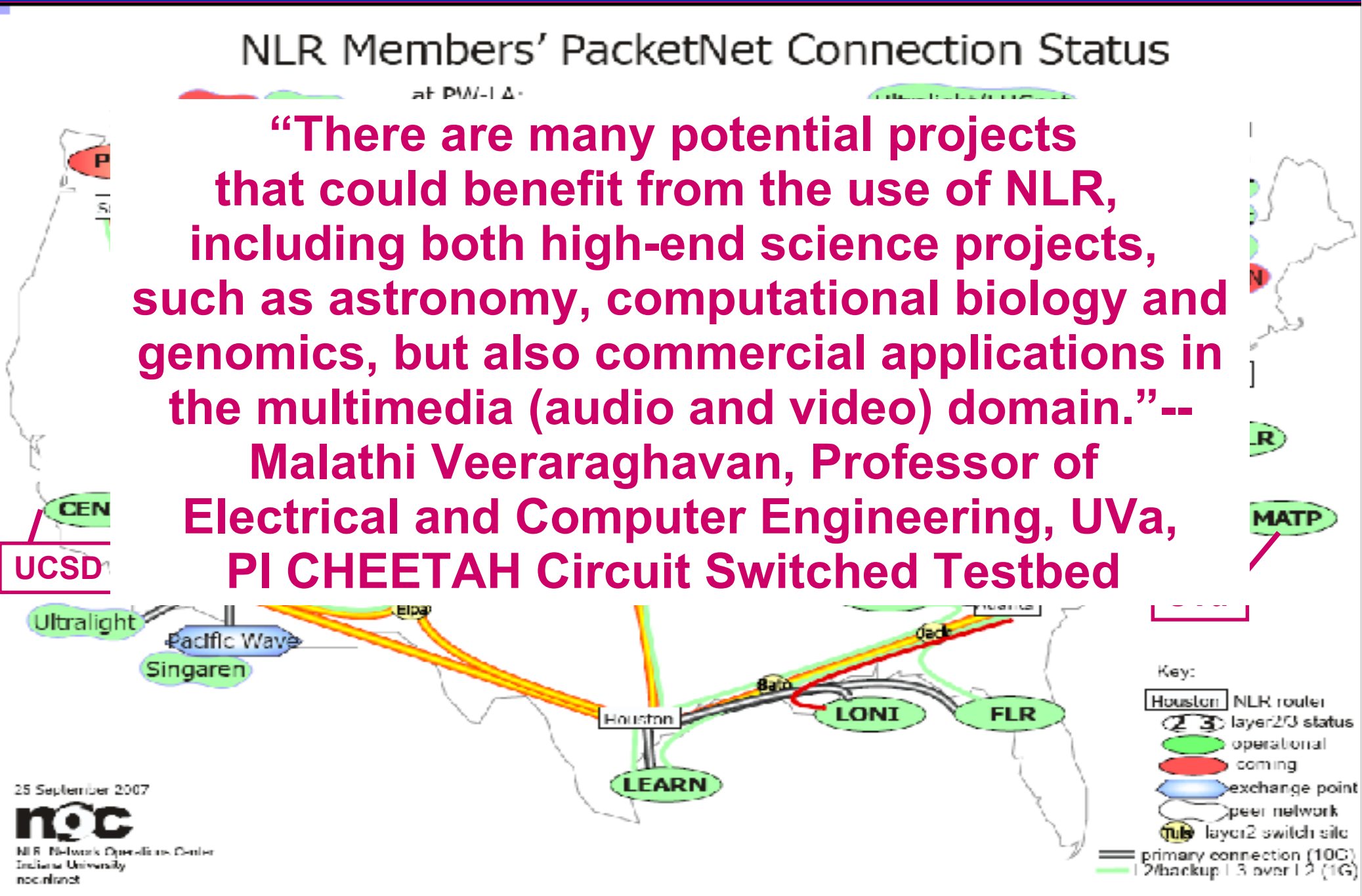
CHIARO

UCSD
UCIrvine

Parallel Lambdas are Driving Optical Networking The Way Parallel Processors Drove 1990s Computing

it²

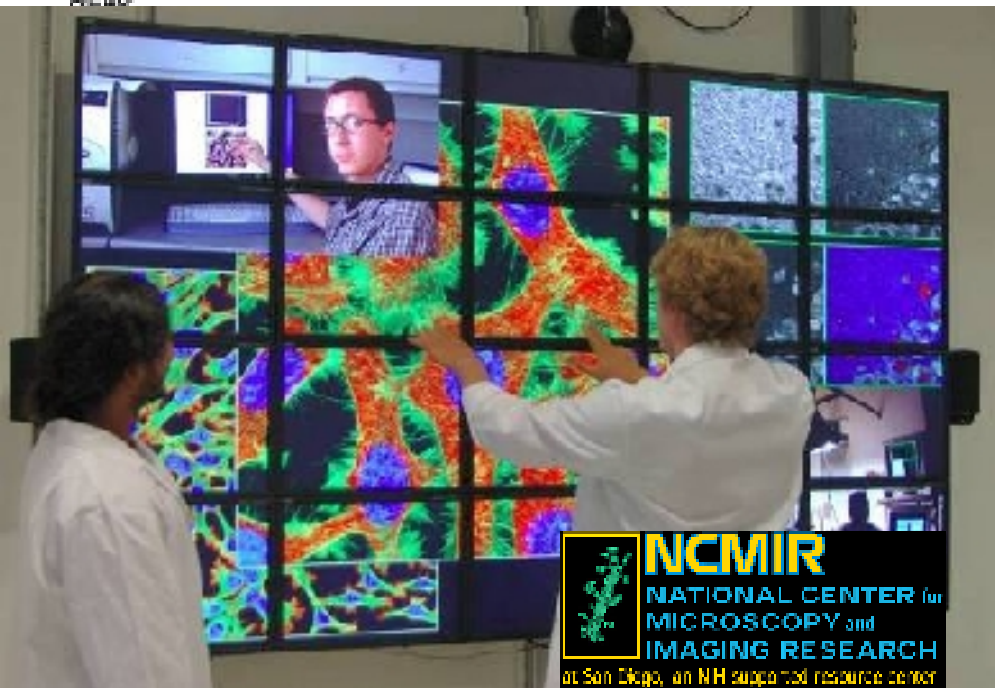
National LambdaRail
Serves the University of Virginia



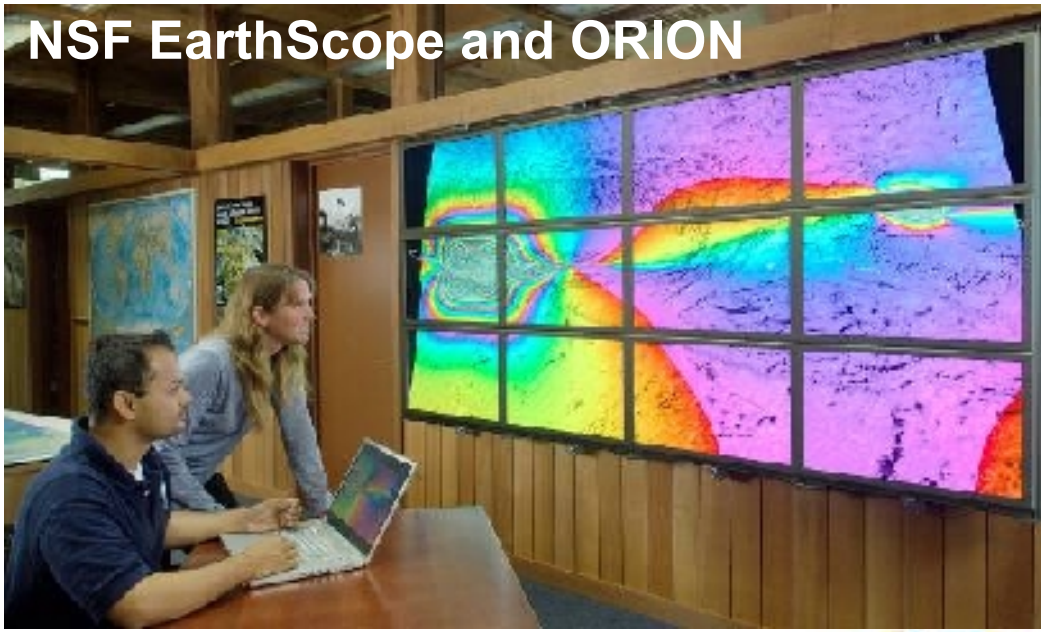
The OptIPuter Project – Creating High Resolution Portals Over Dedicated Optical Channels to Global Science Data

- NSF Large Information Technology Research Proposal
  - Calit2 (UCSD, UCI) and UIC Lead Campuses—Larry Smarr PI
  - Partnering Campuses: SDSC, USC, SDSU, NCSA, NW, TA&M, UvA, SARA, NASA Goddard, KISTI, AIST, CRC(Canada), CICESE (Mexico)
- Engaged Industrial Partners:
  - IBM, Sun, Telcordia, Chiaro, Calient, Glimmerglass, Lucent
- $13.5 Million Over Five Years—Now In the Six and Final Yea..

NIH Biomedical Informatics Research Network



NSF EarthScope and ORION



SCRIPPS INSTITUTION OF OCEANOGRAPHY

My OptIPortal™ – Affordable Termination Device for the OptIPuter Global Backplane

- 20 Dual CPU Nodes, 20 24” Monitors, ~$50,000
- 1/4 Teraflop, 5 Terabyte Storage, 45 Mega Pixels--Nice PC!
- Scalable Adaptive Graphics Environment (SAGE) Jason Leigh, EVL-UIC





UCSD UCIrvine

Powered by Rocks ROCKS

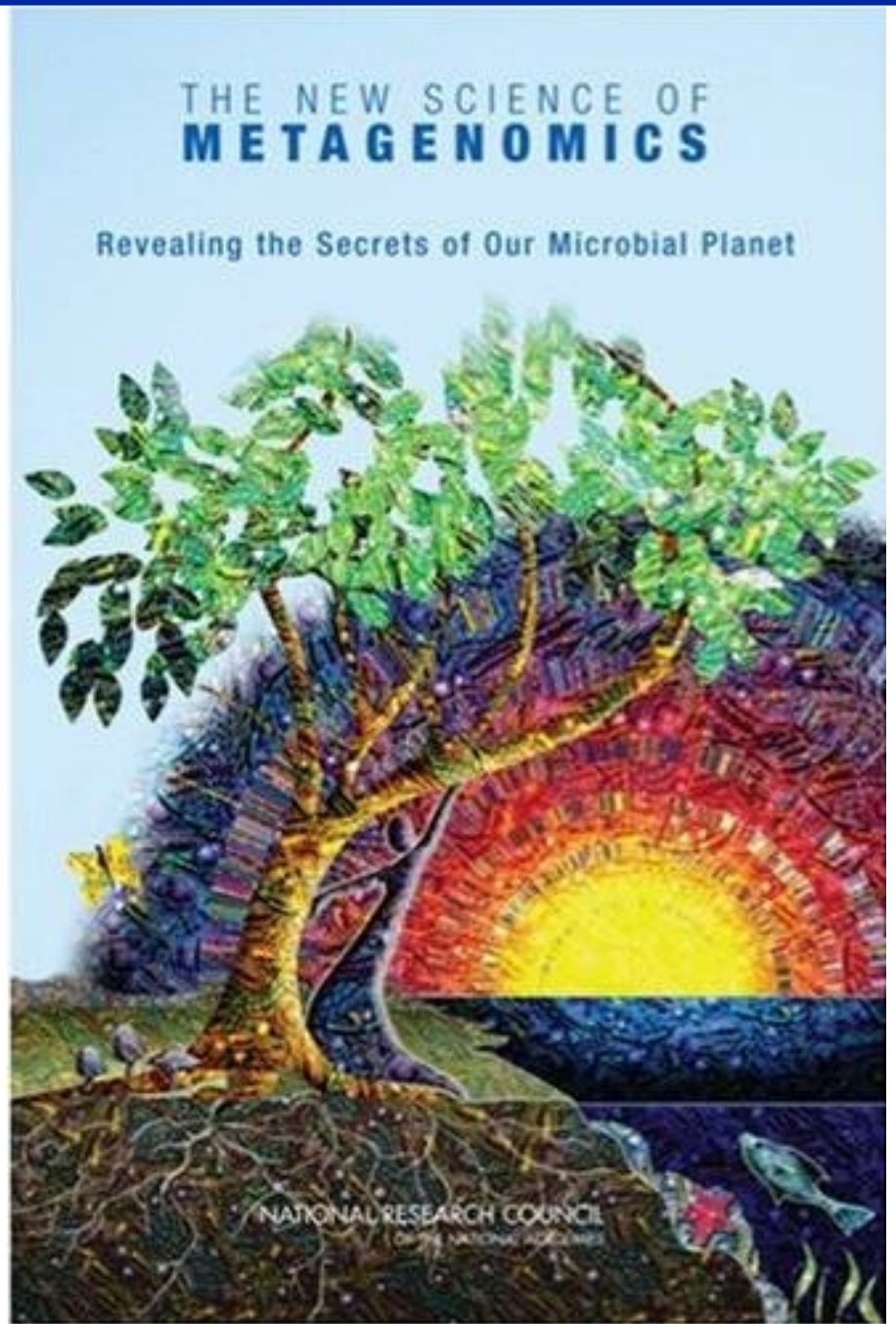
Source: Phil Papadopoulos SDSC, Calit2

it²

The New Science of Metagenomics

NRC Report:

Metagenomic data should be made publicly available in international archives as rapidly as possible.



"The emerging field of metagenomics, where the DNA of entire communities of microbes is studied simultaneously, presents the greatest opportunity -- perhaps since the invention of the microscope – to revolutionize understanding of the microbial world." –

National Research Council
March 27, 2007

UCSD
UCIrvine

it²

Marine Genome Sequencing Project – Measuring the Genetic Diversity of Ocean Microbes



MARINE MICROBIOLOGY INITIATIVE

Gordon and Betty MOORE FOUNDATION

Sorcerer II Data Will Double Number of Proteins in GenBank!

UCSD
UCIrvine

J. Craig Venter INSTITUTE

it²

Calit2's Direct Access Core Architecture Will Create Next Generation Metagenomics Server

Sargasso Sea Data

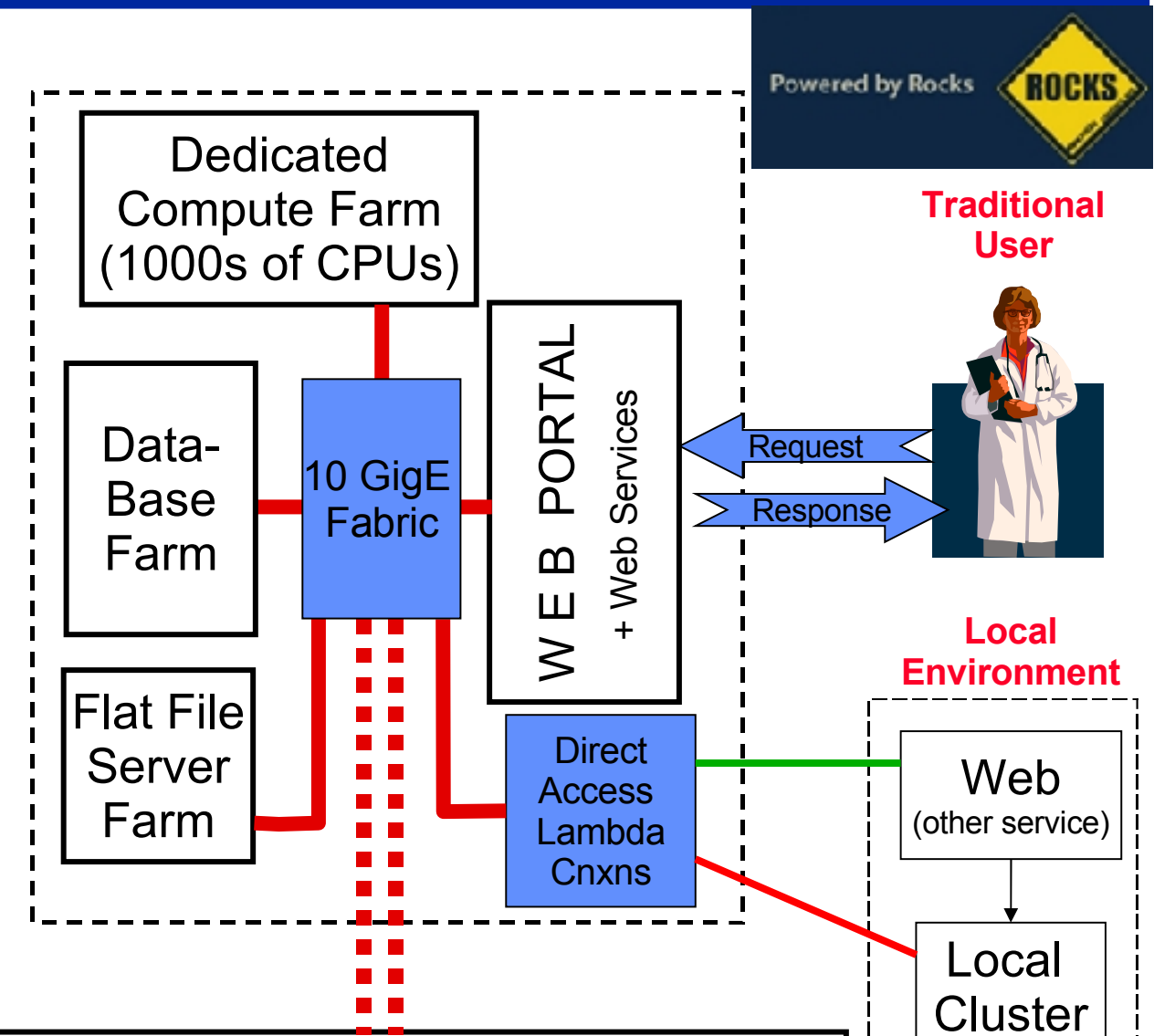
Sorcerer II Expedition (GOS)

JGI Community Sequencing Project

Moore Marine Microbial Project

NASA and NOAA Satellite Data

Community Microbial Metagenomics Data



Source: Phil Papadopoulos, SDSC, Calit2

“Instant” Global Microbial Metagenomics CyberCommunity

Over 1300 Registered Users From 48 Countries



USA	761
United Kingdom	64
Germany	54
Canada	46
France	44
Brazil	33

camera

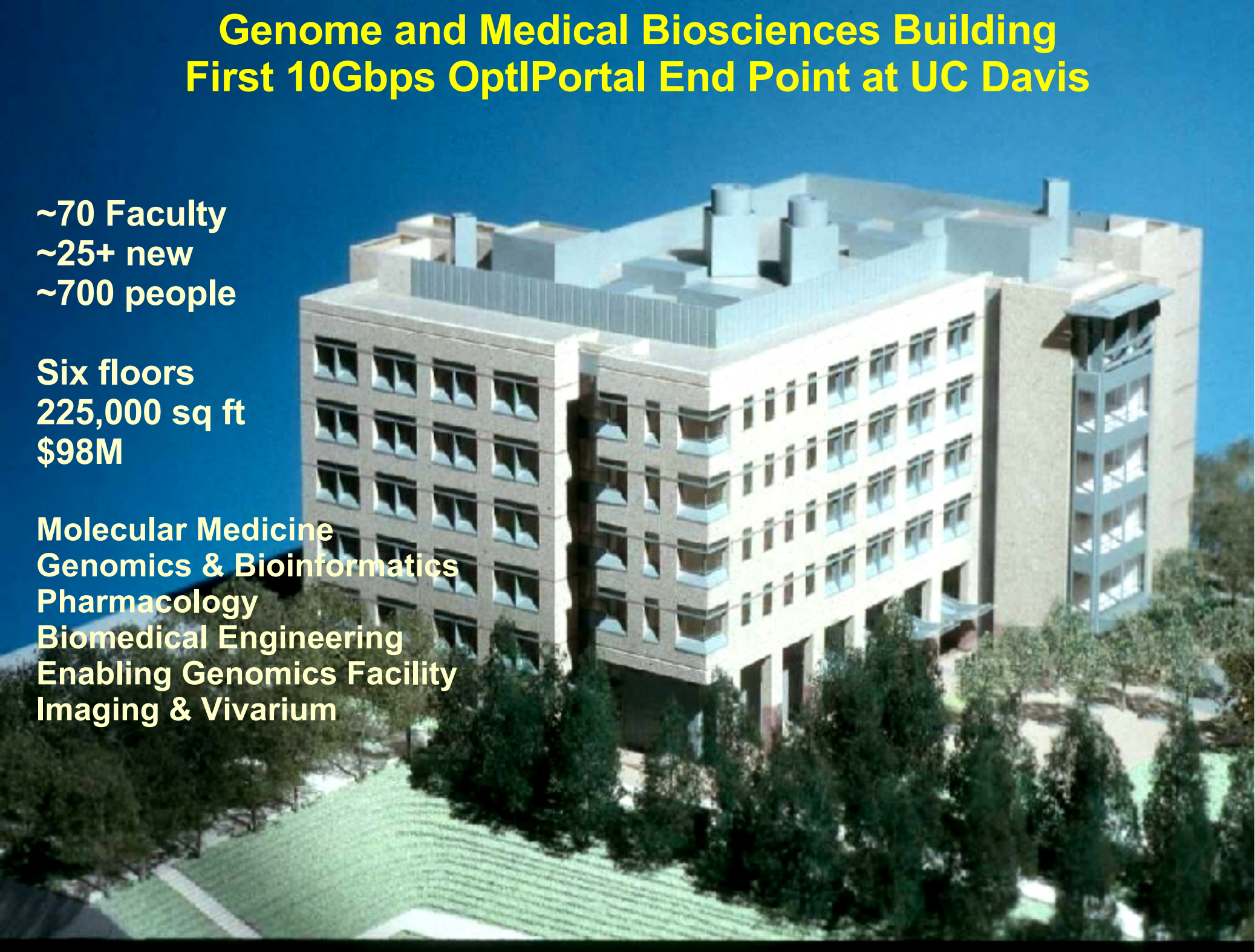
J. Craig Venter INSTITUTE

it²

An Emerging High Performance Collaboratory for Microbial Metagenomics



Genome and Medical Biosciences Building
First 10Gbps OptIPortal End Point at UC Davis

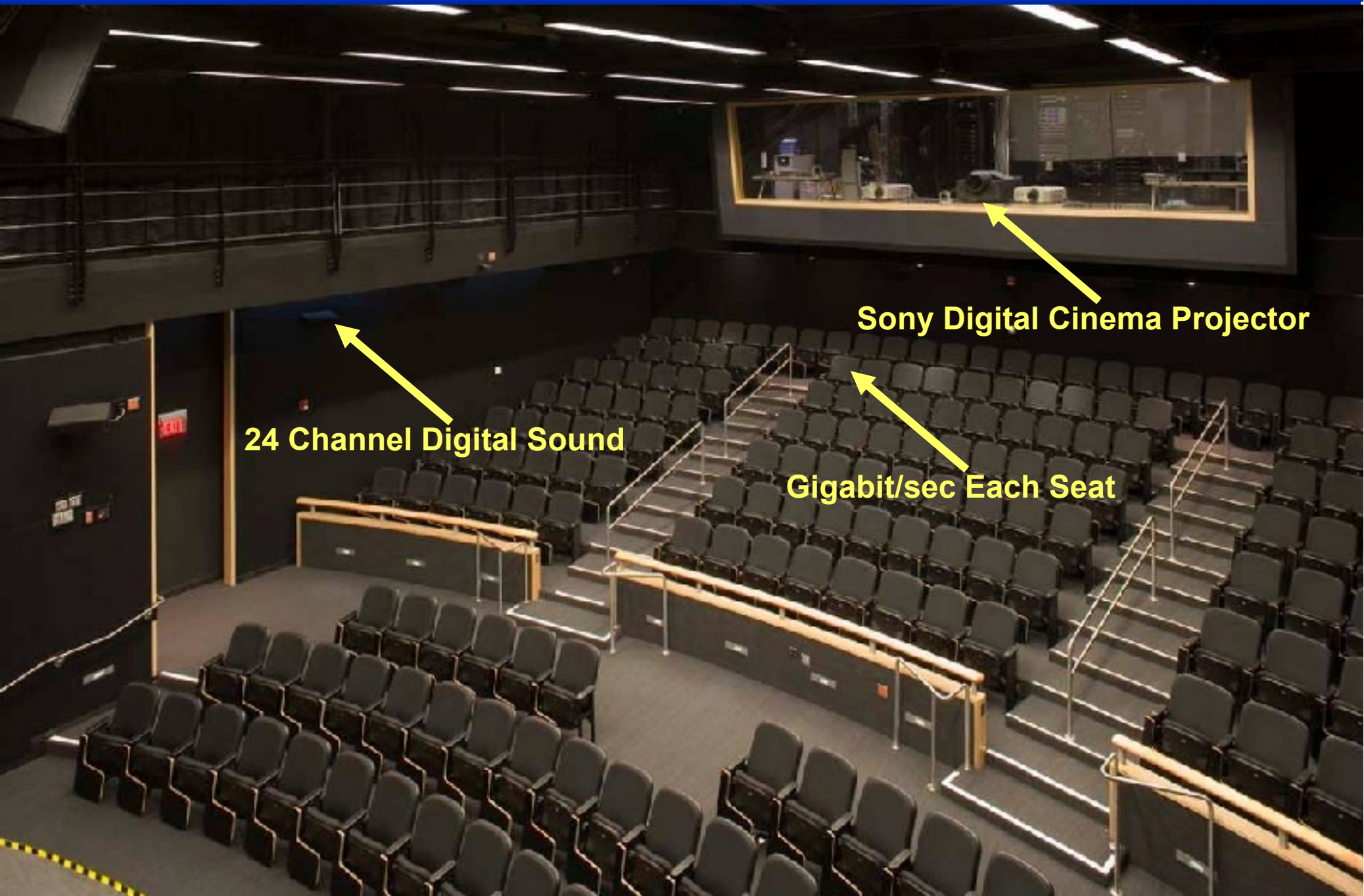


~70 Faculty
~25+ new
~700 people

Six floors
225,000 sq ft
$98M

Molecular Medicine
Genomics & Bioinformatics
Pharmacology
Biomedical Engineering
Enabling Genomics Facility
Imaging & Vivarium

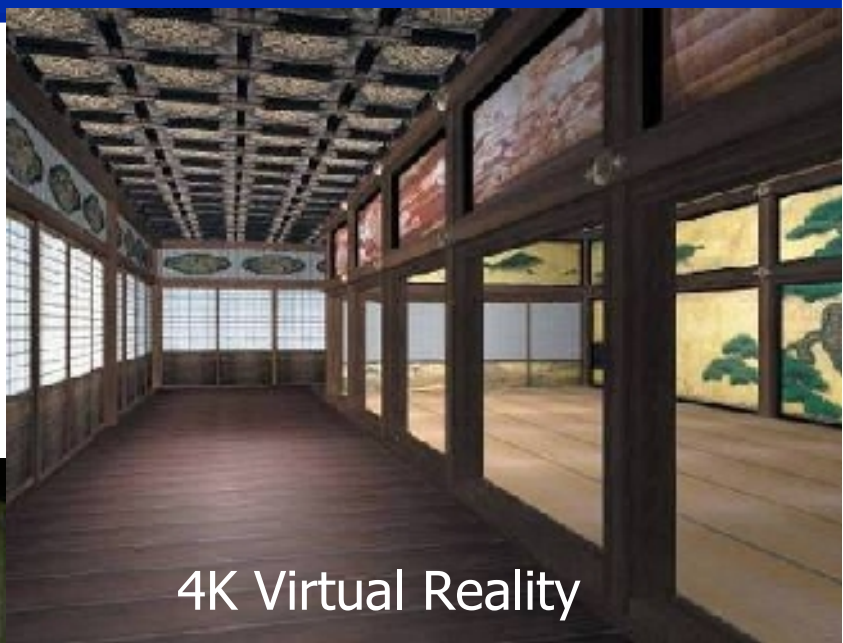
Building a Global Collaboratorium



Digital Cinema Standard 4k—Four Times HD!
CineGrid @ iGrid2005

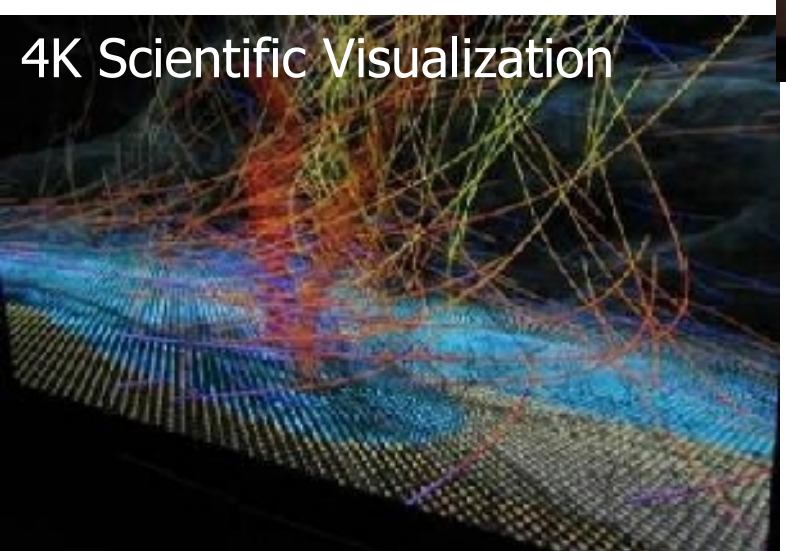


CineGrid



A New Digital Medium for Art, Science, and Collaboration



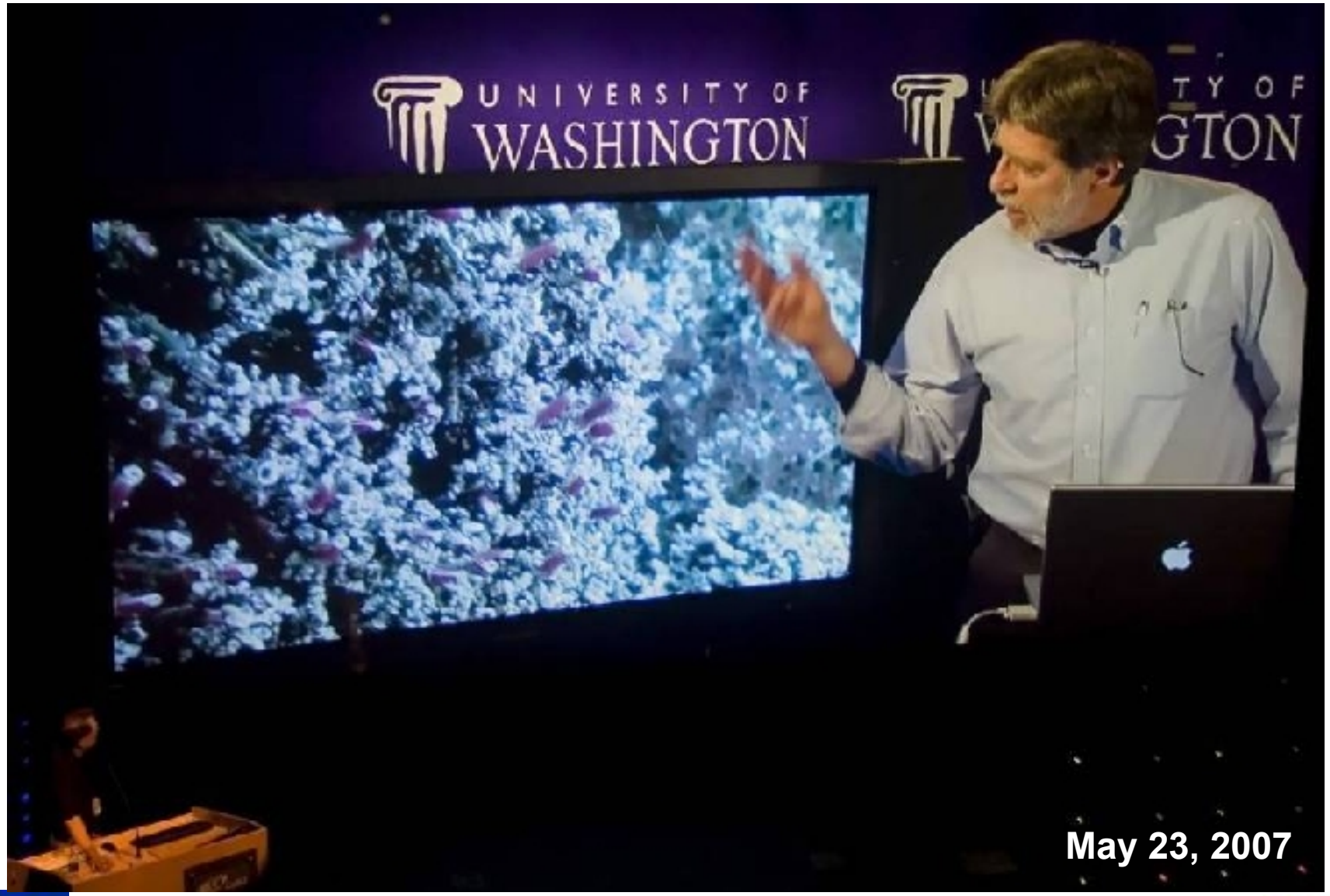




Source: Laurin Herr

e-Science Collaboratory Without Walls Enabled by Uncompressed HD Telepresence

1500 Mbits/sec Calit2 to UW Research Channel Over NLR



John Delaney, PI LOOKING, Neptune

Photo: Harry Ammons, SDSC

UCSD
UCIrvine

Goal for SC'07
iHDTV Integrated into OptIPortal

SC07 is the International Conference for High Performance Computing, Networking, Storage and Analysis

Moving from Compressed HD to Uncompressed iHDTV Reno to UW in Seattle



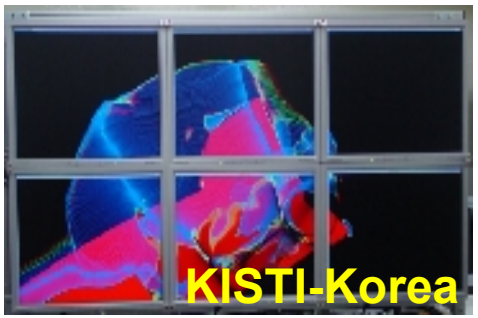


Source: Michael Wellings
Research Channel
Univ. Washington

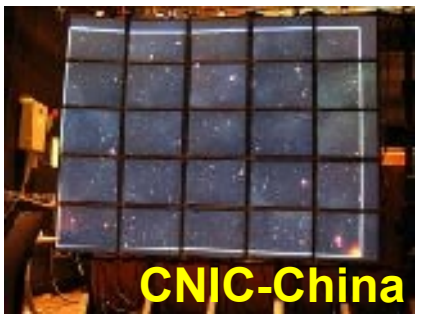
UCSD
UCIrvine

Rocks / SAGE OptIPortals Are Being Adopted Globally

























UCSD
UCIrvine

Powered by Rocks ROCKS

evl electronic visualization laboratory

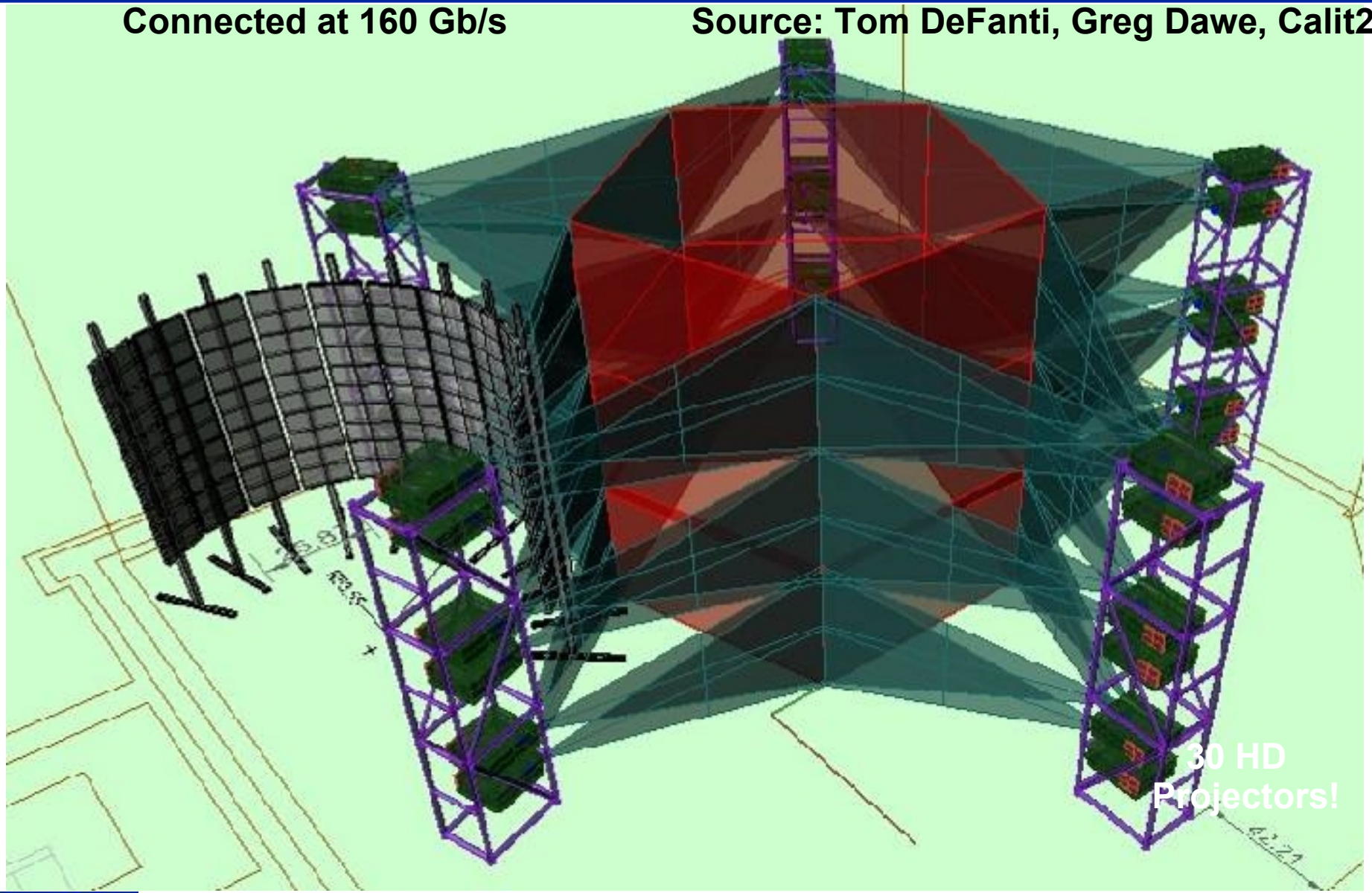
it2

EVL's SAGE Global Visualcasting to Europe September 2007



3D OptIPortals: Calit2 StarCAVE and Varrier Alpha Tests of Telepresence "Holodecks"

Connected at 160 Gb/s **Source: Tom DeFanti, Greg Dawe, Calit2**



UCSD
UCIrvine

60 GB Texture Memory,
Renders Images 3,200 Times the Speed of Single PC

it²

StarCAVE Panoramas









How Do You Get From Your Lab to the National LambdaRail?

"Research is being stalled by 'information overload,' Mr. Bement said, because data from digital instruments are piling up far faster than researchers can study. In particular, he said, campus networks need to be improved. High-speed data lines crossing the nation are the equivalent of six-lane superhighways, he said. But networks at colleges and universities are not so capable. "Those massive conduits are reduced to two-lane roads at most college and university campuses," he said. Improving cyberinfrastructure, he said, "will transform the capabilities of campus-based scientists."

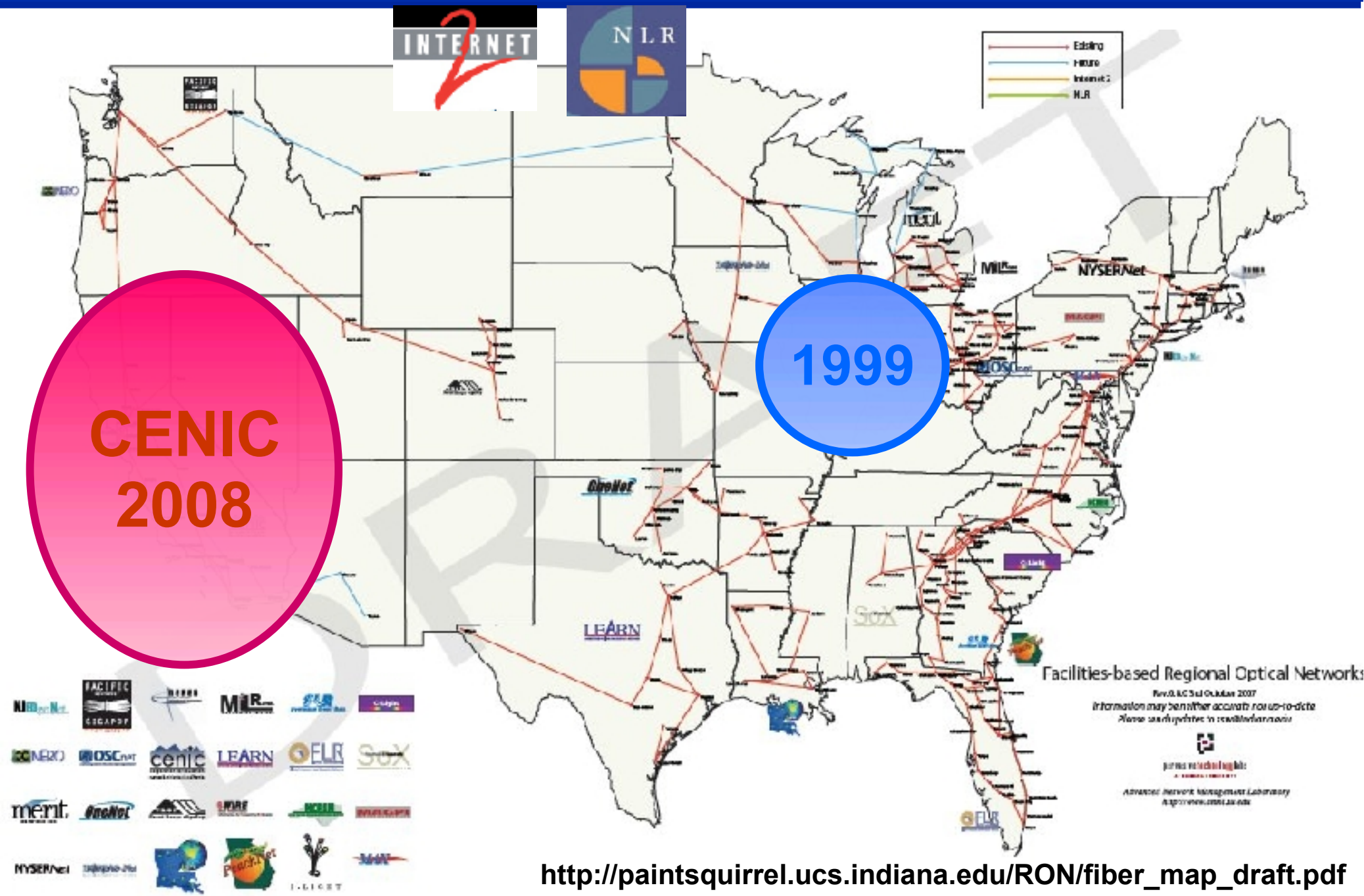
-- Arden Bement, the director of the National Science Foundation



www.ctwatch.org

UCSD
UCIrvine

Interconnecting Regional Optical Networks Is Driving Campus Optical Infrastructure Deployment



California (CENIC) Network Directions

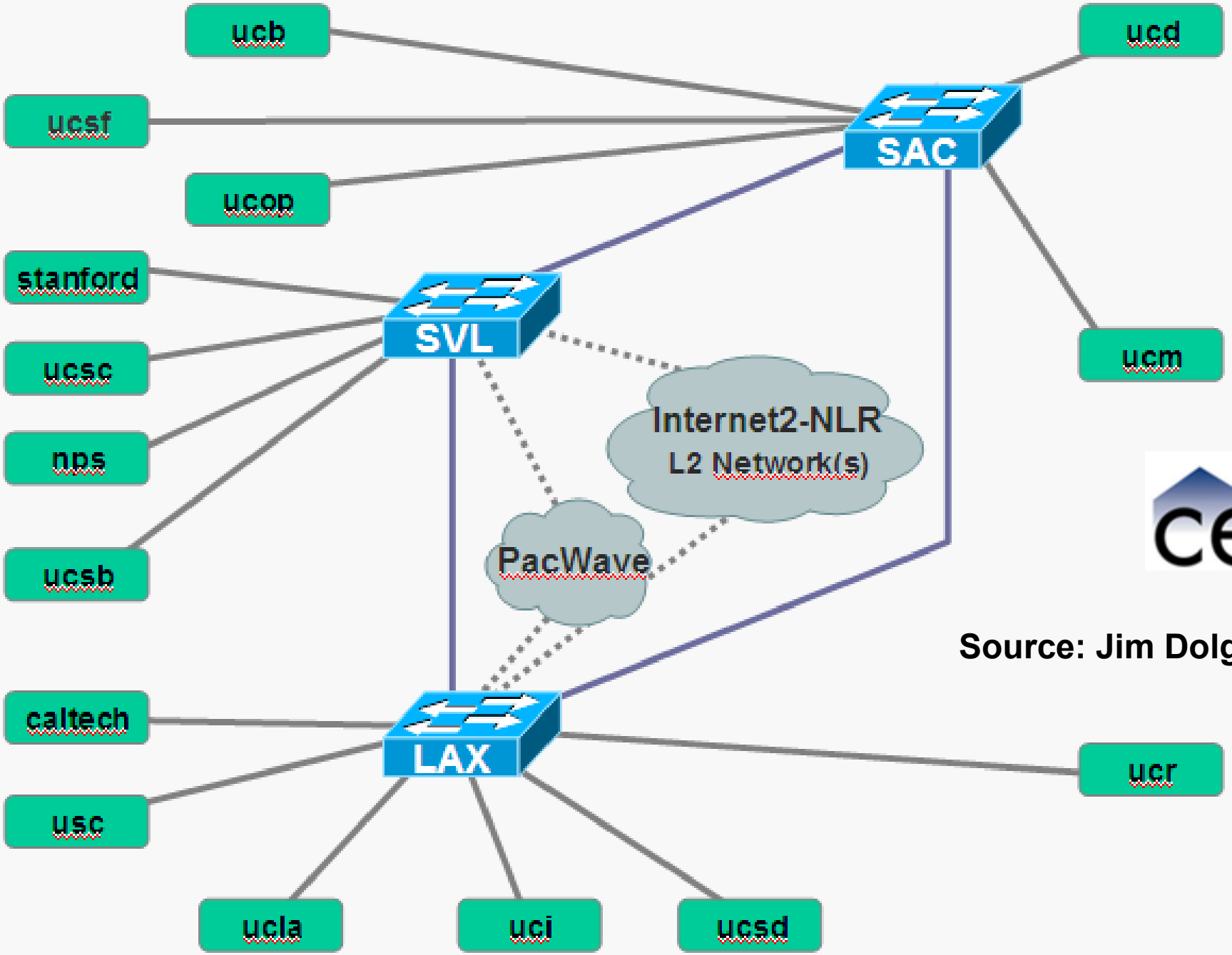
- More Bandwidth to Research University Campuses
  - One or Two 10GE Connections to Every Campus
- More Bandwidth on the Backbone
  - 40Gbps Or 100Gbps
- Support for New Protocols and Features
  - IPv6 Multicast
  - Jumbo Frames: 9000 (or More) Bytes
- "Hybrid Network" Design, Incorporating Traditional Routed IP Service and the New Frame and Optical Circuit Services:
  - "HPRng-L3" = Routed IP Network
  - "HPRng-L2" = Switched Ethernet Network
  - "HPRng-L1" = Switched Optical Network

} CalREN-XD

UCSD UCIrvine

Source: Jim Dolgonas, CENIC

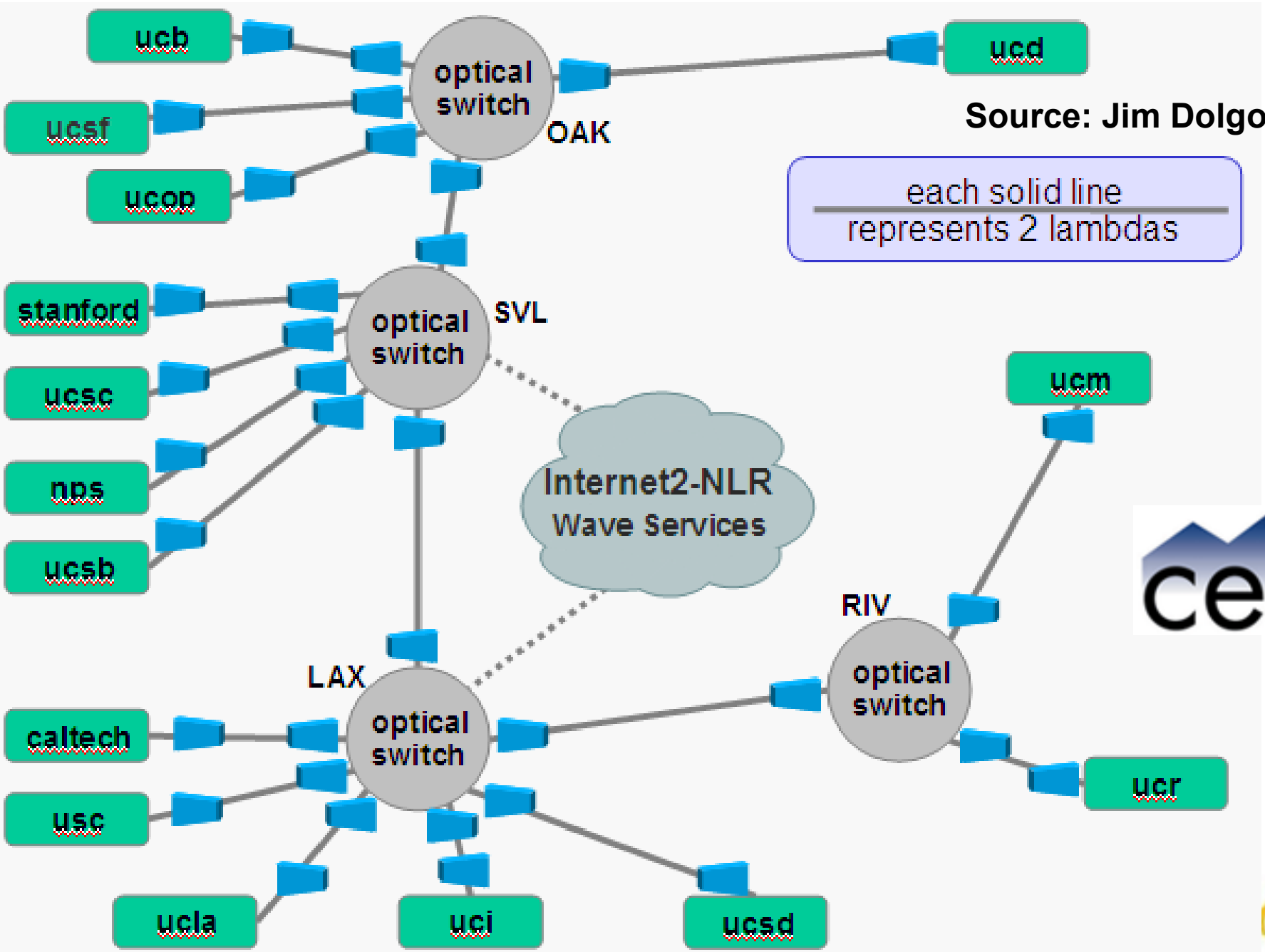
CENIC Switched Ethernet Network
HPRng-L2 Design



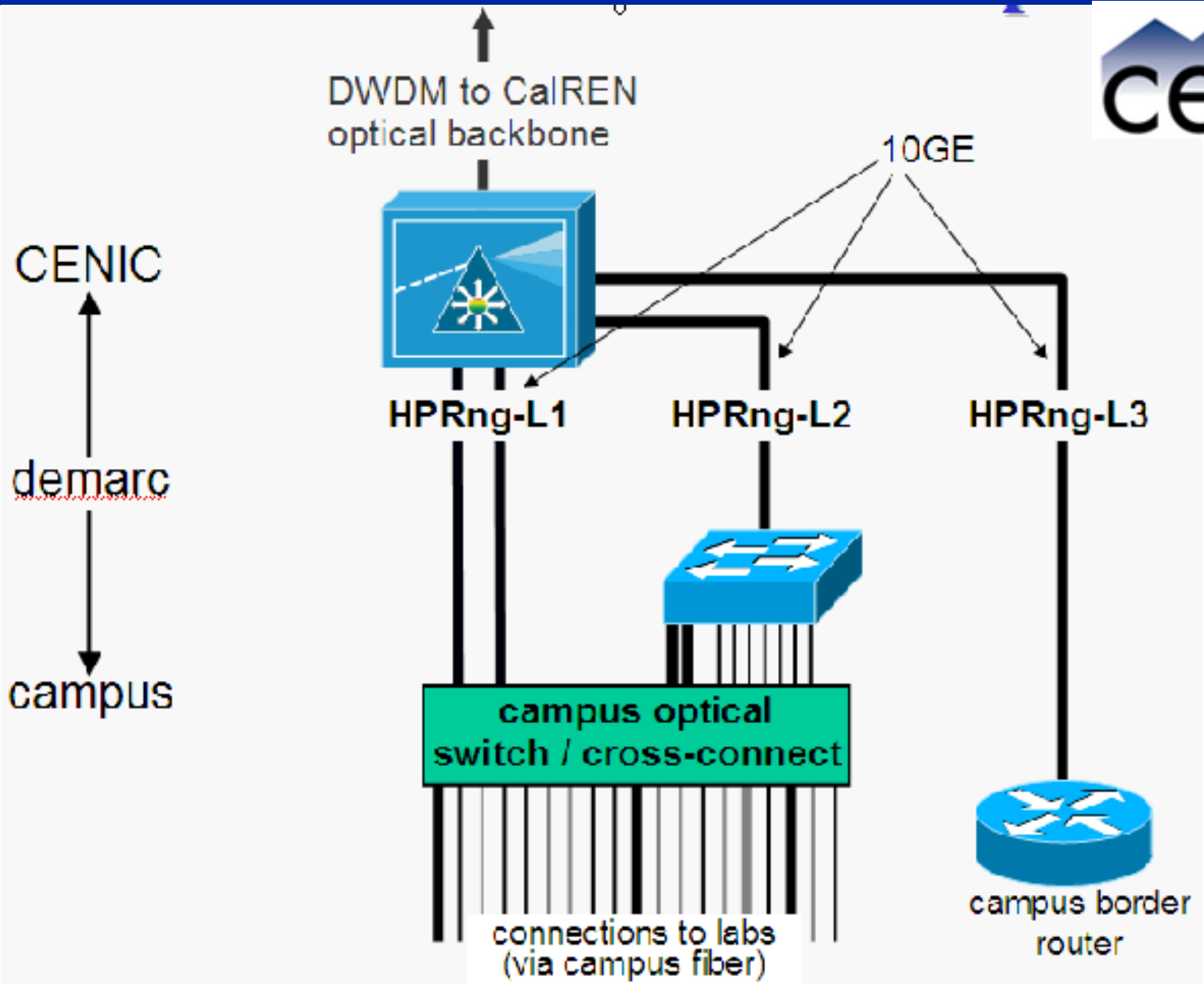
cenic

Source: Jim Dolgonas, CENIC

CENIC Switched Optical Network
HPRng-L1 design

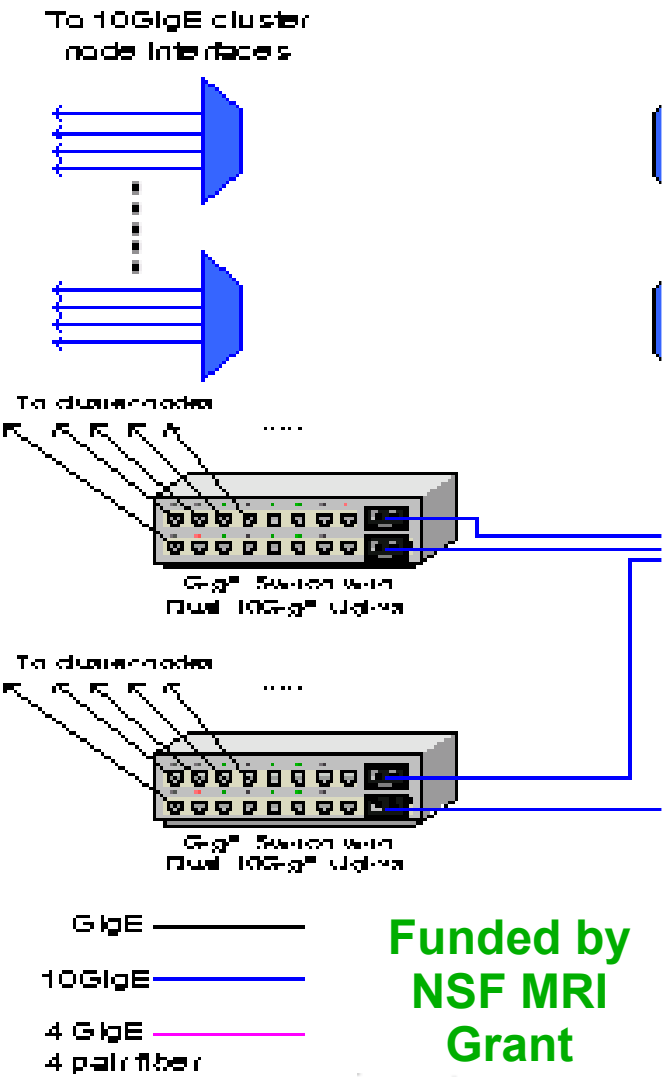


Campus Preparations Needed to Accept CENIC CalREN Handoff to Campus



Source: Jim Dolgonas, CENIC

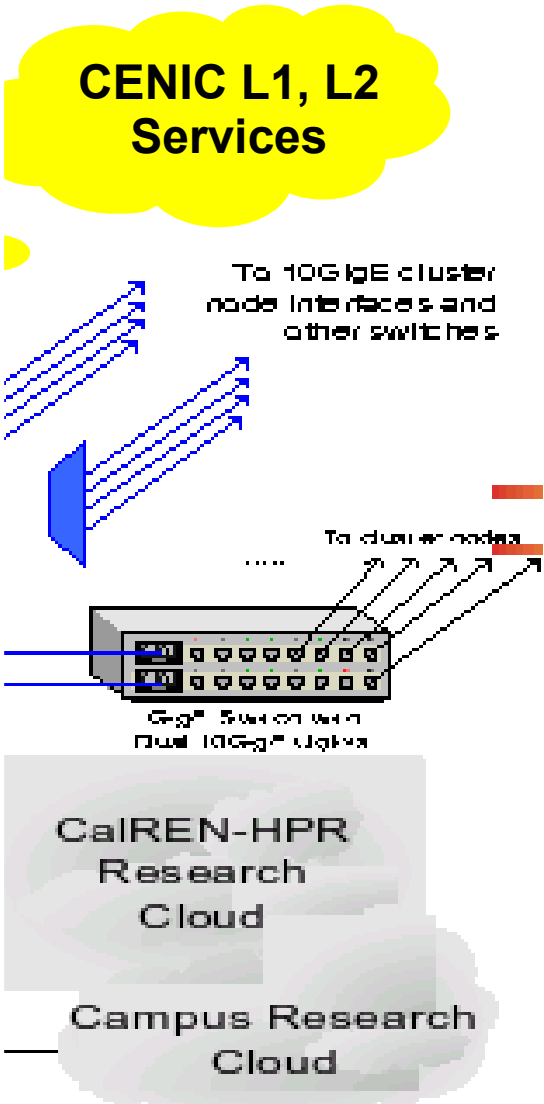
Current UCSD Experimental Optical Core: Ready to Couple to CENIC L1, L2, L3 Services



Goals by 2008:

- >= 50 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 will be a Hybrid Combination of:
Packet, Lambda, Circuit -- OOO and Packet Switches Already in Place

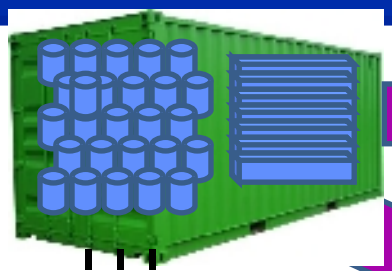


Funded by NSF MRI Grant

OptIPuter Border Router

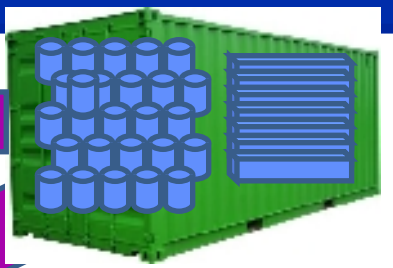
Source: Phil Papadopoulos, SDSC/Calit2 (Quartzite PI, OptIPuter co-PI)

Planned UCSD Production Campus Cyberinfrastructure Supporting Data Intensive Biomedical Research



Active Data Replication

N x 10 Gbit



Eco-Friendly Storage and Compute

N x 10 Gbit

N x 10 Gbit



10 Gigabit L2/L3 Switch

"Network in a box"
- > 200 Connections
- DWDM or Gray Optics

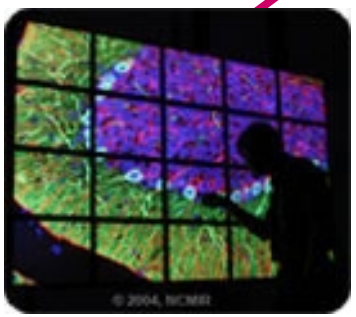
On-Demand Physical Connections

Wide-Area 10G
- CENIC/HPRng
- NLR Cavewave
- I2 NewNet
- Cinegrid
- …

Single 10 Gbit



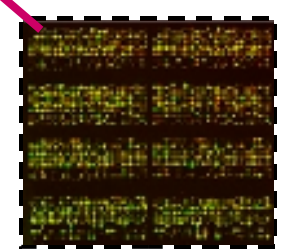
Data Repositories



Microscopes



Mass Spec



Microarray

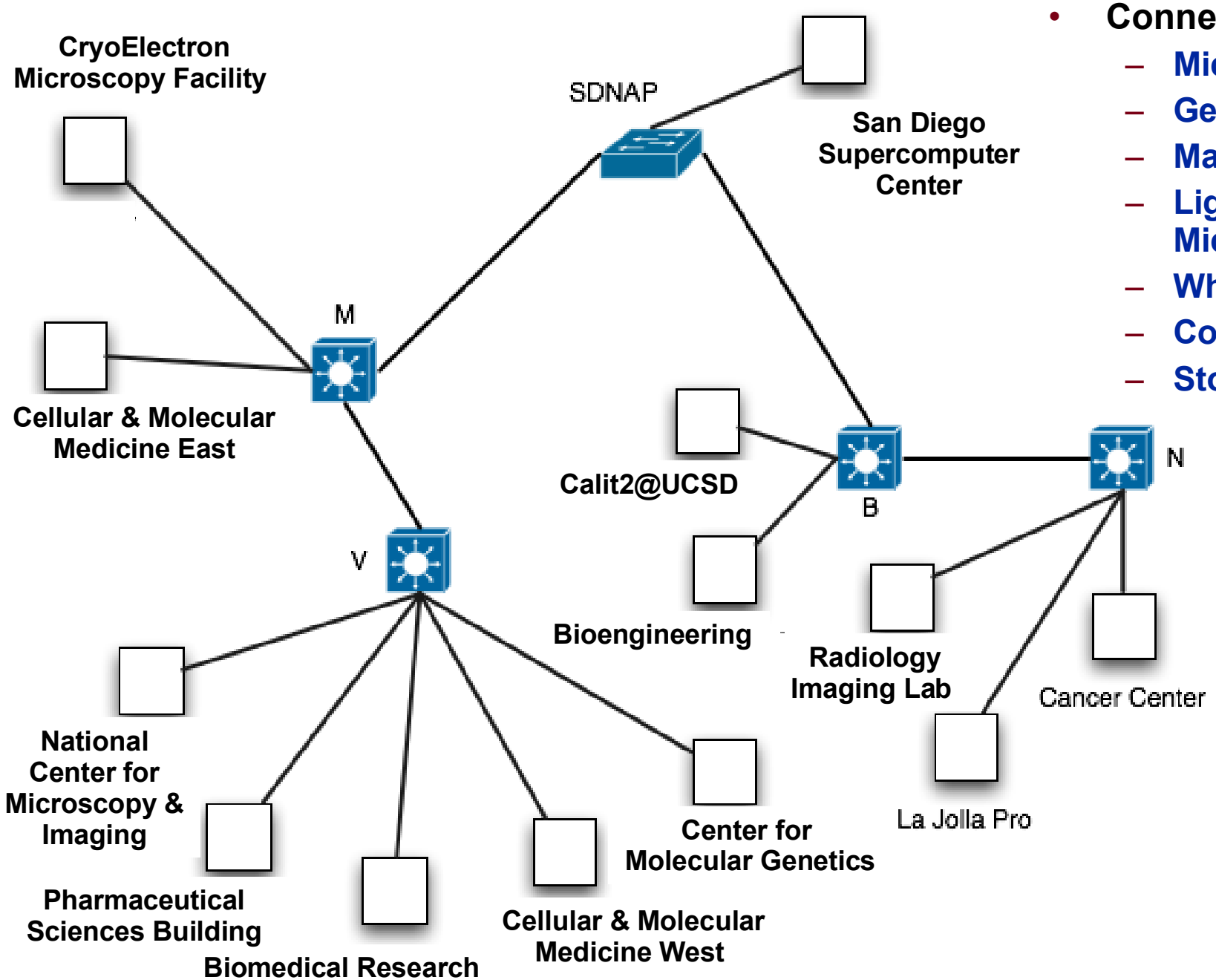
Your Lab Here

UCSD
UCIrvine

Source: Phil Papadopoulos, SDSC/Calit2; Elazar Harel, UCSD

it²

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**

Calit2/SDSC Proposal to Create a UC Cyberinfrastructure of OptIPuter "On-Ramps" to TeraGrid Resources

OptIPuter + CalREN-XD + TeraGrid = "OptiGrid"





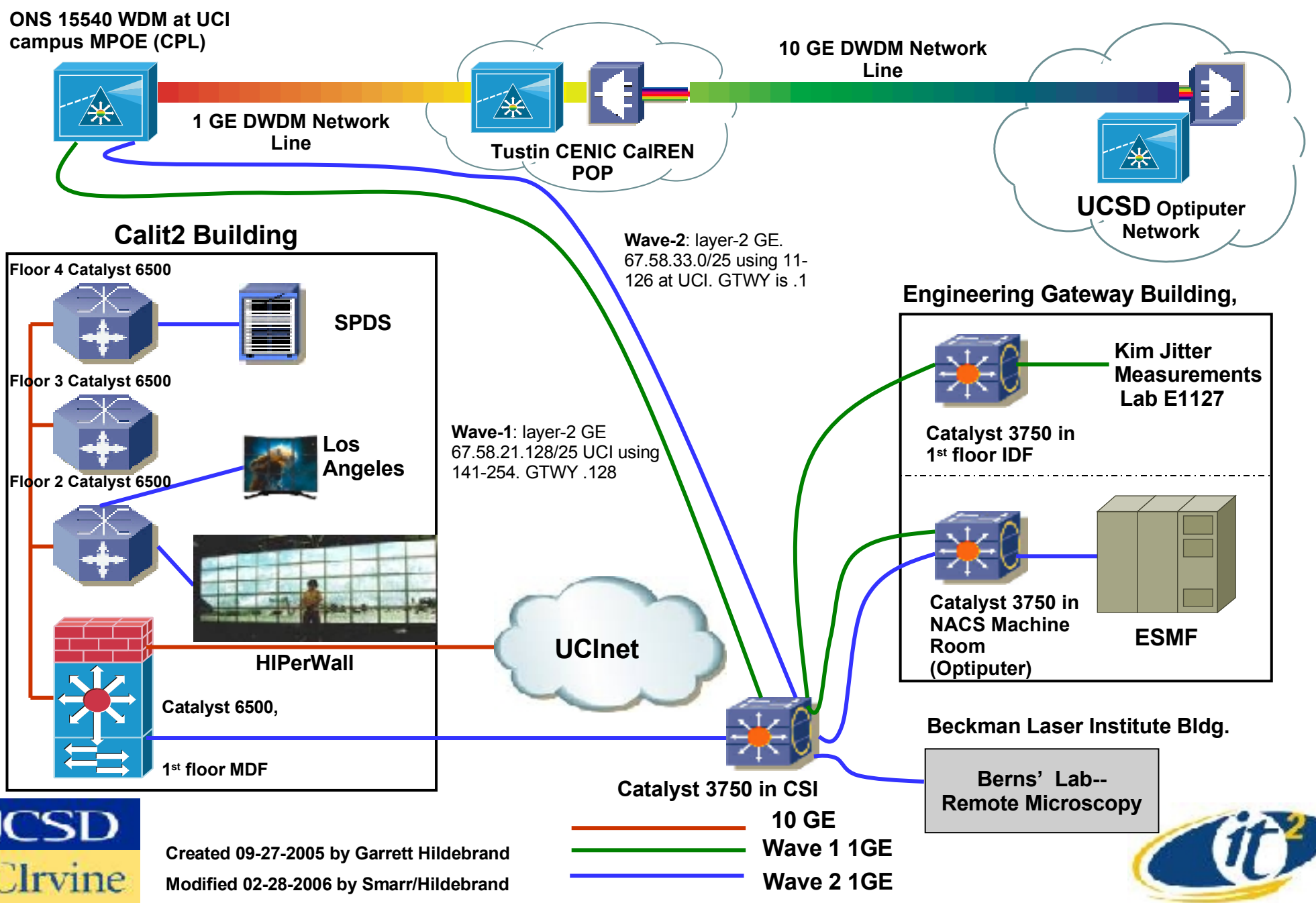
Creating a Critical Mass of End Users on a Secure LambdaGrid

UCSD
UCIrvine

Source: Fran Berman, SDSC , Larry Smarr, Calit2

it2

OptIPuter@UCI is Up and Working

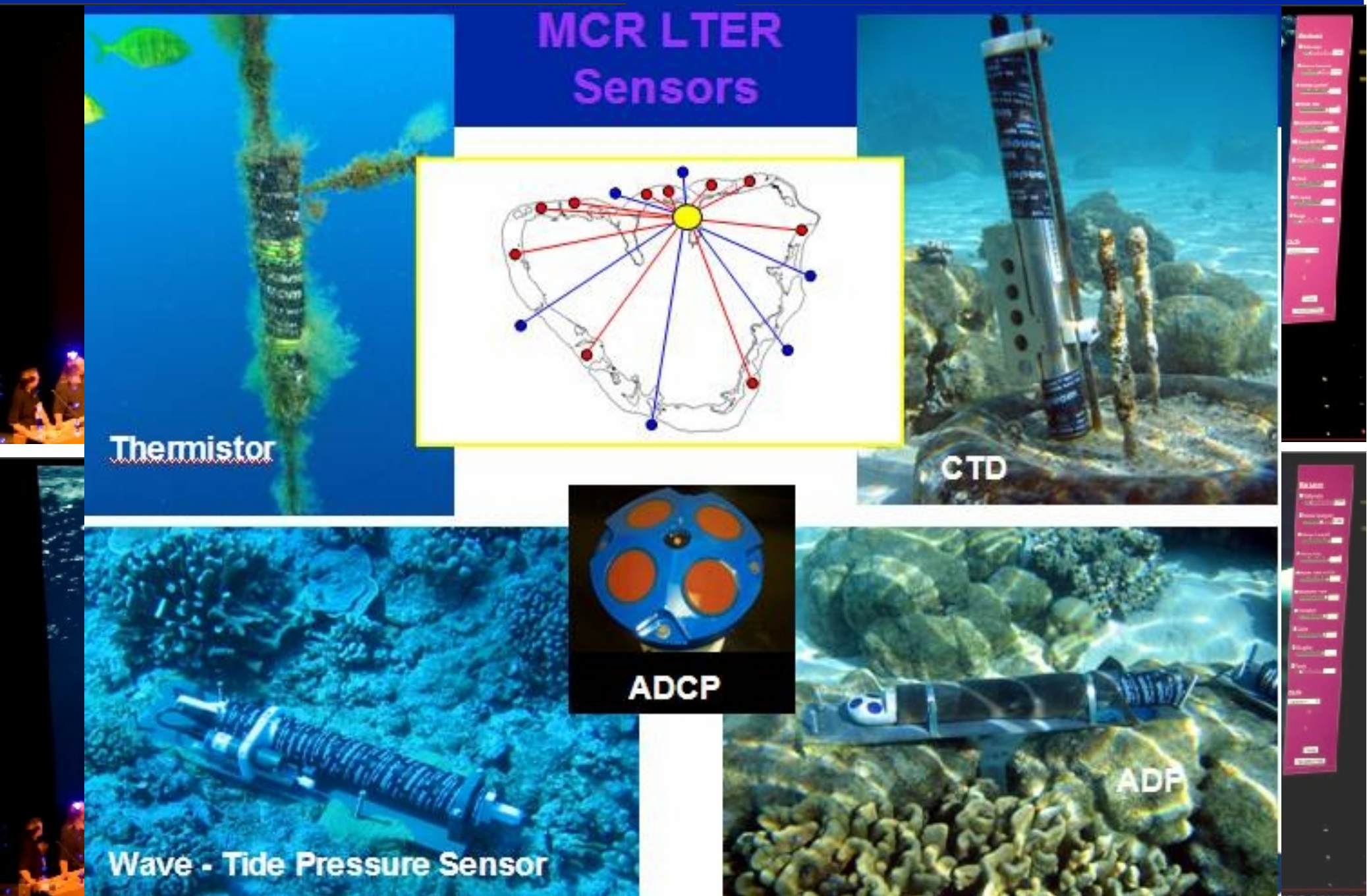


UCSD UCIrvine

Created 09-27-2005 by Garrett Hildebrand

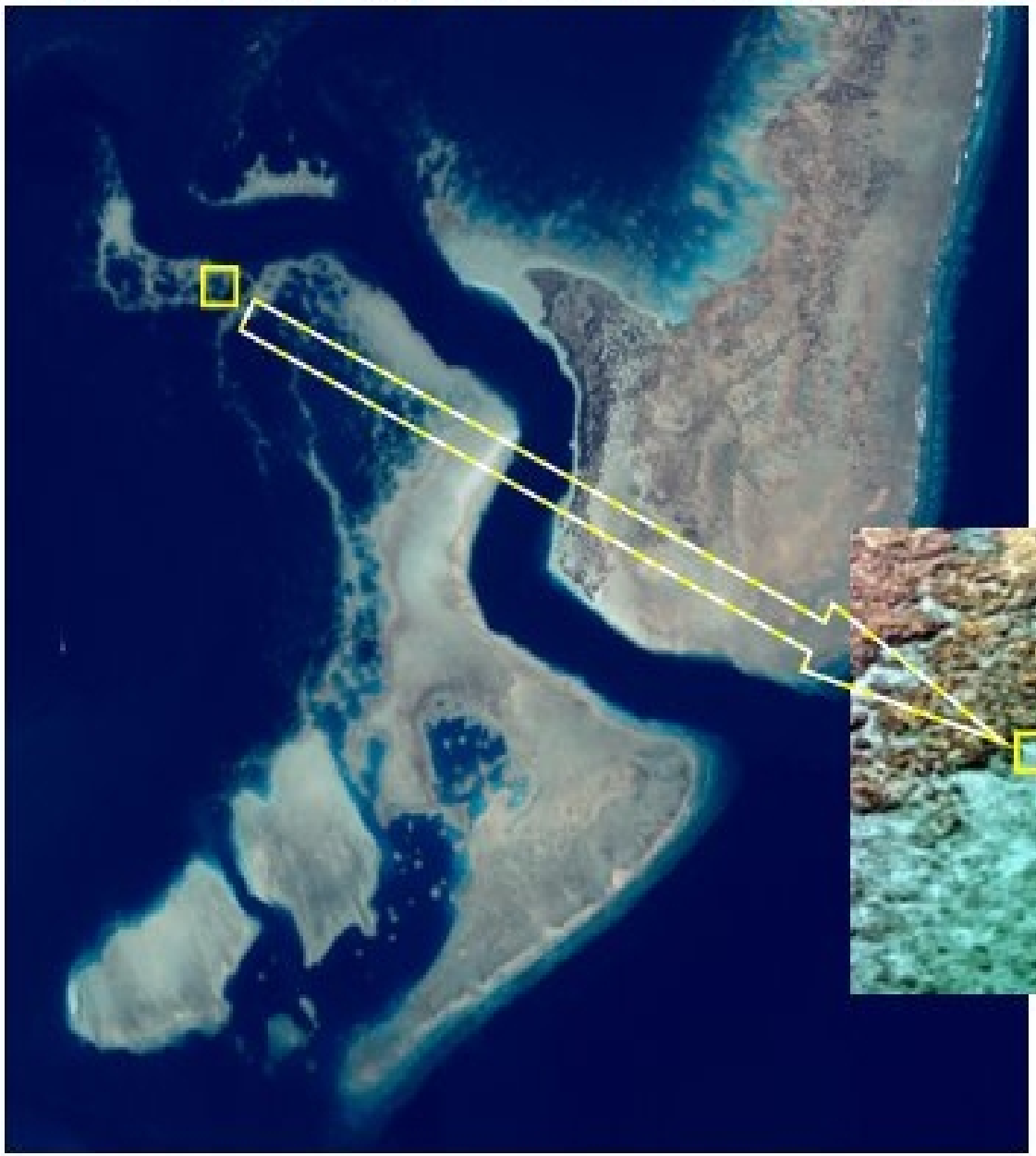
Modified 02-28-2006 by Smarr/Hildebrand

Creating a Digital Moorea
Calit2 Collaboration with UC Gump Station (UCB, UCSB)

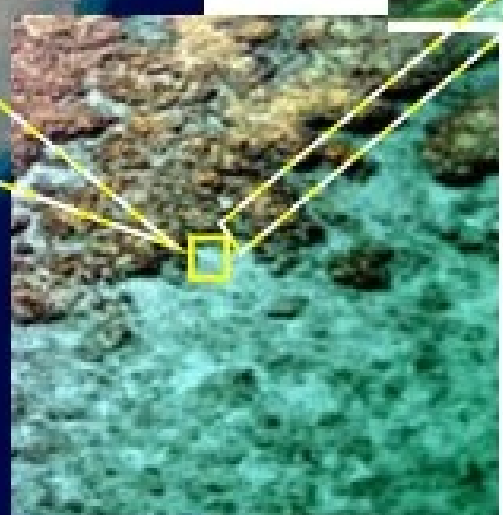


Calit2 ReefBot Design for Digital Reef Mapping

Explore Every Square Meter of the Moorea Coral Reef from a Remote Desktop



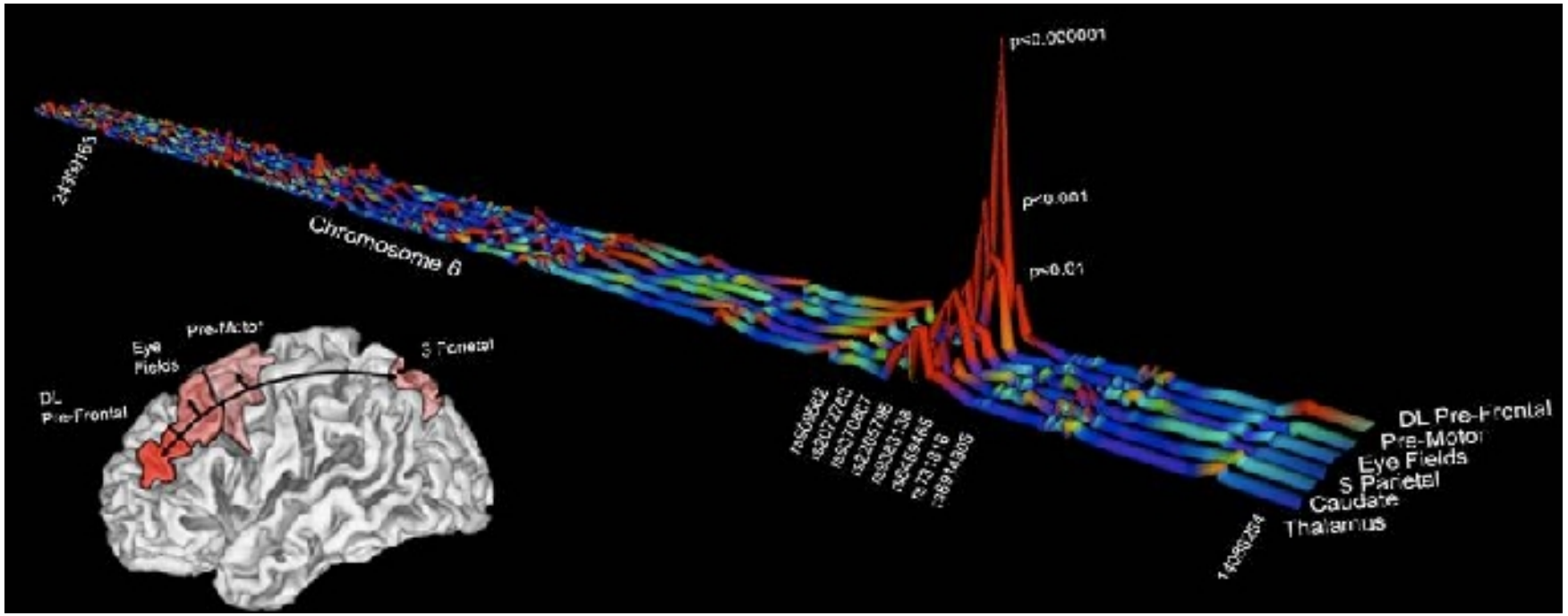




Use The Reef GIS
to Zoom Into
Any Portion
of the Shallow Reef
to View Details Down
to ~ Centimeter

UC Systems Collaborating in Studies Integrating Brain Imaging and Genetics

- Institutional Research Program: Imaging Genetics
- UCI, UCLA, UCSD, UCSF, UCDavis
- Transdisciplinary: molecular to clinical
- Collaborate with CalIT2 for infrastructure



Source: Steven Potkin, UCI