“21st Century e-Knowledge Requires a High Performance e-Infrastructure”

Keynote Presentation

40-year anniversary Celebration of SARA

Amsterdam, Netherlands

December 9, 2011

Dr. Larry Smarr

Director, California Institute for Telecommunications and Information Technology

Harry E. Gruber Professor,

Dept. of Computer Science and Engineering

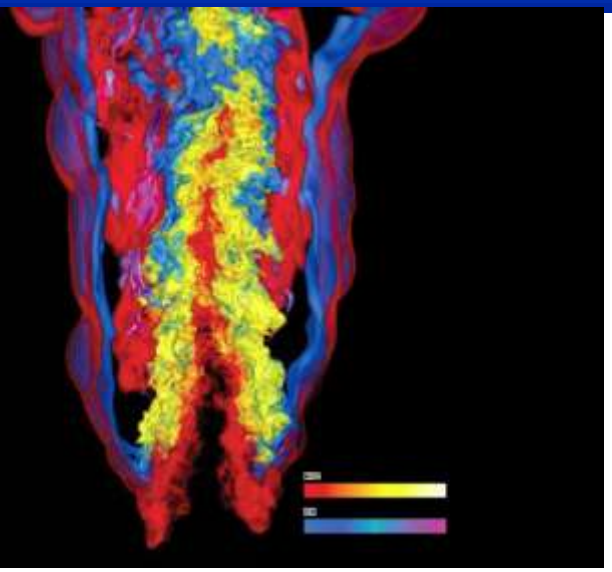
Jacobs School of Engineering, UCSD

http://lsmarr.calit2.net

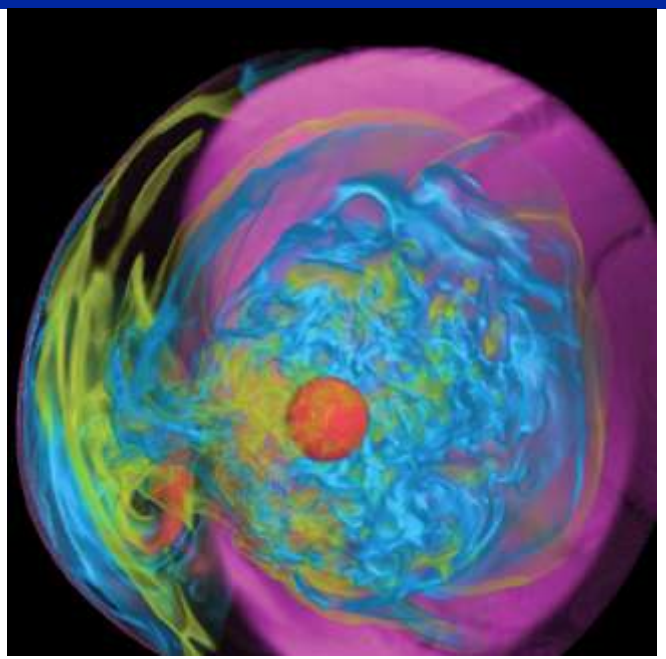
Abstract

Over the next decade, advances in high performance computing will usher in an era of ultra-realistic scientific and engineering simulation-- in fields as varied as climate sciences, ocean observatories, radio astronomy, cosmology, biology, and medicine. Simultaneously, distributed scientific instruments, high-resolution video streaming, and the global computational and storage cloud all generate terabytes to petabytes of data. Over the last decade, the U.S. National Science Foundation funded the OptIPuter project to research how user-controlled 10Gbps dedicated lightpaths (or "lambdas") could provide direct access to global data repositories, scientific instruments, and computational resources from "OptIPortals," PC clusters which provide scalable visualization, computing, and storage in the user's campus laboratory. All of these components can be integrated into a seamless high performance e-infrastructure required to support a next generation e-knowledge data-driven society. In the Netherlands SARA and its partner SURFnet has taken a global leadership role in building out and supporting such a future-oriented e-infrastructure, enabling powerful computing, data processing, networking, and visualization e-science services, necessary for the pursuit of solutions to an increasingly difficult set of scientific and societal challenges

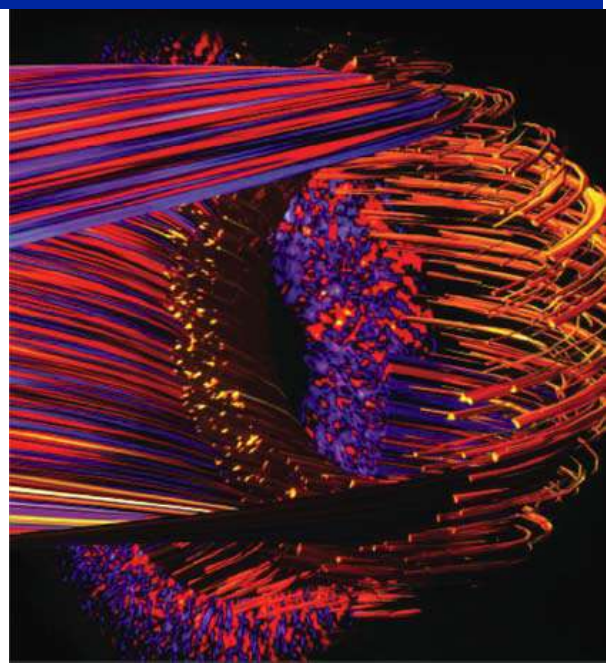
Leading Edge Applications of Petascale Computers Today Are Critical for Basic Research and Practical Apps



Flames

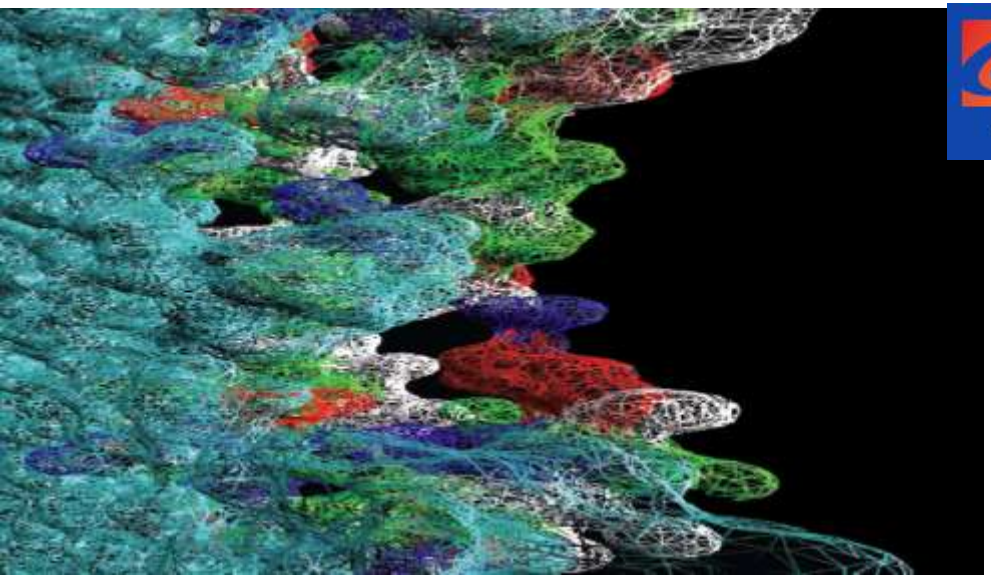


Supernova

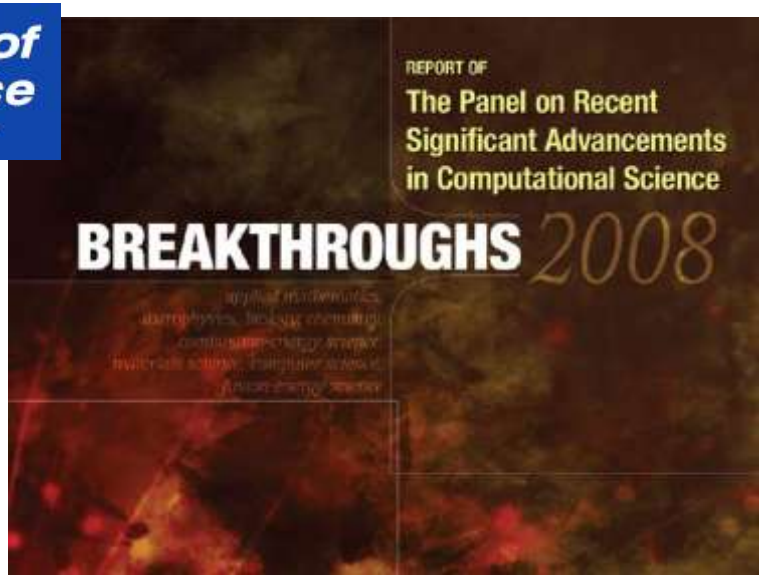


Fusion

Parkinson's

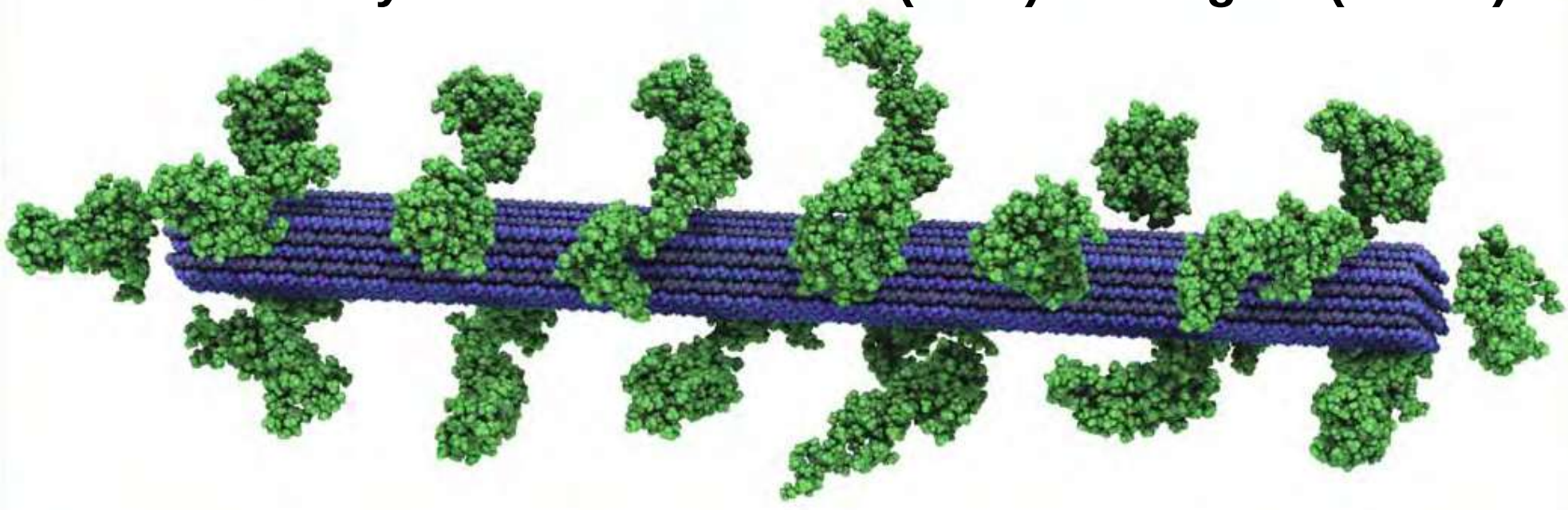


Office of Science
U.S. DEPARTMENT OF ENERGY



Supercomputing the Future of Cellulosic Ethanol Renewable Fuels

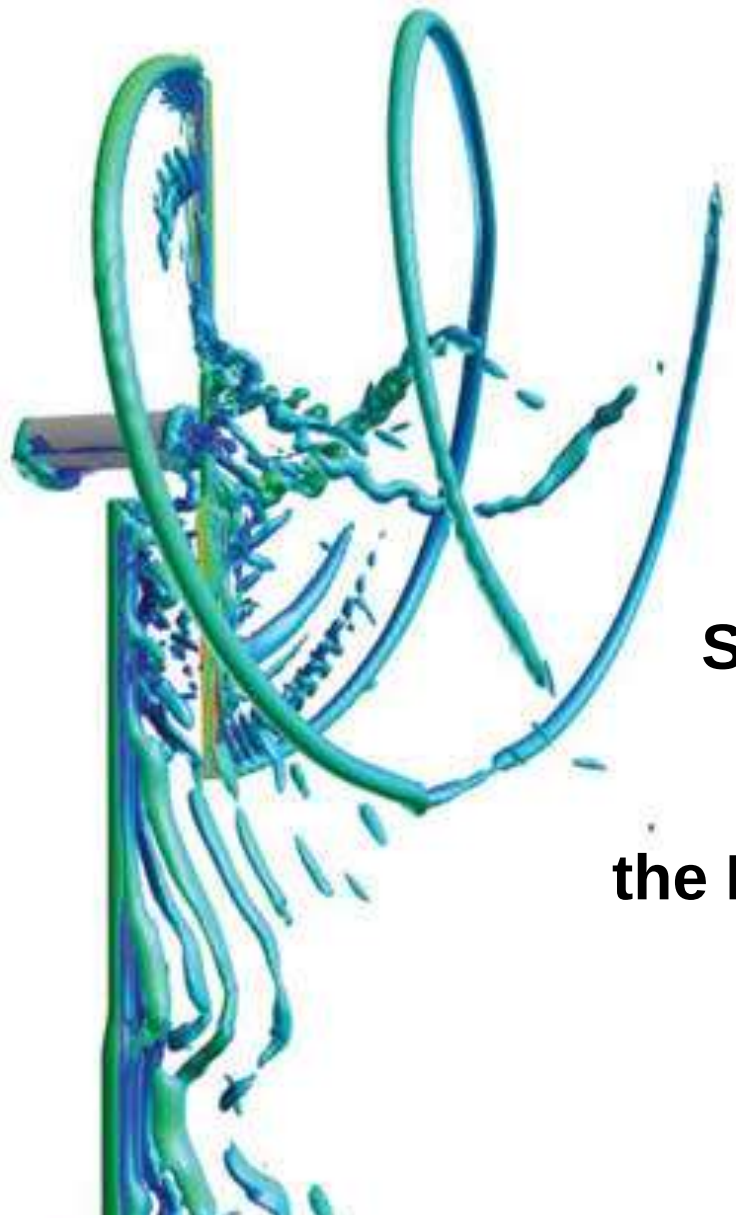
Atomic-Detail Model of the Lignocellulose of Softwoods. The model was built by Loukas Petridis of the ORNL CMB Molecular Dynamics of Cellulose (Blue) and Lignin (Green)

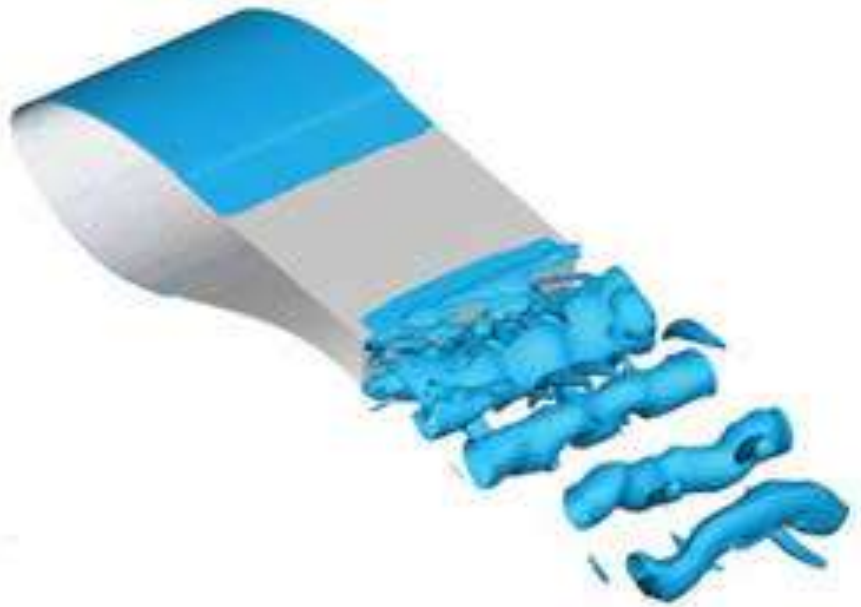


Computing the Lignin Force Field & Combining With the Known Cellulose Force Field Enables Full Simulations of Lignocellulosic Biomass

www.scidacreview.org/0905/pdf/biofuel.pdf

Supercomputers are Designing Quieter Wind Turbines





Simulation of an Infinite-Span “Flatback" Wind Turbine Airfoil Designed by the Netherlands Delft University of Technology Using NASA's FUN3D CFD Code Modified by Georgia Tech to Include a Hybrid RANS/LES Turbulence model

NCSA

Georgia Institute of Technology Professor Marilyn Smith

www.ncsa.illinois.edu/News/Stories/Windturbines/

it²

Increasing the Efficiency of Tractor Trailers Using Supercomputers

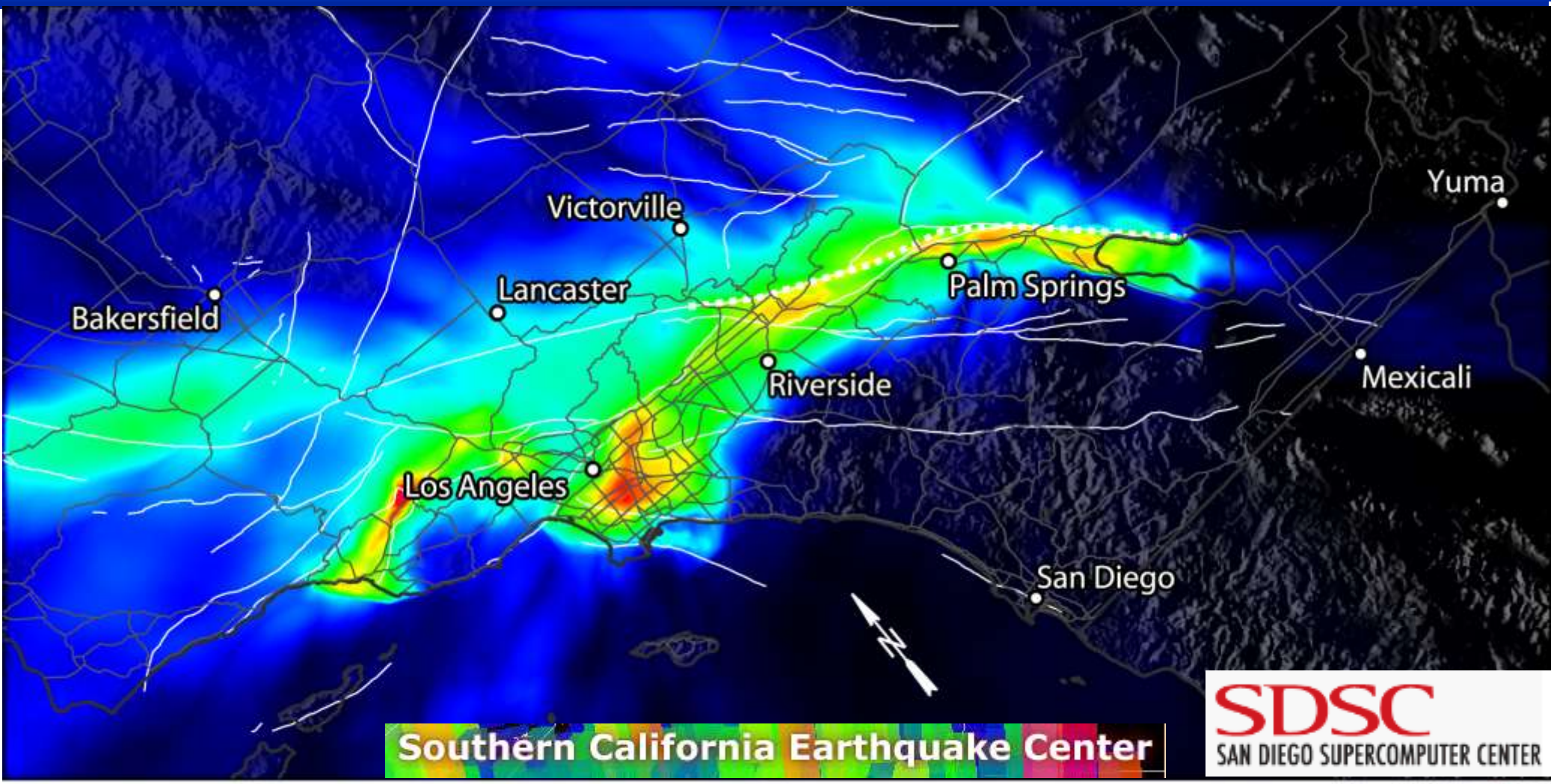
Oak Ridge Leadership Computing Facility & the Viz Team
(Dave Pugmire, Mike Matheson, and Jamison Daniel)





BMI Corporation, an engineering services firm has teamed up with ORNL, NASA, and several BMI corporate partners with large trucking fleets

Realistic Southern California Earthquake Supercomputer Simulations



SDSC Vis-Services

100 km

Peak Spectral Acceleration at 10 sec

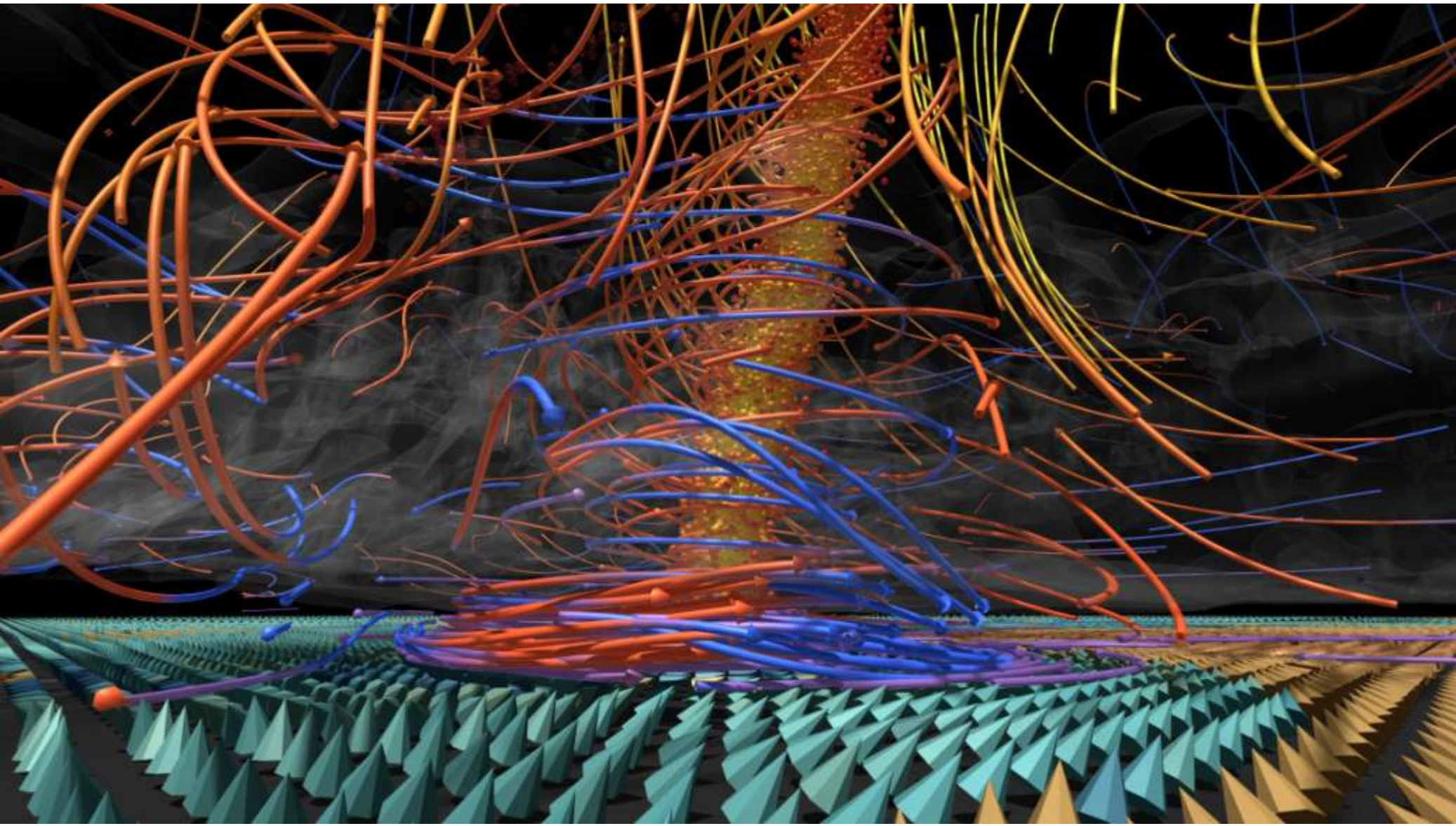
SCEC an NSF+USGS center

Magnitude 7.7 Earthquake

http://visservices.sdsc.edu/projects/scec/terashake/2.1/

it²

Tornadogenesis From Severe Thunderstorms Simulated by Supercomputer

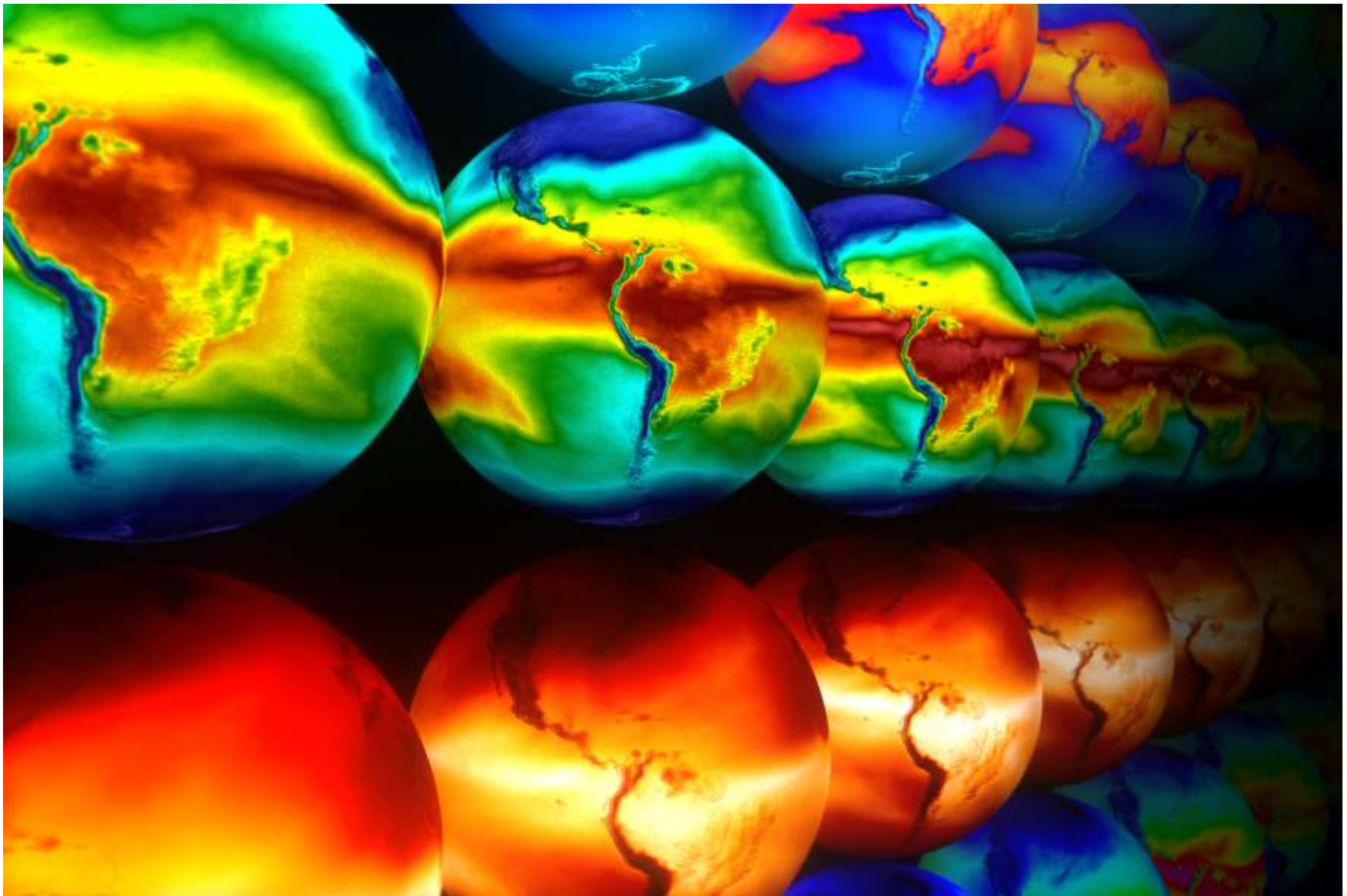


NCSA

Source: Donna Cox, Robert Patterson, Bob Wilhelmson, NCSA

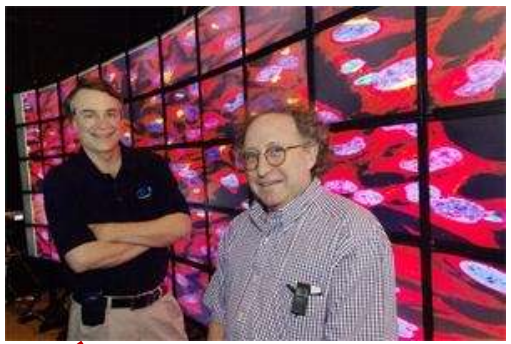
it²

Improving Simulation of the Distribution of Water Vapor in the Climate System



ORNL Simulations by Jim Hack; Visualizations by Jamison Daniel
http://users.nccs.gov/~d65/CCSM3/TMQ/TMQ_CCSM3.html

21st Century e-Knowledge Cyberinfrastructure: Built on a 10Gbps "End-to-End" Lightpath Cloud



End User OptIPortal





Local or Remote Instruments



HD/4k Live Video

TeraGrid





10G Lightpaths

sara HPC Cloud Computing

amazon web services

Campus Optical Switch



Data Repositories & Clusters



HD/4k Video Repositories

SDSC

it2

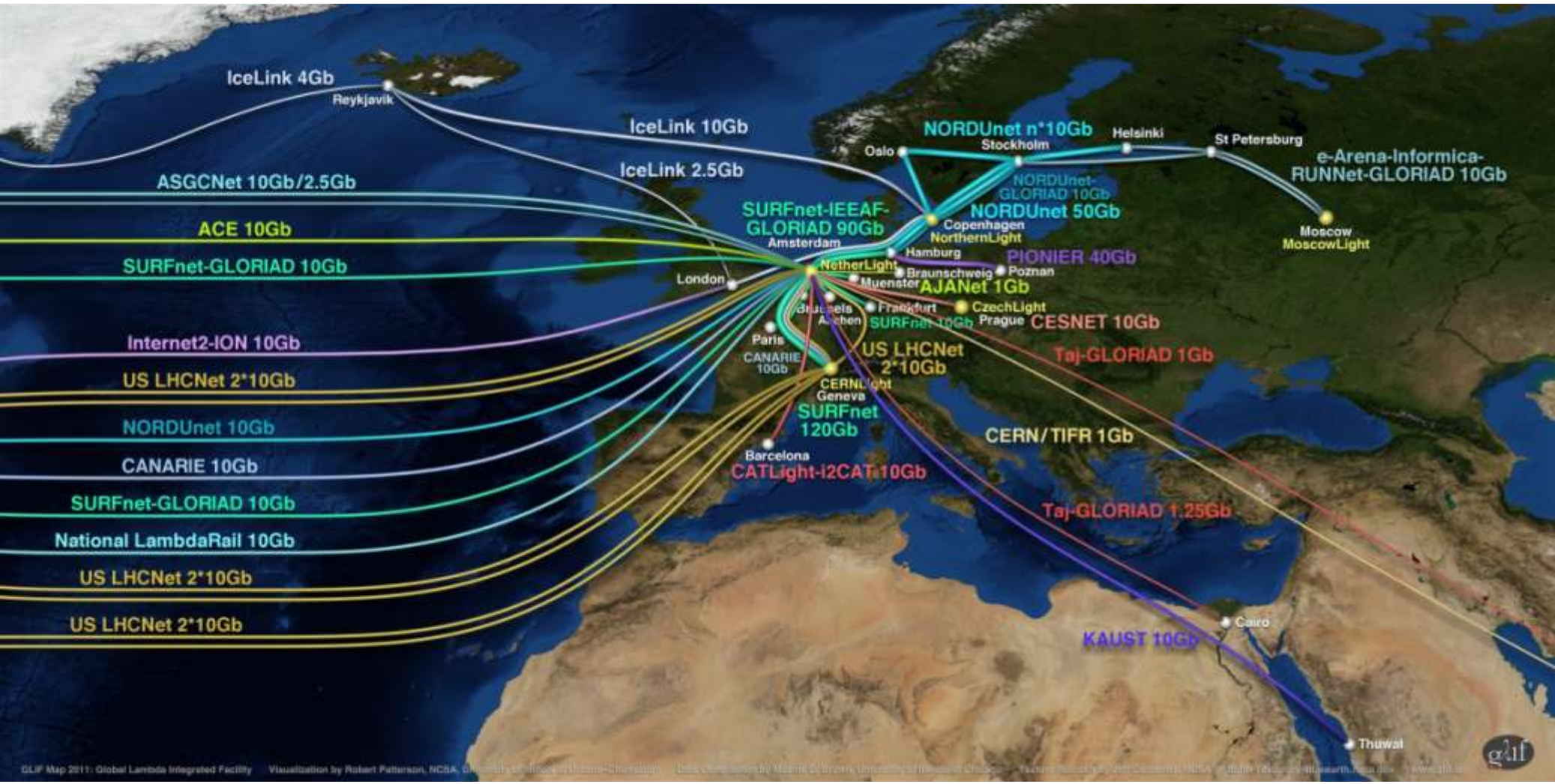
The Global Lambda Integrated Facility--
Creating a Planetary-Scale High Bandwidth Collaboratory

Research Innovation Labs Linked by 10G Dedicated Lambdas



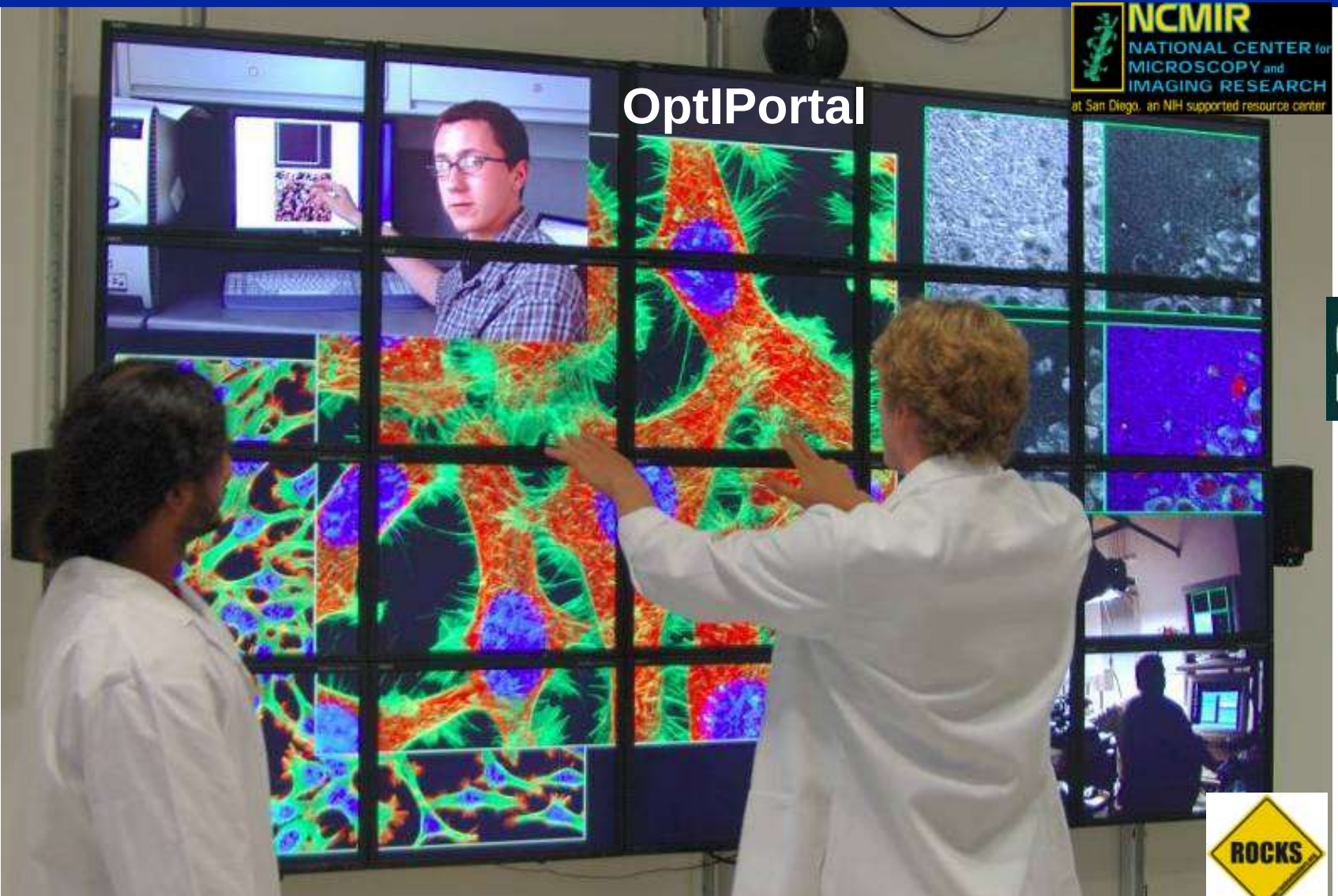
www.glif.is/publications/maps/GLIF_5-11_World_2k.jpg

SURFnet – a SuperNetwork Connecting to the Global Lambda Integrated Facility



Visualization courtesy of
Donna Cox, Bob Patterson, NCSA.

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



NCMIR NATIONAL CENTER for MICROSCOPY and IMAGING RESEARCH at San Diego, an NIH supported resource center

Scalable Adaptive Graphics Environment (SAGE)

evl electronic visualization laboratory

NSF

ROCKS

Picture Source: Mark Ellisman, David Lee, Jason Leigh

OptIPuter

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

it²

The Latest OptIPuter Innovation: Quickly Deployable Nearly Seamless OptIPortables

45 minute setup, 15 minute tear-down with two people (possible with one)



Image From the Calit2 KAUST Lab

The OctIPortable
Calit2/KAUST at SIGGRAPH 2011

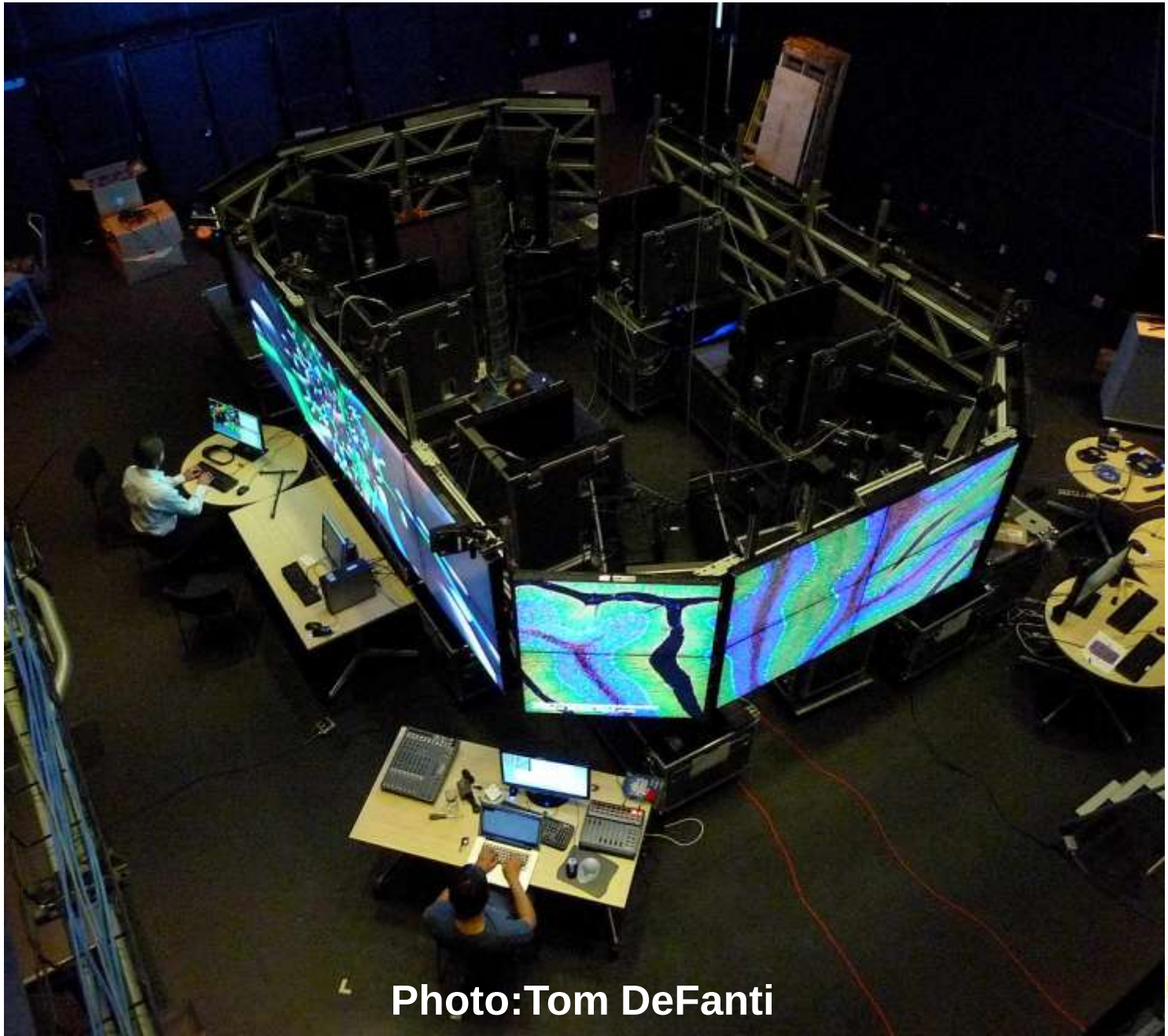


Photo:Tom DeFanti

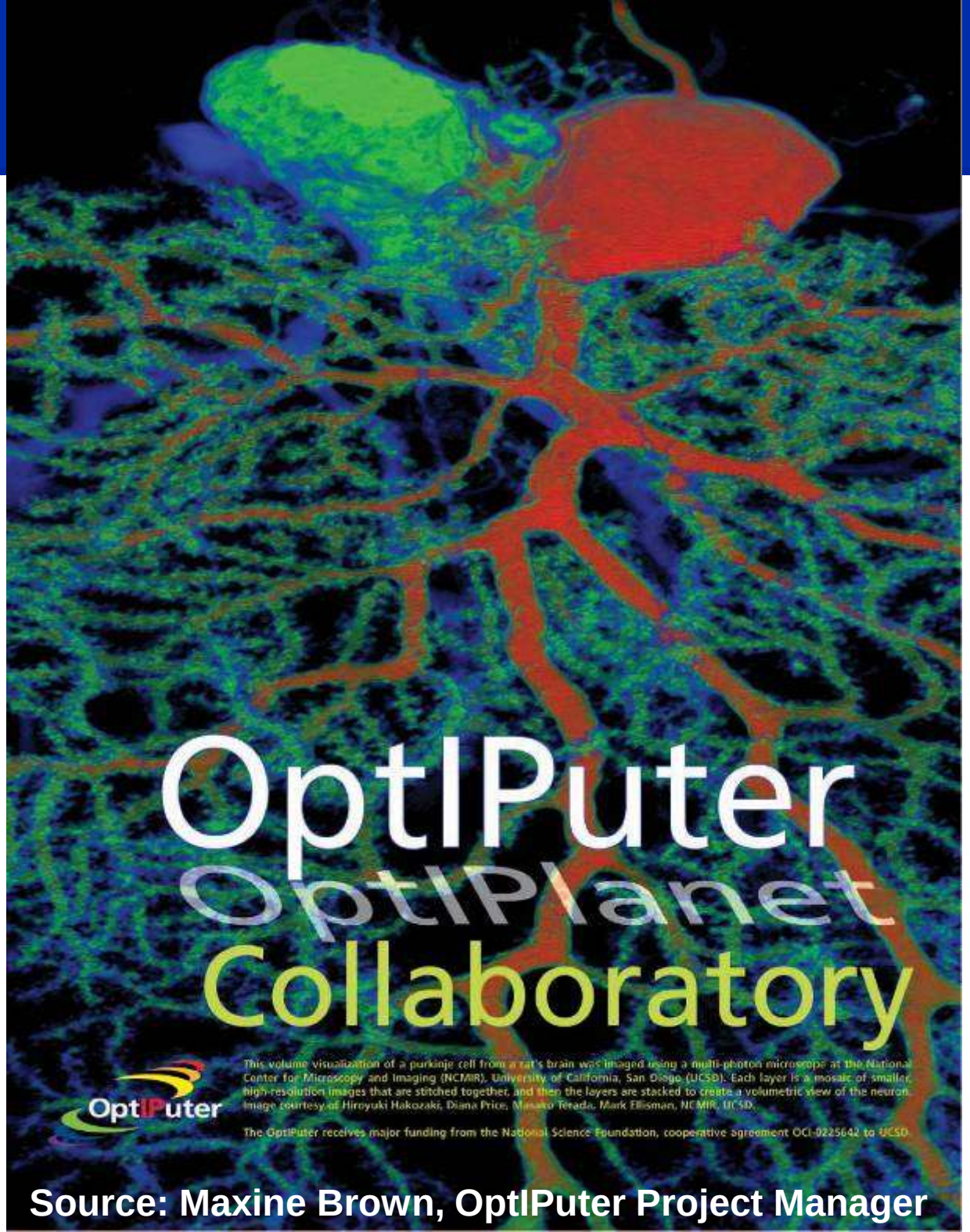
3D Stereo Head Tracked OptIPortal: NexCAVE



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

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

Source: Tom DeFanti, Calit2@UCSD

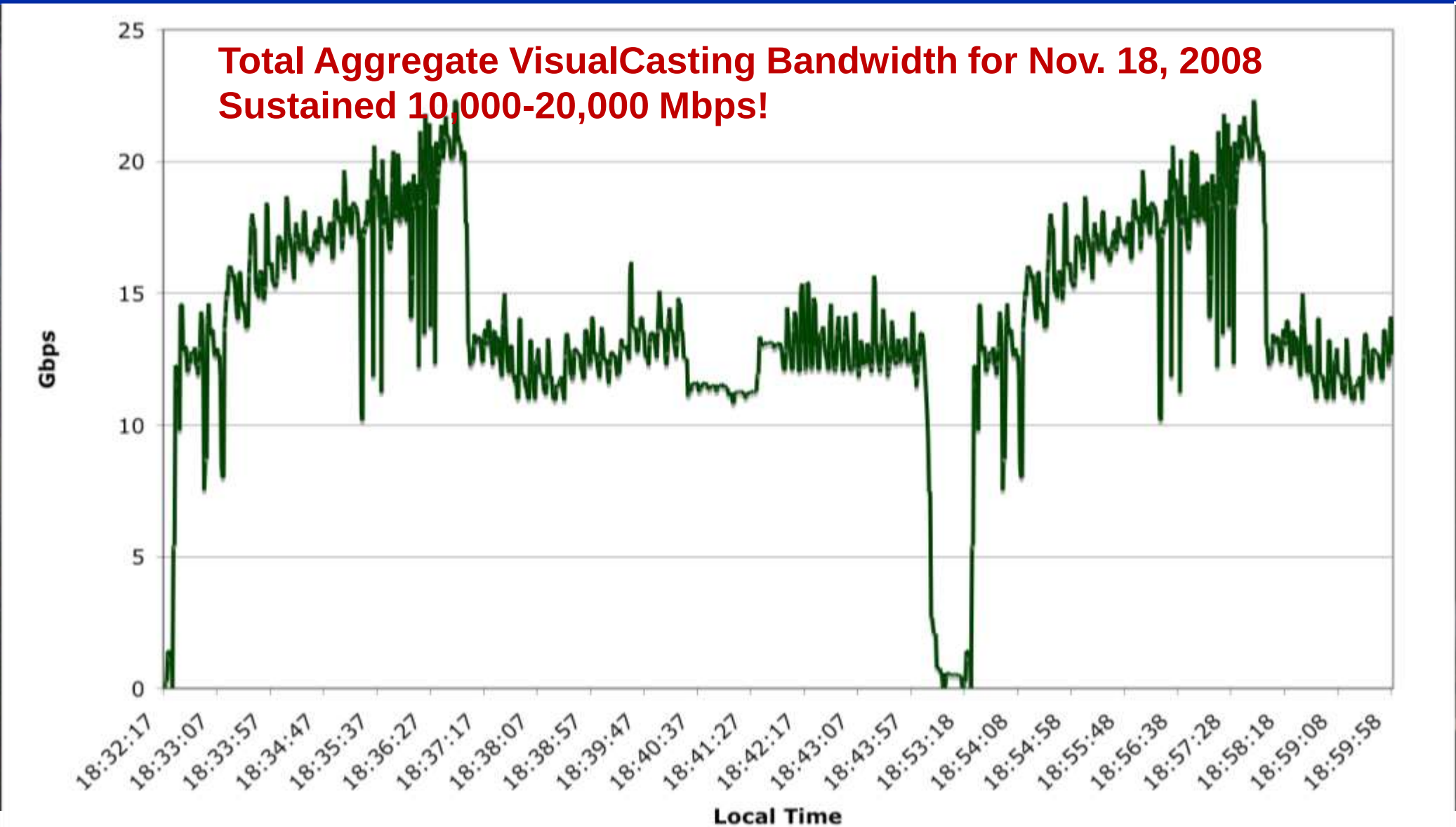


Source: Maxine Brown, OptIPuter Project Manager

Green Initiative:

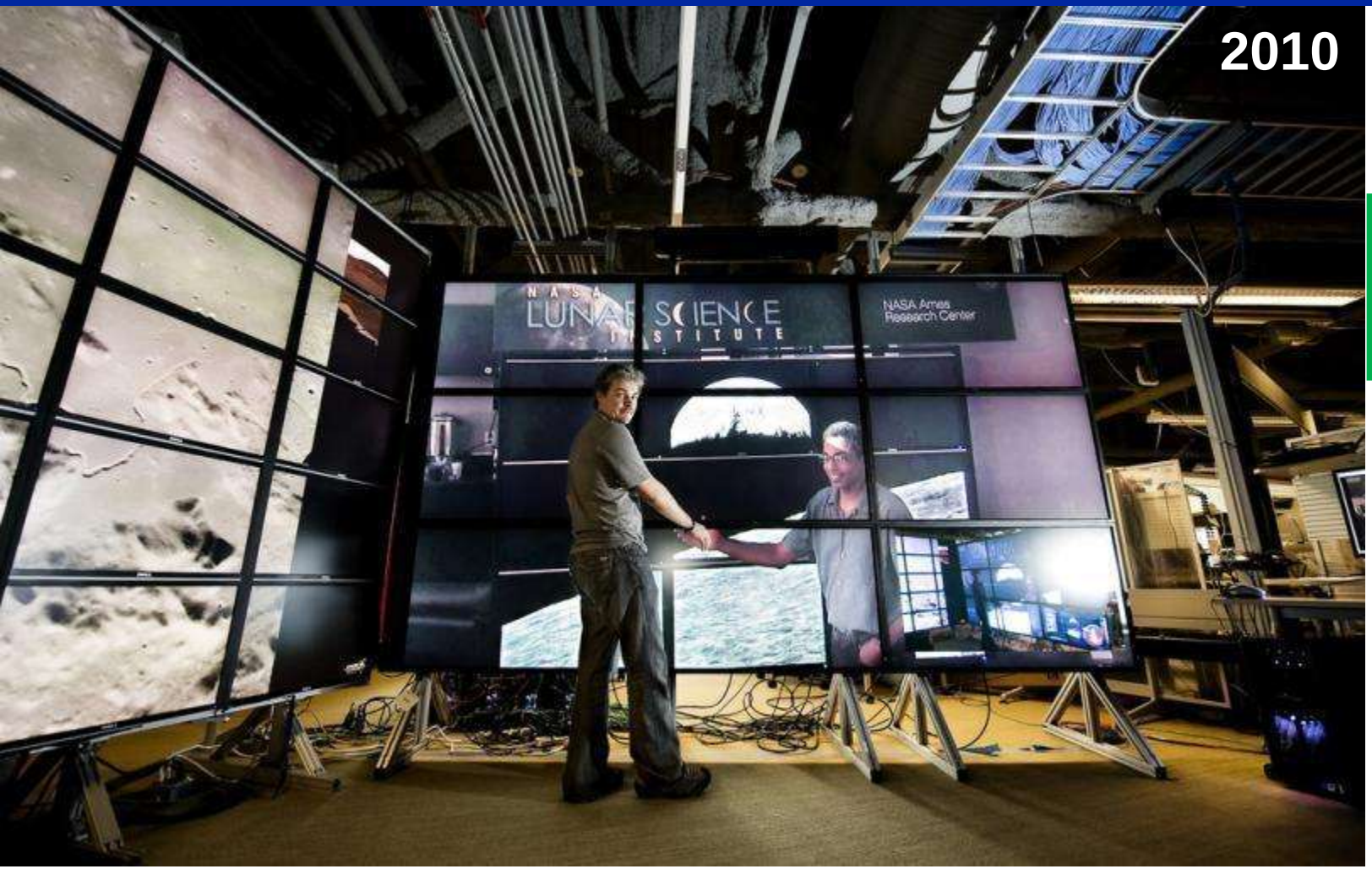
Can Optical Fiber Replace Airline Travel for Continuing Collaborations ?

EVL's SAGE OptIPortal VisualCasting Multi-Site OptIPuter Collaboratory



Source: Jason Leigh, Luc Renambot, EVL, UI Chicago

High Definition Video Connected OptIPortals: Virtual Working Spaces for Data Intensive Research



NASA Supports Two Virtual Institutes

CGLX

LifeSize HD

Calit2@UCSD 10Gbps Link to NASA Ames Lunar Science Institute, Mountain View, CA

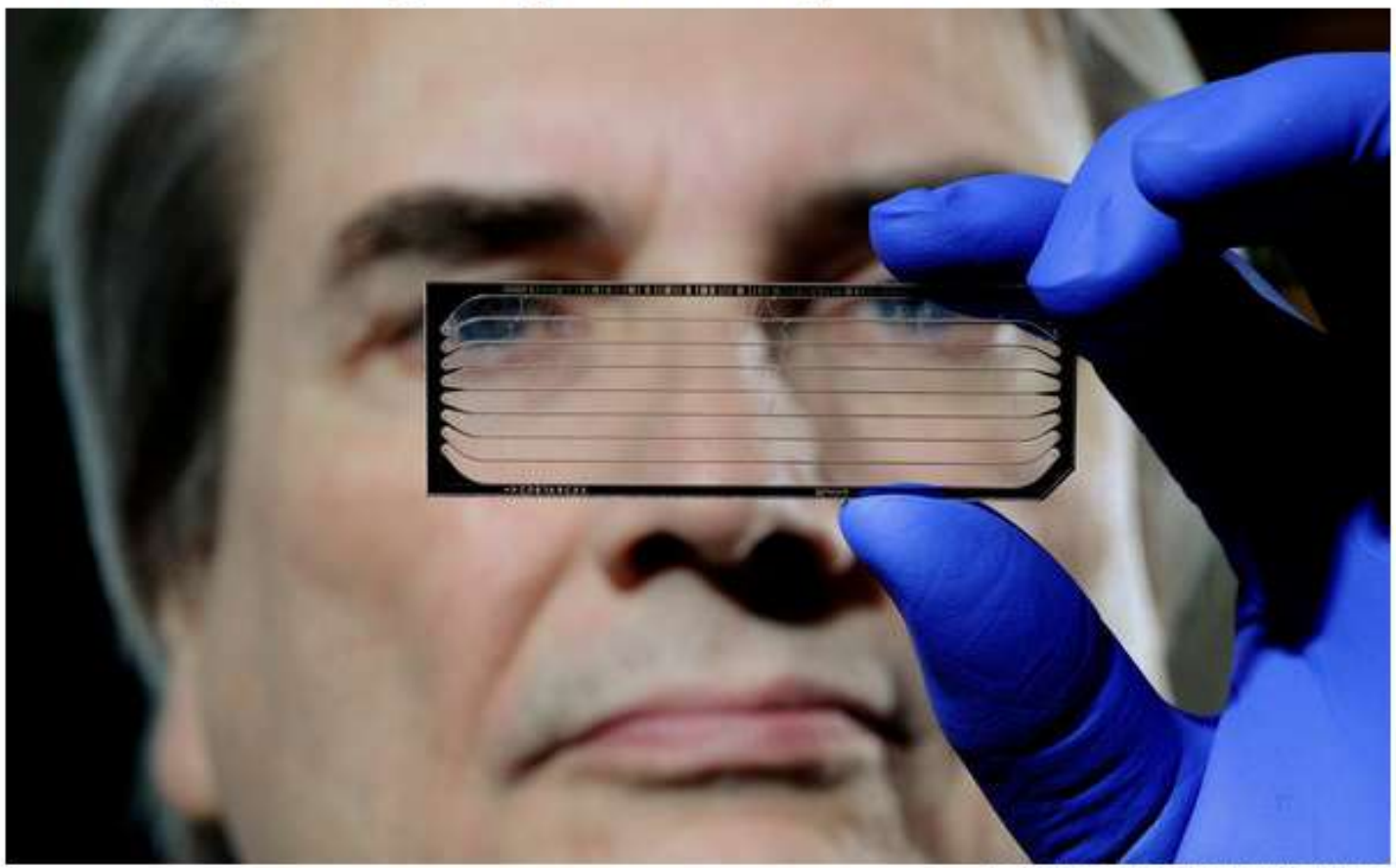
OptIPuter

Source: Falko Kuester, Kai Doerr Calit2; Michael Sims, Larry Edwards, Estelle Dodson NASA

it²

Genomic Sequencing is Driving Big Data

DNA Sequencing Caught in Deluge of Data



Kathy Kmonicek for The New York Times

The New York Times

Business Day

November 30, 2011

BGI—The Beijing Genome Institute is the World's Largest Genomic Institute





- **Main Facilities in Shenzhen and Hong Kong, China**
  - **Branch Facilities in Copenhagen, Boston, UC Davis**
- **137 Illumina HiSeq 2000 Next Generation Sequencing Systems**
  - **Each Illumina Next Gen Sequencer Generates 25 Gigabases/Day**
- **Supported by Supercomputing ~160TF, 33TB Memory**
  - **Large-Scale (12PB) Storage**

Using Advanced Info Tech and Telecommunications to Accelerate Response to Wildfires



Early on October 23, 2007, Harris Fire San Diego

Photo by Bill Clayton, http://map.sdsu.edu/

NASA's Aqua Satellite's MODIS Instrument Pinpoints the 14 SoCal Fires

Calit2, SDSU, and NASA Goddard Used NASA Prioritization and OptIPuter Links to Cut time to Receive Images from 24 to 3 Hours



Moderate Resolution Imaging Spectroradiometer (MODIS)

NASA/MODIS Rapid Response

www.nasa.gov/vision/earth/lookingatearth/socal_wildfires_oct07.html

High Performance Sensornets



Hans-Werner Braun, HPWREN PI

PSAP WIDC KYVW COTD KNW B08 1 BDC PFO GVDA Santa Rosa WMC RDM AZRY CRY SND BZN FRD KSW SMER P474 SO MPO LVA2 DHL SLMS SCS GLRS BVDA P486 P478 MTGY MVFD P510 CRRS WLA GMPK P483 RMNA DSME CWC USGC P506 P499 P480 P509 CE MONP UCSD MLO DESC P497 P494 70+ miles to SCI P473 SDSU IID2 P500 POTR P066

155Mbps FDX 6 GHz FCC licensed
155Mbps FDX 11 GHz FCC licensed
45Mbps FDX 6 GHz FCC licensed
45Mbps FDX 11 GHz FCC licensed
45Mbps FDX 5.8 GHz unlicensed
45Mbps-class HDX 4.9GHz
45Mbps-class HDX 5.8GHz unlicensed
~8Mbps HDX 2.4/5.8 GHz unlicensed
~3Mbps HDX 2.4 GHz unlicensed
115kbps HDX 900 MHz unlicensed
56kbps via RCS network

dashed = planned

approximately 50 miles:

HPWREN Topology, August 2008

HPWREN

- Backbone/relay node
- Astronomy science site
- Biology science site
- Earth science site
- University site
- Researcher location
- Native American site
- First Responder site

Situational Awareness for Wildfires: Combining HD VTC with Satellite Images, HPWREN Cameras & Sensors

Ron Robers, San Diego County Supervisor





Howard Windsor, San Diego CalFIRE Chief

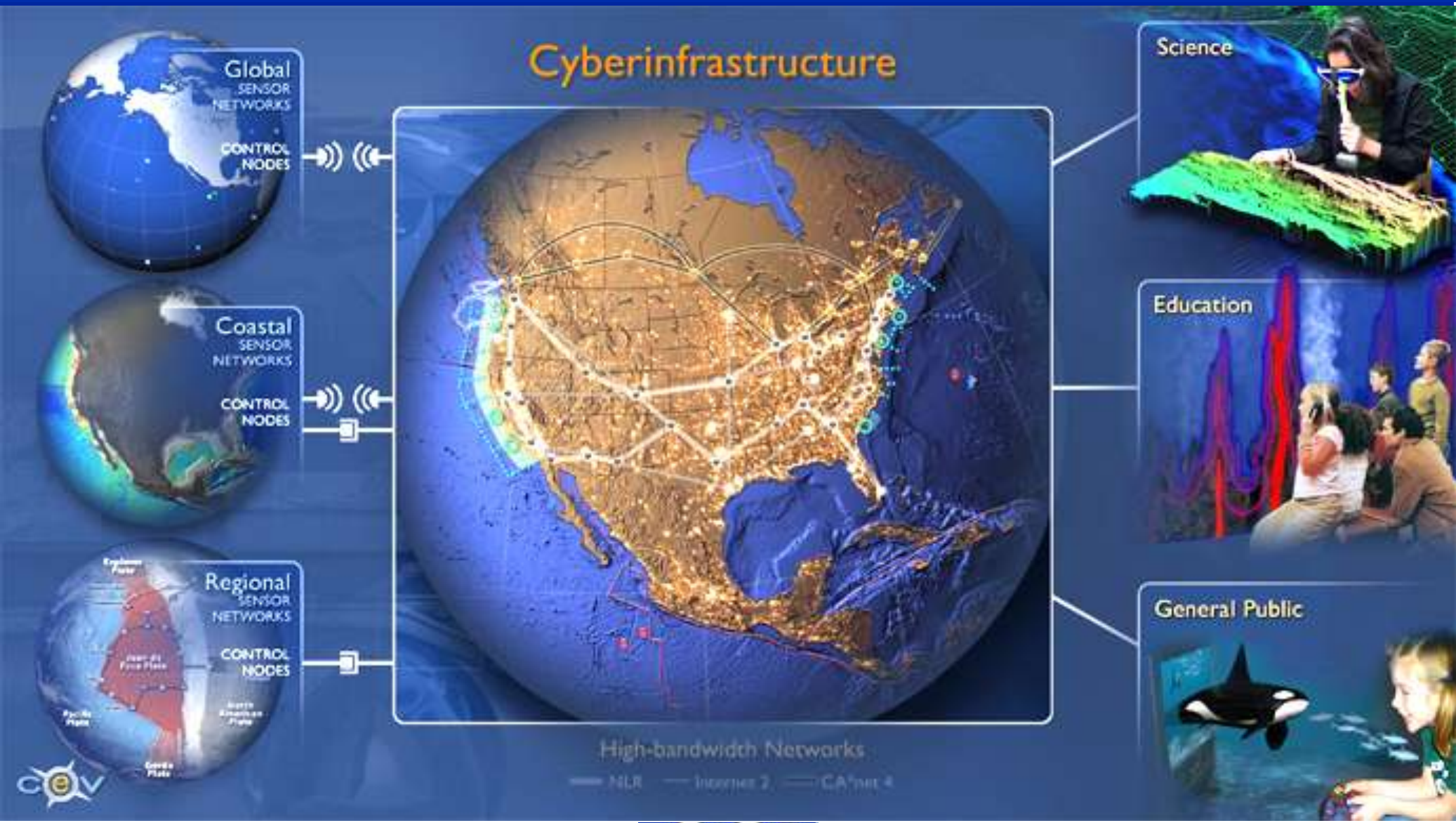
Powered by Rocks ROCKS

Source: Falko Kuester, Calit2@UCSD

CGLX

it²

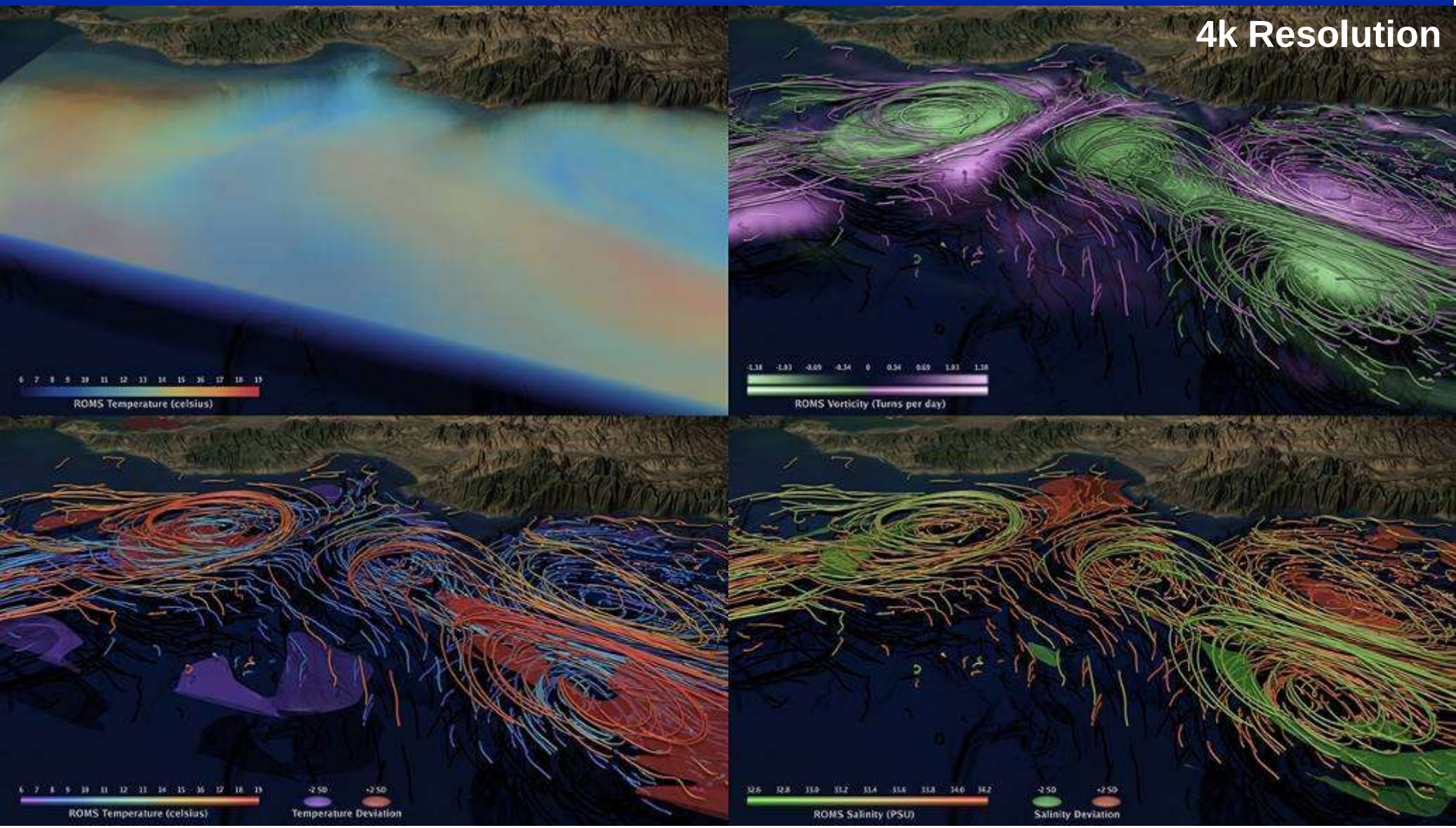
The NSF-Funded Ocean Observatory Initiative With a Cyberinfrastructure for a Complex System of Systems



OCEAN OBSERVATORIES INITIATIVE

Source: Matthew Arrott, Calit2 Program Manager for OOI CI

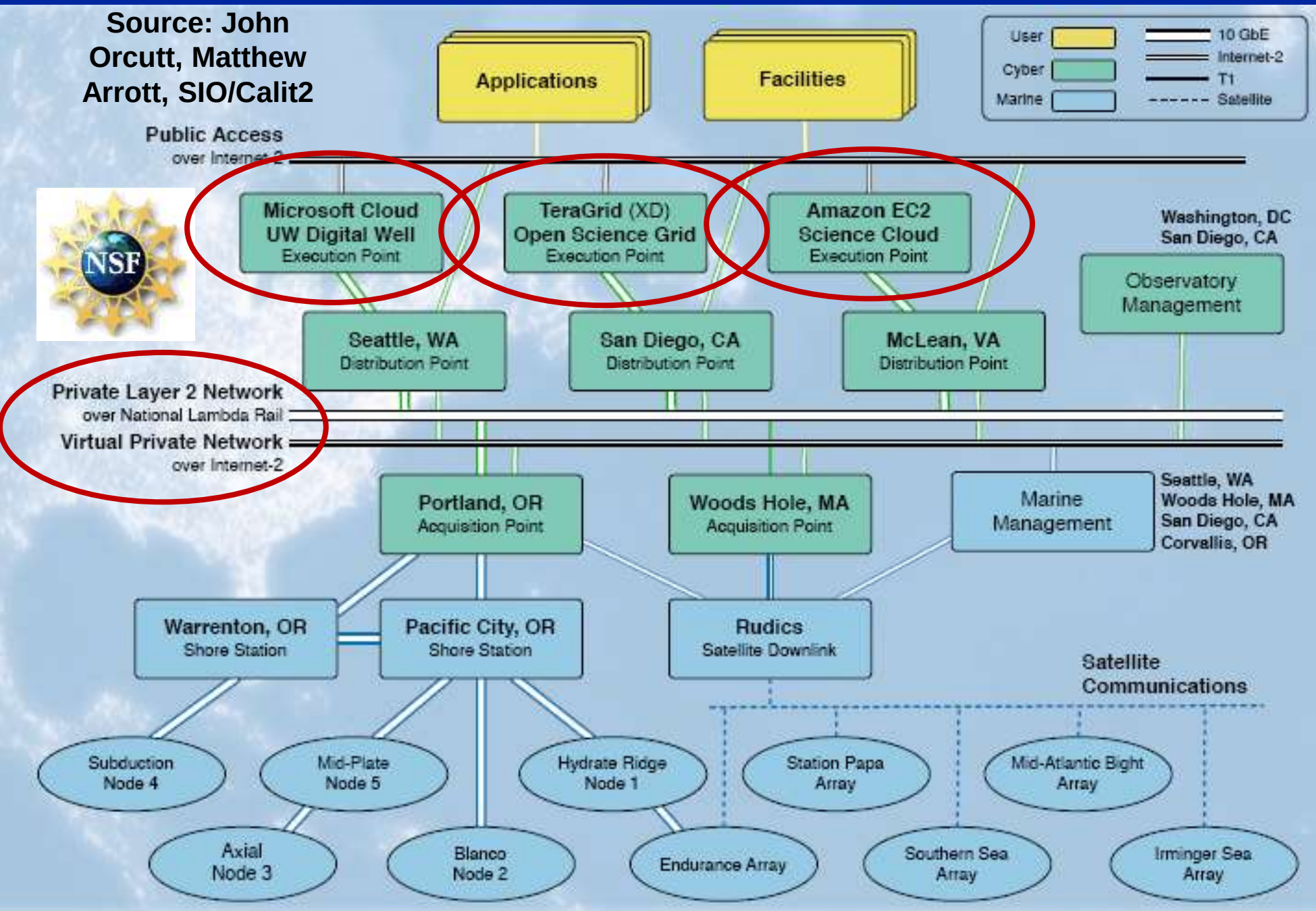
From Digital Cinema to Scientific Visualization: JPL Simulation of Monterey Bay



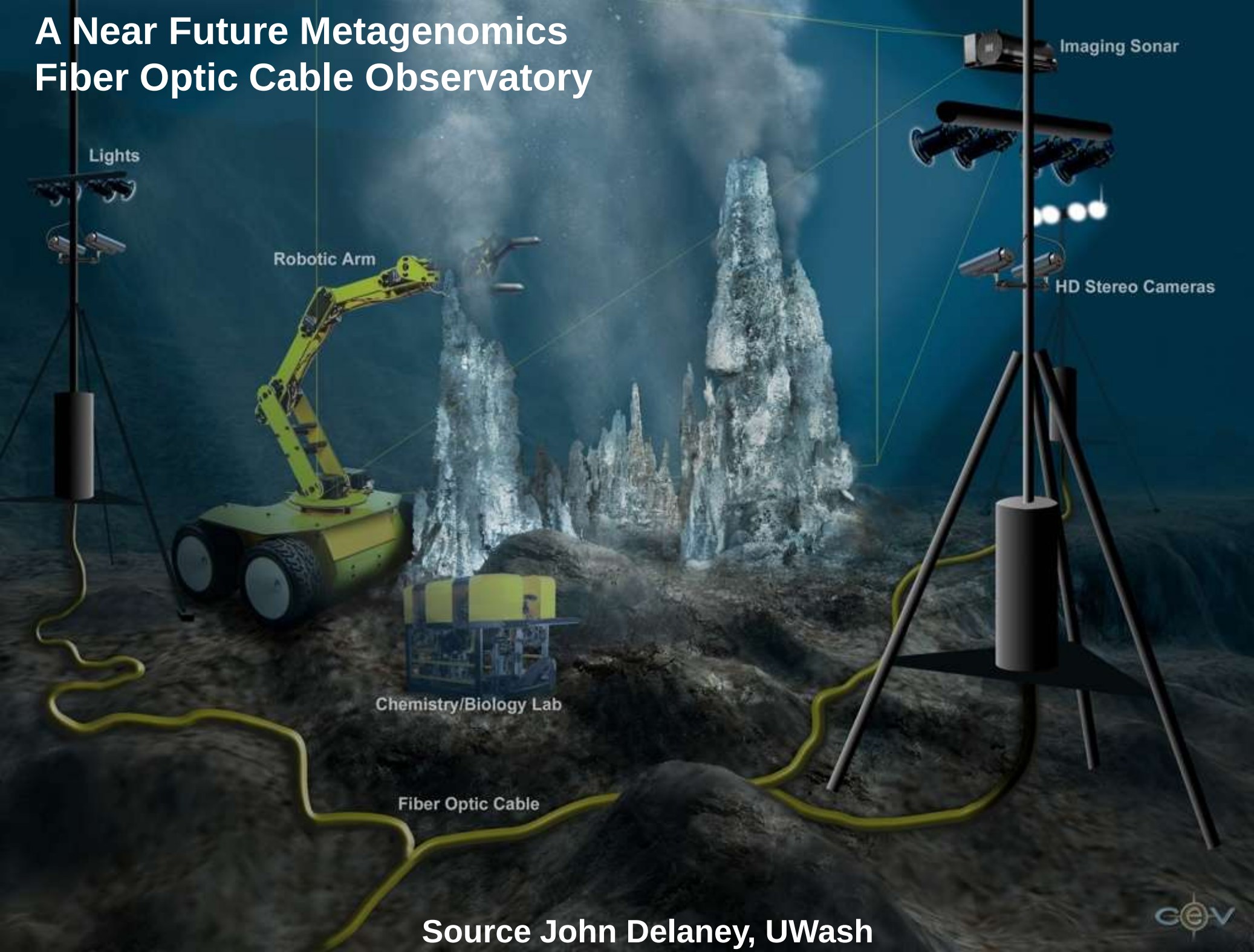
Source: Donna Cox, Robert Patterson, NCSA
Funded by NSF LOOKING Grant

OOI CI
is Built on NLR/I2 Optical Infrastructure

OCEAN OBSERVATORIES INITIATIVE



A Near Future Metagenomics Fiber Optic Cable Observatory



Source John Delaney, UWash

NSF Funds a Big Data Supercomputer: SDSC's Gordon-Dedicated Dec. 5, 2011

- **Data-Intensive Supercomputer Based on SSD Flash Memory and Virtual Shared Memory SW**
  - *Emphasizes MEM and IOPS over FLOPS*
  - *Supernode has Virtual Shared Memory:*
    - *2 TB RAM Aggregate*
    - *8 TB SSD Aggregate*
    - *Total Machine = 32 Supernodes*
    - *4 PB Disk Parallel File System >100 GB/s I/O*



- **System Designed to Accelerate Access to Massive Data Bases being Generated in Many Fields of Science, Engineering, Medicine, and Social Science**

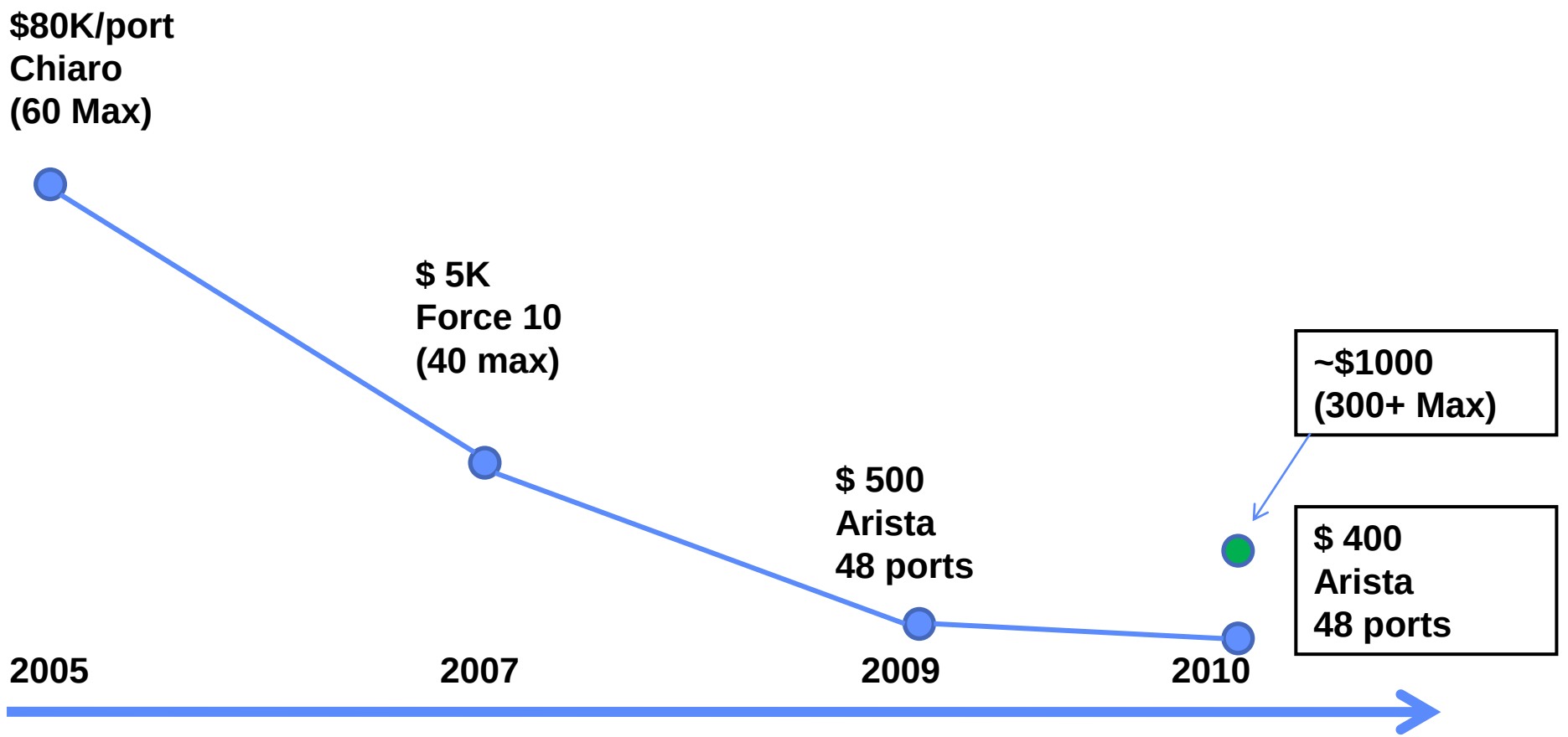
SDSC

Source: Mike Norman, Allan Snavely SDSC

it²

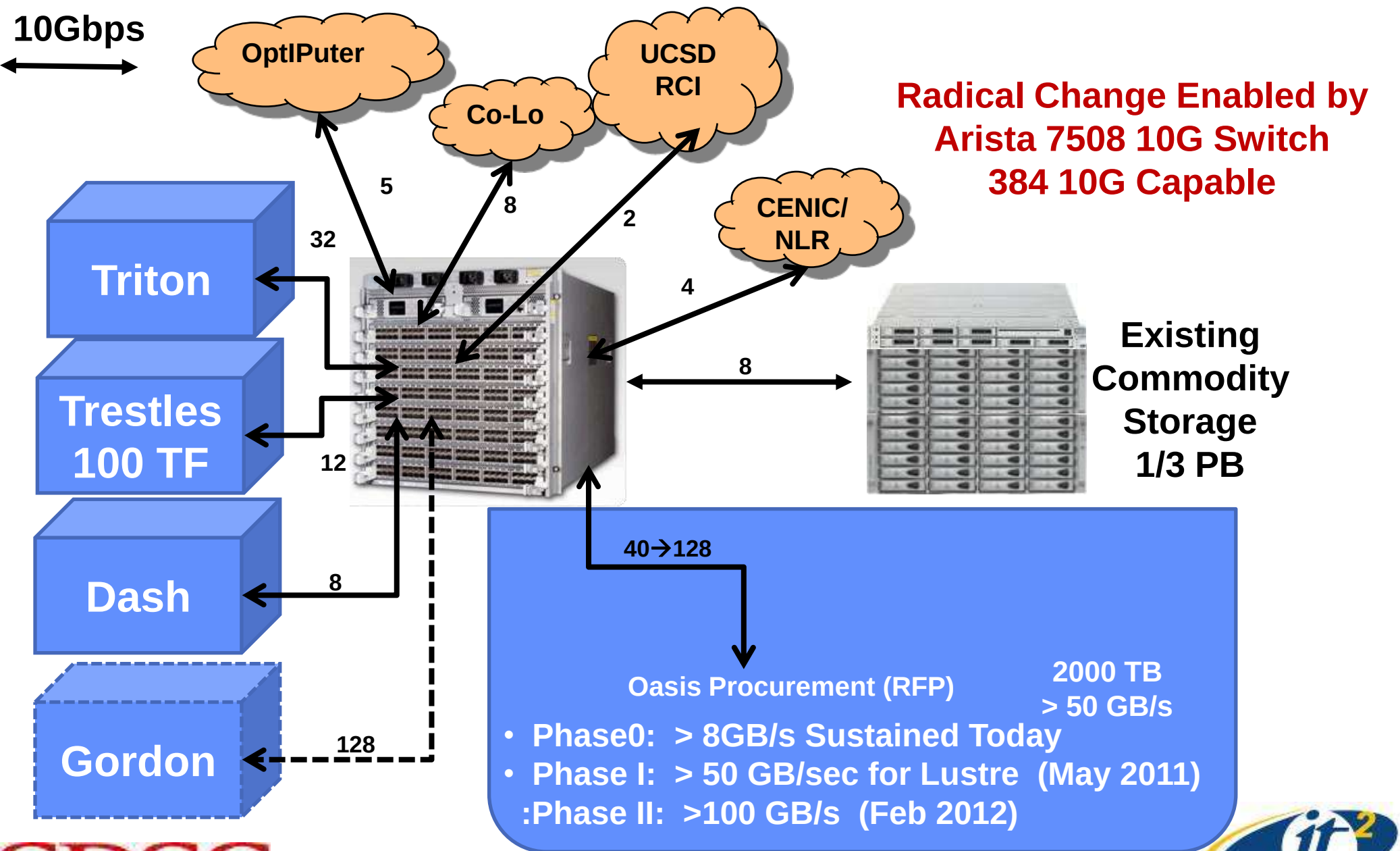
Rapid Evolution of 10GbE Port Prices Makes Campus-Scale 10Gbps CI Affordable

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



Source: Philip Papadopoulos, SDSC/Calit2

Arista Enables SDSC's Massive Parallel 10G Switched Data Analysis Resource



Source: Philip Papadopoulos, SDSC/Calit2

The Next Step for Data-Intensive Science: Pioneering the HPC Cloud



HPC Cloud Computing

Self Service and Dynamically Scalable

High Performance Computing (HPC) on









