"The UCSD Big Data Freeway System"

Invited Short Talk
Workshop on "Enriching Human Life and Society"
UC San Diego
February 6, 2014

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

The Data-Intensive Discovery Era Requires High Performance Cyberinfrastructure

- Growth of Digital Data is Exponential
  - "Data Tsunami"
- Driven by Advances in Digital Detectors, Computing, Networking, & Storage Technologies
- Shared Internet Optimized for Megabyte-Size Objects
- Need Dedicated Photonic Cyberinfrastructure for Gigabyte/Terabyte Data Objects
- Finding Patterns in the Data is the New Imperative
  - Data-Driven Applications
  - Data Mining
  - Visual Analytics
  - Data Analysis Workflows

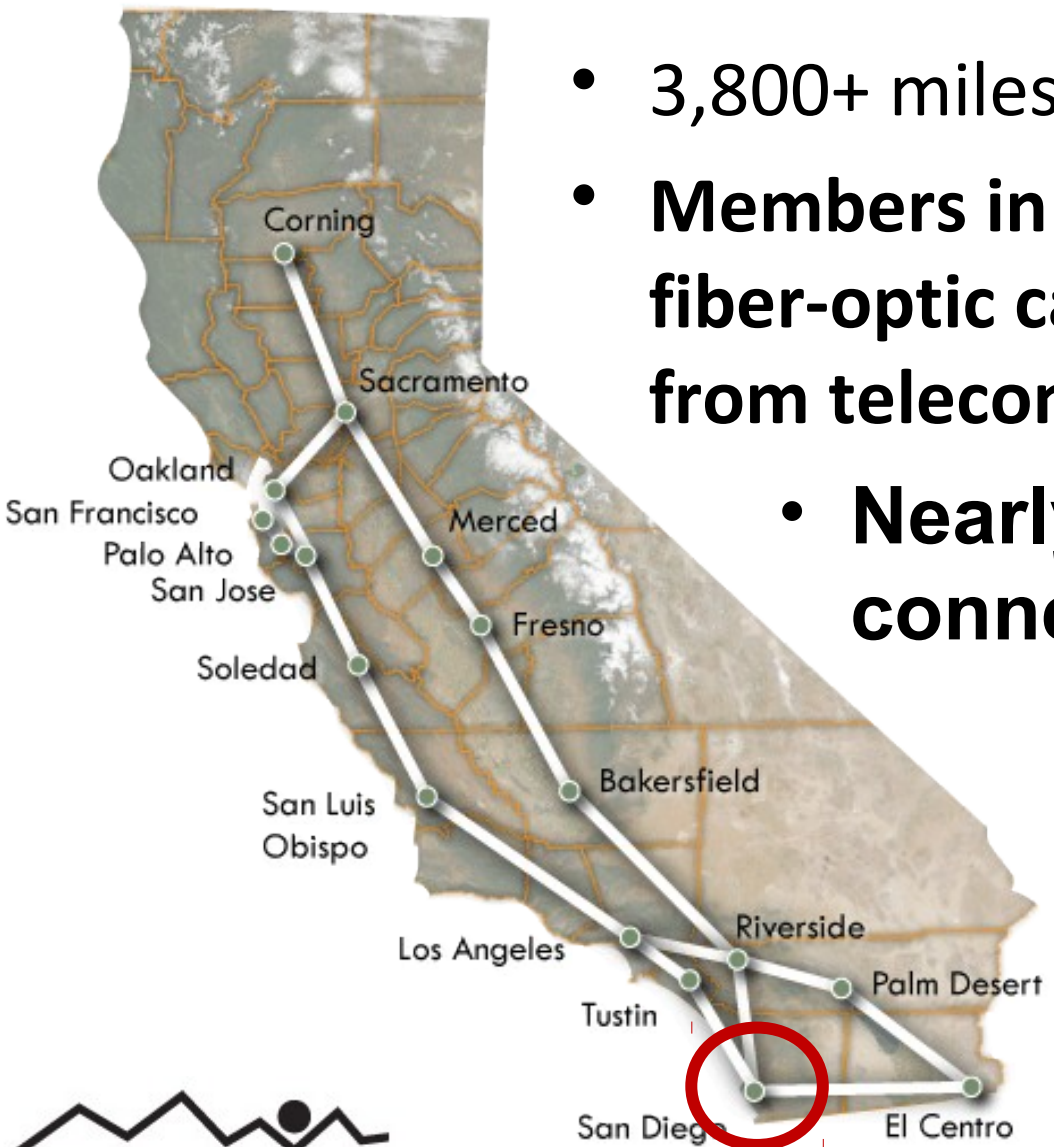


SDSC

Source: SDSC

it²

California Has the Most Advanced Big Data Optical Network in the Nation



- 3,800+ miles of optical fiber
- **Members in all 58 counties connect via fiber-optic cable or leased circuits from telecom carriers**
  - **Nearly *10,000* sites connect to CENIC**
    - *10,000,000+* Californians use CENIC each day
    - Governed by members on the segmental level

cenic

it²

CENIC is Rapidly Moving to Connect at 100 Gbps Across the State and Nation

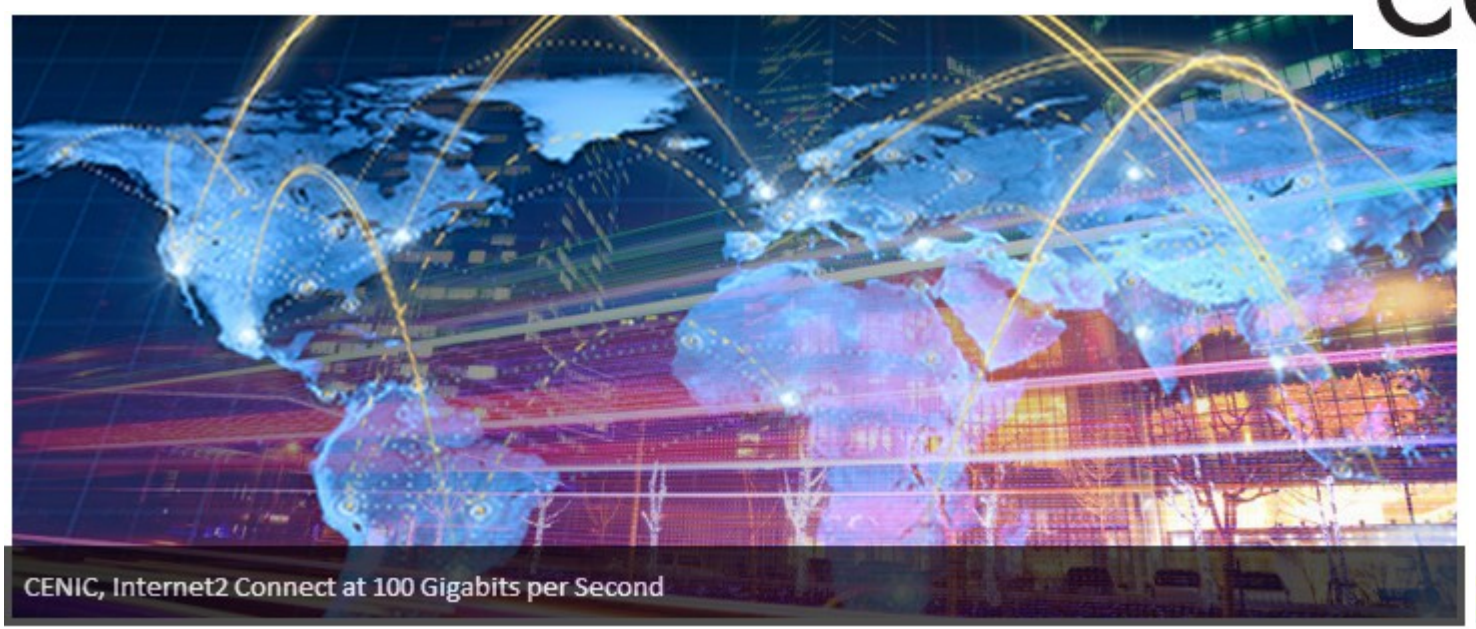
DOE



CENIC and ESnet Connected at 100 Gigabits per Second

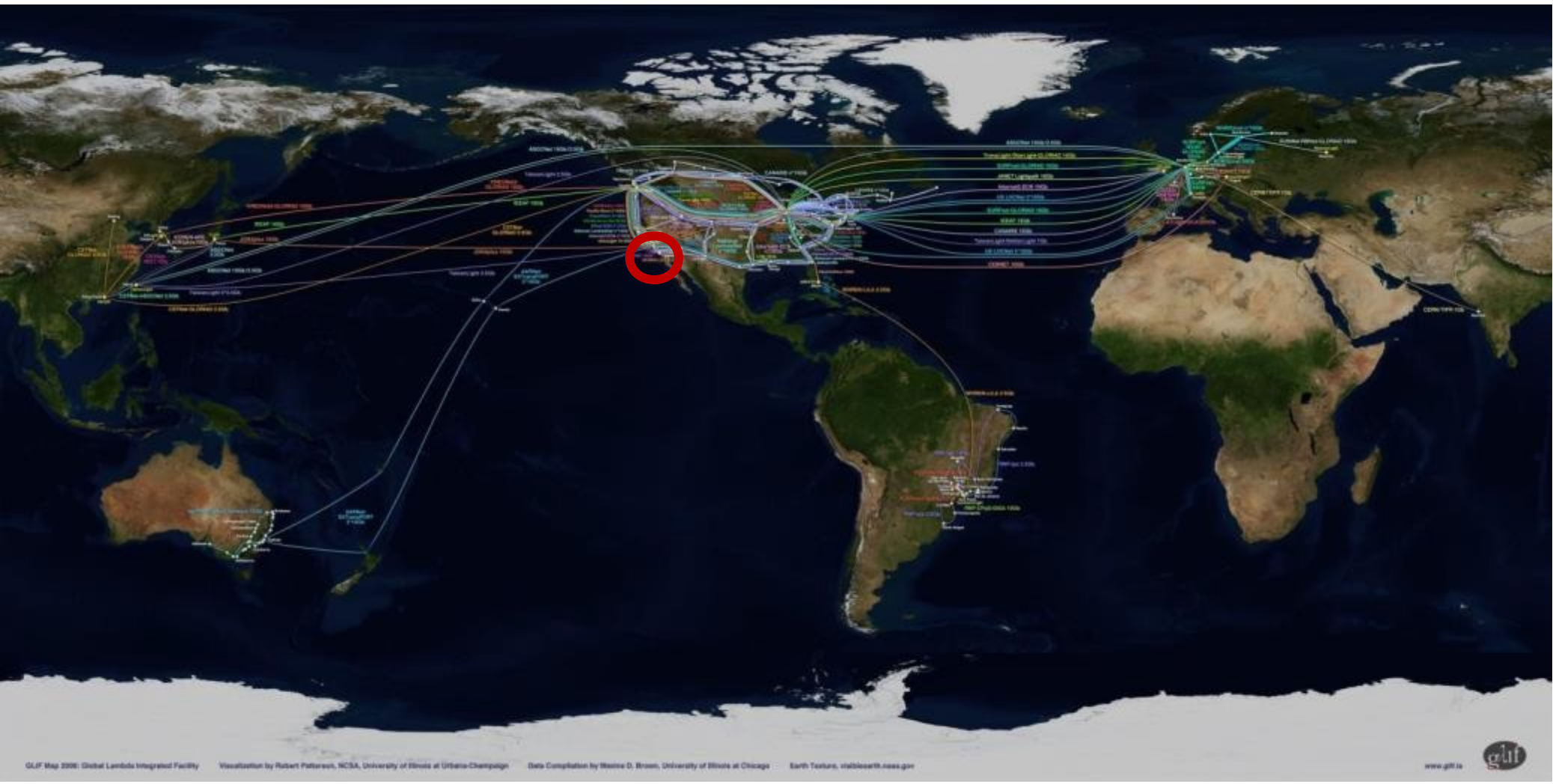
cenic

Internet2


CENIC, Internet2 Connect at 100 Gigabits per Second

it²

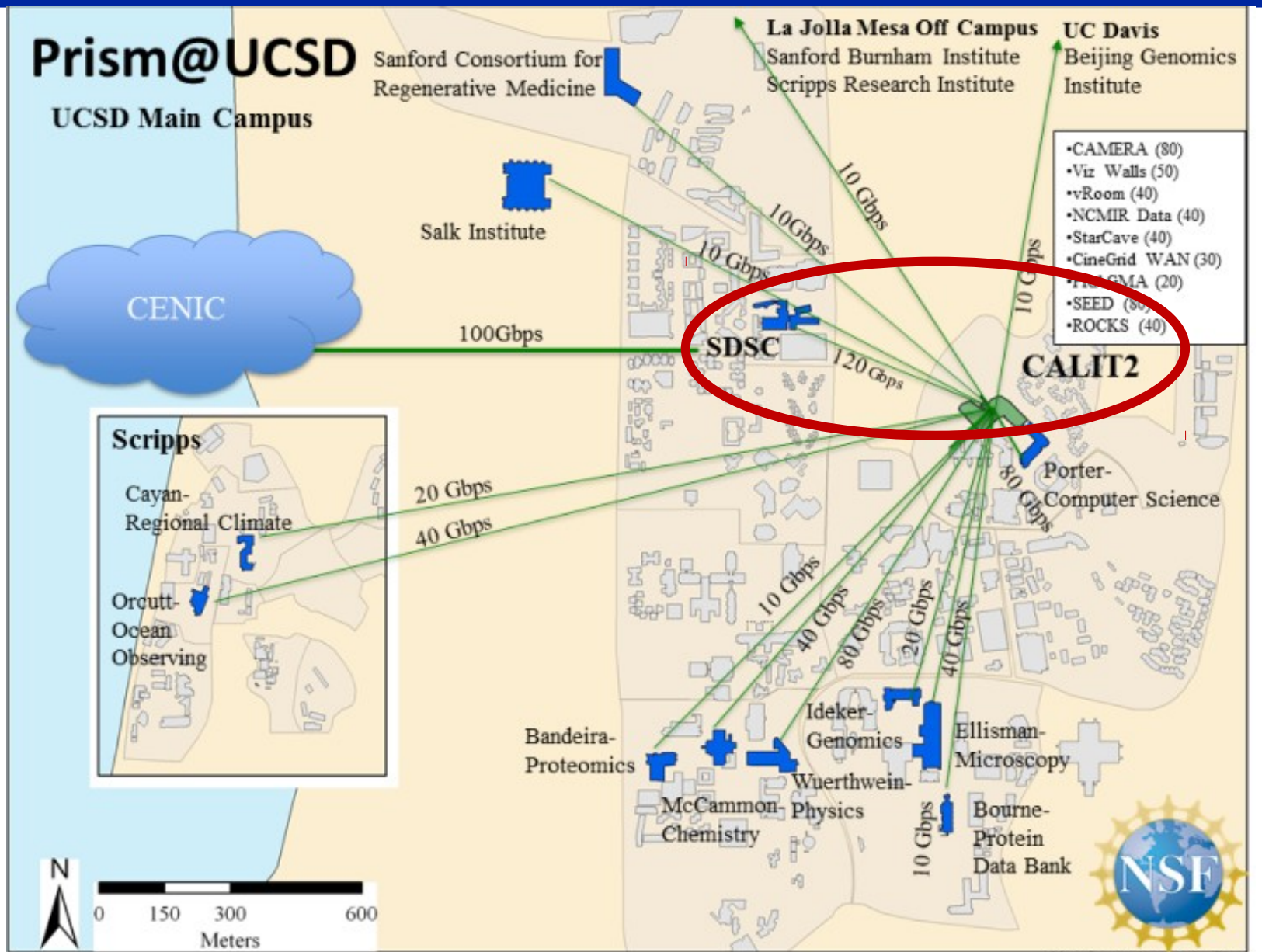
Global Innovation Centers are Being Connected with 10,000 Megabits/sec Clear Channel Lightpaths



Members of This Global Network Meet Annually at the Qualcomm Institute

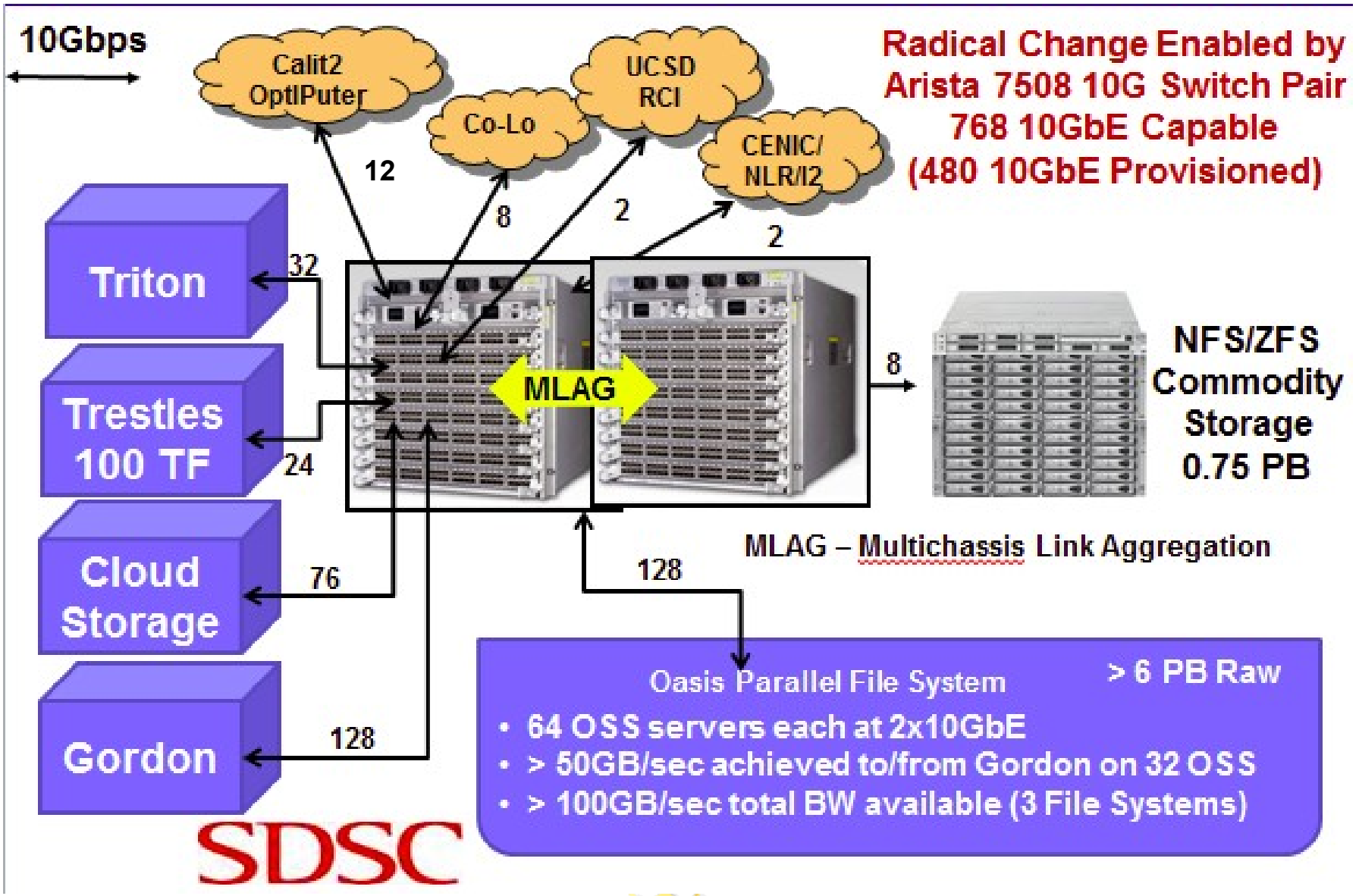
Source: Maxine Brown, UIC and Robert Patterson, NCSA

Connecting YOU on UCSD Campus with the World By Creating a Big Data Freeway System

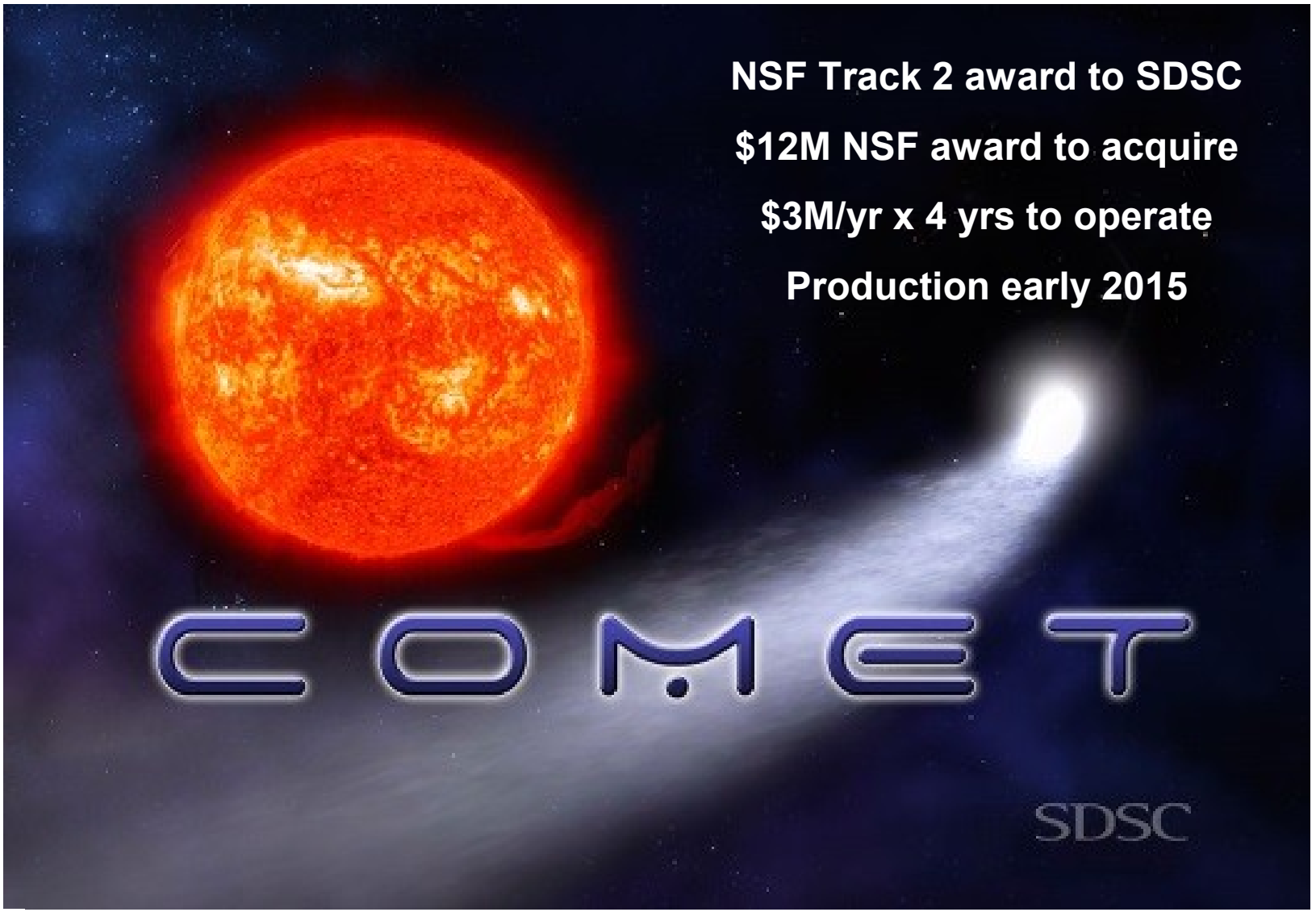


NSF CC-NIE Has Awarded Prism@UCSD Optical Switch
Phil Papadopoulos, SDSC, Calit2, PI

PRISM Puts SDSC's Big Data Gordon Supercomputer and Data Oasis Storage Into Your Lab



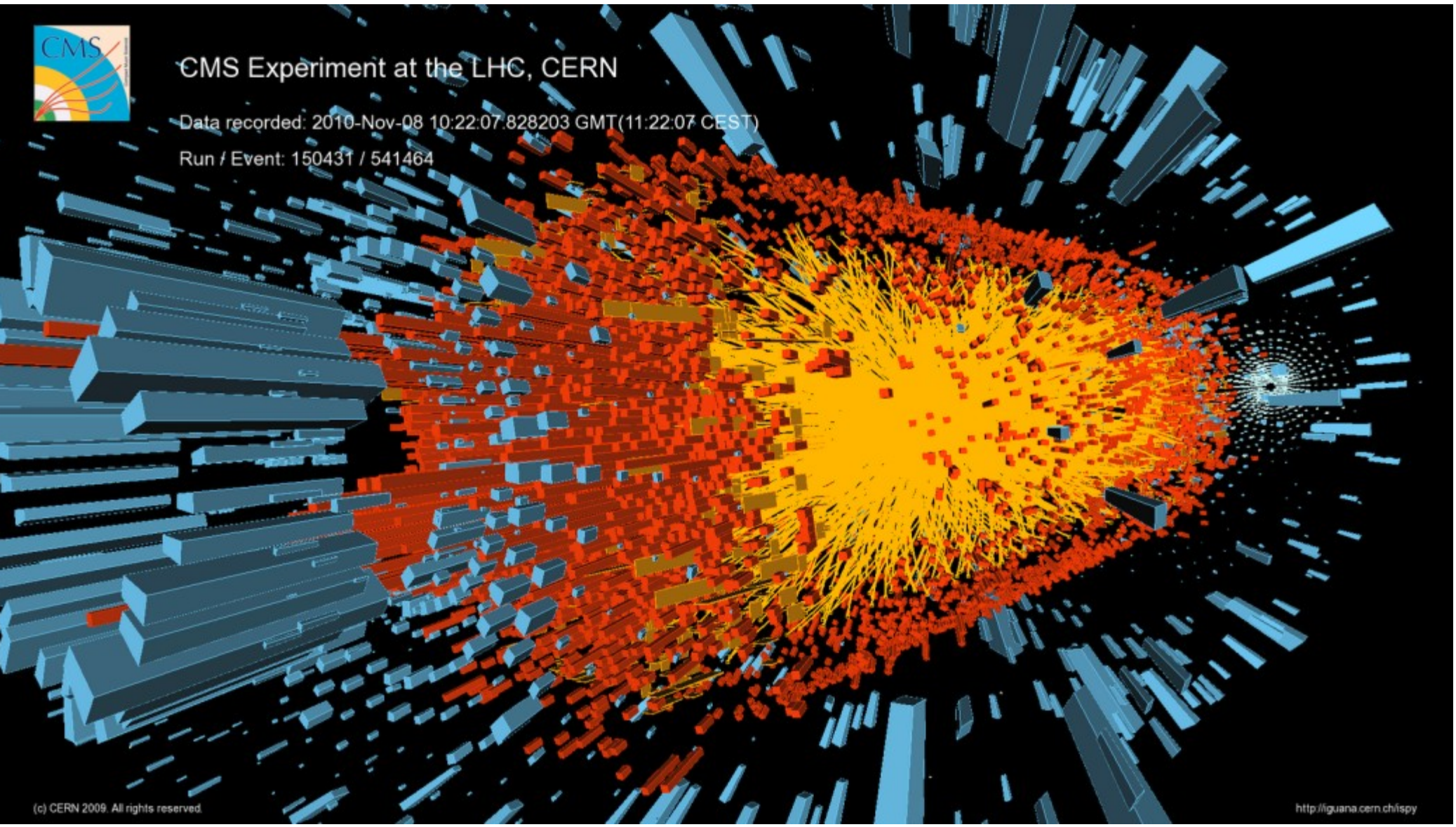
Comet is a ~2000TeraFLOP System Architected for the "Long Tail of Science"



UCSD

SDSC

PRISM is Connecting CERN's CMS Experiment To Our Physics Department



80 Gbps PRISM Connection Has Been Made

Planning for climate change in California

substantial shifts on top of already high climate variability

SIO Campus Climate Researchers Need to Download Results from Remote Supercomputer Simulations to Make Regional Climate Change Forecasts

Dan Cayan

USGS Water Resources Discipline

Scripps Institution of Oceanography, UC San Diego

much support from Mary Tyree, Mike Dettinger, Guido Franco and other colleagues

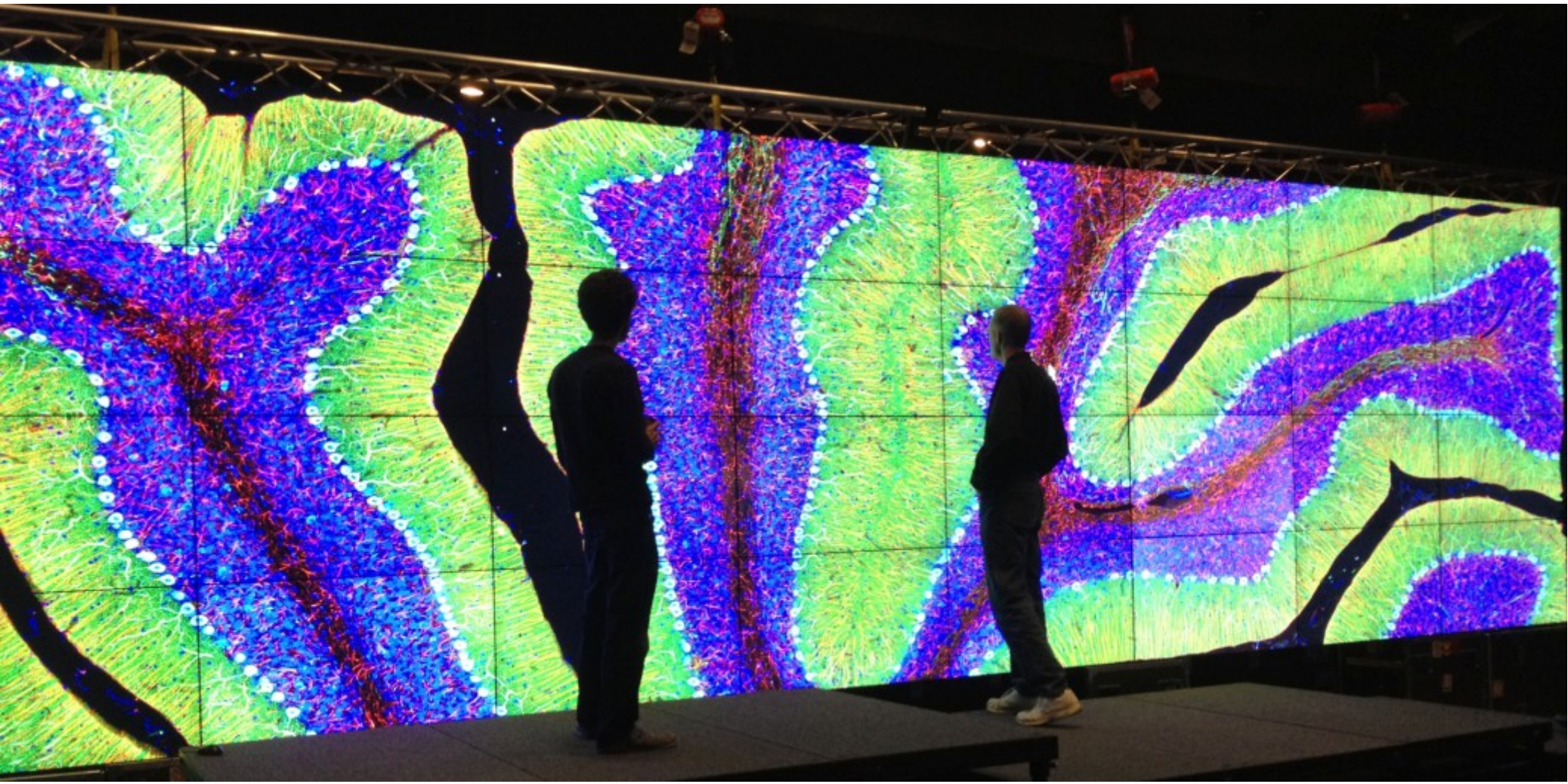
Sponsors:

California Energy Commission

NOAA RISA program

California DWR, DOE, NSF

PRISM Links Calit2's VROOM to NCMIR to Explore Confocal Light Microscope Images of Rat Brains



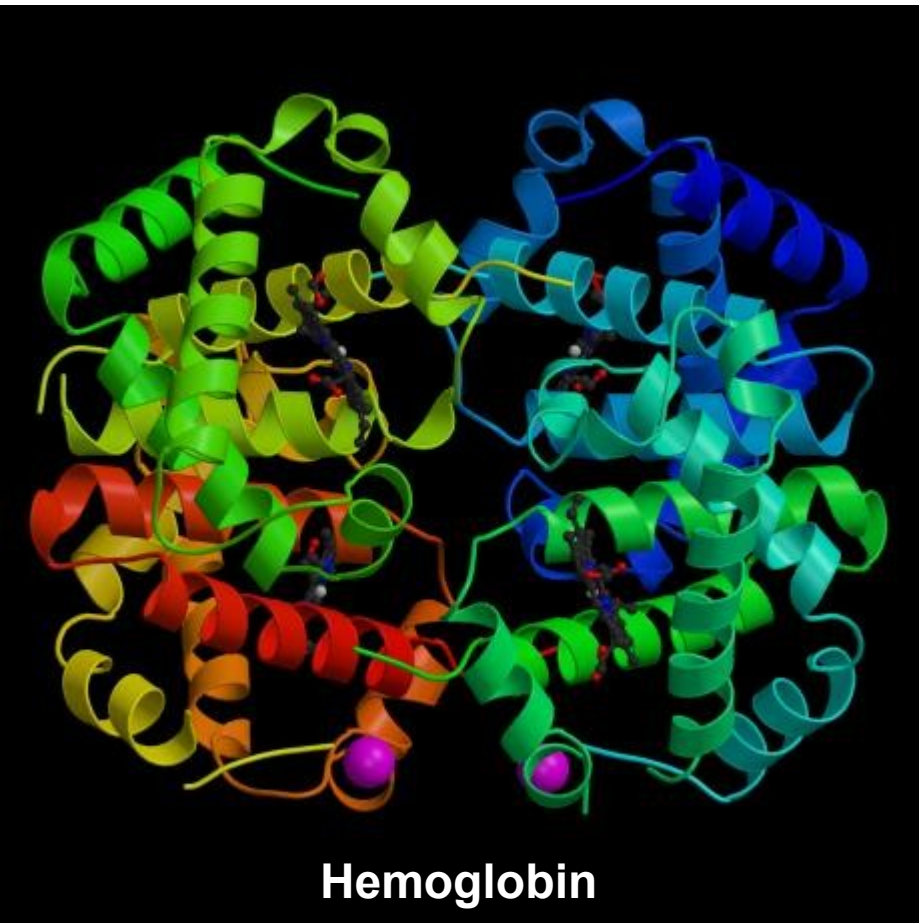


NCMIR National Center for Microscopy and Imaging Research

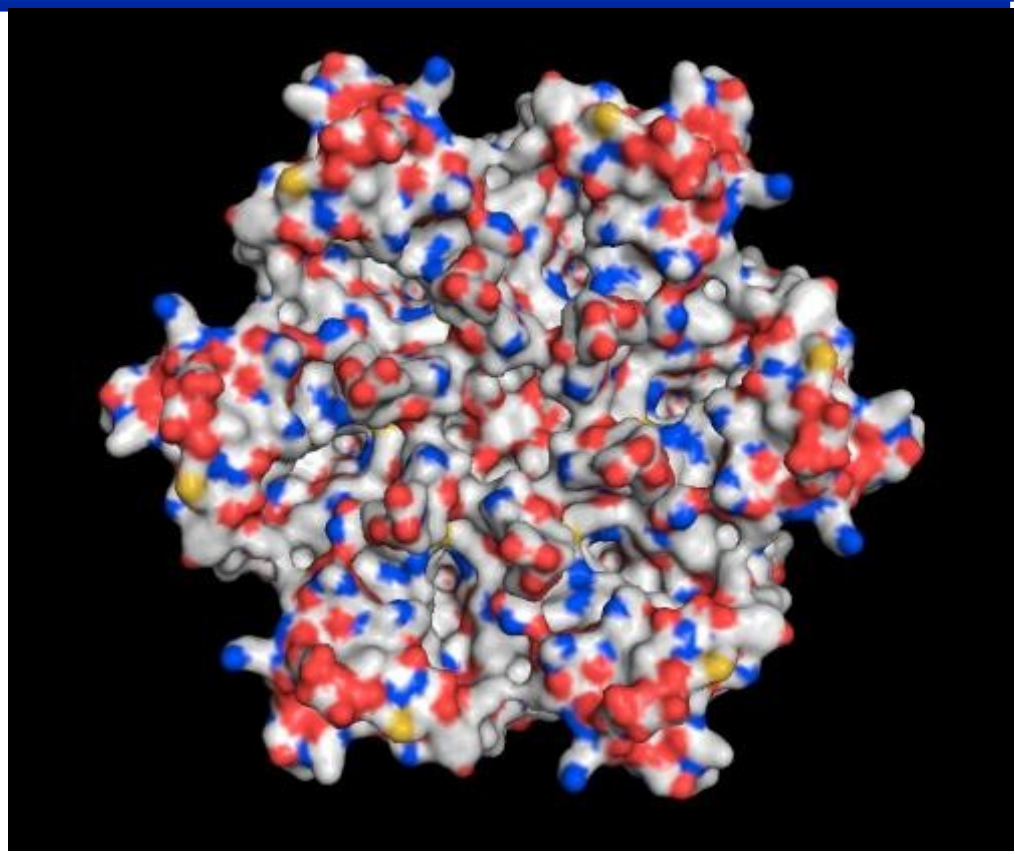
it²

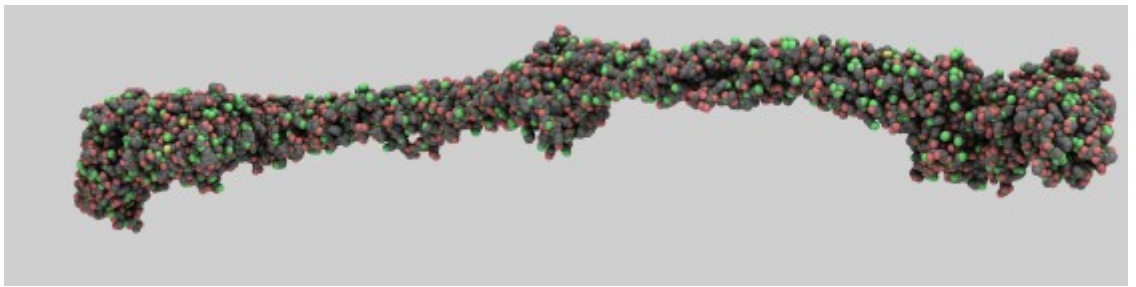
Protein Data Bank (PDB) Needs Bandwidth to Connect Resources and Users

- Archive of experimentally determined 3D structures of proteins, nucleic acids, complex assemblies
- One of the largest scientific resources in life sciences



Hemoglobin





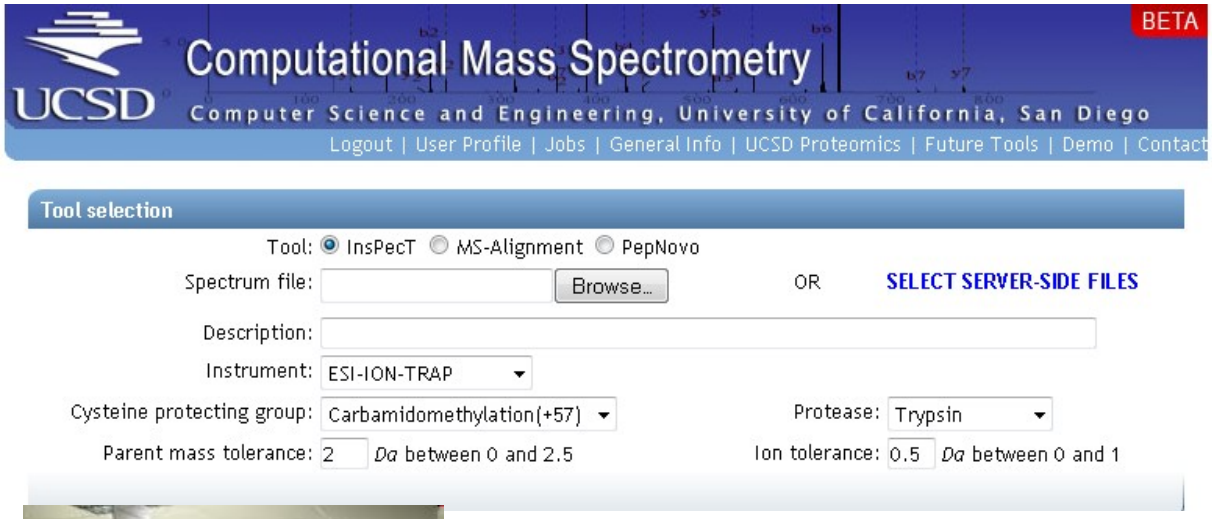
Source: Phil Bourne and Andreas Prlić, PDB

PRISM Will Link Computational Mass Spectrometry and Genome Sequencing Cores to the Big Data Freeway

Source: proteomics.ucsd.edu

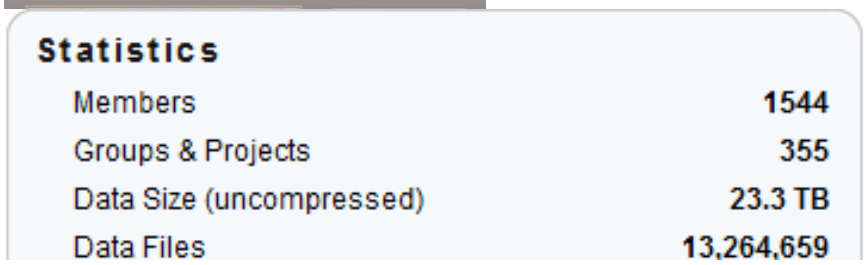
***ProteoSAFe*: Compute-intensive discovery MS at the click of a button**

***MassIVE*: repository and identification platform for all MS data in the world**



Molecular & Cellular Proteomics

Journal of proteome research







IGM Institute for Genomic Medicine

National Institutes of Health
National Institute of General Medical Sciences
BASIC DISCOVERIES FOR BETTER HEALTH

it²



Virtual Jazz-Coupling UCI with UCSD With a Gbps Lightwave

Collaboration Between EVL's CAVE2 and Calit2's VROOM Over 10Gb Wavelength



Source: NTT Sponsored ON*VECTOR Workshop at Calit2 March 6, 2013