"Driving Applications on the UCSD Big Data Freeway System"

Keynote Lecture

Cubic and UC San Diego Innovation Workshop

UC San Diego

February 26, 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

The White House Announcement Has Galvanized U.S. Campus CI Innovations

News

White House launches big data R&D effort

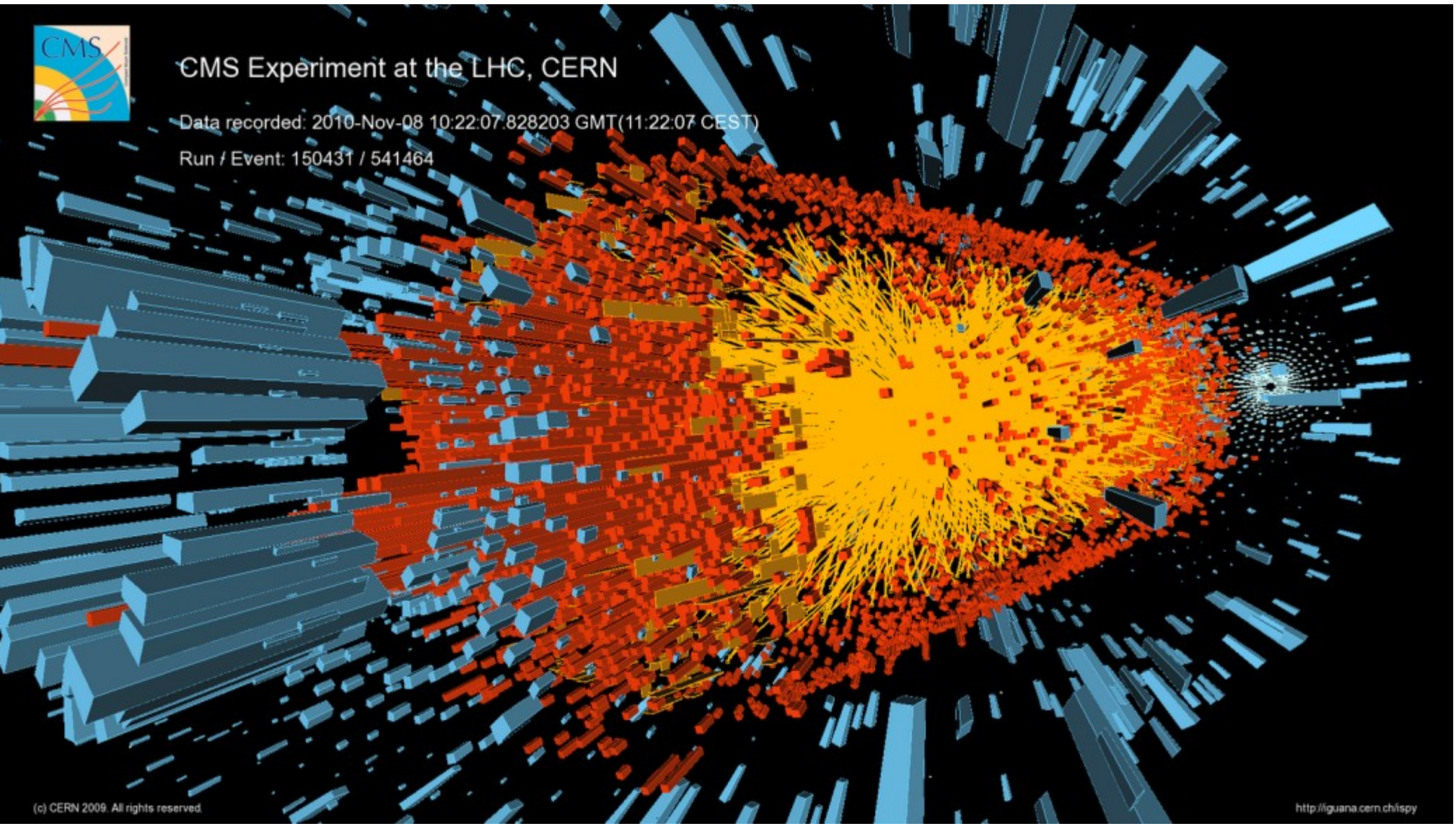
NSF and DARPA are among the agencies looking to harness the power of large volumes of data

By Grant Gross

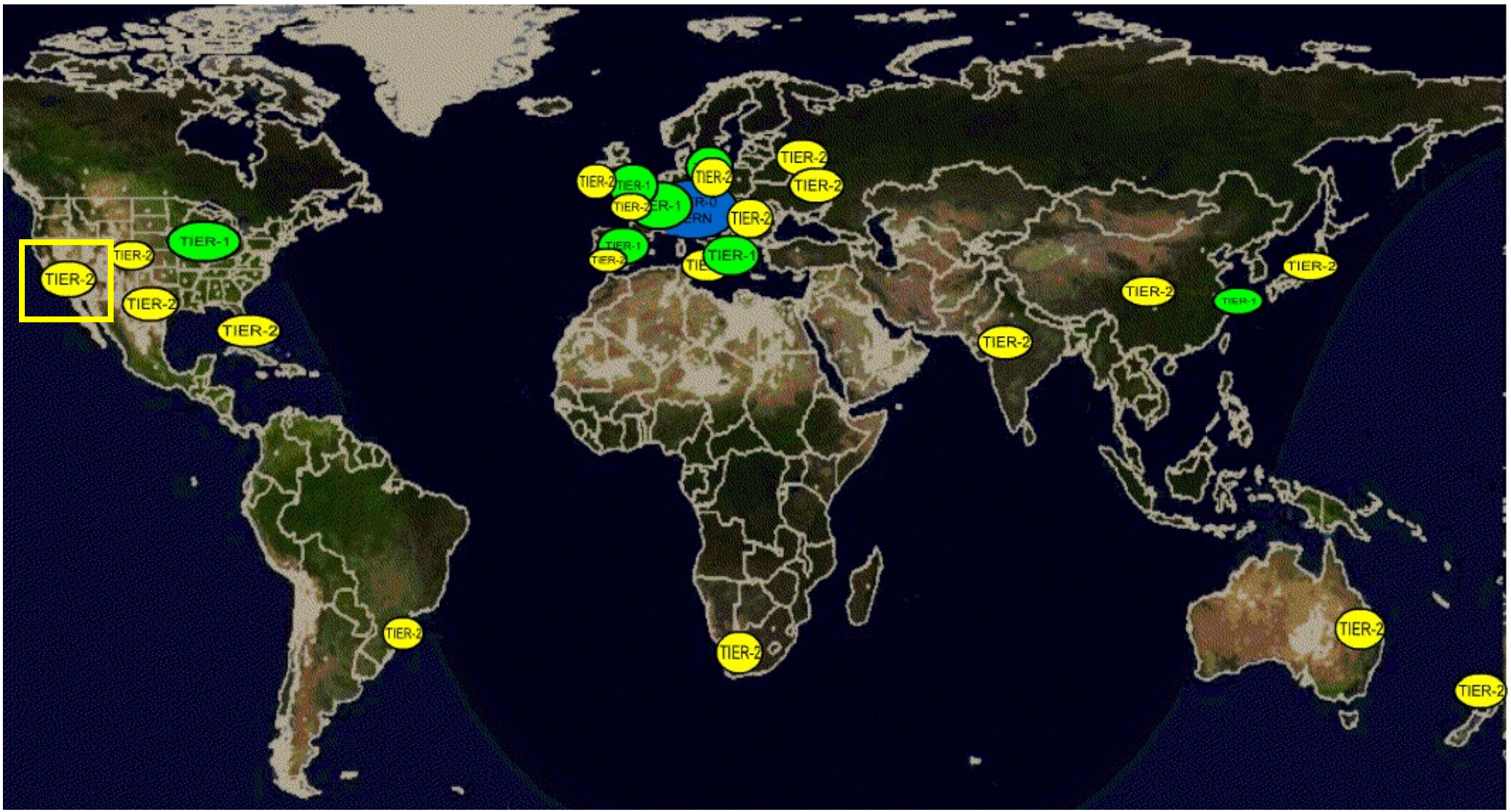
March 29, 2012 12:34 PM ET

5 Comments + Briefcase What's this?

CERN's CMS Experiment Generates Massive Amounts of Data



UCSD is a Tier-2 LHC Data Center: CMS Flow into UCSD Physics Dept. Peaks at 2.4 Gbps



Source: Frank Wuerthwein, Physics UCSD

Planning for climate change in California

substantial shifts on top of already high climate variability

UCSD 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

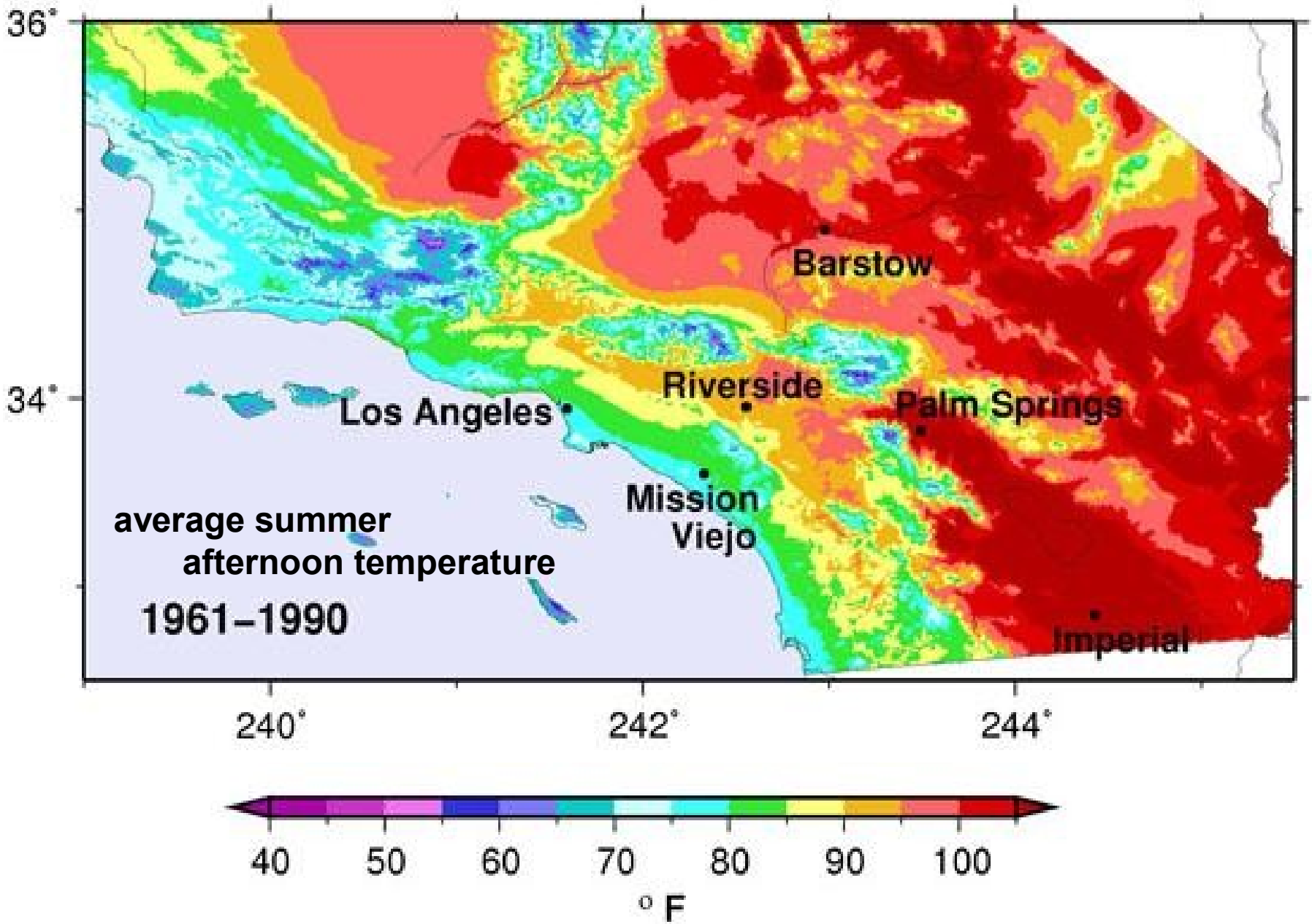
much support from Mary Tyree, Mike Dettinger, Guido Franco and other colleagues

Sponsors:

California Energy Commission

NOAA RISA program

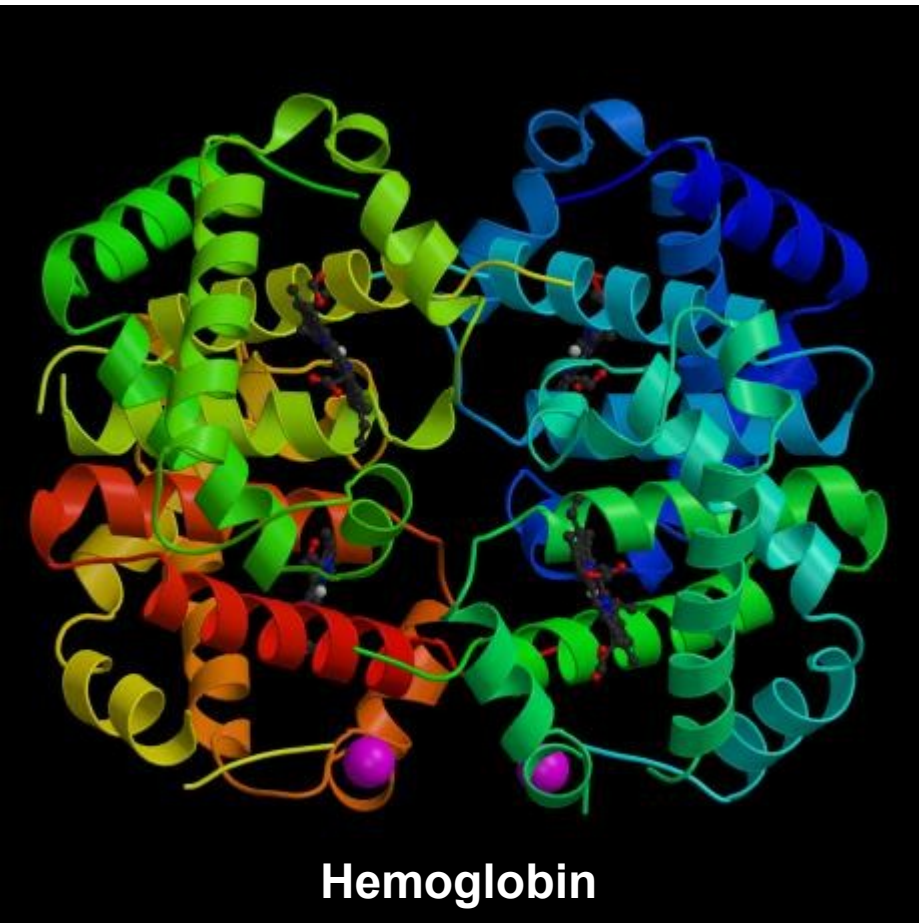
California DWR, DOE, NSF



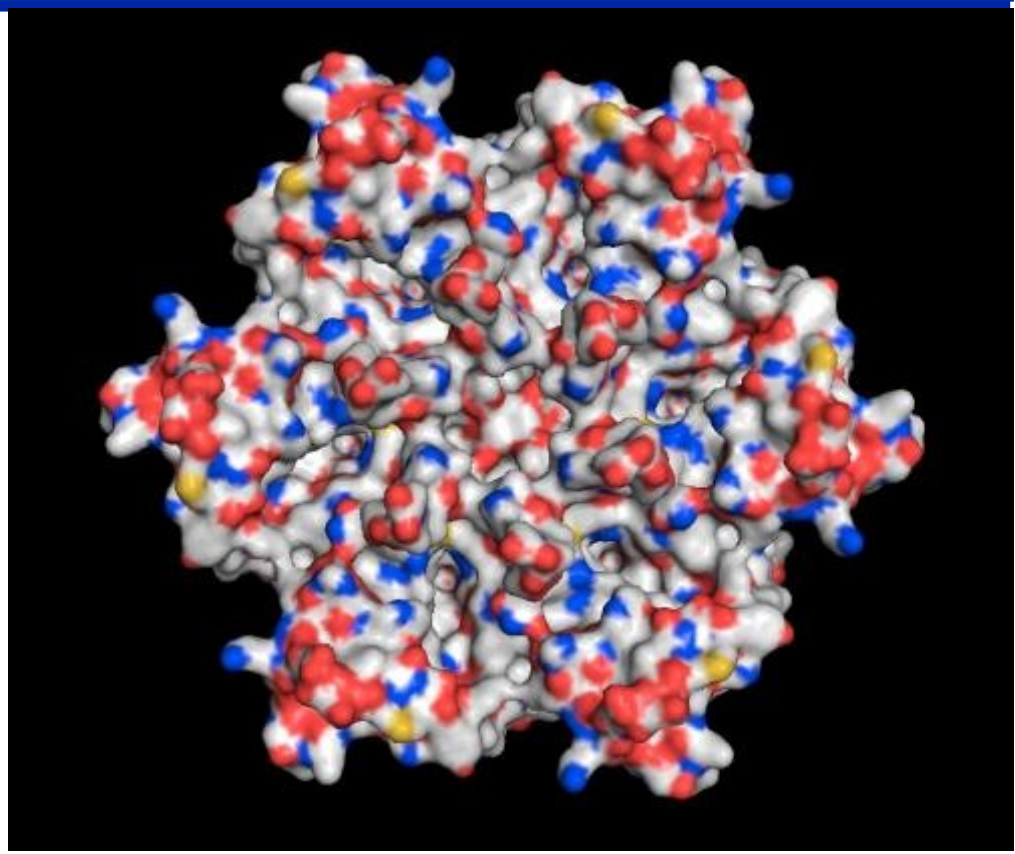
GFDL A2 1km downscaled to 1km
Hugo Hidalgo Tapash Das Mike Dettinger

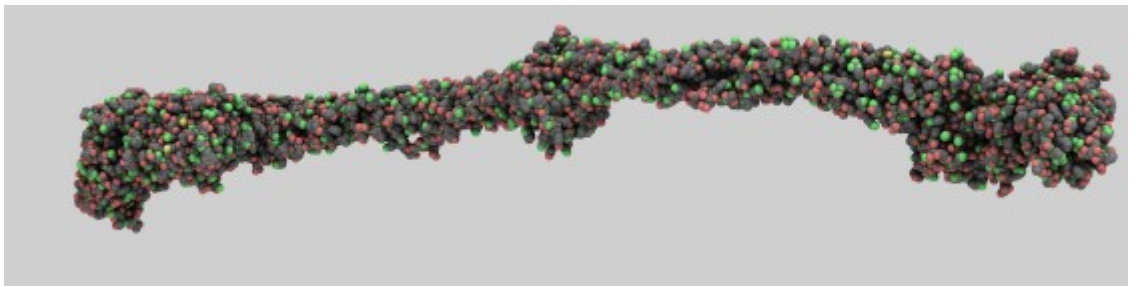
Protein Data Bank (PDB) Needs Bandwidth to Connect Resources and Users

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



Hemoglobin



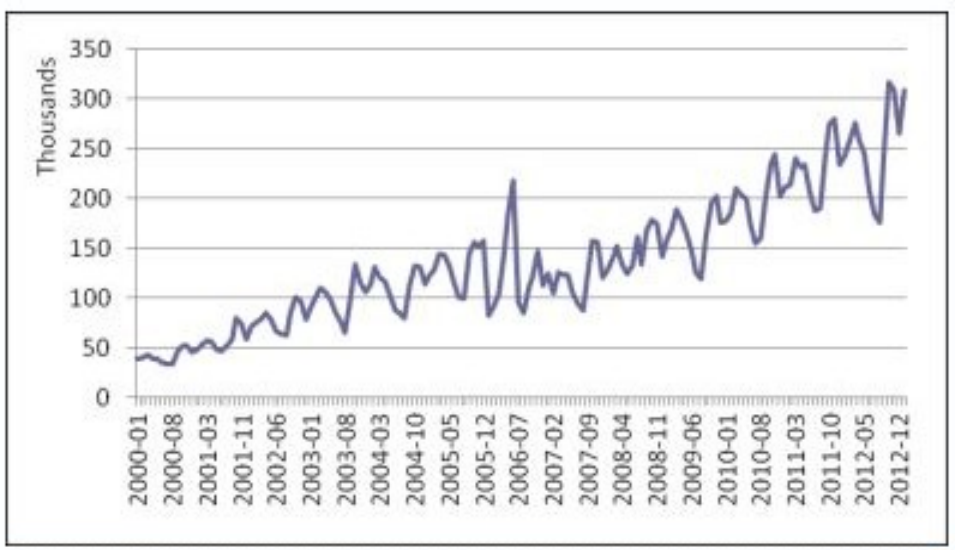


Source: Phil Bourne and Andreas Prlić, PDB

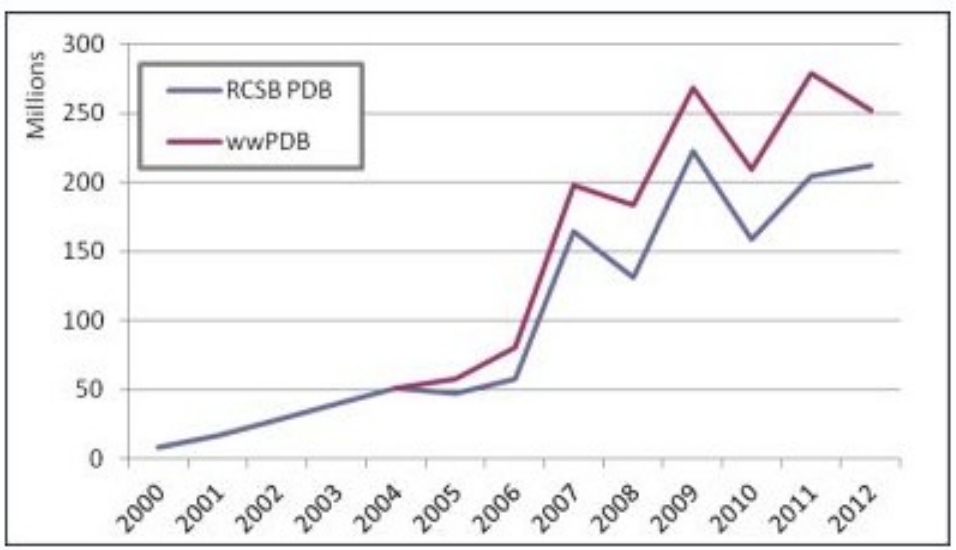
Protein Data Bank Usage Is Growing Over Time

- More than 300,000 Unique Global Visitors per Month
- Up to 300 Concurrent Users
- ~10 Structures are Downloaded per Second 7/24/365
- Increasingly Popular Web Services Traffic

RCSB PDB Website Unique Users by Month



FTP Downloads by Year



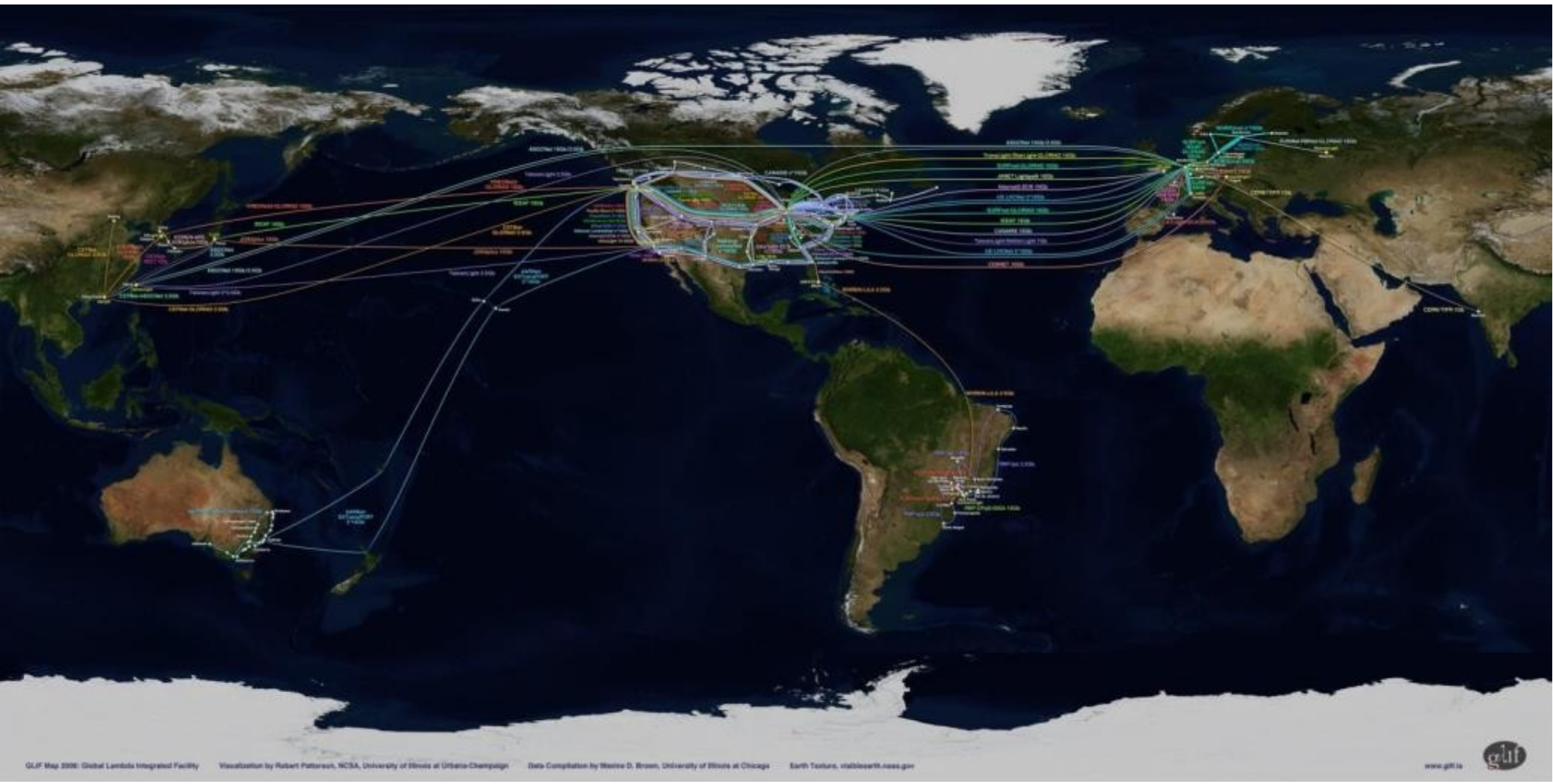
Source: Phil Bourne and Andreas Prlić, PDB

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



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

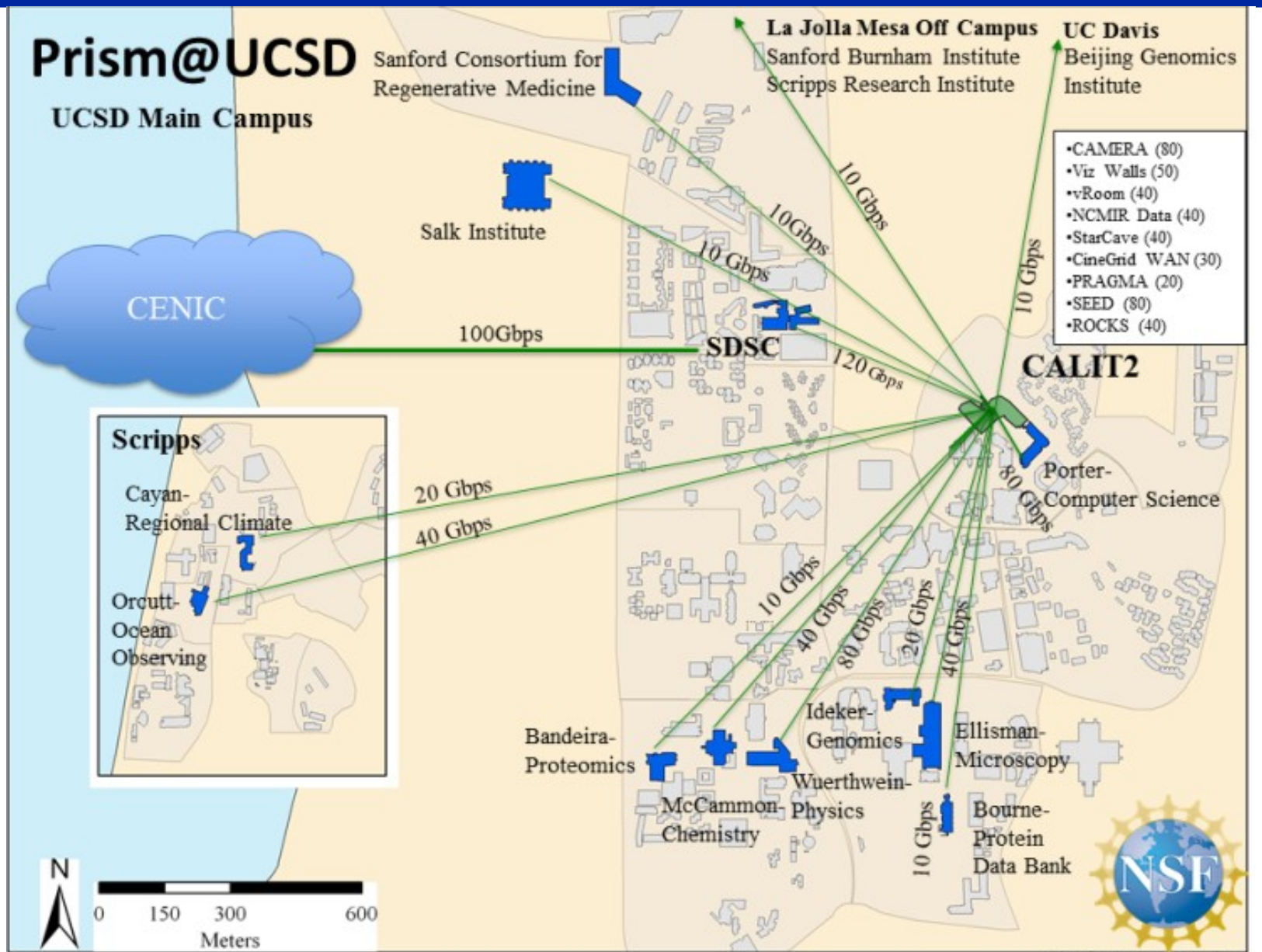
Global Innovation Centers are Being Connected with 10,000 Megabits/sec Clear Channel Lightpaths



100 Gbps Commercially Available;
Research on 1 Tbps

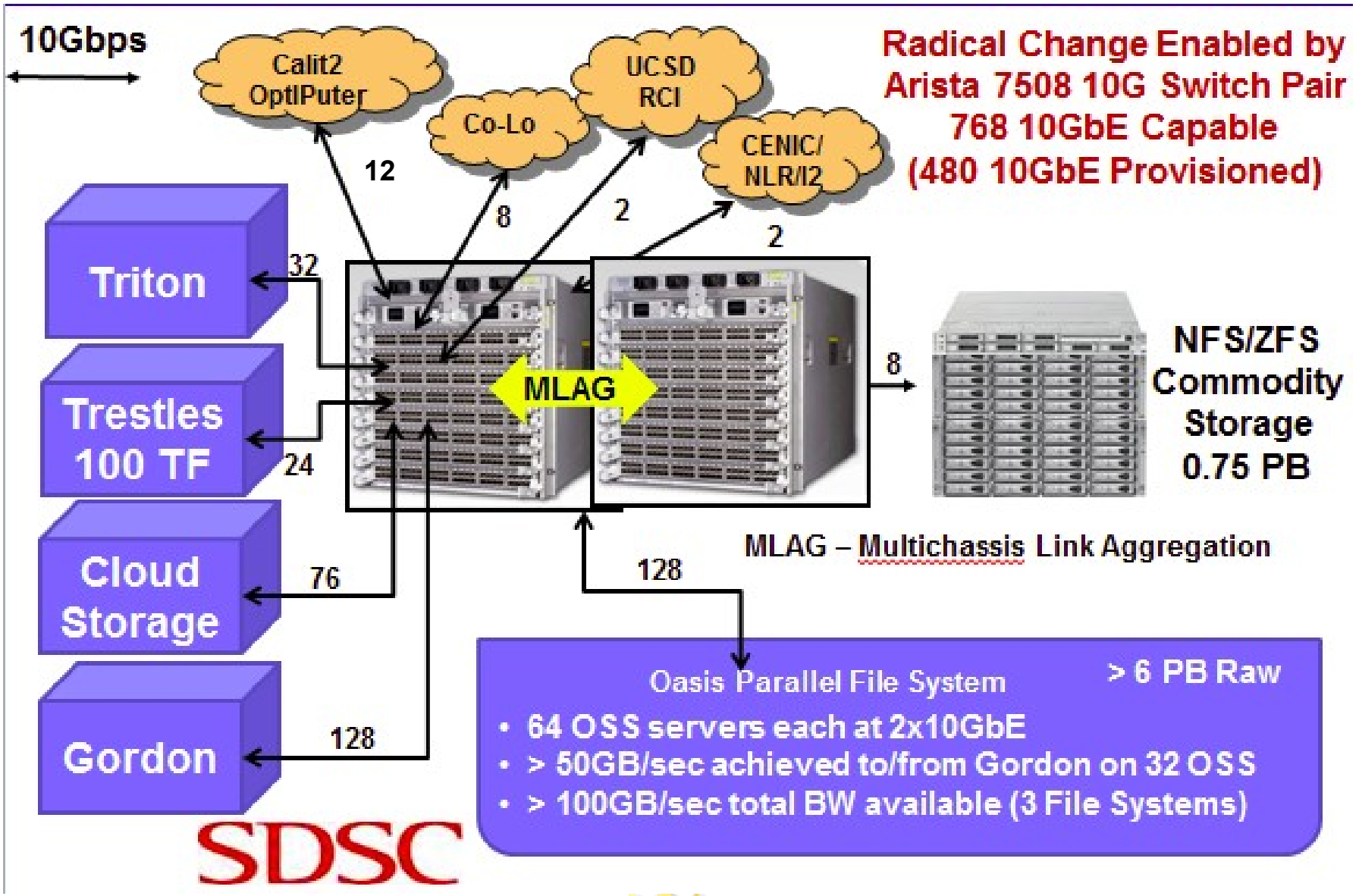
Source: Maxine Brown, UIC and Robert Patterson, NCSA

Creating a Big Data Freeway System: Use Optical Fiber with 1000x Shared Internet Speeds



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

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





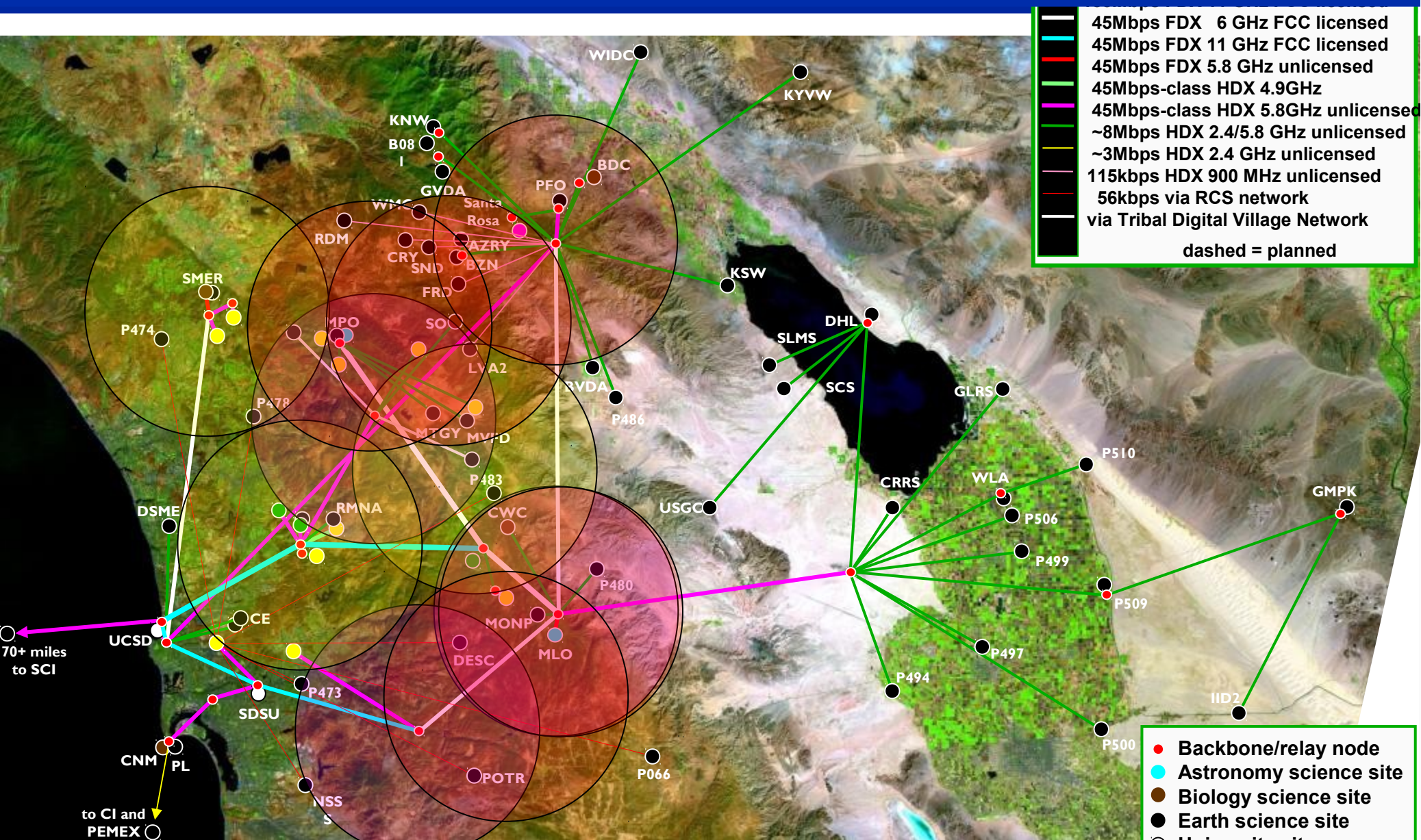
High Performance Wireless Research and Education Network

http://hpwren.ucsd.edu/

National Science Foundation awards 0087344, 0426879 and 0944131

HPWREN UCSD SDSC SAN DIEGO SUPERCOMPUTER CENTER it² SAN DIEGO STATE UNIVERSITY NATIONAL SCIENCE FOUNDATION

HPWREN Topology, 360 Degree Cameras

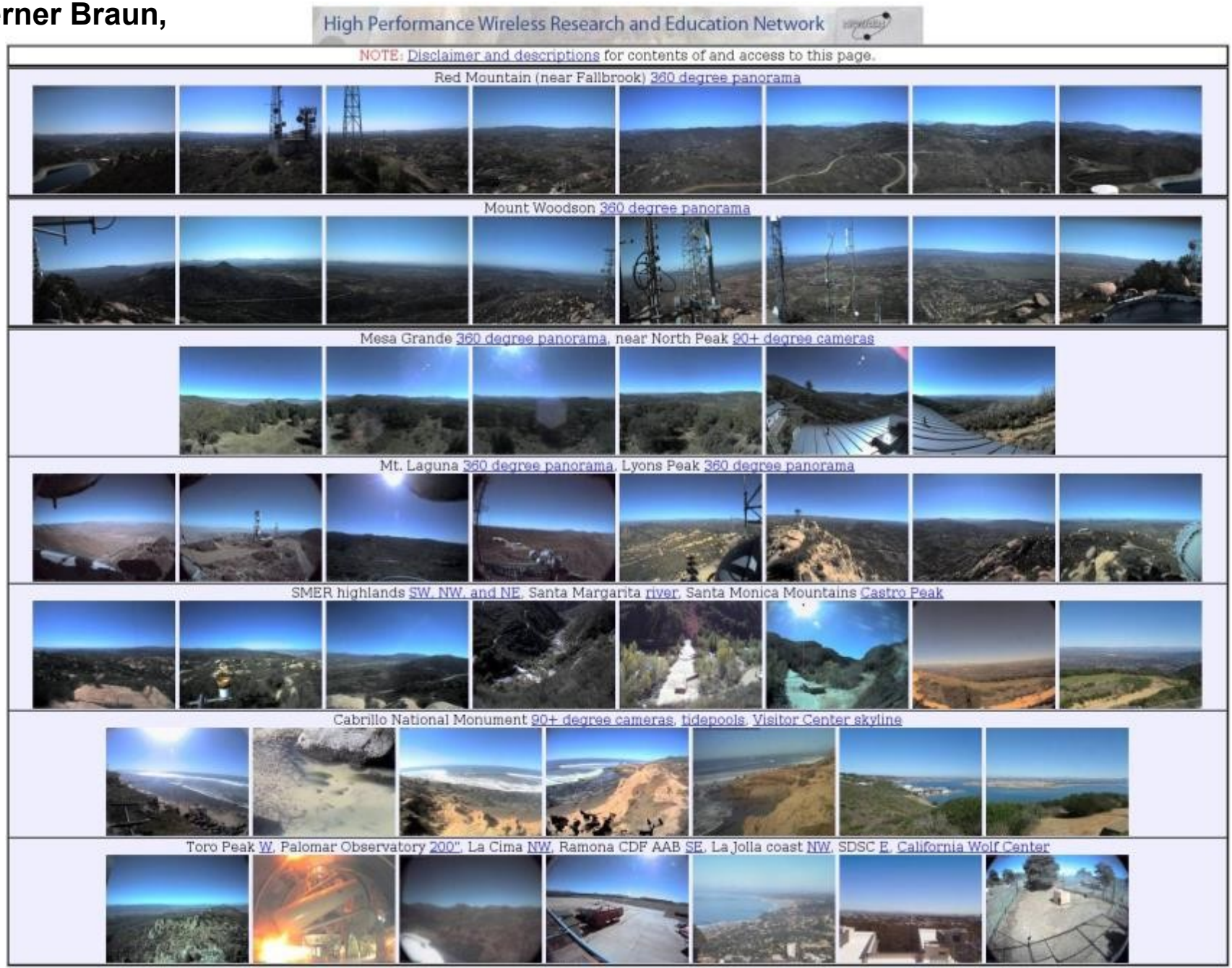


Source: Hans Werner Braun, HPWREN PI

Note: locations are approximate

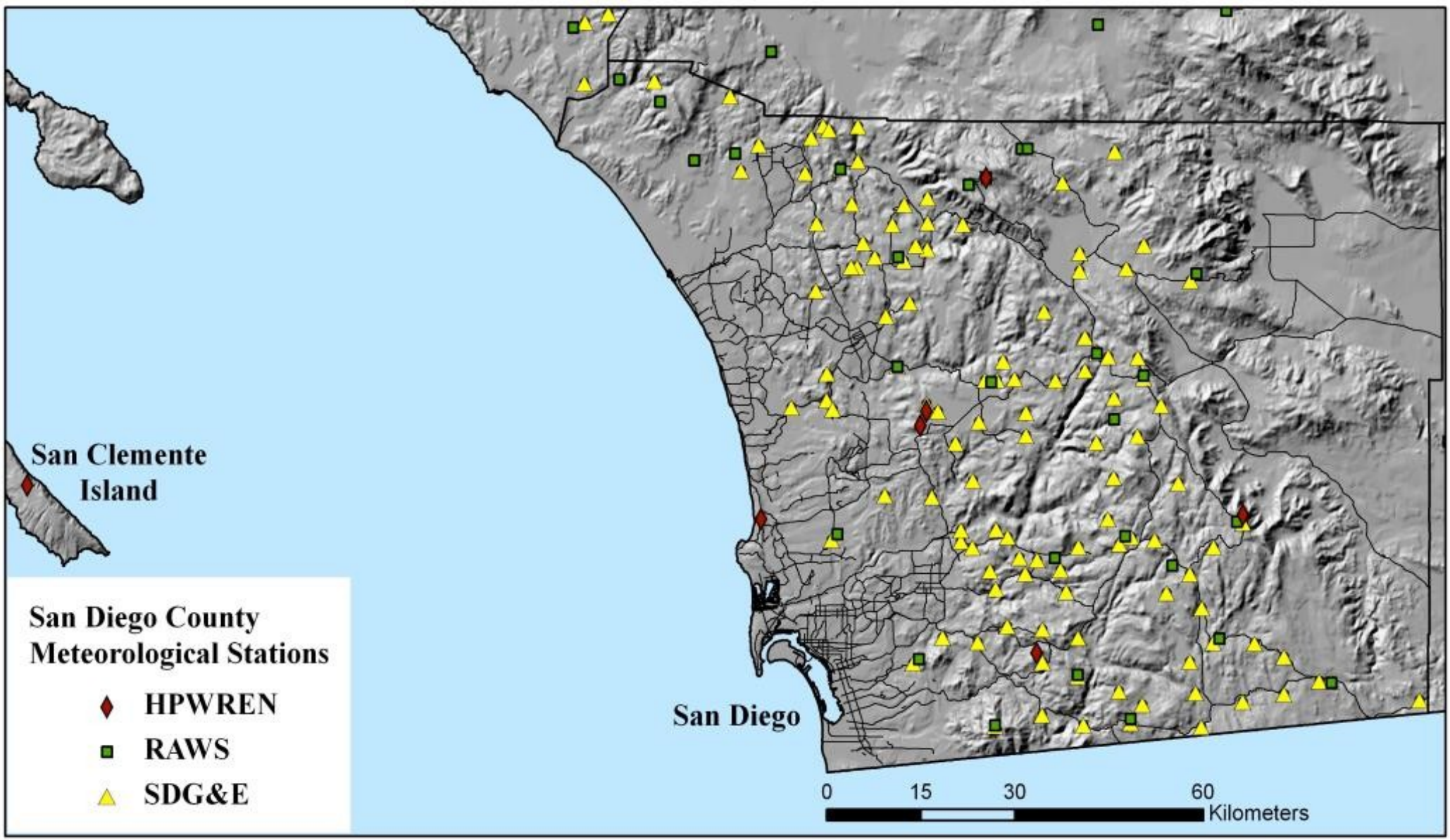
Various Real-Time Network Cameras for Environmental Observations

Source: Hans Werner Braun, HPWREN PI

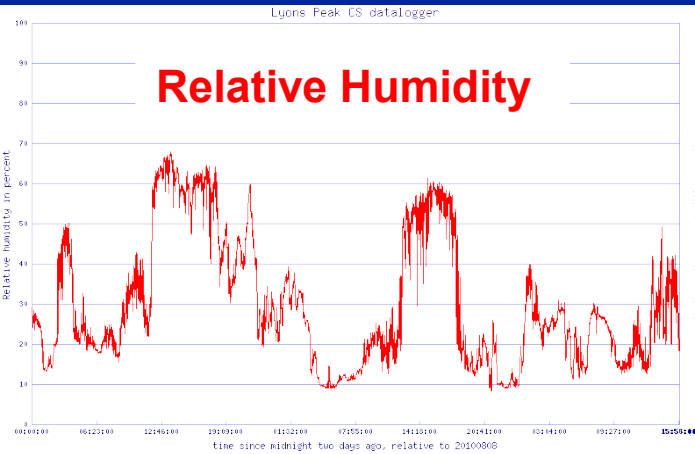


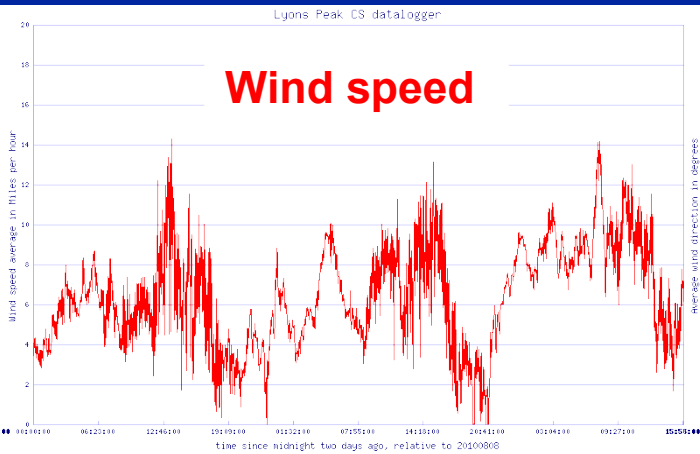
San Diego County Digital Weather Stations:
High Spatial Density Reads Out Time-Changing Atmosphere

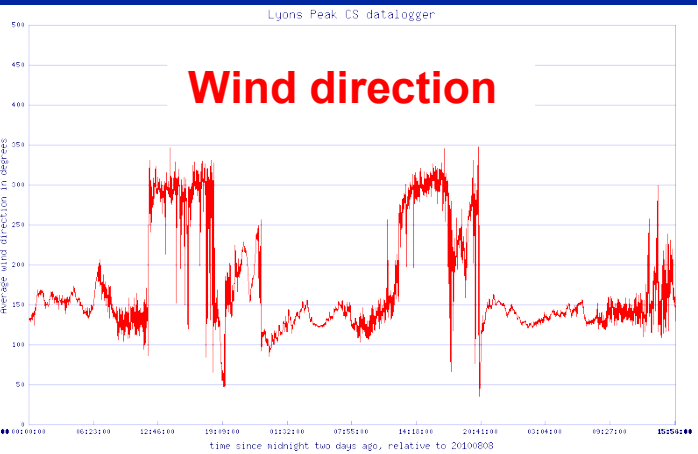
Source: Jessica Block, Calit2



Real-Time Data-Based Alerts to First Responders

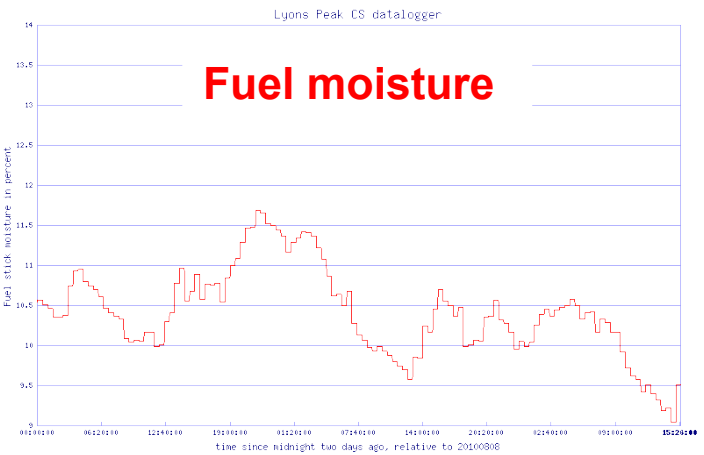






Trigger real-time computer-generated alerts, if:

condition "A" **AND** condition "B" **AND** condition "C"
OR condition "D"



exists, in which case several San Diego emergency officers are being paged or emailed during such alert conditions, based on HPWREN data parameterization by a CDF Division Chief. This system has been in operation since 2004.

Date: Wed, 4 Aug 2010 09:31:05 -0700 **Source: Hans Werner Braun, HPWREN PI**
Subject: URGENT weather sensor alert

LP: RH=26.1 WD=135.2 WS=1.9 FM=6.8 AT=80.7 at 20100804.093100
More details at http://hpwren.ucsd.edu/Sensors/

I Arrived in La Jolla in 2000 After 20 Years in the Midwest and Decided to Move Against the Obesity Trend

Age 41



Age 51



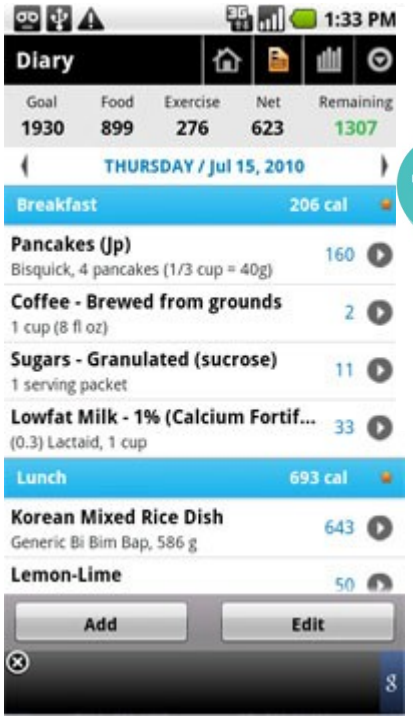
Age 61



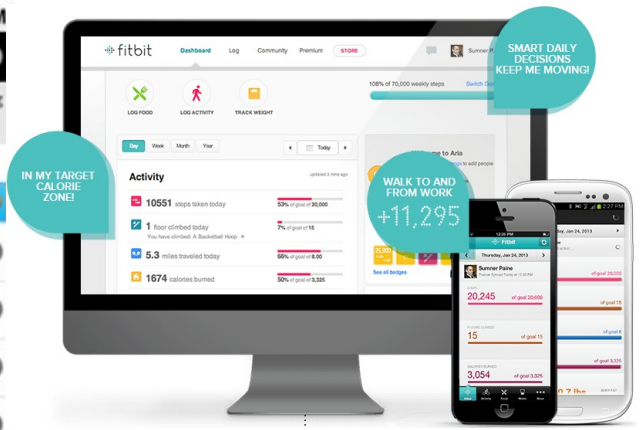
I Reversed My Body's Decline By Quantifying and Altering Nutrition and Exercise

http://lsmarr.calit2.net/repository/LS_reading_recommendations_FiRe_2011.pdf

I Used a Variety of Emerging Personal Sensors To Quantify My Body & Drive Behavioral Change



MyFitnessPal-
Calories Ingested



FitBit -
Daily Steps &
Calories Burned



Withings/iPhone-
Blood Pressure





Withings WiFi Scale -
Daily Weight

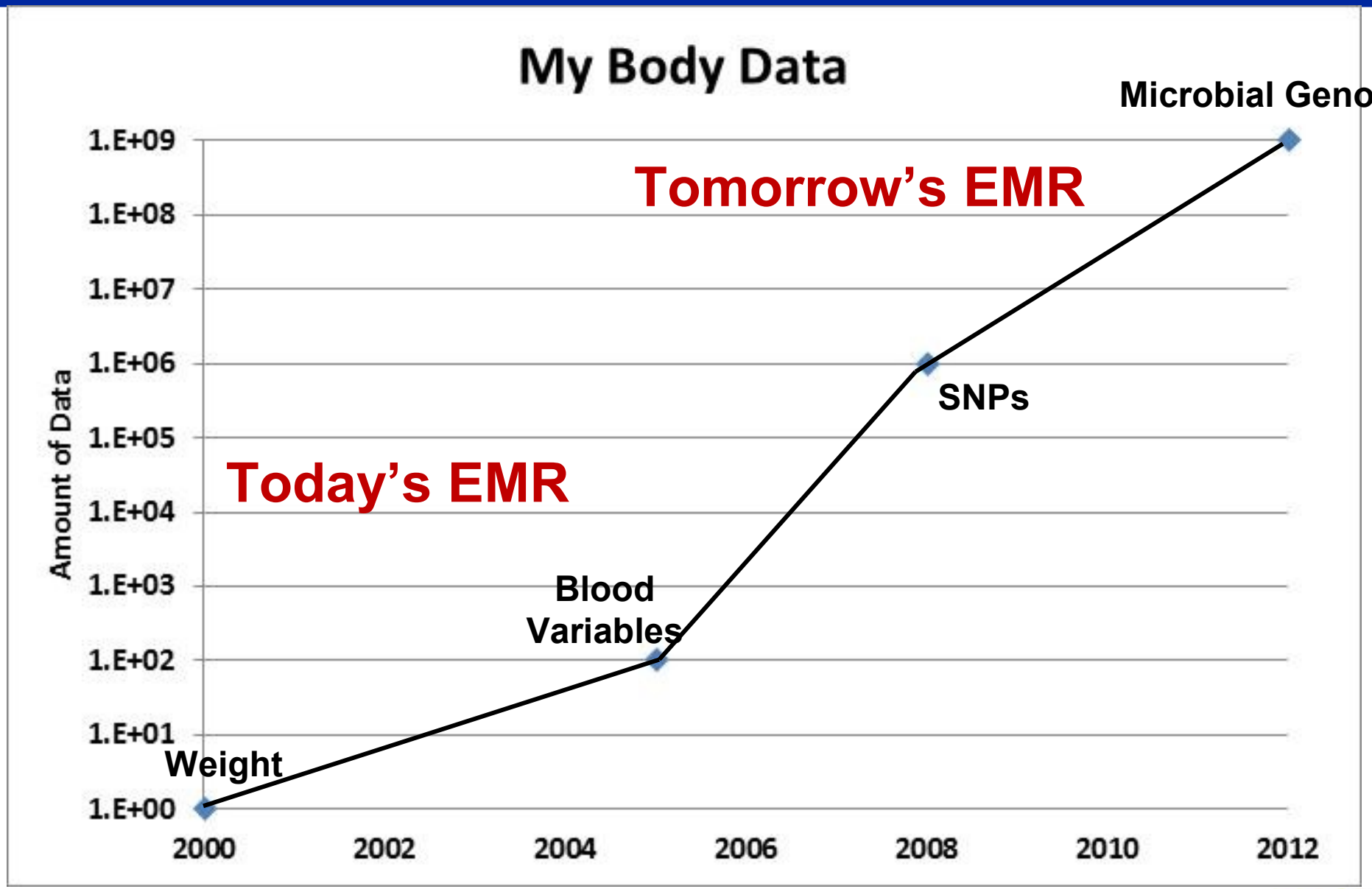


Zeo-Sleep

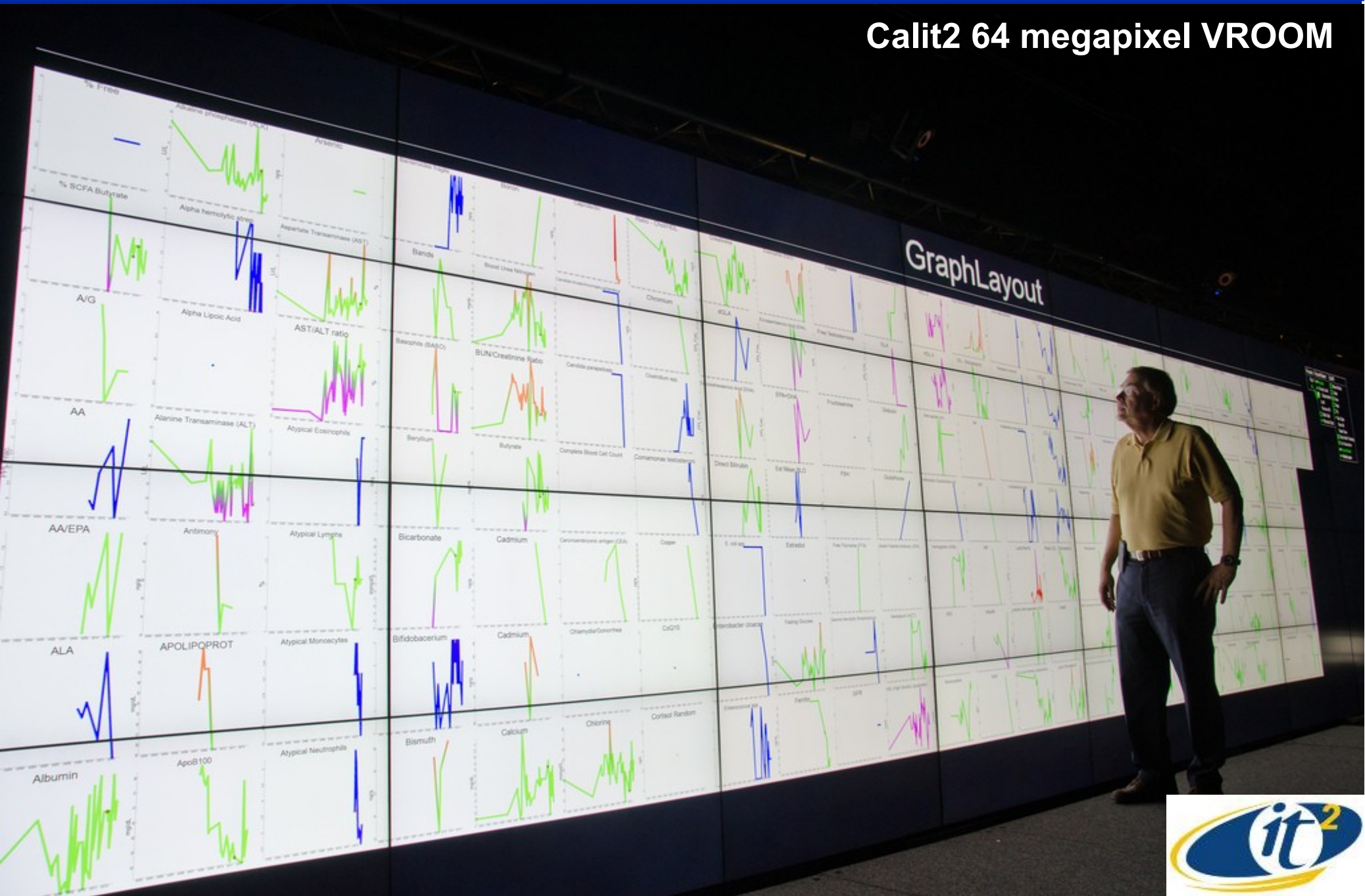


Azumio-Heart Rate

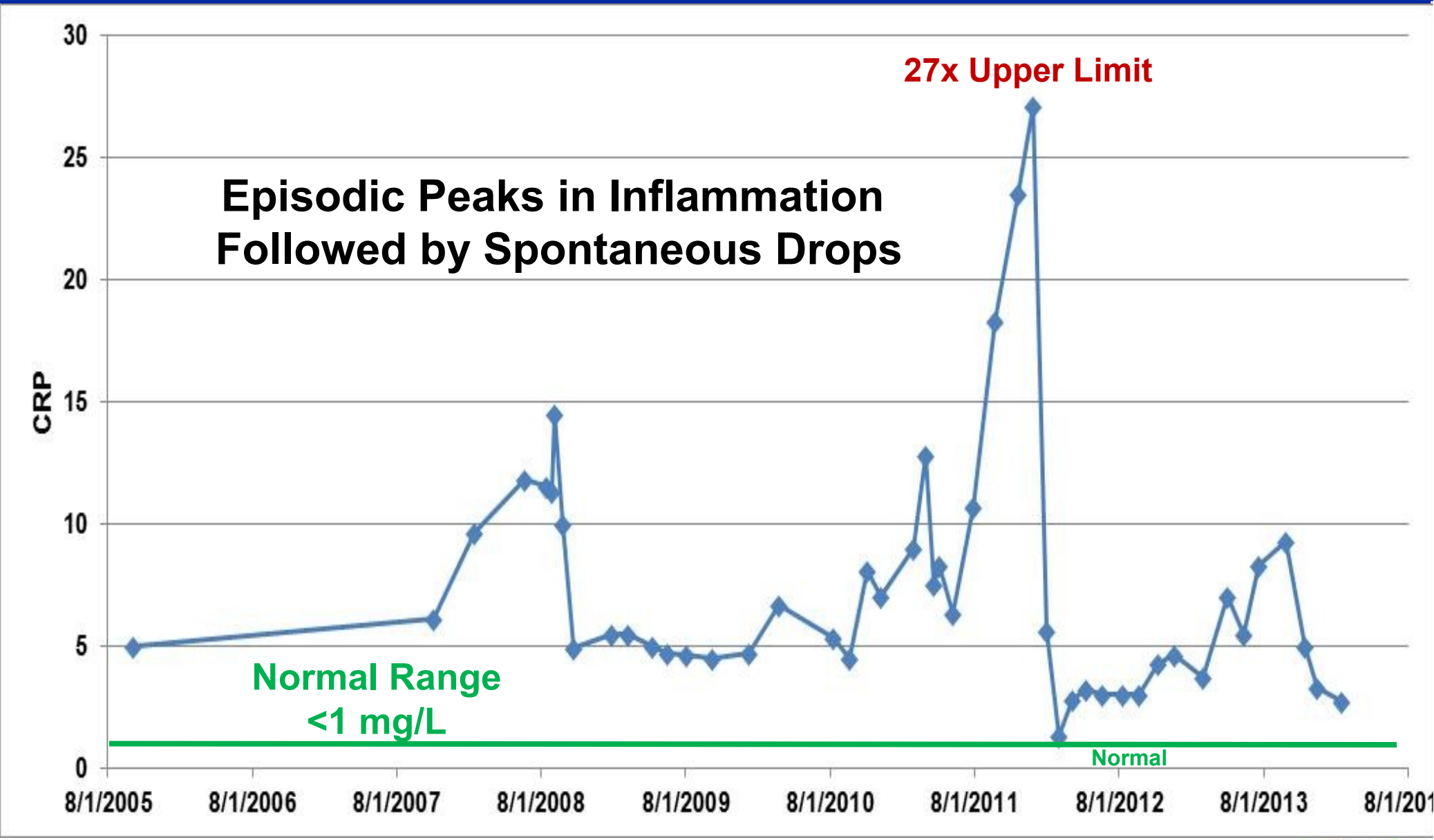
From One to a Billion Data Points Defining Me: Big Data Coming to the Electronic Medical Record (EMR)



Visualizing Time Series of 150 LS Blood and Stool Variables, Each Over 5-10 Years



Only One of My Blood Measurements Was Far Out of Range--Indicating Chronic Inflammation



Complex Reactive Protein (CRP) is a Blood Biomarker for Detecting Presence of Inflammation

Consumer Self Measurement is Exploding Totally Outside of the Medical Complex

From the First San Francisco QS Meetup in 2008
To 116 Cities in 37 Countries in Four Years



Groups	Members	Interested	Cities	Countries
157	27,814	7,373	116	37

The Self-Monitoring Business Has Reached Market Takeoff

- **MyFitnessPal**
  - 40 Million Users
  - Aug 2013 Raised $18M Series A, Led by Kleiner Perkins
- **Fitbit**
  - Has Raised ~$70M
- **BodyMedia Was Bought by Jawbone**
  - For ~$100M
- **Zeo Sleep Monitor**
  - Closed Down in 2013

More Mergers Likely as the Shakeout Continues

mHealth Technology Progression

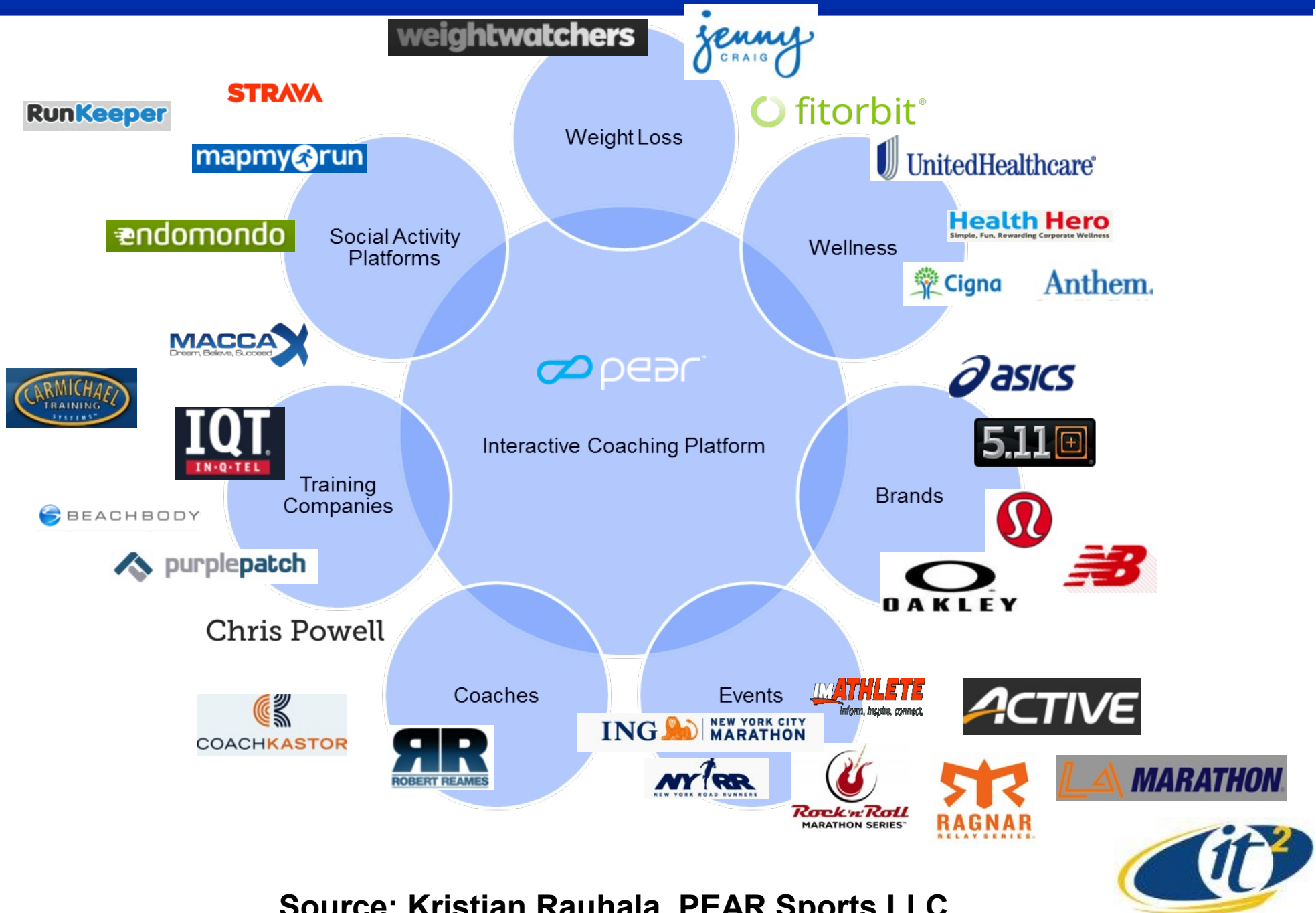


Mobile Health Market Projected to be $30B-$60B by 2015

Source: Rick Valencia, Qualcomm Life

QUALCOMM LIFE
a Qualcomm company

Platforms Enable Expanding Ecosystems Empowering Many to Serve Diverse Customer Sets



Source: Kristian Rauhala, PEAR Sports LLC