“The Pacific Research Platform: Building a Distributed Big-Data Machine-Learning Cyberinfrastructure”

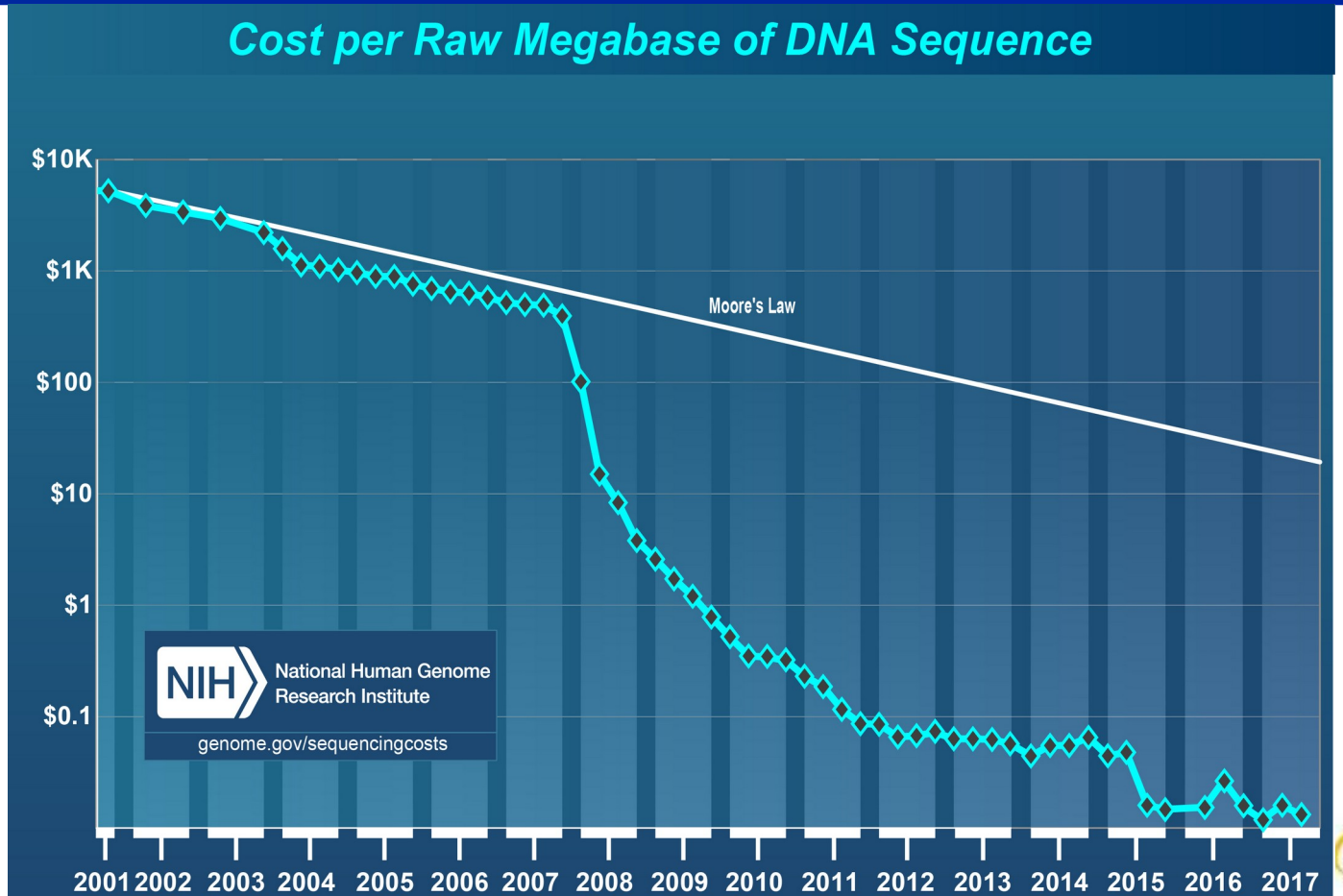
Briefing
Chancellor’s Council
University of California San Diego
May 13, 2019

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

QUALCOMM INSTITUTE
UC San Diego

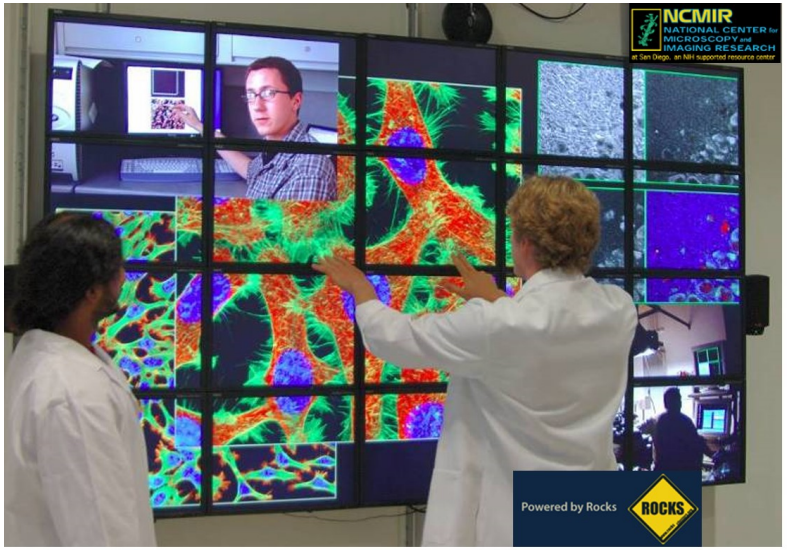
The Unrelenting Exponential Decrease in Cost of Generating Data Has Led to the Need for a Big Data Cyberinfrastructure

One Million Times Cheaper

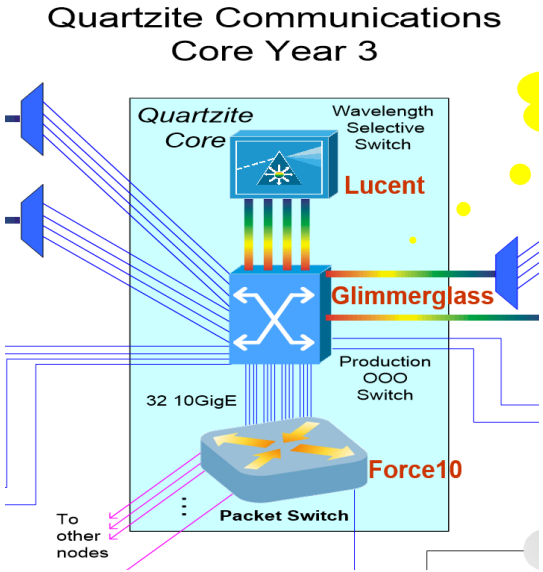


QUALCOMM INSTITUTE
UC San Diego

UC San Diego's Calit2 & SDSC Have Pioneered Big-Data Cyberinfrastructure for 17 Years 2002-2009: OptIPuter and Quartzite



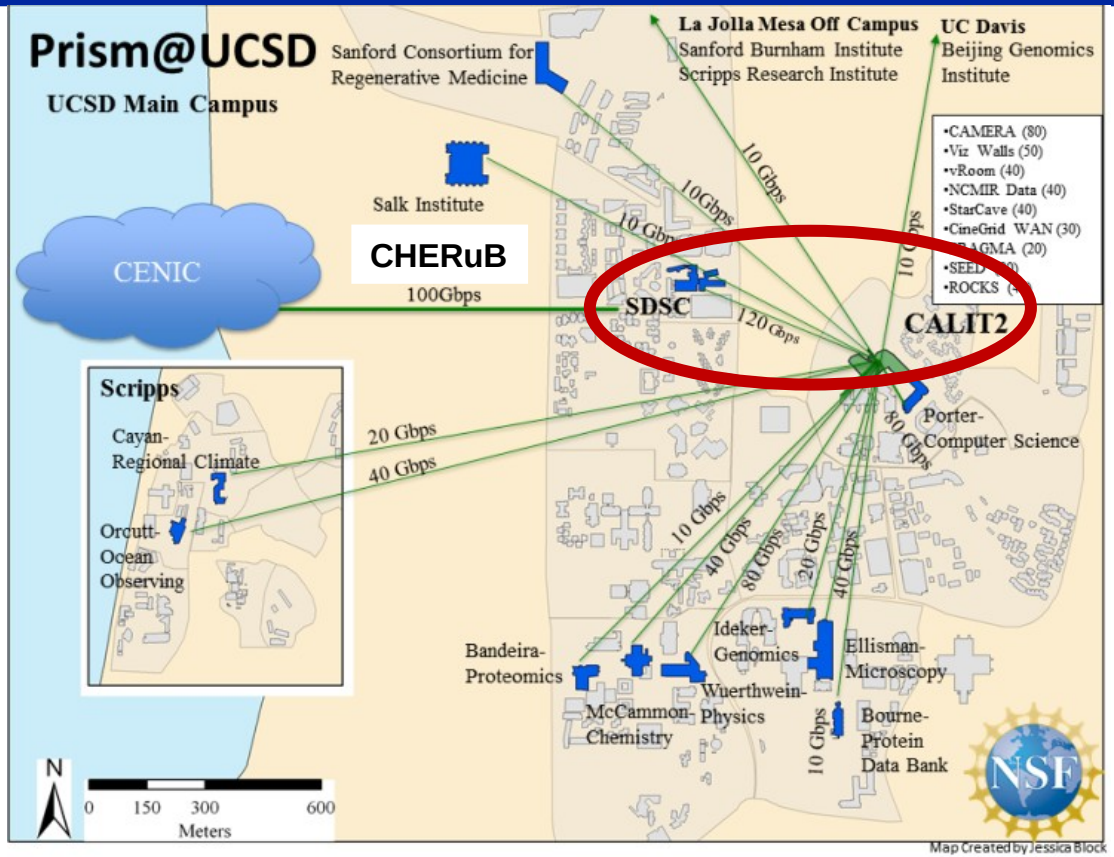
Calit2 (UCSD, UCI), SDSC, and UIC Leads—Larry Smarr PI
Univ. Partners: NCSA, USC, SDSU, NW, TA&M, UvA, SARA, KISTI, AIST
Industry: IBM, Sun, Telcordia, Chiaro, Calient, Glimmerglass, Lucent



OptIPuter
$13.5M
PI Smarr,
Co-PI DeFanti
Co-PI Papadopoulos, Ellisman
2002-2009

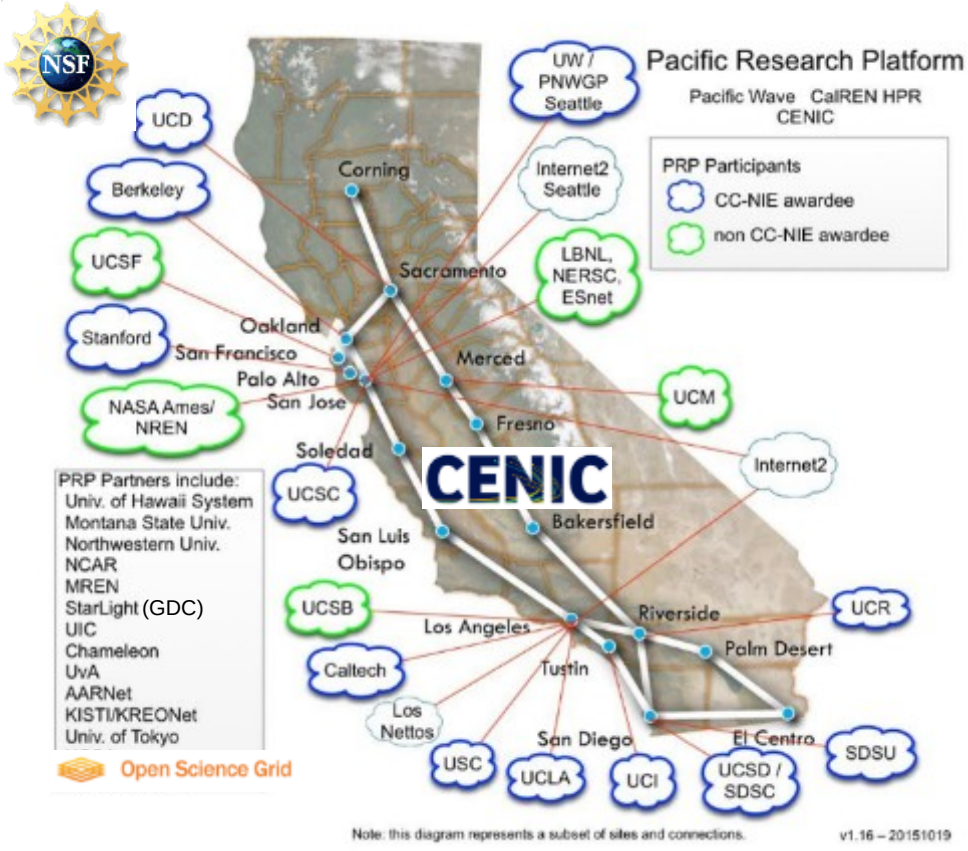
Quartzite
$1.2M
PI Papadopoulos,
Co-PI Smarr
2004-2007

2013-2015: Creating a "Big Data" Backplane on Campus: NSF Funded Prism@UCSD and CHERuB



Prism@UCSD, $500,000, Phil Papadopoulos, SDSC, Calit2, PI; Smarr co-PI
CHERuB, $500,000, Mike Norman, SDSC PI

2015-2020: The Pacific Research Platform Connects Campus "Big Data Freeways" to Create a Regional End-to-End Science-Driven "Big Data Superhighway" System



Source: John Hess, CENIC

NSF CC*DNI Grant
$6M 10/2015-10/2020

PI: Larry Smarr, UC San Diego Calit2

Co-PIs:

- **Camille Crittenden, UC Berkeley CITRIS,**
- **Tom DeFanti, UC San Diego Calit2/QI,**
- **Philip Papadopoulos, UCSD SDSC,**
- **Frank Wuerthwein, UCSD Physics and SDSC**

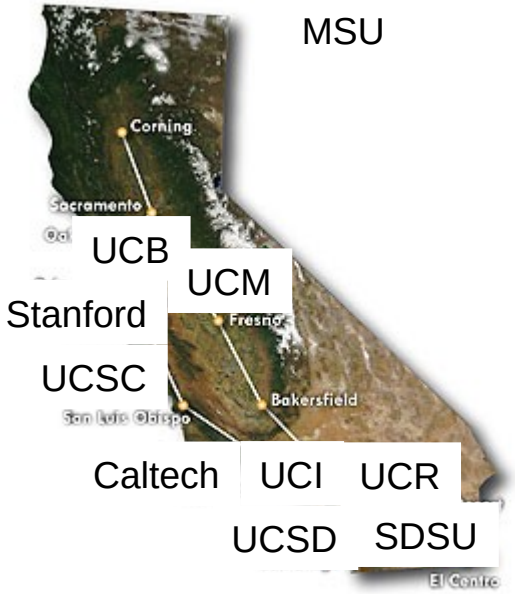
Letters of Commitment from:

- 50 Researchers from 15 Campuses
- 32 IT/Network Organization Leaders

UCOP CIO Tom Andriola
Provided Funds and ITLC Support
for Using Ten UC Campuses
For Advanced Technology Testing

UC San Diego
SDSC
CITRIS BANATAO INSTITUTE

2017-2020: CHASE-CI Adds Machine-Learning to the Data-Science Community Cyberinfrastructure



CI-New: Cognitive Hardware and Software Ecosystem Community Infrastructure (CHASE-CI)

For the Period September 1, 2017 – August 31, 2020

SUBMITTED – January 18, 2017

PI: Larry Smarr, Professor of Computer Science and Engineering, Director Calit2, UCSD
Co-PI: Tajana Rosing, Professor of Computer Science and Engineering, UCSD
Co-PI: Ken Kreutz-Delgado, Professor of Electrical and Computer Engineering, UCSD
Co-PI: Ilkay Altintas, Chief Data Science Officer, San Diego Supercomputer Center, UCSD
Co-PI: Tom DeFanti, Research Scientist, Calit2, UCSD

NSF Grant for 256 High Speed "Cloud" GPUs
For 32 ML Faculty & Their Students at 10 Campuses
To Train AI Algorithms on Big Data

QUALCOMM INSTITUTE
UC San Diego

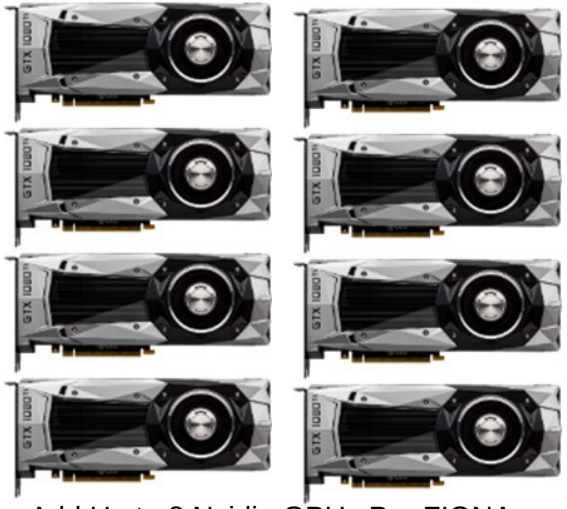
SDSC UC San Diego

PRP Engineers Designed and Built Several Generations of Optical-Fiber Big-Data Flash I/O Network Appliances (FIONAs)

UCSD-Designed FIONAs Solved the Disk-to-Disk Data Transfer Problem *at Near Full Speed* on Best-Effort 10G, 40G and 100G Networks



Two FIONA DTNs at UC Santa Cruz: 40G & 100G
Up to 200 TeraByte Rotating Storage



Add Up to 8 Nvidia GPUs Per FIONA
To Add Machine Learning Capability



FIONette—
1G, $250
Used for
Training 50
Engineers in
2018-2019

Over 100 FIONAs Now Deployed on PRP

FIONAs Designed by UCSD's Phil Papadopoulos, John Graham, Joe Keefe, and Tom DeFanti

QUALCOMM INSTITUTE
UC San Diego

NSF

SDSC

it²

UCSD Adding >350 Game GPUs to Data Sciences Cyberinfrastructure - Devoted to Data Analytics and Machine Learning

Toward an "AI University"

UC San Diego

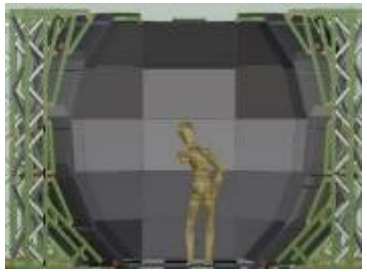
IT SERVICES

108 GPUs for Students

32 GPUs for Research ECE Dept

Open Science Grid

48 GPUs for OSG Applications



SunCAVE 70 GPUs

UCSD Cognitive Science

JACOBS SCHOOL OF ENGINEERING

HALICIOĞLU DATA SCIENCE INSTITUTE

SDSC it²

QUALCOMM INSTITUTE



Plus 288 64-bit GPUs On SDSC's Comet



FIONA with 8-Game GPUs

CHASE-CI Grant :
96 GPUs at UCSD
for Training AI Algorithms on Big Data

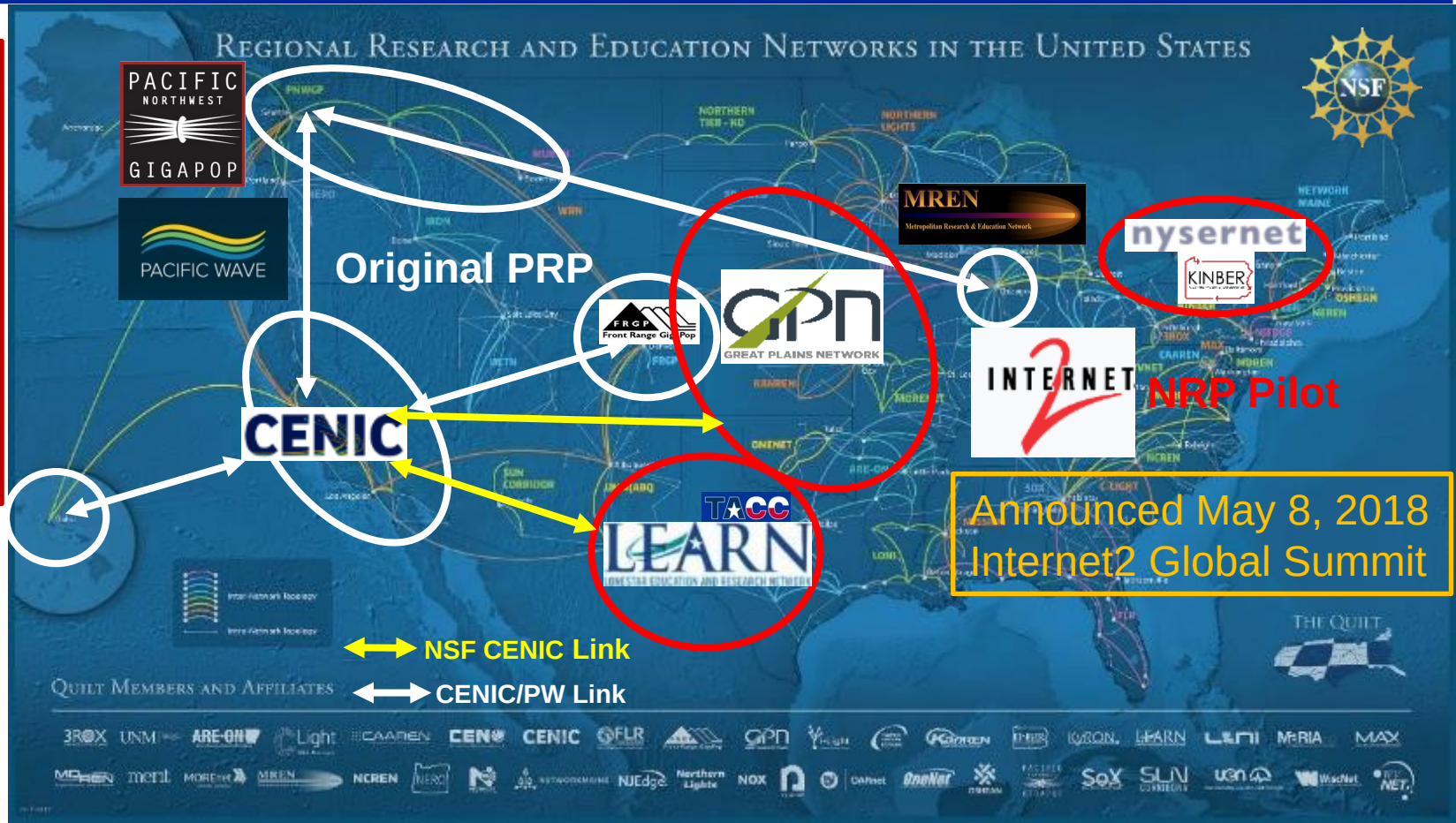
QUALCOMM INSTITUTE
UC San Diego

it²

2018-2019: National-Scale Pilot - Using CENIC & Internet2 to Connect Quilt Regional R&E Networks

"Towards The NRP"
3-Year Grant Funded by NSF
$2.5M
October 2018

PI Smarr
Co-PIs Altintas Papadopoulos Wuerthwein Rosing



SDSC

QUALCOMM INSTITUTE
UC San Diego

2018-2019: PRP Game Changer!
Using Kubernetes to Orchestrate Containers Across the PRP

CADE METZ BUSINESS 06.10.14 01:15 PM

GOOGLE OPEN SOURCES ITS SECRET WEAPON IN CLOUD COMPUTING

WIRED

"Kubernetes is a way of stitching together a collection of machines into, basically, a big computer,"
--Craig Mcluckie, Google
and now CEO and Founder of Heptio

"Everything at Google runs in a container."
--Joe Beda,Google

How Kubernetes Conquered 2017 (and is Positioned for 2018)





Dan Meyer
December 25, 2017

sdx central

Kubernetes, an open-source software project that started at Google, has exploded in popularity and is now used by at least 54% of the Fortune 500. BUSINESS INSIDER **Rosalie Chan** Jan. 27, 2019,

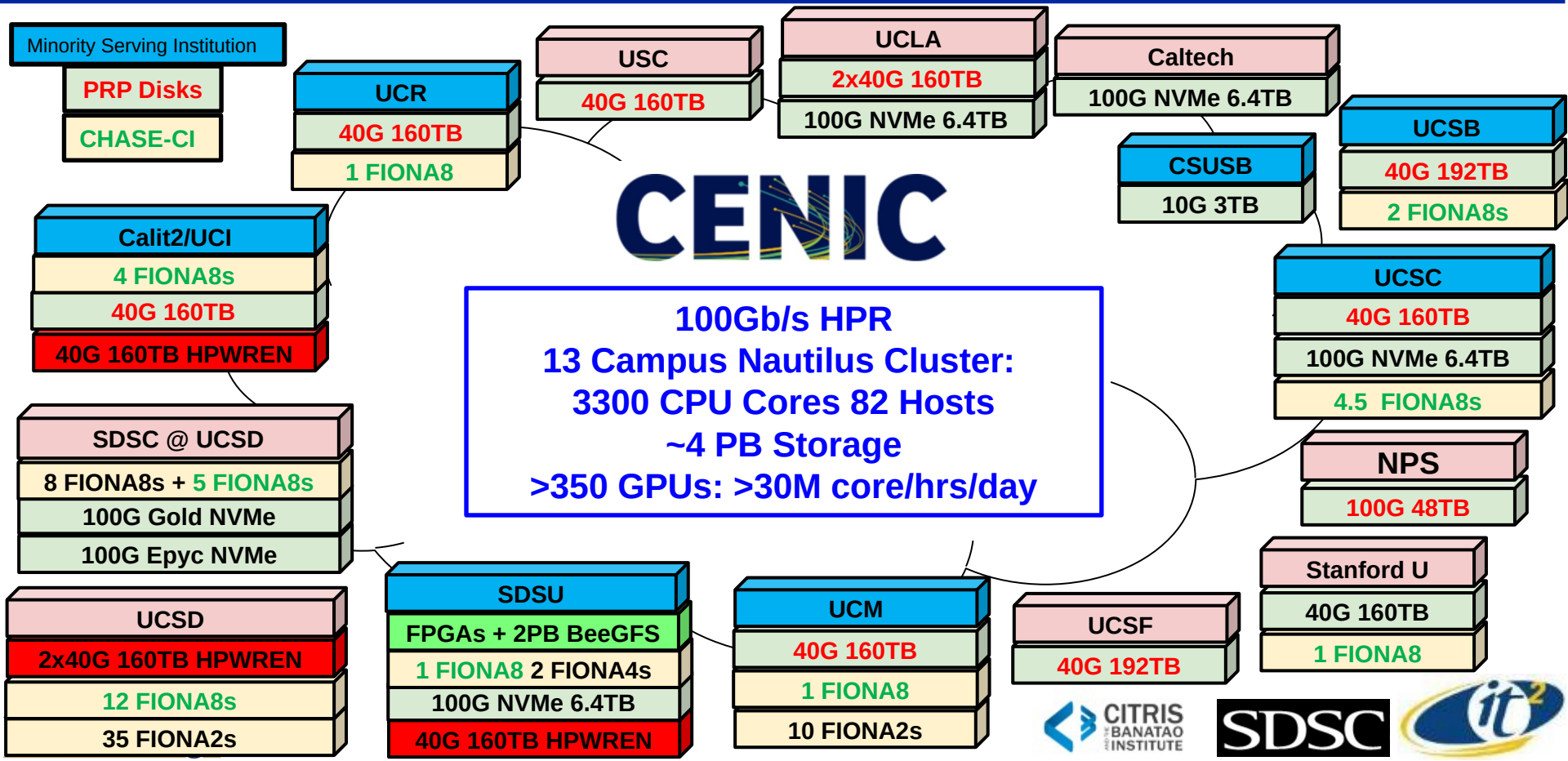
QUALCOMM INSTITUTE
UC San Diego

CLOUD NATIVE COMPUTING FOUNDATION

SDSC

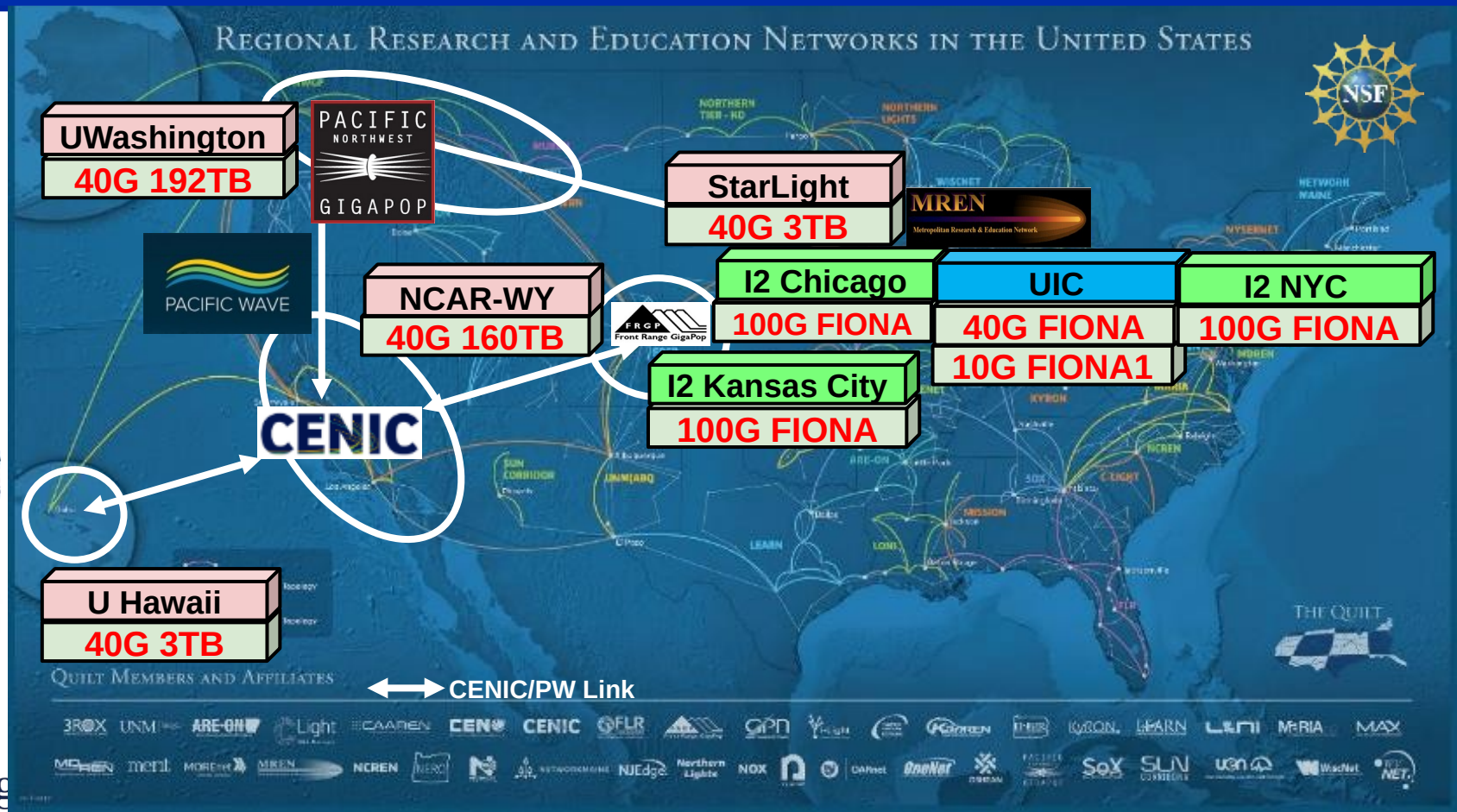
it²

California-Connected by CENIC
PRP Kubernetes "Nautilus" Hypercluster



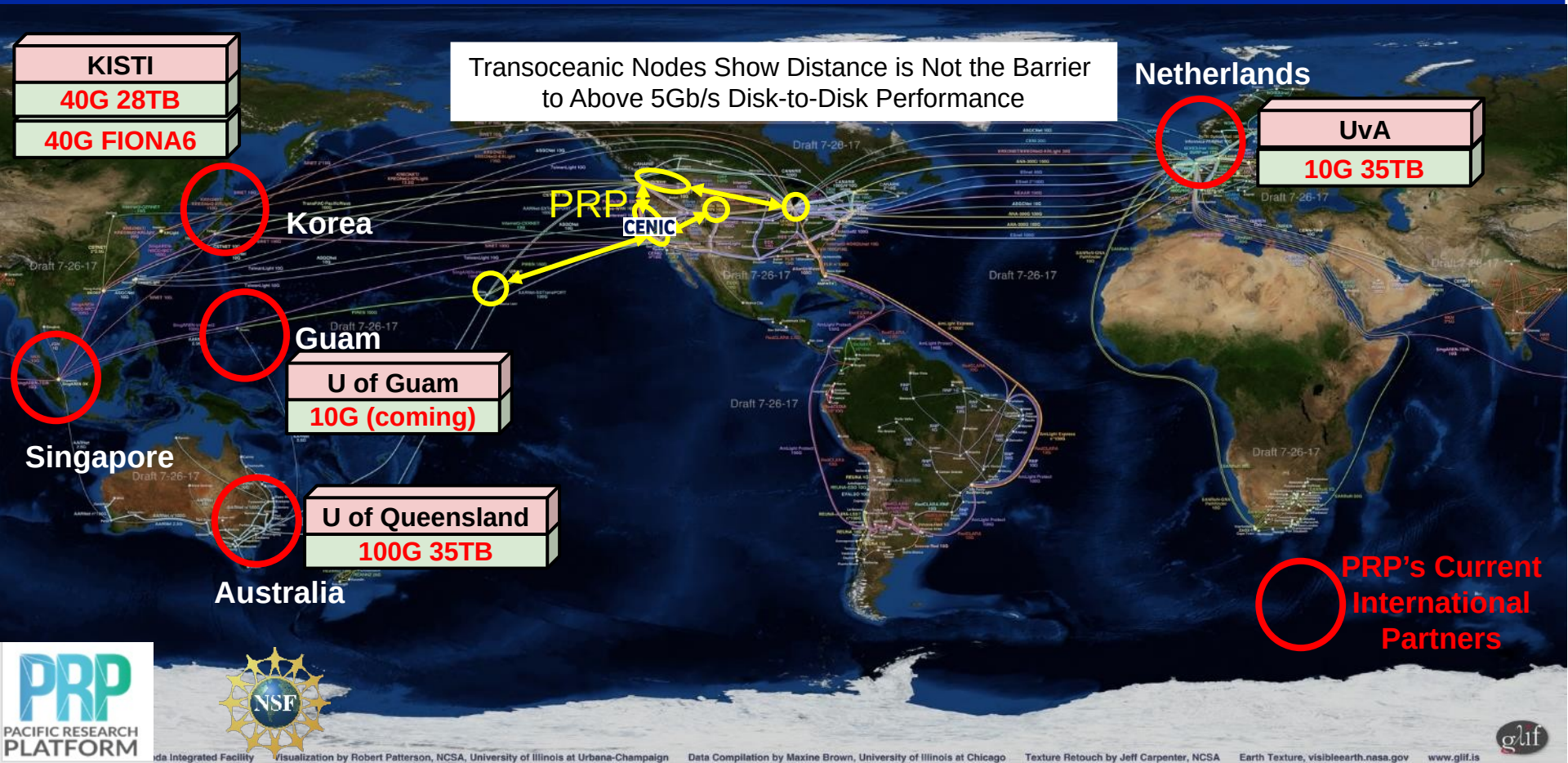
United States PRP/TNRP Nautilus Hypercluster
Now Connects 3 More Regionals and 3 Internet2 Sites

Nautilus



ESnet
CENIC
SDSC
CITRIS BANATAO INSTITUTE
QUALCOMM INSTITUTE
UC San Diego

Global PRP Nautilus Hypercluster Is Rapidly Adding International Partners Beyond Our Original Partner in Amsterdam



PRP is Science-Driven:
Connecting Multi-Campus Application Teams and Devices



Particle Physics

Biomedical 'omics

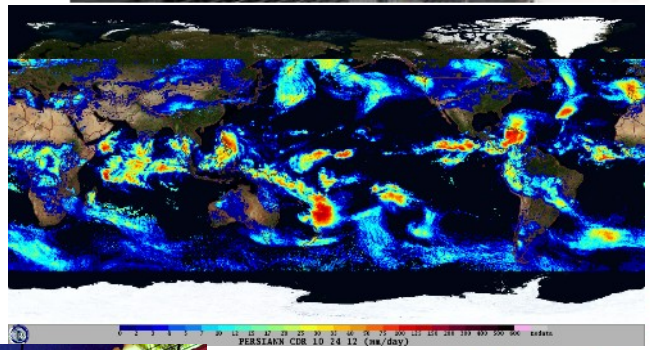


NSF

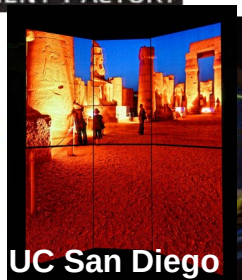


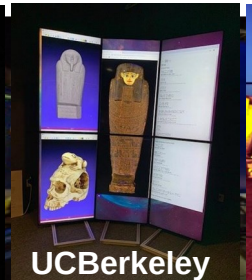
Telescope Surveys

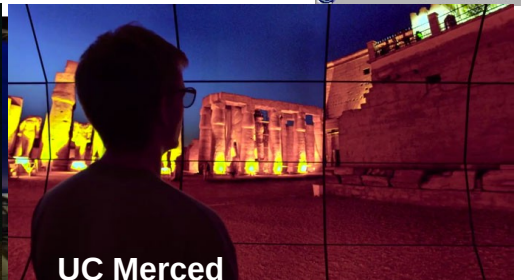
Earth Sciences



Visualization, Virtual Reality, Collaboration





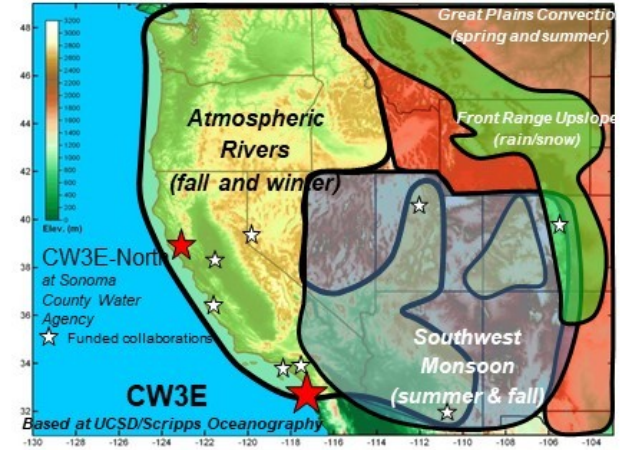


CITRIS THE BAN INST

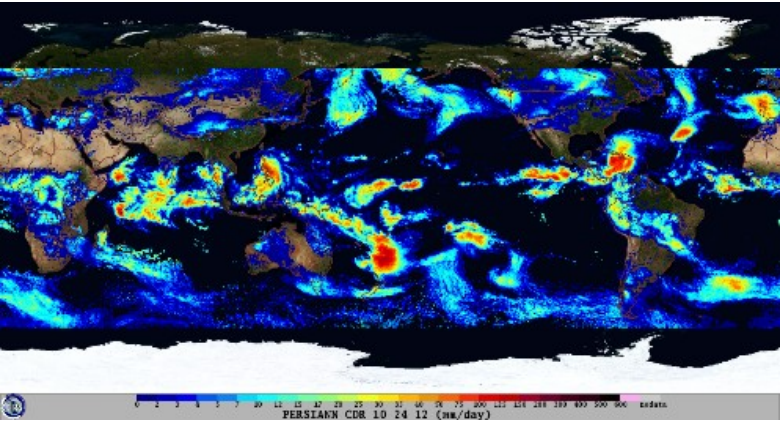
SDSC

it²

Collaboration on Atmospheric Water in the West Between UC San Diego and UC Irvine



Key Phenomena Causing Extreme Precipitation in the Western U.S. (Ralph et al. 2014)



Center for Western Weather and Water Extremes
SCRIPPS INSTITUTION OF OCEANOGRAPHY AT UC SAN DIEGO

Director: **F. Martin Ralph**

Big Data Collaboration with:

Center for Hydrometeorology & Remote Sensing
University of California, Irvine

Director, Soroosh Sorooshian, UCSD

Source: Scott Sellers, PhD CHRS; Postdoc CW3E

PRP Sped Up Scott Sellar's Workflow by Over 500 Times!

Pacific Research Platform (10-100 Gb/s)

Complete Workflow Time: 19.2 Days→52 Minutes!

UC, Irvine

CENTER FOR HYDROMETEOROLOGY & REMOTE SENSING — CHRS — UCIrvine

GPUs



Calit2's FIONA

QUALCOMM INSTITUTE
UC San Diego

SDSC's COMET



Source: Scott Sellers, US State Dept.

UC, San Diego

GPUs



Calit2's FIONA

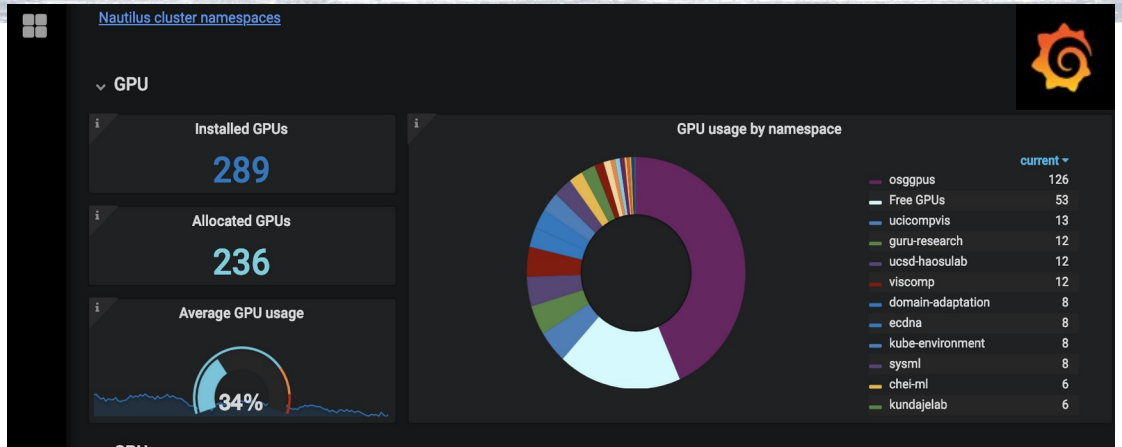
it²

OSG IceCube Usage on PRP (Purple Segment) 3/9/19: Using 126 GPUs + 142 CPUs + 49 GB RAM

GPU Simulations Needed *to Improve Ice Model*.
=> Results in Significant Improvement in Pointing Resolution for Multi-Messenger Astrophysics



IceCube

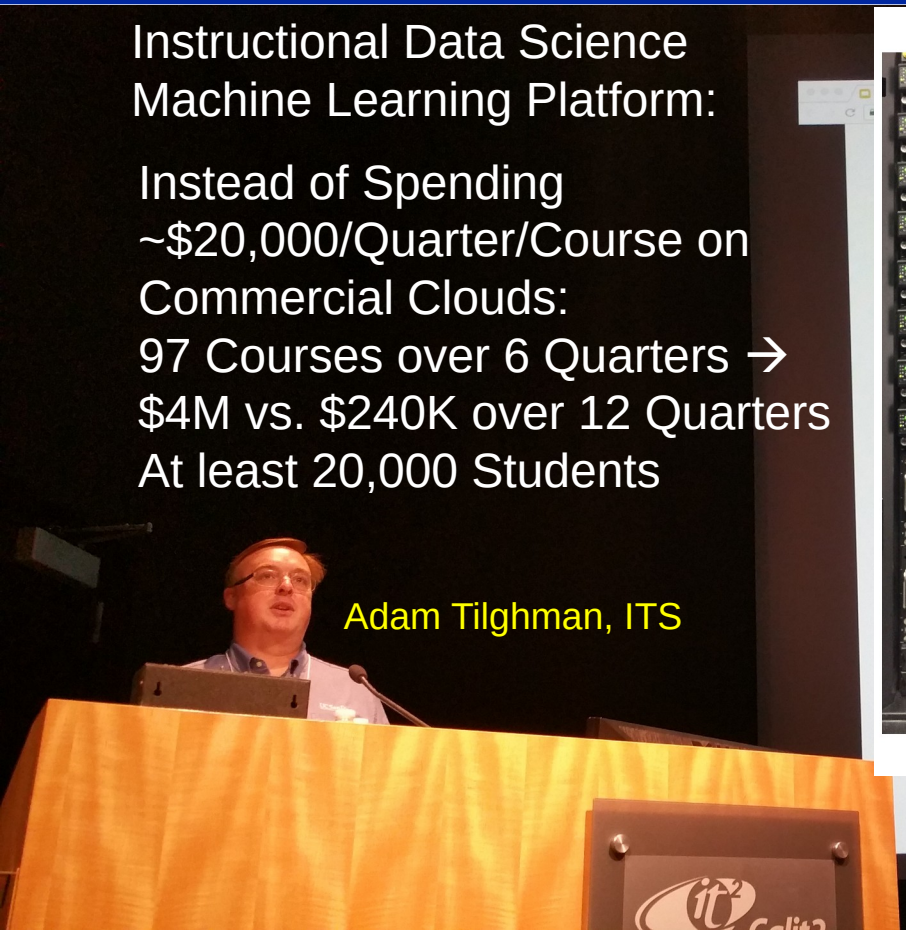


QUALCOMM INSTITUTE
UC San Diego

UCSD's ITS Adapted PRP FIONA8s To Support Data Science Courses

Instructional Data Science Machine Learning Platform:

Instead of Spending ~$20,000/Quarter/Course on Commercial Clouds:
97 Courses over 6 Quarters → $4M vs. $240K over 12 Quarters
At least 20,000 Students



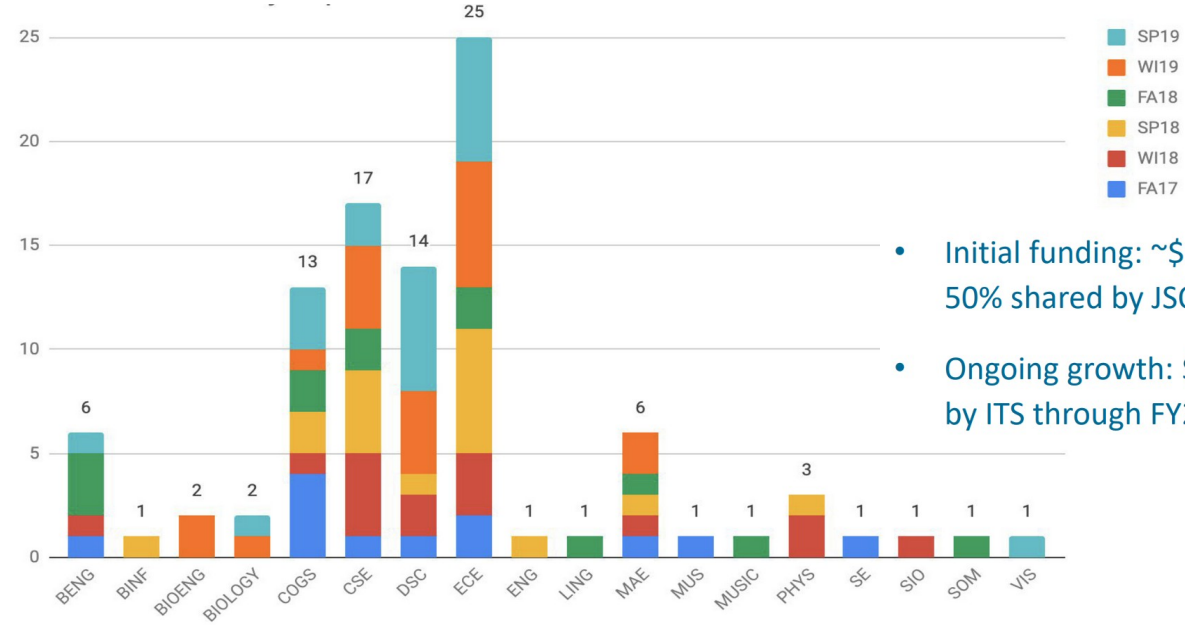


- 104 GPUs
  - 80 GTX 1080Ti
  - 16 RTX 2080Ti
  - 1.05 Petaflops
- 28 nodes
  - 544 CPU cores
  - 6.5TB RAM
- 40TB Flash Storage
- 10G networking



Source: UCSD ITS

The Student GPUs
Have Supported a Broad Set of Courses Across Campus



- Initial funding: ~$150K, 50% from ITS & 50% shared by JSOE and Cog Sci
- Ongoing growth: $50K/yr commitment by ITS through FY20/21

Selected Spring 2018 Courses

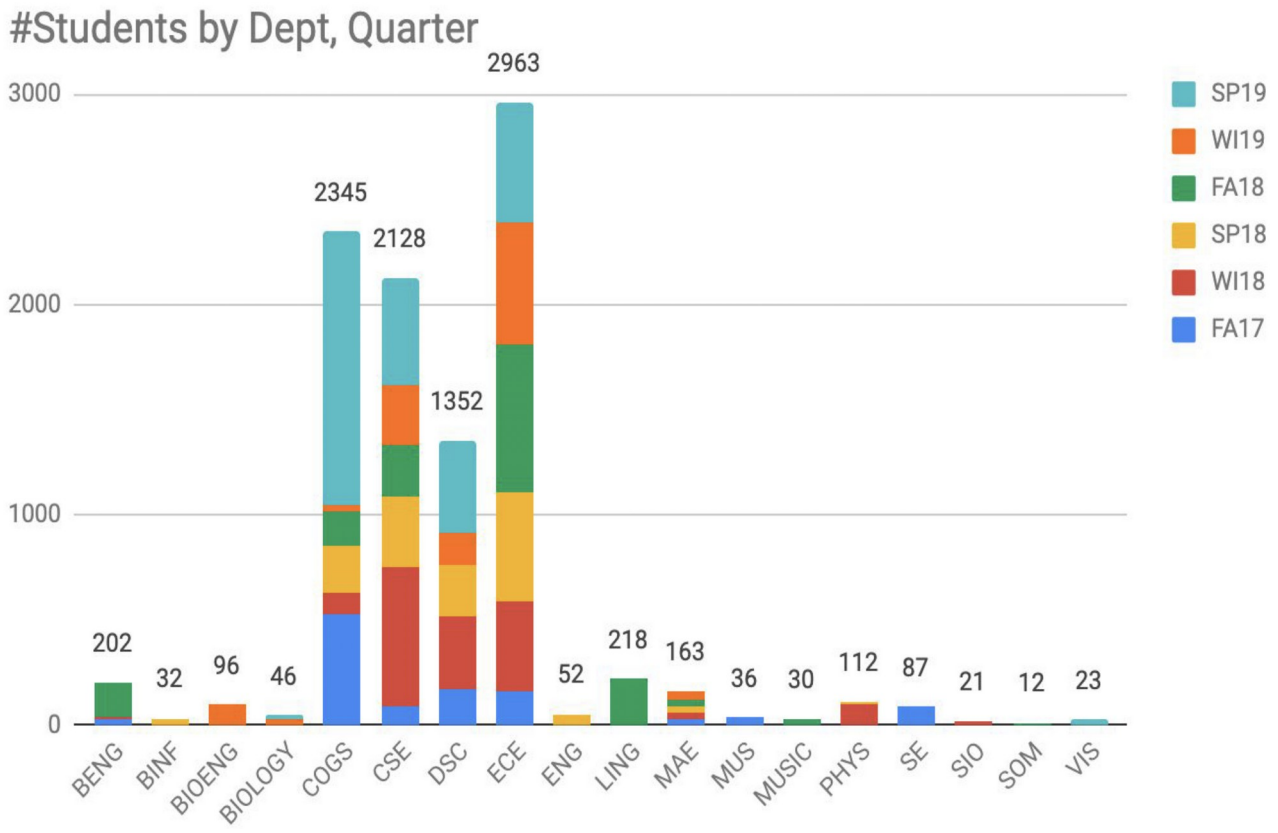
Advanced Statistical NLP
Vision & Learning
Comp Phys II:PDE & Matrix Modeling
Deep Learning and Applications
Robot Reinforcement Learning
Intelligent Vehicles and Assistance Systems
Intro to Autonomous Vehicles
Optimization and Acceleration of DL

QUALCOMM INSTITUTE
UC San Diego

Source: UCSD ITS

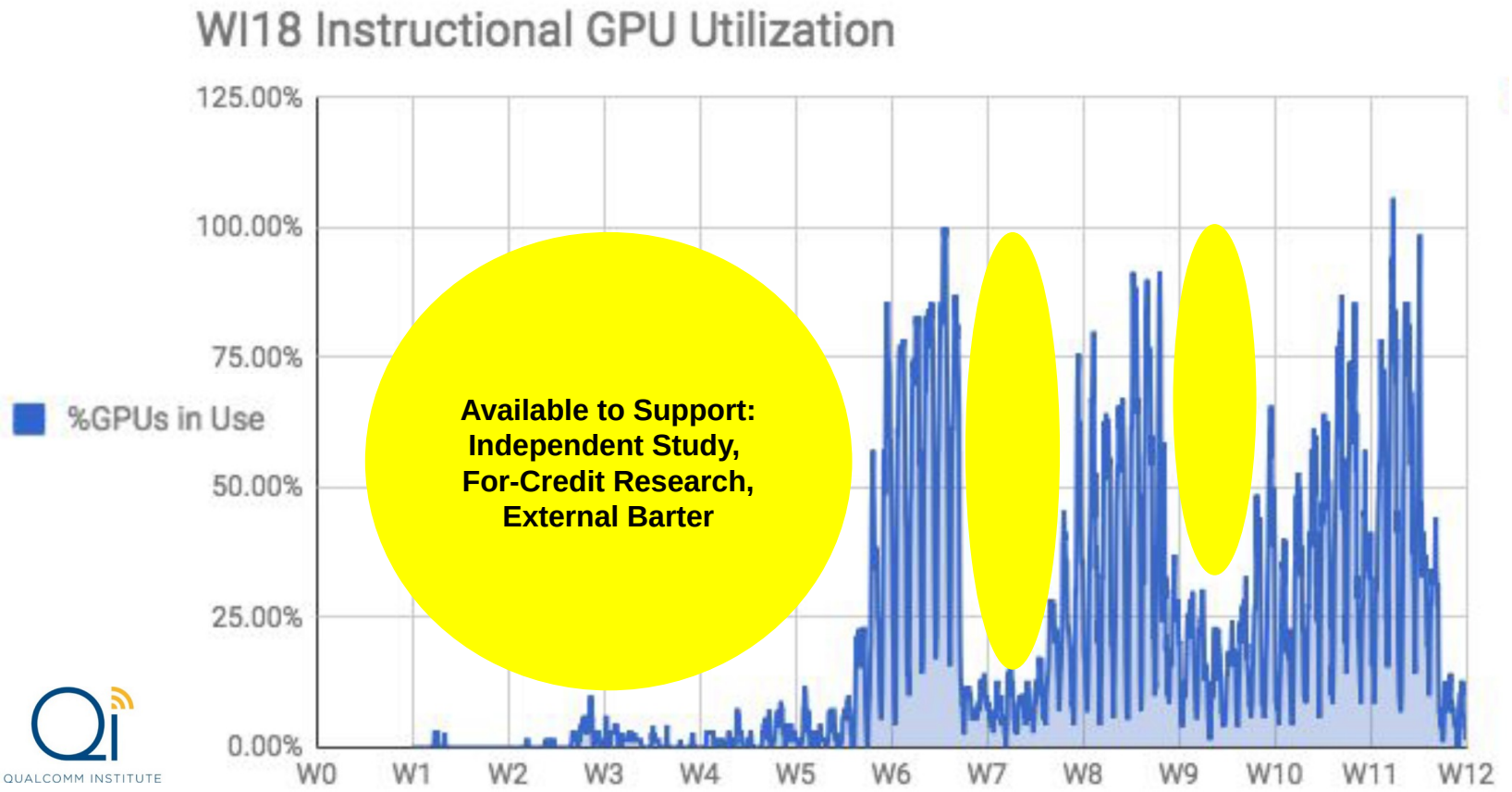
The ITS GPUs
Have Supported Thousands of Students

Source: UCSD ITS



QUALCOMM INSTITUTE
UC San Diego

Student GPU Demand Is Variable Allowing for Other Student Uses



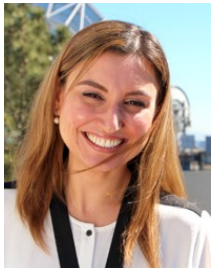
QUALCOMM INSTITUTE
UC San Diego

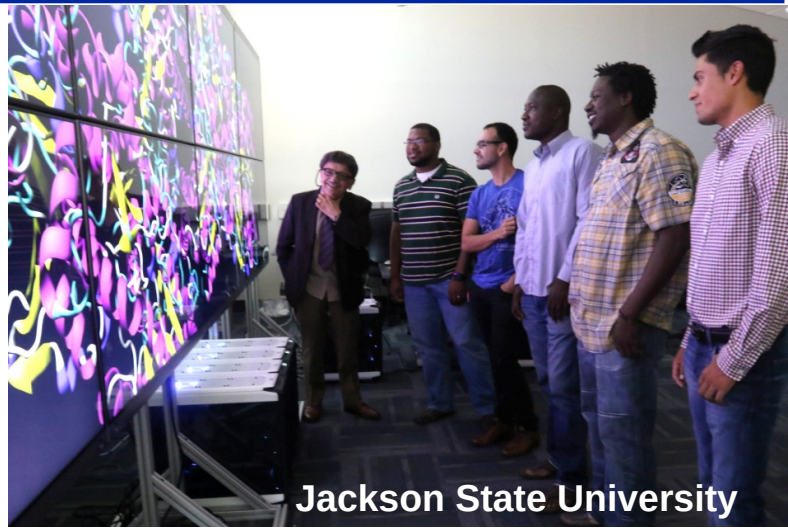
Source: UCSD ITS

PRP Actively Develops Diversity

- **Grants**
  - 3 Female co-PIs
  - 1 Hispanic co-PI
- **Campuses**
  - 8 Minority-Serving Institutions in PRP/CHASE-CI
- **Workshops**
  - NRPII Workshop Steering Committee 80% Female
  - Multiple MSI, EPSCoR Focused Workshops






Jackson State University




PRP MSI Workshop

Presenting FIONettes

QUALCOMM INSTITUTE
UC San Diego