“Introduction to UC San Diego’s Integrated Digital Infrastructure”

Opening Talk

IDI Showcase 2015

University of California, San Diego

May 6-7, 2015

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

IDI

What is IDI?

- **5 Partnering Units:**

The Library UC SAN DIEGO

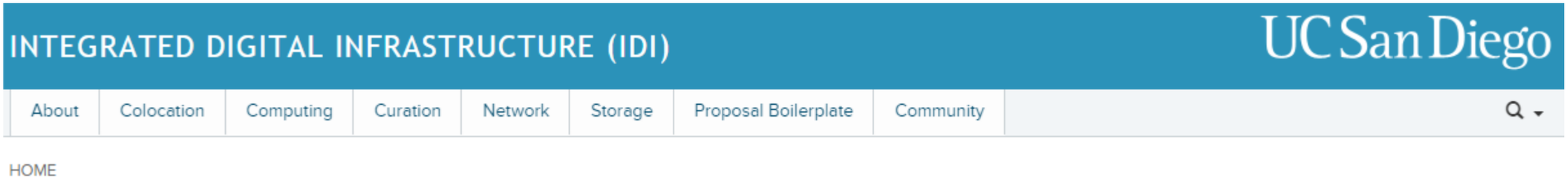
Administrative Computing & Telecommunications (ACT)

acms Academic Computing and Media Services

SDSC

QUALCOMM INSTITUTE

- **A Process to Support UCSD Strategic Plan With an Integrated Digital Infrastructure**
  - **Concierge service to Identify the Right Mix of Services to Meet Faculty and Research Staff Needs**
  - **Coordination Across Units to Ensure Research Needs are Met Efficiently and Effectively**
  - **Transformational Projects and Digital Research Platforms to Enhance Research and Education**
  - **Supporting a Research Data Library and Critical High Performance Cyberinfrastructure**



IDI



IDI Supporting UCSD's Strategic Plan

- **Transformational Projects & big data instructional support give our students hands-on experience with new technology and industry-standard tools (Goal 1)**
- **Collaboration & communications technology, including high-speed networking & electronic lab notebooks, improves interaction & reduces barriers to participation and makes cross-campus & multidisciplinary research practical (Goals 2 & 3)**
- **High-performance computing & networking supports big data research for economic & social improvement, allows us to deploy science for the public good faster (Goal 4)**
- **Digital Research Platforms that serve multiple researchers reduce excess spending, focus technology and research growth, and ensure good stewardship of public funds (Goal 5)**

IDI

Making Critical High Performance Cyberinfrastructure Seamlessly Available to IDI Users Where They Work



CENIC

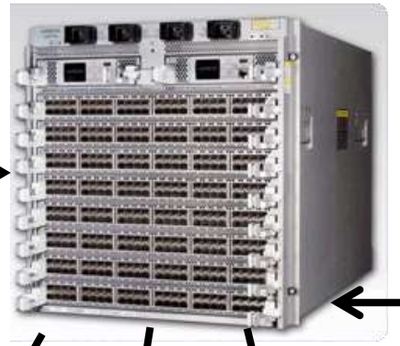
CHERuB

10

1-16

8

Prism@UCSD







288

128

SDSC
Supercomputers



SDSC
SAN DIEGO SUPERCOMPUTER CENTER

4



1

4

8

UCSD IDI Users





384

384 x 10Gbps = 3.8Tbps



NSF

Oasis Data Store

>13,000 TB

> 800 Gbps

# of Parallel 10Gbps
Optical Light Paths

SDSC

IDI

Data Research Library:
Curation, Annotation, Publishing of Researcher Generated Datasets

species	LS001	LS002	LS003	LS004	LS005
Methanobrevibacter smithii	0.184008	0.003628	0.003694	0.038584	0.031065
Parvimonas micra	0.159721	0.000073	0.000048	0.000263	0.00021
Escherichia coli	0.103199	0.099081	0.096903	0.065271	0.001651
Peptostreptococcus stomatis	0.062615	0.000057	0.000054	0.000208	0.000114
Fusobacterium sp. 12_1B	0.055544	0.000021	0.000024	0.000082	0.000024
Solobacterium moorei	0.051153	0.000126	0.000086	0.000257	0.000253
Parvimonas sp. oral taxon 110	0.033583	0.000043	0.00004	0.000207	0.000083
Parvimonas sp. oral taxon 393	0.027749	0.000054	0.000052	0.000192	0.000079
Methanosphaera stadtmanae	0.014501	0.000012	0.000008	0.002937	0.000508
Dialister invisus	0.014365	0.020419	0.007633	0.000792	0.000322
Gemella morbillorum	0.011566	0.000035	0.000035	0.000175	0.000432
Streptococcus thermophilus	0.009034	0.008043	0.024032	0.033324	0.010473
Escherichia sp. 3_2_53FAA	0.008568	0.000372	0.000364	0.000235	0.000006
Collinsella aerofaciens	0.008566	0.082094	0.017921	0.030015	0.045038
Akkermansia muciniphila	0.007203	0.004183	0.021161	0.031123	0.069592

Large Memory High Performance Computing Enables Comparison Across Human Gut Microbiome of Patients with Autoimmune Diseases and Healthy Subjects

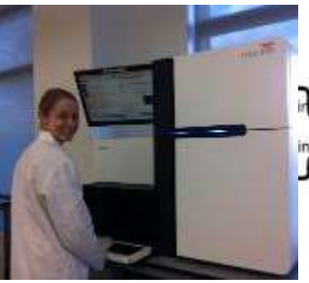
Sitao Wu[1], Weizhong Li[1], Larry Smarr[2, 3], Karen Nelson[4], Shibu Yooseph[5], Manolito Torralba[4]
University of California San Diego, 9500 Gilman Drive, La Jolla, CA, USA; [1]Center for Research for Biological Systems, [2]California Institute for Telecommunication and Information Technology, [3]Department of Computer Science and Engineering; [4]J. Craig Venter Institute, Rockville, MD, USA; [5]J. Craig Venter Institute, San Diego, CA, USA
(SW) siw006@ucsd.edu, (WL) liwz@sdsc.edu, (LS) lsmarr@ucsd.edu, (KN) Kenelson@jcvi.org, (SY) SYooseph@jcvi.org, (MT) MTorralba@jcvi.org

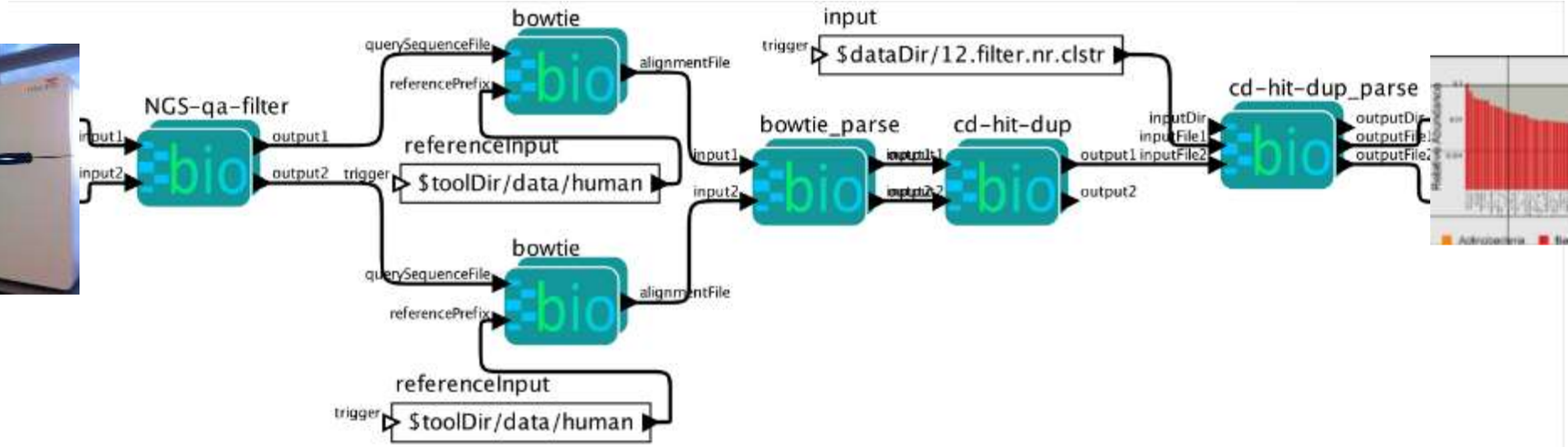
ABSTRACT

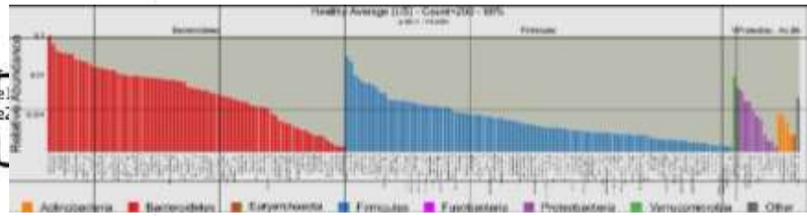
Microbial communities that live on the outside and inside of the human body dramatically influence human health and diseases. In recent years, major progress has been made in understanding the human microbiome communities through projects such as the

1. INTRODUCTION

The microbes that live in and on the human body outnumber the human cells by 10-fold. The collective human microbial communities, known as the human microbiome, play a profound







IDI

Digital Research Platform: Distributed IPython/Jupyter Notebooks:
Cross-Platform, Browser-Based Application Interleaves Code, Text, & Images

IJulia
IHaskell
IFSharp
IRuby
IGo
IScala
IMathics
Ialdor
LuaJIT/Torch
Lua Kernel
IRKernel (for the R language)
IErlang
IOCaml
IForth
IPerl
IPerl6
Ioctave
Calico Project
- kernels implemented in Mono, including Java, IronPython, Boo, Logo, BASIC, and many others



Evolved from the IPython Project

IScilab
IMatlab
ICSharp
Bash
Clojure Kernel
Hy Kernel
Redis Kernel
jove, a kernel for io.js
IJavascript
Calysto Scheme
Calysto Processing
idl_kernel
Mochi Kernel
Lua (used in Splash)
Spark Kernel
Skulpt Python Kernel
MetaKernel Bash
MetaKernel Python
Brython Kernel
IVisual VPython Kernel

Source: John Graham, QI

IDI

Digital Research Platform Software System Being Used by Numerous IDI Transformational Projects





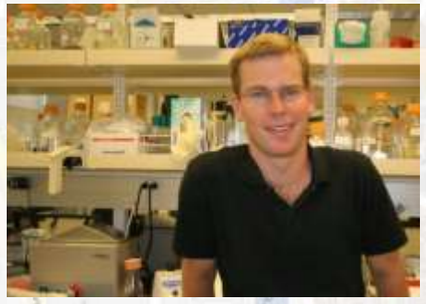














Source: John Graham, QI

IDI Showcase Speakers

WEDNESDAY, MAY 6 10AM-12PM/SCHOOL OF MEDICINE/MET LEARNING CENTER	
Ilkay Altintas, SDSC	WiFIRE/UCSD GIS effort
Jurgen Schulze, QI/CSE & Trey Ideker, Medicine	Creating greatly expanded, scalable visualization capability for graphing gene and cellular networks
Rommie Amaro, Chemistry & Biochemistry	Rational drug design
Falko Kuester, SE	Prototype lab for student access to drones
Lucila Ohno-Machado, Medicine	Establish a scalable Health Sciences HIPAA cloud for human-generated data

IDI

IDI Showcase Speakers

WEDNESDAY, MAY 6, 2015 2-4PM/SCRIPPS INSTITUTION OF OCEANOGRAPHY/4500 HUBBS HALL	
Mark Ellisman, Neuroscience	Crack the living cell nucleus
Frank Wuerthwein, Physics	Large Hadron Collider/CMS Data Tier Two site
Brenda Bloodgood, Biology	Neuronal computation changes in response to interactions with the environment
Greg Hidley, CalIT2	SDSC's High Performance Wireless Research and Education Network
Jules Jaffe, MPL	Underwater imaging of plankton/phytoplankton

IDI

IDI Showcase Speakers

THURSDAY, MAY 7 10AM-12PM/WEST CAMPUS/MEETING ROOMS ON FIFTEEN/THE VILLAGE 15TH FLOOR	
Kim Albizati, Chemistry & Biochemistry	Undergraduate instruction in upper-division Chemistry
Thomas Levy, Anthropology	Evolution of societies in the southern Levant from the Neolithic to Islamic periods
Mehrdad Yazdani, QI	Twitter Big Data study measuring happiness of metro areas
Alison Marsden, MAE	Graduate instruction on computational fluid dynamics
Rob Knight, Pediatrics	Radically advance UCSD's capabilities in multi-omic integration of the human microbiome

IDI