



CEPH in HPC Environments

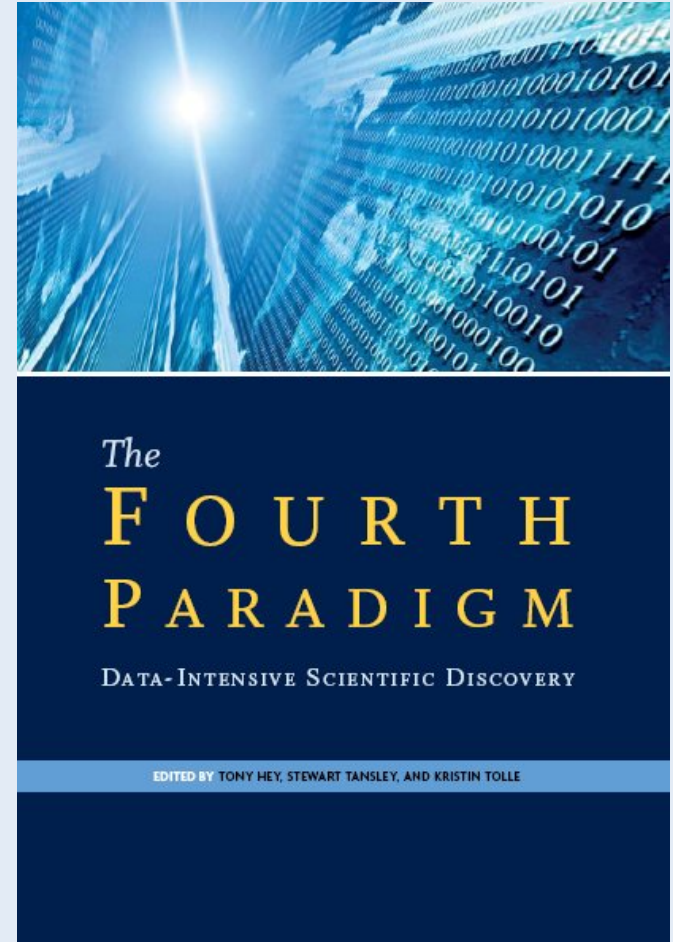
November 18, 2015

SC15 Birds of a Feather

<https://www.msi.umn.edu/sc15Ceph>

Jim Gray 2007

“Experimental, theoretical, and computational science are all being affected by the data deluge, and a fourth, “data-intensive” science paradigm is emerging.”



UNIVERSITY OF MINNESOTA

© 2015 Regents of the University of Minnesota. All rights reserved.

Minnesota Supercomputing Institute
MSI

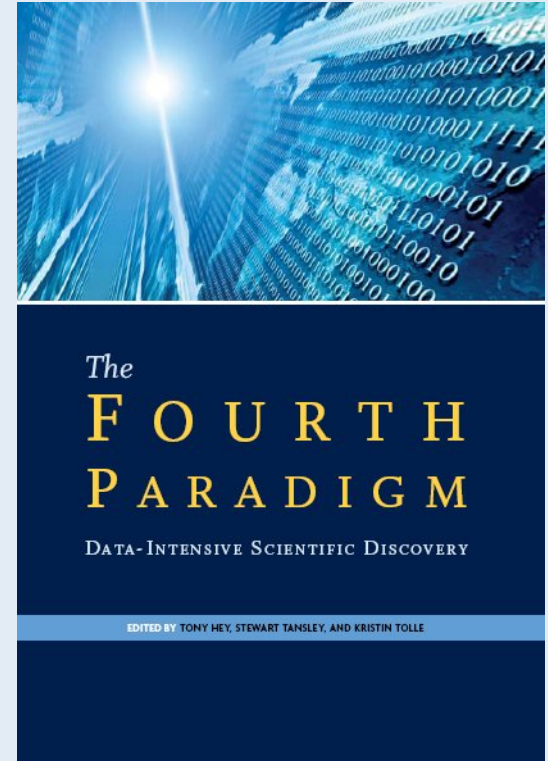
Science Paradigms

(*a la* Jim Gray)

- Experiments
 - Describe natural phenomena
- Theory
 - Generalize these phenomena
- Computations
 - Simulate complex phenomena
- Large-scale Data Exploration
 - unify experimentation, theory, and simulation

Jim Gray 2007

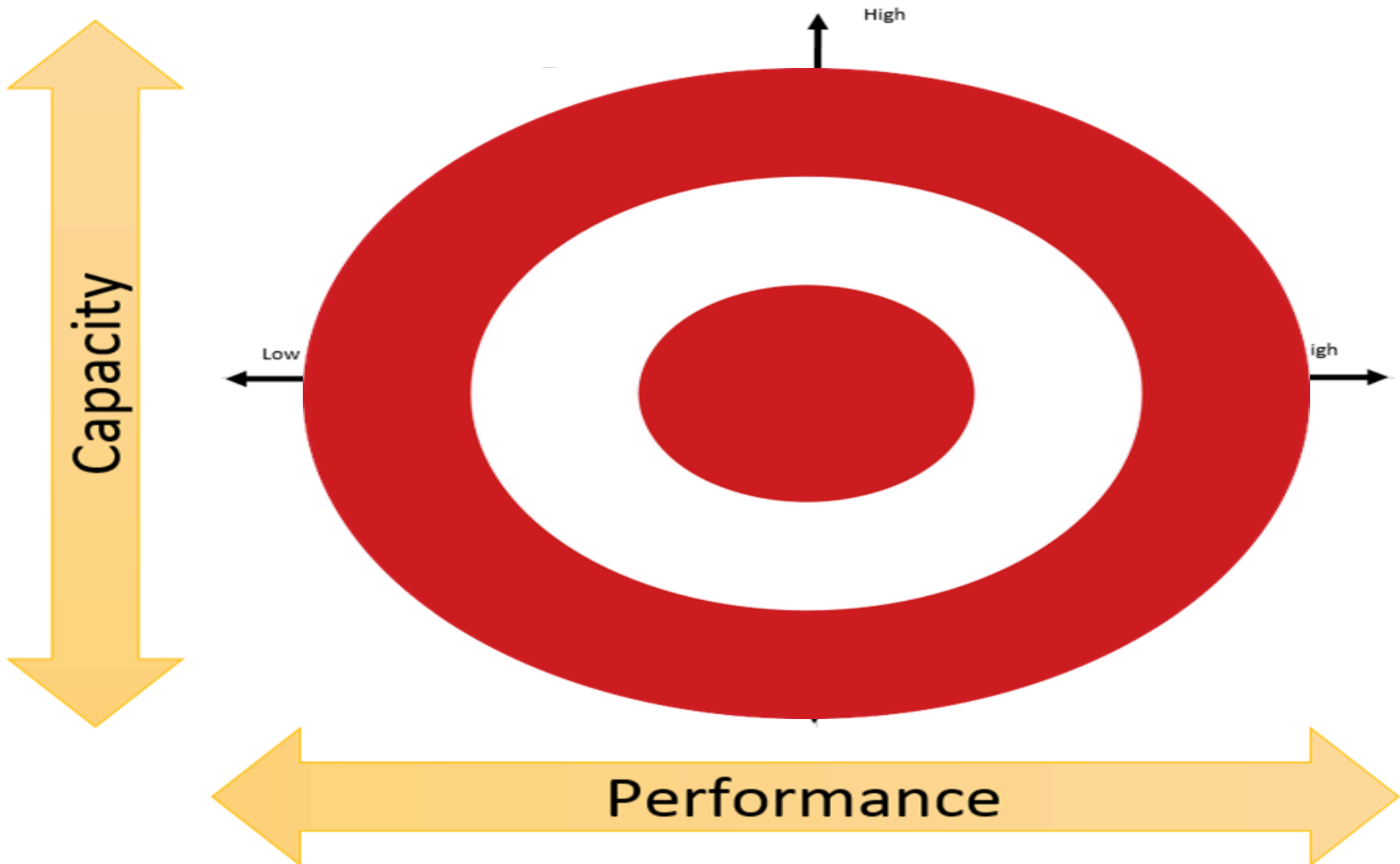
“Today, the tools for capturing data both at the mega-scale and at the milli-scale are just dreadful.”



UNIVERSITY OF MINNESOTA

© 2015 Regents of the University of Minnesota. All rights reserved.

Minnesota Supercomputing Institute
MSI



Funding



Office of Science and Technology Policy
Executive Office of the President
New Executive Office Building
Washington, DC 20502

FOR IMMEDIATE RELEASE

March 29, 2012

Contact: Rick Weiss 202 456-6037 rweiss@ostp.eop.gov
Lisa-Joy Zgorski 703 292-8311 lisajoy@nsf.gov

**OBAMA ADMINISTRATION UNVEILS “BIG DATA” INITIATIVE:
ANNOUNCES \$200 MILLION IN NEW R&D INVESTMENTS**

UNIVERSITY OF MINNESOTA

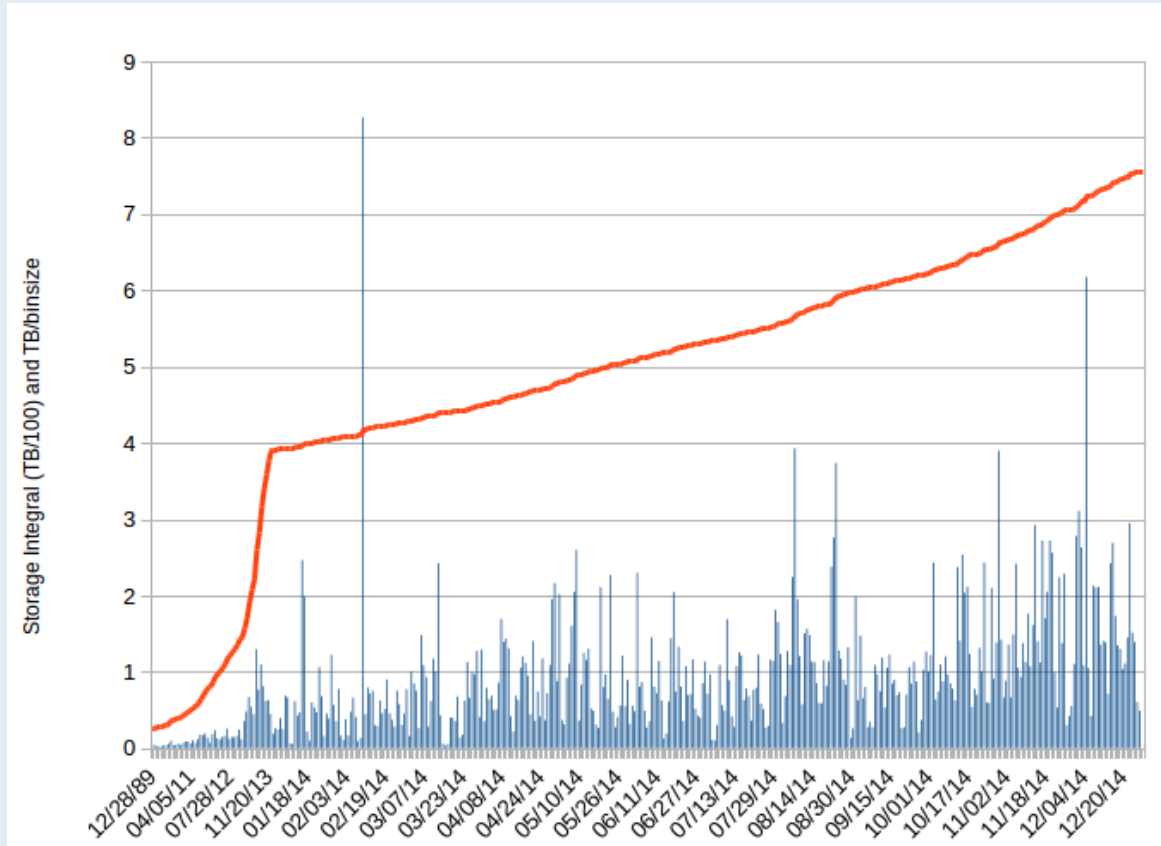
© 2015 Regents of the University of Minnesota. All rights reserved.

Minnesota Supercomputing Institute
MSI

BIG DATA Challenges

- Data generators are getting better
- Example data generators (and data aggregators) on UMN's campus:
 - Genetic Sequencing (UM Genomics Center)
 - Proteomics (Center for Mass Spectrometry & Proteomics)
 - Demographic Information (Minnesota Population Center)
 - Large Computers (e.g., UMN's Mesabi) generate simulation results faster

BIG DATA Challenges



UNIVERSITY OF MINNESOTA

© 2015 Regents of the University of Minnesota. All rights reserved.

Minnesota Supercomputing Institute
MSI

BIG DATA Challenges

“... in 2012, I revised my request from 75 GB to 5 TB to 35 TB, and then in 2014 I requested 120TB.”

Speakers

- Jim Wilgenbusch, Minnesota Supercomputing Institute, University of Minnesota
- Douglas Fuller, Red Hat, Inc.
- Dan Ferber, Intel Corporation
- Ben Lynch, Minnesota Supercomputing Institute, University of Minnesota
- BJ Lougee, Center for the Advancement of Data and Research in Economics (CADRE)
- Dr. Hong Ong, Advanced Computing Lab and Accelerative Technology Lab, MIMOS Berhad
- John-Paul Robinson, University of Alabama at Birmingham
- Greg Farnum, Red Hat, Inc.
- Mark Nelson, Red Hat, Inc.