An Overview of Selected NSF Programs relevant to Knowledge Discovery from Data Analytics on High-Performance Computing Platforms

Presentation to
"*Delivering Science on Petascale Computers*",
Session on "*Knowledge Discovery thru Data Analytics*"
Tuesday, September 9, 2008

Douglas H. Fisher
Program Director
National Science Foundation (NSF)
Directorate for Computer & Information Science & Engineering (CISE)
Division of Information & Intelligent Systems (IIS)
Robust Intelligence Program (RI)

It's not hard to find NSF support of research on Conference topics

http://www.google.com/search?hl=en&q=Delivering+Science+on+Petascale+Computers&btnG=Search

Results **1** - **10** of about **18,300** for **Delivering Science on Petascale Computers**. (**0.15** seconds)
Search Results

Massive $208 million ***petascale computer*** gets green light **...**
Sep 5, 2008 **...** Massive $208 million ***petascale computer*** gets green light **...** Blue Waters is expected to ***deliver*** sustained performance of more than one **...** **NSF**

National ***Science*** Board Approves Funds For ***Petascale*** Computing Systems
Aug 14, 2007 **...** This "***petascale***" system is expected to be able to make arithmetic **...** the world's fastest ***computers***, the National ***Science*** Foundation's (NSF) **...**

Notice 06-26
Aug 17, 2006 **...** Accelerating ***Delivery*** of ***Petascale*** Computing Environment **...** ***computer science***, and computational ***science*** in the physical, biological, **...**

Amazon.com: ***Petascale*** Computing: Algorithms and Applications **...**
Featuring contributions from the world's leading experts in computational ***science***, this edited collection explores the use of ***petascale computers*** for **...**
Cached - Similar pages

[DOC] ***Petascale*** computational systems
File Format: Microsoft Word - View as HTML
Abstract: Computational ***science*** is changing to be data intensive. Super-***Computers*** must be balanced system, not just CPU farms but also ***petascale*** IO and **...**

About the Blue Waters project **NSF**
Read the National ***Science*** Foundation's solicitation for ***Petascale*** Computing **...** The system will ***deliver*** sustained performance of one to two petaflops for **...**

[PDF] ***Delivering Science on Petascale Computers***
File Format: PDF/Adobe Acrobat - View as HTML
Delivering Science on Petascale Computers. http://www.ccs

Douglas H. Fisher

HPC news from NSF's "front page"
http://www.nsf.gov, September 9, 2008

Press Release 08-153
Climate Computer Modeling Heats Up

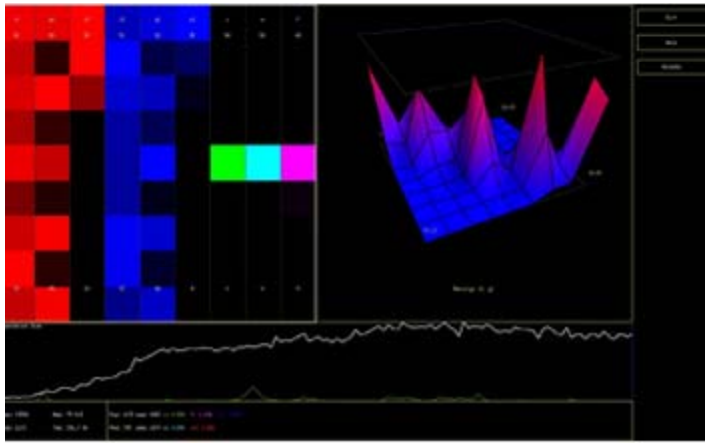
September 4, 2008
New "petascale" computer models lead to better understanding of weather-climate links

"The limiting factor to more reliable climate predictions at higher resolution is not scientific ideas, but computational capacity to implement those ideas," said Jay Fein, NSF program director in NSF's Division of Atmospheric Sciences. "This project is an important step forward in providing the most useful scientifically-based climate change information to society for adapting to climate change."

Douglas H. Fisher

HPC news from NSF's "front page"
http://www.nsf.gov, September 9, 2008

Press Release 08-152
NSF Funds New Center to Bring Together Biologists, Mathematicians
Power of mathematics and modeling to be applied to large-scale questions in biology



Anolis lizard diversity, shown in a simulation model's output, is among the center's subjects.

September 3, 2008

"The interface between the mathematical and physical sciences and the biological sciences represents an important strategic investment area for MPS [NSF's Directorate for Mathematical and Physical Sciences]," said Tony Chan, NSF assistant director for MPS. Chan noted that many divisions within MPS have a history of significant investments at this interface, and new programs continue to be developed and initiated.

This grant follows on the heels of the NSF creation of the National Institute for Computational Sciences at UT Knoxville, a $65 million award to build and operate a supercomputer to assist scientists nationwide.

UT and Oak Ridge National Laboratory experts in high-performance computing will work with NIMBioS scientists to apply the power of supercomputing to the difficult questions NIMBioS will answer.

Douglas H. Fisher

Cyberinfrastructure Vision for 21st Century Discovery

National Science Foundation
Cyberinfrastructure Council
March 2007

The National Science Foundation's Cyberinfrastructure Council (CIC)1, based on extensive input from the research community, has developed a comprehensive vision to guide the Foundation's future investments in cyberinfrastructure (CI). In 2005, four multi-disciplinary, cross-foundational teams were created and charged with drafting a vision for cyberinfrastructure in four overlapping and complementary areas: *1) High Performance Computing, 2) Data, Data Analysis, and Visualization, 3) Cyber Services and Virtual Organizations, and 4) Learning and Workforce Development*. Draft versions of the document were posted on the NSF website and public comments were solicited from the community. These drafts were also reviewed for comment by the National Science Board. The National Science Foundation thanks all of those who provided feedback on the Cyberinfrastructure Vision for 21st Century Discovery document. Your comments were carefully reviewed and considered during preparation of this version of the document, which is intended to be a living document, and will be updated periodically.
http://www.nsf.gov/pubs/2007/nsf0728/nsf0728.pdf

Douglas H. Fisher

Cyberinfrastructure Vision for 21st Century Discovery

The Next Five Years: Creating a **High Performance Computing** Environment for Petascale Science and Engineering

The following **principles** will guide the agency's FY 2006 through FY 2010 investments:

• Research and education priorities will drive HPC investments.

• Collaborative activities … are needed to ensure that HPC systems and services … support petascale scientific computing.

• Researchers and educators require access to reliable, robust, production-quality HPC resources and services.

• HPC-related R&D advances (public and private; domestic and foreign) must be leveraged to enrich HPC capabilities.

• The development, implementation and annual update of an effective multi-year HPC strategy is crucial to the timely introduction of R&D outcomes and innovations in HPC systems, software and services.

NSF's **implementation plan** to create a petascale environment includes the following three interrelated components:

1). Specification, Acquisition, Deployment and Operation of Science-Driven HPC Systems Architectures

2). Development and Maintenance of Supporting Software: New Design Tools, Performance Modeling Tools, Systems Software, and Fundamental Algorithms.

3). Development and Maintenance of Portable, Scalable Applications Software

Douglas H. Fisher

Cyberinfrastructure Vision for 21st Century Discovery

The Next Five Years: Towards a **National Digital Data Framework**

The following **principles** will guide the agency's FY 06 through FY10 investments:

Data generated with NSF funding will be readily accessible and easily usable, and will be appropriately, responsibly and reliably preserved.

Broad community engagement is essential to the prioritization and evaluation of the utility of scientific data collections, …

Exploitation of data in the creation of new knowledge requires that investigators have access to the tools/services …

Strong, reciprocal, international, interagency and public-private partnerships is essential

Mechanisms will be created to share data stewardship best practices between nations, communities, organizations and individuals.

Mechanisms essential to the development of both statistical and technical ways to protect privacy and confidentiality will be supported.

1) A Coherent Organizational Framework - Data Collections and Managing Organizations

NSF will foster the establishment of interagency, public-private and international consortia charged with providing stewardship for digital data collections to promote interoperability across data collections.

NSF will continue to promote the coalescence of appropriate collections with overlapping interests, approaches and services.

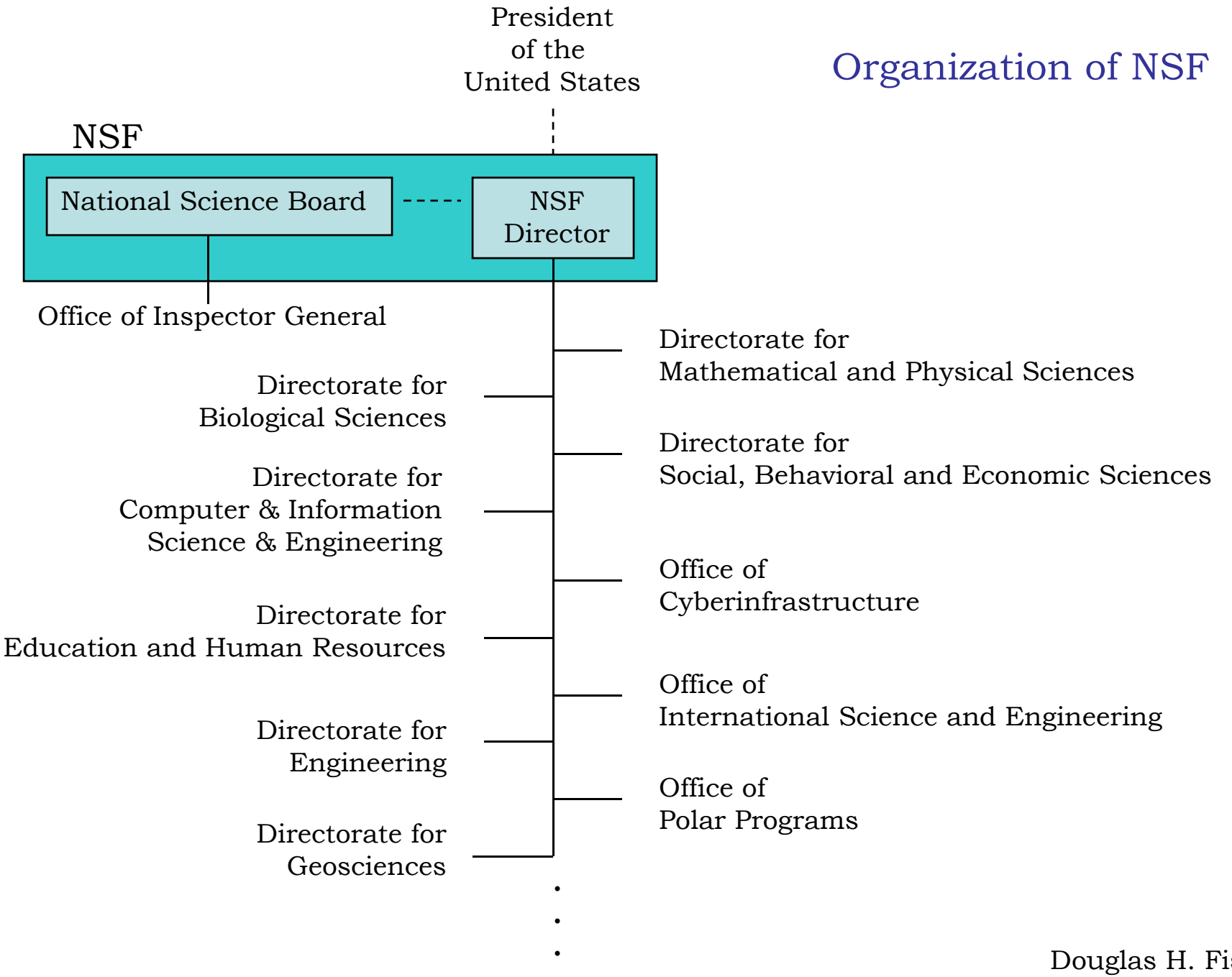
2) Developing A Flexible Technological Architecture

The architecture must use standard open protocols and interfaces,…, facilitate user acquisition, access, analysis, mining, and visualization, integration of data, addressing issues such as authentication, authorization and other security concerns.. It must also support complex workflows enabling data discovery….metadata, ontology

3) Developing and Implementing Coherent Data Policies

Douglas H. Fisher

Organization of NSF



Douglas H. Fisher

Where is data analytics for knowledge discovery with HPC found at NSF?

a) Where does "petascale computing" reside – research and applications?
b) ....."high-performance computing" .....?
c) ..... "knowledge discovery" .....?
d) ..... "data mining" .....?
e) ..... "data analysis" .....?
f) ..... "data visualization" .....?
Search (active) NSF Award titles and abstracts for X (http://www.nsf.gov)

Directorate for
Biological Sciences (BIO)
a)0 b)10 c)3 d)21 e)53 f)2

Directorate for
Computer & Information
Science & Engineering (CSE)
a)5 b)93 c)29 d)188 e)100 f)14

Directorate for
Engineering (ENG)
a)0 b)21 c)8 d)30 e)44 f)5

Directorate for
Geosciences (GEO)
a)2 b)7 c)1 d)5 e)181 f)12

Directorate for
Mathematical and Physical Sciences (MPS)
a)4 b)52 c)1 d)46 e)233 f)6

Office of
Cyberinfrastructure (OCI)
a)18 b)34 c)0 d)4 e)18 f)1

The ***number*** (not magnitude) of awards for each search key is given and is ***suggestive*** of ***relative*** support level in each NSF Directorate and Office

Douglas H. Fisher

About Office of Cyberinfrastructure (OCI)

http://www.nsf.gov/dir/index.jsp?org=OCI

The Office of Cyberinfrastructure coordinates and supports the acquisition, development and provision of state-of-the-art cyberinfrastructure resources, tools and services essential to the conduct of 21st century science and engineering research and education.

OCI supports cyberinfrastructure resources, tools and related services such as supercomputers, high-capacity mass-storage systems, system software suites and programming environments, scalable interactive visualization tools, productivity software libraries and tools, large-scale data repositories and digitized scientific data management systems, networks of various reach and granularity and an array of software tools and services that hide the complexities and heterogeneity of contemporary cyberinfrastructure while seeking to provide ubiquitous access and enhanced usability.

OCI supports the preparation and training of current and future generations of researchers and educators to use cyberinfrastructure to further their research and education goals, while also supporting the scientific and engineering professionals who create and maintain these IT-based resources and systems and who provide essential customer services to the national science and engineering user community.

ACCELERATING DISCOVERY IN SCIENCE AND ENGINEERING THROUGH PETASCALE SIMULATIONS AND ANALYSIS (PetaApps)

http://www.nsf.gov/funding/pgm_summ.jsp?pims_id=501015&from=fund

Full Proposal Deadline Date: October 30, 2008

This solicitation seeks proposals to develop the future simulation and analysis tools that can use petascale computing to advance the frontiers of scientific and engineering research. Proposals are sought from researchers aiming to capitalize on emerging petascale computing architectures, catalyzing potentially transformative research. NSF's emphasis is on implementation and exploitation of forefront techniques. Proposers must be prepared to demonstrate that they have a research problem that requires and can exploit petascale computing capabilities. Proposals from or including junior researchers are encouraged as one of the goals of this solicitation is to build a community capable of using petascale computing.

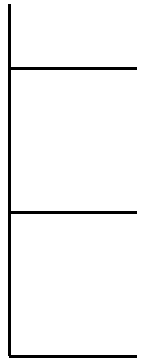
ACCELERATING DISCOVERY IN SCIENCE AND ENGINEERING THROUGH PETASCALE SIMULATIONS AND ANALYSIS (PetaApps)

Name	Dir / Div	Name	Dir / Div
Abani Patra		Almadena Chtchelkanova	CISE / CCF
Clark Cooper	ENG / CMMI	Daryl Hess	MPS / DMR
Eric C. Itsweire	GEO / OCE	Hans Kaper	MPS / DMS
Raima M. Larter	MPS / CHE	Ping Li	SBE / BCS
Stephen Meacham	OD / OCI	Scott F. Midkiff	ENG / ECCS
Barry Schneider	MPS / PHY	Nigel Sharp	MPS / AST
Phillip R. Westmoreland	ENG / CBET	Eva Zanzerkia	GEO / EAR

Where is data analytics for knowledge discovery with HPC found in **CSE**?

a) “petascale computing” b) “high-performance computing” c) “knowledge discovery”
d) “data mining” e) “data analysis” f) “data visualization

Directorate for Computer & Information Science & Engineering (CSE)
a)5 b)93 c)29 d)188 e)100 f)14



Computer and Communication Foundations (CCF)
a)1 b)36 c)7 d)44 e)42 f)2

Computer and Network Systems (CNS)
a)4 b)54 c)6 d)54 e)16 f)8

Information and Intelligent Systems (IIS)
a)0 b)3 c)16 d)90 e)42 f)4

How much research is supported at the intersection:
interactive data mining on HPC

high-performance & "data mining": CNS (2)

high-performance & visualization: NSF (69), CSE (32), OCI (16)

"data mining" & visualization: NSF(0)

Approaching NSF on individual proposals:
do your homework and approach relevant Program Director(s) with a white paper

NSF is guided by bottom-up (i.e., community) and by top-down influences, so you can:

Approach NSF as a community:
do your homework and approach relevant Program and/or Division Directors with a Workshop white paper, proposing investigations into new, emerging fields

"interactive data mining on HPC" is consistent with CISE and NSF themes featuring "human in the loop", "data to knowledge", "understanding and managing complexity", "computational thinking", and "high-performance computing"

Douglas H. Fisher

Another example of NSF responsiveness:

Update on Computer Graphics and Visualization

http://www.nsf.gov/news/news_summ.jsp?cntn_id=112033

“Advances in such areas as computer vision and image retrieval, information visualization and data mining, and multimodal interfaces and human-computer interaction, among others, have led to a portfolio of NSF investments now spanning both CCF and IIS.”

“…we encourage PIs to submit proposals in computer graphics to the Human Centered Computing Program in IIS and include "computer graphics" as one of their keywords, and to submit proposals in visualization to the Information Integration and Informatics Program in IIS and include "visualization" as one of their keywords. Computational geometry proposals should continue to be submitted to the Algorithmic Foundations Program within CCF.”

Haym Hirsh, Division Director, Information and Intelligent Systems
Sampath Kannan, Division Director, Computer and Communications Foundations
Jeannette M. Wing, Assistant Director

Computer and Information Science and Engineering Directorate
National Science Foundation

Douglas H. Fisher

Continuing (from FY08 Initiatives)

- Expeditions in Computing (CISE) – September 10 pre-proposals
- Cyber-enabled Discovery and Innovation (CDI, NSF-wide)
  (Nov-Dec preliminary proposal window – see solicitation)

Enhanced Initiatives

- Network Science and Engineering
- Trustworthy Computing

New Initiatives

- Data-Intensive Computing (CISE) (Med: Oct^{2008}; Lar: Nov; Sma: Dec)
- Cyber-Physical Systems (joint with ENG)

Expeditions in Computing

- Bold, creative, visionary, high-risk ideas
- Whole >> $\sum$ part **I**
- Solicitation is deliberately under constrained
  - Tell us what YOU want to do!
  - Response to community
- Loss of ITR Large, DARPA changes, support for high-risk research, large experimental systems research, etc.
- FY08: 4 awards, each at about $10M for 5 years
  - 122 LOI, 75 prelim, 20 final, 7 reverse site visits

Adapted from Jeannette Wing's Snowbird 2008 presentation (http://www.cra.org/Activities/snowbird/2008/slides/Wing_Snowbird08.pdf)

Expeditions FY 2008 Awards

http://www.nsf.gov/news/news_summ.jsp?cntn_id=112075&org=NSF&from=news

- *Computational Sustainability: Computational Methods for a Sustainable Environment, Economy and Society*
  Lead PI: Carla Gomes, Cornell University

- *Understanding, Coping with and Benefiting from Intractability*
  Lead PI: Sanjeev Arora, Princeton University

- *The Molecular Programming Project*
  Lead PI: Erik Winfree, California Institute of Technology

- *Open Programmable Mobile Internet 2020*
  Lead PI: Nick McKeown, Stanford University

Expeditions in Computing (CISE)

Computational Sustainability:

Computational Methods for a Sustainable Environment, Economy, and Society

Lead PI: Carla Gomes, Cornell University

Sustainability: "development that meets the needs of the present without compromising the ability of future generations to meet their needs." Our Common Future, Brundtland Report,1987

Vision: *Computer scientists can — and should —play a key role in increasing the efficiency and effectiveness of the way we manage and allocate our natural resources, while enriching and transforming Computer Science.*

Goals for Sustainability

To inject *computational thinking* into Sustainability,

- establishing *computational sustainability* as a new field
- bringing new insights to sustainability challenges
- preparing a new generation to grapple with long-term sustainability

Goals for Computational Studies

- to motivate *transformative synthesis* and new methodologies across computing sub-disciplines
- to broaden participation in the computing sciences and engineering, in part by
- broadening the public image of computing science, as a field of great societal importance

The Institute for Computational Sustainability (ICS)

Establishing a vibrant research community in Computational Sustainability, operating under principles of inclusiveness, extending its influence and resources far beyond its members

CDI: Cyber-Enabled Discovery and Innovation

Computational Thinking for Science and Engineering

http://www.nsf.gov/funding/pgm_summ.jsp?pims_id=503163&from=fund

- Paradigm shift
  - Not just our metal tools (transistors and wires) but also our mental tools (abstractions and methods)
- Three Themes
  - From Data to Knowledge: *enhancing human cognition and generating new knowledge from a wealth of heterogeneous digital data;*
  - Understanding Complexity in Natural, Built, and Social Systems: *deriving fundamental insights on systems comprising multiple interacting elements;* and
  - Virtual Organizations: *enhancing discovery and innovation by bringing people and resources together across institutional, geographical and cultural boundaries.*
- FY08: $47.9M for ~30 awards
  - 1900 LOIs, 1300 preliminary proposals, 200 final proposals

Adapted from Jeannette Wing's Snowbird 2008 presentation (http://www.cra.org/Activities/snowbird/2008/slides/Wing_Snowbird08.pdf)

Data Intensive Computing

http://www.nsf.gov/funding/pgm_summ.jsp?pims_id=503324&org=CNS

How Much Data?

- NOAA has ~1 PB climate data (2007)
- Wayback machine has ~2 PB (2006)
- CERN's LHC will generate 15 PB a year (2008)
- HP is building WalMart a 4PB data warehouse (2007)
- Google processes 20 PB a day (2008)
- "All words ever spoken by human beings" ~ 5 EB
- Int'l Data Corp predicts 1.8 ZB of digital data by 2011

Adapted from Jeannette Wing's Snowbird 2008 presentation
(http://www.cra.org/Activities/snowbird/2008/slides/Wing_Snowbird08.pdf)

Data Intensive Computing: Convergence in Trends

- Drowning in data
- Data-driven approach in computer science research
  - graphics, animation, language translation, search, …, computational biology
- Cheap storage
  - Seagate Barracuda 1TB hard drive for $195
- Growth in huge data centers
- Data is in the "cloud" not on your machine
- Easier access and programmability by anyone
  - e.g., Amazon EC2, Google+IBM cluster, Yahoo! Hadoop

Adapted from Jeannette Wing's Snowbird 2008 presentation (http://www.cra.org/Activities/snowbird/2008/slides/Wing_Snowbird08.pdf)

Data-Intensive Computing Sample Research Questions

- Science
  - What new abstractions (including models, languages, algorithms) are needed for data-intensive, rather than process-intensive computing?
  - What new metrics are needed to evaluate performance of data-intensive computations?
- Technology
  - How can we automatically manage the hardware and software of these data-intensive computing systems at scale?
  - How can we provide security and privacy for simultaneous mutually untrusted users, for both processing and data?
  - How can we reduce these systems' power consumption?
- Society
  - What (new) uses and users might arise from our ability to process large scale datasets?

Adapted from Jeannette Wing's Snowbird 2008 presentation (http://www.cra.org/Activities/snowbird/2008/slides/Wing_Snowbird08.pdf)

Other (Selected) Continuing Programs (see website http://www.nsf.gov for full list).

• Creative IT (CISE, SBE)

• Collaborative Research in Computational Neuroscience (CRCNS) (CISE, BIO, SBE, MPS) + NIH

• DataNet

• High End Computing University Research Activity (HECURA) (CISE)

• Computing Research Infrastructure (CISE, Sept. 22, 2008)

• Major Research Instrumentation (NSF-wide)

• High Performance Computing System Acquisition: Towards a Petascale Computing Environment for Science and Engineering (OCI, Nov. 28, 2008)

• TeraGrid Phase III (OCI, November 4, 2008)

• Education: CPATH, BPC

• Science and Technology Centers (NSF-wide)

FY08 and FY09 CISE Funding

- FY08 (FY began 10/1/07)
  - CISE Request was $574 million, a **9%** increase over FY07
  - CISE Appropriate is $535 million, only a 1.5% increase
  - Missed opportunities of $39 million (e.g, ~325 awards or 400 grad students
- FY09 (FY begins 10/1/08)
  - CISE Request totals $639 million
  - Reflects a **$104** million increase, or **19.5%** over FY08 level.

Adapted from Jeannette Wing's Snowbird 2008 presentation
(http://www.cra.org/Activities/snowbird/2008/slides/Wing_Snowbird08.pdf)

The 2009 NSF Budget request

http://www.nsf.gov/about/budget/fy2009/pdf/01_fy2009.pdf

Cross-Foundation Investments. The FY 2009 Request includes four major cross-foundation investments that aim to have a transformative impact across science and engineering, especially in such areas of national priority as manufacturing, computing, energy, cybersecurity, sensors, and materials.

• **Cyber-enabled Discovery and Innovation (CDI),** initiated in FY 2008, increases to $100.0 million (from $47.9 million) in FY 2009 to advance science and engineering along fundamentally new pathways opened by computational capabilities.

• **Science and Engineering Beyond Moore's Law (SEBML)** aims to position the U.S. at the forefront of communications and computation capability beyond the physical and conceptual limitations of current systems. This $20.0 million, NSF-wide effort … Activities in FY 2009 will encourage transformational activities as well as creating partnering opportunities with the private sector and national laboratories to accelerate innovation.

•**Adaptive Systems Technology (AST)** focuses on generating creative pathways and natural interfaces between human and physical systems that will revolutionize the development of novel adaptive systems. This investment of $15.0 million in FY 2009 … AST is essential to advances in highly-innovative adaptive control systems, hybrid computer architectures, improved electronic PDAs, and computer-based, self-paced, learning and training tools.

• **Dynamics of Water Processes in the Environment (WATER)** is a $10.0 million, NSF-wide investment that aims to increase fundamental understanding of the Earth's freshwater systems and provide the scientific basis for decision-making about water resources. Major efforts in FY 2009 include fundamental research on the complex processes and feedbacks that affect the vulnerability and resilience of freshwater systems to climate and environmental change. …

Douglas H. Fisher

The 2009 NSF Budget request

http://www.nsf.gov/about/budget/fy2009/pdf/01_fy2009.pdf

INTERAGENCY R&D PRIORITIES

(Dollars in Millions)

	FY 2007 Actual	FY 2008 Estimate	FY 2009 Request	Change over FY 2008 Estimate Amount	Percent
National Nanotechnology Initiative	$388.69	$388.69	$396.79	$ 8.10	2.1%
Climate Change Science Program	206.63	205.25	220.60	15.35	7.5%
Climate Change Technology Program	21.00	21.00	23.50	2.50	11.9%
Networking and Information Technology R&D	908.45	931.48	1,090.25	158.77	17.0%
Homeland Security	388.76	368.41	379.17	10.76	2.9%

Douglas H. Fisher

Concluding Remarks

NSF will play a leadership role in the development and support of a comprehensive cyberinfrastructure essential to 21st century advances in science and engineering research and education. (**C****yberinfrastructure* ***V****ision for* ***21****st* ***C****entury* ***D****iscovery,* http://www.nsf.gov/pubs/2007/nsf0728/nsf0728_2.pdf)

Examine the NSF Web site (all of it) for current solicitations, supplement possibilities (e.g., Communicating Research to Public Audiences, http://www.nsf.gov/funding/pgm_summ.jsp?pims_id=5362&from=fund), and longer-term visions.

Call/email program director**S** with questions – white papers by email help

Consider service as a PD, a DD, an AD

Douglas H. Fisher

Auxiliary Slides

Searched for **2008 Petascale "knowledge discovery" visualization** in **NSF** NEWS
901 documents found in 2394 documents searched
1 2 3 4 5 6 7 8 9 10
Ranger Supercomputer Dedicated by NSF and Texas Advanced Computing Center
Update on Computer Graphics and Visualization
National Science Board Approves Funds for Petascale Computing Systems
National Science Foundation Requests $6.43 Billion for Fiscal Year 2008
Visualize This
LSU Center for Computation & Technology Director To Head NSF's OCI
Lasers, Software and the Devil's Slide
Climate Computer Modeling Heats Up
Back to School: Five Myths about Girls and Science
July 15 Deadline to Apply for High-Performance Computational Time on Teragrid
Powerful Tool Crunches Commutes
New Technologies Could Make Airport Screening More Effective and Less Cumbersome
National Science Foundation Presents FY 2008 Budget Proposal on February 5
Attend Capital Science 2008, a Premier Showcase of Scientific Talent, at NSF
National Science Board Releases Science and Engineering Indicators 2008

Douglas H. Fisher

http://www.networkworld.com/community/node/32152

Massive $208 million petascale computer gets green light

The 200,000 processor core system known as Blue Waters got the green light recently as the University of Illinois at Urbana-Champaign and its National Center for Supercomputing Applications (NCSA) said it has finalized the contract with IBM to build the world's first sustained petascale computational system.

Blue Waters is expected to deliver sustained performance of more than one petaflop on many real-world scientific and engineering applications. A petaflop equals about 1 quadrillion calculations per second. They will be coupled to more than a petabyte of memory and more than 10 petabytes of disk storage. All of that memory and storage will be globally addressable, meaning that processors will be able to share data from a single pool exceptionally quickly, researchers said. Blue Waters, is supported by a $208 million grant from the National Science Foundation and will come online in 2011.

Douglas H. Fisher

http://www.sciencedaily.com/releases/2007/08/070810212829.htm

National Science Board Approves Funds For Petascale Computing Systems

ScienceDaily (Aug. 14, 2007) — Today the National Science Board (NSB) approved a resolution authorizing the National Science Foundation (NSF) to fund the acquisition and deployment of the world's most powerful "leadership-class" supercomputer, proposed in response to NSF's "Track 1" supercomputing solicitation. This "petascale" system is expected to be able to make arithmetic calculations at a sustained rate in excess of a sizzling 1,000-trillion operations per second (a "petaflop" per second) to help investigators solve some of the world's most challenging science and engineering research problems.

Douglas H. Fisher

DataNet Vision

• "Science and engineering digital data are routinely deposited in a well-documented form, are regularly and easily consulted and analyzed by specialists and non-specialists alike, are openly accessible while suitably protected, and are reliably preserved."• NSF Cyberinfrastructure Vision for 21st Century Discovery

• Catalyze development of a system of science and engineering data collections that is open, extensible and evolvable.

HECURA

The HECURA initiative invites research and education projects focused on novel programming models, languages, compilers and underlying communication libraries to enable extreme-scale highly computation and data-intensive scientific and engineering applications in high-end computing (HEC) environments.

Researchers in a wide range of science and engineering fields increasingly depend upon computation and simulation to augment theoretical and experimental studies. In fact, modern HEC systems comprised of thousands to tens-of-thousands of processors now allow researchers to solve previously intractable scientific problems. However, researchers depend not only on access to state-of-the-art HEC systems, but on the availability of state-of-the-art tools that allow them to exploit HEC system capabilities to the fullest extent. To simplify the challenges of developing scientific and engineering codes that are portable, scale well on HEC systems, and effectively utilize computational resources, further research focused on programming models and languages for HEC is needed.

The gap between HEC processing power and HEC application performance points to a need to develop practical parallel programming models that provide ease of use, increase programmer productivity, and improve application performance and scalability, while taking advantage of underlying hardware, communication interconnects and operating systems.

This solicitation is concerned exclusively with research in novel programming models, languages, compilers and underlying communication libraries for extreme-scale highly computation- and data-intensive scientific and engineering applications. The target hosts for these tools are next generation systems comprised of tens-of-thousands to hundreds-of-thousands of processors.