

*Cyberinfrastructure  
and the  
Office of Cyberinfrastructure  
(OCI)*

Campus Research Computing Cyberinfrastructure

Workshop

José Muñoz and Guy Almes  
Office of Cyberinfrastructure



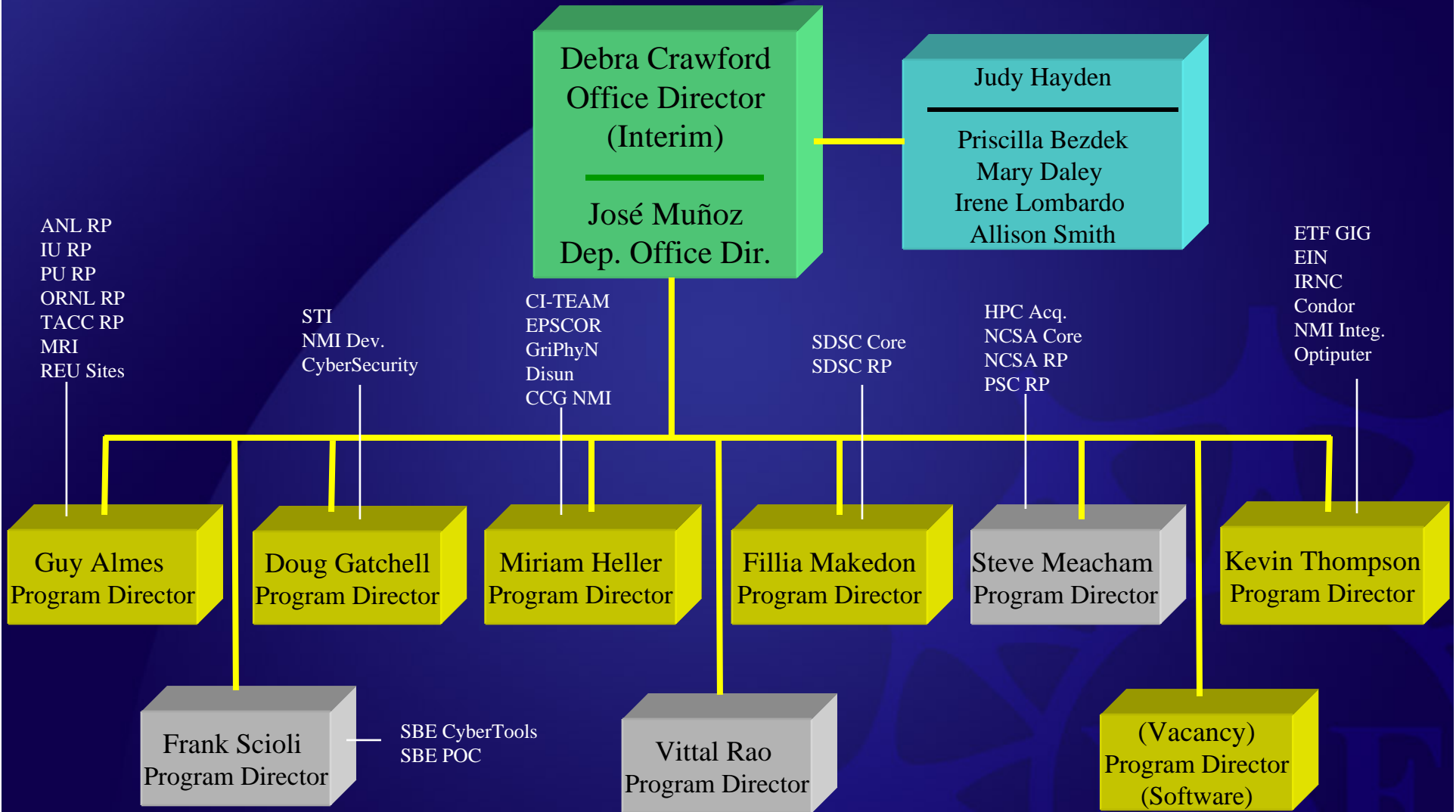


# Cyberinfrastructure (CI) Governance

- Dr. Dan Atkins (Univ Michigan) selected as new OCI Director
  - chaired 2003 Blue-Ribbon Advisory Panel on Cyberinfrastructure: *Revolutionizing Science and Engineering Through Cyberinfrastructure*
  - tenure begins June 2006
- Cyberinfrastructure Council (CIC)
  - NSF ADs and ODs, chaired by Dr. Bement (NSF Director)
  - CIC oversees NSF's broad cyberinfrastructure portfolio
- Advisory Committee for NSF's CI activities and portfolio
  - first meeting planned June 2006



# Office of Cyberinfrastructure





# Strategic Plan (FY 2006 – 2010)

- Ch. 1: Call to Action  
Strategic Plans for:
- Ch. 2: High Performance Computing
- Ch. 3: Data, Data Analysis & Visualization
- Ch. 4: Colaboratories, Observatories and Virtual Organizations
- Ch. 5: Learning and Workforce Development

Completed in Summer 2006

## NSF'S CYBERINFRASTRUCTURE VISION FOR 21<sup>ST</sup> CENTURY DISCOVERY

NSF Cyberinfrastructure Council

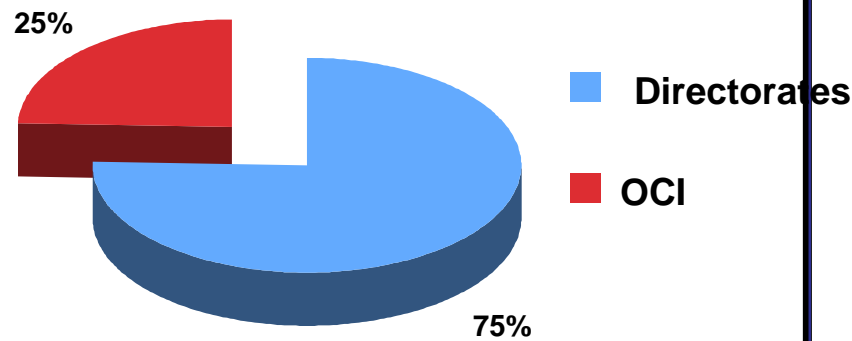


National Science Foundation  
September 26, 2005  
Version 4.0

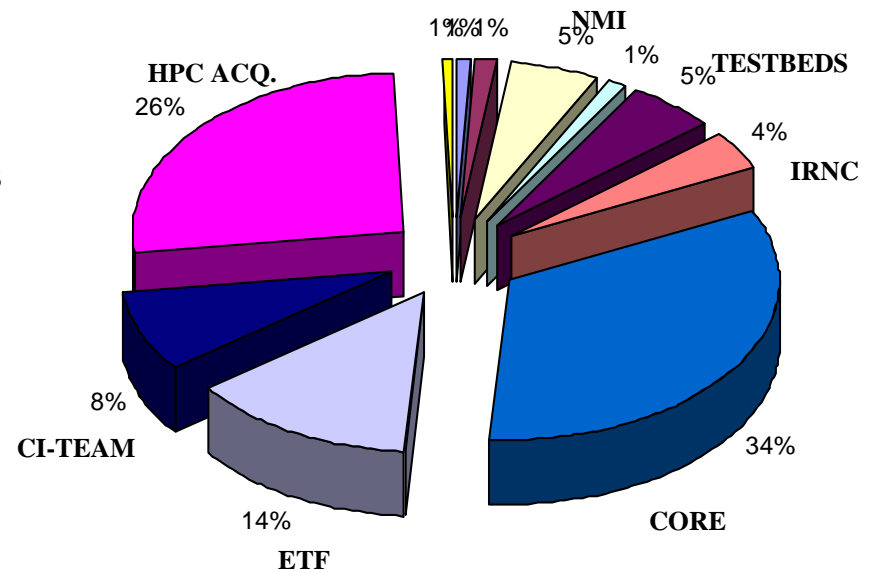


# CyberInfrastructure Budgets

NSF 2006 CI Budget



OCI Budget: \$127M (FY06)

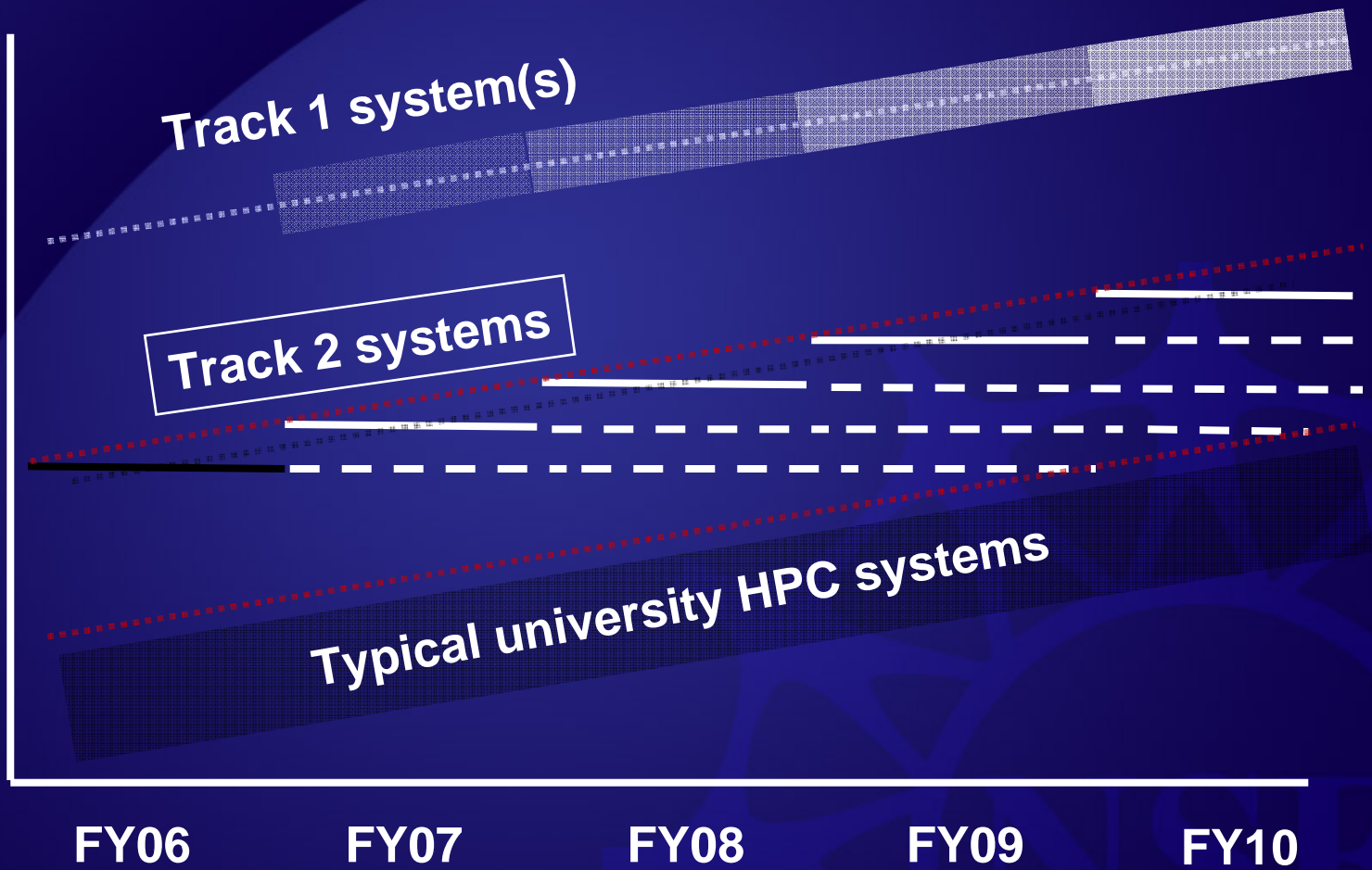


**OCI Request: \$182.42 (FY07)**



# Acquisition Strategy

Science and engineering capability  
(logarithmic scale)





# TeraGrid: What is It?

## TeraGrid:

- (1) Provides a unified user environment to support high-capability, production-quality cyberinfrastructure services for science
- (2) Provides new S&E opportunities – by making new ways of using distributed resources and services

- GIG responsible for integration:

- Software integration
- Base infrastructure
- User support
- Community engagement

- 8 Resource Providers:

- PSC, TACC, NCSA, SDSC, ORNL, Indiana, Purdue, Argonne



Examples of services include:

- HPC
- Data collections
- Visualization servers
- Portals

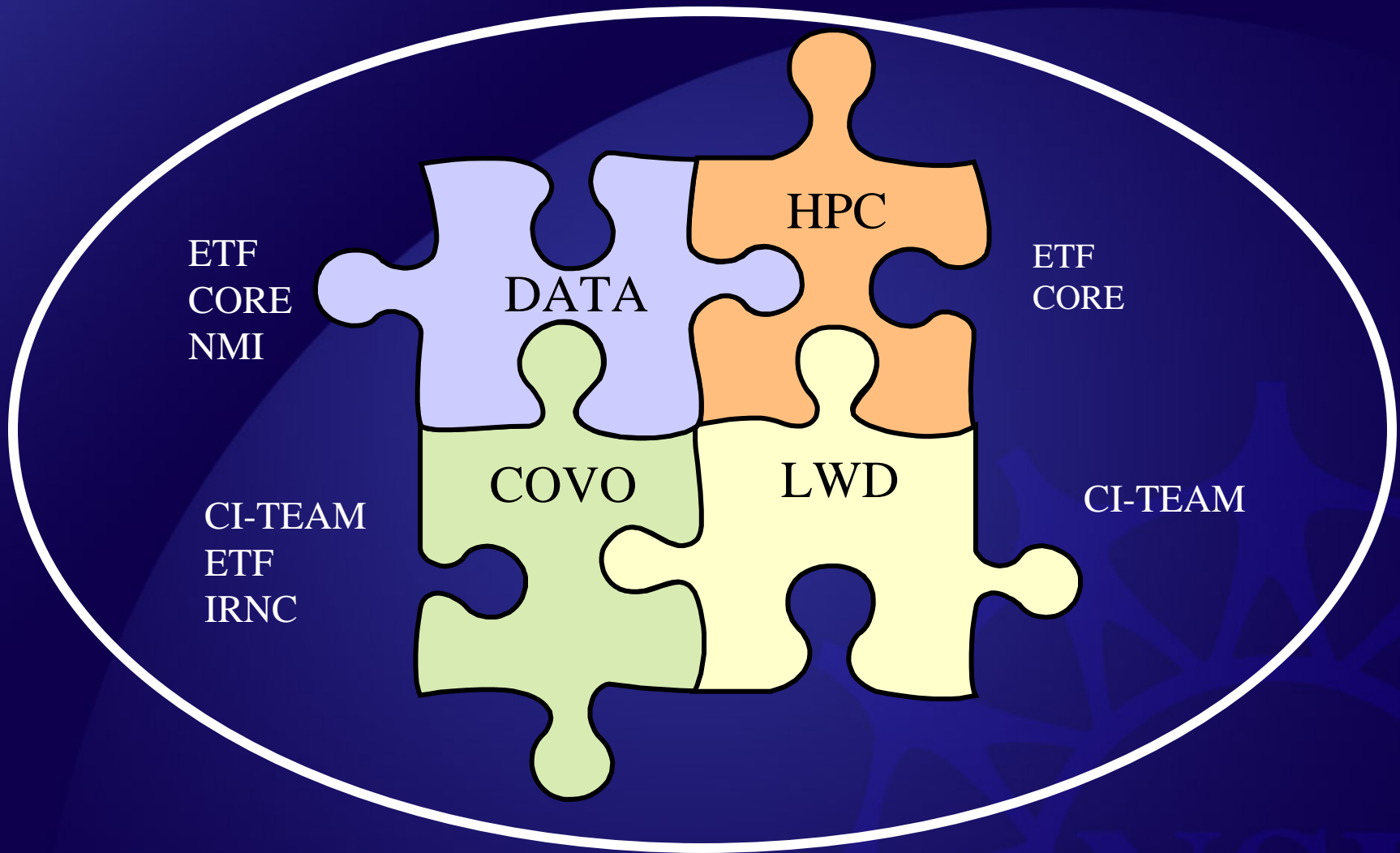


# Other Major OCI Programs

- NSF Middleware Initiative
  - ☞ Supports development, deployment, and sustaining of key grid/other software
  - ☞ Globus, Shibboleth, Condor
- International Research Network Connections
  - ☞ International connectivity to Europe, Asia, Australia, and Latin America
- CI Team
  - ☞ Workforce development
  - ☞ Recent solicitation



# How it all fits together...





# HECURA – 2005/2006 SCOPE (CISE)

- File Systems Research
- Future File Systems related protocols
- I/O middleware
- Quality of Service
- Security
- Management, reliability, and availability at scale
- Archives/Backups as extensions to file systems
- Novel storage devices for the IO stack
- I/O Architectures
- Hardware and software tools for design, simulation of I/O, file and storage systems.
- Efficient benchmarking, tracing, performance measurement and tuning tools of I/O, file and storage systems



# HPC Benchmarking

- HPC Challenge Benchmarks (<http://icl.cs.utk.edu/hpcc/>)
  1. HPL - the Linpack TPP benchmark which measures the floating point rate of execution for solving a linear system of equations.
  2. DGEMM - measures the floating point rate of execution of double precision real matrix-matrix multiplication.
  3. STREAM - a simple synthetic benchmark program that measures sustainable memory bandwidth (in GB/s) and the corresponding computation rate for simple vector kernel.
  4. PTRANS (parallel matrix transpose) - exercises the communications where pairs of processors communicate with each other simultaneously. It is a useful test of the total communications capacity of the network.
  5. RandomAccess - measures the rate of integer random updates of memory (GUPS).
  6. FFTE - measures the floating point rate of execution of double precision complex one-dimensional Discrete Fourier Transform (DFT).
  7. *Communication bandwidth and latency* - a set of tests to measure latency and bandwidth of a number of simultaneous communication patterns; based on b\_eff (effective bandwidth benchmark).



# DARPA HPCS

- Partitioned Global Address Space (PGAS) programming paradigm
  - ☞ Intended to support scaling to 1000s of processors
  - ☞ Co-Array Fortran
  - ☞ Unified Parallel C
  - ☞ Cray's Chapel
  - ☞ IBM's X10
  - ☞ Sun's Fortress
- DARPA HPCS productivity activities
  - ☞ HPC specific programming environments?



# HPC Acquisition - Track 1

- Increased funding will support first phase of a petascale system acquisition
- Over four years NSF plans to spend \$200M
- Acquisition is critical to NSF's multi-year plan to deploy and support world-class HPC environment
- Collaborating with sister agencies with a stake in HPC
  - ☞ DARPA, HPCMOD, DOE/OS, DOE/NNSA, NIH
- Even the electric bill is impressive



# Cyberinfrastructure Summary

## “The Tide that Raises All Boats”

- Cyberinfrastructure enables a broad spectrum of science and engineering activities
- NSF Cyberinfrastructure deployment/acquisition activities must be driven by the needs of the science, engineering and education communities
- Many opportunities to work with other federal agencies
- Work required in all kinds of Cyberinfrastructure software
  - ☞ application and middleware for petascale systems
  - ☞ systems software
- Cyberinfrastructure also broadens participation



In the end...

**It's all about the SCIENCE**