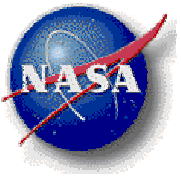




NCCS User Forum

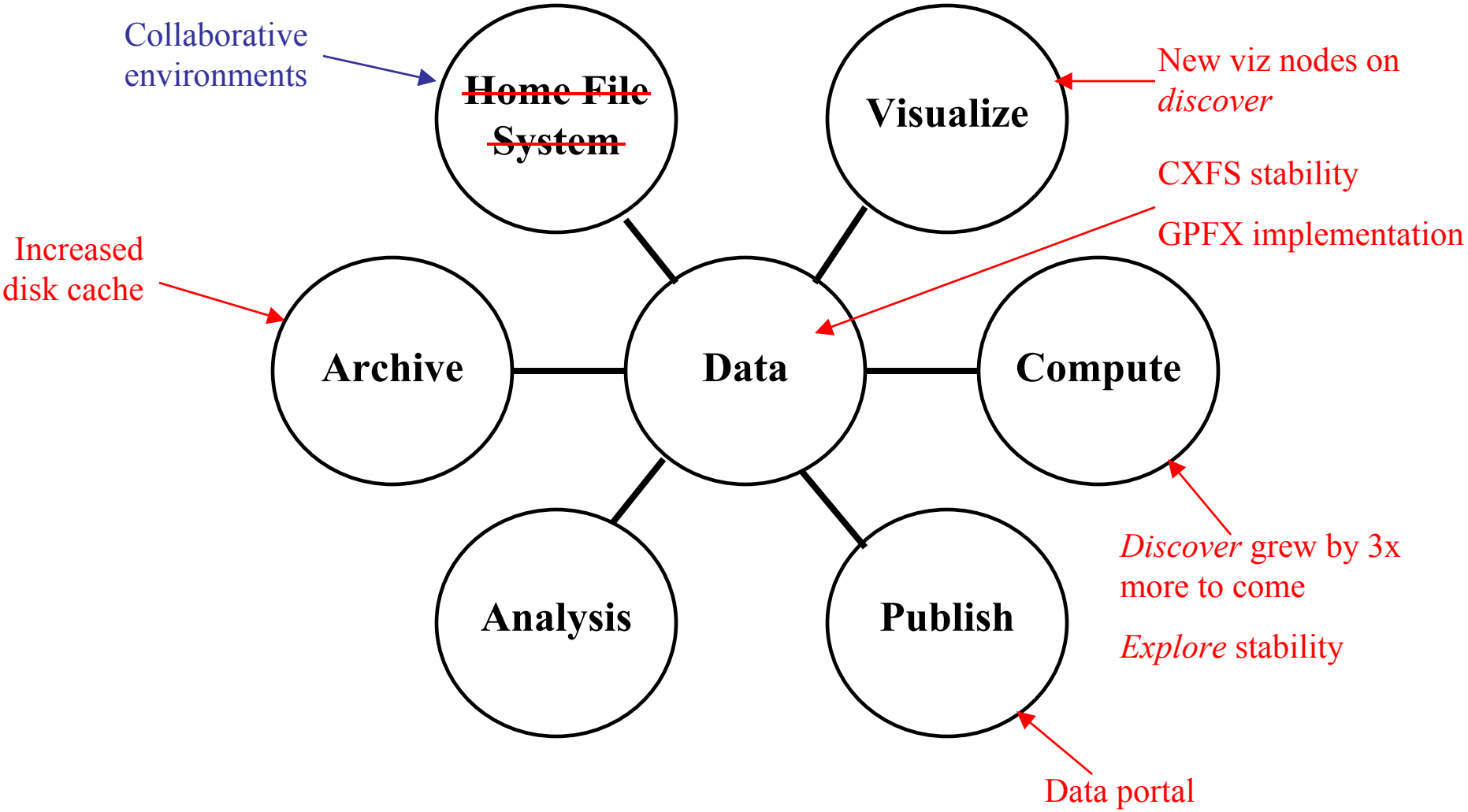
26 April 2007

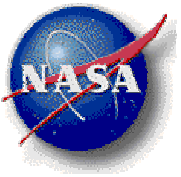
Agenda

- Introduction — Phil Webster
- Systems Status—Mike Rouch
- NREN to NISN—Phil Webster
- New Data Sharing Services—Harper Pryor
- User Services—Sadie Duffy
- Questions or Comments

Conceptual Architecture







Halem



- Halem will be retired 1 ~~April 2007~~ May 1
  - Four years of service
  - Self maintained for over 1 year
- Replaced by Discover
  - Factor of 3 capacity increase
  - Migration activities completed
  - "We need the cooling & power"
- Status
  - Up and running during "excess" process
  - Un-supported, and files are not backed up
  - Disk may be removed, software licenses moving to discover
  - Efforts will not be made to recover the system in the event of a major system failure

Last User Forum

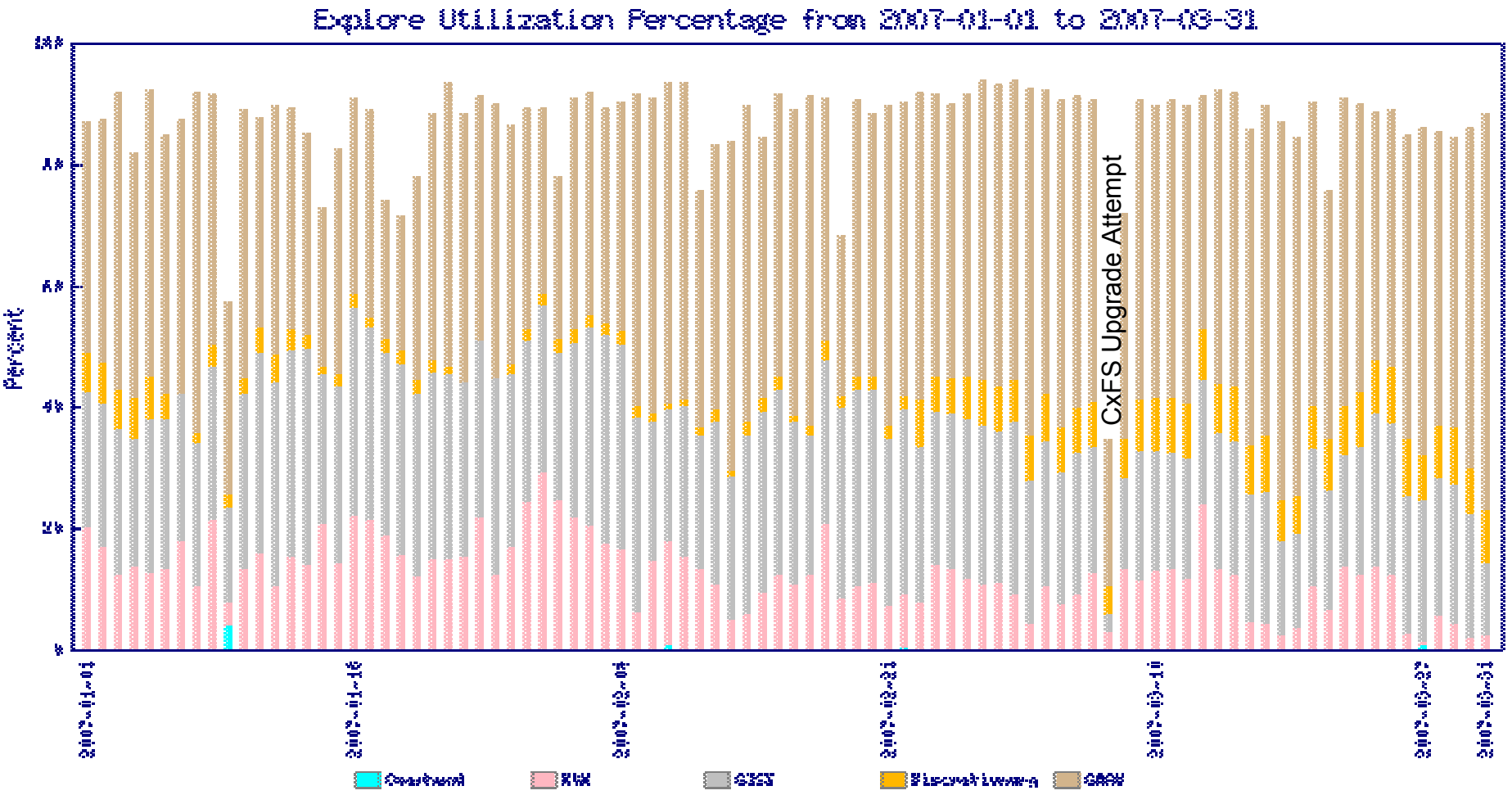
Agenda

- Introduction — Phil Webster
- Systems Status—Mike Rouch
- NREN to NISN—Phil Webster
- New Data Sharing Services—Harper Pryor
- User Services—Sadie Duffy
- Questions or Comments

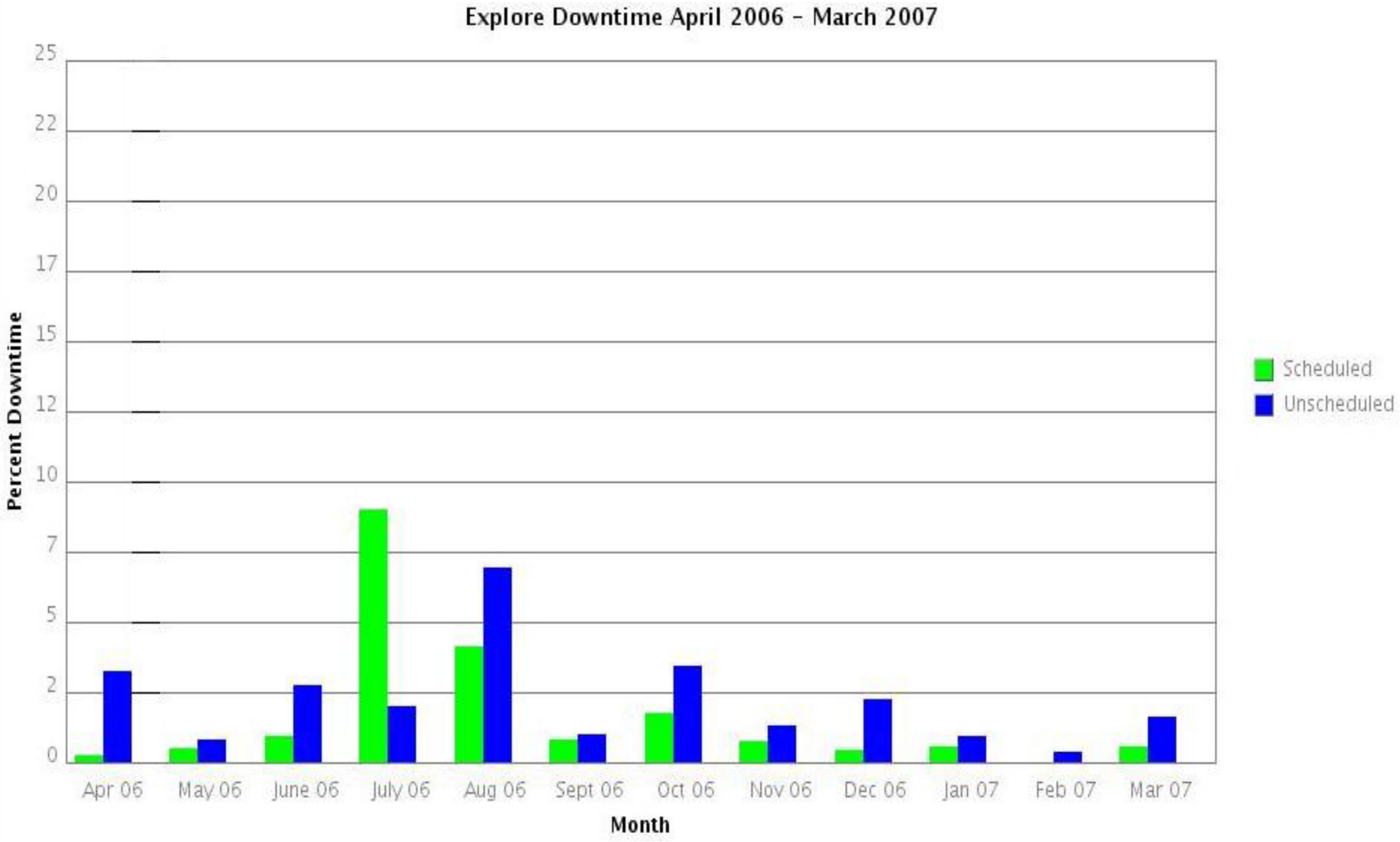
Explore Utilization
Jan – March 2007



NASA Center for Computational Sciences



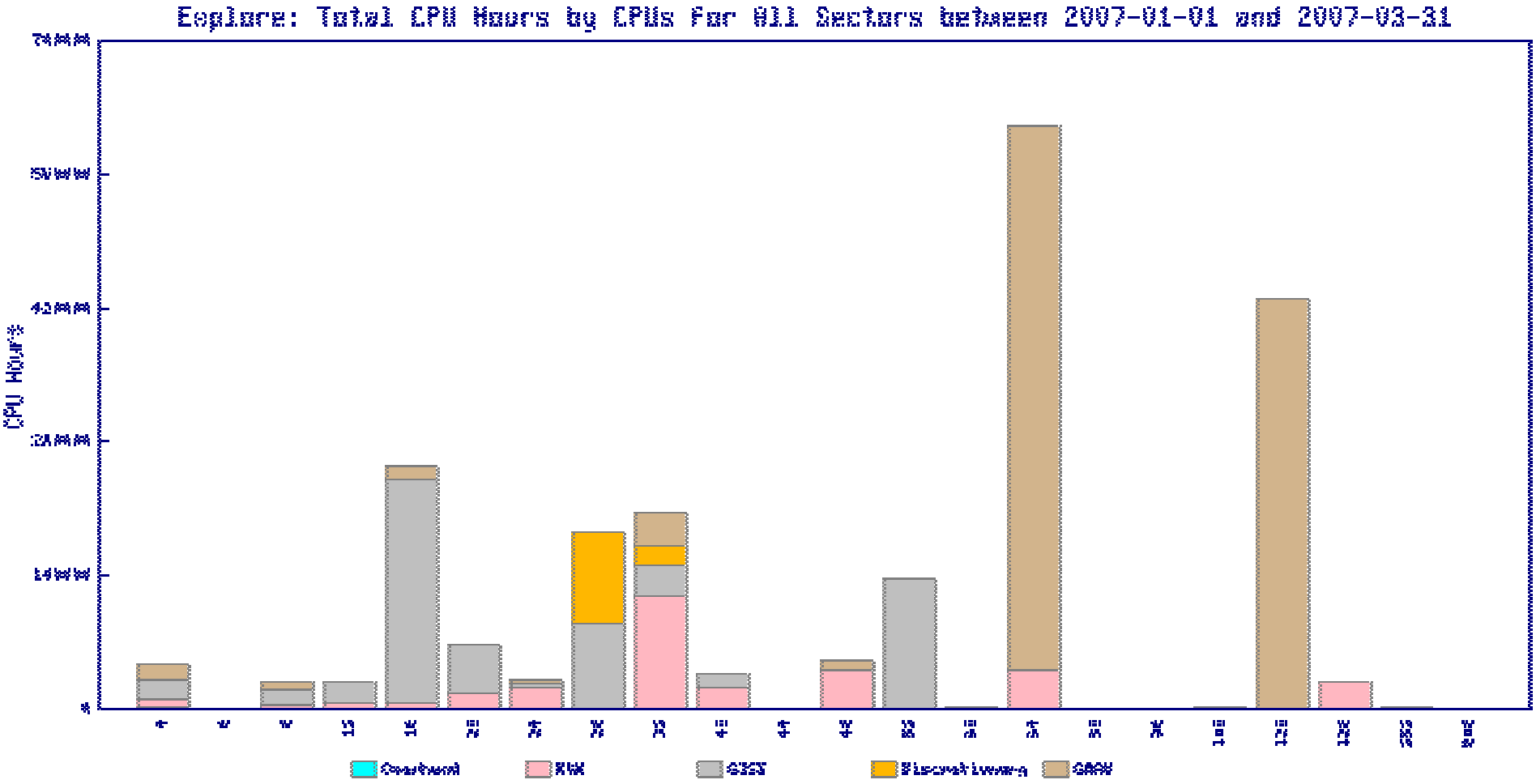
SGI Explore Downtime



Explore CPU Usage
Jan – March 2007



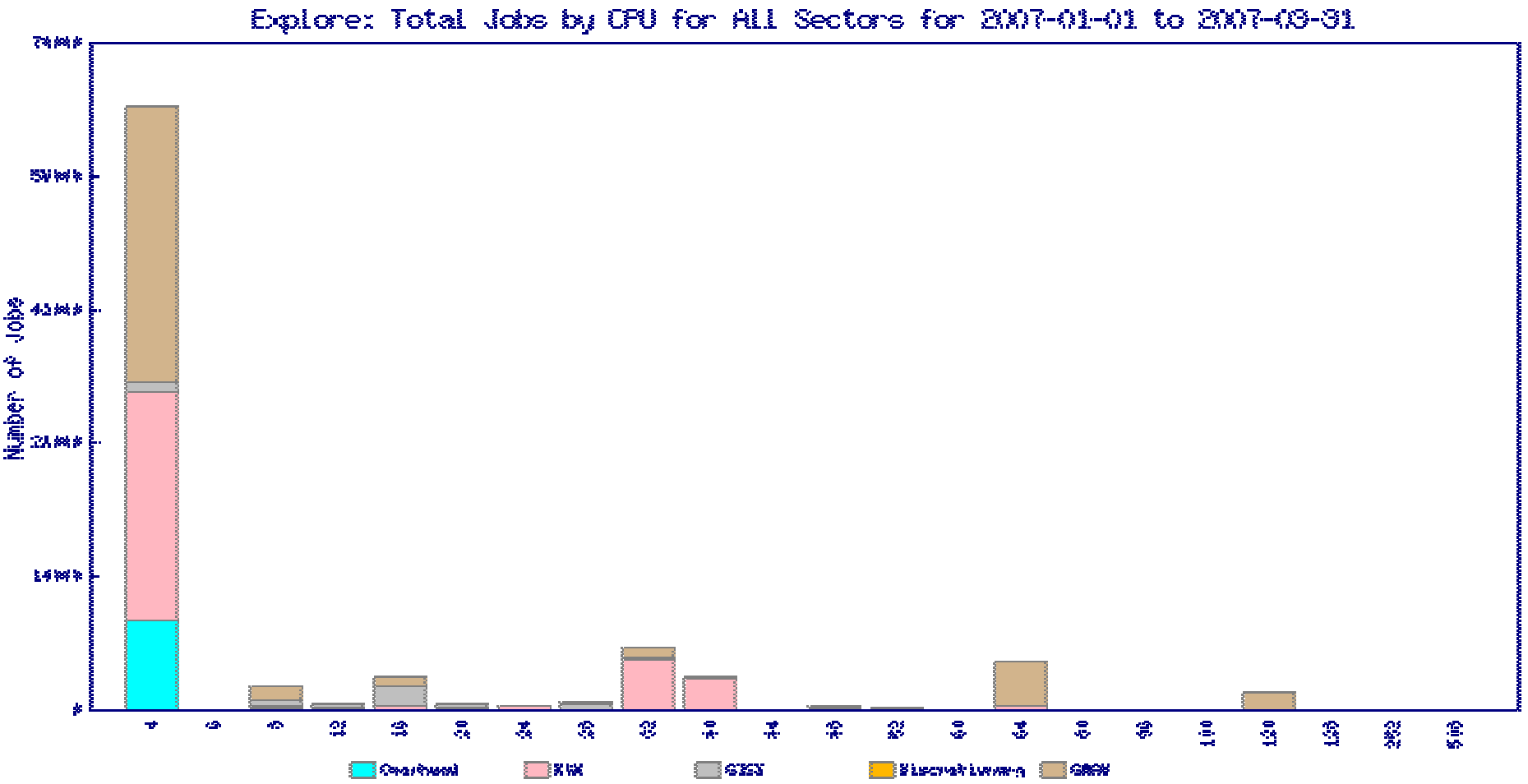
NASA Center for Computational Sciences



Explore Job Mix
Jan – March 2007

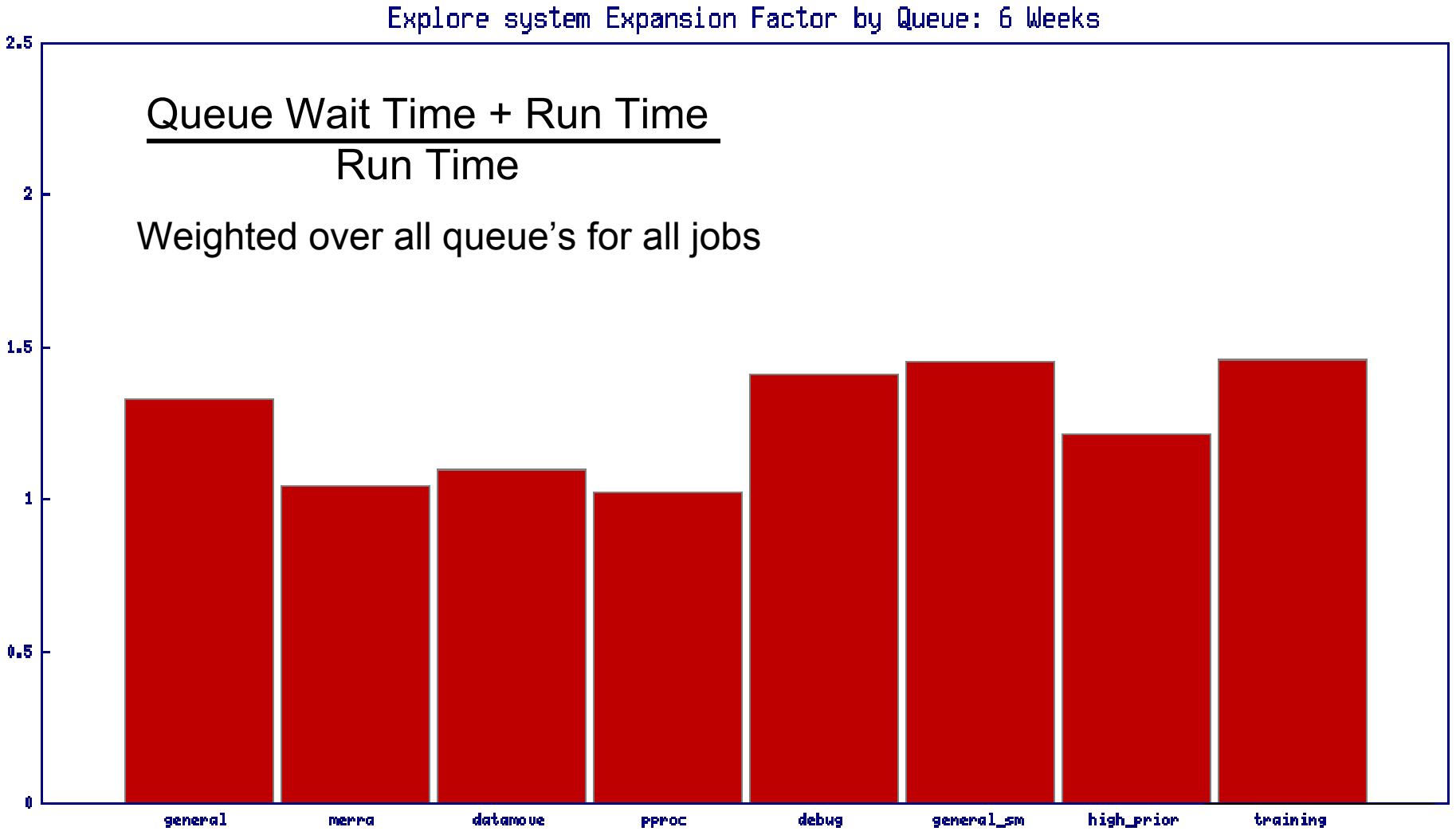


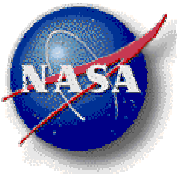
NASA Center for Computational Sciences



Explore Queue Expansion Factor







Explore Issues



Eliminate Data Corruption – SGI Systems
On Going Process

Issue: Files being written at the time of an SGI system crash **MAY** be corrupted. However, files appear to be normal.

Interim Steps: Careful Monitoring

- ✓ Install UPS – ***COMPLETED 4/11/2007***
- ✓ Continue Monitoring
- ✓ Daily Sys Admins scan files for corruption and directly after a crash
- All affected users are notified

Fix: SGI will provide XFS file system patch

- Awaiting fix
- Will schedule installation after successful testing

Explore Improvements



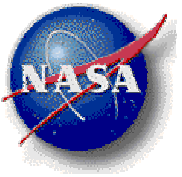
Reduced Impact of Power Outages
COMPLETED

Issue: Power fluctuations during thunderstorms

Effect: Systems lose power and crash; Reduce system availability; Lower system utilization; Reduce productivity for users

Fix: Acquire & install additional UPS systems

- ✓Mass Storage Systems - Completed
- ✓New LNXI System - Completed
- ✓SGI Explore Systems - **COMPLETED 4/11/2007**



Explore Improvements



Enhanced NoBackup Performance on Explore
On Going Process

Issue: NoBackup Shared file system poor I/O performance

Effect: Slow job performance

Fix: From the Acquired additional disks discussed last quarter

- ✓ Creating More NoBackup File Systems
- ✓ Spread out the load across more file systems
- ✓ Upgraded System I/O hba's – 4GB
- ✓ Implementing New FC Switch – 4GB
- – **On Going Process** – Improvements have been made; striving for more

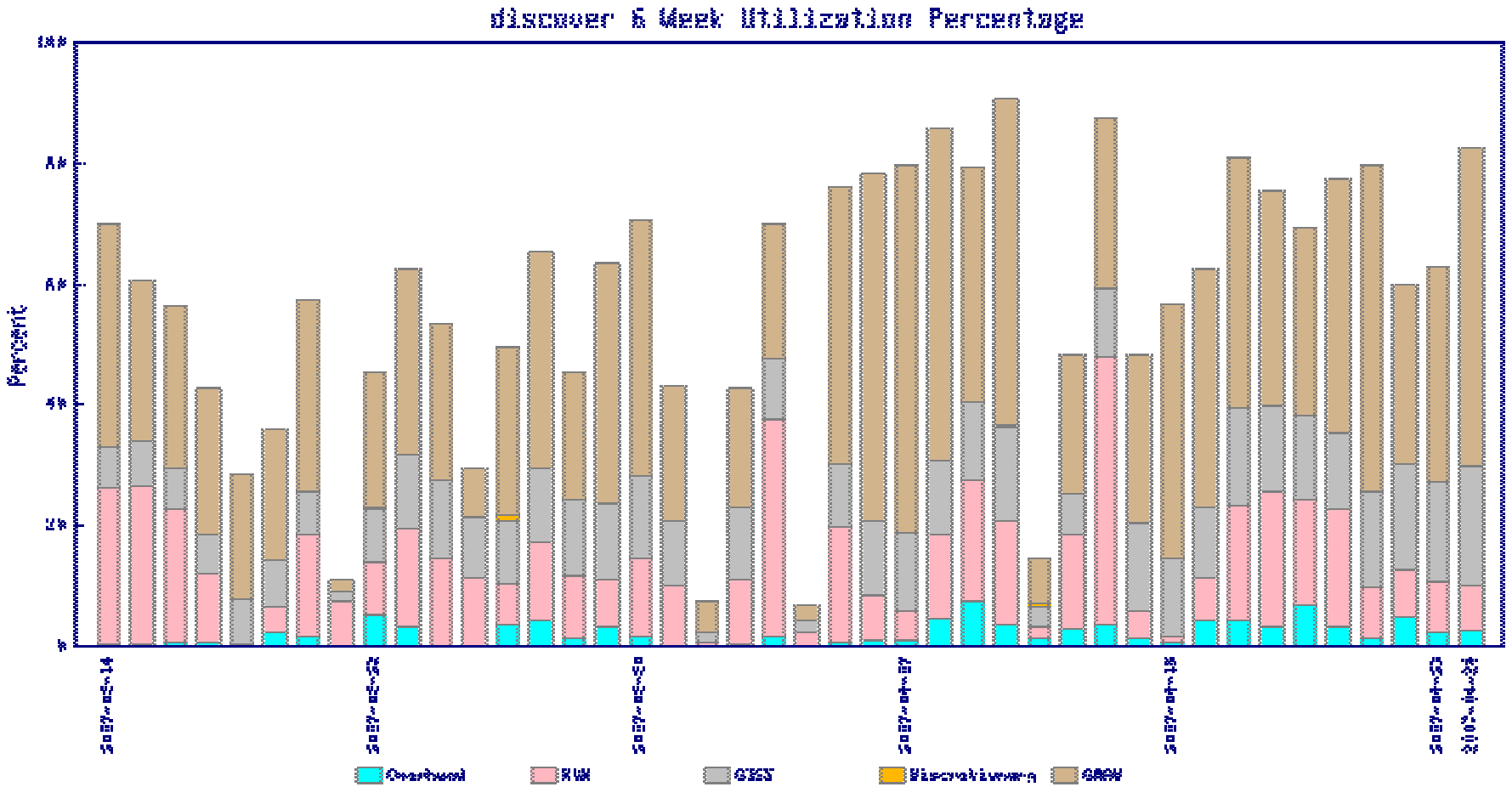
Explore Improvements

- **Improving File Data Access – Q2 2007**
  - Increase File System Data Residency from Days to Months
  - ✓ Analysis completed; New File System being created
  - ✓ Scheduling with users to move data into new file systems
- **Increasing Tape Storage Capacity – Q1 2007**
  - ✓ New STK SLA8500 (2 x 6500 slot library) (Jan 07)
  - ✓ 12 new Titanium tape drives (500 GB Tape) (Jan 07)
  - ✓ ~6PB Total Capacity
  - ✓ **Completed 3/2007**
- **Enhancing Systems – Q2 2007**
  - Software OS & CxFS upgrades to Irix (**May 9, 2007**)
  - Software OS & CxFS upgrades to Altix (**May 9, 2007**)

Discover Utilization
Last 6 Weeks

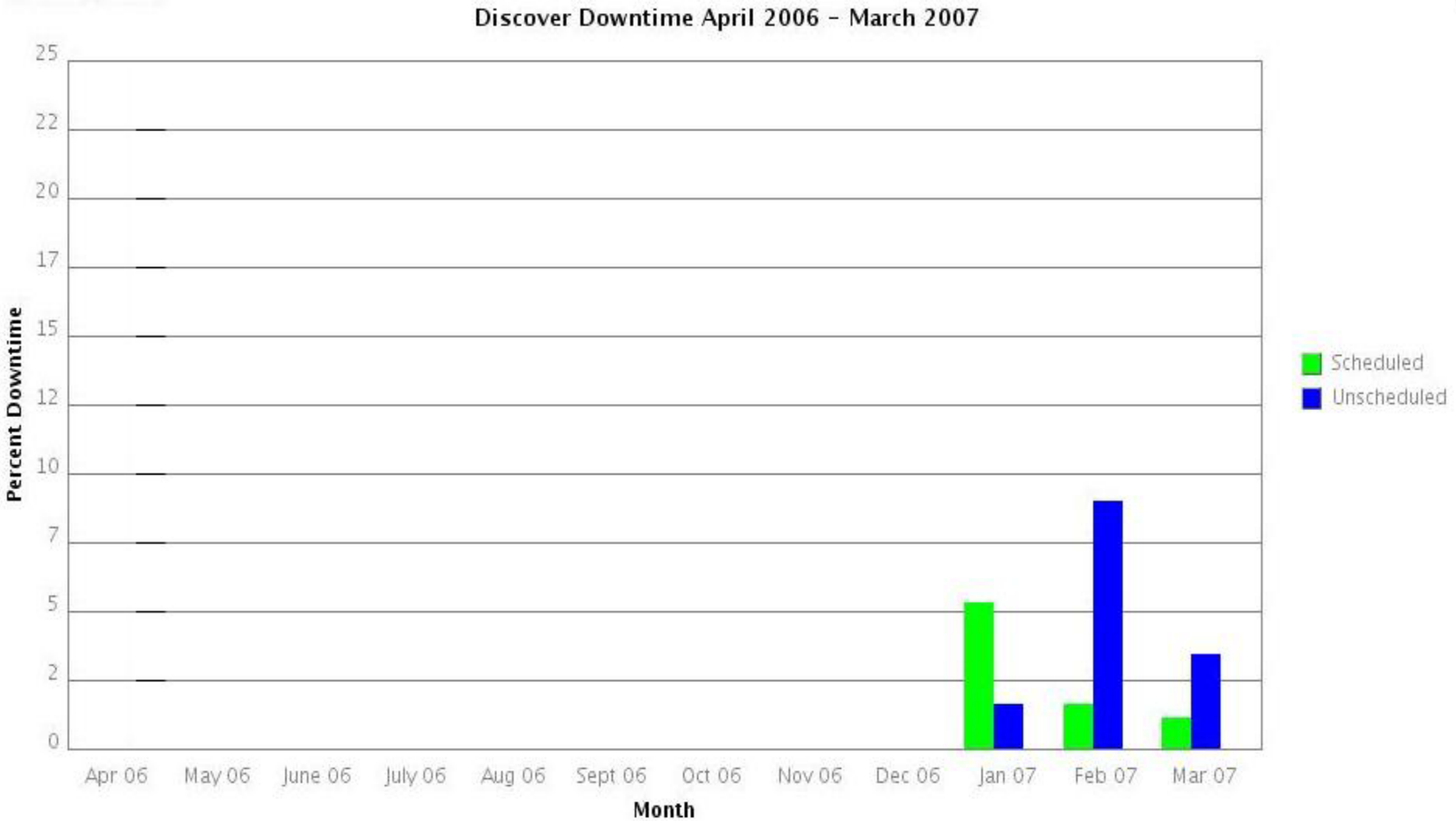


NASA Center for Computational Sciences



Discover Cluster Downtime

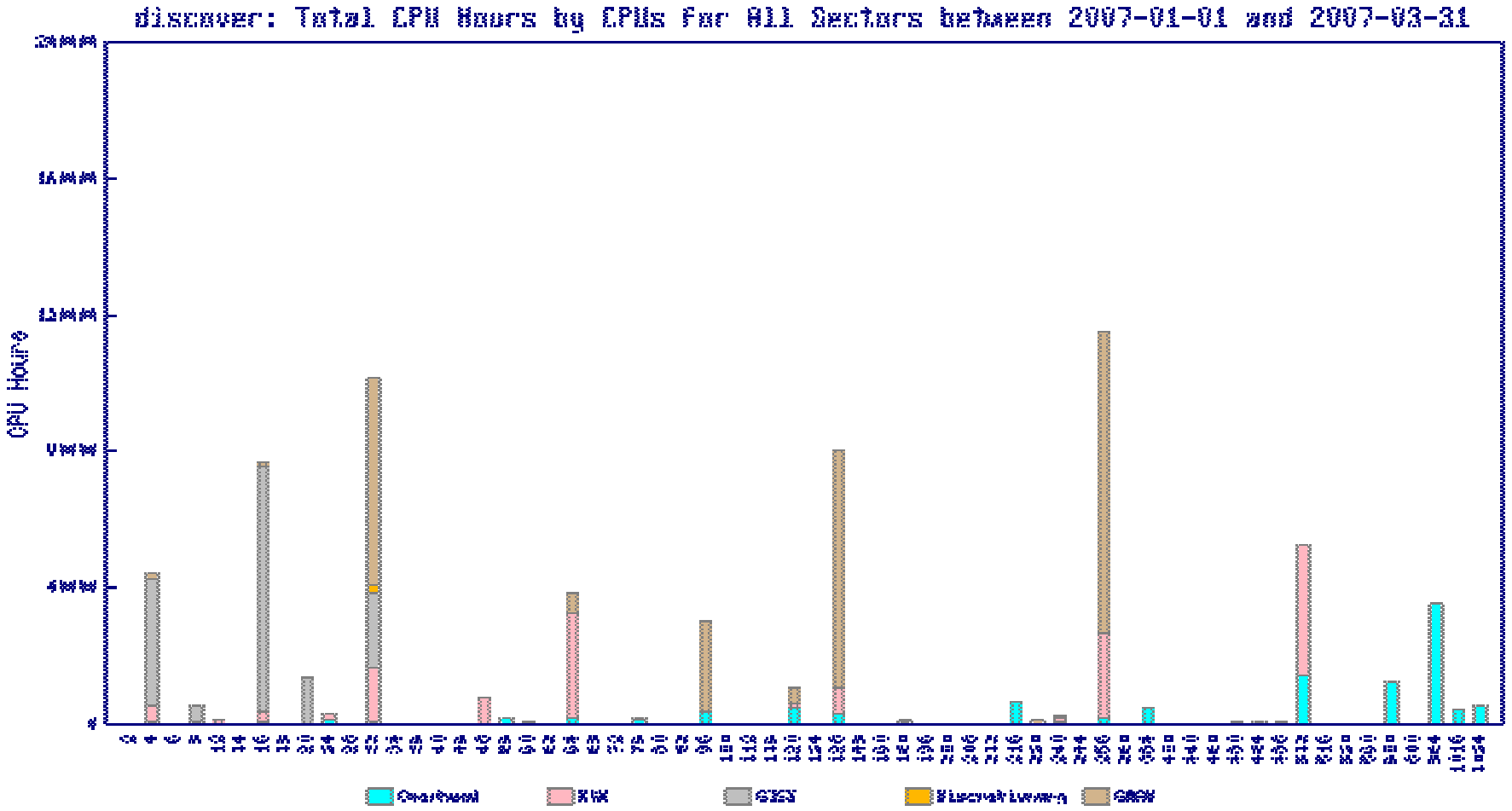




Discover CPU Usage
Jan – March 2007



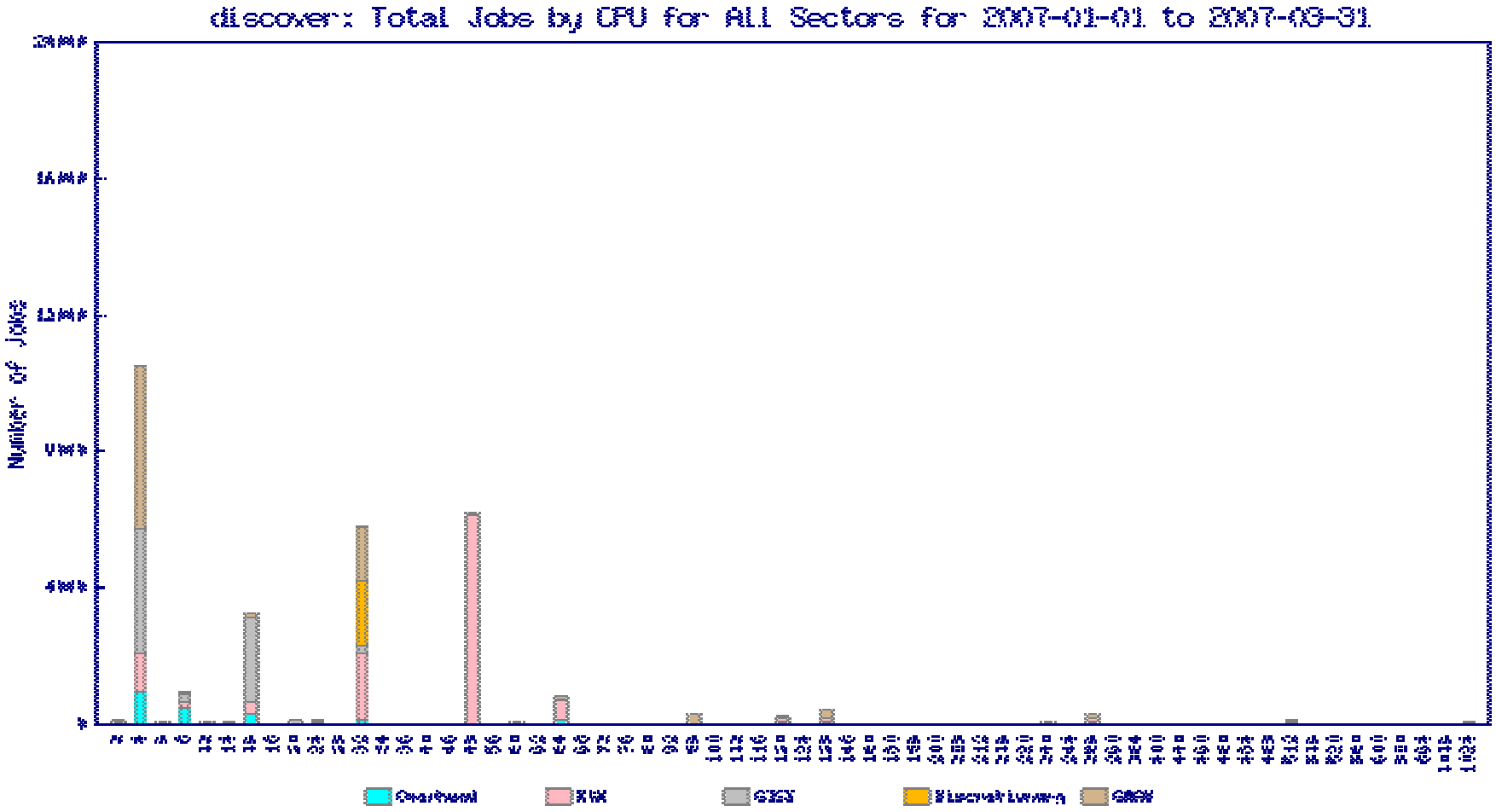
NASA Center for Computational Sciences



Discover Job Mix
Jan – March 2007



NASA Center for Computational Sciences



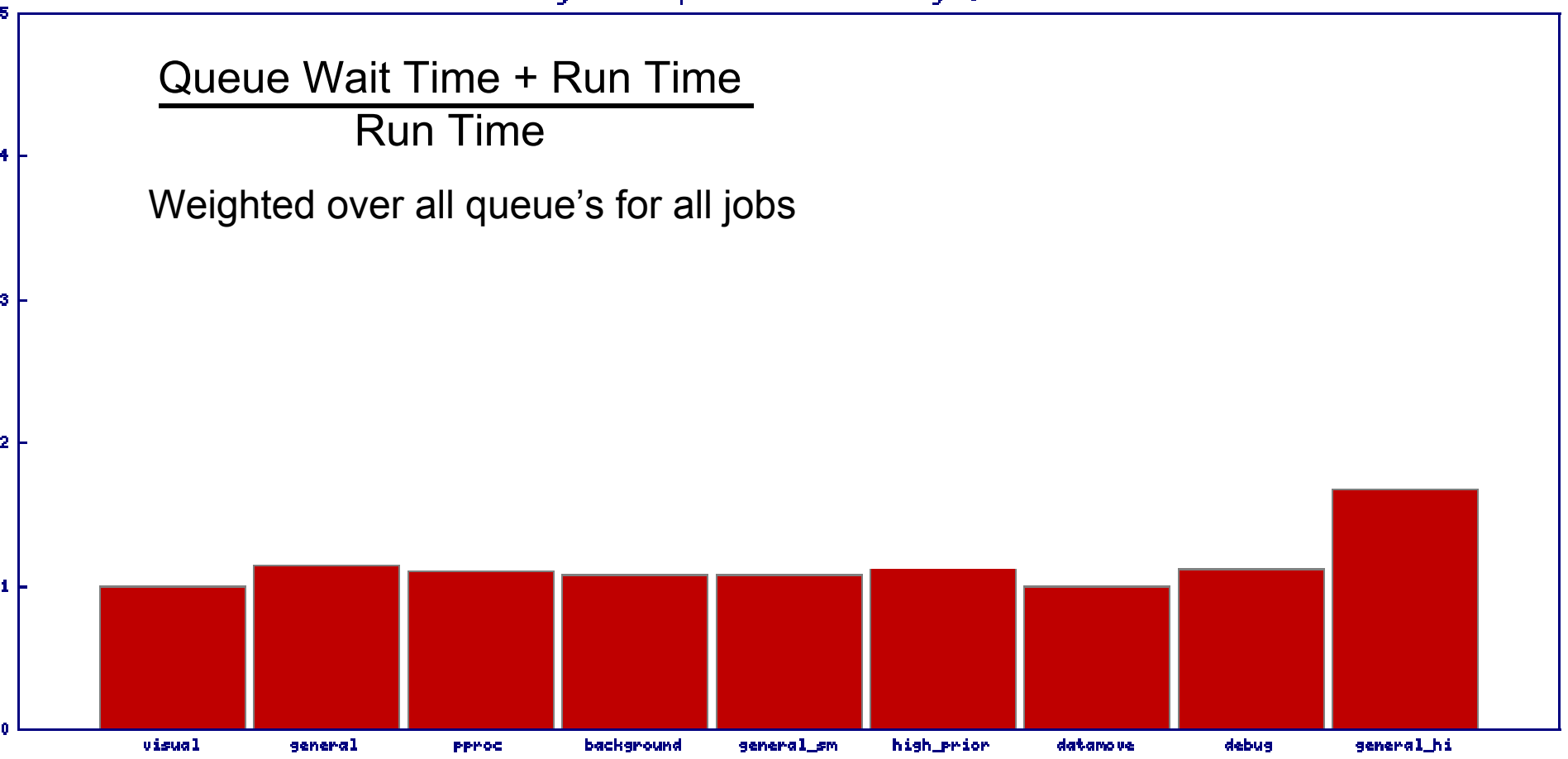
Discover Queue Expansion Factor



discover system Expansion Factor by Queue: 6 Weeks

$$\frac{\text{Queue Wait Time + Run Time}}{\text{Run Time}}$$

Weighted over all queue's for all jobs

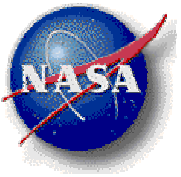


Discover Status



- **SCU1 unit accepted**
  - General Availability
  - **User environment still evolving**
    - ✓Tools: IDL, TotalView – **DONE**
    - Libraries: different MPI versions – Intel very close
    - Other software: sms, tau, papi – Work in progress
- **PBS queues up and running jobs!**
  - New Submit Option
    - -r y means it IS rerunable
    - -r n means it's NOT rerunable

If you need anything – please call User Services

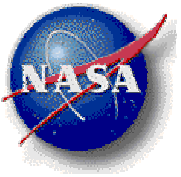


Recent Issues Discover



COMPLETED

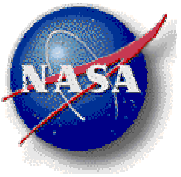
- **Memory leak**
  - **Symptom:** Total memory available to user processes slowly decreases
  - **Outcome:** The same job will eventually run out of memory and fail
  - **Fix:** Silverstorm released a fix and it has been implemented
- **10 GbE problem**
  - **Symptom:** 10 GbE interfaces on gateway nodes are not working
  - **Outcome:** Intermittent access to cluster and Altix systems
  - **Fix:** The infiniband manufacturer released a firmware fix for the problem and currently 10GbE enabled



Recent Issues Discover



- **PBS** ***COMPLETED***
  - **Symptom:** When a new node is added to the PBS server list of known nodes, information about that node including its IP address and naming information must be sent to all the nodes causing a reboot to take many hours longer than normal.
  - **Outcome:** Altair generated a new start up procedure and a fix
  - **Fix: Completed** The new startup procedure is working
- **DDN Controller Hardware**
  - **Symptom:** After a vendor recommended firmware upgrade a systemic problem was identified with the hardware causing the file systems to become unavailable
  - **Outcome:** Vendor replaced it with newer generation hardware
  - **Fix: Completed**

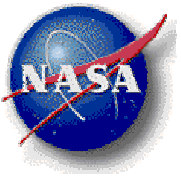


Current Issues Discover



On Going Process

- **Job goes into Swap**
  - **Symptom:** When a job is running, one or more nodes goes into a swap condition
  - **Outcome:** The processes on those nodes runs very slow causing the total job to run slower.
  - **Progress:** Monitoring is in place to trap this condition. The monitoring is working for majority instances. As long as the nodes do not run out of swap, the job should terminate normally.



Current Issues Discover



- **Job Runs Out of Swap**
  - **Symptom:** When a job is running, one or more nodes run out of swap
  - **Outcome:** The nodes become hopelessly hung, requires a reboot and the job dies.
  - **Progress:** Monitoring in place to catch this condition, kill the job before it runs out of swap, notify the user and examine the job. The monitoring is working for majority instances. Also, scripts are in place to cleanup after this condition and it is also working for majority instances.
  - **NOTE:** If your job fails abnormally please open a ticket so we can analyze why the monitoring scripts did not catch the condition so it can be updated with the new error checking.

What's New?



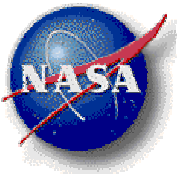
- Addition of viz nodes (16)
  - Opteron based with viz tools
    - IDL – Working through PBS queue called "visual"
  - Access to all the same GPFS file systems as the Discover Cluster
  - Viz environment still evolving
  - Pioneer use available by sending a request to User Services (support@nccs.nasa.gov)
- Addition of test system – May 2007

Agenda

- Introduction — Phil Webster
- Systems Status—Mike Rouch
- NREN to NISN—Phil Webster
- New Data Sharing Services—Harper Pryor
- User Services—Sadie Duffy
- Questions or Comments

NASA HEC WAN Migration

- HEC Program Office made a strategic decision to migrate from NASA Research and Engineering Network (NREN) to NASA Integrated Services Network (NISN)
  - NREN joined the National Lambda Rail (NLR) project to provide 10 Gbps WAN services to a number of NASA centers
  - NISN is upgrading their WAN infrastructure to provide 10 Gbps service between NCCS & NAS in 6 months, with 10 Gbps service to all NASA centers in ~24 months
- High speed WAN access maintained to universities and research centers
- The HEC Program is working with NISN to implement a practical transition strategy to ensure minimal disruption to users



Phased NISN-HEC Upgrades

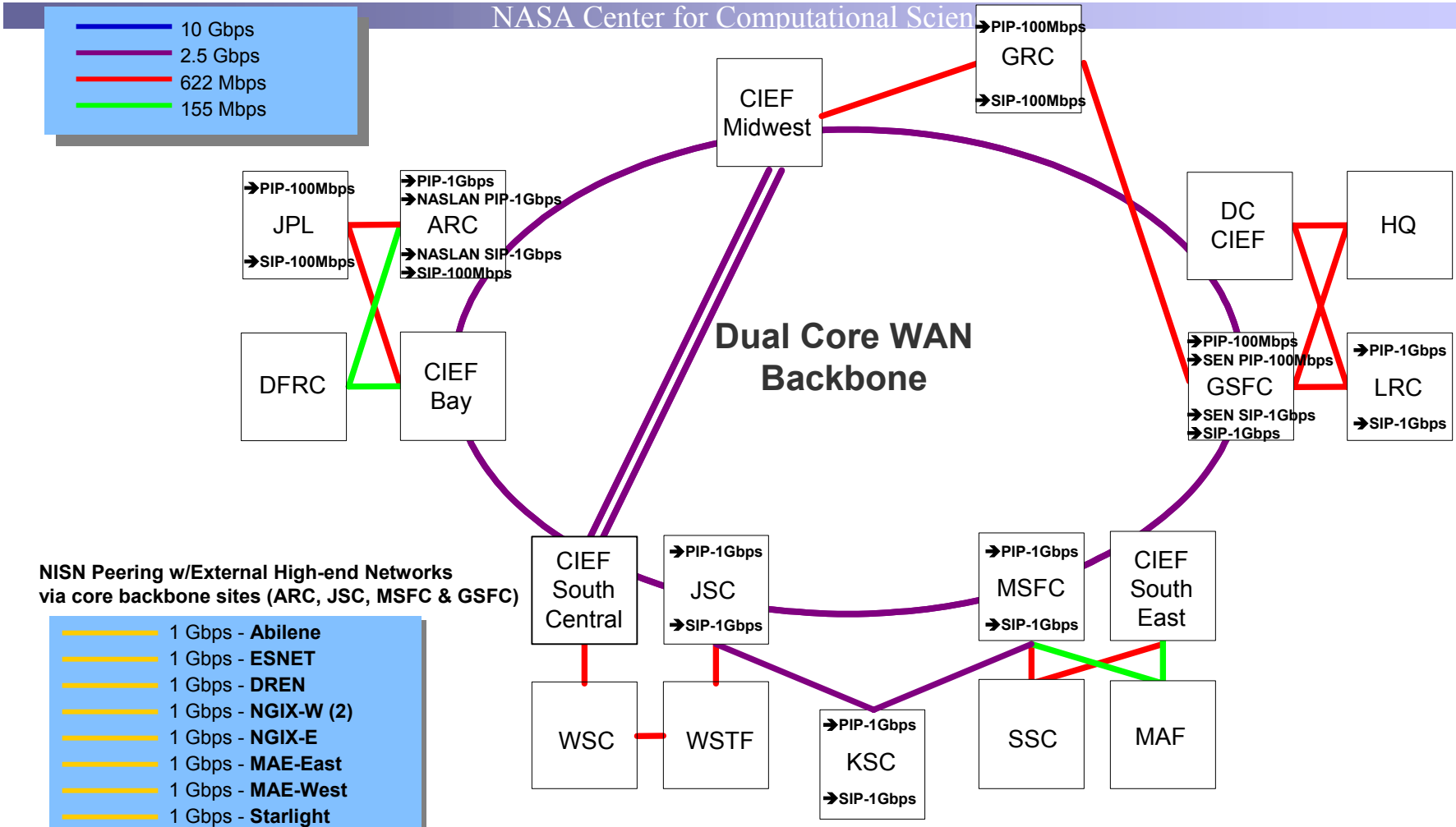


- NISN Backbone (Today)
  - GSFC PIP 100 Mbps
  - GSFC/SEN PIP 100 Mbps
  - GSFC/SEN SIP 1 Gbps
  - GSFC SIP 1 Gbps
  - Core backbone 2.5 Gbps
- NISN-HEC Step 2 (6 Months)
  - Establishes direct 10 Gbps link between ARC & GSFC
  - GSFC PIP upgrade to 1 Gbps
  - GSFC/SEN PIP upgrade to 10 Gbps
- NISN-HEC Step 3 (24 Months)
  - Core backbone upgrade to 10 Gbps

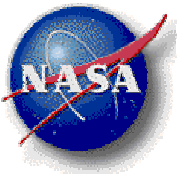
***** Highlights for GSFC users do not represent all planned NISN upgrades *****

NISN Backbone (Today)





NISN-HEC Step 2 (6 Months)



10 Gbps
2.5 Gbps
622 Mbps
155 Mbps

NASA Center for Computational

CIEF Midwest

GRC
PIP-100Mbps
SIP-100Mbps

JPL
PIP-100Mbps
SIP-100Mbps

ARC
PIP-10Gbps
SIP-1Gbps

DFRC

CIEF Bay

Dual Core WAN Backbone

DC CIEF

HQ

GSFC
PIP-1Gbps
SEN/PIP-10Gbps
SIP-1Gbps

LRC
PIP-1Gbps
SIP-1Gbps

CIEF South Central

JSC
PIP-1Gbps
SIP-1Gbps

MSFC
PIP-1Gbps
SIP-1Gbps

CIEF South East

WSC

WSTF

KSC
PIP-1Gbps
SIP-1Gbps

SSC

MAF

NISN Peering w/External High-end Networks via core backbone sites (ARC, JSC, MSFC & GSFC)

1 Gbps - **Abilene**
1 Gbps - **ESNET**
1 Gbps - **DREN**
1 Gbps - **NGIX-W (2)**
1 Gbps - **NGIX-E**
1 Gbps - **MAE-East**
1 Gbps - **MAE-West**
1 Gbps - **Starlight**
1 Gbps - **PAIX (2)**

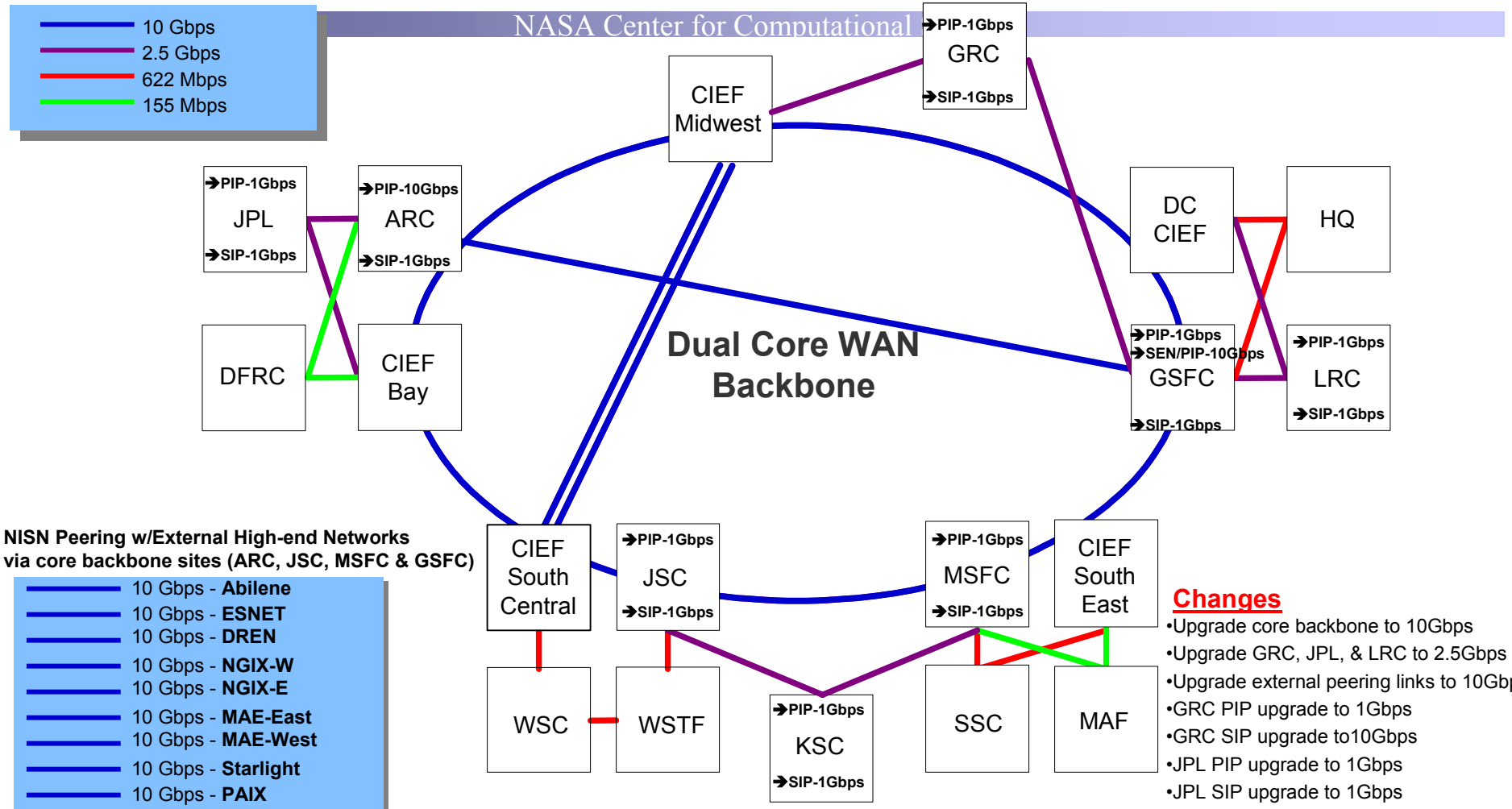
➔Center's local access to NISN WAN

Changes

- Add direct 10Gbps circuit between ARC (NAS) & GSFC (NCCS)
- GSFC PIP upgrade to 1Gbps
- GSFC/SEN PIP upgrade to10Gbps
- ARC PIP upgrade to 10Gbps
- ARC SIP upgrade to 1Gbps

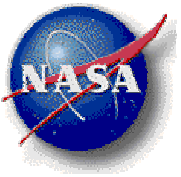
NISN-HEC Step 3 (24 Months)





Changes

- Upgrade core backbone to 10Gbps
- Upgrade GRC, JPL, & LRC to 2.5Gbps
- Upgrade external peering links to 10Gbps
- GRC PIP upgrade to 1Gbps
- GRC SIP upgrade to10Gbps
- JPL PIP upgrade to 1Gbps
- JPL SIP upgrade to 1Gbps
- PAIX peering upgraded to 10Gbps
- NGIX-East peering upgraded to 10Gbps
- NGIX-West peering upgraded to 10Gbps
- MAE-East & MAE-West peering terminated



NISN-HEC Focus & Planning



- HEC Program Office is dedicated to supporting current and future HEC WAN requirements
- Engaged in more detailed requirements gathering and analyses to determine if additional investments are needed
- Jerome Bennett is leading the GSFC migration effort for the HEC Program
- Question and concerns can be directed to:
  - Jerome.D.Bennett@NASA.gov 301-286-8543
  - Phil.Webster@NASA.gov 301-286-9535

Agenda

- Introduction — Phil Webster
- Systems Status—Mike Rouch
- NREN to NISN—Phil Webster
- New Data Sharing Services—Harper Pryor
- User Services—Sadie Duffy
- Questions or Comments

NCCS Support Services



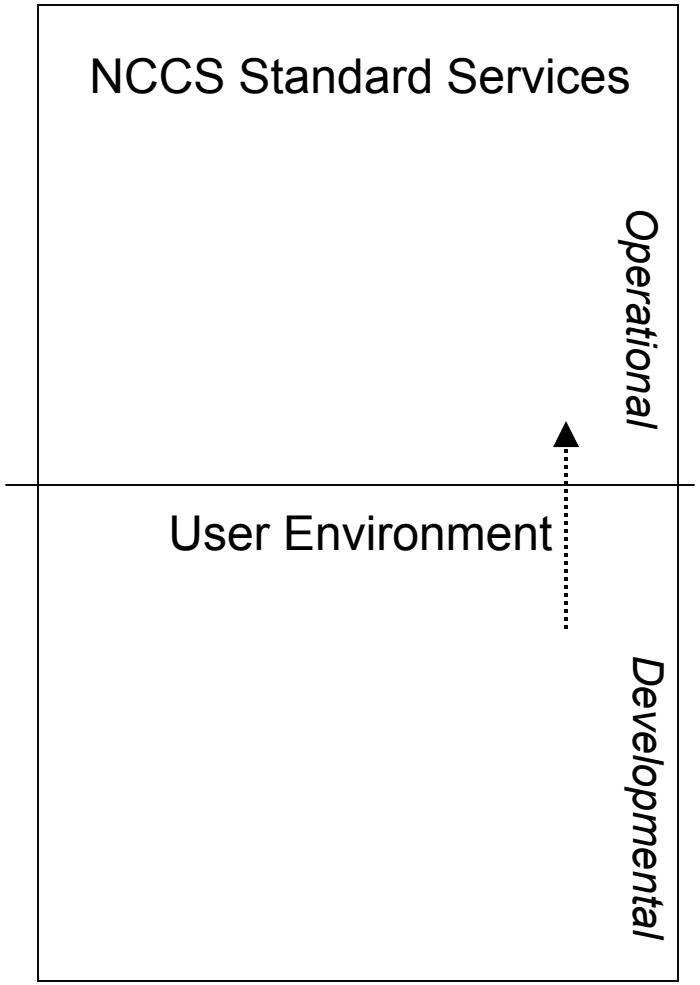
- Range of service offerings to support modeling and analysis activities of SMD users:
  - Production Computing
  - Data Archival & Stewardship
  - Code Development Environment
  - Analysis & Visualization
  - Data Sharing & Publication
- Data Sharing services
  - Share results with collaborators without requiring NCCS accounts
  - Capabilities include: web access to preliminary data sets with limited viewing and data download

Data Sharing Services

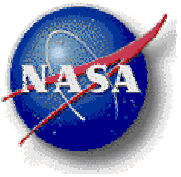
- General Characteristics
  - Data created by NCCS users
  - Support to active SMD projects
  - Not an on-line archive (will provide access to NCCS archived data)

- Approach
  - Develop capabilities for specific projects and generalize for public use
  - Development environment for project use
  - Resources managed by the NCCS
  - Software developed by SIVO and SMD users

Data Portal Service Model



- Projects may develop specific capabilities in a user environment.
- Used as an environment to assess customer needs.
- Promote to a standard service when production ready.
- Collaborators may access the user environment via unadvertised FTP & URL



State of the Data Portal



- History
  - Datastage
  - MAP06 data portal prototype
  - Data portal prototype – extended
- Current Platform
  - 8 blade Opteron
  - 32 TB GPFS managed storage
- Services
  - Web registration
  - Usage monitoring & reporting
  - Directory listings
  - Data download
  - Limited data viewing/display (GrADs, IDL)
- Projects under development
  - OSSE - MAP/ME
  - Cloud Library - Coupled Chemistry
  - MAP WMS - GMI



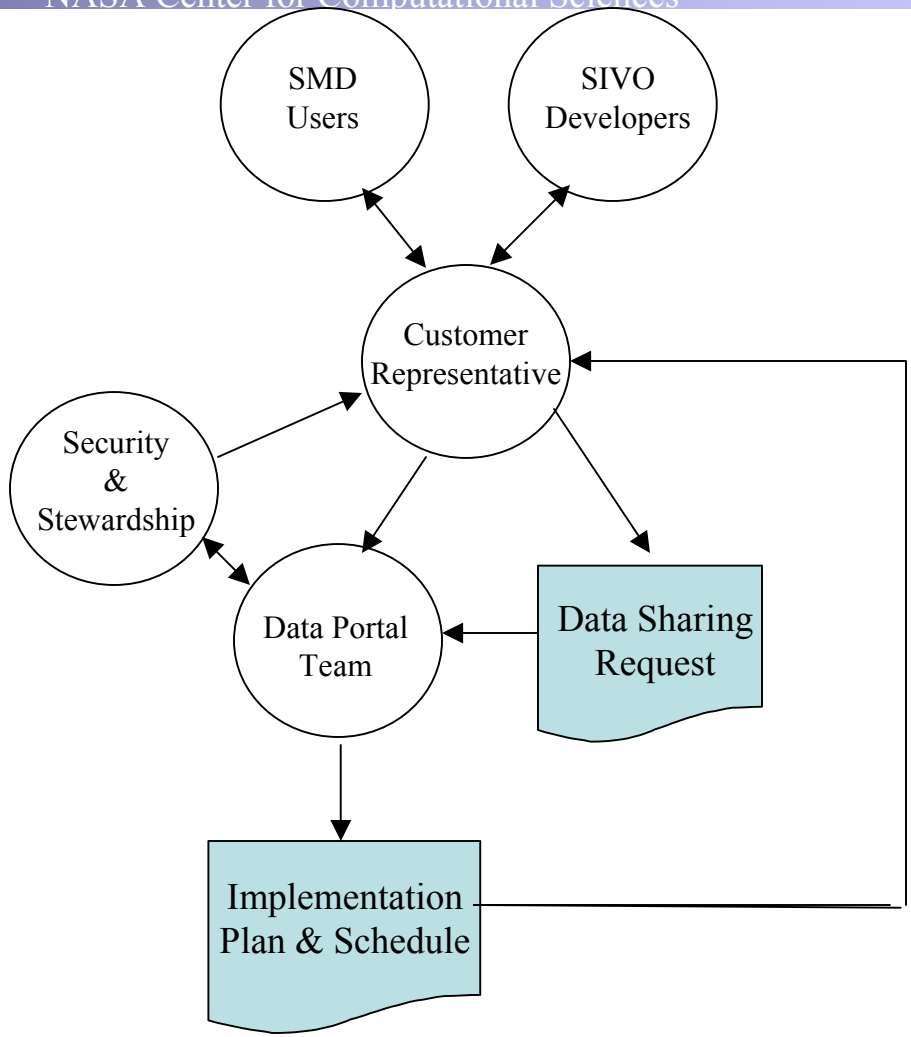
Data Sharing Service Request

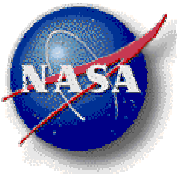
- **Project:** *SMD Project Name*
- **Sponsor:** *Sponsor Requesting Data Sharing Service*
- **Date:** *Date of Request*

- **Overview:** *Description of the specific SMD project producing data that are needed by collaborators outside of NCCS.*
- **Data:** *Information about data types, owners, and expected access methods to support data stewardship & protection planning.*
- **Access:** *Define collaborators eligible to access data.*
- **Resources:** *Estimate required data volumes and CPU resources.*
- **Duration:** *Define project lifecycle and associated NCCS support.*
- **Capability:** *Description of incremental service development.*
  Example:
  - » Web interface to display directory listings & download data
  - » Evaluate usage & data demands
  - » Add thumbnail displays to better identify data files
  - » Implement data subsetting capabilities to reduce download demands on remote users
  - » Reach back into NCCS archive for additional data holdings

Planning & Communication Paths







Discussion



- Let us know if you have a project that could benefit from data sharing services – so we can plan for it.
- Contact us if you want to explore opportunities.

- Your Point of Contact is:

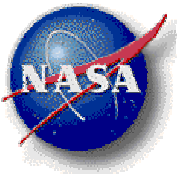
Harper.Pryor@gsfc.nasa.gov 301-286-9297

Agenda

- Introduction — Phil Webster
- Systems Status—Mike Rouch
- NREN to NISN—Phil Webster
- New Data Sharing Services—Harper Pryor
- User Services—Sadie Duffy
- Questions or Comments

Allocations

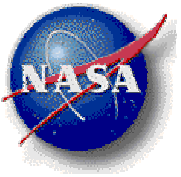
- FY07Q3—May 1st, 2007 allocation requests are with NASA headquarters for review
  - Expected award on, or shortly after, May 1st
- Next opportunity begins August 1st, 2007
- If you have a need between now and then, call the help desk



User Services Updates



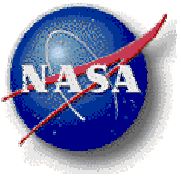
- Reminder of opportunities
  - User Telecon every Tuesday at 1:30pm 866-903-3877 participant code 6684167
  - USG staff available from 8am to 8pm to provide assistance
  - Online tutorials at http://nccs.nasa.gov/tutorials.html
  - Quarterly User Forum
- Feedback
  - Let us know if we can make these experiences more relevant (content, delivery method, venue, etc.)
  - Call at 301-286-9120 or email at support@nccs.nasa.gov



New Ticketing System



- NCCS uses a ticketing system to track issues reported to the help desk
  - Current system is very basic
  - Difficult to find old issues for reference
  - Users have no insight into their tickets
- New system is called Footprints by Numara
  - Provides NCCS staff with a much better tool for tracking and escalating user issues
  - Lots of extras to help us become more efficient



New Ticketing System



- How it affects you:
  - Provides the ability to open and view your personal tickets through an online interface
  - Escalation capability to ensure no issues are ever missed
  - On-line Peer to Peer chat capability with support staff
  - Quick access to broadcast alerts for system issues
  - Access to searchable knowledge base to help solve problems faster

- Questions?
- Comments?