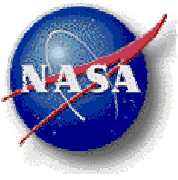


NASA Center for Computational Sciences

# *NCCS User Forum*

12 January 2007

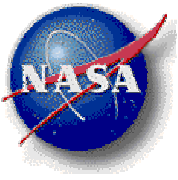


# Agenda



NASA Center for Computational Sciences

- Introduction—Phil Webster
- Systems Status—Mike Rouch
- Linux Networx Cluster update—Dan Duffy
- Preliminary Cluster Performance—Tom Clune
- New Services—Secure Unattended Proxy—Ellen Salmon
- User Services Updates—Sadie Duffy
- Questions or Comments

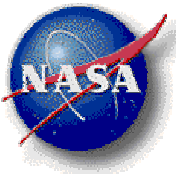


# Introducing the User Forum



NASA Center for Computational Sciences

- The NCCS User Forum is a quarterly meeting designed to facilitate *dialogue* with the NCCS users
- Topics will vary and may include:
  - Current NCCS services and systems
  - Suggestions for system utilization
  - Future services or systems
  - Questions and discussion with the user community
- Meeting will be available via remote access
- We are seeking your feedback
  - [Support@nccs.nasa.gov](mailto:Support@nccs.nasa.gov)
  - [Phil.Webster@NASA.gov](mailto:Phil.Webster@NASA.gov)

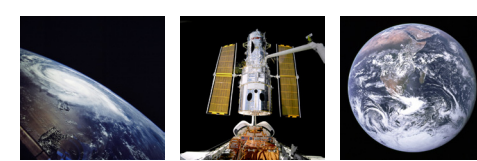
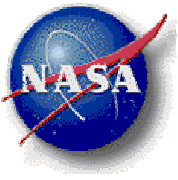


# Introduction



NASA Center for Computational Sciences

- Changes in HPC at NASA
  - [www.hpc.nasa.gov](http://www.hpc.nasa.gov)
- New HPC system
- Lots of work of to make the NCCS systems more robust
- Old systems will be retiring
- New tape storage
- Expanding data services



## NASA Center for Computational Sciences



NATIONAL AERONAUTICS  
AND SPACE ADMINISTRATION

[+ NASA Home](#)

FIND IT @ NASA:

Keywords

+ GO



# High-End Computing Program

## HEC Program

### APPLYING FOR RESOURCES

NASA's four Mission Directorates, the NASA Engineering and Safety Center, and the National Leadership Computing System (NLCS) initiative each have procedures for awarding time on our high-end computers. All but NLCS require prior NASA funding.

- [Aeronautics Research Mission Directorate](#)
- [Exploration Systems Mission Directorate](#)
- [Science Mission](#)

### ABOUT OUR PROGRAM

NASA's High-End Computing (HEC) Program is committed to maintaining a stable, service-oriented computing environment for the Agency's Mission Directorates, the NASA Engineering and Safety Center, external collaborators, and the nation. Program funding flows through the Shared Capability Assets Program and the Science Mission Directorate. Combined, HEC Program resources at Ames Research Center and Goddard Space Flight Center provide researchers with over 13,000 processors and 75 teraflops (peak) of computing power.

[+ Read more](#)

### NEW & NOTABLE



**AISR Call for Proposals Expanded**  
The Applied Information Systems Research (AISR) program is seeking proposals to "deploy enhancements to persistent software tools and environments that assist the productivity of the scientific users of high-end computing resources." Proposals are due January 30.

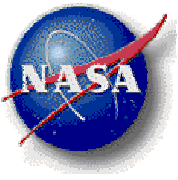
### PROGRAM FACILITIES



**NASA Advanced Supercomputing Division**  
Ames Research Center,  
Moffett Field, California  
[+ Visit this website](#)



**NASA Center for Computational Sciences**  
Goddard Space Flight Center,  
Greenbelt, Maryland

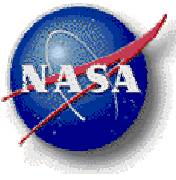


# Halem



NASA Center for Computational Sciences

- *Halem* will be retired 1 April 2007
  - Four years of service
  - Un-maintained for over 1 year
- Replaced by *Discover*
  - Factor of 3 capacity increase
  - Migration activities underway
  - “*We need the cooling*”

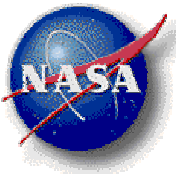


# HPC in NASA



NASA Center for Computational Sciences

- HPC Program Office has been created
- Managed for NASA by SMD
  - Tsengdar Lee - Program Manager
  - Located in the Science Division
- Two Components
  - NCCS - Funded by SMD for SMD users
  - NAS - Funded by SCAP (Shared Capability Asset Program) for all NASA users
  - Managed as a coherent program
    - One NASA HPC Environment
    - Single Allocation request



# OneNASA HPC Activities



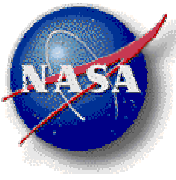
NASA Center for Computational Sciences

- OneNASA HPC Initiatives

Streamline processes and improve interoperability between NCCS and NAS

- Create a common *NASA HPC Environment*
- Account/Project initiation
- Common UID/GID
- File Transfer improvements
- More flexible job execute opportunities



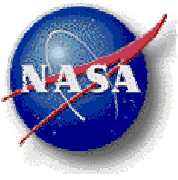


# Agenda



NASA Center for Computational Sciences

- Introduction—Phil Webster
- Systems Status—Mike Rouch
- Linux Networx Cluster update—Dan Duffy
- Preliminary Cluster Performance—Tom Clune
- New Services—Secure Unattended Proxy—Ellen Salmon
- User Services Updates—Sadie Duffy
- Questions or Comments

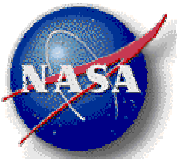


# System Status



NASA Center for Computational Sciences

- Statistics
  - Utilization
  - Usage
  - Service Interruptions
- Known Problems



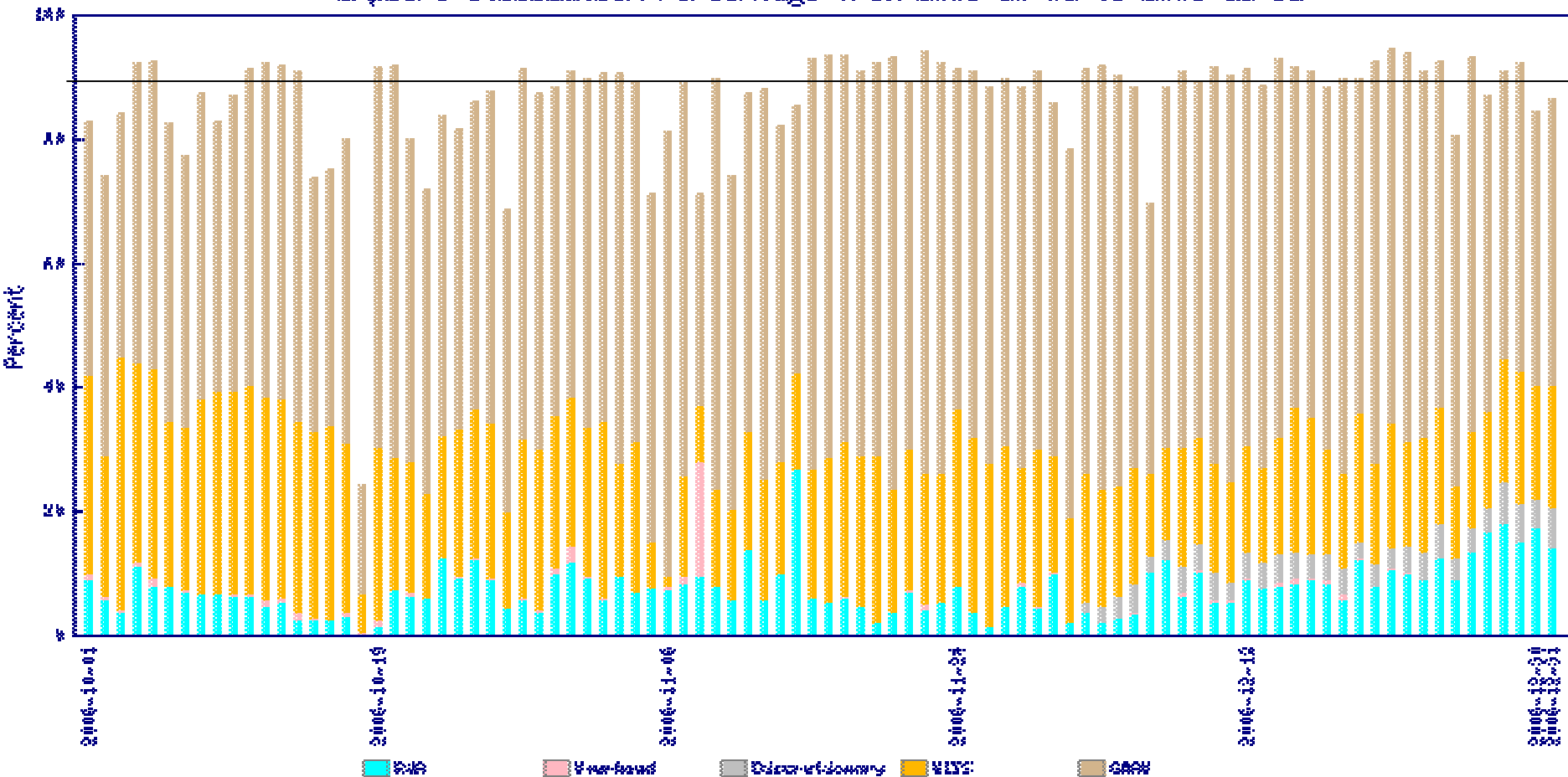
# Explore Utilization

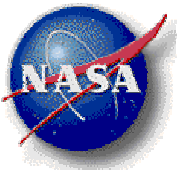
## Oct–Dec 2006



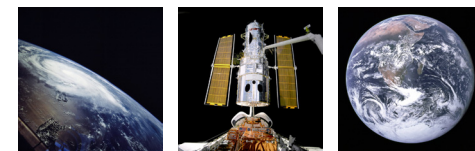
NASA Center for Computational Sciences

Explore Utilization Percentage from 2006-10-01 to 2006-12-31



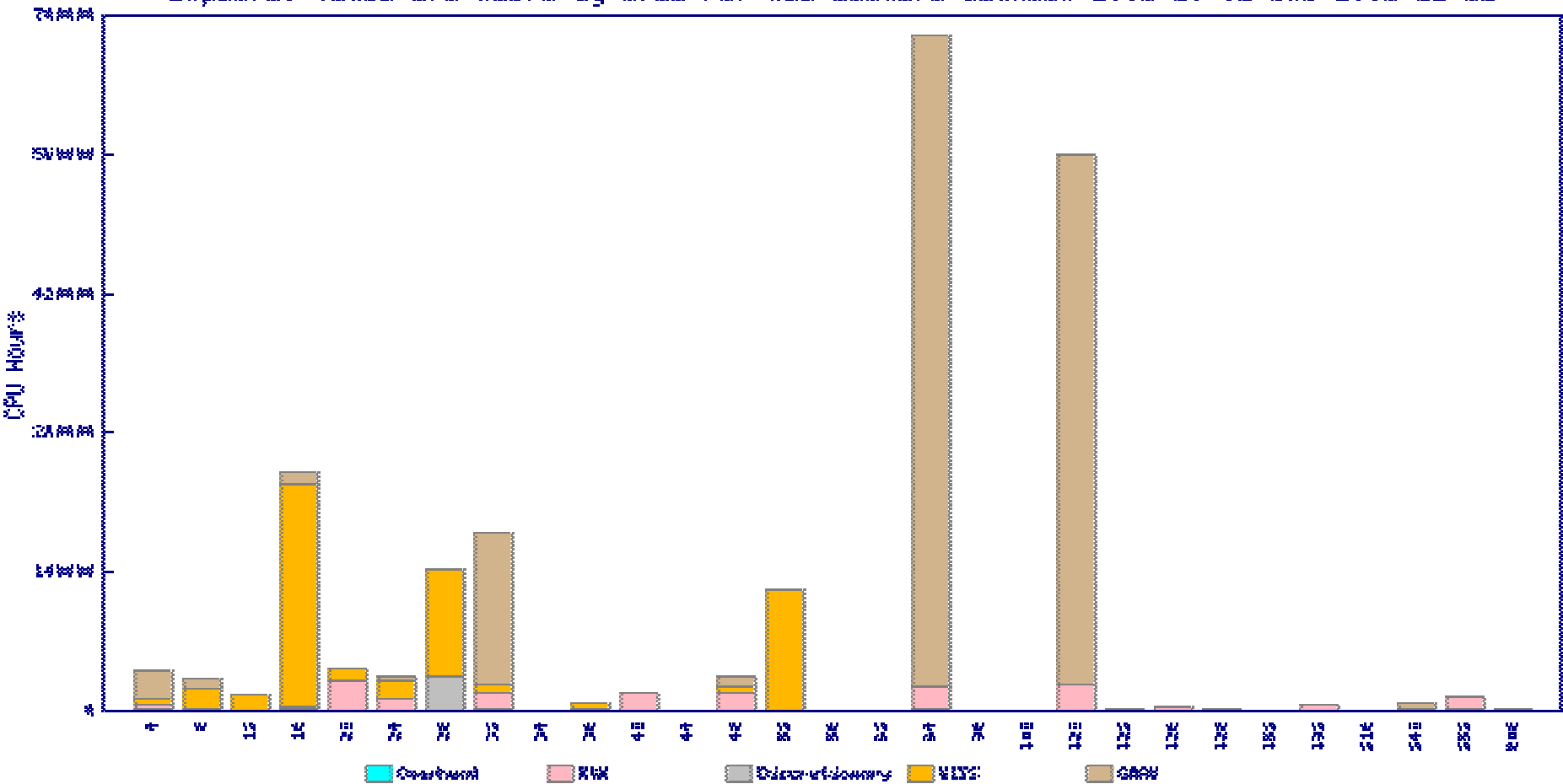


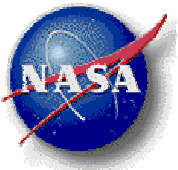
# Explore CPU Usage Oct–Dec 2006



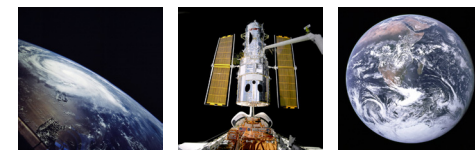
NASA Center for Computational Sciences

Explore: Total CPU Hours by CPUs for All Sectors between 2005-10-01 and 2005-12-31



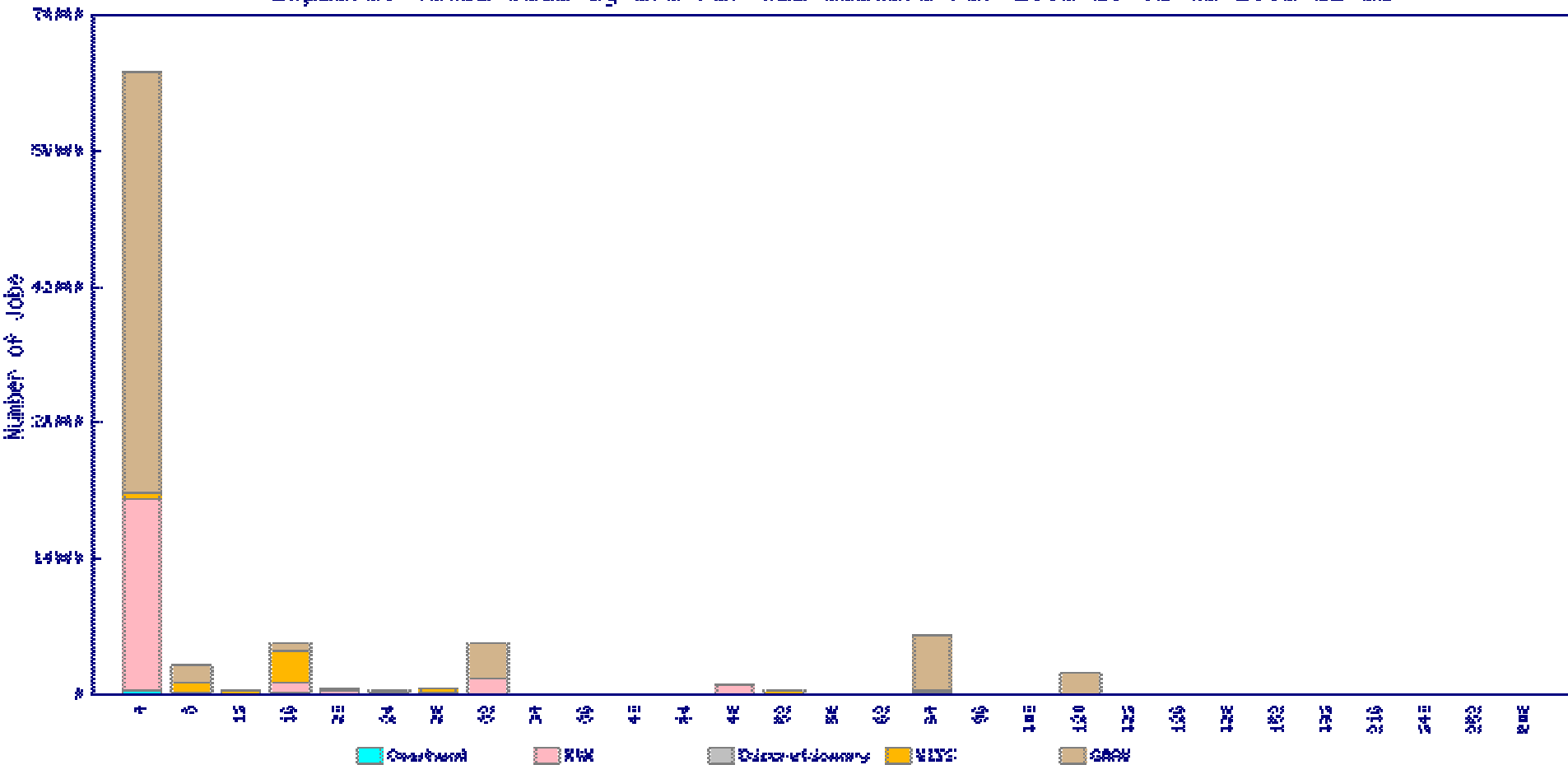


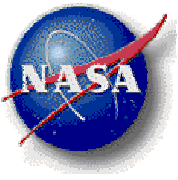
# Explore Job Mix Oct-Dec 2006



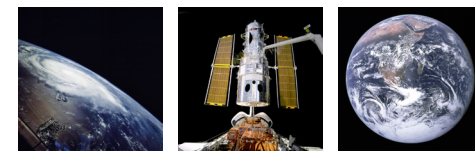
NASA Center for Computational Sciences

Explore: Total Jobs by CPU For All Sectors For 2006-10-01 to 2006-12-31



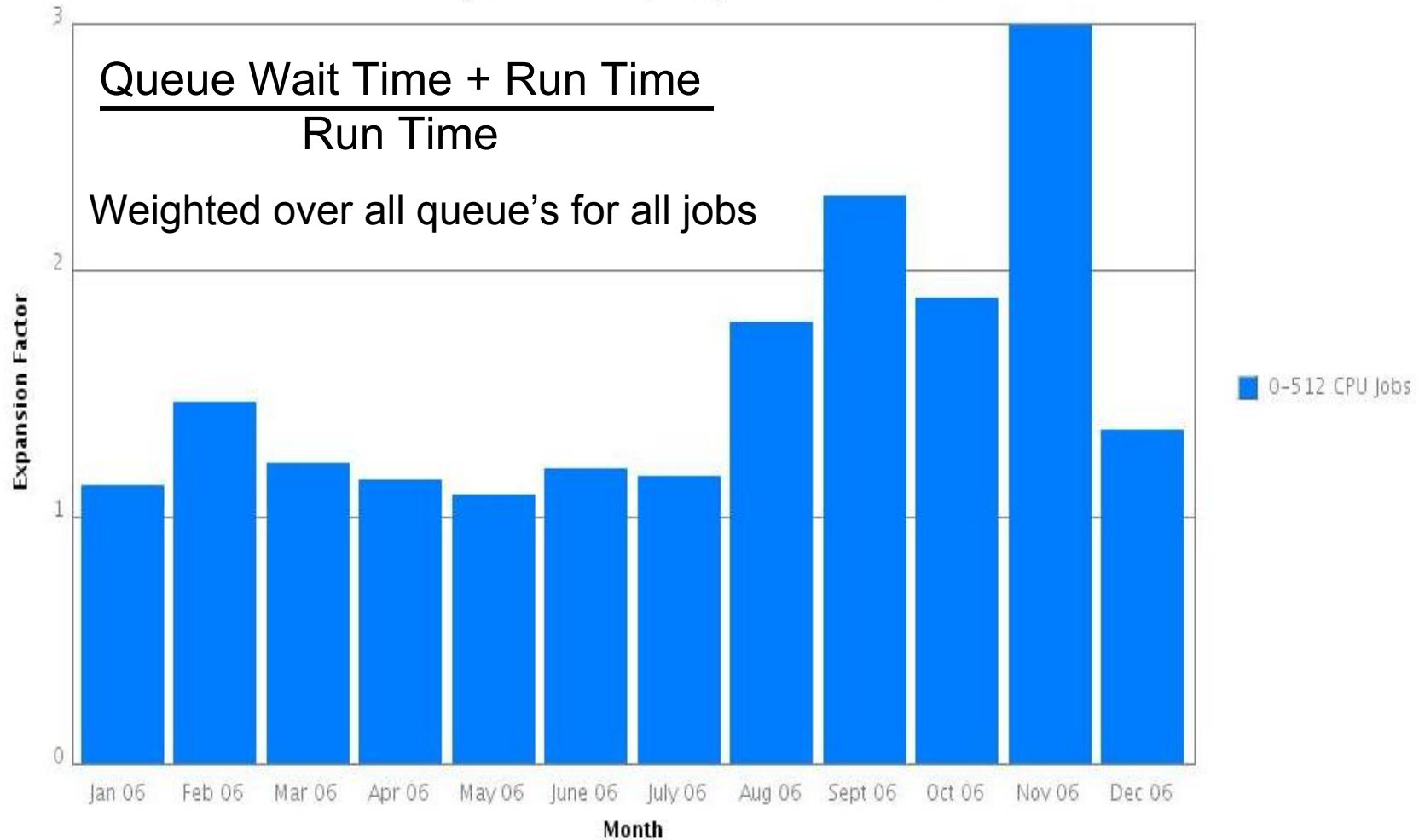


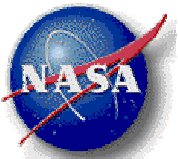
# Explore Queue Expansion Factor



NASA Center for Computational Sciences

Expansion Factors January 2006 - December 2006





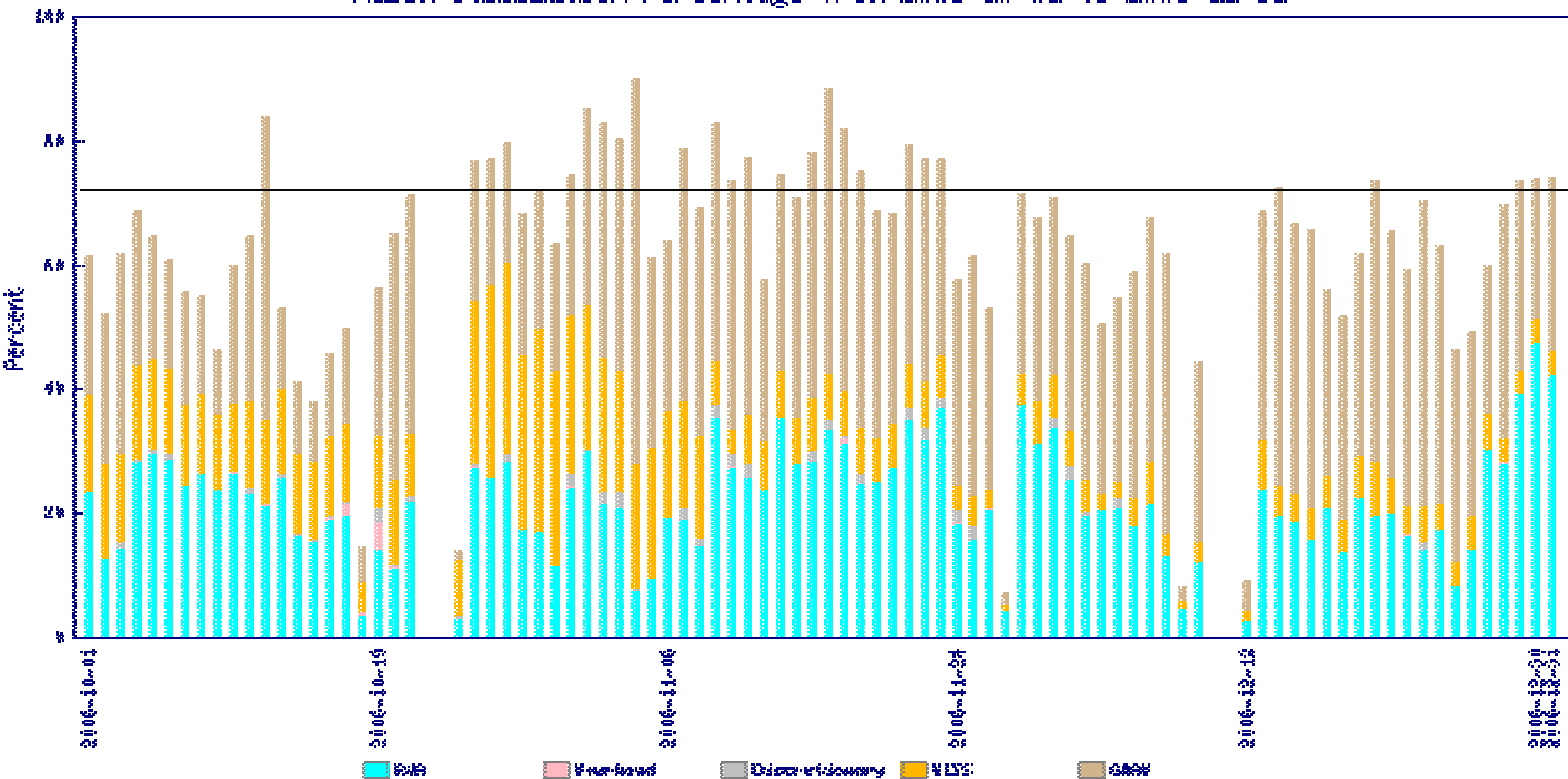
# Halem Utilization

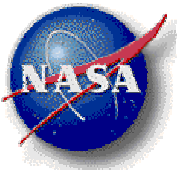
## Oct–Dec 2006



NASA Center for Computational Sciences

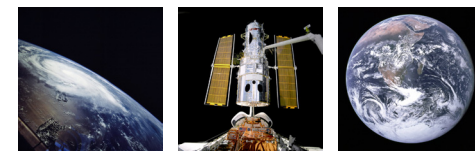
halem Utilization Percentage from 2006-10-01 to 2006-12-31





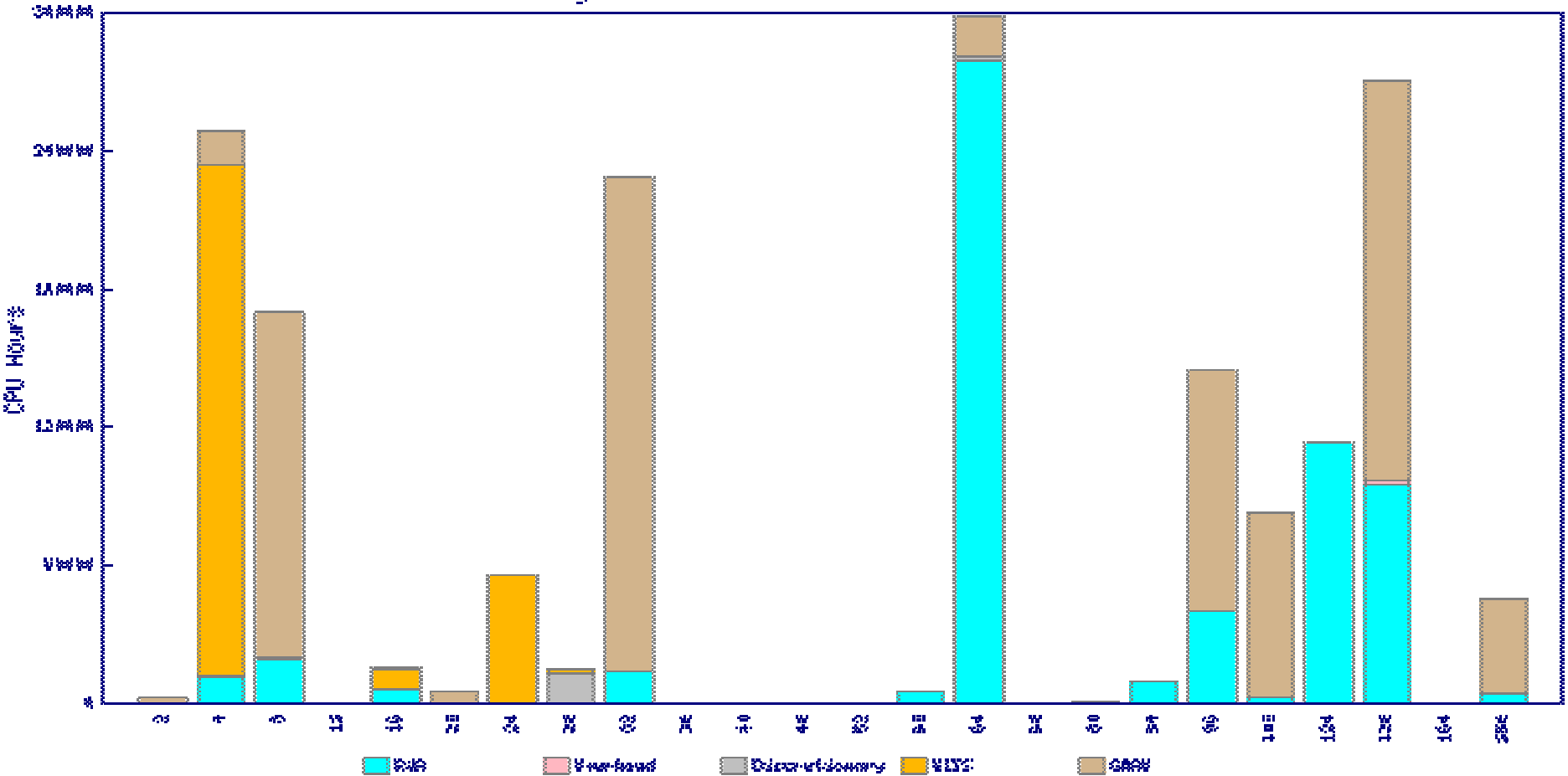
# Halem CPU Usage

## Oct–Dec 2006

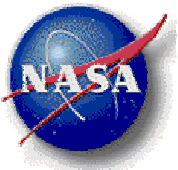


NASA Center for Computational Sciences

hallex: Total CPU Hours by CPUs for All Sectors between 2005-10-01 and 2005-12-31

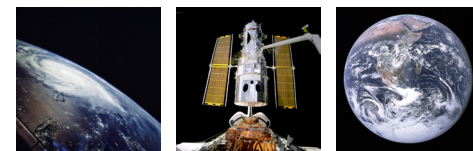






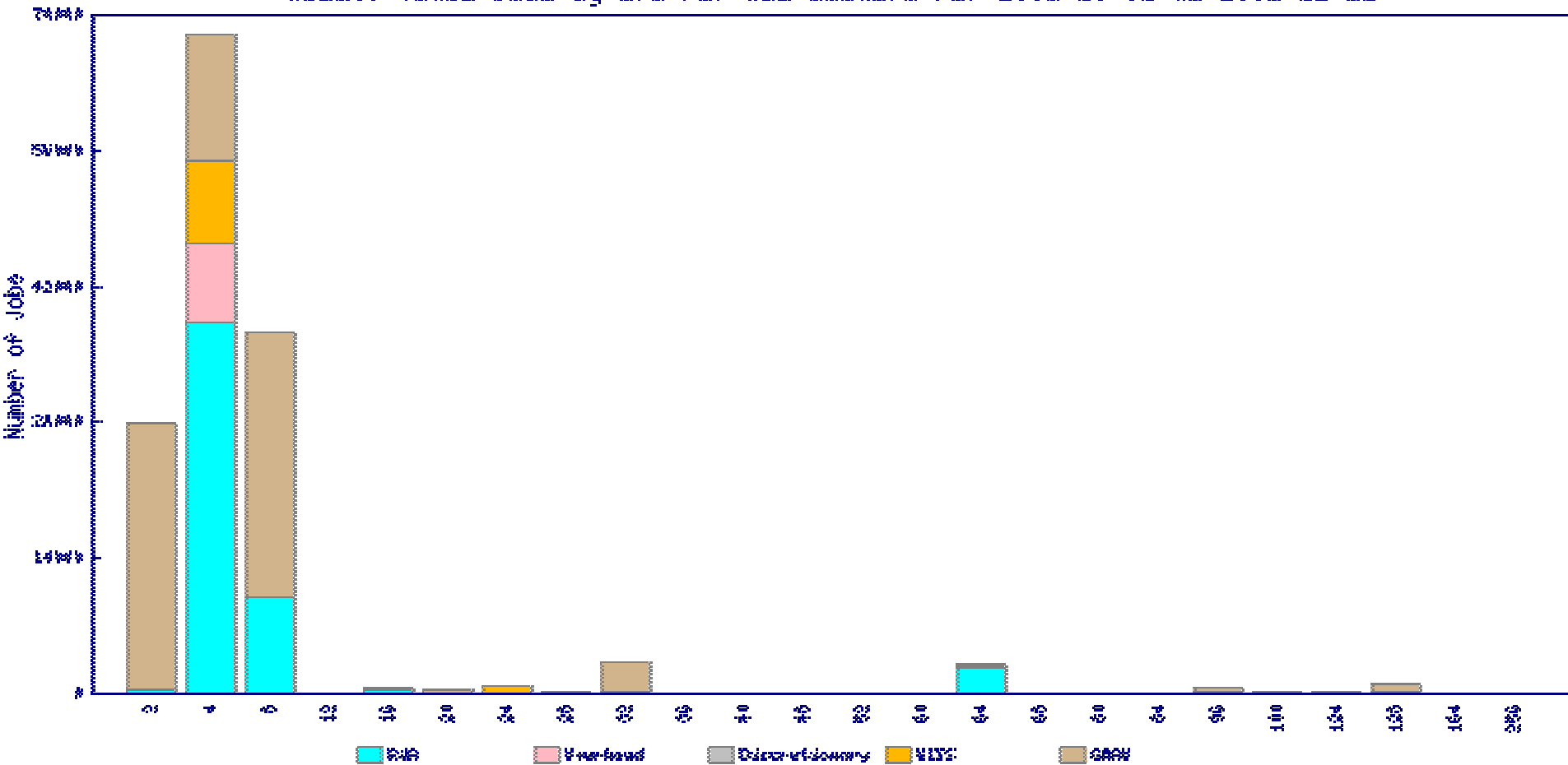
# Halem Job Mix

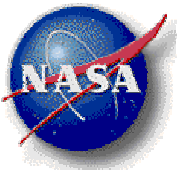
## Oct–Dec 2006



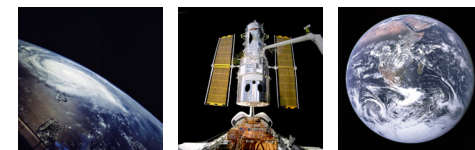
NASA Center for Computational Sciences

halem: Total Jobs by CPU For All Sectors For 2006-10-01 to 2006-12-31





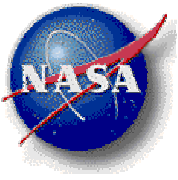
# DMF Mass Storage Growth



NASA Center for Computational Sciences

Date	# Files	PB	# of Files Growth	Growth TB
Jun	49,856,083	1.34		
July	51,656,396	1.40	1,800,313	66.84
Aug	53,730,549	1.46	2,074,153	59.78
Sep	55,989,279	1.54	2,258,730	83.78
<b>Oct</b>	<b>58,990,934</b>	<b>1.64</b>	<b>3,001,655</b>	<b>102.50</b>
<b>Nov</b>	<b>61,503,059</b>	<b>1.74</b>	<b>2,512,125</b>	<b>106.25</b>
<b>Dec</b>	<b>64,168,613</b>	<b>1.84</b>	<b>2,665,554</b>	<b>93.05</b>

***Adequate Capacity for Continued Steady Growth***



# DMF Mass Storage File Distribution



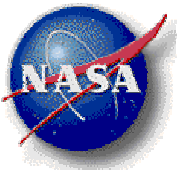
NASA Center for Computational Sciences

Total Number of Files	64.5 Million	
$\leq 1\text{mb}$	28.6 Million	44%
$1 < \leq 10\text{mb}$	18.5 Million	29%

**73% of all files are  $\leq 10\text{mb}$**

*Please tar up files when appropriate*

- Improves file recall performance
- Reduce total number of files and improves DMF performance
- May increase burden to cleanup of files in /nobackup

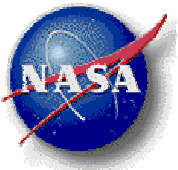


# SGI Explore Incidents



NASA Center for Computational Sciences

	<b>Oct</b>	<b>Nov</b>	<b>Dec</b>
Hardware	2		1
Power	2		2
Application	1	4	
O/S Bugs	3		3
FS CxFS		3	1
FS Hardware			
FS lscr			
NFS			
Network			1
Admin			1
<b>Total</b>	<b>8</b>	<b>7</b>	<b>9</b>

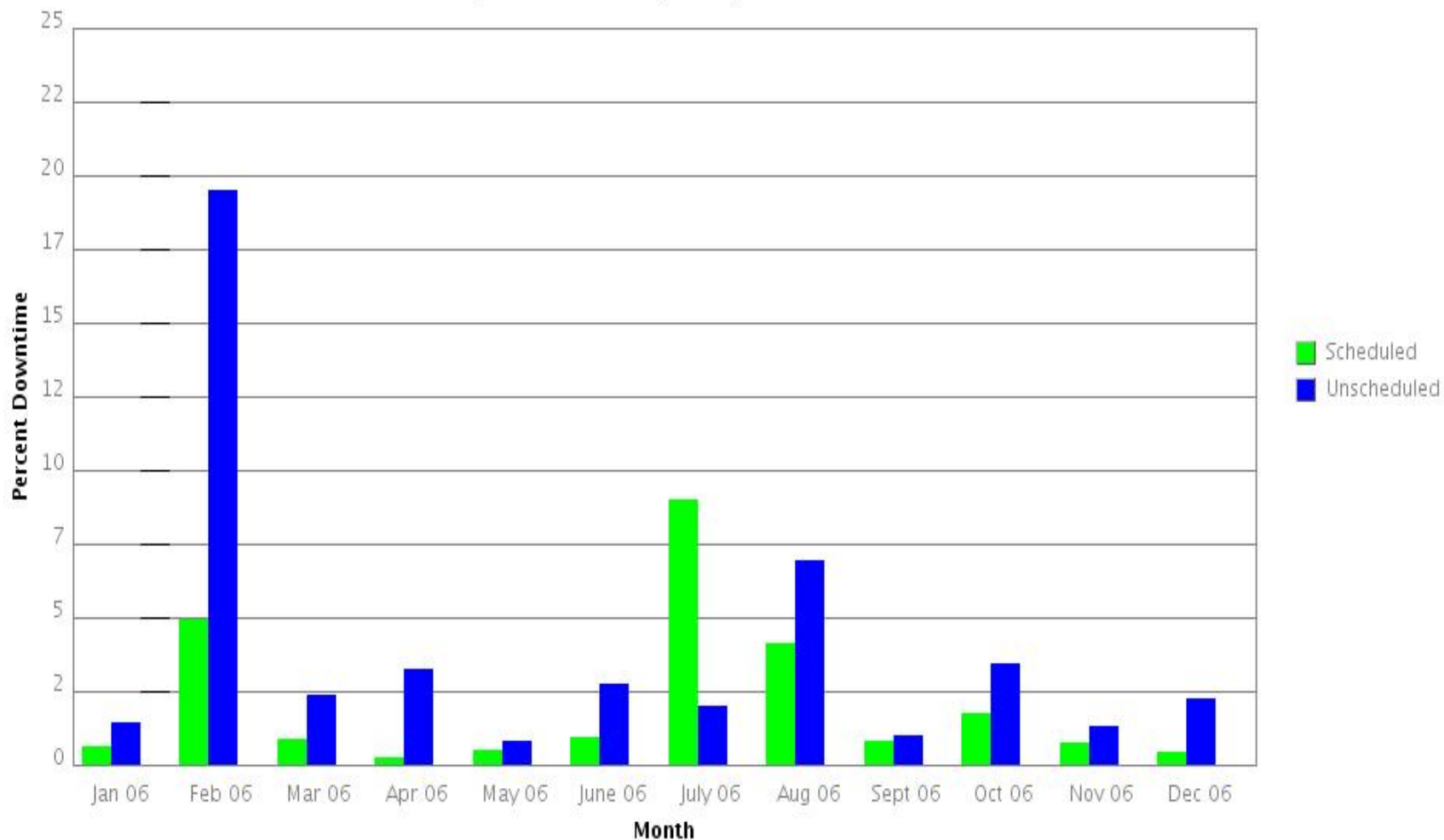


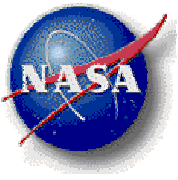
# SGI Explore Downtime



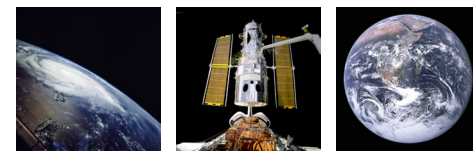
NASA Center for Computational Sciences

Explore Downtime January 2006 - December 2006



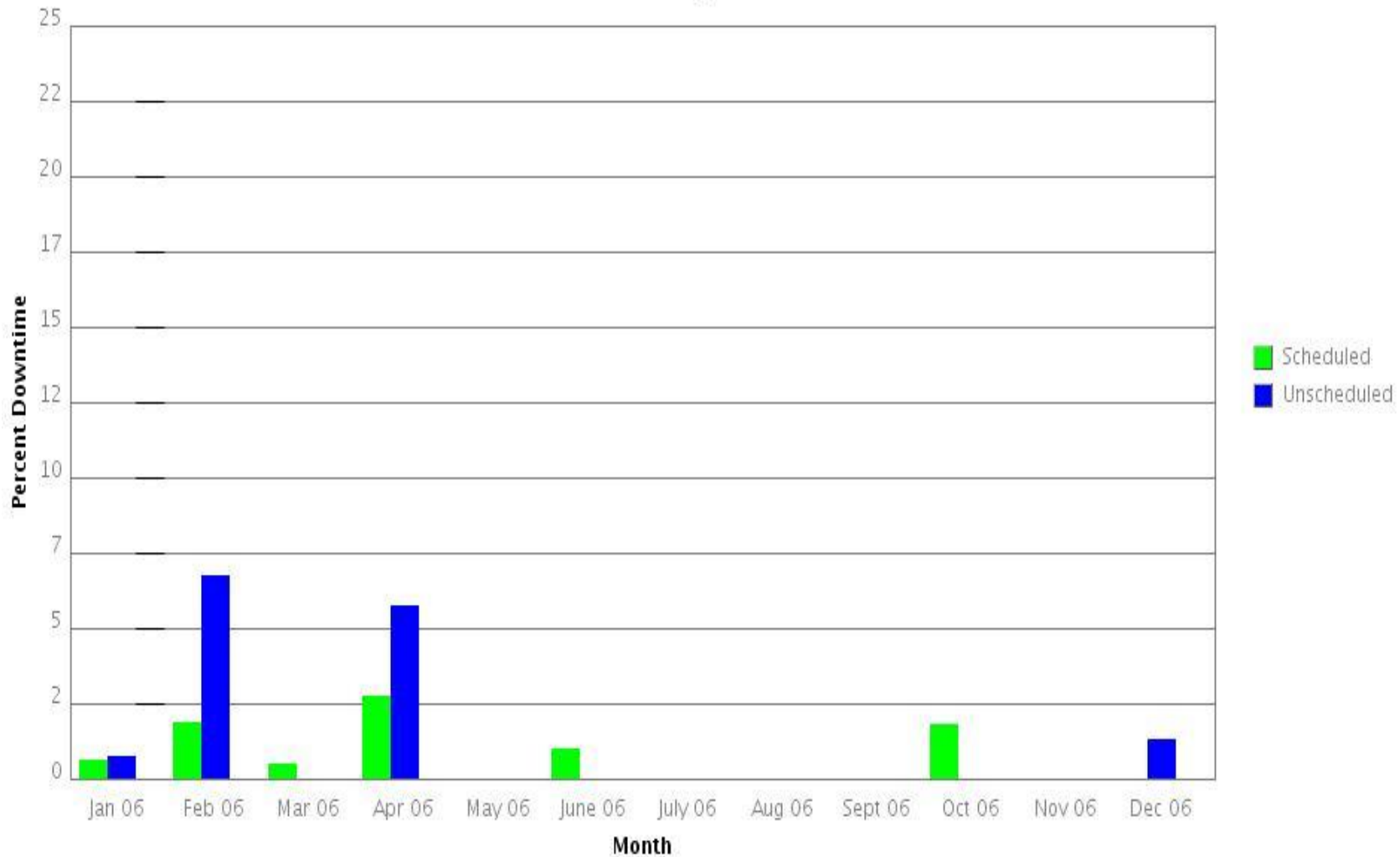


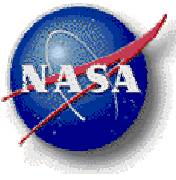
# Halem Downtime



NASA Center for Computational Sciences

Halem Downtime January 2006 - December 2006





# Issues



NASA Center for Computational Sciences

## Eliminate Data Corruption – SGI Systems

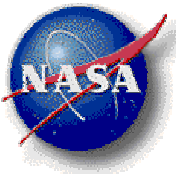
**Issue:** Files being written at the time of an SGI system crash **MAY** be corrupted. However, files appear to be normal.

### Interim Steps: Careful Monitoring

- Install UPS
- 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



# Improvements



NASA Center for Computational Sciences

## Reduced Impact of Power Outages Q1 2007

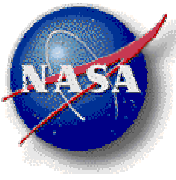
**Issue:** Power fluctuations during thunderstorms

**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 - Q1 2007





# Improvements



NASA Center for Computational Sciences

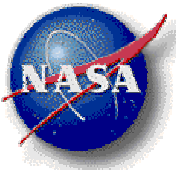
## Enhanced NoBackup Performance on Explore Q1 2007

**Issue:** NoBackup Shared file system poor I/O performance

**Effect:** Slow job performance

**Fix:** From the Acquired additional disks discussed last quarter

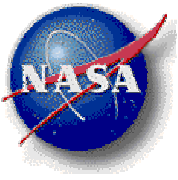
- Creating More NoBackup File Systems
- Spread out the load across more file systems
- Upgraded System I/O hba's – 4GB
- Implementing New FC Switch – 4GB



# Improvements

NASA Center for Computational Sciences

- **Improving File Data Access – Q1 2007**
  - Increase DMF disk cache to 100 TB
  - Increase NoBackup file system sizes for Users and Projects, as required
- **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
- **Enhancing Systems – Q1 2007**
  - Software OS & CxFS upgrades to Irix (Feb 07)
  - Software OS & CxFS upgrades to Altix (Feb 07)

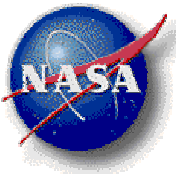


# Explore Upgrade



NASA Center for Computational Sciences

- **Upgrading Explore Operating System**
  - Suse10 & PP5 – Q1 2007
  - Ames Testing - successful
  - NCCS Testing Jan 07
  - Ames Upgrade Schedule (tentative)
    - 2048 - Feb 2007
    - Rest of the Altix - Mar 2007
  - NCCS Upgrade Schedule
    - March 2007

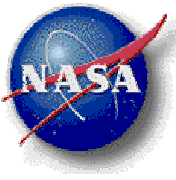


# Agenda



NASA Center for Computational Sciences

- Introduction—Phil Webster
- Systems Status—Mike Rouch
- Linux Networx Cluster update—Dan Duffy
- Preliminary Cluster Performance—Tom Clune
- New Services—Secure Unattended Proxy—Ellen Salmon
- User Services Updates—Sadie Duffy
- Questions or Comments

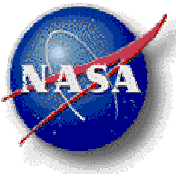


# Current Discover Status



NASA Center for Computational Sciences

- Base unit accepted – What does that mean?
  - Ready for production
    - Well, sort of... ready for general availability
  - User environment evolving
    - Tools: NAG, absoft, IDL, TotalView
    - Libraries: different MPI versions
    - Other software: sms, tau, papi, etc
    - If you need something, please open up a ticket with User Services
- PBS queues up and running waiting for jobs!

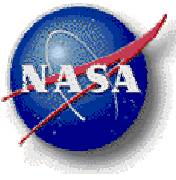


# Quick PBS Lesson



NASA Center for Computational Sciences

- How do I specify resources on the Discover cluster using PBS?
  - Recall there are 4 cores/node
  - You probably only want to run at most 4 processes/node
- Example 1: Run 4 processes on a single node
  - #PBS -l select=1:ncpus=4
  - mpirun -np 4 ./myexecutable
- Example 2: Run a 16 process job across multiple nodes
  - # PBS -l select=4:ncpus=4
  - mpirun -np 16 ./myexecutable
- Example 3: Run 2 processes per node across multiple nodes
  - # PBS -l select=4:ncpus=2
  - mpirun -np 8 ./myexecutable

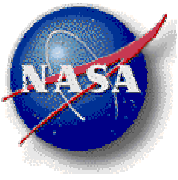


# Current Issues



NASA Center for Computational Sciences

- Memory leak
  - Symptom: Total memory available to user processes slowly decreases
  - Outcome: The same job will eventually run out of memory and fail
  - Progress:
    - Believe it is in Scali MPI version
    - Have fixed a problem in the Infiniband stack
    - Looking for different MPI options, namely Intel MPI
    - Will provide some performance and capability comparison when the MPI versions become available
- 10 GbE problem
  - Symptom: 10 GbE interfaces on gateway nodes are not working
  - Outcome: Intermittent access to cluster and Altix systems
  - Progress:
    - Currently using 1 GbE interfaces
    - Vendors actively working on the problem



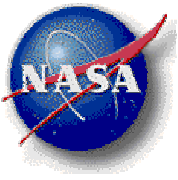
# What Next? (Within a couple of months)



NASA Center for Computational Sciences

- Addition of the first scalable unit
  - 256 compute nodes
    - Dual socket, dual core
    - Intel Woodcrest 2.67 GHz
    - Initial results are showing a very good speedup, anywhere between 15% to 50% depending on the application – no recompilation is required
  - Will require about 12 hours of outage to merge the base unit and scalable unit
- Additional disk space available under GPFS
  - To be installed at the time of the scalable unit merger
- Addition of viz nodes
  - Opteron based with vis tools, like IDL
  - Will see all the same GPFS file systems
- Addition of test system



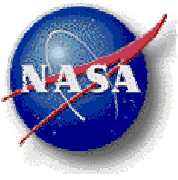


# Agenda



NASA Center for Computational Sciences

- Introduction—Phil Webster
- Systems Status—Mike Rouch
- Linux Networx Cluster update—Dan Duffy
- ➔ Preliminary Cluster Performance—Tom Clune
- New Services—Secure Unattended Proxy—Ellen Salmon
- User Services Updates—Sadie Duffy
- Questions or Comments

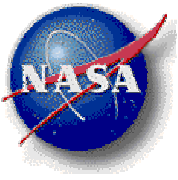


# Tested Applications



NASA Center for Computational Sciences

- modeLE - Climate simulation. (GISS)
- Cubed-Sphere FV dynamical core. (SIVO)
- GEOS - Weather/climate simulation.(GMAO)
- GMI - Chemical transport model (Atmospheric Chemistry and Dynamics Branch)
- GTRAJ - Particle trajectory code (Atmospheric Chemistry and DynamicsBranch)

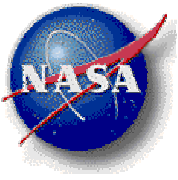


# Performance Factors



NASA Center for Computational Sciences

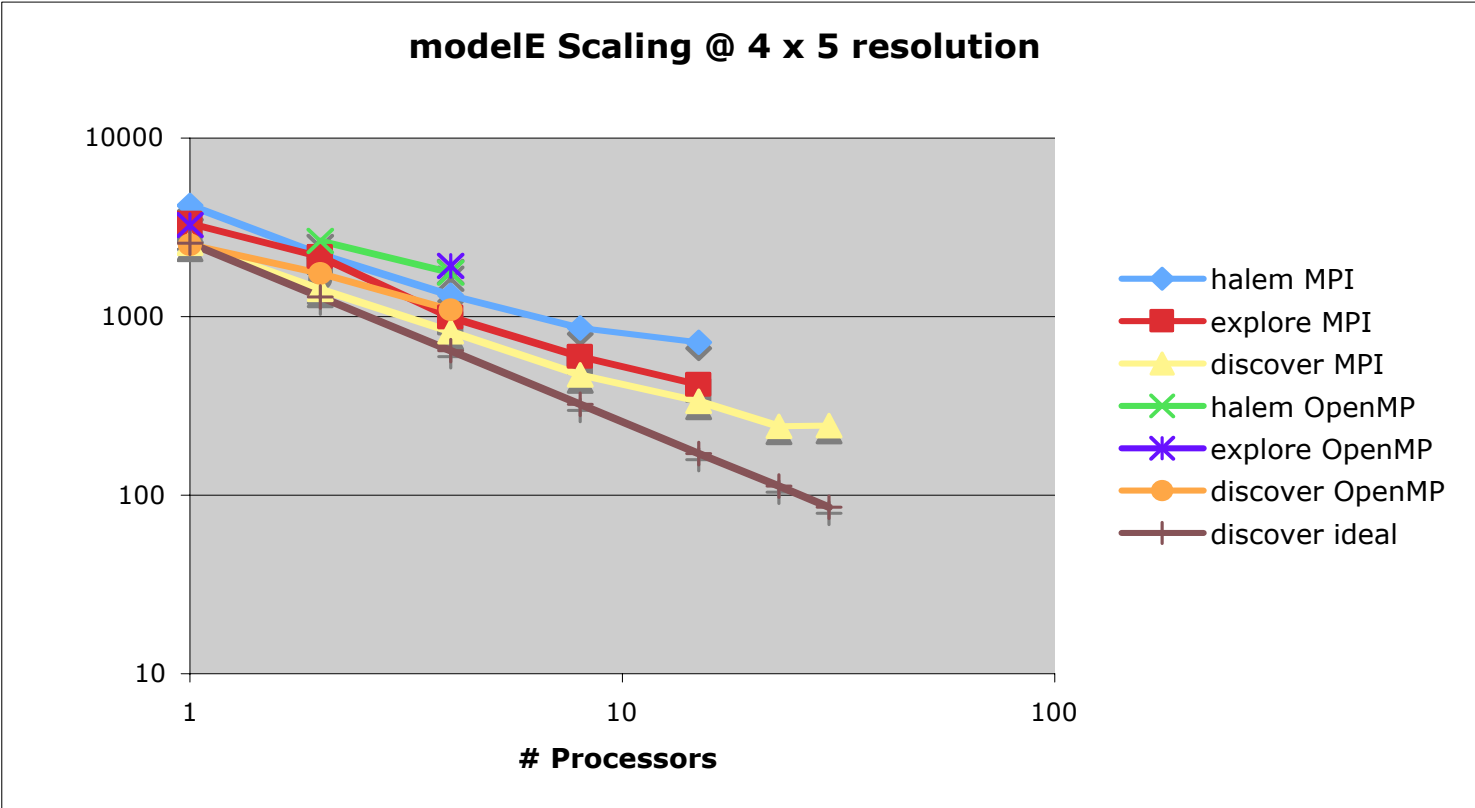
	Halem	Explore/ Columbia	Discover
Clock	1.25 Ghz	1.5-1.6 Ghz	3.2 Ghz
Peak	2.5 GF	6.0-6.4 GF	6.4 GF
L2/L3 cache	16 MB	6-9 MB	2 MB
Memory BW per core	1.3 GB/s	6.4 GB/s	4 GB/s
Network Bandwidth	200 MB/s	3.2 GB/s	2.5 GB/s

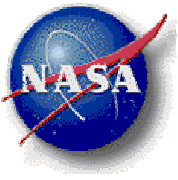


# ModelE

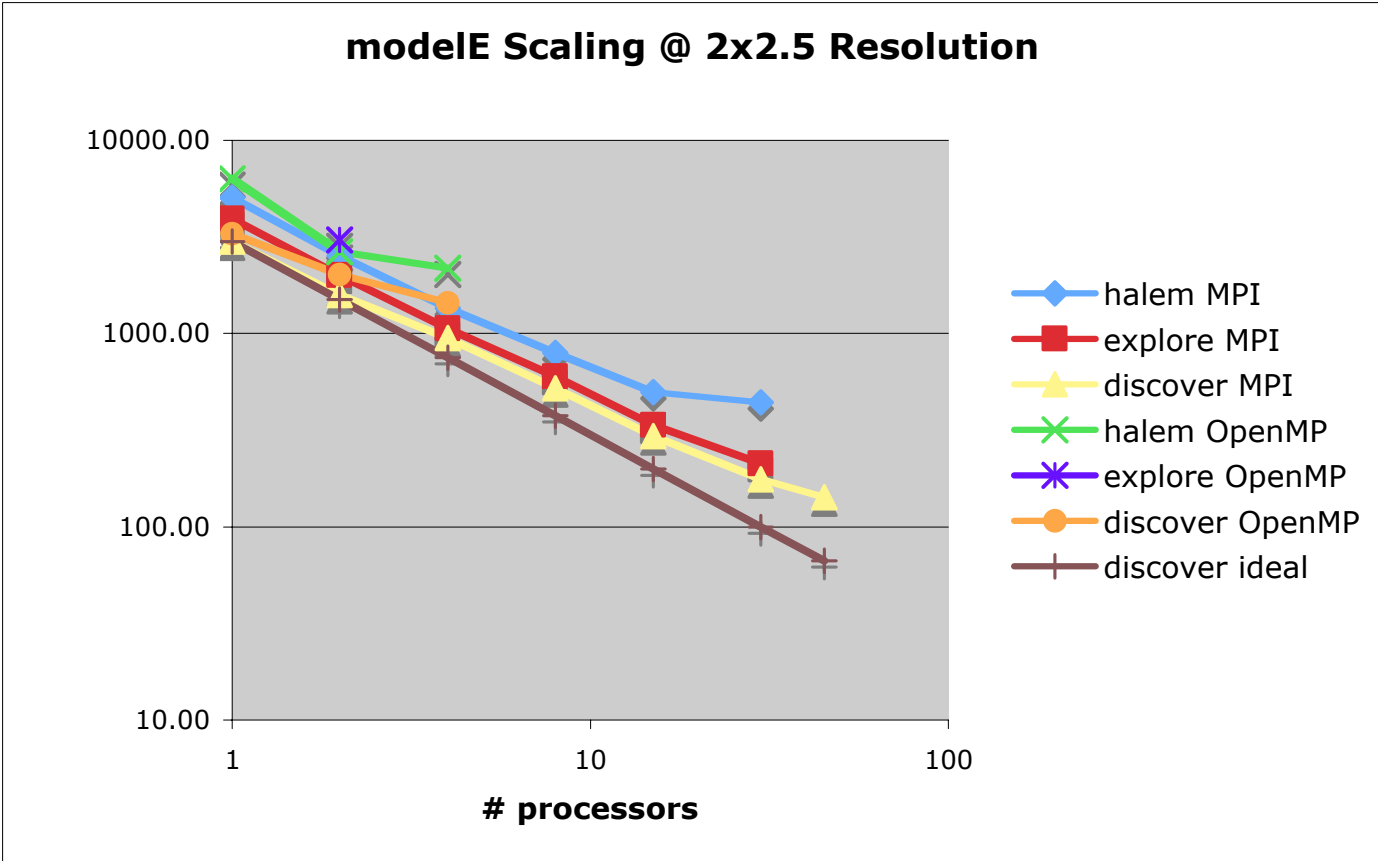
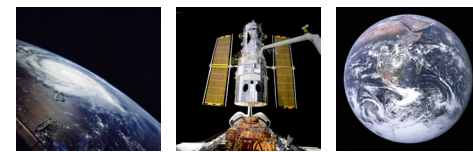


NASA Center for Computational Sciences



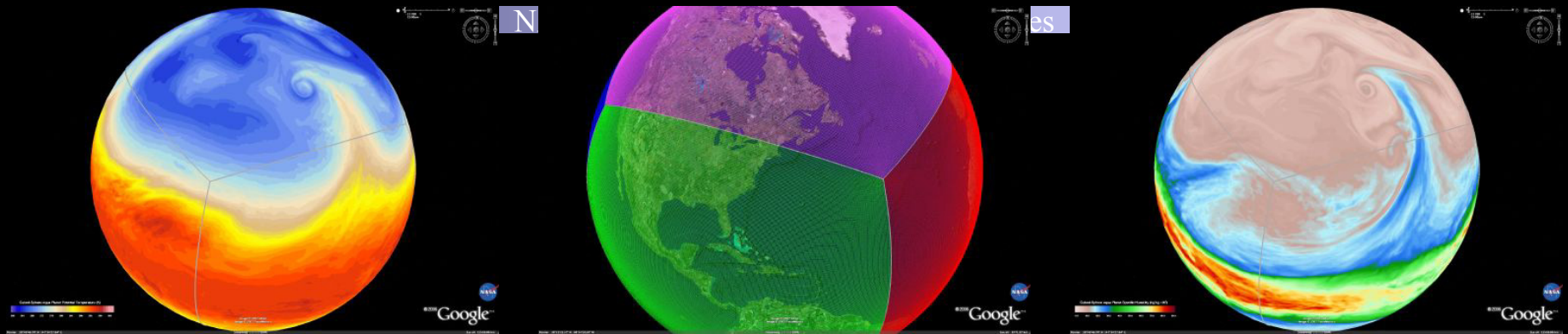


# ModelE



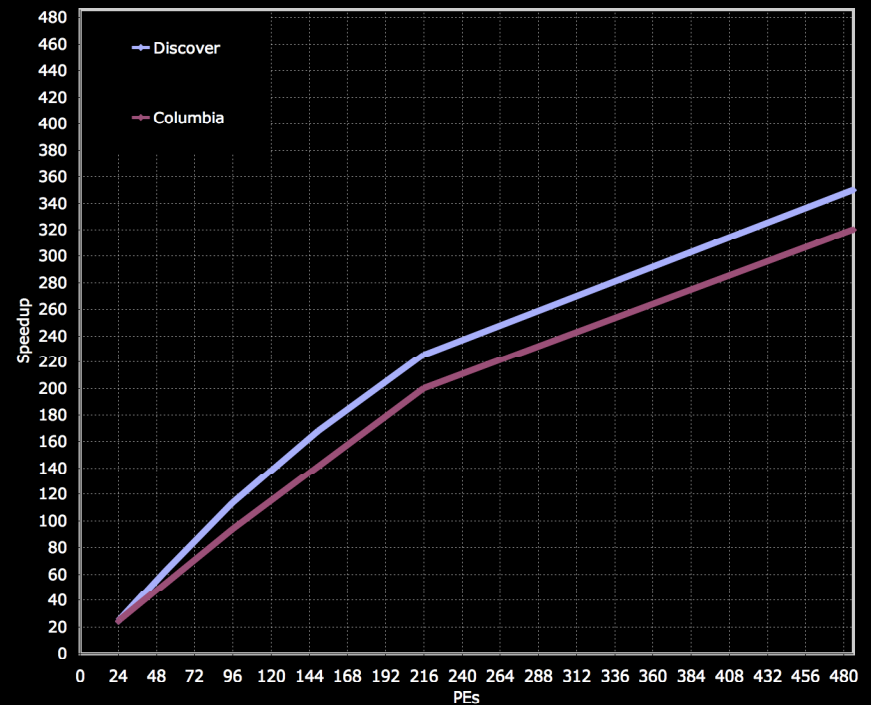
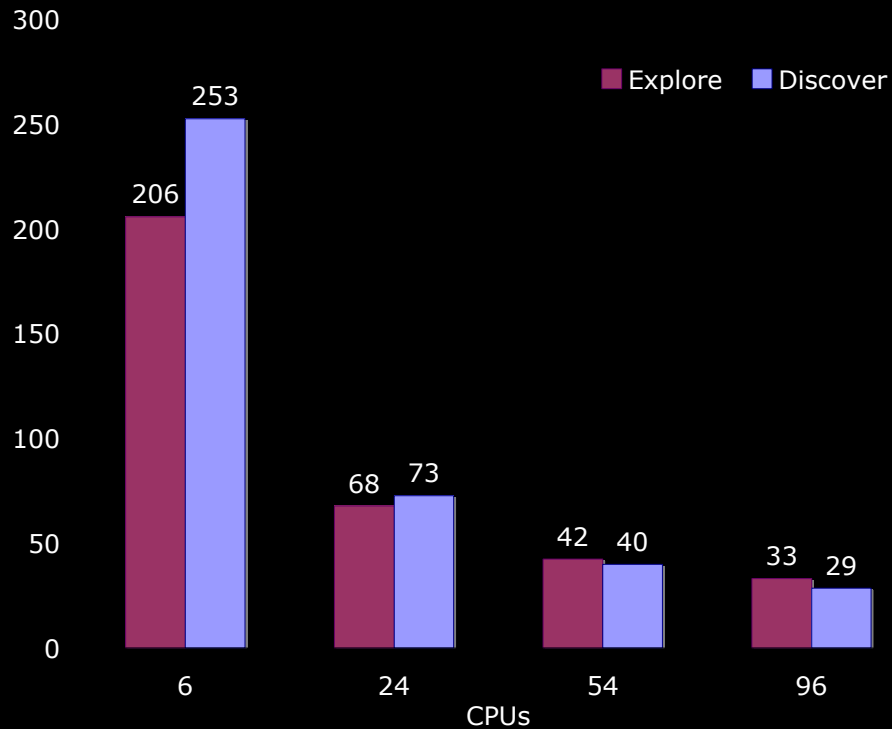


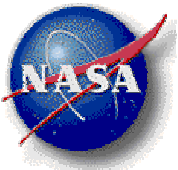
# The Cubed-Sphere FV CORE on Discover



c44 (2-deg) 32-Level Cubed-Sphere FV CORE Benchmarks

Cubed-Sphere Finite-Volume Dynamical Core Scaling c180 (0.5-deg) 32-Levels



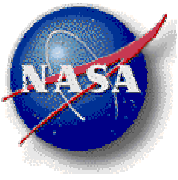


# Other Cases



NASA Center for Computational Sciences

Application	# Cores	Halem	Explore	Discover
GTRAJ	40	42s	47s	23s
GEOS GCM 1 deg	64		2783s	2972s
GEOS GCM 0.5 deg	120		5233s	2p 4874s
GEOS GCM 0.5 deg	180		???	2p 3561s
			< 4000s	3p 3978s
			???	4p 5104s
GMI Aerosol 4 deg	16		630s	537s
GMI Combo 2 deg	61	1813s		1p 1035s
				2p 2640s



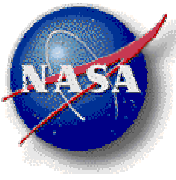
# Performance Expectations



NASA Center for Computational Sciences

- Many applications should see significant performance improvement over halem out-of-the-box.
- A few applications may see modest-to-severe performance degradation.
  - Still investigating causes.
  - Possibly due to smaller cache.
  - Using fewer cores per node may help absolute performance in such cases, but is wasteful.
  - Expect training classes for performance profiling and tuning.
- Please any performance observations or concerns to USG.



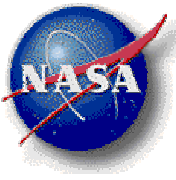


# Agenda



NASA Center for Computational Sciences

- Introduction—Phil Webster
- Systems Status—Mike Rouch
- Linux Networx Cluster update—Dan Duffy
- Preliminary Cluster Performance—Tom Clune
- ➔ New Services—Secure Unattended Proxy—Ellen Salmon
- User Services Updates—Sadie Duffy
- Questions or Comments

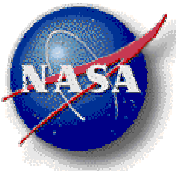


# What's SUP?



NASA Center for Computational Sciences

- Allows for unattended (i.e., batch job, cron) transfers directly to dirac and palm (discover coming soon)
- SUP development targeted Columbia/NAS $\Leftrightarrow$ NCCS transfers, but works from other hosts also
- After one-time setup, use your RSA SecurID fob to obtain one-week “keys” for subsequent fob-less remote commands:
  - scp, sftp
  - qstat
  - rsync
  - test
  - bbftp

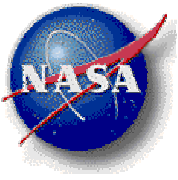


# Using the SUP - Setup



NASA Center for Computational Sciences

- Detailed online documentation available at:  
<http://nccs.nasa.gov/sup1.html>
- One-time setup:
  - Contact NCCS User Services to become activated for SUP
  - Copy desired ssh authorized\_keys to NCCS SUP
  - Create/tailor .meshrc file on palm, dirac (and soon discover) to specify file transfer directories
- Obtain one-week key from NCCS SUP using your RSA SecurID fob



# Using the SUP

## Once You Have Your SUP Key(s)



NASA Center for Computational Sciences

- Start an ssh-agent and add to it your one-week SUP key(s), e.g.:

```
eval `ssh-agent -s`  
ssh-add ~/.ssh/nccs_sup_key_1 ~/.ssh/nccs_sup_key_2
```

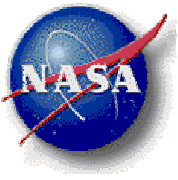
- Issue your SUP-ified commands: simpler examples

– qstat

```
ssh -Ax -oBatchMode=yes sup.nccs.nasa.gov \  
ssh -q palm qstat
```

– test

```
ssh -Ax -oBatchMode=yes sup.nccs.nasa.gov \  
ssh -q palm test -f /my/filename
```



# Using the SUP

## Once You Have Your SUP Key(s)



NASA Center for Computational Sciences

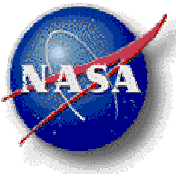
- Issue your SUP-ified commands: via wrapper script
  - Edit and make executable a ‘supwrap’ script containing:  

```
#!/bin/sh  
exec ssh -Ax -oBatchMode=yes sup.nccs.nasa.gov ssh -q $@
```
  - Perform file transfers using your ‘supwrap’ script, e.g.:  

```
scp -S ./supwrap file1 dirac.gsfc.nasa.gov:/my/file1
```
- Other techniques, tips (e.g., command-line aliases) are in the online documentation
  - Includes strategies for using the SUP from batch job scripts

<http://nccs.nasa.gov/sup1.html>

- *Questions?*

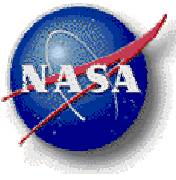


# Agenda



NASA Center for Computational Sciences

- Introduction—Phil Webster
- Systems Status—Mike Rouch
- Linux Networx Cluster update—Dan Duffy
- Preliminary Cluster Performance—Tom Clune
- New Services—Secure Unattended Proxy—Ellen Salmon
- User Services Updates—Sadie Duffy
- Questions or Comments



# Allocations

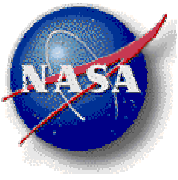


NASA Center for Computational Sciences

- NASA HEC Program Office working on formalizing the process for requesting additional hours
- New projects expecting to start on FY07Q2 (Feb. 1<sup>st</sup>) must have requests posted to e-Books by January 20<sup>th</sup>, 2007.

<https://ebooks.reisys.com/submission>.

- Quarterly reports due February 1<sup>st</sup>, 2007 to the e-Books system, Annual reports due for projects ending in February.



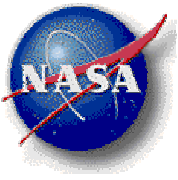
# Service Merge with NAS



## NASA Center for Computational Sciences

- One of the Unified NASA HEC initiatives
- Goal is to have a common Username, and Groupname space within the NASA HEC program
- What does this mean to you?
  - Same login at each center
  - Data created at each center within a given project will have the same number, making transfers easier
  - Update user information in one place
  - All users will be able to use online account request interface
  - Helps to streamline the allocation request process



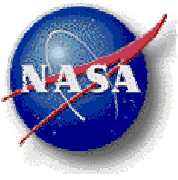


# Explore Queue Changes



NASA Center for Computational Sciences

- **datamove**
  - 2 jobs per user at one time to run.
  - Job size is limited to 2 processors.
  - Only 10 processors in total set aside for this queue and run only on backend systems.
  - Must specify queue in qsub parameters
- **pproc**
  - Jobs now limited to no more than 16 processors, 6 jobs per user at one time
  - Must specify queue in qsub parameters
- **general\_small**
  - Job size is limited to 18 processors.

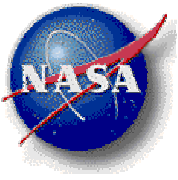


# NCCS Training



NASA Center for Computational Sciences

- Training is provided both by the NCCS and our partners at SIVO
- Looking for input from users
  - What topics?
  - What type?
    - Online Tutorials
    - One-on-one
    - Group training
  - What frequency?
- Issues with training received in the past?

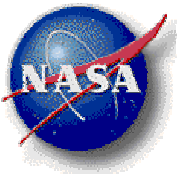


# Porting Help



NASA Center for Computational Sciences

- Hands-on training for porting to Discover.
- Up to 4 users in each session.
- The sessions will be 9:00-12:00 or 1:30-4:30.
- Locations are anywhere that has a Guest-CNE network, but typically in B33.
- Users should submit requests to USG, and ASTG will try to schedule the sessions as frequently as possible. (Probably 1-2 per week at first.)
- Requirements: laptop with wireless connectivity and access to CNE guest network. If users cannot meet the above requirement ASTG will work with the user to find an accommodation. E.g. we have a couple of spare laptops.



- Questions?
- Comments?