



# NCCS Brown Bag Series Introduction to Discover

October 22 2020

# Today's Program

- [Discover \(the Borg\)](#)
- [Systems and Components](#)
- [RSA Tokens and Passcodes](#)
- [NCCS LDAP Passwords](#)
- [System Login](#)
- [Files and Data](#)
- [Compute](#)
- [Useful Links and Contact Info](#)
- [Q & A](#)



# Discover The Borg (History)

Bruce Pfaff

Lead HPC Systems Engineer  
NASA Center For Climate Simulation  
Goddard Space Flight Center

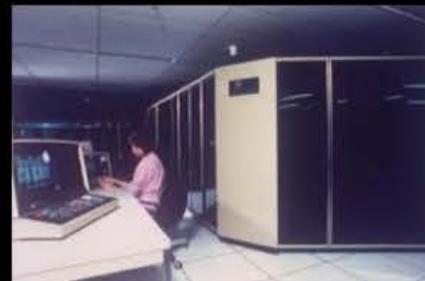


# We've Been Here A LONG Time



## SCF – Scientific Computing Facility

- IBM 360 & 370 class systems



## VPF – Vector Processing Facility

- CDC Cyber 205 + Amdahl V6 & V7 + ETA-10



## NSESCC – NASA Space and Earth Science Computing Center

- Cray Y-MP, C98, T3D, T3E, J90, Y-MP EL, IBM-SP, Convex



## NCCS – NASA Center for Computational Sciences

- HP/Compaq Tru64 Cluster, Linux Cluster

## NCCS – NASA Center for Climate Simulation (est. 2010)

- Linux Cluster, ADAPT, Cloud, etc.



# Discover SCU14 – 2017/2018



# Where Discover All Started - 2006



I'm An Old Cluster



I'm A New Cluster



# Comparison of the old and the new



## I'm a 4-year old cluster

- I have 100 cabinets
- I have 384 nodes
- I have 24 TB of storage in 12 racks
- I have over 11 miles of cables and 10 switches
- I require 100 tons of cooling
- Yeah, but my peak computing capacity is 3.2 TF



Picture courtesy of the Apple web site.  
No permission was given by Apple to use  
this picture.

## I'm a brand new cluster

- I have only 5 cabinets
- I have 128 nodes
- I have 60 TB of storage in one-half of a rack
- I have only 0.5 miles of cables and 2 switches
- I require about 30 tons of cooling
- Well, my peak computing capacity is 3.3 TF



# We've Come A Long Way In 14 Years



**July 2006**

**Oct 2020**

**5 Racks**

**103 Racks**

**3.2 Tflops**

**6.8 Pflops**

**60 Tb disk**

**45 Pb disk**

**142 Nodes**

(2 x 2 cores)

**4100+ Nodes**

(2 x 20 cores)

**520 Cores**

**129,000+ Cores**

# The Discover Cluster: 13 Years Of Evolution



Year	Unit	Vendor	Intel Processor
2006	– Base Unit	Linux Networx	Dempsey
2007	– SCU1 & SCU2	Linux Networx	Woodcrest
2008	– SCU3 & SCU4	IBM	Harpertown
2009	– SCU5 & SCU6	IBM	Nehalem
2010	– SCU7	Dell	Westmere
2011	– SCU1,2,3,4+	IBM	Westmere
		<ul style="list-style-type: none"> <li>Decommissioned all the Dempsey, Woodcrest and Harpertown systems</li> </ul>	
2012	– SCU8	IBM	SandyBridge + Intel Phi's
2013	– SCU9	IBM	SandyBridge
		<ul style="list-style-type: none"> <li>Decommissioned SCU5 &amp; SCU6 (all the Nehalem systems)</li> </ul>	
2014	– SCU10	SGI	Haswell
		<ul style="list-style-type: none"> <li>Decommissioned SCU7 (Westmere systems)</li> </ul>	
2015	– SCU11 & SCU12	SGI	Haswell
		<ul style="list-style-type: none"> <li>Decommissioned SCU1-4 (last of the Westmere systems)</li> </ul>	
2016	– SCU13 <small>Decom SCU8</small>	SGI	Haswell
2017	– SCU14 <small>Decom ½ SCU9</small>	Supermicro	Skylake
2019	– SCU15 <small>Decom ½ SCU9</small>	Aspen	Skylake



Currently  
Active  
Hardware



# The Latest Addition... SCU15



**640 compute nodes**

**2 x 20-core processors per node**

**25,600 Intel Skylake processors**

**192 GB RAM per node**

**1.9 Pflops**



# Systems and Components

- Computing
  - [Discover](#) - Interactive and batch processing capabilities
  - [NCCS GitLab](#) - On-premise GitLab instance
- Centralized Storage
  - [Centralized Storage System](#) (CSS) - Hosts curated NASA and related datasets
- Data Services
  - [Dataportal](#) - Provides public access to some NCCS data through various services

# RSA Tokens and Passcodes

- RSA Tokens

- Managed by the NASA Enterprise Service Desk (ESD)
- Two types - “hard” and “soft”



Fob-based hard token



Phone-based soft token

- Passcodes

- When using a hard token, enter your pin and the six-digit token code
- When using a soft token, enter the eight-digit token code

# NCCS LDAP Passwords

- Managed by NCCS Support
- Most NCCS systems and resources use a single password
- Passwords are valid for 60 days
- Passwords will lock after 5 failed attempts
- Passwords cannot be changed again within 24 hours
- If you forget or lock your password, contact NCCS Support

# System Login

- Login mode

`ssh -Y user_id@login.nccs.nasa.gov`

**PASSCODE:** *Enter your PIN and the six-digit token code (when using a hard token)  
or Enter the eight-digit token code (when using a soft token)*

**Host:** *discover*

**Password:** *Enter your NCCS LDAP password*

- Direct mode

- Recommended for command-line users and file transfers
- Requires establishing ssh config settings on your local system; see NCCS web site ([link](#)) for step-by-step instructions

*\*Additional login information can be found in the [Q & A section](#) of this document*

# Files and Data

- [Cluster-Wide File System](#)
- [Quota Limits](#)
- [Node-Specific Temporary Storage](#)
- [Centralized Storage System \(CSS\)](#)
- [Data Management Plans](#)
- [Managing Your Files](#)
- [Data Sharing](#)
- [File Transfer](#)

# Cluster-Wide File System

- General Parallel File System (GPFS)
- Accessible from all Discover login and compute nodes
- Hosts key user directories
  - `$HOME` or `/home/user_id` - your home directory
    - Home directory storage is limited to 1 GB
    - Ideal for storing source code and scripts - backed up daily
  - `$NOBACKUP` or `/discover/nobackup/user_id` - your short-term storage data area
    - Nobackup directory storage is limited to 5 GB and 100,000 inodes (files)
    - **Not backed up!** Long-term storage data should be moved elsewhere

# Quota Limits

- Two kinds of user-specific storage quotas are enforced:
  - Storage space used in \$HOME (1 GB) and \$NOBACKUP (5 GB)
  - Number of inodes (files) in \$NOBACKUP (100,000)
- Two types of quota limits are in place:
  - Hard limits – can never be exceeded. Any attempt to use more than your hard limit will be refused with an error
  - Soft limits – can be exceeded temporarily. When soft limit is exceeded, 7-day grace period goes into effect. You have to bring usage back below the soft limit value within the grace period, or any attempts to use more storage will be refused with an error

# Node-Specific Temporary Storage

- Node-specific Scratch Space
  - [\\$LOCAL\\_TMPDIR](#) - fast performing file system, but NOT global
  - Consider using it if your applications create, read, or write many small-size files
  - Files generated in \$LOCAL\_TMPDIR should be copied to \$NOBACKUP for later access. Files under \$LOCAL\_TMPDIR are scrubbed periodically

# Centralized Storage System (CSS)

- 30 PB of storage
- Provides storage of, and compute on, large NASA curated data sets from our HPC, Cloud, GPU, and Dataportal environments
- Provides data discovery and usage reporting to reduce data duplication and facilitate data deletion
- Manage the data lifecycle through Data Management Plans and policies

# Data Management Plans

- Four types of data:
  - Input – store on Discover or, if a curated dataset, on CSS
  - Intermediate – data created during software runs, store on Discover project space:
    - Not permanent
    - Not to be shared publicly
    - Could be restart/checkpoint files, research results, temporary files
  - Final – used for publications, shared with the science community or collaborators, could be input to other science programs – store on CSS
  - Software – save in a Git repository for re-use

# Managing Your Files

- Use the “*showquota*”\* command to check usage on \$HOME and \$NOBACKUP - see NCCS web site ([link](#)) for step-by-step instructions
- **\$NOBACKUP is NOT backed up.** It is your responsibility to copy valuable data to either \$HOME or to remote systems
- Always use \$HOME or /home/*user\_id*, and \$NOBACKUP or /discover/nobackup/*user\_id* in your scripts to specify paths

*\*Use “showquota -h” for human-readable output (e.g., “5G” instead of “5242880[K]”)*

# Data Sharing

- A common way to share files/directories with group members and others is to change permissions using the “*chmod*” command:

```
$ ls -l
drwx----- 2 cpan2 k3001 8192 2013-01-07 16:17 tmp/
$ chmod -R go+rx tmp | ls -l
drwxr-xr-x 2 cpan2 k3001 8192 2013-01-07 16:17 tmp/
$ chmod -R o-rx tmp | ls -l
drwxr-x--- 2 cpan2 k3001 8192 2013-01-07 16:17 tmp/
$ groups cpan2
cpan2 : k3001 k3002
$ chgrp -R k3002 tmp | ls -l
drwxr-x--- 2 cpan2 k3002 8192 2013-01-07 16:17 tmp
```

- **Do NOT make files/directories world-writable.** If you have a specific need to share data with group members or others, send a ticket to NCCS Support and we will help you!

# File Transfer to and from Discover

- Files can be transferred to and from Discover using scp, sftp, or rsync
- To copy data from a remote system to Discover, a user must use the [Bastion Service Direct Mode](#) - see NCCS web site [\(link\)](#) for step-by-step instructions
- Initiating commands from Discover to pull/push data from a remote system is also possible
- WinSCP users - See the NCCS website [\(link\)](#) for written instructions and an instructional video

# Discover Compute

- [Default Shell](#)
- [Cron Jobs](#)
- [Modules](#)
- [Compilers](#)
- [MPI Libraries](#)
- [Intel Math Kernel Libraries](#)
- [Standard Billing Units](#)
- [Running Compute Jobs via Slurm](#)
- [Debugging and Profiling Tools](#)
- [Licensed Application Software](#)
- [Open Source Software Packages](#)

# Default Shell

- “echo \$SHELL” to check your default shell, default is bash
- To change the default shell, contact NCCS Support

Shell	Startup files to edit
sh or ksh	~/.profile
bash	~/.bashrc if it exists; or ~/.bash_profile if it exists; or ~/.profile if it exists (in that order)
csh	~/.cshrc
tcsh	~/.tcshrc if it exists; or ~/.cshrc if it exists (in that order)

# Cron Jobs

- Manage your cron jobs at [discover-cron](#). [discover-cron](#) is an alias for a Discover login-style node that runs cron. From any of the Discover nodes, run
  - `ssh discover-cron`
- For batch jobs submitted via cron, you will first need to source `/etc/profile` to define bash environment variables:

```
0 1 * * * . /etc/profile ; sbatch myjob.sh 1>> FULLPATH/submit.out 2>&1
```

# Modules

- The “module” command allows you to choose compilers, libraries, and packages to create/change your own personal environment
- When you initially log into the NCCS system, **no modules** are loaded by default
- The module commands can be run in your shell startup file, your job script, or at the command line

Common Commands	Explanation
module avail (av)	Display a complete list of available modules
module list	Display loaded modules
module load <i>module_name1</i> ...	Load new modules
module purge	Unload all loaded modules
module show <i>module_name</i>	Display the environmental variables set by the module

# Compilers

- To accommodate the needs of a broad range of user groups, multiple versions of compilers from different vendors are provided
- Run module avail to see the versions available

Compiler	Access with:
GNU	<code>module load comp/gcc/<i>version</i></code>
Intel	<code>module load comp/intel/<i>version</i></code>
PGI	<code>module load comp/pgi/<i>version</i></code>
NAG	<code>module load comp/nag/<i>version</i></code>

# MPI Libraries

- Prior to loading an MPI module, you will have to load an appropriate module for a supported compiler suite.

Vendor	Modules	Supported Compilers
Intel MPI	mpi/impi	Intel Compiler only
HPCX	mpi/hpcx	GNU and Intel compilers
SGI-MPT	mpi/sgi-mpt	GNU, Intel, and PGI compilers

- MPI libraries are not visible in module avail until you select a compiler
- For new users, we recommend starting with Intel compiler and Intel MPI, for example,

```
module load comp/intel/19.1.2.254
```

```
module load mpi/impi/19.1.2.254
```

# Intel Math Kernel Library (MKL)

- Intel MKL is the primary numerical libraries with comprehensive math functionality, including BLAS, LAPACK, FFTs, Vector math, Statistics, and data fitting.
- MKL libraries are already included in LD\_LIBRARY\_PATH if you use Intel Compiler version  $\geq 17$
- If you use Intel Compiler version 10, PGI, or GNU compiler, and want to use MKL, you will need to load the compiler and an MKL module, lib/mkl-\*, e.g.,

```
module load comp/gcc/10.1.0
```

```
module load lib/mkl/19.1.2.254
```

# Standard Billing Units (SBUs)

- Computer resource allocations are quantified with SBUs. You can no longer run batch jobs if your allocated SBUs are used up.
- Command to check SBU balance and CPU hours used is:

`/usr/local/bin/allocation_check`

# Running Compute Jobs via Slurm

- Slurm is a distributed workload management system that handles the computational workload on Discover.
- Quality of Service (QoS) list:

QoS	Wall Time Limit	Max CPUs per job	Max jobs per user
allnccs (default)	12 hrs	6300	25
debug	1 hr	1120	1
long	24 hrs	560	25
serial	12 hrs	4116	1

- You do not need to specify “allnccs” to use the default QoS
- In order to run multi-node Slurm jobs, you have to set up a [\\$HOME/.ssh/authorized\\_keys](#) file.

# Common Slurm Commands

- Use Slurm commands to request both interactive and batch access to Discover computational resources.

QoS	Explanation
sbatch	Submit a batch job script for queueing and execution
salloc/xalloc	Submit an interactive job request
srun	Run a command within an existing job, on a subset of allocated resources
scancel	Cancel a queued or running job
squeue	query the status of your job(s) or the job queue

- View the [Slurm instructional video](#) on nccs.nasa.gov for a detailed explanation of how to use Slurm on Discover

# Debugging and Profiling Tools

- Debugging Tools – See the NCCS website for additional information
  - Code debugging
    - IDB and GDB
    - Totalview (<http://www.nccs.nasa.gov/images/Totalview-Part1-Doris.pdf>)
    - Allinea DDT
  - Memory debugging
    - Valgrind (<http://www.nccs.nasa.gov/images/Valgrind-brownbag.pdf>)
    - Totalview/MemScape (<http://www.nccs.nasa.gov/images/Totalview-Part2-Doris.pdf>)
    - Intel Inspector XE
- Profiling Tools – See the NCCS website for additional information
  - Gprof
  - MpiP
  - TAU (<http://www.nccs.nasa.gov/images/TAU-brownbag.pdf>)
  - Vtune Amplifier

# Licensed Application Software

- A few licensed applications from different vendors are installed on the NCCS systems:

- **Matlab :**

```
$ module load matlab
```

- **IDL:**

```
$ module load idl
```

```
$ /discover/vis/itt/idl/idl85/bin/lmstat -a
```

- **TOTALVIEW**

```
$ module load tview
```

# Open Source Software Packages

- A variety of open source software packages are installed under:  
[/usr/local/other/](#)
- After each system OS upgrade, some software are recompiled and older versions are retired. Users should always try to use the most recent build of a software.
- With few exceptions (e.g. Python or gcc) you can use most of third-party software directly WITHOUT loading modules
- A user may request installing a new package through NCCS Support

# Commonly Used Open Source Software

- Module environments:
  - **Python** : Python distributions (2.7, 3.x) for scientific computing
- /usr/local/other software:
  - **HDF4** and **HDF5**
  - **Netcdf3** and **Netcdf4**
  - **R**
  - **NCO**
- User maintained:
  - **GrADS** : Version 2.2.1.oga.1 -- /discover/nobackup/projects/gmao/share/dasilva/opengrads/Contents/opengrads

# Useful Links and Contact Info

- NCCS Web Site:

[nccs.nasa.gov](https://nccs.nasa.gov)

- Overviews and In-depth documentation on using Discover:

[nccs.nasa.gov/nccs-users/instructional/using-discover](https://nccs.nasa.gov/nccs-users/instructional/using-discover)

Help is always available by emailing

[support@nccs.nasa.gov](mailto:support@nccs.nasa.gov)

# Q & A (1 of 3)

## Questions, answers and comments from the Teams Chat:

Comment: If you have a NASA managed system that is PIV enabled to access NCCS systems, you also can use PIV based authentication in place of RSA token/soft-token. This is detailed on the setup for proxy mode: <https://www.nccs.nasa.gov/nccs-users/instructional/logging-in/bastion-host>.

Comment: [The General Parallel File System (GPFS) is also know as ] IBM SpectrumScale

Question: Why we can not transfer files from a GOOGLE DRIVE to Discover?

Answer: You can in fact copy data from google drive. The issues are limitations of clients accessing google drive data. Tools like wget/curl are limited in how they can access google drive. There is no native google drive client you can install on discover that I am aware of. There is: <https://github.com/prasmussen/gdrive> but this would need to download/build that software yourself.

Question: If NOBACKUP isn't backed up, what happens in the event of a systemwide hardware failure?

Answer: That is a risk, if we lost the disks that make up the filesystem, data would be lost. However, the disk we are talking about is highly redundant (RAID). If there is important data that you can not reproduce, you should back that data up outside of discover or copy it to other filesystems.

\$NOBACKUP resides on enterprise class, hardware RAID based storage subsystems and is considered to be a reliable storage environment. In the event of multiple hardware failure in the same places at the same time, it is possible that there could be data loss.

Comment: Intel MPI is the preferred MPI version on Discover. SGI-MPT is license restricted to only run on HPE Haswell nodes.

Comment: All brown bag slides (and other previous sessions) are available here: <https://www.nccs.nasa.gov/nccs-users/user-events/brown-bag-sessions>

# Q & A (2 of 3)

## Questions, answers and comments from the Teams Chat (Continued):

Comment: Additional information about using Slurm on Discover can be found here: <https://www.nccs.nasa.gov/nccs-users/instructional/using-slurm>

Comment: There are man pages for all of the native slurm commands. xalloc is a wrapper around salloc that provides X11 access to a SLURM interactive job.

Question: Can you talk about how to get remote window display to be able to use IDL / Matlab?

Answer: For Matlab/IDL, there are limited licenses. You need to limit yourself to a single active session.

Answer: You can only do X11 forwarding of display data from matlab/IDL. You need to ensure that X11 forwarding is functioning through ssh prior to starting IDL/Matlab.

Answer: `ssh -Y -C login.nccs.nasa.gov`

Question: Is it possible to use Conda env's?

Answer: Conda is available via the "python/GEOSpyD" environment modules, which are python distributions based on Anaconda:

```
# module avail python
```

```
----- /usr/local/other/modulefiles/Core -----  
python/GEOSpyD/Ana2018.12_py2.7 python/GEOSpyD/Ana2019.03_py2.7 python/GEOSpyD/Ana2019.10_py2.7 python/GEOSpyD/Min4.8.3_py2.7  
python/GEOSpyD/Ana2018.12_py3.7 python/GEOSpyD/Ana2019.03_py3.7 python/GEOSpyD/Ana2019.10_py3.7 python/GEOSpyD/Min4.8.3_py3.8 (D)
```

Question: Is there a simple way to estimate how many SBUs are needed?

Answer: The SLURM epilog now estimates SBU usage for a job (it's printed at the end after your sbatch job is done). That is a simple way to figure it out.

# Q & A (3 of 3)

## Questions, answers and comments from the Teams Chat (Continued):

Question: Are VisIt or ParaView available to run server-side and connect from client-side?

Answer: We do not have VisIt or ParaView installed. there are challenges to running client side applications. Our security does not allow ssh local port forwarding. I am not aware of any clients for visit or paraview that would be able to function without port forwarding and the login nodes for discover sit behind a firewall. The only access is via ssh through login.nccs

Question: How to use Google Earth Engine as Python library? In the Discover, it gives error message when we try to use GOOGLE Earth Engine

Answer: [https://developers.google.com/earth-engine/guides/python\\_install-conda](https://developers.google.com/earth-engine/guides/python_install-conda)

You can try to leverage that conda env. If you run into issues we would be more than happy to advise you further. Submit the error you are seeing and we can certainly point you in the right direction.