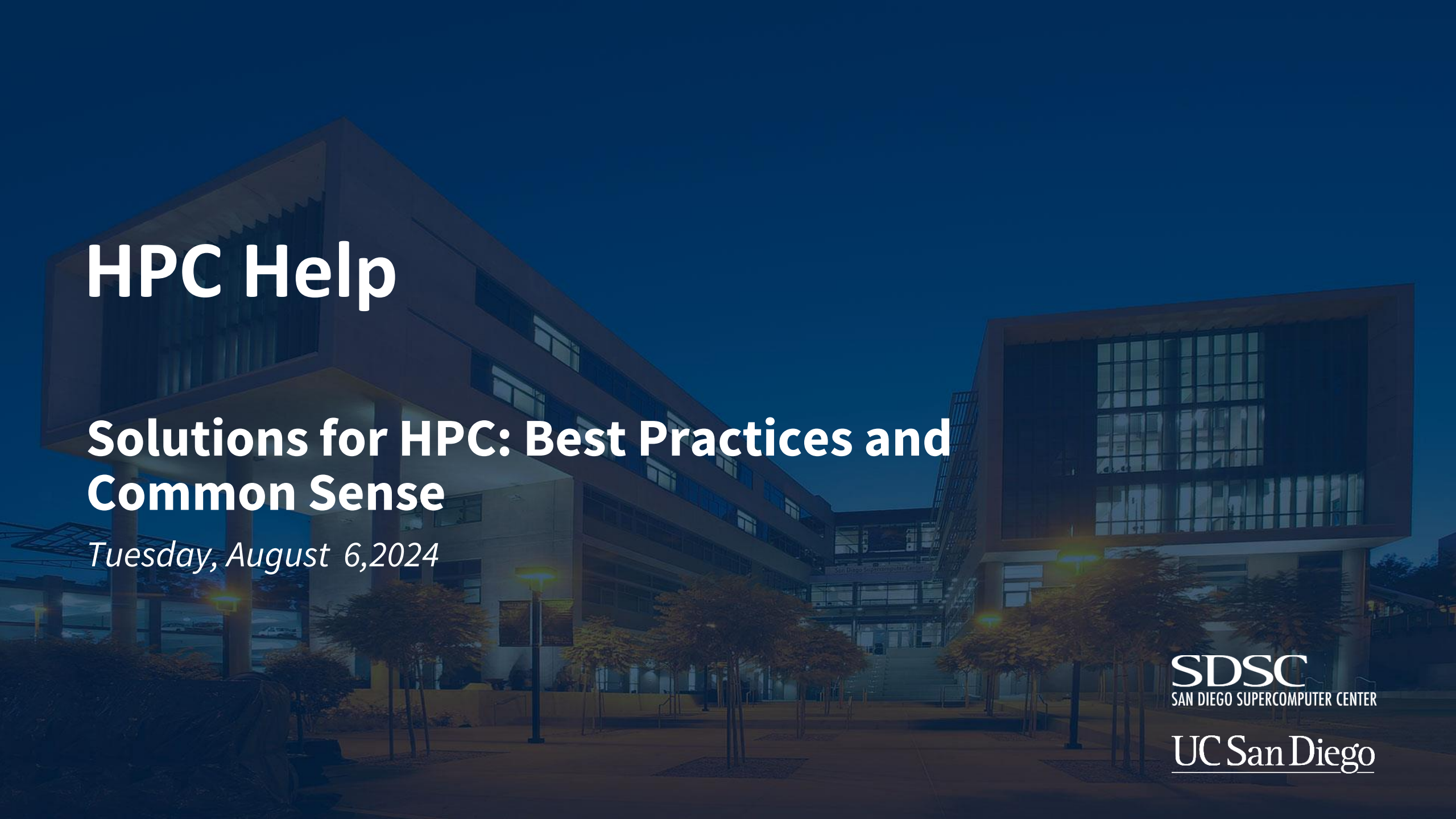


HPC Help

Solutions for HPC: Best Practices and Common Sense

Tuesday, August 6,2024

SDSC
SAN DIEGO SUPERCOMPUTER CENTER

UC San Diego

Agenda

- Motivation
- HPC Ecosystem
- Available Help Resources
- Common Issues & solutions
  - Login, file systems, performance, batch scheduler, allocations
- Summary Best practices

Motivation

- Based on experience with Help tickets
- When you use the Help desk , use it efficiently
- Goal to avoid the necessity to use Help Desk

HPC Ecosystem

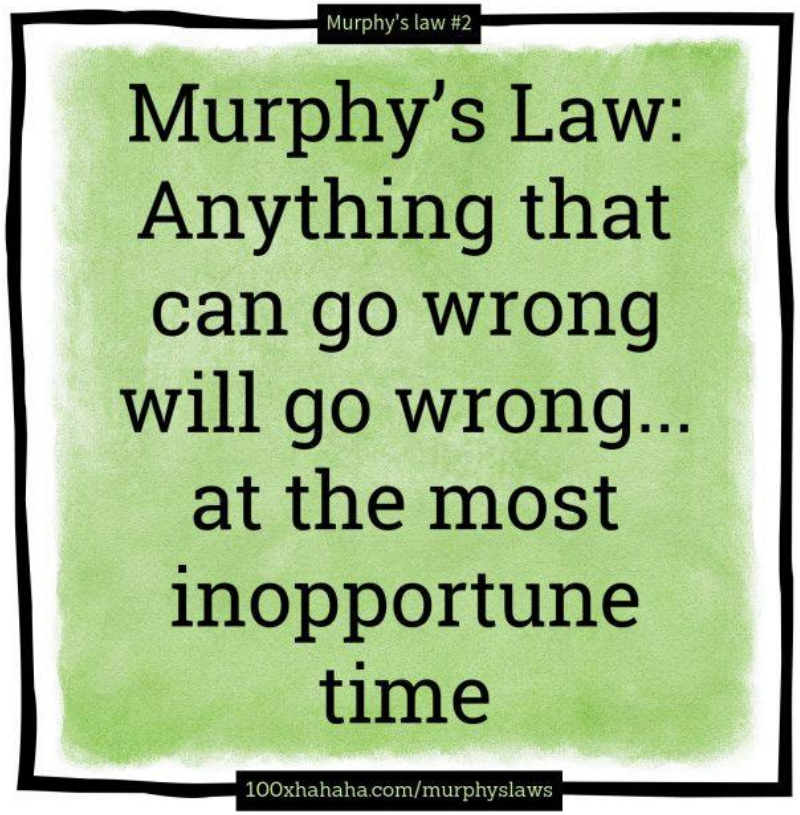
- You are not alone (HPC is a shared resource)
  - As a user (HPC ecosystem)
  - In getting support (available resources)
- What all is shared
  - Login Nodes
  - File systems
  - Compute nodes
  - System Administration, Security, Networking and System support teams
    - System policies

What are the available Help options?

- What are the available resources
  - Project/Colleagues
  - Web (Community forums, User guides, Git repositories)
  - Software tools
  - Helpdesk

General Tips

- Don’t leave HPC tasks to the last minute
- Back up your data
- Use version control



Help Desk: Useful Information

- Clear description of is issue
- Resource, UserID/Username, Account/Project/Allocation
- JobId
  - List of Node(s) that job ran on, Project, Start time, End time, Resources requested
- Working directory (submit script name)
- Location of .err and .out files
- Create new ticket for new issues
- Always be nice to the support desk! ☺

Common Issues & Solutions & Best Practices

- Access/Logging in
- Jobs/Policies/Schedulers
- Allocations/Job charging
- Applications/software
- File systems
- Performance

Common Issues: Resource Access/Logging in

- Different resources have different access protocols, or mechanisms for access, even if its at the same site
  - Portal, direct Access
  - Site local passwords, federated login(CI-logon), MFA, ssh keys
- Password and username issues
  - Problem: Indicator message: Enter verification code
  - Solution: check username https://allocations.access-ci.org/profile (if username is not available for the resource then the account has not been created yet) Generally it can take 1 business day for accounts to be fully functional.

```
 % ssh jombo@login.expnase.sdsc.edu
ssh: Could not resolve hostname login.expnase.sdsc.edu: Name or service not known
% ssh jombo@login.expanse.sdsc.edu
Verification code:
```

Common Issues: Resource Access/Logging in

- ssh keys
  - Problem: Indicator message: Enter password
  - Solution: resend ssh pub key, or use –i option to point to location of private key
- MFA(2FA)
  - Problem new device
  - Solution: Have an administrator delete the entry and start over

Common Issues: Resource Access/Logging In

- Unable to access system
  - Problem: System Maintenance
  - Solution: Patience, stay informed
    - Pay attention to MOTD
    - User News (Stay Subscribed to be notified)
    - Check on User Portal -- https://support.access-ci.org/announcements
- Hanging on login
  - Problem: Users overstimulating the file system
    - Solution: Contact Support staff
    - User Data mover nodes to move data
    - Don’t run compute intensive jobs on the login nodes
  - Problem: Activating Conda in your .bashrc
    - Solution: Do not initialize conda in .bashrc, but only when needed in run scripts

Common Issue: Batch scripts failure

- Job Scripts are used to request resources
  - Allocations
  - Compute resources (cpu, gpu)
    - Queues/Partitions/QOS
  - Time Limits
  - Memory
  - Set up environment and execute applications
  - Reservations
  - Licenses
  - Input files
  - Executables
  - Modules*
  - Location of output and error files

Common Issues: Job Submissions

- Submitting Jobs(sbatch)
  - Missing Allocation
  - Missing Software/License
  - Improper QOS
  - Insufficient Memory
  - Policies
    - Queues/Partitions
    - Time Limits
    - Memory

```
login01 user_support]$ sbatch azton.sb
sbatch: error: Batch job submission failed: Invalid account or
account/partition combination specified
```

```
login01 user_support]$ sbatch azton.sb
sbatch: error: QOSMaxCpuPerJobLimit
sbatch: error: Batch job submission failed: Job violates accounting/QOS
policy (job submit limit, user's size and/or time limits)
```

- Common Error messages
  - sbatch: error: Project balance is not enough to run the job
  - sbatch: error: QOSMaxNodePerJobLimit
- Error Message may be from Slurm or the bank plugin
  - Error messages are not always obvious

Common Issues: Queue and Time Limits

Partition Name	Max Walltime	Max Nodes/Job	Max Running Jobs	Max Running + Queued Jobs	Charge Factor	Notes
compute	48 hrs	32	32	64	1	Exclusive access to regular compute nodes; *limit applies per group*
ind-compute	48 hrs	32	32	64	1	Exclusive access to Industry compute nodes; *limit applies per group*
shared	48 hrs	1	4096	4096	1	Single-node jobs using fewer than 128 cores
ind-shared	48 hrs	1	32	64	1	Single-node Industry jobs using fewer than 128 cores
gpu	48 hrs	4	4	8 (32 Tres GPU)	1	Used for exclusive access to the GPU nodes
ind-gpu	48 hrs	4	4	8 (32 Tres GPU)	1	Exclusive access to the Industry GPU nodes
gpu-shared	48 hrs	1	24	24 (24 Tres GPU)	1	Single-node job using fewer than 4 GPUs
ind-gpu-shared	48 hrs	1	24	24 (24 Tres GPU)	1	Single-node job using fewer than 4 Industry GPUs
large-shared	48 hrs	1	1	4	1	Single-node jobs using large memory up to 2 TB (minimum memory required 256G)
debug	30 min	2	1	2	1	Priority access to shared nodes set aside for testing of jobs with short walltime and limited resources
gpu-debug	30 min	2	1	2	1	Priority access to gpu-shared nodes set aside for testing of jobs with short walltime and limited resources; *max two gpus per job*
preempt	7 days	32		128	.8	Non-refundable discounted jobs to run on free nodes that can be pre-empted by jobs submitted to any other queue
gpu-preempt	7 days	1		24 (24 Tres GPU)	.8	Non-refundable discounted jobs to run on unallocated nodes that can be pre-empted by higher priority queues

https://www.sdsc.edu/support/user_guides/expanse.html#running

Sinfo: Why is my job not running? Queue, wait times, Time Limits

- sinfo –used to view information about nodes and partitions

```
login01 ~]$ sinfo
PARTITION       AVAIL  TIMELIMIT  NODES  STATE NODELIST
compute            up 2-00:00:00      1  inval exp-1-15
compute            up 2-00:00:00      2 drain$ exp-13-[55-56]
compute            up 2-00:00:00      3 drain* exp-2-07,exp-5-50,exp-9-23
compute            up 2-00:00:00      7   drng exp-3-30,exp-4-17,exp-5-
[18,29,39,53],exp-7-45
compute            up 2-00:00:00      1  drain exp-4-55
compute            up 2-00:00:00     12   resv exp-2-[21-24],exp-14-50,exp-16-
[54-56],exp-17-[53-56]
compute            up 2-00:00:00    230    mix exp-1-[01-03,12-14,16,18,20,22-
26,28-39,41,43-44,47-48,51-54,56],
```

https://www.sdsc.edu/support/user_guides/expanse.html#running

Accounting

- Expanse uses expanse-client tool
  - Per user
  - Per Project

```
login01 ~]$ expanse-client user train112 -p

Resource  expanse

NAME      STATE  PROJECT  TG PROJECT     USED  AVAILABLE  USED BY PROJECT
-------------------------------------------------------------------------
train112  allow  gue998   TG-CIE960001S   337     200000         114621
```

- Different resources use different home grown tools to help users evaluate their usage.
  - expanse-client tool(SDSC)
  - TSCC_client tool (TSCC)
  - projects (PSC)
  - taccinfo (TACC)
- ACCESS Portal updated at various intervals

Common Issues: Accounting/Charging

- All systems charge differently
- ACCESS allocates in ACCESS Credits which can be converted to SUs (service Unit)
  - Each resource has a unique definition of an SU
- Allocations are shared
- Charging is generally based on what is requested, not how resources are used
- Do test jobs to evaluate
  - Slurm commands to collect information
    - sacct –u $USER
    - sacct –j $JOBID

Charging

- Charging is based on what is requested, not how resources are used
- Charging is based on the Maximum of memory and CPU core fraction
- Minimum charge for any job is 1SU

Example for CPU

Max [3600 * #CPU cores, 1800 * #Mem in GB] / 3600 * (wallclock time in secs/3600)

- 1 CPU and less then 2GB of memory are charged 1 CPU Service Unit(1SU = 1 core/hour).
- 1 GPU and up to 10 CPUs and 92 GB of memory are charged 1 GPU Service Unit (SU)/hour. This job will be charged 1 GPU SU/hour.
- The minimum charge for any job is 1 SU.
- 1 Expanse SUs = 1 ACCESS Credit
- 1 Expanse GPU SU = 54 Expanse SUs (for conversion)
  - https://allocations.access-ci.org/exchange_calculator

Allocations

- SDSC allocates resources via:
  - ACCESS-CI: (https://access-ci.org/)
    - **https://allocations.access-ci.org/prepare-requests-overview**
  - HPC@UC: https://www.sdsc.edu/collaborate/hpc_at_uc.html
  - HPC@MSI: https://www.sdsc.edu/collaborate/hpc_at_msi.html
  - Industrial Partners: https://www.sdsc.edu/collaborate/industry_and_sponsors.html

  - Trial accounts: **consult@sdsc.edu**

Common Issues: Job Status

- Slurm tools used to monitor and manage resources
  - squeue, scontol, sacct, sinfo
- squeue -- reports status and reason codes
  - Queued Jobs (*ReqNodeNotAvail, Reserved for maintenance)

```
[nickel@login01 ~]$ squeue | more
             JOBID PARTITION     NAME     USER ST       TIME  NODES NODELIST(REASON)
          13574113   compute 80dgree_   yweng3 PD       0:00      2 (MaxMemPerLimit)
          12668967   compute 0-xtensi  kavousi PD       0:00      1 (MaxMemPerLimit)
          14756880   compute job001_p  amytsai PD       0:00     10 (Reservation)
          14800161   compute namd-com sasadian PD       0:00      6 (QOSMaxCpuPerUserLimit)
          14800218   compute namd-com sasadian PD       0:00      6 (QOSMaxCpuPerUserLimit)
          14789098   compute bl_8JHNp    uscms PD       0:00      1 (MaxJobsPerAccount)
```

  - Running jobs

```
14813206   compute post0110 lpegolot  R   16:30:28      1 exp-9-35
14800090   compute namd-com sasadian  R   16:13:01      6 exp-2-29,exp-3-23,exp-4-33,exp-7-20,exp-9-[03,26]
14764467   compute V1WTReRU  aminkvh  R   16:08:56      1 exp-2-54
14773832   compute V4R1639Q  aminkvh  R   15:55:58      1 exp-8-14
14800092   compute namd-com sasadian  R   15:29:28      6 exp-4-29,exp-7-[07,39-40],exp-9-[28,41]
14812166   compute  scratch mlaskow2  R   15:53:59      1 exp-10-20
14812167   compute  scratch mlaskow2  R   15:39:34      1 exp-8-48
14800158   compute namd-com sasadian  R   15:17:18      6 exp-2-[26,50],exp-4-[52-53],exp-7-[42-43]
14812168   compute  scratch mlaskow2  R   15:20:01      1 exp-10-37
```

Common Issues: Jobs Pending

- squeue – Common “reasons” for pending jobs
  - MaxMemPerLimit
    - Max. mem per Node = 243G
  - QOSMaxNodePerUserLimit
  - Priority
  - ReqNodeNotAvail, Unavailable nodes: exp-x-xx
- File system not available
  - We have added a slurm directive #SBATCH –constraint = “lustre” to indicate if your job is using the lustre file system. If this is provided, the scheduler will not launch the job on a node that is missing lustre.
- System Maintenance
  - https://support.access-ci.org/outages

scontrol

- Slurm command to to view or modify Slurm configurations and state on currently queued or running jobs
- scontrol [OPTIONS...] [COMMAND...]
- Show individual job information
  - scontrol show job <<local_jobid>>

```
login01]$ scontrol show job 32658556
JobId=32658556 JobName=NGBW-JOB-BEAST2_XSEDE-43842C4A2D8C4B5085E71285ADC4D5DB
   UserId=testuser(505687) GroupId=use300(6099) MCS_label=N/A
   Priority=4509 Nice=0 Account=sds121 QOS=shared
   JobState=RUNNING Reason=None Dependency=(null)
   Requeue=1 Restarts=0 BatchFlag=1 Reboot=0 ExitCode=0:0
   RunTime=00:46:22 TimeLimit=03:00:00 TimeMin=N/A
   SubmitTime=2024-07-30T12:22:15 EligibleTime=2024-07-30T12:22:15
   AccrueTime=2024-07-30T12:22:15
   StartTime=2024-07-30T12:22:16 EndTime=2024-07-30T15:22:16 Deadline=N/A
   SuspendTime=None SecsPreSuspend=0 LastSchedEval=2024-07-30T12:22:16
Scheduler=Main
   Partition=shared AllocNode:Sid=login01:2295441
   ReqNodeList=(null) ExcNodeList=(null)
   NodeList=exp-16-53
   BatchHost=exp-16-53
   NumNodes=1 NumCPUs=1 NumTasks=1 CPUs/Task=1 ReqB:S:C:T=0:0:*:*
   ReqTRES=cpu=1,mem=1G,node=1,billing=3600
   AllocTRES=cpu=1,mem=1G,node=1,billing=3600
   Socks/Node=* NtasksPerN:B:S:C=1:0:*:* CoreSpec=*
   MinCPUsNode=1 MinMemoryNode=1G MinTmpDiskNode=0
   Features=(null) DelayBoot=00:00:00
   Reservation=cipres-shortjobs
   OverSubscribe=OK Contiguous=0 Licenses=cipres:1 Network=(null)
   Command=./_batch_command.run
   WorkDir=/expanse/projects//NGBW-JOB-BEAST2_XSEDE-
43842C4A2D8C4B5085E71285ADC4D5DB
   StdErr=/expanse/projects//NGBW-JOB-BEAST2_XSEDE-
43842C4A2D8C4B5085E71285ADC4D5DB/_scheduler_stderr.txt
   StdIn=/dev/null
   StdOut=/expanse/projects//NGBW-JOB-BEAST2_XSEDE-
43842C4A2D8C4B5085E71285ADC4D5DB/_scheduler_stdout.txt
   Power=
   MailUser=user@sdsc.edu
MailType=INVALID_DEPEND,BEGIN,END,FAIL,REQUEUE,STAGE_OUT
```

sacct

- sacct [OPTIONS...]
  - View accounting data for completed jobs and job steps

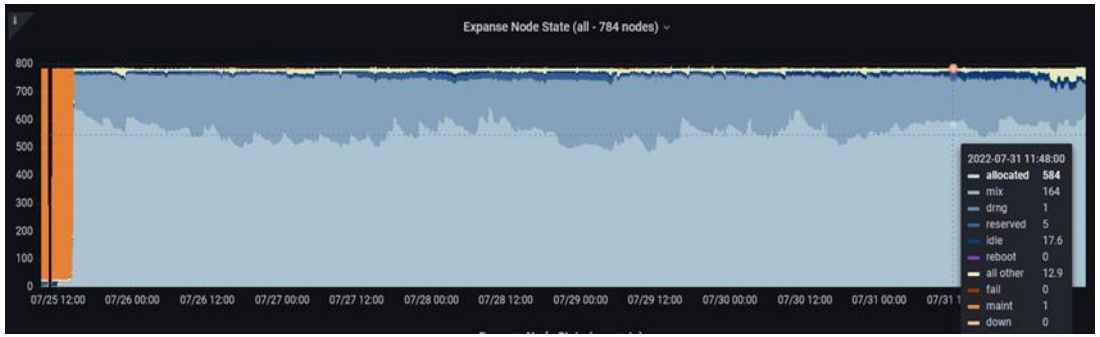
```
l@login01 user_support]$ sacct -j 32659299 -l
JobID        JobIDRaw        JobName  Partition  MaxVMSize  MaxVMSizeNode  MaxVMSizeTask  AveVMSize     MaxRSS MaxRSSNode MaxRSSTask     Ave
RSS MaxPages MaxPagesNode   MaxPagesTask   AvePages     MinCPU MinCPUNode MinCPUTask     AveCPU   NTasks  AllocCPUS    Elapsed      State Ex
itCode AveCPUFreq ReqCPUFreqMin ReqCPUFreqMax ReqCPUFreqGov     ReqMem ConsumedEnergy  MaxDiskRead MaxDiskReadNode MaxDiskReadTask    AveDis
kRead MaxDiskWrite MaxDiskWriteNode MaxDiskWriteTask   AveDiskWrite    ReqTRES  AllocTRES TRESUsageInAve TRESUsageInMax TRESUsageInMaxNode T
RESUsageInMaxTask TRESUsageInMin TRESUsageInMinNode TRESUsageInMinTask TRESUsageInTot TRESUsageOutMax TRESUsageOutMaxNode TRESUsageOutMaxTas
k TRESUsageOutAve TRESUsageOutTot
```

- For formatted output

```
@login01 user_support]$ sacct -j 28582088 --format=JobID%20,Partition,CPUTime,MaxRSS,MaxDiskRead,MaxDiskWrite,NCPUs,ReqNodes,NNodes,ReqMem
               JobID  Partition    CPUTime     MaxRSS  MaxDiskRead MaxDiskWrite      NCPUS ReqNodes   NNodes     ReqMem
-------------------- ---------- ---------- ---------- ------------ ------------ ---------- -------- -------- ----------
            28582088     shared 16-00:27:44                                          64        1        1    112.50G
      28582088.batch            16-00:27:44  12143272K      232.05M      120.89M         64        1        1
     28582088.extern            16-00:27:44      1856K        0.01M        0.00M         64        1        1
```

sacctmgr

- Slurm command to view or modify Slurm account information
- sacctmgr show qos
- Lists policies on queue/partitions, wait times, Time Limits
- User Guide




```
login01 ~]$ sacctmgr show qos format=name%20,MaxWall,MaxTRESPU%20,MaxJobsPU,MaxSubmitPU,MaxTRESPA%20,MaxJobsPA,MaxSubmitPA
                Name     MaxWall             MaxTRESPU MaxJobsPU MaxSubmitPU             MaxTRESPA MaxJobsPA MaxSubmitPA
-------------------- ----------- -------------------- --------- ----------- -------------------- --------- -----------
              normal  2-00:00:00     cpu=8192,node=64        32          64   cpu=16384,node=128        32          64
       shared-normal  2-00:00:00     cpu=8192,node=64      4096        4096   cpu=16384,node=128      4096        4096
 large-shared-normal  2-00:00:00                              2           4                               4           4
     preempt-normal  7-00:00:00                                        128     cpu=4096,node=32                    128
         gpu-normal  2-00:00:00 cpu=160,gres/gpu=16+           4           8   gres/gpu=32,node=8        8          16
   gpu-shared-normal  2-00:00:00 cpu=240,gres/gpu=24+          24          24 cpu=320,gres/gpu=32+        24          24
  gpu-preempt-normal  7-00:00:00   gres/gpu=24,node=6          12          16  gres/gpu=48,node=12        16          20
       debug-normal    01:00:00       cpu=256,node=2           2           2       cpu=256,node=2         2           2
   gpu-debug-normal    00:30:00    gres/gpu=8,node=2           2           2    gres/gpu=8,node=2         2           2
```

Common Issues: Software

- Software Installs
  - Modules
  - Compile (Build your own)
  - Containers
- Most HPC system use module to manage their software stacks
- Login nodes, compute nodes are different
  - Compiling codes need to happen on the nodes on which they will run
  - Modules need to be loaded on the nodes on which applications will run
- Software managed by the resource provider has usually been tested and fine tuned for the specific resources

Environment Modules

- Environment Modules help manage software incompatibilities, versioning, and dependencies
- Environment Modules provide for dynamic modification of your shell environment
- Module commands set, change, and/or delete environment variables
- Modules manage software versions
- Module manage dependencies by loading or unloading other modules
  - Check for dependencies with module spider <module_name>
- Module list - lists all the currently loaded modules

Navigating with Modules

Command	Description
`module list`	Currently Loaded Modules
module avail	List of available software modules based on your current module path
module spider	List all available software and versions on the system
Module spider <application name>	List available application specific modules and module details, including versions and dependencies
module load [module file]	Load module(s) or specify unresolved dependencies
module show [module file]	Display information about loaded modules including changes, dependencies, versions and paths
module unload [module file]	Unloads a specified module form the environment
module purge	Unloads all the loaded modules
module reset	Reset modules to default settings

Common Issues: File systems

- Common File systems and their utilization
  - Home – Usually limited in space
  - Scratch - Large space, limited persistence, no backup
  - Node local scratch(SSD)- good performance, only available during job
  - Persistent storage – good performance, limited persistence
  - Archival- Slow, backup
- Types of file systems
  - Lustre
  - nsf
  - gpfs
  - ceph

Common Issues: File systems

- Cant write to filesystem
  - Problem: quota reached
  - Solution: Clean up or archive old files
    - Review usage with du command
- Review file system usage
  - lfs (Lustre)
  - du (nfs)

Common Issues: File sharing

- Problem: collaborators cant see, edit files
- Solution: Use Unix commands chmod, chown, chgrp to manage file permissions
  - Be careful
- Review file/directory ownership and permissions
  - ls : list file contents
  - ls –laht(Attribute), User, Group, Other
  - Default 755 (User(read, write, execute) execute): Other(read, execute)
- Find common group for sharing
  - id : review available group
  - groups: review associated groups

```
login01 ~]$ ls -laht | more
total 11M
drwxr-xr-x 1092 root   root      0 Jul 31 14:15 ..
drwxr-x--x   43 nickel sdsc     94 Jul 31 12:54 .
-rw-------    1 nickel use300  22K Jul 31 12:54 .viminfo
-rw-------    1 nickel use300  47K Jul 30 20:26 .bash_history
drwxr-xr-x    8 nickel use300   29 Jul 30 20:03 user_support
drwx------    2 nickel sdsc      5 Jul  1 09:10 .ssh
drwxr-xr-x   11 nickel use300   46 Jun  6 08:58 comet
drwxr-xr-x    4 nickel use300   13 Jun  3 07:52 support
drwxr-xr-x    4 nickel use300    7 May 28 11:20 test
-rw-r--r--    1 nickel use300  104 May 28 11:19 err.30886666
-rw-r--r--    1 nickel use300    0 May 28 11:19 out.30886666
```

Common Issue: system performance

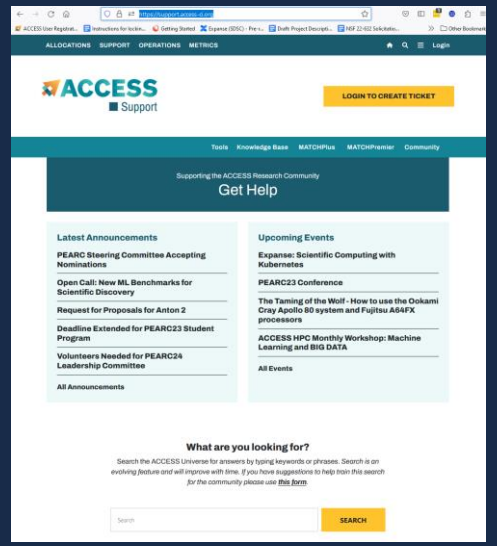
- GPU
- CPU
- Login Nodes
- Lustre

Monitoring Resource Utilization

- System evaluation Tools
  - top, htop, atop : display Linux processes
  - mpstat : display processors related statistic
  - sar : display system activity report
  - free: display memory statics
  - ps: display active processes
  - nvidia-smi: display gpu activity and statistic
- Job evaluation tools
  - (slurm efficiency) after job is complete
  - Slurm tools: sacct, sacctmgr
- Profiling tools
- Debuggers

Questions

- How to Reach Support
  - consult@sdsc.edu
  - https://support.access-ci.org/




NSF

Best Practices: Getting help

- Understand your problem
- Engage with appropriate support tools
- Help Desk
  - Provide relevant and adequate information for helpdesk to reduce iterations
    - Username, Account, System, Jobid, specific error message if available, etc.
    - The user with the problem should submit the help ticket
- Always be nice to the support desk! ☺

Best Practices: Secure your credentials

- Passwords
  - Don’t reuse passwords
  - Longer is better
  - Don’t keep digital plaintext copies of passwords
  - Don’t Hard code password in files
  - Don’t share passwords
- Use password-manager program
- Use SSH keys, ssh-agent
- Multi-factor Authentication(MFA)

Best Practices: Software

- Review user guide for tools available
- Use system installed applications when available
- Use containers to manage out of date software

Best Practices: jobs and job charging

- Check user guide for accounting policies
- Use system tools for most up to date accounting information
  - Slurm for individual job details
    - Sacctmgr
    - Sacct
    - Squeue
    - scontrol
    - Sinfo
  - Home grown tools for accounting information

Best practices: Have a backup plan

- Copy critical data of system regularly
- Version Control: Git
- Checksum data transfers to ensure no corruption
- ACCESS to HPC system and file retention is usually limited. Transferring data takes time
- Data on the system is the users responsibility
- Plan ahead for data transfers
- Convert many small files into a single archive file before transfer

Best Practices: Backups

- Backups should be done at regular intervals that make sense to your project
  - Frequency
  - Ensure backups are made on “good” versions
  - Perhaps retain a few versions just in case
- Don’t back up everything
  - Clean up unnecessary files
  - Backup files not easily reproduced or replaced
    - source code, scripts, config files and large output files
- Backups should be on a different resource
- Ensure credentials would not allow hackers to get onto the external resource
- Test the backup plan with the restore process