

HPC Hardware : Obtaining hardware information and monitoring performance

COMPLECS Team

<https://bit.ly/COMPLECS>

<https://github.com/sdsc-complecs>

SDSC
SAN DIEGO SUPERCOMPUTER CENTER

UC San Diego

About COMPLECS

COMPLECS (COMPrehensive Learning for end-users to Effectively utilize CyberinfraStructure) is a new SDSC program where training will cover non-programming skills needed to effectively use supercomputers. Topics include parallel computing concepts, Linux tools and bash scripting, security, batch computing, how to get help, data management and interactive computing.

*COMPLECS is supported by
NSF award 2320934.*



Table of Contents

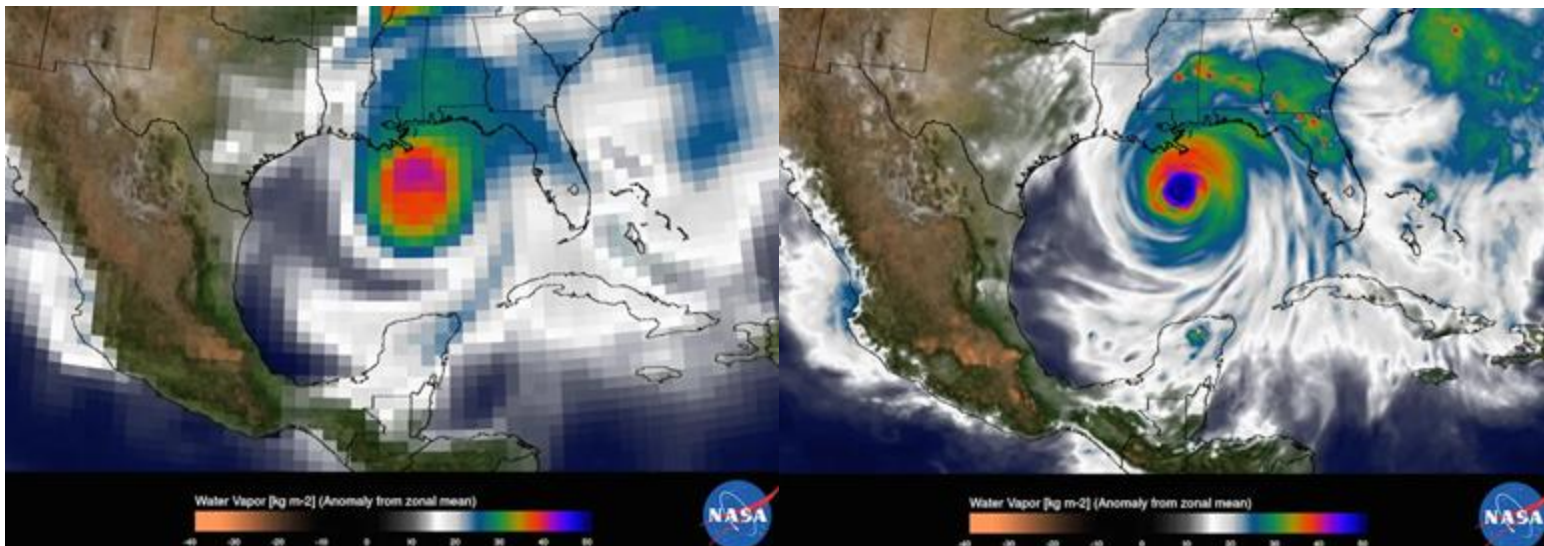
- What makes up an HPC system, and how is each associated to your work
- Hardware information
 - OS, Kernel
 - CPU (lscpu, /proc/cpuinfo)
 - GPU (....., nvidia-smi)
 - Memory (/proc/meminfo)
 - Cache
 - SCSI
 - File Systems (df, free, lfs, ...)
 - Network
- Hardware monitoring
 - top, htop, atop tools
 - mpstat
 - sar (System Activity Report)
 - Free
 - ps
- Job Monitoring
 - seff, sacct

Agenda

- Identify the Hardware components of HPC systems
- Understand why they matter to you
- Get information about your system
- Use some common tools to :
 - Monitor HPC resources
 - Monitor HPC Jobs on resources

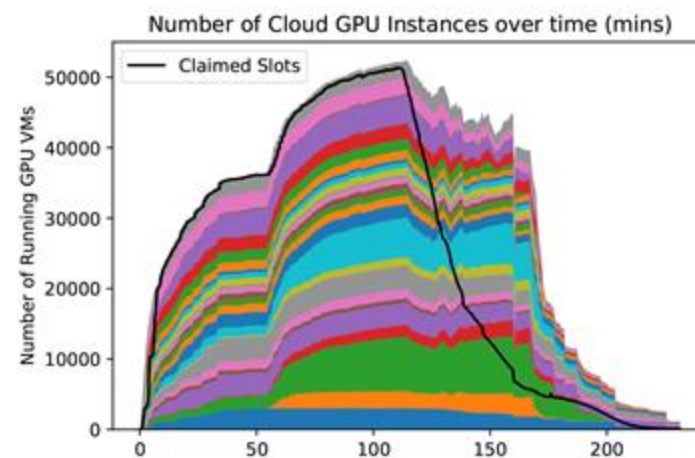
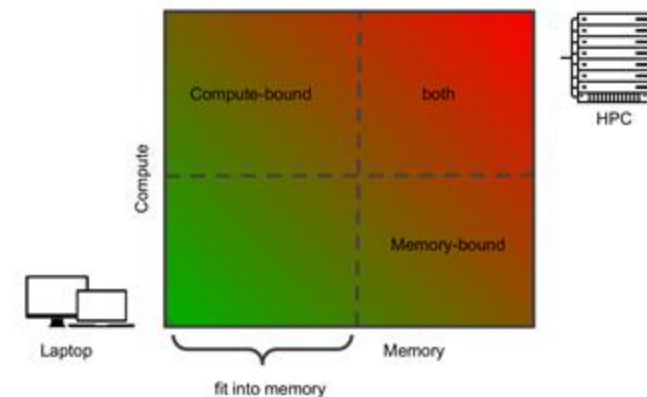
Introduction: High-Performance Computing (HPC)

High-performance computing (HPC) uses supercomputers and computer clusters to solve advanced computation problems.



Key advantages of leveraging HPC for your research:

- **Speed** - Solve a problem more quickly
- **Scale** - Solve a larger, more complex problem
- **Throughput** - Solve many (simple) problems more quickly



Finding OS and kernel information

- `uname` (unix name) – prints system information
 - `uname [OPTION]...`
- Use `uname` to get information on the Linux kernel

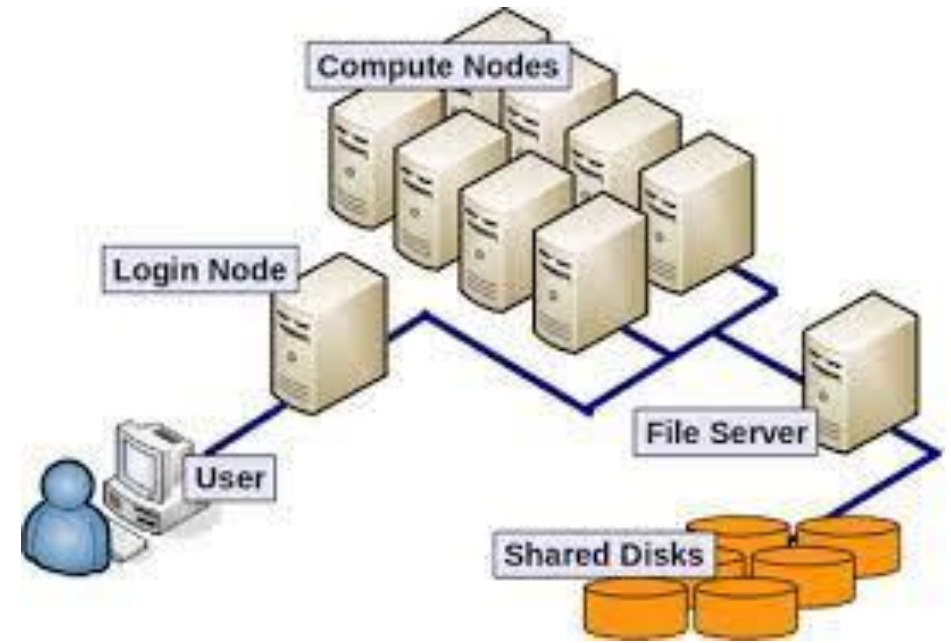
```
@login02]$ uname -a
Linux login02 4.18.0-477.15.1.el8_8.x86_64 #1 SMP Wed Jun 28 15:04:18
UTC 2023 x86_64 x86_64 x86_64 GNU/Linux
```

- Look in `/etc/centos-release` to get the Linux distribution (will vary by Linux distro)

```
@login02]$ cat /etc/centos-release
Rocky Linux release 8.8 (Green Obsidian)
```

Overview of Compute Nodes

- Nodes
 - Processors and accelerators
 - Sockets
 - Memory
 - Attached devices
- Networking
 - Interconnect
- Software and tools
- Storage



Getting hardware information – why do I care?

- Reporting
 - You may be asked to report details of your hardware in a manuscript, presentation, proposal or request for computer time
- Access
 - Resource Selection
 - Based on available resources will your job be able to run?
 - Estimating performance relative to another system. *All else being equal*, jobs will run at least as fast on hardware with
 - Faster CPU clock speeds
 - Larger caches
 - Faster local drives
- Identifying/Addressing system behavior
 - Job Compilation
 - Tune application for hardware
 - Efficient utilization
 - Save money and resources

Using the correct tool

- Myth 1: A supercomputer will improve code performance
- Myth 2: Requesting more hardware will reduce time to solution and improve efficiency
- HPC applications need to be explicitly programmed to take advantage of the hardware



CPU



GPU/CPU



GPUs



GPU&CPU



Large Memory

Obtaining hardware information

- User Guide
- Linux tools for collecting architecture information
 - **lscpu**
 - **/proc/ file system – Various Linux kernel statistics**
 - pseudo-file systems (/proc /sys)

Processor specifications: **lscpu**

On Linux systems, the **lscpu** command lists key processor information

- Number of processors (sockets)
- Processor type or model
- Nominal clock speed
- Number of cores per processor
- Cache sizes
- Instruction set architecture
- NUMA nodes

Processor specifications: (lscpu) *Expanse* compute node

```
Architecture:          x86_64
CPU op-mode(s):        32-bit, 64-bit
Byte Order:            Little Endian
CPU(s):                128
On-line CPU(s) list:   0-127
Thread(s) per core:    1
Core(s) per socket:    64
Socket(s):             2
NUMA node(s):          8
Vendor ID:             AuthenticAMD
CPU family:            23
Model:                49
Model name:            AMD EPYC 7742 64-Core Processor
Stepping:              0
CPU MHz:               3257.493
BogoMIPS:              4491.71
Virtualization:        AMD-V
L1d cache:             32K
L1i cache:             32K
L2 cache:              512K
L3 cache:              16384K
```

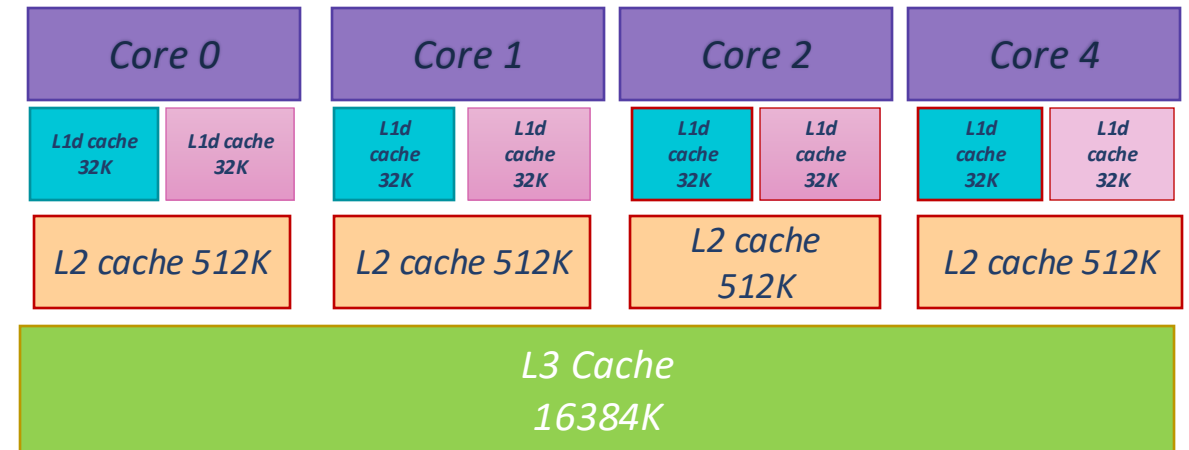
```
NUMA node0 CPU(s):    0-15
NUMA node1 CPU(s):    16-31
NUMA node2 CPU(s):    32-47
NUMA node3 CPU(s):    48-63
NUMA node4 CPU(s):    64-79
NUMA node5 CPU(s):    80-95
NUMA node6 CPU(s):    96-111
NUMA node7 CPU(s):    112-127
Flags:                 fpu vme de pse tsc msr pae mce
                      cx8 apic sep mtrr pge mca cmov pat pse36 clflush mmx
                      fxsr sse sse2 ht syscall nx mmxext fxsr_opt pdpe1gb
                      rdtscp lm constant_tsc rep_good nopl xtopology
                      nonstop_tsc cpuid extd_apicid aperfmperf pni
                      pclmulqdq monitor ssse3 fma cx16 sse4_1 sse4_2
                      x2apic movbe popcnt aes xsave avx fl6c rdrand
                      lahf_lm cmp_legacy svm extapic cr8_legacy abm sse4a
                      misalignsse 3dnowprefetch osvw ibs skinit wdt tce
                      topoext perfctr_core perfctr_nb bpext perfctr_llc
                      mwaitx cpb cat_l3 cdp_l3 hw_pstate sme ssbd mba sev
                      ibrs ibpb stibp vmmcall fsgsbase bmi1 avx2 smep bmi2
                      cqm rdt_a rdseed adx smap clflushopt clwb sha_ni
                      xsaveopt xsavec xgetbv1 xsaves cqm_llc cqm_occup_llc
                      cqm_mbm_total cqm_mbm_local clzero irperf xsaveerptr
                      wbnoinvd arat npt lbrv svm_lock nrip_save tsc_scale
                      vmcb_clean flushbyasid decodeas
```

<https://git.kernel.org/cgit/linux/kernel/git/stable/linux-stable.git/tree/arch/x86/include/asm/cpufeatures.h>

Processor specifications: (lscpu) *Exppanse* compute node

```
@exp-9-55 ~]$ lscpu -e
CPU NODE SOCKET CORE L1d:L1i:L2:L3 ONLINE
0 0 0 0 0:0:0:0 yes
1 0 0 1 1:1:1:0 yes
2 0 0 2 2:2:2:0 yes
3 0 0 3 3:3:3:0 yes
4 0 0 4 4:4:4:1 yes
5 0 0 5 5:5:5:1 yes
6 0 0 6 6:6:6:1 yes
7 0 0 7 7:7:7:1 yes
8 0 0 8 8:8:8:2 yes
9 0 0 9 9:9:9:2 yes
10 0 0 10 10:10:10:2 yes
11 0 0 11 11:11:11:2 yes
12 0 0 12 12:12:12:3 yes
13 0 0 13 13:13:13:3 yes
14 0 0 14 14:14:14:3 yes
15 0 0 15 15:15:15:3 yes
16 1 0 16 16:16:16:4 yes
17 1 0 17 17:17:17:4 yes
18 1 0 18 18:18:18:4 yes
19 1 0 19 19:19:19:4 yes
20 1 0 20 20:20:20:5 yes
. . . .. " .....
```

- CPU : CPU Index
- Node: NUMA node
- Socket: Index of physical socket to which the cpu belongs
- Core: Index of physical core
- L1d:L1i:L2:L3: Index of L1 Data Cache, L1 Instruction Cache, Index of L2 cache, Index of L3 cache.
 - Indicates that Exppanse has 128 L1d, L1i, L2 caches, once for each physical core
 - And 32 L3 Caches shared by 4 CPUs



Simultaneous multithreading (SMT) / hyperthreading

Thread(s) per core: 1

Simultaneous multithreading, abbreviated as SMT, is the process of a CPU splitting each of its physical cores into virtual cores, which are known as threads. This is done in order to increase performance and allow each core to run two instruction streams at once.

Intel branded this process as hyper-threading, but hyper-threading is the same thing as simultaneous multithreading. For example, AMD CPUs with four cores use simultaneous multithreading to provide eight threads, and most Intel CPUs with two cores use hyper-threading to provide four threads.

<https://www.tomshardware.com/reviews/simultaneous-multithreading-definition,5762.html>

SDSC does not enable hyperthreading on its systems. When hyperthreading is enabled, core count will appear to be doubled.

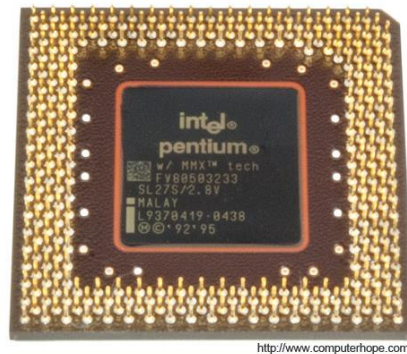
Advanced Vector Extensions (AVX, AVX2, AVX512)

Flags: fpu vme de **avx** tsc msr pae mce cx8 apic sep mtrr ... **avx2** mca ...

The Advanced Vector Extensions (AVX) are an extension to the x86 microprocessor architecture that allows a compute core to perform up to 8 floating point operations per cycle. Previous limit was 4/core/cycle

- AVX2 improves this to 16 Flops/cycle/core (Comet, Expanse)
- AVX512 further improves to 32 Flops/cycle/core (Intel \geq Skylake)

These were developed partially in response to challenges in increasing CPU clock speeds



March 6, 2000 8:00 AM PST

AMD makes move to 1-GHz chip

By Joe Wilcox and Michael Kanellos
Staff Writers, CNET News

Advanced Vector Extensions (AVX, AVX2, AVX512)

- Can theoretically obtain a 2x speedup when going from a non-AVX processor to an AVX capable processor (all else being equal)
 - And another 2x from AVX to AVX2
 - And another 2x from AVX2 to AVX512
- But don't get too excited (or worried that *Expanse* doesn't have AVX512)
 - It's difficult enough to make good use of AVX and even harder to make good use of AVX2 or AVX512.
 - Need long loops with vectorizable content. Memory bandwidth not keeping up with gains in computing power.
 - On Skylake, clock speed scaled down when executing AVX512 instructions

HPC System Architecture: Conceptual Model

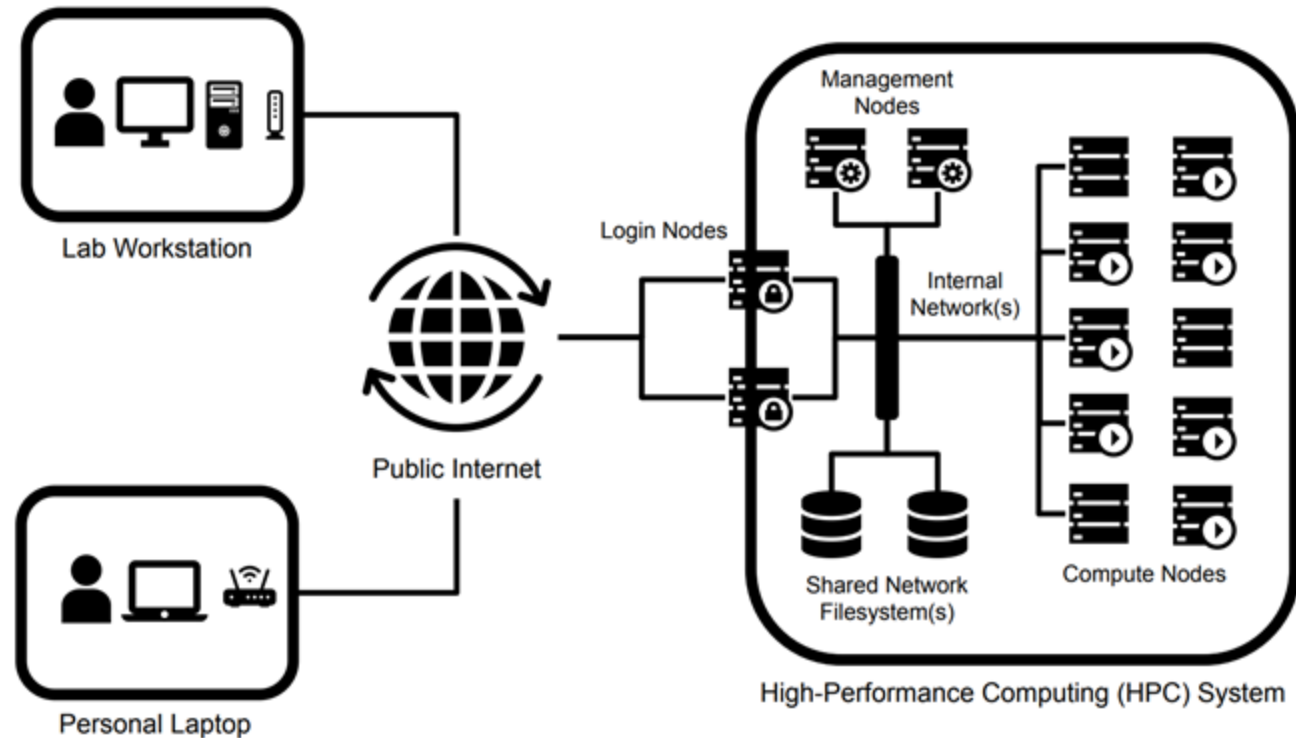
Login node(s): Provide remote access to an HPC system; use only for simple tasks such as editing files, limited data transfers to and from the system, and batch job submission

Compute nodes: Run computational workloads: simulations, data analysis and visualization

Data Mover nodes: Dedicated and configured specifically for transferring data

Internal Network(s): Provide high-bandwidth, low-latency communication between compute nodes ; access to shared (parallel) filesystems; system management

Shared Network Filesystem(s): Provide input/output (I/O) access to data storage systems from any compute node



Management node(s): Run core system services such as cluster management software, system monitoring software, *batch job scheduler*, etc

CPU vs. GPU

- CPU (Central Processing Unit)
 - large and broad instruction sets
 - managing every input and output of a computer
 - Use RAM
- GPU
 - More ALUs (processors)
 - Use VRAM
 - <https://www.heavy.ai/technical-glossary/cpu-vs-gpu>

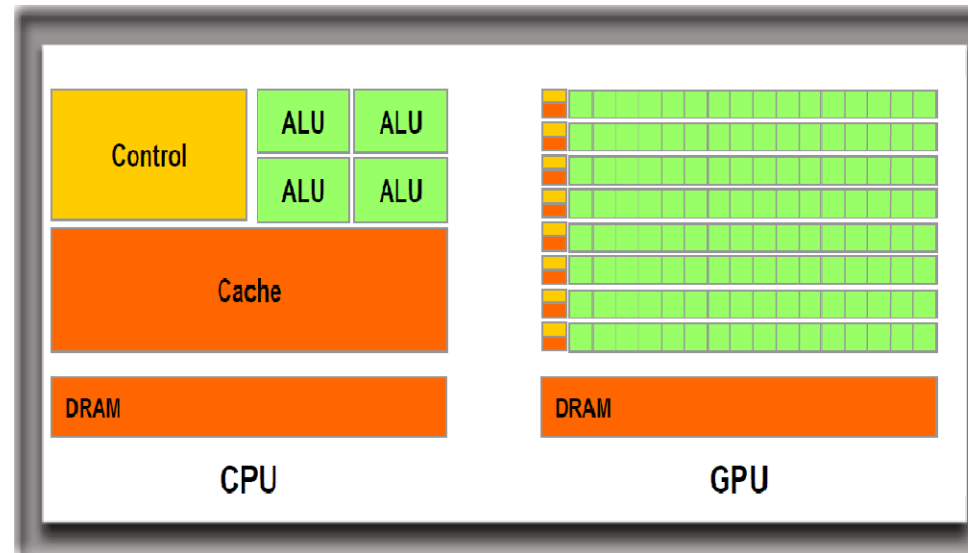


Image from NVIDIA

Processor Specifications: (lscpu) Expanse gpu node

```
Architecture:           x86_64
CPU op-mode(s):         32-bit, 64-bit
Byte Order:             Little Endian
CPU(s):                 40
On-line CPU(s) list:    0-39
Thread(s) per core:     1
Core(s) per socket:     20
Socket(s):              2
NUMA node(s):          2
Vendor ID:              GenuineIntel
CPU family:             6
Model:                 85
Model name:             Intel(R) Xeon(R) Gold 6248 CPU
@ 2.50GHz
Stepping:              7
CPU MHz:               1719.434
CPU max MHz:           3900.0000
CPU min MHz:           1000.0000
BogoMIPS:              5000.00
L1d cache:             32K
L1i cache:             32K
L2 cache:              1024K
L3 cache:              28160K
```

- NUMA node0 CPU(s):
0,2,4,6,8,10,12,14,16,18,20,22,24,26,28,30,32,34,36,38
- NUMA node1 CPU(s):
1,3,5,7,9,11,13,15,17,19,21,23,25,27,29,31,33,35,37,39
- Flags: fpu vme de pse tsc msr pae mce
cx8 apic sep mtrr pge mca cmov pat pse36 clflush dts
acpi mmx fxsr sse sse2 ss ht tm pbe syscall nx
pdpe1gb rdtscp lm constant_tsc art arch_perfmon pebs
bts rep_good nopl xtopology nonstop_tsc cpuid
aperfmpperf pni pclmulqdq dtes64 monitor ds_cpl smx
est tm2 ssse3 sdbg fma cx16 xtpr pdcm pcid dca sse4_1
sse4_2 x2apic movbe popcnt tsc_deadline_timer aes
xsave avx f16c rdrand lahf_lm abm 3dnowprefetch
cpuid_fault epb cat_l3 cdp_l3 invpcid_single
intel_ppin ssbd mba ibrs ibpb stibp ibrs_enhanced
fsgsbase tsc_adjust bmi1 avx2 smep bmi2 erms invpcid
cqm mpx rdt_a avx512f avx512dq rdseed adx smap
clflushopt clwb intel_pt avx512cd avx512bw avx512vl
xsaveopt xsavec xgetbv1 xsaves cqm_llc cqm_occup_llc
cqm_mbm_total cqm_mbm_local dtherm ida arat pln pts
pku ospke avx512_vnni md_clear flush_lld
arch_capabilities

Node information: (gpu) lscpu -e

```
[nickel@exp-7-59 ~]$ lscpu -e
```

CPU	NODE	SOCKET	CORE	L1d:L1i:L2:L3	ONLINE	MAXMHZ	MINMHZ
0	0	0	0	0:0:0:0	yes	3900.0000	1000.0000
1	1	1	1	1:1:1:1	yes	3900.0000	1000.0000
2	0	0	2	2:2:2:0	yes	3900.0000	1000.0000
3	1	1	3	3:3:3:1	yes	3900.0000	1000.0000
4	0	0	4	4:4:4:0	yes	3900.0000	1000.0000
5	1	1	5	5:5:5:1	yes	3900.0000	1000.0000
6	0	0	6	6:6:6:0	yes	3900.0000	1000.0000
7	1	1	7	7:7:7:1	yes	3900.0000	1000.0000
8	0	0	8	8:8:8:0	yes	3900.0000	1000.0000
9	1	1	9	9:9:9:1	yes	3900.0000	1000.0000
10	0	0	10	10:10:10:0	yes	3900.0000	1000.0000
11	1	1	11	11:11:11:1	yes	3900.0000	1000.0000
12	0	0	12	12:12:12:0	yes	3900.0000	1000.0000
13	1	1	13	13:13:13:1	yes	3900.0000	1000.0000

Getting GPU information

On GPU nodes, you can use nvidia-smi (NVIDIA System Management Interface program) to get GPU information (type, count, etc.)

When running in gpu-shared partition, will only see the GPUs you had requested (typically one)

```
$ nvidia-smi
Mon Aug  2 10:16:29 2021

+-----+
| NVIDIA-SMI 460.32.03      Driver Version: 460.32.03      CUDA Version: 11.2      |
+-----+-----+-----+
| GPU  Name           Persistence-M| Bus-Id        Disp.A | Volatile Uncorr. ECC |
| Fan  Temp  Perf    Pwr:Usage/Cap|      Memory-Usage | GPU-Util  Compute M. |
|                               |                  |              MIG M. |
+=====+=====+=====+
|   0   Tesla V100-SXM2...  On      | 00000000:18:00:0 Off  |              0      |
| N/A   40C    P0     67W / 300W |  0MiB / 32510MiB |    0%      Default  |
|                               |                  |              N/A    |
+-----+-----+-----+

+-----+
| Processes:                                                       |
| GPU  GI    CI           PID    Type    Process name                        GPU Memory |
|          ID    ID                                   Usage   |
+=====+
| No running processes found                                     |
+-----+
```

Processor specifications: (lspci) *Expanse* gpu node

- A graphics processing unit (**GPU**)
- lspci: information about PCI buses in the system and devices connected to them
 - `grep -E "VGA|3D"`
- lshw: List hardware
 - `-numeric -C display` (or video)

```
lspci | grep -E "VGA|3D"
03:00.0 VGA compatible controller: Matrox
Electronics Systems Ltd. Integrated Matrox G200eW3
Graphics Controller (rev 04)
18:00.0 3D controller: NVIDIA Corporation GV100GL
[Tesla V100 SXM2 32GB] (rev a1)
3b:00.0 3D controller: NVIDIA Corporation GV100GL
[Tesla V100 SXM2 32GB] (rev a1)
86:00.0 3D controller: NVIDIA Corporation GV100GL
[Tesla V100 SXM2 32GB] (rev a1)
af:00.0 3D controller: NVIDIA Corporation GV100GL
[Tesla V100 SXM2 32GB] (rev a1)
```

```
[nickel@exp-7-59 gpu]$ lshw -numeric -C display
WARNING: you should run this program as super-user.
*-display
  description: VGA compatible controller
  product: Integrated Matrox G200eW3 Graphics Controller [102B:536]
  vendor: Matrox Electronics Systems Ltd. [102B]
  physical id: 0
  bus info: pci@0000:03:00.0
  logical name: /dev/fbo
  version: 04
  width: 32 bits
  clock: 66MHz
  capabilities: vga controller bus master cap_list fb
  configuration: depth=32 driver=Matrox latency=0 maxlatency=32 minqnt=16 resolution=1024,768
  resources: irq:16 memory:91000000-91ffffff memory:95008000-9500bfff memory:94800000-94ffffff memory:c0000-dfff
*-display
  description: 3D controller
  product: GV100GL [Tesla V100 SXM2 32GB] [10DE:1DB5]
  vendor: NVIDIA Corporation [10DE]
  physical id: 0
  bus info: pci@0000:18:00.0
  version: a1
  width: 64 bits
  clock: 33MHz
  capabilities: bus master cap_list
  configuration: driver=nvidia latency=0
  resources: iomemory:38100-380ff iomemory:38180-3817f irq:249 memory:9e000000-9effffff memory:38100000000-3817ffffff memory:381800000000-381801ffffff
*-display
  description: 3D controller
  product: GV100GL [Tesla V100 SXM2 32GB] [10DE:1DB5]
  vendor: NVIDIA Corporation [10DE]
```

A brief aside on pseudo-files

Up to this point, we've been using the term pseudo-file without defining what it is. Recall that in the UNIX/Linux world, /proc and /sys are just interfaces to the Linux kernel data structures in a convenient and familiar file system format

everything is treated as a file (files, directories, devices, etc.)

```
$ ls -ld /proc
dr-xr-xr-x 2258 root root 0 Jul 28 09:27 /proc
[sinkovit@login01 ~]$ ls -ld /proc/cpuinfo
-r--r--r-- 1 root root 0 Jul 28 16:56 /proc/cpuinfo

$ head /proc/cpuinfo
processor      : 0
vendor_id     : AuthenticAMD
cpu family    : 23
Model.        : 49
model name    : AMD EPYC 7742 64-Core Processor
stepping      : 0
microcode     : 0x8301038
```

Processor specifications: /proc/cpuinfo

On Linux systems, the **/proc/cpuinfo** pseudo-file contains pretty much the same information that you get from `lscpu`, but with a few differences

- Information is listed per core
- Access to instantaneous clock speeds
- Bugs detected / addressed
see discussion: <https://unix.stackexchange.com/questions/456425/what-does-the-bugs-section-of-proc-cpuinfo-actually-show>
- Microcode, TLB size , power management, cache line flush sizes and other low-level details that you probably don't need to know about

Processor specifications: /proc/cpuinfo

```
$ grep 'cpu MHz' /proc/cpuinfo | head -5
cpu MHz      : 3325.325
cpu MHz      : 2239.263
cpu MHz      : 3374.887
cpu MHz      : 2360.551
cpu MHz      : 2393.493
```

```
--- selected output ---
microcode    : 0x8301038
bugs         : sysret_ss_attrs spectre_v1 spectre_v2 spec_store_bypass
TLB size     : 3072 4K pages
clflush size : 64
cache_alignment : 64
address sizes : 43 bits physical, 48 bits virtual
power management : ts ttp tm hwpstate cpb eff_freq_ro [13] [14]
```

Getting memory information: /proc/meminfo

On Linux machines, the /proc/meminfo pseudo-file lists key memory specs. More information than you probably want, but at least one bit of useful data

Regular Memory Node

MemTotal:	263698228 kB	(total physical memory)
MemFree:	251035032 kB	
MemAvailable:	250623760 kB	
Buffers:	12824 kB	
Cached:	3126364 kB	
SwapCached:	0 kB	
Active:	1301564 kB	(good approx used memory)
Inactive:	2990668 kB	
Active(anon) :	1240284 kB	
Inactive(anon) :	2890076 kB	
Active(file) :	61280 kB	
Inactive(file) :	100592 kB	
Unevictable:	0 kB	
Mlocked:	0 kB	
SwapTotal:	0 kB	
SwapFree:	0 kB	
Dirty:	32 kB	
Writeback:	0 kB	
AnonPages:	1151660 kB	

Large Memory Node

MemTotal:	2113365024 kB
MemFree:	2003813192 kB
MemAvailable:	2003837576 kB
Buffers:	1816060 kB
Cached:	3765288 kB
SwapCached:	19588 kB
Active:	18699308 kB
Inactive:	74523312 kB
Active(anon) :	15751904 kB
Inactive(anon) :	73244676 kB
Active(file) :	2947404 kB
Inactive(file) :	1278636 kB
Unevictable:	3072 kB
Mlocked:	0 kB
SwapTotal:	16759804 kB
SwapFree:	15768188 kB
Dirty:	0 kB
Writeback:	0 kB
AnonPages:	57923904 kB
Mapped:	1313932 kB
Shmem:	1355308 kB

GPU Node

MemTotal:	394817856 kB
MemFree:	390316552 kB
MemAvailable:	389003936 kB
Buffers:	8244 kB
Cached:	1028748 kB
SwapCached:	0 kB
Active:	176572 kB
Inactive:	2139400 kB
Active(anon) :	2308 kB
Inactive(anon) :	1410160 kB
Active(file) :	174264 kB
Inactive(file) :	729240 kB
Unevictable:	3072 kB
Mlocked:	0 kB
SwapTotal:	0 kB
SwapFree:	0 kB
Dirty:	3256 kB
Writeback:	0 kB
AnonPages:	1280240 kB
Mapped:	741024 kB
Shmem:	133256 kB

For more details, see <http://www.redhat.com/advice/tips/meminfo.html>

Finding Cache Information

On Linux systems, can obtain cache properties through the /sys pseudo filesystem. Details may vary slightly by O/S version and vendor, but basic information should be consistent. Goes deeper into cache properties than lscpu

```
$ pwd
/sys/devices/system/cpu

$ ls
cpu0    cpu17  cpu25  cpu33  cpu41  cpu5   cpu58  cpu9    offline
cpu1    cpu18  cpu26  cpu34  cpu42  cpu50  cpu59  cpufreq online
cpu10   cpu19  cpu27  cpu35  cpu43  cpu51  cpu6   cpuidle possible
cpu11   cpu2   cpu28  cpu36  cpu44  cpu52  cpu60  hotplug power
...

$ cd cpu0/cache
$ ls
index0  index1  index2  index3  power  uevent

$ cd index0
$ ls
coherency_line_size  physical_line_partition  size
id                   power                     type
level                shared_cpu_list          uevent
number_of_sets       shared_cpu_map            ways_of_associativity
```

Finding SCSI device information

SCSI (Small Computer System Interface) is a common interface for mounting peripheral, such as hard drives and SSDs. The `lsscsi` command or the `/proc/scsi/scsi` file will provide info on SCSI devices.

```
$ lsscsi
[3:0:0:0]    disk      ATA          SSDSC2KB480G8R  DL67  /dev/sda
[N:0:0:1]    disk      Dell Express Flash NVMe P4510 1TB SFF__1 /dev/nvme0n1
```



Dell 1TB PCIe
NVMe Read

File system information


- Pseudo filesystems
 - /proc/filesystem
 - /proc/mounts & /etc/mtab
- Linux commands
 - df – disk free
 - du – disk usage
- File sytem specific commands
 - lfs – lustre file system

/proc/filesystem and /etc/mtab

- /proc/filesystem
 - lists file system types currently supported by the kernel
- /etc/mtab
 - List currently mounted filesystems

df command: Information on filesystem usage

```
$ df -h
```

Filesystem	Size	Used	Avail	Use%	Mounted
on /dev/nvme0n1p1 	916G	77M	870G	1%	/scratch
ps-071.sdsc.edu:/ps-data/community-sw /expanse/community	1.0T	102G	923G	10%	
10.21.0.21:6789,10.21.11.7:6789,10.21.11.8:6789:/ /cm/shared	1.7T	553G	1.2T	33%	
10.22.101.123@o2ib:10.22.101.124@o2ib:/expanse/projects /expanse/lustre/projects	11P	1.4P	9.3P	13%	
10.22.101.123@o2ib:10.22.101.124@o2ib:/expanse/scratch /expanse/lustre/scratch	11P	1.4P	9.3P	13%	
10.22.100.113:/pool3/home/sinkovit /home/sinkovit	209T	4.1T	205T	2%	

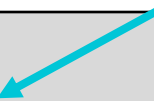
--- only selected filesystems shown ---

df command: information on filesystem usage

```
$ df -h
Filesystem                                Size  Used Avail Use% Mounted on
/dev/nvme0n1p1                            916G   77M   870G   1% /scratch
ps-071.sdsc.edu:/ps-data/community-sw     1.0T  102G   923G  10% /expanse/community
10.21.0.21:6789,10.21.11.7:6789,10.21.11.8:6789:/ 1.7T  553G   1.2T  33% /cm/shared
10.22.101.123@o2ib:10.22.101.124@o2ib:/expanse/projects 11P   1.4P   9.3P  13% /expanse/lustre/projects
10.22.101.123@o2ib:10.22.101.124@o2ib:/expanse/scratch 11P   1.4P   9.3P  13% /expanse/lustre/scratch
10.22.100.113:/pool3/home/sinkovit        209T   4.1T  205T   2% /home/sinkovit

--- only selected filesystems shown ---
```

Community and SDSC maintained software stacks

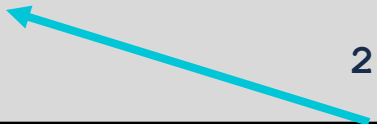


df command: Information on filesystem usage

```
$ df -h
```

Filesystem	Size	Used	Avail	Use%	Mounted on
/dev/nvme0n1p1	916G	77M	870G	1%	/scratch
ps-071.sdsc.edu:/ps-data/community-sw	1.0T	102G	923G	10%	
/expanse/community					
10.21.0.21:6789,10.21.11.7:6789,10.21.11.8:6789:/	1.7T	553G	1.2T	33%	/cm/shared
10.22.101.123@o2ib:10.22.101.124@o2ib:/expanse/projects	11P	1.4P	9.3P	13%	
/expanse/lustre/projects					
10.22.101.123@o2ib:10.22.101.124@o2ib:/expanse/scratch	11P	1.4P	9.3P	13%	
/expanse/lustre/scratch					
10.22.100.113:/pool3/home/sinkovit	209T	4.1T	205T	2%	
/home/sinkovit					

--- only selected filesystems shown ---



Lustre scratch and project filesystems

df Command: Information on filesystem usage

```
$ df -h
```

Filesystem	Size	Used	Avail	Use%	Mounted on
/dev/nvme0n1p1	916G	77M	870G	1%	/scratch
ps-071.sdsc.edu:/ps-data/community-sw	1.0T	102G	923G	10%	
/expanse/community					
10.21.0.21:6789,10.21.11.7:6789,10.21.11.8:6789:/	1.7T	553G	1.2T	33%	/cm/shared
10.22.101.123@o2ib:10.22.101.124@o2ib:/expanse/projects	11P	1.4P	9.3P	13%	
/expanse/lustre/projects					
10.22.101.123@o2ib:10.22.101.124@o2ib:/expanse/scratch	11P	1.4P	9.3P	13%	
/expanse/lustre/scratch					
10.22.100.113:/pool3/home/username	209T	4.1T	205T	2%	/home/username

Home filesystem

--- only selected filesystems shown ---

du command

- Disk usage
- --time : the time of the last modification to any file in the directory or subdirectory that you run it against
- -s: summary
- -h: “human readable”
- -a: lists the sizes of all files and directories in the given file path

Networking/interconnects

- Ethernet
- 10GbE
- Infiniband



Finding Network Information

The ip command (/sbin/ip) is normally used by sys admins, but regular users can use it to learn about networking information

```
$ /sbin/ip link
1: lo: <LOOPBACK,UP,LOWER_UP> mtu 65536 qdisc noqueue state UNKNOWN mode DEFAULT
group default qlen 1000
    link/loopback 00:00:00:00:00:00 brd 00:00:00:00:00:00
2: eno1: <BROADCAST,MULTICAST> mtu 1500 qdisc noop state DOWN mode DEFAULT group
default qlen 1000
    link/ether 6c:2b:59:bb:61:24 brd ff:ff:ff:ff:ff:ff
3: eno33: <BROADCAST,MULTICAST,UP,LOWER_UP> mtu 9000 qdisc mq state UP mode DEFAULT
group default qlen 1000
    link/ether 1c:34:da:62:a8:50 brd ff:ff:ff:ff:ff:ff
4: eno34: <BROADCAST,MULTICAST> mtu 1500 qdisc noop state DOWN mode DEFAULT group
default qlen 1000
    link/ether 1c:34:da:62:a8:51 brd ff:ff:ff:ff:ff:ff
5: ib0: <BROADCAST,MULTICAST,UP,LOWER_UP> mtu 4092 qdisc mq state UP mode DEFAULT
group default qlen 256
    link/infiniband 20:00:11:07:fe:80:00:00:00:00:00:00:1c:34:da:03:00:5d:53:90 brd
00:ff:ff:ff:ff:12:40:1b:ff:ff:00:00:00:00:00:00:00:ff:ff:ff:ff
6: eno33.450@eno33: <BROADCAST,MULTICAST,UP,LOWER_UP> mtu 9000 qdisc noqueue state
UP mode DEFAULT group default qlen 1000
    link/ether 1c:34:da:62:a8:50 brd ff:ff:ff:ff:ff:ff
```

Prefix	Network Type
en	Ethernet
ib	InfiniBand
sl	Serial line IP
wl	WLAN (wireless local area network)
ww	WWAN (wireless wide area network)

Monitoring Resource Utilization

- 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
- seff (slurm efficiency) after job is complete

top, htop, atop

- Used to analyze the load of a server running Linux
 - **top** - a standard utility, installed in all Linux versions by default
 - **htop** - easier to use than top, interactive
 - **atop** – easier then top, interactive, allows logging

Using the Linux top utility

The top utility is found on all Linux systems and provides a high-level view of running processes. Does not give any information at the source code level (profiling), but can still be very useful for answering questions such as

- How many of my processes are running?
- What are the states of the processes (running, sleeping, etc.)?
- Which cores are being utilized?
- Are there any competing processes that may be affecting my performance?
- What fraction of the CPU is each process using?
- How much memory does each process use?
- Is the memory usage growing over time? (Useful for identifying memory leaks)
- How many threads are my processes using?

top utility control

By Default:

- Reports on Processes only (no threads)
 - “H” - To toggle threads display, type “H” while top is running
- Displays information for all users
 - Can restrict to a single user by launching with “top -u username”
 - “u” – sort by username
 - “n” – limit number of process
- Updates information every 3 seconds
 - Change refresh rate by launching with “top -d n”
- Ordered by CPU usage
 - “M” - order by memory usage
- Process ID, priority, ‘nice’ level, virtual memory, physical memory, shared memory, state, %CPU, %memory, CPU time, command
 - “f” - to manage fields and toggle fields
 - “c” – display full path of process in COMMAND column

Top utility: default output

stivoknis — sinkovit@gcn-17-57:~ — ssh — 94x33

```
top - 08:37:00 up 60 days, 14:23, 1 user, load average: 15.32, 10.36, 6.12
Tasks: 624 total, 17 running, 607 sleeping, 0 stopped, 0 zombie
Cpu(s): 68.7%us, 1.3%sy, 0.0%ni, 29.9%id, 0.1%wa, 0.0%hi, 0.0%si, 0.0%st
Mem: 66054160k total, 37885796k used, 28168364k free, 8808k buffers
Swap: 2097144k total, 13400k used, 2083744k free, 32927192k cached
```

PID	USER	PR	NI	VIRT	RES	SHR	S	%CPU	%MEM	TIME+	COMMAND
70388	sinkovit	20	0	194m	76m	1612	R	100.0	0.1	1:31.06	lobfaster.pl
72547	sinkovit	20	0	120m	2976	1612	R	100.0	0.0	0:01.49	lobfaster.pl
72516	sinkovit	20	0	127m	9.9m	1608	R	100.0	0.0	0:02.09	lobfaster.pl
72526	sinkovit	20	0	121m	3388	1612	R	100.0	0.0	0:01.84	lobfaster.pl
72535	sinkovit	20	0	121m	4208	1612	R	100.0	0.0	0:01.73	lobfaster.pl
72565	sinkovit	20	0	120m	3212	1612	R	100.0	0.0	0:01.01	lobfaster.pl
72268	sinkovit	20	0	130m	12m	1612	R	98.9	0.0	0:11.96	lobfaster.pl
72359	sinkovit	20	0	123m	5976	1612	R	98.9	0.0	0:09.77	lobfaster.pl
72460	sinkovit	20	0	127m	10m	1612	R	98.9	0.0	0:08.38	lobfaster.pl
72481	sinkovit	20	0	131m	13m	1612	R	98.9	0.0	0:07.44	lobfaster.pl
72529	sinkovit	20	0	122m	4576	1612	R	98.9	0.0	0:01.82	lobfaster.pl
72439	sinkovit	20	0	130m	12m	1612	R	97.0	0.0	0:08.64	lobfaster.pl
72590	sinkovit	20	0	120m	3140	1612	R	71.7	0.0	0:00.37	lobfaster.pl
72602	sinkovit	20	0	120m	2576	1612	R	38.8	0.0	0:00.20	lobfaster.pl
72605	sinkovit	20	0	120m	2528	1600	R	34.9	0.0	0:00.18	lobfaster.pl
72608	sinkovit	20	0	119m	2340	1600	R	21.3	0.0	0:00.11	lobfaster.pl

Shown are 16 processes, each using anywhere from 21.3% to 100% of a compute core.

Memory footprint (RES) is minimal, with each process only using up to 76 MB.

CPU times ranging from 0.11s (just started) to 1:31

htop utility

- System monitoring
- Process Viewing
- Process Management
- Interactive commands
 - F1: help
 - M: sort by memory
 - P: sort by processor
 - T: sort by time
 - U: display for specific user

```

0[12] 4[27] 8[19] 12[27] 16[7.] 20[6.] 24[9.] 28[18] 32[13.] 36[26.] 40[15.] 44[22.] 48[10.] 52[16.] 56[35.] 60[18.]
1[9.] 5[6.] 9[10] 13[10] 17[9.] 21[17] 25[9.] 29[16] 33[10.] 37[12.] 41[13.] 45[11.] 49[11.] 53[16.] 57[8.9] 61[14.]
2[8.] 6[10] 10[17] 14[7.] 18[11] 22[17] 26[47] 30[10] 34[14.] 38[19.] 42[12.] 46[16.] 50[18.] 54[29.] 58[13.] 62[15.]
3[23] 7[12] 11[26] 15[8.] 19[27] 23[25] 27[13] 31[7.] 35[10.] 39[8.3] 43[10.] 47[14.] 51[10.] 55[22.] 59[5.6] 63[17.]
Mem[|||||] [94.6G/125G] Tasks: 6599, 6512 thr, 1235 kthr; 7 running
Swp[|||||] [11.5G/12.0G] Load average: 40.10 42.89 41.96
Uptime: 46 days, 01:02:49

Main 1/6
PID USER PRI NI VIRT RES SHR S CPU%MEM% TIME+ Command
627557 zheting 20 0 29024 6484 2356 I 46.8 0.0 0:02.59 cp CHG CHGCAR CONTCAR DOSCAR EIGENVAL IBZKPT INCAR KPOINTS OSZICA
3635 root 20 0 13.5G 192M 188M S 44.6 0.2 177h3m /usr/sbin/rsyslogd -n
3647 root 20 0 13.5G 192M 188M S 44.6 0.2 176h41m /usr/sbin/rsyslogd -n
3653092 keyajoshi 20 0 136M 9204 4692 S 18.9 0.0 1:21.19 sshd: keyajoshi@notty
2153 root 20 0 4160M 350M 87436 D 17.8 0.3 50h46:34 /usr/lib/systemd/systemd-journald
2270065 keyajoshi 20 0 138M 11376 4688 S 17.3 0.0 7:52.13 sshd: keyajoshi@notty
534104 nickel 20 0 38024 20024 3136 R 16.2 0.0 0:29.02 htop
3630 nsld 20 0 436M 3772 1764 S 9.5 0.0 21h15:39 /usr/sbin/nsld
630825 uscms 20 0 67760 23092 3592 S 8.9 0.0 0:00.16 /cm/shared/apps/slurm/current/bin/scontrol show job
2539706 atsushim 20 0 136M 8704 4640 S 8.4 0.0 1h10:00 sshd: atsushim@notty
764133 x4zou 20 0 136M 8940 4636 S 6.1 0.0 56:57.24 sshd: x4zou@notty
2270208 keyajoshi 20 0 38040 9172 447 htop 3.2.1 - (C) 2004-2019 Hisham Muhammad. (C) 2020-2022 htop dev team.
3653187 keyajoshi 20 0 38040 9184 448 Released under the GNU GPLv2+. See 'man' page for more info.
631364 zhanglab 20 0 54924 11628 489 CPU usage bar: [low/normal/kernel/guest used%]
764197 x4zou 20 0 21796 3380 224 Memory bar: [used/buffers/shared/cache used/total]
260002 vdommes 20 0 332M 280M 818 Swap bar: [used/cache used/total]
1096453 mghazvinia 20 0 6903M 6259M 435 Type and layout of header meters are configurable in the setup screen.
2539790 atsushim 20 0 38456 9400 443 Process state: R: running; S: sleeping; t: traced/stopped; Z: zombie; D: disk sleep

```

```

Tab: switch to next screen tab S-Tab: switch to previous screen tab
Arrows: scroll process list Space: tag process
Digits: incremental PID search c: tag process and its children
F3 /: incremental name search U: untag all processes
F4 \: incremental name filtering F9 k: kill process/tagged processes
F5 t: tree view F7 l: higher priority (root only)
p: toggle program path F8 [: lower priority (+ nice)
m: toggle merged command a: set CPU affinity
Z: pause/resume process updates e: show process environment
u: show processes of a single user i: set IO priority
H: hide/show user process threads l: list open files with lsOF
K: hide/show kernel threads x: list file locks of process
F: cursor follows process s: trace syscalls with strace
+ - *: expand/collapse tree/toggle all w: wrap process command in multiple lines
N P M T: sort by PID, CPU%, MEM% or TIME F2 C S: setup
I: invert sort order F1 h ?: show this help screen
F6 > .: select sort column F10 q: quit

Press any key to return.

```

atop utility

- Tool for monitoring system resources in Linux.
 - System load information at the process level.
 - Default shows cpu usage
 - -m shows memory usage
 - -d shows disc usage
 - -au processes per user

ATOP - login02		2024/03/29		11:20:25		-----										46d1h13m59s elapsed	
PRC	sys	41d08h	user	13d06h	#proc	8233	#trun	5	#tslpi	13e3	#tslpu	1301	#zombie	14	no	procacct	
CPU	sys	555%	user	201%	irq	19%	idle	5552%	wait	73%	steal	0%	guest	0%	curf	2.47GHz	
CPL	avg1	57.77	avg5	42.04	avg15	40.35			csw	273717e6	intr	22369e7			numcpu	64	
MEM	tot	124.9G	free	12.5G	cache	11.0G	dirty	161.3M	buff	111.9M	slab	18.0G	shrss	10.9M	numnode	4	
WFI	tot	12.0G	free	0.0M	swcac	503.6M							vmcom	973.2G	vmlim	74.4G	
PAG	scan	13455e7	steal	1219e8	compact	44e5	numamig	25e8	migrate	65e8	swin	2406411	swout	1171e4	oomkill	0	
DSK	sda	busy	9%	read	16936e3	write	8764e5	discrd	0	MBr/s	0.3	MBw/s	13.8	avio	0.42 ms		
DSK	sdb	busy	0%	read	34600	write	62099	discrd	0	MBr/s	0.0	MBw/s	0.0	avio	0.52 ms		
NFM	/pool13/home	srv	10.22.10	read	6.0T	write	2.7T	nread	9.9T	nwrit	8.4T	dread	0.0K	dwrit	0.0K		
NFM	ome/ccompute	srv	10.22.10	read	6.0T	write	2.7T	nread	9.9T	nwrit	8.4T	dread	0.0K	dwrit	0.0K		
NFC	rpc	462179e3	read	36938e4	write	1717e6	retxmit	21e3	autref	462e6							
NET	transport	tcpi	40540e6	tcpo	10453e7	udpi	7638184	udpo	9461204	tcpao	2930e4	tcppo	2724e3	tcprs	6003e4		
NET	network	ipi	405568e5	ipo	35180e6e5	ipfrw	0	deliv	4055e7			icmpi	1192e3	icmpo	21637		
NET	ensif0n	1%	pcki	84765e6	pcko	82463e6	sp	25 Gbps	si	461 Mbps	so	424 Mbps	erri	0	erro	0	
NET	ib0	0%	pcki	10231e6	pcko	20912e6	sp	100 Gbps	si	36 Mbps	so	134 Mbps	erri	0	erro	0	
FB	lx5_0/1	0%	pcki	21750e7	pcko	11141e7	sp	100 Gbps	si	802 Mbps	so	348 Mbps	lanes	2			
Number of variable resources limited to fit in this window																	
PID	SYSCPU	USRCPU	RDELAY	VGROW	RGROW	RUID	EUID	ST	EXC	THR	S	CPUNR	CPU	CMD	1/344		
3635	6d03h	29h48m	84m47s	9.9G	84.2M	root	root	N-	-	3	S	63	16%	rsyslogd			
2153	28h10m	22h37m	69m06s	8.6G	360.4M	root	root	N-	-	1	S	63	5%	systemd-journ			
516	33h02m	0.00s	22m26s	0B	0B	root	root	N-	-	1	R	48	3%	kswapd3			
515	25h49m	0.00s	15m20s	0B	0B	root	root	N-	-	1	S	36	2%	kswapd2			
1	17h32m	4h56m	62m46s	263.1M	33.6M	root	root	N-	-	1	S	33	2%	systemd			
3630	12h57m	8h19m	8h42m	436.8M	3.7M	nslcd	nslcd	N-	-	6	S	51	2%	nslcd			
4073	17h43m	2m40s	18.04s	1.2G	6.9M	root	root	N-	-	8	S	10	2%	automount			
10153	7h58m	8h12m	2h14m	151.0M	5.2M	root	root	N-	-	3	S	37	1%	auditd			
2990	2h04m	12h39m	2h32m	292.5M	137.5M	dbus	dbus	N-	-	1	S	34	1%	dbus-daemon			
514	13h26m	0.00s	8m15s	0B	0B	root	root	N-	-	1	S	16	1%	kswapd1			
513	8h52m	0.00s	5m51s	0B	0B	root	root	N-	-	1	S	0	1%	kswapd0			
10155	4h53m	1h53m	27m56s	10.9M	664.0K	root	root	N-	-	1	S	63	1%	audisp-syslog			
2974	2h31m	1h44m	74m17s	117.1M	33.7M	root	root	N-	-	1	S	3	0%	systemd-logind			
10577	3h37m	0.00s	20m49s	0B	0B	root	root	N-	-	1	S	27	0%	kiblnsd_sd_01_0			
10575	3h37m	0.00s	20m48s	0B	0B	root	root	N-	-	1	S	29	0%	kiblnsd_sd_01_0			
10576	3h37m	0.02s	20m49s	0B	0B	root	root	N-	-	1	S	28	0%	kiblnsd_sd_01_0			
10574	3h37m	0.01s	20m45s	0B	0B	root	root	N-	-	1	S	30	0%	kiblnsd_sd_01_0			
3268574	3h18m	15m42s	37m59s	546.6M	12.4M	root	root	N-	-	5	S	40	0%	udisksd			
3539973	32m17s	2h19m	70.90s	135.9M	8.3M	x4zou	x4zou	N-	-	1	S	53	0%	sshd			
464	2h28m	0.00s	19m55s	0B	0B	root	root	N-	-	1	S	39	0%	kauditd			
10726	2h27m	0.00s	11m55s	0B	0B	root	root	N-	-	1	I	59	0%	ptlrpcd_03_08			
10720	2h27m	0.00s	11m53s	0B	0B	root	root	N-	-	1	I	49	0%	ptlrpcd_03_02			
10730	2h27m	0.00s	11m54s	0B	0B	root	root	N-	-	1	I	50	0%	ptlrpcd_03_12			
10718	2h27m	0.00s	11m54s	0B	0B	root	root	N-	-	1	I	59	0%	ptlrpcd_03_00			

mpstat

- Report processors related statistics
- Usage: mpstat [options] [<interval> [<count>]]

```
login02 ~]$ mpstat -A
Linux 4.18.0-477.15.1.el8_8.x86_64 (login02)      03/29/2024      _x86_64_      (64 CPU)

08:19:33 AM  CPU      %usr   %nice    %sys %iowait    %irq   %soft  %steal  %guest  %gnice   %idle
08:19:33 AM  all       3.07    0.06    8.66    1.14    0.16    0.13    0.00    0.00    0.00   86.77
08:19:33 AM    0       3.97    0.14   10.38    0.90    0.19    0.08    0.00    0.00    0.00   84.34
08:19:33 AM    1       2.78    0.13    9.61    1.00    0.16    0.07    0.00    0.00    0.00   86.25
08:19:33 AM    2       2.36    0.02    8.81    0.95    0.15    0.06    0.00    0.00    0.00   87.65
08:19:33 AM    3       2.39    0.02    8.20    0.87    0.14    0.05    0.00    0.00    0.00   88.33
08:19:33 AM    4       2.96    0.14    9.57    0.64    0.15    0.05    0.00    0.00    0.00   86.48
.
.
.
```

sar

sar – (system activity report)

- Collect, report, or save system activity information
- CPU usage(default) –u (similar to iosatt)
- Memory usage -r
- File system usage -F

```
[nickel@login02 ~]$ sar
Linux 4.18.0-477.15.1.el8_8.x86_64 (login02)      03/29/2024      _x86_64_      (64 CPU)

12:00:02 AM      CPU      %user      %nice      %system      %iowait      %steal      %idle
12:10:02 AM      all       2.26       0.00       9.51       2.59       0.00      85.64
12:20:01 AM      all       2.10       0.00       9.37       3.73       0.00      84.80
12:30:01 AM      all       2.19       0.00       9.12       2.30       0.00      86.39
12:40:01 AM      all       2.19       0.00       9.36       1.91       0.00      86.54
12:50:02 AM      all       1.96       0.00       8.81       1.65       0.00      87.57
01:00:03 AM      all       1.93       0.00       9.55       3.20       0.00      85.32
01:10:03 AM      all       2.68       0.00      13.83       1.88       0.00      81.60
01:20:01 AM      all       2.17       0.00       8.71       1.67       0.00      87.46
01:30:00 AM      all       3.00       0.00       9.59       1.65       0.00      85.75
01:40:03 AM      all       2.67       0.00       9.42       1.69       0.00      86.21
01:50:02 AM      all       2.12       0.00       8.77       3.54       0.00      85.57
02:00:03 AM      all       3.60       0.00       8.83       1.70       0.00      85.87
```

free

- Evaluate memory Usage statics of the Linux operating system
- Usage: free [options]

```
@login02 ~]$ free -h
```

	total	used	free	shared	buff/cache	available
Mem:	124Gi	89Gi	10Gi	72Mi	25Gi	34Gi
Swap:	11Gi	11Gi	0B			

ps

- Returns the users current processes
- `man ps` - will provide all the options
- `-e`: returns every process on system

```
login01 gpu]$ ps -aux | head -10
USER      PID %CPU %MEM    VSZ   RSS TTY      STAT START   TIME COMMAND
root         1  1.1  0.0 253436 22956 ?        Ss   Mar21 225:43 /usr/lib/systemd/systemd
root         2  0.0  0.0    0     0 ?         S   Mar21   3:44 [kthreadd]
user1 1790480  0.0  0.0  12736  3156 ?        Ss   Mar29   0:00 /bin/bash /usr/libexec/dbus-1/dbus-kill-process-with-session
1790477
user1 1790487  0.0  0.0 149644  9224 ?        Sl   Mar29   0:29 gio monitor -f /run/systemd/sessions/76616 /tmp/dbus-
session-monitor.NGdpD
user2 1794352  0.0  0.0  12884  2944 pts/1    T   Mar26   0:00 /bin/bash /home/jitsuk/gene-dev/tools/runassign 01
user3 1797357  0.0  0.0 136692 10396 ?        Ss   Apr02   0:00 sshd: cyang2 [priv]
user4 1797526  0.2  0.0  98640 16760 ?        Ss   Mar25  28:10 /usr/lib/systemd/systemd --user
user5 1797539  0.0  0.0  307424 1648 ?         S   Mar25   0:00 (sd-
```

Getting GPU information

If you're using GPU nodes, you can use nvidia-smi (NVIDIA System Management Interface program) to get GPU information (type, count, etc.)

When running in gpu-shared partition, will only see the GPUs you had requested (typically one)

```
$ nvidia-smi
Mon Aug  2 10:16:29 2021

+-----+
| NVIDIA-SMI 460.32.03      Driver Version: 460.32.03      CUDA Version: 11.2      |
+-----+-----+-----+-----+
| GPU  Name           Persistence-M| Bus-Id        Disp.A | Volatile Uncorr. ECC |
| Fan  Temp  Perf  Pwr:Usage/Cap|      Memory-Usage | GPU-Util  Compute M. |
|                               |                  |              MIG M. |
+=====+=====+=====+=====+
|    0  Tesla V100-SXM2...  On      | 00000000:18:00:0 Off  |              0      |
| N/A   40C    P0     67W / 300W |      0MiB / 32510MiB |      0%      Default |
|                               |                  |              N/A     |
+-----+-----+-----+-----+

+-----+
| Processes:                                                       |
| GPU  GI    CI          PID    Type    Process name                        GPU Memory |
|          ID    ID                                   Usage     |
+=====+
| No running processes found                                         |
+-----+
```

Obtaining job hardware usage info (seff and sacct)

- seff
 - Slurm Job Efficiency Report
 - seff <<completed_jobid>>
- sacct:
 - Usage: sacct [options]
 - Slurm command to evaluate pending, running, and finished jobs
- Unfortunately, no details on GPU usage, just cpu and memory

In Conclusion

- Machine Information overkill?
- Use the correct tools
 - System tools
 - Scheduler tools
- Avoid wasting resources and money

Machine info – cheat sheet

File or command	Information provided
<code>less /proc/cpuinfo</code> or <code>lscpu</code>	CPU specs
<code>less /proc/meminfo</code>	Memory specs and usage
<code>nvidia-smi</code>	GPU specs and usage
<code>cd /sys/devices/system/cpu/cpu0/cache</code> ... then look at directory contents	Cache configuration
<code>less /proc/scsi/scsi</code> or <code>lsscsi</code>	Peripherals (e.g. SSDs)
<code>less /etc/mtab</code>	Mounted file systems
<code>df -h</code>	File system usage (readable format)
<code>du -sh</code>	Disk usage (human readable)
<code>/sbin/ip link</code>	Networking information
<code>uname -a</code>	OS information
<code>less /etc/centos-release</code>	Centos version

Machine and job monitoring – cheat sheet

File or command	Information provided
top	Monitor process activity
atop	Monitor process activity
htop	Monitor process activity
sar	System utilization
free	Memory utilization
ps	Process
mpstat	Processors
seff	Slurm Job efficiency data
sacct	https://slurm.schedmd.com/sacct.html

Review and helpful links

- Slurm
 - <https://slurm.schedmd.com/sacct.html>
- seff and sacct:
 - https://csc-training.github.io/csc-env-eff/handson/batch_resources/tutorial_sacct_and_seff.html
- Training Catalog
 - https://www.sdsc.edu/education_and_training/training_hpc.html#catalog
- And do not forget the man pages



Questions

- consult@sdsc.edu
- <https://support.access-ci.org/>

