

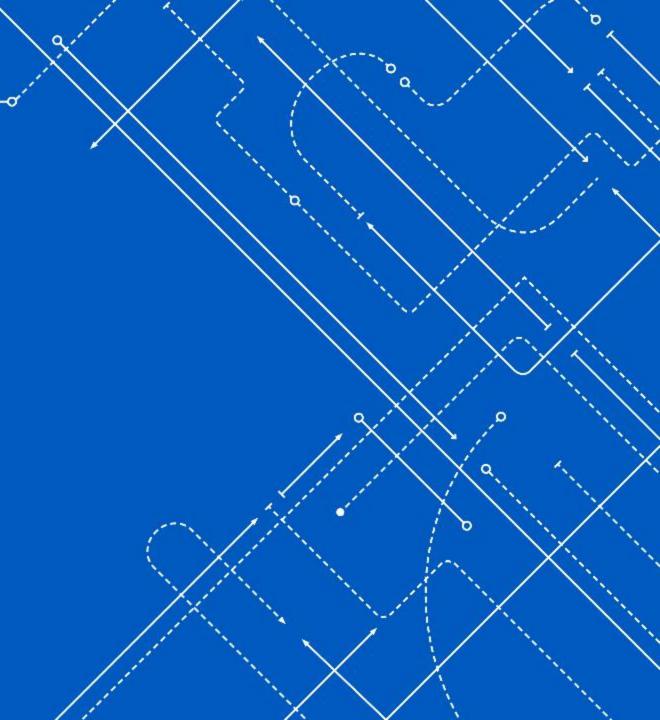
L. Shawn Matott, Ph.D.

UB Center for Computational Research

University at Buffalo The State University of New York

Computational Scientist

OVERVIEW



Workshop Objectives

- To learn some basic concepts of high performance computing (HPC)
- To learn about the UB Center for Computational Research (CCR)
- To learn how to use the CCR On Demand portal



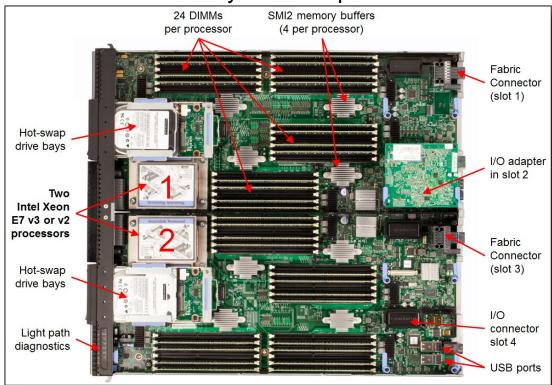
source: UB Photo Database

BASIC HPC CONCEPTS

Some Basic HPC Terminology

- Compute Node a physical self-contained compute unit consisting of an operating system (OS), processors, memory, storage and networking interfaces. Also known as a "blade". Compute nodes can contain multiple processors.
- Processor a computing chip that can contain multiple processing cores.
- Core a single processing unit within a processor.
 These are what ultimately perform computations.
 - Cores on the same compute node must share the node resources (e.g. memory, storage, and networking).
 - Within software, cores are often referred to as cpus or processors

Anatomy of a compute node



Source: lenovopress.com

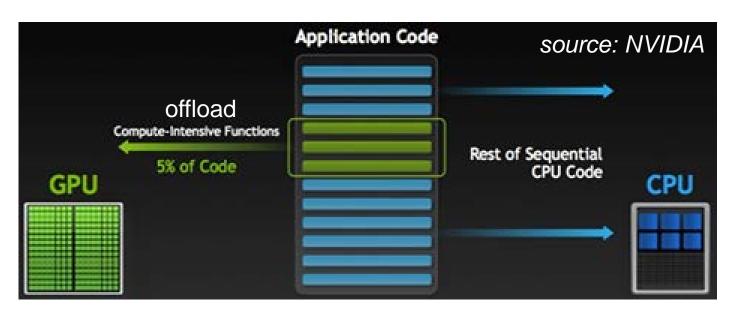


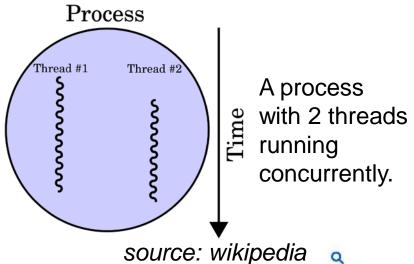
- Node Rack a cabinet that stores a stack of interconnected nodes and other compute equipment
- Compute Cluster a set of interconnected racks.
- Shared Memory Computing an HPC application where each core (all residing on the same node) uses a shared block of memory.
- Distributed Memory Computing an HPC application where each core (possibly residing on different nodes) uses separate blocks of memory and work is coordinated using message passing.

Two Rows of Node Racks at UB CCR



- *Thread* short for "thread of execution", a sequence of instructions that is run concurrently with other threads of the same process. Threads within a process can share memory.
- Accelerator a compute device that a connected processor can use to offload certain computations.





UB CCR: HISTORY AND MISSION

What is UB CCR?

- CCR provides UB researchers and affiliated partners, including industry, with access to advanced computing resources.
 - Academic, Industrial, and Faculty Compute Nodes
 - High Performance Storage and Networking
 - Visualization and Cloud Computing Resources

History

- 1999 UB CCR founded, 60 GFLOPS peak performance
- 2006 UB CCR relocated from North Campus to downtown Medical Campus
- 2007 Major Upgrades, 13 TFLOPS, 30 TB Storage
- 2010 Major Upgrades, 170 TFLOPS, 500 TB Storage, GPU Nodes
- 2014 Add Industry Cluster, 250 TFLOPS
- 2015 Expand Storage to 4.2 PB
- 2016 to 2018 Expand Faculty Clusters, 550 TFLOPS
- 2018 Expand Industry and Academic Clusters, 1.3 PFLOPS, new GPU Nodes

Mission

- Enable research and scholarship among UB faculty
- Provide hi-tech workforce training
- Foster economic development and job creation among area industries



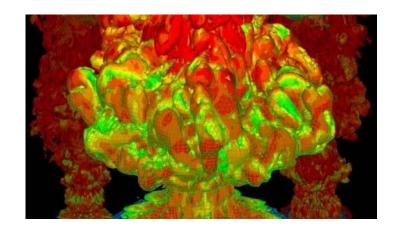




Research

- UB CCR serves all decanal units at UB
 - o 565 total users in 2017
 - 1200 publications since 2016
- No cost for faculty groups to use CCR compute resources
- Online account requests:
 <u>buffalo.edu/ccr/support/ccr-help/accounts.html</u>
- Research topics (not comprehensive!):
 - Precision Agriculture
 - Bioinformatics
 - Fluid Dynamics
 - Molecular Modeling
 - Stormwater Management

- Computer Vision
- Computational Chemistry
- Crystallography
- Volcanology
- Many other areas!





Training

- UB CCR provides free training to faculty and students on the use of its compute resources
 - By request or drop-in at our Bell Hall satellite office
- Helpdesk
 - Online ticket system for requesting technical support
 - Also provides searchable self-help knowledgebase
- Other Outreach Activities
 - Tours of the facility
 - VIDIA Virtual Infrastructure for Data Intensive Analysis
 - Summer Workshops for High School Students
 - STEAM activities at local K-12 schools
 - ICDS/CDSE "short course" modules and workshops
 - Guest lectures for numerous UB courses
 - Hackathon support (e.g. BlockChain, TopCoder, etc.)





Solutions / CCR Services / High Performance Cluster Computing



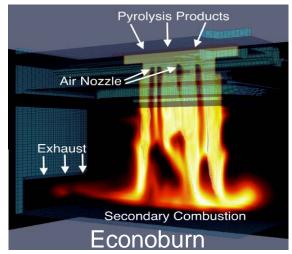


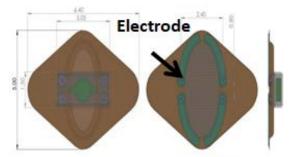
Economic Development

- UB CCR provides local industry with access to a 72 TFLOP (3,456 processors) cluster
- In process of expanding to include additional processors and a batch of NVIDIA Volta V100 GPU nodes
 - Will allow industry to leverage machine learning and deep neural networks
- Since 2013, CCR has assisted over 24 companies
 - o 68 jobs created, 27 retained, \$52 million in economic impacts (savings, investment, revenue, etc.)



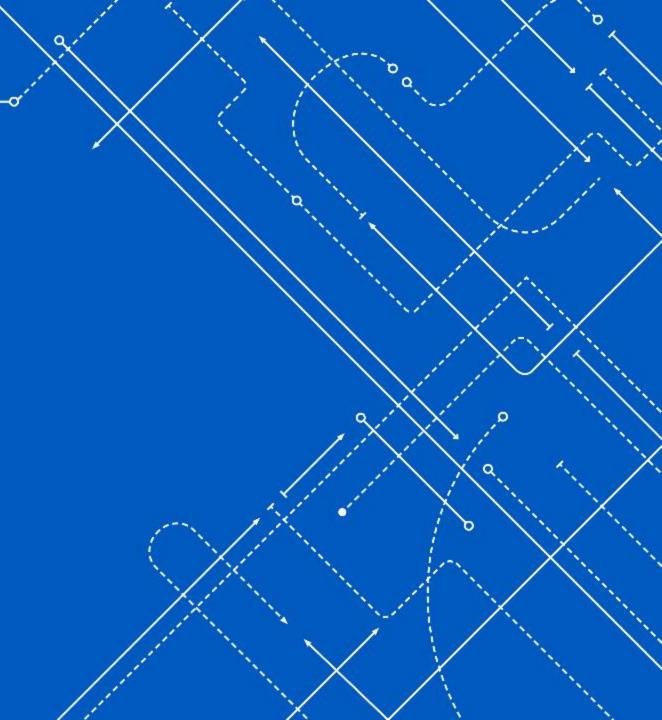








UB CCR: COMPUTING RESOURCES



Compute Resources (as of April 2018)

- Total Compute Capability:
 - 1,373 Compute Nodes
 - 20,044 Processing Cores
- Academic Compute Capability:
 - 684 Compute Nodes
 - 7,828 Processing Cores
- Faculty* Compute Capability:
 - 471 Compute Nodes
 - 8,760 Processing Cores
- Industry* Compute Capability:
 - 216 Compute Nodes
 - o 3,456 Processing Cores

Configuration of UB CCR's Academic Computing Cluster (ub-hpc)

Type of Node	Approximate # of Nodes	# Cores per Node	Clock Rate	RAM	Network*	SLURM TAGS	Local /scratch
Compute	32	16	2.20GHz	128GB	Infiniband (M)	IB CPU- E5-2660	773GB
Compute	372	12	2.40GHz	48GB	Infiniband (QL)	IB CPU- E5645	884GB
Compute (Dell)	128	8	2.13GHz	24GB	Infiniband (M)	IB CPU- L5630	268GB
Compute (IBM)	128	8	2.27GHz	24GB	Infiniband (M)	IB CPU- L5520	268GB
High Memory Compute (INTEL CPUs)	8	32	2.13GHz	256GB	Infiniband (QL)	IB CPU- E7-4830	3.1TB
High Memory Compute (AMD CPUs)	8	32	2.20GHz	256GB	Infiniband (QL)	IB CPU- 6132HE	3.1TB
High Memory Compute (INTEL CPUs)	2	32	2.13GHz	512GB	Infiniband (QL)	IB CPU- E7-4830	3.1TB
GPU Compute	26**	12	2.67GHz	48GB	Infiniband (M)	IB CPU- X5650	0.9TB

^{*} NETWORK: Infiniband (M) = Mellanox and Infiniband (QL) = Q-Logic

^{* -} Unused Faculty and Industry cylces can be scavenged via pre-emptive scheduling

Data Storage

- Unix/Linux File Permissions
 - All nodes run CentOS Linux
- 5 GB of user space
 - backed up daily
- up to 1 TB of projects space
 - backed up daily
- 'unlimited' scratch space
 - deleted after 21 days
 - NOT backed up





user

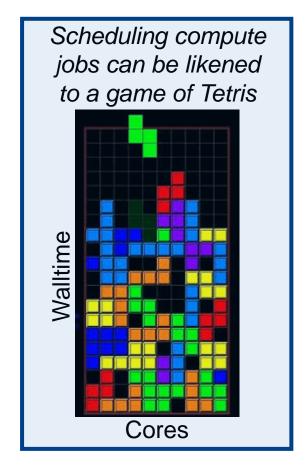






Batch Computing

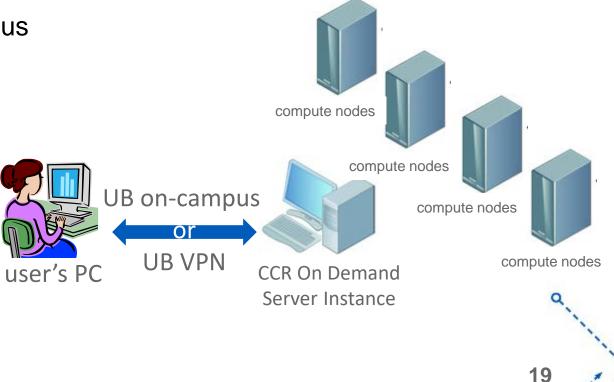
- UB CCR's compute resources are organized as a batch computing system
 - Users submit compute requests (jobs) to the scheduler software (SLURM)
 - The scheduler examines requested resources (cores, memory, walltime, etc.) and determines when jobs should run and on what nodes
 - When jobs finish early, "holes" form in the schedule. The scheduler "backfills" jobs to fill these holes, kind of like a game of Tetris.



scheduler

Cluster Access

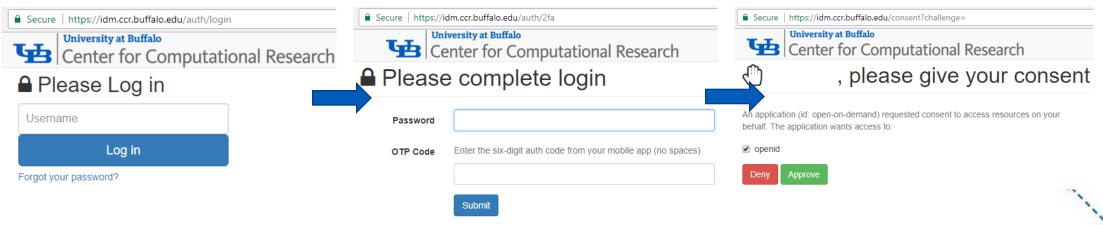
- UB CCR provides secure remote access to its resources
 - Hardware and software firewalls
 - Virtual Private Network for Off-Campus
 - Secure Shell (ssh) login node
- Connect to "CCR On Demand" to:
 - Request compute nodes
 - Load software
 - Launch batch or interactive jobs
 - Monitor job status
 - Transfer and edit files



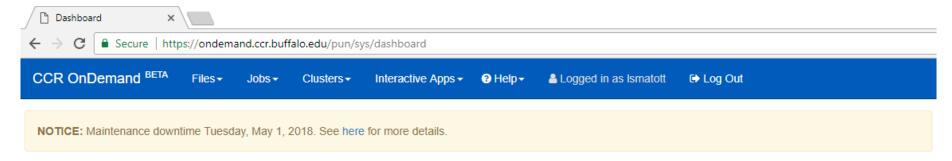
UB CCR ON DEMAND

Login to CCR On Demand

- If off-campus, must first connect to VPN
 - UB CIT provides "AnyConnect" software (see <u>ubit.buffalo.edu</u>)
- Point a web-browser to ondemand.ccr.buffalo.edu
 - Login using your UB CCR username and password
 - Answer a security question or provide a two-factor code
 - Give consent for resource access



On Demand Dashboard



OnDemand provides an integrated, single access point for CCR's HPC resources

Users can transfer files, access a shell environment on the cluster front-end login server, launch interactive and remote visualization jobs, and monitor jobs all without installing any client software or web plug-ins. Access these features using the menus at the top of this page. Note that many of the apps will launch in a new tab or new browser window but the dashboard will remain open in the original window.

For documentation on the CCR OnDemand portal, please read these articles:

- How to use CCR OnDemand
- Remote Visualization in OnDemand

Additional CCR web resources

- CCR website
- · Knowledgebase and help portal
- · Identity management portal password, one-time tokens and SSH key management
- Coldfront subscription management & account request portal for PIs
- · XDMoD job statistics and usage information
- · WebMO Chemistry application portal



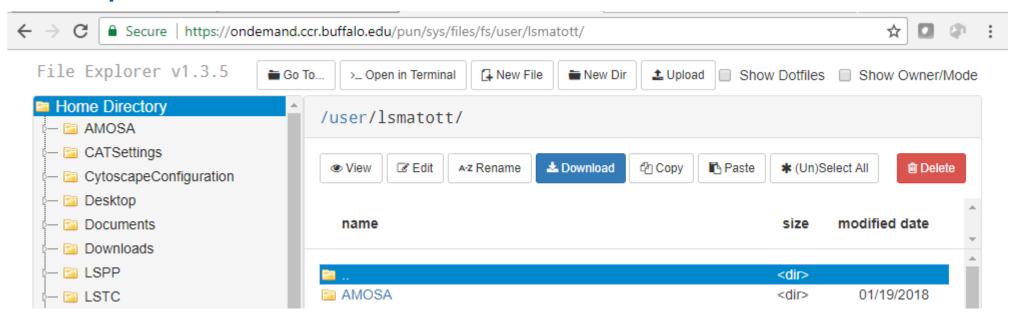
On Demand Dashboard



- Files
 - Launches "File Explorer"
 - Navigate storage
 - Transfer files to/from PC
 - Edit, Delete, Copy, & Rename files
- Jobs
 - Check status of active jobs
 - Compose and submit new job scripts (under development)

- Clusters
 - Command line shell terminal
- Interactive Apps
 - Remote Desktop Sessions (for GUI applications)

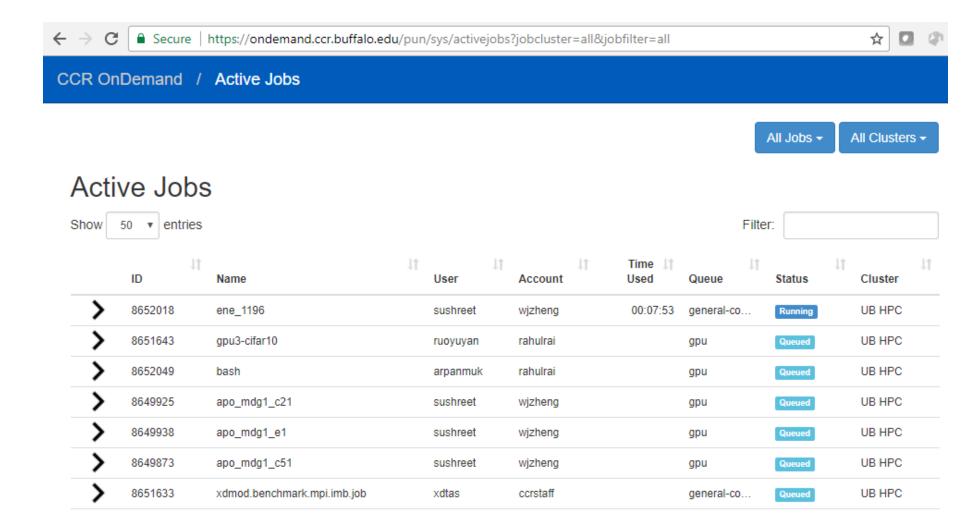
File Explorer



- "Go To ..." Navigate to project or scratch areas
- "Open in Terminal" Open a command line shell in the currently viewed directory
- "Upload" Transfer files from PC to CCR storage
- "Download" Transfer files from CCR storage to PC



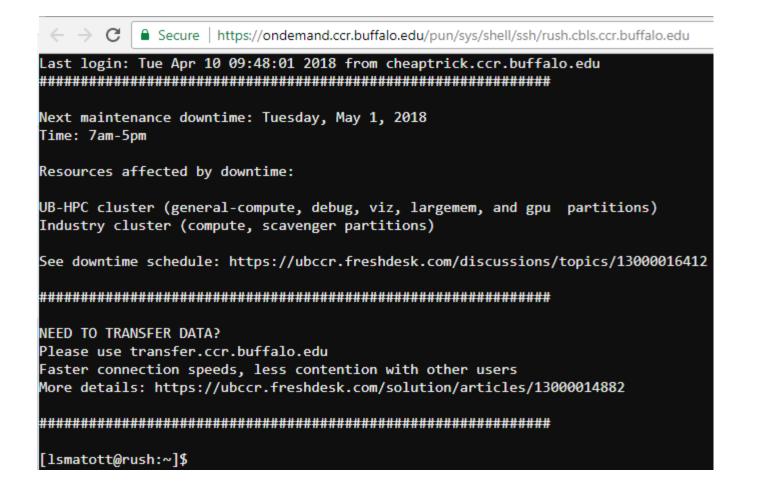
Active Jobs



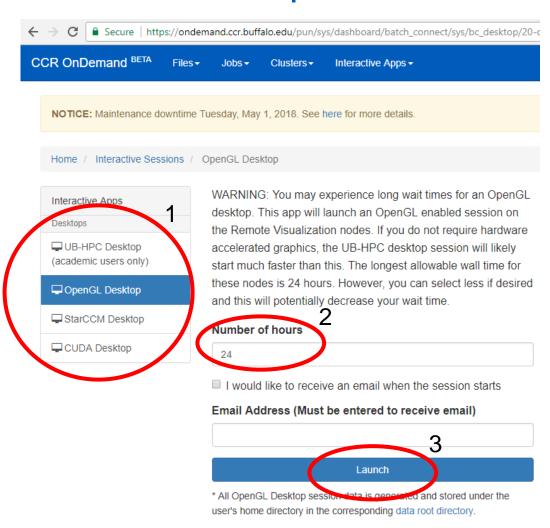


0

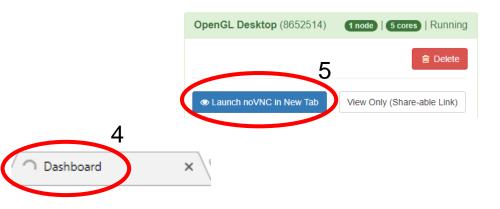
Cluster Shell



Remote Dekstop

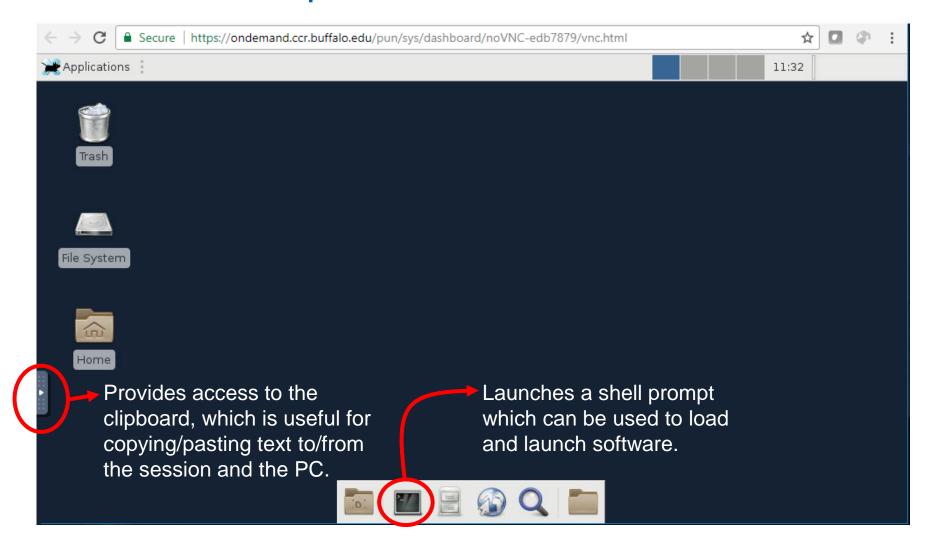


- 1. Choose the type of desktop.
- 2. Choose the duration of session.
- 3. Click Launch.
- 4. Wait for Session to Start.
- 5. When running, click "Launch noVNC in New Tab".





Remote Dekstop



UB CCR: SOFTWARE MODULES

Software Modules

- CCR supports 300+ software packages
- Installed as "modules"
- Important module commands:

```
module avail - what's installed?
module load - load module
```

Example (try from On Demand):

```
$ rstudio
-bash: rstudio: command not found
$ module load R
$ rstudio
(successfully launches rstudio)
```





Engineering	Machine Learning	Hadoop/MapReduce	Math/Physics	Image Processing	Utilities
ABAQUS	Caffe	Hadoop	BandUP	EMAN2	7zip
ANSYS	Caffe2	HBase	CrystFEL	FFMPEG	Ant
AVL FIRE	Keras	Hive	FEniCS	OpenCV	СМАКЕ
CFX	MikeNet	Pig	GAP	OpenEXR	DMTCP
COMSOL	TensorFlow	Spark	Gnuid	OpenSlide	gnu-parallel
CPLEX	Theano	ZooKeeper	Gurobi	Environmental	HDF/HDF5
DAKOTA	Torch	GPU Programming	kmos	Cantera	KchmViewer
iso2mesh	Data Analytics	CUDA	ROOT	GRASS	lxml
LSDYNA	edgar	Intel-OpenCL	VMTK	HYSPLIT	Mono
NETGEN	Jags	OpenCL	Visualization	LANDIS-II	netCDF, NCO, CDO
OpenFOAM	MCL	PyCUDA	de.caff	MODFLOW	pbzip2
OpenSees	ms	Editors	NCL	NaSt3DGP	pigz
pyFormex	R	Emacs	ParaView	puffin	TauBench
QUCS	SAS	Idle	TecPlot	TELEMAC	tmux
SALOME	Stan	NetBeans	VTK	TITAN	xclip
StarCCM	STATA	Vim	xmgrace	WRF-WPS	xfig

Molecular Dynamics	Quantum Chemistry	Libraries (cont.)	Programming	Programming (cont.)
Amber , AmberTools	ABINIT	Elemental	Cilk	Julia
AutoDock	Dalton	GLPK	DDT/MAP	Mathematica
CHARMM	Elk	GMP	DMD	MATLAB
CP2k	GAMESS	Intel: MKL, MPI, TBB	gcc , gdb	Maven
GROMACS	GPAW	libgd	golang	Mercurial
LAMMPS	NWChem	libgpuarray	google-api	mpiP
MDAnalysis	OpenBabel	libgtextutils	IDL	Octave
MMTSB	Orca	libmatheval	Intel-Advisor	PAPI
MODELLER	Q-Chem	libmesh	Intel-CLCK	Perl
NAMD	QUANTUM ESPRESSO	mvapich2	Intel-Compiler	PGI
OpenMM	Siesta	OpenBLAS	Intel-DAAL	PP (Parallel Python)
PLUMED	Libraries	OpenMPI	Intel-Inspector	Python
pyRosetta , Rosetta	Armadillo	PETSc , SLEPc	Intel-IPP	Rdesktop
Schrodinger	ARPACK	scikit-tensor	Intel-ITAC	Scala
supercell	BOOST	TetGen	Intel-vTune	SCons
VMD	CGAL	Tioga	Java	Valgrind

Bioinformatics / Genomics (130 modules)									
AdmixTools	BedTools	CNVnator	fusioncatcher	Kallisto	miRDeep2	PARE	qctool	seqtk	TopHat
Admixture	bio3d	Cromwell	gatk	Kraken	MixMapper	PATSER	QIIME	SnpEff	Tracer
AlignGraph	bioconda	Cufflinks	GCTA	lobSTR	mothur	pBlat	Randfold	SnpSift	TreeMix
AML	bioperl	cutadapt	genipe	lumpy-sv	MrBayes	picard	RAxML	Snptest	TrimGalore
ANGSD	BLASR	DADI	GMAP	MACS	MSMC	Pindel	RELION	SomaticSniper	Trinity
ANNOVAR	BLAST	deepTools	GTOOL	MACS2	MuMmer	PLINK	RSem	speedseq	VarDict
BadiRate	Blat	deFuse	HLAreporter	mapDamage	NCBI	PLINK/SEQ	RSeQC	Squid	VCFtools
Bambino	Bowtie	dice	HOMER	MEME	NgsAdmix	Polysolver	Sailfish	SRA Toolkit	VEGAS2 / VEGAS2v02
bam-readcount	Bowtie2	EIGENSOFT	HTSeq	MERLIN	ngsPopGen	Preplot	Salmon	Strelka	Velvet
BamTools	BreakDancer	EMBOSS	HTSlib	METAL	NgsRelate	Preseq	SAMtools	StringTie	VEP
bcl2fastq	BWA	EPACTS	IgBLAST	MetAMOS	nucleo	ProDy	schmutzi	SurvivalGWAS_SV	ViennaRNA
Beagle	CDK	FastQC	Jellyfish	Minimac3	OpenGATE	psmc	SeqAn	svtyper	WDLtool
BEAST2	CGHub	FASTX	JointSNVMix	MIRA	PAML	PyBedTools	SEQLinkage	tabix	XWAS

UB CCR: SUBMITTING BATCH JOBS

About the Scheduler

- SLURM
 - Simple Linux Utility for Resource Management
- Useful scheduler commands (issue these from an On Demand "Clusters" shell prompt):
 - sbatch submit a job script
 - o squeue/sqstat check the status of a job
 - scancel delete a job
 - o snodes show node info and status
 - sranks show job priorities
 - stimes show when jobs are expected to start



Decomposing An Example Job Script

```
#!/bin/sh
#SBATCH --partition=debug
#SBATCH --time=00:15:00
#SBATCH --nodes=2
#SBATCH --ntasks-per-node=12
##SBATCH --mem=48000
#SBATCH --constraint=IB&CPU-E5645
#SBATCH --job-name="hello_test"
#SBATCH --output=test-mpi-debug-%j.out
#SBATCH --error=test-mpi-debug-%j.err
#SBATCH --mail-user=cdc@buffalo.edu
#SBATCH --mail-type=ALL
```

Job scripts begin with a list of SBATCH directives. These directives inform the scheduler about the resources needed for the job along with who to contact when the job completes and what to do with output that would normally be printed on the screen.

The next step in the job script is to configure the desired computing environment by loading required modules and setting relevant variables.

We typically also print (echo) various useful SLURM variables, like the job id, the nodelist, and the working directory. These can come in handy if there is a need to troubleshoot a completed job.

Decomposing An Example Job Script

```
#export I_MPI_DEBUG=4
export I_MPI_PMI_LIBRARY=/usr/lib64/libpmi.so
srun ./helloworld
#
echo "All Done!"
```

The final step in the job script is to launch the desired application. How an application is launched depends on the application. **srun** is the preferred launcher for multi-node applications. For single node applications, just launch the application directly (e.g. **python** ./helloworld). Most applications also require various command line arguments.

CCR provides example job scripts for various applications. These are located in the /util/academic/slurm-scripts folder. You can copy the example script of your choice and edit to suit your needs.

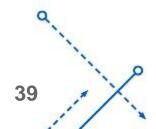
An Example Python Multiprocessing Job Script

```
#!/bin/bash
#SBATCH --job-name=pi mp
#SBATCH --output=pi_mp.out
#SBATCH --error=pi mp.err
#SBATCH --mail-user=your user name@buffalo.edu
#SBATCH --mail-type=END
#SBATCH --time=00:10:00
#SBATCH --nodes=1
#SBATCH --cpus-per-task=1
#SBATCH --exclusive
#SBATCH --partition=general-compute
#SBATCH --constraint=CPU-E5645
#SBATCH --mem=48000
#SBATCH --tasks-per-node=12
module load python/anaconda
ulimit -s unlimited
python pi mp.py
```

This script requests:

1 node, 12 processors, and 48 GB RAM

It loads the python/anaconda module and then launches a python script. The pi_mp.py script contains some Python code that takes advantage of the multiprocessing module to use the 12 requested processors.



An Example MPI for Python Job Script

```
#!/bin/bash
#SBATCH --job-name=mpi4py
                                                    This script requests:
#SBATCH --output=mpi4py.out
#SBATCH --error=mpi4py.err
                                                              2 nodes, 24 processors, and 48 GB RAM per node
#SBATCH --mail-user=your user name@buffalo.edu
#SBATCH --mail-type=END
#SBATCH --time=00:10:00
                                                    It loads the python/anaconda and intel-mpi modules and then
#SBATCH --nodes=2
#SBATCH --cpus-per-task=1
                                                    adds the location of the mpi4py installation to the Python
#SBATCH --exclusive
                                                    path. Finally, it sets up the intel-mpi environment and
#SBATCH --partition=general-compute
#SBATCH --constraint=IB&CPU-E5645
                                                    launches a python script using the "srun" launcher. srun is
#SBATCH --mem=48000
                                                    the preferred MPI launcher (instead of mpirun or mpiexec)
#SBATCH --tasks-per-node=12
                                                    on the CCR system.
# load modules
module load python/anaconda
module load intel-mpi
ulimit -s unlimited
# enable mpi4py module
export PYTHONPATH=/util/academic/python/mpi4py/v2.0.0/lib/python2.7/site-packages:$PYTHONPATH
# launch app
export I MPI FABRICS LIST=tcp
export I MPI DEBUG=4
export I MPI PMI LIBRARY=/usr/lib64/libpmi.so
srun -n $SLURM NPROCS python mpi4py pi.py
```

40

CLOSING REMARKS

What have we learned?

- Some basic HPC terms
- About UB CCR its history and mission
- UB CCR's current resources
- CCR's On Demand access portal
- Working with modules and SLURM

