High Performance Computing at NIU

METIS compute cluster

CRCD Compute Cluster front view.

CRCD team

←NIU.edu A-Z Index Calendar Directory Libraries Make a Gift Info For Quick Links Search

NORTHERN ILLINOIS UNIVERSITY
Division of Research and Innovation Partnerships

Center for Research Computing and Data

About Prospective Users Current Users CRCD at Work

Center for Research Computing Data / About / Leadership

About

- Leadership
- At A Glance
- Documents
- Collaborations

Prospective user?

Request an account.

Leadership

CRCD Helpdesk

Contact our Helpdesk

Research Computing Team

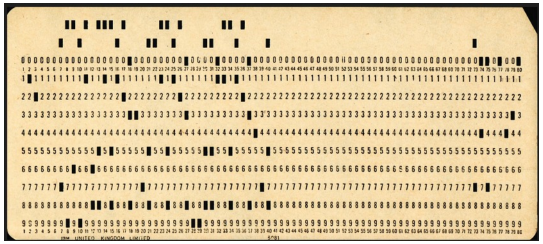
Name	Title	Email
Bela Erdelyi	Professor of Physics, CRCD Director	berdelyi@niu.edu
Sergey A. Uzunyan	Senior Research Scientist, CRCD Director of Science and Engineering	suzunyan@niu.edu
Eric Biletzky	RIPS IT Manager, CRCD Associate	crcdhelpdesk@niu.edu
Andrew Johnson	Senior IT Director	crcdhelpdesk@niu.edu

Small team, efficient HPC management!

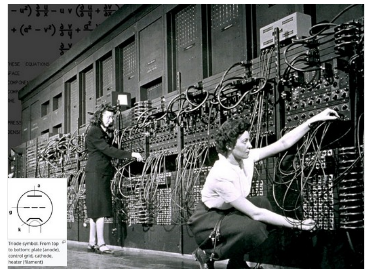
An extra brief history of computers

Binary data format
120 bytes below

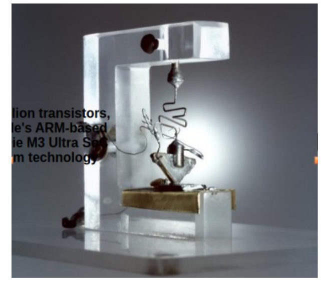
Electronic data processing



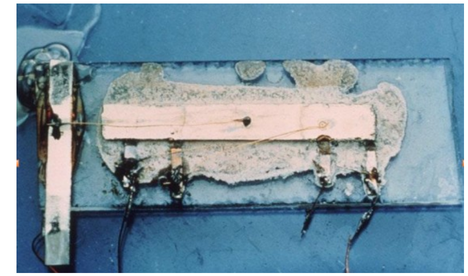
Punch records, year of 1725+



Vaccum tubes logic (~1940-1960)
ENIAC- 1946, ~17000 tubes

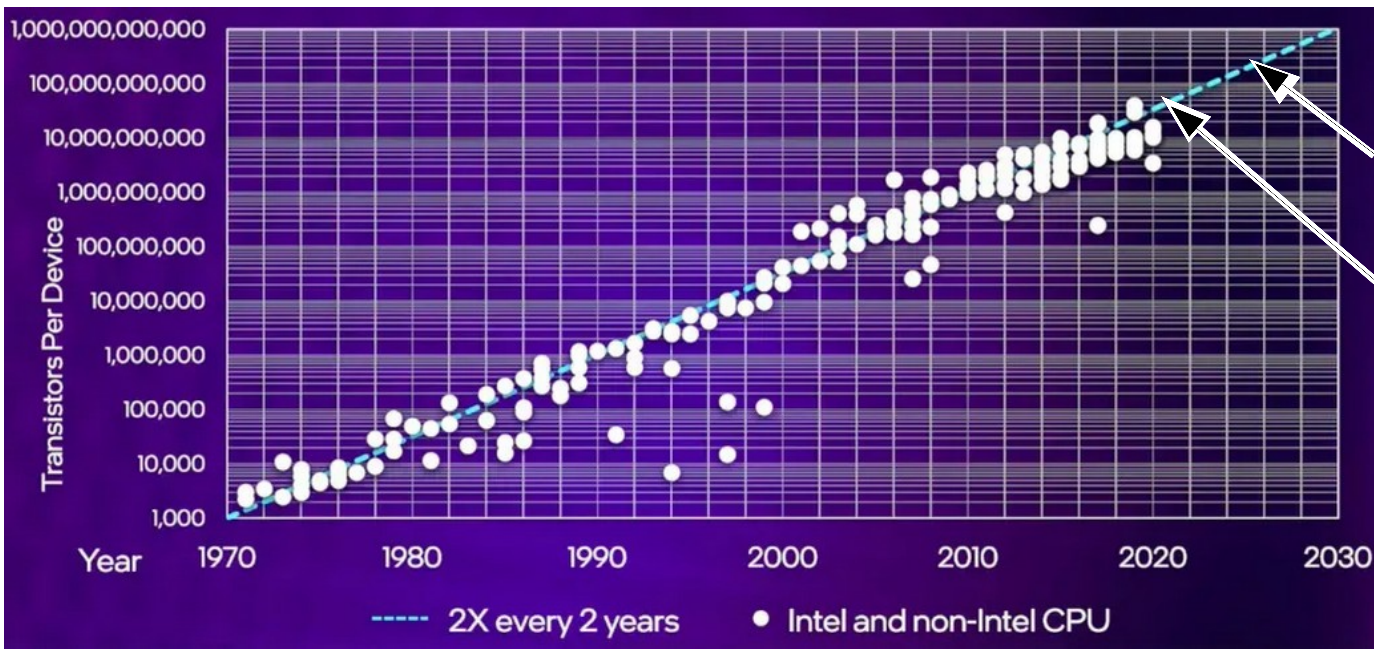


First transistor, 1947



First IC, 3 transistors, 1958

(Gordon) Moore's (Intel co-founder) law: number of transistors in a device will double each 2 years



Today:
184 billion transistors,
in Apple's ARM-based
dual-die M3 Ultra SoC
3nm technology

NIU HPC chips: ~50 billions transistors (A100 GPU)
~32 billions (AMD EPYC CPUs), 7nm

Scale:
~0.15 nm – silicon atom size

A lot of compute power for HPC menu

The origin of modern HPC

HPC cluster @ 1994 == Beowulf cluster: a number of computers (nodes) assembled to run as a single system

HPC cluster today:

- **An assembly of compute nodes designed to run as single system**

- A powerful compute nodes (desktops in a rack friendly form-factor)

+

- Fast interconnect (200 GBit/s)

+

- Large (1 PB+) parallel shared disk system

Becker, Donald J; Sterling, Thomas; Savarese, Daniel; Dorband, John E; Ranawak Udaya A; Packer, Charles V (1995). "BEOWULF: A parallel workstation for scier omputation". Proceedings, International Conference on Parallel Processing. 95.

National Aeronautics and Space Administration

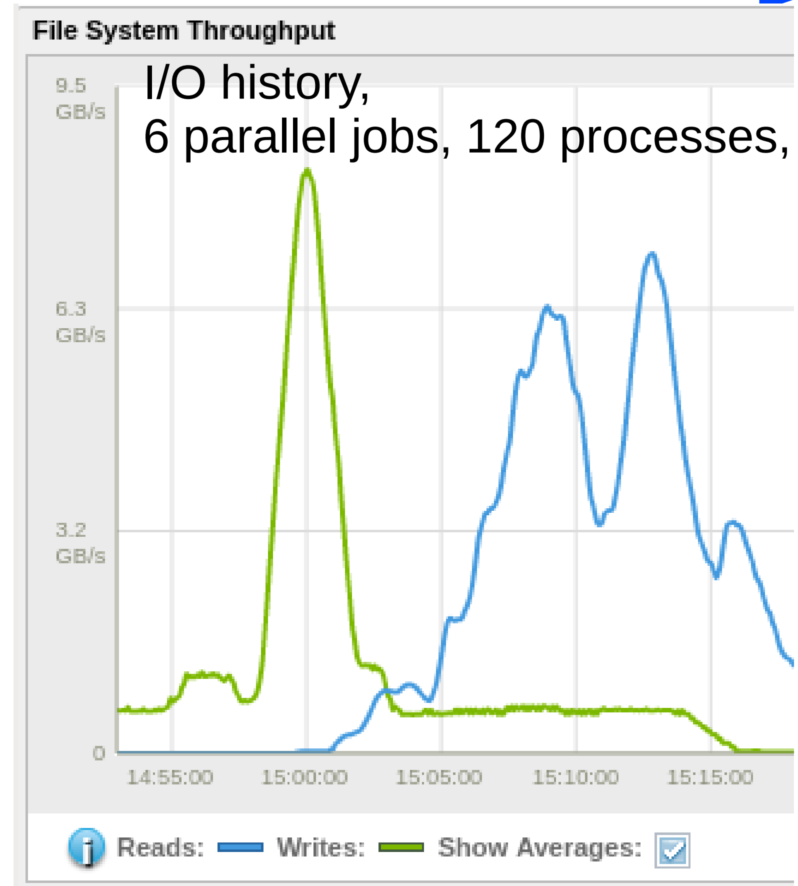
HPC applications

When to use HPC systems at NIU?

- a personal system (laptop or desktop) is “too slow”
- an application benefits from parallelization
- an application needs to be run multiple times
- an application needs extra memory
- an application needs a powerful GPU
- an application requires Linux OS to run

(and a different OS is needed on a personal system)
- fast access to large input data
- the results have to be easy accessible
- convenience: access to shared software libraries

and CRCD team support



Minimal requirements

- **read the CRCD documentation**
- **beginners knowledge of Linux**

Efficiency of shared HPC system depends on users accuracy

Job	User	Account	Class	Remaining time	Used %	Nodes	Average load
344368	z1862058	wheeler1	extra	1:23:09:15	90.6%	12	12.47
344369	z1862058	wheeler1	extra	1:23:10:09	90.6%	12	12.51
344750	z1962831	moisture	extra	10:48:29	94.6%	1	10.38
344800	z1962831	moisture	extra	2:00:28:45	71.5%	1	11.23
344801	z1962831	moisture	extra	3:06:40:44	60.7%	1	11.04
344815	z1962831	moisture	extra	5:10:20:00	34.8%	1	10.44
344816	z1962831	moisture	extra	5:10:20:13	34.8%	1	10.83
344833	kpittman3	climlab	extra	1:23:26:59	47.3%	10	12.48

Jobs: running=8 | all=8
Nodes: used=39 | free=19 | down=2 | all=60

Why we use Linux?

Administrators

- Stability
- Security
- Assess to OpenSource
- Built for development
- Customizable
- Supported by the hardware vendors
- Easy to administer

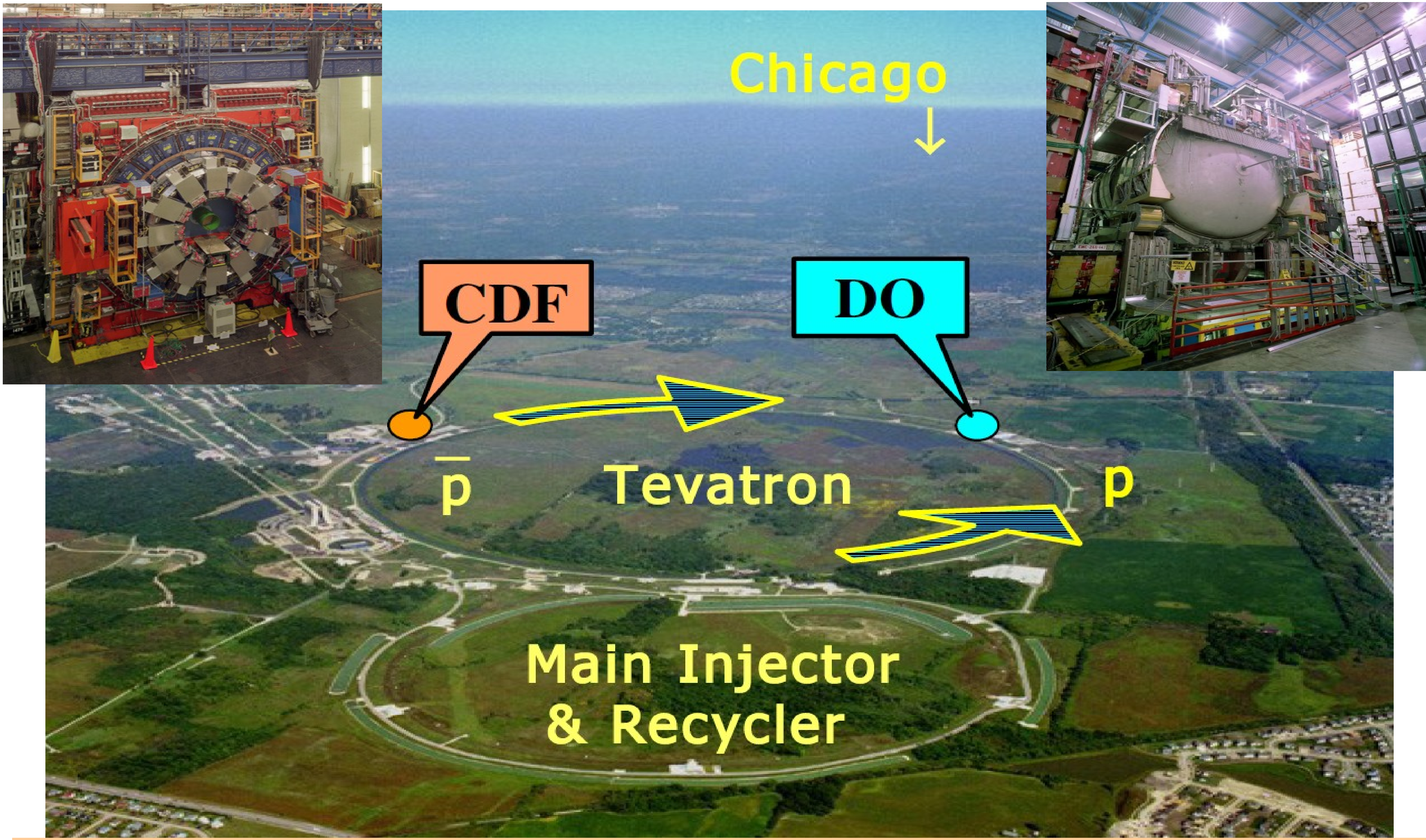
Users

- Community support
- A lot of free distributions
- Can run on older hardware
- Easy to install
- Tons of applications
- A decent skill in resume (if you plan to work in Fortune 500 list)
- Fun to use

Recommended desktop Operating systems to work with NIU HPC

+More than 600 supported Linux distributions

The HPC start at NIU -D0 experiment RunII (2000 - 2013)



Large experiments (650+ scientists, 90 institutions, 18 countries in D0 alone), a lot of data to work with == need of powerful computers

The HPC start at NIU - first compute clusters

Year 2000 – 2 NIU Phys. Dep. Servers - niuhep.niu.edu and nicadd.niu.edu

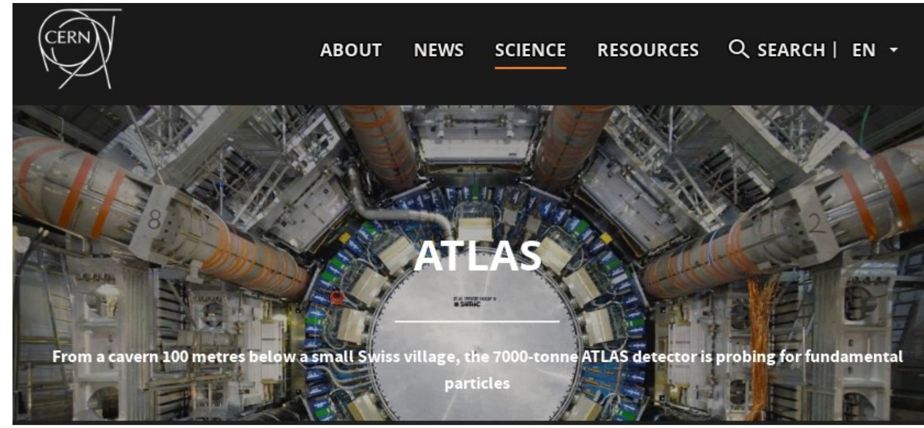
Fermi National Accelerator Laboratory

The DØ Experiment

For the Public | DØ Results | DØ Collaboration | DØ at Work

The DØ Experiment consists of a worldwide collaboration of scientists conducting research on the fundamental nature of matter. The experiment is located at the world's premier high-energy physics laboratory the Fermi National Accelerator Laboratory (Fermilab) in Batavia, Illinois, USA. The research is focused on precise studies of interactions of protons and antiprotons at the highest available energies provided by the Tevatron collider. It involves an intense search for subatomic clues that reveal the character of the building blocks of the universe. The DØ experiment finished data collection in 2011 when the Tevatron collider run ended and now is analyzing the collected data set.

Note: Some of the pages on this site are legacy pages and are no longer updated.



BEAM NICADD group

ClueD0 desktop cluster
~400 desktops in 2004
D0 NICADD group

ATLAS Tier T3 clusters
ATLAS collaborators
ATLAS NICADD group

ClueD0 design

- Current copy of D0 software.
- Access to cluster-wide batch queues.
- Security patches and updates for your machines
- Local root access available on your machines... Sys-admins available during the day to fix "supported" features
- Home directory backup every night...
- Centralized account management.

NICADD HPC
~20 nodes
700 processor slots (1.8-2.6 GHz) cluster under the HT CONDOR batch system running Scientific Linux OS
+
desktops and data servers

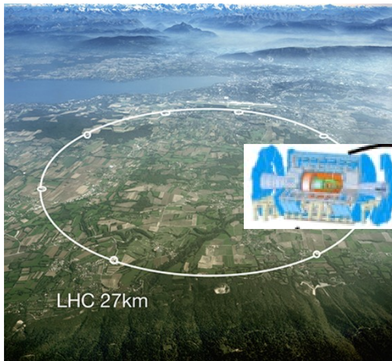
Tier 3g design/Philosophy

- Design a system to be flexible and simple to setup (1 person < 1 week)
- Simple to operate - < 0.25 FTE to maintain
- Scalable with Data volumes
- Fast - Process 1 TB of data over night
- Relatively inexpensive
  - Run only the needed services/process
  - Devote most resources to CPU's and Disk
- Using common tools will make it easier for all of us
  - Easier to develop a self supporting community.

ATLAS T3 project

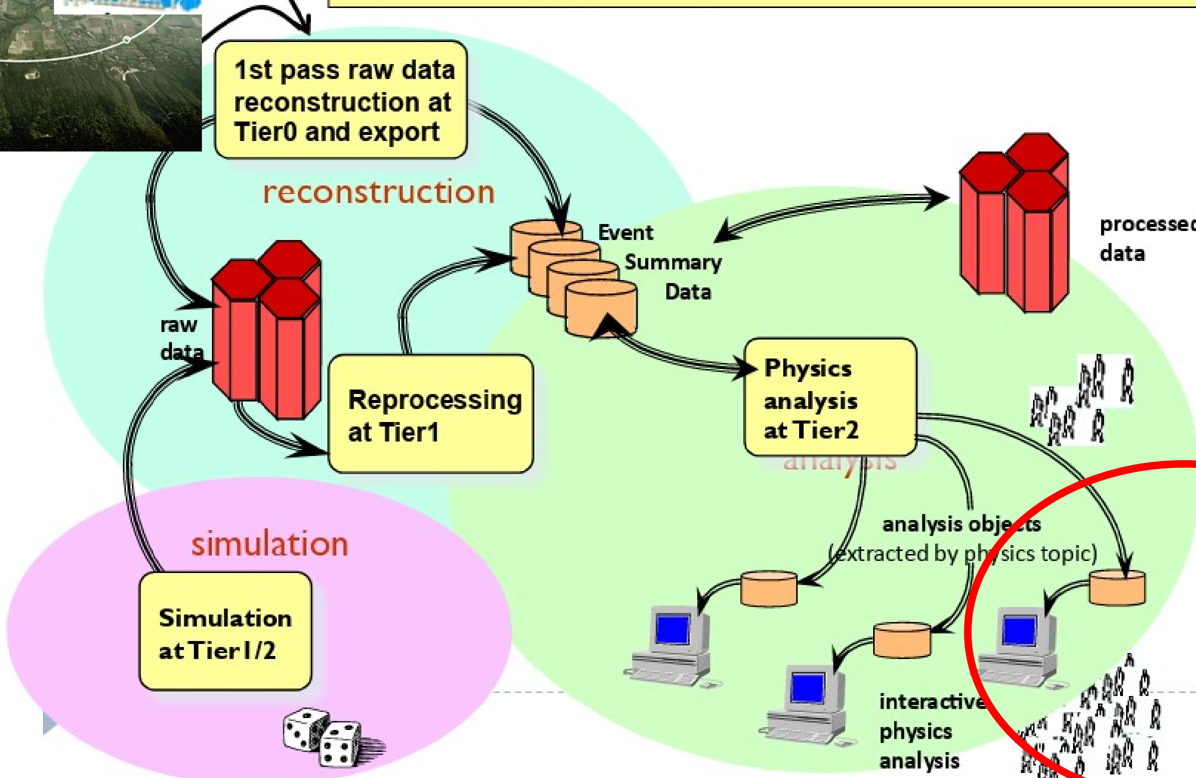
Doug Benjamin, Duke University - T3 model at time of creation

Simplified View - Atlas Computing Model



4 main computing operations according to the Computing Model:

- Initial processing of Raw data at CERN Tier0 - data export to Tier-1's/Tier-2's
- Data re-processing at Tier-1's using updated calibration constants
- Simulation of Monte Carlo samples at Tier-1's and Tier-2's
- (Distributed) physics analysis at Tier-2's and at more local facilities (Tier-3's)



Fresno State Tier 3 CERN / Atlas Cluster Hardware Overview

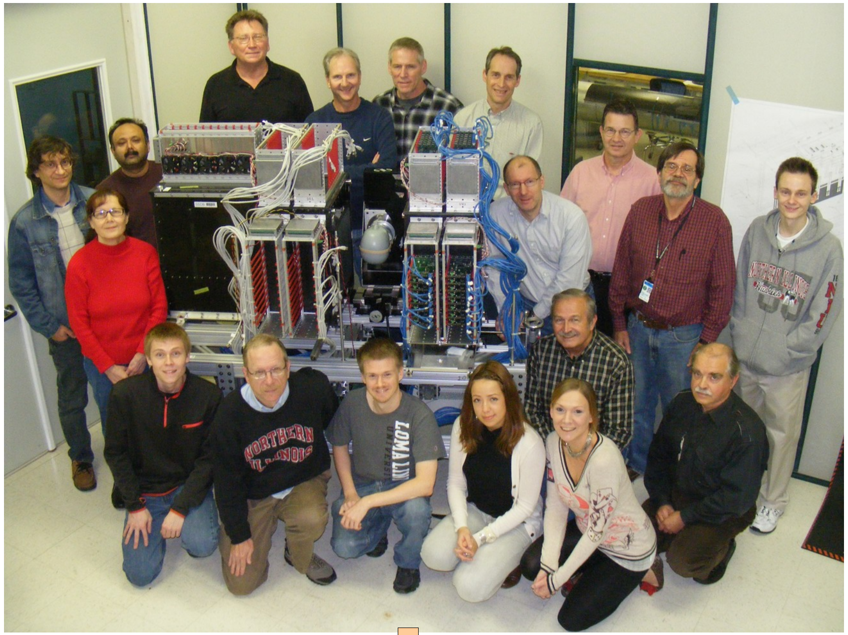
- Cisco 6500 Series
- Dell PowerConnect 6248
- Dell MD1200
- Dell MD1200
- Dell R510 - NFS
- Dell MD1200
- KVM
- Dell R510 - Work
- Dell R510 - Work
- Dell R710 - Head/Int
- Dell R710 - Work
- Dell R710 - Work
- Dell R710 - Work
- Dell R710 - Work
- Dell R710 - Work

NICADD HPC

Collaboration with Fresno University Atlas Team, 2011-2013

NICADD HPC

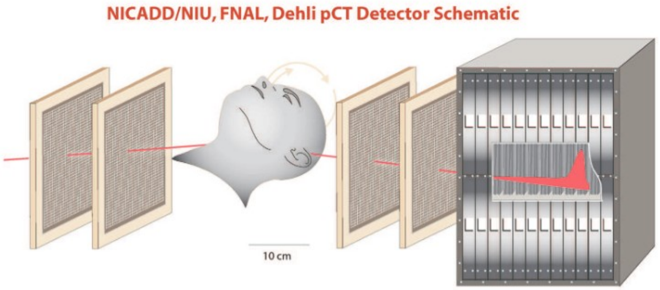
Proton Computed Tomography (pCT) project



- Gaea built in 2012 as a GPU-based system to process 200 GB "images" in 10 min

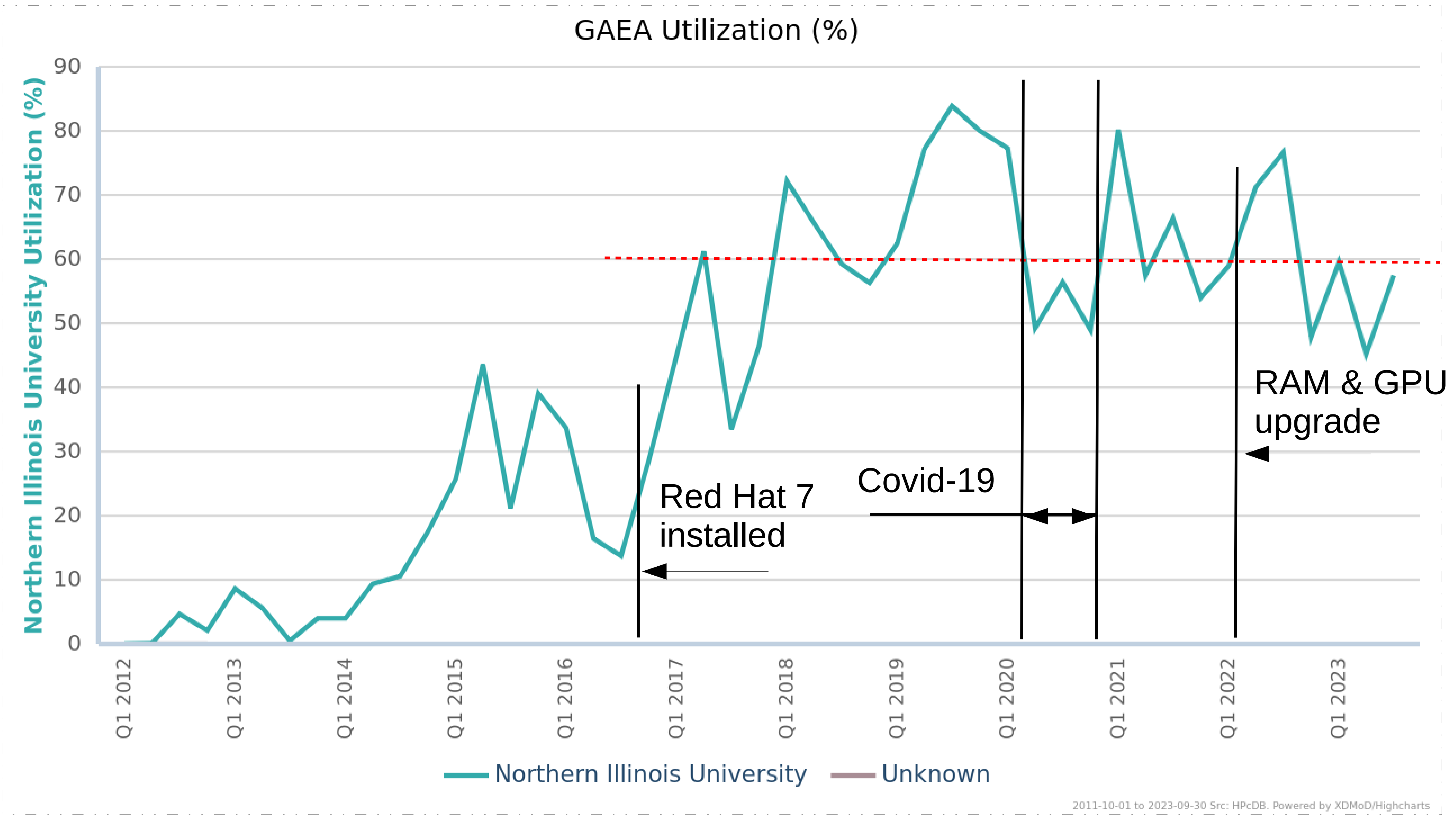
pCT project PIs
George Coutrakon and Nick Karonis

- Converted into a shared HPC system in 2014

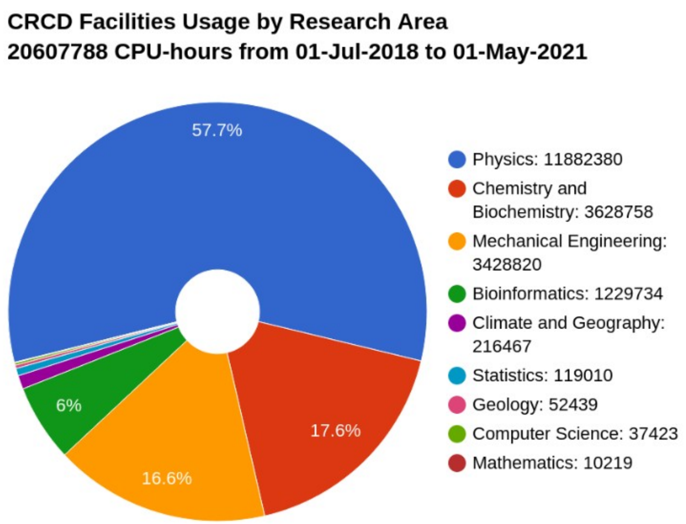


GAEA HPC

NIU HPC utilization history

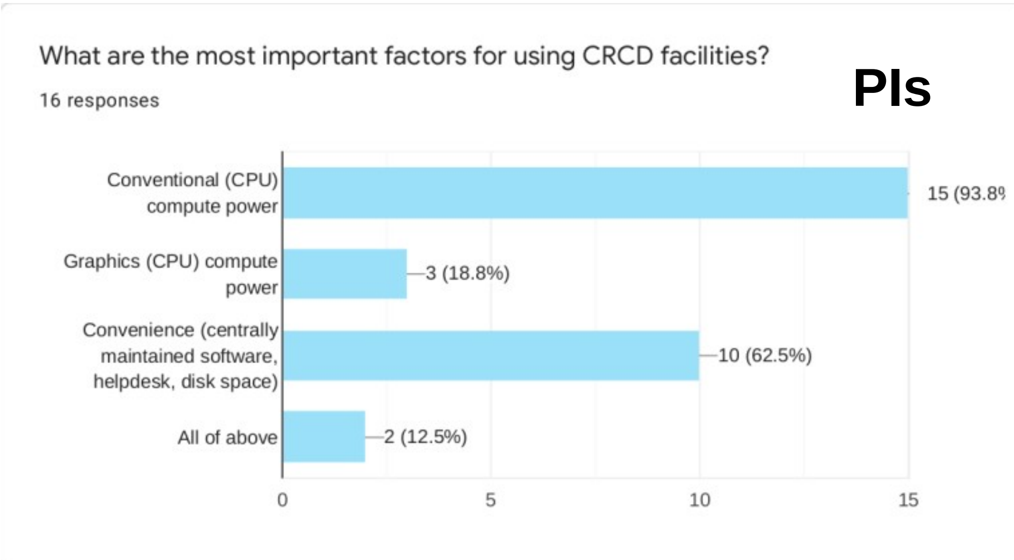


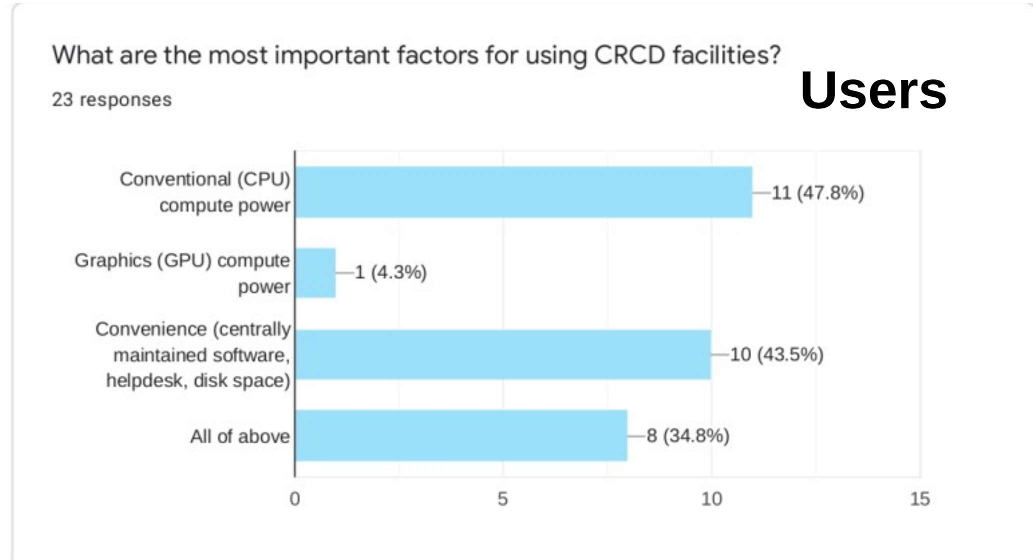
CRCD statistics (including NICADD), 2018-2021



NIU total (Jul 2018 -Nov 2021)

1. Submitted research proposals: 50 requesting $20.5M
2. Funded research: 20 totaling $6.2M
3. Proposals still under review 3 totaling $3.1M
4. Journal papers: 71
5. Ph.D. thesis: 13 (10 completed/3 in progress)
6. Master thesis: 12 (10 completed/2 in progress)
7. Conference presentations: 38
8. External collaborations: 25





The success of GAEA in attracting external funds allowed obtain NIU's financing for the Metis system (Gerald Blazey, CRCD team, DoIT)

NIU HPC Systems specification

Gaea Cluster

Gaea is a 60-node CPU/GPU hybrid cluster running Red Hat Enterprise Linux 7.x operating system. Each compute node is an HP SL380s G7 equipped with:

- 2 x Intel X5650 2.66 GHz 6-core processors
- 144 GB RAM, 4 x 500 GB 2.5" SATA disk drives in RAID10 configuration (i.e., 1 TB each node)
- 2 x NVIDIA TESLA P4 GPU, Pascal™ architecture, 8 GB RAM each card
- All 60 nodes are connected via Full 1:1 non-blocking Infiniband and Ethernet switch connectors.

The cluster also has two storage servers, each an HP Proliant DL380G7 server, and an HP P2000 disk storage array with 192 TB of effective storage space (i.e., after RAID6). The storage array is connected to the storage servers via 6 Gigabit per second SAS connections.

Metis Cluster

Metis (commissioned in September 2023) is a 32-node CPU/GPU hybrid cluster running Red Hat Enterprise Linux 8.x operating system. Each compute node is an HPE DL 385 Gen 10+ V2 server equipped with:

- 2x AMD EPYC 7713 CPUs 2.0 GHz 64-core processors
- 256-1TB GB RAM, 1 x 1TB SSD scratch disk drives
- 1 x NVIDIA A100 GPU, Amper™ architecture, 40 GB RAM each card
- All 32 nodes are connected via 200 Gbps Infiniband network

A Cray ClusterStor E1000 storage server provides the cluster with 1 PB of shared disk space.

Metis system highlights

- 32 nodes, (128 2.0 GHz cores, A100 GPU, 1.2 TB RAM, 4.4 TB SSDscratch)/node
- RHEL 8.x Linux, PBSpro batch system, 1 PB Lustre scratch disk, 200 Gbps network
- Maximum theoretical performance in 64-bit TeraFlops (131 GPUs, 310 GPUs)
- For some applications be treated as a single computer with 4096 CPU cores

Metis favorites

▪ Tasks optimized OpenMPI-OpemMP-GPUs (can use all resources simultaneously)	▪ Single CPU instances of less than 2GB RAM running simultaneously

NIU HPC software installations

CRCD supported software at Metis, RHEL 8.x, October 2025

Module	Description	License/cost	Projects
GCC v4.9.3-12.3.0	GNU Compiler Collections (c,c++,fortran)	GPL/free	all
intel-oneapi-2023.1.0	Intel Compiler Collection (c,c++,fortran)	Intel/free	all
boosts-1.83.0	Portable C++ source libraries	GPL/free	all
openmpi, v1.8.1-4.1.5	OpenMPI libraries	GPL/free	all
Python3,R,Lua,Java	Script languages	GPL,Oracle/free	all
CUDA v7.5-12.2	NVIDIA GPU Libraries	NVIDIA/free	all
netcdf-4.9.2	Scientific data format library	UNIDATA/free	marssim, climlab, aard
hdf5-1.10.10,phdf5	High performance data management	HDF5/free	marssim, climlab, aard
ROOT,Octave	Physics Analysis	GPL/free	physics,pct
PYTHIA,MADGRAPH	Particle collisions simulations	GPL/free	HEP
GEANT4 via cvmfs	Physics Detectors simulations	CERN/free	HEP
g16-rC02avx2	Chemical processes modeling	Gaussian/$$	wheeler1,tglab
LAMMPS	Molecular dynamics simulator	GPL/free	wheeler1
namd-2.12	Parallel molecular dynamics	UIUC/free	moisture, wheeler1
orca-5.0.4	Quantum chemistry program	Academic/free	tglab
qmcpack-3.9.2	Quantum chemistry Monte-Carlo package	Academic/free	wheeler1
qp2-2.1.2	Quantum chemistry, wave function methods	Academic/free	wheeler1
opal-2022.1	Parallel Accelerator Library	GPL/free	aard,fast
ACE3P-2023	Electromagnetic, thermal and mechanical modeling	Stanford/free	aard,fast
WarpX-2023	Lasers and particle beams propagation	WarpX/free	aard,fast
trelis-16.5	High-quality mesh generation	Coreform/$$	aard,fast
richdem-0.0.3,taudem2023	Hydrologic analysis tools	Academic/free	marssim
cm1r21.0	Atmospheric Research	MIT/free	climlab
WRF,MPAS	Weather Research and Forecasting Models	Academic/free	climlab
OSRM	Open Source Routing Machine	GPL/free	marssim

Both CPU and GPU based packages supported up to the most recent versions
(79 unique packages and 274 accounting for different versions).

HPC software use

https://www.niu.edu/crcd/current-users/crnt-users-software.shtml

CRCD installations

- Accessible via environment modules

module av; module load; module purge

Personal installations

- Can be build from source under

/opt/metis/el8/ucontrib

Python modules

- pip3 manager

  pip3 install pkgName

- conda manager
- conda create -name=p39tst python=3.9
- conda activate p39tst
- conda install pkgName

Jupyter notebooks

- Install Jupyter notebook

  pip3 install jupyter

  pip3 install urllib3=1.26.6

- Launch a notebook at a port xxxxx

  jupyter notebook –no-browser \

  –port=xxxxx -ip=0.0.0.0

- Connect via instructions at CRCD page

CVMFS based software libraries

- Pre-mounted for ATLAS (/opt/atlas) and CMS (/opt/osg/cmssoft) repositories

CRCD installs and supports “tagged” versions
and provides resources for users applications under development

Current Operations - crcd.niu.edu

About · Prospective Users · Current Users · CRCD at Work

Center for Research Computing and Data

Search CRCD

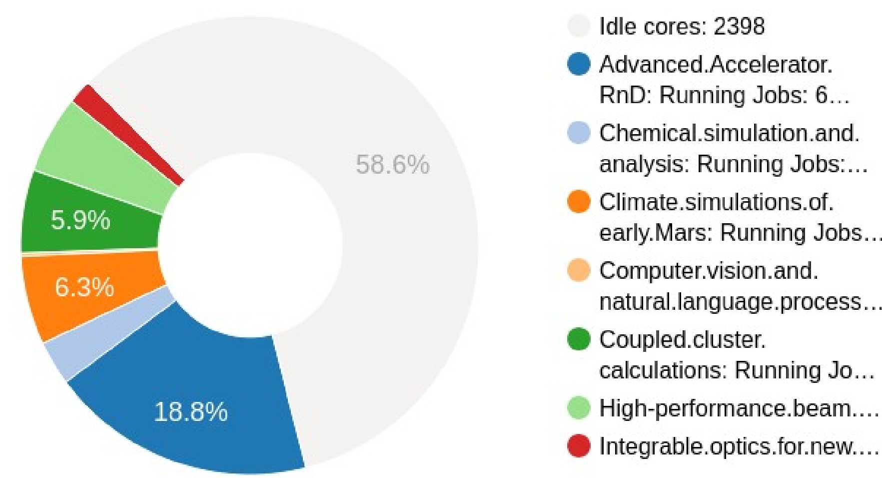
CRCD web site provides:

- real-time system status
- cluster usage policies
- detailed access instructions for beginners
- hardware and software documentation
- quick-start examples
- job monitoring tools
- contact information
- CRCD helpdesk app

Metis Cluster Status, Sep 09 2025, 10:51:16

Running Jobs	Queued Jobs	Running Nodes	Idle Nodes	Reserved Nodes (R|I)	CPU Cores (R|I)	GPU Cards (R|I)	CPU Usage	GPU Usage	Node Usage
37	23	32	0	0 (0|0)	1698|2398	29|3	**41%**	**90%**	**100%**

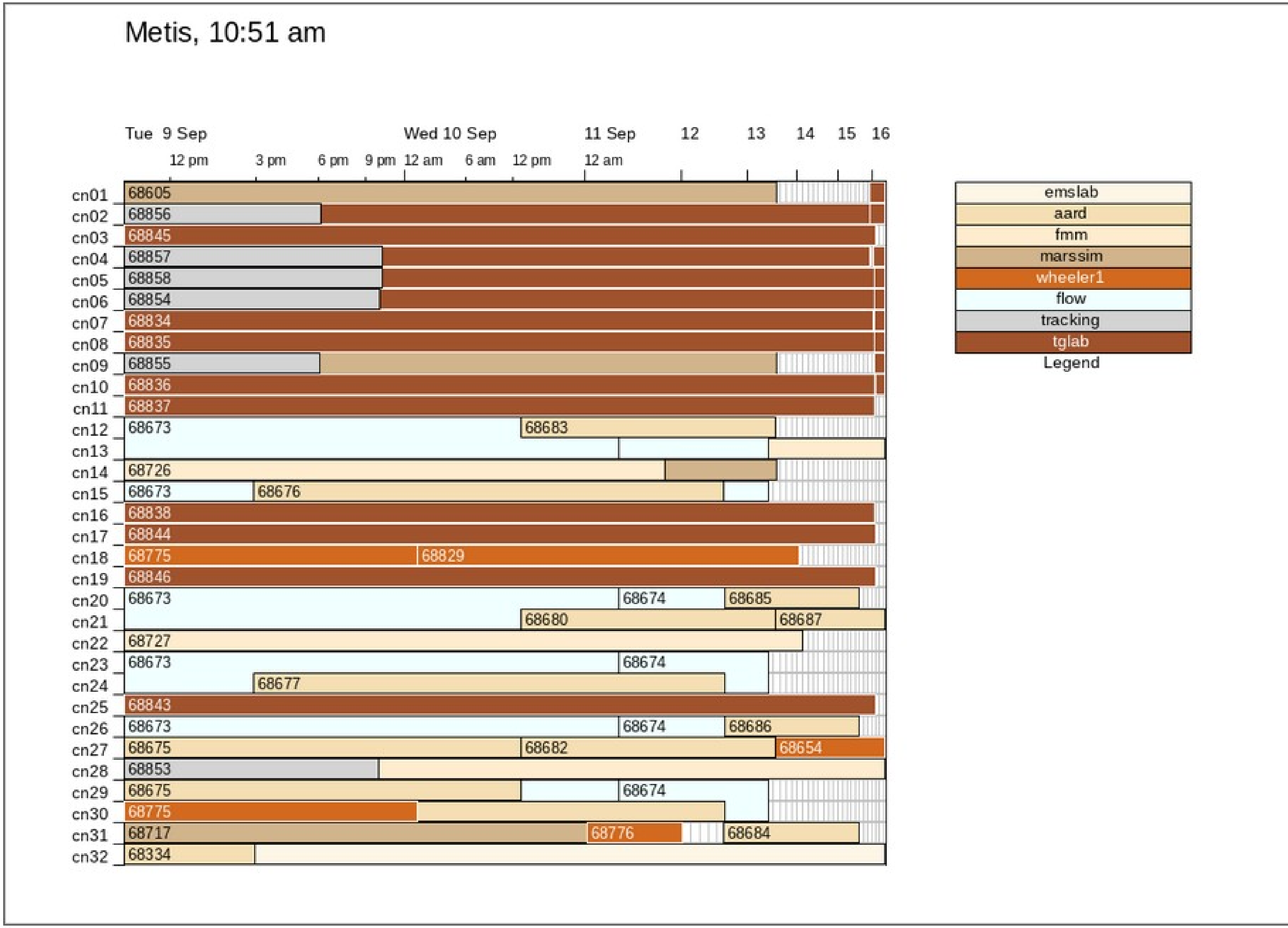
Running Projects (CPUs in use,%)



The site accumulates the CRCD experience since its foundation; the first stop for new users.

Current Operations - crcd.niu.edu

Batch System Monitor



The system is busy, we must rely on user’s cooperation

Types of computational workflow - virtualization

We support Virtual Machines for special applications

Supported projects

Course Projects

▪ A PI submits a short description of the project(s) and a list of students. A PI creates and manages course materials and shared folders on the system

CRCD provides accounts and software support and the intro tutorial.

Research Projects

▪ A PI submits a short description of the project(s) and a list of researchers.

CRCD provides accounts and software support

The new user best practice

▪ Follow instructions in the “Welcome e-mail”

read the corresponding documentation pages

▪ Google for a beginner’s Linux tutorial

learn how to edit files, setup the environment, write bash scripts

▪ Make an application run at the login node

▪ Go over the batch job system tutorial, run the supplied test job, modify it, learn how to use the batch sytem monitoring tools.

▪ If runninng into a problem, submit a detailed help request to the crcdhelpdesk@niu.edu

Types of computational workflow

Tasks classification (critical resources are the memory and scratch space at the compute nodes):

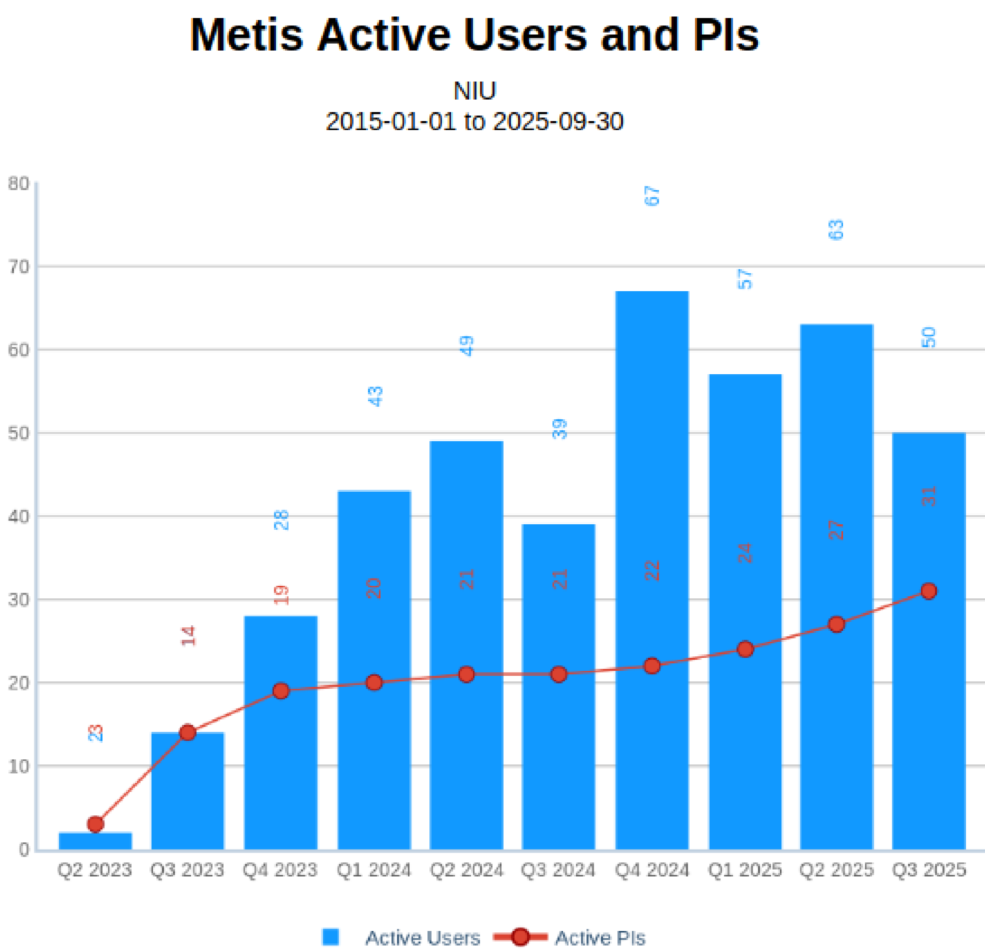
- **Sequential jobs**
  Each job uses only one CPU or GPU; multiple jobs process separate chunks of data. Can use up to 1256 GB of memory and up to 4.4 TB of scratch space

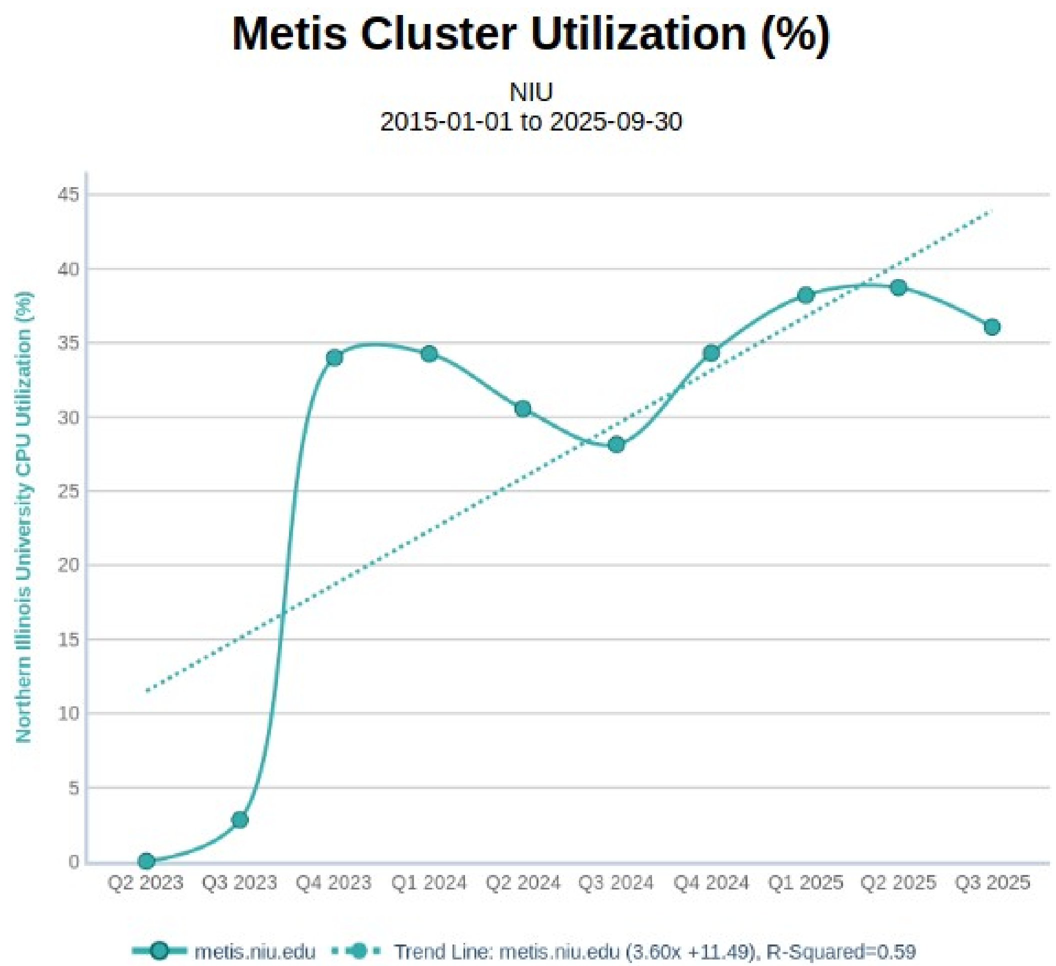
- **MPI (parallel) jobs**
  Each job uses several CPUs or GPU; 1 CPU per MPI process; total memory consumption of a job restricted by the 1256GB

- **OpenMP (parallel) jobs**
  Each job uses several CPUs or GPU; total memory consumption of a job restricted by the 1256GB. But each MPI process can use several CPU cores.

It is tempting to assume a linear run time decrease with an increased number of CPUs or GPUs. But the best practice is to run a scan and find the optimal number of CPUs, GPUs, and nodes for each workflow

Usage history





31 active research projects today, 35% mean system utilization

Summary

Years of successful operations of CRCD facilities

We are welcoming new users

A note on a code quality

Nvidia A100 GPU on a Metis worker node, runtime 3.3 sec
(down from 23 sec for non-optimized code)

```
module load  openmpi/openmpi-4.1.5-gcc-12.3.0-cuda-12.2
nvcc -arch=sm_80 -o jacobi_step6 -x cu -lnvToolsExt 6_cudaswap.cpp
nsys profile --stats=true -o jacobi_step6 -f true ./jacobi_step6 > metis_profile.txt
Success!
Run time = 3.304 seconds
.......
.......
NVTX Range Statistics:

 Time (%)  Total Time (ns)  Instances     Avg (ns)       Med (ns)       Min (ns)     Max (ns)     StdDev (ns)    Style        Range
 --------  ---------------  ---------  -------------  -------------  -----------  -----------  -------------  -------  ---------------
     86.5    2,852,220,373        288    9,903,543.0    9,605,938.5    9,561,175   26,891,544   1,360,082.1  PushPop  Jacobi step
     10.5      345,243,947          1  345,243,947.0  345,243,947.0  345,243,947  345,243,947           0.0  PushPop  Allocate memory
      1.8       60,256,911          1   60,256,911.0   60,256,911.0   60,256,911   60,256,911           0.0  PushPop  Initialize data
      1.1       36,484,125        288      126,681.0      123,310.5      121,106      221,464      12,729.8  PushPop  Swap data
      0.0        1,375,482          1    1,375,482.0    1,375,482.0    1,375,482    1,375,482           0.0  PushPop  Free memory
#=====================================================================================================================================
```

Nvidia P4 GPU on Gaea worker node, runtime 4.2 sec
(down from 100 sec for non-optimized code)

```
module load openmpi/openmpi-4.0.2-gcc-9.2.0-cuda-11.5
nvcc -arch=sm_80 -o jacobi_step6 -x cu -lnvToolsExt 6_cudaswap.cpp
nsys profile --stats=true -o jacobi_step6 -f true ./jacobi_step6 > gaea_profile.txt
Success!
Run time = 4.15 seconds
.......
.......
NVTX Range Statistics:

 Time(%)  Total Time (ns)  Instances  Average (ns)   Minimum (ns)  Maximum (ns)  StdDev (ns)   Style        Range
 -------  ---------------  ---------  -------------  ------------  ------------  -----------  -------  ---------------
    83.7    3,452,441,162        288   11,987,642.9    11,627,775    29,261,697  1,246,594.0  PushPop  Jacobi step
     6.1      253,039,111          1  253,039,111.0   253,039,111   253,039,111          0.0  PushPop  Allocate memory
     5.1      209,138,746          1  209,138,746.0   209,138,746   209,138,746          0.0  PushPop  Initialize data
     5.1      208,789,359        288      724,963.1       707,671       897,991     41,133.5  PushPop  Swap data
     0.1        2,221,107          1    2,221,107.0     2,221,107     2,221,107          0.0  PushPop  Free memory
#=====================================================================================================================
```

Programming skills matter – a small run-time difference for a well optimized code

Backup Slides

Top HPC systems

2025, https://www.top500.org/lists/top500/2025/06/highs/

Rank	Site	System	Cores	Rmax (TFlop/s)	Rpeak (TFlop/s)	Power (kW)
1	DOE/NNSA/LLNL United States	**El Capitan** - HPE Cray EX255a, AMD 4th Gen EPYC 24C 1.8GHz, AMD Instinct MI300A, Slingshot-11, TOSS HPE	11,039,616	1,742.00	2,746.38	29,581
2	DOE/SC/Oak Ridge National Laboratory United States	**Frontier** - HPE Cray EX235a, AMD Optimized 3rd Generation EPYC 64C 2GHz, AMD Instinct MI250X, Slingshot-11, HPE Cray OS HPE	9,066,176	1,353.00	2,055.72	24,607
3	DOE/SC/Argonne National Laboratory United States	**Aurora** - HPE Cray EX - Intel Exascale Compute Blade, Xeon CPU Max 9470 52C 2.4GHz, Intel Data Center GPU Max, Slingshot-11 Intel	9,264,128	1,012.00	1,980.01	38,698

Top HPC systems

2023, https://www.top500.org/lists/top500/2023/06/

Rank	System	Cores	Rmax (PFlop/s)	Rpeak (PFlop/s)	Power (kW)
1	**Frontier** - HPE Cray EX235a, AMD Optimized 3rd Generation EPYC 64C 2GHz, AMD Instinct MI250X, Slingshot-11, HPE DOE/SC/Oak Ridge National Laboratory United States	8,699,904	1,194.00	1,679.82	22,703
2	**Supercomputer Fugaku** - Supercomputer Fugaku, A64FX 48C 2.2GHz, Tofu interconnect D, Fujitsu RIKEN Center for Computational Science Japan	7,630,848	442.01	537.21	29,899
3	**LUMI** - HPE Cray EX235a, AMD Optimized 3rd Generation EPYC 64C 2GHz, AMD Instinct MI250X, Slingshot-11, HPE EuroHPC/CSC Finland	2,220,288	309.10	428.70	6,016