HPC at Colby College

* Where we started here at Colby
* Current status of our three clusters at Colby
* How they are accessed & used (not just ssh!
* Where we started here at Colby
* Scientific reproducibility and the future
* Three usage examples

(Docker and remote computation)

(Nick, Stephanie, Dave)
HPC at Colby College

In 2007

advocacy by Chemistry/Physics Faculty.
When it ended Colby picked it up because of
My position was originally grant funded.

Schupflab (9 SGIs and 10 Mac pros)
Startup funded machines (hidden in offices)
NSF funded machine (PacificTech AMD)

In 2007

HPC at Colby College
Mission of this liberal arts college: Undergraduate research matters and faculty have to do original research for tenure consideration. No engineering/material science labs, DOE's Oakridge, with repeated workflows. We are not an R1 University; Pharmaceutical, Jackson major research institute (Broad Institute, Jackson). We are not an R1 University; Pharmacutical or... but our students are going there...

HPC at Colby College
HPC at Colby College

Student's peak year (240+ compute hrs)

- 2008: Charley White
- 2009: Martha Witick, Eric Braunstein
- 2010: John B. Goods
- 2011: Zach Cecera
- 2012: Titobiloluwa O. Awe, David J. Cain
- 2013: David M. Brazel

- 2010: Amanda Towne
- 2011: Jesus S. Vidaurri-Martinez, Justin Ambrose Sperry
- 2012: William M. Randall, Zhanar Seitmagzimova
- 2013: Roxana A. Kessell, Karthik A. Moore

- 2014: Iavor V. Dekov, Bethany A. Alther
- 2015: Aurora V. Hellowell, Claire A. Herbig
- 2016: Daniel R. Cover, Mary A. Fletcher
- 2017: Triobiodioma O. Awe, David J. Cain

- 2018: Sara Harmon, Allison Cheever
- 2019: Zach Cecera
- 2020: Eric White
- 2021: Charley Towne

HPC at Colby College

Student's peak year (240+ compute hrs)

- Joseph D. DeAngelis
- William J. Linn
- Andrew Y. Kim
- Megan S. Michie
- Alexander T. Lato
- Audrey E. Lyman
- Scott T. Hansen
- Devon S. Cormack
- Ian M. Tibbetts
- Adam A. Lavenu
- William T. Kearney
- Cameron A. Matricks
- Olivia W. Lang
- Abebu A. Kassie
- James L. Epstein
- Jonathan C. Brink-Roby

2014
HPC at Colby College

Three clusters

Mission/support and who to talk to

Teaching cluster the Faculty/TA

Experimental cluster the student

Production cluster (nscc) the PI
HPC at Colby College
Experimental Cluster
Redhat 7, Fedora 22 (25-30 nodes)
Mixed hardware (cpus and gpus)
Uptime in terms of days, hands on.
Local storage, redundancy, remote storage
* * * stuff breaks here * * *

HPC at Colby College
Experimental Cluster
Redhat 7, Fedora 22 (25-30 nodes)
Mixed hardware (cpus and gpus)
Uptime in terms of days, hands on.
Local storage, redundancy, remote storage

* * * stuff breaks here * * *
HPC at Colby College

Teaching cluster

- 10s of days uptime
- Redhat 7/Centos 7
- Well behaved software, no code without testing

Visualization (VMD, MOE, IDL)

Gaussian, Webmo, Xray (Apex3), 12-24 nodes as machines come and go

GRID compute cluster

Gaussian, Webmo, Xray (Apex3), 12-24 nodes as machines come and go

Teaching cluster
HPC at Colby College

NSCC (National Science and Compute Center) production cluster

32 Terabyte direct storage
Commodity hardware with redundancy
Redhat 6/7 and CentOS 7
Locked in the machine room behind ID & keys
Uptime in 100s of days

no cost recovery; other funding models
a community cluster
(tanked by grants, faculty startup, & capital items)
HPC at Colby College
HPC at Colby College

- 96 (40) terabyte quick
- 28 terabyte fast
- 2 terabyte very fast
- 200-500 Giga Wokspace
- 128-48 Giga RAM per node
- 300+ CPUs
- more blades
- 17 nodes w/ room for 9
- Cool side
HPC at Colby College

- Hot side
- Network back plane
- Everything fiber except management
- 2/3 40gig Ethernet
- 1/3 10gig Ethernet
- Distance in cm to storage arrays
HPC at Colby College

Software:

NSCC (Production) 75 regularly used
open source MATLAB, R, Genious, GROMACS, Gaussian, MOE, original compiled code, Trinity (RNASeg), BLAST[*], IDL and Python

In order of usage by discipline: CS(s)ystems Biology modeling, Computational Chemistry, Bioinformatics, Astronomy, Geology

(this changes by semester)

Biointormatics, Astronomy, Geology

MATLAB, R, Genious, GROMACS, Gaussian,
MOE, original compiled code, Trinity (RNASeg),
BLAST[*], IDL and Python

HPC at Colby College
HPC at Colby College

Departmental usage: Chemistry and Astronomy

Heavy on the visualizations of larger data sets.

Teaching cluster: Gaussian, R, MOE, VMD, IDL, IDL

Software
Departmental usage: CompSci, Statistics, Biology

Original code, MATLAB, Octave, R, GPUs

Experimental Cluster: GCC, CUDA, Quartus

Software

HPC at Colby College
HPC at Colby College

How they are accessed

ssh
ssh
ssh

(Webmo, Galaxy, rstudio, etc)
(Web servers in the virtual space)

(Short demo after if you want)
Three real examples

Nick, Stephanie, Dave

If time permits other resources

HPC at Colby College
Google & Azure not as mature but coming.

Many external options: Amazon, iPlant, UMN, FSU

Atom, Docker and Containers

Reproducibility and remote deployment (cloud)

HPC at Colby College