orcd@mit:~# echo “office of research computing and data”

Basic Computing Services In The Physics Department

Dr. James Andrew Cuff

6th January, 2023

orcd@mit:~$ background

University of Oxford, UK, D. Phil. Molecular Biophysics,

- *Neural networks for protein structure prediction*

Wellcome Trust Sanger Institute, Group Leader

- *Human Genome, built the Informatics Systems Group*

Broad Institute of MIT & Harvard, Group Leader

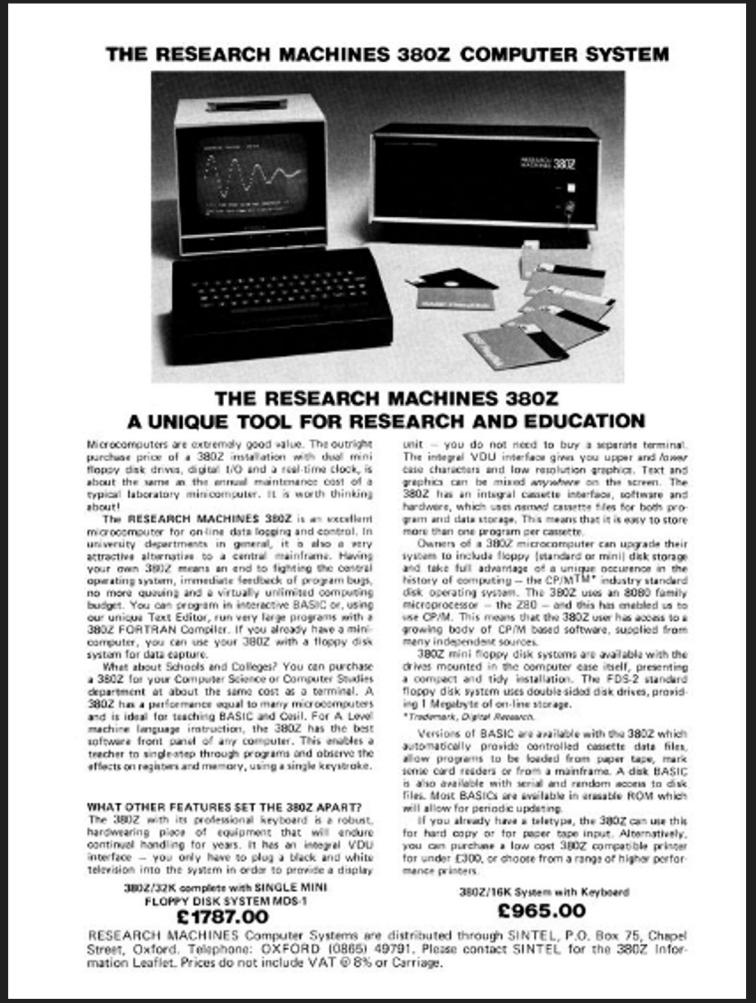
- *Built the Applied Production Systems Group*

Harvard, Assistant Dean & Distinguished Engineer for Research Computing

- *100,000+ CPU cluster & MGHPCC from scratch*

Independent Consultant

- 5+ years learning the “cheat-codes” in .com and .edu



THE RESEARCH MACHINES 380Z COMPUTER SYSTEM

THE RESEARCH MACHINES 380Z
A UNIQUE TOOL FOR RESEARCH AND EDUCATION

Microcomputers are extremely good value. The outright purchase price of a 380Z installation with dual mini floppy disk drives, digital I/O and a real-time clock, is about the same as the annual maintenance cost of a typical laboratory minicomputer. It is worth thinking about!

The RESEARCH MACHINES 380Z is an excellent microcomputer for on-line data logging and control. In university departments in general, it is also a very attractive alternative to a central mainframe. Having your own 380Z means an end to fighting the central operating system, immediate feedback of program bugs, no more queuing and a virtually unlimited computing budget. You can program in interactive BASIC or, using our unique Text Editor, run very large programs with a 380Z FORTRAN Compiler. If you already have a mini-computer, you can use your 380Z with a floppy disk system for data capture.

What about Schools and Colleges? You can purchase a 380Z for your Computer Science or Computer Studies department at about the same cost as a terminal. A 380Z has a performance equal to many microcomputers and is ideal for teaching BASIC and Cesil. For A Level machine language instruction, the 380Z has the best software front panel of any computer. This enables a teacher to single-step through programs and observe the effects on registers and memory, using a single keystroke.

WHAT OTHER FEATURES SET THE 380Z APART?

The 380Z with its professional keyboard is a robust, hardwearing piece of equipment that will endure continual handling for years. It has an integral VDU interface – you only have to plug a black and white television into the system in order to provide a display unit – you do not need to buy a separate terminal. The integral VDU interface gives you upper and lower case characters and low resolution graphics. Text and graphics can be mixed anywhere on the screen. The 380Z has an integral cassette interface, software and hardware, which uses named cassette files for both program and data storage. This means that it is easy to store more than one program per cassette.

Owners of a 380Z microcomputer can upgrade their system to include floppy (standard or mini) disk storage and take full advantage of a unique occurence in the history of computing – the CP/M™* industry standard disk operating system. The 380Z uses an 8080 family microprocessor – the Z80 – and this has enabled us to use CP/M. This means that the 380Z user has access to a growing body of CP/M based software, supplied from many independent sources.

380Z mini floppy disk systems are available with the drives mounted in the computer case itself, presenting a compact and tidy installation. The FDS-2 standard floppy disk system uses double sided disk drives, providing 1 Megabyte of on-line storage.

*Trademark, Digital Research.

Versions of BASIC are available with the 380Z which automatically provide controlled cassette data files, allow programs to be loaded from paper tape, mark sense card readers or from a mainframe. A disk BASIC is also available with serial and random access to disk files. Most BASICs are available in erasable ROM which will allow for periodic updating.

If you already have a teletype, the 380Z can use this for hard copy or for paper tape input. Alternatively, you can purchase a low cost 380Z compatible printer for under £300, or choose from a range of higher performance printers.

380Z/32K complete with SINGLE MINI FLOPPY DISK SYSTEM MDS-1
£1787.00

380Z/16K System with Keyboard
£965.00

RESEARCH MACHINES Computer Systems are distributed through SINTEL, P.O. Box 75, Chapel Street, Oxford. Telephone: OXFORD (0865) 49791. Please contact SINTEL for the 380Z Information Leaflet. Prices do not include VAT @ 8% or Carriage.

orcd@mit:~$ history

james cuff fermilab

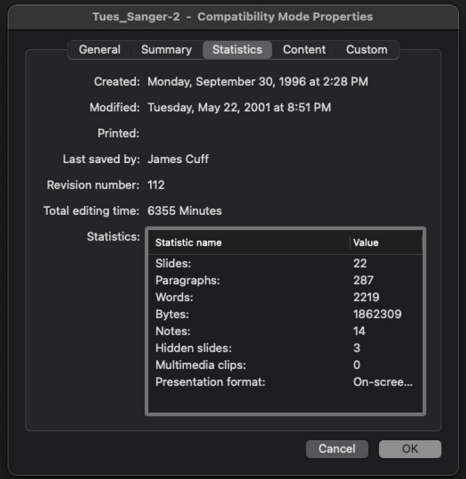
All | Images | News | Shopping | Videos | More | Tools

About 290,000 results (0.44 seconds)

https://conferences.fnal.gov › lccws › info › talks

Large Scale Cluster Computing Workshop Talks

1:45 **Fermi Lab** Overview and Welcome - Matthias Kasemann (FNAL) ... **James Cuff** (Sanger Inst); Doug Thain (Wisconsin) (Condor I/O); Kors Bos Talk (NIKHEF) ...



Last Updated by AGS, 13 Sep, 2001

http://conferences.fnal.gov/lccws/papers/tues/Tues_Sanger.ppt

conferences.fnal.gov/lccws/info/talks.html

Talks From The Large Scale Cluster Computing Workshop

Tuesday

12:00-1:30 Registration - Wilson Hall Atrium South End

1:30 Opening Session - Chaired by Dane Skow (FNAL)

Opening - Goals of the Workshop - Alan Silverman (CERN) and Dane Skow (FNAL)

1:45 Fermi Lab Overview and Welcome - Matthias Kasemann (FNAL)

2:15 LHC Computing Needs - Wolfgang Von Rueden (CERN)

3:00 The IEEE Task Force on Cluster Computing and Directions in Scalable Clusters - Bill Gropp (ANL)

3:30 Break

4:00 Panel on Usage Cases - FNAL, BNL, BaBar, Non-HEP

Panelists:

Panel Chair: Dane Skow

- Tom Yanuklis Talk (RHIC)
- Charles Young Talk (BaBar)
- Steve Wolbers Talk (FNAL)
- James Cuff Talk (The Sanger Centre)
- Ralf Gerhards Talk (H1 at DESY)
- Atsushi Manabe Talk (Kek)
- Jim Simone Talk (TH QCD at Fermilab)

For Each Site:

- A Brief Description of Each Cluster
- Its Size
- Its Architecture
- What it is used for
- Any Special Features
- What decisions/accomodations (architecture, hardware etc) did you make because of any special nature of the applications to be run.
- What Optimizations Did You Make

5:30 End of Day 1

Wednesday

9:00 - 10:30 Clusters at Large Sites

Chairman - Steve Wolbers

- CERN Clusters of Today - Tim Smith (CERN)
- BNL and other large clusters - Steve DuChene (VA Linux)
- The SLAC Computer Centre - Chuck Boeheim

10:30 Break

10:50 Panel on Hardware Issues

Selection Criteria, Life Cycles, Cluster Heterogeneity Etc...

Chairwoman: Lisa Giachetti

Panelists

- Tom Yanuklis Talk (BNL)

orcd@mit:~# echo “dec alpha”

orcd@mit:~$:
greybeard.exe

The 60's | The 70's | The 80's | The 90's | The 00's | Beyond

*From **centralized** to **decentralized**, **collaborative** to **independent** and right back again!*

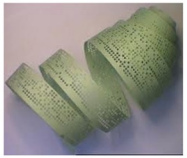
Mainframes
~ 0Mbit

VAX
~ 1Mbit

The PC
~ 10Mbit

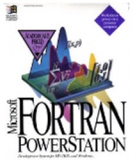
Beowulf Clusters
~ 1000 Mbit

Central Clusters
~ 10,000 Mbit



Centers provide access to compute

The supercomputing famine, funding gap



Individual computing

Computing is too big to fit under desk, Linux explodes

Clouds/VMware
IaaS, SaaS, PaaS

SHARING

100%	60%	0%	40%	???%

Bigger, better but further and further away from the scientist's laboratory and desktop

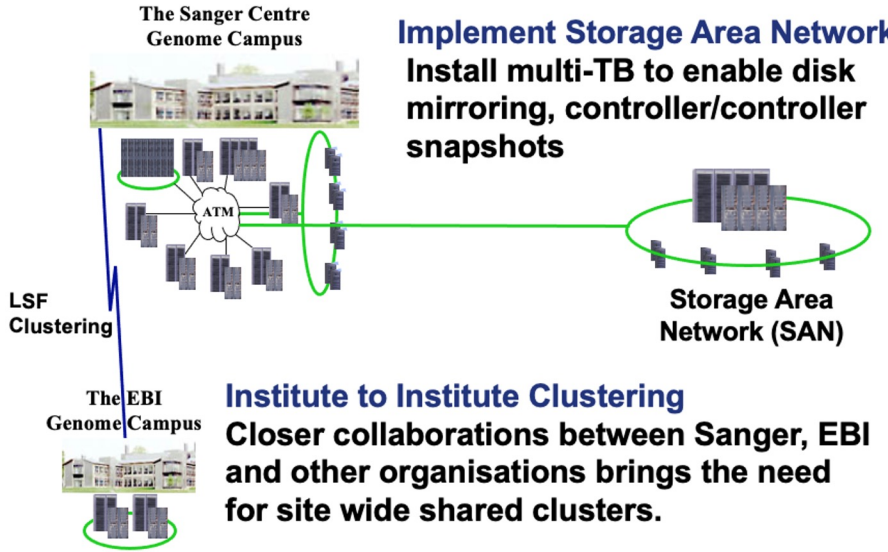
```
orcd@mit:~$: echo 'SELECT YEAR from PPT;'| mysql
+--------------------+
|               YEAR |
+--------------------+
|               2001 |
+--------------------+
```

Longer Term Future

- **Wide Area Clusters**
  Needed for large scale collaborations.
- **GRID Technology - Global Distributed Computing**
  International Cluster collaborations with other scientific institutes
- **Sanger is keen to keep abreast of this emerging technology**

GLOBAL COMPUTE ENGINES

Immediate Future



Implement Storage Area Network
Install multi-TB to enable disk mirroring, controller/controller snapshots

Storage Area Network (SAN)

Institute to Institute Clustering
Closer collaborations between Sanger, EBI and other organisations brings the need for site wide shared clusters.

orcd@mit:~$ jobs

Started Sept 26th 2022

Chris, Paul, Mike, Lauren, Kurt

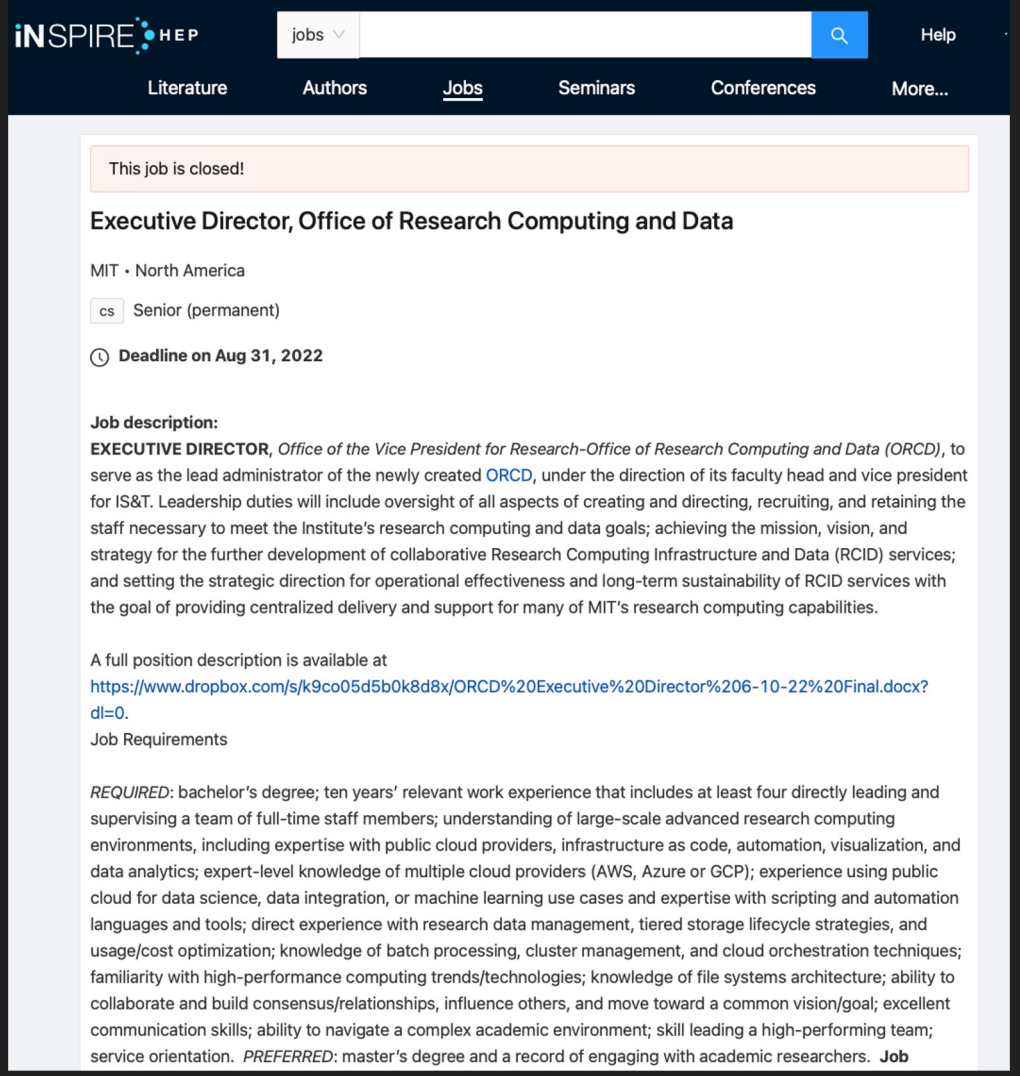
TechSquare

SuperCloud (signif. $ and FTE)

Satori

Engaging

CC3DB



iNSPIRE HEP | jobs | Help

Literature | Authors | Jobs | Seminars | Conferences | More...

This job is closed!

Executive Director, Office of Research Computing and Data

MIT • North America

cs Senior (permanent)

Deadline on Aug 31, 2022

Job description:

EXECUTIVE DIRECTOR, *Office of the Vice President for Research-Office of Research Computing and Data (ORCD)*, to serve as the lead administrator of the newly created ORCD, under the direction of its faculty head and vice president for IS&T. Leadership duties will include oversight of all aspects of creating and directing, recruiting, and retaining the staff necessary to meet the Institute's research computing and data goals; achieving the mission, vision, and strategy for the further development of collaborative Research Computing Infrastructure and Data (RCID) services; and setting the strategic direction for operational effectiveness and long-term sustainability of RCID services with the goal of providing centralized delivery and support for many of MIT's research computing capabilities.

A full position description is available at https://www.dropbox.com/s/k9co05d5b0k8d8x/ORCD%20Executive%20Director%206-10-22%20Final.docx?dl=0.

Job Requirements

REQUIRED: bachelor's degree; ten years' relevant work experience that includes at least four directly leading and supervising a team of full-time staff members; understanding of large-scale advanced research computing environments, including expertise with public cloud providers, infrastructure as code, automation, visualization, and data analytics; expert-level knowledge of multiple cloud providers (AWS, Azure or GCP); experience using public cloud for data science, data integration, or machine learning use cases and expertise with scripting and automation languages and tools; direct experience with research data management, tiered storage lifecycle strategies, and usage/cost optimization; knowledge of batch processing, cluster management, and cloud orchestration techniques; familiarity with high-performance computing trends/technologies; knowledge of file systems architecture; ability to collaborate and build consensus/relationships, influence others, and move toward a common vision/goal; excellent communication skills; ability to navigate a complex academic environment; skill leading a high-performing team; service orientation. *PREFERRED*: master's degree and a record of engaging with academic researchers. **Job**

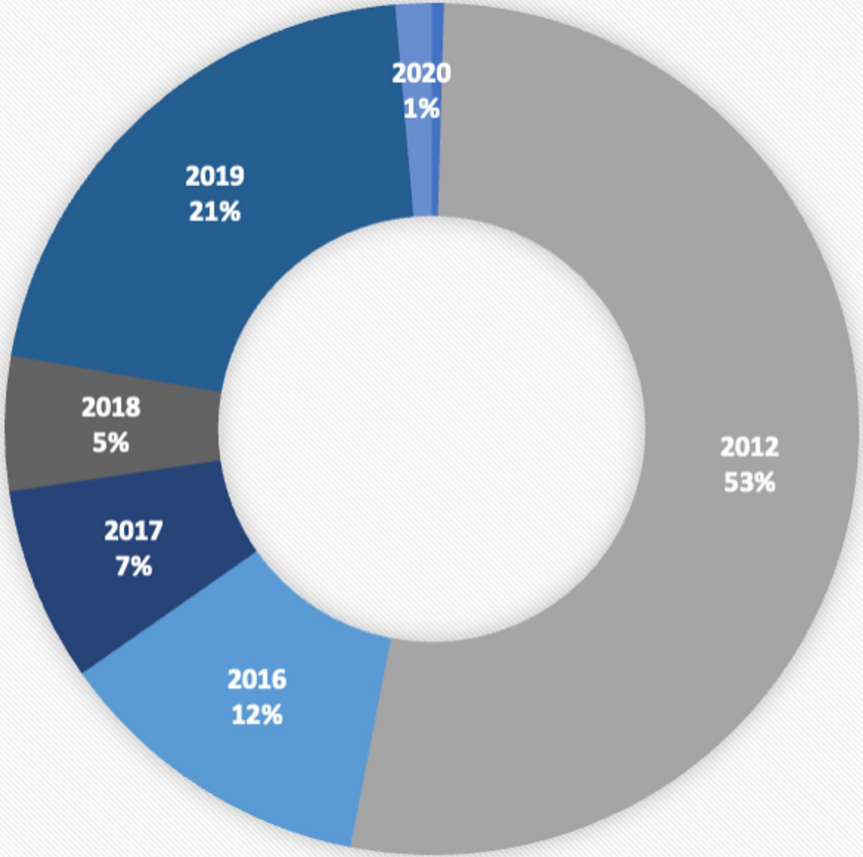
orcd@mit:~# uptime -p

Engaging-1 – an example shared ORCD cluster

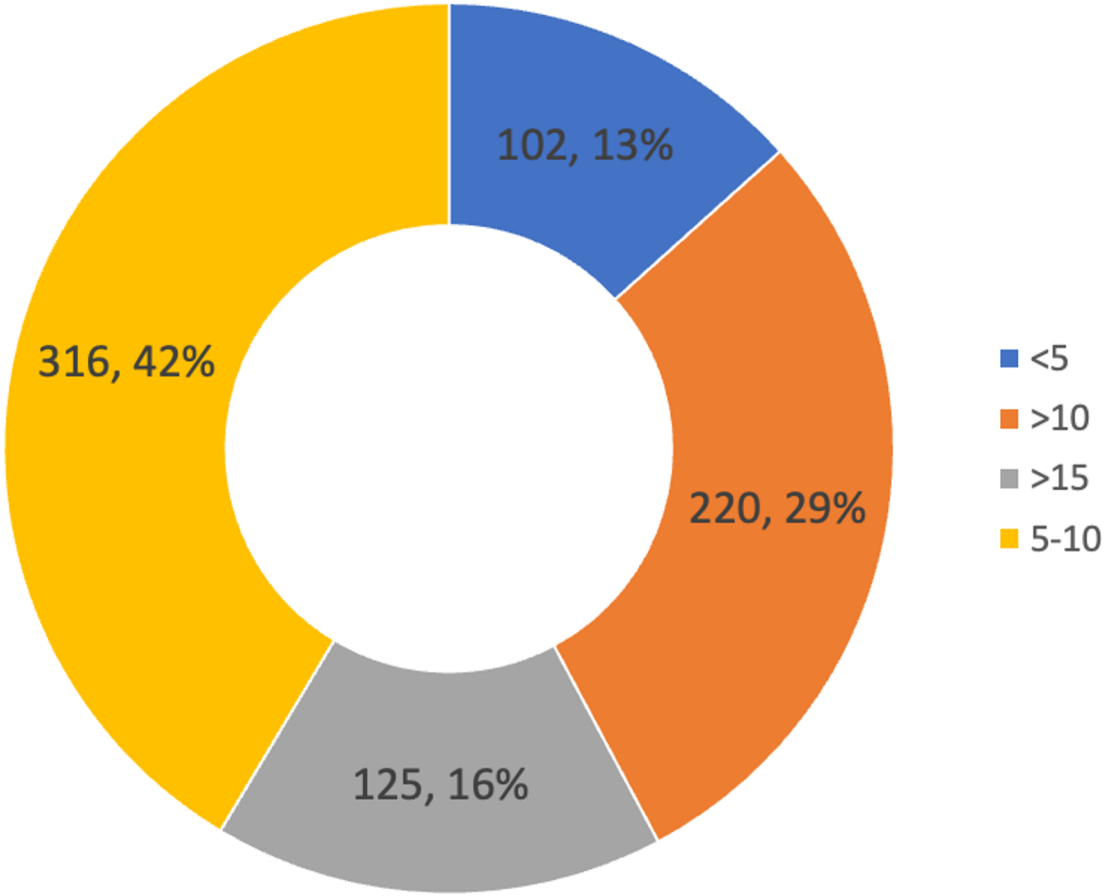
As of Q4 2022, there now exist over **1,463** machines in the computer cluster called “Engaging-1”

Of those **1,463** machines, **1,097** are currently available for processing. Of those **1,097** available, **797** are between 5 and 11 years old, **300** are less than 5 years old, and **244** remain under some form of warranty and are available for repair.

Of the total machines available, those considered *viable* for modern computing represent less than **16%** of the total available fleet, after taking into consideration damaged, out of warranty, legacy and currently unavailable systems.

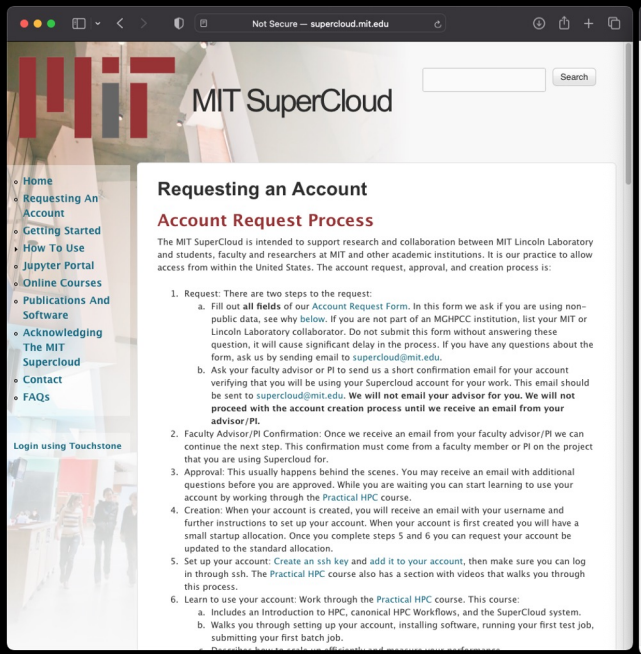


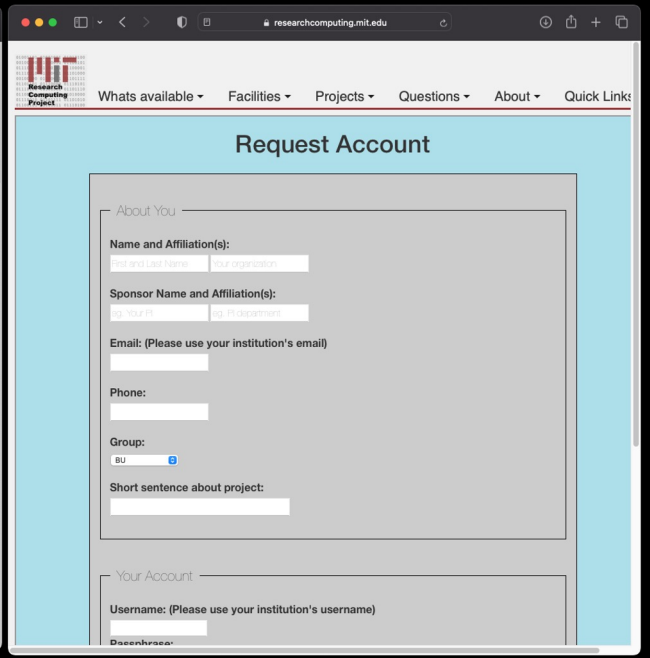
orcd@mit:~# !!

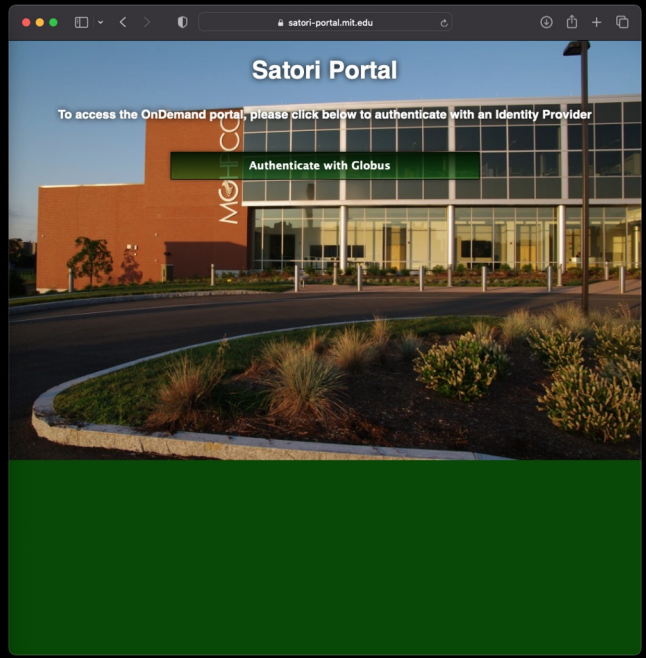


(Yes, “!!” is nerd for “run command again”)

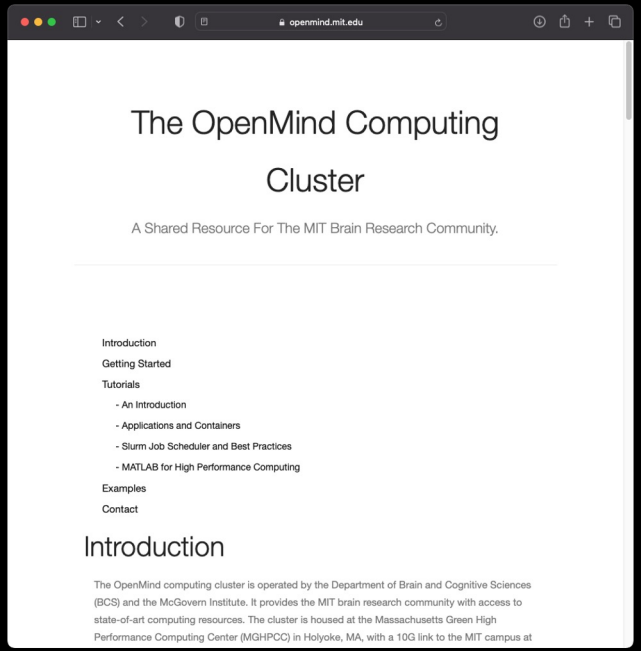
orcd@mit:~# sinfo -a

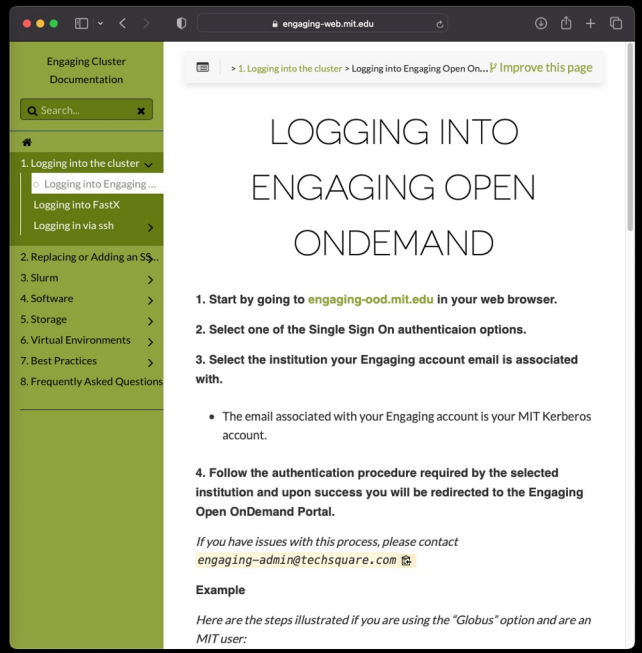


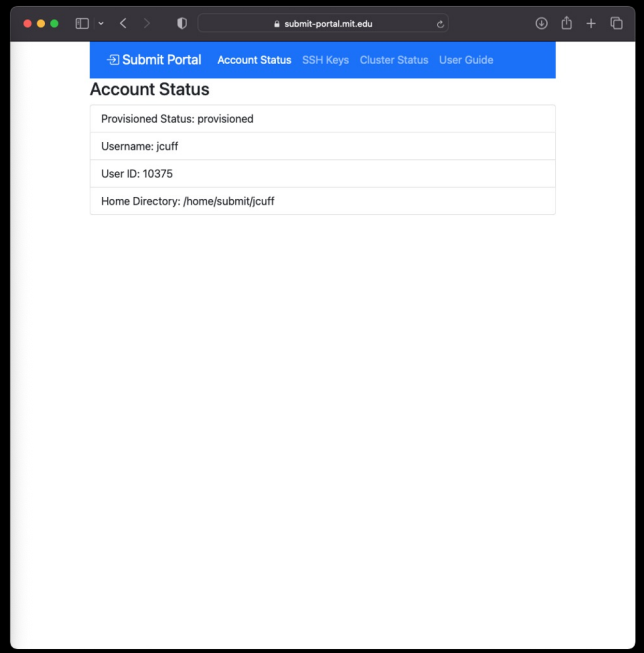




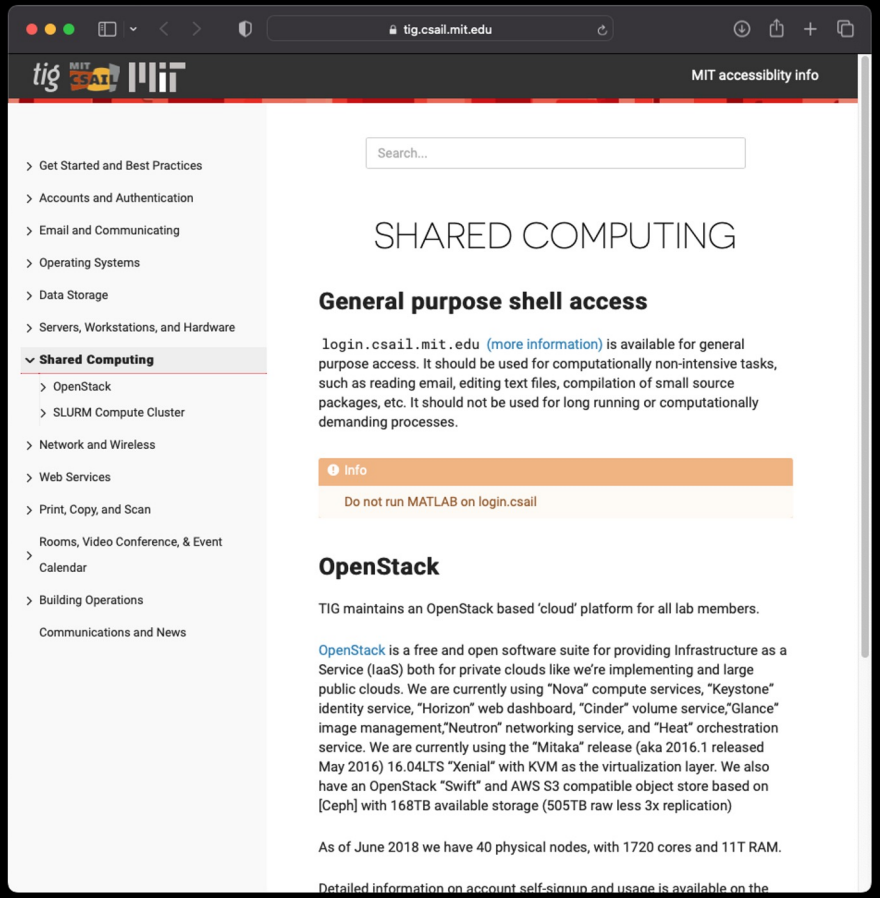
orcd@mit:~# !!

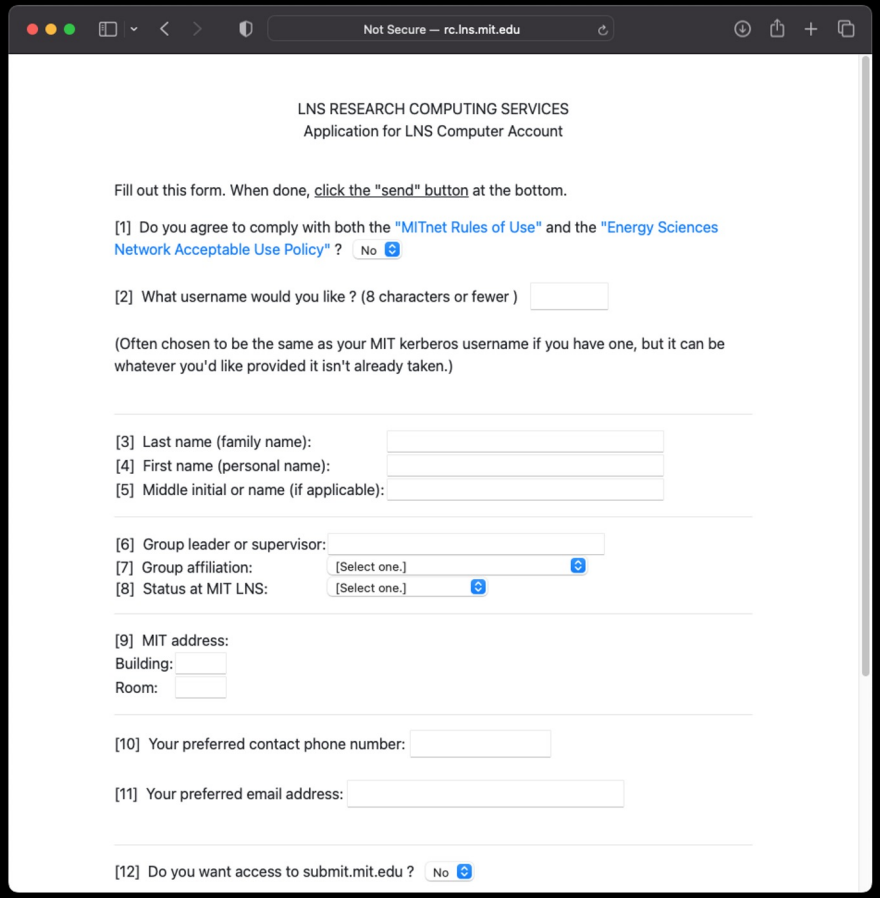

openmind.mit.edu
The OpenMind Computing Cluster
A Shared Resource For The MIT Brain Research Community.
Introduction
Getting Started
Tutorials
- An Introduction
- Applications and Containers
- Slurm Job Scheduler and Best Practices
- MATLAB for High Performance Computing
Examples
Contact
Introduction
The OpenMind computing cluster is operated by the Department of Brain and Cognitive Sciences (BCS) and the McGovern Institute. It provides the MIT brain research community with access to state-of-art computing resources. The cluster is housed at the Massachusetts Green High Performance Computing Center (MGHPCC) in Holyoke, MA, with a 10G link to the MIT campus at


engaging-web.mit.edu
Engaging Cluster Documentation
Search...
1. Logging into the cluster
Logging into Engaging ...
Logging into FastX
Logging in via ssh
2. Replacing or Adding an SS...
3. Slurm
4. Software
5. Storage
6. Virtual Environments
7. Best Practices
8. Frequently Asked Questions
> 1. Logging into the cluster > Logging into Engaging Open On...
Improve this page
LOGGING INTO ENGAGING OPEN ONDEMAND
1. Start by going to engaging-ood.mit.edu in your web browser.
2. Select one of the Single Sign On authenticaion options.
3. Select the institution your Engaging account email is associated with.
The email associated with your Engaging account is your MIT Kerberos account.
4. Follow the authentication procedure required by the selected institution and upon success you will be redirected to the Engaging Open OnDemand Portal.
If you have issues with this process, please contact engaging-admin@techsquare.com
Example
Here are the steps illustrated if you are using the "Globus" option and are an MIT user:


submit-portal.mit.edu
Submit Portal
Account Status
SSH Keys
Cluster Status
User Guide
Account Status
Provisioned Status: provisioned
Username: jcuff
User ID: 10375
Home Directory: /home/submit/jcuff

orcd@mit:~# !!


tig.csail.mit.edu
MIT accessiblity info
Search...
Get Started and Best Practices
Accounts and Authentication
Email and Communicating
Operating Systems
Data Storage
Servers, Workstations, and Hardware
Shared Computing
OpenStack
SLURM Compute Cluster
Network and Wireless
Web Services
Print, Copy, and Scan
Rooms, Video Conference, & Event Calendar
Building Operations
Communications and News
SHARED COMPUTING
General purpose shell access
login.csail.mit.edu (more information) is available for general purpose access. It should be used for computationally non-intensive tasks, such as reading email, editing text files, compilation of small source packages, etc. It should not be used for long running or computationally demanding processes.
Info
Do not run MATLAB on login.csail
OpenStack
TIG maintains an OpenStack based 'cloud' platform for all lab members.
OpenStack is a free and open software suite for providing Infrastructure as a Service (IaaS) both for private clouds like we're implementing and large public clouds. We are currently using "Nova" compute services, "Keystone" identity service, "Horizon" web dashboard, "Cinder" volume service,"Glance" image management,"Neutron" networking service, and "Heat" orchestration service. We are currently using the "Mitaka" release (aka 2016.1 released May 2016) 16.04LTS "Xenial" with KVM as the virtualization layer. We also have an OpenStack "Swift" and AWS S3 compatible object store based on [Ceph] with 168TB available storage (505TB raw less 3x replication)
As of June 2018 we have 40 physical nodes, with 1720 cores and 11T RAM.
Detailed information on account self-signup and usage is available on the


Not Secure — rc.lns.mit.edu
LNS RESEARCH COMPUTING SERVICES
Application for LNS Computer Account
Fill out this form. When done, click the "send" button at the bottom.
[1] Do you agree to comply with both the "MITnet Rules of Use" and the "Energy Sciences Network Acceptable Use Policy" ? No
[2] What username would you like ? (8 characters or fewer)
(Often chosen to be the same as your MIT kerberos username if you have one, but it can be whatever you'd like provided it isn't already taken.)
[3] Last name (family name):
[4] First name (personal name):
[5] Middle initial or name (if applicable):
[6] Group leader or supervisor:
[7] Group affiliation: [Select one.]
[8] Status at MIT LNS: [Select one.]
[9] MIT address:
Building:
Room:
[10] Your preferred contact phone number:
[11] Your preferred email address:
[12] Do you want access to submit.mit.edu ? No

orcd@mit:~# !!

svante.mit.edu

Svante
Updated 11/16/22

Search docs

CONTENTS:

» 2. General Information View page source

2. General Information

2.1. How to log into Svante

To login into the Svante cluster, from a terminal window on your local computer, type:

```
ssh -Y «username»@svante-login.mit.edu
```

a password will be requested; use your athena password (your svante «username» above will be the same your athena «username»). After you log in, you should receive a terminal prompt with your local working directory set to `/home/«username»`; you will logged into the Svante cluster login (or "head") node `svante-login`. See Section 6 for discussion of proper uses of the head node, file server nodes, and compute nodes.

2.2. `/home` spaces

`/home/«username»/`: we have about 100 terabytes (TB) of total 'home space' for general purpose usage, source code, plots and figures, model builds, etc. Every svante user is given disk space here; quotas on home space are 500 gigabytes (GB) per user. `/home` is backed up daily (offsite) and protected from disk failure via a RAID array. Svante `/home` space is mounted to all nodes in the cluster. Home space is not intended as a repository for large (or large numbers of) data files for analyses, or to be used as disk space for large model runs.

2.3. File servers

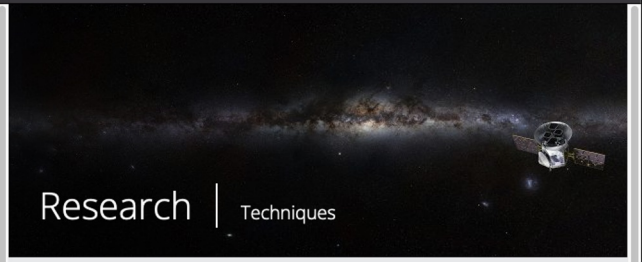
Svante file servers are named fsxx (see Table 2.1), currently `fs01` - `fs11`, with total capacity presently roughly 4 PB. To get onto a file server node, for example, typing

```
ssh fs02
```

(from `svante-login` or any other node in the Svante cluster) will give you a shell on file server `fs02`. Once you `ssh` to a file server, local disks should be accessed as `/d0`, `/d1`, `/d2`, and `/d3`; note some partitions not present on all fileservers. From all other nodes, this space can be reached through 'remote' mounts in which your access paths would be, for example, `/net/fs02/d0` to access the `/d0` partition on `fs02`. Storage on these file servers is for runs, experiments, downloaded data, etc. that will be accessed and kept for longer-term periods, and these spaces are backed up (weekly offsite backup, although in periods of heavy use, it may take longer for backups to complete). These machines can also be accessed externally (i.e. from outside svante): `fs02` is `svante2.mit.edu`, `fs03`

space.mit.edu

MIT Kavli Institute for Astrophysics and Space Research



High Performance Computing

High performance computing (HPC) is a necessary component of modern astrophysics research. MKI has used 3 HPC clusters over the past decade. The first cluster has been used by the Chandra space telescope for data analysis and modeling associated with the HETG instrument contract. It was retired in October 2013.

The second cluster was used by the LIGO project, and was recently retired. As LIGO undergoes upgrades to Advanced LIGO sensitivity, the computing and analysis infrastructure will be completely overhauled.

The third cluster was originally built by Professors Edmund Bertschinger and Scott Hughes, and largely used by their research groups. It was significantly upgraded in 2008, replacing the original 32-bit compute nodes with 64-bit nodes. It has been further upgraded to use CentOS Enterprise Linux with SLURM workload management software. Slurm is an open source, fault-tolerant, and highly scalable cluster management and job scheduling system. The HPC cluster currently consists of 51 compute nodes and one master node, plus 4 storage servers providing researchers 1,048 CPU processors (cores) with 2,832 GB memory and over 1 PB data storage.

orcd@mit:~# setenforce 1

Federal Bureau of Investigations

MIT, Harvard, etc.

Risk profiles

Academic integrity

Nation state actors

Soft underbelly of Higher Education

orcd@mit:~# mghpcc -list -orcd -today

64 empty racks
@ ca. 50KW
Liquid cooled racks

2.4-3.2 MW of power
$3MM spent

In place

$0 for network…

orcd@mit:~$ play hua_rong_dao.wav

In the year 208 Cao Cao led 220,000 troops of the Wei army to fight against an army of 50,000 Shu troops in a mountainous area near Chibi in today's Hubei province. Because of some strategic errors, Cao Cao's troops were badly defeated in the Battle of Chibi, and he fled with only a handful of his soldiers.

The opening at the bottom of the board is Huarong Pass. Initially the blocks are arranged as shown here, with Cao Cao's block trapped by the other nine. The player's job is to slide blocks horizontally and vertically so that Cao Cao can eventually escape through the pass.

(Hint: The game can be solved in 81 moves of 25,955)

orcd@mit:~# cat /etc/motd

Do no harm

Continue investigations and discovery

Complete work on capital planning, MOUs and space planning

Staffing planning, develop budget and org chart

Explore some “cloudy” bridging options

Run many, many Hua Rong Dao simulations

orcd@mit:~# finger orcd

```
Login: orcd                         Name: Office of Research Computing
Directory: /mit/vpr/orcd            Shell: /bin/tcsh
Last login Thur Jan  5 04:56 2023 (EDT) on pts/0 from 192.168.99.5
New mail received Fri Jan  6 05:25 2023 (EST)
     Unread since Fri Jan  6 05:24 2023 (EST)

Plan:
Build a great, sustainable, reliable and professional “One MIT”
research computing experience for faculty, staff and students

URL:                    https://orcd.mit.edu
RESOURCES:              https://orcd.mit.edu/orcd-resources
EMAIL:                  jcuff@mit.edu; fisherp@mit.edu
MAILING LIST:           orcd-admin@mit.edu
```

```
orcd@mit:~$ echo ‘echo anyone?’ \
> questions.sh
```

```
orcd@mit:~$ while true; \
do sh ./questions.sh; done
```

Did anyone spot the root commands?