

Home

TTS Research Computing Resources



For additional information, please contact Lionel Zupan, Director of Research and Geospatial Technology Services (RGTS), at x74933 or via email Lionel.Zupan@Tufts.edu.

Tufts Technology Services(TTS) Research computing options

- High-performance computing research cluster
- Bioinformatics server
- Research Storage
- Visualization Center
- GIS Center

See [overview](#) for additional information about TTS Research and Geospatial Technology Services.

1. Tufts High-performance computing research cluster

What is a Cluster?

Cluster computing is the result of connecting many local computers (nodes) together via a high speed connection to provide a single shared resource. Its distributed processing system allows complex computations to run in parallel as the tasks are shared among the individual processors and memory. Applications that are capable of utilizing cluster systems break down the large computational tasks into smaller components that can run in serial or parallel across the cluster systems, enabling a dramatic improvement in the time required to process large problems and complex tasks. For more specifics check IBM's website for IDataPlex and BladeCenter systems.

Typical Cluster Usage at Tufts

Faculty, Research Staff and students use this resource in support of a variety of research projects. [See how.](#)

Tufts Linux Research Cluster

Tufts Technology Services **TTS** provides a wide array of services in support of Tufts research community including a High Performance Computing IBM research cluster (IBM HS21 and HS22 blades, M3 and M4 iDataplexes, nVidia GPUs, 10Gb/s interconnect) with 173 compute nodes, 2100 cores and a peak performance of roughly 52 Teraflops (47 from CPU and 5 from GPU). In this HPC environment, TTS also provides researchers with access to commercial and open-source software applications, tools for bioinformatics, and networked and secure storage for research data (400+ TB CIFS and NFS storage on NetApp appliances). Finally, TTS also maintains a Center for Scientific **Visualization**. Tufts is connected via 10Gb links to Internet2 sites.

Each cluster node has 8, 12 or 16 cores with several generations of Intel CPUs and 16, 24,32,48, 96 or 128 gigabytes of memory. This includes the recent addition of 60 faculty contributed IBM iDataplex M4 nodes by Prof. Lin from Chemistry.

The Linux operating system on each node is configured identically across every machine. In addition there is a login node and a management node supporting the compute nodes. Client/user workstations access the cluster via the Tufts Network using ssh based connection client

software. The user login node has an additional network interface that connects to the compute nodes using private IP addressing via 10Gig network hardware. This scheme allows the compute nodes to be a "virtual" resource managed by the **LSF** queueing software. This approach also allows the cluster to scale to a large number of nodes thus providing the structure for future growth. The login node of the cluster is reserved for the use of compilers, running shell tools, and launching and submitting programs to compute nodes. The login node is not intended for long running programs, or for general computing purposes, and all jobs are to be submitted to compute nodes using **bsub** via various LSF submission queues.

New Cluster Specific Information

Click [New Cluster](#).

Grant Applications related information

Content to support applications can be found [here](#).

Cluster User Accounts

Click [Account Information](#) for additional information about cluster accounts.

Orientation for new cluster users

This [content](#) is for someone that has never used linux or mainframes or super computing centers.

Research Cluster Restrictions

Conditions and use of the research cluster include and are not limited to the following expectations. Additional related details may be found throughout this page.

Expectations
no user root access
supported OS is RedHat 6 Enterprise version
no user ability to reboot node(s)
all cluster login access is via the login headnode
no user machine room access to cluster hardware
no alternative linux kernels other than the current RED HAT version
no access to 10Gig Ethernet network hardware or software
no user cron or at access
no user servers/demons such as: HTTP(apache), FTP, etc.
Cluster quality of service is managed through LSF queues and priorities
all user jobs destined for compute nodes are submitted via LSF's bsub command
all compute nodes follow one naming convention
only Tufts Technology Services NFS approved storage is supported
unused contributed node CPU time reverts to cluster user community
no user contributed direct connect storage such as usb memory, or external disks
only limited outgoing Internet access from the headnode will be allowed; exceptions must be reviewed
allow approximate 2-week turn around for software requests
whenever possible, commercial software limit to the two most recent versions
Only user home directories and optional research NFS mounted storage is backed up
temporary public storage file systems have no quota and are subject to automated file deletions

Cluster does not export file systems to user desktops
Cluster does not support Virtual Machine instances

Software request policy

Please send your request via email to cluster-support@tufts.edu and address the following questions:

- What is the the name of the software?
- Where can additional information about the software be found?
- Who are the intended users of the software?
- When is it needed by?
- Will it be used in support of a grant and if so what grant?
- What if any special requirements are needed?

Note: A software request normally may take up to 2 weeks. However depending on the installation complexity and number of packages requested it may take longer. When it appears that an assessment of the tasks suggest longer than 2 weeks we will contact you with an estimate so that prioritization can be made.

Recent Cluster News

Click [News](#)

Cluster Storage Options

Click [here](#) for details.

Network Concurrent Software Licenses

[Click here](#)

Support venue

If you have any questions about cluster related usage, applications, or assistance with software, please contact cluster-support@tufts.edu.

MODULES: Cluster software environment

[Click here](#)

Installed Cluster Software

[Click here](#)

Compilers, Editors, etc...

[Click here](#)

Frequently Asked Questions - FAQs:

Cluster Connections/Logins

[Click here](#)

Parallel programming related information

[Click here](#)

User Account related FAQs:

[Click here](#)

X based graphics FAQs

[Click here](#)

Application specific Information FAQs

[Click here](#)

Linux and LSF information FAQs

[Click here](#)

Compilation FAQs

[Click here](#)

Miscellaneous FAQs

[Click here](#)

How do Tufts students and faculty make use of the cluster?

[See How](#)

2. Bioinformatics services

A separate server is used to support these services in some cases. However some software may require installation on the linux research cluster. Check the [Installed Software](#) for Bioinformatic software available on the cluster. To make a special request for software installation, please follow the instructions as noted elsewhere on this page.

Emboss services can be found [here](#)

3. Tufts Center for Scientific Visualization (or VisWall)

A description may be found [here](#). The user guide is available [here](#). Also, visit the Tufts Visualization [Awards](#) program.

See the attached EduCause document on Visualization for general information and direction this technology provides. You may find it under this wiki page "Page Operations" section above.

The unique facilities at the Tufts Center for Scientific Visualization allow researchers to display 2D or 3D high-resolution images and video at 4K resolution (4096x2160) on a rear-projected, 15 by 8 foot screen. Users can also bring their personal laptop to the Viswall and display their images or video at a quarter of the maximum resolution, which is standard HD (2048x1080).

Applications are run from one of three operating systems: Red Hat Linux, Windows, or Mac OS X. The HP workstation, which runs Red Hat Linux 5 or Windows XP, has 8 Intel processor cores and 16GB of RAM. The Apple MacPro, which runs Mac OS X 10.6, has 8 Intel processor cores and 16GB of RAM. Visualization processing can be handled on one of these two machines, or remotely at Tufts' cluster computing facility which is connected to the Center for Scientific Visualization via 10 Gigabit Ethernet.

The Viswall has recently been upgraded to improve usability, which is now handled by a simple touchscreen interface. A toolkit is also being developed for the use of the Microsoft Kinect motion capture sensor at the Viswall, with the aim of providing additional resources to the visualization research community at Tufts.

The research cluster is available to VisWall users for additional computational resources. Current connectivity follows standard practices using ssh and x11 forwarding. Viswall users with a cluster account may forward cluster based application graphic output for display on the VisWall. As of mid-May 2009, a new and dedicated 10Gig fiber network connects the cluster nodes with the Viswall. This provides for a novel and very high speed connection for moving massive amounts of data.

If you intend to use this facility, a short mandatory training class is required. Contact via email cluster-support@Tufts.edu to make arrangements.

4. Tufts GIS Center

Tufts GIS Center and resources can be found [here](#).

Tufts GeoPortal

Many organizations and institutions are developing large spatial data repositories. Discovering and accessing these data sets pose many challenges. As a result, Tufts and Harvard are collaboratively developing an open source, federated web application to discover, preview, and retrieve geospatial data as part of global and national spatial data infrastructure. **The Open Geoportal** combines an intuitive, map-based search interface along with traditional text-based metadata search tools for rapid data discovery. Tufts instance of The Open Geoportal can be found [here](#).

Tufts Research Cluster indirectly supports GIS spatial statistical computation with the availability of modern spatial statistics programs as found in **R**. This is a useful resource when faced with either complex estimation tasks, long runtimes or access to more memory than is often available on desktop workstations. R programs such as the following are available:

fields, ramps, spatial, geoR, geoRglm, RandomFields, sp, spatialCovariance, spatialkernel, spatstat, spBayes, splancs,

For additional information please contact cluster-support@tufts.edu.