

Research Computing

Information for New and Junior Faculty,
Researchers and Graduate Students
August, 2016

Erin Shaw
Advanced Cyberinfrastructure, Research & Education Facilitator (ACI-REF)
USC Center for High-Performance Computing

USC ITS
Information Technology Services

University of Southern California

Information Technology Services (ITS)



Douglas Shook
USC CIO



Randolph Hall
USC VP of Research &
Faculty Executive
Director, HPC*



Maureen Dougherty
HPC Director

*The HPC Faculty Advisory
committee advises the CIO
about the faculty's research
needs related to the
university's HPC resources.



ITS Data Center

USC ITS
Information Technology Services

University of Southern California

ITS Research Computing Initiative

- **Software**
 - Campus-wide software licenses
- **Storage**
 - Data storage options
- **High-Performance Computing**
 - Overview and resources
- **HPC Support and Outreach**
 - Faculty support, outreach, and education

USC ITS
Information Technology Services

University of Southern California

Software

USC ITS
Information Technology Services

University of Southern California

Statistical Computing Software

- **Campus-wide licenses**
 - itservices.usc.edu/stats
 - Available for low/no-cost through USC campus-wide licenses
 - SAS and STATA installed on HPC
 - Available for free through your HPC account
- **Open source on HPC**
 - R Statistical computing software
 - Python scientific computing packages
 - These software packages are pre-installed on HPC

USC ITS
Information Technology Services

University of Southern California

Technical Computing Software

- **Technical software available for download**
 - MATLAB
 - itservices.usc.edu/matlab
 - Mathematica
 - itservices.usc.edu/mathematica
- **GIS and remote sensing software**
 - spatial.usc.edu/index.php/software
 - ArcGIS and CityEngine
 - TerrSet Constellation
 - Data/Geoportal
- **Software available on HPC**
 - MATLAB and Mathematica
 - Compilers: **Gnu, Intel, PGI**
 - Numerical Libraries: **mpich, Openmpi, CUDA, Fftw, Petsc**
 - Molecular Simulation: **NAMD, Gromacs, Amber**
 - Quantum Chemistry: **Gaussian '03, QChem**

USC ITS
Information Technology Services

University of Southern California

Qualitative Data Analysis Software

- **Data analysis**
software.usc.edu
 - ATLAS.ti
 - NVivo
- **Data collection**
itservices.usc.edu/qualtrics
 - Qualtrics

USC ITS Information Technology Services University of Southern California

Bioinformatics Software
(nml.usc.edu/bioinformatics/#software)

- **Comprehensive functional analysis and advanced literature/data mining**
 - Partek Genomics Suite
 - Partek Flow
 - Galaxy
 - Vector NTI Advance
 - Microarray R US
- **Statistical analysis of high-throughput data and DNA/protein sequence analysis**
 - Ingenuity Pathway Analysis
 - BIOBASE
 - Nextbio
 - OncoPrint
 - GeneInvestigator
 - Quetzal

USC ITS Information Technology Services University of Southern California

Research Management Software
(Available through other sources to USC community)

- **Citation management**
 - RefWorks, Mendeley, EndNote, and Zotero
 - libraries.usc.edu/research/citation-management
 - norris.usc.libguides.com/CitationManagement
- **Project management**
 - RedCap
 - redcap.sc-ctsi.org

USC ITS Information Technology Services University of Southern California

UNIX Computing Utilities
(itservices.usc.edu/sftp, itservices.usc.edu/unix)

- **UNIX computing utilities**
 - Secure file transfer
 - FileZilla, Fetch, SCP
 - Secure shell
 - XWin-32

(Pre-configured with connections to USC networks, including HPC)
- **UNIX documentation**
 - Accounts, logging in, tutorials, etc.
 - itservices.usc.edu/unix

USC ITS Information Technology Services University of Southern California

Storage

USC ITS Information Technology Services University of Southern California

Cloud Data Storage

- **MS Office 365/OneDrive for Business**
 - 5 TB capacity
 - Approved for storing some restricted information, including HIPAA- and FERPA-related information
 - All faculty and staff with ITS-supported O365 accounts receive access
 - Graduate students that need access should inquire with their respective schools
- **Google Drive**
 - Unlimited capacity
 - Not approved for storing HIPAA-related information
 - All faculty, staff, and students receive access

USC ITS Information Technology Services University of Southern California

USC Data Storage

- **USC Digital Repository**
 - repository.usc.edu
 - 40 PB of total storage capacity
 - High-speed, high-performance disk storage, and low-demand tape storage
 - Available for Windows, Mac, and Linux (including HPC)
 - Highly secure and encrypted cloud storage (Approved for storing HIPAA-related information)
 - Fee-based usage policy
- **USC Drive**
 - itservices.usc.edu/uscdrive
 - High-speed storage that allows users to upload large data files quickly
 - For faculty and graduate students
 - Initial storage space of 50 GB per person
 - Additional storage can be purchased
 - Not approved for storing HIPAA-related information

USC ITS Information Technology Services University of Southern California

HPC Data Storage

- Group projects can request up to 2 TB of disk space
 - Backed up daily
- Users have access to a 328 TB /staging directory
 - No data backup
 - Cleaned during semi-annual downtimes or when capacity reached
- Users also have access to temporary directories
 - /tmp and /scratch while running job on compute node(s)
 - Cleared at end of each job
- HPC is not approved for storing HIPAA information but compliancy is under development

USC ITS Information Technology Services University of Southern California



HPC Condo Data Storage

HPC allows PIs with active HPC accounts to purchase storage on its condominium-style compute environment.

- HPC obtains approved vendor quotes for storage arrays based on user needs.
- All storage requires a mandatory minimum 3-year MOU with 3 years of vendor support.
- HPC provides limited backup of data (up to 100 TB) at no cost to the condo owner.
- The condo storage is only available on the HPC cluster; the file system is not shared outside of HPC.

USC ITS Information Technology Services University of Southern California

High-Performance Computing



High-Performance Computing (HPC)

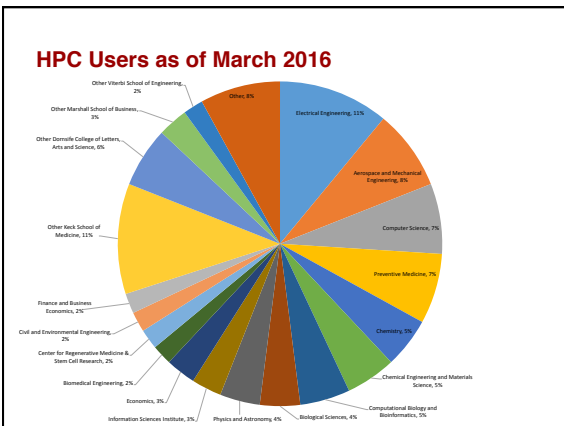
HPC advances USC's mission by providing the infrastructure and support necessary for research computing

- Many of HPC's resources are available at no charge to USC faculty, researchers, and graduate students
- As a part of USC's Information Technology Services (ITS) division, HPC is housed within the ITS data center and is monitored around-the-clock by ITS staff

World-class super-computing center


- LINPACK benchmarking, published by Top500 Supercomputers
- 12th fastest academic supercomputer in U.S., June 2016 listing



HPC Resources

- **Compute cycles on two low-latency bandwidth Linux clusters**
 - Infiniband 56-gigabit (1,025 nodes)
 - Myricom 10-gigabit (1,741 nodes)
- **Disk arrays**
 - Part of a storage area network presented via networked file systems from multiple file servers
- **Data and disk resources**
 - Limited storage of research data and output from cluster-run jobs
- **Available to faculty, researchers, grad students, and iVIPs**



USC ITS
Information Technology Services

University of Southern California

Condo Compute Nodes

HPC allows PIs with active HPC accounts to purchase compute nodes on its condominium-style compute environment

- Condo nodes can be purchased through HPC at any time, as long as nodes are available in our existing pool of nodes
 - Only HPC-purchased nodes are available for condo
- Condo nodes are priced at the amount that HPC purchased them, including taxes
 - This is a one time fee
 - Condo fees do not include infrastructure, which is covered by HPC
- Contact hpc@usc.edu for current configuration and pricing information

USC ITS
Information Technology Services

University of Southern California

HPC Support and Outreach

USC ITS
Information Technology Services

University of Southern California

HPC Support, Facilitation, and Collaboration

- Proposal Support**
 - Help build startup packages, provide quotes for condo purchases, and write letters of support for grant submissions
 - Provide a description of ACI resources and support for proposals
 - hpc.usc.edu/files/2016/01/Support-for-Grant-Applications1.pdf
- ACI-REF (Facilitation)**
 - HPC director is co-PI of NSF-sponsored, 6-university grant
 - Advancing scientific discovery through a national network of Advanced Cyberinfrastructure Research and Education Facilitators (ACI-REFs)
- Collaboration**
 - Collaborates on HPC projects with researchers on all USC campuses

USC ITS Information Technology Services University of Southern California

HPC Outreach, Education, and Assistance

Outreach and Education <ul style="list-style-type: none"> New faculty presentations Classroom presentations Workshops <ul style="list-style-type: none"> Intro Linux, Intro to HPC Advanced HPC, Parallel Matlab Pegasus workflows Sponsor guest talks <ul style="list-style-type: none"> Mathworks, NVIDIA ITS data center tours 	Direct Assistance <ul style="list-style-type: none"> Email hpc@usc.edu Attend office hours <ul style="list-style-type: none"> Tuesdays@2:30 (UPC LVL) Mondays@2:30 (HSC NML) Request consultation <ul style="list-style-type: none"> Lab consultations Individual meetings
---	---

USC ITS Information Technology Services University of Southern California

The screenshot shows the HPC website interface. At the top, there's a navigation menu with 'About', 'Research', 'User Support', 'New User Guide', 'Frequently Asked Questions', and 'Contact Us'. A search bar is on the right. The main content area features a news article titled 'National Science Foundation Award To Benefit USC Researchers' with a photo of a group of people. Below the article are four staff member portraits with their names: Lowell Stott, Frank Alber, Veronica Ellason, and Arieh Warshel. A 'Popular Topics' sidebar on the right lists items like 'Applying for an HPC Account' and 'Managing and Renewing Your Account'.
