Center for High-Performance Computing
NEWS
THE UNIVERSITY OF UTAH
SPRING 2015

Research

CI-WATER Provides Insights into Water Futures in the Intermountain West

By Court Strong, Steven Burian, Steve Corbató, Laura Hunter, and Jason Scalzitti

The availability of clean and affordable water is an increasingly pressing environmental sustainability challenge across the western United States as a result of both population growth and climate change. An interdisciplinary team of Utah and Wyoming university researchers is collaborating on a four-year, $6 million project to develop a better understanding of the interconnectivity of natural and human water systems and how these interactions will change in the future.

Funded through the National Science Foundation's EPSCoR program, the CI-WATER project (see http://ci-water.org) has enabled the team to develop new computer models and obtain new high performance computational resources – cyberinfrastructure – to support the simulation and investigation of how increased population, varying land use patterns, and changes in climate will impact water storage and availability in the Intermountain West. In Utah, the collaborative team draws from the three research universities – BYU, USU, and the University of Utah. Here at the U, the team includes members from Atmospheric Sciences (Court Strong, lead), Civil and Environmental Engineering (Steve Burian), the Utah Education Network (Laura Hunter), and CHPC (Steve Corbató).

One of CI-WATER's computationally intensive components made possible by CHPC is the simulation of climate at spatial and temporal scales adequate for studying future changes in water availability. Adam Kochanski and Court Strong have adapted the Weather Research & Forecasting (WRF – see http://www.wrf-model.org for more details) model for simulation of climate in Utah and Wyoming on a 4-km horizontal grid (high-resolution for climate studies spanning a century). They added a lake model to WRF to simulate important effects from the Great Salt Lake and also incorporated an urban irrigation scheme to provide realistic variations in soil moisture resulting from watering of commercial and residential lawns. Figure 1 shows climate change indicated by their most recently completed WRF simulations for Representative Concentration Pathway

(Continued on Page 2)

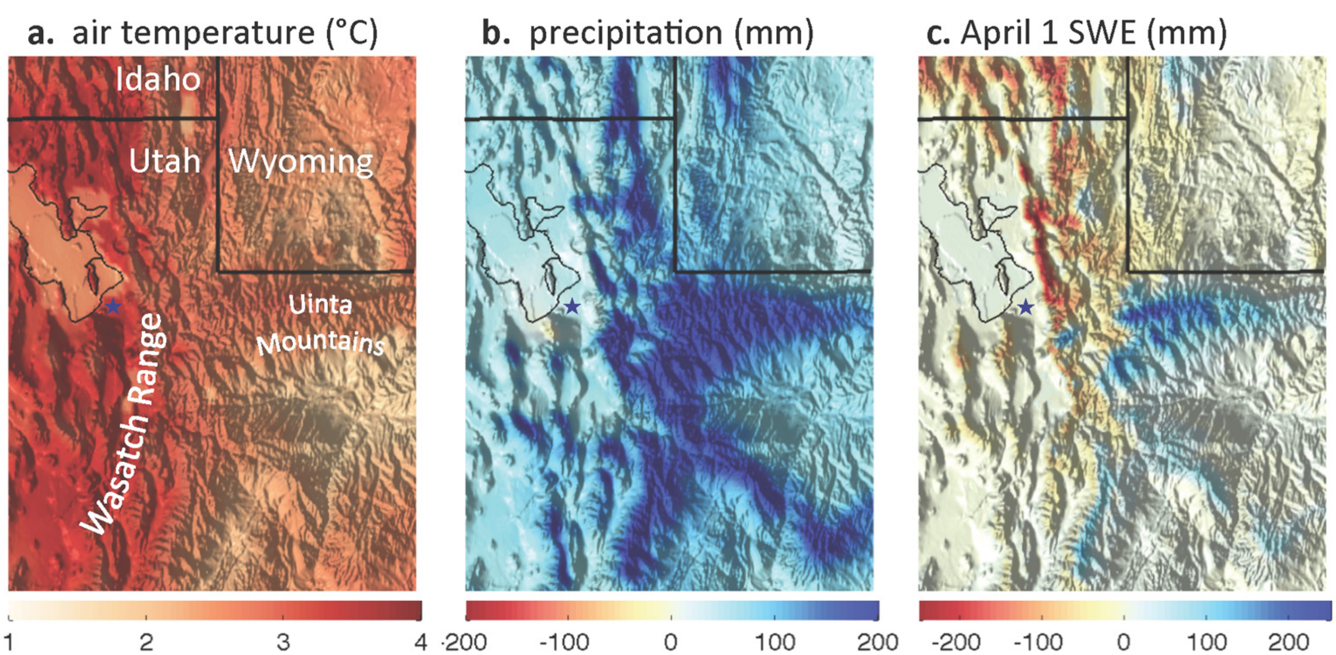


Figure 1. Changes in climate by the end of the current century, meaning results from a simulation of years 2085-2094 minus a historical reference simulation of years 1985-2004. In each panel, the Great Salt Lake is outlined toward the middle left of the map with Salt Lake City indicated by a star to its southeast. Bold black lines are state boundaries, and gray shading represents illumination of the topography from the southwest. The changes shown in each panel correspond to the following: a. October through March air temperature two meters above the ground, b. October through March total precipitation (depth of water falling as rain plus depth of water that would result if falling snow were theoretically melted to water), and c. snowpack on April 1 expressed as a snow water equivalent (SWE; the depth of water that would result if the April 1 snowpack was theoretically melted).

(Continued from Page 1)

(RCP) 6.0, which is a moderate greenhouse gas emission scenario. By the end of the current century, projected cool-season (October-March) air temperature increases of 4°C are common across much of Utah, especially west of the Wasatch Range and away from water bodies (red shading, Fig. 1a). A warmer atmosphere tends to be moister, and many climate models project increases in cool-season precipitation for northern portions of the Great Basin, especially at the higher elevations (blue shading, Fig. 1b). More water coming into the system would be good news for water managers, but warming means that the fraction of precipitation arriving as snow will decline. Thus, the snow on the ground will melt faster and earlier. As a result, the model projects only modest snowpack increases for the high Uinta Mountains and some of the highest peaks of the Wasatch Range (blue shading, Fig. 1c), with much of the lower elevations suffering snowpack reductions (red shading, Fig. 1c). The representation of the historical climate (1985-2004) in this comparison comes from a retrospective WRF simulation with the same configuration as the future projections, constrained by a reanalysis of past atmospheric conditions. An ongoing analysis of these historical data not only indicates a good overall model performance with respect to past observations but also seeks to synthesize the scope and source of noted biases.

Figure 2. The CI-WATER storage array housed in the research computing area within the University's Downtown Data Center. Photo by Sam Liston

The WRF model uses a nested domain configuration, meaning that the high-resolution representation of Utah and Wyoming is embedded in two progressively coarser simulation domains extending west to encompass the Pacific Ocean and east to cover the Central Plains for a total of 5 million spatial grid points. Horizontal boundary conditions for the coarsest domain are based on output from a global climate model run in advance at coarser resolution (the Community Climate System Model Version 4 developed by the National Center for Atmospheric Research or NCAR). Model development, testing, and preparation of boundary conditions were performed on CHPC resources, and decade-long simulations were performed on the National Center for Atmospheric Research's 1.5-Petaflop, 72,576-processor Yellowstone supercomputer (see www2.cisl.ucar.edu/resources/Yellowstone). WRF is written in parallelized FORTRAN. Running with a 13-second time step on 624 cores, it takes approximately one week of wall clock time to simulate a year of climate. The hourly model output is thus valuable, and rather large – approximately 2 Terabytes (TB) per year. CHPC has installed a 1.0-Petabyte storage array (1,000 TB) to support the storage needs of CI-WATER's atmospheric and hydrological researchers, including the input and output files associated with the climate simulations. The climate simulations also are backed up on a tape archiving system.

One extension of the simulation results is the study of changes in climate and land use on the reliability, resiliency, and vulnerability of Salt Lake City's water supply. Steve Burian and Erfan Goharian of the Department of Civil and Environmental Engineering use their colleagues' climate simulation results to drive an integrated water system model representing the water supply, stormwater drainage, and wastewater collection infrastructure in the Salt Lake City metropolitan area. The integrated water model has been developed through a 2-year collaboration with Salt Lake City Department of Public Utilities facilitated by Tim Bardsley of Western Water Assessment. Stakeholder participation helped to create and validate the integrated water model and to design the future scenarios used in the study. The combined climate simulation, population growth, and water system response results show snowmelt trending earlier and with less yield than historical, leading to lowered reliability and resiliency, and higher vulnerability of the Salt Lake City water system. The team is currently working with Salt Lake City Public Utilities to evaluate a range of feasible management actions and infrastructure construction choices to maintain historical system performance.

Communicating the research and its impact has been a constant focus of the project. UEN produced ten videos that explain CI-WATER to the broader community, hosted annual symposia with the CI-WATER team and industry experts, and broadcast water week programs on UEN-TV. New curricula for high schools, hydrology toolboxes for teachers, and a multi-institution online hydroinformatics course have been lauded by evaluators and community partners. Campus partners for this include the Genetic Science Learning Center and Natural History Museum of Utah.

The CI-WATER research partnership has helped CHPC and the campus researchers whom it supports through the early deployment of Petascale data storage capabilities, testing of high-speed data transport, and the ability to work more closely with corresponding research computing support teams at USU, BYU, the University of Wyoming, and the NCAR-Wyoming Supercomputing Center.

CI-WATER's impact will expand this summer when the team shares its findings at the new National Water Center in Tuscaloosa, AL and their National Flood Interoperability Experiment to help build a new high resolution, near real-time hydrologic simulation and forecasting model for the United States that will sustain innovation into the future.

CHPC Users

The Move to SLURM for Batch Scheduling

By Anita Orendt

Recently CHPC moved to SLURM – Simple Linux Utility for Resource Management – for batch job scheduling. This transition coincided with our recent downtime and the start of the new quarter of allocation. The motivation for this move is both financial and practical. SLURM is an open source product, whereas the Moab/Torque that we were using is licensed. For the transition we have a SLURM support contract; however, there is no charge for the use of the resource manager and the support contract is not required for access to the scheduler. Also SLURM is widely used at other national and campus computational resources.

As part of this transition we are also moving to a new model for the accounting of usage. Previously the metric was one core hour equalled one SU on all clusters except kingspeak, where one core hour equalled 1.5SU. Moving forward, all usage will be reported in only core hours. Allocations will also be given in core hours.

As you transition to SLURM, you need to focus on two aspects of your batch scripts: the #PBS lines at the start; and, the PBS environmental variables that may be used in your script. In order to access the resources, you always need to specify the account and the partition, making sure that these two match, i.e., use an account that is allowed access to the partition. The partitions are cluster, cluster-freecycle, pi-cl, and cluster-guest, where cluster is the full name of the cluster and cl is the abbreviation of the cluster name (kingspeak and kp, ember and em, ash and ash, lonepeak and lp, apexarch and aa).

Accounts are usually your unix group name, typically your PI's lastname. If your group has owner nodes, the account is usually <unix_group>-<cluster_abbreviation> (where cluster abbreviation is =kp, lp, em, ash). There is also the owner-guest account; all users have access to this account to run on the owner nodes when they are idle. Jobs run as owner-guest are preemptable. The format to specify these is:

#SBATCH --account=<youraccount>
#SBATCH --partition=<yourpartition>

A few basic commands are: **sbatch** for job submission, **scancel** to delete a job, **squeue** to see jobs in the queue, and **sinfo** to obtain information about the nodes and partitions. You will find more useful SLURM aliases that provide the output of sinfo and squeue in a more informative fashion, as well as other usage information, on the CHPC SLURM documentation page:

https://www.chpc.utah.edu/documentation/software/slurm.php.

For nearly a year, Steve Corbató served as CHPC's interim director. Rather than biding time until the appointment of our new director, Tom Cheatham, Steve was actively involved in the management of CHPC while also serving as deputy and interim CIO of UIT.

While at CHPC he opened new funding opportunities for the center by bringing in grant money, and encouraged CHPC staff to assist research groups in purchasing their own nodes and storage for CHPC to manage. He oversaw the move of CHPC's dispersed resources to the new Downtown Data Center,and he reorganized the staff to better meet the needs of the users.

Steve has accepted the position of chief technology officer at the Oregon Health & Science University (OHSU) in Portland. He will be leaving the University of Utah at the end of May.

CHPC thanks him for the leadership he provided during the transition. We wish him all the best in his new adventure. May it be peaceful.

Modules at CHPC

By Anita Orendt

The Lmod modules package is an environment management tool which makes modifying the user's shell environment simple and dynamic. The advantage of modules primarily comes from the capability to load and unload the environment needed for a given software package, allowing users to quickly and easily start using programs or switch development environments. CHPC has implemented modules on all of its clusters to improve the user experience and to align with other national computing centers which are now using modules. CHPC is using the Lmod version of modules as implemented by the Texas Advanced Computer Center (TACC).

All new accounts are being provisioned to use modules. In the near future, CHPC will start to phase out the old system of sourcing different shell scripts for the applications of interest when starting sessions. A timetable for this will be announced.

CHPC User Services has created modules for quite a few of the applications we support, but not yet for all. However, if you want to use a module that has not yet been created please contact us. In addition, there are two document pages – the first is on how to get started and the second covers topics such as creating modules for personal installations: https://www.chpc.utah.edu/documentation/software/modules.php https://www.chpc.utah.edu/documentation/software/modules-advanced.php

Setting up Modules

To enable modules, we recommend that users have the following five files in their home directory: the module version of the .bashrc and .tchsrc, a .custom.sh and .custom.csh file, and an .aliases file. The first four of these are required (even if you use bash, you may run programs that require csh, or vica versa) and CHPC provides these files. The .bashrc and .tcshrc files are files that should not need to be changed. Any customization needed for interactive sessions should be put in the .custom files; as we get new or retire old clusters we will need to add or remove sections from this file. The provided .custom files only load one module, namely the module for the latest version of the intel compiler. This is the place users can load additional modules that they always want as part of their environments on a per cluster basis. Finally, users can have an optional .aliases file in which they can add any command aliases that they wish to have. If this file exists it will be sourced upon login. Users should not make any additions to paths in the .aliases file as this will interfere with the proper function of modules.

How to set up your module environment

First step: make a backup copy of your current .tcshrc and .bashrc before copying over the new module version. This is especially crucial if you have any customizations in your current dot files.

Second step: copy over the new module based CHPC bashrc and tcshrc to their home directory by doing the following.

cp /uufs/chpc.utah.edu/sys/modulefiles/templates/bashrc ~/.bashrc

cp /uufs/chpc.utah.edu/sys/modulefiles/templates/tcshrc ~/.tcshrc

Third step: copy the custom.sh and custom.csh scripts into your home directory.

cp /uufs/chpc.utah.edu/sys/modulefiles/templates/custom.sh ~/.custom.sh

cp /uufs/chpc.utah.edu/sys/modulefiles/templates/custom.csh ~/.custom.csh

Fourth step (optional): Finally, if you have any command aliases, create the .aliases file to contain these. This file, if it exists, will automatically be used.

If you have any questions or need assistance, please send in a request. CHPC staff is available to work with individual users and groups in order to help with this transition.

Modules Primer

Once you have done the above four steps you are now configured to use modules – so the next question is how do you manage your environment. There is only one command: module, with option.

Useful commands to get you started:

- **module** – lists options
- **module avail** – modules available to be loaded; note that this changes based on which modules are already listed
- **module list** – lists modules currently loaded
- **module spider** – lists all modules
- **module spider modulename** – lists information on versions available; if version included, lists dependency
- **module load modulename** – load single application module
- **module unload modulename** – unload specified module
- **module purge** – unload all modules

News

What's New at CHPC?

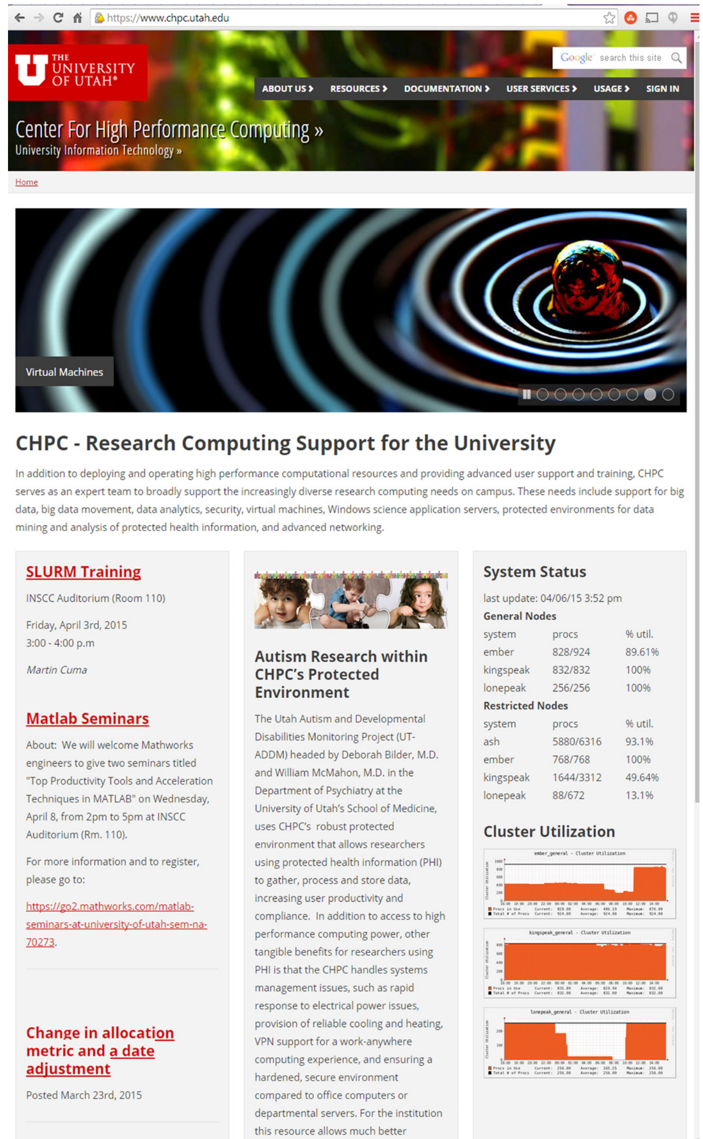
1) The kingpeak cluster is still growing. There are now 233 compute nodes. CHPC has recently added new Haswell based nodes to the owner node pool on kingspeak. These nodes have 24 physical cores and 128GB memory. While these are nodes owned by individual research groups, if idle they are available (with preemption) to the general user via the owner-guest account. If you need any information on the cost of nodes and storage, please contact us via issues@chpc.utah.edu.

2) You may have seen the recent email requesting information about CHPC usage in 2014. Our need for this information – both stats on publications, presentations, dissertations that make use of CHPC resources, and on funding that requires access to CHPC resources to complete – is for reporting required by NSF funding of the ACIREF project. As part of the allocation request process we already collect this information for groups who request a general allocation; however, we have not been gathering this information for groups with owner nodes and those groups who use other CHPC resources, such as kuchina, swasey, the VM farms, or the protected environment. In addition, short narratives on the impact on your research program are greatly appreciated! Please send any contribution directly to CHPC's director, at tec3@utah.edu.

3) CHPC has made a number of changes in the operation of the clusters over the last two months. Two of these changes are the adoption of modules for managing user environments and the use of SLURM for batch scheduling. Both of these are highlighted in this newsletter.

4) Hadoop – Do you have any applications or problems for which Hadoop might be useful? CHPC is setting up a test Hadoop cluster. If you are interested in exploring this new resource, contact us via issues@chpc.utah.edu .

5) This semester CHPC has three student interns: **Alan Navarro** is helping with the active network monitoring using PerfSonar, **Christyn Phillipp** is working on a database of software installed on the clusters, and **Paul Fischer** is helping CHPC staff evaluate the open source package SpaceWalk as a replacement for the commercial RedHat Network Satellite in order to manage and update the many linux machines CHPC administers.



6) We have a new website -- www.chpc.utah.edu. To comply with the University of Utah's recommended format, Walter Scott headed a team of CHPC staff that restructured, reorganized, took photos, and fussed over content until the new site went online in March. All the utilities of the old site are included in the new site, with some significant improvements. As you can see in the screen shot above, the usage information is highlighted on the right. In the center column we have rotating highlights of the research being done with CHPC computing resources. (If you would like your research highlighted, please email Janet Ellingson.) On the left you will see the latest news about training, presentations, and downtimes. Because we have created this new site for you we are very interested in your feedback. If you have difficulty finding the information you need, please let us know via issues@chpc.utah.edu so we can improve this site.

Selection of Recent Research Using CHPC Resources

Brown, S., Thorne, M.S., Miyagi, L., Rost, S. (2015). A composition origin to ultralow-velocity zones. Geophys. Res. Lett, 42. doi: 10.1002/2014GL062097

Fatehi, S., Steele, R.P. (2015). Multiple-Timestep ab initio molecular dynamics based on two-electron integral screening J Chem Theor Comput, 11, 884 - 898.

Galindo-Murillo, R., Roe, DR, Cheatham III, T.E. (2015). Convergence and reproducibility in molecular dynamics simulations of the DNA duplex d(GCACGAACGAACGAACGC). Biochem. Biophys. Acta., 1850(5), 1041-1058. doi: 10.1016/j.bbagen.2014.09.007

Hamian, S., Yamada, T., Faghri, M., Park, K. (2015). Finite Element Analysis of Transient Ballistic-Diffusive Phonon Heat Transport in Two-Dimensional Domains. International Journal of Heat and Mass Transfer, 80, 781 - 788.

Lund, A., Pagola, G., Orendt, A., Ferraro, M., Facelli, J. (2015). Crystal structure prediction from first principles: The crystal structures of glycine. Chemical Physics Letters, 626, 20 - 24.

Milo, A., Neel, A.J., Toste, F.D., Sigman, M.S. (2015). A data-intensive approach to mechanistic elucidation applied to chiral anion catalysis. Science, 347(6223), 737 - 743.

Shen, P. S., Park, J., Qin, Y., Li, X., Parsawar, K., Larson, M.H., Cox, J., Cheng, Y., Lambowitx, A.M., Weissman, J.S., Brandman, O., Frost, A. (2015). Rqc2p and 60S ribosomal subunits mediate mRNA-independent elongation of nascent chains. Science, 347(6217), 75 - 78.

Simpson, J. J. (2015). A summary of the major global 3-D FDTD modeling capabilities to-date. Paper presented at the International Conference on Computational Electromagnetics, Hong Kong, China.

SC15 Austin, TX | hpc transforms.

November 15 - 20, 2015

In the News...

CHPC's director, Tom Cheatham, was highlighted in the current edition *of HPCwire:*

"April 8 — NCSA's Blue Waters supercomputer will be used by three research teams to gain new understanding of the deadly Ebola virus, thanks to allocations provided through the National Science Foundation's Rapid Response Research program...

"Thomas Cheatham, University of Utah, aims to reduce the time needed to find promising drugs to fight Ebola. A promising type of antiviral drug is one that inhibits the entry of the Ebola virus into the cell. In order for this drug to work well, it must be made up of Ebola inhibitor molecules that bind strongly to the Ebola virus. Cheatham's team will analyze potential inhibitors and improve their binding affinity by simulating and optimizing their behavior on Blue Waters. Cheatham's research team has previously used Blue Waters to improve molecular dynamics methods/codes and to study biomolecules."

http://www.hpcwire.com/off-the-wire/blue-waters-to-help-researchers-tackle-ebola/

What is CHPC?

The Center for High Performance Computing (CHPC) serves as an expert team to broadly support the increasingly diverse research computing needs on campus. These needs include support for big data, big data movement, data analytics, security, virtual machines, Windows science application servers and advanced networking.

CHPC also operates a protected environment (PE) for researchers who work with data that is sensitive in nature. These resources have been reviewed and vetted by the Information Security Office and the Compliance Office as being an appropriate place to work with Protected Health Information (PHI).

These computing resources are available to all faculty at the University of Utah, their students and research staff.

CHPC Staff Directory

Administrative Staff	Title	Phone*	Email	Location
Tom Cheatham	Director	585-6318	tec3@utah.edu	414 INSCC
Julia D. Harrison	Associate Director	585-1869	julia.harrison@utah.edu	430 INSCC
Guy Adams	Assistant Director, Systems & Network	554-0125	guy.adams@utah.edu	424 INSCC
Anita Orendt	Assistant Director, Research Consulting & Faculty Engagement	231-2762	anita.orendt@utah.edu	422 INSCC
Janet Ellingson	Administrative Manager & Newsletter Editor	585-3791	janet.ellingson@utah.edu	405 INSCC

Scientific Staff	Expertise	Phone*	Email	Location
Wim Cardoen	Scientific Applications	971-4184	wim.cardoen@utah.edu	420 INSCC
Martin Cuma	Scientific Applications	652-3836	martin.cuma@utah.edu	418 INSCC
Sean Igo	Natural Language Processing	N/A	sean.igo@utah.edu	405-31 INCSS
Albert Lund	Graduate Assistant	N/A	albert.lund@utah.edu	405-14 INSCC

Project Leads		Phone*	Email	Location
Wayne Bradford	Security	243-8655	wayne.bradford@utah.edu	426 INSCC
Joe Breen	Advanced Network Initiatives	550-9172	joe.breen@utah.edu	416 INSCC
Steve Harper	Virtualization	541-3514	s.harper@utah.edu	405-30 INSCC
Brian Haymore	HPC & Storage	558-1150	brian.haymore@utah.edu	428 INSCC
Sam Liston	Special Events/SC	232-6932	sam.liston@utah.edu	405-41 INSCC

Technical Support Staff		Phone*	Email	Location
Amanda Allen	Help Desk Assistant	N/A	N/A	405-3 INSCC
Irvin Allen	Systems Admin	231-3194	irvin.allen@utah.edu	405-40 INSCC
Nathaniel Ellingson	Tech Assistant	N/A	N/A	405-28 INSCC
Jake Evans	Network Engineer	718-1526	jake.evans@utah.edu	405-32 INSCC
Paul Fischer	Intern - HPC			
David Heidorn	Systems Admin	303-987-3072	david.heidorn@utah.edu	405-18 INSCC
Alan Navarro	Tech Assistant			
Christyn Phillippi	Intern - Web			
David Richardson	Systems Admin	550-9518	david.richardson@utah.edu	405-29 INSCC
Walter Scott	User Services	309-0763	walter.scott@utah.edu	405-39 INSCC
Alan Wisniewski	Network Support	580-5835	alan.wisniewski@utah.edu	405-38 INSCC

*All phone numbers are preceded by area code 801 unless otherwise noted.

THE UNIVERSITY OF UTAH

Center for High Performance Computing
155 South 1452 East, RM #405
SALT LAKE CITY, UT 84112-0190

Welcome to CHPC News!

If you would like to be added to our mailing list, please fill out this form and return it to:

Janet Ellingson
THE UNIVERSITY OF UTAH
Center For High Performance Computing
155 S 1452 E ROOM 405
SALT LAKE CITY, UT 84112-0190
FAX: (801)585-5366

Name:
Phone:

Department or Affiliation:
Email:

Address:
(UofU campus or U.S. Mail)

Thank you for using our Systems!

Please help us to continue to provide you with access to cutting edge equipment.

ACKNOWLEDGEMENTS

If you use CHPC computer time or staff resources, we request that you acknowledge this in technical reports, publications, and dissertations. Here is an example of what we ask you to include in your acknowledgements:

"A grant of computer time from the Center for High Performance Computing is gratefully acknowledged."

Please submit copies or citations of dissertations, reports, pre-prints, and reprints in which the CHPC is acknowledged to: Center for High Performance Computing, 155 South 1452 East, Rm #405, University of Utah, Salt Lake City, Utah 84112-0190