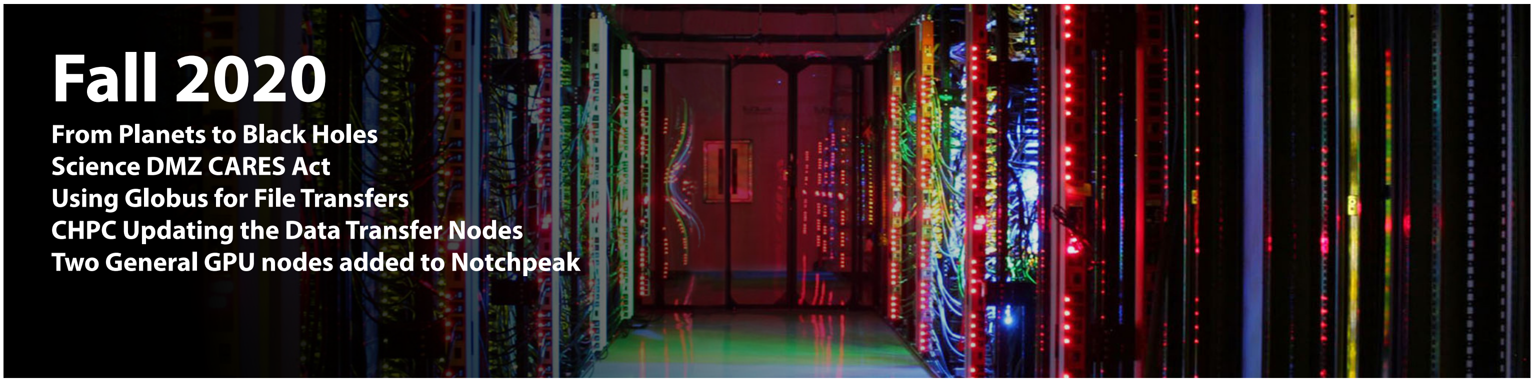
Center for
HIGH PERFORMANCE COMPUTING
THE UNIVERSITY OF UTAH

Newsletter



From Planets to Black Holes: Tracing the History of the Universe

Research Highlight · Gail Zasowski and Joel Brownstein · Department of Physics and Astronomy

Humans have carefully observed and tracked objects in the night sky since before recorded history (Figure 1). In the absence of light pollution, there are about 10,000 stars visible to the naked human eye -- only about 5,000 at one time, of course, the Earth being famously opaque to optical light. When the telescope was developed in the early 1600s, we learned that there were far more things in the sky than our simple eyes can see; as instruments were developed that could detect light in other parts of the electromagnetic spectrum, we realized that there were objects and energy patterns in the sky that we could never see with our eyes, no matter how sensitive. And the Universe is not static -- binary stars whip around each other, black holes flare brightly as they gobble up gas, and galaxies merge in billion-year-long collisions. How do we observe and analyze all these patterns to understand the nature of the Universe?

Enter astronomical surveys, and in particular, the newly-launched fifth generation of the Sloan Digital Sky Survey (SDSS). The first four generations of SDSS were each, in their own way, ground-breaking in terms of scientific goals and methods, and the new SDSS-V continues that tradition. As the data storage and processing hub of the survey, the University of Utah plays a very important role in the discoveries that will arise from SDSS-V's massive data trove throughout the 2020s.

SDSS-V is a spectroscopic survey, which means it collects high-resolution spectra of objects rather than taking images (Figure 2). With spectra, astronomers can infer

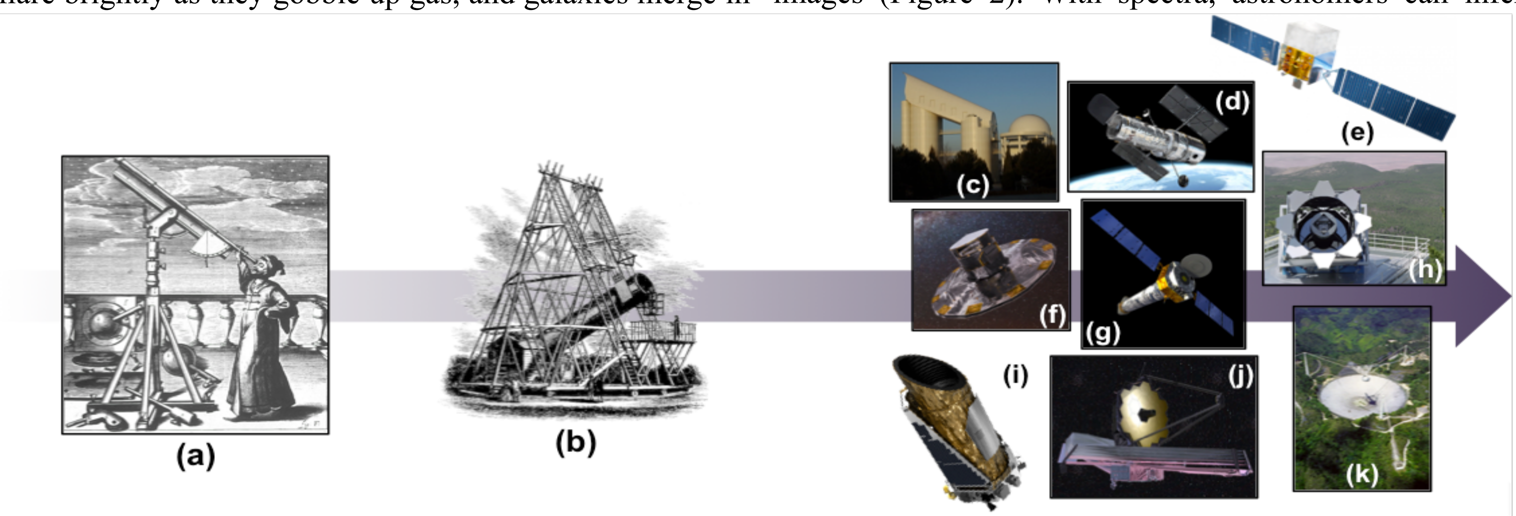


Figure 1: Telescopes, then and now. (a) Early 1600s: Popularized by Galileo Galilei, early telescopes revealed the moons of Jupiter and the fact that many seemingly "fuzzy" parts of the sky were in fact made of distinct stars. (b) Mid 1800s: Large lens-based telescopes (shown is the "Leviathan of Parsonstown", completed in 1845 in Ireland) were used to resolve distant galaxies, though these were thought to be part of the Milky Way at the time. (c)-(k) shows a range of modern day telescopes that capture light across the electromagnetic spectrum: (c) LAMOST Telescope (Hebei Province, China, visible light), (d) Hubble Space Telescope (in Earth orbit, UV+visible+IR light), (e) Fermi Telescope (in Earth orbit, gamma rays), (f) Gaia (in Earth orbit, visible light), (g) Chandra X-Ray Observatory (in Earth orbit, X-rays), (h) Sloan Telescope (NM, USA, optical+IR light), (i) Kepler Space Telescope (in Earth orbit, visible light), (j) James Webb Space Telescope (to be launched to L2 in 2021, IR light), (k) Arecibo Telescope (PR, USA, radio waves).

(among other properties) the line-of-sight velocity, temperature, and composition of a star; the redshift, average age and stellar motions of a chunk of galaxy; the density and composition of intergalactic gas clouds; and the density and temperature of expanding supernova remnants. In many cases, the changes in these properties over time is even more exciting -- e.g., the change in density and temperature of a supernova remnant (which happens over scales of hours and days) tells us how the explosion material and energy is being deposited back into the interstellar gas and dust. A period shift in a star's line-of-sight velocity can tell us about unseen companions to the star, from planets to stellar-mass black holes. SDSS-V is providing spectroscopic information for millions of objects in the night sky, and in many cases time-resolved spectroscopic information, which is a necessary complement to the massive imaging datasets from space satellites like TESS (NASA), Gaia (ESA), and eROSITA (ESA), and from ground-based projects like the LSST. The span of its dataset, both in number of objects and in time, is what sets it apart

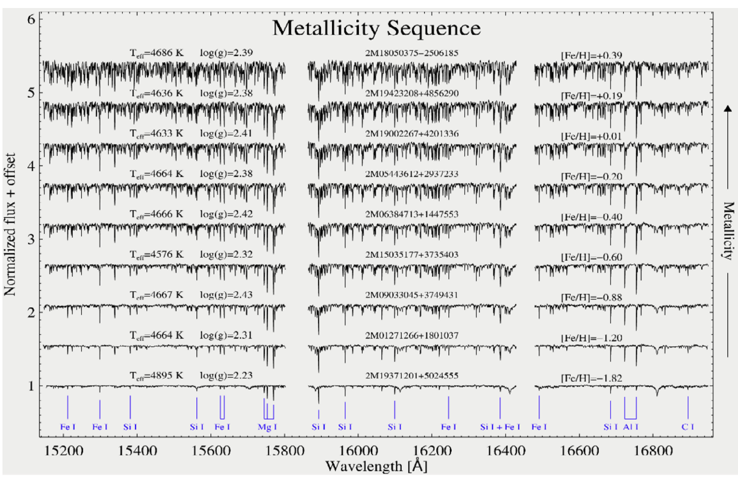


Figure 2: Examples of stellar spectra, that is, the relative energy emitted by the star as a function of wavelength. The spectra here are sorted in order of heavy element fraction. The stars poorest in heavy elements are at the bottom (with about 2% the fraction of heavy elements found in the Sun), and the ones richest in heavy elements at the top (with 250% the Solar value). The elements corresponding to different features are indicated at the bottom.

from other astronomical surveys happening this decade.

The primary science goals of SDSS-V can be sorted into a few broad themes that span the early Universe to the Sun's (relatively speaking) immediate backyard.

One key focus is on supermassive black holes (SMBHs): ultradense objects that sit at the heart of most galaxies in the Universe. The energy injected by SMBHs into their host galaxies -- via the light coming from their superhot gas disks, or from their enormous plasma jets -- plays an enormous role in shaping how those galaxies evolve .Understanding how this energy is produced, and when and how the SMBHs grow accretion disks and launch jets, requires 1) finding a lot of systems at different stages of the Universe's timeline, and 2) tracking how they start and stop growing (and spitting out energy) on both very short and very long timescales. SDSS-V is tackling both of these goals by collecting spectra of hundreds of thousands of bright X-ray sources, many of which will turn out to be SMBHs (many others will be distant clusters of hundreds or thousands of galaxies). We'll then be able to trace how average SMBH mass and behavior change over billions of years!

Another key part of SDSS-V looks closer to home, collecting data for more than 5 million stars in our own Milky Way Galaxy and in a handful of our smaller galactic neighbors. These spectra are analyzed to get the stars' temperatures, compositions, surface gravities, and line-of-sight velocities, which can be used to infer distances, ages, and orbit within the galaxy. By mapping these properties across the Milky Way -- for example, ages and chemical compositions -- we will learn an enormous amount about when, where, and how the Milky Way formed its stars. All of the atoms in the Universe of elements heavier than hydrogen and helium were produced by stars, so retracing this stellar history helps us understand not only galaxy evolution but also the chemical enrichment of the Universe and the formation of the periodic table as we know it, so crucial to life here on Earth.

By repeatedly observing the same star, we can also measure changes in its line-of-sight velocity due to its orbit around a companion object. Many stars in the sky have at least one stellar companion (the Sun is actually in the minority, in not having one), or a companion that was once a star but is now a white dwarf or a black hole. Many, perhaps most, stars also have planets! So these orbital wobbles tell us about the architecture of multi-object stellar systems -- how many companions, their mass, their distance from the star. By measuring these quantities for hundreds of thousands of stars of different ages, masses, and compositions, SDSS-V will provide huge insights into how stars form and what affects those formation processes, which will improve our understanding of star formation (and the likelihood of finding life-bearing exoplanets!) throughout the Universe's history.

Tackling these and SDSS-V's other primary science questions requires a large team -- to build the instruments, operate the telescope, collect and store the data, write the software to process the data, analyze the data, etc. The SDSS-V Collaboration currently spans nearly 600 individuals at more than 60 institutions across the planet. The University of Utah and the Center for High Performance Computing (CHPC) play a key role in this collaboration. Joel Brownstein, the Principal Data Scientist for all of SDSS-V, is a research professor in the Department of Physics & Astronomy. Gail Zasowski, also a professor in

Physics & Astronomy, serves as the survey's Spokesperson. Numerous Utah astronomy faculty, staff, postdocs, and students have actively participated in research and publication using data from earlier generations of SDSS.

On top of this, the U's CHPC has served as the data processing and distribution center for SDSS since 2014. It takes tens of millions of core-hours to process all of the millions of spectra that SDSS uses to do research and releases to the world. More than 150 astronomers from SDSS's member institutes around the world have CHPC accounts. These people develop all of the data reduction software, analysis pipelines, data products, and value-added catalogs that are ultimately released to the public. The latest of these is Data Release 16, which was released in 2019, as documented on the *www.sdss.org* website. The new SDSS-V survey is described at *www.sdss5.org*, and anticipates releasing its first data in 2022.

Earlier generations of SDSS directed light from the sky to the spectrographs using fiber optic cables hand-plugged into large aluminum plates connected to the telescope. Because SDSS-V is replacing this system with robot-positioned fibers, both the volume of data and the number of times we can visit the same object will be increasing dramatically. This leads to a variety of challenges in how we organize, search, and visualize the data. We are also very fortunate at the U to have the new Utah Center For Data Science, *datascience.utah.edu*, which brings cutting edge technologies, such as machine learning and artificial intelligence, to the forefront of astronomical data analysis.

Science DMZ CARES act: The next network wave supporting science

Joseph Breen, CHPC Senior IT Architect and Anita Orendt, CHPC Assistant Director for Research Consulting

The next network wave is upon us. Performance speeds of 400Gigabit/sec for backbones and 100-200Gb/s for end nodes are here. Petabyte datasets are happening now with Exabyte datasets expected in the very near future. Research is growing rapidly, especially to support COVID-19 investigations. These type of investigations require more resiliency and security. In response to these changes, the University of Utah is positioning itself to handle this new wave and to support the burgeoning data requirements of our research community.

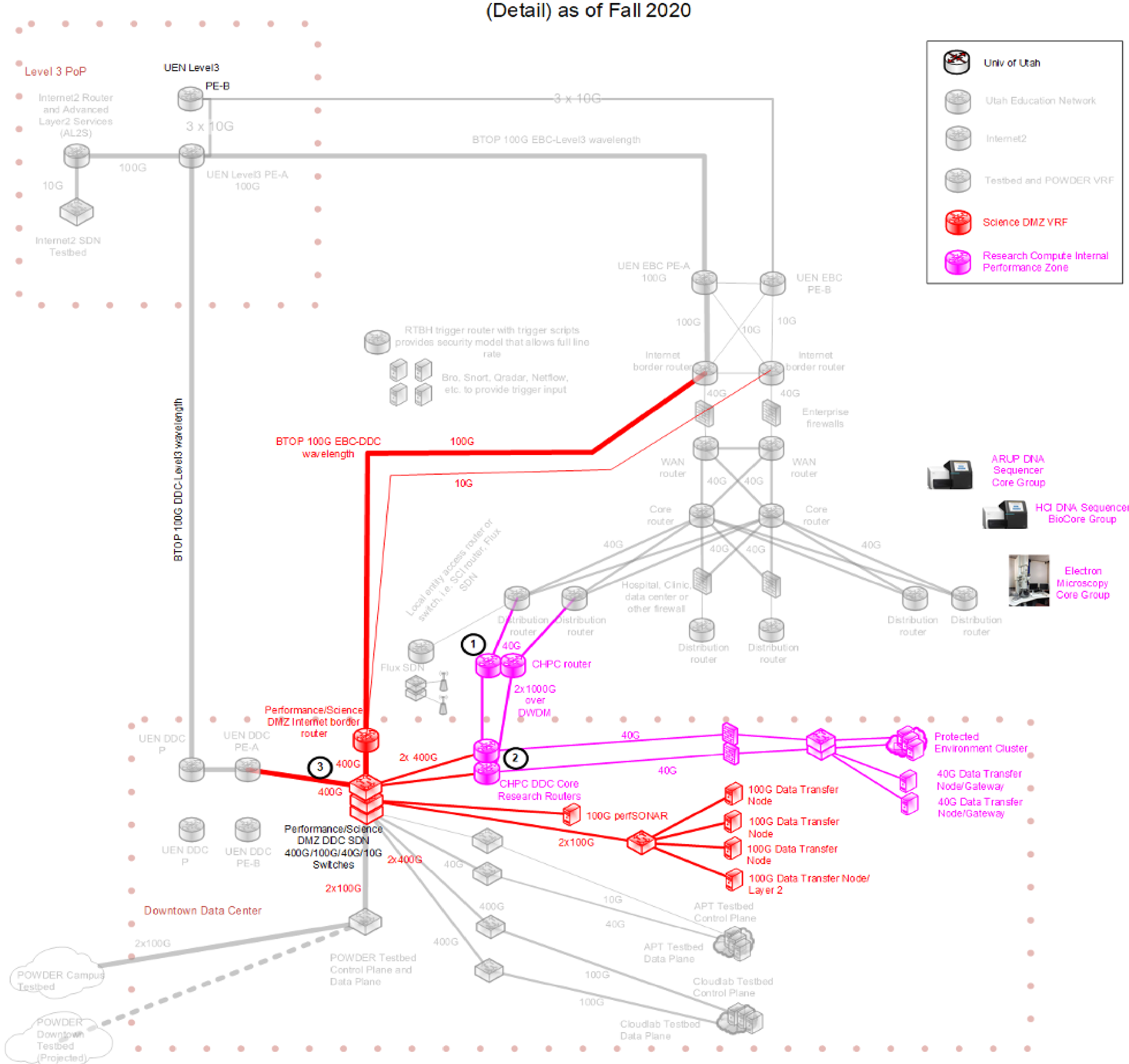
In July 2020, the state of Utah received a CARES act[1,2] stimulus package, some of which it allocated to schools throughout the state and to the Utah Education and Telehealth Network (UETN)[3] to stimulate learning and support investigations into COVID-19. The University of Utah's portion included specific funding to advance the abilities of the research infrastructure. A portion of the University's funds were provided to the Center for High Performance Computing Center (CHPC) with the focus of supporting COVID-19 research, specifically in the context of the University's Science DMZ[4] and Research Computing infrastructure. Through strategic design decisions, the opportunistic use of rapidly changing technologies, and with collaborative work with partner vendors partners such as Cisco, Dell, Atipa, Supermicro, Palo Alto, Mellanox, Arista, FS.com, and Globus, CHPC was able to stretch this stimulus funding to replace not only aging Science DMZ infrastructure, but to look at the Science DMZ and Research Computing environment holistically.

The University of Utah first installed its Science DMZ in 2013 as part of a National Science Foundation (NSF) grant #1341034, CC-NIE *Integration: Science Slices Converting Network Research Innovation into Enhanced Capability for Computational Science and Engineering at the University of Utah*.[5] The equipment at the time supported up to 100Gb/s speeds and dynamic software defined network capabilities for enabling specific areas of research. The Science DMZ equipment continues to support diverse research from Genomics, Astronomy (i.e., Sloan Digital Sky Survey), BGP (Border Gateway Protocol) peering testbed, national systems/network testbeds (Cloudlab, POWDER, Apt) and more. Through the years, the Science DMZ has supported over $120 million in different science grants. However, the equipment has begun to age and has gone out of warranty.

Traditionally, high-end features and performance only came in large chassis units, which were used in the original Science DMZ. However, new technologies and chipsets are now available that support 400G/s speeds in small 1 or 2 rack unit (1U or 2U, with 1 rack unit = 1.75in) footprints, as well as also providing additional features for monitoring traffic such as active telemetry. The research needs have also changed as support of real-time health research and other protected research (i.e., over 5000 research projects in the University of Utah CCTS Biomedical Informatics Core REDCap[6]) has grown in significance. The growth in this type of research has changed the support models and has created the need to build more resilient infrastructure in key strategic points.

The Science DMZ CARES act funding along with the exploration of new designs and technologies is allowing the growth and expansion of the original Science DMZ concepts. High performance, flexible software programmability, and the ability to support isolated network segments were key components of the original Utah Science DMZ and Research Computing network implementation. For the new iteration, strategic redundancy, fast storage, program-

University of Utah/UEN
Performance/Science DMZ/Testbed
Science DMZ CARES Physical Implementation
(Detail) as of Fall 2020



mable network interface offloading, and active telemetry have also risen in importance.

The figure above shows a schematic of the proposed new Science DMZ and Research Network, with some of the new network features highlighted in color. Highlights of the new network components include:

•Redundancy in key areas of the research network interfacing to campus and the core routers of the research network, designated in the schematic by ①and ②, respectively.

•a change of the 100Gb/s uplink to a 400Gb/s uplink, indicated by ③ in schematic for the Science DMZ network to external sites, in parallel to the UETN CARES act which is upgrading the regional portion of the network to 400Gb/s.

•the addition of ports of higher speed so the team could connect new research projects, such as Cloudlab3, via 400Gb/s.

Since the smaller units cost models are only about a third of the cost of traditional chassis, budget also was available to adopt a very holistic approach beyond just the core network components by allowing for the consideration of other aspects of the Science DMZ and Research Computing network, including:

- *The replacement of the Data Transfer Nodes (DTNs).*[7] (See the separate article in this newsletter for more about the new DTNs.) The new DTNs for the general environment will connect at 100 Gb/s and will allow disk-to-disk transfers at wire rate.
- *The addition of a telemetry node with a tiered data store.* This node, which connects at 100Gb/s, has internal NVMe disk capacity with additional directly connected spinning storage allowing for a tiered storage solution. The telemetry node will collect ongoing metrics of the network and the systems and allow for exploration on how to store the data in an open source, tiered fashion.
- *The addition of new perfSONAR*[8] *active measurement modes.* These nodes, also connect at 100Gb/s, have NVMe storage large enough to conduct disk-to-disk tests and memory-to-memory tests, to allow for validation of the network infrastructure and its ability to transfer large datasets in/out of the university and within the university.
- *The addition of advanced, smart network interface cards (NICs)*. The new DTNs, perfSONAR nodes and the telemetry node cards include advanced NICs which provide additional capabilities in programmability, network manipulation, and monitoring. These capabilities provide ability for exploratory work with active data and flow management.
- *The replacement of the core Infiniband Gateways.* The new gateway will replace the existing Infiniband gateways which are nearly 8 years old, and are completely new devices and designs which will provide new functionality beyond just performance. They will interface the Infiniband network with the Ethernet one in order to transport control plane services and storage networks.
- *The extension of the CHPC globus license for five years.* This license aids the research groups in moving their data (both protected and unprotected) to locations around the world. This software allows the multiple DTNs to act as a parallel transfer cluster to move data from all of the machines simultaneously, thus more fully leveraging the high performance network.
- *The addition of firewall licenses.* Additional Palo Alto firewall licenses help secure the protected enclave research that the University of Utah houses. The redundant licenses allow for more granularity and real time monitoring of the environment.
- *The addition of 102.4 TB of fast NVMe storage for the Cloudlab testbed.* This storage is to support real-time workflow research for handling large data sets and moving them around to different resources. This type of workflow research has implication on domain science research such as medical.

The team hopes that this holistic approach will further COVID-19 research and other research for many years in ways that accurately reflect the workflows of the research and the science.

[1]Coronavirus Aid, Relief, and Economic Security (CARES) act: *https://www.congress.gov/bill/116th-congress/senate-bill/3548*

[2]CARES act provides assistance for state, local and tribal governments: *https://home.treasury.gov/policy-issues/cares/state-and-local-governments*

[3]UETN CARES Act: *https://www.uen.org/caresact/*

[4]Science DMZ motivation: *https://fasterdata.es.net/science-dmz/motivation/*

[5]*https://www.nsf.gov/awardsearch/showAward?AWD_ID=1341034&HistoricalAwards=false*

[6]REDCap: *https://redcap01.brisc.utah.edu/ccts/redcap/*

[7]CHPC Data Transfer Nodes: *https://www.chpc.utah.edu/documentation/data_services.php#Data_Transfer_Nodes*

[8]PerfSONAR: *https://www.perfsonar.net*

Using Globus for File Transfers

Martin Cuma, CHPC Scientific Consultant

Doing research these days often requires moving data, and sometimes large amounts of data, not just within an institution or between different institutions, but also to and from researchers' homes. We have a help page devoted to Data Transfer Services, *https://www.chpc.utah.edu/documentation/data_services.php*, where we list about a dozen ways how to transfer data. Out of these approaches, Globus is standing out for its ease of use, performance and flexibility. In this article we introduce Globus and detail its use for data transfers to and from CHPC, and how can one share their data located at CHPC with external collaborators.

Globus service, endpoints, collections and app

Globus, *www.globus.org*, is a subscription based non-profit service for research data management. It provides tools for moving, sharing and discovering data. The University of Utah is one of the subscribers of the Globus service, which allows CHPC to connect to the Globus infrastructure.

File systems, such as CHPC file servers, or drives on your personal laptop, connect to the Globus service by the use of endpoints. CHPC maintains Globus endpoints both for the general and protected environment, by the use of our Data Transfer Nodes (DTNs, *https://www.chpc.utah.edu/documentation/data_services.php#Data_Transfer_Nodes*). All but one of the general environment DTNs are attached to the Science DMZ network which bypasses the campus firewalls; these DTNs are secured in a less traditional way that allows for intrusion detection and mitigation to happen without impacting network performance. For all off-campus transfers using a DTN on the Science DMZ is recommended. CHPC currently also has one DTN, dtn03.chpc.utah.edu, in the general environment not on the Science DMZ; this DTN is on campus network space and is designed for intra-campus transfers. In the future CHPC plans to add additional intra-campus DTNs to facilitate better, fast transfers between resources on campus. As there is not a DMZ for our protected environment (PE) the DTNs in the PE follow a path off campus through the campus firewalls. This should be kept in mind as you make transfers in and out of the PE that performance will be impacted somewhat by having to traverse the campus firewalls. As a side note, we are in the process of replacing the existing DTNs with new servers, purchased with the CARES act funding mentioned in the previous article. A description of the new DTNs follows this article. In the Globus web app, the endpoints are denoted with the icon.

Access to endpoints is provided by the collections. *Mapped Collections* are created by the endpoint administrators, and CHPC provides mapped collections to CHPC file servers through the CHPC defined collections listed below. *Guest Collections* are created by users and can be used for sharing data. The collections in the Globus web app are displayed with the icon.

Globus' central usage interface is the web app, *app.globus.org*. Having accounts at multiple institutions may be managed by clicking the Account on the toolbar found on the right-hand side. Linked multiple identities allow one to have access to endpoints from different institutions and initiate data transfers between them. Multi-institutional account may be best organized by having a Globus ID, *https://www.globusid.org/create*, which the user can use as a master account covering all the institutions with which the user may be affiliated.

After logging into the app with the University credentials, one can search the endpoints to find those that belong to CHPC. We recommend to use "CHPC Utah" as a search phrase, as there is also a CHPC in South Africa. At the time of this writing, for the general environment access, use the Mapped Collections University of Utah – CHPC DTN01 or University of Utah – CHPC DTN02; for the protected environment, use University of Utah - CHPC PE-DTN01 or University of Utah - CHPC PE-DTN02 Some minor endpoint reconfiguring will take place in the near future with new hardware purchases. Our Globus documentation at *https://www.chpc.utah.edu/documentation/software/globus.php* provides up to date information.

Data transfers are initiated in the Globus app and run in the background, with Globus tuning the performance parameters, maintaining security, monitoring progress and validating correctness. Once the transfer is finished the user receives an e-mail notification.

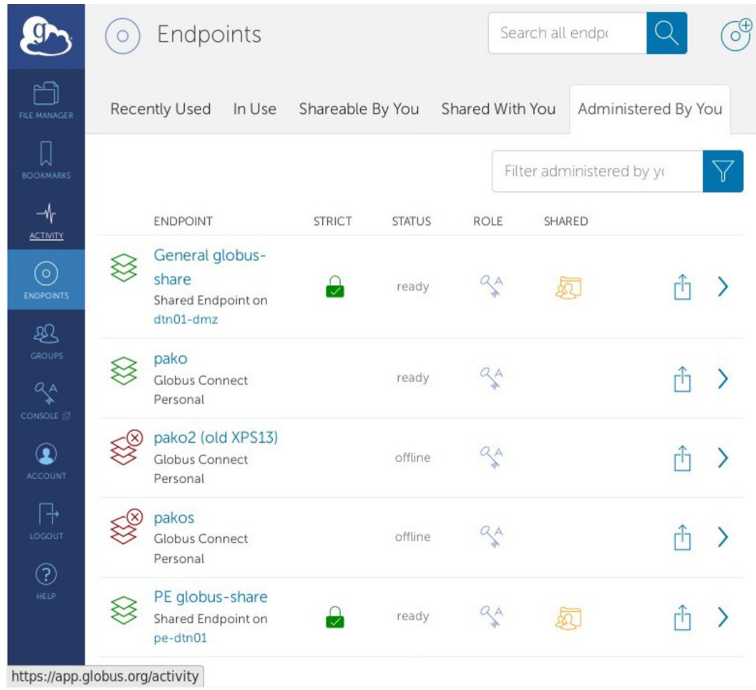


*Figure 1. Globus web app listing user administered endpoints. Both shared endpoints for sharing data on CHPC systems and Globus Connect Personal (GCP) endpoint for connecting personal computer are shown. The two red personal endpoints are author's other personal computers that are currently offline. Also note the **STRICT** column, which denotes support for protected information. GCP also supports this, but, user account at the personal machine must be using the University of Utah's authentication to be able to connect to it from the Globus app authenticated with the University credentials. Clicking on the **Shared With You** tab will display data collections other users shared*

Copying data from local machine to CHPC

For moving data between personal computers and CHPC, we have been recommending file transfer programs that implement the secure copy (scp) protocol, such as WinSCP for Windows. A Globus equivalent is the Globus Connect Personal (GCP). It is a Windows, Mac or Linux program that one installs on a personal computer, which creates a personal endpoint. One can then use the web app to transfer files between this personal endpoint and any other endpoints the user has access to, including the CHPC endpoints. Detailed documentation on how to install GCP is at *https://www.globus.org/globus-connect-personal*. Once the local GCP application is installed, one can see the personal endpoint in the web app by going to Endpoints – *Administered By You* (Figure 1)

To transfer data from local machine to CHPC, choose the *File Manager* in the left side menu, and choose two panes in the upper right. Click on the search on the right pane to display your connections and select the GCP to display data on your personal computer. In the left pane,

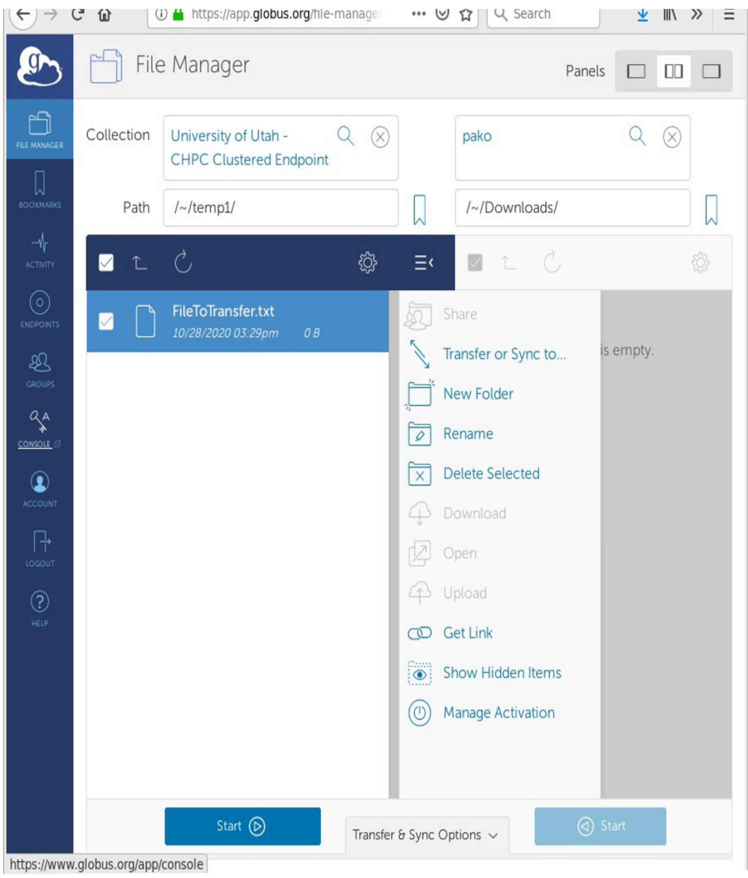


*Figure 2. Globus web app **File Manager** with two panels, CHPC home directory on the left, personal computer on the right. Clicking on the Start button initiates the transfer of FileToTransfer.txt from CHPC to the personal computer.*

search for "CHPC Utah" and choose the appropriate endpoint, e.g. *University of Utah – CHPC DTN01*, which will bring your CHPC home directory contents. Navigate to the appropriate directories in both the local and CHPC endpoints, click the files or directories to transfer and click the *Start* button at the bottom of the page to initiate the transfer (Figure 2). The transfer will run in the background, and you will receive an e-mail once the transfer is finished. Statistics about the transfer can be obtained by clicking on the *Activity* menu on the left hand side toolbar, and choosing the appropriate transfer item.

The data transfers can be done between any endpoints one has access to, such as accounts at other institutions, or multiple personal computers running either Windows, Mac OS or Linux operation systems, or data collections shared by other Globus users. Globus web app thus provides a unified interface for essentially any data transfer one may need. This also includes access to cloud storage providers, such as Google Drive or Box – however note that this access requires additional costly subscription that CHPC currently does not have.

The main advantages of using Globus instead of other data transfer tools are improved performance and resilience of the transfer, which becomes more important

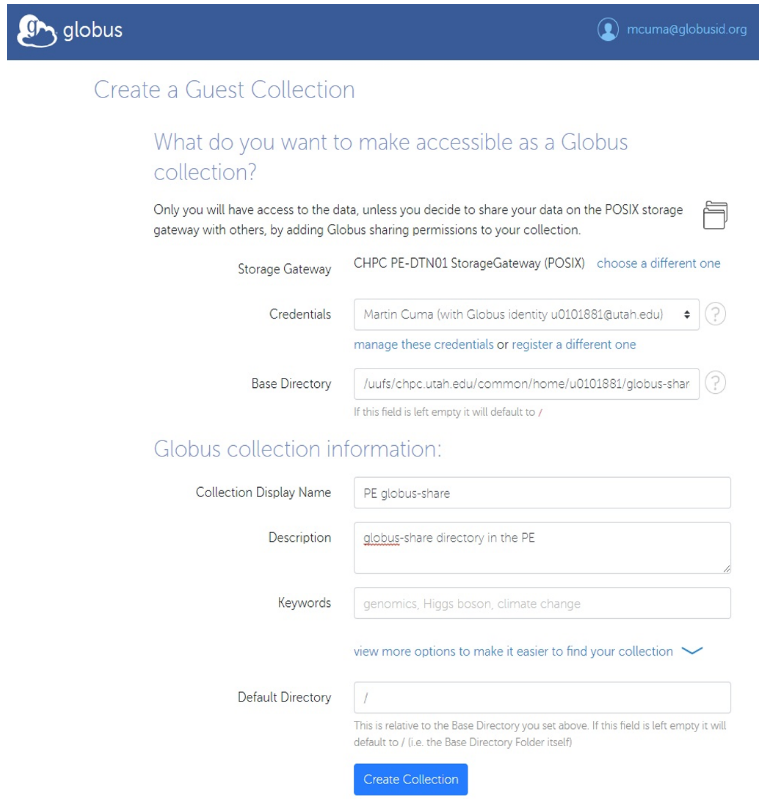


*Figure 3. Create Guest Collections Permissions window; make sure to use the absolute path for the **Base Directory**.*

with increased size of the transferred data. In a simple test we have transferred a 1GB file from a home directory on CHPC clusters to a Windows personal laptop connected to the UConnect wifi network using Globus and WinSCP. The Globus achieved about 12 MB/s speed, transferring the file in a few minutes. WinSCP averaged only about 0.5 MB/s and it took half an hour.

Sharing data with collaborators

Within the Globus endpoints, one can choose directories to share with others. These shared directories are called Guest Collections. Guest collections are built on top of CHPC Globus endpoints. In the general environment, these endpoints are dtn01-dmz or dtn04-dmz for the general environment and pe-dtn01 or pe-dtn02 for the PE.

Globus provides an e-mail invitation and a web link that the collaborators can use to access the data. The advantage of this approach is that it eliminates the need for data duplication, either in the cloud or in the CHPC Temporary Guest Transfer service (*https://www.chpc.utah.edu/documentation/data_services.php#Guest_Transfers*). One can directly share data that are located on the CHPC storage, rather than having to copy it elsewhere for sharing.

To initiate a share, one has to create a *Guest Collection*. In the *Endpoints* menu, search for *CHPC Utah* to find the endpoints, click on the endpoint name, and navigate to the *Collections* tab. Here click on the *Add a Guest Collection* button to create the collection. You will be prompted to re-authenticate with the University credentials, and if doing this for the first time, allow the Collections App to access certain information and register your account (your uNID) with the Globus Connect Server. After this initial setup, the *Create Guest Collection* screen appears. The *Base Directory* has to be the absolute path to the data you want to share, e.g., */uufs/chpc.utah.edu/common/home/u0123456/sharing*. Also enter the *Collection Display Name*, optionally the *Description*, *Keywords* and *Default Directory* (Figure 3). Then click the *Create Collection* button to create the collection. Once the collection is created, you have several immediate options available, including *Share data on this new collection with others* to add collaborators to share the data in the collection, or to get a link for sharing, that can be sent to the collaborator.

If you add collaborators to the collection, they will receive a notification e-mail with a web link, or you can send them the link for sharing. When collaborators input that link in a web browser, it will bring up the Globus web app, prompting them to authenticate – either with their own institution's credentials, or to create a new free Globus account. Once the collaborators authenticate with Globus, the Guest Collection will be accessible to them under the Endpoints menu, Shared With You tab.

Multi-user data collaboration is facilitated by the Globus Groups. The *Group* menu item lists the groups one belongs to. One can also create new groups. To share data with a group, choose the Guest Collection, and in the *Permissions* tab, *Add Permissions – Share With*, select the newly created group. Here one can also modify the data access to be read or read/write.

Conclusions

Over the recent years, Globus has evolved into a complex solution for data transfer and sharing needs. We believe that it is currently the best and easiest tool for data management. Note that the instructions listed above are for general, non-protected, data. For protected (high-assurance) data, the process is similar but the identity management and data access policies are stricter.

Aside from the data transfer and sharing options described above, we are also working on creating dedicated cluster resources for data transfers schedulable by SLURM, that could be incorporated into workflow managers such as Snakemake or Nextflow.

Globus also provides a command line interface (CLI) that can be used to automate data management in scripts, and an application programming interface (API) that allows to programmatically access Globus features either in user programs, or in creating data web portals.

CHPC Updating the Data Transfer Nodes

Sam Liston, CHPC Storage System Administrator

As part of the funding secured through the CARES Act, CHPC purchased six new DTN nodes, four for the general environment and two for the protected environment (PE). These nodes are an improvement over the existing DTNs, with more and faster cores (each node has 24 physical cores at 3.2GHz base clock speed) and more memory at 128 GB per node.

The four nodes in the general environment will be attached to the Science DMZ and internally to CHPC networks with a 100gb ethernet link as well as having direct links to the Notchpeak and Kingspeak Infiniband fabric. The two nodes in the protected environment will be connected to CHPC networks at 40gb and to the redwood infiniband fabric.

All of the new DTNs are outfitted with local NVMe storage. We are still experimenting with how to best utilize this local storage. We are currently pursuing unifying the local disks of all four of the general environment DTNs into a single, larger, very fast scratch space through the use of Ceph and a CephFS file system.

For use with Globus, these DTNs will be configured into a clustered setup so that transfers can take advantage of the available CPU, memory, and network resources of across all four general and the two PE nodes. As part of the upgrade the Globus clustered endpoint will also be installed with Globus Connect Server version 5. The upgrade to version 5 will introduce improvements to security and data sharing.

As a new feature, CHPC also plans to make these DTNs schedulable via SLURM. The idea is to provide users the ability to schedule a dedicated portion of a DTN as part of their SLURM jobs to facilitate data transfer to or from CHPC as part of their workflow. The implementation of this feature is still being developed; we will announce when it is ready for testing.

Two General GPU nodes added to Notchpeak

Anita Orendt, CHPC Assistant Director for Research Consulting

Two additional GPU servers were added to notchpeak using CARES act funding. The first is notch271 with eight RTX2080Ti GPUs, 40 compute cores (two Intel Cascade Lake 6230 20-core processors), and 192 GB of memory. The second is notch293 with four Nvidia Data Center A100 GPUs. The node also has 64 cores (two AMD EPYC 7502 32-core processors) and 512 GB of memory. Both are in the notchpeak-gpu partition. This addition brings the total number of nodes in the notchpeak-gpu partition to nine, with 36 GPUs. Please see the CHPC documentation *https://www.chpc.utah.edu/documentation/guides/gpus-accelerators.php* for additional details on using the GPU nodes.

Center for
HIGH PERFORMANCE COMPUTING
THE UNIVERSITY OF UTAH

The University of Utah
University Information Technology
Center for High Performance Computing
155 South 1452 East, Room 405
SALT LAKE CITY, UT 84112–0190

Please acknowledge the use of CHPC resources

Welcome to CHPC News!

If you would like to be added to our mailing list, please provide the following information and via the contact methods described below.

Name:
Phone:
Email:

Department
or Affiliation:

Address:
(campus
or U.S. mail)

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

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

If you make use of the CHPC Protected Environment, please also acknowledge the NIH shared instrumentation grant:

"The computational resources used were partially funded by the NIH Shared Instrumentation Grant 1S10OD021644-01A1."

Thank you for using CHPC resources!

Please submit copies or citations of dissertations, reports, pre-prints, and reprints in which CHPC is acknowledged in one of the following ways:

Electronic responses

By email: *helpdesk@chpc.utah.edu*
By fax: (801) 585–5366

Paper responses

By U.S. mail: 155 South 1452 East, Rm 405
Salt Lake City, UT 84112–0190