UC San Diego
UC San Diego News Center

March 10, 2020 | By Jan Zverina

SDSC Announces Comprehensive Data Sharing Resource

'HPC Share' powered by SeedMeLab web-based data management system

The San Diego Supercomputer Center (SDSC) at the University of California San Diego has announced the launch of 'HPC Share', a data sharing resource that will enable users of the Center's high-performance computing resources to easily transfer, share, and discuss their data within their research teams and beyond.

"HPC (High Performance Computing) users face a range of hurdles to share their data," said SDSC Visualization Group Leader Amit Chourasia, also SeedMeLab's Principal Investigator (PI). "Their collaborators may not have adequate context of the computed data, or may not be able to find, access, or fetch the data from HPC system. Collaborators also often have a pressing need to repeatedly find, access, and review data products in a fast-paced environment. These hurdles inevitably create a bottleneck at best, or in the worst case can cripple the scientific discovery process and also burden the principal HPC user to devote considerable time and effort to develop ad-hoc and poorly maintained data sharing methods rather than focusing on the research."

HPC Share solves these hurdles by letting users accelerate the research pace and information exchange via a ready-to-use infrastructure. Its key capabilities include easy data transfer, accessibility and sharing via a web browser on any device, an ability to add annotation to any file or folder, discuss and visualize tabular data. *HPC Share* is available to all users of SDSC's *Comet* supercomputer, with a potential expansion to SDSC's *Triton Shared Computing Cluster (TSCC)* and its *Expanse* supercomputer slated to enter production later this year.

"*HPC Share* bridges a major gap in our current cyberinfrastructure by offering a turnkey system that eliminates barriers among collaborating researchers so that they can quickly access and review scientific results with context," said SDSC Director Michael Norman. "Users will benefit from reduced complexity, ubiquitous accessibility, and more importantly rapid knowledge exchange."

HPC Share is powered by SDSC's open-source *SeedMeLab* software that was developed with support from National Science Foundation (NSF). Its built-in web services, coupled with an API extension, make it a versatile platform to create branded data repositories for small research groups to large communities or integrate with existing research data flow while also serving as an important stepping stone for researchers to realize FAIR data management in practice.

March 19 Webinar

Researchers interested in learning more about *HPC Share* and *SeedMeLab* are invited to participate in a free webinar on March 19 that will cover the use of the data management content system on SDSC's *Comet* supercomputer.

"We have set up a dedicated instance of *SeedMeLab* for our HPC users," said Chourasia, who will lead the webinar. "We will provide hands-on training to move, share, and present your data to and from *Comet* that will organize and store your data in one place while making it accessible to other researchers." Details of the webinar can be found here.

In addition to Chourasia, the SeedMeLab project includes SDSC Director Michael Norman as co-PI and David Nadeau as a technical architect with SDSC. Master's and undergraduate interns at UC San Diego's Computer Science and Engineering Department, as well as regional high school students, have assisted in the project with prototype extensions and comprehensive quality assurance of the software. SeedMeLab is supported by NSF grant #OAC-1443083. More information about SeedMeLab can be found here.

MEDIA CONTACT

Jan Zverina, 858-534-5111, jzverina@sdsc.edu
Amit Chourasia, 8585-822-3656 achourasia@ucsd.edu