High Performance Computing

Welcome the Smithsonian Institution High Performance Computing Wiki.

This Wiki holds information for the use of HPC resources at the Smithsonian.

Central to the SI HPC resources is the Smithsonian Institution High Performance Cluster (SI/HPC), named Hydra.

- High performance computing is administered by the Office of Research Computing.
- The OCIO data center in Herndon, VA houses the high performance computing cluster, Hydra.

- The documentation is organized as follow:
  1. A quick start guide;
  2. Reference pages (technical information about the cluster);
  3. Cluster Upgrade documentation (past upgrades);
  4. Cookbooks pages (didactic information on how to run specific tools); (NYA: not yet available)
  5. FAQs, or frequently asked questions

What's New

- **Sep 15 2020**
  - Scrubbing on /scratch has resumed: files older than 180 days are scrubbed
  - The GPFS s/w is in the process of being upgraded from v4 to v5, on a rolling basis and transparent to the users
  - IDL v 8.8[.0] is available

- **Apr 3 2020** - Hydra DOI
  - We have setup a DOI for Hydra (https://doi.org/10.25572/SIHPC)
  - You are now able and encouraged to cite Hydra whenever research has benefited from its use and add a DOI link to that citation or acknowledgment.

- **Mar 24 2020** - Hydra status while teleworking
  - Hydra remains up and running.
  - We will address problems that require on-site staff as fast as possible.
  - We will answer people's questions and requests as promptly as possible.
  - Access to Hydra via VPN:
    - users are asked to limit the strain on the institutional VPN resources when ever appropriate.
    - Hydra can be accessed without VPN:
      - use the "Hydra" link under "It Tools" (or use RDP) at telework.si.edu,
      - SAO users can use login.cfa.harvard.edu to ssh to Hydra.
      - Access to the self serve password page is now working.
    - How to use Dropbox or Firefox Send (ffsend) to copy files to/from Hydra is documented on the Wiki.
  - The scrubbing policy has been modified as follows:
    - the scrubber will run on /pool/sao and /pool/genomics as usual, but
    - the scrubbed content will not be deleted for at least 21 days, and
    - we will accept requests to preserve what was scrubbed (beyond 21 days) as long as needed.
    - To get your files restored, follow the usual instructions.
  - Users are asked to remain in contact via their SI email.

- Reminders
  - All public disks (/pool and /scratch only) are scrubbed once a week.
  - Reasonable requests to restore scrubbed files must be sent no later than the following Friday, by 5pm.
  - If you have any questions or encounter any problems, email:
    - SI-HPC-Admin@si.edu for sys-admin related issues,
    - SI-HPC@si.edu for Bioinformatics/Genomics questions,
    - hpc@cfa.harvard.edu for SAO users who need help.

Quick Links

- Cluster monitoring page
- QSub Generator (only accessible within the SI network)

FYI

- References to Hydra (publications, proposals, etc) should mention the cluster as:

  the Smithsonian Institution High Performance Cluster (SI/HPC).

Last updated 14 Jan 2020  SGK/MPK.