Beams-doc-8245 — April 23, 2020

Data storage options for IOTA/FAST experiments

Giulio Stancari*

Fermi National Accelerator Laboratory, P.O. Box 500, Batavia, Illinois 60510, USA

As the research program ramps up and after the completion of Run 2b, on behalf of the IOTA/FAST Scientific Committee, I outline data storage, backup and documentation options to support IOTA/FAST experiments at the Integrable Optics Test Accelerator (IOTA) and its injector, the Fermilab Accelerator Science and Technology (FAST) superconducting electron linac [1–3].

Each experimental collaboration is responsible for specifying in their proposal and in subsequent reports which Fermilab-managed resources are used to store and backup the data, to document the experiment, and to ensure reproducibility of the results.

There are of course many options. Here are suggestions that have worked well and that experimenters may find useful. The emphasis is on long-term solutions for small experiments. Solutions should be easy to access and to share.

Shared network drive beamssrv1.fnal.gov/iota-fast.bd: This is the recommended option. It is a new shared network drive with a capacity of 8 TB dedicated to raw and pre-processed data. The network path for Windows users is `\\beamssrv1.fnal.gov\iota-fast.bd`, whereas Linux and Mac OS users can access it at `smb://beamssrv1.fnal.gov/iota-fast.bd`. On the accelerator consoles at NML and ESB, the drive is mounted as `I:` to make it easily accessible. On the drive, files are organized by data-taking run and by experiment in directories with names like `yyyy-mm-dd_dataset`, containing `README` files with descriptions of the data. Most members of the IOTA/FAST group should have permissions to access this drive. For details on usage, please contact myself, Dan Broemmelsiek or Alexander Valishev. Like other shared network drives, it is maintained and backed up by the Servers team within the Networking Group of the Accelerator Division Controls Department. Daily backups can be restored by the users: on Windows, choose Properties > Previous Versions. Starting May 1, 2020, the AD Servers team will also save weekly incremental backups to tape. The tapes will be kept for up to 2 years. Please contact `ad-servers@fnal.gov` if you need to access the archives. A similar network drive, `beamssrv1.fnal.gov/nmlscrf.bd`, has been used for several years to collect documentation regarding IOTA/FAST activities. It has limited disk space and it is not specifically dedicated to storage of experimental data, but it can be used for file exchange. If you are a Mac user, please consider disabling the creation of `.DS_Store` files on the shared drives. (Instructions are widely available on the web. In summary, from a terminal, issue the command `defaults write com.apple.desktopservices DSDontWriteNetworkStores -bool TRUE`, log out and log back in.)

Fermi Redmine cdcvs.fnal.gov/redmine/projects: This is the recommended option for documentation. It supports documents, files, wikis, repositories, etc. Several experiments are already using it and help is available to get started. This is not meant as a storage space for data sets of tens of gigabytes or more, but smaller data sets of the order of hundreds of megabytes can be easily managed here. There is a limit of 300 MB for each individual file.

* E-mail: stancari@fnal.gov

Fermilab Office 365 OneDrive office365.com: Fermilab has invested in the Microsoft Office suite of products and this is an extra resource that is available to us. Currently, Fermilab users have about 5 TB of disk space available. This is not the preferred option, as support and quotas may change in the future and are out of our control (CERN is moving away from Microsoft products because of cost, for instance), but it is a good option nonetheless.

Accelerator Division Controls Cluster clxNN.fnal.gov: These machines reside on the restricted Controls network. They are useful for short- and long-term storage of ACNET data collected through ACL scripts, Python scripts, and other front-end programs.

Open Science Framework osf.io: This is a free platform funded by various institutions to support research and collaborations. It includes files, wikis, etc. It is particularly useful when not all members of a collaboration have Fermilab accounts. I have used it with colleagues at Brookhaven and CERN on several projects, for instance.

Google Drive: Several institutions and universities such as IEEE, UChicago, etc. offer this option to their members, students and employees with a typical quota of 15 GB or more. It is a good option for short-term storage and for file exchange. However, copies of all relevant data should eventually reside on a Fermilab-managed resource.

For all cloud-based solutions, besides web access and desktop applications, a very useful tool for file management and synchronization is `rclone`.

Comments and suggestions on needs or solutions that can help IOTA/FAST experiments with data collection, storage and documentation are of course very welcome.

---

[1] S. Antipov et al., IOTA (Integrable Optics Test Accelerator): Facility and Experimental Beam Physics Program, JINST 12, T03002 (2017).

[2] General information on the IOTA/FAST facility can be found at fast.fnal.gov.

[3] The web page of the IOTA/FAST Scientific Committee contains information on the experimental program, with links to experiment proposals and web pages: cdcvs.fnal.gov/redmine/projects/ifsc/wiki.