

Data Capacitor Policy

The Data Capacitor is a uniquely architected high-end system for distributed high volume data processing and flow. At the lowest level it is a high speed, high capacity filesystem architected for the short to mid term storage of large research data sets. Due to the nature of the system, all storage on the system will be transient in nature, will not be backed up, and all responsibility for data management within this space will belong to the user. The Data Capacitor has been architected to protect against disk failure, controller failure and server failure. However, we cannot account for every possible failure mode and in the event of catastrophic failure, data recovery will be attempted, but no guarantee of success can be implied.

Because the Data Capacitor was funded by a grant submitted to the NSF, system use will be governed by an allocation committee consisting of the grant's principle investigator (PI) and co-principle investigators (CoPIs). This committee will review all project requests and allocation changes. Decision criteria during the period of the grant will be heavily weighted towards research outcomes that will strengthen the success of the NSF grant. Once the grant period has passed, the criteria will be reevaluated. Priority use of the system during the grant period will be given to PI, Co-PIs and senior investigators as listed in the grant. Once the grant period has passed, this too will be reevaluated. Due to the transitory nature of the resource, requests for project allocations must include a plan for archiving the data to a persistent store, such as to CDs, HPSS, or to a local backed-up file system in your department or school.

Users must respond promptly to requests for text and information needed for annual and final reports to the NSF about the Data Capacitor and its use. This has proven critical in the past to IU's continuing ability to win awards from the NSF that support such unique high-end systems. Some requests may come even after the funding period ends, when the NSF asks IU for additional information.

All users of the Data Capacitor are required to cite the use of the facility in publications and presentations as stated in the NSF grant policy guide which is available here:

http://www.nsf.gov/pubs/2002/nsf022/nsf0202_6.html

An acknowledgment of NSF support and a disclaimer must appear in publications (including Web pages) of any material, whether copyrighted or not, based on or developed under NSF-supported projects:

This material is based upon work supported by the National Science Foundation under Grant No. CNS-0521433.

NSF support also must be orally acknowledged during all news media interviews, including popular media such as radio, television and news magazines.

Except for articles or papers published in scientific, technical or professional journals, the following disclaimer must be included:

Any opinions, findings, and conclusions or recommendations expressed in this material are those

of the author(s) and do not necessarily reflect the views of the National Science Foundation.

Data Capacitor file space is divided into two categories: Project and Scratch.

Project Space

The Data Capacitor projects directory is dedicated to long-term projects with storage requirements that cannot be met with other existing systems. Requests for project space will be submitted to Team Data Capacitor and evaluated by the Data Capacitor Allocation Committee. Project requests will include:

- Project proposal and justification
- Project participants
- TB requested
- Data rate required
- Special requests (additional mounts?)
- Project name for directory name - /N/dc/projects/project_name

The default size for project requests will be 10TB. If space requirements are greater than 10TB, a written request will be required and reviewed by the Data Capacitor Allocation Committee.

Files in project space with access times greater than 30 days may be purged.

Scratch Space

The Data Capacitor scratch directory is a temporary workspace currently available to all users of Big Red. Scratch space is not allocated and its total capacity will fluctuate based on project space requirements.

Files in scratch space may be purged after 14 days.