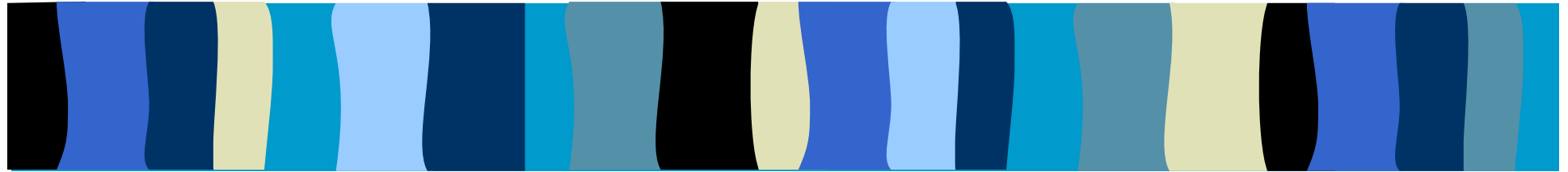


The NCAR Mass Storage System



George Williams

Scientific Computing Division

National Center for Atmospheric Research

1850 Table Mesa Drive

Boulder, Colorado 80303

gwilliam@ncar.ucar.edu

Phone: (303) 497-1843; Fax: (303) 497-1848

Presented at the THIC Meeting at the Denver Southeast Marriott

July 23, 1997

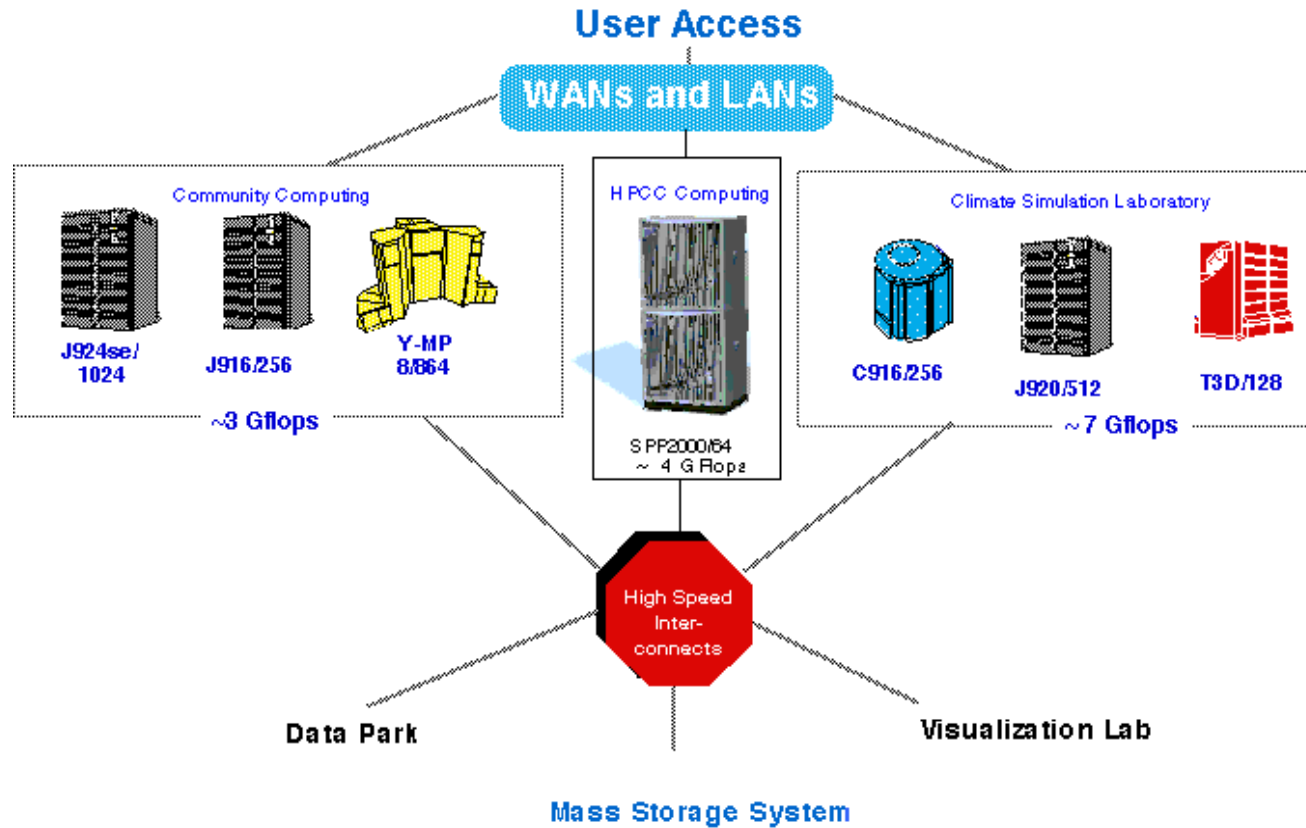


What is NCAR?

The National Center for Atmospheric Research plans, organizes, and conducts atmospheric and related research programs in collaboration with universities; provides state-of-the-art research tools and facilities to the entire atmospheric sciences community; supports and enhances university atmospheric research education; and facilitates the transfer of technology to both the public and private sectors.

NCAR is sponsored by the National Science Foundation.

Functional Diagram of the NCAR Computing Facility



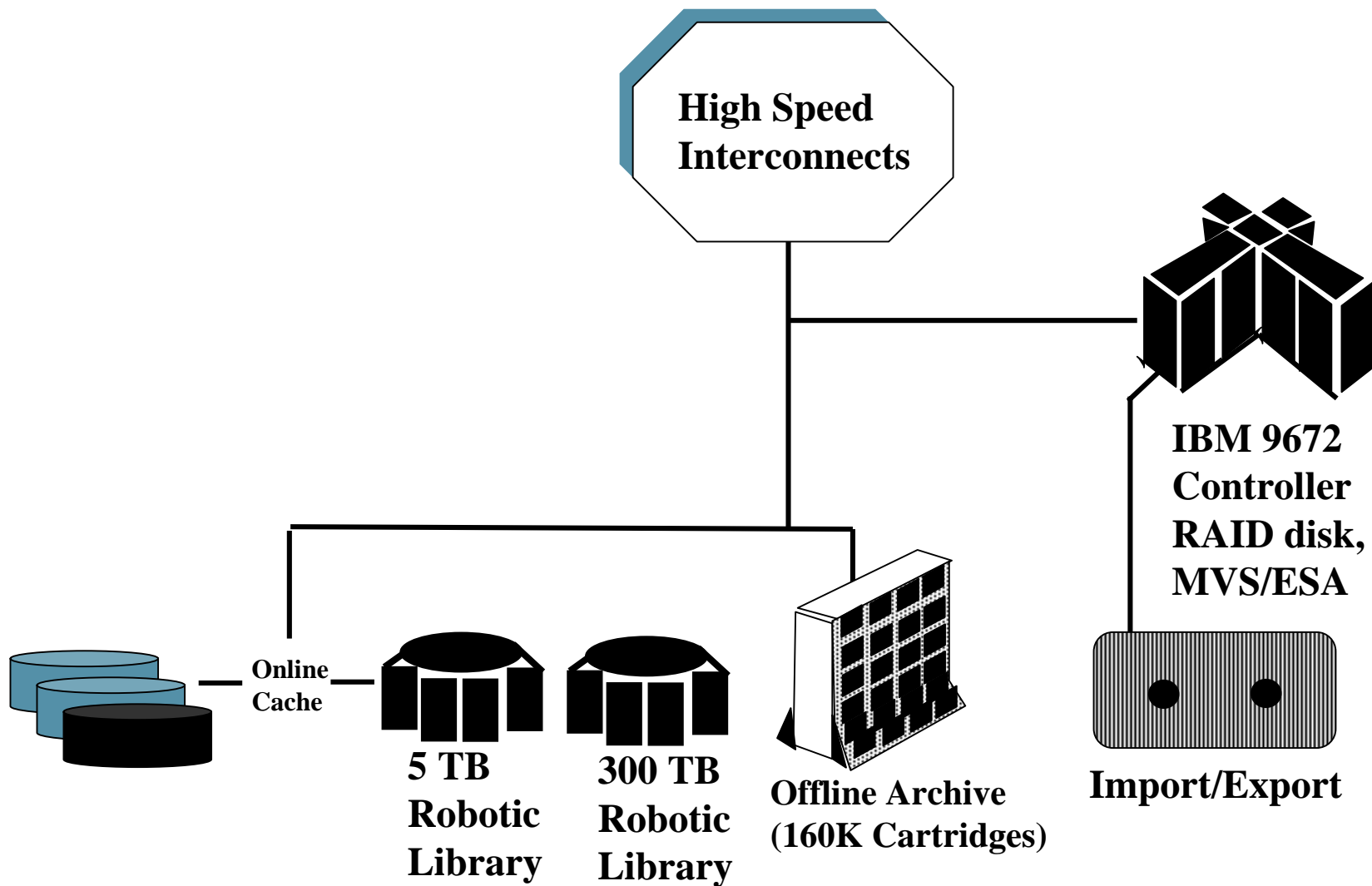
June 1997



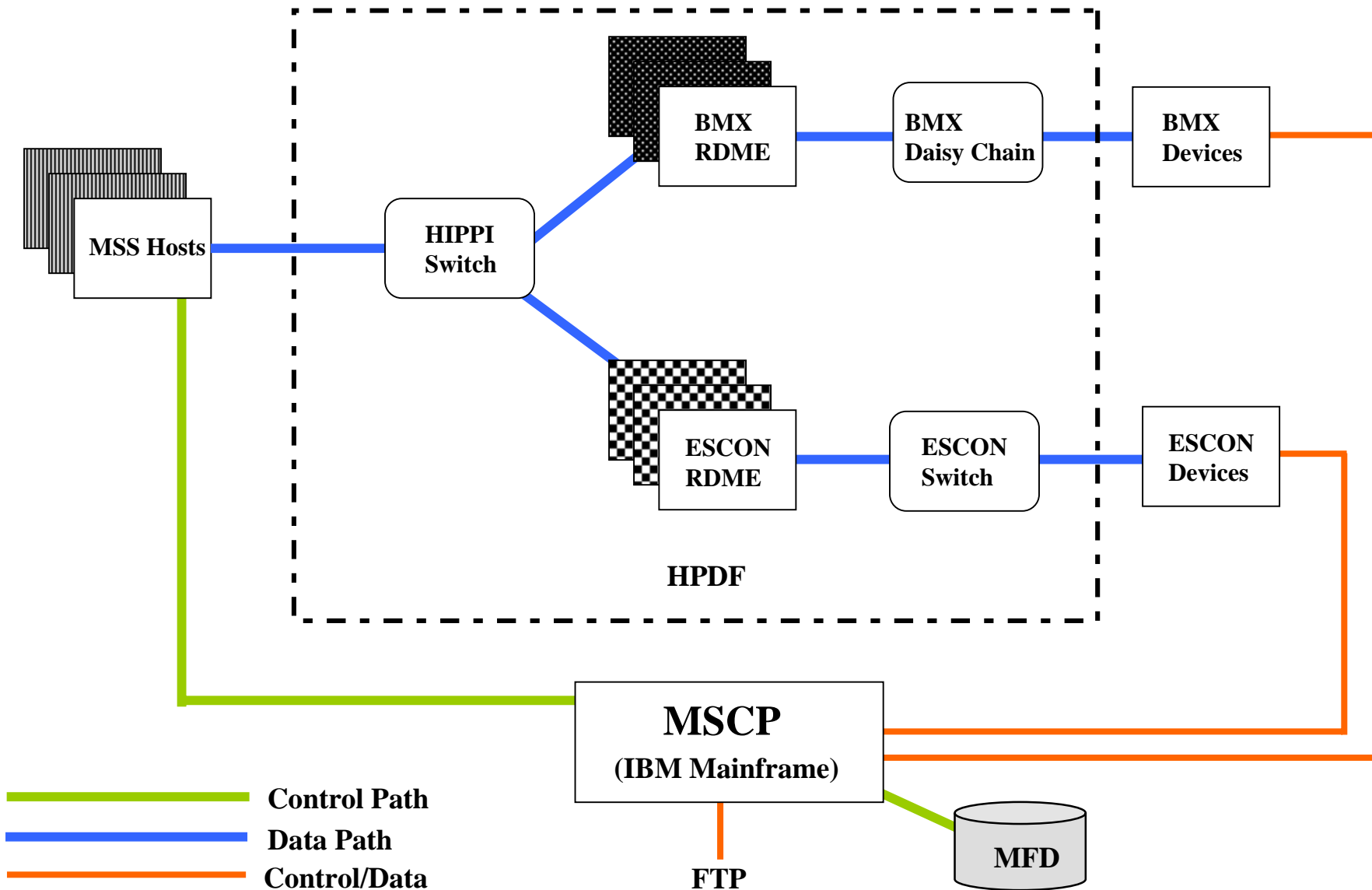
The NCAR Mass Storage System

The NCAR Mass Storage System (MSS) is a central, large-scale archive that stores data used and generated by climate models, and other programs executed on NCAR's compute servers; as well as data collected by field experiments and observations.

NCAR Mass Storage System



The High Performance Data Fabric



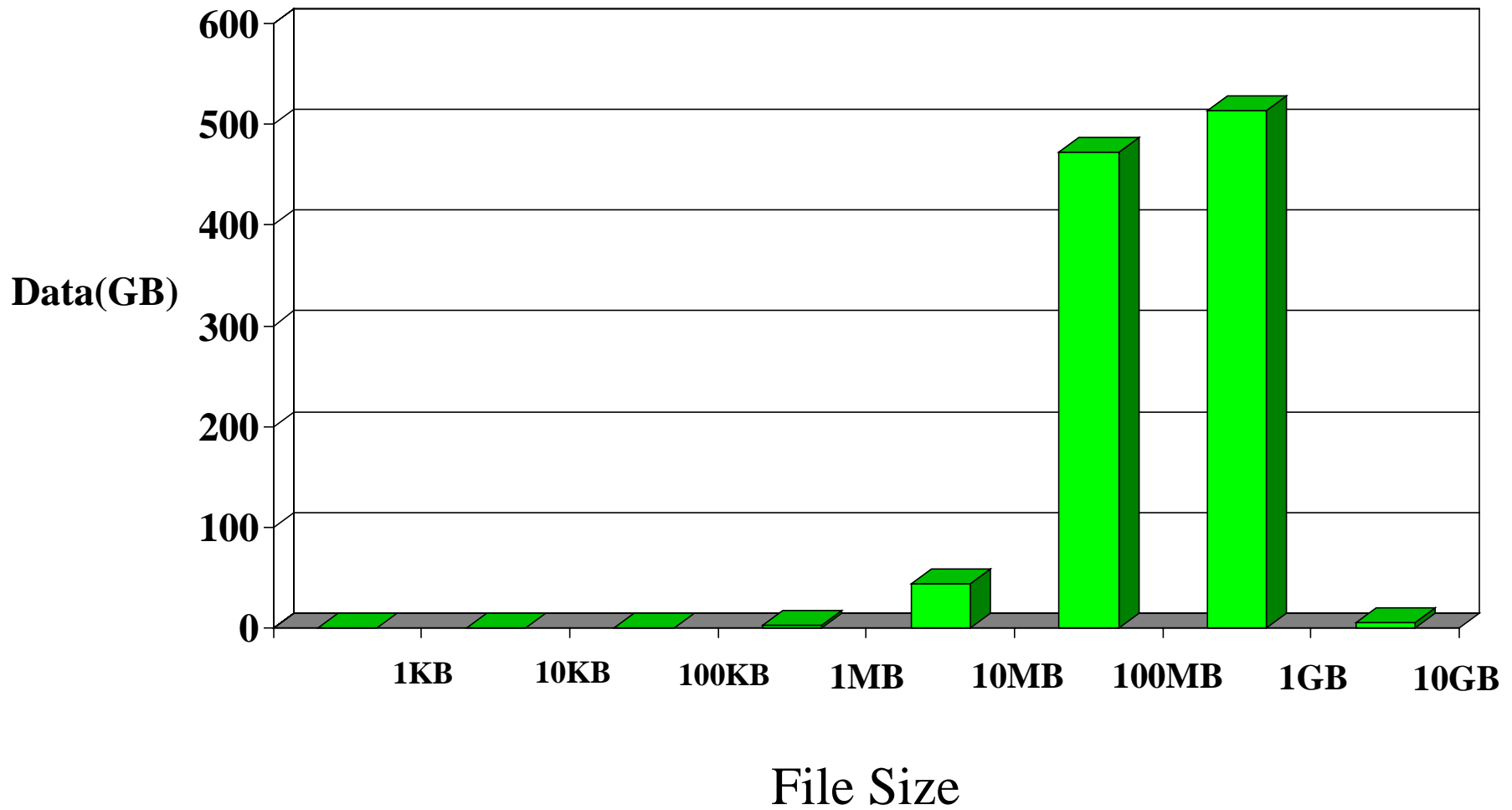


MSS Statistics

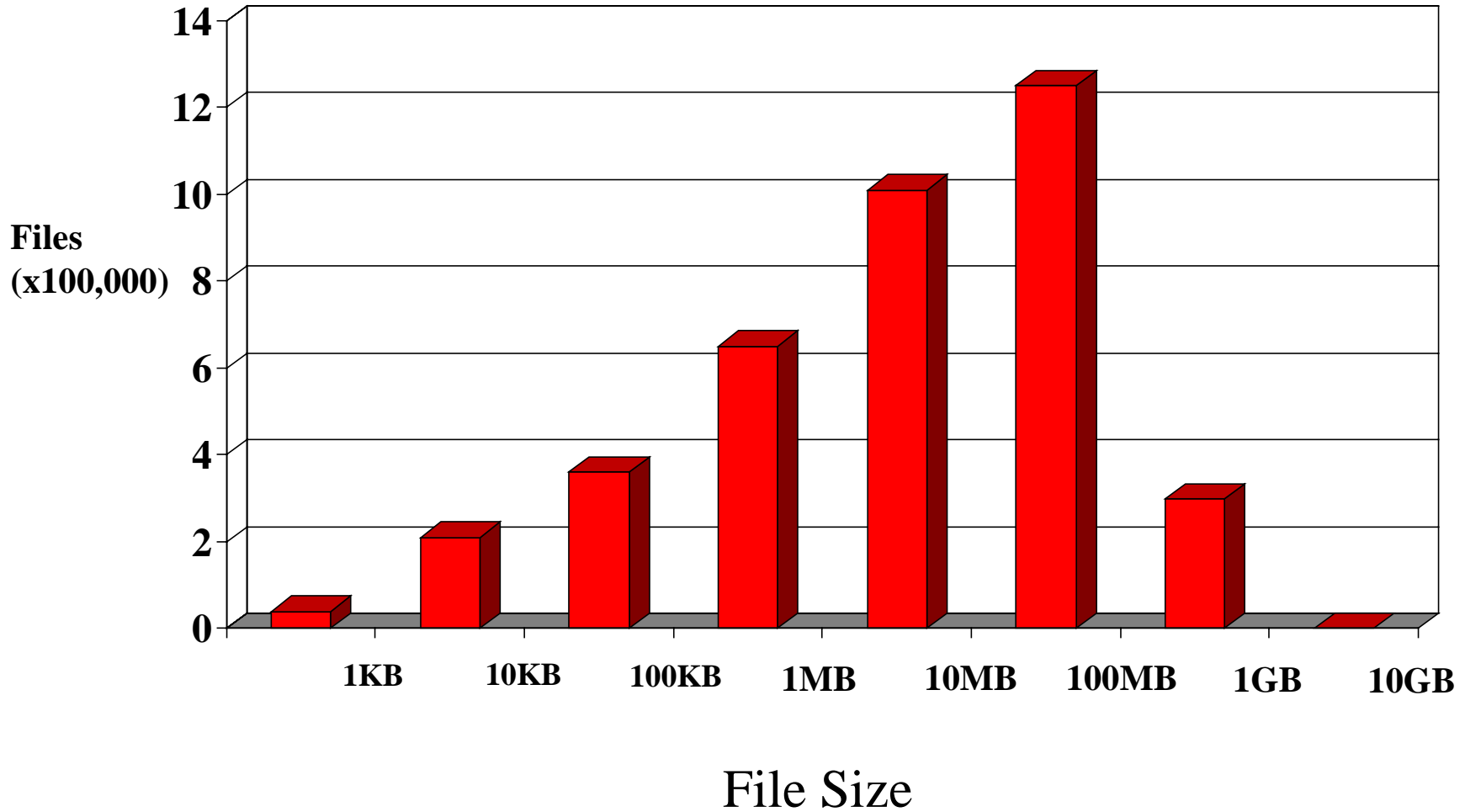
- 104 TBytes in ~3.8 million files
- Net growth of 3 TBytes/month
- Data flow of 16 TBytes/month (client reads & writes)
 - 6 MB/sec avg., ~25 MB/sec peak
- Additional 16 TBytes/month of internal data movement
- 160,000+ cartridges in the archive (manual mount)
- 60,000 manual mounts/month
- 50,000 silo mounts/month

Amount of data moved is proportional to GFLOPS available (~0.6 MB/GFLOP for client reads & writes)

Distribution of Data



Distribution of Files





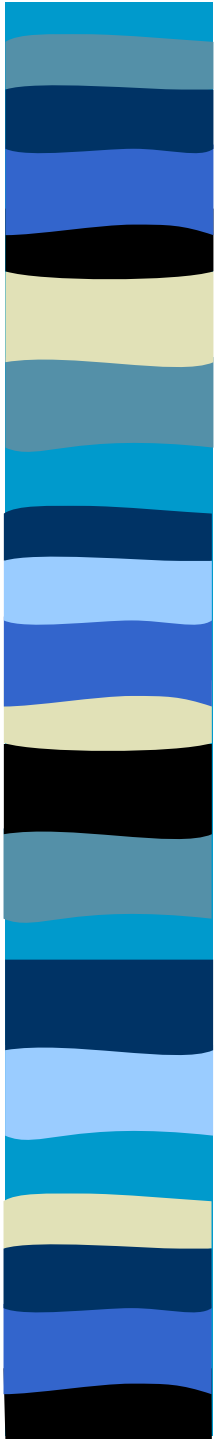
Challenges

- Accommodate increases in compute power
 - 10x increase over the next 3-4 years implies:
 - Net growth of 30 TBytes/month
 - Data flow of 160 TBytes/month
 - 60 MB/s average
 - Additional 160 TBytes/month for internal movement
 - 600,000 manual mounts/month
 - *Meltdown!*
- Adapt to Changing Environments
 - WAN/MAN/Web access
 - Heterogeneous client platforms (UNIX, Windows NT)
 - Transparency and consistency
- Cope with Budgetary Trends



Strategies

- New tape/disk/channel technologies
 - Capacity/Performance/Reliability - Choices!
- New Operational Methods
 - Sneakernet vs. Electronic copying
- Data Park - front-end file service
 - Reduce number of files sent to the archive
 - Speed matching
 - Visibility
- MSS-IV
 - Scalability
 - Adaptability
 - Maximal use of COTS products



Summary

- Keep thinking
- Keep moving
- Keep our options open
- Keep our heads above water