



## Spring Semester, 2007 — CS 780: Special Topics -- Storage Systems and Storage Area Networks

Robert D. Russell  
[rdr@unh.edu](mailto:rdr@unh.edu)

Tuesdays and Thursdays — 2:10–3:30pm  
Nesmith Hall Room 326, Durham, NH

Information is often the most important asset of an organization — without secure, reliable, uninterrupted access to huge repositories of information, most modern business and scientific endeavors would simply stop functioning.

Storage systems deal with reading, writing and permanently retaining all types of digital information. They are a dominant component in the design, operation, performance, management and cost of a computer system. Storage area networks have become the preferred means of connecting storage to applications using the storage — such networks permit consolidation and better utilization of storage, easier management of storage, and greater access at lower cost than ever before.

Today, disks for PCs are rated in tens of gigabytes ( $10^9$ ) bytes, but high-end storage systems are rated in tens or hundreds of terabytes ( $10^{12}$  bytes), and demand for multi-petabyte ( $10^{15}$  bytes) storage systems already exists. Similarly, transfer rates for simple disks are currently measured in Megabytes per second, but rates on storage area networks are measured in Gigabits per second, with higher rates already available in development labs. In spite of these impressive numbers, storage access is often a major performance bottleneck, so continued improvements in storage technology are crucial in many application areas.

Storage systems are based on components that utilize magnetic recording technology (disks and tapes), but they also deal with many related issues: security, backup and recovery, management, sharing, network access. The storage industry has been expanding at an accelerating rate, with many companies devoted exclusively to storage and/or storage networks. To remain competitive, traditional computer companies also have incorporated storage as a significant part of their business — just look at <http://www.xxx.com/storage>, where “xxx” can be ibm, microsoft, sun, hp, dell, etc.

In academia and industrial research labs there is intense activity in storage and related issues: object store, file system design, self-organizing storage, active disks, storage security, management, modeling, emerging technologies, etc. See for example, the Parallel Data Lab at Carnegie Mellon University (<http://www.pdl.cmu.edu>).

This course will cover the concepts, design, implementation and performance of storage systems and the networks used to access them. Topics will include disks, disk arrays (RAID), disk layout, file systems, network file systems (NFS), cluster file systems (GFS, GPFS), storage protocols (SCSI), storage area network technologies (Fibre Channel, iSCSI), object store. For each topic we will consider issues of design, performance, reliability, security, management, standards and cost. Hardware will be discussed, but emphasis is on software.

Students should have a basic knowledge of computer operating systems and computer networks (i.e., CS620 Operating Systems Fundamentals or equivalent).