Chapter 6
Data Storage and Storage Networks

Direct Attached Storage

- Used where data sharing is limited to specific departments
- Used where centralized management is not an issue
- Consists of data storage devices physically attached to a single server
  - Same enclosure, or connected by cabling in a separate enclosure
Direct Attached Storage

- Ideal for expanding the capacity of individual servers
- Easily installed and configured
- More drives can be added as capacity requirements grow

Direct Attached Storage

- Limitations
  - Capacity is limited by the physical structure of the server and external enclosures
  - Best used by small or departmental networks, in which a fixed number of users require shared access
  - Data is inaccessible if the server crashes
Direct Attached Storage

- Limitations
  - Unused storage space is unavailable to other departments
  - Each server requires its own backup solution
  - Network data traffic can slow down the rest of the network
  - Does not scale well

Direct Attached Storage

- RAID
  - Stands for redundant array of independent disks
  - One practical implementation of DAS
Direct Attached Storage

- RAID 0
  - Stripes data across two or more disks for improved Read/Write performance
  - No data redundancy - single drive failure results in loss of data from whole set

Direct Attached Storage

- RAID 1
  - Replicates and maintains identical copies of data on two sets of drives
  - Redundant and fault-tolerant
  - Commonly called data mirroring
Direct Attached Storage

- RAID 5
  - Stripes data across three or more drives
  - Also stripes parity information (for error correction)
  - Fault-tolerant, but not redundant

Direct Attached Storage

- RAID 6
  - RAID 5 plus an additional drive just for parity
  - Can recover from the failure of two drives
  - Not redundant
Direct Attached Storage

- RAID 1 + 0
  - Mirrors two striped arrays
  - Redundant and fault-tolerant

Network Attached Storage

- A server computer that is optimized to provide fast and efficient shared data-storage services to network clients
- No network services other than file sharing
- Specialized operating system designed for optimum file-sharing performance
Network Attached Storage

- **Common Internet File System (CIFS)**
  - Originally known as Server Message Block (SMB)
  - Developed by IBM in the early 1980s
  - Functions within the upper three layers of the OSI model
  - Common on Windows operating systems
  - Linux systems can connect to CIFS shares using Samba

Network Attached Storage

- **Network File System (NFS)**
  - Developed by Sun Microsystems; released in 1984
  - Standard file-sharing system in all UNIX/Linux systems
Network Attached Storage

● Network File System (NFS)
  ● Originally designed for encapsulation in UDP packets
  ● Switched to TCP packets in version 3 for improved performance and more efficient use of bandwidth

Network Attached Storage

● Other NAS Protocols
  ● HTTP
  ● FTP
  ● WebNFS
  ● NetWare Core Protocol (NCP)
  ● AppleTalk Filing Protocol (AFP)
Network Attached Storage

- Why use NAS?
- Why not just set up a Windows server that supports CIFS, or a Unix or Linux server that supports NFS?

Network Attached Storage

- Why use NAS?
  - Different solutions require different sets of expertise
  - General-purpose operating systems provide sub-optimal file-sharing support
  - Other servers perform better once they no longer have to share files
Network Attached Storage

- NAS Terminology
- NAS Appliance
  - NAS device in which file-storage devices are self-contained in one computer that can be connected to the network
- NAS Head
  - NAS device to which users connect
  - Connects to storage devices through network storage fabric

Network Attached Storage

- NAS Terminology
- Network Storage Fabric
  - Specialized network that connects NAS head with storage devices
- NAS Gateway
  - Allows network users to connect to a SAN
Network Attached Storage

- NAS Configuration Options
- NAS devices are generally configured through a web interface, or through SSH

Network Attached Storage

- NAS Configuration Options
  - **High-Availability NAS Clustering**
    - Provides continuous access to data even if a NAS server crashes
    - Two or more NAS heads are linked in a cluster, which is connected to a shared set of storage disks
Network Attached Storage

- NAS Configuration Options
  - High-Availability NAS Clustering
    - Any head can process any request -- allows for active/active failover
    - Data is not fault-tolerant without mirroring
  - Network Data Management Protocol (NDMP)
    - Automatically backs up NAS data

Network Attached Storage

- Advanced NAS Technologies
  - Non-Volatile RAM (NVRAM)
    - Cache that retains its contents even during crash, reboot, or power loss
  - Direct Access File System
    - RAM in the NAS device performs as an extension of client's RAM
Network Attached Storage

- Advanced NAS Technologies
- iSCSI
  - Maps the SCSI protocol on top of TCP; allows transfers to avoid translation to CIFS/NFS first

Storage Area Networks

- A Storage Area Network (SAN) is a network that uses Serial SCSI to store and retrieve vast amounts of data
Storage Area Networks

- Historical Reasons for SAN Usage
  - Originally, to separate backup traffic from LAN traffic and to provide shared access to centralized data storage
  - As data storage and backup needs grew, with the advent of large databases, backups had to be performed continuously
  - SANs moved data storage and backups off the LAN while still providing shared access to data

Storage Area Networks

- Contemporary Reasons for SAN Usage
  - High availability
  - Disaster recovery
  - Business continuation
Storage Area Networks

- **SCSI** (Small Computer System Interface)
  - Provides parallel data transport mechanism for data delivery
  - An alternative to IDE, SATA, etc.

Storage Area Networks

- **SCSI** (Small Computer System Interface)
  - Limitations:
    - Insufficient when connecting many computers to a single set of shared storage devices
    - Only 16 storage devices may be connected to the same cable
    - Limitations on cable length
Storage Area Networks

- **SCSI** (Small Computer System Interface)
  - Limitations are overcome by **Serial SCSI**
    - Theoretical limits: 16 million hosts can connect to 16 million storage devices located miles away

Storage Area Networks

- **Data Block Transfer Service**
  - Transfers data between devices without the assistance of a file system
  - Therefore, it can move large amounts of data as quickly as possible
  - Data blocks are processed by a file system such as NFS or CIFS for presentation to the user
Storage Area Networks

- Fibre Channel
  - A commonly-used Serial SCSI protocol that provides Data Block Transfer Service

- Fibre Channel
  - Architectures:
    - Point-to-point: one server and one storage device
    - Fibre Channel Arbitrated Loop: connects up to 126 servers and storage devices with a hub
    - Fibre Channel Fabric Topology: connects up to 16 million devices with a FC switch
Storage Area Networks

• Fibre Channel
  • Fibre Channel makes IP Storage Networks possible using FCIP, iFCP, or iSCSI