

# Data Transfer

- Many programs use a disk file paradigm for I/O.
- Before networks, transferring data from one computer to another required the use of a removable medium (disk or tape) and the "sneakernet."
- A network allows direct communication.
  - File transfer - the equivalent of transferring data by disk or tape.
  - Remote file system - access to files on a networked computer through the same interface as access to local files.

# Two Historical Problems

- Historically, there were two problems with distributed computations.
  - Coordination of applications on many computers.
  - Recovering from a failure.
- Both problems are solved by saving intermediate results.
  - Components of a distributed computation need not be run concurrently.
  - Intermediate results can be used to restart a failed computation.

# Features of File Transfer

- Allow transfer of arbitrary files.
- Accommodate different file types.
- Convert between heterogeneous systems.
  - Different data types.
  - Different word lengths.
  - Different rules for file names.
- Allow user login.

# Batch & Interactive Transfer

- Batch transfer:
  - A user creates a list of files to be transferred through an interface program.
  - The interface program drops the request into a queue.
  - A transfer program reads the request and performs the transfer.
  - Good for slow or unreliable transfers.

# Batch & Interactive Transfeer

- Interactive transfer:
  - A user starts a transfer program.
  - Actions include listing contents of directories and transferring files.
  - The user can find and transfer files immediately.
  - More convenient.
  - Provides quick feedback in case of errors.

# FTP (RFC 959)

- File Transfer Protocol (FTP) - a protocol used to transfer a complete file from one computer to another.
- Enormously popular.
- One of the oldest protocols still in use. Dates back to 1971.
- Predates TCP/IP. Adapted to TCP/IP later.
- A general-purpose protocol.
  - OS and hardware independent.
  - Transfers arbitrary files.
  - Accommodates file ownership and access restrictions.

# FTP Model & Interface

- FTP permits either batch or interactive use. Most use interactive.
- MIME and HTTP can use FTP directly.
- FTP actions include:
  - Listing the contents of a directory.
  - Changing to a different working directory.
  - Getting a file.
  - Putting a file.

# FTP Commands

- See page 401 for a list of FTP commands.
- The FTP client interface from BSD UNIX is the *de facto* standard.
- In UNIX, FTP is invoked with **ftp**.
- Many commands are archaic and no longer used.
- The most frequently used commands are **cd**, **dir**, **ls**, **get**, **put**.
- A typical FTP session might go as follows:
  - Launch **ftp**.



# FTP Commands

- Connect to a remote host.

Involves logging into a user account on a remote host.

Some FTP servers provide anonymous FTP.

- Navigate through the directory hierarchy to find the directory you want.

- Download or upload a file.

Use **get** to download a file, **mget** for multiple files.

Use **put** to upload a file, **mput** for multiple files.

- Quit the session.

# An FTP Session

- Try the sample session.
- Note that each message from the server includes a 3-digit decimal number.
- Verbose mode shows the messages; quiet mode suppresses the messages.
- Various applications have been written to make FTP-ing easier.

Example: Fetch

- Can also use a Web browser to FTP.

Example: **ftp://ftp.sun.com**

# File Types & Transfer Modes

- There are many different file types.
- FTP does only two basic types of transfer to accommodate most files:
  - Text for text files.
  - Binary for everything else.

# FTP Client-Server Interaction

- FTP uses the client-server model.
- When a user starts an FTP session, the FTP client requests and the FTP server accepts a TCP control connection on port 21.
  - The client side uses the control connection to send the user ID, password, and commands to the server.
  - The control connection remains open during the entire session.

# FTP Client-Server Interaction

- When a user requests a file transfer, the FTP server opens a TCP data connection on port 20.
  - FTP sends exactly one data file over a data connection.
  - After a file is sent, a data connection is closed.
  - If another file is to be sent, another data connection is opened.
- The two functions of sending commands and data transfer are separated.

## TFTP (RFC 1350)

- Trivial File Transfer Protocol (TFTP) - a simple protocol used to transfer a file from one computer to another.
- Differences between FTP and TFTP:
  - TFTP uses UDP instead of TCP.
  - TFTP supports only file transfer—no interaction. For example, there's no directory listing.
  - TFTP does not have user authorization.
- TFTP is often used for bootstrapping a diskless hardware device over a network. All the device needs is a small ROM into which TFTP, UDP, and IP are hardwired.

# NFS

- Network File System (NFS) - a remote file access mechanism to allow applications on one computer to gain access to files on a remote computer.
- NFS is the TCP/IP standard. Developed by Sun.
- There are other distributed file systems:
  - AppleShare
  - Microsoft/NETBIOS
- NFS is based on the client-server model, but has an interface quite unlike FTP.
- Instead of a separate client application, NFS is integrated into a computer's file system.

# NFS

- A special directory is created in a computer's file system and associated with a remote computer.
- When an application performs an operation on a file in the special directory, NFS client software performs the operation on the file in the remote system.
- The remote file system appears to be on the local file system, and remote files can be processed just as if they were local.
- Our system uses NFS, and all files are kept on Triton. We say that your user file system on Triton is mounted onto the local directory of any machine you log into.



# NFS

- Question: Why would one want to do this?