DATA STORAGE TECHNOLOGIES & NETWORKS (CS C446 & IS C446)

LECTURE 32 – DAS, NAS, SAN, FIBRE CHANNEL

Storage

- Separation of Concern:
  - Computation vs. Storage
- Components
  - Client (data user)
  - Server (data processor)
  - File System (data manager)
  - Disks (passive data repository)

Alternative: Database System

- 80% of data is unstructured / semi-structured
  - File systems are
    - still a better option than relational database for storing such data.
    - and often simpler even under complex processing requirements.

Storage on the Network

- What does a Computer Network achieve?
  - Communication
  - Complexity management for computation
    - Local computation vs. Non-local Computation
  - Collaboration
- Are these applicable for (persistent) storage as well?
  - Yes
  - Then, storage should be on the network as well.
- Recall from Lecture 1
  - Persistent Shared Data

Storage Requirements [3]

- Persistent Shared Data
  - Data is accessible to (or accessible through) multiple computers and persistent across computations
    - Storage is shared by multiple computers i.e. available on a network
    - Question: Is the network same as the "network of computers"?
    - Question: Is the network "transparent" to the computers?

Network Attached Storage (NAS) -

uses Computer Network for file access

Storage Area Networks (SAN) -

uses (separate) Storage Network for (raw) data access

I/O Architecture - Evolution

- Up to the 80s

	Interconnect Distance	Bandwidth (typical)
I/O Channels	**<10^2 meters**	**10 MBps**
LAN	**>10^2 meters**	**1 MBps**

I/O Architecture - Evolution

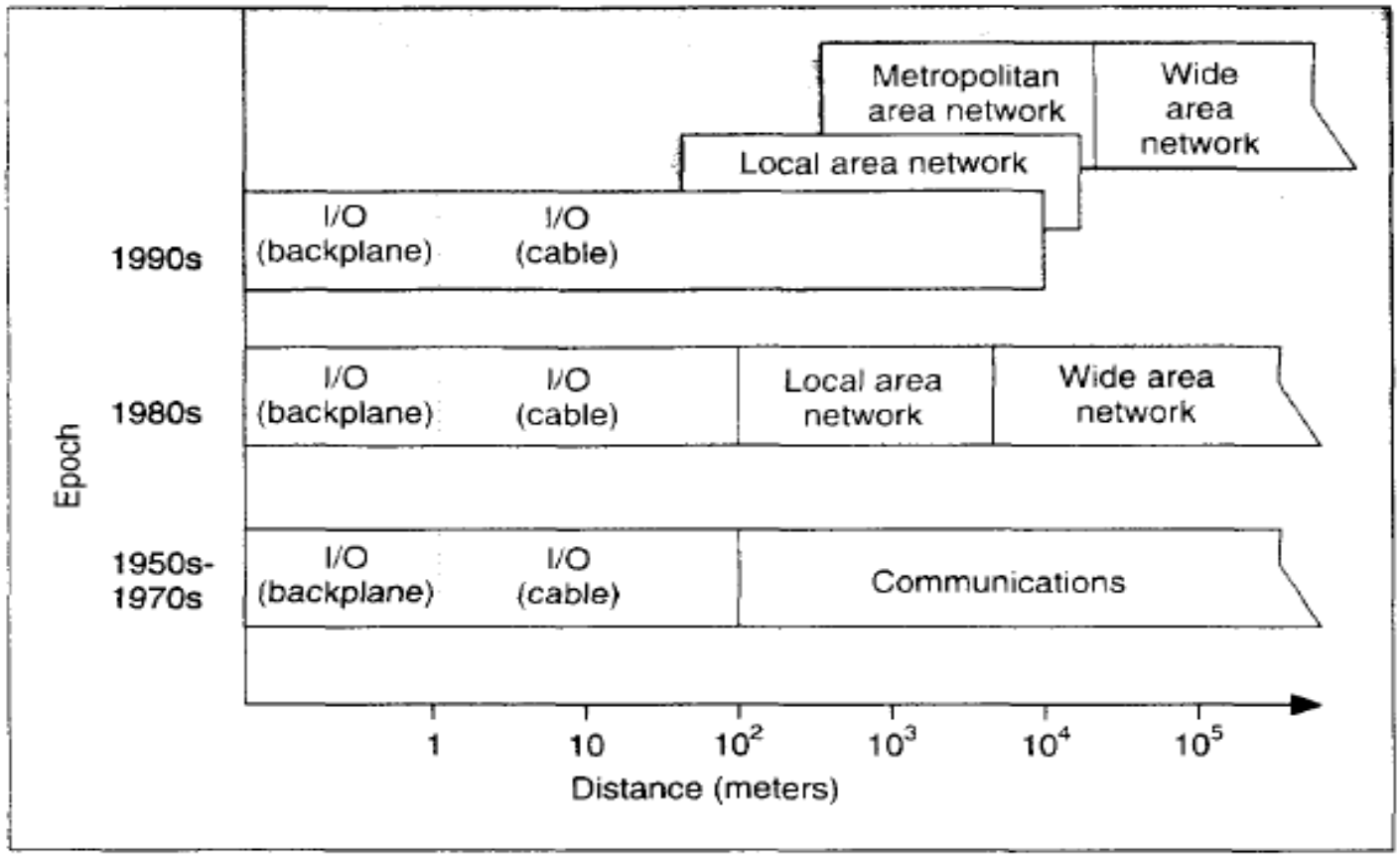


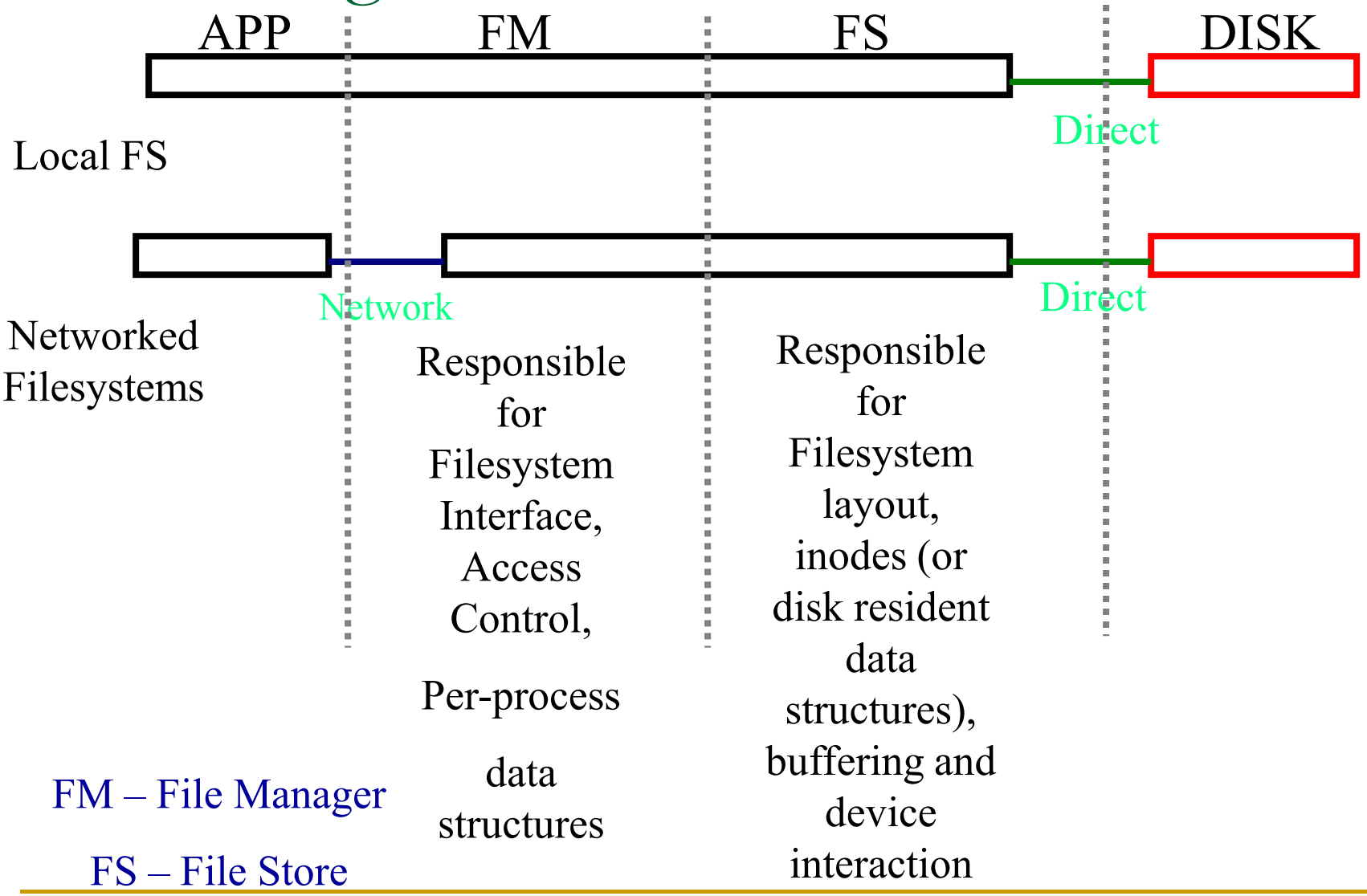
Figure 1. Interconnect type and interconnect distance for various epochs.

Figure from Sachs, Leff, and Sevigny

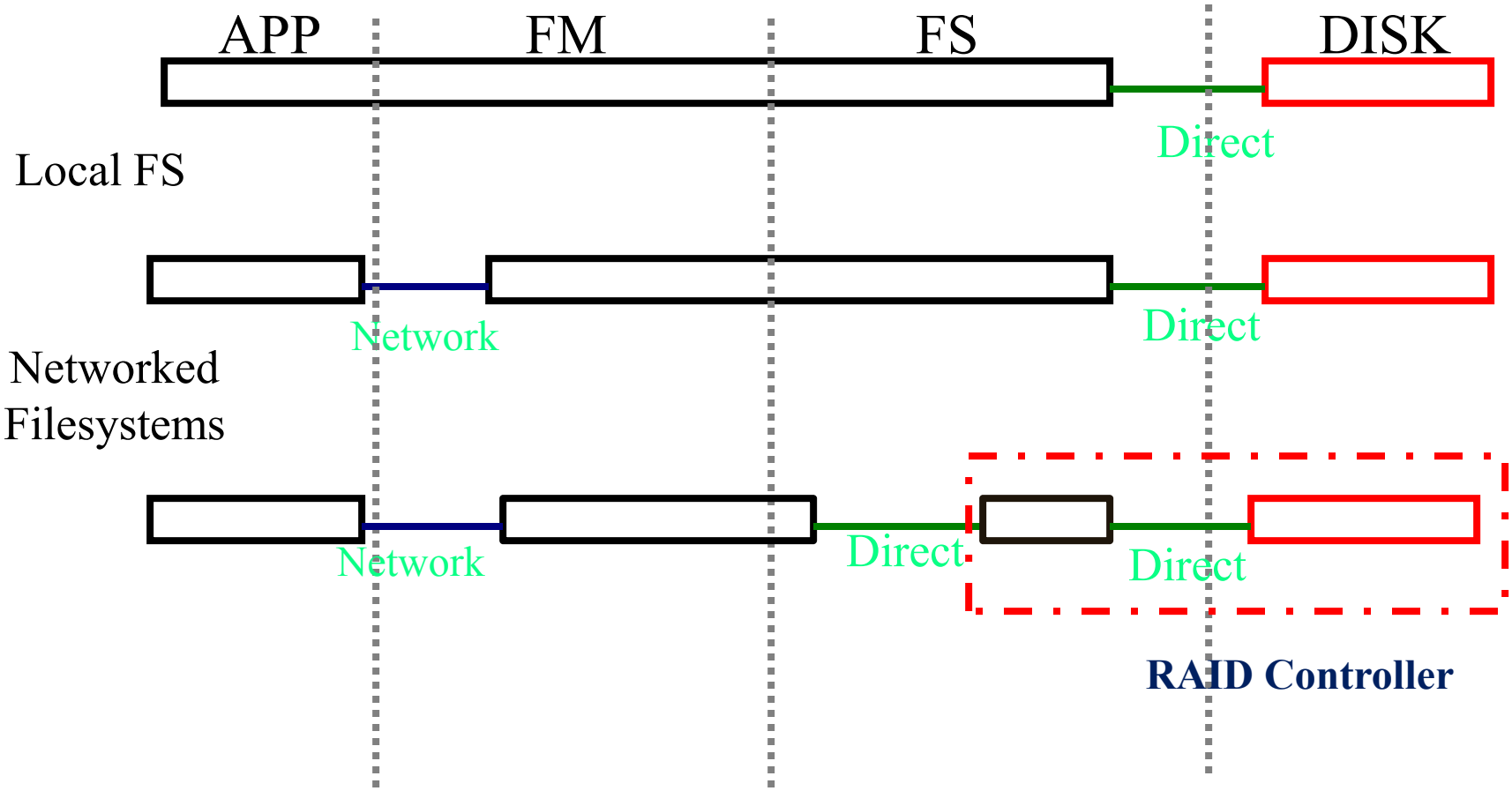
I/O Architecture - Evolution

- Blurring of Technology
  - I/O channels for longer interconnect distances and LANs with higher bandwidth
  - Networks for smaller areas (e.g. DAN, PAN, CAN)
    - LAN, WLAN, WAN, MAN, SAN (Storage / System / Server / Small), CAN (Campus / Controller / Cluster), PAN (Personal), DAN (Desk)
  - Application impact on I/O
    - Larger (sized) I/O operations
      - E.g. Multimedia applications
    - More concurrent/distributed computations
      - Require remote I/O
        - E.g. Distributed Shared Memory/Storage

Storage Architecture - Evolution



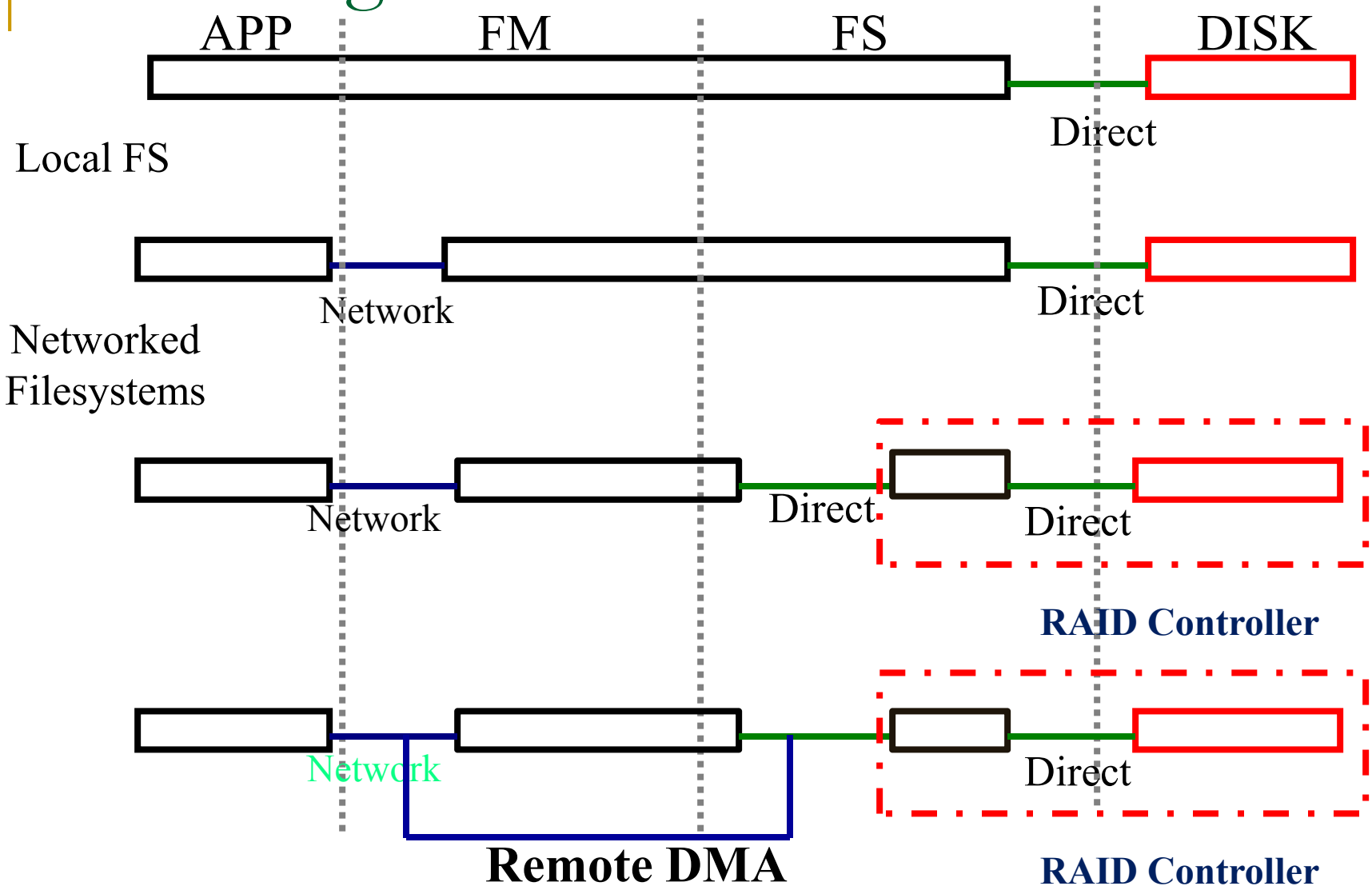
Storage Architecture - Evolution



FM – File Manager

FS – File Store

Storage Architecture - Evolution



RDMA – enables the network adapter to transfer data directly to or from application memory, eliminating the need to copy data between application memory and the data buffers in the operating system

DAS

- Storage connects directly to servers
- Applications access data from DAS using block level protocols
- Examples: internal HDD, tapes, directly connected external HDD
- DAS communicates with each other by using predefined protocols
  - IDE/ATA, SATA, SAS,SCSI and FC [implemented on the HDD controller]

Types of DAS

- Types of DAS
  - Internal
    - Connected through parallel internal bus. Internal buses has limitation on # of devices connecting to it, distance and occupies large amount of space inside host.
  - External
    - Communication btw host and storage device takes place over SCSI or FC protocol
    - Overcomes the distance and device count limitations and provides centralized management of storage devices

DAS – Benefits and Limitations

- Benefits
  - Simple and Lower initial investment
  - Configuration is simple and can be deployed easily and rapidly
  - Setup is managed using host based tools such as host OS
- Limitations
  - Does not scale well. Has limited number of ports
  - Limited bandwidth [restricts I/O processing capability]
  - Distance limitation [can overcome by fibre channel]
  - Does not make optimal use of resources due to limited ability to share front end ports
  - Unused resources can not be reallocated easily
  - Disk utilization, throughput, cache memory, virtual memory, RAID – level configuration, storage controller protocols and efficiency of the buses affect DAS performance.

Small History of NAS and SAN

- In 1979, Shugart Associates defined a way to have shared disk devices. Named it as SASI
  - Shugart Associates System Interface
- Leads to SCSI (in 1982 by ANSI task group X3T9.3)
  - Small Computer System Interface
- In 1980, Sun Microsystems developed NFS
- 1984, Syntec developed NetBIOS for IBM
  - NetBIOS → SMB → CIFS
- CIFS is the predominant method of sharing files in Windows environment
- SAN is next evolution of SCSI
- NAS is next evolution of NFS and CIFS

Storage on the Network - Pragmatics

- Primary goal – sharing of data / storage
  - NAS – sharing at filesystem level
  - SAN – sharing at disk(s) level