Session T08

NAS and iSCSI -- Which to deploy?

John L. Hufferd

IBM @server xSeries Technical Conference

Aug. 9 - 13, 2004

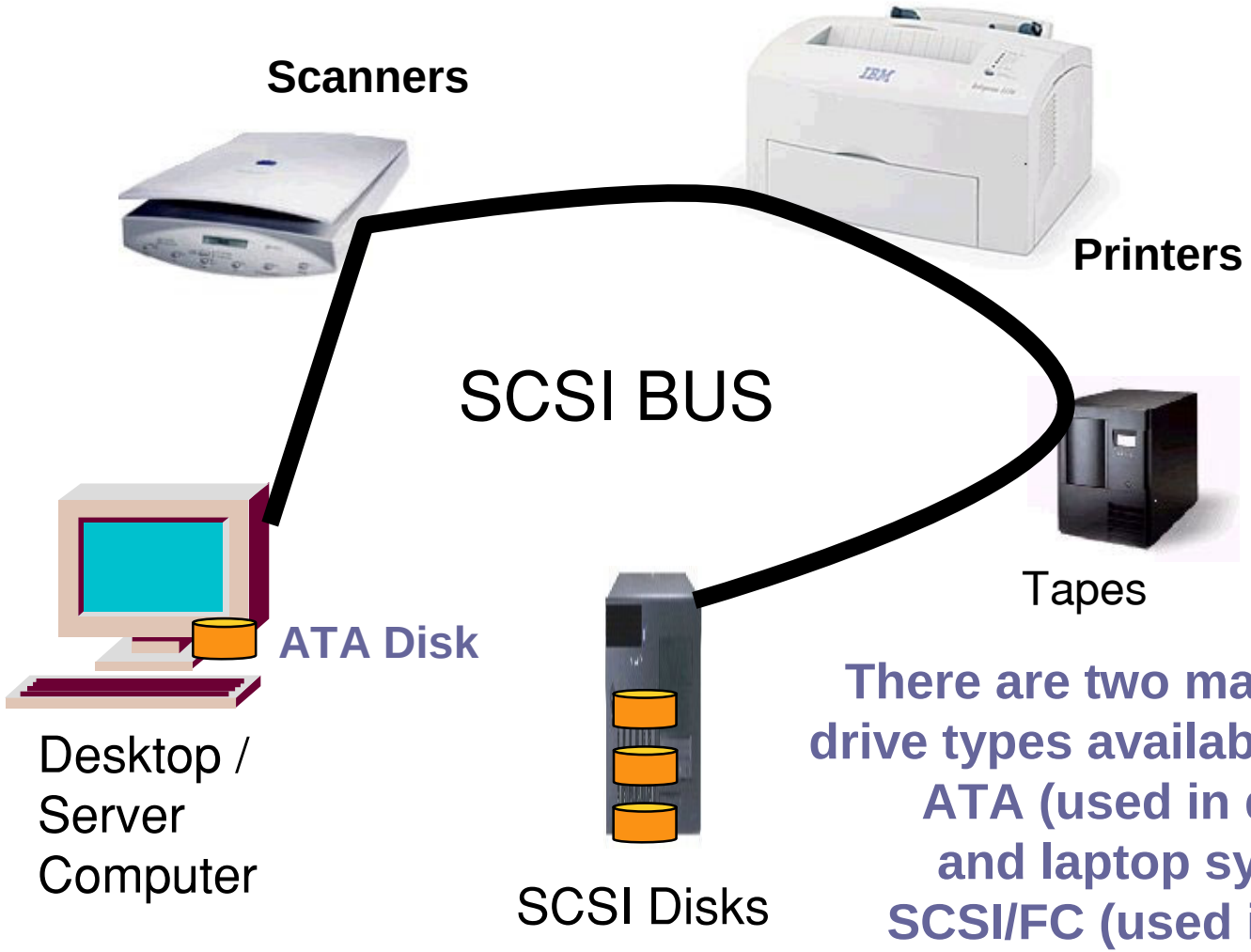
Chicago, IL

RETURN TO INDEX

Terms

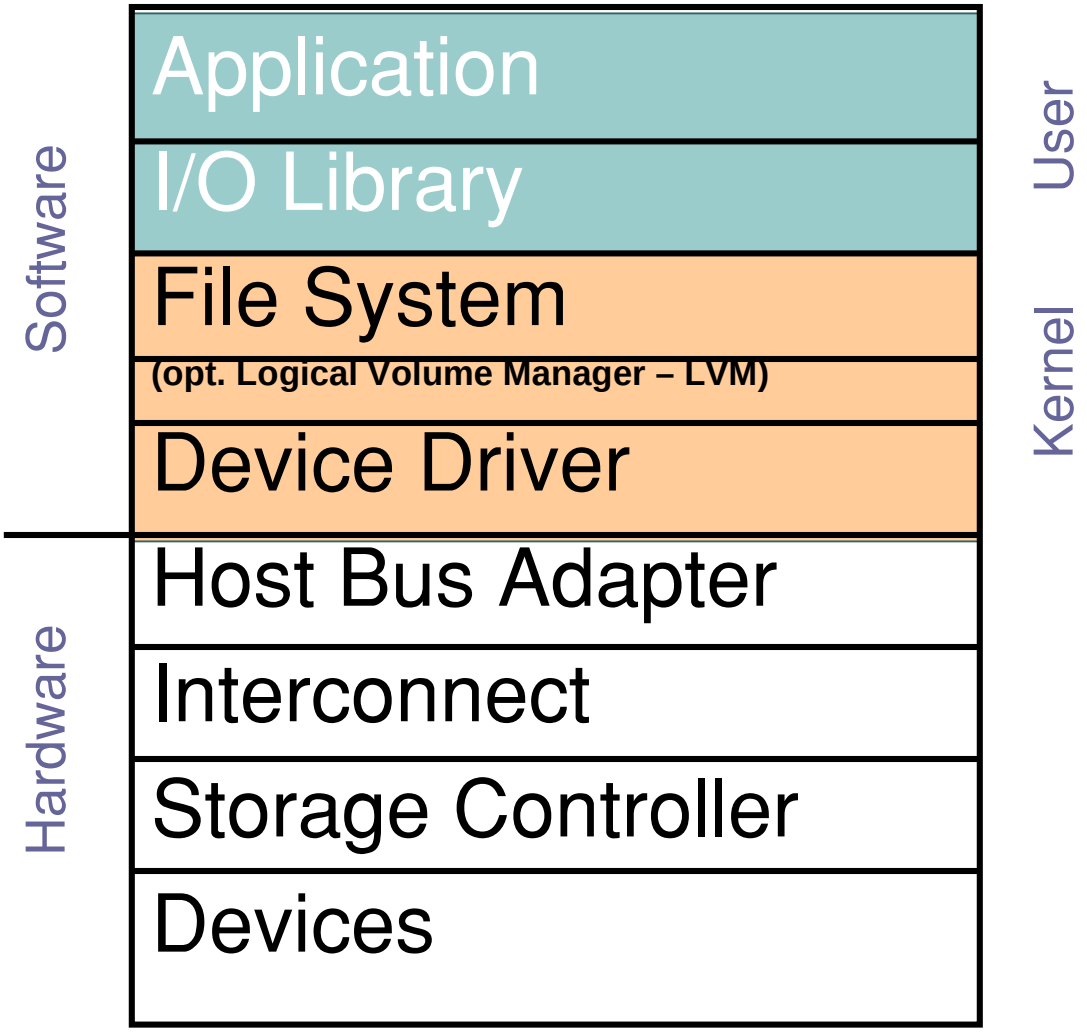
- ATA – Advanced Technology Attached (desktop disk attachment)
- CIFS – Common Internet File System
- CLI – Command Line Interface
- FC - Fibre Channel
- FS – File System
- GUI – Graphical User Interface
- HBA - Host Bus Adapter
- iSAN - iSCSI Storage Area Network
- iSCSI -- Internet SCSI
- NAS -- Network Attached Storage
- NFS – Network File System
- SAN - Storage Area Network
- SAS – Serial Attached SCSI (new SCSI disk attachment)
- S-ATA – Serial ATA (new desktop disk attachment)
- SCSI - Small Computer System Interconnect
- TOE - TCP/IP Offload Engine
- VFS – Virtual File System

Small Computer System Interconnect (SCSI)



There are two main hard drive types available today:
ATA (used in desktop and laptop systems)
SCSI/FC (used in server-class systems)

Storage System Components



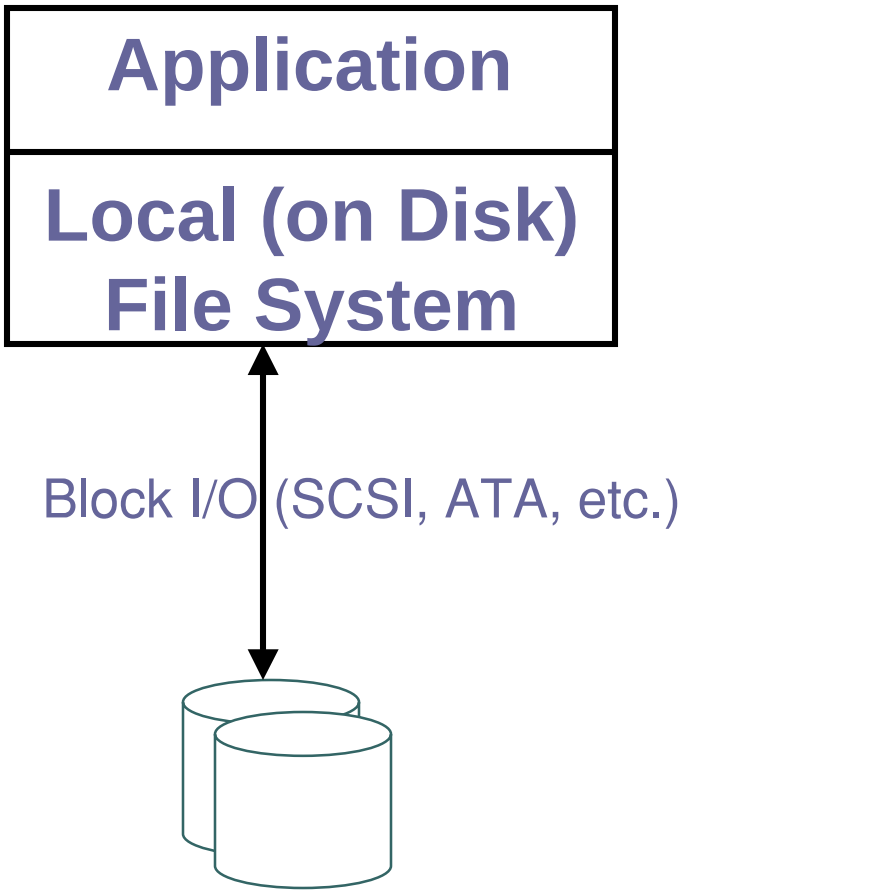
What is a File System?

- A (kernel) process that
  - Organizes storage into “clumps”, called Files
  - Keeps organizational information (Metadata) about the Files
    - Location of pieces of the files
    - Creation and access dates & times
    - etc.
- Applications have interfaces, via the File System, to access the files into which application data can be placed

File System Access

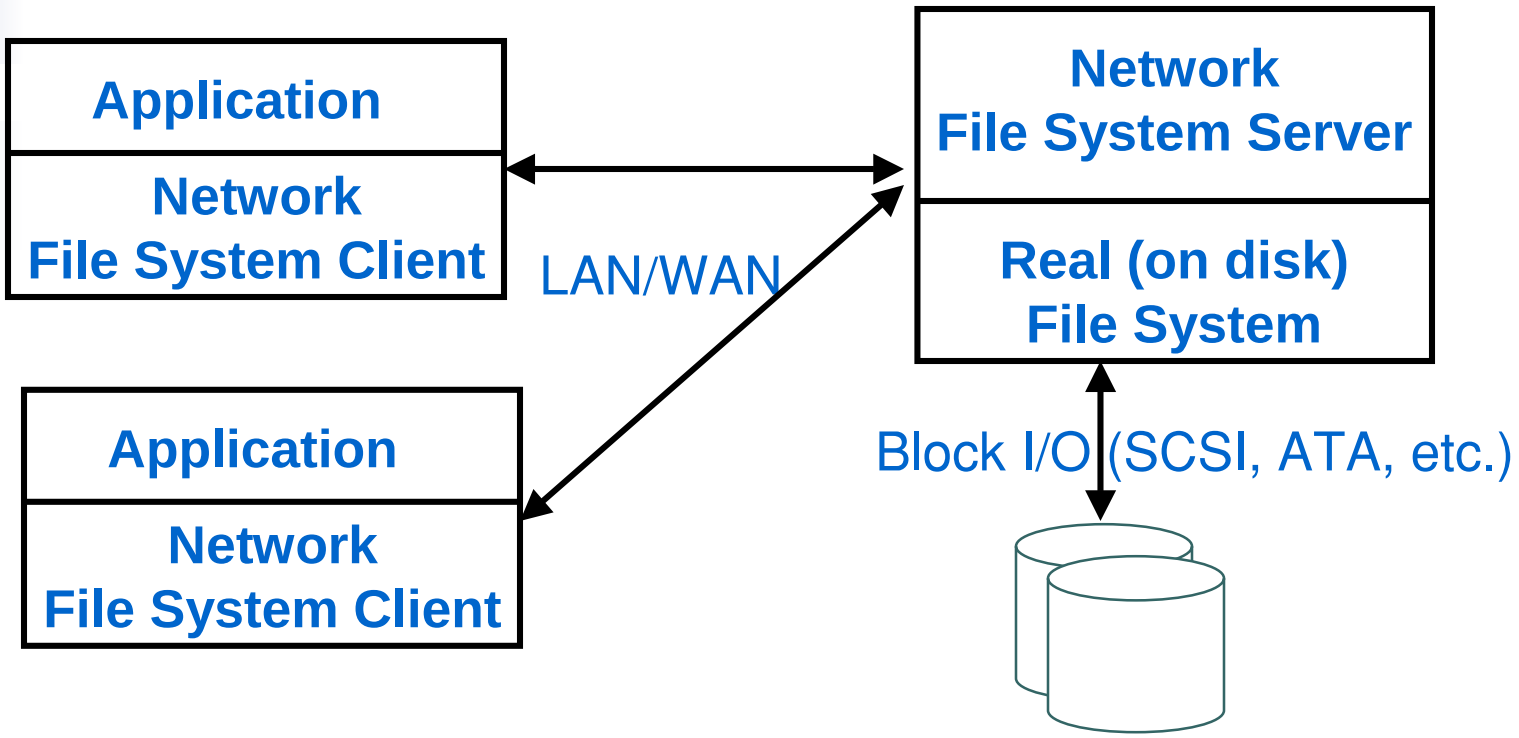
- **FS Called through “standard” interface**
  - VFS layer in Unix
  - FS interface in Windows
  - Usually has standard API (e.g., POSIX)
  - Windows has their own APIs
- **Talks to lower layers**, usually:
  - Class drivers & device drivers
  - (optionally logical volume manager- LVM)

Local File System



Direct Attached or Network (SAN) Storage

Network File System



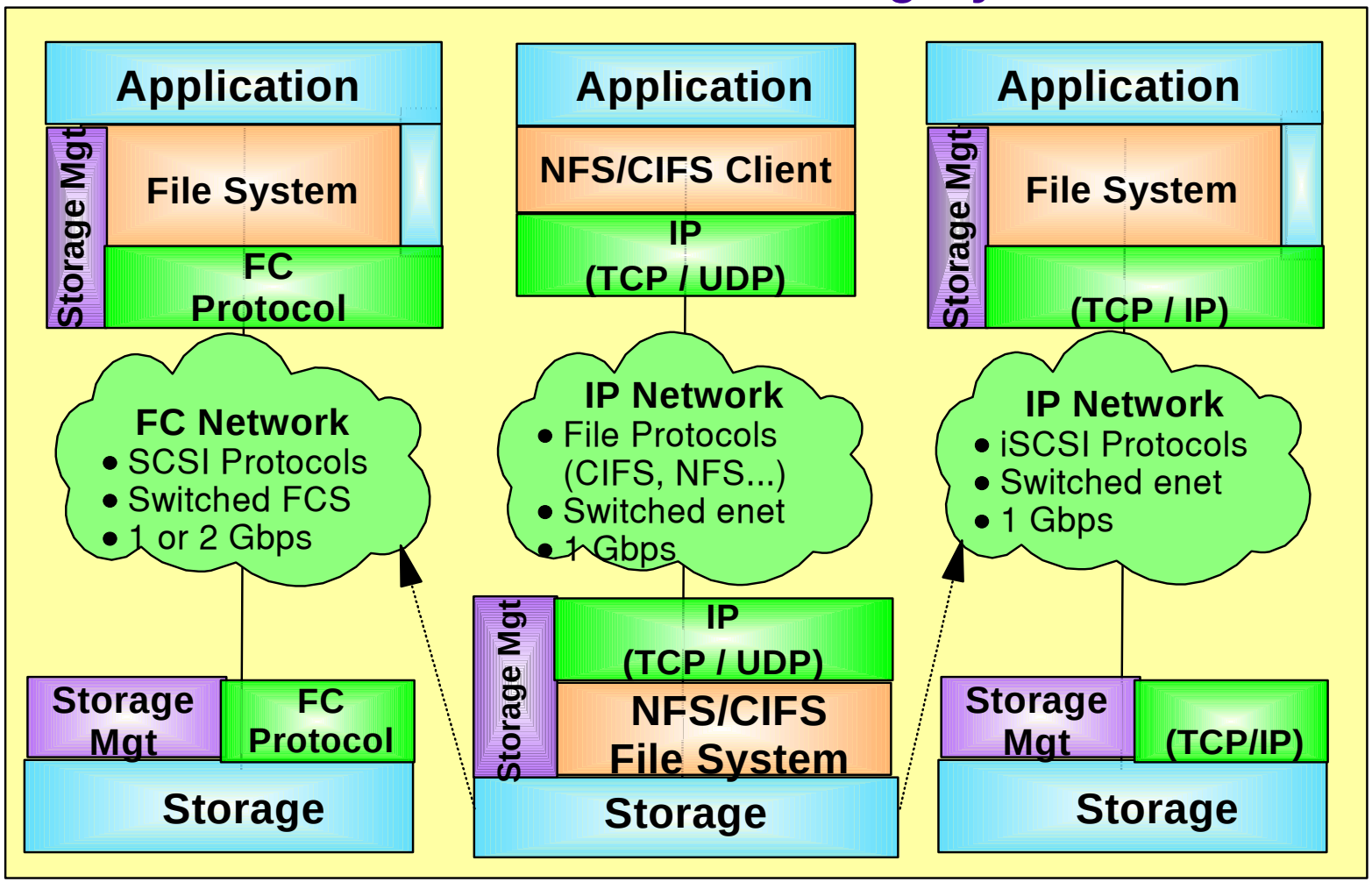
Direct Attached or Network (SAN) Storage

Types of Storage Attachments

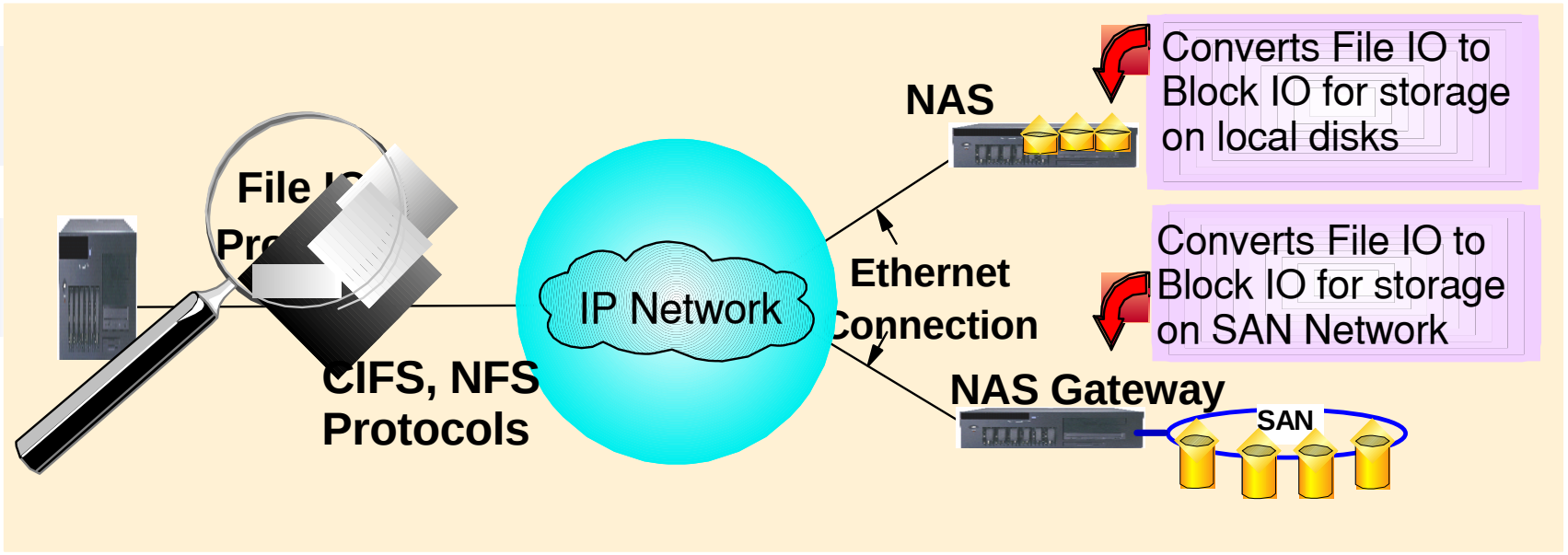
- **Locally Attached**
  - ATA
  - SCSI, FCAL
  - S-ATA, SAS, …
- **External (Pooled) Storage**
  - Attached via a Network
    - Fibre Channel Block I/O
    - TCP/IP
      - Network File System (NFS, or CIFS) aka NAS
      - iSCSI Block I/O

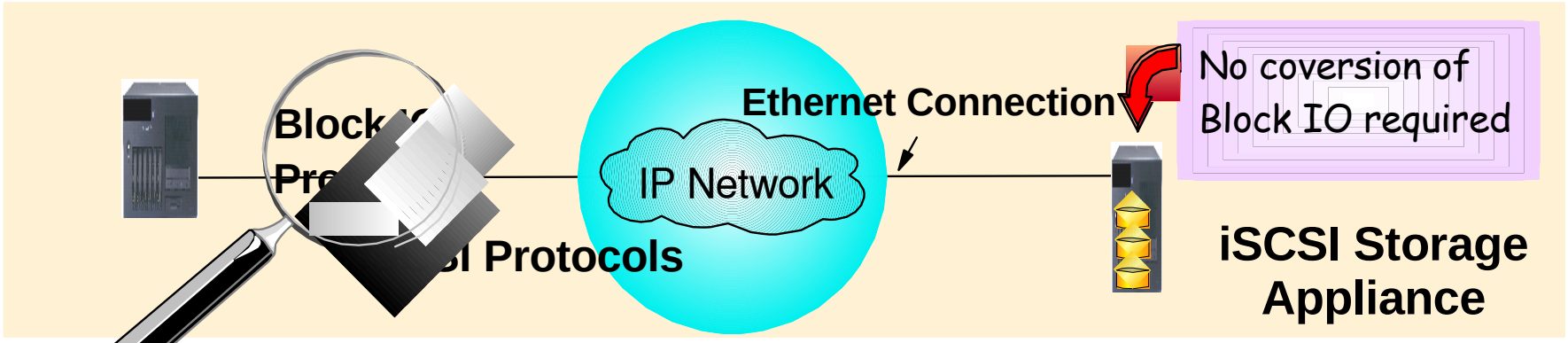
Contrasting Storage networking Technologies

Functional Placement / Processing Cycles



Network Topology Comparison





Block I/O vrs NAS

- **Two different Core requirements**
  - File I/O
  - Block I/O

  (But there are competitive overlaps)

- **NAS is deployed today for two reasons**
  1. End User Needs it (they have files to share)
  2. The Administrator Needs it (to simplify their Work)

- **Block I/O is the basic element of I/O**
  - Most valuable in unstructured I/O to storage
  - Valuable in Non Shared File Environments

Who Needs File Sharing?

- **Engineers Need File Sharing**
  - Common Macros, Development Libs, Simulation of Designs, etc
  - This is true for **other collaborating creators**
  - But even Engineers do not share the majority of their files

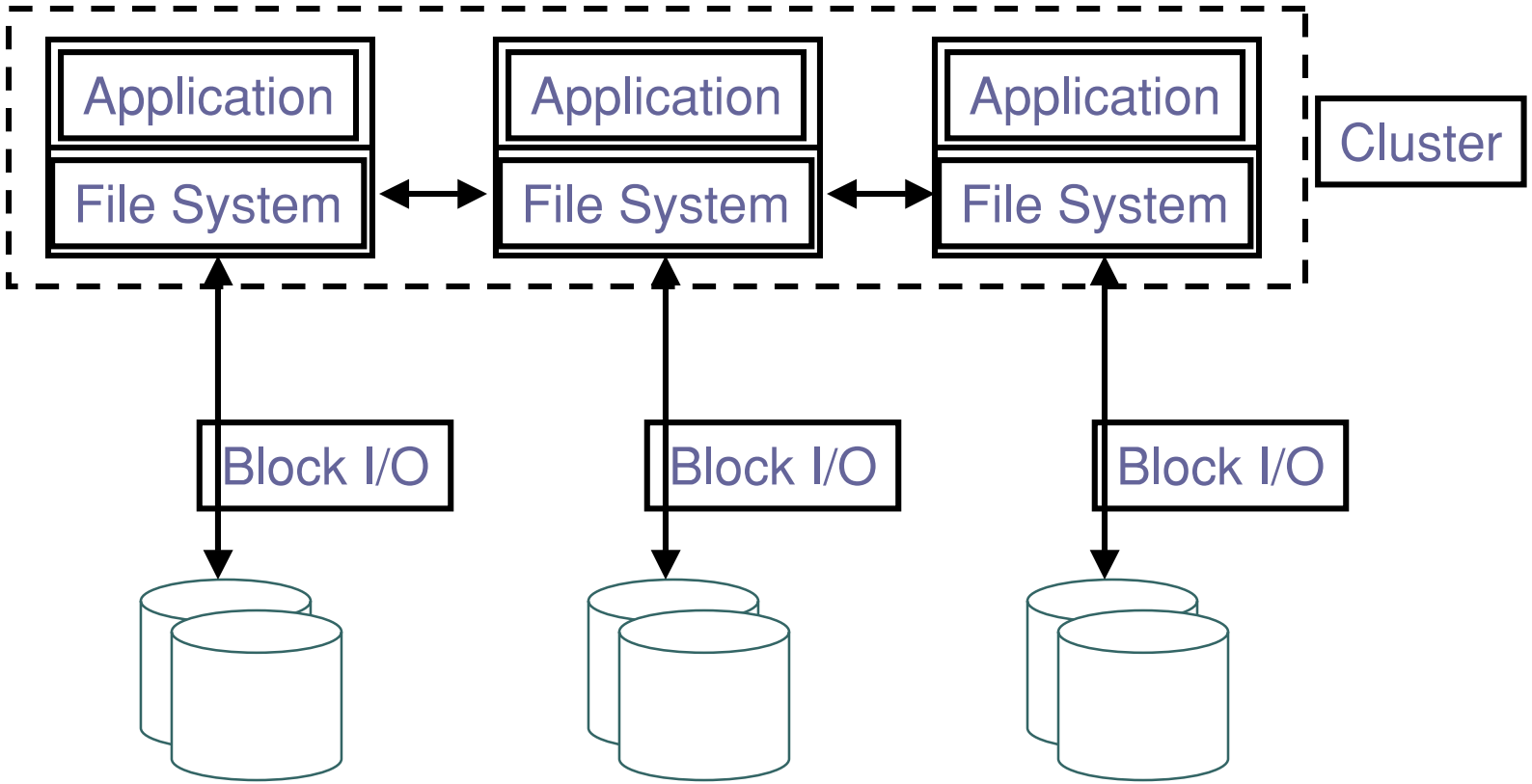
- **Most Non Engineers do not need Shared Files**
  - Files are generally shared serially via e-Mail
  - Shared things are often in a Data Base
  - Probably less than 10% of non engineers need real File sharing
    - And then only for a very few files

New Evolving Applications that Need Shared Files

- **Web Catalogs**
- **Content Servers**
- **…**

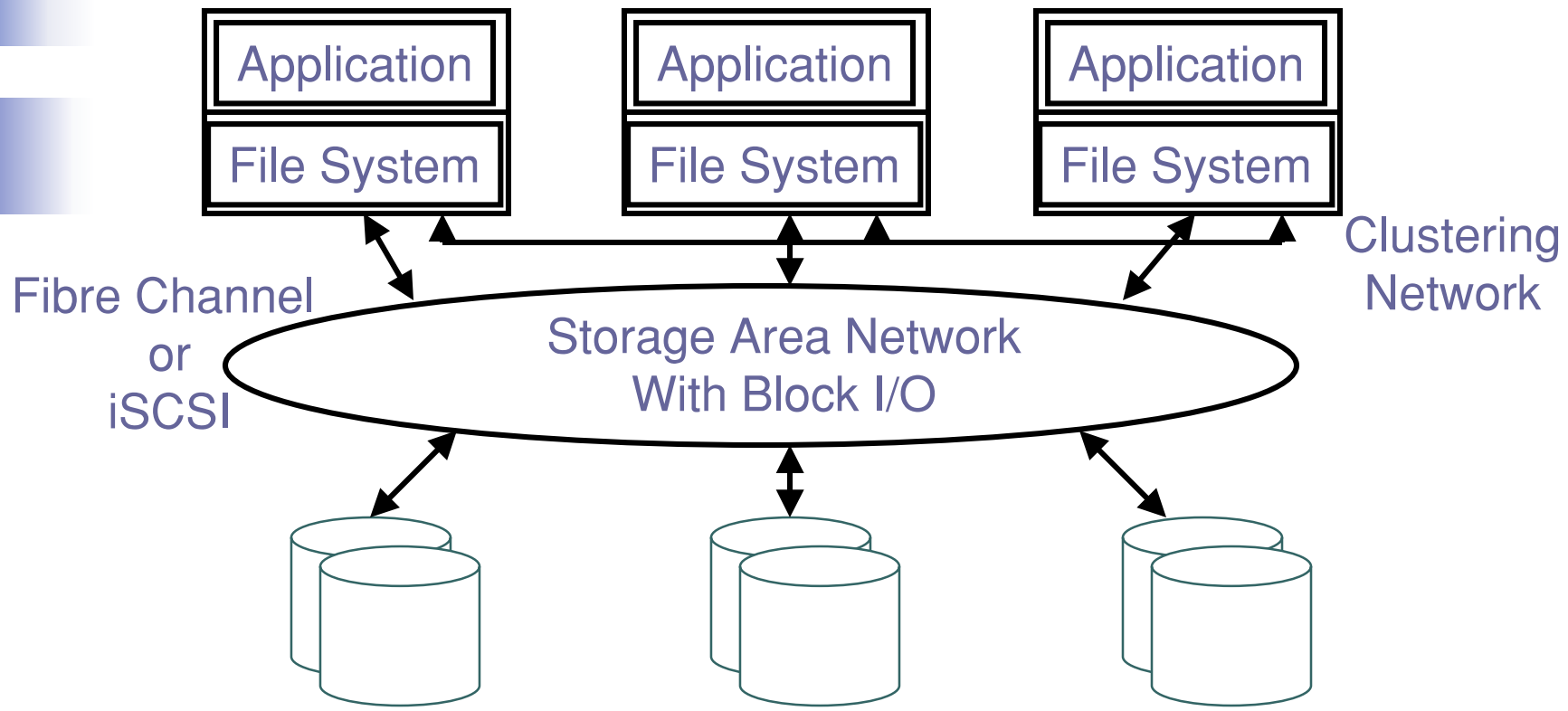
- **Can be addressed via**
  - NAS
  - Special (Shared File) Environments
    - Shared (Clustered) File Systems and Block I/O
    - SAN File Systems and Block I/O

Special (Shared File) Environments - Clustered File System – (Shared Nothing)



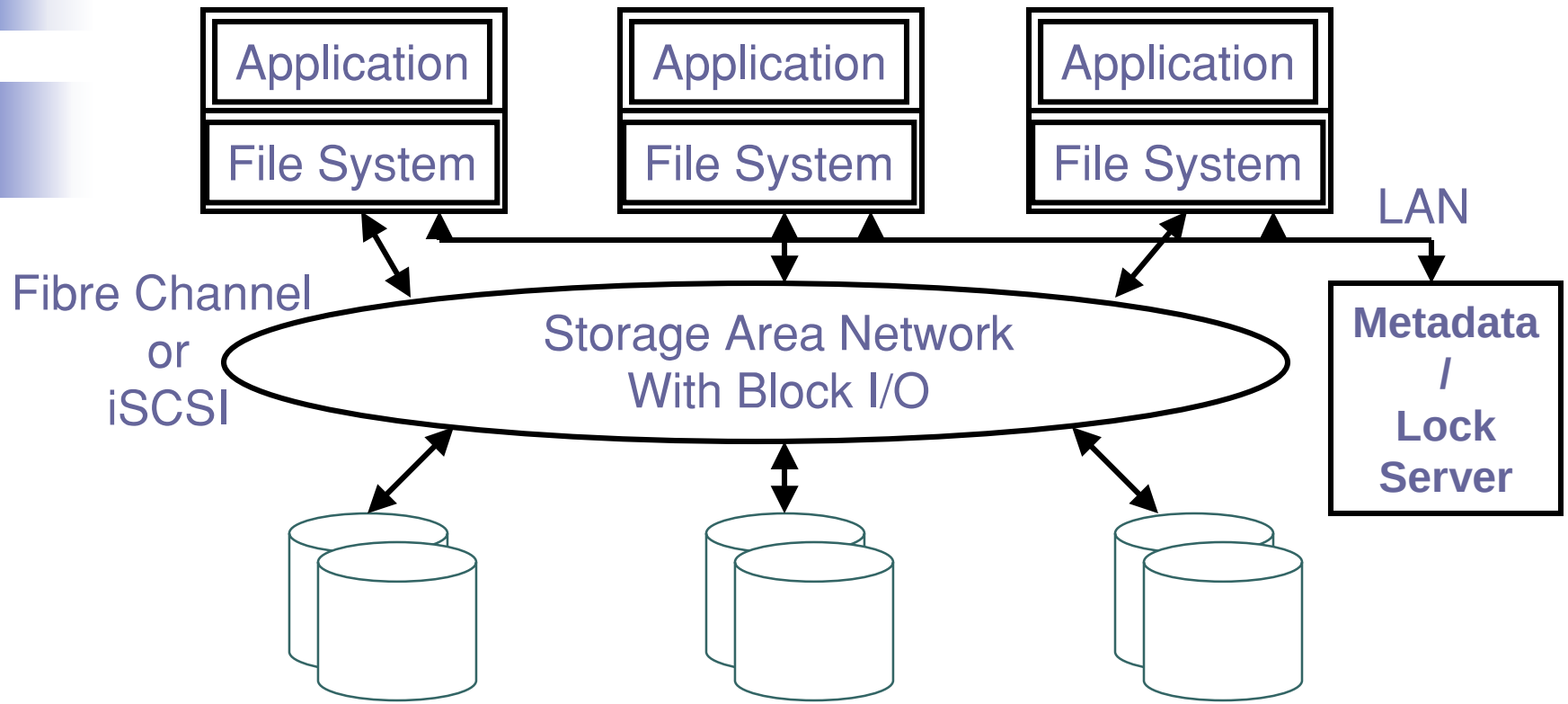
Data sent from owning system to using system on Cluster network

Special (Shared File) Environments - Clustered File System – (with Shared Storage)



Only Meta Data and Locking/State information sent on Clustering Network
Application data is accessed directly via the (i)SAN

Special (Shared File) Environments - SAN File System -



iSCSI Management
(Is iSCSI easy to Administrate?)

- **iSCSI vendors claim to be able to join a system to a Virtual Disk as easily as NAS client can be joined to a file system**

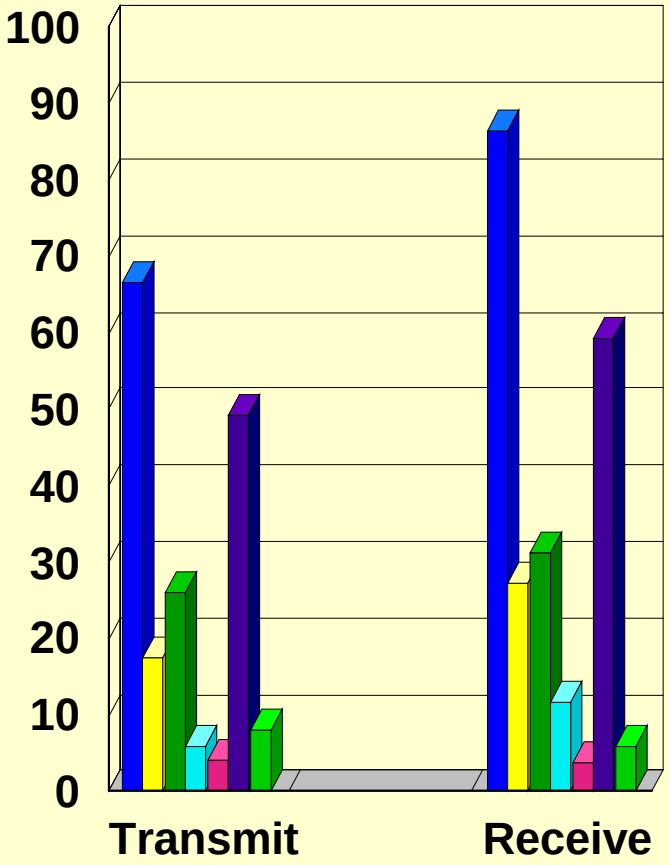
- **New iSCSI Storage Controllers have Admin similar to NAS**
  - Permit Admin to create LUs of default sizes
    - Similar to Admin operations on NAS Servers
  - Permit Admin to assign Authentication strings to Client Systems
    - Similar to NAS
  - Permit Admin to assign LUs to Systems
    - NAS assigns Systems to Filers (NAS Servers)
  - Automatically assign LUs to File Systems and Format the LU
    - Looks to the OS like a local disk implicitly mounted at boot or on request
    - NAS mounted on request, can also be mounted by script at boot
  - Many have very easy to use GUIs and CLIs
    - Each one claims to be better than the other
    - Both NAS and iSCSI can be simple and easy to use
    - Both NAS and iSCSI can add complexity dealing with advanced features
      - Security features
      - RAID types, and LU sizes
      - Snapshots,…

iSCSI Performance

- **At least as good as NAS**
  - probably better cost performance
- With software TCP/IP, Block I/O Server has 3-4 times less overhead then software TCP/IP based NAS
- With TCP/IP HW Acceleration (both iSCSI & NAS)
  - iSCSI Block I/O is 1/12 -1/16 the overhead of NAS

NAS vrs iSCSI
(on the Storage Controller)

Percent CPU Overhead



With unmodified TCP/IP, iSCSI is 1/3 the overhead of NFS

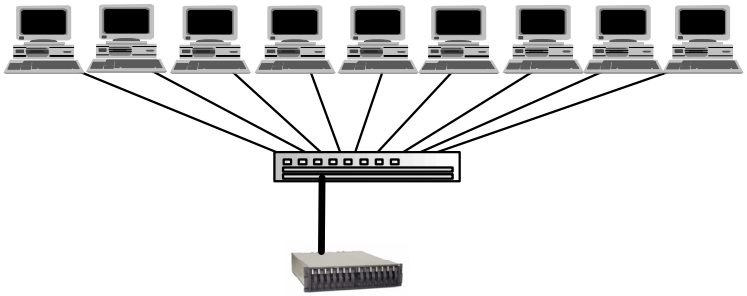
With all TCP/IP copy overhead offloaded (0 Copy) iSCSI was 1/12 the overhead of NFS

- NFS
- SCSI over GE-TCP/IP
- % of NFS cpu used by SCSI over GE-TCP/IP
- SCSI over GE-TCP/IP 1 copy
- SCSI over GE-TCP/IP 0 copy
- NFS (with 0 TCP/IP data copies) est.
- % of NFS cpu (with 0 TCP/IP data copies) used by SCSI over GE-TCP/IP (0 copies)

Goal: Use NAS for Sharing Files, and iSCSI for everything else

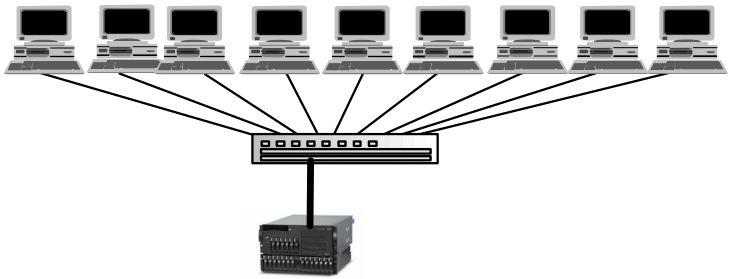
Spreading vrs Centralizing the File System Overhead

Block I/O (including iSCSI) spread the File System overhead across all the Clients



Block I/O (including iSCSI) Storage Controllers just store the I/O blocks where the Client File System requests (perhaps with Virtualizing LUN Mapping)

NAS Clients move the File System overhead to the NAS server



NAS Servers centralizes the File System functions (and overhead) for all its clients into the NAS Server

Plus the NAS Server still must map the resultant Blocks onto the Storage (perhaps with Virtualizing LUN Mapping)

The non TCP/IP Server side overhead can be 12- 16 times higher in NAS Servers than Block I/O (iSCSI) Storage Controllers

Therefore use NAS for File Sharing and iSCSI for other IP Storage Requirements

Decision Points (in General Environments)

Choice of NAS vrs iSCSI depends on file sharing needs

- Do I need to share files?
  - If **no** ➔ Use a Local File System with iSCSI Storage
  - If **yes** ➔ Use NAS for those files
- Do I need to share **all** my files?
  - If **yes** ➔ Use only NAS
  - If **no** ➔ also use a Local File System with iSCSI Storage or consider a **Dual Dialect storage controller** (both iSCSI and NAS)

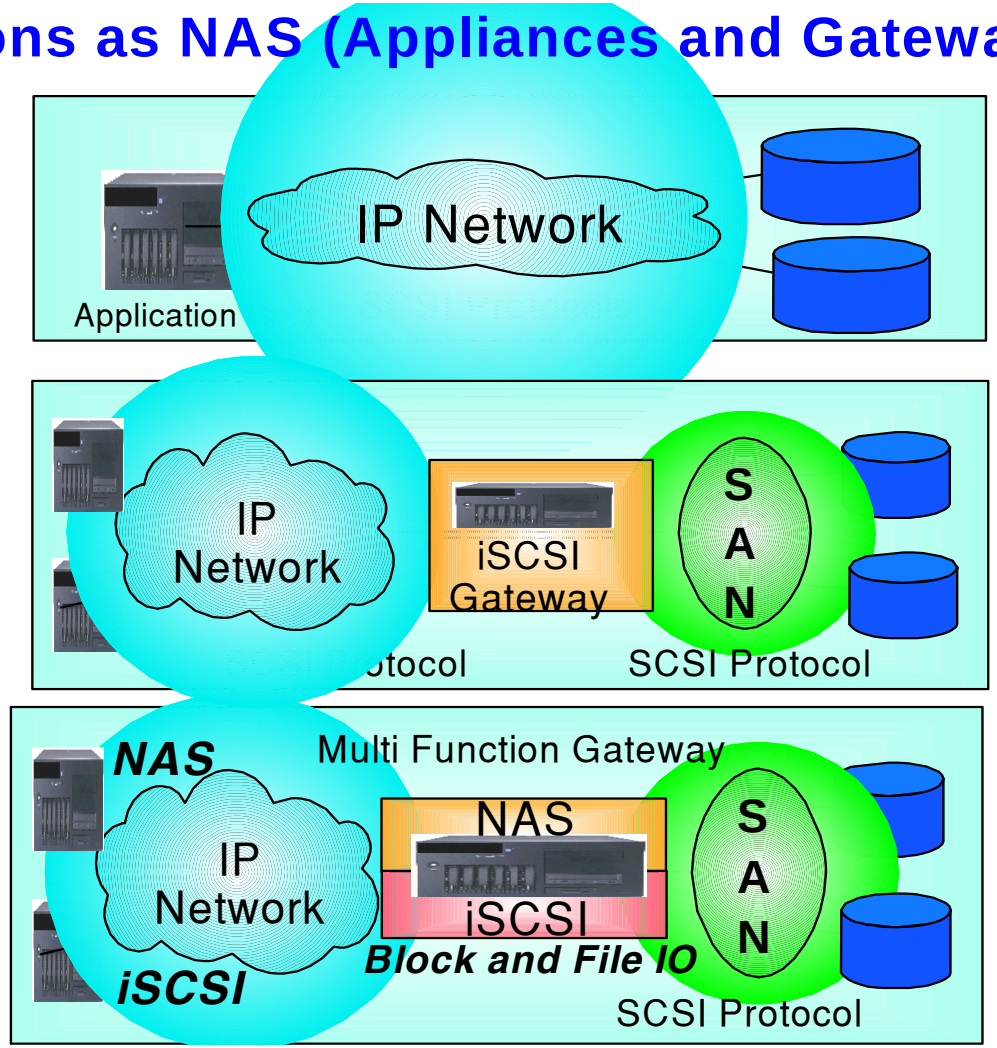
iSCSI Deployments

Same HW Configurations as NAS (Appliances and Gateways)

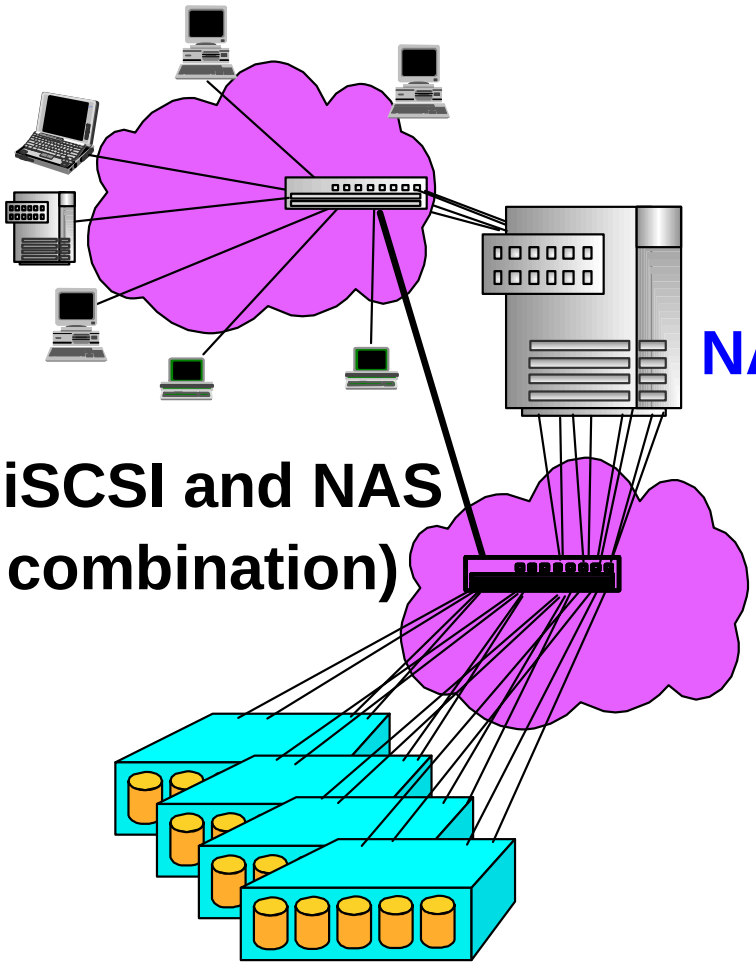
Independent iSCSI Deployment

Extending the SAN

In Combination with NAS (Dual Dialect)



Peaceful Co-existence
iSAN & NAS

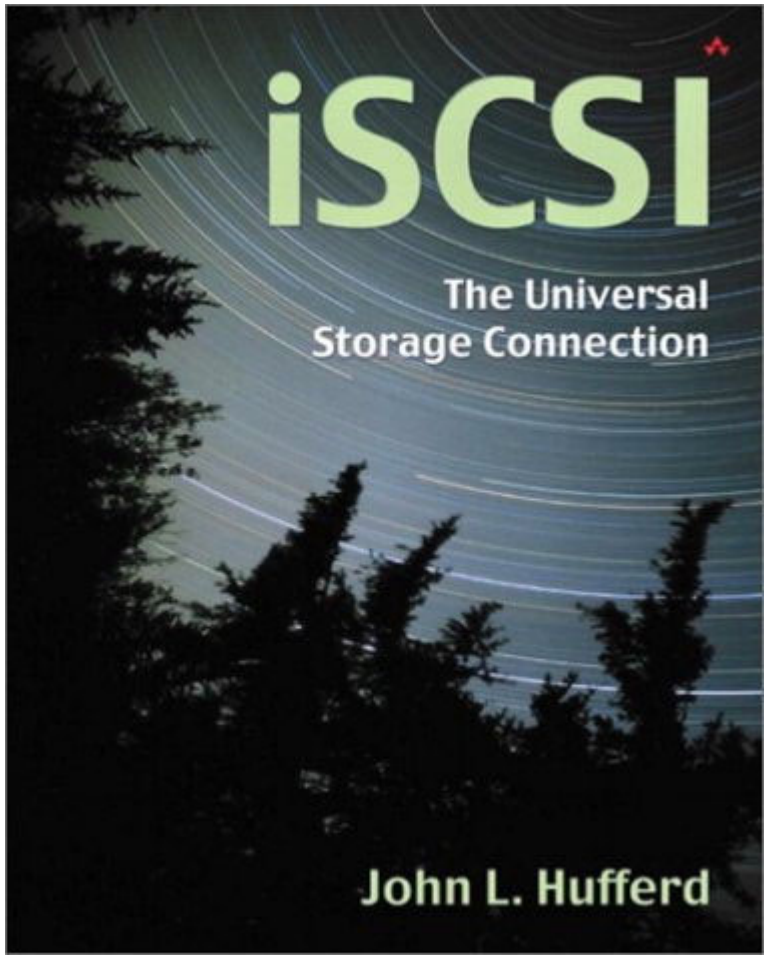


NAS iSCSI-Gateway

Supports both iSCSI and NAS
(a Dual Dialect combination)

iSCSI Raid Ctlrs

iSCSI Reference Book



Published by Addison-Wesley

Available in Book Stores

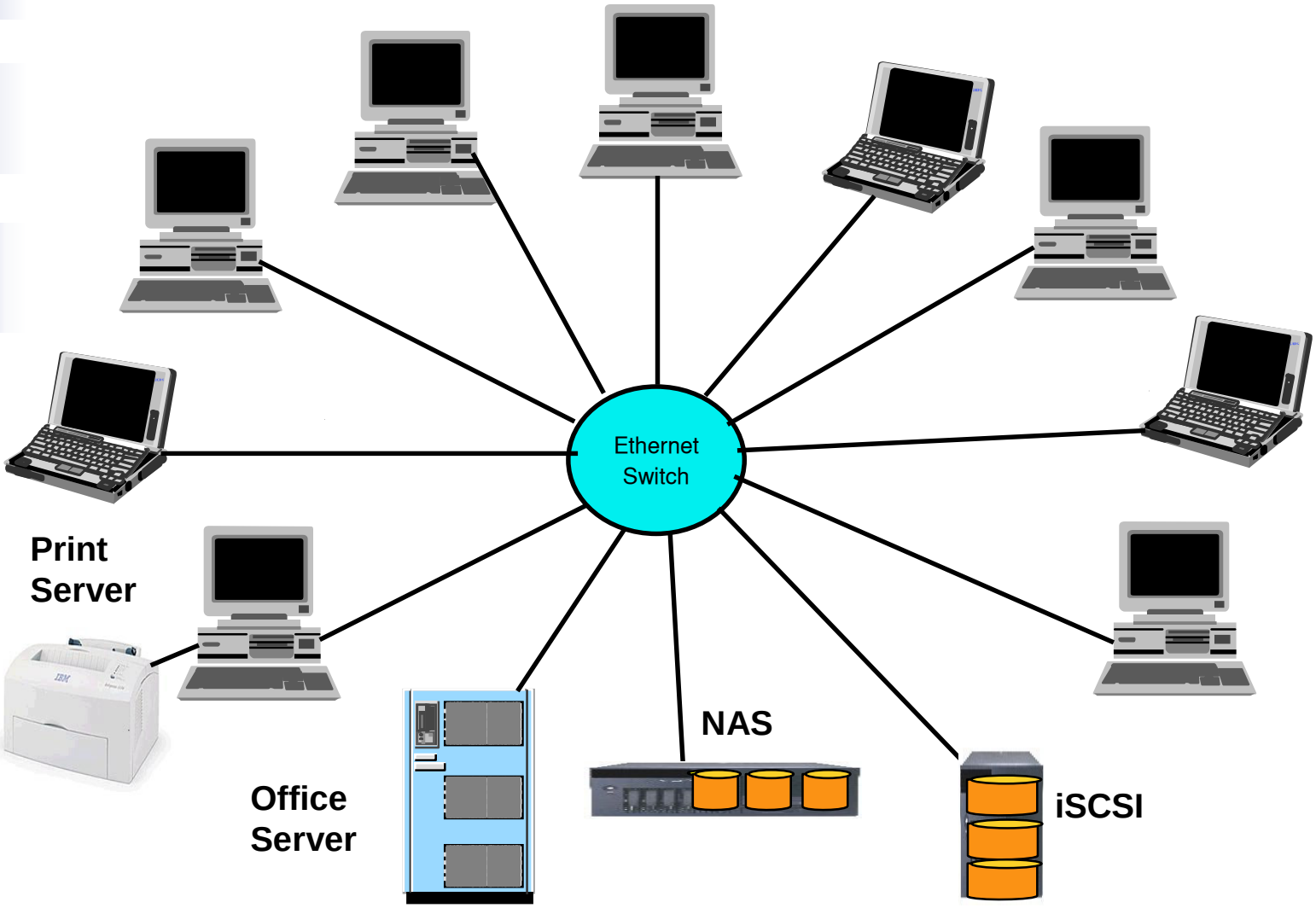
and Amazon.com

Volume purchases available

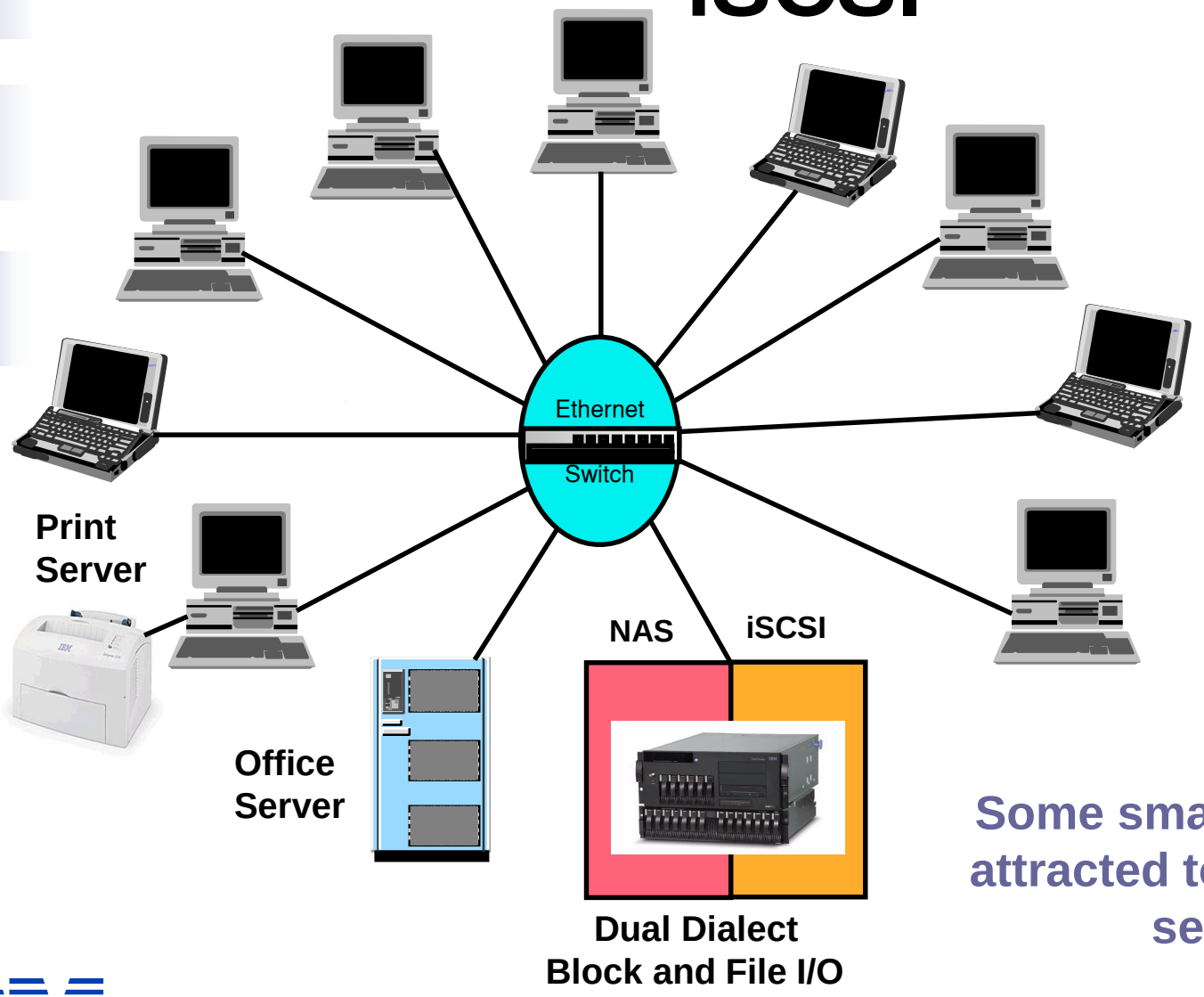
Appropriate for Marketing, Sales, Engineering, and IT personnel

Backup Slides

Small Office Interconnect

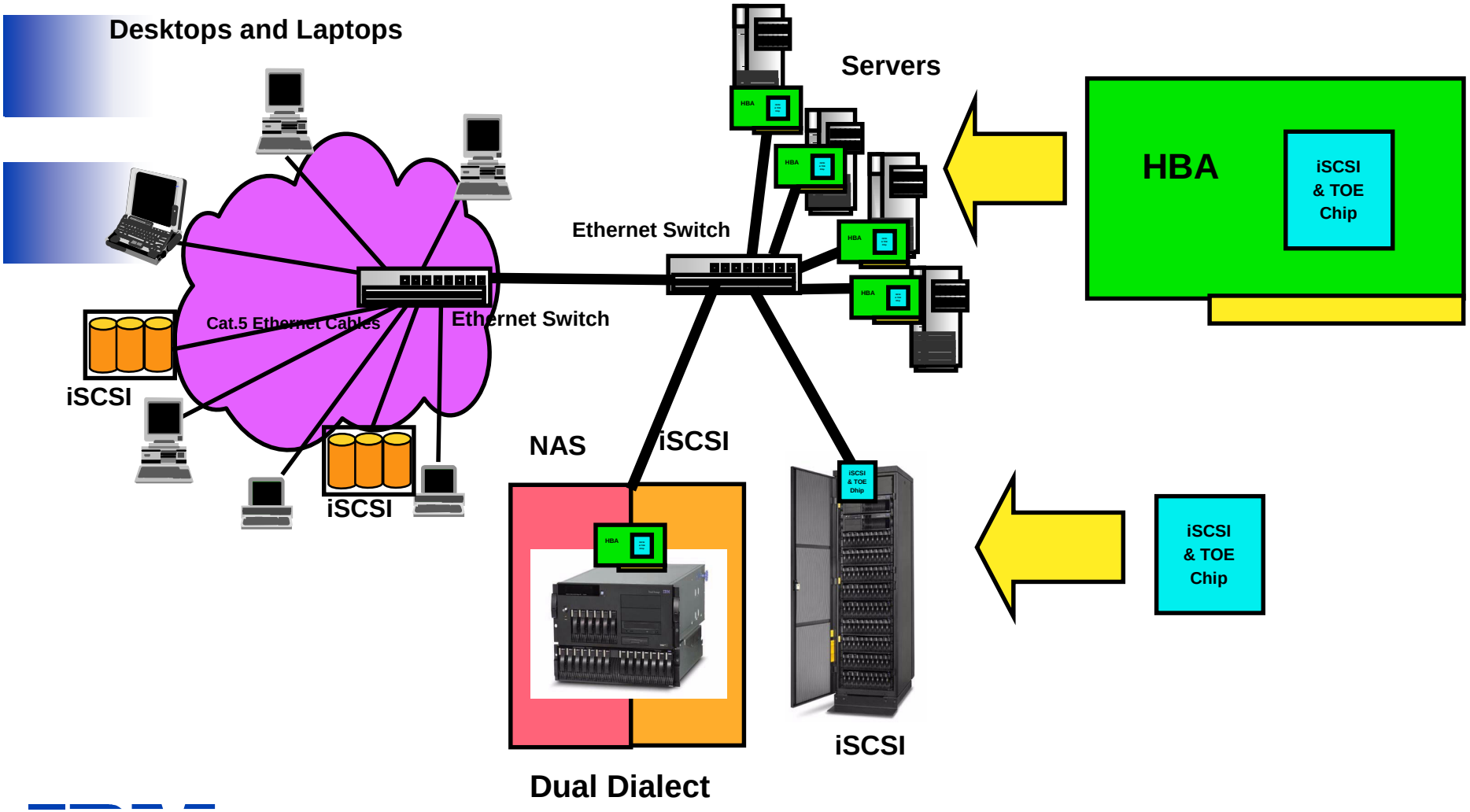


IP Storage Combo -- NAS & iSCSI



Some small Offices are attracted to dual dialect servers

Midrange Environments



RETURN TO INDEX