IBM GLOBAL SERVICES

Session T07

iSCSI Tutorial

John L. Hufferd

IBM @server xSeries Technical Conference

Aug. 9 - 13, 2004

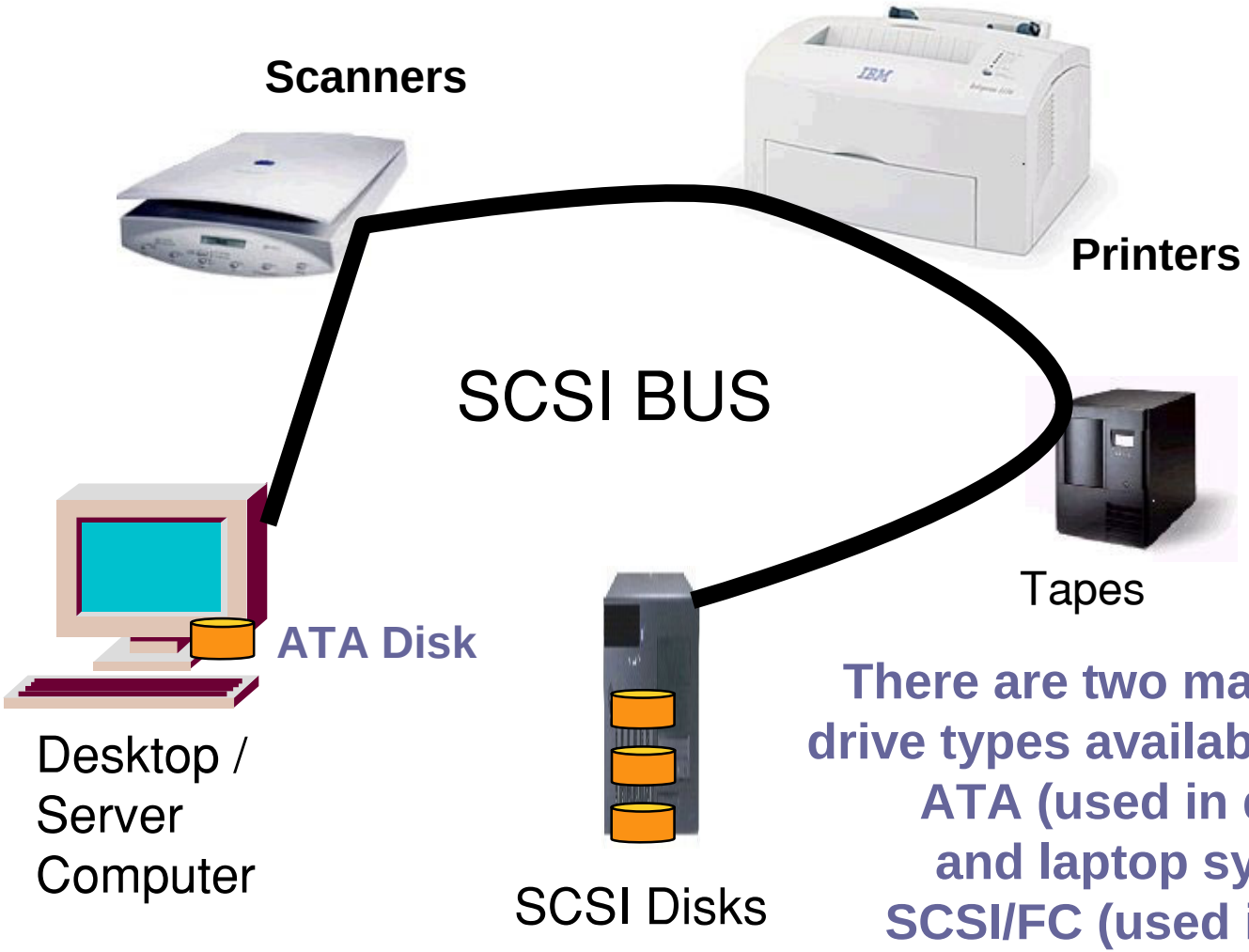
Chicago, IL

RETURN TO INDEX

Terms

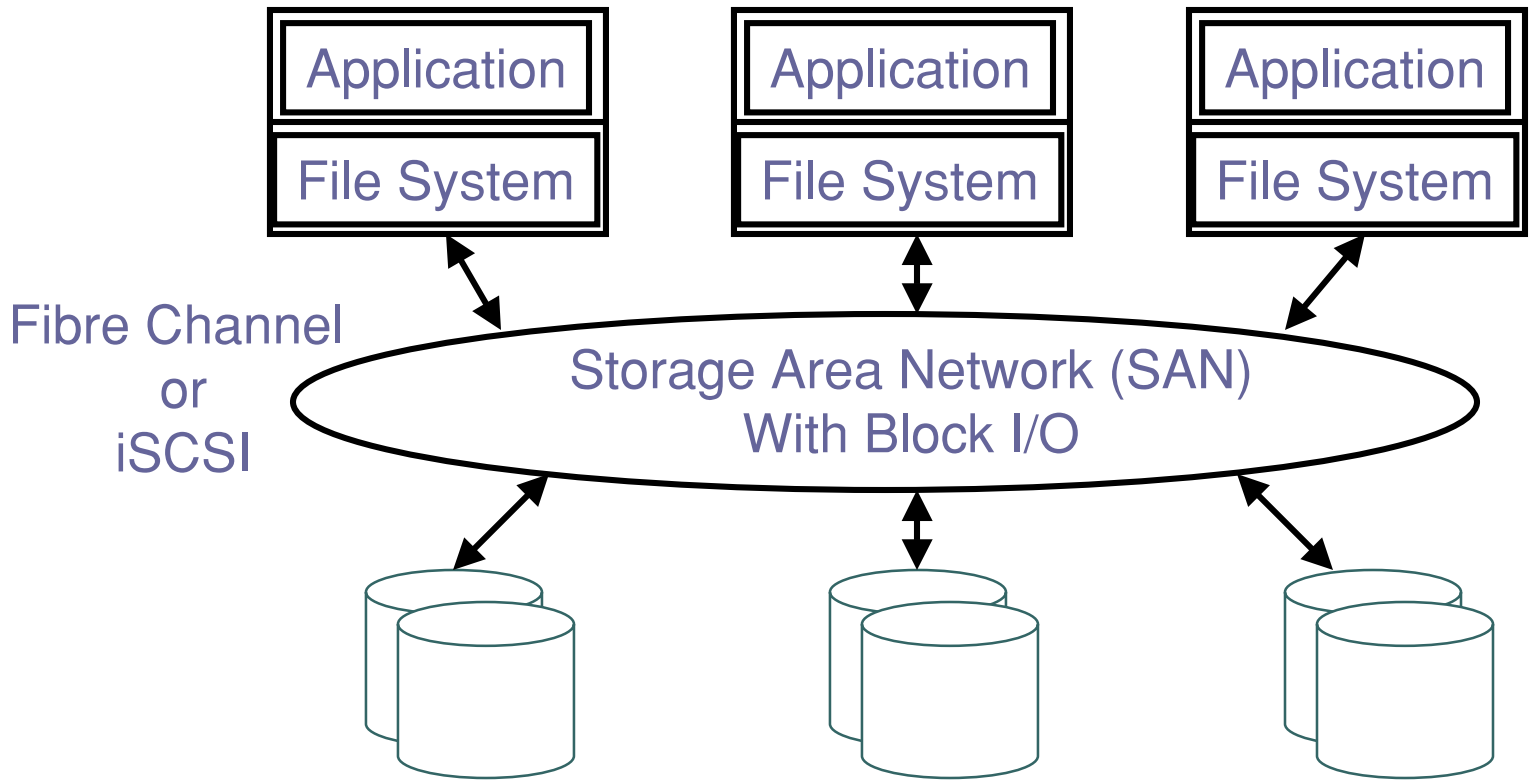
- iSCSI - Internet SCSI
- NAS - Network Attached Storage
- HBA - Host Bus Adapter
- TOE - TCP/IP Offload Engine
- FC - Fibre Channel
- SAN - Storage Area Network
- iSAN - iSCSI Storage Area Network
- PDU - Protocol Data Unit
- WWN - World Wide Name

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)

Systems with SCSI over Networks



Both Fibre Channel and iSCSI can makeup a SAN

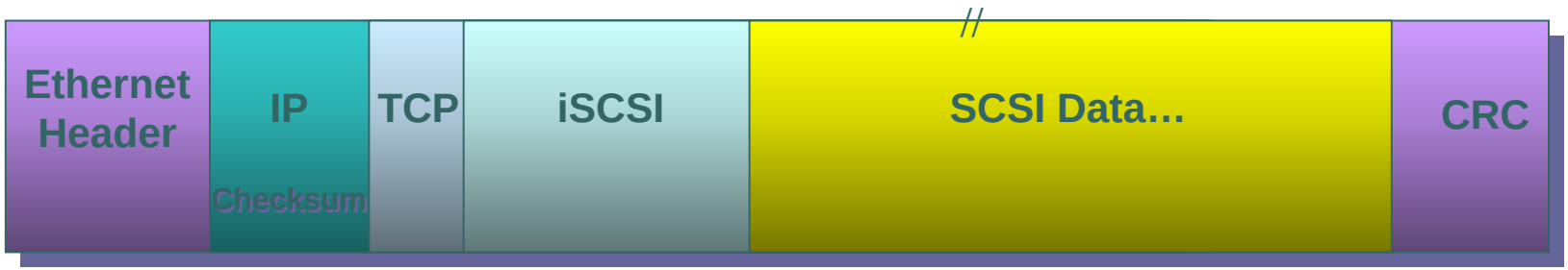
Replaces shared bus with switched fabric

iSCSI is:

- **Internet SCSI**
  - **i**nternet **S**mall **C**omputer **S**ystem **I**nterconnect
  - SCSI over TCP/IP
    - On Ethernet LANs
      - Copper
      - Optical
    - On ATM WANs
    - On SONET WANs
    - Etc.

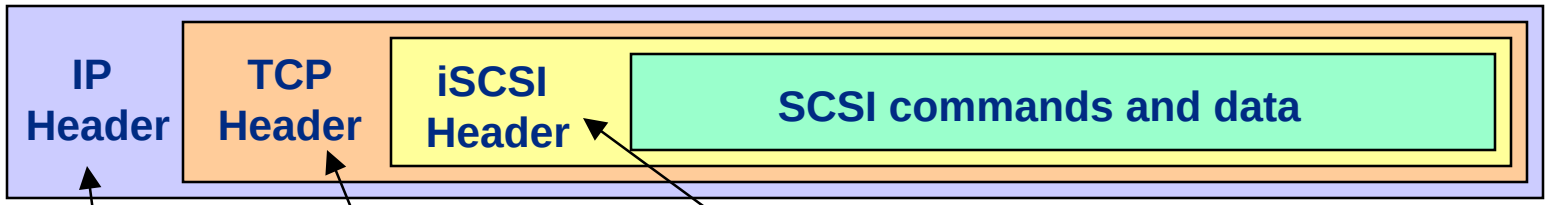
iSCSI Protocol Structure

- iSCSI is a SCSI transport protocol for mapping of block-oriented storage data over TCP/IP networks
- The iSCSI protocol enables universal access to storage devices and Storage Area Networks (SANs) over standard TCP/IP networks



iSCSI Protocol Structure – cont.

- **iSCSI (Internet SCSI) specifies a way to "encapsulate" SCSI commands in a TCP/IP network connection:**

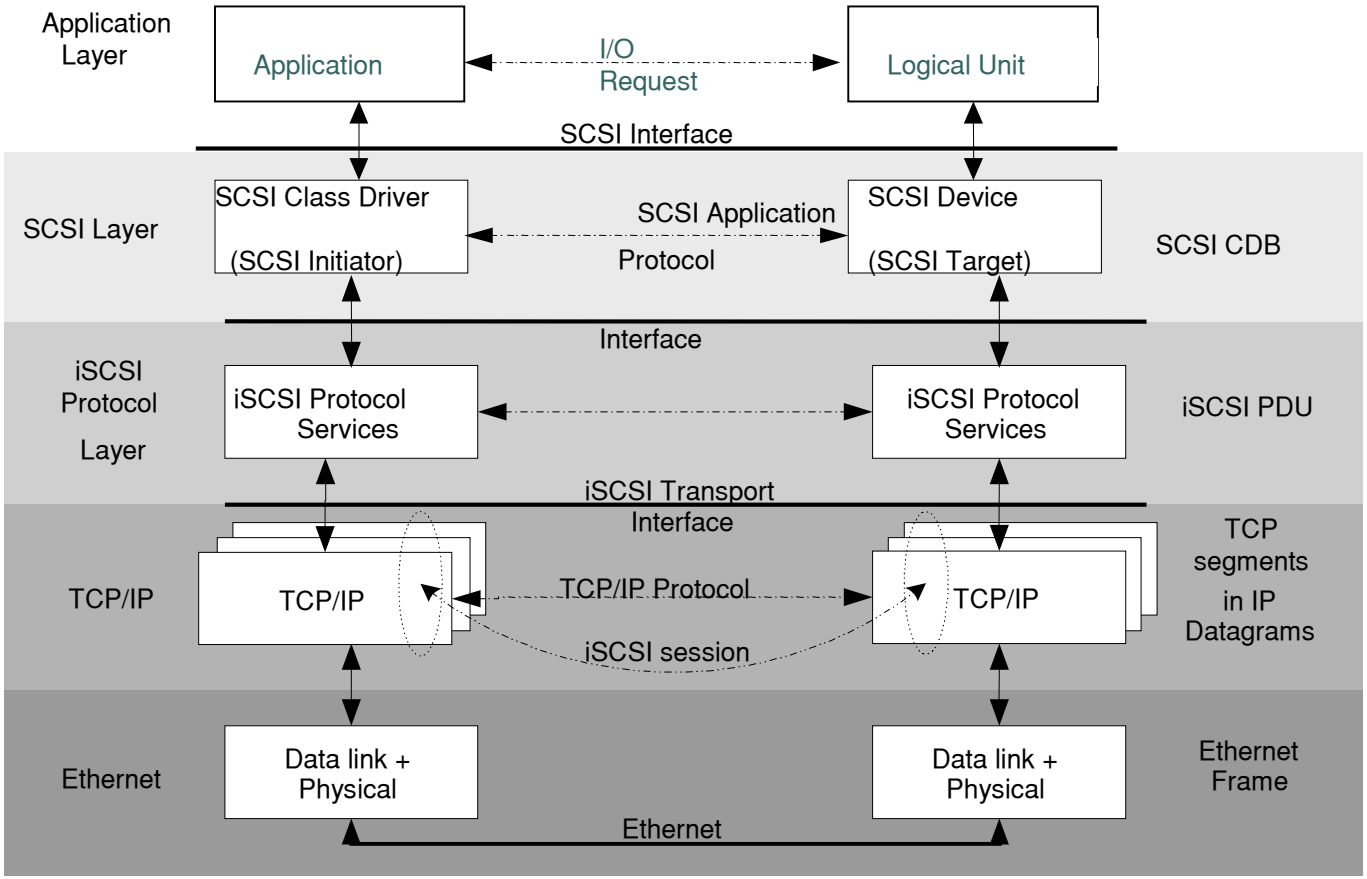


Explains how to extract SCSI commands and data

Provides information necessary to guarantee delivery

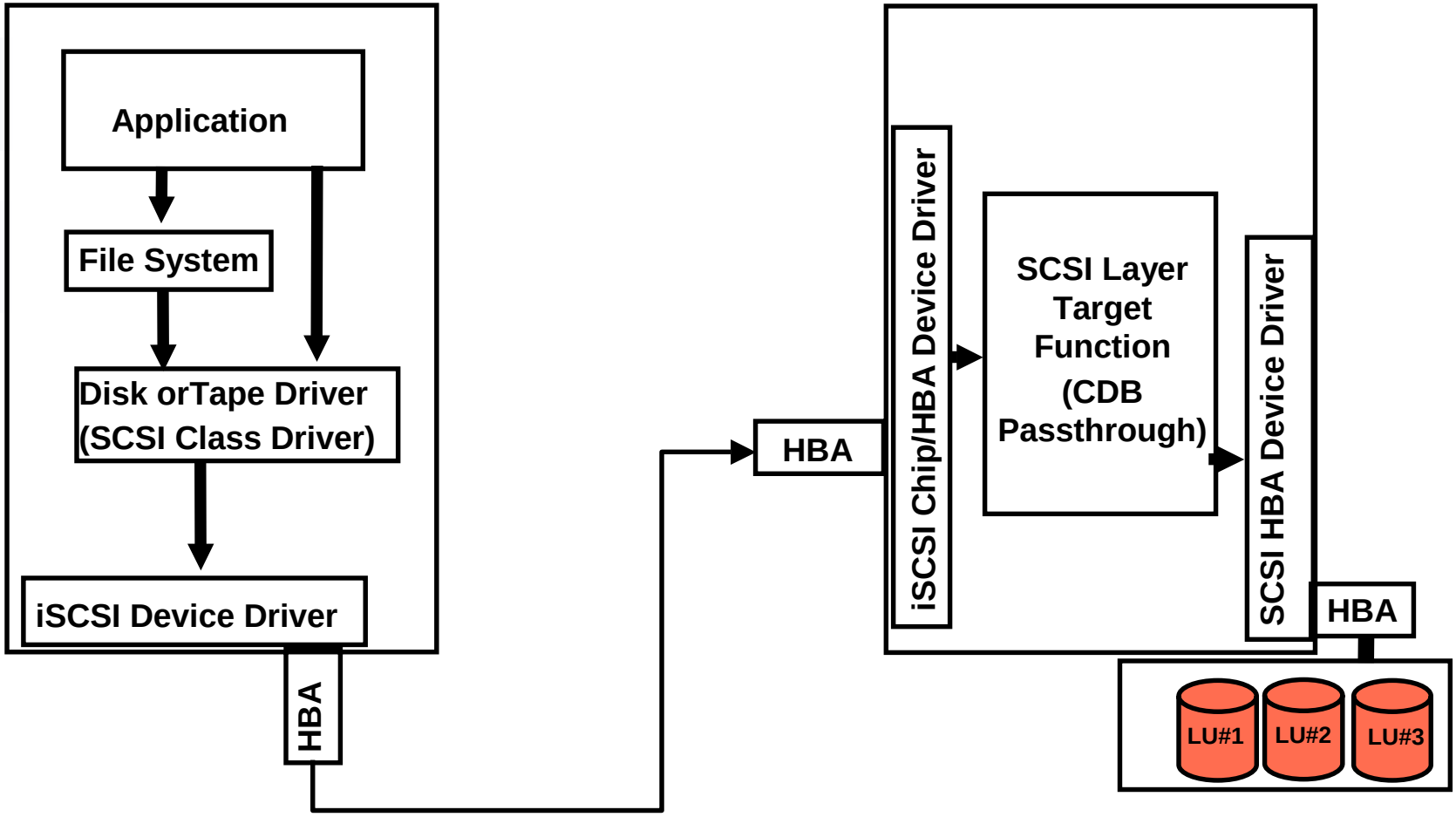
Contain "routing" information So that the message can find its Way through the network

iSCSI - Layered Model

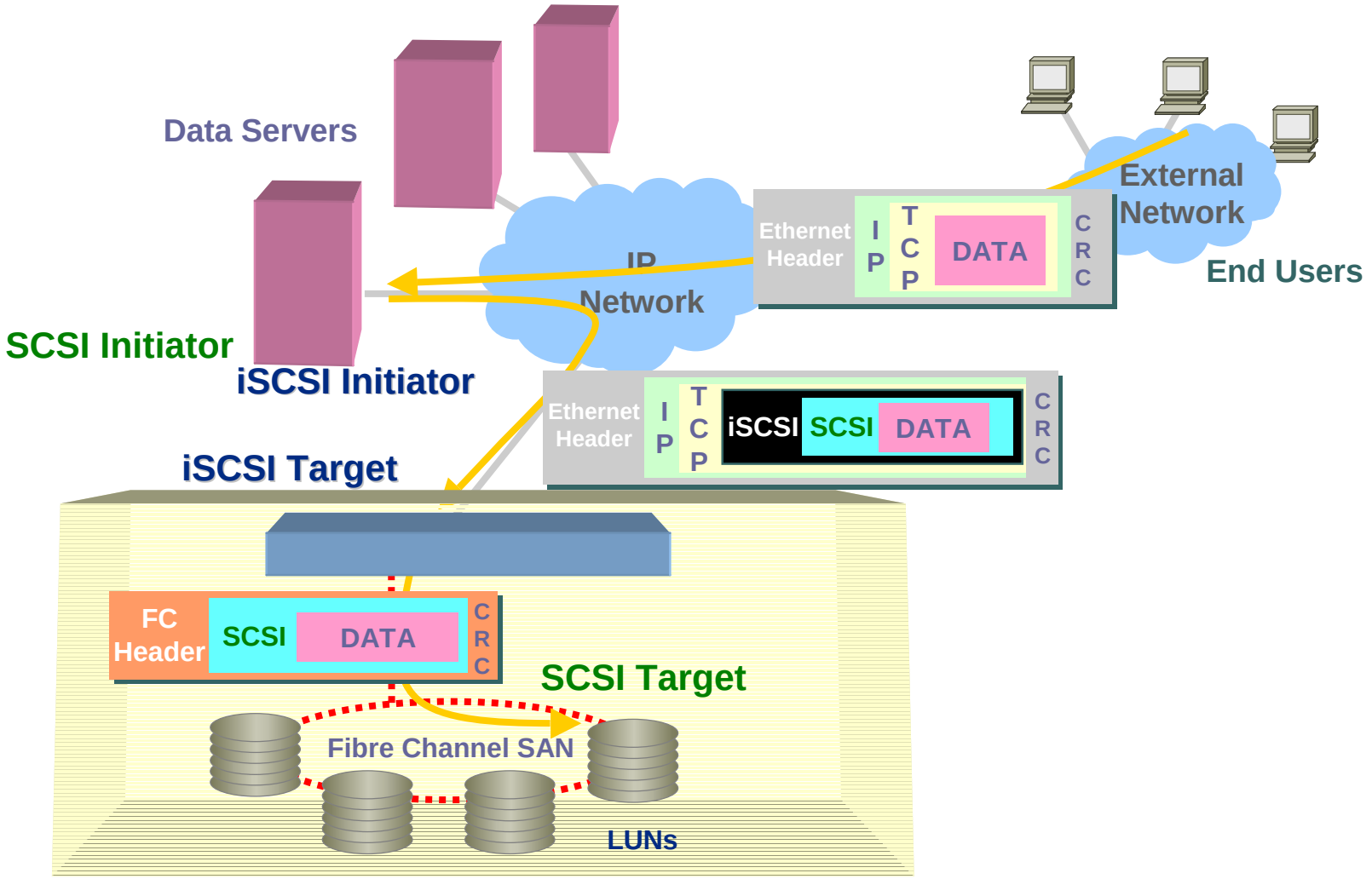


Transparently encapsulates SCSI Command Descriptor Blocks (CDBs)

Application to LU Command Flow



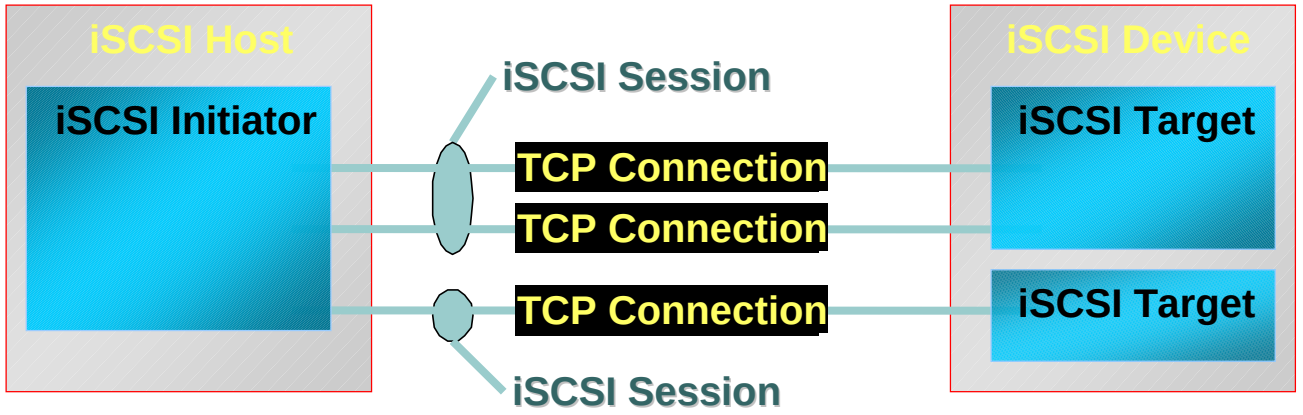
Application Data Flow



iSCSI Structure

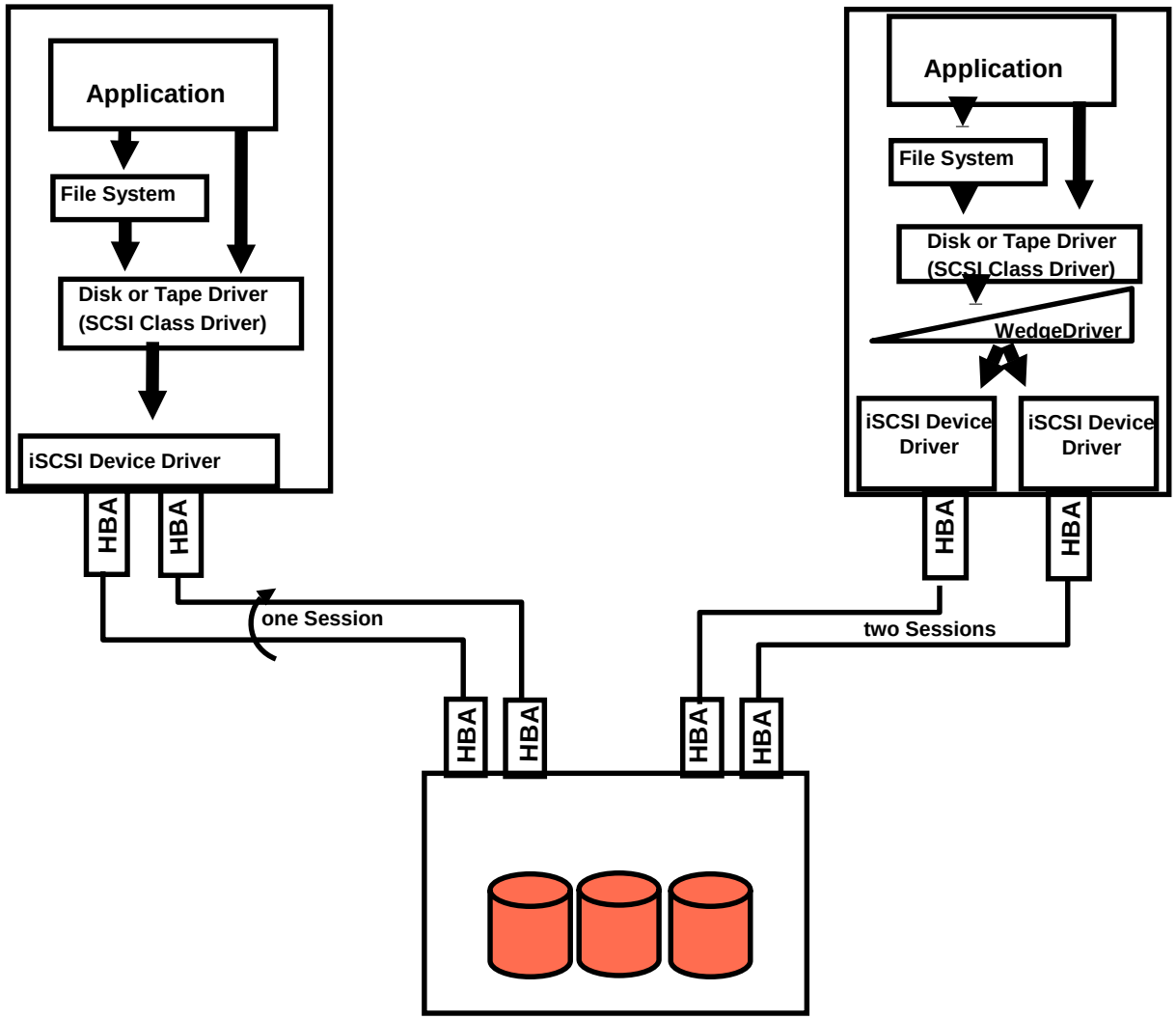
- **iSCSI has the concept of a Session**
  - A session maybe made up of one or more TCP/IP connections
  - The Session can be thought of as a SCSI Port
  - The Session is started after Login is complete

iSCSI Sessions



- Session between initiator and target
  - One or more TCP connections per session
  - Login phase begins each connection
- Deliver SCSI commands in order
- Recover from lost connections

Multiple Connections Between Hosts and Storage Controllers



iSCSI Integrity

- **iSCSI adds Cyclic Redundancy Check (CRC)**
  - CRC-32C - A 32 bit check word algorithm
  - End to End Checking
  - In addition to TCP/IP Checksums
  - In addition to Ethernet Link level CRCs
- **CRC "check word" is called a "Digest"**
- **iSCSI Digests for iSCSI Headers and Data**
  - Header Digest is optional to use (MUST implement)
    - Insures correct operation and data placement
  - Data Digest is optional to use (MUST implement)
    - Insures data is unmodified through-out network path

iSCSI Message Types

Called Protocol Data Units (PDUs)

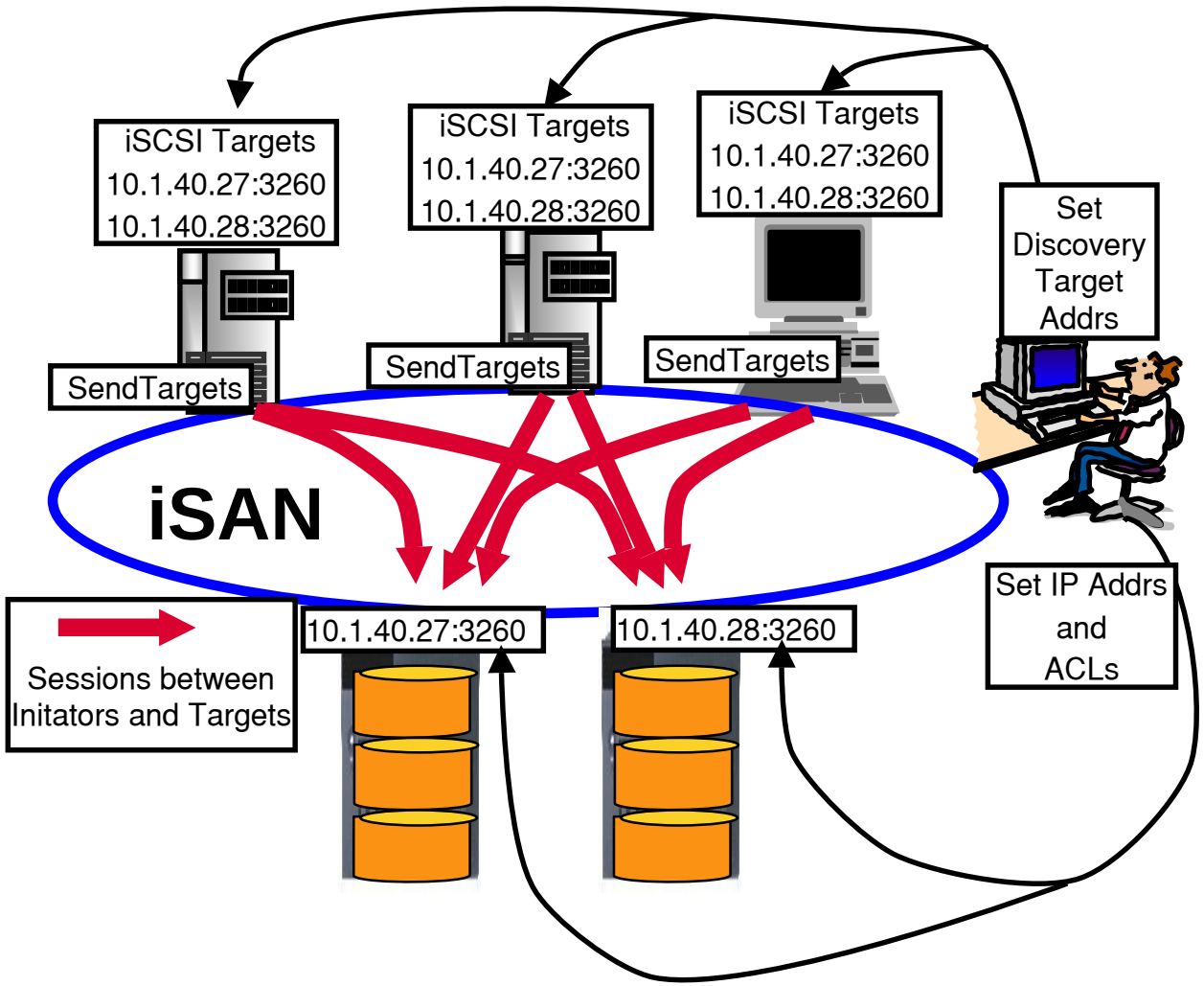
- Initiator to Target
  - NOP-out
  - SCSI Command
    - Encapsulates a SCSI CDB
  - SCSI Task Mgmt Cmd
  - Login Command
  - Text Command
    - Including SendTargets
      - Used in iSCSI Discovery
  - SCSI data
    - Output Data for Writes
  - Logout Command

- Target to Initiator
  - NOP-in
  - SCSI Response
    - Can contain status
  - SCSI Task Mgmt Rsp
  - Login Response
  - Text Response
  - SCSI data
    - Input Data from Reads
  - Logout Response
  - Ready to Transfer
    - R2T
  - Async Event

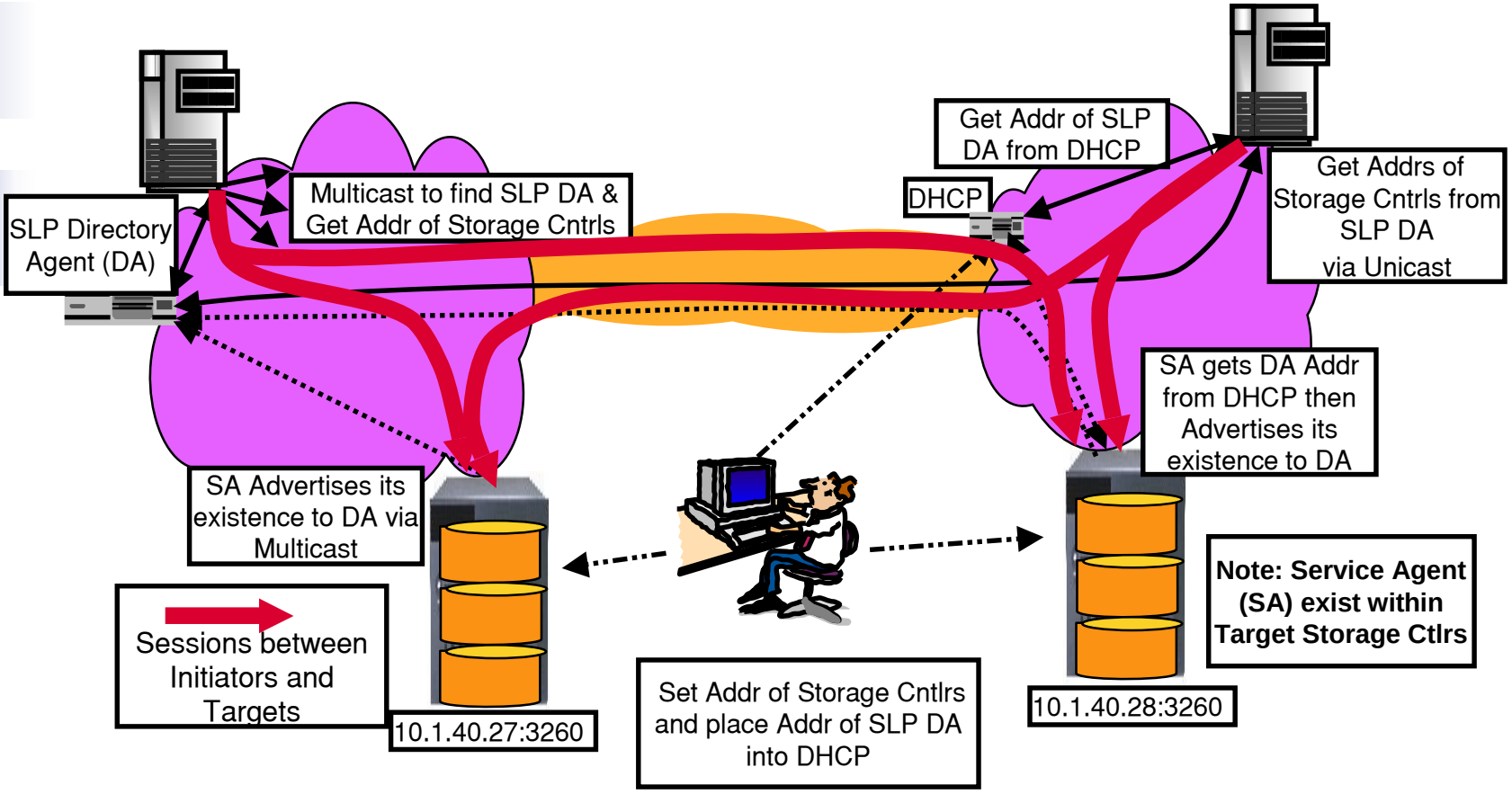
iSCSI Error Handling

- **ErrorRecoveryLevel = 0**
  - When iSCSI detects errors it will bring down the TCP connection and restart it
  - iSCSI will let the SCSI layer retry the operation
- **ErrorRecoveryLevel = 1**
  - Detected errors (Header or Data) causes PDUs to be discarded
  - iSCSI will retransmit discarded commands
  - iSCSI will retransmit discarded data
- **ErrorRecoveryLevel = 2**
  - Caused by loss of the TCP/IP connection
  - Connection & Allegiance reestablishment
  - Uses ErrorRecoveryLevel 1 to recover lost PDUs

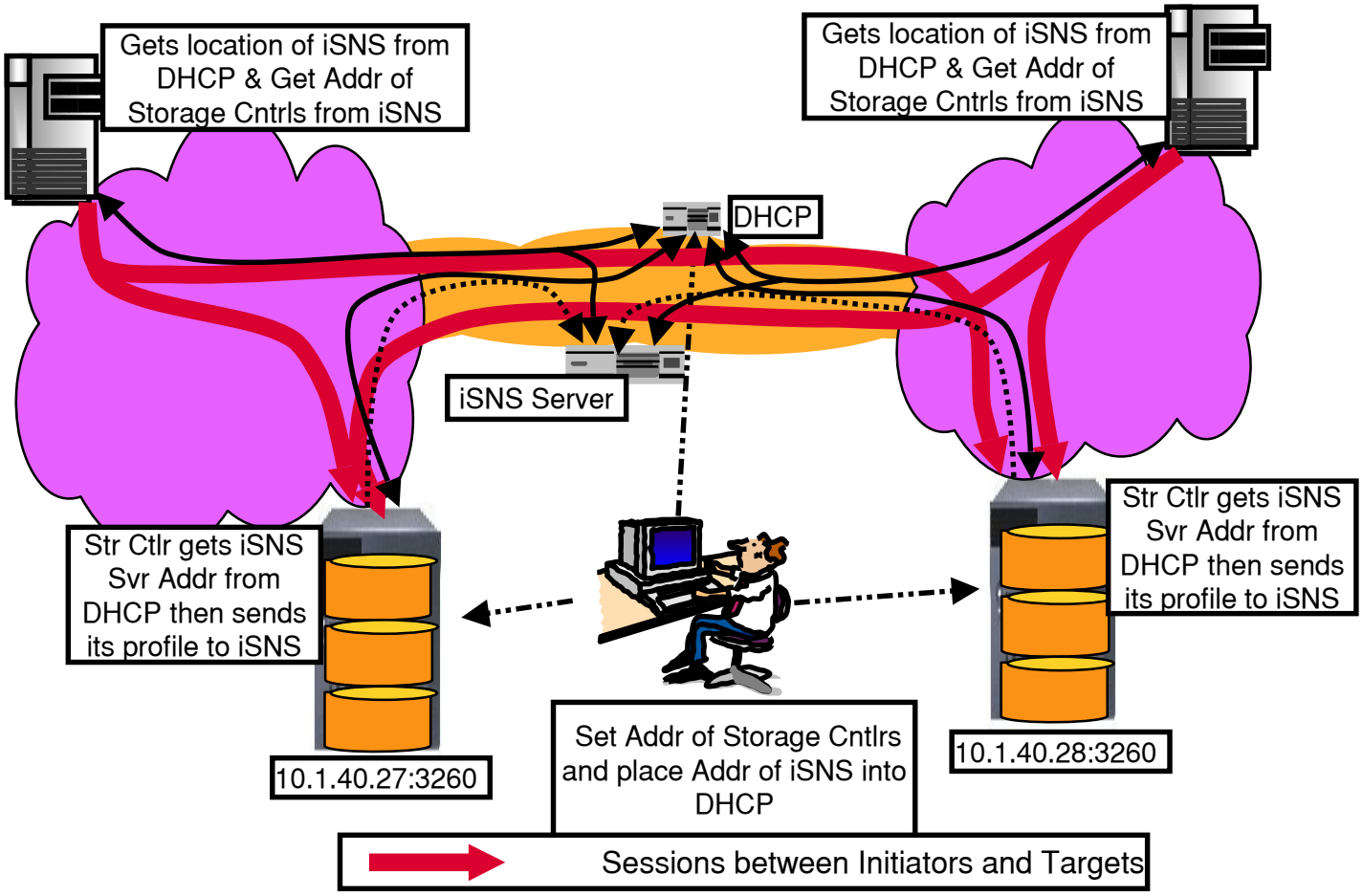
Discovery via SendTargets



Discovery via SLP



Discovery via iSNS



iSCSI Relocation

- **After attempting to Login at specified location:**
  - The specified Target may signal a relocation when Normal Login attempted
    - Temporary relocation
    - Permeate relocation
- **Relocation used for:**
  - Corrections between Discovery DB updates
  - Admin or automatic Hardware disablement
    - for Service
    - Because of HW problems
  - For load balancing

iSCSI Boot

- **Static configuration information for Boot**
  - Admin sets authorized iSCSI Target Node Name and iSCSI Address, Optional LUN
    - Default LUN is 0
- **Dynamic configuration via use of DHCP, SLP, iSNS**
  - DHCP can be used by Host to get an IP address
  - DHCP can hold the iSCSI Boot Service Option (Admin Set)
    - May contain all that is needed to reach the Boot device
    - May only contain iSCSI Target Node Name, then use SLP/iSNS to resolve to iSCSI address
  - SLP, or iSNS can also be used to find the Boot location
- **The Boot load process**
  - The Admin. or DHCP, SLP or iSNS can enable the access
  - BootP/PXE is also possible as part of a SW two phase process
  - HW HBA can act as a normal SCSI HBA for system BIOS use

iSCSI Security Considerations

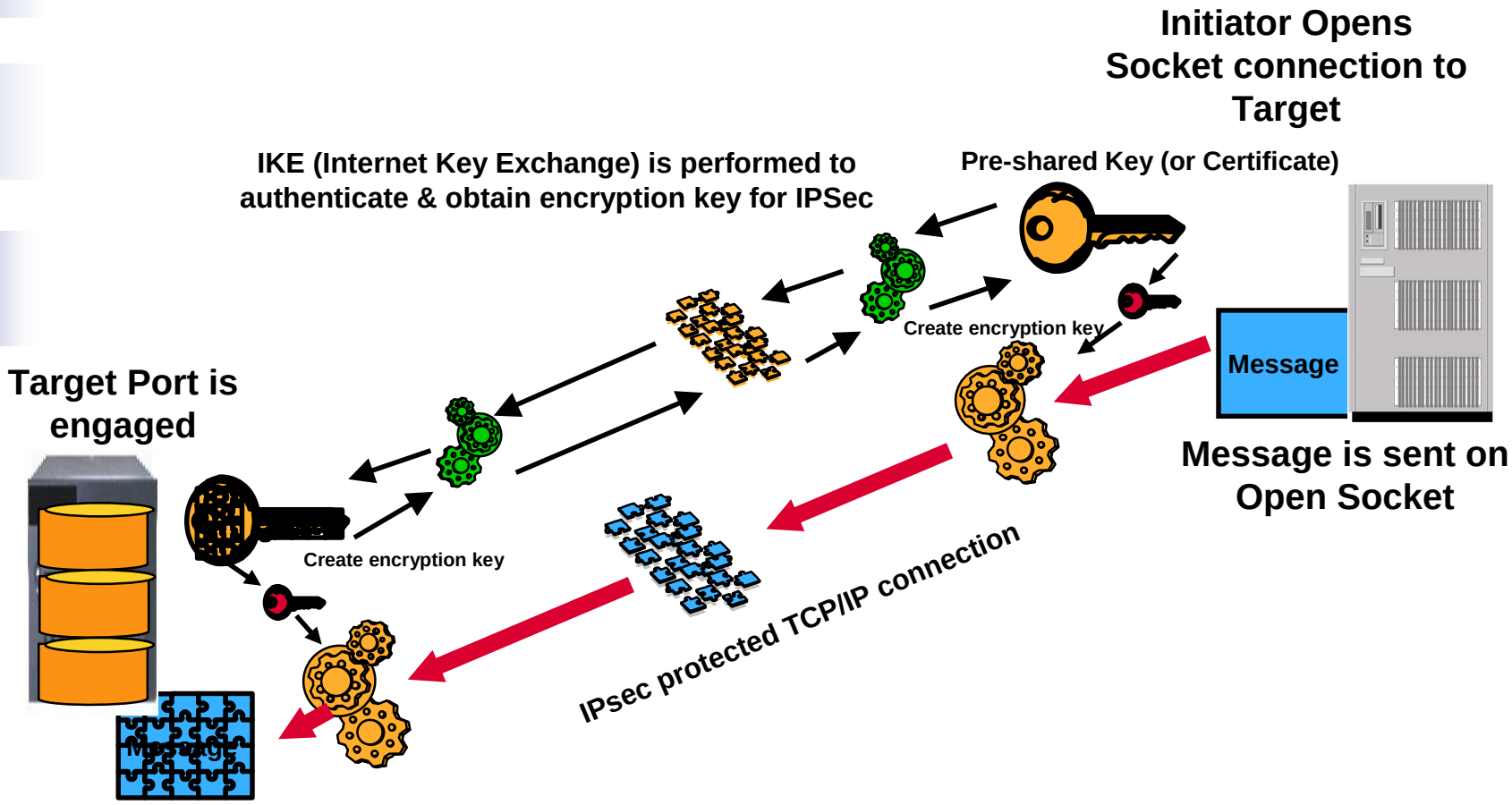
- **Authentication is iSCSI way to determine trustworthiness via**

  Connection Security

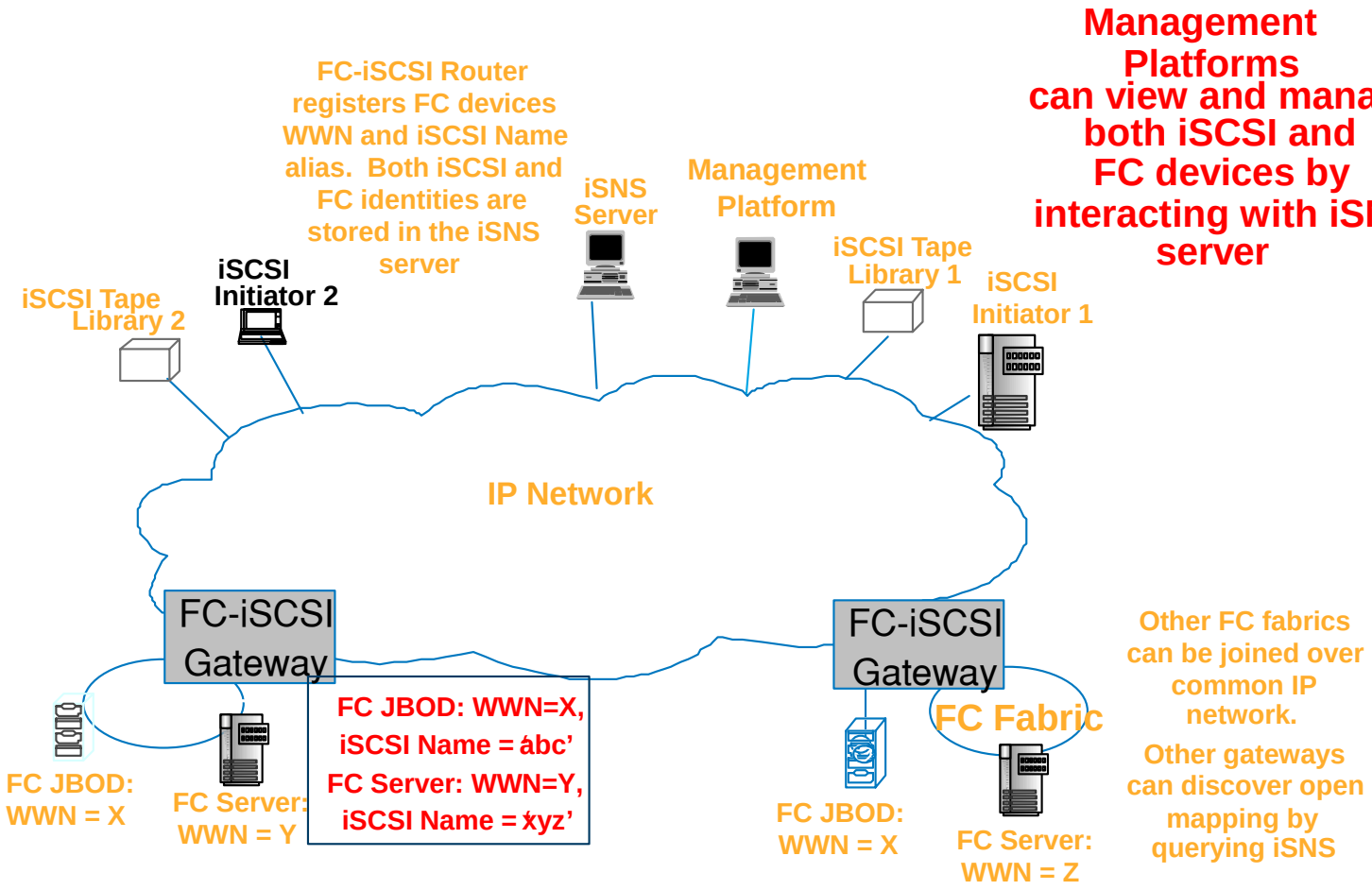
  - CHAP -- Challenge Handshake Authentication Protocol
  - SRP -- Secure Remote Password
  - Kerberos -- A Third Party Authentication protocol
  - SPKM-1,SPKM-2 -- Simple Public Key Mechanism

- **Connection Security may be used in addition to IPsec's:**
  - Packet Authentication
    - Origin assurance
    - Anti-Reply protection
  - Privacy
    - Encryption

iSCSI with IPsec

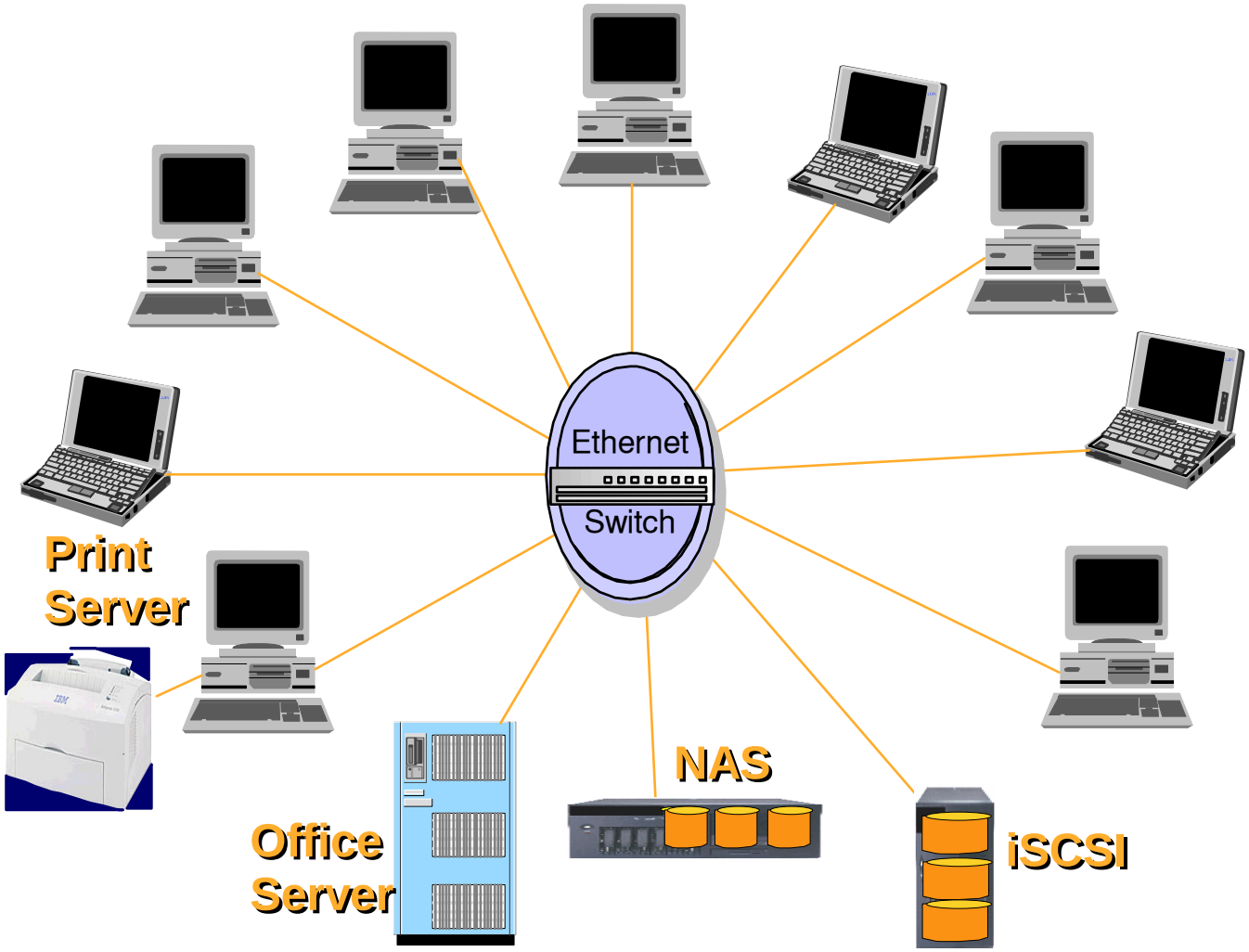


Combining of FC and iSCSI

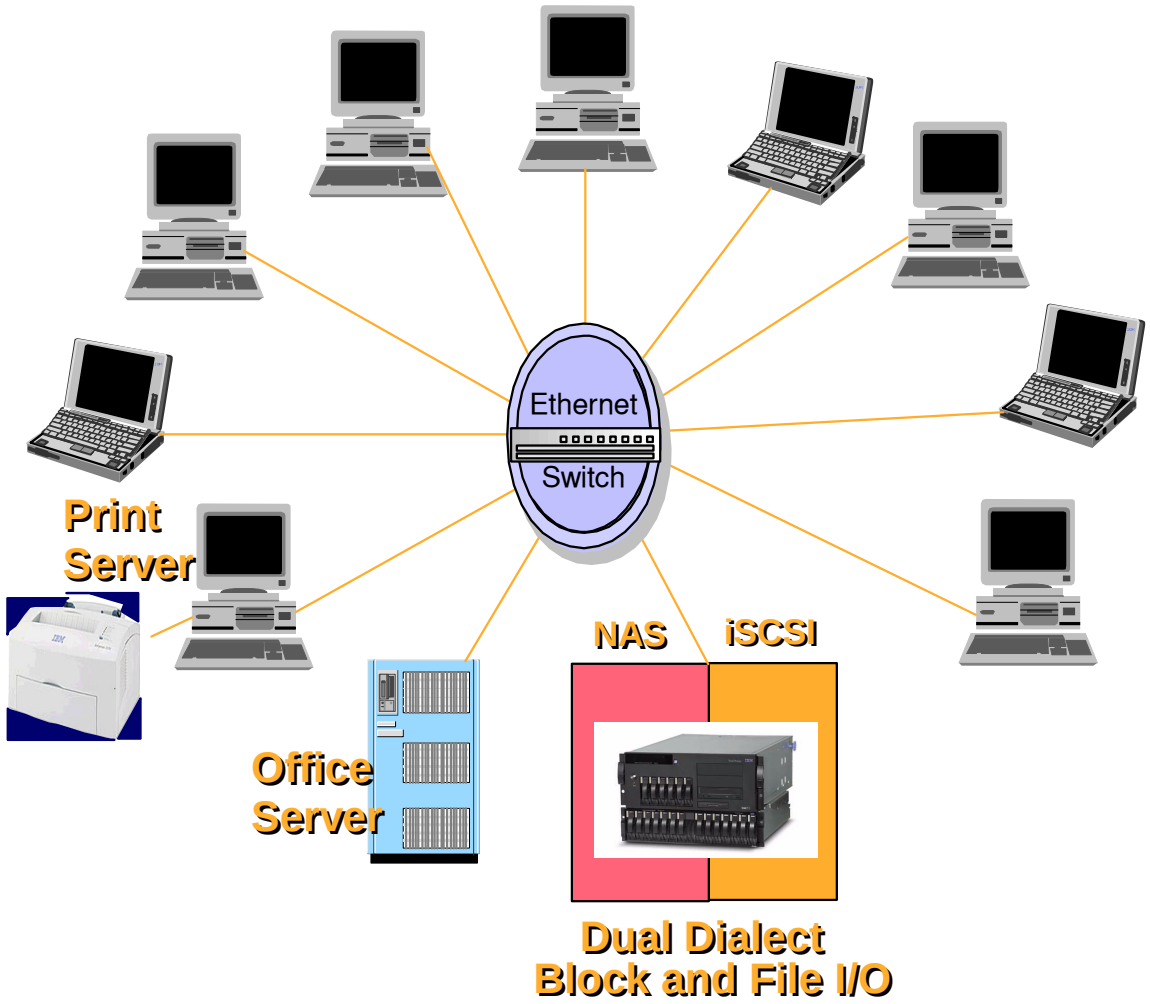


Now lets look at the various environments where iSCSI is appropriate

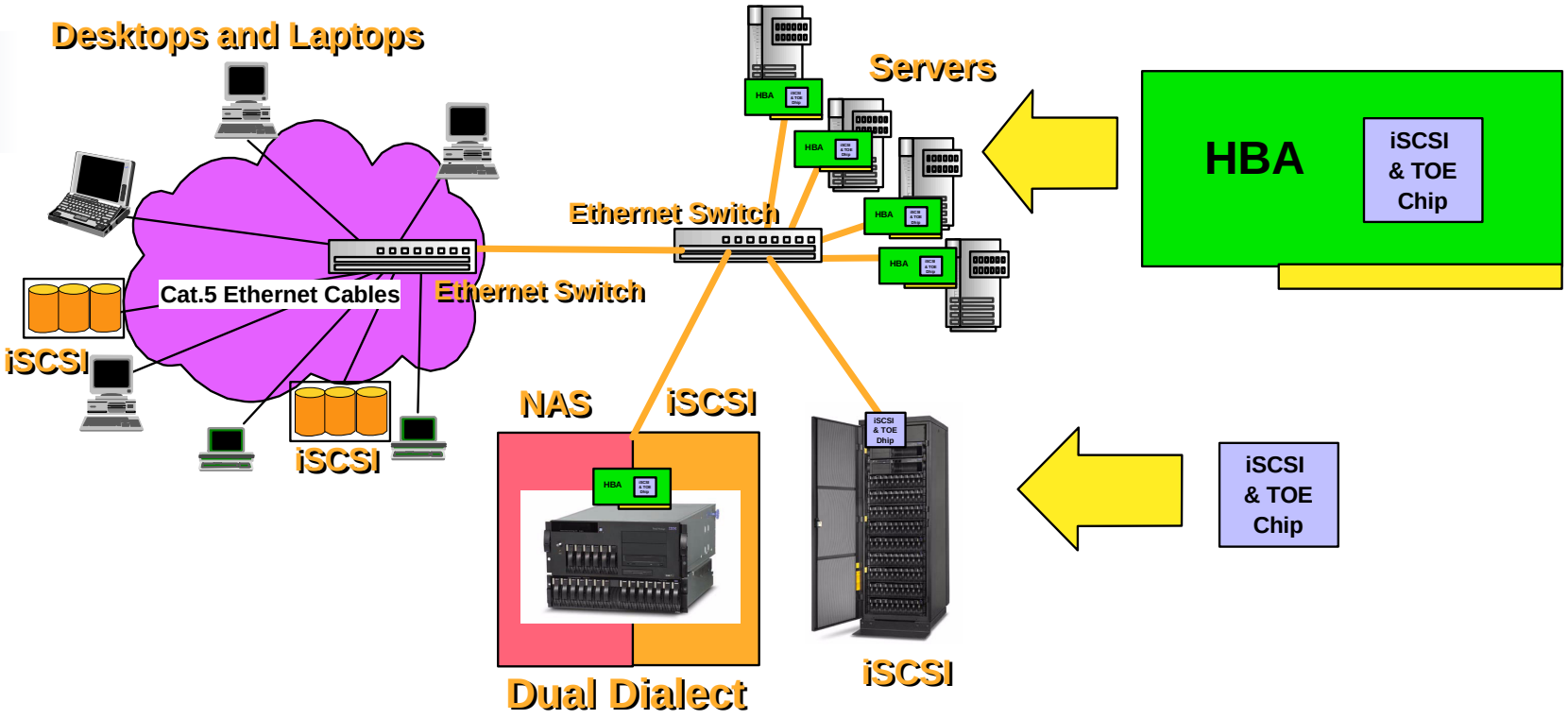
Small Office Interconnect



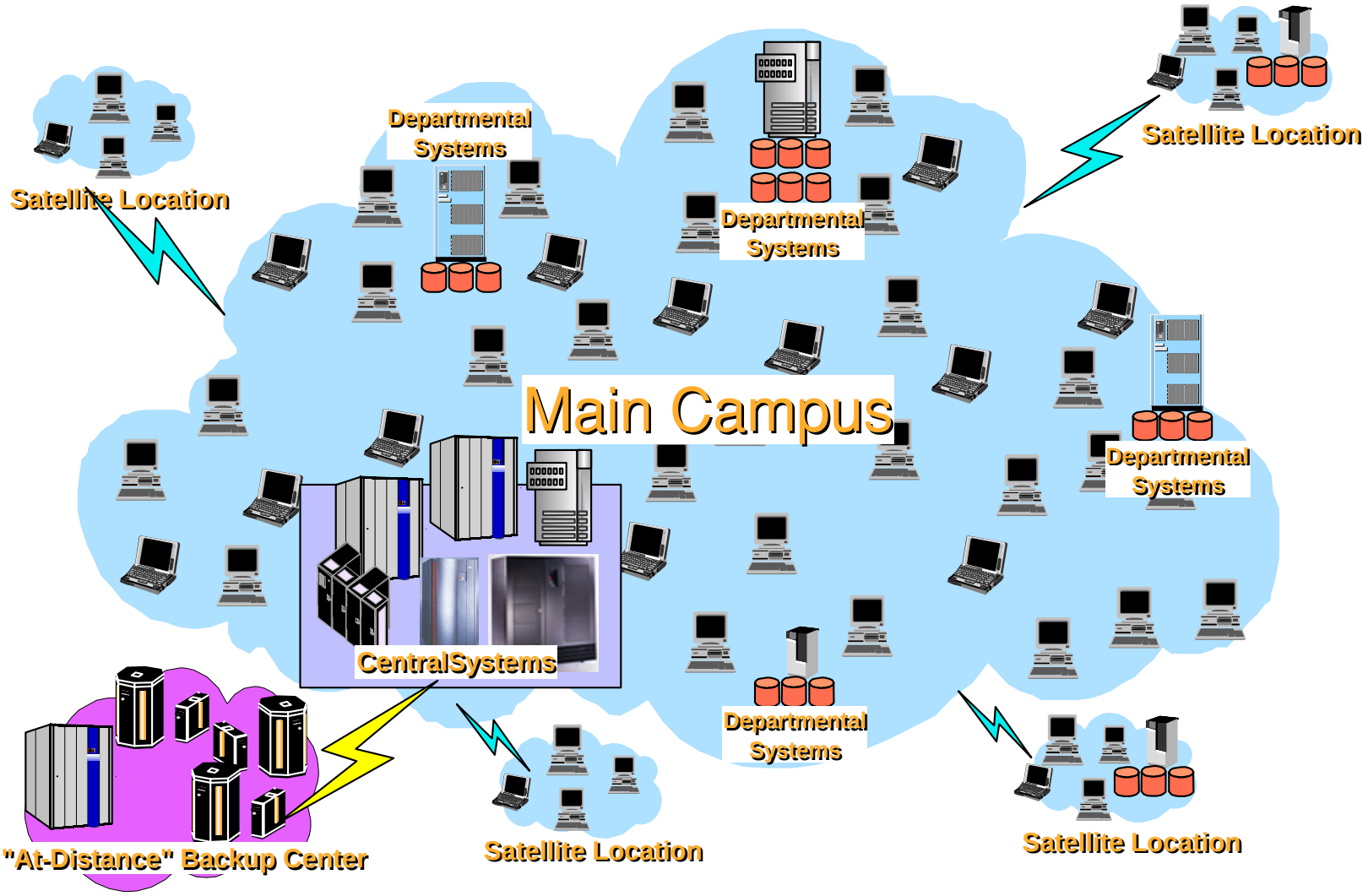
IP Storage Combo -- NAS & iSCSI



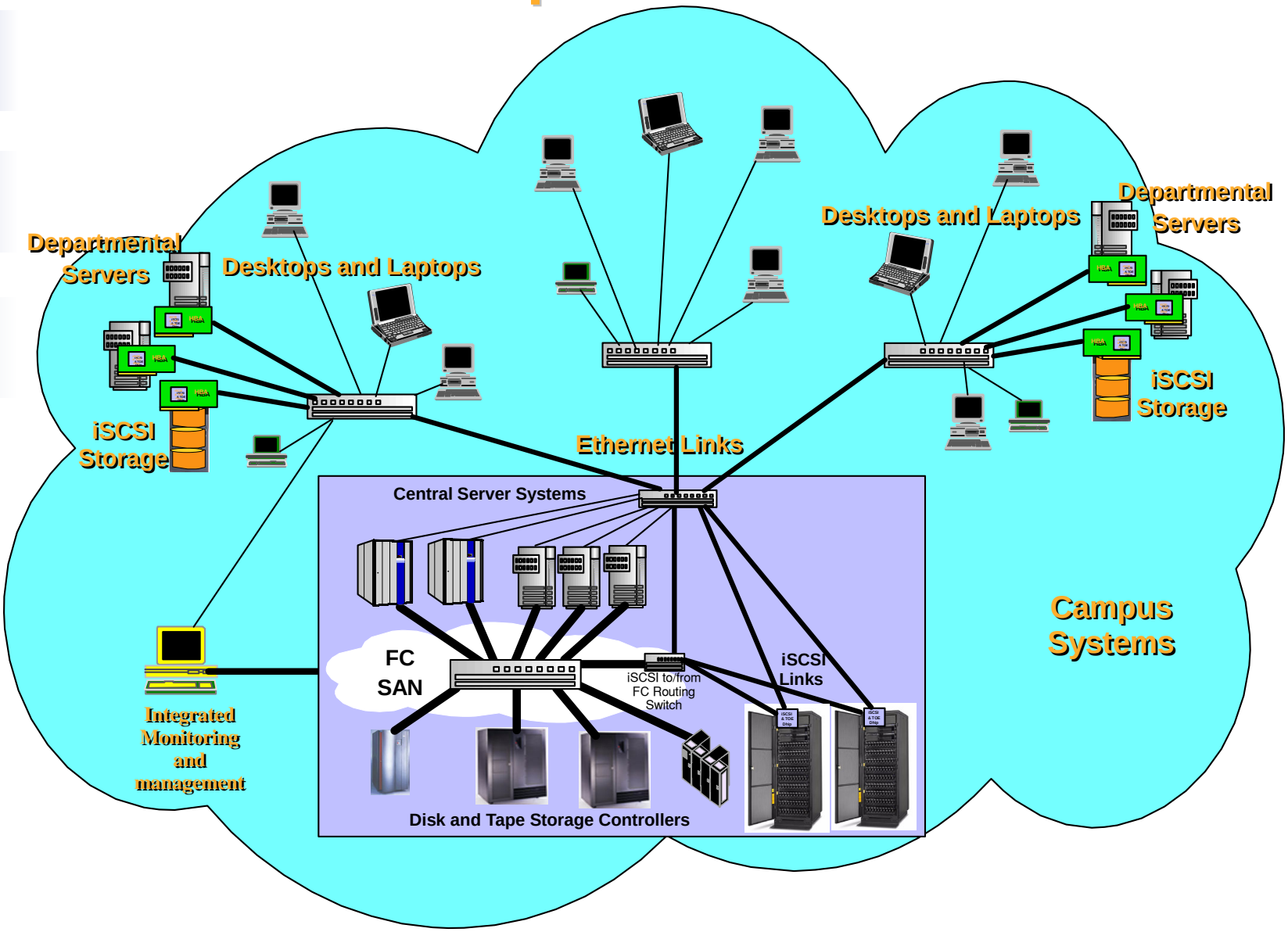
Midrange Enviornment



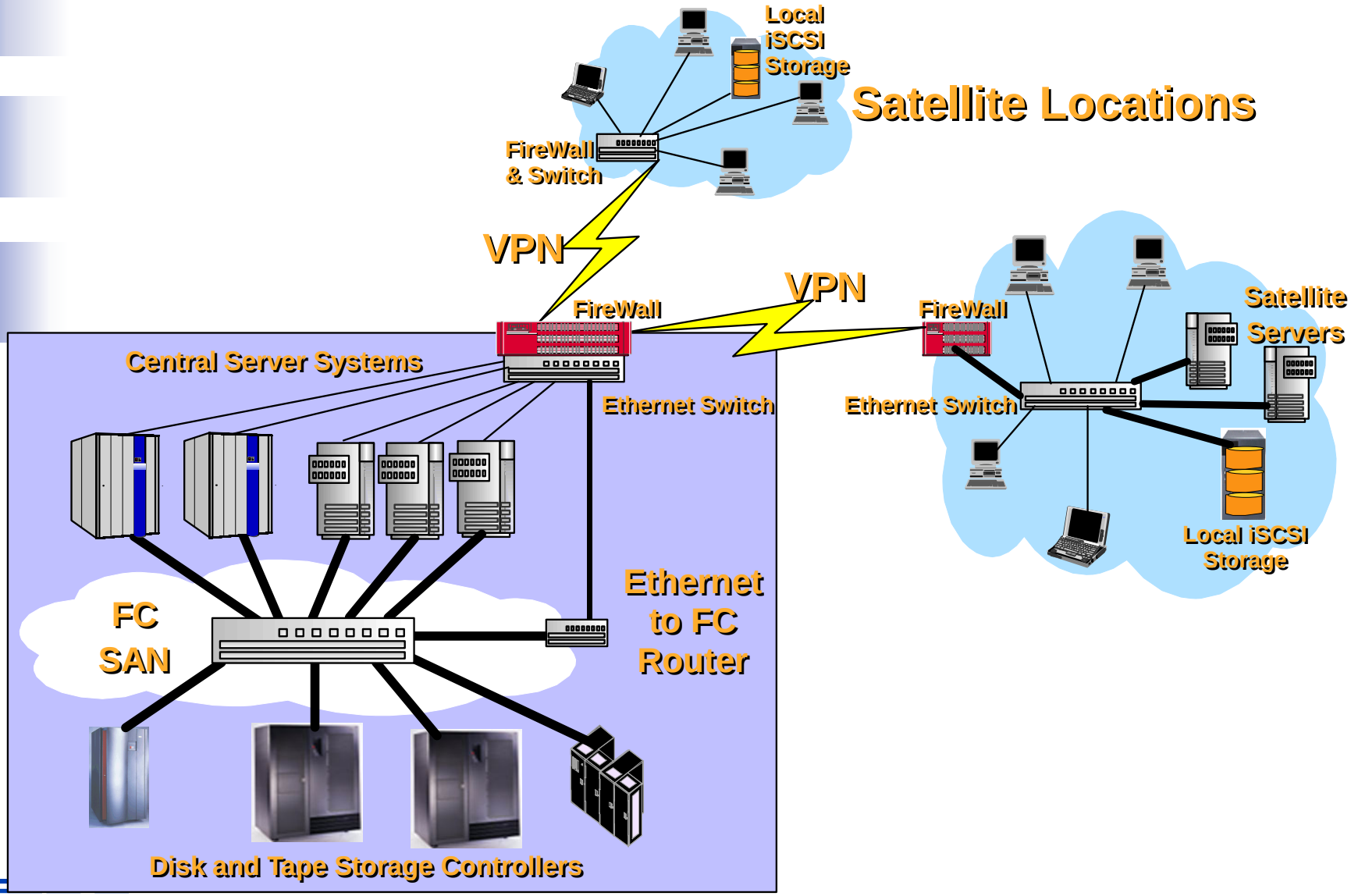
High-End Enviornment



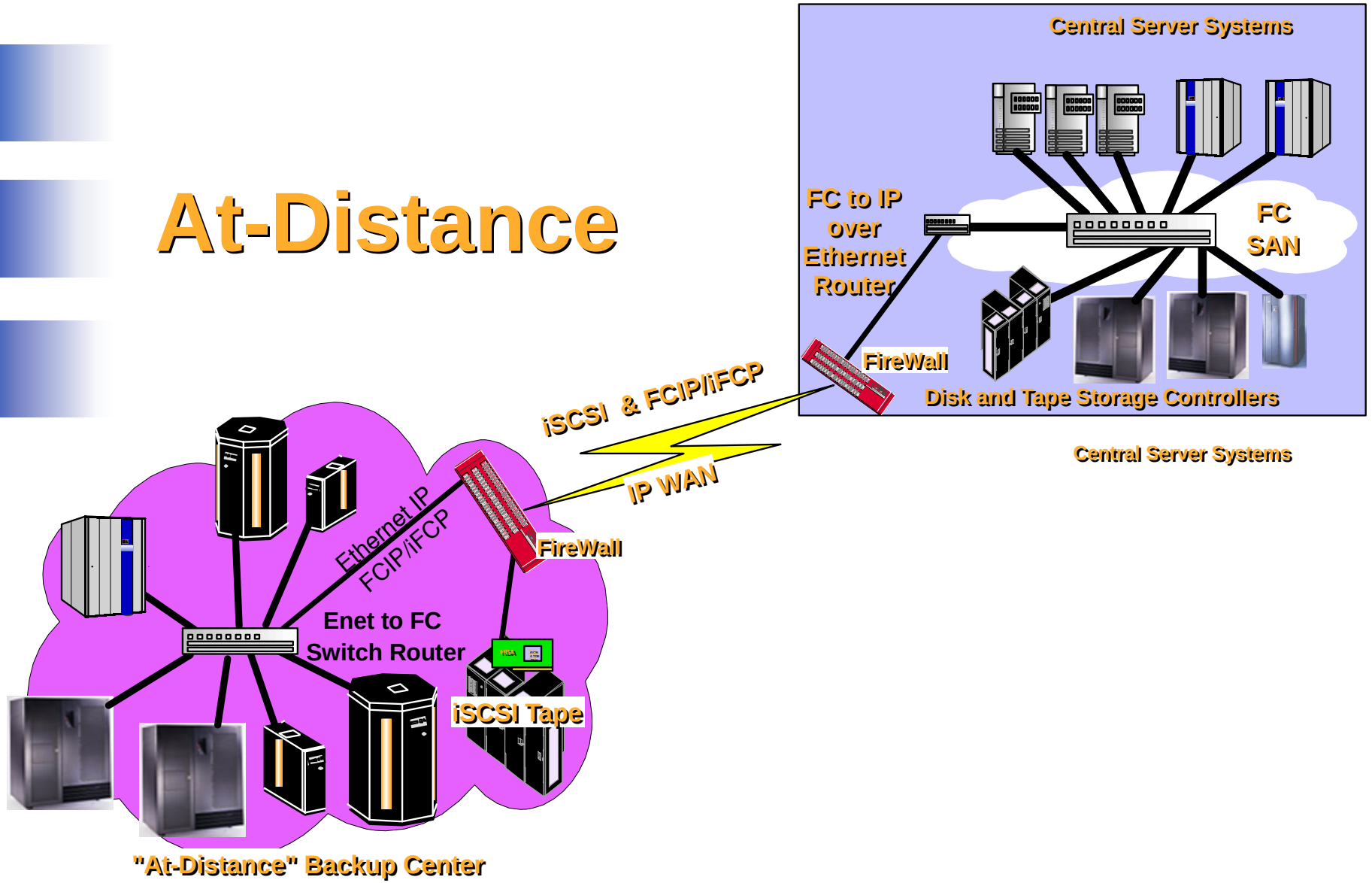
Campus Network



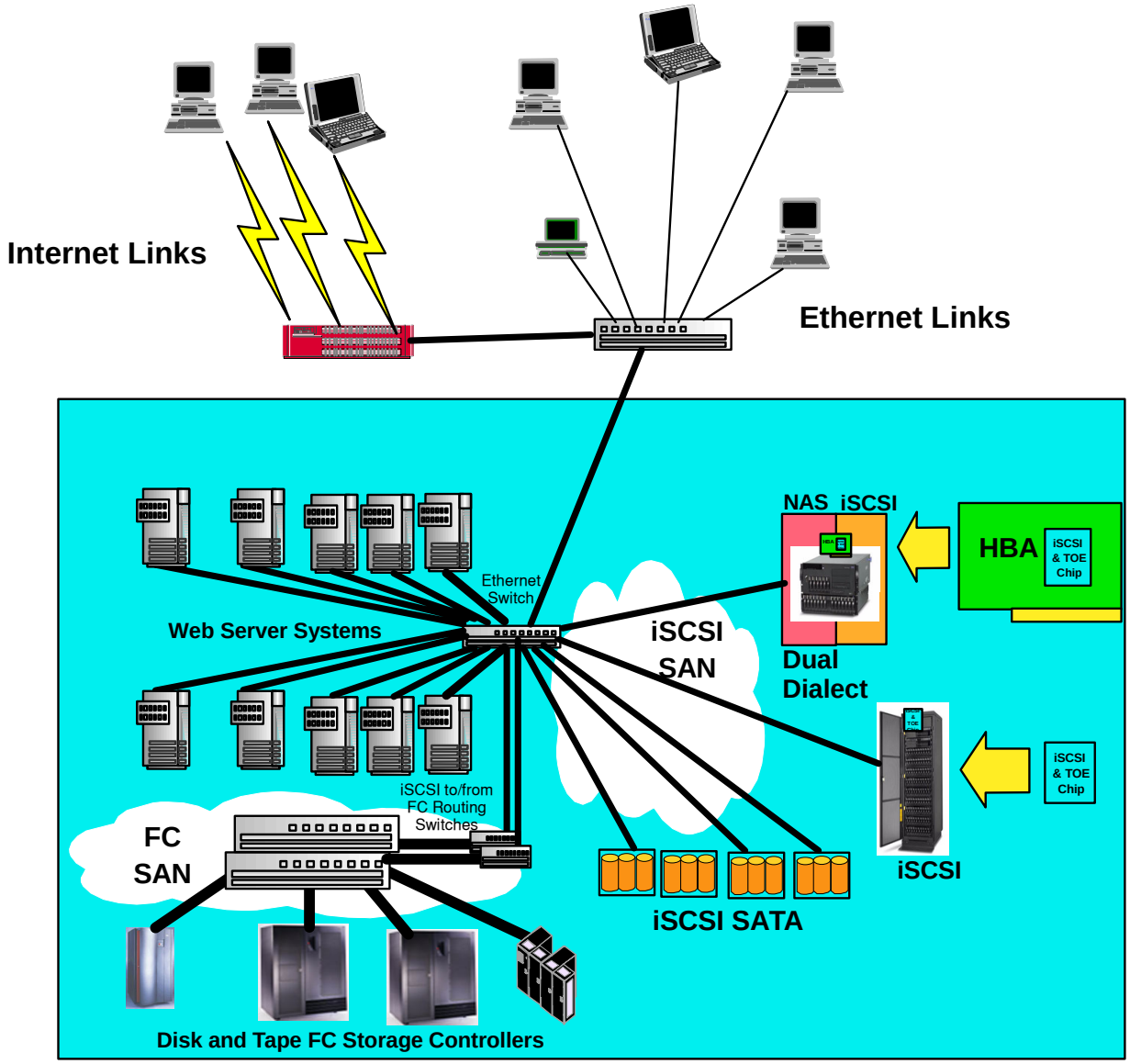
Satellite and Central System/Storage



At-Distance

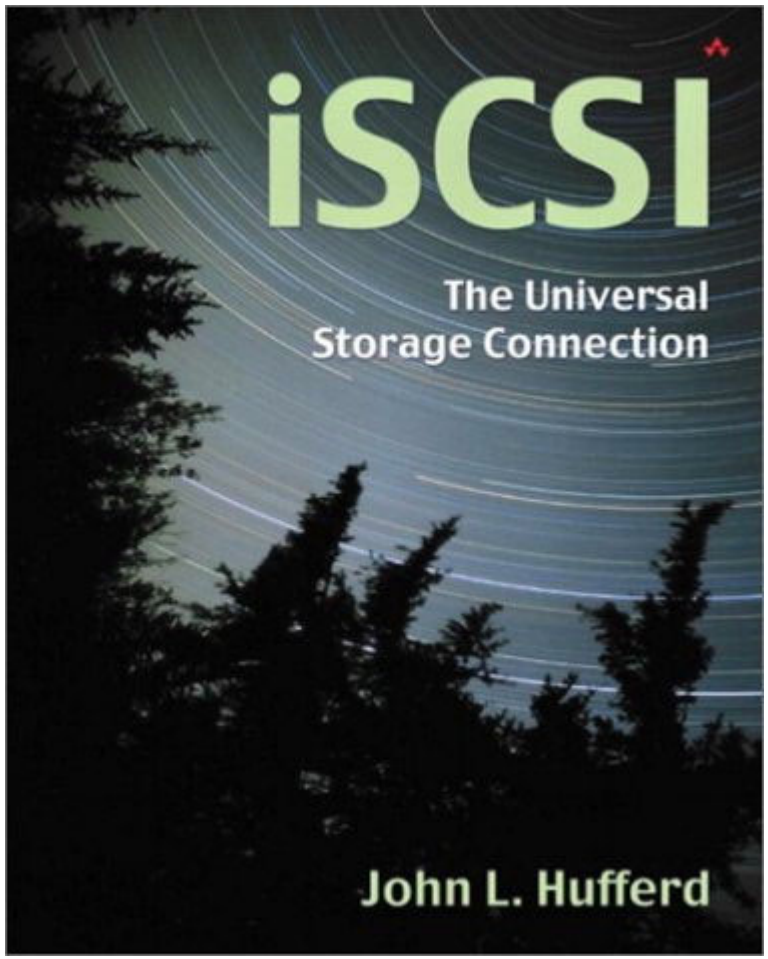


Web Server Installation



Conclusions

iSCSI Reference Book



Published by Addison-Wesley

Available in Book Stores

and Amazon.com

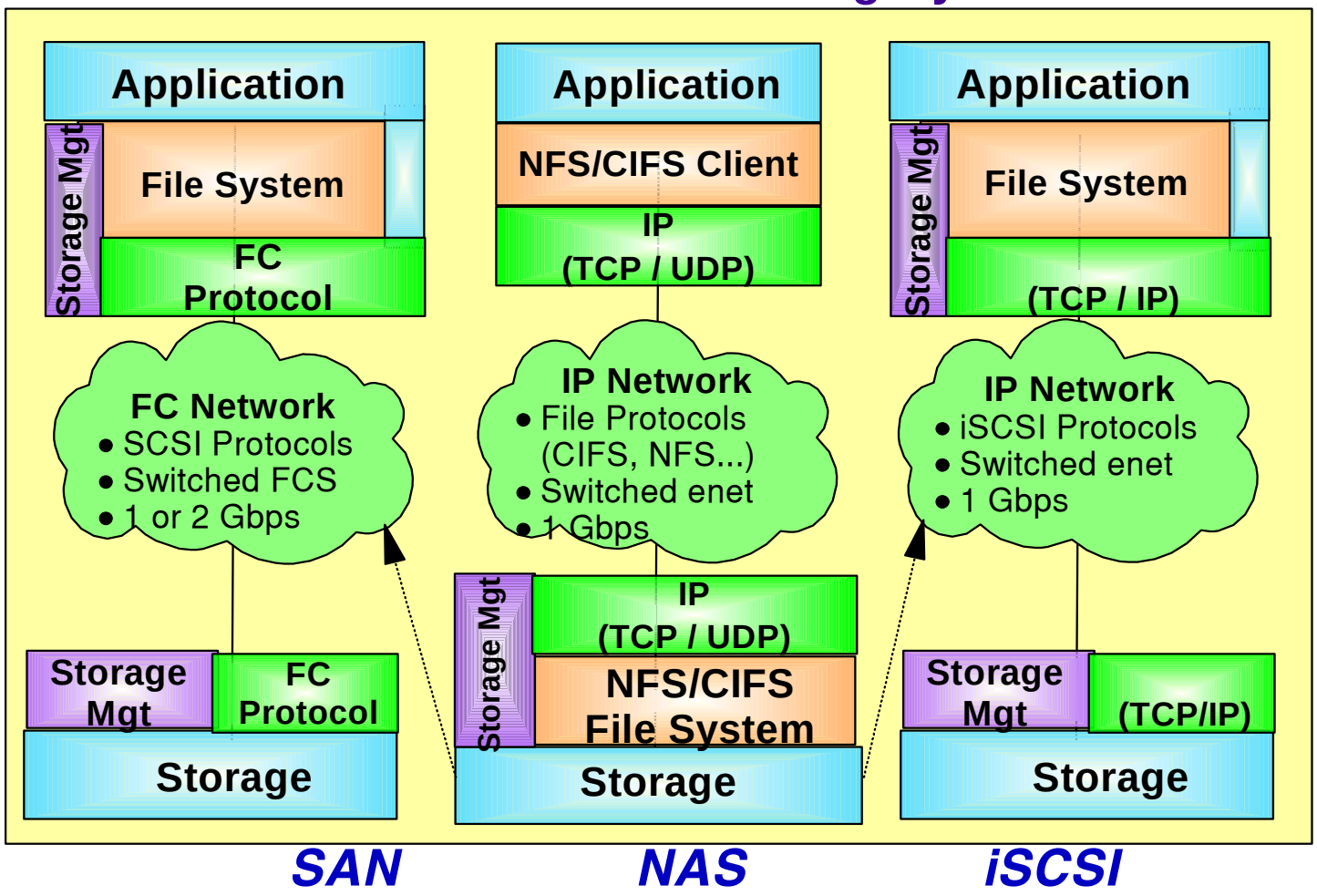
Volume purchases available

Appropriate for Marketing, Sales, Engineering personnel, and IT personnel

Additional Information

Contrasting Storage networking Technologies

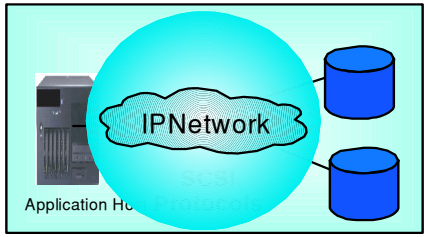
Functional Placement / Processing Cycles



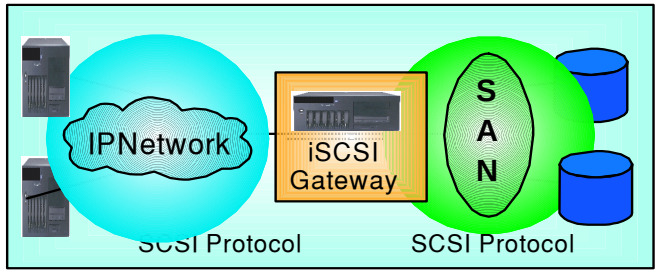
iSCSI Deployment

Same HW Configurations as NAS
Workgroup, Departmental, & Enterprise
(Appliances and Gateways)
GAs throughout 2003 & 2004

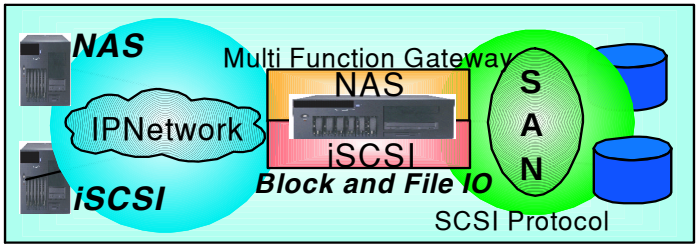
Independent iSCSI Deployment



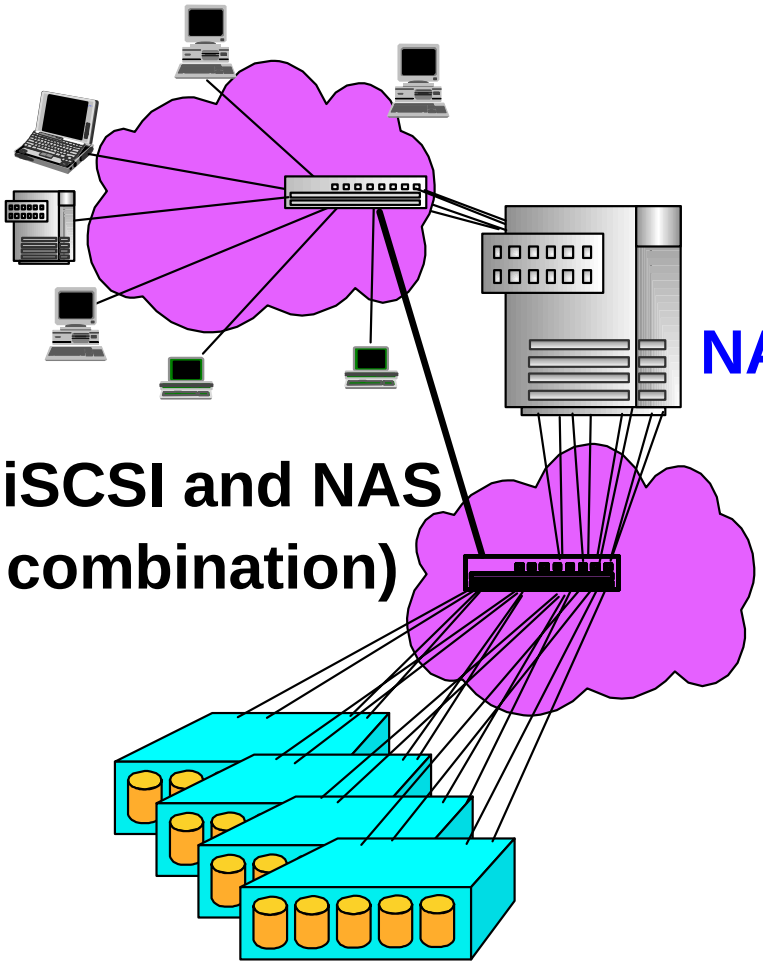
Extending the SAN



In Combination with NAS



Peaceful Co-existence
iSAN & NAS



NAS iSCSI-Gateway

Supports both iSCSI and NAS
(a Dual Dialect combination)

iSCSI Raid Ctlrs

RETURN TO INDEX