iSCSI for Proliant

Gene Chesser

Technical Director
HP Network Storage Solutions

HP WORLD 2003
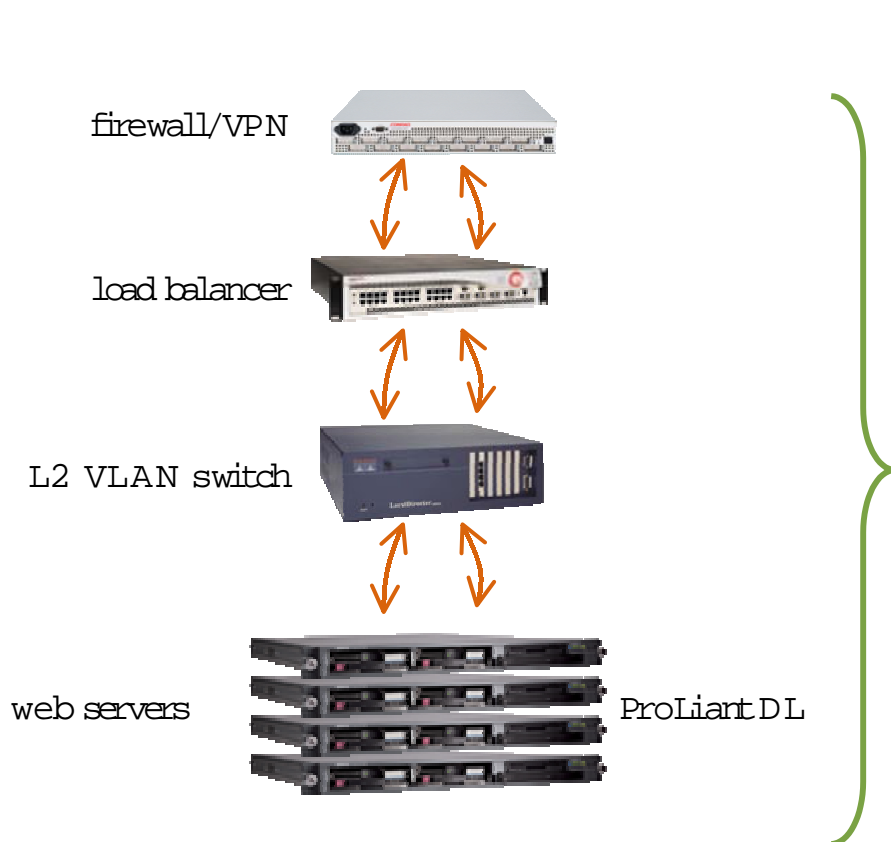Solutions and Technology Conference & Expo

Agenda

1. Value Proposition for iSCSI on Proliants
2. Focus on Blade Servers
3. p-Class GbE2 Switch
4. SR 2122 iSCSI Storage Router
5. Routing Architectures
6. iSCSI Challenges
7. Feedback

Value Proposition for iSCSI on Proliants
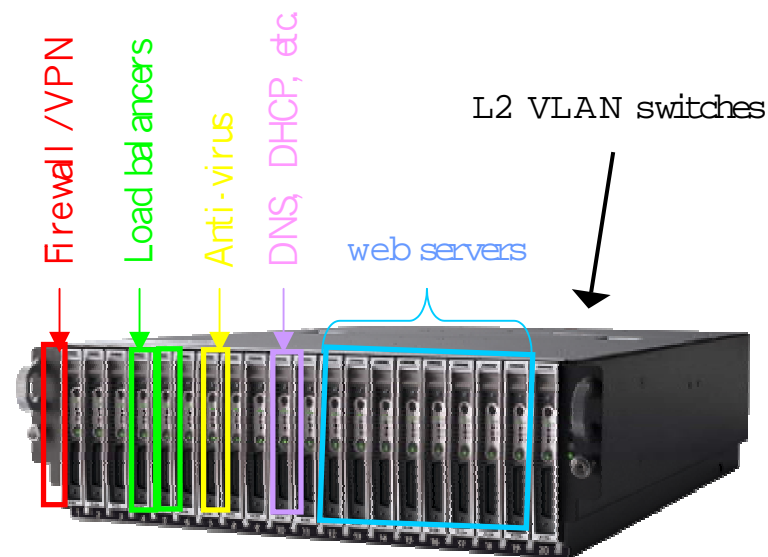

- IT and Server Consolidation Needs
  - Space constrained datacenters
  - Need for reduction of cabling
  - Personnel shortages
  - Managing multiple & remote sites
- Exploit the advantage of networked storage
  - Fibre Channel provides the storage backbone today
  - iSCSI bridging provides access to Ethernet based servers
  - iSCSI is cost effective

Web & Infrastructure Front-Ends

firewall/VPN

load balancer

L2 VLAN switch

web servers   ProLiant DL

Firewall /VPN

Load balancers

Anti-virus

DNS, DHCP, etc

web servers

L2 VLAN switches

rack-mounted server architecture

blade server enclosure

Focus on Blade Servers

edge server blades

BL e-Class (BL10e)

static web server

infrastructure apps

(firewall, gateway, security, DNS, DHCP)

computational cluster node

2P performance front & mid-tier server blades

BL p-Class (BL20p, BL20p G2)

dynamic web / ASP hosting

computational cluster node

terminal server farm

AV, media streaming
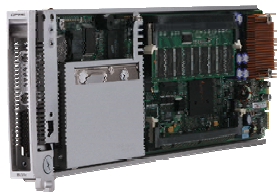
4P performance back-end server blades

BL p-Class (BL40p)

database server

mail / messaging server

HA failover cluster

ProLiant BL e-Class

❶ ProLiant BL10e server blade　**❷** ProLiant BL e-Class server blade enclosure

Rapid Deployment Pack

•Each Blade has Two Ethernet
　　　Ports 10/100
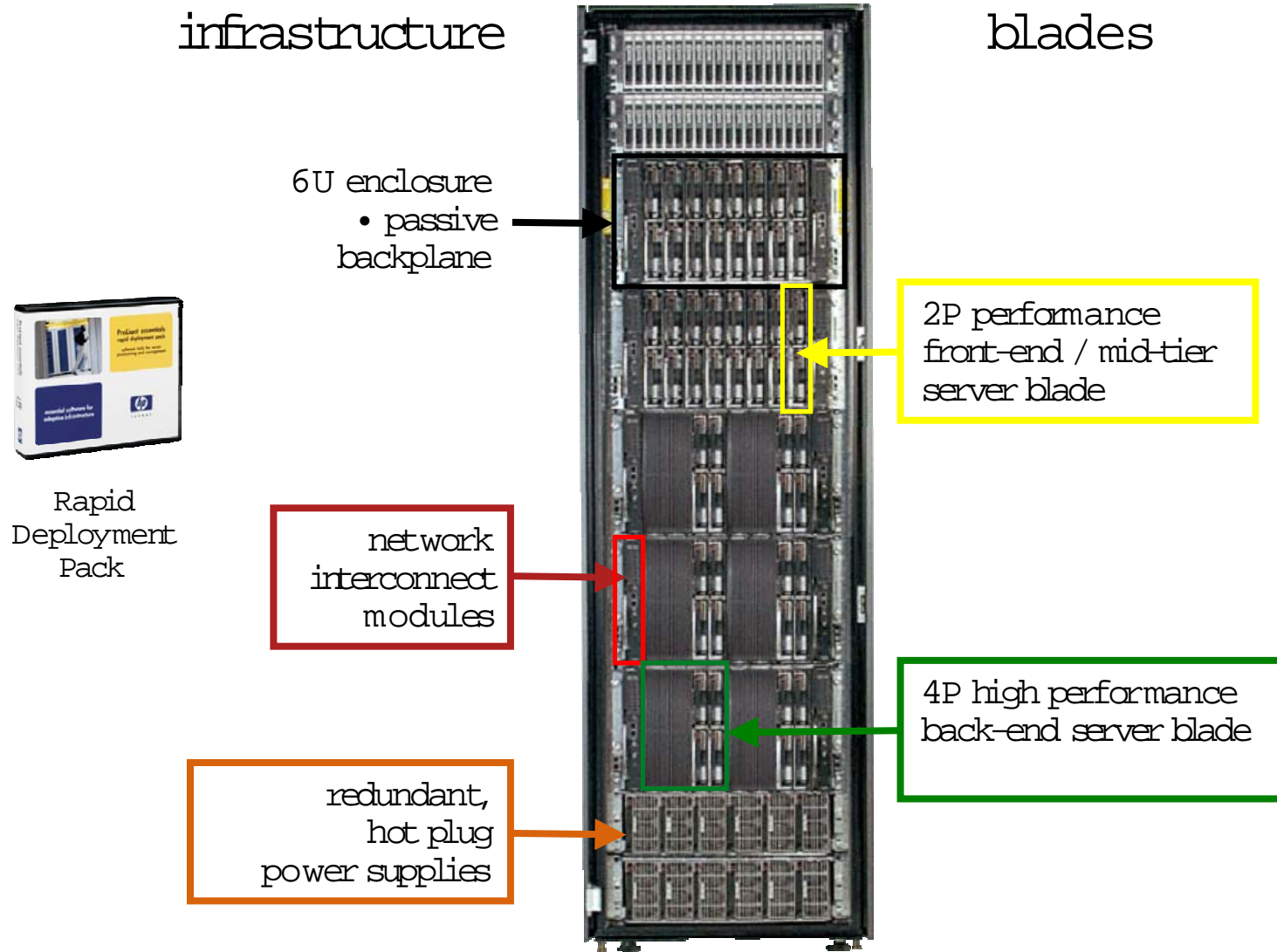
Interconnect tray features

- Integrated Administrator – Enables *serial* text remote console to each individual blade

Interconnect tray options

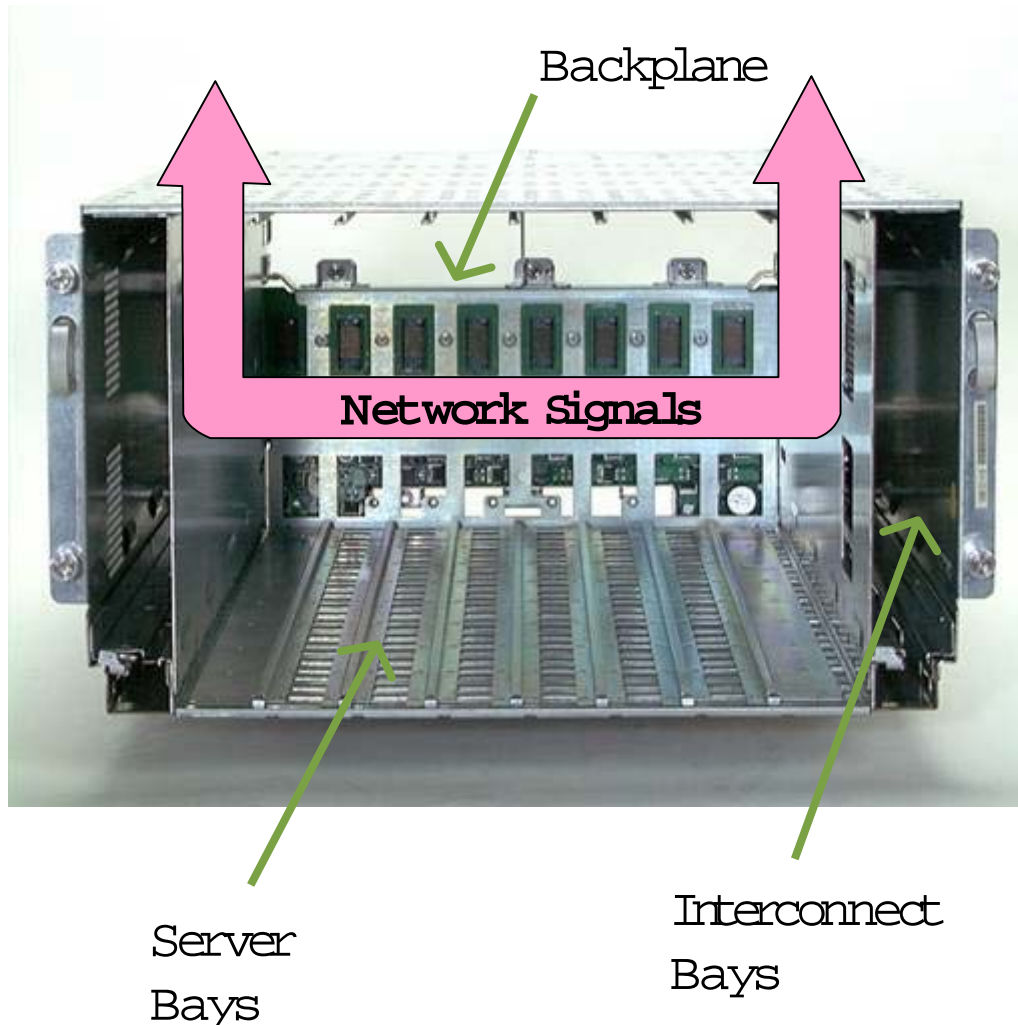- Integrated redundant switches to reduce network cables

- Patch panel

ProLiant BL p-Class
anatomy

infrastructure   blades

6U enclosure
• passive
backplane

Rapid
Deployment
Pack

2P performance
front-end / mid-tier
server blade

network
interconnect
modules

4P high performance
back-end server blade

redundant,
hot plug
power supplies

BL p-Class Network Enclosure Routing

Backplane

Network Signals

Server
Bays

Interconnect
Bays

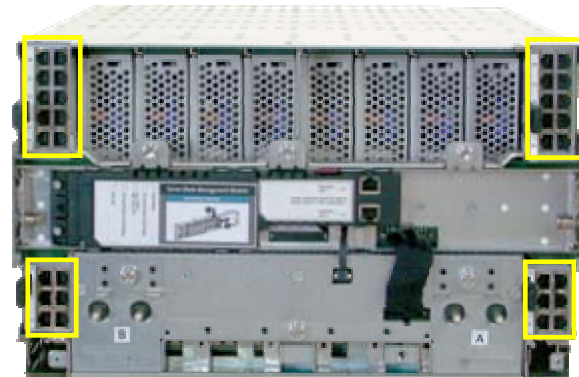Server network signals routed from backplane to interconnects and out rear of enclosure

signals are divided between the two interconnects

- improves availability
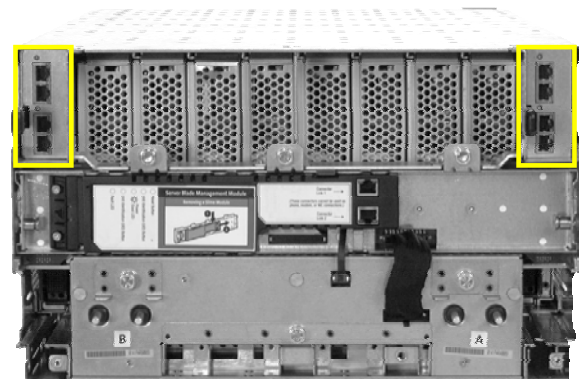
BL p-Class network cabling
interconnect options

– option 1: patch panel with one RJ-45 port per each network adapter

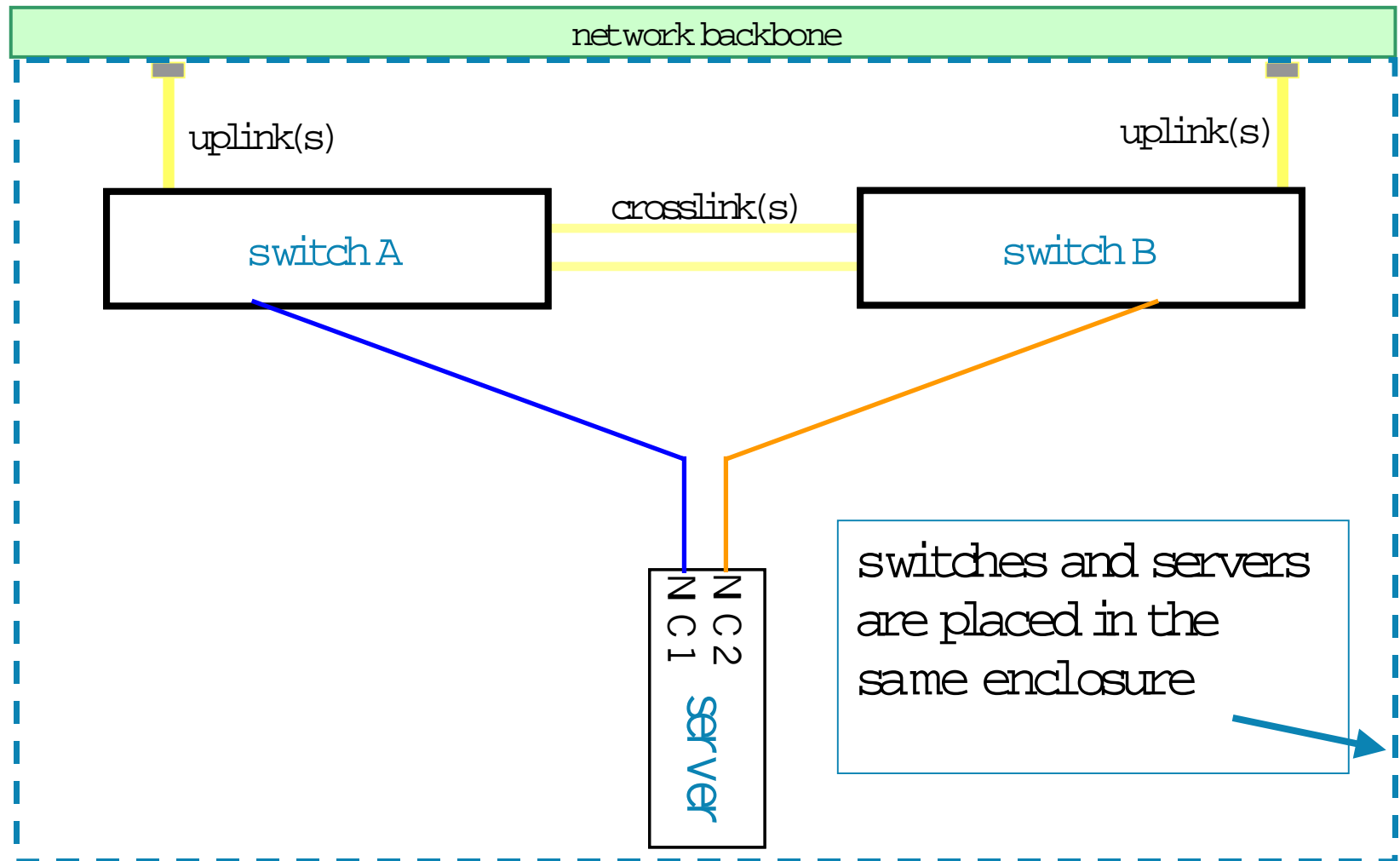rear view of p-Class enclosure with patch panel

– option 2: GbE Interconnect Switch with a few uplink ports to reduce network cabling

rear view of p-Class enclosure with interconnect switch

Standard Server Network Configuration

(with redundancy)

network backbone

uplink(s)   uplink(s)

switch A   crosslink(s)   switch B

N C 2
N C 1 server

switches and servers are placed in the same enclosure

BL p-Class GbE2 Interconnect Switch
overview

GbE2 will be shipping in September.

– Partnership with Nortel

– all gigabit layer 2 Ethernet switch

- 32 10/100/1000T downlinks
- 2 10/100/1000T crosslinks
- 8 1000SX or 10/100/1000T uplinks
- 4 10/100/1000T ports at front

– architecture supports future options

- Layer 3-7 load balancing option
- 10GbE uplink option

– additional network features

– BL20p G2 Fibre Channel signal pass-through option

BL p-Class GbE interconnect switch

redundant modular design

- left and right interconnects

- front mounted blade, rear mounted modules

- hot-pluggable

- rapidly remove blade (no disconnecting cables)

- redundant cross-links between switches.

- Half of signals route to one side; half to other, so always have connection to each blade even if 1 switch fails

rear mounted interconnect module

front mounted interconnect blade
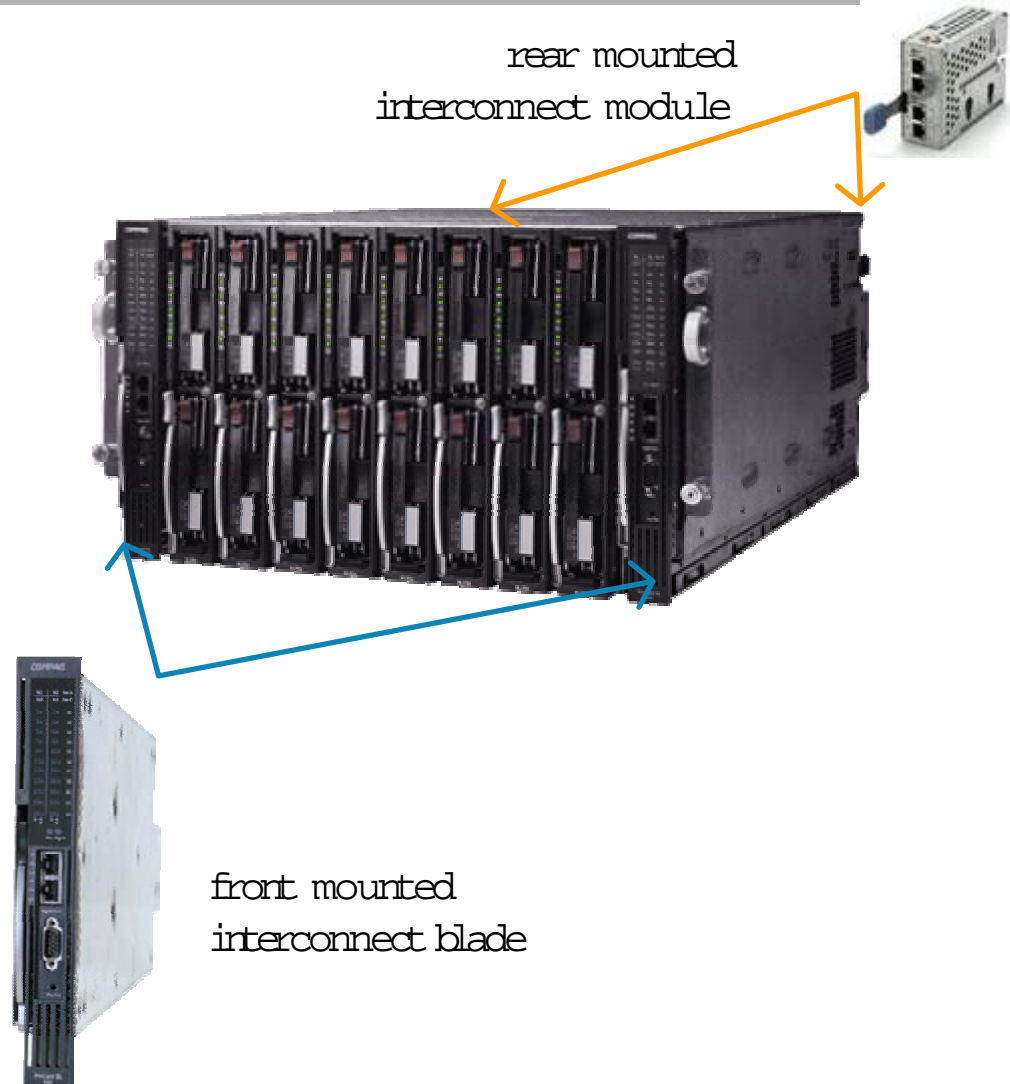
p-Class GbE2 switch overview

- **same form factor as GbE switch**
  - except LAN is moved to smaller (lower) cube, SAN is top cube

- **supported on existing p-Class infrastructure**
  - use with any mix of ProLiant BL p-Class servers

- **24-port Gigabit Ethernet, layer 2 enterprise switch**
  - 16 internal GbE "downlink" ports to the server NICs
  - 2 internal GbE "crosslink" ports linking the two switches
  - 4 external GbE copper- or fiber-based "uplink" ports on rear hot-plug cube
  - 2 external GbE "management" ports on front

- **enhanced set of network protocol and standards**

- **architected to support future options**
  - 10 GbE uplink and layer 3-7 switching

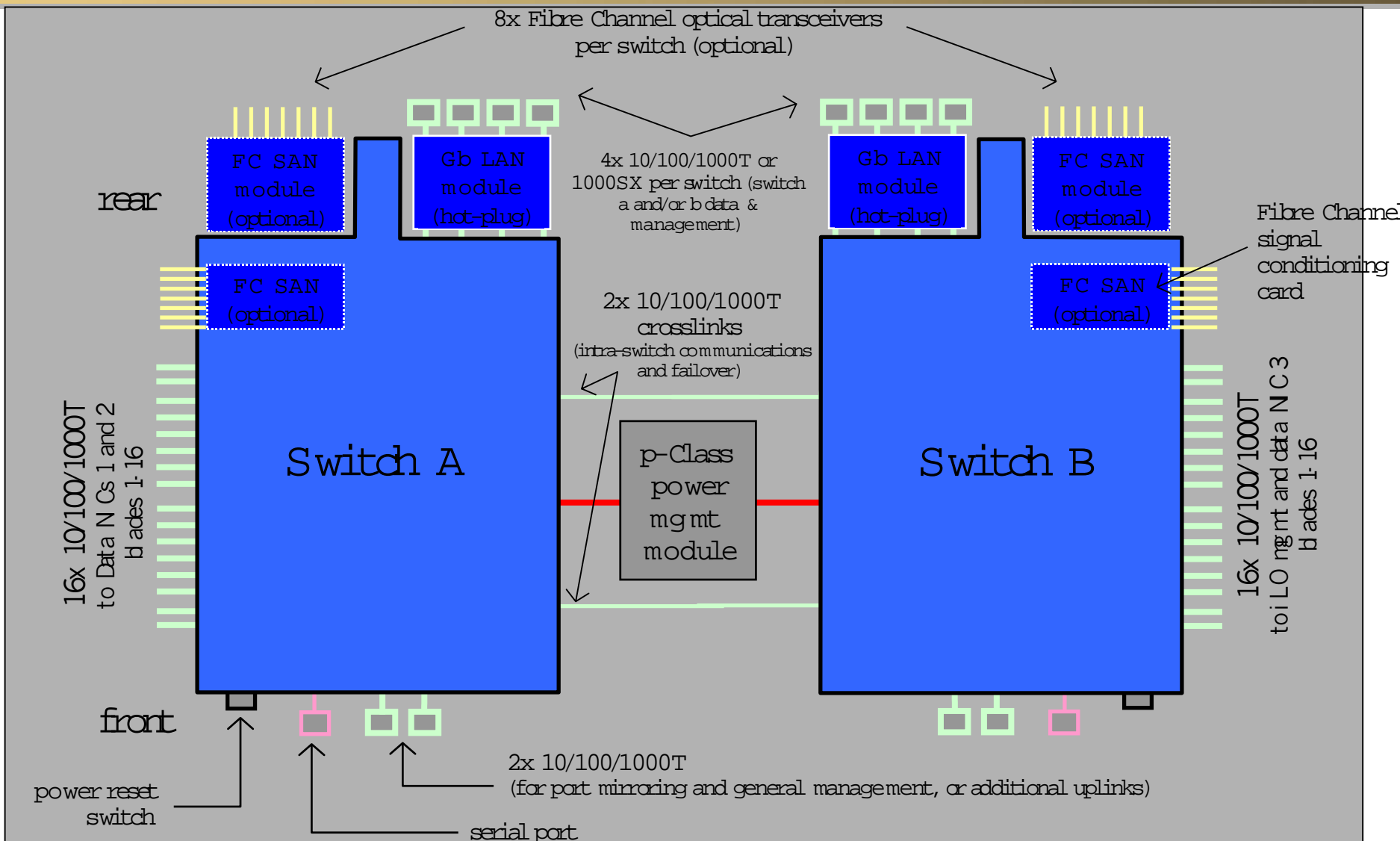- **BL20p G2 Fibre Channel pass-through option**

p-Class GbE2 switch future upgrade options

- HP/ Nortel Partnering

- layer 3-7 switching option
  - Nortel-only offering, Future
  - two versions (standard and high performance)
  - add L3-7 switch mezz card + f/w upgrade per switch
  - user installable in field

- 10GbE uplink option
  - HP offering, Future
  - GbE2 switch includes a 4-port 10GbE bridge ASIC
  - add 10GbE ASIC mezz card per switch
  - replace 4x 1Gb uplink cube with new 10GbE cube
  - user installable in field
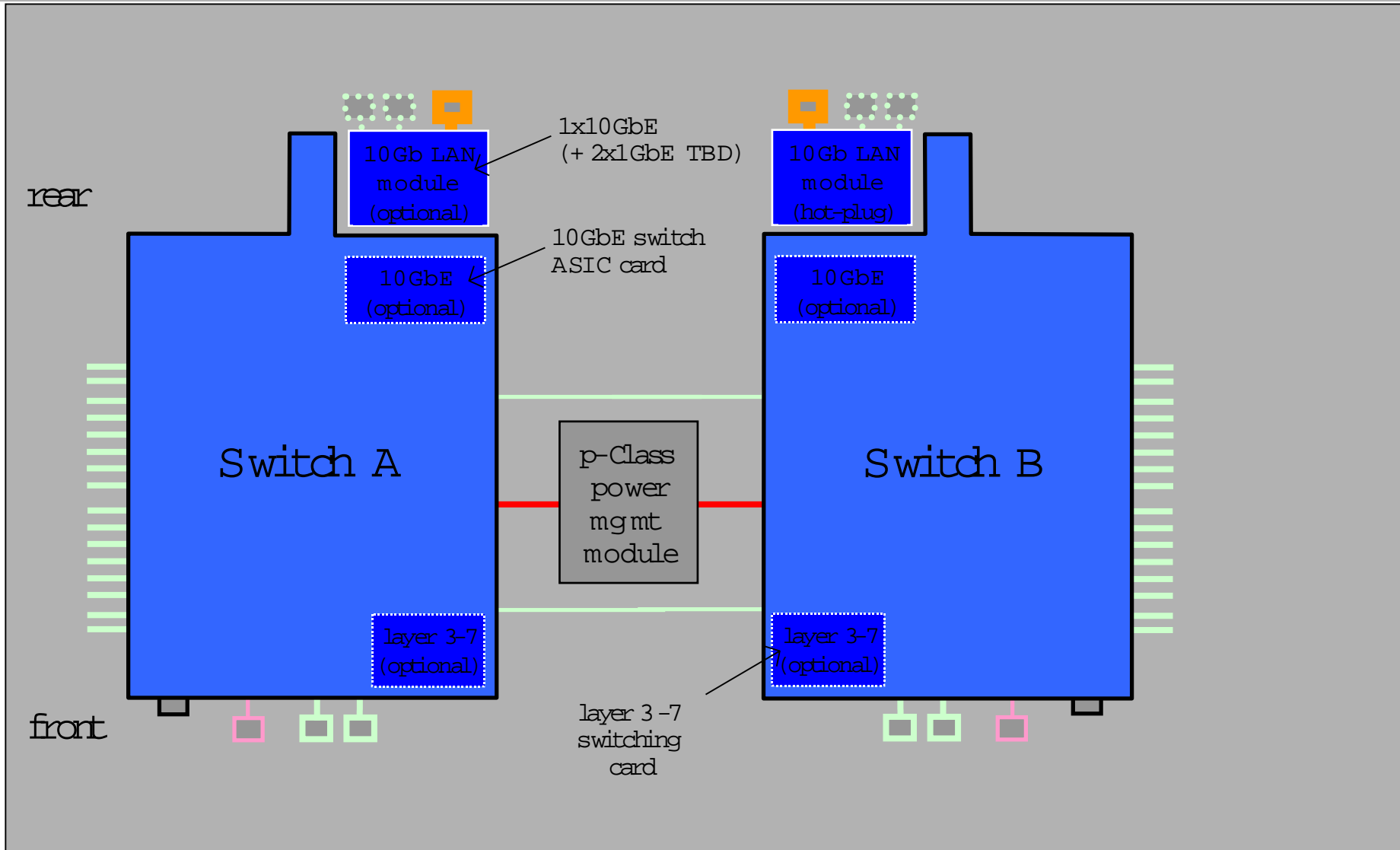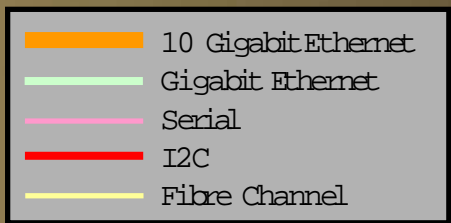
p-Class GbE2 switch architecture
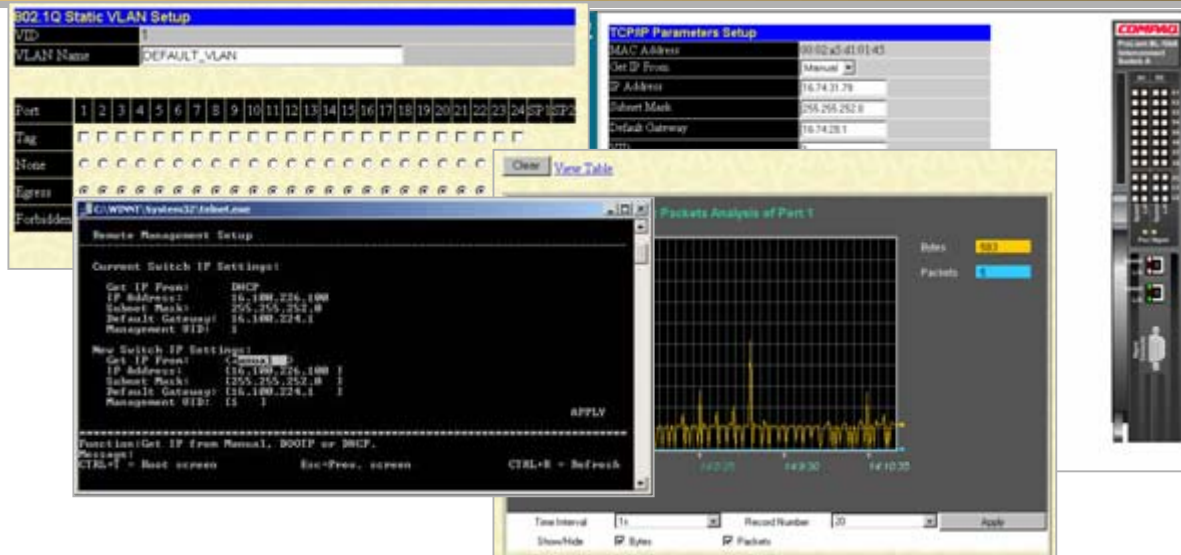## (standard configuration with FC SAN option)

Gigabit Ethernet
Serial
I2C

8x Fibre Channel optical transceivers per switch (optional)

rear

FC SAN module (optional)

Gb LAN module (hot-plug)

4x 10/100/1000T or 1000SX per switch (switch a and/or b data & management)

Gb LAN module (hot-plug)

FC SAN module (optional)

Fibre Channel signal conditioning card

FC SAN (optional)

FC SAN (optional)

16x 10/100/1000T to Data NCs 1 and 2 blades 1-16

Switch A

2x 10/100/1000T crosslinks
(intra-switch communications and failover)

p-Class power mgmt module

Switch B

16x 10/100/1000T toi LO mgmt and data NC3 blades 1-16

front

power reset switch

2x 10/100/1000T
(for port mirroring and general management, or additional uplinks)

serial port

p-Class GbE2 switch architecture
(future add-on options – L3-7; 10GbE uplinks)

rear

1x10GbE
(+ 2x1GbE TBD)

10Gb LAN module (optional)

10Gb LAN module (hot-plug)

10GbE switch ASIC card

10GbE (optional)

10GbE (optional)

Switch A

p-Class power mgmt module

Switch B

layer 3-7 (optional)

layer 3-7 (optional)

layer 3-7 switching card

front

p-Class GbE interconnect switch management

- included web-based and console menu driven user interfaces

- command line interface (CLI) available as free firmware upgrade

- SNMP & RMON with SNMP scripting and recommended example scripts

- TFTP support to save and download interconnect switch configuration
  - rapidly deploy multiple systems; provides backup and restore capabilities

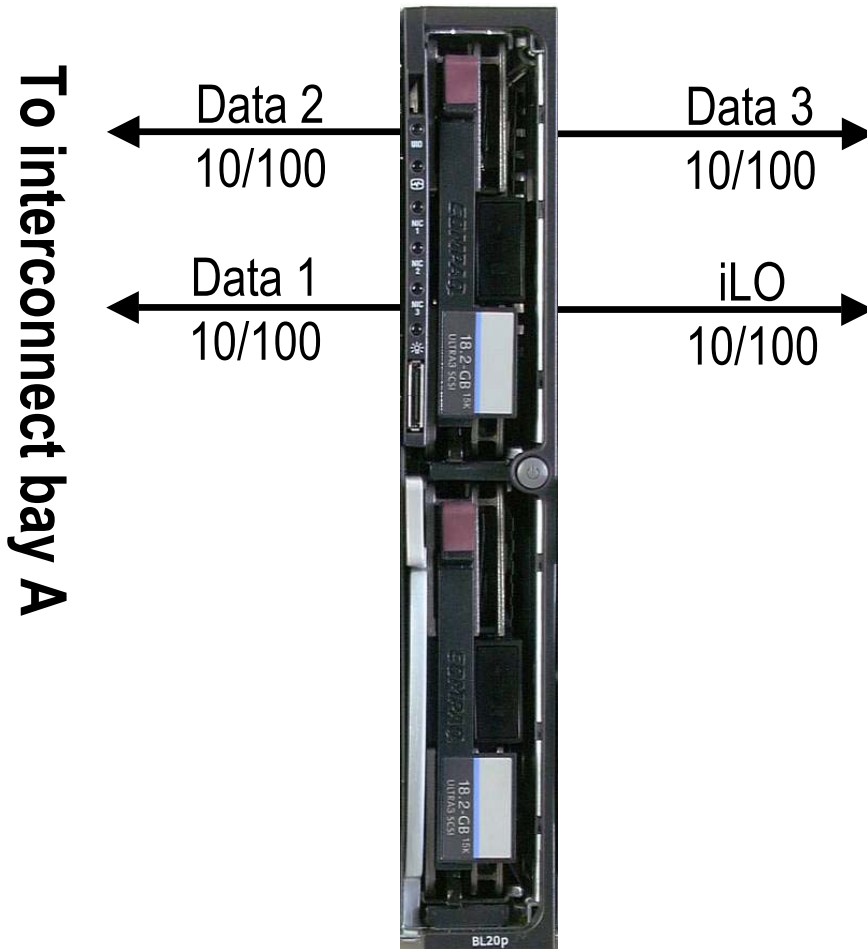- fully pre-configured and operational at power-up
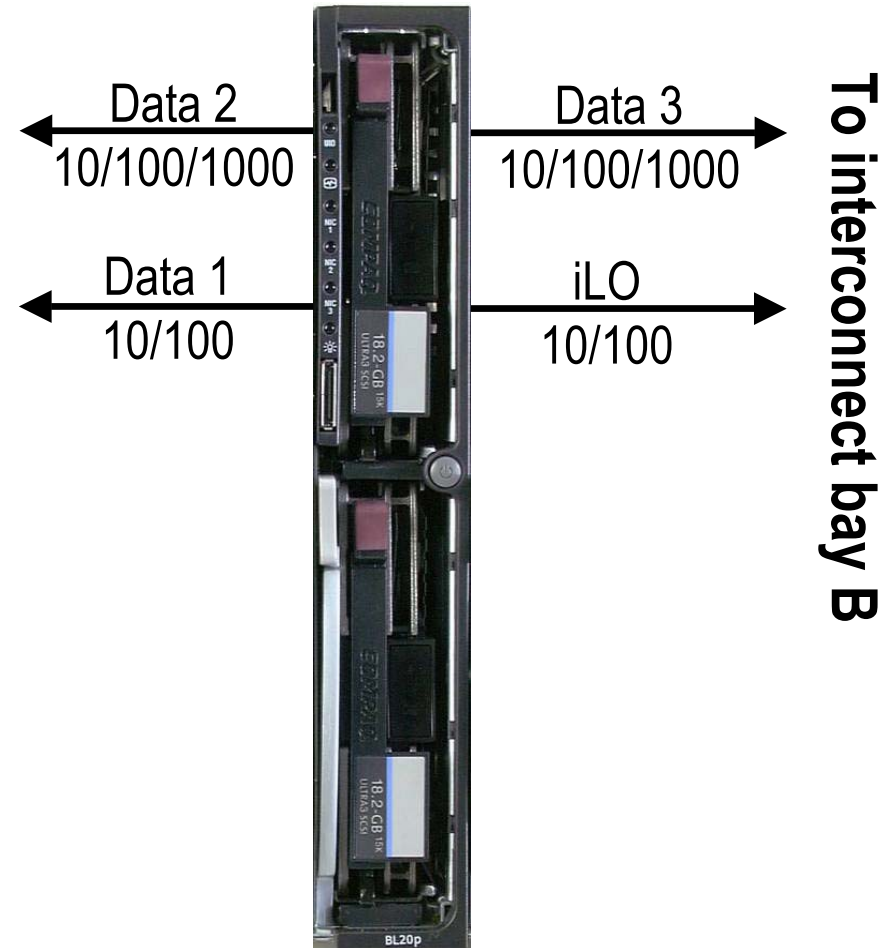
ProLiant BL Storage Connectivity
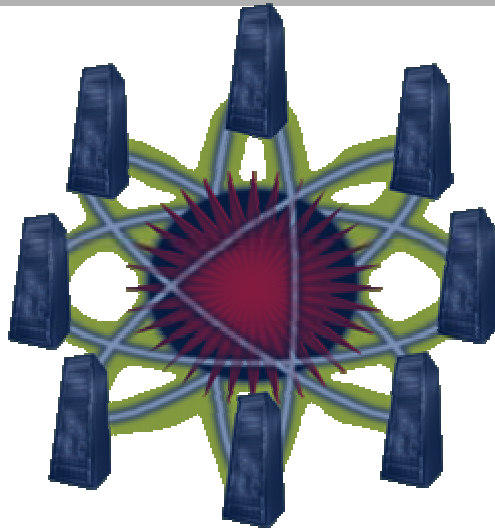
	ProLiant BL10e	ProLiant BL20p	ProLiant BL20p G2	ProLiant BL40p
NAS (Network Attached Storage)	✓	✓	✓	✓
iSCSI to FC (Using StorageWorks SR2122 iSCSI to FC router)	✓	✓	✓	✓
SAN (Fibre channel option)			✓	✓

ProLiant BL20p signal routing

Standard 10/100 configuration

Dual NC7780 gigabit option

To interconnect bay A

Data 2
10/100

Data 3
10/100

Data 1
10/100

iLO
10/100

Data 2
10/100/1000

Data 3
10/100/1000

Data 1
10/100

iLO
10/100

To interconnect bay B
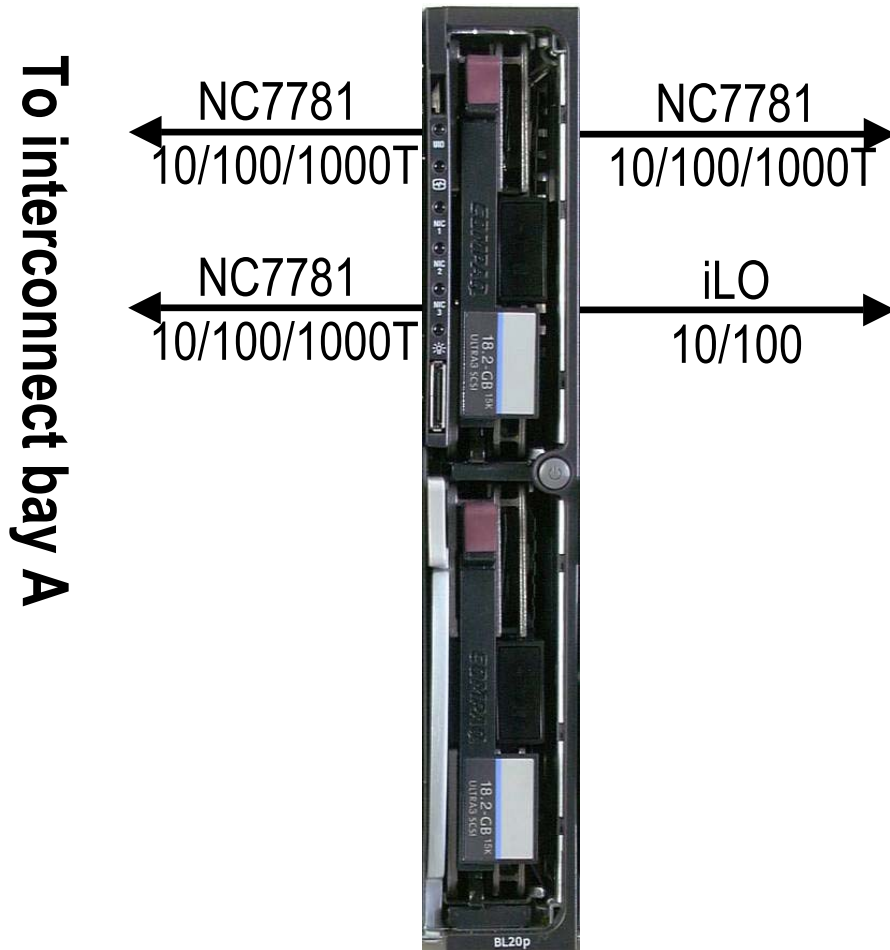
ProLiant BL20p G2 signal routing

Standard network mezzanine card **Fibre Channel mezzanine card**

To interconnect bay A

To interconnect bay B

NC7781
10/100/1000T
←

NC7781
10/100/1000T
→

NC7781
10/100/1000T
←

NC7781
10/100/1000T
→

NC7781
10/100/1000T
←

iLO
10/100
→

NC7781
10/100/1000T
←

iLO
10/100
→

Fibre
Channel
←
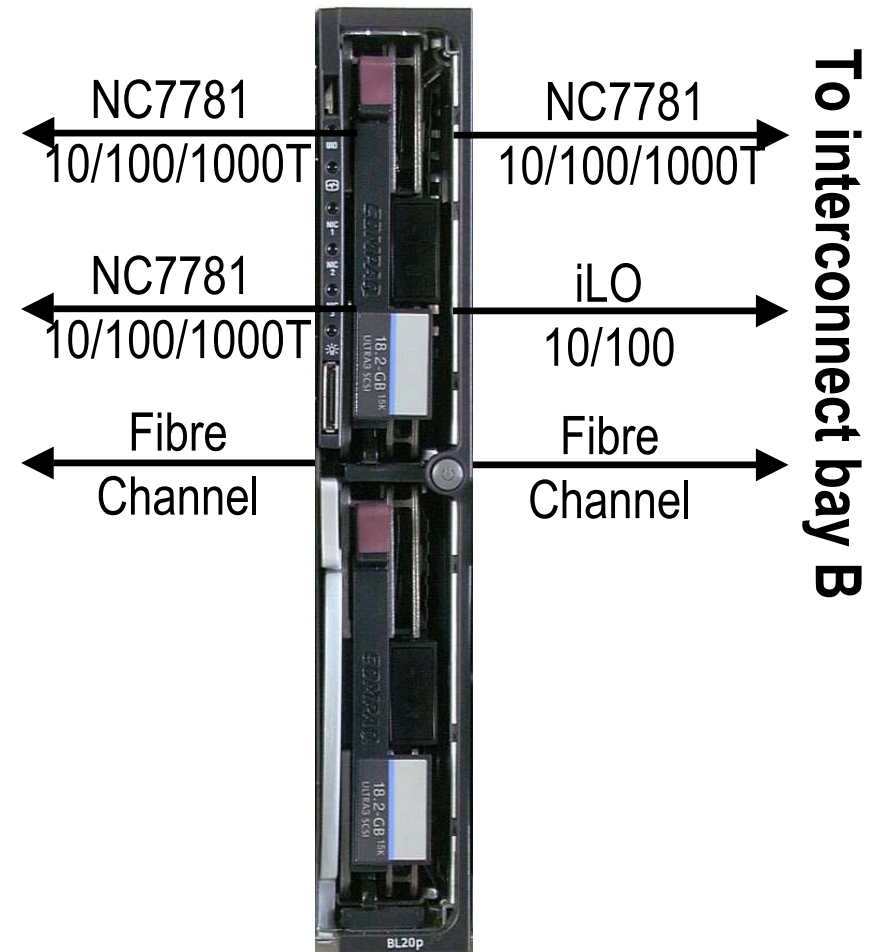
Fibre
Channel
→

ProLiant BL40p signal routing

To interconnect bay A

NC7781
10/100/1000T

NC7781
10/100/1000T

NC7781
10/100/1000T

To interconnect bay B

iLO
10/100T

NC7781
10/100/1000T

NC7781
10/100/1000T
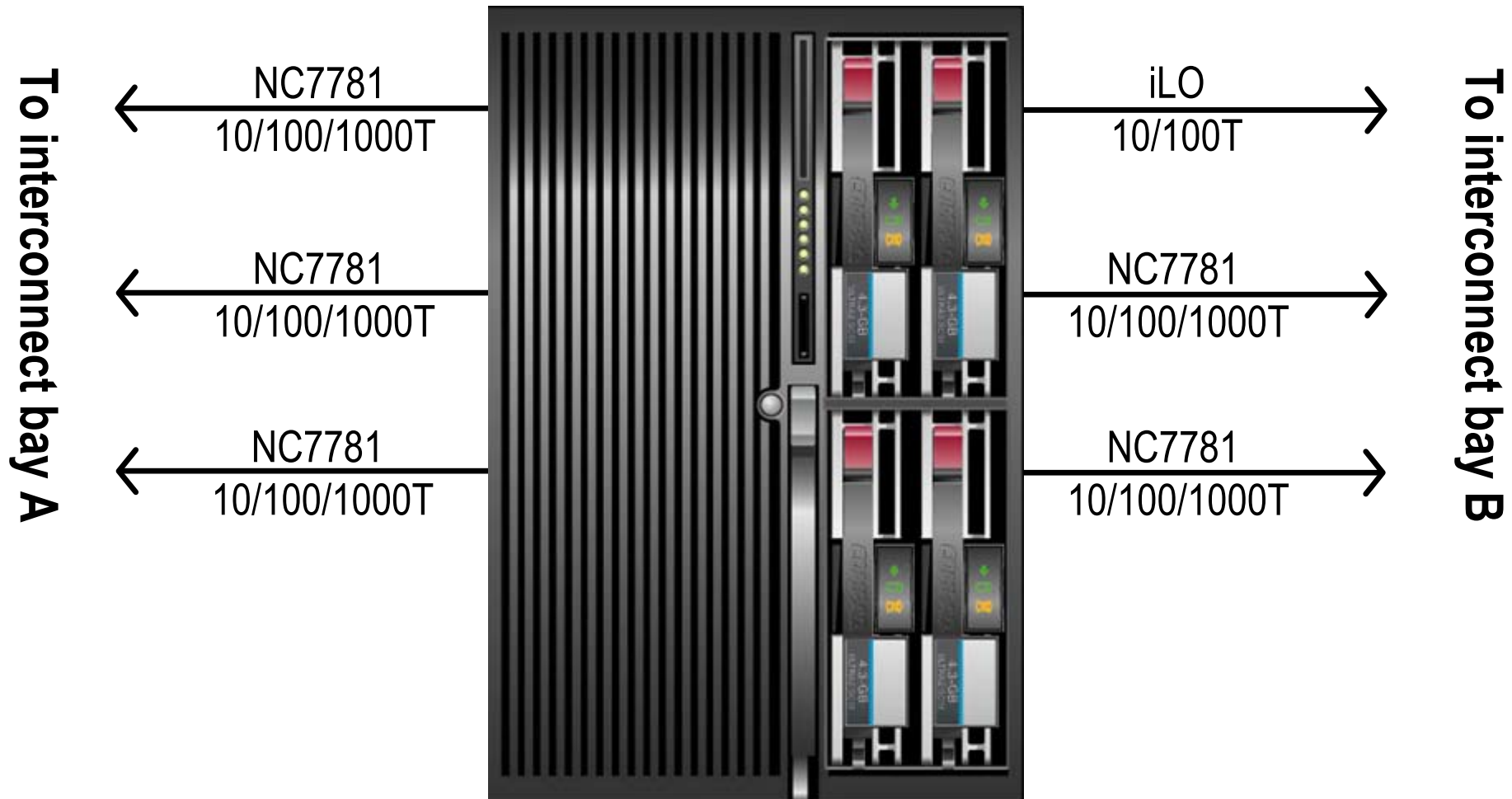
HP StorageWorks SR2122 iSCSI to FC Storage Router

Enables access to block storage on a Fibre Channel SAN across an Ethernet network

SR2122 iSCSI Storage Router bridge expands SAN benefits and reduces cost of Storage consolidation

SR2122 iSCSI Storage Router bridging to FC SAN

SR2122 iSCSI Storage Router bridging to Backup

SAN

IP Network

To Enterprise Backup System (EBS)
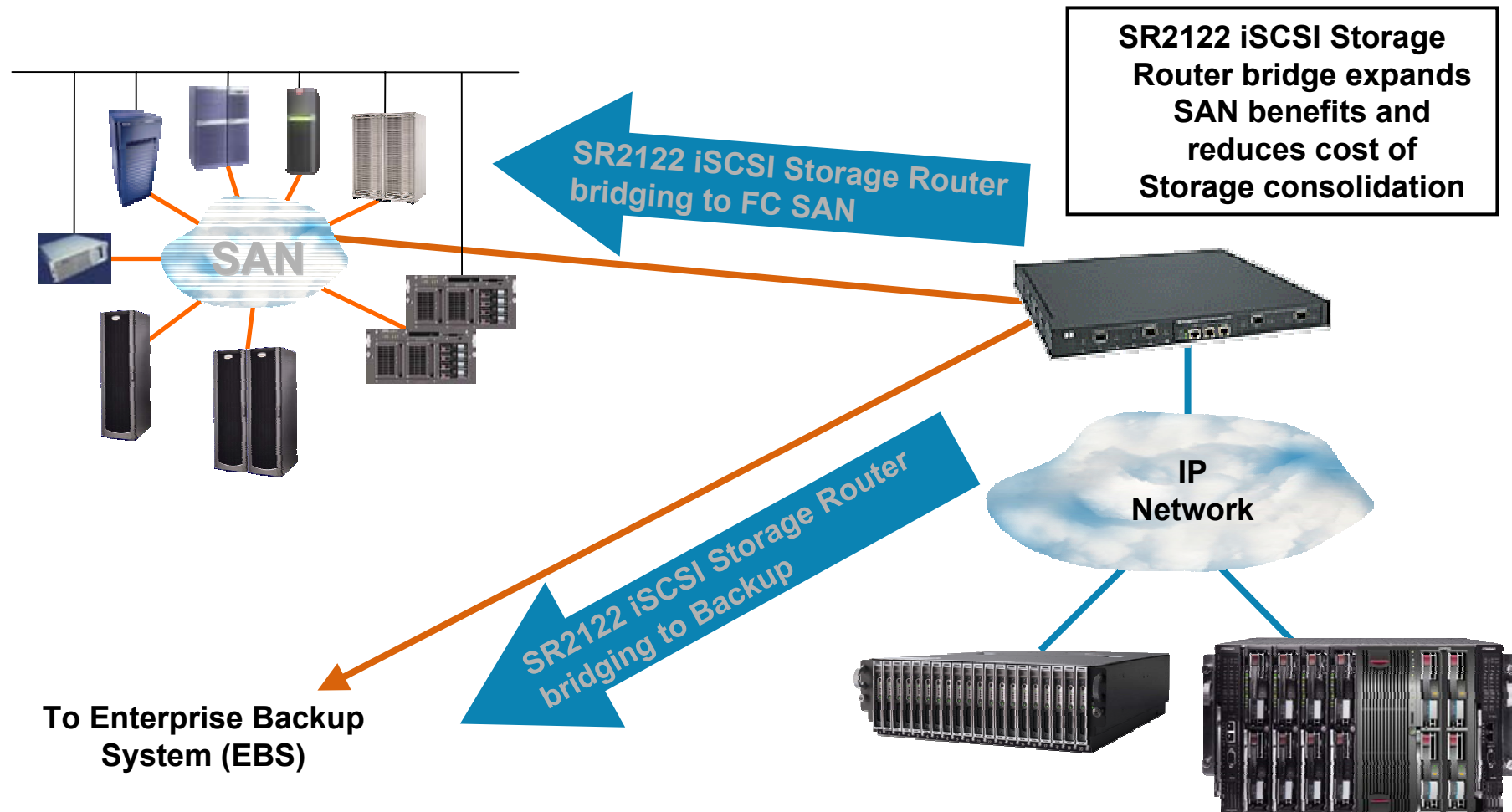
HP StorageWorks SR2122 iSCSI Storage Router


Proven iSCSI bridge to fibre channel solution

- Logically transparent
  - Servers see SCSI disks
  - Applications see blocks
- Up to 140 MBs Full Duplex
- HA Configuration Supported

- **Flexible:** supports two Ethernet and two Fibre Channel connections

- **Affordable:** delivers access to the SAN without hardware reconfiguration and at potentially lower cost

- **Integrated:** fully tested to hp standards to work with key subsystems

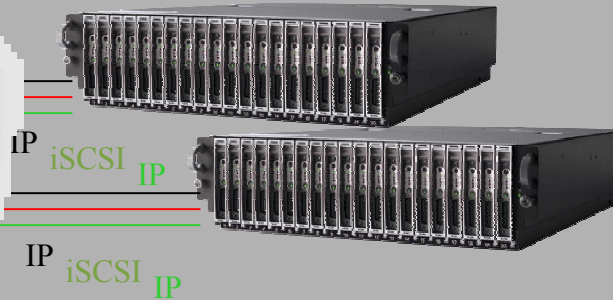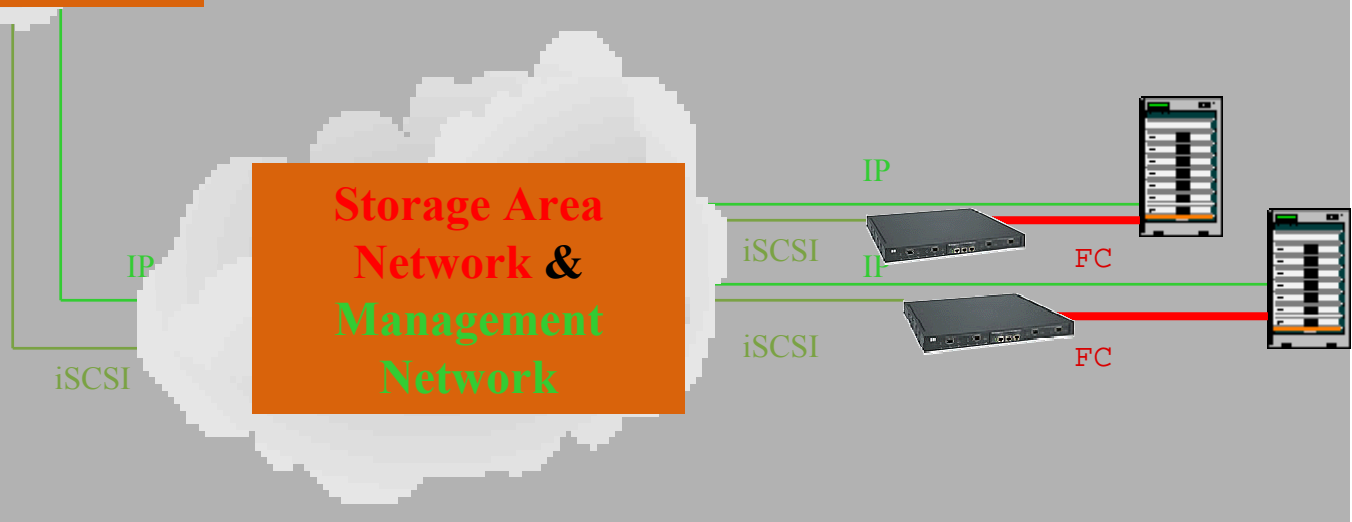# iSCSI SAN Isolated from General IP Network

Routing Architecture
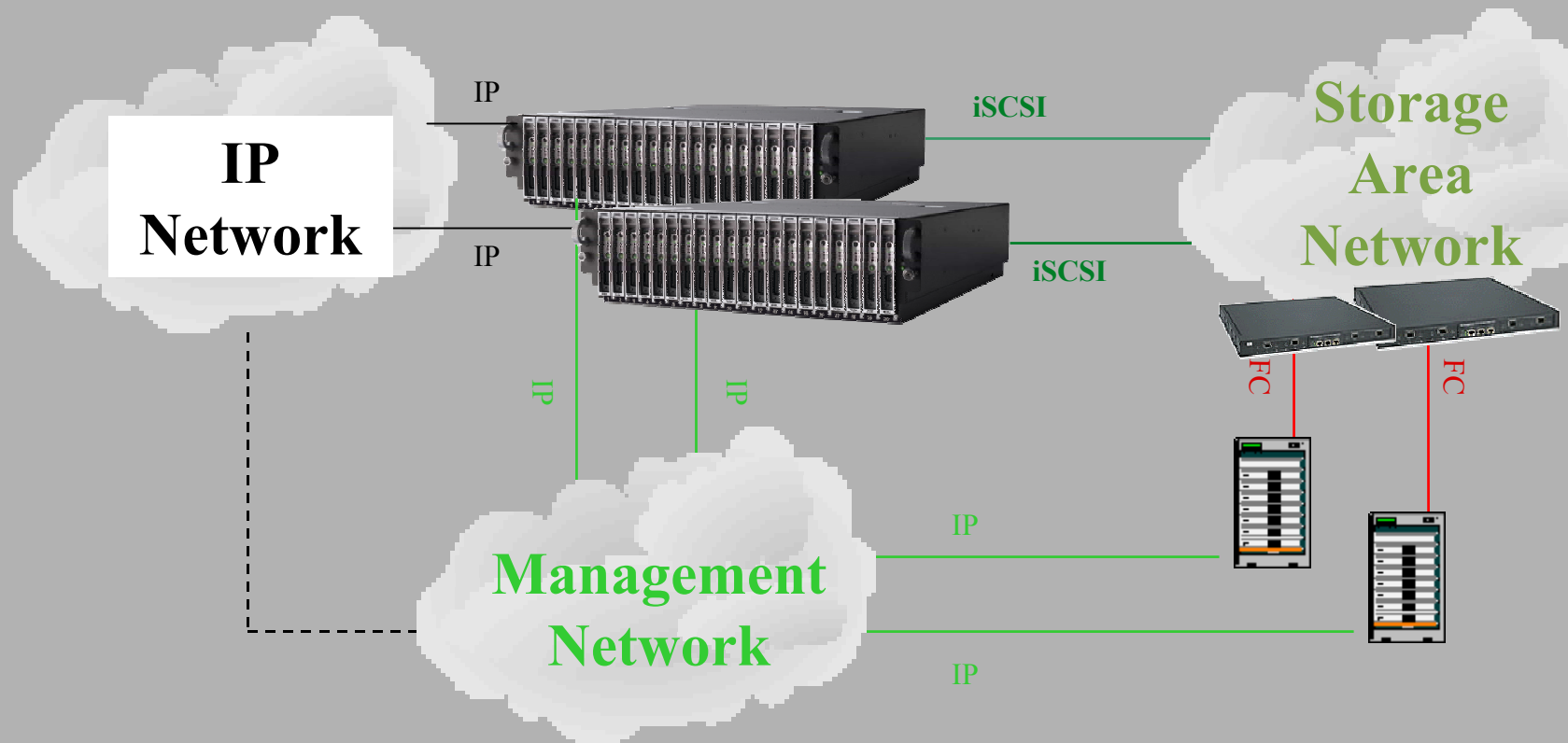
iSCSI SAN Grafted onto General IP Network

Routing Architecture
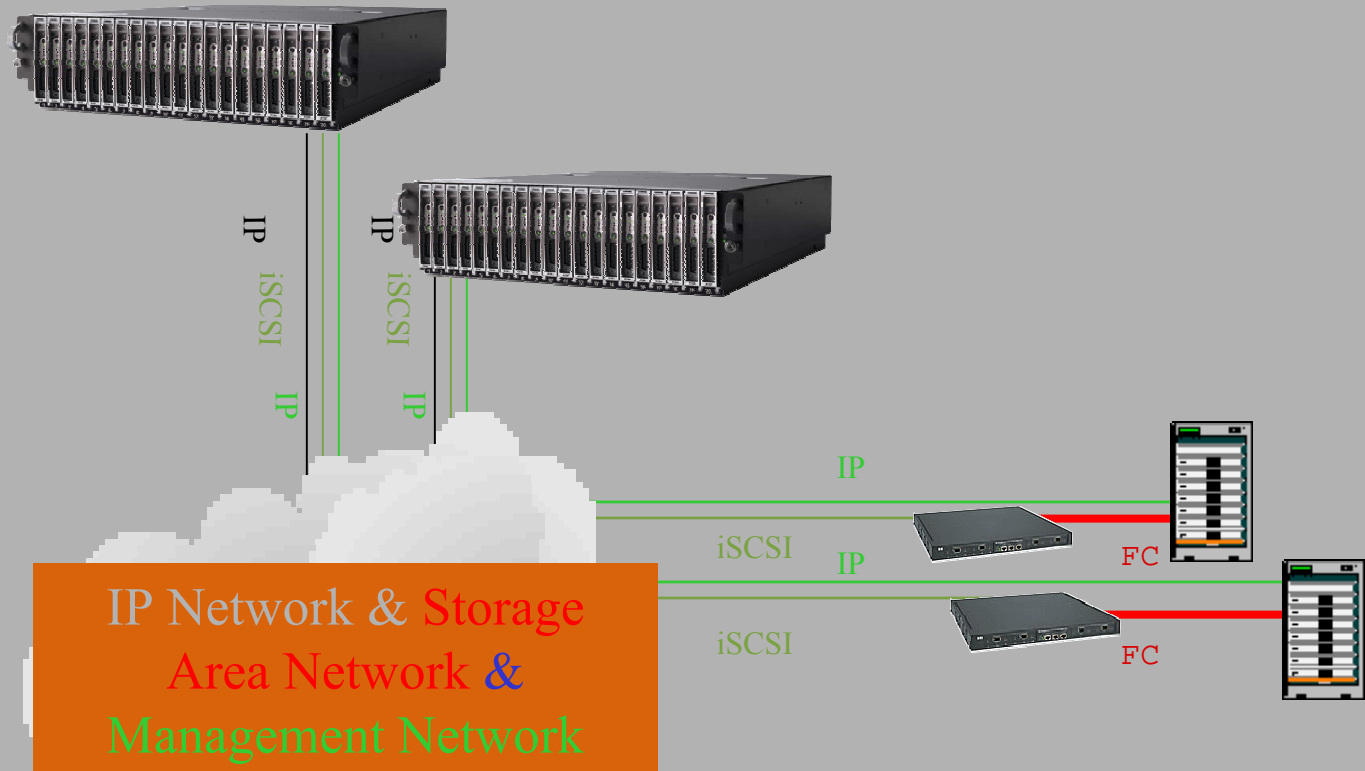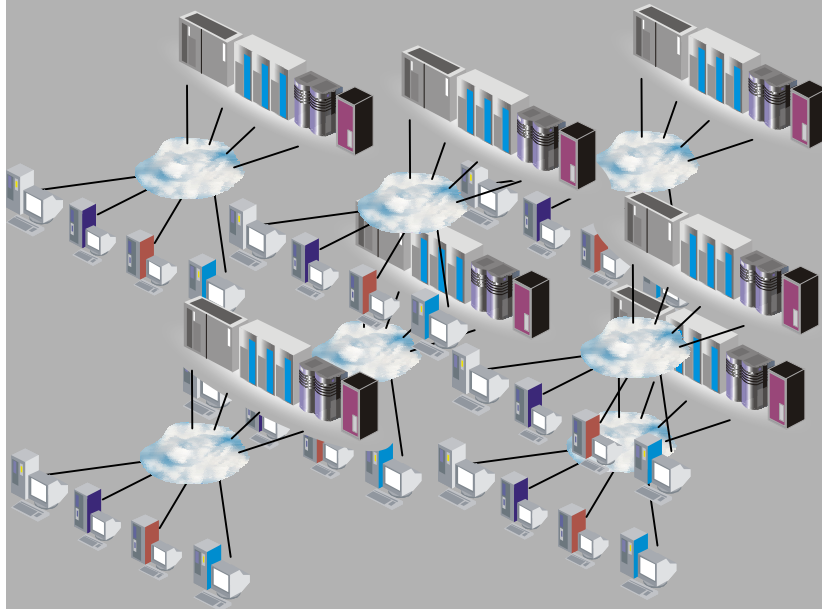
IP Network &
Storage Area
Network &
Management
Network

IP   iSCSI   IP

IP   iSCSI   IP

Storage Area
Network &
Management
Network

IP

iSCSI

iSCSI

IP

iSCSI   IP

iSCSI   II

iSCSI

FC

FC

iSCSI SAN Overlapping General IP Network

Routing Architecture

IP

IP

iSCSI

iSCSI

IP

IP

IP

iSCSI   IP

iSCSI

FC

FC

IP Network & Storage Area Network & Management Network

- desire to leverage network storage efficiencies leads to greater numbers of SANs
  - needed: efficient, manageable ways to connect storage islands
- one size does not fit all
  - needed: wider range of network storage cost/performance options
- accommodating technology change
  - needed: architectural support for adding new technologies

IP Interconnect Convergence

- providing a unified IP fabric for servers, storage and networking

fabrics

storage

IP Fabric Performance with RDMA/TCP
- **end-to-end efficiency for practical scaling**
- **broad vendor adoption**
- **breakthrough economics**

iSCSI (Storage over IP)
- **storage and network over a unified infrastructure**
- **geographic flexibility and disaster recovery**

remote management

Lights-out Management (KVM over IP)
- **remote server management and virtual presence**
- **out-of-band, hardware based solutions**

networking today

Pervasive Datacenter Network
- **existing infrastructure**
- **established IT knowledge base**
- **tested and proven technology**
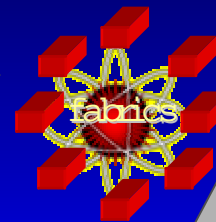
Fabric Strategy & Customer Benefits

Fabric strategy – Drive fabric enhancements to expand infrastructure capabilities, increase solution performance and enable platform innovations with a focus on Ethernet/IP

Ethernet pervasiveness

- Volume & cost leadership
- Knowledge & experience
- Existing, proven tools

Key fact: Operational savings & lower acquisition cost

1 Gbit / 10 Gbit

Simplify infrastructure

- Consolidate ports
- Converge functions
- Centralize management

Key fact: Reduce switch ports & save up to $1000 each

KVM / iSCSI

Just better networking

- Lower CPU utilization, free memory bandwidth
- Increased overall solution performance

Key fact: Low 10 Gbit Ethernet utilization w/o RDMA/TOE

TOE / RDMA

Over time … a unified fabric infrastructure

How RDMA/TCP Technology Works

Fast and secure communications

- remote direct memory access (RDMA) provides efficient memory to memory transfers between systems
  - much less CPU intervention needed
  - true "zero copy" between systems, data placed directly in final destination
  - makes CPU available for other tasks
- maintains current, robust memory protection semantics
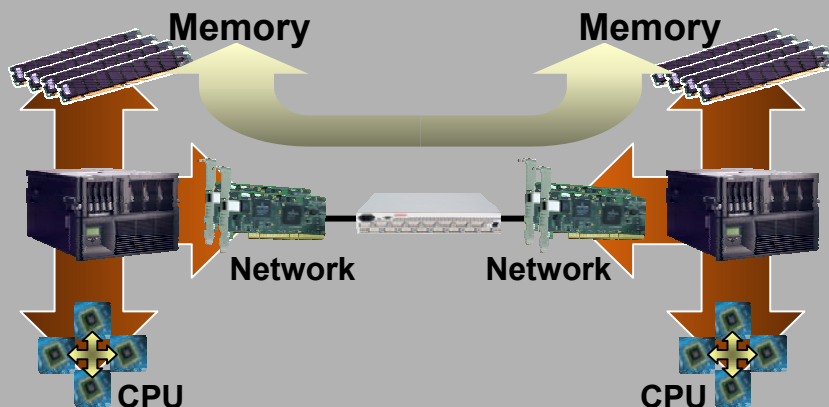
Protocol overhead & efficiency

- TCP/IP protocol overhead an increasing percent of CPU & memory workload
  - ~1Ghz CPU power per Gb/s
  - memory bandwidth ~3x wire speed
- 3 options to solve the problem:
  - faster CPUs (paced by Moore's Law)
  - move work to the NIC (TCP/IP offload)
  - create a more efficient, compatible protocol (RDMA/TCP)

Applications

Operating System

Network (TCP/IP)   Storage (FC, iSCSI, etc.)

RDMA Fast Path

IB   IP

SDP, SRP, NFS, DAFS

iSCSI & Proliant Feedback

Please ask questions or share comments!

Sample Questions

- HP Technical Marketing
    1. Why use iSCSI and not NFS/SMB/CIFS/FC SAN?
    2. Cost for implementation compared to FC SAN?
    3. Additional advantages over other techniques (correlates to answer from Q1)
    4. What OS can access iSCSI?
    5. Which apps will benefit from iSCSI

Sample Questions

- HP Senior SAN Technologist

  1. Do you see the eventual convergence of you storage network and your regular network? Or is there value in having a special network dedicated to storage?

  2. Why would you want to use iSCSI instead of NAS (NFS/CIFS)? If NAS offered the performance and cost of iSCSI, would it be preferable or not?

  3. Moving off of Fibre Channel to go to iSCSI will require quite a bit of work and a lot of new equipment. Is it worth it?

  4. When do you think that iSCSI will be as "solid" as Fibre Channel, with the same features and performance for dollar), so that it could be used to solve the same sorts of problems. For example, when would you feel comfortable using an iSCSI storage network to run your financial system?

Sample Questions

- HP Senior Post-sales support Specialist
  1. Do you think you will employ iSCSI in your environment
  2. For which environment do you see iSCSI fit
  3. Do you think it is a supplement or replacement for NAS
  4. If you plan to employ iSCSI, what timeframe
  5. On which platform do you see iSCSI fit most

Sample Questions

- HP Technical Director

1. Would you plan to use iSCSI to consolidate networking infrastructure by physically combining storage and local area networks?

2. In your view, does iSCSI need 10 Gbit network infrastructure in order to be competitive with FC?

3. As you transition to iSCSI will you train storage administrators to manage IP networks, or would you ask existing IP network administrators to perform this function? Would you train IP network administrators to manage storage?

4. What percentage of the cost of managing FC networks is related to network infrastructure as opposed to other storage specific work?

5. Which FC networking problem offers the most compelling opportunity for iSCSI; (lower cost infrastructure, lower cost of management, or better interoperability?

Questions?

Interex, Encompass and HP bring you a powerful new HP World.