Mac Enterprise - Apple Products in Scientific Computing

Key Topic in HEP: Storage/Storage Access

Agenda

- Introduction

  Apple delivers ...
  - serving Compute Cluster & Storage

  HEP Computing SW/HW Examples
  - OSX and the LHC Experiments
  - HEP Software – LHC, Linear Coll., Grid, JetWeb
- Storage Implementation
  - Critical Components
  - Storage Planning
  - Storage architecture knowledge

    RAID, DAS, NAS, SAN
- Why SAN?
- Apple Storage Products
  - SAN setup (even mixed platform)
  - HW/SW for initial system ~100TB – (€)
  - sufficient for us?
- Q & A

Scientific Computing

Apple delivers:

Portables
64-bit Desktop Systems
high-performance compute clusters and storage

Full-featured PowerBook G4 – portable UNIX solution

Power Mac G5 – 64-bit performance desktop system

Server — rack-optimised Xserve G5
(fully-featured standard and cost-efficient cluster configurations)

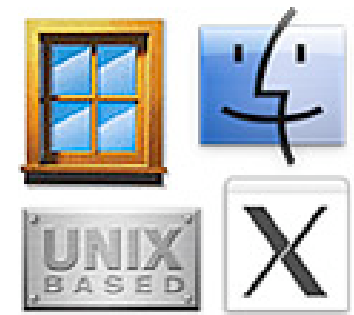
Xserve RAID — high-capacity storage solution offering up to 7TB of high-performance storage

Scientific Computing

Apple delivers:

UNIX-based operating system
- **MacOSX based on Darwin (Open Source) complete UNIX environment (X11 and POSIX services - like for Linux/FreeBSD)**
- **Developer tools - allowing cross-platform development**

Only on Mac you can run native
- **Mac application, UNIX, Windows, X11 and command line applications at the same time**



The Mac simplifies and streamlines IT support
- **MacOSX-Server, Apple Remote Desktop**

Compute Clusters & Storage

Server — rack-optimised Xserve G5 (fully-featured standard and cost-efficient cluster configurations)

Supercomputer:
— Virginia Tech University
— COLSA Coll.
MACH5 Multiple Advanced Computers for Hypersonics

MACH5 - Apple XServe, 2.0 GHz, Myrinet

System Summary

System Name	MACH5
Site	COLSA
System Family	NOW - PowerPC
System Model	XServe Cluster
Computer	Apple XServe, 2.0 GHz, Myrinet
Vendor	Self-made
Application area	CFD
Main Memory	5376 GB
Installation Year	2005

Performance/Linpack Data

Processors	Rmax(GFlops)	Rpeak(GFlops)	Nmax	Nhalf
3072	16180	24576	750000	160000

System

Overview

- System Details
- Performance/Linpack Data
- Ranking History

Sites with a similar system

- University of Illinois
- Bowie State University

Place 15

System X - 1100 Dual 2.3 GHz Apple XServe/Mellanox Infiniband 4X/Cisco GigE

System Summary

System Name	System X
Site	Virginia Tech
System Family	NOW - PowerPC
System Model	XServe Cluster
Computer	1100 Dual 2.3 GHz Apple XServe/Mellanox Infiniband 4X/Cisco GigE
Vendor	Self-made
URL	http://www.tcf.vt.edu/systemX....
Application area	Not Specified
Installation Year	2004

Performance/Linpack Data

Processors	Rmax(GFlops)	Rpeak(GFlops)	Nmax	Nhalf
2200	12250	20240	620000	0

System

Overview

- System Details
- Performance/Linpack Data
- Ranking History

Place 20

Super Computing

HEP Computing - Example

MANCHESTER 1824

The University of Manchester

Storage Space 26 TB

14 TB Apple Xraid Servers

5 TB Babar	3ware
5 TB D0	3ware
1 TB Users	3ware

MANCHESTER 1824

The University of Manchester

Farm 3 Apple G5

No need for Linux

This is new 7 X-server with 14 TB Disks
We are using it to run simulation
work have started testing Atlas and D0
simulation and analysis

27/04/05 Sabah Salih

http://hepwww.rl.ac.uk/sysman/april2005/talks/manchester.ppt

MANCHESTER 1824

The University of Manchester

OS X and the experiments

- LHC experiments and LCG show recent development activities on OS X
  Little effort by single individuals (even on spare time)
  - ALICE: the whole offline software runs on
  - OS X CMS: Iguana, Orca, initial port of external libraries
  - ATLAS: CMT, Atlantis, Atlas distribution kit (based on CMT & Pacman), single packages of the Athena framework
  - LHCb: Gaudi, Panoramix, Da Vinci
  - LCG: SEAL, PI, externals libraries on Xserve

Sabah@hep.man.ac.uk

http://hepwww.rl.ac.uk/sysman/Nov2004/talks/apple-hepsysman.ppt

MANCHESTER 1824

The University of Manchester

HEP Software:

1- General

- a- CERN Library
- b- Root
- d- Geant4
- e- LHAPDF "Les Houches Accord Parton Density Function"
- f- Hdcay
- g- Pythia
- h- ktjet (Code by Manchester)
- i- ExHuME (Monte Carlo Generator Code by Manchester)
- j-CPX+ (CPX SUSY code by Manchester)
- k-pomwig (Monte Carlo Generator Code by Manchester)
- l- **HERWIG**

2- Experiment specific

See next page

Sabah@hep.man.ac.uk

http://hepwww.rl.ac.uk/sysman/Nov2004/talks/apple-hepsysman.ppt

MANCHESTER 1824

The University of Manchester

CLHEP, CERNLIB

On Fink - (OpenSource)

http://fink.sourceforge.net

Geant4

From SLAC and on Fink

http://geant4.slac.stanford.edu/g4cd/July2004/Documentation/WorkshopExercises/Prerequisities/MacOSX.html

ROOT

From ROOT (since long time) and on Fink

Sabah@hep.man.ac.uk

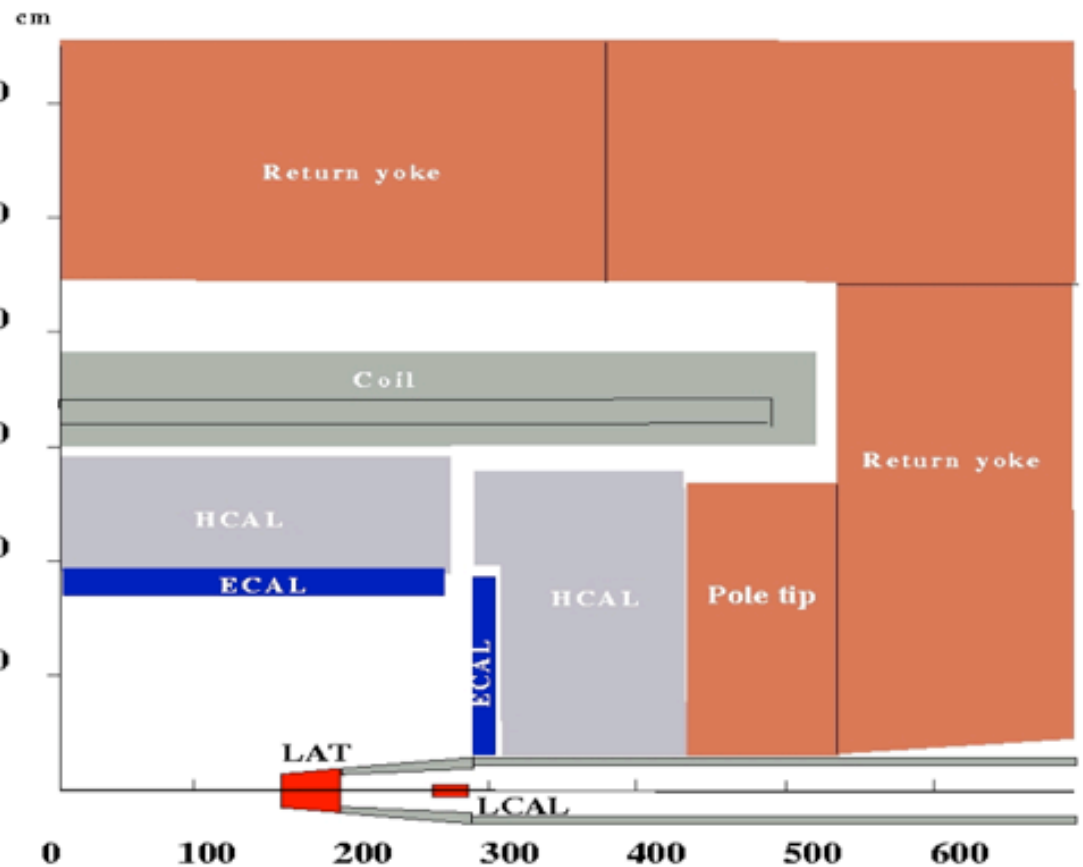
http://hepwww.rl.ac.uk/sysman/Nov2004/talks/apple-hepsysman.ppt

Linear Collider

-

- SIMDET V.4.0 -

-A Fast Simulation Tool for Linear Collider Detector Studies



Sabah@hep.man.ac.uk

http://hepwww.rl.ac.uk/sysman/Nov2004/talks/apple-hepsysman.ppt

MANCHESTER 1824

The University of Manchester

Grid Tools

Globus, pyGlobus, Globus Job Manager, Condor, LSCdataFind, ...

http://www.lsc-group.phys.uwm.edu/~duncan/computing/mac_info.html

Built-in "gridification"

Xgrid, Technology Preview 2

free d/l from Apple - Grid-ready architecture

Sabah@hep.man.ac.uk

http://hepwww.rl.ac.uk/sysman/Nov2004/talks/apple-hepsysman.ppt

MANCHESTER 1824

The University of Manchester

```
sh-2.05b# /usr/local/bin/rpm -i edg-voms-vo-wp6-0.0.1-1.noarch.rpm
      package edg-voms-vo-wp6-0.0.1-1 is intended for a linux operating system
sh-2.05b# /usr/local/bin/rpm -i edg-voms-vo-iteam-0.0.1-1.noarch.rpm
      package edg-voms-vo-iteam-0.0.1-1 is intended for a linux operating system

sh-2.05b# /usr/local/bin/rpm -i ca_UKeScience-0.18-1.noarch.rpm
      package ca_UKeScience-0.18-1 is intended for a Linux operating system


sh-2.05b# rpm2cpio ca_CERN-0.23-1.noarch.rpm | cpio -ivd
./etc/grid-security/certificates/fa3af1d7.0
./etc/grid-security/certificates/fa3af1d7.crl_url
./etc/grid-security/certificates/fa3af1d7.signing_policy
0 blocks
sh-2.05b# ls etc/
grid-security
sh-2.05b# ls etc/grid-security/
certificates
sh-2.05b# ls etc/grid-security/certificates/
fa3af1d7.0          fa3af1d7.crl_url      fa3af1d7.signing_policy
sh-2.05b#
```

Sabah@hep.man.ac.uk

http://hepwww.rl.ac.uk/sysman/Nov2004/talks/apple-hepsysman.ppt

MANCHESTER 1824

The University of Manchester

JetWeb

Automated Data Comparisons for High Energy Physics

18/02/04: This server is currently read-only. See news for details.

best fits, all data

- HERWIG
- PYTHIA

summaries, all fits

- HERWIG latest
- PYTHIA latest
- HERWIG all
- PYTHIA all

documentation, downloads

- Latest News
- Bibliography
- Generator Parameters
- Developer Resources

simulations

- HERWIG
- PYTHIA

experiments

- HERA(H1,ZEUS)
- LEP (OPAL)
- Tevatron (CDF, D0)
- HEPDATA

Search the DataBase **Maintenance**

Selected Results

- Studies for a Future Linear Collider
- Minimum PT of hard scatters
- Intrinsic KT photon/proton
- PYTHIA parton showers PARP67
- Parton Distribution Functions in Photon
- HERWIG Soft Underlying Event
- HERWIG Photon Radius
- HERWIG fragmentation parameters (CLMAX,PSPLT)

If you do use any results from here, please quote Comp. Phys. Comm. vol 153/2 164-178 (2003)

The current focus of this project is on jet and heavy flavour production in hadron-like collisions (which includes hadron-photon and photon-photon). There is no reason why other data shouldn't be incorporated though.

If you'd like join in, or have any comments or suggestions please contact us at jetweb@hep.ucl.ac.uk

The story so far:
6410 jobs submitted to Manchester PBS, 5782 completed
2800 jobs submitted to UCL PBS, 2598 completed
171 jobs submitted to UCL NQS, 142 completed
68 jobs submitted to GridPP, 35 completed
641 jobs submitted to Sheffield PBS, 536 completed

These pages are automatically regenerated frequently as new fits are performed and more data are included. Thanks to Manchester University (especially Brian Cox and Sabah Salih) and UCL (especially Gordon Crone) for CPU.

JetWeb: J. M. Butterworth, S. Butterworth, B. M. Waugh, University College London;
jetweb@hep.ucl.ac.uk
Pages remade 18-Feb-2004 at 11:13:18

Sabah@hep.man.ac.uk

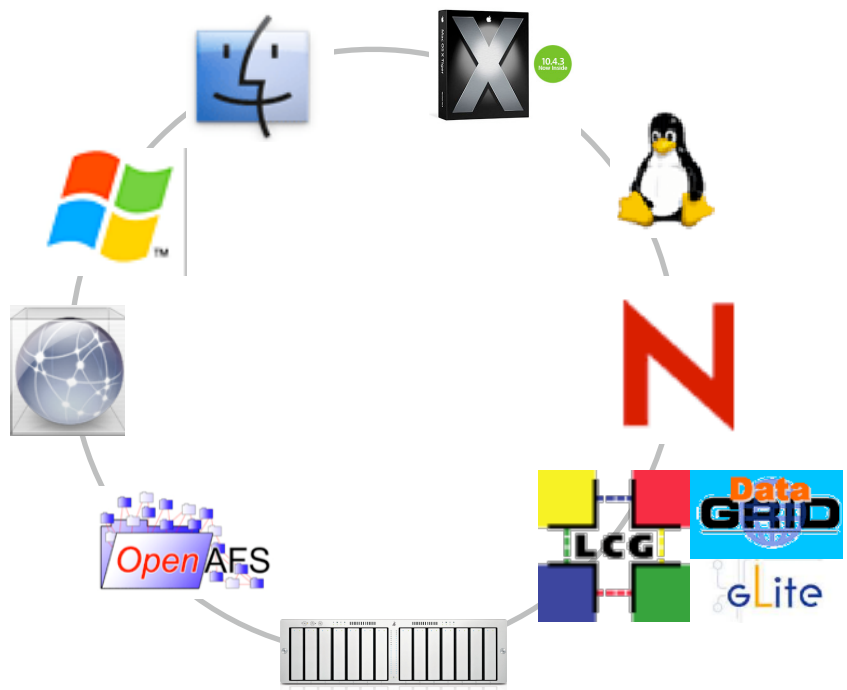
http://jetweb.hep.ucl.ac.uk/

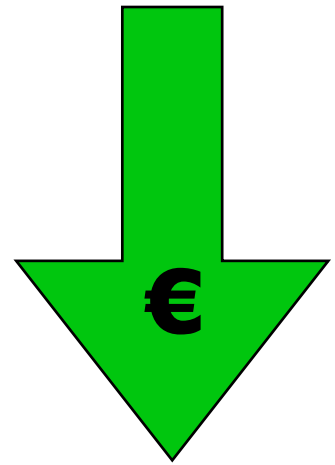
http://hepwww.rl.ac.uk/sysman/Nov2004/talks/apple-hepsysman.ppt

The ideal Storage Implementation!?

Critical Components

- Scalable
- Secure & reliable
- Easy to manage
- High throughput
- Cross-platform
- Affordable





I'm not going to answer all in detail
- discussion among users
- Apple dominated talk ⇒ but I'm open for any good solution

Storage Planning

Proper storage planning must consider

- Existing infrastructure AFS, home directories, ~200 node farm...
  - How old is the current deployment
  - Storage does wear out...
  - Do you already possess management tools
- True capacity requirements - planing for today and tomorrow
- Throughput requirements - Stephan W.: 1-2GB/s now - peak 10GB/s
  - Application-driven
  - IOPS or MB/s
  - Configuration driven
  - Dedicated storage DAS*
  - Network attached storage NAS*
  - Shared storage SAN*
- Availability requirements
- Disaster recovery
  - Remote Replication
  - Doubles the cost—budget?
  - Off-site back up service
  - Lower initial cost—sometimes more labour required
- Budget
  - Complete overhaul or staged deployment
  - Hidden costs
  - Infrastructure
  - Management software
  - Service contracts
  - Expansion costs
  - A trusted vendor

IOPS is the standard unit of measurement for *I/O* (Input/Output) *operations per second*. This measurement is a performance-based measurement and is usually seen written as:

- **Total IOPS**: Average number of I/O operations per second.
- **Read IOPS**: Average number of read I/O operations per second.
- **Write IOPS**: Average number of write I/O operations per second.

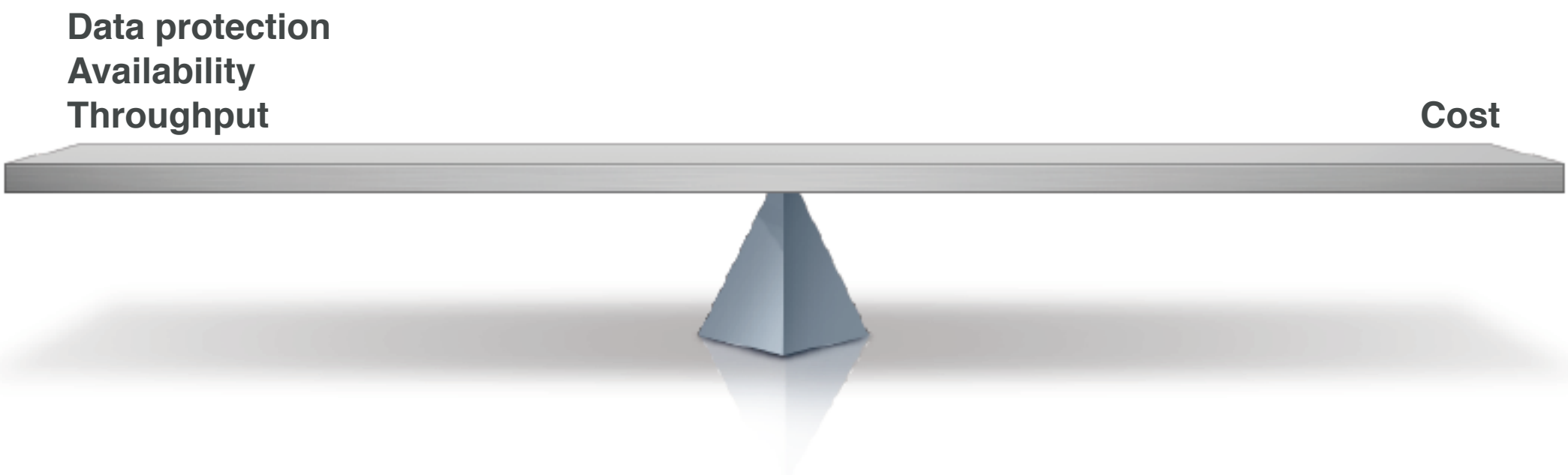
* will be explained later

Some rough ideas - mainly informative

⇓

Building a Smart Storage Strategy

Balancing features and cost based on application requirements



RAID

(rād) Short for ***R**edundant **A**rray of **I**ndependent (or **I**nexpensive) **D**isks,*
a category of disk drives that employ two or more drives in combination for fault tolerance and performance. RAID disk drives are used frequently on servers but aren't generally necessary for personal computers.

There are number of different RAID levels:

Level 0 -- Striped Disk Array without Fault Tolerance: Provides *data striping* (spreading out blocks of each file across multiple disk drives) but no redundancy. This improves performance but does not deliver fault tolerance. If one drive fails then all data in the array is lost.

Level 1 -- Mirroring and Duplexing: Provides disk mirroring. Level 1 provides twice the read transaction rate of single disks and the same write transaction rate as single disks.

Level 2 -- Error-Correcting Coding: Not a typical implementation and rarely used, Level 2 stripes data at the bit level rather than the block level.

Level 3 -- Bit-Interleaved Parity: Provides byte-level striping with a dedicated parity disk. Level 3, which cannot service simultaneous multiple requests, also is rarely used.

Level 4 -- Dedicated Parity Drive: A commonly used implementation of RAID, Level 4 provides block-level striping (like Level 0) with a parity disk. If a data disk fails, the parity data is used to create a replacement disk. A disadvantage to Level 4 is that the parity disk can create write bottlenecks.

Level 5 -- Block Interleaved Distributed Parity: Provides data striping at the byte level and also stripe error correction information. This results in excellent performance and good fault tolerance. Level 5 is one of the most popular implementations of RAID.

Level 6 -- Independent Data Disks with Double Parity: Provides block-level striping with parity data distributed across all disks.

Level 0+1 – A Mirror of Stripes: Not one of the original RAID levels, two RAID 0 stripes are created, and a RAID 1 mirror is created over them. Used for both replicating and sharing data among disks.

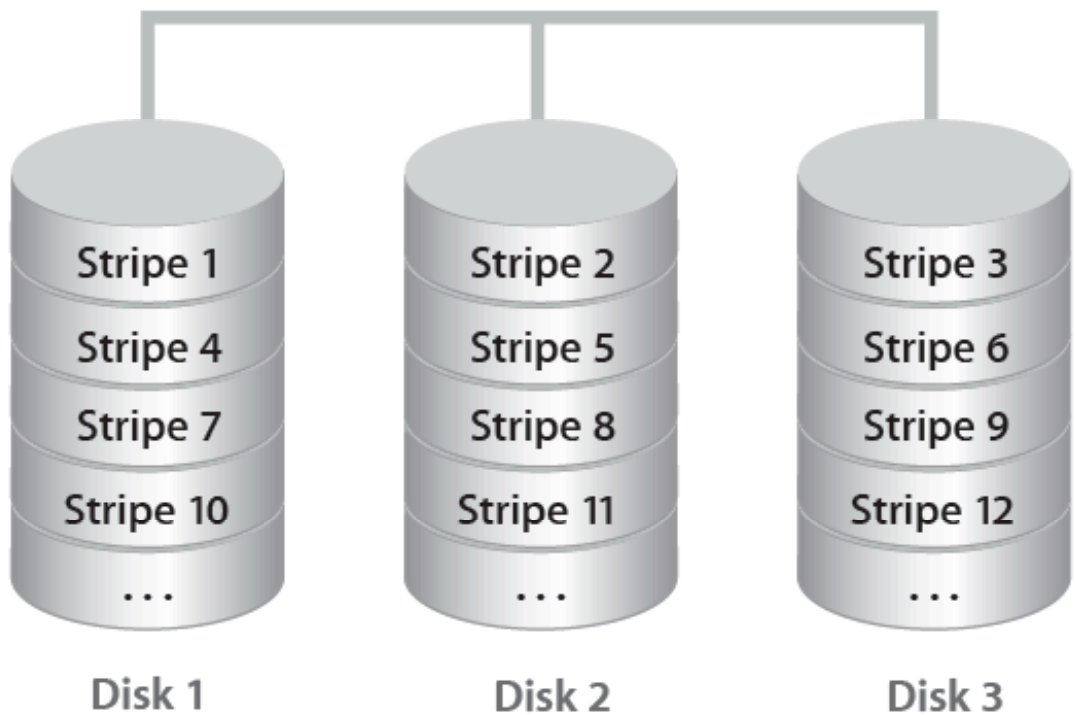
Level 10 – A Stripe of Mirrors: Not one of the original RAID levels, multiple RAID 1 mirrors are created, and a RAID 0 stripe is created over these.

Level 7: A trademark of Storage Computer Corporation that adds caching to Levels 3 or 4.

RAID S: EMC Corporation's proprietary striped parity RAID system used in its Symmetrix storage systems.

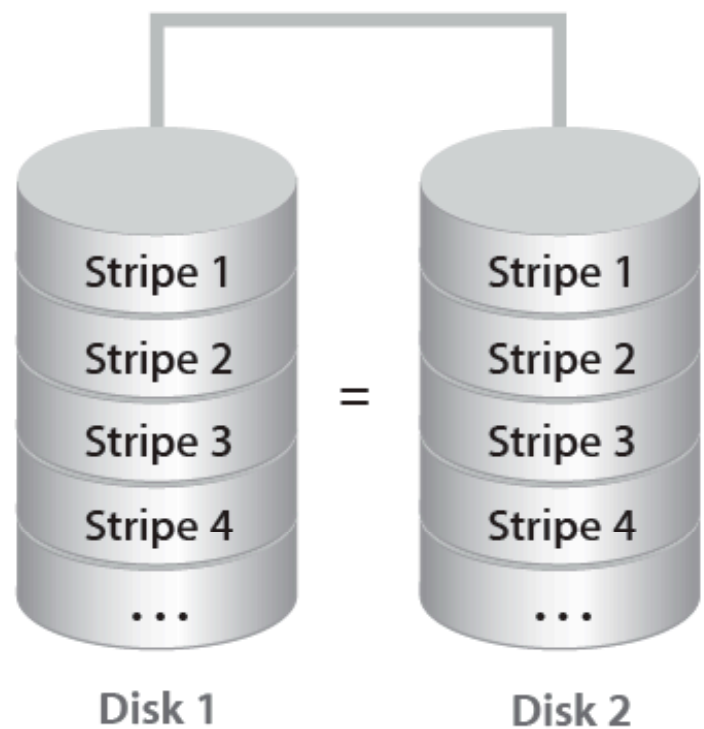
RAID 0: Striping

Distribute data across 2 or more drives



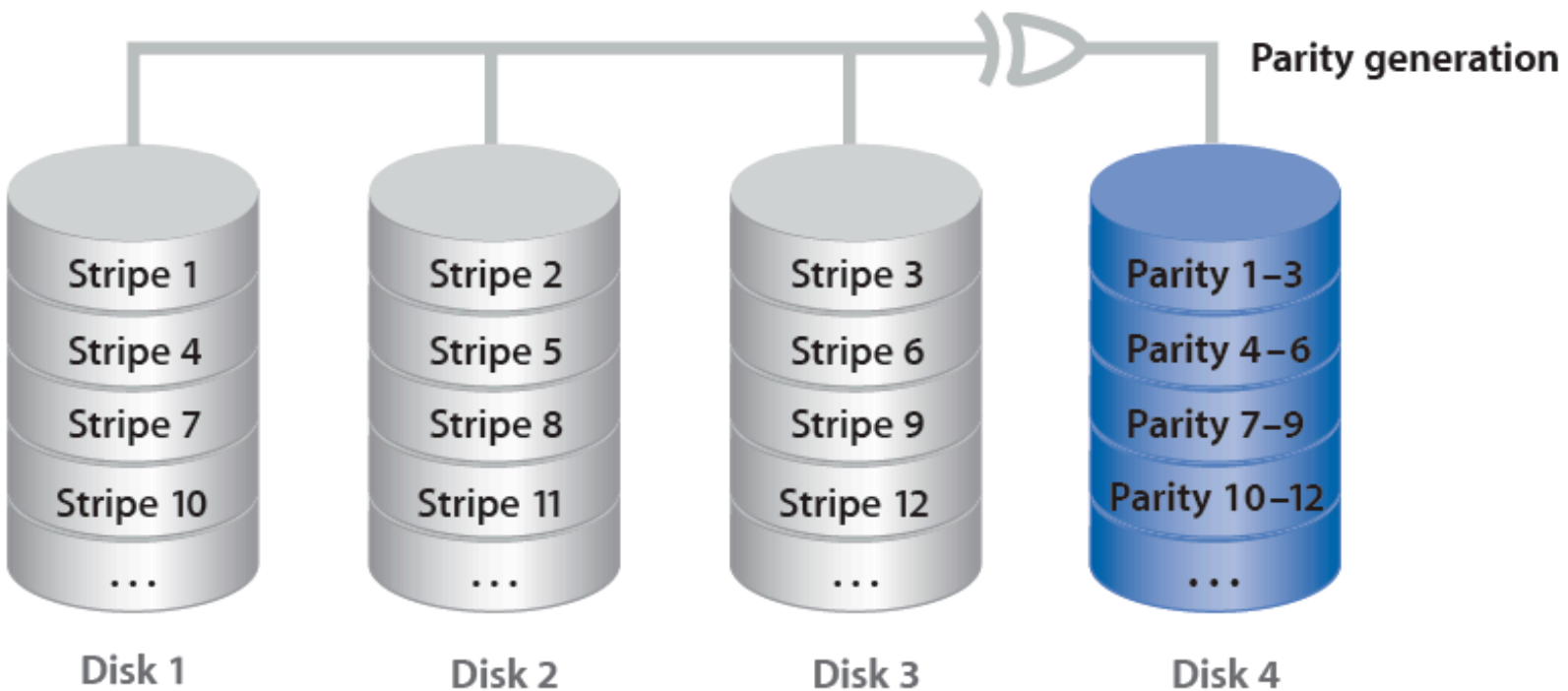
RAID 1: Mirroring

Data duplicated across 2 drives



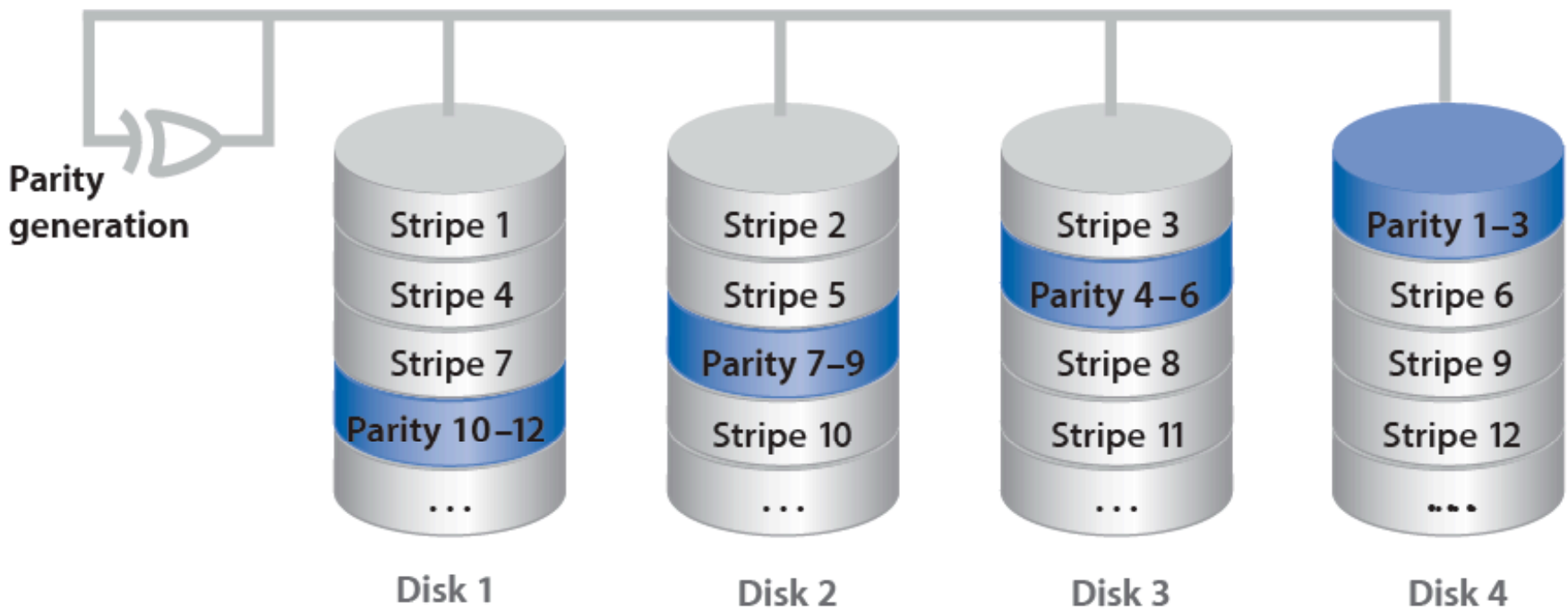
RAID 3: Striping with parity

Strip across 2 or more drives with a dedicated parity drive



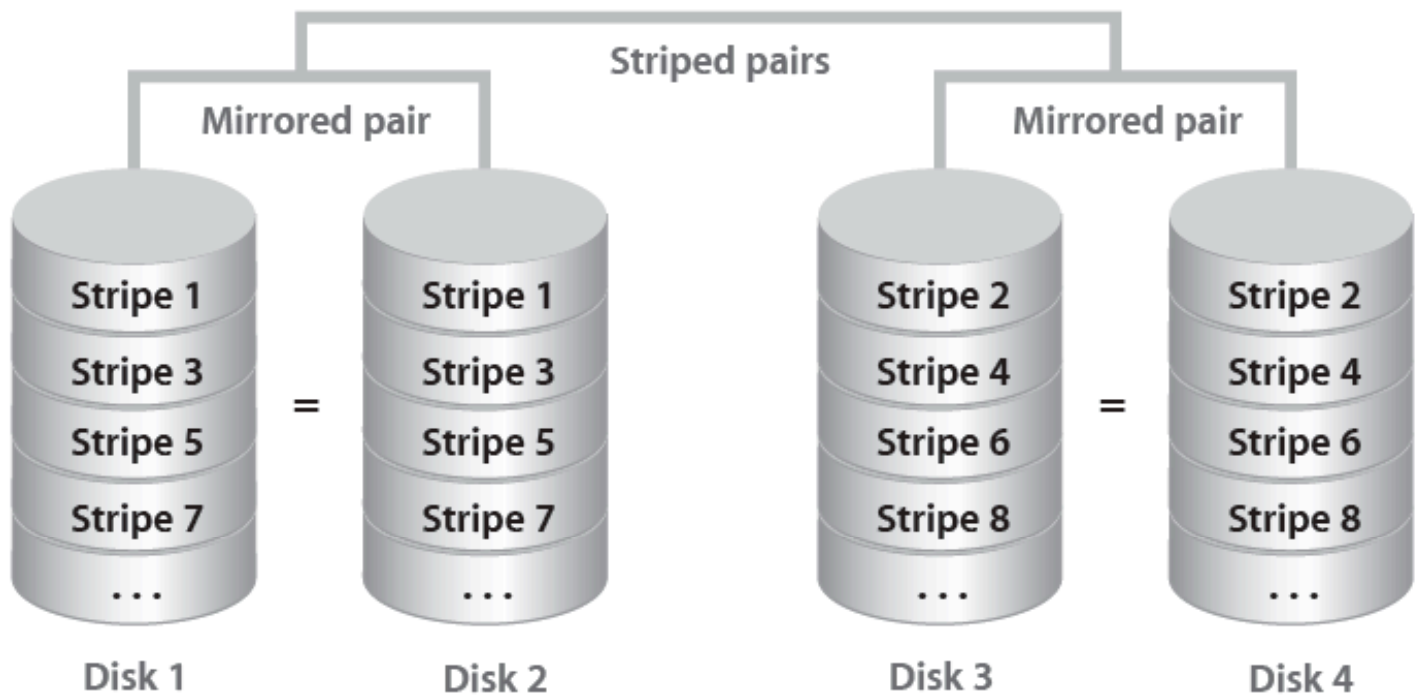
RAID 5: Striping w/ Distributed Parity

Distribute data and parity across all drives



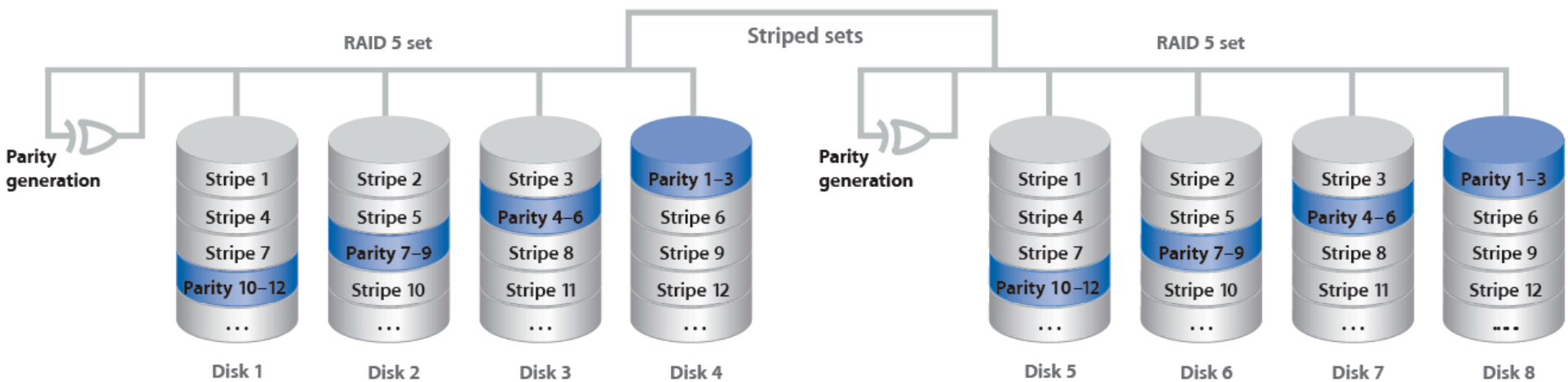
RAID 0+1: Striping over Mirroring

Mirrors drives then stripes data across mirrored set



RAID 10, 30, 50: Hybrid Raid

Using software and hardware RAID to produce hybrid 50



Comparison of RAID levels

Six basic levels

RAID Level	Redundancy	Performance	Capacity	Application	# of Drives
RAID 0—Striping	None	Fast Reads, Fast Writes, High IOPS	Full	Static Data easily replicated—scratch area	2
RAID 1—Mirroring	100% Plus	Good Reads, Single Drive Writes, Low IOPS	50%	Financial Data, Low Usage Archive in smaller data sets	2
RAID 0+1	100%	Good Reads, Good Writes, Medium IOPS	50%	Database, Financial, smaller data sets	3
RAID 3—Parity RAID	100%	Fast Reads, Single Drive Writes, High Read IOPS	100%—1 Drive	Realtime Applications, VOD	3
RAID 5—Distributed Parity	100%	Fastest Reads, Fastest Writes, Highest IOPS	100%—1 Drive	All Applications	4
RAID 10,11,50,51... —Hybrid	Varies	varies with implementation	variable		4

Types of Storage Implementations

DAS	NAS	SAN
Direct Attached Storage	Network Attached Storage	Storage Area Network

DAS in Research Computing

Direct-Attached Storage

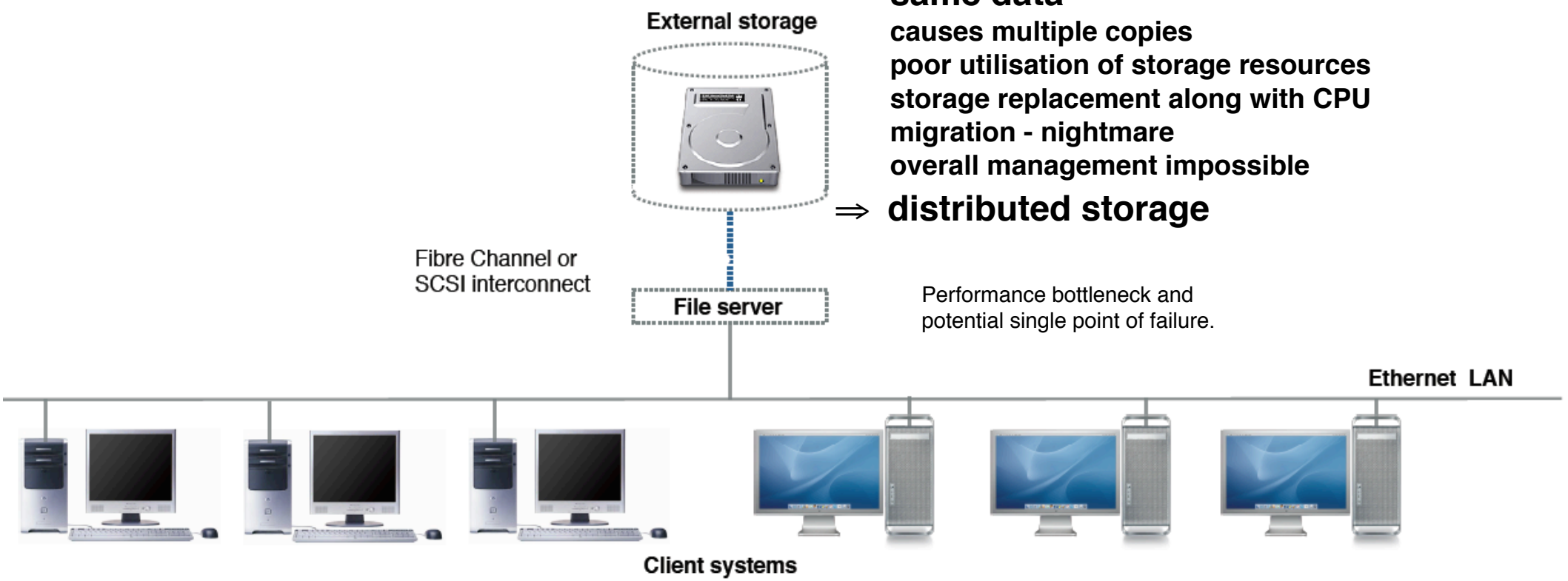
Server is single point of access to storage

Every server has it own dedicated storage

⇒ multiple computer accessing similar or same data

causes multiple copies
poor utilisation of storage resources
storage replacement along with CPU migration - nightmare
overall management impossible

⇒ distributed storage

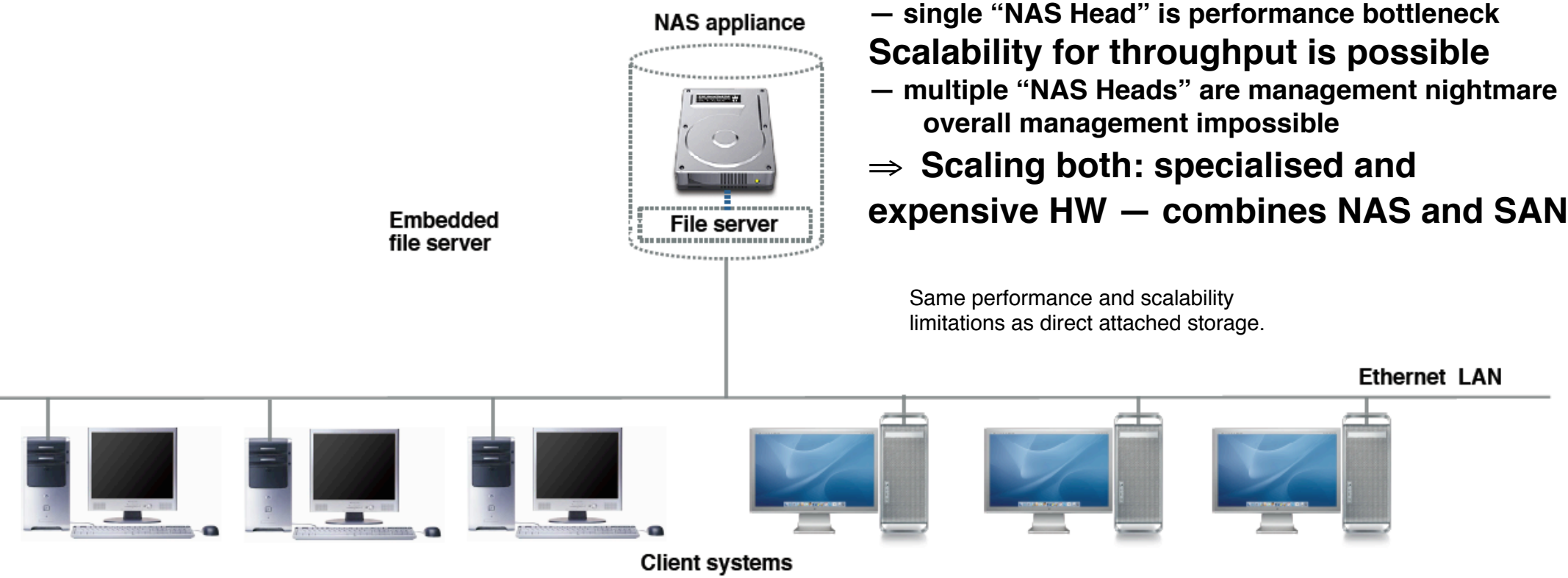


NAS in Research Computing

Network-Attached Storage

Purpose-built shared storage appliance

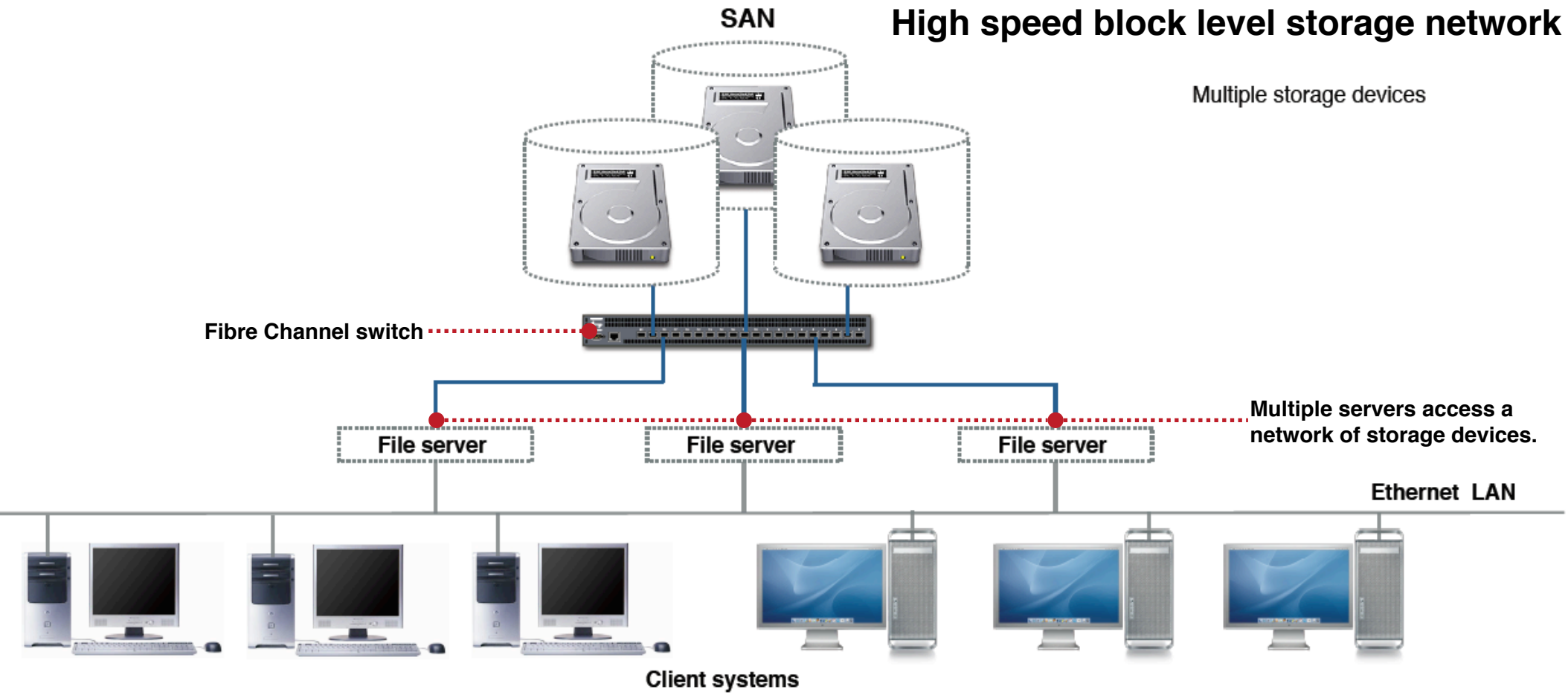
Entry cost for NAS is good
Utilisation of storage resources is good
Scalability for capacity is possible
— single "NAS Head" is performance bottleneck
Scalability for throughput is possible
— multiple "NAS Heads" are management nightmare
overall management impossible
⇒ Scaling both: specialised and expensive HW — combines NAS and SAN



SAN in Research Computing

Storage Area Network

Multiple storage devices in a Fibre Channel network



SAN in Research Computing

Storage Area Network

Multiple computers access the same storage device

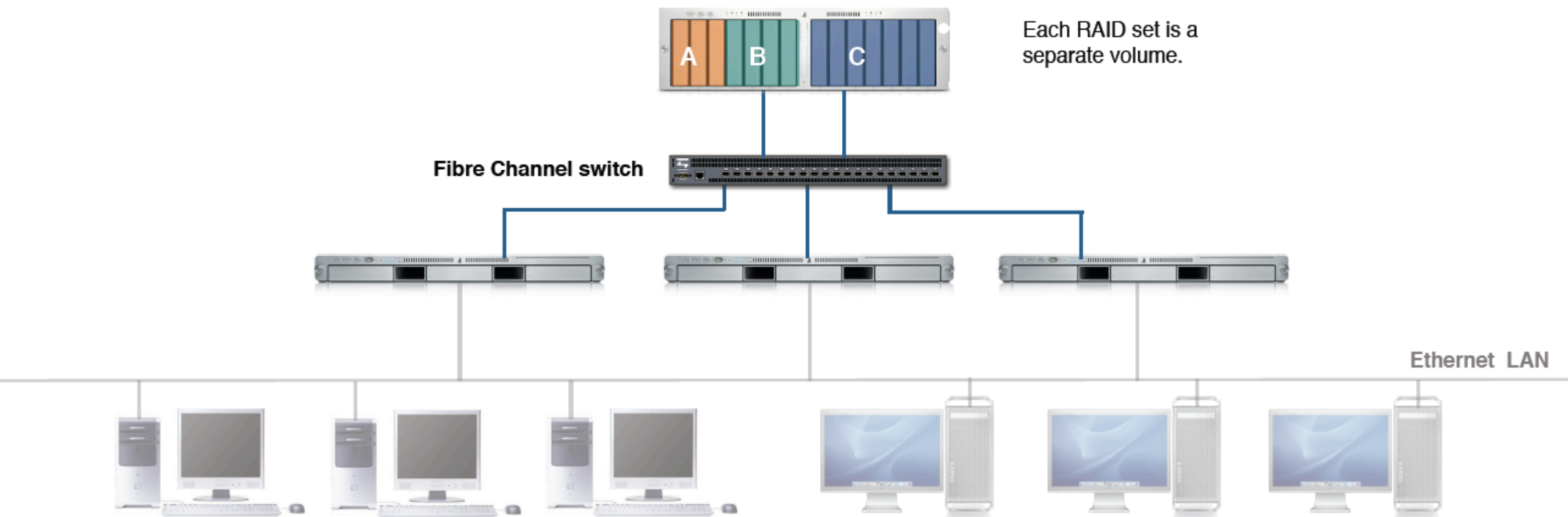
Deployment can scale in capacity & performance - no penalty

Fibre Channel switch

Ethernet LAN

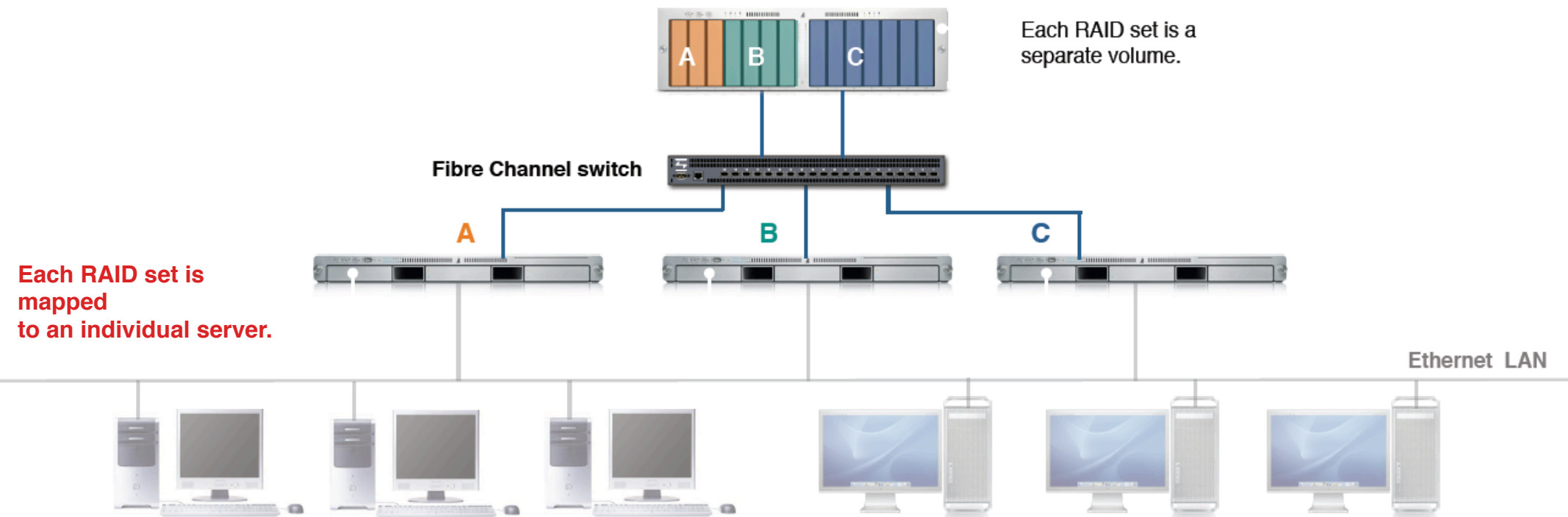
Before SAN File System

Data is separated into storage silos



Before SAN File System

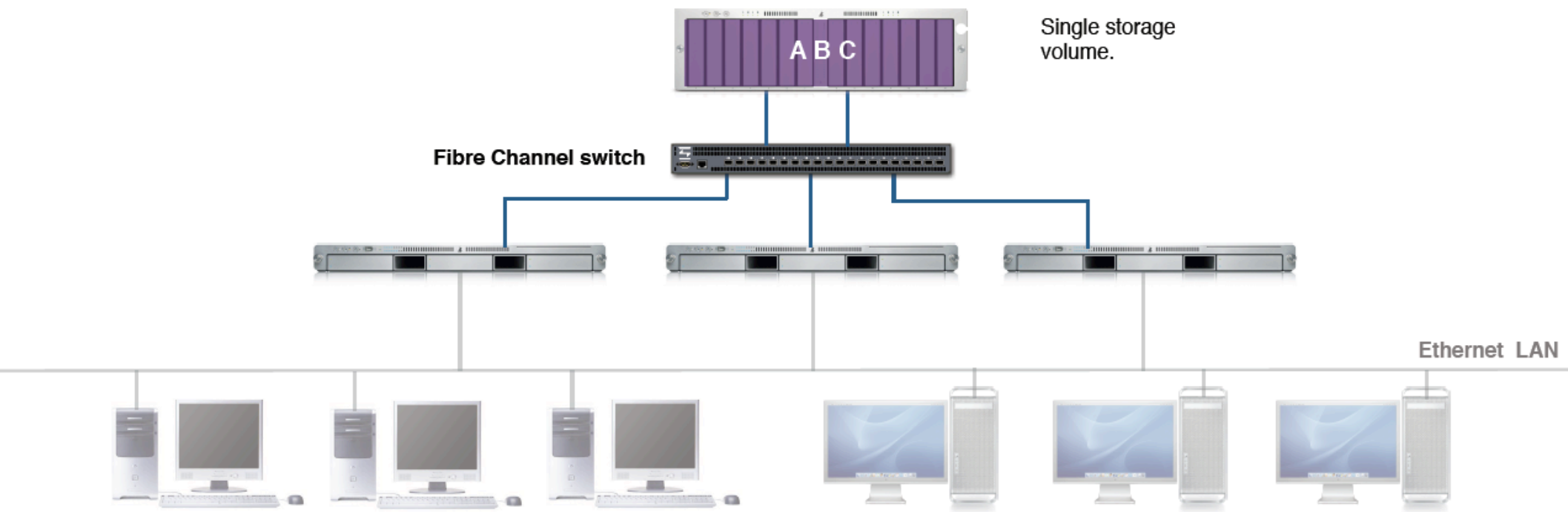
Data is separated into storage silos



After SAN File System

Data is consolidated into a single volume

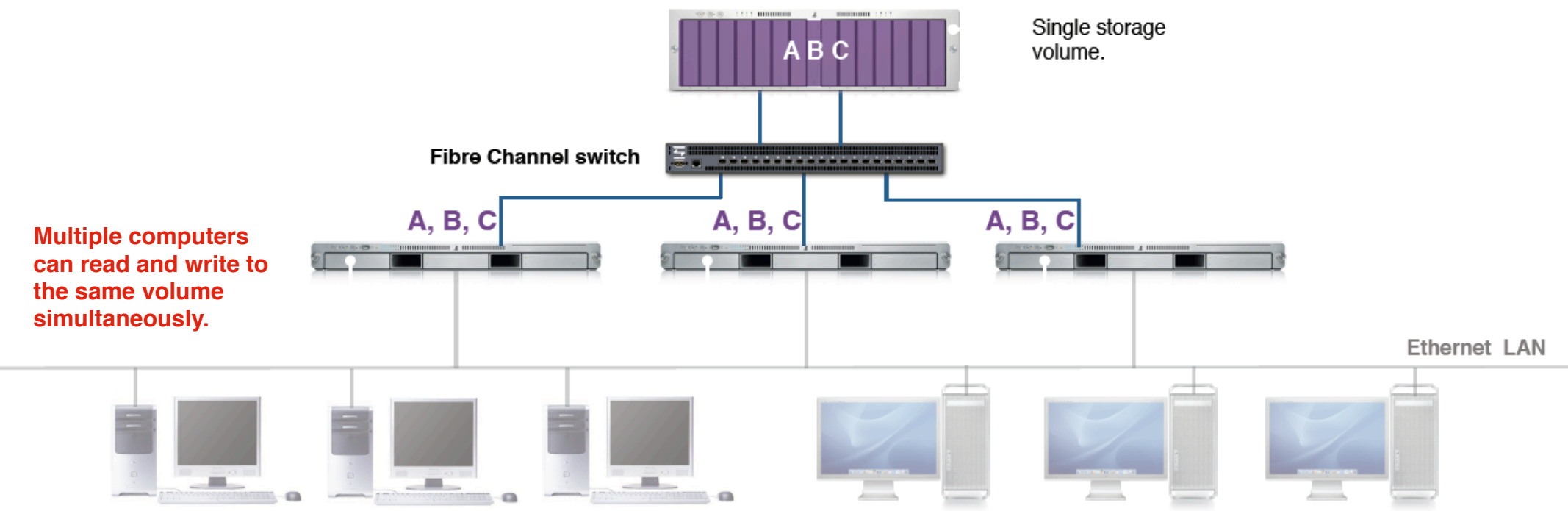
Centralised management allows easier provisioning - easy growth and max. utilisation



After SAN File System

Data is consolidated into a single volume

Centralised management allows easier provisioning - easy growth and max. utilisation

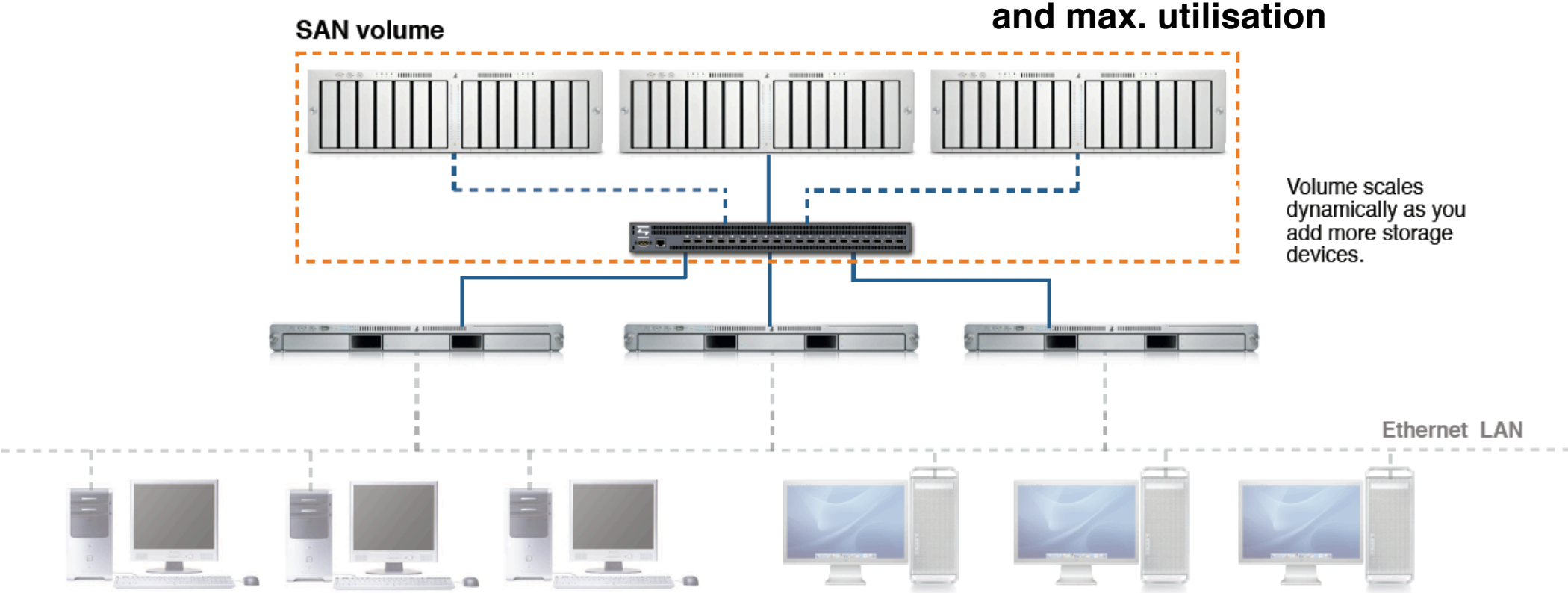


Multiple computers can read and write to the same volume simultaneously.

SAN File System

Flexible topology for easy scalability

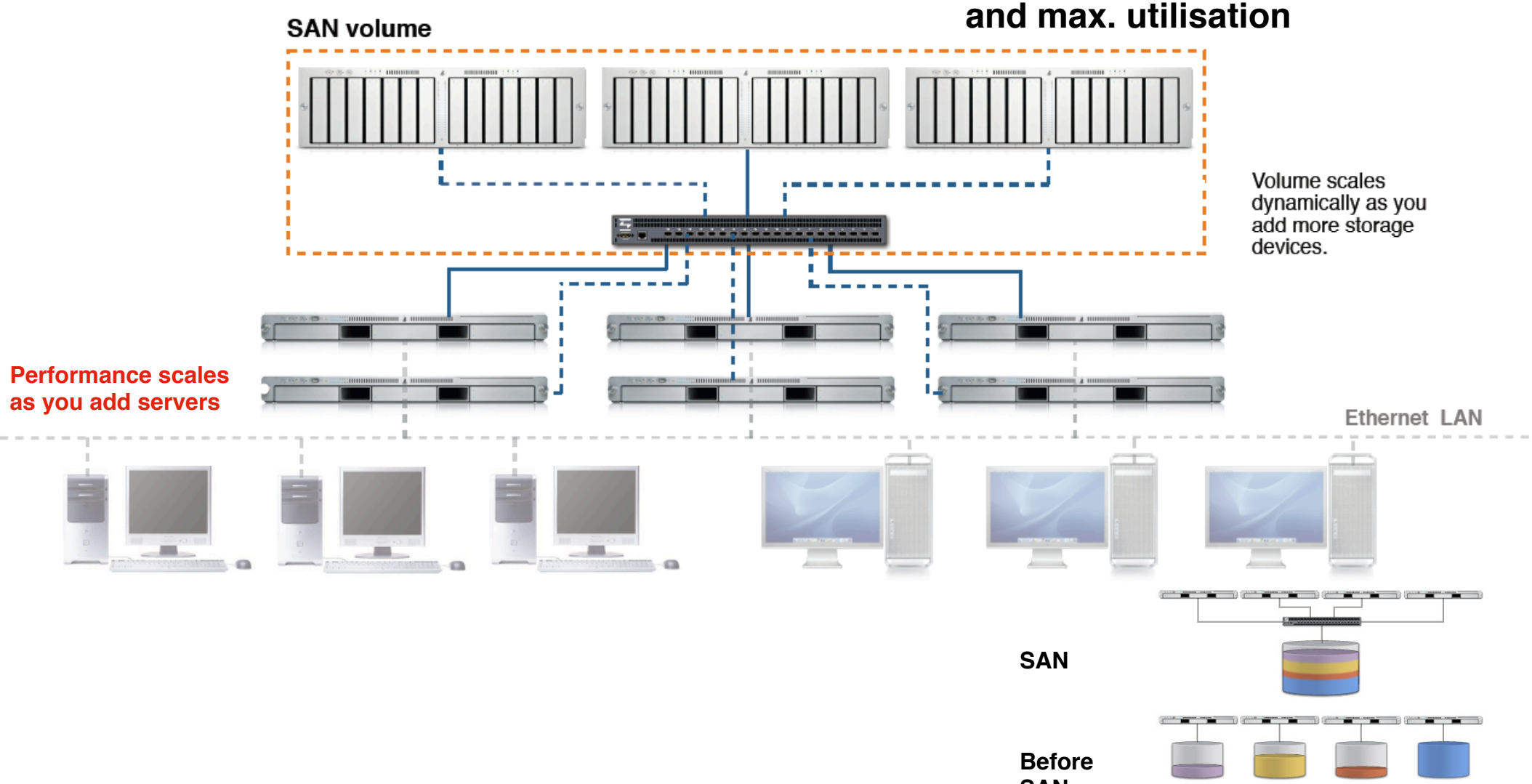
Centralised management allows easier provisioning - easy growth and max. utilisation



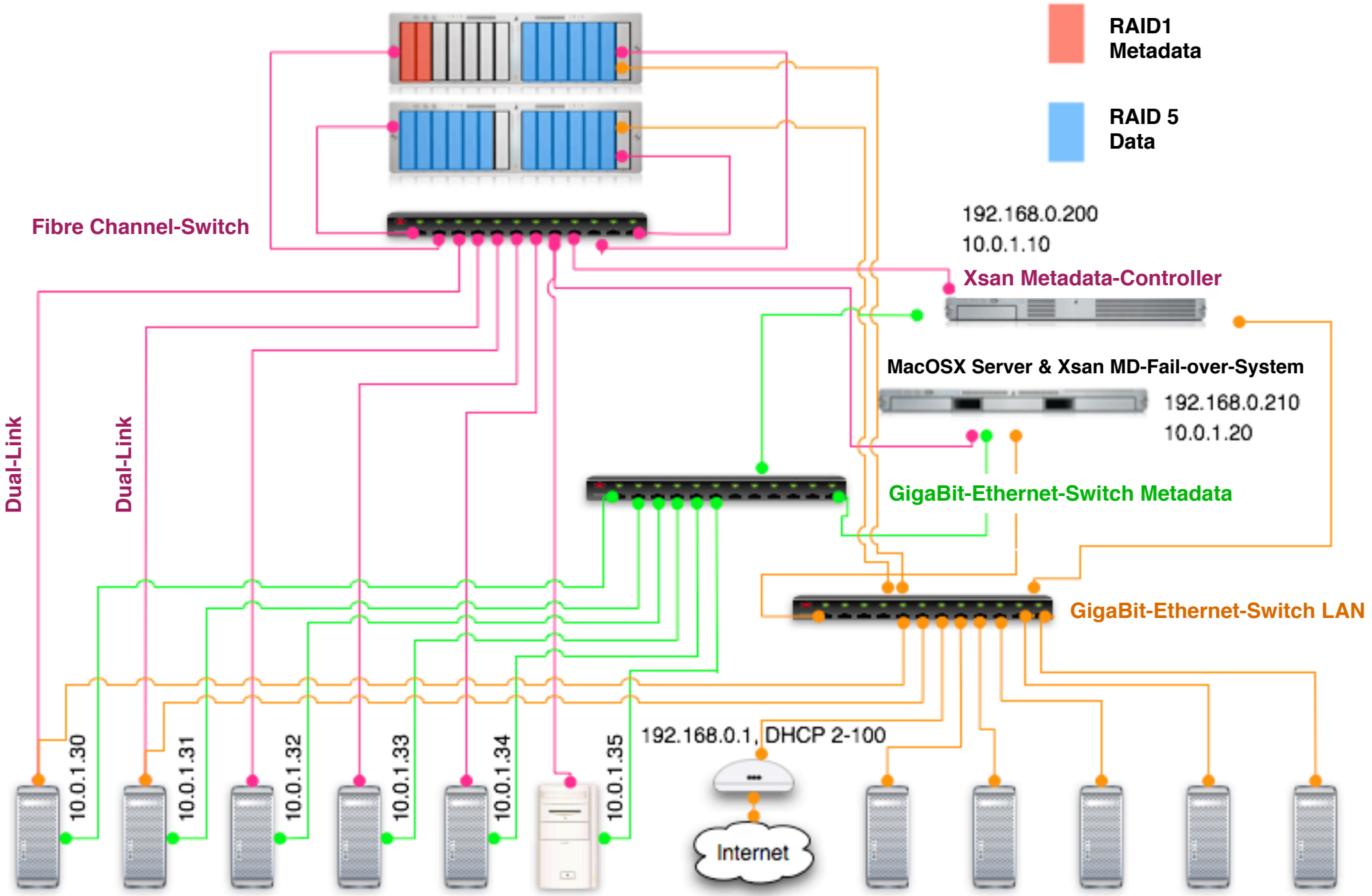
SAN File System

Flexible topology for easy scalability

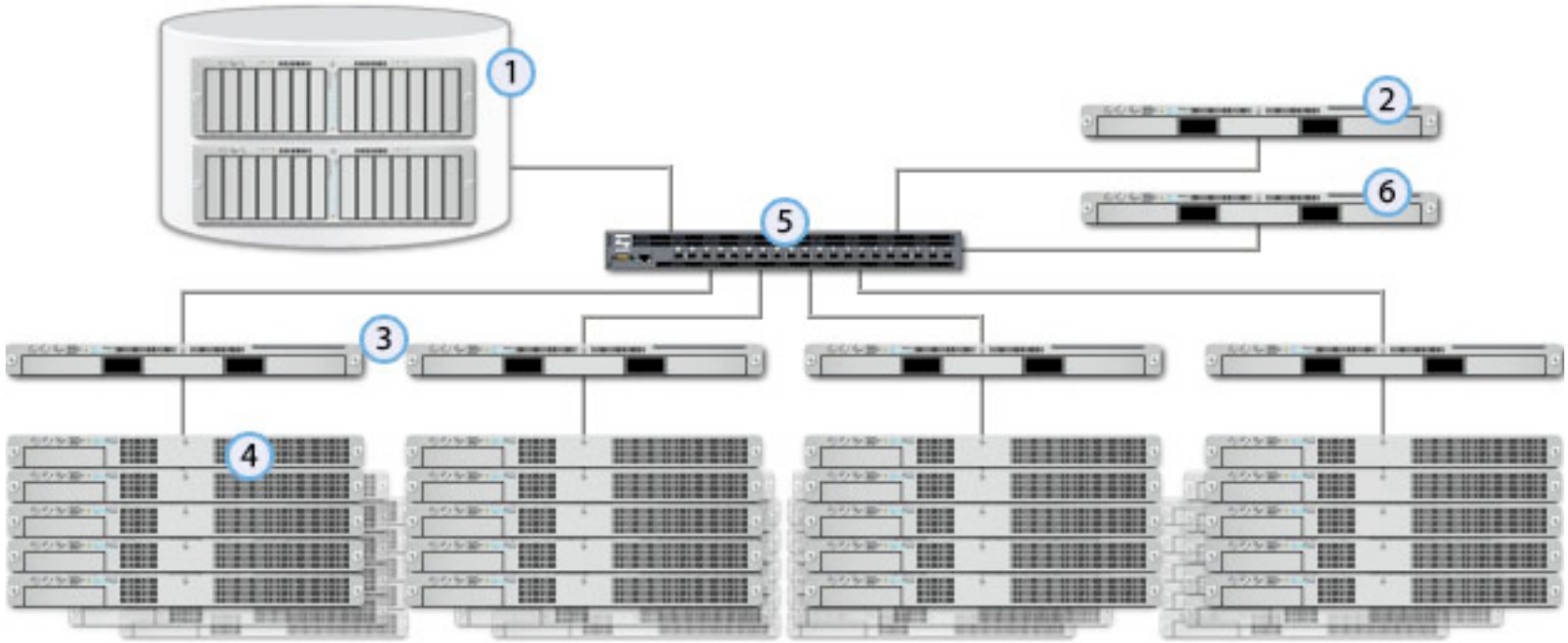
Centralised management allows easier provisioning - easy growth and max. utilisation



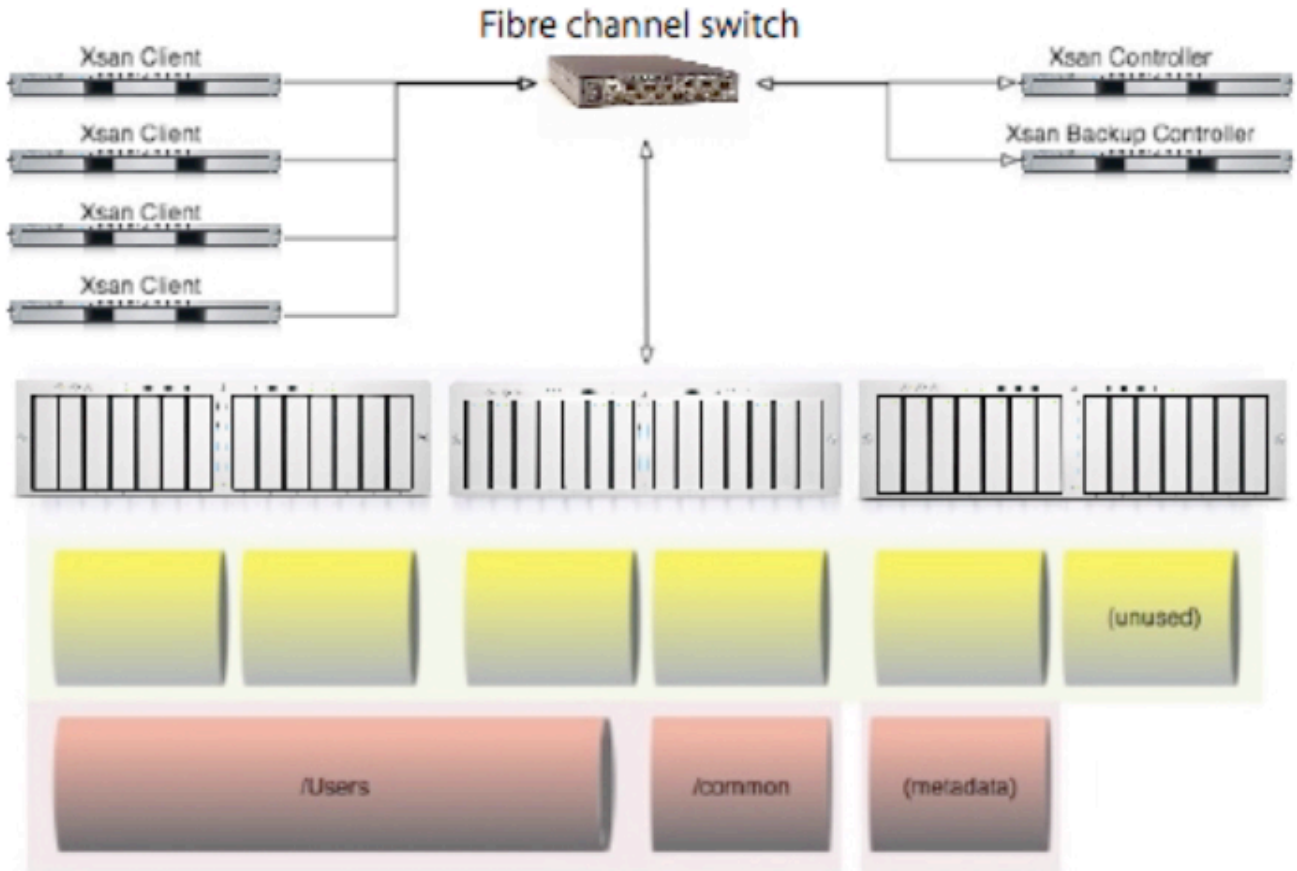
SAN - Fibre Channel & Gb-Ethernet



Storage for large computational cluster.



1	**Xserve RAID Storage Pool.**	Manage data centrally.
2	**Metadata Controller.**	Directs servers to files in pool.
3	**Head Nodes.**	Retrieve data over Fibre Channel.
4	**Client Nodes.**	Process data received from head node – farm nodes
5	**Fibre Channel Switch.**	Gives high-speed access to storage.
6	**Standby Controller.**	Provides high-availability insurance.



Benefits of such a SAN Setup (simplified)

- Metadata isolated to dedicated RAID controller (and ethernet network)
- User space and shared directory space on separate RAID controller
- User space easily expandable
- Xsan controller isolated from NFS server duty
- NFS server expandability

After all, why a SAN File System?

- Access shared volumes concurrently from multiple computers
- Consolidate data for greater efficiency
- Scale capacity and performance as needs grow
- Eliminate the need for storage provisioning separately
  (e.g. H1-Zeuthen: h1zeut0x - many TB on 4 nodes NFS shared)

NAS: simultaneous access solution too expensive

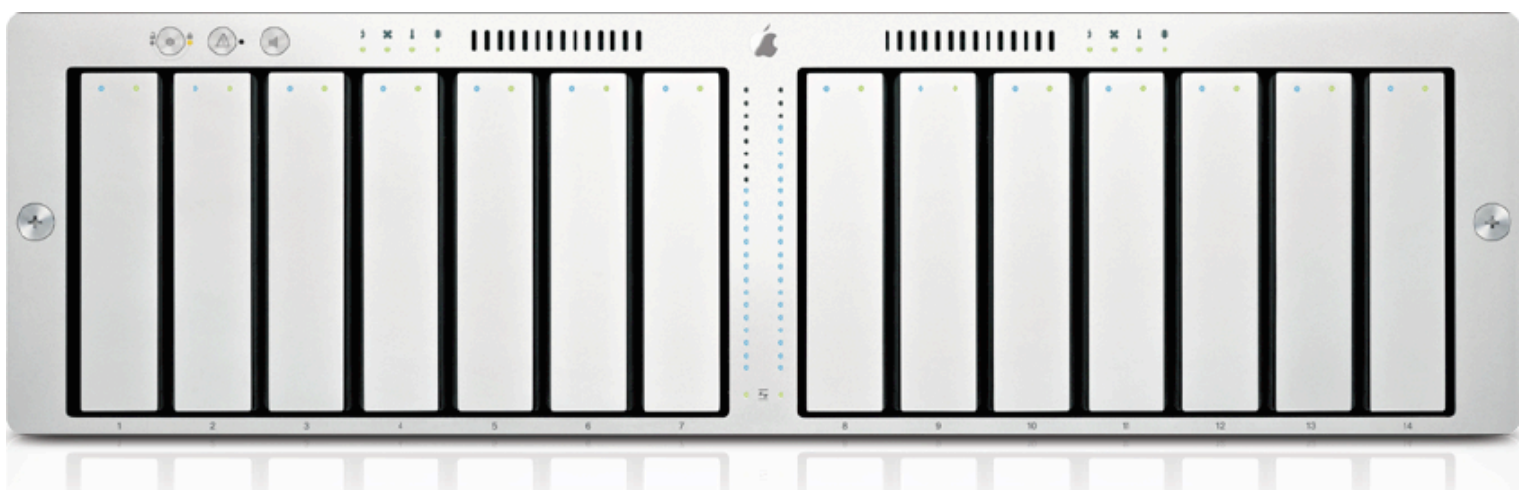
Could serve:

- Home directories (simultaneous access?)(AFS, Windows ...), extended ACL
- Experiment data storage
- dCache
- LCG ready (it seems to be)
  remote access - access/research on the go
- Backup, archives
  storage planning - tiered solution?!
- ...

Storage Capabilities

- Apple introduced Xserve RAID in February 2003
  - Massive storage capacity
  - High-availability design
  - Advanced architecture for data protection and performance
  - Easy setup and remote management

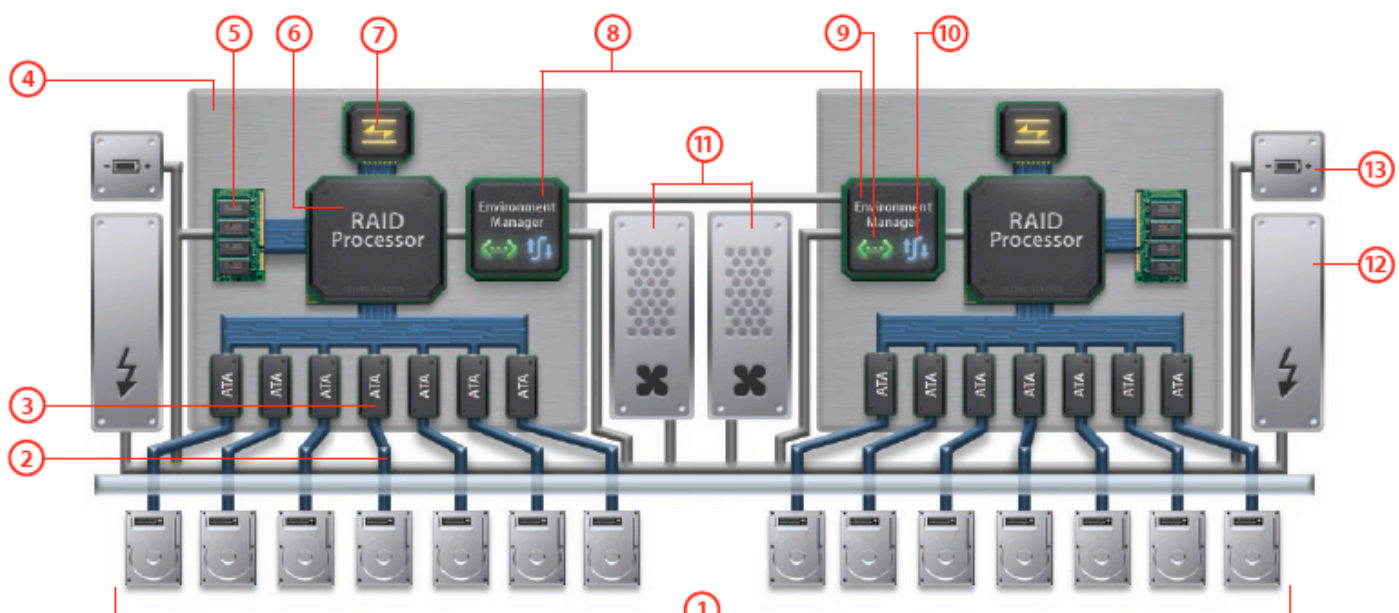
http://www.apple.com/xserve/raid/specs.html



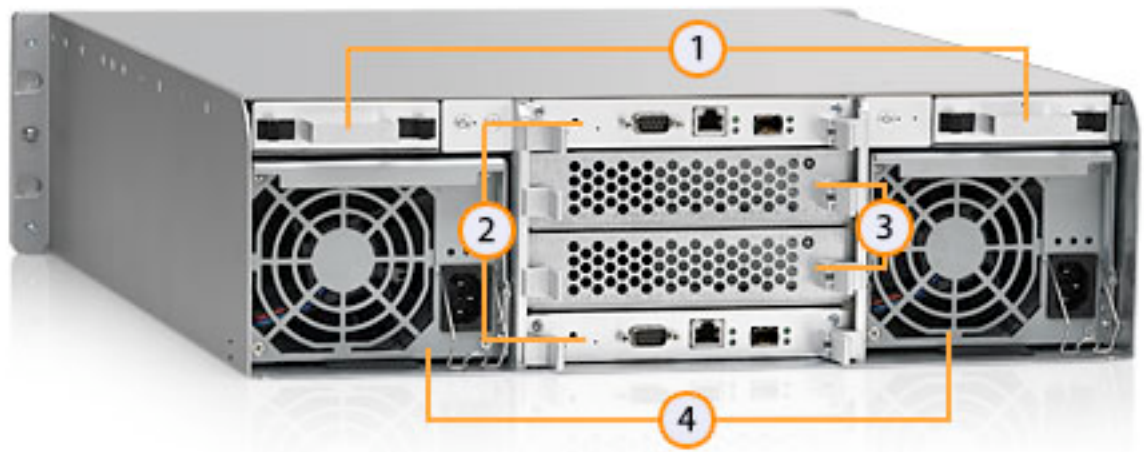
Xserve RAID

- Up to 7TB of redundant online storage
  - High performance - up to 380MB per second
- Flexible deployment options
  - Certified by leading Fibre Channel infrastructure vendors including Cisco
  - Qualified for use in Linux, Windows, NetWare, and multi-platform networks
- Pricing: below 1,70€ per gigabyte

1. Apple Drive Modules
2. Independent ATA drive channels
3. Drive controllers
4. RAID controller modules
5. Cache memory
6. Dual independent RAID processors
7. Fibre Channel ports
8. Redundant environment managers
9. Ethernet ports
10. Serial ports
11. Redundant cooling modules
12. Redundant power supplies
13. Cache Backup Battery Modules (optional)



Xserve RAID



Convenient replacement modules AppleCare Service Parts Kits for Xserve RAID (sold separately). Each kit has an Apple Drive Module, power supply module, cooling module, and RAID controller module.

1. **Battery backup module.** Optional Cache Backup Battery Modules that provide more than 72 hours of backup power to protect data integrity.
2. **Dual independent RAID controllers.** Two independent storage processor units manage RAID functions, data transfers and failure protection for each set of seven drives.
3. **Redundant cooling modules.** Redundant, hot-swap cooling modules provide automatic front-to-rear cooling for rack environments.
4. **Redundant power supplies.** Either of the redundant, load-sharing, hot-swappable power supplies can power Xserve RAID alone in case one fails.

Up to 14 7200 RPM Ultra ATA hard drive modules. Because each hard drive is isolated on its own bus, a drive failure doesn't degrade the accessibility or performance of the surviving drives.

Xserve G5
Performance in a 1U Server

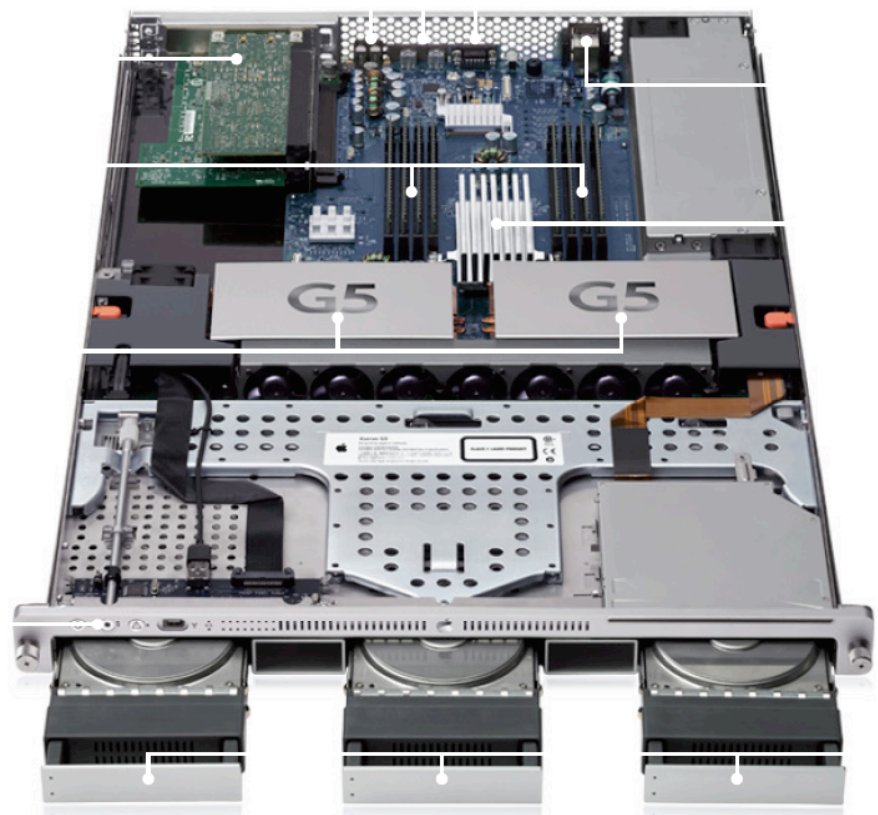
Xserve G5
Performance in a 1U Server

Dual PCI-X slots at up to 133MHz

Up to 8GB ECC DDR main memory

Dual 2.3GHz PowerPC G5 processors

Enclosure and drive lock



FireWire 800, USB 2.0, and DB-9 serial ports

Dual onboard Gigabit Ethernet ports

Advanced system controller with dual 1.15GHz system buses

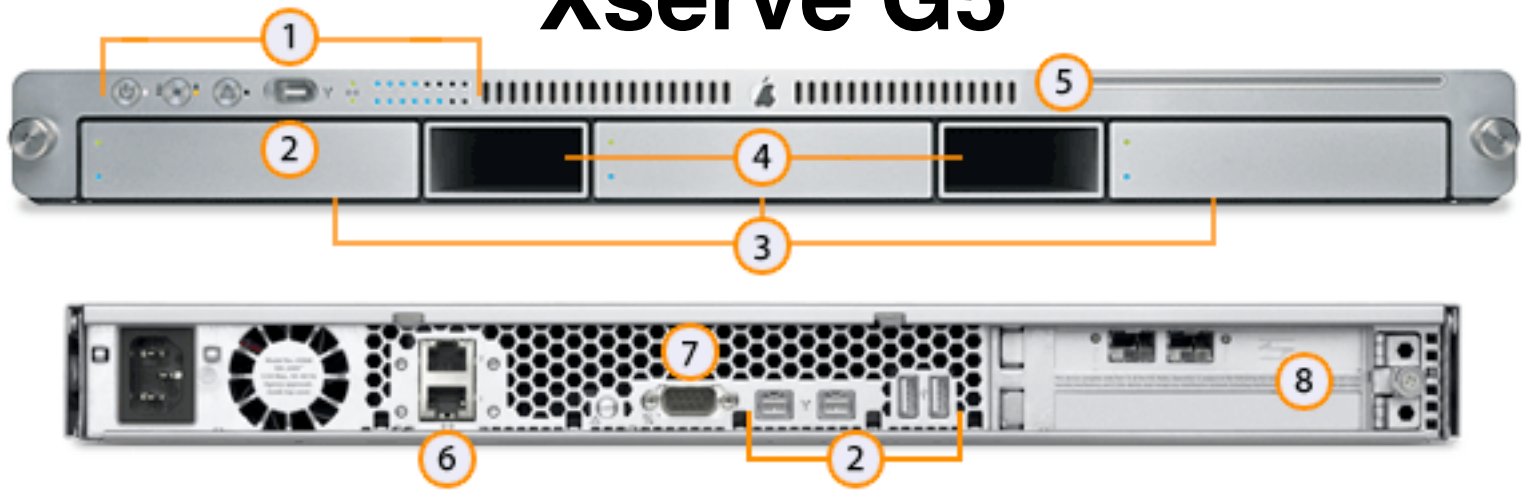
Up to three 400GB drive modules—1.2TB in 1U

At 2.3 GHz the G5 consumes at most 55W per processor. Compare that to 89W for an Opteron or 110W for a Xeon

Xserve offers three independent Serial ATA drive channels*, which can be configured with up to 1.2TB of internal storage per 1U server with an optional hardware RAID adapter for hardware RAID 0,1 and 5.

* Specifications differ for Xserve cluster configuration.

Xserve G5



1	**Status row:**	Indicator lights for power, enclosure lock, Ethernet links, and system identifier and activity. A lock secures ports, disk and optical drives and enclosure.
2	**Peripherals:**	FireWire 400 and 800 hard drives, two USB ports.
3	**Get a handle on it:**	Three serial ATA drive bays – up to 1.2TB of hot-plug internal storage. LEDs keep you informed of hard disk status and activity.
4	**Goes with the flow:**	Two well-placed ducts for better air circulation and improved performance
5	**Optical drive:**	Use this slot-load Combo drive (DVD-ROM/CD-RW) to add or reinstall software. Optional SuperDrive (DVD-R/CD-RW)
6	**Gigabit for two:**	Dual Gigabit Ethernet on the main logic board.
7	**All access pass:** is	DB-9 serial port – System access through a serial console session, even when the network down.
8	**Expansion cards:**	2 slots for PCI-X cards. Add one card running at 133MHz or two running at 100MHz. (Fibre channel card(s))

High-speed fibre transport.

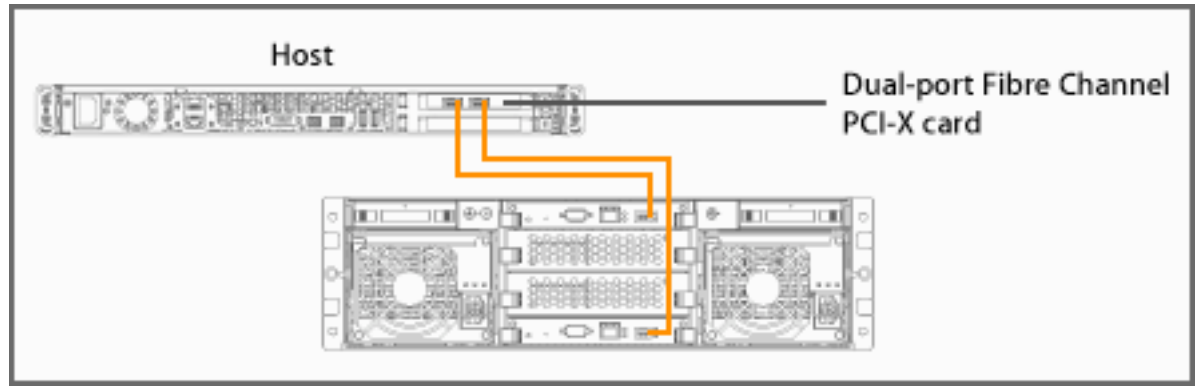
Dual 2Gb Fibre Channel transport
Industry standard SFP connectors
Works with Mac, Windows, Linux and NetWare

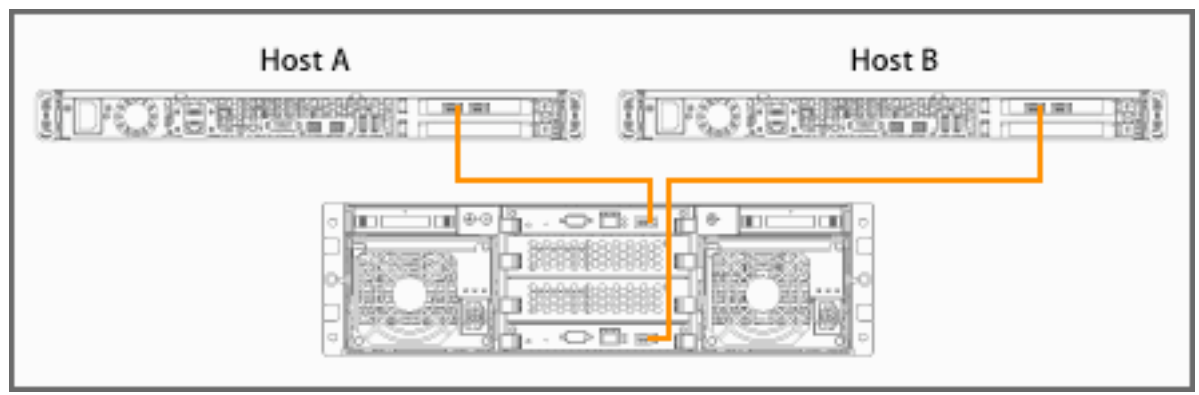
Fibre Channel Interconnect

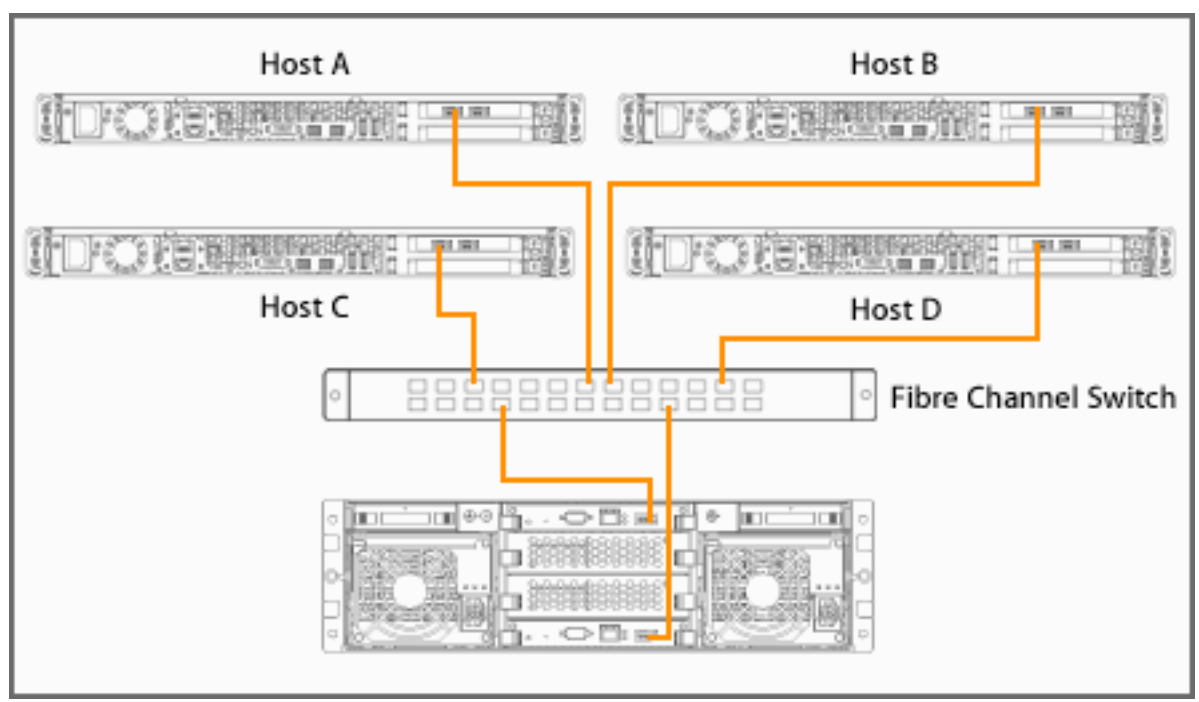
Dual independent 2Gb Fibre Channel ports -
provides superior bandwidth, availability, and deployment flexibility over SCSI technology.
Xserve RAID uses the latest industry-standard SCSI 3
protocol and work with industry-standard Fibre Channel switches.

400MB/s throughput
Each 2Gb Fibre Channel port offers 200MB/s bandwidth for a total throughput of up to 400MB/s. Fibre Channel is the only storage connectivity technology that provides guaranteed bandwidth, so the host computer receives data at the same high speed as the RAID system sends it.

Superior scalability
Fibre Channel technology can address up to 126 devices per loop,
(In contrast, SCSI allows for a maximum of only 15 devices per channel)


Host
Dual-port Fibre Channel
PCI-X card


Host A
Host B


Host A
Host B
Host C
Host D
Fibre Channel Switch

Mac OS X Server v10.4

Xsan

The SAN file system for Mac OS X

The dual 2.3GHz Xserve G5

Single 2.0 GHz or Dual 2.3 GHz G5 processors
Up to 1.5TB hot-swappable storage
Dual Gigabit Ethernet and FireWire 800 onboard
Up to 16GB of DDR SDRAM with ECC

Xserve RAID

Up to 7TB in 14 hot-swap drives
High-perfomance 2Gb Fibre Channel

Apple Server and Storage Products

Lowering the cost and increasing the value of IT solutions

Xserve G5

Xserve G5 Cluster Node

Xserve RAID

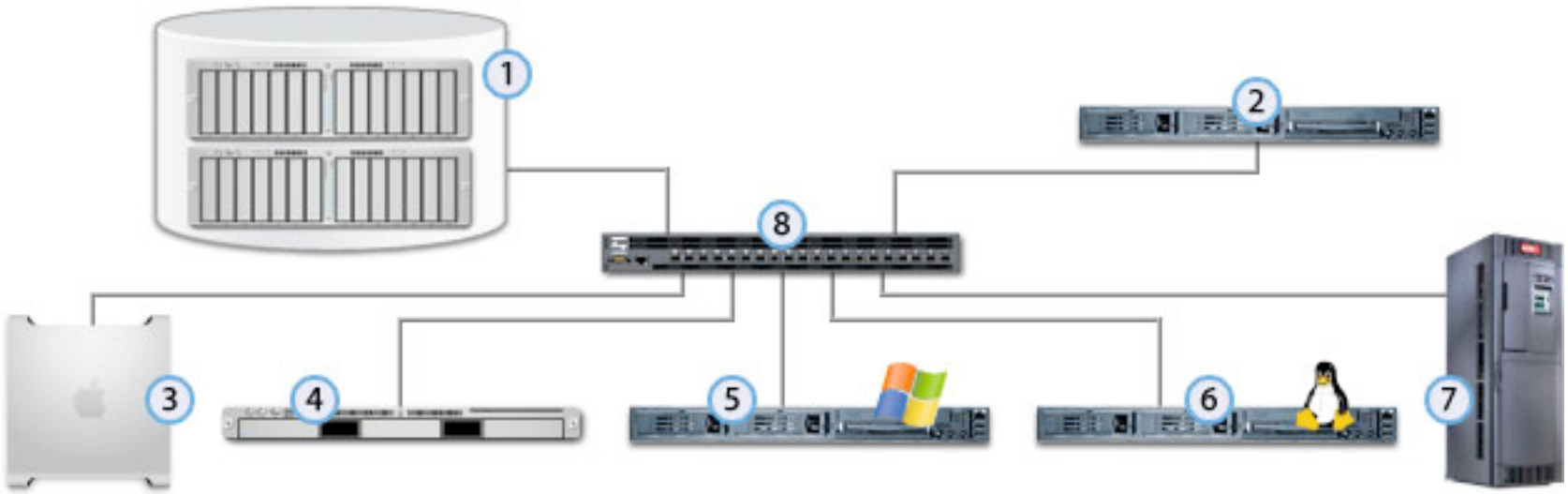
Mac OS X Server

Xsan

Apple Remote Desktop

AppleCare Support Programs

Mixed Platform XSAN



1	**Storage Pool.**	Use Xserve RAID in volumes up to 16TB.
2	**Metadata Controller.**	Directs servers to files in pool.
3	**Mac Workstation.**	Processes data in storage pool.
4	**Mac Server.**	Provides network services to client machines.
5	**Windows Server.**	Add to SAN via ADIC StorNext File System.
6	**Linux Server.**	Add to SAN via ADIC StorNext File System.
7	**Tape Library.**	Back up data for reliability and regulations.
8	**Fibre Channel Switch.**	Connects nodes to the storage area network.

Mixed Platform XSAN

Xsan – create a storage area network compatible with third-party products

Connect the Macs – Fibre Channel switch.
Apple has qualified the

- Brocade Silkworm 3200, Silkworm 3800, Silkworm 3900 and Silkworm 12000 series;
- QLogic SANbox2-8, SANbox 2-16, SANbox2-64 and SANbox 5200;
- Emulex SAN Switch 355, 375; Emulex 9200;
- Cisco MDS 9000 family

to work with Xsan and Xserve RAID via the Apple Fibre Channel PCI-X card.

Connect Everything – heterogeneous network
add non-Mac OS X clients to your Xsan-based SAN using

- ADIC's StorNext File System

Add any combination of
Windows XP-, Windows 2003-, Windows NT-, Windows 2000-,
Linux-, IRIX-, Solaris-, or AIX-based computers
as nodes on your storage area network.
Xsan and the StorNext File System work together seamlessly.

Certifications

Designed for Microsoft® Windows® Server 2003

ORACLE

Oracle ASM
(Automatic Storage Manager)

QLOGIC

Fibre Channel switches &
– host bus adapters

YES CERTIFIED SUSE® LINUX
Novell®

Linux Enterprise
Server 9

YES TESTED & APPROVED
Novell®

NetWare

Cisco Systems

MDS 9000
directors &
switches

BROCADE

Switches

redhat.

Enterprise Linux

McDATA®

Switches

LSI LOGIC

Fibre Ch. PCI-X
Fibre Ch. PCI-Express

EMULEX

Switches

VERITAS™

Volume Manager™

ATTO
ATTO Technology, Inc.

Fibre Channel
infrastructure

adic
Intelligent Storage™

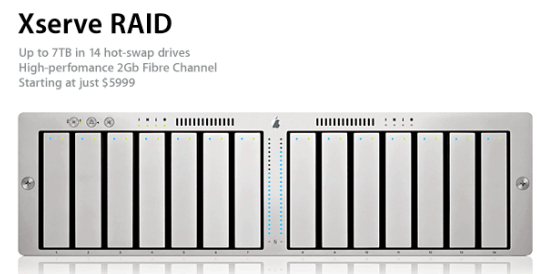
StorNext solutions

SAN
Hardware

Mass Storage - RAID

14 x EDU Apple Xserve RAID 7000GB

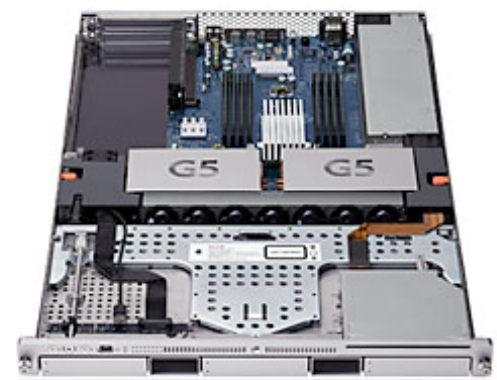
- 14x500GB Apple drives — 98TB
- 2x512MB RAID Controller-Cache
- redundant power & cooling modules
- 2 RAID Controller



SAN
Hardware

Cluster Server

4 x EDU Apple Xserve G5 2,3 GHz dual processor,
512KB L2 Cache, 1,15GHz system bus per processor,
4GB (4*1024) DDR ECC SDRAM at 400MHz,
160GB (2*80) Serial ATA,
DVD/CDRW Combo drive,
Hardware RAID PCI-card,
Fibre Channel PCI-X-card,
without graphic card,
2*Gigabit Ethernet,
2*PCI-X slots,
2*FireWire 800,
1*FireWire 400,
2*USB 2.0,
1*DB9 serial port (terminal/terminal server)



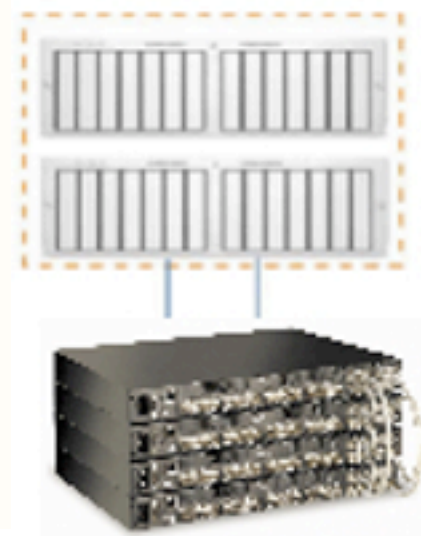
SAN
Hardware

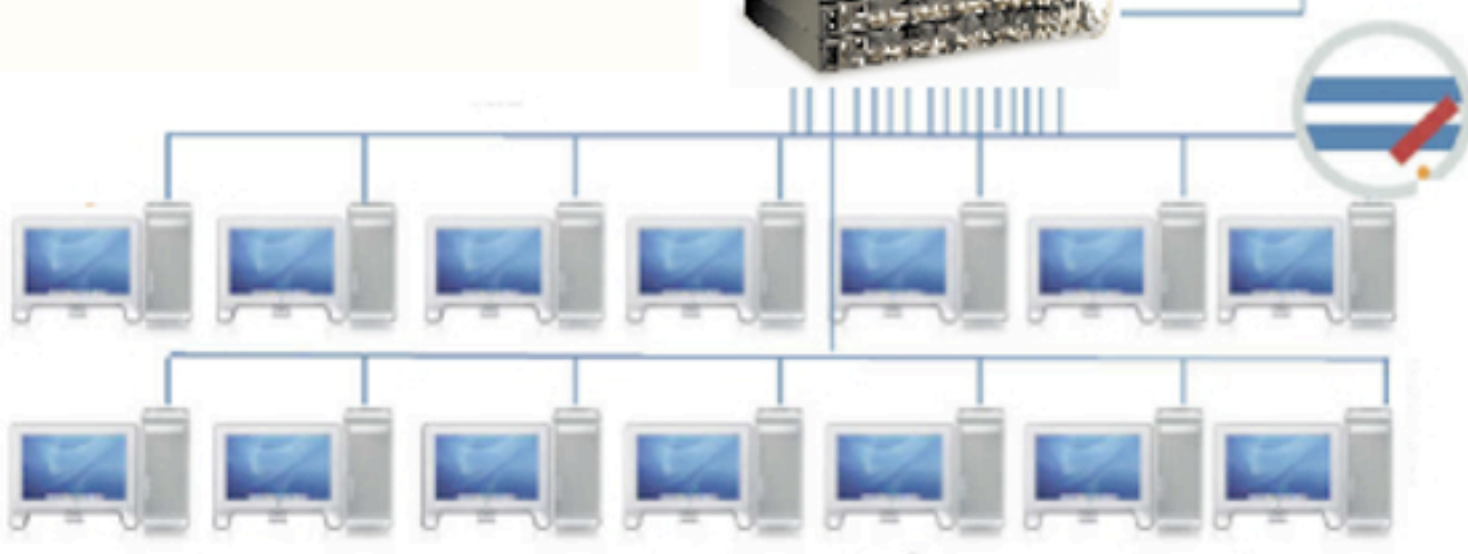
Fibre Channel Switches

4 x QLogic SANbox 5202 Stackable Fibre Channel Switch
16 x 2 Gbit ports and 4 x 10 Gbit stack-ports,
redundant power supply
SANsurfer Management Suite
warranty: 2 years

4 x 10 Gbit stacking-cable (Cu)
XPAK-plugs for
QLogic SANbox 5200 Fibre Channel Interswitch-Links

4 x Rackmount-Kit for
QLogic SANbox 5202 stackable Fibre Channel Switch





SAN
Software

Xserve
1 x Apple Mac OS X Server (unlimited number of clients)

4 x EDU Apple Xsan 1.1
Storage Area Network Software
Compatible with Mac OS X v.10.4
64 Bit File System access on 2 PetaBytes (PB) of data
scalable and fast access onto data

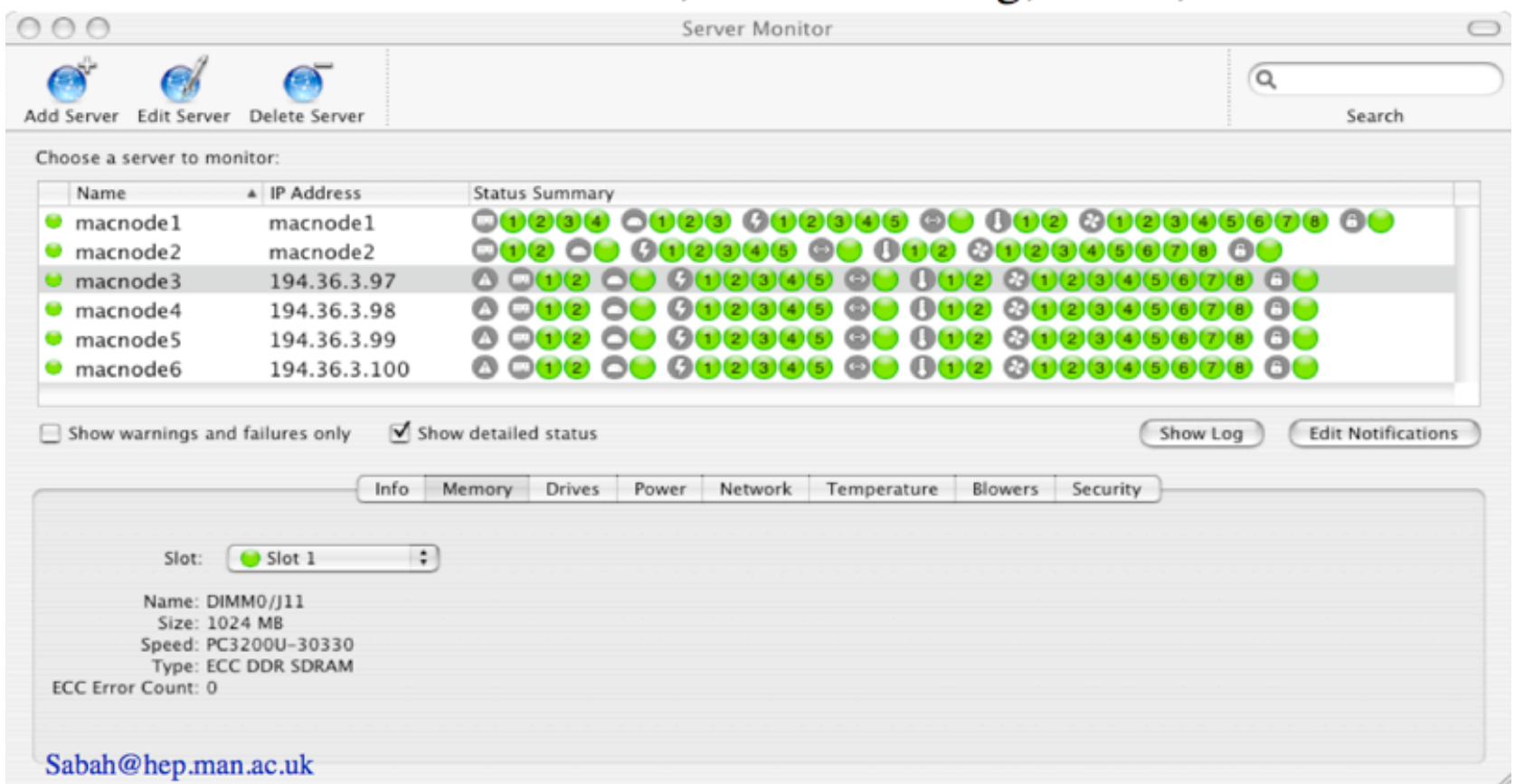
Remote control and monitoring
1 x EDU Apple Remote Desktop 2.2 (10 Client-unlimited:+200€)

SANsurfer Management Suite (QLogic)
Fibre Channel • iSCSI • MacOSX • Windows • Linux • NetWare • Solaris • HBAs • Switches

Fixed support cost include:
Software maintenance - 3 years all major upgrade
Premium support - 3 year on-site service, 24/7 phone & email
Service spare kits for Xserve and Xserve RAID

OS X at a glance

- OS X platform is increasing in popularity amongst developers and users on both the LHC experiments at CERN and the Tevatron at Fermilab
- Familiar Unix OS, no dual-booting, robust, stable



http://hepwww.rl.ac.uk/sysman/Nov2004/talks/apple-hepsysman.ppt

Summary

- Existing infrastructure AFS, home directories, ~200 node farm... √
  SAN could be set up to serve all needs

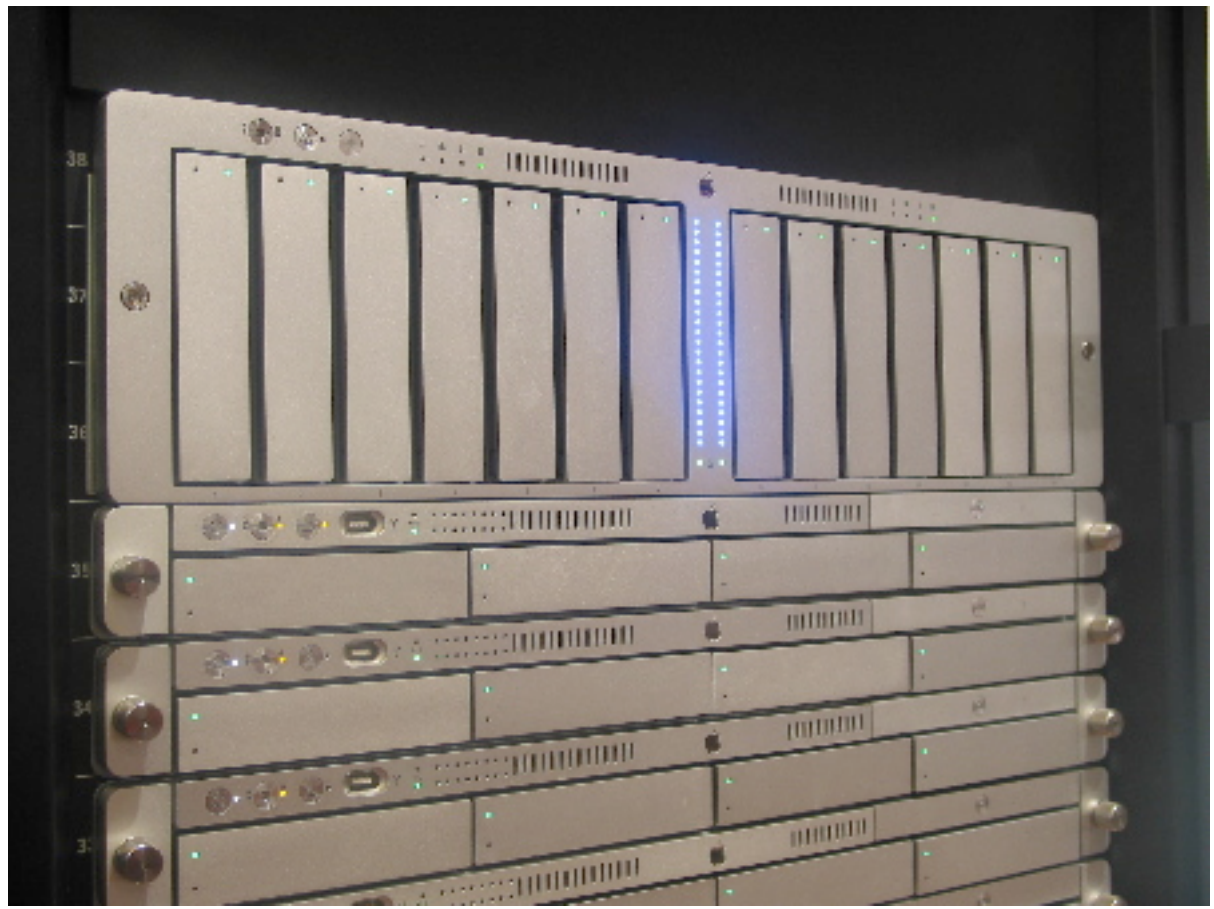
- True capacity requirements √
  limited by budget – but scalable – should we start now?

- Throughput requirements – Stephan W.: 1-2GB/s now – peak 10GBs ?
  point to point to farm node?
  SAN is the only system with all freedoms of extension – but again €
  Apple only solution does not meet these throughput requirements.
  Tiered storage with different requirements could be adopted – € √

- Availability requirements √
  The level of redundancy, safety margins for data etc. – € limited again
  could be very high level – to be discussed tiered data storage – including backup/archive/disaster scenario; Hierarchical Storage Management (HSM)

- Budget
  - Staged deployment √
  - Management software √
  - Service contracts √
  - A trusted vendor ? (√)

- Storage is a key topic – fast, flexible access from almost everywhere



Q & A