Mass Storage at SARA

Peter Michielse (NCF)

Mark van de Sanden, Ron Trompert (SARA)
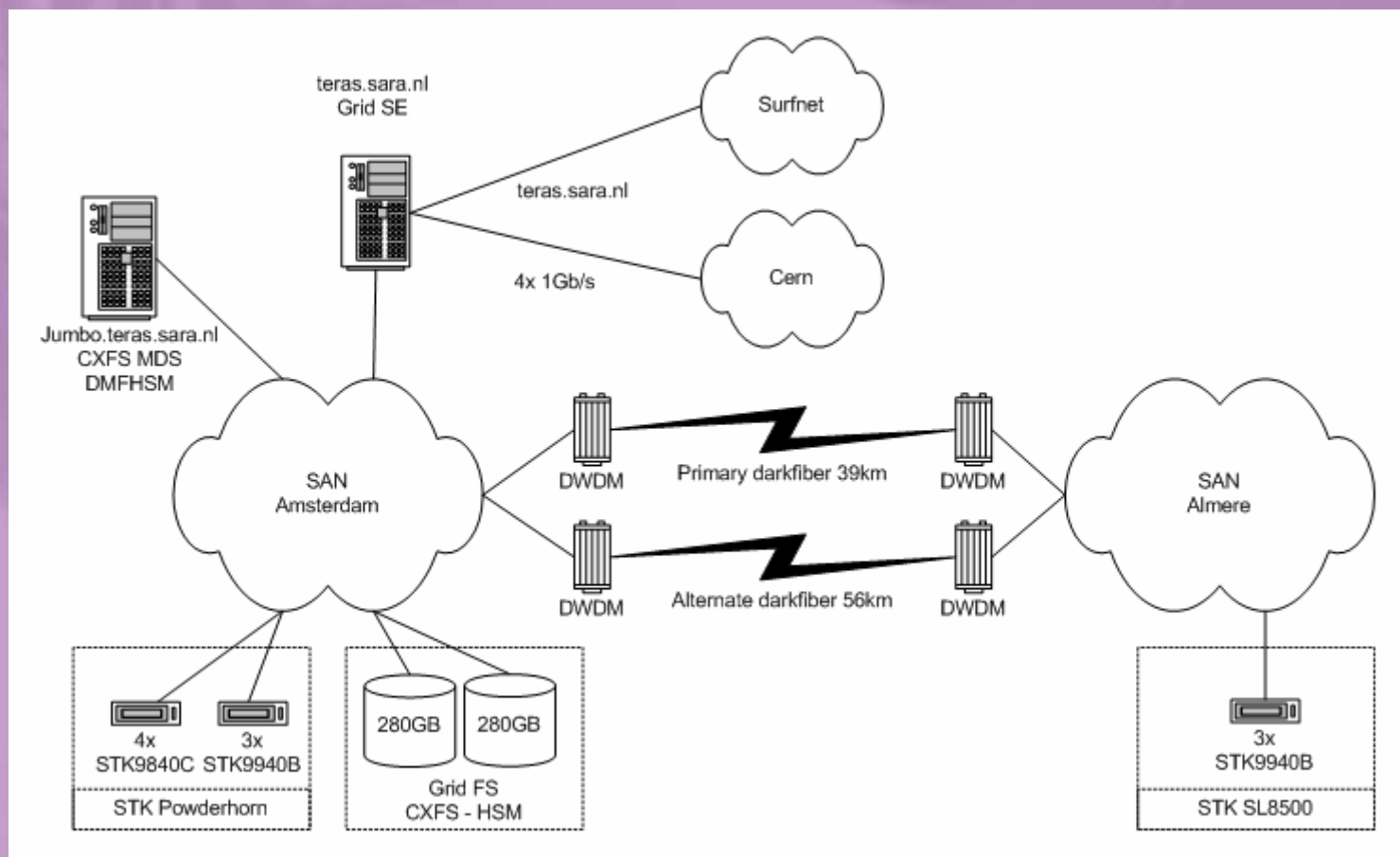
GDB – CERN – January 12, 2005

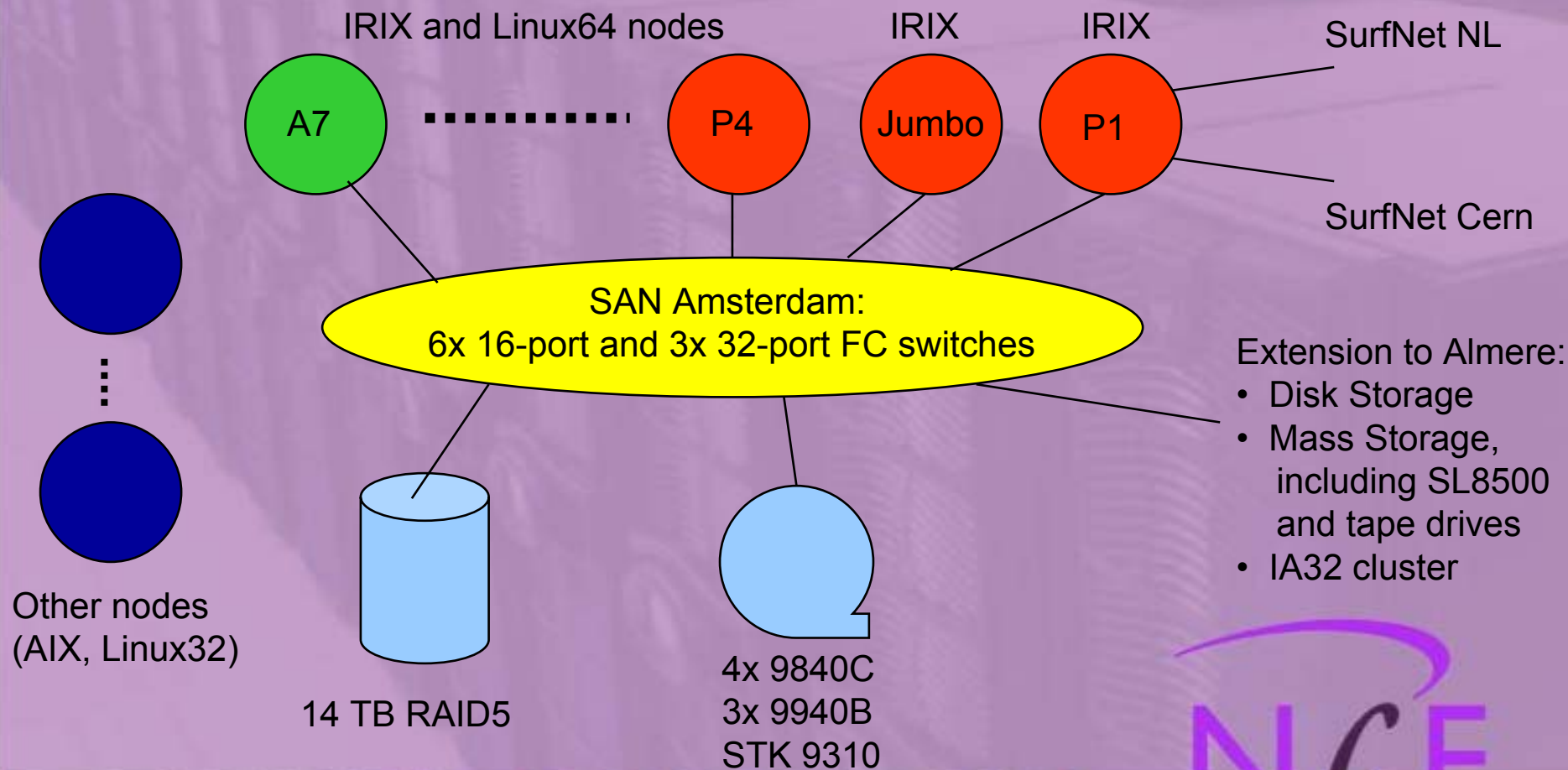Netherlands National Computing Facilities Foundation

Contents

- **High-level set-up at SARA**

- **Storage details**

- **December 2004 situation for Service Challenge**

- **Service Challenge tests and possible configurations (with 50-80 MB/s in mind for July 2005)**
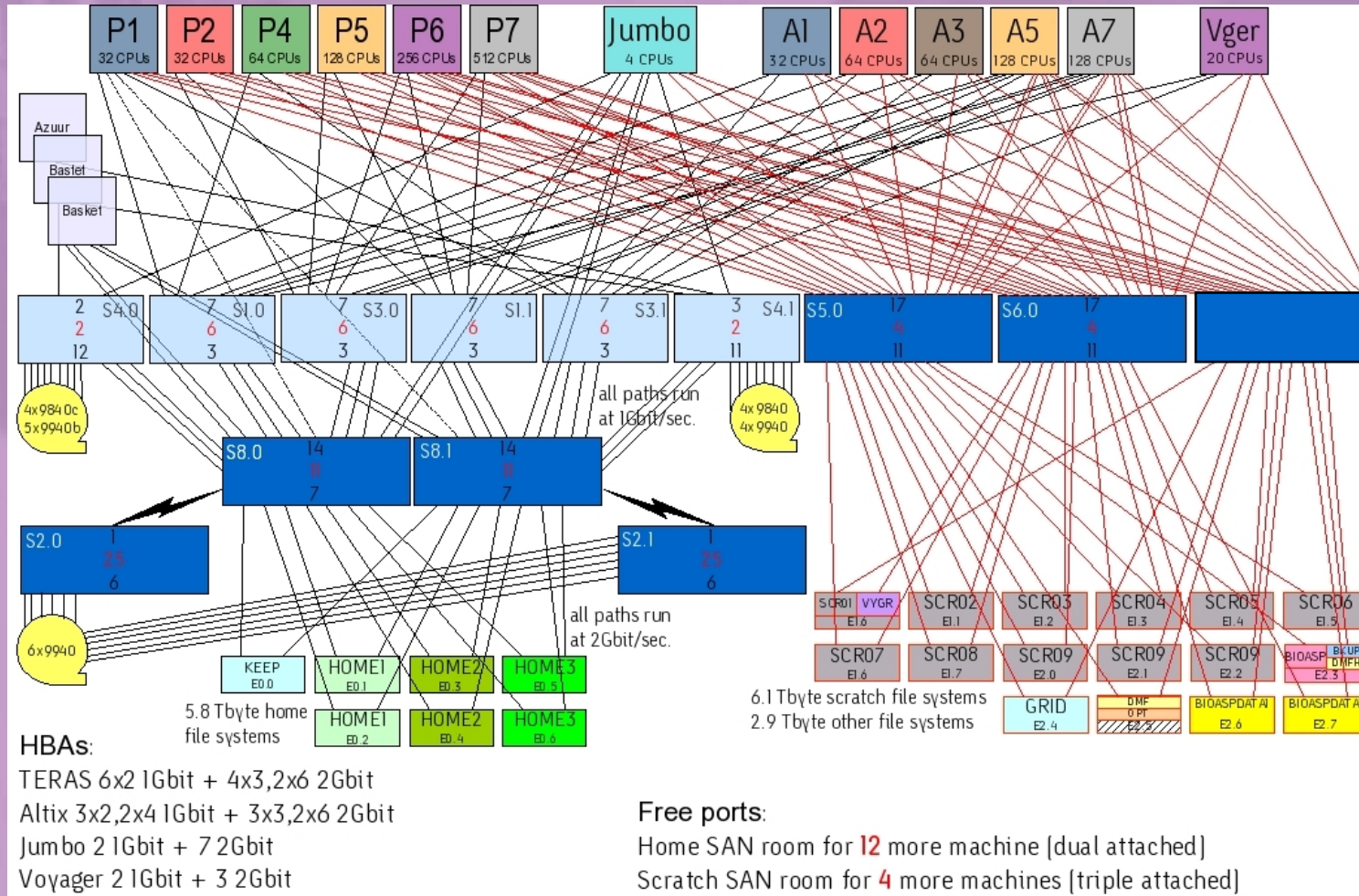
2

Netherlands National Computing Facilities Foundation

High-level set-up at SARA - 1

High-level set-up at SARA - 2

IRIX and Linux64 nodes

IRIX

IRIX

SurfNet NL

A7 •••••••••• P4 Jumbo P1

SurfNet Cern

SAN Amsterdam:
6x 16-port and 3x 32-port FC switches

Extension to Almere:
- Disk Storage
- Mass Storage,
  including SL8500
  and tape drives
- IA32 cluster

Other nodes
(AIX, Linux32)

14 TB RAID5

4x 9840C
3x 9940B
STK 9310

Netherlands National Computing Facilities Foundation
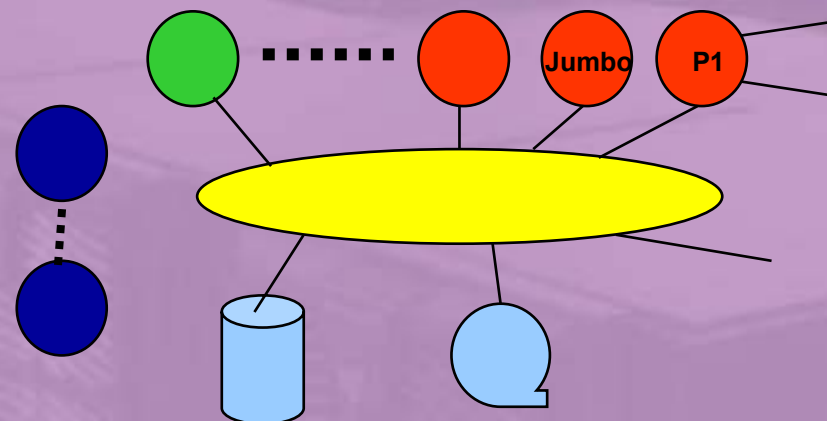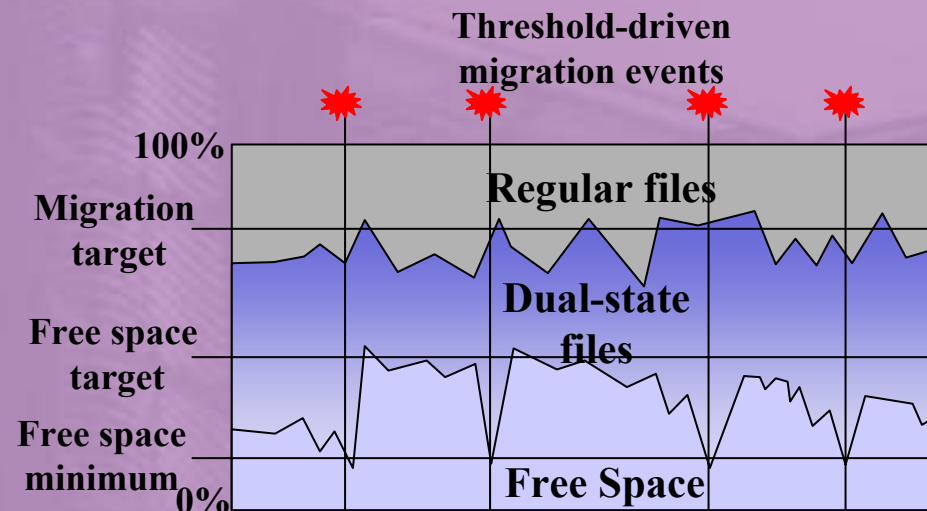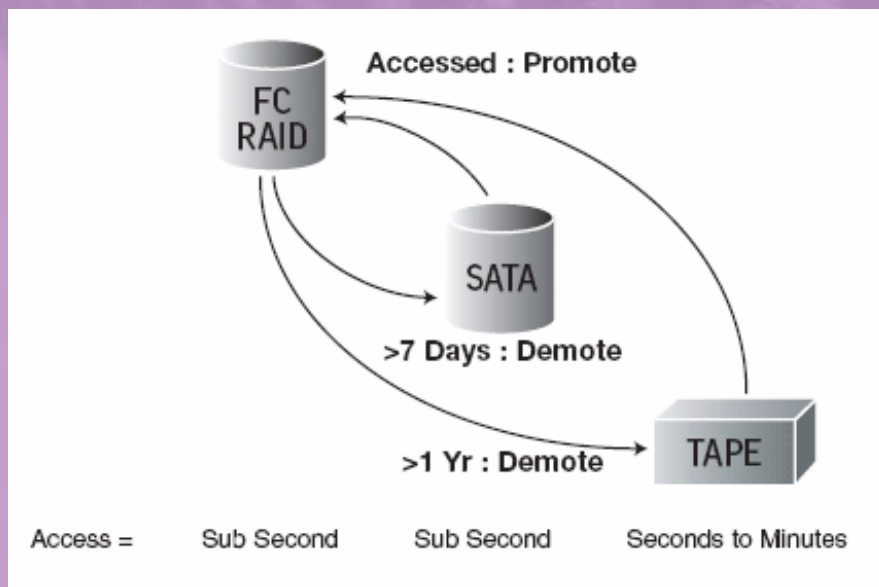
Storage Details - 1

Storage Details - 2

- P1 is SGI Origin3800
  - 32p (MIPS), 32 GB
  - Part of 1024p TERAS
  - IRIX (SGI Unix)
  - CXFS since 2001 (SGI's shared file system)
  - Interactive node for users to test and submit their jobs
  - Has been used so far as the Grid Storage Element

- Jumbo is SGI Origin350
  - 4p (MIPS), 4 GB
  - IRIX (SGI Unix)
  - CXFS MetaDataServer
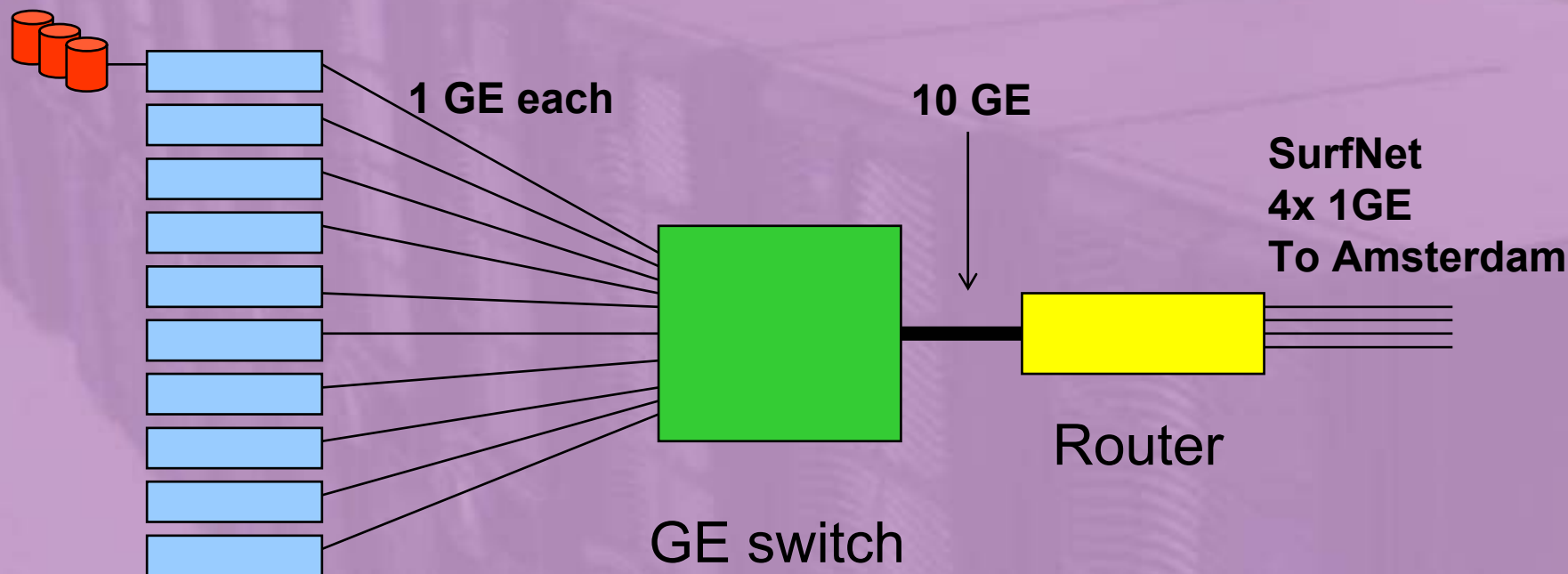 - DMF/TMF (SGI's hierarchical storage manager)

Netherlands National Computing Facilities Foundation

Storage Details - 3

- DMF/TMF (SGI's hierarchical storage manager)

Netherlands National Computing Facilities Foundation

December 2004 situation for Service Challenge - 1

1 GE each

10 GE

SurfNet
4x 1GE
To Amsterdam

Router

GE switch

10x HP rx2600
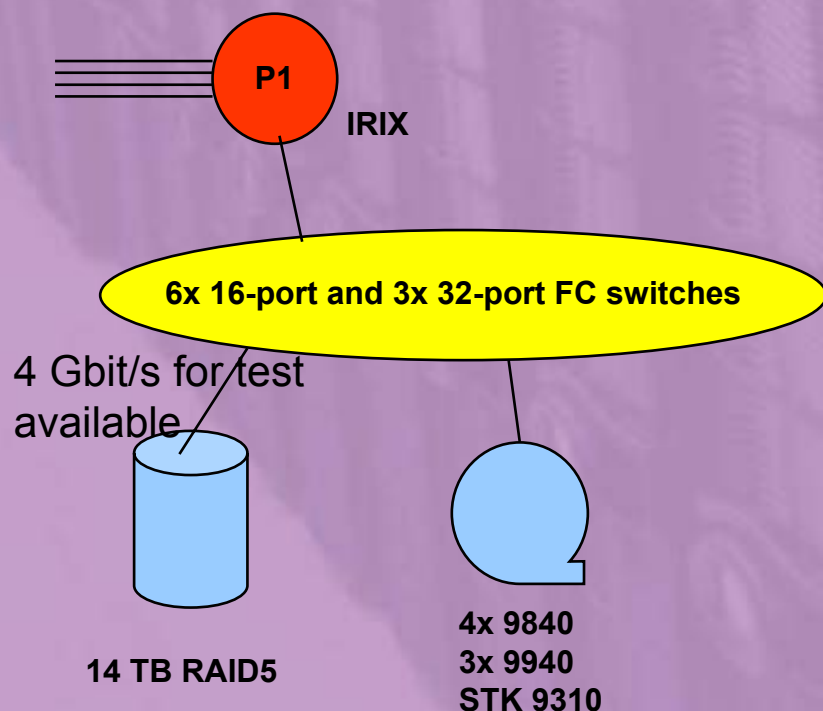**2x 1.5GHz Itanium-2**
1 GE to LAN
1 GE to WAN

Coming out from CERN T0 .....

December 2004 situation for Service Challenge - 2

SurfNet, 4x 1GE from CERN

.... Arriving in Amsterdam T1

P1

IRIX

6x 16-port and 3x 32-port FC switches

4 Gbit/s for test available

14 TB RAID5

**4x 9840
3x 9940
STK 9310**

**This was the situation so far:**

- **Disk-to-disk**
- **P1 (the Grid SE) is part of a general purpose production facility, so difficult to experiment with**
- **Uses proprietary OS (SGI IRIX), while tools are based on RH 7.3**
- **Limited by disk storage**
- **Relatively old Gbit cards in P1**
- **Planning not optimal**
- **DMF HSM seems to be an advantage**

Service Challenge tests until July 2005 - 1

- **Timeline:**

| When	CERN T0	SARA/NIKHEF T1
January-March 05	500 MB/s	Set up and test initial configurations
April-June 05		Set up preferred configuration
		50-80 MB/s in place
July 05	1000 MB/s	Demonstrate: 50-80 MB/s – 1 month
		CERN Disk to SARA/NIKHEF Tape

- **We need to figure out an alternative to:**
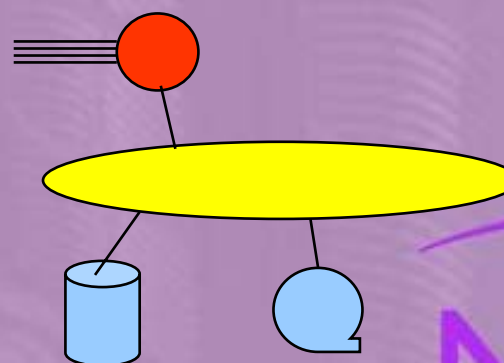
Service Challenge tests until July 2005 – 2: Ideas

- **Separation of tasks:**
  - Servers that handle incoming data traffic from T0
  - Servers that handle outgoing data traffic to T2's
  - Servers that handle mass storage (tape)
- **Consequences for storage environment:**
  - Direct Attached Storage
  - Integrated storage: SAN, global file system
  - Layers of disks
- **Hardware and software choices**

Netherlands National Computing Facilities Foundation

Service Challenge tests until July 2005 - 3
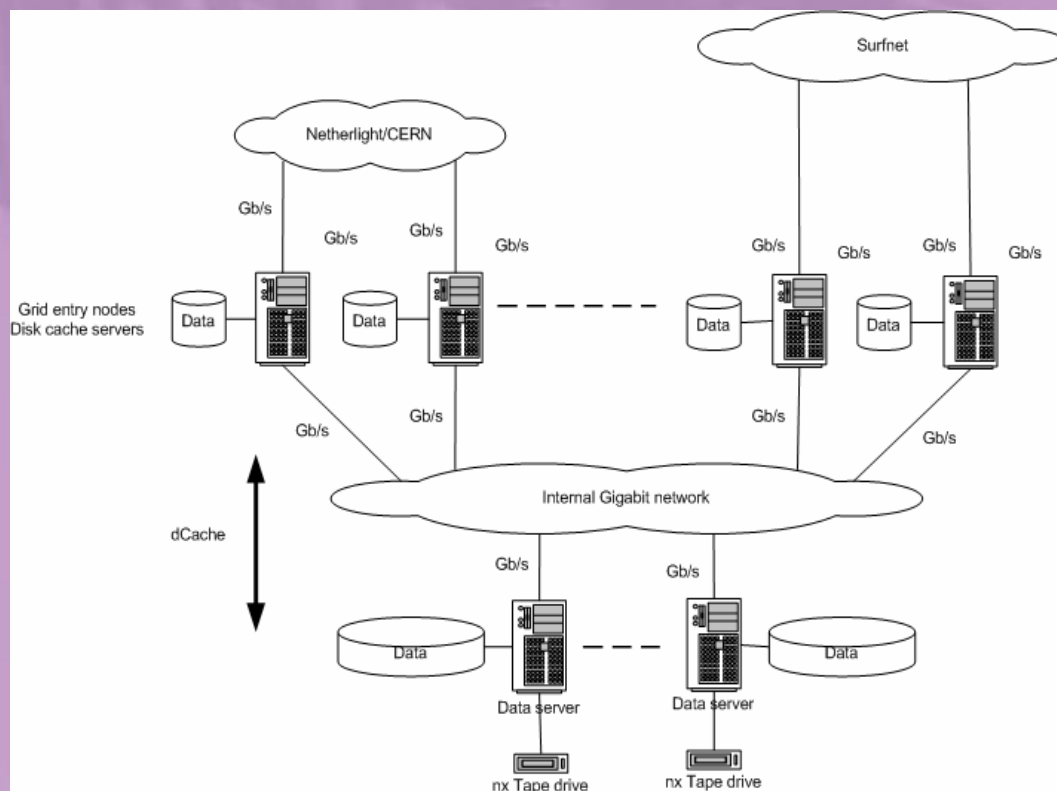
- **Several alternatives - 1**
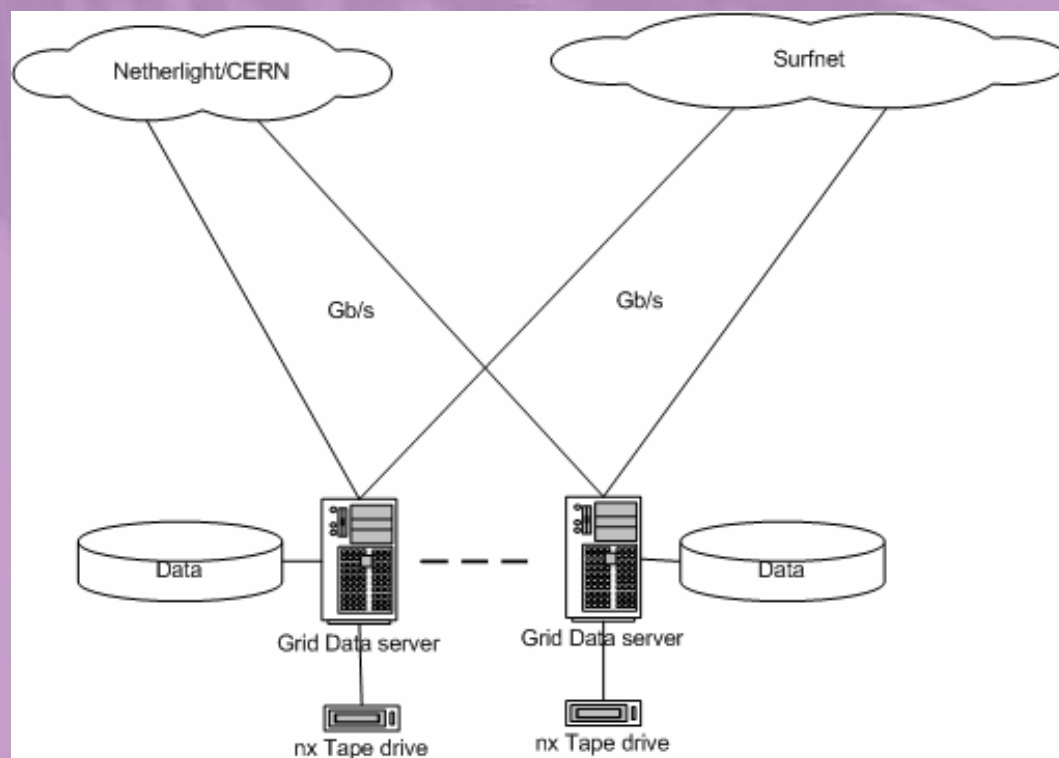
Pro:
- Simple hardware
- Some scalability

Con:
- No separation of incoming (T0) and outgoing (T2) data
- No integrated storage environment
- Tape drives fixed to host/ file system

Netherlands National Computing Facilities Foundation

Service Challenge tests until July 2005 - 4

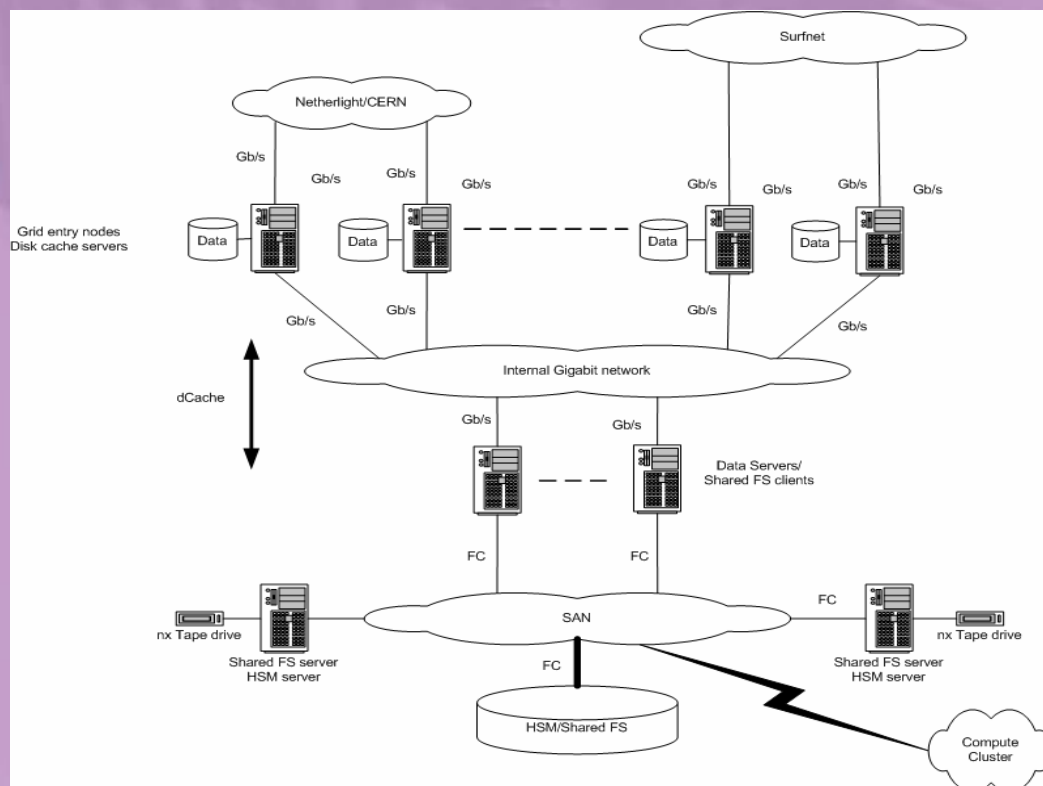- **Several alternatives - 2**

Pro:
- Simple hardware
- More scalability
- Separation of incoming (T0) and outgoing (T2) data

Con:
- More complex setup, so more expensive
- No integrated storage environment
- Tape drives fixed to host/ file system

Service Challenge tests until July 2005 - 5

- **Several alternatives - 3**

Pro:
- Simple hardware
- More scalability
- Separation of incoming (T0) and outgoing (T2) data
- HSM involved, tape drives not fixed to host/file system
- Integrated storage environment

Con:
- More complex setup, so more expensive
- Possible dependance on HSM

Netherlands National Computing Facilities Foundation

Summary

- **Choices to be made**
  - Based on tests in 1HCY2005
  - Hardware, software and configuration
  - First tests with available Opteron cluster, as in alternative 1, not yet with tape drives
  - Subsequently testing of other alternatives

Netherlands National Computing Facilities Foundation