Virtual Observatory
Paris Data Centre

Storage Virtualisation

Pierre Le Sidaner Claude Huc Paul Kopp
Observatoire de Paris
CNES

Context and Goals

Context

Storage is an element of an Archive

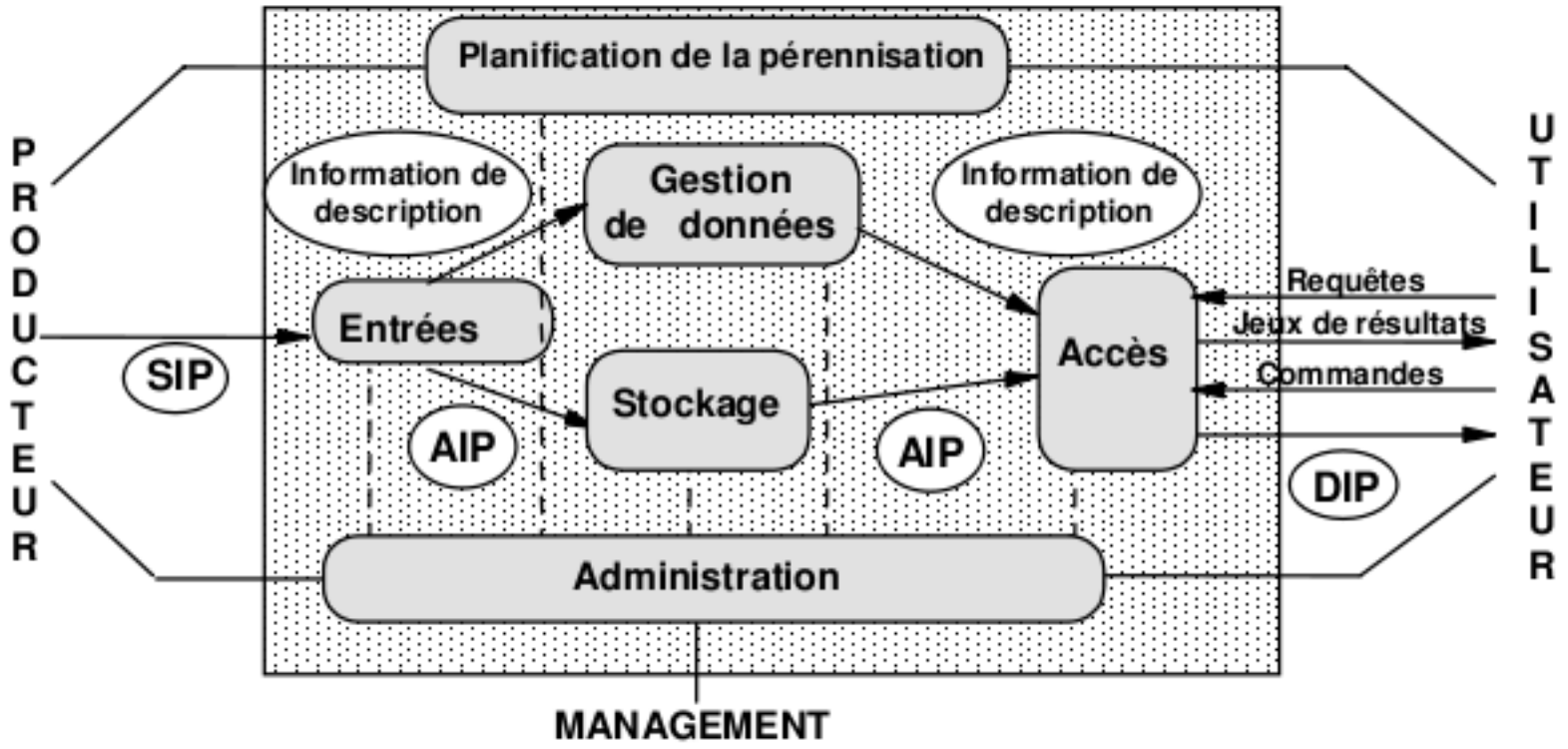
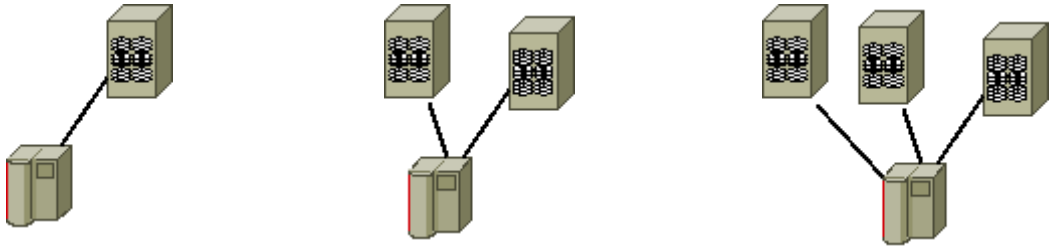


Schéma 4-1 : Entités fonctionnelles OAIS

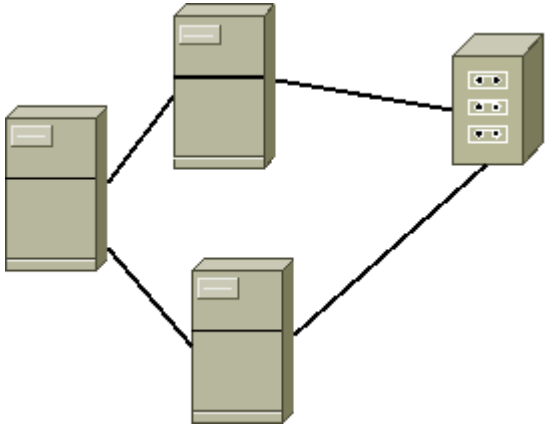
Context and Goals

- **storage must be capable in dealing with the increasing data volume**



The LVM (Logical volume manager) software can be use to concatenate the nodes

- As you become rich, the data volume and data storage quality increase. There is still need to manage hot-swap replecement and tackle the sclability



Context and Goals

- **Storage has tobe operating system independa**

- What is the tipical OS lifetime ?

- How does the filesystem depend on the OS ?

- How do OS update and migration affect/modify the data ?

- **storage has to deal with limited-lifetime storage elements**

- Can different storage element having different age be part of the same storage system ?
Can data migration be « au fil de l'eau »?

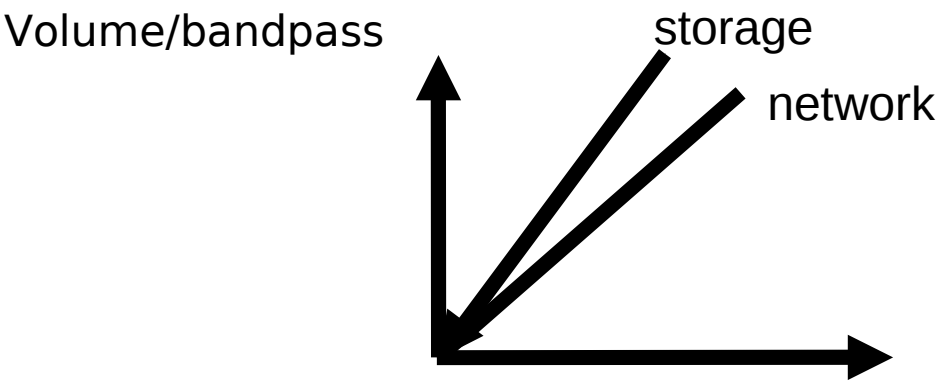
Context and Goals

- **Storage must deal with limited network bandwidth**

- usual network bandwidth is from 1Gb to 10Gb (for rich). Time transfer for 20 TB around 3day theoretically.

. 1980 Ethernet – 10 to 100 Mbps
[1996-1998] : Gbps
[2003-2005] : 10 Gbps

Compare it to storage volume evolution !



Context and Goals

- **Storage should be resistant to and easily recoverable from failures**

- How many days will it take to recover from your tape backup?

- How long will it take to make full backup ?

- How long will it take to synchronise two storage elements ?

- How to make replication (multiple instance of data) and be sure of consistency ?

- does you RAID array realy protect you from all these problems ?

Virtual Observatory
Paris Data Centre

Context and Goals

- **Storage should be easy to distribute geographically.**

- For security reason (major crash)
- For availability
- For geographically distributed project

- How to synchronise ?

- How to deal transparently with geographicaly distributed ?

Distributed Storage Virtualisation

- **An answer to Constraints**
  - Security, redundancy, quick access …
- **Presentations of different solutions and feedback from users**
- **Interactive session with a discussion about difficulties arising in different projects with storage management in term of data distribution and preservation**

Virtual Observatory
Paris Data Centre

Agenda

10h Introduction from : Pierre Le Sidaner Claude Huc Paull Kopp

10h15 Storage Virtualisation at IN2P3 : Jean Yves Nief

11h Break

11h 15 data storage virtualisation at Virtual Observatory Paris Data Centre: Pierre Le Sidaner

12h00 Methodic approach to designing and building scalable data management systems : Pawel Plaszczak

12h45 Lunch at IAS

14h30 Storage Virtualisation at CDS for VO storage space :A. Schaaff

15h15 Storage Virtualisation on SHAMAN Project : Adil hasan,

15h45 Break

16:00 Discussion and Conclusion