Scalable Preservation and Archiving Repository (SPAR)

National Library of France
Bibliothèque nationale de France
### Missions of the library

**Missions:**

- to build up the collections,
- to preserve and communicate them to the public,
- to produce a reference catalog,
- to cooperate with other institutions,
- to participate to research programs.

---

### Legal deposit:

- legal deposit since 1537 for printed materials
- 1648: engravings and maps
- 1793: musical scores
- 1925: photos
- 1938: phonograms
- 1975: videograms
- 1992: electronic documents
- 2006: Web legal deposit
Technical issue: deal with the volume

Initial digitization
B&W - 300dpi - TIFF G4
1 page ~ 200Kb

Today digitization
Color (24 bits) - 400dpi - TIFF uncompressed
1 page ~ 80Mb

More than x500 !!!
Functional issue: make a digital preservation copy

Digitization as a mean for preservation and dissemination
Business issue: native digital document

More and more documents exist only in digital form.

This is the form for the legal deposit and then for preservation.
How to manage digital archives

• Build a central repository to reduce the diversity (media, formats, departments …)
• To take inspiration from good practices and standards

⇒ Key requirements:
  – OAIS compliance
  – modularity and distributivity
  – abstraction
  – use of well known formats and standards
  – use of open-source technical building blocks
Digital archiving at BnF

Preservation digitization
Audiovisual
WEB Archiving

Preservation planning
Data management
Storage
Ingest
Administration
Access
Storage Abstraction Service (SAS)

SPAR - Realization
SPAR - Infrastructure

Dissemination applications
...
Choice of the infrastructure

• Openness of the solution
  – Openness to multiple environments
  – Compatible with multiple providers

• Availability
  – Reliability of all parts
  – Usage of the system 24h./day, 7d./week, 365d./year
  – Maximum down time: 2h per month

• Upgradeability
  – Great potential in increase of volume and power
  – Coverage of the need for the next 8 years
Infrastructure

Main site

Online storage

Secondary storage

Backup secondary storage

Backup servers

Backup site

Primary storage

Secondary storage

Lookup storage

Infrastructure
Primary and backup storage

Oracle StorageTek SL8500

- up to 64 tape drives
- up to 8500 tapes
- up to 8 hand pickers
- up to 32 linked libraries

Primary storage
- 2 libraries
- 16 PB maximum

Backup storage
- 2 libraries
- 16 PB maximum
Tape library in-situ
Medias

Primary storage
LTO5

Capacity: 1.5 TB
Transfer rate: 140 MB/s

Backup storage
T10000B

Capacity: 1 TB
Transfer rate: 120 MB/s

(previously: 9840C – 40GB)
(previously: T10000A – 500GB)
Servers

Oracle Sun M4000

Use the zoning capabilities of Solaris

4 Proc 2.15 GHz
64 GB RAM
2 HBA dual FC
Solaris 10
Operating platform

- JBoss
  - Pré Versement
  - Zone de travail

- JBoss ADM1
  - Administration
  - DM
  - Storage

- JBoss ING1
  - Ingest
  - DM
  - Storage

- JBoss ACC1
  - Access
  - DM
  - Storage

- JBoss ADM2
  - Administration
  - DM
  - Storage

- JBoss ING2
  - Ingest
  - DM
  - Storage

- JBoss ACC2
  - Access
  - DM
  - Storage

Storage unit
- SAM/FS
  - SAM-FS stage area
- iRods
  - iCAT
- PostgreSQL - SQL

Operating platform
Validation platform

Treatment unit

JBoss ADM1
Administration
DM
Storage

JBoss ACC1
Access
DM
Storage

JBoss ING1
Ingest
DM
Storage

JBoss ING2
Ingest
DM
Storage

JBoss ING3
Ingest
DM
Storage

Work areas
Solaris Zone

Storage unit
SAM/FS
SAM-FS stage area
iRods

Solaris

PostgreSQL - SQL
iCAT
Thanks for your attention

Questions?

Thomas Ledoux
thomas.ledoux_AT_bnf.fr