



{ BnF

IFLA International Newspaper Conference

**“Newspaper Digitization and Preservation.
New prospects.
Stakeholders, Practices, Users and Business Models”**

**11-13 April 2012
BnF, Paris**

With the support of:





Long-term preservation at BnF

**Scalable Preservation and
Archiving Repository
(SPAR)**

{ BnF

Agenda

- Context of digital preservation at BnF
- Overall schedule of the project
- Overall architecture and main features
- Service Level Agreements: how to build on mutual trust
- Conclusion

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

Le Monde



Le Monde.fr

France-soir



More and more documents exist **only** in digital form



This is the form for the legal deposit and then for preservation

Initial goals of the preservation

- Take the responsibility of digital data preservation
- Provide a storage facility for archives
- Reach a minimal critical mass to reduce costs (hardware, software and human)
- Allow for the pooling of archive among several institutions

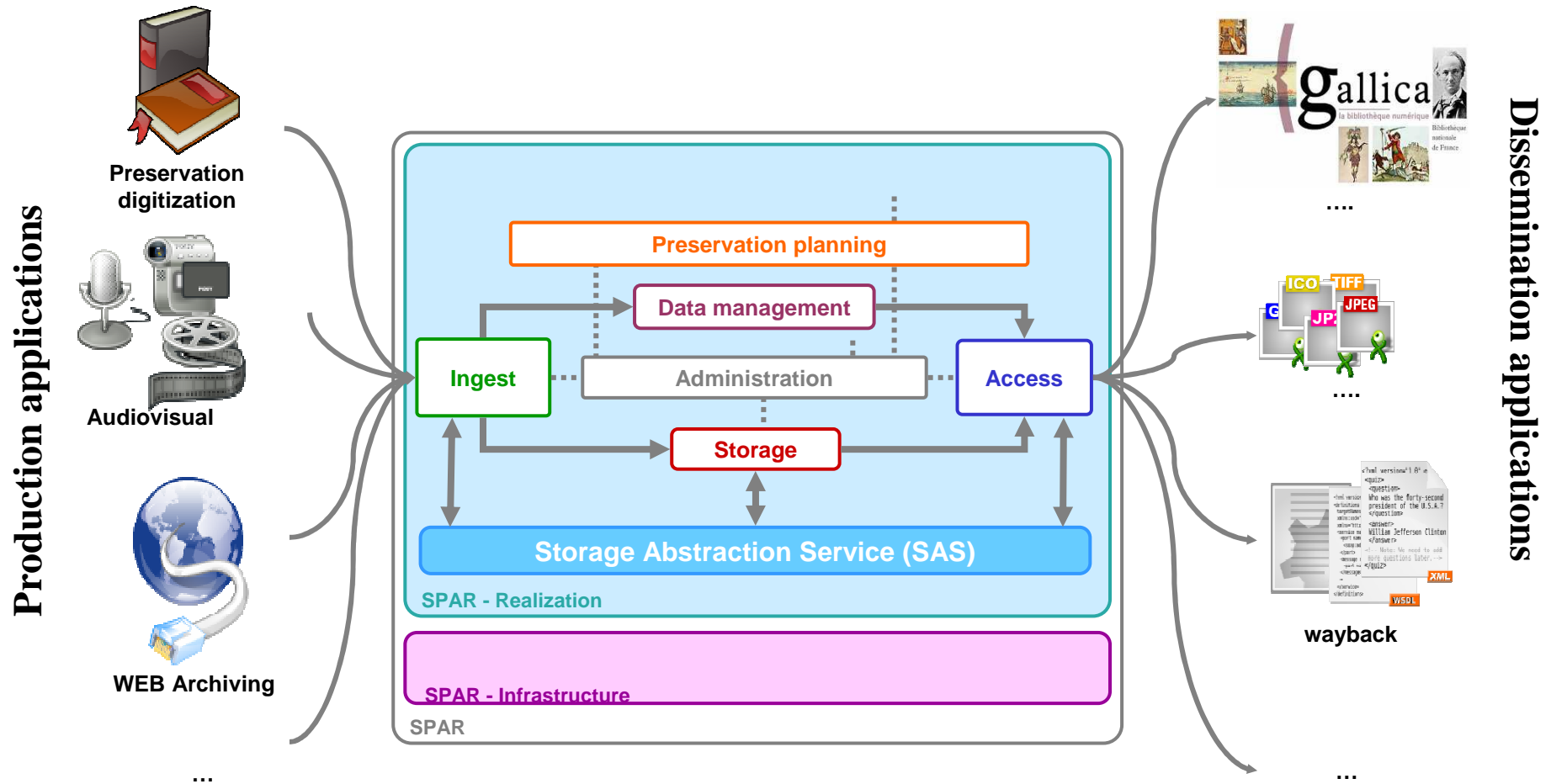
How to manage digital archives

- Build a central repository to diminish 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



Decomposition in channels

- To deal with the variability and heterogeneity of the data, definition of **channels**
- build on the relation between the digital objects and the archival system, independently of any given organization:
 - Preservation digitization
 - Audiovisual material
 - Negotiated legal deposit (dark Web, regional press)
 - Automatic legal deposit (surface Web)
 - Administrative production
 - Deposit / Third party archiving
 - Acquisition / Donation

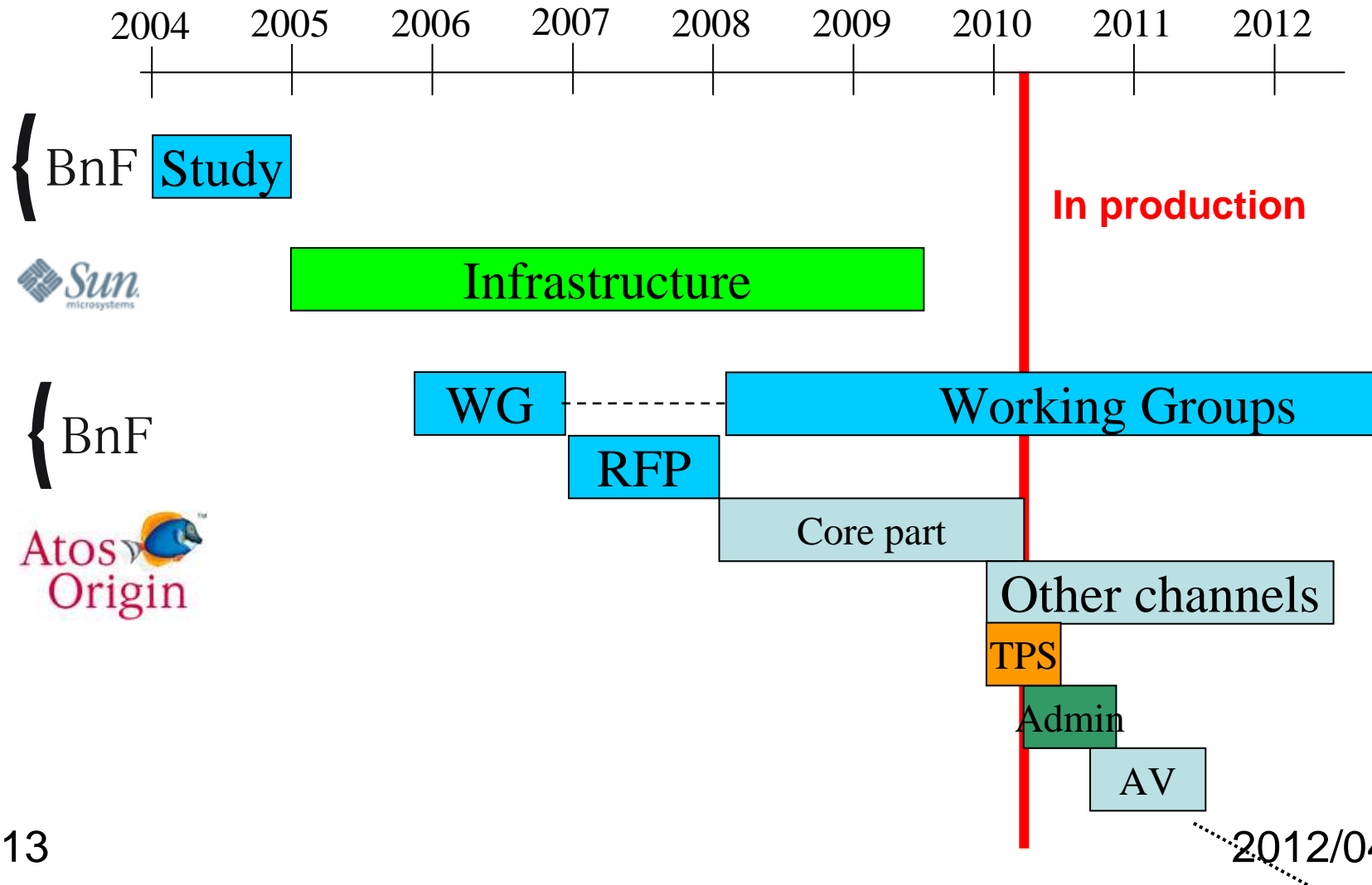
Agenda

- Context of digital preservation at BnF
- Overall schedule of the project
- Overall architecture and main features
- Service Level Agreements: how to build on mutual trust
- Conclusion

Schedule

- Requirements definition (Working Groups) 2006
- Technical study of existing solutions 2006
- Implementation RFP 2007
- Iterative development 2008 - 2009
- Validation starts december 2009
- Production starts march 2010

Planning

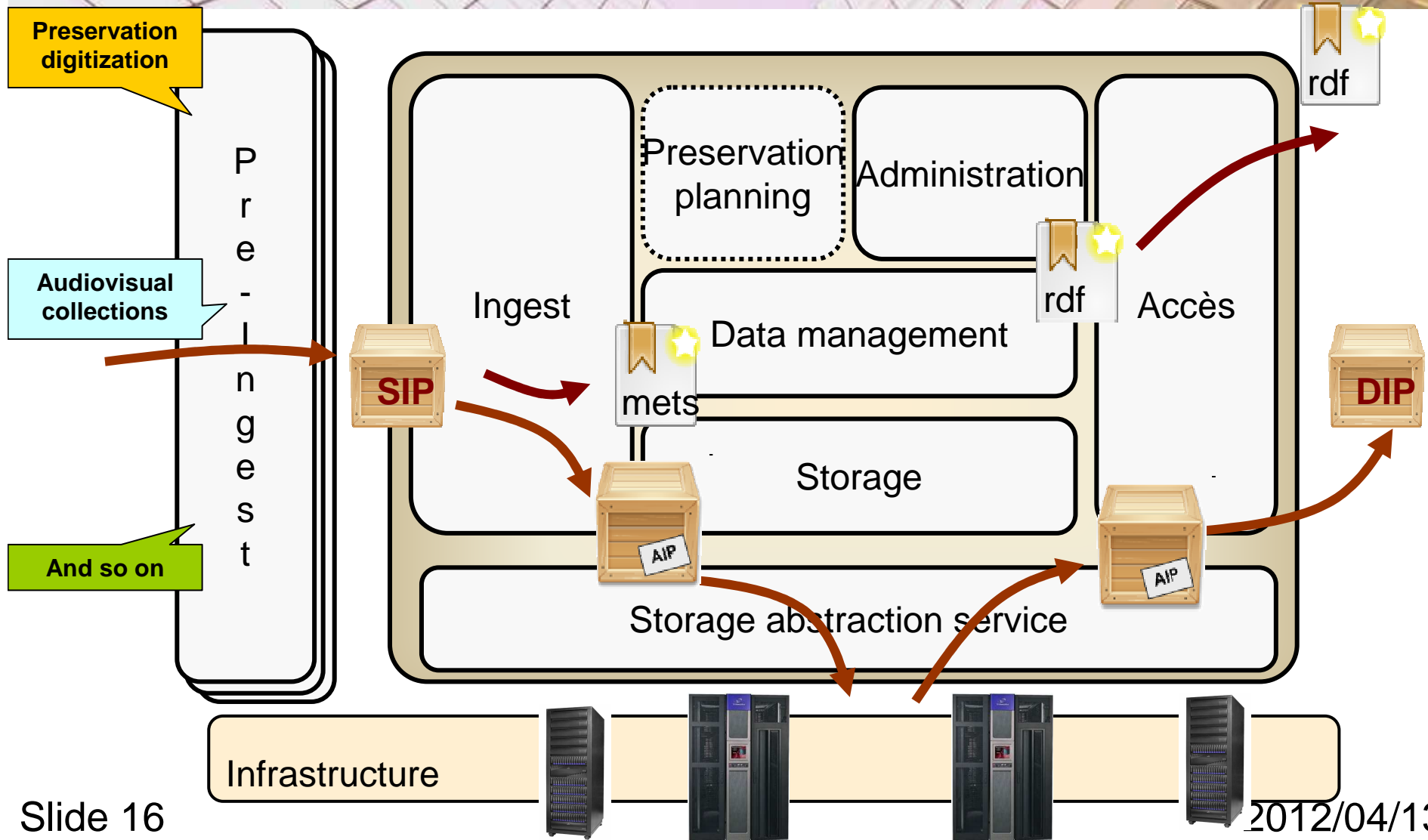


Agenda

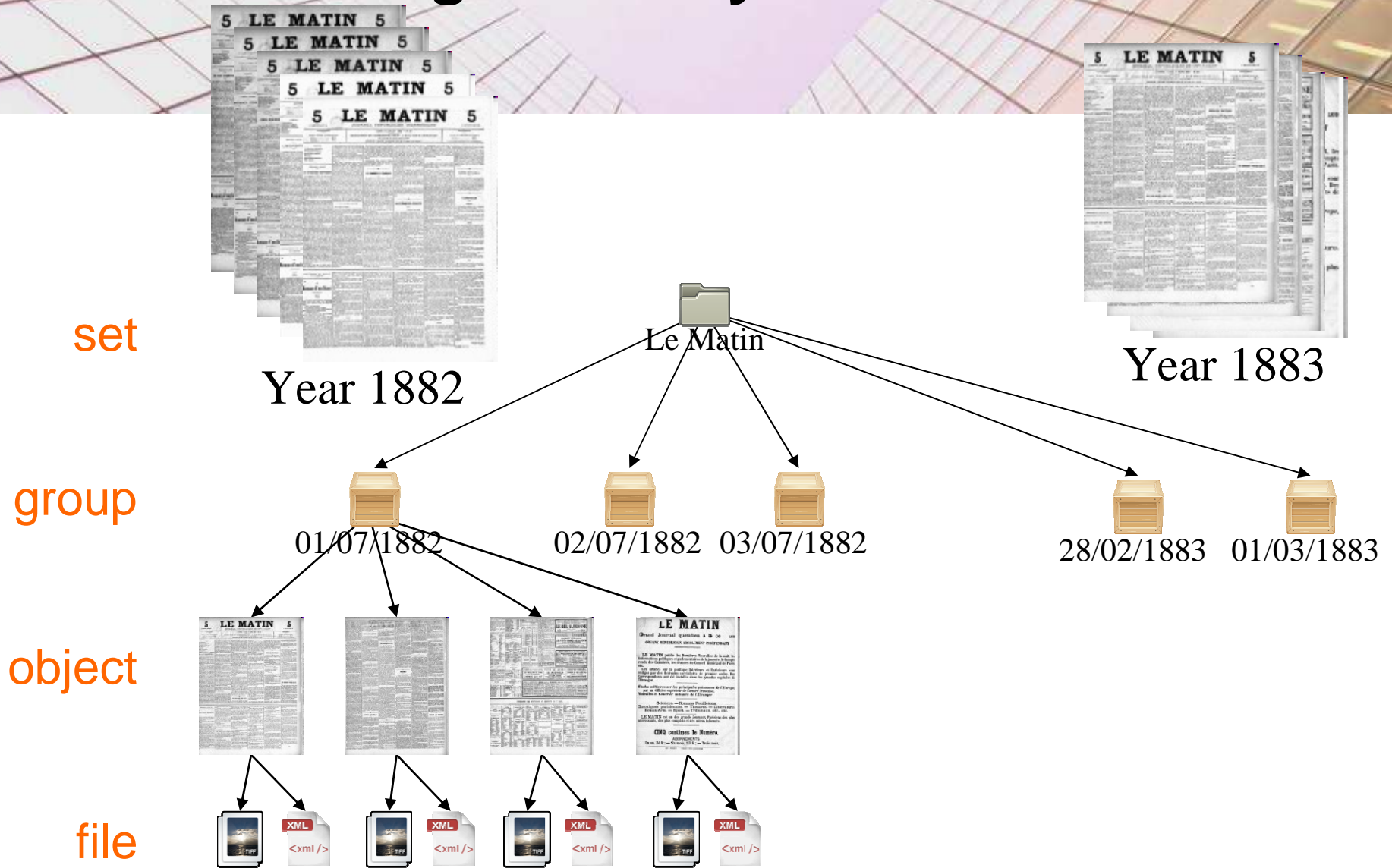
- Context of digital preservation at BnF
- Overall schedule of the project
- Overall architecture and main features
- Service Level Agreements: how to build on mutual trust
- Conclusion

SPAR

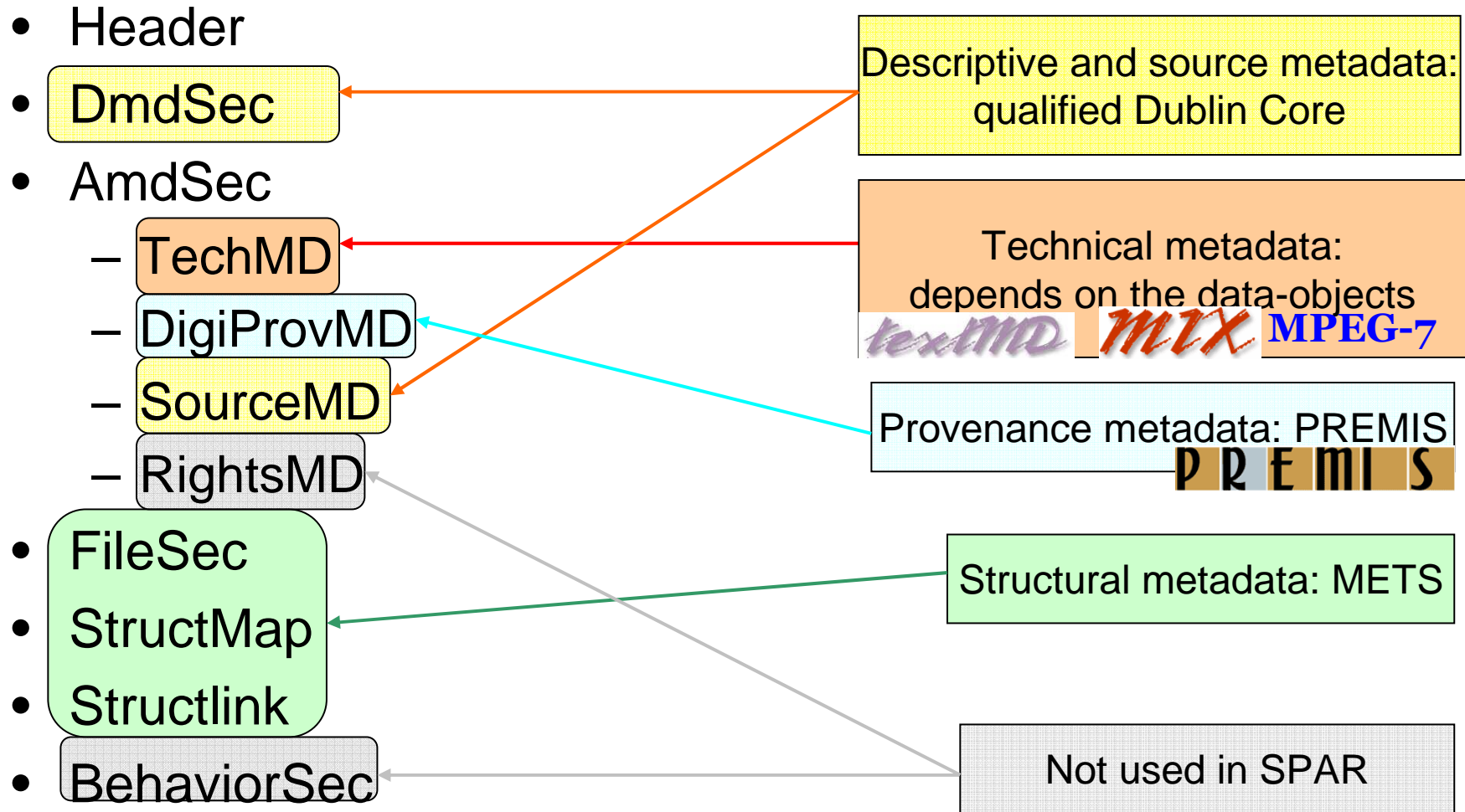
a generic repository solution at BnF



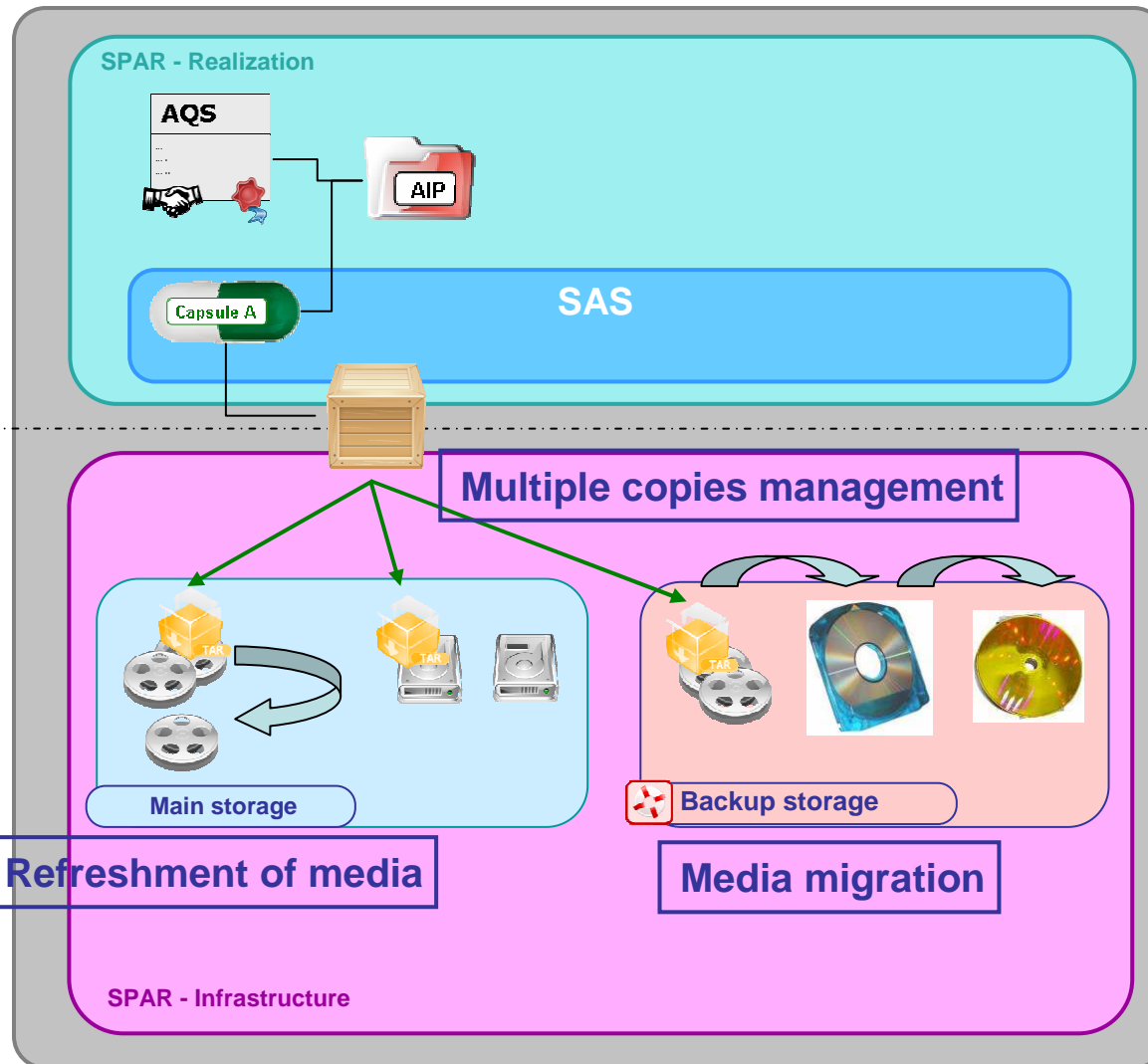
The levels of granularity



Highlights from our information model: the seven deadly sections of METS



Storage abstraction



Application viewpoint
Infrastructure viewpoint

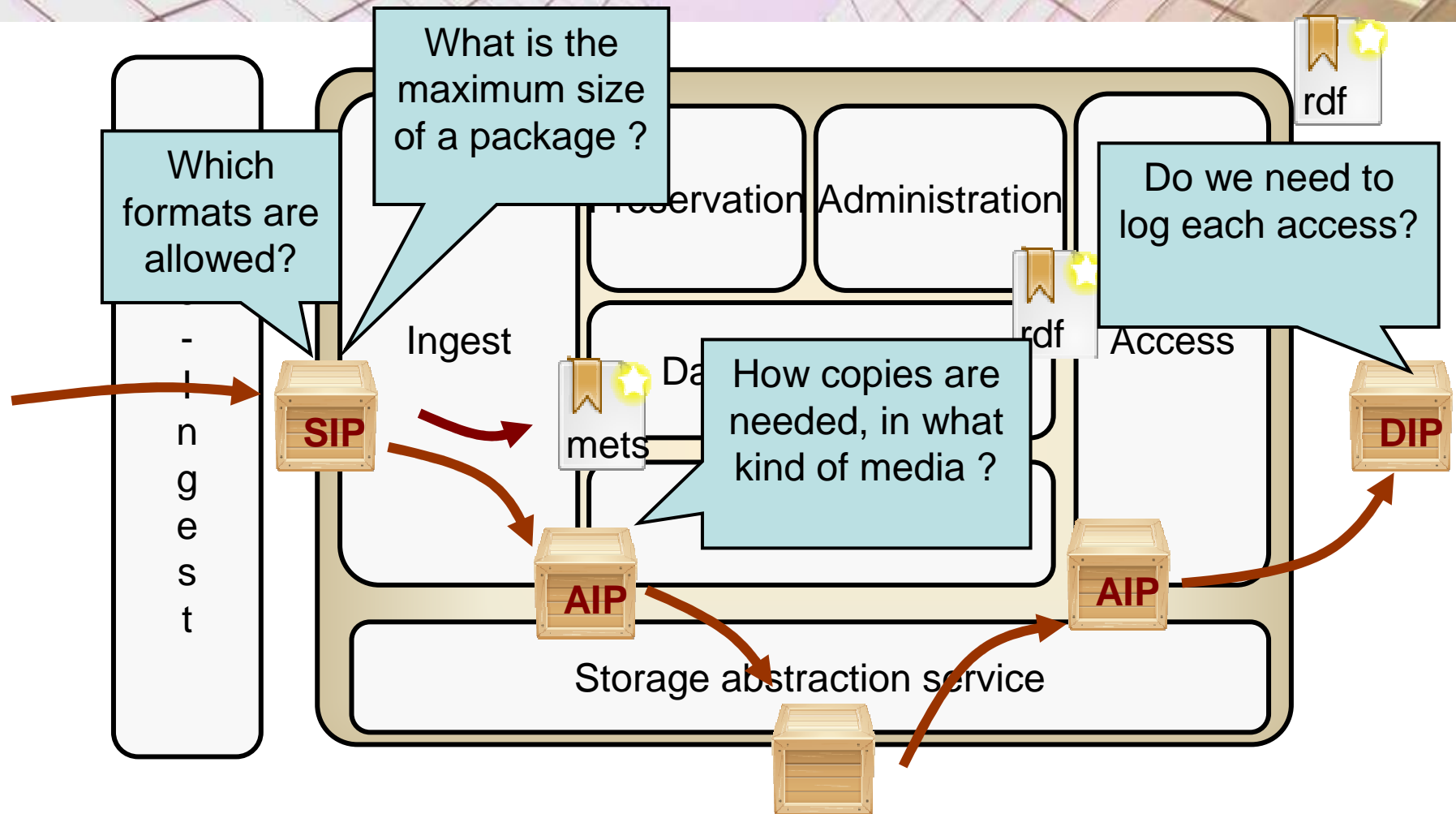
Agenda

- Context of digital preservation at BnF
- Overall schedule of the project
- Overall architecture and main features
- Service Level Agreements: how to build on mutual trust
- Conclusion

The Service Level Agreements

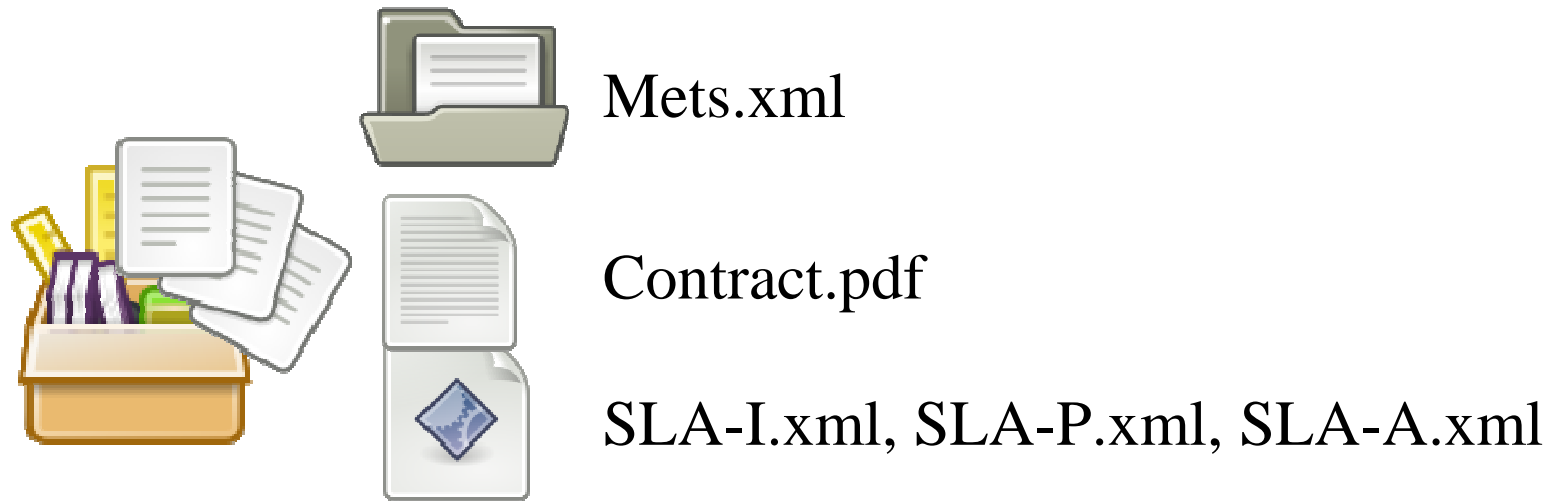
- The Service Level Agreements: contractualize with the users to offer a more transparent system
- The system is no longer a black box only known by technical experts

Taken into account in the system



SLA: definition

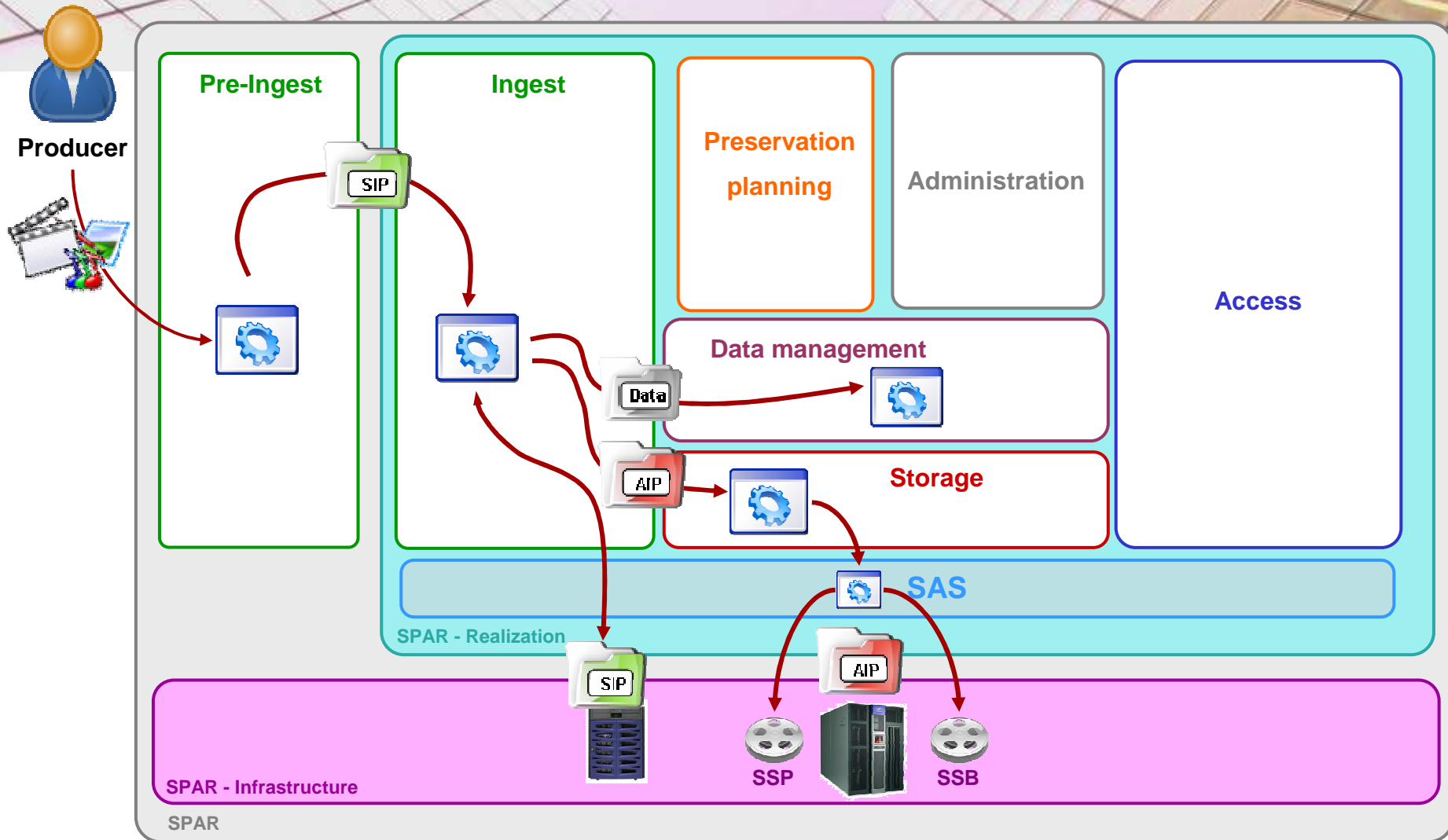
- 3 SLA: Ingest, Preservation, Access
- Formalize in XML the ways of managing the packages
- Those 3 SLA are recorded in a reference package that describes the channel



Agenda

- Context of digital preservation at BnF
- Overall schedule of the project
- Overall architecture and main features
- Service Level Agreements: how to build on mutual trust
- Conclusion

Overview of an ingest



Information about a package

Nom : ark-12148-btv1b9054864m.version0.release0.tar

Catégorie : AIP

Statut : OK

MD5 : 74a8b5e689462ba56d7ccd5524ec29de

Informations de stockage :

Copie	Élément de stockage	Chemin	volume ID
copie 1	capsAIP01	/sas/elemAIP01-1/AIP/2792 /2188/2294/ark-12148- btv1b9054864m.version0.release0.tar	A01706 74a8b5e689
copie 2	elemAIP01-2	/sas/elemAIP01-2/AIP/2792 /2188/2294/ark-12148- btv1b9054864m.version0.release0.tar	B00024 74a8b5e689

Processus associés au paquet : [Afficher toutes les actions](#)

3 items trouvé(e)s, affichage des items.

Jeton	Titre	Statut	Date de début	Date de fin
a91ae050-62ad-11e0-b49e-00144feac0aa	ING_1	completed	2011/04/09 15:31:22	2011/04/09 15:39:29
b832bf80-62ae-11e0-a473-00144feac0aa	STO_1	completed	2011/04/09 15:38:57	2011/04/09 15:39:29
cafc2a70-62ae-11e0-8dfb-00144feac0aa	DM_2	completed	2011/04/09 15:39:29	2011/04/09 15:39:29



Thanks for your attention

Questions?

Thomas Ledoux
thomas.ledoux_AT_bnf.fr

{ BnF