



Traitement de la vidéo dans la base de données XML native XediX TS XediX

Didier Courtaud
Pierre Brochard

CEA DAM Ile-de-France

Courriels : didier.courtaud@cea.fr
pierre.brochard@cea.fr

XediX : la genèse au CEA DAM Ile-de-France



● Besoin de conservation de connaissances

- Expériences passées
- Fond hétérogène
 - ◆ Papier
 - ◆ Textes électroniques sous différents formats (Word 6, Word 97, PDF)
 - ◆ Vidéos
- Volume important : de l'ordre du Téraoctet !

● Besoin de lier document source - document électronique

- Gérer des copies d'écran issues du papier de manière synchrone avec leurs copies électroniques
- Gérer des documents multimédia (vidéo, images)

● Assurer la pérennité du fond documentaire

● Besoin d'accès à la granularité la plus fine

XediX



● La plateforme XediX est l'association de :

- Un filtre d'import qui valide les documents vis-à-vis de leur DTD déclarée
 - ◆ XediX gère n'importe quelle DTD
- Un gestionnaire de base qui assure le stockage physique des éléments XML éclatés après le parsing
 - ◆ Il s'appuie sur les filesystems Unix en sécurisant l'accès
 - ◆ Possibilité d'utiliser une base répartie sur plusieurs nœuds d'une grille
 - ◆ Possibilité d'utiliser simultanément des systèmes hétérogènes

XediX (II)



- D'un moteur de recherche
 - ◆ ReX autorisant la recherche par expression régulière sur des structures XML identifiées par XPath
 - ◆ Exemple :
 - Recherche de tous les auteurs dont le nom commence par D comme premier mot du champ AUTEUR
 - ▲ `^D* <DANS> AUTEUR`
 - ◆ Framework autorisant le branchement d'autres moteurs
- Une interface Web de consultation
 - ◆ Conversion dynamique des documents XML en HTML paramétrable pour chaque structure de document (DTD)
 - ◆ Indépendant du navigateur Web : Mozilla, Netscape, IE, ...
 - ◆ Export en XML
 - A différents niveaux de granularité
 - ▲ tout le document
 - ▲ un fragment XML identifié par sa balise
 - A grande vitesse pour *dumper* la base

XediX (III)



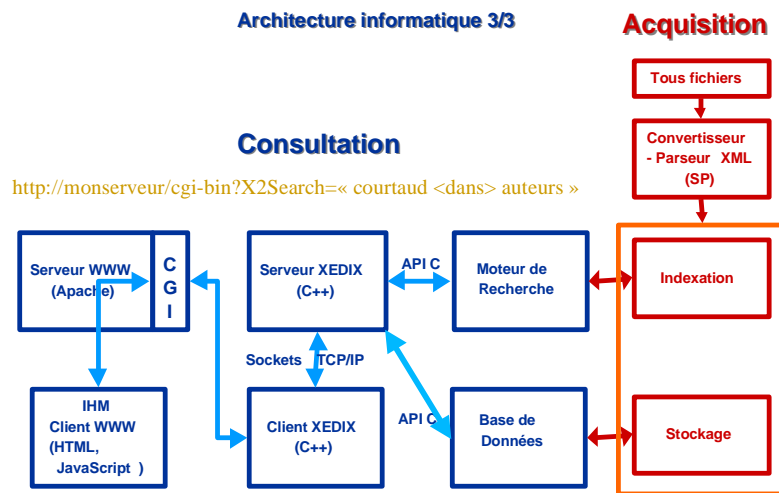
- **Gestion intégrée des droits d'accès**
 - ◆ **Au niveau des utilisateurs**
 - Nommé individuellement
 - Par groupe constitué
 - ◆ **Au niveau des documents**
 - Par identifiant de document
 - Par classes de documents
 - ◆ **Pris en compte**
 - De *ranges* : 1, 2-8, 11-, -5
 - Masquage total
 - ▲ *Des documents*
 - ▲ *Des parties de document*

XediX (IV)



- **Gestion des rétroconversions (*legacy*)**
 - ◆ Mise en base de la version XML ET de la source
 - ◆ Synchronisation de la vue XML et de la source
- **Gestion du multimédia**
 - ◆ Gestion fine des vidéos (dans la base)
- **Personnalisation de l 'interface, du transcodage en HTML DTD par DTD**
 - ◆ Par fichier de configuration
- **Un paquetage de procédures d 'import et de transformation en XML**
 - ◆ Possibilité d 'import par lot de grands volumes
 - ◆ Possibilité d 'import off-line par ligne de commande

Architecture de XediX



Architecture de XediX (II)



- **Possibilité de dialoguer avec la base au travers d'appels au client XediX**
 - Par script
 - Par langage de programmation
- **Possibilité de constituer des documents dynamiques**
 - Par encapsulation de requêtes au client XediX au sein des documents

La vidéo : pourquoi



● Fond documentaire nativement en vidéo

- Films d'expériences
- Films de communication

● Production des méthodes de captation des connaissances

- Interviews des experts
- Mise en place de grains de connaissance

● Nécessité

- D'une identification fine
 - ◆ Permettant de repérer des plans individuels
 - ◆ Susceptibles d'être recherché
- D'une sécurisation forte

La vidéo : comment



● Toutes les vidéos sont physiquement stockés dans la base

- Sécurité
- Portabilité

● On leur affecte un URI (Uniform Resource Identifier)

- Atteignables par cet URI comme n'importe quel élément XML

● L'utilisation de formats normalisés de la famille MPEG permet d'isoler des plans et de travailler au niveau des VOBs (Video Blocks)

- Possibilité d'isoler des séquences correspondant à un grain de connaissance



Démonstration par Pierre Brochard

Conclusion



- **XediX est une implémentation mature du concept de NXD pour gérer des données /documents**
 - Permettant une gestion aisée de très grands volumes de données
 - Autorisant une gestion fine des documents multimédia
 - Offrant de nombreux services pour une utilisation industrielle de documents ou de données au format XML
 - Pouvant facilement se coupler à d'autres produits

- **Commercialisée par la société AM2 Systems**
 - Président : Alain Michard
 - Site WWW d'AM2 Systems : <http://www.am2systems.com>



Questions ?

didier.courtaud@cea.fr
pierre.brochard@cea.fr

<http://www.am2systems.com>