

# Evolution des Serveurs de Données vers XML

Georges GARDARIN  
Laboratoire PRiSM, UVSQ  
Versailles, France

# Plan

- Serveurs patrimoniaux
- Serveurs relationnels
- Serveurs objet-relationnel
- Serveurs XML
- Intégration et Médiation
- Et SOA ?

# 1. Réseau et Hiérarchique

- La première génération :
  - BD = Ensemble de fichiers reliés par des pointeurs
  - Modèles hiérarchique et réseau
  - Langages navigationnels (Codasyl DML, IMS DL1)
- Dates clés
  - 1966 : IDS.I, IMS.I
  - 1971 : RECOMMANDATIONS DU CODASYL
  - 1978 : CODASYL REVISITE



# Avantages et Inconvénients

- Modélisation assez riche
- Performance des accès
- Langage procédural intégré à la programmation
- Très proche du physique
- Manque d'indépendance programme-données
- Pas de langage déclaratif
- Evolution de schéma difficile

# 2. Le Relationnel

- 2e GENERATION DE SGBD
- LANGAGE SQL NON PROCEDURAL
  - 1970 : THE RELATIONAL MODEL (CODD- ACM)
  - 1974 : DEBUT DES PROJETS R et Ingres
  - 1980 : COMMERCIALISATION SQL/DS, Oracle et Ingres
  - 1986 : NORMALISATION DE SQL1
  - 1989 : EXTENSION SQL1 (INTEGRITE)
  - 1992 : NORMALISATION DE SQL2 (COMPLETUDE)

# Avantages et Inconvénients

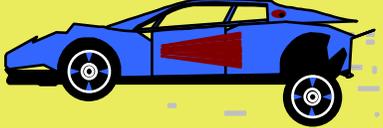
- Un modèle simple et facile
  - Table, Lignes et Colonnes
- Langage assertionnel puissant
  - SQL2
- Une théorie bien assise
  - Relation, Logique
- Le standard des données
  - Depuis 1986
- Gestion et décisionnel
- De nombreux produits
- Couplage LP un peu difficile
  - Impédance mismatch
- Ne supporte pas les données complexes
  - Table en 1NF
- Peu de sémantique
  - Contraintes

# 3. Objet-Relationnel

- 1990 : The Object-Oriented Manifesto
- 1991 : The Third Generation Manifesto
- 1996 : L'objet-relationnel s'impose dans l'industrie
- 1999 : Normalisation du langage SQL3
- Fonctionnalités :
  - extension du relationnel avec les concepts de l'objet
  - support de types de données abstraits extensibles
  - extension de SQL pour le support des objets
  - intégration des règles (triggers)

# Extension des tables relationnelles

- Peuvent contenir des attributs de type complexe
- Peuvent être des tables d'objets

Table Assurance			
Police#	Accident#	Rapport	Photo
24	134		
24	219		
24	037		

# Avantages et Inconvénients

- Un modèle riche dérivé du génie logiciel et des BD
- Garde les avantages du relationnel
- Un modèle très complet
- Un modèle extensible
- Un standard de requêtes
  - SQL3 encore appelé SQL99
- Un modèle complexe
- La mariage de la carpe et du lapin
- Un langage de requêtes peu élégant et complexe (SQL3)
- Les types complexes ne sont pas standardisés

# 4. XML: Rappels



```
<Vin>
  <Cru>Volnay</Cru>
  <Commentaire>Un des plus fameux</Commentaire>
  <Region>
    <Nom>Bourgogne</Nom>
    <Pays>France</Pays>
  </Region>
  <Millésime>1995</Millésime>
  <Millésime>1996</Millésime>
  <Prix Unité="Euro">10.5</Prix>
</Vin>
```

*Éléments balisés et attributs sont la bases*

*Les éléments peuvent être imbriqués et répétitifs*

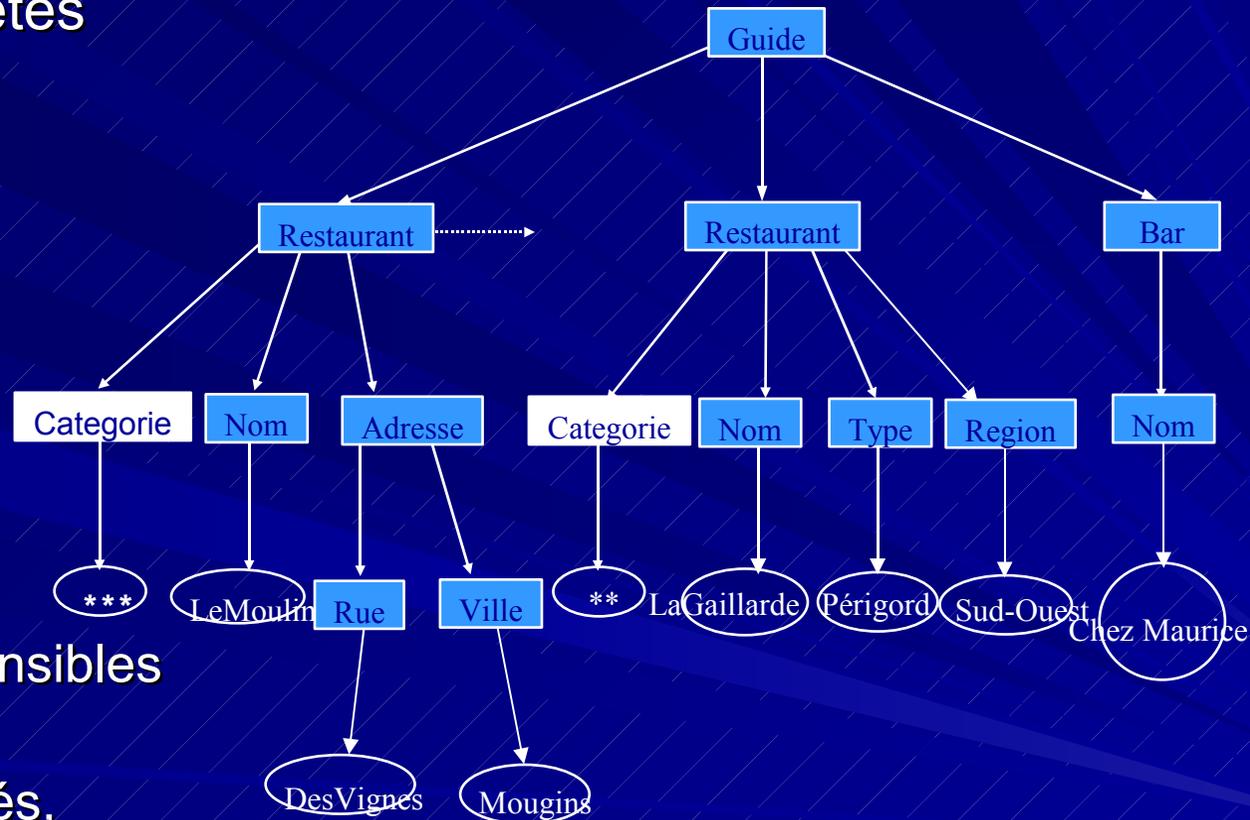
*Une grammaire des balises peut être imposée (DTD, Schema)*

*Le vocabulaire des balises peut être issu d'un référentiel*

# Modèle de données XML

## ■ Graphe à nœuds étiquetés

- Nœuds internes
  - balises et attributs
- Nœuds feuilles
  - données
- Arcs
  - Imbrications, hyperliens



## ■ Schéma XML

- Types simples variés
- Types complexes extensibles
  - séquence, choix, tas
- Contraintes (cardinalités, patterns)

## ■ Le schéma est optionnel

# Langage d'interrogation pour XML

## ■ SQL/XML

- Extension de SQL avec XPath

## ■ XQuery

- Proposition du W3C

## ■ Caractéristiques essentielles

- Documents et collections
- Langage fonctionnel
- Expressions de chemins XPath
- Expressions FLWR
- Tris et agrégats
- Expression conditionnelles
- Fonctions externes

```
for $R in
    doc("Guide.xml")/Restaurant
where $R/@categorie = "****"
return
<Resultat>
    <Nom> {$R/Nom} </Nom>
    <Adresse> {$R/Adresse}
    </Adresse>
</Resultat>
```

# La puissance de XQuery

**Puissance  
de SQL**

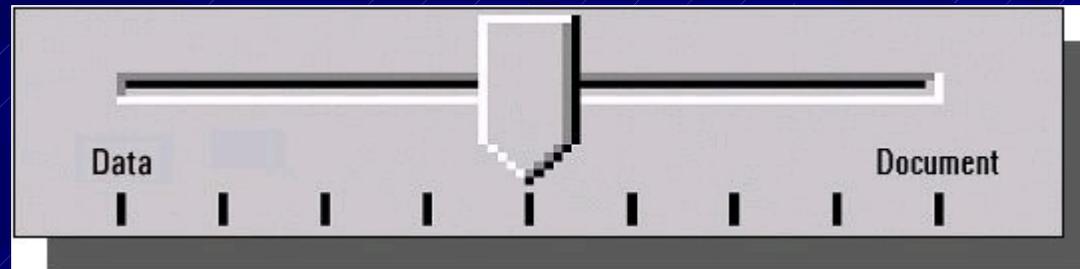
**Langage  
Fonctionnel**

Types de données extensibles  
(XML Schema)

**Structure  
d'arbre**

**Recherche  
plein texte**

# Avantages : XML est partout !



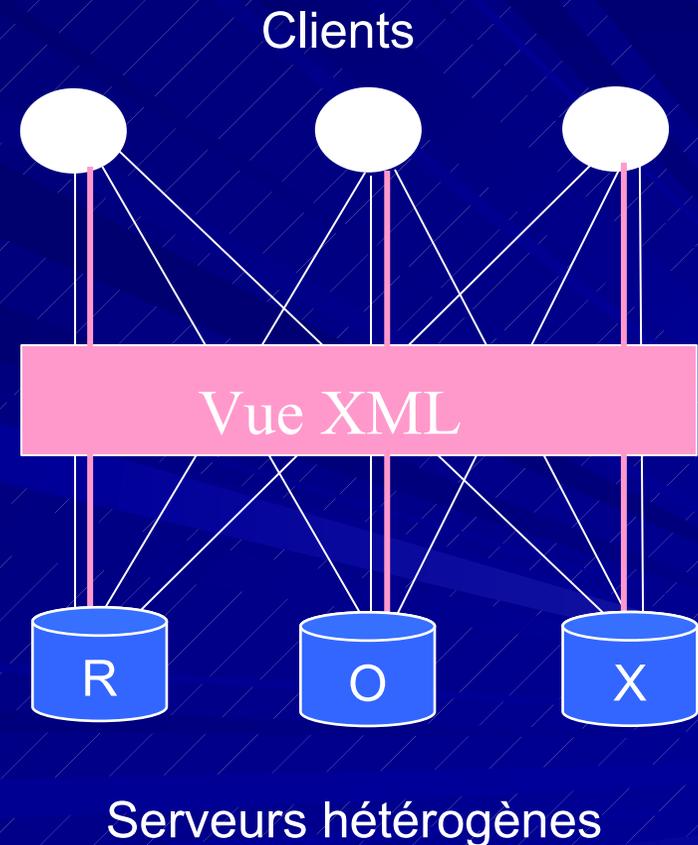
- Structuration forte
  - Tables en bases relationnelles
  - Mises à jour en place
  - Intégrité référentielle
  - SQL
  - Indexation exacte
  - OLTP et OLAP
  - Grandes bases
- Structuration faible
  - Objets complexes
  - Versions
  - Liens hypertexte
  - Recherche plein texte
  - Indexation approchée
  - Documentaire
  - Très grandes bases

# Requêtes sur structure et contenu

- Recherche sur mots et phrases dans des éléments
  - Trouvez tous les livres avec la phrase "volcan d'Auvergne" et le mot "pouzzolane" dans le résumé.
- Recherche via thésaurus
  - Trouvez tous les livres avec des mots phonétiquement similaires à "pouzzolane".
  - Vérifier si les articles de Georges Gardarin incluent des travaux sur l'optimisation de requêtes XQuery et d'autres langages.
- Requêtes avec classement par pertinence
  - Trouvez les trois livres les plus complets sur XML et XQuery.

# 5. Intégration: Médiation d'information

- L'intégration d'information hétérogène nécessite un modèle pivot (global, fédérateur, d'échange)
- Le relationnel a d'abord été utilisé
- XML est conçu pour cela
- XQuery est le langage de requêtes associé
- Il permet l'interrogation de vues XML intégrées



# XML et SOA

- Le rôle de XML est central
  - Toutes les données/messages sont codés en XML
  - Toutes les métadonnées sont aussi en XML
    - Schema : définition des types de messages
    - RDF : description des ressources
    - OWL : définition d'ontologies
    - BPEL: coordination des business process
- Des outils pour construire des Web services
  - Serveurs de données XML
  - Médiateurs de données XML

# 6. Conclusion

- Les modèles de données sont multiples et variés
- XML est LE standard émergeant pour l'intégration de données en SOA
- Les grands SGBD intègrent relationnel, objet et XML
- La médiation propose une technologie d'intégration de données hétérogènes basée sur XML
- XQuery vise à devenir le langage de requêtes universel pour les BD XML et le Web (Sémantique) ...
- Le Web Sémantique devrait apporter la couche sémantique nécessaire pour intégrer les données selon des ontologies