

## Extensions à la spécification XML

*Le "plus petit cercle" XML*

Yves Bekkers

Mise à jour : 9 octobre 2003

## Spécification d'XML

- Deux niveaux de spécification

Document XML valide				Niveau logique
DTD	Schéma	Espace de nom	Xlink	
Document XML bien formés				Niveau physique
Unicode	Balises	Entités		

Xpath - Yves Bekkers - IFSIC

2

## Les normes incontournables

- Les schéma de données (DTD et XMLSchema)
- Les espaces de noms : XML namespace
- L'inclusion de documents : XInclude
- Résolution d'URL relatives : XML Base
- Langage d'adressage d'arbre : XPath
- Extension des URLs : XPointer
- Liens entre ressources : XLink
- Langage de transformation : XSL (XSLT, XSLFO)
- Langage d'interrogation : XQuery
- Programmation XML : DOM, SAX, JDOM, JAPX ...

Xpath - Yves Bekkers - IFSIC

3

## Espaces de noms

Xpath - Yves Bekkers - IFSIC

4

## Motivations

- Faire cohabiter deux dialectes XML dans un même document

```
<?xml version="1.0" encoding="ISO-8859-1"?>
<facture>
  <client>
    <nom>Ducroc</nom>
    <prénom>Jean</prénom>
  </client>
  <produits>
    <produit>
      <nom>Crayon</nom>
      <qte>100</qte>
    </produit>
  </produits>
</facture>
```

← Dialecte client

← Dialecte Produit

Ambiguïté sur l'élément nom

Xpath - Yves Bekkers - IFSIC

5

## Le problème et sa solution

- Le problème
    - Un espace de nommage unique !
  - La solution
    - Qualifier les noms d'éléments avec un nom unique (universel) : Une URI par exemple
- ```
<{http://www.ss2i.fr/client}nom>
```
- URI qualifiante      Nom local
- Mise en œuvre
    - Déclaration de préfixes

Xpath - Yves Bekkers - IFSIC

6

## URI - URL

- URI Uniform Resource Identifier
  - Moyen d'identifier une ressource
  - Ce n'est pas une localisation
- URL Uniform Resource Locator
  - Localisation physique d'une ressource sur Internet
- Une URL peut être utilisée comme URI mais pas l'inverse

## Changer l'espace de noms par défaut

```
<?xml version="1.0" encoding="ISO-8859-1"?>
<facture>
  <client xmlns="http://www.ss2i.fr/client">
    <nom>Ducroc</nom>
    <prénom>Jean</prénom>
  </client>
  <produits>
    <produit xmlns="http://www.ss2i.fr/produit">
      <nom>Crayon</nom>
      <qte>100</qte>
    </produit>
  </produits>
</facture>
```

- Signification : les noms d'éléments *sous la portée de la déclaration* sont qualifiés par la valeur de l'attribut

## Portée d'une déclaration d'espace de noms

- La portée est lexicale, sont concernés :
  - L'élément contenant la déclaration
  - Les éléments contenus par l'élément déclarant
  - Les noms d'éléments mais aussi les noms d'attribut
- L'espace de nom peut être changé à nouveau par n'importe quel descendant

## Expliciter des espaces de noms

- On utilise des préfixes

```
<?xml version="1.0" encoding="ISO-8859-1"?>
<facture xmlns:cl="http://www.ss2i.fr/client"
  xmlns:pr="http://www.ss2i.fr/produit">
  <cl:client>
    <cl:nom>Ducroc</cl:nom>
    <cl:prénom>Jean</cl:prénom>
  </cl:client>
  <produits>
    <pr:produit>
      <pr:nom>Crayon</pr:nom>
      <pr:qte>100</pr:qte>
    </pr:produit>
  </produits>
</facture>
```

Trois espaces de noms  
• par défaut (sans préfixe)  
• `xmlns:cl`  
• `xmlns:pr`

Les identificateurs de préfixes sont libres ...

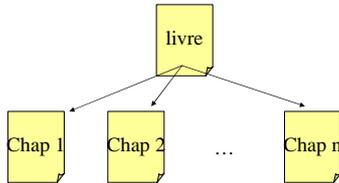
## Avantages des espaces de noms

- Résolution des conflits de noms
- Encourage la réutilisation et la modularité des définitions de dialectes XML

## Construction modulaire de documents

## Le problème

- Construire un document par réutilisation d'autres documents



Xpath - Yves Bekkers - IFSIC

13

## Héritage de SGML

- On utilise les entités (un système de macro)

```
<?xml version="1.0" standalone="yes"?>
<!DOCTYPE livre [
  <!ELEMENT livre (html*)>
  <!ENTITY chapitre1 SYSTEM "cours1/cours.xml">
  <!ENTITY chapitre2 SYSTEM "cours2/cours.xml">
]>
<livre>
  &chapitre1;
  &chapitre2;
</livre>
```

- Le problème : nécessite une dtd
- Éviter cette solution autant que possible ...

Xpath - Yves Bekkers - IFSIC

14

## XInclude

- Directive d'inclusion `xi:include`

```
<?xml version="1.0" standalone="yes"?>
<livre xmlns:xi="http://www.w3.org/2001/XInclude">
  <xi:include href="cours1/cours.xml" />
  <xi:include href="cours2/cours.xml" />
</livre>
```
- Il s'agit d'une substitution d'arbre qui est faite juste après la compilation du document principal
  - Le document source principal et le résultat de la substitution sont équivalents
  - Le compilateur doit faire appel à un processeur XInclude
- Attention : les boucles sont interdites !

Xpath - Yves Bekkers - IFSIC

15

## XInclude – espace de nomage

- Nom officiel de l'espace de noms XInclude  
`http://www.w3.org/2001/XInclude`
- Identificateur de préfixe - généralement : `xi`

Xpath - Yves Bekkers - IFSIC

16

## Outils XInclude

Rubrique "*nous avons essayer pour vous*"

- XInclude Engine [[lien](#)]
  - Un ensemble de classe java pour interpréter XInclude  
`http://sourceforge.net/projects/xincluder/`
- Ant XInclude task [[lien](#)]
  - Une tâche XInclude pour Jajarta Ant  
`http://people.freebsd.org/~znerd/xincluder-task/`

Xpath - Yves Bekkers - IFSIC

17

## XInclude – attribut href

- XInclude permet de spécifier un document en entier ou une partie de document
  - Attribut `href`
- Exemple

```
<xi:include href="monFic.xml#partiel" />
```
- On utilise une extension des URL appelée XPointer (*voir plus loin*)

Xpath - Yves Bekkers - IFSIC

18

## XInclude – attribut parse

- XInclude permet de spécifier le type de traitement en entrée
  - Attribut `parse`
  - Deux traitements sont possibles
    - XML traitement par défaut
    - Texte
- Exemple

```
<xi:include href="monFic.xml" parse="text" />
```

## XInclude – Attribut encoding

- XInclude permet de spécifier le type d'encodage du document
  - Attribut `encoding`
- Exemple

```
<xi:include href="monFic.xml" encoding="UTF8" />
```

## XInclude – valeur par défaut

- XInclude permet de donner une valeur par défaut dans le cas où la ressource est absente

```
<xi:include href="monFic.xml">  
  <xi:fallback>  
    ...  
  <xi:fallback>  
</xi:include>
```

## XLink

## Problèmes avec les liens HTML

- Les liens HTML sont binaires unidirectionnels



- Ils ne sont pas assez génériques
  - Leur sémantique est celle des liens hypertexte et elle est associée à celle des éléments `<a>` et `<img>`
- Une ancre doit être placée sur la destination
  - Une cible ne peut pas être définie de l'extérieur

## Le liens dans les documents XML

- XLink Une généralisation des liens HTML
  - Ils ne sont pas limités aux liens hypertexte
    - Ils permettent de représenter des liens abstraits tels que des liens de parenté, des liens de propriété ...
  - Ils sont plus souples
    - Multidirectionnels
    - Pas limités aux relations binaires
    - Ils utilisent XPointer pour repérer les ressources

## Lien étendu un exemple

```
<facturation xmlns:xlink="http://www.w3.org/1999/xlink">
  <facture xlink:type="extended">
    <leClient xlink:type="locator" xlink:href="#p1"
      xlink:label="client"/>
    <leProduit xlink:type="locator" xlink:href="#prod1"
      xlink:label="produit"/>
    <aAcheté xlink:type="arc" xlink:from="client"
      xlink:to="produit"><qte>100</qte></aAcheté>
  </facture>
  <client id="p1">
    <nom>Ducroc</nom>
    <prénom>Jean</prénom>
  </client>
  <produits>
    <produit id="prod1">
      <nom>Crayon</nom>
    </produit>
  </produits>
</facturation>
```

Élément de liaison

Xpath - Yves Bekkers - IFSIC

25

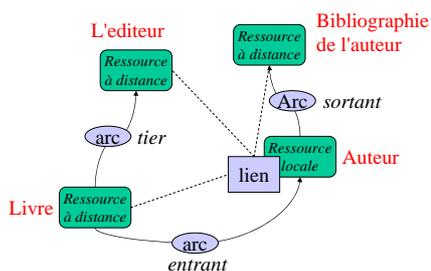
## Les liens de XLink

- **Lien XLink**
  - une relation explicite entre 2 ou plusieurs ressources, locale ou à distance
  - Il peut être caractérisé
- **Élément de liaison**
  - Un élément qui définit l'existence d'un lien et décrit les caractéristiques du lien
- **Repère (locator)**
  - L'identification d'une ressource (à distance) qui participe à un lien
- **Arc**
  - Un élément de liaison définit un ensemble d'arcs traversables entre les ressources

Xpath - Yves Bekkers - IFSIC

26

## Les arcs définis par un élément de liaison



Xpath - Yves Bekkers - IFSIC

27

## Types de liens

- Une ressource *locale* vient avec l'élément de liaison les autres sont à *distance*
- Type d'un élément de liaison
  - Sortant "*Outbound*" : la source est la ressource locale
  - Entrant "*Inbound*" : la destination est la ressource locale
  - Tier "*Third-party*" : aucune des ressources n'est locale
- Deux versions
  - Étendu (*extended*) la forme générale
  - Simple Une version restreinte des liens étendus
    - Valable seulement pour les liens binaires unidirectionnels (tels que ceux de HTML)

Xpath - Yves Bekkers - IFSIC

28

## XLink comment ?

- XLink est un ensemble d'attributs dans un espace propre
  - xlink:type, xlink:href, xlink:label, ...
- Ils permettent de décorer les éléments de l'utilisateur
- Nom officiel de l'espace de noms XLink
  - http://www.w3.org/1999/xlink
- Préfixe recommandé xlink :

Xpath - Yves Bekkers - IFSIC

29

## Lien étendu un exemple

```
<facturation xmlns:xlink="http://www.w3.org/1999/xlink">
  <facture xlink:type="extended">
    <leClient xlink:type="locator" xlink:href="#p1"
      xlink:label="client"/>
    <leProduit xlink:type="locator" xlink:href="#prod1"
      xlink:label="produit"/>
    <aAcheté xlink:type="arc" xlink:from="client"
      xlink:to="produit"><qte>100</qte></aAcheté>
  </facture>
  <client id="p1">
    <nom>Ducroc</nom>
    <prénom>Jean</prénom>
  </client>
  <produits>
    <produit id="prod1">
      <nom>Crayon</nom>
    </produit>
  </produits>
</facturation>
```

Élément de liaison

Xpath - Yves Bekkers - IFSIC

30

## Analyse de l'exemple

```
<facture xlink:type="extended">
  <leClient xlink:type="locator" xlink:href="#p1"
    xlink:label="client"/>
  <leProduit xlink:type="locator"
    xlink:href="#prod1"
    xlink:label="produit"/>
  <aAcheté xlink:type="arc" xlink:from="client"
    xlink:to="produit"><qte>100</qte></aAcheté>
</facture>
```

- Élément de liaison : `xlink:type="extended"`
- Ressource à distance : `xlink:type="locator"`
- Un mode de traversée : `xlink:type="arc"`

Xpath - Yves Bekkers - IFSIC

31

## Attributs Xlink classés

- Un élément de liaison
  - `xlink:type="extended"`
- Ressources Locale
  - `xlink:type="ressource"`
  - `xlink:label="ID"` ID un identificateur pour les arcs
- Ressource à distance
  - `xlink:type="locator"`
  - `xlink:href="XP"` XP un XPointer
  - `xlink:label="ID"`
- Arc (mode de traversée) :
  - `xlink:type="arc"`
  - `xlink:from="ID"`
  - `xlink:to="ID"`

Xpath - Yves Bekkers - IFSIC

32

## Associer un "comportement" à un arc

- Les arcs peuvent être annotés d'information de *comportement*
  - Il s'agit de comportement abstraits
  - La sémantique est donnée par les applications
- Deux types d'information de comportement
  - Le comment (ou le quoi) : `xlink:show`
    - Valeurs : `embed`, `new`, `replace`, `other`, `none`
  - Quand : `xlink:actuate`
    - Valeurs : `onLoad`, `onRequest`, `other`, `none`

Xpath - Yves Bekkers - IFSIC

33

## XInclude et XLink

- `xi:include` de XInclude est différent d'un liens XLink avec l'attribut `show="embed"`
  - XLink ne définit pas un modèle de traitement
    - Demande l'inclusion d'une ressource graphique généralement
    - Est là pour faciliter la recherche d'une ressource aux outils de traitements
  - XInclude spécifie la transformation d'un arbre XML en un autre arbre XML

Xpath - Yves Bekkers - IFSIC

34

## XPointer

- Adresser une partie de document à partir d'un autre document
  - Poser des étiquettes dans le document cible
  - `<xi:include href="monFic.xml#partiel"/>`
  - Utiliser XPointer une extension des URL avec XPath qui permet de ne pas poser d'étiquette dans la cible
  - Exemple les 5 premières section d'un article
- C'est utilisé par XLink et XInclude

```
<xi:include
  href="monFic.xml#XPointer(article/section[position
    ()<=5])"/>
```

Xpath - Yves Bekkers - IFSIC

35

## Repères XPointer

- Exemple de repère XPointer
  - `href="monFic.xml#partieadroite"`
- La partie droite peut être
  - Une valeur d'attribut ID
  - Une séquence d'entiers qui dénote le chemin dans l'arbre en partant de la racine. Par exemple `/1/4/5`
  - Une séquence de la forme `xpointer(E)` où `E` est une expression Xpointer
- Exemples
  - `monFic.xml#part1`,
  - `monFic.xml#/1/4/5`
  - `monFic.xml#xpointer(/ads/ad[notel])`

Xpath - Yves Bekkers - IFSIC

36

## XPointer une généralisation de XPath

- Une expression XPointer est une expression XPath



- XPointer offre des fonctionnalités supplémentaires par rapport à XPath
  - par exemple on peut adresser au niveau du caractère dans un contenu textuel

## Caractères spéciaux dans les repères XPointers

- Certains caractères doivent être échappés, par exemple l'espace %20
- Les caractères spéciaux d'XML tels que &, <, >, " et ' doivent aussi être échappés, par exemple &lt;

## XML Base

- Deux types d'URLs
  - URL absolue  
`http://quelquePart/unFichier.xml`
  - URL relative  
`unFichier.xml`
- Définir une base : attribut prédéfini `xml:base`  
`<aaa xml:base=http://quelquePart>`  
`<bbb href="unFichier.xml" />`  
`</aaa>`
- C'est utilisé dans XLink par exemple

## Conclusion

## Le "plus petit cercle" XML

- Vous avez vu
  - Les schéma de données (DTD et XMLSchema)
  - Les espaces de noms : XML namespace
  - L'inclusion de documents : XInclude
  - Résolution d'URL relatives : XML Base
  - Extension des URLs : XPointer
  - Liens entre ressources : Xlink
  - Langage d'adressage d'arbre : XPath
- Ce que vous devez voir encore
  - Langage de transformation : XSL (XSLT, XSLFO)
  - Langage d'interrogation : XQuery
  - Programmation XML : DOM, SAX, JDOM, JAPX ...