



Apache Cocoon Framework d'applications XML

Sylvain Wallez
Anyware Technologies

www.anyware-tech.com

Sylvain Wallez

- Anyware Technologies
 - Cofondateur, directeur R&D
 - Architecte, expert Java, web et XML
 - Formation, expertise, support
 - sylvain.wallez@anyware-tech.com
- Apache Software Foundation
 - Committer Cocoon, Avalon, Jackrabbit
 - Membre et Vice-Président de la Fondation Apache
 - sylvain@apache.org



→ *Qu'est-ce que Cocoon ?*

- Principes techniques
- La communauté Cocoon
- Qui utilise Cocoon ?

Qu'est-ce que Cocoon ?

Historique

- Origine : génération du site java.apache.org
→ Contenu en XML, formaté avec XSLT
- 1999 : Cocoon 1.0
 - Transformations de DOM
- 2001 : Cocoon 2.0
 - Sitemap : définition centralisée des traitements
 - Pipelines d'événements SAX (performance)
- 2003 : Cocoon 2.1
 - Solution d'applications web complète
 - Gestion de la cinématique par *flowscript*

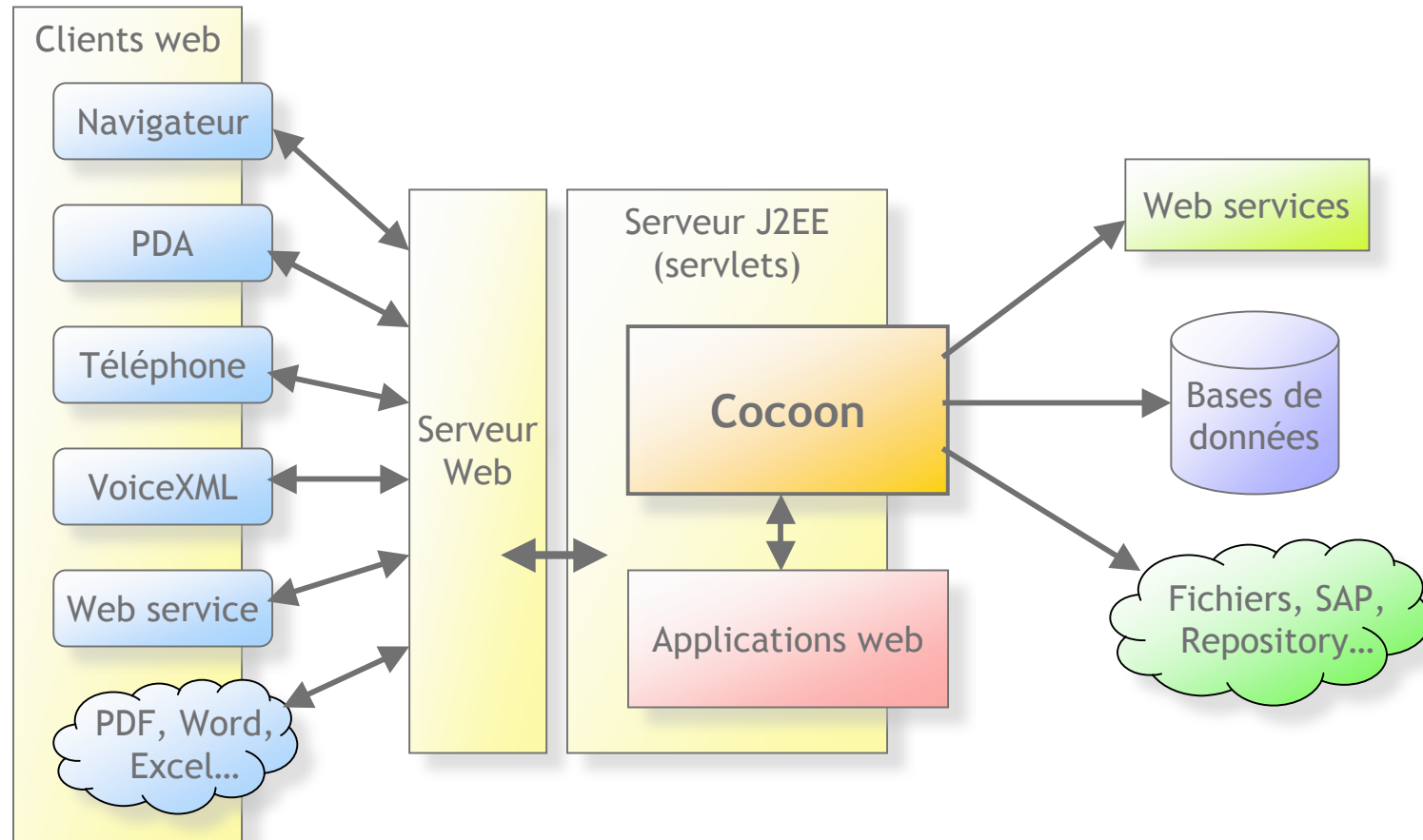
Qu'est-ce que Cocoon ?

Séparation des domaines

- Différents rôles et techniques associées
 - Développeurs : Java
 - Auteurs : XML
 - Designers : HTML, XSLT
- Moyen de la séparation : chaîne de traitement XML
 - Chaque étape de la chaîne a une responsabilité spécifique
- Modèle MVC strict
 - Contrôleur : sitemap et flowscript
 - Définition de la vue : pipeline XML

Qu'est-ce que Cocoon ?

Cocoon et son environnement



- 
-
- Qu'est-ce que Cocoon ?

→ *Principes techniques*

- La communauté Cocoon
- Qui utilise Cocoon ?

Principes techniques

Architecture en composants

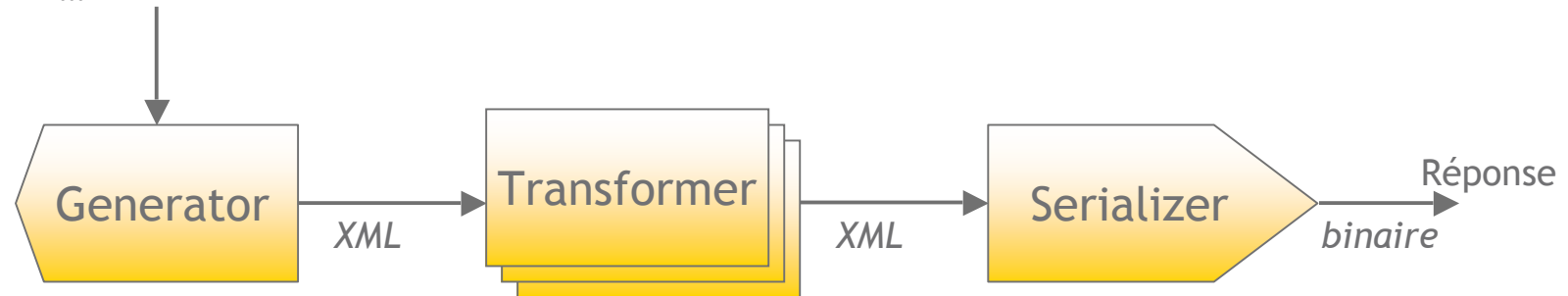
- Cocoon est un ensemble de composants
 - Composant = une interface, des implémentations
 - Forte isolation des composants
 - Grande flexibilité (composant = plugin)
- Support des composants : Avalon
 - Projet Apache (<http://avalon.apache.org>)
 - Infrastructure et services de bases

Principes techniques

Principe des pipelines

Requête :

- http
- soap
- ligne de commande
- ...



- Fichiers
- Bases de données
- Classes métier
- XSP
- Velocity, JSP
- ...

- XSLT
- Moteurs de requêtes
- Internationalisation
- Agrégation
- ...

- HTML
- XML, WML, SVG, SOAP
- PDF, RTF, XLS, MIF
- GIF, JPEG
- ...

Principes techniques

Construction des pipelines : la sitemap

- Exemple : les news en HTML et PDF

```
<map:match pattern="news/*.html">
  <map:generate type="file" src="news-data/{1}.xml"/>
  <map:transform type="xslt" src="news2html.xsl"/>
  <map:serialize type="html"/>
</map:match>
```



```
<map:match pattern="news/*.pdf">
  <map:generate type="file" src="news-data/{1}.xml"/>
  <map:transform type="xslt" src="news2xslfo.xsl"/>
  <map:serialize type="fo2pdf"/>
</map:match>
```



Substitution
du "*" du match

Principes techniques

Les matchers et selectors : routage dans la sitemap

- Matcher : test simple, selector : tests multiples
- Implémentations variées :
 - Pattern sur l'URL demandée
 - Valeurs de paramètres
 - Préférences linguistiques (i18n)
 - Type de navigateur (IE, Netscape, WAP, etc.)
 - Nom du serveur appelé (virtual hosts)
 - Autorisations de l'utilisateur
 - ...

Principes techniques

Selectors: les news en multi-canal

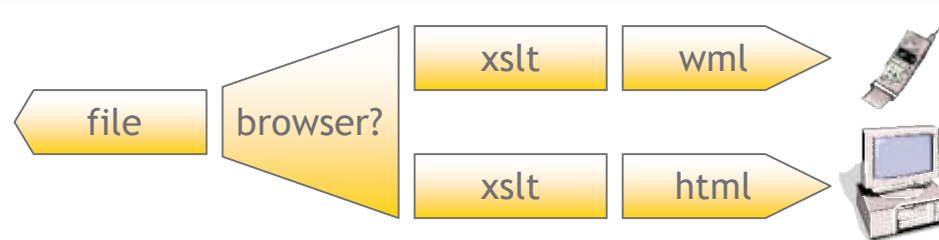
```
<map:match pattern="news/*">
  <map:generate src="news-data/{1}.xml"/>

  <map:select type="browser">

    <map:when test="wap">
      <map:transform type="xslt" src="news2wml.xsl"/>
      <map:serialize type="wml"/>
    </map:when>

    <map:otherwise>
      <map:transform type="xslt" src="news2html.xsl"/>
      <map:serialize type="html"/>
    </map:when>

  </map:select>
</map:match>
```



Principes techniques

Les "sources" dans Cocoon

- Notion d'URL étendue

- Sources fichier

```
<map:generate type="file" src="myfile.xml"/>
```

- Sources externes

```
<map:generate type="file" src="http://www.yahoo.com"/>
```

- Sources internes

```
<map:generate type="file" src="cocoon:/pipeline.html"/>
```

Appel récursif de la sitemap

- Sources spécialisées : Blob

```
<map:generate type="file" src="blob:/db/docs/req[id='43']"/>
```

Extrait le contenu de *SELECT req FROM docs WHERE id='43'*

Principes techniques

Générateur : contenu XML initial

- File
 - Toute "source" (URLs étendues)
- JXTemplate
 - Pages serveur nativement XML
- Directory
 - Listing (récursif) du contenu d'un répertoire
- Script
 - Génération par scripts : JavaScript, VB, Perl, Python, etc.
- Servlet
 - Utilisation de n'importe quel servlet
- MP3Directory
 - Directory + infos sur le morceau musical
- ...

Principes techniques

Transformers : transformation du contenu

- XSLT
- I18N (internationalisation)
 - Traduction à l'aide de dictionnaires de termes
- SQL, LDAP
 - Exécution de requêtes "à la volée"
- XInclude
 - Inclusion (portails, documents composites)
- RoleFilter
 - Filtrage par rapport au rôle de l'utilisateur
- ...

Principes techniques

Serializers : production du résultat

- HTML/XML
 - HTML, XML, SVG, WML, VoiceXML...
- SVG
 - Images JPEG dynamiques avec Batik
- PDF
 - Documents Acrobat avec FOP ou iText
- Zip
 - Archives Zip, documents OpenOffice
- ...

Principes techniques

Exemple avancé : agrégation

- Obtenir l'édito de avril 2004 :
→ http://serveur/edito_04_2004.html

```
<map:match pattern="edito_*_.html">

  <map:aggregate element="page">
    <map:part src="cocoon:/menu.xml"/>
    <map:part src="blob:/db/editos/texte[annee='{2}' and mois='{1}']"/>
  </map:aggregate>

  <map:transform type="xslt" src="page2html.xsl"/>
  <map:serialize type="html"/>

</map:match>

<map:match pattern="menu.xml">
  <map:generate type="jxt" src="menugenerator.jx"/>
  <map:serialize type="xml"/>
</map:match>
```

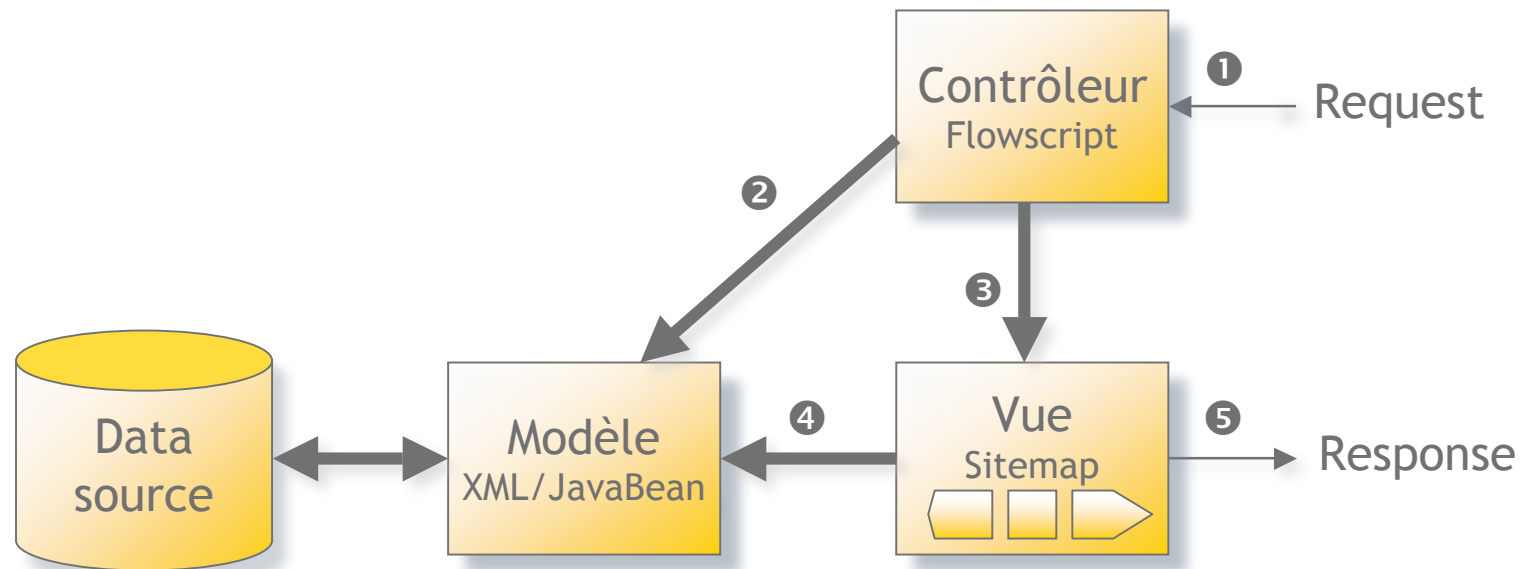
Appel du pipeline
"menu.xml"

Lecture directe en
base de données

Principes techniques

Applications web avec Cocoon

- Contrôleur : flowscript
 - Gestion de la cinématique (enchaînements)
- Vue : sitemap
- Modèle : XML ou JavaBeans



Principes techniques

Flowscript : continuations

- Le script est "gelé" lors de l'envoi d'une page
 - Interaction décrite en code séquentiel
- Simplification énorme (pas d'automate d'état)

```
var cart;
var user;

function checkout()
{
    while(user == null) {
        cocoon.sendPageAndWait("login.html");
        user = UserRegistry.getUser(cocoon.request.get("name"));
    }
    cocoon.sendPageAndWait("creditCard.html");
    var creditCard = cocoon.request.get("creditCard");
    cocoon.sendPageAndWait("confirmOrder.html");
    EnterpriseSystem.placeOrder(user, cart, creditCard);
    cocoon.sendPage("orderPlaced.html");
}
```


Appel de la vue sur
la sitemap et
suspension du script

Appel de la vue
sur la sitemap et
fin du script

Principes techniques

Les "blocks" : solutions thématiques

- Portal : système de portail complet
- CForms : gestion des formulaires de saisie
- Axis : services web SOAP
- Lucene : indexation du contenu et recherche
- Web3 : connexion avec SAP
- WebDAV
 - Source WebDAV : accès à un repository
 - Sitemap WebDAV : Cocoon est le serveur WebDAV
- Mail : webmail complet
- ...

- 
-
- Qu'est-ce que Cocoon ?
 - Grands principes techniques

→ *La communauté Cocoon*

- Qui utilise Cocoon ?

La communauté Cocoon

La Fondation Apache

- Organisation à but non lucratif
- Logiciel Open Source de qualité commerciale
 - httpd : 70% des sites web de la planète !
 - Tomcat : implémentation de référence des servlets
 - Xerces, Xalan : standards dans le monde XML
- "La communauté est plus importante que le code"
 - Un logiciel génial sans utilisateurs est mort
 - Un logiciel moyen avec beaucoup d'utilisateurs va s'améliorer



La communauté Cocoon

Les développeurs

- Committers (accès en écriture au CVS)
 - 60 depuis l'origine
 - 30 actifs
- Des profils variés
 - Visionnaire charismatique, architecte, designer web, enseignant-chercheur, développeur...
 - La plupart dans un cadre professionnel
 - Lié à l'esprit de la licence Apache
- Projet principalement européen

La communauté Cocoon

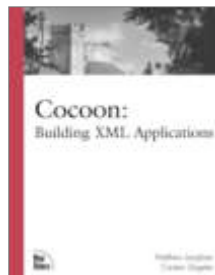
Mailing-lists


- users@cocoon.apache.org
 - Support utilisateur
 - Réponses rapides (parfois dans l'heure!)
 - 1100 inscrits, 1000 à 1500 messages par mois
- users-fr@cocoon.apache.org
 - Liste francophone nouvellement créée (01/2005)
- dev@cocoon.apache.org
 - Là où tout le développement se fait
 - Pas de discussions privées, tout le monde peut participer
 - 500 inscrits, 1500 à 2000 messages par mois

La communauté Cocoon

Autour de Cocoon

- Orixo : alliance européenne autour de Cocoon
 - Anyware en est un des fondateurs
 - <http://www.orixo.com/>
- Le "Cocoon Get Together"
 - Rencontre annuelle des utilisateurs
 - 3ème édition : 140 participants de 17 pays
 - Fait rare dans l'opensource (Linux, Zope, PHP)
- Nombreux articles et conférences
- 3 livres : New Riders, Sams, Sybex



- 
-
- Qu'est-ce que Cocoon ?
 - Grands principes techniques
 - La communauté Cocoon

→ *Qui utilise Cocoon ?*

Qui utilise Cocoon ?

Sites internet

- Plus de 100 sites listés dans les "live sites"
 - Scoot.be (annuaire)
 - Sun-catalogue (vente en ligne)
 - Warner bros Suisse (cinéma)
 - Librairie du congrès US
 - Salon.com
- Beaucoup d'autres sites non listés

Qui utilise Cocoon ?

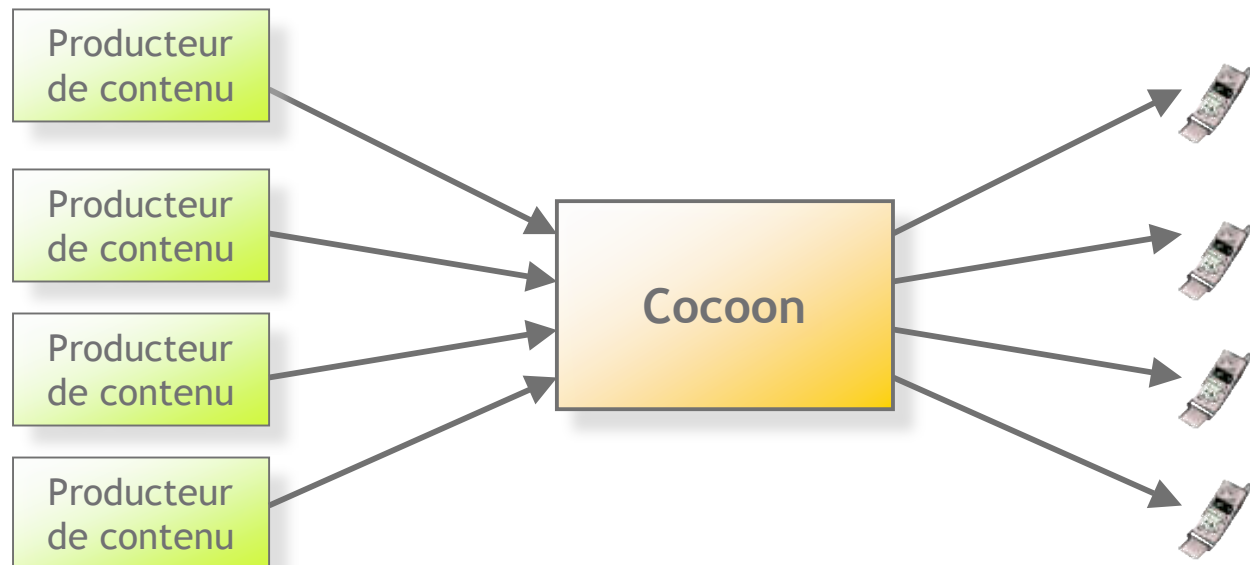
Portails pour téléphones mobiles

- RAI en Italie, 3 opérateurs en Allemagne
- Très forts trafics (Vodafone Live Allemagne)
 - 3,5 millions de requêtes/jour
 - Prise en compte de 90 types de terminaux
 - Temps de réponse moyen : 50 ms

Qui utilise Cocoon ?

Portails pour téléphones mobiles

- Deux missions pour Cocoon
 - Adapter le contenu aux terminaux
 - Agréger les sources de contenu



Qui utilise Cocoon ?

Jouve

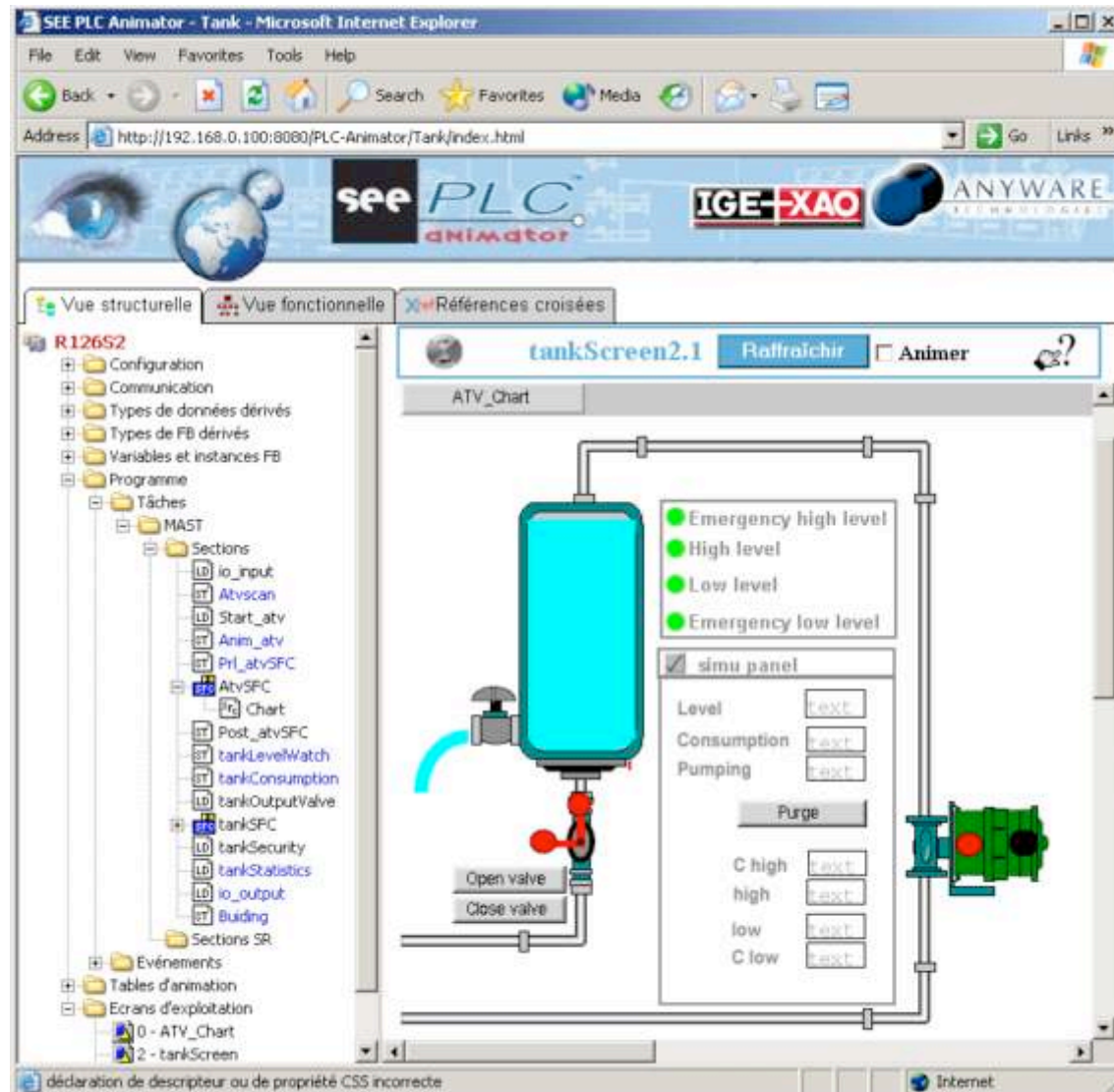
- Grosse gestion documentaire
 - Administrations
 - Universités
 - Grandes entreprises
- Choix stratégique sur Cocoon
 - Premier projet : Office Européen des Brevets
 - Utilisé maintenant sur de nombreux nouveaux projets

Qui utilise Cocoon ?

Schneider Electric

- Édition et gestion de données techniques
 - Saisie de la description des équipements
 - Production de catalogues, d'applications de configuration, etc.
- Supervision de processus industriel
 - SVG dynamique :
 - Programme d'automatisme
 - Diagrammes de supervision
 - Schématique électrique (lien avec la CAO)

Qui utilise Cocoon ?



Qui utilise Cocoon ?

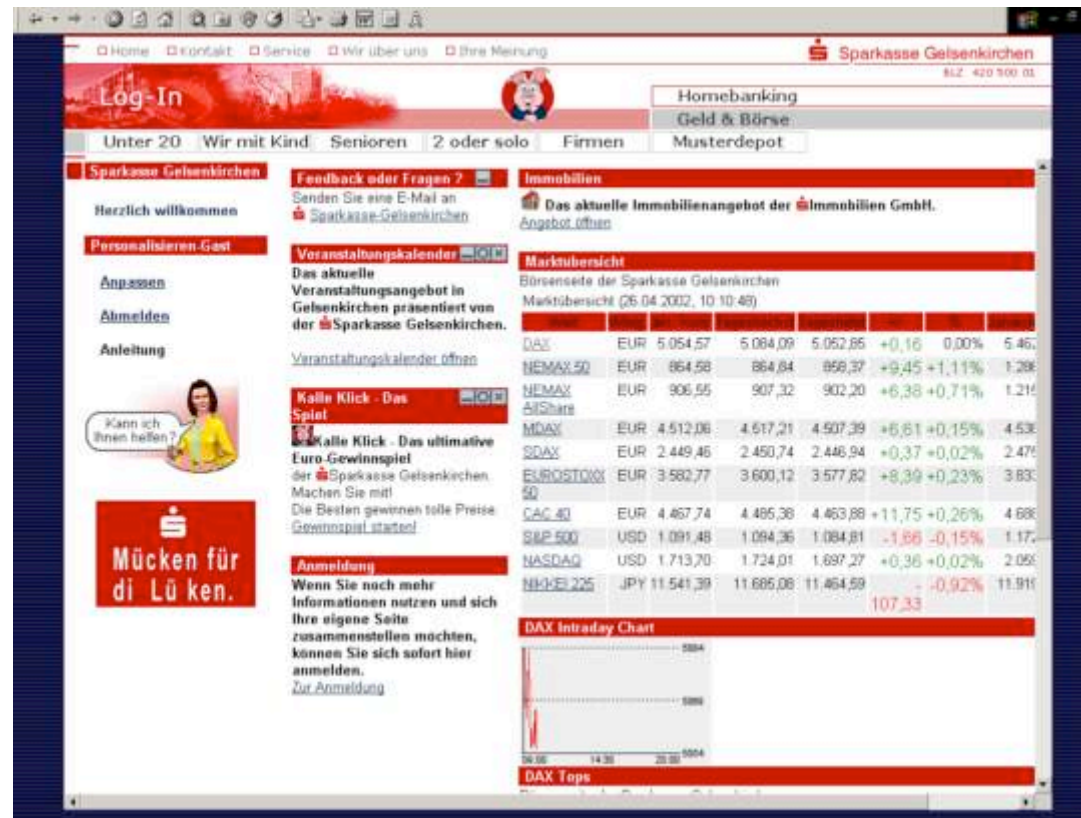
Airbus

- Documents de certification
 - Extraction base de données
 - Inclusion de commentaires HTML
 - Documents PDF de 1500 pages
- Méta-moteur de recherche (knowledge management)
 - Connecteurs vers des bases métiers
 - Traduisent la requête vers le système cible
 - Publient le résultat en XML
 - SQL, Lotus Notes, Fichiers, Web services
 - Agrégation et présentation des les résultats

Qui utilise Cocoon ?

Banque

- Sparkasse Gelsenkirchen
 - Portail Internet des clients



Qui utilise Cocoon ?

Banque

- Swiss Interbank Clearing
 - Échanges interbancaires
 - 75 banques reliées
 - 600 000 000 CHF échangés par jour
 - Fonctionne 7/7 depuis 2 ans
 - Cocoon comme moteur de webservices
 - Passerelle EDIFACT <-> SOAP
 - Étudié par le Gartner Group
 - Un des premiers succès des webservices
 - http://www.gartner.com/DisplayDocument?doc_cd=108916

Conclusion

Cocoon, framework complet

- Publication multi-canal
- Intégration de nombreuses sources de données
- Les blocks, solutions fonctionnelles
- Communauté active
- Nombreuses références industrielles



<http://cocoon.apache.org>