

Introduction aux technologies du Web



Mercredi 12 décembre 2007
Patrice Pilot
<http://www.toulibre.org>



Plan

- Définitions
- Les fondements du Web
- Problèmes et solutions
- Le Web *tendance*

Définitions



Le World Wide Web

- Définition Wikipedia
 - un système de documents hypertextuels, liés entre eux, accédés via le réseau Internet.
- Croisement de technologies ou concepts fondamentaux
 - Internet (au sens large)
 - langages de balisage
 - hypertextualité



Internet

- Réseau basé sur l'utilisation du protocole Internet (Internet Protocol - IP)
 - réseau de machines, pas de documents
 - protocole réseau ?
 - protocole IP ou HTTP ?



Les langages de balisages

- Encodage des caractères
 - les ordinateurs utilisent des valeurs numériques
 - ⇒ différents codes
- Encodage de la mise en page

Première·**phrase**¶
Deuxième·*phrase*¶

```
<page>  
· <ligne>Première <texte style="gras">phrase</texte></ligne>  
· <ligne>Deuxième <texte style="italique">phrase</texte></ligne>  
</page>
```



L'hypertextualité

- S'affranchir de la linéarité du texte imprimé
- années 60, projet NLS de D. Engelbart
 - souris, liens hypertextes



Le World Wide Web

- années 90, formalisation du WWW par T. Berners-Lee et R. Cailliau au CERN
 - 1993, la technologie passe dans le domaine public
 - 1994, fondation du W3C



Interlude pratique

- Quelques erreurs courantes et leur signification
 - Erreur *page avec des hiéroglyphes*
 - Erreur « Serveur non trouvé »
 - Erreur « Le serveur ne répond pas »

Les fondements du Web



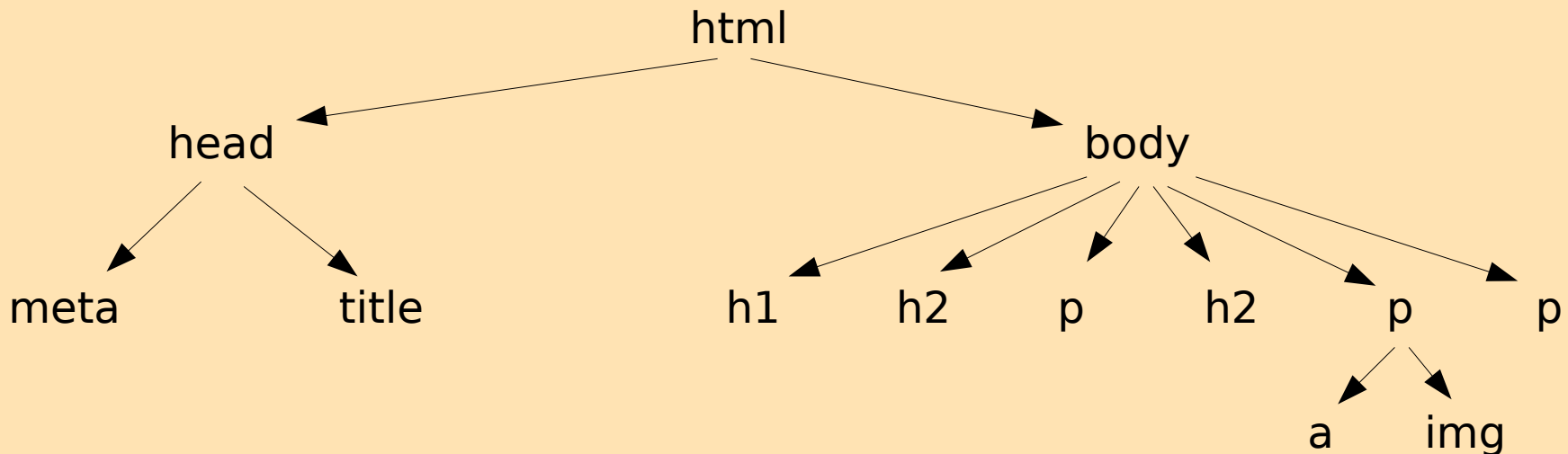
(X)HTML : les balises du Web

- HTML, XHTML, SGML, XML ???????
 - HTML : dialecte SGML
 - XHTML : dialecte XML
- SGML et XML définissent les règles d'écriture des balises (<, >, /, ...)
- HTML et XHTML définissent les noms des balises, leur agencement et leur sémantique,



Les balises (X)HTML

```
<!DOCTYPE html PUBLIC "-//W3C//DTD HTML 3.2//EN">
<html lang="fr_FR">
  <head>
    <meta http-equiv="Content-Type" content="text/html; charset=UTF-8" />
    <title>Découverte du HTML</title>
  </head>
  <body>
    <h1>Découverte du HTML</h1>
    <h2>Introduction</h2>
    <p>Ceci n'est pas une introduction.</p>
    <h2>Histoire</h2>
    <p>Le langage HTML a été inventé au <a href="http://www.cern.ch">CERN</a> par.
    . . . Tim Berners-Lee dont voici la photo : </p>
    <p>La suite plus tard ....</p>
  </body>
</html>
```





HTTP

HyperText Transport Protocol

- Protocole client/serveur
 - Un programme client : le navigateur (fureteur, butineur, explorateur, ...) ou *browser*
 - Un programme serveur : le *serveur* Web
 - un cycle *requête / réponse*
 - des ressources accessibles via des URL :
 - <http://toulibre.org/index.html>
 - requêtes GET vs. POST
 - <http://truc.org/search?mot=Thomas&mot=Alex>



HTTP

HyperText transport Protocol

- Une version sécurisée (HTTPS)
 - important dans les transactions financières
 -
- D'autres usages que le Web *stricto sensu*
 - WebDAV
 - Subversion
 - Feed RSS
 - ...



Interlude pratique

- Encore quelques erreurs courantes et leur signification
 - Erreur « page non trouvée » (404)
 - Erreur *page blanche*
 - gestion des certificats
 - Les requêtes POST et les boutons de navigation

Problèmes et solutions



Un Web avec du *style*

- Balises sémantiques
- Aux origines, formatage par ajout d'attributs
- Feuilles de styles (CSS - Cascading Style Sheets)
 - séparation structure // présentation
 - adaptation en fonction du lecteur



Retrouver la mémoire

- HTTP, un protocole *stateless* (sans état = sans mémoire)
 - Quelques en-têtes particuliers
 - Referer
 - Set-Cookie / Cookie (extension au protocole)



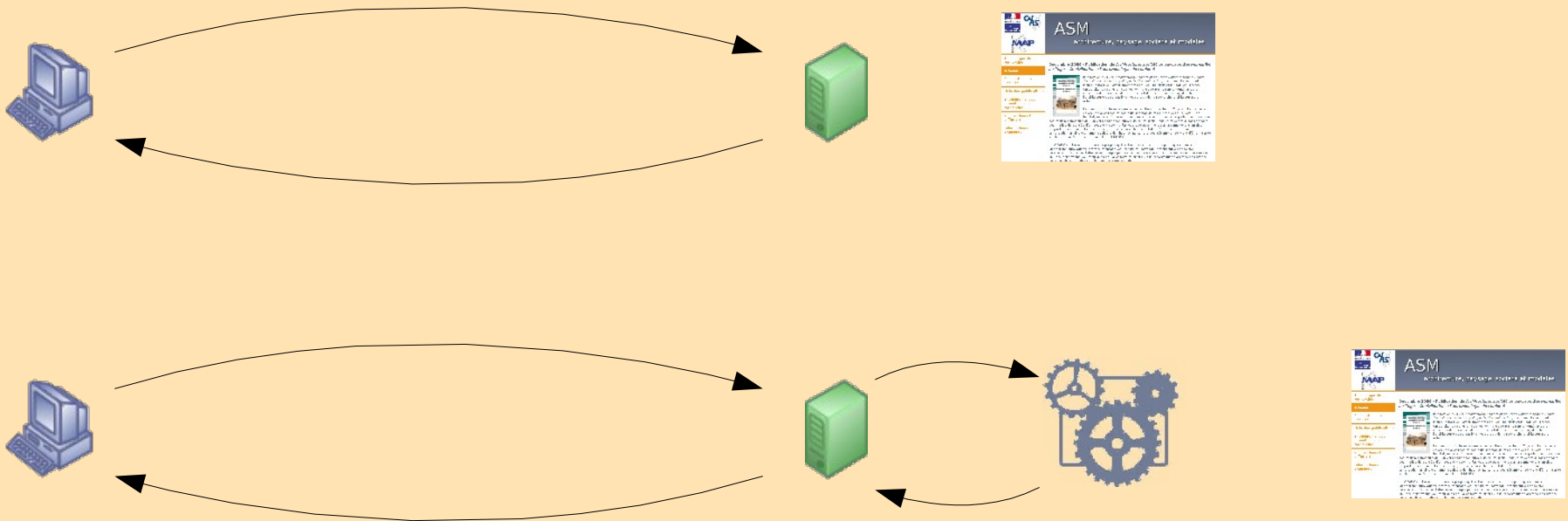
Modularisation des sites

- Diverses techniques côté serveur
 - Server Side Includes (SSI) et utilisation basique des langages de script
 - Génération des pages par des programmes
- Côté utilisateurs, les Frames
 - un document « maître »
 - plusieurs sous-documents
 - posent de gros problèmes



Les *applications* web (1)

- Découplage du serveur et du *générateur* de pages





Les *applications* web (2)

- Une plus grande souplesse de développement
- Différents types d'*applications* web
 - CMS, blogs, wikis, ...
 - Galeries photos
 - Forges
 - Webmail
 - ...
- C'est mieux que du statique ... mais c'est plus dur !



L'interactivité côté client

- Améliorer l'interactivité
 - Manipuler les documents via JavaScript *et al.*
 - Remplacer le contenu du document
 - Applets Java
 - Animations Flash
 - Gérer le multimédia
 - Lecteurs embarqués (Flash, ...)
 - Lecteurs externes (xine, Realplayer, ...)

Le Web à la mode



De nouvelles techniques

- Syndication RSS, Atom, ...
- AJAX
 - l'alliance des frames et du JavaScript
- Frameworks JavaScript
- Les clients riches
 - les documents ne sont plus que des conteneurs de données



De nouveaux usages

- Sites collaboratifs
 - wikis, ...
- Sites « sociaux »
 - galeries photos
 - blogs
 - sites d'aggrégations (Netvibes ...)
 - sites de réseaux sociaux (FaceBook, ...)
- Remote applications



Attention danger

- stockage des données personnelles
 - Flickr, Google, webmails, ...
- publication des données confidentielles
 - FaceBook, LiveSpace, ...