



# Technologies du Web

## Créer et héberger un site Web



## Planification

Choisir une solution d'hébergement

Administration

Développement du site



- **Objectifs** du site
- **Fonctionnalités** :
  - Zones de contenu (nombre de pages, etc.)
  - Accès à des données d'une base existante
  - Fonctionnalités de recherche, de mise à jour, etc.
  - Interactions utilisateur : authentification, commentaires, profils, vente, etc.
  - Contenu multimédia
  - Besoins spéciaux d'interface (jeux, cartographie, etc.)
- Définir l'**arborescence** du site
- **Calendrier** de lancement
- **Public** visé : tous clients Web, navigateurs graphiques récents uniquement, nécessité d'un site spécial ou d'une application pour smartphone, etc.





# Maquette fonctionnelle

- Se baser sur le cahier des charges pour faire une maquette fonctionnelle (mock-up) de ce à quoi ressemblera le site
- Sans programmer :
  - **Structure** principale de chaque type de page
  - Contenu des **formulaires**
  - **Informations** présentées à chaque niveau
- Peut-être fait sur papier, avec des mock-ups HTML rapide, avec un logiciel de dessin vectoriel
- Inutile de se préoccuper des aspects graphiques à ce niveau, mais réflexion sur l'**ergonomie**



# Choix des technologies : en fonction de quoi ?

En se basant sur :

- le cahier des charges
- la maquette
- les compétences à disposition
- les contraintes d'hébergement
- les contraintes externes (p. ex., intégration avec un logiciel tiers)
- le public visé

26 octobre 2012





# Choix des technologies

Choisir :

- Une **plate-forme** (Apache / IIS, Linux / Windows)
- Un format de **stockage des données** (fichiers plats, XML, base Oracle, base MySQL, etc.)
- Une ou plusieurs technologies **côté serveur** (CMS, framework d'applications, langage)
- Les technologies **côté client** (version d'HTML visée, framework JavaScript, technologies pour applications multimédia riches)
- **HTTPS ?**





# Dimensionnement

Calcul approximatif de la **charge** du serveur :

- Nombre de requêtes par seconde (en moyenne, en pleine charge)
- Volume de données transmis par seconde (en moyenne, en pleine charge)
- Calculs coûteux côté serveurs





# Exemple de dimensionnement

## ■ Paramètres :

- 1 000 visiteurs attendus par jour, avec des pointes à 10 000
- Chaque visiteur visite une dizaine de pages en moyenne, et accède à une vidéo en moyenne dans 10% des cas
- Chaque page fait appel (images, scripts, etc.) à une dizaine de fichiers et pèse de l'ordre de 500 kilo-octets
- Une vidéo pèse de l'ordre de 200 méga-octets
- Pas de traitements coûteux côté serveur

## ■ Estimation :

- $\frac{1\,000 \times 10 \times 10}{3\,600 \times 24} = 1,15$  accès par seconde en moyenne
- $\frac{1\,000 \times (10 \times 500 + 0,1 \times 200\,000)}{3\,600 \times 24} = 289$  kilo-octets par seconde en moyenne (soit 2 mégabits/seconde)
- 100 fois plus en charge (10 fois plus de visiteurs certains jours, non répartis uniformément sur la journée)





- Dépôt de **marque** ?
- Constitution d'une **société** ?
- D'où vient le contenu du site ? **Propriété intellectuelle**
- Réfléchir aux **conditions** d'utilisation, de vente, de conservation des informations, etc.
- Déclaration à la **CNIL** en cas de base d'utilisateurs avec informations personnelles



Planification

Choisir une solution d'hébergement

Administration

Développement du site





# Auto-hébergement

Un serveur Web peut être hébergé directement sur un ordinateur de l'entreprise ou d'un particulier connecté à Internet, mais attention aux aspects suivants :

- **Débit** suffisant (en particulier en upload) ?
- Garanties de pouvoir garder le serveur Web allumé **24h/24** ? (refroidissement, interventions en cas de panne, etc.)
- **Visibilité** de l'ordinateur sur Internet : IP publique ? NAT ?





## Recours à un hébergeur

Les hébergeurs (p. ex., OVH, Gandi, Amen, Amazon, Orange, Free) proposent en général trois types d'offre :

- Hébergement Web **mutualisé** : le site est un hôte virtuel hébergé sur la même instance d'Apache ou IIS que d'autres sites. Restrictions d'espace disque, de débit, et de logiciels potentiellement utilisables. Convient pour des sites avec un petit nombre d'accès (moins d'un accès par seconde) et sans contrainte particulière.
- Serveur **virtualisé** : accès complet à une machine virtuelle physiquement située sur la même machine que d'autres serveurs virtuels. Apparaît comme une machine indépendante, mais typiquement moins puissante qu'un vrai serveur dédié. Convient à la plupart des sites.
- Serveur **dédié**. Pour les sites avec gros trafic ou les serveurs avec autre chose que du Web.

26 octobre 2012





# Enregistrement du nom de domaine

- Le nom de domaine désiré doit être **enregistré** (acheté) auprès d'un *registrar* (*Go Daddy, Network Solutions, etc.*)
- La plupart des hébergeurs proposent également ce service
- Attention à la **durée** d'engagement
- Attention à ne pas manquer les dates de **renouvellement**
- Réfléchir au contenu de l'entrée **Whois**. Certains registrars permettent de la masquer en payant un supplément.
- Éventuellement une bonne idée d'enregistrer des noms de domaines proches





# Gestion des noms de domaine

- Besoin de (au moins) deux **serveurs de noms** qui répondront aux requêtes pour le nom de domaine enregistré
- Plusieurs possibilités :
  - le serveur Web lui-même, plus un autre serveur de l'entreprise
  - hébergeurs DNS dédiés, gratuits ou bon marché (FreeDNS, CloudDNS, ZoneEdit, etc.)
  - l'hébergeur Web fournit souvent gratuitement ou bon marché l'hébergement DNS
- Les serveurs de nom de domaine sont ensuite **déclarés** auprès du registrar



En cas de site HTTPS, besoin d'un **certificat SSL** :

- Peut être **autosigné**, mais provoque des alertes du navigateur à la visite
- Peut être acheté à une **autorité de certification** (VeriSign, GeoTrust, DigiCert, etc.)
- Les hébergeurs Web fournissent souvent la possibilité de s'en occuper
- Valide seulement pour un nom de domaine donné !



Planification

Choisir une solution d'hébergement

Administration

Développement du site







# Installation du serveur

- Pas dans le cas d'un hébergement mutualisé
- Choix (et éventuellement achat de la license) de la version du système d'exploitation
- Installation (et éventuellement achat) de tous les logiciels nécessaires :
  - Serveur Web (Apache, IIS)
  - Accès distant (SSH, Remote Desktop, VNC, FTP)
  - SGBD (MySQL, Oracle, PostgreSQL, etc.)
  - Ce qui est requis par le langage côté serveur (module PHP, serveur d'applications Java, etc.)
  - CMS, frameworks côté serveur, etc.
  - Services non Web hébergés par le serveur (p. ex., email)





# Configuration d'Apache : Comment ?

Apache peut être configuré :

- par ses fichiers de configurations **globaux** (sous Linux, dans `/etc/apache2`) : uniquement pour les administrateurs du serveur. Seul endroit où de nouveaux hôtes virtuels peuvent être créés, ou où la configuration SSL peut être effectuée.
- par des fichiers **.htaccess** dans chacun des répertoires du site Web : doit être activé dans la configuration du serveur (`AllowOverride All`) – parfois impossible chez certains hébergeurs mutualisés.





# Configuration d'Apache : Quoi ?

- Jeu de caractère du site : `AddDefaultCharset utf-8`
- Redirections HTTP : `RedirectMatch permanent`
- Fichier d'index d'un répertoire : `DirectoryIndex toto.php`
- Authentification HTTP : `AuthUserFile`
- Négociation de contenu
- Types MIME
- cf. <http://httpd.apache.org/docs/>





# Serveur d'applications Java

Pour JSP, servlets Java, nécessité d'installer un serveur d'applications Java :

- **Tomcat**, JBoss, Geronimo (surcouche de Tomcat) : libres
- WebLogic Server, WebSphere : commerciaux
- Peuvent aussi servir de serveurs Web, ou connectés avec Apache :  
Tomcat se connecte avec Apache via mod\_jk



- Optimiser la base de données (**index**)
- **Optimiser le code** côté serveur
- Réduire le **volume** des pages téléchargées
- Ajouter un **cache** des pages Web les plus demandées :
  - Au niveau du serveur Web (mod\_cache d'Apache)
  - Sur un serveur intermédiaire entre Internet et le serveur Web, servant de proxy
  - En utilisant des sociétés tiers pour héberger certains contenu (p. ex., vidéo) ou pour servir de proxy pour le site tout entier : Akamai, Google Web Accelerator
- **Répartir** la charge entre plusieurs serveurs Web :
  - Plusieurs adresses IP pour le même nom de domaine
  - Une adresse IP pointant vers un **load balancer** qui se contente de rediriger la requête vers de multiples serveurs
  - Possibilité d'optimisations géographiques



# Maintenance régulière

- **Mises à jour** de sécurité :
  - du système d'exploitation
  - du serveur Web
  - des modules du serveur Web et logiciels tiers
  - des CMS et frameworks d'application Web
  - du SGBD
- **Surveillance** régulière de la charge et de l'activité du serveur Web
- Surveillance de la structure de liens du site (pas de liens cassé, internes ou externes)
- **Migrations** ponctuelles : nouveau serveur, nouvelles versions de logiciels, etc.



Planification

Choisir une solution d'hébergement

Administration

Développement du site





# Développement par couches

- **Successivement** (entrelacement possible, mais chaque étape correspond à une progression logique dans le développement) :
  - Conception de la base de données
  - Développement de la logique de l'application côté serveur (modèle)
  - Développement des vues et du contrôleur côté serveur ; possible annotation sémantique dans les vues
  - Conception graphique, CSS
  - Interactions JavaScript
  - Contenu multimédia riche
  - Méta-informations : `robots.txt`, sitemap
- À chaque étape :
  - Démarche **qualité** : validation W3C, accessibilité, etc.
  - Tests utilisateurs, retours, pour infléchir le développement au besoin





- S'assurer de résister à la charge
- Basculement en production
- Marketing
- Surveiller l'activité très attentivement les premiers jours pour être prêts à corriger : performance, bugs, ergonomie, etc.



- S'assurer que le site est accessible aux robots avant le lancement
- Respecter les bonnes pratiques (cf. cours moteur de recherches)
- Marketing pour créer des liens vers le site
- Publicité ciblée : dans les moteurs de recherche, sur les sites tiers





# Licence de droits d'usage



Contexte public } avec modifications

***Par le téléchargement ou la consultation de ce document, l'utilisateur accepte la licence d'utilisation qui y est attachée, telle que détaillée dans les dispositions suivantes, et s'engage à la respecter intégralement.***

La licence confère à l'utilisateur un droit d'usage sur le document consulté ou téléchargé, totalement ou en partie, dans les conditions définies ci-après et à l'exclusion expresse de toute utilisation commerciale.

Le droit d'usage défini par la licence autorise un usage à destination de tout public qui comprend :

- le droit de reproduire tout ou partie du document sur support informatique ou papier,
- le droit de diffuser tout ou partie du document au public sur support papier ou informatique, y compris par la mise à la disposition du public sur un réseau numérique,
- le droit de modifier la forme ou la présentation du document,
- le droit d'intégrer tout ou partie du document dans un document composite et de le diffuser dans ce nouveau document, à condition que :
  - L'auteur soit informé.

Les mentions relatives à la source du document et/ou à son auteur doivent être conservées dans leur intégralité.

Le droit d'usage défini par la licence est personnel et non exclusif.

Tout autre usage que ceux prévus par la licence est soumis à autorisation préalable et expresse de l'auteur : [sitopedago@telecom-paristech.fr](mailto:sitopedago@telecom-paristech.fr)

26 octobre 2012

