

Infrastructures du Commerce Électronique (NFE102)

1 – Bases du Web

Auteurs: Raphaël Fournier-S'niehotta, Philippe Rigaux et Nicolas Travers
(philippe.rigaux,fournier,traversn@cnam.fr)

Département d'informatique
Conservatoire National des Arts & Métiers, Paris, France

Plan du cours

S1 Bases sur le fonctionnement du Web

S2 URLs et URIs

S3 HTTP et HTTPS

S4 HTML et XHTML

S5 Applications dynamiques

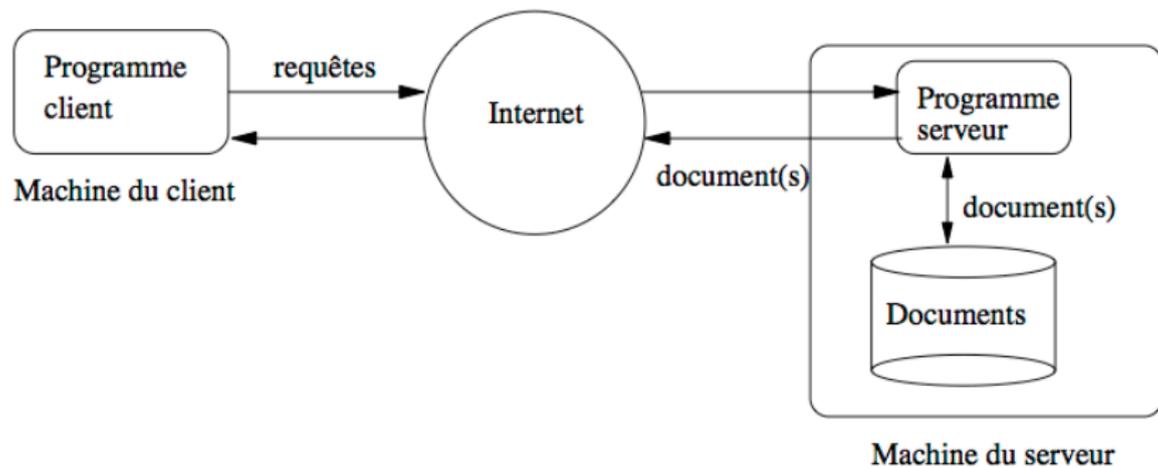
Le World-Wide Web

- Le **World-Wide Web** (ou WWW, ou Web) est un très grand système d'information, réparti sur un ensemble de **machines** connectées par le réseau Internet.
- Ce système est essentiellement constitué de **documents hypertextes** et de documents **multimedia**: textes, images, sons, vidéos, etc.
- Chaque machine propose un ensemble plus ou moins important de documents qui sont transmis sur le réseau par l'intermédiaire d'un **programme serveur**.
⇒ Exemple : Apache, un serveur **open source** très couramment utilisé pour des sites simples (blogs, sites statiques)
- Ce programme serveur dialogue avec un **programme client** qui peut être situé n'importe où sur le réseau
⇒ Le programme client prend le plus souvent la forme d'un **navigateur** (Firefox, Internet Explorer, Safari, Opera)

Fonctionnement du Web

- Le dialogue entre un programme serveur et un programme client s'effectue selon des règles précises qui constituent un **protocole**
⇒ Le protocole du Web est HTTP
- Tout site web est constitué, matériellement, d'une machine connectée à l'Internet équipée du programme serveur tournant en permanence
- Le programme serveur est en attente de requêtes transmises à son attention sur le réseau par un programme client
- Quand une requête est reçue, le programme serveur l'analyse afin de déterminer quel est le document demandé, recherche ce document et le transmet au programme client
⇒ Le programme client peut demander au programme serveur d'exécuter un programme, en fonction de paramètres, et de lui transmettre le résultat

Fonctionnement du Web



Page Web

← → ↻ 🏠 📄 www.bdpedia.fr/nfe102-document-de-travail/

Données, applications, systèmes

Cours, supports, pratique en ligne



Accueil Bases relationnelles Systèmes relationnels Bases documentaires et NoSQL Applications orientées données

NFE102 – Document de travail

Objectifs et contenus

Cette UE vise à donner à ses auditeurs un panorama des composants méthodologiques, techniques et de valorisation liés à la mise en place d'une plate-forme de commerce électronique. Elle est destinée à des informaticiens souhaitant intervenir, dans le cadre d'une maîtrise d'ouvrage ou maîtrise d'œuvre, sur la conception, le développement et la gestion d'un tel système. Elle permet d'acquérir les compétences suivantes:

- **Méthodologiques:** conception, architecture, choix d'un environnement de développement et d'exploitation, frameworks, évaluation d'un site e-commerce.
- **Techniques:** caractéristiques des applications Web et mobiles, sécurité, services et interfaces, passage à l'échelle (NoSQL).
- **Valorisation:** monétisation et paiement électronique, certification électronique, analyse du trafic, référencement, gestion des utilisateurs, signature numérique, profils et recommandations.

Les acquis du cours doivent permettre d'intervenir dans la conception et la mise en place d'une application de commerce électronique et comprenant les différents enjeux et en

Autre ressources

- [A propos](#)
- [Enseignements au Cnam](#)
- [Blog](#)
- [WebScope – SQL en ligne](#)
- [Lien vers le RSS du site](#)

Liste des forums

- [Forum "Bases de données documentaires et distribuées"](#)

Derniers billets

- [Ouverture des inscriptions](#)
- [Apprendre Hibernate et JPA](#)
- [Fin de la session 2013: bonnes vacances à tous](#)
- [Vidéos disponibles pour les trois premiers chapitres](#)
- [Début de la formation en ligne](#)

Connexion

Username:

le cnam

Plan du cours

S1 Bases sur le fonctionnement du Web

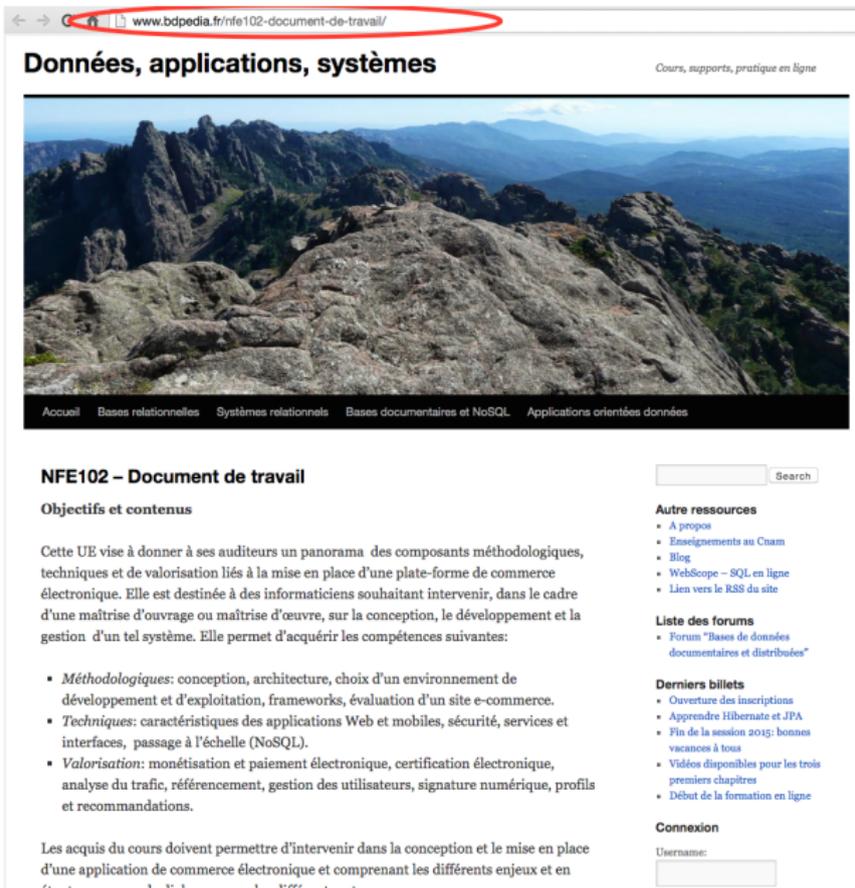
S2 URLs et URIs

S3 HTTP et HTTPS

S4 HTML et XHTML

S5 Applications dynamiques

Page Web - URL



← → www.bdpedia.fr/nfe102-document-de-travail/

Données, applications, systèmes

Cours, supports, pratique en ligne



Accueil Bases relationnelles Systèmes relationnels Bases documentaires et NoSQL Applications orientées données

NFE102 – Document de travail

Objectifs et contenus

Cette UE vise à donner à ses auditeurs un panorama des composants méthodologiques, techniques et de valorisation liés à la mise en place d'une plate-forme de commerce électronique. Elle est destinée à des informaticiens souhaitant intervenir, dans le cadre d'une maîtrise d'ouvrage ou maîtrise d'œuvre, sur la conception, le développement et la gestion d'un tel système. Elle permet d'acquérir les compétences suivantes:

- **Méthodologiques:** conception, architecture, choix d'un environnement de développement et d'exploitation, frameworks, évaluation d'un site e-commerce.
- **Techniques:** caractéristiques des applications Web et mobiles, sécurité, services et interfaces, passage à l'échelle (NoSQL).
- **Valorisation:** monétisation et paiement électronique, certification électronique, analyse du trafic, référencement, gestion des utilisateurs, signature numérique, profils et recommandations.

Les acquis du cours doivent permettre d'intervenir dans la conception et la mise en place d'une application de commerce électronique et comprenant les différents enjeux et en

Autre ressources

- [A propos](#)
- [Enseignements au Cnam](#)
- [Blog](#)
- [WebScope – SQL en ligne](#)
- [Lien vers le RSS du site](#)

Liste des forums

- [Forum "Bases de données documentaires et distribuées"](#)

Derniers billets

- [Ouverture des inscriptions](#)
- [Apprendre Hibernate et JPA](#)
- [Fin de la session 2013: bonnes vacances à tous](#)
- [Vidéos disponibles pour les trois premiers chapitres](#)
- [Début de la formation en ligne](#)

Connexion

Username:

le cnam

Les URLs

- Toutes les **ressources** sur le Web sont identifiées par une **URL** (**Uniform Resource Locator**), de la forme suivante :
http://www.toto.com:443/chemin/vers/doc?code=NFE102&type=latex#fragment
- Le **protocol** d'échange associé, ici HTTP pour les pages Web. Peut être mailto:, ftp:, news:, ...
- l'**hôte** (*hostname*, ou domaine, avec le sous-domaine www)
⇒ Un des services du réseau Internet (le DNS) convertit ce nom d'hôte en adresse IP, permettant la communication entre machines.
- le **port**. Chacun des ports correspondant à l'un des services gérés par le serveur. Pour HTTP, le port par défaut est 80 (8080 pour Tomcat) et peut être omis.

Les URLs

- Toutes les **ressources** sur le Web sont identifiées par une **URL (Uniform Resource Locator)**, de la forme suivante :

http://www.toto.com:443/chemin/vers/doc?code=NFE102&type=latex#fragment

- le **chemin vers la ressource**. Il suit la syntaxe d'un chemin d'accès à un fichier dans un système de fichiers. S'il est omis, l'URL désigne la **racine** du site web.
 - ⇒ Dans les sites simples, "statiques", ce chemin correspond de fait à un emplacement physique vers le fichier contenant la ressource.
 - ⇒ Dans des applications dynamiques, les chemins sont virtuels et conçus pour refléter l'organisation logique des ressources offertes par l'application.
- la liste des **paramètres (query string)**
- le **fragment** (une sous-partie) du document

URI

- Une URI est une identification unique de ressource (**Uniform Resource Identifier**¹)
- La ressource (page Web, image, vidéo, texte, paragraphe, tweet...) est caractérisée par son unicité sur le réseau.
- Elle doit permettre de l'identifier de manière permanente, même si la ressource est déplacée ou supprimée.
- Types d'URI :
 - Physique avec une **URL** :
<http://www.bdpedia.fr/nfe102-document-de-travail/>
 - Abstraite avec une **URN** :
<URN:DOI:10.1145/2790755.2790768>
 - **Uniform Resource Name** est défini à l'aide du schéma : `urn:namespace:namespace string`
 - Le namespace est code de schéma : ISBN, ISAN, DOI, ietf,...
 - Indépendant de la localisation

¹URI : RFC 1630, Juin 1994

Plan du cours

S1 Bases sur le fonctionnement du Web

S2 URLs et URIs

S3 HTTP et HTTPS

S4 HTML et XHTML

S5 Applications dynamiques

Le protocole HTTP

- HTTP, pour **HyperText Transfer Protocol**, est un protocole extrêmement simple, basé sur TCP/IP (ne contient pas les informations de connexion)
- Protocole de transfert de données : texte, hypertexte, images, vidéos, ...
- Echange entre le client (i.e., navigateur) et le serveur (i.e., serveur web)
- Voici par exemple une requête envoyée à un serveur Web :

```
GET /nfe102-document-de-travail/ HTTP/1.1
Host: http://www.bdpedia.fr
```

- Et voici la réponse à cette requête :

```
HTTP/1.1 200 OK
Content-Type: text/html; charset=UTF-8

<!DOCTYPE html>
<html lang="en-US">
  <head>
    <meta charset="UTF-8" />
    <title>NFE102 &#8211; Document de travail | Données, applications, systèmes</title>
    <meta name="generator" content="WordPress 4.0.8" />
  </head>
  <body class="page page-id-252 page-template-default">
    <div id="wrapper" class="hfeed">
      <div id="header">
```

Requête HTTP

GET /nfe102-document-de-travail/ HTTP/1.1

Host: http://www.bdpedia.fr

- Envoyée par le navigateur lorsque l'on demande le contenu d'une URL (par exemple)
- Le **message** doit contenir :
 - Une **méthode** : GET, HEAD, POST, PUT, DELETE
 - GET : Demande la ressource
 - HEAD : Informations sur la ressource
 - POST : Envoi d'information au serveur
 - PUT : Envoi d'une ressource pour remplacement
 - DELETE : Suppression de ressource
 - L'**URI** de référence
 - Le **protocole** et la version utilisée
 - L'**hôte**
 - Autre : User Agent (browser), Accept (Mime types, password), If-Modified-Since (answer with content if necessary)

Réponse HTTP

HTTP/1.1 200 OK

Content-Type: text/html; charset=UTF-8

```
<!DOCTYPE html>
<html lang="en-US">
<head>
  <meta charset="UTF-8" />
  <title>NFE102 &#8211; Document de travail | Données, applications, systèmes</title>
  <meta name="generator" content="WordPress 4.0.8" />
</head>
<body class="page page-id-252 page-template-default">
```

- La réponse HTTP est constituée de deux parties séparées par une ligne blanche :
 - L'entête, avec :
 - Le protocole et la version
 - Le code de confirmation
 - Le type de la ressource
 - Le corps.
- Le programme client affiche le corps du message retourné, en fonction des informations contenues dans l'entête. Si le code HTTP est 200 par exemple, il procède à l'affichage. Un code 404 indique une ressource manquante
- La requête peut contenir la protection par mot de passe, la négociation de contenu (type de document demandé), des cookies, informations déposées côté client pour mémoriser les séquences d'échanges (sessions)

Code de retour HTTP

■ 1xx : Information

■ 2xx : Réponse correcte

- 200 OK
- 201 Commande POST
- 202 Requête acceptée
- 203 Requête GET ou HEAD réalisée avec succès
- 204 Pas de contenu dans la ressource

■ 3xx : Redirection

- 300 La ressource est à plusieurs endroits
- 301 Déplacée de manière permanente
- 302 Déplacée temporairement
- 304 Pas de modification

■ 4xx : Erreur *client*

- 400 Requête invalide
- 401 Requête non autorisée
- 402 Requête demandant un paiement
- 403 Accès interdit à la ressource
- 404 Ressource non trouvée
- 405 Méthode non autorisée
- 406 Type de ressource interdit

■ 5xx : Erreur *serveur*

- 500 Erreur interne
- 501 Méthode non implémentée
- 502 Problème de passerelle ou surcharge serveur
- 503 Service indisponible ou temps de réponse trop long [504] Passerelle secondaire sans réponse

HTTP versions

0.9 (1990) Méthode + URI

1.0 (1995) Types MIME, code retour, cache de données, autorisations

1.1 (1999) Segments de ressources, streaming, qualité de la ressource, contrôle d'intégrité, transactions (multiples requêtes pour une connexion)

HTTPS + Verrou + certificat

The screenshot shows a Google Maps page with a certificate popup. The browser address bar displays the URL `https://www.google.fr/maps/place/292+Rue+Saint-Martin,+75003+Paris/@48.8665992,2.3548777,17z/data=!3m1!4m2!3m1!1s0x47e666666666666666:0x47e666666666666666`. The popup displays the following information:

- GeoTrust Global CA** (Issuer)
- Google Internet Authority G2** (Issued by)
- *.google.com** (Domain)
- *.google.com** (Subject)
- Délicivé par: Google Internet Authority G2**
- Expire le mercredi 23 décembre 2015 01:00:00 heure normale d'Europe centrale**
- Ce certificat est valide** (Valid)

Détails

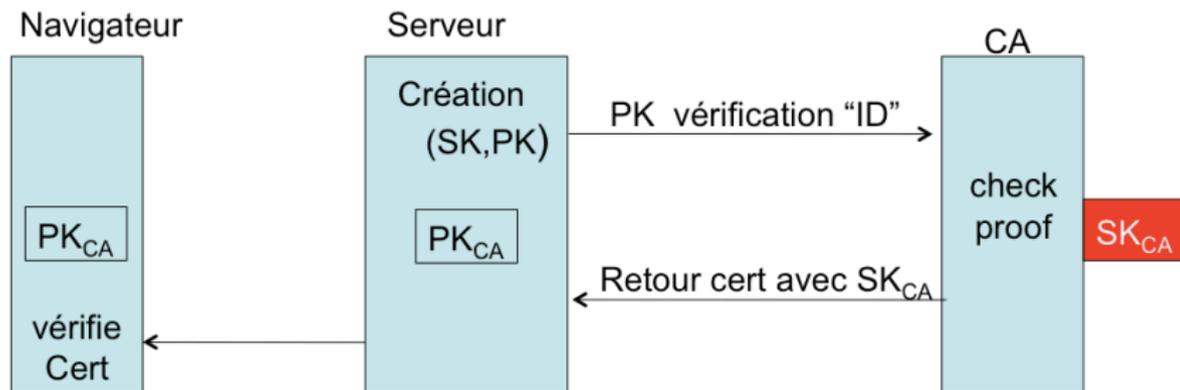
Sujet	
Pays	US
Région/Province	California
Localité	Mountain View
Organisation	Google Inc
Nom	*.google.com
Nom de l'émetteur	
Pays	US
Organisation	Google Inc
Nom	Google Internet Authority G2

OK

- Protocole sécurisé d'échange de données : SSL/TLS

SSL/TLS

- Le serveur crée la clé PK d'**encryptage**
- Échange et vérification de la clé avec le serveur d'authentification (CA - Certificate Authorities)
 - ⇒ Génération de la clé SK de **décryptage**
- Envoi de PK et CA au navigateur (certification)
- Les messages sont cryptés avec PK, et décryptés avec SK

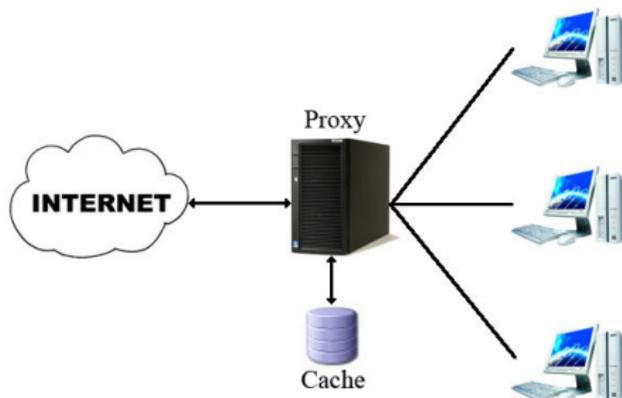


Soucis du HTTPS

- Problème avec les proxy (protocole coupé)
- Problème avec le *virtual hosting* (même IP)
- Pourquoi ne pas utiliser toujours HTTPS ?
 - Ralentissement des serveurs et du trafic
 - Pas de caching possible

Principe sur les Proxys

- **Proxy** : Serveur utilisé comme intermédiaire pour les requêtes clients
- But principaux :
 - **Anonymiser** les machines associées (sécurité)
 - Accélérer l'accès à une ressource via **gestion de cache**
Si le cache existe et récent, il est retourné. Sinon, requête au serveur
 - **Server Web intermédiaire**
Peut réécrire les requêtes, anonymise, gestion multi-thread, équilibrage de charge
 - **Filtrage** de contenu



Plan du cours

S1 Bases sur le fonctionnement du Web

S2 URLs et URIs

S3 HTTP et HTTPS

S4 HTML et XHTML

S5 Applications dynamiques

Page Web et HTML

← → ↻ 🏠 www.bdpedia.fr/nfe102-document-de-travail/

Données, applications, systèmes

Cours, supports, pratique en ligne



images : ``

Accueil Bases relationnelles Systèmes relationnels Bases documentaires et NoSQL Applications orientées données

Entête : `<H1>` Formulaires : `<form> + <input>`

NFE102 – Document de travail

Objectifs et contenus Paragraphe : `<p>`

Cette UE vise à donner à ses auditeurs un panorama des composants méthodologiques, techniques et de valorisation liés à la mise en place d'une plate-forme de commerce électronique. Elle est destinée à des informaticiens souhaitant intervenir, dans le cadre d'une maîtrise d'ouvrage ou maîtrise d'œuvre, sur la conception, le développement et la gestion d'un tel système. Elle permet d'acquérir les compétences suivantes:

- *Méthodologiques*: conception, architecture, choix d'un environnement de développement et d'exploitation, frameworks, évaluation d'un site e-commerce.
- *Techniques*: caractéristiques des applications Web et mobiles, sécurité, services et interfaces, passage à l'échelle (NoSQL).
- *Valorisation*: monétisation et paiement électronique, certification électronique, analyse du trafic, référencement, gestion des utilisateurs, signature numérique, profils et recommandations.

Liste d'items : ` + `

Les acquis du cours doivent permettre d'intervenir dans la conception et la mise en place d'une application de commerce électronique et comprenant les différents enjeux et en

Formulaires : `<form> + <input>`

Search

Autres ressources

- A propos
- Enseignements au Cnam
- Blog
- WebScope – SQL en ligne
- Cliquez vers le RSS du site

Liens/ancre : `<a href>`

Liste des forums

- Forum "Bases de données documentaires et distribuées"

Derniers billets

- Ouverture des inscriptions
- Apprendre Hibernate et JPA
- Fin de la session 2013: bonnes vacances à tous
- Vidéos disponibles pour les trois premiers chapitres
- Début de la formation en ligne

Connexion

Username:

Le langage HTML

- Sur le Web, le type de document le plus présent est : **document hypertexte**
- HTML est le langage permettant de spécifier à la fois **la structure** et **la présentation** des pages Web. Ce n'est pas un langage de programmation
- Le style d'affichage peut être couplé avec une feuille de **style CSS**
- Il existe de nombreux éditeurs de documents HTML qui facilitent la saisie des balises et fournissent une aide au pré-positionnement des différentes parties du document

HTML & Balises

- HTML est un langage structuré à l'aide de **balises**,
- Une balise est définie par :
“<” + nom + (attribut=“...”)* + “>” ...“</” + nom + “>”
- **XHTML** est une forme normalisée et standardisée :
 - Toute balise ouverte est fermée
 - Pas de croisement de balises
 - Le nom des balises et les attributs font parti du schéma XHTML
- Les noms de balise standards sont :
 - <html>,<body>,<head> : Entêtes de structures
 - <head> : contient des méta-informations non affichées (encodage, fichiers liés - CSS, scripts...)
 - <body> : structure de la page, ce qui est affiché
 - <div> : bloc d'affichage
 - : lien hypertexte (contient généralement une URL)
 - / + : liste d'items
 - : lien vers une image
 - <h1>/<h2>/... : titres de différentes importances
 - ,<i>,, : formatage de texte
 - ...
- Les balises peuvent avoir des attributs, permet de spécifier des méta-informations ou un formatage :
 - Contient une valeur, le nom de l'attribut est unique pour la balise
 - href, title, class, id, font...

Exemple de page HTML

```
<html lang="en-US">
  <head>
    <meta charset="UTF-8" />
    <title>NFE102 &#8211; Document de travail | Données, applications, systèmes</title>
    <link rel="profile" href="http://gmpg.org/xfn/11" />
    <link rel="stylesheet" type="text/css" href="http://www.bdpedia.fr/wp-content/themes/twentyten/style.css" />
  </head>
  <body class="page page-id-252 page-template-default">
    <div id="wrapper" class="hfeed">
      <div id="header">
        <div id="masthead">
          <div id="branding" role="banner">
            <div id="site-title">
              <span><a href="http://www.bdpedia.fr/" title="Données, applications, systèmes" rel="home">
                Données, applications, systèmes
              </a></div>
            <div id="site-description">Cours, supports, pratique en ligne</div>
            
          </div>
          <div id="access" role="navigation">
            <div class="skip-link screen-reader-text"><a href="#content">Skip to content</a></div>
            <div class="menu-header">
              <ul id="menu-navigation" class="menu">
                <li id="menu-item-12"><a href="http://www.bdpedia.fr/">Accueil</a></li>
                <li id="menu-item-15"><a href="http://www.bdpedia.fr/bd-relationnelles/">Bases relationnelles</a></li>
                <li id="menu-item-28"><a href="http://www.bdpedia.fr/systemes-relationnels/">Systèmes relationnels</a></li>
                <li id="menu-item-69"><a href="http://www.bdpedia.fr/bases-documents/">Bases de données</a></li>
                <li id="menu-item-31"><a href="http://www.bdpedia.fr/applications/">Applications orientées</a></li>
              </ul>
            </div>
          </div>
        </div>
      </div>
      <div id="main"><div id="container"><div id="content"><div id="post-252">
        <h1 class="entry-title">NFE102 &#8211; Document de travail</h1>
        <div class="entry-content">
          <p><strong>Objectifs et contenus</strong><br />
          </p>
        </div>
      </div>
    </div>
  </div>
</body>
</html>
```

CSS - Feuille de style

- La feuille de style permet de préciser des informations de formatage et d'affichage pour les balises
- Évite de recopier une information pour toutes les balises de même type, sur l'ensemble du site Web

```
html, body, div, span, applet, object, iframe,
h1, h2, h3, h4, h5, h6, p, blockquote, pre,
a, abbr, acronym, address, big, cite, code,
del, dfn, em, font, img, ins, kbd, q, s, samp,
small, strike, strong, sub, sup, tt, var,
b, u, i, center,
dl, dt, dd, ol, ul, li,
fieldset, form, label, legend,
table, caption, tbody, tfoot, thead, tr, th, td {
    background: transparent;
    border: 0;
    margin: 0;
    padding: 0;
    vertical-align: baseline;
}
body {
    line-height: 1;
}
h1, h2, h3, h4, h5, h6 {
    clear: both;
    font-weight: normal;
}
ol, ul {
    list-style: none;
}
blockquote {
    quotes: none;
}
blockquote:before, blockquote:after {
    content: '';
    content: none;
}
del {
    text-decoration: line-through;
```

Plan du cours

S1 Bases sur le fonctionnement du Web

S2 URLs et URIs

S3 HTTP et HTTPS

S4 HTML et XHTML

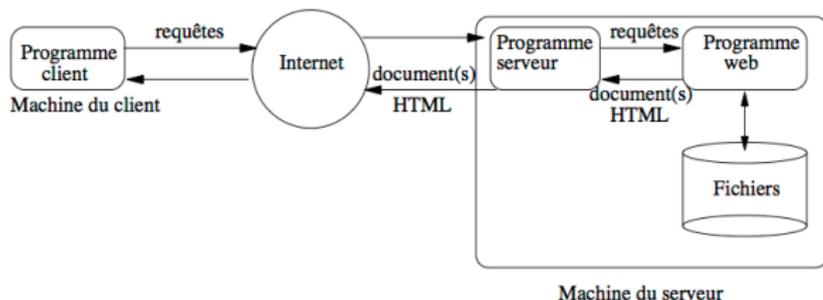
S5 Applications dynamiques

Sites Web dynamiques

- HTML est un langage permet de définir le contenu et l'affichage de manière **statique**
- Il est possible de créer des pages dynamiques :
 - Le contenu de la page est généré **dynamiquement** par le serveur
 - Le contenu de la page est modifié de manière **interactive**

Sites Web dynamiques

- HTML est un langage permet de définir le contenu et l'affichage de manière **statique**
- Il est possible de créer des pages dynamiques :
 - Le contenu de la page est généré **dynamiquement** par le serveur :
 - **PHP** : Hyperlink Preprocessor (OpenSource)
 - **ASP** : Active Server Pages (Microsoft)
 - **JSP** : Java Servers Pages (Java - SUN)
 - **Perl**
 - ...



- Le contenu de la page est modifié de manière **interactive**

Sites Web dynamiques

- HTML est un langage permet de définir le contenu et l'affichage de manière **statique**
- Il est possible de créer des pages dynamiques :
 - Le contenu de la page est généré **dynamiquement** par le serveur
 - Le contenu de la page est modifié de manière **interactive** :
 - **Javascript** : Scripts intégrés aux pages Web
 - Exécuté par le navigateur, il permet de modifier le code HTML&CSS de la page en cours de manière interactive
 - Il peut communiquer avec le serveur (i.e., **Ajax** via HTTP) pour récupérer des données (format **JSON**)

