

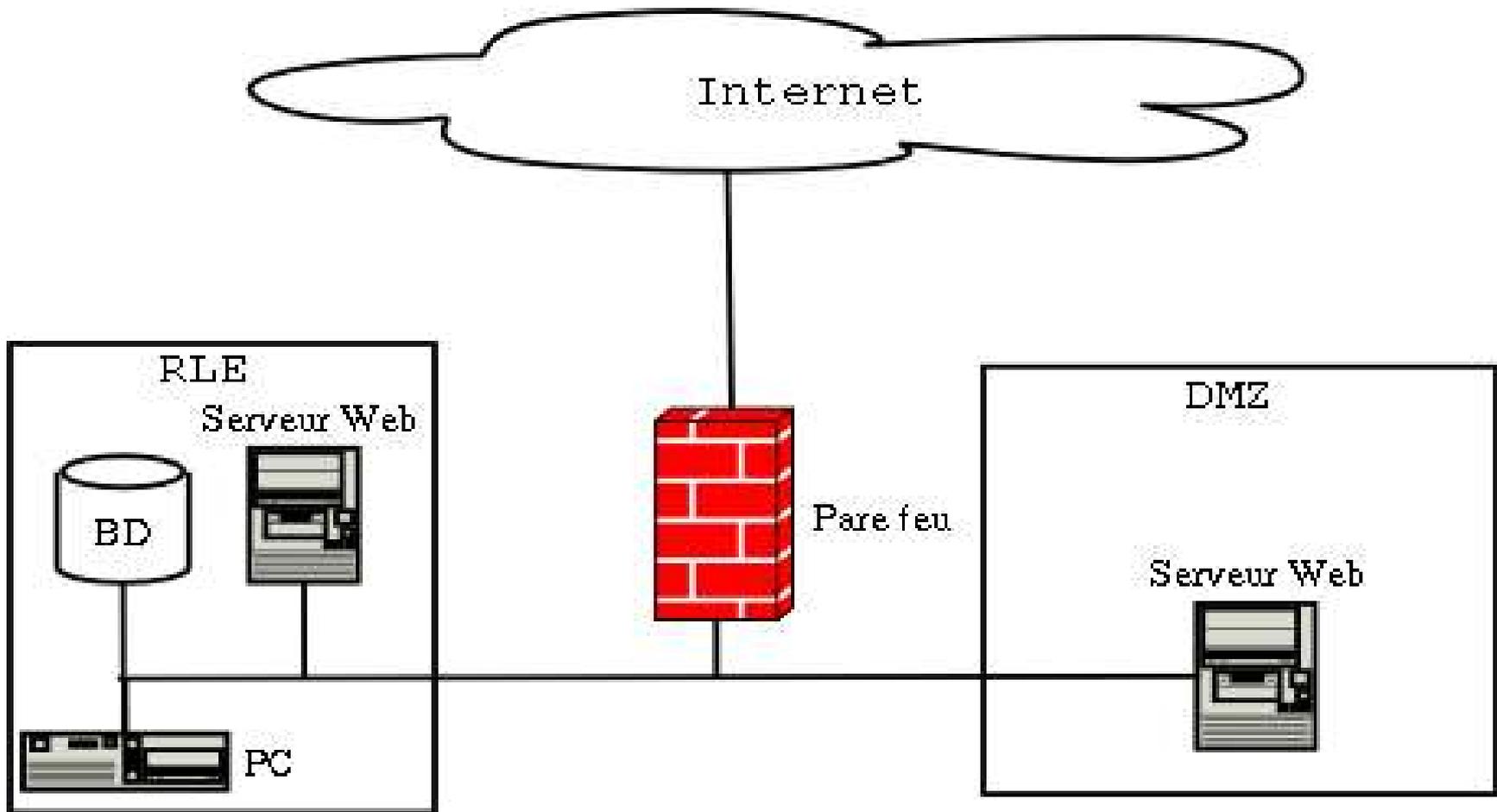
Diffusion du droit

Module 6: Architecture de
systèmes de diffusion libre du droit

Ernst Perpignand

ernst@lexum.umontreal.ca

Matériel informatique



Logiciels

- Système d'exploitation
 - Linux (Debian, Red Hat)
- Pare-feu:
 - Inclus dans linux
 - Difficile à configurer
 - Logiciel commercial
 - Astaro
 - Solution réseau
 - Routeur – pare feu

Logiciels

- Serveur Web (DMZ)
 - Serveur HTTP
 - Apache (www.apache.org)
 - Awstats (www.awstats.org)
 - Serveur JSP
 - Jakarta Tomcat (<http://jakarta.apache.org>)
 - Moteur de recherche
 - Lucene (<http://jakarta.apache.org>)
 - LiiSA (LexUM)
 - Sino (AustLII, <http://www.austlii.edu.au>)

Logiciels

- Serveur Web (RLE)
 - Serveur HTTP
 - Apache
 - Awstats (logiciel de statistiques d'accès au site)
 - Serveur JSP
 - Tomcat
 - Logiciels développés à l'interne basés sur logiciels libres
 - Système de gestion du contenu
 - Système de publication

Systeme de gestion du contenu

- XML + XHTML
 - Les méta informations sont organisées dans des fichiers XML qui eux même sont organisés selon l'organisation hiérarchique de la ressource
 - Mise à jour manuelle
 - Utilitaires de mise à jour
 - Les documents originaux sont archivés, convertis au format XHTML et organisés dans la hiérarchie

Systeme de gestion du contenu

- Base de données relationnelle
 - Les méta informations sont organisées dans les tables d'un SGBD
 - Schéma de la base de données
 - Utilitaire de mise à jour de la base de données
 - Les documents originaux sont archivés, convertis au HTML dépouillé et organisés dans la hiérarchie de la ressource

Systeme de publication

- XML + XSLT
 - Des feuilles de styles XSLT sont appliquées aux fichiers XML afin de générer les pages HTML du site
 - Génération statique
 - Librairie Xalan incluse dans la distribution du langage java
 - Génération dynamique
 - Interprète XSLT dans les navigateurs (IE, Mozilla)
 - XSLT est un langage de transformation pour le XML

Systeme de publication

- Base de données relationnelle
 - Les méta informations sont chargées à partir de la base de données et utilisées dans la création des pages HTML
 - Génération statique
 - Des utilitaires java utilisent la librairie JDBC pour accéder à la BD
 - Génération dynamique
 - Utiliser des pages JSP

Moteur de recherche

- Indexation automatique
 - Les mots du document sont extraits, normalisés et organisés dans un index conçu pour un accès rapide
 - Un poids, représentant l'importance d'un mot pour un document donné, est associé à chaque mot présent dans le document en question

Moteur de recherche

- Recherche
 - L'utilisateur formule une requête en spécifiant les mots (concepts) qui l'intéressent
 - L'engin de recherche interroge l'index afin de retrouver tous les documents contenant les mots contenus dans la requête

Moteur de recherche

- Type de requêtes
 - Simple
 - Fiscal
 - Transport automobile
 - Booléenne
 - Alcool **ET** conduite
 - Vol **OU** délit
 - Phrase
 - « droit administratif »

Moteur de recherche

- Présentation des résultats
 - Modèle booléen
 - Retient seulement si le mot est présent ou non dans un document
 - Résultats non ordonnés
 - Modèle vectoriel
 - Associe un poids à chaque mot pour le document en question
 - Résultats ordonnés par ordre de pertinence

Moteur de recherche

- Lucene (<http://jakarta.apache.org>)
 - Logiciel libre jouissant d'un bon support de la communauté
 - Java
 - facile à modifier et à intégrer
 - Implémentation flexible
 - Format supportés
 - Texte
 - Word
 - Base de données
 - ...

Moteur de recherche

- SINO
 - Développé par l'équipe AustLII
 - Performant
 - Langages de programmation
 - C
 - Perl
 - Pas adapté pour le français
 - Pas vraiment libre