
Accès à distance aux ressources électroniques avec EZproxy

30 avril 2009

**Bibliothèque
et Archives
nationales**

Québec 

Accès à distance aux ressources électroniques disponibles à BAnQ

Objectif :

- Permettre l'utilisation à distance des ressources électroniques de façon transparente au plus grand nombre d'utilisateurs de BAnQ.
 - Trois contraintes principales :
 - Les utilisateurs doivent établir leur identité sur le portail, mais aucune autre identification n'est nécessaire pour l'accès aux ressources électroniques (SSO – Single Sign-On)
 - L'accès aux ressources électroniques ne doit pas nécessiter l'installation de logiciel(s) ou la modification de la configuration de l'ordinateur de l'utilisateur
 - Le contrôle de l'accès aux ressources électroniques doit permettre d'assurer aux fournisseurs que seuls les utilisateurs de BAnQ peuvent accéder à leur contenu
-

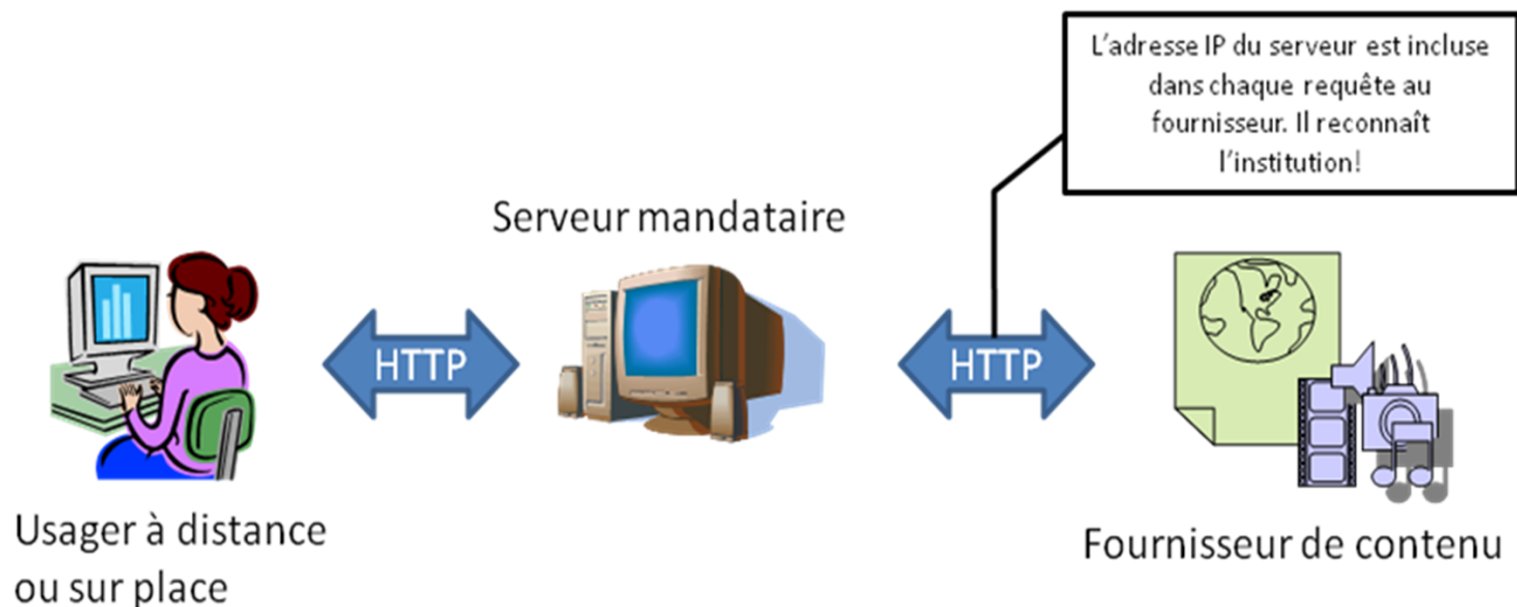
Comment les fournisseurs gèrent les accès de leur côté (utilisation sur place ou à distance)

- Méthode utilisant un code d'utilisateur : devient rapidement ingérable avec plusieurs fournisseurs et des milliers d'utilisateurs. Ne répond pas à la contrainte d'identification unique.
- Méthode utilisant le référent HTTP : ne fonctionne pas avec les signets/favoris. Première solution utilisée à BAnQ. C'est souvent bloqué par les logiciels (de plus en plus populaires) d'antivirus ou de contrôle de la confidentialité installés sur les ordinateurs des utilisateurs.
- Méthode utilisant l'identification de l'adresse IP : c'est idéal pour l'accès sur place; on communique aux fournisseurs la ou les adresses IP de sortie* du réseau des postes de la bibliothèque. Comment faire pour utiliser cette méthode pour l'accès à distance? Les fournisseurs ne peuvent pas identifier les adresses IP de tous nos utilisateurs!

* L'adresse IP de sortie doit être permanente (elle l'est habituellement pour les organismes). Pour connaître l'adresse de sortie d'un poste informatique : <http://www.whatismyip.org/>

Identification par adresse IP pour l'accès à distance

- C'est possible grâce à un type de logiciels que l'on appelle serveur mandataire (proxy server). Un serveur mandataire réside sur un ordinateur faisant partie du réseau de la bibliothèque et c'est son adresse IP qui sert aux fournisseurs à nous authentifier. L'accès à distance et même sur place peut se faire via le serveur mandataire.



Identification par adresse IP pour l'accès à distance (suite)

- Un seul problème : pour que le serveur mandataire puisse relayer les requêtes de l'utilisateur, ce dernier doit modifier la configuration de son navigateur Internet. Ceci ne respecte donc pas nos contraintes.
-

EZproxy – un serveur mandataire « transparent »

- Il existe un type de serveurs mandataires dit « transparent » qui ne demande pas de modification à la configuration du poste de l'utilisateur. Ce n'est pas la configuration qui change, mais la façon de publier les liens.
 - Il existe plusieurs logiciels de ce type, mais nos recherches d'une solution nous ont permis de découvrir EZproxy. Ce qui le distingue des autres solutions c'est qu'il a été conçu expressément pour l'utilisation en bibliothèques afin de gérer l'accès aux ressources électroniques des fournisseurs de contenus.
-

EZproxy – un serveur mandataire « transparent » (Suite)

- Par exemple sur le site Web de la bibliothèque, au lieu de publier le lien : <http://www.maressource.com/>
 - On publiera le lien : <http://ezproxy.mabiblio.qc.ca/login?url=http://www.maressource.com/>
-

EZproxy en bref

- Publié en 1999 par Useful Utilities (Chris Zagar) et acquis en 2008 par OCLC. Utilisé dans plus de 2500 institutions à travers le monde.
 - Fonctionne en interceptant les pages des fournisseurs de contenus pour réécrire les liens (URL) afin de les faire pointer vers le serveur EZproxy avant de transmettre le contenu à l'utilisateur.
 - Fonctionne avec la très grande majorité des fournisseurs de contenus.
 - Se connecte à une grande variété de services d'identification (LDAP, SIP, etc.).
 - Génère des fichiers de journalisation sous le format standard des serveurs Web (IIS, Apache) et permet d'utiliser les outils d'analyse standards sur ces fichiers journaux.
 - Il est facile à maintenir et très stable. Les seuls arrêts depuis 4 ans étaient dus à des changements de configuration seulement; aucune panne.
 - Une version d'évaluation illimitée est disponible.
-

Les requis

- Supporte les systèmes d'exploitation Windows (configurable comme un service) : Linux et Solaris.
 - Nécessite un ordinateur avec une configuration très minimale. Pour Windows et Linux : un Pentium II 400 et 256 Mb de mémoire vive.
 - Doit être relié au réseau Internet (!!!) et avoir un nom de serveur déclaré publiquement (ex. : ezproxy.mabiblio.qc.ca).
-

Au sujet de la configuration réseau...

Deux facteurs de succès :

- Le serveur doit avoir une adresse publique statique; on ne peut pas utiliser un système de noms de domaine dynamique. Les fournisseurs de contenus utiliseront cette adresse pour identifier la bibliothèque.
- EZproxy a deux modes de fonctionnement : celui qui est grandement recommandé nécessite l'utilisation d'un nom de serveur avec masquage (« wildcard DNS »).

Par exemple, avec le DNS BIND, l'administrateur réseau pourrait ajouter des lignes telles que :

ezproxy.mabiblio.qc.ca. IN A (adresse IP du serveur)

*.ezproxy.mabiblio.qc.ca. IN A (adresse IP du serveur)

Installation et configuration d'EZproxy

- L'installation est très simple. En suivant la documentation sur le site d'OCLC, on peut installer le logiciel en quelques minutes. Par exemple, sous Windows, il suffit de suivre les instructions sur la page :
<http://www.oclc.org/support/documentation/ezproxy/download/install-win32.htm>
- La configuration du logiciel est plus complexe, mais une fois complétée, on ne devrait plus avoir à intervenir sur le logiciel sauf pour ajouter de nouvelles ressources.

Deux fichiers permettent de gérer la configuration du logiciel :

- config.txt : permet de définir les paramètres globaux du logiciel et les définitions des ressources permises par le serveur mandataire.
 - user.txt : permet de définir la méthode d'identification des usagers. Dans sa forme la plus simple, il contiendra une liste de codes d'utilisateurs et de mots de passe.
-

Fichier config.txt

- Option ProxyByPort : Par défaut, le logiciel est configuré pour utiliser le mode « ProxyByPort ». Sous ce mode, la réécriture se fait en spécifiant un port de communication dans l'URL (<http://ezproxy.mabiblio.qc.ca:2051/index.html>). Ce mode de fonctionnement n'est pas recommandé, car il peut poser problème si un coupe-feu est utilisé.
 - Option ProxyByHostname : On utilisera plutôt, le mode « ProxyByHostname » qui ne nécessite pas de configuration spéciale dans les coupe-feux. Par contre, comme mentionné plus tôt, le serveur doit être joignable par un nom d'hôte avec masquage. On obtiendra alors des URL du genre : <http://www.maressource.com.ezproxy.mabiblio.qc.ca/index.html>.
 - Important : il faut redémarrer le logiciel après avoir effectué des changements au fichier config.txt.
-

Recommandation pour l'utilisation du mode « ProxyByHostname »

- À ajouter au début du fichier config.txt :
#Nom du serveur Ezproxy
Éditer pour mettre le nom du serveur
Name ezproxy.mabiblio.qc.ca
Spécifie que les entrées/sorties sont redirigées sur le port 80
Peut entrer en conflit avec un serveur Web (IIS, Apache)
Peut cohabiter, voir la documentation
Loginport 80
L'option suivante est utilisée en conjonction avec le "wildcard"
DNS
qui a été défini. Cette option permet de réécrire les adresses
sous le format:
www.maressource.com.ezproxy.qc.ca une fois qu'Ezproxy a
effectué la connexion
Option ProxyByHostname

Définition des ressources (database)

- Entrées ajoutées au fichier config.txt pour définir quels sites seront permis et réécrits par EZproxy.
 - Des définitions existent pour une grande quantité de ressources de fournisseurs américains et sont partagées par OCLC et les utilisateurs d'EZproxy. Il existe plusieurs ressources francophones dont BAnQ a déjà effectué et testé la définition.
-

Définition des ressources (database) (suite)

- Définition de base d'une ressource :

Title: Ma ressource

URL: <http://www.maressource.com/index.htm>

Domain: maressource.com

- Il existe d'autres clauses et les définitions de certaines ressources peuvent être assez complexes. Par contre, le partage de l'information fonctionne très bien dans ce domaine : tout le monde utilise à peu près les mêmes ressources.
-

Fichier user.txt

- EZproxy peut utiliser une grande variété de méthodes d'identification des usagers. Certaines sont simples et d'autres plus complexes. Il peut entre autres se connecter sur des systèmes utilisant les protocoles LDAP (incluant Active Directory), SIP/SIP2 (protocole 3M, disponible au catalogue), ODBC (Windows seulement) et bien d'autres.
 - Au plus simple, le contenu du fichier user.txt pourra être constitué d'une liste d'usagers et de mots de passe :
Usager1 : motdepasse1
Usager2 : motdepasse2
UsagerX : motdepasseX
 - Par exemple, il pourrait être produit quotidiennement à partir d'informations extraites d'une base de données (SQL).
-

Autres détails

- Les pages présentées par EZproxy (identification, erreur, menu, etc.) sont en anglais. Par contre, il est facile de les éditer : ce sont de simples fichiers HTML contenant certains codes de remplacement utilisés par EZproxy.
 - EZproxy peut aussi gérer l'accès à partir de postes à l'intérieur de la bibliothèque (idéalement c'est ce qu'on fera pour ne publier qu'un lien d'accès). Dans ce cas, on peut vouloir que les utilisateurs de ces postes soient automatiquement admis ou que la réécriture soit désactivée : les clauses `AutoLoginIP` et `ExcludeIP` peuvent alors être utilisées dans le fichier `config.txt`.
 - EZproxy supporte la notion de groupe pour les ressources. On pourra grouper les ressources et attribuer des groupes à certains usagers selon certains critères.
-

Stratégie d'implantation

- En principe, la version d'évaluation pourra être utilisée jusqu'à la fin de l'implantation.
 - Installer et se familiariser avec le logiciel.
 - Définir les besoins en ce qui concerne l'identification et l'accès aux ressources. Par exemple, est-ce que certaines ressources ne sont disponibles que pour les usagers sur place et est-ce qu'on veut gérer ces accès avec EZproxy?
 - Implanter la méthode d'identification.
 - Effectuer la configuration réseau pour permettre l'utilisation du mode « ProxyByHostname ».
 - Définir et tester les entrées des ressources auxquelles on veut donner accès.
-

Conclusion

■ Bénéfices :

- ❑ Simple à installer et à maintenir (logiciel d'une très grande stabilité)
- ❑ Très grande flexibilité en ce qui a trait à la gestion de l'identification
- ❑ Évolue constamment pour ajouter de nouvelles fonctionnalités
- ❑ Excellent soutien technique et bonne collaboration des autres utilisateurs du produit
- ❑ Coût en matériel peu élevé. Le prix de base pour EZproxy est de 495 \$ US (on peut contacter OCLC pour obtenir une tarification pour un achat de groupe)

■ Inconvénients :

- ❑ Ralentissement par rapport à une connexion directe aux fournisseurs de contenu
 - ❑ Augmentation de la consommation de la bande passante
 - ❑ Pour le moment, il n'y a pas de groupe d'utilisateurs francophones
-

Références

- Site d'EZproxy chez OCLC :
<http://www.oclc.org/ezproxy/default.htm>
 - Liste de discussion EZproxy
 - Détails sur l'abonnement :
<http://www.oclc.org/support/documentation/ezproxy/list.htm>
 - Archives : <http://ls.suny.edu/read/?forum=ezproxy>
-

Questions?
