World Library and Information Congress:
70th IFLA General Conference and Council

22-27 August 2004
Buenos Aires, Argentina

Programme: http://www.ifla.org/IV/ifla70/prog04.htm

Code Number: 055-F
Meeting: 155. Information Technology
Simultaneous Interpretation: -

Shibboleth[1], système open source de partage de ressources

Marianne Afifi
Directrice des ressources électroniques et chargée des projets spécifiques
Division des Sevices d'Information, LVL 113C
Université de Californie du Sud
650 W 35th Street
Los Angeles, CA 90089-2571
Tel: 213.740.8817
Fax: 213.740.7713
Email: afifi@usc.edu
Web: http://isd.usc.edu/~afifi

RESUME *:*
Shibboleth est un projet mené par le groupe d'initiative des logiciels médiateurs de l'Internet2, NSF. Son objectif est de « développer une solution ouverte et normalisée, adaptée aux besoins des institutions en matière d'échanges des informations concernant leurs usagers, selon des protocoles sécurisés et dans le respect de la confidentialité des données. » Les bibliothèques et les organismes de recherche constatent des attentes croissantes d'un accès validé et protégé aux ressources en ligne qu'ils produisent, diffusent sous licence, acquièrent ou partagent. Les fournisseurs et les éditeurs sont également intéressés par de tels modèles d'accès, différents des authentifications habituelles, réalisées sur la base des numéros IP, les serveurs proxy et du système identifiant / mot de passe. Les chercheurs et enseignants qui font un usage intensif de l'Internet ont des exigences toujours croissantes en matière d'échanges d'informations, et se montrent soucieux de la gestion des niveaux d'accès et du respect de la confidentialité des données.

Shibboleth a été développé dans le but de permettre le partage des ressources pour tous ces protagonistes d'une manière simple et normalisée. Dans ma présentation, je résumerai les origines de Shibboleth, donnerai un aperçu de ses développements, et expliquerai son fonctionnement.

[1] En français, l'orthographe correcte est Schibboleth, mais le nom de baptême américain a été conservé (n. d. t.)

J'explorerai également les raisons pour lesquelles c'est un outil important pour le partage des ressources dans les bibliothèques et les environnements fédérés.

J'ai le plaisir de m'adresser à vous aujourd'hui pour évoquer un projet intéressant qui touche au thème de l' « authentification et [de l'] autorisation appliquées à l'offre de services en bibliothèque »[i], thème que la section Technologies de l'information de l'IFLA a proposé pour cette conférence. Ce projet s'appelle Shibboleth. En premier lieu, je voudrais proposer quelques définitions pour les termes que je vais utiliser au long de mon intervention. Ensuite, je donnerai un aperçu des raisons qui ont conduit à la création de Shibboleth, comment il fonctionne, et je montrerai la pertinence de son utilisation en bibliothèque, singulièrement dans le cas de portails de recherche intégrés, comme par exemple le Scholars Portal.[2]

Le développement de Shibboleth a été soutenu par de nombreuses institutions. Certains d'entre vous ne les connaissent peut-être pas, c'est pourquoi je vais vous les énumérer. Le consortium Internet2 de la National Science Foundation est le commanditaire du projet. L'étude des logiciels médiateurs[3] est l'un des buts que le consortium Internet2 s'est fixés.

Définitions et concepts

NSF
Internet2
Logiciels médiateurs
Authentification
Autorisation

La National Science Foundation

A été créée par la loi « NSF »de 1950 et la législation afférente (Code des Etats-Unis, titre 42, sections 1861 et suivantes). Ses compétences ont été étendues par la loi Science and Engineering Equal Opportunities (Code des Etats-Unis, titre 42, section 1885) ainsi que par le titre I de la loi « Education for Economic Security » (Code des Etats-Unis, titre 20, sections 3911 et 3922).

Cette loi fixe ainsi les missions de la NFS :
« Promouvoir les avancées scientifiques ; œuvrer pour le progrès en matière de santé, de prospérité et de bien-être des citoyens ; garantir la défense nationale. »[ii]
http://www.nsf.gov/home/about/creation.htm

Internet2

Est un consortium composé de 206 universités qui travaillent en partenariat avec le secteur de l'industrie et l'Etat pour développer et mettre en œuvre des applications et des technologies de pointe en matière de réseaux, et jouent ainsi un rôle moteur dans la création de l'Internet de demain. Internet2 recrée le partenariat entre les acteurs de l'enseignement supérieur et de la recherche, ceux de l'industrie, et l'Etat, partenariat qui a constitué le terreau de l'Internet à ses débuts. Les buts essentiels d'Internet2 sont les suivants :

- Créer une infrastructure réseau de pointe en faveur de la communauté nationale des chercheurs ;
- Mettre en œuvre des applications Internet révolutionnaires ;
- Permettre dans un délai court la mise à disposition de nouveaux services et de nouvelles applications réseau à l'ensemble des internautes.

[2] Cf. http://www.arl.org/access/scholarsportal/ (n. d. t.)
[3] La traduction du terme est empruntée à http://www.alaide.com/dico.php?q=middleware (n. d. t.)

http://www.internet2.edu/about/aboutinternet2.html[iii]

Logiciels médiateurs

Egalement appelés « glues ». Il s'agit de couches logicielles intercalées entre le réseau et les applications. Ces programmes offrent des services tels que l'identification, l'authentification, l'autorisation, des répertoires, et des outils de sécurité. Dans l'Internet actuel, les applications sont le plus souvent obligées d'inclure ces services, ce qui entraîne la mise en concurrence de standards incompatibles entre eux. Les logiciels médiateurs, en promouvant la normalisation et l'interopérabilité, rendront les applications réseau de pointe beaucoup plus faciles à utiliser.
L'Internet2 Middleware Initiative (I2-MI)[4] travaille en ce moment au développement de logiciels médiateurs de base au sein des universités membres d'Internet2.
http://middleware.internet2.edu/[iv]

Il y a deux termes dont la définition me semble essentielle à la bonne compréhension de cette présentation : authentification et autorisation. J'ai retenu les définitions utilisées par Clifford Lynch, membre de la Coalition for Networked Information[5] (CNI) :

- Authentification : « Le fait, pour l'utilisateur d'un réseau, de se voir octroyer le droit d'utiliser une identité –essentiellement un nom. »[v]
- Autorisation : « Le fait de déterminer si l'on permet à une identité donnée (à laquelle sont associés une série d'attributs) de réaliser certaines actions, comme par exemple accéder à une ressource. »[vi]

Il y a souvent confusion entre ces termes, mais il est important de les distinguer même si aujourd'hui, dans les bibliothèques, ils recouvrent souvent de fait la même réalité. Par exemple, quand on accède à une ressource soumise à une validation de l'adresse IP, les utilisateurs sont authentifiés sur la foi de leur adresse IP, ce qui signifie automatiquement qu'ils sont également autorisés. Dans un environnement Shibboleth, les utilisateurs peuvent très bien être authentifiés comme faisant partie d'un groupe donné, mais ne pas être autorisés à utiliser telle ou telle ressource.

L'Oxford English Dictionary donne, dans sa seconde édition de 1989, plusieurs définitions du mot Shibboleth. J'en ai sélectionné deux, qui s'appliquent au projet dont je vous entretiens aujourd'hui.[6]

« 1. Mot hébreu utilisé par Jephté comme signe de ralliement afin de distinguer les membres de la tribu d'Ephraïm (incapables de prononcer le son « sh ») de ceux de sa propre tribu, les Galaadites. (Juges xii. 4-6)
2. (Au fig.) Slogan ou devise adopté par un groupe ou par une secte, qui permet d'en reconnaître les membres ou les sympathisants, ou d'en exclure toute autre personne. »

On le voit, ce nom est adapté à la désignation d'un procédé de gestion des accès.

Le projet logiciel Shibboleth est né d'un besoin exprimé par les membres d'Internet2 et par d'autres organismes dotés de structures et d'un environnement similaires, de trouver un vecteur plus efficace pour communiquer des informations sur eux-mêmes ou sur leurs utilisateurs. Shibboleth a été conçu pour faciliter la communication et l'échange d'informations entre plusieurs protagonistes, comme par exemple des universités, des organismes gouvernementaux et des sociétés commerciales. A cause de la diversité des membres d'Internet2, le cahier des charges de Shibboleth précisait qu'il devrait reposer sur des normes, être un logiciel libre, garantir la confidentialité des données et favoriser un esprit fédératif parmi les institutions participantes. Nous verrons plus loin en quoi les aspects fédératifs sont d'une importance toute particulière en matière de partage de ressources. Les infrastructures qui servent

[4] En français : Groupe de travail « logiciels médiateurs » du consortium Internet2 (n. d. t.)
[5] Coalition pour la mise en réseau de l'information. (n. d. t.)
[6] Nouveau Petit Robert (éd. 1995) : « Epreuve décisive qui fait juger de la capacité d'une personne » (n. d. t.)

aujourd'hui à la communication entre institutions membres d'Internet2 s'appuient sur des technologies qui ont une fâcheuse tendance à être complexes et lourdes à administrer. Shibboleth permet un échange d'informations affranchi des contraintes liées à des méthodes d'authentification et d'autorisation multiples, non normalisées et délicates à mettre en œuvre. Une fois complètement déployé, Shibboleth permettra une circulation d'information basée sur des normes et garantissant la confidentialité des données individuelles des utilisateurs rattachés aux institutions participantes. Même si le déploiement complet chez les membres et dans les communautés concernées n'est pas encore d'actualité, beaucoup d'institutions ont commencé à tester Shibboleth. De plus, certaines organisations ont également entrepris des démarches ambitieuses de création d'architectures et d'infrastructures pour pouvoir utiliser Shibboleth.

Les universités peuvent installer Shibboleth pour gérer la communication entre chercheurs, dans les bibliothèques, comme portail, et bien sûr comme systèmes de gestion de cours en ligne.[7] Il est hautement souhaitable de rendre tous ces systèmes connectables et interopérables. La pertinence de Shibboleth pour les bibliothèques ne peut manquer de sauter aux yeux de quiconque a travaillé à la gestion de ressources électroniques. L'accès aux sites Web des éditeurs, des fournisseurs et des agrégateurs nécessite le plus souvent que la bibliothèque fournisse une liste des adresses IP ou qu'elle gère en interne une liste de combinaisons nom d'utilisateur / mot de passe. Il est demandé à certaines bibliothèques de mettre en œuvre des serveurs proxy pour identifier les utilisateurs des services. D'autres établissements utilisent un RPV[8] pour permettre à leurs utilisateurs de se connecter au réseau du campus. En outre, quelques fournisseurs imposent des restrictions strictes pour l'accès aux ressources (obligeant par exemple le bibliothécaire à saisir à la place d'un lecteur un nom d'utilisateur et un mot de passe sur un poste précis). Nous sommes au XXIe siècle, et de tels types d'accès surchargent les tâches des bibliothécaires, et pénalisent l'institution et les utilisateurs ; ces méthodes sont de plus en plus dépassées.

La diversité des publics des bibliothèques universitaires s'accroît en même temps que la diversité des populations universitaires. La communauté de nos utilisateurs n'est plus seulement constituée des enseignants et enseignants chercheurs, des étudiants et du personnel universitaire sur le site même de l'université. L'enseignement hors site, le réseau des anciens étudiants, l'implication de plus en plus importante dans la vie de la cité ainsi que les activités universitaires de nature coopérative on généré des catégories d'utilisateurs très variées, à tel point qu'il est aujourd'hui malaisé de leur assurer un accès aux ressources, notamment celles de la bibliothèque.

Shibboleth peut faciliter la gestion des accès aux ressources en ligne, mais cela implique que les établissements assurent la gestion d'une base d'utilisateurs qui soit normalisée. Le plus souvent, la mise en place de telles bases ne repose sur aucune norme, même si cet aspect a fait l'objet de certains efforts. On peut par exemple citer EduPerson, projet défini par ses promoteurs comme « une classe auxiliaire d'objets destinée aux annuaires des universités afin d'améliorer la communication entre les établissements d'enseignement supérieur. Il s'agit d'un ensemble de descriptions de données ou d'attributs applicables aux personnes relevant de l'enseignement supérieur, ainsi que de recommandations syntaxiques ou sémantiques pour la saisie de données dans ces attributs. »[vii] Une fois les attributs applicables aux personnes clairement établis, et grâce à la création et à la gestion régulière d'un annuaire normalisé, il devient beaucoup plus aisé de gérer les accès. Et pour les utilisateurs, la confidentialité des données les concernant est beaucoup mieux protégée dans un environnement Shibboleth que dans les systèmes d'authentification actuels.

Pour la bibliothèque, l'avantage c'est que les ressources peuvent être réservées à des groupes donnés, avec à la clé la possibilité de faire des économies. Par exemple, l'accès à une revue onéreuse peut être accordé uniquement aux personnes participant à un projet de recherche donné plutôt qu'à l'ensemble de la communauté universitaire, ce qui peut permettre de dégager des moyens pour l'acquisition

[7] Le sigle CMS (Content Management System) est ici réutilisé dans le texte anglais pour Course Management System. (n. d. t.)
[8] Réseau Privé Virtuel (Virtual Private Network en anglais) (n. d. t.)

d'autres ressources documentaires. De plus, les aspects fédératifs de Shibboleth peuvent présenter des avantages pour les consortia de bibliothèques, pour les universités comportant plusieurs campus, pour celles gérées au niveau d'un état, pour les projets collaboratifs au sein d'une fédération, mais aussi pour les fournisseurs. Du point de vue de ces derniers, Shibboleth permet d'harmoniser les modes d'accès qu'ils ont actuellement à gérer. La mise en place initiale de Shibboleth ne sera pas instantanée, mais ensuite l'ajout de nouveaux clients Shibboleth sera très simple. Shibboleth permettra également aux fournisseurs d'ajuster et de moduler leur offre plus précisément en fonction de la demande de leurs clients. Cela impliquera probablement aussi la création de nouveaux modèles tarifaires.

L'application Shibboleth modifie considérablement la manière dont on accède aux ressources. Avec les schémas anciens, les utilisateurs se connectent au système du propriétaire des données, par exemple au site Web d'un éditeur, ou à OCLC First Search, et c'est là que se fait leur authentification. Avec Shibboleth, c'est l'institution à laquelle l'utilisateur est rattaché qui vérifie qu'il est présent dans l'annuaire, et transmet au fournisseur de contenu les informations dont il a besoin. Les étapes du processus sont les suivantes :

Imaginons qu'une utilisatrice se connecte au site d'un fournisseur de contenu, disons par exemple JSTOR, qui propose des archives de revues savantes.

1. JSTOR ne sait pas à quelle institution est rattachée l'utilisatrice.
2. Le SHIRE[9] de JSTOR demande au service WAYF[10] à quelle institution est rattachée l'utilisatrice.
3. Le service WAYF demande à l'utilisatrice à quelle institution elle est rattachée.
4. L'utilisatrice redirige la question au serveur de *handles*[11] Shibboleth de son institution de rattachement.
5. Le serveur de *handles* effectue une recherche dans l'annuaire de l'institution de rattachement afin d'y vérifier la présence de l'utilisatrice.
6. Une fois la vérification faite, le serveur de *handles* renvoie une requête à l'utilisatrice pour lui demander de s'identifier.
7. L'utilisatrice s'identifie auprès du serveur de *handles*.
8. Le serveur renvoie un *handle* (également appelé alias) au SHIRE de JSTOR, qui indique le rattachement de l'utilisatrice à l'institution.
9. Le SHIRE fait suivre le *handle* au SHAR[12], qui envoie à l'AA (Autorité Attributs, localisée au sein de l'institution de rattachement) une requête pour obtenir des attributs concernant l'utilisatrice.
10. L'Autorité Attributs fournit les attributs de l'utilisatrice au SHAR.
11. Le SHAR fournit les attributs de l'utilisatrice au système de gestion des ressources de JSTOR, qui autorise l'utilisatrice à accéder aux ressources.

[Insérer le Schéma ici]

Fig.1 : Schéma fourni par le Dr Ken Klingenstein d'Internet2, extrait d'une présentation faite à une réunion de la CNI (cf. note 5) en avril 2004, à partir d'une démo SWITCH (Swiss Education and Research Network : Réseau Suisse pour l'Education et le Recherche).

[9] SHIRE : Shibboleth Indexical Reference Establisher
[10] "Where Are You From" : Où êtes-vous ? (n. d. t.)
[11] "Un objet (souvent réduit à un simple entier) contrôlant l'accès à d'autres objets ou structures de données. Souvent, un handle contrôle également l'acquisition et la libération des ressources (mémoire, accès réseau, etc.). Une utilisation courante des handles et le contrôle d'accès à des structures de données de longueur variable" (http://www.alaide.com/dico.php?q=handle (n. d. t.)
[12] SHAR : Shibboleth Attribute Requester

Même si la description de ce qui se passe dans les coulisses peut sembler complexe, les différents services interagissent dans un environnement web, ce qui permet un accès rapide et transparent aux ressources demandées.

Applications pour les bibliothèques et les portails

Gestion des ressources
Gestion des attributs
- La Segmentation des utilisateurs
- Des portails condidérés comme cibles
- Des portails considérés comme utilisateurs d'autres services
- Le Scholars Portal

Deux concepts cruciaux apparaissent dans ce schéma : l'origine et la cible.
Ces concepts sont importants en ceci que, comme nous allons le voir, les portails peuvent jouer alternativement l'un et l'autre rôle. L'origine, c'est l'endroit où se trouvent le serveur de handles et l'Autorité Attributs. La cible, c'est la ressource en ligne. Le serveur de handles et l'Autorité Attributs utilisent des données stockées dans l'annuaire de l'institution de rattachement de l'utilisatrice, qui gère et contrôle ces données. Ces données sont strictement limitées à ce que la ressource en ligne a besoin de savoir ; aucune information supplémentaire n'est communiquée, contrairement à ce qui se passe avec d'autres méthodes d'authentification. Le schéma préserve la confidentialité des données concernant l'utilisatrice, car le gestionnaire des ressources en ligne n'a besoin de connaître que certains des attributs attachés à l'utilisatrice, alors que les méthodes d'authentification actuelles fournissent plus d'informations concernant les utilisateurs que le strict nécessaire.

Shibboleth et les portails

Dans un portail, les informations et les données proviennent de systèmes et de sources variés, pour être combinés dans un environnement web, de telle sorte que les informations pertinentes pour un groupe ou pour un utilisateur donnés sont présentés soit automatiquement soit sur demande. De plus, les utilisateurs se voient offrir, via le portail, des services comme une messagerie électronique, un service de *chat*[13], des services d'information. Les bibliothèques mettent de plus en plus souvent en œuvre de tels portails, afin d'offrir un guichet unique d'accès à leurs ressources et à leurs services (recherches par mots-clés, méta-recherches, services de renseignements en ligne).

Don Gourley, du Washington Research Library Consortium[viii], a publié récemment un article sur Shibboleth et les portails des bibliothèques, dans lequel il propose quelques scénarios pour leur utilisation. Dans cet article, il suggère d'utiliser le portail de la bibliothèque comme cible Shibboleth.

Le Scholars Portal est un exemple de portail utilisé comme cible. Il s'agit d'un projet ARL qui concerne sept universités américaines qui ont mis au point un portail de méta-recherches en collaboration avec un fournisseur, Fretwell Downing[ix]. Aucune des sept universités n'utilisait à l'origine la même méthode pour gérer les authentifications et les autorisations. Fretwell Downing a travaillé avec les différentes institutions afin d'adapter leurs besoins en matière d'authentification des utilisateurs du portail. L'Université de Californie du Sud (USC), qui est membre d'Internet2 et a déployé Shibboleth, s'est intéressée à l'utilisation de ce dernier dans le but d'interroger le Scholars Portal.

La mise en œuvre du service est en cours : l'USC souhaite que le Scholars Portal (un portail ZPortal développé par Fretwell Downing) puisse se comporter en cible, ce qui implique que le ZPortal doit

[13] Espace de discussion sur Internet, dans lequel les participants conversent au moyen d'outils de messagerie instantanée, les réponses étant généralement publiques. A la différence du forum de discussion, les échanges sont synchrones et, de ce fait, le domaine d'intervention d'un éventuel modérateur est très réduit. (http://www.alaide.com/dico.php?q=Chat&ix=298) (n. d. t.)

héberger le SHIRE, le SHAR et le système de gestion des ressources, et qu'il puisse contacter les serveurs de l'USC. Les utilisateurs qui voudront bénéficier des fonctionnalités de personnalisation proposées par le Scholars Portal passeront par Shibboleth. Cependant, l'accès aux ressources via le Scholars Portal sera toujours géré de la même manière (principalement par adresses IP avec un accès à distance –hors campus– au travers d'un RPV[14]). Ce n'est que lorsque ZPortal lui-même prendra le rôle d'origine, et que la plupart des fournisseurs des ressources proposées à l'USC seront en mesure de déployer des services cibles, que nous pourrons accéder à ces ressources selon la méthode proposée par Shibboleth.

Shibboleth et les structures fédératives

Le projet Shibboleth a été lancé en 2000, et la version 1.1 est sortie en 2003. Elle est en cours de déploiement chez un nombre sans cesse croissant d'universités et de fournisseurs. Le travail sur la nouvelle version, dont la sortie est prévue cette année, se poursuit. Parmi les nouveautés et les développements en projet, on peut citer la recherche de compatibilité avec les portails. Un groupe de travail GUI Shibboleth contribue à l'élaboration d'un éditeur permettant de mettre en œuvre une politique de diffusion d'attributs[15] dans Shibboleth. Cet éditeur permettra aux bibliothèques, au sein d'une fédération, de gérer les attributs concernant les ressources et les services qu'ils gèrent actuellement.

Pour tirer le meilleur parti de Shibboleth, des environnements fédératifs doivent être créés entre les acteurs participant au projet, afin de mettre en place une infrastructure qui améliore la gestion et l'accès des ressources en ligne. C'est dans cette optique qu'a été formé InCommon au sein d'Internet2. Pour citer la présentation qui en est faite sur le site http://www.incommonfederation.org/index.cfm, « InCommon est une fédération d'organisations dont l'objectif est de créer une structure commune venant en renfort de la recherche et de l'enseignement. Le but premier de la fédération est de faciliter la collaboration grâce au partage de ressources dont l'accès est protégé, au moyen d'une structure commune. »[x] Nonobstant le peu de fédérations qui existent à ce jour, et le fait qu'elles n'en soient qu'à leurs balbutiements, nous espérons qu'il s'en créera de plus en plus, car ce mode d'association promet de faciliter les accès aux ressources tout en assurant la confidentialité des informations personnelles concernant les utilisateurs.

Shibboleth modifiera les modes d'interactions qui existent entre les différentes communautés qui utilisent le réseau, ainsi que la manière dont on accèdera à de nouveaux types de ressources. On peut aisément imaginer que de nouveaux schémas de communication verront le jour dans cet environnement privé et sécurisé.

---

[i] Voir la Foire aux Questions : http://shibboleth.internet2.edu/shib-faq.html#001
[ii] La Création de la NSF (fondation natinale pour la Science) et ses missions : http://www.nsf.gov/home/about/creation.htm
[iii] Pour plus d'information sur Internet2 ® : http://www.internet2.edu/about/aboutinternet2.html
[iv] Voir sur le site Internet de l'Internet2 : http://middleware.internet2.edu
[v] Coalition pour l'accès de l'information en réseau, voir http://www.cni.org/projects/authentication-wp.html
[vi] Ibid.
[vii] Spécification d'EduPerson (200312) : voir http://www.nmi-edit.org/eduPerson/internet2-mace-dir-eduperson-200312.html
[viii] Don Gourley, *Le rôle des portails documentaires dans une fédération Shibboleth*, Consortium de la bibliothèque de rechrche de Washington (WRLC), 30 octobre 2003 : http://shibboleth.internet2.edu/docs/gourley-shibboleth-library-portals-200310.html
[ix] Sur le projet Scholars Portal : voir http://www.arl.org/access/scholarsportal/
[x] Page d'accueil de InCommon : http://www.incommonfederation.org/index.cfm

---

[14] Cf. note 8
[15] ARP (Attribute Release Policy)