Cahier des charges pour la mise en compatibilité Shibboleth de nuxeo

Table des matières

1 Introduction

Le but de ce cahier des charges est de spécifier ce que l'on entend par la mise en compatibilité Shibboleth de nuxeo afin d'estimer la charge de travail (et donc le coût) d'une telle opération.

Il est demandé à la société nuxeo de chiffrer le coût d'une prestation permettant de répondre à ce cahier des charges.

Note : Les mentions « nous demandons » et « nous souhaitons » déterminent les éléments obligatoires et optionnels du livrable.

1.1 Shibboleth

Shibboleth est le moyen technique utilisé par les universités françaises (mais aussi beaucoup d'universités à travers le monde) pour répondre à la problématique de fédération d'identités.

Actuellement c'est le GIP RENATER (http://www.renater.fr/) qui opère la fédération d'identités pour l'ensemble de la communauté enseignement/recherche française.

RENATER a un site web dédié à la fédération (https://federation.renater.fr)

On y trouve notamment toute une partie introductive (https://federation.renater.fr/introduction/a-quoi-ca-sert) en plus de toutes les informations techniques et contractuelles liées à la fédération.

Note : On peut aussi se référer à une tutoriel JRES 2005 de référence sur le sujet (http://2005.jres.org/tutoriels.html).

Dans ce document, nous manipulerons notamment les notions de fournisseur d'identités (IdP pour Identity Provider ; considéré comme préexistants dans les établissements) et de fournisseur de services (SP pour Service Provider). Dans le cadre de la mise en compatibilité Shibboleth de nuxeo, nuxeo est un fournisseur de services.

2 Adaptation de nuxeo à Shibboleth

2.1 Introduction

Sur le site de la fédération RENATER on trouve des documents techniques sur la façon d'installer la brique technique Shibboleth sur un serveur Web (https://federation.renater.fr/docs/installation-sp).

Il est à noter que tout le travail de dialogue SAML est entièrement pris en charge par le serveur apache en frontal de l'application. L'application n'a plus qu'à manipuler des entêtes http pour lire l'ensemble des attributs véhiculés depuis le fournisseur d'identités correspondant à l'utilisateur courant.

Néanmoins, il reste, côté application, tout un travail fonctionnel pour savoir quand demander l'authentification, quels attributs rapatrier, pour quel usage, etc.

Sur le site de la fédération RENATER on trouve un support de formation donnant des exemples de « Shibbolisation » d'applications (https://federation.renater.fr/formations/index)

2.2 User et Groupes

Shibboleth permet d'authentifier un utilisateur mais aussi de récupérer des attributs caractérisant cet utilisateur.

2.2.1 Authentification utilisateur

Shibboleth permet de garantir qu'un utilisateur s'est bien authentifié sans pour autant fournir une quelconque information nominative au fournisseur de services. Dans le cadre de la shibbolisation de nuxeo nous ne nous plaçons pas dans un tel cas d'utilisation. Au contraire, nous partons du principe que nous avons un attribut précis permettant de clairement identifier l'utilisateur. Il faut prévoir un paramétrage permettant de spécifier le nom de l'attribut à utiliser pour ce besoin. Nous utiliserons, en première approche l'email. En effet, l'email est l'information la plus facile à manipuler et elle est bien connue des utilisateurs. De plus, dans l'environnement universitaire, la gestion de cet attribut est bien maitrisée, évitant tout risque d'usurpation de cet attribut.

2.2.2 Notion de groupe shib

Nous introduisons ici la notion de « groupe shib ». Cette notion n'existe pas formellement dans la documentation shibboleth. Néanmoins, dans la mesure où l'on peut, pour un utilisateur donné, récupérer, via les mécanismes shibboleth, des attributs caractérisant l'utilisateur, on peut regrouper des utilisateurs ayant les mêmes caractéristiques.

Exemple : On peut donc définir un groupe « étudiants » dans lequel on pourrait rattacher, dynamiquement à la connexion, tout utilisateur arrivant avec un attribut *affiliation* valant *student*.

Dans l'exemple, ci-dessus, j'ai volontairement mis l'expression « dynamiquement à la connexion ».

En effet, contrairement à ce qui existe actuellement dans nuxeo pour les groupes LDAP où il est possible de connaître tous les membres d'un groupe donné, il n'est pas possible de connaître les membres d'un « groupe Shib ». Ceci de part la nature même des mécanismes shibboleth qui ne permet pas une interrogation centralisée ou même toute forme de « provisioning ».

2.3 Shibbolisation de nuxeo

Dans le cadre de la shibbolisation de nuxeo nous identifions plusieurs actions :

- Authentification de l'utilisateur
- Saisie de la définition des « groupes shib »
- Adaptation de la saisie des droits d'accès sur les documents nuxeo
- Autorisation de l'utilisateur

2.3.1 Authentification de l'utilisateur

2.3.1.1 Limitations du système actuel

Aujourd'hui nous utilisons nuxeo avec une authentification invité et CAS.

Dans le cadre d'un passage à shibboleth nous imaginons ne plus utiliser CAS mais le remplacer par shibboleth.

Note : En effet, shibboleth est un mécanisme plus général que CAS mais qui ne fait pas perdre le bénéfice de SSO apporté par CAS. En effet, dans la pratique, l'authentification auprès du fournisseur d'identités Shibboleth est réalisé grâce à un mécanisme CAS.

Le fonctionnement actuel (testé en nuxeo 5.2 GA) de nuxeo utilisant une authentification invité et CAS n'est pas satisfaisant.

Ceci pour deux raisons :

1. Il n'est pas possible actuellement de donner un lien à un utilisateur qui lui permette de se retrouver dans un environnement nuxeo en étant authentifié CAS. En effet, les authentifications invité et CAS étant chainées, la solution pour se connecter CAS est de donner une adresse de logout (du mode invité) à l'utilisateur. Malheureusement, si ce dernier est déjà connecté CAS il se retrouve alors dans un mode invité. Un même URL ne donne donc pas le même résultat suivant l'état initial.
2. Il n'est pas possible de donner un lien vers un document privé. Si on le fait, que l'utilisateur n'est pas encore authentifié, alors ce dernier se trouve redirigé, après authentification, vers la page par défaut de nuxeo mais pas sur le document ciblé.

Dans le cadre de ce cahier des charges nous demandons à ce que ces limitations soient levées.

De plus, quand on se connecte à un fournisseur de services shibboleth on est d'abord redirigé vers un service demandant à l'utilisateur courant de spécifier son fournisseur d'identités. Ce service est généralement nommé WAYF (pour Where Are You From). Il est néanmoins possible de forger un « URL direct » permettant de forcer l'accès à un fournisseur de services shibboleth via un fournisseur d'identités spécifique. Ceci est particulièrement utile quand on veut donner accès à un fournisseur de services depuis un portail sur lequel on est déjà authentifié.

Nous demandons que ce mécanisme de « URL direct » puisse fonctionner dans le cadre de la shibbolisation de nuxeo.

Note : Dans la mesure où la première limitation mentionnée ci-dessus sera résolue (cf. tickets SUPNXP-1081 et SUPNXP-1565) il ne devrait pas y avoir de difficultés à répondre à cette

demande.

2.3.1.2 Authentification d'un utilisateur sans IdP

Comme précisé précédemment nous considérons que tous les utilisateurs disposent d'un fournisseur d'identités afin de pouvoir accéder à nuxeo. Ceci notamment parce-que, dans le cadre de la fédération RENATER, il existe un fournisseur d'identités spécifique qui est le SAC (Service d'Authentification du CRU). Il est accessible librement à toute personne suite à un enregistrement en ligne. Une fois enregistré un utilisateur peut donc s'authentifier pour accéder à un fournisseur de services. Bien entendu, cet utilisateur ne sera pas obligatoirement autorisé à interagir avec le fournisseur de services en fonctions du paramétrage de ce dernier.

Note : Le SAC apparaît dans la liste des IdP proposés par le service WAYF ce qui permet à un utilisateur de sélectionner l'IdP de son établissement ou le SAC suivant son cas.

Ceci nous conduit à NE PAS vous demander de gérer dans nuxeo un mécanisme d'authentification basé sur l'envoi d'un mot de passe par email à tout utilisateur en faisant la demande. Néanmoins, ce mécanisme existe dans bien d'autres produits que nuxeo (alfresco par exemple) et semble très intéressant. Vous avez bien compris que dans le contexte de la shibbolisation de nuxeo et grâce à l'existence du SAC nous n'avons pas besoin d'un tel mécanisme. Mais nous avons conscience que notre demande sur shibboleth va, peut-être, dans le sens d'une refonte des mécanismes d'authentification de nuxeo et nous souhaitons attirer votre attention sur ce besoin qu'il convient sans doute d'intégrer -à terme- si une telle refonte devait être envisagée.

2.3.1.3 Exemple de mise en œuvre

Le site Web des projets ORI-OAI (http://ori-oai.org/) et ESUP-Portail (http://esup-portail.org/) utilisent le produit confluence. Ce dernier, a un comportement, vis à vis de shibboleth conforme à nos attentes :

- Possibilité de naviguer librement sur le site pour toute information publique
- Demande d'authentification pour tous accès à un URL pointant vers un document non public et retour sur ce document, une fois la phase d'authentification réussie.

2.3.2 Définition des « groupes shib »

Nuxeo offre aujourd'hui la possibilité de définir des groupes et leurs membres dans un annuaire LDAP. Une interface d'administration est prévue pour cela.

Nous demandons qu'il soit possible d'aussi définir des « groupes shib » via une interface d'administration dédiée.

2.3.2.1 Syntaxe d'une définition de groupes

Aujourd'hui le module de WorkFlow ORI-OAI permet aussi de définir des « groupes shib ». Pour des raisons de cohérence nous souhaitons que la syntaxe de définition de ces groupes soit la même.

Voici un exemple de définition de « groupe shib » du module de WorkFlow ORI-OAI :

```
shibAttrsMap.get("Shib-EP-UnscopedAffiliation").equals("member") ||
shibAttrsMap.get("Shib-EP-PrimaryAffiliation").equals("professor")
```

Note : Les expressions sont compilées avec Janino (http://www.janino.net/) ce qui offre une grande souplesse dans l'expression de la règle.

2.3.2.2 Fonctionnalité de l'interface

Nous ne spécifions pas précisément la forme et l'ergonomie de l'interface de définition des groupes. Nous demandons que nuxeo nous fasse des propositions suffisantes en terme d'ergonomie et conformes au reste de la plateforme.

Nous spécifions néanmoins les fonctionnalités attendues :

- Possibilité de saisir la définition d'un groupe
- Possibilité de retrouver, éditer, supprimer une définition préalablement saisie (prévoir le cas de quelques dizaines à quelques centaines de définitions)
- La saisie d'une définition peut se limiter à un simple champ de saisie. Il conviendrait néanmoins d'offrir un mécanisme permettant de se prémunir de saisies syntaxiquement incorrectes.

2.3.3 Saisie des droits d'accès

Aujourd'hui, la saisie des droits d'accès dans nuxeo se fait via une zone de saisie avec proposition, via une liste de choix, fonction des 3 premiers caractères saisis pour les users (bizarrement il nous semble que ce soit 5 caractères pour les groupes).

Note : Les utilisateurs et les groupes sont représentés dans la liste de choix avec un pictogramme différent.

2.3.3.1 Sélection des groupes shib

Pour les « groupes shib » nous demandons à ce que ces groupes apparaissent dans la liste de choix, comme les groupes LDAP mais, éventuellement, avec un pictogramme différent (ex : symbole actuel pour les groupes + une planète).

De plus, l'ergonomie actuelle de sélection des groupes nous semble insuffisante ou mal adaptée au volume de groupes que nous avons potentiellement à traiter. Dans la mesure où l'amélioration de l'ergonomie de la sélection des groupes n'est pas spécifique aux « groupes shib » merci de vous reporter au paragraphe 2.4.2 pour plus de précision sur la demande.

2.3.3.2 Sélection des utilisateurs shib

Pour les utilisateurs shibboleth nous demandons à ce qu'il soit possible de saisir l'attribut shibboleth de référence (typiquement une adresse email).

2.3.4 Autorisation de l'utilisateur

Aujourd'hui, pour contrôler l'accès à une ressource dans nuxeo on vérifie que l'utilisateur -ou un des groupes auxquels appartient l'utilisateur- a bien le droit d'accéder à cette ressource (en lecture, écriture ou autre).

Dans la mesure où la shibbolisation de nuxeo proposée ici s'appuie sur les mêmes notions d'utilisateurs et de groupes le fonctionnement ne doit pas évoluer. Néanmoins il est à noter que la détermination de l'appartenance d'un utilisateur à un « groupe shib » se fait en évaluant une règle contenue dans la définition du « groupe shib ». Ceci a peut-être un impact sur la logique interne de nuxeo que nous ne maitrisons pas (Notification, etc.). Nous tenons donc à attirer l'attention de nuxeo sur ce point. Il semble qu'évaluer l'appartenance ou non de l'utilisateur aux « groupes shib » pourrait se faire au moment de la connexion.

Info : C'est le fonctionnement actuel de U-Portal avec la notion de groupes PAGS. U-Portal est la solution de portail JSR168 utilisée dans le cadre du projet ESUP-Portail.

2.4 Divers

2.4.1 Compatibilité shibboleth/LDAP

Nous avons vu dans ce cahier des charges comment gérer des utilisateurs shib et des « groupes shib ». Néanmoins, il reste pratique de pouvoir continuer à travailler avec des utilisateurs et des groupes LDAP pour les raisons suivantes :

- Non besoin de définir des groupes Shib via des règles si ces derniers ne contiennent que des utilisateurs LDAP
- Facilité de sélection des utilisateurs dans les listes de choix (à comparer avec la saisie d'un email)
- Etc.

On aurait donc un mécanisme LDAP pour les utilisateurs et groupes « locaux à l'établissement » (établissement où est installé nuxeo) et un mécanisme shibboleth pour les utilisateurs et les groupes « externes à l'établissement ».

Néanmoins, si on décide d'utiliser un mécanisme d'authentification shibboleth, tous les utilisateurs, qu'ils soient locaux ou externes à l'établissement, s'authentifieront via shibboleth. Il faut donc trouver un mécanisme garantissant une compatibilité shibboleth/LDAP.

En première approche il est proposé le mécanisme suivant :

- Enregistrement, dans la configuration du module d'authentification shibboleth pour nuxeo, des ID du (ou éventuellement des) fournisseur d'identités considéré comme étant le fournisseur d'identités correspondant à l'établissement d'installation.
- Si l'utilisateur courant se connecte via ce (éventuellement ces) fournisseur d'identités alors on recherche dans les attributs véhiculés via shibboleth l'attribut correspondant à son UID local (nom d'attribut configurable) et c'est lui qui est utilisé comme UID nuxeo.
- Si l'utilisateur courant se connecte via un autre fournisseur d'identités alors on recherche dans les attributs véhiculés via shibboleth l'attribut shibboleth de référence (généralement l'email. cf. 2.2.1) et c'est lui qui est utilisé comme UID nuxeo.

Nous demandons à ce que la compatibilité LDAP/Shib existe.

2.4.1.1 Compatibilité avec d'autres modules

Le mécanisme d'authentification shibboleth est un mécanisme basé sur des redirections http. Néanmoins, nous imaginons que, certains autres modules (live edit ?, SharePoint ?) nécessitent une authentification client/serveur classique avec transmission d'un user/password.

Nous demandons que ces autres modules restent opérationnels pour les utilisateurs « locaux à l'établissement » (établissement où est installé nuxeo).

2.4.2 Sélection des groupes

L'ergonomie actuelle de sélection des groupes nous semble insuffisante ou, du moins, mal adaptée au volume de groupes que nous avons potentiellement à traiter.

2.4.2.1 Comment fonctionnent ou vont fonctionner les universités ?

Conformément à la norme supAnn 2008 (cf. http://www.cru.fr/documentation/supann/) les universités mettent leurs groupes au même niveau dans leur LDAP (ou=groups) et utilisent un nommage permettant d'organiser ces groupes.

Les établissements qui utilisent ou vont utiliser *grouper* (Cf. http://www.internet2.edu/grouper/) ont des groupes dont le nom est du type « Tous:Personnels:Service1:Direction ». Le caractère « : » permettant de séparer chaque branche d'un arbre de groupes.

Note : ESUP-Portail initie un travail afin d'inciter les établissements à déployer *grouper*.

2.4.2.2 Comment améliorer la sélection des groupes ?

Nous demandons à ce qu'il soit possible de sélectionner un groupe en utilisant une IHM sous forme d'un arbre. Un peu à la manière dont on sélectionne une section lors de la publication.

L'arbre serait construit en utilisant un ou n caractères permettant de découper le nom du groupe. Dans l'exemple ci-dessous le caractère permettant le découpage est « : » mais il est intéressant d'utiliser, en complément, 0, 1 ou n autres caractères (par ex. « _ »). Ceci afin de permettre la plus grande souplesse d'adaptation à un existant en terme de groupes LDAP.

Exemple : Tous:Personnels:Service1:Direction

Devient :

- Tous
  - Personnels
    - Service1
      - Direction

Dans l'exemple ci-dessus, « Direction » peut être sélectionné par l'utilisateur car il constitue une feuille de l'arbre. « Service1 » ne l'est pas, sauf si on trouve aussi dans le LDAP un groupe de nom « Tous:Personnels:Service1 »

Compatibilité avec les « groupes Shib »

Nous demandons à ce qu'il soit possible de rattacher les « groupes shib » à cet arbre de groupes. Ceci à un emplacement donné (configurable) de l'arbre. Soit, par exemple :

- Tous:Extérieurs (afin d'avoir une branche « Extérieurs » parallèle à « Personnels »)
- Autres:TousLesAutres (afin d'avoir une nouvelle racine « Autres » parallèle à « Tous »)

De même, la règle de découpage applicable aux noms des groupes LDAP doit pouvoir s'appliquer au « groupes shib » définis dans nuxeo. Ceci afin de permettre leur sélection sous forme arborescente par l'utilisateur final.

2.4.3 Mixage de groupes

Nuxeo permet actuellement de gérer des sous groupes. Il nous semble intéressant d'explorer la possibilité de définir des groupes contenant à la fois des groupes LDAP et des « groupes shib ». Ceci afin de permettre de facilement mixer des utilisateurs locaux ou externes à l'établissement d'installation.

Exemple : groupe de travaille « projet web » intégrant le groupe LDAP « équipe Web » et le groupe shibboleth « équipe de développement de la SSII truc ».

Nous demandons à ce que nuxeo chiffre cette possibilité sous forme d'une variante.