La gestion des identités dans l'Éducation Nationale, état des lieux et perspectives

Alexandre Guyot
Pôle de compétences
DSI - Rectorat d'Orléans-Tours, 10 rue Molière 45000 Orléans

Nicolas Romero
Pôle de Compétences, Equipe Gestion des Identités
DSI - Rectorat d'Orléans-Tours, 10 rue Molière 45000 Orléans

Résumé

La gestion des identités dans le Système d'Information de l'Éducation Nationale est un projet dont les bases ont été posées à la fin des années 90 avec le projet @melouvert*, visant à fournir à tous les personnels une adresse de courriel, et donc un couple identifiant/mot de passe. Le passage des principales applications de gestion (*Sconet, gestion de la scolarité du 2nd degré*) du mode client-serveur, installées dans chaque établissement, au web, mutualisées au niveau académique, ainsi que l'augmentation du porte-feuille des applications nationales (BE1D, gestion des élèves dans le 1er degré) a ensuite rendu indispensable une réflexion poussée sur la gestion des identités. Les objectifs étaient d'élever le niveau général de sécurité, de masquer la complexité du SI du point de vue utilisateur, et d'uniformiser l'implémentation et l'administration des politiques d'accès aux applications.*
La prise en compte de ces problématiques ne pouvait être que globale, pérenne, et comprise par toutes les parties prenantes du SI : maîtrises d'ouvrages, maîtrises d'œuvres, équipes de développements, équipes de production et utilisateurs. Dans une organisation à la fois très centralisée et très hétérogène (un ministère, 30 académies, 65000 établissements scolaires, un million d'agents), une démarche aussi structurante ne pouvait s'envisager qu'en s'appuyant sur des choix organisationnels et techniques forts. Pour cela, l'effort a porté sur la standardisation à tous les niveaux : référentiels, technologies, infrastructures, méthodes.
Cette démarche a permis d'intégrer progressivement SSO, authentification forte, et fédération d'identité. Mais pensé et mis en œuvre dans un périmètre interne, le projet doit aujourd'hui évoluer pour faire face à un nouveau défi : celui de l'ouverture à des publics dont les identités ne sont plus entièrement maîtrisées par l'Éducation Nationale, tels que parents, élèves, collectivités locales, ou prestataires privés.

Mots clefs

SSO, habilitations, authentification forte, OTP, fédération d'identités, SAML

1 Introduction

Les premières réflexions sur la gestion des identités dans le Système d'Information de l'Éducation Nationale ont débuté à la fin des années 90 avec la généralisation de la messagerie et la migration progressive des applications de gestion d'un mode client-serveur à un mode web. La diffusion à grande échelle des identifiants et mots de passe de messagerie, combinée à la multiplicité des applications ont alors conduit à envisager de donner aux utilisateurs la possibilité d'utiliser les mêmes authentifiants pour y accéder.

Au cours des dix dernières années, l'Éducation Nationale a progressivement intégré le SSO, puis l'authentification forte et la fédération au cœur de son informatique, rendant la gestion des identités incontournable dans le développement de son Système d'Information. Pour cela, l'effort a porté sur la standardisation à tous les niveaux : référentiels, technologies, infrastructures, méthodes.

Ainsi, la gestion des autorisations et des habilitations et, éventuellement, la fédération des identités est réalisée de manière uniforme quelle que soit l'application, et repose sur :

- une architecture technique construite à l'identique dans chaque académie,
- un ensemble d'attributs Ldap normalisés et utilisés par l'intermédiaire d'une bibliothèque Java ou d'un service Web spécifique,
- un document standard de définition des politiques d'habilitations :la fiche de politique d'habilitations, décliné pour toutes les applications nationales.

Toute application amenée à être intégrée dans le SI de l'Éducation Nationale doit être compatible avec ce cadre.

Cet article présente la démarche suivie par l'Éducation Nationale pour gérer les identités dans son Système d'Information, dresse un état des lieux de l'existant, et décrit les évolutions nécessaires aujourd'hui.

2 Contexte

2.1 La structure de l'Éducation Nationale

L'organisation complexe et pyramidale, et la taille de l'institution Éducation Nationale ont influencé la constitution de son Système d'Information.

Au sommet, le ministère (« La Centrale ») représente le niveau national. Le territoire français est découpé en une trentaine d'académies (correspondant approximativement aux régions et territoires d'Outre-Mer). Chaque académie est gérée par un Rectorat, qui s'appuie sur les Inspections Académiques au niveau départemental. Les 65000 établissements scolaires et le million d'agents (enseignants et personnel administratif et technique) forment la base de cette pyramide et s'adressent à un public de 12 millions d'élèves et à leurs parents.

Ainsi, le SI de l'Éducation Nationale s'est construit comme une juxtaposition de SI, la Centrale et chaque Rectorat gérant chacun la population de son périmètre avec sa propre infrastructure. Dans ce contexte, afin de garantir une cohérence forte, le Ministère pilote les infrastructures et les applications informatiques nationales, les académies étant chargées de la mise en œuvre.

2.2 Un peu d'histoire

Le SI de l'Éducation Nationale s'est longtemps limité à l'informatique de gestion. Chaque application était soit constituée d'un client lourd, installé sur le poste de travail local, avec parfois,un mode client-serveur, ou à défaut une synchronisation périodique, soit accédée en mode caractère au travers de terminaux. A l'échelle de l'Éducation Nationale, la population d'utilisateurs était alors assez réduite, et les identités gérées de manière indépendante à l'intérieur de chaque application. La plupart des comptes étaient fonctionnels, et partagés.

La généralisation de la messagerie professionnelle *@melouvert*, visant à fournir à tous les personnels une adresse de courriel, et donc un couple identifiant/mot de passe, a remis en question cet état de fait. La migration en Java ou l'écriture des principales applications de gestion des établissements scolaires (Sconet, BE1D), devenues accessible par le web et mutualisées au niveau académique, a ensuite rendu indispensable une réflexion poussée sur la gestion des identités. D'une part, les comptes de connexion aux applications étant les mêmes que pour l'accès aux messageries professionnelles, il n'était plus question de permettre leur partage, tout en permettant d'utiliser le compte personnel pour accéder aux différentes applications. D'autre part, il devenait plus facile d'atteindre les applications en utilisant un simple navigateur : le contrôle des accès devenait crucial. De plus, certaines applications sensibles nécessitaient une sécurisation accrue : il fallait trouver le moyen de hausser le niveau d'authentification, sans gêner l'expérience utilisateur. Enfin, certaines applications devenant mutualisées au niveau national plutôt qu'académique, il fallait permettre la transmission de l'identité depuis le guichet d'authentification académique vers ces applications.

Une démarche aussi structurante ne pouvait s'envisager qu'en s'appuyant sur des choix organisationnels et techniques forts. L'Éducation Nationale a commencé par étudier les différents produits leaders dont Netegrity Siteminder, Sun Identity Server, RSA Cleartrust mais aussi des produits du libre comme Shibboleth ou CAS. Cette étude a permis de recenser les possibilités des différents produits pour voir comment ils pouvaient répondre aux besoins, et vérifier que l'expression de besoin était réaliste. Cette phase a servi aussi à se familiariser avec les notions de gestion d'identités.

En 2004, les produits libres CAS et Shibboleth ne remplissaient pas tous les critères techniques et fonctionnels que s'imposait l'Éducation Nationale, avec comme priorités le respect des standards (SAML 1.1, Liberty Alliance), la volumétrie (1 million d'utilisateurs), la possibilité de gérer les habilitations et l'intégration de la solution avec les technologies utilisées par ailleurs dans le Système d'Information (J2EE et Redhat Linux). Notamment, CAS ne permettait que le SSO mais sans gestion centralisée des habilitations, et Shibboleth n'était pas compatible Liberty Alliance. C'est pourquoi l'Éducation Nationale a lancé cette année-là un appel d'offres qui a abouti au choix du produit RSA Cleartrust.

Plus récemment, RSA Federated Identity Manager (SAML 2.0) et RSA Authentication Manager (Securid) ont complété le portefeuille des briques techniques pour apporter les fonctionnalités de fédération d'identités et l'authentification forte par OTP (One Time Password).

3 Aspects techniques

3.1 Architecture

Le Système d'Information de l'Éducation Nationale étant maintenant presque exclusivement constitué d'applications web, l'architecture générale des Systèmes d'Information académiques est décomposée de manière classique en couches :

- la couche **Présentation** est le point d'accès aux applications ; elle est constituée de boîtiers de répartition de charge et de serveurs frontaux embarquant des agents interconnectés avec la couche suivante,
- la couche **Gestion des Identités** interagit avec la couche présentation afin de contrôler l'accès aux applications,
- les couches **Traitements** et **Données** hébergent les applications et leurs données.

D'un point de vue réseau, on distingue trois zones d'accès aux Systèmes d'Information, avec des niveaux de confiance différents :

- la zone **DMZ**, accessible depuis Internet, est considérée comme peu sûre : les accès aux applications sensibles sont interdits, ou doivent impérativement être chiffrés (https) et protégés par authentification forte. Nous pouvons retrouver dans cette zone l'ensemble des personnels du ministère, utilisant par défaut une authentification faible, mais également un certain nombre d'acteurs nomades et/ou situés sur des sites exogènes, qui peuvent utiliser une authentification forte
- la zone **Agriates**, accessible depuis le réseau administratif des établissements scolaires du second degré (EPLE), est considérée comme assez sûre : seules les applications à destination des agents des EPLE sont présentées dans cette zone; leur accès peut être chiffré (https) et protégé par authentification forte
- la zone **Racine**, accessible depuis le réseau administratif des Rectorats, Inspections Académiques et Ministère, est considérée comme sûre : l'accès aux applications présentées dans cette zone peut être chiffré (https) et protégé par authentification forte. Les applications accessibles dans cette zone sont principalement des applications orientées vers les services de gestion des services.

3.2 Infrastructure de gestion des Identités

Les principaux composants techniques et applicatifs impliquées dans la gestion des Identités sont d'une part des briques d'infrastructures s'appuyant le plus possible sur des standards (LDAP, SAML 2.0, J2EE, SOAP) et d'autre part des applications ou des bibliothèques développées par l'Éducation Nationale afin de garantir la cohérence des développements et des traitements liés à la gestion des identités au niveau académique.

L'infrastructure de gestion des identités est ainsi constituée :

- d'un référentiel des Identités pour les personnels de l'Éducation Nationale, qui est un annuaire Ldap (Annuaire Académique des Agents), présent dans chaque académie et alimenté depuis le SIRH académique ou par les applications de création de comptes,
- de briques techniques RSA :
    - serveurs *Cleartrust/Access Manager* (identification, autorisations, habilitations, SSO),
    - agents *Cleartrust/Access Manager*, embarqués dans les serveurs web de présentation,
    - serveurs *FIM* (fédération des identités),
    - serveurs *Authentication Manager* (authentification forte par OTP),
- d'un composant de gestion d'identité, sous forme de bibliothèque Java ou service Web, permettant aux applications d'utiliser les informations transmises sur les utilisateurs,
- de portails utilisateurs,
- d'applications de création de comptes,
- d'outils de délégation et d'assistance.

Les fiches utilisateurs présentes dans le référentiel d'identités LDAP comportent des attributs spécifiques à l'Éducation Nationale. Ces attributs permettent de définir les habilitations et peuvent être transmis aux applications. Les applications de ressources humaines et d'alimentation des référentiels utilisateurs ont évolué et des outils spécifiques ont été développé pour prendre cela en compte.

Avec l'apparition des applications hébergées au niveau national, il a fallu repenser cette infrastructure pour les centres d'hébergement nationaux. La Figure 1 schématise les principales briques techniques mises en œuvre, la partie gauche étant répliquée pour chaque académie, et celle de droite pour chaque centre d'hébergement national (moins d'une dizaine à ce jour).

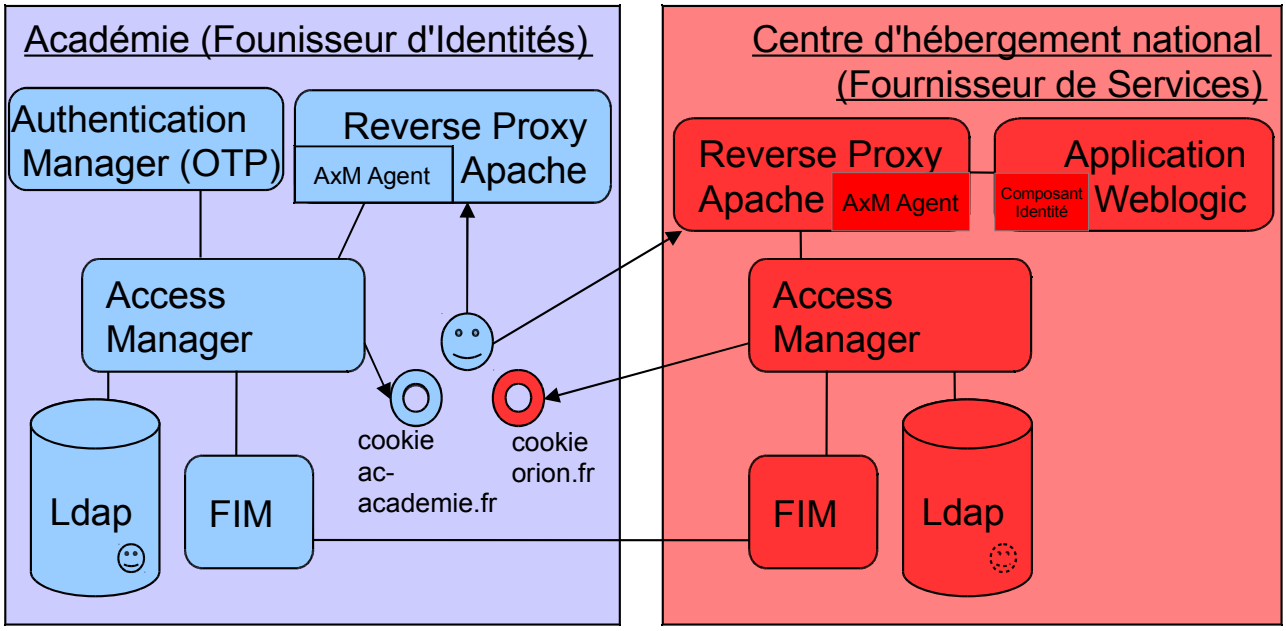

Figure 1: briques techniques de gestion des identités

4 Aspects organisationnels

Une gestion des identités ne se résume pas qu'à la technique et sa mise en œuvre. L'objectif à atteindre est de mettre des applications à disposition des utilisateurs en fonction de leur profil pour élever le niveau général de sécurité tout en masquant la complexité du Système d'Information du point de vue de l'utilisateur.

Pour cela il est nécessaire d'uniformiser la définition, l'implémentation et l'administration des politiques d'accès aux applications. L'aspect organisationnel a donc été considéré comme important dès le début du projet. Cela a nécessité beaucoup de sensibilisation auprès de toutes les parties prenantes du Système d'Information : les maîtrises d'ouvrages, les maîtrises d'œuvres, les équipes de développements, les équipes de production et bien sûr les utilisateurs.

Une équipe nationale, le Pôle de Compétences Gestion des Identités, a été créé afin d'apporter une expertise aux maîtrises d'ouvrage et maîtrises d'œuvre et de coordonner les travaux en collaboration avec le ministère et les académies, tandis que les équipes de développement ont été sensibilisées et formées et que des correspondants académiques ont été nommés et chargés de la mise en œuvre.

La gestion des Identités consiste en premier lieu à identifier l'utilisateur, afin de déterminer son profil, et d'en déduire ce qu'il peut faire. Un travail de profilage a été mené afin d'identifier les principales populations d'utilisateurs et les critères permettant de les déterminer.

Toute application amenée à être intégrée dans le Système d'Information de l'Éducation Nationale, et ayant vocation à être couplée à l'infrastructure de gestion des identités ou à tout autre équivalent technique, doit prendre en compte la gestion des identités dès le début de son développement. Les différents modules applicatifs, le niveau d'authentification requis (anonyme, mot de passe, authentification forte, ...), les zones d'accès réseaux et les profils autorisés sont étudiés afin de définir la politique d'habilitations de l'application.

La fiche de politique d'habilitations est un document standardisé décliné pour toutes les application nationales qui décrit l'ensemble des dispositifs à mettre en œuvre en académie et en administration centrale pour permettre l'accès à l'application. Elle contient notamment :

- les règles d'habilitation qui sont mises en œuvre pour gérer l'authentification et les droits d'accès à une application,
- les attributs utilisateurs sur lesquels doit se baser l'application pour déterminer le profil et le périmètre de responsabilité de chaque utilisateur,
- les niveaux d'authentifications requis selon les modules applicatifs et les zones d'accès.

Les scripts de configuration des différentes briques d'infrastructure sont déduits à partir des fiches de politique d'habilitations, qui forment un élément central du dispositif entre besoins fonctionnels et implémentation technique.

C'est pourquoi un workflow organisationnel a été élaboré pour établir et valider les fiches de politique d'habilitations. Celui-ci est représenté ci-dessous. Toute création ou modification d'une application nationale nécessite la mise en œuvre de ce workflow en vue de la validation de la fiche. Cela permet d'impliquer toutes les parties prenantes d'un projet : maîtrise d'ouvrage, maîtrise d'œuvre, équipe de développement et pas seulement les équipes techniques.

Workflow de validation des fiches de politique d'habilitations	
Etape 1 : Chefs de projets MOA/MOE *Définition de la politique d'habilitation avec la MOA et rédaction de la fiche de politique d'habilitation*	D'après le modèle type Présentation de l'application et du projet Liste des acteurs, des droits, des profils Définition des libellés dans le portail utilisateur Périmètre de responsabilité de chaque acteur Délégations
Etape 2 : CPN MOE, Responsable Dév. *Rédaction de la fiche de politique d'habilitation*	Périmètre de responsabilité de chaque acteur Liste des attributs concernés dans l'annuaire Ldap Liste des ressources avec leur adresse URL Libellés dans le portail utilisateur et dénomination détaillée Utilisation de la fédération Version du composant de gestion d'identité utilisée Définition des règles d'accès
Etape 3 : CPN MOE, Responsable Dév., Pôle Identité *Rédaction de la fiche de politique d'habilitations*	Habilitations déclaratives
Etape 4 : Pôle Identité, Dév. Identité, Diffusion Identité *Etude des besoins d'évolution des outils et des attributs LDAP*	Evolutions réalisées si nécessaire après validation du ministère
Etape 5 : CPN MOE, Responsable Dév., Pôle Identité *Validation de la fiche de politique d'habilitations*	Tant que la fiche n'est pas validée, retour à l'étape 2
Etape 6 : Pôle Identité, Diffusion Identité *Création et diffusion des scripts de configuration*	Seulement si la fiche est validée

5 L'ouverture aux ENT

Depuis les lois de décentralisation, certains personnels et certaines compétences dévolues à l'Éducation Nationale ont été transférées aux collectivités territoriales (régions, départements, communes). L'équipement 'informatique pédagogique, notamment, est maintenant un élément supplémentaire à prendre en compte par les collectivités pour assurer le bon fonctionnement des établissements.. Ainsi, les espaces numériques de Travail (ENT), projet initié par le Ministère de l'Éducation Nationale et portés en commun par les académies et par les collectivités de rattachement, à destination des enseignants, des élèves et de leurs parents, connaissent un essor certain depuis deux ans. Parallèlement à cette mise en place, l'Éducation Nationale poursuit la modernisation et l'enrichissement de son système d'information et propose de plus en plus de services, appelés Télé-services (TS), tels que l'accès aux notes ou aux absences, à ces publics.

Les télé-services et les ENT ont des profils d'utilisateurs cibles en commun, mais les élèves et leurs parents n'entraient pas dans le périmètre de la gestion d'identités de l'Éducation Nationale. En effet, ces publics ne disposent pas d'une identité dans les annuaires et référentiels d'identité académiques. De la même manière, les personnels Éducation Nationale en établissement ont vocation à accéder à la fois aux ENT et au Système d'Information de l'Éducation Nationale. Il a donc été nécessaire de proposer des solutions techniques et organisationnelles permettant de gérer au mieux l'identité de ces personnes et l'interaction entre les ENT et les Télé-services, de façon à pouvoir naviguer de l'un à l'autre sans ré-authentification. Cette situation, avec des personnes pouvant disposer d'identités gérées par des acteurs différents, fait apparaître la notion de guichet d'identité ou de guichet d'authentification, avec des positionnements forts du ministère en fonction des publics ciblés. Ces éléments sont décrits dans la version 3 du Schéma Directeur des Espaces numériques deTravail (SDET) [1]. Il faut alors considérer les éléments suivants

- Un agent doit avoir son identité gérée par son administration de tutelle. Ainsi, un agent de l'éducation nationale doit utiliser une identité délivrée par le ministère,
- Un élève peut utiliser une identité fournie par l'ENT ou par son ministère de « rattachement » (l'académie pour les élèves de l'éducation nationale, l'agriculture pour les élèves des établissements agricoles, …)
- Un parent d'élève peut utiliser une identité fournie par l'ENT ou par le ministère de « rattachement » de son enfant (l'académie pour les élèves de l'éducation nationale, l'agriculture pour les élèves des établissements agricoles, …)
- Toute autre personne devant être acteur de l'ENT peut disposer d'une identité gérée par l'ENT ou par son organisme, sous réserve de mettre en oeuvre les accords de confiance adéquats.

Le SDET propose 3 modèles qui permettent de réaliser les différentes articulations précédentes, le choix du modèle étant laissé libre aux porteurs du projet ENT. Ce cadre technique décrit également les flux d'informations nécessaires (notamment les vecteurs d'identité) et les scénarios interdits.

Parmi ces trois modèles, le modèle dit « Modèle 2 », décrit dans la figure 2, explique les principaux mécanismes mis en œuvre, ainsi que leur séquencement, lorsque l'éducation nationale est considérée comme le guichet d'identité pour les élèves ou parents d'élèves de son périmètre. La gestion des comptes est alors réalisée par une brique technique nommée ATEN (Accès aux Télé-services de l'Éducation Nationale), brique fortement couplée au système d'information de gestion de la scolarité du 2nd degré. Dans ce modèle, l'acteur (élève ou parent d'élève), est libre de choisir son portail d'accueil. Par contre, il doit forcément utiliser le guichet d'identité de l'Éducation Nationale pour s'authentifier et accéder aux service auxquels il a le droit d'accéder. Dans ce cas, les droits d'accès et le périmètre d'intervention des acteurs est très fortement couplé au système d'information de l'EPLE (Établissement Public Local d'Enseignement). Ainsi, c'est un acte de gestion au sein de l'EPLE qui pourra permettre d'autoriser, ou non, l'accès aux différents services.

La figure 2 illustre un des modèles d'articulation retenu dans le cadre du SDET v3. Il consiste à utiliser le guichet d'authentification de l'Éducation Nationale pour authentifier les élèves ainsi que leur responsable, qu'ils accèdent à l'ENT (première partie : Accès aux modules ENT) ou aux télé-services de l'Éducation Nationale (deuxième partie : Accès aux TS ou applications externes hébergées).

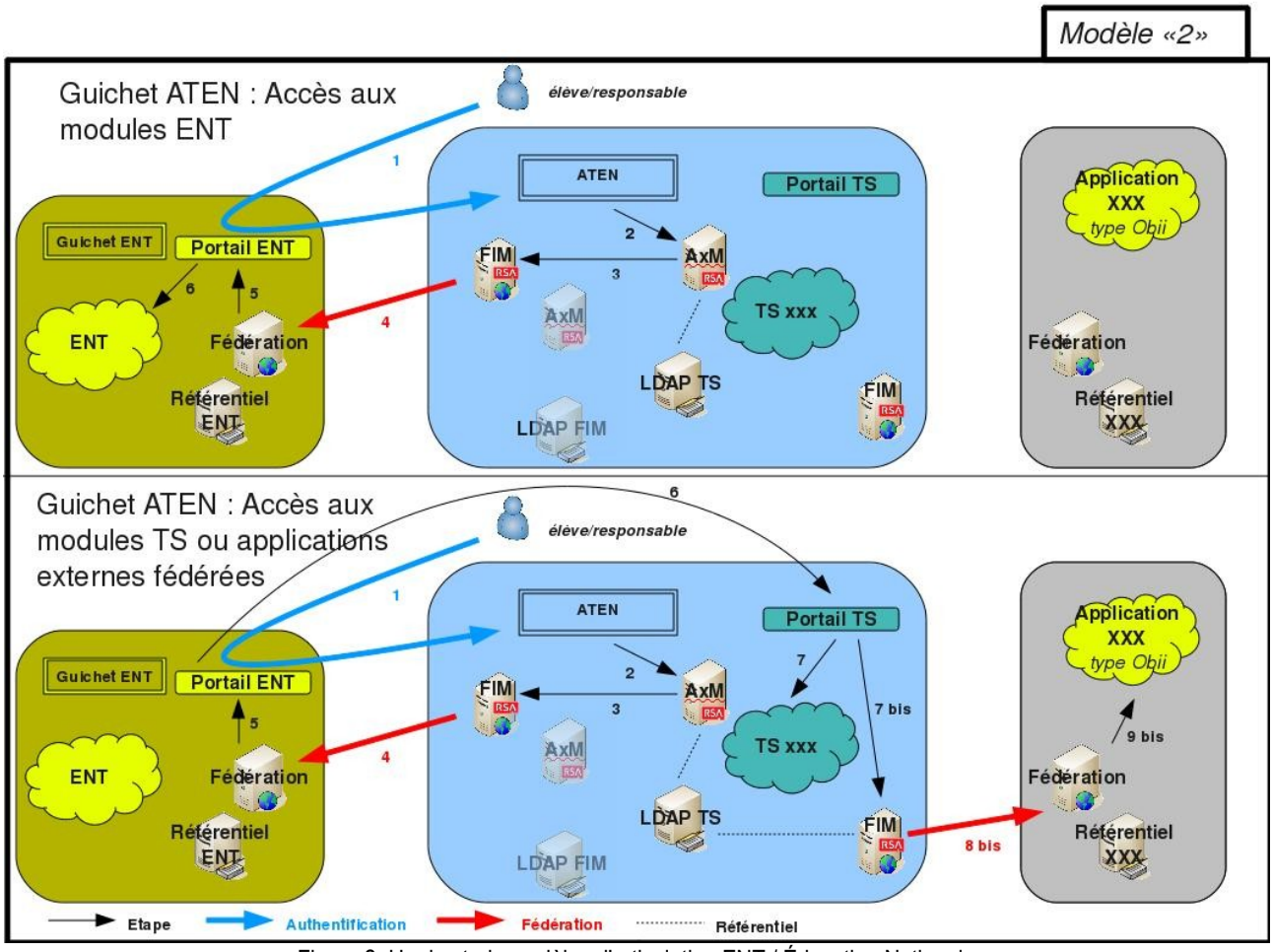

Figure 2: Un des trois modèles d'articulation ENT / Éducation Nationale

La cinématique est alors la suivante, dans un mode de type SP Initiated, à savoir où l'utilisateur essaie d'accéder au service avant toute action d'authentification :

(1) L'utilisateur accède à l'ENT : il choisit sa catégorie (« Elève»), catégorie mémorisable. Il est redirigé vers le guichet ATEN ;

(2) L'élève s'authentifie sur ATEN (IDP, Identityp Provider ou fournisseur d'identité associé à l'élève)

(3 - 4) Une fédération d'identité est alors mise en œuvre

(5) L'utilisateur accède au portail ENT en étant authentifié

(6) L'utilisateur accède aux différents SP (Service Provider ou fournisseur de Service) possibles, soit l'ENT, soit le portail télé-services, soit tout autre SP définit dans le cercle de confiance.

(7 – 9) Dans le cas de services hébergés nationalement à l'Éducation Nationale, une fédération entre le guichet ATEN académique et le SP national est alors mise en œuvre.

Ce fonctionnement s'appuie sur les principes de base des modèles de fédération d'identité entre un fournisseur d'identités (ici l'académie), et un ou plusieurs fournisseurs de services (ici l'ENT, l'académie ou tout autre fournisseur de service étant entré dans le cercle de confiance de l'ENT). Dans ce contexte, les notions d'IDP discovery et de « Where Are You From » sont les bases techniques de la mise en œuvre de l'ensemble des modèles décrits dans le SDET. L'IDP discovery permet à un utilisateur d'être redirigé sur le bon guichet d'identité en fonction de son appartenance à telle ou telle entité. La figure suivante montre un exemple de mire de connexion utilisée par l'académie de Nantes dans un projet mené en partenariat avec les collectivités.

Figure 3: Exemple de mire de connexion de l'ENT « Nantais »

L'articulation présentée ici, comme ses variantes décrites dans le SDET, permettent de replacer chaque acteur de l'ENT au cœur du dispositif d'authentification et de bien afficher les responsabilités et limitations de chacun.

A l'heure actuelle, un travail d'accompagnement des académies et des porteurs de projets ENT est mené pour mettre en œuvre ces recommandations et devrait aboutir avant fin 2012 à une généralisation des ENT et des Télé-services pour tous les établissements du second degré.

6 Conclusion

Pensée et mise en œuvre dans un périmètre interne à l' Éducation Nationale, alors que l'organisation était encore très centralisée, la gestion des identités doit aujourd'hui évoluer. L'ouverture à des publics dont les identités ne sont plus entièrement maîtrisées par l'Éducation Nationale, tels que parents, élèves, collectivités locales, prestataires privés, Enseignement Supérieur doit se poursuivre.

D'autre part, la ré-urbanisation de l'informatique de l'Éducation Nationale, qui a commencé avec la migration vers les applications web J2EE, amène à de plus en plus de mutualisation. De plus en plus souvent, les nouvelles applications sont hébergées au niveau national, et fédérées avec les niveaux académiques.

Enfin, la refonte du Système d'Information de gestion des Ressources Humaines, avec le projet SIRHEN, bouleverse le cycle de vie des identités, des profils et des données utilisateurs et amène à repenser l'architecture et le contenu des référentiels d'identités.

L'Éducation Nationale a intégré progressivement SSO, authentification forte, et fédération en mettant toujours l'accent sur la standardisation à tous les niveaux : référentiels, technologies, infrastructures, méthodes. Ces choix ont permis aujourd'hui de doter l'ensemble de ses agents d'une identité numérique, couplée, pour 250 000 d'entre eux, d'un dispositifs d'authentification forte (OTP ou certificat) et ont permis également de pouvoir proposer ses différents guichets d'identités à ses partenaires dans le cadre notamment des ENT. Ce sont également ces choix organisationnels et techniques forts et la maturité acquise dans ces domaines qui lui permettront de relever ces nouveaux défis.

7 Bibliographie

[1] Schéma directeur des Espaces Numériques de Travail version 3, http://eduscol.Éducation.fr/cid56994/preconisations-techniques.html