Un Hub pour les fédérer tous

Sophie Schaal
DSI de l'académie de Rennes
96 rue d'Antrain
35000 Rennes

Olivier Adam
DSI de l'académie de Rennes
96 rue d'Antrain
35000 Rennes

Sébastien Simenel
DSI de l'académie de Rennes
96 rue d'Antrain
35000 Rennes

Long Ya
DSI de l'académie de Rennes
96 rue d'Antrain
35000 Rennes

Résumé

Dans l'académie de Rennes, depuis 2007, nous mettons à disposition un espace numérique de travail (ENT) « Toutatice » pour tous les acteurs de l'éducation du premier et second degré en Bretagne, soit environ 1,5 million d'utilisateurs.

Comment identifier simplement les agents, les élèves et leurs représentants légaux pour l'accès aux offres de services numériques portées par le ministère, l'académie, les collectivités et les établissements en tenant compte de la diversité des protocoles d'authentification et de l'APIsation des applications ?

En 2015, nous présentions une solution permettant de fédérer des applications SAML2 et CAS derrière un même fournisseur de services avec un module CASShib permettant de « shibbolétiser » un serveur CAS.

Depuis 2017, dans le cadre du projet PIA[1] IPANEMA, nous faisons évoluer cette solution pour prendre en compte le protocole OIDC - OpenID Connect, le consentement utilisateur et renforcer sa disponibilité.

La solution réalisée, bâtie sur Shibboleth Identity Provider (IdP) et Service Provider (SP), est interopérable avec l'ensemble des fournisseurs d'identité académiques et nationaux. Elle est performante, hautement disponible et supervisée. Elle est compatible avec les protocoles standards actuels : CAS, SAML2 et OpenId Connect. Elle a permis de connecter l'ENT Toutatice à EduConnect, nouveau fournisseur d'identité de l'Éducation Nationale, ainsi qu'au GAR, le service d'accès aux ressources numériques porté par le Ministère.

1. Programme d'investissement d'avenir : https://www.gouvernement.fr/le-programme-d-investissements-d-avenir

Nous vous présentons ici le résultat de nos travaux en détaillant l'architecture, l'implémentation et le cheminement ayant conduit à la mise en œuvre de ce véritable Hub d'identité multi-protocolaires.

Mots-clefs

SSO, protocoles, SAML2, OpenId Connect, CAS, Shibboleth, fédération d'identité, Hub, Identity Provider, Service Provider

1 Contexte de la fédération d'identité de Toutatice

L'académie de Rennes met en œuvre l'espace numérique de travail (ENT) « Toutatice » pour tous les acteurs de l'éducation du premier et second degré en Bretagne.

1.1 Toutatice, un environnement numérique de confiance pour l'éducation en Bretagne (2006 à 2011)

Les débuts de Toutatice remontent à septembre 2006. Son socle est mis en œuvre par la DSI de l'académie pour déployer l'ENT avec, comme premier service fourni, le logiciel de gestion des centres de documentation des CDI des établissements scolaires de l'académie, PMB (www.sigb.net).

Pour satisfaire le besoin d'identification et d'autorisation, la DSI de Rennes choisit de mettre en œuvre :

- un annuaire LDAP spécifique approvisionné à partir des référentiels identité existants, des bases de gestion des élèves et des personnels et du référentiel des établissements au moyen d'un outil d'ETL spécifiquement développé pour le projet : Alambic ;
- un serveur et protocole SSO éprouvé dans le cadre du déploiement des ENT universitaires : CAS.

L'environnement numérique de confiance alors mis en œuvre pour PMB (environ 340 instances applicatives pour les besoins des collèges et lycées) permet de déployer d'autres services numériques à valeur ajoutée pour les usagers des établissements scolaires. La compatibilité avec le schéma d'annuaire ENT préconisé dans le SDET et avec le serveur CAS permet de facilement intégrer de nouveaux services dans l'environnement numérique de confiance Toutatice : Educhorus, Pronote, Moodle, SPIP, Nuxeo, services numériques de collectivités…

Le périmètre applicatif s'étend, l'environnement numérique de Toutatice prend corps progressivement jusqu'à juin 2011. À cette date, une trentaine de services numériques distincts sont « CASsifiés ». Certaines applications exploitent le mécanisme de ProxyCas afin d'accéder aux webservices, d'autres en authentifiant l'usager connecté par le SSO CAS.

L'annuaire de Toutatice devient le référentiel des comptes, groupes, applications et habilitations, tandis que le serveur CAS devient le dispositif permettant d'authentifier les utilisateurs et de fournir les données d'identité nécessaires aux applications intégrées. Le mécanisme d'enrichissement du pivot d'identité CAS, le service « Person Directory », est utilisé pour transmettre des données supplémentaires aux applications. Il permet de faciliter l'interconnexion avec des services hébergés en mode SaaS, ne pouvant accéder à l'annuaire de Toutatice.

Dans le même temps, le ministère de l'agriculture met en œuvre son fournisseur d'identité Educ@gri tandis que les services numériques de l'éducation développés par le ministère se structurent autour de deux environnements :

- ARENA : Accès aux Ressources de l'Éducation Nationale et Académiques. Les services sont destinés aux agents de l'éducation nationale travaillant en académie ;
- ATEN : Accès aux Téléservices de l'Éducation Nationale. Les services numériques sont destinés aux élèves et parents d'élèves pour dématérialiser l'inscription des enfants, permettre le suivi du

B2I et permettre le suivi de la scolarité : accès aux outils de gestion de notes et d'absence édité par l'éducation nationale.

La coexistence de ces différents environnements numériques pour des mêmes usagers complexifie l'offre de service académique. Les usagers se voient attribuer au moins deux identifiants différents pour accéder aux applications mises à disposition par l'académie.

1.2 Fédération de Toutatice : CASShib ou comment shibbolethiser mon CAS (2011 à 2017)

En septembre 2011, le contexte change : la région Bretagne et l'académie signent une convention dans laquelle ils s'engagent à ouvrir l'ENT à des acteurs extérieurs. L'ENT Toutatice et toutes ses applications doivent alors passer derrière un fournisseur de service afin d'intégrer une fédération d'identité.

Le remaniement une par une des applications pour être compatibles SAML2 dépend de la disponibilité d'opérateurs internes et externes et s'annonce compliqué. Le choix est alors fait de mettre en œuvre un pont SAML2 vers CAS pour simplifier cette transformation. La solution CASShib permettant de « Shibbolethiser » un serveur CAS est mise en œuvre.

Présentation de la solution CASShib aux JRES 2015 :

https://conf-ng.jres.org/2015/document_revision_2097.html?download

1.3 FranceConnect, EduConnect et IPANEMA (Depuis 2014)

L'écosystème change rapidement et les besoins utilisateurs tendent à la simplification de la gestion de leurs identités numériques.

FranceConnect est lancé depuis 2014.

L'éducation nationale investit également dans son fournisseur d'identité national avec le projet EduConnect.

L'académie de Rennes et le pôle national identité d'Orléans-Tours (IH2M) s'associent alors en 2015 afin d'imaginer et de défendre un projet PIA « Programme d'Investissement d'Avenir » dans le domaine « Identité numérique et relation usagers » : le projet IPANEMA.

Retenu par la DINSIC et débuté en 2017, IPANEMA a pour objectif principal l'expérimentation de solutions permettant l'amélioration de la gestion des identités numériques dans le périmètre de l'académie de Rennes afin de pouvoir mutualiser les résultats et retours d'expérience au niveau national de l'éducation nationale. Il est réalisé en collaboration entre une équipe dédiée au sein de l'académie de Rennes et le pôle national identité (IH2M).

Depuis le lancement de ce projet, de nombreux travaux ont permis l'évolution de la fédération d'identité de l'académie de Rennes jusqu'à la mise en œuvre d'un véritable Hub d'identité territorial.

2 Pourquoi faire évoluer la fédération d'identité Toutatice ?

2.1 Enjeux

En 2017, la DSI de l'académie de Rennes administre donc sa propre fédération d'identité permettant de fédérer plusieurs fournisseurs d'identité nationaux et académiques, et permettre l'accès à des services numériques en CAS et en SAML2 avec enrichissement et personnalisation des vecteurs d'identité à partir de l'annuaire ENT Toutatice.

Mais le monde de la fédération d'identité continue à évoluer avec de nouveaux protocoles devenant incontournables et la fin de la maintenance de la brique centrale de cette fédération : CASShib.

La volonté de moderniser cette architecture vieillissante, mais aussi de mettre en œuvre les moyens d'aujourd'hui pour authentifier les accès aux applications, a permis de constituer les objectifs précis suivants :

- Avoir une fédération d'identité compatible avec les protocoles standards actuels : CAS, SAML2 et OpenId Connect ;
- Étudier les différentes solutions du marché : serveur CAS, IdP/SP Shibboleth, etc. ;
- Choisir une solution performante dotée d'une communauté assurant son maintient opérationnel de façon pérenne ;
- Avoir une architecture hautement disponible ;
- Avoir une architecture supervisée ;
- Pouvoir inter-opérer l'ensemble des fournisseurs d'identité académiques et nationaux.

2.2 Le projet IPANEMA

En tant que projet PIA, le projet IPANEMA est accompagné d'un certain nombre d'objectifs auxquels le projet de Hub d'identité territorial a permis de répondre.

L'objectif principal est la mutualisation des résultats des travaux réalisés dans le cadre d'IPANEMA.

Pour cela, les objectifs suivants s'ajoutent aux précédents :

- Mettre à disposition des éléments réutilisables : scripts Ansible, documentations, etc. ;
- Accompagner d'autres académies dans la mise en œuvre d'une architecture similaire.

Aujourd'hui, ces objectifs sont atteints :

- Le principe général du Hub d'identité territorial été reversé au niveau national, notamment au pôle national identité d'Orléans-Tours, le pôle IH2M, afin de permettre l'accrochage d'applications CASsifiées derrière le Hub d'identité national ;
- Il a également été reversé, « à la demande », comme par exemple à l'académie de Bordeaux, pour leur permettre de s'accrocher au projet national GAR (Gestionnaire d'Accès aux Ressources) en parallèle de leur fédération d'identité en place ;
- Un accompagnement personnalisé a été assuré par l'équipe IPANEMA à chaque sollicitation pour des échanges et partages d'expérience sur ce domaine.

3 Présentation du Hub d'identité territorial breton

3.1 Choix de la solution : Shibboleth

En 2017, les premiers tests ont concerné la brique CASV5 qui promettait la propagation des identités aussi bien en CAS qu'en SAML2. L'expérimentation fut rapidement limitée par des problèmes d'implémentation de la partie SAML2.

Le choix s'est alors porté sur l'IdP Shibboleth V3 que notre équipe connaissait mieux et pour lequel l'implémentation a semblé plus simple à prendre en main et à paramétrer pour permettre la propagation des identités aussi bien en SAML2 (natif) qu'en CAS.

Le choix a donc été fait d'appuyer le Hub d'identité territorial sur la solution Shibboleth V3.

3.2 Architecture

3.2.1 Principe général : les fédérer tous !

Le Hub d'identité mis en œuvre dans l'environnement de l'académie de Rennes permet de fournir l'identité des utilisateurs venant de plusieurs fournisseurs d'identité, à une multitude de services numériques.

Pour cela, il est centralisé sur un « Identity Provider Toutatice » (IdP Toutatice) fédéré avec l'ensemble des services numériques. Cet « IdP Toutatice » délègue son authentification à un « Service Provider Toutatice » (SP Toutatice) fédéré avec l'ensemble des fournisseurs d'identité régionaux, académiques et nationaux.

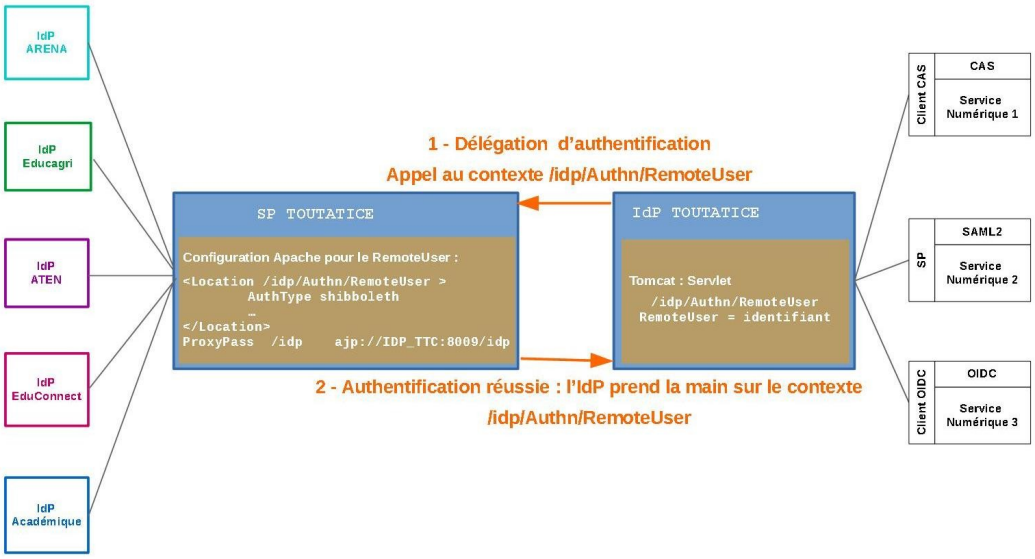

Figure 1 - Architecture en rebond du Hub d'identité territorial

Pour permettre cette délégation entre l'IdP Toutatice et le SP Toutatice, il faut configurer ce « couple IdP-SP » en RemoteUser tel que présenté par la Figure 1.

Le principe de cette configuration est que toute demande d'authentification d'un service numérique auprès de l'IdP Toutatice déclenche un appel au contexte

/idp/Authn/RemoteUser. Ce contexte est configuré sous Apache comme étant protégé par une authentification shibboleth SP. C'est donc le SP Toutatice qui va réaliser l'authentification de l'utilisateur auprès d'un des IdP auxquels il est lui-même associé (EduConnect, ATEN, ARENA,…).

Une fois l'authentification réalisée, le SP Toutatice va positionner sur le contexte, l'attribut RemoteUser avec une valeur correspondant à l'identité de l'utilisateur authentifié, puis laisser à l'IdP Toutatice la main sur le contexte grâce au ProxyPass.

D'un point de vue utilisateur, les échanges entre les différentes briques techniques vous sont présentés dans le diagramme de séquence suivant :

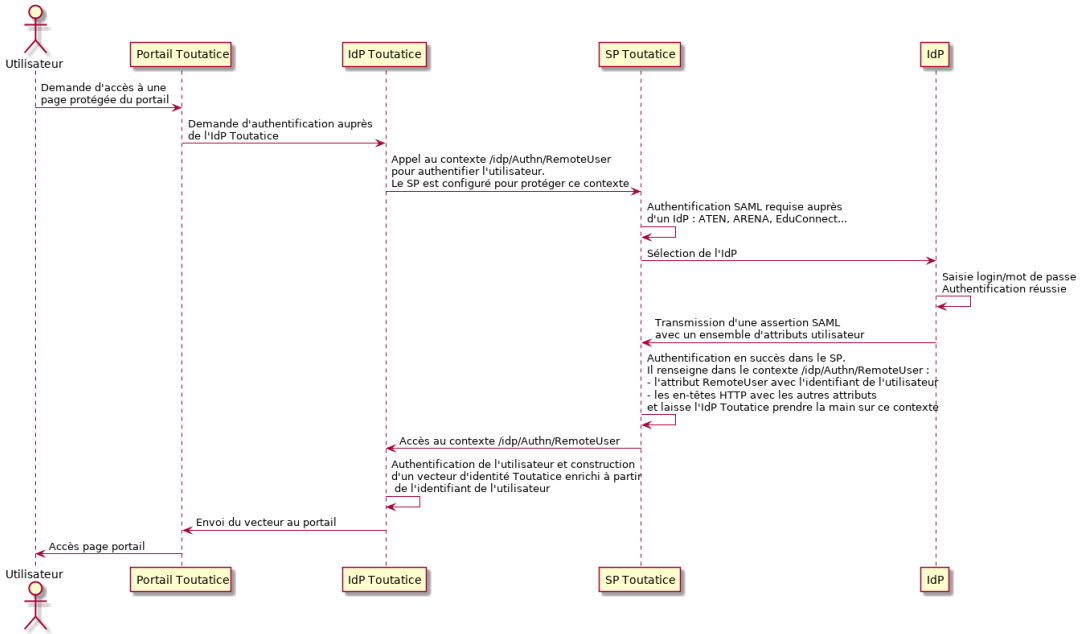

Figure 2 - Diagramme de séquence d'un utilisateur de l'éducation national accédant au portail Toutatice à partir de son identité (ATEN, ARENA, EduConnect...)

Cette architecture permet donc au Hub d'identité territorial de mettre en avant 3 capacités intéressantes :

- Être compatible avec tous les fournisseurs d'identité institutionnels au format SAML2 ;
- Être fédéré à un seul fournisseur d'identité d'un point de vue applications ;
- Permettre l'enrichissement des identités avec des données du référentiel de l'ENT (LDAP Toutatice).

3.2.2 Haute disponibilité

Afin d'être en capacité de répondre aux importantes sollicitations des utilisateurs et d'assurer un service hautement disponible, le choix s'est porté sur une architecture redondée en mode cluster composée actuellement de deux membres.

La répartition de charge sur les différents membres est gérée par paramétrage d'un F5 BIG-IP.

Chaque membre est composé :

– D'un Service Provider (SP) Shibboleth ;

– D'un Identity Provider (IdP) Shibboleth V3.3.2 couplé au SP ;

– D'une instance Memcached permettant le stockage des sessions SAML2 et CAS : Les instances Memcached sont configurées pour permettre une lecture croisée des sessions entre elles.

Chaque membre interagit avec :

– Un cluster MariaDB pour le stockage des sessions OpenId Connect ;

– Un cluster d'annuaire LDAP (ODSEE 11g) en réplication Maître / Maître.

Tous deux également répartis par le F5 BIG-IP.

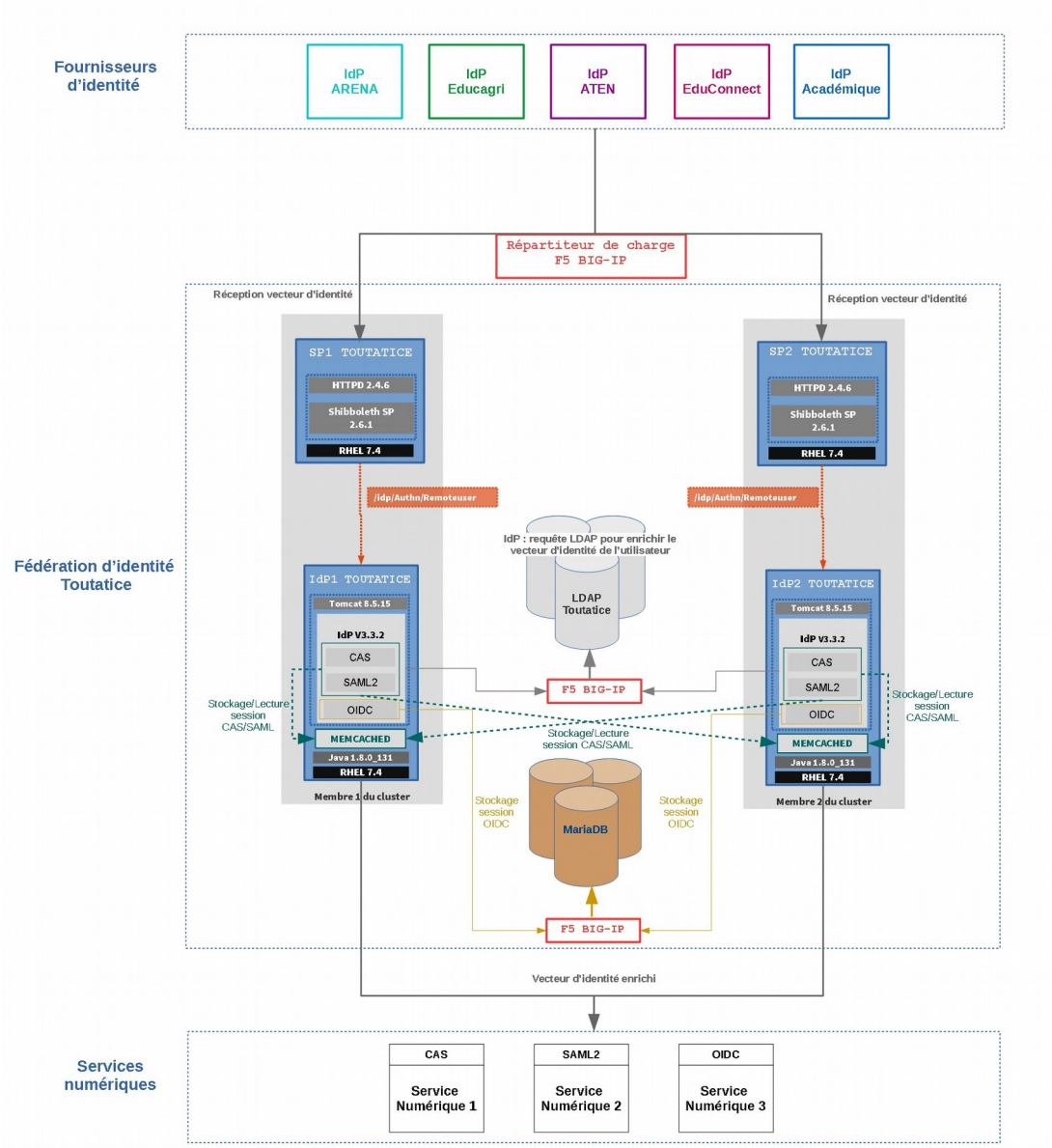

Figure 3 - Architecture de haute disponibilité du Hub d'identité territorial

3.2.3 Multi-protocolaires

Le choix de l'IdP Shibboleth V3 a permis d'utiliser nativement les protocoles CAS et SAML2.

Afin de pouvoir ajouter le protocole OpenId Connect, protocole devenu incontournable notamment pour tous les usages autour de la mobilité, le choix s'est porté sur l'extension de l'IdP Toutatice avec le plugin de l'université de Chicago s'appuyant sur MITREid Connect (https://github.com/uchicago/shibboleth-oidc).

L'ensemble permet donc d'obtenir un Hub pouvant recevoir des identités en SAML2, de les enrichir au besoin à partir du LDAP Toutatice, et de les distribuer dans le protocole CAS, SAML2 ou OpenId Connect.

3.2.4 Supervision

En plus de l'utilisation de la supervision de l'académie de Rennes avec l'outil Check_Mk qui permet le suivi du maintien en conditions opérationnelles des différentes briques techniques du Hub identité territorial, la supervision du Hub d'identité territorial est complété par l'utilisation du projet académique de centralisation des logs de la DSI basée sur la suite Elastic.

Les produits Shibboleth SP et IdP permettent de produire des logs dans un format facilement exploitable par la suite Elastic.

Format de log configuré sur le SP Toutatice :

```
<date>|<IdP_source>|<attributs_vecteur>|<identifiant>|<status>
```

Exemple de ligne de log :

```
2019-09-26 15:24:42|idp-arena|ctdn(1),uid(1)|martin.dupont|
Success
```

Format de log configuré sur l'IdP Toutatice :

```
<date>|<identifiant>|<adresse_ip>|<url_service_numerique>|
<user_agent>|<service_ticket_cas_ou_assertion_id_saml>
```

Exemple de ligne de log :

```
2019-09-26
15:24:43|martin.dupont|127.0.0.1|https://www.toutatice.fr/portail
/MonEspace|Mozilla/5.0 (Windows NT 10.0; Win64; x64; rv:69.0)
Gecko/20100101 Firefox/69.0|ST-1569503129575
```

L'exploitation de ces logs permet ainsi que mettre en place des tableaux de bord complets pour le suivi des différents briques techniques, mais également des parcours utilisateurs.

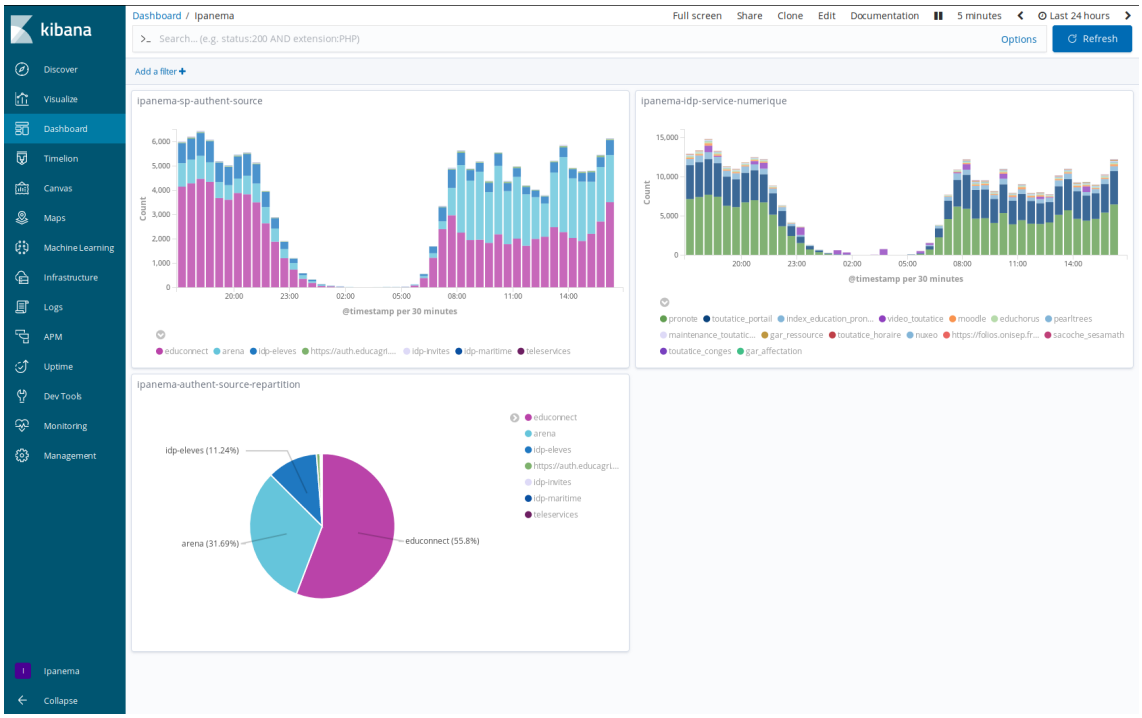

Figure 4 - Exemple d'écran de supervision obtenu par la centralisation des logs du Hub d'identité territorial

3.3 Accrochage d'EduConnect

Avec l'arrivée d'EduConnect, fournisseur d'identité national de l'Éducation Nationale pour les élèves et leurs représentants légaux, un réel besoin d'adaptations spécifiques de notre Hub d'identité territorial apparaît.

Le choix d'une solution Open Source telle que Shibboleth permet le développement de plugins additionnels afin de répondre à des besoins spécifiques. Ces plugins peuvent répondre à des besoins d'enrichissement de données, à des traitements particuliers, voire à la gestion de nouveaux protocoles.

Pour EduConnect, plusieurs développements ont été réalisés, notamment un plugin spécifique associé à notre IdP Toutatice.

3.3.1 Contexte

Dans le cadre d'IPANEMA, l'académie de Rennes a assuré le rôle de démonstrateur d'EduConnect. Nos travaux ont donc été menés en étroite collaboration avec l'équipe projet EduConnect d'Orléans-Tours. Ils ont permis de démontrer la faisabilité d'une bascule partielle en cours d'année scolaire et d'une généralisation en début d'année pour toutes les populations visées.

Pour ces travaux, plusieurs adaptations spécifiques ont dues être mises en place sur le Hub d'identité territorial. Ces adaptations ne sont donc pas nécessaires au bon fonctionnement nominal du Hub en tant que passerelle d'identité multi-protocolaires et hautement disponible, mais peuvent être utile en cas d'accrochage du fournisseur d'identité EduConnect à une fédération d'identité.

3.3.2 Cas d'usage

Le changement d'identité numérique d'une population, d'autant plus lorsqu'il s'agit d'enfants mineurs ou de représentants légaux non acculturés aux environnements numériques complexes, n'est pas quelque chose d'anodin ou de facile à accompagner. Un objectif clairement affiché dans nos travaux était de faire porter le moins possible ce changement important aux personnels administratifs des établissements. Pour cela, deux cinématiques ont été mises en place :

- Pour les représentants légaux : création de leur nouvelle identité EduConnect en autonomie à partir de leur numéro de téléphone portable ;
- Pour les élèves, création de leur nouvelle identité EduConnect en autonomie à partir de leur ancienne identité académique en passant par une phase nommée « phase de transition ».

Une autre volonté était de formater un vecteur d'identité compatible avec les exigences du RGPD et ne portant donc plus de données sensibles et facilement identifiable.

De ces choix découlent plusieurs besoins devant être traités au niveau du Hub identité :

- Pouvoir définir quels utilisateurs doivent être « poussés » vers la phase de transition EduConnect quand ils arrivent authentifiés avec leur identité académique ;
- Pouvoir créer/mettre à jour « à la volée » les comptes applicatifs des utilisateurs lors qu'ils arrivent avec une identité qu'ils ont créée en autonomie ;
- Pouvoir assurer un accueil personnalisé dans les services cibles à partir d'un vecteur d'identité ne présentant aucune distinction de profil.

3.3.3 Développements spécifiques

Pour répondre à ces différents besoins, plusieurs développements et adaptation de configuration ont du être réalisés :

- Développement d'un plugin « EduConnect » intégré à l'IdP Toutatice afin de gérer les utilisateurs devant être « poussés » vers la phase de transition ;
- Développement d'un webservice spécifique à la création / mise à jour des comptes applicatifs à partir des éléments présents dans le vecteur d'identité EduConnect reçu ;

– Configuration de la servlet IpanemaRemoveUser, point d'entrée de l'authentification dans l'IdP Toutatice pour permettre le traitement des différents cas d'usage en fonction des vecteurs d'identité reçus :
  – Poursuite de l'authentification « classique » ;
  – Création de comptes / Mise à jour « à la volée » dans le LDAP Toutatice avant poursuite de l'authentification ;
  – Affichages spécifiques (Pages d'information / erreur, Loader, …) ;
  – …

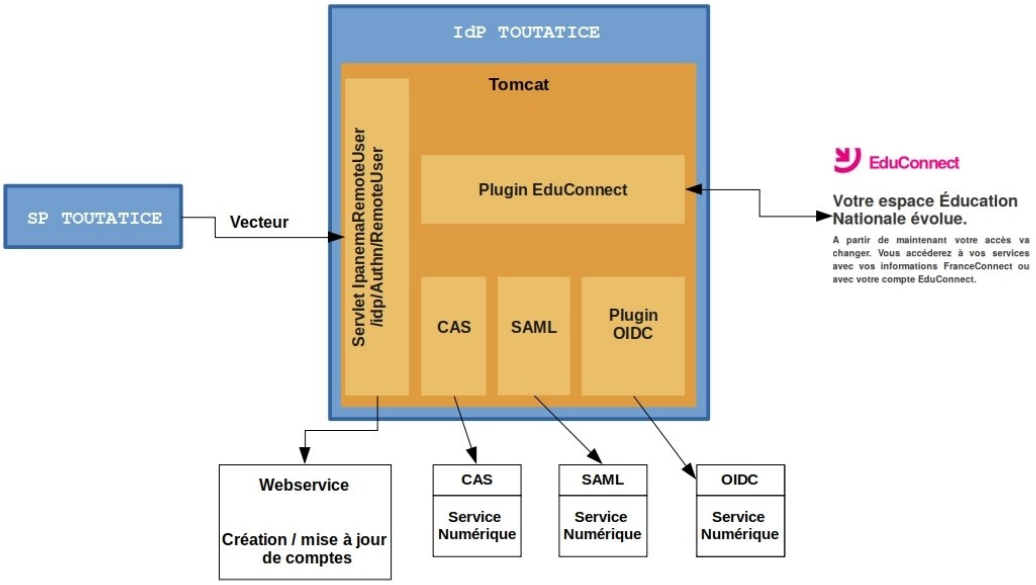

Figure 5 - IdP Toutatice avec les développements spécifiques et le plugin de l'université de Chicago

3.4 Le Hub en production : quelques chiffres

L'académie de Rennes a la responsabilité de la gestion des identités du 1er et du 2nd degré en Bretagne, soit un potentiel de 1 522 000 utilisateurs (Personnels, élèves, représentants légaux, …).

Notre Hub identité territorial absorbe chaque jour plus de 500 000 demandes de sessions avec des pics pouvant aller jusqu'à 100 000 demandes de sessions / heure.

Chaque semaine, c'est environ 2 600 000 demandes de sessions à traiter.

Ci-dessous quelques graphiques générés grâce à la centralisation des logs des SP / IdP :

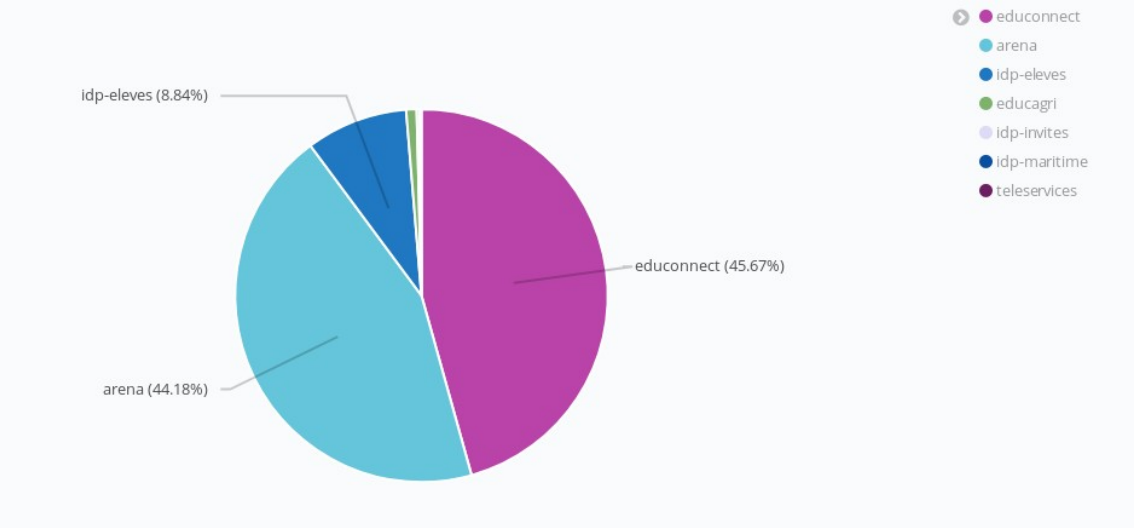

Figure 6 - Exemple de répartition des authentifications sur le Hub par IdP

Ce graphique datant de septembre 2019, quelques jours après la généralisation d'EduConnect pour les élèves et les représentants légaux de l'académie de Rennes, montre la répartition des connexions à travers les différents IdP « sources » :

- IdP EduConnect pour les élèves et représentants légaux utilisant leur identité EduConnect ;
- IdP Arena pour les personnels de l'Éducation Nationale ;
- IdP eleves (IdP académique de l'environnement Toutatice) pour les élèves utilisant la « phase de transition » pour activer leur identité EduConnect ;
- …

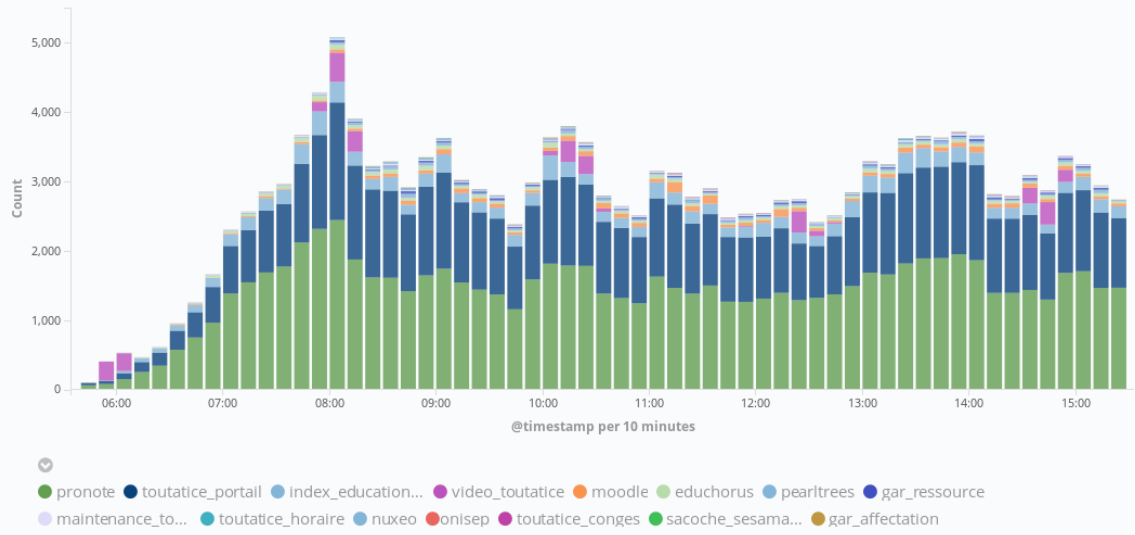

Figure 7 - Exemple du nombre de sessions demandé au Hub d'identité en fonction du service numérique cible

Ce graphique datant de septembre 2019 présente la répartition des demandes de sessions par service numérique. Il montre clairement la cible principale des accès durant la phase de rentrée scolaire :

- L'accès aux applications « vie scolaire » (avec l'application pronote majoritairement déployée dans l'académie) ;
- L'accès au portail Toutatice permettant la consultation de nombreuses informations et l'accès aux applications numériques proposées par le ministère, l'académie, l'établissement, les collectivités…

3.5 Sources et documentations

3.5.1 Sources et documentation Shibboleth

Le Hub identité s'appuie sur des briques open sources Shibboleth disponibles sur le site http://shibboleth.net/downloads/

La documentation de référence pour leur installation est disponible sur le site https://wiki.shibboleth.net/

3.5.2 Architecture du Hub identité

Afin d'automatiser le déploiement du « cœur » de l'architecture du Hub identité territorial, un playbook Ansible a été mis en place. Ce playbook permet le déploiement d'un couple SP Shibboleth / IdP Shibboleth paramétré en mode remoteUser et intégrant la partie multi-protocolaires au niveau de l'IdP (intégration du plugin de l'université de Chicago).

3.5.3 Documentation du projet

L'ensemble des travaux du projet IPANEMA, dont le Hub identité territorial, a été documenté au fil de l'eau sur le wiki du projet protégé par la Fédération Education – Recherche : https://ipanema.education.fr/xwiki/

4 Et après ?

Comme tout environnement technique, il faut maintenant assurer le maintien en conditions opérationnelles et conditions de sécurité de ce Hub d'identité territorial breton.

Pour cela, il faut continuer à investir dans sa supervision, mais également planifier annuellement les mises à jour nécessaires sur les différentes briques techniques.

Il va également falloir rester attentif aux prochaines versions de Shibboleth afin de pouvoir simplifier son architecture avec les évolutions proposées, voire enrichir ses fonctionnalités et les protocoles proposés.

À plus court terme, les travaux d'amélioration suivants vont être poursuivis :

- Rendre opérationnel le consentement des utilisateurs pour l'ensemble des protocoles avec adaptation des pages « standards » proposées ;
- Persistance des choix des utilisateurs en base de données.

Et pourquoi pas une présentation aux JRES 2021 ?

Remerciements

Nous souhaitons remercier particulièrement notre ancien collègue, Fabien Berteau, pour son implication et sa participation importante aux premières versions de notre fédération. Notre HUB n'aurait pas pu atteindre cette maturité sans ses compétences hors normes et son inventivité inépuisable.

Un grand merci également à l'équipe nationale identité de la Direction du Numérique pour l'Éducation, le pôle IH2M à Orléans-Tours, pour les nombreux échanges et retours d'expérience nous ayant régulièrement aidés dans notre feuille de route.