evidian

Synchronisation des identités pour un référentiel d'identités multi-annuaires

Bull
atos technologies

Introduction

La construction d'un référentiel d'identité est au cœur des approches de gestion des identités et des accès.

En effet, quelle que soit la qualité de la politique de sécurité des accès, quelle que soit l'efficacité de sa mise en œuvre et quelle que soit la finesse des outils de reporting, si la donnée définissant l'utilisateur n'est pas fiable, c'est tout l'édifice qui s'écroule.

Au cœur de la gestion des identités et des accès

La construction d'un référentiel d'identité est au cœur des approches de gestion des identités et des accès. En effet, quelle que soit la qualité de la politique de sécurité des accès, quelle que soit l'efficacité de sa mise en œuvre et quelle que soit la finesse des outils de reporting, si la donnée définissant l'utilisateur n'est pas fiable, c'est tout l'édifice qui s'écroule.

La mise en place d'un tel référentiel peut se heurter à bien des obstacles, comme par exemple la définition d'un modèle trop complexe, une approche purement technique ou encore le mélange entre les informations qui définissent l'identité de l'utilisateur et celles qui définissent ses droits. Avant de se lancer dans un tel projet, il est important d'utiliser une approche qui permettra lors de sa mise en œuvre de définir des modèles et des processus simples, robustes et efficaces.

Un référentiel cohérent est indispensable. Pourquoi ?

Evidian recommande la mise en place d'un modèle qui permet de bien séparer les informations d'identité, les informations définissant les accès ainsi que les règles et procédures permettant de les construire.

Figure 1. Des données autoritatives aux données attribuées

Les informations définissant l'identité sont dites « données autoritatives » alors que les informations définissant les droits d'accès font partie des « données attribuées ».

	Définition	Exemples
Autoritatives	Données qui définissent l'identité, au sens large, de l'utilisateur.	Nom, prénom, identifiant entreprise, organisation, métier, position hiérarchique, site, bureau, n° de téléphone,...
Attribuées	Données qui définissent les droits d'accès de l'utilisateur sur les systèmes et applications cibles.	Identifiants et mot de passe applicatifs, privilèges, plage horaire d'utilisation, appartenance à un groupe LDAP, ...

Tableau 1. Données autoritatives et attribuées

Les données attribuées d'un utilisateur sont en général obtenues en appliquant des règles d'une politique de sécurité aux données autoritatives du même utilisateur. Par exemple, un ingénieur commercial (données autoritatives) aura un identifiant et un mot de passe (données attribuées) pour accéder à l'application dédiée aux vendeurs (Politique de sécurité) « vente ».

Situées en début de processus, les données autoritatives sont donc bien au cœur des procédures de gestion des identités et des accès. Toute création, modification ou suppression d'une donnée autoritative peut avoir un impact immédiat sur une donnée attribuée.

La constitution d'un référentiel d'identité

Pour une organisation, il y a plusieurs manières de mettre en place un référentiel d'identité.

Par exemple, le système de gestion des données d'identités déjà en place peut être suffisamment simple pour qu'un workflow de gestion des identités puisse alimenter un annuaire unique partagé par toutes les organisations (Plus d'informations voir notre site www.evidian.com).

Mais, dans la plupart des cas, les organisations ont déployé de multiples annuaires collectant des données d'identité aux formats souvent incompatibles et difficiles à partager. De plus, des groupes d'administrateurs dédiés sont en charge de leur administration et appliquent leurs propres modèles de données. Par exemple, au sein des systèmes de ressources humaines, les administrateurs gèrent les noms des utilisateurs, leurs métiers et les attributs liés à leurs organisations ; dans la base de données du PABX, les administrateurs ont en charge la gestion des numéros de téléphone ; dans les annuaires LDAP, les administrateurs doivent gérer les adresses mail des utilisateurs, etc..

Cette situation incohérente génère d'importants coûts de non qualité :

- Une perte de productivité des équipes informatiques qui doivent gérer la même donnée dans plusieurs endroits et sous plusieurs formats.
- Des identités non contrôlées et exposées à des attaques.
- La prolifération de données d'identités non fiables qui rend complexe le contrôle et la gestion des habilitations.

Pour résoudre cette problématique, Evidian a introduit une fonction de synchronisation des identités (ID Synchronization) afin de construire un référentiel unique des identités.

La synchronisation des identités

La synchronisation des identités permet de créer un **référentiel d'identité** fiable et cohérent.

Le module ID Synchronization qui offre cette fonction interagit avec les autres fonctions de la solution de gestion des identités et des accès **d'Evidian Identity Governance and Administration :**

Le gestionnaire de politique s'appuie sur le référentiel ainsi créé pour initialiser les opérations de création des droits d'accès en fonction de la politique de sécurité.

Le workflow complète les mécanismes de consolidation en mettant en œuvre des procédures de validations pour les applications les plus critiques.

Le provisionnement utilise les données d'identité pour créer, modifier ou supprimer les droits d'accès dans les systèmes et applications cibles.

L'authentification unique (single sign-on - SSO) d'entreprise s'appuie sur les données pour valider l'identité d'un utilisateur.

Figure 2. La synchronisation des identités

Les principales fonctions de la synchronisation des identités

La synchronisation des identités met en œuvre les principes de synchronisation auxquels s'adjoignent les fonctions d'interface des modules de gestion des identités et des accès.

Les fonctions de synchronisation

Il s'agit de la définition et de l'application d'une politique de synchronisation multi-référentiels.

L'activation des fonctions en aval

Il s'agit de l'activation des opérations de mise à jour des droits via les mécanismes de règles « métier » au sein du gestionnaire de politique. Ces opérations peuvent mener à l'activation des processus de provisionnement sur les applications cibles du système d'information.

La journalisation

Il s'agit de la journalisation des opérations de gestion des identités à des fins d'analyse et de reporting.

La politique de synchronisation

De manière générale, le périmètre de la politique de synchronisation doit être restreint aux données autoritatives de définition de l'utilisateur.

Cette politique va permettre d'appliquer des règles :

- de **consolidation** d'une identité à partir d'enregistrements multiples, les données relatives à une identité pouvant être dans différents référentiels,
- de **réconciliation** en cas d'incohérence des données,
- de **création ou suppression** des informations liées à un utilisateur,
- de **gestion des seuils** qui empêche l'altération du référentiel d'identité dans le cas où les sources de données sont corrompues.

La synchronisation peut se faire soit en batch soit de manière continue ; elle ne s'appuie de toutes les manières que sur les dernières modifications.

Si le nombre d'opération à effectuer dépasse le seuil, le traitement n'est pas réalisé et un administrateur est aussitôt notifié. Il peut alors rejouer le flux en mode simulation pour analyser l'origine de l'anomalie. Les sources corrompues pourront alors être corrigées et le traitement relancé.

Afin de pérenniser la maintenabilité de la solution mise en place, il est très important de bien séparer les données autoritatives d'identités des données attribuées et de n'appliquer la synchronisation qu'aux données autoritatives. Par exemple, les mécanismes de réconciliation, bien qu'apparemment similaires pour les 2 familles de données, n'obéissent pas aux mêmes règles :

- Les règles de réconciliation entre données autoritatives relèvent d'un arbitrage entre le niveau de confiance ou de fiabilité accordé à chaque donnée. Si le numéro de téléphone d'un utilisateur est différent dans deux annuaires, il suffit de définir une règle permettant de décider quelle est la donnée de référence.
- Les règles de réconciliation pour les données attribuées (données en général provisionnées sur les systèmes cibles) comparent la réalité d'une donnée sur un système cible et la valeur qu'elle devrait avoir en fonction de la politique de sécurité des accès. Elles doivent donc s'interfacer avec le moteur de politique. Le résultat d'une réconciliation en faveur de la donnée réelle peut être la modification de la politique de sécurité des accès.

Les principales fonctions de la synchronisation des identités

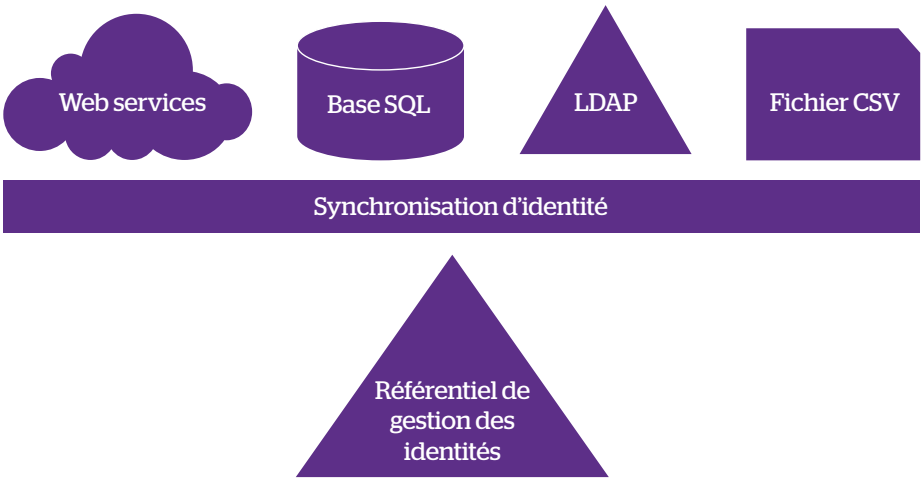


Figure 3. Les référentiels à prendre en compte

Les référentiels

La synchronisation des identités travaille sur différents types de sources de données.

Les technologies les plus courantes sont les annuaires LDAP, les bases de données relationnelles, les fichiers plats (csv, ldif,...) ou encore au travers de Web services.

De manière complémentaire, il est possible de trouver des interfaces vers les bases de données des applications de Ressources Humaines. Ces bases RH peuvent en effet contenir des informations nécessaires à la définition des utilisateurs ou encore peuvent initialiser des évènements de création ou de suppression de l'identité d'un utilisateur au sein du SI.

Si la solution globale de gestion des identités et des accès possède sa propre base d'identité, la synchronisation des identités devra naturellement intégrer cette base dans ses mécanismes.

Un cas particulier : les référentiels des applications et systèmes cibles

Dans certains cas, une donnée d'identité peut se trouver dans le référentiel interne de l'application elle-même. La synchronisation des identités peut alors utiliser les mécanismes techniques, comme les agents et connecteurs généralement utilisés par le provisionnement, pour intégrer les données cibles au sein du référentiel d'identité.

En effet, ces connecteurs et agents utilisent des interfaces (API) publiques et stables dans le temps fournies par le fournisseur de l'application. Plutôt que de mettre en place un accès direct au référentiel interne de l'application, l'utilisation de ces API publiques permet de garantir une plus grande stabilité à la solution mise en place.

Une organisation multi-niveaux

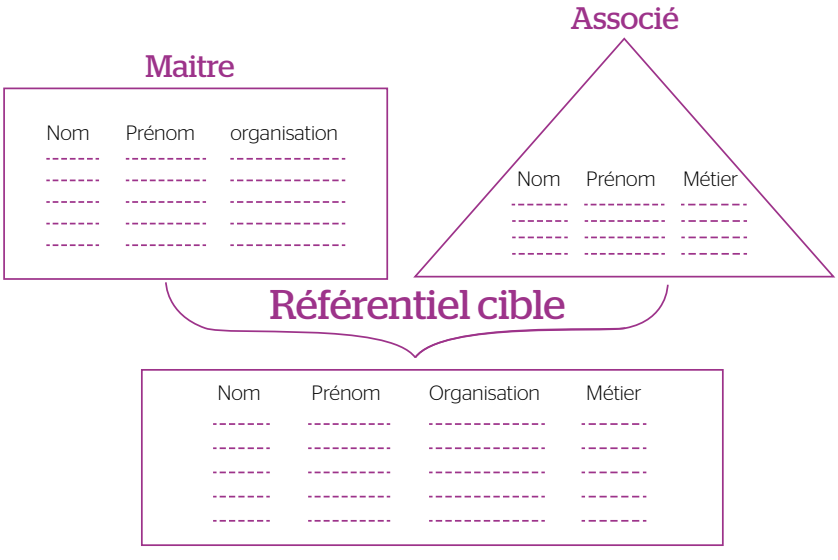


Figure 4. Les jointures

Afin de créer des ensembles cohérents de données sources, il est possible d'associer un ensemble de référentiels sources « Associées » à un référentiel source « Maître » (à travers des mécanismes de jointure). Une source « Maître » contient l'enregistrement auquel on pourra raccrocher les données des sources associées. La suppression d'un enregistrement « Maître » fera donc disparaître complètement l'enregistrement, par contre la suppression d'enregistrement « Associé » sera prise comme une absence de donnée à traiter, si nécessaire, via les règles de synchronisation.

Ce premier niveau d'agrégation permet d'organiser les données afin d'obtenir la même visibilité pour des données venant de sources différentes. Les règles de synchronisation peuvent alors être de haut niveau et s'appliquer à l'ensemble des données accessibles.

La mise à jour des référentiels cibles

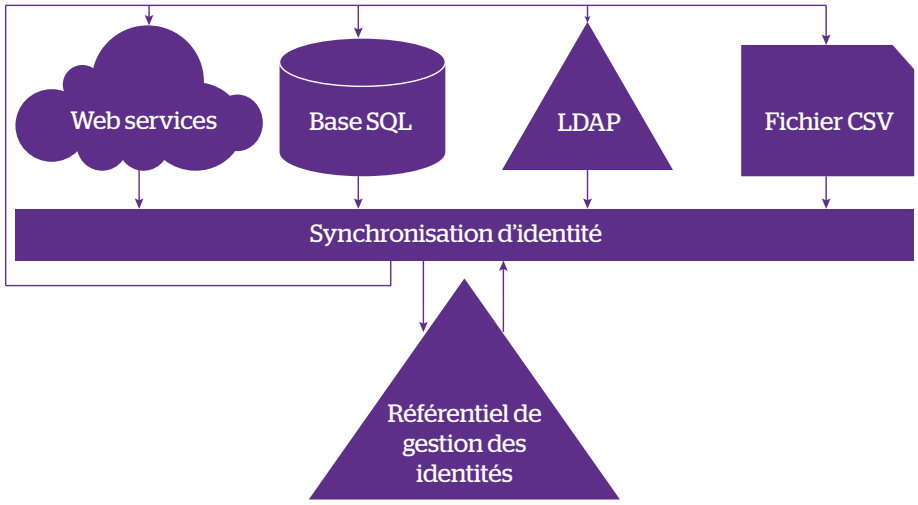


Figure 5. Les référentiels cibles

Les règles de synchronisation permettent de créer un référentiel unique centralisé, c'en est d'ailleurs l'objectif principal.

Elles peuvent aussi mettre à jour un groupe de référentiels cibles et ainsi créer une identité cohérente sur un ensemble distribué de référentiels.

Cette distribution de l'identité permet de répondre de manière simple à une problématique d'architecture d'annuaire, d'optimisation des flux réseaux ou encore d'homogénéisation de données au sein de référentiels différents et/ou incompatibles.

La typologie des règles

Assez paradoxalement, il est possible de mettre en place un mécanisme de synchronisation des identités en s'appuyant simplement sur 4 règles de base :

Join

Cette règle permet d'adosser un enregistrement d'une source « Associée » à une ou plusieurs sources « Maître ».

Attribute Mapping

Cette règle permet de définir une relation de correspondance entre attributs de différentes sources. Cette relation est ordonnée et mène à la mise à jour des attributs entre eux. Cette mise à jour peut également faire appel à l'application d'une fonction de transformation intermédiaire.

Enfin cette relation de correspondance permet de mettre à jour des attributs multi-valués à partir de multiples sources. Un exemple simple d'attribut multi-valué est l'attribut adresse email ; en effet un employé peut avoir plusieurs adresses emails au sein d'une organisation : prenom.nom@organisation.com, p.nom@organisation.com, pnom@organisation.com, prenom.com@country.organisation.com, Prénom Nom/Org/Pays.... Chaque adresse email est disponible dans un annuaire particulier. La consolidation au sein d'un seul référentiel des adresses emails peut se faire via un attribut multi-valué.

Creation

Lorsqu'un utilisateur existe dans un référentiel « Maître » mais pas dans un référentiel « Associé », cette règle permet de créer dans le référentiel « Associé » l'ensemble des attributs associés à cet utilisateur.

Deletion

Cette règle permet de supprimer automatiquement les attributs d'un utilisateur dans un référentiel « Associé » si cet utilisateur a été supprimé du référentiel « Maître ».

Les autres règles

Il est possible d'imaginer beaucoup d'autres règles pour la création d'attributs dans les référentiels. Ces autres règles sont en général relatives à la création des données attribuées. Elles s'intègrent donc naturellement dans le gestionnaire de politique qui, à partir des données d'identités et en application de la politique, permettra de créer et provisionner sur les applications et systèmes cibles les droits d'accès de l'utilisateur.

Exemples d'architectures techniques

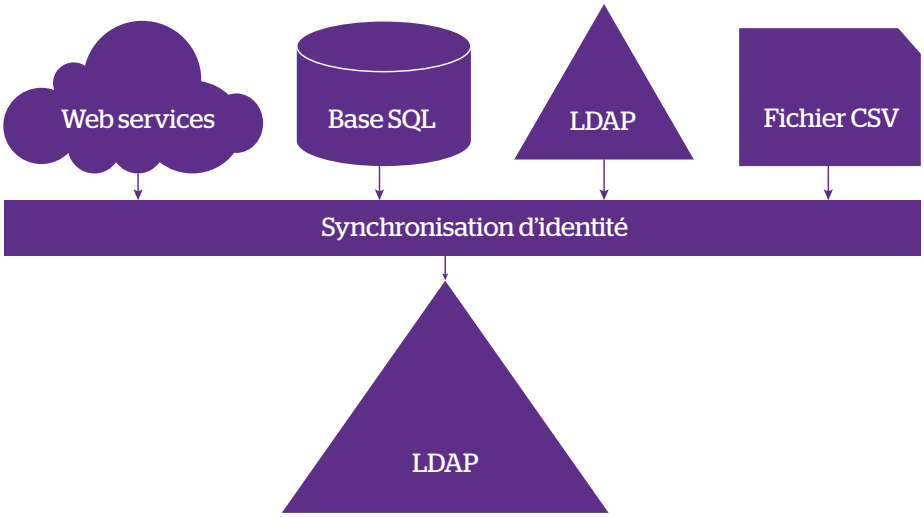


Figure 6. La création d'un annuaire LDAP de référence

Cas 1 : Création d'un annuaire LDAP de référence

Dans ce cas, l'annuaire LDAP des identités de référence est construit et maintenu de manière cohérente à partir de sources d'informations de différentes catégories (base de données SQL, fichiers CSV, annuaires LDAP, Web services...).

Ces sources d'informations sont en général sous la responsabilité d'un département RH, d'un responsable des partenaires ou font partie d'un ERP. Les données au sein de l'annuaire de référence résultent de l'agrégation des données sources consolidées lors d'opérations de jointure.

Il est ainsi possible de créer un référentiel qui permettra, par exemple, de définir un annuaire LDAP pour le SSO d'entreprise.

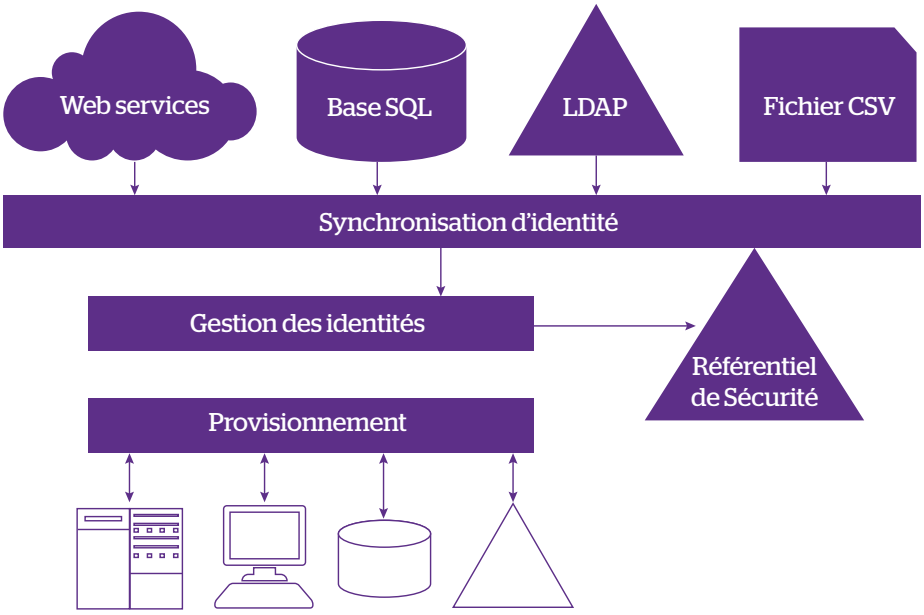


Figure 7. Utilisation de la base de la solution de gestion des identités et des accès comme référentiel d'identité unique.

Cas 2 : La création d'un référentiel d'identité en utilisant la base d'une solution de gestion des identités et des accès.

Dans ce cas, la synchronisation des identités est utilisée pour peupler et maintenir de manière cohérente la base d'identité d'un outil de gestion des identités et des accès.

Cette base est en général gérée par un module de gestion des identités et est utilisée par les modules de provisionnement.

Cette base peut-être appelée Référentiel de Sécurité. Elle peut contenir l'ensemble des données des utilisateurs ou seulement un sous-ensemble (cf. cas 3 suivant).

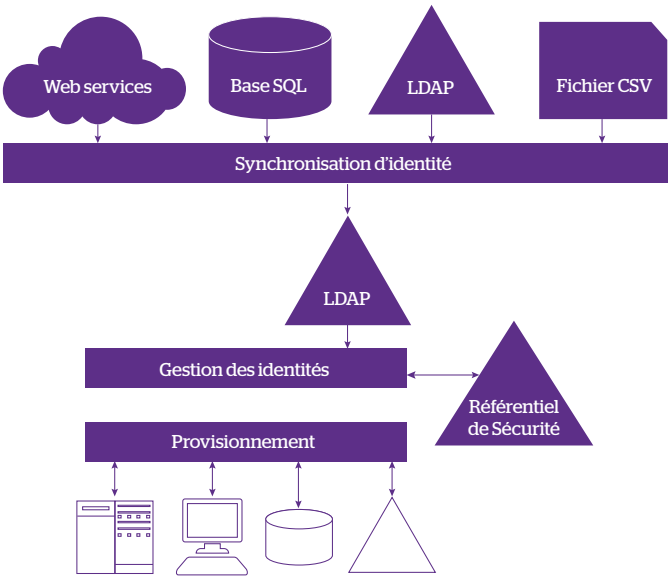


Figure 8. Utilisation de l'annuaire LDAP comme source d'identité unique

Cas 3 : Utilisation de l'annuaire LDAP comme source d'identité.

Ce cas est une extension du précédent. C'est le cas où la solution de gestion des identités et des accès utilise un annuaire LDAP externe comme source d'identité. Le Référentiel de Sécurité peut contenir alors, par exemple, les données définissant les droits d'accès, ou encore des données d'identités complémentaires.

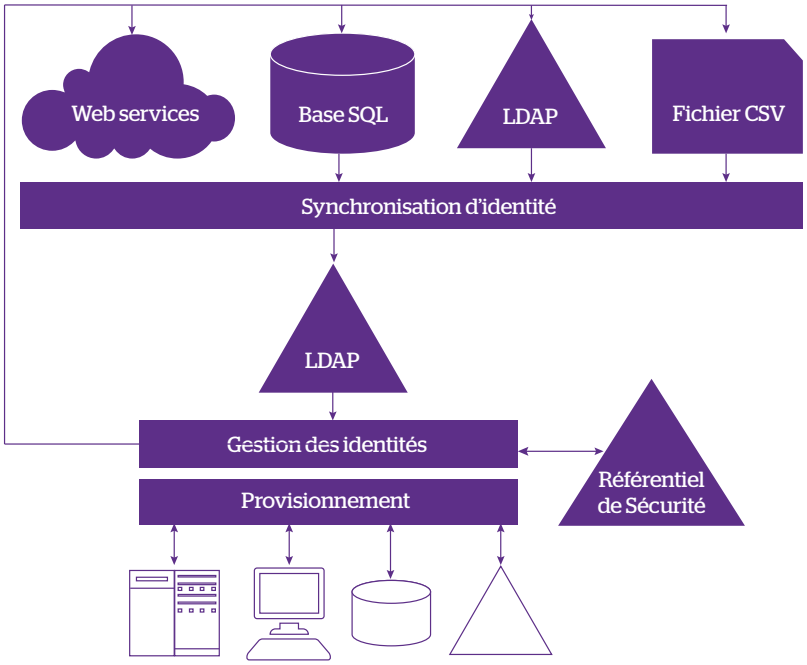


Figure 9. La synchronisation avec des données des applications cibles

Cas 4 : La synchronisation avec des données des applications cibles.

C'est le cas le plus complet pour lequel certaines données utilisateurs des référentiels internes des applications cibles peuvent être utilisées comme données sources autoritatives.

Il est alors possible d'utiliser des agents et connecteurs, pour synchroniser les données des applications cibles avec les différentes autres sources autoritatives.

Un exemple de mise en œuvre opérationnelle

L'exemple de cas d'utilisation ci-après est un cas réel mis en œuvre avec la solution IAM d'Evidian. Il utilise les possibilités de la synchronisation des identités ainsi que le couplage annuaire LDAP / Référentiel de Sécurité. En s'appuyant sur les caractéristiques techniques de la synchronisation des identités, il est possible de décrire le cas opérationnel suivant.

Les utilisateurs sont définis dans 4 bases sources :

- le référentiel Mainframe qui contient le prénom, le nom et l'identifiant d'un employé.
- le référentiel RH qui contient des données complémentaires comme l'adresse email, la localisation, le numéro de téléphone, le numéro de Fax.
- le fichier CSV des exceptions qui contient la liste des adresses email d'un utilisateur. Cette liste contient l'historique des adresses emails d'un utilisateur suite aux différentes acquisitions, consolidations et changements de domaine.
- Une base accessible en Web services contenant le numéro de mobile professionnel.

La base cible principale est un annuaire Active Directory qui contient la liste des groupes locaux Active Directory. Le référentiel Maître est le référentiel Mainframe. La présence, ou l'absence, d'un utilisateur dans ce référentiel décide de la présence ou de l'absence d'un utilisateur dans Active Directory.

Les règles de mise à jour des données s'appuient sur quatre types de flux

Création des utilisateurs

Les utilisateurs sont créés dans Active Directory avec tous leurs attributs. Le prénom, le nom et l'identifiant viennent du référentiel Mainframe, les données complémentaires telles que l'adresse email, la localisation ou le numéro de téléphone proviennent de la base RH. Le numéro de mobile professionnel vient d'une base accessible par Web services. Lorsqu'un utilisateur possède plusieurs adresses emails, le champ multi-valué « adresse email » dans Active Directory récupère l'ensemble des adresses.

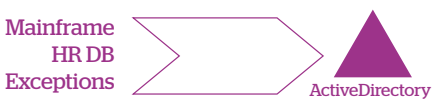


Figure 10. Le peuplement d'Active Directory avec les données des bases sources

Création des groupes globaux

Les groupes de l'organisation qui existent dans le référentiel du Mainframe sont créés dans Active Directory. Ce sont ces groupes qui permettent de définir un ensemble de ressources applicatives associées.

Création des groupes locaux Active Directory dans le référentiel du Mainframe

Pour pouvoir associer un groupe global à un groupe local Active Directory, l'administrateur du Mainframe a besoin de récupérer les groupes locaux Active Directory sur le Mainframe. Il faut donc recopier ces groupes d'Active Directory vers le Mainframe.

Figure 11. La synchronisation avec des données d'Active Directory

Création des relations d'appartenance

Les relations d'appartenance des utilisateurs aux groupes locaux Active Directory et aux groupes globaux ainsi que les relations d'appartenance des groupes globaux aux groupes locaux Active Directory sont définies dans le référentiel du Mainframe par les administrateurs. Ces relations sont recopiées dans Active Directory afin de permettre un contrôle des accès aux ressources de l'entreprise au niveau de Windows

Définition et application d'une politique de sécurité à base de groupes

Bien que la définition la plus complète des utilisateurs se trouve dans Active Directory, la gestion des groupes globaux est définie dans le mainframe : les relations d'appartenance sont définies par l'administrateur du Mainframe. Cette politique est ensuite « recopiée » dans l'annuaire Active Directory.

Le référentiel du Mainframe contient un sous-ensemble d'attributs permettant de réaliser cette gestion des groupes.

Les applications utilisant Active Directory

Certaines applications d'infrastructure peuvent utiliser nativement les données disponibles dans Active Directory. Ainsi l'annuaire Active Directory permet de gérer l'accès aux ressources partagées disques et imprimantes. De même, l'application Page Blanche s'appuie sur les données définies dans cet annuaire pour publier les informations publiques des utilisateurs. Un SSO d'Entreprise peut aussi s'appuyer sur cette base utilisateur pour contrôler l'accès aux applications cibles.

Pas de changement du schéma de l'annuaire

Dans ce cas d'utilisation, l'organisation souhaitait ne pas changer le schéma de l'annuaire Active Directory. Le Référentiel de Sécurité de la solution de gestion des identités et des accès a donc été utilisé pour étendre le schéma en dehors de l'annuaire Active Directory
(cf. « Cas 3 » des exemples techniques.

Une solution qui s'adapte à l'organisation

La mise en œuvre d'une solution de synchronisation des identités s'adapte aux contraintes techniques et organisationnelles.

Par exemple, une politique de sécurité basée sur les groupes peut s'appliquer sur n'importe quel référentiel de l'entreprise. Il suffit qu'il contienne les informations autoritatives nécessaires à la définition et à l'application de cette politique.

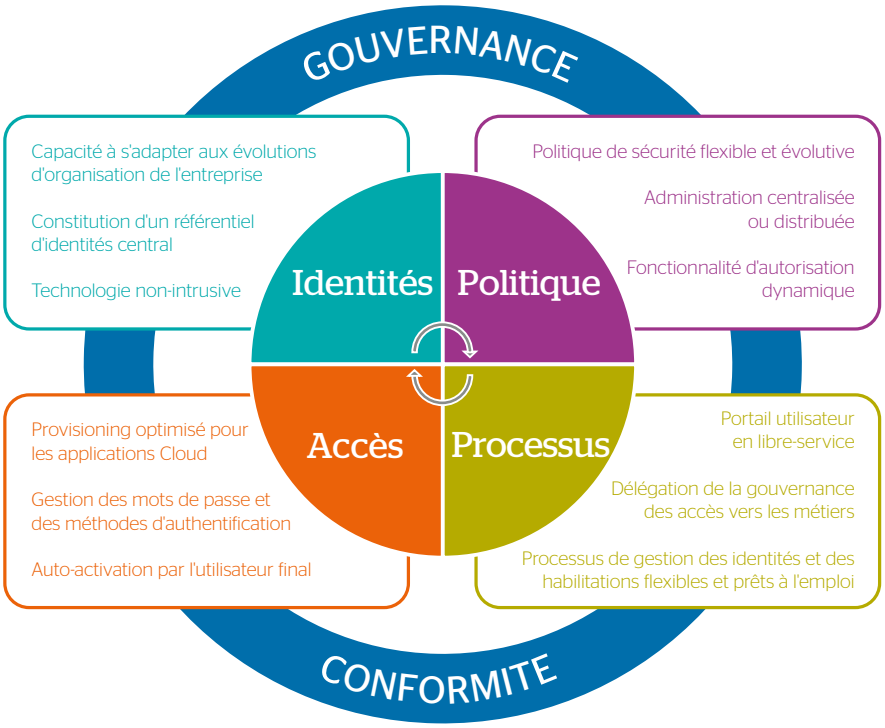
Les flux de mise à jours sont multidirectionnels; chaque référentiel peut donc être à la fois client et fournisseur de l'information, en fonction du caractère autoritatif ou non des données qu'il contient.

Une solution de provisionnement amont intégrée

Il y a deux manières de réfléchir à la mise en place d'une solution de création d'un référentiel d'identité. Soit de manière autonome et désynchronisée des autres fonctions de gestion des identités et des accès, soit de manière intégrée au sein d'une solution complète. Cette deuxième option permet d'intégrer immédiatement les problématiques globales et gérer les quatre piliers du marché de la gouvernance des identités et des accès :

- **Identités :** tout d'abord, identifier et gérer les utilisateurs qui ont des droits d'accès au système d'information ; les employés mais aussi les sous-traitants, les partenaires, voire les clients.
- **Politique :** puis définir et mettre à jour les droits des identités identifiées, en ayant les moyens de gérer l'évolution de ces droits dans le temps.
- **Processus :** permettre aux utilisateurs d'être eux-mêmes acteurs dans la gestion du cycle de vie des droits par la mise en œuvre de processus de workflow orientés métier.
- **Accès :** et pour finir, appliquer et vérifier la politique de sécurité définie pour les identités et validée via les processus sur les systèmes cible. Et permettre à chaque utilisateur de gérer ses comptes.

Evidian propose sa solution Evidian Identity Governance and Administration pour répondre aux objectifs de ces quatre piliers, comme illustré dans le schéma.



Le composant **ID Synchronisation** facilite la création du référentiel d'identités. Ce **référentiel d'identités** permet d'avoir une vue d'ensemble de toutes les identités ayant des accès au système d'information, qu'elles représentent des employés, des sous-traitants, des partenaires, voire des clients le cas échéant. Ces identités peuvent provenir de différentes sources (Excel, CSV, SQL...), être créées manuellement, ou bien par l'utilisateur final lui-même grâce à la fonction d'auto-inscription. La communication entre les sources d'identités et le référentiel d'identités peut être bidirectionnelle : la simulation et la gestion des seuils permettent de prévenir la corruption du référentiel d'identités. L'unicité des identités est assurée par la fonction de détection des doublons, même pour les identités qui sont en période de rétention. La déclaration anticipée des utilisateurs à l'aide d'un identifiant temporaire est également possible.

A propos d'Atos & Bull

Atos est un leader international de la transformation digitale avec environ 100 000 collaborateurs dans 72 pays et un chiffre d'affaires annuel de l'ordre 12 milliards d'euros. Numéro un européen du Big Data, de la Cybersécurité, des supercalculateurs et de l'environnement de travail connecté, le Groupe fournit des services Cloud, solutions d'infrastructure et gestion de données, applications et plateformes métiers, ainsi que des services transactionnels par l'intermédiaire de Worldline, le leader européen des services de paiement. Grâce à ses technologies de pointe et son expertise digitale & sectorielle, Atos accompagne la transformation digitale de ses clients dans les secteurs Défense, Finance, Santé, Industrie, Médias, Énergie & Utilities, Secteur Public, Distribution, Télécoms, et Transports. Partenaire informatique mondial des Jeux Olympiques et Paralympiques, le Groupe exerce ses activités sous les marques Atos, Atos Consulting, Atos Worldgrid, Bull, Canopy, Unify et Worldline. Atos SE (Societas Europea) est une entreprise cotée sur Euronext Paris et fait partie de l'indice CAC 40.

Pour plus d'informations **atos.net**

Bull est la marque Atos dédiée aux produits et logiciels de technologies distribués dans plus de 50 pays à travers le monde. Avec un héritage riche de plus de 80 années d'innovations technologiques, 2000 brevets et plus de 700 experts R&D soutenus par la Communauté scientifique d'Atos, Bull propose aux clients du Groupe Atos des produits et logiciels à forte valeur ajoutée afin de les accompagner dans leur transformation digitale pour répondre aux défis du Big Data et aux cybermenaces.

Leader européen du Calcul Haute Performance (HPC), Bull est à l'origine de nombreuses solutions reconnues dont bullx, le supercalculateur à faible consommation énergétique grâce à un système breveté par Bull, bullion l'un des serveurs x86 les plus puissants au monde pour répondre aux enjeux du Big Data, Evidian, les solutions logicielles de sécurité pour la gestion des identités et des accès (IAM), TrustWay, les modules cryptographiques sécurisés (HSM), Hoox, le smartphone ultra sécurisé.

Pour plus d'informations **bull.com**

Trusted partner for your **Digital Journey**

CT_170621_JG_Evidian-ID-Synchronization-factsheet-EN_01