11

La vie d'un annuaire d'entreprise

Les différentes étapes de la vie d'un annuaire

Les annuaires d'entreprise doivent évoluer constamment afin de répondre aux nouveaux besoins. Plus ils sont centralisés et partagés par un ensemble d'applications et de services, et plus ils doivent pouvoir évoluer rapidement tout en restant cohérents avec les applications et l'organisation de l'entreprise.

En effet, si un annuaire partagé n'évolue pas, les nouvelles applications auront tendance à gérer les données dont elles ont besoin dans leurs propres bases de données, même si ces données peuvent être communes avec d'autres applications. Ceci nuirait aux avantages de l'annuaire et peut rapidement rendre obsolète les investissements consentis pour sa conception et sa mise en œuvre.

Mais toute évolution doit être appliquée avec soin. Le changement du schéma pour les besoins d'une application peut altérer le fonctionnement des autres applications s'il n'est pas effectué correctement. De même, il est important de s'assurer que les données partagées par de nouvelles applications respectent la stratégie de sécurité de l'entreprise.

A fortiori, un contrôle trop rigide de l'annuaire et un processus d'évolution complexe ne faciliteront pas la tâche des utilisateurs, et risquent de nuire à la rapidité de mise en œuvre de nouveaux services.

L'annuaire doit donc pouvoir évoluer rapidement, mais de façon cohérente et tout en respectant la stratégie de sécurité mise en œuvre.

Nous allons décrire dans ce chapitre les évolutions possibles d'un annuaire et comment assurer la mise en œuvre de celui-ci efficacement.

Les bonnes questions à se poser

Nous allons prendre un exemple simple d'une nouvelle application qui va s'appuyer sur un annuaire d'entreprise existant, afin de mettre en évidence l'impact sur celui-ci. Dans ce cas, les différentes questions à se poser sont les suivantes :

- Quelles sont les données qui existent déjà dans l'annuaire d'entreprise et qui peuvent être utiles pour la nouvelle application ?
- Quelles sont les données qu'il faut rajouter à l'annuaire d'entreprise, susceptibles d'être réutilisées par d'autres applications ?
- Dans le cas de rajout de données, comment modifier le schéma de l'annuaire en réduisant l'impact sur les applications existantes ?
- Quels sont les droits d'accès aux données, aussi bien en lecture qu'en mise à jour, qu'il faut attribuer à la nouvelle application ?
- Quel va être l'impact sur les performances de l'annuaire ?
- Quel va être l'impact sur la disponibilité de l'annuaire (par exemple vingt-quatre heures sur vingt-quatre et sept jours sur sept) ?
- Dans la mesure où la nouvelle application contient sa propre base de données avec des informations qui se trouvent aussi dans l'annuaire, comment synchroniser les deux environnements ?

Pour répondre à ces questions il est important de mettre en place une organisation et des outils que nous décrivons ci-dessous.

Les acteurs

Pour répondre à ces questions, il faut avant tout identifier ceux qui sont les mieux placés pour le faire et mettre en place une organisation adéquate.

Généralement, il existe trois catégories de personnes concernées par l'annuaire :

- *Les gestionnaires du schéma* : ils sont responsables de toute modification du schéma de l'annuaire (classes d'objets, attributs, syntaxes) et du DIT. Ils doivent aussi avoir une bonne connaissance de la sémantique des données (attributs et classes d'objets), afin d'assurer une cohérence « métier » du schéma et être en mesure de communiquer aux utilisateurs de l'annuaire la signification des attributs et des classes.
- *Les gestionnaires du contenu* : ils sont responsables du contenu lui-même. Leur rôle consiste à s'assurer de la qualité des données qui se trouvent dans l'annuaire, et à définir les règles de sécurité qui s'y appliquent. Ils doivent aussi s'assurer que ces règles sont bien appliquées. Pour cela, il est nécessaire qu'ils aient une bonne connaissance des applications qui utilisent l'annuaire, des besoins pour chacune d'elles, y compris les contraintes « métier » de sécurité.
- *Les administrateurs des plates-formes* : ils sont responsables de la gestion des plates-formes matérielles et logicielles qui prennent en charge l'annuaire d'entreprise. Ils doivent s'assurer qu'elles répondent aux besoins des applications en termes de disponibilité et de sécurité, et doivent être en mesure de piloter l'activité des serveurs et d'intervenir en cas de problème.

La principale difficulté dans la gestion d'un annuaire d'entreprise est qu'aucun de ces acteurs ne peut agir sans l'autre. En effet, l'annuaire ne peut être géré que par une équipe mixte, car les aspects fonctionnels, organisationnels et techniques doivent être pris en compte pour toute évolution de celui-ci. Comme nous l'avons vu précédemment, toute nouvelle application nécessite de se poser des questions sur le plan fonctionnel (quelles données ?), sur le plan organisationnel (quelles acteurs et quels droits ont-ils ?) et sur le plan technique (quels impacts sur les plates-formes ?).

Favoriser la réutilisation de l'annuaire

Un des écueils à éviter lors du développement d'applications est de négliger ce qui existe déjà et de recréer un annuaire LDAP pour les besoins propres à leurs applications. Les problèmes qui peuvent alors surgir au moment de la mise en production sont un schéma qui peut être incompatible avec l'existant, une duplication des données nécessitant une synchronisation de différents objets dans l'annuaire lui-même, et un volume de données qui peut s'accroître rapidement et dégrader les performances globales de l'annuaire.

Pour éviter ceci, il est important de communiquer à l'intérieur de l'entreprise sur le contenu de l'annuaire. Pour cela, il faut disposer de documents décrivant ce que contient l'annuaire, régulièrement mis à jour. Il s'agit essentiellement de bien décrire le schéma, la sémantique de chaque attribut, les règles de sécurité, ainsi que le DIT.

Un des moyens pourrait être d'interroger l'annuaire lui-même afin d'y lire le schéma. Mais il arrive souvent que pour des raisons de sécurité il ne soit pas accessible simplement par tous. D'autre part, le résultat de la lecture n'est pas souvent très parlant pour les développeurs et concepteurs fonctionnels, car il contient des informations techniques comme l'OID ou des attributs techniques propres aux serveurs d'annuaire.

Le meilleur moyen pour documenter le schéma est de rédiger sa description dans un document au format bureautique accessible à tous ou encore mieux dans un document HTML publié sur l'intranet de l'entreprise.

L'ensemble des rubriques importantes qu'il est nécessaire de documenter est présenté ci-dessous. Pour plus de détails sur ces rubriques, veuillez vous référer au chapitre 7, relatif à la conception fonctionnelle de l'annuaire.

Les attributs

Un tableau de ce type pourra être établi afin d'identifier et de décrire les attributs.

Attribut	Standard ou spécifique	Syntaxe	Mono-valeur	Appartenance	Fréquence d'utilisation
Nom de l'attribut	Il s'agit de préciser si l'attribut appartient déjà au serveur d'annuaire ou si c'est un nouvel attribut.	Dans le cas d'un attribut spécifique, il faut choisir sa syntaxe parmi celles qui sont supportées par le serveur d'annuaire.	Dans le cas d'un attribut spécifique, il faut préciser s'il peut avoir plusieurs valeurs ou non.	Il s'agit de décrire l'origine de l'attribut (application et utilisateur).	Faible, fréquente ou aléatoire

Les classes d'objets

Un tableau de ce type pourra être établi afin d'identifier et de décrire les classes d'objets.

Classe	Standard ou spécifique	Classe mère	Type de la classe	Attributs obligatoires	Attributs facultatifs
Nom de la classe	Il s'agit de préciser si la classe appartient déjà au serveur d'annuaire ou si c'est une nouvelle classe.	S'il s'agit d'une classe spécifique, il faut choisir la classe dont elle dérive parmi celles supportées par le serveur d'annuaire.	S'il s'agit d'une classe spécifique, il faut choisir son type parmi : structurel, abstrait ou auxiliaire.	S'il s'agit d'une classe spécifique, il faut définir la liste des attributs obligatoires (elle peut ne pas en avoir).	S'il s'agit d'une classe spécifique, il faut définir la liste des attributs autorisés et facultatifs (elle peut ne pas en avoir).

On pourra aussi utiliser le diagramme de classe de la modélisation UML pour représenter les relations entre celles-ci. La figure suivante illustre par un exemple ce type de diagramme, dans lequel on trouve :

- les classes et leurs attributs (dans cette étape, on ne liste pas les méthodes, elles seront ajoutées au modèle lors de la phase de réalisation) ;
- les relations d'héritage entre classes, représentées par des flèches vers le haut ;
- les associations avec leurs cardinalités, représentées par des liens.

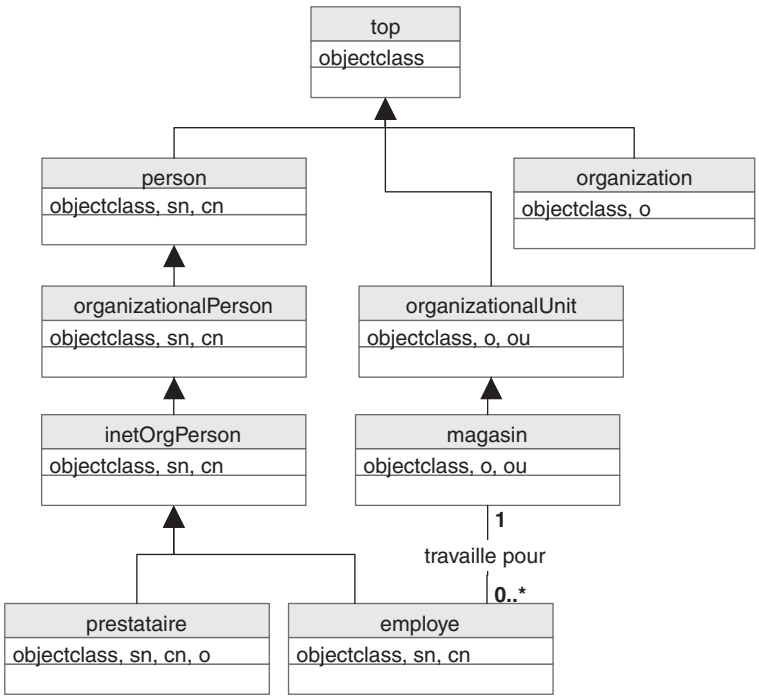


Figure 11.1
Exemple de diagramme de classe UML appliqué à LDAP

Les acteurs

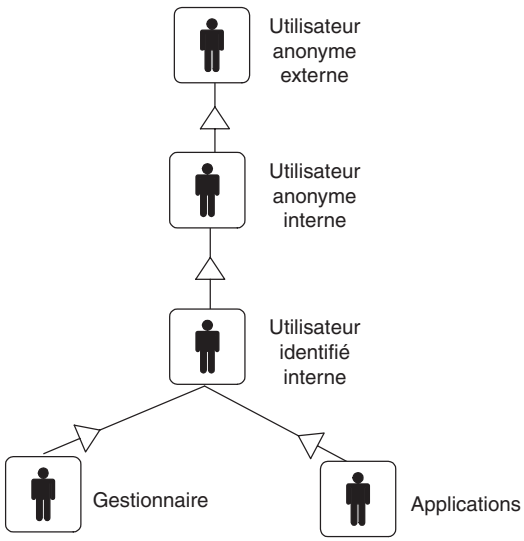
Les acteurs sont les différentes personnes ou applications autorisées à se connecter à l'annuaire pour y lire des données ou les mettre à jour.

Un tableau de ce type pourra être établi afin d'identifier et de décrire les utilisateurs :

Utilisateur	Type	Nombre d'utilisateurs	Fréquence d'utilisation de l'annuaire	Temps moyen d'utilisation
Anonyme, identifié ou caractérisé par un critère (par exemple, utilisateur externe ou appartenant à un groupe)	Personne physique ou application	Nombre d'utilisateurs potentiels, et nombre d'utilisateurs simultanés	Peu, une fois par jour, de façon aléatoire (cas d'une messagerie par exemple), etc.	Il faut différencier le temps moyen de consultation du temps moyen de recherche requis.

On peut aussi utiliser la modélisation UML pour représenter les liens d'héritage entre les acteurs (à la fois les utilisateurs, les gestionnaires et les administrateurs).

Figure 11.2
Exemple de diagramme d'acteurs



La définition des droits

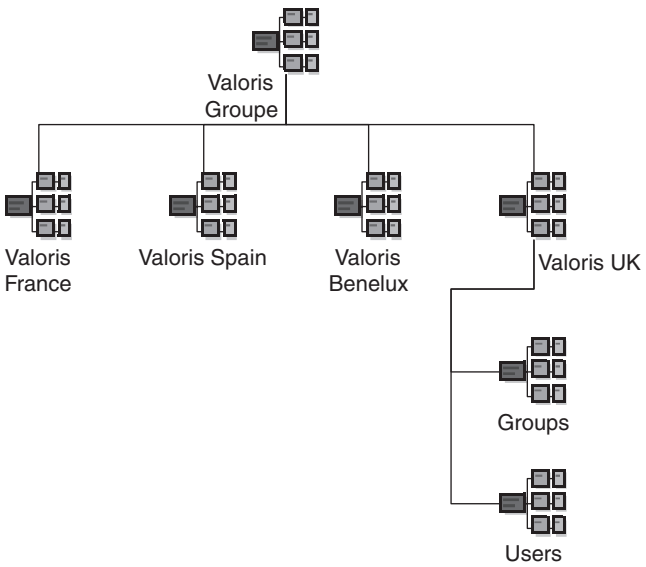
Un tableau de ce type pourra être établi afin de décrire les droits des utilisateurs sur les données de l'annuaire :

Acteur	Droit	Classe d'objet et attribut
Utilisateur, gestionnaire ou administrateur	Rechercher et lire des données Comparer des données Modifier un objet Supprimer un objet Ajouter un objet Renommer le DN d'un objet	Nom de la classe d'objet et/ou de l'attribut concerné (un droit peut s'appliquer à un attribut, quelle que soit la classe à laquelle il appartient)

Le DIT ou l'arbre des données

Le DIT peut être dessiné à l'aide de Microsoft Visio 2000/2003 comme le montre la figure suivante.

Figure 11.3
Exemple DIT avec Microsoft Visio



Microsoft Vision 2000/2003 version Entreprise possède en standard un connecteur LDAP qui permet d'interroger en temps réel un annuaire et dessiner son arborescence automatiquement. Il permet aussi de dessiner des diagrammes UML et des objets LDAP à l'aide de symboles graphiques dédiés à cet effet.

Modifier le schéma

Modifier le schéma d'un annuaire est une opération délicate, qui peut s'avérer fastidieuse et avoir des effets importants sur les applications existantes. Nous allons décrire les principaux cas qui peuvent survenir et décrire comment procéder dans chacun d'eux.

Pour modifier le schéma, il est conseillé d'utiliser la console d'administration de votre serveur d'annuaire. Il est en effet possible de le faire avec des commandes LDIF ou avec un outil du marché, mais la console d'administration offre une interface plus conviviale et effectue des contrôles de cohérence indispensables à ce type d'opération.

Notons également que la modification du schéma sur un serveur peut se répercuter sur les éventuels serveurs répliqués. C'est ce qui se passe lorsqu'on a plusieurs serveurs avec une stratégie de réplication entre eux, comme par exemple des serveurs Sun Java System Directory Server, ou entre plusieurs serveurs Microsoft Active Directory faisant partie d'une même forêt (il n'y a qu'un seul schéma pour l'ensemble des contrôleurs de domaine d'une même forêt).

Rajouter un attribut à une classe d'objet

Pour cela, il faut d'abord rajouter l'attribut à la liste des attributs de l'annuaire, puis il faut mettre celui-ci dans la classe d'objet voulue. Ceci peut être fait avec la console d'administration du serveur d'annuaire, comme nous l'avons montré dans le chapitre 9, et n'a pas d'impact sur les données existantes. La majorité des serveurs d'annuaire autorise cette opération durant le fonctionnement du serveur : il n'est donc pas nécessaire de l'arrêter. Certains d'entre eux nécessitent de recharger le schéma afin qu'il soit pris en compte.

Si des recherches doivent être effectuées sur cet attribut, il peut être utile d'activer son indexation dans le serveur. Ceci a pour effet de mettre à jour la base d'index en y rajoutant l'attribut en question, et d'accélérer ainsi les temps de réponse lors d'une recherche. Pour cela, il faut utiliser la console d'administration de l'annuaire, comme le montre la figure 11.4.

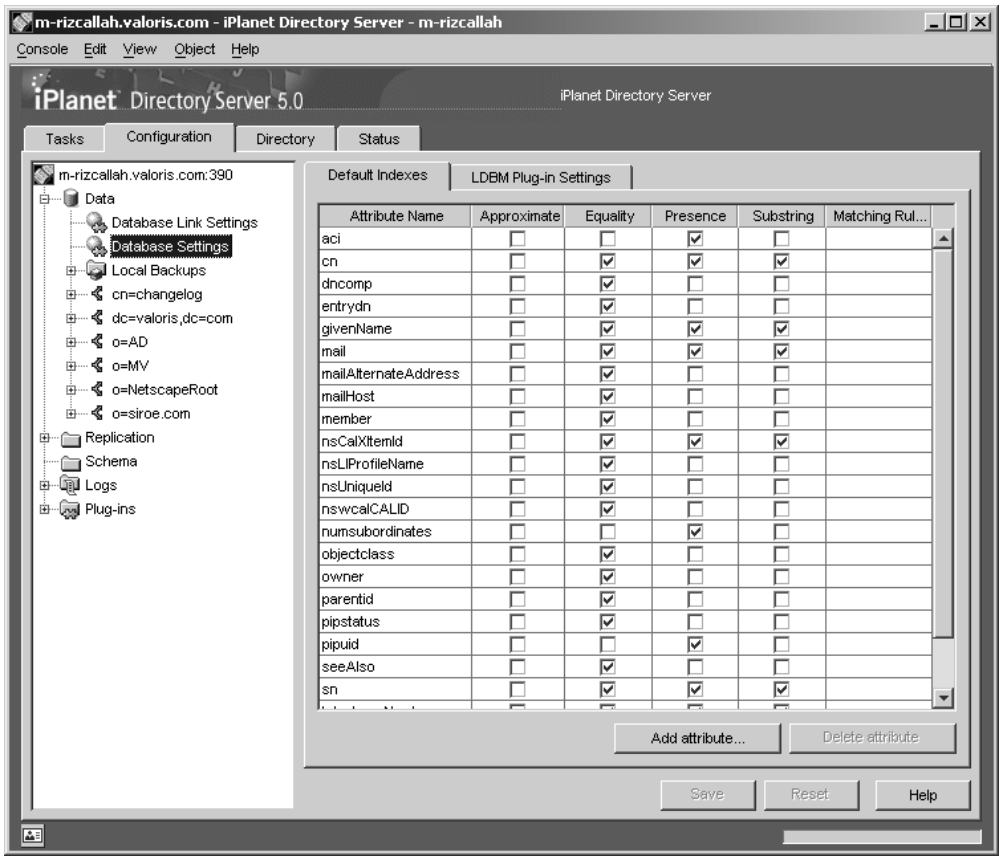


Figure 11.4

Exemple d'indexation d'attributs avec iPlanet Directory Server 5.0 de Sun

Notons que l'indexation augmente les temps de mise à jour, et n'est utile que pour la lecture des données. Dans le cas d'une architecture maître-esclave, il sera plus judicieux de mettre en œuvre l'indexation des attributs sur les serveurs esclaves uniquement.

Supprimer un attribut à une classe d'objet

Certains serveurs d'annuaire, tel Microsoft Active Directory pour Windows 2000 (ce n'est plus le cas avec Windows Serveur 2003), n'autorisent pas la suppression d'un attribut d'une classe d'objet, ni de l'annuaire pour l'ensemble des classes. Cela provient principalement de la nécessité de supprimer les valeurs de cet attribut pour la classe d'objet ; ceci n'est pas fait automatiquement par le serveur.

Pour supprimer un attribut, il faut donc suivre les étapes suivantes :

1. Vérifier si le serveur d'annuaire que vous utilisez autorise la suppression d'attributs et s'il supprime les valeurs de cet attribut automatiquement ou pas. Pour cela il suffit de faire l'essai avec une classe d'objet de test.
2. Si c'est le cas, identifier toutes les classes d'objets ainsi que les règles de sécurité (ACL) qui font référence à cet attribut. Pour trouver les classes d'objets, il suffit d'extraire le schéma dans un fichier LDIF et d'y rechercher l'attribut en question. Pour trouver les règles de sécurité, il faut soit utiliser la console d'administration de l'annuaire, soit extraire celles-ci dans un fichier LDIF si le serveur d'annuaire utilise la syntaxe de l'IETF (voir le chapitre 4), puis rechercher l'attribut dans ce fichier.

   L'exemple suivant montre comment extraire les règles de sécurité dans un format LDIF :

   ```
   ldapsearch -b "o=entreprise.com" -p 391 -D "cn=Directory Manager" -w password
   objectclass=* aci
   ```

   Rappelons que dans cet exemple les paramètres `–b` et `–p` permettent de désigner respectivement la base de la recherche et l'adresse du port IP où se trouve l'annuaire. Les paramètres `–D` et `–w` désignent respectivement l'identifiant et le mot de passe de l'administrateur.
3. Écrire un programme qui va supprimer toutes les valeurs de cet attribut pour la classe d'objet en question. Ce programme peut être écrit avec un langage de script comme Perl, ou encore en VB avec ADSI/.Net.
4. Adapter les règles de sécurité afin de ne plus faire référence à l'attribut supprimé.
5. Supprimer l'attribut de la classe d'objet à l'aide de la console d'administration de l'annuaire.
6. Supprimer l'attribut du schéma de l'annuaire s'il n'est pas utilisé par une autre classe d'objet.

Modifier la syntaxe d'un attribut

Pour modifier la syntaxe d'un attribut, il suffit d'utiliser la console d'administration de l'annuaire. Mais, comme dans la suppression d'un attribut, certains serveurs d'annuaire n'autorisent pas la modification de la syntaxe. Il faut alors le supprimer et en ajouter un autre.

En revanche, lorsqu'il est possible de le faire, comme avec Sun Java Directory Server, il n'y a pas de vérification automatique de la syntaxe pour les valeurs déjà enregistrées dans l'annuaire. Celle-ci se fera uniquement pour les nouvelles entrées. Pour trouver les valeurs existantes qui peuvent ne pas correspondre à la nouvelle syntaxe, il suffit de lire et d'écrire toutes les valeurs de cet attribut ; en cas d'erreur de syntaxe une erreur sera générée par l'interface LDAP utilisée lors de la mise à jour.

Ajouter une classe d'objet

C'est l'opération la plus simple : il suffit d'utiliser la console d'administration de l'annuaire. Il faut vérifier au préalable que les attributs de la nouvelle classe existent déjà dans le schéma de l'annuaire. Pensez aussi à dériver la nouvelle classe d'une classe faisant partie du standard LDAP ou propre au serveur d'annuaire et qui soit sémantiquement la plus proche. Par exemple, si la nouvelle classe concerne une personne, que ce soit un client ou un employé, il faut qu'elle dérive de la classe `organizationalperson`.

Supprimer une classe d'objet

Avant de supprimer une classe d'objet, il faut s'assurer qu'il n'existe plus d'instance de cette classe dans l'annuaire. L'exemple suivant montre comment le faire à l'aide de la commande `ldapsearch` :

```
ldapsearch -b "o=entreprise.com" -p 391 -D "cn=Directory Manager" -w password
objectclass=classeasupprimer
```

Pour supprimer tous les objets de la classe, on peut soit utiliser la console d'administration de l'annuaire et supprimer les objets un par un, soit écrire un programme en VB ou en Perl en appliquant un filtre de recherche sur cette classe.

Avec Microsoft Active Directory, il est possible d'utiliser l'outil ADSI Edit. Il permet de trier les enregistrements par classe d'objet dans une branche donnée. Mais il faut quand même supprimer les objets un à un, car cet outil ne permet pas d'en sélectionner plusieurs à la fois.

Là aussi, il est important de vérifier qu'il n'existe pas de règle de sécurité qui fasse référence à cette classe d'objet.

Tester et analyser les performances d'un annuaire

Comment tester les performances d'un annuaire LDAP ?

Une premier moyen est d'utiliser tout simplement une des applications de l'entreprise accédant à l'annuaire et d'en tester les performances manuellement. On peut encore utiliser un outil générique intégrant un client LDAP, comme le produit LDAP Browser/Editor que nous avons cité dans le chapitre 10. Ce type de test permet de savoir si l'annuaire est toujours opérationnel, et de connaître rapidement les temps de réponse de celui-ci pour un utilisateur connecté à la fois.

Une autre façon de faire est d'utiliser un outil spécifique du marché, comme l'outil LDAP Optimiser de la société NCC-USA (outil pour Solaris, voir l'adresse *www.nccgroupusa.com/ldap/optimizer.htm* pour plus de renseignements) ou encore l'outil LoadRunner de la société Mercury Interactive *(www.merc-int.com)*. Ces outils permettent de générer du trafic entre un client LDAP et un serveur, et d'analyser les temps de réponse. Mais ils peuvent parfois être trop onéreux pour l'usage que l'on veut en faire, et constituer un investissement difficile à justifier.

Enfin, il est possible d'écrire son propre outil de test. Celui-ci peut être fait avec un langage de script, comme Perl, ou à l'aide d'un programme en C ou C++. Il est conseillé d'au moins tester les temps de réponse pour s'identifier à l'annuaire, pour effectuer une recherche et pour effectuer une mise à jour.

Comment analyser l'activité d'un annuaire LDAP ?

En général, les serveurs d'annuaire offrent en standard des outils de monitoring qui permettent de connaître les temps de réponse du serveur. Ces temps de réponses sont analysés à partir de journaux qu'il est possible de générer ou pas à la demande.

À titre d'exemple, le serveur iPlanet Directory Server 5.0 de Sun permet, *via* la console de l'outil, de consulter le nombre moyen d'entrées par minute, renvoyées par l'annuaire aux clients qui l'interrogent, comme le montre la figure 11.5 dans le paramètre « Entries Sent To Clients ».

Ce chiffre n'est pas significatif en soit : il permet de connaître la capacité de réponse du serveur et de savoir si la charge augmente avec le temps (encore faut-il mémoriser cette valeur dans le temps), mais ne permet pas de dire si celle-ci est suffisante pour les utilisateurs.

C'est pour cela qu'il est souvent plus utile d'envoyer ces informations à un outil d'administration proprement dit, comme Tivoli d'IBM ou UniCenter de TNG. L'intégration entre le serveur d'annuaire et ces outils d'administration se fait à l'aide du protocole SNMP. Chaque serveur d'annuaire est capable d'envoyer des informations sur son activité comme par exemple le nombre d'identifications anonymes ou le nombre d'identifications inabouties à l'aide de ce protocole. Il est alors possible de générer des alertes dans l'outil d'administration lorsque l'un de ces paramètres atteint une valeur donnée.

Sous Windows NT et 2000/2003, certains serveurs d'annuaire offrent la possibilité d'analyser les performances à l'aide de la console MMC. Par exemple, lorsqu'on installe iPlanet Directory Server 5.0 de Sun sous Windows 2000, ou bien AD/AM de Microsoft (Active Directory Application Mode), on peut analyser un certain nombre de paramètres : nombre de recherches par seconde, nombre de suppressions par seconde, etc. Pour cela, il suffit de lancer l'outil de suivi des performances, accessible dans le menu des outils d'administration de Windows 2000/2003.

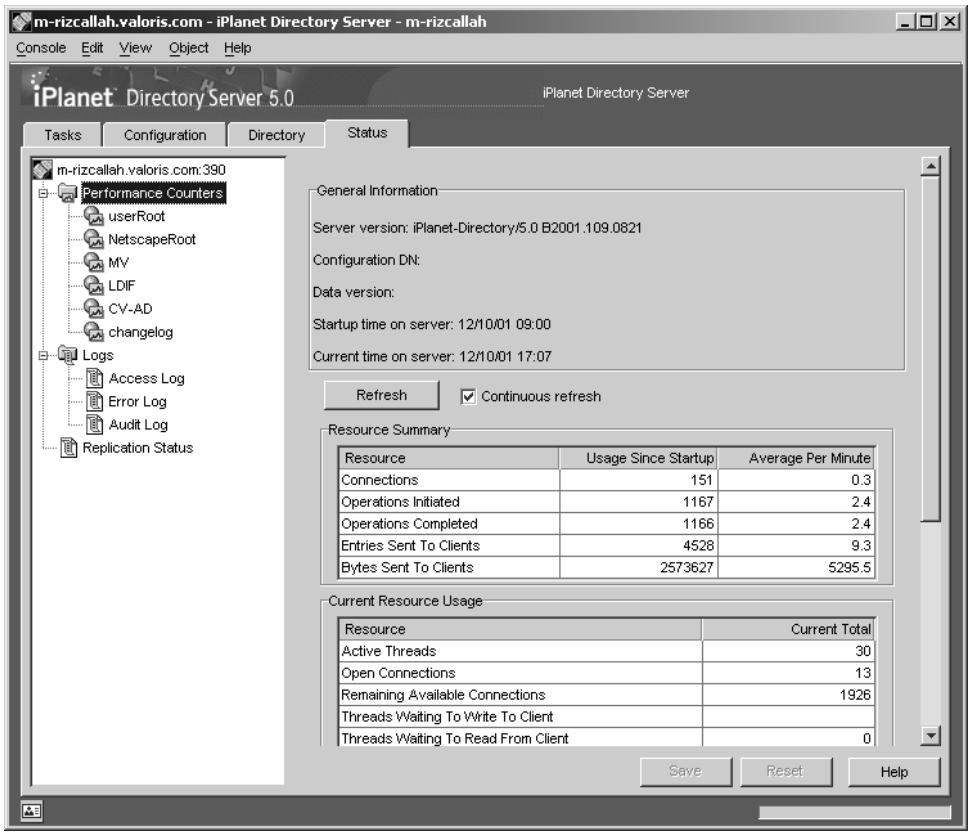
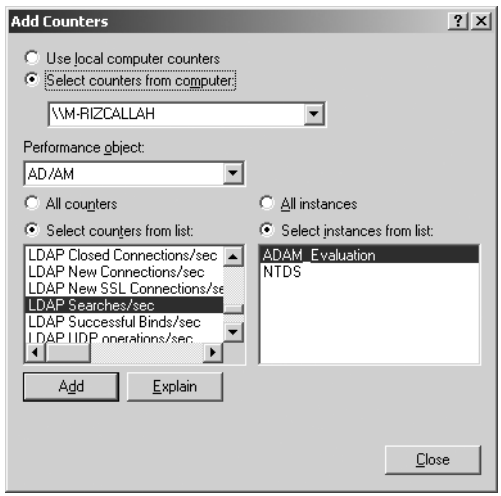


Figure 11.5

Exemple de monitoring avec iPlanet Directory Server 5.0 de Sun

Figure 11.6

Performances de AD/AM avec le moniteur de performances de Windows 2000/2003



Mais quels sont les paramètres significatifs qui permettent d'analyser l'activité d'un annuaire ? À notre avis, les principaux paramètres qu'il faut être en mesure de suivre sont les suivants :

- Le nombre moyen de connexions simultanées : ce nombre permet de savoir combien de clients se connectent en moyenne par jour, par heure ou par minute à l'annuaire. Il est en général fourni par l'annuaire.
- Le nombre moyen de requêtes traitées par l'annuaire : ce nombre permet de savoir combien de requêtes sont envoyées à l'annuaire par jour, par heure ou par minute, comme la lecture ou l'écriture d'un enregistrement. Il est en général fourni par l'annuaire.
- Le nombre moyen d'octets renvoyés par l'annuaire en réponse à ces requêtes : ce nombre permet d'avoir une idée sur la volumétrie des données associées aux requêtes. Il est en général fourni par l'annuaire.
- Le nombre moyen de requêtes par client : ce nombre peut être déduit des deux premiers.
- Le temps moyen de réponses à une requête par client : ce nombre est fourni par le serveur d'annuaire en standard, ou doit être calculé en analysant les fichiers de traces générés par l'annuaire. Notons que pour cela, il est nécessaire d'activer ces traces, ce qui peut dégrader les performances du serveur.
- Le nombre de renvois de référence traités par l'annuaire : ce nombre permet de mettre en évidence l'impact du renvoi de référence sur les performances finales vues par le client. Si ce nombre est faible, les performances sont essentiellement dépendantes de la capacité du serveur interrogé. Sinon, elles peuvent dépendrent de celles des autres serveurs auxquels il est fait référence.

Viennent ensuite les paramètres de sécurité, qui sont les suivants :

- le nombre de connexions anonymes ;
- le nombre d'identifications erronées (mot de passe ou identifiant incorrects) ;
- le nombre de requêtes violant les règles de sécurité, comme la mise à jour dans une branche non autorisée, ou la lecture d'un attribut non autorisé. Certaines applications peuvent utiliser ce moyen une première fois afin de savoir si l'utilisateur connecté a droit d'accès aux données ou pas, et adapter en conséquence l'interface utilisateur présentée.

Optimiser les performances d'un annuaire

Il existe plusieurs façons d'améliorer les performances d'un annuaire. La façon la plus simple consiste à augmenter la mémoire et la puissance de la plate-forme sur laquelle repose l'annuaire. Mais ceci est généralement coûteux ; ceux qui en ont les moyens n'hésiteront pas à s'équiper par exemple d'une machine Sun E10 000, pouvant supporter jusqu'à soixante-quatre processeurs. Heureusement, il existe des moyens moins onéreux pour obtenir des résultats probants.

Ces moyens dépendent essentiellement des capacités offertes par le serveur d'annuaire choisi. Ils doivent pour cela offrir des fonctions particulières, prises en charge par une grande majorité des outils, et que nous répertorions ci-dessous.

L'indexation des attributs

Tous les attributs ne sont généralement pas indexés par le moteur de base de données de l'annuaire. Les attributs les plus couramment utilisés le sont par défaut, comme le `cn` et `l'uid`. Si les recherches effectuées dans l'annuaire utilisent couramment un attribut, il est conseillé d'indexer celui-ci. En revanche, il n'est pas conseillé d'indexer tous les attributs car ceci augmente considérablement les volumes disques occupés par la base de données et peut réduire les performances de mise à jour de l'annuaire (généralement les serveurs d'annuaire créent les index lors de la mise à jour).

Mais comment savoir quels attributs sont les plus couramment utilisés par les applications ? Si vous maîtrisez le code source des applications qui accèdent à l'annuaire, il est possible d'analyser celui-ci et d'en déduire les attributs les plus couramment utilisés. Mais ceci ne sera pas généralement le cas : ceux chargés de l'exploitation de l'annuaire ne sont pas ceux qui réalisent les applications.

Le seul moyen consiste donc à analyser les fichiers de trace de l'annuaire. Voici un exemple extrait du journal des accès à Sun Java System Directory Server :

```
[20/Oct/2001:15:37:34 +0100] conn=86 op=0 BIND dn="cn=Directory Manager"
method=128 version=3
[20/Oct/2001:15:37:34 +0100] conn=86 op=0 RESULT err=0 tag=97 nentries=0
etime=0 dn="cn=directory manager"
[20/Oct/2001:15:37:34 +0100] conn=86 op=1 SRCH base="dc=valoris,dc=com"
scope=2 filter="(objectClass=person)" attrs=ALL
[20/Oct/2001:15:37:34 +0100] conn=86 op=1 RESULT err=0 tag=101 nentries=6
etime=0
[20/Oct/2001:15:37:34 +0100] conn=86 op=2 UNBIND
```

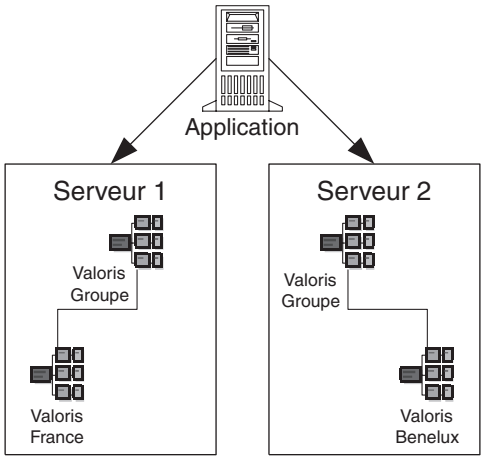
Pour retrouver à partir de ce fichier les attributs les plus utilisés pour la recherche, il faut analyser les lignes contenant le mot-clé `SRCH` (fonction de recherche), puis analyser la chaîne de caractères associée au mot-clé `filter` (filtre de la recherche). Un script en Perl pourrait par exemple extraire les noms d'attributs des chaînes de filtre et compter pour chacun d'eux le nombre de fois où ils sont utilisés. Il faut ensuite comparer les attributs apparaissant un nombre de fois élevé avec les attributs indexés dans l'annuaire, et décider s'il faut en rajouter ou pas.

La répartition de charge

Celle-ci peut se faire de deux façons : asymétriquement en répartissant la charge de façon statique sur plusieurs serveurs contenant chacun un sous-ensemble des branches de l'annuaire, ou symétriquement en répartissant la charge de façon aléatoire (ou dynamique) sur plusieurs annuaires identiques.

Dans le premier cas, il s'agit tout simplement de répartir les données sur plusieurs serveurs et d'accéder au serveur voulu à partir de l'application, comme le montre la figure 11.7.

Figure 11.7
Exemple de répartition de charge asymétrique



L'inconvénient de cette méthode est que l'application doit savoir quel serveur interroger à chaque requête et répartir la charge de façon statique entre ceux-ci. En outre, si pour des raisons de performance il est nécessaire de répartir à nouveau l'un des serveurs en plusieurs autres serveurs, il faudra modifier toutes les applications qui y accèdent en conséquence.

La deuxième méthode consiste tout d'abord à dupliquer les serveurs avec un contenu identique puis à répartir la charge en fonction des capacités restantes de chaque serveur.

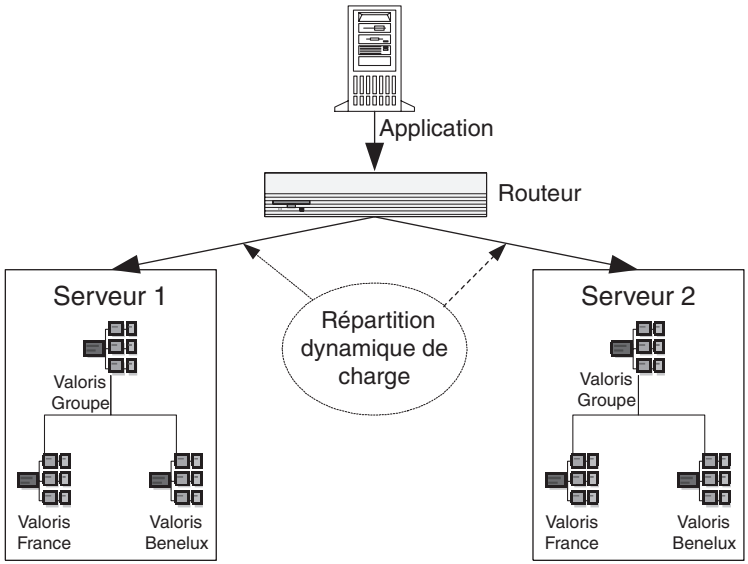
Les différentes techniques de réplication sont décrites en détail dans le chapitre 8. La réplication multimaître est la plus courante et la plus simple à mettre en œuvre. Elle consiste à avoir au moins deux serveurs identiques qui se synchronisent automatiquement dès qu'on met à jour des données sur l'un d'eux. La plupart des serveurs d'annuaires du marché supportent la réplication multimaître dans leur dernière version. Il est aussi possible de dédier des serveurs à l'écriture et d'autres à la lecture. Cette dernière étant généralement plus fréquente, on pourra augmenter les capacités de ces derniers ainsi que leur nombre en fonction des besoins, tout en ayant un ou deux serveurs dédiés à l'écriture.

Une fois la réplication effectuée, il faut répartir la charge sur les différents serveurs. Ceci nécessite un outil spécifique, généralement un routeur réseau, situé en amont des différents serveurs, chargé d'analyser les requêtes LDAP et de les transmettre au serveur le moins chargé.

Ce type d'outil peut être soit un routeur réseau comme LocalDirector de Cisco, ou un routeur Alteon de Nortel, soit un logiciel spécifique de type proxy inversé, que nous décrivons ci-dessous. Certains outils sont capables de répartir la charge sur des critères « intelligents » comme :

- La disponibilité des serveurs : si l'un des serveurs tombe en panne, aucun flux n'est redirigé vers lui jusqu'à ce qu'il soit remis en marche.

Figure 11.8
Exemple de répartition de charge dynamique



- Les temps de réponse : si un serveur met du temps à répondre aux requêtes qui lui sont envoyées, il sera sollicité plus rarement par le routeur.
- La capacité des serveurs : si un serveur a des capacités supérieures à d'autres, il est possible de lui attribuer un « poids » afin qu'il reçoive plus de requêtes.

La désactivation du contrôle du schéma

Un des moyens permettant d'accélérer les performances en écriture d'un annuaire est tout simplement de désactiver le contrôle du schéma si le serveur d'annuaire offre cette fonctionnalité. À chaque mise à jour d'une donnée, le serveur va vérifier la syntaxe des attributs et la cohérence des classes d'objets, ce qui consomme des ressources machines. Si on désactive le contrôle du schéma, ces contrôles ne sont plus effectués, et ceci accélère notablement les performances du serveur.

En contrepartie, on risque de mettre à jour des données de façon incohérente dans l'annuaire et donc de polluer celui-ci à terme. Il est donc fortement conseillé de désactiver le contrôle du schéma pour le serveur de production uniquement : en effet, les erreurs sont plus fréquentes lors de la programmation des applications. Pour cela il faut dédier un serveur d'annuaire à l'environnement de développement.

L'utilisation d'un proxy LDAP inversé

Nous décrivons en détail ce qu'est un proxy LDAP inversé dans le chapitre 8 de ce livre. Parmi les fonctionnalités qu'il offre, celles qui concernent les performances sont le traitement du renvoi de référence et le « cache » mémoire du résultat des requêtes.

Modifier l'arborescence de l'arbre LDAP (DIT)

Il arrive souvent que les branches de l'arbre nécessitent d'être modifiées pour répondre à de nouveaux besoins. Par ailleurs, les applications qui accèdent à l'annuaire peuvent créer des branches pour des besoins propres, qui nécessitent d'être maintenues par des administrateurs en accédant directement à l'annuaire. Nous allons décrire dans ce paragraphe comment manipuler les branches d'un arbre LDAP et les précautions à prendre dans les différents cas de figure.

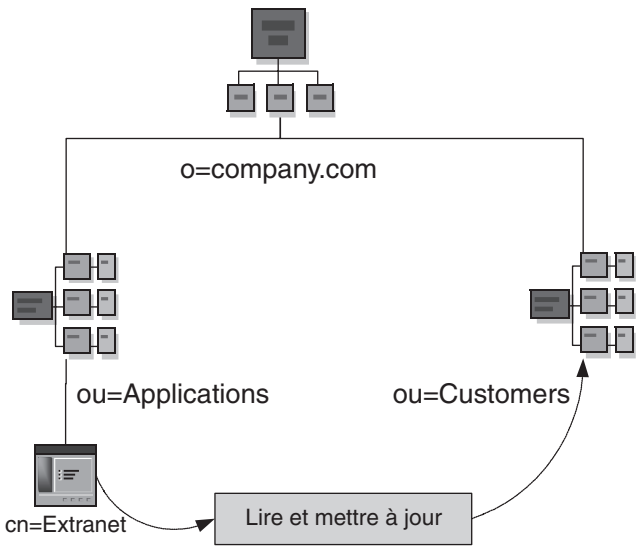
Ajouter une branche

C'est une opération assez simple car elle consiste à rajouter un objet dans l'annuaire. Cet objet deviendra une branche lorsqu'on ajoutera d'autres objets sous celui-ci. Elle peut être réalisée avec n'importe quel outil, y compris la console d'administration de l'annuaire.

Néanmoins, il est nécessaire de prendre quelques précautions lors de cette opération :

- Il est conseillé d'utiliser la classe d'objet `organizationalunit` pour créer le nœud de la branche.
- Il faut vérifier que la branche est accessible aux applications : pour cela il est nécessaire de s'assurer que les droits d'accès appliqués sur le nœud père autorisent bien l'accès aux objets de cette branche ou pas en fonction des besoins. Si ce n'est pas le cas, il faut créer une règle d'accès spécifique à cette branche, comme le montre la figure 11.9. Par exemple, si on crée une branche qui va contenir les clients de l'entreprise, il faut s'assurer que les applications extranets peuvent bien lire et écrire dans cette branche.

Figure 11.9
Exemple de règle d'accès associée à une nouvelle branche



Déplacer une branche

Déplacer une branche est une opération fastidieuse si vous n'utilisez pas un outil dédié à cet effet. L'interface LDAP n'offre pas de fonction prête à l'emploi pour réaliser cette opération. Il faut créer la nouvelle branche, recopier les objets de l'ancienne branche vers la nouvelle un à un, puis supprimer ces objets de l'ancienne branche et enfin supprimer le nœud associé.

Heureusement, la plupart des consoles d'administration des annuaires offrent en standard la possibilité de déplacer une branche. Ceci est offert à l'aide de l'interface graphique de la console en faisant un glisser/déplacer de la branche, ou tout simplement un copier/coller.

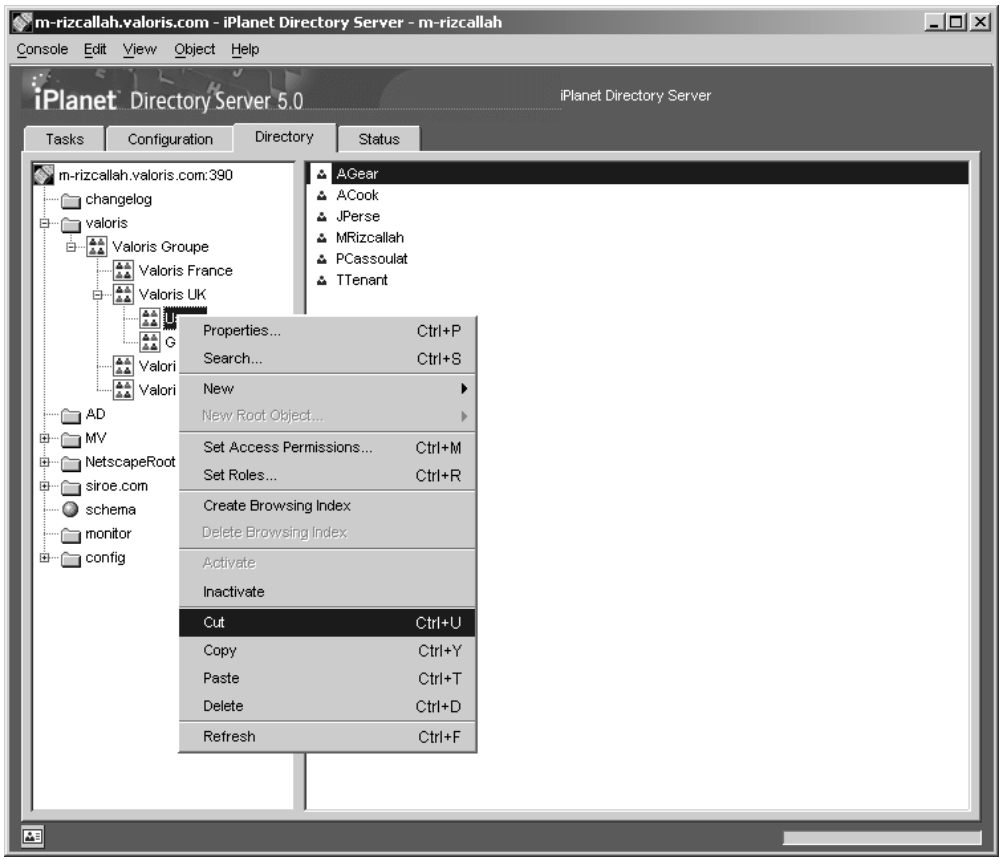


Figure 11.10

Comment déplacer une branche avec la console l'annuaire de Sun

Il est important de savoir que les droits d'accès positionnés sur le nœud déplacé peuvent ne pas être pris en compte lors d'une opération effectuée. En effet, dans la plupart des

serveurs d'annuaires du marché, les règles de sécurité sont décrites dans un attribut attaché au nœud où elles s'appliquent, qui contient les DN de l'objet à qui on attribue les droits, ainsi que le DN de ceux sur qui ces droits s'appliquent.

Par exemple, dans la figure 11.10, on cherche à déplacer la branche `ou=Users` du nœud `ou=Valoris UK` vers le nœud `ou=Valoris France`. Supposons qu'on ait appliqué une règle attribuant les droits d'accès en écriture à cette branche à l'utilisateur identifié par `cn=MRizcallah`. Cette règle, contenue dans le paramètre `aci` de l'objet `ou=Users`, aura la syntaxe suivante avec le serveur Sun Java System Directory Server :

```
(targetattr = "*") (target = "ldap:///ou=Users, ou=Valoris UK, ou=Valoris
Groupe,dc=valoris,dc=com") (version 3.0; acl "Read and Write access"; allow
(all) (userdn = "ldap:///uid=MRizcallah,ou=Users,ou=Valoris UK,ou=Valoris
Groupe,dc=valoris,dc=com");)
```

On remarque dans cette règle que les paramètres `target` et `userdn` contiennent respectivement le DN des objets sur lesquels les droits s'appliquent et à qui ils sont attribués. Lorsque la branche sera déplacée, les DN des objets changeront (Valoris UK est remplacé par Valoris France), et donc la règle ne sera plus valable. Si vous utilisez la console d'administration de Sun Java System Directory Server pour déplacer le nœud, la règle est supprimée. Il faudra donc la recréer manuellement par la suite.

Supprimer une branche

Pour supprimer un nœud de l'arbre, il est nécessaire au préalable de supprimer tous les objets sous-jacents à celui-ci. La plupart des consoles d'administration des annuaires offrent cette possibilité en standard.

Les règles de sécurité sont aussi détruites puisqu'elles sont décrites dans des attributs rattachés à chacun des objets supprimés. Les règles de sécurité, héritées des nœuds supérieurs, restent bien entendu.

Synchroniser les données

Une des principales difficultés dans la vie d'un annuaire d'entreprise est la mise en cohérence des données qu'il contient avec les applications et les autres annuaires de l'entreprise. Il est malheureusement rare de pouvoir fédérer tous les annuaires de l'entreprise et toutes les sources de données relatives aux personnes dans un annuaire unique. En effet, les applications existantes qui ont leur propre base de données ou annuaire ne sont généralement pas modifiées pour s'appuyer sur un annuaire centralisé. La solution préférée des entreprises est la synchronisation des données, car elle peut être faite sans modifier les applications existantes et sans en arrêter la production.

En quoi consiste la synchronisation des données ? Elle contient essentiellement quatre étapes :

- La première étape consiste à *extraire* les informations à synchroniser d'une source de données comme l'annuaire, une base de données ou une application.

- La deuxième étape consiste à *transformer* les données à échanger afin de les mettre dans un format compréhensible par le système destinataire, mais aussi afin de transformer les données elle-mêmes pour les adapter au modèle de données système cible.
- La troisième étape consiste à *transporter* les données du système source vers le système destinataire.
- La quatrième étape consiste à *importer* les données dans le système destinataire.

Il existe plusieurs méthodes pour effectuer la synchronisation des données que nous allons décrire ci-dessous.

L'import et export de fichiers

La plus évidente est celle qui s'appuie sur l'import et l'export de fichiers entre l'annuaire et les systèmes périphériques. Le format de ces fichiers peut s'appuyer sur LDIF, propre au standard LDAP, dont l'avantage est d'être lu par la majorité des serveurs. Mais l'inconvénient est qu'il n'est pas compatible avec les bases de données. Il n'est pas possible d'importer un fichier LDIF dans une base de données de type Oracle par exemple. Pour cela, il faut développer une application qui va traiter ce fichier et transformer les requêtes en langage SQL. De même, une extraction d'une base de données peut se faire au format CSV par exemple, mais il faudra transformer celui-ci au format LDIF avant de l'importer dans l'annuaire.

Une fois extraites du système source, les données peuvent être transportées vers le système destinataire par transfert de fichier FTP ou à l'aide de la messagerie électronique, par exemple.

Il est également possible d'utiliser le format DSML, dérivé de XML, comme format pivot. Ce format n'étant pas compatible directement avec les serveurs d'annuaire, il faudra utiliser un outil en amont chargé d'interpréter ce format et de dialoguer avec l'annuaire à l'aide du protocole LDAP, comme DSML Services for Windows de Microsoft ou DSML for eDirectory de Novell.

L'avantage de cette solution est qu'elle est simple à mettre en œuvre lorsqu'il s'agit de la synchronisation entre annuaires LDAP. L'import et l'export de données de l'annuaire est une fonction offerte par la plupart des logiciels du marché, et le transport des fichiers peut se faire par messagerie, par FTP ou tout simplement par disquette (à condition qu'ils ne soient pas trop fréquents !).

En revanche, si tel n'est pas le cas, elle nécessite la réalisation de programmes complexes afin de transformer des formats de fichiers différents et d'exécuter l'import/export de façon automatique si besoin.

Mais son principal inconvénient réside dans le fait que le chargement des données de l'annuaire peut prendre un temps considérable dans le cas d'un nombre élevé d'entrées et lorsqu'il faut effectuer des traitements spécifiques sur les données (quelques heures pour plus de 500 000 entrées par exemple), les serveurs d'annuaires étant généralement plus lents en écriture qu'en lecture.

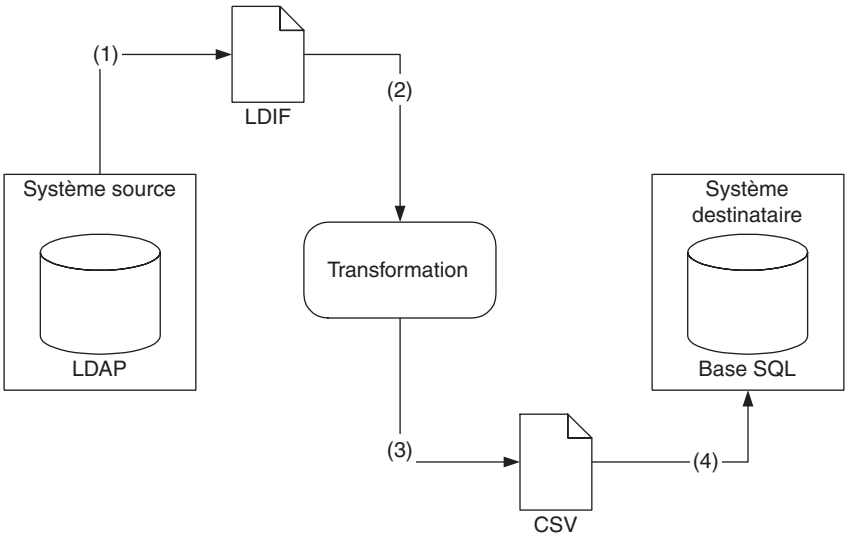


Figure 11.11
Exemple de synchronisation par fichiers

Pour pallier cet inconvénient, il sera nécessaire de redimensionner la machine hébergeant le serveur, ce qui peut s'avérer coûteux, ou de développer des modules n'exportant que les entrées récemment modifiées des applications sources et de l'annuaire. La maintenance de tels modules peut se montrer complexe et avoir un impact non négligeable sur les applications sources, notamment lorsqu'il s'agit de progiciels du marché comme des PGI (Progiciels de gestion intégrés ou ERP comme SAP ou HR Access d'IBM).

Les modules développés de façon spécifique

C'est la méthode la plus utilisée à ce jour par les entreprises qui démarrent la mise en œuvre de leur annuaire. Elle consiste à développer une application spécifique, en Java ou en Visual Basic par exemple, qui d'une part va se connecter à l'annuaire LDAP *via* le protocole associé, et d'autre part va accéder au système à synchroniser *via* un protocole pris en charge par celui-ci, comme SQL pour une base de données. L'extraction, l'importation et la transformation sont programmées de façon spécifique dans cette application.

L'avantage de cette solution est qu'elle peut sembler peu coûteuse et rapide à mettre à œuvre. Son inconvénient vient essentiellement des coûts élevés requis par cette maintenance, notamment lorsqu'il faut synchroniser plus de deux systèmes, et qu'il ne faut tenir compte que des changements survenus depuis la dernière synchronisation afin d'optimiser les temps de traitements.

Les modules peuvent rapidement devenir complexes, et leur nombre peut s'accroître très rapidement faute de coordination et de communication entre leurs concepteurs.

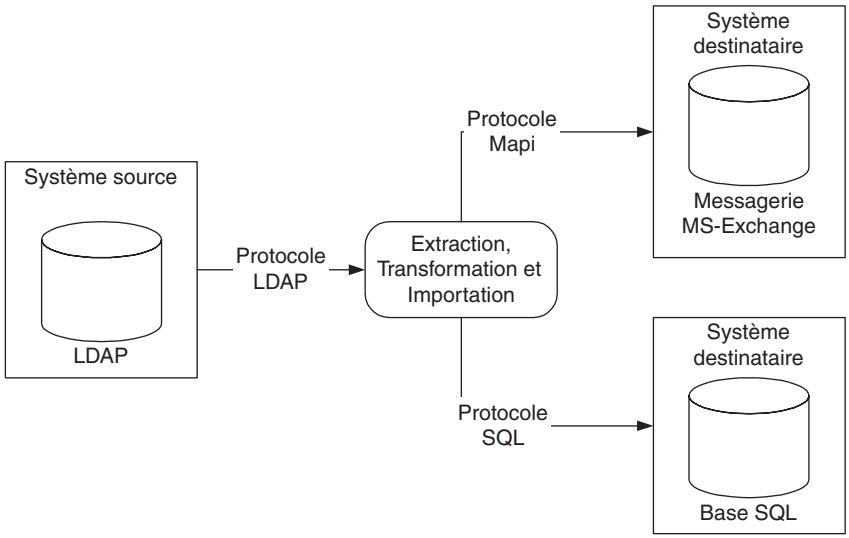


Figure 11.12

Exemple de synchronisation à l'aide d'une application spécifique

Méta-annuaire et EAI

Cette méthode consiste à utiliser un outil de synchronisation du marché dédié à cet effet. Il existe deux catégories d'outils qui peuvent répondre au besoin de synchronisation : les méta-annuaires et les EAI *(Enterprise Application Integration).* Nous allons décrire les particularités de chacun d'eux afin de savoir dans quel cas il faut utiliser l'un ou l'autre.

Qu'est-ce qu'un méta-annuaire ? Nous avons décrit longuement ce type d'outil dans le chapitre 10 de cet ouvrage. Nous allons simplement en rappeler la définition : c'est un outil spécifique chargé d'agréger dans un annuaire unique et de synchroniser un ensemble de données relatives à l'identité des personnes et des ressources, provenant de différents systèmes d'information de l'entreprise.

Un méta-annuaire contient principalement les composants suivants :

- Un référentiel de données agrégées, basé sur le protocole LDAP, communément nommé *Meta Vue*. Il permet d'avoir une vue unifiée des données provenant de différentes sources à travers le protocole LDAP.
- Des connecteurs à différents types de systèmes dont essentiellement des annuaires comme les annuaires de systèmes de messagerie (MS-Exchange, Lotus Notes, etc.) et les annuaires de systèmes d'exploitation (Windows NT, Windows 2000/2003, Unix, etc.), des bases de données relationnelles (Oracle, MS SQL Server, Sybase, etc.) et des formats de fichiers texte comme le format LDIF, CSV ou XML.
- Un moteur de jointure dont le rôle est d'associer les entrées se trouvant dans différents systèmes, accessibles à travers les connecteurs, à celles de la Meta Vue, et ceci à l'aide de règles qui peuvent être prédéfinies par l'administrateur.

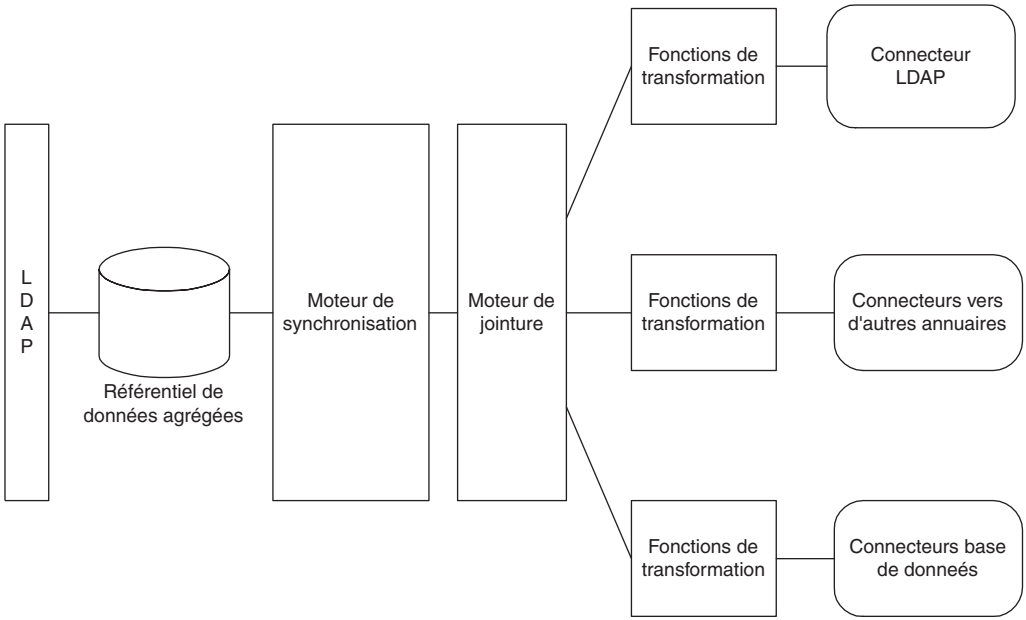


Figure 11.13

Vue d'ensemble des fonctions d'un méta-annuaire

- Des fonctions de transformation qui permettent de traiter ou de calculer certains attributs avant leur mise à jour dans la Meta Vue ou dans l'un des systèmes à synchroniser.
- Des fonctions de gestion de l'appartenance des entrées et des attributs, permettant de résoudre les conflits de mises à jour. Ainsi, si un attribut, comme l'adresse e-mail, se trouve dans différents systèmes à synchroniser, il sera possible de préciser celui qui en sera maître et n'autoriser les mises à jour que vers les autres systèmes.
- Un moteur de synchronisation bidirectionnel s'appuyant sur les fonctionnalités citées précédemment, chargé d'exécuter les synchronisations soit de façon régulière à des périodes prédéfinies par l'administrateur, soit en fonction d'événements survenus dans les différents systèmes à synchroniser, comme l'ajout d'une entrée dans une base de données, le changement de la valeur d'un attribut dans un annuaire LDAP, ou la suppression d'une boîte aux lettres dans un système de messagerie.

Qu'est ce qu'un EAI ? C'est un middleware chargé d'échanger des flux de données et d'exécuter des transactions entre différents systèmes d'information. Son principal usage est l'intégration des applications et des données qui résident dans l'entreprise, soit l'intégration de celles-ci avec celles de partenaires ou fournisseurs (cas du BtoB ou *business to business*). Par exemple, lorsqu'une entreprise industrielle veut intégrer son système avec celui de son fournisseur, un EAI peut être utilisé afin d'intégrer automatiquement le système d'approvisionnement de l'entreprise au système de commande du fournisseur.

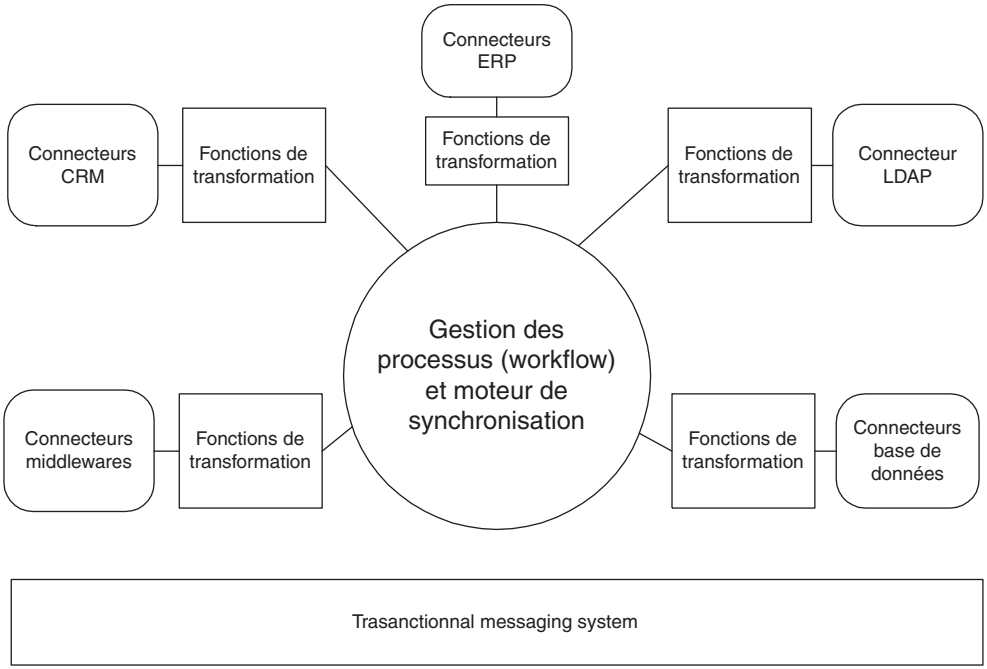


Figure 11.14
Vue d'ensemble des fonctions d'un EAI

Un EAI contient principalement les composants suivants :

- Un système de messagerie transactionnel, chargé de transporter les données et les requêtes de façon fiable et sécurisée, et de les router vers les différents systèmes destinataires de façon asynchrone ou synchrone. Les EAI peuvent avoir leur propre système de messagerie intégré, ou bien ils peuvent s'appuyer sur des technologies du marché comme le modèle COM de Microsoft ou Corba, ou encore utiliser des middlewares comme MQ-Series.
- Des fonctions de transformation qui permettent de traiter ou de calculer les données transportées.
- Des connecteurs à différents types de systèmes, comme des bases de données, des annuaires LDAP, des applications du marché comme des progiciels de gestion intégrés (SAP) ou des progiciels de CRM (Siebel), des plates-formes transactionnelles comme CICS, et autres.
- Des fonctions de gestion des processus métiers et de *workflow,* qui permettent de définir et d'exécuter en séquence les requêtes émises par les différents systèmes. Par exemple, un EAI peut automatiquement exécuter un ordre de vérification du stock chez le fournisseur avant d'exécuter l'ordre de passage de commande.
- Un moteur d'échange bidirectionnel s'appuyant sur les fonctionnalités citées précédemment, chargé d'effectuer les requêtes en fonction d'événements survenus dans les différents systèmes à synchroniser, comme l'ajout d'une entrée dans une base de données.

Voici quelques exemples d'outils disponibles sur le marché considérés comme des EAI : les produits webMethods et SeeBeyond des sociétés de même nom, MQ-Series Integrator de la société IBM, Tibco RendezVous de Reuters, etc.

On voit bien que les méta-annuaires et les EAI ont des vocations très différentes. Les méta-annuaires ont été conçus pour synchroniser des annuaires (LDAP et autres) et des bases de données, alors que les EAI ont été conçus pour échanger des flux d'information entre systèmes et exécuter des transactions à distance. Les EAI peuvent être utilisés pour synchroniser un annuaire LDAP avec d'autres systèmes, s'ils offrent un connecteur LDAP. En revanche, ils ne possèdent en général pas de connecteurs pour d'autres types d'annuaires comme ceux des messageries électroniques ou des systèmes d'exploitation.

Par ailleurs, les méta-annuaires contiennent un référentiel des données agrégées offrant une interface LDAP, alors que les EAI n'en ont pas.

Le tableau suivant résume les fonctions offertes par ces deux catégories d'outils.

Fonctions	Méta-annuaire	EAI
Référentiel de données agrégées offrant une interface LDAP	Oui	Non
Connecteurs à des bases de données	Oui	Oui
Connecteurs à des progiciels (ERP, CRM, etc.)	Non	Oui
Connecteurs à des annuaires (messagerie, système d'exploitation, etc.)	Oui	Non
Moteur de jointure	Oui	Non
Fonctions de transformation	Oui	Oui
Fonctions de gestion des processus métiers et de *workflow*	Non	Oui
Moteur de synchronisation bidirectionnel entre annuaires et bases de données, initiée périodiquement et en fonction d'événements	Oui	Non
Moteur d'échanges bidirectionnel entre systèmes dont des progiciels, des bases de données et des annuaires LDAP, initiés en fonction d'événements	Non	Oui
Système de messagerie transactionnel asynchrone et synchrone	Non	Oui
Gestion de l'appartenance des entrées et des attributs	Oui	Non

Dans quelles circonstances faut-il utiliser le méta-annuaire et dans quelles circonstances faut-il utiliser l'EAI ? Les méta-annuaires ont été conçus pour synchroniser des annuaires entre eux y compris des bases de données. Mais il arrive souvent qu'un EAI soit déjà en place dans l'entreprise. Il est alors légitime de se poser la question de l'achat et du déploiement d'un outil supplémentaire pour le méta-annuaire.

Il faut commencer par faire un état des lieux des besoins de synchronisation afin de mettre en évidence les requis technologiques associés. L'EAI saura répondre à certaines fonctionnalités comme le montre le tableau ci-dessus. Par exemple, il sera en mesure d'effectuer une synchronisation entre une base de données et un annuaire LDAP. Dans tous les cas, si l'on souhaite utiliser un EAI, il faudra au moins construire son propre annuaire LDAP agrégé, car il n'est pas offert par ce type d'outil. Mais l'EAI ne saura pas

effectuer de jointure entre différentes sources de données, et ne saura pas accéder à des annuaires autres que LDAP.

Il faut aussi noter que la plupart des méta-annuaires du marché sont vendus avec un annuaire intégré. Il n'y a donc pas de coût de licences additionnel pour l'annuaire. En revanche, pour un EAI, il faudra acquérir de toute façon les coûts de licence de l'annuaire.

Les étapes à suivre pour ajouter une nouvelle application

Nous allons décrire les différentes étapes à suivre pour ajouter une application à l'annuaire. Cette liste est une sorte de liste de contrôle qui permet de s'assurer que les principaux impacts sur l'annuaire ont été pris en compte.

Identifier les attributs, les classes et les données réutilisables

La première étape, une fois les besoins de l'application bien définis, consiste à identifier les attributs, les classes d'objets et les données qui existent déjà dans l'annuaire d'entreprise et qui peuvent être réutilisées par cette application.

Il est important pour cela d'avoir une description claire de la sémantique de chaque attribut et des classes d'objets afin d'éviter les redondances et le mauvais usage des données. Par exemple, il faut identifier la classe qui permet de décrire les personnes référencées dans l'annuaire. Est-ce la classe standard du serveur d'annuaire comme `inetorgperson` ou `user`, ou est-ce une classe dérivée spécifique ? Y a-t-il des classes différentes pour décrire les clients, les employés et les partenaires ? Comment sont décrites les différentes sociétés auxquelles appartiennent éventuellement ces personnes ? Y a-t-il une branche par société cliente ou par département de l'entreprise contenant l'ensemble des personnes qui s'y trouvent ?

Concevoir et rajouter les nouveaux types de données

Il s'agit ici de comparer les besoins de l'application par rapport à ce qui existe déjà dans l'annuaire. Par exemple, dans le cas d'un extranet concernant les revendeurs, il faudra ajouter une classe d'objet qui permet de les décrire, différente de celle qui permet de décrire les forces de vente internes qui sont des employés.

Une fois les classes d'objets à rajouter identifiées, il faudra trouver pour chacune d'elles la classe dont elle va dériver. Par exemple, pour un extranet revendeurs, on peut créer une classe décrivant *l'entreprise* qui revend et une classe pour chaque *personne* de cette entreprise pouvant accéder à l'annuaire. La première pourra dériver de la classe `organizationunit` et la deuxième de la classe `inetorgperson` ou `user`.

Identifier les acteurs et rajouter les objets associés

Il s'agit ici d'identifier les applications et les utilisateurs qui pourront accéder aux données rajoutées afin de pouvoir contrôler les droits d'accès à celles-ci. Il est toujours possible de donner les droits de l'administrateur à toute nouvelle application pour accéder à la totalité de l'annuaire, aussi bien en lecture qu'en écriture, sans se soucier de la sécurité.

Mais ceci présente des risques importants lorsque l'application est accessible à des personnes externes à l'entreprise *via* Internet.

Nous conseillons donc de créer systématiquement un objet pour chaque nouvelle application, et d'imposer aux développeurs d'utiliser cet objet pour s'identifier à l'annuaire. L'administrateur pourra ainsi contrôler les droits à sa guise ; il pourra même attribuer tous les droits dans un premier temps afin de permettre une mise en œuvre rapide du service, puis les restreindre par la suite aux seules données autorisées.

Il peut aussi y avoir des cas plus complexes où l'on souhaite déléguer la gestion d'un ensemble d'utilisateurs à un utilisateur donné. Par exemple, s'il s'agit d'un extranet revendeurs et que l'on souhaite pouvoir déléguer la gestion des utilisateurs d'un revendeur donné (création, modification, suppression) à un gestionnaire propre à celui-ci. L'objet associé au gestionnaire sera alors créé quand on ajoutera un revendeur dans l'annuaire à l'aide de l'application extranet. Les utilisateurs seront créés lorsque le gestionnaire les ajoutera, toujours à l'aide de l'extranet. Voir l'étude de cas MyPizza dans le chapitre 9 en exemple.

Adapter le DIT

Une fois les acteurs identifiés et les données à rajouter décrites, il faudra probablement adapter l'arborescence de l'annuaire. Nous avons décrit précédemment comment procéder pour rajouter un nœud ou déplacer des données d'une branche à l'autre.

Il est important de tenir compte de l'impact sur les applications existantes en cas de changement de DIT. En effet, certaines applications peuvent utiliser le DN des objets pour les lire, et ne fonctionneront plus si l'objet est déplacé. La seule solution consiste à mettre en paramètres, et non dans le code source, les différents noms de nœuds (ou RDN) utilisés pour calculer un DN.

Identifier et attribuer les droits d'accès des acteurs sur les données

C'est une des étapes les plus délicates de la mise en œuvre d'un annuaire ; il s'agit en effet de trouver le bon compromis entre une gestion complexe et trop rigoureuse des droits, et entre une sécurité trop faible par rapport aux besoins des applications et aux enjeux de l'entreprise.

Voici quelques recommandations à ce sujet :

- Créer un objet dans l'annuaire pour toute nouvelle application, qui sera utilisé pour s'identifier à l'annuaire, comme le montre la figure 11.9.
- Ne pas attribuer les droits d'accès aux données de l'annuaire à cet objet directement, mais passer plutôt par un groupe. En effet, il arrive souvent qu'un ensemble d'applications partage les mêmes droits sur un sous-ensemble de l'annuaire. Par exemple, plusieurs applications extranet destinées aux revendeurs, comme la gestion du réapprovisionnement ou le service clients, peuvent partager des droits sur les objets décrivant les utilisateurs de l'extranet. Le fait d'attribuer les droits à un groupe permettra de les fédérer et d'éviter

de créer autant de règles que d'applications. Voir le chapitre 8 pour plus d'informations sur la gestion des habilitations.

- Créer les groupes et attribuer les droits d'accès à ceux-ci *via* la console d'administration de l'annuaire. Quelquefois, la gestion des groupes doit être accessible aux utilisateurs. Il n'est pas conseillé de leur donner accès à la console d'administration. Vous pouvez alors soit développer une application spécifique à cet effet, soit utiliser des produits du marché offrant ce type de service comme avec COREid d'Oblix.
- Éviter la création d'un nombre important de groupes et la gestion complexe de groupes imbriqués. En fonction de la complexité des applications, le nombre de groupes peut devenir rapidement très élevé, voire atteindre et même dépasser le nombre d'utilisateurs dans l'annuaire ! L'utilisation de groupes imbriqués réduit le nombre de groupes, mais peut nuire à la visibilité du contenu des groupes. Il est aussi conseillé d'utiliser des groupes dynamiques soit des ACL utilisant des attributs indirects (voir l'étude de cas MyPizza dans le chapitre 9). Les groupes dynamiques sont une facilité offerte par les récents serveurs d'annuaire. On les constitue en désignant un filtre LDAP au lieu d'une liste de membres désignés par leur DN. La requête est exécutée dynamiquement lorsqu'on cherche à en lire le contenu. Voir le chapitre 10 pour plus d'informations sur les groupes dynamiques.
- Mettre à jour la documentation aussi bien pour les attributs, les classes d'objets que les droits rajoutés.

Mesurer l'impact sur la volumétrie et les performances

L'ajout d'une nouvelle application peut avoir des répercussions sur les volumes de données gérés par l'annuaire. Par exemple, un extranet clients va nécessiter le rajout de l'ensemble des profils clients dans l'annuaire. Ceci peut s'élever à quelques millions d'enregistrements s'il s'agit d'un produit grand public. Ainsi, un opérateur de téléphonie mobile qui souhaite offrir à ses clients la possibilité de gérer leur abonnement en ligne devra prendre en charge autant d'entrées dans son annuaire que d'abonnements en cours, soit plusieurs millions.

La première chose à faire est de s'assurer que la volumétrie ne va pas dégrader les performances, même si peu d'utilisateurs se connectent à l'annuaire. Ceci dépend essentiellement de l'outil utilisé pour l'annuaire et de la plate-forme sur laquelle il fonctionne. Ceux qui en ont les moyens pourront d'emblé s'équiper d'une plate-forme performante et évolutive comme une machine Sun E10K ou F15K, pouvant supporter jusqu'à soixante-quatre processeurs et équipée de disques durs rapides. Les autres devront s'appuyer sur des topologies permettant de répartir les volumes de données sur plusieurs machines. Ceci peut être fait soit en répliquant l'annuaire sur différentes machines et en utilisant des outils de répartitions de charge, soit en répartissant les branches de l'annuaire sur plusieurs machines reliées par des mécanismes de renvoi de référence. Voir le chapitre 8 pour plus de détails sur ces différents cas.

Ensuite, il faut mesurer l'impact du nombre de requêtes générées par la nouvelle application. Il est difficile de prévoir ce nombre avant la mise en production de l'application, mais il

est aussi risqué de ne rien faire et de mettre en production celle-ci sans s'assurer de la qualité de service offerte aux utilisateurs. Cela nécessite la mise en œuvre des outils de simulation qui permettent de générer un nombre élevé de requêtes, ainsi que des outils d'audit qui permettent de mesurer les temps de réponse de l'annuaire. Pour plus d'informations à ce sujet, voir le paragraphe sur l'optimisation et les tests de performance plus haut dans ce chapitre.

Étudier les besoins de synchronisation

Il s'agit ici d'identifier les systèmes avec lesquels il faudra mettre en place des processus de synchronisation de données. En effet, une fois les types de données requis par la nouvelle application identifiés, il est important de vérifier s'il existe d'autres systèmes dans l'entreprise possédant les mêmes informations. Par exemple, si l'on ajoute une nouvelle application qui nécessite le rajout dans l'annuaire de l'adresse de messagerie des clients, il faudra probablement assurer la cohérence entre celui-ci et les systèmes de gestion de la relation client qui contiennent aussi l'adresse de messagerie.

Si c'est le cas, il faudra décrire les données à synchroniser, mettre en évidence leur appartenance, et décrire les processus et les outils permettant d'assurer la synchronisation. Nous avons présenté les différentes solutions possibles précédemment dans ce chapitre.

Documenter et communiquer

Un des facteurs clés du succès d'un annuaire d'entreprise est que son contenu soit bien documenté. En effet, c'est ce qui va permettre à de nouvelles applications de réutiliser celui-ci. Il est aussi important de documenter les droits d'accès attribués sur ces données.

Enfin, une fois la documentation mise à jour, il est utile de communiquer sur les nouveaux services mis en place. Un espace « annuaire » sur l'intranet de l'entreprise, destiné aux concepteurs d'applications, est un atout indispensable à son cycle de vie.

Le cadre légal

Plusieurs de mes lecteurs m'ont demandé de décrire la façon dont les annuaires s'inscrivent dans le cadre légal relatif à la protection des données sur les individus. Notons que ceci n'est pas propre aux annuaires : toute base de données contenant des informations sur des personnes doit respecter les lois relatives à l'informatique et aux libertés individuelles du pays où elle se trouve.

Mais quelles sont ces règles ou ces lois et que faut-il faire dans les annuaires d'entreprise pour les respecter ?

Droits et obligations

Chaque pays dans le monde a défini un ensemble de lois qui permettent de protéger les libertés individuelles relatives à l'usage des données informatiques. Des organismes comme la CNIL (Commission nationale de l'informatique et des libertés, www.cnil.fr)

en France et le National Telecommunications & Information Administration aux États-Unis sont chargés d'édicter ces lois. Elles sont généralement constituées de droits pour les individus et d'obligations pour les acteurs (les organismes ou personnes qui collectent et traitent les informations sur les individus). On retrouve les droits des individus dans la plupart des législations sur la protection des données personnelles en Europe et dans le monde. La Convention n° 108 du Conseil de l'Europe, ratifiée par de nombreux États dont la France, les a consacrés au plan international en 1981.

Les droits des individus

Face aux potentialités quasi infinies qui résultent des technologies de l'information, la loi « Informatique et libertés » du 6 janvier 1978 en France a prévu de solides garde-fous pour protéger l'individu des dangers liés à la multiplication des données informatiques le concernant. Cette loi n'interdit pas la création de fichiers nominatifs. Ce n'est pas un outil de lutte contre l'informatique, bien au contraire, c'est un moyen d'en réglementer l'usage afin d'en limiter les effets liberticides.

La loi du 6 janvier 1978 reconnaît essentiellement sept droits aux personnes :

- *Le droit à l'information préalable* : les fichiers ne doivent pas être créés à l'insu des individus concernés. Ceux qui créent des traitements ne doivent pas laisser les individus dans l'ignorance de l'utilisation qu'ils vont faire de ces données. Sinon, la loi « Informatique et libertés » est purement et simplement violée.
- *Le droit de curiosité* : pour pouvoir accéder aux données qui concernent les personnes, chaque personne a le droit de demander à tout organisme s'il détient des informations sur elle.
- *Le droit d'accès direct* : toute personne peut obtenir la communication des informations qui la concernent en les demandant directement à l'organisme qui détient le fichier dans lequel elle figure. C'est un droit fondamental qu'il ne faut pas hésiter à exercer.
- *Le droit d'accès indirect* : pour certaines données nominatives, la loi prévoit un intermédiaire entre l'individu et l'organisme qui détient le traitement. Par exemple, pour les données médicales, un médecin de votre choix, pour les données figurant dans des traitements intéressant la sûreté de l'Etat, la défense et la sécurité publique, un commissaire de la CNIL.
- *Le droit de rectification* : si vous avez constaté des erreurs lorsque l'organisme qui détient le fichier vous a communiqué les données vous concernant, vous pouvez les faire corriger. La loi va même plus loin puisqu'elle oblige l'organisme à rectifier d'office et de lui-même les informations dès lors qu'il a connaissance de leur inexactitude.
- *Le droit d'opposition* : si vous avez des raisons légitimes pour ne pas figurer dans tel ou tel fichier, vous pouvez vous opposer à votre fichage. La loi garantit un droit d'opposition que l'on peut exercer au moment de la collecte ou plus tard, en demandant par exemple la radiation des données contenues dans les fichiers commerciaux. Bien sûr, ce droit ne s'applique qu'aux fichiers qui n'ont pas été rendus obligatoires par une loi.

- *Le droit à l'oubli* : l'informatique permet de conserver indéfiniment les données personnelles. La loi a donc prévu un droit à l'oubli, afin que les personnes ne soient pas marquées à vie par tel ou tel événement.

Le non-respect de ces droits par les acteurs responsables des données lorsque vous souhaitez les exercer est le plus souvent sanctionné pénalement. Il est possible de porter plainte auprès du procureur de la République, ou plus simplement auprès de la CNIL en France, par simple courrier, en vue d'un règlement amiable entre les parties.

Les obligations de ceux qui collectent et traitent les données sur les individus

Ceux qui collectent et traitent les données sur les individus doivent respecter les droits cités précédemment et ont des obligations, comme :

- s'assurer que le traitement ne fait pas l'objet d'un détournement de finalité ;
- ne pas substituer l'ordinateur à l'homme pour la prise de décision ;
- s'assurer que la collecte des informations n'est ni frauduleuse, ni déloyale, ni illicite et qu'elle s'accompagne d'une bonne information des personnes ;
- s'assurer que les informations sensibles (nationalités, opinions politiques, philosophiques ou religieuses, mœurs et condamnations pénales) éventuellement recueillies le sont conformément à la loi, que le numéro de Sécurité sociale n'est pas utilisé sans autorisation ;
- que les informations ne sont pas conservées au-delà de la durée prévue, qu'elles sont bien mises à jour lorsqu'elles sont périmées, et que les tiers qui auraient pu y avoir accès ont bien été informés de cette mise à jour ;
- que les traitements font l'objet d'une sécurité optimale, afin qu'aucun détournement ne puisse avoir lieu ;
- que les informations ne sont pas communiquées à des personnes non autorisées ;
- que la commercialisation éventuelle de ces données se réalise bien dans le cadre légal ;
- que l'établissement de flux transfrontaliers de données est bien conforme au droit du pays en vigueur où se trouvent les données.

La déclaration auprès de la CNIL

Afin de s'assurer du respect des droits et des obligations associées aux libertés individuelles, certains organismes nationaux, comme la CNIL en France, obligent les entreprises à déclarer les informations qu'elles récoltent sur les individus. Il suffit pour cela de remplir des formulaires que l'on trouve généralement sur leurs sites Internet, et de les renvoyer signés.

Cette déclaration permet au responsable de la base de données de communiquer à l'organisme ses intentions : quelle sera la finalité de la base données, quelles informations vont être enregistrées, pendant combien de temps, qui y aura accès, à quel service les personnes peuvent-elles s'adresser pour exercer leur droit d'accès, etc. La consultation du « fichier des fichiers » ainsi constitué par l'organisme permet aux personnes fichées de

prendre connaissance de ces informations, et de répondre facilement à la question : « telle société, telle administration me fiche, pour quoi faire ? ».

Ainsi les personnes référencées peuvent-elles s'adresser à l'organisme de contrôle et non pas à l'entreprise en cas de problème, ce qui simplifie les démarches administratives associées. En outre, cela permet à l'entreprise de subir des contrôles adéquats pour garantir au tiers la conformité de son système avec aux lois en vigueur.

Le non-accomplissement de ces formalités est sanctionné pénalement en France (article 226-16 du code pénal) : trois ans d'emprisonnement et 45 000 euros d'amende.

Les différentes étapes à suivre

La CNIL (*www.cnil.fr*) publie sur son site des documents décrivant les différentes étapes à suivre pour monter votre site Internet. Vous trouverez en annexe de cet ouvrage un document, intitulé « Je monte un site Internet » issu de ce site et détaillant point par point ce qu'il faut faire. Nous vous conseillons de consulter régulièrement le site de la CNIL afin de vous assurer d'avoir la dernière version de ce type de document.

Impact sur les annuaires d'entreprise

Quels sont les impacts des droits et obligations cités précédemment sur les annuaires LDAP ?

Les obligations concernant la protection des données sur les individus afin d'éviter les détournements et la communication d'informations à des personnes non autorisées, nécessitent la mise en place d'un coupe-feu (ou *firewall)* pour la protection de l'annuaire, d'un mécanisme de contrôle de la sécurité qui soit compatible avec la stratégie adoptée, comme la gestion des habilitations (ou ACL) dans l'annuaire, soit un contrôle dans toutes les applications accédant à celui-ci. Par exemple, si l'annuaire contient le numéro de Sécurité sociale des personnes, il faut s'assurer que celui-ci n'est accessible que par l'individu ou par l'administrateur de l'annuaire, mais en aucun cas par les autres individus référencés dans l'annuaire. La mise en place d'un coupe-feu n'est pas suffisante, car il autorise (ou non) l'accès à l'annuaire, mais ne permet pas de différencier les données accessibles par deux personnes différentes.

L'exemple à ne pas suivre est de donner à toute application accédant à l'annuaire les droits de l'administrateur. En effet, certaines applications peuvent demander l'identifiant et le mot de passe à l'utilisateur afin de vérifier son identité, mais elles utilisent l'identifiant d'un administrateur pour accéder à l'annuaire. Le risque encouru dans ce cas est que la protection des données ne dépende plus de la gestion des habilitations (ou ACL) de l'annuaire, mais du développeur de l'application. Voir dans le chapitre 10, le paragraphe sur les serveurs d'applications pour plus d'informations sur ce sujet.

Ces obligations nécessitent aussi de s'assurer que le serveur d'annuaire est protégé dans un local accessible uniquement aux personnes autorisées, et d'autre part que seules ces personnes possèdent le mot de passe de l'administrateur. S'il s'agit de l'hébergement du serveur chez un prestataire de services, comme un ISP *(Internet Service Provider)*, il est important de vérifier que les mêmes règles sont bien appliquées.

Les droits d'accès aux informations, comme les droits d'accès direct ou indirect et le droit de curiosité concernant les individus nécessitent que l'administrateur de l'annuaire dispose d'un outil permettant d'extraire toutes les informations concernant une personne donnée. Il ne s'agit pas ici uniquement de l'objet `person` et des classes qui en dérivent, mais aussi de tout autre objet rattaché à celui-ci. Ceci peut être intégré dans une application accessible à travers le réseau Internet, soit faire partie d'un outil d'administration interne à l'entreprise. Dans ce dernier cas, il doit être possible d'imprimer ou d'exporter les données dans un fichier texte afin de pouvoir les remettre à l'individu.

Le droit de rectification doit permettre à l'individu de modifier les données le concernant. Le moyen le plus efficace est de permettre à l'individu de modifier lui-même ces informations à travers un site Internet ou un intranet. Ceci nécessite bien entendu une identification préalable de la personne. Dans certaines circonstances, les modifications ne pourront pas prendre effet immédiatement. Par exemple, si l'utilisateur modifie son adresse de facturation sur le site Internet d'un commerçant, il faudra que celle-ci soit vérifiée, c'est-à-dire que la nouvelle adresse doit effectivement exister, puis il faudra qu'elle soit prise en compte par le système de facturation qui en général ne s'appuie pas sur l'annuaire LDAP. Si l'on veut automatiser l'enchaînement de ces différentes étapes, on pourra utiliser des outils intégrant des fonctions de workflow, comme les outils de e-provisioning que nous avons cités dans le chapitre 10. Sinon, il faudra transmettre la demande de l'utilisateur à un administrateur qui se chargera de l'exécuter manuellement.

Le droit à l'oubli et le droit à l'opposition, lorsqu'ils sont exercés, doivent normalement détruire tout enregistrement concernant l'individu dans l'annuaire. Or il arrive souvent que les entreprises invalident l'entrée correspondante plutôt que de la détruire, car ceci leur permet d'effectuer des statistiques *a posteriori* et de conserver un historique des données. Mais il doit quand même être possible de supprimer définitivement une entrée afin de respecter ces droits.