Chapitre 12
Implémentation de la haute disponibilité

1. Introduction

Les services AD LDS ont permis aux entreprises d'implémenter facilement et rapidement des annuaires LDAP qui reposent sur les mêmes fondamentaux que l'Active Directory sans encourir de risques de corruption du schéma.

Au fur et à mesure que les usages AD LDS se généralisent, la demande pour faire évoluer leur implémentation et garantir des niveaux de disponibilité plus élevés augmente. Le code source de l'AD LDS est identique à celui de l'AD DS. Les deux disposent du même moteur de réplication et donc des mêmes caractéristiques de performance, sauf que les mécanismes de haute disponibilité sont différents.

Dans un environnement Active Directory, les opérations d'équilibrage de charge se déroulent en arrière-plan, grâce au processus de découverte automatique, qui permet de localiser les contrôleurs de domaine. Ce n'est pas le cas pour les services AD LDS.

AD LDS n'est pas à l'abri d'une interruption de service, il est donc important d'étudier l'intégration de cette solution à une topologie adaptée au service qui doit être rendu. Pour ce faire, Microsoft propose NLB (*Network Load Balancing*), qui permet d'accroître la disponibilité et l'évolution d'une architecture d'annuaire.

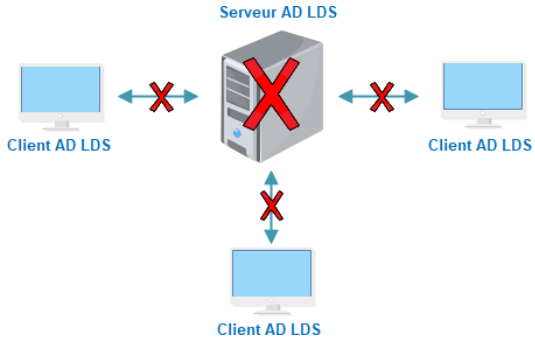
Ce chapitre commencera par exposer les risques d'une architecture centralisée puis présentera l'infrastructure à équilibrage de charge fondée sur la technique NLB de Microsoft Windows Server. Il abordera également la mise en pratique d'un cas classique de déploiement des serveurs AD LDS à équilibrage de charge.

2. Les risques d'une architecture centralisée

Aujourd'hui, les entreprises sont confrontées à de nouveaux challenges dans la gestion de leurs systèmes informatiques. Les utilisateurs sont de plus en plus exigeants, ils s'attendent à ce que les services soient toujours disponibles et surtout plus performants.

Cependant, il peut arriver que ces services s'interrompent de temps à autre, mais cela ne devra porter atteinte ni au système d'information ni à l'activité de l'entreprise.

Une infrastructure qui repose sur une architecture centralisée paraissait peut-être une bonne idée à l'époque – n'avoir en général à maintenir qu'un seul système facilitait le travail des administrateurs lors du déploiement et de l'administration des applications.



À l'heure actuelle, ce type d'infrastructure soulève plusieurs difficultés.

Tout d'abord, avoir un point unique de défaillance (SPOF - *Single Point Of Failure*) présente un risque potentiel pour les organisations. Si le serveur central tombe en panne, les services qui lui sont associés ne peuvent plus traiter les demandes des utilisateurs.

Deuxièmement, étant donné que toutes les applications sont gérées par un seul serveur, la scalabilité reste limitée. La seule façon de faire évoluer l'infrastructure est d'ajouter plus de mémoire, plus de capacité de stockage et donc plus de puissance de traitement. Cela peut ne pas s'avérer être une solution rentable à long terme.

Enfin, la performance des applications peut être affectée par un manque de bande passante. Un seul serveur peut rapidement devenir le goulot d'étranglement d'une architecture qui implique plus de trafic. De ce fait, il serait difficile de faire face à l'afflux massif d'utilisateurs.

La haute disponibilité (HA, pour *High Availability*) reste la solution idéale pour protéger les systèmes contre les interruptions de service et la perte de données. Cela consiste à mettre en place des serveurs qui agissent comme un système unique et qui fournissent une disponibilité continue.

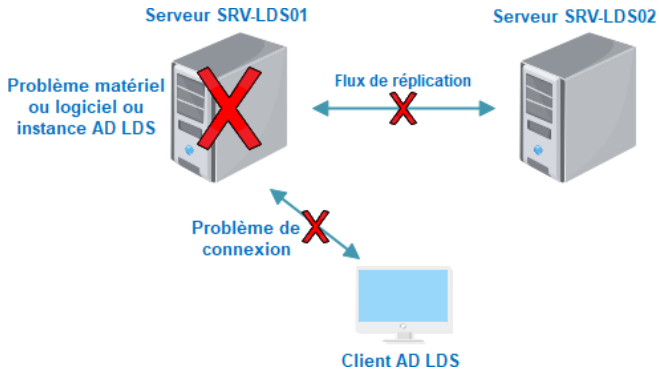
Il existe deux techniques permettant de mettre en œuvre la haute disponibilité : l'équilibrage de charge réseau (NLB, pour *Network Load Balancing*), qui permet d'équilibrer le trafic entrant sur plusieurs serveurs, et le cluster à basculement, qui consiste à surveiller en permanence l'application afin de garantir qu'elle soit toujours disponible.

Seule la technique d'équilibrage de charge réseau peut être utilisée dans une infrastructure AD LDS. La technique du cluster à basculement n'est pas abordée dans ce chapitre.

3. La haute disponibilité de l'annuaire AD LDS

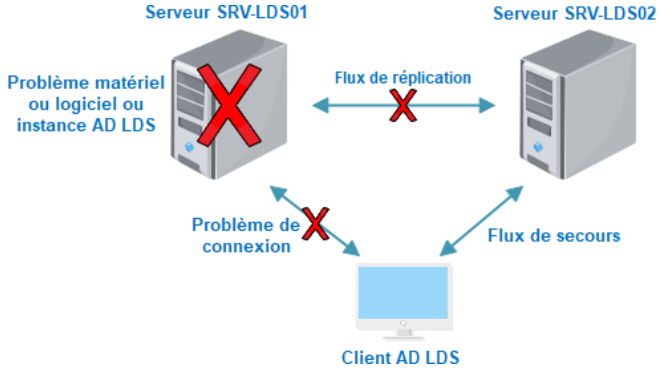
Le fait de configurer deux serveurs AD LDS en réplication "jeu de configuration" permet d'assurer la redondance de l'instance et des données de l'annuaire. Cependant, le serveur AD LDS ne dispose d'aucune fonction de redirection du service LDAP si un des deux annuaires tombe ou devient défaillant.

Il convient donc de déléguer la vérification de l'état de santé du serveur AD LDS à l'agent (ou utilisateur) ou bien à un équipement réseau entre le serveur AD LDS et ses utilisateurs. Ce scénario est illustré à la figure suivante, où l'utilisateur se trouve dans l'incapacité de communiquer avec le serveur SRV-LDS01.



3.1 Configuration de plusieurs serveurs AD LDS

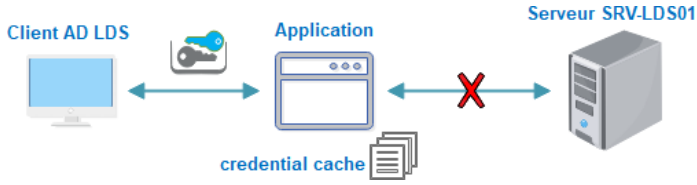
Pour régler le problème de l'exemple précèdent, nous pouvons spécifier l'existence d'un serveur redondant aux clients AD LDS. Par exemple, les applications ayant une adhérence avec l'annuaire AD LDS permettent la configuration de plusieurs services d'annuaire dans leurs fichiers de configuration. De ce fait, si l'utilisateur n'arrive plus à accéder à son serveur, il contactera (après plusieurs tentatives) le deuxième serveur d'annuaire de sa liste et ainsi de suite. Une fois la communication rétablie, l'utilisateur reprend la main sur l'application.



3.2 Configuration de la mise en cache AD LDS

Il est également possible d'améliorer la fonction de haute disponibilité en configurant le mode de connexion hors ligne sur les applications ayant une adhérence avec l'annuaire AD LDS. Ainsi, l'authentification reste tout de même possible même si le serveur est injoignable.

En général, la mise en cache des informations d'identification (*cached credentials*) implique l'utilisation des API qui retiennent les réponses du serveur AD LDS lors d'une tentative de connexion réussie.

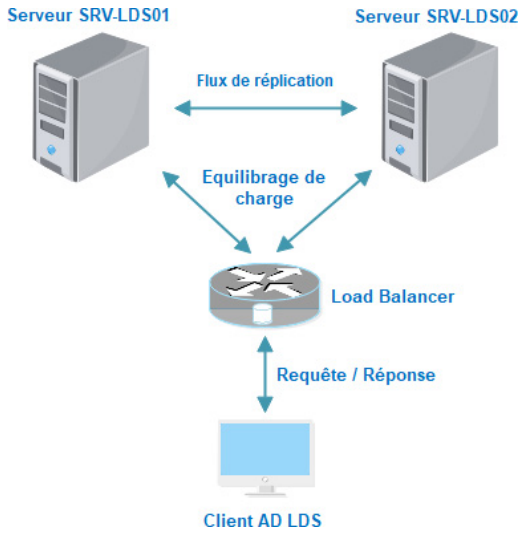


Ce mode d'authentification doit être utilisé avec précaution. Certes, il est très pratique pour son aspect fonctionnel, mais présente des risques pour la sécurité des annuaires, puisqu'il peut arriver que le compte d'un employé ayant quitté l'entreprise soit toujours utilisable, bien qu'effacé de l'annuaire AD LDS. Il est important, donc, de paramétrer les applications pour qu'elles suppriment le cache, par exemple toutes les semaines.

3.3 Mise en place d'une solution d'équilibrage de charge

Certaines applications peuvent rencontrer des difficultés pour définir plusieurs serveurs d'annuaire AD LDS dans leur configuration… et dans ce scénario, il faudra alors une autre solution (logicielle ou matérielle) qui se chargera de l'interception des requêtes LDAP et de leur acheminement vers une instance active de l'infrastructure AD LDS.

Cette option de haute disponibilité fournit également la possibilité de faire appel à la fonction d'équilibrage de charge, permettant ainsi de distribuer les requêtes LDAP sur de multiples serveurs AD LDS, ou bien même, dans une configuration avancée, de surveiller leur disponibilité.



On voit, dans la figure ci-dessus, que l'utilisateur ne peut s'adresser qu'à un seul serveur AD LDS et ceci est vrai dans cette configuration. Cependant, l'adresse IP ou l'alias défini est celle/celui d'un équilibreur de charge qui joue le rôle d'un agent de circulation qui achemine les demandes des utilisateurs vers les serveurs AD LDS de manière équilibrée, ou unilatérale si l'un d'entre eux est injoignable.

4. L'équilibrage de charge

L'équilibrage de charge est une méthode permettant d'améliorer la disponibilité et les performances des applications. Elle vise à répartir le traitement et le trafic de manière égale sur plusieurs serveurs d'un groupe, en veillant à ce qu'aucun d'entre eux ne soit submergé. Cette méthode est employée dans plusieurs domaines, en particulier avec les serveurs web frontaux qui interceptent les requêtes http émises par les utilisateurs et qui nécessitent de la persistance ou du traitement.