

L051

Administration des bases de données et systèmes d'exploitation

Médian : Administration des systèmes d'exploitation

CORRIGÉ

Conditions de réalisation :

Pas de documents autorisés, Pas de calculatrice

Durée : 2h

Répondre sur les copies d'examen, en précisant le numéro de la question.

Rédacteurs : Hélène Béhague, François Beuraud
Toute reproduction interdite sans autorisation écrite

Remarques générales :

Vous devriez plus vous appliquer à répondre aux questions posées. Par exemple sur la question 7 qui ?, quand ? ont souvent été ignorées.

Vous devriez lire les questions. Par exemple sur la question 11, où il est question de démarrage de l'ordinateur, plusieurs ont parlé de l'utilisateur loggué.

1) Rappelez la méthode à utiliser pour rattacher une machine windows seven à un domaine Active Directory. Quel utilisateur a le droit de faire cette opération ? Y a-t-il lieu de mettre en place des restrictions à ce sujet ?

2 pts - 5 lignes max

Panneau de configuration > Système et sécurité > Système > Modifier les paramètres...

Par défaut, tout utilisateur du domaine a le droit d'introduire des machines dans le domaine. Dans le cadre d'une politique sérieuse d'admin système, ce droit doit être limité aux administrateurs du domaine ou à quelques utilisateurs (staff technique). Ceci se paramètre par les GPO.

2) Sur un domaine analogue à celui vu en TP, un de vos utilisateurs se logue sur un poste ubuntu et reçoit le message "No home directory - logging with HOME = /". Que faites-vous ?

2 pts - 15 lignes max

Cela signifie que son répertoire d'accueil n'est pas monté.

Points à vérifier :

Si le problème se produit pour les autres utilisateurs, sur d'autres postes, il y a sans doute un problème du côté du serveur de fichiers. Celui-ci est-il en service, la partition est-elle accessible sur le serveur, et l'export est-il actif (commande exports pour vérifier) ?

Si le problème est limité à ce poste, il y a sans doute un problème du côté de la configuration de ce poste : le montage nfs est-il effectif, vérifier si, en tant que root, on peut avoir accès à la partition, si les droits sont corrects (l'utilisateur peut écrire). Vérifier aussi que ldap est correctement configuré ; le login s'est bien passé, donc a priori le ldap est bien configuré concernant le login, mais il peut y avoir un problème sur la configuration concernant la session.

Si le problème est limité à cet utilisateur, il y a sans doute un problème du côté de sa configuration. Points à vérifier : Sur le serveur AD, son répertoire d'accueil est-il correctement spécifié. Sur le serveur de fichier, ce répertoire existe-t-il, avec les bons droits d'accès. L'utilisateur est-il connu sur le serveur de fichiers (wbinfo -u) ?

Certains ont parlé de Samba, concernant le partage de fichiers. Non : le partage de fichiers se fait en NFS, entre deux machines linux. Samba pourrait être en cause uniquement par son complément winbind, s'il s'avérait que le serveur de fichiers ne connaisse pas l'utilisateur, puisque notre serveur de fichiers connaît les utilisateurs par l'intermédiaire de winbind.

3) En vous plaçant dans la suite logique du TP, vous voulez remplacer la machine serveur2008 sous windows par des équivalents sous linux. Expliquez, dans les grandes lignes, la façon dont vous allez procéder.

3 pts - 20 lignes max

Un serveur de domaine active directory remplit les rôles de :

serveur LDAP

serveur kerberos

serveur DNS

serveur de temps

Il faut donc remplacer chacun de ces services. Cela peut être effectué sur une seule machine, ou plusieurs. On choisira des machines debian, pour rester sur une distribution connue qui convient parfaitement. On pourrait même utiliser le serveur de fichiers, mais cela est à déconseiller.

On commencera par le serveur de temps, la synchronisation étant indispensable pour un bon fonctionnement général. Puis on installera le serveur DNS, qui est la brique de base de la communication en réseau. Pour ce faire, on devra configurer le fichier de zone, qui est le fichier de configuration du serveur DNS. Ensuite on mettra en place un serveur LDAP, et on terminera par le serveur kerberos.

En ce qui concerne DNS et LDAP, il faudra prendre soin de récupérer les informations existantes, soit les resaisir si elles ne sont pas trop volumineuses, soit effectuer des recherches pour trouver comment procéder (comment générer le DIT de l'AD, et comment le réintégrer dans le nouveau serveur).

Pour nos machines clientes windows, nous ajouterons sur la machine serveur LDAP un serveur Samba qui jouera le rôle de serveur de domaine. En rattachant nos machines à ce domaine, nous bénéficierons de services identiques à ceux fournis par un serveur windows : accès à la liste des utilisateurs et à leur configuration (profil).

Bien entendu, il s'agit d'un travail complexe, à mener à bien étape par étape.

4) Vous tombez face à un stagiaire qui confond profil et répertoire d'accueil. Expliquez lui la différence en quelques phrases.

1 pt - 5 lignes max

Ces notions représentent des données propres à chaque utilisateur. Profils et répertoires d'accueil sont tous deux stockés sur des serveurs de fichiers (qui peuvent être le même, ou pas), et sont, en principe et sauf erreur de configuration, inaccessibles aux autres utilisateurs.

Le profil est un ensemble de données de configuration comme par exemple l'aspect du bureau, l'historique de navigation web, et la partie de la base de registre concernant l'utilisateur. Comme **il transite sur le réseau lors de chaque login/logout**, il est très important de limiter sa taille au strict minimum. Sur le poste de travail, il se stocke sous C:\Documents and Settings. Il n'est pas forcément effacé lors de la déconnexion de l'utilisateur, cela dépend de plusieurs facteurs (la politique d'administration mise en place, l'encombrement du réseau au moment de la déconnexion), ce qui est un problème pour la confidentialité des données.

Le répertoire d'accueil est un espace de stockage personnel beaucoup plus vaste où l'utilisateur peut stocker des données qui ne seront pas transférées vers les disques du poste de travail, les accès se font par le réseau au moment où les données doivent être lues ou écrites.

Il était indispensable de dire que le contenu du profil transite sur le réseau lors de chaque login/logout.

En présentant en cours la notion générale de profil, j'en ai parlé en disant que dans certains contextes (sites web, messageries...) le profil contient des données élémentaires caractérisant l'utilisateur, comme le nom, l'adresse. Il est bien évident que le profil windows (dont il est implicitement question ici) est assez différent, et ne contient pas le nom ou l'adresse.

5) Dans DHCP, que signifie la lettre D ? Qu'est-ce qu'un DHCP statique ? Quel est son intérêt ?

1 pt - 5 lignes max

Le "D" signifie "Dynamic". DHCP = Dynamic Host Configuration Protocol.

Il peut donc sembler paradoxal de parler de DHCP statique.

Le DHCP classique attribue à un poste client une adresse IP disponible prise dans un intervalle défini dans la configuration du serveur DHCP. Le DHCP statique attribue également une adresse IP au poste client, mais une adresse non pas prise au hasard, mais déterminée précisément sur la base de l'adresse MAC du poste client.

Son intérêt est de faciliter l'administration des machines. Au redémarrage, l'adresse IP de chaque machine sera celle définie de façon centralisée dans la configuration du serveur DHCP. Il est inutile de passer sur chaque poste pour vérifier et/ou remettre à jour une adresse IP qui aurait été modifiée par erreur.

On utilise le DHCP statique pour des postes "clients", postes de bureaux, salles de TP. En ce qui concerne les serveurs, on leur attribue généralement une adresse IP fixe (configuration directe sur la machine, sans passer par un DHCP).

6) Dans une commande **mount -t truc machin chose**, quel est le rôle de chacun des mots truc, machin et chose ?

2 pts

L'option -t est suivie d'un mot clé qui indique le type de système de fichier. On peut trouver par exemple ext2, ext3, nfs, vboxsf, vfat... ou beaucoup d'autres. En TP, nous avons vu vboxsf et nfs. vboxsf, dans le cas de machine virtuelle sous virtualbox, indique qu'il s'agit d'un répertoire situé sur la machine hôte. nfs indique qu'il s'agit d'un répertoire accessible via le protocole NFS. Cette option est facultative, par exemple si le second mot commence par un nom suivi de ":", ce nom sera interprété comme nom de serveur, et le type sera par défaut "nfs".

Le second mot indique ce qui doit être monté. Nous en avons vu 2 exemples : ce peut être une chaîne contenant le nom d'un serveur et le répertoire exporté, comme "debian:/home/donnees", ou le nom de partage pour un répertoire partagé entre machine hôte et machines virtuelles.

Le dernier mot est le chemin du point de montage. Le point de montage est un répertoire dans lequel apparaîtra le contenu du répertoire monté. Ce répertoire doit exister et il est préférable qu'il soit vide. Si il n'est pas vide, son contenu sera inaccessible, masqué par le contenu du répertoire monté. Le point de montage peut être situé n'importe où dans l'arborescence, mais, pour une bonne organisation, il est conseillé de le mettre sous /mnt.

7) **tune2fs -o +acl /dev/sd...**

Quel est le rôle de cette commande ?

Qui la lance et quand ?

Equivalent sous windows ?

2 pts

Cette commande sert à doter une partition des extensions pour gérer les acl.

Elle est émise une seule fois dans la vie de la partition, en principe juste après un formatage. Bien entendu, seul un administrateur système (le compte root ou un sudoer) peut l'émettre.

Il n'y a pas d'équivalent sous windows, vu que les acl sont prises en compte nativement dans les partitions windows (ntfs).

Notes complémentaires :

Cette commande est très particulière, liée à un système de fichier comme celui utilisé en TP sur la machine debian (ext3) non doté nativement des extensions pour acl, mais capable d'en être doté par une extension adéquate.

Avec des systèmes de fichiers plus archaïques (fat, ext...), cette commande n'aurait pas lieu d'être émise : aucune extension ne peut doter ces FS de fonctionnalités ACL.

A l'inverse, avec des FS plus performants, cette commande n'aurait pas lieu d'être émise non plus : ces FS gèrent nativement les ACL (NTFS, XFS...)

Par ailleurs, noter que tout ce qui précède concerne uniquement l'option -o acl de la commande tune2fs. La commande tune2fs possède de nombreuses autres options pour affiner la configuration d'une partition.

8) Voici un extrait du fichier smb.conf sur la machine dont le hostname est serv02.

[global]

workgroup = PROD

[homes]

path = /home/donnees/%G

L'utilisateur dont le login est yming (uid 567), du groupe staff (gid 34), a son répertoire d'accueil sur cette machine. Dans quel répertoire de serv02 se trouvent ses données ?

/home/donnees/staff/yming (on suppose que le nom du répertoire correspond au nom de login, comme c'est très généralement l'usage mais rien en fait n'est obligatoire à ce niveau-là)

Quel est le chemin réseau à utiliser pour connecter ce répertoire sur une machine windows ?

\\nom_serveur\nom_partage\nom_login (même remarque que ci-dessus), c'est-à-dire ici :

\\serv02\homes\yming

Quel est le nom du démon qui gère ces partages ?

smbd

2 pts

9) Quel est le rôle de la commande "su" ? Quelle est la différence entre "su" et "su -" ?

1 pt

La commande su permet de prendre temporairement l'identité d'un autre utilisateur :

su toto => je deviens toto

Si le nom de l'utilisateur est omis, la valeur par défaut est "root". Si le mécanisme du sudoer est mis en place, seuls les sudoers peuvent devenir root. Sur une machine sans mécanisme du sudoer, le compte root existe et n'importe qui peut faire un "su" (sans login) ou un "su root" (ce qui est la même chose) à condition de connaître le mot de passe de root.

L'utilisateur (nouveau) est invité à saisir son mot de passe (exception possible : on est root au départ)

La commande id renvoie alors le uid et le gid du nouvel utilisateur. Le répertoire courant et les variables d'environnement ne sont pas affectés par ce changement.

L'option "-" est une (curiosité unix) abréviation de l'option la plus courante de cette commande : "-l" ou "--login". Tout se passe comme si l'utilisateur venait d'effectuer un login : les variables d'environnement sont mises à jour, le répertoire d'accueil de l'utilisateur devient le répertoire courant et son fichier .bashrc (ou autre) est exécuté.

10) Rappelez la définition des RAID 0, 1, 5, 6 et 51.
1 pt au total - 2 lignes max à chaque définition

RAID0 : "Data Striping". Les clusters sont stockés alternativement sur deux (ou plusieurs) disques. La vitesse de lecture et d'écriture est améliorée ; aucun gain en terme de sécurité. Utilisation typique : clacul : stockage de résultats intermédiaires.

RAID1 : "Miroir". Deux disques (rarement plus) portent les mêmes données. Vitesse de lecture légèrement améliorée. Si un disque tombe, les données sont toujours présentes sur l'autre. Inconvénient : il faut prévoir 2 fois plus de disque que ce qui sera disponible.

RAID5 : "Parité distribuée". Les données sont réparties sur plusieurs disques (typiquement 3). Un peut tomber en panne. Si un disque tombe, les données seront extraites des autres. Lorsqu'on remplacera le disque défectueux, les données y seront réécrites à partir des données présentes sur les autres disques.

RAID6 : "double parité distribuée". Les données sont réparties sur plusieurs disques (typiquement 6). Deux peuvent tomber en panne simultanément. Même comportement que pour RAID5.

Nota : en RAID 5 et 6, les disques sont complètement équivalents, il n'y en a pas un ou deux dédiés à la parité, comme on peut parfois le lire.

RAID51 : miroir de 2 groupes organisés chacun en RAID5

11) Quel mécanisme est mis en oeuvre lorsque vous voyez le message "Activation des paramètres de l'ordinateur" au démarrage d'une machine windows ?

Il s'agit de la mise en oeuvre des stratégies de groupes, plus particulièrement des règles concernant l'ordinateur. Les règles concernant l'utilisateur seront mises en oeuvre au moment du login.

Comment se nomme l'outil qui permet de gérer ce mécanisme.

L'éditeur de stratégies de groupes.

1 pt

12) La déduplication peut être mise en oeuvre au niveau fichier ou au niveau cluster (bloc). Le fichier f1 contient 4 clusters de données, dont le dernier est incomplet. On fait 2 copies de f1, nommées f2 et f3. Combien de clusters de données sont occupés par les 3 fichiers f1, f2 et f3

a/ si la déduplication n'est pas mise en oeuvre ?

b/ dans le cas de déduplication au niveau fichier ?

c/ dans le cas de déduplication au niveau cluster ?

a/ 12.

3 fichiers de 4 clusters chacun = 12 clusters ;

Pas compris pourquoi tout le monde n'avait pas *au moins* indiqué ça !

b/ et c/ 4

La déduplication comme son nom l'indique, évite de dupliquer des "choses" (fichiers ou clusters)

identiques. Le résultat est le même dans les 2 cas : les 4 clusters suffisent pour représenter les données, qui sont communes aux 3 fichiers.

f3 grossit et occupe désormais 5 clusters de données. Après cette opération, combien de clusters de données sont occupés par les 3 fichiers f1, f2 et f3

d/ si la déduplication n'est pas mise en oeuvre ?

e/ dans le cas de déduplication au niveau fichier ?

f/ dans le cas de déduplication au niveau cluster ?

d/ 13 : f1 et f2 occupent toujours leurs 4 clusters chacun. f3 en occupe 5 désormais. $(2 \times 4) + 5 = 13$

e/ 4 clusters pour les 2 fichiers qui n'ont pas bougé, et sont toujours représentés par une seule copie et 5 pour le nouveau fichier, soit 9 au total, car $4 + 5 = 9$.

f/ f3 a les mêmes 3 premiers clusters que les 2 autres fichiers, ces 3 clusters n'occupent pas de place supplémentaire. f3 a besoin de 2 clusters supplémentaires, pour ses 2 derniers clusters qui lui sont propres. Il faut donc 4 clusters (pour f1 et f2, et dont les 3 premiers servent aussi à f3) plus 2 clusters propres à f2 (les 2 derniers), soit au total : 6 clusters.

lpt

Manifestement, beaucoup n'ont pas compris le rôle de la déduplication. La déduplication permet de gagner de l'espace. La déduplication au niveau cluster est plus fine (donc plus efficace en terme de gain de place) que la déduplication au niveau fichier. En ce qui concerne les valeurs numériques, sans aucun calcul, on ne pouvait qu'avoir $a \geq b \geq c$ et $d \geq e \geq f$.

La précision "dont le dernier est incomplet" est là juste pour dire que lorsque le fichier grossit et passe à 5 clusters, 2 clusters sont affectés. Dans le cas très improbable où le dernier cluster aurait été complet, un seul cluster aurait différencié f3 des 2 autres.

13) Plusieurs techniques contribuent à la haute disponibilité. Définissez-les en quelques mots.

Heartbeat

Les machines testent périodiquement que les autres sont toujours actives, par exemple par des ping. En cas de réponse négative, on tente de relancer automatiquement la machine défaillante.

Agrégation de liens

Chaque machine possède non pas une, mais plusieurs cartes réseau. Le débit en est augmenté. Si une carte tombe en panne, le système continue à fonctionner, en mode dégradé.

Réplication

Les données sont dupliquées et disponibles sur deux machines différentes.

Multipath I/O

Il existe plusieurs (au moins 2) chemins physiques complètement séparés par lesquelles les données peuvent transiter (cartes réseau, commutateurs).