5 - LE SYSTEME NFS

NFS (Network File System) permet d'offrir un accès aussi transparent que possible à l'espace disque d'un environnement distant. Il a été développé dès les années 1980 par Sun. Il a été depuis intégré par de nombreux constructeurs.

NFS utilise un ensemble de procédures RPC qui utilisent le protocole XDR. C'est un protocole utilisant UDP. Un serveur NFS ne garde aucun historique des requêtes concernant un fichier. Par suite, si un serveur NFS n'est plus en mesure de donner des réponses, les processus clients sont bloqués.

La transparence par NFS provient

- d'opérations génériques réalisées sur un système virtuel de fichier (VFS)
- de la définition de noeuds virtuels (Vnode) qui cachent la structure physique du fichier.

L'ensemble forme l'interface vnode.

exemple : l'appel système chown() possède un analogue dans l'interface vnode qui affecte les attributs de fichiers. Si un client utilise cette fonction sur un système de fichier DOS, l'interface rejettera l'opération.

De la même manière qu'UNIX nécessite un descripteur de fichier pour accéder à la ressource, pour effectuer un appel système NFS, on doit tout d'abord obtenir un descripteur vnode. La couche VFS traduira le descripteur de fichier au niveau du noyau.

Sur un serveur NFS, il existe un démon nfsd qui accepte les appels RPC en provenance de clients. Pour traiter les requêtes de montage et de traduction des chemins d'accès un démon mountd fonctionne. En général, il existe plusieurs nfs et plusieurs mountd pour éviter la saturation des files d'attentes.

Exemple sur sauron.emi, il y a 4 démons nfsd et 4 démons mountd.

Sur le site client, tout processus qui utilise des fichiers montés par NFS effectue des appels RPC qui ne sont que des extensions des procédures de bases sur un système de gestion de fichier. En d'autre terme, de la même manière que la fonction getpwuid() sait comment utiliser NIS pour localiser l'information, les appels systèmes de base destinés au système de fichiers savent comment contacter le système distant.

Les procédures sont les suivantes :

- opérations sur les répertoires : mkdir,rmdir, readdir, rename, remove, create
- opération de traduction : lookup
- opérations sur les données : read, write, getattr, setattr, link, symlink, readlynk
- information sur les sytèmes de fichiers montés : statfs

Les opérations de montages sont traitées par le démon mountd et ne font pas partie du protocole de montage car elles sont liées au système d'exploitation. Le protocole de montage doit donc faire des traduction des chemins d'accès.

Les fonctions sont les suivantes :

- mnt, ajout d'un répertoire à la table de montage
- dump, valeur de la table de montage
- umnt, suppression d'une entrée de la table des montages
- umntall, supprime toutes les entrées de la table des montages
- export, donne la liste des répertoire exportés.

Certaines de ces opérations ne sont pas idempotentes. Un serveur NFS possède un cache qui lui permet de rejeter une requête identique à une autre très récente. Il contient plusieurs centaines d'entrées soit quelques secondes d'activité ... Il ne possède pas de mécanisme de reprise sur panne. Le démon mountd maintient un fichier /etc/rmtab des clients et de leurs requêtes lorsque celles ci sont acceptées. Les entrées sont supprimées losque le client effectue un démontage explicite.

Côté client, le noyau garde un cache de recherche de noms de répertoire afin d'éviter des requêtes inutiles.

Exemple Après la séquence

mount serv1:/usr/local /usr/local

mount serv2:/usr/local/bin-local /usr/local/bin

une requête à /usr/local/bin se fera bien sur serv2.

Exemple Suposons qu'on ait effectué la commande

mount sauron: /opt/gnu /usr/local nfs ro,hard 0 0

l'exécution de la commande /usr/local/bin/emacs va commencer par la conversion du chemin d'accès.

1 - le client NFS demande au démon mountd du serveur sauron un identificateur pour le système de fichier monté sur /usr/local. Soit X1 cet identificateur.

2 - En utilisant l'identificateur du point de montage, le client demande le composant suivant en effectuant une requête lookup sur le serveur sauron avec l'identificateur X1 et l'argument bin. Le serveur renvoit un identificateur X2.

3 - le client travaille alors sur le dernier composant emacs, envoit une requête lookup au serveur avec pour argument X2 et emacs. L'identificateur retourné sera utilisé pour toutes les opérations futures sur ce fichier.

Sur un client, il est possible de faire tourner un démon biod pour améliorer les performances de lecture et écriture de fichiers. Ce démon utilise le principe de cache pour anticiper les opérations de lecture ou au contraire retarder les opérations d'écritures.

NFS ne prend pas en charge le vérouillage des fichiers, ce sont deux autres démons lockd et statd qui gèrent ce problème.

6 – INTALLATION DE NFS

Pour utiliser NFS sur un client les démons lockd et statd et éventuellement biod doivent être lancés. Sur un serveur nfs on doit lancer les démons nfsd et mountd ainsi que les démons de gestion des vérouillage lockd et statd.

Le démon nfsd utilise un fichier /etc/exports qui détermine les système de fichiers qui pourront être montés par un client.

Exemple

/usr/bin ro -root=fermat

La syntaxe des lignes de ce fichier est la suivante :

<répertoire> <liste_option>

où les options sont les suivantes :

-rw:host:host	limiter les sites qui peuvent écrire
-ro	interdiction d'écriture
access=host:host	accès limité aux sites spécifiés
root=host:host	les utilisateurs root en provenance des sites specifiés restent root
anon=uid	faire correspondre l'utilisateur anonyme à uid
secure	accès uniquement par secure rpc

Ce fichier doit être exporté par la commande exportfs qui produit un fichier /etc/xtab. L'exportation de système de fichier est soumise à des règles :

1 - Tout système de fichier ou sous-ensemble peut être exporté.

Exemple Si /usr est un système contenant /usr/local. Ce dernier est un sous ensemble

2 - Il est impossible d'exporter un sous-système d'un système ou sous-système exporté à moins qu'il ne soit sur un disque différent.

Exemple Si /usr est un système contenant /usr/local et /usr est exporté on ne peut exporter /usr/local.

3 - Il est impossible d'exporter le père d'un système exporté à moins qu'il ne soit sur un disque différent.

4 - Seuls les systèmes de fichiers locaux peuvent être exportés.

La commande df permet de vérifier la validité des exportations car elle indique sur quel système de fichiers local le réperoire réside.

Côté client, le montage se fait par des commandes habituelles mount et le système de fichier figure en général dans le fichier /etc/fstab. L'option soft (resp. hard) de mount permet d'effectuer l'arrêt d'une exécution de requête par abandon (resp. kill).

Quelques outils permettent d'observer le comportement de nfs :

- exportfs : permet de connaître le contenu de /etc/xtab
- showmount : permet de connaître les paires clients-répertoire
- nfsstat : affiche les statistiques du client (resp. serveur) avec l'option -c (resp. -s)

7 – L'AUTOMONTEUR

L'automonteur est un outil qui monte automatiquement des systèmes de fichiers lorsqu'ils sont référencés et les démonte quand ils ne sont plus utilisés. Il applique les principe de NIS aux fichiers de configuration de NFS. Les avantages sont les suivants :

- fichiers /etc/fstab moins complexes sur chaque hôte
- tâche d'administration plus légère
- risque de suspension d'un processus suite à une panne réduit
- extension du protocole de montage au serveur le plus proche dans le cas de dupplication de ressources communes.

Ce démon amd intercepte les requête NFS en direction des volumes montés. Il effectue alors réellement le montage, le démontage se fait automatiquement après une période d'inactivité. Il utilise des tables pour sa configuration. Ce sont ces tables qui seront gérées par NIS si elles sont introduites dans le Makefile du /var/yp.

La table principale /etc/auto.master précise pour chaque chaque répertoire, le fichier contenant le plan de montage associé.

Ce fichier s'appelle une table directe ou indirecte. Elle est directe si le point de montage correspond au répertoire distant et indirecte sinon.

Une entrée de ce fichier contient un nom de répertoire suivi du répertoire distant a monté avec éventuellement les options nécessaires au montage.

Exemple (extrait de Moreno):

fichier /etc/automaster

```
/-          /etc/auto.direct
/sources    /etc/auto.sources -rw,soft
```

fichier /etc/auto.direct

```
/usr/src    -rw    thermidor:/usr/src
/usr/man    -rw    amertume:/usr/man
```

fichier /etc/auto.sources

```
musique     sweet:/dic/sources/musique
images      soft   prairial:/images
```