LA CONTINUITÉ DE SERVICE

INTRODUCTION

Si la performance est un élément important de satisfaction de l'utilisateur de réseau, la permanence de la disponibilité des ressources l'est encore davantage.

Les éléments «sensibles» d'un réseau sont par ordre d'importance:

- la connectique qui est responsable de 70% des pannes réseau
- le serveur qui, lors d'une panne, provoque l'arrêt de tout le réseau
- l'électronique active (HUB, cartes).

LES SOLUTIONS POUR ÉVITER UNE PANNE DU SERVEUR

LOCALISATION DU SERVEUR

Il faut installer le serveur dans un lieu protégé, à l'abri des malveillances. Le local l'abritant doit être suffisamment ventilé pour éviter en hiver comme en été une température trop élevée, ou est climatisé.

ENVIRONNEMENT ÉLECTRIQUE

L'alimentation électrique du serveur doit être réalisée avec une ligne autonome, ou tout au moins sans interférence avec des dispositifs "polluants" de type systèmes industriels. Le serveur est protégé par un onduleur de type "en ligne", c'est-à-dire sans délai d'attente, le courant distribué étant toujours "propre" (en provenance de la batterie); d'une puissance suffisante (si la consommation de l'ensemble serveur + écran est, par exemple, de 800 VA (Volt-Ampère), il est recommandé de s'équiper d'un onduleur de 1000 VA, ce qui représente une marge de sécurité de l'ordre de 25%).

L'onduleur est équipé d'un dispositif "powerchute". Ce dispositif est constitué d'une interface de communication implantée sur l'onduleur, d'un câble de jonction reliant l'interface à un port du serveur (souris, COM1,.), d'un logiciel fonctionnant sur le serveur et exploitant les informations en provenance de l'onduleur. En liaison avec le système d'exploitation du serveur, ce dispositif permettra de détecter et d'informer les utilisateurs: d'une coupure secteur, d'une fin imminente d'autonomie des batteries avec clôture immédiate des fichiers, d'une reprise de l'alimentation secteur (éventuellement), d'un arrêt immédiat du serveur. Ce dispositif permet également un redémarrage automatique et programmé du serveur.

ARCHITECTURE DU SERVEUR

En sélectionnant rigoureusement des composants de qualité, en organisant de façon redondante certains dispositifs, on obtiendra un produit fiable tendant vers le «0 défaut» (0 panne).

Cette sécurité a un prix et chacun, compte tenu du niveau souhaité de disponibilité de son système et de son budget, fera des choix.

Le serveur peut être constitué ou organisé avec des éléments qui vont favoriser la sécurité du dispositif. Les principaux éléments à prendre en compte sont : La mémoire, les disques, le processeur, le bloc d'alimentation électrique, la surveillance de la température.

TOLÉRANCE DE PANNE MÉMOIRE

Un système de tolérance de panne mémoire peut être mis en oeuvre par l'utilisation de mémoire ECC (Error Checking and Correcting Memory). La mémoire ECC utilise des bits redondants pour permettre à la mémoire système de détecter et de corriger les erreurs d'un simple bit (qui représentent 95% des erreurs mémoire). On peut encore améliorer la fiabilité de la mémoire en utilisant une mémoire "advanced ECC", qui peut corriger les erreurs de 4 bits adjacents.

Ceci permet au serveur de rester opérationnel même si un composant entier de module mémoire se trouve hors d'usage.

TOLÉRANCE DE PANNE DISQUE

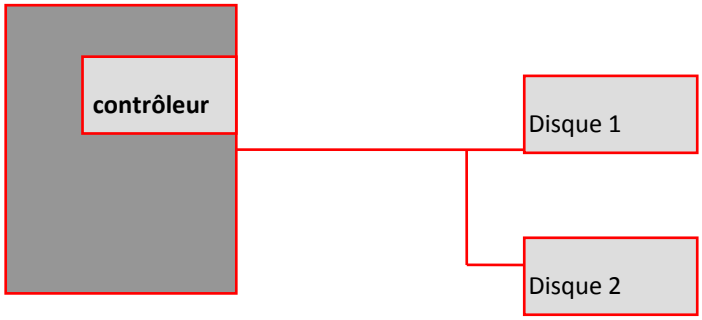
Différentes techniques peuvent être employées pour assurer la protection des données et offrir des systèmes à tolérance de panne. Une classification de l'organisation des données sur les disques a été proposée par l'Université de Berkeley sous le nom de RAID (Redundant Arrays of Inexpensive Disk).La classification d'origine propose 5 niveaux de RAID (RAID-1 à 5), auxquels se sont ajoutés d'autres niveaux : RAID-0, RAID-5 orthogonal, RAID-6, RAID-10.

Nous ne présenterons ici que les niveaux de RAID assurant un bon débit d'E/S et une grande disponibilité des données.

RAID-1

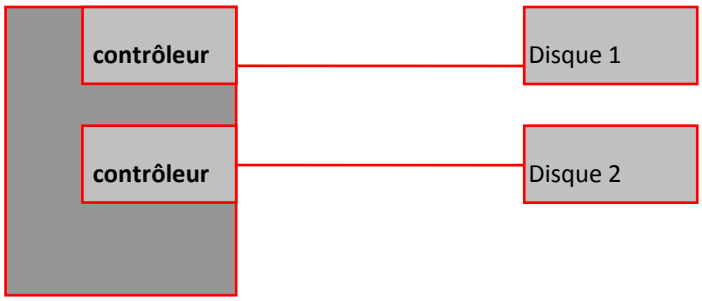
Cette technique est aussi connue sous le nom de "**mirroring/duplexing**". Chaque fois qu'une modification est enregistrée sur un disque, cette modification est également réalisée sur un deuxième disque appelé disque miroir du premier. Ainsi, l'un ou l'autre des deux disques en miroir peut tomber en panne, l'information reste toujours accessible. De plus, une requête de lecture peut être satisfaite plus rapidement dans certains cas, car elle est réalisée à partir de deux disques.

Cette technique très satisfaisante sur le plan de la sécurité est coûteuse, puisqu'elle double la quantité de disque nécessaire (seule la moitié de la capacité installée est disponible pour enregistrer les données).



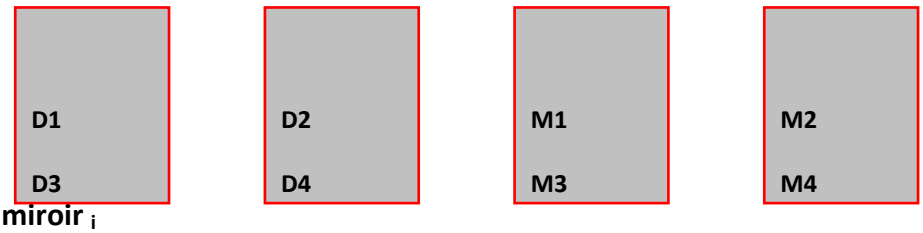
"Mirroring"

Le "**duplexing**" est identique dans son fonctionnement au "**mirroring**" avec, en plus, l'attachement de chaque disque à un contrôleur différent. Ceci améliore encore la fiabilité (et la performance) de l'ensemble, puisqu'un contrôleur peut tomber en panne sans que cela n'arrête le serveur.



"Duplexing"

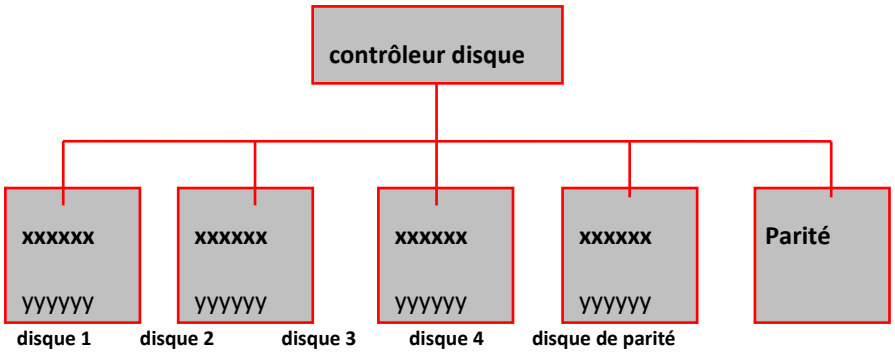
La généralisation du RAID-1 à plus de deux disques, s'appelle, suivant les constructeurs RAID-6 ou RAID-10 ou RAID "hybride". Elle apporte le même niveau de sécurité que le RAID-1 et renforce les performances.



D_i : donnée$_i$ **M_i : miroir $_i$**

RAID-4

Les données sont réparties sur plusieurs disques de telle sorte que les performances soient optimales lors de l'accès en lecture à de grands fichiers. Les données de parité sont stockées sur un disque dédié.

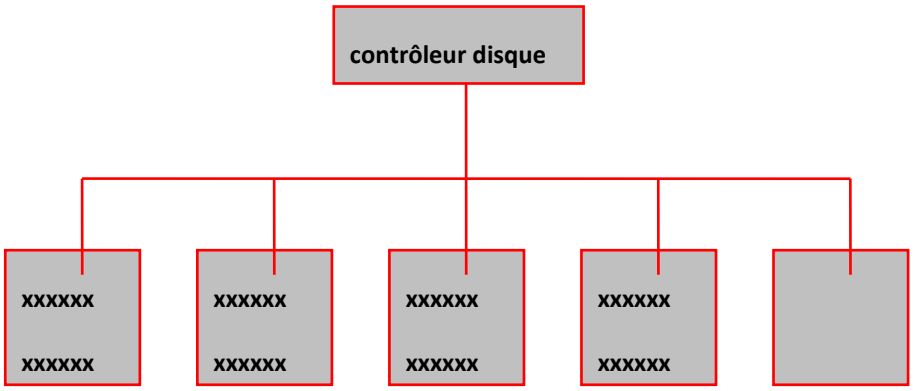


xxxxx et yyyyy = blocs appartenant à de grands fichiers.

En cas de panne d'un disque on peut reconstruire la structure du fichier de données grâce aux informations de parité et aux données enregistrées sur les autres disques. Par contre on mobilise un disque entier uniquement pour enregistrer ces informations de parité.

RAID-5

Les données sont réparties sur plusieurs disques de telle sorte que les performances soient optimales lors de l'accès en lecture à de grands fichiers. Les données de parité, à la différence du RAID-4, au lieu d'être stockées sur un disque dédié, sont réparties sur plusieurs disques, évitant ainsi la création d'une file d'attente sur le disque de parité et donc une baisse de performances.



Cette technique est aussi particulièrement adaptée aux environnements transactionnels intensifs nécessitant de nombreuses opérations d'E/S. Si un disque tombe en panne, le contrôleur pourra reconstituer le disque endommagé, à partir des données et des parités enregistrées sur les autres disques. Il faut donc disposer d'un disque de rechange (ou disque relais). Ce disque doit être de capacité supérieure ou égale à la plus grande des capacités des disques de données. Si ce disque relais est déjà installé, le contrôleur pourra donc automatiquement, sans arrêt du système, reconstituer le disque endommagé.

De plus, si les disques installés sont de type "hot-plug" (extractible à chaud), on pourra échanger le disque endommagé, sans arrêt du système, par un nouveau disque qui deviendra le nouveau disque relais.

RAID-5 orthogonal

Cette organisation reprend le principe de l'organisation RAID-5, et regroupe par paquet de 4 les disques sur lesquels s'appliquent le RAID-5. Les disques d'un même paquet sont reliés à des contrôleurs différents (généralisation de la technique de "duplexing"), ce qui améliore les performances et la fiabilité du sous-système disque.

TOLÉRANCE DE PANNE PROCESSEUR

Ceci est réalisé par certains constructeurs qui proposent dans leur système un processeur non connecté, de sauvegarde. Si le processeur actif tombe en panne, le serveur s'arrête, puis redémarre automatiquement avec le processeur de sauvegarde. Une intervention de dépannage (avec arrêt du serveur) pourra être ainsi planifiée, de telle sorte qu'elle pénalise le moins possible les utilisateurs.

LE BLOC D'ALIMENTATION ÉLECTRIQUE

Disposer de deux blocs d'alimentation électrique sur un serveur, permet, dans l'hypothèse ou l'un tombe en panne, à l'autre de prendre immédiatement le relais, et donc d'éviter l'interruption du serveur.

LA SURVEILLANCE DE LA TEMPÉRATURE

Une sonde thermique présente dans le serveur permettra, de lancer des alertes si la température s'élève anormalement, et ainsi de diagnostiquer une avarie (par exemple une panne de ventilateur) qui pourra être réparée rapidement.

LES CLUSTERS DE SERVEURS

Il suffit de prendre l'exemple d'un serveur de messagerie qui ne fonctionne plus pendant une demi-journée suite à un disque dur défectueux, le temps de le remplacer et de faire la restauration ; le service commercial ainsi que la direction ne peuvent plus répondre aux appels d'offres, le service clients ne peut pas faire le suivi des réclamations, le service technique ne peut plus passer de commandes. La structure est paralysée pour tous les échanges de courrier électronique, qui représente un pourcentage conséquent de la gestion des activités au sein d'une structure commerciale. Les résultats sont radicaux puisque l'entreprise cumule des contrats perdus, des

bénéfices en moins, des heures de travail perdues, des sanctions pour l'équipe de commerciaux et pour le service informatique.

Pour éviter ce genre de scénario catastrophe certains architectes ou administrateurs de systèmes d'informations décident d'implémenter un service de cluster sur les serveurs hébergeant les applications critiques : serveur de messagerie, serveur ERP, serveur Web commerce électronique, serveur de base de données, serveur de fichiers ou autres.

TERMINOLOGIE

La technologie de clustering permet d'avoir une haute disponibilité des ressources publiées. On utilise cette technologie pour avoir une disponibilité et stabilité des ressources proche de 100 %. Tolérance zéro pour les pannes matérielles ou logicielles. Il y a également une répartition des charges entre les nœuds d'un cluster.

Un serveur de cluster est un groupe de serveurs gérant des ressources stockées sur des disques partagés. Les nœuds et les disques sont connectés par un bus de liaison (SCSI ou Fibre Channel).

Un serveur dans le cluster est appelé nœud dit node en anglais. Les données publiques sont appelées ressources, chaque disque du bus partagé représente un groupe de ressources ; pour publier un groupe de ressources accessible par les clients externes, il est nécessaire de créer un serveur virtuel en lui adressant une adresse IP virtuelle et un nom d'hôte.

Lorsqu'un client externe se connecte pour faire une requête sur les données, celle-ci transite par le serveur virtuel, qui fait office de « passerelle » entre les nœuds connecté aux disques partagés du cluster et le client.

Par défaut chaque groupe de ressources est attribué à un nœud. Dans le cas où le nœud a une défaillance , l'autre nœud prend en charge les groupes de ressources de son homologue, et répond aux requêtes distantes.

C'est la phase de basculement entre les 2 nœuds, appelé **failover**, en conséquent la mise en place d'un cluster permet d'avoir une disponibilité des ressources proche de 100%.

Haute disponibilité (Availability) des ressources sur le cluster, celles-ci sont garanties disponibles à 99,9 % du temps. Dans le cas où un des nœuds ne pourrait plus fournir des réponses aux requêtes des clients, alors les autres nœuds du cluster prennent le relais. Ainsi la communication avec les clients et l'application hébergée ou autres ressources sur le cluster ne subit pas d'interruption ou une très courte interruption.

Adaptabilité : (Scalability) : il est possible d'ajouter un à plusieurs nœuds, ou d'ajouter des ressources physiques (disques, processeurs, mémoire vive) à un nœud du cluster. En effet, il est

possible que de part les trop nombreuses requêtes sur le serveur que celui-ci soit en saturation au niveau de la charge processeur, mémoire ou autre, dans quel cas il est nécessaire d'ajouter des éléments, voir un autre nœud.

Évolutivité : Lorsque la charge totale excède les capacités des systèmes du cluster, d'autres systèmes peuvent lui être ajoutés. En architecture multiprocesseur, pour étendre les capacités du système, il faut dès le départ opter pour des serveurs haut de gamme coûteux autorisant l'ajout d'autres processeurs, de lecteurs et de la mémoire supplémentaires.

TECHNOLOGIE

Sur le plan technique, le clustering consiste à mettre en grappe des serveurs qui partagent des périphériques communs en se répartissant la charge du traitement. Les unités de stockage doivent être communes, et il faut un bus de communication entre les différents serveurs.

Les systèmes à répartition de charge permettent de distribuer l'exécution de processus systèmes ou réseaux à travers les nodes du cluster.

Le node server se voit ainsi attribuer la tâche de réceptionner le processus et de le répartir sur la machine adéquate. Cette dernière est en fait choisie car sa charge est faible et donc elle peut traiter le processus entrant de manière quasi instantanée. Elle peut aussi être choisie en fonction de sa spécialisation, c'est à dire qu'elle seule pourra traiter la demande sur l'ensemble des nodes du cluster.

CLUSTERS MICROSOFT

Depuis la version NT4 de son système d'exploitation (OS) Windows, Microsoft propose de mettre en place un cluster constitué de serveurs (Microsoft ! bien entendu !) pour répondre aux besoins croissants des entreprises en terme de messagerie électronique, de base de données et depuis quelques années de serveurs WEB ou FTP.

Microsoft a implémenté deux technologies de clustering sur ses serveurs Windows.

Le service de cluster MSCS

Le service MSCS fournit une haute disponibilité pour les applications critiques, telles que les bases de données, les serveurs de messagerie, serveur de fichier et d'impression.

Network Load Balancing

NLB permet d'équilibrer le trafic IP entrant. A travers différentes règles établies les connexions entrantes sont réparties entre les différents nœuds du cluster, il peut y avoir jusqu'à 32 nœuds pour équilibrer la charge IP en mode Network Load Balancing. Le service d'équilibrage de

charge de réseau augmente la disponibilité et la montée en charge des applications serveur basées sur l'accès Internet, tels que des serveurs WEB, des serveurs médias streaming, serveur Windows Terminal serveur ou autres.

Il existe une troisième technologie de clustering implémentée sur les serveurs Application Center.

Component Load Balancing - Application Center 2000

Equilibrage de Composants, le service CLB est intégré à Application Center 2000 (ou versions antérieures), ce type de clustering permet de répartir la charge sur plusieurs nœuds du cluster, pour les applications basées sur la technologie des objets COM et COM+ , une mise a jour pour les objets WMI et la gestion du framework .NET est désormais disponible. On parle de clustering d'application dit clustering de puissance.

L'architecture CLB est souvent couplé à la l'architecture de cluster NLB, dans le cas de serveur WEB basé sur le commerce électronique.

LES RÔLES:

NLB - répartition des connexions IP et requêtes distantes, répartition de la charge et bande passante sur les nœuds, connexion au site web.

CLB - répartition des appels sur les modules d'applications hébergés sur le cluster, accès à l'application commerciale distribuée.