Sécurité et sûreté des données

Celui qui n'a pas peur de perdre des données est un inconscient.

Plan du cours

- Enjeux et niveaux de perte de données
- Politique de sauvegarde et de restauration
- Technique robuste d'accès au données
- Protection des données système
- Protection des données utilisateurs
- Protection des SGBD

Introduction :

- Il s'agit d'un problème de politique d'entreprise plus que technique
- Il s'agit avant tout d'un problème d'architecture et d'organisation
- Il faut protéger mais toujours prévoir le pire

Chapitre 1 :

Enjeux et niveaux de perte de données

Méthodes pour perdre des données :

- Effacement accidentel
- Corruption volontaire
- Dysfonctionnement d'applications
- Panne matériel
- Perte ou vol de support
- Inondation
- Incendie

Coût des pertes de données :

Conséquence = Blocage du fonctionnement de l'entreprise

- Coût humain : travail à refaire !
- Coût financiers :
  - Perte de commande
  - Litige commerciaux
  - Salaire à repayer
- Certaines données ne peuvent être reconstruite !

Coût de la protection :

- Matériel spécifique
  - Centralisation
  - Redondance
  - sauvegarde
- Temps :
  - De conception
  - Surveillance réparations

Balance gain/perte

Il s'agit d'un calcul de risque/coût :

- Le coût de la protection est connus
- Le coût de la perte est difficile à évaluer dépend très fortement des situations

⇒Les coût des solutions rebute les décideurs !

⇒Attention au responsabilté !

⇒La variable d'ajustement c'est la politique et le logiciel !

Type de disque

Type	SATA	SCSI	SAS	FC	SSD
Largeur de bande par canal	133 Mo/s	320 Mo/s	300 MO/s	200-400Mo/s	133 MO/S (SATA)
Performance par disque	65 MO/s	90-105 Mo/s	105 Mo/s	105 Mo/s	20-80 Mo/s
Vitesse de rotation	7 200 T/min	10 000-15000 T/min			Pas de rotation
Temps d'accès	8 ms	4 ms			0,1 ms
Organisation E/S	NCQ 32b	TCQ 256 bits			NCQ 32b
Limitation par canal		A partir de 4 disques			
MTBF	1,2 M d'heures	1,4 millions d'heures			2 millions
Température	60 °C	60 °C			85 °C

Type de données à protéger

- Donnée d'application
- Fichier utilisateurs
- Données de mesures
  - Sécurité
  - Accès
  - Mesure physique
  - Messagerie
- Annuaire
- Données système

Technique de protection

- Sauvegarde
  vise à limité toutes consséquences des pertes !
- Archivage
  vise à garantir la pérennité et l'utilisation à long terme des données
- Redondance
  vise à garantir la disponibilité à cours terme

Politique de sécurité des données

C’est le cahiers des charges de la gestions des données :

- Une politique de stockage
- Une politique de sauvegarde
- Une politique d’archivage

Les politiques sont des suites de procédure décrite sous forme : but/action/limite

Un exemple pratique à discuté !

Chapitre 2 :

Politique de sauvegarde et de restauration

Politique de sauvegarde :

- Que sauvegarde t'on
- Avec quelle fréquence
- Quelles sont les moyens
- Quels sont les rôles
- Quelles sont les responsabilités
- Quelles sont les limites

Que sauvegarde t'on ?

Tout ce qui ne peut-être reconstruit automatiquement :

- Les données des applications d'entreprises
- Les fichiers utilisateurs
- La messagerie
- Les paramétrages systèmes/annuaire
- Les logiciels

On fait particulièrement attention aux index !

Matériel

Librairie :

- Disque : baie NAS
- Bande :
  - Type de K7 : DLT/SDLT/DAT/TK
  - Lecteur standlone
  - Robot chargeur
- Appliance (Synerbox par exemple)

Logiciel

Plusieurs types :

- Sauvegarde par poste
- Sauvegarde serveur
- Intégrer à un NAS
- Développer soit même ou acheter.

Il faut faire attention a :

- Matériel géré en particulier robot chargeur
- Licence …

Technique propre au données utilisateurs

Les données utilisateur sont :

- Hors contrôle pour le volume (outlook, téléchargement, …)
- Modifier aléatoirement
- Il n’y a pas d’état atomique

Ceci implique :

- Il faut les bornées (quota, antivirus,centralisation)
- Il faut les sauvegardé fréquement

Elle doivent être retenus sur disque préférentiellement

Technique propre au données systèmes

Fichier de configuration :

- Windows/MAC/OS400 outils spécifique
- Linux/Unix : /etc

Fichier de log :

- Linux/Unix /var/log
- Windows y a pas !

Annuaire :

- Unix/LINUX /etc ….
- Windows active directory Controleur BDC
- Mac Time/Machine

Technique propre au SGBD

- SGDB = DATA + structure
- Fichier d'index
- Volume impacté car gros fichiers

⇒Pas de structure pas de données !

⇒On procéde par réplique !

⇒Ne pas confondre avec l'historisation et l'archivage = datawarhouse !

Politique de sauvegarde

- Calendrier
  que sauvegarde t-on et quand ?
- Durée de rétention
  pendant combien de temps ?
- Rôle - documentation
  qui gère les K7, gestion des absences ?
- Scénario
  perte simple, virus, catastrophe ?

Calendrier de sauvegarde

S/J	1	2	3	4	5	6	7	8	1	2	3	4	5	6	7
L	A0 F	A2 F	A1 F	A2 F	A1 F	A2 F	A1 F	A2 F	A10 F	A2 F	A1 F	A2 F	A1 F	A2 F	A1 F
M	A1 F	A2 I1	A1 I1	A2 I1	A1 I1	A2 I1	A1 I1	A2 I1	A1 F	A2 I1	A1 I1	A2 I1	A1 I1	A2 I1	A1 I1
M	A1 I1	A2 I2	A1 I2	A2 I2	A1 I2	A2 I2	A1 I2	A2 I2	A1 I1	A2 I2	A1 I2	A2 I2	A1 I2	A2 I2	A1 I2
J	A1 I2	A2 I3	A1 I3	A2 I3	A1 I3	A2 I3	A1 I3	A2 I3	A1 I2	A2 I3	A1 I3	A2 I3	A1 I3	A2 I3	A1 I3
V	A1 I3	A2 I4	A1 I4	A2 I4	A1 I4	A2 I4	A1 I4	A2 I4	A1 I3	A2 I4	A1 I4	A2 I4	A1 I4	A2 I4	A1 I4
S	A1 I4	A2 I5	A1 I5	A2 I5	A1 I5	A2 I5	A1 I5	A2 I5	A1 I4	A2 I5	A1 I5	A2 I5	A1 I5	A2 I5	A1 I5

Exemple sur deux mois !

Contrôle des flux

Il faut réduire les volumes :

- Durée des sauvegardes
- Durée en restauration
- Sollicitation du matériel
- Impact d'une perte

Il faut donc trié les données :

- De mesure/simulation/calcul
- De téléchargement/cache/machine virtuel
- De log

Il faut morcelé au maximum !

Restauration des données utilisateurs

Les problèmes sont de deux ordres :

- Psychologique (sentiment de faute)
- Technique (retrouvé des fichiers)

Interrogatoire en règle :

- Quand le fichiers a t-il été perdu
- Ou se trouvait-il et quel est sont nom
- Quand à t-il été modifier la dernière fois

Travailler en off (sans l'utilisateur) !

Technique propre au mail

Le mail est répartie entre :

- Données de spool
- Fichiers utilisateurs

Les risques sont multiples :

- Virus
- Effacement accidentelle

=> La solution est la centralisation (IMAP)

Reprise en cas de désastre

Ce plan doit permettre la restauration en cas :

- Destruction par incendie
- Erreur matériel majeur
- Fausse manipulation majeur

Il doit inclure :

- Le remplacement rapide du matériel
- La restauration de tout les fichiers et données
- Plusieurs personnes (rôles)
- Une documentation détaillée sur papier

=> Il doit être connus et répété !

Ordre de restauration :

- Système
- Application
- Paramétrages
- Données d’annuaire
- Messagerie
- Fichiers utilisateurs

Attention à :

⇒Gestion de communication

⇒Blocages des modification pendant le procéssus (web, utilisateur)

Quelques conseils et évidences

- Une sauvegarde doit être testée
- Attention aux incendies
- Le budget est un problème
- Attention aux négligences

Gestion des portables

Problèmes :

- Ils ne sont pas connecté en permanence
- Ils ne sont pas sure
- Ils sont facile à voler

Solutions :

- Ils doivent initié leur propre sauvegarde
- La duplication en local est utile
- Les utilisateurs doivent être sensibilisés
- L’utilisation de tunnels (VPN ou ssh) est une solution !

Chapitre 3 :

Technique robuste de stockage et d'accès au données

Pourquoi centraliser les données

- Faciliter les sauvegarde
  - Moins de client
  - Moins de transfert réseaux
- Facilité la restauration
- Moins de destruction accidentel
- Accélérer les transferts entre serveur
- Eviter la duplication sauvage

Architecture possible

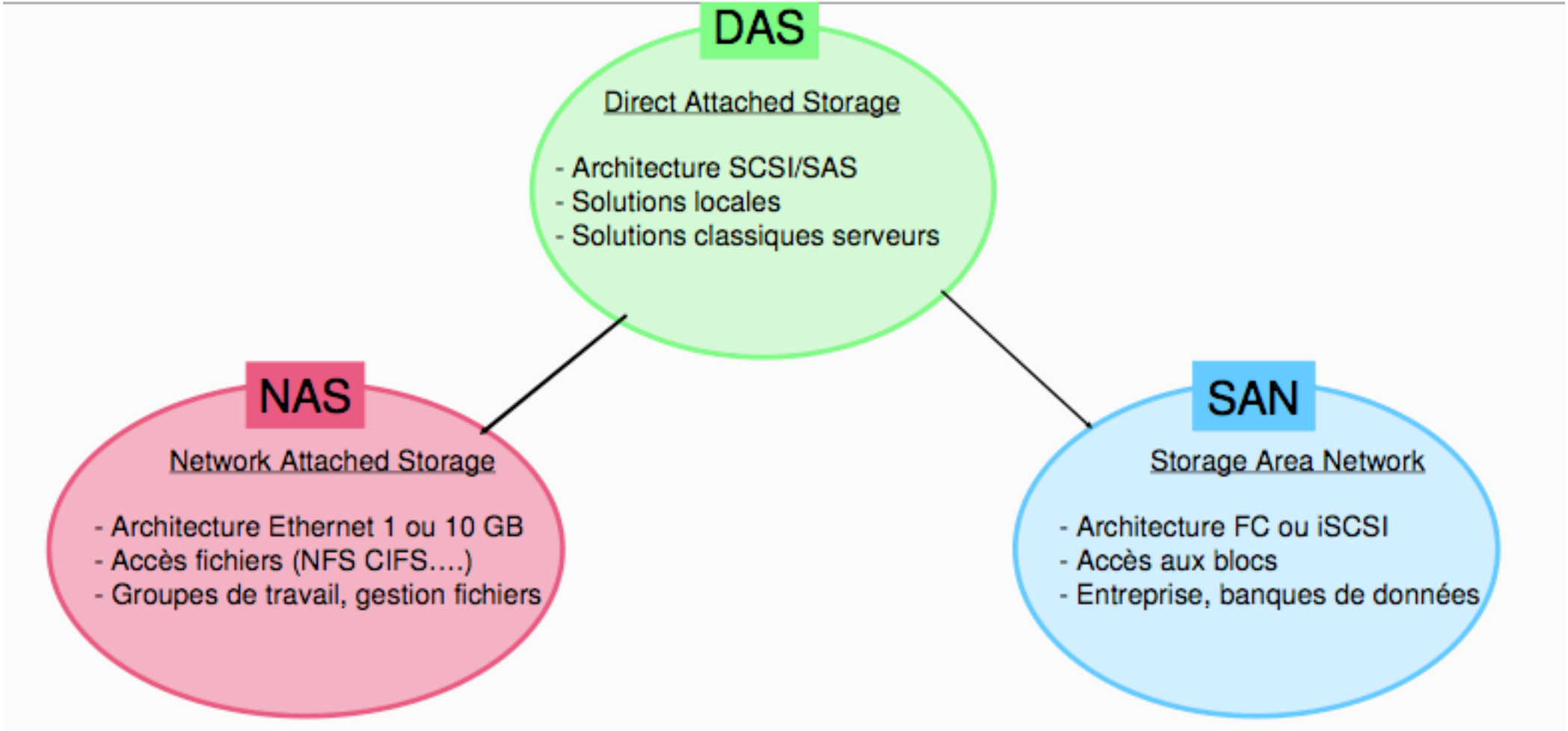
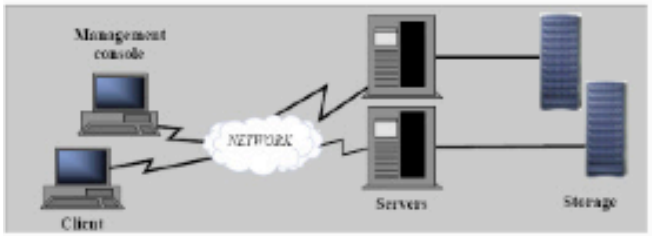


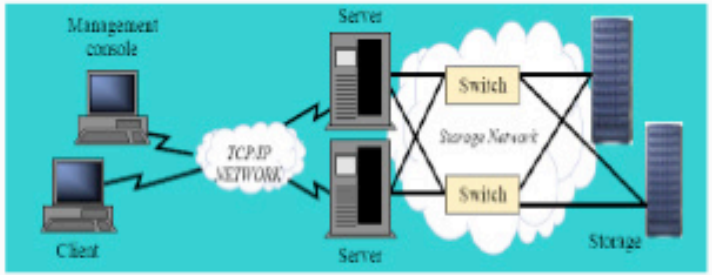
Schéma de centralisation

DAS

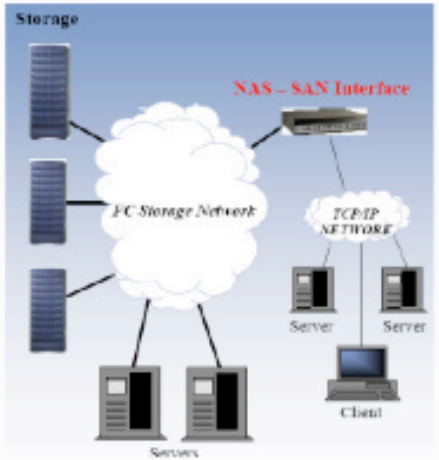


NAS

SAN



SAN avec pont NAS



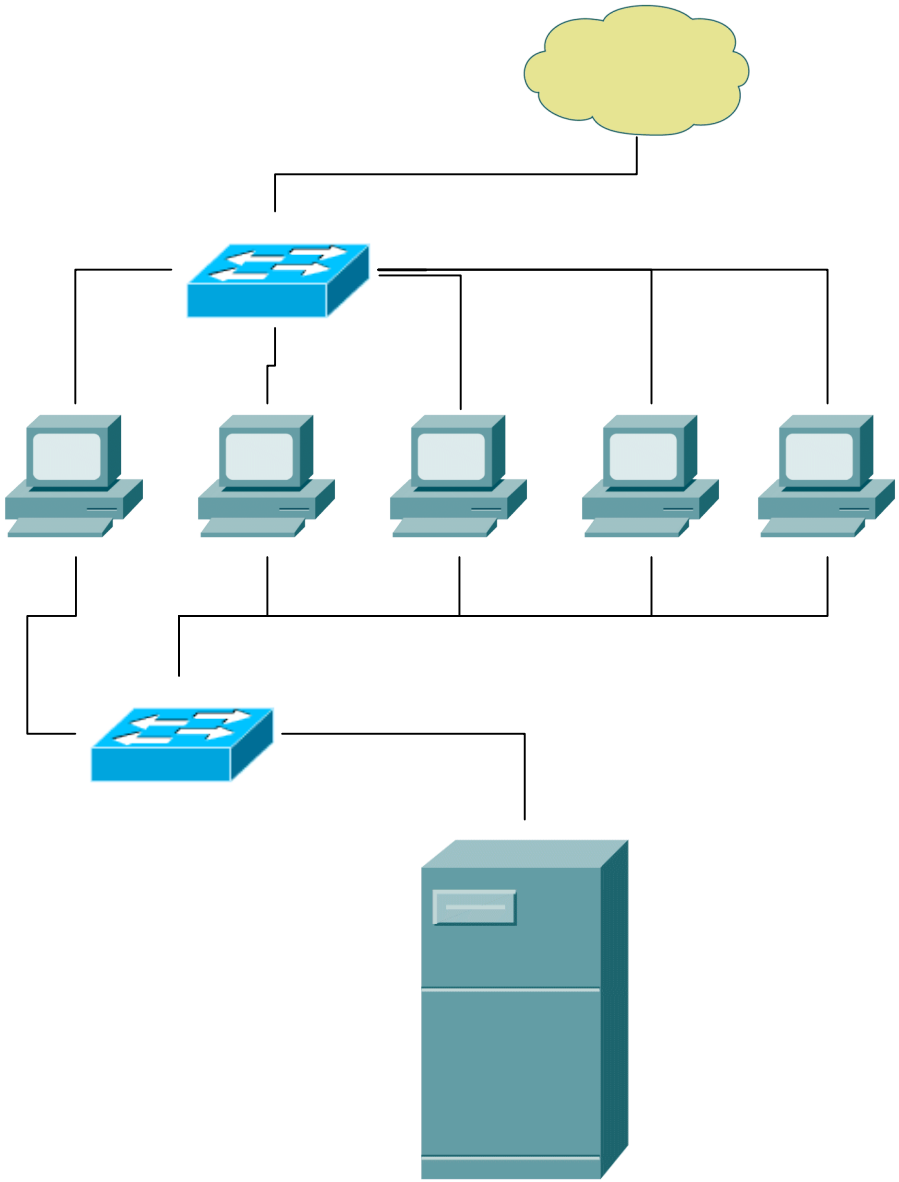
Niveau fichier – NAS

Serveur de fichiers dédier !

- Protocole CIFS
- Protocole NFS
- Protocole HTTP
- L’authentification se fait sur un serveur :
  - Active directory
  - NIS
  - Kerberos

Organisation service de fichiers.

- Réseaux dédier :
  - VLAN
  - Adresse IP
  - Switch dédié pour les serveurs principaux
- Système de quota

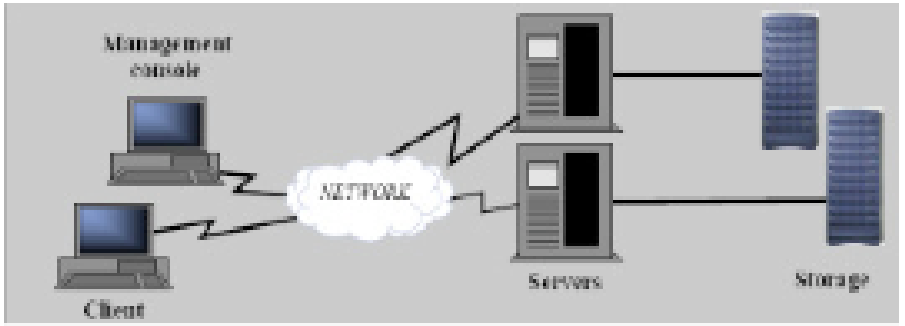


Niveau block – SAN

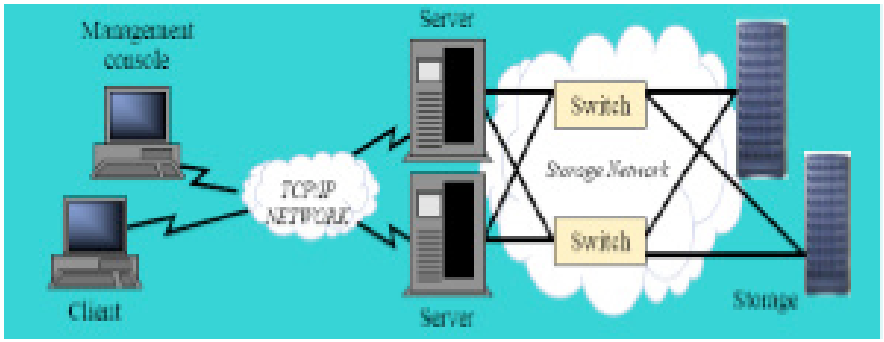
Il s'agit de baie de disque sur un réseau spécialisé :

- transports en mode block
- Disque virtualisé (LUN)

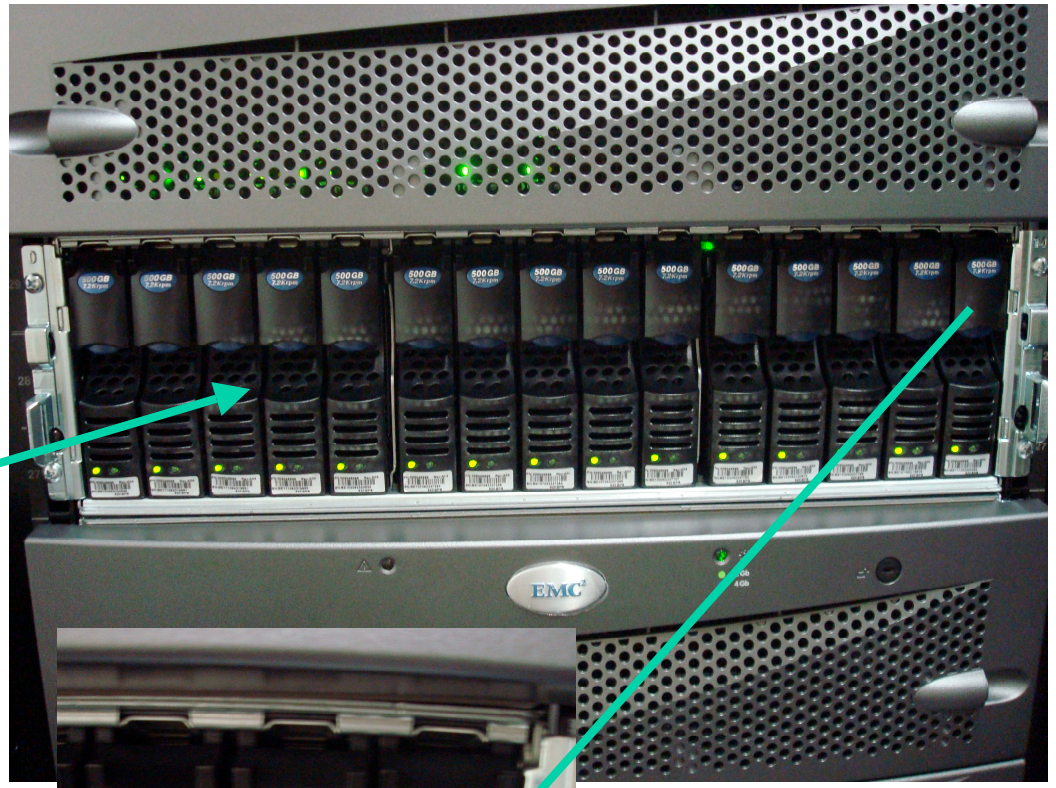
DAS



SAN



Concrètement



Avantages des SAN

- Accès commun aux appareils de stockage
- Performance Fiber Channel (4 GBit)
- Nombre d’appareils et de serveurs
- Utilisation des technologies réseau classique pour ISCSI ou FC
- Possibilité de câblage redondant
- Remplacement facile du serveur
- Sanctuarisation des données

Défauts des SAN

- Coût de l’espace
- Coût des connexions (carte HBA)
- Maintenance complexe
- Obsolescence
- Risque de black out

=> Il faut réfléchir à deux fois ce choix.

Mise en redondance

On veut évité les interruptions de service :

- Redondance d'alimentation (batterie double)
- Redondance de processeur (double processeur)
- Redondance de disque (raid)
- Redondance de chemin (double attachement)
- Redondance de performance (surcharge)
- Redondance de site
- Redondance de service ?

Redondance disque

Système RAID :

- Raid 0 = striping
- Raid 1 = mirroring
- Raid 2 = repartition bit à bit
- Raid 3 = répartition par secteur
- Raid 5 = répartition par bande de stockage
- Raid 10 = raid 1 + raid 5

Schéma RAID 1

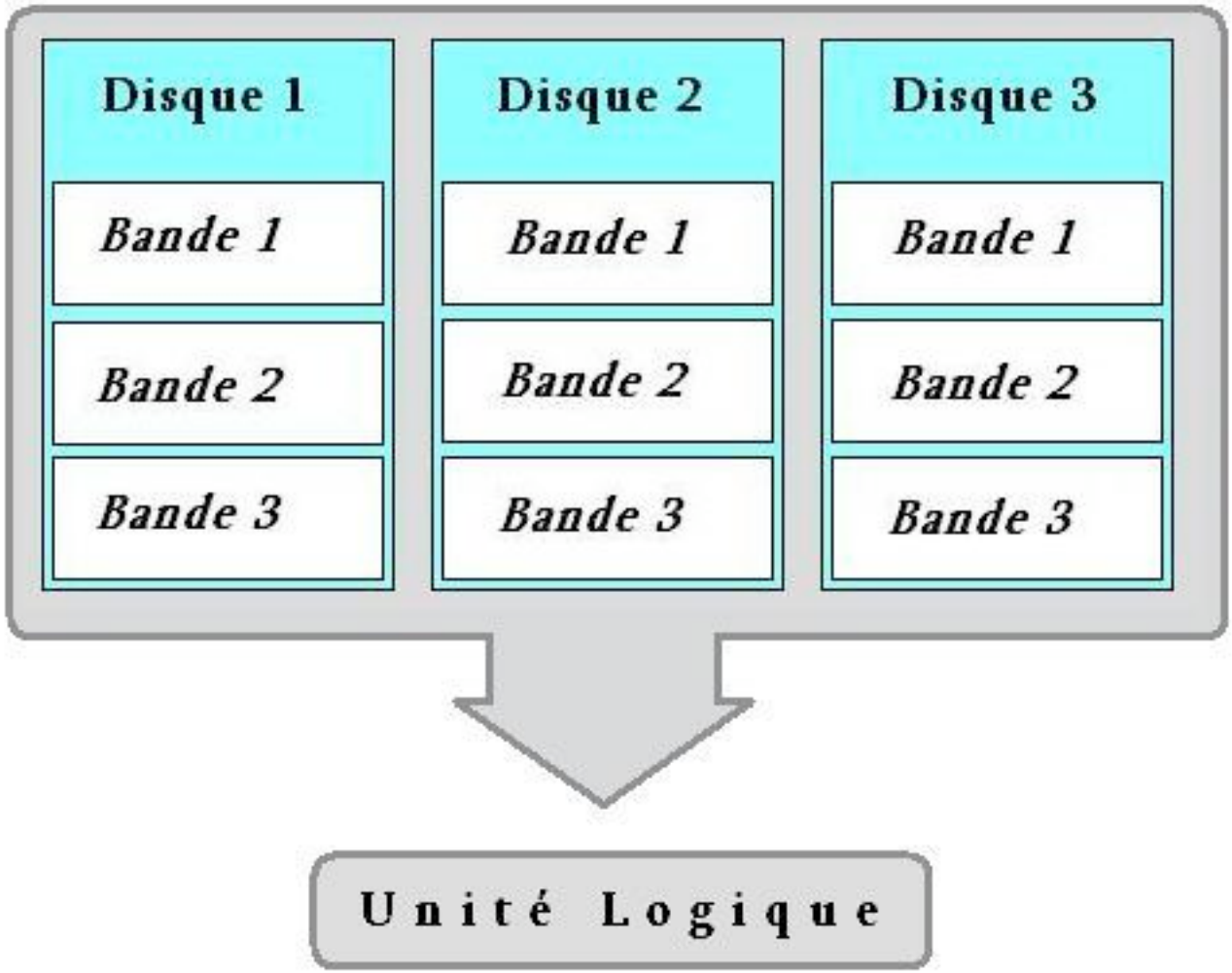
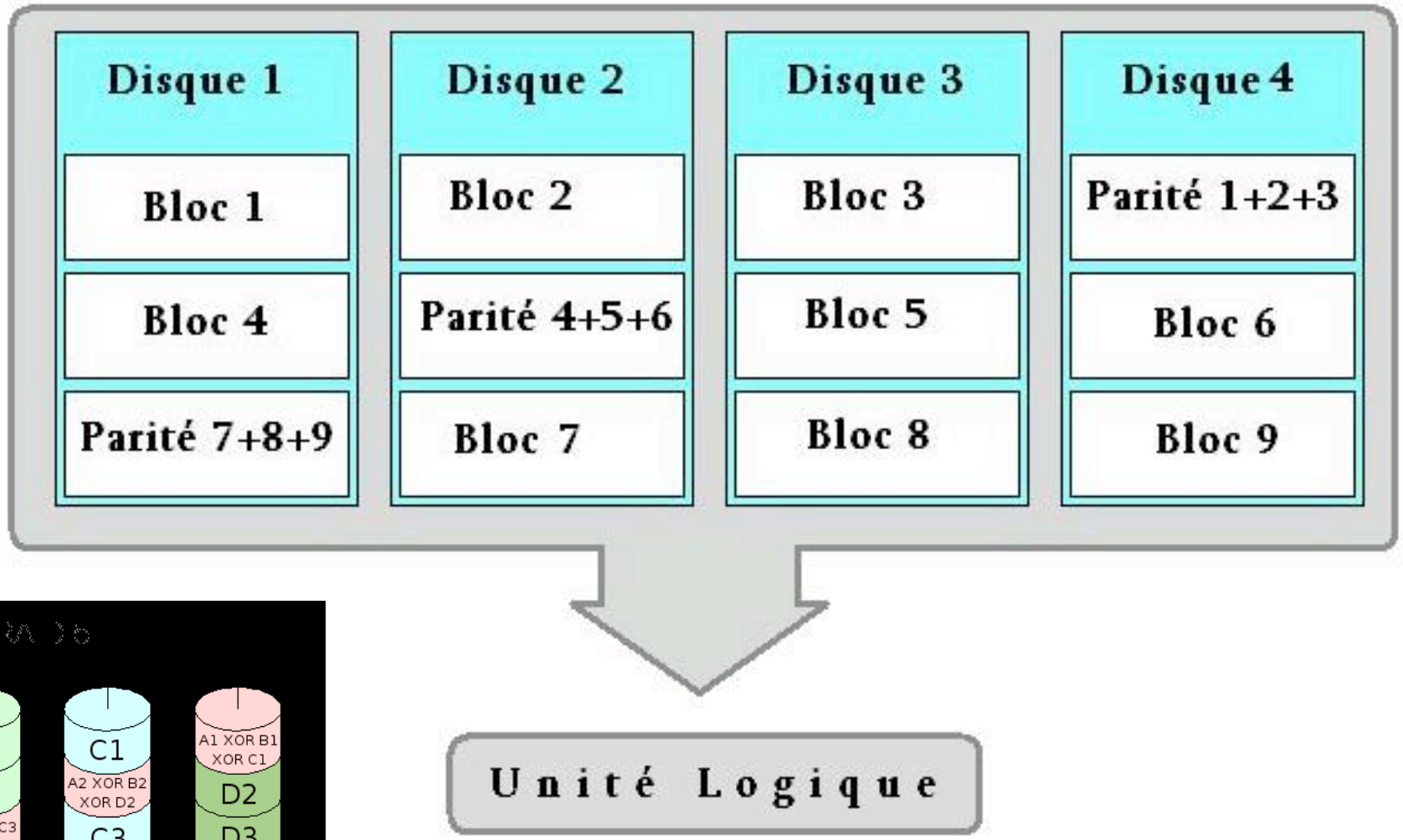
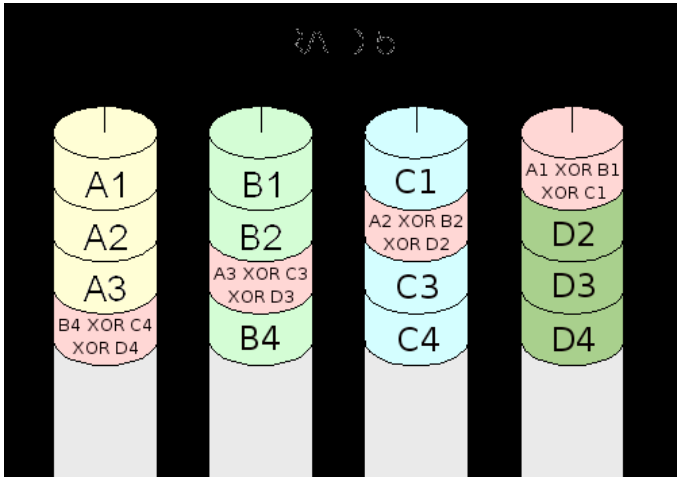
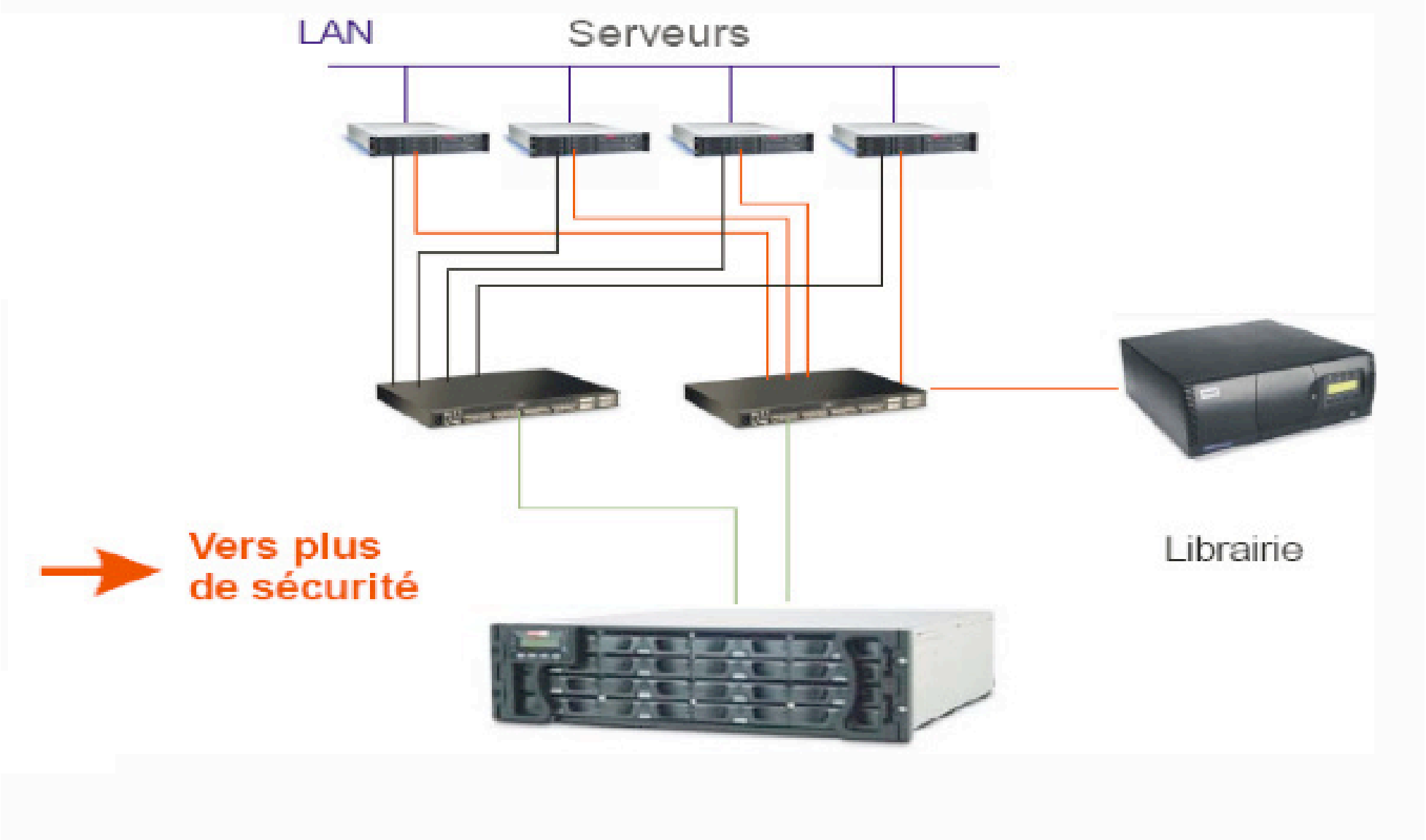


Schéma RAID 5

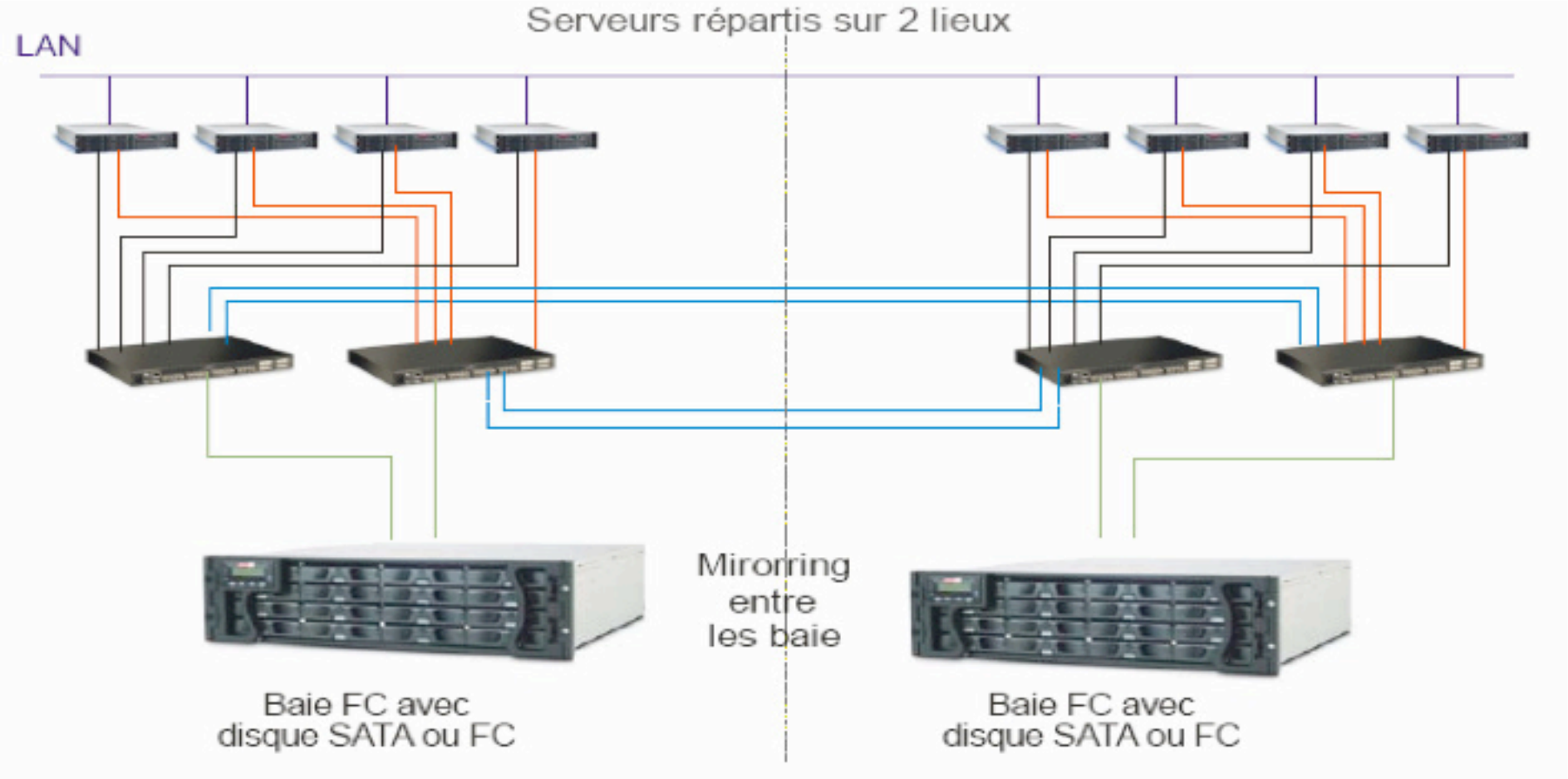




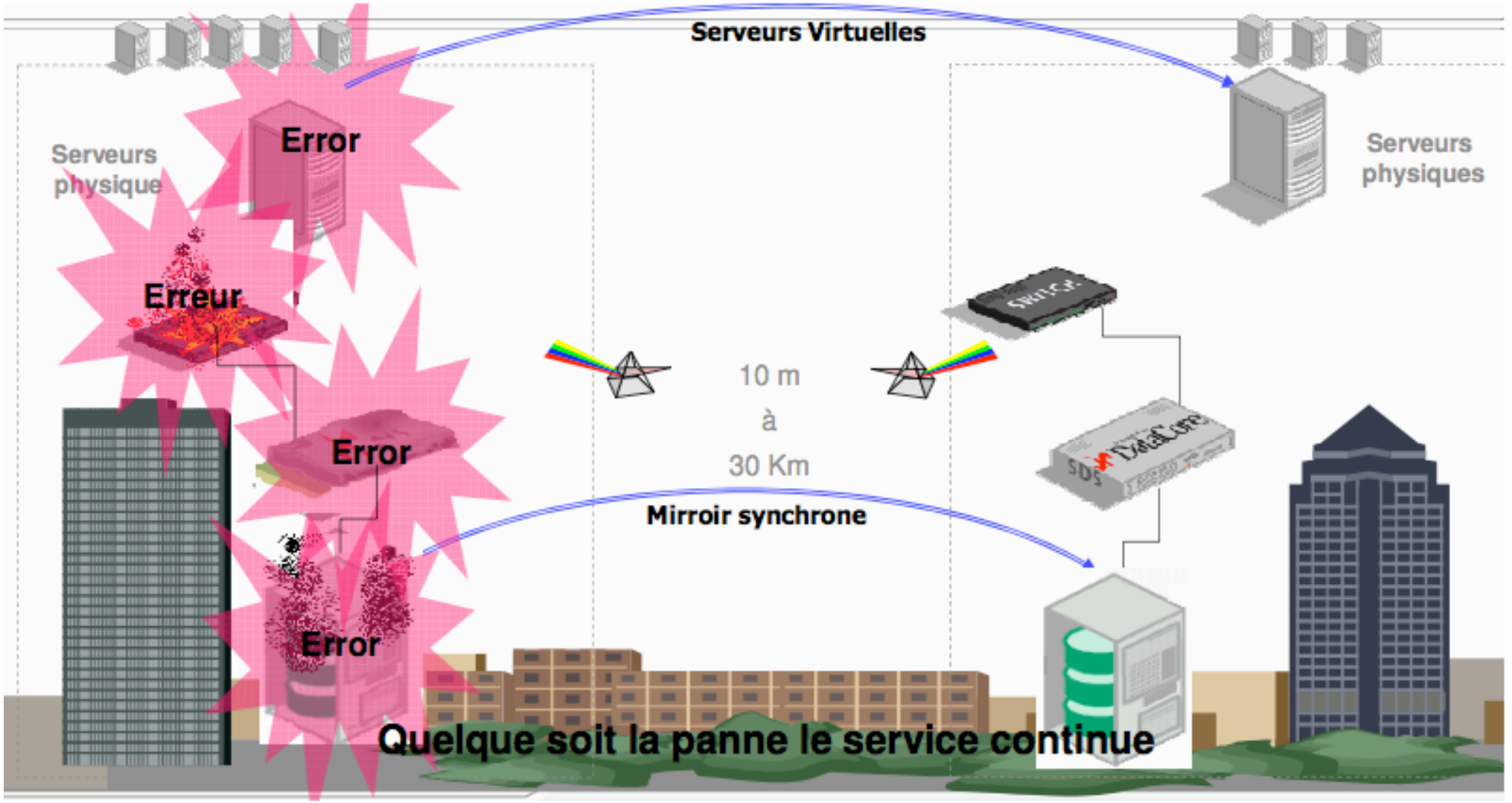
Redondance d'accès



Réplication de site



Haute disponibilité



Quelques éléments de choix

Type de disque

- SATA (consomation)
- SAS (perf prix)
- Fibre chanel (attention au 2Gb)

Type d’attachement

- Fibre chanel
- ISCSI

Comparaison FC/ISCSI

	FC	iSCSI
Bande passante	4 GB par lien 400 MO/s 8 GB par lien 800 MO/s	1 GB par lien – 80 MO/s 10 GB par lien – 1000 MO/s
Latence	10 % plus performant	Optimisation nécessaire Égal sur 10 GB
Coût	cher	Switch Layer 3 conseillé (4 x moins cher) et cartes Ethernet IP
Utilisation du serveur	0 % avec HBA	10 % (Dual core) – 0 % avec HBA
Mise en place	Complexe, Modification à froid Peu d'automatisation	Simple Modification à chaud Beaucoup de fonctionnalités
Future	10 GB	Généralisation du 10 GB
Il est possible d'avoir un réseau mixte FC et iSCSI		

Exemple des systèmes EMC - Clarion

Pour un san EMC :

- Des processeurs = SP
- Une alimentation secourue redondante
- Une baie de disque = DAE
- Des disques
- Mais aussi
- Des cartes HBA ou Gigabit
- Un switch dédier (Fibre chanel ou giga ethernet)

Chapitre 4 :

Protection des données système

Système windows

- Créer un fichier qui contient le mbr
- Créer des point de restauration du système
- Utiliser plusieur controleur active directory (attention au DNS)
- Exporter l'active directory en mode texte
  csvde -f test.csv
  difde -f test.ldf

Système Unix

- Le mbr
  dd bs=512 count=1 if=/dev/sda
  of=/$HOSTNAME.mbr
- La table des partitions
- Les fichiers système sont dans :

  /etc => hosts,passwd,shadow,exportfs,fstab, ….

  /boot =>Menu.lst, grub.conf, ….

Cas particulier des VM

Deux méthodes :

- Snapshot + backup classique
  Gros volume de donnée
  restauration simple
- Sauvegarde de l'intérieur
  logiciel de sauvegarde qui tourne sur le système hote
  petit volume de données
  restauration + complexe
  => on peut utiliser les deux …

Chapitre 5 :

Protection des données utilisateurs

Centralisation des données

- Elle est nécessaire pour la sauvegarde
- Elle protége l'utilisateur contre lui même
- Elle doit être imposé ou contrôlé
- Elle permet une reprise rapide
- Elle permet un contrôle global des droits
- Il faut un réseau performant

Gestion des droits utilisateurs !

L’organisation décide de tout !

Il faut trois type d’espace :

- Un espace privé
- Un espace public en lecture
- Un espace de travail en groupe

On utilise pour cela les ACL

Protection de la messagerie

La seul solution est serveur UNIX + IMAP :

- Les boites sont dupliqué entre client et serveur
- Les boites sont stocker dans les /home sur UNIX
- Les mails en attente sont dans /var/spool/mail
- Les mails en spool /var/spool/mqueue

=> Avec les trois éléments la sauvegarde est intègre !

Chapitre 6 :

Protection des SGBD

Rôle et risques

Coeur du Système d'informations :

- Stockage de données importantes/métiers :
  - données clients (concurent)
  - données produits (respect des commandes)
  - gestion de la paie (problème avec le FISC)
  - comptabilité (notes de frais).
- Peu d'interruption de services souhaitables :
  - ne pas prendre de risques avec des mises à jour logiciel
  - ne pas redémarrer le SGBD sans raison.

Enjeux

Les SGDB ajoute de nombreuse fonctionalité aux système :

- Ils permettent un accès à distance au données
- Ils permettent des recherche rapide et ciblé
- Ils sont exploité par plusieurs utilisateurs
- Ils accèdent aux données en lecture est écriture

Leurs protection est donc multiple !

Attaques sur les SGBD

- Utilisation du système
  mauvaise sécurisation des accès
- Utilisation du mode client serveur
  utilisateur par défaut/pas de filtrage réseau
- Attaque brute
  uilisation de bug faute de mise à jours
- Rebond vers le système
  utilisation de jointure vers des fichiers :
  SELECT nom from blabla where id = 1 union select password from users LIMIT 2,1

Sécurisation des accès :

- Suppression des mots de passe par défaut
- Recompilation du logiciel ou patch
- Utilisation d'un utilisateur système sans droit
- Restriction des droits sur les fichiers via ACL
- Serveur de bd en accès restreind au DBA
- Restriction des accès au réseau
  pas de routeur, skip-networking, iptable
- Restriction de l'usage du DNS
  - Dépendance externe
  - Injection sql

Sécurisation des data

- Utilisation de redondance disque :
  - RAID5 local
  - SAN en chemin redondant
- Fichier de log
- Snapshot
- Serveur réplique
- Limitation des tailles => datawarehouse

Sauvegarde des sgbd

- Utilisation de réplique
  blocage réplique/sauvegarde brut/ratrapage
- Verrouillage en écriture/dump
- Sauvegarde incluse dans l'application
- Snapshot en SQL

De préférences :

- Utilisez les outils du SGBD
- Puis les outils du système

Exemples : mysql snapshot

Verrouillages des tables

```
mysql> FLUSH TABLES WITH READ LOCK;
```

Création du snapshot

```
bash% mysqldump --all-databases --
lock-all-tables >dbdump.db
```

Déverrouillage

```
mysql> UNLOCK TABLES;
```

Création d'une réplique mysql

- installer deux mysql
- Créer un utilisateur de réplication
- Configurer le slave
- Récupérer l'état des logs binaires
- Dumper les bases
- Injecter les bases sur le slave
- Lancer la réplication
  `CHANGE MASTER TO`

http://dev.mysql.com/doc/refman/5.1/en/replication-howto.html