

Les systèmes de fichier.

Outils pour enregistrer et organiser
les données électroniques.

- Les système de fichier locaux (hfs+, ntfs, xfs, zfs ...)
- Les système de fichier réseaux (NFS, CIFS, AFP)
- Les système de fichier pour support optique ISO9660 (Joliet, Rock ridge, ...) et le nouveau UDF (RW)
- Les système de fichier de cluster (pvfs2, Lustre, cfs, gfs ...)
- Les système de fichier virtuels (procfs, vmfs, VFS)
- ...

Ce dont je vais parler

Préambule: Les données sont ce que vous avez de plus précieux

... Faites des sauvegardes sur d'autres média!

Les technologies de disque

Comment est organisé un système de fichier

Les différences fondamentales

Les nouveautés

Le futur

Interfaces

– SCSI

- Haut de gamme en production (15 trs/mn)
- Ordonanceur E/S

– ATA

- prix

Flash (SSD) vs technologie magnétique

- Robustesse (portable)
- Temps d'accès
- Taille / capacité (400 Go sur CF) / consommation
- Il faudra penser de nouveaux systèmes de fichiers (JFFS2?)

Bus parallèle vs série

De 150 à 300 E/S par seconde

Lors du formatage les disques durs sont segmentés en bloc (4ko -32ko)

Les données (fichiers) sont identifiés par un numéro unique (le numéro d'inode).

Les métadonnées (Inodes) (\neq ressource fork)

- Droit d'accès
- Dates
- Propriétaires
- Les liens
- Les type (fichier, lien, répertoire)
- Tailles de bloc

Table d'allocation (lien entre fichiers et bloc disques)

- 16 bits HFS vs 32 bits HFS+

Différents catalogue d'indexation (B-Tree)

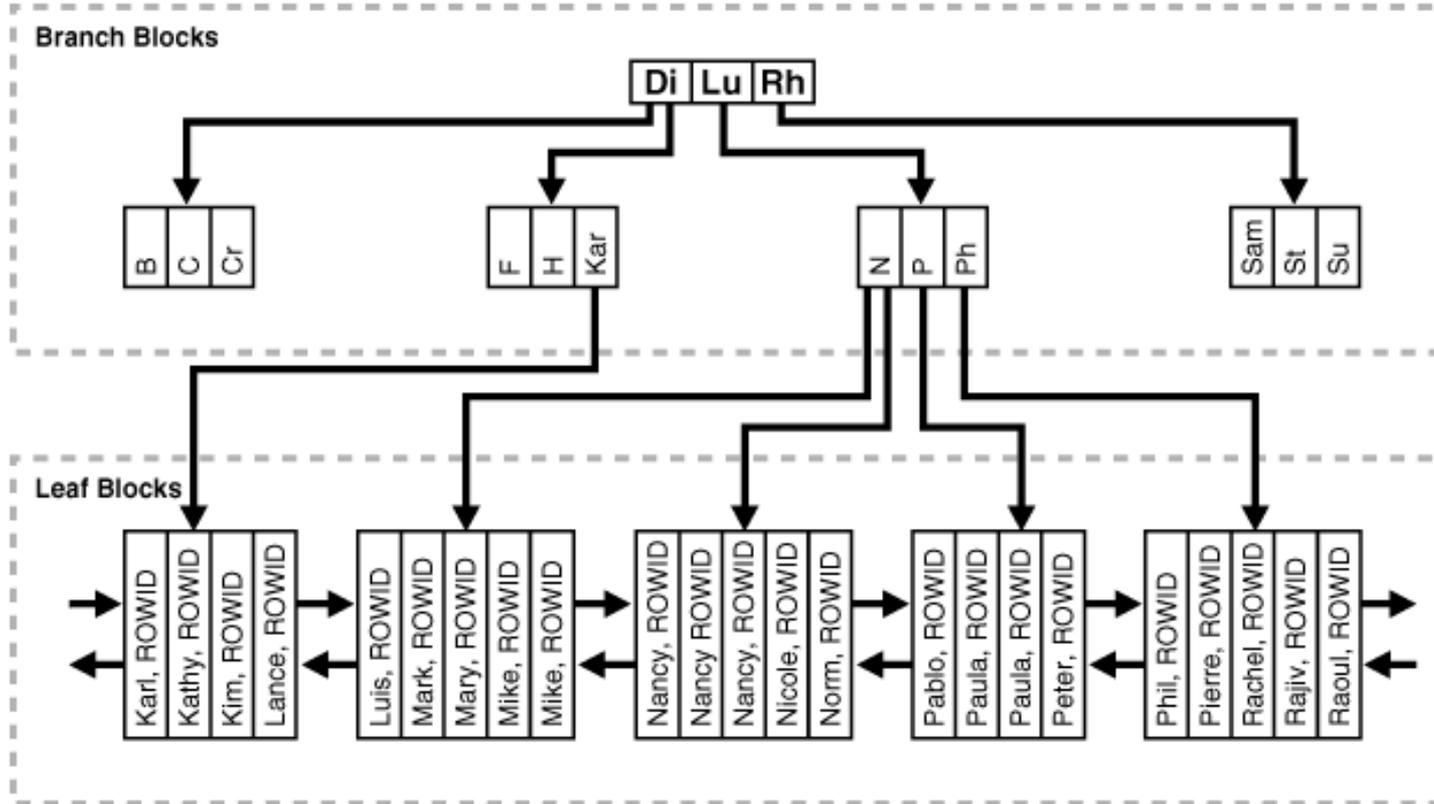
Les inodes : exemple Unix XFS

```

madrid:/tmp # stat network
  File: `network'
  Size: 4096          Blocks: 8          IO Block: 4096   directory
Device: fd00h/64768d Inode: 1049154     Links: 2
Access: (0755/drwxr-xr-x)  Uid: (  0/      root)   Gid: (  0/      root)
Access: 2007-12-17 21:33:36.678447245 +0100
Modify: 2007-12-04 14:16:38.556869599 +0100
Change: 2007-12-04 14:16:38.556869599 +0100
madrid:/tmp # stat sync.log
  File: `sync.log'
  Size: 29707180      Blocks: 58024      IO Block: 4096   regular file
Device: fd00h/64768d Inode: 430          Links: 1
Access: (0644/-rw-r--r--)  Uid: (  0/      root)   Gid: (  0/      root)
Access: 2007-11-20 13:44:49.062474805 +0100
Modify: 2007-11-20 13:47:45.940169798 +0100
Change: 2007-11-20 13:47:45.940169798 +0100
madrid:/tmp # █

```

Les B-tree



La robustesse

- Attention avec les dernières nouveautés
- La journalisation

L'intégration dans l'OS

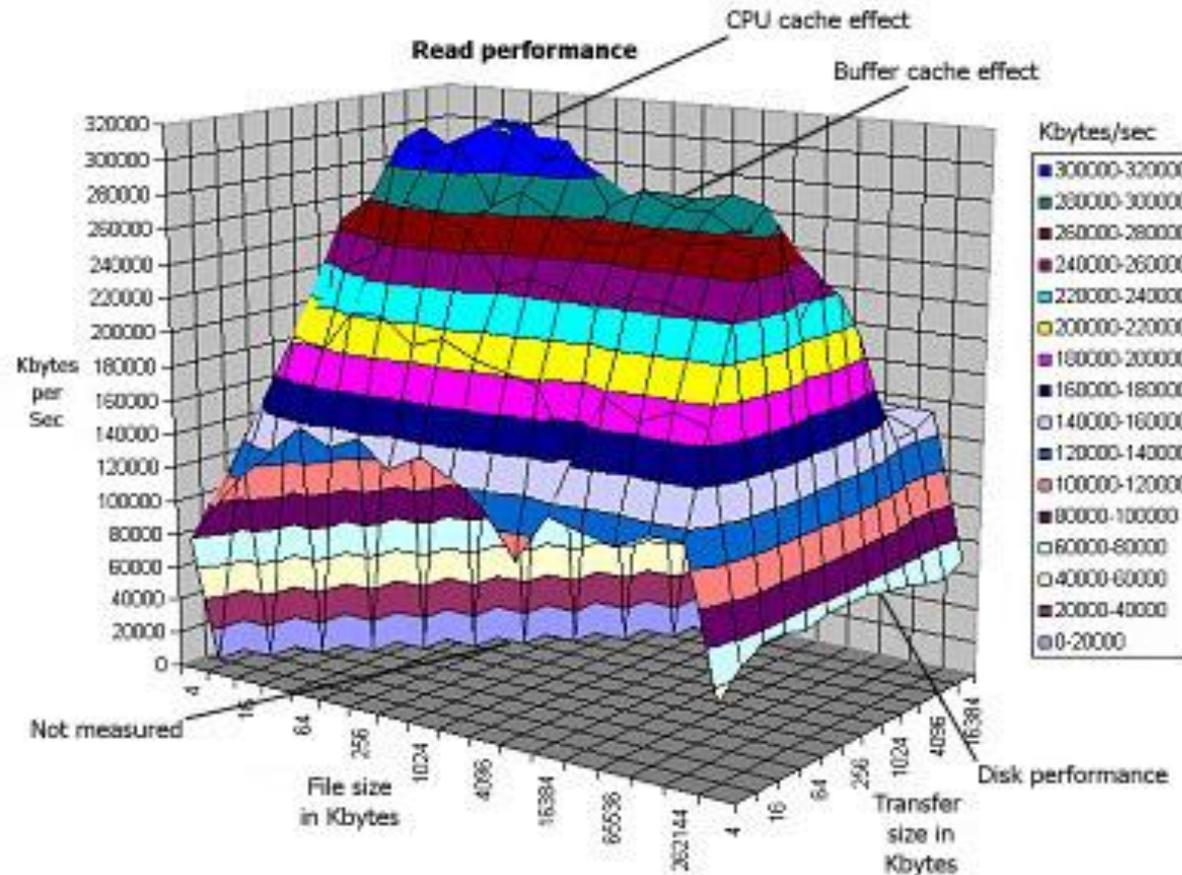
Les outils d'accompagnement

Les performances

- La fragmentation
- L'indexation
- La rapidité des opérations (benchmark)

Les capacités (12, 16, 32 - HFS+, 64 XFS, 128 bits - ZFS)

Tests avec un outil spécialisé : iotzone



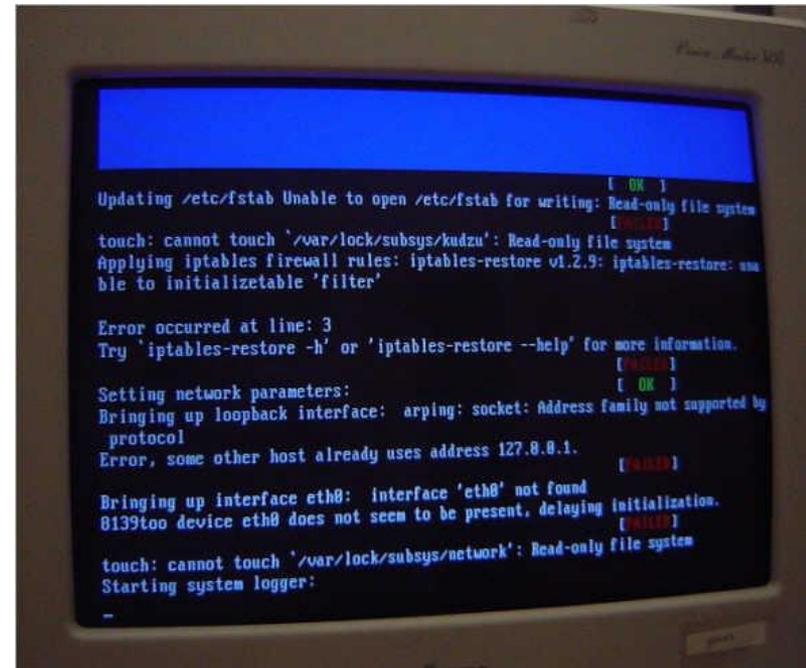
La journalisation

Technologie qui vient des bases de données

Toutes les opérations de modifications sont écrites dans un journal et « committée » on parle alors de transaction.

C'est la fin des fsck! et la plus intéressante des améliorations.

Journalisation
+ Raid
= fin des problèmes
de corruption.



Cryptage des données (NTFS5 / EFS sous windows)

- Bonne idée ? Beaucoup de pleurs

Redimensionnement à chaud (Volume Manager)

- Passer de 4Go à 40Go sans arrêter les opérations, très utilisé en entreprise
- Aggrégation de disque

Les ACLs étendues

Les raids

Les snapshots

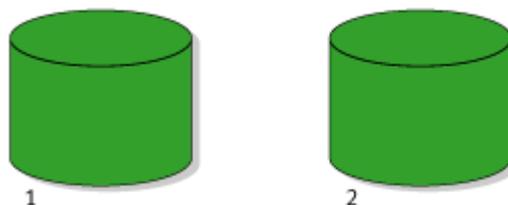


Cela n'est pas comparable à une véritable sauvegarde

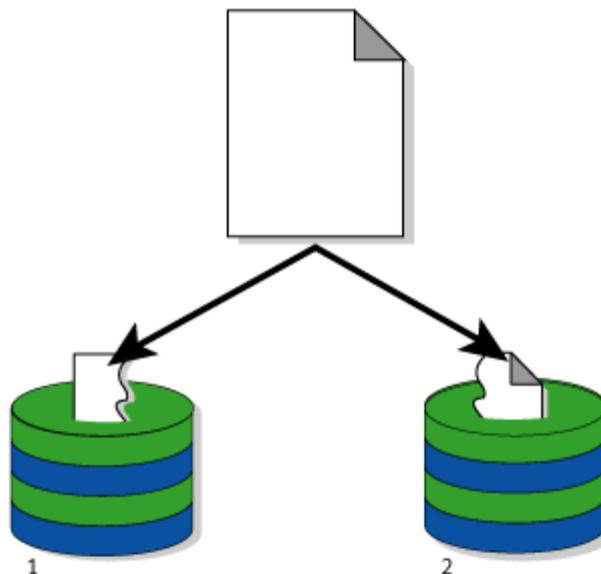
Privilégié le Raid matériel

Différents niveaux de Raid

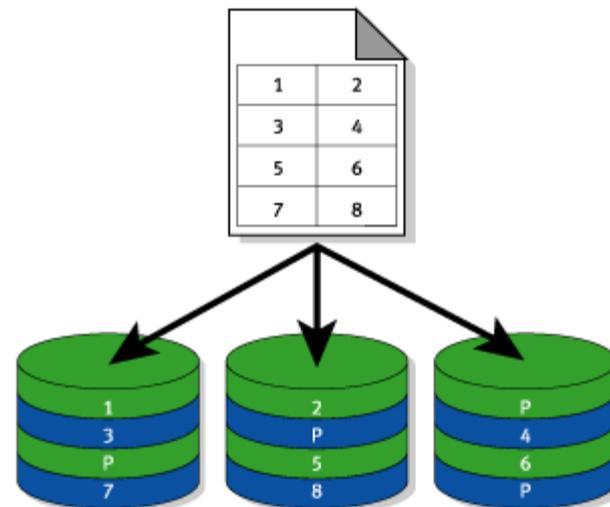
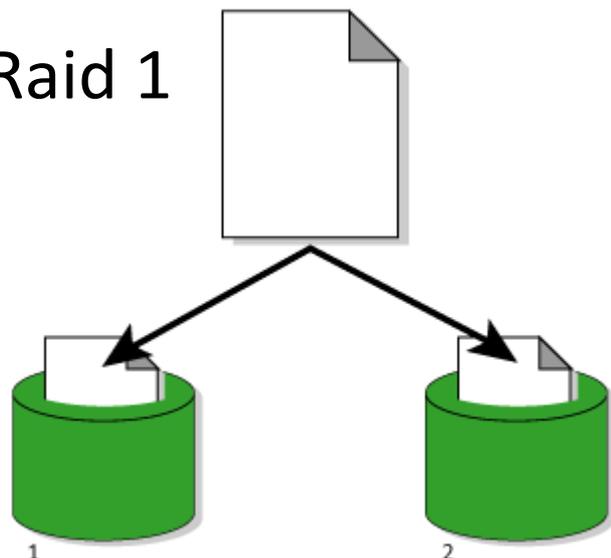
– Pas de Raid



– Raid 0



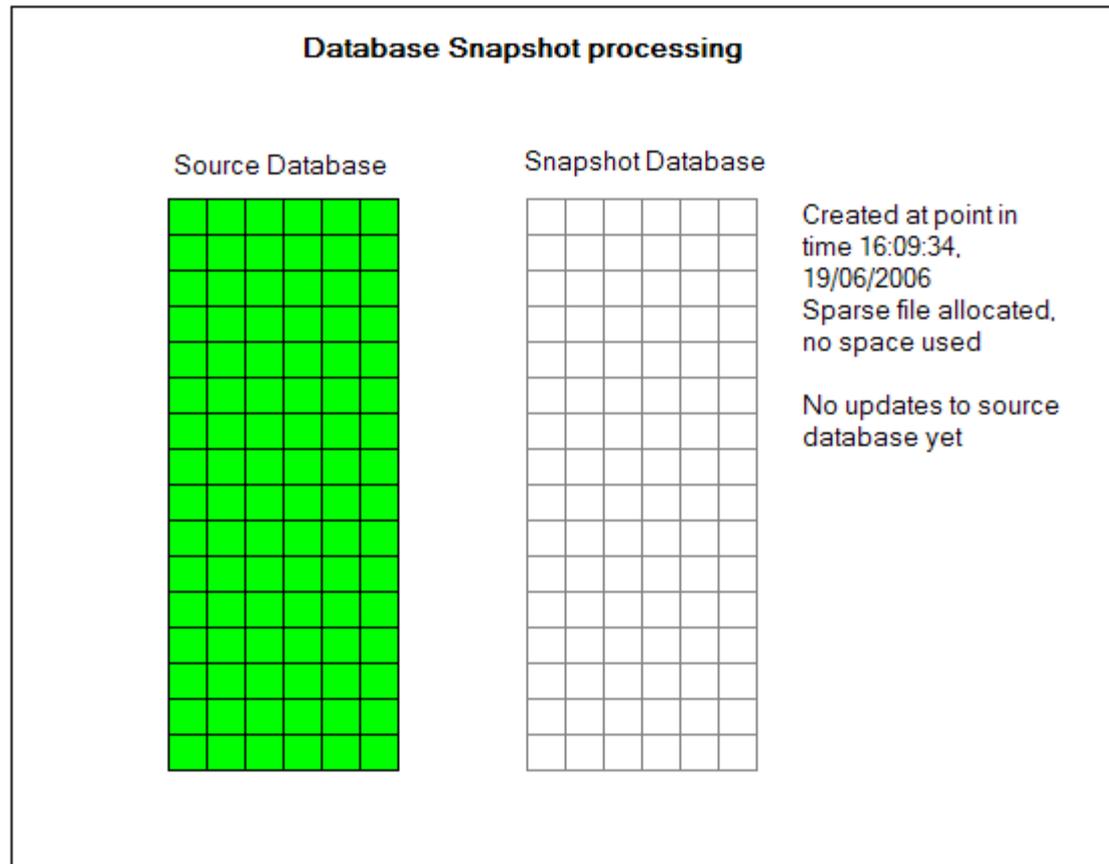
– Raid 1



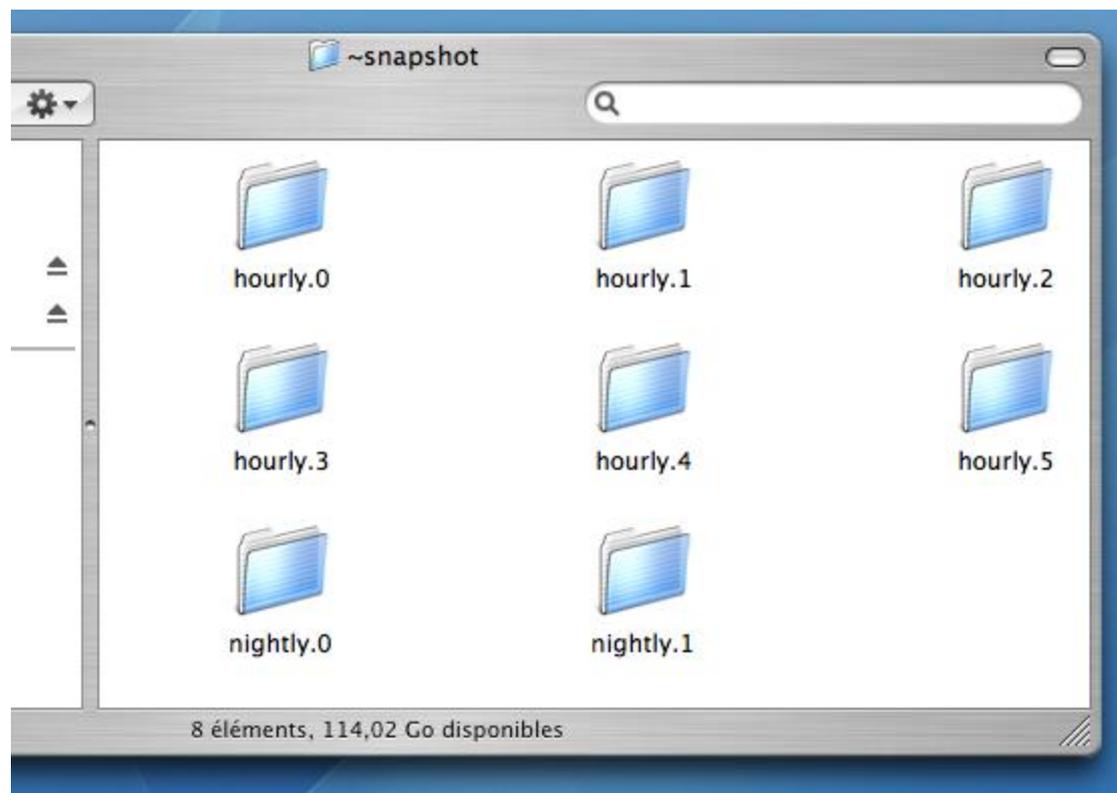
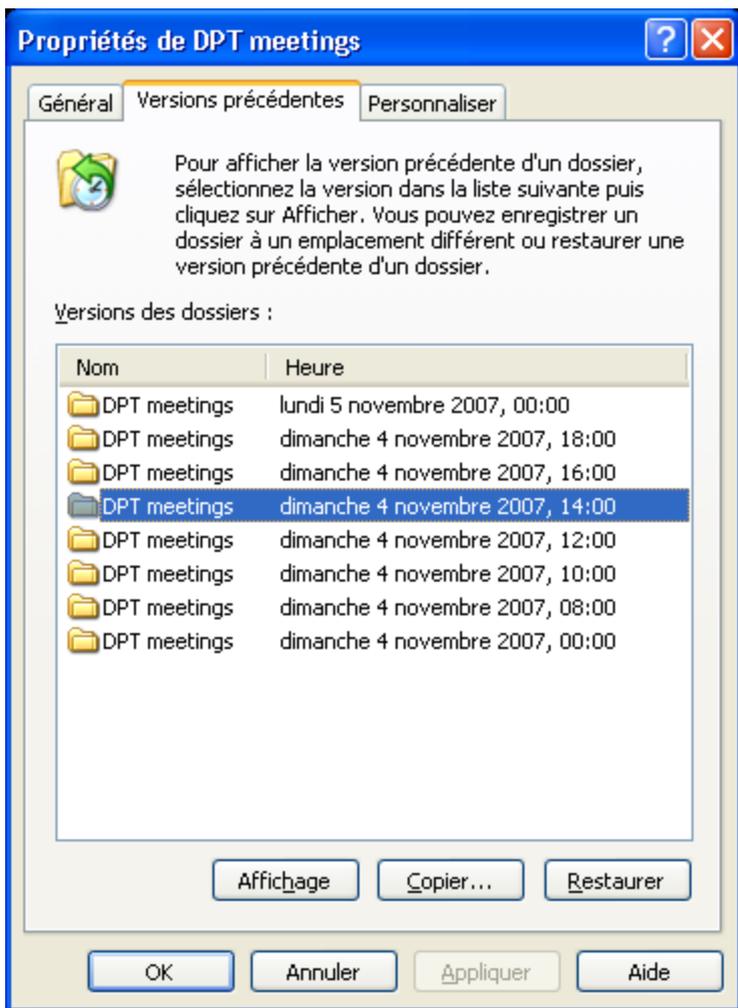
– Raid 5

- Basé sur XOR consiste à déterminer si sur n bits de données considérés, le nombre de bits à l'état 1 est pair ou impair. Si le nombre de 1 est pair, alors le bit de parité vaut 0. Si le nombre de 1 est impair, alors le bit de parité vaut 1

Outil de sauvegarde des modifications.



Interfaces utilisateur



Le future du Mac ZFS ?

Mac OS 8.1 : HFS+

Mac OS 10.4 : Journalisation par défaut + HFSX

Mac OS 10.5 : ZFS alpha, en lecture seule

ZFS (Sun Solaris) :

- Full 128 bits
- Journalisation
- Raid-Z (Raid 5 à géométrie variable – chaque bloc a son propre groupe raid), tous les niveaux de Raid sont possibles
- VM inclut (ajout de disque pour augmenter l'espace ou les performances a chaud)
- Snapshots (copy-on-write)
- Checksum puis réparation automatique des blocs

Merveilleux mais ... que fait on pour les disques Flash ?