

Maxtor®



NEC/CI Angers

8 Mars 2001

Technologie disque

Master Boot Record

Partition(s)

Commandes ATA

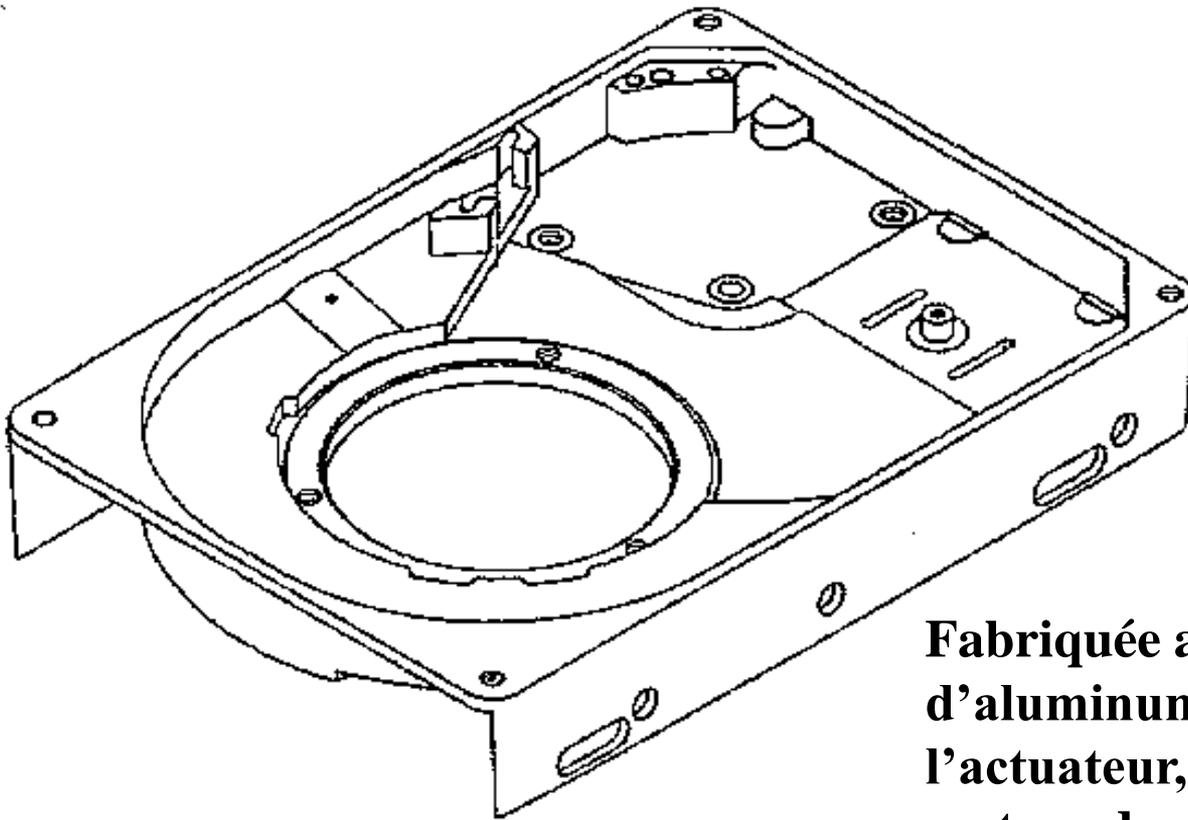
Technologie Disque

Vue globale



Technologie Disque

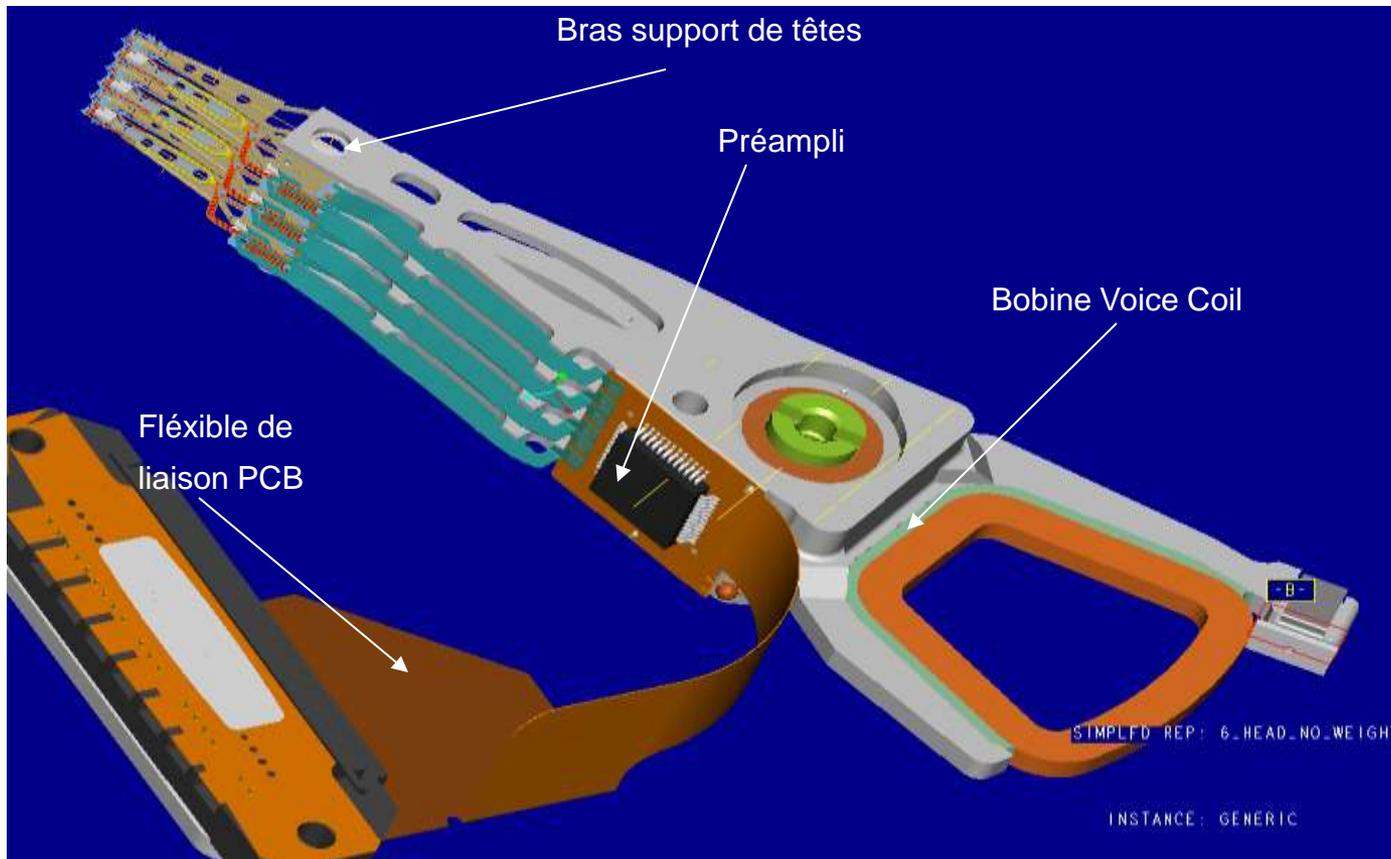
Base mécanique



Fabriquée avec un alliage d'aluminium, elle reçoit l'actuateur, l'ensemble têtes et le moteur de rotation.

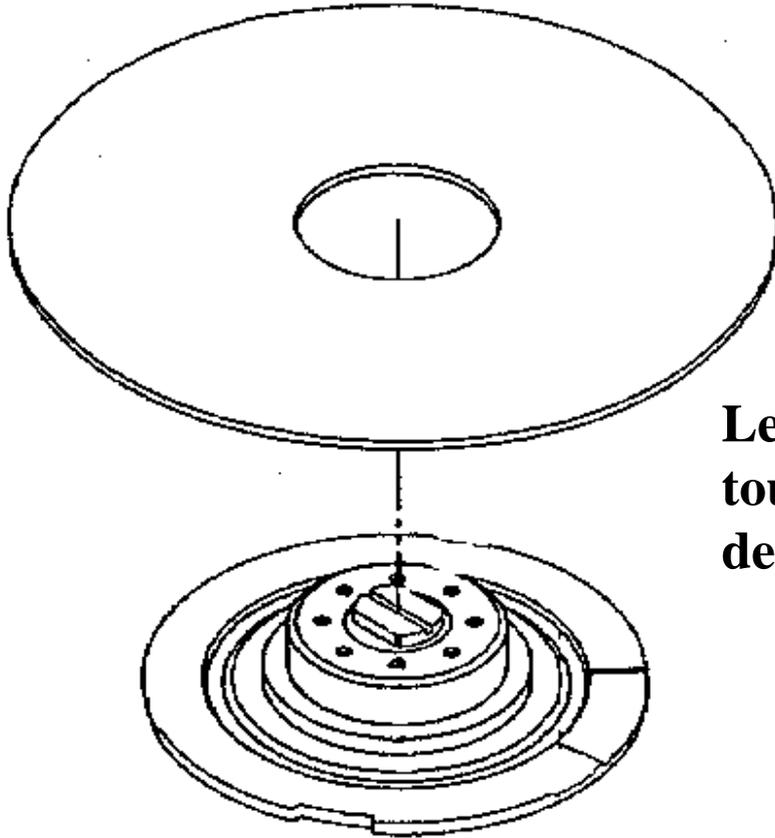
Technologie Disque

Ensemble actuateur



Technologie Disque

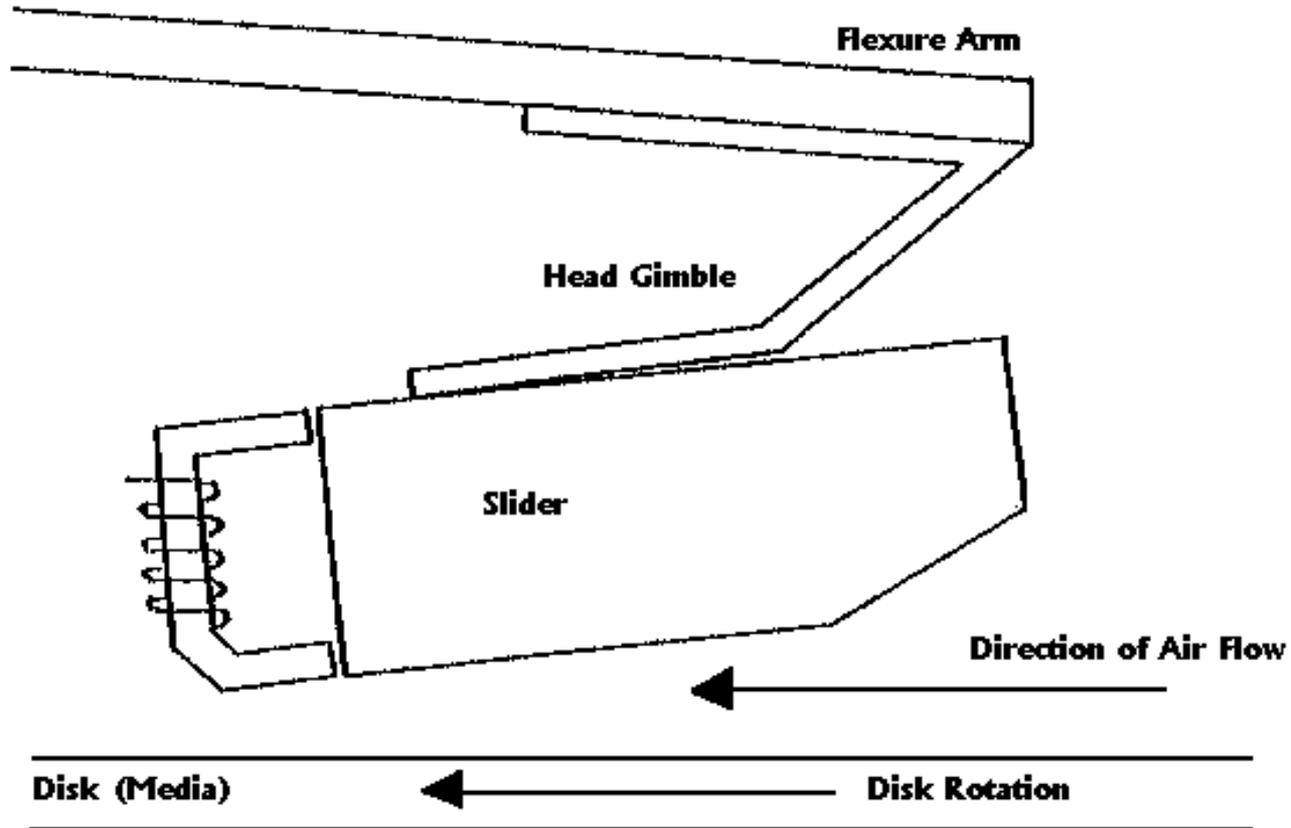
Moteur / Plateau(x)



**Le moteur (roulements à billes)
tourne à 5400 Tpm ou 7200 Tpm,
dependant de la famille du produit.**

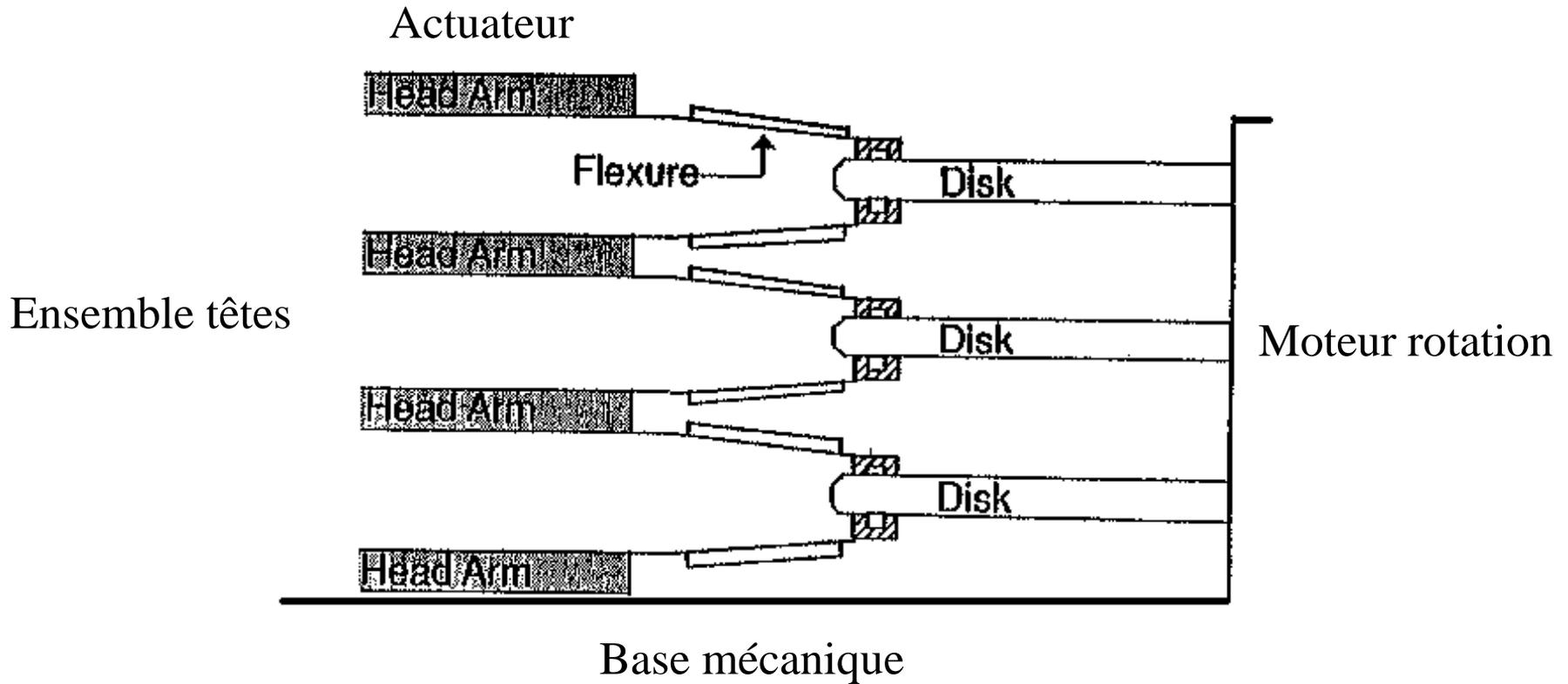
Technologie Disque

Ensemble têtes



Technologie Disque

Ensemble têtes



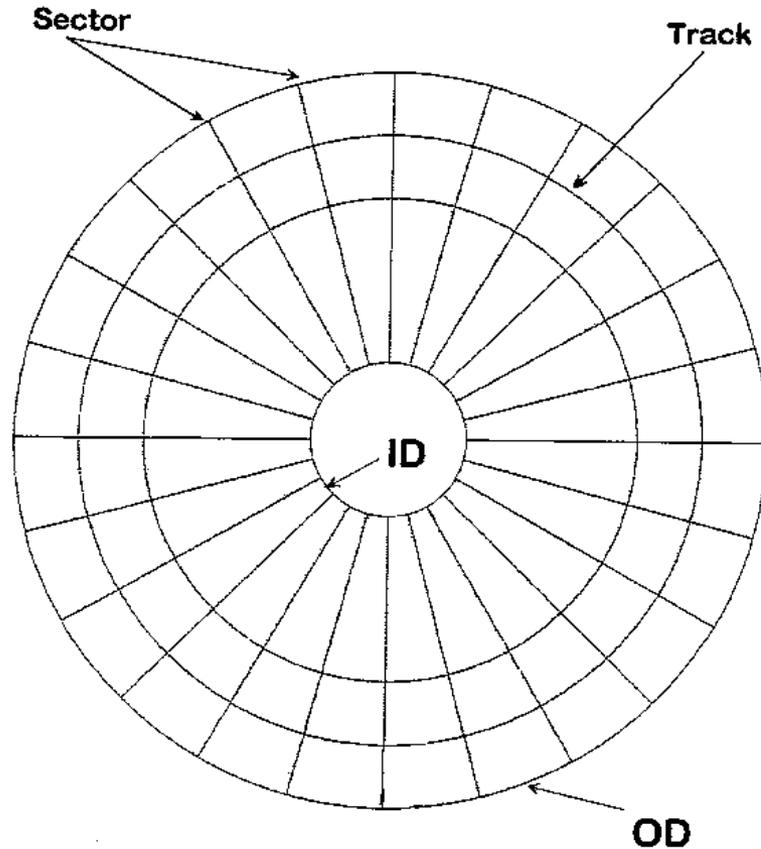
Technologie Disque

Printed Circuit Board Assembly



Technologie Disque

Adressage des données



Le disque est découpé en pistes (ou cylindres si plusieurs têtes) et secteurs pouvant être adressés directement. Pour lire, on déplace la tête sur la piste cherchée (seek), puis on attend que le secteur passe sous la tête (latency).

Master Boot Record

- **Situé sur le Cylindre 0, Tête 0, Secteur 1.**
- **Chaque OS peut créer un “Master” secteur de boot (FDISK).**
- **Mais chaque OS a son propre format de secteur de boot !**
 - situé dans “sa” partition
- **Après le POST, l’interruption 19h est appelée :**
 - Essai de lecture sur le floppy (cyl 0, t 0, s 1)
 - Si pas de secteur de boot, lecture sur le disque (cyl 0, t 0, s 1)
 - Si secteur de boot, chargement en 0000:7C00
 - Recherche d’une partition active dans la table
 - Si elle existe, le secteur de boot de CETTE partition est lu en mémoire et le MBR lui “donne le contrôle”.

Master Boot Record

- **Séquence du “Master” secteur de boot :**
 - Chargement en 0000:7c00, puis relogement en 0000:0600
 - Recherche d’une partition active
 - si aucune active, message d’erreur **“Rom Basic”**
 - si plusieurs actives, message **“Invalid partition table”**
 - Lecture du secteur de boot de la partition active (en 0000:7c00)
 - si erreur de lecture à ce niveau, **“Error loading system”**
 - Verification de la signature (“AA55”)
 - si pas de signature, **“Missing operating system”**
 - Toute autre erreur lance une boucle infinie
 - `HANG: JMP HANG ;Boucle infinie`
 - Si ok, jump en 0000:7c00 et exécution du secteur de boot de la partition (différent du “Master” secteur de boot)

Master Boot Record

```

■   OFFSET 0 1 2 3 4 5 6 7 8 9 A B C D E F *0123456789ABCDEF*
■   000000 fa33c08e d0bc007c 8bf45007 501ffbf0 * .3. . . . | . . P . P . . *
■   000010 bf0006b9 0001f2a5 ea1d0600 00bebe07 * . . . . . . . . . . *
■   000020 b304803c 80740e80 3c00751c 83c610fe * . . . < . t . . < . u . . . . *
■   000030 cb75efcd 188b148b 4c028bee 83c610fe * . u . . . . . I . . . . . *
■   000040 cb741a80 3c0074f4 be8b06ac 3c00740b * . t . . < . t . . . . < . t . *
■   000050 56bb0700 b40ecd10 5eebf0eb febf0500 * V . . . . . ^ . . . . . *
■   000060 bb007cb8 010257cd 135f730c 33c0cd13 * . . | . . W . . _ s . 3 . . . *
■   000070 4f75edbe a306ebd3 bec206bf fe7d813d * O u . . . . . . . . . } . = *
■   000080 55aa75c7 8bf5ea00 7c000049 6e76616c * U . u . . . . . | . . Inval *
■   000090 69642070 61727469 74696f6e 20746162 * i d p a r t i t i o n t a b *
■   0000a0 6c650045 72726f72 206c6f61 64696e67 * l e . E r r o r l o a d i n g *
■   0000b0 206f7065 72617469 6e672073 79737465 * o p e r a t i n g s y s t e *
■   0000c0 6d004d69 7373696e 67206f70 65726174 * m . M i s s i n g o p e r a t *
■   0000d0 696e6720 73797374 656d0000 00000000 * i n g s y s t e m . . . . . *
■   0000e0 00000000 00000000 00000000 00000000 * . . . . . . . . . . *
■   0000f0 TO 0001af (Rempli de 000000)
■   0001b0 00000000 00000000 00000000 00008001 * . . . . . . . . . . *
■   0001c0 0100060d fef83e00 00000678 0d000000 * . . . . . > . . . . x . . . *
■   0001d0 00000000 00000000 00000000 00000000 * . . . . . . . . . . *
■   0001e0 00000000 00000000 00000000 00000000 * . . . . . . . . . . *
■   0001f0 00000000 00000000 00000000 000055aa * . . . . . . . . . . U . *

```



Debut du code
(CLI, XOR AX,AX ...)



Fin du code (Jmp 7C00)



Debut messages



Debut de la partition
en 0000:07be. Ici, une
seule partition, active.

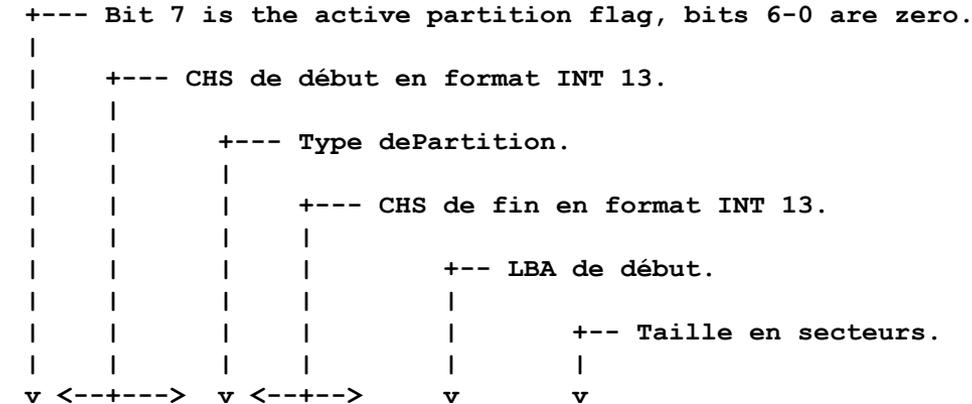


Signature

Partition Table

- Aucun document ne couvre complètement le fonctionnement de FDISK, mais tous suivent à peu près les mêmes idées.
- Toutes les tables de partition ont le même format :
 - situé dans le cyl 0, tête 0, secteur 1
 - Les 66 derniers octets du secteur contiennent la table de partition, suivis de 2 octets de signature (voir description du MBR)
 - Les partitions étendues sont “imbriquées” les unes dans les autres
- Chaque entrée de la table comprend 16 octets :
 - Début et fin de la partition en CHS (!?)
 - Début de la partition en LBA (!)
 - Taille de la partition en secteurs
 - “Type” de partition et “Active/Inactive” status
- 4 partitions primaires dans MBR, ou 3 primaires et une étendue
- La partition étendue peut contenir une “étendue” ou une “secondaire”
- Une seule partition est active à la fois

Partition Table



0	1	2	3	4	5	6	7	8	9	A	B	C	D	E	F	
DH	DL	CH	CL	TB	DL	CH	CL	LBA.....	SIZE.....							
80	01	01	00	06	0e	be	94	3e000000	0c610900							élément n°1
00	00	81	95	05	0e	fe	7d	4a610900	724e0300							élément n°2
00	00	00	00	00	00	00	00	00000000	00000000							élément n°3
00	00	00	00	00	00	00	00	00000000	00000000							élément n°4

- Ici, une partition DOS FAT, qui commence au CHS 0/1/1 (LBA 3E) et finit au CHS 294/E/3E avec une taille de 9610Ch secteurs
- Un 2eme (étendue) qui commence au CHS 295/0/1, se termine au CHS 37D/E/3E avec 34E72h secteurs
- 3 et 4 sont inutilisées...
- Les partitions sont alignées sur des cylindres !! Ce qui laisse des secteurs inutilisés, **utilisés par Disk Manager ou certains virus !**
- Le remplissage de la table peut se faire par le haut ou par le bas et l'effacement peut laisser des "trous"

Partition Table

```

CHS=0,0,1 | Master Boot Record contenant |
| la table de partition et le programme |
| de recherche de celle-ci |
| +-----+ |
| | description partition DOS FAT | | pointe vers CHS=0,1,1
| +-----+ | | pointe vers CHS=a
| | description partition OS/2 HPFS | |
| +-----+ |
| | entrée table inutilisée | |
| +-----+ |
| | entrée partition étendue | | pointe vers CHS=b
| +-----+ |
+-----+
CHS=0,0,2 | Le reste de la "piste 0" -- c'est là | :
to | que les drivers comme Disk Manager | : en principe
CHS=0,0,n | ou autre sont placés. Peut aussi | : inutilisé
| contenir des virus !!!!!!!!!!! | :
+-----+
CHS=0,1,1 | Secteur de Boot pour partition DOS | : System de fichier Dos FAT
| FAT | |
+-----+
.....

```

```

|<----- le disque entier ----->|
|
|M<---C:--->
|
|      E<---D:---><---reste de la 1ere étendue----->|
|
|              E<---E:---><---reste de la 2eme----->|
|
|                      E<-----F:----->|
|

```

Disque contenant 4
partitions FAT (C, D, E, F)

Lettres et Partitions

- **A: et B: sont réservées aux floppies, même s'il n'y en a qu'un...**
 - **Les disques durs démarrent en C: , avec une lettre pour chaque partition**
 - **Mais le DOS assigne les lettres aux partitions primaires en premier !**
 - **Seulement ensuite, il affecte les partitions logiques**
- ⇒ **Peut provoquer des changements d'affectation si on ajoute un disque ...**

Partition	HDD1 : Partition Primaire	HDD1 : Partition Logique	HDD2 : Partition Primaire	HDD2 : Partition Logique
Avant	C:	D: , E:		
Après	C:	E: , F:	D:	G:

- ⇒ **Pour éviter cela, il est légal de créer seulement une partition étendue avec tous les disques logiques dedans.**

FAT, Cluster, Secteur

- Taille des secteurs : 512 octets -> gérer un grand nombre de secteurs est difficile et la performance est faible, d'où l'introduction du Cluster
- Cluster : Regroupe un ensemble de secteur, facilite la gestion
- FAT : stocke l'information relative aux clusters, une entrée par cluster.

	FAT12	FAT16	FAT32
Utilisation	Floppy	Petit disque	Gros disque
Taille des entrées	12 bits	16 bits	32 bits
Nbre max cluster	4086	65526	268435456
Taille cluster	0,5kB à 4 kB	2 kB à 32 kB	4 kB à 32 kB
Taille max Volume	16736256	2147123200	2 TB

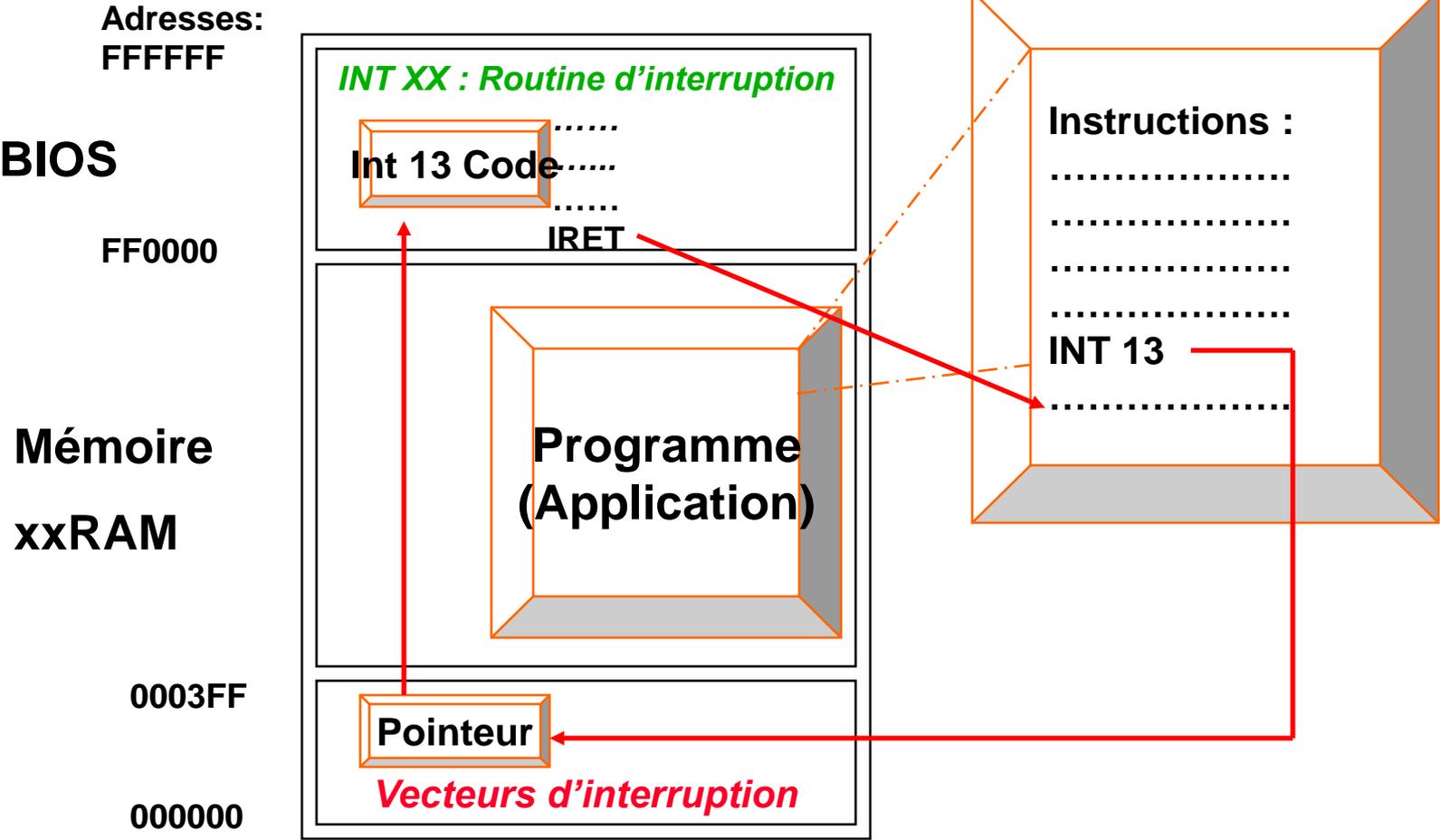
FAT vs Taille cluster

- Le cluster de petite taille évite de perdre trop d'espace disque, si l'on utilise des petits fichiers
- Mais la taille de la FAT augmente avec la taille du disque à taille de cluster constante
- Or, la FAT est accédée très souvent (pour tout accès à un fichier) et donc mise dans le cache, si sa taille le permet.
- D'où le tableau suivant pour la FAT32 :

Taille Cluster	Partition Min	Partition Max
4KB	0,5 GB	8 GB
8KB	8 GB	16 GB
16KB	16 GB	32 GB
32KB	32 GB	64 GB ?

- Ceci permet de garder la taille de la FAT à 8MO pour la partition maximum.

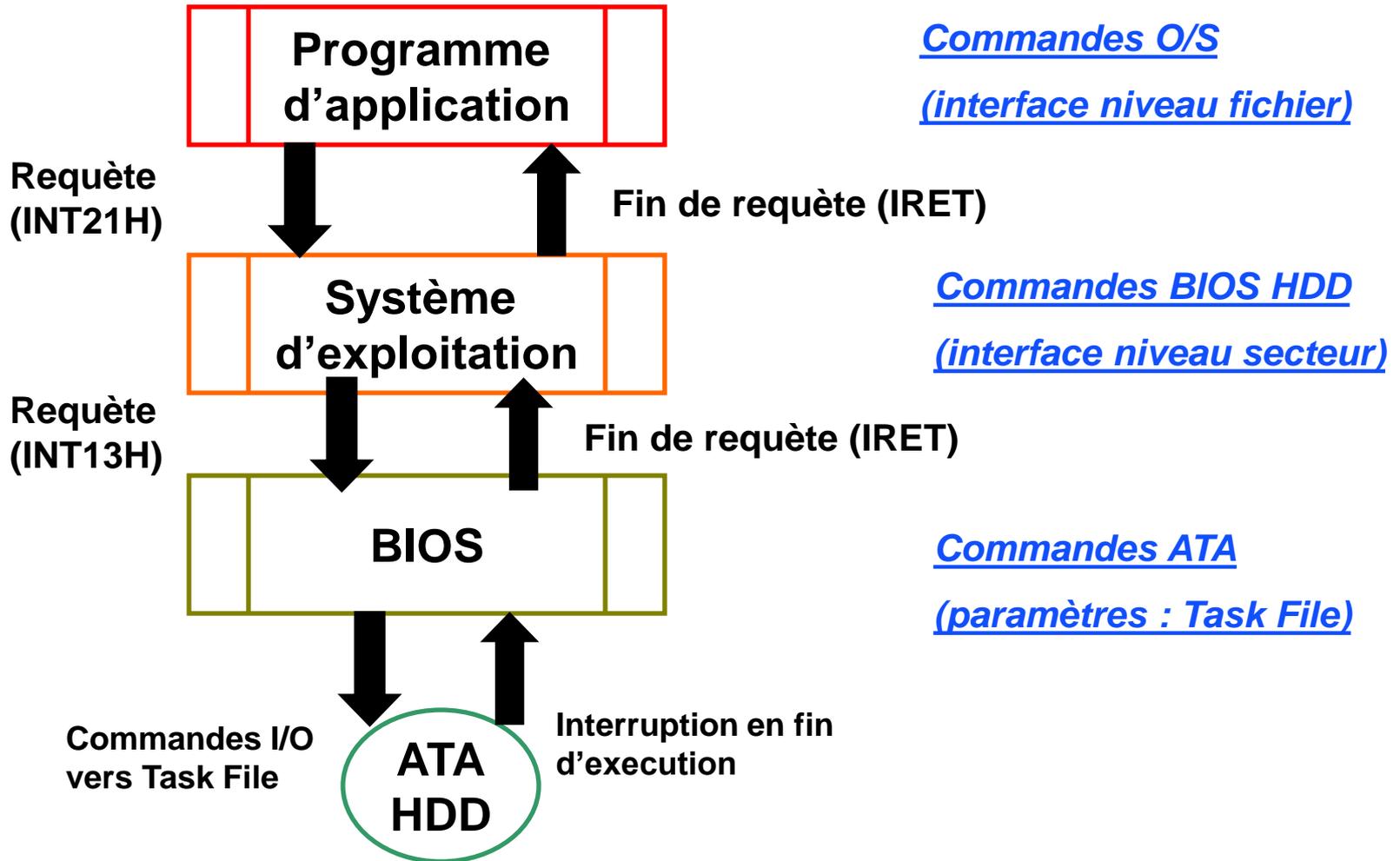
ATA et PC : Interruptions (rappel)



ATA et PC : Interruptions (rappel)

- **La table des vecteurs d'interruption (1024 octets) contient les pointeurs (double mot de 16 bits) vers les routines BIOS qui gèrent chaque élément.**
- **Attention, la valeur segment/offset est à la mode Intel, offset en premier**
- **Les vecteurs peuvent pointer sur le BIOS (voir ci-dessous) ou sur des routines de l'OS (interruption 21H, par exemple, non incluse dans BIOS)**
- **Quelques interruptions BIOS typiques :**
 - Clavier : INT 09H
 - Vidéo : INT 10H
 - Disque Dur : INT 13H
 - Port Série : INT 14H
 - Port // : INT 17H
 - Boot Strap Loader : INT 19H

ATA et PC : Accès au disque dur



ATA et PC : Accès au disque dur

- Commandes Application -> O/S (interface niveau fichier) :
 - langage de haut niveau du type -> “open, read, write, close, ...” orienté fichier
- Commandes O/S -> BIOS HDD (interface niveau secteur) :

- MOV AH, 02h	Code de commande BIOS (lecture,écriture, seek,..)
- MOV AL, 05h	Nombre de secteurs
- MOV CH, 02h	Numéro de cylindre (poids faibles)
- MOV CL, 41h	Numéro de cylindre (2 bits poids fort) + numéro de secteur
- MOV DH, 00h	Numéro de tête
- MOV DL, 80h	Sélection du disque
- LES DX,Buffer	Pointeur pour les données
- INT 13h	Interruption pour lancer l'exécution
- Commandes BIOS -> ATA (paramètres : Task File) : 8 registres

ATA et PC : Registres TASK File Mode CHS

Port I/O	Lecture	Ecriture	Taille
1F0h	Data	Data	16 bits
1F1h	Error	Feature	8 bits
1F2h	Nombre	secteurs	8 bits
1F3h	Numéro de	Secteur	8 bits
1F4h	Cylindre	(poids faib)	8 bits
1F5h	Cylindre	(poids fort)	8 bits
1F6h	Disque	Tête	8 bits
1F7h	Status	Commande	8 bits

ATA et PC : Registres TASK File Mode LBA

Port I/O	Lecture	Ecriture	Taille
1F0h	Data	Data	16 bits
1F1h	Error	Feature	8 bits
1F2h	Nombre	secteurs	8 bits
1F3h	Bits LBA	0 - 7	8 bits
1F4h	Bits LBA	8 - 15	8 bits
1F5h	Bits LBA	16 - 23	8 bits
1F6h	Disque	Bits 24 - 27	8 bits
1F7h	Status	Commande	8 bits

ATA et PC : Limites BIOS/Task File

	Limites BIOS (INT13h)	Limites Task File
Nombre de cylindres	10 bits	16 bits
Nombre de têtes	8 bits	4 bits
Nombre de secteurs	6 bits	8 bits
Total des Secteurs	16,777,072	268,435,456
Taille des secteurs	512 octets	512 octets
Taille disque max	8 Goctets	128 Goctets

La vieille barrière de 528 Moctets venait de la combinaison de ces 2 limites au minimum (cylindre = 10 bits, têtes = 4 bits, secteurs = 6 bits)

ATA et PC : Commandes ATA

- **Commandes Lecture :**
 - Read, Read Verify (sans transfert), Read Buffer, Read Multiple, Read DMA
- **Commandes Ecriture :**
 - Write, Write Verify (sans transfert), Write Buffer, Write Multiple, Write DMA
- **Commandes “Set/Check”**
 - Set Features, Set Multiple Mode, Set Max Mode
- **Commandes “Power” :**
 - Standby, Standby Immediate, Idle, Idle Immediate, Check Power Mode, Set Sleep Mode
- **Commandes Initialisation :**
 - Identify, Init Drive Param, Recalibrate
- **Commandes speciales :**
 - **Seek, Format, Exec Drive Diag, SMART**

ATA et PC : Commandes ATA

- **Read ou Write Multiple :**
 - transfère les données de plusieurs secteurs sans générer une interruption vers le système
 - Le nombre de secteurs transférés à chaque fois est défini par la commande “Set Multiple”
 - permet de gagner en performance, car une seule interruption et donc un seul traitement est nécessaire pour un ensemble de secteurs
- **Read ou Write DMA :**
 - permet le transfert en mode DMA, jusqu’à 100 Mo/sec
 - nécessite d’initialiser le boîtier DMA
- **Identify :**
 - permet pendant le POST de définir de quoi est capable le disque (taille, mode DMA/PIO supportés,...)

ATA et PC : Mode PIO et (U)DMA

- **Sont les deux modes de transfert entre le disque et le système**
 - PIO = Programmed Input/Output
 - UDMA = Ultra Direct Memory Access
- **PIO utilise les instructions assembleur IN et OUT (INSW, OUTSW) du processeur**
 - d'où une importante charge CPU
- **UDMA est exécuté par un contrôleur séparé**
 - il faut l'initialiser à chaque transfert, mais la charge CPU est moindre
- **UDMA ne sert que lors des transferts de données, les commandes sont exécutées en PIO**

ATA et PC : Registre “Status”

- **Bit 0** : à 1 si la commande précédente était en erreur
- **Bit 1 et 2** : Non utilisé (=0)
- **Bit 3** : à 1 indique que le buffer est prêt à transmettre
- **Bit 4** : à 1 indique que le seek est terminé
- **Bit 5** : à 1 indique une erreur d'écriture
- **Bit 6** : à 1 indique que le disque est prêt pour une commande
- **Bit 7** : passe à 1 lorsque l'on écrit dans l'octet de commande (port 1F7h), indique que le disque est occupé

- **Attention** : les registres ne peuvent être lus que si le Bit 7 est à 0

ATA et PC : Registre “Error”

- **Registre en lecture seule, détails dans manuel**
 - Bit 0 : Marque d’adresse non trouvée
 - Bit 1 : Piste 0 non trouvée
 - Bit 2 : Commande abortée
 - Bit 3 : Non utilisé (=0)
 - Bit 4 : Identifieur non trouvé
 - Bit 5 : Non utilisé (=0)
 - Bit 6 : Erreur d’ECC données
 - Bit 7 : Erreur de CRC en UDMA
- **Permet de recueillir les informations sur la commande précédente**

Quelques chiffres étonnants

- Kilo : 10^3
- Tera : 10^{12}
- Zetta : 10^{21}
- Mega : 10^6
- Peta : 10^{15}
- Yotta : 10^{24}
- Giga : 10^9
- Exa : 10^{18}
- (Pas de préfixe après Yotta)
- 10 To = Tous les imprimés de la Bibliothèque du Congrès Américain
- 8 Po = les données disponibles sur le Web
- 20 Po = L'ensemble des disques durs produits en 1995
- 2 Eo = Volume annuel des informations générées dans le monde
- 5 Eo = tous les mots prononcés depuis le début de l'humanité