TD2

Démarrage et arrêt du système LINUX

Le BIOS
(Basic Input Output System)

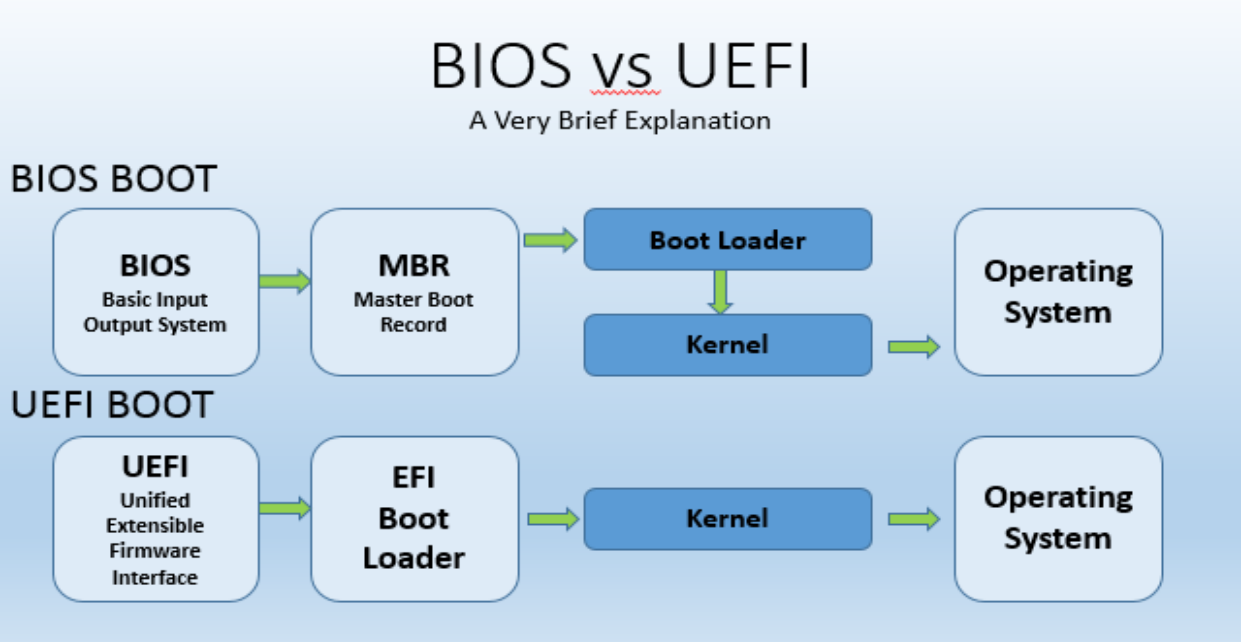
- L'objectif du BIOS est de rendre transparent, à tout système d'exploitation le matériel (hardware). En utilisant les mêmes fonctions du BIOS sur deux cartes mères différentes, on obtiendra le même résultat.
- Le BIOS comprend également le logiciel nécessaire à l'amorçage de l'ordinateur. La première phase de l'amorçage (boot) est l'auto-configuration à l'allumage (POST, Power-On Self-Test), qui compte la quantité de mémoire, teste les disques et configure les composants.
- La séquence d'amorçage continue avec la recherche d'un système d'exploitation, avant de le lancer.
- Le BIOS prend en charge à bas niveau les communications avec les périphériques. Parmi les prises en charge offertes par le BIOS, il y a le clavier et un mode d'affichage simplifié.

Les deux types de « BIOS »

- Le BIOS « historique » est remplacé depuis quelques années par l'UEFI (Unified Extensible Firmware Interface).
- L'UEFI apporte des améliorations par rapport au BIOS notamment une gestion améliorée des partitions de disque (utilisation du format GPT -Globally unique identifier Partition Table). Le BIOS ne gère que les partitions MBR.
- L'UEFI est capable de charger directement le noyau sans bootloader
- L'UEFI apporte aussi des « sécurités » comme secure boot (UEFI valide la signature du noyau avant de le démarrer –> pose des problèmes avec les OS alternatifs comme Linux).

BIOS VS UEFI

- Le démarrage est très différent entre le BIOS et l'UEFI
- Le BIOS historique va chercher un disque pour lancer un gestionnaire de démarrage (bootloader). C'est ce gestionnaire qui va lui se charger de démarrer l'OS. Le BIOS ne sait pas lire un disque.
- L'UEFI se comporte comme un « mini OS » et il va directement charger le noyau via un driver UEFI. l'UEFI est capable de lire différents format de disques.
- L'UEFI semble donc plus souple mais finalement le risque est la compatibilité des drivers et la complexité globale du système.



Le gestionnaire d'amorçage (bootloader) :

Le chargeur est un petit programme chargé de mettre en mémoire l'image du noyau linux puis de lui passer la main.

Le bootloader se trouve dans un endroit précis du périphérique d'amorçage le MBR (Master Boot Record).(rappelons que le BIOS ne sait pas lire les Disques)

Un bootloader peut charger, au choix, des OS différents (cas des systèmes multi-boot).

Gestionnaire de démarrage : GRUB

Il existe plusieurs programmes qui font office de bootloader. Dans le monde Linux deux sont principalement utilisés LILO et GRUB.

Nous allons regarder GRUB qui est le plus utilisé :

GRUB, ou Grand Unified BootLoader est un programme segmenté en 2 parties.

- La première partie de 444 octets est dans le MBR
- La deuxième partie (la plus grosse) se trouve dans /boot/grub/

Les options de démarrage sont dans le fichier /boot/grub/menu.lst (ou grub.conf ou grub.d/ suivant les distributions)

GRUB va alors charger le « noyau » de LINUX : le cœur du système d'exploitation.

GRUB va également charger les modules du noyaux linux via l'initrd (initial RAM Disk) qui est un fichier compressé contenant tous les modules utilisables par le noyau.

GRUB2 est compatible avec le système de démarrage UEFI.

Gestionnaire de démarrage : GRUB

Quand GRUB charge le noyau Linux, il lui donne aussi des paramètres, inscrits dans le menu.lst (sous la forme d'une ligne de commandes).

Ces paramètres incluent la partition à monter sur le dossier racine (sous la forme /dev/sdxx ou /dev/disk/by-uuid/xxxx).

Grub permet à l'utilisateur d'intercepter le démarrage pour changer les options de chargement du noyau et, par exemple, passer en mode « single-user » (mode de secours).

Exercices : Démarrage

Ne pas jamais éditer grub.cfg « à la main », ce fichier est mis à jour par update-grub

- Changer le temps d'attente au démarrage
  - Édition de /etc/default/grub
  - Changer le TIMEOUT
  - Recharger la configuration
- Redémarrer et entrer en mode single

Le système d'initialisation System V

Le système d’initialisation dit « System V » est le procédé « historique » du démarrage des Unix.

Une fois le chargement terminé le noyau lance le processus init (/sbin/init), Il possède le PID 1 en tant que « premier » processus.

Les tâches d’initialisation comprennent entre autres :

- montage de de /proc et /sys
- mise à l’heure
- chargement de LVM
- l’activation du swap
- ...

La notion de runlevel

System V utilise des niveaux de démarrage qui configurent globalement la machine. A un runlevel correspond un dossier /etc/rcx.d qui contient les liens pour le lancement de certains services.

Les runlevels ne sont pas tous normalisés, mais en général la signification suivante est adoptée :

0 : Arrêt

1 : Mode mono-utilisateur ou maintenance (mode single-user)

2 à 5 : dépend du système d'exploitation

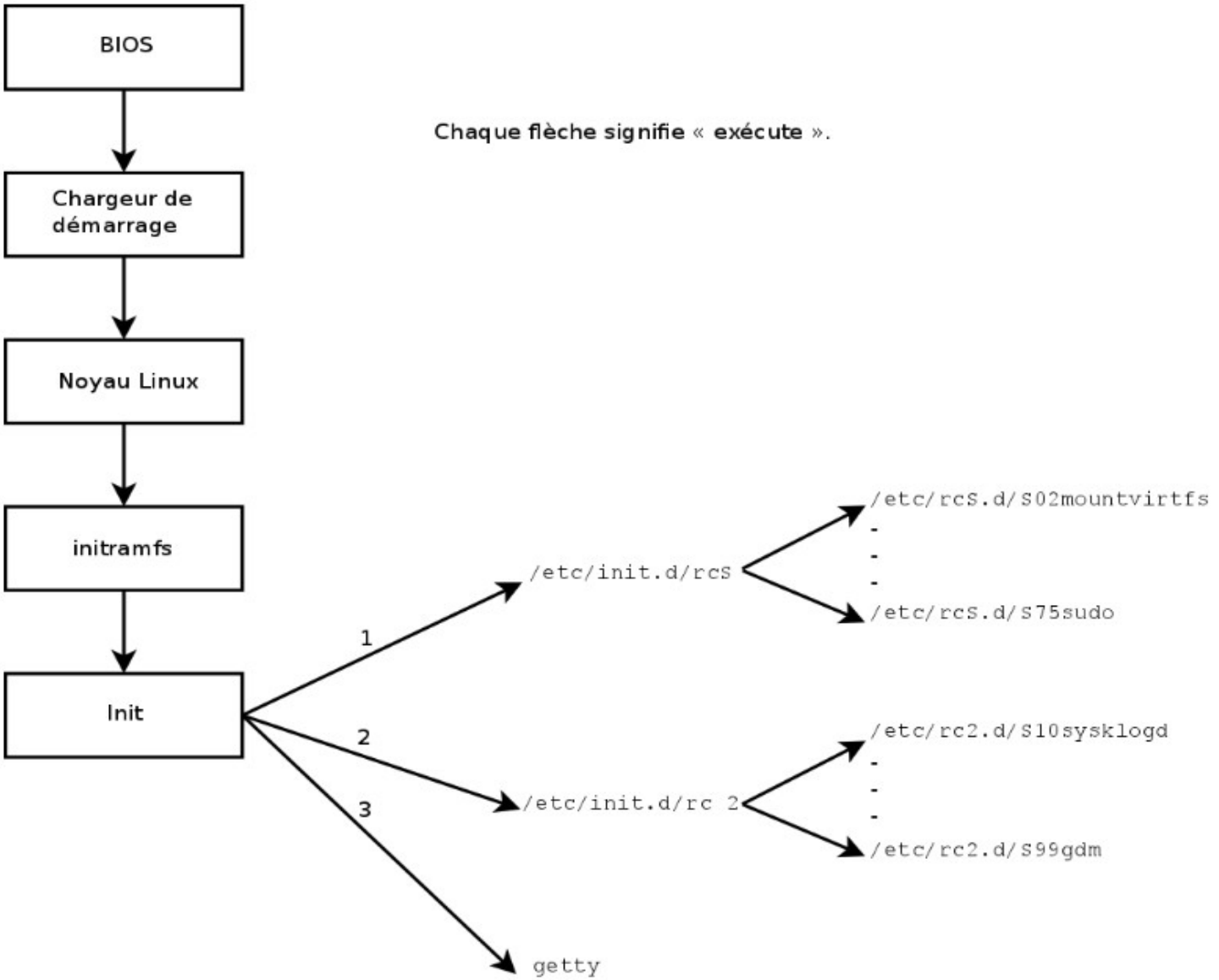
6 : Redémarrage

2 peut correspondre à un mode multi-utilisateur sans serveur applicatif.

3 correspond alors à un environnement multi-utilisateur avec serveurs applicatifs.

4 ou 5 est parfois utilisé pour lancer l'environnement graphique.

Schéma du démarrage



Initialisation de System V

- Le système exécute un ensemble de commandes listées dans le fichier /etc/inittab
- Les dossiers rcX.d contiennent des liens symboliques vers les différents scripts à exécuter et se trouvant dans le dossier /etc/init.d/.

  Par convention les noms commençants par Sxx (ou xx est un nombre entre 0 et 99) indique que le lien sera exécuté au démarrage (Start) et le nombre indique l'ordre de démarrage. Maintenant cette numérotation est remplacée par des informations placées dans l'entête des scripts et gérées par l'utilitaire: insserv.
- Les liens commençant par K (Kill) seront exécutés pour l'arrêt.

systemd

Systemd est l'alternative au démarrage « System V ».

Il offre une meilleure gestion du processus de démarrage en étant plus précis dans la gestion des dépendances entre services.

- le réseau doit être opérationnel avant de démarrer le pare-feu, lui même actif avant de démarrer le serveur WEB, ...

Meilleure journalisation (authentifiés, depuis le boot)

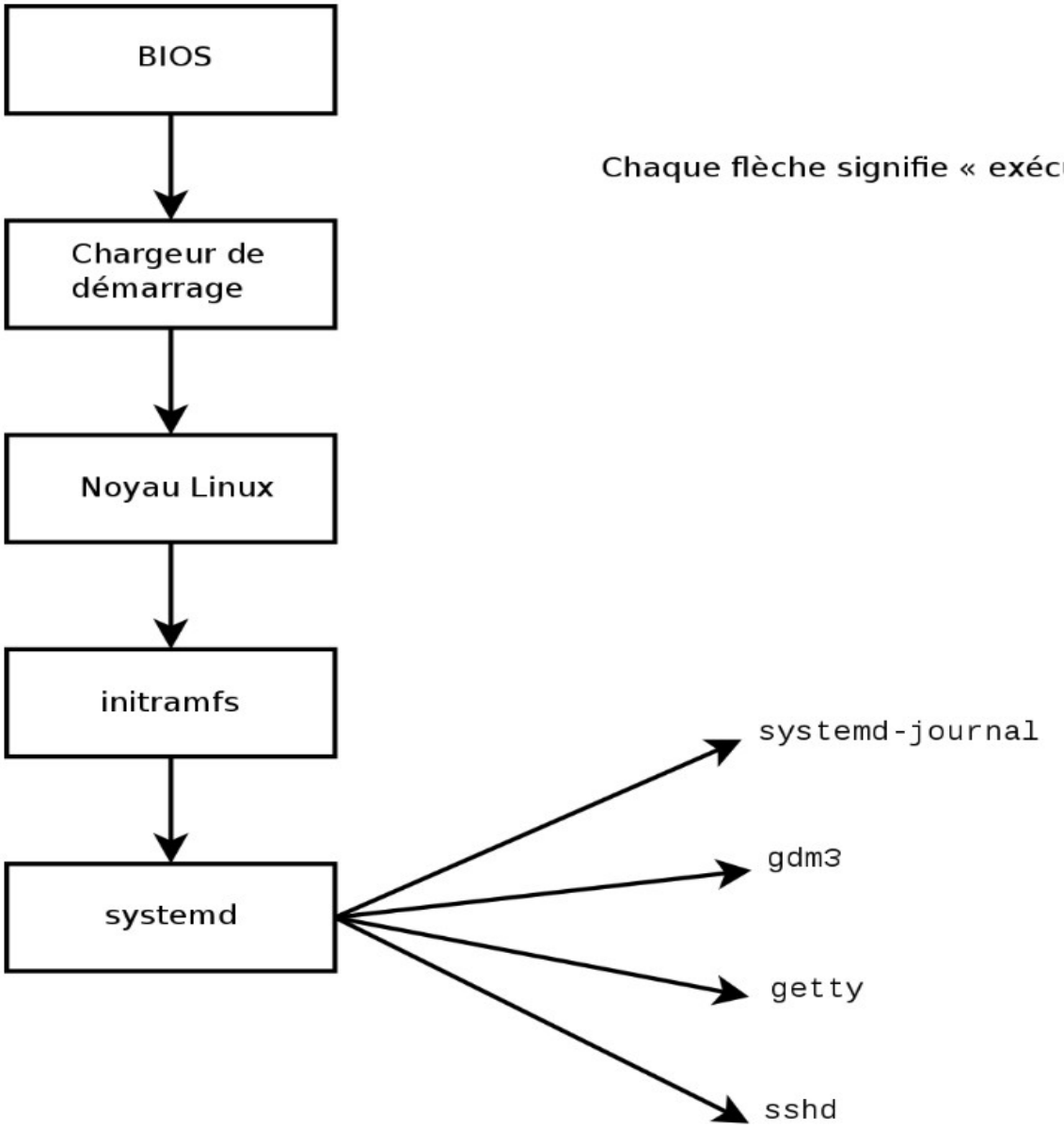
Systemd n'a plus besoin de scripts exécutables.

Il a été largement (et rapidement) adopté par les distributions car il simplifie les maintenances du système.

systemd

- Systemd est un ensemble de 80+ binaires dont :
  - /sbin/init (systemd) ne s'occupe que de l'init ;
  - journalctl ne s'occupe que de consulter le journal ;
  - udev s'occupe de découvrir les périphériques et peupler /dev ;
  - hostnamectl s'occupe de la gestion du nom d'hôte (hostname) de la machine ;
  - machinectl s'occupe de la gestion des conteneurs lancés par systemd ;
  - coredumpctl s'occupe des vidanges système (core dump) ;
  - systemd-analyze s'occupe d'inspecter la vitesse du démarrage ;
  - systemctl permet de gérer les services.
  - …
- Depuis 2015 systemd est le mode par défaut des OS des familles Debian et RED HAT.

Schéma du démarrage



Chaque flèche signifie « exécute ».

systemd

Systemd introduit la notion de target au sein de ses unités. Une target permet de regrouper dans un seul paquet plusieurs autres unités et de retrouver la notion de runlevel (pour la compatibilité des scripts System V).

Tableau de correspondance

Runlevel	Systemd Target	Notes
0	runlevel0.target, poweroff.target	Arrête le système
1	runlevel1.target, rescue.target	Mode single user
3	runlevel3.target, multi-user.target	Mode multi-utilisateur, non graphique
2, 4	runlevel2.target, runlevel4.target, multi-user.target	Mode défini par l'utilisateur, identique au 3 par défaut.
5	runlevel5.target, graphical.target	Mode graphique multi-utilisateur
6	runlevel6.target, reboot.target	Redémarre
emergency	emergency.target	Shell d'urgence

En System V, pour arrêter un ordinateur, on fait shutdown -h (ou -r).

- Le programme shutdown demandait à init de passer en runlevel 0 ou 6.

Avec systemd, l'utilitaire shutdown devient un lien symbolique sur /sbin/systemctl

Autres fonctions de systemd

Cartographie des services lancés au démarrage avec leur temps de lancement

systemd-analyze blame

Liste des services

systemctl list-unit-files

Etat des logs

journalctl -f (pour avoir les log du système en temps réel)

Désactiver un service

systemctl disable <Nom_du_service>.service

Activer une service

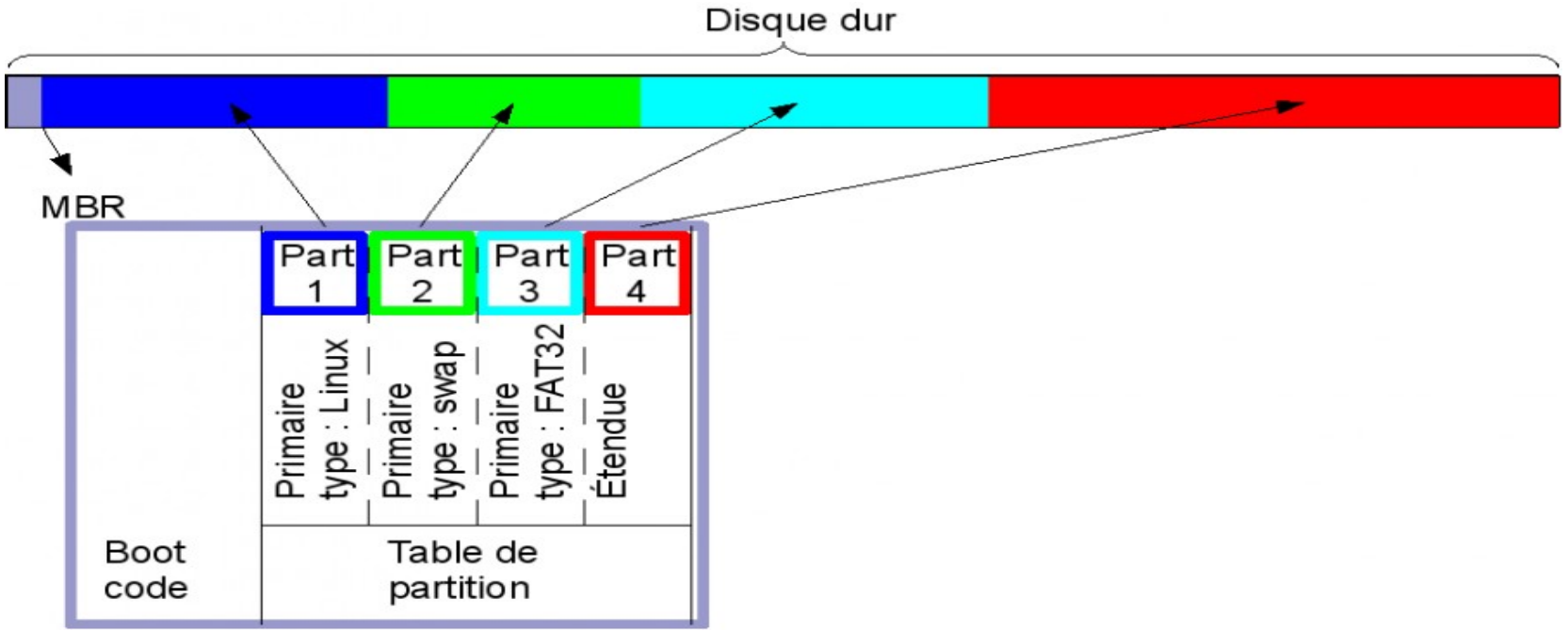
systemctl enable <Nom_du_service>.service

Exercices : systemd

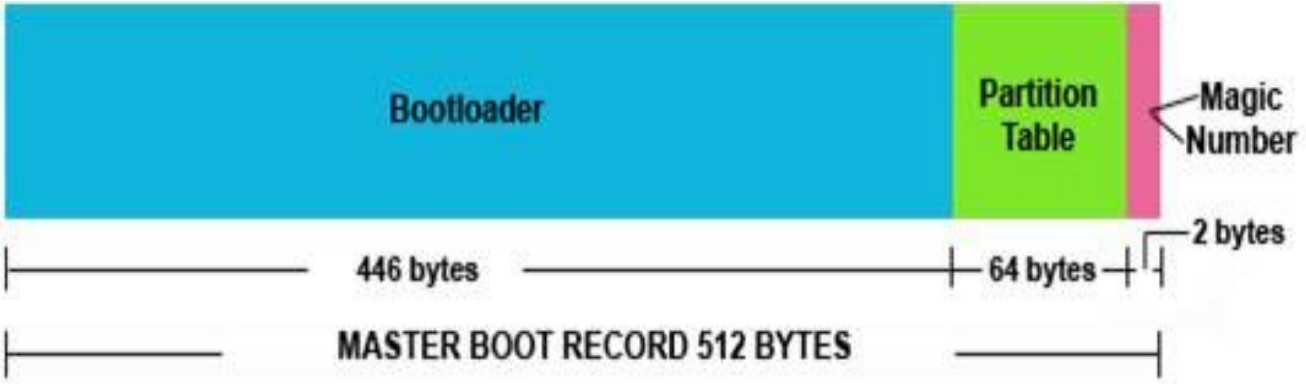
- Démarrez le service ssh
- Donnez l'état du service ssh
- Arrêtez ce service
- Proposez trois façons de rebooter la machine depuis le shell

La table de partition de type MBR

- Une Table de partition MBR n'autorise que 4 partitions.
  - Partitions primaires
  - Partition étendue
    - Partition utilisée par les partitions secondaires
  - Partitions secondaires (logiques)
    - Ces partitions sont contenues à l'intérieur d'une partition étendue
- Au maximum un disque peut gérer 2,2 To



Le MBR (Master Boot Record)



Le premier bloc de données de 512 Octets répartis en :

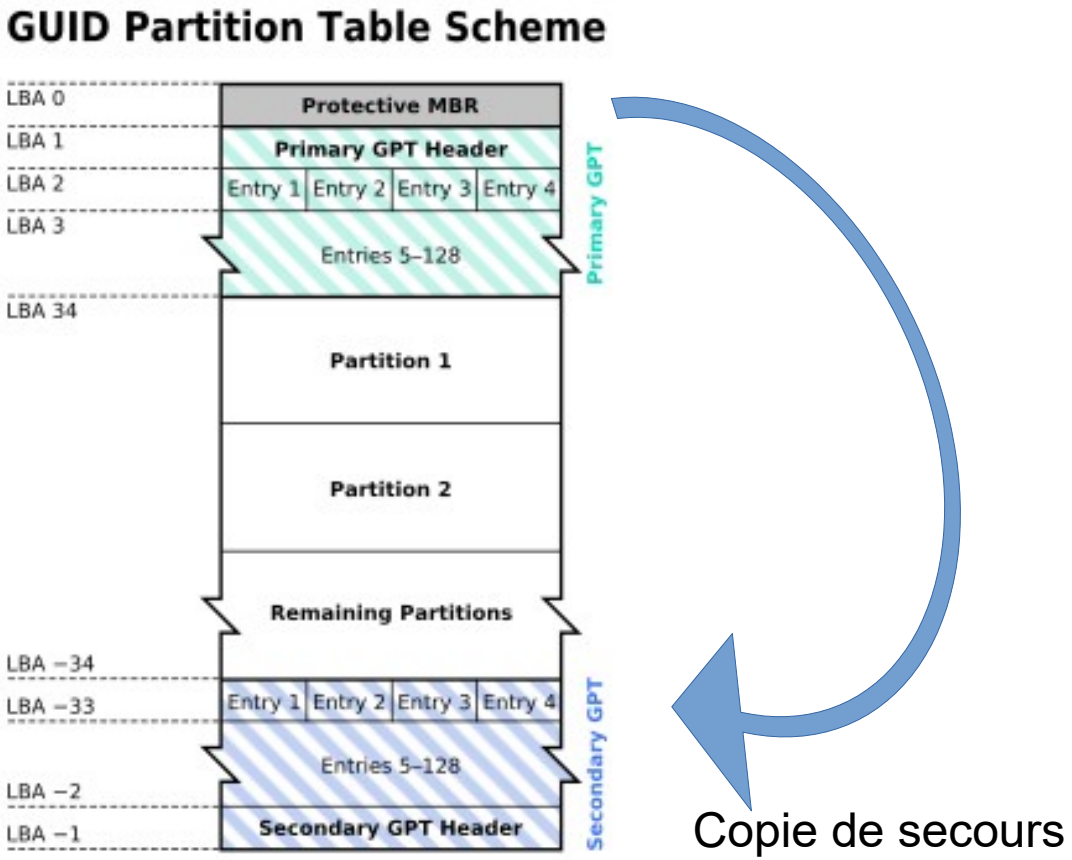
446 octets pour stocker la première partie du binaire du programme bootloader.

64 octets pour la table de partition (16 octets par table → 4 tables)

Magic number AA55 de 2 octets (pour vérifier l'intégrité de la table)

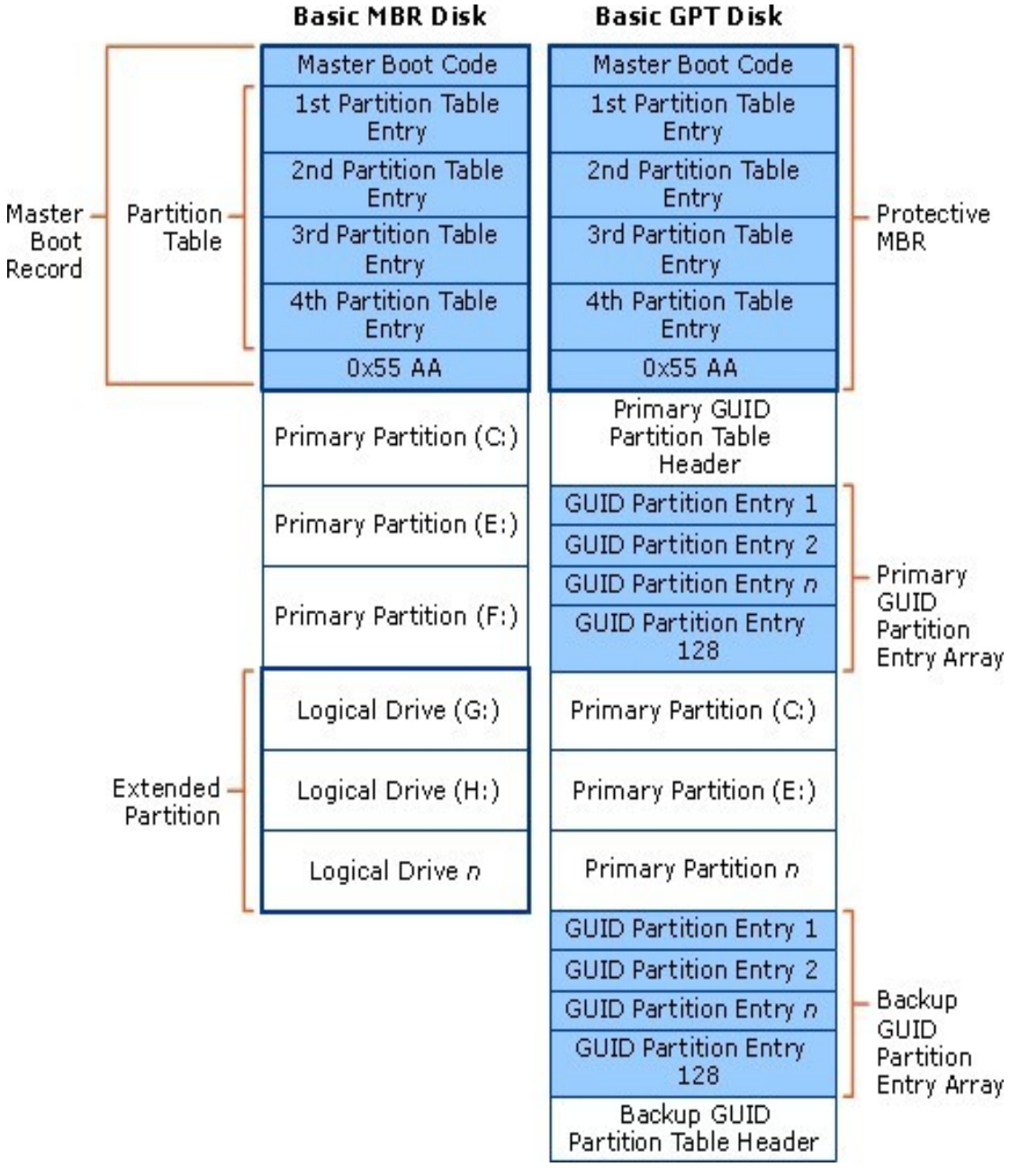
La table de partition de type GPT

- Jusqu'à 128 partitions (primaires)
- 256 To par partition
- Un format beaucoup plus complexe



format MBR vs GPT

- Il existe bien 2 systèmes différents mais GPT a calqué le format MBR sur la partie du boot. Cela pour simplifier les problèmes de compatibilité.
- Que l'on soit en MBR ou en GPT, les 512 premiers octets ont le même formatage.



Exercices : MBR

Faire une copie du MBR, l'afficher en hexadécimal et repérer le « magic number »:

1) Repérer si le disque est utilisé en MBR ou GPT

2) Faire une copie bit à bit des 512 premiers octets

3) Utiliser od (un éditeur hexadécimal) pour repérer le « magic number »

Correction : Démarrage

- Changer le temps d'attente au démarrage
  - Édition
    - sudo vi /etc/default/grub
  - Changer
    - GRUB_TIMEOUT=5 en GRUB_TIMEOUT=30
  - Recharger la config :

    sudo update-grub
- Lors du reboot
- Entrer en mode single
  - Menu : Options avancées
  - Puis Menu ...(recovery mode)
    - Lance la commande « runlevel »
    - Pour sortir du mode maintenance et continuer le boot

Correction : systemd

- Suivi du service ssh
  - systemctl status sshd.service
- Démarrage du service ssh
  - systemctl start sshd.service
- État du service
  - systemctl status sshd.service
- Arrêt du service
  - systemctl stop ssh.service
- Proposez trois façons de redémarrer la machine depuis le shell
  - sudo shutdown -r
  - sudo systemctl reboot

  on peut aussi demander à systemd de changer de « target » (runlevel) et passer en 6 pour redémarrer
  - systemctl isolate runlevel6.target
- Recharger une nouvelle configuration
  - systemctl reload nom_du_service.service
- Lister les services démarrés
  - systemctl list-units --type=service

Correction : MBR

Faire une copie du MBR, l'afficher en hexadécimal et repérer le « magic number »:

1) Vérifier le format du disque
   sudo fdisk -l /dev/sda

2) Faire une copie bit à bit dans le fichier boot.mbr
   sudo dd if=/dev/sda of=boot.mbr bs=512 count=1

3) Utiliser od, un éditeur hexadécimal
   od -w1 -h -Ad boot.mbr

L'exercice a pour but de rendre « visible » la théorie.

La copie du MBR était aussi une solution utilisable pour corriger l'action de certains virus qui modifiaient le boot (surtout sous Windows mais le principe reste le même).