IVC

Infrastructures Virtuelles et Conteneurs

Legond-Aubry Fabrice

fabrice.legond-aubry@parisnanterre.fr

Infrastructures Virtuelles et Conteneurs

Principes Pratiques

Plan du Cours

virtualbox

Podman

buildah

vagrant

Virtualisation: l'exemple virtualbox
Présentation

- Logiciel s'exécutant sur MacOSx, Linux, Windows
- Avant de créer une machine, il faut toujours configurer les répertoires pour savoir où seront stockées les données (disques) et les métadonnées (configuration, logs)
  - ✓ File > Préférences (CTRL-G) > Général (Default Machine Folder)
  - ✓ Variables d'environnement sous linux (variable env. VBOX_USER_HOME)
- Valeur par défaut :
  - ✓ Linux and Oracle Solaris: $HOME/.config/VirtualBox
  - ✓ Windows: $HOME/.VirtualBox
  - ✓ Mac OS X: $HOME/Library/VirtualBox
- Données :
  - ✓ Default machines : $HOME/VirtualBox VMs
  - ✓ Default disk image location : In each machine's folder
  - ✓ Machine settings file extension : .vbox
  - ✓ Media registry ; Each machine settings file Media registration is done automatically when a storage medium is attached to a VM

Virtualisation: l'exemple virtualbox
Présentation

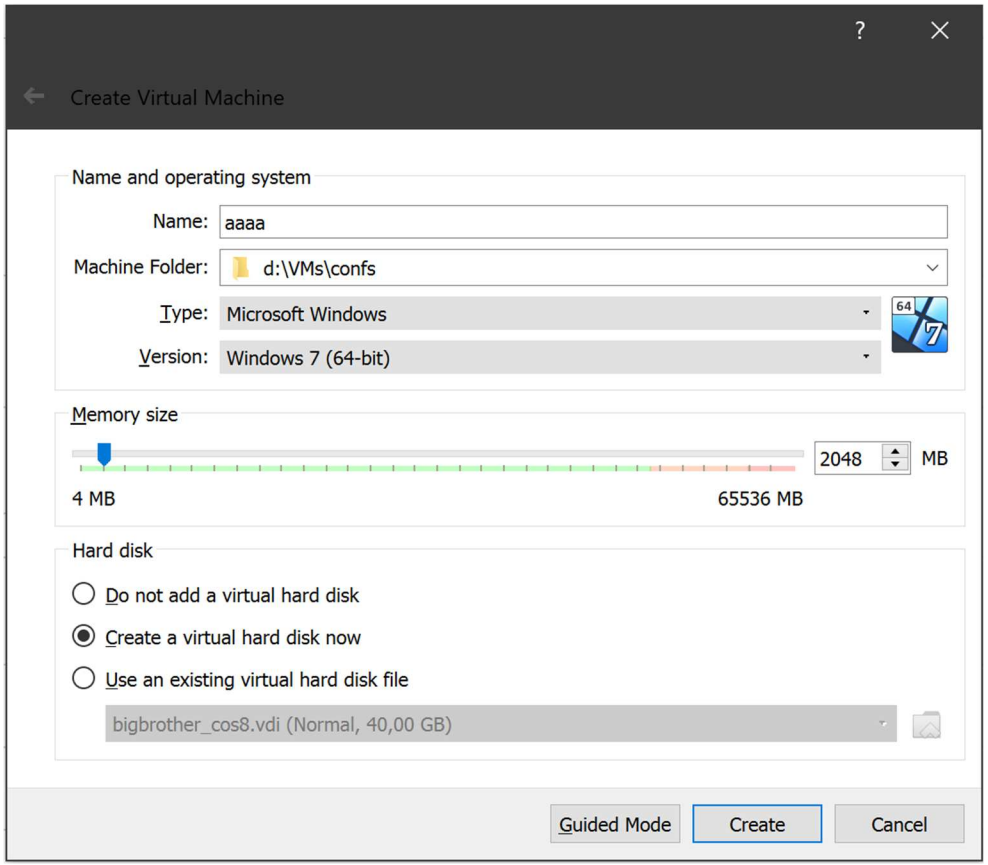
- Les configurations sont stockés dans des fichiers xml lisibles et modifiables
- La machine virtuelle va afficher un écran "virtuel" (invité) sur votre écran physique (hôte)
- L'application capture le clavier

  - ✓ Pour sortir de la prison "VM", il existe une touche qui apparait en bas à droite de la fenêtre de vos vm
  - ✓ Elle peut être changée via File > Preferences > Input > Virtual Machine
  - ✓ Touche "Host Key Combination"
  - ✓ Il y a aussi la liste des autres raccourcis de la VMs

Virtualisation: l'exemple virtualbox
Création

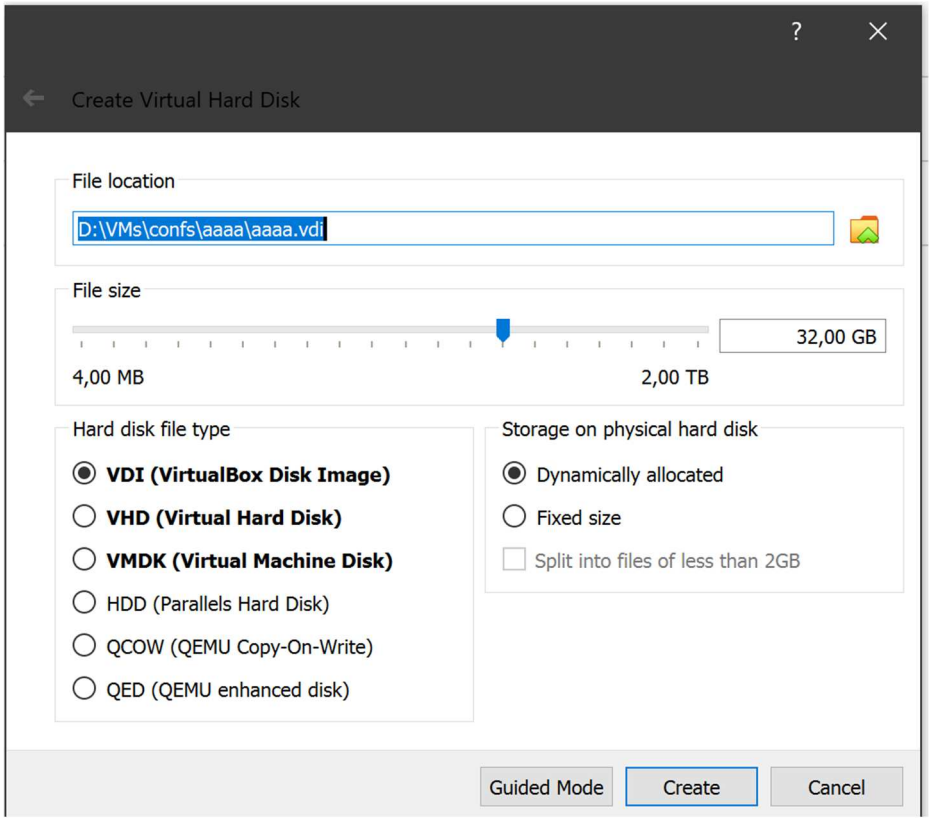
- Manipulation des machines via la GUI
  - ✓ Ou via un programme en ligne de commande: VBoxManage
- Création d'une machine
  - ✓ Menu : Machine > New
    - ➢ Fixer le Nom
    - ➢ Fixer le répertoire de stockage des données
    - ➢ Fixer le type d'OS + Version (spécifie le CPU 32/64bits). Active ou non des options
    - ➢ Fixer le montant de la RAM maximum
    - ➢ Choisir un disque (nouveau ou existant)
  - ✓ On remarquera que vous devez créer l'équivalent d'une machine physique

Virtualisation: l'exemple virtualbox
Création (paramètres de base)



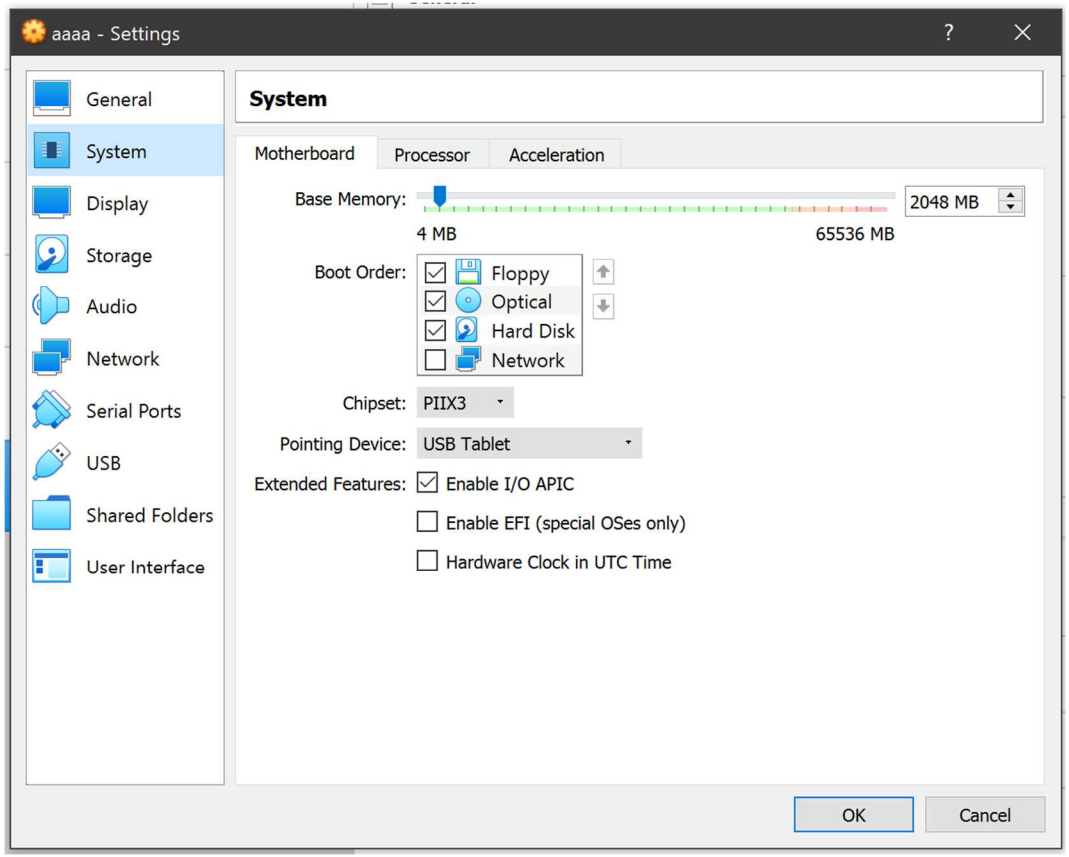
Virtualisation: l'exemple virtualbox
Création (disque)

- Gestion du disque :
  - ✓ Un disque invité est représenté par un fichier sur le système hôte
  - ✓ Il existe plusieurs formats
    - ➢ Qcow/qed → Qemu, VHD → virtual pc, VDI → virtualbox, vmdk → vmware
  - ✓ Il existe des fichiers de taille variable (plus lent mais plus adaptable), ou fixe (préallocation, plus rapide)



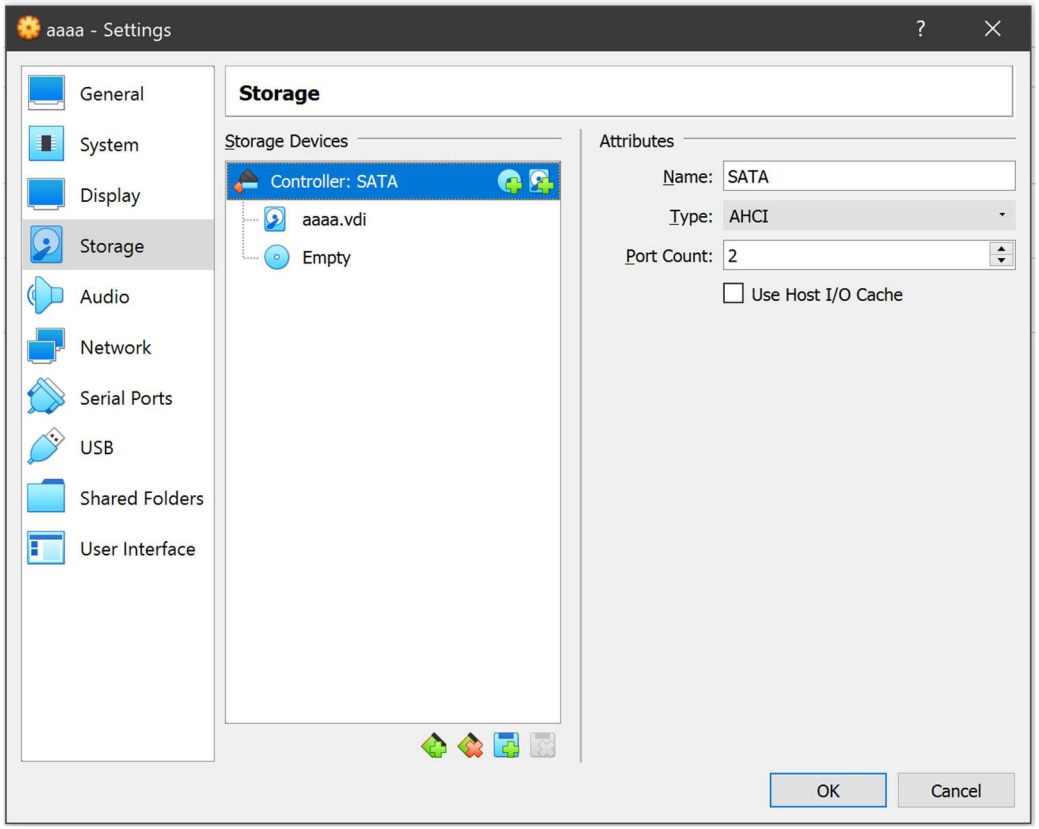
Virtualisation: l'exemple virtualbox
Paramètres

- On peut fixer plus précisément les paramètres de chaque machine
  - ✓ Machine > Setting (CTRL-S)
- Page "System"
  - ✓ Enable I/O APIC, Hardware clock
    - ➢ Délégation des appels a certains drivers
  - ✓ Onglet "Processor"
    - ➢ On fixe le nombre de core
    - ➢ On retrouve les options de paravirtualisation (accélartion avec VTx/AMDv et pour le MMU (PAE/NX)
  - ✓ Onglet "Accélération" (paravirtualisation)
    - ➢ Fixer le module utiliser pour accélérer (HyperV, KVM)



Virtualisation: l'exemple virtualbox
Paramètres

- Page "Display"
  - ✓ Fixer la carte graphique émulée
  - ✓ On remarque que les cartes proposés sont des cartes non réélles pour virtualbox
- Page "Storage"
  - ✓ On fixe l'ensemble des supports de stockage
  - ✓ Lecteurs CD + HDD/SDD
  - ✓ Le type de contrôleur émulé
  - ✓ Possibilité d'activer le cache hôte

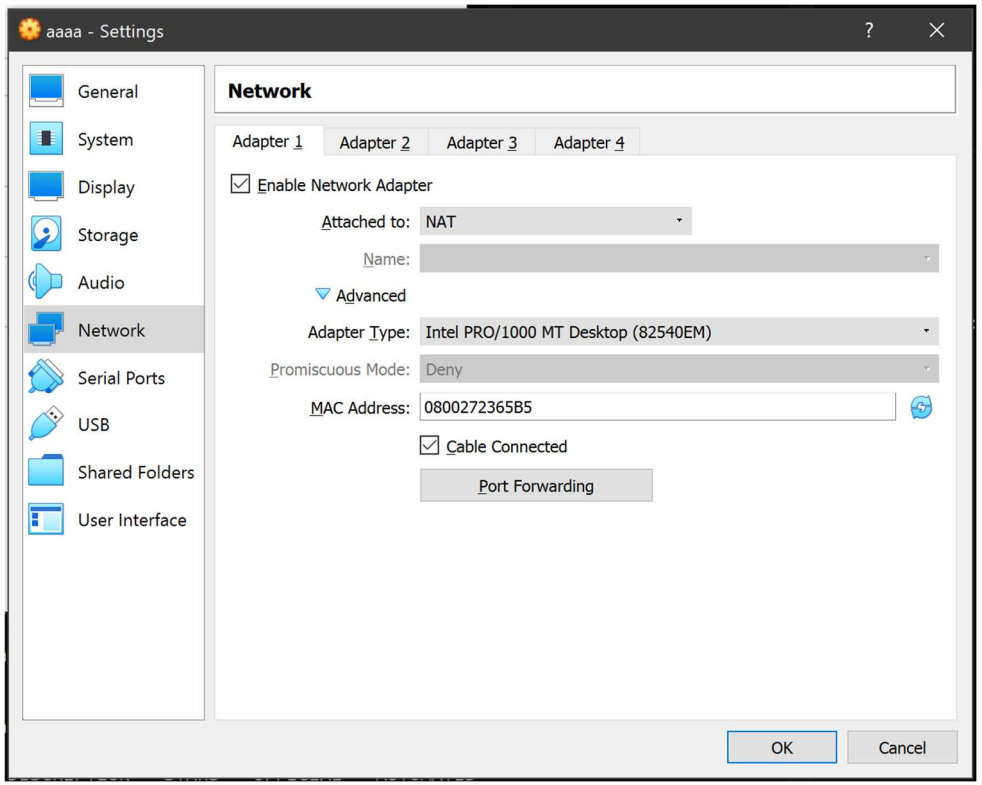


Virtualisation: l'exemple virtualbox
Paramètres

- Page "Audio"
  - ✓ Fixer la carte son et les ports sériels
- Page "Serials Ports"
  - ✓ Fixer la carte son invité (micro+ HP)
  - ✓ Fixer le liens carte son invité vers carte son hôte
- "Page "USB"
  - ✓ Fixer la délégation des périphériques USB
  - ✓ Note: virtualbox ne sait pas booter sur USB (version 6.1.18)

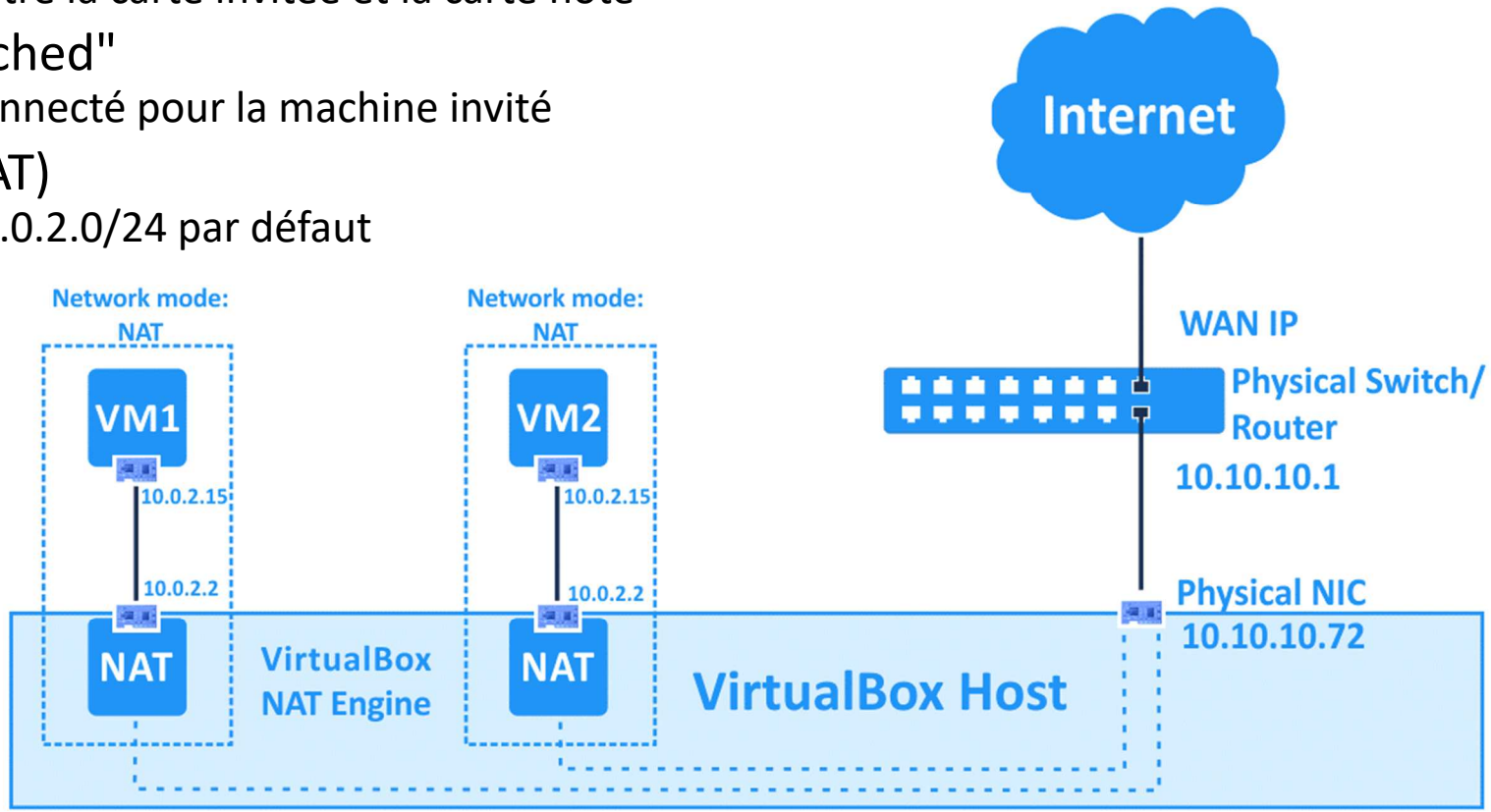
Virtualisation: l'exemple virtualbox
Paramètres (réseaux)

- Page "Network"
  - ✓ Jusqu'à 4 cartes réseaux possibles
  - ✓ Fixer le type de carte réseaux
  - ✓ Emulation de carte réelles (Intel, Pcnet)
  - ✓ Emulation d'une carte virtuelle (paravirtualisation) → seulement niveau réseau L3+ (IP et sup) nommée virtio-net
  - ✓ Fixer l'@MAC (voir cours réseau)
  - ✓ Promiscuous : capture de tout le traffic réseau
- Virtio-net: https://projectacrn.github.io/latest/developer-guides/hld/virtio-net.html



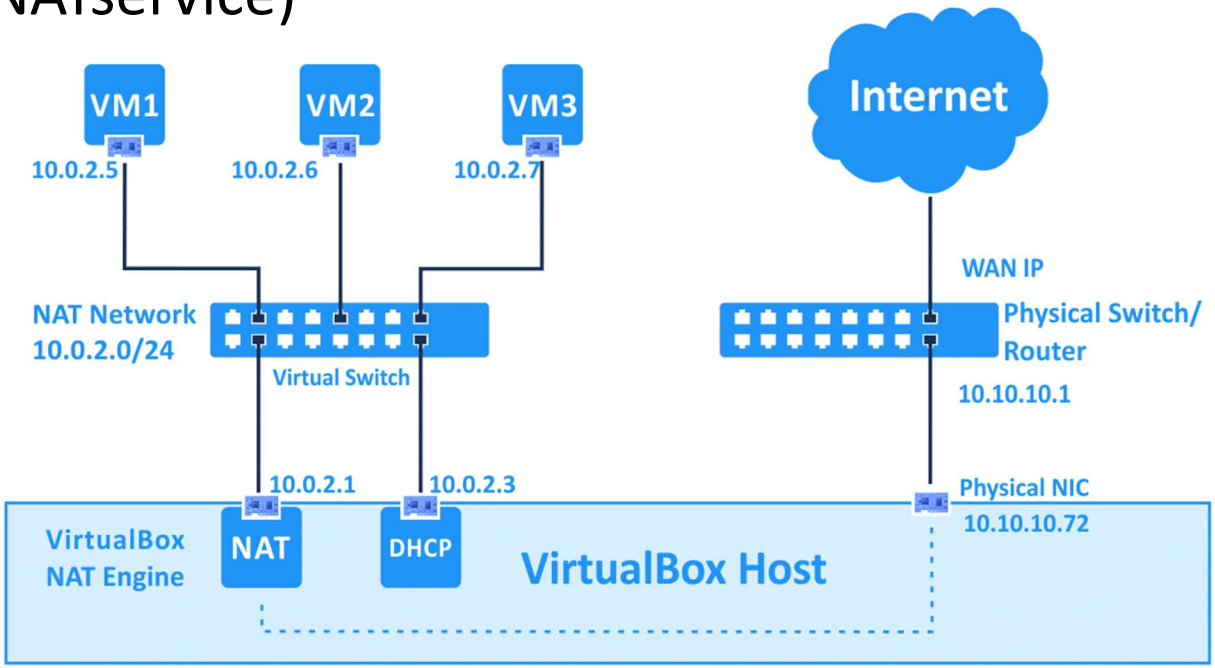
Virtualisation: l'exemple virtualbox
Paramètres (réseaux)

- Page "Network"
  - ✓ Fixer Type de liens entre la carte invitée et la carte hôte
- Type de lien "Not attached"
  - ✓ Simule un câble déconnecté pour la machine invité
- Type de lien "NAT" (NAT)
  - ✓ Réseau par défaut 10.0.2.0/24 par défaut
  - ✓ Sortie possible
  - ✓ Entrée impossible
  - ✓ **Non visibilité entre VMs**



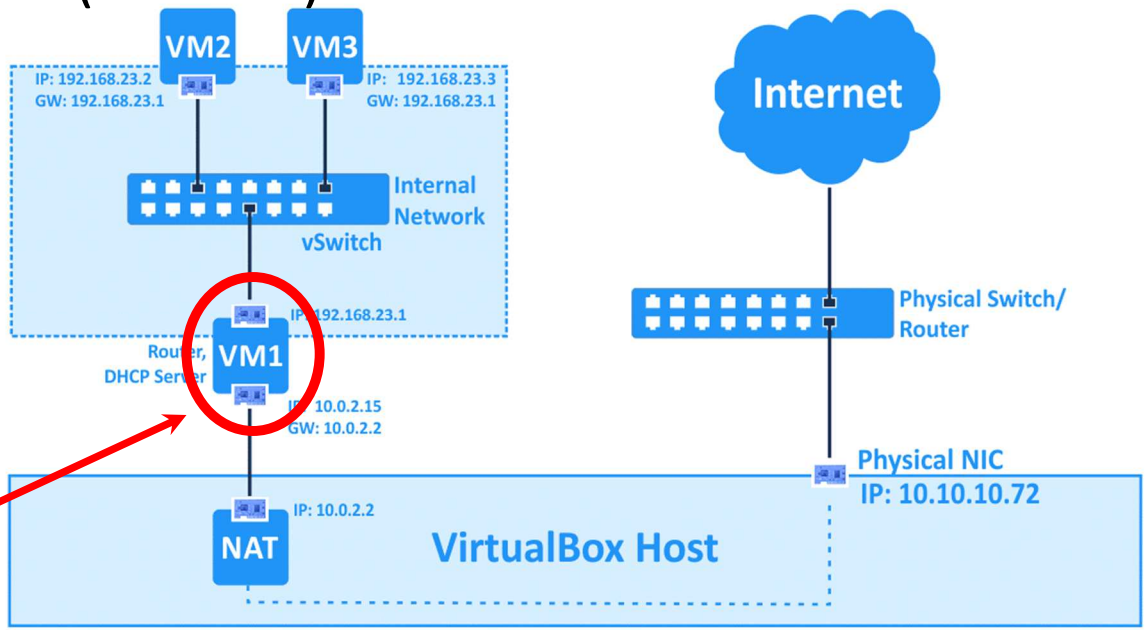
Virtualisation: l'exemple virtualbox
Paramètres (réseaux)

- Page "Network"
  - ✓ Fixer Type de liens entre la carte invitée et la carte hôte
- Type de lien "NAT Network" (NATservice)
  - ✓ Réseau à créer uniquement via "VBoxManage natnetwork"
  - ✓ VMs peuvent sortir
  - ✓ VMs peuvent se voir
  - ✓ Entrée impossible



Virtualisation: l'exemple virtualbox
Paramètres (réseaux)

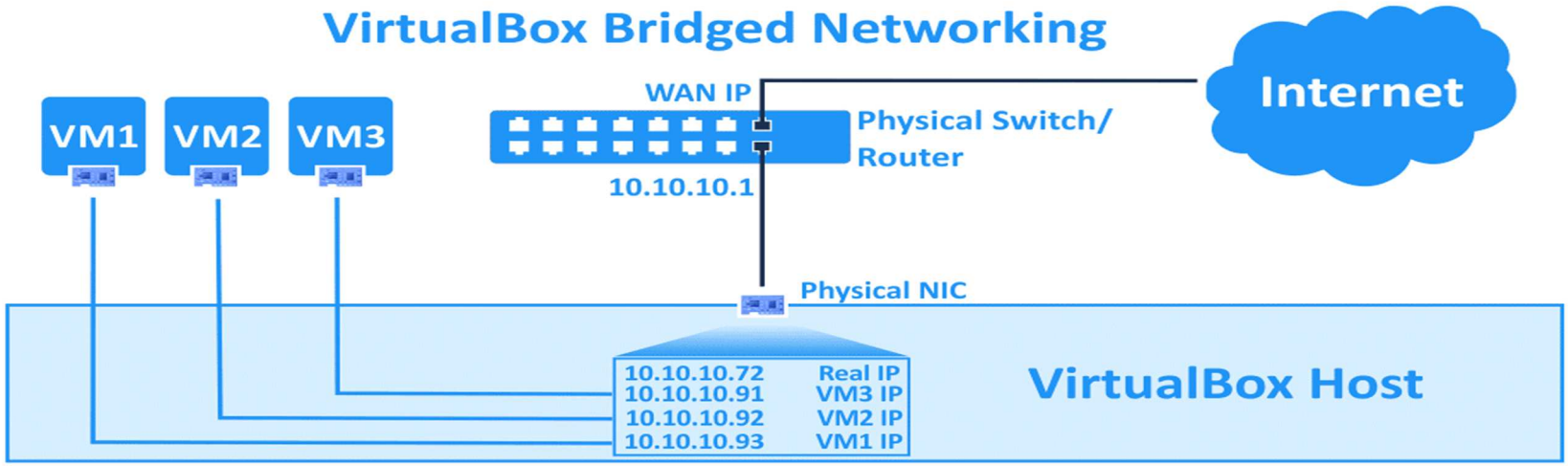
- Page "Network"
  - ✓ Fixer Type de liens entre la carte invitée et la carte hôte
- Type de lien "Internal Network" (Internal)
  - ✓ VBoxManage modifyvm "VM name" --nic<x> intnet …"
  - ✓ Impossible de sortir
  - ✓ Impossible d'entrée
  - ✓ Les VMs se voient entre elles



Pour sortir, il faut une VM à double carte réseau pour simuler un routeur

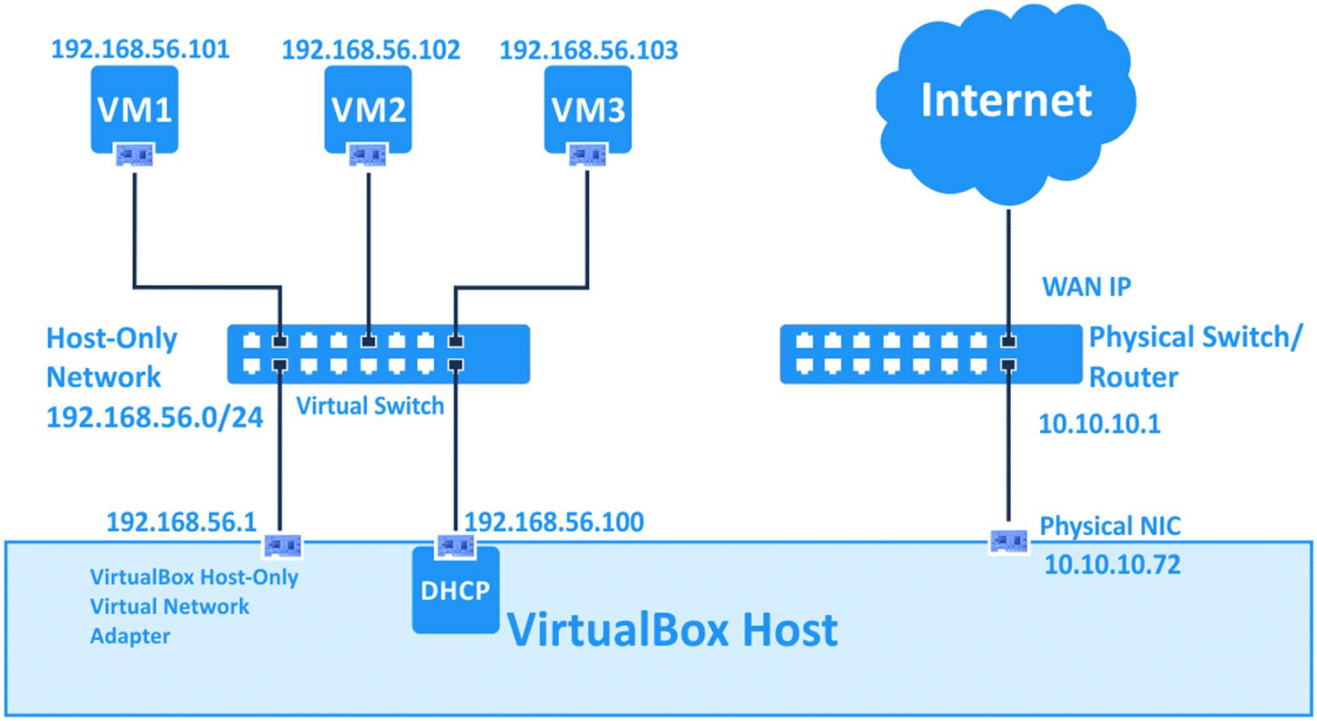
Virtualisation: l'exemple virtualbox
Paramètres (réseaux)

- Page "Network"
  - ✓ Fixer Type de liens entre la carte invitée et la carte hôte
- Type de lien "Bridged Adapter" (bridged)
  - ✓ Comme si 2 machines (VM+hôte) étaient sur le réseau physique
  - ✓ Offre le plus de possibilités dans la VMs



Virtualisation: l'exemple virtualbox
Paramètres (réseaux)

- Onglet Réseaux
  - ✓ Fixer Type de liens entre la carte invitée et la carte hôte
- Type de lien "Host-Only Network"
  - ✓ Les VMs se voient
  - ✓ Hôtes ←→ VMs OK
  - ✓ Impossible de sortir
  - ✓ Impossible d'entrée



Virtualisation: l'exemple virtualbox
Paramètres (réseaux)

- Table de capacité de chaque mode réseaux standard de virtualbox
  - ✓ Note: commun à beaucoup de système de virtualisation !!!
  - ✓ Important : attention en cas de VM serveur la dernière colonne est importante !!!
  - ✓ On peut fixer le port forwarding (PAT=Port Address Translation) sur certains modes
- Il existe un dernier mode réseau nommé "Generic Driver"
  - ✓ Il est fait pour se connecter à un mode réseau VDE switch dont on peut tout contrôler. Les capacités dépendent de la configuration du réseau.
  - ✓ Nécessite une configuration poussée

Mode	VM→Host	VM←Host	VM1↔VM2	VM→Net/LAN	VM←Net/LAN
Host-only	+	+	+	–	–
Internal	–	–	+	–	–
Bridged	+	+	+	+	+
NAT	+	Port forward	–	+	Port forward
NATservice	+	Port forward	+	+	Port forward

Virtualisation: l'exemple virtualbox
Gestion des médias

- Gestion des médias
  - ✓ Un média peut être mutable ou non
  - ✓ Se modifie par "VBoxManage modifymedium … --type"
  - ✓ Type: "normal"(mutable), "readonly", "immutable"
  - ✓ Un type non mutable fonctionne donc par COW et peut être remis à 0 à chaque redémarrage (paramètre " autoreset" "on" ou "off").
- On peut refusionner toute les modifications pour re-créer un disque homogène
  - ✓ Commande "VBoxManage clonemedium …"

Virtualisation: l'exemple virtualbox
Gestion des médias

- Immutabilité créer des hiérarchies de disques
  - ✓ Les disques COW apparaissent comme des sous médium dans la GUI
  - ✓ Le parent des disques COW apparaissent grâce à "VBoxManage showmediuminfo" ou "VBoxManage list hdds"
- Les images COW (differentielles)
  - ✓ Empêche le détachement ou l'effacement de l'image originale

Virtualisation: l'exemple virtualbox
Gestion des médias

- Exemple :

```
UUID:           5bbc033e-dcc7-456f-ad2b-82dbbe591d8f
Parent UUID:    base
State:          created
Type:           immutable
Location:
D:\VMs\disks\VMSecu_UbuntuSVC16LTS_v2021_updt_wo_corbof_pmaster.vdi
Storage format: VDI
Format variant: dynamic default
Capacity:       18432 Mbytes
Size on disk:   15013 Mbytes
Encryption:     disabled
Property:       AllocationBlockSize=1048576
Child UUIDs:    e4f7cea0-0e2f-4e1e-b0e3-fc7b9e06673b
```

Virtualisation: l'exemple virtualbox
Gestion des médias

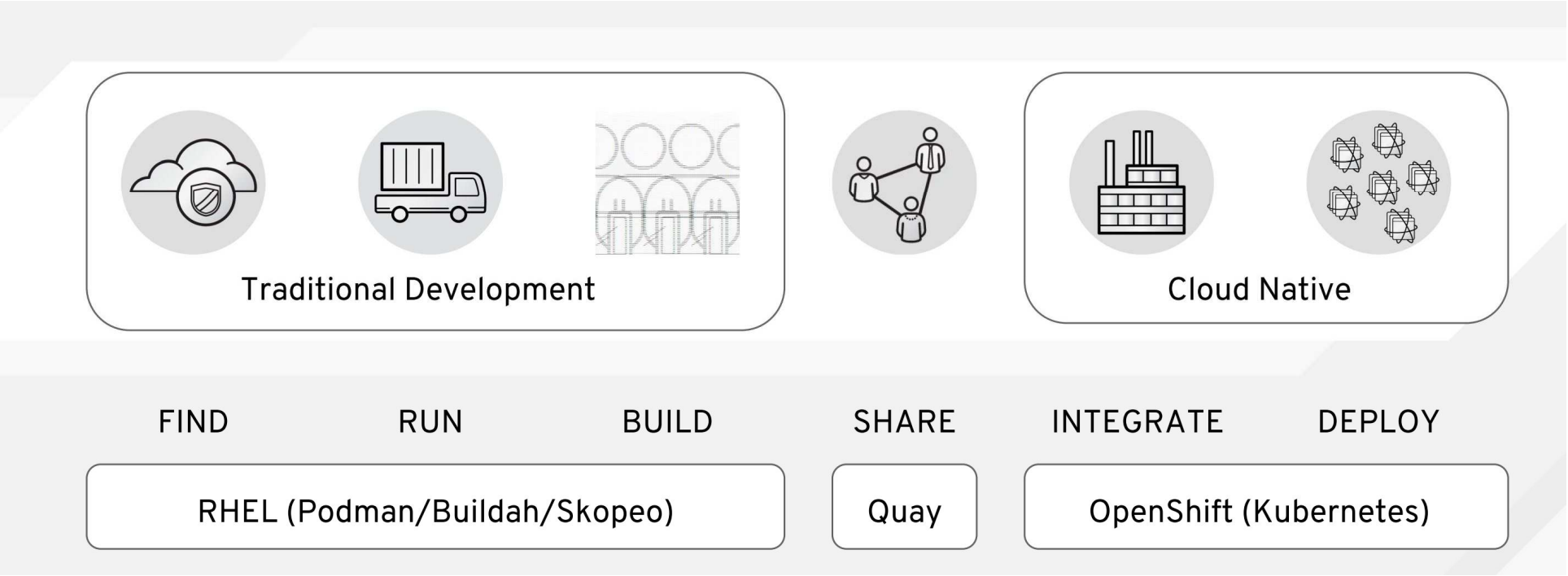
- Exemple :

```
UUID:           e4f7cea0-0e2f-4e1e-b0e3-fc7b9e06673b
Parent UUID:    5bbc033e-dcc7-456f-ad2b-82dbbe591d8f
State:          created
Type:           normal (differencing)
Auto-Reset:     on
Location:       D:\VMs\confs\ssi\Snapshots\{e4f7cea0-0e2f-4e1e-b0e3-
fc7b9e06673b}.vdi
Storage format: VDI
Format variant: differencing default
Capacity:       18432 Mbytes
Size on disk:   1744 Mbytes
Encryption:     disabled
Property:       AllocationBlockSize=1048576
In use by VMs:  ssi (UUID: 0f781ec9-3e82-4a03-a408-2acbe4932f2c)
```

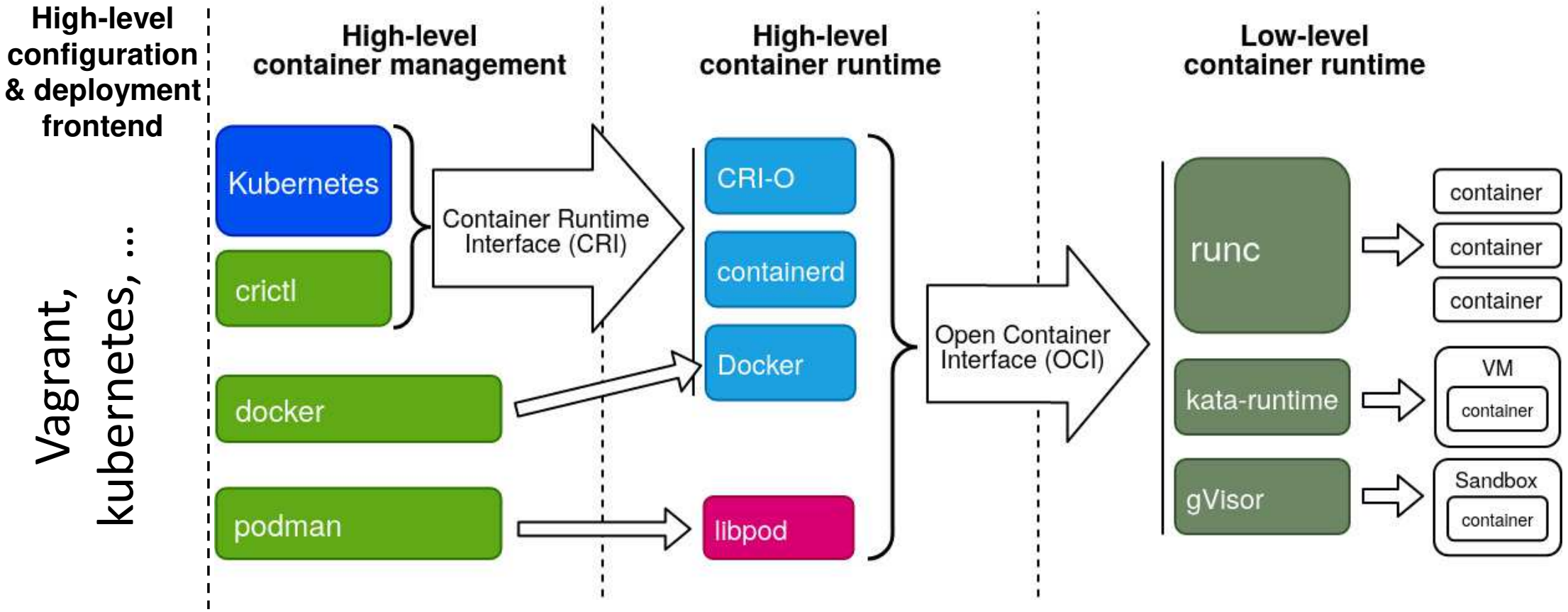
Conteneur: création et déploiement
Présentation

- Créer un conteneur peut se faire à la main
  - ✓ Lourd, long.
  - ✓ Voir TD (création partielle)
- Il existe des outils pour automatiser les déploiements et la création
  - ✓ Déploiement: podman, docker, vagrant
  - ✓ Création : dockers (une partie des outils), openshift, buildah

Conteneur: création et déploiement
Petite vision synthétique



Conteneur: création et déploiement
Petite vision synthétique



Conteneur: podman/buildah

- Couple buildah / podman
  - ✓ Podman → déploiment / gestion des conteneurs
    - ➢ https://podman.io/releases/
  - ✓ Buildah → construction des images OCI
    - ➢ https://buildah.io/
  - ✓ Liens github: https://github.com/containers
- Solution pour Conteneur "rootless" et "daemonless"
- Utilise la librairie "runc" pour la mise en place de l'exécution
- **Note: podman peut aussi gérer des conteneur "rootful"**

Conteneur: podman/buildah

Présentation

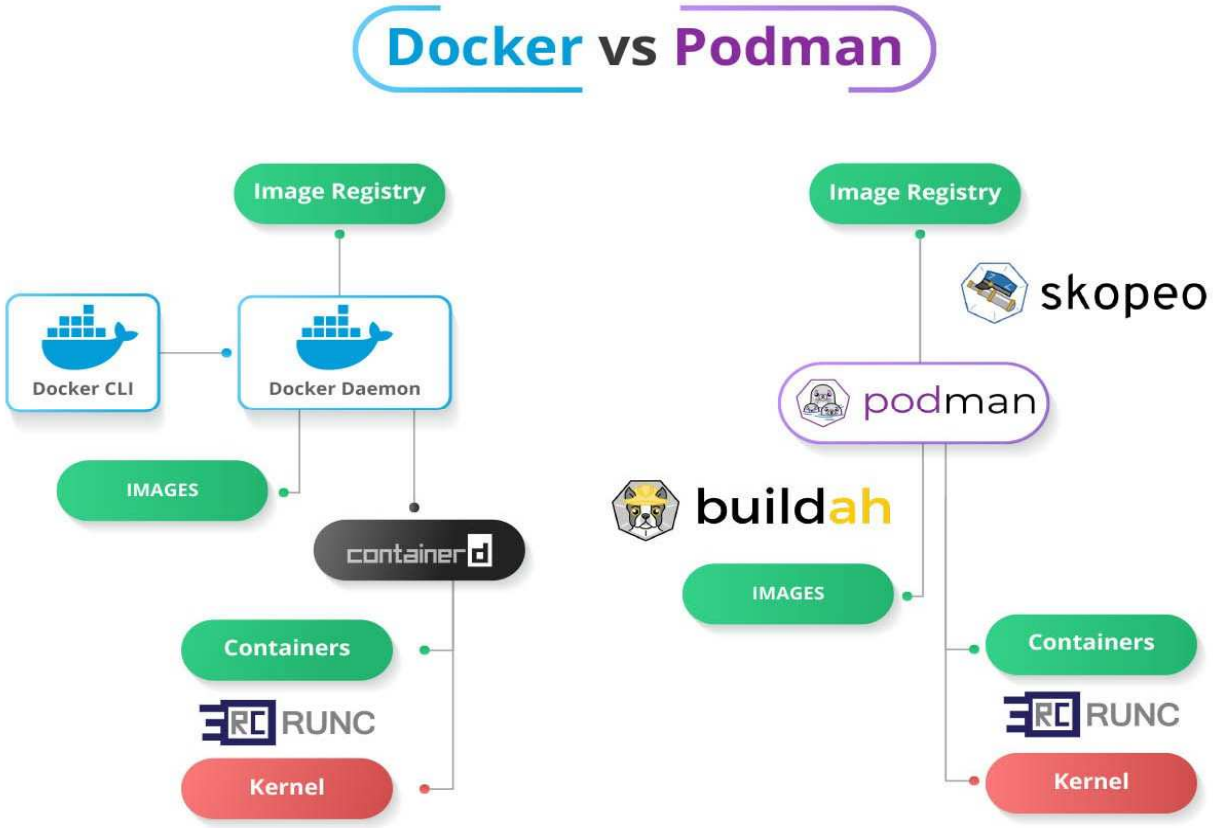
- Rappel: Solution pour Conteneur "rootless" et "daemonless"
  - ✓ Mitigation des vulnérabilités potentielles lors de l'exécution du conteneur et du contrôleur du conteneur
  - ✓ Permet aux utilisateurs de partager des machines puissantes en limitant les risques d'escalade de privilège
  - ✓ Meilleure isolation des conteneurs imbriqués (Container-In-Container)
  - ✓ Ex: Docker CVE-2014-9357, CVE-2014-3519, CVE-2017-1002101, CVE-2019-5736, CVE-2020-13295
  - ✓ Ne résiste pas aux failles de noyaux / drivers matériels
  - ✓ Voir la partie sécurité
- Note: "Docker v20.10+" supporte le "rootless" mais reste "daemonful"
- Note: "LXC" supporte le rootless depuis 2013 mais reste "daemonful"

Conteneur: podman/buildah
Présentation

- Podman à un avantage.
  - ✓ Il est particulièrement compatible avec docker dans ses commandes
  - ✓ Alias docker='podman'
  - ✓ Basé sur le langage "go" (https://golang.org/)
- Gestionnaire sans deamon de conteneur (daemonless)
  - ✓ Contrairement à Docker (deamonful)
- Dépendances:
  - ✓ La librairie "runc"
  - ✓ Voir le package manager de la distribution linux (apt/yum/dnf/…)
  - ✓ slirp4netns (réseaux), fuse-overlayfs
  - ✓ Peut utiliser en complément: Systemd (lancement conteneurs)
  - ✓ Noyaux : cgroup, namespace

Conteneur: podman/buildah

Présentation: daemonless



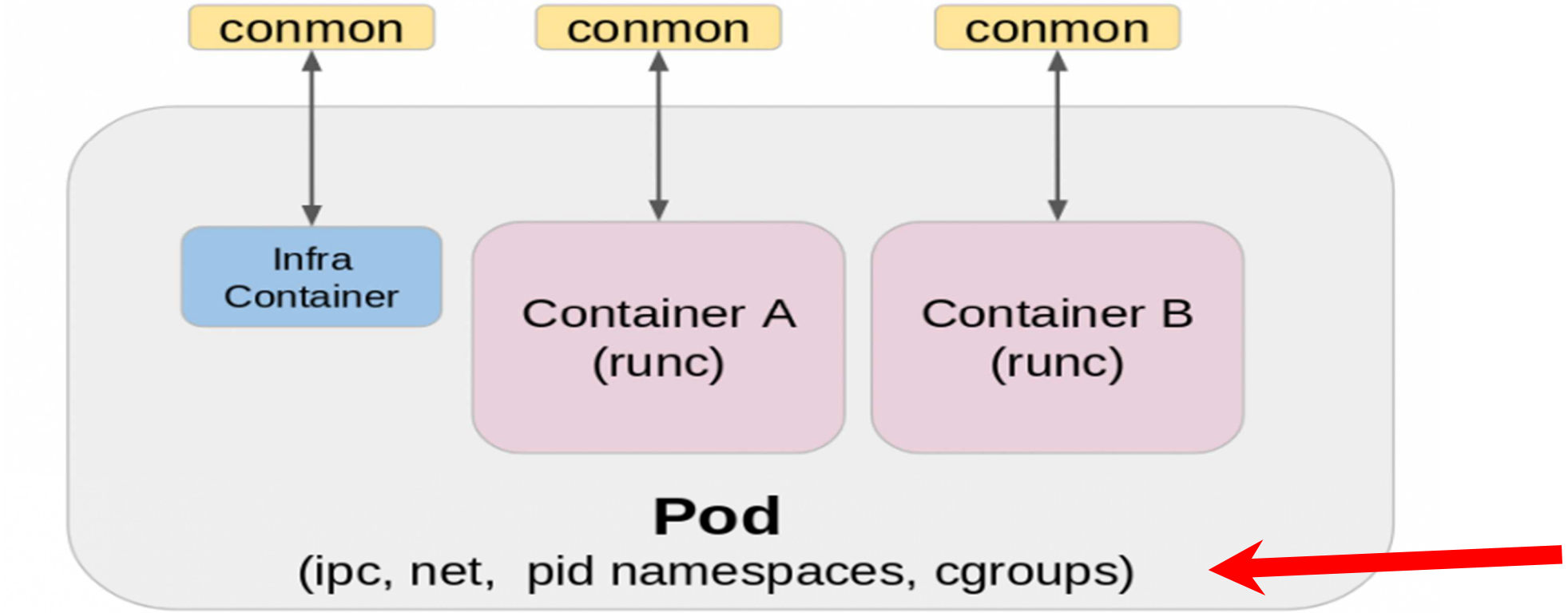
Conteneur: podman/buildah

Présentation

- Il existe des bind python (pypodman)
- Un Conteneur ne peut accéder à un namespace d'un autre conteneur
- Création d'espaces de ressources partagées (pods)
  - ✓ Même principe que pour Kubernetes et ses Pods
  - ✓ Un groupe (pod) de conteneur peut partager les mêmes ressources

Conteneur: podman/buildah

Présentation : pods



Conteneur: podman/buildah

Présentation. Podman ne fait pas …

- Le (re) démarrage automatique des conteneurs
  - ✓ Utilisation des initd (systemd) de l'OS hôte
- Swarm (inondations)
  - ✓ Utilisation de l'orchestrateur de Kubernetes
- Contrôle de l'état de santé des conteneurs
  - ✓ En dev … systemd ? Side-conteneur de monitoring ? Démon ?
- Support des Docker API
  - ✓ Pas prévu

Conteneur: podman/buildah

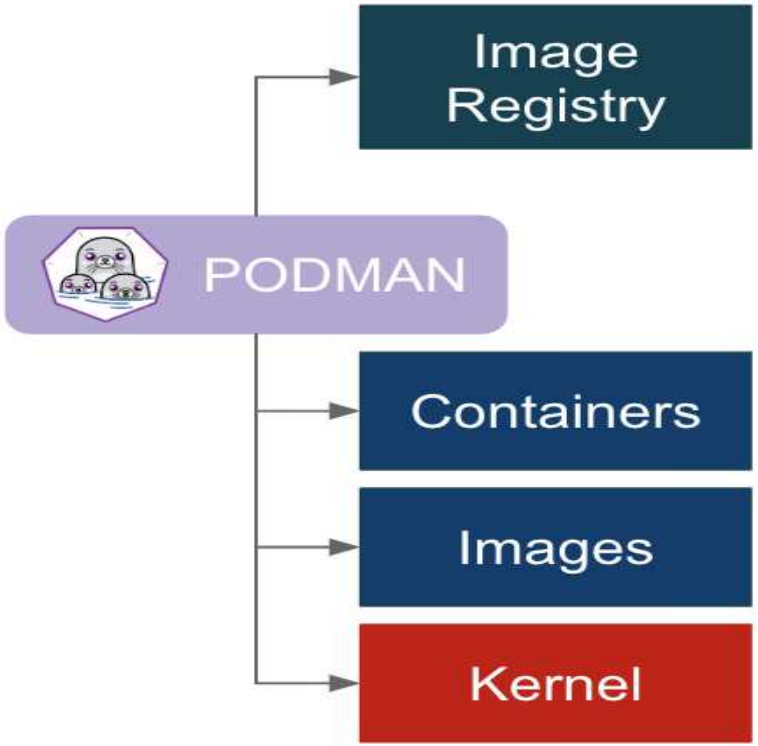
Outils

- Définition des images à déployer
  - ✓ Utilise les OCI (définition standard et opensources des images)
  - ✓ https://opencontainers.org/ et https://github.com/opencontainers
  - ✓ **Cet élément n'est pas spécifique à podman !**
  - ✓ **github → containers (podman/buildah) et opencontainer (OCI)**
- Mécanisme pour télécharger une image OCI à partir d'un registre
  - ✓ https://github.com/containers/image
  - ✓ Ne permet pas la gestion des images dans un registre
- Outils pour déployer une image sur un système de fichier virtuel COW
  - ✓ https://github.com/containers/storage
  - ✓ Rappel système: COW = Copy-On-Write
- Outils pour exécuter un conteneur (lancement)
  - ✓ Spécification OCI pour l'exécution
  - ✓ https://opencontainers.org/release-notices/v1-0-2-runtime-spec

Conteneur: podman/buildah
Outils

- Outils pour exécuter un conteneur (lancement)
  - ✓ Spécification OCI pour l'exécution
  - ✓ https://opencontainers.org/release-notices/v1-0-2-runtime-spec
  - ✓ Outils runc (de la spec OCI). Même outil que pour docker.
- Un moyen standardisé de fixer / configurer le réseau
  - ✓ "Container Networking Interface"
  - ✓ Et outils réseaux classiques (bridge, …)
- Outils pour surveiller les conteneurs
  - ✓ Habituel. Commun. Outils d'administrations, …
- Un outil client (compatible pour les paramètres avec docker)
  - ✓ https://github.com/containers/podman
- **PAS DE SERVEUR (daemonless)**

Conteneur: podman/buildah
Outils



podman image, images
skopeo

Note: Il existe des raccourcis pour certaines commandes très utilisées.
Ex: "podman image pull …" ET "podman pull …"

podman container

podman image + buildah

Réseaux rootless : slirp4netns
Rootful : vxlan, veth, vde

Conteneur: podman/buildah

Commandes: les aides

- La base.
  - ✓ RTFM (http://docs.podman.io/en/latest/).
  - ✓ "man podman"
  - ✓ "podman –help" ou "podman commande --help"

- Pour les manipulations, voir le TD podman/buildah

- "podman info"
  - ✓ Informations sur le système hôte et la configuration podman

Conteneur: podman/buildah

Commandes: qqs manipulations d'images

- "podman images"
  - ✓ Liste des images stockées localement
- "podman image search texte" ou "podman search texte"
  - ✓ Chercher une image contenant 'texte'
  - ✓ "podman search httpd --filter=is-official"
- "podman image pull nom_image" ou "podman pull nom_image"
  - ✓ Charger une image localement, elle apparait avec "podman images"
- La destruction d'une image ne peut se faire que s'il n'y a plus d'instance en cours d'exécution
  - ✓ Pour détruire, il faut stopper tous les conteneurs qui en dépendent
  - ✓ Il faut détruire tous les conteneurs qui en dépendent
  - ✓ Puis utiliser la commande "rmi" (rm image)

Conteneur: podman

Commandes: qqs manipulations de conteneur

- Gestion des conteneurs
  - ✓ "podman container"
- Les exécutions des conteneur se font par les commandes
  - ✓ "run", "start", "stop"
- On peut inspecter les conteneurs par "inspect"
- On se connecte à un conteneur
  - ✓ via la commande "exec" qui permet d'exécuter une commande root
  - ✓ Via la commande attach qui permet de s'attacher à un conteneur "détaché"
- On peut "checkpointer", "pauser", "restaurer", "migrer" les conteneurs
  - ✓ Utile pour figer un état, faire des backups ou des audits
  - ✓ Commandes "checkpoint", "restore", "pause", "stop"
- La destruction de conteneur (PAS DE L'IMAGE) → "rm"

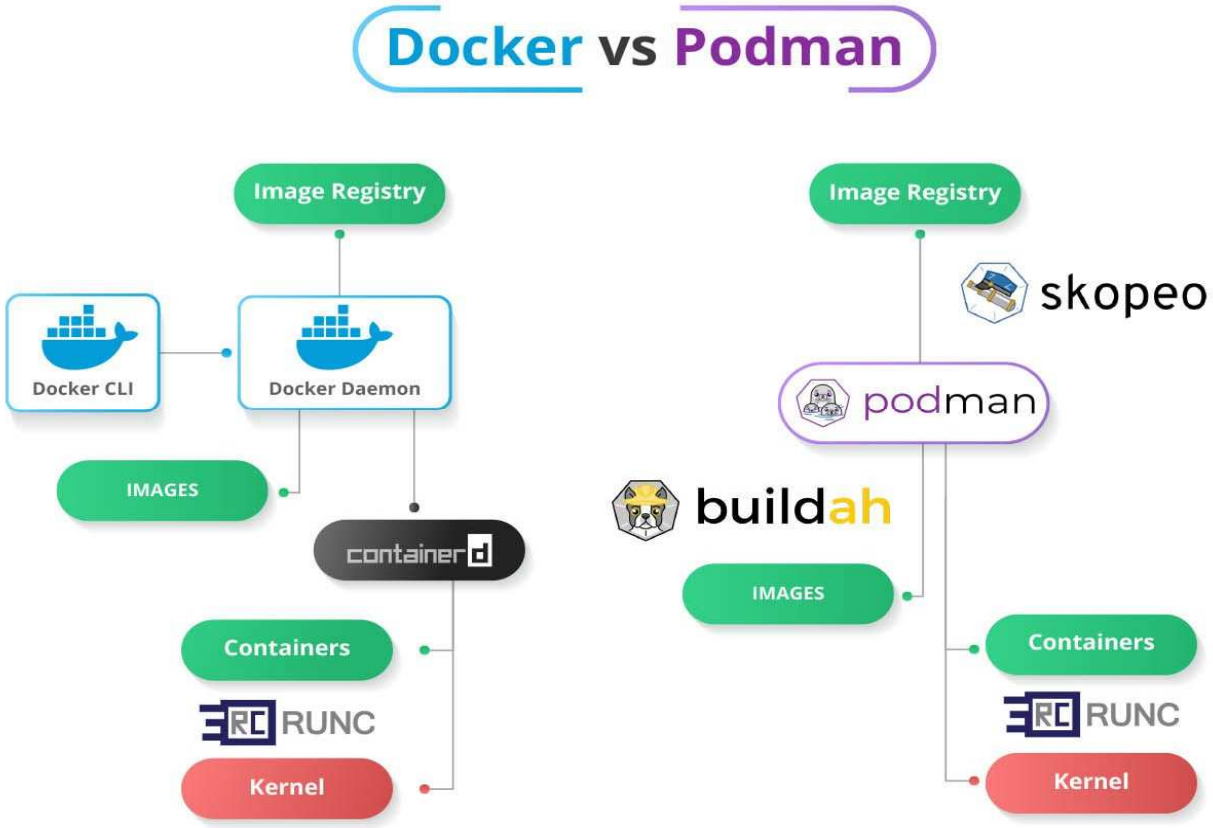
Conteneur: buildah

- Buildah
  - ✓ Une bonne partie des commandes sont compatibles avec podman
  - ✓ C’est-à-dire qu'on peut soit utiliser podman soit buildah
- Permet de
  - ✓ valider (commit) des images dans des registres ("registry") OCI distantes
  - ✓ Exporter des images (structuré en OCI)
- Buildah peut utiliser les dockerfiles pour créer des images
  - ✓ https://docs.docker.com/engine/reference/builder/
  - ✓ Commande " buildah build-using-dockerfile"

Conteneur: buildah

Conteneur: docker

Présentation: daemonful (rappel)

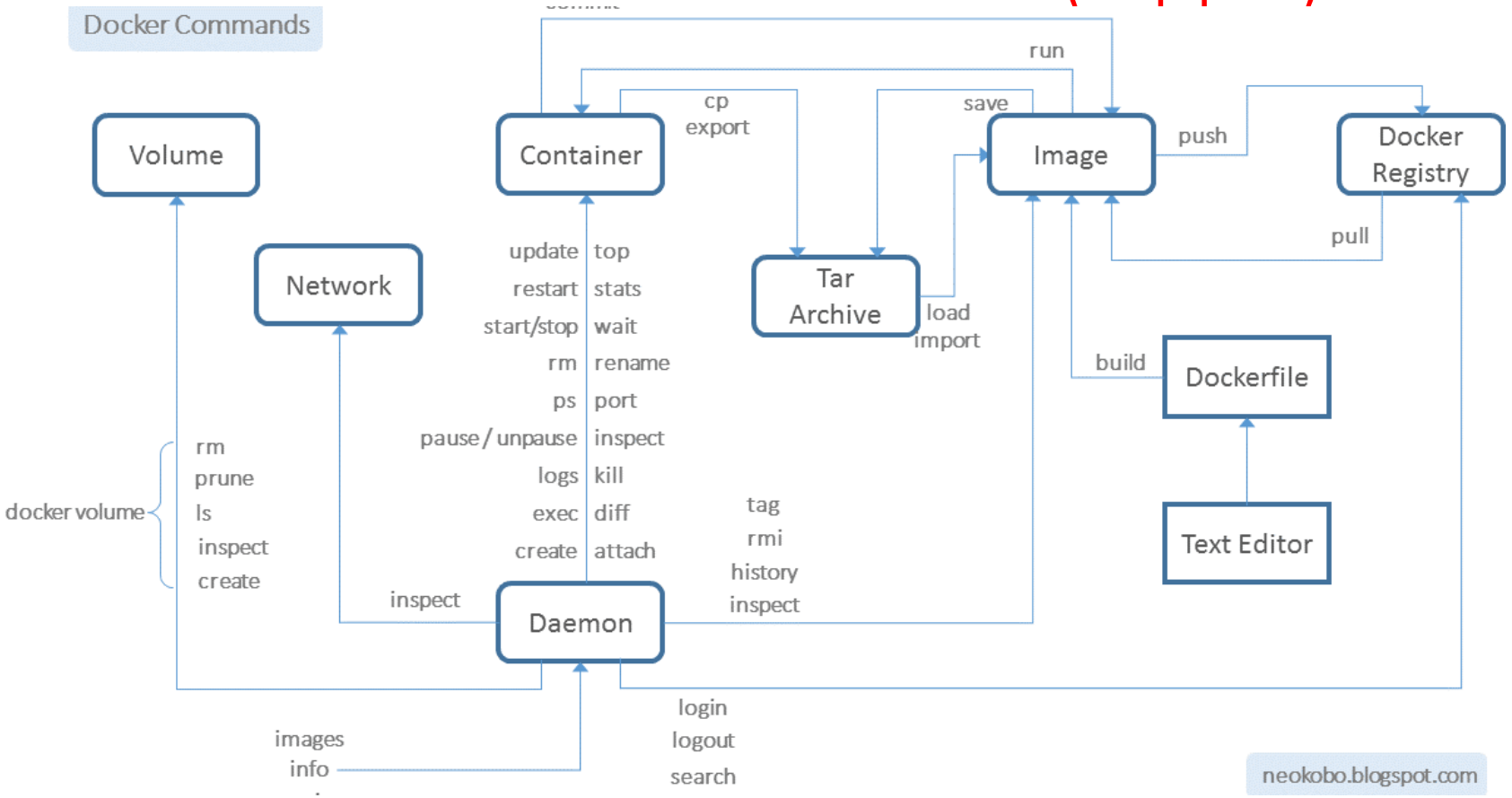


Conteneur: Docker
Présentation: daemonful (rappel)

- Docker
  - ✓ L'ancien, peut être le plus connu
  - ✓ Utilise un fichier de déploiement
    - ➢ Dockerfile
    - ➢ https://docs.docker.com/engine/reference/builder/
- Commandes similaires à podman
  - ✓ Rappel: la pub de podman dit " faire alias docker='podman' "
  - ✓ 75% de ce qui a été dit avant s'applique à Docker
- La différence : docker est "deamonful"
  - ✓ Daemon "containerd" (moteur de déploiement des conteneurs)
  - ✓ Daemon "dockerd" (execution des commandes clientes)

Conteneur: docker

Présentation: daemonful (rappel)



Conteneur: vagrant

- vragant