paul.millet@lameduse.fr

LaMeDuSe

MKU : Kubernetes, mise en œuvre

Version : 10/06/2024

Prérequis

- Connaissances dans l’administration de système Linux
- Connaissances générales en conteneurisation

Objectifs

- Comprendre le positionnement de Kubernetes et la notion d’orchestration
- Installer Kubernetes et ses différents composants
- Utiliser les fichiers descriptifs YAML
- Définir les bonnes pratiques pour travailler avec Kubernetes

Tour de table

- Présentation des membres
- Vos attentes vis-à-vis de la formation

Programme de la formation

1. Base de kubernetes
2. Installation & Configuration
3. Architecture de Kubernetes
4. API & Object API
5. Securité des API
6. Gestion d'un état avec des déploiements
7. Mise en production : Services
8. Volumes et données
9. Ingress
10. Planification
11. Logging, dépannage & surveillance
12. Définition de ressources personnalisées
13. Helm (bonus)
14. Sécurité

paul.millet@lameduse.fr

LaMeDuSe

UBE : Kubernetes, mise en œuvre

Bases de Kubernetes

1. Définition de Kubernetes
2. Structure de cluster
3. Ecosystème & OCI (Open Container Initiative)
4. Gouvernance de projet et la Cloud Native Computing Foundation (CNCF).

Un orchestrateur c’est quoi ?

- Coordination, gestion et automatisation
  - de la configuration (pré-requis, paramètres...)
  - du déploiement (allocation, provisionnement...)
  - de la maintenance (mise à jour, pannes...)

=> Augmentation de l'efficacité opérationnelle

\+ réduction des erreurs

Etude de cas : “Borg, L’orchestrateur de Google”

- Google utilisait Borg depuis les années 2000
- Problèmes de Borg
  - Très adapté aux besoins internes / difficile à adapter pour d’autres cas d’usage => Pas d’utilisateur extérieur.
  - Complexité de gestion : Exploitation difficile par des équipes non formée et courbe d’apprentissage raide.
  - Pas / peu standardisée.
  - Conception monolithique : Une base de code pour plusieurs fonctions métier.

Problème : Complexité d’utilisation, de maintenance, et d’évolution.

Kubernetes : La solution de Google

- Automatisation
- Standardisation des interfaces
- Standardisation des composants
- Open-Source = Utilisation par des entreprises autres que Google
- Faible courbe d’apprentissage
- Bonne documentation
- Séparation des composants : Conception en micro-service

=> Facile à mettre en place et à maintenir

Qui développe Kubernetes

kubernetes

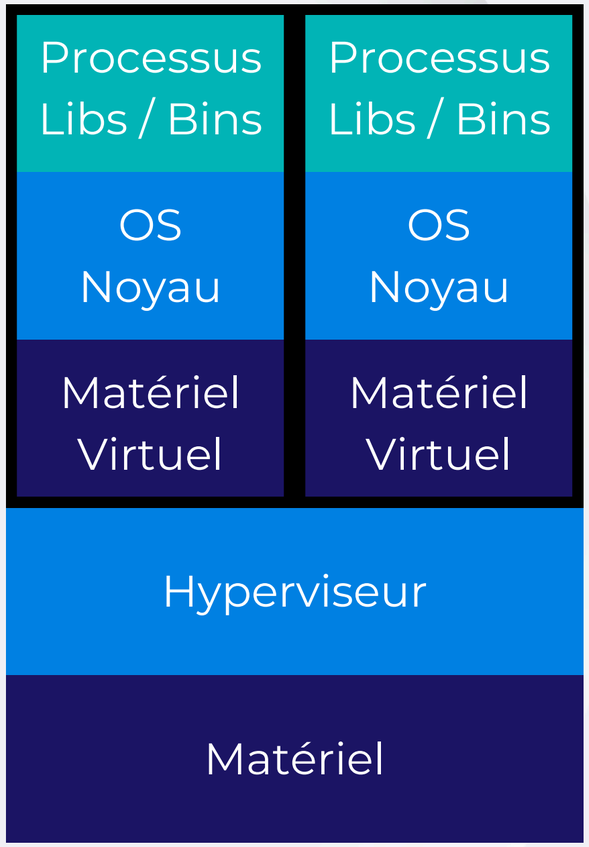
Google + Red Hat + CoreOS

CLOUD NATIVE
COMPUTING FOUNDATION

Structure d’un cluster

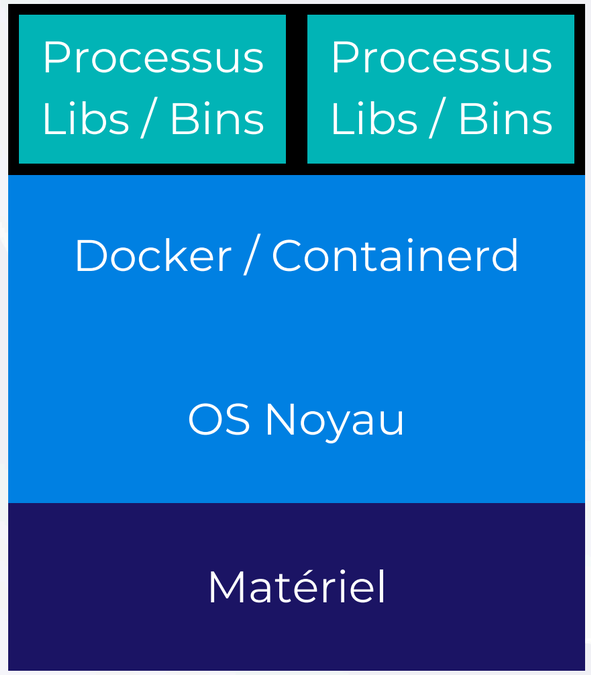
- Roles :
  - ETCD : Base de donnée
  - Control plane : Surveillance & gestion du cluster
  - Worker node : Composant de run
- Différence architecture
  - Ancienne : les composants déployé directement sur la machine
  - Nouvelle : les composants déployé dans des containers
- Maintenant : Les control plane peuvent executer des pods (kubelet & kube proxy)

Virtualisation et Conteneurisation



Virtualisation

Stockage persistant (Statefull)	Stockage non persistant par défaut (Stateless)
Séparation des fichiers et librairie	Séparation des fichiers et librairie
Système d'exploitation indépendant	Utilise les ressources du système d'exploitation (kernel)
Matériel virtualisé	Pas de virtualisation matérielle
Consommation importante Isolation totale	Consommation réduite Isolation partielle



Conteneurisation

Solutions d'installation de kubernetes

	Développement Local	**Production Cloud Public**	**Production Cloud Privé (3rd party)**	**Production Cloud Privé (Sans 3rd party)**
Solutions	MiniKube	Google Kubernetes Engine (AWS) Elastic Kubernetes Engine Azure Kubernetes Service	Rancher Kubernetes Engine 2 K3s Open-Shift	KubeAdm
Avantages	Simple d'utilisation	Simple d'utilisation Support commercial	Simple d'utilisation Support commercial	Customisable No vendor lock-in
Inconvénient	Non utilisable en production	Onéreux Vendor Lock-In	Vendor Lock-in	Complexité d'utilisation Support communautaire

Installation de Docker

- Dockerd / Containerd
  - Dockerd : moteur de commande de docker
  - Containerd : moteur d’execution de docker
- Depuis la séparation Dockerd / Containerd
  - Containerd est recommandé

Moteur de container

- Interface standard CRI (Container Runtime Interface)
  - Containerd : Moteur de Docker
  - CRI-O : Moteur léger développé a la base par Redhat
  - Gvisor : Moteur développé par Google axé sécurité et sandbox

Moteur de container

- Interface standard CNI (Container Networking Interface)
  - Cilium
  - Calico
  - Canal
  - Flannel

Moteur de container

- Interface standard CSI (Container Storage Interface)
  - Longhorn
  - Ceph (Rook)
  - Cubefs
  - Spécifique (e.g. baie SAN)

CNCF (Cloud Native Computing Foundation

- Sous fondation de la Linux Foundation
- https://landscape.cncf.io/

- Stade de maturité :
  - Graduated : Stable, adopté en production
  - Incubating : En croissance
  - Sandbox : Émergent / Projets récents

SIGs (Special Interest Groups)

- SIG-AP : Machinery / API, CRD, admission controllers
- SIG-Node : kubelet / container runtime
- SIG-Network : CNI, services, DNS
- SIG-Storage : volumes, CSI
- SIG-Security : RBAC, policies

KEPs (Kubernetes Enhancement Proposals) : processus formel pour proposer de nouvelles fonctionnalités.

paul.millet@lameduse.fr

LaMeDuSe

MKU : Kubernetes, mise en œuvre

Bases de Kubernetes

1. Débuter avec Kubernetes (Déploiement, Pod, Service)
2. Solution de tests : Minikube.
3. Solution de production : AKS
4. Kubectl : commandes de base
5. TP : Installation d’un environnement de test

Les trois objets fondamentaux

Pod : Un ou plusieurs container, unité atomique

```
apiVersion: v1
kind: Pod
metadata:
  name: mon-pod
  labels:
    app: nginx
spec:
  containers:
  - name: nginx
    image: nginx:1.25
    ports:
    - containerPort: 80
```

Les trois objets fondamentaux

Deployment : gère le cycle de vie des Pods (réplication, rolling update, rollback)

```
apiVersion: apps/v1
kind: Deployment
metadata:
 name: mon-app
spec:
 replicas: 3
 selector:
  matchLabels:
   app: mon-app
 template:
  metadata:
   labels:
    app: mon-app
  spec:
   containers:
   - name: app
    image: mon-image:v1.0
```

Les trois objets fondamentaux

Service : expose des Pods via un point d'accès stable (IP virtuelle + DNS)

```
apiVersion: v1
kind: Service
metadata:
 name: mon-service
spec:
 selector:
  app: mon-app
 ports:
 - port: 80
  targetPort: 8080
 type: ClusterIP
```

Flux de déploiement

1. kubectl apply -f deployment.yaml
2. API Server reçoit la requête
3. etcd stocke l'état désiré
4. Controller Manager crée le ReplicaSet
5. Scheduler assigne les Pods aux nœuds
6. kubelet démarre les conteneurs

Flux de déploiement

1. kubectl apply -f fichier.yaml # Créer/mettre à jour
2. kubectl get pods # Lister les pods
3. kubectl get pods -o wide # Avec plus de détails
4. kubectl describe pod mon-pod # Détails d'un pod
5. kubectl logs mon-pod # Logs d'un pod
6. kubectl exec -it mon-pod -- /bin/bash # Shell interactif
7. kubectl delete -f fichier.yaml # Supprimer

Minikube : Installation sur Linux (Debian/Ubuntu)

curl -LO https://storage.googleapis.com/minikube/releases/latest/minikube-linux-amd64
sudo install minikube-linux-amd64 /usr/local/bin/minikube
minikube version
minikube dashboard # ouvre une dashboard kubernetes

Cliquer sur default en haut a gauche, puis sur tout les espaces de noms, puis pods.

Utilisation de base

```
minikube start              # Démarrer
minikube status             # Statut
minikube dashboard          # Dashboard K8s
minikube stop               # Arrêter
minikube delete             # Supprimer
```

Addons

```
minikube addons list
minikube addons enable metrics-server # Pour le hpa
minikube addons enable ingress # Pour l’ingress
minikube addons enable dashboard # Pour la dashboard kube
```

Solution de production : AKS

Azure Kubernetes Service (AKS) est le service Kubernetes **managé** de Microsoft Azure. Le Control Plane est entièrement géré par Azure (gratuit), seuls les worker nodes sont facturés.

Solution de production : AKS

Control Plane managé : pas de gestion de etcd, kube-apiserver…
Mises à jour automatiques du Control Plane
Intégration Azure : ACR, Azure AD, Monitor, Key Vault…
Node pools : mix de VM différentes (GPU, Spot, ARM64…)

Création d'un cluster AKS

```
# Connexion
az login
# Créer le resource group
az group create --name rg-k8s-formation --location francecentral
```

Création d'un cluster AKS

```
# Créer le cluster
az aks create \
 --resource-group rg-k8s-formation \
 --name aks-formation \
 --node-count 2 \
 --node-vm-size Standard_D2s_v3 \
 --enable-managed-identity \
 --generate-ssh-keys

# Récupérer les credentials kubectl
az aks get-credentials \
 --resource-group rg-k8s-formation \
 --name aks-formation
```

Cluster Autoscaler

```
az aks nodepool update \
 --resource-group rg-k8s-formation \
 --cluster-name aks-formation \
 --name nodepool1 \
 --enable-cluster-autoscaler \
 --min-count 2 \
 --max-count 10
```

Intégration Azure Container Registry

```
# Créer un ACR
az acr create \
 --resource-group rg-k8s-formation \
 --name acrformation \
 --sku Basic

# Attacher ACR à AKS
az aks update \
 --resource-group rg-k8s-formation \
 --name aks-formation \
 --attach-acr acrformation
```

Accès au cluster kubernetes : CLI (Kubectl)

- kubectl est l'outil en ligne de commande pour interagir avec les clusters Kubernetes.
- Il vous permet de déployer et de gérer des applications sur Kubernetes ainsi que de visualiser et diagnostiquer les ressources et les clusters.

Accès au cluster kubernetes : CLI (Kubectl)

- kubectl get <type_ressource> [nom_ressource] :
  - Affiche les informations sur les ressources Kubernetes. Si aucun nom n'est spécifié, toutes les ressources de ce type seront affichées.
  - Exemples :
    - kubectl get pods - Liste tous les pods.
    - kubectl get services - Liste tous les services.
  - Options :
    - -o <format>
      - yaml : sortie en yaml
      - json : sortie en json
      - wide : colonnes étendue
    - -n <namespace>

Accès au cluster kubernetes : CLI (Kubectl)

- kubectl describe <type_ressource> <nom_ressource> :
  - Affiche des informations détaillées sur une ressource spécifique, y compris ses événements et son état actuel.
  - Exemple :
    - kubectl describe pod my-pod - Affiche les détails du pod nommé "my-pod".

Accès au cluster kubernetes : CLI (Kubectl)

- kubectl create -f <fichier_yaml> :
  - Crée une ressource Kubernetes à partir d'un fichier de configuration YAML.
    - Exemple :
      - kubectl create -f deployment.yaml - Crée une ressource à partir du fichier deployment.yaml.

Accès au cluster kubernetes : CLI (Kubectl)

- kubectl apply -f <fichier_yaml> :
  - Applique les modifications définies dans un fichier YAML en effectuant un strategic merge pour mettre à jour les ressources existantes ou en créer de nouvelles.
  - Exemple :
    - kubectl apply -f service.yaml - Applique les configurations du fichier service.yaml.

Accès au cluster kubernetes : CLI (Kubectl)

- kubectl replace -f <fichier_yaml> :
  - Applique les modifications définies dans un fichier YAML en effectuant un remplacement pour mettre à jour les ressources existantes.
  - Exemple :
    - kubectl replace -f service.yaml - Applique les configurations du fichier service.yaml.

Accès au cluster kubernetes : CLI (Kubectl)

- kubectl delete <type_ressource> <nom_ressource> :
  - Supprime une ressource spécifique. Vous pouvez également utiliser un fichier YAML pour supprimer des ressources.
  - Exemples :
    - kubectl delete pod my-pod - Supprime le pod nommé "my-pod".
    - kubectl delete -f deployment.yaml - Supprime les ressources définies dans le fichier deployment.yaml.

Accès au cluster kubernetes : CLI (Kubectl)

- kubectl logs <nom_pod> [-c <nom_container>] :
  - Affiche les logs du pod spécifié. Vous pouvez également spécifier un conteneur particulier dans un pod multi-containers.
  - Exemple :
    - kubectl logs my-pod - Affiche les logs du pod nommé "my-pod".
    - kubectl logs my-pod -c my-container - Affiche les logs du conteneur nommé "my-container" dans le pod "my-pod".

Accès au cluster kubernetes : CLI (Kubectl)

- kubectl exec -it <nom_pod> -- <commande> :
  - Exécute une commande dans un conteneur en cours d'exécution. Utilisez -it pour obtenir un terminal interactif.
  - Exemple :
    - kubectl exec -it my-pod -- /bin/sh - Ouvre un shell interactif dans le pod nommé "my-pod".
  - Note :
    - Option --container pour spécifier le container

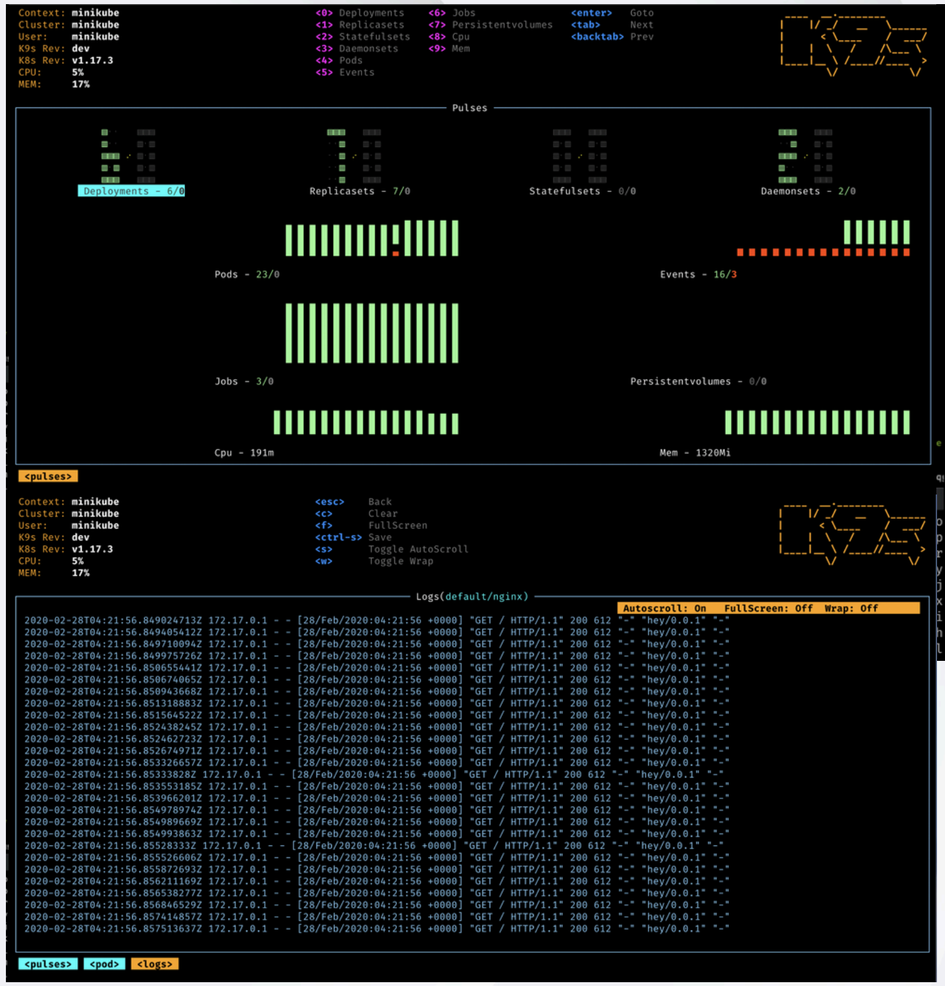
Accès au cluster kubernetes : CLI (Kubectl)

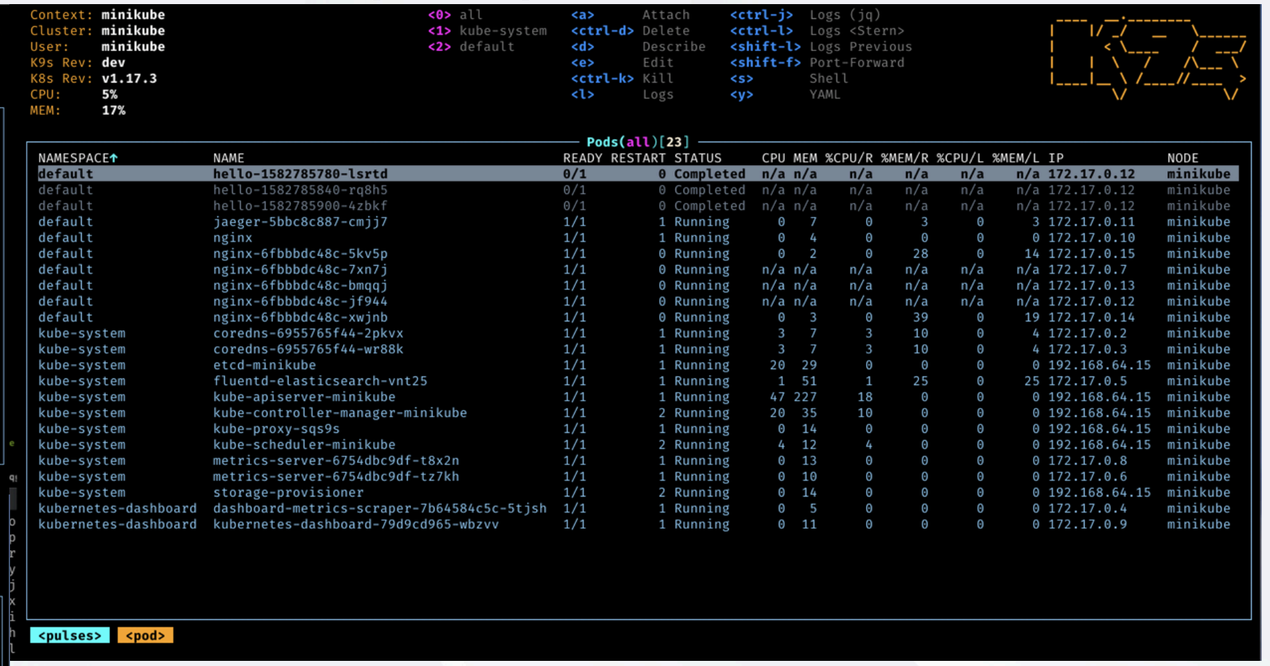
- kubectl cluster-info :
  - Affiche des informations sur les services principaux du cluster, y compris l'URL de l'API et des composants comme le dashboard.
  - Exemple :
    - kubectl cluster-info - Affiche des informations sur le cluster Kubernetes.

Accès au cluster kubernetes : CLI (Kubectl)

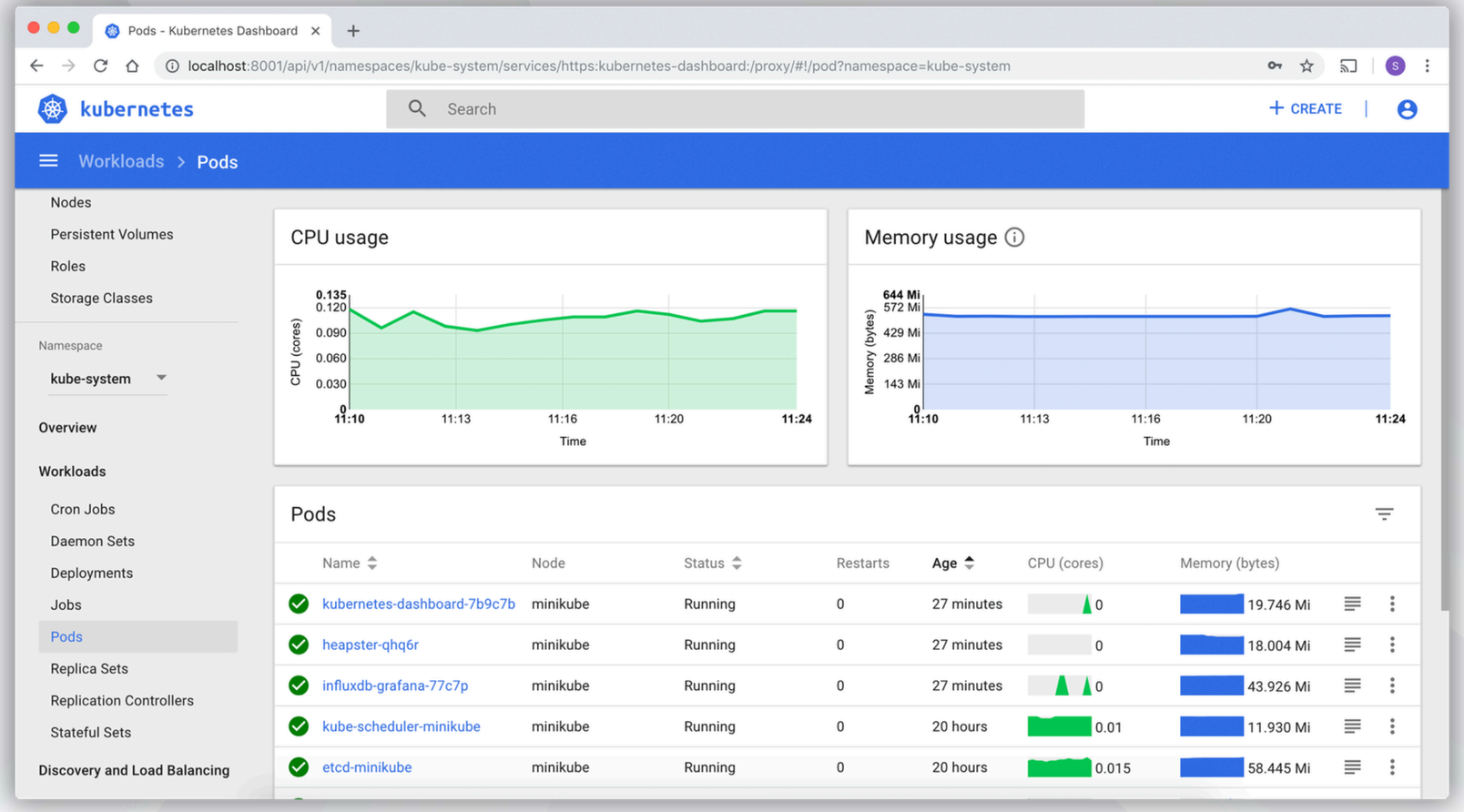
- kubectl config current-context :
  - Affiche le contexte Kubernetes actuellement utilisé. Le contexte détermine le cluster, l'utilisateur et le namespace par défaut.
  - Exemple :
    - kubectl config current-context - Affiche le contexte actuel configuré dans kubectl

Accès au cluster kubernetes : CLI (K9s)

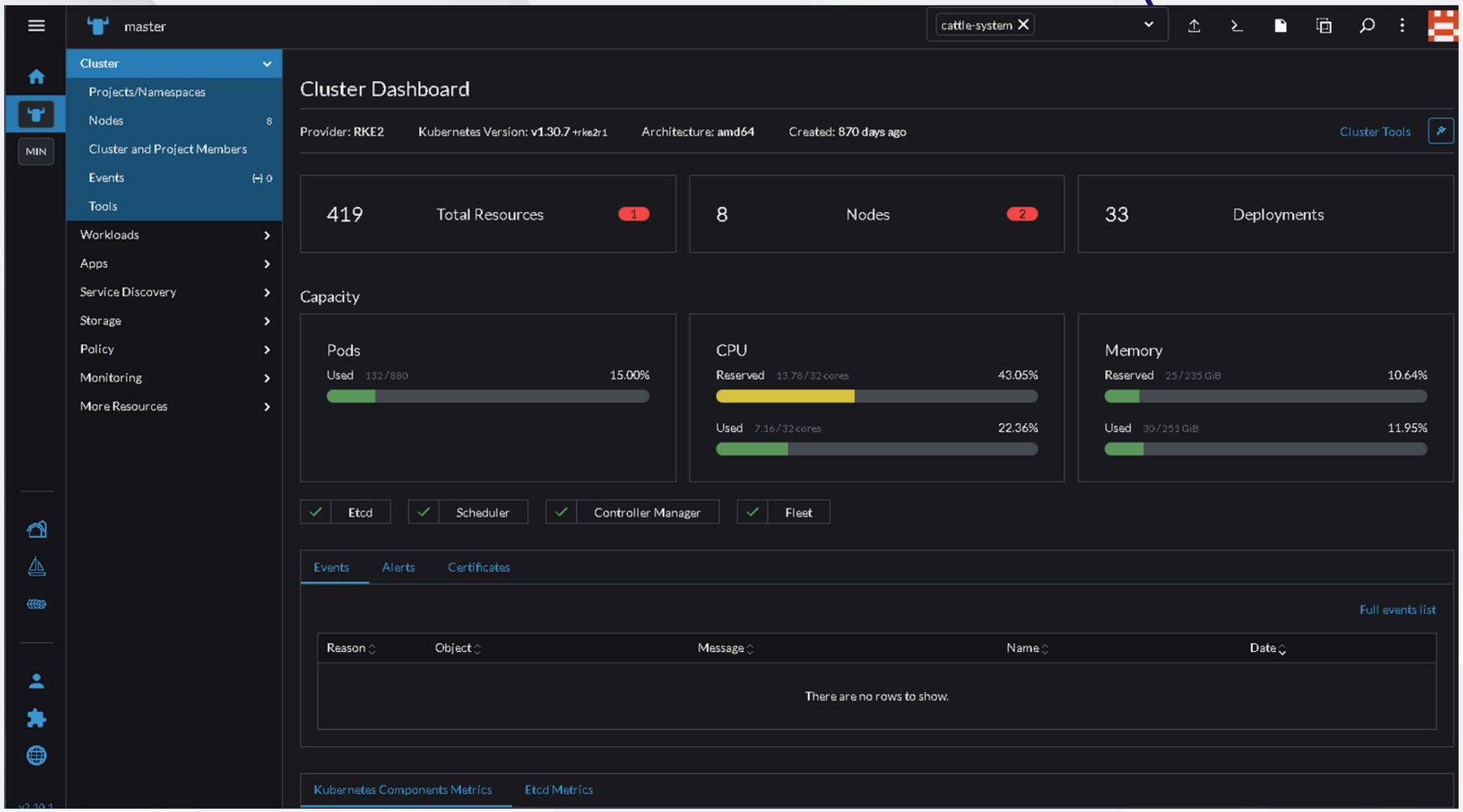




Accès au cluster kubernetes : GUI (Dashboard)



Accès au cluster kubernetes : GUI Rancher (Dashboard)



paul.millet@lameduse.fr

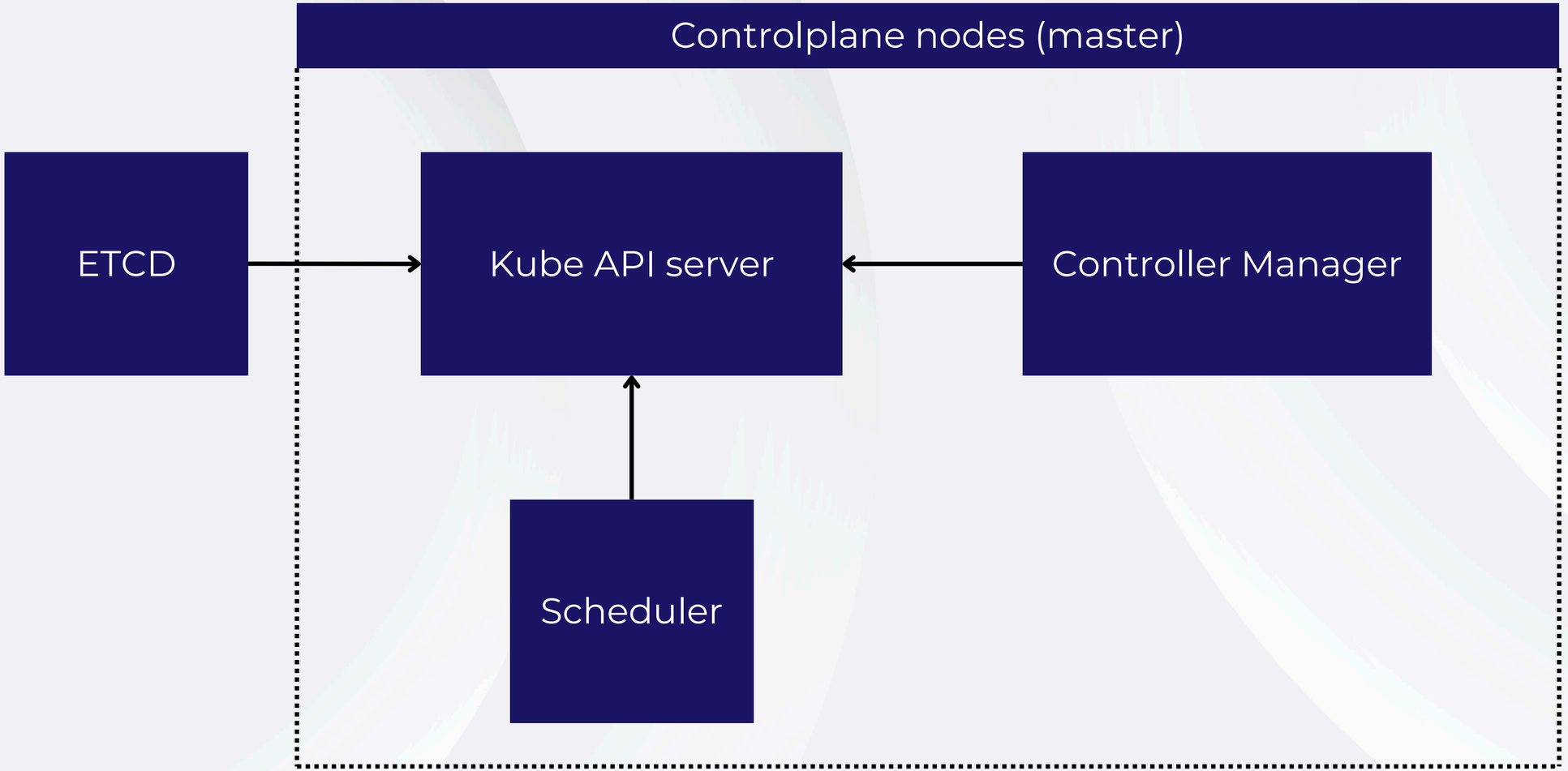
LaMeDuSe

MKU : Kubernetes, mise en œuvre

Architecture Kubernetes

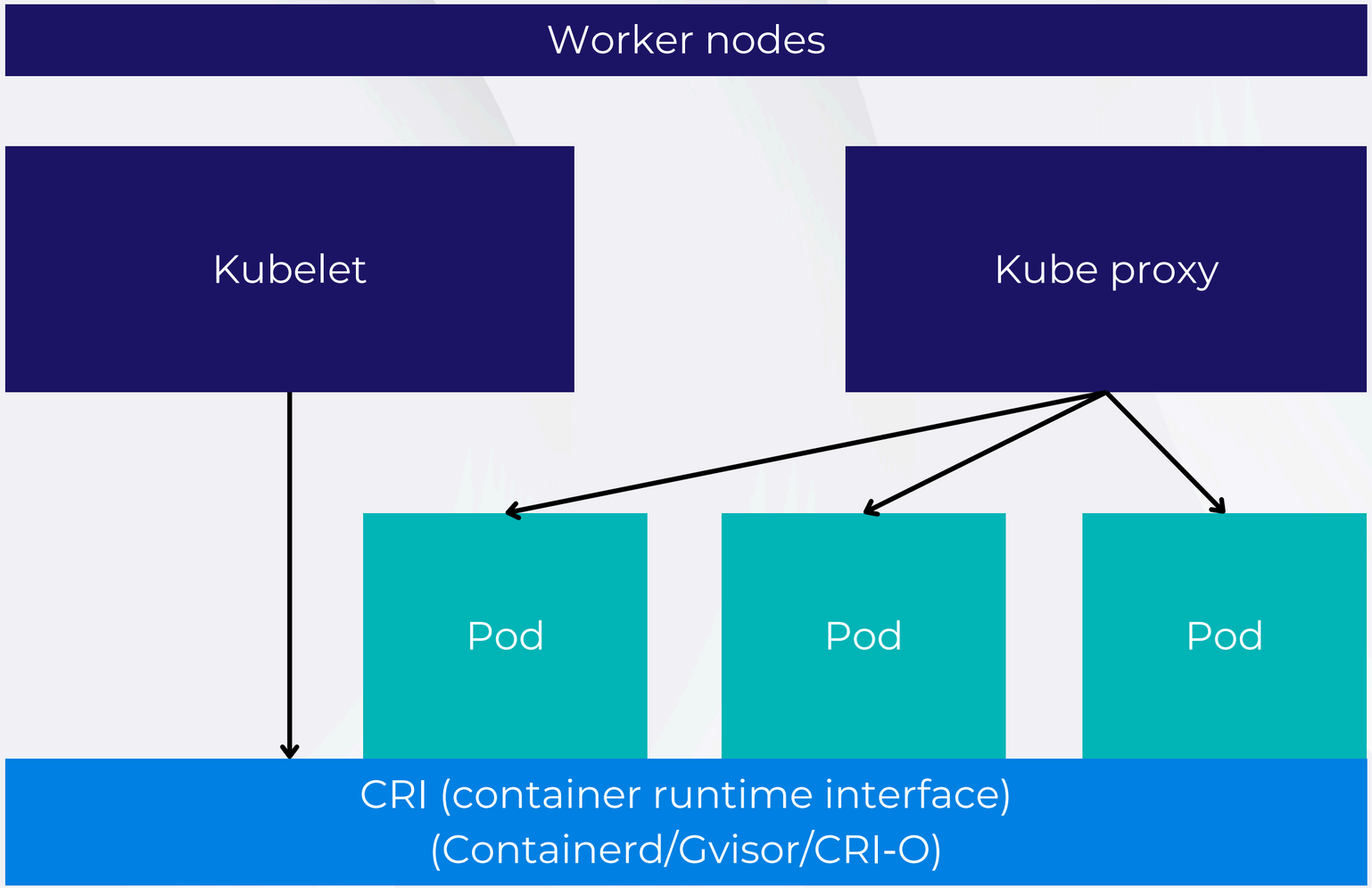
1. Composants du master node : API server, scheduler, controller manager, etc.
2. Architecture d'un noeud : Kubelet, le moteur de conteneur (docker), Kube-proxy.
3. Mise en réseau (calico, cilium)
4. Bonnes pratiques de productions
5. TP : Déploiement première application avec Kubernetes

Composants du master node

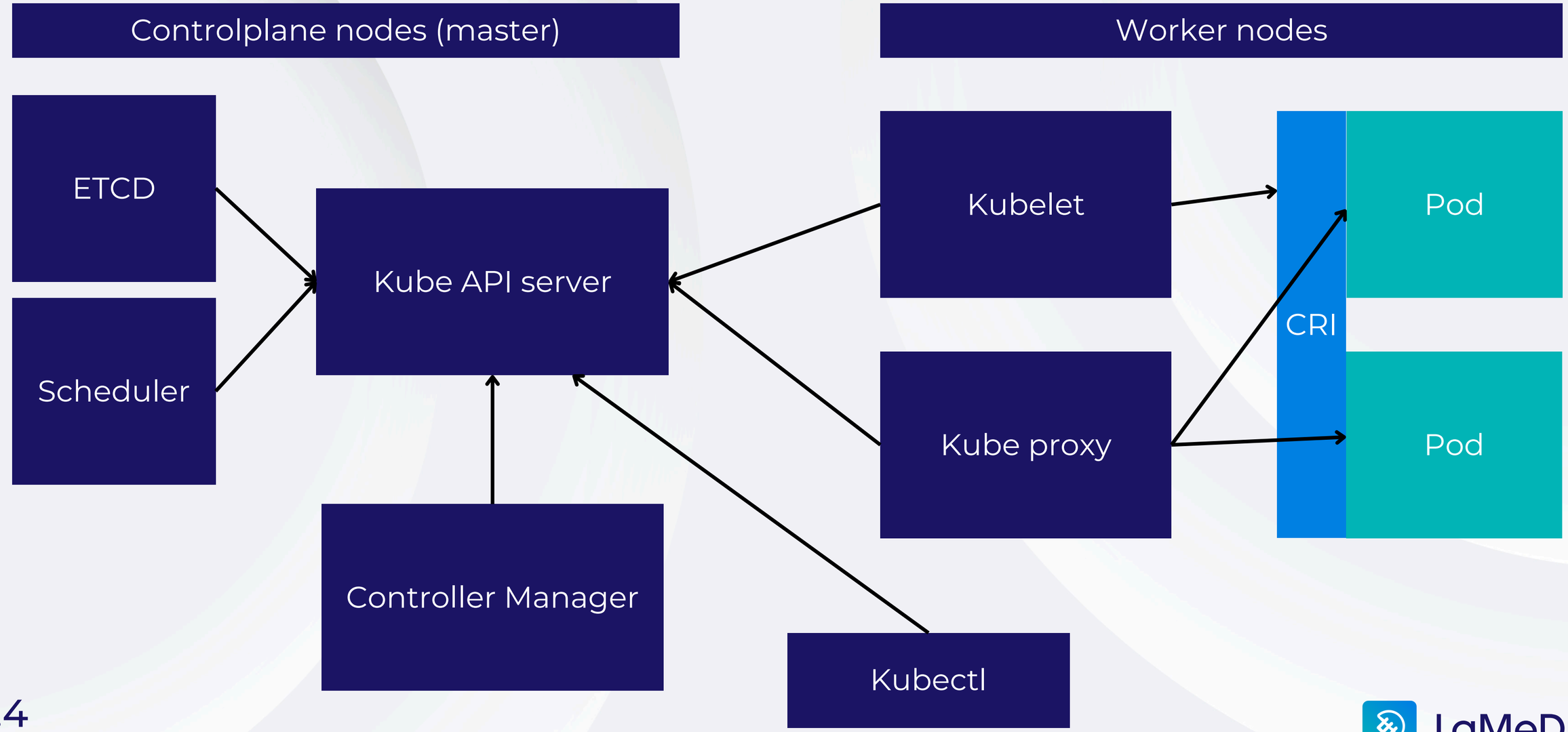


ETCD peut être déployé à l'extérieur du cluster kubernetes ou dans celui-ci

Architecture d'un noeud



Architecture Kubernetes



kube-apiserver

Point d'entrée **unique** pour toutes les opérations.

- RESTful API
- validation
- authentification

etcd

Base de données **distribuée** de type clé-valeur (basée sur Raft pour le consensus).

Stocke : tout l'état du cluster (Pods, Services, ConfigMaps, Secrets, RBAC...).

etcd est le composant le plus critique du cluster. Sa perte = perte de tout l'état. **Sauvegarder etcd régulièrement !**

kube-scheduler

Assigne les nouveaux Pods aux nœuds en deux phases :

1. **Filtrage** : quels nœuds peuvent accueillir ce Pod ? (ressources, taints, affinités)
2. **Score** : quel nœud est le meilleur candidat ? (ressources dispo, spreading, affinité, topologie)
nager

kube-controller-manager

Regroupe plusieurs **boucles de contrôle** (controllers) :

Controller => Rôle

Deployment Controller => Gère les ReplicaSets
ReplicaSet Controller => Maintient le nombre de Pods
Node Controller => Surveille l'état des nœuds
Job Controller => Gère les Jobs jusqu'à completion
Endpoint Controller => Peuple les Endpoints des Services
ServiceAccount Controller => Crée les ServiceAccounts par défaut

Boucle de réconciliation (Reconciliation Loop)

C'est le principe fondamental de Kubernetes : tout est convergence vers l'état désiré.

Haute disponibilité du Control Plane

Quorum etcd : minimum **3 nœuds** (tolérance à 1 panne), **5 nœuds** recommandé.

Le modèle réseau Kubernetes

Règles fondamentales :

1. Chaque Pod a sa propre adresse IP
2. Tous les Pods peuvent communiquer entre eux sans NAT
3. Les agents sur les nœuds peuvent communiquer avec tous les Pods

Container Network Interface (CNI)

Interface standard pour les plugins réseau. Le kubelet appelle le plugin CNI pour :

- Créer l'interface réseau du Pod
- Assigner une IP
- Configurer le routage

Calico

Solution réseau et politique réseau **layer 3** (routage BGP).

Cilium

Solution réseau basée sur eBPF : plus performante qu'iptables.

Avantages :
- Performances supérieures (bypass iptables)
- Observabilité native (Hubble)
- Network policies L7 (HTTP, gRPC, Kafka)
- Encryption transparente (WireGuard)

Comparaison Calico vs Cilium

Critère	Calico	Cillium
Technologie	iptables/BGP	eBPF
Performances	Bonne	Excellentes
Network Policy	Standard k8s	Standard k8s + L7
Observabilité	Basique	Hubble
Maturité	Très mature	Mature
Complexité	Modéré	Elevé

Bonne pratique : Ressources (Requests & Limits)

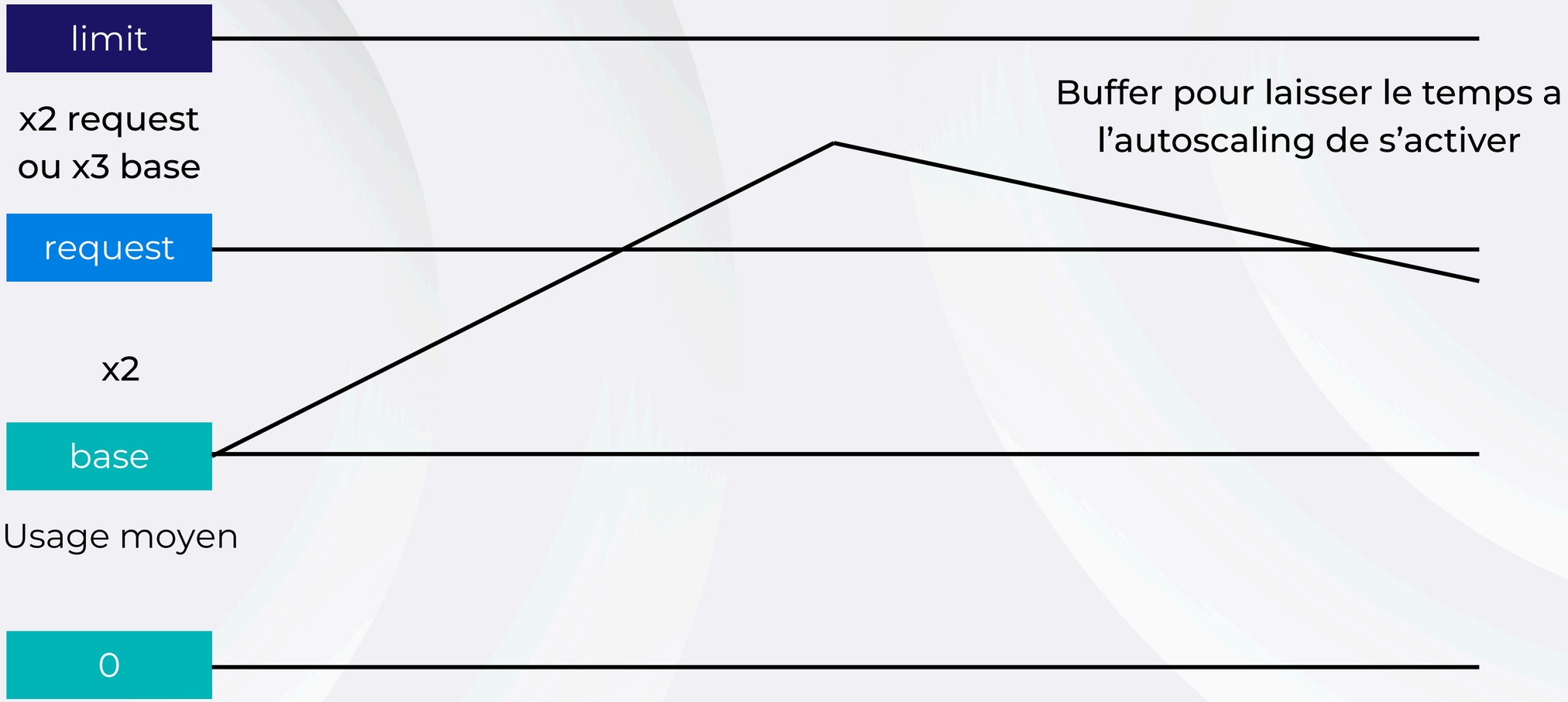
Toujours définir des **requests** et **limits** pour chaque conteneur.

```
resources:
 requests:
  cpu: 50m
  memory: 64Mi
 limits:
  cpu: 100m
  memory: 128Mi
```

Requètes & Limites

- Concepts Clés :
  - Requêtes de Ressources : Quantité minimale de CPU et de mémoire garantie pour un conteneur.
  - Limites de Ressources : Quantité maximale de CPU et de mémoire qu'un conteneur peut utiliser.

Requètes & Limites



Les namespaces et les quotas.

- Qu'est-ce qu'un Quota de Ressources ?
  - Les quotas de ressources limitent la consommation de ressources (CPU, mémoire, etc.) dans un namespace pour éviter la surutilisation et assurer une gestion équitable des ressources.
- Objectifs des Quotas :
  - Prévention de la Surutilisation : Éviter qu'un namespace consomme toutes les ressources disponibles du cluster.
  - Gestion Équitable : Assurer que les ressources sont partagées équitablement entre différents namespaces.
  - Contrôle Budgétaire : Aider à respecter les contraintes budgétaires en contrôlant l'utilisation des ressources.

Les namespaces et les quotas.

```
apiVersion: v1
kind: ResourceQuota
metadata:
  name: my-resource-quota
  namespace: my-namespace
spec:
  hard:
    requests.cpu: "2"
    requests.memory: "4Gi"
    limits.cpu: "4"
    limits.memory: "8Gi"
    pods: "10"
    services: "5"
```

Les namespaces et les quotas.

```
apiVersion: v1
kind: LimitRange
metadata:
  name: enforce-resource-requests
spec:
  limits:
    - type: Container
      default:  # default limit
        cpu: "100m"
        memory: "128Mi"
      defaultRequest:  # default request
        cpu: "100m"
        memory: "128Mi"
      min:  # minimum resources for both requests and limits
        cpu: "100m"  # 100 mili-cores (100m = 0.1 cores)
        memory: "128Mi"  # 128 MB
      max:  # maximum resources for both requests and limits
        cpu: "4000m"  # 4 cores
        memory: "8Gi"  # 8 GB
      maxLimitRequestRatio:  # how much the limit can exceed the request (default is 2x)
        cpu: 2
        memory: 2
```

paul.millet@lameduse.fr

LaMeDuSe

UBE : Kubernetes, mise en œuvre

Accès au cluster kubernetes : APIs

- L'API de Kubernetes est le cœur du système Kubernetes.
- Elle permet aux utilisateurs, applications et outils de communiquer avec le cluster.
- Toutes les interactions avec le cluster, que ce soit via kubectl, des outils tiers ou des services internes, passent par cette API.

Accès au cluster kubernetes : APIs

- L'API expose différents types de ressources qui représentent des objets dans le cluster, tels que des Pods, Services, Deployments, ConfigMaps, etc.
- Chaque ressource a une représentation en JSON ou YAML, avec des champs spécifiques pour configurer son état et son comportement.

Accès au cluster kubernetes : APIs

- L'API de Kubernetes est organisée en groupes et versions pour assurer la compatibilité et l'évolution du système.
  - core/v1 : Le groupe API de base contenant les ressources fondamentales (ex : Pods, Services).
  - apps/v1 : Contient des ressources pour les déploiements, les DaemonSets, etc.
  - batch/v1 : Utilisé pour les Jobs et CronJobs.
- Chaque groupe API peut évoluer indépendamment avec des versions comme v1, v1beta1, etc.

Accès au cluster kubernetes : APIs

L'API de Kubernetes supporte les opérations CRUD (Create, Read, Update, Delete) sur les ressources.

- POST : Créer une nouvelle ressource.
- GET : Lire les détails d'une ressource ou lister des ressources.
- PUT/PATCH : Mettre à jour une ressource existante.
- DELETE : Supprimer une ressource.

Ces opérations sont exécutées via des appels HTTP RESTful.

Accès au cluster kubernetes : APIs

Les endpoints de l'API Kubernetes sont des URL spécifiques utilisées pour interagir avec les ressources.

- /api/v1/namespaces/{namespace}/pods : Lister ou créer des pods dans un namespace donné.
- /apis/apps/v1/deployments : Gérer les déploiements.
- /api/v1/nodes : Obtenir des informations sur les nœuds du cluster.

Les requêtes peuvent être effectuées en utilisant des outils comme curl, ou à travers des SDKs dans différents langages de programmation.

Accès au cluster kubernetes : APIs

L'accès à l'API Kubernetes est protégé par des mécanismes d'authentification et d'autorisation.

- Authentification : Via des tokens, certificats client ou OAuth.
- Autorisation : Basée sur des rôles (RBAC - Role-Based Access Control) qui définissent quelles actions un utilisateur ou un service peut effectuer sur quelles ressources.

Accès au cluster kubernetes : APIs

- Les labels sont des paires clé-valeur attachées aux ressources Kubernetes.
- Ils permettent de filtrer et de sélectionner des ensembles de ressources via des sélecteurs (label selectors).
- Exemple : Sélectionner tous les pods avec le label app=frontend.
  - Endpoint : /api/v1/namespaces/{namespace}/pods?labelSelector=app=frontend.
- Les labels sont essentiels pour la gestion des déploiements, la mise en place de services, et l'organisation des ressources

Accès au cluster kubernetes : APIs

- Les annotations sont similaires aux labels, mais servent à attacher des métadonnées non modifiables par les systèmes internes de Kubernetes.
- Elles sont utilisées pour stocker des informations supplémentaires sur les ressources, comme des traces, des configurations ou des informations liées à des outils tiers.

Accès au cluster kubernetes : APIs

- Kubernetes utilise des contrôleurs qui implémentent des boucles de réconciliation pour garantir que l'état actuel des ressources correspond à l'état désiré spécifié par les utilisateurs.
  - Exemple : Le Deployment Controller s'assure que le nombre souhaité de réplicas d'un pod est toujours en cours d'exécution.
- Les contrôleurs sont des consommateurs de l'API Kubernetes qui surveillent les ressources, prennent des décisions, et effectuent les actions nécessaires.

Accès au cluster kubernetes : APIs

- Les Admission Controllers sont des composants qui interceptent les requêtes à l'API avant que les objets ne soient stockés. Ils peuvent modifier, rejeter ou valider des requêtes selon des règles définies.
  - Exemple : L'Admission Controller PodSecurityPolicy peut bloquer la création de pods non conformes aux politiques de sécurité.
- Les Admission Controllers jouent un rôle crucial dans la sécurisation et la gouvernance du cluster.

Accès au cluster kubernetes : APIs

- Kubernetes permet d'étendre l'API avec des Custom Resource Definitions (CRDs) et API Aggregation.
  - CRDs : Permettent aux utilisateurs de définir leurs propres types de ressources.
  - API Aggregation : Permet l'ajout de nouveaux groupes d'API, servant des ressources custom via des serveurs d'API supplémentaires.
- Ces mécanismes permettent aux développeurs d'étendre Kubernetes pour répondre à des besoins spécifiques ou créer des opérateurs.

Accès au cluster kubernetes : APIs

L'API de Kubernetes expose plusieurs endpoints pour le monitoring et le debugging du cluster.

- /metrics : Expose les métriques Prometheus pour les composants du cluster.
- /logs : Permet de récupérer les logs des nœuds et composants.
- /healthz : Points de vérification de la santé des composants.

Ces outils sont essentiels pour assurer la stabilité et la performance du cluster.

Versioning des APIs

v1alpha1 : Instable, peut être supprimée, Tests uniquement
v1beta1 , Relativement stable , Staging
v1 , Stable, rétrocompatibilité garantie , Production

Ne jamais utiliser des API Alpha en production.

Objets Kubernetes : volume, service, pod, etc.

- Pod
  - Unité de base dans Kubernetes, représentant un ou plusieurs conteneurs qui partagent des ressources (stockage, réseau).
  - Utilisé pour exécuter des applications conteneurisées.
- Deployment
  - Gère des ensembles de pods réplicas, permettant des mises à jour sans interruption et le scaling.
  - Fournit des fonctionnalités comme les rolling updates et les rollbacks.

Objets Kubernetes : volume, service, pod, etc.

- Service
  - Expose les pods pour permettre la communication réseau, interne ou externe.
  - Types principaux : ClusterIP, NodePort, LoadBalancer.
- Namespace
  - Permet de diviser les ressources d'un cluster en environnements logiques isolés.
  - Utilisé pour organiser et gérer des applications distinctes ou des équipes.

Objets Kubernetes : volume, service, pod, etc.

- ConfigMap
  - Stocke des configurations sous forme de paires clé-valeur, indépendamment du code d'application.
  - Injecte des configurations dans les pods via des variables d'environnement ou des volumes.
- Secret
  - Stocke des informations sensibles comme des mots de passe, des clés SSH, des certificats.
  - Similaire à ConfigMap mais sécurisé (base64 encodé).

Objets Kubernetes : volume, service, pod, etc.

- PersistentVolume (PV)
  - Représente un stockage physique dans le cluster, utilisé pour persister les données des pods.
  - Exemples : stockage local, disques réseau, NFS, etc.
- PersistentVolumeClaim (PVC)
  - Demande de stockage faite par les utilisateurs pour consommer des PersistentVolumes.
  - Les pods utilisent les PVC pour accéder au stockage.

Objets Kubernetes : volume, service, pod, etc.

- ServiceAccount
  - Fournit une identité pour les pods afin d'interagir avec l'API Kubernetes.
  - Utilisé pour limiter les permissions d'accès aux ressources du cluster.
- Role et RoleBinding
  - Role : Définit des permissions spécifiques (lecture, écriture) sur des ressources dans un namespace.
  - RoleBinding : Associe un rôle à un utilisateur ou un groupe.

Objets Kubernetes : volume, service, pod, etc.

- ClusterRole et ClusterRoleBinding
  - ClusterRole : Similaire à Role mais s'applique à l'ensemble du cluster.
  - ClusterRoleBinding : Associe un ClusterRole à des utilisateurs ou groupes à l'échelle du cluster.

Objets Kubernetes : volume, service, pod, etc.

- ReplicaSet
  - Assure un nombre spécifié de réplicas de pods en cours d'exécution à tout moment.
  - Géré principalement par des déploiements.
- DaemonSet
  - Exécute un pod sur chaque nœud du cluster (ou sous-ensemble de nœuds).
  - Utilisé pour des services de cluster comme les agents de surveillance ou de logging.

Objets Kubernetes : volume, service, pod, etc.

- StatefulSet
  - Gère des applications nécessitant un état stable ou un identifiant réseau unique.
  - Utilisé pour des bases de données, des caches, ou des applications à état.
- Job et CronJob
  - Job : Exécute une tâche jusqu'à sa complétion (pods temporaires).
  - CronJob : Planifie des Jobs à exécuter à des intervalles réguliers (similaire aux cron jobs Linux).

Objet statefull, objet stateless

- Stateful : Applications qui maintiennent un état ou des données persistantes entre les sessions ou les redémarrages.
  - Base de donnée, stockage de donnée
- Stateless : Applications qui ne conservent pas d'état entre les requêtes et peuvent être recréées ou déplacées sans impact.
  - Serveur web, serveur applicatif

Solution du deployment.

- Deployment : gère la création et la mise à jour de groupes de pods sans état (stateless).
- StatefulSet : utilisé pour les applications stateful qui nécessitent des identités stables et un stockage persistant.
- DaemonSet : garantit que tous (ou certains) nœuds exécutent une copie d’un pod donné.
- ReplicaSet : assure un nombre spécifique de pods en cours d'exécution à tout moment.
- Job : utilisé pour les tâches courtes et exécutées une seule fois (one-off).
- CronJob : planifie des tâches périodiques ou à exécution unique basées sur des expressions cron.

paul.millet@lameduse.fr

LaMeDuSe

UBE : Kubernetes, mise en œuvre

Les 3 étapes d'accès

1. Requête
2. Authentication
3. Authorization
4. Admission Control
5. API

Méthodes d'authentification

Certificats X.509 :

- Cert client signé par CA Kubernetes
- Admins, composants

Bearer Tokens

- Tokens d'accès
- Service Accounts

OIDC

- OAuth2/OpenID Connect
- Azure AD, Keycloak

Webhook

- Service externe
- Auth custom

RBAC (Role-Based Access Control)

Modèle d'autorisation basé sur des **rôles** assignés à des **sujets** (users, groups, service accounts).

Les 4 objets RBAC

Role : permissions dans un namespace :

ClusterRole : permissions cluster-wide :

RoleBinding : assigner un Role à un sujet :

ClusterRoleBinding : assigner un ClusterRole globalement (même syntaxe, scope cluster).

Verbes RBAC

get (GET) Lire une ressource spécifique
list (GET) Lister les ressources
watch (GET+watch) Surveiller les changements
create (POST) Créer
update (PUT) Mettre à jour
patch (PATCH) Modifier partiellement
delete (DELETE) Supprimer

Service Accounts

Principe du moindre privilège : n'accorder que les permissions strictement nécessaires.

Gestion des accès.

```
apiVersion: rbac.authorization.k8s.io/v1
kind: Role
metadata:
  name: pod-reader
  namespace: default
rules:
- apiGroups: [""]
  resources: ["pods"]
  verbs: ["get", "list", "watch"]
- apiGroups: [""]
 resources: ["services"]
 verbs: ["get", "list", "create"]
```

```
apiVersion:
rbac.authorization.k8s.io/v1
kind: ClusterRole
metadata:
  name: cluster-admin
rules:
- apiGroups: [""]
  resources: ["*"]
  verbs: ["*"]
```

Gestion des accès.

```
apiVersion:
rbac.authorization.k8s.io/v1
kind: RoleBinding
metadata:
  name: read-pods
  namespace: default
subjects:
- kind: User
  name: alice
  apiGroup: rbac.authorization.k8s.io
roleRef:
  kind: Role
  name: pod-reader
  apiGroup: rbac.authorization.k8s.io
```

```
apiVersion:
rbac.authorization.k8s.io/v1
kind: ClusterRoleBinding
metadata:
  name: admin-binding
subjects:
- kind: User
  name: bob
  apiGroup: rbac.authorization.k8s.io
roleRef:
  kind: ClusterRole
  name: cluster-admin
  apiGroup: rbac.authorization.k8s.io
```

paul.millet@lameduse.fr

LaMeDuSe

UBE : Kubernetes, mise en œuvre

replicas et deployment.

- Objectif de la réplication :
  - Assurer la disponibilité, la tolérance aux pannes et la scalabilité des applications en répartissant les charges sur plusieurs instances (pods) au sein du cluster.
- Rôles des Replicas et Déploiements :
  - Replicas : Instances identiques d'un pod pour équilibrer la charge et assurer la disponibilité.
  - Déploiements : Gèrent les replicas, permettent les mises à jour continues, le scaling et les rollbacks.
  - ReplicasSet: Objet piloté par le déploiement, un replicasSet = 1 version

Scalabilité d'un déploiement. (yaml)

```yaml
apiVersion: apps/v1
kind: Deployment
metadata:
  name: my-app
spec:
  replicas: 5 #modifier le nombre de réplicas
  selector:
    matchLabels:
      app: my-app
  template:
    metadata:
      labels:
        app: my-app
    spec:
      containers:
      - name: my-app-container
        image: my-app-image:v1
```

Scalabilité d'un déploiement. (commande)

Sinon par commande :

- “kubectl scale deployment my-app --replicas=5"

Les fichiers descriptifs - Déploiement

```
apiVersion: apps/v1
kind: Deployment
metadata:
  name: my-app
spec:
  replicas: 1
  selector:
    matchLabels:
      app: my-app
  template:
    metadata:
      labels:
        app: my-app
    spec:
      containers:
      - name: my-app-container
        image: nginx
        ports:
        - containerPort: 8080
        env:
        - name: NAME
          value: "value"
```

2.4-full

Stratégie de mise à jour sans interruption (update/rollback).

- Mise à jour sans interruption :
  - Kubernetes permet de déployer des mises à jour d'applications sans interrompre les services.
  - Cela est réalisé à l'aide de déploiements gérés par des stratégies de mise à jour, garantissant que les utilisateurs continuent d'accéder aux applications pendant les modifications.

Stratégie de mise à jour sans interruption (update/rollback).

- Rolling Update : Cette stratégie met à jour progressivement les pods avec une nouvelle version de l'application, en remplaçant les anciens pods par de nouveaux, un ou plusieurs à la fois.
  - Avantages : Pas de temps d'arrêt, possibilité de surveiller chaque étape.
  - Fonctionnement :
    - Un nouveau pod est créé avec la mise à jour.
    - Un ancien pod est supprimé.
    - Le processus continue jusqu'à ce que tous les pods soient remplacés.

Stratégie de mise à jour sans interruption (update/rollback).

- Rollback : Si une mise à jour cause des problèmes, Kubernetes permet de revenir rapidement à une version antérieure du déploiement.
- Pourquoi faire un rollback ? : Corriger une erreur, éviter un downtime prolongé.
- Fonctionnement :
  - Kubernetes garde un historique des déploiements. Un rollback remet la configuration à une version antérieure.

Stratégie de mise à jour sans interruption (update/rollback).

- Déclencher un rollback :
  - Revenir à la version précédente :
    - kubectl rollout undo deployment my-app
  - Rollout vers une version spécifique :
    - Voir les révisions disponibles :
      - kubectl rollout history deployment my-app
    - Rollback vers une révision spécifique :
      - kubectl rollout undo deployment my-app --to-revision=<numéro_de_revision>
    - Vérifier l'état après rollback :
      - kubectl get deployment my-app
  -

Configurer la miise a jour en roulement

```
strategy:
  type: RollingUpdate
  rollingUpdate:
    maxSurge: 50%        # Pods supplémentaires pendant la mise à jour
    maxUnavailable: 1  # Pods indisponibles pendant la mise à jour
```

Daemons set.

- Qu'est-ce qu'un DaemonSet ?
  - Un DaemonSet assure qu'une copie d'un pod particulier tourne sur chaque nœud (ou un sous-ensemble de nœuds) du cluster.
  - Utilisé pour déployer des tâches de maintenance ou des services critiques, comme des agents de monitoring ou des collecteurs de logs.
- Utilisations Courantes :
  - Monitoring (ex. : Prometheus Node Exporter).
  - Collecte de logs (ex. : Alloy, Fluentd, Filebeat).
  - Gestion des ressources (ex. : Daemons de stockage).

DaemonSet

```
apiVersion: apps/v1
kind: DaemonSet
metadata:
 name: node-exporter
 namespace: monitoring
spec:
 selector:
  matchLabels:
   app: node-exporter
 template:
  metadata:
   labels:
    app: node-exporter
  spec:
   containers:
   - name: node-exporter
    image: prom/node-exporter:latest
```

StatefulSet

Pour les applications **stateful** qui nécessitent :
- Des identités stables (pod-0, pod-1, pod-2)
- Un stockage persistant individuel
- Un ordre de démarrage/arrêt garanti

Cas d'usage : bases de données (MySQL, PostgreSQL, MongoDB), caches distribués (Redis Cluster), messagerie (Kafka).

StatefulSet

```
apiVersion: apps/v1
kind: StatefulSet
metadata:
 name: postgres
spec:
 serviceName: postgres-headless  # Service headless requis
 replicas: 3
 selector:
  matchLabels:
   app: postgres
 template:
  metadata:
   labels:
    app: postgres
  spec:
   containers:
   - name: postgres
     image: postgres:15
```

paul.millet@lameduse.fr

LaMeDuSe

UBE : Kubernetes, mise en œuvre

Types de services.

- ClusterIP (par défaut)
  - Usage : Expose le service à l'intérieur du cluster, accessible uniquement par d'autres services ou pods.
  - Utilisation courante : Communication interne entre les microservices.

```
apiVersion: v1
kind: Service
metadata:
  name: my-app-clusterip
spec:
  type: ClusterIP
  selector:
    app: my-app
  ports:
    - port: 80
      targetPort: 8080
```

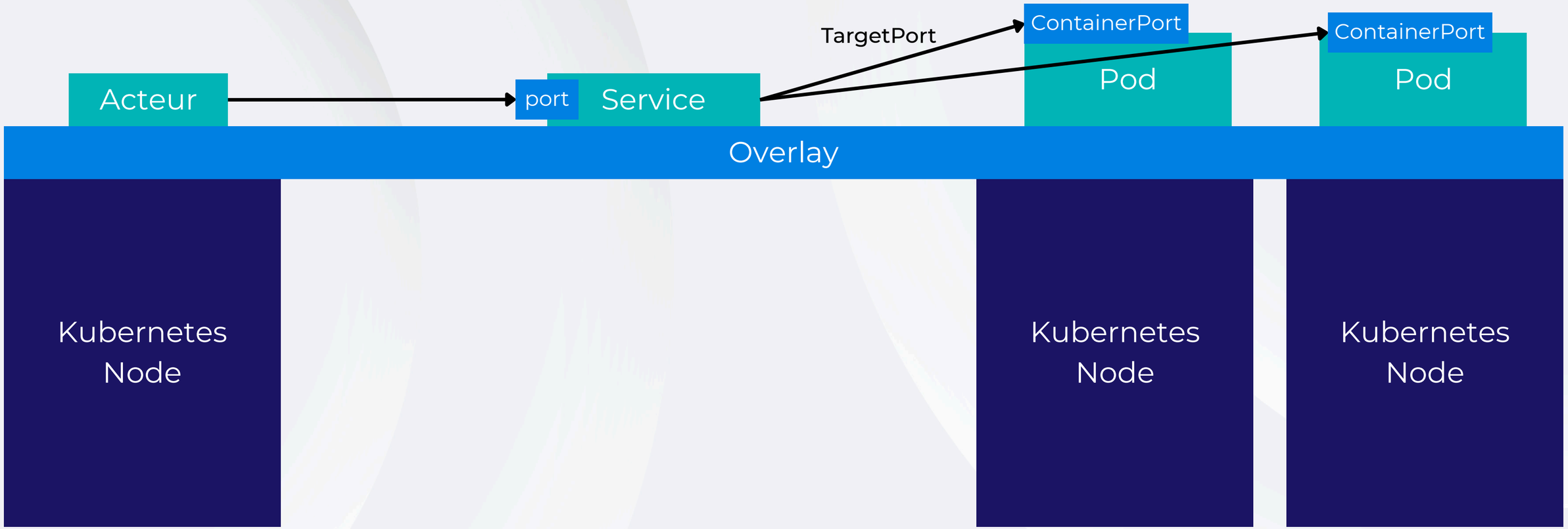
Service = configuration de la couche réseau

Acteur = initiateur de connexion

TargetPort doit être égal a ContainerPort

Types de services.

- ClusterIP (par défaut)



Types de services.

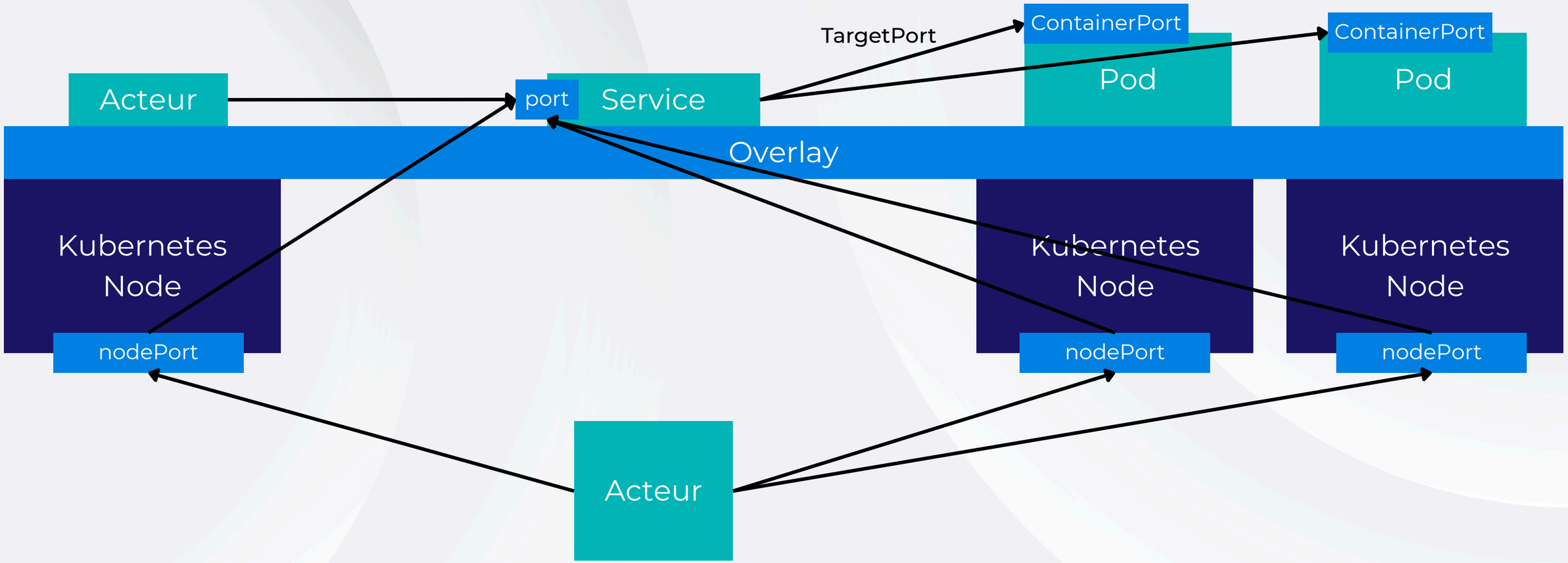
- NodePort
  - Expose le service sur une IP de chaque nœud du cluster, accessible via un port statique (port du nœud).
  - Exposer des services à l'extérieur du cluster pour le développement ou des tests.

```
apiVersion: v1
kind: Service
metadata:
  name: my-app-nodeport
spec:
  type: NodePort
  selector:
    app: my-app
  ports:
    - port: 80
      targetPort: 8080
      nodePort: 30007  # Port entre 30000 et 32767
```

nodePort écouté par le kubeproxy
choisi par vous (si défini) ou par kubernetes sinon

Types de services.

- NodePort



Types de services.

- LoadBalancer
  - Service qui as pour vocation d'être utiliser avec un composant externe pour rentrer du traffic (une gateway)

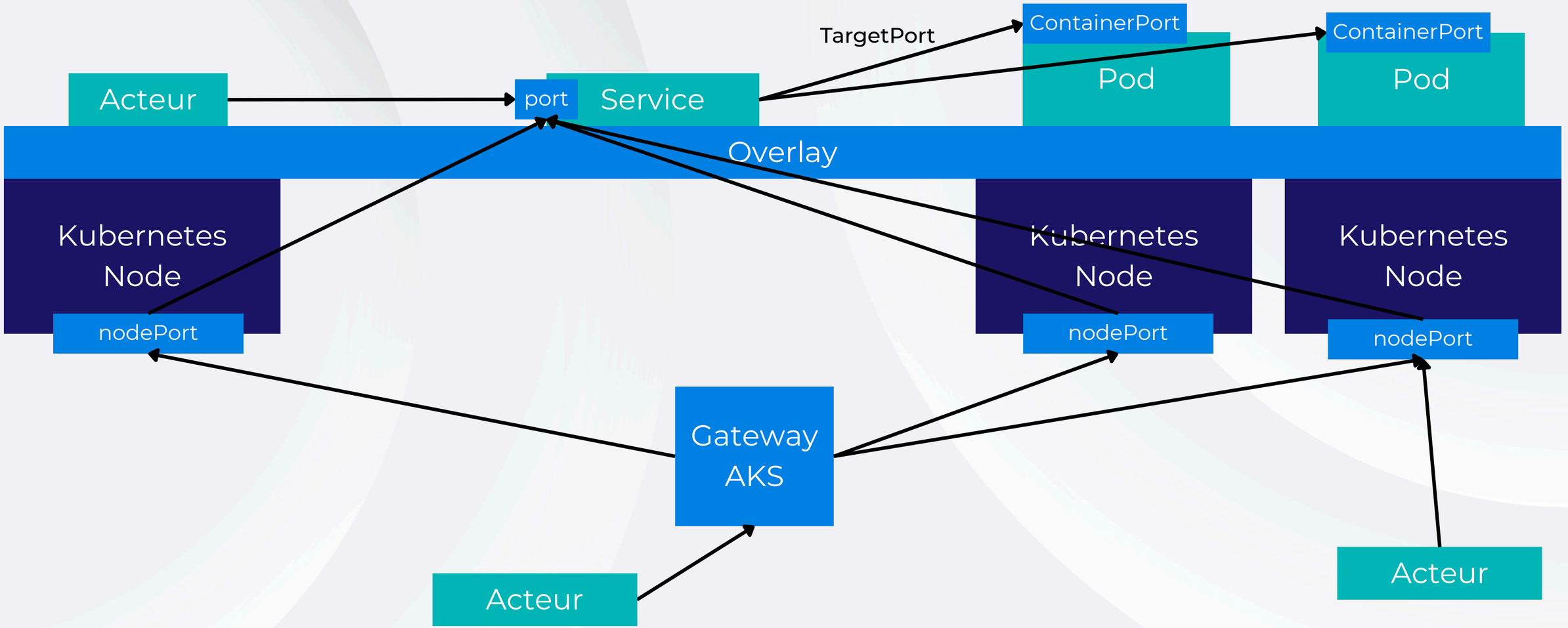
```
apiVersion: v1
kind: Service
metadata:
  name: hello-service
spec:
  selector:
    app: hello
  type: LoadBalancer
  ports:
    - protocol: TCP
      port: 80       # exposed port
      targetPort: 5678  # container port
```

Attention : Service non fonctionnel sans implémentation externe

Le “service provisioner” est alerté qu’un service loadbalancer a été créé et envoie un signal à AKS

Types de services.

- LoadBalancer



Types de services.

- ExternalName
  - Mappe un service à un nom DNS externe, redirige les requêtes vers des services en dehors du cluster.
  - Accéder à des services externes (bases de données, API externes).

```
apiVersion: v1
kind: Service
metadata:
  name: my-app-externalname
spec:
  type: ExternalName
  externalName: my.external.service.com
```

Types de services.

- Headless (ClusterIP sans IP)
  - Un service headless est un service Kubernetes qui n’a pas d’adresse IP virtuelle (clusterIP: None). Contrairement à un service classique, il ne fait pas de load balancing ni de proxy entre les clients et les Pods.

Usage : découverte de pods, clusterisation des pods (statefulset), communication directe.

```
apiVersion: v1
kind: Service
metadata:
  name: my-headless-service
spec:
  clusterIP: None  # <--- devient un svc headless
  selector:
    app: my-app
  ports:
    - name: http
      port: 80
      targetPort: 8080
```

Types de services.

- Quand utiliser ClusterIP ?
  - Pour les communications internes entre microservices au sein du cluster.
- Quand utiliser NodePort ?
  - Pour des accès externes simples ou temporaires, développement, et tests.
- Quand utiliser LoadBalancer ?
  - Applications en production nécessitant une haute disponibilité et un accès depuis l'extérieur du cluster via un IP public.
- Quand utiliser ExternalName ?
  - Pour rediriger les requêtes à des services externes qui ne sont pas gérés par Kubernetes.

Types de services.

- Quand utiliser un service Headless ?
  - Pour la découverte des pods, la clusterisation, une communication directe avec ceux-ci

Service Discovery (DNS).

- Qu'est-ce que la Découverte de Services ?
  - La découverte de services permet aux applications de localiser et communiquer avec d'autres services dans un cluster Kubernetes sans avoir à connaître les adresses IP spécifiques.
- Kubernetes fournit un service DNS interne pour résoudre les noms des services en adresses IP.

Service Discovery (DNS).

- Service DNS Kubernetes :
  - Kubernetes fournit un service DNS interne (coredns) qui résout les noms de services en adresses IP.
  - Les noms de service sont résolus au format
    - service-name.namespace.svc.cluster.local.
    - service-name (même namespace)
    - service-name.namespace (autre namespace)
- Format de Résolution DNS :
  - Un service nommé my-service dans le namespace default peut être résolu en utilisant :
    - my-service.default.svc.cluster.local
    - Pour un service avec un port spécifique :
      - my-service.default.svc.cluster.local:80

Service Discovery (DNS).

- Service Headless
  - Le DNS renvoie des entrées A pour chaque pod

```
;; QUESTION SECTION:
;my-headless-service.default.svc.cluster.local. IN A

;; ANSWER SECTION:
my-headless-service.default.svc.cluster.local. 30 IN A 10.244.1.12
my-headless-service.default.svc.cluster.local. 30 IN A 10.244.2.45
my-headless-service.default.svc.cluster.local. 30 IN A 10.244.3.78
```

Service Discovery (DNS).

- Pod DNS
  - pod-ipv4-address.my-namespace.pod.cluster.local
  - 172-17-0-3.default.pod.cluster.local

Service Discovery (DNS).

- Service Headless + Statefulset
  - Le DNS renvoie des entrées A pour chaque pod sur le nom du service (monservice.monnamespace.svc.cluster.local)
  - Chaque pod a un nom :
    - (web-0.monservice.monnamespace.svc.cluster.local)

```
curl http://web-0.web-headless.default.svc.cluster.local
curl http://web-1.web-headless.default.svc.cluster.local
curl http://web-2.web-headless.default.svc.cluster.local
```

```
apiVersion: apps/v1
kind: StatefulSet
metadata:
  name: web
spec:
  serviceName: "web-headless"  # lien avec le service headless
  replicas: 3
  selector:
    matchLabels:
      app: web
  template:
    metadata:
      labels:
        app: web
    spec:
      containers:
        - name: nginx
          image: nginx:stable
          ports:
            - containerPort: 80
```

paul.millet@lameduse.fr

LaMeDuSe

UBE : Kubernetes, mise en œuvre

Persistent Volumes et Persistent Volumes Claim.

- Qu'est-ce qu'un Persistent Volume (PV) ?
  - Un PV est une ressource de stockage dans Kubernetes qui abstrait les détails de l'infrastructure de stockage (NFS, iSCSI, cloud storage, etc.).
  - Il est provisionné par un administrateur ou dynamiquement par Kubernetes.
- Caractéristiques :
  - Indépendant du cycle de vie des pods : les données persistent même après la suppression des pods.
  - Défini par des capacités (taille, type d'accès).

Persistent Volumes et Persistent Volumes Claim.

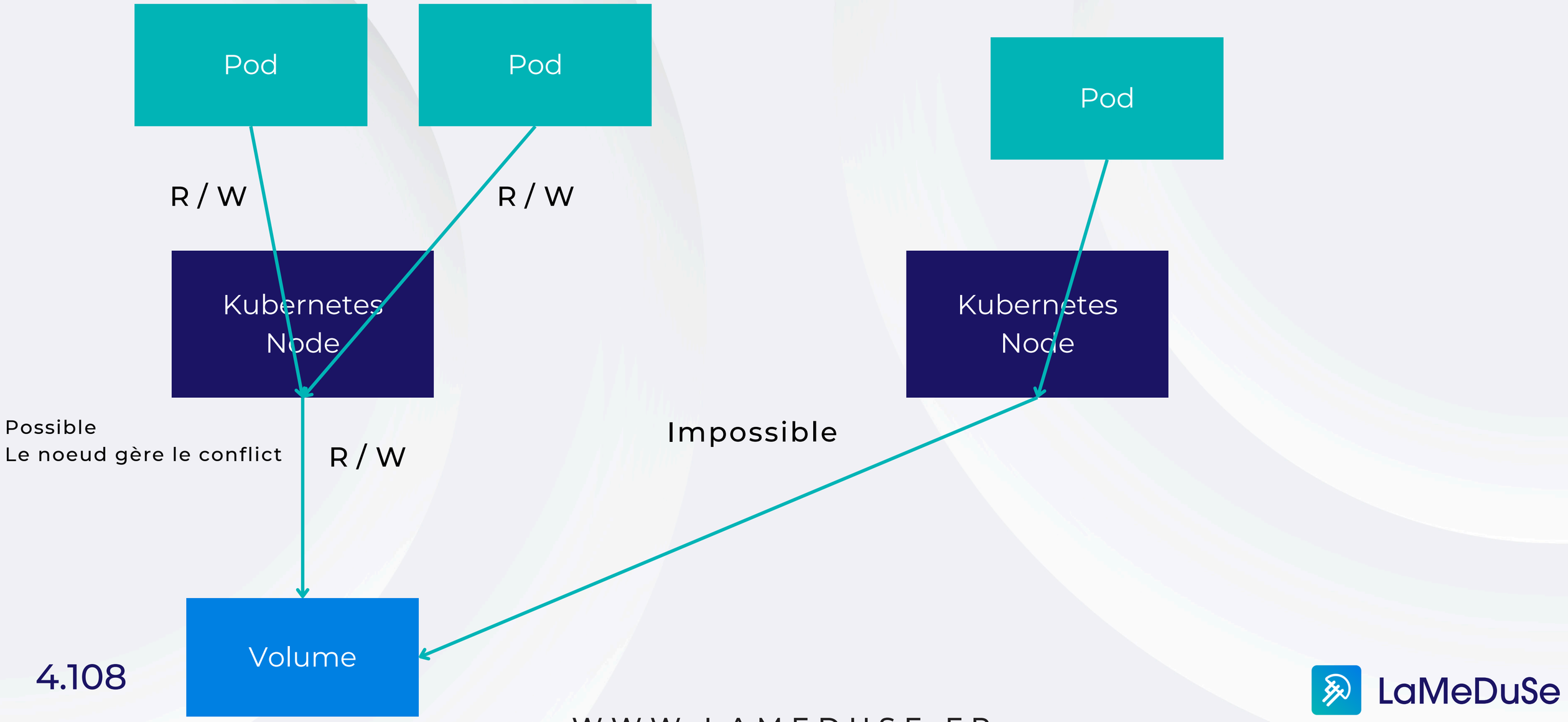
- Types d'accès :
  - ReadWriteOnce (RWO) : Peut être monté en lecture/écriture par un seul nœud.
  - ReadOnlyMany (ROX) : Peut être monté en lecture seule par plusieurs nœuds.
  - ReadWriteMany (RWX) : Peut être monté en lecture/écriture par plusieurs nœuds.
  - ReadWriteOncePod (RWOP) : Peut être monté en lecture/écriture par un seul pod.
- **NB : C’est la couche de stockage qui traite le conflict**

Mode : RWO

Seul un noeud a accès au volume

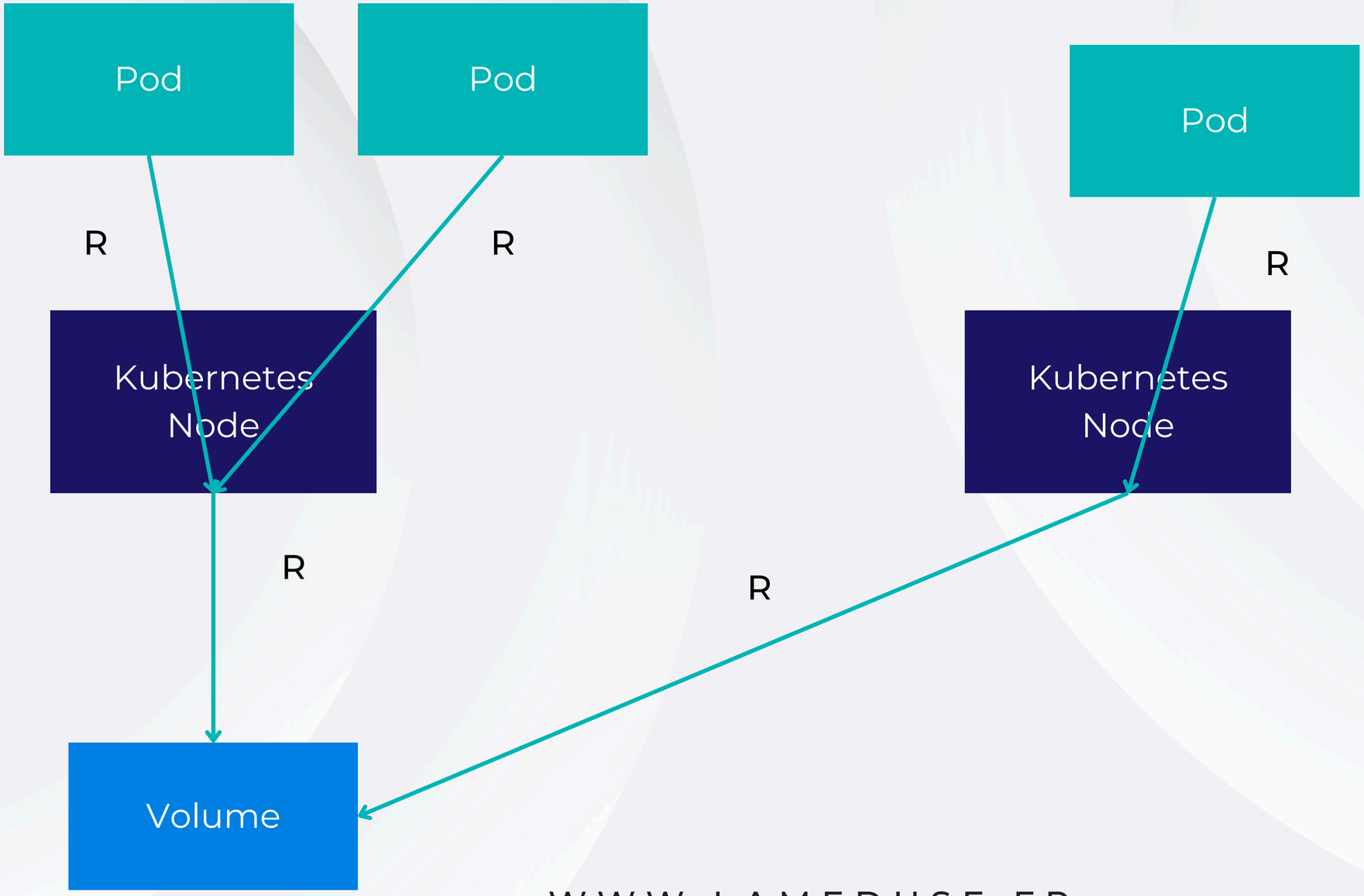
Mode : RWO

Seul un noeud a accès au volume



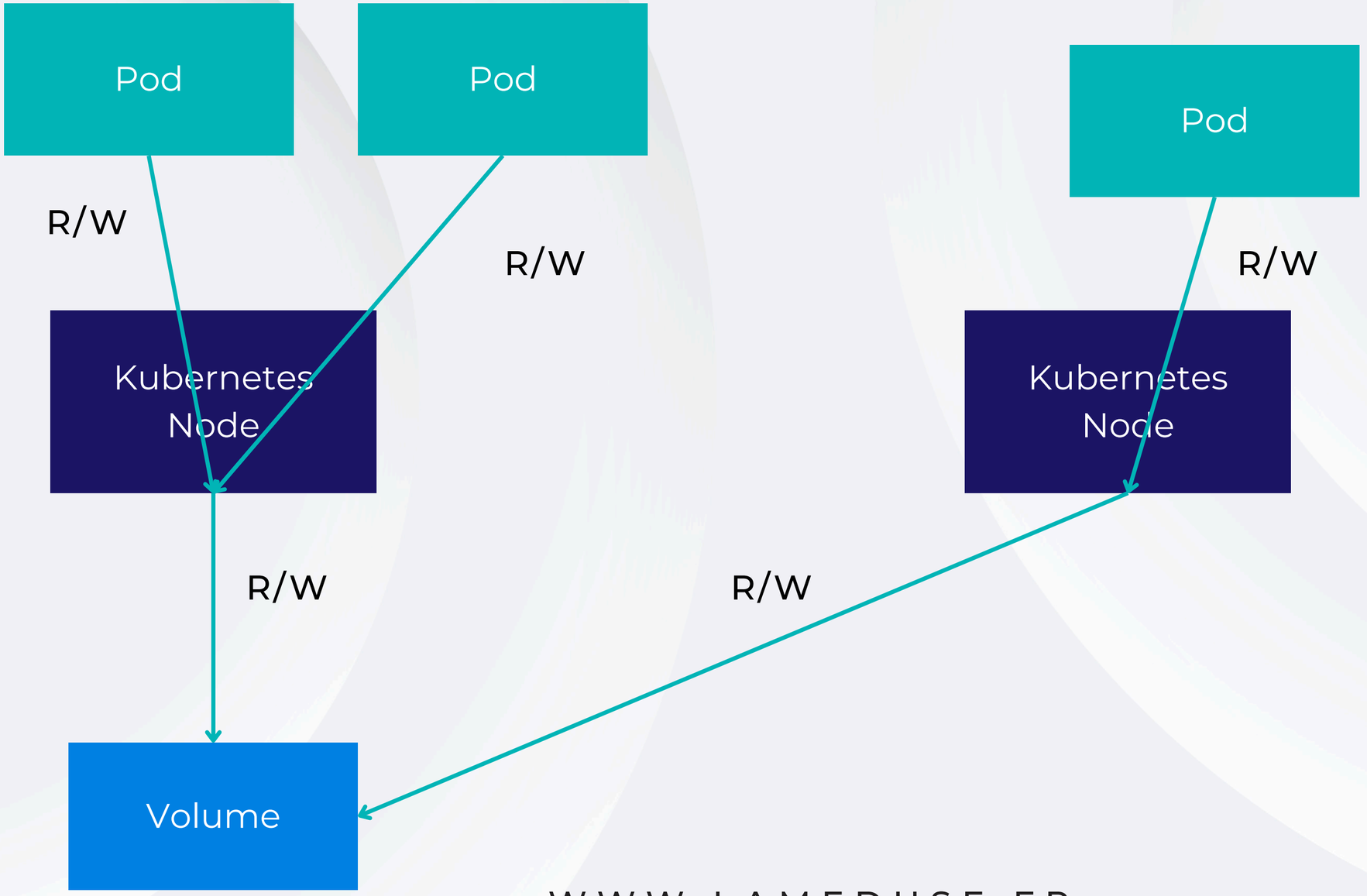
Mode : ROX

Lecture sur plusieurs noeud



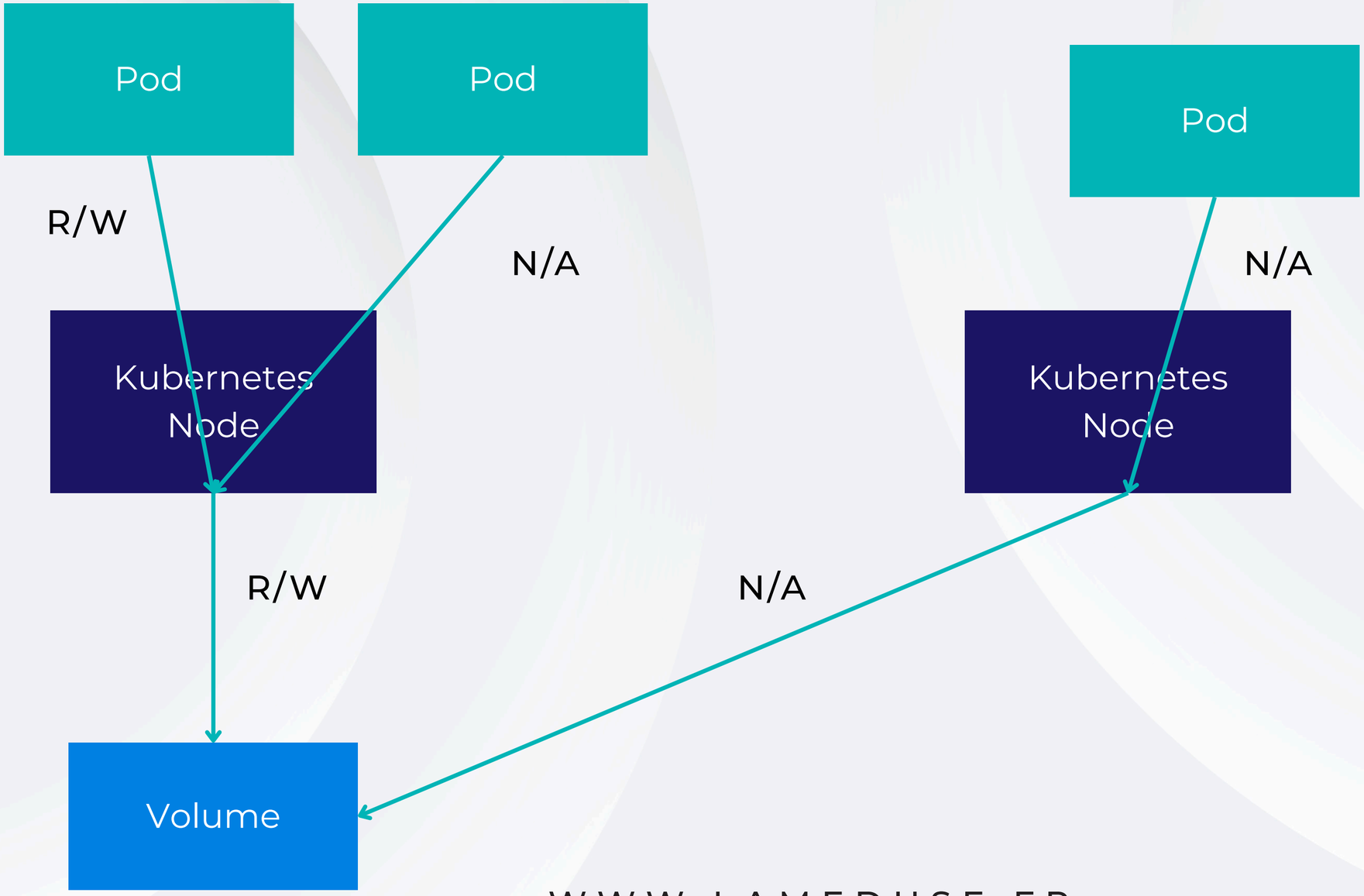
Lecture et écriture multi-noeud

Mode : RWX



Mode : RWOP

Seul un seul pod a accès au volume



Persistent Volumes et Persistent Volumes Claim.

```
apiVersion: v1
kind: PersistentVolume
metadata:
  name: pv-example
spec:
  capacity:
    storage: 10Gi
  accessModes:
    - ReadWriteOnce
  persistentVolumeReclaimPolicy: Retain # Retain = on conserve / Recycle = rm -rf / Delete = suppression
  hostPath:
    path: /mnt/data
```

Persistent Volumes et Persistent Volumes Claim.

- Qu'est-ce qu'un Persistent Volume Claim (PVC) ?
  - Un PVC est une requête d'un utilisateur pour un PV avec des spécifications spécifiques (taille, modes d'accès).
  - Kubernetes associe un PVC à un PV approprié qui répond aux exigences de la demande.
- Fonctionnement :
  - Les utilisateurs créent des PVCs, et Kubernetes les associe automatiquement à des PVs disponibles.
  - PVCs permettent aux utilisateurs de demander du stockage sans avoir à connaître les détails de l'infrastructure sous-jacente.

Persistent Volumes et Persistent Volumes Claim.

```
apiVersion: v1
kind: PersistentVolumeClaim
metadata:
  name: pvc-example
spec:
  storageClass: ceph
  accessModes:
    - ReadWriteOnce
  resources:
    requests:
      storage: 5Gi
```

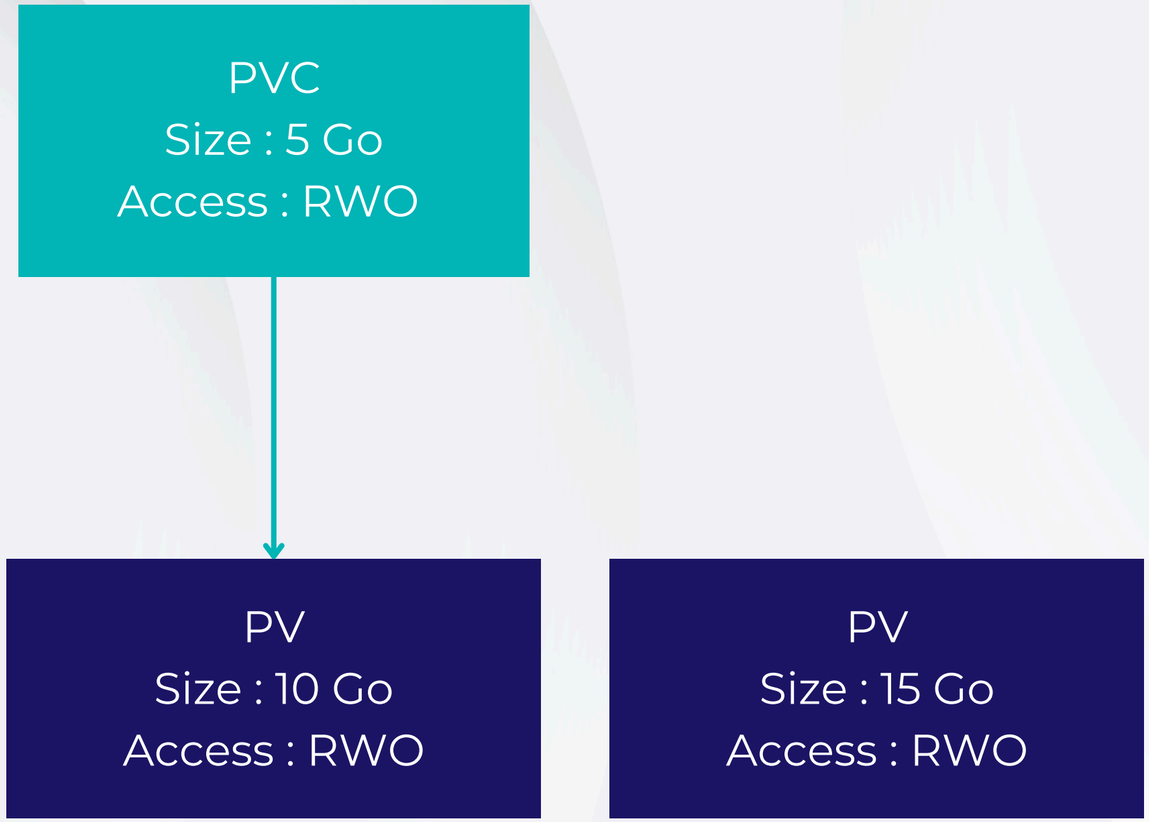
Association PV / PVC (one - one)
Provisionnement Manuel

PVC
Size : 5 Go
Access : RWO

PV
Size : 10 Go
Access : RWO

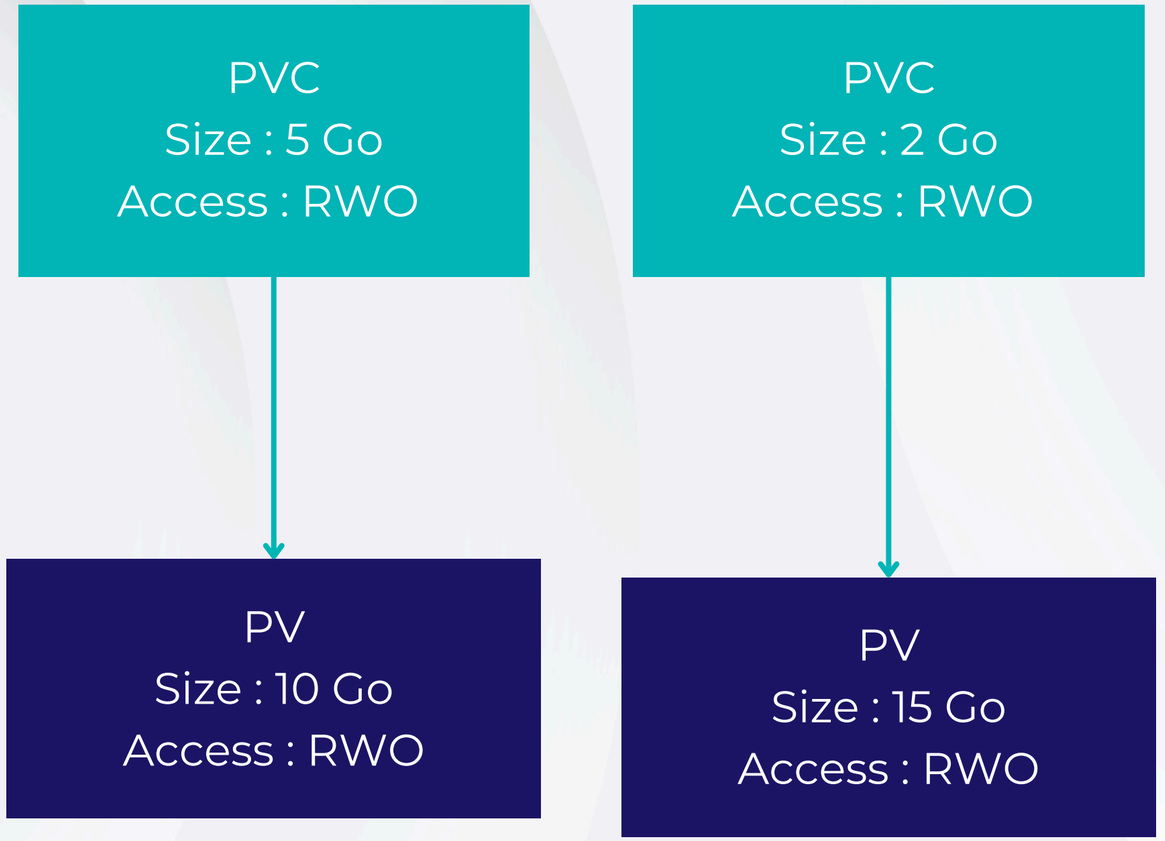
PV
Size : 15 Go
Access : RWO

Association PV / PVC (one - one)
Provisionnement Manuel



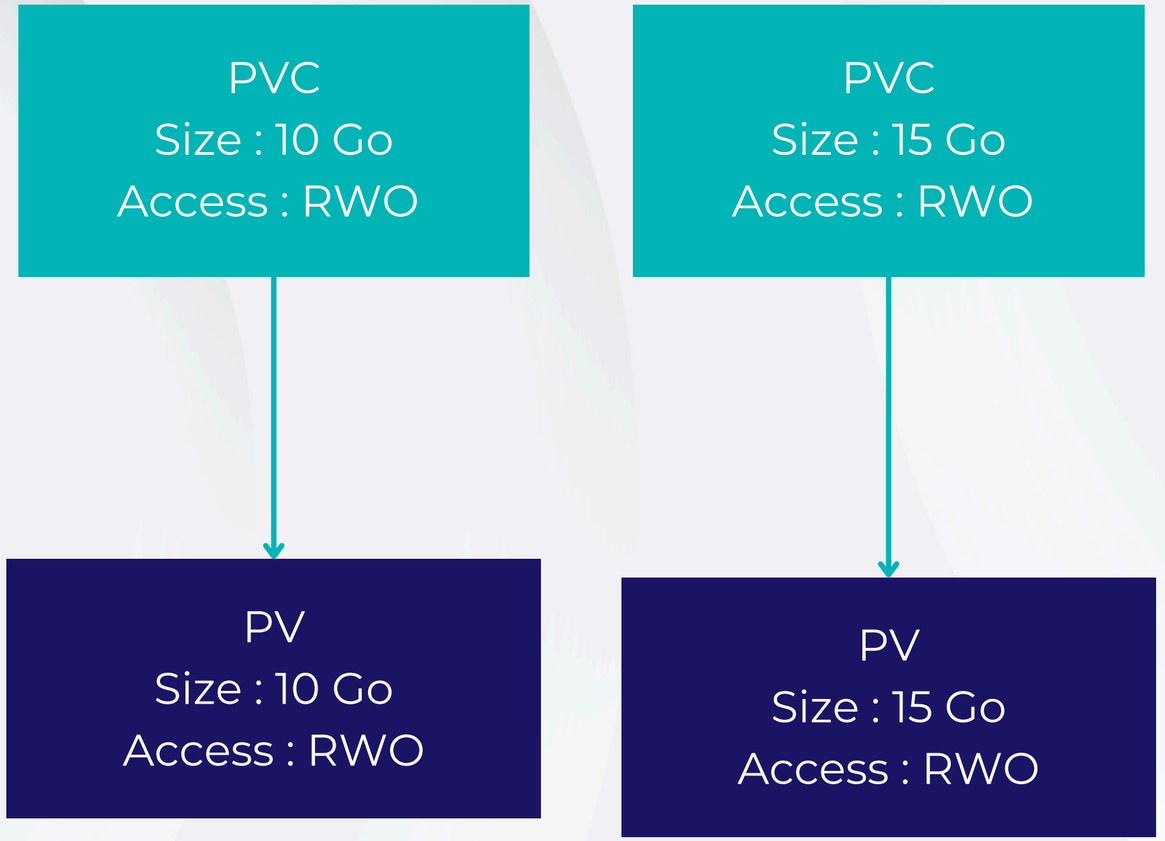
Association PV / PVC (one - one)
Provisionnement Manuel

Association PV / PVC (one - one)
Provisionnement Manuel



Attention : Relation PVC / PV est unique
Un seul PVC pour un PV et inversement

Association PV / PVC (one - one)
Provisionnement Manuel



Attention : Relation PVC / PV est unique
Un seul PVC pour un PV et inversement
redimensionne le PVC pour correspondre a la taille du PV

Association PV / PVC (one - one)
Provisionnement Manuel

PVC
Size : 5 Go
Access : RWO
StorageClass : ceph

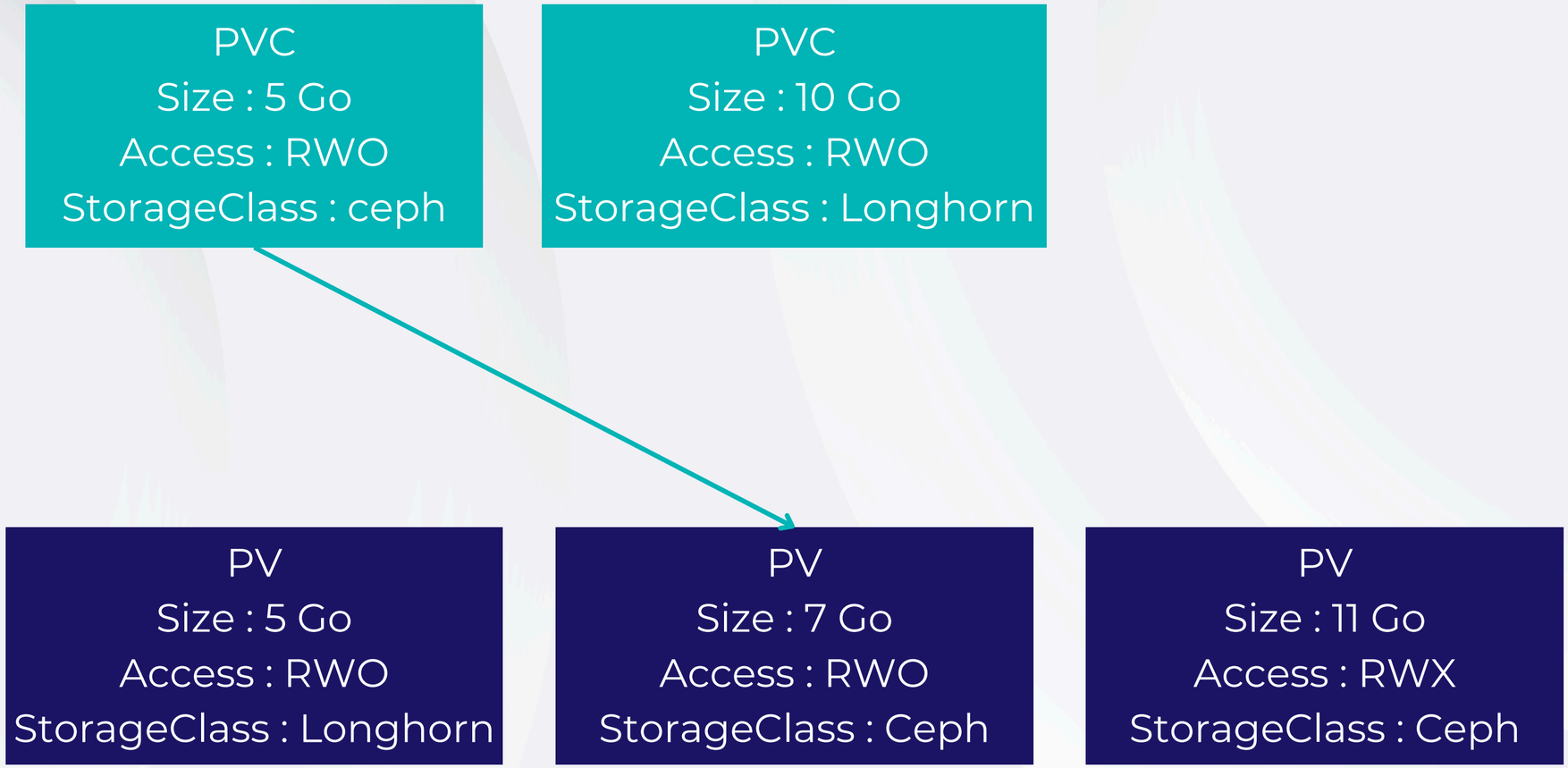
PVC
Size : 10 Go
Access : RWO
StorageClass : Longhorn

PV
Size : 5 Go
Access : RWO
StorageClass : Longhorn

PV
Size : 7 Go
Access : RWO
StorageClass : Ceph

PV
Size : 11 Go
Access : RWX
StorageClass : Ceph

Association PV / PVC (one - one)
Provisionnement Manuel



Association PV / PVC (one - one)
Provisionnement Manuel

PVC
Size : 5 Go
Access : RWO
StorageClass : ceph

PVC
Size : 10 Go
Access : RWO
StorageClass : Longhorn

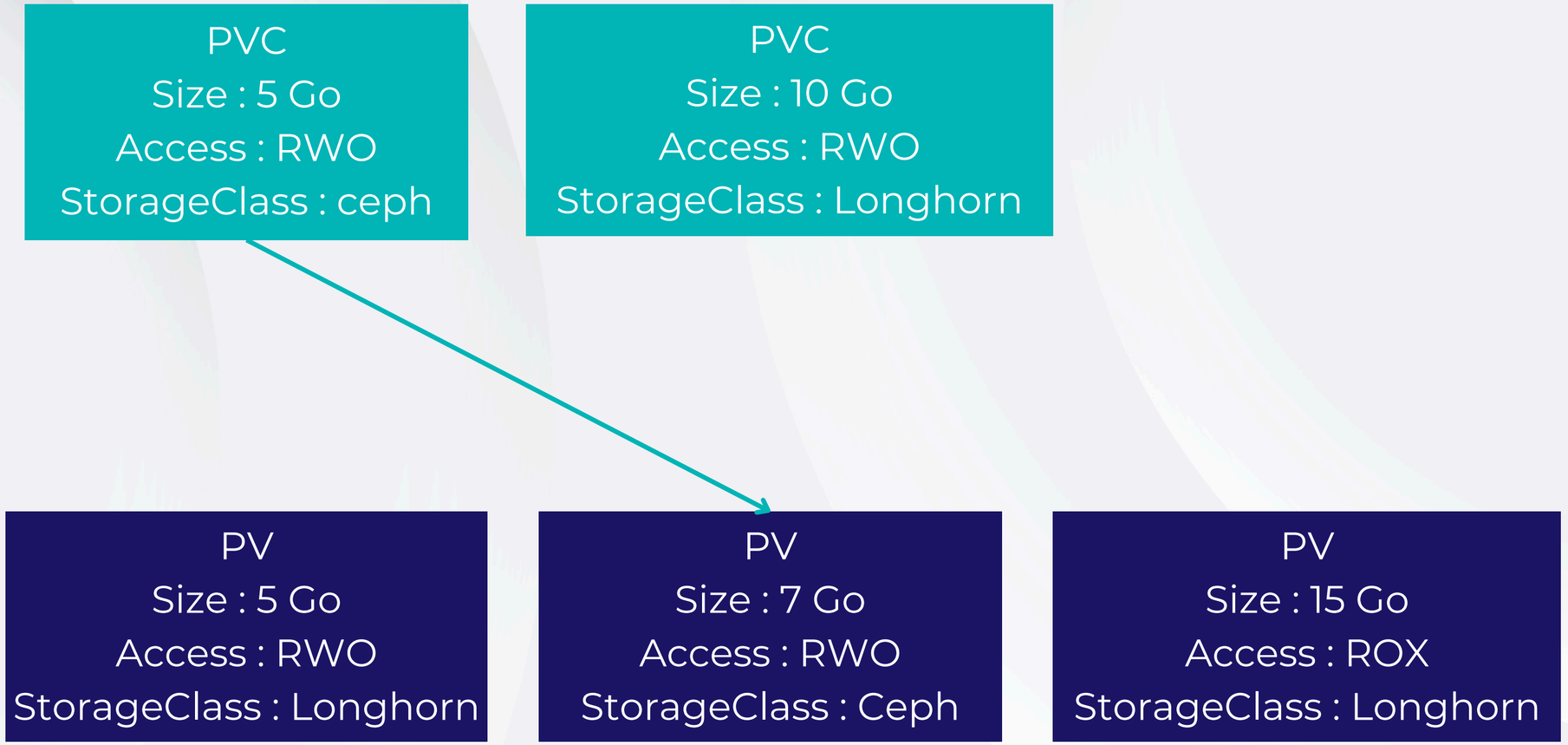
Taille non respecté donc aucune association

PV
Size : 5 Go
Access : RWO
StorageClass : Longhorn

PV
Size : 7 Go
Access : RWO
StorageClass : Ceph

PV
Size : 11 Go
Access : RWX
StorageClass : Ceph

Association PV / PVC (one - one)
Provisionnement Manuel



Association PV / PVC (one - one)
Provisionnement Manuel

PVC
Size : 5 Go
Access : RWO
StorageClass : ceph

PVC
Size : 10 Go
Access : RWO
StorageClass : Longhorn

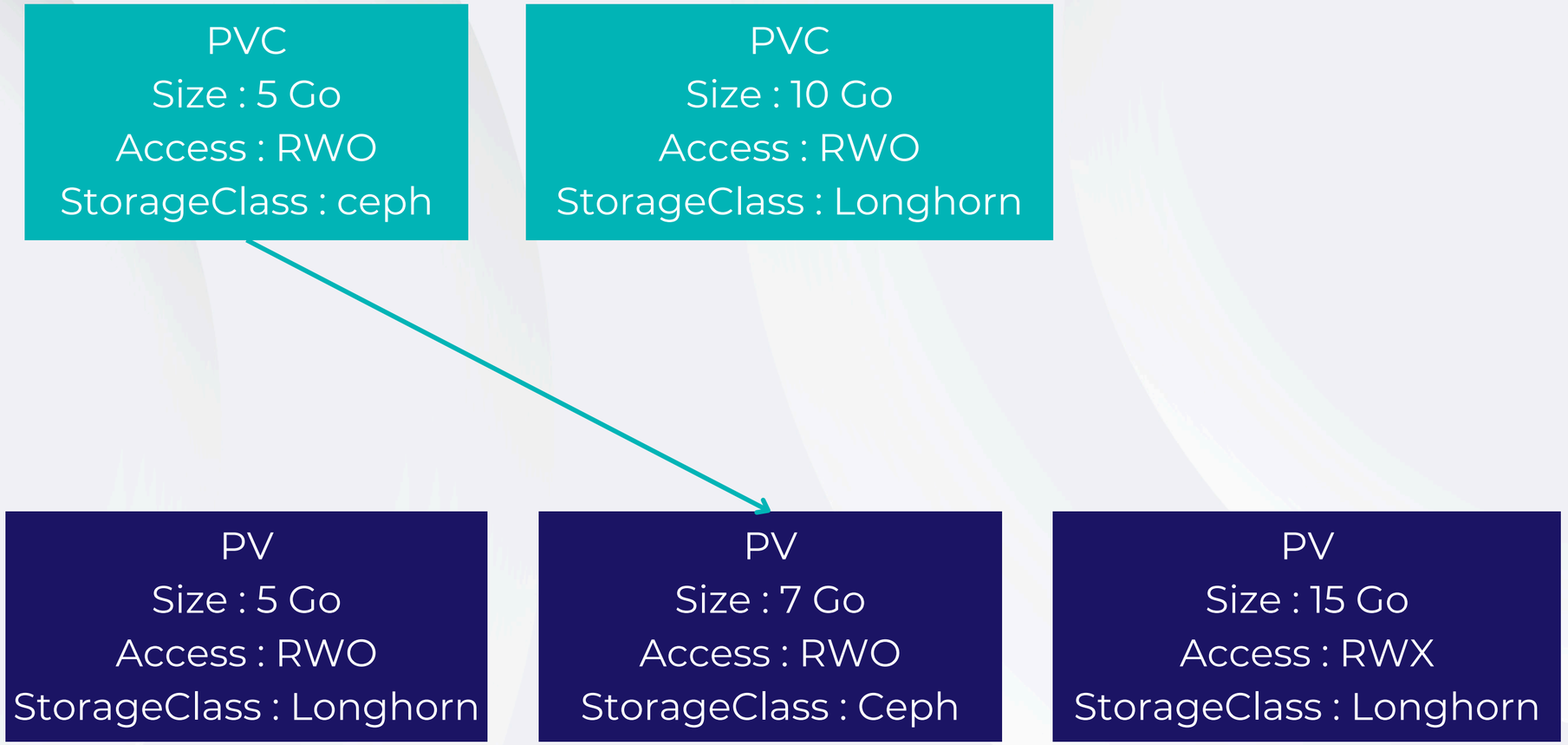
Mode non respecté donc aucune association

PV
Size : 5 Go
Access : RWO
StorageClass : Longhorn

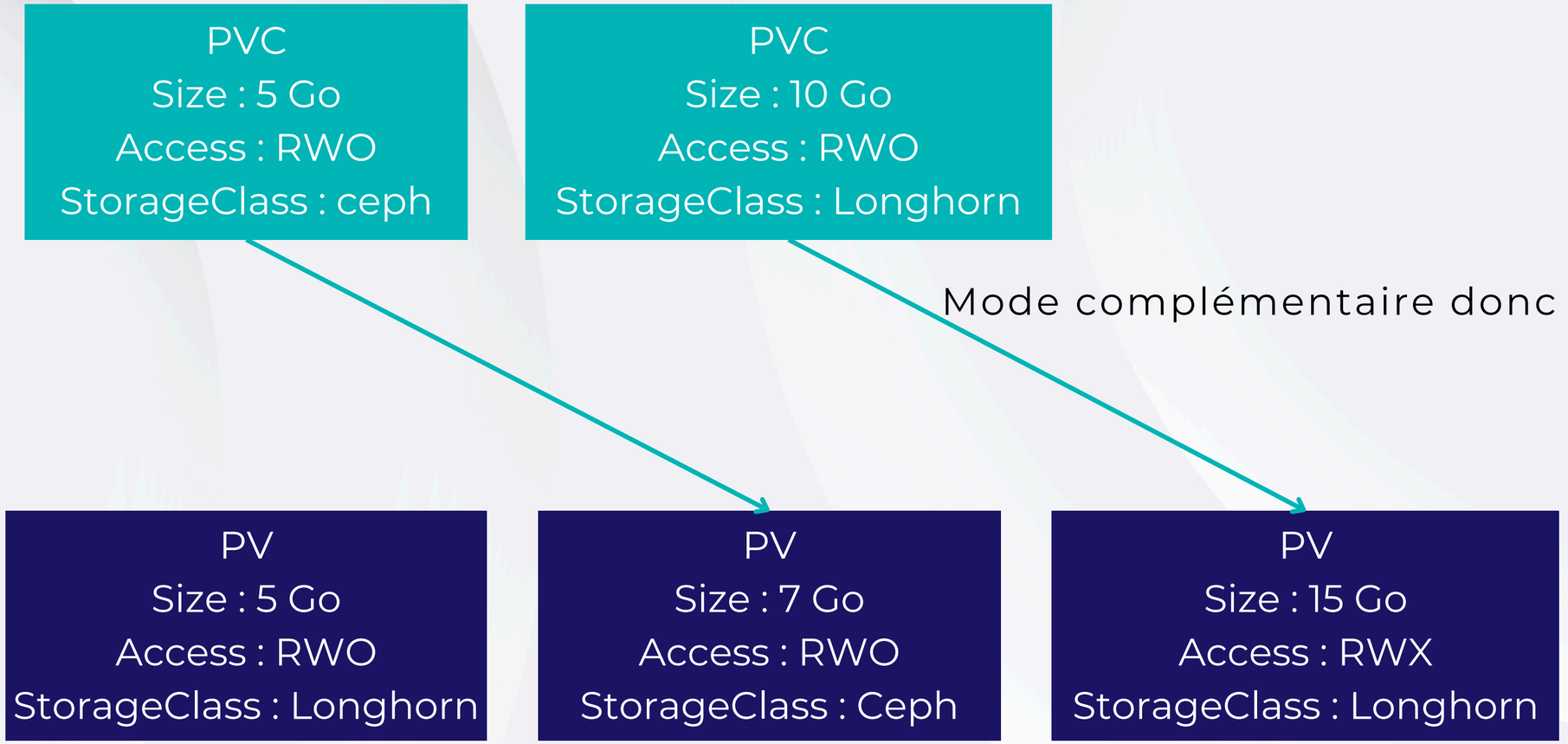
PV
Size : 7 Go
Access : RWO
StorageClass : Ceph

PV
Size : 15 Go
Access : ROX
StorageClass : Longhorn

Association PV / PVC (one - one)
Provisionnement Manuel



Association PV / PVC (one - one)
Provisionnement Manuel



Association PV / PVC (one - one)

PVC / PV	RWOP	RWO	ROX	RWX
RWOP				
RWO				
ROX				
RWX				

Le PV doit avoir les modes demandé par le PVC

Les modes sont cumulable

Persistent Volumes et Persistent Volumes Claim.

```
apiVersion: v1
kind: Pod
metadata:
  name: task-pv-pod
spec:
  volumes:
    - name: task-pv-storage
      persistentVolumeClaim:
        claimName: pvc-example
  containers:
    - name: task-pv-container
      image: nginx
      ports:
        - containerPort: 80
          name: "http-server"
      volumeMounts:
        - mountPath: "/usr/share/nginx/html"
          name: task-pv-storage
```

```
apiVersion: apps/v1
kind: StatefulSet
metadata:
  name: web
spec:
  replicas: 3
  selector:
    matchLabels:
      app: web
  template:
    metadata:
      labels:
        app: web
    spec:
      containers:
        - name: nginx
          image: nginx:stable
          ports:
            - containerPort: 80
          volumeMounts:
            - name: www
              mountPath: /usr/share/nginx/html
  volumeClaimTemplates:
    - metadata:
        name: www
      spec:
        accessModes: [ "ReadWriteOnce" ]
        resources:
          requests:
            storage: 1Gi
```

Config map et secrets.

- Qu'est-ce qu'un ConfigMap ?
  - Un ConfigMap permet de stocker des configurations non sensibles sous forme de paires clé-valeur.
  - Permet de décorréler les paramètres de configuration du code de l'application.
- Utilisations Courantes :
  - Fournir des fichiers de configuration, des variables d'environnement.
  - Injecter des paramètres de configuration dans les pods à l'exécution.

Config map et secrets.

```
apiVersion: v1
kind: ConfigMap
metadata:
  name: my-config
data:
  DATABASE_URL: "postgres://user:password@host:5432/dbname"
  config.json: |
    {
      "setting1": "value1",
      "setting2": "value2"
    }
```

```
spec:
  containers:
  - name: my-app-container
    image: my-app-image:v1
    env:
    - name: DATABASE_URL
      valueFrom:
        configMapKeyRef:
          name: my-config
          key: DATABASE_URL
```

Config map et secrets.

```
apiVersion: v1
kind: Pod
metadata:
  name: env-demo
spec:
  containers:
    - name: busybox
      image: busybox
      command: ["sh", "-c", "env; sleep 3600"]
      envFrom:
        - configMapRef:
            name: my-config
```

Config map et secrets.

```
# pod.yaml
apiVersion: v1
kind: Pod
metadata:
  name: configmap-demo
spec:
  containers:
    - name: demo-container
      image: busybox
      command: [ "sleep", "3600" ]
      volumeMounts:
        - name: config-volume
          mountPath: /etc/config
  volumes:
    - name: config-volume
      configMap:
        name: my-config
```

Config map et secrets.

```
apiVersion: v1
kind: Pod
metadata:
  name: file-mount-demo
spec:
  containers:
    - name: busybox
      image: busybox
      command: ["sleep", "3600"]
      volumeMounts:
        - name: config-volume
          mountPath: /etc/myconfig/database_endpoint # Le point de montage devient un fichier
          subPath: DATABASE_URL # avec le contenu de la clef
  volumes:
    - name: config-volume
      configMap:
        name: my-config
```

Config map et secrets.

```
apiVersion: v1
kind: Pod
metadata:
  name: file-mount-demo
spec:
  containers:
    - name: busybox
      image: busybox
      command: ["sleep", "3600"]
      volumeMounts:
        - name: config-volume
          mountPath: /etc/myconfig/config-file.json # Le point de montage devient un fichier
          # (clef config.json dans config-file.json)
          subPath: config.json # avec le contenu de la clef
  volumes:
    - name: config-volume
      configMap:
        name: my-config
```

Config map et secrets.

```
apiVersion: v1
kind: Pod
metadata:
  name: file-mount-demo
spec:
  containers:
    - name: busybox
      image: busybox
      command: ["sleep", "3600"]
      volumeMounts:
        - name: config-volume
          mountPath: /etc/myconfig
  volumes:
    - name: config-volume
      configMap:
        name: single-file-config
        items: # Monte uniquement les fichiers sélectionné
        - key: config.json
          path: config-file.json # clef “config.json” dans config-file.json
```

Config map et secrets.

- Qu'est-ce qu'un Secret ?
  - Un Secret est similaire à un ConfigMap, mais il est conçu pour stocker des données sensibles, comme des mots de passe, des clés API, ou des certificats.
  - Les données sont encodées en base64 et sont mieux sécurisées en mémoire et en transit que les ConfigMaps.
- Utilisations Courantes :
  - Stocker des mots de passe de base de données, des clés SSH, des tokens d'API.
  - Fournir des informations sensibles à des applications sans les exposer dans le code ou les configurations.

Config map et secrets.

```
apiVersion: v1
kind: Secret
metadata:
  name: my-secret
type: Opaque
data:
  username: dXNlcg==         # base64 for "user"
  password: cGFzc3dvcmQ=     # base64 for "password"
```

```yaml
apiVersion: v1
kind: Pod
metadata:
  name: secret-env-demo
spec:
  containers:
    - name: busybox
      image: busybox
      command: ["sh", "-c", "env && sleep 3600"]
      env:
        - name: SECRET_USERNAME
          valueFrom:
            secretKeyRef:
              name: my-secret
              key: username
        - name: SECRET_PASSWORD
          valueFrom:
            secretKeyRef:
              name: my-secret
              key: password
```

Config map et secrets.

```
apiVersion: v1
kind: Pod
metadata:
  name: secret-envfrom-demo
spec:
  containers:
    - name: app
      image: busybox
      command: ["sh", "-c", "env && sleep 3600"]
      envFrom:
        - secretRef:
            name: my-secret
```

Config map et secrets.

```
apiVersion: v1
kind: Pod
metadata:
  name: secret-volume-demo
spec:
  containers:
    - name: busybox
      image: busybox
      command: ["sleep", "3600"]
      volumeMounts:
        - name: secret-volume
          mountPath: /etc/secret
  volumes:
    - name: secret-volume
      secret:
        secretName: my-secret
```

paul.millet@lameduse.fr

LaMeDuSe

UBE : Kubernetes, mise en œuvre

LaMeDuSe

Frontal administrable Ingress.

- Qu'est-ce qu'un Ingress ?
  - Ingress est un objet Kubernetes qui gère l'accès externe aux services du cluster, généralement via HTTP ou HTTPS.
  - Il offre un moyen de configurer des règles de routage pour le trafic entrant vers les services internes du cluster.
- Pourquoi utiliser Ingress ?
  - Pour exposer plusieurs services sous un seul point d'entrée (ex. : une seule adresse IP ou un seul DNS).
  - Pour gérer des règles de routage complexes, des terminaux TLS, et des équilibrages de charge.
  - Fournit des fonctionnalités de gestion de trafic avancées, comme les redirections et les réécritures d'URL.

Frontal administrable Ingress.

- Ingress Controller :
  - Composant responsable de l'implémentation des règles Ingress.
  - Différents contrôleurs disponibles (NGINX, Traefik, HAProxy, etc.).
  - Nécessaire pour faire fonctionner l'objet Ingress ; sans contrôleur, l'Ingress ne fonctionnera pas.
- Ingress Resource :
  - Définit les règles de routage pour les requêtes HTTP/HTTPS vers les services Kubernetes.
  - Utilise des annotations pour configurer des options spécifiques (ex. : authentification, redirection HTTPS).
- TLS Support : (Cert Manager)
  - Ingress peut gérer les certificats TLS pour sécuriser le trafic entre les clients et les services du cluster.

Frontal administrable Ingress.

```
apiVersion: networking.k8s.io/v1
kind: Ingress
metadata:
  name: my-ingress
  annotations:
    nginx.ingress.kubernetes.io/rewrite-target: /
spec:
  rules:
  - host: example.com
     http:
     paths:
     - path: /app1
       pathType: Prefix
       backend:
         service:
         name: app1-service
         port:
           number: 80
  tls:
  - hosts:
     - example.com
     secretName: tls-secret
```

paul.millet@lameduse.fr

LaMeDuSe

UBE : Kubernetes, mise en œuvre

Labels et choix d'un nœud pour le déploiement.

- Qu'est-ce qu'un Label ?
  - Un label est une paire clé-valeur attribuée à des objets Kubernetes (pods, nœuds, services, etc.).
  - Utilisé pour organiser, sélectionner et filtrer les objets.
- Pourquoi utiliser des Labels ?
  - Pour catégoriser et identifier les objets pour des requêtes spécifiques.
  - Pour la sélection des pods dans les services, les déploiements, et d'autres configurations.
- Exemples de Labels :
  - “env: production”
  - “app: frontend”

Labels

- Sélection de Pods avec Labels :
  - Les services, déploiements, et autres ressources utilisent les labels pour cibler les pods spécifiques.

```
apiVersion: apps/v1
kind: Deployment
metadata:
  name: my-app
spec:
  replicas: 3
  selector:
    matchLabels:
      app: my-app
  template:
    metadata:
      labels:
        app: my-app
    spec:
      containers:
      - name: my-app-container
        image: my-app-image:v1
```

Labels

- Sélection de Pods avec Labels :
  - Les services, déploiements, et autres ressources utilisent les labels pour cibler les pods spécifiques.

```
apiVersion: v1
kind: Service
metadata:
  name: my-app-service
spec:
  selector:
    app: my-app
  ports:
    - port: 80
      targetPort: 8080
```

Choix d'un nœud pour le déploiement.

- Choix d'un Nœud : Pourquoi ?
  - Pour déployer des pods sur des nœuds spécifiques en fonction de leurs caractéristiques (CPU, mémoire, emplacement, etc.).
  - Pour respecter des contraintes de sécurité, de performances, ou d'autres exigences spécifiques.
- Affinité de Nœud (Node Affinity) :
  - Déploye les pods sur les nœuds qui correspondent à certains labels.
- Types d'affinité :
  - Required (obligatoire)
  - Preferred (préférée).

Choix d'un nœud pour le déploiement.

```
spec:
  affinity:
    nodeAffinity:
      requiredDuringSchedulingIgnoredDuringExecution:
        nodeSelectorTerms:
        - matchExpressions:
          - key: disktype
            operator: In
            values:
            - ssd
  containers:
  - name: my-app-container
    image: my-app-image:v1
```

Choix d'un nœud pour le déploiement.

```
spec:
  affinity:
    nodeAffinity:
      preferredDuringSchedulingIgnoredDuringExecution:
      - weight: 1
        preference:
          matchExpressions:
          - key: disktype
            operator: In # In / NotIn / Exist / NotExist
            values:
            - ssd
      - weight: 50
        preference:
          matchExpressions:
          - key: disktype
            operator: In
            values:
            - nvme
  containers:
  - name: nginx
    image: nginx
```

Capacité max = 1pod

Affinity : NodeAffinity
Required
Selection
- disktype = ssd

Pod

Pod

Pod

Kubernetes Node

disktype=ssd

Kubernetes Node

disktype=ssd

disktype=nvme

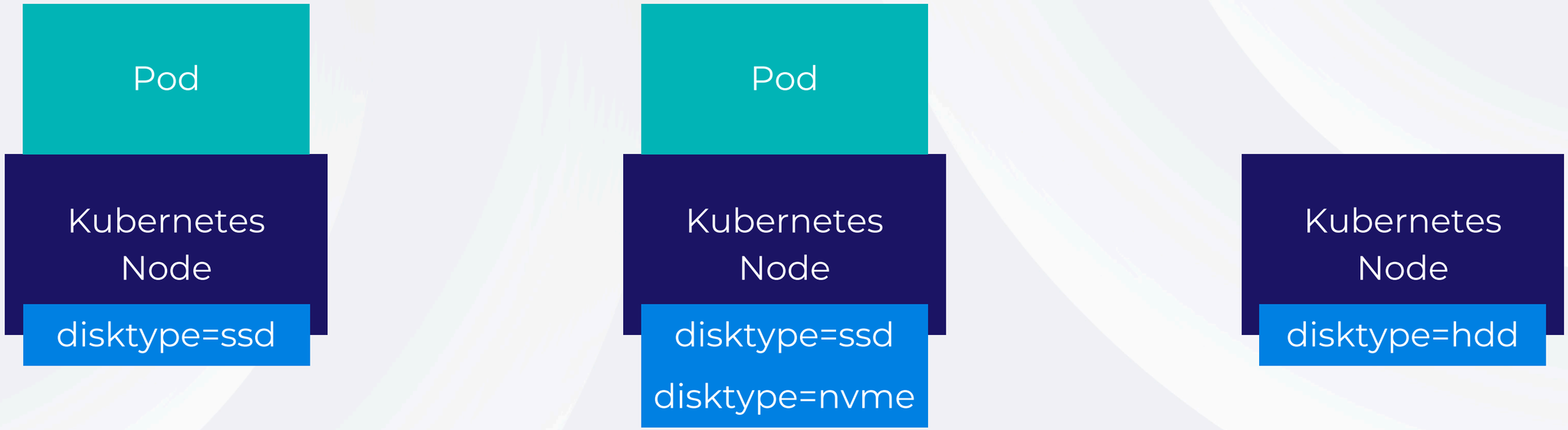
Kubernetes Node

disktype=hdd

Capacité max = 1pod

Affinity : NodeAffinity
Required
Selection
- disktype = ssd

Pod



Capacité max = 1pod

Pod
(ECHEC)

Affinity : NodeAffinity
Required
Selection
- disktype = ssd

Pod

Kubernetes
Node

disktype=ssd

Pod

Kubernetes
Node

disktype=ssd

disktype=nvme

Kubernetes
Node

disktype=hdd

Capacité max = 1pod

Affinity : NodeAffinity
Preferred
Selection

- disktype = ssd = 1
- disktype = nvme = 50

Pod

Pod

Pod

Kubernetes Node

disktype=ssd

Kubernetes Node

disktype=ssd

disktype=nvme

Kubernetes Node

disktype=hdd

Capacité max = 1pod

Affinity : NodeAffinity
Preferred
Selection

- disktype = ssd = 1
- disktype = nvme = 50

Po

Pod

Poids = 1

Kubernetes
Node

disktype=ssd

Poids = 51

Kubernetes
Node

disktype=ssd

disktype=nvme

Poids = 0

Kubernetes
Node

disktype=hdd

Capacité max = 1pod

Affinity : NodeAffinity
Preferred
Selection
- disktype = ssd = 1
- disktype = nvme = 50

Po

Pod

Pod

Poids = 1

Kubernetes
Node

disktype=ssd

Poids = 51

Kubernetes
Node

disktype=ssd

disktype=nvme

Poids = 0

Kubernetes
Node

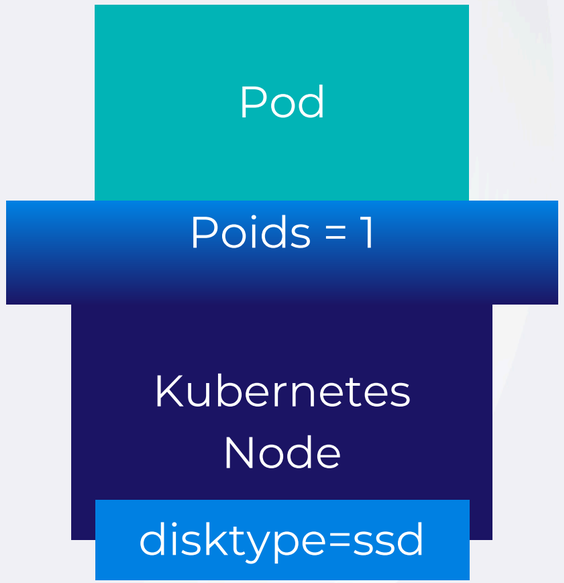
disktype=hdd

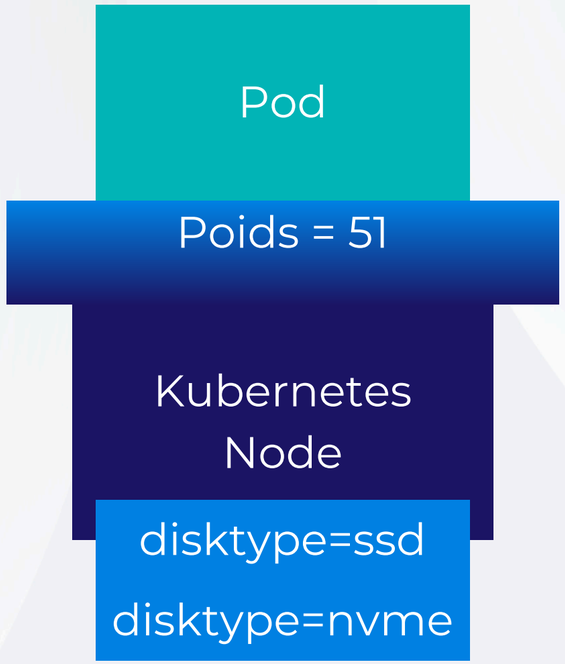
Capacité max = 1pod

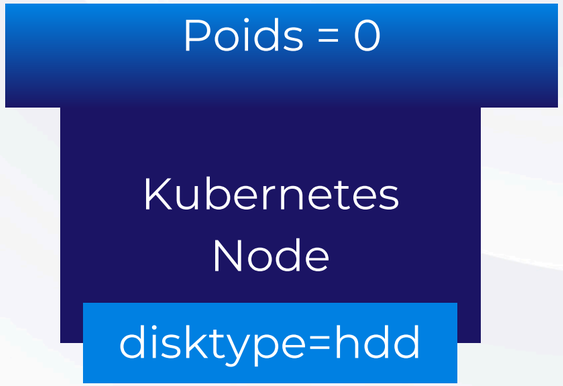
Affinity : NodeAffinity
Preferred
Selection

- disktype = ssd = 1
- disktype = nvme = 50

Pod



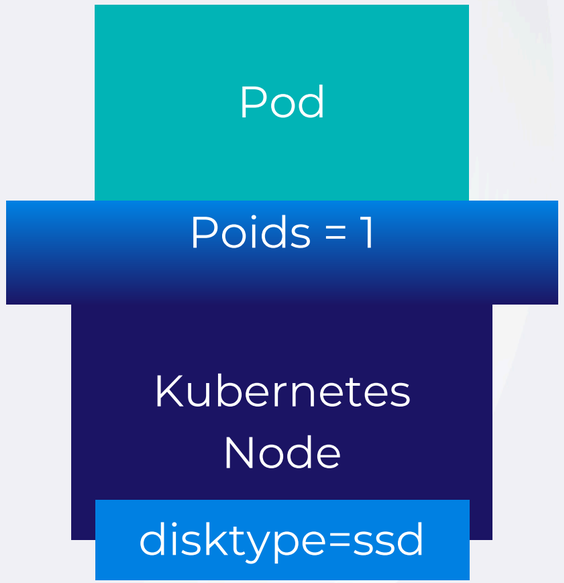


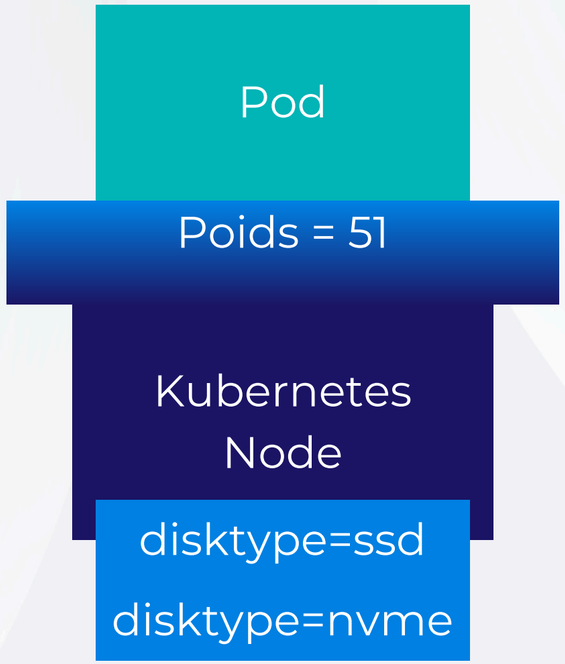


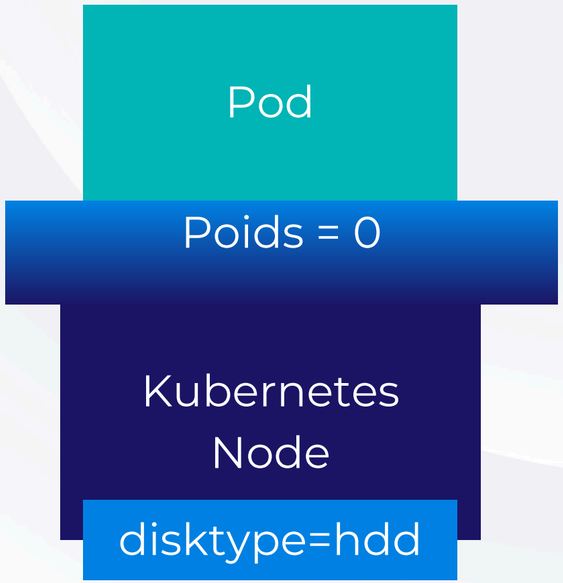
Capacité max = 1pod

Affinity : NodeAffinity
Preferred
Selection

- disktype = ssd = 1
- disktype = nvme = 50







Affinité et anti-affinité.

- Objectif :
  - Contrôler le placement des pods dans le cluster Kubernetes en fonction de règles prédéfinies.
  - Utilisé pour améliorer la distribution des workloads, optimiser l'utilisation des ressources, et répondre à des exigences spécifiques de performance ou de sécurité.
- Types :
  - Affinité de Pod: Pour spécifier où les pods doivent (ou ne doivent pas) être planifiés.
  - Anti-Affinité de Pod : Pour spécifier quels pods ne doivent pas être placés ensemble.

Affinité et anti-affinité.

- Affinité de Pod :
  - Encourage les pods à être placés à proximité d'autres pods pour améliorer la performance et la communication intra-cluster.
  - Utilise des règles basées sur des labels de pods.
  - TopologyKey : Que représente une zone dans le cluster.

```
spec:
  affinity:
    podAffinity:
      requiredDuringSchedulingIgnoredDuringExecution:
        - labelSelector:
            matchLabels:
              app: frontend
          topologyKey: "kubernetes.io/hostname"
  containers:
  - name: my-app-container
    image: my-app-image:v1
```

Le pod sera programmé sur le même nœud que les pods avec le label app=frontend

Affinité et anti-affinité.

```
spec:
 affinity:
  podAffinity:
   preferredDuringSchedulingIgnoredDuringExecution:
   - weight: 100
     podAffinityTerm:
      labelSelector:
       matchExpressions:
       - key: app
         operator: In
         values:
         - frontend
       topologyKey: kubernetes.io/hostname
 containers:
- name: my-app-container
  image: my-app-image:v1
```

Capacité max = 2pod

Pod

Pod

Pod

Affinity : PodAffinity
Required
Topology = hostname
Label = (app = db)

Pod
app=db

Kubernetes
Node

hostname=n1

room=r1

Kubernetes
Node

hostname=n2

room=r1

Kubernetes
Node

hostname=n3

room=r2

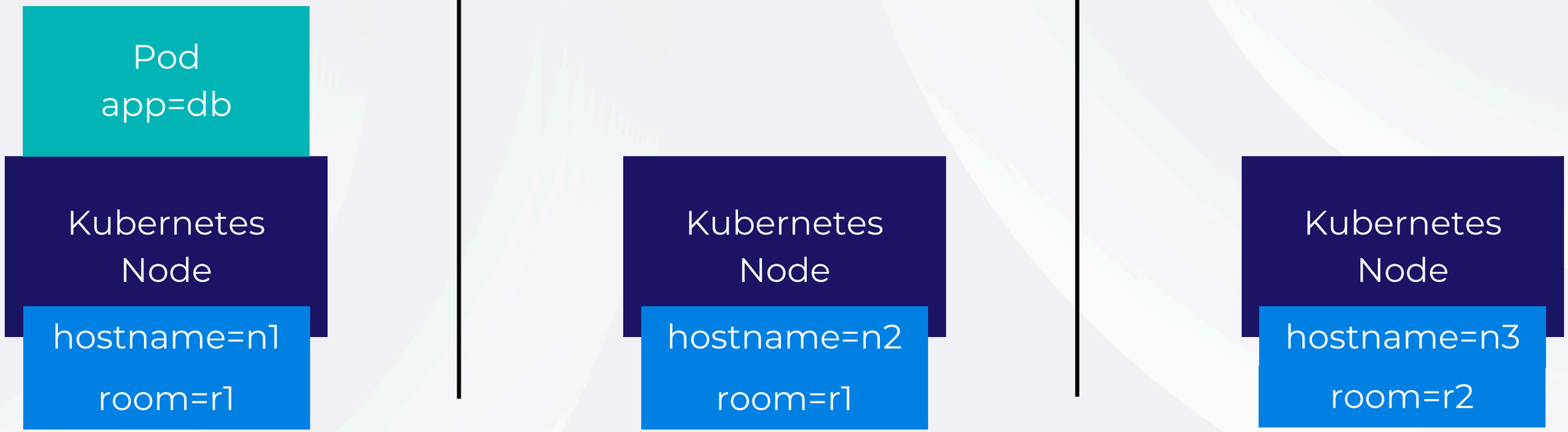
Capacité max = 2pod

Pod

Pod

Pod

Affinity : PodAffinity
Required
Topology = hostname
Label = (app = db)

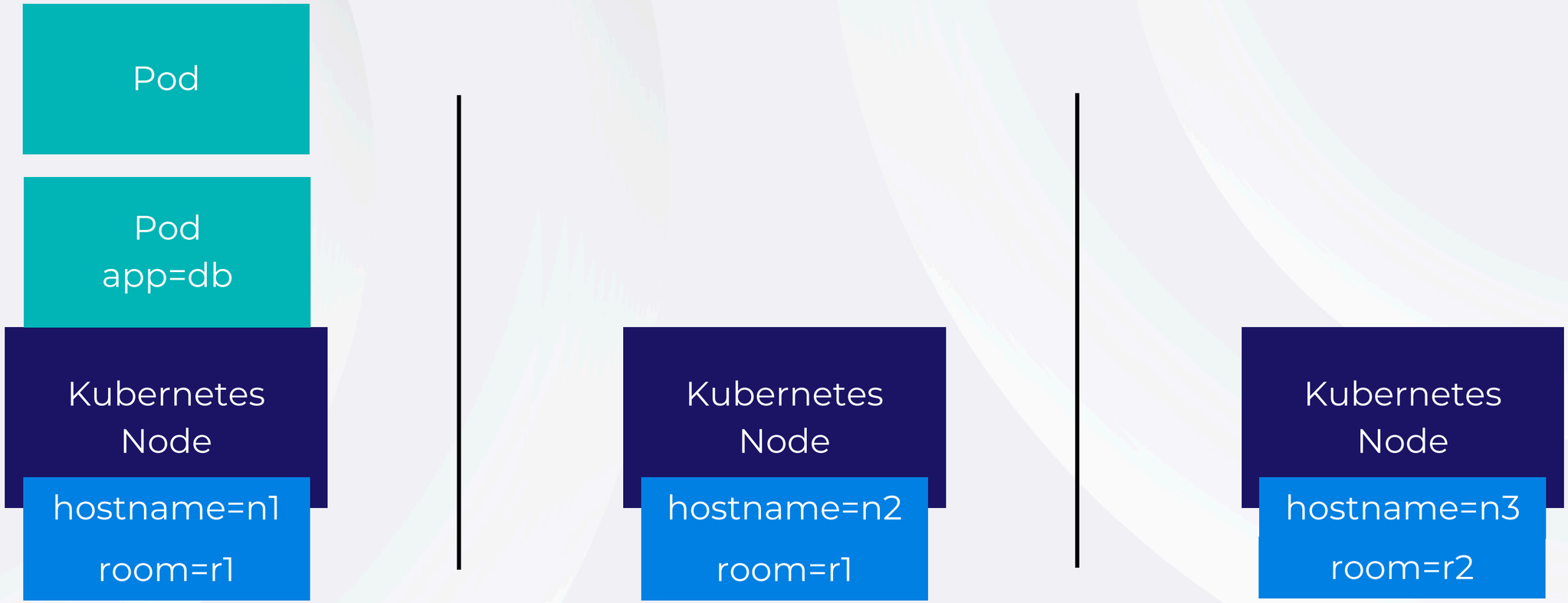


Capacité max = 2pod

Pod

Pod

Affinity : PodAffinity
Required
Topology = hostname
Label = (app = db)

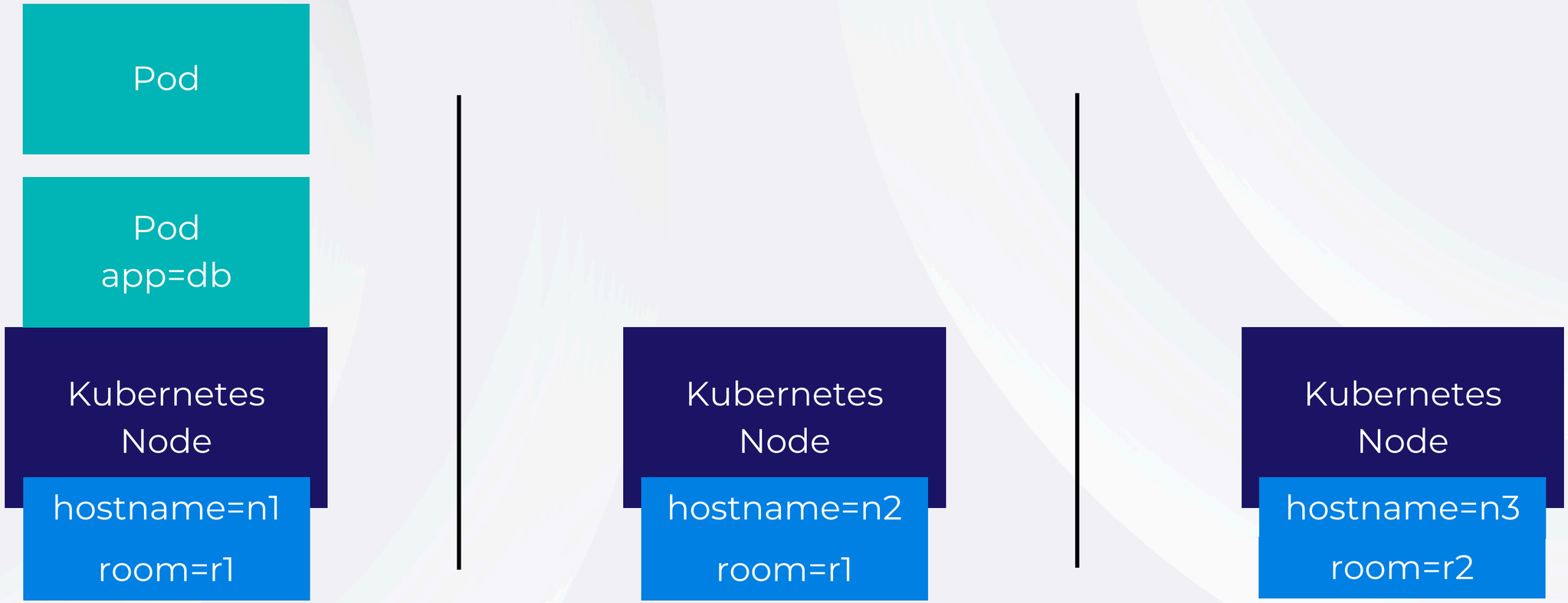


Capacité max = 2pod

Pod
(REJET)

Pod
(REJET)

Affinity : PodAffinity
Required
Topology = hostname
Label = (app = db)



Capacité max = 2pod

Pod

Pod

Pod

Affinity : PodAffinity
Required
Topology = hostname
Label = (app = db)

Kubernetes Node

hostname=n1

room=r1

Kubernetes Node

hostname=n2

room=r1

Kubernetes Node

hostname=n3

room=r2

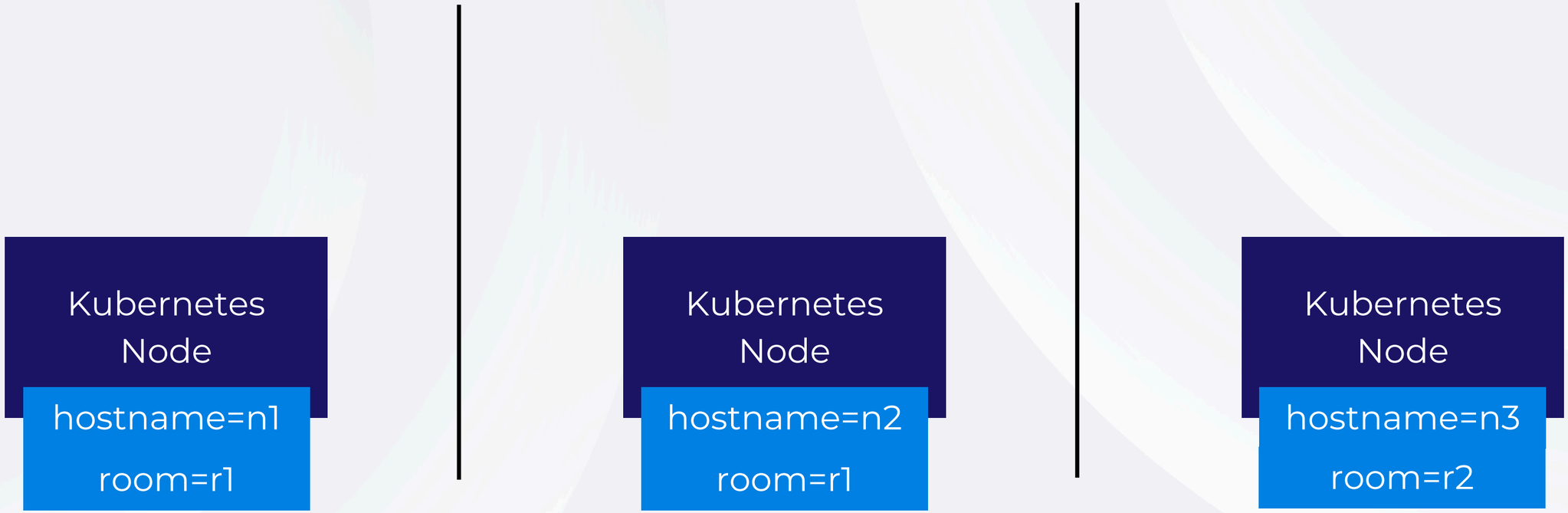
Capacité max = 2pod

Pod

Pod

Pod

Affinity : PodAffinity
Required
Topology = hostname
Label = (app = db)



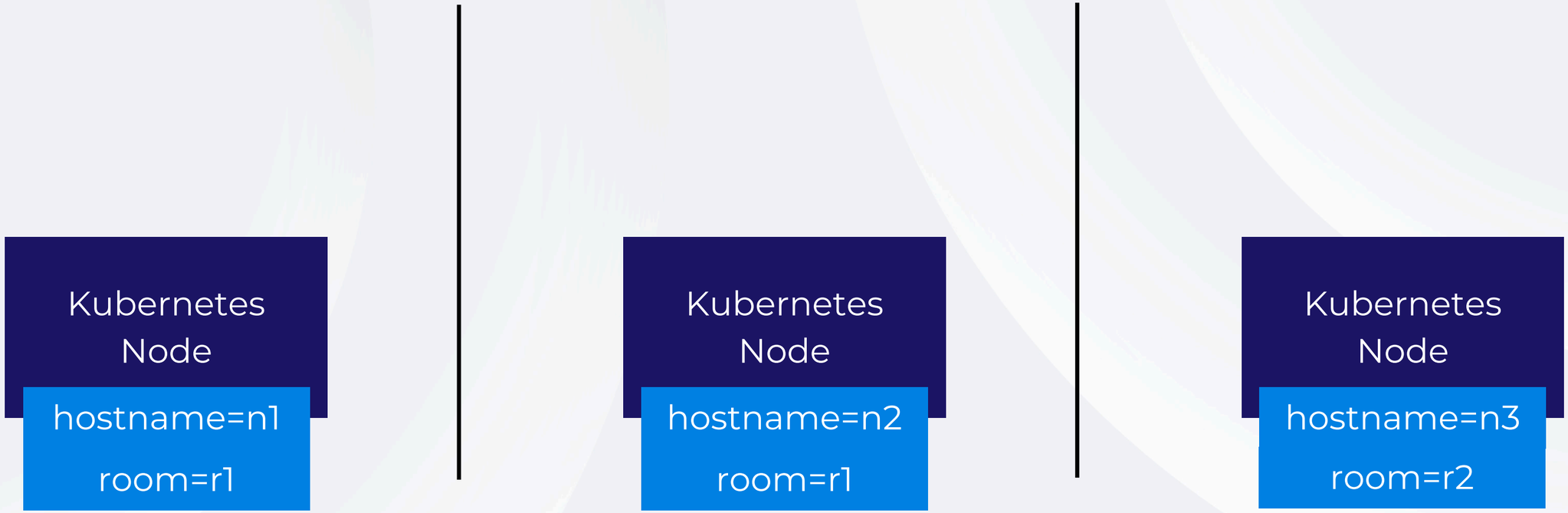
Capacité max = 2pod

Pod
(REJET)

Pod
(REJET)

Pod
(REJET)

Affinity : PodAffinity
Required
Topology = hostname
Label = (app = db)



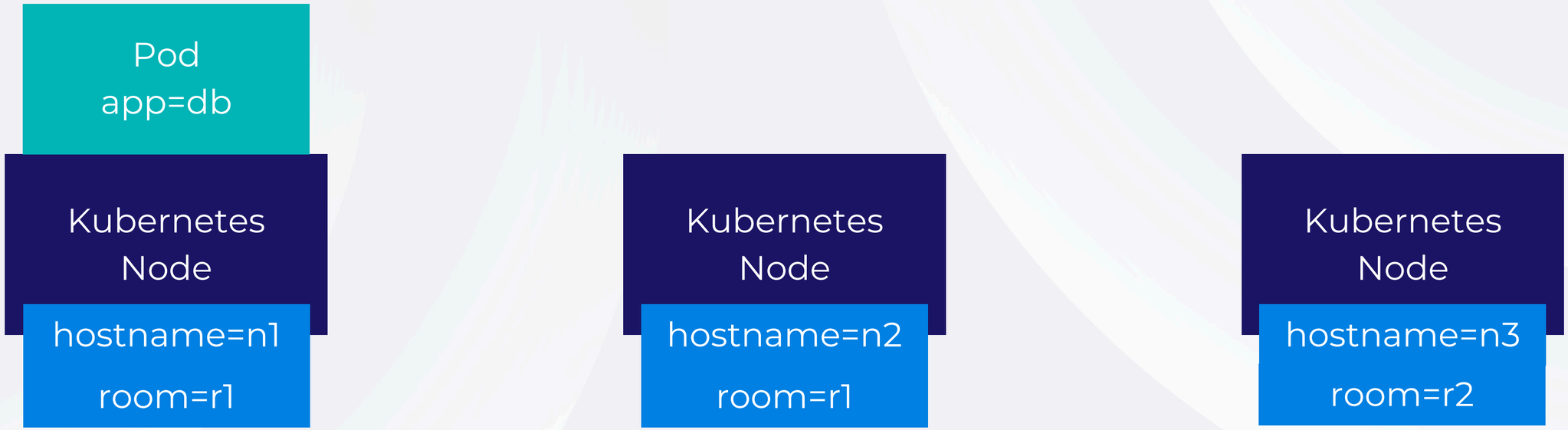
Capacité max = 2pod

Pod

Pod

Pod

Affinity : PodAffinity
Required
Topology = room
Label = (app = db)



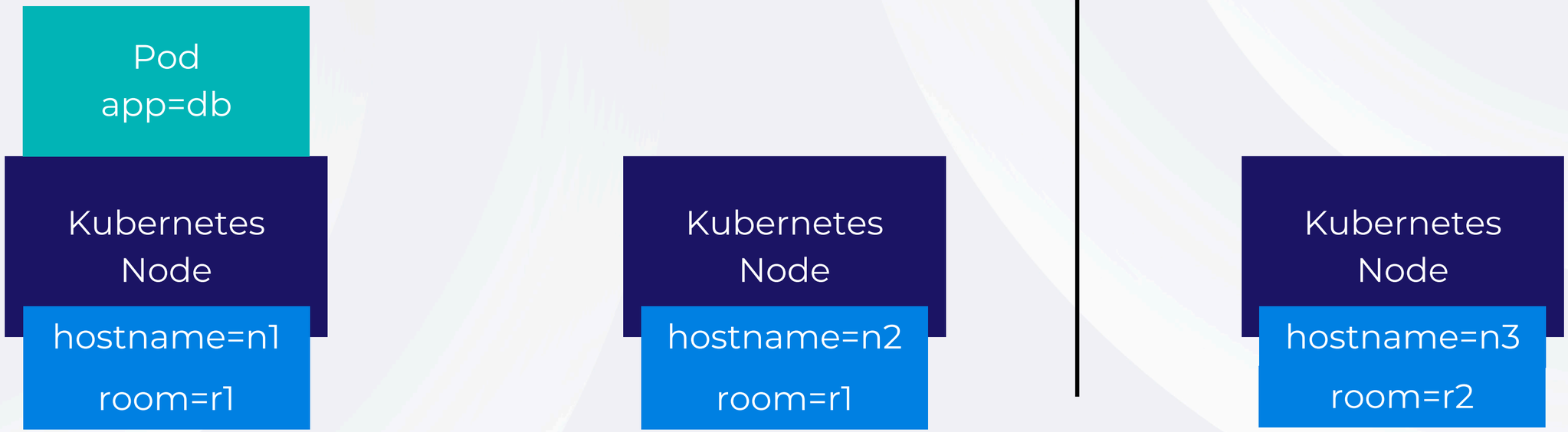
Capacité max = 2pod

Pod

Pod

Pod

Affinity : PodAffinity
Required
Topology = room
Label = (app = db)

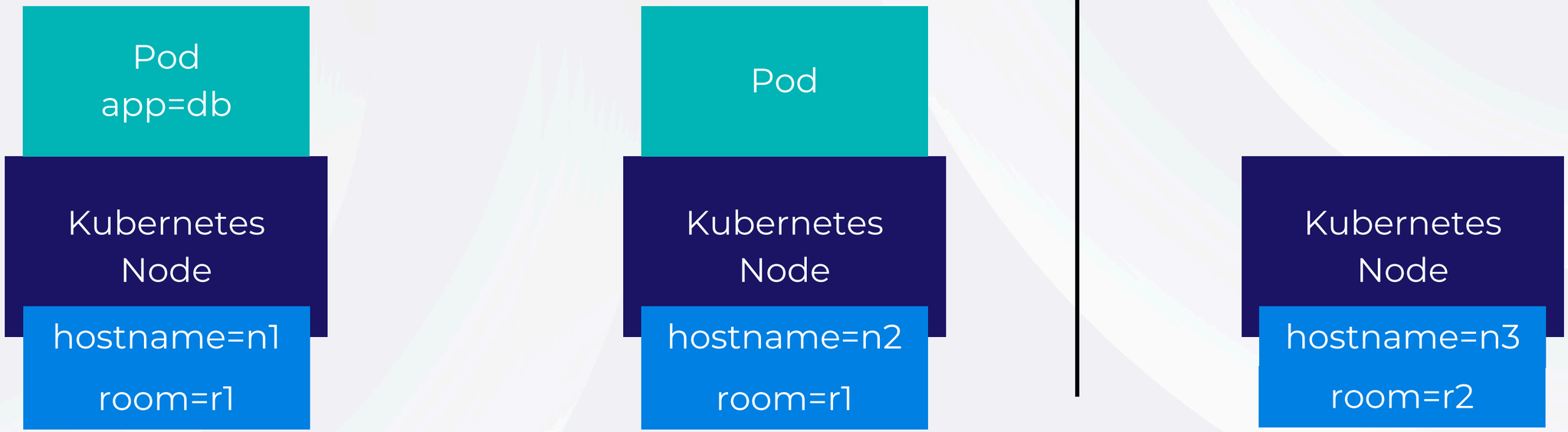


Capacité max = 2pod

Pod

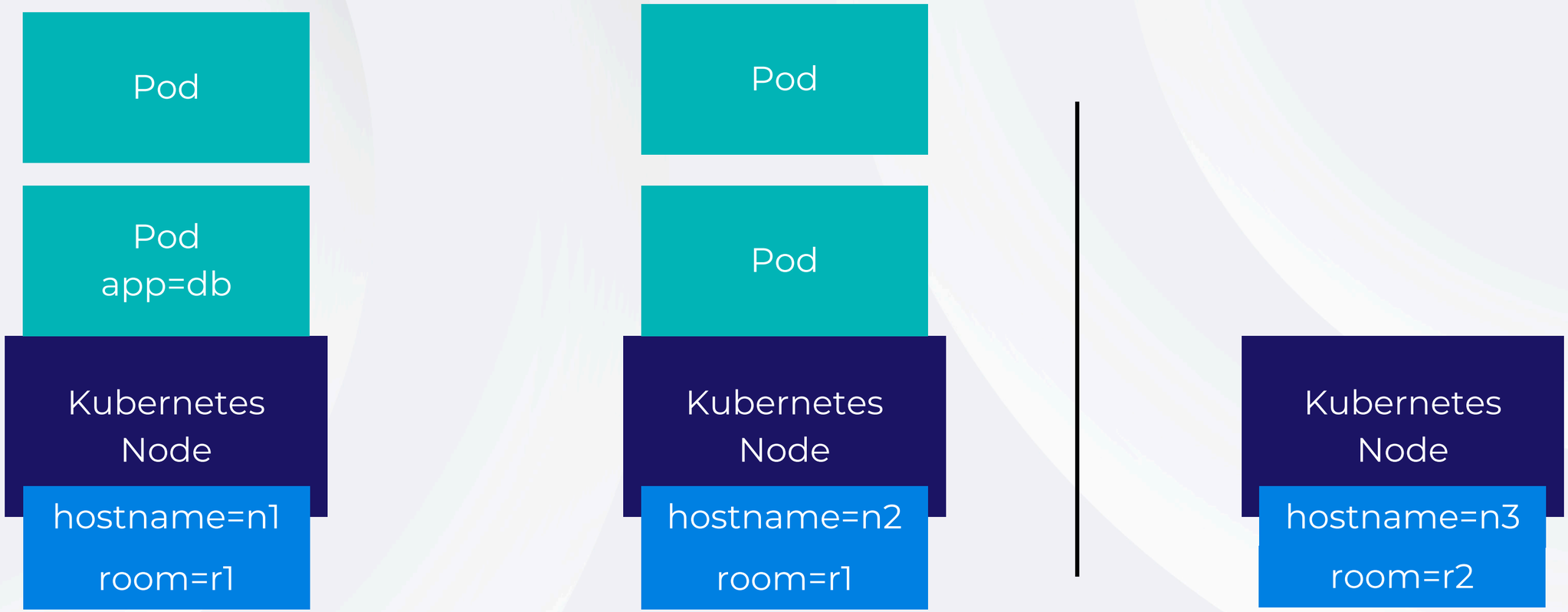
Pod

Affinity : PodAffinity
Required
Topology = room
Label = (app = db)



Capacité max = 2pod

Affinity : PodAffinity
Required
Topology = room
Label = (app = db)



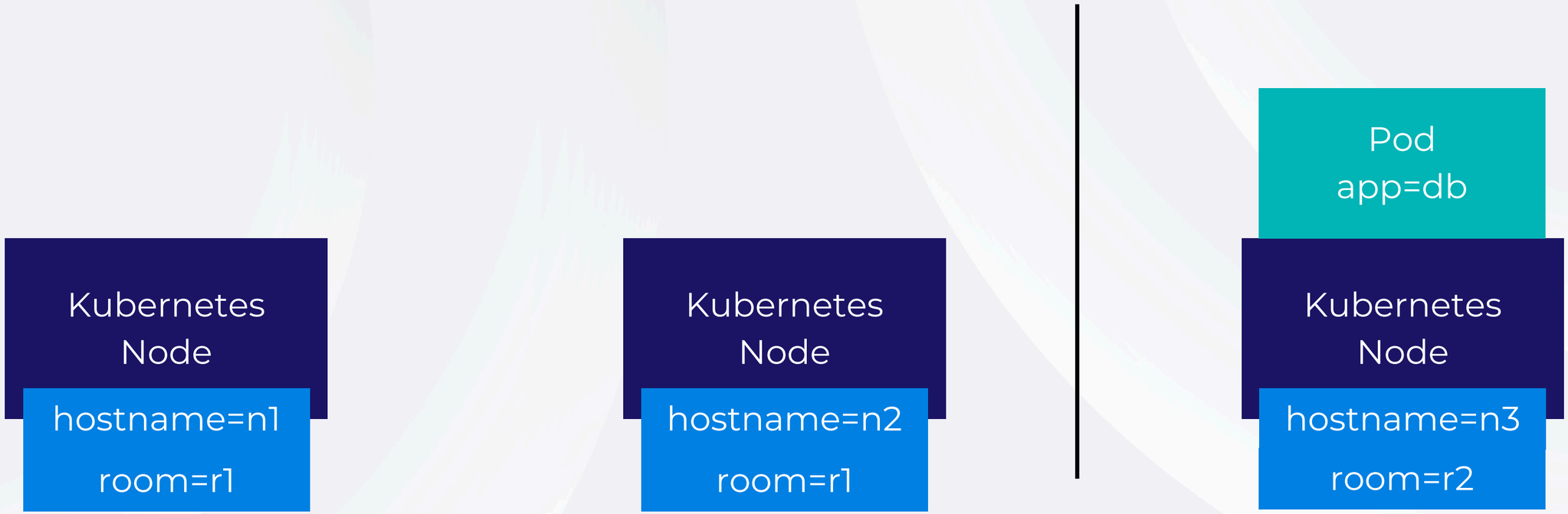
Capacité max = 2pod

Pod

Pod

Pod

Affinity : PodAffinity
Required
Topology = room
Label = (app = db)



Capacité max = 2pod

Pod

Pod

Affinity : PodAffinity
Required
Topology = room
Label = (app = db)

Pod

Pod
app=db

Kubernetes
Node

hostname=n1

room=r1

Kubernetes
Node

hostname=n2

room=r1

Kubernetes
Node

hostname=n3

room=r2

Capacité max = 2pod

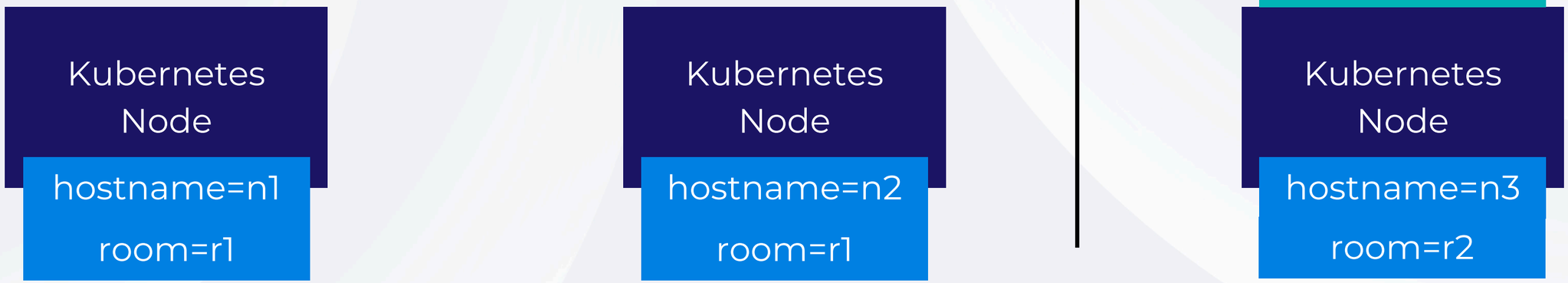
Pod
(REJET)

Pod
(REJET)

Affinity : PodAffinity
Required
Topology = room
Label = (app = db)

Pod

Pod
app=db



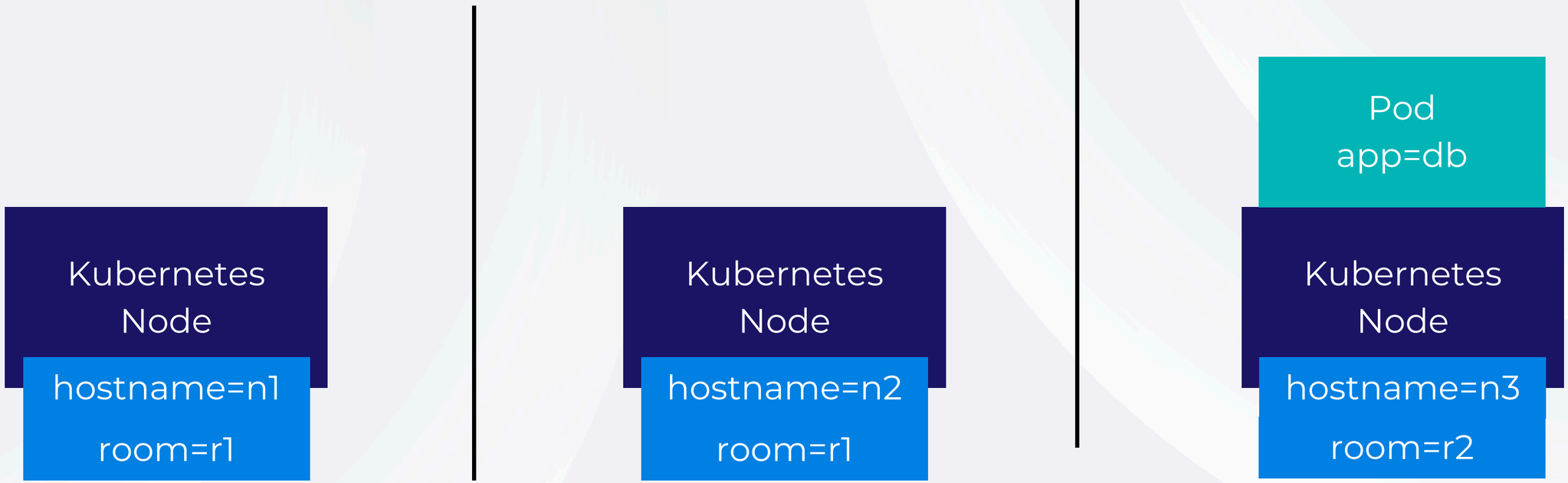
Capacité max = 2pod

Pod

Pod

Pod

Affinity : PodAffinity
Preferred
Topology = hostname
Label = (app = db) = 50



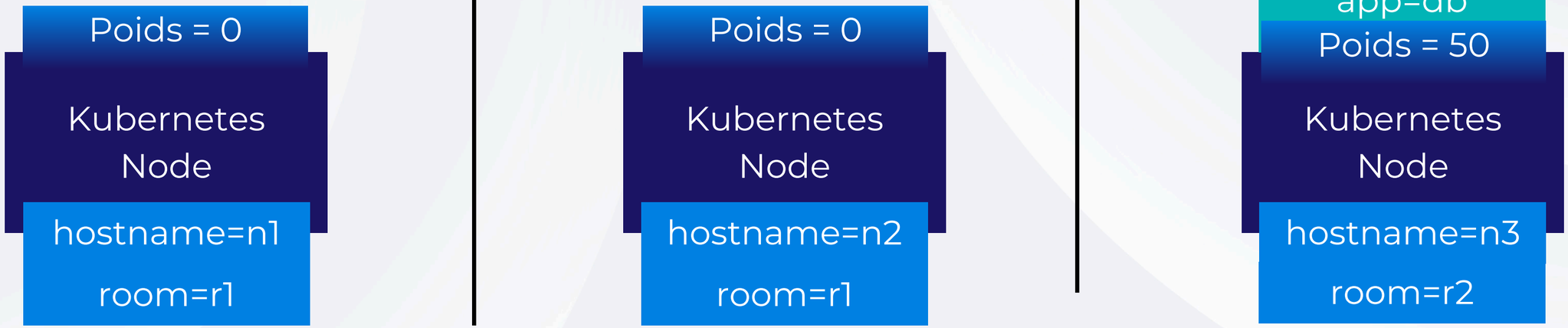
Capacité max = 2pod

Pod

Pod

Pod

Affinity : PodAffinity
Preferred
Topology = hostname
Label = (app = db) = 50

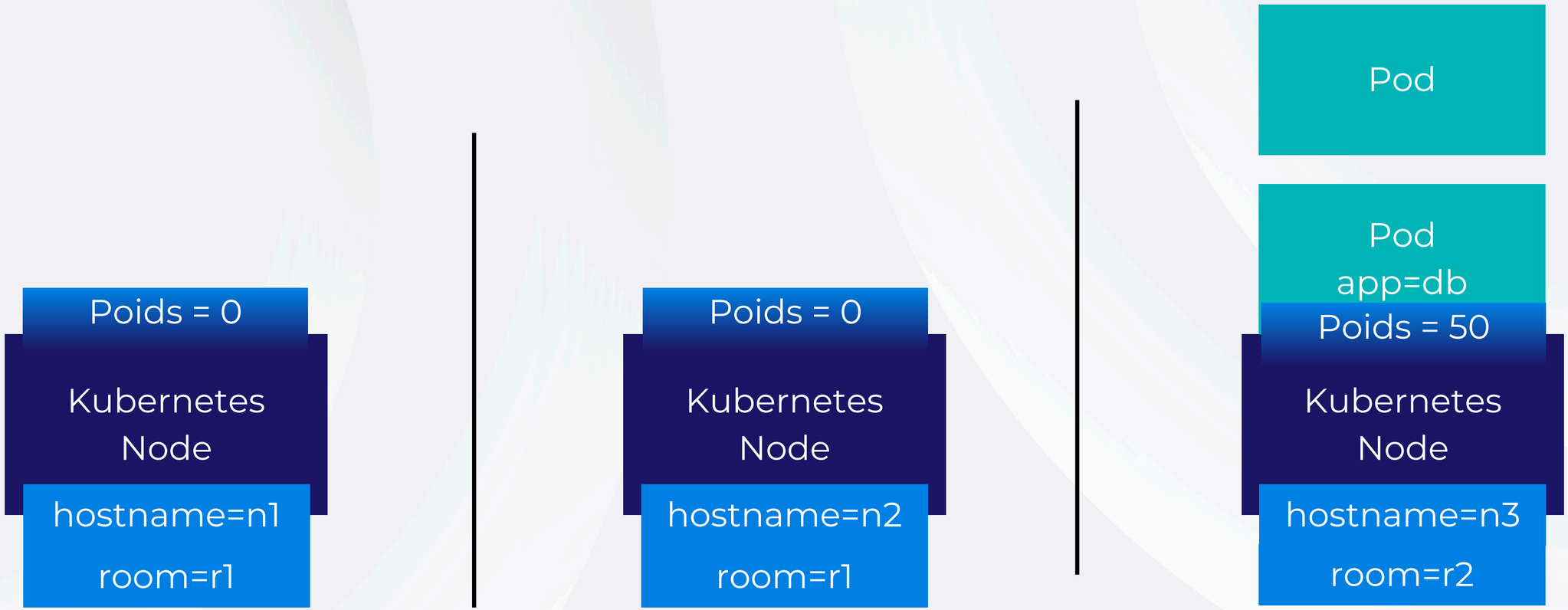


Capacité max = 2pod

Pod

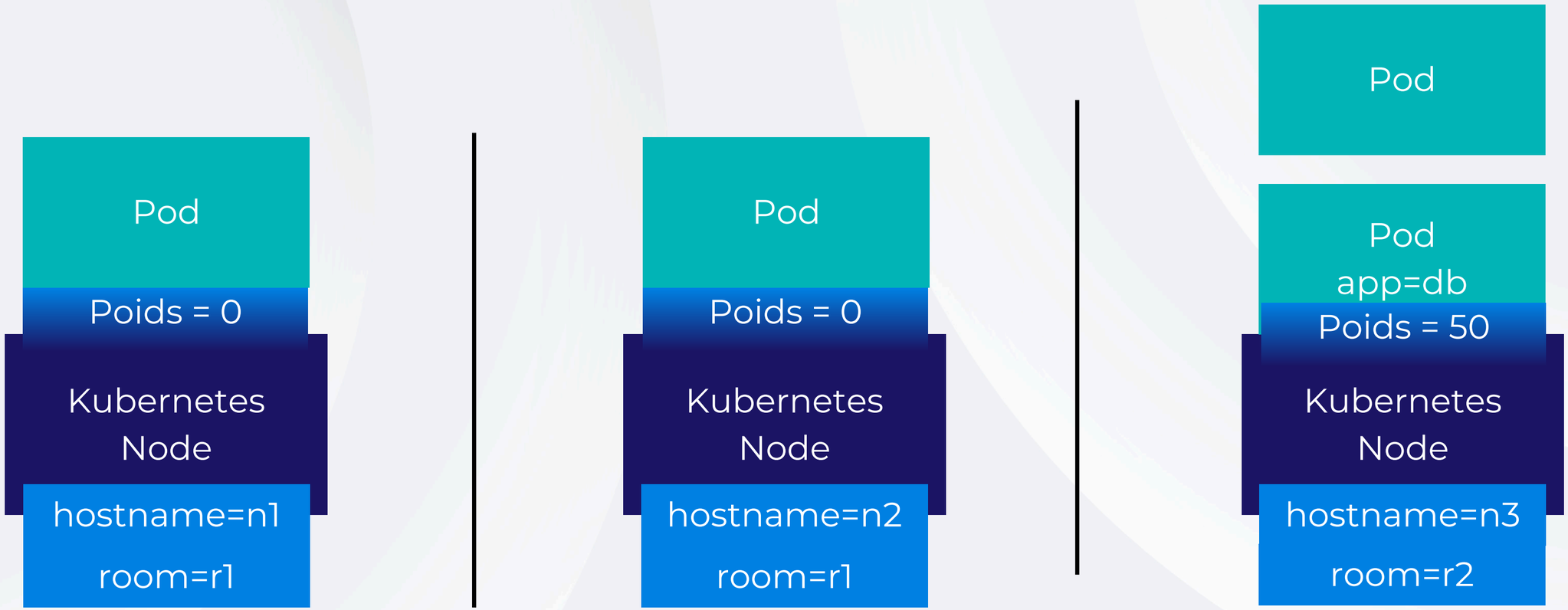
Pod

Affinity : PodAffinity
Preferred
Topology = hostname
Label = (app = db) = 50



Capacité max = 2pod

Affinity : PodAffinity
Preferred
Topology = hostname
Label = (app = db) = 50



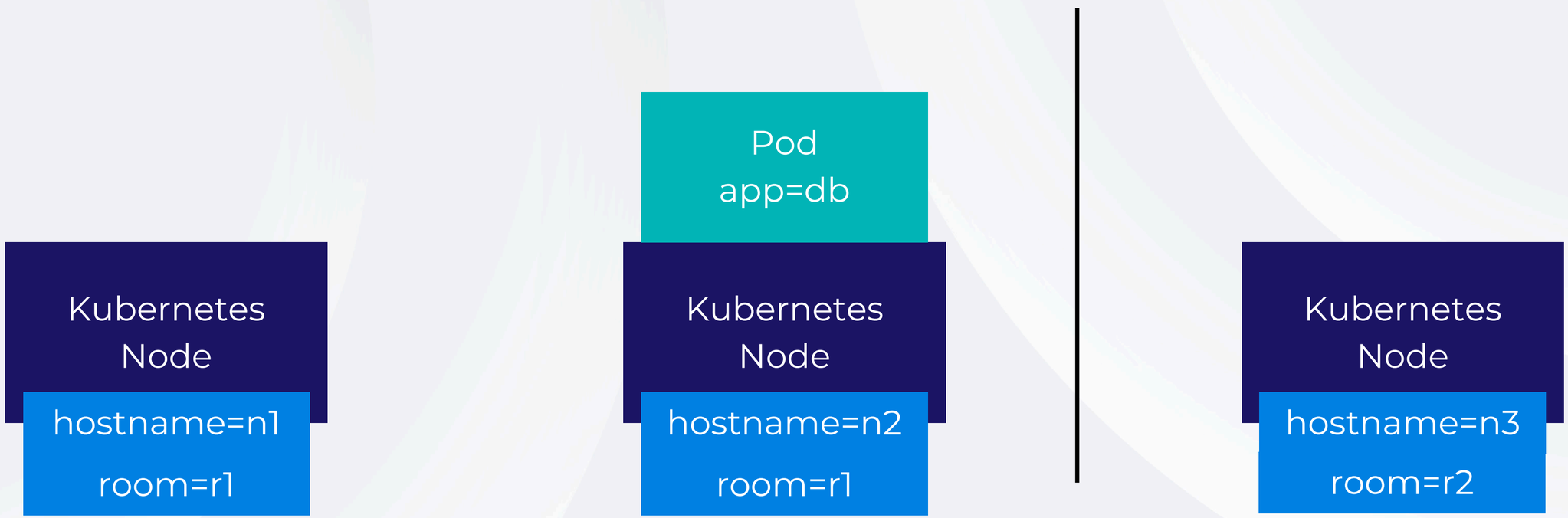
Capacité max = 2pod

Pod

Pod

Pod

Affinity : PodAffinity
Preferred
Topology = room
Label = (app = db) = 50



Capacité max = 2pod

Pod

Pod

Pod

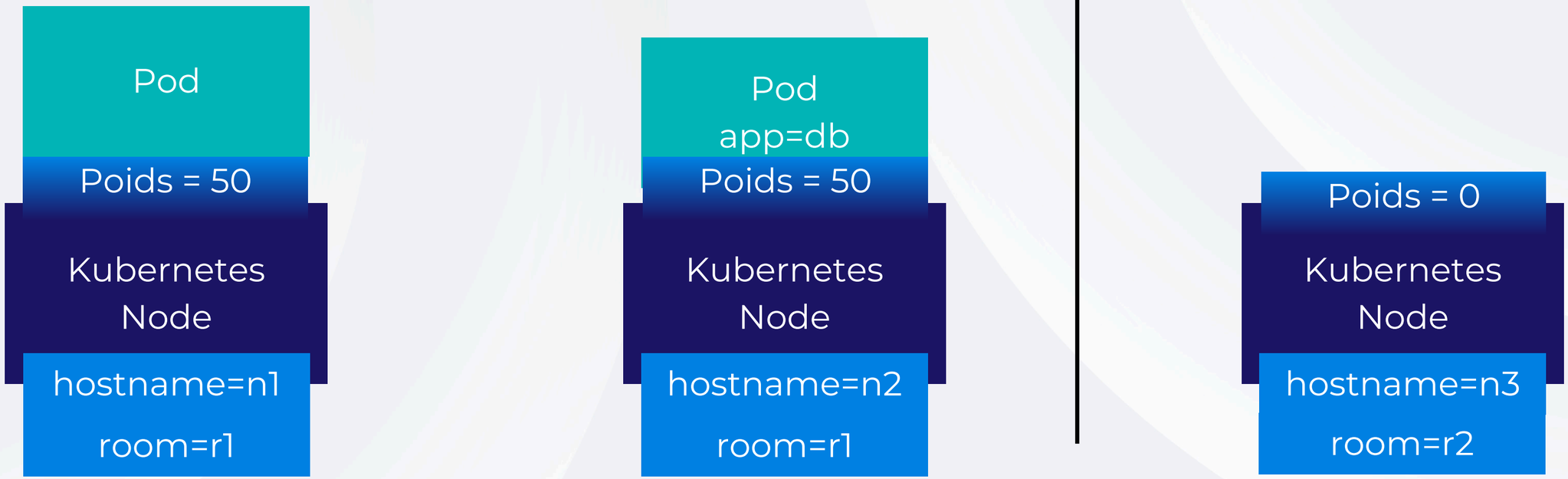
Affinity : PodAffinity
Preferred
Topology = room
Label = (app = db) = 50

Capacité max = 2pod

Pod

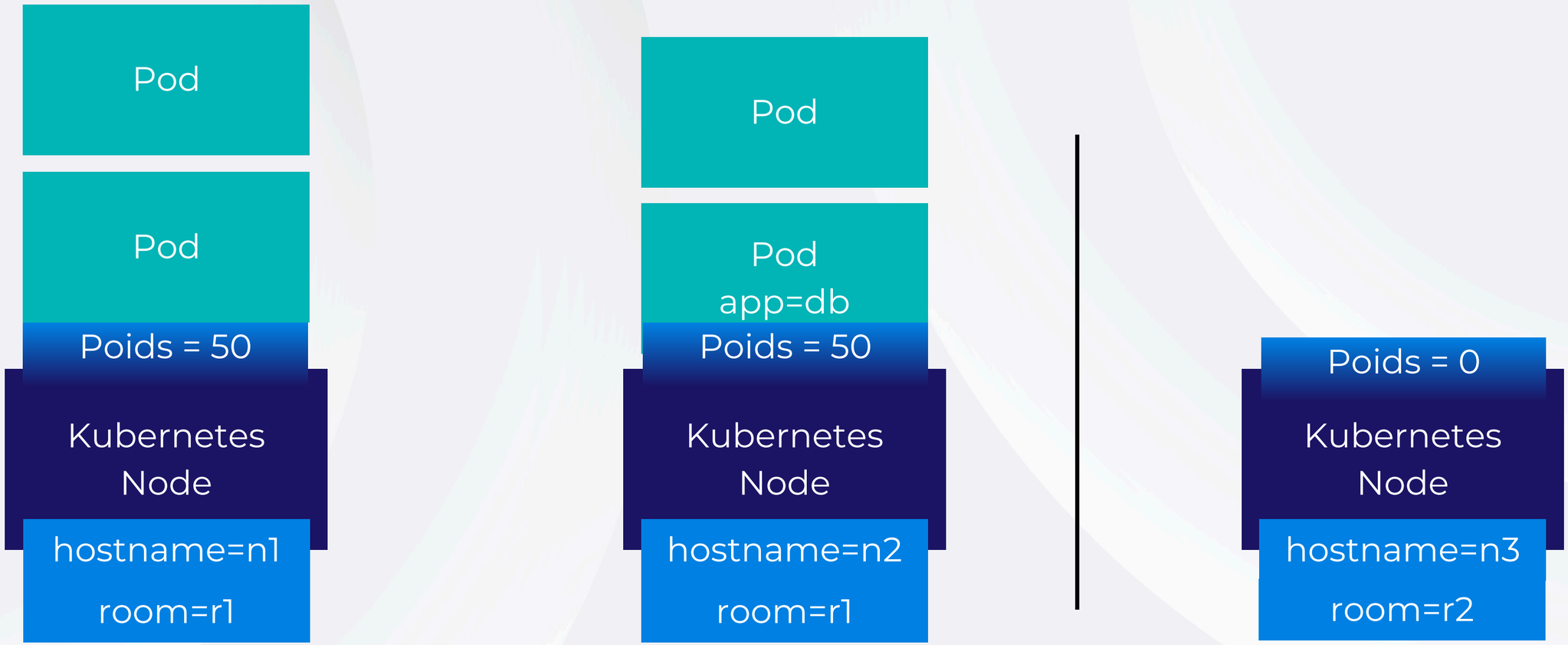
Pod

Affinity : PodAffinity
Preferred
Topology = room
Label = (app = db) = 50



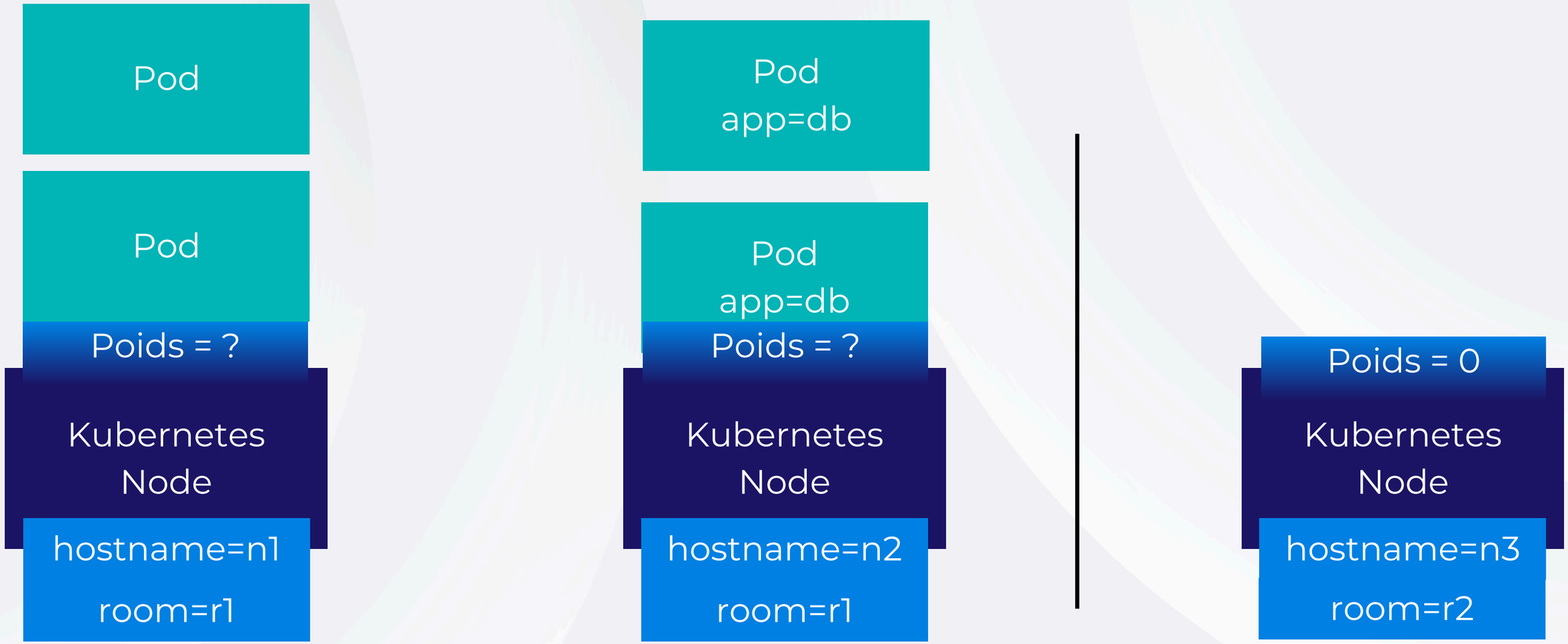
Capacité max = 2pod

Affinity : PodAffinity
Preferred
Topology = room
Label = (app = db) = 50



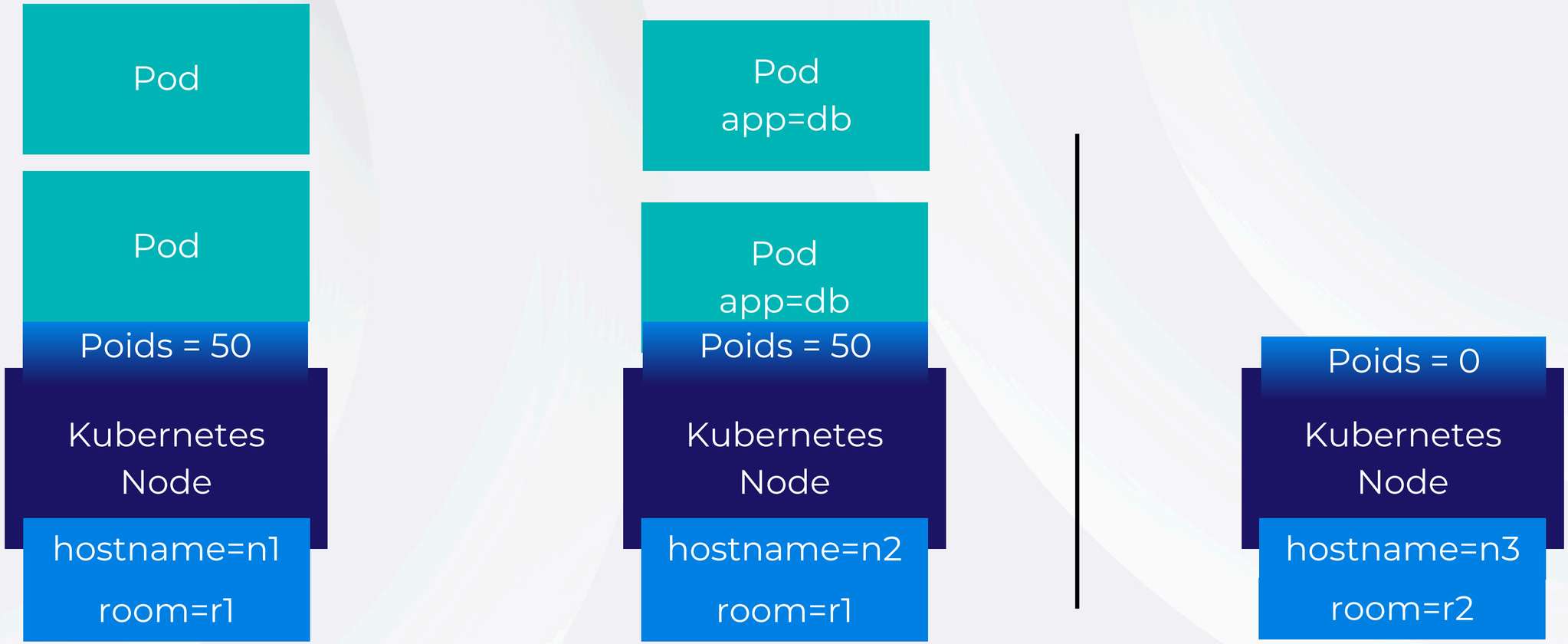
Capacité max = 2pod

Affinity : PodAffinity
Preferred
Topology = room
Label = (app = db) = 50



Capacité max = 2pod

Affinity : PodAffinity
Preferred
Topology = room
Label = (app = db) = 50



Affinité et anti-affinité.

- Anti-Affinité de Pod :
  - Empêche les pods d’être planifiés ensemble sur les mêmes nœuds pour répartir la charge ou pour des raisons de tolérance aux pannes.
  - Utile pour éviter les « points chauds » ou les dépendances critiques sur un seul nœud.

```
spec:
  affinity:
    podAntiAffinity:
      requiredDuringSchedulingIgnoredDuringExecution:
        - labelSelector:
            matchLabels:
              app: frontend
          topologyKey: "kubernetes.io/hostname"
  containers:
  - name: my-app-container
    image: my-app-image:v1
```

Le pod ne sera pas programmé sur un nœud qui exécute déjà des pods avec le label app=frontend.

Affinité et anti-affinité.

```
spec:
 affinity:
  podAntiAffinity:
   preferredDuringSchedulingIgnoredDuringExecution:
   - weight: 100
     podAffinityTerm:
      labelSelector:
       matchExpressions:
       - key: app
         operator: In
         values:
         - frontend
       topologyKey: topology.kubernetes.io/zone
 containers:
- name: my-app-container
  image: my-app-image:v1
```

Capacité max = 2pod

Pod
app=db

Pod
app=db

Pod
app=db

Pod
app=db

Affinity : PodAntiAffinity
Required
Topology = hostname
Label = (app = db)

Kubernetes
Node

hostname=n1

room=r1

Kubernetes
Node

hostname=n2

room=r1

Kubernetes
Node

hostname=n3

room=r2

Capacité max = 2pod

Affinity : PodAntiAffinity
Required
Topology = hostname
Label = (app = db)

Pod
app=db

Pod
app=db

Pod
app=db

Pod
app=db

Kubernetes
Node

hostname=n1

room=r1

Kubernetes
Node

hostname=n2

room=r1

Kubernetes
Node

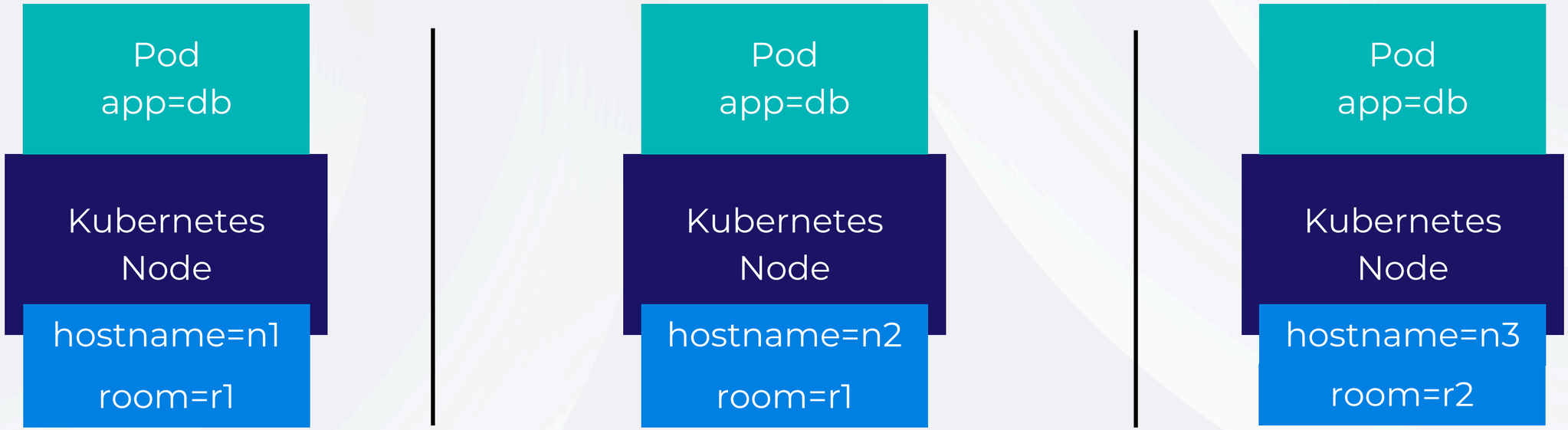
hostname=n3

room=r2

Capacité max = 2pod

Affinity : PodAntiAffinity
Required
Topology = hostname
Label = (app = db)

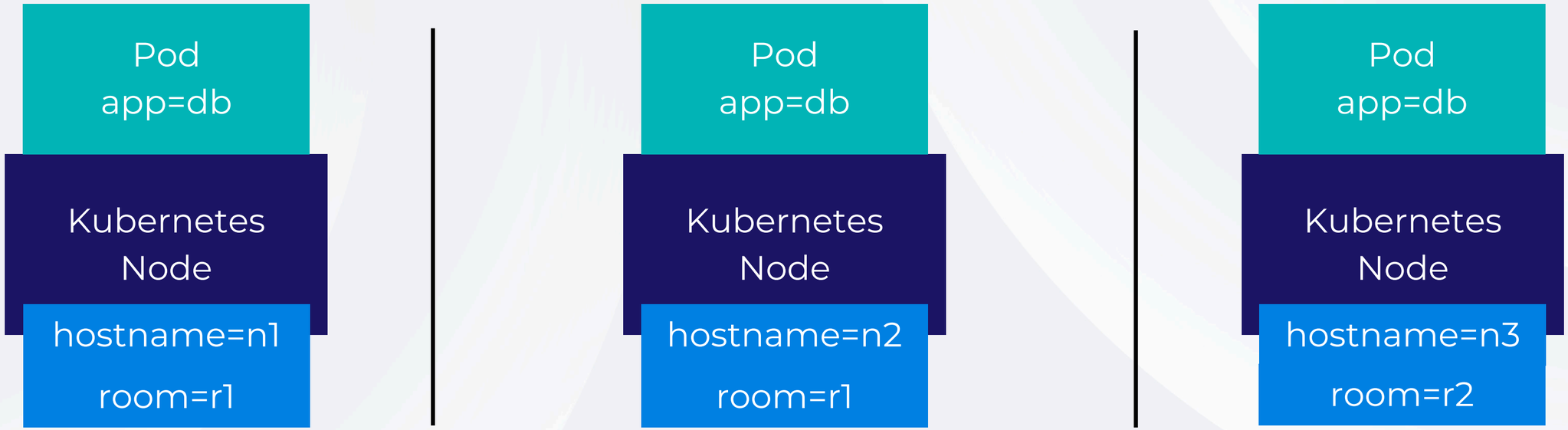
Pod
app=db



Capacité max = 2pod

Pod
app=db
(REJET)

Affinity : PodAntiAffinity
Required
Topology = hostname
Label = (app = db)



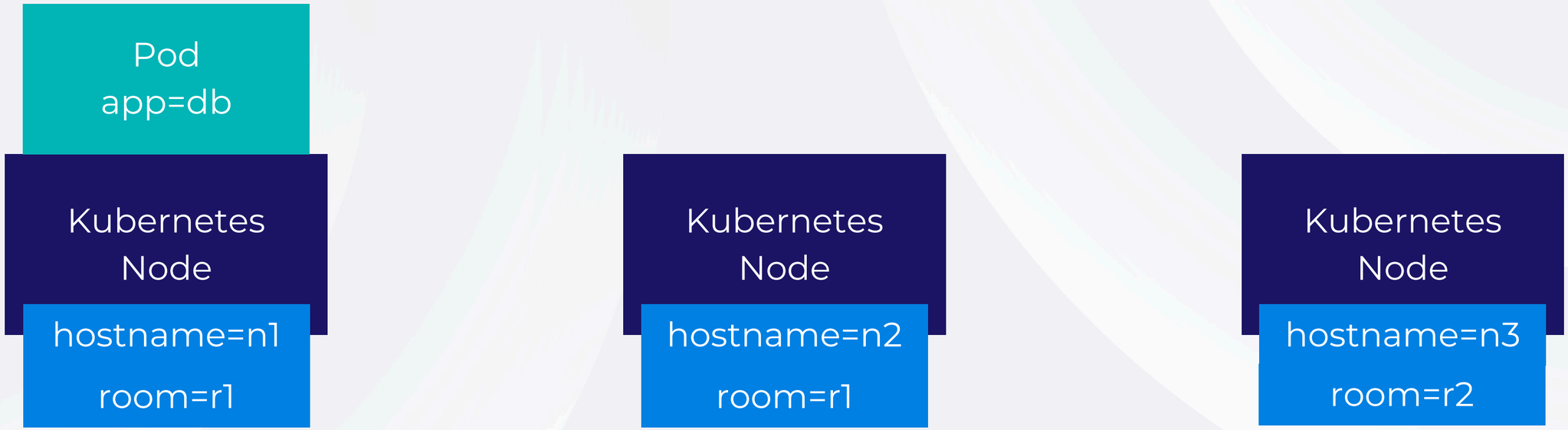
Capacité max = 2pod

Pod
app=db

Pod
app=db

Pod
app=db

Affinity : PodAntiAffinity
Required
Topology = room
Label = (app = db)



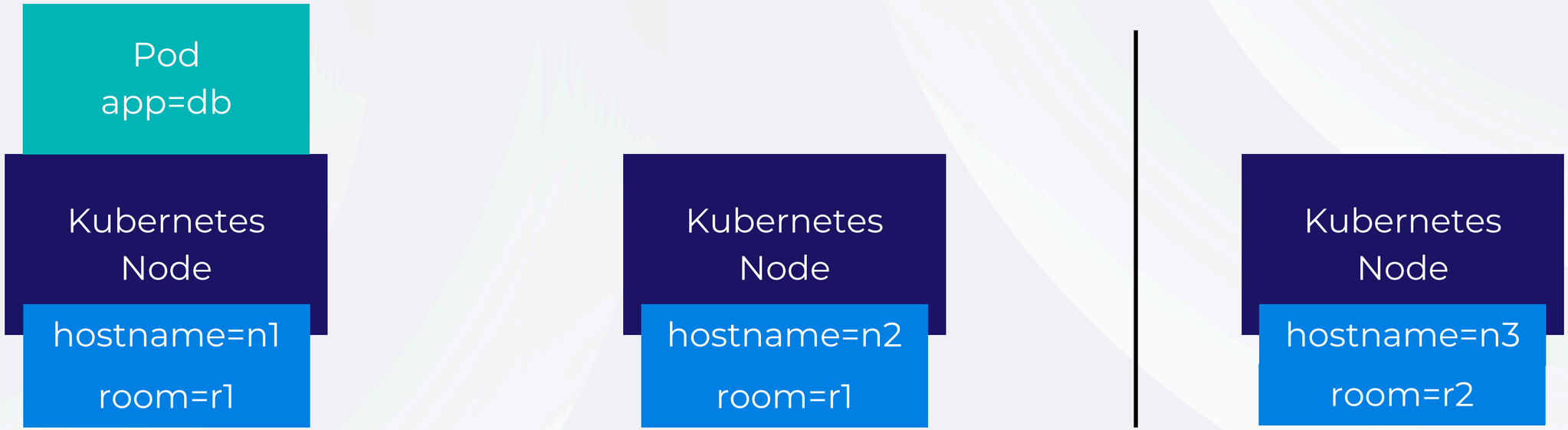
Capacité max = 2pod

Pod
app=db

Pod
app=db

Pod
app=db

Affinity : PodAntiAffinity
Required
Topology = room
Label = (app = db)

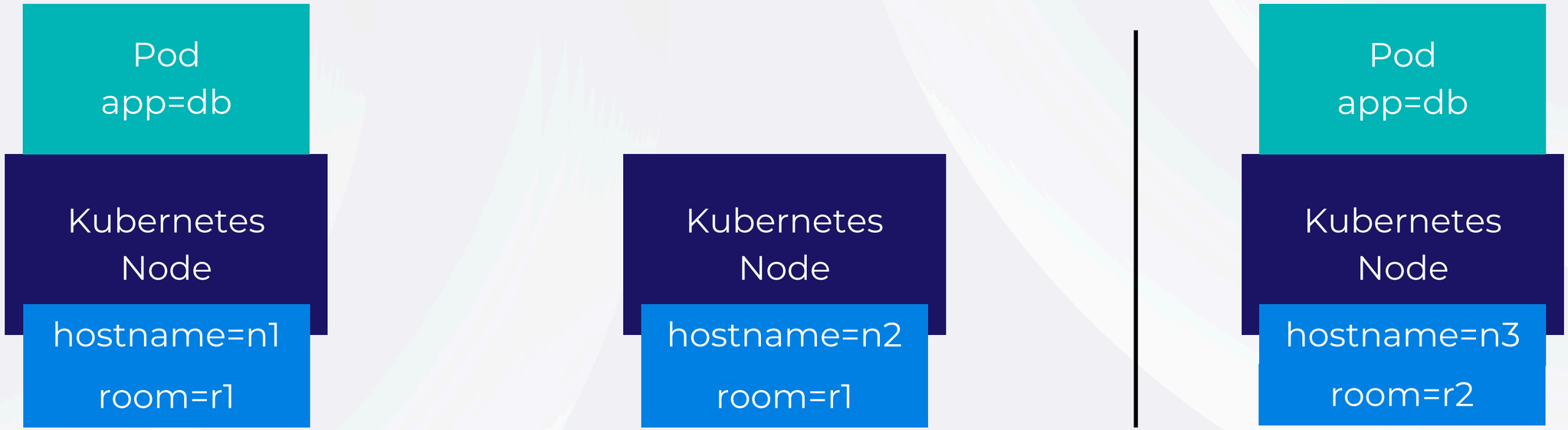


Capacité max = 2pod

Pod
app=db

Pod
app=db

Affinity : PodAntiAffinity
Required
Topology = room
Label = (app = db)

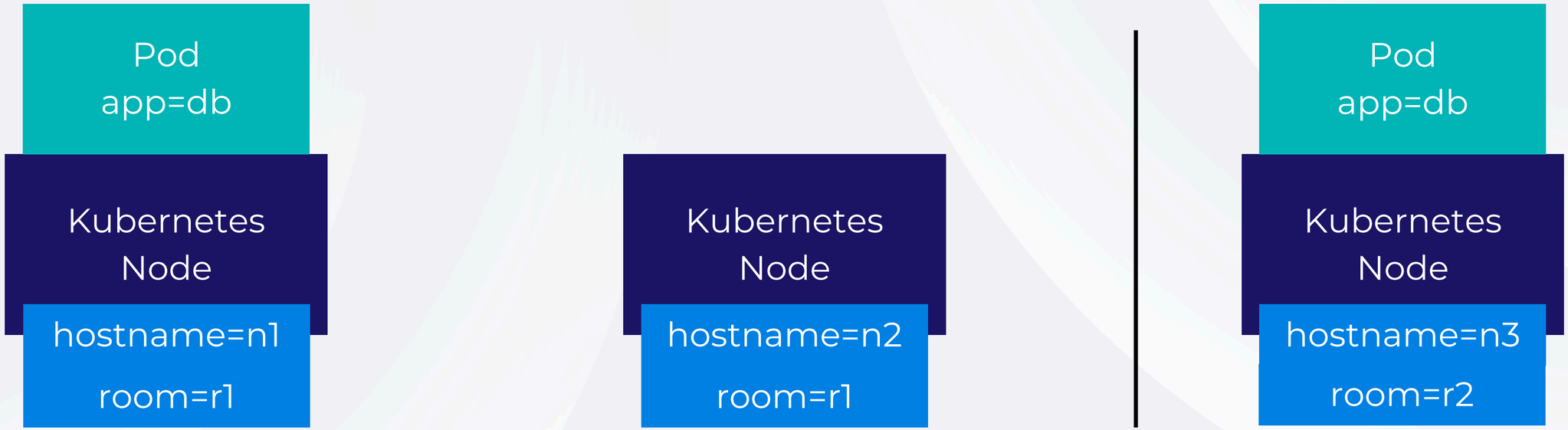


Capacité max = 2pod

Pod
app=db
(REJET)

Pod
app=db
(REJET)

Affinity : PodAntiAffinity
Required
Topology = room
Label = (app = db)



Capacité max = 2pod

Pod
app=db

Pod
app=db

Pod
app=db

Pod
app=db

Affinity : PodAntiAffinity
Preferred
Topology = hostname
Label = (app = db) = 50

Kubernetes
Node

hostname=n1

room=r1

Kubernetes
Node

hostname=n2

room=r1

Kubernetes
Node

hostname=n3

room=r2

Capacité max = 2pod

Pod
app=db

Pod
app=db

Pod
app=db

Pod
app=db

Affinity : PodAntiAffinity
Preferred
Topology = hostname
Label = (app = db) = 50

Poids = ?

Kubernetes
Node

hostname=n1

room=r1

Poids = ?

Kubernetes
Node

hostname=n2

room=r1

Poids = ?

Kubernetes
Node

hostname=n3

room=r2

Capacité max = 2pod

Pod
app=db

Pod
app=db

Pod
app=db

Pod
app=db

Affinity : PodAntiAffinity
Preferred
Topology = hostname
Label = (app = db) = 50

Poids = 50

Kubernetes
Node

hostname=n1

room=r1

Poids = 50

Kubernetes
Node

hostname=n2

room=r1

Poids = 50

Kubernetes
Node

hostname=n3

room=r2

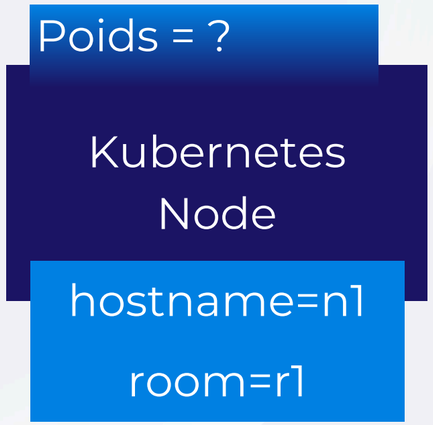
Capacité max = 2pod

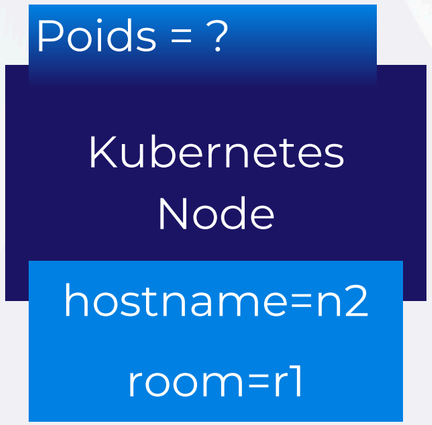
Pod
app=db

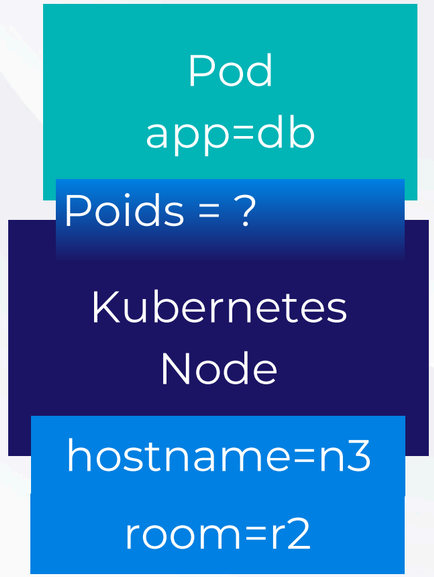
Pod
app=db

Pod
app=db

Affinity : PodAntiAffinity
Preferred
Topology = hostname
Label = (app = db) = 50







Capacité max = 2pod

Pod
app=db

Pod
app=db

Pod
app=db

Affinity : PodAntiAffinity
Preferred
Topology = hostname
Label = (app = db) = 50

Poids = 50

Kubernetes
Node

hostname=n1

room=r1

Poids = 50

Kubernetes
Node

hostname=n2

room=r1

Pod
app=db

Poids = 0

Kubernetes
Node

hostname=n3

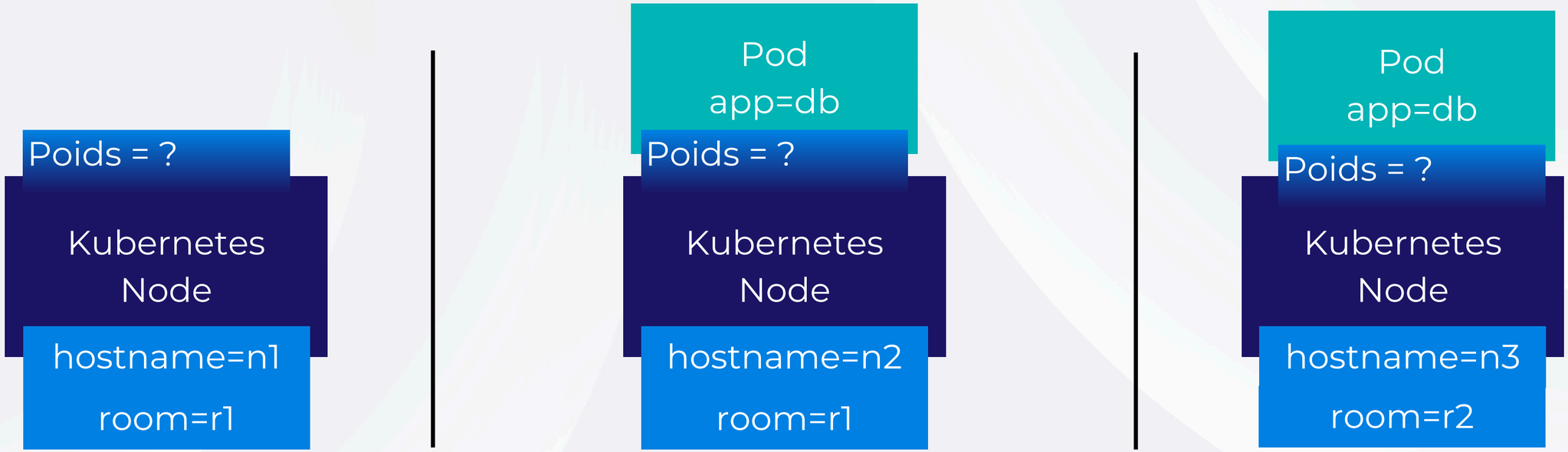
room=r2

Capacité max = 2pod

Pod
app=db

Pod
app=db

Affinity : PodAntiAffinity
Preferred
Topology = hostname
Label = (app = db) = 50

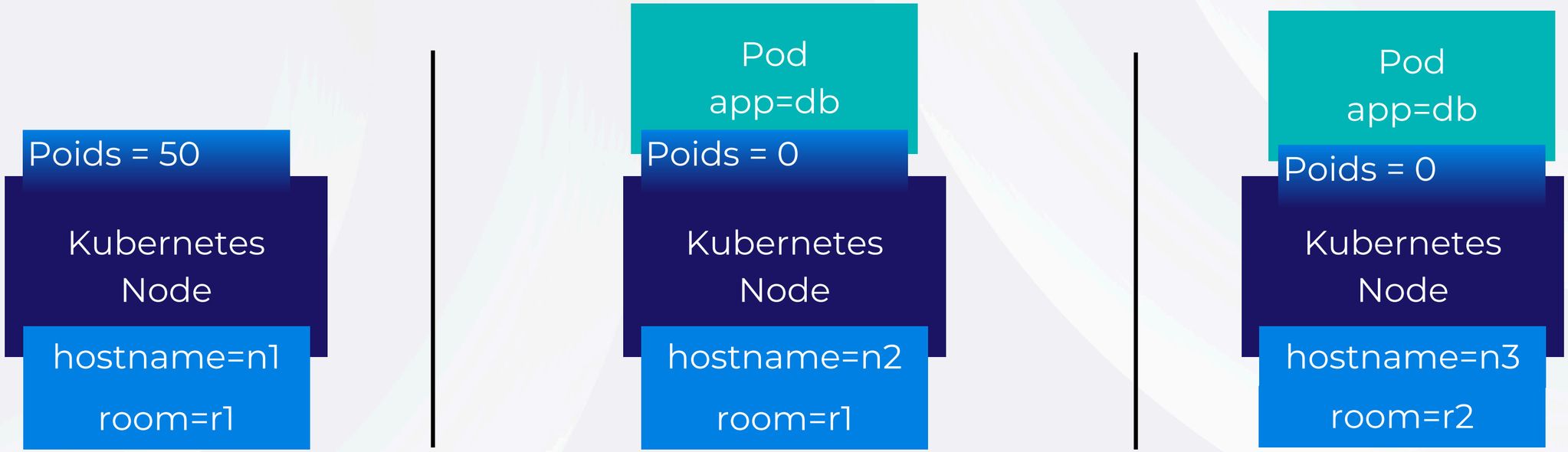


Capacité max = 2pod

Pod
app=db

Pod
app=db

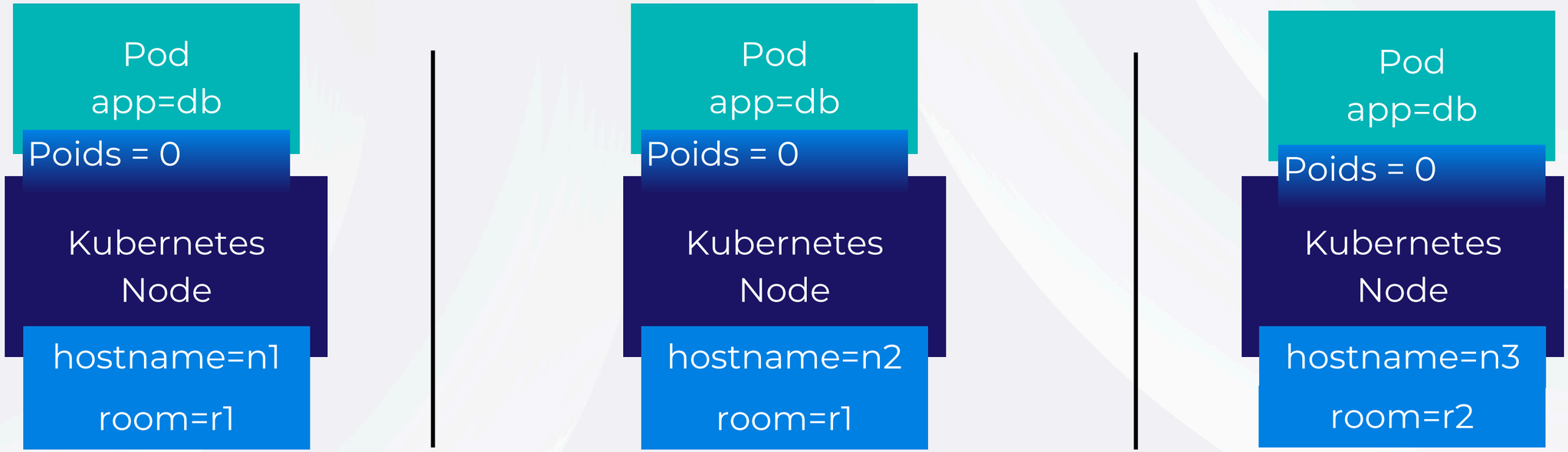
Affinity : PodAntiAffinity
Preferred
Topology = hostname
Label = (app = db) = 50



Capacité max = 2pod

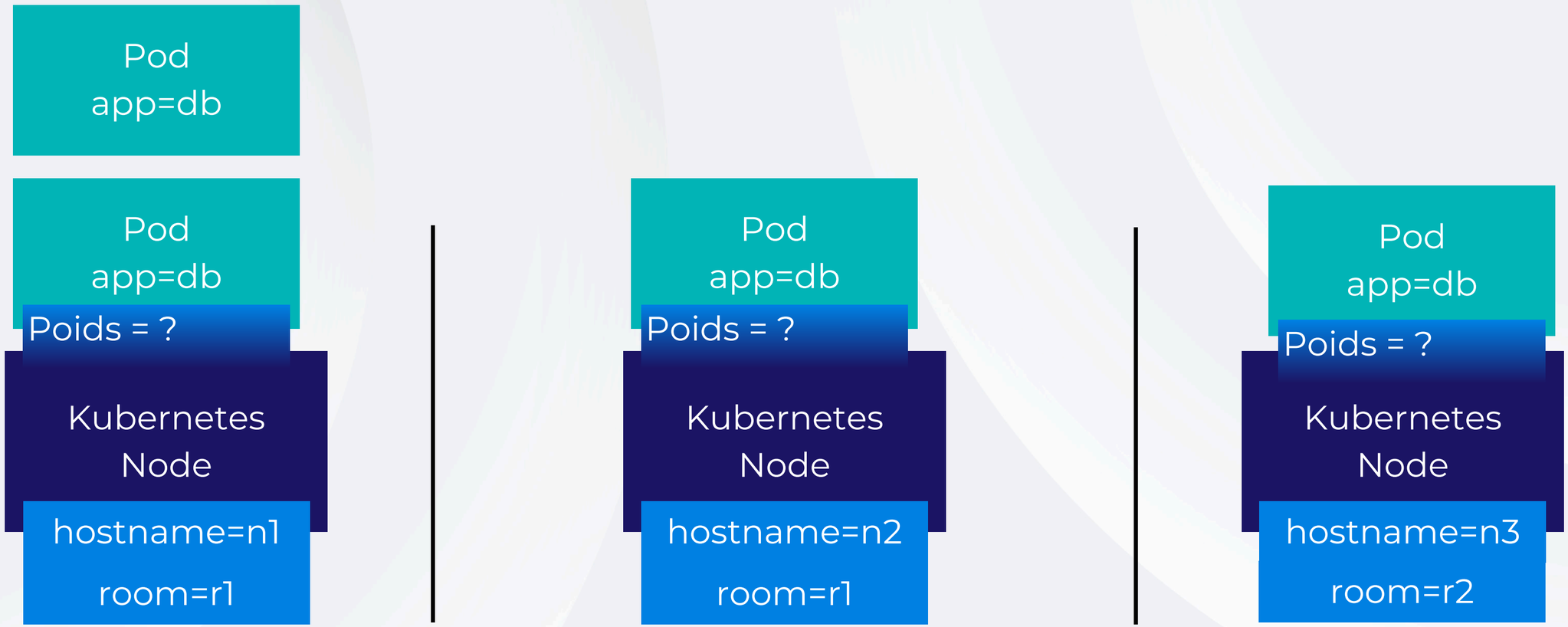
Pod
app=db

Affinity : PodAntiAffinity
Preferred
Topology = hostname
Label = (app = db) = 50



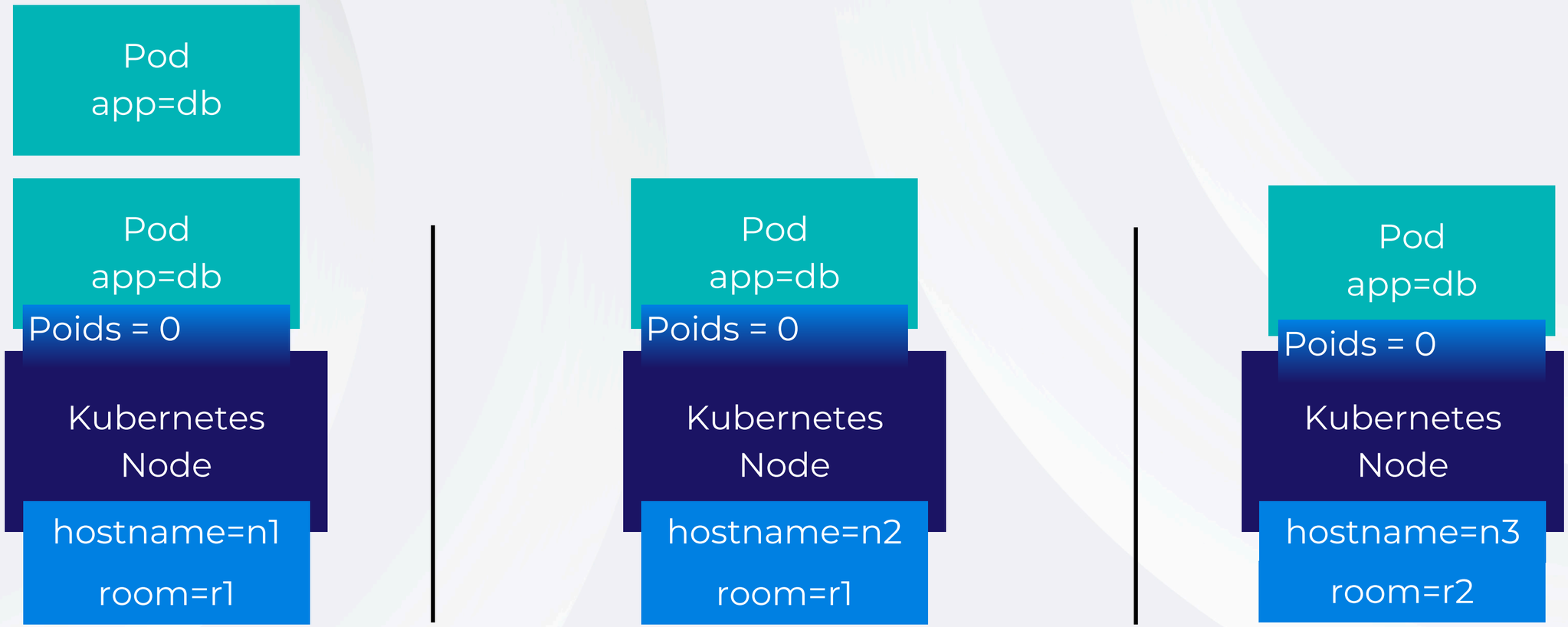
Capacité max = 2pod

Affinity : PodAntiAffinity
Preferred
Topology = hostname
Label = (app = db) = 50



Capacité max = 2pod

Affinity : PodAntiAffinity
Preferred
Topology = hostname
Label = (app = db) = 50



Cas particulier Pod Affinity Cyclique

Capacité max = 2pod

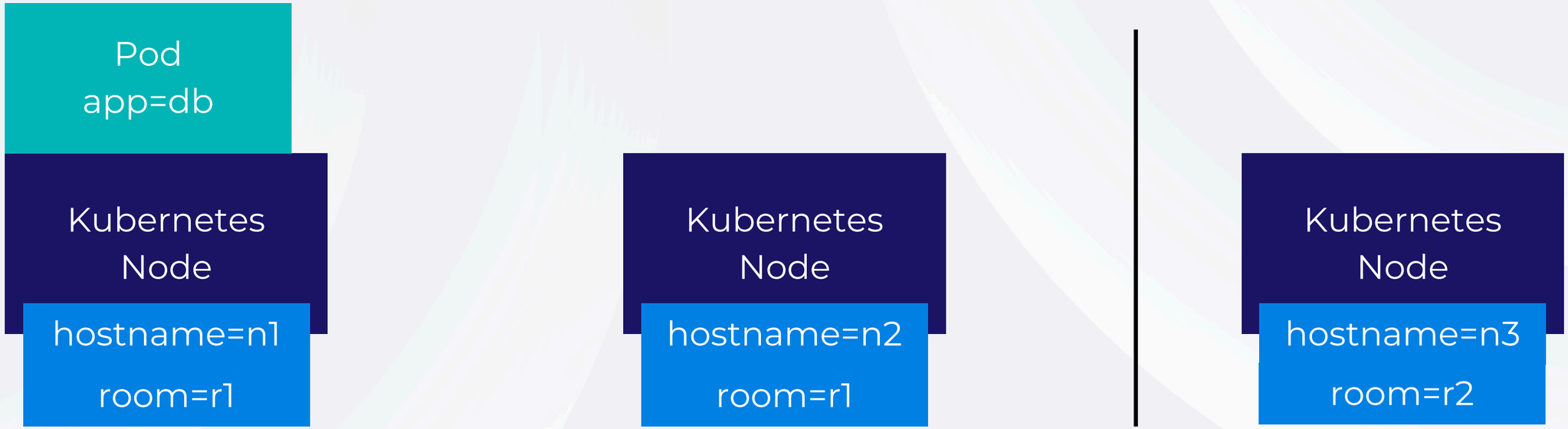
Pod
app=api

Pod
app=api

Pod
app=db

(Pod = API)
Affinity : PodAffinity
Required
Topology = room
Label = (app = db)

(Pod = BDD)
Affinity : PodAffinity
Required
Topology = room
Label = (app = api)



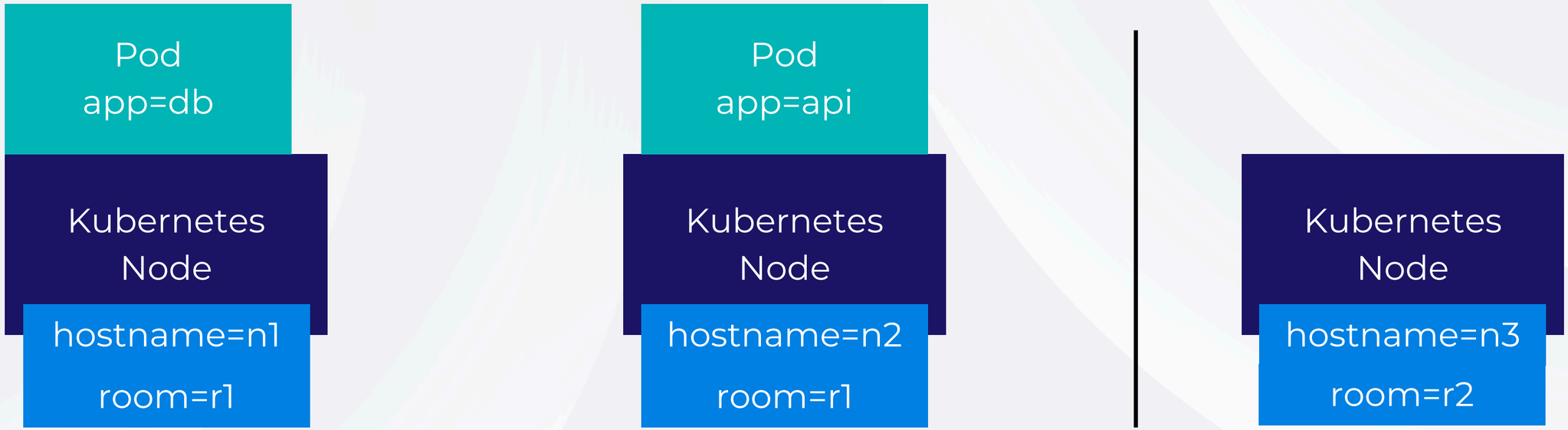
Capacité max = 2pod

(Pod = API)
Affinity : PodAffinity
Required
Topology = room
Label = (app = db)

(Pod = BDD)
Affinity : PodAffinity
Required
Topology = room
Label = (app = api)

Pod
app=api

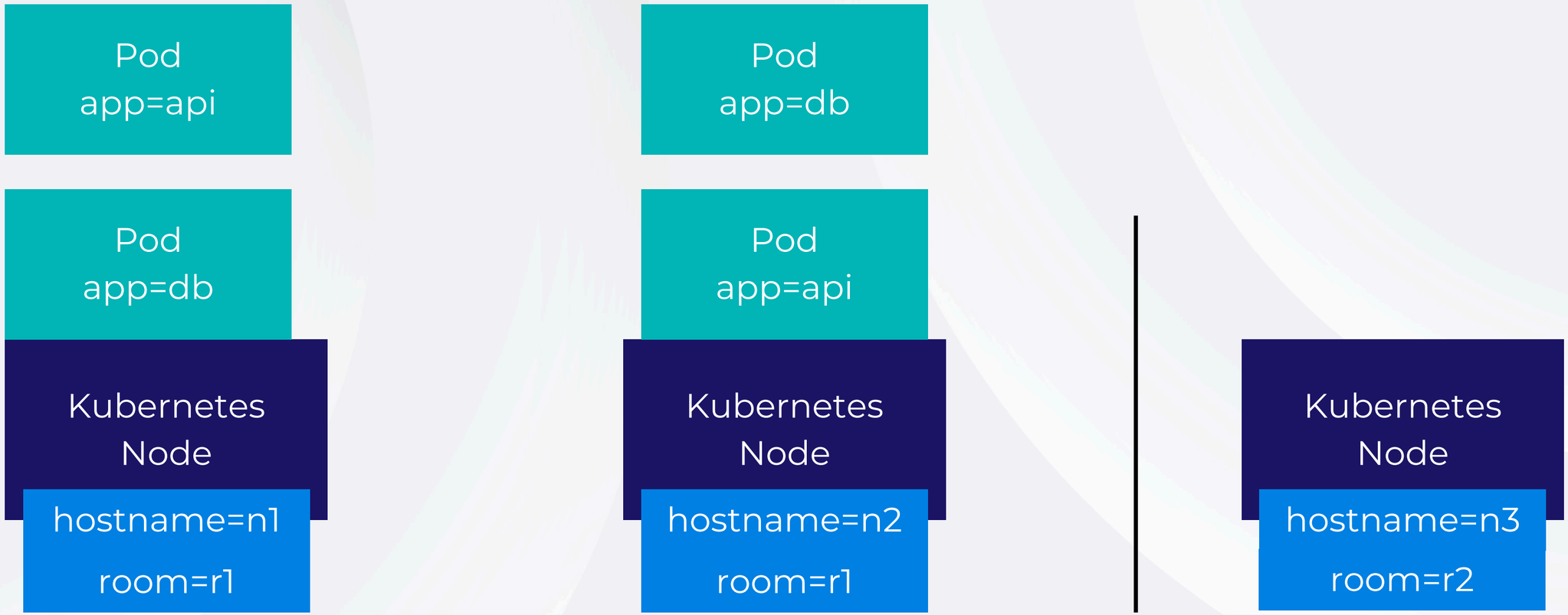
Pod
app=db



Capacité max = 2pod

(Pod = API)
Affinity : PodAffinity
Required
Topology = room
Label = (app = db)

(Pod = BDD)
Affinity : PodAffinity
Required
Topology = room
Label = (app = api)



Capacité max = 2pod

Pod
app=api

Pod
app=db

Pod
app=api

Pod
app=db

(Pod = API)
Affinity : PodAffinity
Required
Topology = room
Label = (app = db)

(Pod = BDD)
Affinity : PodAffinity
Required
Topology = room
Label = (app = api)

Kubernetes
Node
hostname=n1
room=r1

Kubernetes
Node
hostname=n2
room=r1

Kubernetes
Node
hostname=n3
room=r2

Capacité max = 2pod

Pod
app=api
(REJET)

Pod
app=db
(REJET)

Pod
app=api
(REJET)

Pod
app=db
(REJET)

(Pod = API)
Affinity : PodAffinity
Required
Topology = room
Label = (app = db)

(Pod = BDD)
Affinity : PodAffinity
Required
Topology = room
Label = (app = api)

Kubernetes
Node

hostname=n1

room=r1

Kubernetes
Node

hostname=n2

room=r1

Kubernetes
Node

hostname=n3

room=r2

Capacité max = 2pod

Pod
app=api

Pod
app=db

Pod
app=api

Pod
app=db

(Pod = API)
Affinity : PodAffinity
Preferred
Topology = room
Label = (app = db) = 10

(Pod = BDD)
Affinity : PodAffinity
Preferred
Topology = room
Label = (app = api) = 10

Kubernetes
Node
hostname=n1
room=r1

Kubernetes
Node
hostname=n2
room=r1

Kubernetes
Node
hostname=n3
room=r2

Capacité max = 2pod

Pod
app=api

Pod
app=db

Pod
app=api

Pod
app=db

(Pod = API)
Affinity : PodAffinity
Preferred
Topology = room
Label = (app = db) = 10

(Pod = BDD)
Affinity : PodAffinity
Preferred
Topology = room
Label = (app = api) = 10

Poids = 0 (API)
Poids = 0 (BDD)
Kubernetes
Node
hostname=n1
room=r1

Poids = 0 (API)
Poids = 0 (BDD)
Kubernetes
Node
hostname=n2
room=r1

Poids = 0 (API)
Poids = 0 (BDD)
Kubernetes
Node
hostname=n3
room=r2

Capacité max = 2pod

Pod
app=api

Pod
app=api

Pod
app=db

(Pod = API)
Affinity : PodAffinity
Preferred
Topology = room
Label = (app = db) = 10

(Pod = BDD)
Affinity : PodAffinity
Preferred
Topology = room
Label = (app = api) = 10

Pod
app=db

Poids = 10 (API)

Poids = 0 (BDD)

Kubernetes
Node

hostname=n1

room=r1

Poids = 10 (API)

Poids = 0 (BDD)

Kubernetes
Node

hostname=n2

room=r1

Poids = 0 (API)

Poids = 0 (BDD)

Kubernetes
Node

hostname=n3

room=r2

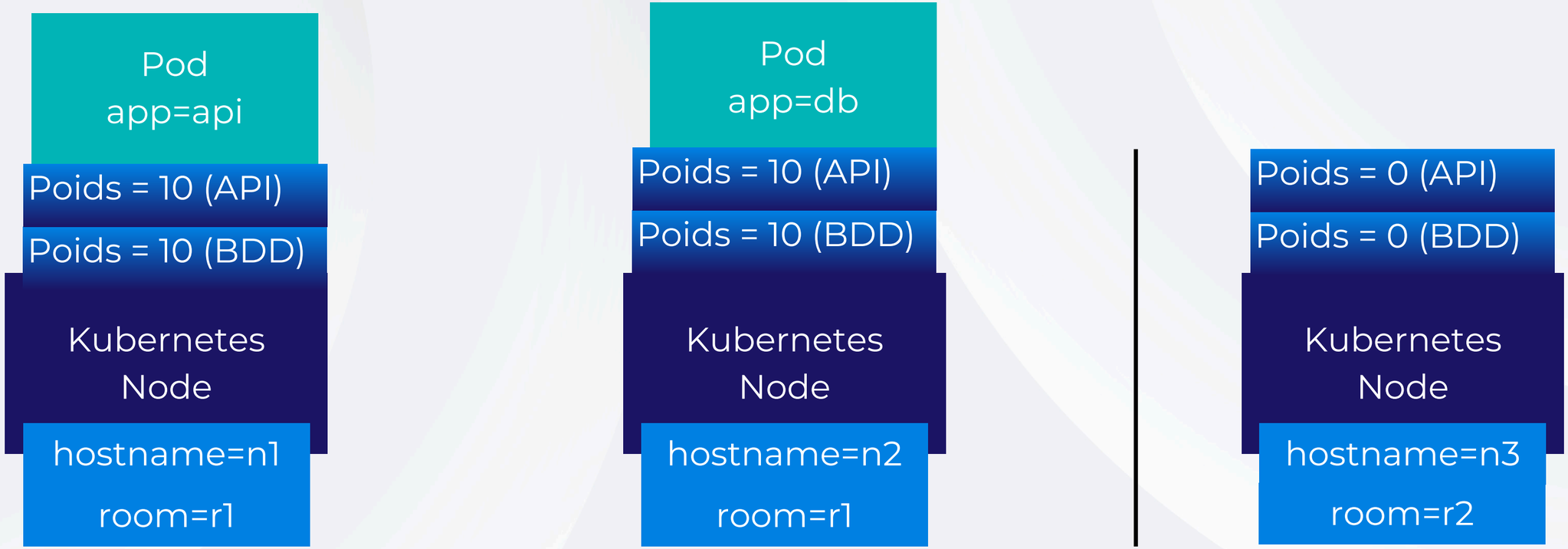
Capacité max = 2pod

Pod
app=api

Pod
app=db

(Pod = API)
Affinity : PodAffinity
Preferred
Topology = room
Label = (app = db) = 10

(Pod = BDD)
Affinity : PodAffinity
Preferred
Topology = room
Label = (app = api) = 10



Capacité max = 2pod

Pod
app=api

Pod
app=db

Pod
app=api

Poids = 10 (API)

Poids = 10 (BDD)

Kubernetes
Node

hostname=n1

room=r1

(Pod = API)
Affinity : PodAffinity
Preferred
Topology = room
Label = (app = db) = 10

(Pod = BDD)
Affinity : PodAffinity
Preferred
Topology = room
Label = (app = api) = 10

Pod
app=db

Poids = 10 (API)

Poids = 10 (BDD)

Kubernetes
Node

hostname=n2

room=r1

Poids = 0 (API)

Poids = 0 (BDD)

Kubernetes
Node

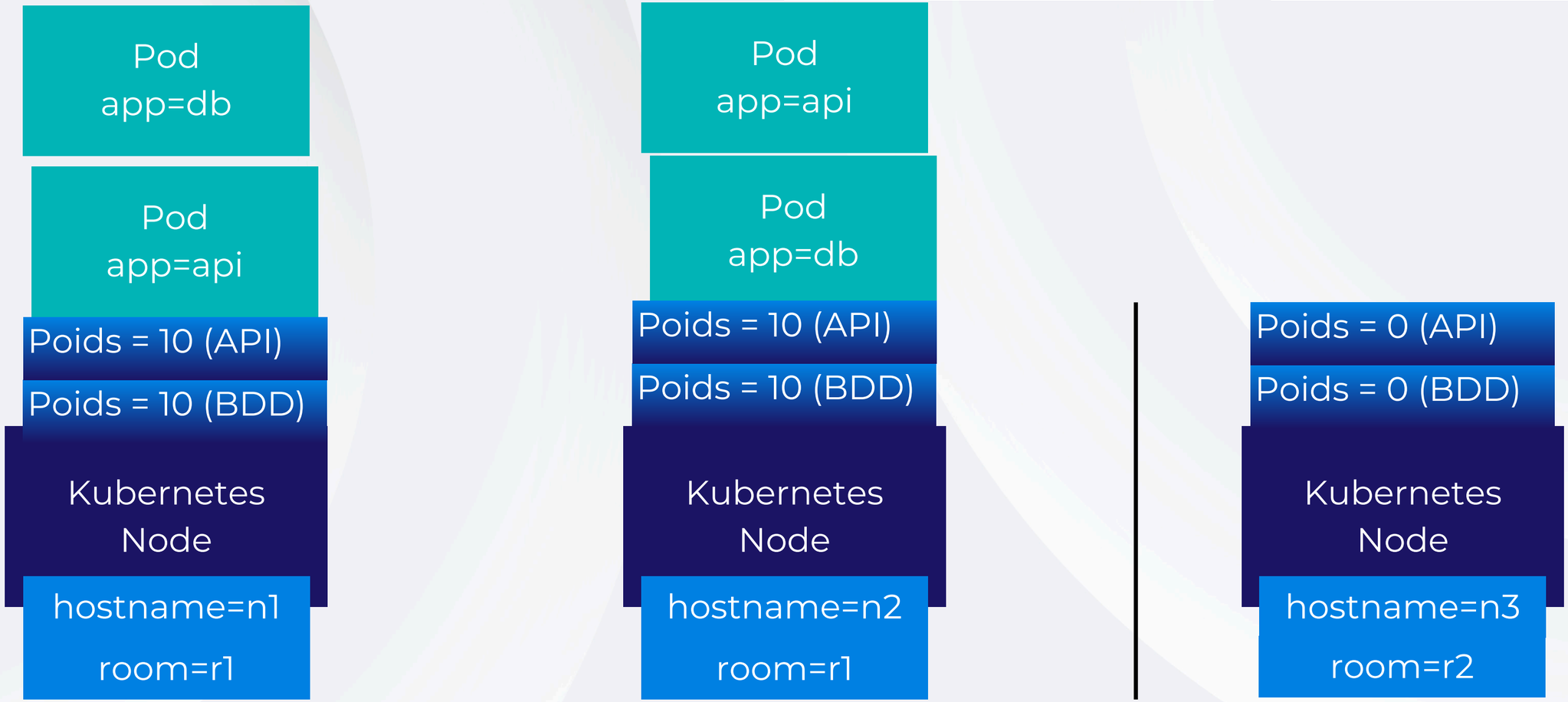
hostname=n3

room=r2

Capacité max = 2pod

(Pod = API)
Affinity : PodAffinity
Preferred
Topology = room
Label = (app = db) = 10

(Pod = BDD)
Affinity : PodAffinity
Preferred
Topology = room
Label = (app = api) = 10



Capacité max = 2pod

Pod
app=api

Pod
app=db

Pod
app=api

Pod
app=db

(Pod = API)
Affinity : PodAffinity
Preferred
Topology = room
Label = (app = db) = 10

(Pod = BDD)
Affinity : PodAffinity
Preferred
Topology = room
Label = (app = api) = 10

Poids = 0 (API)
Poids = 0 (BDD)
Kubernetes
Node
hostname=n1
room=r1

Poids = 0 (API)
Poids = 0 (BDD)
Kubernetes
Node
hostname=n2
room=r1

Poids = 0 (API)
Poids = 0 (BDD)
Kubernetes
Node
hostname=n3
room=r2

Capacité max = 2pod

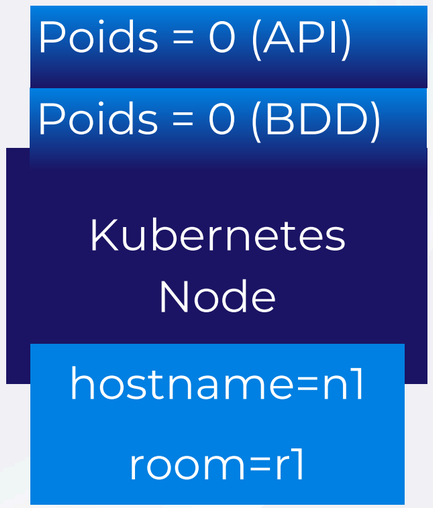
Pod
app=api

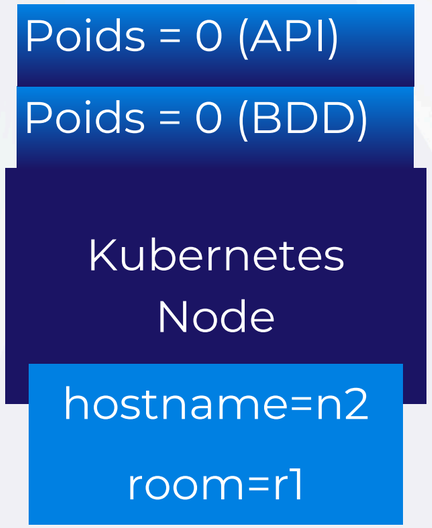
Pod
app=api

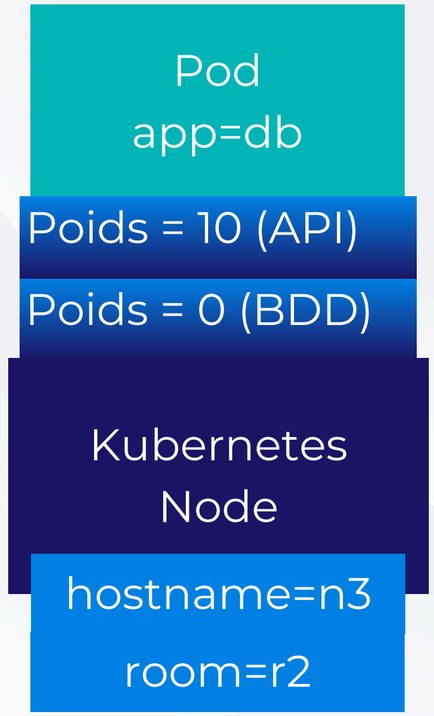
Pod
app=db

(Pod = API)
Affinity : PodAffinity
Preferred
Topology = room
Label = (app = db) = 10

(Pod = BDD)
Affinity : PodAffinity
Preferred
Topology = room
Label = (app = api) = 10







Capacité max = 2pod

Pod
app=api

Pod
app=db

(Pod = API)
Affinity : PodAffinity
Preferred
Topology = room
Label = (app = db) = 10

(Pod = BDD)
Affinity : PodAffinity
Preferred
Topology = room
Label = (app = api) = 10

Pod
app=api

Pod
app=db

Poids = 0 (API)

Poids = 0 (BDD)

Kubernetes
Node

hostname=n1

room=r1

Poids = 0 (API)

Poids = 0 (BDD)

Kubernetes
Node

hostname=n2

room=r1

Poids = 10 (API)

Poids = 10 (BDD)

Kubernetes
Node

hostname=n3

room=r2

Capacité max = 2pod

Pod
app=api

(Pod = API)
Affinity : PodAffinity
Preferred
Topology = room
Label = (app = db) = 10

(Pod = BDD)
Affinity : PodAffinity
Preferred
Topology = room
Label = (app = api) = 10

Pod
app=api

Pod
app=db

Poids = 10 (API)

Poids = 0 (BDD)

Kubernetes
Node

hostname=n1

room=r1

Pod
app=db

Poids = 10 (API)

Poids = 0 (BDD)

Kubernetes
Node

hostname=n2

room=r1

Poids = 10 (API)

Poids = 10 (BDD)

Kubernetes
Node

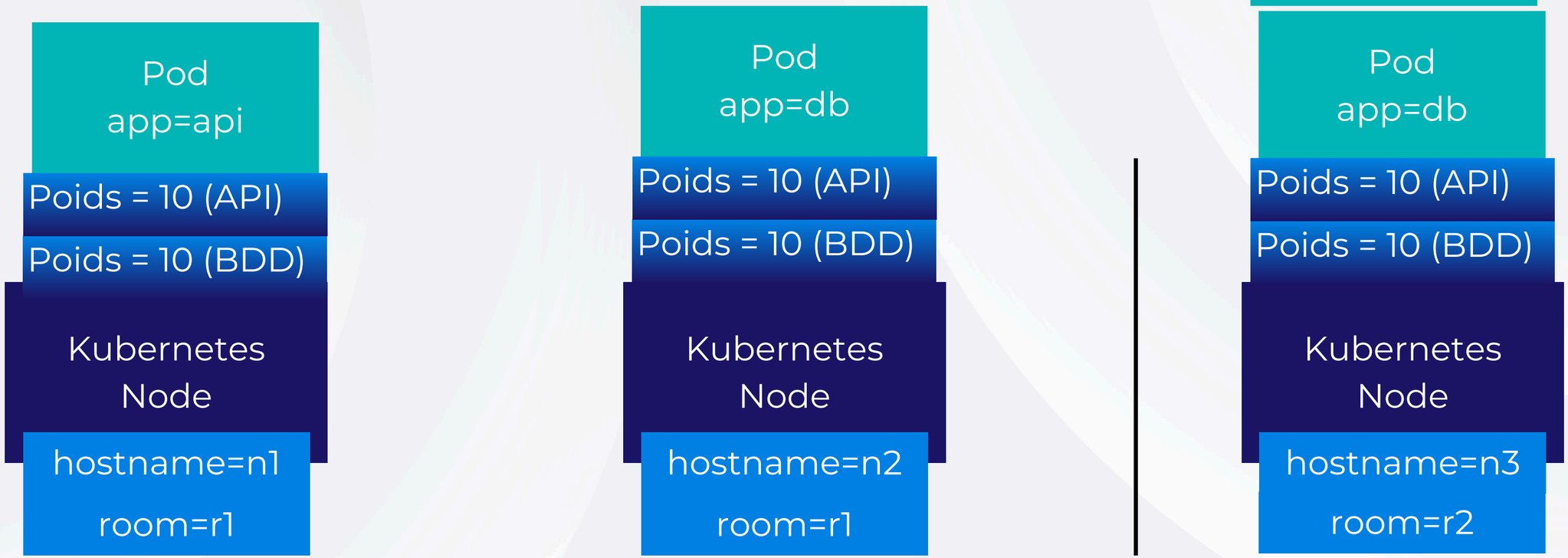
hostname=n3

room=r2

Capacité max = 2pod

(Pod = API)
Affinity : PodAffinity
Preferred
Topology = room
Label = (app = db) = 10

(Pod = BDD)
Affinity : PodAffinity
Preferred
Topology = room
Label = (app = api) = 10



Taint & Tolerence

Capacité max = 2pod

Pod

Pod

Pod

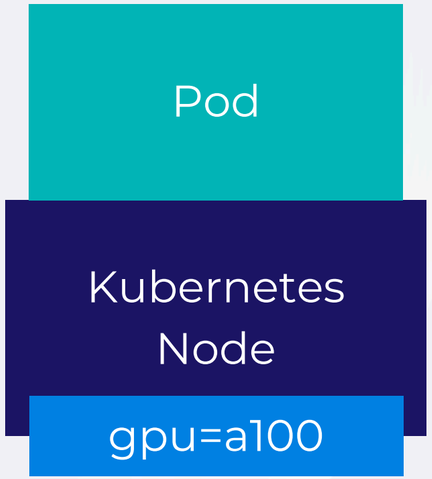
Kubernetes Node

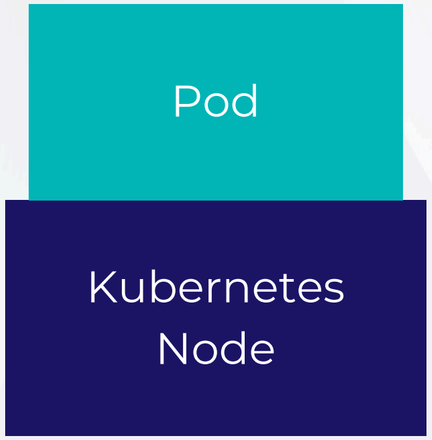
gpu=a100

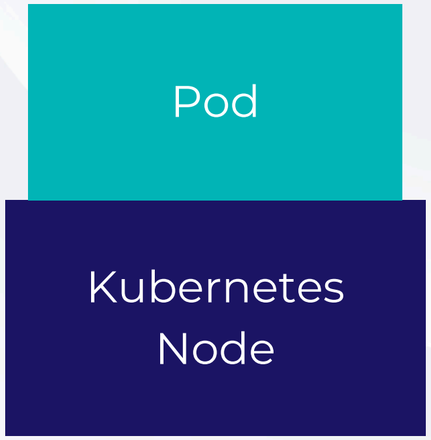
Kubernetes Node

Kubernetes Node

Capacité max = 2pod







Capacité max = 2pod

Pod
Require GPU

Pod

Kubernetes
Node

gpu=a100

Pod
Require GPU

Pod

Kubernetes
Node

Pod

Kubernetes
Node

Capacité max = 2pod

Pod

Pod

Kubernetes Node

Pod

Kubernetes Node

Kubernetes Node

gpu=a100

taint requiregpu,no execute

Capacité max = 2pod

Pod
Require GPU

Pod

Pod
Require GPU

Pod

Pod

Kubernetes
Node

Kubernetes
Node

Kubernetes
Node

gpu=a100

taint requiregpu,no
execute

Capacité max = 2pod

Pod

Pod

Kubernetes Node

Pod Require GPU

Pod

Kubernetes Node

Pod Require GPU

Pod

Kubernetes Node

gpu=a100

taint requiregpu,no execute

Azure

Pour AKS, utiliser les zones de disponibilité Azure (`topology.kubernetes.io/zone`) pour une haute disponibilité réelle.

Effets des taints

NoSchedule : Aucun Pod sans toleration ne peut être schedulé
PreferNoSchedule : Évite de scheduler sans toleration (soft)
NoExecute : Expulse les Pods existants sans toleration

paul.millet@lameduse.fr

LaMeDuSe

UBE : Kubernetes, mise en œuvre

Pourquoi la surveillance est critique

Dans un cluster Kubernetes, les applications sont distribuées sur de nombreux nœuds. Sans outillage adapté, diagnostiquer un problème devient rapidement impossible.

Sources de logs dans Kubernetes

Pods/Conteneurs : stdout/stderr applicatif (kubectl logs, Loki)
Nœuds : kubelet, container runtime (journalctl, node-problem-detector)
Control Plane : apiserver, scheduler, etcd (journalctl, cloud provider)
Audit : Accès API enregistrés (Audit policy K8s)
Events : Events K8s (kubectl get events)

Architecture de collecte

Stdout/Stderr uniquement : Les conteneurs doivent écrire leurs logs sur la sortie standard. Ne jamais écrire dans des fichiers à l'intérieur du conteneur.

Grafana Alloy : L'agent collecteur

Alloy est le successeur de Promtail et Grafana Agent. C'est un collecteur OpenTelemetry-compatible.

Loki : Le stockage de logs

Loki est conçu pour être économique : il n'indexe que les labels, pas le contenu des logs.

Grafana : Visualisation

Datasources à configurer :

Dashboards recommandés :
- **Kubernetes / Compute Resources / Namespace (Pods)** : ID: 17781
- **Loki / Chunks** : monitoring Loki lui-même
- **Node Exporter Full** : ID: 1860

Exemple d'alerte Grafana sur les logs :

Qu'est-ce que Rancher ?

Rancher Manager est une plateforme de gestion multi-cluster Kubernetes. Elle offre une interface graphique complète pour gérer des clusters AKS, EKS, GKE ou on-premise.

Qu'est-ce que Rancher ?

Multi-cluster : Gérer N clusters depuis une interface
Monitoring intégré : Prometheus + Grafana préconfigurés
Alerting : Alertmanager intégré
RBAC centralisé : Gestion des droits cross-clusters
Logging : Intégration Loki/Elasticsearch
Istio : Service mesh intégré

Le problème de l'éphémérité

Par défaut, les logs d'un conteneur sont perdus lorsque le pod est supprimé ou redémarré. Kubernetes conserve les logs des containers terminés sur le nœud, mais uniquement jusqu'à la rotation (généralement 100 Mo ou 5 fichiers).

Stratégies de persistance

1. Sidecar logging (logs dans un volume partagé) :

- Un container a côté du container applicatif

2. DaemonSet logging (recommandé) :

C'est l'approche recommandée : un seul agent par nœud collecte tous les logs des pods.

Commandes kubectl essentielles

Inspecter les pods :

- kubectl describe pod <nom-du-pod>

Inspecter les événements :

- kubectl events

Diagnostiquer les nœuds :

- kubectl describe node <nom-du-node>

États des pods et causes fréquentes

Pending :

- Pas de nœud disponible, quota dépassé
- kubectl describe pod → Events

ImagePullBackOff

- Image introuvable, credentials
- Vérifier image name, imagePullSecret

CrashLoopBackOff

- Application crash, OOM
- kubectl logs --previous

OOMKilled

- Limite mémoire dépassée
- Augmenter `resources.limits.memory

États des pods et causes fréquentes

Evicted :

- Nœud sous pression (disque/mémoire)
- kubectl describe pod → Message

Terminating

- bloqué / Finalizer non résolu
- kubectl patch pod … -p '{"metadata":{"finalizers":null}}'

ContainerCreating long :

- Volume non disponible
- ConfigMap manquant
- kubectl describe pod → Events

paul.millet@lameduse.fr

LaMeDuSe

UBE : Kubernetes, mise en œuvre

LaMeDuSe

Pourquoi étendre l'API Kubernetes ?

Kubernetes est conçu pour être extensible. Les ressources natives (Pods, Deployments, Services...) couvrent les cas généraux, mais les équipes ont souvent besoin de représenter leurs propres concepts métier.

Pourquoi étendre l'API Kubernetes ?

Exemples de cas d'usage :

- Gérer des bases de données : PostgreSQLCluster (CloudNativePG)
- Gérer des certificats TLS : Certificate, Issuer (cert-manager)
- Gérer des règles de monitoring : PrometheusRule (Prometheus Operator)
- Gérer un ingress avancé : IngressRoute (Traefik)
- Gérer un service mesh : VirtualService (Istio)
- Déployer des apps : Application (Argo CD)

Créer une CRD

Une CRD définit un nouveau type de ressource dans l'API Kubernetes.

Qu'est-ce qu'un Operator ?

Un Operator combine une CRD avec un **contrôleur** qui réagit aux changements pour automatiser des tâches opérationnelles complexes.

Qu'est-ce que l'API Aggregation Layer ?

L'API Aggregation Layer permet d'étendre l'API Kubernetes avec des s**erveurs API externes** qui apparaissent comme faisant partie de l'API native.

Différence CRD vs API Aggregation

Complexité	Faible (YAML)	**Élevée (serveur complet)**
Stockage	ETCD K8s	Base propre possible
Validation	OpenAPI Schema	Custom

metrics-server : Exemple d'API agrégée

metrics-server expose l'API metrics.k8s.io utilisée par kubectl top et HPA :

Prometheus Adapter : Custom Metrics

Le Prometheus Adapter permet à l'HPA de scaler sur des métriques custom

paul.millet@lameduse.fr

LaMeDuSe

UBE : Kubernetes, mise en œuvre

LaMeDuSe

Le problème sans Helm

Déployer une application en production nécessite souvent de nombreux fichiers YAML :

Gérer ces fichiers pour **dev**, **staging** et **production** avec des valeurs différentes devient vite ingérable.

Helm : Le gestionnaire de paquets Kubernetes

Ce que Helm apporte :

- Templating : Variables dans les YAML
- Packaging : Un chart = une application complète
- Versioning : Historique des releases
- Rollback : Revenir à une version précédente
- Dépendances : Un chart peut dépendre d'autres charts
- Repositories : Partager et réutiliser des charts

Structure d'un chart

Chart.yaml :

- Définie les informations de l’application

values.yaml :

- Définie les valeurs de l’application

Templates Helm

Les templates utilisent le langage Go template avec des fonctions Sprig.

Commandes essentielles

```
# Ajouter un dépôt
helm repo add bitnami https://charts.bitnami.com/bitnami

# Mettre à jour les dépôts
helm repo update

# Lister les dépôts
helm repo list

# Chercher un chart
helm search repo postgresql
helm search repo grafana/loki --versions
```

Commandes essentielles

```
# Installation simple
helm install my-release bitnami/postgresql \
 --namespace production \
 --create-namespace

# Installation avec valeurs custom
helm install my-app ./my-chart \
 --namespace production \
 --values values-production.yaml \
 --set image.tag=1.5.0 \
 --set replicaCount=3
```

Commandes essentielles

```
# Mise à jour
helm upgrade my-app ./my-chart \
 --namespace production \
 --values values-production.yaml \
 --set image.tag=1.6.0

# Install ou upgrade (idempotent)
helm upgrade --install my-app ./my-chart \
 --namespace production \
 --create-namespace \
 --values values-production.yaml
```

paul.millet@lameduse.fr

LaMeDuSe

UBE : Kubernetes, mise en œuvre

LaMeDuSe

Les couches de sécurité Kubernetes

La sécurité Kubernetes se pense en défense en profondeur : plusieurs couches indépendantes, chacune limitant l'impact d'une compromission.

Les 4Cs de la sécurité Cloud Native

Cloud : Azure RBAC, Private AKS, Azure Firewall
Cluster : K8s RBAC, Admission Controllers, Audit
Conteneur : Non-root, read-only FS, capabilities
Code : Dépendances à jour, SAST, secrets gérés

Kubeconfig et contextes

```
apiVersion: v1
kind: Config
clusters:
 - name: my-aks-cluster
  cluster:
   server: https://my-aks.hcp.westeurope.azmk8s.io:443
   certificate-authority-data: <base64-ca>

users:
 - name: my-user
  user:
   client-certificate-data: <base64-cert>
   client-key-data: <base64-key>

contexts:
 - name: production
  context:
   cluster: my-aks-cluster
   user: my-user
   namespace: production

current-context: production
```

Gestion des accès.

- Pod Security Policies (PSP) : (cluster wide)
  - Définissent des politiques de sécurité pour les pods, comme les capacités, les utilisateurs, et les volumes.
  - PSP est remplacé par les nouvelles fonctionnalités telles que PodSecurity Admission (PSA).
- PodSecurityAdmission (PSA) : (namespace)
  - Remplace PSP et fournit une manière plus flexible et simplifiée de gérer les politiques de sécurité des pods.

Gestion des accès.

```
apiVersion: policy/v1
kind: PodSecurityPolicy
metadata:
  name: example-psp
spec:
  privileged: false
  capabilities:
    add: ["NET_ADMIN"]
  volumes:
  - "configMap"
  - "secret"
```

Gestion des accès.

- Modes de Sécurité PSA :
  - Privileged : Permet des configurations de sécurité moins restrictives, mais ce mode est généralement déconseillé pour la production. (root)
  - Baseline : Applique des règles de sécurité de base, incluant des configurations minimales recommandées pour une sécurité raisonnable.
  - Restricted : Applique des règles de sécurité plus strictes pour des environnements plus sécurisés.

Gestion des accès.

```
apiVersion: policy/v1beta1
kind: PodSecurity
metadata:
  name: example-policy
  namespace: default
spec:
  enforce: baseline
```

Vue d'ensemble

Par défaut, tous les pods d'un cluster peuvent communiquer entre eux. Les **NetworkPolicies** permettent de définir des règles de pare-feu au niveau L3/L4.

Les NetworkPolicies nécessitent un CNI compatible : Calico ou **Cilium** (Flannel ne les supporte pas).

Vue d'ensemble

```
apiVersion: networking.k8s.io/v1
kind: NetworkPolicy
metadata:
 name: default-deny-all
 namespace: production
spec:
 podSelector: {}     # Sélectionne tous les pods
 policyTypes:
  - Ingress
  - Egress
```

paul.millet@lameduse.fr

LaMeDuSe

DOK : Docker, créer et administrer ses conteneurs virtuels d'applications

Ressources des TP

1. TP : Déploiement d'une plateforme de test
2. TP : Déploiement, publication et analyse d'un déploiement
3. TP : Utilisation de deployment
4. TP : Déploiement d'une base de données et d'une application
5. TP : Déploiement de conteneur et gestion de la montée en charge
6. TP : Déploiement d'un cluster

1- TP: Déploiement d'une plateforme de test

Sujet 1 : Utilisation de Minikube

1. Installation de Docker
2. Installation de Minikube
3. Démarer un cluster de 2 noeuds
4. Utiliser Kubectl pour explorer les différentes ressources
5. Lancer la dashboard

2- TP: Déploiement, publication et analyse d'un déploiement.

Sujet 1 : Déploiement

1. Créer un fichier de configuration YAML pour un Deployment.
    a. image : nginxdemos/nginx-hello (8080
2. Déployer l'application avec kubectl apply -f deployment.yaml
3. Exposer le déploiement
    a. kubectl expose deployment <deployment-name> --type=NodePort --port=8080
    b. (optionnel) essayer de ne pas utiliser la commande et de réaliser le yaml soi-même

3- TP: Utilisation de deployment

Sujet 1 : Déploiement

1. Déployer une Application
    a. Écrire un fichier YAML pour le Deployment avec les spécifications de l'application (image : nginxdemos/nginx-hello) (port : 8080)
2. Mise à Jour de l'Application : Modifier l'image de l'application
    a. image: nginx
    b. (port : 80)
3. Scalabilité et Gestion des Réplicas
    a. Utiliser la commande scale pour modifier le nombre de répliques.
    b. Observer la montée ou la descente en charge via kubectl get pods.
4. Rollback d'un Déploiement :
    a. utiliser rollout undo pour revenir à la version stable précédente.

4- TP: Déploiement d'une base de données et d'une application

Informations

- variable pour la base de donnée
  - MYSQL_ROOT_PASSWORD=somewordpress
  - MYSQL_DATABASE=wordpress
  - MYSQL_USER=wordpress
  - MYSQL_PASSWORD=wordpress
- variable pour le serveur wordpress
  - WORDPRESS_DB_HOST=db
    - note : changer cette variable pour qu'elle corresponde au nom de votre service
  - WORDPRESS_DB_USER=wordpress
  - WORDPRESS_DB_PASSWORD=wordpress
  - WORDPRESS_DB_NAME=wordpress
- Images
  - mariadb:10.6.4-focal
    - pour la base de donnée (port : 3306)
  - wordpress:6.6.1-php8.2-apache
    - pour l'application (port : 80)

4- TP: Déploiement d'une base de données et d'une application

Etape 1 : Déploiement de la base de donnée

- Création du fichier de déploiement
  - Utiliser un StatefulSet
  - Créer un service (cluster IP)
- Application du fichier de déploiement
- Vérification du déploiement

4- TP: Déploiement d'une base de données et d'une application

Etape 2 : Déploiement de l’application

- Création du fichier de déploiement
  - Utiliser un deployment
  - Créer un service (cluster IP)
- Application du fichier de déploiement
- Vérification du déploiement

5- TP: Déploiement de conteneur et gestion de la montée en charge

Sujet 1 : Partie 2 : Déploiement de l’application

- Création du fichier de déploiement
  - Utiliser un deployment
  - Créer un service (type: NodePort)
- Création d’une politique d’autoscaling
- Test avec un outil comme “hey” pour faire monter la charge
  - hey -z 5m -c 1000 http://<node-ip>:<node-port>

6- TP: Déploiement d'un cluster

Sujet 1 :

- Création des machines virtuelles
- Installation des prérequis
- Mise en place du noeud maitre
  - Démarrage du service
  - Récupérer le token
- Mise en place du noeud worker
  - Configurer le noeud avec le token récupéré
  - Démarage du service
- Mise en place du noeud d’aministration
  - Installer kubectl
  - Mettre le fichier kubeconfig

6- TP: Déploiement d'un cluster

Sujet 2 : Utiliser le cluster

- déployer wordpress et la base de donnée sur le cluster

paul.millet@lameduse.fr

LaMeDuSe

UBE : Kubernetes, mise en œuvre

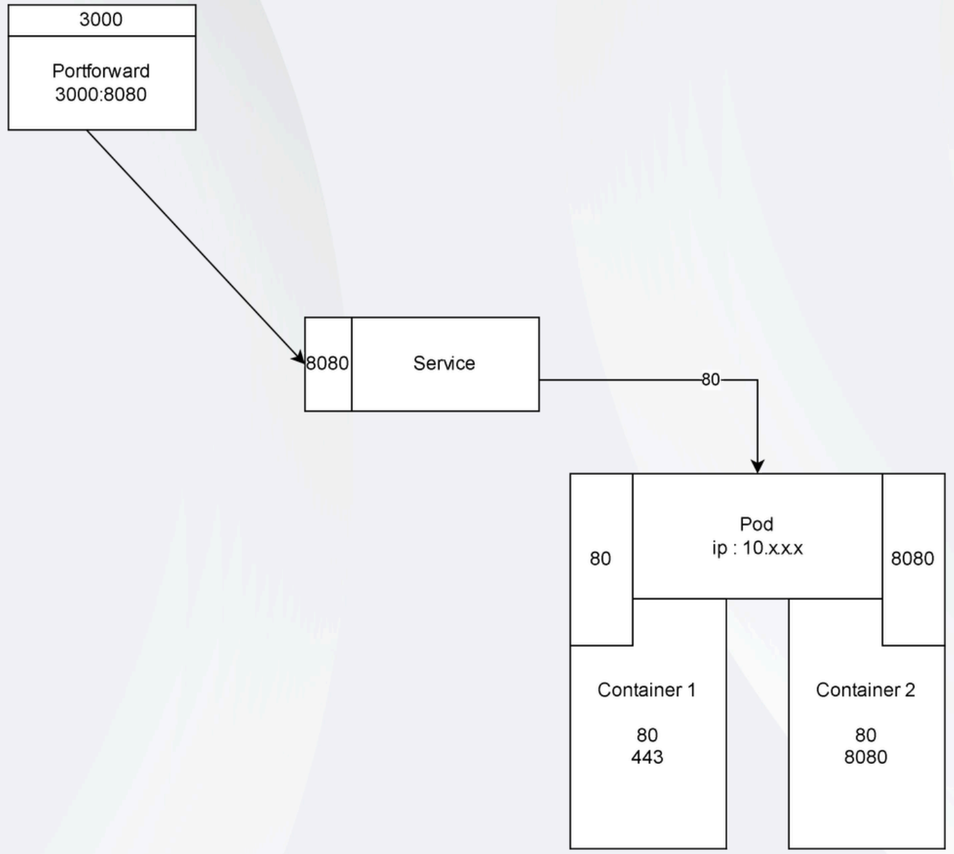
LaMeDuSe

Introduction à Kubernetes

Annexes :

- A. Schemas diverses

Introduction à Kubernetes



Introduction à Kubernetes