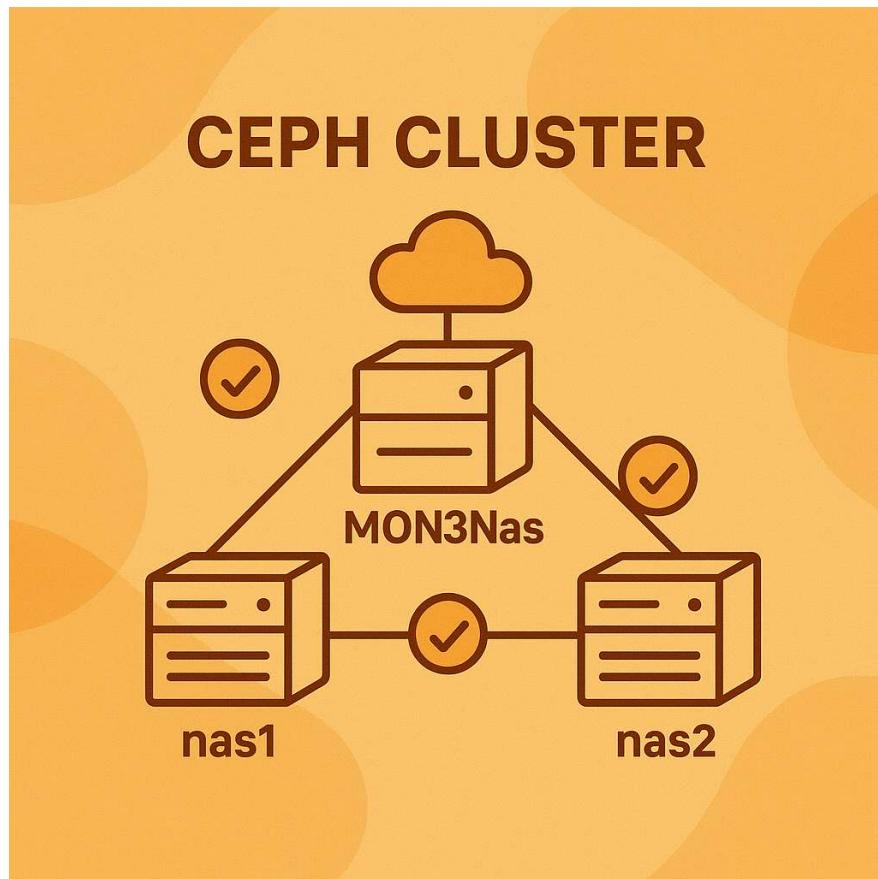

MISE EN PLACE D'UN CLUSTER CEPH AVEC 3 MACHINES – EPREUVE E6 BTS SIO SISR



CONTEXTE DU PROJET - MISE EN PLACE D'UNE INFRASTRUCTURE DE STOCKAGE HAUTE DISPONIBILITE

Dans le cadre du besoin du cabinet comptable Perlier d'assurer la continuité de service et la sécurité des données des clients, nous avons conçu et déployé une infrastructure de stockage redondante, basée sur **Ceph**, une solution distribuée, résiliente et évolutive.

L'environnement repose sur un **cluster haute disponibilité**, réparti sur trois nœuds :

- **NAS1 (192.168.30.13)** – Nœud principal avec 500 Go de stockage (OSD)
- **NAS2 (192.168.30.14)** – Nœud secondaire avec 500 Go de stockage (OSD)
- **MON3Nas (192.168.30.21)** – Nœud moniteur dédié, utilisé comme **arbitre** (tiebreaker) pour garantir le **quorum** même en cas de panne d'un des deux nœuds de stockage

L'objectif était de mettre à disposition une **solution de stockage centralisée, redondante et tolérante aux pannes**, capable d'assurer :

- Une **accessibilité continue des données**, même en cas de défaillance d'un nœud
- Une **gestion simplifiée** grâce à l'outil intégré **cephadm**
- Une **scalabilité** permettant de rajouter facilement des nœuds ou du stockage

Ce projet s'inscrit dans une démarche de professionnalisation orientée **qualité de service, résilience et modernisation de l'infrastructure**.

1. PRE-REQUIS ET PREPARATION

Les machines doivent être sur un réseau commun, synchronisées via NTP (ou configuration de l'horodatage manuelle), et disposées ainsi :

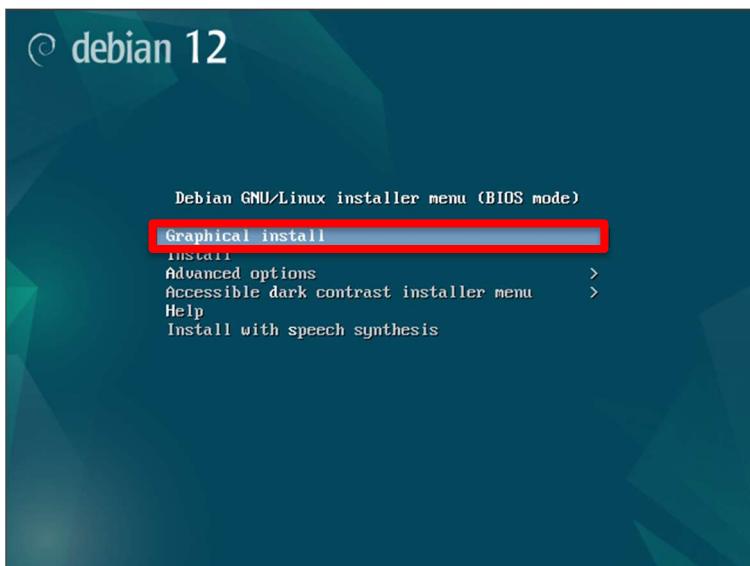
- Debian 12 installé sur chaque machine
- Paquets nécessaires : podman ou docker, curl, chrony, openssh-server, sudo

Installation de Debian 12 :

Télécharger le .iso afin d'installer Debian 12 sur les deux machines

→ <https://www.debian.org/index.fr.html>

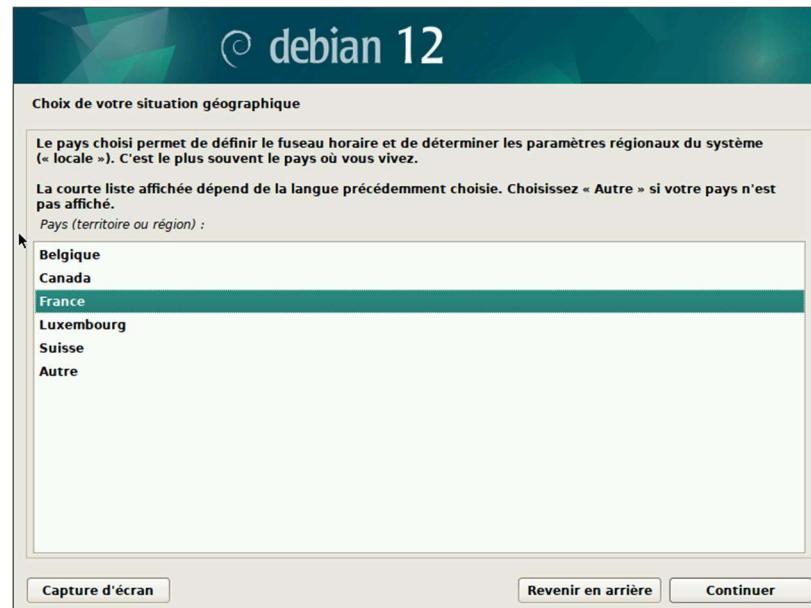
Lancer le .iso



Sélectionner « Graphical install »



Sélectionner la langue



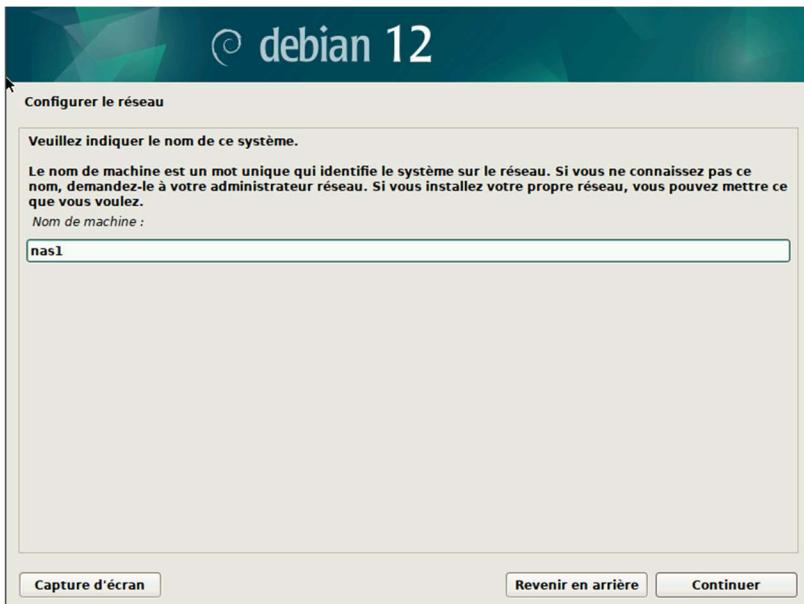
Sélectionner le pays



Sélectionner la disposition du clavier



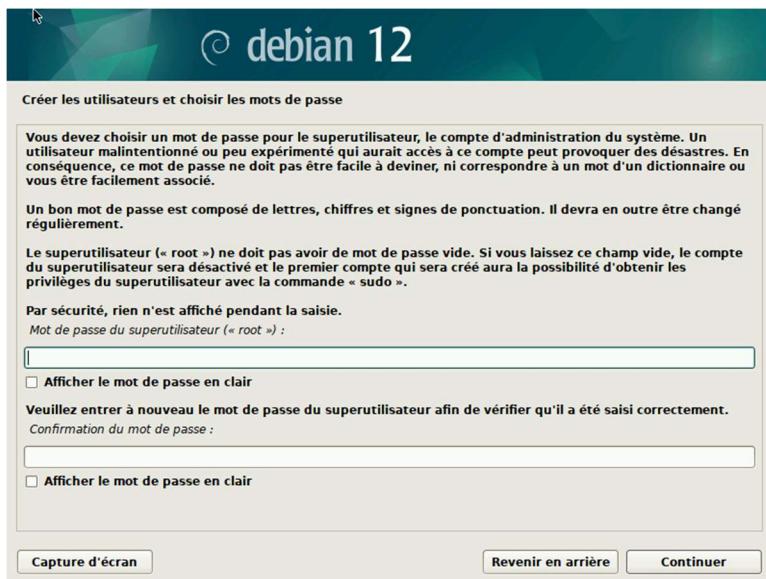
Si un DHCP est configuré l'interface principale récupère une configuration IP



Choisir le hostname



Laisser le nom de domaine vide à cette étape



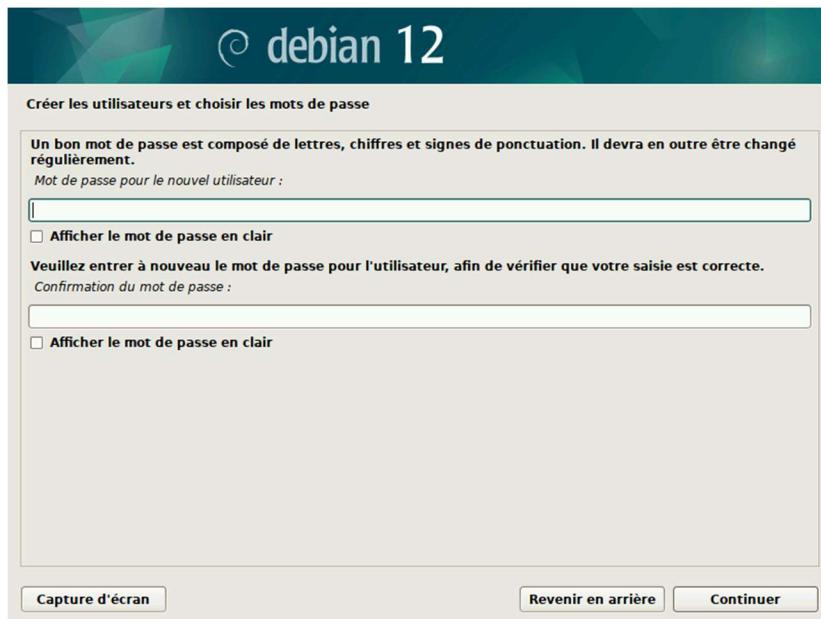
Entre un mot de passe pour le super-utilisateur **root**



Créer un utilisateur (sudoer) → ex : user ou ceph



Lui donner un nom d'utilisateur



Définir un mot de passe pour l'utilisateur



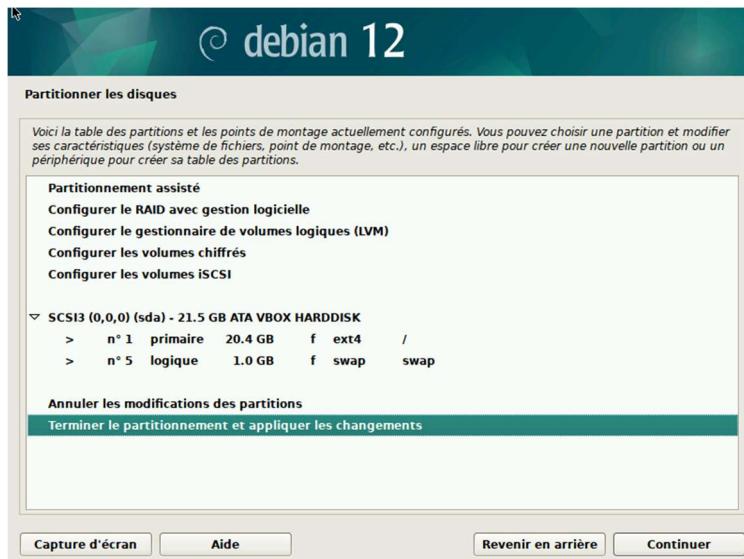
Sélectionner « Assisté – utiliser un disque entier »



Sélectionner le disque pour l'installation de l'OS



Sélectionner « Tout dans une seule partition »



Terminer le partitionnement



Sélectionner « Oui »



Les partitions se formattent



Debian s'installe

Sur chaque machine :

Mettre à jour le système :

```
apt update && apt upgrade -y
```

Installer les dépendances nécessaires

```
apt install -y curl gnupg2 lsb-release podman
```

2. INSTALLATION DE CEPHADM (SUR NAS1 UNIQUEMENT)

***curl --silent --remote-name
<https://download.ceph.com/keys/release.asc>***

```
root@nas1:~# curl --silent --remote-name https://download.ceph.com/keys/release.asc
```

***gpg --no-default-keyring --keyring ./ceph-release.gpg --import
release.asc***

```
root@nas1:~# gpg --no-default-keyring --keyring ./ceph-release.gpg --import release.asc
gpg: le trousseau local « ./ceph-release.gpg » a été créé
gpg: répertoire « /root/.gnupg » créé
gpg: /root/.gnupg/trustdb.gpg : base de confiance créée
gpg: clef E84AC2C0460F3994 : clef publique « Ceph.com (release key) <security@ceph.com> » importée
gpg:      Quantité totale traitée : 1
gpg:                           importées : 1
```

***gpg --no-default-keyring --keyring ./ceph-release.gpg --export >
/etc/apt/trusted.gpg.d/ceph-release.gpg***

```
root@nas1:~# gpg --no-default-keyring --keyring ./ceph-release.gpg --export > /etc/apt/trusted.gpg.d/ceph-release.gpg
```

echo deb https://download.ceph.com/debian-reef/ \$(lsb_release -sc) main > /etc/apt/sources.list.d/ceph.list

```
root@nas1:~# echo deb https://download.ceph.com/debian-reef/ $(lsb_release -sc) main > /etc/apt/sources.list.d/ceph.list
```

apt update && apt install -y cephadm

```
root@nas1:~# apt update && apt install -y cephadm
Atteint :1 http://security.debian.org/debian-security bookworm-security InRelease
Atteint :2 http://deb.debian.org/debian bookworm InRelease
Atteint :3 http://deb.debian.org/debian bookworm-updates InRelease
Réception de :4 https://download.ceph.com/debian-reef bookworm InRelease [8 547 B]
Réception de :5 https://download.ceph.com/debian-reef bookworm/main amd64 Packages [17,6 kB]
26,2 ko réceptionnés en 1s (22,0 ko/s)
Lecture des listes de paquets... Fait
Construction de l'arbre des dépendances... Fait
Lecture des informations d'état... Fait
Tous les paquets sont à jour.
Lecture des listes de paquets... Fait
Construction de l'arbre des dépendances... Fait
Lecture des informations d'état... Fait
Les NOUVEAUX paquets suivants seront installés :
  cephadm
0 mis à jour, 1 nouvellement installés, 0 à enlever et 0 non mis à jour.
Il est nécessaire de prendre 330 ko dans les archives.
Après cette opération, 370 ko d'espace disque supplémentaires seront utilisés.
```

3. BOOTSTRAP DU CLUSTER SUR NAS1

```
cephadm bootstrap --mon-ip 192.168.30.13 --initial-dashboard-user admin --initial-dashboard-password admin
```

```
root@nas1:~# cephadm bootstrap --mon-ip 192.168.30.13 --initial-dashboard-user admin --initial-dashboard-password admin
Creating directory /etc/ceph for ceph.conf
Verifying podman|docker is present...
Verifying lvm2 is present...
Verifying time synchronization is in place...
Unit systemd-timesyncd.service is enabled and running
Repeating the final host check...
podman (/usr/bin/podman) version 4.3.1 is present
systemctl is present
lvcreate is present
Unit systemd-timesyncd.service is enabled and running
Host looks OK
Cluster fsid: c79d6alc-0e0f-11f0-9d88-a85e45c6070d
Verifying IP 192.168.30.13 port 3300 ...
Verifying IP 192.168.30.13 port 6789 ...
Mon IP '192.168.30.13' is in CIDR network '192.168.30.0/24'
Mon IP '192.168.30.13' is in CIDR network '192.168.30.0/24'
Internal network (--cluster-network) has not been provided, OSD replication will default to the public_network
Pulling container image quay.io/ceph/ceph:v18...
Ceph version: ceph version 18.2.4 (e7ad5345525c7aa95470c26863873b581076945d) reef (stable)
Extracting ceph user uid/gid from container image...
Creating initial keys...
Creating initial monmap...
Creating mon...
Waiting for mon to start...
Waiting for mon...
mon is available
Assimilating anything we can from ceph.conf...
Generating new minimal ceph.conf...
Restarting the monitor...
Setting public_network to 192.168.30.0/24 in mon config section
Wrote config to /etc/ceph/ceph.conf
Wrote keyring to /etc/ceph/ceph.client.admin.keyring
Creating mgr...
Verifying port 0.0.0.0:9283 ...
Verifying port 0.0.0.0:8765 ...
Verifying port 0.0.0.0:8443 ...
Waiting for mgr to start...
Waiting for mgr...
mgr not available, waiting (1/15)...
mgr not available, waiting (2/15)...
mgr not available, waiting (3/15)...
mgr not available, waiting (4/15)...
mgr not available, waiting (5/15)...
mgr is available
Enabling cephadm module...
Waiting for the mgr to restart...
Waiting for mgr epoch 5...
mgr epoch 5 is available
Setting orchestrator backend to cephadm...
Generating ssh key...
Wrote public SSH key to /etc/ceph/ceph.pub
Adding key to root@localhost authorized_keys...
Adding host nas1...
Deploying mon service with default placement...
Deploying mgr service with default placement...
Deploying crash service with default placement...
Deploying ceph-exporter service with default placement...
Deploying prometheus service with default placement...
Deploying grafana service with default placement...
Deploying node-exporter service with default placement...
Deploying alertmanager service with default placement...
Enabling the dashboard module...
Waiting for the mgr to restart...
Waiting for mgr epoch 9...
mgr epoch 9 is available
Generating a dashboard self-signed certificate...
Creating initial admin user...
Fetching dashboard port number...
Ceph Dashboard is now available at:
    URL: https://nas1:8443/
    User: admin
    Password: admin
Enabling client.admin keyring and conf on hosts with "admin" label
Saving cluster configuration to /var/lib/ceph/c79d6alc-0e0f-11f0-9d88-a85e45c6070d/config directory
Enabling autotune for osd_memory_target
You can access the Ceph CLI as following in case of multi-cluster or non-default config:
```

4. AJOUT DES AUTRES NŒUDS AU CLUSTER

Copier la clé SSH générée par cephadm sur chaque hôte :

```
ssh-copy-id -f -i /etc/ceph/ceph.pub root@nas2
ssh-copy-id -f -i /etc/ceph/ceph.pub root@MON3Nas
```

```
ceph orch host add nas2 192.168.30.14
```

```
ceph orch host add MON3Nas 192.168.30.21
```

5. DEPLOIEMENT DES SERVICES CEPH

Déployer les moniteurs :

```
ceph orch apply mon 'nas1,nas2,MON3Nas'
```

Déployer les OSD :

1. Identifier les disques à utiliser comme OSD :

```
lsblk
```

2. Ajouter l'OSD sur nas1 (par exemple sur /dev/sdb) :

```
ceph orch daemon add osd nas1:/dev/sdb
```

3. Ajouter l'OSD sur nas2 (par exemple sur /dev/sdb) :

```
ceph orch daemon add osd nas2:/dev/sdb
```

Une fois ces commandes exécutées, Ceph configure automatiquement les disques, crée les partitions nécessaires et déploie les services OSD correspondants. Il est possible de vérifier leur bon fonctionnement avec la commande :

```
ceph osd status
```

6. DEPLOIEMENT DU MDS (METADATA SERVER)

Le service MDS (Metadata Server) est requis pour le système de fichiers CephFS. Il gère les métadonnées (noms, hiérarchie, permissions) tandis que les OSDs gèrent les blocs de données.

- Voici les étapes de déploiement :

1. Créer un volume CephFS :

```
ceph fs volume create cephfs
```

2. Vérifier la création :

```
ceph fs ls
```

3. Déployer un MDS sur le cluster :

```
ceph orch apply mds cephfs --placement='1 host'
```

Pour assurer une haute disponibilité, ajouter un MDS en standby :

```
ceph orch apply mds cephfs --placement='2 hosts'  
ceph orch apply mds cephfs --placement='nas1,nas2'
```

5. Vérifier l'état des MDS :

```
ceph mds stat
```

7. ACCES A L'INTERFACE WEB DU CLUSTER

L'interface Web d'administration du cluster est accessible via : <https://192.168.30.13:8443> (ou <https://192.168.30.14:8443> si nas2 devient le manager du cluster)

8. VÉRIFICATION DU CLUSTER

```
ceph -s  
ceph osd status  
ceph health detail
```

9. CONCLUSION

Ce projet permet de mettre en œuvre un cluster de stockage tolérant aux pannes en utilisant les bonnes pratiques modernes de déploiement via cephadm. MON3Nas, bien qu'il ne dispose pas de stockage, permet de maintenir un quorum même si un des nœuds OSD tombe en panne.

10. ACCES CEPHFS DEPUIS UN CLIENT WINDOWS VIA SMB

Ceph ne fournit pas de service SMB (partage Windows) nativement.

Pour permettre à un client Windows d'accéder au CephFS, il est nécessaire de passer par un serveur Samba intermédiaire.

1. Installer Samba sur une machine Linux (ex: nas1 ou une VM intermédiaire)

apt update

apt install samba ceph-common -y

2. Monter le CephFS localement sur cette machine

Créer un point de montage :

mkdir -p /mnt/cephfs

Monter le FS avec l'utilisateur admin :

mount -t ceph 192.168.30.13:6789:/ /mnt/cephfs -o name=admin,secretfile=/etc/ceph/admin.secret

Le fichier /etc/ceph/admin.secret doit contenir la clé de l'utilisateur admin :

echo 'AQB...cle_base64...' > /etc/ceph/admin.secret

chmod 600 /etc/ceph/admin.secret

3. Partager le dossier CephFS avec Samba

Modifier */etc/samba/smb.conf* et ajouter à la fin :

[cephfs]

path = /mnt/cephfs

browsable = yes

read only = no

guest ok = yes

force user = root

Redémarrer Samba :

systemctl restart smbd

4. Depuis Windows

Dans l'explorateur de fichiers :

\|192.168.30.13\cephfs

Ce chemin donne accès au contenu de CephFS via SMB.

11. CONCLUSION DU PROJET

Ce projet m'a permis de concevoir une solution de stockage redondante, sécurisée et hautement disponible, adaptée à un environnement professionnel. Après plusieurs essais sur des solutions prêtées à l'emploi comme **TrueNAS** puis **Synology**, j'ai fait le choix de basculer vers une approche **entièrement maîtrisée et open source** en utilisant **Ceph** et **Cephadm**.

J'ai ainsi pu bâtir une infrastructure robuste reposant sur :

- Un **cluster Ceph haute disponibilité** avec quorum
 - Une **bascule automatique** via Keepalived (IP flottante)
 - Une **réPLICATION MariaDB** en mode actif/passif
 - Une interface **Nextcloud** synchronisée entre les deux nœuds
 - Une **supervision SNMP** avec intégration dans **Observium**
-

Enrichissement personnel et technique

Ce projet m'a permis de :

- Me confronter à un **environnement de production réaliste**
 - Appréhender des notions avancées comme le **quorum**, le **failover**, la **résilience**, ou encore la **consistance des données**
 - Automatiser des tâches critiques pour garantir une **continuité de service**
-

Perspectives d'amélioration

Dans un cadre professionnel, plusieurs axes pourraient encore renforcer cette solution :

- Déploiement de **Prometheus/Grafana** pour des métriques temps réel
 - Mise en place d'un **outil de failover avancé** comme **Orchestrator**
 - Ajout d'un **reverse proxy** (HAProxy/NGINX) pour une gestion fine des accès
 - Intégration avec un **annuaire LDAP** pour une gestion centralisée des utilisateurs
-

Application professionnelle

Ce type d'infrastructure est parfaitement adapté pour :

Une PME ou un établissement souhaitant **garantir la disponibilité de ses données**, tout en conservant **la main sur sa solution** et en évitant des coûts de licence élevés.