



Retour d'expérience sur PROXMOX et HA

une solution de virtualisation et de disponibilité de services

Richard Ferrere
Romain Pacé

Retour d'expérience au Laboratoire de Mathématiques de Besançon

Mathrice Grenoble - 8 au 10 avril 2014

Retour
d'expérience sur
PROXMOX et HA

Richard Ferrere
Romain Pacé

Un peu de
terminologie

PROXMOX et DRBD ?

La haute disponibilité

Infrastructure du
laboratoire

Notre laboratoire à
Besançon

Notre cahier des charges

Système et topologie cluster
retenus

Installation des
différentes briques

IPMI, réseau

Proxmox, DRBD et LVM

Briques liées à la HA

Défaillances et HA

Capacité réseau...

Matérielle, énergétique...

Démo

Conclusion

Liens et références
internet

Logs liés à la haute
disponibilité

Plan de la présentation

Un peu de terminologie

PROXMOX et DRBD ?

La haute disponibilité

Infrastructure du laboratoire

Notre laboratoire à

Besançon

Notre cahier des charges

Système et topologie

cluster retenus

Installation des différentes briques

IPMI, réseau

Proxmox, DRBD et LVM

Briques liées à la HA

Défaillances et HA

Capacité réseau..

Matérielle, énergétique...

Démo

Conclusion

Liens et références internet

Retour
d'expérience sur
PROXMOX et HA

Richard Ferrere
Romain Pacé

Un peu de terminologie

PROXMOX et DRBD ?

La haute disponibilité

Infrastructure du laboratoire

Notre laboratoire à
Besançon

Notre cahier des charges

Système et topologie cluster
retenus

Installation des différentes briques

IPMI, réseau

Proxmox, DRBD et LVM

Briques liées à la HA

Défaillances et HA

Capacité réseau..

Matérielle, énergétique...

Démo

Conclusion

Liens et références internet

Logs liés à la haute disponibilité

Un peu de terminologie

Un peu de terminologie

PROXMOX et DRBD ?

La haute disponibilité

Infrastructure du laboratoire

Notre laboratoire à
Besançon

Notre cahier des charges

Système et topologie cluster
retenus

Installation des différentes briques

IPMI, réseau

Proxmox, DRBD et LVM

Briques liées à la HA

Défaillances et HA

Capacité réseau..

Matérielle, énergétique...

Démo

Conclusion

Liens et références internet

Logs liés à la haute disponibilité

Server View Datancenter

Search Summary Options Storage Backup Users Groups Pools Permissions Roles Authentication HA

Type	Description	Disk usage	Memory usage	CPU usage	Uptime
node	boole	65.2%	39.8%	2.0% of 32CPUs	120 days 06:57:24
node	cardan	1.8%	7.1%	3.3% of 32CPUs	120 days 06:38:14

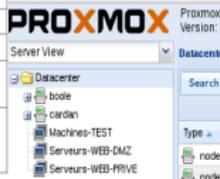
- ▶ Distribution linux basée sur DEBIAN en 64 bits (en version 3.2 actuellement !)
- ▶ Virtualisation : modules KVM et support OpenVZ
- ▶ Outils de sauvegarde, de restauration et bien d'autres
- ▶ Interface web de gestion centralisée, accès à l'adresse `https://NomServeur:8006/`
- ▶ Fonction de clustering avec regroupement de noeuds
- ▶ Mode haute disponibilité
- ▶ Et également une version payante possible avec assistance et service

Proxmox VE Subscriptions

	COMMUNITY	BASIC	STANDARD	PREMIUM
Stable updates via Enterprise repository	Yes	Yes	Yes	Yes
Technical support	via community forum	via Proxmox Customer Portal	via Proxmox Customer Portal	via Proxmox Customer Portal
Support tickets included	None *	3 per year	10 per year	Unlimited
Response time	n/a	1 business day	1 business day	1 business day
Remote login via SSH	No	No	Yes	Yes
Qualified for the purchase of additional support tickets	No	Yes	Yes	n/a
Pricing	€ 4,16 per month & CPU-socket	€ 16,58 per month & CPU-socket	€ 33,17 per month & CPU-socket	€ 66,33 per month & CPU-socket

* Support via public support forum

<https://172.16.1.80:8006/#v>



Retour
d'expérience sur
PROXMOX et HA

Richard Ferrere
Romain Pacé

Un peu de
terminologie

PROXMOX et DRBD?

La haute disponibilité

Infrastructure du
laboratoire

Notre laboratoire à
Besançon

Notre cahier des charges

Système et topologie cluster
retenus

Installation des
différentes briques

IPMI, réseau

Proxmox, DRBD et LVM

Briques liées à la HA

Défaillances et HA

Capacité réseau..

Matérielle, énergétique...

Démo

Conclusion

Liens et références
internet

Logs liés à la haute
disponibilité

Beaucoup d'actions possibles avec Proxmox!

The image shows a screenshot of the Proxmox VE interface. On the left, a tree view shows a VM named '106 (VM 338)' selected, with a context menu open. The menu options include: Start, Migrate, Shutdown, Stop, Clone, Convert to template, and Console. On the right, a table lists recent tasks performed on this VM.

Time	Start	End	Status	User	VM
Mar 26 14:18:07	Mar 26 14:59:24		boote	root@pve	VMCT 106 - Console
Mar 26 10:11:38	Mar 26 11:34:54		boote	root@pve	VMCT 106 - Console
Mar 26 09:18:21	Mar 26 09:20:45		boote	root@pve	VMCT 106 - Console
25 08:19:28	Mar 25 09:21:54		boote	root@pve	VMCT 106 - Console
25 08:12:25	Mar 25 08:12:35		boote	root@pve	VMCT 106 - Console
24 17:27:42	Mar 24 18:30:22		boote	root@pve	VMCT 106 - Console
24 17:25:11	Mar 24 17:27:02		boote	root@pve	VMCT 106 - Console
23 11:42:37	Mar 23 12:43:18		boote	root@pve	VMCT 106 - Console
21 09:42:53	Mar 21 10:25:28		boote	root@pve	VMCT 106 - Console
21 09:37:22	Mar 21 09:29:17		boote	root@pve	VMCT 106 - Console
20 18:40:39	Mar 20 18:53:37		boote	root@pve	VMCT 106 - Console
20 18:39:48	Mar 20 18:39:51		boote	root@pve	VM 106 - Start
20 18:02:03	Mar 20 18:34:59		boote	root@pve	VM 106 - Clone
Mar 20 17:52:36	Mar 20 18:01:17		boote	root@pve	VMCT 106 - Console

- ▶ Démarrage/Arrêt, Migration, Clonage, Conversion en template
- ▶ Bonne vision réelle sur les 2 noeuds et l'état des VMs avec graphiques de charge CPU, MEM
- ▶ Sauvegardes : Vzdump des VMs
- ▶ Riche en historique (tâches et logs) → historique des actions dans /var/log/ et ../pve/tasks
- ▶ Assistants de création de VMs avec configuration



Retour
d'expérience sur
PROXMOX et HA

Richard Ferrere
Romain Pacé

Un peu de
terminologie

PROXMOX et DRBD?

La haute disponibilité

Infrastructure du
laboratoire

Notre laboratoire à
Besançon

Notre cahier de charges

Système et topologie cluster
retenus

Installation des
différentes briques

IPMI, réseau

Proxmox, DRBD et LVM

Briques liées à la HA

Défaillances et HA

Capacité réseau...

Matérielle, énergétique...

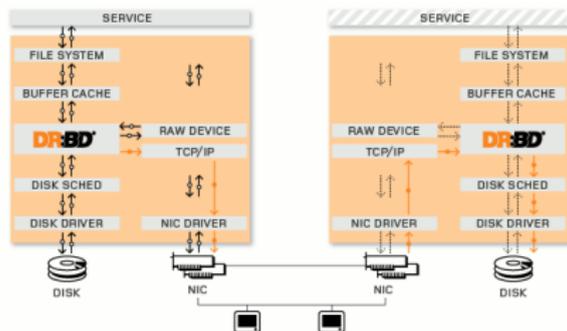
Démo

Conclusion

Liens et références
internet

Logs liés à la haute
disponibilité

DRBD : Distributed Replicated Block Device



- ▶ Outil de réplication en mode bloc par le réseau
- ▶ RAID1 sur IP ou miroir entre 2 noeuds du cluster
- ▶ Synchronisation permanente synchrone et asynchrone des données
- ▶ Réplication de disques/volumes LVM au travers du réseau
- ▶ Module noyau et driver du système de fichiers

Le mode par défaut est :

mode primary/secondary avec système de fichiers GFS ou OCFS2

Notre choix :

mode primary/primary avec système de fichiers EXT3/4

Retour
d'expérience sur
PROXMOX et HA

Richard Ferrere
Romain Pacé

Un peu de
terminologie

PROXMOX et DRBD?

La haute disponibilité

Infrastructure du
laboratoire

Notre laboratoire à
Besançon

Notre cahier des charges

Système et topologie cluster
retenus

Installation des
différentes briques

IPMI, réseau

Proxmox, DRBD et LVM

Briques liées à la HA

Défaillances et HA

Capacité réseau...

Matérielle, énergétique...

Démo

Conclusion

Liens et références
internet

Logs liés à la haute
disponibilité

HA : High Availability

Retour
d'expérience sur
PROXMOX et HA

Richard Ferrere
Romain Pacé

Une analyse à faire :

1. Evaluer la criticité des services en production
2. Evaluer le coût d'une défaillance ou d'une indisponibilité **coût en € = P (probabilité) x C (coût horaire journalier) x T (temps)**
3. Evaluer les impacts...

Agir sur la formule !

- ▶ PCA (Plan de de continuité de l'activité, PRA (Plan de reprise de l'activité)
- ▶ HA = PCA (moyens matériels, logiciels) + PRA (procédures)
- ▶ Disponibilité : données stockées et services en production
- ▶ Assurer la continuité des services aux utilisateurs en cas de défaillance

Un peu de
terminologie

PROXMOX et DRBD ?

La haute disponibilité

Infrastructure du
laboratoire

Notre laboratoire à
Besançon

Notre cahier des charges

Système et topologie cluster
retenus

Installation des
différentes briques

IPMI, réseau

Proxmox, DRBD et LVM

Briques liées à la HA

Défaillances et HA

Capacité réseau..

Matérielle, énergétique...

Démo

Conclusion

Liens et références
internet

Logs liés à la haute
disponibilité

HA : High Availability - suite...

Retour
d'expérience sur
PROXMOX et HA

Richard Ferrere
Romain Pacé

Ce qui revient à :

1. Détecter automatiquement la défaillance (moyens mis en oeuvre)
2. Réagir rapidement par un ou plusieurs PRA (procédures adaptées et testées)

Pour notre laboratoire avec Proxmox et DRBD

→ Migrer rapidement et automatiquement les services indisponibles d'un noeud du cluster à l'autre

Un peu de
terminologie

PROXMOX et DRBD ?

La haute disponibilité

Infrastructure du
laboratoire

Notre laboratoire à
Besançon

Notre cahier des charges

Système et topologie cluster
retenus

Installation des
différentes briques

IPMI, réseau

Proxmox, DRBD et LVM

Briques liées à la HA

Défaillances et HA

Capacité réseau..

Matérielle, énergétique...

Démo

Conclusion

Liens et références
internet

Logs liés à la haute
disponibilité

Principe de fonctionnement de la HA avec Proxmox

Prérequis

- ▶ 2 noeuds minimum et disque de quorum
- ▶ Un stockage répliqué avec DRBD
- ▶ Dispositif de *fencing* : l'**agent fence**

Plusieurs démons communiquent entre eux

- **Corosync** : Coeur du système, communication entre les services
- **CMAN** : Gère les communications entre les noeuds
- **Disque de quorum** : Agit en tant qu'arbitre
- **RGManager** : Répartit et redistribue automatiquement les services

En résumé...

- ▶ Redémarrer automatiquement sur l'autre noeud les VMS du noeud défectueux
- ▶ S'assurer que le noeud défectueux ne soit plus présent dans le cluster !

Retour
d'expérience sur
PROXMOX et HA

Richard Ferrere
Romain Pacé

Un peu de
terminologie

PROXMOX et DRBD?

La haute disponibilité

Infrastructure du
laboratoire

Notre laboratoire à
Besançon

Notre cahier des charges

Système et topologie cluster
retenus

Installation des
différentes briques

IPMI, réseau

Proxmox, DRBD et LVM

Briques liées à la HA

Défaillances et HA

Capacité réseau..

Matérielle, énergétique...

Démo

Conclusion

Liens et références
internet

Logs liés à la haute
disponibilité

Retour
d'expérience sur
PROXMOX et HA

Richard Ferrere
Romain Pacé

Infrastructure du laboratoire

Un peu de
terminologie

PROXMOX et DRBD?

La haute disponibilité

Infrastructure du
laboratoire

Notre laboratoire à
Besançon

Notre cahier des charges

Système et topologie cluster
retenus

Installation des
différentes briques

IPMI, réseau

Proxmox, DRBD et LVM

Briques liées à la HA

Défaillances et HA

Capacité réseau..

Matérielle, énergétique...

Démo

Conclusion

Liens et références
internet

Logs liés à la haute
disponibilité

Quelques chiffres...

- ▶ 5 équipes de recherche
- ▶ environ 100 personnels : 70 enseignants-chercheurs, plus de 30 doctorants et des invités
- ▶ Parc informatique hétérogène avec plus de 100 postes :
 - 1 pare-feu gérant 3 zones réseaux séparées (privée, DMZ et invités)
 - 11 serveurs physiques
 - 40 postes sous Windows
 - 60 postes Linux (dont clients légers)
- ▶ différents services communs regroupés sur le cluster : DHCP, DNS, Web, Samba,...

Un peu de terminologie

PROXMOX et DRBD ?

La haute disponibilité

Infrastructure du laboratoire

Notre laboratoire à Besançon

Notre cahier des charges

Système et topologie cluster retenus

Installation des différentes briques

IPMI, réseau

Proxmox, DRBD et LVM

Briques liées à la HA

Défaillances et HA

Capacité réseau...

Matérielle, énergétique...

Démo

Conclusion

Liens et références internet

Logs liés à la haute disponibilité

Grandes lignes de l'existant

- ▶ Parc de serveurs vieillissant et nombreuses pannes de disques
- ▶ Pannes répétées de la climatisation de la salle machine
- ▶ Sauvegardes sur lecteurs/bandes non fiables !
- ▶ Budget de renouvellement limité à 13 K€(pas 1 euro de plus!)

Objectifs à plusieurs niveaux :

- ▶ Disposer en permanence des services et des données sur les différentes zones réseau (privée et DMZ)
- ▶ Faire de la haute disponibilité autant que possible
- ▶ Intégrer chacun des services sur une infrastructure de virtualisation
- ▶ Bénéficier d'une souplesse et d'une facilité pour la gestion et l'administration
- ▶ Disposer d'une configuration matérielle et logicielle redondante
- ▶ Rester sur une solution libre (pas d'appel d'offres à passer !)
- ▶ Respecter l'environnement en utilisant moins d'énergie
- ▶ Déployer des services à la demande et déléguer des tâches d'administration

Retour
d'expérience sur
PROXMOX et HA

Richard Ferrere
Romain Pacé

Un peu de
terminologie

PROXMOX et DRBD ?

La haute disponibilité

Infrastructure du
laboratoire

Notre laboratoire à
Besançon

Notre cahier des charges

Système et topologie cluster
retenus

Installation des
différentes briques

IPMI, réseau

Proxmox, DRBD et LVM

Briques liées à la HA

Défaillances et HA

Capacité réseau..

Matérielle, énergétique...

Démo

Conclusion

Liens et références
internet

Logs liés à la haute
disponibilité

Matériel retenu

- ▶ Achat de 2 serveurs frontaux DELL R720XD Xeon E5-2650 2Ghz comprenant :
 1. 32 coeurs et 64Go de mémoire et 7 disques attenants
 2. RAID 1 avec 2 disques très rapides de 300Go (système Proxmox)
 3. RAID 5 avec 5 disques de 3To (stockage et volume DRBD)
 4. 6 ports réseaux : 4 ports Gigas + 2 ports 10 Gigas
- ▶ Pas de baie SAN de stockage à part entière : séparation des flux trop coûteuse ! 42 K€

Système retenu

- ▶ Proxmox et DRBD
- ▶ Mode DRBD primary/primary ou dual-primary
- ▶ Système de fichier sur le RAID 1 en EXT3
- ▶ DRBD et LVM

Retour
d'expérience sur
PROXMOX et HA

Richard Ferrere
Romain Pacé

Un peu de
terminologie

PROXMOX et DRBD ?

La haute disponibilité

Infrastructure du
laboratoire

Notre laboratoire à
Besançon

Notre cahier des charges

Système et topologie cluster
retenus

Installation des
différentes briques

IPMI, réseau

Proxmox, DRBD et LVM

Briques liées à la HA

Défaillances et HA

Capacité réseau..

Matérielle, énergétique...

Démo

Conclusion

Liens et références
internet

Logs liés à la haute
disponibilité

Topologie retenue et en production avec les noeuds Boole et Cardan

- ▶ Coeur de réseau en Gigabits (ressource du CRI)
- ▶ architecture des 2 serveurs connectée directement (noeud Cardan lié au noeud Boole)
- ▶ Le client ne voit toujours qu'un noeud du cluster (soit Boole, soit Cardan)

PROXMOX Proxmox Virtual Environment
Version: 3.1-24/060bd5a6 You are logged

Server View Datacenter

Search Summary Options Storage Backup Users Groups Pools Permissions Roles Authentication HA

Type	Description	Disk usage	Memory usage	CPU usage	Uptime
node	boole	65.2%	39.8%	2.0% of 32CPUs	120 days 06:57:24
node	cardan	1.8%	7.1%	3.3% of 32CPUs	120 days 06:38:14

Un peu de
terminologie

PROXMOX et DRBD ?

La haute disponibilité

Infrastructure du
laboratoire

Notre laboratoire à
Besançon

Notre cahier de charges

Système et topologie cluster
retenus

Installation des
différentes briques

IPMI, réseau

Proxmox, DRBD et LVM

Briques liées à la HA

Défaillances et HA

Capacité réseau..

Matérielle, énergétique...

Démo

Conclusion

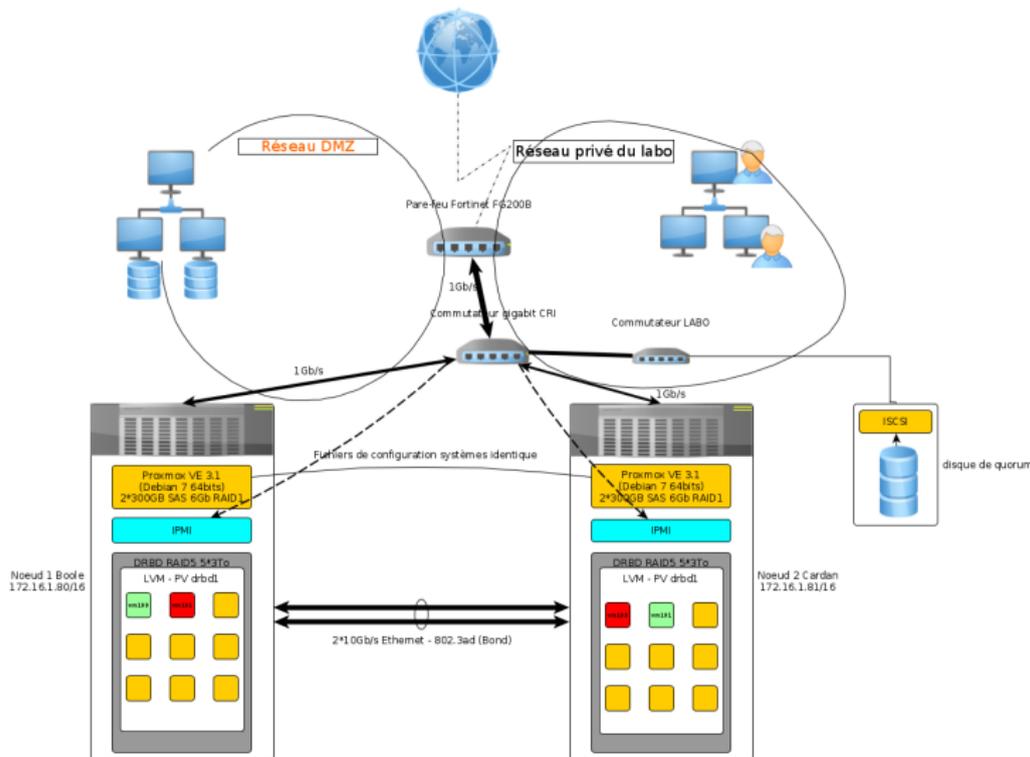
Liens et références
internet

Logs liés à la haute
disponibilité

Schéma de la topologie cluster : 2 noeuds Proxmox avec HA en DRBD

Retour
d'expérience sur
PROXMOX et HA

Richard Ferrere
Romain Pacé



Un peu de
terminologie

PROXMOX et DRBD ?

La haute disponibilité

Infrastructure du
laboratoire

Notre laboratoire à
Besançon

Notre cahier des charges

**Système et topologie cluster
retenus**

Installation des
différentes briques

IPMI, réseau

Proxmox, DRBD et LVM

Briques liées à la HA

Défaillances et HA

Capacité réseau...

Matérielle, énergétique...

Démo

Conclusion

Liens et références
internet

Logs liés à la haute
disponibilité

Système et topologie retenus - suite...

Retour
d'expérience sur
PROXMOX et HA

Richard Ferrere
Romain Pacé

Timing : déroulement et évolutions

1. Avril 2013 : démarrage du projet
2. Mai - septembre 2013 : tests
3. Octobre 2013 : production, intégration des services (KVM)
4. Janvier - avril 2014 : plate-forme de tests dédiée HA
5. Mai -juin 2014 : basculement de la HA en production
6. Juillet - octobre 2014 : serveur de secours Proxmox indépendant, écritures de scripts de surveillance
7. Novembre 2014 : achat de disques de 3To → volume DRBD supplémentaire (r1)

Un peu de
terminologie

PROXMOX et DRBD?
La haute disponibilité

Infrastructure du
laboratoire

Notre laboratoire à
Besançon

Notre cahier des charges

**Système et topologie cluster
retenus**

Installation des
différentes briques

IPMI, réseau

Proxmox, DRBD et LVM

Briques liées à la HA

Défaillances et HA

Capacité réseau..

Matérielle,énergétique...

Démo

Conclusion

Liens et références
internet

Logs liés à la haute
disponibilité

Retour
d'expérience sur
PROXMOX et HA

Richard Ferrere
Romain Pacé

Un peu de
terminologie

PROXMOX et DRBD ?

La haute disponibilité

Infrastructure du
laboratoire

Notre laboratoire à
Besançon

Notre cahier des charges

**Système et topologie cluster
retenus**

Installation des
différentes briques

IPMI, réseau

Proxmox, DRBD et LVM

Briques liées à la HA

Défaillances et HA

Capacité réseau..

Matérielle, énergétique...

Démo

Conclusion

Liens et références
internet

Logs liés à la haute
disponibilité

Installation des différentes briques

Briques IPMI et réseau

- ▶ Activation de l'IPMI dans le BIOS
- ▶ Activation de la mise sous tension automatique après une coupure électrique

- ▶ Initialiser l'IPMI (paquet ipmitool) et tester le fonctionnement :

```
# ipmitool lan set 1 ipaddr 172.16.8.110  
# ipmitool lan set 1 netmask 255.255.0.0  
# ipmitool user set password 2 secret  
# ipmitool lan set 1 access on  
# ipmitool -H 172.16.8.210 -U root -P secret power  
reset
```

Un peu de
terminologie

PROXMOX et DRBD?

La haute disponibilité

Infrastructure du
laboratoire

Notre laboratoire à
Besançon

Notre cahier des charges

Système et topologie cluster
retenus

Installation des
différentes briques

IPMI, réseau

Proxmox, DRBD et LVM

Briques liées à la HA

Défaillances et HA

Capacité réseau..

Matérielle, énergétique...

Démo

Conclusion

Liens et références
internet

Logs liés à la haute
disponibilité

Briques IPMI et réseau

```
auto lo
iface lo inet loopback
```

```
iface eth0 inet manual
iface eth4 inet manual
iface eth5 inet manual
```

```
auto vubr0
iface vubr0 inet static
    address 172.16.1.80
    netmask 255.255.0.0
    gateway 172.16.0.1
    bridge_ports eth0
    bridge_stp off
    bridge_fd 0
```

```
auto bond0
iface bond0 inet manual
    slaves eth4 eth5
    bond-mode 802.3ad
    bond-miimon 100
    bond-downdelay 200
    bond-updelay 200
```

```
auto vubr1
iface vubr1 inet static
    address 10.10.10.1
    netmask 255.255.255.0
    bridge_ports bond0
    bridge_stp off
    bridge_fd 0
```

Retour
d'expérience sur
PROXMOX et HA

Richard Ferrere
Romain Pacé

Un peu de
terminologie

PROXMOX et DRBD?

La haute disponibilité

Infrastructure du
laboratoire

Notre laboratoire à
Besançon

Notre cahier des charges

Système et topologie cluster
retenus

Installation des
différentes briques

IPMI, réseau

Proxmox, DRBD et LVM

Briques liées à la HA

Défaillances et HA

Capacité réseau..

Matérielle, énergétique...

Démo

Conclusion

Liens et références
internet

Logs liés à la haute
disponibilité

Briques Proxmox et DRBD

- ▶ à l'identique sur les 2 noeuds :
 - Configuration réseau : bond 2*10Gb pour la synchro DRBD
 - Création du volume DRBD à l'identique sur les 2 noeuds
 - Définition de la ressource DRBD dans `/etc/drbd.d/r0.res`
 - Modification du taux de transfert pour le bond DRBD
`/etc/drbd.d/global_common.conf`
-> 800M (augmentation de la bande passante)
- ▶ Toujours sur les 2 noeuds, 3 commandes :
`# service drbd start`
`# drbdadm create-md r0`
`# drbdadm up r0`
- ▶ Sur le 1er noeud, synchronisation initiale avec la commande :
`# drbdadm -- --overwrite-data-of-peer primary r0`
==> De 4 à 5h pour 8To !!

Un peu de
terminologie

PROXMOX et DRBD ?

La haute disponibilité

Infrastructure du
laboratoire

Notre laboratoire à
Besançon

Notre cahier des charges

Système et topologie cluster
retenus

Installation des
différentes briques

IPMI, réseau

Proxmox, DRBD et LVM

Briques liées à la HA

Défaillances et HA

Capacité réseau..

Matérielle, énergétique...

Démo

Conclusion

Liens et références
internet

Logs liés à la haute
disponibilité

Briques Proxmox, DRBD et LVM suite...

- ▶ 3 cafés plus tard, 2 commandes à faire sur les 2 noeuds :

```
# service drbd stop  
# service drbd start
```

==> Nécessaire pour être en primary/primary

==> *# cat /proc/drbd* doit afficher **ro :Primary/Primary**

- ▶ Configurer LVM pour être au dessus de DRBD puis créer le PV, à faire sur les 2 noeuds :

```
# By default we accept every block device:  
filter = [ "r|/dev/sdb1|", "r|/dev/disk/|", "r|/dev/block/|", "a/.*/" ]
```

```
# pvcreate /dev/drbd0
```

- ▶ Création du VG sur un noeud :

```
# vgcreate drbdvg /dev/drbd0
```

- ▶ Ajouter le VG dans Proxmox et lui assigner le stockage des images

Un peu de
terminologie

PROXMOX et DRBD?

La haute disponibilité

Infrastructure du
laboratoire

Notre laboratoire à
Besançon

Notre cahier des charges

Système et topologie cluster
retenus

Installation des
différentes briques

IPMI, réseau

Proxmox, DRBD et LVM

Briques liées à la HA

Défaillances et HA

Capacité réseau..

Matérielle, énergétique...

Démo

Conclusion

Liens et références
internet

Logs liés à la haute
disponibilité

Brique Proxmox, installation et contrôle

- ▶ Création sur le noeud 1 (Boole)
pvecm create CLUSTER_NAME
- ▶ Ajout du noeud 2 (Cardan)
pvecm add IP_ADDRESS_NOEUD2
- ▶ Contrôles : présence du cluster des noeuds Boole - Cardan, status des noeuds
pvecm nodes
pvecm status

Retour
d'expérience sur
PROXMOX et HA

Richard Ferrere
Romain Pacé

Un peu de
terminologie

PROXMOX et DRBD?

La haute disponibilité

Infrastructure du
laboratoire

Notre laboratoire à
Besançon

Notre cahier des charges

Système et topologie cluster
retenus

Installation des
différentes briques

IPMI, réseau

Proxmox, DRBD et LVM

Briques liées à la HA

Défaillances et HA

Capacité réseau..

Matérielle, énergétique...

Démo

Conclusion

Liens et références
internet

Logs liés à la haute
disponibilité

Briques liées à la HA

► Fichier générale du cluster `/etc/pve/cluster.conf`

```
1 <?xml version="1.0"?>
2 <cluster config_version="13" name="hacluster">
3   <cmn expected_votes="3" keyfile="/var/lib/pve-cluster/corosync.authkey"/>
4   <quorum allow_kill="0" interval="1" label="pve_quorum" tko="10" votes="1"/>
5   <totem token="54000"/>
6   <clusternodes>
7     <clusternode name="hanode1" nodeid="1" votes="1">
8       <fence>
9         <method name="1">
10          <device action="reboot" name="ipminode1"/>
11        </method>
12      </fence>
13    </clusternode>
14    <clusternode name="hanode2" nodeid="2" votes="1">
15      <fence>
16        <method name="1">
17          <device action="reboot" name="ipminode2"/>
18        </method>
19      </fence>
20    </clusternode>
21  </clusternodes>
22  <fencedevices>
23    <fencedevice agent="fence_ipmilan" ipaddr="172.16.8.110" login="root" name="ipminode1" passwd="Labo"/>
24    <fencedevice agent="fence_ipmilan" ipaddr="172.16.8.210" login="root" name="ipminode2" passwd="Labo"/>
25  </fencedevices>
26 </rm>
27 <pvevm autostart="1" vmid="100"/>
28 <pvevm autostart="1" vmid="101"/>
29 <pvevm autostart="1" vmid="102"/>
30 <pvevm autostart="1" vmid="103"/>
31 </rm>
32 </cluster>
```

Un peu de
terminologie

PROXMOX et DRBD?

La haute disponibilité

Infrastructure du
laboratoire

Notre laboratoire à
Besançon

Notre cahier des charges

Système et topologie cluster
retenus

Installation des
différentes briques

IPMI, réseau

Proxmox, DRBD et LVM

Briques liées à la HA

Défaillances et HA

Capacité réseau..

Matérielle, énergétique...

Démo

Conclusion

Liens et références
internet

Logs liés à la haute
disponibilité

Autoriser le fencing

- ▶ Modification de `/etc/default/redhat-cluster-pve`

```
FENCE_JOIN="yes"  
fence_tool join
```

Installation du disque de quorum

- ▶ Un simple machine hautement disponible suffit
- ▶ Une target ISCSI montée sur les 2 noeuds
- ▶ Initialisation du disque sur un noeud

```
# parted -a optimal /dev/sdc1  
# mklabel msdos  
# mkpart primary 1 100%
```
- ▶ Création du disque dans le cluster

```
# mkqdisk -c /dev/sdc1 -l pve_quorum
```

Redémarrage des noeuds!!!

Retour
d'expérience sur
PROXMOX et HA

Richard Ferrere
Romain Pacé

Un peu de
terminologie

PROXMOX et DRBD?

La haute disponibilité

Infrastructure du
laboratoire

Notre laboratoire à
Besançon

Notre cahier des charges

Système et topologie cluster
retenus

Installation des
différentes briques

IPMI, réseau

Proxmox, DRBD et LVM

Briques liées à la HA

Défaillances et HA

Capacité réseau..

Matérielle, énergétique...

Démo

Conclusion

Liens et références
internet

Logs liés à la haute
disponibilité

Défaillances et HA

Retour
d'expérience sur
PROXMOX et HA

Richard Ferrere
Romain Pacé

Un peu de
terminologie

PROXMOX et DRBD ?

La haute disponibilité

Infrastructure du
laboratoire

Notre laboratoire à
Besançon

Notre cahier des charges

Système et topologie cluster
retenus

Installation des
différentes briques

IPMI, réseau

Proxmox, DRBD et LVM

Briques liées à la HA

Défaillances et HA

Capacité réseau..

Matérielle, énergétique...

Démo

Conclusion

Liens et références
internet

Logs liés à la haute
disponibilité

Cas de défaillances possibles : réseau

Retour
d'expérience sur
PROXMOX et HA

Richard Ferrere
Romain Pacé

Cas 1 : probabilité très haute (75-90%) -> coupure réseau sans agrégation de liens sur un des noeuds

- ▶ Le lien dédié à DRBD et l'interface réseau coupés sur le noeud 1
- ▶ La synchronisation entre les 2 noeuds interrompue !
- ▶ Le noeud 1 hébergeant des machines virtuelles est à jour
- ▶ Les VMs migrent automatiquement sur le noeud 2
-> L'état de la synchronisation : Primary/Unknown –
UpToDate/Unknown
-> Les services sont arrêtés et démarrés sur le noeud 2
même si son état n'est plus à jour !

Un peu de
terminologie

PROXMOX et DRBD ?

La haute disponibilité

Infrastructure du
laboratoire

Notre laboratoire à
Besançon

Notre cahier des charges

Système et topologie cluster
retenus

Installation des
différentes briques

IPMI, réseau

Proxmox, DRBD et LVM

Briques liées à la HA

Défaillances et HA

Capacité réseau..

Matérielle, énergétique...

Démo

Conclusion

Liens et références
internet

Logs liés à la haute
disponibilité

Cas de défaillances possibles : en terme d'efficience

Cas 2 : probabilité moyenne (40-60%) -> maintenance/durée

- ▶ Le noeud 1 revient dans le groupe cluster
- ▶ Le noeud 2 doit avoir terminé sa resynchronisation !
- ▶ Arrêt du noeud 2 pour maintenance
- ▶ Durée d'arrêt longue = temps de synchronisation élevé
-> Les services sont arrêtés et démarrés sur le noeud 1
- ▶ Mais si le noeud 1 tombe entretemps, alors :
-> blocs de données erronés et split-brain !

Cas 3 : probabilité très faible (1-10%) -> agrégation de liens, électrique

- ▶ Coupure de l'agrégation de liens réseaux (Bonding)
- ▶ Plus d'alimentation électrique sur 1 noeud !
-> La situation est bloquée, il ne se passe rien le temps que...

(Loading Video...)

Retour
d'expérience sur
PROXMOX et HA

Richard Ferrere
Romain Pacé

Un peu de
terminologie

PROXMOX et DRBD ?

La haute disponibilité

Infrastructure du
laboratoire

Notre laboratoire à
Besançon

Notre cahier des charges

Système et topologie cluster
retenus

Installation des
différentes briques

IPMI, réseau

Proxmox, DRBD et LVM

Briques liées à la HA

Défaillances et HA

Capacité réseau..

Matérielle, énergétique...

Démo

Conclusion

Liens et références
internet

Logs liés à la haute
disponibilité

Conclusion et références

Un peu de terminologie

PROXMOX et DRBD ?

La haute disponibilité

Infrastructure du laboratoire

Notre laboratoire à Besançon

Notre cahier des charges

Système et topologie cluster retenus

Installation des différentes briques

IPMI, réseau

Proxmox, DRBD et LVM

Briques liées à la HA

Défaillances et HA

Capacité réseau..

Matérielle, énergétique...

Démo

Conclusion

Liens et références internet

Logs liés à la haute disponibilité

CONCLUSION

- ▶ Proxmox et DRBD : véritable binôme pour un gain de temps appréciable et beaucoup de possibilités...
- ▶ Outil simple de par son interface web, opérationnel, stable et nécessitant peu de maintenance mais de la configuration
- ▶ Très grande souplesse à plusieurs niveaux : déploiement de services à la demande, sauvegardes, restaurations et autres...
- ▶ Nécessite une attention particulière pour mettre en place la HA pour une continuité de service
- ▶ Beaucoup de lectures et de recherches bibliographiques pour appréhender la terminologie HA
- ▶ La HA a ses limites face à certaines défaillances
→ Fencing OK pour la couche réseau et système ; Pas toujours OK pour les défauts électriques
- ▶ Développement de scripts possibles pour la surveillance de la synchronisation DRBD et autres...
- ▶ Oblige à avoir un parc de serveurs plus homogène et facile à gérer
- ▶ L'intégrité des données est-elle toujours préservée ?

Retour
d'expérience sur
PROXMOX et HA

Richard Ferrere
Romain Pacé

Un peu de
terminologie

PROXMOX et DRBD ?
La haute disponibilité

Infrastructure du
laboratoire

Notre laboratoire à
Besançon
Notre cahier des charges
Système et topologie cluster
retenus

Installation des
différentes briques

IPMI, réseau
Proxmox, DRBD et LVM
Briques liées à la HA

Défaillances et HA

Capacité réseau...
Matérielle, énergétique...

Démo

Conclusion

Liens et références
internet

Logs liés à la haute
disponibilité

Quelques références

- ▶ http://pve.proxmox.com/wiki/Proxmox_VE_2.0_Cluster
- ▶ <http://forum.proxmox.com/forum.php>
- ▶ <http://pve.proxmox.com/wiki/DRBD>
- ▶ <https://pve.proxmox.com/wiki/Fencing>
- ▶ <https://aresu.dsi.cnrs.fr/spip.php?article198>
- ▶ <http://blog.héry.com/article11/cluster-proxmox-distant-le-concept>
- ▶ https://access.redhat.com/site/documentation/en-US/Red_Hat_Enterprise_Linux/6/html/High_Availability_Add-On_Overview/ch.gfscs.cluster-overview-CSO.html

Retour
d'expérience sur
PROXMOX et HA

Richard Ferrere
Romain Pacé

Un peu de
terminologie

PROXMOX et DRBD?

La haute disponibilité

Infrastructure du
laboratoire

Notre laboratoire à
Besançon

Notre cahier des charges

Système et topologie cluster
retenus

Installation des
différentes briques

IPMI, réseau

Proxmox, DRBD et LVM

Briques liées à la HA

Défaillances et HA

Capacité réseau...

Matérielle, énergétique...

Démo

Conclusion

Liens et références
internet

Logs liés à la haute
disponibilité

Quelques références - suite...

- ▶ https://pve.proxmox.com/wiki/Service_daemons
- ▶ <http://miao5.blogspot.fr/2008/11/what-are-recommend-settings-when-using.html>
- ▶ https://alteeve.ca/w/AN!Cluster_Tutorial_2
- ▶ http://pve.proxmox.com/wiki/Two-Node_High_Availability_Cluster
- ▶ <http://linux.die.net/man/5/cluster.conf.../qdisk.../cman.../corosync.conf.../rgmanager>

Retour
d'expérience sur
PROXMOX et HA

Richard Ferrere
Romain Pacé

Un peu de
terminologie

PROXMOX et DRBD?

La haute disponibilité

Infrastructure du
laboratoire

Notre laboratoire à
Besançon

Notre cahier des charges

Système et topologie cluster
retenus

Installation des
différentes briques

IPMI, réseau

Proxmox, DRBD et LVM

Briques liées à la HA

Défaillances et HA

Capacité réseau..

Matérielle, énergétique...

Démo

Conclusion

Liens et références
internet

Logs liés à la haute
disponibilité

Les logs se trouvent à la suite de la présentation, 6 pages détaillées de logs des différentes actions liées au mécanisme de la HA !

Bonne lecture et bonne chance 😊

Retour
d'expérience sur
PROXMOX et HA

Richard Ferrere
Romain Pacé

Un peu de
terminologie

PROXMOX et DRBD ?

La haute disponibilité

Infrastructure du
laboratoire

Notre laboratoire à
Besançon

Notre cahier des charges

Système et topologie cluster
retenus

Installation des
différentes briques

IPMI, réseau

Proxmox, DRBD et LVM

Briques liées à la HA

Défaillances et HA

Capacité réseau..

Matérielle, énergétique...

Démo

Conclusion

Liens et références
internet

Logs liés à la haute
disponibilité

Merci de votre attention !
Avez-vous des questions ?

Un peu de terminologie

PROXMOX et DRBD ?

La haute disponibilité

Infrastructure du laboratoire

Notre laboratoire à
Besançon

Notre cahier des charges

Système et topologie cluster
retenus

Installation des différentes briques

IPMI, réseau

Proxmox, DRBD et LVM

Briques liées à la HA

Défaillances et HA

Capacité réseau..

Matérielle, énergétique...

Démo

Conclusion

Liens et références internet

Logs liés à la haute disponibilité

Logs du nœud 1 en « vie »

```
#####  
#COUPURE RESEAU A ENVIRON 16:55#  
#####
```

```
##### /var/log/cluster/corosync.log  
#####
```

```
Mar 27 16:55:59 corosync [TOTEM ] A processor failed, forming new configuration.  
Mar 27 16:56:01 corosync [CLM ] CLM CONFIGURATION CHANGE  
Mar 27 16:56:01 corosync [CLM ] New Configuration:  
Mar 27 16:56:01 corosync [CLM ] r(0) ip(172.16.8.100)  
Mar 27 16:56:01 corosync [CLM ] Members Left:  
Mar 27 16:56:01 corosync [CLM ] r(0) ip(172.16.8.200)  
Mar 27 16:56:01 corosync [CLM ] Members Joined:  
Mar 27 16:56:01 corosync [QUORUM] Members[1]: 1  
Mar 27 16:56:01 corosync [CLM ] CLM CONFIGURATION CHANGE  
Mar 27 16:56:01 corosync [CLM ] New Configuration:  
Mar 27 16:56:01 corosync [CLM ] r(0) ip(172.16.8.100)  
Mar 27 16:56:01 corosync [CLM ] Members Left:  
Mar 27 16:56:01 corosync [CLM ] Members Joined:  
Mar 27 16:56:01 corosync [TOTEM ] A processor joined or left the membership and a new membership was formed.  
Mar 27 16:56:01 corosync [CPG ] chosen downlist: sender r(0) ip(172.16.8.100) ; members(old:2 left:1)  
Mar 27 16:56:01 corosync [MAIN ] Completed service synchronization, ready to provide service.  
Mar 27 16:58:35 corosync [CLM ] CLM CONFIGURATION CHANGE  
Mar 27 16:58:35 corosync [CLM ] New Configuration:  
Mar 27 16:58:35 corosync [CLM ] r(0) ip(172.16.8.100)  
Mar 27 16:58:35 corosync [CLM ] Members Left:  
Mar 27 16:58:35 corosync [CLM ] Members Joined:  
Mar 27 16:58:35 corosync [CLM ] CLM CONFIGURATION CHANGE  
Mar 27 16:58:35 corosync [CLM ] New Configuration:  
Mar 27 16:58:35 corosync [CLM ] r(0) ip(172.16.8.100)  
Mar 27 16:58:35 corosync [CLM ] r(0) ip(172.16.8.200)  
Mar 27 16:58:35 corosync [CLM ] Members Left:  
Mar 27 16:58:35 corosync [CLM ] Members Joined:  
Mar 27 16:58:35 corosync [CLM ] r(0) ip(172.16.8.200)  
Mar 27 16:58:35 corosync [TOTEM ] A processor joined or left the membership and a new membership was formed.  
Mar 27 16:58:35 corosync [QUORUM] Members[2]: 1 2  
Mar 27 16:58:35 corosync [QUORUM] Members[2]: 1 2  
Mar 27 16:58:35 corosync [CPG ] chosen downlist: sender r(0) ip(172.16.8.100) ; members(old:1 left:0)  
Mar 27 16:58:35 corosync [MAIN ] Completed service synchronization, ready to provide service.
```

```
##### /var/log/cluster/fenced.log  
#####
```

```
Mar 27 16:56:01 fenced fencing node hanode2  
Mar 27 16:56:15 fenced fence hanode2 success
```

```
##### /var/log/cluster/rgmanager.log  
#####
```

```
Mar 27 16:56:01 rgmanager State change: hanode2 DOWN  
Mar 27 16:56:16 rgmanager Starting stopped service pvevm:100  
Mar 27 16:56:16 rgmanager Starting stopped service pvevm:101  
Mar 27 16:56:16 rgmanager [pvevm] VM 102 is running  
Mar 27 16:56:16 rgmanager [pvevm] Move config for VM 100 to local node  
Mar 27 16:56:16 rgmanager [pvevm] Move config for VM 101 to local node  
Mar 27 16:56:17 rgmanager Service pvevm:100 started  
Mar 27 16:56:17 rgmanager Service pvevm:101 started
```

```
##### /var/log/cluster/qdiskd.log  
#####
```

```
Mar 27 16:41:17 qdiskd qdisk cycle took more than 1 second to complete (1.270000)  
Mar 27 16:41:19 qdiskd qdisk cycle took more than 1 second to complete (1.230000)  
Mar 27 16:41:21 qdiskd qdisk cycle took more than 1 second to complete (1.290000)  
Mar 27 16:41:23 qdiskd qdisk cycle took more than 1 second to complete (1.320000)  
Mar 27 16:41:26 qdiskd qdisk cycle took more than 1 second to complete (1.200000)  
Mar 27 16:43:01 qdiskd qdisk cycle took more than 1 second to complete (1.300000)  
Mar 27 16:43:03 qdiskd qdisk cycle took more than 1 second to complete (1.310000)  
Mar 27 16:43:05 qdiskd qdisk cycle took more than 1 second to complete (1.260000)  
Mar 27 16:43:08 qdiskd qdisk cycle took more than 1 second to complete (1.240000)  
Mar 27 16:55:22 qdiskd Writing eviction notice for node 2  
Mar 27 16:55:23 qdiskd Node 2 evicted
```

/var/log/messages

#####

logs en rapport avec DRBD et RGMANAGER

```
Mar 27 16:55:13 hanode1 kernel: block drbd1: peer( Primary -> Unknown ) conn( Connected -> NetworkFailure ) pdsk( UpToDate -> DUnknown )
Mar 27 16:55:13 hanode1 kernel: block drbd1: new current UUID
DE4ABC80C15BA0AB:999A64E8DAD9173D:C647447650C9216F:C646447650C9216F
Mar 27 16:55:13 hanode1 kernel: block drbd1: asender terminated
Mar 27 16:55:13 hanode1 kernel: block drbd1: Terminating asender thread
Mar 27 16:55:13 hanode1 kernel: block drbd1: Connection closed
Mar 27 16:55:13 hanode1 kernel: block drbd1: conn( NetworkFailure -> Unconnected )
Mar 27 16:55:13 hanode1 kernel: block drbd1: receiver terminated
Mar 27 16:55:13 hanode1 kernel: block drbd1: Restarting receiver thread
Mar 27 16:55:13 hanode1 kernel: block drbd1: receiver (re)started
Mar 27 16:55:13 hanode1 kernel: block drbd1: conn( Unconnected -> WFConnection )
Mar 27 16:55:13 hanode1 kernel: bnx2 0000:07:00.0: eth1: NIC Copper Link is Up, 100 Mbps full duplex, receive & transmit flow control ON
Mar 27 16:58:55 hanode1 rgmanager[2907]: State change: hanode2 UP
```

Logs du nœud 2 avec un problème réseau

```
#####  
#COUPURE RESEAU A ENVIRON 16:55#  
#####
```

```
##### /var/log/cluster/corosync.log  
#####
```

```
Mar 27 16:55:59 corosync [TOTEM ] A processor failed, forming new configuration.  
Mar 27 16:55:59 corosync [TOTEM ] The network interface is down.  
Mar 27 16:56:01 corosync [CLM ] CLM CONFIGURATION CHANGE  
Mar 27 16:56:01 corosync [CLM ] New Configuration:  
Mar 27 16:56:01 corosync [CLM ] r(0) ip(127.0.0.1)  
Mar 27 16:56:01 corosync [CLM ] Members Left:  
Mar 27 16:56:01 corosync [CLM ] r(0) ip(127.0.0.1)  
Mar 27 16:56:01 corosync [CLM ] Members Joined:  
Mar 27 16:56:01 corosync [QUORUM] Members[1]: 1  
Mar 27 16:56:01 corosync [CLM ] CLM CONFIGURATION CHANGE  
Mar 27 16:56:01 corosync [CLM ] New Configuration:  
Mar 27 16:56:01 corosync [CLM ] r(0) ip(127.0.0.1)  
Mar 27 16:56:01 corosync [CLM ] Members Left:  
Mar 27 16:56:01 corosync [CLM ] Members Joined:  
Mar 27 16:56:01 corosync [TOTEM ] A processor joined or left the membership and a new membership was formed.  
Mar 27 16:56:01 corosync [QUORUM] Members[2]: 1 2  
Mar 27 16:56:01 corosync [QUORUM] Members[2]: 1 2  
Mar 27 16:56:01 corosync [CPG ] chosen downlist: sender r(0) ip(127.0.0.1) ; members(old:2 left:1)  
Mar 27 16:56:01 corosync [MAIN ] Completed service synchronization, ready to provide service.  
Mar 27 16:58:35 corosync [MAIN ] Corosync Cluster Engine ('1.4.5'): started and ready to provide service.  
Mar 27 16:58:35 corosync [MAIN ] Corosync built-in features: nss  
Mar 27 16:58:35 corosync [MAIN ] Successfully read config from /etc/cluster/cluster.conf  
Mar 27 16:58:35 corosync [MAIN ] Successfully parsed cman config  
Mar 27 16:58:35 corosync [MAIN ] Successfully configured openais services to load  
Mar 27 16:58:35 corosync [TOTEM ] Initializing transport (UDP/IP Multicast).  
Mar 27 16:58:35 corosync [TOTEM ] Initializing transmit/receive security: libtomcrypt SOBER128/SHA1HMAC (mode 0).  
Mar 27 16:58:35 corosync [TOTEM ] The network interface [172.16.8.200] is now up.  
Mar 27 16:58:35 corosync [QUORUM] Using quorum provider quorum_cman  
Mar 27 16:58:35 corosync [SERV ] Service engine loaded: corosync cluster quorum service v0.1  
Mar 27 16:58:35 corosync [CMAN ] CMAN 1364188437 (built Mar 25 2013 06:14:01) started  
Mar 27 16:58:35 corosync [SERV ] Service engine loaded: corosync CMAN membership service 2.90  
Mar 27 16:58:35 corosync [SERV ] Service engine loaded: openais cluster membership service B.01.01  
Mar 27 16:58:35 corosync [SERV ] Service engine loaded: openais event service B.01.01  
Mar 27 16:58:35 corosync [SERV ] Service engine loaded: openais checkpoint service B.01.01  
Mar 27 16:58:35 corosync [SERV ] Service engine loaded: openais message service B.03.01  
Mar 27 16:58:35 corosync [SERV ] Service engine loaded: openais distributed locking service B.03.01  
Mar 27 16:58:35 corosync [SERV ] Service engine loaded: openais timer service A.01.01  
Mar 27 16:58:35 corosync [SERV ] Service engine loaded: corosync extended virtual synchrony service  
Mar 27 16:58:35 corosync [SERV ] Service engine loaded: corosync configuration service  
Mar 27 16:58:35 corosync [SERV ] Service engine loaded: corosync cluster closed process group service v1.01  
Mar 27 16:58:35 corosync [SERV ] Service engine loaded: corosync cluster config database access v1.01  
Mar 27 16:58:35 corosync [SERV ] Service engine loaded: corosync profile loading service  
Mar 27 16:58:35 corosync [QUORUM] Using quorum provider quorum_cman  
Mar 27 16:58:35 corosync [SERV ] Service engine loaded: corosync cluster quorum service v0.1  
Mar 27 16:58:35 corosync [MAIN ] Compatibility mode set to whitetank. Using V1 and V2 of the synchronization engine.  
Mar 27 16:58:35 corosync [CLM ] CLM CONFIGURATION CHANGE  
Mar 27 16:58:35 corosync [CLM ] New Configuration:  
Mar 27 16:58:35 corosync [CLM ] Members Left:  
Mar 27 16:58:35 corosync [CLM ] Members Joined:  
Mar 27 16:58:35 corosync [CLM ] CLM CONFIGURATION CHANGE  
Mar 27 16:58:35 corosync [CLM ] New Configuration:  
Mar 27 16:58:35 corosync [CLM ] r(0) ip(172.16.8.200)  
Mar 27 16:58:35 corosync [CLM ] Members Left:  
Mar 27 16:58:35 corosync [CLM ] Members Joined:  
Mar 27 16:58:35 corosync [CLM ] r(0) ip(172.16.8.200)  
Mar 27 16:58:35 corosync [TOTEM ] A processor joined or left the membership and a new membership was formed.  
Mar 27 16:58:35 corosync [QUORUM] Members[1]: 2  
Mar 27 16:58:35 corosync [QUORUM] Members[1]: 2  
Mar 27 16:58:35 corosync [CPG ] chosen downlist: sender r(0) ip(172.16.8.200) ; members(old:0 left:0)  
Mar 27 16:58:35 corosync [MAIN ] Completed service synchronization, ready to provide service.  
Mar 27 16:58:35 corosync [CLM ] CLM CONFIGURATION CHANGE  
Mar 27 16:58:35 corosync [CLM ] New Configuration:  
Mar 27 16:58:35 corosync [CLM ] r(0) ip(172.16.8.200)  
Mar 27 16:58:35 corosync [CLM ] Members Left:  
Mar 27 16:58:35 corosync [CLM ] Members Joined:  
Mar 27 16:58:35 corosync [CLM ] CLM CONFIGURATION CHANGE  
Mar 27 16:58:35 corosync [CLM ] New Configuration:  
Mar 27 16:58:35 corosync [CLM ] r(0) ip(172.16.8.100)  
Mar 27 16:58:35 corosync [CLM ] r(0) ip(172.16.8.200)
```

```
Mar 27 16:58:35 corosync [CLM ] Members Left:
Mar 27 16:58:35 corosync [CLM ] Members Joined:
Mar 27 16:58:35 corosync [CLM ] r(0) ip(172.16.8.100)
Mar 27 16:58:35 corosync [TOTEM ] A processor joined or left the membership and a new membership was formed.
Mar 27 16:58:35 corosync [CMAN ] quorum regained, resuming activity
Mar 27 16:58:35 corosync [QUORUM] This node is within the primary component and will provide service.
Mar 27 16:58:35 corosync [QUORUM] Members[2]: 1 2
Mar 27 16:58:35 corosync [QUORUM] Members[2]: 1 2
Mar 27 16:58:35 corosync [CPG ] chosen downlist: sender r(0) ip(172.16.8.100) ; members(old:1 left:0)
Mar 27 16:58:35 corosync [MAIN ] Completed service synchronization, ready to provide service.
Mar 27 16:58:48 corosync [CMAN ] quorum device registered
Mar 27 16:58:48 corosync [QUORUM] Members[2]: 1 2
```

/var/log/cluster/fenced.log

#####

```
Mar 27 16:56:01 fenced fencing node hanode1
Mar 27 16:56:02 fenced fence hanode1 dev 0.0 agent fence_ipmilan result: error from agent
Mar 27 16:56:02 fenced fence hanode1 failed
Mar 27 16:56:05 fenced fencing node hanode1
Mar 27 16:58:49 fenced fenced 1364188437 started
```

/var/log/cluster/rgmanager.log

#####

```
Mar 27 16:58:52 rgmanager I am node #2
Mar 27 16:58:52 rgmanager Resource Group Manager Starting
Mar 27 16:58:52 rgmanager Loading Service Data
Mar 27 16:58:54 rgmanager Initializing Services
Mar 27 16:58:54 rgmanager [pvevm] VM 101 is already stopped
Mar 27 16:58:54 rgmanager [pvevm] VM 100 is already stopped
Mar 27 16:58:54 rgmanager [pvevm] VM 102 is already stopped
Mar 27 16:58:54 rgmanager [pvevm] VM 103 is already stopped
Mar 27 16:58:55 rgmanager Services Initialized
Mar 27 16:58:55 rgmanager State change: Local UP
Mar 27 16:58:55 rgmanager State change: hanode1 UP
Mar 27 16:58:55 rgmanager Migration: pvevm:100 is running on 1
Mar 27 16:58:55 rgmanager Migration: pvevm:101 is running on 1
Mar 27 16:58:56 rgmanager Migration: pvevm:102 is running on 1
Mar 27 16:58:56 rgmanager Migration: pvevm:103 is running on 1
```

/var/log/cluster/qdiskd.log

#####

```
Mar 27 16:55:17 qdiskd qdiskd: read (system call) has hung for 5 seconds
Mar 27 16:55:17 qdiskd In 5 more seconds, we will be evicted
Mar 27 16:58:39 qdiskd Quorum Partition: /dev/block/8:33 Label: pve_quorum
Mar 27 16:58:39 qdiskd Quorum Daemon Initializing
Mar 27 16:58:44 qdiskd Node 1 is the master
Mar 27 16:58:48 qdiskd Initial score 1/1
Mar 27 16:58:48 qdiskd Initialization complete
Mar 27 16:58:48 qdiskd Score sufficient for master operation (1/1; required=1); upgrading
```

/var/log/messages

#####

Logs en rapport avec DRBD, RGMANAGER, QDISKD

```
Mar 27 16:55:13 hanode2 kernel: block drbd1: peer( Primary -> Unknown ) conn( Connected -> NetworkFailure ) pdsk( UpToDate -> DUnknown )
Mar 27 16:55:13 hanode2 kernel: block drbd1: new current UUID
E3B0A4E8C3DC0D1F:999A64E8DAD9173D:C647447650C9216F:C646447650C9216F
Mar 27 16:55:13 hanode2 kernel: block drbd1: asender terminated
Mar 27 16:55:13 hanode2 kernel: block drbd1: Terminating asender thread
Mar 27 16:55:13 hanode2 kernel: block drbd1: Connection closed
Mar 27 16:55:13 hanode2 kernel: block drbd1: conn( NetworkFailure -> Unconnected )
Mar 27 16:55:13 hanode2 kernel: block drbd1: receiver terminated
Mar 27 16:55:13 hanode2 kernel: block drbd1: Restarting receiver thread
Mar 27 16:55:13 hanode2 kernel: block drbd1: receiver (re)started
Mar 27 16:55:13 hanode2 kernel: block drbd1: conn( Unconnected -> WFCConnection )
Mar 27 16:55:13 hanode2 kernel: block drbd1: conn( WFCConnection -> Disconnecting )
Mar 27 16:55:13 hanode2 kernel: block drbd1: Discarding network configuration.
Mar 27 16:55:13 hanode2 kernel: block drbd1: Connection closed
Mar 27 16:55:13 hanode2 kernel: block drbd1: conn( Disconnecting -> StandAlone )
Mar 27 16:55:13 hanode2 kernel: block drbd1: receiver terminated
Mar 27 16:55:13 hanode2 kernel: block drbd1: Terminating receiver thread
Mar 27 16:55:17 hanode2 qdiskd[2560]: qdiskd: read (system call) has hung for 5 seconds
Mar 27 16:55:17 hanode2 qdiskd[2560]: In 5 more seconds, we will be evicted
Mar 27 16:58:35 hanode2 kernel: block drbd1: Resync done (total 1 sec; paused 0 sec; 80224 K/sec)
Mar 27 16:58:35 hanode2 kernel: block drbd1: updated UUIDs
DE4ABC80C15BA0AB:0000000000000000:999B64E8DAD9173D:999A64E8DAD9173D
```

Mar 27 16:58:35 hanode2 kernel: block drbd1: conn(SyncTarget -> Connected) disk(Inconsistent -> UpToDate)
Mar 27 16:58:35 hanode2 kernel: block drbd1: helper command: /sbin/drbdadm after-resync-target minor-1
Mar 27 16:58:35 hanode2 kernel: block drbd1: helper command: /sbin/drbdadm after-resync-target minor-1 exit code 0 (0x0)
Mar 27 16:58:35 hanode2 kernel: block drbd1: bitmap WRITE of 1066 pages took 8 jiffies
Mar 27 16:58:35 hanode2 kernel: block drbd1: 0 KB (0 bits) marked out-of-sync by on disk bit-map.
Mar 27 16:58:35 hanode2 corosync[2461]: [TOTEM] The network interface [172.16.8.200] is now up.
Mar 27 16:58:35 hanode2 corosync[2461]: [QUORUM] Using quorum provider quorum_cman
Mar 27 16:58:35 hanode2 corosync[2461]: [SERV] Service engine loaded: corosync cluster quorum service v0.1
Mar 27 16:58:35 hanode2 corosync[2461]: [CMAN] CMAN 1364188437 (built Mar 25 2013 06:14:01) started
Mar 27 16:58:35 hanode2 corosync[2461]: [SERV] Service engine loaded: corosync CMAN membership service 2.90
Mar 27 16:58:35 hanode2 corosync[2461]: [SERV] Service engine loaded: openais cluster membership service B.01.01
Mar 27 16:58:35 hanode2 corosync[2461]: [SERV] Service engine loaded: openais event service B.01.01
Mar 27 16:58:35 hanode2 corosync[2461]: [SERV] Service engine loaded: openais checkpoint service B.01.01
Mar 27 16:58:35 hanode2 corosync[2461]: [SERV] Service engine loaded: openais message service B.03.01
Mar 27 16:58:35 hanode2 corosync[2461]: [SERV] Service engine loaded: openais distributed locking service B.03.01
Mar 27 16:58:35 hanode2 corosync[2461]: [SERV] Service engine loaded: openais timer service A.01.01
Mar 27 16:58:35 hanode2 corosync[2461]: [SERV] Service engine loaded: corosync extended virtual synchrony service
Mar 27 16:58:35 hanode2 corosync[2461]: [SERV] Service engine loaded: corosync configuration service
Mar 27 16:58:35 hanode2 corosync[2461]: [SERV] Service engine loaded: corosync cluster closed process group service v1.01
Mar 27 16:58:35 hanode2 corosync[2461]: [SERV] Service engine loaded: corosync cluster config database access v1.01
Mar 27 16:58:35 hanode2 corosync[2461]: [SERV] Service engine loaded: corosync profile loading service
Mar 27 16:58:35 hanode2 corosync[2461]: [QUORUM] Using quorum provider quorum_cman
Mar 27 16:58:35 hanode2 corosync[2461]: [SERV] Service engine loaded: corosync cluster quorum service v0.1
Mar 27 16:58:35 hanode2 corosync[2461]: [MAIN] Compatibility mode set to whitetank. Using V1 and V2 of the synchronization engine.
Mar 27 16:58:35 hanode2 corosync[2461]: [CLM] CLM CONFIGURATION CHANGE
Mar 27 16:58:35 hanode2 corosync[2461]: [CLM] New Configuration:
Mar 27 16:58:35 hanode2 corosync[2461]: [CLM] Members Left:
Mar 27 16:58:35 hanode2 corosync[2461]: [CLM] Members Joined:
Mar 27 16:58:35 hanode2 corosync[2461]: [CLM] CLM CONFIGURATION CHANGE
Mar 27 16:58:35 hanode2 corosync[2461]: [CLM] New Configuration:
Mar 27 16:58:35 hanode2 corosync[2461]: [CLM] #011r(0) ip(172.16.8.200)
Mar 27 16:58:35 hanode2 corosync[2461]: [CLM] Members Left:
Mar 27 16:58:35 hanode2 corosync[2461]: [CLM] Members Joined:
Mar 27 16:58:35 hanode2 corosync[2461]: [CLM] #011r(0) ip(172.16.8.200)
Mar 27 16:58:35 hanode2 corosync[2461]: [TOTEM] A processor joined or left the membership and a new membership was formed.
Mar 27 16:58:35 hanode2 corosync[2461]: [QUORUM] Members[1]: 2
Mar 27 16:58:35 hanode2 corosync[2461]: [QUORUM] Members[1]: 2
Mar 27 16:58:35 hanode2 corosync[2461]: [CPG] chosen downlist: sender r(0) ip(172.16.8.200) ; members(old:0 left:0)
Mar 27 16:58:35 hanode2 corosync[2461]: [MAIN] Completed service synchronization, ready to provide service.
Mar 27 16:58:35 hanode2 corosync[2461]: [CLM] CLM CONFIGURATION CHANGE
Mar 27 16:58:35 hanode2 corosync[2461]: [CLM] New Configuration:
Mar 27 16:58:35 hanode2 corosync[2461]: [CLM] #011r(0) ip(172.16.8.200)
Mar 27 16:58:35 hanode2 corosync[2461]: [CLM] Members Left:
Mar 27 16:58:35 hanode2 corosync[2461]: [CLM] Members Joined:
Mar 27 16:58:35 hanode2 corosync[2461]: [CLM] CLM CONFIGURATION CHANGE
Mar 27 16:58:35 hanode2 corosync[2461]: [CLM] New Configuration:
Mar 27 16:58:35 hanode2 corosync[2461]: [CLM] #011r(0) ip(172.16.8.100)
Mar 27 16:58:35 hanode2 corosync[2461]: [CLM] #011r(0) ip(172.16.8.200)
Mar 27 16:58:35 hanode2 corosync[2461]: [CLM] Members Left:
Mar 27 16:58:35 hanode2 corosync[2461]: [CLM] Members Joined:
Mar 27 16:58:35 hanode2 corosync[2461]: [CLM] #011r(0) ip(172.16.8.100)
Mar 27 16:58:35 hanode2 corosync[2461]: [TOTEM] A processor joined or left the membership and a new membership was formed.
Mar 27 16:58:35 hanode2 corosync[2461]: [CMAN] quorum regained, resuming activity
Mar 27 16:58:35 hanode2 corosync[2461]: [QUORUM] This node is within the primary component and will provide service.
Mar 27 16:58:35 hanode2 corosync[2461]: [QUORUM] Members[2]: 1 2
Mar 27 16:58:35 hanode2 corosync[2461]: [QUORUM] Members[2]: 1 2
Mar 27 16:58:35 hanode2 corosync[2461]: [CPG] chosen downlist: sender r(0) ip(172.16.8.100) ; members(old:1 left:0)
Mar 27 16:58:35 hanode2 corosync[2461]: [MAIN] Completed service synchronization, ready to provide service.
Mar 27 16:58:39 hanode2 qdiskd[2520]: Quorum Partition: /dev/block/8:33 Label: pve_quorum
Mar 27 16:58:39 hanode2 qdiskd[2520]: Quorum Daemon Initializing
Mar 27 16:58:44 hanode2 qdiskd[2520]: Node 1 is the master
Mar 27 16:58:48 hanode2 qdiskd[2520]: Initial score 1/1
Mar 27 16:58:48 hanode2 qdiskd[2520]: Initialization complete
Mar 27 16:58:48 hanode2 corosync[2461]: [CMAN] quorum device registered
Mar 27 16:58:48 hanode2 corosync[2461]: [QUORUM] Members[2]: 1 2
Mar 27 16:58:48 hanode2 qdiskd[2520]: Score sufficient for master operation (1/1; required=1); upgrading
Mar 27 16:58:49 hanode2 fenced[2691]: fenced 1364188437 started
Mar 27 16:58:49 hanode2 dlm_controld[2712]: dlm_controld 1364188437 started
Mar 27 16:58:52 hanode2 rgmanager[2953]: I am node #2
Mar 27 16:58:52 hanode2 rgmanager[2953]: Resource Group Manager Starting
Mar 27 16:58:52 hanode2 rgmanager[2953]: Loading Service Data
Mar 27 16:58:54 hanode2 rgmanager[2953]: Initializing Services
Mar 27 16:58:54 hanode2 pvesh: <root@pam> starting task UPID:hanode2:00000ED8:00001EEA:53344ABE:startall::root@pam:
Mar 27 16:58:54 hanode2 pvesh: <root@pam> end task UPID:hanode2:00000ED8:00001EEA:53344ABE:startall::root@pam: OK
Mar 27 16:58:54 hanode2 rgmanager[3807]: [pvevm] VM 101 is already stopped
Mar 27 16:58:54 hanode2 rgmanager[3842]: [pvevm] VM 100 is already stopped
Mar 27 16:58:54 hanode2 rgmanager[3841]: [pvevm] VM 102 is already stopped
Mar 27 16:58:54 hanode2 rgmanager[3882]: [pvevm] VM 103 is already stopped
Mar 27 16:58:55 hanode2 rgmanager[2953]: Services Initialized

Mar 27 16:58:55 hanode2 rgmanager[2953]: State change: Local UP
Mar 27 16:58:55 hanode2 rgmanager[2953]: State change: hanode1 UP
Mar 27 16:58:55 hanode2 rgmanager[2953]: Migration: pvevm:100 is running on 1
Mar 27 16:58:55 hanode2 rgmanager[2953]: Migration: pvevm:101 is running on 1
Mar 27 16:58:56 hanode2 rgmanager[2953]: Migration: pvevm:102 is running on 1
Mar 27 16:58:56 hanode2 rgmanager[2953]: Migration: pvevm:103 is running on 1