

Konfiguracja klastra Proxmox

Konfiguracja Klastra Proxmox

Tworzenie klastra

Tworzenie maszyn wirtualnych

Konfiguracja sieci w GNS3

Bonding połączeń sieciowych w proxmox

Konfiguracja VLAN na przełącznikach

Konfiguracja klastra Active - Passive

DRBD

Pacemaker

Wysoka dostępność VM

iSCSI

Przyłączenie zasobu iSCSI

Migracja dysku na zasób iSCSI

Testowa awaria

Jesteś administratorem sieci w firmie X.

Otrzymał(a)eś od przełożonego prośbę, aby mając do dyspozycji 3 nowe serwery (VM1, VM2, VM3), wdrożyć wysoko dostępny klaster dla maszyn wirtualnych zvirtualizowanego środowiska XCP-ng lub Proxmox-VE.

Otrzymałeś również poniższe założenia, jakie mają zostać spełnione dla klastra jaki masz wdrożyć:

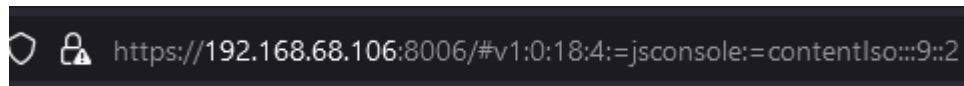
1. Klaster ma zostać wdrożony jako klaster wysokodostępny z wykorzystaniem trzech węzłów. Ma on również wykorzystywać na potrzeby realizacji wysokiej dostępności współdzielony zasób dyskowy w postaci współdzielonej jednostki LUN w ramach udostępnionego celu iSCSI (uruchomionego na osobnym jakimś wybranym systemie). W gotowym klastrze należy utworzyć dwie wirtualne maszyny z systemem Windows Server 2022, i przetestować na nich prawidłowe działanie klastra i mechanizmów wysokiej dostępności.
2. Wdrożony klaster wysokodostępny powinien mieć skonfigurowane połączenie do lokalnej sieci komputerowej na takiej zasadzie, że powinny istnieć dwa osobne połączenia do dwóch różnych przełączników Cisco Catalyst, działających na takiej zasadzie, że gdy jedno podstawowe

połączenie ulegnie awarii, to drugie przejmie jego rolę. Ponadto biorąc pod uwagę, że w firmie są wykorzystywane VLAN'y (VLAN 20, oraz VLAN 30), to klaster środowiska zwirtualizowanego ma obsługiwać te VLAN'y, w taki sposób iż jeden z Windows Server ma być przyłączony do VLAN20, a drugi do VLAN 30.

3. Wdrożyć wysokodostępny cel iSCSI, bazujący na systemie operacyjnym GNU/Linux, na którym będzie udostępniona jednostka LUN wykorzystywana jako współdzielony zasób dyskowy przez utworzony wcześniej klaster środowiska zwirtualizowanego. Tenże klaster pracy awaryjnej iSCSI ma działać na takiej zasadzie, że składać się ma z dwóch węzłów, synchronizujących dane między sobą w locie, natomiast sama usługa iSCSI ma zostać udostępniona jako klaster Active<>Passive (jeżeli jeden węzeł ulegnie awarii, to drugi posiadający wszystkie dane, przejmie jego rolę).

Konfiguracja Klastra Proxmox

Po instalacji proxmoxa na 3 maszynach (każda z nich posiada 2 interfejsy sieciowe), można połączyć się do panelu webowego za pomocą ip serwera.



Domyślnie proxmox korzysta z samopodpisanego certyfikatu ssl. Dlatego widnieje ostrzeżenie o niezabezpieczonym połączeniu.

Tworzenie klastra

Przechodząc w panelu webowym do zakładki "Klaster", można utworzyć klaster. Korzystając z wygenerowanych danych możemy przyłączyć pozostałe maszyny do utworzonego klastra.

Create: Virtual Machine

General

OS

System

Disks

CPU

Memory

Network

Confirm

Node:

proxmox1

VM ID:

100

Name:

win

Resource Pool:

Start at boot:

☐

Start/Shutdown order:

any

Startup delay:

default

Shutdown timeout:

default

Help

Advanced

☒

Back

Next

Create: Virtual Machine

General

OS

System

Disks

CPU

Memory

Network

Confirm

☒ Use CD/DVD disc image file (iso)

Storage:

local

ISO image:

17763.737.190906-2324.rs5_rele

Guest OS:

Type:

Microsoft Windows

Version:

10/2016/2019

☐ Use physical CD/DVD Drive

☐ Do not use any media

Advanced

☒

Back

Next

Create: Virtual Machine

General

OS

System

Disks

CPU

Memory

Network

Confirm

Graphic card:

Default

SCSI Controller:

VirtIO SCSI single

Machine:

Default (i440fx)

Qemu Agent:

☐

Firmware

BIOS:

Default (SeaBIOS)

Add TPM:

☐

Create: Virtual Machine

General

OS

System

Disks

CPU

Memory

Network

Confirm

ide0

Disk

Bandwidth

Bus/Device:

IDE

0

Cache:

Default (No cache)

Storage:

local-lvm

Discard:

☐

Disk size (GiB):

20

IO thread:

☐

Format:

Raw disk image (raw)

SSD emulation:

☐

Backup:

☒

Read-only:

☐

Skip replication:

☐

Async IO:

Default (io_uring)

+

Add

?

Help

Advanced

☒

Back

Next

Create: Virtual Machine

General

OS

System

Disks

CPU

Memory

Network

Confirm

Sockets: 1

Type: x86-64-v2-AES

Cores: 2

Total cores: 2

VCPUs: 2

CPU units: 100

CPU limit: unlimited

Enable NUMA: ☐

CPU Affinity: All Cores

Extra CPU Flags:

Default	- <input type="radio"/> <input checked="" type="radio"/> <input type="radio"/> <input type="radio"/> +	md-clear	Required to let the guest OS know if MDS is mitigated correctly
Default	- <input type="radio"/> <input checked="" type="radio"/> <input type="radio"/> <input type="radio"/> +	pcid	Meltdown fix cost reduction on Westmere, Sandy-, and IvyBridge Intel CPUs
Default	- <input type="radio"/> <input checked="" type="radio"/> <input type="radio"/> <input type="radio"/> +	spec-ctrl	Allows improved Spectre mitigation with Intel CPUs
Default	- <input type="radio"/> <input checked="" type="radio"/> <input type="radio"/> <input type="radio"/> +	ssbd	Protection for "Speculative Store Bypass" for Intel models
Default	- <input type="radio"/> <input checked="" type="radio"/> <input type="radio"/> <input type="radio"/> +	ibpb	Allows improved Spectre mitigation with AMD CPUs

Help

Advanced ☒

Back

Next

Create: Virtual Machine

General

OS

System

Disks

CPU

Memory

Network

Confirm

Memory (MiB): 1024

Minimum memory (MiB): 1024

Shares: Default (1000)

Ballooning Device: ☒

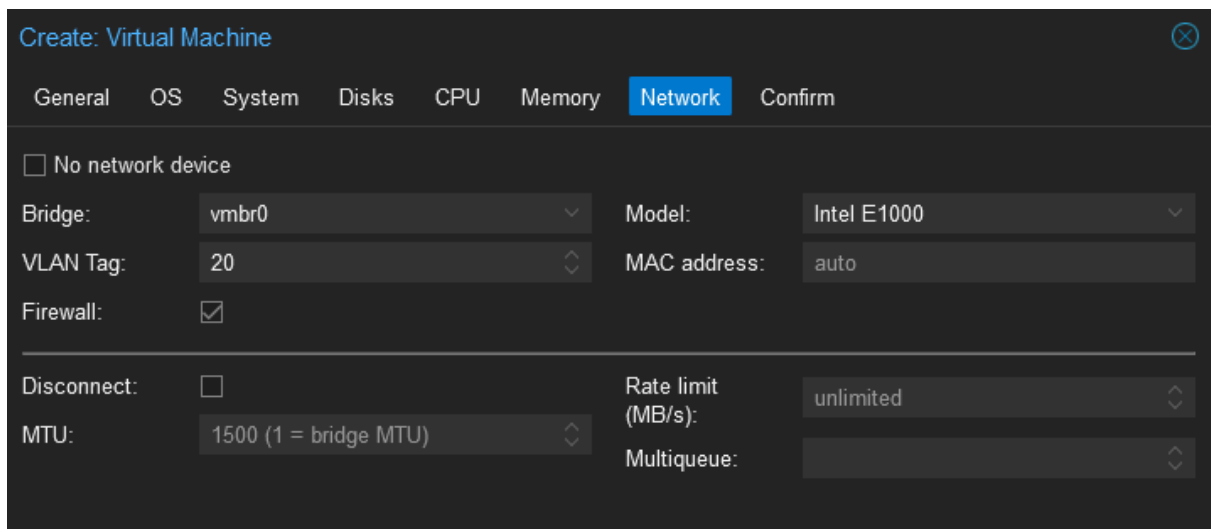
Help

Advanced ☒

Back

Next

podczas tworzenia maszyn można wybrać VLAN do którego ma należeć maszyna:



The screenshot shows the 'Create: Virtual Machine' window with the 'Network' tab selected. The configuration includes:

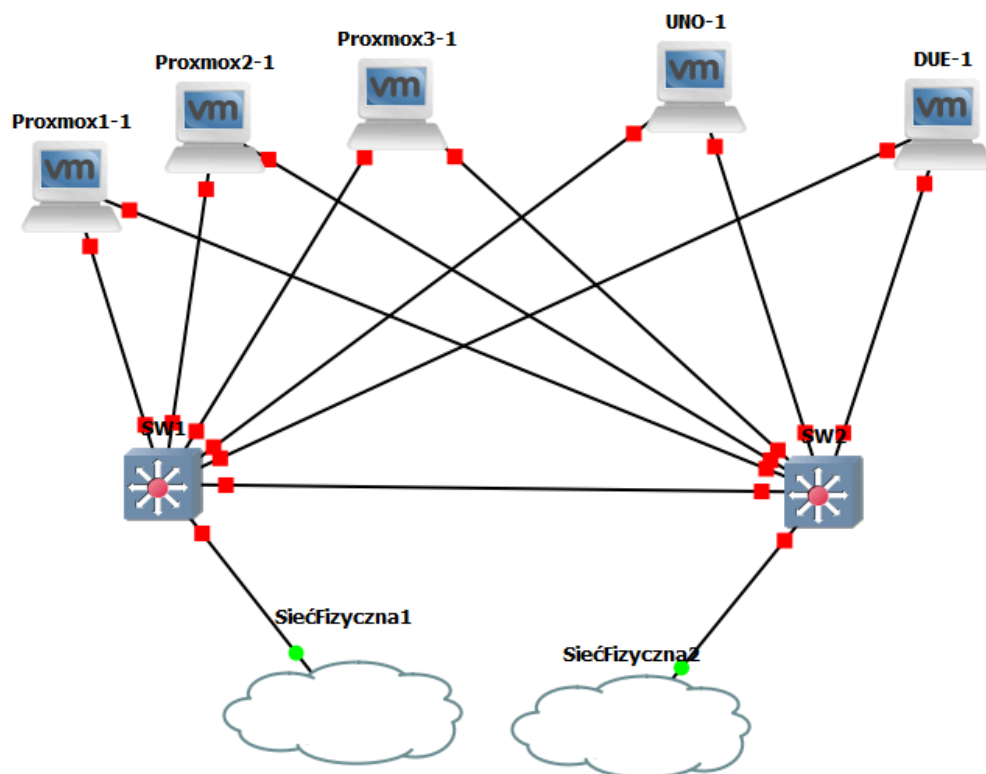
- ☐ No network device
- Bridge: vmbr0
- Model: Intel E1000
- VLAN Tag: 20
- MAC address: auto
- Firewall: ☒
- Disconnect: ☐
- Rate limit (MB/s): unlimited
- MTU: 1500 (1 = bridge MTU)
- Multiqueue: (empty)

Po zakończeniu konfiguracji można przejść do instalacji systemu. Na obu maszynach zainstalowałem windows server core 2019. Dodatkowo pierwsza maszyna (100) jest w VLANie 20, a druga (101) w VLANie 30.

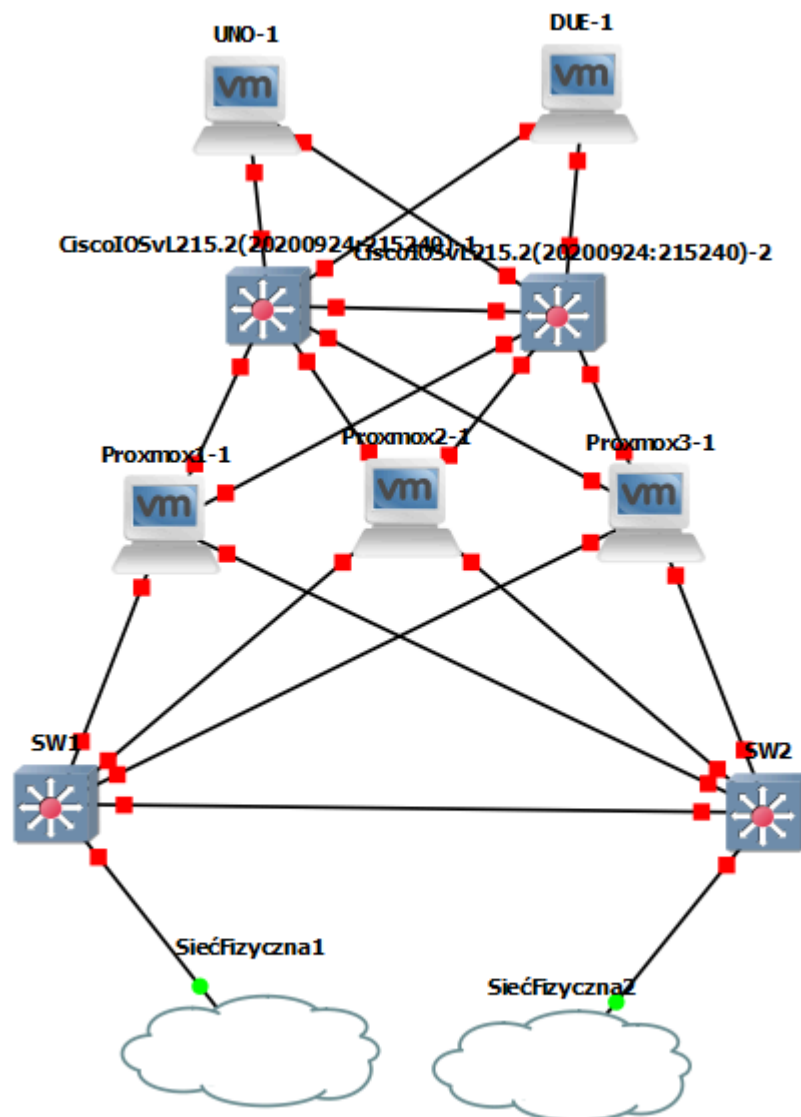
(Utworzenie celu iSCSI i migracja zasobów vm w ostatniej sekcji)

Konfiguracja sieci w GNS3

Utworzona konfiguracja sieci w gns3. Zgodnie z treścią zadania.



Jednakże, w normalnych warunkach, raczej powinien istnieć odrębny segment fizyczny sieci wyznaczony dla iSCSI. Przykładowa konfiguracja:



W takim przypadku klaster z celem iSCSI, jest odłączony od sieci lokalnej. Można też dodać po dwa interfejsy sieciowe do węzłów klastra i podłączyć je do przełączników sieci lokalnej.

Bonding połączeń sieciowych w proxmox

Połączenie dwóch interfejsów sieciowych w trybie active-backup, czyli komunikacja odbywa się tylko na jednym z nich, a drugi funkcjonuje jako interfejs awaryjny.

Edit: Linux Bond

Name:

bond0

IPv4/CIDR:

Gateway (IPv4):

IPv6/CIDR:

Gateway (IPv6):

Autostart:

☒

Slaves:

ens33 ens34

Mode:

active-backup

Hash policy:

bond-primary:

ens33

Comment:

MTU:

1500

Advanced ☒

OK

Reset

Edit: Linux Bridge

Name:

vmbr0

IPv4/CIDR:

192.168.68.106/24

Gateway (IPv4):

192.168.68.1

IPv6/CIDR:

Gateway (IPv6):

Autostart:

☒

VLAN aware:

☒

Bridge ports:

bond0

Comment:

MTU:

1500

Advanced ☒

OK

Reset

Trzeba pamiętać o utworzeniu takiej konfiguracji na każdym z węzłów klastra.

Konfiguracja VLAN na przełącznikach

Informacje, nt. VLAN:

https://www.youtube.com/watch?v=cjFzOnm6u1g&ab_channel=Jeremy'sITLab

https://www.youtube.com/watch?v=Jl9OOzNaBDU&ab_channel=Jeremy'sITLab

https://www.youtube.com/watch?v=OkPB028l2eE&ab_channel=Jeremy'sITLab

trunk porty do klastra proxmox. dla testów można ustawić jakieś VPCs i do nich przypisać access porty z określonym VLANem.

Konfiguracja klastra Active - Passive

Na dwóch systemach debian11 instalacja drdb i pacemaker w celu utworzenia

Aby utworzyć wysokodostępny cel iSCSI należy najpierw utworzyć i zsynchronizować nośnik DRBD między dwoma linuxami (do tego celu wykorzystałem dodatkowe dyski 80GB każdy, dołączone do obu maszyn).

Następnie trzeba utworzyć klastrer z obu maszyn, do tego celu wykorzystałem pacemaker.

Przy takiej konfiguracji można utworzyć wysokodostępny cel iSCSI.

DRBD

Konfiguracja utworzona na podstawie Pana filmów oraz artykułu z linux journal. Odnośniki w ostatniej sekcji.

Pamiętać o utworzeniu wpisów do /etc/hosts dla obu linuxów.

Konfiguracja zasobu

```
UNO DUE
/etc/drbd.d/r0.res [----] 1 L:[ 1+20 21/
resource drbd0 {
on UNO {
device /dev/drbd0;
disk /dev/sdb;
address 192.168.68.132:7788;
meta-disk internal;
}
on DUE {
device /dev/drbd0;
disk /dev/sdb;
address 192.168.68.133:7788;
meta-disk internal;
}
syncer {
rate 100M;
}
net {
shared-secret "Zaq12wsx";
cram-hmac-alg "sha256";
}
}|
```

Pierwsza synchronizacja

```
UNO DUE
Co 1,0s: cat /proc/drbd

version: 8.4.11 (api:1/proto:86-101)
srcversion: 2CC17D07553A98E96473D42
0: cs:SyncSource ro:Primary/Secondary ds:UpToDate/Inconsistent C r-----
   ns:29180800 nr:0 dw:0 dr:29181056 al:8 bm:0 lo:0 pe:1 ua:2 ap:0 ep:1 wo:f oos:54703580
   [=====>.....] sync'ed: 34.8% (53420/81916)M
   finish: 0:56:40 speed: 16,068 (13,744) K/sec
```

```
UNO DUE
root@UNO:~# drbdadm status
drbd0 role:Primary
      disk:UpToDate
      peer role:Secondary
          replication:Established peer-disk:UpToDate

root@UNO:~# |
```

Pacemaker

Instalacja i konfiguracja pacemaker-a na podstawie:

https://clusterlabs.org/pacemaker/doc/deprecated/en-US/Pacemaker/2.0/html/Clusters_from_Scratch/index.html

Na samym starcie wywołanie komendy `#pcs cluster destroy`, aby łatwiej było skonfigurować nowy klaster.

```
root@UNO:~# pcs host auth UNO DUE
Username: hacluster
Password:
UNO: Authorized
DUE: Authorized
root@UNO:~# |
```

```
root@UNO:~# pcs cluster setup pawlaczkowo UNO DUE
No addresses specified for host 'UNO', using 'uno'
No addresses specified for host 'DUE', using 'due'
Destroying cluster on hosts: 'DUE', 'UNO'...
DUE: Successfully destroyed cluster
UNO: Successfully destroyed cluster
Requesting remove 'pcsd settings' from 'DUE', 'UNO'
DUE: successful removal of the file 'pcsd settings'
UNO: successful removal of the file 'pcsd settings'
Sending 'corosync authkey', 'pacemaker authkey' to 'DUE', 'U
DUE: successful distribution of the file 'corosync authkey'
DUE: successful distribution of the file 'pacemaker authkey'
UNO: successful distribution of the file 'corosync authkey'
UNO: successful distribution of the file 'pacemaker authkey'
Sending 'corosync.conf' to 'DUE', 'UNO'
DUE: successful distribution of the file 'corosync.conf'
UNO: successful distribution of the file 'corosync.conf'
Cluster has been successfully set up.
root@UNO:~# |
```

Będzie potrzebne ponowne uruchomienie corosync i pacemaker. po uruchomieniu klastra:

```
root@UNO:~# pcs status
Cluster name: pawlaczkowo

WARNINGS:
No stonith devices and stonith-enabled is not false

Cluster Summary:
* Stack: corosync
* Current DC: UNO (version 2.0.5-ba59be7122) - partition with quorum
* Last updated: Sun Jul  2 22:53:08 2023
* Last change:  Sun Jul  2 22:52:57 2023 by hacluster via crmd on UNO
* 2 nodes configured
* 0 resource instances configured

Node List:
* Online: [ DUE UNO ]

Full List of Resources:
* No resources

Daemon Status:
corosync: active/enabled
pacemaker: active/enabled
pcsd: active/enabled
```

Dodanie ip klastra

```
root@UNO:~# pcs resource create ClusterIP ocf:heartbeat:IPaddr2 \
> ip=192.168.68.200 cidr_netmask=24 op monitor interval=30s
```

```
root@UNO:~# pcs status
Cluster name: pawlaczkowo
Cluster Summary:
  * Stack: corosync
  * Current DC: UNO (version 2.0.5-ba59be7122) - partition with quorum
  * Last updated: Sun Jul  2 23:07:08 2023
  * Last change:  Sun Jul  2 23:06:07 2023 by root via cibadmin on UNO
  * 2 nodes configured
  * 1 resource instance configured

Node List:
  * Online: [ DUE UNO ]

Full List of Resources:
  * ClusterIP   (ocf::heartbeat:IPaddr2):          Started DUE

Daemon Status:
  corosync: active/enabled
  pacemaker: active/enabled
  pcsd: active/enabled
root@UNO:~# ping 192.168.68.200
PING 192.168.68.200 (192.168.68.200) 56(84) bytes of data.
64 bytes from 192.168.68.200: icmp_seq=1 ttl=64 time=0.560 ms
64 bytes from 192.168.68.200: icmp_seq=2 ttl=64 time=0.405 ms
^C
--- 192.168.68.200 ping statistics ---
2 packets transmitted, 2 received, 0% packet loss, time 1031ms
rtt min/avg/max/mdev = 0.405/0.482/0.560/0.077 ms
```

Teoretycznie już mamy klaster typu Active - Passive.

Dodanie drbd jako zasobu klastra.

```

root@UNO:~# pcs status
Cluster name: pawlaczkowo
Cluster Summary:
  * Stack: corosync
  * Current DC: DUE (version 2.0.5-ba59be7122) - partition with quorum
  * Last updated: Sun Jul  2 23:38:05 2023
  * Last change:  Sun Jul  2 23:37:02 2023 by root via cibadmin on UNO
  * 2 nodes configured
  * 4 resource instances configured

Node List:
  * Online: [ DUE UNO ]

Full List of Resources:
  * ClusterIP      (ocf::heartbeat:IPaddr2):      Started DUE
  * Clone Set: p_drbd_drbd0-clone [p_drbd_drbd0] (promotable):
    * p_drbd_drbd0      (ocf::linbit:drbd):      Starting UNO
    * p_drbd_drbd0      (ocf::linbit:drbd):      Starting DUE
  * drbdfs          (ocf::heartbeat:Filesystem):      Stopped

Daemon Status:
  corosync: active/enabled
  pacemaker: active/enabled
  pcsd: active/enabled
root@UNO:~# |

```

Pamiętać o połączeniu interfejsów sieciowych jak na maszynach "proxmox"

Wysoka dostępność VM

iSCSI

Udostępnienie zasobu iSCSI:

```

/backstores/fileio> create disk1 /dane/disk1.img 70G
Created fileio disk1 with size 75161927680

```

```

/backstores/fileio> cd /iscsi
/iscsi> create iqn.2023-06.pawlak:storage.debiantarget01
Created target iqn.2023-06.pawlak:storage.debiantarget01.
Created TPG 1.
Global pref auto_add_default_portal=true
Created default portal listening on all IPs (0.0.0.0), port 3260.

```

```

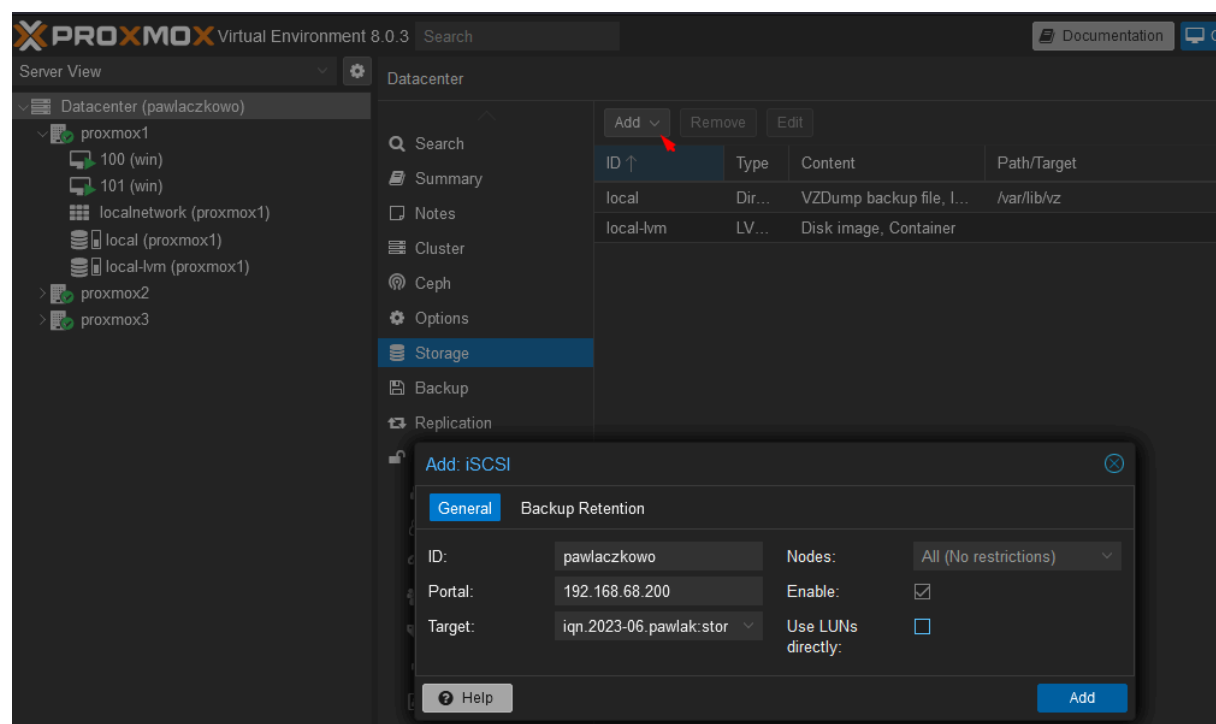
/iscsi/iqn.20...t01/tpg1/luns> create /backstores/fileio/disk1
Created LUN 0.

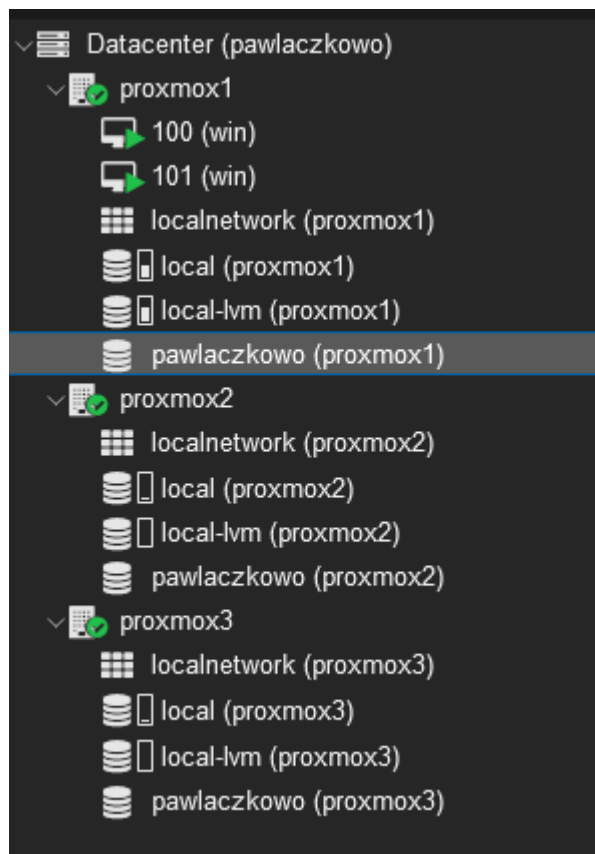
```


Pobranie wszystkich iqn inicjatora z pliku `/etc/iscsi/initiatorname.iscsi`, i przekazanie ich do listy acl

```
/iscsi/iqn.20...t01/tpg1/acls> ls /
o- / ..... [..]
o- backstores ..... [..]
o- block ..... [Storage Objects: 0]
o- fileio ..... [Storage Objects: 1]
o- disk1 ..... [/dane/disk1.img (70.0GiB) write-back activated]
o- alua ..... [ALUA Groups: 1]
o- default_tg_pt_gp ..... [ALUA state: Active/optimized]
o- pscsi ..... [Storage Objects: 0]
o- ramdisk ..... [Storage Objects: 0]
o- iscsi ..... [Targets: 1]
o- iqn.2023-06.pawlak:storage.debian:target01 ..... [TPGs: 1]
o- tpg1 ..... [no-gen-acls, no-auth]
o- acls ..... [ACLs: 3]
o- iqn.1993-08.org.debian:01:612fb8a693c2 ..... [Mapped LUNs: 1]
o- mapped_lun0 ..... [Lun0 fileio/disk1 (rw)]
o- iqn.1993-08.org.debian:01:dc36418a8918 ..... [Mapped LUNs: 1]
o- mapped_lun0 ..... [Lun0 fileio/disk1 (rw)]
o- iqn.1993-08.org.debian:01:f91359391cd8 ..... [Mapped LUNs: 1]
o- mapped_lun0 ..... [Lun0 fileio/disk1 (rw)]
o- luns ..... [LUNs: 1]
o- lun0 ..... [fileio/disk1 (/dane/disk1.img) (default_tg_pt_gp)]
o- portals ..... [Portals: 1]
o- 0.0.0.0:3260 ..... [OK]
o- loopback ..... [Targets: 0]
o- vhost ..... [Targets: 0]
o- xen-pvscsi ..... [Targets: 0]
```

Przyłączenie zasobu iSCSI



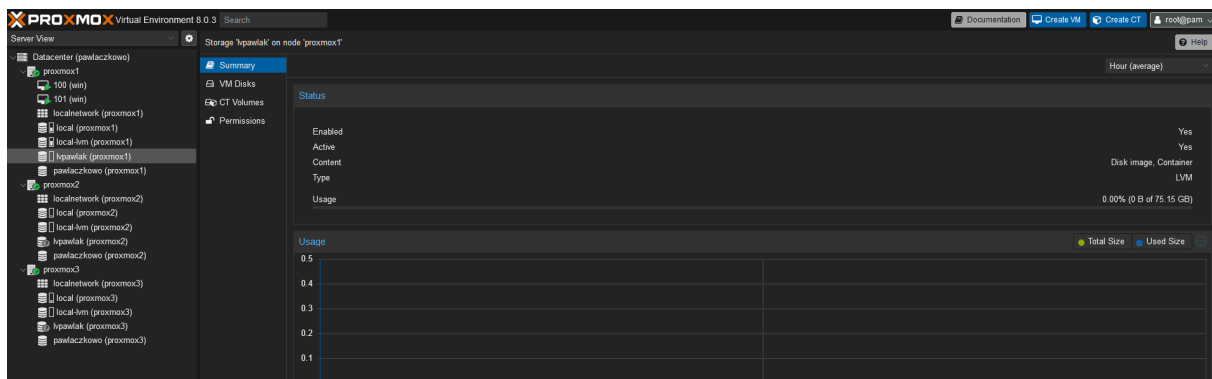


Add: LVM ⓧ

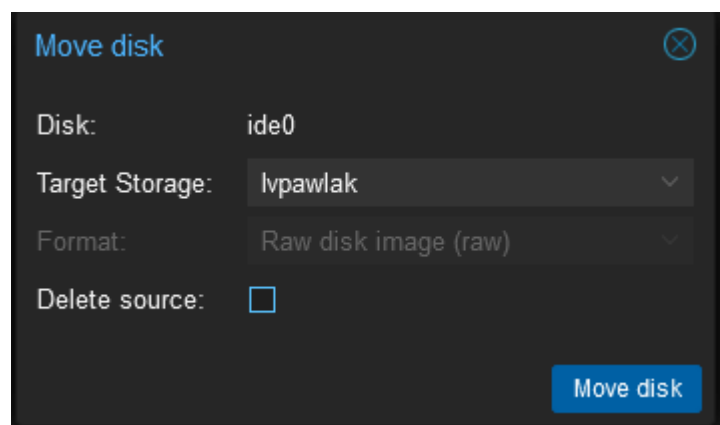
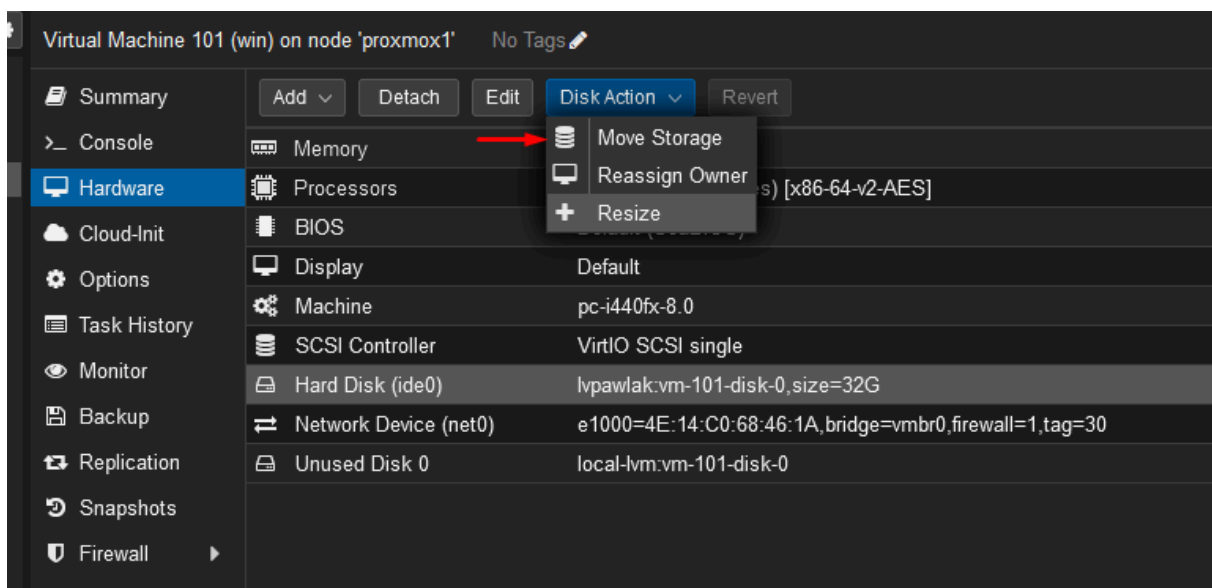
General Backup Retention

ID:	lvpawlak	Nodes:	All (No restrictions) ▾
Base storage:	pawlaczkowo (iSCSI) ▾	Enable:	<input checked="" type="checkbox"/>
Base volume:	CH 00 ID 0 LUN 0 ▾	Shared:	<input checked="" type="checkbox"/>
Volume group:	lvpawlak		
Content:	Disk image, Container ▾		

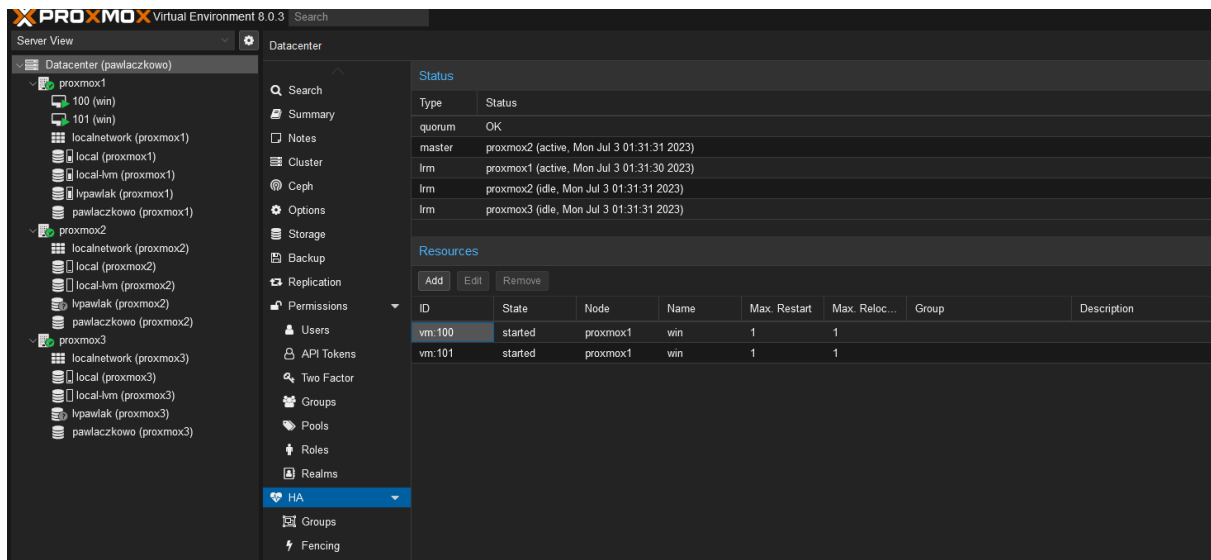
ⓘ Help Add



Migracja dysku na zasób iSCSI

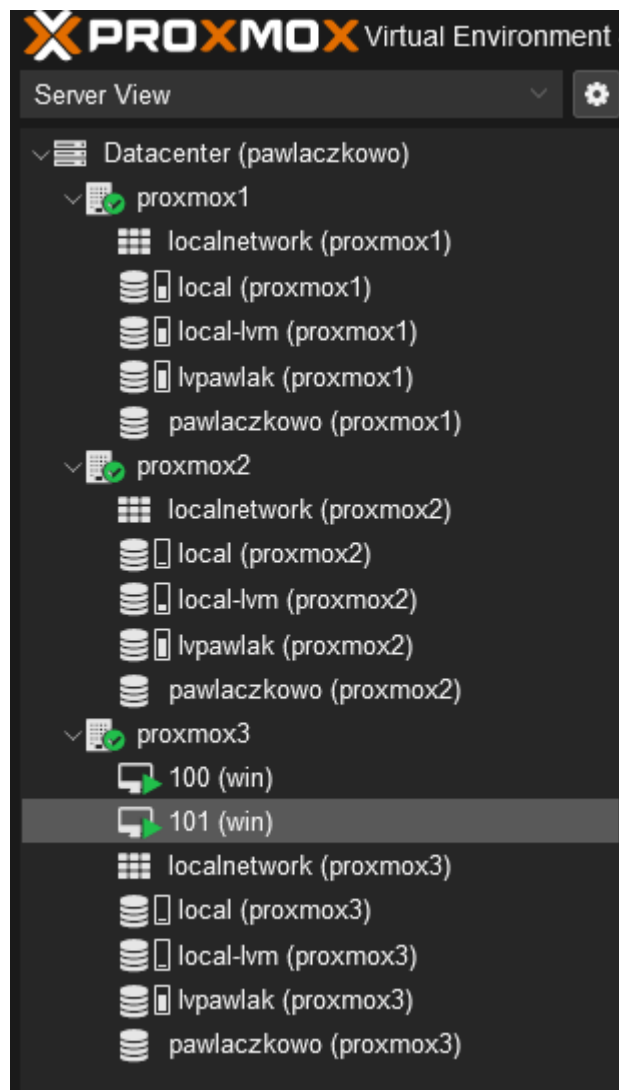


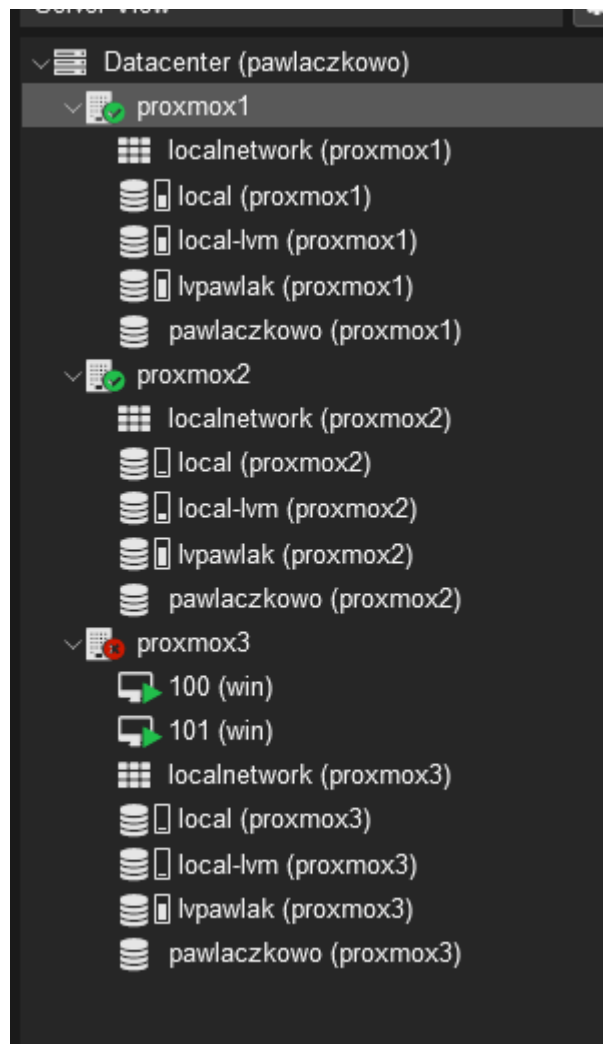
Trzeba pamiętać jeszcze o dodaniu tych maszyn do sekcji "HA"



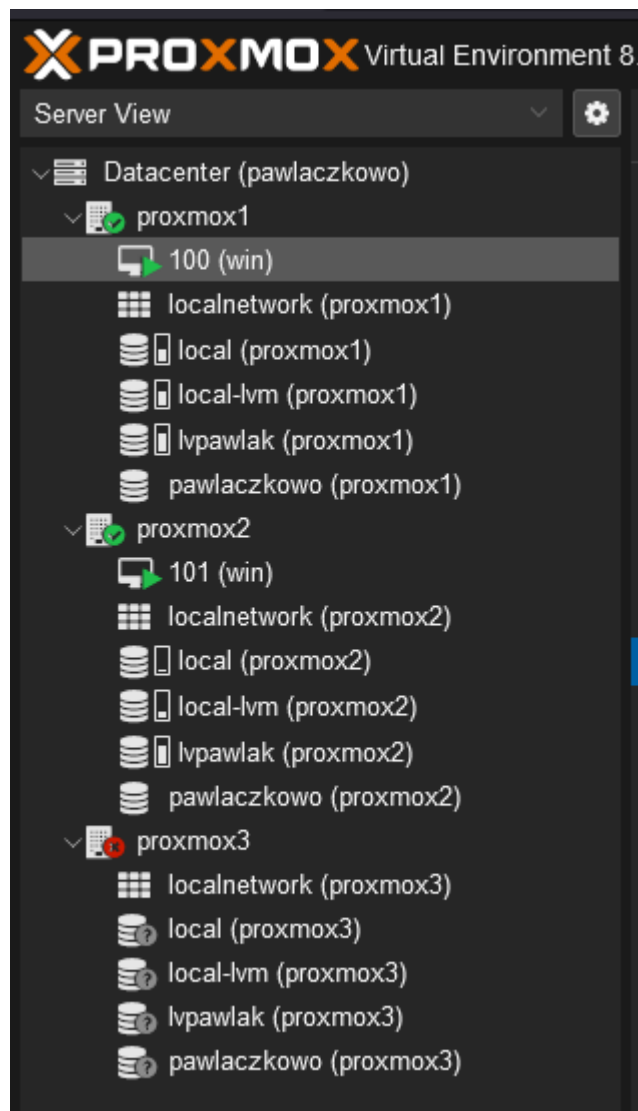
Testowa awaria

Wyłączenie proxmox3 w trakcie migracji maszyny.





Stan już po chwili



Działą

