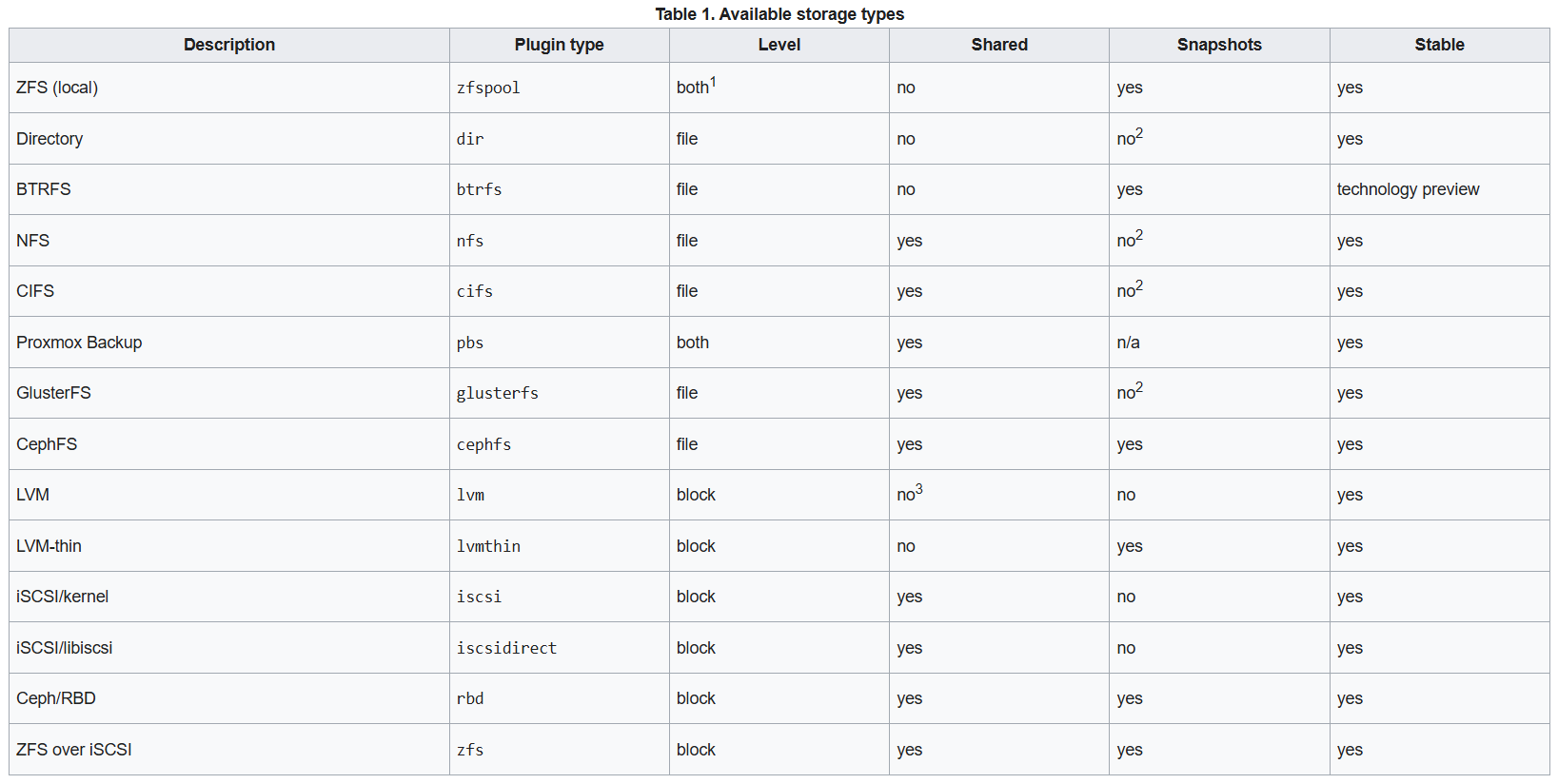
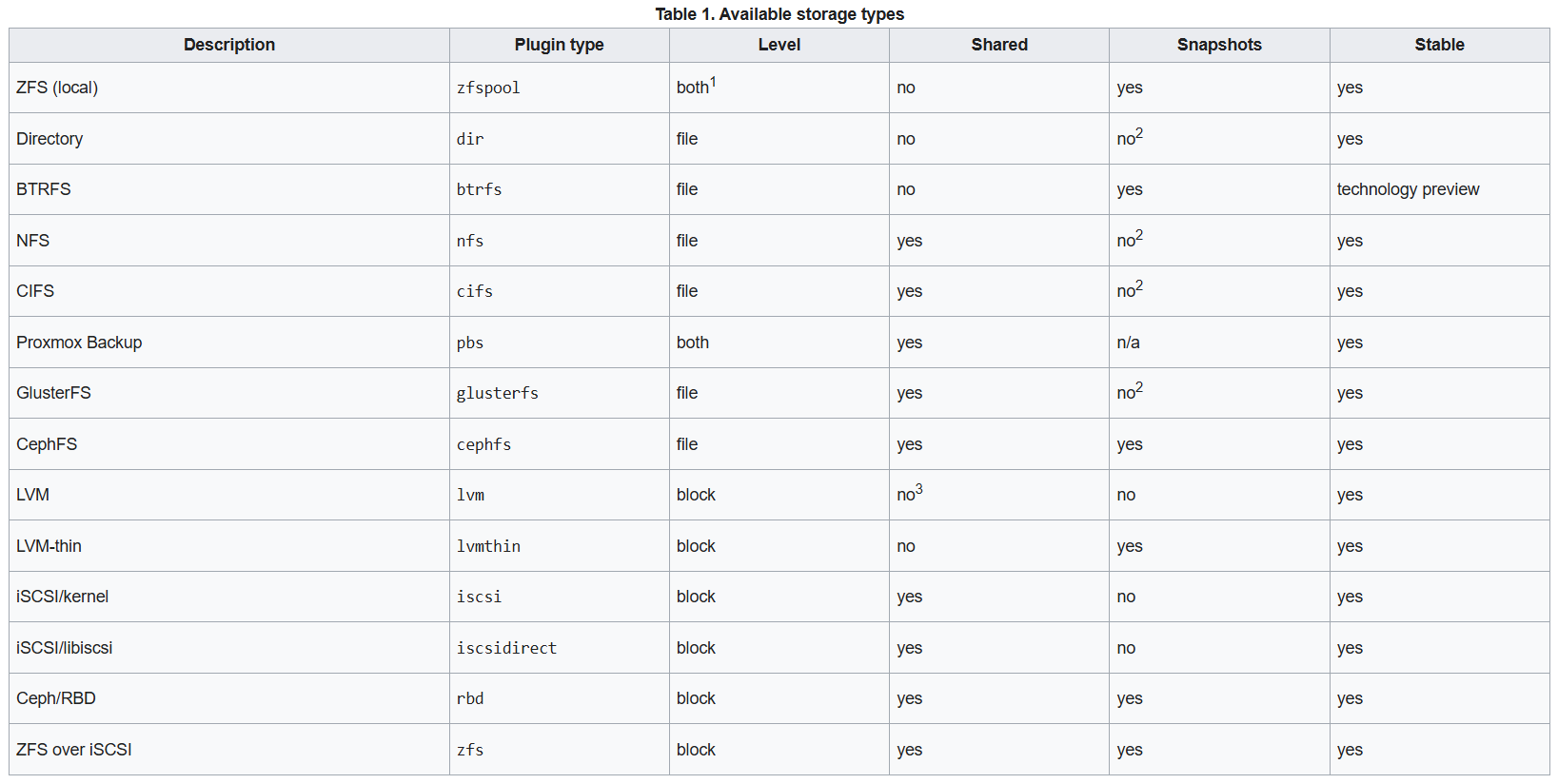
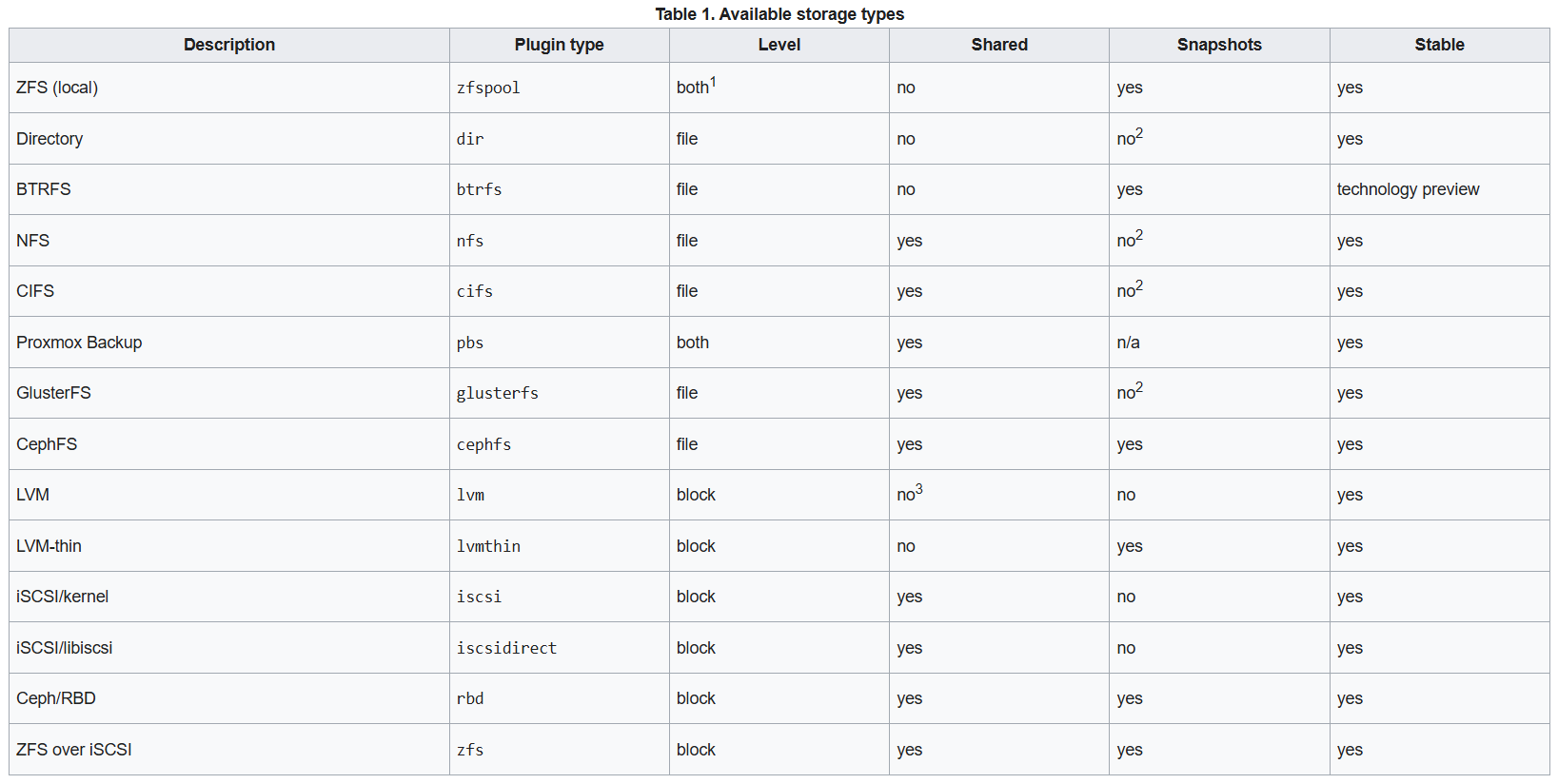
**Armazenamento Proxmox**

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
|  | **Directory (Diretoria)** | **LVM (Logical Volume Manager)** | **LVM-Thin (Thin Provisioning)** | **ZFS** | **Ceph** |
| **Descrição** | Armazena os dados numa pasta normal dentro de um file system (ext4, xfs). | Cria volumes lógicos dentro de um grupo de volume (VG), aloca espaço fixo. | Similar ao LVM, mas permite alocação dinâmica. | Sistema de ficheiros com gestão de volume integrado | Armazenamento distribuído para clusters, utilizado em ambientes Proxmox HA. |
| **Vantagens** | Fácil de configurar  Suporte para backups | Melhor desempenho que diretorias | Melhora uso de espaço  Snapshots rápidos e eficientes  Ideal para VMs | Snapshots eficientes  RAID integrado  Verificação de integridade automática | Redundância automática  Boa escalabilidade |
| **Desvantagens** | Sem suporte nativo para snapshots | Sem snapshots | Risco de sobrecarga se o espaço físico acabar  Mais lento que LVM normal | Alto consumo de RAM  Menor desempenho em algumas situações | Requer múltiplos nós  Complexo de configurar |
| **Thin Provisioning** | Suporta thin provisioning no formato de imagem QEMU (qcow2). | Não suporta thin provisioning. | Suporta thin provisioning. | Suporta thin provisioning. | Suporta thin provisioning. |
| **Snapshots** | Não suporta snapshots | Não suporta snapshots. | Suporta Snapshots | Suporta Snapshots | Suporta Snapshots |



|  |  |  |
| --- | --- | --- |
| **Característica** | **Ceph RBD (RADOS Block Device)** | **CephFS (Ceph File System)** |
| Tipo de armazenamento | Block Storage | File Storage |
| Uso principal | Armazenar discos de VMs e containers | Armazenar ficheiros, ISOs, backups |
| Desempenho | Alto desempenho para VMs | Mais lento para VMs (acesso via file system) |
| Snapshots | Suportado | Não tem suporte nativo |
| Thin Provisioning | Sim (permite alocação dinâmica de espaço) | Sim, mas menos eficiente para VMs |
| Live Migration | Sim, rápido e sem cópias | Não recomendado para VMs |
| Alta disponibilidade (HA) | Sim, distribuído entre os nós | Sim, mas mais indicado para ficheiros |
| Uso no Proxmox | Armazena discos de VMs de forma otimizada | Armazena ISOs, backups e ficheiros partilhados |

**Ainda não testámos o Ceph**

* pve-root(local criado por omissão)  
  O pve-root é um volume lógico (LV) do LVM dentro do grupo de volume (VG) chamado pve.
* vgs-> lista todos os grupos de volumes com os seus atributos
* lvs-> lista todos os volumes lógicos (LVs) no sistema com os seus atributos

|  |  |  |  |
| --- | --- | --- | --- |
| **#PV** | | |  | | --- | | Número de **Physical Volumes (PV)**, ou seja, quantos discos ou partições estão alocados nesse VG. | |
| **#LV** | |  | | --- | | Número de **Logical Volumes (LVs)** dentro do VG. Cada LV é como uma "partição" dentro do LVM. | |
| **#SN** | |  | | --- | | Número de **snapshots** criados dentro do VG. | |

**Comandos usados:**

* lvextend -l +100%FREE /dev/pve/root -> expandir o disco
* resize2fs /dev/pve/root

**Criar uma partição**

* fdisk /dev/sdx
* n → Nova partição
* primária (p)
* último setor (ex:+100G)
* w(salvar)

**Criar um Volume Group (VG)**

* vgcreate nomeVolume /dev/sdx

**Criar um LVM-Thin dentro do VG**

* lvcreate --name nomePool --size 50G --type thin-pool nomeVolume

**Criar Logical Volumes**

* lvcreate -l 100%FREE --name testDirect vg\_storage
* mkfs.ext4 /dev/nomeVolume/lv\_nome -> diretoria

**Editar repositórios Proxmox**

* nano /etc/apt/sources.list.d/pve-enterprise.list
* nano /etc/apt/sources.list.d/ceph.list

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  | **IPs** | **Disco Principal (sda)** | **Disco Secundário (sdb)** | **RAM** |
| **PC1** | 172.79.79.2 | 250GB | 250GB | 8GB |
| **PC2** | 172.79.79.3 | 500GB | 500GB | 8GB |
| **PC3** | 172.79.79.4 | 350GB | 250GB | 8GB |

**Especificações dos servidores:**

|  |  |  |  |
| --- | --- | --- | --- |
| Característica | ZFS | Ceph | NFS |
| Live Migration | Não (Apenas com replicação e desligamento da VM) | Sim (Storage distribuído) | Sim (Armazenamento compartilhado) |
| Alta Disponibilidade (HA) | Não (Cada nó tem seu próprio storage) | Sim (Storage acessível por todos os nós) | Sim (Se o servidor NFS tiver redundância) |
| Thin Provisioning | Sim, mas precisa ativar sparse=1 | Sim (Automaticamente thin provisioned) | Sim (Automaticamente thin provisioned) |
| Snapshots Rápidos | Sim (Integrado no ZFS) | Sim, mas com impacto na performance | Sim, mas pode ser mais lento que ZFS |
| Redundância de Dados | Sim (via RAID-Z ou Mirror) | Sim (Os dados são replicados entre os nós) | Não (Depende do storage físico) |
| Escalabilidade | Limitado ao armazenamento local | Alta (Facilmente escalável adicionando nós) | Média (Depende do servidor NFS) |
| Requisitos de RAM | 8GB mínimo, recomendado 16GB+ | 16GB mínimo, recomendado 32GB+ | Baixo (4GB é suficiente) |
| Requisitos de Rede | 1Gbps suficiente | 10Gbps recomendado para bom desempenho | 1Gbps pode funcionar, mas 10Gbps recomendado |
| Complexidade de Configuração | Média | Alta | Fácil |

**8/04**

Uma imagem com captura de ecrã, Teclas

Os conteúdos gerados por IA poderão estar incorretos.

|  |  |
| --- | --- |
| **Campo** | **Explicação** |
| **ID** | Identificador do OSD. Tens 3: 0, 1 e 2. |
| **CLASS** | Tipo de dispositivo. Neste caso, todos são hdd (discos rígidos). |
| **WEIGHT** | Peso do OSD no algoritmo CRUSH. Indica quantos dados ele deve receber. |
| **REWEIGHT** | Reajuste manual do peso. Está a 1.00000 para todos (valor padrão). |
| **SIZE** | Tamanho total físico do disco. Tens dois discos de 233 GiB e um de 466 GiB. |
| **RAW USE** | Espaço total físico utilizado. Todos usam cerca de 21 GiB. |
| **DATA** | Dados reais armazenados (sem metadados/omap). |
| **OMAP** | Objetos armazenados no OMAP (útil para RADOS/cephfs). Pouco espaço aqui. |
| **META** | Metadados usados pelo Ceph. Varia entre 680 MiB e 1.2 GiB. |
| **AVAIL** | Espaço disponível. Os de 233 GiB têm ~212 GiB livres; o de 466 GiB tem ~445 GiB. |
| **%USE** | Percentagem de utilização de cada OSD. |
| **VAR** | Variação de uso comparado com a média (ajuda no balanceamento). |
| **PGS** | Placement Groups atribuídos ao OSD. |
| **STATUS** | Estado do OSD |

O **volume de bloco virtual** é um dispositivo que foi associado ao sistema operativo. Ou seja, a imagem RBD associada é um **volume de disco lógico**, armazenado no Ceph, mas que se comporta como um **dispositivo de bloco local** dentro do sistema.

**# Cria uma imagem RBD com 100 GiB**

rbd create Storage1/test-image1 --size 100G

**# Isto torna a imagem visível como um dispositivo de bloco**  
rbd map Storage1/test-image1

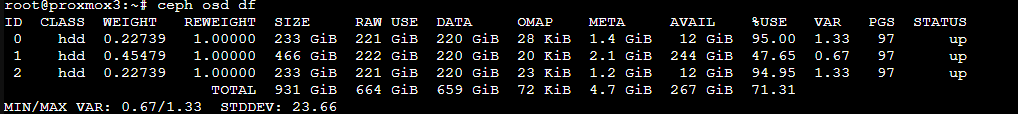
**# Formatar**  
mkfs.ext4 /dev/rbd0

**# Montar o volume**  
mkdir /mnt/test-rbd

mount /dev/rbd0 /mnt/test-rbd  
  
**# Usar /dev/zero para escrever dados reais:**  
dd if=/dev/zero of=/mnt/rbd-test/fill.img bs=1G count=100 status=progresso

**Teste storage**

**Uma imagem com texto, captura de ecrã, Gráfico, file

Os conteúdos gerados por IA poderão estar incorretos.Uma imagem com captura de ecrã, Teclas

Os conteúdos gerados por IA poderão estar incorretos.Objetivo:** Verificar o comportamento do Ceph quando os discos atingem a capacidade máxima.

**Antes do teste**



**Conclusão:** À medida em que os dados sendo armazenados nos OSDs, verificamos que o tamanho total dos mesmos diminui simultaneamente entre os 3 OSDs e com este resultado confirmámos que está a ser armazenado uma réplica em cada.



**Depois do teste**

**Teste disco**

**Objetivo:** Verificar o comportamento do ceph quando um disco falha. **Uma imagem com texto, captura de ecrã, software, Software de multimédia

Os conteúdos gerados por IA poderão estar incorretos.**

**Uma imagem com texto, captura de ecrã, Software de multimédia, Software gráfico

Os conteúdos gerados por IA poderão estar incorretos.Conclusão:** Ao removermos o cabo de alimentação do disco, o OSD do proxmox3 ficou down/in e passado uns minutos passa para down/out. O OSD está inativo e foi removido logicamente do cluster, o que significa que o Ceph vai começar a redistribuir os dados para outros OSDs e não vai usar mais esse OSD para novos dados.

Como podemos ver na imagem abaixo o armazenamento total passou para os 434.99GB, porque agora tem dois ODS em funcionamento. Portanto só consegue garantir duas réplicas dos dados.

Depois conectar o disco com a máquina em funcionamento, a partir da saída do comando “dmesg | grep -i sd”, podemos observar que o disco sdb (250 GB) foi reconhecido e apresentou um problema de I/O (entrada/saída) durante a escrita e erros relacionados ao Ceph.

Solução: reiniciámos a máquina, o OSD passado uns minutos voltou ao estado up/in. Podemos observar que o Ceph ainda não tinha processado completamente, mostrando valores do armazenamento elevados, mas depois foram normalizados.

Uma imagem com captura de ecrã, texto, Software de multimédia, Software gráfico

Os conteúdos gerados por IA poderão estar incorretos.

Valores ao reiniciar

Uma imagem com captura de ecrã, texto

Os conteúdos gerados por IA poderão estar incorretos.

Valores normalizados

Uma imagem com texto, captura de ecrã, Tipo de letra

Os conteúdos gerados por IA poderão estar incorretos.

Próximos testes:

* Remover 2 discos (em princípio vai ser possível ler os dados, mas não vai ser possível escrever).
* Remover 1 disco (já fizemos, mas temos de repetir)
* Colocar um novo disco.
* …………………..

down/in: O OSD está inativo (não responde), mas ainda está dentro do cluster e é considerado para armazenamento de dados.

down/out: O OSD está inativo e foi removido logicamente do cluster, o que significa que o Ceph vai começar a redistribuir os dados para outros OSDs e não vai usar mais esse OSD para novos dados.