

# Diagnóstico y resolución de fallos en snapshots LVM durante backups en Proxmox

## Contexto del problema

### Síntomas iniciales

Durante la ejecución de backups automatizados en Proxmox 7.2.14, se producían fallos intermitentes al procesar el disco secundario (disco-1) de la VM 473, lo que causaba error en el tamaño de la copia.

```
Processing backup of disk 473-0
Transferring: kvm-473-0.2025-10-27-05-41-56.gz
25442648064 bytes (25 GB, 24 GiB) copied, 129 s, 197 MB/s
dd: error reading '/dev/lvm/snap-473-0': Input/output error
3051+1 records in
3051+1 records out
25594953728 bytes (26 GB, 24 GiB) copied, 130.739 s, 196 MB/s
```

## Características del sistema

- **Proxmox VE:** 7.2.14
- **Kernel:** 5.15.74-1-pve
- **Storage:** LVM sobre NVMe
- **Método backup:** Scripts personalizados (no vzdump estándar)
- **Patrón:** Primera copia OK, segunda copia FAIL

## Proceso de análisis

### Fase 1: Verificación de espacio en VG

Descartado rápidamente. El Volume Group tenía espacio abundante:

```
pvs -o +pv_name,vg_name,pv_size,pv_free
PV          VG   Fmt  Attr PSize  PFree
/dev/nvme1n1p1 lvm  lvm2 a--  <1.75t  757.49g
```

**Conclusión:** 757GB libres, suficiente para snapshots.

## Fase 2: Revisión de logs del sistema

```
dmesg -T | grep -i "error\|fail\|i/o" | tail -50
```

### Hallazgos:

```
[Mon Oct 27 05:38:02 2025] Buffer I/O error on dev dm-54, logical block 5923136, async page read
[Mon Oct 27 05:38:02 2025] Buffer I/O error on dev dm-54, logical block 5923136, async page read
[Mon Oct 27 05:38:02 2025] Buffer I/O error on dev dm-54, logical block 5923136, async page read
[Mon Oct 27 05:44:05 2025] Buffer I/O error on dev dm-51, logical block 6248768, async page read
```

- Errores en dispositivos **dm-XX** (device mapper → snapshots LVM)
- Tres reintentos por operación (patrón típico del kernel)
- Tipo: **async page read** (lectura desde snapshot)

**Sospecha inicial:** Problema de hardware en disco NVMe.

## Fase 3: Verificación de salud del disco NVMe

```
smartctl -a /dev/nvme1n1p1
```

### Resultado crítico:

```
SMART overall-health self-assessment test result: PASSED
Media and Data Integrity Errors:      0
Error Information Log Entries:        0
Available Spare:                      100%
Percentage Used:                      8%
```

**Conclusión:** El disco NVMe está **perfectamente sano**. No es problema de hardware.

## Fase 4: Análisis de device mapper

Intentamos localizar los dispositivos dm-54 y dm-51 reportados en dmesg:

```
dmsetup ls --tree | grep -E "dm-5[14]"  
ls -la /dev/dm-54 /dev/dm-51  
# cannot access '/dev/dm-54': No such file or directory  
# cannot access '/dev/dm-51': No such file or directory
```

**Interpretación:** Los snapshots temporales ya fueron destruidos tras el backup (comportamiento normal). Los errores ocurrieron durante la existencia efímera del snapshot.

## Hipótesis evaluadas

### Hipótesis 1: Sectores defectuosos en disco físico

- **Estado:**  DESCARTADA
- **Motivo:** SMART sin errores, 0 sectores realojados

### Hipótesis 2: Bug del kernel 5.15 con snapshots bajo carga

- **Estado:**  POSIBLE (no confirmada)
- **Motivo:** Kernel antiguo con bugs conocidos en snapshot mechanism

### Hipótesis 3: Snapshot COW lleno durante backup

- **Estado:**  CONFIRMADA
- **Método de validación:** Monitorización en tiempo real

## Diagnóstico definitivo

# Herramienta de diagnóstico utilizada

Durante la ejecución del backup, se monitorizó en tiempo real el estado de los snapshots:

```
watch -n1 'lvs -a | grep snap'
```

## Output revelador:

```
Every 1.0s: lvs -a | grep
snap
  snap-473-0    lvm  swi-aos---  1.00g    vm-473-disk-0  7.61
  snap-473-1    lvm  swi-I-s---  1.00g    vm-473-disk-1 100.00
  snap-999-0    vg   swi-a-s---  3.00g    vm-999-disk-0  9.09
```

# Interpretación de los flags LVM

Campo	Valor	Significado
snap-473-0	swi-aos---	Snapshot activo, OK (7.61% uso)
snap-473-1	swi-I-s---	Snapshot <b>INVÁLIDO</b> (letra I mayúscula)
snap-473-1	100.00	<b>COW al 100%</b> - completamente lleno

## Conclusión definitiva:

El snapshot del disco-1 alcanzó el 100% de su espacio Copy-On-Write (1GB), provocando:

1. Marcado del snapshot como inválido por LVM
2. Cualquier lectura posterior genera "Buffer I/O error"
3. Fallo del proceso de backup

# Causa raíz

## ¿Por qué el disco-1 y no el disco-0?

```
snap-473-0: 7.61% usado → VM escribe poco en este disco durante backup
snap-473-1: 100% usado → VM escribe intensivamente en este disco
```

El **disco-1** contiene elementos con alta tasa de escritura:

- Bases de datos con logs activos
- Swap en uso
- Directorios /tmp o /var con actividad constante
- Logs del sistema operativo

# El problema del tamaño de snapshot

## Configuración problemática:

```
# Snapshot de 1GB para disco de 35GB = ~3% del tamaño del disco  
lvcreate -L1G -s -n snap-473-1 /dev/lvm/vm-473-disk-1
```

Durante el backup (que puede tardar 2-3 minutos con 35GB):

- VM continúa funcionando
- Disco-1 recibe escrituras constantes
- Cada bloque modificado consume espacio COW
- 1GB se llena antes de completar el backup → snapshot inválido

# Solución implementada

## Ajuste del tamaño de snapshot en scripts de backup

Al ser un proceso solapado de hosts *snapshots* con creación aleatoria de marca de tiempo, y de *mysqldump* en el KVM, el cual por tamaños era causa del problema (> 6G) no procede el aumento tan significativo de **porcentaje de snapshot**

Dado que hay incidentes similares, y ya había algunas modificaciones al script de backups de mysql, [mysqldump\\_remote\\_utility.sh](#) se procede a realizar una mejora en el script, para poder realizar los backup al vuelo, (directamente mysqldump + compresión sobre la conexión).

Consultar [README\\_mysqldump\\_remote.md](#)

# Datos relevantes para otros cálculos

## Cálculo del tamaño óptimo de snapshot

Tamaño disco	Actividad VM	Snapshot recomendado
--------------	--------------	----------------------

35GB	Baja (logs, configs)	10% = 3.5GB
35GB	Media (servicios web)	20% = 7GB
35GB	Alta (BD, swap activo)	30% = 10GB
100GB+	Alta	20-30GB (límite práctico)

## Verificación del tamaño durante backup

Añadir monitorización al script de backup:

```
#!/bin/bash
# En paralelo durante el backup
watch -n5 "lvs -a | grep snap-473 | awk '{print \$1, \$6}'"

# 0 con alertas:
while true; do
    USAGE=$(lvs --noheadings -o snap_percent snap-473-1 2>/dev/null | xargs)
    if (( $(echo "$USAGE > 80" | bc -l) )); then
        echo "WARNING: Snapshot at ${USAGE}%"
    fi
    sleep 5
done
```

## Comandos útiles para diagnóstico

### Monitorización en tiempo real

```
# Ver estado de snapshots activos
watch -n1 'lvs -a | grep snap'

# Ver uso específico con más detalle
lvs -a -o +chunksize,snap_percent | grep snap

# Logs en tiempo real
dmesg -wT | grep -i "error\|snapshot"
```

## Post-mortem tras fallo

```
# Revisar logs del kernel
dmesg -T | grep -i "error\|fail\|i/o" | tail -50

# Estado de los VG
vgs -o +lv_metadata_size,metadata_percent

# Snapshots huérfanos o colgados
lvs -a | grep snap

# Salud del storage
smartctl -a /dev/nvmeXnY
```

## Lecciones aprendidas

1. **Los errores de I/O en dm-XX no siempre son hardware:** Primero verificar el estado de snapshots LVM.
2. **El tamaño de snapshot no es porcentaje fijo:** Depende de la tasa de escritura durante el backup, no del tamaño del disco.
3. **Herramienta clave:** `watch -n1 'lvs -a | grep snap'` permite ver en tiempo real el llenado del COW.
4. **Pattern recognition:** "Primera copia OK, segunda FAIL" → indica recurso que se agota (no hardware defectuoso).
5. **Flag 'I' en lvs:** Snapshot inválido por COW lleno. Crítico para diagnóstico rápido.

## Prevención futura

### Script de validación pre-backup

```
#!/bin/bash
# validate_snapshot_size.sh

VM_ID=$1
DISK_NUM=$2
DISK_LV="/dev/lvm/vm-{$VM_ID}-disk-{$DISK_NUM}"

# Obtener tamaño del disco
DISK_SIZE_G=$(lvs --noheadings -o lv_size --units g "$DISK_LV" | tr -d 'g' | xargs)
```

```

# Calcular snapshot recomendado (30%)
RECOMMENDED_G=$(echo "$DISK_SIZE_G * 0.3 / 1" | bc)

# Obtener espacio libre en VG
VG_NAME=$(lvs --noheadings -o vg_name "$DISK_LV" | xargs)
FREE_G=$(vgs --noheadings -o vg_free --units g "$VG_NAME" | tr -d 'g' | xargs)

echo "Disco: $DISK_LV"
echo "Tamaño: ${DISK_SIZE_G}G"
echo "Snapshot recomendado: ${RECOMMENDED_G}G"
echo "Espacio libre en VG: ${FREE_G}G"

if (( $(echo "$FREE_G < $RECOMMENDED_G" | bc -l) )); then
    echo "ERROR: Espacio insuficiente en VG"
    exit 1
fi

echo "OK: Suficiente espacio disponible"

```

## Monitorización automatizada

```

#!/bin/bash
# monitor_snapshots.sh
# Ejecutar en paralelo durante backups

THRESHOLD=85 # Alertar al 85% de uso

while true; do
    lvs -a --noheadings -o lv_name,snap_percent 2>/dev/null | grep snap | while read name
percent; do
        if [ -n "$percent" ]; then
            percent_int=$(echo "$percent" | cut -d'.' -f1)
            if [ "$percent_int" -ge "$THRESHOLD" ]; then
                logger -t backup-monitor "WARNING: Snapshot $name at ${percent}%"
                # Opcional: enviar alerta email/telegram
            fi
        fi
    fi
fi

```

```
done
sleep 10
done
```

# Referencias

- LVM Snapshot documentation: `/usr/share/doc/lvm2/`
- Device mapper errors: `man dmsetup`

---

**Fecha de resolución:** 27 de octubre de 2025

**Sistema:** Proxmox VE 7.2.14

**Causa raíz:** Snapshot COW undersized para tasa de escritura de la VM

**Solución:** Mejora script cliente KVM para *mysqldump* en solución **on-fly**

## Aviso

Esta documentación y su contenido, no implica que funcione en tu caso o determinados casos. También implica que tienes conocimientos sobre lo que trata, y que en cualquier caso tienes copias de seguridad. El contenido el contenido se entrega, tal y como está, sin que ello implique ningún obligación ni responsabilidad por parte de [Castris](#)

Si necesitas soporte profesional puedes contratar con Castris [soporte profesional](#).

---

Revision #1

Created 2025-10-27 09:05:47 UTC by Abkrim

Updated 2025-10-27 09:28:34 UTC by Abkrim