

Conversación con Gemini

Exported: 15/6/2026 9:12:11

Jordi H.

Has dicho he pedido un iostat sin el intervalo 2. me saca un solo valor que es "desde el inicio". se reinicia cada semana. creo que es más representativo que el ultimo valor que te he mostrado. el nuevo valor es

vg--dades-lv--dades	65,11	2347,63	554,25	471246829	1112557

Gemini

Ese valor es, efectivamente, mucho más sólido para tomar una decisión. Al ser un promedio acumulado desde el último reinicio (hace una semana), elimina los "valles" de inactividad y nos da una media real de carga. Analizando tus nuevos datos

- 554,25 kB/s es tu media real de escritura (kB_wrtn/s).
- Esto equivale a ~1,9 GB por hora.
- Con un snapshot de 25 GB, tienes margen para unas 13 horas de backup antes de que se llene.

¿Es suficiente para 14TB?

Aquí es donde debemos ser realistas: copiar 14TB de datos a través de una red gigabit o a un disco externo suele tardar bastante más de 13 horas (a 100 MB/s constantes, tardaría unas 40 horas). Si tu backup va a durar 13 horas, tienes tres estrategias:**

Estrategia A: El Snapshot "Quirúrgico" (Recomendada)

No hagas el backup de los 14TB de una sola vez si no es necesario.

1. Crea el snapshot.
2. Usa rsync para copiar solo las carpetas críticas o los cambios recientes.
3. Borra el snapshot en cuanto termines esa carpeta.

Esto minimiza el tiempo que el snapshot está "vivo" y acumulando escrituras.

Estrategia B: Ampliar el VG (La más segura)

Como te sobran muy pocos GB en tu RAID5 actual, lo ideal para dormir tranquilo sería añadir un disco más al sistema (aunque sea un SSD barato o un disco viejo de 500GB) solo para que actúe como "colchón" de espacio para el snapshot.

- Al añadirlo al vg-dades, tu VFree subiría y podrías crear un snapshot de, por ejemplo, 200GB, dándote días enteros de margen.

Estrategia C: Reducir la actividad durante el backup

Si usas Python para analítica, evita ejecutar procesos que escriban muchos temporales o logs en /dev/vg-dades/lv-dades mientras el backup esté corriendo. Si la escritura baja de 554 kB/s a 100 kB/s, tu tiempo de seguridad sube drásticamente.

Cómo proceder ahora

Si decides arriesgarte con los 25GB porque tu backup es incremental o rápido, usa este comando para crearlo (es ligeramente diferente al anterior para asegurar que use el espacio exacto disponible):

```
Bas
```

CODE

```
lvcreate -l 100%FREE -s -n snap_backup_ymir /dev/vg-dades/lv-dades
```

- -l 100%FREE: Utiliza todo ese sobrante de 26.67GB automáticamente.

¿Qué herramienta vas a usar para mover los datos? (Si usas rsync, puedo darte los parámetros exactos para que la copia sea eficiente y no sature el disco)

Jordi H.

Has dicho vamos a mirarlo al revés. tengo una tasa de ocupación en este momento del 50% [jordih@ymir ~]\$ df -h S. fitxers Mida En ús Lliure %Ús Muntat a /dev/mapper/vg--dades-lv--dade
s 11T 4,6T 5,8T 45% /dades si mis compañeros del centro de cálculo me ofrecen un storage exterior no donde volcar el snapshot, pongamos un dispositivo de archiving, teniendo en cuenta que voy conectado a 1 gigabit y que comparto ese gigabit con más sistemas, deberíamos estimar el menor caudal posible en estas circunstancias y calcular el tiempo de b

Gemini

Este es el enfoque correcto. Vamos a trabajar con el **peor escenario posible** para ver si tus 26,67 GB de "colchón" (VFree) aguantarán mientras los datos viajan por la red

1. Estimación del Caudal (Ancho de Banda)

Si tienes una red de 1 Gbps pero está compartida y el destino es un storage de archiving (que suelen ser más lentos escribiendo), una estimación conservadora y realista sería un aprovechamiento del **25-30%** de la línea

- Velocidad teórica (1 Gbps): ~110 MB/s.
- Velocidad estimada "peor escenario": 30 MB/s (aprox. 108 GB por hora).

2. Tiempo de Backup para 4,6 TB

Si tienes que mover 4,6 TB de datos reales (el espacio "En ús"):

$$\text{Tiempo} = \frac{4.600 \text{ GB}}{108 \text{ GB/hora}} \approx 42,6 \text{ horas}$$

3. El Cálculo de Supervivencia del Snapshot

Ahora crucemos este tiempo con tu tasa de escritura media de **554,25 kB/s** (~1,9 GB/hora) que obtuvimos de iostat

- Escritura total durante el backup: $42,6 \text{ horas} \times 1,9 \text{ GB/hora} = \mathbf{80,94 \text{ GB}}$
- Tu espacio disponible (VFree): 26,67 GB

Conclusión: El snapshot se rompería

En el escenario actual, **el snapshot se llenará mucho antes de que termine el backup** (aguantaría unas 14 horas, pero el backup necesita 42). Cuando el snapshot llega al 100% de su capacidad (los 26 GB), se invalida y el proceso de lectura desde /mnt/snapshot fallará con un error de entrada/salida (I/O Error)

¿Cómo proceder para que funcione?

Tienes tres opciones técnicas para salvar este desfase:

1. Reducir el volumen lógico (Arriesgado pero efectivo):

Como solo usas 4,6 TB de un total de 11 TB, tienes más de 6 TB "vacíos" dentro de tu LV. Si reduces el sistema de archivos y el LV a 6 TB, liberarías 5 TB hacia el VFree, permitiéndote crear un snapshot gigante que duraría semana

- Nota: Esto requiere `resize2fs` (si es `ext4`) o `xfs_growfs` (XFS no permite reducir), y normalmente debe hacerse offline

2. Backup por bloques o incremental (Recomendado):

En lugar de volcar los 4,6 TB cada vez, usa el snapshot para hacer un `rsync` incremental. La primera vez tardará mucho, pero si el snapshot solo tiene que "vivir" lo que tarde en copiar los archivos modificados del día, tus 26 GB serán más que suficiente

3. Añadir "Espacio de Maniobra":

Pide a tus compañeros del centro de cálculo si pueden mapear temporalmente un disco virtual pequeño (ej. 100-200 GB) a tu sistema Ymir. Lo añades al Volume Group (`vgextend`) y así tendrás espacio de sobra para que el snapshot respire durante las 42 horas

¿Te interesa que miremos cómo reducir el volumen actual para ganar ese espacio libre permanentemente? Es la solución más limpia si no puedes añadir discos físicos

Jordi H.

Has dicho otro cambio de enfoque. mis datos en `/dades` son de tres tipos: 5 ficheros gordísimos de los discos de mis máquinas virtuales de 800GB cada uno. El drama es que, a su vez, estos discos están al 50% de uso. los volúmenes de datos de los contenedores los directorios de overlay de los containers en concreto: `[root@ymir dades]# du -h --max-depth=1`

```
19G  ./vol-containers
1,4T  ./vms
273G  ./download
16K   ./elasticsearch_data
28K   ./lost+found
3,7G  ./infraestructura
348K  ./mongo8
14G   ./container_registry
4,6T  ./var
```

Gemini

Este desglose de datos cambia las reglas del juego por completo. Tienes un escenario de **"Escritura Pesada"**. **El problema no es solo cuánto escriben tus compañeros, sino cómo** escriben.** Al tener 5 archivos de 800 GB (discos de VMs), cualquier pequeño cambio interno en la VM (un log, una actualización) obliga a LVM a copiar bloques del archivo de 800GB al snapshot. Aquí tienes el análisis estratégico de tus datos para el backup:

1. El Peligro de los "Ficheros Gordos" (`./vms`)

Los archivos de 800GB son **ficheros dispersos (sparse files)** o imágenes de disco.

- El problema del Snapshot: Si una de tus VMs hace una desfragmentación o escribe un archivo de 10GB dentro de su disco virtual, LVM detectará cambios en esos bloques y usará tu preciado espacio de 26GB del snapshot a una velocidad de vértigo
- Los 2,9T de `/vms` son tu mayor riesgo de desborde.

2. Clasificación de Prioridad para el Backup

No todos tus datos necesitan el mismo tratamiento. Podemos dividir el backup en tres fases para que el snapshot sobreviva

Carpeta	Tamaño	Estrategia	Riesgo para el Snapshot
<code>./vol-containers</code>	19G	Snapshot + <code>rsync</code>	Muy bajo
<code>./elasticsearch_data</code>	273G	Snapshot + <code>rsync</code>	Medio (muchas escrituras)

./vms	2,9T	Excluir del snapshot inicial	Crítico
./downloads	1,4T	Copia directa (sin	Nulo

3. Propuesta Técnica: Backup Híbrido

Dado que el snapshot de 26GB es un "traje que te queda pequeño" para 4,6TB, mi recomendación es **no meter** el mismo saco**:

Fase A: Los datos "vivos" y pequeños (BBDD y Config)

Usa el snapshot para /vol-containers, /elasticsearch_data, /mongo8 y /var.

- Ventaja: Como son pocos GB, el rsync terminará en minutos. El snapshot de 26GB aguantará de sobra porque el tiempo de exposición es mínimo

Fase B: Los Gigantes (./vms)

Para las máquinas virtuales, el snapshot es peligroso. Tienes dos opciones mejores:

1. Suspender/Pausar la VM: Si puedes permitirte 5 minutos de parada, pausa la VM, copia el archivo de 800GB directamente al storage externo (sin snapshot) y reanuda
2. Snapshot interno de la VM: Si las VMs usan QCOW2, es mejor hacer el snapshot a nivel de hypervisor que a nivel de LVM

Fase C: Lo prescindible (./downloads)

No uses snapshots para esto. Si un fichero de descargas cambia mientras se copia, no pasa nada, se reintenta luego. Ahorras espacio de snapshot para lo que sí es crítico

4. El comando que yo usaría en tu lugar

Para maximizar las posibilidades de éxito, usa rsync con limitación de ancho de banda para no saturar ese Gigabit compartido y dar estabilidad al flujo: Bas

CODE

```
# Crear el snapshot solo para la fase rápida
lvcreate -l 100%FREE -s -n snap_quick /dev/vg-dades/lv-dades

# Montar
mount -o ro /dev/vg-dades/snap_quick /mnt/backup_snap

# Backup de carpetas críticas (ejemplo a 50MB/s para no molestar)
rsync -av --bwlimit=50000 /mnt/backup_snap/vol-containers /mnt/backup_snap/var /destino/externo/

# Desmontar y borrar snapshot INMEDIATAMENTE
umount /mnt/backup_snap
lvremove /dev/vg-dades/snap_quick
```

¿Qué te parece este enfoque segmentado?

Si me dices qué software gestiona esas VMs (¿KVM/QEMU, Proxmox?), puedo decirte cómo copiar esos 2,9TB de forma consistente sin depender del snapshot de LVM, que es tu cuello de botella actual

Has dicho vale. para los discos de las máquinas virtuales, hay alguna estrategia que el mundo de IT haya consolidado? con software libre

Gemini

En el mundo de IT profesional, el backup de máquinas virtuales (VMs) de gran tamaño (como tus discos de 800GB) se ha consolidado bajo un estándar: **no copiar el archivo entero cada vez**, sino gestionar los bloques que cambian. Para hacerlo con software libre en un entorno como el tuyo, estas son las tres estrategias "reina"

1. Dirty Bitmaps (El estándar de QEMU/KVM)

Si usas **Libvirt/KVM** (muy probable en Linux), el sistema puede llevar un registro de qué partes del disco de 800GB se han escrito desde el último backup. Esto se llama Dirty Bitmap

- Cómo funciona: El hipervisor marca los bloques "sucios" (modificados). Al hacer el backup, solo extraes esos bloques
- Software Libre: Virt-back o scripts que usan virsh backup-begin.
- Ventaja: No necesitas snapshots de LVM. La VM sigue corriendo y el backup es increíblemente rápido porque solo transmites, por ejemplo, 5GB de cambios en lugar de 800GB

2. Proxmox Backup Server (PBS) - La joya de la corona

Aunque no uses Proxmox como hipervisor, puedes usar su tecnología de backup. Es, actualmente, la solución de software libre más avanzada para este problema

- Deduplicación en el origen: PBS divide los 800GB en trozos (chunks). Si en el segundo backup el 95% de los trozos son iguales, solo envía el 5% por la red
- Estrategia: Instalas el proxmox-backup-client en tu sistema Ymir. Él se encarga de leer los archivos de /dades/vms y enviarlos de forma eficiente
- Resultado: Tus 2.9TB de VMs podrían tardar horas la primera vez, pero solo minutos las siguientes, sin saturar tu red de 1Gbps

3. Backy2 / BorgBackup (Deduplicación a nivel de bloque)

Si no quieres montar un servidor de backups dedicado y solo quieres volcar los datos al almacenamiento de archiving que te ofrecen tus compañeros

- BorgBackup: Es extremadamente popular en IT. Es capaz de detectar que dentro de esos 800GB hay muchísimos datos repetidos o que no han cambiado
- Funcionamiento: Bash

CODE

```
borg create --progress /destino/archiving::backup-{now} /dades/vms
```

- Ventaja: Maneja de forma nativa los "archivos dispersos" (sparse files) y solo guarda las diferencias. Es ideal para conexiones de 1Gbps porque comprime y deduplica antes de enviar el dato por el cable

¿Cuál elegir para Ymir?

Dada tu limitación de **26GB de espacio libre en LVM**, la estrategia consolidada es **olvidarse de los snapshots de duplicación****:

1. Si quieres algo simple y robusto: Usa BorgBackup. Te permitirá copiar esos 2.9TB de VMs de forma que,

tras la primera copia, las siguientes sean instantáneas. Además, cifra los datos, lo cual es un plus si el storage de archiving es compartido

2. Si buscas una solución "Enterprise": Mira Proxmox Backup Server. Es imbatible en eficiencia de almacenamiento

Comparativa de tiempo estimado (después de la 1ª copia)

Método	Datos enviados	Tiempo (1Gbps)	Impacto en LVM
Copia manual	2.9 TB	~30-40 horas	Riesgo de colapso
Borg/PBS (Incremental)	~10-20 GB (cambios)	10-15 minutos	Cero (sin snapshot)

¿Te gustaría que te ayude a montar un comando de prueba con BorgBackup para uno de esos discos de estructuras.

Jordi H.

Has dicho yo quiero hacer backup para eliminar el riesgo de pérdida de datos y también minimizar el tiempo de reconstrucción. Las dos cosas en caso de accidente grave. Los accidentes de perder un disco del raid lo afrontaría con la reconstrucción automática los de perder dos, con el backup los de perder la configuración del sistema o corrupción del s.o. (no te he contado que el s.o + home + swap + root + uefi + la raíz viven en un raid1 de 450GB que mis compis harán backup con legato. para el raid5 no tienen infraestructura para usar legato, por tamaño. por eso estaba buscando alternativas: ahora que tienes más visión de conjunto, qué harías para hacer b

Gemini

Con esta visión de conjunto, el panorama está mucho más claro. Tienes el sistema operativo protegido por tus compañeros (Legato en RAID1), pero el "corazón" de tus datos (los 14 TB en RAID5) es responsabilidad tuya y es demasiado grande para los métodos tradicionales. Para cumplir tus dos objetivos (**evitar pérdida y minimizar la reconstrucción**), **yo implementaría una estrategia de Backup por Capas****. Olvidaría el snapshot de LVM para los 4,6 TB completos (es un suicidio técnico con solo 26 GB libres) y haría lo siguiente:

1. La Estrategia: Backup de Bloques con Deduplicación

Para el RAID5, la herramienta consolidada que mejor se adapta a tu entorno de 1 Gbps y almacenamiento de archiving es **BorgBackup** (o su interfaz gráfica **Vorta** si prefieres)

¿Por qué esta solución?

- Deduplicación: Si tus discos de VM de 800 GB solo cambian en un 2%, Borg solo enviará esos 16 GB por la red, no los 800 GB
- Eficiencia en RAID5: Al no usar snapshots de LVM de larga duración, no estresas la escritura del RAID5 mientras el backup viaja por la red
- Velocidad de Reconstrucción: Si pierdes el RAID5 (falla de 2 discos), Borg permite montar el backup como si fuera una carpeta local (borg mount). Puedes levantar las VMs directamente desde el storage de backup mientras copias los datos de vuelta

2. Plan de Acción Técnico

Paso A: Preparación de las VMs

Para que los archivos de 800 GB sean consistentes, antes de lanzar el backup de la carpeta /dades/vms, debes asegurar que los discos no se estén moviendo

- Opción Pro: Usa virsh snapshot-create-as (snapshot interno de QEMU). Esto desvía las escrituras a un archivo temporal pequeñito, dejando el disco de 800 GB "quieto" para que Borg lo lea perfectamente
- Opción Simple: Si el backup se hace de madrugada, puedes programar un script que suspenda las VMs (virsh suspend), lance el backup y las reanude (virsh resume)

Paso B: El comando de Backup (Ejemplo)

Instalas Borg y ejecutas:Bash

CODE

```
# Inicializas el repositorio en el storage de tus compañeros (solo la primera vez)
borg init -e repokey /mnt/archiving/backup_ymir

# Lanzas el backup de las carpetas críticas
borg create --compression lz4 --progress \
/mnt/archiving/backup_ymir::{now} \
/dades/vol-containers \
/dades/elasticsearch_data \
/dades/vms
```

3. ¿Cómo afrontar los desastres?

Escenario	Acción de Reconstrucción	Tiempo Estimado
Fallo de 1 disco	El RAID5 reconstruye solo al insertar	Horas (automático).
Fallo de 2 discos (Muerte del RAID5)	Recrear el RAID5, montar el backup	~40-50 horas (limitado por la red).
Corrupción de una VM	No restauras todo; con Borg	Minutos.

4. Mi recomendación final para Ymir

1. No redimensiones el LVM: Es un riesgo innecesario si puedes usar herramientas que no dependen de snapshots
 2. Usa BorgBackup: Es software libre, extremadamente robusto en entornos de investigación y analítica, y gestiona los "ficheros gordos" mejor que nadie
 3. Automatiza: Crea un cron job que se ejecute a las 02:00 AM.
 - Primero las carpetas pequeñas (containers, mongo, elastic).
 - Después las VMs (usando el truco de virsh suspend o el snapshot de QEMU).
- ¿Te gustaría que te ayude con el script de Bash para automatizar este flujo de "parar VMs -> Borg Backup ->

Jordi H.

Has dicho escribe un documento de toda la estrategia, en catalán, y que no sea una conversación entre tú y yo sino un tono impersonal para presentarlo a mis compañeros

Aquest és el document tècnic amb la proposta d'estratègia de còpies de seguretat per al sistema **Ymir**, dissenyat per ser presentat al departament d'IT o al centre de càlcul

Proposta d'Estratègia de Backup: Sistema Ymir (RAID5 LVM)

1. Introducció i Context

El sistema **Ymir** gestiona un volum de dades crític de **14,5 TB** configurat sobre un LUN LVM en **RAID5** (4 discs de 4 TB). Actualment, l'ocupació real és de **4,6 TB**, repartits entre imatges de màquines virtuals (2,9 TB), dades de contenidors i bases de dades (Elasticsearch, MongoDB). L'objectiu d'aquest document és definir una metodologia de backup que minimitzi el temps de recuperació davant un error catastròfic (pèrdua de més d'un disc) i garanteixi la integritat de les dades sense dependre exclusivament dels snapshots de LVM, atesa la limitació d'espai lliure al Volume Group **26,67**

2. Anàlisi de Riscos i Limitacions

- Capacitat de Snapshot limitada: Amb només 26 GB de VFree, un snapshot de tot el volum es desbordaria en menys de 14 hores sota la taxa d'escriptura mitjana actual (554 kB/s)
- Fitxers de gran volum: Les imatges de disc de les VM (800 GB cadascuna) generen una alta amplificació d'escriptura en snapshots, invalidant els mètodes de còpia tradicionals bit a bit
- Coll d'ampolla de xarxa: La connexió de 1 Gbps limita la transferència teòrica a ~100 MB/s, la qual cosa implica un temps de còpia inicial de 40-50 hores per als 4,6 TB

3. Estratègia Tècnica Proposada: Backup per Capes

Es proposa abandonar la idea d'un snapshot global de LVM en favor d'una solució de **Deduplicació a nivell de backup** (programari lliure).

Capa A: Dades Dinàmiques (Containers i BBDD)

- Objectiu: /vol-containers, /elasticsearch_data, /mongo8.
- Mètode: Snapshot de LVM de curta durada.
- Procediment: Es crea un snapshot de 20 GB, es munta en mode lectura, s'executa la sincronització i s'elimina immediatament el snapshot. Això garanteix la consistència de les bases de dades amb un temps d'exposició mínim

Capa B: Màquines Virtuals (Imatges de 800 GB)

- Objectiu: /vms.
- Mètode: Deduplicació i incrementalitat amb Borg.
- Procediment: En lloc de copiar tot el fitxer de 800 GB, Borg analitza quins blocs han canviat des de l'última sessió. Només els blocs modificats s'envien pel canut de 1 Gbps cap a l'emmagatzematge d'archiving
- Consistència: Per garantir la integritat dels sistemes de fitxers de les VM, s'executarà una pausa temporal (virsh suspend) o un snapshot intern de QEMU durant la lectura del fitxer

4. Arquitectura del Flux de Treball

L'execució se sincronitzarà en hores de baixa activitat mitjançant un script automatitzat:

1. Preparació: Suspensió temporal de serveis crítics o escriptures en Python.
2. Snapshotting: Creació del volum lògic de snapshot per a dades petites.

3. Execució Borg: Enviu de dades deduplicades al storage extern.
4. Neteja: Desmuntatge de snapshots i represa de serveis.

5. Temps de Recuperació (RTO) i Objectiu (RPO)

- RPO (Pèrdua de dades): Màxim 24 hores (freqüència de la tasca programada).
- RTO (Temps de restauració): Error parcial: Restauració d'un fitxer o VM individual en minuts mitjançant el muntatge directe del repositori Borg (borg mount). Error total del RAID5: Reconstrucció del sistema de fitxers i bolcat de 4,6 TB a 30-40 MB/s (estimat ~42 hores fins a la recuperació total)

6. Conclusions i Requeriments

Aquesta estratègia permet protegir el sistema **Ymir** sense necessitat d'ampliar el maquinari actual ni posar en risc l'estabilitat del RAID5 amb snapshots persistents. Es requereix que el centre de càlcul faciliti

1. Un punt de muntatge (NFS/SMB) al storage d'archiving amb un espai inicial de 5 TB.
2. Persistència de la xarxa de 1 Gbps durant la finestra de backup.

Voleu que prepari el script de Bash que implementa aquests passos per incloure'l com a annex tècnic al

Jordi H.

Has dicho crec que en aquest document en català has mantingut l'estratègia del snap de lvm, que no podem fer, vols redactar-ho de nou amb la estratègia correcta

Gemini

Tens tota la raó. He mantingut el snapshot de LVM per a les dades petites per inèrcia, però donada la limitació crítica de **26 GB** i el risc d'amplificació d'escriptura (especialment amb bases de dades com Elasticsearch o MongoDB), el més segur és prescindir totalment dels snapshots de LVM i utilitzar una estratègia basada en la **consistència a nivell d'aplicació i fitxer**. Aquí tens la versió corregida i professional del document, eliminant qualsevol dependència dels snapshots de

Proposta d'Estratègia de Backup: Sistema Ymir (RAID5 LVM)

1. Introducció i Context

El sistema **Ymir** gestiona un volum de dades de **14,5 TB** en un entorn LVM sobre **RAID5**. Actualment, l'ocupació és de **4,6 TB**, amb una càrrega de dades heterogènia que inclou imatges de màquines virtuals (2,9 TB), volums de contenidors i bases de dades (Elasticsearch, MongoDB). L'objectiu és establir un protocol de backup robust que garanteixi la recuperació davant una fallada de més d'un disc físic, optimitzant l'ús de la xarxa de **1 G** cap al storage d'archiving

2. Limitacions Tècniques Detectades

- Inviabilitat de Snapshots LVM: El Volume Group només disposa de 26,67 GB lliures. Aquest espai és insuficient per mantenir un snapshot durant el temps necessari per transferir TB de dades, ja que qualsevol escriptura al volum original ompliria el snapshot ràpidament, provocant la seva ruptura i errors d'I/O
- Fitxers de Gran Tamany (Sparse Files): Les imatges de 800 GB de les VM no es poden copiar de forma tradicional cada vegada sense saturar la xarxa i el storage de destinació
- Cabal de Xarxa: Amb una xarxa compartida de 1 Gbps, el temps estimat per a una transferència completa

de 4,6 TB és de més de 40 hores

3. Estratègia Proposada: Deduplicació i Consistència d'Aplicació

Es proposa una solució basada en **BorgBackup**, que permet fer còpies incrementals a nivell de bloc sense necessitat de snapshots de LVM

Capa A: Consistència de Bases de Dades i Contenidors

- Objectiu: /vol-containers, /elasticsearch_data, /mongo8.
- Metodologia: En lloc de fer un snapshot de disc, s'utilitzaran les eines pròpies de cada servei per garantir la consistència: MongoDB/Elasticsearch: Comandaments de "freeze" o "dump" puntuals, o bé aprofitar que Borg llegeix el sistema de fitxers de forma seqüencial. Deduplicació: Borg només enviarà al storage d'archiving els blocs de dades que hagin canviat des de l'últim backup, reduint dràsticament el trànsit de xa

Capa B: Màquines Virtuals (Imatges de 800 GB)

- Objectiu: /vms.
- Metodologia: Backup incremental directe. Borg és capaç de detectar quines parts dels fitxers de 800 GB han variat
- Procediment: Per evitar la corrupció de dades en els discos de les VM durant la lectura: Es llança un comandament virsh suspend (pausa) per a cada VM. Borg realitza la indexació i còpia dels blocs modificats (procés ràpid gràcies a la deduplicació). Es reprimeix l'activitat de la VM amb virsh resum

4. Avantatges de l'Estratègia

1. Seguretat de l'Estructura LVM: No es modifica el Volume Group ni es creen volums lògics temporals, eliminant qualsevol risc de fragmentació o bloqueig del RAID5
2. Optimització del Storage d'Archiving: Gràcies a la deduplicació, el storage extern no s'omplirà amb múltiples còpies de 4,6 TB, sinó que només emmagatzemarà les diferències diàries
3. Restauració Atòmica: Borg permet restaurar fitxers individuals o imatges de VM concretes sense haver de descarregar tot el backup de 4,6 TB

5. Resum del Flux Operatiu

1. Fase de Quiescència: Pausa controlada de les màquines virtuals i bases de dades.
2. Sincronització Borg: Transferència incremental de blocs cap al storage extern (NFS/SMB).
3. Represa: Activació immediata dels serveis un cop finalitzada la fase crítica de lectura de blocs.
4. Verificació: Comprovació periòdica de la integritat del repositori de backup.

6. Conclusions

Aquesta estratègia transforma un backup inviable de 40 hores en una tasca incremental que, en el dia a dia, només mourà els pocs GB de dades que realment canvien en el sistema **Ymir**, mantenint la integritat de les dades i la disponibilitat del servei dins dels paràmetres de la infraestructura actual **Vols que inclogui ara** e comandaments de BorgBackup per mostrar-los com s'implementaria tècnicament?*

Jordi H.

Has dicho voldria que valoressis si, tenint en compte que el volum de s.o. (raid1) el gestionarem els backups amb legato, podria afegir subtasques també amb legato pels directoris mencionats (excepte el /dades/vms) ??
? llavors el panorama seria fer el backup de legato per al disc de sistema, fer les tasques també de legato pe ls altres directoris i finalment fer les tasques de borg escrivint a un mount de archiv

Gemini

Aquesta és una reflexió molt encertada i té tot el sentit des d'un punt de vista de gestió d'IT. Utilitzar **Legato** per a tot allò que sigui fitxers estàndard i deixar Borg** exclusivament per als "monstres" de 800 GB és una estratègia de segmentació molt intel·ligent. Aquí tens la valoració d'aquesta estratègia mixta:

Valoració de l'Estratègia Mixta (Legato + Borg)

1. Avantatges d'afegir sub-tasques de Legato:

- Consolidació: Tindràs la majoria de les dades del sistema sota la mateixa eina i interfície que el S.O. Això facilita la recuperació ràpida de configuracions o volums de contenidors petits
- Suport oficial: Si el centre de càlcul ja gestiona Legato, ells es fan responsables de la retenció i la disponibilitat d'aquestes dades
- Granularitat: Legato és excel·lent per recuperar un sol fitxer de /dades/vol-containers sense haver de muntar un repositori sencer

2. Per què mantenir Borg només per a /dades/vms?

- El problema del "Full Backup" de Legato: Legato sol tenir dificultats amb fitxers de 800 GB que canvien diàriament. Forçaria la xarxa de 1 Gbps a moure TB de dades cada vegada que vulgui fer un backup complet
- La potència de la Deduplicació: Borg brilla exactament on Legato pateix. Com que les VMs tenen moltes dades redundants entre elles, Borg estalviarà moltíssim espai al storage de destinació i temps de xarxa

Escenari Final de Backup per a Ymir

Així quedaria el teu mapa de protecció de dades:

Component	Emmagatzematge	Eina de Backup	Destinació
S.O. / Root / Home	RAID1 (450GB)	Legato	Infraestructura Central
Dades Críiques Petites	RAID5 (14TB)	Legato (Sub-tasques)	Infraestructura Central
Bases de Dades	RAID5 (14TB)	Legato	Infraestructura Central
Imatges de VM	RAID5 (14TB)	BorgBackup	Mount d'Archiving
Descàrregues / Temporals	RAID5 (14TB)	Cap (Opcional)	-

Annex Tècnic: Comandaments per a la Tasca de Borg

Per completar el document per als teus companys, aquí tens la implementació de la part que tu gestionaràs directament (les VMs):

1. Inicialització del repositori (només un cop):

Bas

CODE

```
# Suposant que el storage d'archiving està muntat a /mnt/archiving
borg init -e repokey /mnt/archiving/borg_vms_ymir
```

2. Script de backup programat (cron diari):

Bash

CODE

```
#!/bin/bash
# 1. Pausa de seguretat de les VMs (opcional per consistència)
virsh suspend vm-recerca-01
virsh suspend vm-analitica-02

# 2. Execució del backup deduplicat
borg create --compression lz4 --progress \
/mnt/archiving/borg_vms_ymir::{now} \
/dades/vms

# 3. Represa de les VMs
virsh resume vm-recerca-01
virsh resume vm-analitica-02

# 4. Neteja de backups antics (mantenir els últims 7 dies i 4 setmanes)
borg prune -v --list /mnt/archiving/borg_vms_ymir --keep-daily=7 --keep-weekly=4
```

Conclusió

Aquesta estratègia mixta és la més equilibrada: **Legato et dona la pau mental** per a la configuració i les dades petites, i **Borg et dona l'eficiència tècnica** per gestionar els 3 TB de màquines virtuals sense col·lapsar la xarxa del centre de càlcul. **T'agradaria que t'ajudi a ajustar algun paràmetre del script de Borg o vols que revisem c**

Jordi H.

Has dicho doncs torna a escriure el document impersonal perquè jo pugui presentar la proposta als meus companys del datacenter amb aquest nou enfo

Gemini

Aquest és el document tècnic definitiu, redactat en un estil formal i impersonal, integrant l'estratègia híbrida (Legato + Borg) per optimitzar la protecció del sistema **Ymir**

Proposta d'Estratègia de Backup Híbrida: Sistema Ymir (RAID5 LVM)

1. Introducció i Objectiu

El sistema **Ymir** disposa d'un volum de dades de **14,5 TB** (RAID5 LVM) amb una ocupació real de **4,6 TB**.

Donada la naturalesa crítica i el gran volum d'algunes dades (imatges de màquines virtuals de 800 GB), es proposa una estratègia de backup segmentada que combina les eines corporatives existents amb solucions de deduplicació eficients per a grans volums. L'objectiu és garantir la recuperació total davant un desastre catastròfic, minimitzant l'impacte en la xarxa de **1 Gbps** i optimitzant l'espai d'emmagatzematge ext

2. Limitacions del Sistema

- Espai Lliure en LVM (VFree): El Volume Group només disposa de 26,67 GB lliures. Aquesta limitació impedeix l'ús de snapshots de LVM de llarga durada, ja que es desbordarien ràpidament durant la transferència de dades massives
- Fitxers Monolítics: Les imatges de disc de les VM (2,9 TB en total) són fitxers de gran tamany que no permeten backups tradicionals complets de forma diària sense saturar el canut de xarxa

3. Estratègia de Backup Segmentada

Es proposa dividir la protecció de dades en dues metodologies diferenciades segons el tipus de dada:

A. Gestió amb Legato (Networker)

Per a les dades que requereixen una gestió centralitzada i una recuperació granular de fitxers, s'utilitzaran sub-tasques de **Legato**

- Àmbit d'aplicació: * Sistema Operatiu, configuracions i directoris /home. Volumes de contenidors (/vol-containers). Bases de dades (/elasticsearch_data, /mongo8). Directoris d'infraestructura i logs (/var, /infraestructura
- Avantatges: Integració amb la política de retenció corporativa i facilitat de recuperació de fitxers individuals.

B. Gestió amb BorgBackup (Deduplicació de Blocs)

Per a les imatges de màquines virtuals que sumen prop de 3 TB, es proposa l'ús de **BorgBackup** muntat sobre un recurs d'archiving extern

- Àmbit d'aplicació: Directori /dades/vms (imatges de disc de 800 GB).
- Metodologia: Borg realitza una deduplicació en origen. En lloc de transmetre el fitxer sencer, només s'envien els blocs de dades que han canviat des de l'última sessió
- Consistència: S'implementaran scripts de pausa/represa (virsh suspend/resume) per garantir la integritat dels sistemes de fitxers de les VM durant la lectura de blocs

4. Arquitectura del Flux de Treball

El procés es coordinarà en finestres d'activitat mínima per reduir la latència:

1. Sincronització de Legato: Execució de les tasques programades per a la dades estructurades i fitxers petits
2. Preparació de VMs: Suspensió temporal de l'activitat de les màquines virtuals.
3. Execució Borg: Transferència incremental de blocs de les imatges de disc cap al storage d'archiving.
4. Finalització: Represa de les VMs i verificació de la integritat del repositori.

5. Beneficis de la Proposta

1. Integritat del RAID5: Es prescindeix totalment dels snapshots de LVM, eliminant el risc de bloqueig o corrupció per falta d'espai al Volume Group
2. Eficiència de Xarxa: La deduplicació permet reduir el trànsit de TB a GB en les còpies diàries, respectant l'amplada de banda compartida de 1 Gbps
3. Flexibilitat de Recuperació: Es manté el suport corporatiu per a les dades crítiques (Legato) mentre es gestiona tècnicament la complexitat de les grans imatges de disc (Borg)

6. Requeriments Tècnics

- Configuració de les sub-tasques de backup a l'agent de Legato ja existent.
- Punt de muntatge (NFS/SMB) a la infraestructura d'archiving per al repositori de Borg.
- Espai inicial estimat al storage d'archiving: 3-4 TB (amb creixement reduït gràcies a la deduplicació).

Necessites que t'ajudi a detallar el procediment de recuperació (Restore) per si els teus companys pregunten

Has dicho nou gir de guió. les VMs contenen un cluster de StarRocks. Aquesta base de dades OLAP té el seu propi sistema de backup. És molt similar al dump de mysql. Ho farem servir en comptes del Borg. el nostr e "servei d'informàtica" ens ofereix usar un bucket de s3a. refés el document amb aquesta nova propos

Gemini

Aquest nou escenari amb **StarRocks** i l'ús de **S3a** (protocol S3 per a entorns de fitxers distribuïts) simplifica enormement la infraestructura, ja que utilitzem eines natives de la pròpia base de dades. Això elimina la necessitat de gestionar fitxers de 800 GB a nivell de sistema de fitxers i garanteix una consistència de dades absoluta. Aquí teniu la proposta actualitzada per presentar als vostres companys

Proposta d'Estratègia de Backup Híbrida: Sistema Ymir (Legato + StarRocks S3)

1. Introducció i Objectiu

El sistema **Ymir** allotja un clúster de base de dades OLAP **StarRocks** sobre una estructura de RAID5 LVM. Donada la naturalesa distribuïda d'aquesta càrrega de treball i el gran volum de dades (4,6 TB en ús), es proposa una estratègia de protecció de dades que utilitzi els mecanismes nadius de la base de dades cap a un emmagatzematge d'objectes (Object Storage) i el servei corporatiu per a la resta del sistema

2. Anàlisi del Volum de Dades

L'ocupació del disc de dades de 14,5 TB es segmenta de la següent manera:

- Entorn de Serveis: Volumes de contenidors, configuracions, logs i bases de dades auxiliars (Elasticsearch, MongoDB)
- Entorn de Dades OLAP: Imatges de màquines virtuals que contenen els nodes de StarRocks (approx. 3 TB).

3. Estratègia de Backup Segmentada

Es proposa l'eliminació total de l'ús de snapshots de LVM (degut a la limitació de **26 GB** de VFree) en favor d'una estratègia de "dump" i còpia directa

A. Gestió amb Legato (Networker)

S'utilitzarà l'agent corporatiu per a tota la informació no gestionada per StarRocks.

- Àmbit: Sistema operatiu (RAID1), /home, configuracions de contenidors (/vol-containers), i dades auxiliars (/elasticsearch_data, /mongo8)
- Metodologia: Sub-tasques programades de Legato per a una recuperació granular de fitxers i configuracions del sistema

B. Gestió Nativa de StarRocks (Broker Load / Backup a S3)

En lloc de realitzar còpies de seguretat de les imatges de disc de les VM, s'utilitzarà la funció nativa de backup de StarRocks cap a un **Bucket S3a**

- Àmbit: Tota la base de dades OLAP allotjada a les VMs.
- Metodologia: StarRocks utilitza un "Snapshot" intern a nivell de base de dades (sense bloqueig) i exporta les dades de forma consistent cap al bucket S3 proveït pel servei d'informàtica
- Avantatges: * Consistència: Es garanteix la integritat de les dades OLAP sense pausar les VMs. Eficiència: StarRocks només exporta les dades útils, ignorant l'espai buit dels discos de 800 GB. S3a: L'ús de l'API S3 permet una transferència directa i escalable, ideal per al volum de TB gestionat

4. Flux Operatiu de Recuperació

Davant d'un desastre (pèrdua de més d'un disc del RAID5), el procediment de restauració seria:

1. Recuperació de l'Host: Reinstal·lació del sistema operatiu i configuracions mitjançant Legato.
2. Desplegament de Nodes: Aixecar les màquines virtuals (templates nets o base).
3. Restauració de Dades: Execució de la sentència RESTORE de StarRocks des del bucket S3 per rehidratar el clúster OLAP

5. Beneficis de la Proposta

1. Optimització del Storage: El backup a S3 només ocupa l'espai de les dades reals, estalviant els TB d'espai buit que contenen les imatges de disc de les VM
2. Independència del Hardware: En cas de fallada total del RAID5, les dades estan fora de l'entorn local, disponibles al bucket S3 per a una reconstrucció en qualsevol altra infraestructura
3. Mínim Impacte: S'elimina el risc de saturació de l'espai de snapshots de LVM del sistema Ymir.

6. Requeriments per al Servei d'Informàtica

Per a la implementació, es requereix:

- Credencials d'accés (Access Key / Secret Key) i Endpoint per a un Bucket S3a.
- Quota d'emmagatzematge de 5 TB al bucket S3.
- Configuració de les sub-tasques de Legato per als directoris de dades auxiliars.

Voldries que t'ajudi a preparar un exemple de la sentència SQL de StarRocks per configurar el "Repository" de
