KVM / DR

TECHNICAL GUIDE

Sizing Storage for KVM DR

How much storage a disaster recovery estate really needs — the full replica, simulation staging, and long-term archive — and a simple rule of thumb to sanity-check the total: keep it under ~3.5× your primary.

Published by **KVMDR** · kvmdr.ai · 2026

A VENDOR-NEUTRAL GUIDE FOR INFRASTRUCTURE AND PLATFORM TEAMS

EXECUTIVE SUMMARY

DR storage is four buckets, not one

The most common sizing mistake is budgeting for a single copy of production. A real DR estate carries more: a live replica, a journal of recent change for point-in-time recovery, staging space so recovery can be tested without disruption, and a long-term archive for retention and compliance.

Each bucket scales differently — some with your data, some with your change rate, some with your retention policy. Sizing them as one number either under-provisions (and breaks recovery) or over-provisions (and wastes budget). This guide breaks down all four, gives the math behind each, and closes with a quick heuristic to validate the total.

1. **Full replica.** A complete, recovery-ready copy of protected data at the recovery site. Scales with primary *used* storage.
2. **Retention journal.** Recent changed blocks kept for point-in-time recovery. Scales with change rate and retention window.
3. **Simulation staging.** Space to spin up workloads for non-disruptive recovery testing. Scales with what you test concurrently.
4. **Long-term archive.** Deduplicated history for compliance and long retention. Scales with unique change over months.

BUCKETS 1 & 2

Full replica and retention journal

1 Full replica

The replica is a complete copy of protected data, kept continuously up to date by changed-block replication so it trails production by seconds rather than hours. Size it against **used** capacity, not provisioned — thin-provisioned and sparse volumes only consume what they hold.

Rule: Full replica ≈ 1.0 × primary used storage.

2 Retention journal

To recover to a point in time — not just the latest state — the system keeps a rolling journal of changed blocks over a retention window. Its size is driven by how much data changes per day and how many days you keep, not by total capacity. A modest daily change rate over a week-long window adds far less than a second full copy.

Math: Journal ≈ primary × daily-change-rate × retention-days.
Example: 10 TB × 3% / day × 7 days ≈ 2.1 TB (~0.21×).

> *Replication bandwidth is sized from the same daily-change figure: spread evenly across 24 hours, minimum throughput ≈ daily-change-GB × 8000 ÷ 86,400 Mbps. A higher change rate costs both journal space and link capacity.*

BUCKETS 3 & 4

Simulation staging and long-term archive

3 Simulation staging

Non-disruptive recovery testing spins workloads up in an isolated sandbox while live replication keeps running. That sandbox needs its own space: a working copy of each VM under test, plus room for the writes the test itself generates. Approaches that maintain a **persistent full sim-disk copy per VM** trade extra steady-state space for instant, repeatable testing.

Math: Sim staging ≈ (used storage of VMs tested concurrently) + ~1 week of test writes. Budgeting a full copy plus a small write buffer (~5%) is a safe default: ≈ 1.05×.

4 Long-term archive

Archive holds deduplicated history for compliance and long retention — measured in months, not days. Because only *unique* change accumulates, it grows far slower than a naive "copy per month" estimate. Seed it with one full dataset, then add each month's unique change.

Math: Archive ≈ data + monthly-unique × months, where monthly-unique = data × $(1 - (1 - \text{daily-rate})^{30})$. Example: 10 TB at 3%/day → ~6 TB unique/month; 6 months ≈ 10 + small-compounded ≈ 0.4–0.7× after dedup/compression.

Plus overhead: metadata, bitmaps, and filesystem overhead add roughly **1%** on top of the four buckets.

PUTTING IT TOGETHER

The 3.5× sanity check

Add the buckets and a well-designed KVM DR estate typically lands comfortably under **3.5× primary used storage**. That figure is not a hard limit — it's a sanity check. If your sizing materially exceeds it, the cause is usually an over-long retention window, an over-broad archive policy, or testing every VM at once.

BUCKET	TYPICAL MULTIPLE OF PRIMARY
Full replica	≈ 1.0×
Retention journal	≈ 0.2 – 0.6×
Simulation staging	≈ 1.0 – 1.05×
Long-term archive	≈ 0.4 – 0.8×
Metadata & overhead	≈ 0.01× (1%)
Total	**≈ 2.6 – 3.5×**

> *Over 3.5×? Revisit retention days, archive months, and how many VMs you stage for testing at once — before you buy disk. Under 2.5×? Confirm you've actually budgeted simulation staging; it's the bucket most often forgotten.*

The free DR Storage Calculator at **kvmdr.ai/storage-calculator.html** applies all of this math interactively — adjust VMs, change rate, retention, and archive months to see each bucket and the total multiple live.

CONCLUSION

Budget the buckets. Then check the multiple.

DR storage isn't one copy of production — it's a replica, a journal, staging for testing, and an archive, each scaling to a different driver. Size them separately, account for the ~1% overhead, and validate the total against the 3.5× rule of thumb.

Plan capacity deliberately at design time and the estate stays affordable and recoverable. Skip a bucket — most often simulation staging — and you discover the gap the first time you try to test recovery under pressure.

KVM / DR

About KVMDR

KVMDR is enterprise disaster recovery built natively for the KVM ecosystem — oVirt, RHV, and OLVM. It provides agentless, near-sync replication, one-click failover and failback, non-disruptive recovery testing, immutable recovery copies, and AI-assisted ransomware detection. Its sizing follows exactly this model — replica, retention journal, simulation staging, and long-term archive — so capacity is predictable from day one.

Learn more and claim a free pilot at kvmdr.ai →