maxRAID Grid · Block storage for value-conscious clouds · November 2026

More sellable flash, that flies.

Beta Q4 2026 · general availability Q1 2027

2.4× or more
the sellable capacity of three copies, from the same flash[1]
More terabytes to sell. Same SSDs.

1.25 TB
of flash per terabyte stored at 8+2, surviving any two failures[1]
Before thin provisioning and compression.

Nine 9s
of durability, excluding correlated failures[2]
Reliability is math, not a wish.

20 GB/s
and 2 million IOPS per pool, read and write[3]
Every pool performs like a local array.

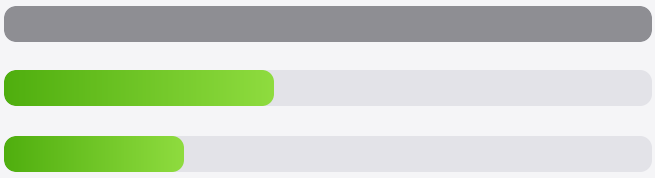
Summary

A cloud sells block storage by the usable terabyte and buys it by the raw one. The common open-source answer, Ceph RBD, keeps three full copies of every block, so a host buys three terabytes of flash for each one it sells, and every write lands on three drives.

maxRAID Grid protects volumes with erasure coding across servers instead. An 8+2 pool survives any two drive or server failures, as three copies do, from 1.25 terabytes of flash per terabyte stored. That is 2.4 times the sellable capacity from the same flash, before inline compression adds more.

Volumes reach clients over NVMe/TCP, which Linux, QEMU/KVM and Proxmox already speak, with QoS on every volume, a Proxmox driver, a Kubernetes CSI plugin, and as much as nine nines of durability.

Flash to buy for 1 PB sellable, at 80% fill[1]



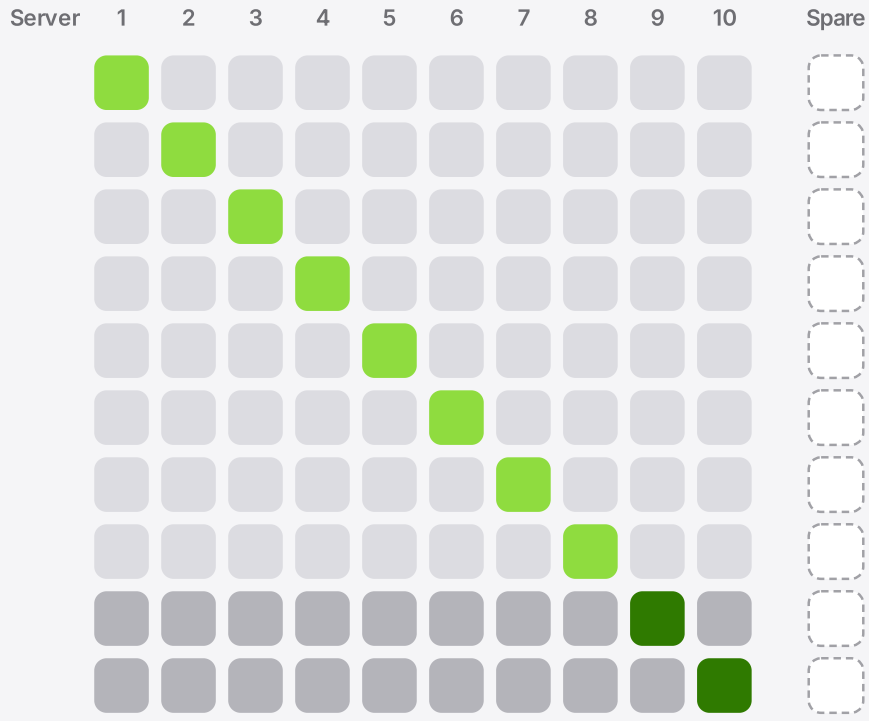
About 2.2 PB less flash to buy for every petabyte you sell — before thin provisioning.

© 2026 EasyCo LLC dba WildFire Storage · maxRAID Grid

1 Less flash per sellable terabyte

Three copies cost three times the flash. Ceph can place RBD volumes on erasure-coded pools such as 4+2 or 6+2, but it erasure-codes each object on its own: a small write becomes a read-modify-write across the whole set, block performance drops further, and published tests show less space saved than the code promises. Block clouds therefore almost always run three copies.

The grid is a set of pools. Each pool is a K+M erasure-coded set with its own software flash layer, and its drives are spread across the grid: one drive in every server. Writes are gathered in memory and laid down as whole stripes, so parity never needs a read-modify-write.



A 10-server grid with one hot-spare server. Each 8+2 pool puts one drive in every server, on its own diagonal (one pool in green; darker cells hold parity). If a server fails, the spare takes its place.

Flash for 1 PB sellable, at 80% fill[1]

	Overhead	Raw flash	Survives
Ceph RBD, three copies	3.0×	3.75 PB	any two copies lost
maxRAID Grid, 8+2	1.25×	1.56 PB	any two drives or servers
maxRAID Grid, 10+2	1.2×	1.50 PB	any two drives or servers
maxRAID Grid, 9+3	1.33×	1.67 PB	any three drives or servers

Four more things stretch each terabyte further

Inline compression, always on.

Every 4 KiB block is compressed with LZ4, fast enough to keep up with the network. Because the grid always writes whole stripes, compression adds no read-modify-write: it only cuts the bytes that reach the flash, saving space and wear together. At 1.5:1, 1 PB sellable on 8+2 needs about 1.04 PB of flash. Ceph, including Proxmox's Ceph, leaves compression off by default.

Thin provisioning.

A volume uses flash only for what its tenant writes, so a host sells provisioned capacity and buys written capacity.

Low flash wear.

With three copies, every byte a tenant writes lands on three SSDs, and each typically writes about 2.5 bytes of flash for every byte it is sent. The grid writes whole stripes of compressed data, so its total write amplification is often below 1:1. Byte for byte, SSDs can last as much as 10 times longer.

Dense, ordinary SSDs.

QLC and PLC drives, including those with large indirection units; SATA, SAS or NVMe; white-box x86 servers with modest memory.

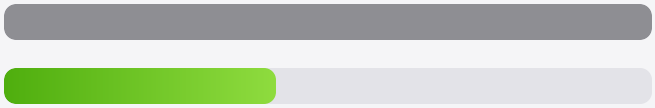
2 Nine nines, by the numbers

Nine nines means a one-in-a-billion chance per year of losing a pool. Our sizing guide, *Sizing a maxRAID Grid for Nine Nines*, shows the arithmetic. Plenty of events sit outside what it models, such as correlated failures, but reliability is math, not a wish.

3 Fast, and fair between tenants

Each pool runs on its own, with no central metadata server, so it performs like a local array: about 20 GB/s and 2 million IOPS per pool, read and write. Volumes striped across pools scale further: ten servers with ten SSDs each reach about 200 GB/s and 20 million IOPS. With three copies every client write costs three drive writes; at 8+2 it costs 1.25, so the same drives carry about 2.4 times the client write bandwidth.

Drive writes for every client write



The same drives carry about 2.4 times the client write bandwidth — and wear more slowly.

QoS that belongs to the volume

QoS is set per volume and enforced by the storage. Read and write IOPS, and read and write bandwidth, are each limited separately. Bursts work like a bank balance: a volume running under its limit banks credit, and spends it later up to a burst ceiling for a set time. The limits live with the volume on the server, so one setting covers every client that mounts it — VMs, containers and bare-metal hosts alike. The same volumes serve AI clusters too: NVMe-oF over RDMA with GPUDirect, and volumes that several clients mount at once.

© 2026 EasyCo LLC dba WildFire Storage · maxRAID Grid

4 Fits what you run, and runs itself

- **NVMe/TCP** through the initiator built into Linux and used by QEMU/KVM, over standard Ethernet; NVMe-oF over RDMA where the fabric has it, and iSCSI for older clients. Multiple paths to every volume.
- **Proxmox** has a storage driver for the grid. **OpenStack** volumes attach over NVMe/TCP, and a Cinder driver is planned. **Kubernetes** has a CSI plugin.
- **One drive per server in every pool,** so losing a drive, or a whole server, costs each pool at most one member. Failed SSDs are replaced automatically, and rebuilds spread across the grid.
- **A failed server fails over to a hot spare,** which takes over its target IP addresses, so clients carry on without reconfiguration. Growth is adding servers, on white-box hardware. Managed with maxRAIDOne.

Volume limits, per server

Block devices, each exportable on a target	65,536
Logical volumes, snapshots included	1 million

Compared with Ceph RBD

	Ceph RBD, three copies	maxRAID Grid
Protection	three copies, 3.0×	erasure coding across servers: 8+2 is 1.25×
Flash for 1 PB sellable	3.75 PB	1.56 PB at 8+2, less with compression
Inline compression	BlueStore, off by default	LZ4 on every 4 KiB block, always on
Snapshots and clones	snapshots; clones	1,000+ deep read-only snapshots; clones
QoS per volume	QEMU throttling per hypervisor	set on the storage, for every client, with banked bursts
Client protocol	QEMU's RBD driver (librbd)	NVMe/TCP, NVMe/RDMA and iSCSI, from the Linux kernel
Proxmox, OpenStack, Kubernetes	yes, yes, ceph-csi	Proxmox driver; NVMe/TCP attach, Cinder driver planned; CSI plugin

Talk to us about sellable flash.

wildfire-storage.com/grid · +1-707-maxRAID · (707) 629-7243 · sales@wildfire-storage.com

1. Capacity: arithmetic from the stated assumptions. Ceph RBD on three-way replicated pools; both at 80% fill.
2. Durability from *Sizing a maxRAID Grid for Nine Nines*: independent SSD and server failures; correlated failures excluded.
3. maxRAID Grid is pre-release; performance is a design target and depends on SSDs, CPUs, network and how well the data compresses. Proxmox is a registered trademark of Proxmox Server Solutions GmbH; Kubernetes of The Linux Foundation; OpenStack of the OpenInfra Foundation; Ceph of Red Hat, Inc.

© 2026 EasyCo LLC dba WildFire Storage · maxRAID Grid