把4 × 1.2T 10k SAS 做 RAID10 和 3 × 1.2T 10k SAS 做 RAID5 性能对比
很好 — 我把 4 × 1.2T 10k SAS 做 RAID10（两对镜像再条带）和 3 × 1.2T 10k SAS 做 RAID5
在常见场景下的读写性能、延迟、空间利用、重建行为、适用性都算清楚并做对比。并给出结论与推荐。为便
于理解，我用一个常见的物理盘性能假设作为基线并把计算过程写清楚 ——
你可以把基线值替换成你实际盘的数字来重新套算。

假设（基线，便于定量比较）
每块 1.2T、 10k RPM SAS 磁盘的 随机 4K IOPS我取： 150 IOPS/盘（行业常见区间约 120–180， 10k 常取 ~150）
每块盘 顺序带宽 我取： 150 MB/s/盘（ 10k SAS 顺序读写常见为 120–180 MB/s，取 150
便于计算）

（若你有厂商给的确切 IOPS/带宽数值，我可以用那些数值重算并给精确结论。）

——下面所有算术我会显示计算步骤，数字逐位计算确保正确。

一、容量（可用空间）
RAID10 (4盘)： usable = (总盘数 / 2) × 单盘容量 = (4 / 2) × 1.2T = 2 × 1.2T = 2.4 TB RAID5 (3盘)： usable = (总盘数 − 1) × 单盘容量 = (3 − 1) × 1.2T = 2 × 1.2T = 2.4 TB

=> 两者在此盘数下可用容量一致 = 2.4TB。

二、随机 4K IOPS（小随机读 / 写，数据库 /虚拟机最敏感）

1) RAID10（4 磁盘，做成两镜像对后条带）

- 每盘 150 IOPS。
- 读：理论上 RAID10 可以并行从所有镜像盘读取，最大读 IOPS ≈ 4 × 150 = 4 × 150 = 600 → **600 IOPS (read)**
- 写：每逻辑写需要写入镜像对的两块盘（写入必须到两块），但条带把写分散到两对镜像上，所以写并行度 = 镜像对数 = 4 / 2 = 2 对 → 写 IOPS ≈ 2 × 150 = 2 × 150 = 300 → **300 IOPS (write)**

2) RAID5（3 磁盘）

- 每盘 150 IOPS，总物理 IOPS = 3 × 150 = 3 × 150 = 450 → **450** 物理 **IOPS** 可用
- 读：读可以从数据盘并行读取（不读取 parity），读并行度 = n − 1 = 2 → 读 IOPS ≈ 2 × 150 = 2 × 150 = 300 → **300 IOPS (read)**
- 写（小随机写，最坏情形，采用读-修改-写）：RAID5 写惩罚通常为 **4** 个物理 **I/O**（读旧数据 + 读旧 parity + 写新数据 + 写新 parity）。所以 逻辑写 **IOPS ≈** 物理总 **IOPS / 4** = 450 / 4 = 450 ÷ 4 = 112.5 → 取整 **≈112 IOPS (small random write)**

> 小结（随机 4K）：
>
> - RAID10：读 **600 IOPS** / 写 **300 IOPS**
> - RAID5：读 **300 IOPS** / 写 **≈112 IOPS**

结论：**RAID10** 在随机读/写下显著优于 **RAID5**（尤其是随机写，约 **2.7×** 更快）。

三、顺序吞吐（大文件顺序读 / 写）

假设每盘顺序带宽 150 MB/s。

RAID10（4盘）

- 顺序读：可从四盘读取 → 4 × 150 = 4 × 150 = 600 → **600 MB/s (seq read)**
- 顺序写：写入镜像对并条带，条带数 = 2 → 2 × 150 = 2 × 150 = 300 → **300 MB/s (seq write)**

RAID5（3盘）

- 顺序读：数据条带在 n−1 磁盘 → 2 × 150 = 2 × 150 = 300 → **300 MB/s (seq read)**
- 顺序写（full-stripe writes，能做到不读旧数据）：也能写入到 n−1 磁盘 → 2 × 150 = 2 × 150 = 300 → **300 MB/s (seq write, full-stripe optimal)**

> 小结（顺序）：
>
> - RAID10：读 **600** / 写 **300 MB/s**
> - RAID5：读 **300** / 写 **300 MB/s**（在 **full-stripe** 场景）

结论：顺序读 **RAID10** 有明显优势；顺序写在 **full-stripe** 情况下两者相当。
但注意：真实文件复制往往不是 perfect full-stripe 写，且文件系统与缓存行为会影响实际数值。

四、延迟与尾延迟

- RAID10：写延迟较小（镜像写入能并行），随机写延迟明显优于 RAID5。尾延迟较小，较稳定。
- RAID5：随机小写入有写惩罚，导致平均延迟与尾延迟显著上升；在高并发 / 随机写场景下会产生长尾（卡顿、响应慢）。

五、重建（Rebuild）与可靠性/风险

- RAID10（4盘）：如果一块盘故障，只需重建该盘到其镜像对上；重建只需复制镜像对的数据 → 重建速度快，恢复期间性能影响较小，风险较低。
- RAID5（3盘）：故障后需要重建整个阵列（重建时需要读取剩余所有盘并计算校验），重建时间长且在重建期间阵列处于降级状态，若再有一块盘故障则数据丢失 → 更高风险。

六、CPU/控制器开销与写开销

- RAID5 有 parity 计算（如果是软件 RAID，会消耗 CPU；如果是硬件 RAID，会用 RAID 卡），所以在写密集场景下 CPU/控制器开销更高。
- RAID10 CPU 开销极低（仅做镜像复制，不做校验计算）。

七、总结对比（表格）

项目	**4×1.2T 10k RAID10**	**3×1.2T 10k RAID5**
可用容量	2.4 TB	2.4 TB
随机 4K 读（估）	**600 IOPS**	300 IOPS
随机 4K 写（估）	**300 IOPS**	**~112 IOPS**
顺序读（估）	**600 MB/s**	300 MB/s
顺序写（full-stripe）	300 MB/s	300 MB/s
写惩罚（小写）	低（镜像）	高（4x）
重建速度与风险	快，风险低	慢，风险高（双盘故障会丢数据）
CPU/Parity 负担	低	高（如果是软件 RAID）
适合场景	随机写/数据库/虚拟机/低延迟需求	顺序读写/归档/成本敏感场景

八、建议（基于业务场景）

- 如果主要工作负载是 数据库、虚拟机、随机小 **IO**、多并发写 （延迟敏感），选择 **RAID10** 。它能提供更高的随机 IOPS、稳定性更好、重建影响更小。
- 如果主要工作负载是 大文件存储、备份、顺序读写 ，且你更看重磁盘利用率与成本（在盘数更多时 RAID5 更节省），可以考虑 **RAID5**，但需接受随机写性能较差与重建风险。
- 如果你使用的是 软件 **RAID5**（ **Windows** 动态磁盘 **/ mdadm**），要特别注意 parity 计算对 CPU 的消耗与重建期间的性能抖动；对关键业务，不推荐软件 RAID5 架构（ RAID10 更稳妥）。

九、举例说明（把表格数字代入实际）

用上面基线（ 150 IOPS/盘， 150 MB/s/盘）：

- 随机写： RAID10 ≈ **300 IOPS** vs RAID5 ≈ **112 IOPS** → RAID10 大约 **2.7×** 更快（小随机写场景）。
- 顺序读： RAID10 ≈ **600 MB/s** vs RAID5 ≈ **300 MB/s** → RAID10 恰好 **2×** 更快（顺序读取大量数据时）。

如果你愿意，我可以：

1. 用你实际的单盘 IOPS/带宽（或厂商型号）把上面的数字重算成精确值；
2. 给出在 特定并发 **/ block size / queue depth** 下的估算（比如 4K QD32 的估算）；
3. 给出 软硬件配置建议 （如：若继续用 RAID5，建议使用带 BBU/写缓存的硬件 RAID 卡或使用 RAID6/RAID10 替代以降低风险）。

你想用厂商的单盘指标（ IOPS/MB/s）重算，还是我用上面基线继续给 QD/BS 更细的估算？

Revision #1
Created 2025-12-09 07:38:26 UTC by Admin
Updated 2025-12-09 07:38:56 UTC by Admin