

唤醒沉睡的存储：提升利用率， 增强成本效益



普通SSD的局限导致存储容量未能得到充分利用。

存储利用率在存储管理中的作用

数据基础设施和运营团队会密切关注存储系统的多个关键指标，包括总存储容量、当前使用率（以及剩余的空闲空间）、数据生成速度和存储资源的消耗情况，从而准确预测未来的存储需求并合理安排存储设备的部署计划。他们还会对存储系统中出现的任何问题或潜在风险保持警觉。其中一个特别关注的指标是存储利用率阈值（FRT），或者称为最大利用率，这涉及到监控“每个存储设备的容量中有多少百分比已经被数据填满”。

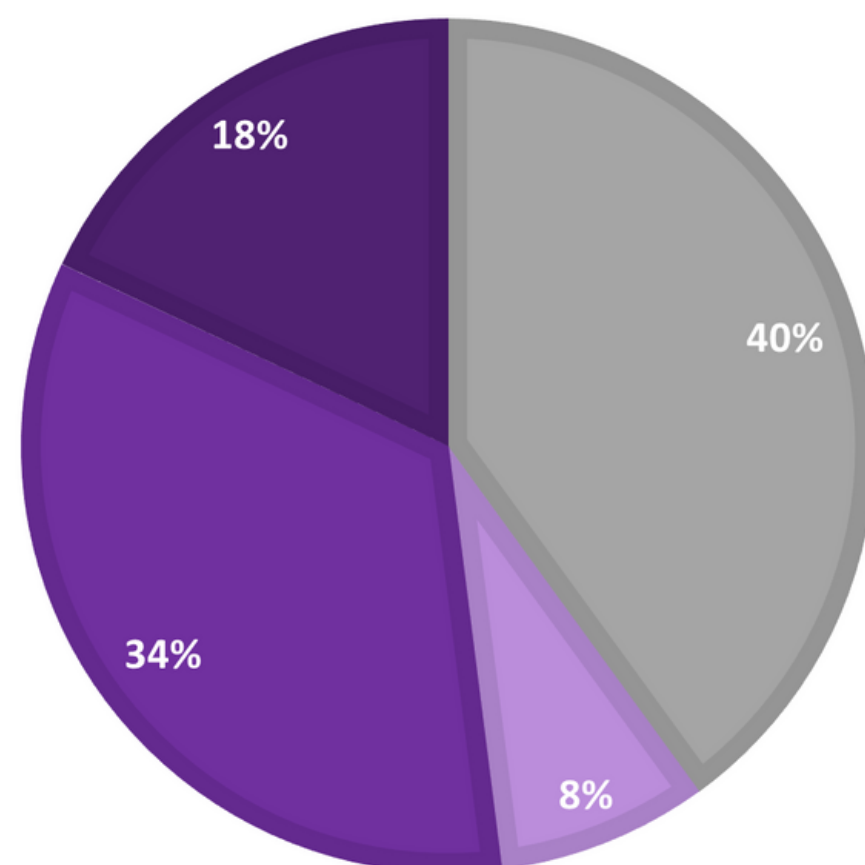
如何设定存储利用率，这种设定对成本有何影响？

在我们于2024年进行的一项调查中，有52%的受访者表示他们将存储容量阈值设定在70%或50%。

如果我们将95%的存储利用率阈值视为最合理的上限，这个上限可以确保在存储空间耗尽之前有足够的时间来增加更多的存储资源。那么，与95%的标准相比，70%的利用率阈值意味着有25%的存储容量未能得到利用，这可能会使你在存储每TB数据时的成本增加到原来的1.36倍。

FILL RATE THRESHOLDS

■ <90% ■ <80% ■ <70% ■ <50%



如果设定一个较为保守的50%存储利用率阈值，就意味着你的存储潜力有将近一半没有得到充分发挥，而且存储数据的成本也会随之增加到原来的1.9倍。这还没考虑到你需要的存储设备数量也得翻倍。

所以，受访者认为提高存储利用率阈值能够减少成本支出和降低设备故障风险，这是很有道理的。

基础设施和运营决策者为何对提高存储利用率持谨慎态度？

当被问及为何选择特定的存储利用率上限时，仅有32%的人担心是否能够及时增加存储容量，而其他68%的人则更关心存储设备在高利用率下的性能表现、存储设备的耐用性和整个存储集群的可靠性。

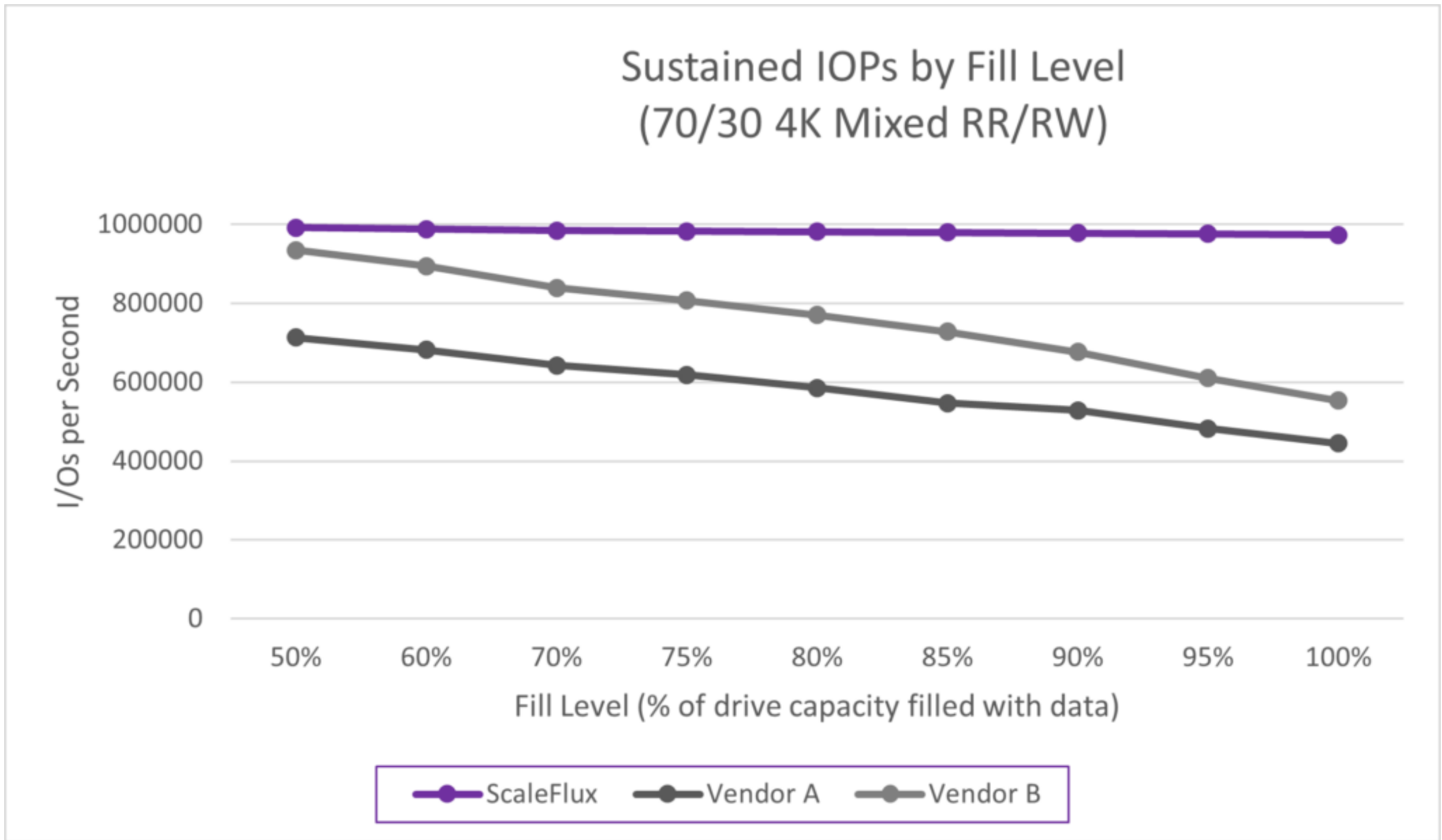
这些担忧当然是有道理的，当SSD存储设备接近满载时，它们的性能会受到影响，耐用性问题也变得更加突出。简单来说，存储设备在高利用率下需要处理更多的数据，导致更多的后台操作，这些操作会占用宝贵的I/O资源并降低前端性能，还会增加每次写入数据时对NAND Flash的磨损。企业级NVMe存储设备测试也验证了这一点。

有哪些策略可以帮助企业在不牺牲性能的前提下提高存储利用率？

搭载可计算存储的SSD即使在较高的利用率下也能保持稳定的每秒输入/输出操作数（IOPS）、吞吐量（GB/s）和延迟，有效避免了通常与高利用率阈值（FRT）相关的缺点。

为了验证此类SSD的性能，我们用FIO工具对三款存储设备进行了测试，模拟了企业级和数据中心环境中常见的4KB随机读写操作，读写比例为70%读和30%写。测试特别关注了存储设备在不同利用率（从50%到100%的利用率阈值）下的性能表现。测试结果显示，即使是市面上领先的企业级NVMe SSD，当存储利用率提高时，其性能也会有所下降。

ScaleFlux可计算存储设备CSD3000在利用率达到 100% 时仍保持了性能和延迟的一致性（见下图）。



ScaleFlux的CSD 3000利用其内置的硬件压缩引擎，在达到满额存储利用率时仍能维持性能和延迟的稳定性。它在数据写入NAND闪存之前对数据进行压缩，这减少了对存储设备的频繁写入需求，释放了更多未使用空间，并降低了写入放大效应。因此，CSD 3000即使在95%的高存储利用率下也能提供比传统存储设备在50%利用率下更优的性能和耐用性。用户可以完全利用他们购买的存储容量，同时简化了系统管理，无需担心存储设备的过度磨损或性能降低。

关于ScaleFlux锐钲:

ScaleFlux锐钲是大规模部署计算存储的领导者，旨在帮助其客户利用数据增长作为竞争优势，提供企业级计算存储芯片解决方案，其硬件计算加速引擎极大优化了NVMe SSD，提升了存储的能力。有效加速应用程序并优化数据中心、企业和边缘网络的基础设施资源。让客户在处理数据库、分析、物联网和5G等工作负载时获得更大的竞争优势。更多内容请访问www.scaleflux.cn