

IBM Cloud Object Storage 通过高级存储引擎改善性能

亮点

最高可达：

300%
GET 吞吐量提升幅度

180%
PUT 吞吐量提升幅度

30%
跨地域分散节点的延迟的降低幅度

55GB/s
12 节点配置

区域分片存储 (ZSS) 支持原始磁盘存储的直接管理，进而提供更有效的数据读取和写入，显著提高性能

执行概要

随着组织在对象存储及其适用的工作负载方面的需求不断增长，IBM 已经投入了大量资金来应对这一挑战。具体来说，IBM 已对区域切片存储 (ZSS) 的基础架构进行了改进，并将其作为 IBM Cloud Object Storage (COS) 的新存储引擎。ZSS 大幅提升了 COS 的性能，使系统能够更有效地读写数据，同时使其能够适应新的存储媒介技术。

存储引擎是一种在软件堆栈上执行最低级别功能的软件 - 物理存储媒介数据读写。作为软件堆栈的最低级别，ZSS 对于使用 IBM COS 的应用是不可见的。在结合采用 ZSS 时，COS 如今所支持的所有功能均不会受到影响。它能够与现有存储引擎实现方式“打包切片存储 (PSS)”同时运行，这使您可以继续使用现有存储。ZSS 存储引擎既可在部署新的 COS 时使用，也可以在增量式存储容量扩展时使用。

ZSS 的主要优点在于提升性能，即每秒操作数 (OPS)、吞吐量和延迟。此外，它还为其两个非常重要的未来收益奠定了基础：

- 主机托管的叠瓦式磁记录 (SMR) 磁盘驱动器支持，该技术能够提高每个驱动器的存储密度，而不会影响磁盘托管 SMR 驱动器所带来的性能优势
- 延长 SSD 的设备使用寿命

之所以能够实现性能优势，其原因在于：

能够顺序写入大批数据，因此能够减少在向磁盘写入相同数量的数据时的写入操作量。从写入吞吐量和 OPS 方面来说，这无疑会增加系统带宽。在许多情况下，ZSS 仅需要一次搜寻便可完成读取，而 PSS 最多需要 4 次搜寻（而 ext4 等传统文件系统技术则需要 4 次或更多次的搜寻），因此 ZSS 能够增加 COS 系统在读取吞吐量和 OPS 方面的带宽。ZSS 还有助于降低 COS 系统的读取延迟，尤其是对于地域分散的 COS 系统。

对于那些可能在运行时出现磁盘硬件故障的系统，或者遇到意外断电或不当关机的系统，ZSS 还可为其提供一些运行改善。此外，ZSS 能够让磁盘压缩变得更加有效。在许多存储系统中，压缩是一个后台进程，能够在磁盘驱动器中的数据由于随机写入和数据删除而变得碎片化之后，将其合并到磁盘驱动器之中。

ZSS 如何实现该功能的？

ZSS 通过应用软件直接控制磁盘。它不像 PSS 那样依赖于 Linux 内核进行读写。ZSS 的工作流程允许应用确定最适合与存储媒体交互的内容。COS 的工作负载都是唯一的，因此，如果由应用控制事务处理而非依赖于 Linux ext4，磁盘读写便会更有效。

区域切片存储有助于提高性能

在本节中，我们将会介绍一些性能测量结果，用于量化 ZSS 相比于现有存储引擎“打包存储引擎”的一些性能改善。我们根据读取 (GET) 和写入 (PUT) 吞吐量的测量数据得出了对比结果。为了简化性能对比，我们使用一系列大小固定的对象分别运行了 PUT 和 GET 操作。

此外，我们还提供了 IBM Cloud Object Storage 两种不同操作模式（保险库模式和容器模式）下的性能测量结果。保险库模式通常用于 COS 的传统备份和归档用例。容器模式主要用于高度可扩展的用例，例如服务即存储、云服务和移动服务以及大型企业数据海洋。

第一组测试所用的 Slicestor™ 设备是配备了 53 个 4 TB 磁盘驱动器的 SS53。同时通过一组 4105 Accesser™ 设备托管负载驱动程序。性能测试的重点是衡量 ZSS 相比于 PSS 的吞吐量提升。在下文的对比中，我们对结果进行了按比例缩放，以便对 12-Slicestor 系统进行建模。系统的某些方面（扫描、重建、压缩）未进行建模，也未在结果中体现。请注意，在下文图表中，吞吐量数字越大表示性能越高。

配置：

- (1) 台配备 4 TB 驱动器和 40G NIC 的 SS53 完全填充 Slicestor 设备
- 保险库模式
- 12/8/10 IDA（已建模）
- 索引关闭
- SecureSlice™ 已禁用
- 分段大小 4 MiB
- 每次测试的固定对象大小：100 KB、1.25 MB、3.5 MB、100 MB

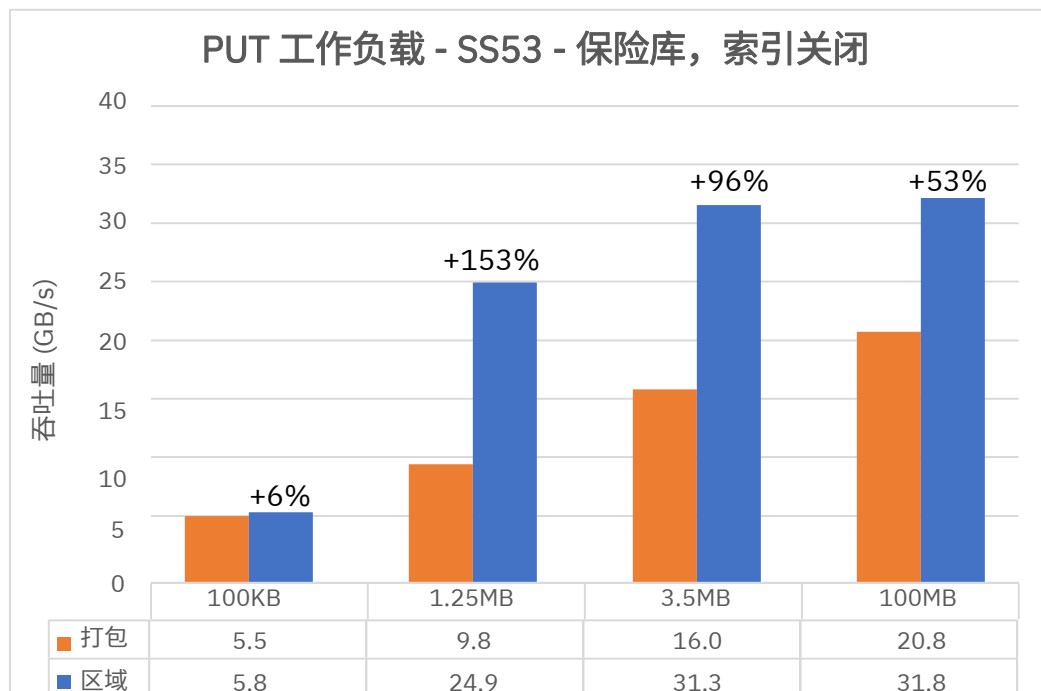


图 1: 数字越大表示性能越高

采用相同系列的固定大小对象进行了 GET 请求测试。COS 系统的配置与 PUT 测试时相同。

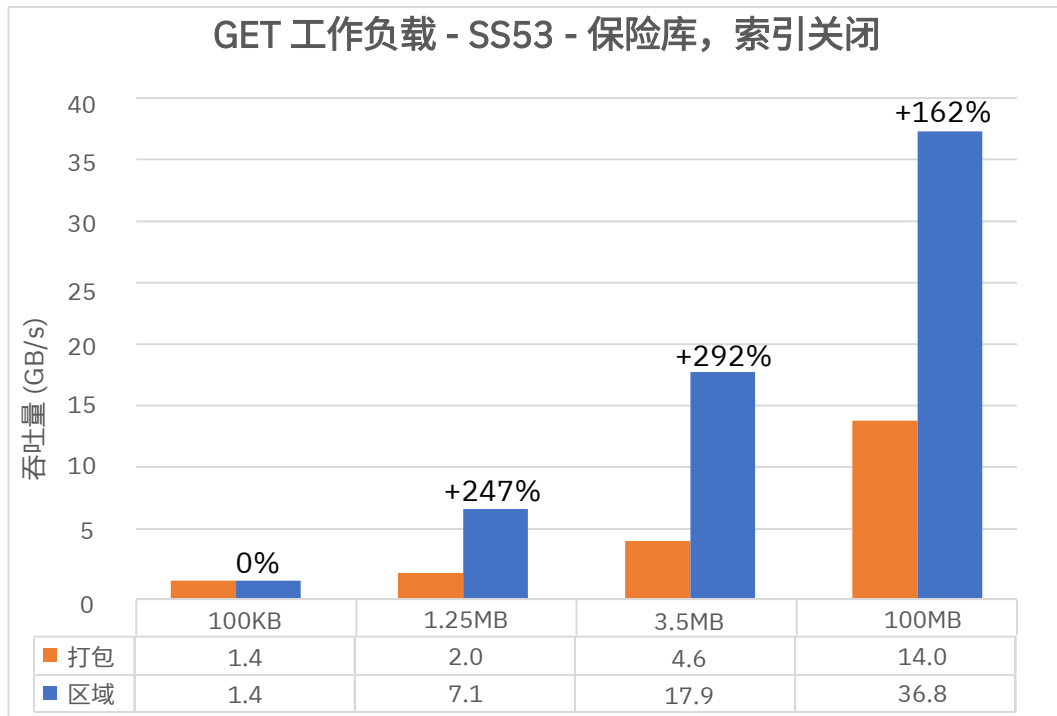


图 2: 数字越大表示性能越高

在保险库模式下, 无论是 PUT 还是 GET 工作负载, 相比于 PSS, ZSS 均可显著提升吞吐量。对于大于 100KB 的对象, 尤其是大小介于 1.25 到 3.5 MB 之间的对象, 提升程度非常显著。

SS106 Slicestor 设备的保险库模式吞吐量测量结果:

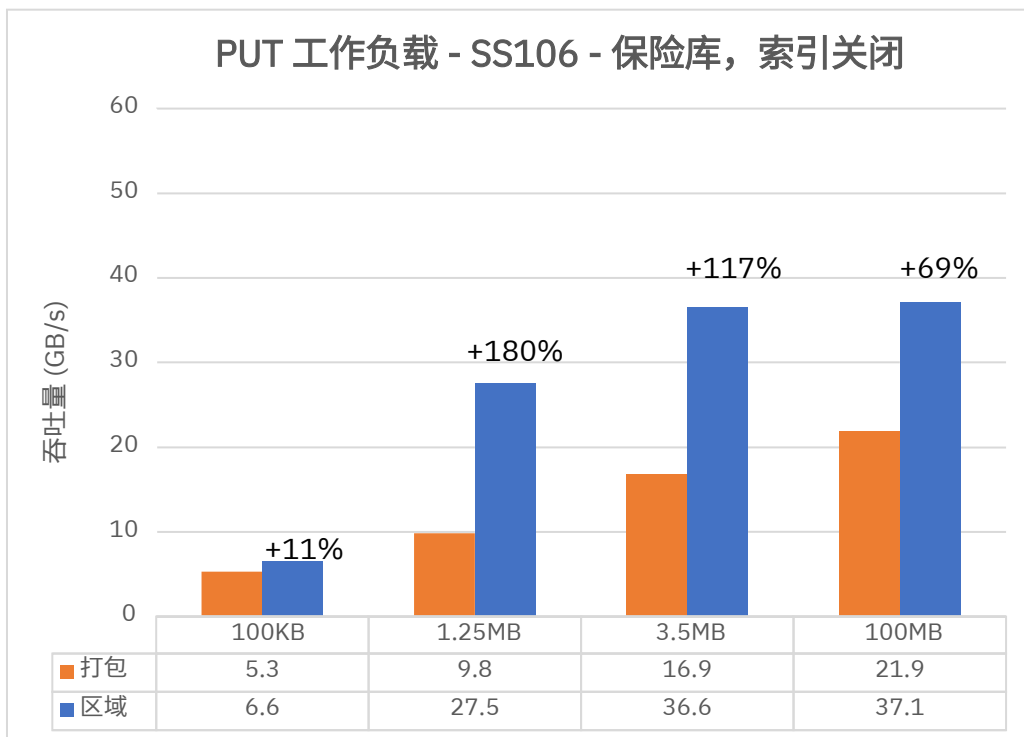


图 3: 数字越大表示性能越高
[预约专家](#)

尤其是 GET 工作负载，随着主轴数量的增加（即 SS53 与 SS106），性能得到了显著改善。从大对象的测量结果来看，PUT 和 GET 工作负载均让 40G NIC 达到饱和。

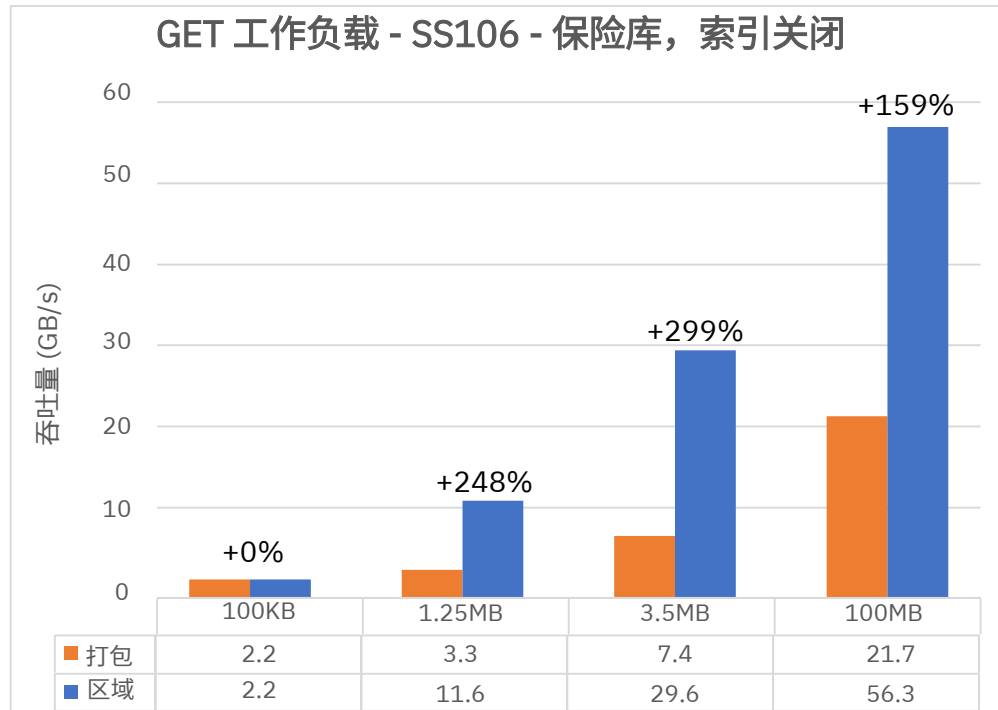


图 4: 数字越大表示性能越高

容器模式测试

第二组测试与第一组相似，区别在于测试系统被重新配置为容器模式而非保险库模式。同样要注意，在下文图表中，吞吐量数字越大表示性能越高。

配置：

- (1) 台配备 4 TB 驱动器和 40G NIC 的 SS53 完全填充 Slicestor 设备
- 容器模式
- 12/8/10 IDA（已建模）
- SecureSlice 已禁用
- 分段大小 4 MiB
- 每次测试的固定对象大小：100 KB、1.25 MB、3.5 MB、100 MB

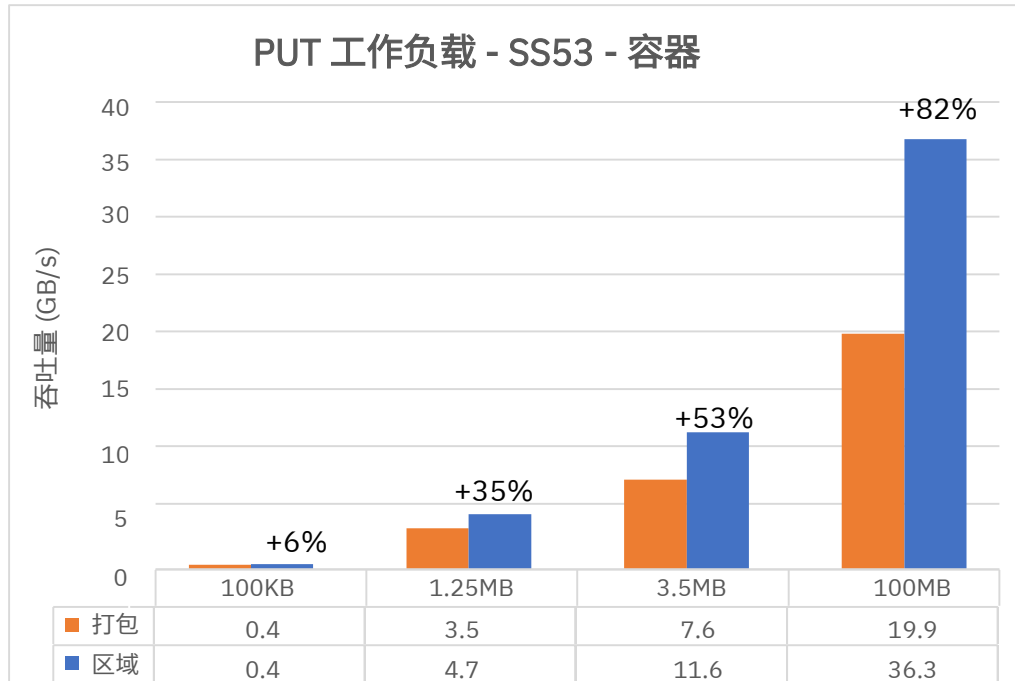


图 5: 数字越大表示性能越高

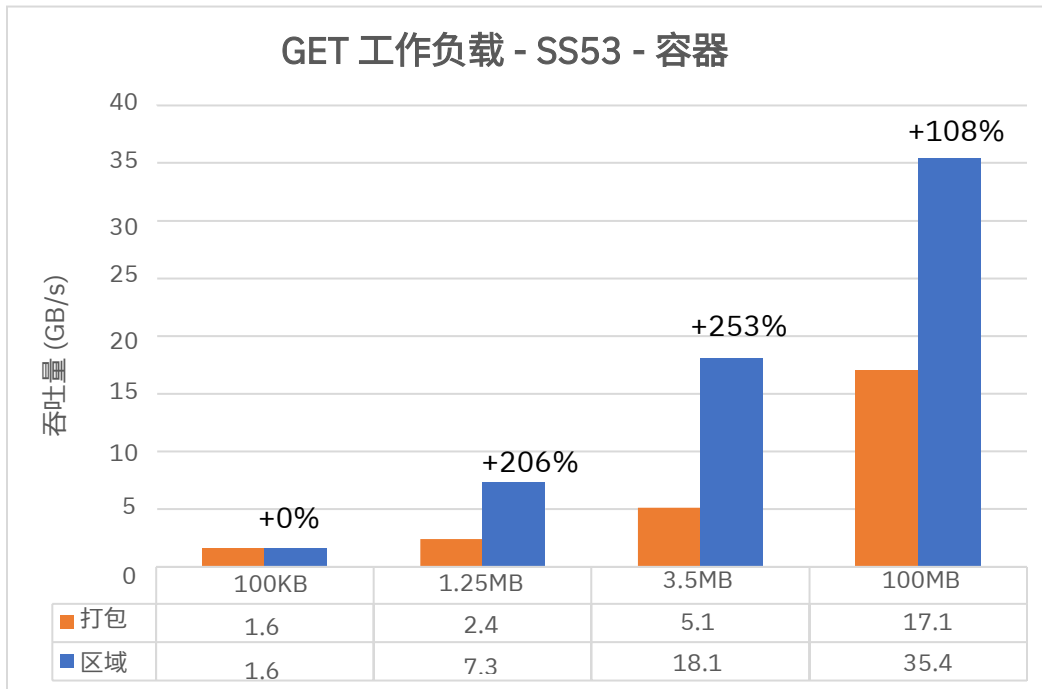


图 6: 数字越大表示性能越高

PUT 和 GET 的性能均有显著提升，尤其是大小介于 3.5MB 到 100MB 的文件。这表明了 ZSS 通过顺序写入/读取磁盘数据所带来的优势。

SS106 Slicestor 设备的容器模式吞吐量测试（使用相同的配置）：

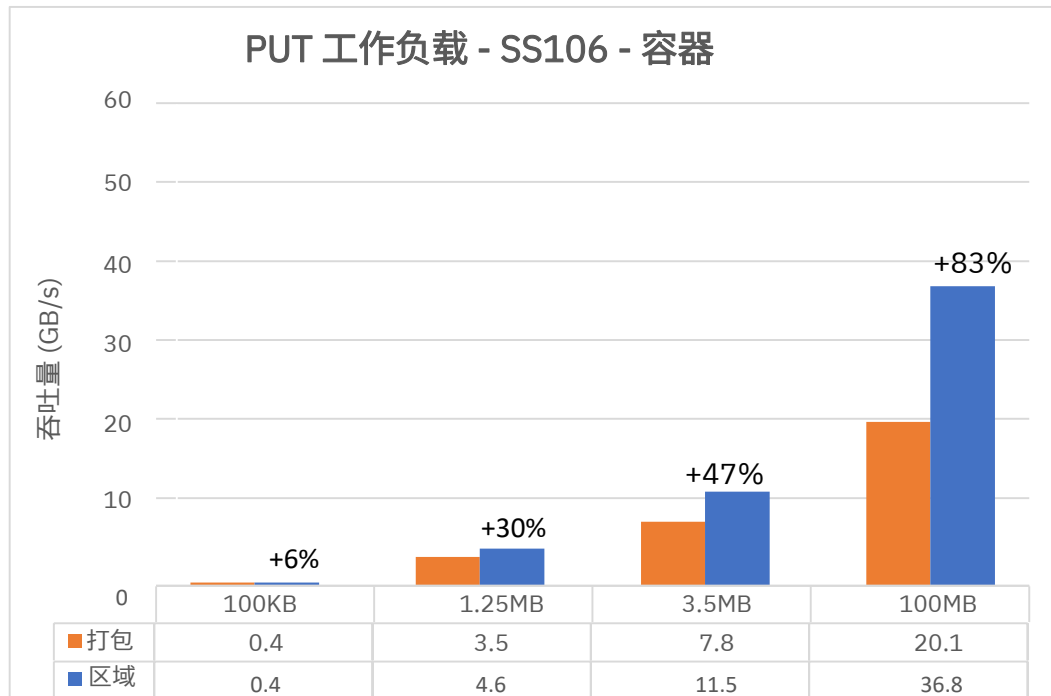


图 7: 数字越大表示性能越高

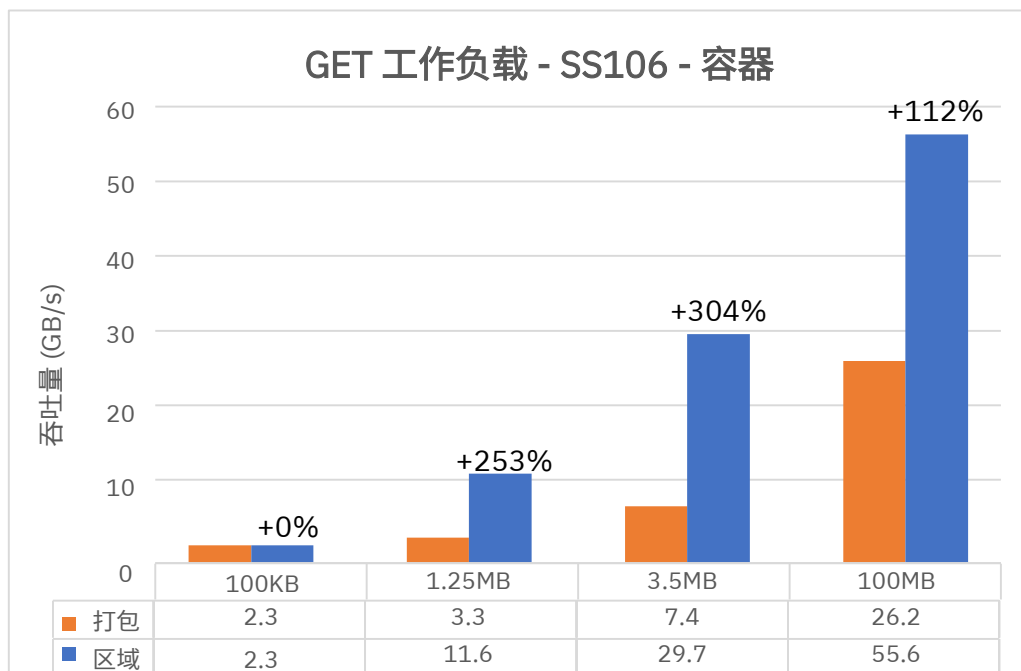


图 8: 数字越大表示性能越高

与保险库模式下的测试一样，在容器模式下，随着主轴数量的增加，PUT 和 GET 的吞吐量均有显著提升。同样，在这些测试中，大对象的 PUT 和 GET 测试结果与 NIC 密切相关。

其余测试的目的在于比较 ZSS 相比于 PSS 在 PUT 和 GET 延迟方面的改善。这些测试在 2212 Slicestor 设备的 12 宽系统上执行，其中采用 4105 Accessor 设备托管负载驱动程序。该系统被拆分为三个站点，每个站点之间注入固定的网络延迟。

配置：

- 三个站点
- (12) 台配备 4 TB 驱动器和 10G NIC 的 2212 Slicestor 设备
- (3) 台 4105 Accessor 设备
- 保险库模式
- 12/8/10 IDA
- 固定的站点间网络延迟
- SecureSlice 已禁用
- 分段大小 4 MiB
- 每次测试的固定对象大小：100 KB、1.25 MB、3.5 MB、100 MB

该系统的 ZSS PUT 延迟结果与 PSS 相当（未显示）。不过，ZSS 的 GET 结果显示，各种大小的测试对象的延迟都有所改善。请注意，在下文图表中，延迟（负百分比）数字越小表示性能越高。

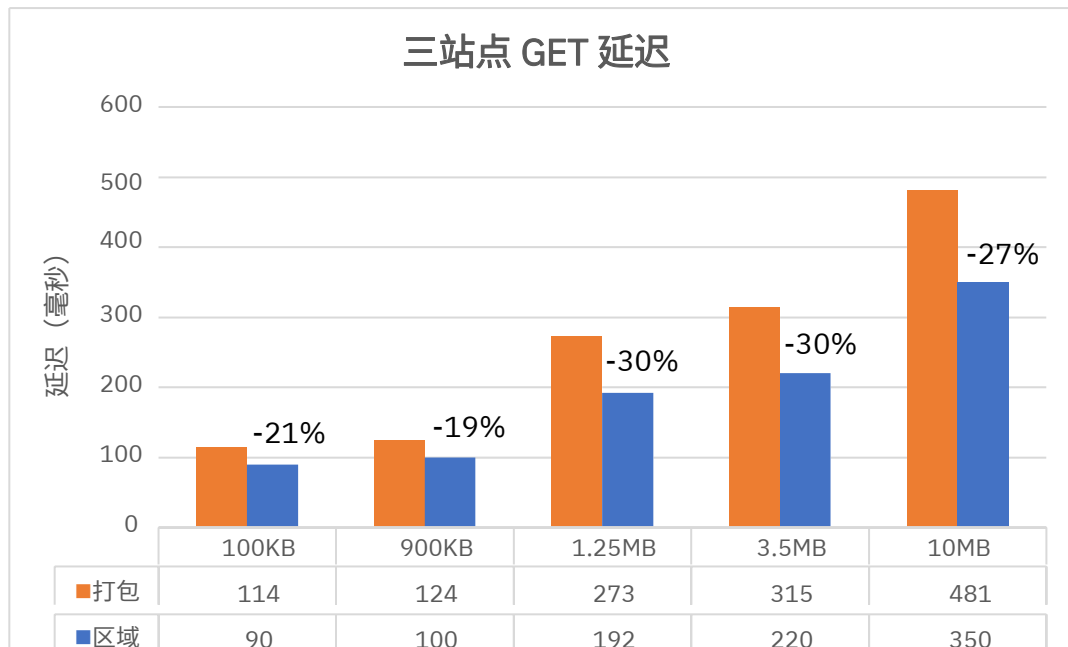


图 9：数字越小表示性能越高

总结

Zone Slice Storage (ZSS) 是 IBM Cloud Object Storage 系统的首选新存储引擎，测试结果也表明，ZSS 可为客户提供真正的价值。ZSS 能够显著提升读写吞吐量，尤其是对于较大的对象而言。PUT 吞吐量的提升幅度高达 180%，GET 吞吐量的提升幅度高达 300%。除了吞吐量提升之外，GET 延迟的降低幅度高达 30%。

这些改善均不会对使用 IBM Cloud Object Storage 的应用造成影响。在部署新的 COS 系统时，或在现有 COS 系统中添加新的增量存储并进行并发运行时均可以使用 ZSS，进而确保它对最终用户或应用透明。除了这些性能和操作优势之外，配备了 ZSS 的 COS 还支持使用更高密度的 SMR 驱动器，而且可以通过基于 SSD 的 Slicestor 设备提高性能。



为什么选择 IBM?

创新技术、开放标准、卓越性能及久经验证的存储软件与硬件解决方案组合，均由公认的行业领导力提供支持，这些就足以令贵企业选择 IBM 的存储解决方案。此外，IBM 可交付业内最佳的存储产品、技术、服务和解决方案，贵企业无需与不同的硬件和软件供应商接洽，便可满足所有需求。

有关更多信息

如欲了解有关 IBM Spectrum Storage 的更多信息，请联系您的 IBM 代表或 IBM 业务合作伙伴，或访问以下网站：
ibm.com/cn-zh/it-infrastructure/storage



© Copyright IBM Corporation 2020.

IBM、IBM 徽标及 [ibm.com](https://www.ibm.com) 是 International Business Machines Corp. 在世界各地司法辖区的注册商标。其他产品和服务名称可能是 IBM 或其他公司的商标。Web 站点 <https://www.ibm.com/legal/us/en/copytrade.shtml> 包含了 IBM 商标的最新列表；Web 站点 https://www.ibm.com/legal/us/en/copytrade.shtml#section_4 包含了可能在本文档中提及的所选第三方商标列表。

本文档中包含了与以下 IBM 产品（IBM Corporation 的商标和/或注册商标）相关的信息：

IBM®、IBM Aspera®、IBM FASP®、IBM POWER®、NVIDIA®、IBM Cloud™、IBM Linear Tape File System™

有关 IBM 未来发展方向及意图的声明如有变更或撤销，恕不另行通知，且仅用于说明目标之用。