



目录:

- 1 概述
 - 2 面临的问题
 - 3 根本解决方案 – IBM Aspera 云直连传输
 - 5 第三方云存储平台验证
 - 5 结论
-

IBM Aspera 云直连存储

本技术白皮书介绍一流高速传输云直连存储和对第三方云存储平台的支持

概述

Aspera FASP® 高速传输平台支持向/从一些领先第三方云存储平台以及在平台之间进行高性能、安全的 WAN 文件、目录和其他大型数据集传输。该实施是 Aspera 服务器软件中一个增强型传输堆栈和虚拟服务系统层，支持使用 FASP 协议和特定第三方文件系统的本机 I/O 功能通过 WAN 进行对象存储直连传输。该堆栈存在于所有正式发布的 Aspera 服务器软件产品，支持与所有正式发布的 Aspera 客户端软件进行互操作传输。

Aspera 随着市场需求发展不断增添对新的第三方存储平台的支持，非常荣幸地在 3.4 版中提供对所有领先云存储平台的支持，包括面向 IBM® SoftLayer® 和 Rackspace 的 OpenStack Swift (v 1.12)、Amazon S3、Windows Azure BLOB、Akamai NetStorage、Google Storage 和 Limelight Cloud Storage。本白皮书概要介绍该平台的开发动机 – 向/从云环境传输大型数据集的根本问题 – 详细介绍了平台功能，并描述了性能和功能测试，包括各存储平台验证。



面临的问题

主流“云”存储平台是“对象存储”架构，其设计源自于由领先网络搜索公司开发的早期横向扩展存储系统，例如 Hadoop File System (HDFS)、Google File System (GFS) 和 Amazon Dynamo。这些对象存储系统的关键设计原则是将文件数据以及名称、权限、访问时间等关联元数据组织为“对象”，并以分离的形式存储文件数据及引用文件数据的元数据，从而实现极高规模和吞吐量。文件数据以冗余副本存储在分布式商品存储设备中以实现可靠性，并通过单一命名空间（其中主表存储着对象标识符哈希和对磁盘上文件数据副本的引用）实现规模，因而可以快速、全面地处理整个分布式平台上的各个对象（请参见图 1）

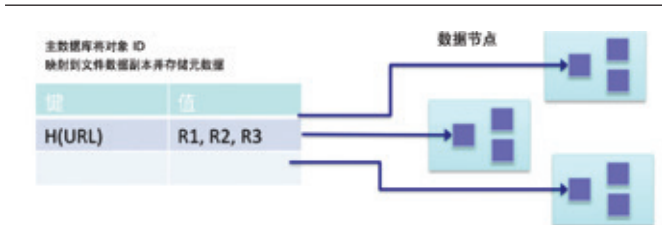


图 1：云对象存储将文件数据与用作标识的元数据分离开，并在整个底层存储分发文件数据

这种方法非常适合用作可扩展 Web 搜索等应用的存储，因为它允许应用利用极大数据集，实现非常高的批处理聚合吞吐量，并将低廉的磁盘用于底层存储。

这种“对象存储”平台的可扩展性看起来似乎也非常适合存储大型非结构化数据类型，例如大型文件和目录。不过，对象存储设计的核心是假定文件数据以小型“块”（通常为 64 MB 至 128 MB）写入存储系统，并以冗余方式存储到众多物理磁盘。每次写入需要将每个块的多个冗余副本写入磁盘，并在主元存储器中创建对这些副本的引用。类似地，要“读取”对象，只能通过查询包含该对象的块、从存储设备中检索并重组。

一个应用程序要上传或下载任何超过块大小（例如 64 MB）的单个项目，必须将对象分割并重组为合适的块；除非以高度并行方式进行，否则这本身就是一项繁琐的工作，并在局域传输速度中存在瓶颈。例如，对于 64 MB 的块，写入一个 1 TB 文件需要分割成超过 10,000 块，实际实施中的吞吐量最高不足每 I/O 流 100 Mbps。我们将这称作局域存储瓶颈（请参见图 2）

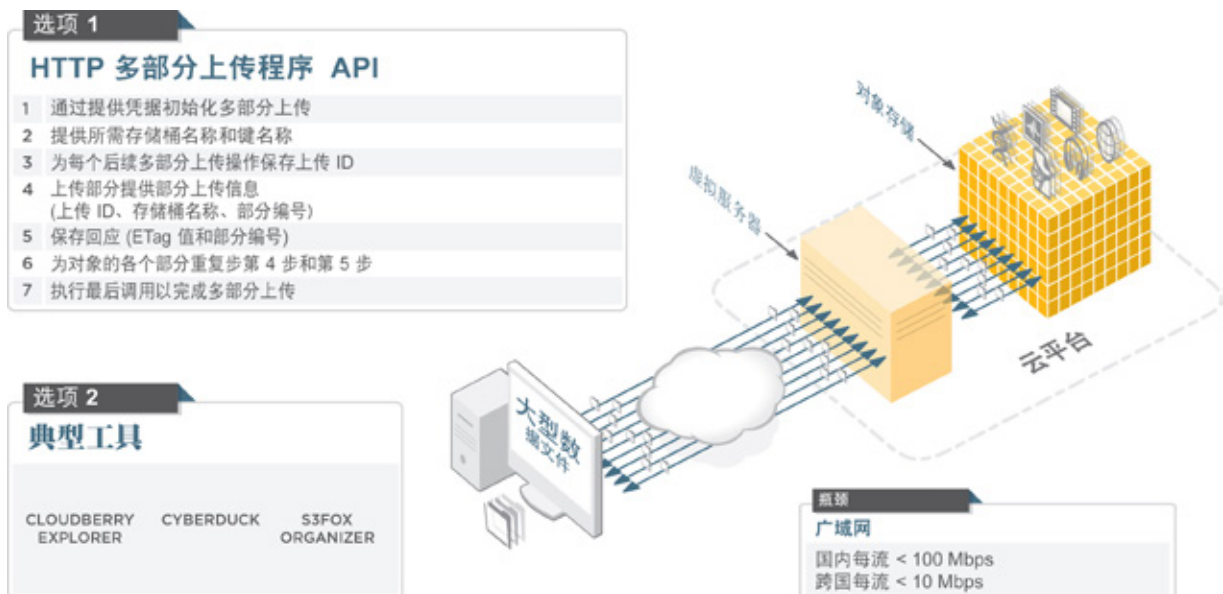


图 2：多部分 HTTP 传输 API 面临局域 I/O 瓶颈和广域传输瓶颈

顾名思义，云存储平台通常与上传或下载的应用程序的 WAN 距离进行定义，因此这种块形式的传输还受制于 WAN 上 TCP 基本性能限值。具体而言，由主流云存储系统实施的 S3 兼容“多部分”对象写入和读取 API 把 HTTP 用作 PUT 或 GET 各对象块的可靠传输机制。在典型的国内 WAN 距离上，往返延迟和数据包丢失足以将可实现的吞吐量限制为 <100 Mbps，而跨国 WAN 将可实现的吞吐量限制为 <10 Mbps。我们将这称作 WAN 传输瓶颈（请参见图 2）

除了存储和传输瓶颈，“多部分”API 不支持在活动会话中断的情况下恢复上传/下载，而是将该工作留给应用管理。虽然 HTTPS 传输将确保网络“块”传输安全性，但多数云存储没有静态加密选项，或要求应用使用云文件系统中的加密选项，而其速度可能会非常慢，因而会对高速上传或下载造成另一个瓶颈。最后，浏览对象存储以查看大型文件和目录之类的免费功能要求在对象存储 API 的基础上构建，呈现给最终用户的不是令人熟悉的文件系统层次结构。

为了绕过使用多部分 API 的编程和部署挑战，一些应用转向“s3fs”等虚拟文件系统驱动程序，以虚拟形式“安装”对象存储。s3fs 是基于 FUSE 的文件系统，由 Amazon S3 提供支持。这可以方便地将对象存储作为分层传统文件系统呈现给应用，但代价是极其缓慢的吞吐速度。s3fs 的大型文件读写速率受限于不足 100 Mbps。

我们需要一款根本性解决方案，能够支持大型文件和目录直接上传和下载到对象存储，同时保持高速、安全性和可靠性，然而云存储平台本身并不存在这种解决方案。Aspera 的云直连传输功能作为一款根本性解决方案进行了全新设计，经过拓展现已支持所有主要的商用云存储平台。

根本解决方案 – IBM Aspera 云直连传输

Aspera 云直连传输平台是独一无二的根本性解决方案，用于向/从云存储以及在云存储之间传输文件和目录。它基于 FASP 传输技术而构建并与对象存储深入集成，为云存储带来了 Aspera 传输平台的所有特性：在单个传输流中，无论网络距离如何，都能以最大速度将文件和目录上传到云、从云下载并在云之间传输 – 无需并行流处理，同时支持存储平台所允许的最大文件和目录大小¹。传输速率通过 Aspera 专利

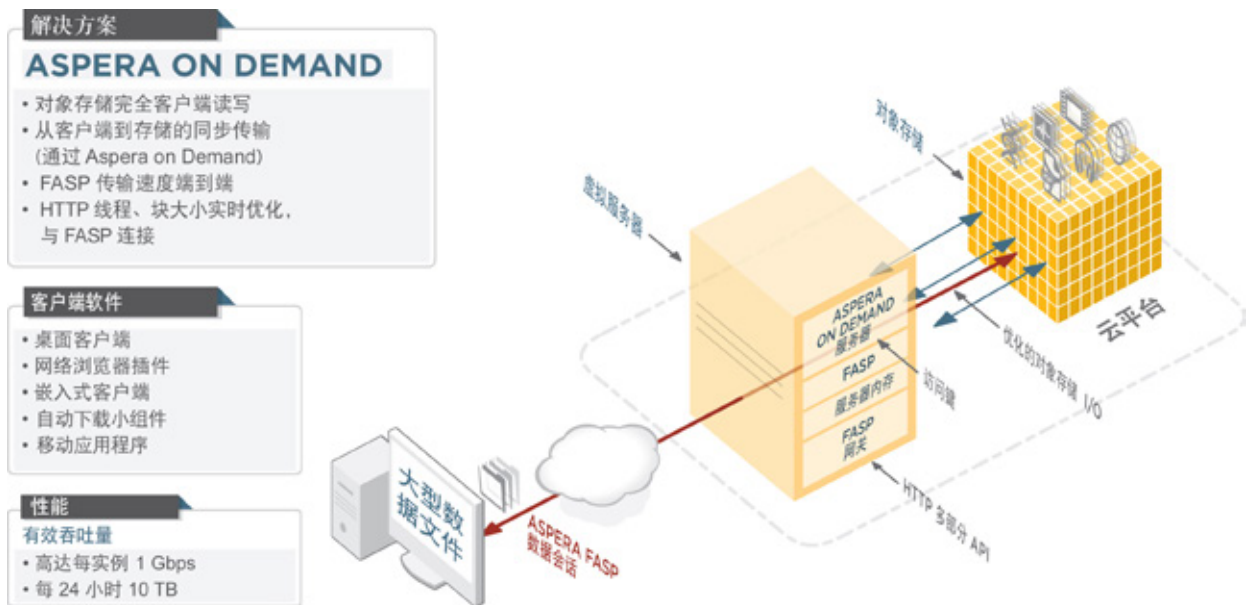


图 3：Aspera 云直连传输是利用云对象存储传输大型文件和目录的根本解决方案，借助与对象存储的深度集成提供本机端到端 FASP 传输功能。

动态速率控制自动适应可用网络带宽和存储带宽，多个传输的聚合带宽通过 Aspera 的 vlink 技术进行精确控制。该平台借助网络和静态加密解决云中数据相关的基本安全问题，并使用本机存储凭据对所有传输和浏览操作进行验证，从而在多租户存储环境中提供隐私保护。中断的传输自动重启并从中断点恢复。所有 Aspera 客户端都支持安全文件浏览和传输，包括浏览器、桌面、CLI 和嵌入式 / SDK 模式。

功能详细信息如下文所述。

- **性能不受距离影响** – 最高速单流传输，不受往返延迟和数据包丢失（500 毫秒 / 30% 包丢失+）影响，最高为平台的 I/O 极限。
- **大型文件大小** – 单次传输会话所支持的文件和目录大小高达采用默认 64 MB 多部分块大小的特定平台所支持的最大对象大小，例如 AWS S3 上单次会话为 0.625 TB。（最新软件版本拥有可配置的块大小，可传输平台所支持的最大对象大小。）
- **小型文件的大型目录** – 借助高速支持包含任意数量单个文件的目录，包括超大数量的极小型文件（通过 WAN 以 100 Mbps 速度传输大小为 1-10 KB 的文件集）。
- **自适应带宽控制** – 基于网络和磁盘的拥塞控制，可让传输速度自动适应向/从存储平台的可用网络带宽和可用 I/O 吞吐量，以避免拥塞和负载过度。
- **自动恢复** – 支持几乎任何传输（单文件和目录）从中断点进行自动重试和检查点恢复。
- **内置加密和静态加密** – 内置网络加密和静态加密 (AES 128)。
- **安全身份验证和访问控制** – 内置对使用私有云凭据实施的已验证 Aspera docroots 的支持。支持按用户帐户进行可配置的读取、写入和列表访问。支持基于平台特定角色的访问控制，包括 Amazon IAMS 和 Microsoft Secure SaaS URL。
- **无缝、全功能的 HTTP 回滚** – 在受限的网络环境中无缝回滚到 HTTP(s)，包括支持完全支持加密、静态加密以及自动重试和恢复。
- **并发传输支持** – 并发传输支持可扩展为环境中每个 VM 实例约 50 个并发传输。（云存储平台支持并发会话的能力存在差异，具体取决于平台成熟度和特定 VM 主机到云文件系统架构的容量。）
- **文件属性保留** – 对于 ES 3.5 和更高版本，传输可以配置为依据 AWS S3 和 Swift 保留文件创建、修改时间。（发行前版本可用时间早于 3.5）。
- **完全兼容 Aspera Client** – 完全互操作的传输支持，所有核心 Aspera 产品作为云存储的传输对等点，包括 IBM® Aspera® Enterprise Server、IBM® Aspera® Connect Server、IBM® Aspera® Connect、IBM® Aspera® Desktop Client、IBM® Aspera® Point-to-Point Client、IBM® Aspera® Cargo、IBM® Aspera® 移动客户端、IBM® Aspera® Add-in for Microsoft® 和 IBM® Aspera® Drive。
- **全功能传输模式** – 对这些产品中的所有传输模式提供全面互操作传输支持，包括命令行 (CLI)、互操作点击式 GUI、浏览器、热文件夹自动化和 SDK 自动化。
- **全面的服务器功能** – 全面支持所有 Aspera 服务器功能，包括安全 docroots、BW 控制台配置、安全性和文件处理策略以及上报 Aspera Console。
- **支持正向和逆向代理** – 向/从云环境的传输支持正向或逆向模式的客户端 Aspera 代理。
- **全面的 SDK 功能** – 服务器端软件支持所有核心 Aspera 传输和管理 SDK，包括 Connect JavaScript API、faspmanager、用于作业初始化的 SOAP 和 REST Web 服务、可靠查询、通过 stats collector 进行报告聚合以及自动后处理脚本。

第三方云存储平台验证

为了在我们发行的软件中提供对新对象存储平台的支持并验证对存储平台的支持，Aspera 执行一整套自动和手动测试，以验证 WAN 条件性能、大型文件大小和数量、文件完整性、并发性、负载测试、包括加密和访问控制的安全性以及版本之间的向下兼容性。Aspera 旨在在平台所提供测试主机的数量、种类和网络连接限制范围内对所有平台运行相同的测试集合和条件。运行 Aspera 服务器软件的单个虚拟主机计算机的测试实例和性能功能的参数将在下页的表 1 中按平台详细介绍。

当前认证和受支持的平台

从 Aspera 核心产品 3.4 版上线开始，Aspera 在正式发布版本中对以下云存储平台提供官方支持：

- Amazon AWS S3
- Swift 对象存储，例如 IBM Softlayer 和 Rackspace 1.12 和更高版本
- Microsoft Azure BLOB
- Akamai NetStorage
- Google Storage
- 对 Limelight Cloud Storage 提供 Beta 支持

结束语

当今市场多数基于云的存储基于对象存储。对象存储架构的关键设计原则是分离文件数据与元数据、在分布式商品存储中复制数据，以及跨分布式节点和集群进行统一访问。相比基于块的传统存储，这些原则可实现具有更高冗余性和耐用性的具有成本效益的横向扩展。

不过，同样是这些属性在存储大型非结构化数据方面带来挑战。大型文件必须被分隔成“块”，并在“写入”时进行独立存储、在“读取”时进行重新组装。加上通过具有高往返时间和数据包丢失的长距离 WAN 连接移动大型文件的传统瓶颈，由于传输速度和吞吐量大幅下降、文件传输和存储延迟加大，对于大型非结构化数据而言，基于云的对象存储变得不符合实际。

Aspera FASP 高速传输平台支持向/从云存储平台以及在这些存储平台之间进行高性能、安全的 WAN 文件、目录和其他大型数据集传输。FASP 克服了 WAN 数据移动瓶颈，并最大限度提高并稳定了到底层对象存储的吞吐量。与对象存储 API 的深入集成可实现最高传输性能，并提供关键传输管理功能，例如暂停、恢复以及网络和静态加密。通过其设计模型和对基于云的领先对象存储平台的性能验证测试，我们可以发现 FASP 是实现至/从云存储和在云存储之间可靠数据传输的理想下一代技术。

测试领域	测试类型	测试范围	范围值	平台限值
并行性	负载测试、 压力测试、 峰值测试、 饱和测试、	负载	服务器：并发会话 [1、10、25、35、50] 客户机：并发会话 [1, 10, 25]	S3 最多并发 25, Azure 最多并发 6, Google 最多并发 6
		方向	上、下、混合	
		文件集	0-10Mb 为 3800, 1MB 为 5000	
		带宽	[25%、50%、75%、100%、125%、150%] (1 Gbps 容量)	
		持续时间	10 分钟、1 小时、8 小时、2 天	
		数据包延迟	0、100毫秒	
		加密	开、关	
		恢复	无、元数据、稀疏、全部	
		操作系统	Windows、*nix	
		流量峰值	35 - 50 个并发会话	
		云存储	负载测试、	文件大小
带宽	10 Mbps、100 Mbps、500 Mbps			Azure 最大带宽 400 Mbps
并行性	2、6 客户机; 8、12 服务器			Azure 最多 6 个并发 Google 最多 6 个并发
加密	开、关			
方向	上、下、混合			
数据集	小型文件 - 0 byte 至 100KB (42 万个文件) 中型文件 - 1MB 至 100MB (9000 个文件) 大型文件 - 1GB 至 100GB			Google 最大文件 62 GB (利用 Google 的全新实验性“10,000 部分”支持实现 625 GB)
传输策略	低、中、高、固定			
压力测试	文件大小		0 byte 至 400 GB	s3 最大文件为 625 GB (利用默认配置, 单个会话可提高至 5 TB) Google 最大会话 62 GB (利用 Google 的全新实验性“10,000 部分”支持实现 625 GB) Azure 最大会话 200 GB
	带宽		500 Mbps 至 1 Gbps	Azure 最大带宽 400 Mbps
	并行性		12、15、20	Azure 最多 6 个并发 Google 最多 6 个并发
	加密		开、关	
	方向		上、下、混合	
	数据集		小型文件 - 0 byte 至 100 KB (42 万个文件) 中型文件 - 1MB 至 100MB (9000 个文件) 大型文件 - 1GB 至 100GB	Google 最大文件 62 GB (利用 Google 的全新实验性“10,000 部分”支持实现 625 GB)
	传输策略		低、中、高、固定	
向下兼容性测试	产品版本		ES 3.3.4、ES 3.4.0	
饱和测试	文件大小		0 byte 至 10 GB	s3 最大文件为 625 GB (利用默认配置, 单个会话可提高至 5 TB) Google 最大会话 62 GB (利用 Google 的全新实验性“10,000 部分”支持实现 625 GB) Azure 最大会话 200 GB
	带宽		10 Mbps、300 Mbps	Azure 最大带宽 400 Mbps
	并行性		4, 6	
	方向		上、下、混合	
	传输策略		低、中、高、固定	
	持续时间		100 小时	
文件完整性测试	文件大小		10 byte、4MB、64MB、100MB、1GB	
	方向		上、下	
	加密	开、关		
系统测试	产品	faspex™、Console、Shares		
	文件大小	10 byte 至 1 GB (各种/现实世界)		
	方向	上、下、混合		

表 1: Aspera 云存储验证测试。请注意: 所有测试都在功能与 EC2 m3.xlarge AOD 相当的环境中的单个虚拟机主机上运行 Aspera 服务器软件 (3.4.3 版), 相同地区采用 16 GB 内存和 S3 存储桶

IBM 软件
技术白皮书

测试领域	测试类型	测试范围	范围值	平台限值
WAN	性能测试	带宽	512Kbps、1Mbps、10Mbps、155Mbps、622Mbps、1Gbps、3Gbps、10Gbps	
		往返时间	0 毫秒、2 毫秒、10 毫秒、100 毫秒、300 毫秒、500 毫秒、1000 毫秒	
		数据包丢失率	0%、0.1%、1%、5%、10%、20%	
		平均文件大小 (数据集)	1KB、10KB、100KB、1MB、5MB (小型媒体文件)、10MB	
		并发性	1、10 (更高并发性将在 ssh 负载测试中进行测试)	
		负载过度	2, 10 和 100	
		加密	启用和禁用	
		块大小 - 读取和写入	16KB、64KB、128KB、256KB、512KB、1MB、4MB	
		路由器缓冲大小 (队列深度)	10 毫秒、100 毫秒、250 毫秒	
		方向	上传、下载	
		操作系统	主要操作系统	
安全性	功能测试	传输加密	开、关 (全面的用例)	在受控的实验室环境中执行
		EAR	上传、下载、FASP/HTTP	
		文件校验和	MD5、SHA1、无	
		ssh 指纹	ascp、HTTP 回滚	在受控的实验室环境中执行
		HTTP 代理	访问控制、令牌	
		DNAT 代理	配置、并发性 (20)、恢复、http 回滚	在受控的实验室环境中执行
		令牌授权	上传/下载、文件、列表和 pair-list、FASP 和 HTTP 回滚、令牌密码	
		HTTP 回滚	令牌授权、伪造的请求	

术语表

测试类型	定义
负载测试	验证目标负载条件下的产品行为
压力测试	评估运行负载大幅高于目标水平时的产品行为
峰值测试	验证反复出现压力水平峰值时的产品行为
饱和测试	验证一段时间内运行在目标负载条件下的产品行为
系统测试	验证功能在 IBM Aspera 产品集成环境中的功能性
性能测试	评估并衡量指示测试范围中文件传输性能
文件完整性测试	验证指示测试范围中的完整性和传输的文件数据
功能测试	验证指示功能范围中产品功能性行为
向下兼容性测试	验证针对早期产品发行版的功能性和非功能性产品行为

关于 IBM 旗下的 Aspera 公司

IBM 旗下的 Aspera 是下一代传输技术的发明者，无论文件大小、传输距离和网络条件如何，该技术能够以最高速度传输数据。基于其已获专利的 Emmy® 屡获殊荣的 FASP® 协议，Aspera 软件充分利用现有基础架构提供最快、最具预测性的文件传输体验。Aspera 的核心技术提供前所未有的带宽控制、全面的安全性和不折不扣的可靠性。六大洲各行业的组织依靠 Aspera 对数字化资产进行业务关键型传输。

有关更多信息

如需有关 IBM Aspera 解决方案的更多信息，请访问：

ibm.com/software/aspera 并在 Twitter

[@asperasoft](https://twitter.com/asperasoft) 上关注我们。



© Copyright IBM Corporation 2015

IBM Corporation
Route 100
Somers, NY 10589

2015 年 1 月

IBM、IBM 徽标、ibm.com 和 Aspera 是 International Business Machines Corporation 在美国和/或其他国家/地区的商标或注册商标。如果这些和其他 IBM 带商标的术语在本文中第一次出现时带有商标符号 (® 或 ™)，则这些符号表示在本文发布时它们是 IBM 的美国注册或习惯法商标。此类商标在其他国家/地区也可能是注册商标或习惯法商标。IBM 商标的最新列表可在下述网页的“版权和商标信息”中查看：ibm.com/legal/copytrade.shtml。

Java 和所有基于 Java 的商标与徽标均为 Oracle 和/或其子公司的商标或注册商标。

Microsoft、Windows、Windows NT 和 Windows 徽标是 Microsoft Corporation 在美国和/或其他国家/地区的商标。

SoftLayer® 和 SoftLayer® 设备是 IBM 旗下公司 SoftLayer, Inc. 的商标或注册商标。

其他产品、公司或服务名称是其他公司的商标或服务标记。

本文档为初始发布日时的最新文档，IBM 可能随时对其进行更改。IBM 并未在每个开展业务的国家/地区提供所有产品/服务。

本文档中引用的性能数据和客户示例仅作参考用途。实际性能结果可能会有所不同，具体取决于特定的配置和操作条件。对于与 IBM 产品和程序配合使用的其他任何产品或程序，用户应负责相关的评估与验证工作。本文档中的信息“按原样”提供，不带任何明示或暗示的保证，包括不带任何适销性、对特定用途的适用性的保证以及任何不侵权的保证或条件。IBM 产品根据提供这些产品时所依据协议的条款与条件进行保证。

1 使用默认 Aspera 软件配置所支持的最大对象大小取决于存储平台的功能。特定平台的限值在表 1 中列出。



请回收利用