

数据加速器 GooseFS

产品选型指引



腾讯云

【 版权声明 】

©2013–2025 腾讯云版权所有

本文档（含所有文字、数据、图片等内容）完整的著作权归腾讯云计算（北京）有限责任公司单独所有，未经腾讯云事先明确书面许可，任何主体不得以任何形式复制、修改、使用、抄袭、传播本文档全部或部分内容。前述行为构成对腾讯云著作权的侵犯，腾讯云将依法采取措施追究法律责任。

【 商标声明 】



及其它腾讯云服务相关的商标均为腾讯云计算（北京）有限责任公司及其关联公司所有。本文档涉及的第三方主体的商标，依法由权利人所有。未经腾讯云及有关权利人书面许可，任何主体不得以任何方式对前述商标进行使用、复制、修改、传播、抄录等行为，否则将构成对腾讯云及有关权利人商标权的侵犯，腾讯云将依法采取措施追究法律责任。

【 服务声明 】

本文档意在向您介绍腾讯云全部或部分产品、服务的当时的相关概况，部分产品、服务的内容可能不时有所调整。您所购买的腾讯云产品、服务的种类、服务标准等应由您与腾讯云之间的商业合同约定，除非双方另有约定，否则，腾讯云对本文档内容不做任何明示或默示的承诺或保证。

【 联系我们 】

我们致力于为您提供个性化的售前购买咨询服务，及相应的技术售后服务，任何问题请联系 4009100100或 95716。

产品选型指引

最近更新时间：2025-03-18 17:58:02

产品形态对比

维度	GooseFS	GooseFSx	GooseFS-Lite
核心定位	分布式缓存加速服务，面向数据湖存储性能优化。	高性能并行文件存储服务，提供全托管、POSIX 兼容的文件系统。	轻量级本地挂载工具，针对大文件高吞吐读取场景，提供低成本的 COS 数据接入。
架构设计	分布式缓存系统，提供近计算端缓存。	基于分布式架构的文件系统，性能随容量线性扩展，支持多客户端多节点并行访问。	单机轻量级工具，直接通过 FUSE 挂载 COS 桶，无分布式组件依赖。
部署方式	支持全托管、Master 托管、管控面托管三种部署方式。	全托管云服务，一键购买自动扩容，无需运维。	需手动安装依赖（如 FUSE 库）并挂载，无托管选项。
协议支持	支持 HDFS、FUSE、POSIX 协议。	完全兼容 POSIX 语义，支持Windows/Linux 系统挂载。	支持基础 POSIX 操作。

核心功能对比

GooseFS

- 分层缓存能力：通过内存/SSD/HDD 分级存储，将热数据智能调度到计算节点本地，提升数据吞吐。
- 统一命名空间：通过透明的命名机制，可以融合多种不同的底层存储系统访问语义，为用户提供了数据统一管理能力。
- Page Store 缓存：采用内存分页式缓存机制，可以极大地优化离散 IO 访问模型的缓存空间利用率以及冷读效率。

GooseFSx

- 数据流动：数据在数据加速器 GooseFSx 与对象存储 COS 之间按需流动。
- 数据预热：从 COS 存储桶预热数据到 GooseFSx 目录，将您指定的数据（整个目录、某个子目录或清单）自动、完整、增量地预热到 GooseFSx 中。
- 数据沉降：将 GooseFSx 目录的数据沉降到 COS 存储桶，将您指定的数据（整个目录、某个子目录或清单）自动、完整、增量地沉降到 COS 中。

- 云盘多挂：将一块云盘同时挂载到多个 GooseFSx 存储节点。云盘多挂容忍任意多个节点同时故障，业务不中断、数据不丢失，极大提升产品可用性（从99.9%提升到99.9999999%）。

GooseFS-Lite

轻量级挂载：支持将对象存储 COS 桶挂载到本地文件系统。兼容 POSIX 文件操作（顺序读/写、目录操作），但不支持随机写、truncate 操作及软硬链接。

优势对比

GooseFS

- 性能强劲：基于分布式缓存架构，为用户提供近计算端的高性能数据访问能力，显著降低数据访问延迟。
- 成本集约：充分利用计算节点的闲置本地盘资源提供数据访问加速能力，提升资源利用率。
- 生态亲和性：深度适配主流计算框架生态，支持 Spark、TensorFlow 等与大数据与 AI 计算框架无缝对接。
- 易用性：提供全托管、Master 托管、管控面托管三种部署方式，根据用户实际情况选择。全托管模式无需用户运维集群。
- 稳定性：集成 CLS 日志服务与云原生 Prometheus 监控体系，构建多维度的实时健康监测系统，简化运维流程，提升稳定性。

GooseFSx

- 超高性能：可提供每秒数百 GB 吞吐量、百万级 IOPS 及亚毫秒级延迟。
- 与计算生态无缝融合：完全支持 POSIX 文件语义，无需修改代码即可适配 HPC、AI 训练等场景。支持自动批量挂载功能将存储空间映射为本地目录。
- 数据流动：支持从 COS 快速预热训练数据集至 GooseFSx，并可将生成结果自动沉降回 COS。
- 冷热分层弹性高效：GooseFSx 与 COS 相互解耦，各自弹性扩展且深度融合。
- 简单易用：全托管服务，控制台一键部署，无需进行集群运维。

GooseFS-Lite

- 轻量化部署：采用客户端工具形式部署，无需独立部署缓存集群或分布式系统，仅需在计算节点上安装即可。
- 低成本与资源复用：GooseFS-Lite 直接利用计算节点本地磁盘或内存资源进行数据缓存，既节省了硬件投入，又避免了跨节点数据同步带来的带宽消耗。

适用场景对比

GooseFS

- AI 训练与推理：加速数据预处理（如数据清洗、小文件加载），减少 GPU 等待时间。
- 大数据分析：提升 Spark/Flink 等框架对 COS 数据的访问性能，降低作业延迟。

- 自动驾驶：加速本地 IDC 与云上数据交互，优化路采数据预处理效率以及自动驾驶训练业务。
- AI 内容生成（AIGC）：缓存热数据（如模型参数、向量数据集），提升多模态模型训练效率。

GooseFSx

- AI 训练与推理（C50/C60/C70）：支持 Checkpoint 高速写入、模型训练输出沉降，提供均衡读写性能。
- 自动驾驶（C50/C60/C70）：提供端到端解决方案，提供采（路采车数据上云入湖）、算（立即训练）、存（长期持久化保存）全过程全周期的一体化服务。
- 高性能计算（C50/C60/C70）：提供高性能、低延迟、大吞吐的并行文件服务，充分满足 HPC 计算的高吞吐、低延时的需求；与数据湖底座 COS 联动，提供超高性能、超大规模、超低成本存储服务。
- 基因分析（C50/C60）：加速基因测序/局部比对等阶段的高性能存储需求；与数据湖底座 COS 自由流动数据，实现组学平台立即访问 COS 的样本，实现基因分析结果自动归档至 COS，通过 COS 互联网分发能力交付给终端用户。
- CAE/CAD（C60）：加速 CAE/CAD 小文件读写，与数据湖底座 COS 联动，提供超高性能、超大规模、超低成本存储服务。
- 影视渲染（C60）：提供一体化存储服务，渲染素材低成本归档在 COS，按需拉起数据加速器 GooseFSx C60 进行渲染，成片沉降到 COS 长期保持，通过 COS 互联网分发能力交付给终端用户。

GooseFS-Lite

需要单机环境下快速挂载 COS 桶（如仿真大文件读取），对复杂操作要求低的场景。

规格与限制对比

❗ 说明：

GooseFS-Lite 为客户端本地化部署模式，其规格与限制依赖本地节点，不参与对比。

对比项	GooseFS	GooseFSx
容量扩展	Master 托管和管控面托管无固定起步容量，缓存空间依赖计算节点本地盘	C50起步9TiB，步长3TiB
		C60 T2起步4.5TiB，步长1.5TiB C60 T12起步36TiB，步长12TiB
		C70起步14TiB，步长4.5TiB
读带宽	全托管模式200MB/s每TiB	C50：120MB/s每TiB

	Master 托管和管控面托管模式吞吐随 Worker 节点数弹性扩展，可支持 Tbps 级别带宽	C60：200MB/s每TiB
		C70：600MB/s每TiB
写带宽	全托管模式与 COS 保持一致	C50：120MB/s每TiB
	Master 托管和管控面托管模式随 Worker 节点数弹性扩展，可支持 Tbps 级别带宽	C60：200MB/s每TiB
		C70：200MB/s每TiB
读 IOPS	全托管模式最大支持20W ops	C50：10,000每TiB
	Master 托管模式 Meduim 机型：10W ops；Large 机型：20W ops XLarge 机型：30W ops	C60：20,000每TiB
	管控面托管根据客户自购 CVM 规格决定	C70：30,000每TiB
写 IOPS	全托管模式与 COS 保持一致	C50：10,000每TiB
	Meduim 机型：10W ops，Large 机型：20W ops； XLarge 机型：30W ops	C60：20,000每TiB
	管控面托管根据客户自购 CVM 规格决定	C70：20,000每TiB
文件数	全托管：最大支持10亿	部署容量不足4万GiB时，每GiB容量支持4万文件数
	Master 托管： Meduim机型：1亿级别 Large 机型：10亿级别 XLarge 机型：100亿级别	
	管控面托管：根据客户自购的 CVM 规格决定	
时延	亚毫秒级	亚毫秒级
支持操作系统	Linux	Linux/Windows