

存储架构详解

—— DAS / NAS / SAN / 分布式存储 对比与选型指南

1 概述

存储架构是数据中心与 IDC 基础设施的核心组成部分。按照对外提供数据访问接口的不同，可分为块存储、文件存储、对象存储三种类型；按照组网连接方式的不同，又可分为 DAS（直连存储）、NAS（网络附加存储）、SAN（存储区域网络）以及分布式存储四大类。

需要明确一个关键概念：块存储、文件存储、对象存储是存储对外提供的接口类型；DAS、NAS、SAN 是存储的组网连接架构，二者分属不同维度，不应混淆。

2 传统集中式存储架构

2.1 DAS（Direct Attached Storage，直连存储）

DAS 是最基础的存储架构，存储介质（硬盘/RAID 卡）直接挂载在服务器内部，或通过 SAS 线缆直连服务器，不存在独立的存储网络。

- 典型形态：服务器本地硬盘、RAID 卡组阵列（如戴尔 R740 服务器 H730P 阵列卡）。
- 访问形式：块存储，操作系统识别为本地磁盘。
- 优点：部署简单、延迟低、成本低，无需额外交换机。
- 缺点：存储资源无法跨服务器共享；单机故障后本地盘数据无法被其他服务器接管；容量与性能扩容受限于单台服务器。
- 适用场景：单台服务器业务、临时算力节点、小型测试环境。

2.2 NAS（Network Attached Storage，网络附加存储）

NAS 是独立的存储设备，直接接入以太网（普通 TCP/IP 局域网），对外提供文件级访问服务。

- 典型协议：NFS（Linux 客户端）、SMB/CIFS（Windows 客户端）。
- 工作原理：存储设备自身管理文件系统，客户端通过网络挂载共享目录。
- 优点：多台服务器可同时挂载同一文件目录，文件共享方便；部署简单，复用现有以太网。

- 缺点：文件协议开销大，IO 性能一般；不适合数据库、虚拟机系统盘等低延迟块业务。
- 适用场景：文档共享、日志存储、素材文件、数据备份。

2.3 SAN (Storage Area Network, 存储区域网络)

SAN 构建独立的存储专用网络，存储机头（控制器）连接磁盘柜，服务器通过 SAN 网络获取裸块设备。对外提供块存储服务，服务器可自行创建文件系统。

SAN 分为两种主流实现方式：

- FC-SAN：基于光纤通道，使用专用光纤交换机，延迟低、性能稳定，是传统企业核心存储的主流方案，但硬件成本高昂。
- IP-SAN (iSCSI)：将 SCSI 命令封装在 TCP/IP 协议中，跑在普通以太网上，成本更低，是中小机房广泛采用的方案。
- 优点：裸块设备，适合数据库与虚拟机磁盘；存储资源池化，多服务器共享；支持快照、克隆、RAID、热备、双活等高级特性。
- 缺点：FC-SAN 硬件昂贵；机头（控制器）是单点瓶颈，扩容到一定规模后性能增长缓慢；独立 SAN 网络维护成本高。

NAS 与 SAN 的核心区别可一句话概括：NAS 给你的是“文件夹”（文件存储，走以太网），SAN 给你的是“一块裸硬盘”（块存储，走专用存储网）。

3 分布式存储架构

分布式存储是当前云计算与智算中心的主流架构。它不依赖专用存储机头，而是将大量普通 x86 服务器的本地硬盘通过软件整合为统一存储池。数据以分片方式结合多副本或纠删码打散到不同节点，不存在中心机头瓶颈，一套集群可同时对外提供块、文件、对象三种接口。

3.1 分布式块存储

- 典型产品：Ceph RBD。
- 用途：虚拟机磁盘、云主机系统盘、数据库数据盘，广泛应用于 VMware/ESXi 与 K8s 容器平台。

3.2 分布式文件存储

- 典型产品：CephFS、GlusterFS。

- 用途：多节点共享目录，AI 训练数据集、HPC 高性能计算，多个算力节点同时读写同一文件。

3.3 分布式对象存储

- 典型产品：MinIO、Ceph RGW。
- 访问方式：基于对象（Object），每个对象由数据加元数据组成，通过 HTTP REST 接口访问，无传统目录树结构。
- 用途：图片、视频、备份、AI 模型权重、日志归档等海量非结构化数据存储。

分布式存储的核心优势：横向扩展（增加节点即扩容容量与性能，规模越大优势越明显）；高可用（多副本/纠删码，节点或硬盘损坏后自动重建数据）；一套集群同时提供块/文件/对象三种服务；可基于通用 x86 服务器构建，硬件成本远低于专用存储设备。其局限在于小规模集群性能不如集中式 SAN，软件栈复杂、运维门槛较高，三副本模式存在 3 倍容量开销。

4 三种存储接口类型对比

表 1 块存储、文件存储、对象存储对比

类型	对外形态	典型协议	典型架构	典型场景
块存储	裸磁盘	iSCSI、RBD	SAN、分布式块	虚拟机、数据库、系统盘
文件存储	共享目录	NFS、SMB、CephFS	NAS、分布式文件	日志、素材、AI 数据集共享
对象存储	对象（Key-Value）	HTTP S3	分布式对象	图片、模型、备份、海量静态文件

5 集中式存储与分布式存储对比

表 2 集中式存储与分布式存储架构对比

对比维度	集中式存储（SAN/NAS 机头）	分布式存储
扩容方式	纵向扩容：升级机头、加硬盘柜，到上限后性能难以提升	横向扩容：新增存储节点，容量与性能同步增长
架构瓶颈	控制器机头为瓶颈	无中心机头，瓶颈分散在各节点
硬件成本	专用存储设备，价格昂贵	通用 x86 服务器，硬件成本更低
可用性	双控机头，但机头故障风险依然存在	多副本/纠删码，跨节点冗余
运维难度	以硬件为主，软件相对简单	软件栈复杂，集群运维要求高
适用规模	中小机房，TB 至几十 PB	大规模 IDC、智算中心，几十 PB 至 EB 级

6 选型建议

在 IDC 与智算场景下，可按以下原则选择存储架构：

- 单台服务器、少量业务：选用 DAS 本地盘 + RAID 方案。
- 多机文件共享、日志与素材存储：选用 NAS。
- 传统数据库、虚拟化业务，要求低延迟与稳定性，中小规模：选用集中式 SAN。
- AI 智算集群、海量数据、需要弹性扩容，同时需要块/对象/文件多种接口：选用分布式存储（Ceph/MinIO）。

7 存储常用高级特性

- 副本：一份数据存储多份，Ceph 集群默认采用三副本策略。
- 纠删码：类似 RAID 机制，空间利用率高，但重建开销较大，常用于冷数据存储。
- 快照：某一时间点的数据只读副本，用于备份与快速恢复。
- 精简置备：只在实际写入数据时才占用物理空间，预先分配虚拟容量，提升空间利用率。
- 分层存储：热数据自动放置 SSD，冷数据自动下沉至 HDD，兼顾性能与成本。