

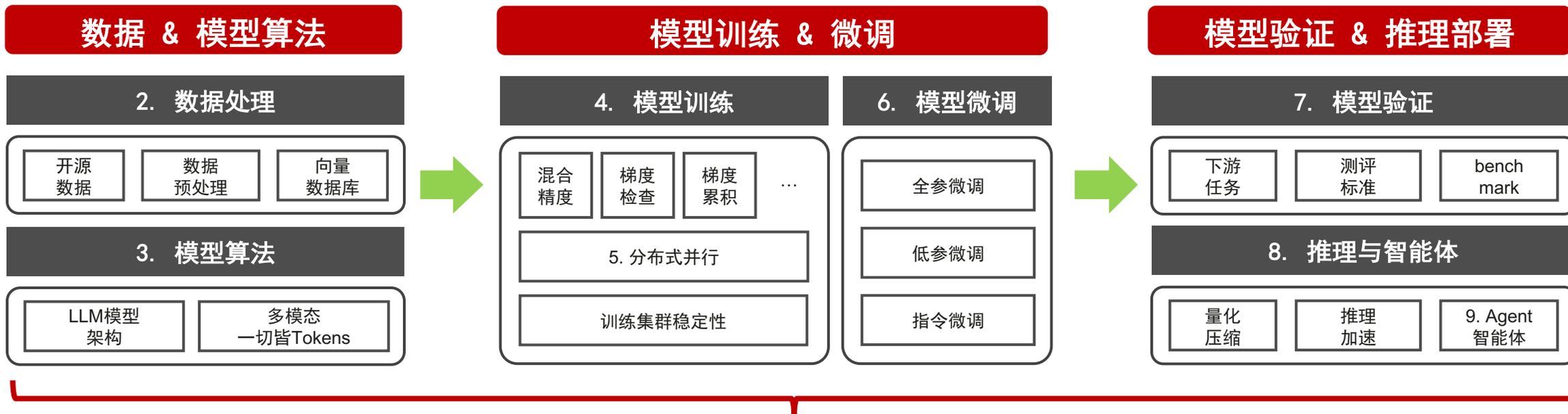
大模型-AI集群 (存)

数据存储 硬件 & 连接



ZOMI

大模型业务全流程



关于本内容

• 内容背景

- AI 集群 + 大模型

• 具体内容

- **数据存储现状和场景**：存储软件类型、存储硬件类型的发展
- **大模型对存储的挑战**：存储性能指标、存储遇到大模型挑战与新机会点
- **大模型训练CKPT优化**：大模型训练过程、CKPT过程分解、CKPT优化
- **大模型时代对存储的思考**：什么样的存储架构才是AI大模型时代的选择？

增长的存储主要靠数据中心和存储服务器拉动

存储 技术 架构

大模型&应用

LLAMA等LLM

多模态大模型

AI Agent智能体

大模型训练

多级存储

CheckPoint

训练数据

存储对象 存储协议

文件存储

NFS

HDFS

块存储

NVMe

SCSI

对象存储

Tensor

Vector

其他

Table

连接方式

SAN

NAS

DAS

存储架构

集中式存储

分布式存储

硬件介质

磁带

光盘

机械硬盘

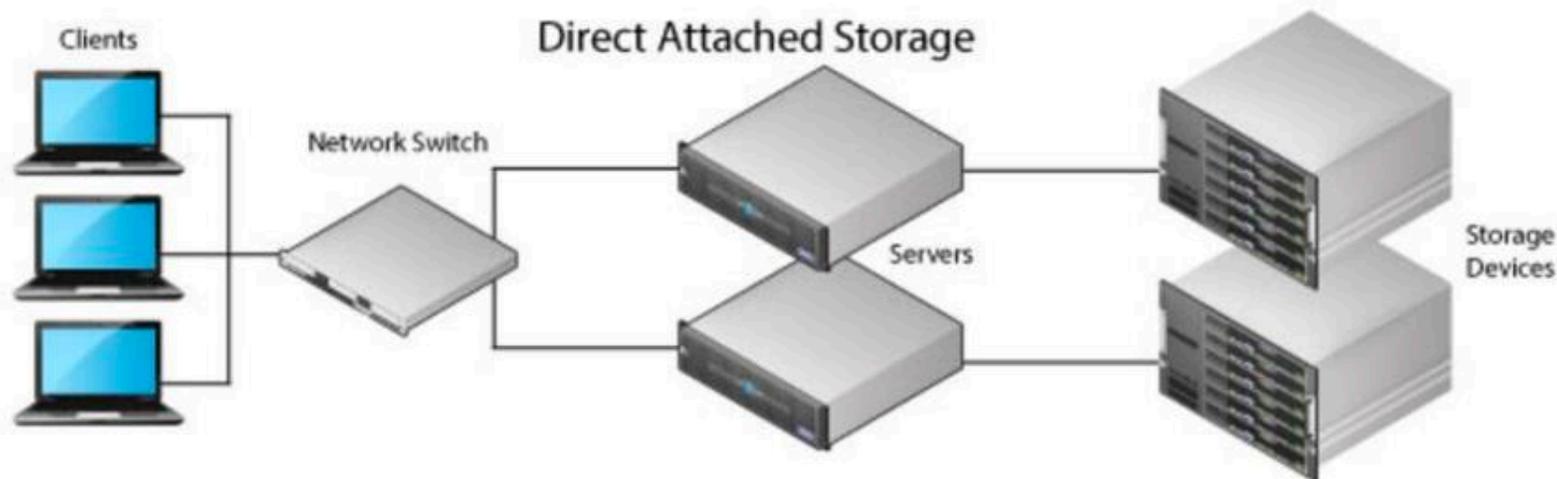
固态硬盘

1. 数据存储の 连接方式



DAS Direct Attached Storage , 直连存储

- 比较熟悉的两种主流存储类型是块存储和文件存储，块存储中 DAS (Direct-Attached Storage 直连存储) 是主流的企业存储方案。直连存储就是存储设备直接与主机服务器连接，其他主机不能使用这个存储设备 (广义上，日常生活使用的 HDD 硬盘可以理解为最简单 DAS 存储)



DAS Pons and Cons

- **DAS 优点**

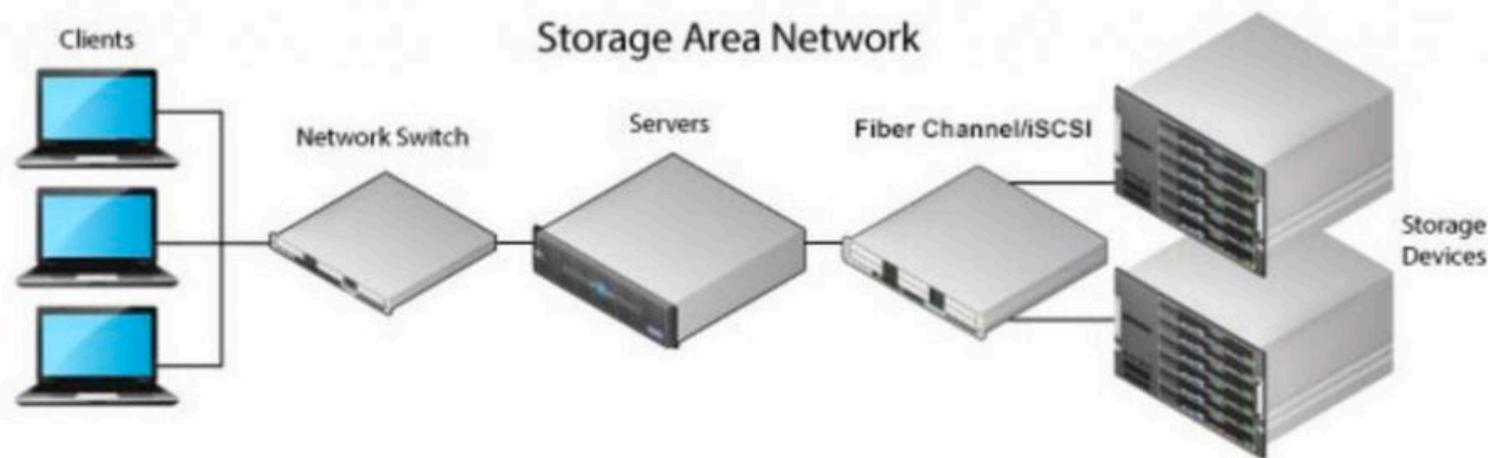
1. **实施简单**：无须专业人员操作和维护，节省用户投资。
2. **大容量存储**：内置较多磁盘和较大存储空间，后端可以挂接磁盘扩展柜，满足海量存储需求。
3. **提高存取性能**：操作单个应用数据，同时有多个物理磁盘并行工作，运行速度比单个磁盘运行速度高。
4. **数据和 OS 分离**：OS 一般存放在服务器硬盘中，应用数据放置于存储阵列中，实现数据和 OS 分离，操作系统升级和应用数据运行相互没有直接影响。

- **DAS 缺点**

1. **故障恢复难**：服务器本身容易成为系统瓶颈，服务器发生故障时，整个存储数据将不可访问。
2. **管理复杂**：对于存在多服务器的系统来说，设备分散，不便于管理；数据备份时的操作会比较复杂。
3. **动态分配**：同时多台服务器使用 DAS，存储空间不能在服务器之间动态分配，可能造成相当的资源浪费。

SAN Storage Area Network , 存储区域网络

- 通过交换机将磁盘阵列等存储设备与相关服务器连接起来的高速专用存储网络。1) FC-SAN , 即基于 FC 设备及通信协议的存储区域网络。2) 光纤网络成本高 , 又发展出IP SAN , 即基于以太网 SAN 存储形式。IP SAN 利用 TCP/IP 协议实现了对 SCSI 协议封装 , 主机端相当于 CS 架构客户端 , 存储端则是服务端。



SAN 优缺点

- **SAN 优点**

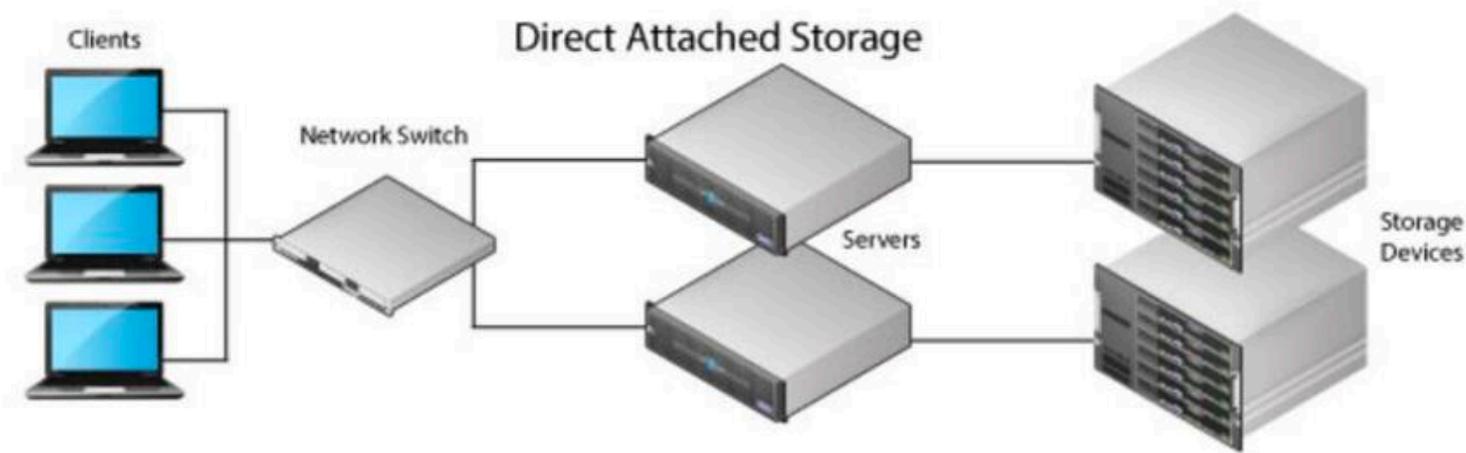
1. **易于扩容**：SAN 提供一个共享存储池；存储空间不足时，可以进行空间扩容，SAN 可以随时增减服务器。
2. **集中管理**：SAN 把前端服务器数据集中到一起，使得数据管理和系统运维都得到极大的简化。
3. **优异存取性能**：SAN 使用光纤传输数据，数据读取性能非常高效非常高速。
4. **方便灵活备份**：重要数据集约化，更有效备份来实施数据保护，运行备份操作时无需考虑对网络总体性能影响。

- **SAN 缺点**

- **价格昂贵**：FC SAN 应用光纤作为通信信道，使得 SAN 整体构架成本居高不下，但是因为SAN高性能、易管理等诸多优点，使得SAN仍然是目前应用最普遍的存储解决方案。

NAS Network Attached Storage , 网络接入存储

- 存储系统直接接入网络，通过网络交换机，将服务器与存储连接在一起，用户通过 TCP/IP 协议访问数据，并通过标准文件共享协议（CIFS、NFS等）实现目录级共享。
- NAS 和 SAN 都构成一个存储网络，可以实现设备集中管理。NAS 访问需要经过文件系统格式转换，所以是文件级访问。SAN 提供是裸设备，适合块存储相关的应用（如数据库）。



NAS 优缺点

- **NAS 优点**

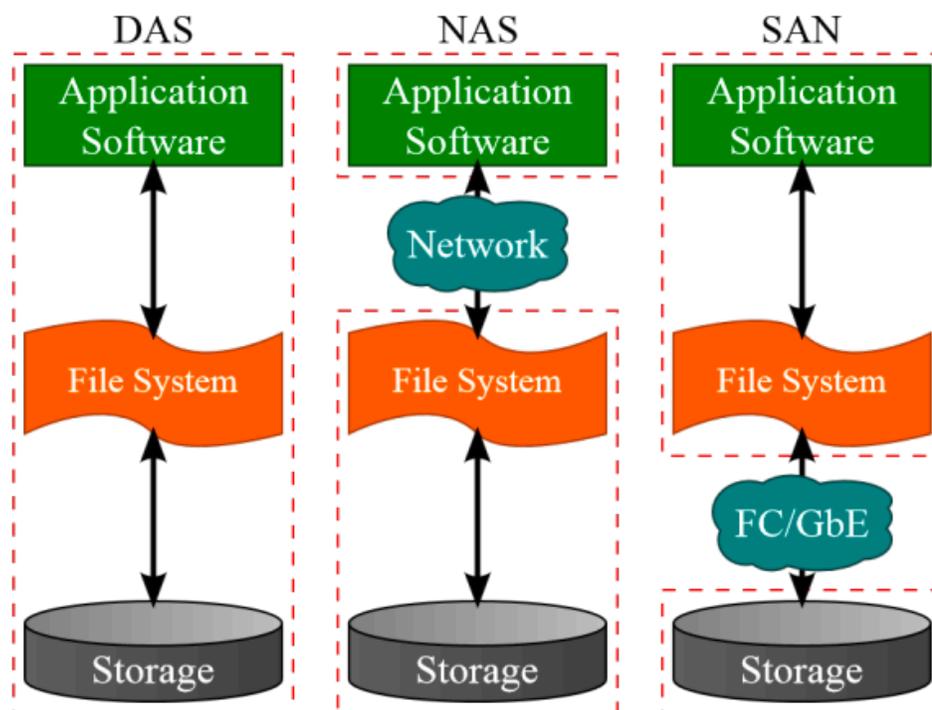
1. **可访问**：所有联网设备都可以访问 NAS。
2. **高性能**：NAS 专用于提供文件服务，其他联网设备无需再提供文件服务。
3. **容错性**：可以对 NAS 进行格式化，以支持复制磁盘、独立磁盘冗余阵列或纠删码，进而确保数据的完整性。
4. **易于设置**：NAS 架构可通过简化脚本，或以简化 OS 预装设备形式交付，大幅缩短存储设置和系统管理的时间。
5. **横向扩展功能**：NAS 增加存储容量简单。不必升级或更换现有服务器，在不中断网络下启用新存储。

- **NAS 缺点**

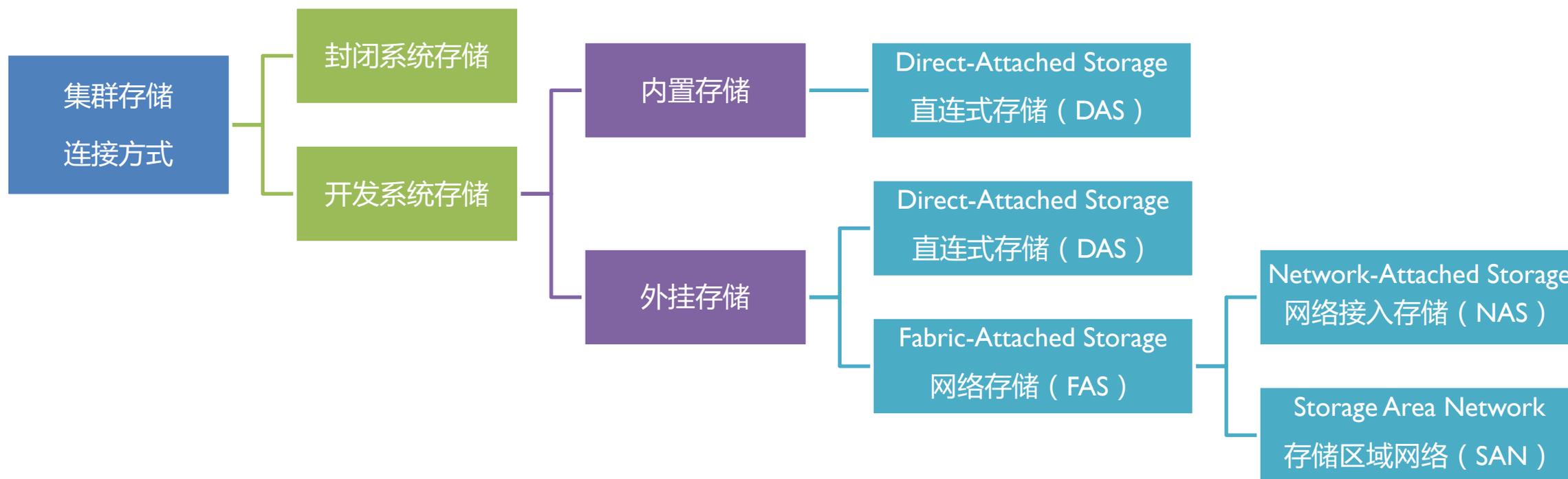
- 比如前期安装和设备成本比较高，可扩展性受到设备大小限制，存储性能有局限性
- NAS受限企业网络带宽，可能出现当多台客户端访问 NAS 导致性能下降，最终不能满足用户需求等。

DAS、SAN、NAS 关系

- DAS 存储一般应用在中小企业，与计算机采用直连方式，NAS 存储则通过以太网添加到计算机上，SAN 存储则使用 FC 接口，提供性能更加的存储。NAS 与 NAS 主要区别体现在操作系统所在位置。



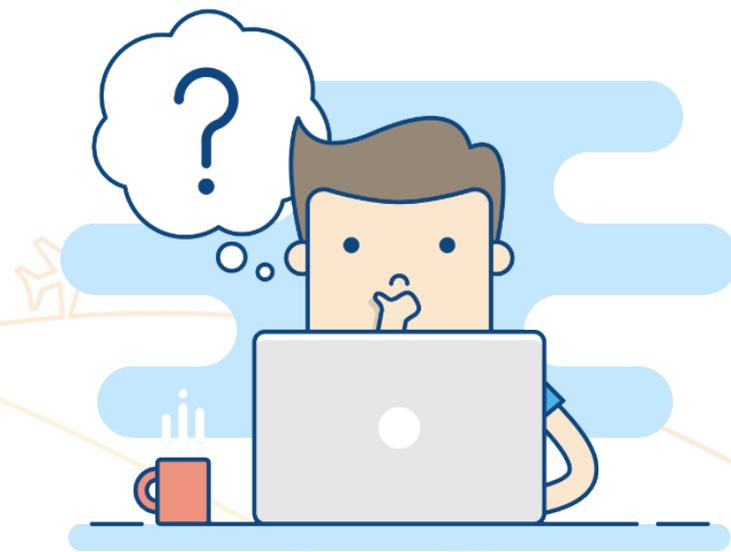
连接方式分类



AI 集群的服务器存储方式 I

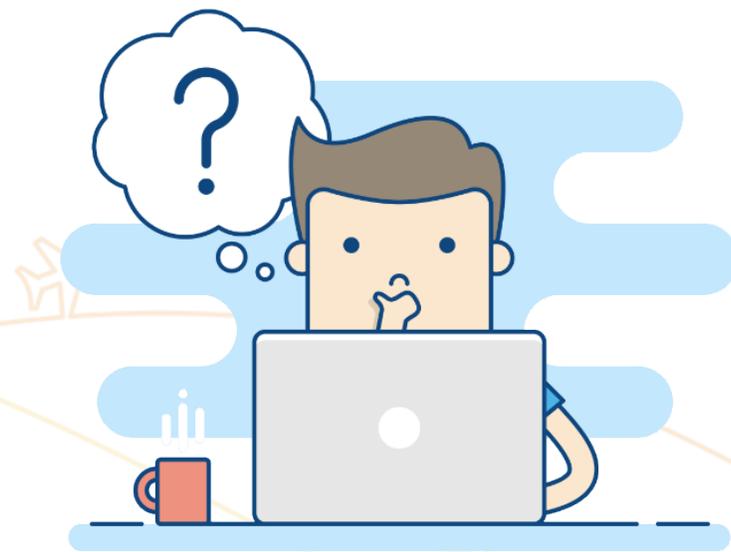
I. 实现统一存储系统？

- 前端主机接口可支持FC 8Gb、iSCSI 1Gb和iSCSI 10Gb，后端具备SAS 6Gb硬盘扩展接口，可支持SAS、SATA硬盘及SSD固态硬盘具备极佳的扩展能力。实现FC SAN与IP SAN、各类存储介质的完美融合。
- 有效整合用户现有存储网络架构，实现高性能SAN网络统一部署和集中管理，以适应业务和应用变化动态需求。主机接口及硬盘接口均采用模块化设计，更换主机接口或硬盘扩展接口，无须更换固件，简化升级维护的难度和工作量。



AI 集群的服务器存储方式 II

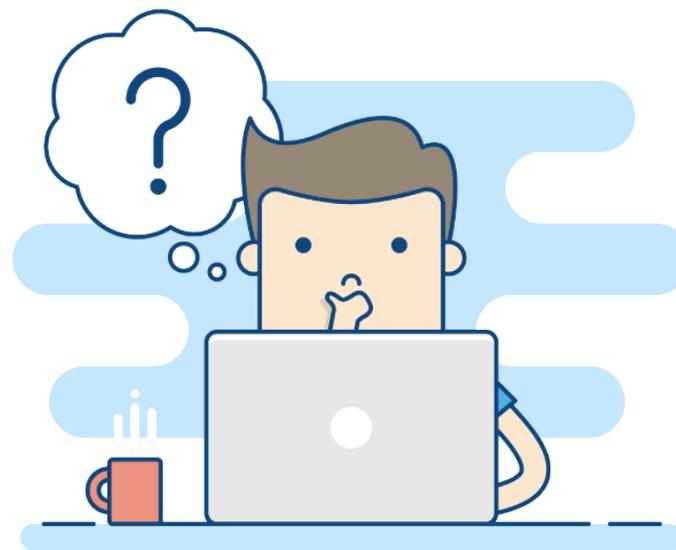
- I. AI 业务会产生 and/or 利用更多非结构化数据，这对 NAS 和 SAN 都是挑战，NAS+SAN 或许将是未来主要的存储解决方案？
 - 作为目前比较热门的统一存储方案。其既然提供集中化磁盘阵列，那么就支持主机系统通过 IP 网络进行文件级数据访问，或光纤协议在 SAN 网络进行块级别数据访问。
 - iSCSI 作为通用 IP 协议，只提供块级别数据访问。磁盘阵列配置多端口的存储控制器和一个管理接口，允许存储管理员按需创建存储池或空间，并将其提供给不同访问类型的主机系统。



2. 小结&思考

小结

1. 了解数据存储的硬件介质和对应的分类级别（RAM、FLASH）
2. 了解数据存储的集群主要提供的连接方式（DAS、NAS、SAN）



Reference 引用&参考

1. <https://www.linkedin.com/pulse/dram-scaling-challenges-grow-aken-cheung>
2. https://nn.wikipedia.org/wiki/Random_Access_Memory
3. <https://www.youtube.com/watch?v=OwtV0NrTLHA>
4. <https://www.agrade.com.cn/IndustryNews/683.html>
5. <https://www.cnblogs.com/FireLife-Cheng/p/17183491.html>
6. <https://www.wpgdadatong.com/blog/detail/43099>
7. https://www.sohu.com/a/476456430_121149658
8. <https://www.programmersought.com/article/66421832717/>
9. <https://blog.router-switch.com/2021/03/storage-architecture-nas-vs-san-vs-das/>
10. <https://www.linkedin.com/pulse/dram-scaling-challenges-grow-aken-cheung>
11. <https://36kr.com/p/1933583910734464>
12. <https://lvxixiao.gitbook.io/blog/os/store/hardware>
13. <https://www.redhat.com/zh/topics/data-storage/network-attached-storage>
14. <https://community.fs.com/cn/case-study/the-common-types-of-storage.html>
15. <https://help.aliyun.com/zh/ecs/user-guide/nvme-protocol>
16. <https://blog.csdn.net/fgf00/article/details/52592651>





Thank you

把AI系统带入每个开发者、每个家庭、
每个组织，构建万物互联的智能世界

Bring AI System to every person, home and
organization for a fully connected,
intelligent world.

Copyright © 2023 XXX Technologies Co., Ltd.
All Rights Reserved.

The information in this document may contain predictive statements including, without limitation, statements regarding the future financial and operating results, future product portfolio, new technology, etc. There are a number of factors that could cause actual results and developments to differ materially from those expressed or implied in the predictive statements. Therefore, such information is provided for reference purpose only and constitutes neither an offer nor an acceptance. XXX may change the information at any time without notice.

 ZOMI

Course [chenzomi12.github.io](https://github.com/chenzomi12)

GitHub github.com/chenzomi12/DeepLearningSystem