

(Part 8)

# 基于对象的存储和 统一存储

# 模块 8：基于对象的存储和统一存储

学完本模块后，您将能够：

- 介绍基于对象的存储模型
- 列出基于对象的存储的关键组件
- 介绍基于对象的存储中的存储和检索过程
- 介绍内容寻址存储
- 列出统一存储的关键组件
- 介绍统一存储中的数据访问过程

# 模块 8：基于对象的存储和统一存储

## 第 1 课：基于对象的存储

本课程将讲述下列主题：

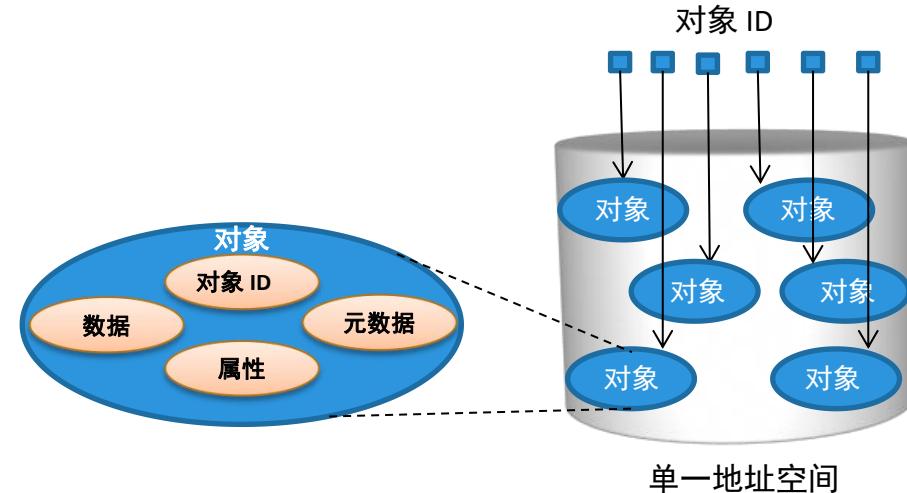
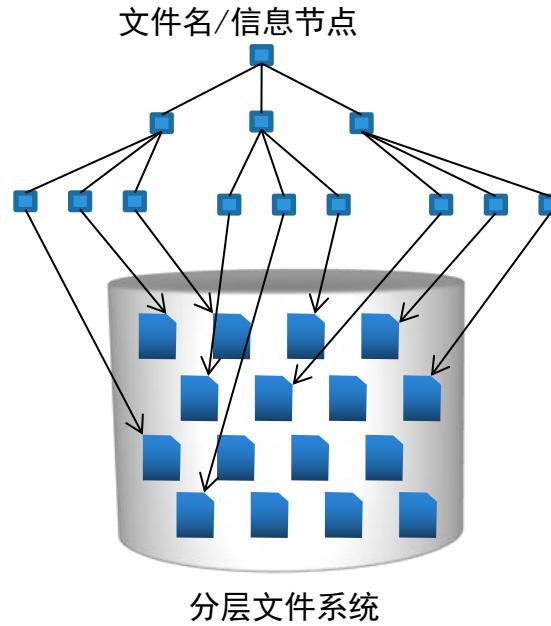
- 分层文件系统与单一地址空间的对比
- 基于对象的存储模型
- 基于对象的存储的关键组件
- 基于对象的存储设备中的存储和检索过程
- 内容寻址存储

# 基于对象的存储的驱动因素

- 超过 90% 正在生成的数据是非结构化数据
- 传统解决方案应对增长效率低下
  - ▶ 由于管理大量权限和嵌套的目录，NAS 开销较大
- 这些挑战需要一种更智能的方法来根据内容管理非结构化数据

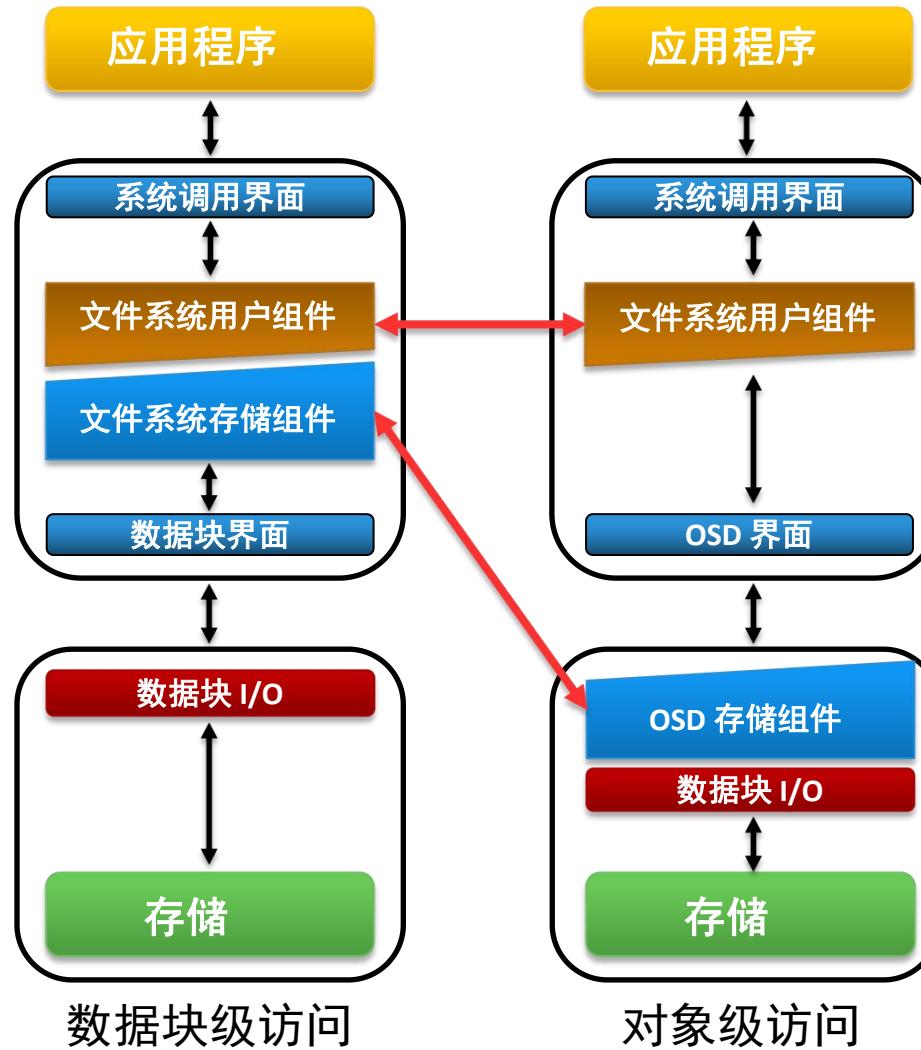
基于对象的存储是一种在单一地址空间上根据文件数据的内容和属性（而不是名称和位置）以对象的形式存储这些数据的方法

# 分层文件系统与单一地址空间

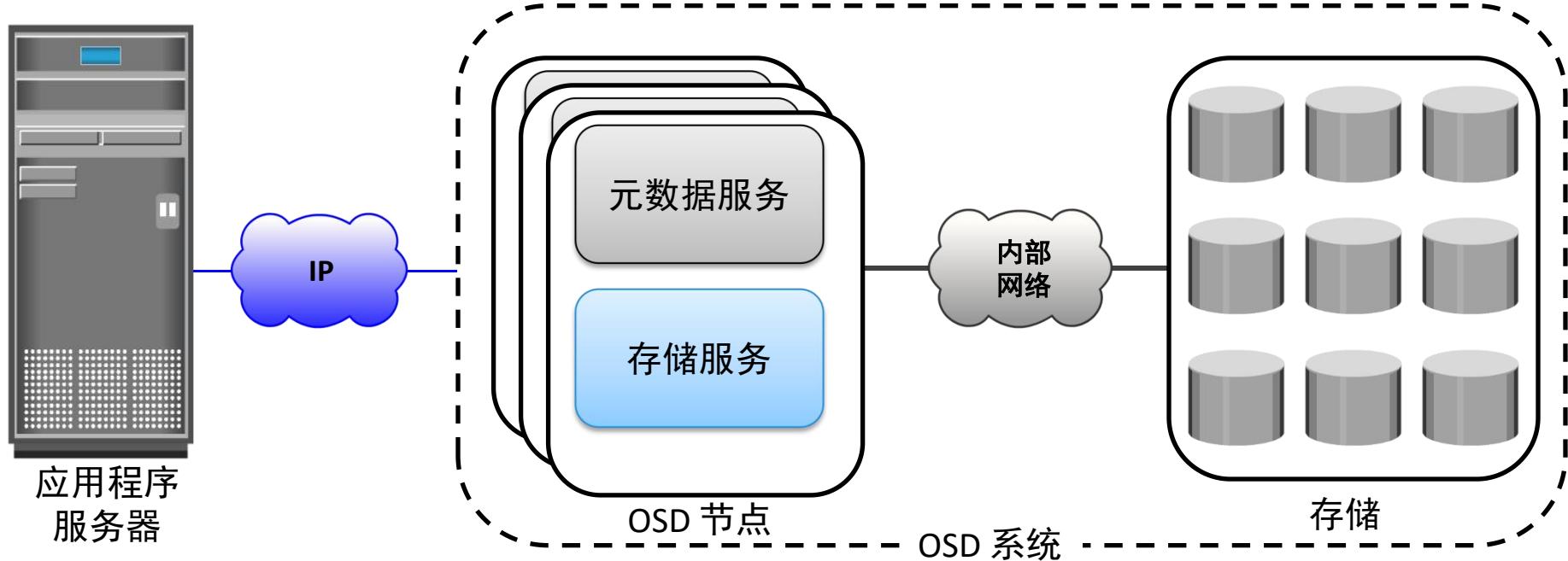


- 分层文件系统以文件和目录的形式组织数据
- 基于对象的存储设备以对象的形式存储数据
  - ▶ 它使用支持存储大量对象的单一地址空间
  - ▶ 对象包含用户数据、相关的元数据和其他属性
  - ▶ 每个对象都具有使用专用算法生成的唯一对象 ID

# 传统与基于对象的存储模型

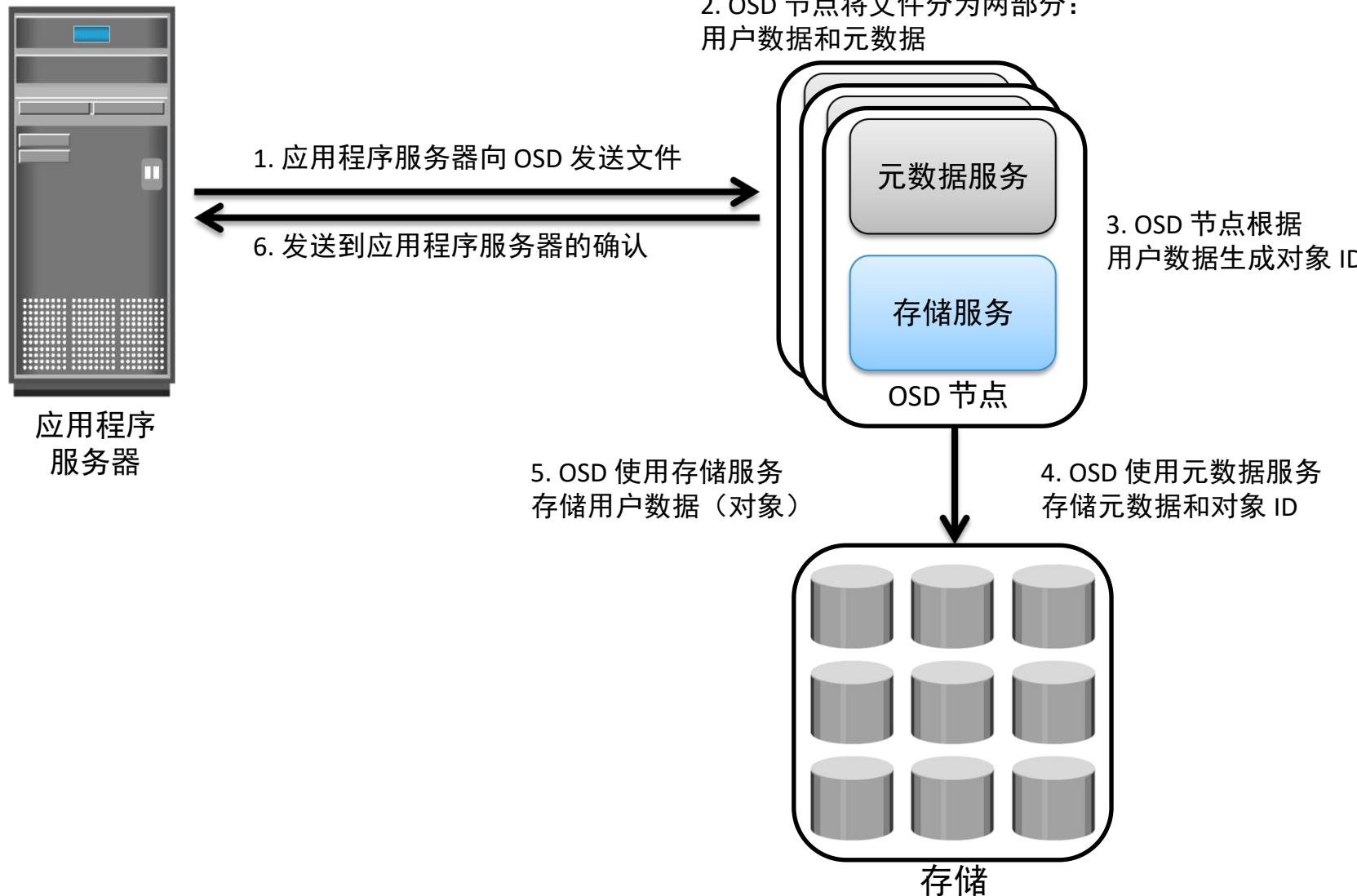


# 基于对象的存储设备的关键组件



- OSD 系统通常包含三个关键组件：
  - ▶ OSD 节点
  - ▶ 内部网络
  - ▶ 存储

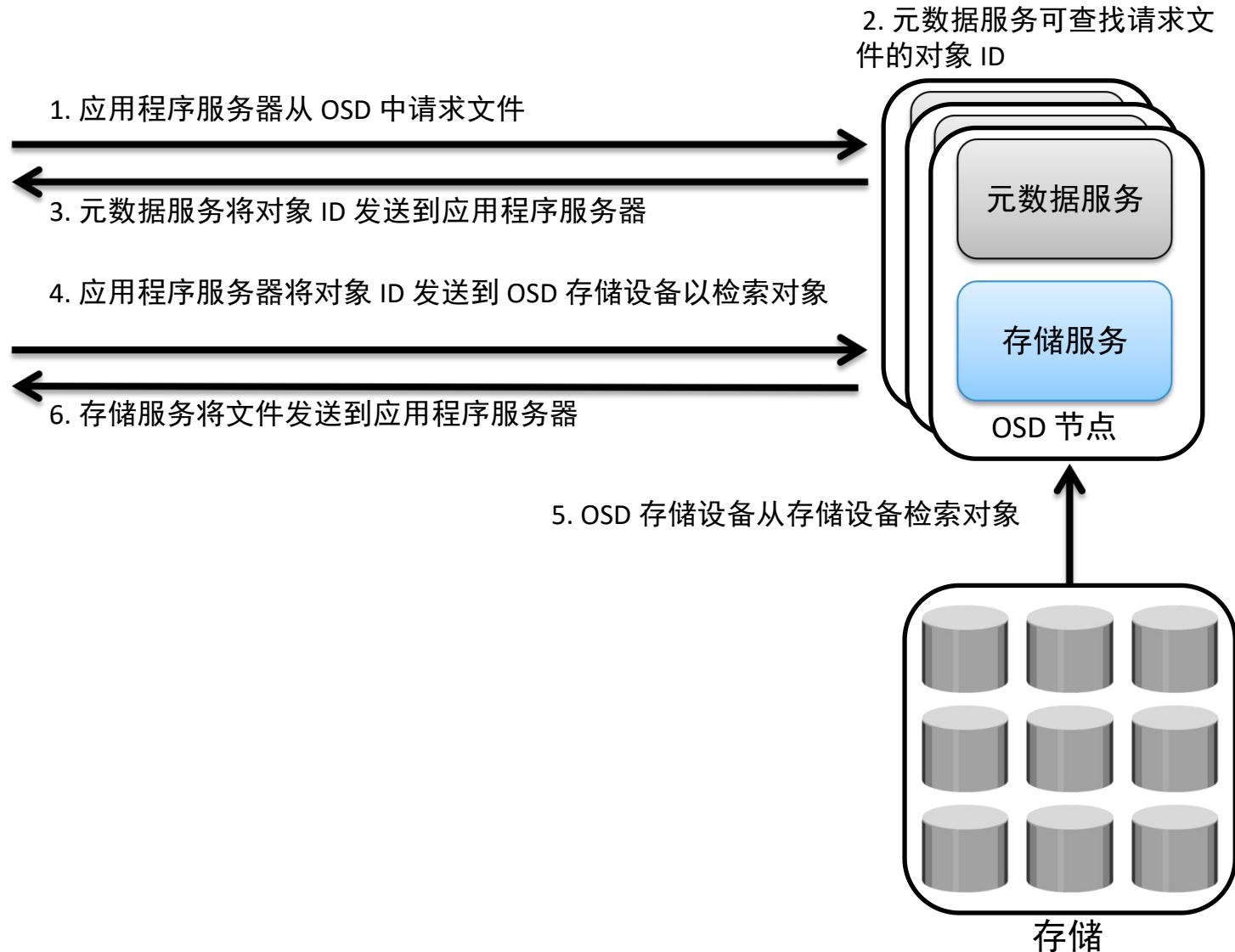
# 在 OSD 中存储对象的过程



# 从 OSD 中检索对象的过程



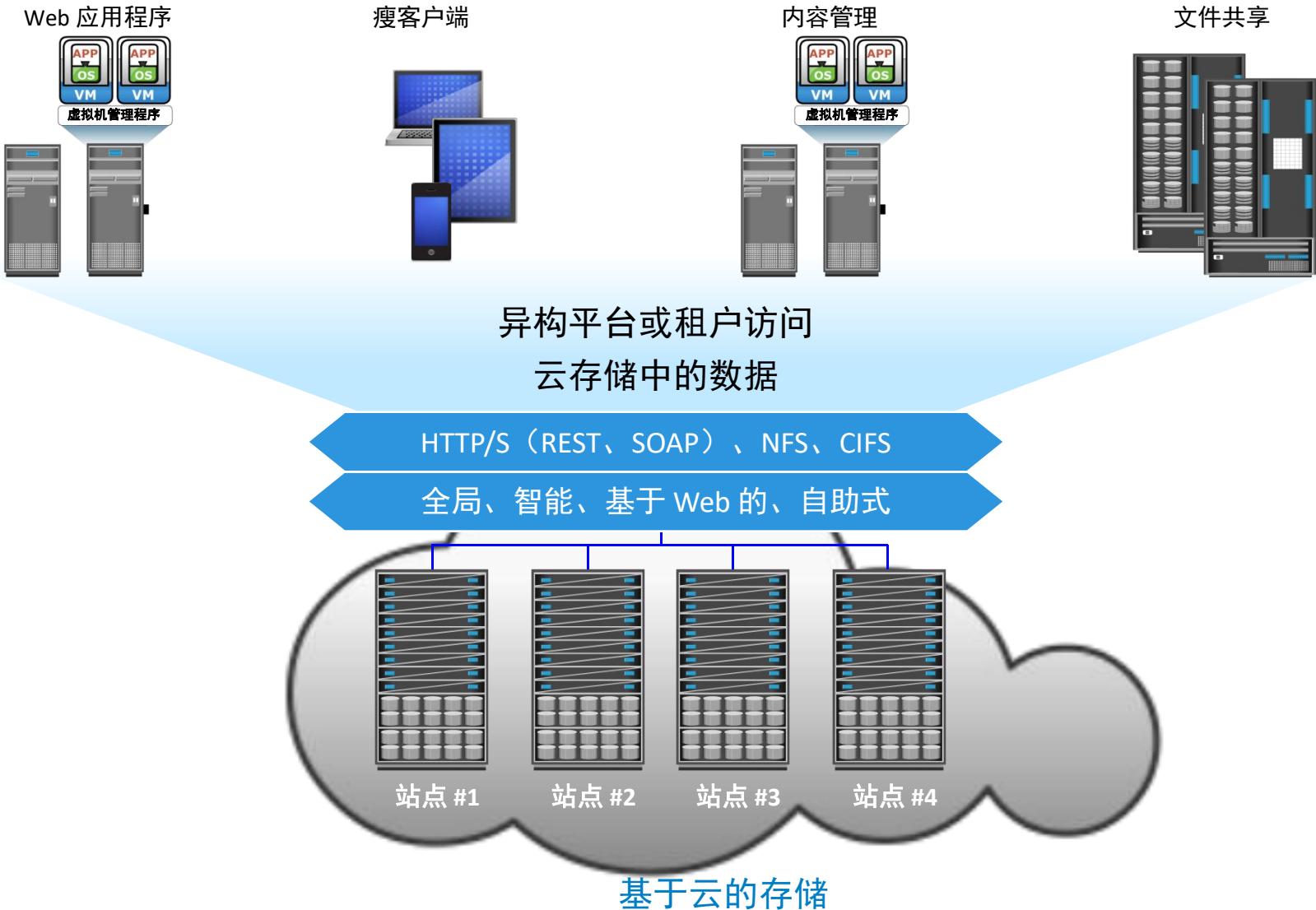
应用程序  
服务器



# 基于对象的存储的主要优势

优势	描述
安全性和可靠性	<ul style="list-style-type: none"><li>通过专用算法生成的唯一对象 ID 可确保数据完整性和内容真实性</li><li>请求身份验证在存储设备中执行</li></ul>
平台独立性	<ul style="list-style-type: none"><li>因为对象是数据的抽象容器，所以它支持跨异构平台共享对象</li><li>此功能使基于对象的存储适用于云计算环境</li></ul>
可扩展性	<ul style="list-style-type: none"><li>OSD 节点和存储均可独立地进行扩展</li></ul>
可管理性	<ul style="list-style-type: none"><li>具有管理对象的固有智能</li><li>具有自我修复功能</li><li>基于策略的管理功能使 OSD 能够自动处理日常作业</li></ul>

# 使用情形 1：基于云的存储



## 使用情形 2：内容寻址存储 (CAS)

- 设计用于存储固定内容的存储
- 将数据作为对象存储
- 会为每个对象分配一个全局唯一标识符，称为内容地址 (CA)
  - ▶ CA 由数据的二进制表示形式派生而来
- CAS 设备可通过应用程序服务器上运行的 CAS API 进行访问

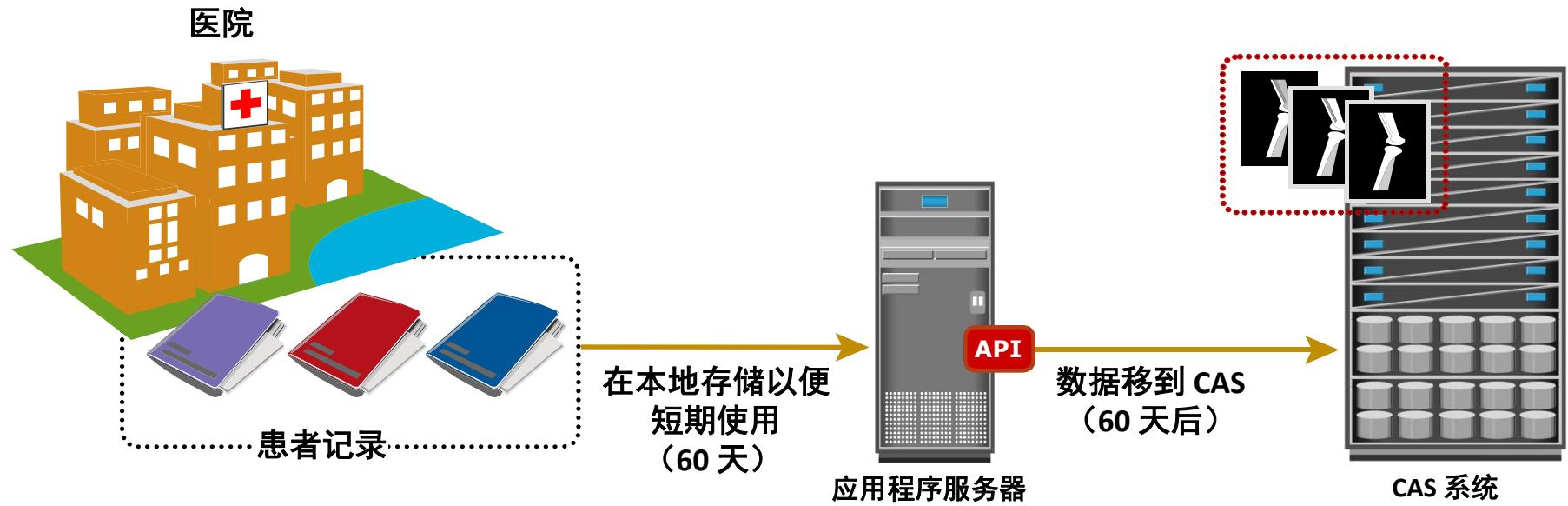
# CAS 的主要功能

- 内容真实性和完整性
- 位置独立性
- 单实例存储
- 强制执行保留
- 数据保护
- 快速检索记录
- 负载平衡
- 可扩展性
- 自我诊断和修复
- 审核追踪和事件通知

# CAS 的主要功能

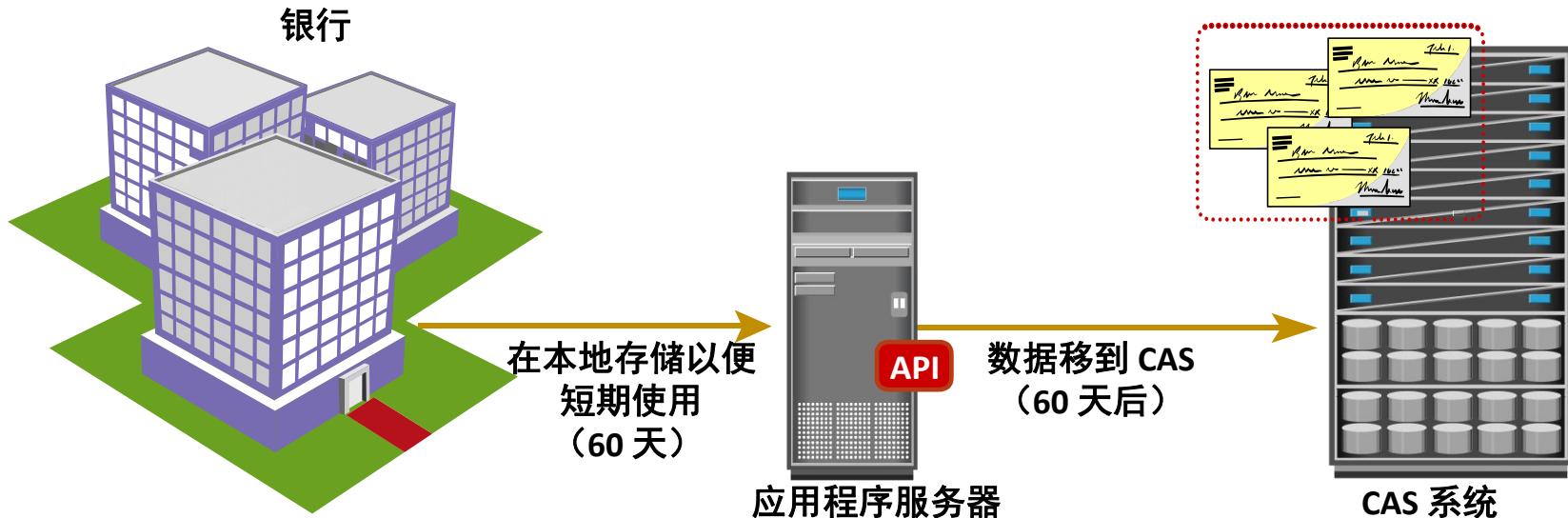
- 内容真实性和完整性
- 位置独立性
- 单实例存储
- 强制执行保留
- 数据保护
- 快速检索记录
- 负载平衡
- 可扩展性
- 自我诊断和修复
- 审核追踪和事件通知

# 使用情形 1：医疗保健解决方案



- 每个 X 射线图像大小大约 15 MB 到 1 GB 以上
- 患者记录在线存储 60 天
- 超过 60 天的患者记录归档到 CAS

## 使用情形 2：金融解决方案



- 每个支票图像大小大约 25 KB
- 支票成像服务提供商每月可能处理大约 9 千万个支票图像
- 支票在线存储 60 天
- 超过 60 天的数据归档到 CAS

# 模块 8：基于对象的存储和统一存储

## 第 2 课：统一存储

本课程将讲述下列主题：

- 统一存储的关键组件
- 统一存储中的数据访问

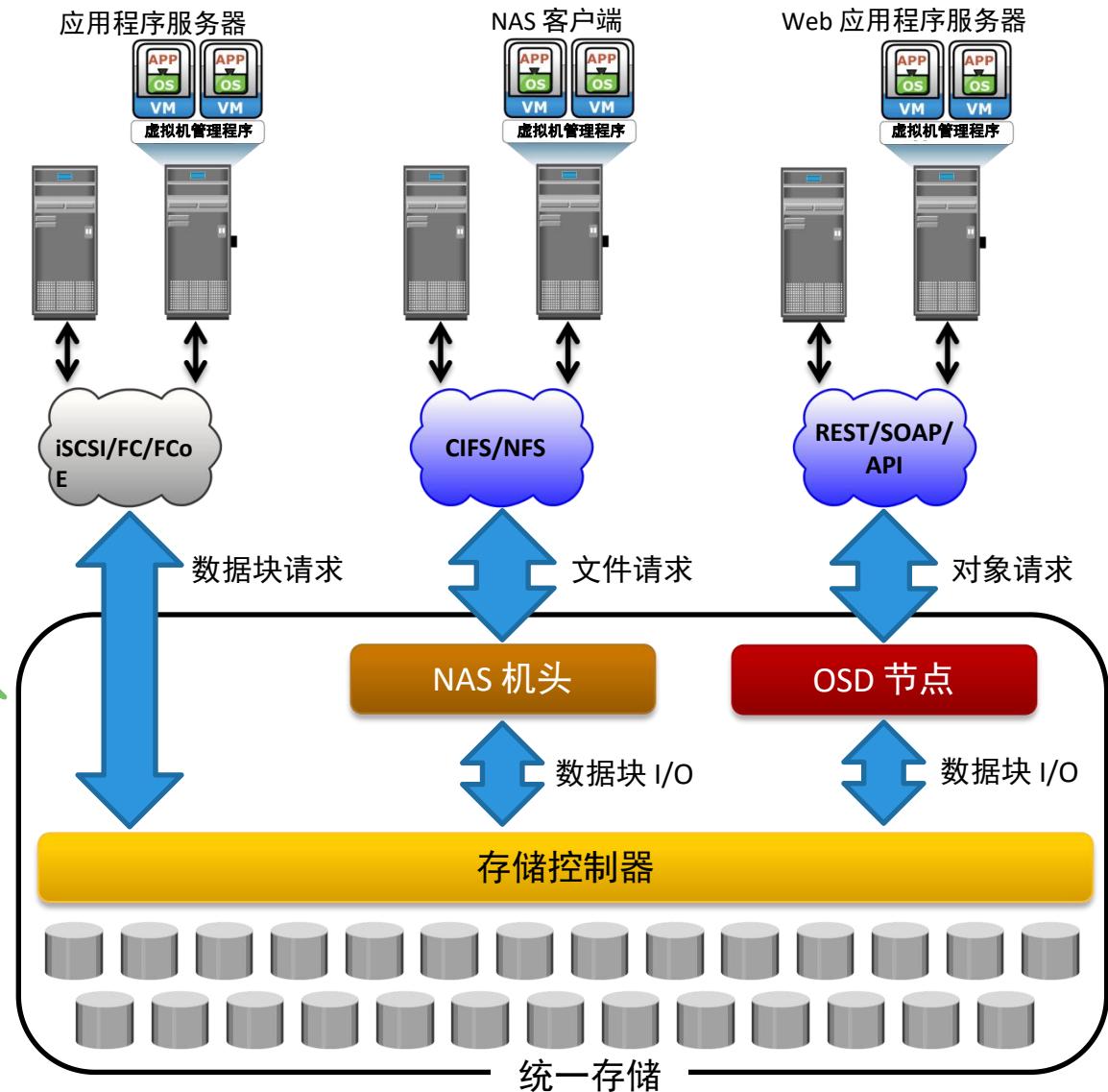
# 统一存储的驱动因素

- 部署不同的存储解决方案（SAN、NAS 和 OSD）会增加管理成本、复杂性和环境开销
- 统一存储可在一个统一平台中整合数据块、文件和基于对象的访问
  - ▶ 支持数据访问的多个协议
  - ▶ 可以通过单个管理界面进行管理

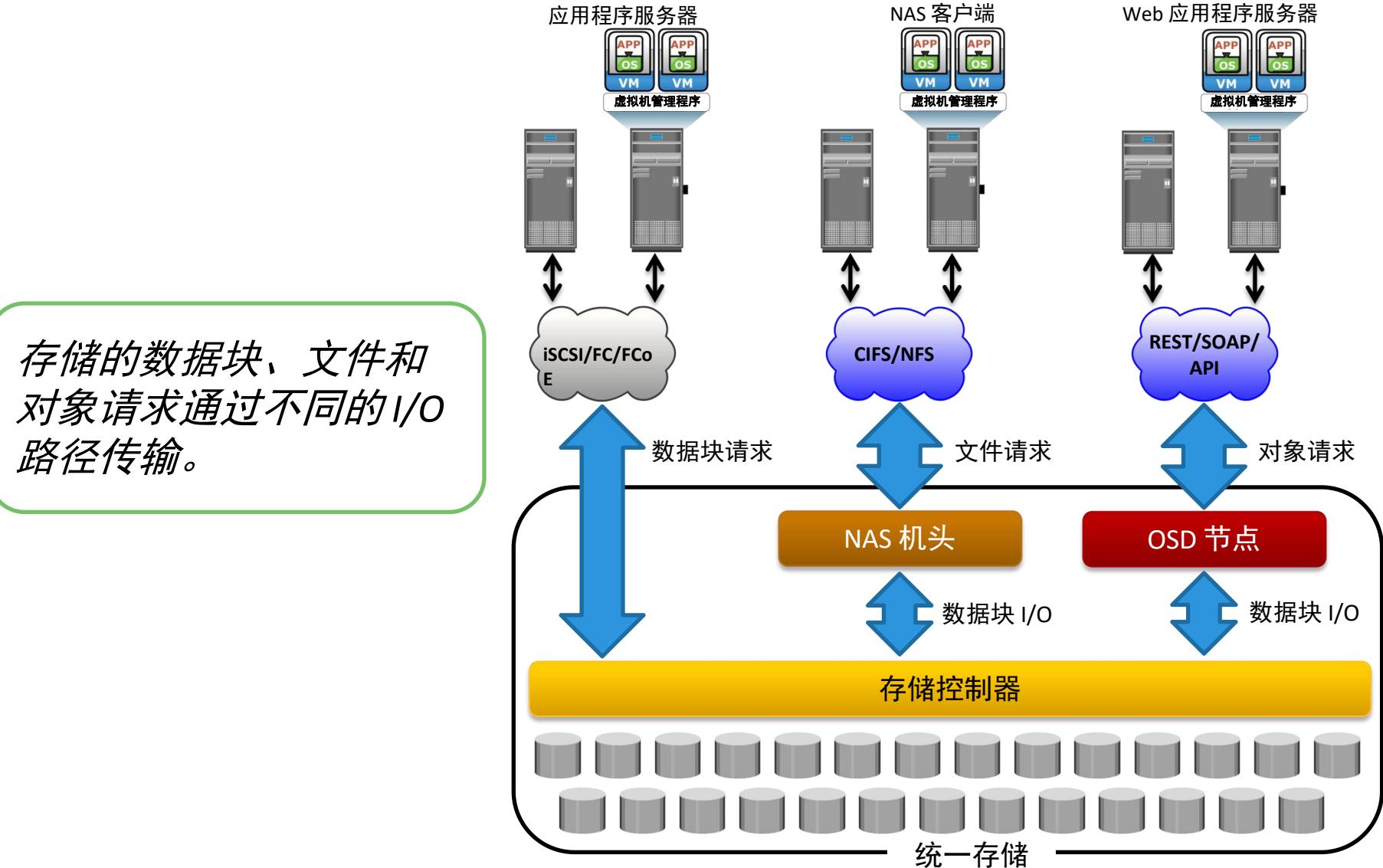
# 统一存储的组件

统一存储的组件有：

- 存储控制器
- NAS 机头
- OSD 节点
- 存储



# 统一存储中的数据访问



# 模块 8：基于对象的存储和统一存储

## 付诸实践的概念

- EMC Atmos
- EMC VNX
- EMC Centera

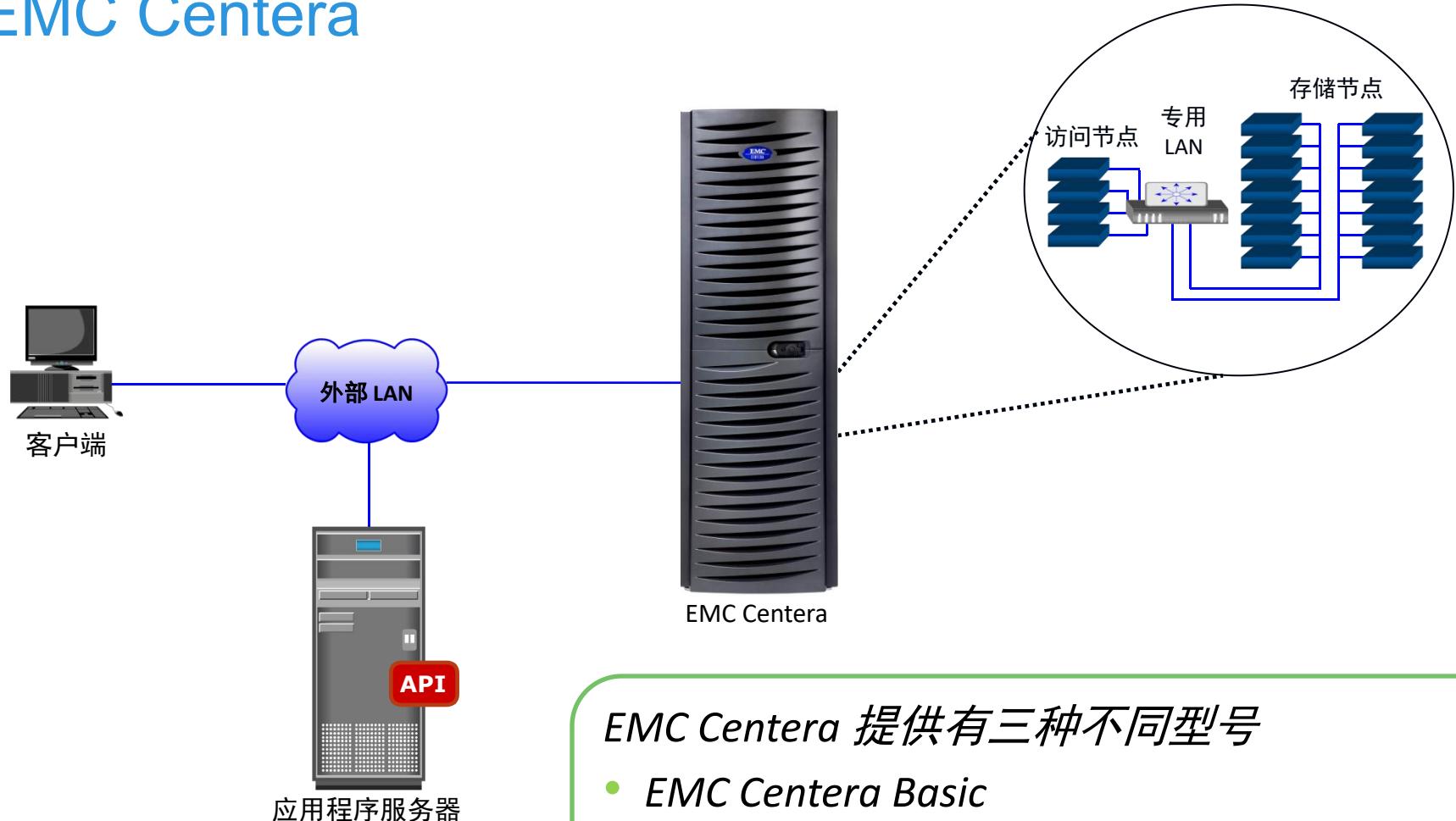
# EMC Atmos

- 具备巨大扩展能力的基于对象的存储
- 可以按两种方式部署：专门构建的硬件应用装置或虚拟机(VM)
- 主要功能
  - ▶ 支持基于策略的管理
  - ▶ 提供复制和奇偶校验保护
  - ▶ 提供诸如压缩和重复数据消除之类的服务
  - ▶ 支持 Web 服务和旧版协议
    - ▶ REST、SOAP、CIFS、NFS 和可安装式文件系统
  - ▶ 支持自动化系统管理
  - ▶ 支持多租户功能
  - ▶ 灵活的管理

# EMC Atmos

- 具备巨大扩展能力的基于对象的存储
- 可以按两种方式部署：专门构建的硬件应用装置或虚拟机(VM)
- 主要功能
  - ▶ 支持基于策略的管理
  - ▶ 提供复制和奇偶校验保护
  - ▶ 提供诸如压缩和重复数据消除之类的服务
  - ▶ 支持 Web 服务和旧版协议
    - ▶ REST、SOAP、CIFS、NFS 和可安装式文件系统
  - ▶ 支持自动化系统管理
  - ▶ 支持多租户功能
  - ▶ 灵活的管理

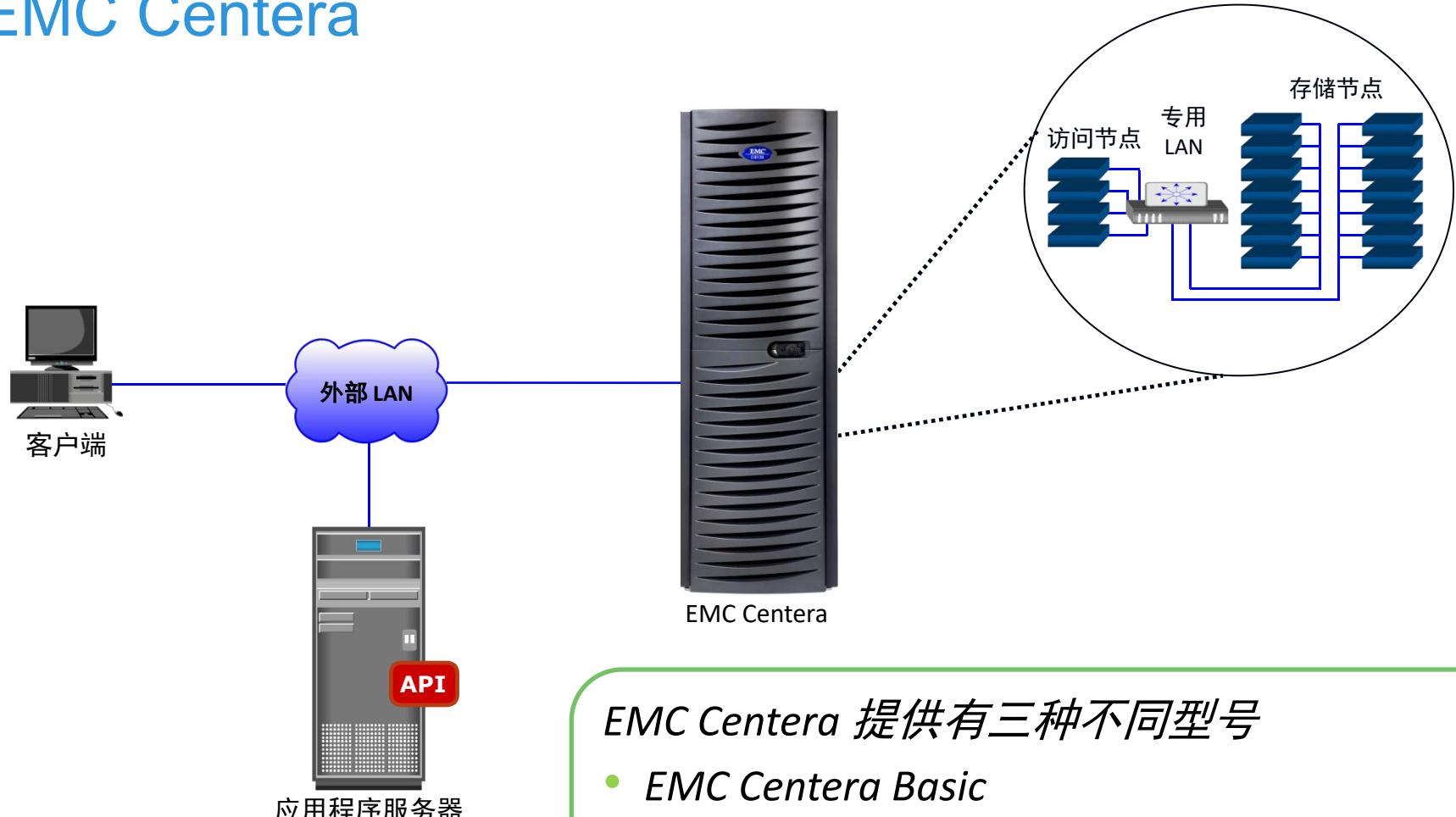
# EMC Centera



*EMC Centera 提供有三种不同型号*

- *EMC Centera Basic*
- *EMC Centera Governance Edition*
- *EMC Centera Compliance Edition Plus (CE+)*

# EMC Centera



EMC Centera 提供有三种不同型号

- EMC Centera Basic
- EMC Centera Governance Edition
- EMC Centera Compliance Edition Plus (CE+)

# EMC VNX

- 在一个解决方案中整合数据块、文件和对象访问的统一存储平台
  - ▶ 支持通过存储处理器的数据块访问
  - ▶ 通过 X-Blade 的文件访问
  - ▶ 通过 EMC Atmos VE 的对象访问
- VNX 的组件有：
  - ▶ 存储处理器
  - ▶ X-Blade
  - ▶ 控制站
  - ▶ 磁盘阵列存储模块
  - ▶ 备用电源



# 模块 8：总结

本模块涵盖以下要点：

- 基于对象的存储模型
- 基于对象的存储的关键组件
- 基于对象的存储的存储和检索过程
- 内容寻址存储
- 统一存储的关键组件
- 统一存储中的数据访问过程

# 知识测验 – 1

- 单一地址空间较层次结构地址空间而言有何优势？
  - A. 高度可扩展且对性能影响最小 ❤
  - B. 根据保留策略提供对数据的访问
  - C. 使用同一接口提供对数据块、文件和对象的访问
  - D. 访问数据时使用的网络带宽更少
- OSD 节点中的元数据服务有何作用？
  - A. 负责以对象的形式存储数据
  - B. 存储为对象生成的唯一 ID ❤
  - C. 存储对象和对象 ID
  - D. 控制存储设备的功能

## 知识测验 – 2

- 使用下面哪一项在 CAS 系统中生成对象 ID？
  - A. 文件元数据
  - B. 源和目标地址
  - C. 数据的二进制表示形式 
  - D. 文件系统类型和所有权
- 下面哪一项准确描述了统一存储中的数据块 I/O 访问？
  - A. I/O 穿过 NAS 机头和存储控制器到达磁盘
  - B. I/O 穿过 OSD 节点和存储控制器到达磁盘
  - C. I/O 穿过存储控制器到达磁盘 
  - D. I/O 会直接发送到磁盘

## 知识测验 – 3

- 下面哪一项准确描述了统一存储?
  - A. 在一个平台中提供数据块、文件和基于对象的访问 
  - B. 使用对象提供数据块和文件存储访问
  - C. 使用单一地址空间支持数据块和文件访问
  - D. 专门构建用于归档的专用存储设备