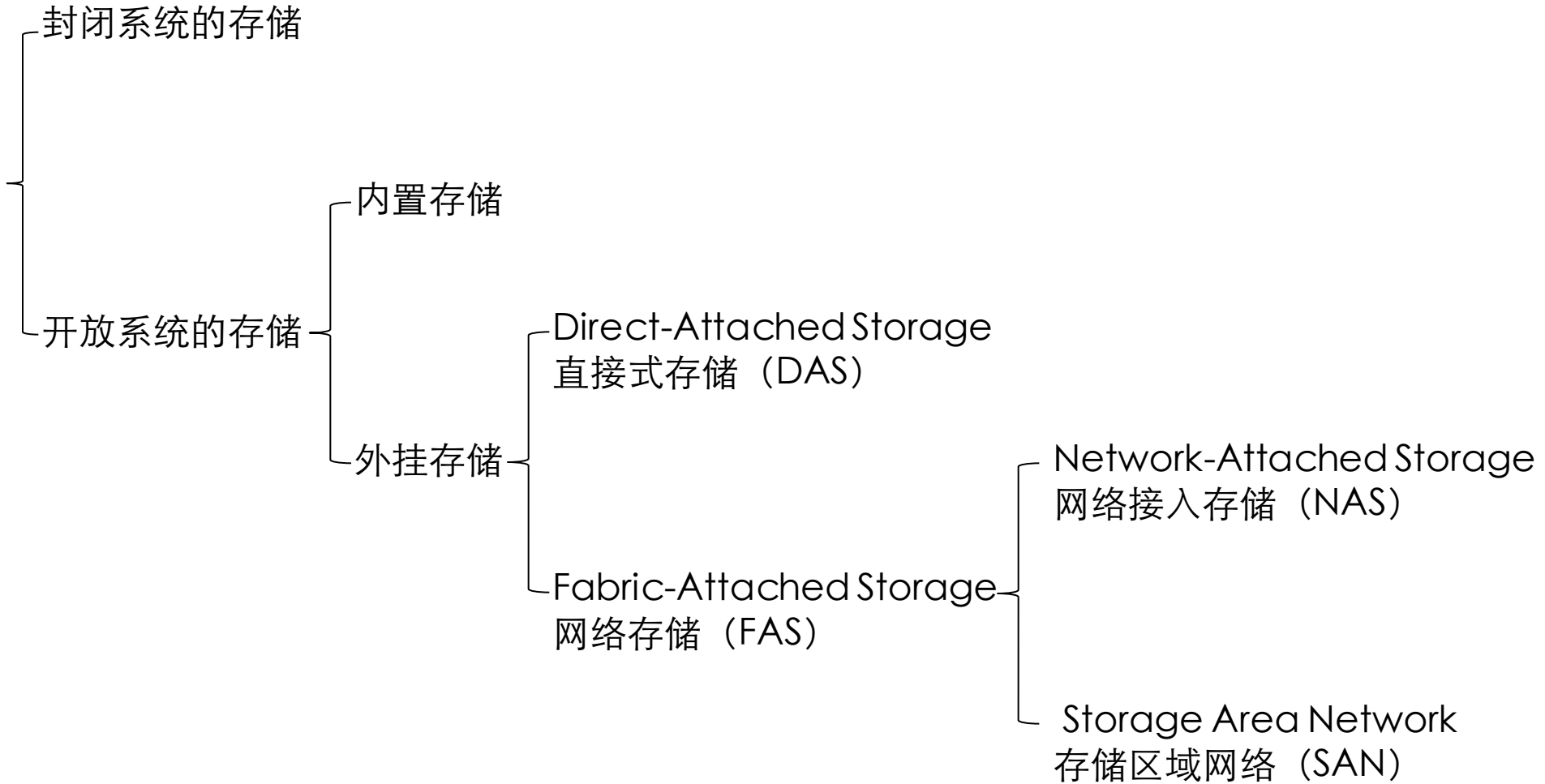


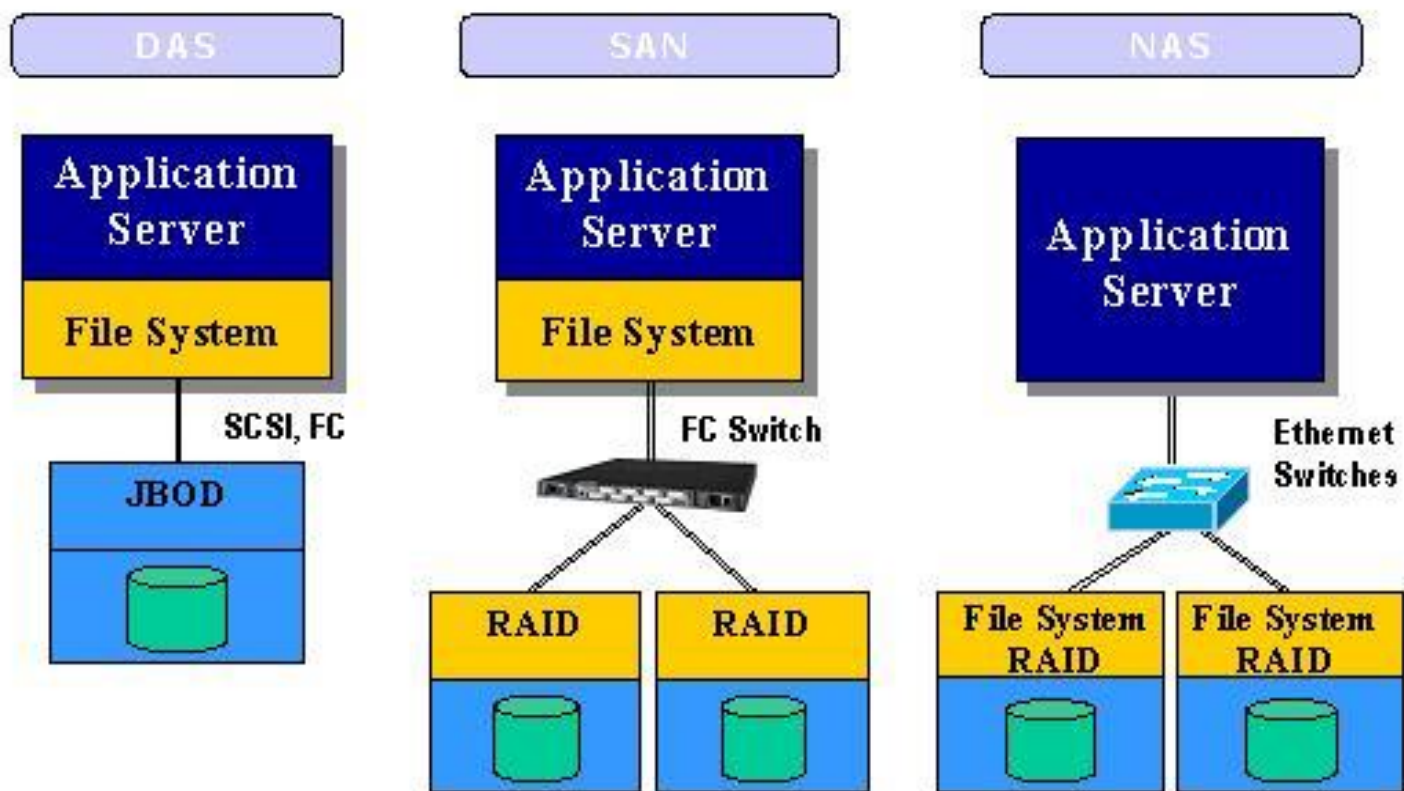
# 目 录

1. 存储组网形态
2. 主要协议
3. RAID
4. 磁盘热备与重构
5. 快照与复制
6. 数据分级存储概念

# 存储分类

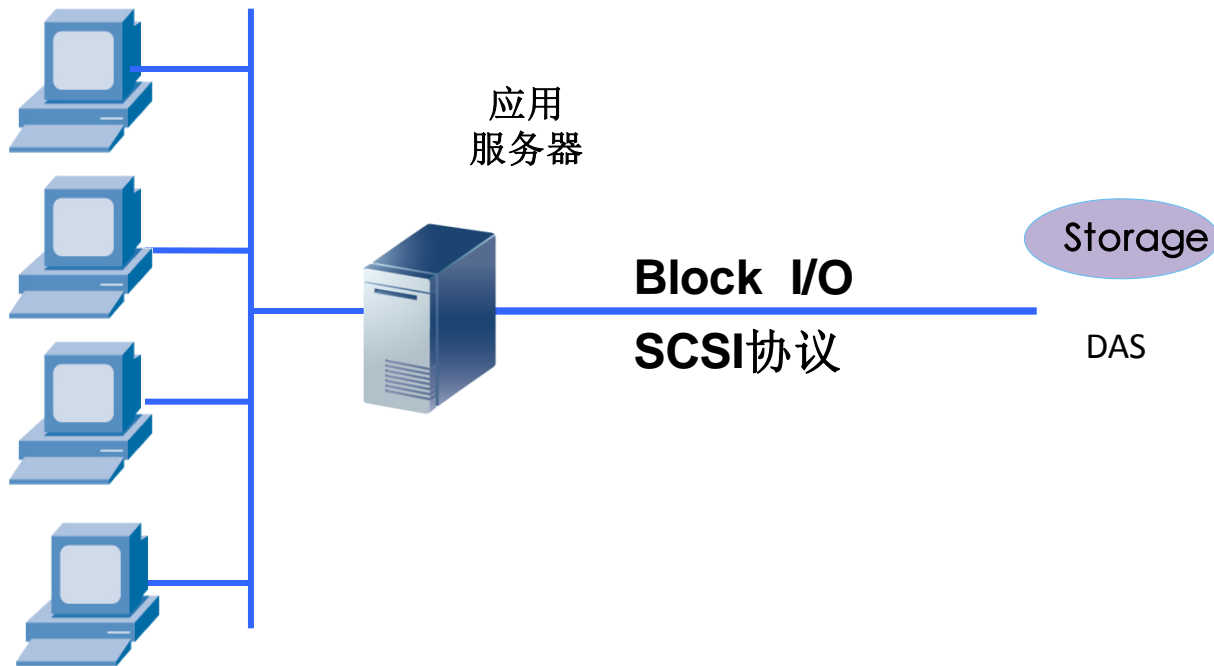


# DAS/SAN/NAS组网示意图

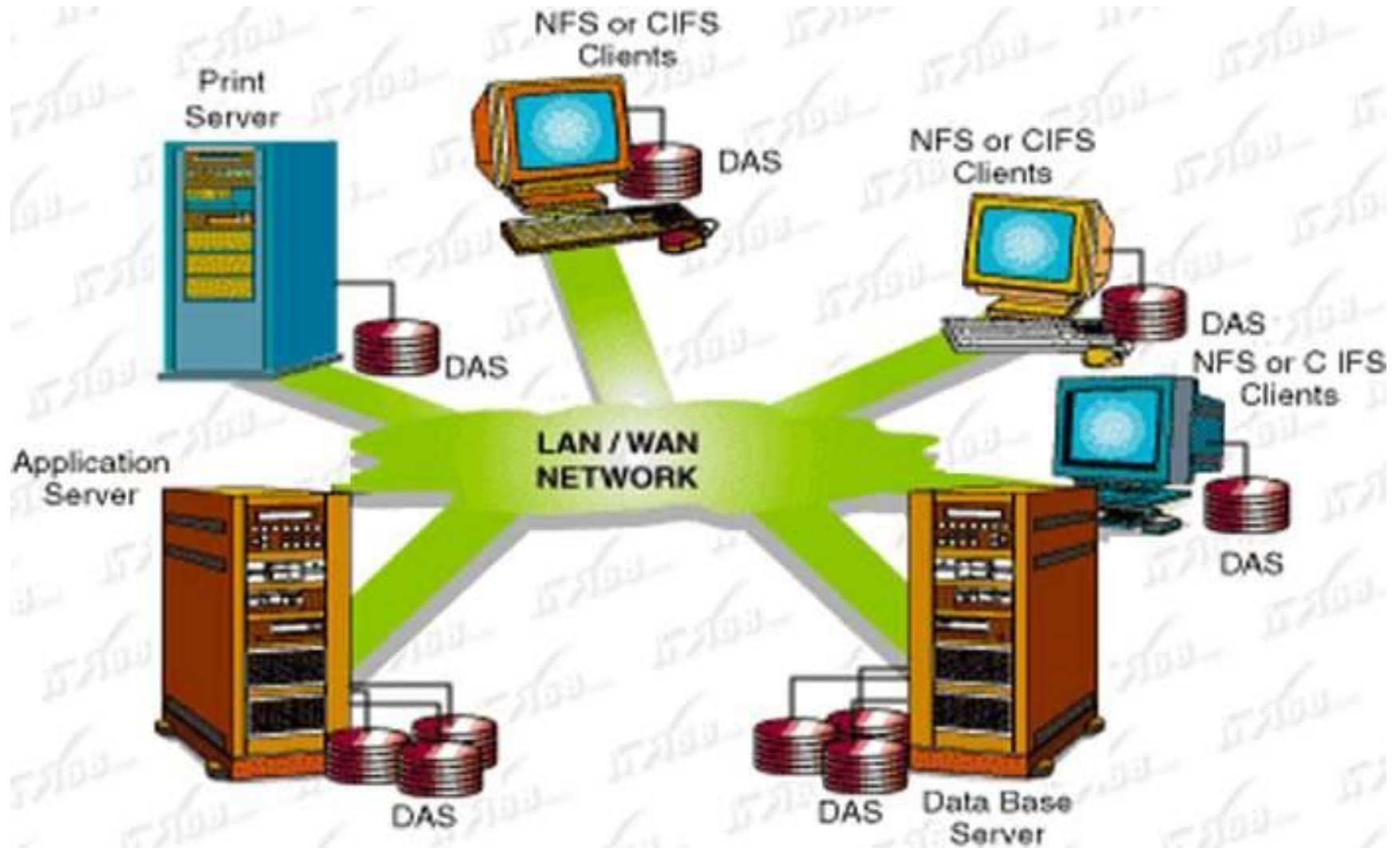


# DAS (Direct Attachment Storage)

**直接连接存储：** 存储设备是通过电缆（通常是SCSI接口电缆）直接连到服务器的。  
I/O请求直接发送到存储设备



# 典型DAS结构



# DAS分析

## DAS优势

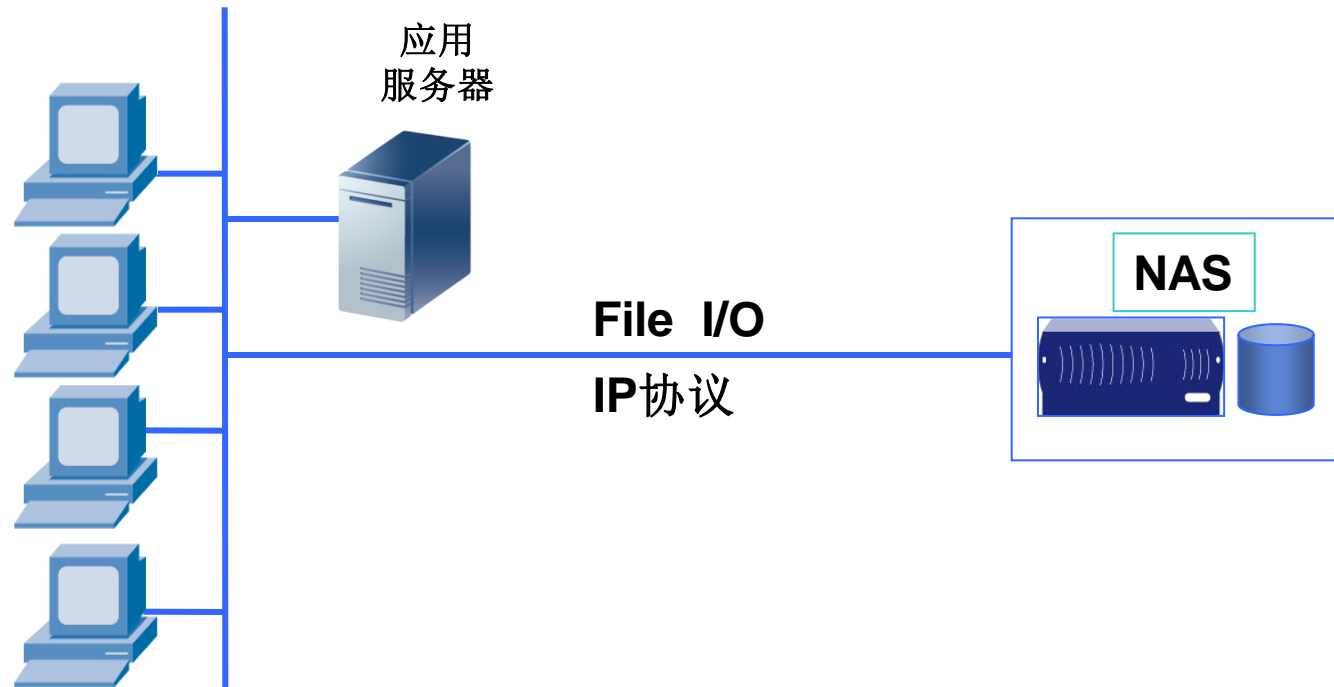
- 连接简单
  - 集成在服务器内部；点到点的连接；距离短；
  - 安装技术要求不高
- 低成本需求
  - SCSI总线成本低
- 较好的性能
- 通用的解决方案
  - DAS的投资低，绝大多数应用可以接受

## DAS劣势

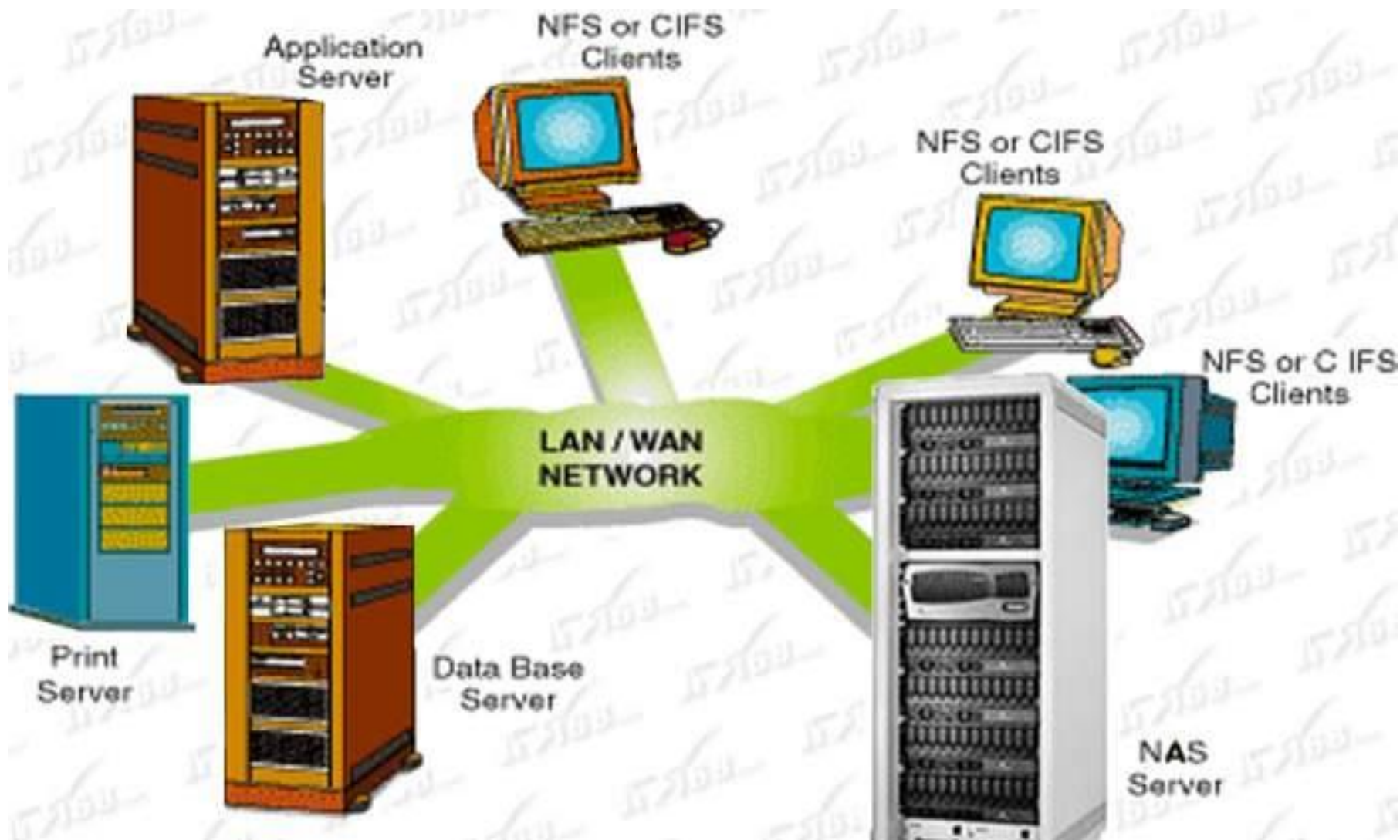
- 有限的扩展性
  - SCSI总线的距离最大25米；最多15个设备
- 专属的连接
  - 空间资源无法与其他服务器共享
- 备份和数据保护
  - 备份到与服务器直连的磁带设备上
  - 硬件失败将导致更高的恢复成本
- TCO（总拥有成本高）
  - 存储容量的加大导致管理成本上升
  - 存储使用效率低

# NAS（Network Attachment Storage）

网络连接存储：存储设备连接到现有的网络上，提供数据和文件服务



# 典型NAS结构





# NAS的优势

## NAS优势

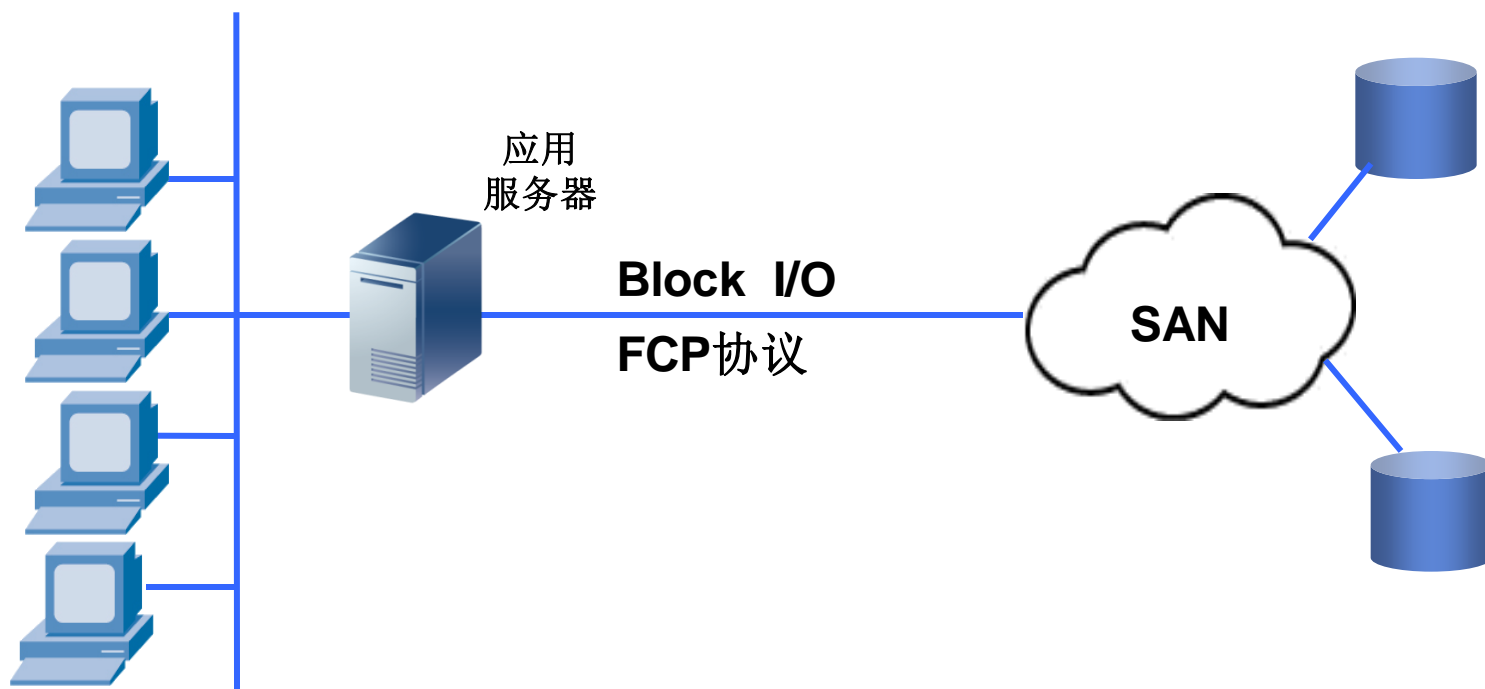
- 资源共享
- 构架于IP网络之上
- 部署简单
- 较好的扩展性
- 异构环境下的文件共享
- 易于管理
- 备份方案简单
- 低的TCO

## NAS劣势

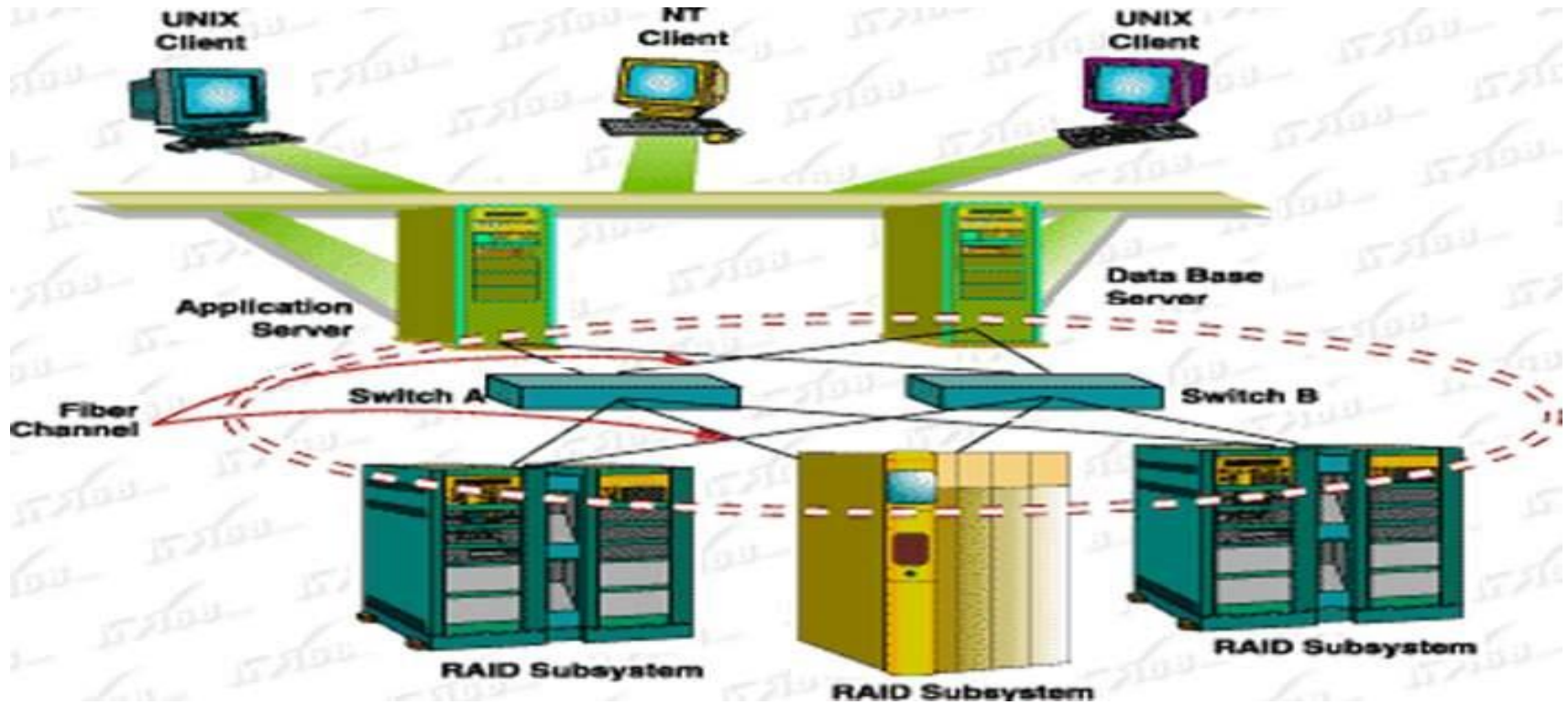
- 扩展性有限
- 一些应用会占用带宽资源
- 不适应某些数据库的应用

# SAN (Storage Area Network)

**存储区域网络：**它是一个用在服务器和存储资源之间的、专用的、高性能的网络体系。它为实现大量原始数据的传输而进行了专门的优化。



# 典型SAN结构



# SAN的优势

## SAN优势

- 实现存储介质的共享
- 非常好的扩展性
- 易于数据备份和恢复
  - 实现备份磁带共享
  - LAN Free和Server Free
- 高性能
- 支持服务器群集技术
- 容灾手段
- 低的TCO

## SAN劣势

- 成本较高
  - 需要专用的连接设备如FC交换机以及HBA
- SAN孤岛
- 技术较为复杂
  - 需要专业的技术人员维护

# DAS/NAS/SAN三种形态比较

	<i>DAS</i>	<i>NAS</i>	<i>FC-SAN</i>	<i>IP-SAN</i>
传输类型	SCSI、FC	IP	FC	IP
数据类型	块级	文件级	块级	块级
典型应用	任何	文件服务器	数据库应用	视频监控
优点	易于理解 兼容性好	易于安装 成本低	高扩展性 高性能 高可用性	高扩展性 成本低
缺点	难以管理，扩展性有限 存储空间利用率不高	性能较低 对某些应用不适合	比较昂贵，配置复杂 互操作性问题	性能较低



# 目 录

1. 存储组网形态
2. 主要协议
3. RAID
4. 磁盘热备与重构
5. 快照与复制
6. 数据分级存储概念

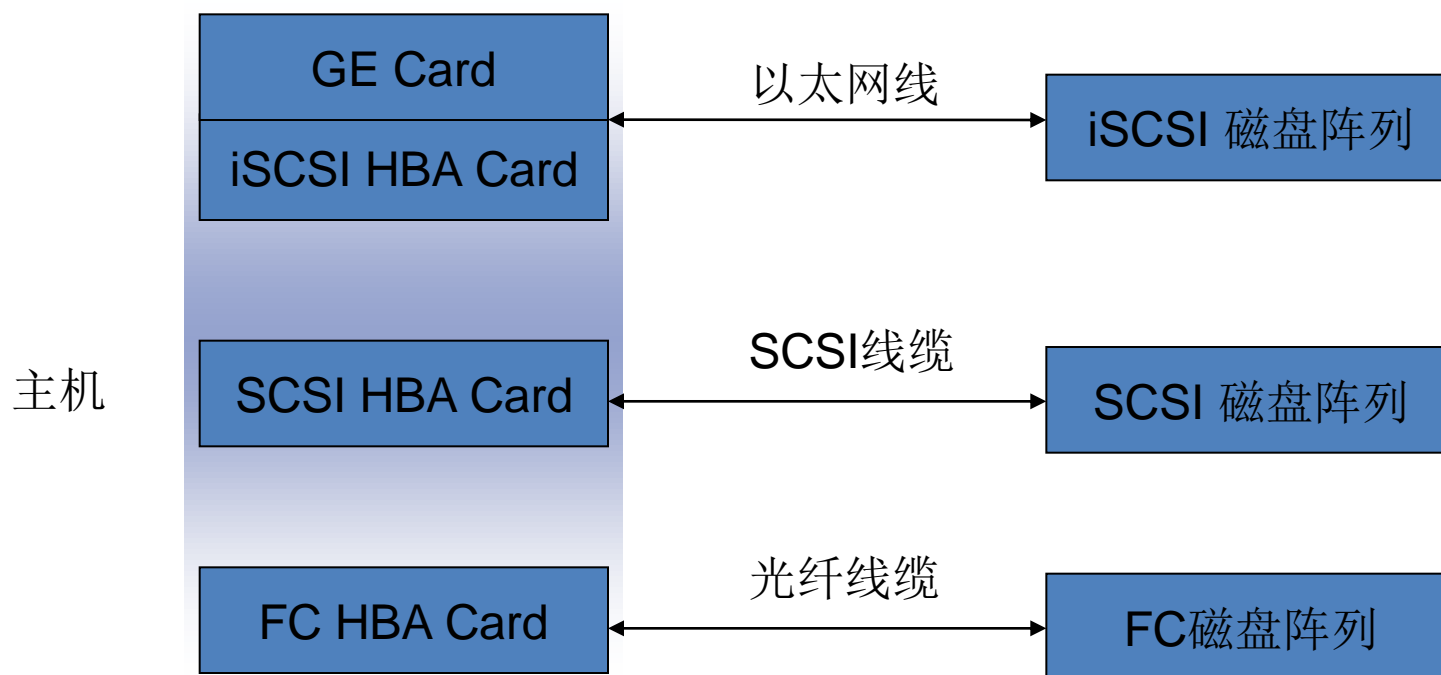
# 传输通道

- 存储传输的硬件通道有三种：

通过以太网线连接

通过SCSI线缆连接

通过光纤线缆连接



# SCSI协议

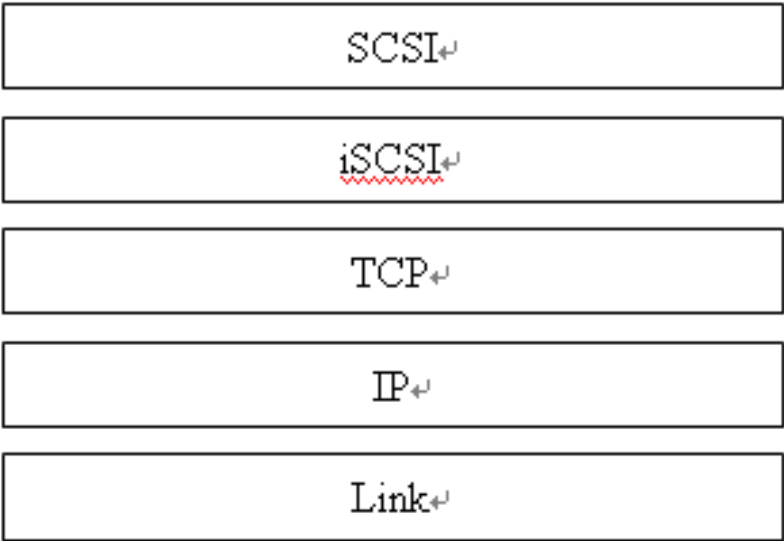


- SCSI是小型计算机系统接口（Small Computer System Interface）的简称，主要用于高端系统的连接，现在已完全普及到了小型机及高端服务器。
- SCSI存储是以主机为中心的；一条SCSI总线上可以串接多台不同类型的设备。

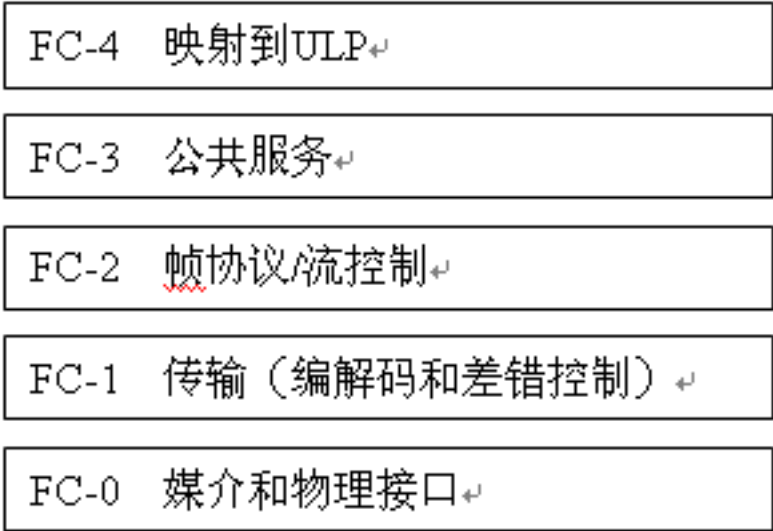
**SCSI协议**可以划分为SCSI-1、SCSI-2、SCSI-3，最新的为SCSI-3，也是目前应用最广泛的SCSI版本。属于SCSI-3的 Ultra 320 SCSI 的最高数据传输率已经达到了320MB/s。



# iSCSI和FC协议层次比较



iSCSI协议的分层结构



FC协议的分层结构

# 目 录

1. 存储组网形态
2. 主要协议
3. RAID
4. 磁盘热备与重构
5. 快照与复制
6. 数据分级存储概念

# RAID简介

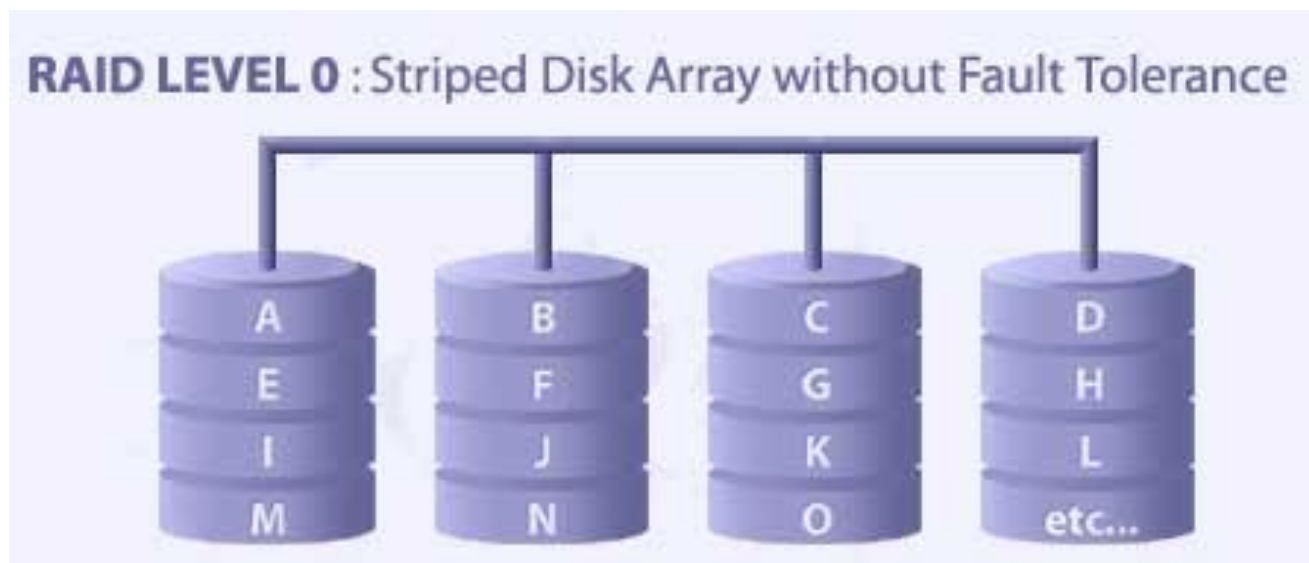
RAID是廉价冗余磁盘阵列(Redundant Array of Inexpensive Disks) 的简称

名词	说明
分区	又称为Extent；是一个磁盘上的地址连续的存储块。一个磁盘可以划分为多个分区，每个分区可以大小不等，有时也称为逻辑磁盘。
分块	又称为Strip；将一个分区分成多个大小相等的、地址相邻的块，这些块称为分块。分块通常被认为是条带的元素。虚拟磁盘以它为单位将虚拟磁盘的地址映射到成员磁盘的地址。
条带	又称为Stripe；是阵列的不同分区上的位置相关的strip的集合，是组织不同分区上条带的单位。
软RAID	RAID 的所有功能都依赖于操作系统（OS）与服务器CPU来完成，没有第三方的控制/处理（业界称其为RAID 协处理器——RAID Co-Processor）与I/O芯片
硬RAID	有专门的RAID 控制/处理与I/O处理芯片，用来处理RAID任务，不需耗用主机CPU资源，效率高，性能好。

# RAID0 (一)

RAID0 : Striped Disk Array without Fault Tolerance( 没有容错设计的条带磁盘阵列，以条带形式将RAID阵列的数据均匀分布在各个阵列中。

总容量=(磁盘数量)\*(磁盘容量)



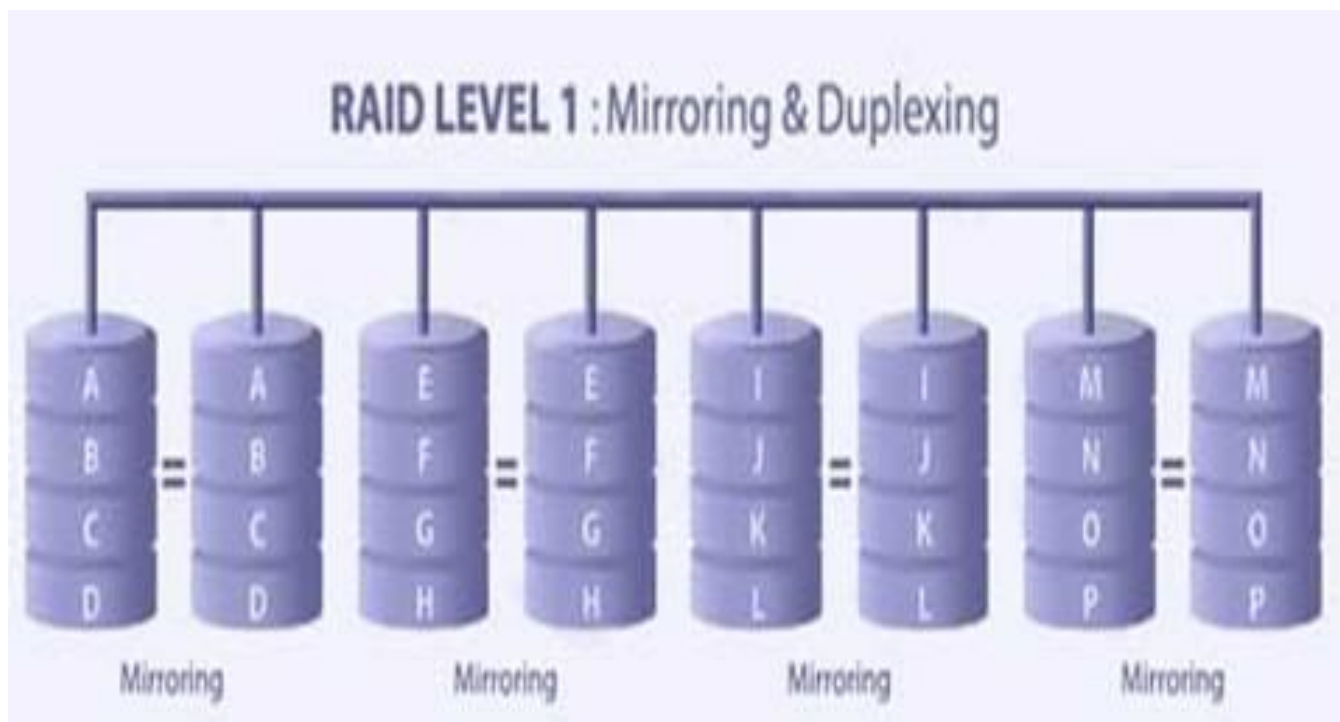
## RAID0（二）

优点	缺点
<ul style="list-style-type: none"><li>➤ 可多I/O操作并行处理，极高的读写效率</li><li>➤ 速度快，由于不存在校验，所以不占用cpu资源</li><li>➤ 设计，使用与配置简单</li></ul>	<ul style="list-style-type: none"><li>➤ 无冗余，一个RAID 0的磁盘失败，那么数据将彻底丢失</li><li>➤ 不能用于关键数据环境</li></ul>
<p><b>适用领域：</b></p> <ul style="list-style-type: none"><li>➤ 视频生成和编辑</li><li>➤ 图像编辑</li><li>➤ 较为“拥挤”的操作</li><li>➤ 其他需要大的传输带宽的操作</li></ul>	
<p>至少需要磁盘数：2个</p>	

# RAID1（一）

RAID 1：以镜像作为冗余手段，虚拟磁盘中的数据有多个拷贝，放在成员磁盘上。

总容量=(磁盘数量 / 2)\*(磁盘容量)



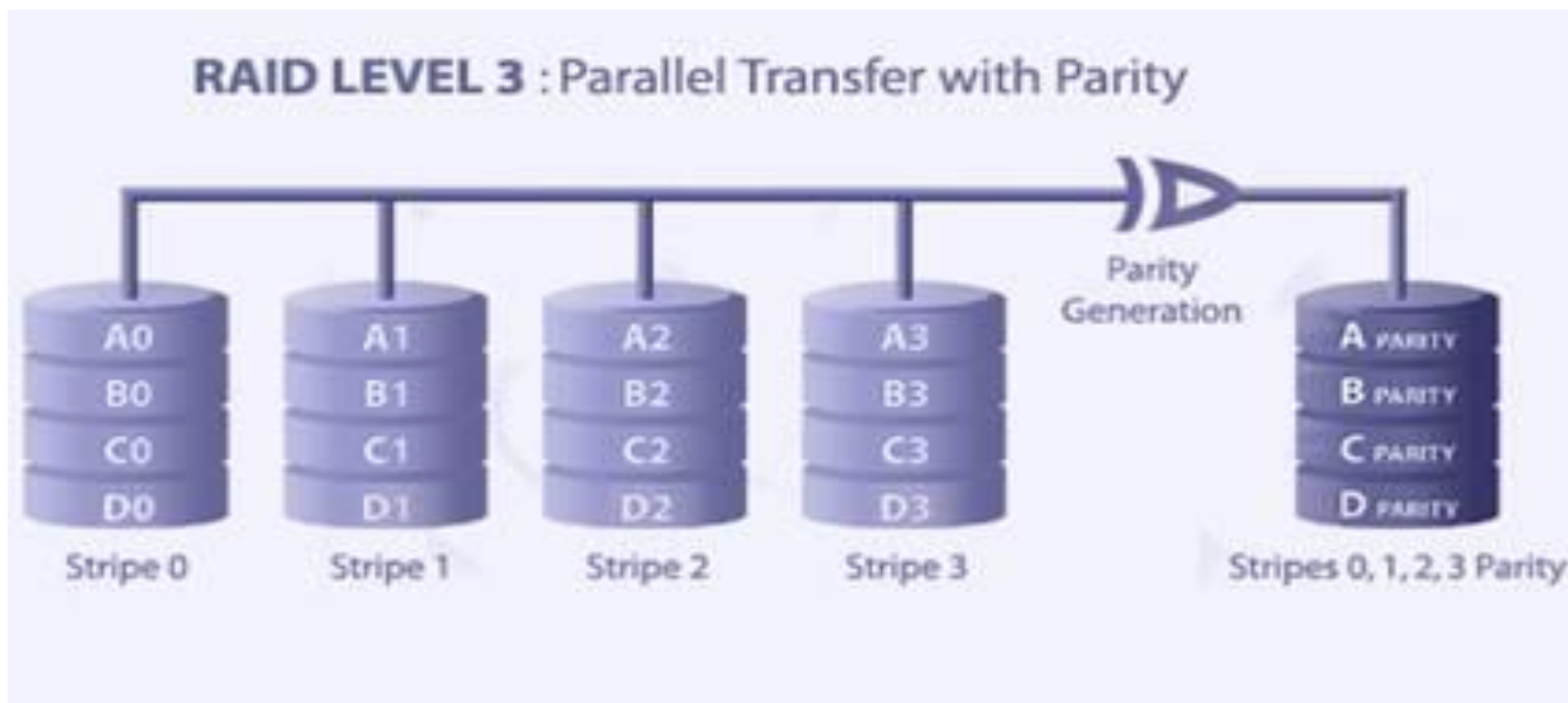
## RAID1（二）

优点	缺点
<ul style="list-style-type: none"><li>➤ 理论上读效率是单个磁盘的两倍；</li><li>➤ 100%的数据冗余；</li><li>➤ 设计、使用简单</li></ul>	<ul style="list-style-type: none"><li>➤ ECC效率低下，磁盘ECC的CPU占用率是所有RAID等级中最高的，在所有成本高；</li><li>➤ 软RAID方式下，很少能支持硬盘的热插拔；</li><li>➤ 空间利用率只有1/2</li></ul>
<p>适用领域：</p> <ul style="list-style-type: none"><li>➤ 财务统计与数据库</li><li>➤ 金融系统</li><li>➤ 其他需要高可用的数据存储环境</li></ul>	
至少需要磁盘数 2个	

## RAID3 (一)

RAID3 (条带分布+专用盘校验): 以xor校验为冗余方式, 使用专门的磁盘存放校验数据, 虚拟磁盘上的数据块被分为更小的数据块并行传输到各个成员物理磁盘上, 同时计算出xor校验数据存放到校验磁盘上。

总容量=(磁盘数量 -1)\*(磁盘容量)





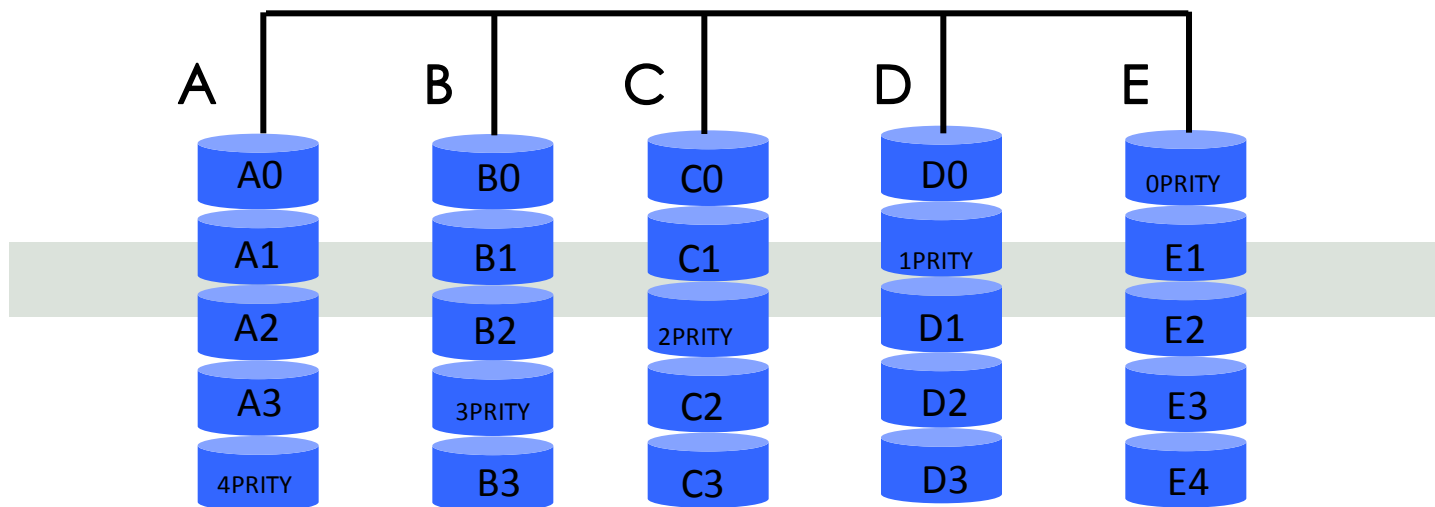
## RAID3（二）

优点	缺点
<ul style="list-style-type: none"><li>➤ 相对较高的读取传输率</li><li>➤ 高可用性，如果有一个磁盘损坏，对吞吐量影响较小</li><li>➤ 高效率的ECC操作</li></ul>	<ul style="list-style-type: none"><li>➤ 校验盘成为性能瓶颈</li><li>➤ 每次读写牵动整个组，每次只能完成一次I/O</li></ul>
<p>适用领域：</p> <ul style="list-style-type: none"><li>➤ 视频生成和在线编辑</li><li>➤ 图像和视频编辑</li><li>➤ 其他需要高吞吐量的场合</li></ul>	
<p>至少需要磁盘数 3个</p>	

# RAID5 (一)

RAID5 (条带技术+分布式校验): 以XOR检验为冗余方式, 校验数据均匀分布在各个数据磁盘上, 对各个数据磁盘的访问为异步操作。

总容量=(磁盘数-1)\*(磁盘容量)



## RAID5（二）

优点	缺点
<ul style="list-style-type: none"><li>➤ 高读取速率</li><li>➤ 中等写速率</li></ul>	<ul style="list-style-type: none"><li>➤ 异或校验影响存储性能</li><li>➤ 磁盘损坏后，重建很复杂</li></ul>
<p>适用领域：</p> <ul style="list-style-type: none"><li>➤ 文件服务器和应用服务器</li><li>➤ OLTP环境的数据库</li><li>➤ WEB， E - MAIL服务器</li></ul>	
<p>至少需要磁盘数 3个</p>	

## RAID6（二）

优点	缺点
<ul style="list-style-type: none"><li>➤ 快速的读取性能</li><li>➤ 更高的容错能力</li></ul>	<ul style="list-style-type: none"><li>➤ 很慢的写入速度</li><li>➤ 成本更高</li></ul>
<p>适用领域：</p> <ul style="list-style-type: none"><li>➤ 高可靠性环境</li></ul>	
<p>至少需要磁盘数： 4个</p>	

## RAID10（二）

优点	缺点
<ul style="list-style-type: none"><li>➤ 高读取速率</li><li>➤ 高写速率，较校验RAID而言，写开销最小</li><li>➤ 至多可以容许N个磁盘同时损坏（2N个磁盘组成的RAID10阵列）</li></ul>	<ul style="list-style-type: none"><li>➤ 贵</li><li>➤ 只有1/2的磁盘利用率</li></ul>
<p>适用领域：</p> <ul style="list-style-type: none"><li>➤ 要求高可靠性和高性能的数据库服务器</li></ul>	
<p>至少需要磁盘数： 4个</p>	

## RAID50（二）

优点	缺点
<ul style="list-style-type: none"><li>➤ 比单个RAID5容纳更多的磁盘</li><li>➤ 比单个RAID5有更好的读性能</li><li>➤ 至多可以容许n个磁盘同时损坏（N个RAID5组成的RAID50阵列）</li><li>➤ 比相同容量的单个RAID5重建时间更短</li></ul>	<ul style="list-style-type: none"><li>➤ 比较难实现</li><li>➤ 同一个RAID5组内的两个磁盘损坏会导致整个RAID50阵列的失效</li></ul>
<p>适用领域：</p> <ul style="list-style-type: none"><li>➤ 大型数据库服务器</li><li>➤ 应用服务器</li><li>➤ 文件服务器</li></ul>	
<p>至少需要磁盘数： 6个</p>	

# 目 录

1. 存储组网形态
2. 主要协议
3. RAID
4. 磁盘热备与重构
5. 快照与复制
6. 数据分级存储概念

# 热备（HotSpare）

- ◆ 所谓热备份是在建立RAID磁盘阵列系统的时候，将其中一磁盘指定为热备磁盘，此热备磁盘在平常并不操作，当阵列中某一磁盘发生故障时，热备磁盘便取代故障磁盘，并自动将故障磁盘的数据重构在热备磁盘上。
- ◆ 热备盘分为：全局热备盘和局部热备盘
  - 全局热备盘：针对整个磁盘阵列，对阵列中所有RAID组起作用。
  - 局部热备盘：只针对某一RAID组起作用。
- ◆ 因为反应快速，加上快取内存减少了磁盘的存取，所以数据重构很快即可完成，对系统的性能影响不大。对于要求不停机的大型数据处理中心或控制中心而言，热备份更是一项重要的功能，因为可避免晚间或无人守护时发生磁盘故障所引起的种种不便。

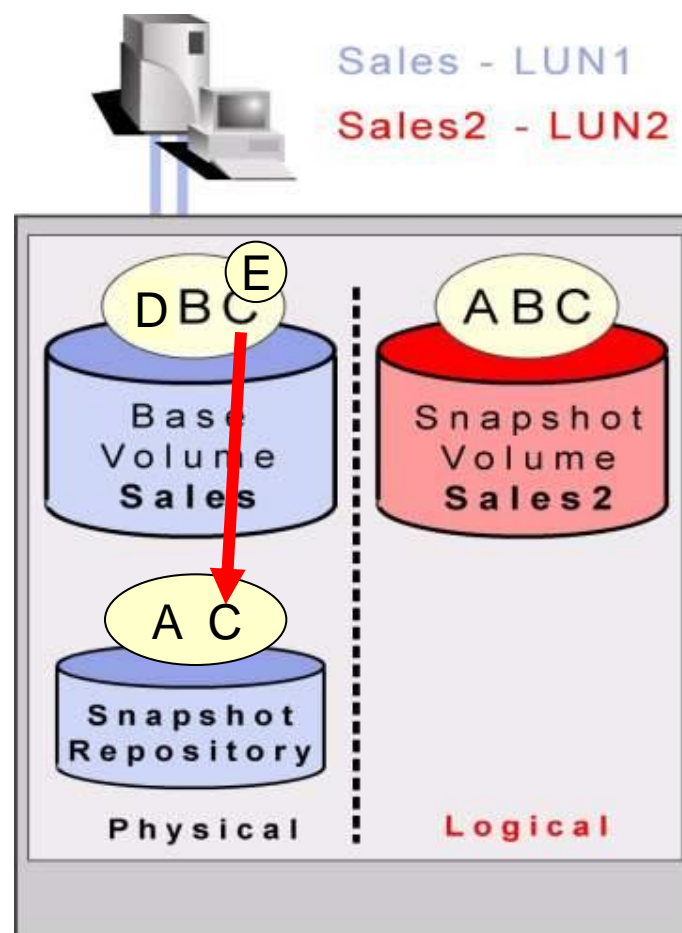


# 目 录

1. 存储组网形态
2. 主要协议
3. RAID
4. 磁盘热备与重构
5. 快照与复制
6. 数据分级存储概念

# 逻辑卷快照(snapshot)

- 基本概念：
  - Base Volume：快照源卷
  - Repository Volume：快照仓储卷，保存快照源卷在快照过程中被修改以前的数据
  - Snapshot Volume：快照卷
- 某一个时间点的逻辑卷映像：
  - 逻辑上相当于整个Base Volume的拷贝
  - 可将Sanpshot Volume分配给任何一台主机
  - Snapshot Volume可读取、写入或拷贝
- 存储空间需求
  - 需要相当于Base Volume 20%的额外空间
- 用途
  - 文件、逻辑卷恢复
  - 备份、测试、数据分析等



# 快照(snapshot)开始

W R I T E S

首先，在快照完成之前控制器是禁止对源卷进行写操作的。

B A S E V O L U M E X Y Z

仓储卷默认是源卷大小的20%

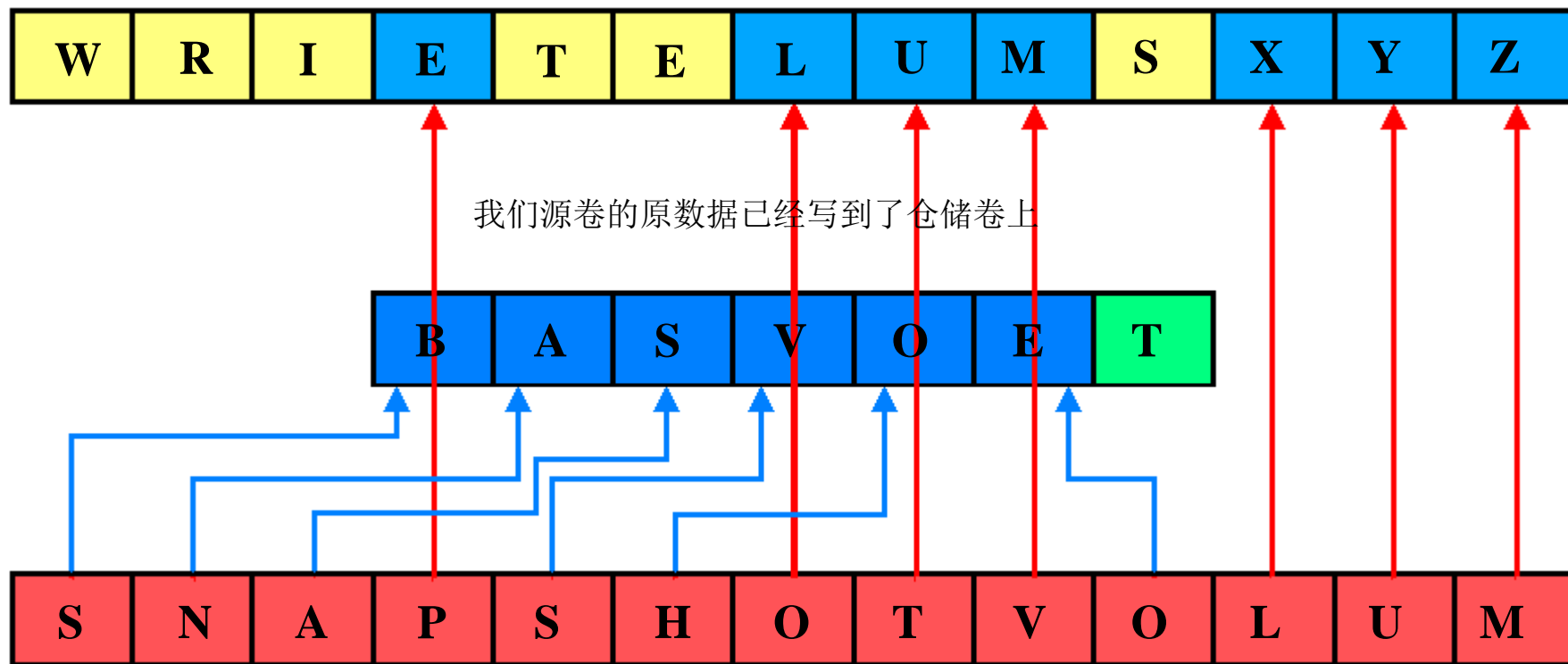
R E P O S I T

满足上面条件后，快照过程开始进行

S N A P S H O T V O L U M E

# 最后一步

源卷数据更新完毕



# 目 录

1. 存储组网形态
2. 主要协议
3. RAID
4. 磁盘热备与重构
5. 快照与复制
6. 数据分级存储概念

# 数据分级存储概念

- 数据分级存储：
- 把数据存放在不同类别的存储设备（磁盘、磁盘阵列、光盘库、磁带）
- 中，通过分级存储管理软件实现数据实体在存储设备之间的自动迁移。根据
- 数据的访问频率、保留时间、容量、性能要求等因素确定最佳存储策略，从
- 而控制数据迁移的规则。
- 分级存储的优点：
  - 最大限度地满足用户需求；
  - 减少总体存储成本；
  - 性能优化；
  - 改善数据可用性；
  - 数据迁移对应用透明。

# 分级存储的存储方式

存储方式	描述	举例
线存储 (On-line storage)	数据存放在磁盘系统上。在线存储一般采用高端存储系统和技术如：SAN、点对点直连技术、S2A。存取速度快，价格昂贵	电视台的在线存储：用于存储即将用于制作、编辑、播出的视音频素材。并随时保持可实时快速访问的状态。在这类应用中，在线存储设备一般采用SCSI磁盘阵列、光纤磁盘阵列等，
离线存储 (Off-line Storage)	数据备份到磁带、磁带库或光盘库上。访问速度低，但能实现海量存储，同时价格低廉。	电视台的离线存储：平时没有连接在编辑/播出系统，在需要时临时性地装载或连接到编辑/播出系统。可以将总的存储做得很大。括制作年代较远的新闻片、专题片等。
近线存储	不经常用到，数据的访问量不大的数据存放在性能较低的存储设备上，同时对这些设备的要求是寻址迅速、传输率高。	近线存储介于在线存储和离线存储之间，既可以做到较大的存储容量，又可以获得较快的存取速度。近线存储设备一般采用自动化的数据流磁带或者光盘塔。近线存储设备用于存储和在线设备发生频繁读写交换的数据，包括近段时间采集的视音频素材或近段时间制作的新闻片、专题片等。