

1. 简单介绍 raid 的级别？

Raid 级别	描述
raid 0	读写性能佳，坏了其中一块，数据挂掉，可靠性低，磁盘利用率 100%
raid 1	镜像备份，同一份数据完整的保存在多个磁盘上，写的性能不佳，可靠性高，读的性能还行，磁盘利用率 50%
raid 5	由多块磁盘做 raid 5，磁盘利用率为 $n-1/n$ ，其中一块放校验数据，允许坏一块盘，数据可以利用校验值来恢复
raid 6	在 raid 5 的基础上再加一块校验盘，进一步提高数据可靠性
raid 1+0	先做 raid 1 再做 raid 0

2. 简述 raid 0、raid 1、raid 5 三种工作模式的工作原理及特点

RAID 0：带区卷，连续以 位 或 字节 为单位分割数据，并行 读/写 于多个磁盘上，因此具有很高的数据传输率但它没有数据冗余。

RAID 0 只是单纯地提高性能，并没有为数据的可靠性提供保证，而且其中的一个磁盘失效将影响到所有数据。因此，RAID 0 不能应用于数据安全性要求高的场合

RAID 1：镜像卷，它是通过磁盘数据镜像实现数据冗余，在成对的独立磁盘上产生 互为备份 的数据，不能提升写数据效率。当原始数据繁忙时，可直接从镜像拷贝中读取数据。

因此 RAID 1 可以提高读取性能

RAID 1 是磁盘阵列中单位成本最高的，镜像卷可用容量为总容量的  $1/2$ ，但提供了很高的数据安全性和可用性。 当一个磁盘失效 时，系统可以自动切换到镜像磁盘上读写，而不需要重组失效的数据

RAID 5：至少由 3 块硬盘组成，分布式奇偶校验 的独立磁盘结构，它的奇偶校验码存在于所有磁盘上 任何一个硬盘损坏，都可以根据其它硬盘上的校验位来重建损坏的数据（最多允许1块硬盘损坏）

所以 raid 5 可以实现数据冗余， 确保数据的安全性，同时 raid 5 也可以提升数据的读写性能

3. 网站有存储下载文件的需求，用什么存储文件比较好

glusterfs、nfs

4. 简单介绍 glusterfs 的特点

glusterfs 是无元数据服务器设计，没有单点故障和性能瓶颈，有很好的扩展性，和稳定性，认为存储是软件的事，不能局限于硬件。以原始数据的形式存储，访问数据简单，迁移容易。有的视频公司将他作为片库。

缺点：数据一致性问题复杂，文件目录遍历效率低，缺乏全局监控，客户端负载大，占用了大量 cpu 和内存，用户空间效率低，与内核空间要经常交换数据，借用 FUSE，有性能损耗

## 5. 还有什么比较好的文件存储方式

ceph

## 6. DAS、NAS、SAN 使用场景

DAS：存储设备通过 `scsi` 或 `fc` 接口直连到服务器，数据与操作系统未分离。

NAS：将存储功能与服务器分离开，用以太网存取数据，有自己的文件管理系统，是文件级存储。多适用于文件服务器，用来存储非结构化数据。

SAN：是通过光纤交换机连接存储阵列和服务器，建立专用数据存储的存储私网。每台服务器独自管理文件。

SAN 允许任何服务器连接到任何存储阵列

（好处是：不管数据放在哪里，服务器都可以直接存取所需的数据哦）适用于大型应用或数据库系统，缺点是成本高、较为复杂。

通俗一点，在 `linux` 里 `lsblk`，看到的就是 SAN 提供的块设备，`windows` 上的共享文件夹，一般是 NAS 的 NFS 服务。

## 7. 存储类型的分类有哪几种？并简单进行描述

存储类型 分类	描述
文件存储	NAS 都属于这一类。简单来说就是 <code>mount</code> 后直接使用的。
块存储	SAN 都属于这一类。简单来说就是类似 <code>/dev/sdb</code> 这种，要分区，格式化后才能 <code>mount</code> 使用
对象存储	简单来说，就是直接使用程序接口去访问

## 8. 请你对各个存储类型的优缺点进行比较？

文件存储：类似一个大的目录，多个客户端都可以挂载过来使用

优点：利于数据共享

缺点：速度较慢

块存储：类似一个 `block` 设备，客户端可以格式化，挂载并使用，和用一个硬盘一样。

优点：和本地硬盘一样，直接使用

缺点：数据不共享

对象存储：一个对象我们可以看成一个文件，综合了文件存储和块存储的优点

优点：速度快，数据共享

缺点：成本高，不兼容现有的模式

## 9. 什么是分布式存储？

分布式存储可以看作拥有多台存储服务器连接起来的存储导出端（多对一，多对多）。把这多台存储服务器的存储合起来做成一个整体再通过网络进行远程共享,共享的方式有目录(文件存储)，块设备(块存储)，对象网关 或者说 一个程序接口(对象存储)。常见的分布式存储开源软件有：GlusterFS、Ceph、HDFS、MooseFS、FastDFS 等

## 10. 分布式存储的优点有哪些？

分布式存储一般都有以下几个优点：

1. 扩容方便，轻松达到 PB 级别或以上
2. 提升读写性能或数据高可用
3. 避免单个节点故障导致整个架构问题
4. 价格相对便宜，大量的廉价设备就可以组成，比光纤 SAN 这种便宜很多。