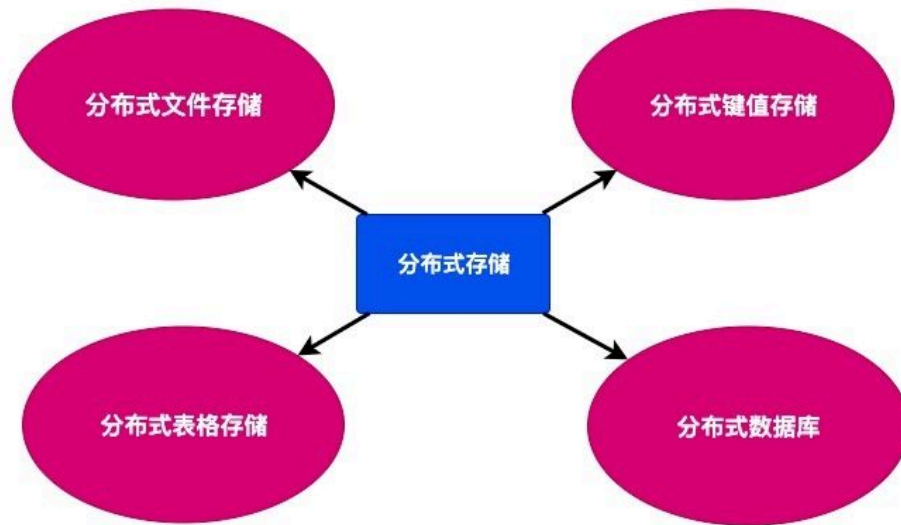


# 分布式存储

前面简单介绍了分布式系统，这里重点介绍下分布式存储。当然分布式存储也分为很多种类，在这里简单分为以下四类：



## 一、分布式文件存储

分布式文件存储主要用来存储各类无结构数据，如文件，数据块等。系统以对象的形式组织，对象之间没有联系，这样的数据称为 Blob，也就是二进制文件。

分布式文件存储一般都是将一个文件以 chunk 的形式切分为很多个小文件。然后通过一致性 hash 以及其变种算法，将文件分别存放在不同的服务器磁盘上。达到海量数据存储，备份的目的。目前市面上使用的分布式文件存储系统已经有很多成熟的产品。如各大云厂商提供的，如阿里云 OSS、AWS S3 等。同时也有一些非常优秀的开源产品，如 openstack swift，ceph 等都是非常优秀的分布式文件存储系统。其中 swift 为跟 OSS 类似，都为对象存储。swift 是通过一致性 hash-ring 的方式将数据分别存放在不同数据中心的数据盘上，并采取一定的冗余策略对数据进行冗余，以防止丢失。

分布式文件存储，往往也会作为其他存储类型的底层存储服务。如数据库的备份等一些主要场景。

## 二、分布式键值系统

说起键值系统，相信大家第一时间会想到 redis，memcache 这样的缓存数据库，但是不管是 redis 还是 memcache 其本身设计并不是分布式的。

像 redis 还是采用的主-从架构。跟分布式系统对等节点的方式有本质的区别。etcd 是一个严格的分布式 k-v 存储系统。在 etcd2.0 的版本，其还是基于内存的内存数据库，可以处理的数据量并不大。但随着 etcd v3 的升级，etcd 将数据写入到了磁盘中，极大的扩展了 etcd 的使用场景。使得 etcd 不仅可以作为配置中心，还可以做缓存，半结构化数据存储，等一系列的场景。当然了，相比与 redis 提供的丰富的数据结构，etcd 还是差的很远。

## 三、分布式表格存储

分布式表格存储是介于 k-v 存储与数据库之前的一种存储方式，除了能够支持简单的 CURD 之外，而且可以扫描某个主键的范围，也就是可以进行简单的范围查找。

分布式表格存储更多的是对于单表的操作，不支持一些特别复杂的操作，如多表关联，多表嵌套，嵌套子查询等。分布式表格系统中，同一个表格的多个数据行也不要求包含相同类型的列。

#### 四、分布式数据库

数据库主要是用来存储结构化的数据，主要有关系型数据库SQL、非关系型数据库NoSQL、还有一些其他的数据库类型。如图数据库，时序数据库等。分布式数据库是为了应对互联网环境下，海量的数据，高速的业务扩张等原因逐步发展起来的。

相信了解数据库的同学都知道，互联网业务一般都是高速增长，很多上周评估迁移的数据资源，这周就已经吃紧了，数据库管理人员就又不es不得不对数据进行迁移跟扩容，这个重复的过程在保证业务可用的前提下，可谓是相当痛苦。正是在这样的大背景下，数据库人员为了解决扩容、迁移、高可用等痛点，提出并开发了分布式数据库。

我们都知道数据库有很多复杂的查询，同时还要支持事务、并发控制等一系列核心功能。所以数据库的分布式是一个很大的挑战，比上面所有提到的分布式系统都要复杂。但是，这些年随着分布式系统的发展跟成熟，行业里面也出现了很多分布式数据库，其中就有蚂蚁的OceanBase，阿里云的DRDS，大数据生态中的HBase，开源的Tidb等，还有一些其他形式的分布式数据库。其中tidb跟我们的etcd又有很多相通的地方。如都是采用Raft一致性算法，都是采用Go语言编写的等等。