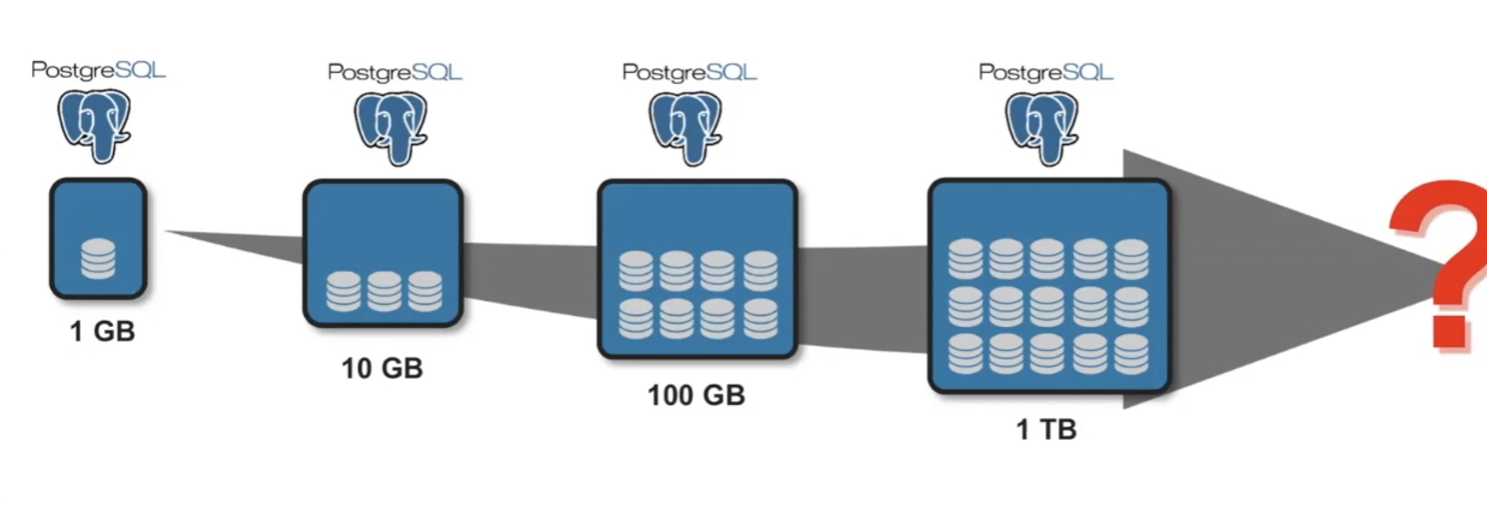
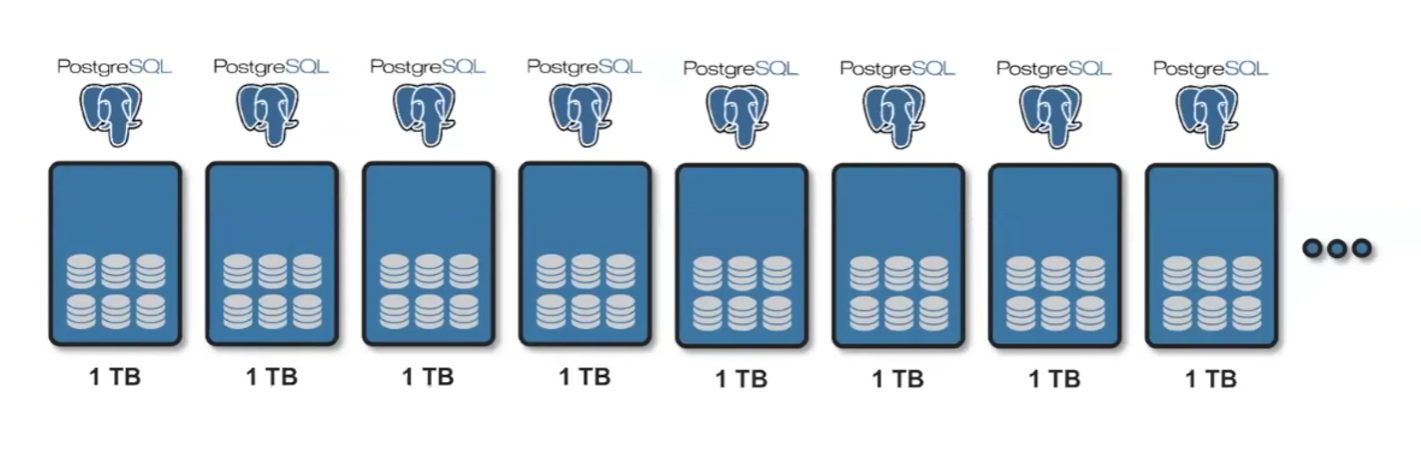
Postgres升级

Scale up(纵向扩展)

主要是利用现有的存储系统，通过不断增加存储容量来满足数据增长的需求



Scale out(横向扩展)架构的升级通常是以节点为单位，每个节点往往将包含容量、处理能力和I / O带宽。一个节点被添加到存储系统，系统中的三种资源将同时升级。



Greenplum基于Scale out架构

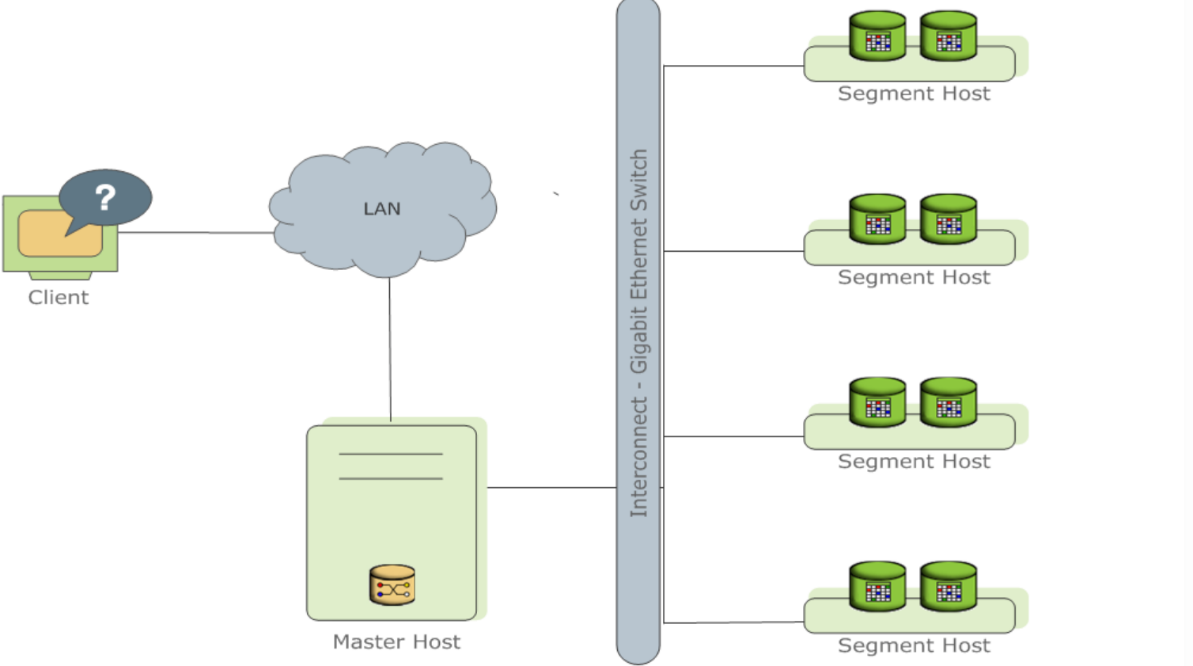
控制节点（master）和多个数据节点（segment Host）构成的集群

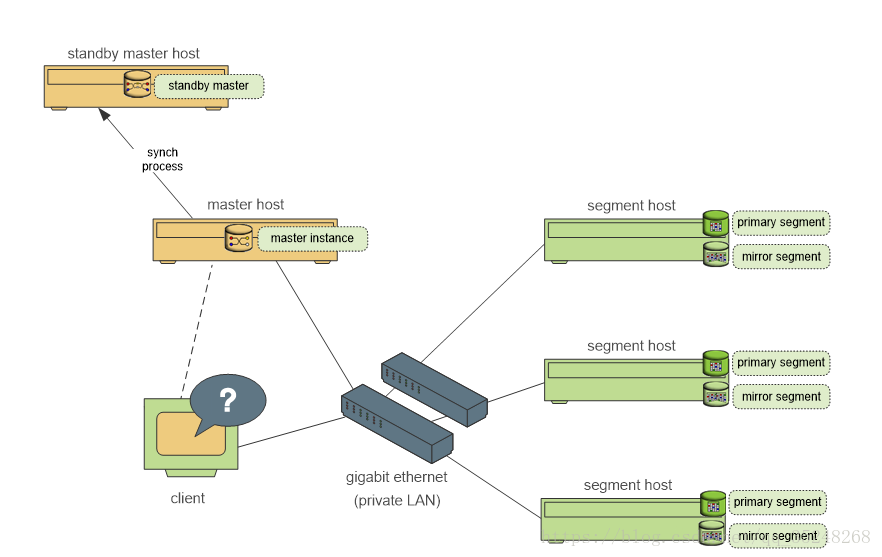
Greenplum主要由Master节点、Segment节点、interconnect三大部分组成。Greenplum master是Greenplum数据库系统的入口，接受客户端连接及提交的SQL语句，将工作负载分发给其它数据库实例（segment实例），由它们存储和处理数据。Greenplum interconnect负责不同PostgreSQL实例之间的通信。Greenplum segment是独立的PostgreSQL数据库，每个segment存储一部分数据。大部分查询处理都由segment完成。

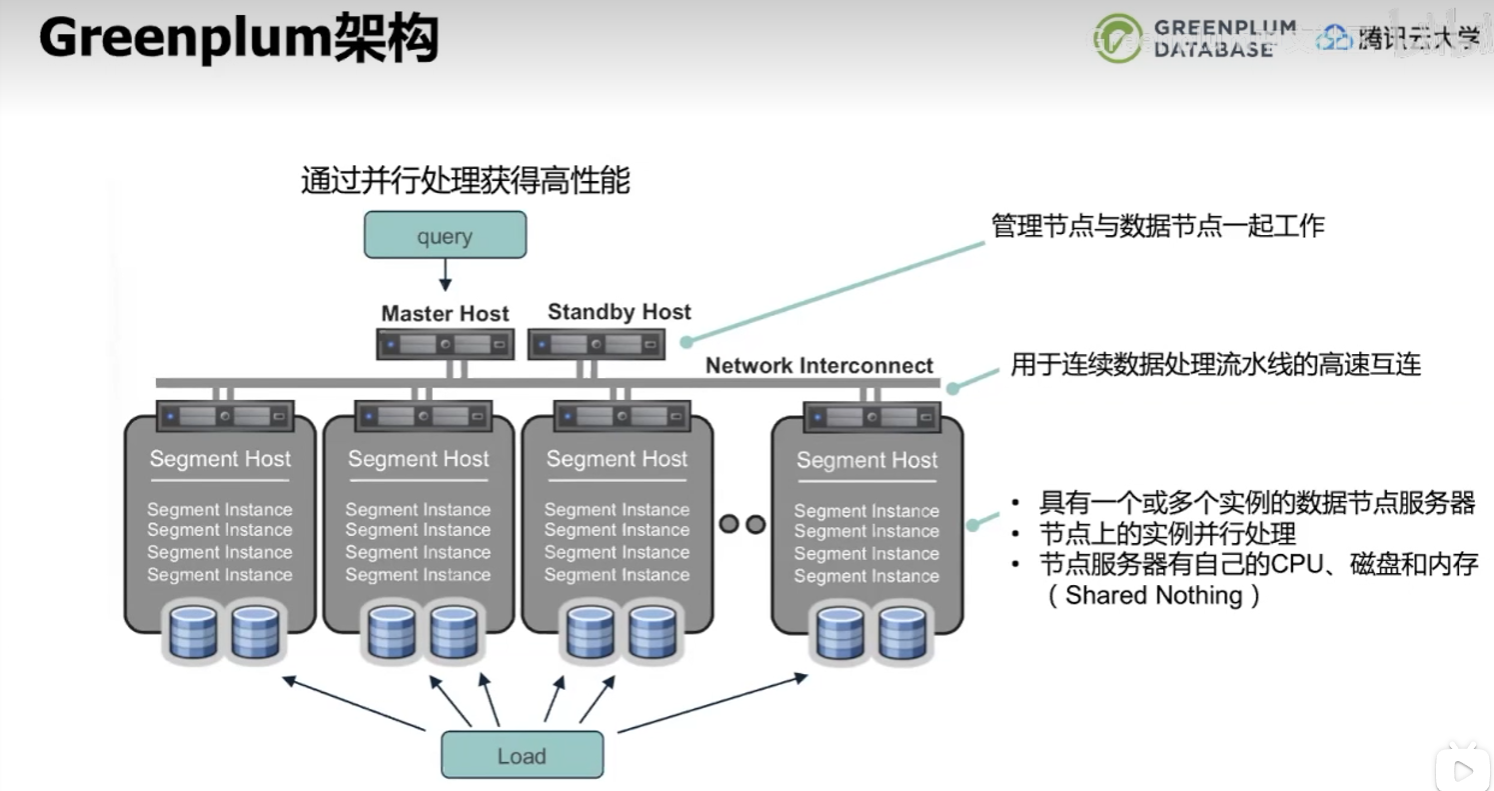
Master节点不存放任何用户数据，只是对客户端进行访问控制和存储表分布逻辑的元数据

Segment节点负责数据的存储，可以对分布键进行优化以充分利用Segment节点的io性能来扩展整集群的io性能

存储方式可以根据数据热度或者访问模式的不同而使用不同的存储方式。一张表的不同数据可以使用不同的物理存储方式：行存储、列存储、外部表

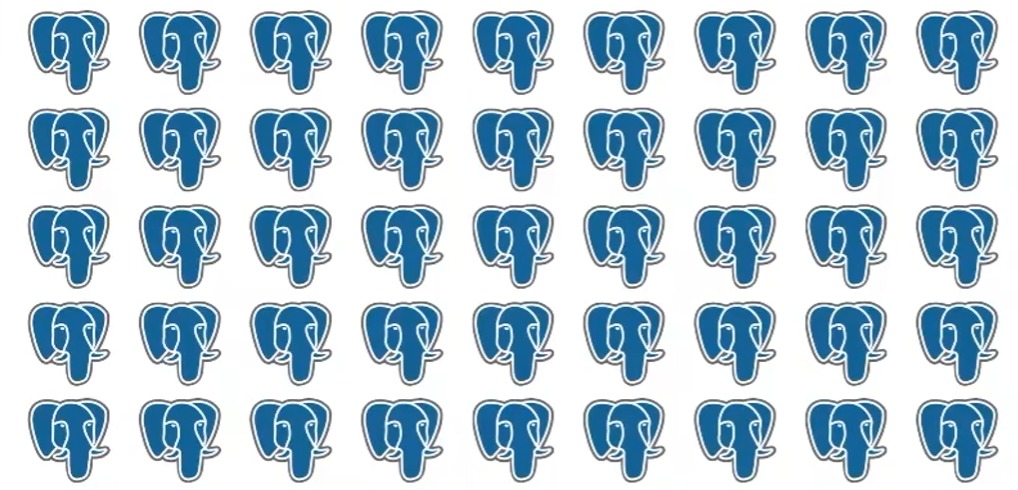


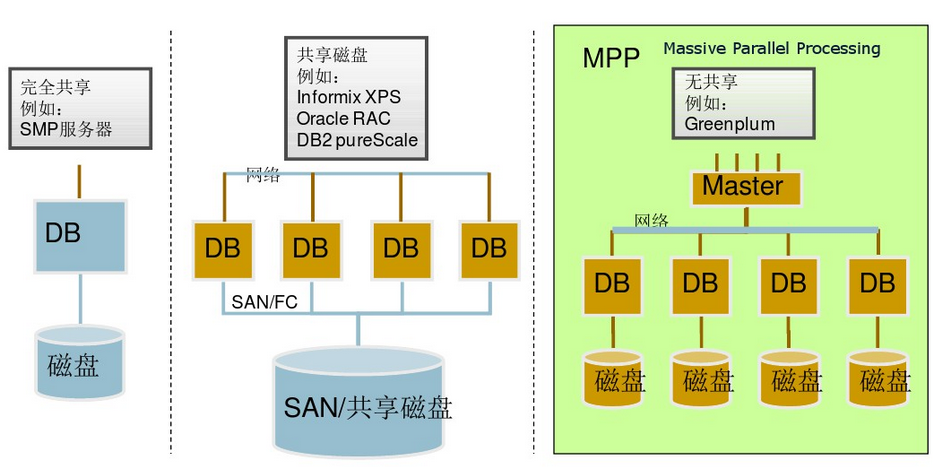






本质上是单机Postges组成的数据库集群





GP的三个关键词

1、shared-nothing

Shared Everthting:一般是针对单个主机，完全透明共享CPU/MEMORY/IO，并行处理能力差，典型的代表SQLServer。 shared-everything架构优点很明显，但是网络，硬盘很容易就会成为系统瓶颈。

Shared Disk：各个处理单元使用自己的私有 CPU和Memory，共享磁盘系统。典型的代表Oracle Rac， 它是数据共享，可通过增加节点来提高并行处理的能力，扩展能力较好。其类似于SMP（对称多处理）模式，但是当存储器接口达到饱和的时候，增加节点并不能获得更高的性能 。

Shared Nothing：各个处理单元都有自己私有的CPU/内存/硬盘等，不存在共享资源，各处理单元之间通过协议通信，并行处理和扩展能力更好。各节点相互独立，各自处理自己的数据，处理后的结果可能向上层汇总或在节点间流转。Share-Nothing架构在扩展性和成本上都具有明显优势。

2、MPP(Massively Parallel Processor大规模并行处理)

大规模并行处理系统是由许多松耦合处理单元组成的，借助MPP这种高性能的系统架构，Greenplum可以将TB级的数据仓库负载分解，并使用所有的系统资源并行处理单个查询。MPP 架构通常由多个节点组成，每个节点都包含了独立的计算资源（例如 CPU、内存和存储等），并且节点之间可以相互通信和协同工作。MPP 架构的优势在于它可以将大型任务分解成多个子任务，并在多个节点上并行执行这些子任务，以提高计算效率和处理速度。例如，在 Greenplum 中，可以将大型数据集分成多个数据块，并将这些数据块分发到多个工作节点上并行处理，以便快速完成数据分析和处理任务。同时，MPP 架构还具有良好的扩展性和容错性，可以根据需要动态地添加或删除节点，并在节点失效时自动恢复和重试任务。

3、MVCC(多版本控制)

与事务型数据库系统通过锁机制来控制并发访问的机制不同， GPDB使用多版本控制(Multiversion Concurrency Control/MVCC)保证数据一致性。 这意味着在查询数据库时，每个事务看到的只是数据的快照，其确保当前的事务不会看到其他事务在相同记录上的修改。据此为数据库的每个事务提供事务隔离。

MVCC以避免给数据库事务显式锁定的方式，最大化减少锁争用以确保多用户环境下的性能。在并发控制方面，使用MVCC而不是使用锁机制的最大优势是， MVCC对查询(读)的锁与写的锁不存在冲突，并且读与写之间从不互相阻塞。

1. gpstate

命令 参数 => 作用

gpstate -b => 显示简要状态

gpstate -c => 显示主镜像映射

gpstart -d => 指定数据目录（默认值：$ MASTER\_DATA\_DIRECTORY）

gpstate -e => 显示具有镜像状态问题的片段

gpstate -f => 显示备用主机详细信息

gpstate -i => 显示GRIPLUM数据库版本

gpstate -m => 显示镜像实例同步状态

gpstate -p => 显示使用端口

gpstate -Q => 快速检查主机状态

gpstate -s => 显示集群详细信息

gpstate -v => 显示详细信息

实验二：down掉其中一个节点  
master hadoop106 segment host hadoop107 hadoop108

现在down掉hadoop108 查看集群状态及执行sql查询功能

故障节点恢复原理-同步原理

Primary master服务器和standby master服务器同步后，standbymaster服务器将通过walsender 和 walreceiver 的复制进程保持最新状态。该walreceiver是standby master上的进程， walsender流程是primary master上的进程。这两个进程使用基于预读日志（WAL）的流复制来保持primary master和standby master服务器同步。在WAL日志记录中，所有修改都会在应用生效之前写入日志，以确保任何进程内操作的数据完整性。

由于primary master不包含任何用户数据，因此只需要在主master和standby master之间同步系统目录表(catalog tables)。当这些表发生更新时，更改的结果会自动复制到备用master上，以确保与主master同步。

当启用Greenplum数据库高可用性时，有两种类型的segment：primary和mirror。每个主segment具有一个对应的mirror segment。主segment接收来自master的请求以更改主segment的数据库，然后将这些更改复制到相应的mirror segment上。如果主segment不可用，则数据库查询将故障转移到mirror segment上。

Primary segment和Mirror segment之间的同步与primary master和standby master相同，也是采用基于WAL流的同步。

group mirroring方式

其中每个主机的primary segment的mirror segment位于另一个主机上。如果单个主机发生故障，则部署该主机的mirror segment主机上的活动primary segment数量会翻倍，从而会加大该主机的负载

spread mirroring方式

每个主机的mirror segment分布在多个主机上，这样，如果任何单个主机发生故障，该主机的mirror segment会分散到其他多个主机上运行，从而达到负载均衡的效果

仅当主机数量多于每个主机的segment数时，才可以使用Spread方式

双重故障(double-fault)

在双重故障情况下，即primary segment和mirror segment都处于失败状态。如果不同segment的主机同时发生硬件故障，则会导致primary segment和mirror segment都处于失败状态。如果发生双重故障，Greenplum数据库将不可用。

[(22条消息) 【实战经验】Greenplum集群Master與Segment节点故障检测与恢复\_greenplum模拟数据节点挂掉\_阿福Chris的博客-CSDN博客](https://blog.csdn.net/chrisy521/article/details/123301971?ops_request_misc=&request_id=&biz_id=102&utm_term=greenplum%E8%8A%82%E7%82%B9%E6%81%A2%E5%A4%8D%E5%8E%9F%E7%90%86&utm_medium=distribute.pc_search_result.none-task-blog-2~all~sobaiduweb~default-0-123301971.142%5ev80%5einsert_down38,201%5ev4%5eadd_ask,239%5ev2%5einsert_chatgpt&spm=1018.2226.3001.4187)

[https://blog.csdn.net/chrisy521/article/details/123301971?ops\_request\_misc=&request\_id=&biz\_id=102&utm\_term=greenplum%E8%8A%82%E7%82%B9%E6%81%A2%E5%A4%8D%E5%8E%9F%E7%90%86&utm\_medium=distribute.pc\_search\_result.none-task-blog-2~all~sobaiduweb~default-0-123301971.142^v80^insert\_down38,201^v4^add\_ask,239^v2^insert\_chatgpt&spm=1018.2226.3001.4187](https://blog.csdn.net/chrisy521/article/details/123301971?ops_request_misc=&request_id=&biz_id=102&utm_term=greenplum%E8%8A%82%E7%82%B9%E6%81%A2%E5%A4%8D%E5%8E%9F%E7%90%86&utm_medium=distribute.pc_search_result.none-task-blog-2~all~sobaiduweb~default-0-123301971.142%5ev80%5einsert_down38,201%5ev4%5eadd_ask,239%5ev2%5einsert_chatgpt&spm=1018.2226.3001.4187)