



2017第八届中国数据库技术大会

DATABASE TECHNOLOGY CONFERENCE CHINA 2017

HashData的设计与实现

简丽荣, HashData

我们是谁

- 由Apache HAWQ（领先的SQL-on-Hadoop解决方案）创始团队于2016年初成立。



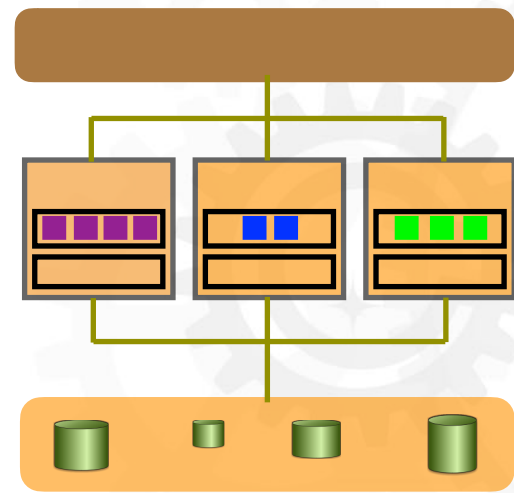
- 专注于SQL-on-Cloud云端数据仓库服务，核心团队由来自Pivotal, Teradata, IBM, Yahoo!等公司资深的分布式数据库、云计算和Hadoop专家组成。



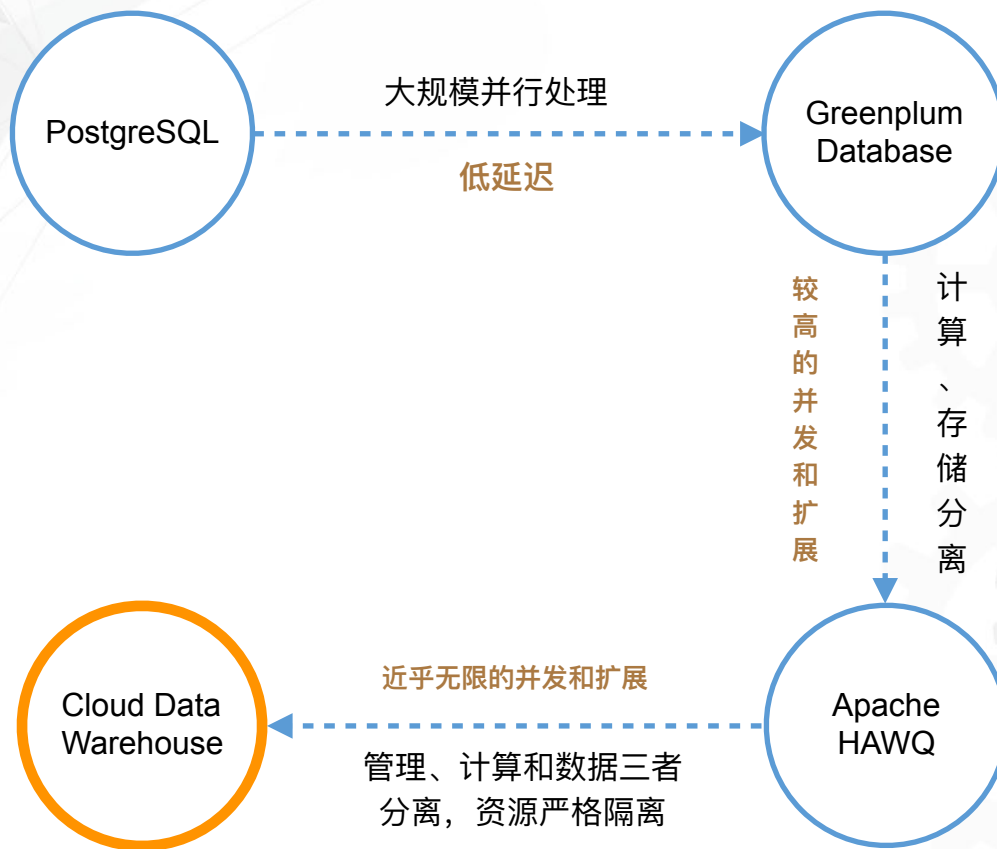
- 10+员工，大部分都是研发工程师（我们正在招人喔～）。

我们的产品

- HashData数据仓库
 - 完全托管的PB级数据仓库服务
 - 数据分析接口100%兼容PostgreSQL和Greenplum Database
 - 云原生
 - 计算和存储分离
 - 对象存储作为数据持久层
 - 元数据服务：在线升级和扩容
 - 一致性hash的数据分布策略：秒级扩容

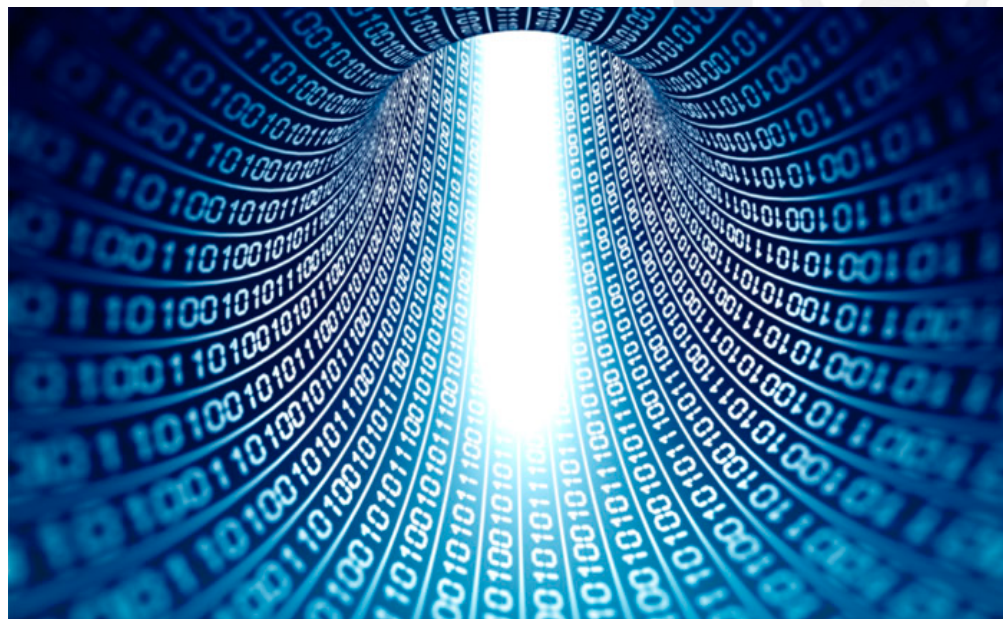


历史回顾



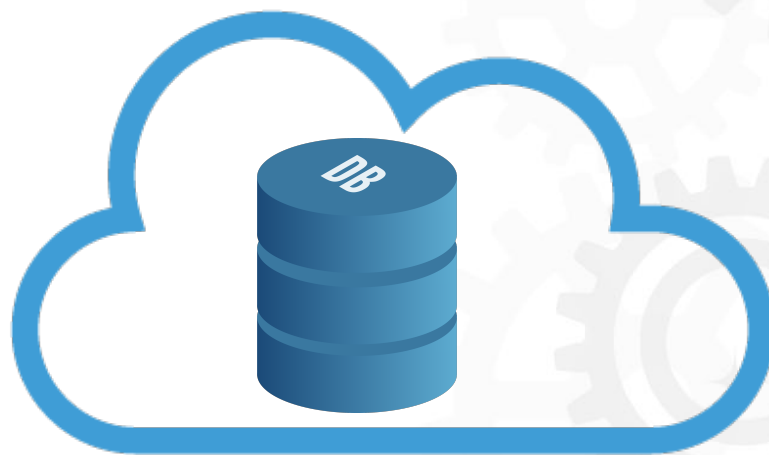
云计算

- 新的计算平台
- 新的操作系统
- 多维度弹性
- 几乎无限的可扩展性
- SaaS的交付模式
- 新的数据平台

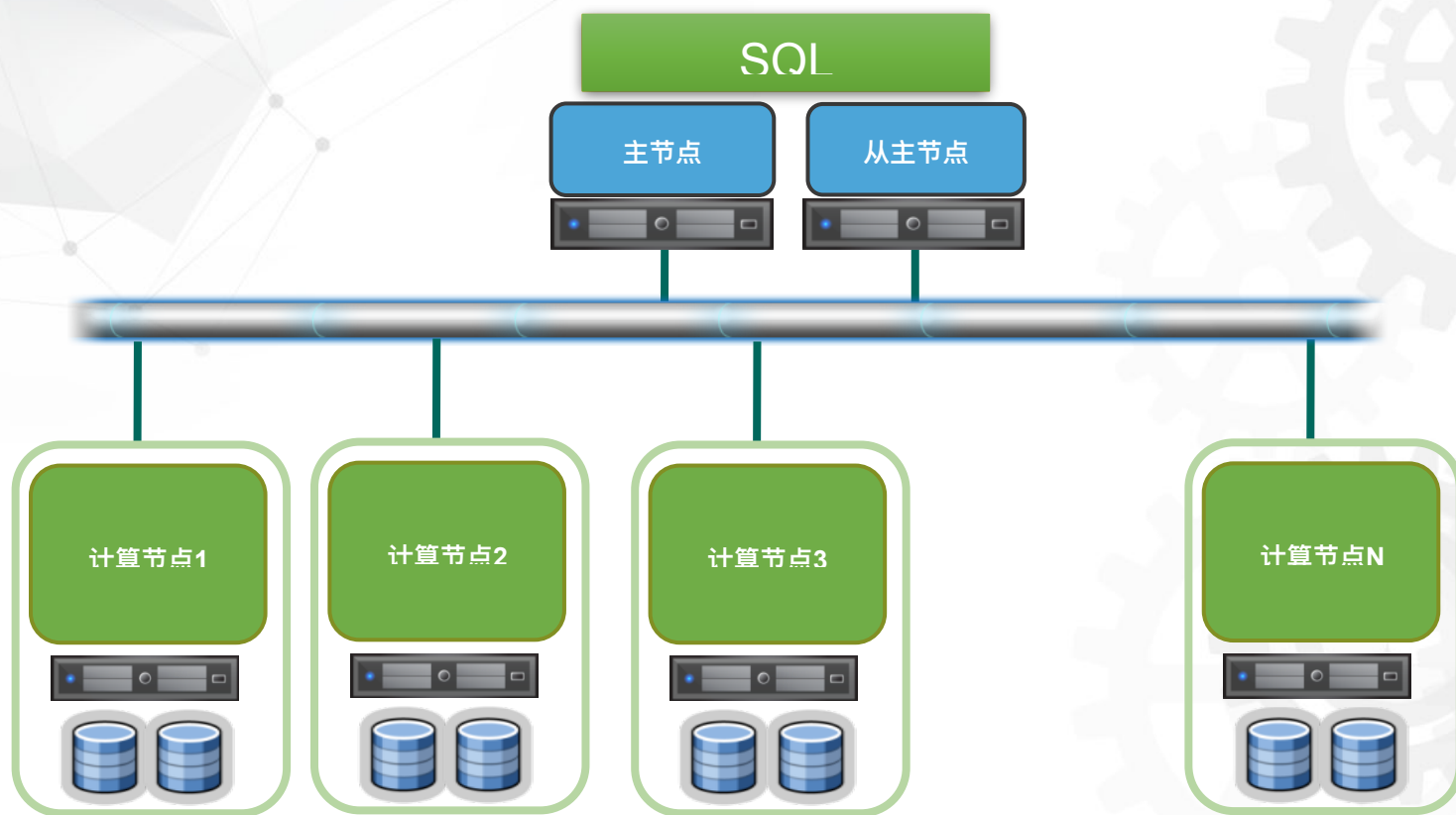


云和数据仓库

- 可行吗?
 - 当然！我们把Greenplum Database部署到阿里云上面。
- 它能很好地工作吗?
 - 动态伸缩?
 - 高可用?
 - 易用性?
 - 可扩展性?

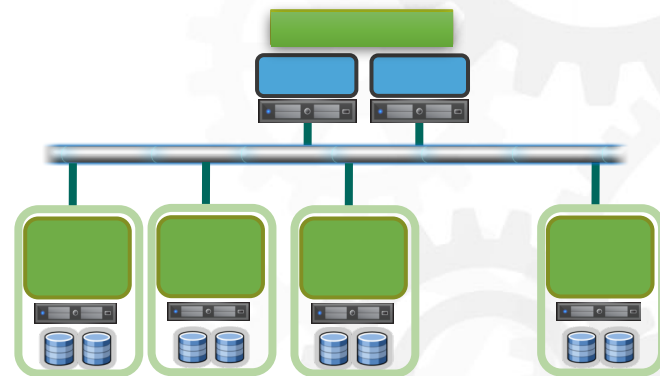


云和数据仓库



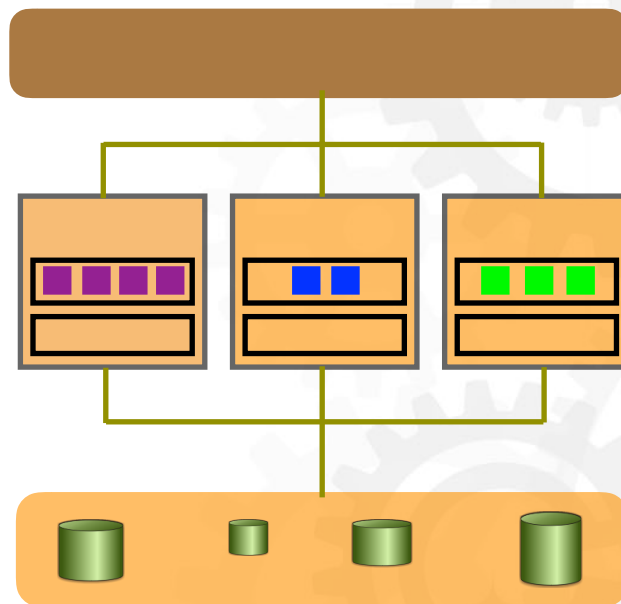
紧耦合的原罪

- 完全无共享MPP架构将计算和存储紧耦合。
- 多维度弹性？
 - 增删节点的时候需要对数据进行重分布
 - 系统可用性不高，扩展性不好
 - 资源浪费，违背pay-as-you-go的原则
 - 很难进行正确的资源配置
- 集群内资源管理非常复杂
 - 同构的资源 vs. 异构的工作负载
 - ETL，跑批生成报表，探索性即席查询

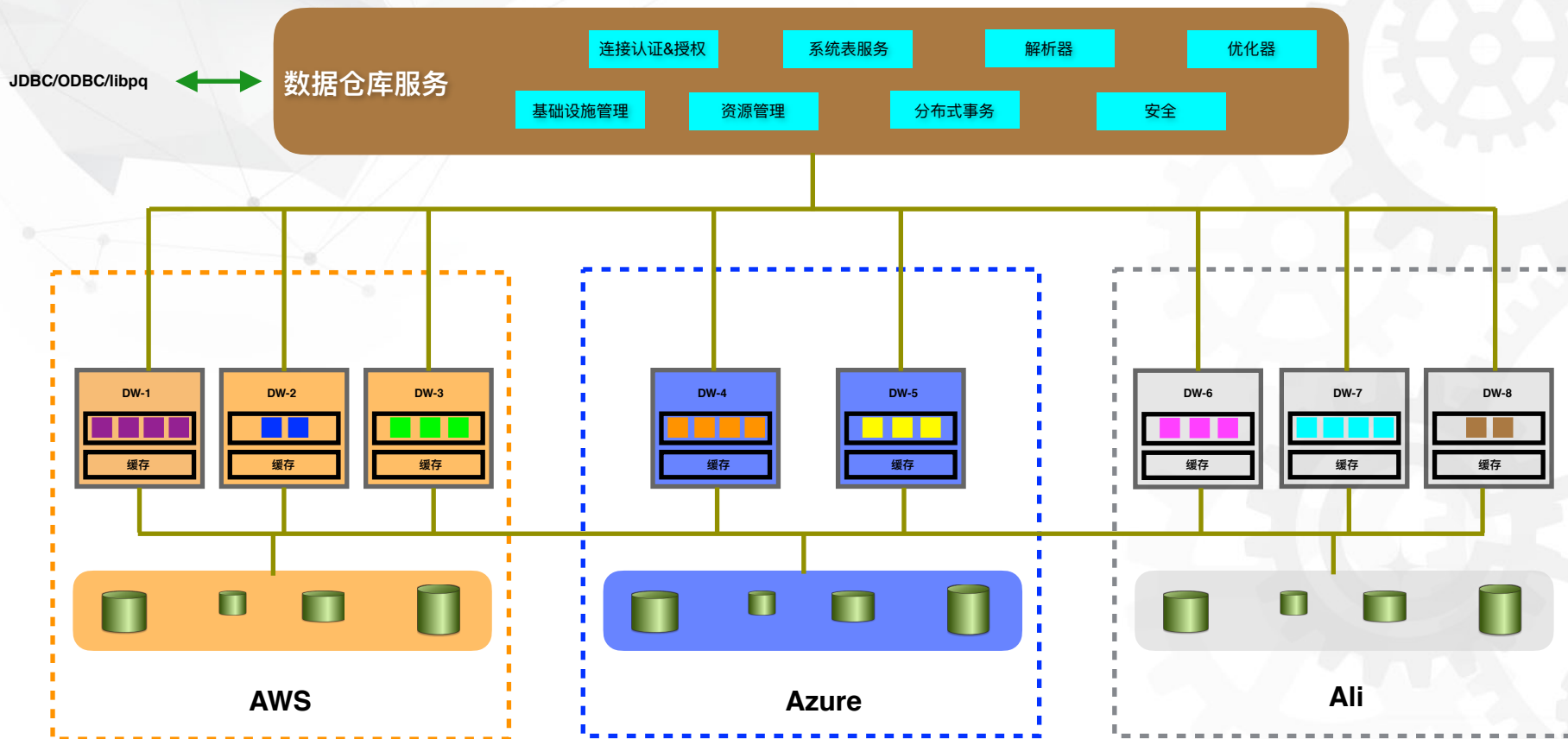


数据仓库服务

- 多维度弹性
 - 高并发
 - 几乎无限的计算和存储能力
 - 秒级的增删节点
- 统一的数据存储和处理平台
 - 结构化、半结构化和非结构化数据
 - 避免数据孤岛
- 真正的服务
 - 用户只需关心加载数据、分析数据
 - 基础设施管理、性能调优、高可用等交给数据仓库本身

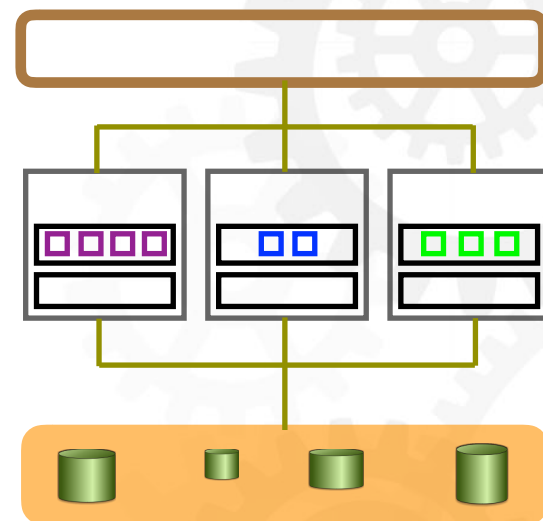


HashData架构



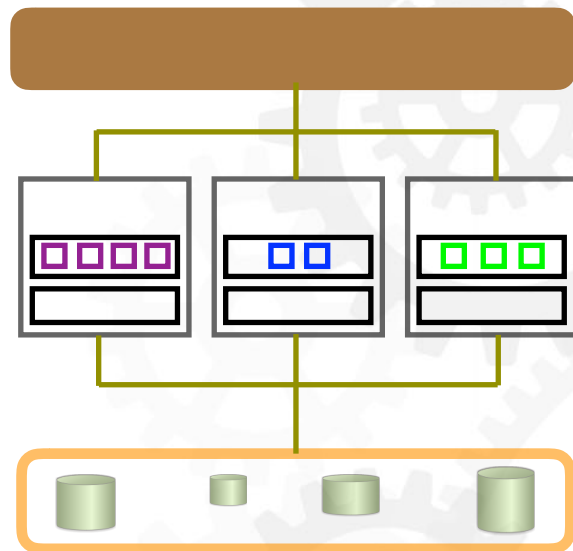
数据存储层

- 存储表数据和查询结果
- 对象存储
 - AWS S3, 青云QingStor, 阿里云OSS等等
 - 提供Restful访问接口
 - 高可用性, 高持久性
 - 按需付费
- 性能的负面影响
 - 直接访问2~3倍的性能下降 (通过本地缓存来降低性能惩罚)
 - 一旦写入不可更改 (精妙设计数据库增删改查操作)
- 数据存储格式和访问方式需要为底下的对象存储服务优化



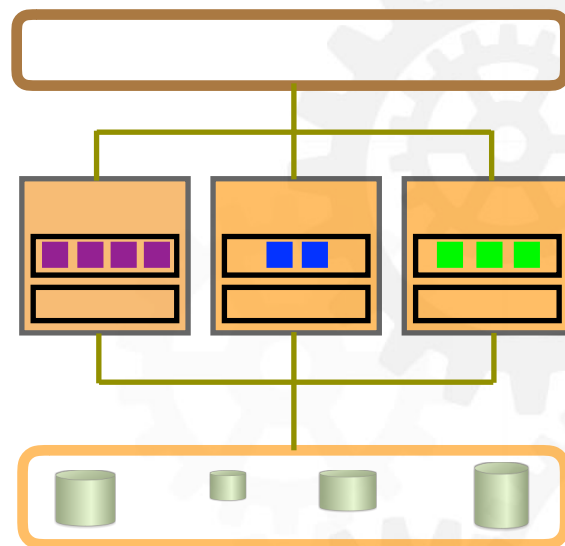
元数据服务

- 使用K-V数据库作为元数据存储层
 - 文件到表的映射关系
 - 表的统计信息（给优化器使用）、锁表、WAL日志等
- 云服务
 - 访问控制、查询优化器、分布式事务、锁管理
 - 标准访问接口，图形化管理界面和命令行管理工具

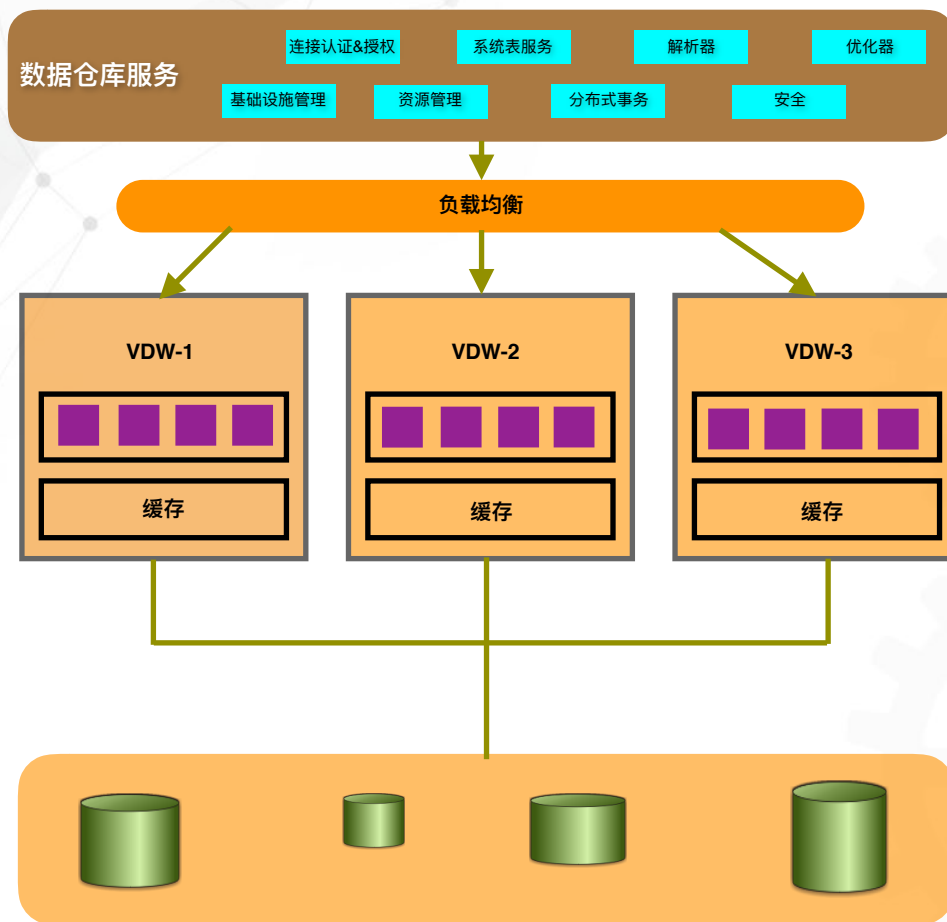


计算资源

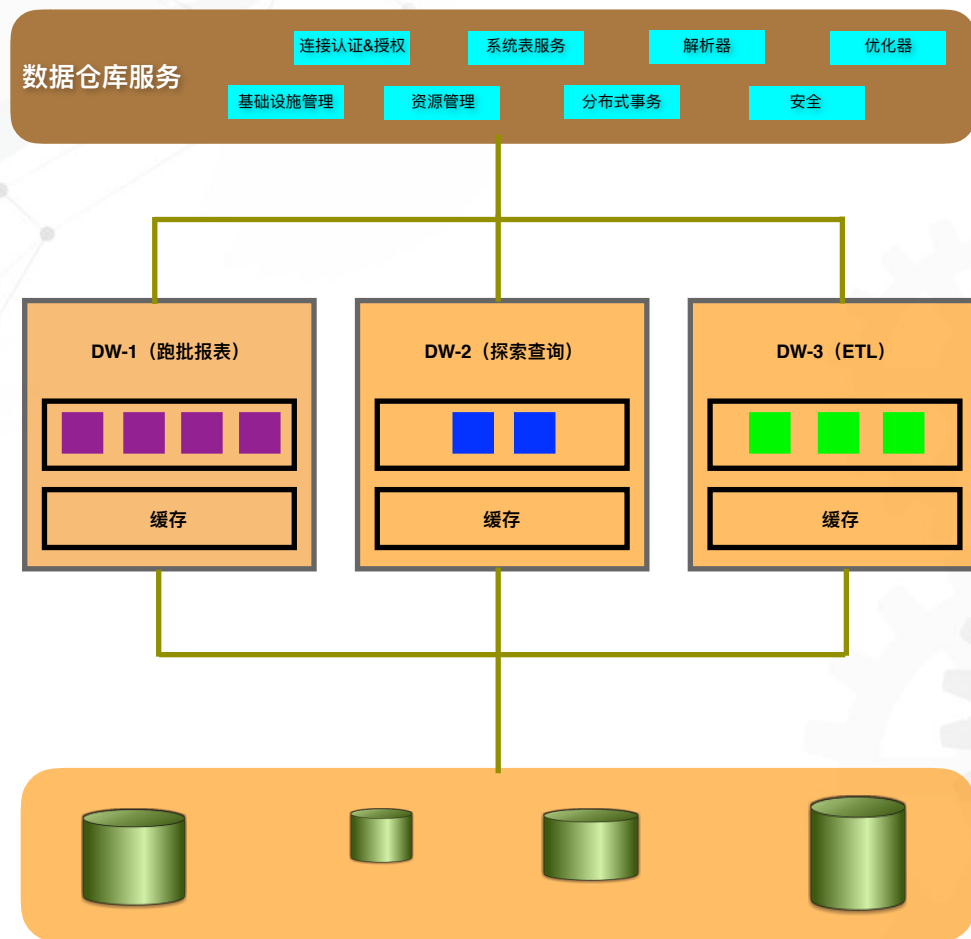
- 云平台提供的虚拟机
- 纯粹的计算资源
 - 按需创建、删除和纵向伸缩
 - 多台虚拟机组成一个数据仓库集群
 - 完全的性能隔离
 - 用户不跑查询的时候可以将所有的虚拟机关掉
- 每台虚拟机有本地的缓存进程
 - 同一个数据仓库集群所有节点的本地缓存进程组成分布式的缓存集群



多集群提供高并发

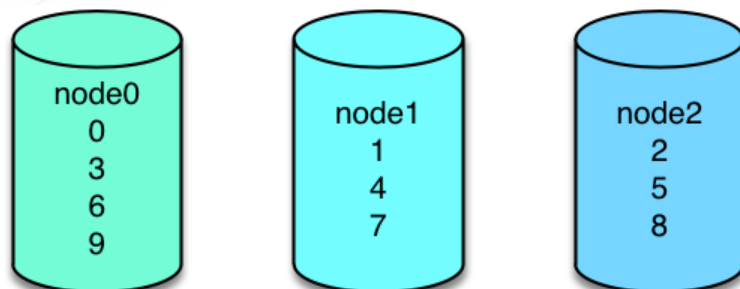


多集群应对异构负载

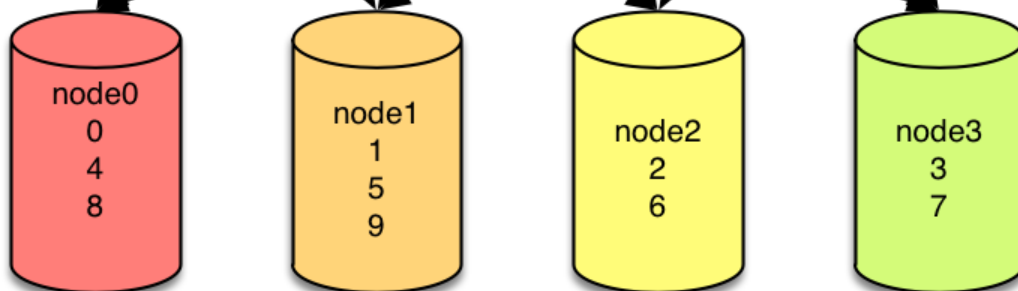


数据分布策略——哈希取模

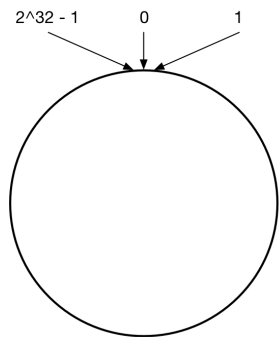
系统扩展前



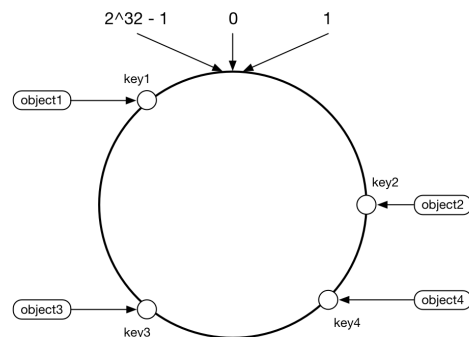
系统扩展后



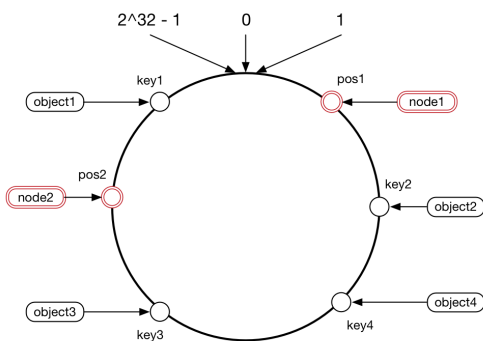
数据分布策略——一致性哈希



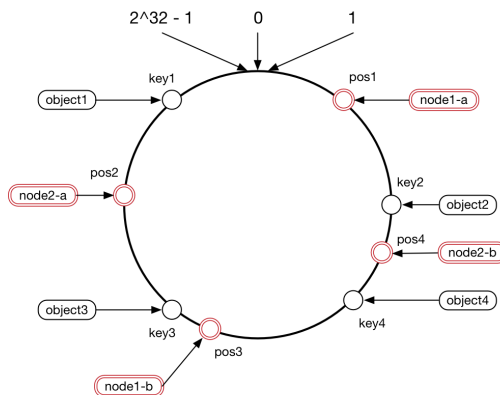
(1)



(2)

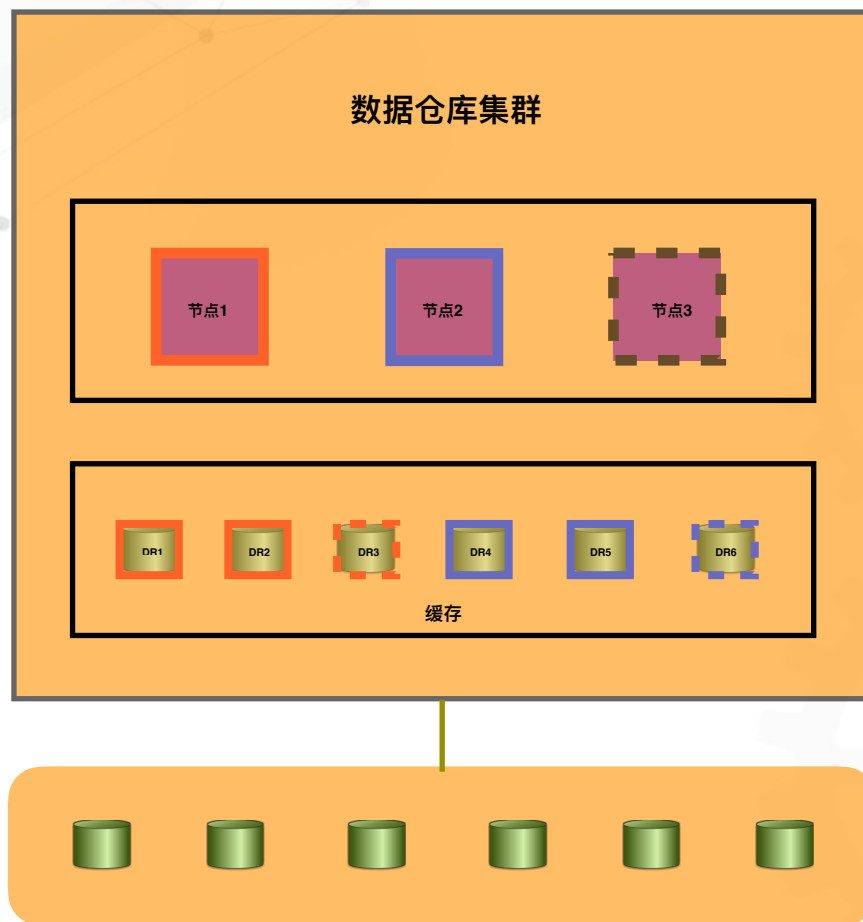


(3)



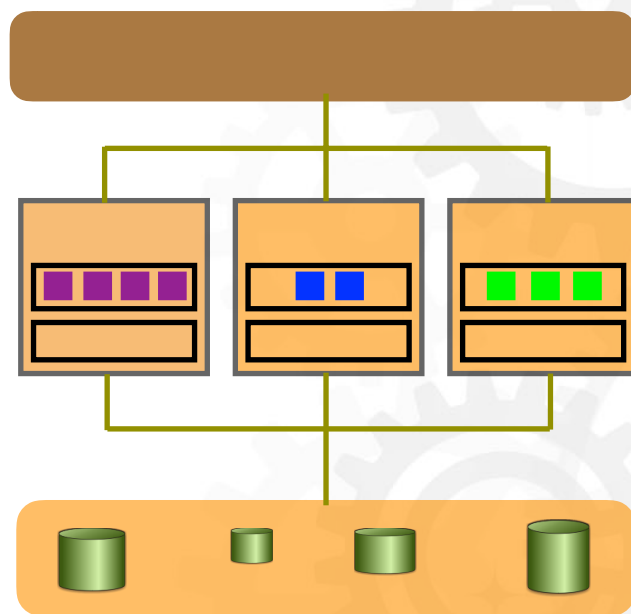
(4)

秒级扩容



其它功能特性

- DBA? NO!!
 - 按需启动数据仓库集群，不要DBA的介入
- 在线升级和扩缩容
 - 独立的元数据服务
- 时间旅行
 - 对象存储服务提供几乎无限的存储容量
 - 数据多版本控制
- 抛弃传统的资源队列
 - 根据不同的工作负载启动不同配置的集群
- 统一的数据平台
 - 开发、测试和生产环境能够完全一致



总结

- 数据仓库已死?
 - 数据仓库需要的不是退出历史舞台，而是根本性的改变。
- 元数据、计算和存储三者分离是关键
 - 多维度弹性
 - 避免数据孤岛
 - pay-as-you-go的商业模式
- 对象存储将替代HDFS
 - 高可用
 - 高并发
 - 扩展性





www.hashdata.cn

We're Hiring!

THANKS

SequeMedia
盛拓传媒

IT168.com

ITPUB

ChinaUnix.net