

微服务下分布式数据库架构的演进实践

SequoiaDB North America Research Lab

巨杉数据库北美研发实验室

Danny Chen



个人经历

- Danny Chen
 - 超过10年 IBM DB2 数据库引擎开发
 - 华为数据库专家组成员
 - 巨杉数据库北美研究所主任



议题

- 云架构与微服务
- 分布式数据库的发展
- 为微服务构建数据平台
- 巨杉数据平台的实践



云架构

- 共享与抽象（虚拟化）
 - 通过平台来降低成本
 - 加快部署周期
 - 通过接口避免重复工作
- 使用服务而专注业务应用开发本身
 - IaaS, PaaS, SaaS
 - 公有云，私有云，混合云，多云

微服务

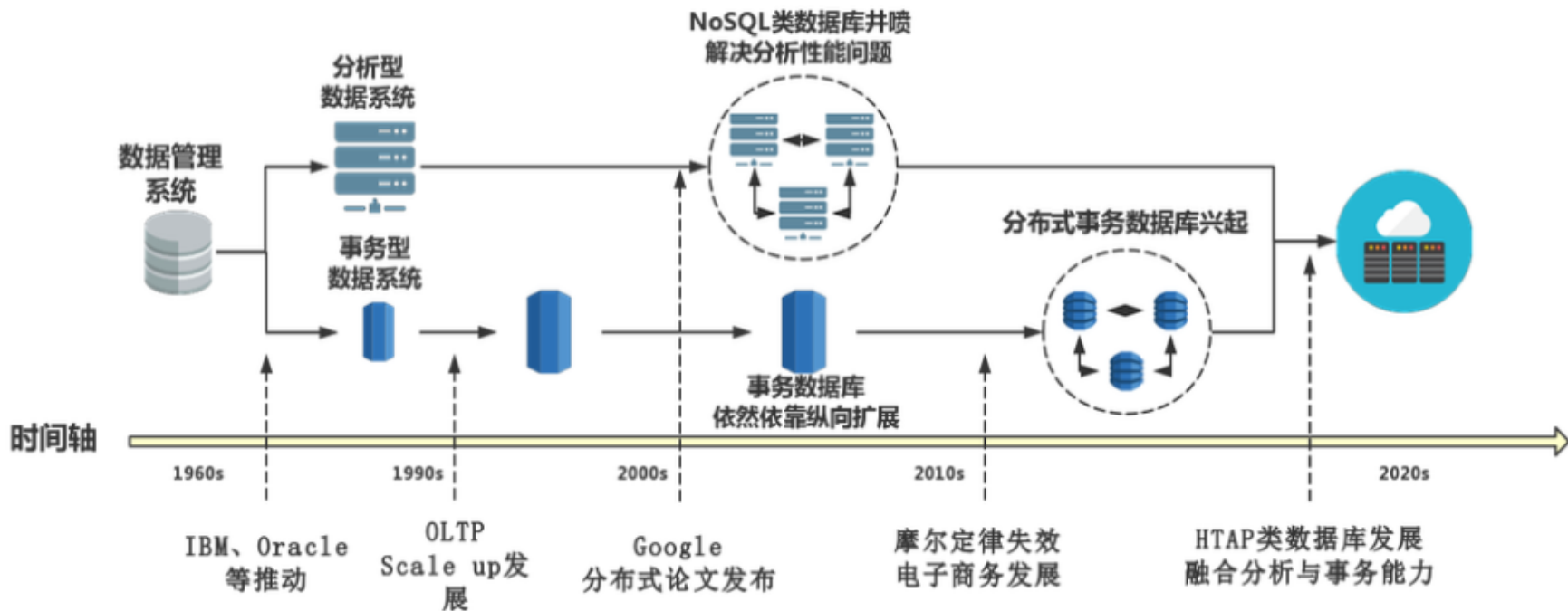
- 设计
 - 将大应用的拆分成多个小应用
 - 功能上轻量独立
 - 小应用间通过服务接口交互集成(使用外部服务并提供对外服务)
- 高敏捷的部署，弹性伸缩
- 资源隔离，安全可靠
- 典型案例： **NFV**

议题

- 云架构与微服务
- 分布式数据库的发展
- 为微服务构建数据平台
- 巨杉数据平台的实践

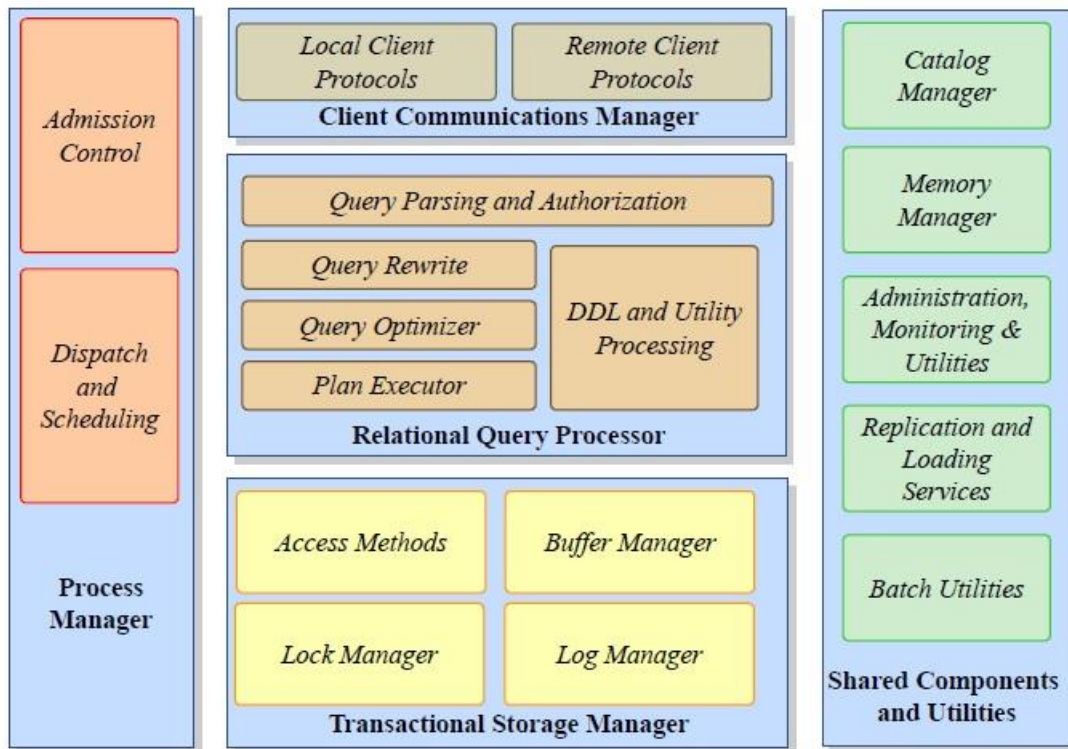


分布式数据库的进化史

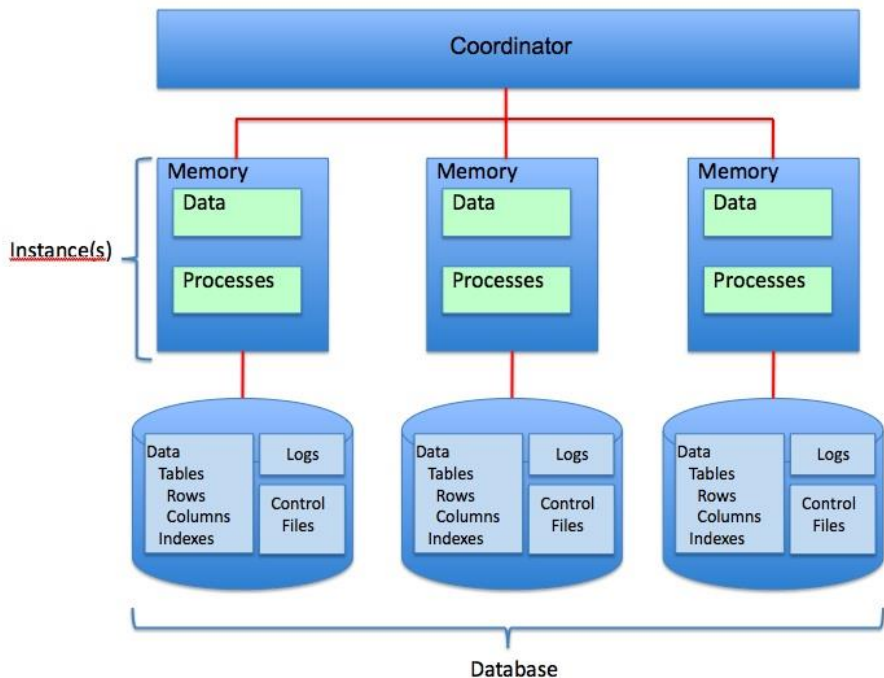


参考GFS和bigtable

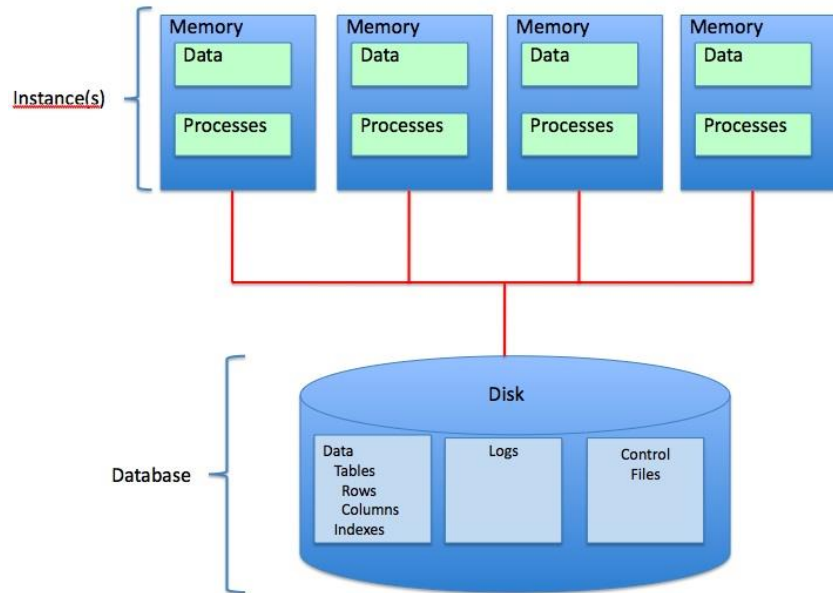
传统数据库的关键组件



分布式数据库的构架



shared nothing (sharding) architecture



shared everything architecture

议题

- 云架构与微服务
- 分布式数据库的发展
- 为微服务构建数据平台
- 巨杉数据平台的实践



数据服务平台vs传统数据库

- 共同：持久，高效，安全，一致
- 存储：结构化/半结构化/非结构化 vs 结构化
- 访问：schema-on-read vs schema-on-write
- 成本：低价通用硬件/云平台 vs 专用硬件
- 应用：灵活使用数据

为微服务构建的数据平台

- 自身设计
 - 内部组件遵从微服务架构设计原则
 - 适合于云部署，通过通用接口摆脱依赖性
 - 对多种应用的兼容（用户习惯与生态）
- 构架
 - 存储和计算分离
 - 支持数据多模（格式，处理，组织）
 - 完善的监控机制甚至自维护

为微服务构建的数据平台

- 数据平台与上层应用的对接
 - 适配不同现有应用的接口（驱动，连接器
 - 对接微应用，提供数据层面的持久化，弹性扩张能力
 - 抽象数据层，屏蔽底层，对上解耦
- 构架
 - 数据的安全与持久
 - 注重动态弹性扩张能力
 - 没有单点故障
 - 灵活对接不同类型应用

议题

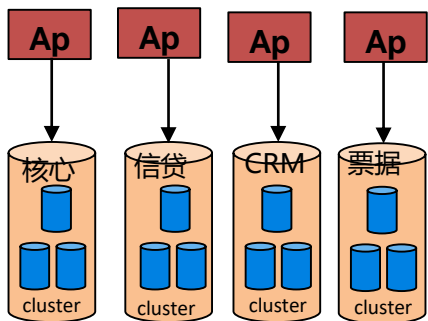
- 云架构与微服务
- 分布式数据库的发展
- 为微服务构建数据平台
- 巨杉数据平台的实践



对比其他的微服务支持方式

应用垂直分库

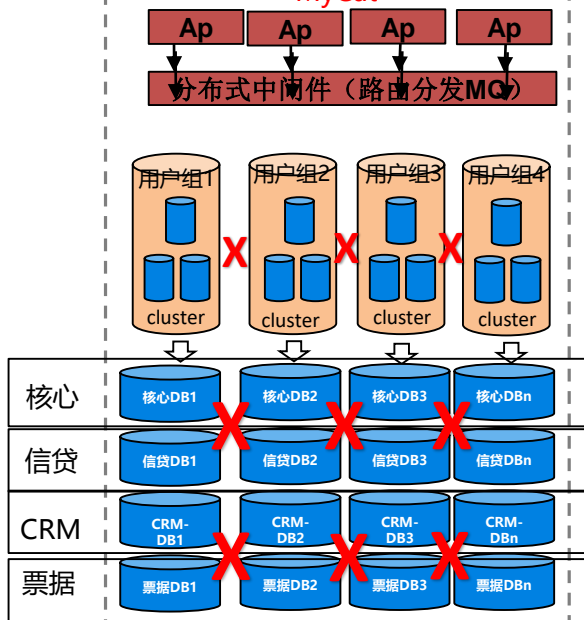
TDDL



将不同模块的数据表分库存储，库间不相互关联查询，如果有，必须通过数据冗余或在应用层二次加工来解决，对应用程序侵入较大。

分库分表

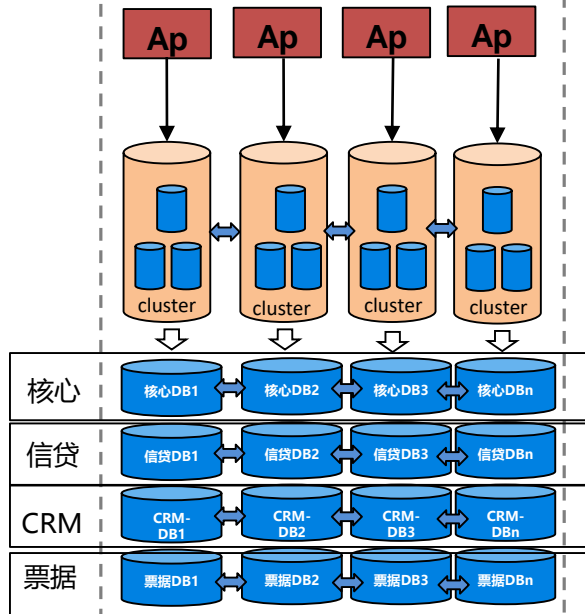
MyCat



通过分布式中间件进行用户ID的路由分发，保证用户的一类操作涉及的表在一个节点上完成，避免分布式事务。如果有跨节点的事务，则通过中间件保证其一一致性。

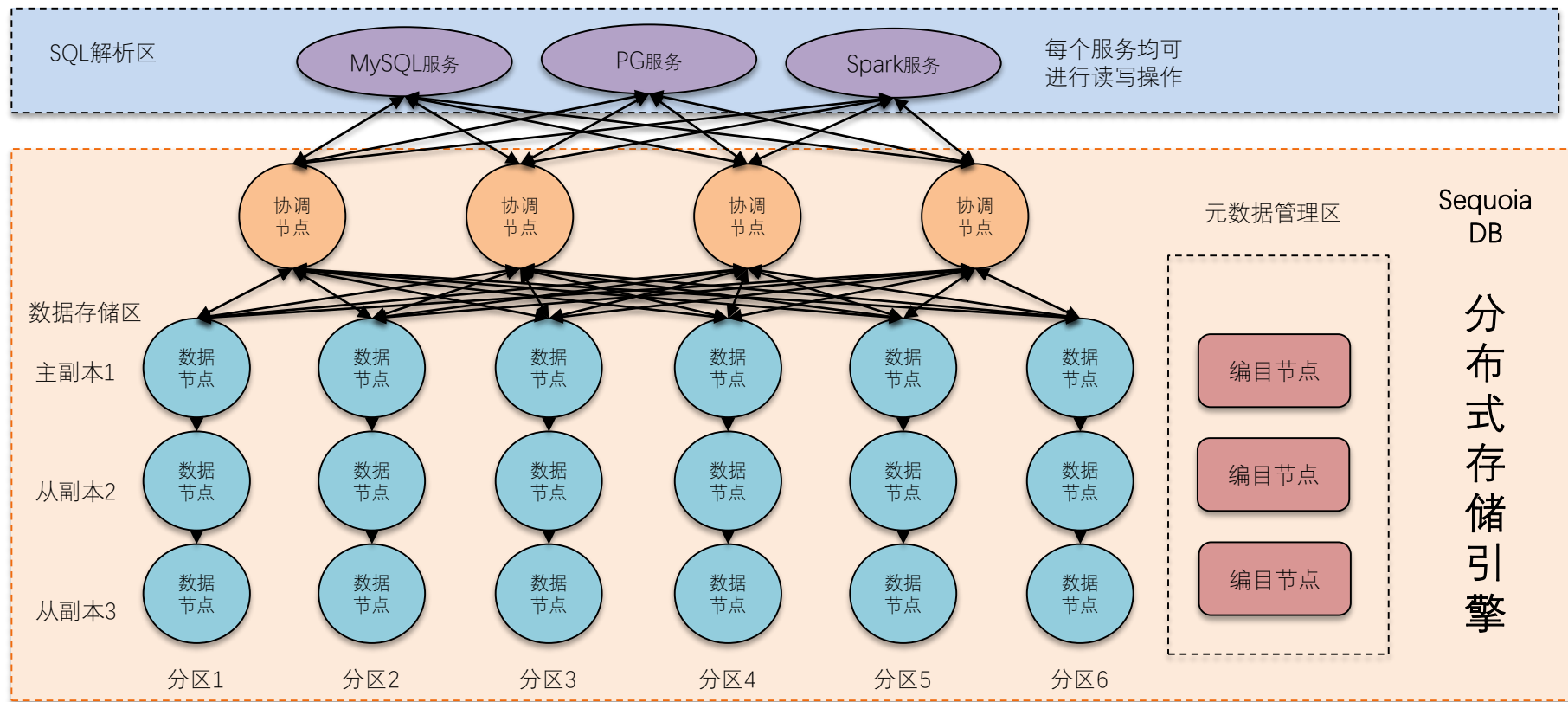
原生分布式数据平台

SequoiaDB



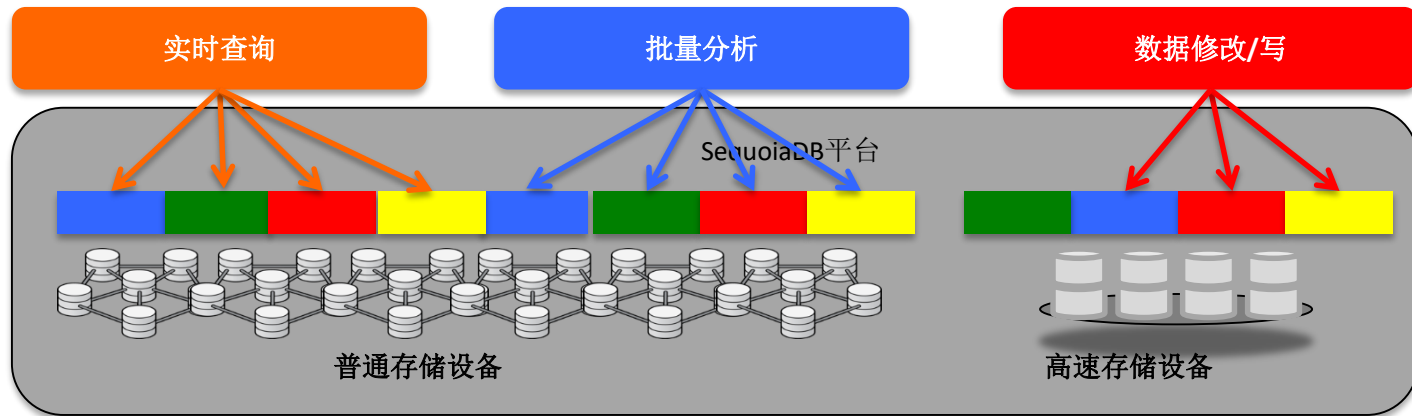
将表分布到不同机器的库上，减轻数据库的压力物理机的CPU、内存、网络IO负载分摊。支持分布式事务。

巨杉数据平台 “计算存储分离” 分布式架构

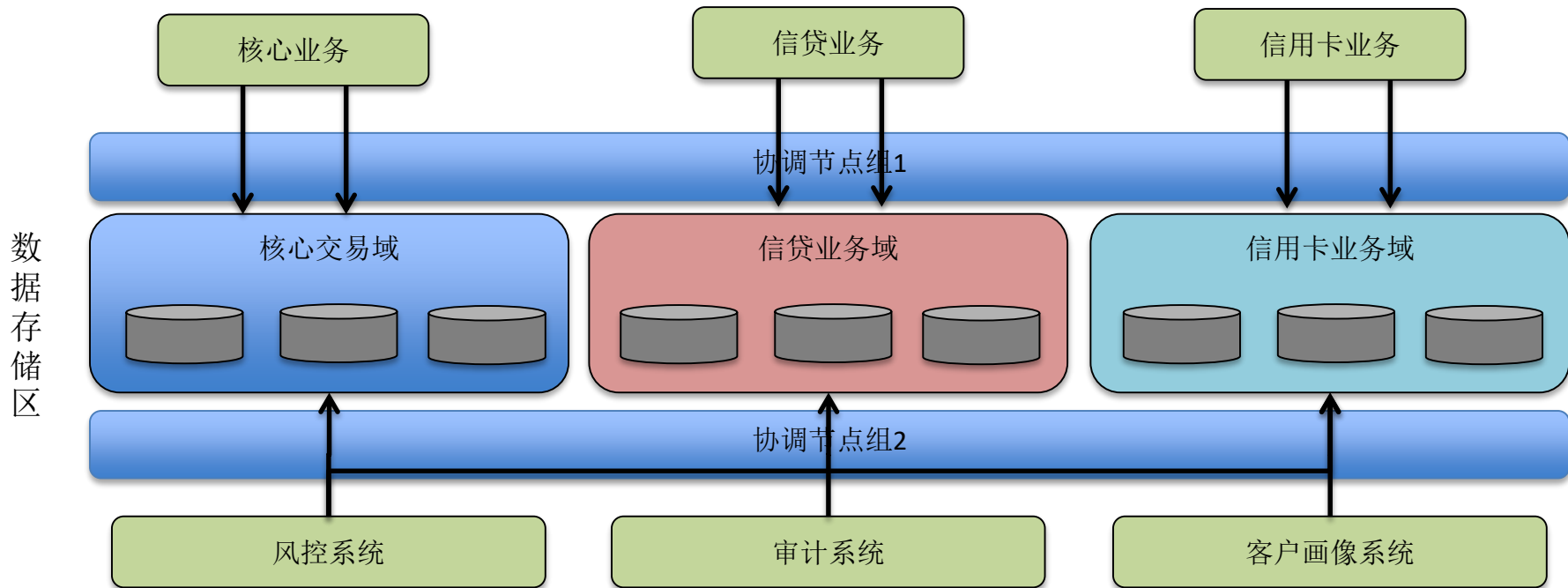


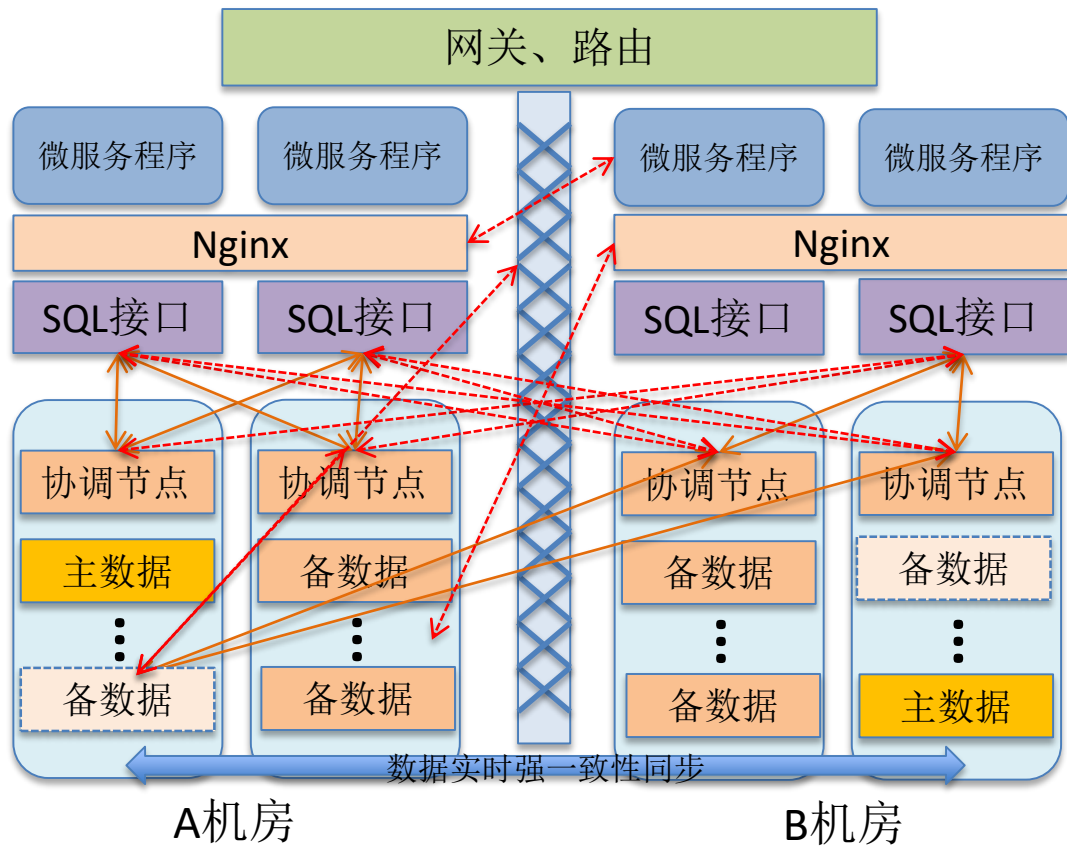
分布式架构优化：读写分离机制

- 数据在多个分布节点内自动复制，并实现写请求和读请求的自动分离，避免读请求对数据写入的影响。
- 此外，可进一步定制数据分布策略，保证不同类型业务可以运行在同一平台上，但同时又不会互相干扰，比如：
 - 冷/热数据区分离
 - 写交易的“强一致性”和“弱一致性”分离
 - 查询/批量分离



分布式架构优化：数据域逻辑与物理隔离

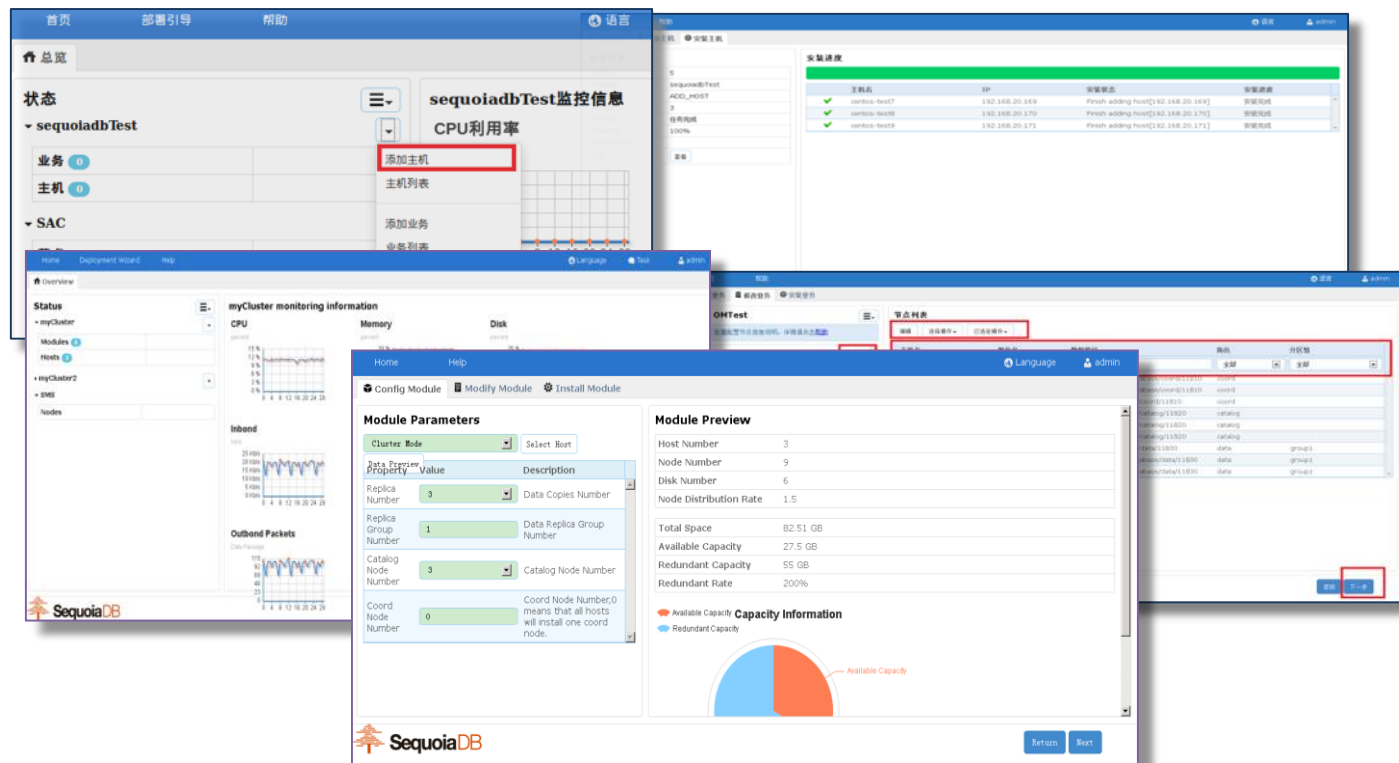




交叉访问原则

- 微服务程序优先连接本机房Nginx服务，只有当本机房Nginx无法提供服务后，再去连接同城机房Nginx服务
- Nginx服务只连接本机房MySQL服务，一个机房拥有多个MySQL服务，确保SQL服务高可用
- 如果一个机房所有MySQL服务均停止服务，则该机房Nginx服务也会停止，微服务程序自动选择同城机房Nginx服务进行连接
- MySQL优先连接本机房协调节点，避免请求在同城机房中交叉访问
- 协调节点在执行写入操作时，自动路由数据库分区的主节点，执行操作
- 协调节点在执行查询操作时，优先选择本机房数据节点进行访问，避免请求在同城机房中交叉访问

平台统一管理监控方案



➤ SequoiaDB提供全面的图形化集群监控功能

- 集群健康状态
- 资源使用状态

➤ SequoiaDB提供图形化操作界面

- 集群部署、扩容
- 集群管理
- 数据CURD操作
- 数据管理，切分、均衡等

结束语

- 巨杉吸引了众多国内外数据库领域的老司机
- 我们倾听客户的声音，拥抱社区
- 我们坚持自主研发，自己掌控未来



谢谢!
Q&A

SequoiaDB Website:
www.sequoiadb.com

Github:
[SequoiaDB/SequoiaDB](#)
[SequoiaDB/sequoiasql-mysql](#)



Join SequoiaDB Community
加入SequoiaDB社区!