

## 5、云数据库

- 系统架构（以阿里UMP为例）

- 设计原则

- 包含的组件

- Mnesia

- 一个分布式数据库管理系统
        - 支持事务、支持透明的数据分片
        - 运行于controller之上

- RabbitMQ

- LVS

简单理解就是多个服务器组合起来变成一个超级服务器，由一个代理服务器统一对外提供访问接口

- ZooKeeper

ZooKeeper是高效和可靠的协同工作系统，提供分布式锁之类的基本服务（比如统一命名服务、状态同步服务、集群管理、分布式应用配置项的管理等）

- 在UMP中作为全局的配置服务器
        - 提供分布式锁（在多个备选controller中选一个当真正的管家）
        - 监控所有的MySQL实例（是否正常运行）

- 包含的角色

- Controller服务器

Controller服务器向UMP集群提供各种管理服务，实现集群成员管理、元数据存储、MySQL实例管理、故障恢复、备份、迁移、扩容等功能。Controller服务器上运行了一组Mnesia分布式数据库服务，其中存储了各种系统元数据，主要包括集群成员、用户的配置和状态信息，以及用户名到后端MySQL实例地址的映射关系（或称为路由表）。当其他服务器组件需要获取用户数据时，可以向controller服务器发送请求获取数据。为了避免单点故障，UMP系统中部署了多态controller服务器，由zookeeper的分布式锁功能来帮助选出一个总管，负责各种系统任务的调度和监控。

- Proxy服务器

向用户提供访问MySQL数据库的服务，它完全实现了MySQL协议，用户可以用已有的MySQL客户端连接到Proxy服务器，Proxy服务器通过用户名获取到用户的认证信息，资源配额的限制（例如qps、IOPS、最大连接数），以及后台MySQL实例的地址，然后，用户的SQL查询请求会被转发到相应的MySQL实例上。除了数据路由的基本功能外，Proxy服务器中还实现了很多的功能，主要包括屏蔽MySQL实例故障、读写分离、分库分表、资源隔离、记录用户访问日志等。

- Agent服务器

部署在运行MySQL进程的机器上，用来管理每台物理机上的MySQL实例，执行主从切换、创建、删除、备份、迁移等操作

- Web控制台

向用户提供系统管理界面

- 日志分析服务器

- 信息统计服务器

- 愚公系统

是一个全量复制结合bin-log分析进行增量复制的工具，可以实现不停机的情况下动态扩容、缩容和迁移。

- 系统功能

- 容灾

为了实现容灾，UMP为每个用户创建两个MySQL实例，一个是主库，一个是从库。主从切换的状态是由zookeeper负责维护的。

- 主从切换过程

- 1、zookeeper监测到主库故障，通知controller服务器
    - 2、controller服务器启动主从切换时，会修改“路由表”，即用户名到后端MySQL实例地址的映射关系
    - 3、把主库标记为不可用
    - 4、借助于RabbitMQ通知所有Proxy服务器修改用户名到后端MySQL实例地址的映射关系

- 读写分离

充分利用主从库实现用户读写操作的分离、实现负载均衡，UMP系统实现了对于用户透明的读写分离功能，当整个功能被开启时，负责向用户提供访问MySQL数据库的Proxy服务器，就会对用户发起的SQL语句进行解析，如果属于写操作，就直接发送到主库，如果是读操作，就会被均衡地发送到主库和从库上执行。

- 分库分表

- 1、proxy服务器重写用户sql语句，得到多个针对相应mysql实例的子语句
  - 2、把子语句分发到对应的mysql实例上执行
  - 3、接收来自多个mysql实例的sql语句执行结果，合并得到最终结果

- 资源调度

- 资源管理

- 资源隔离

- 数据安全