构建高可用分布式 Key-Value Store

使用LedisDB + xcodis + redis-failover

Agenda

1. What：Key-Value Store简介
2. Why：再造Key-Value Store的理由
3. How：如何构建一个Key-Value Store

关于我

1，WPS资深程序员
2，热爱开源
3，Go重度用户
4，LedisDB，xcodis等作者

What

什么是Key-Value Store

Use the associative array as fundamental data model.

现有的栗子

1. Memcached
2. MySQL
3. Redis
4. MongoDB, Couchbase, 特多了！

Why?

Why？

1. 程序员天生的造轮子情结
2. 不同NoSQL选型学习成本
3. 基础服务的可控性
4. 能学到太多的东西

How?

Feature

1. Language
2. Protocol
3. API
4. High Availability
5. Cluster

Let’s Go!

如果用C++写一个服务？

1. 要学习socket API
2. 了解epoll，kqueue或者IOCP
3. 异步逻辑导致的代码割裂
4. 异步代码对象生存期管理
5. 开发速度慢

Go

1. 天生的服务端并发编程语言
2. 官方自带高性能net package
3. goroutine解决异步编程难题
4. 开发快速，个人感觉 > python

Protocol

Protocol

1. 易于实现
2. 快速解析
3. 可读性好

Redis Serialization Protocol

- Simple String: “+OK\r\n”
- Error: “-Error message\r\n”
- Integer: “:123\r\n”
- Bulk String: “$6\r\nfoobar\r\n”
- Array: “*2\r\n$3\r\nfoo\r\n$3\r\nbar\r\n”

API

API

- String: SET, GET, etc.
- Hash: HSET, HGET, etc.
- List: LPUSH, LPOP, etc.
- SET: SADD, SISMEMBER, etc.
- ZSET: ZADD, ZRANGE, etc.

这货不是Redis？

LedisDB

1. 使用Redis协议
2. 类Redis API
3. 提供KV, Hash, List, Set, ZSet数据结构支持
4. 底层基于RocksDB等，超越内存限制

LedisDB

```
//start ledis server
ledis-server
```

```
//another shell
ledis-cli -p 6380
ledis> set a 1
OK
ledis> get a
“1"
```

High Availability

High Availability

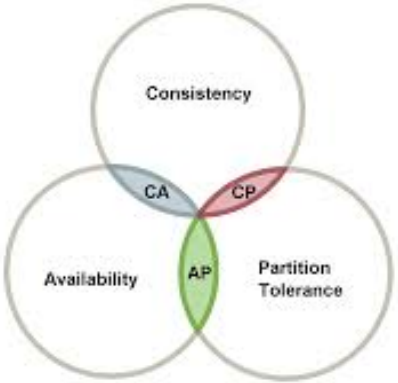
1. CAP
2. Data Security
3. Fault Tolerance

CAP

C: Consistency

A: Availability

P: Partition tolerance



Data Security

不要把鸡蛋放到一个篮子里！

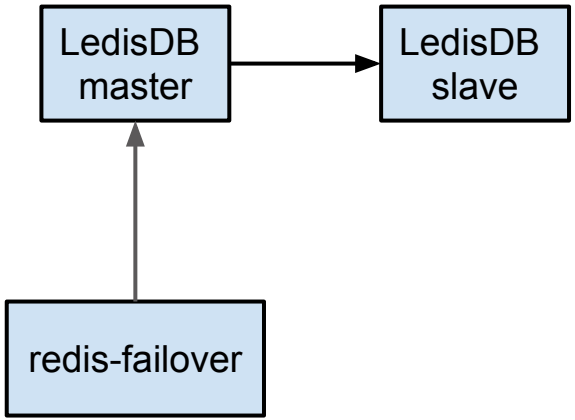
不要把数据放到一台机器上！

Data Security

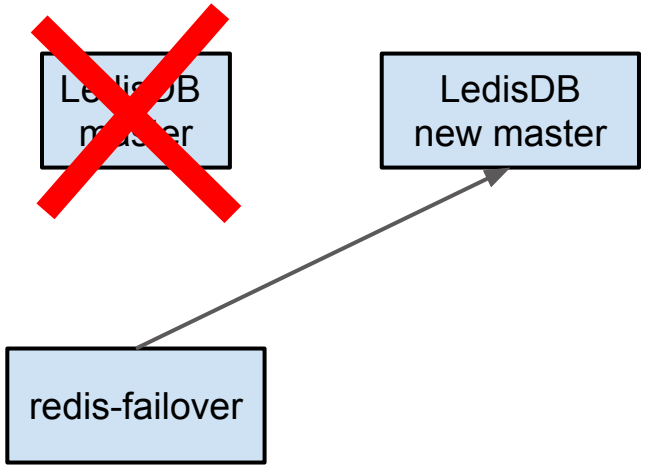
1. Backup
2. BinLog
3. Replication Topology
    a. Synchronous
    b. Asynchronous
    c. Semi-synchronous

Fault Tolerance

Monitor



Failover



Cluster

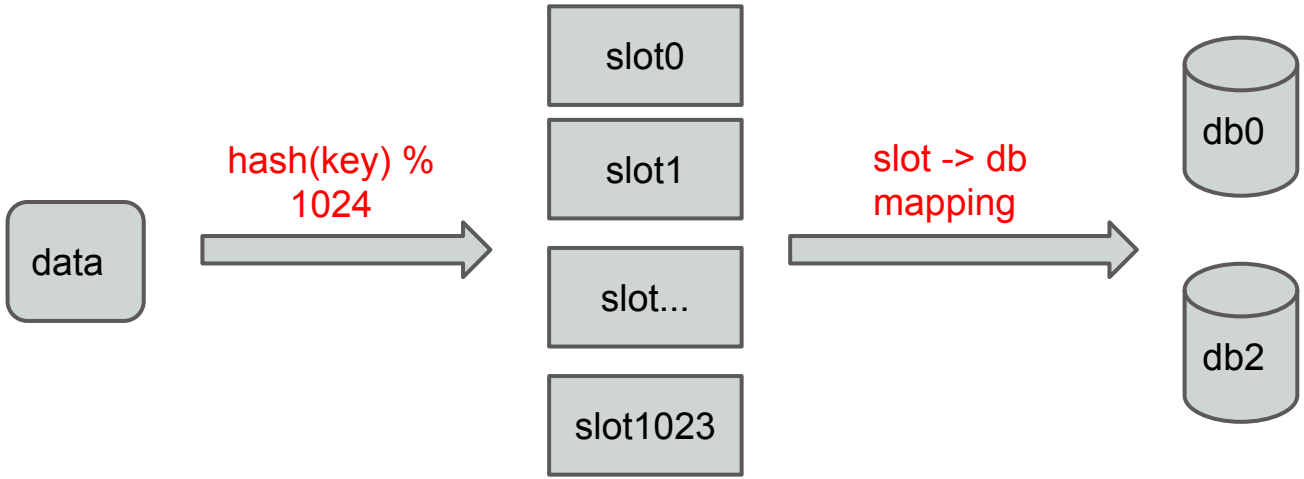
Cluster

1. 如何进行数据切分以及查找
2. 如何动态扩容
3. 如何对外提供服务
4. 如何协调多个服务

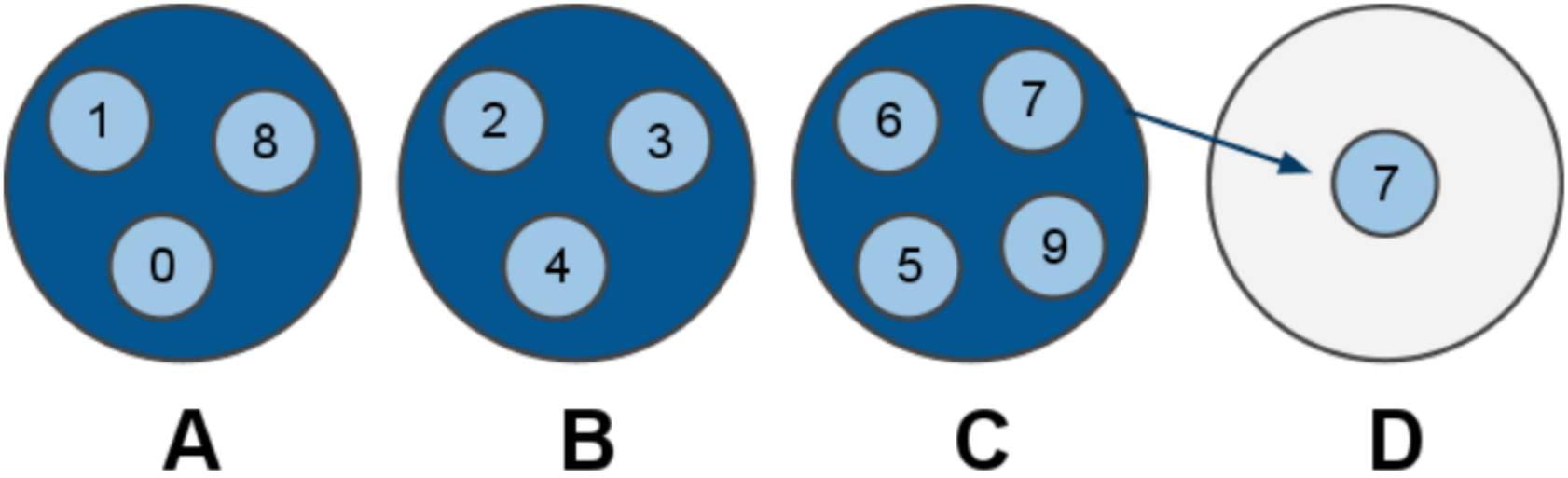
Key Routing

1. Hash
2. Consistent Hash
3. Routing Table

Hash + Routing Table



Resharding



Usage

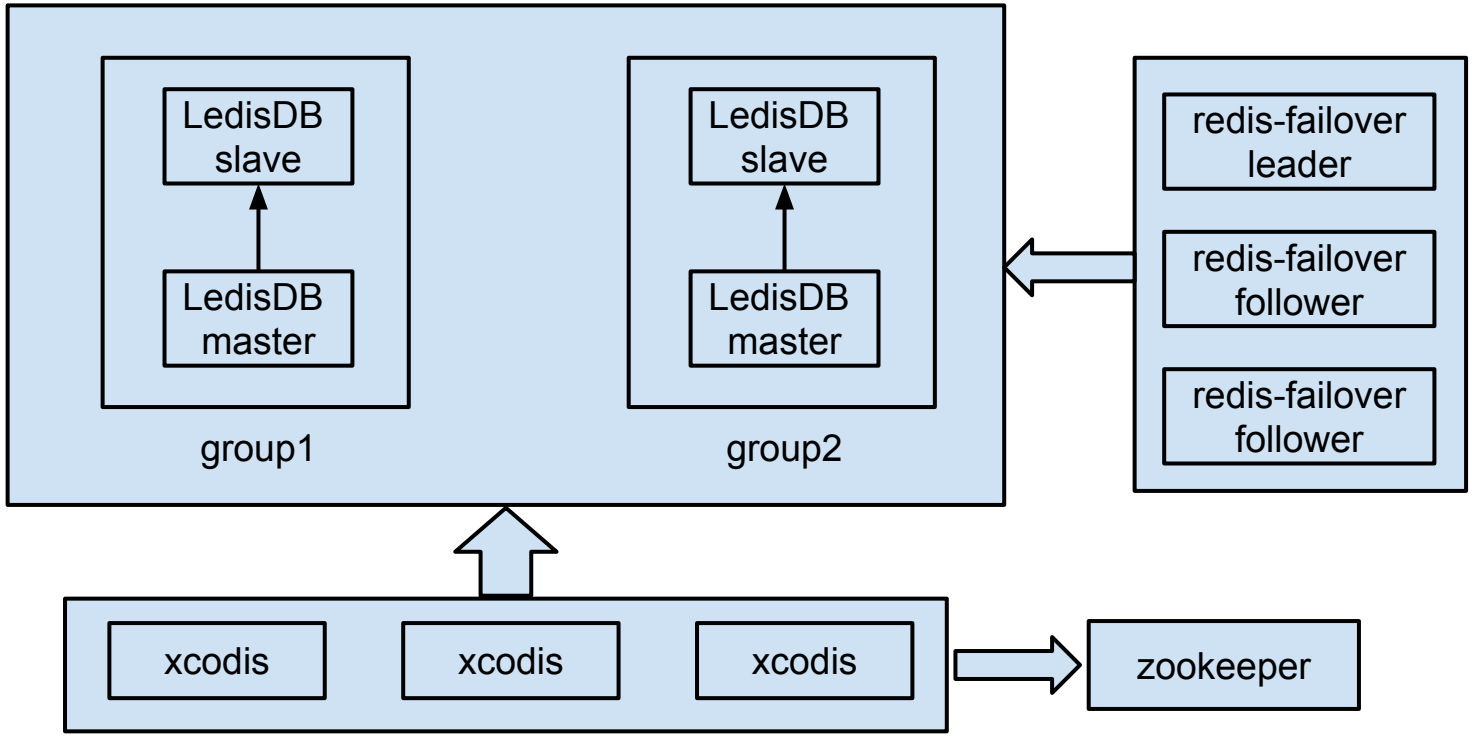
1. 类redis cluster
2. 自定义客户端
3. proxy

Coordination

zookeeper

etcd

Final Architecture



Question?

Thank you!

siddontang@gmail.com
github.com/siddontang
twitter.com/siddontang