



小米MySQL高可用架构演进

小米DBA 张良

全球 2017-11-24

邮箱：zhangliang3@xiaomi.com

大纲

1

历史回顾

2

数据库中间件

3

DBaaS

4

故障自愈

历史回顾

历史回顾

服务

MySQL

Redis

MongoDB



用户

MIUI

云服务

米家智能设备



规模

集群规模 900+

日PV 1000亿次+

数据量 90T+



演进阶段

2017

容器化

2016

接入层

2013

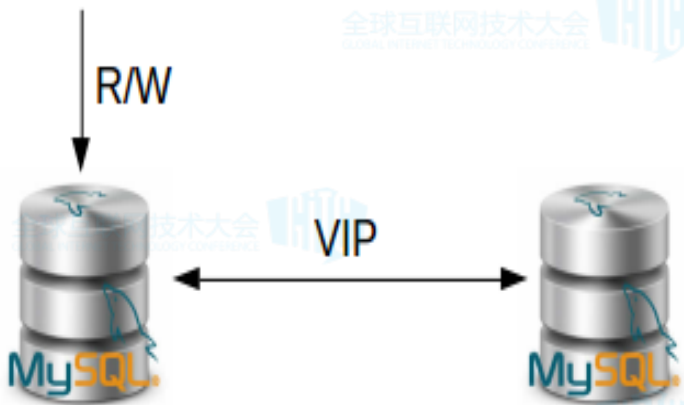
一主多从
+
Sharding

2012

主备



一个人在战斗



数据量较小

流量不高

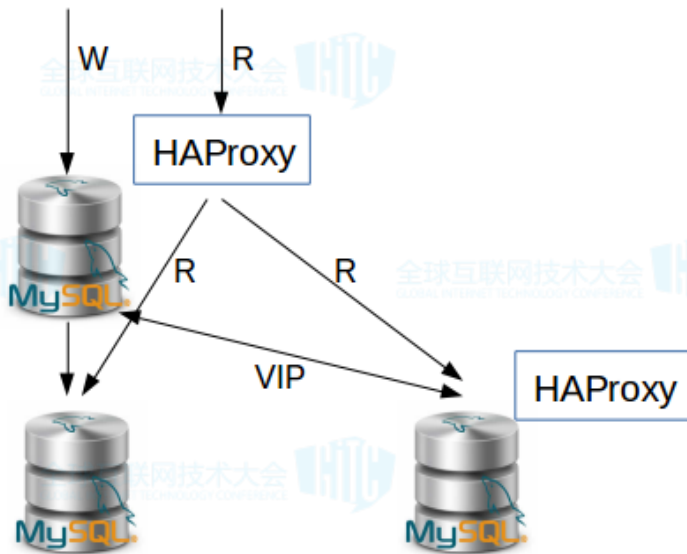
高度合并部署

服务无多机房冗余

一主多从

只读请求负载均衡

从库可扩展



Sharding

全球互联网技术大会
GLOBAL INTERNET TECHNOLOGY CONFERENCE

全球互联网技术大会
GLOBAL INTERNET TECHNOLOGY CONFERENCE

全球互联网技术大会
GLOBAL INTERNET TECHNOLOGY CONFERENCE

0

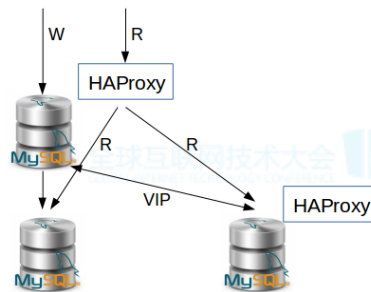
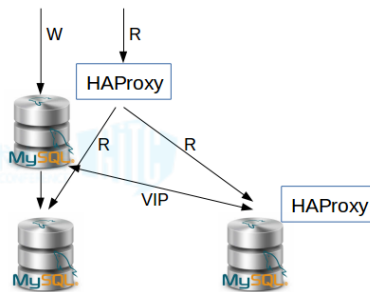
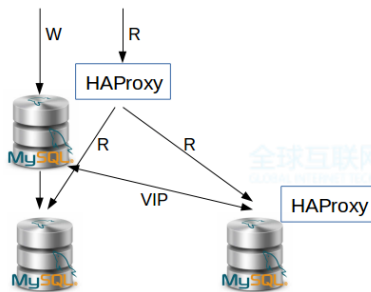
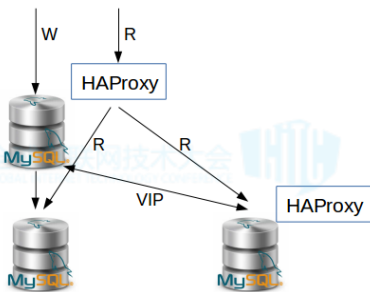
1

2

3

全球互联网技术大会
GLOBAL INTERNET TECHNOLOGY CONFERENCE

全球互联网技术大会
GLOBAL INTERNET TECHNOLOGY CONFERENCE



全球互联网技术大会
GLOBAL INTERNET TECHNOLOGY CONFERENCE

全球互联网技术大会
GLOBAL INTERNET TECHNOLOGY CONFERENCE

▶ 数据横向可扩展

▶ 解决单机容量瓶颈

全球互联网技术大会
GLOBAL INTERNET TECHNOLOGY CONFERENCE

全球互联网技术大会
GLOBAL INTERNET TECHNOLOGY CONFERENCE

Sharding

```
<!-- this configuration is used for management -->
<dbpartition>
  <!-- Here we define all connections -->
  <connections>
    <connection name="master"
      url="jdbc:mysql://10.xxx.xxx.xxx/database?useUnicode=true&characterEncoding=utf-8"
      usr="*****"
      pwd="*****"
      role="master" />
    <connection name="slave"
      url="jdbc:mysql://10.xxx.xxx.xxx/database?useUnicode=true&characterEncoding=utf-8"
      usr="*****"
      pwd="*****"
      role="slave" />
  </connections>

  <!-- Define how tables are partitioned. -->
  <tables>
    <table type="table" partition_column="id" partition_type="number" format="table_%d" >
      <partition min="0" max="9" range="1" database_name="database_0" connection_ref="master;
slave" />
      <partition min="10" max="19" range="1" database_name="database_1" connection_ref="master;
slave" />
    </table>
  </tables>
</dbpartition>
```

▶ hash mod拆表

▶ 账号读写分离

▶ 支持负载均衡

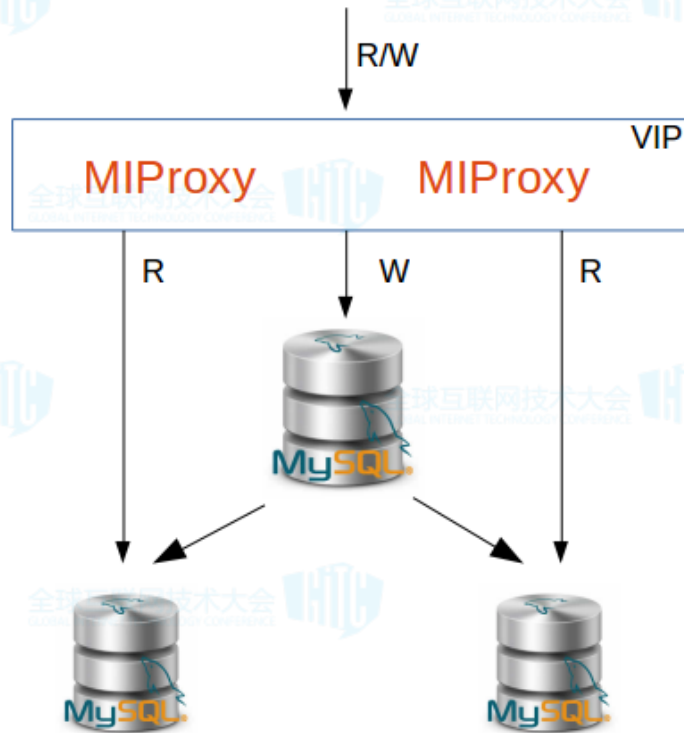
数据库中间件

数据库中间件






读写分离

连接复用

存储对业务透明



中间件选择

	kingshard	Atlas(Qihoo360)	MyCat	ProxySQL	Max-Scale
开发语言	Go	C	JAVA	C++	C
读写分离	Y	Y	Y	Y	Y
分库分表	Y	Y	Y	Y	Y
HA	Y	Y	Y	Y	Y
兼容协议	MySQL	MySQL	MySQL ORACLE Microsoft SQL Server IBM DB2 mongoDB	MySQL	MySQL
部署平台					
连接复用	Y	Y	Y	Y	Y
负载均衡	Y	Y	Y	Y	Y

MIProxy

选择对kingshard适应性改造

- ▶ 兼容性问题及**BUG**修复
- ▶ 连接**session**保持，防止连接状态丢失
- ▶ 多账号管理（**rwsplit, master, slave, admin**）
- ▶ 从库最近邻选择，延迟从自动踢除
- ▶ 快速启动/重启，优化连接池初始化策略
- ▶ 完善监控及日志打印，切割策略
- ▶ 服务发现与节点管理

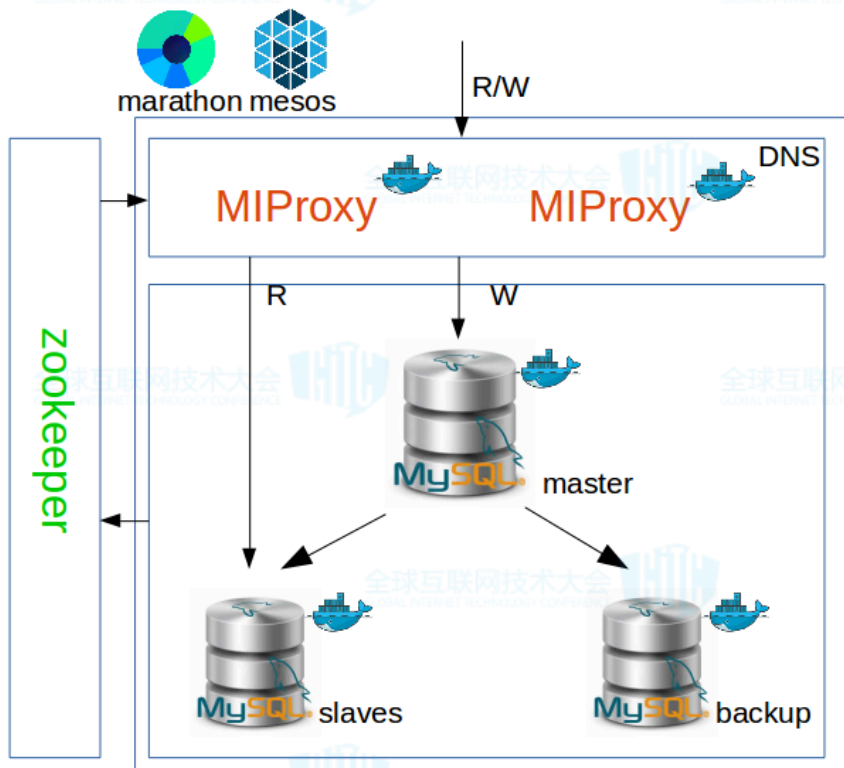
打消用户疑虑



mysql-connector-c
MySQLdb
Spring JDBC
mysql-connector-java
Hibernate
mysql-connector-python
mymysql
PDO_MYSQL
PyMySQL
go-sql-driver
beego

DBaaS(数据库即服务)

DBaaS



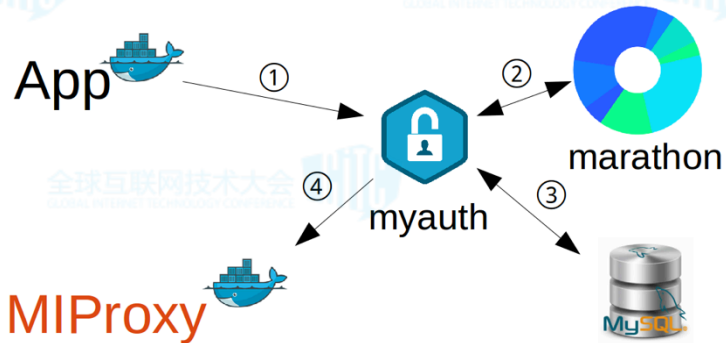
授权自动变更

故障自动转移

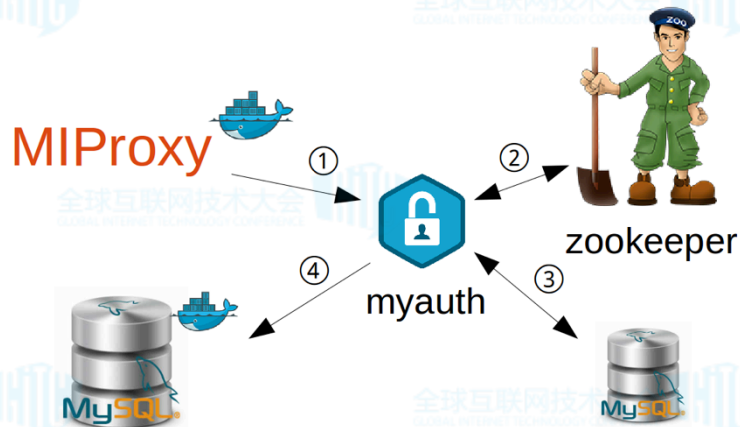
扩容一键完成

动态授权

应用到MIProxy授权



MIProxy到MySQL授权



一键扩容

数据库详情

当前状态

期望实例数: 4

运行实例数: 4

刷新

扩容

实例数目:

确定

从库故障自动扩容

使用当前备库数据恢复

同步完成标记服务上线

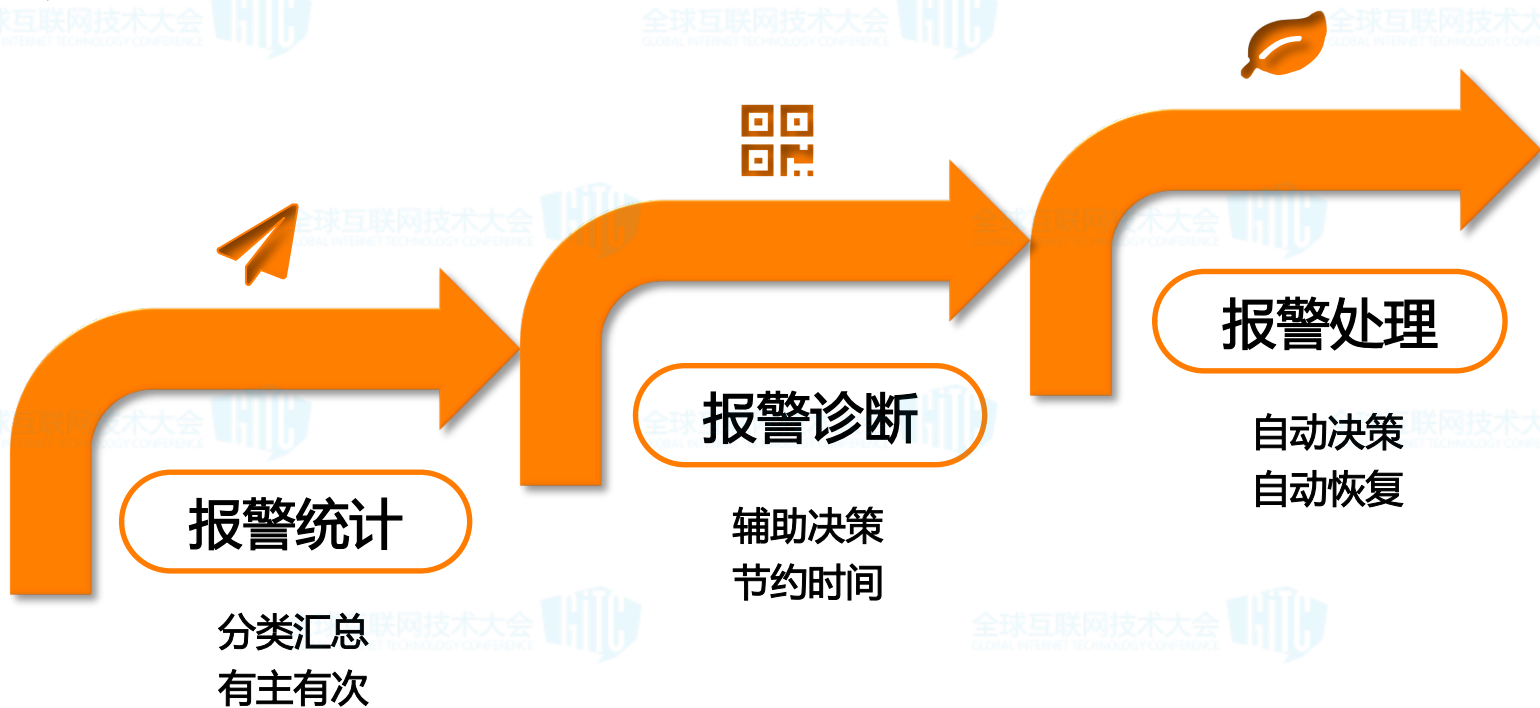
故障自愈

三步走战略

全球互联网技术大会

全球互联网技术大会

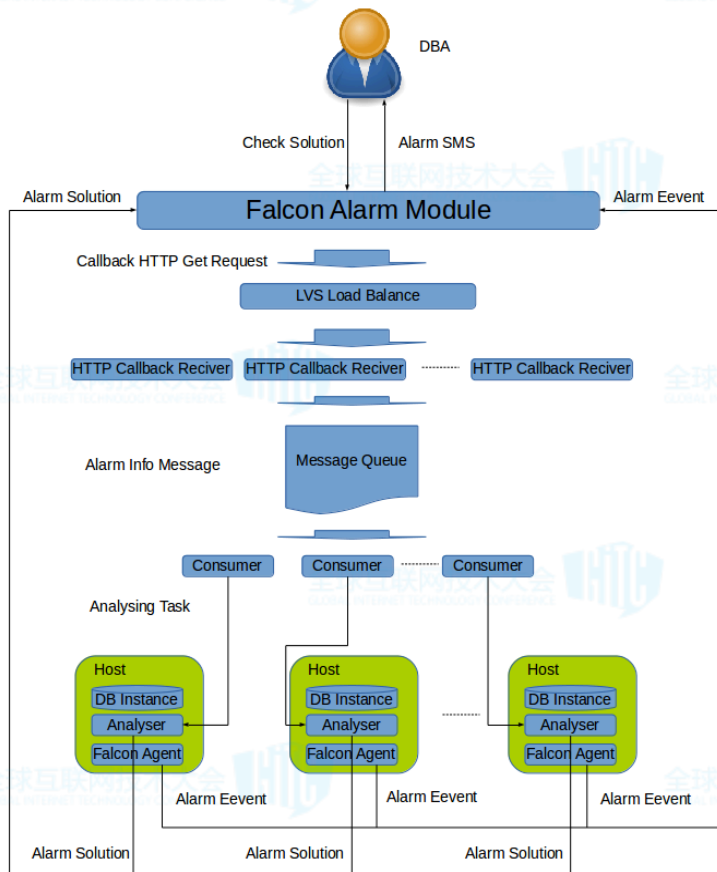
全球互联网技术大会



全球互联网技术大会

全球互联网技术大会

全球互联网技术大会



诊断方式

基础状态通用策略

网络
磁盘
CPU
内存

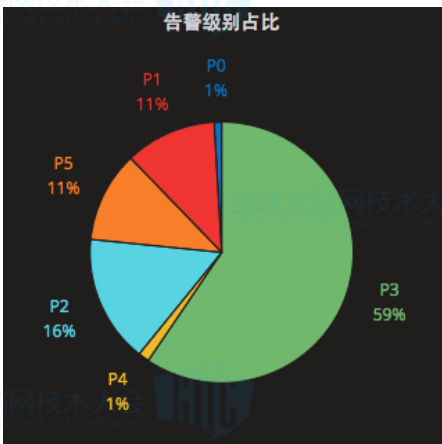
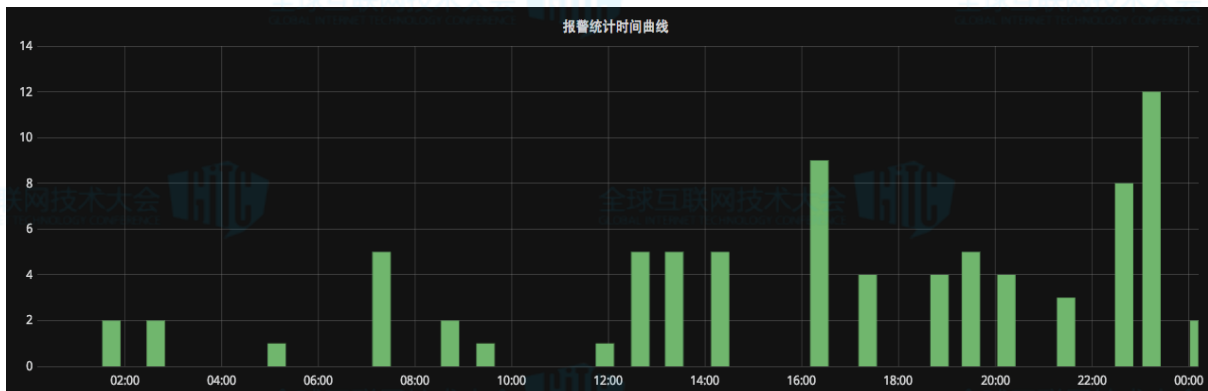
监控项定制策略

插件化
分权重
一票通过

报警统计

报警时间曲线

报警级别占比

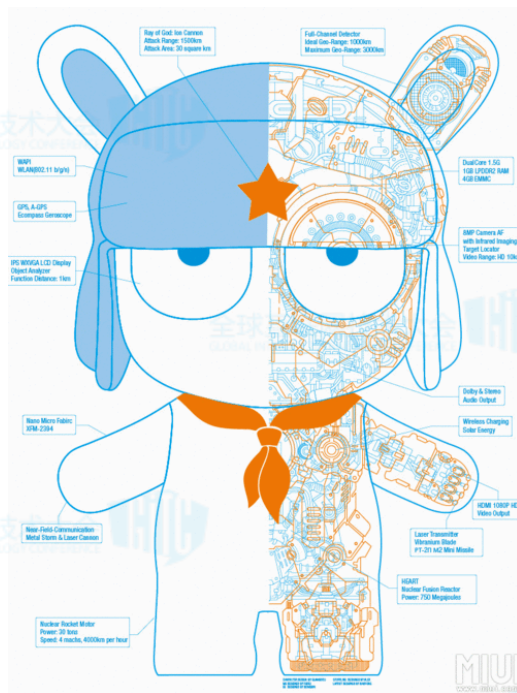


- 报警服务占比
- 报警metric TOP 20
- 报警服务器 TOP 20
- 未恢复的报警

准备做的工作

准备做的工作

- ▶ 提高诊断准确率
- ▶ 机器学习替代传统策略模型
- ▶ 手机APP移动办公



全球互联网技术大会

THANK YOU

全球互联网技术大会

小米DBA: 张良
邮箱: zhangliang3@xiaomi.com