



2017第八届中国数据库技术大会

DATABASE TECHNOLOGY CONFERENCE CHINA 2017

# 魅族数据库平台建设实践

龙启东



传统DB架构面临的问题

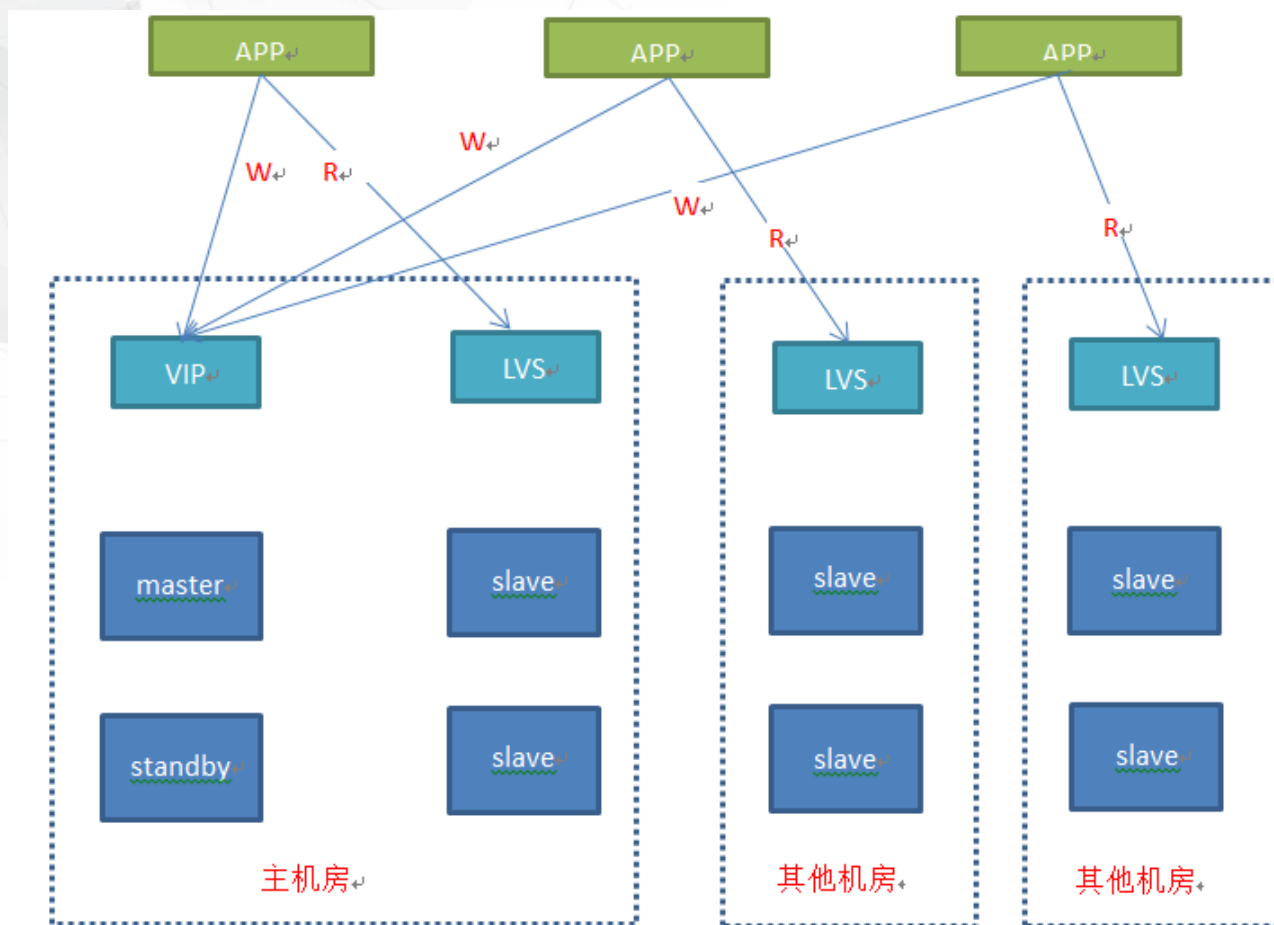
DB平台设计目标

DB平台架构

DB平台主要模块

平台规模和后续发展

# 传统DB架构



# 面临的问题

## 配置复杂

- 读地址、写地址
- 业务做读写分离

## 连接池问题

- DB连接 > 业务单机连接数 \* 机器数量
- 连接数不可控, *too many connections*异常

## 复制延时问题

- Slave延时, 必须在LVS上踢掉slave
- 操作延时、繁琐

## 跨机房切换问题

- 数据双写问题
- 业务更改配置重启, 参与人员较多
- 业务不透明

## 过载保护

- 某几条SQL即可将DB资源耗尽
- 连接数打满
- 整个实例不可用
- 大范围影响业务

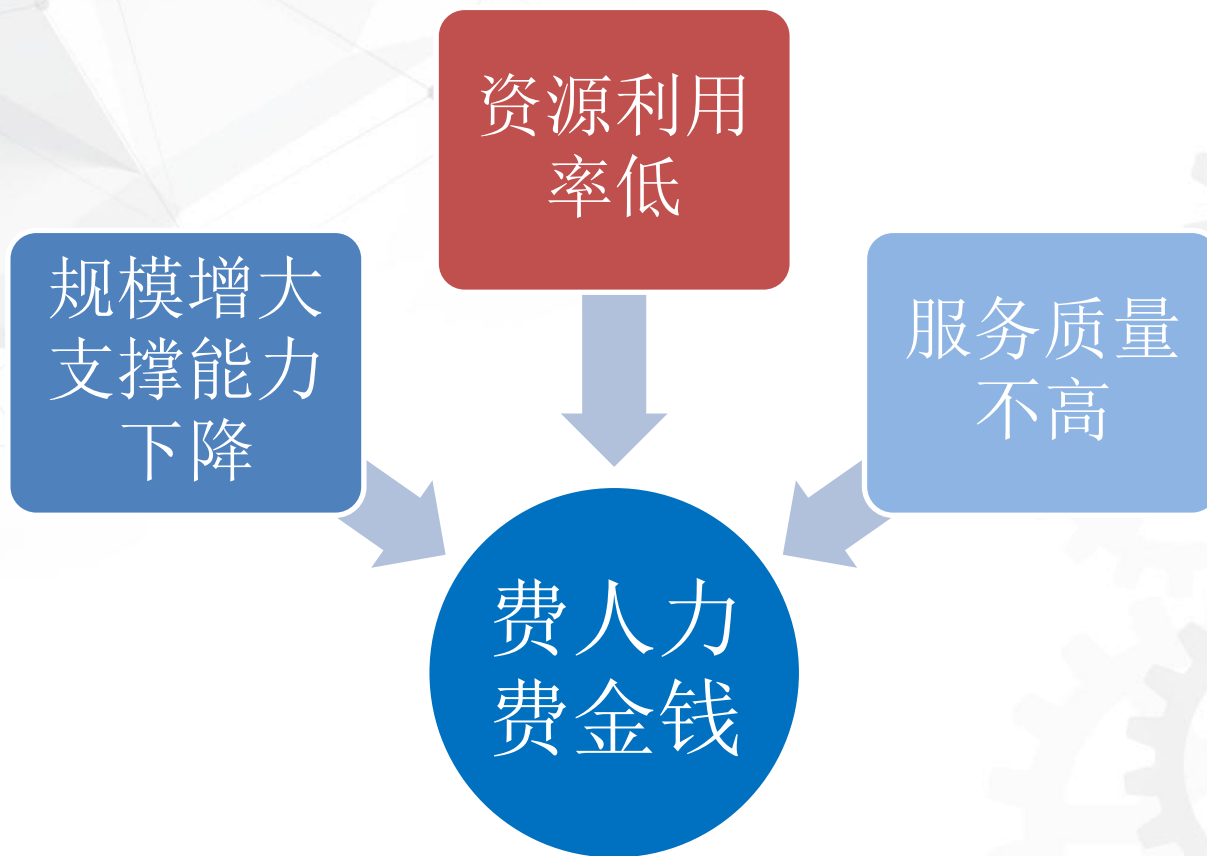
## SQL过滤和拦截

- 不带条件的delete
- 不带条件的update
- ...

## 资源利用率低

- CPU
- IOPS
- 磁盘空间

## 面临的问题





传统DB架构面临的问题

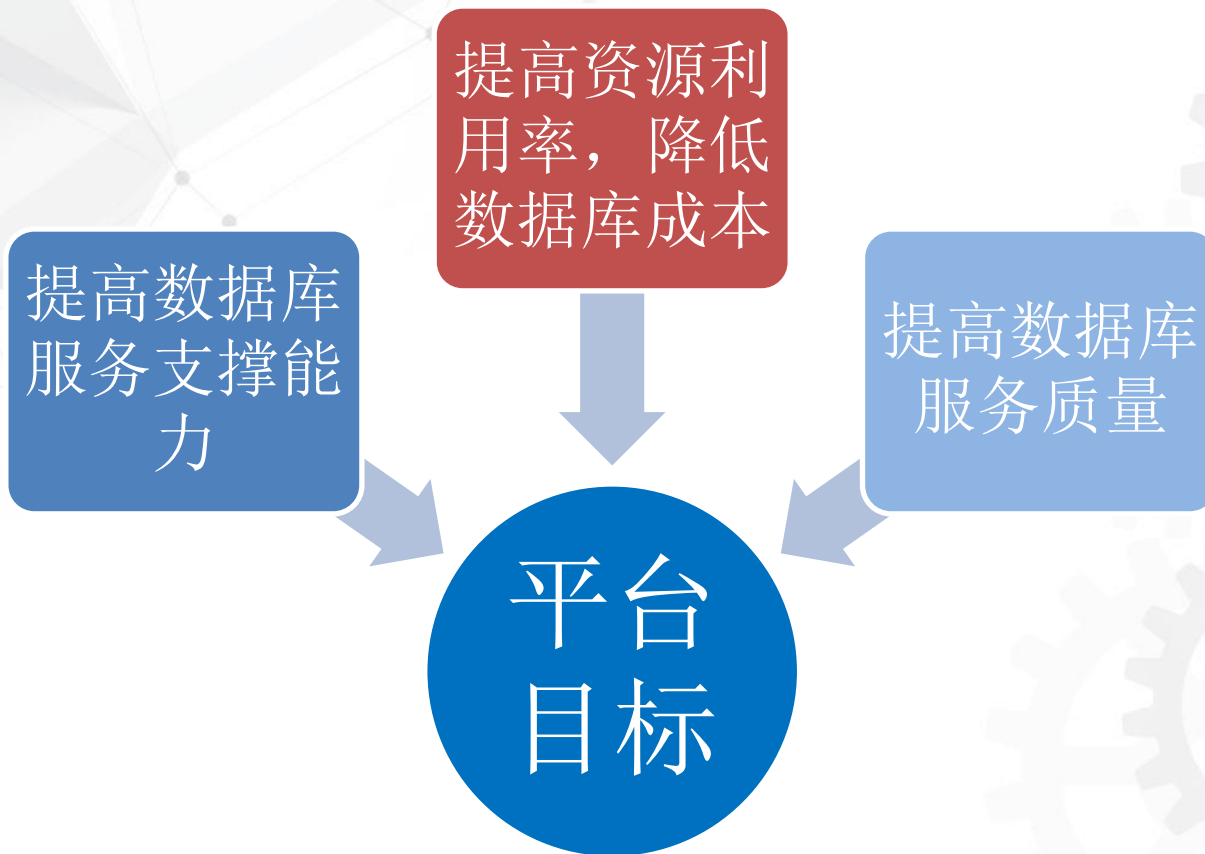
DB平台设计目标

DB平台架构

DB平台主要模块

平台规模和后续发展

# DB平台设计目标



# 平台目标细化

- 易用性
  - 配置简单，每个机房一个地址
  - 快速权限管理，SQL查询等
  - DBA运维简化，运维操作流程化、自动化
- 安全性
  - proxy接入，业务不再直连DB
  - proxy连接数控制，不再受max\_connections限制
  - 高危SQL过滤
  - 突发增长控制，避免DB雪崩，有效保护DB
- 可靠性
  - 数据高可用保障
  - 数据自动备份，支持30天内数据恢复
  - proxy自动检测slave延时，自动剔除和添加



- 可用性
  - 主备架构
  - 自动部署高可用、监控
  - 支持自动、手动快速故障切换
  - 资源隔离
- 可扩展性
  - 弹性扩容、缩容、迁移
  - 自动读写分离，在线增减slave
  - 快速部署跨机房slave
- 成本节约
  - 以实例为单位，严格按照业务需求分配资源
  - 读slave 只需一个实例即可
  - 快速的缩容和扩容，最大程度的控制DB资源成本
  - 成本最大缩减到目前的1/64
  - 根据cpu，IO，mem，磁盘进行计算



传统DB架构面临的问题

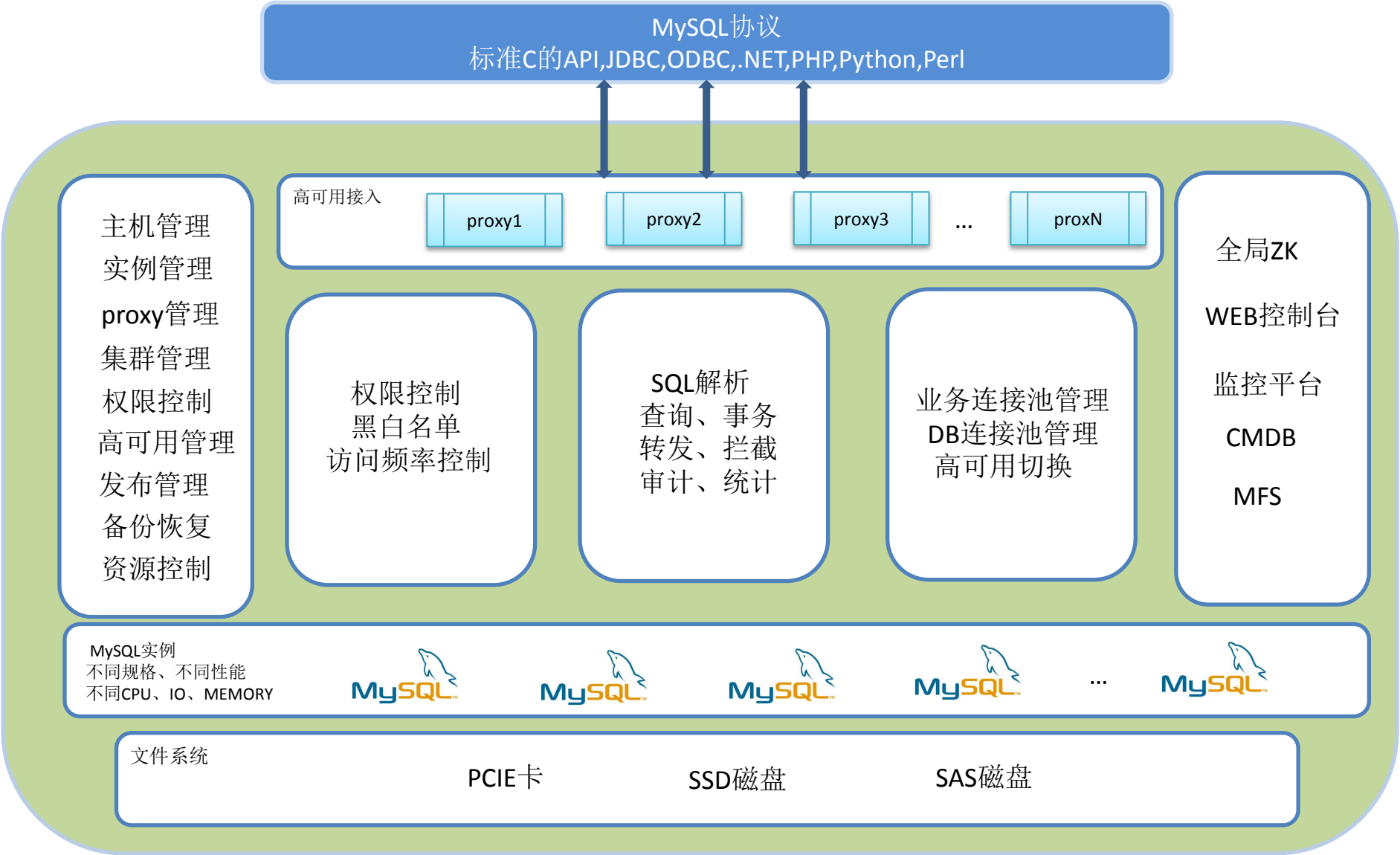
DB平台设计目标

DB平台架构

DB平台主要模块

平台规模和后续发展

# DB平台整体架构



# Proxy主要功能

兼容mysql 协议

连接数控制

高危SQL拦截

自动切换

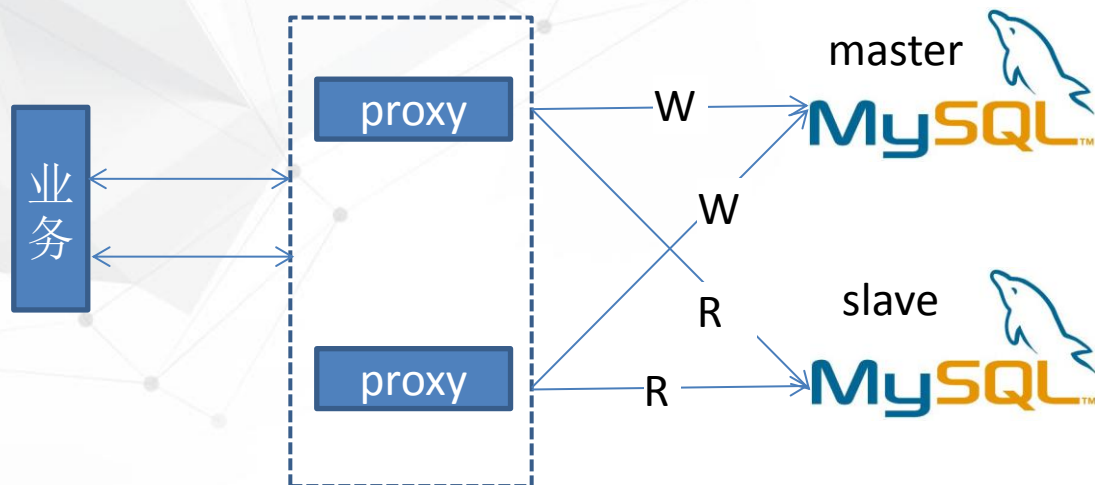
自动读写分离

SQL审计

访问频率控制、访问权限控制

访问延时统计分析

# Proxy访问



- vip方式访问
- zookeeper名字服务方式访问

# Proxy部署

## Proxy部署

每实例一个进程对应一个管理端口、一个访问端口

每实例对应一个MySQL主从

每VM运行多个进程

多个VM为一组DB接入

VIP方式或者zk方式访问

## Proxy性能

CPU负载为后端DB的10%

内存最大<512M

性能损耗5%-8%

延时增加<100us

## Proxy测试

功能测试

性能压测

tcpcopy线上流量测试

# Proxy接口

command	description
proxy help	commands help
proxy echo	echo for test
proxy server status	show proxy server status
proxy connection status	show proxy connection status
proxy show processlist	show proxy processlist
proxy show processlist2	show proxy processlist ignore idle connections
proxy kill (id)	kill connection by special client id
proxy trace	show trace log file
proxy trace start (file)	start trace log with (file)
proxy trace stop	stop trace log
proxy tcpdump	show tcpdump file
proxy tcpdump start (file)	start tcpdump with (file)
proxy tcpdump stop	stop tcpdump
proxy config	show current proxy server config
proxy status	
proxy config2	
proxy log level	
proxy log level (level)	OFF,FATAL,ERROR,WARN,INFO,DEBUG,TRACE,ALL
proxy sql records [maxcount]	
proxy set records [maxcount]	
proxy version	
proxy restart [waitseconds]	
proxy gc	
proxy slowquerytimeout [millis]	
proxy deniedips	
proxy getvars	
proxy setvar var value	set variable only effect in current process
proxy setvarp var value	set variable persist value
proxy getvar var	get variable value
proxy syswarn code message	
proxy syswarncodes	show syswarn known codes

# DB存储(资源池)

SAS、SSD、PCIE按需选择

内存、CPU按需分配

每服务器运行多MySQL实例

服务器按不同机柜、交换机部署

MySQL实例对业务透明

只授权proxy访问



# 备份恢复

备份调度任务管理

全量备份、增量备份

备份校验

创建slave、迁移、扩容、缩容

回档(整实例、某库、某表)、导入导出

备份副本管理

恢复任务管理

# DB性能分析

慢日志监控、分析、推送、跟踪

容量监控、评估

实例访问监控、资源消耗监控、评估

主机性能监控、评估

DB健康状态综合报告



传统DB架构面临的问题

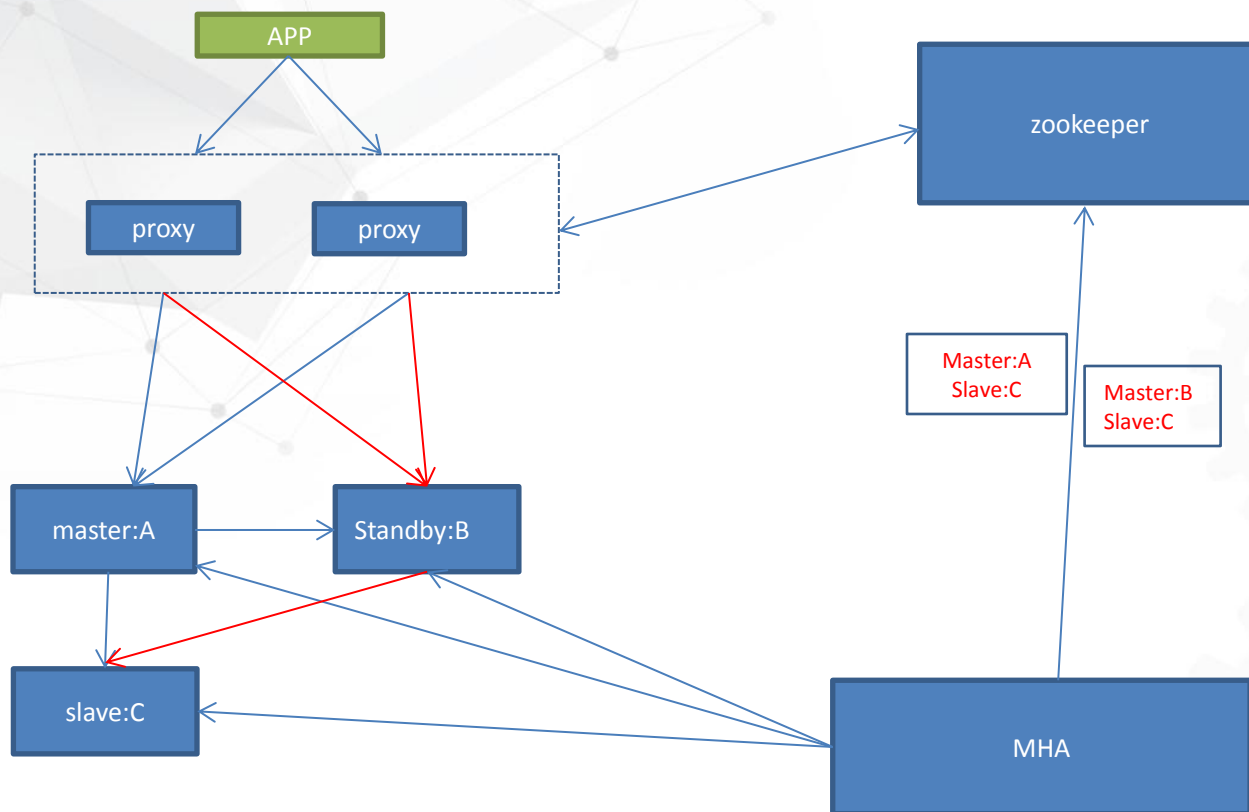
DB平台设计目标

DB平台架构

DB平台主要模块

平台规模和后续发展

# 主从切换



# 切换场景

故障自动切换

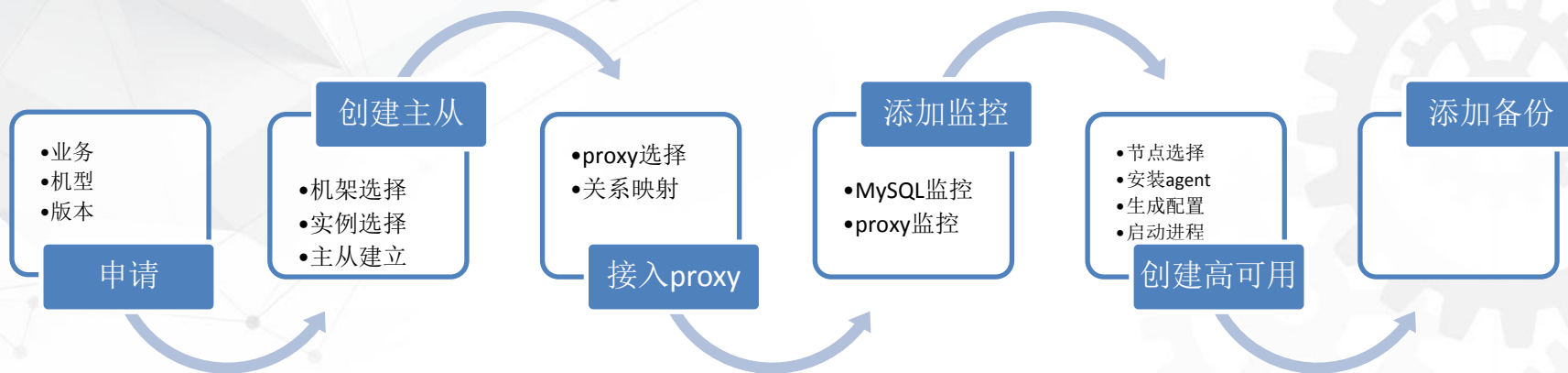
手动切换

迁移

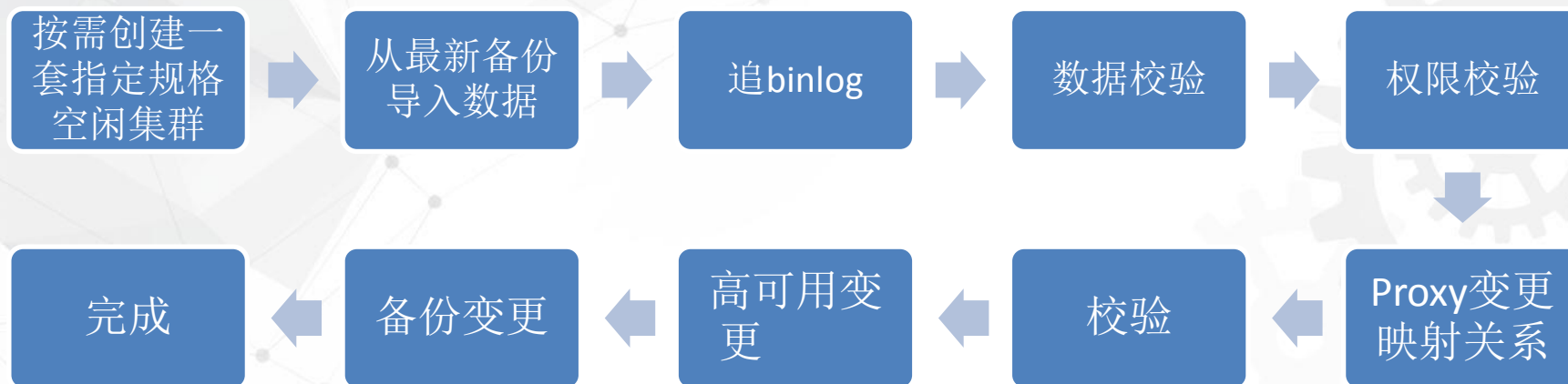
扩容

缩容

# 创建集群流程

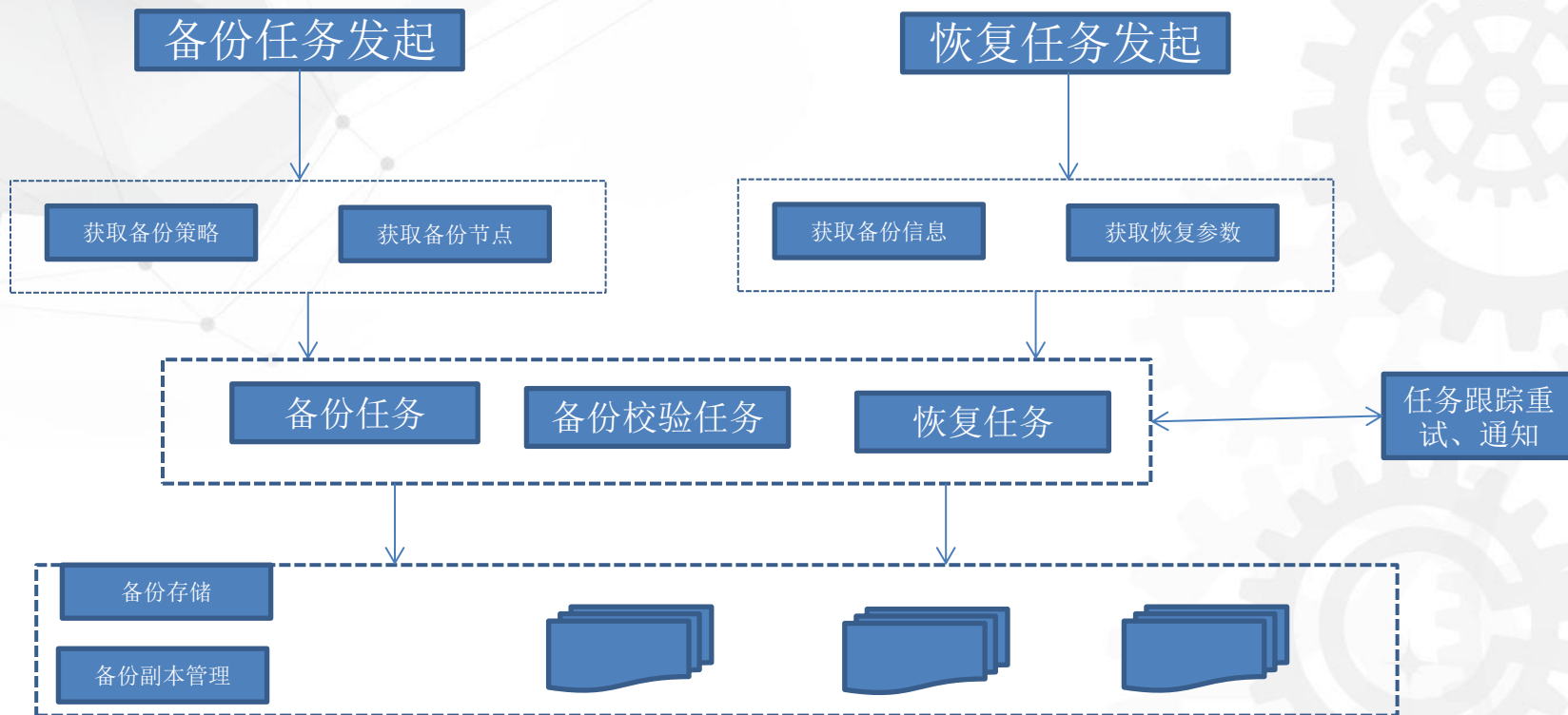


# 扩容、缩容、迁移



1. 整个过程业务无需更改任何配置，影响时间10s左右
2. 业务最大可扩展到128T
3. 业务最小只使用整机的1/64
4. 存储节点的性能差异对业务不可见
5. 根据业务增长周期动态扩展，最大程度利用已有资源

# 备份、恢复

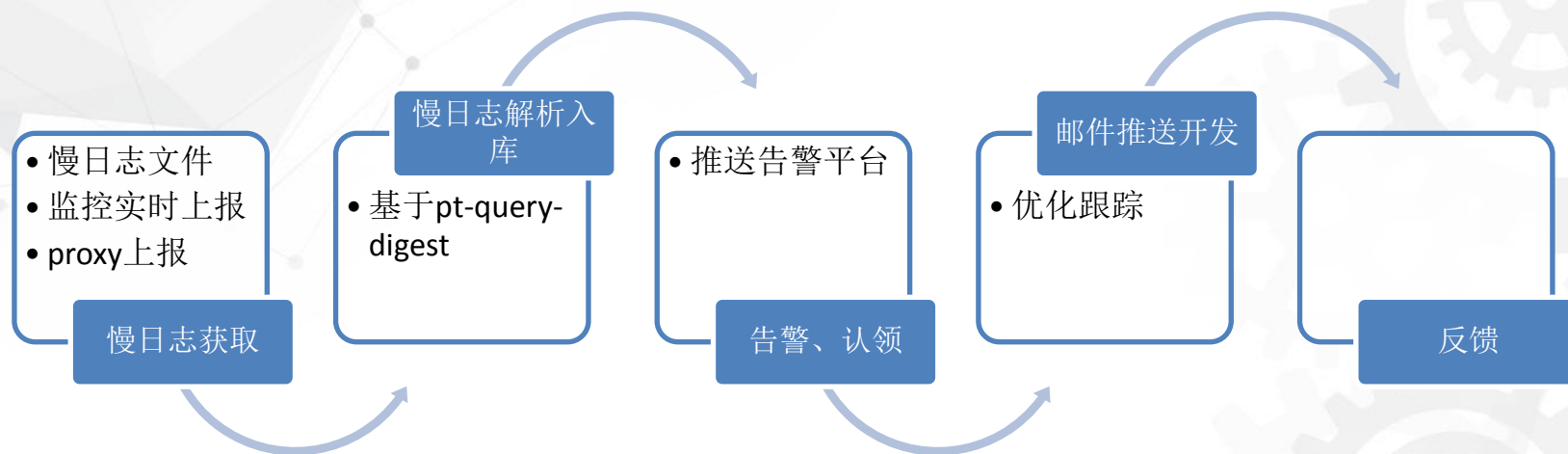




# 资源管理



# 慢日志



# 运维操作入口

## 逻辑单元信息

主机房: 请选择 \* 业务系统名称: 请选择 \* 业务系统别名: \*

MySQL版本: 请选择 \* 实例规格: 请选择 \* 从服务器数: 0 \*

字符集: 请选择 \* 备份天数: 5 \* 机型: 请选择 \*

申请人: \* 业务负责人: \* 运维负责人: \*

☐ 从机是备份机

添加实例

## 操作列表

单元上线 单元下线 单元回收 显示所有备份 主动备份 恢复备份  
代理切换 单元迁移 执行命令 执行命令日志 MHA管理 切换备份

## 恢复备份

逻辑单元编号: Logic\_20160826143023954284 备份ID: 36934 \* 机房: \*

实例ID: Instance\_2016071410545 \* IP地址: \* 实例端口: 3312 \*

选择备份

选择	备份ID	IDC	文件路径	文件大小	BIN_MARK	BIN_POS	错误码 (0为正常)	错误信息	端口
<input checked="" type="radio"/>	36934	ns	/mnt/DBbackup/mysql/20170410	10927205244	000198	12628133	0		3308
<input type="radio"/>	36747	ns	/mnt/DBbackup/mysql/20170409	10895809449	000197	322684891	0		3308
<input type="radio"/>	36558	ns	/mnt/DBbackup/mysql/20170408	10860781285	000196	555604446	0		3308
<input type="radio"/>	36371	ns	/mnt/DBbackup/mysql/20170407	10826615435	000195	812742389	0		3308

## 实例信息

机房: IP地址: 实例规格: 实例类型:

选择	机房	实例编号	IP地址	端口	MySQL版本	实例规格	实例类型	状态	使用状态
<input checked="" type="radio"/>		Instance_20160714105451518404		3312	MySQL-5.5.9	8	普通	正常	未使用
<input type="radio"/>		Instance_20160714105459566450		3310	MySQL-5.5.9	8	备份	正常	未使用

主机: 5307(Pri) Graph: MysqlProxy 5307 Proxy命令计数器

1h 2h 3h 6h 12h 1d 7d 14d 1m All

## MysqlProxy 5307 Proxy命令计数器 (5307)



- MySQLProxy 5307 ClientSqlCommit [avg](min:0,avg:0,max:0)
- MySQLProxy 5307 ClientSqlDelete [avg](min:0,avg:0,max:0)
- MySQLProxy 5307 ClientSqlInsert [avg](min:2,avg:20.0556,max:65)
- MySQLProxy 5307 ClientSqlOther [avg](min:14,avg:25.2118,max:81)
- MySQLProxy 5307 ClientSqlReplace [avg](min:0,avg:0,max:0)
- MySQLProxy 5307 ClientSqlSelect [avg](min:1,484,avg:4,251.2347,max:6,352)
- MySQLProxy 5307 ClientSqlUpdate [avg](min:0,avg:0.3833,max:3)



传统DB架构面临的问题

DB平台设计目标

DB平台架构

DB平台主要模块

平台规模和后续发展

## 后续发展

- 1.proxy访问接入大数据分析，做访问行为、安全扫描
- 2.更完善、更智能的SQL审核、发布机制
- 3.降低proxy自身的性能消耗，降低平台本身成本
- 4.资源隔离方案进一步优化



# THANKS

SequeMedia  
盛拓传媒

IT168.com

ITPUB

ChinaUnix.net