

ACMUG & CRUG 2018

魅族数据库中间件 设计与实现

龙启东



01

可维护性

读写分离配置
延时、故障剔除添加

02

服务质量

too man connections
高危SQL
请求突增
限流、降级

03

扩展性

对业务不透明
vip无法跨网段
硬件升级、过保、迁移、故障

04

成本

单机单实例
单机多实例
读写分离，容灾



01

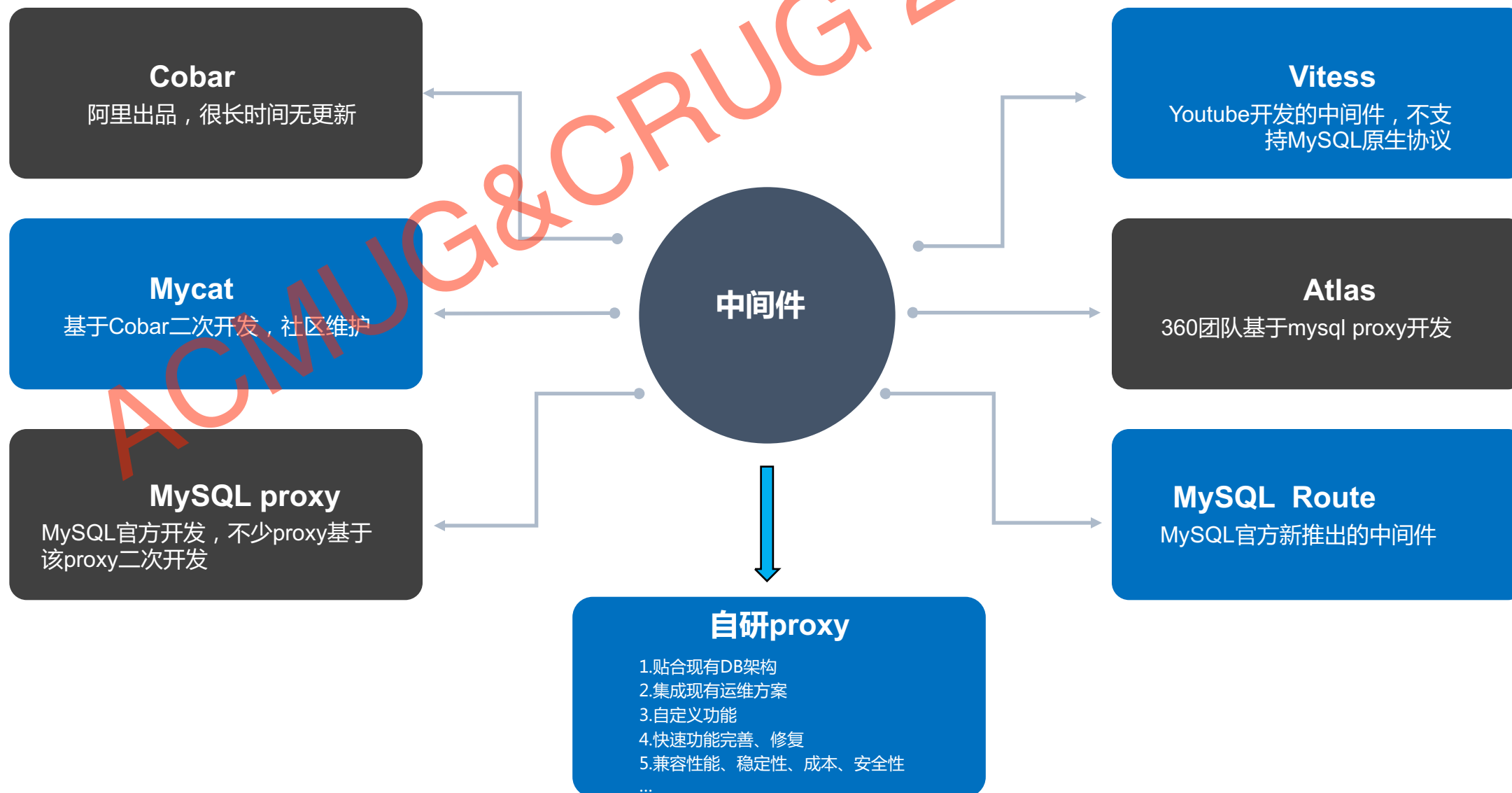
lvs/haproxy

02

域名方式

03

命名服务





易用性

兼容MySQL协议，对业务透明，
配置简单，proxy足够稳定，部署
简单，支持批量部署，已有项目无
缝迁移，集成到DB平台

性能统计

QPS，TPS统计，slow log，
各阈值区间请求统计，故障定
位，性能统计接口

安全性

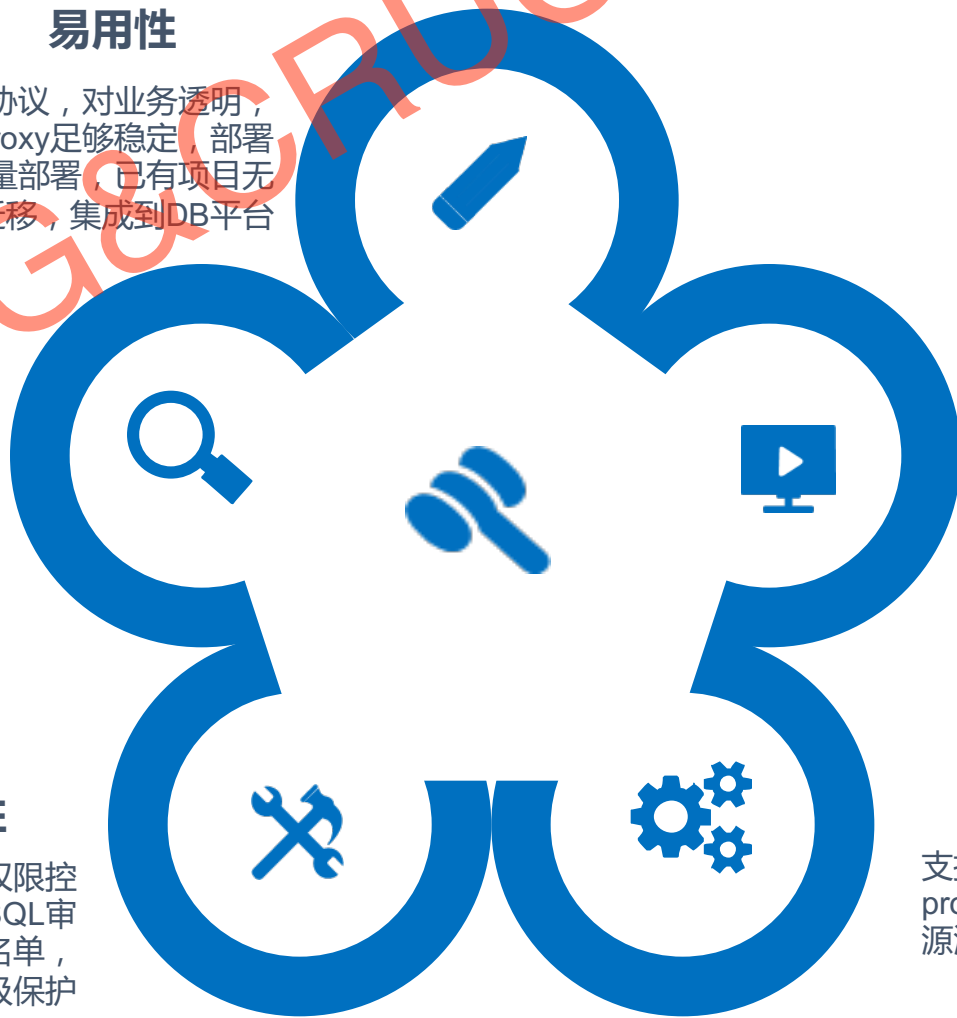
后端DB对业务不可见，权限控
制收缩，高危SQL过滤，SQL审
计，连接池保护，IP黑白名单，
高峰限流，用户优先级保护

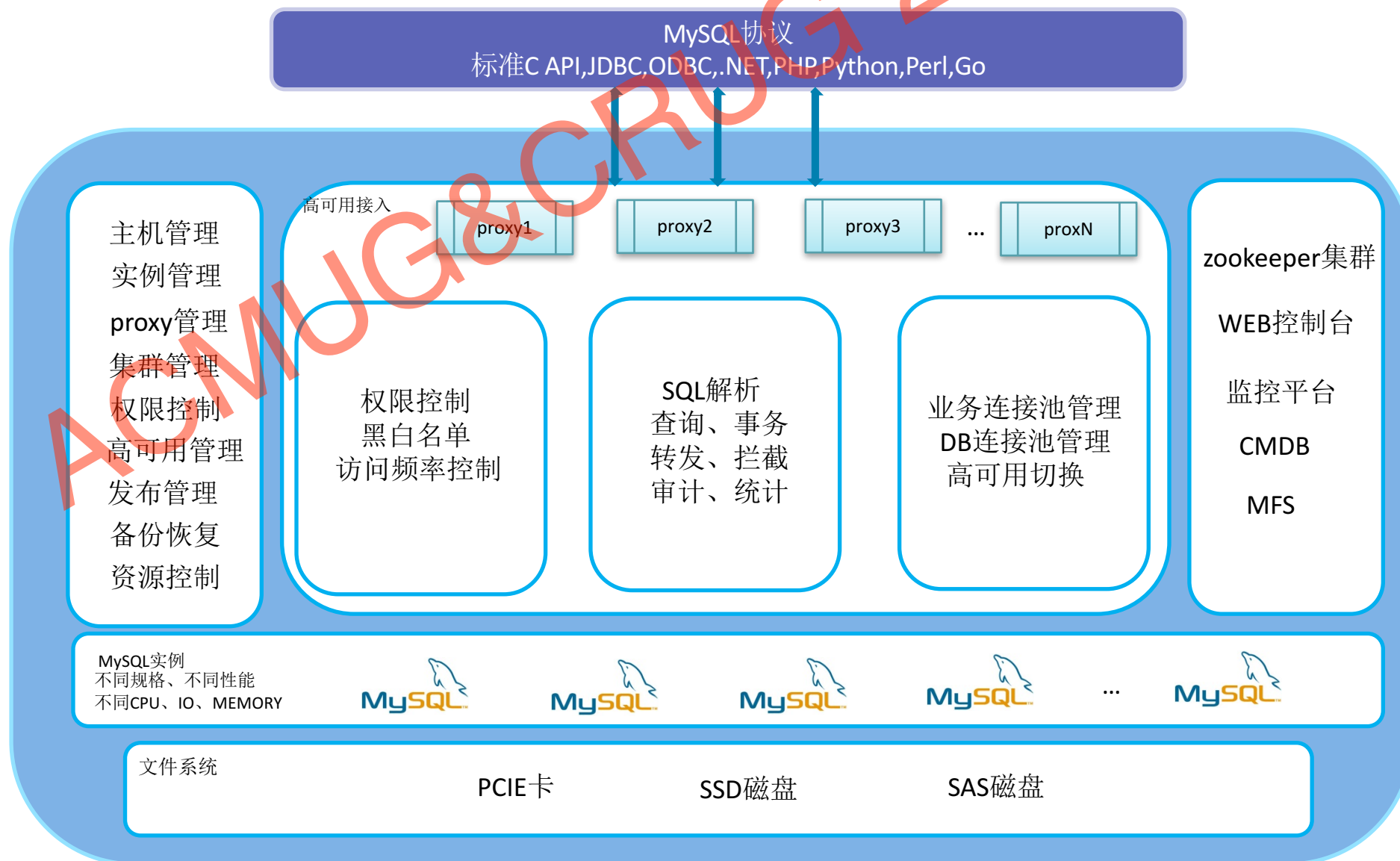
可扩展性

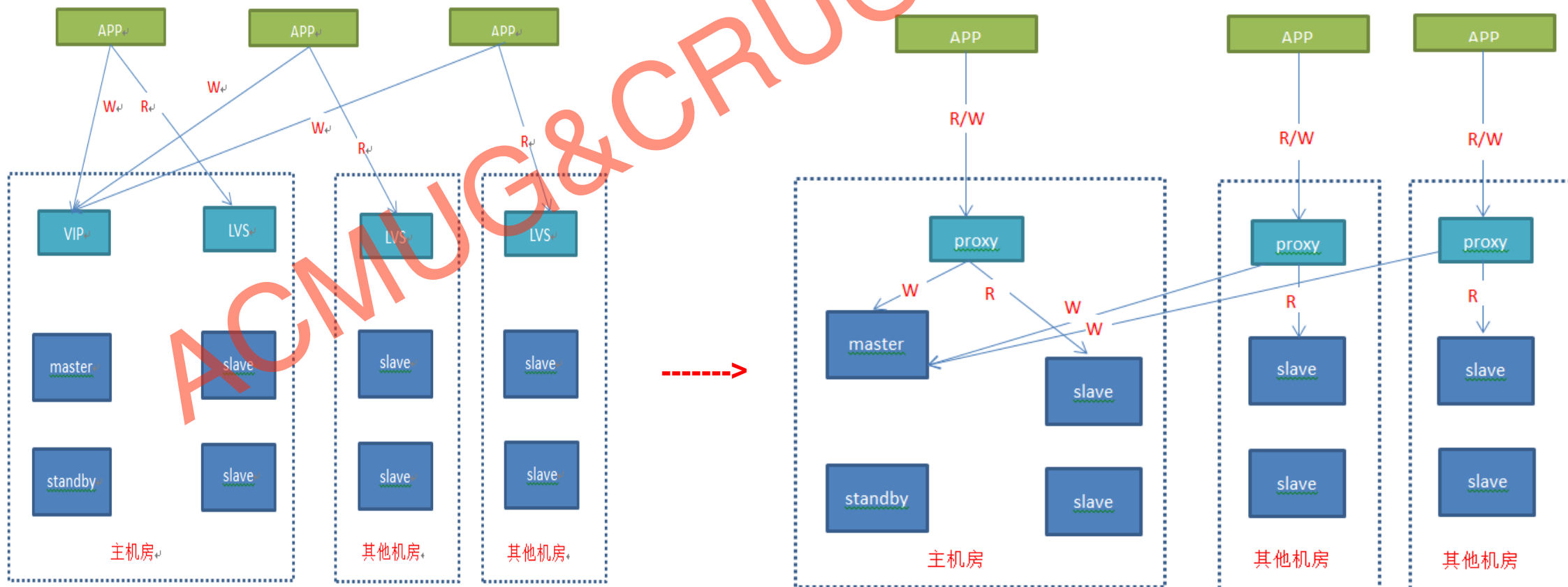
自动读写分离，负载均衡，在
线增减slave，自动延时剔除、
添加，故障自动剔除，恢复自
动添加

可用性

支持MySQL已有高可用方案，
proxy高可用，跨机房容灾，资
源消耗小，性能损耗小

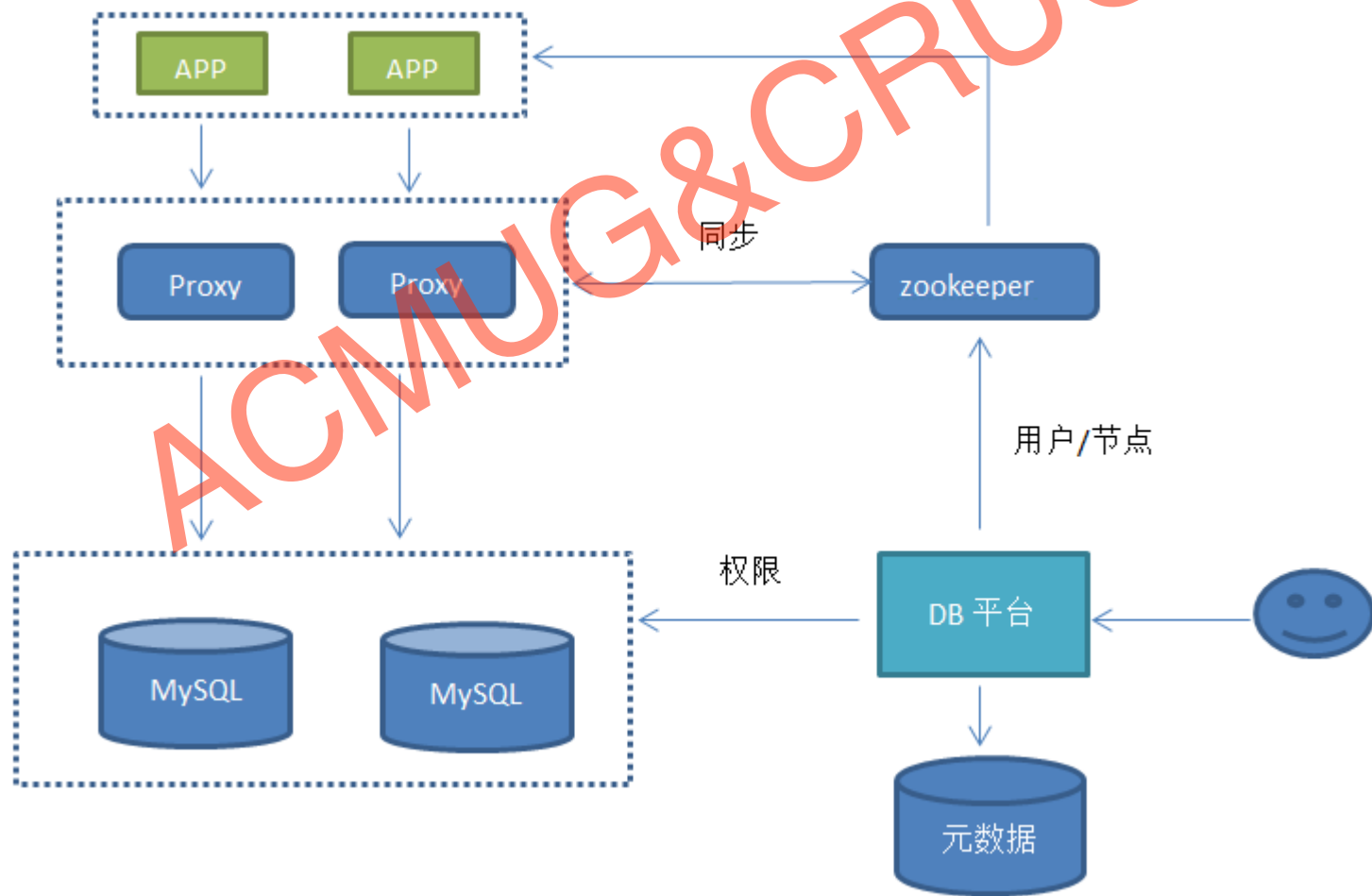




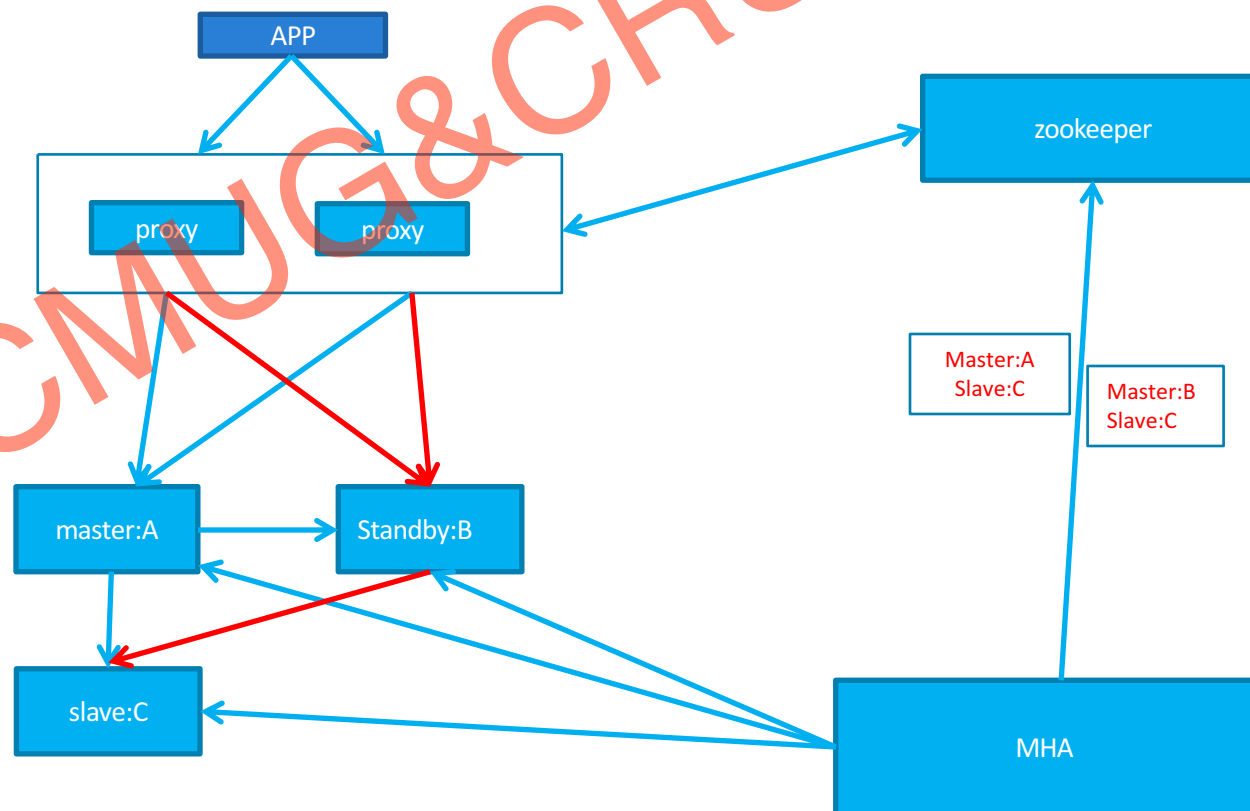


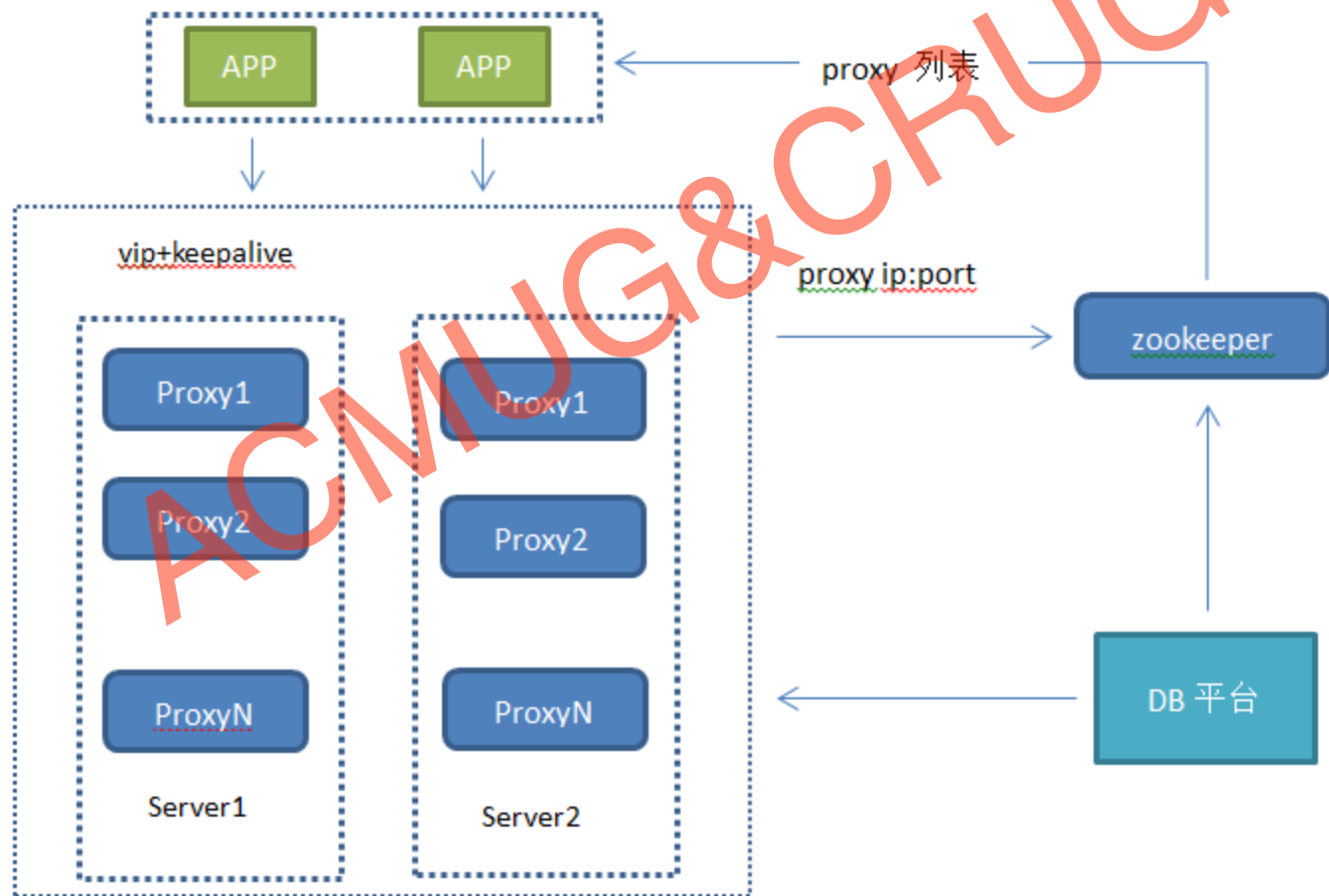
各个机房访问各个proxy
 自动读写分离
 事务转发master
 支持hits
 slave异常, 访问master
 延时自动剔除、恢复

7 Proxy功能设计-权限设计



- 1.proxy不存储密码
- 2.proxy不存储权限
- 3.zk存储用户节点信息
- 4.只能平台发起授权操作
- 5.root直接转发后端mysql
- 6.mysql只对proxy授权





- 1.vip方式
- 2.zookeeper命名服务方式
- 3.主备镜像
- 4.proxy实例池



连接池管理

各用户流量限制

各用户优先级

针对某类SQL拦截

黑白名单

无where条件update/delete 拦截

请求结果集限制



01

实时汇总记录转发到
后端DB出现的slow
log,兼容mysql slow
log格式

慢日志

02

业务请求中出现的异
常日志：字段截断，
表不存在，乱码，主
键冲突，无条件
update/delete语句拦
截

异常SQL

03

设置开关，记录每条
SQL的请求响应明细
包括后端DB,响应时
间，结果集大小，事
务状态等，记录无法
解析的SQL。slave
延时、故障剔除添加
信息。

故障定位



```
+-----+-----+
| 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  |
+-----+-----+
```



proxy show
processlist

根据需要记录SQL请求响应明细，或数据包，故障分析定位很有用

proxy
config/status

查看最近N条SQL请求/去重后的N条SQL请求

proxy set/setp
var value

proxy自身运行的指标告警接口，内存，空指针等

查看当前业务与proxy连接，proxy与后端DB连接，运行SQL，运行时间，运行状态等

proxy
trace/tcpdump

查看proxy配置信息，统计信息

proxy sql/set
records [N]

动态调整proxy自身参数，连接数，请求数，字符集等

proxy syswarn
code message

14 Proxy接口



```
784 rows in set (0.01 sec)
```

Proxy连接数控制，proxy对业务端无限制，对DB端根据配置而定

```
205 rows in set (0.00 sec)
```



proxy SQL请求统计

mysql> proxy sql records 10;

sql	count	totalMillis	totalRequestBytes	totalResponseBytes	aveMillis	aveRequestBytes	aveResponseBytes
INSERT INTO sbtest22(id, k, c, pad) VALUES (...)	902	175	55022	119926	0	61	132
INSERT INTO sbtest23(id, k, c, pad) VALUES (...)	889	471	54229	118157	0	61	132
INSERT INTO sbtest24(id, k, c, pad) VALUES (...)	882	192	53802	117186	0	61	132
INSERT INTO sbtest25(id, k, c, pad) VALUES (...)	872	181	53192	115896	0	61	132
INSERT INTO sbtest26(id, k, c, pad) VALUES (...)	857	404	52277	113901	0	61	132
INSERT INTO sbtest27(id, k, c, pad) VALUES (...)	847	314	51667	112531	0	61	132
INSERT INTO sbtest28(id, k, c, pad) VALUES (...)	840	164	51240	111640	0	61	132
INSERT INTO sbtest29(id, k, c, pad) VALUES (...)	829	175	50569	110057	0	61	132
INSERT INTO sbtest30(id, k, c, pad) VALUES (...)	821	301	50081	109153	0	61	132
SELECT @@version_comment LIMIT ?	2	2	74	198	1	37	99

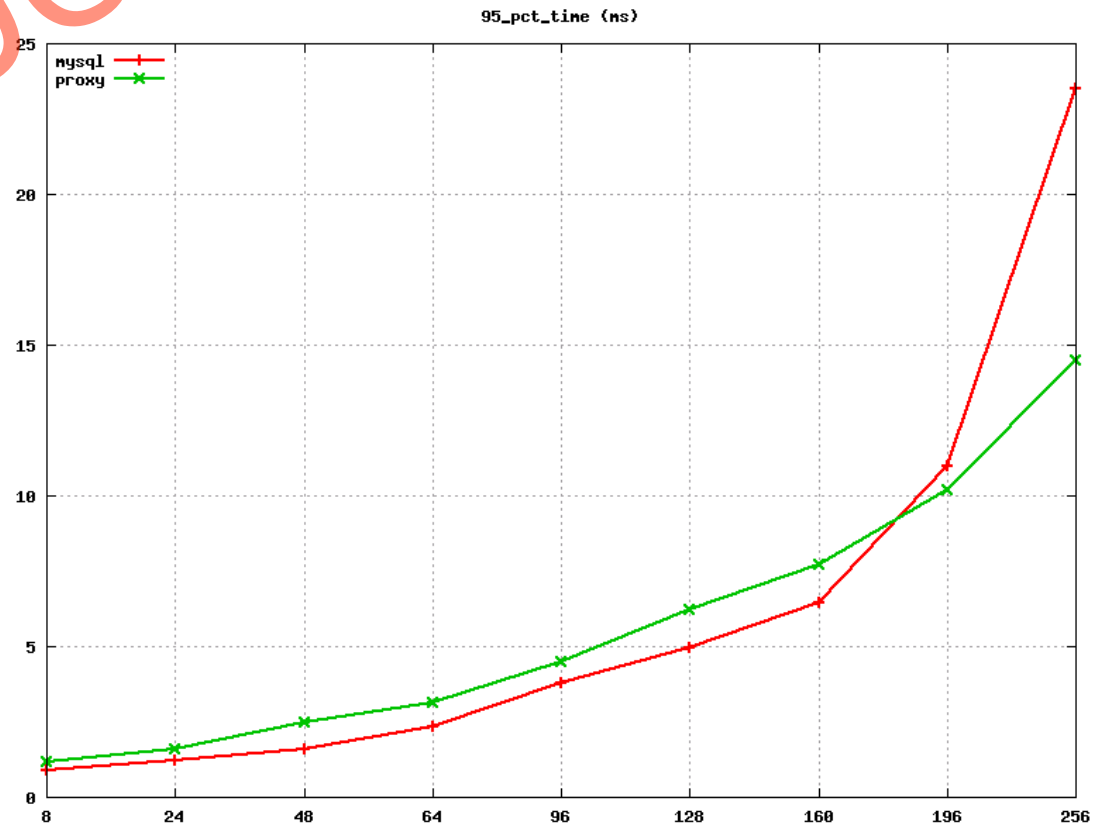
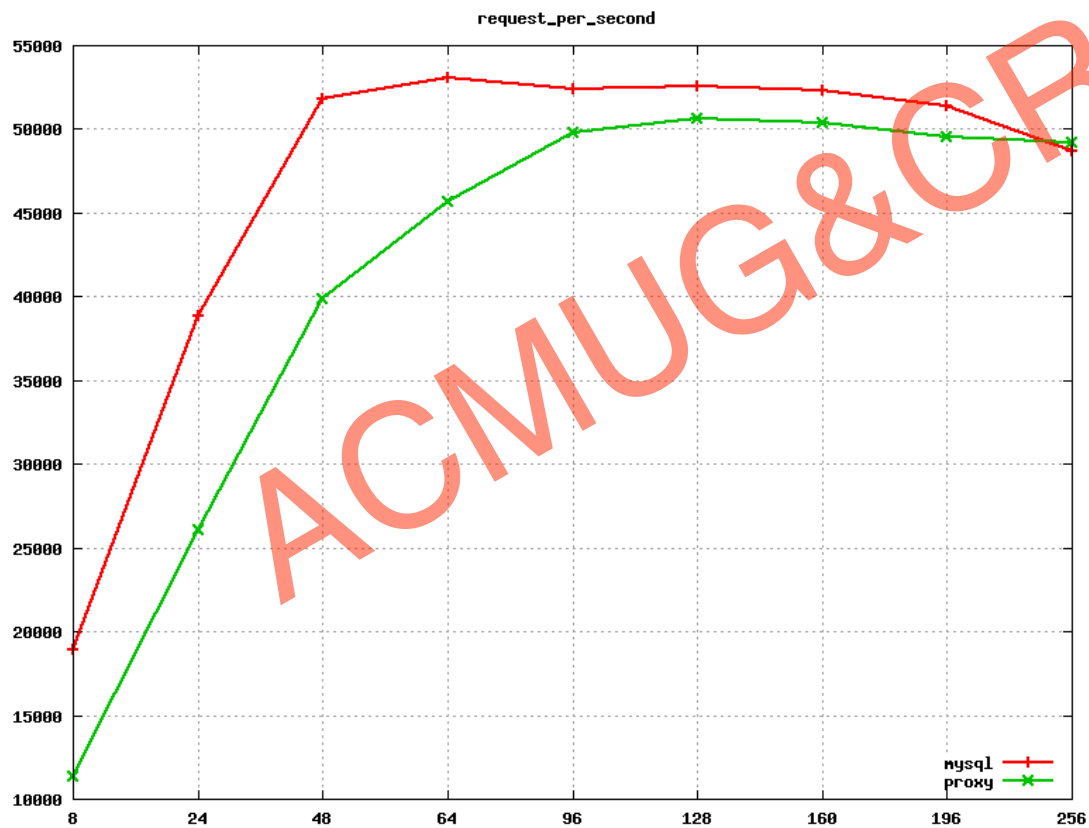
10 rows in set (0.00 sec)

Proxy动态参数

mysql> proxy getvars;

NAME	TYPE	CURRENT	PROCESS	PERSIST	DEFAULT	HELP
sockTcpNoDelay	Boolean	true	NULL	NULL	true	socket tcpNoDelay option
sockSendBufferSize	Integer	NULL	NULL	NULL	NULL	socket sendBufferSize option
sockReceiveBufferSize	Integer	NULL	NULL	NULL	NULL	socket receiveBufferSize option
slowQueryTimeout	Long	1000	NULL	NULL	1000	min slow query time millis
deniedIPs	String	NULL	NULL	NULL	NULL	denied ips
authIPs	String	NULL	NULL	NULL	NULL	authed ips
proxyDirectResponse	Boolean	false	NULL	NULL	false	NULL
syswarnShell	String	syswarn.sh	NULL	NULL	syswarn.sh	system warning shell
maxSqlSetRecords	Integer	10000	NULL	NULL	10000	NULL
maxClients	Integer	8000	NULL	NULL	8000	max client connections
clientCharset	String	NULL	NULL	NULL	NULL	set client charset, ignore client charsets.
defaultClientCharset	String	utf8mb4	NULL	NULL	utf8mb4	default client charset
defaultServerCharset	String	utf8mb4	NULL	NULL	utf8mb4	default server charset
escapeUtf8Mb4Characters	Boolean	false	NULL	NULL	false	escape utf8mb4 characters in log
autoConvertUtf8mb4	Boolean	false	NULL	NULL	false	auto change client charset from utf8 to utf8mb4 when query contains mb4 chars.
logQueryStatements	Boolean	false	NULL	NULL	false	log query statements
maxSqlLogSize	Integer	8192	NULL	NULL	8192	max sql log size
clientCapabilityFlags	Boolean	false	NULL	NULL	false	client capability flags
disablePrepare	Boolean	false	NULL	NULL	false	disable prepare
disableClientKill	Boolean	NULL	NULL	NULL	NULL	disable client kill

20 rows in set (0.00 sec)



MySQL CPU 跑满2300%

Proxy进程CPU使用率220%

高并发性能损耗5%

低并发延时增加0.3ms



- 开发环境测试没问题，线上切换报错（测试不够充分）
- Proxy测试一定时间段的general log
- 采用tcp-copy导流测试
- tcpdump、trace日志分析
- 对于解析失败的SQL语句全部转发master(读写分离失效)
- 权限对已有连接不生效，增加proxy refreshpool触发重连
- Proxy共享连接问题
 - set sql_mode
 - sql_select_limit
 - use db
 - 字符集
 - autocommit
 - last_insert_id
- 99%业务无缝迁移，少数几个业务需要调整

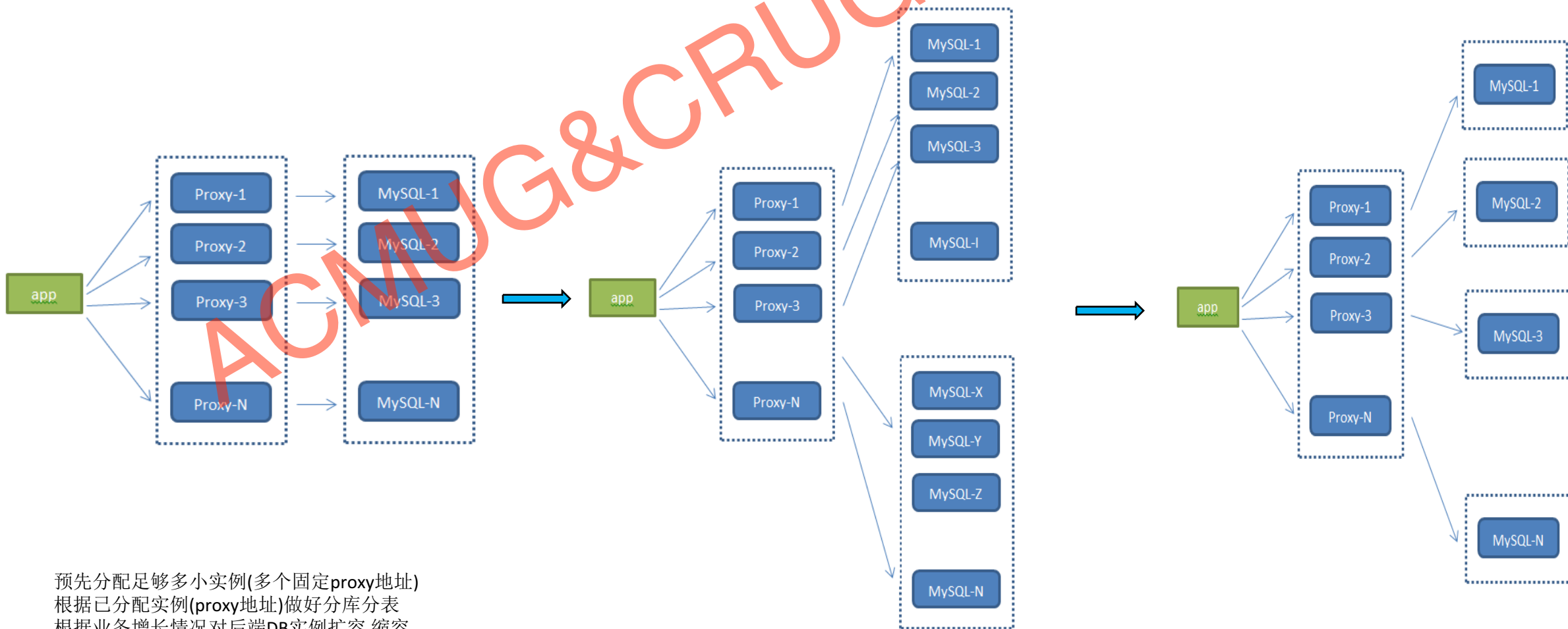


• 客户端处理

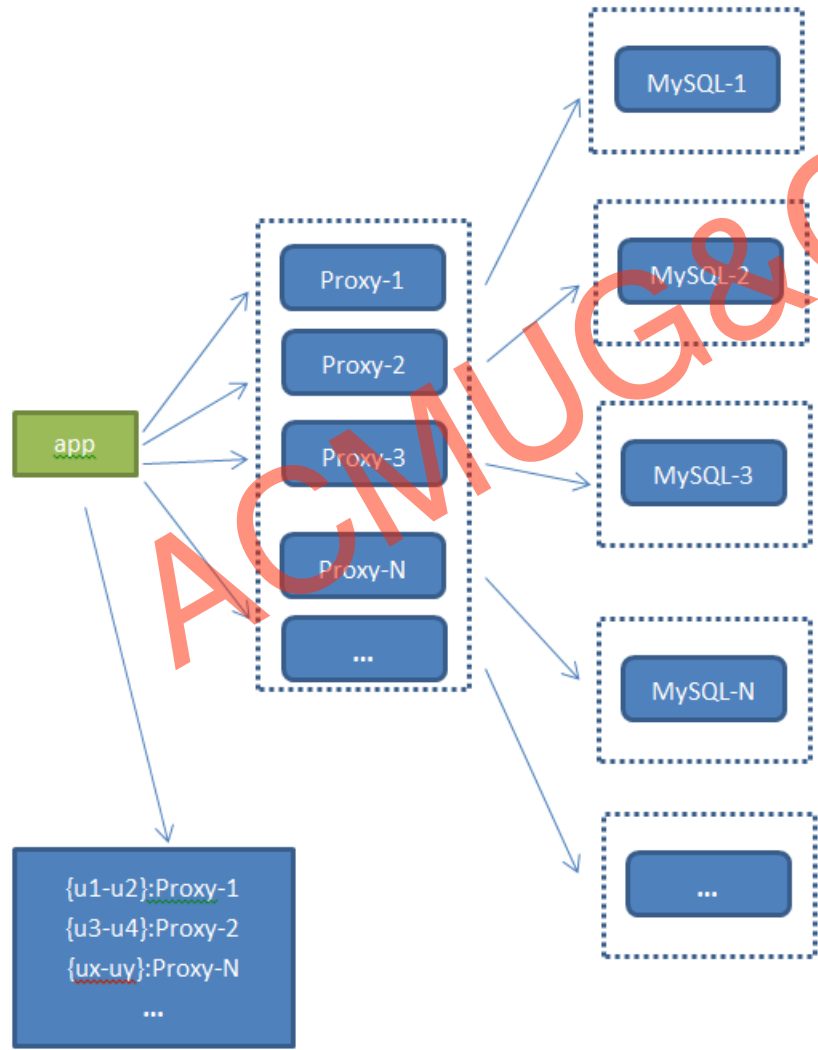
- 现有客户端分库分表方案成熟
- 运维工具支持
- 各个业务分库分表维度不统一
- 减少proxy负载, 降低proxy成本
- 方便快速扩容缩容

• 如何实施

- 预先分配
- 按需分配



预先分配足够多小实例(多个固定proxy地址)
根据已分配实例(proxy地址)做好分库分表
根据业务增长情况对后端DB实例扩容,缩容
扩缩容业务无感知
充分利用资源,避免资源浪费
最大支撑到 $N * 2T$



路由库

按需添加MySQL组
将MySQL组关联的proxy添加到路由库
按需继续上两步



Proxy实例 360+

MySQL实例数量1200+

集群300+

运行时间最长proxy持续运行了400+天

覆盖魅族所有互联网业务



ACMUG&CRUG 2018

谢谢