

16:02:53



# 聚美redis架构及运维实践



JUMEI.COM

# 内容提要

---

1

聚美redis集群架构演进过程

2

聚美redis集群缩容扩容设计

3

聚美redis自动化运维平台设计

4

聚美redis监控设计

# 架构演变

1

1

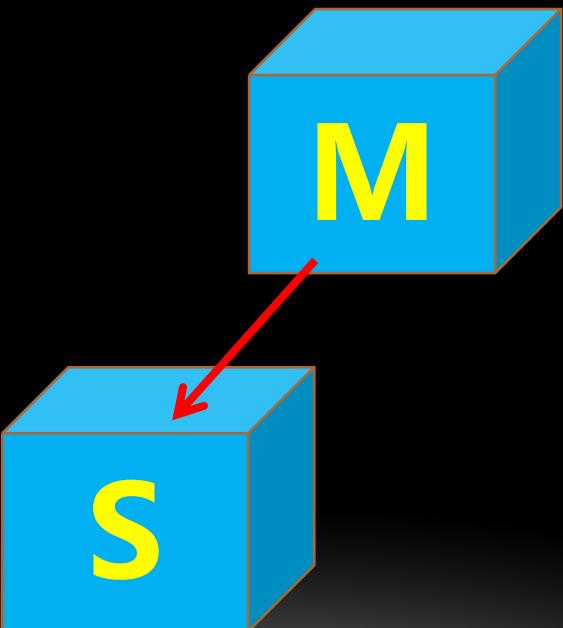
## 主从结构

※ 内存 96G

※ CPU 24C

※ 千兆网卡

※ 普通机械盘



※ 单实例 内存 60G

※ 多业务共享

※ data 和 cache 混用

※ 网卡未做 bond

# 架构演变

2

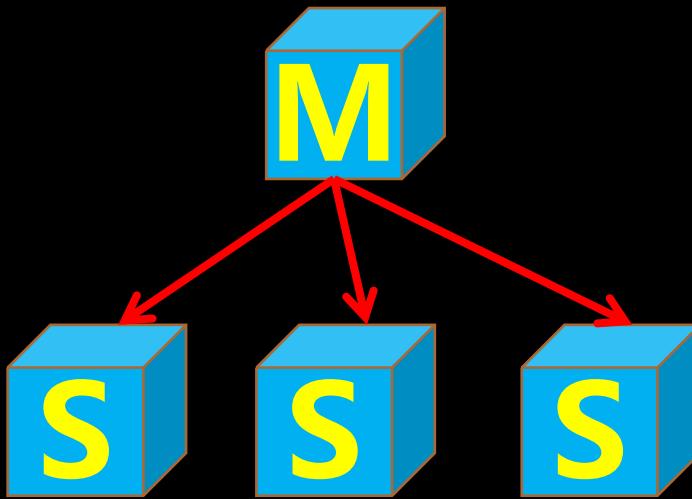
2

## 一主多从

※写master

※不同业务使用不同slave

※网卡bond



# 架构演变

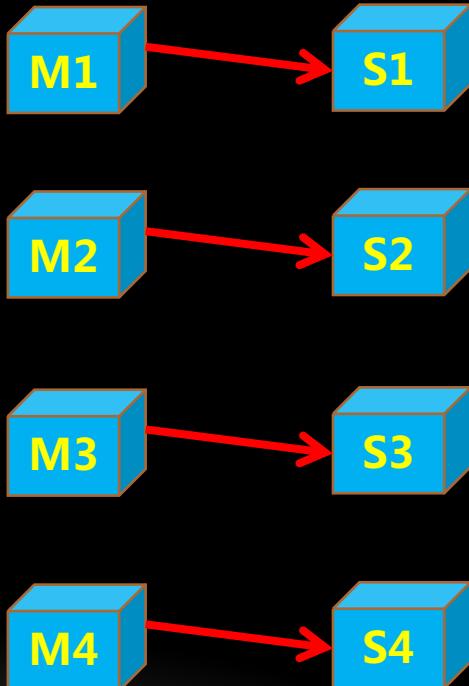
3

3

## 一机多实例

※ 单实例内存10~20G

※ 不同业务使用不同实例



# 架构演变

4

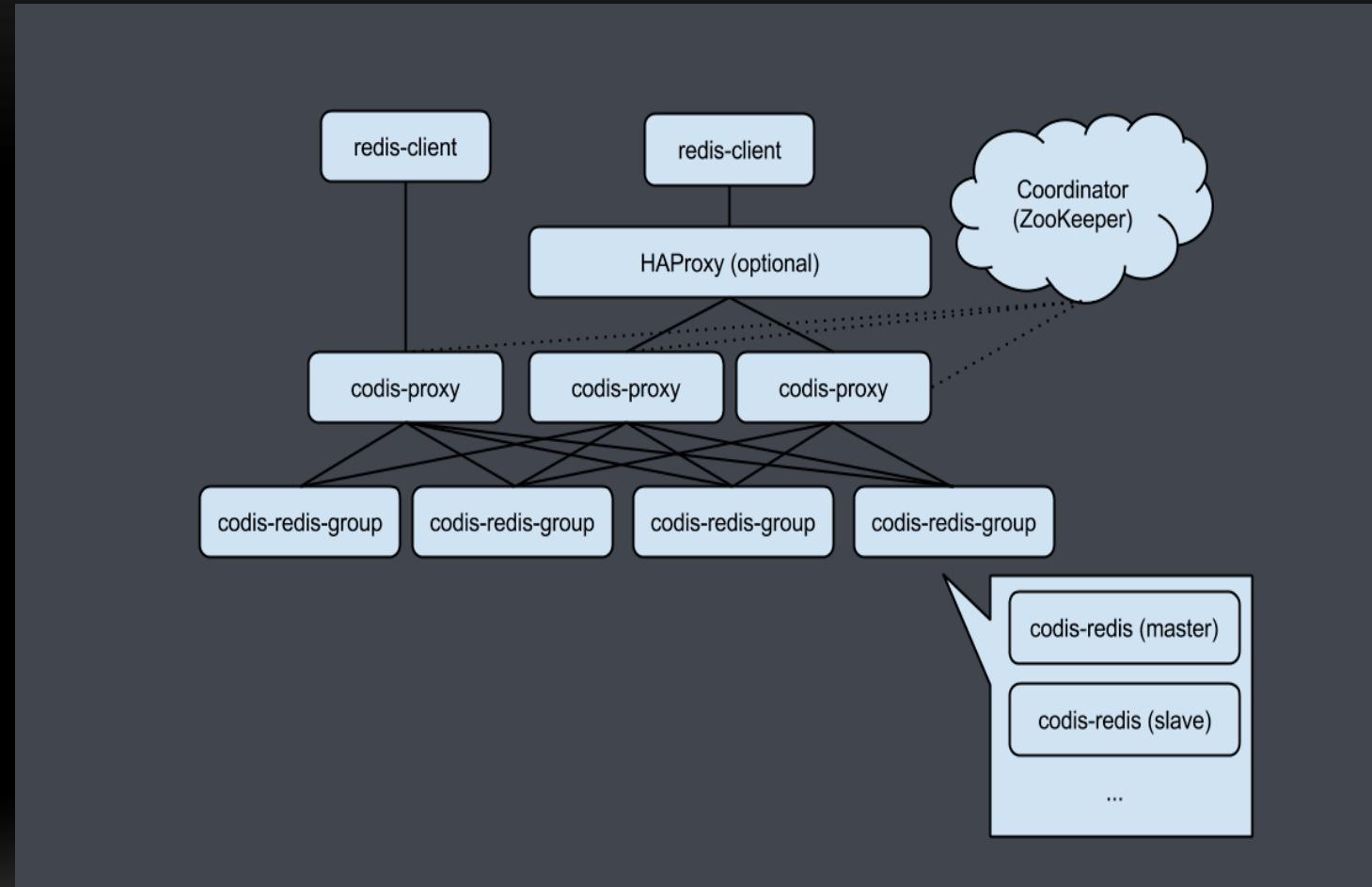
4

## 多实例分片

- ※ Codis
- ※ redis cluster
- ※ jumei redis组件

# codis

- ※ Codis Proxy (codis-proxy)
- ※ Codis Manager (codis-config)
- ※ Codis Redis (codis-server)
- ※ ZooKeeper



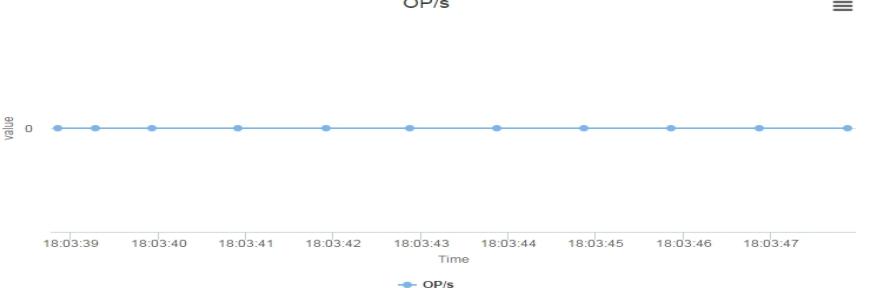
使用  $\text{crc32}(\text{key}) \% 1024$   
计算 Slot id

# codis

Overview 

Product Name:	CodisCluster001
Keys:	0
Mem Used:	446.70 MB
Performance:	0 OP/s

OP/s



Time: 18:03:39, 18:03:40, 18:03:41, 18:03:42, 18:03:43, 18:03:44, 18:03:45, 18:03:46, 18:03:47

OP/s

Server Groups 

+ New Server Group

group\_1

+ Add New Redis Instance 

Addr	Type	Mem Used	Keys
[redacted]	master	29.78M / 1 GB	

group\_2

+ Add New Redis Instance 

Addr	Type	Mem Used	Keys
[redacted]	master	29.78M / 1 GB	

Slot Control

Range Set (set new Group)  Slots Status

Migrate Status 

+ Auto Rebalance  Migrate Slot(s)

Migrate Task Info

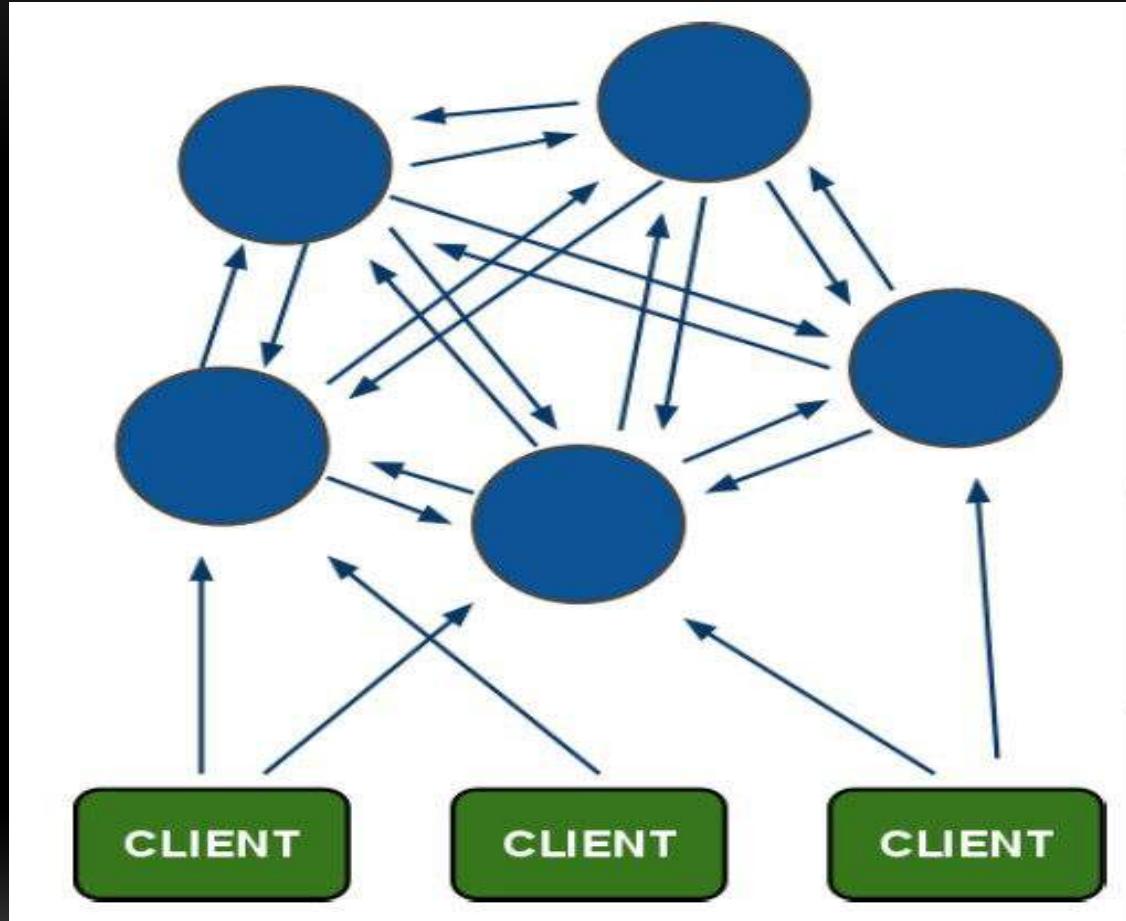
Proxy Status 

Proxy Name	Proxy Addr	Proxy DebugVars Addr	Proxy Status	Mark Online	Mark Offline
proxy_2	l-proxy2.codis.prod.cn1:19000	l-proxy2.codis.prod.cn1:11000	online		
proxy_1	l-proxy1.codis.prod.cn1:19000	l-proxy1.codis.prod.cn1:11000	online		
proxy_3	l-proxy3.codis.prod.cn1:19000	l-proxy3.codis.prod.cn1:11000	online		

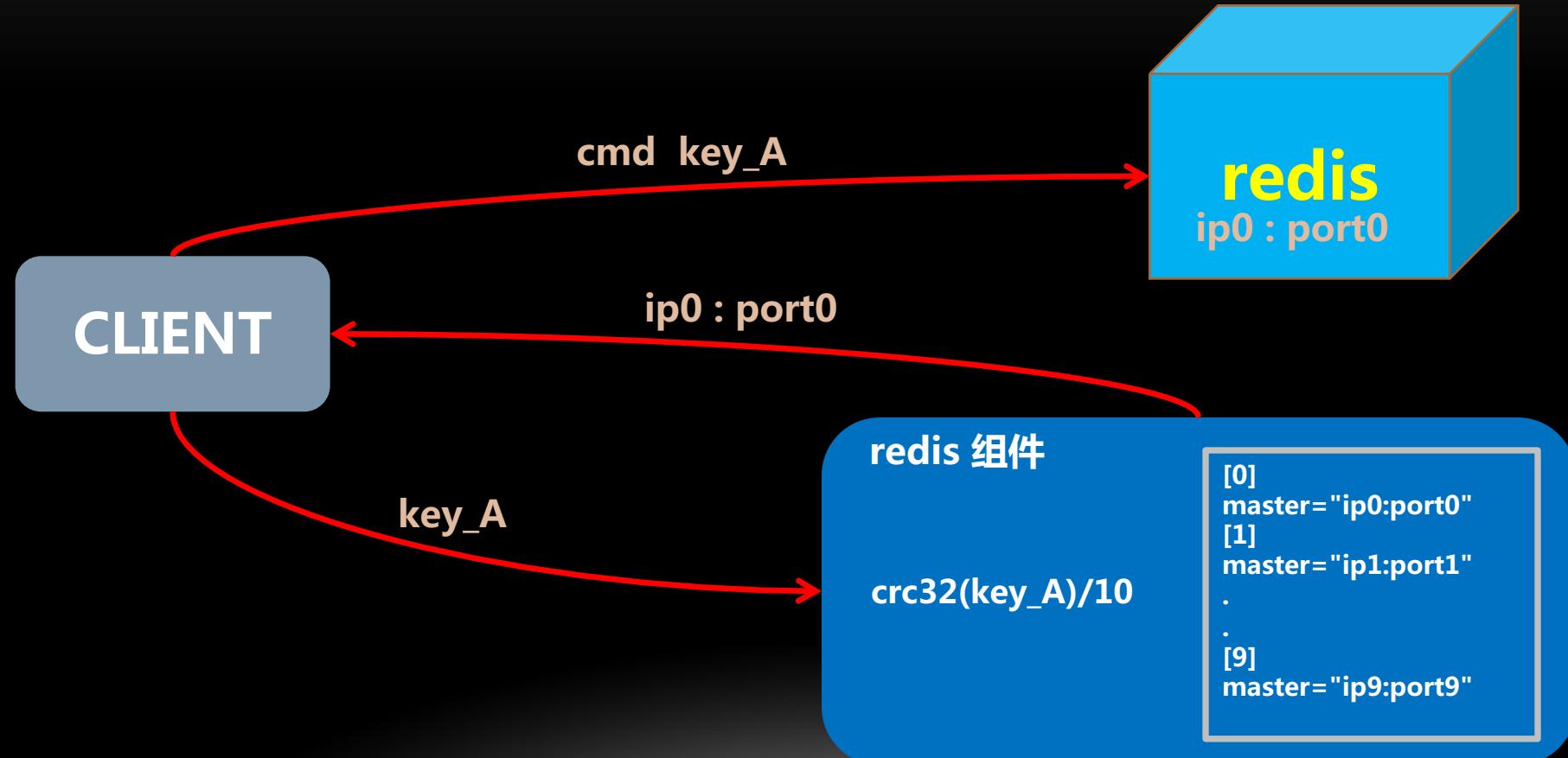
# Redis cluster

结构特点：

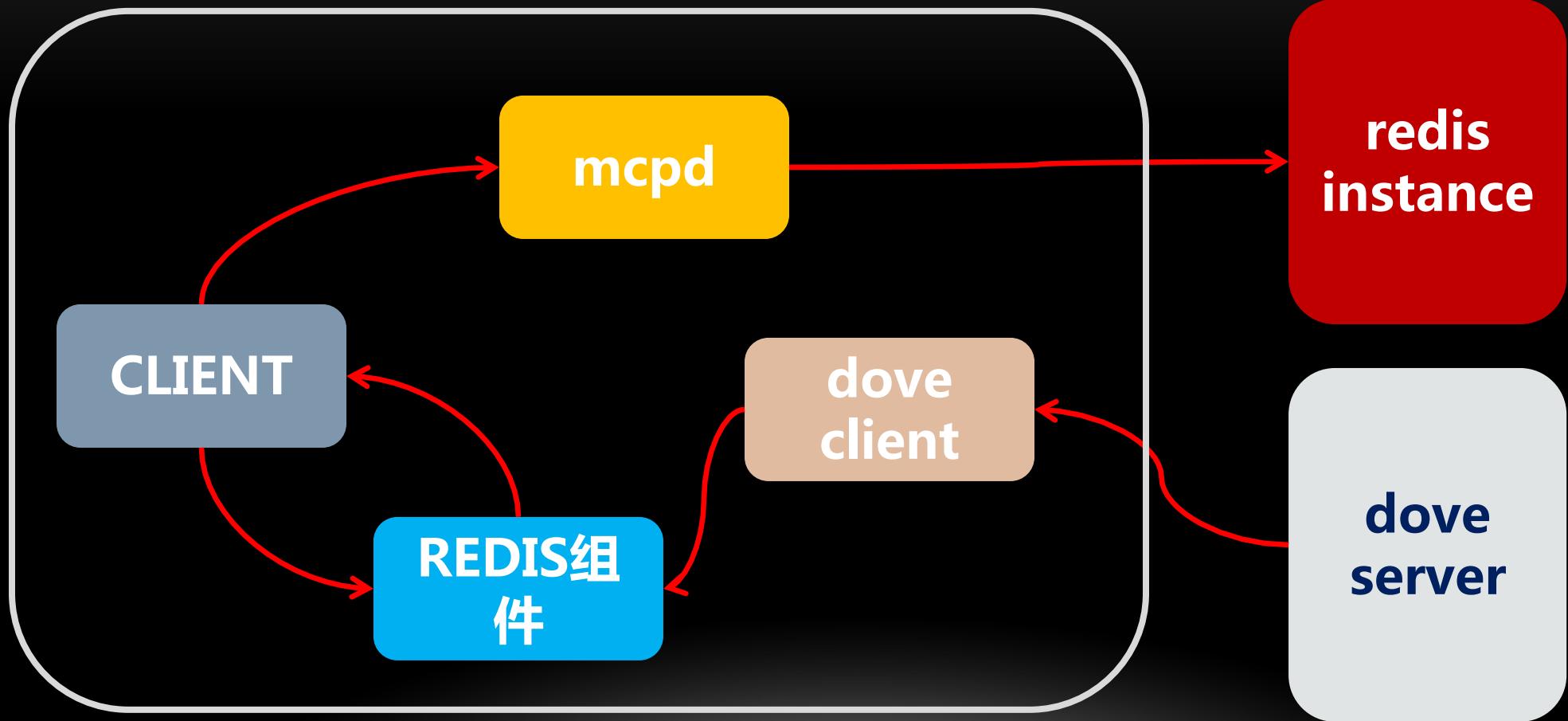
- ※ 所有的redis节点彼此互联(PING-PONG机制)
- ※ 节点的fail是通过集群中超过半数的节点检测失效时才生效
- ※ 客户端与redis节点直连,不需要中间proxy层.
- ※ 客户端不需要连接集群所有节点,连接集群中任何一个可用节点即可
- ※ redis-cluster把所有的物理节点映射到[0-16383]slot上 (不一定是平均分配 )
- ※ Redis集群预分好16384个槽位 , 当需要在 Redis 集群中放置一个 key-value 时 , 根据  $CRC16(key) \bmod 16384$  的值 , 决定将一个key放到哪个槽位中。



# jumei redis组件

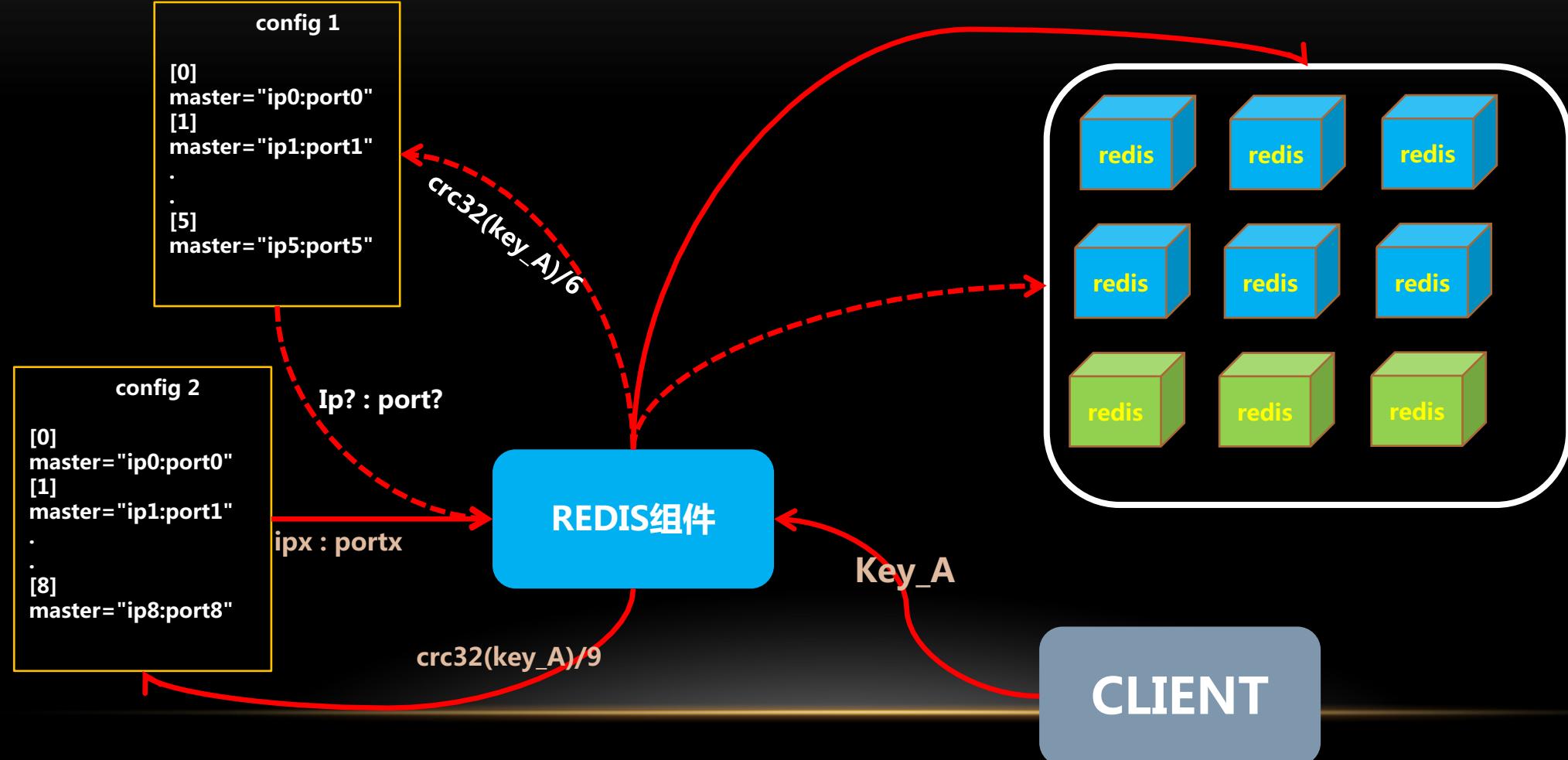


# jumei redis组件



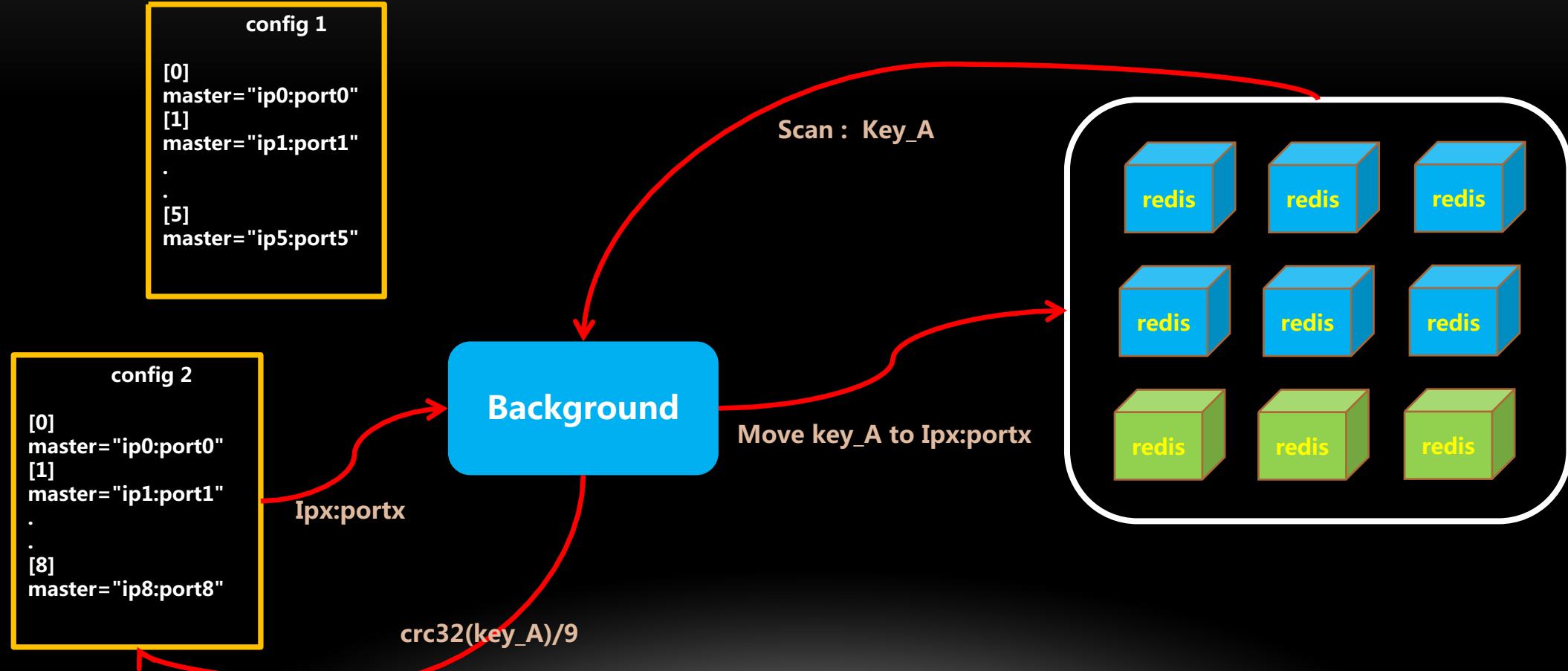
# 扩缩容

## 被动模式



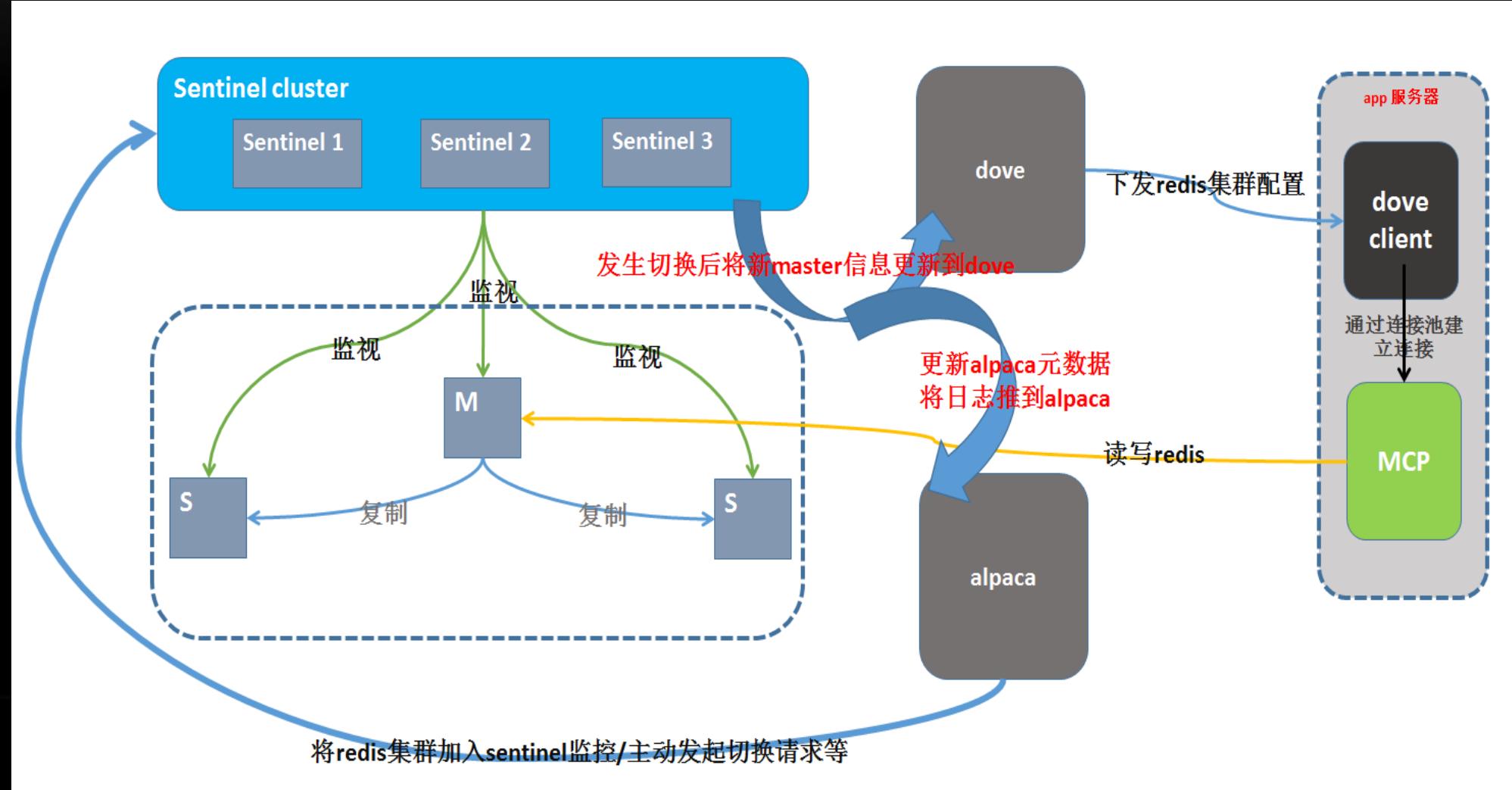
# 扩缩容

## 主动模式



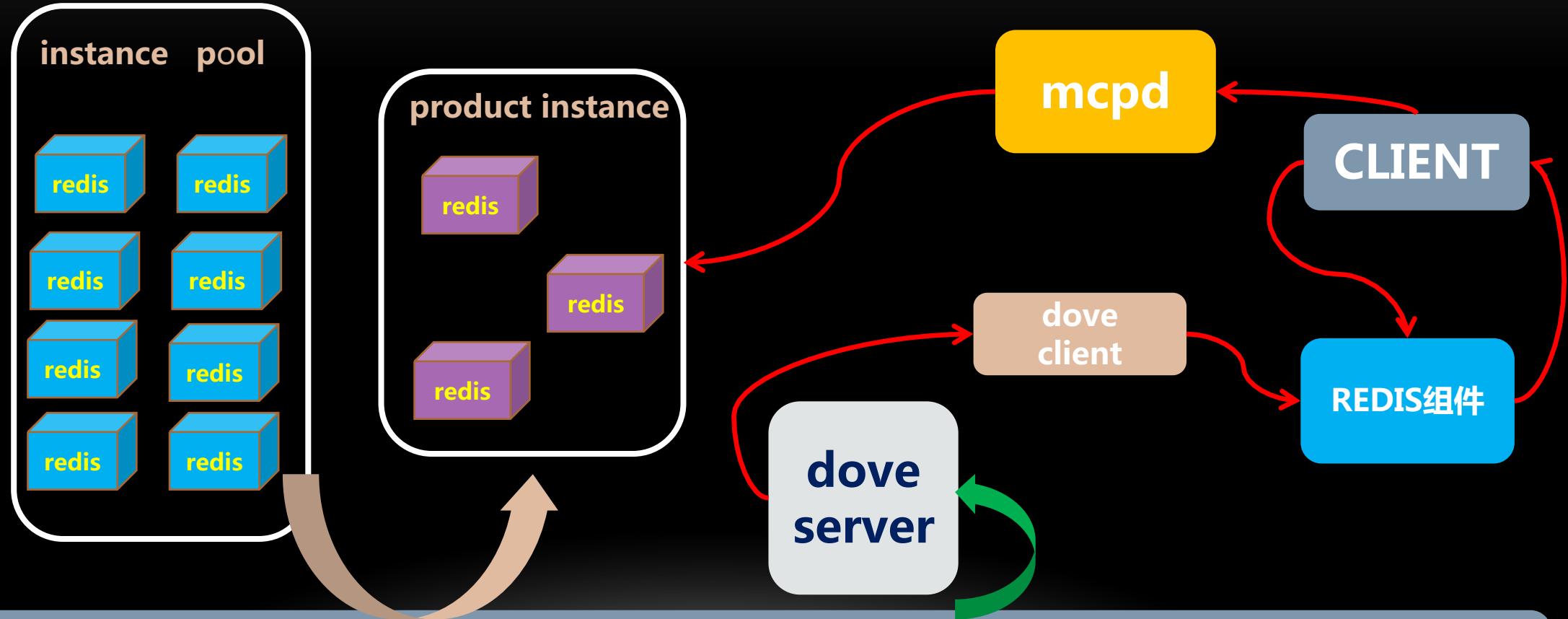
# 高可用

## Data集群高可用



# 高可用

Cache集群高可用

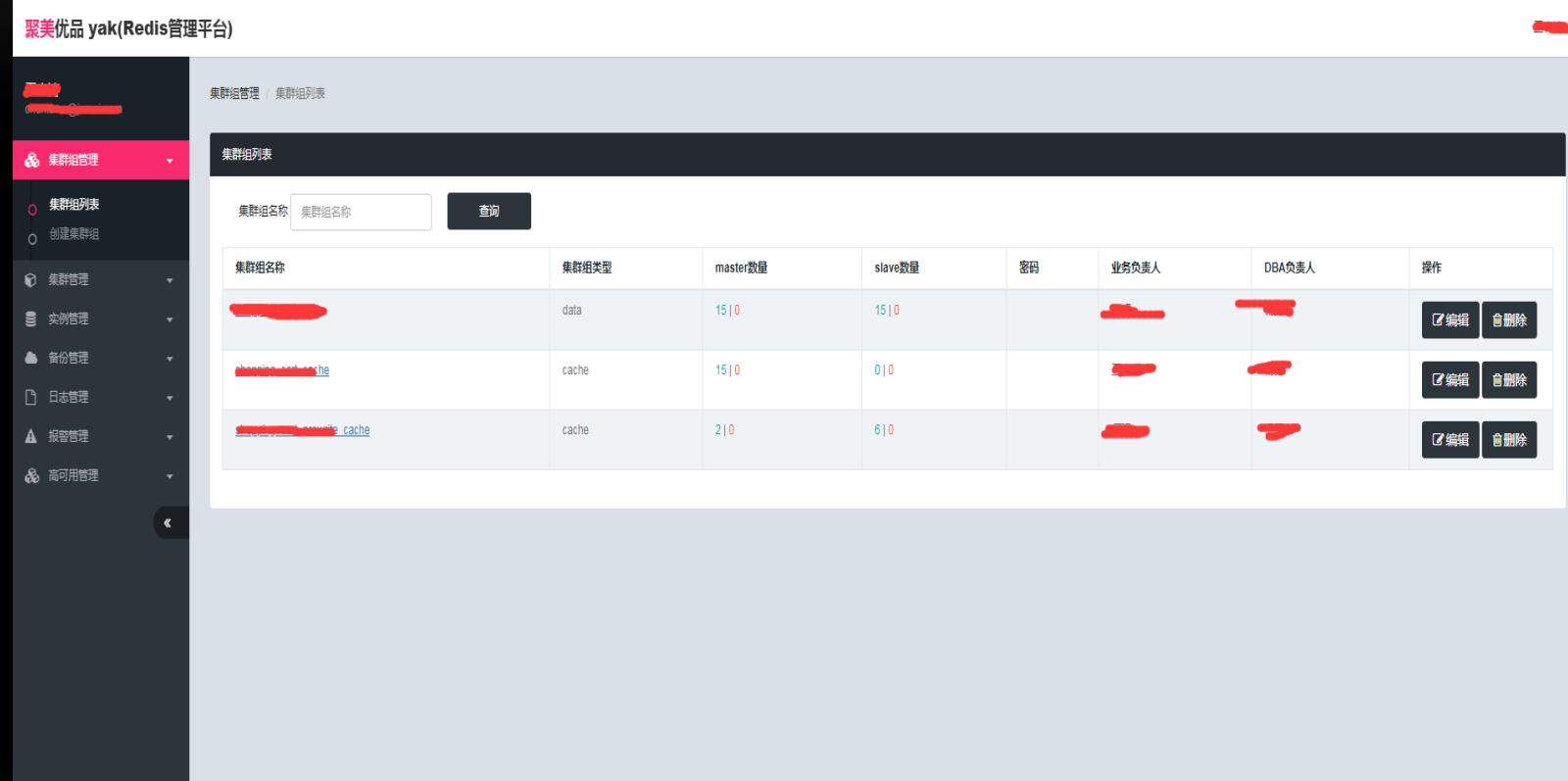


yak

# redis运维平台



■ 集群信息web化



集群组名称	集群组类型	master数量	slave数量	密码	业务负责人	DBA负责人	操作
[REDACTED]	data	15 0	15 0	[REDACTED]	[REDACTED]	[REDACTED]	<button>编辑</button> <button>删除</button>
chenmei-data-cache	cache	15 0	0 0	[REDACTED]	[REDACTED]	[REDACTED]	<button>编辑</button> <button>删除</button>
chenmei-cache	cache	2 0	6 0	[REDACTED]	[REDACTED]	[REDACTED]	<button>编辑</button> <button>删除</button>

■ 部署操作平台化

■ 重要指标实时化

# redis运维平台

聚美优品 yak(Redis管理平台)

王春涛 chuntaow@jumei.com

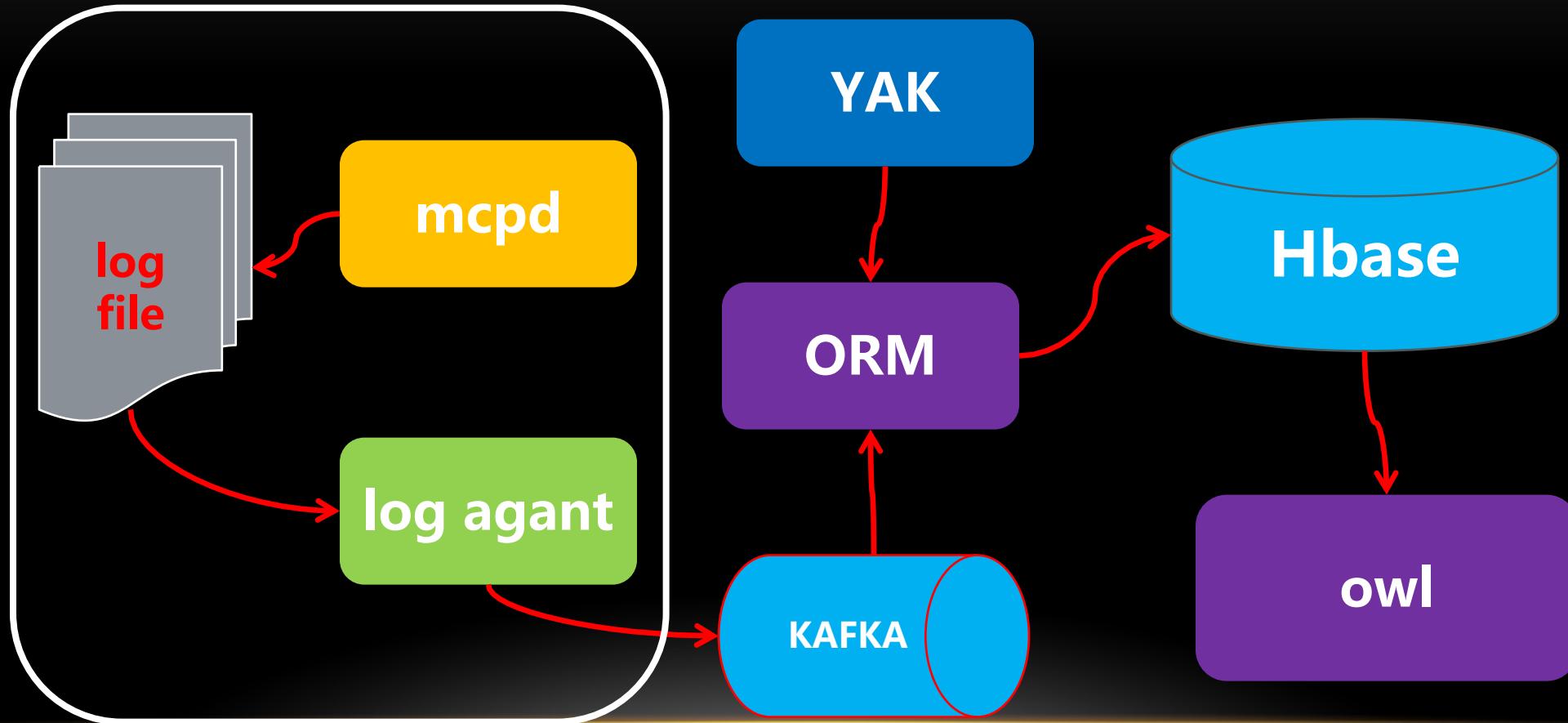
Sentinel管理 / Sentinel切换日志

Sentinel切换日志

Sentinel节点	集群名称	新Master实例	老Master实例	切换时间	详情
10.10.10.10	r_d	10.10.10.204	10.10.10.204	2018-04-20-01-58-57	<a href="#">查看</a>
10.10.10.11	e_d	10.10.10.204	10.10.10.204	2018-04-20-01-58-57	<a href="#">查看</a>
10.10.10.12	a_d	10.10.10.204	10.10.10.204	2018-04-20-01-58-57	<a href="#">查看</a>
10.10.10.13		10.10.10.204	10.10.10.204	2018-04-20-01-58-57	<a href="#">查看</a>
10.10.10.14	a_3	10.10.10.204	10.10.10.204	2018-04-20-01-58-57	<a href="#">查看</a>
10.10.10.15	d	10.10.10.204	10.10.10.204	2018-04-20-01-58-57	<a href="#">查看</a>
10.10.10.16	n	10.10.10.204	10.10.10.204	2018-04-20-01-58-57	<a href="#">查看</a>
10.10.10.17	m	10.10.10.204	10.10.10.204	2018-04-20-01-58-57	<a href="#">查看</a>
10.10.10.18	m	10.10.10.204	10.10.10.204	2018-04-20-01-58-56	<a href="#">查看</a>
10.10.10.19	n	10.10.10.202	10.10.10.202	2018-04-20-01-58-56	<a href="#">查看</a>
10.10.10.20	mob	10.10.10.202	10.10.10.202	2018-04-20-01-58-56	<a href="#">查看</a>
10.10.10.21	mob	10.10.10.202	10.10.10.202	2018-04-20-01-58-55	<a href="#">查看</a>
10.10.10.22	mobata_1	10.10.10.202	10.10.10.202	2018-04-20-01-58-55	<a href="#">查看</a>
10.10.10.23	m	10.10.10.202	10.10.10.202	2018-04-20-01-58-55	<a href="#">查看</a>
10.10.10.24		10.10.10.203	10.10.10.203	2018-04-19-15-38-00	<a href="#">查看</a>
10.10.10.25		10.10.10.20005	10.10.10.20005	2018-04-19-15-38-00	<a href="#">查看</a>
10.10.10.26		10.10.10.203	10.10.10.203	2018-04-19-15-38-00	<a href="#">查看</a>

# 精细化监控

ORM



# 精细化监控

## YAK 红绿灯

聚美优品 大促数据库监控

王春涛 ▾

Mysql 4 Redis 1 自选集群 Mysql慢查询 Redis慢查询 历史趋势 报表统计

机房

单节点 | M

连接数: 0 QPS: 0  
input\_kbps: 0 output\_kbps: 0 内存: 0GB | 0GB  
写入: failed used\_cpu\_user/used\_cpu\_sys: NaN  
监控: ERROR

payments - 0001 | M

连接数: 57 QPS: 9  
input\_kbps: 0.81 output\_kbps: 0.94 内存: 1.49GB | 10.0GB  
写入: ok used\_cpu\_user/used\_cpu\_sys: 1  
监控: OFF

payments - 0002 | M

连接数: 53 QPS: 23  
input\_kbps: 1.53 output\_kbps: 1.24 内存: 1.48GB | 10.0GB  
写入: ok used\_cpu\_user/used\_cpu\_sys: 1  
监控: OFF

payments - 0003 | S

连接数: 2 QPS: 12  
input\_kbps: 0.94 output\_kbps: 0.05 内存: 1.0GB | 10.0GB  
写入: read only used\_cpu\_user/used\_cpu\_sys: 0  
监控: OFF

payments - 0004 | M

连接数: 50 QPS: 15  
input\_kbps: 0.75 output\_kbps: 0.54 内存: 1.49GB | 10.0GB  
写入: ok used\_cpu\_user/used\_cpu\_sys: 1  
监控: OFF

payments - 0005 | S

连接数: 2 QPS: 9  
input\_kbps: 0.45 output\_kbps: 0.05 内存: 1.0GB | 10.0GB  
写入: read only used\_cpu\_user/used\_cpu\_sys: 0  
监控: OFF

payments - 0006 | M

连接数: 57 QPS: 19  
input\_kbps: 1.42 output\_kbps: 1.04 内存: 1.49GB | 10.0GB  
写入: ok used\_cpu\_user/used\_cpu\_sys: 1  
监控: OFF

payments - 0007 | S

连接数: 2 QPS: 15  
input\_kbps: 1.06 output\_kbps: 0.03 内存: 1.0GB | 10.0GB  
写入: read only used\_cpu\_user/used\_cpu\_sys: 0  
监控: OFF

机房

payments - 0008 | M

连接数: 200 QPS: 49  
input\_kbps: 25.34 output\_kbps: 73.52 内存: 8.08GB | 15.0GB  
写入: ok used\_cpu\_user/used\_cpu\_sys: 2  
监控: OFF

payments - 0009 | M

连接数: 1743 QPS: 39  
input\_kbps: 0.98 output\_kbps: 0.87 内存: 0.12GB | 8.0GB  
写入: ok used\_cpu\_user/used\_cpu\_sys: 0  
监控: OFF

payments - 0010 | S

连接数: 9 QPS: 30  
input\_kbps: 26.46 output\_kbps: 0.06 内存: 8.05GB | 15.0GB  
写入: read only used\_cpu\_user/used\_cpu\_sys: 0  
监控: OFF

payments - 0011 | M

连接数: 204 QPS: 61  
input\_kbps: 31.31 output\_kbps: 58.57 内存: 8.06GB | 15.0GB  
写入: ok used\_cpu\_user/used\_cpu\_sys: 2  
监控: OFF

payments - 0012 | S

连接数: 10 QPS: 15  
input\_kbps: 5.32 output\_kbps: 0.06 内存: 8.05GB | 15.0GB  
写入: read only used\_cpu\_user/used\_cpu\_sys: 0  
监控: OFF

payments - 0013 | M

连接数: 200 QPS: 33  
input\_kbps: 7.09 output\_kbps: 35.38 内存: 8.06GB | 15.0GB  
写入: ok used\_cpu\_user/used\_cpu\_sys: 2  
监控: OFF

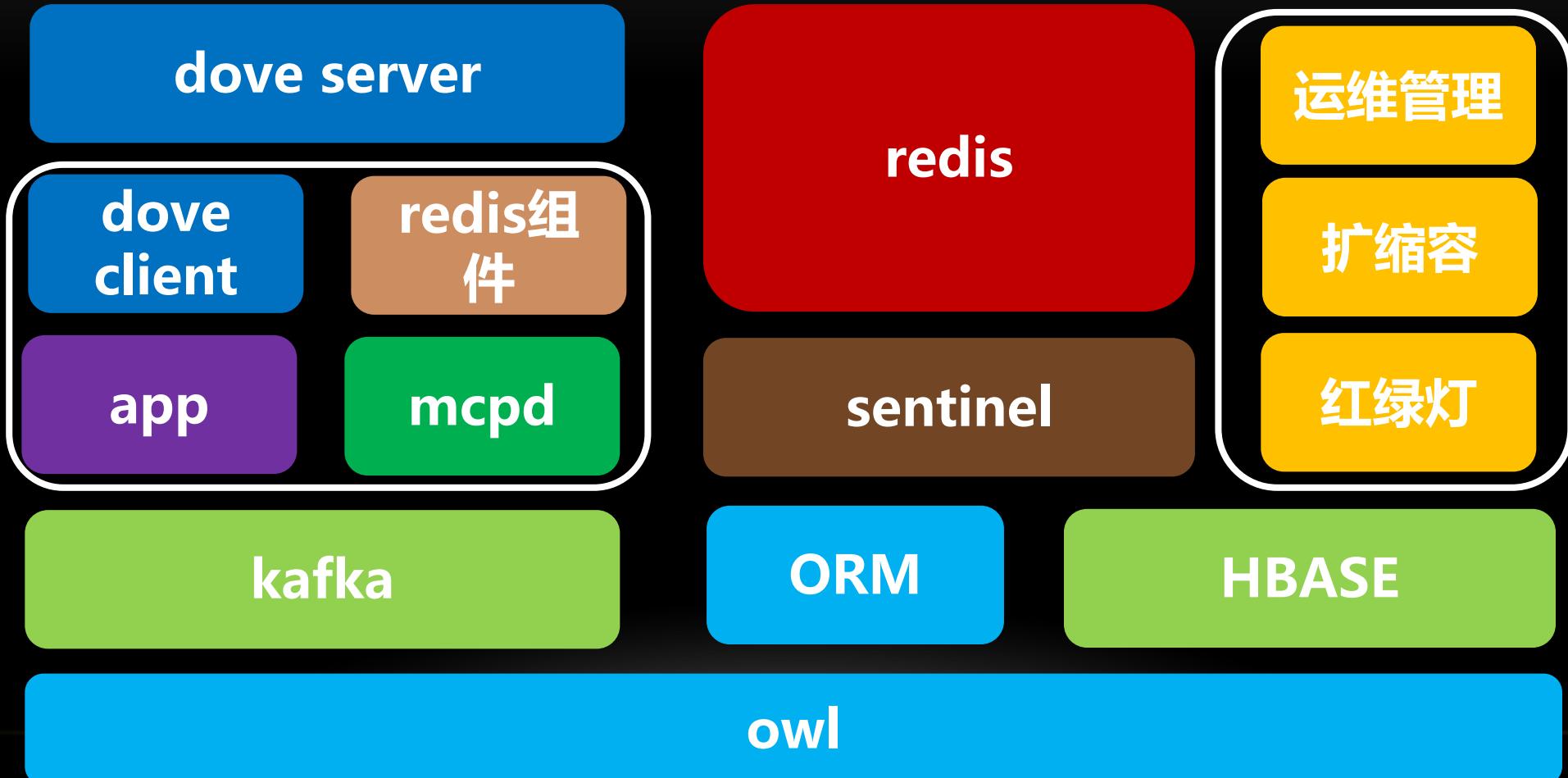
payments - 0014 | S

连接数: 10 QPS: 32  
input\_kbps: 20.44 output\_kbps: 0.06 内存: 8.05GB | 15.0GB  
写入: read only used\_cpu\_user/used\_cpu\_sys: 0  
监控: OFF

payments - 0015 | M

连接数: 213 QPS: 70  
input\_kbps: 23.39 output\_kbps: 71.12 内存: 8.05GB | 15.0GB  
写入: ok used\_cpu\_user/used\_cpu\_sys: 2  
监控: OFF

# 整体结构





THANK YOU