



微信扫码收听演讲音频



Zookeeper 异地多活自动化运维实践

主讲人：赵欢

时间：2018.09.15

拉扎斯网络科技(上海)有限公司
Rajax Network & Technology Co.



微信扫码收听演讲音频

目录

- 1 Zookeeper 初识
- 2 饿了么 异地多活
- 3 运维的挑战&运维实践
- 4 EZK运维管理平台实践



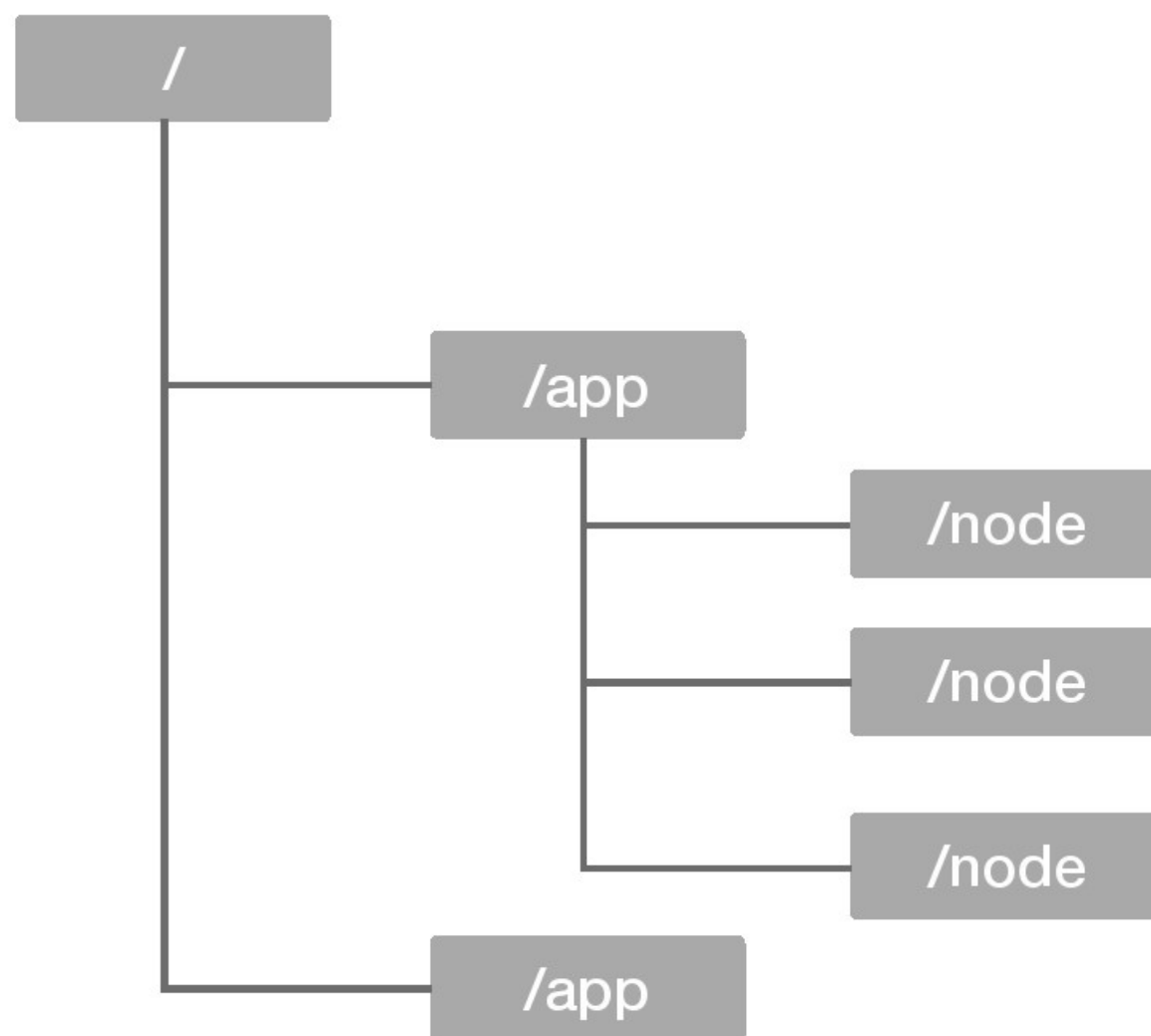
微信扫码收听演讲音频

1 *Zookeeper* 初识



ZooKeeper 简介

- ZooKeeper是一个分布式的，开放开源的高可靠分布式应用程序协调服务
- 使用zab协议数据多副本，Paxos协议改进版一致性的算法
- 数据结构类似于标准文件系统，它是树状数据结构
- 服务注册/发现（配置管理），命名服务，集群管理，分布式锁等





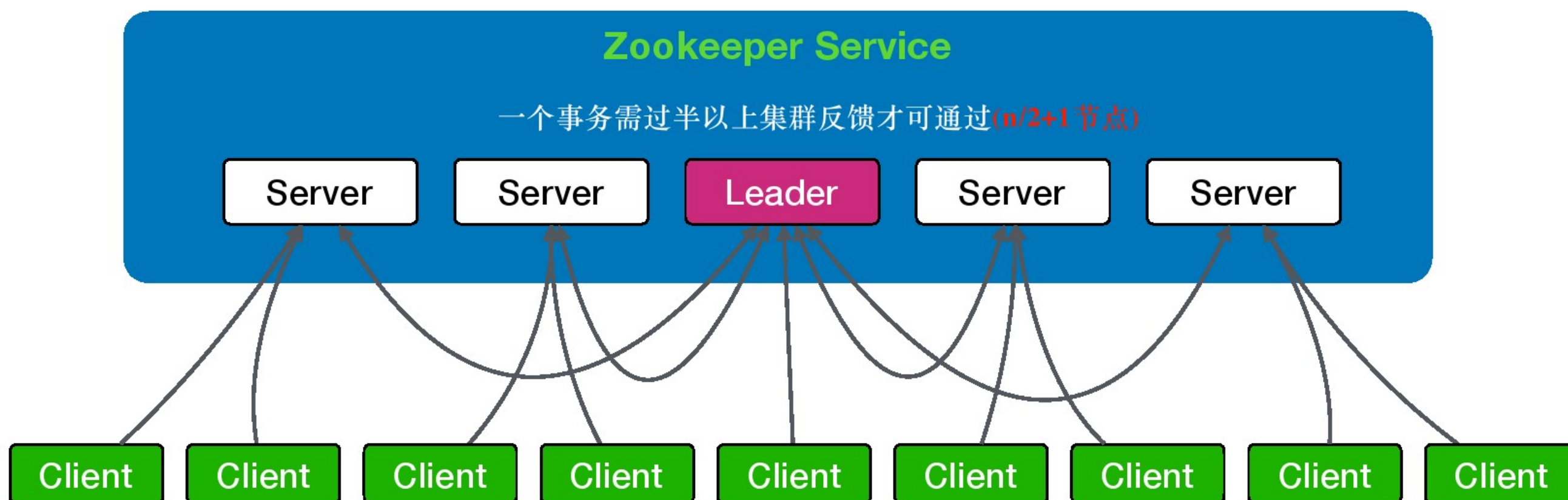
微信扫码收听演讲音频

ZooKeeper 运行机制

Leader (领导者) ZooKeeper 集群主节点，负责整个集群事务处理。

Follower (跟随者) 作为客户端提供读服务，如果是写服务则转发给 Leader，在选举过程中参与投票。

Observer (观察者) 作为客户端提供读服务，如果是写服务则转发给 Leader，不参与选举过程中的投票。





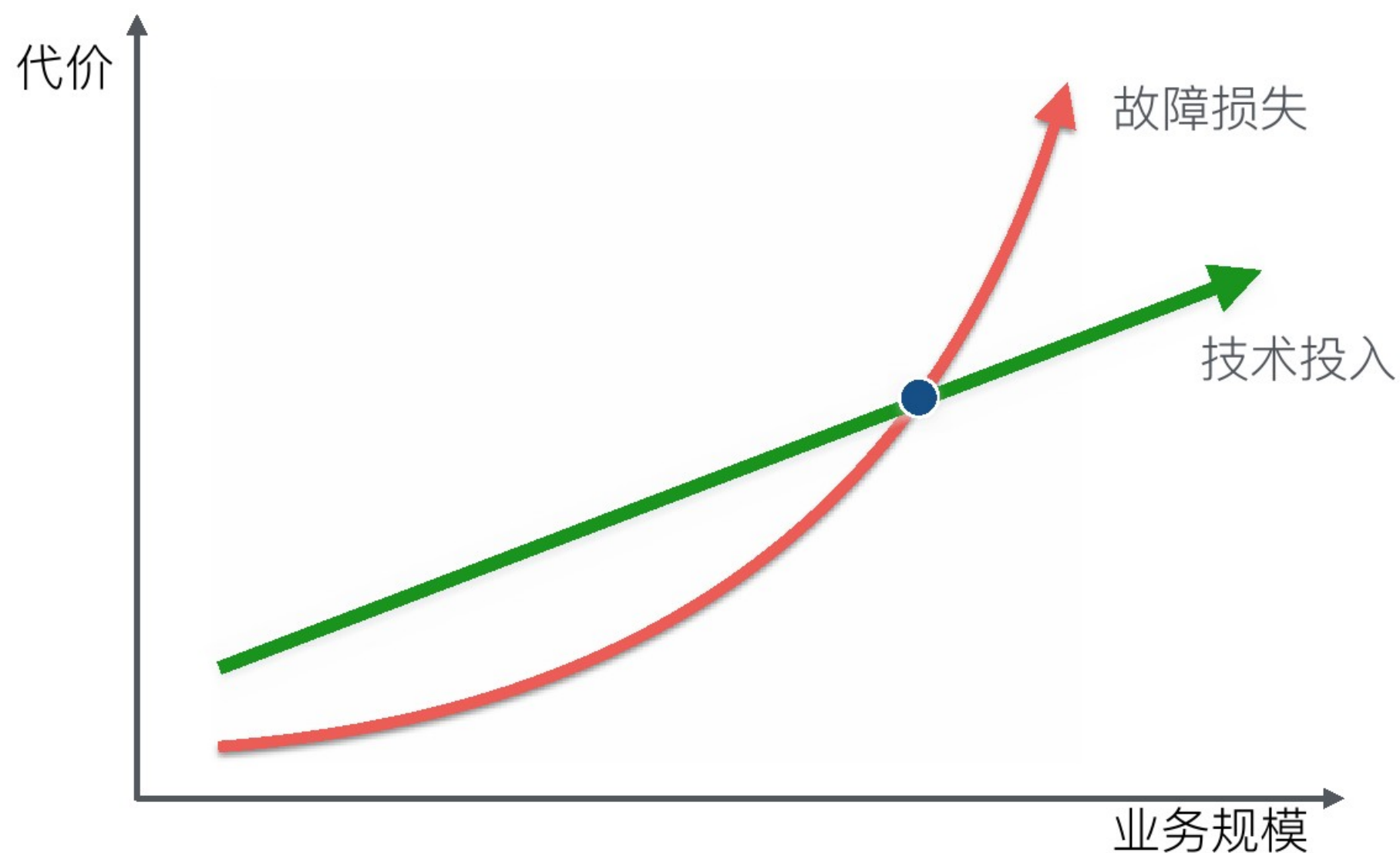
微信扫码收听演讲音频

为什么要做多活？
多机房多活带来那的挑战？

2 饿了么 异地多活



做多活的意义

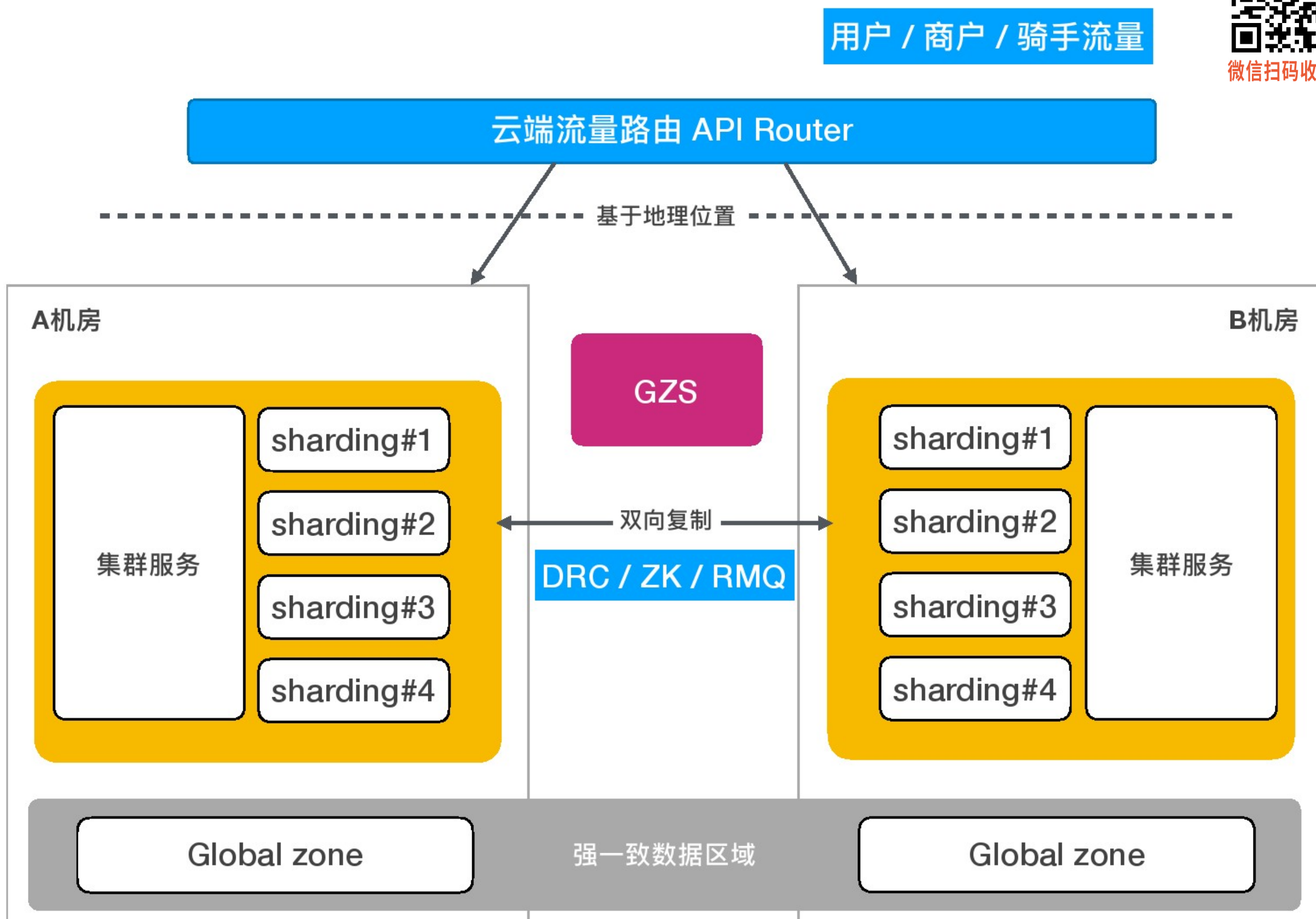


1. 随着业务的增长，会达到一个临界点，故障的损失超过多活的投入
2. 单个机房的容量已经无法承载业务发展，需要突破单机房物理限制

多活整体结构



微信扫码收听演讲音频





微信扫码收听演讲音频

- 单IDC Watch数量 > 亿级数量
- 单IDC TCP有效连接 > 5万+
- 多活节点 Node 节点数量 > 百万+

3

运维的挑战&运维实践



微信扫码收听演讲音频

Zookeeper 多活应用场景

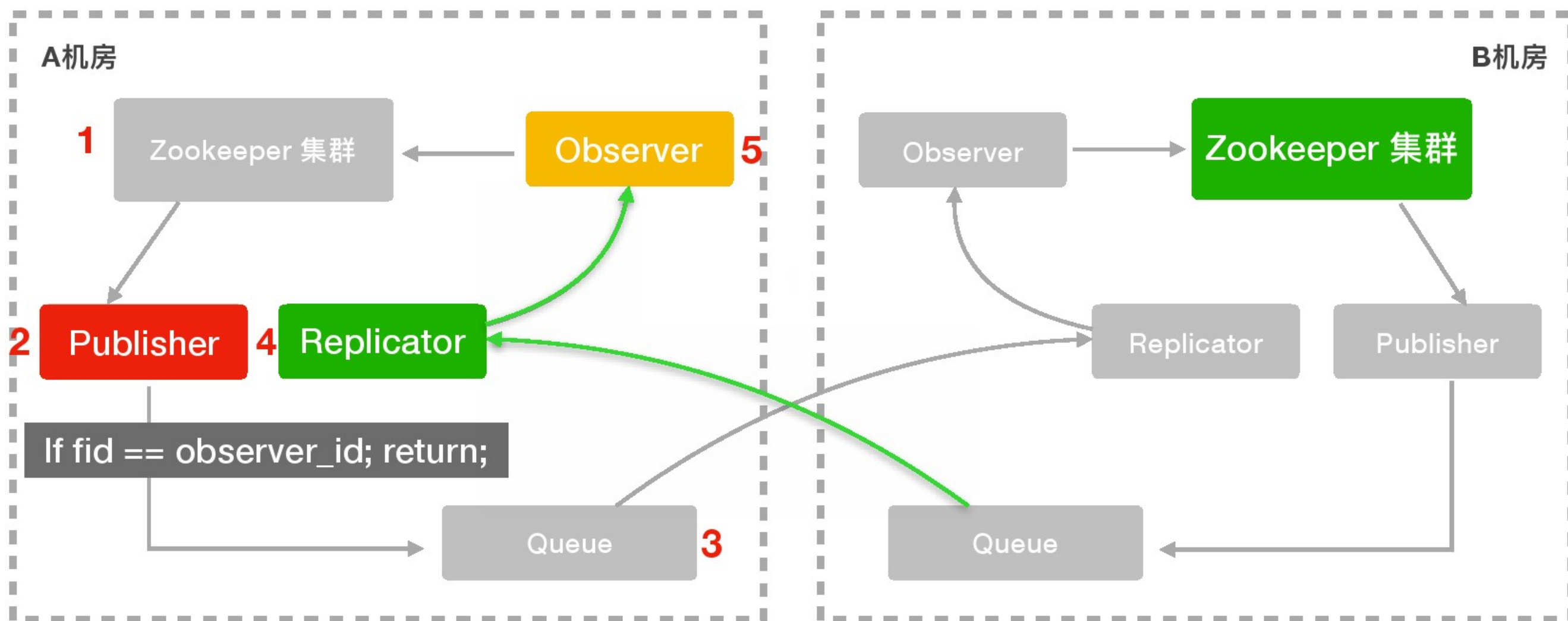
- 数据访问层控制层的源数据
- SOA 服务注册与集群信息
- 负载均衡中间件
- 策略开关/配置开关



微信扫码收听演讲音频

Zookeeper 跨机房复制架构

- 多机房同时写入，如何避免循环复制？



- 1.原生集群对外提供Zookeeper服务
- 2.基于Zookeeper开源代码开发，ZK事物聚合同步组件
- 3.消息列队组件Rabbitmq服务
- 4.基于Java开发的多机房事物消费服组件
- 5.原生集群Observer角色，负责写入其他机房ZK事物



微信扫码收听演讲音频

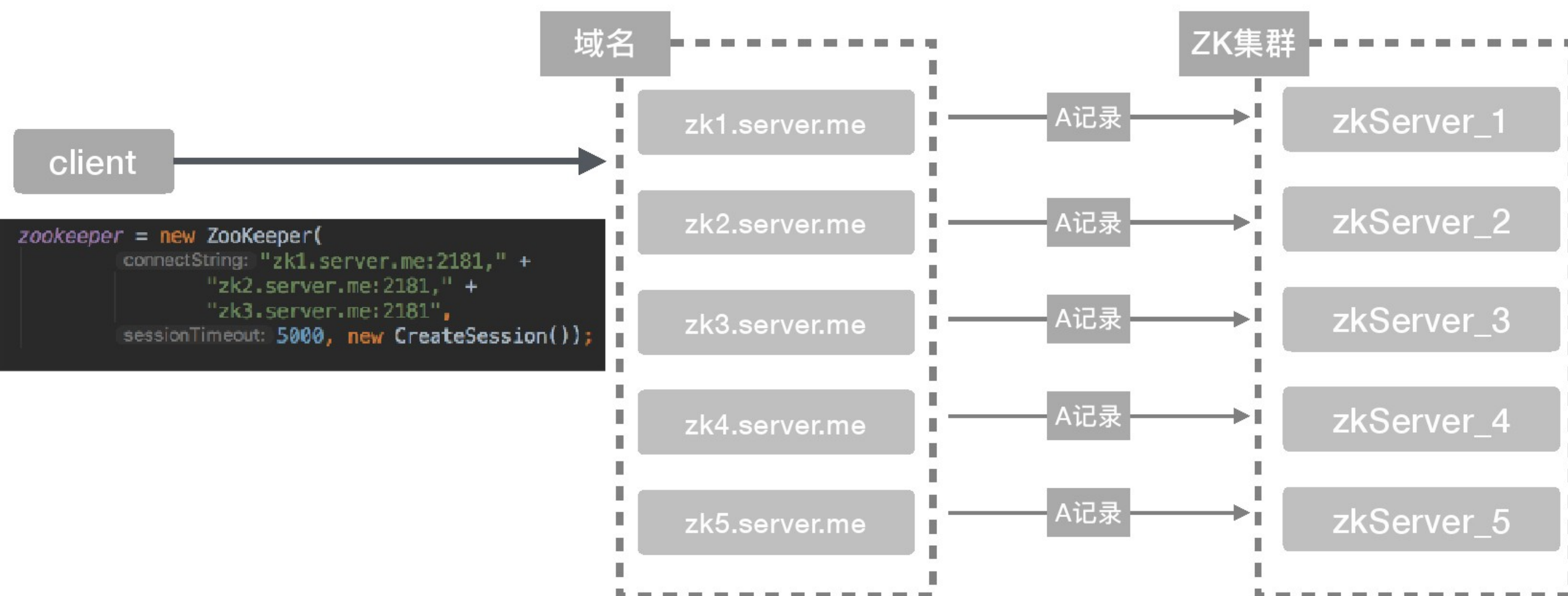
运维实践

- 机柜物理隔离，虚拟机宿主机隔离
- 客户端连接池采用域名配置
- Ha用户无感知切换方案
- 性能/状态监控
- 同步检测/集群功能检测



微信扫码收听演讲音频

域名配置



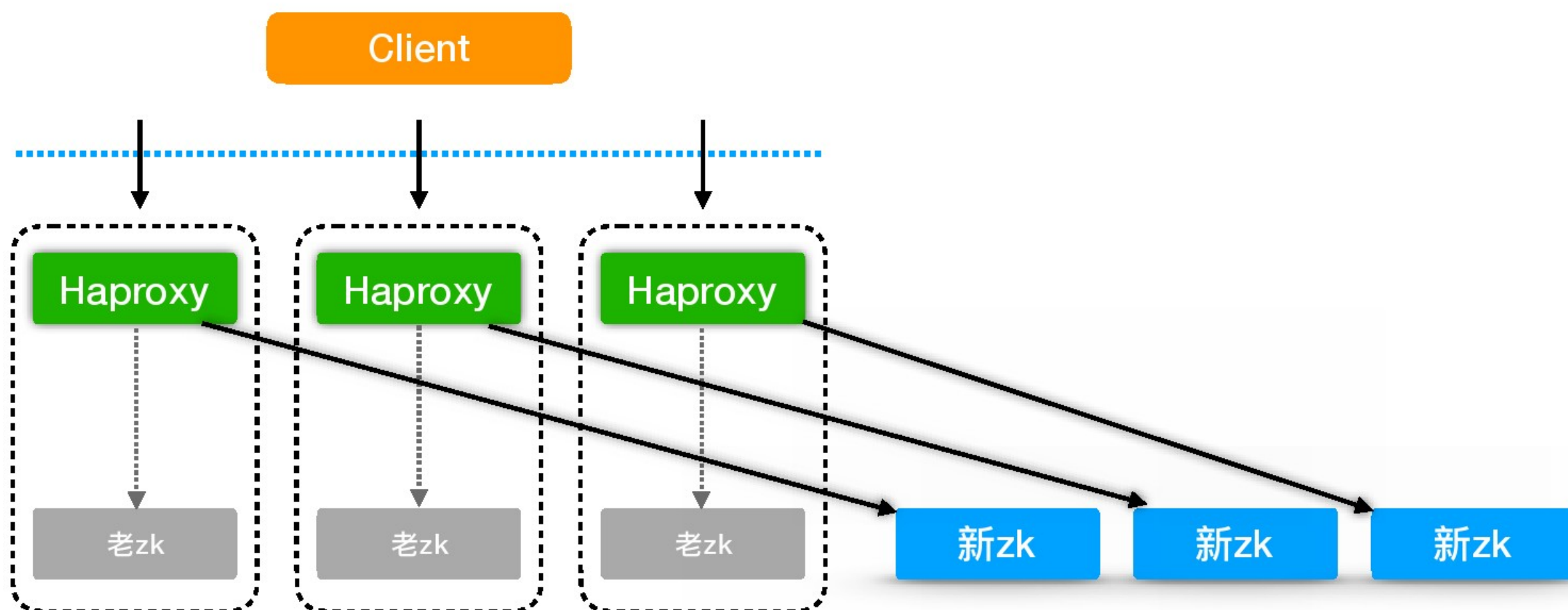


微信扫码收听演讲音频

HaProxy 切换

客户端不重启切集群方案

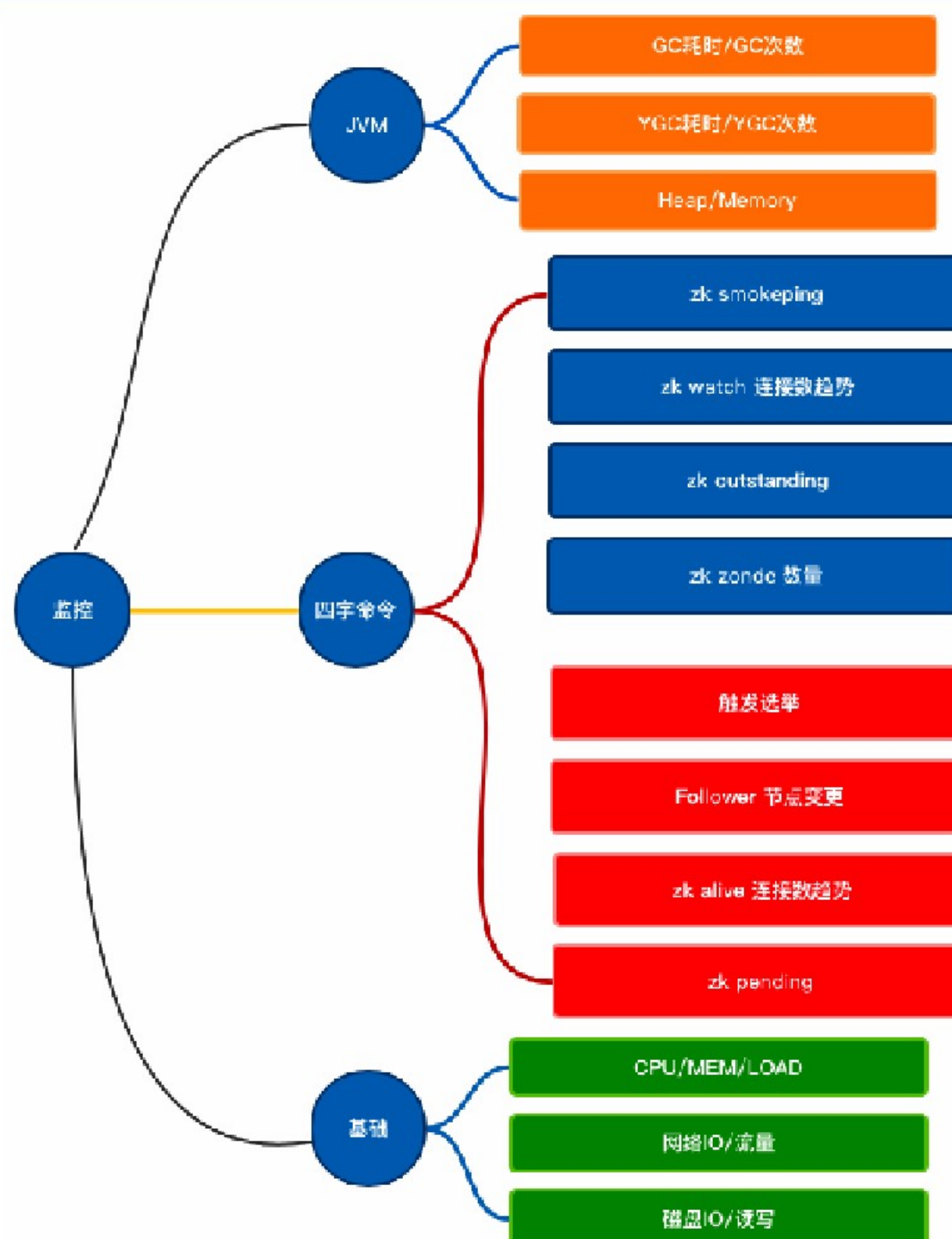
避免线上成千上万个应用重启





微信扫码收听演讲音频

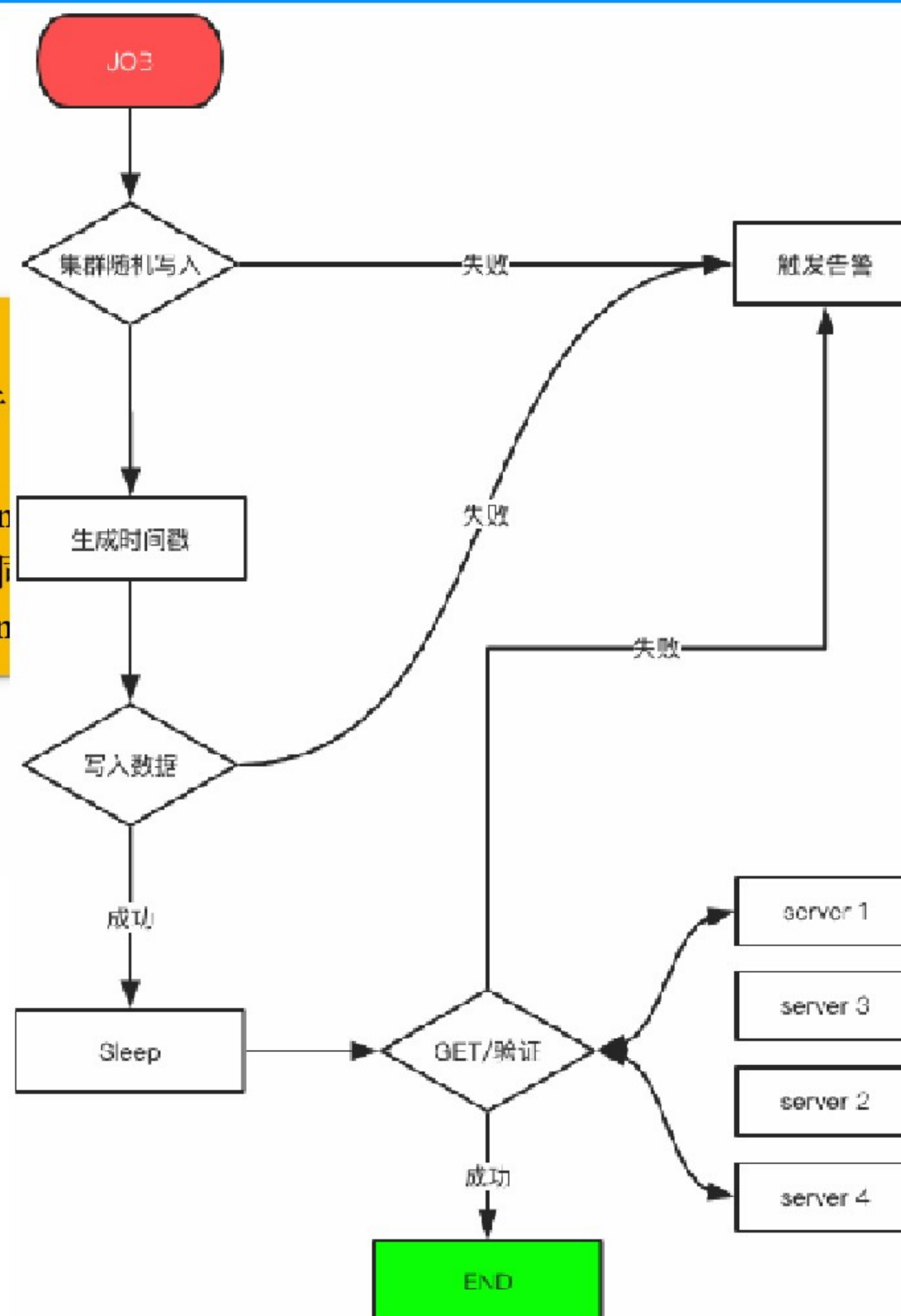
性能/状态





功能监控

1. 定时任务启动
2. 随机选取一台Zk进行
3. 生成当前时间戳 ()
4. 写入数据 (/Monitor/n
5. Sleep /秒, 等待数据同
6. GET数据 (/Monitor/n





微信扫码收听演讲音频

4

*EZK*运维管理平台实践



微信扫码收听演讲音频

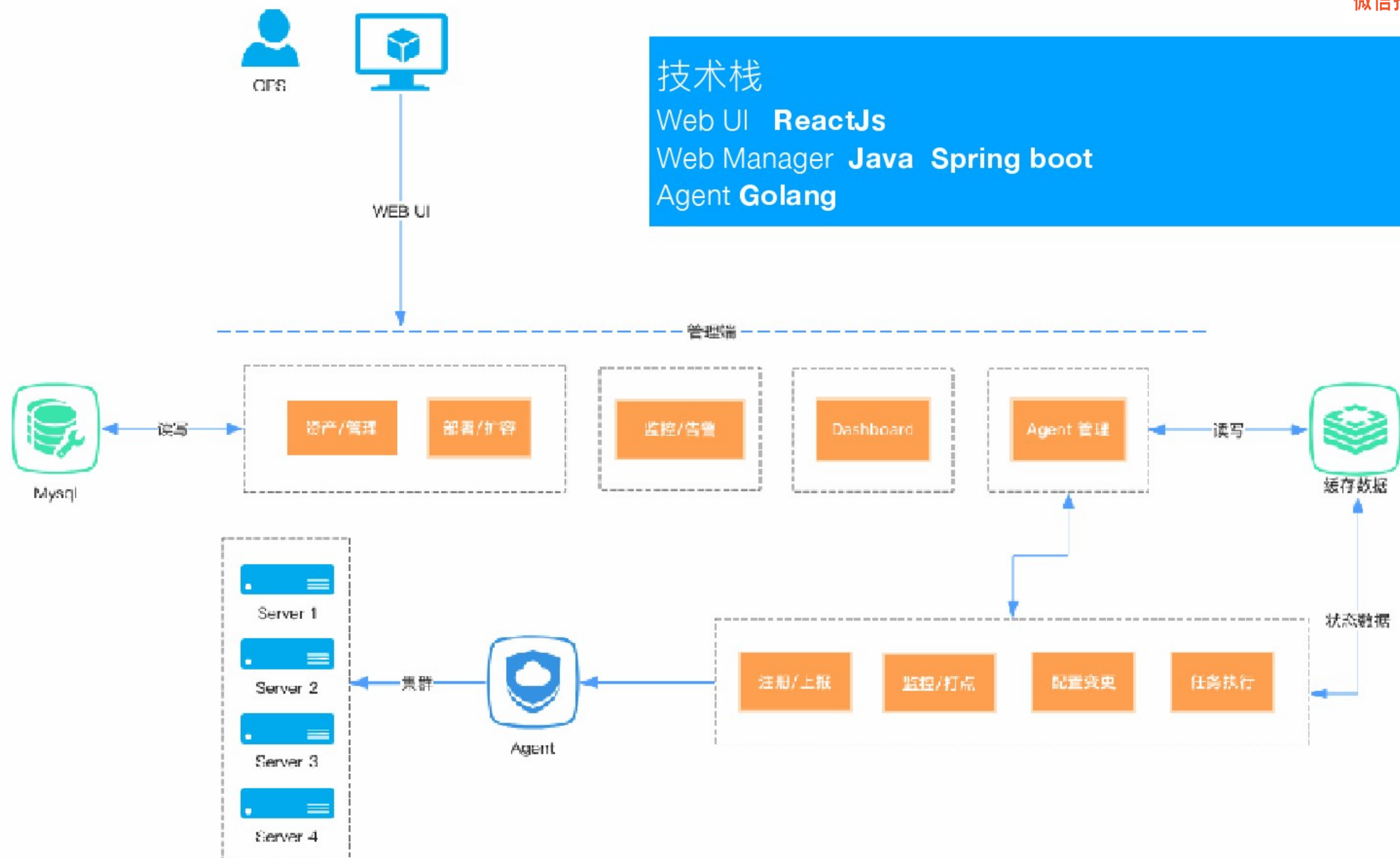
功能简介

- EZK（全称Eleme Zookeeper）它是一套Zookeeper定制化维护管理自动化平台
- 集群管理/资产管理一览
- Zookeeper 运行状态监控/状态信息聚合打点
- 集群快速部署/配置管理

系统架构图



微信扫码收听演讲音频





微信扫码收听演讲音频

测试环境()

4

2

2

2

Role	ID	Heart	Leader	AliveCount	WatchCount	Status	Agreement	状态/位置/结构/连接
follower	102			202	1052455	open	open	
follower	101			408	2032290	open	open	
follower	103			412	2132451	open	open	
leader	104			125	265531	open	open	
follower	105			14	277115	open	open	

conf

ANTR

STAT

CONF

role

alive_count

outstanding

avg_delay

zk_node_count

zk_watch_count

zk_opn_ticount

last_time

uptime

version

follower

14

0

0

310219

27785

49

2018-09-15 13:22:42

108.650548ms

3.4.9-1757319, built on 05/23/2018 06:50 GMT



微信扫码收听演讲音频

1

2

3

4

HostName

IPADDR

ZID

Port

Role

更多参数

chat-undak-1

10.103.101.95

5

2181

FOLLOWER

View parameter Settings

RetTime

1000

...

retLimit

10

...

syncLimit

5

...

dataDir

/data/consensus/etcd

...

dataLogDir

/data/consensus/etcd

...

maxOfEventCounts

100

...

autoPurge_anaPresumCount

1000

...

autoPurge_purgeInterval

1

...

添加节点

节点列表

节点配置

id	HostName	ipaddr	Port	Role	Actions
1	chat-undak-1	10.103.101.95	2181	FOLLOWER	<div><div></div><div></div></div>
2	chat-undak-2	10.103.101.96	2181	FOLLOWER	<div><div></div><div></div></div>
3	chat-undak-3	10.103.101.97	2181	FOLLOWER	<div><div></div><div></div></div>
4	chat-undak-4	10.103.101.98	2181	FOLLOWER	<div><div></div><div></div></div>
5	chat-undak-5	10.103.101.99	2181	FOLLOWER	<div><div></div><div></div></div>

© 2024 极水网络



微信扫码收听演讲音频



Jvm 监控

Total GC Count

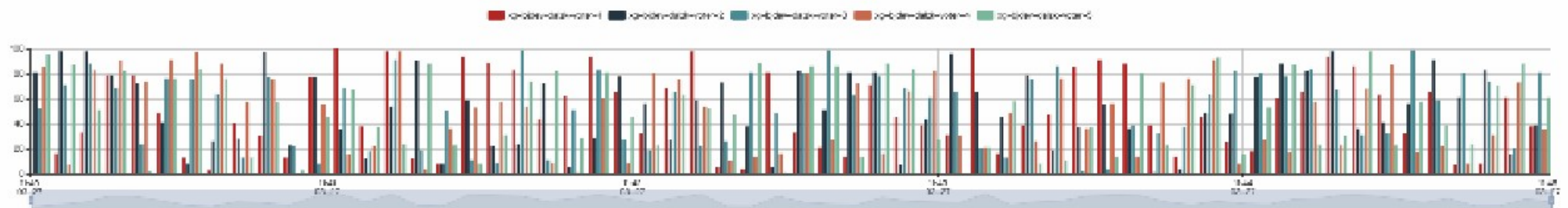


- sg-hjtem-stable-volter-1
- sg-hjtem-stable-volter-2
- sg-hjtem-stable-volter-3
- sg-hjtem-stable-volter-4
- sg-hjtem-stable-volter-5

时段GC数量占比



Total GC time



数据更新



微信扫码收听演讲音频

End<-**THANKS**

