

SERVICE MESH MEETUP #6 广州站

虎牙直播在微服务改造中的实践

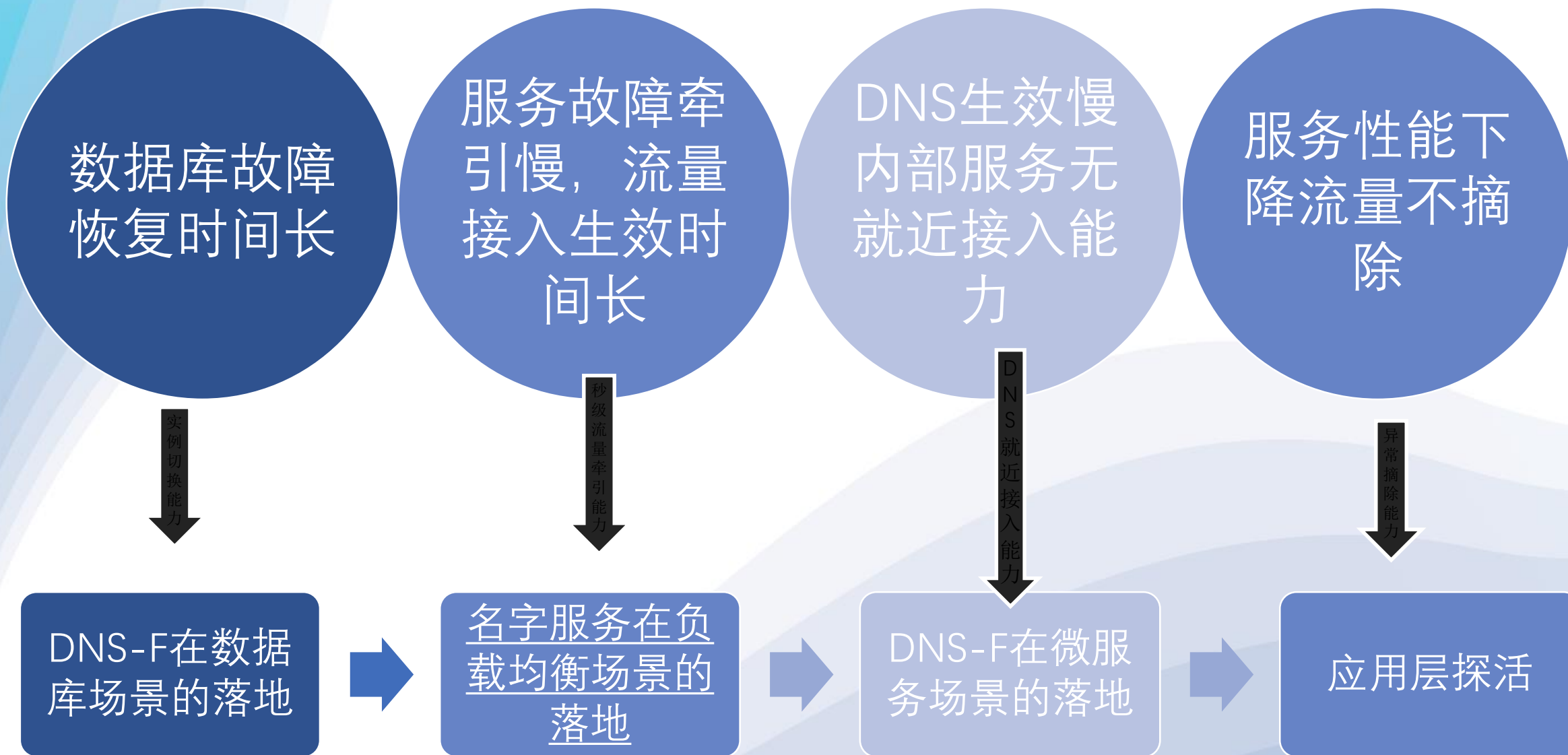
张波

虎牙基础保障部中间件团队负责人

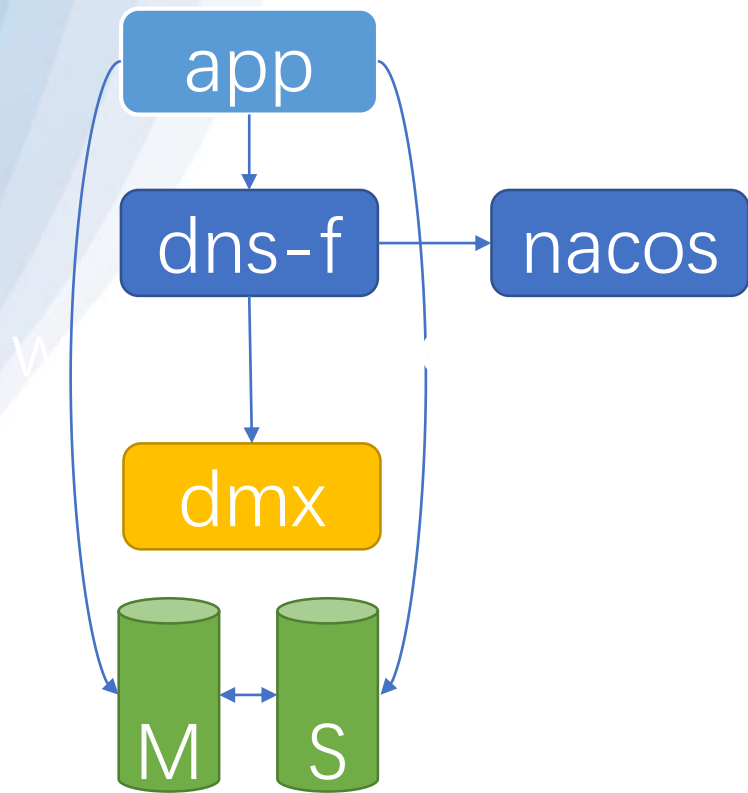
2019.8.11



微服务改造事项



DNS-F在数据库场景的落地



背景

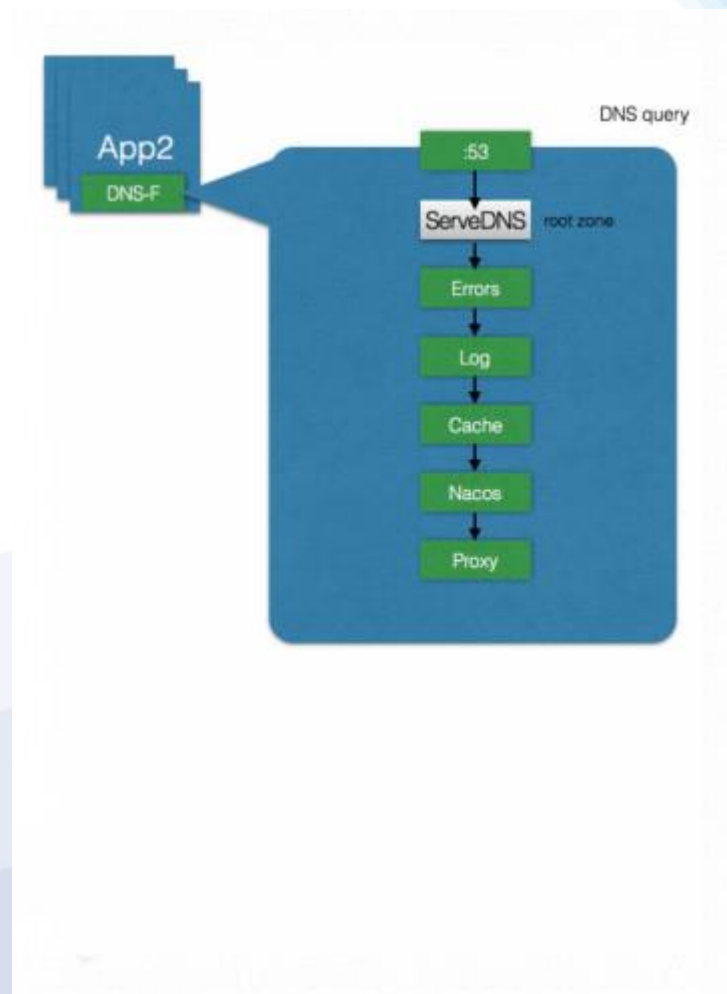
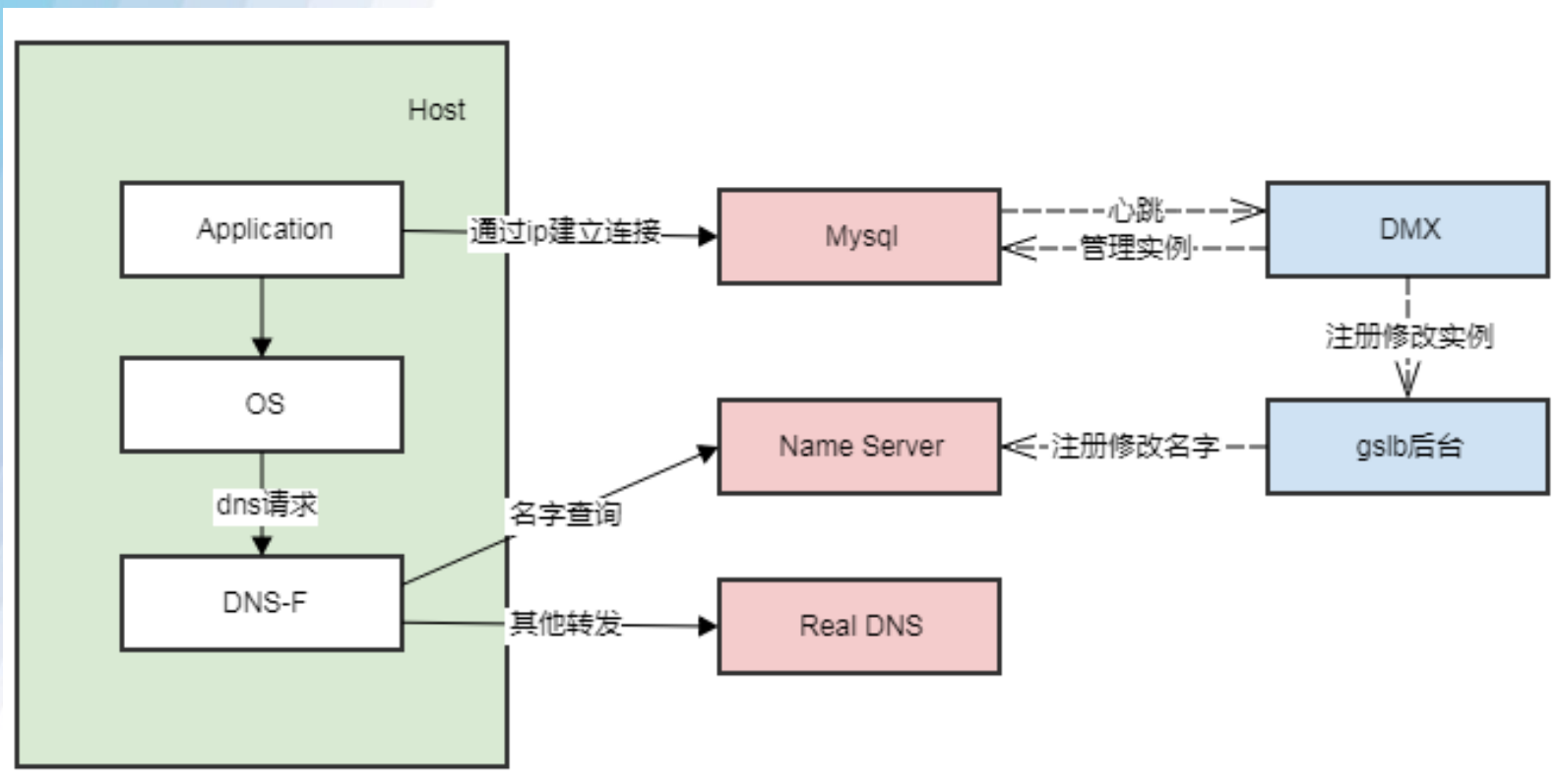
在数据库服务切换时效率低，依赖业务方修改配置，时效不确定，通常需要10分钟以上；



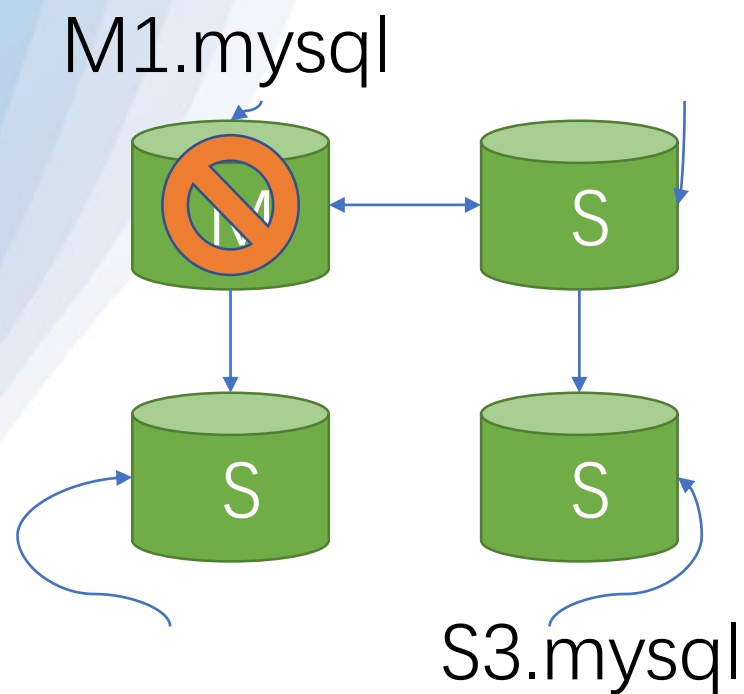
目标

数据库切换秒级生效；

技术特点-DNS-F



切换逻辑



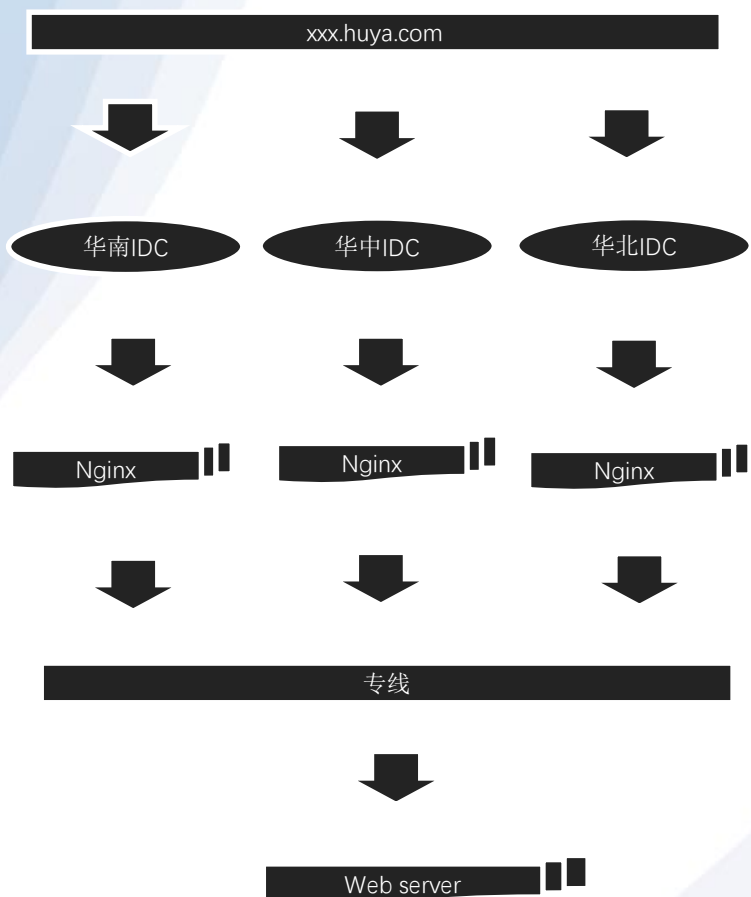
if M1 down and M1 is master:

choose newest slave: S1 or S2, compare slave IO/SQL position
move M1.mysql to newest one(assume S1.mysql)
if group is MySQL 5.7: find all slave on M1:S2 change master to S1 done

if S1 down and S1 is slave:

choose light slave: S3 or M1, compare the instance upload(qps / connection and so on.), i.e: S3
if group is MySQL 5.7: find all slave on S1:S3 change master to M1 done

名字服务在负载均衡场景的落地



背景

负载均衡变更需要更新配置发布后才能生效，节点扩缩容，屏蔽操作需要人工参与，流量接入接出时间长



目标

流量接入接出时长低于30秒

负载均衡配置下发无状态改造

虎牙负载均衡部署的机器数百，配置下发的可靠性保障，一致性保障要求特别严格，如果出现配置下发不及时，或下发配置失败，极大可能出现故障，同时负载均衡服务弹性能力的要求较高，在业务高峰如果不能快速扩容，容易出现全网故障

传统配置下发方式



服务端下发文件更新配置，更新配置生效时间长，由于需要预先知道负责均衡集群的机器信息，扩缩容需要等元信息同步以后才能接入流量，扩容流量接入时间长

配置IP方式

负载均衡配置固话在配置文件中，节点操作需要人工上系统操作增减节点

配置中心监听方式



客户端主动监听配置更新，配置秒级生效，新扩容服务主动拉取全量配置，流量接入时长缩短3分钟+

配置名字服务

负载均衡节点对接名字服务

虎牙负载均衡部署的机器一台机器上可能存在数百的配置，每次节点生效需要重新 reload 配置，且节点新增或删除需人工上系统操作配置，配置生效对服务会有中断

配置IP方式



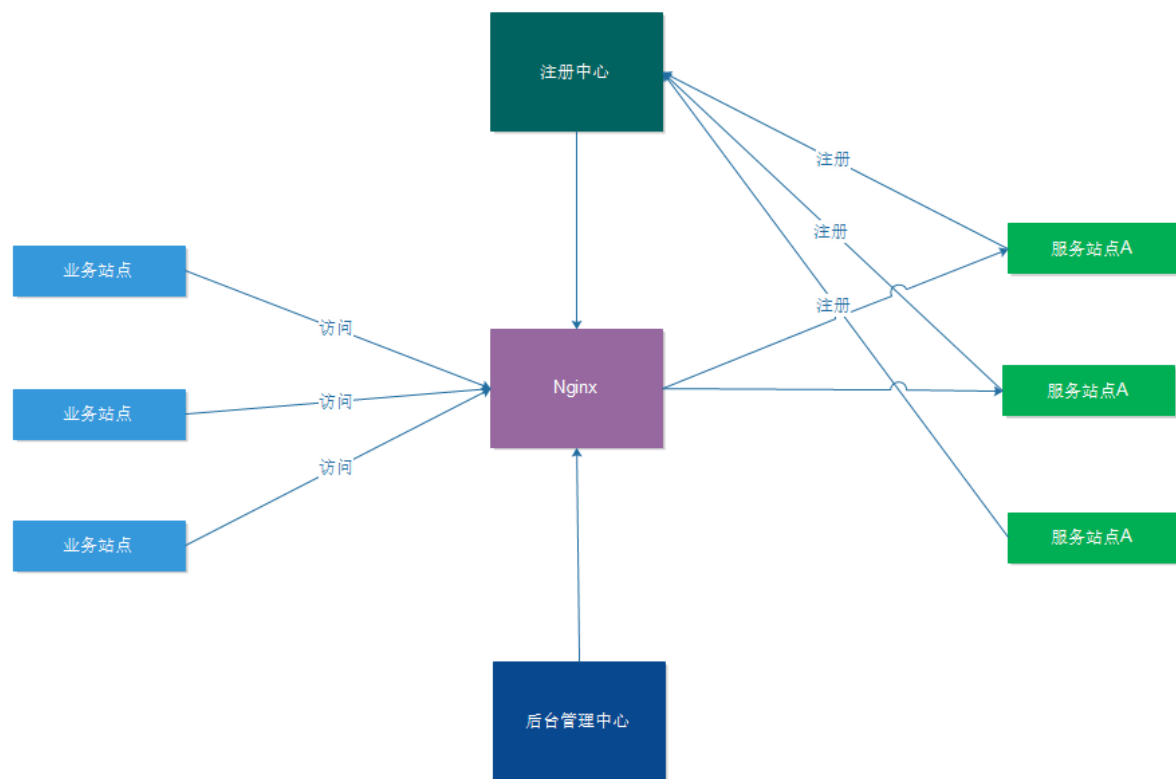
负载均衡配置固话在配置文件中，节点操作需要人工上系统操作增减节点

配置名字服务



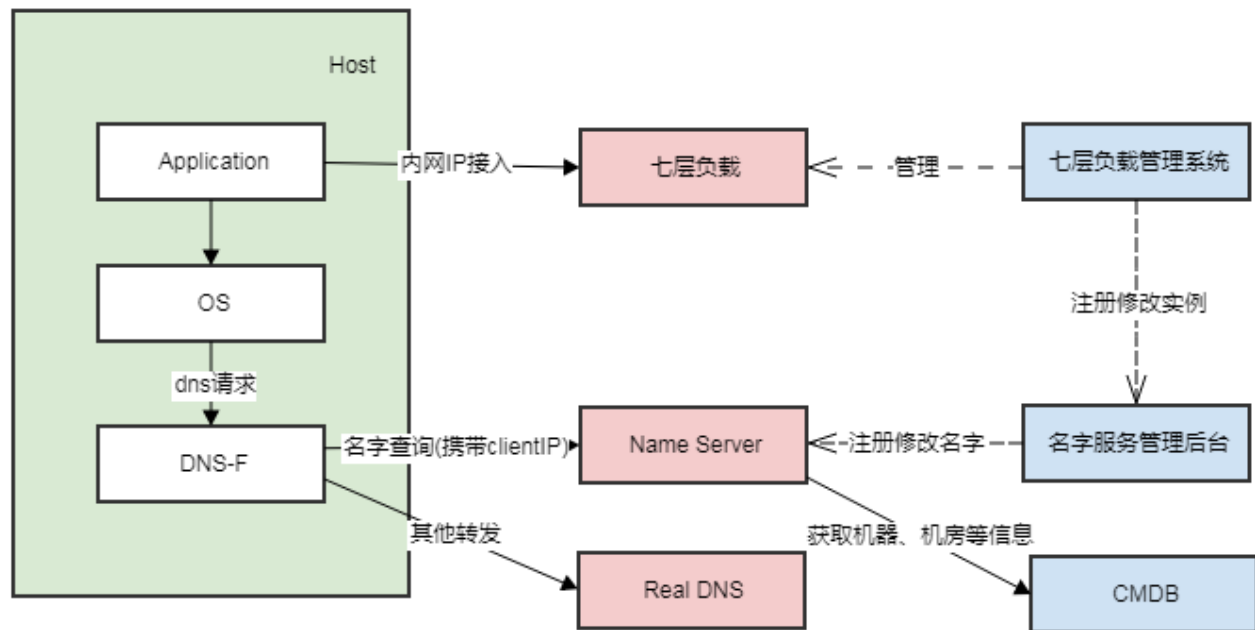
客户端主动监听名字服务更新，节点变更秒级生效

基于Nginx dyups模块的动态上下线



- 基于名字服务感知节点变化快速变更节点
- 基于nginx dyups模块实现节点变更秒级生效

微服务架构中基于DNS的服务注册与发现



背景

公司内部注册中心比较多，常见的有zookeeper、etcd、consul、eureka等。服务注册通常有三种：自注册、第三方注册、注册中心主动同步，异构微服务相互调用困难；公共DNS不稳定，解析结果生效慢，解析失败率高；DNS无法提供内网环境的就近接入能力，内部服务调用无法实现就近访问。



目标

- DNS解析成功率 100%
- 公共DNS分钟级生效
- 内部DNS秒级生效
- 基于IP+CMDB实现就近访问能力

dns当前现状

1、依赖公共localDNS，解析不稳定，延迟大



落地效果



平均解析时间

之前200ms+
现在1ms以下



缓存命中率

之前92%
现在99.8%+



解析失败率

之前0.1%
现在0%

应用探活场景下的落地

背景

- 1、web专区的nginx有健康检查的功能，但是只是基于tcp，根据一段时间请求是否正常，对节点进行临时屏蔽
- 2、web专区和业务都会有自己的拨测，一般只会用于告警，缺乏都能根据业务反馈，动态切换节点的能力
- 3、支付等业务，对故障时间敏感，手工切换无法满足需求

目标

- 1、提供http，tcp的业务拨测能力，能够按照预定条件进行对负载均衡RS节点的摘除，屏蔽故障
- 2、保证拨测任务的SLA，能够按时执行，故障屏蔽秒级生效时间



Home > URL测试 > URL拨测配置

初始化URL拨测配置

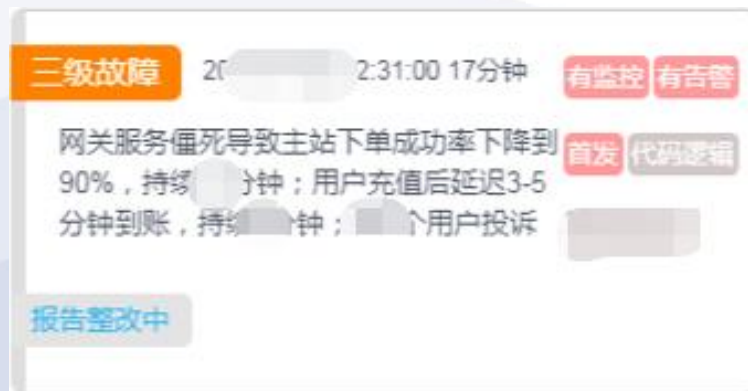
*URI: /index.html

*端口: 8089

*类型: ☒ LVS VIP检查 ☒ Nginx_Proxy检查 ☒ 业务服务器检查

*告警条件: ☐ 返回码不为: 200 多个用分隔
☐ 返回内容不为:
☐ 返回内容大小不为: 字节
☐ 探测时间超过: 秒

*请求方法: GET



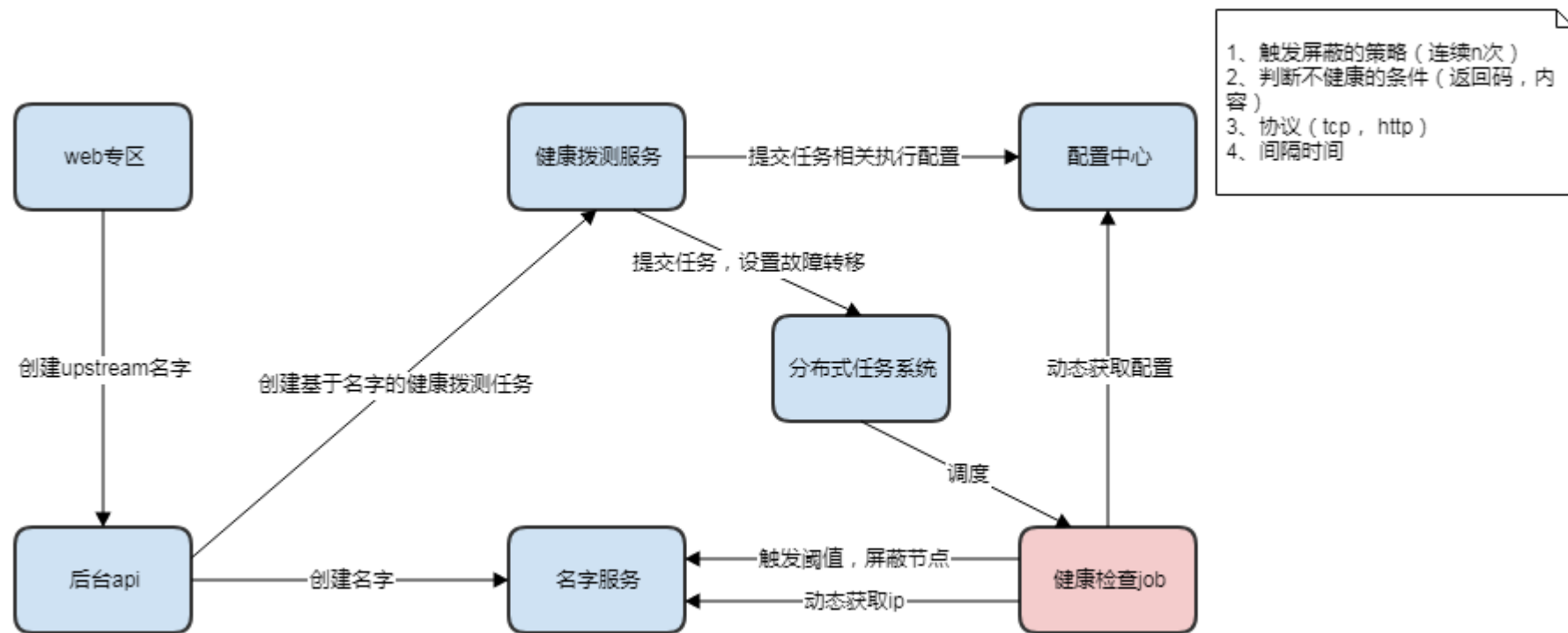
三级故障 20 2:31:00 17分钟 有监控 有告警

网关服务僵死导致主站下单成功率下降到90%，持续 分钟；用户充值后延迟3-5分钟到账，持续 分钟； 个用户投诉

首发 代码逻辑

报告整改中

方案介绍



落地效果



故障节点发现时间
之前分钟级
10s内



流量摘除生效时间
之前10分钟
30s内



故障屏蔽实效
之前10分钟
1分钟内

为什么选择Nacos



Service Mesh

NACOS.

服务注册发现对比

对比项目\注册中心

CAP模型

客户端更新服务信息

伸缩性

健康检查模式/方式

负载均衡

手动上下线服务方式

跨中心同步

k8s集成

分组

权重

厂商

Spring Cloud Nacos

支持AP和CP模型

使用注册+DNS-f+健康检查模式。DNS-F客户端使用监听模式push/pull拉取更新信息

使用Raft选举算法性能、可用性、容错性均比较好，新加入节点无需与所有节点互相广播同步信息

支持服务端/客户端/关闭检查模式，检查方式有tcp、http、sql。支持自己构建健康检查器

支持

通过控制台页面和API

支持

支持

Nacos可用根据业务和环境进行分组管理

Nacos默认提供权重设置功能，调整承载流量压力

阿里巴巴

Spring Cloud Eureka

AP模型

客户端定时轮询服务端获取其他服务ip信息并对比，相比之下服务端压力较大、延迟较大

由于使用广播同步信息，集群超过1000台机器后对eureka集群压力很大

客户端向服务端发送http心跳

支持

通过调用API

不支持

不支持

不支持

不支持

Netflix

NACOS.



蚂蚁金服
ANT FINANCIAL



ServiceMesh

配置中心对比

对比项目/配置中心	apollo	nacos
开源时间	2016.5	2018.6
配置实时推送	支持（HTTP长轮询1s内）	支持（HTTP长轮询1s内）
版本管理	自动管理	自动管理
配置回滚	支持	支持
权限管理	支持	待支持
多集群多环境	支持	支持
监听查询	支持	支持
多语言	Go,C++,Python,Java,.net,OpenAPI	Python,Java,Nodejs,OpenAPI
分布式高可用最小集群数量	Config2+Admin3+Portal*2+Mysql=8	Nacos*3+MySql=4
配置格式校验	支持	支持
通信协议	HTTP	HTTP
数据一致性	数据库模拟消息队列，Apollo定时读消息	HTTP异步通知
单机读（tps）	9000	15000
单机写（tps）	1100	1800



蚂蚁金服
ANT FINANCIAL



ServiceMesh

NACOS.

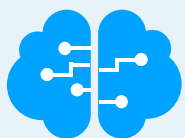
技术选型

- Dubbo生态产品， Nacos源于阿里内部服务的开源实现；
- 支持各种生态环境， 支持k8s， spring cloud， dubbo， service mesh等集成；
- 配置中心与注册中心原生打通；

NACOS.

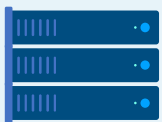


落地的改造和升级



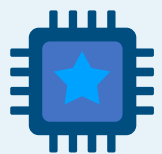
DNS-F

- DNS-F 增加对外部域名的预缓存支持;
- Agent监控数据对接公司内部监控;
- 日志输出对接公司内部日志服务;
- 对接公司CMDB;
- DNS-F Cluster;
- DNS 忽略大小写



Nacos Sync

- 接入TARS注册服务
- 对接K8S注册服务
- 解决多数据中心环形同步



Nacos CMDB

- 扩展对接虎牙CMDB
- 对接内部负载均衡策略

虎牙直播与阿里共建Nacos生态

NACOS.

HOME

DOCS

TEAM

BLOG

COMMUNITY

CAMPUS

DEMO-CONSOLE

中

虎牙直播共建Nacos生态



虎牙直播平台(<https://www.huya.com/>)

虎牙直播是中国一家以游戏直播为主的视频直播网站。网站以YY直播为名创建于2011年，隶属于欢聚时代，2014年11月24日更名为虎牙直播，开始独立运营。其海外版本为以东南亚为市场的Nimo TV。虎牙直播是国内最为资深的以游戏内容为核心的直播平台，其在游戏方面有丰富的独家资源，汇聚目前最为火爆的游戏，如英雄联盟、王者荣耀、球球大作战、守望先锋、炉石传说、绝地求生、绝地求生手游——刺激战场、全民突击等主题的直播内容。

虎牙中间件团队介绍

虎牙中间件团队主要负责虎牙服务注册发现与配置中心，负责均衡，中间件等平台相关的建设，致力于保障改进虎牙用户接入的请求成功率及请求耗时，为虎牙提供稳定可靠全球基础接入服务设施，服务间通信的基础设施。



社区贡献

- Nacos Committer 4人;
- 承担多个模块的开发和维护 (Nacos Service Mesh, Dubbo Mesh, Golang SDK, Nacos Syn等);
- 数十issue, bug修复;

Nacos 支持Service Mesh



ServiceMesh

NACOS.

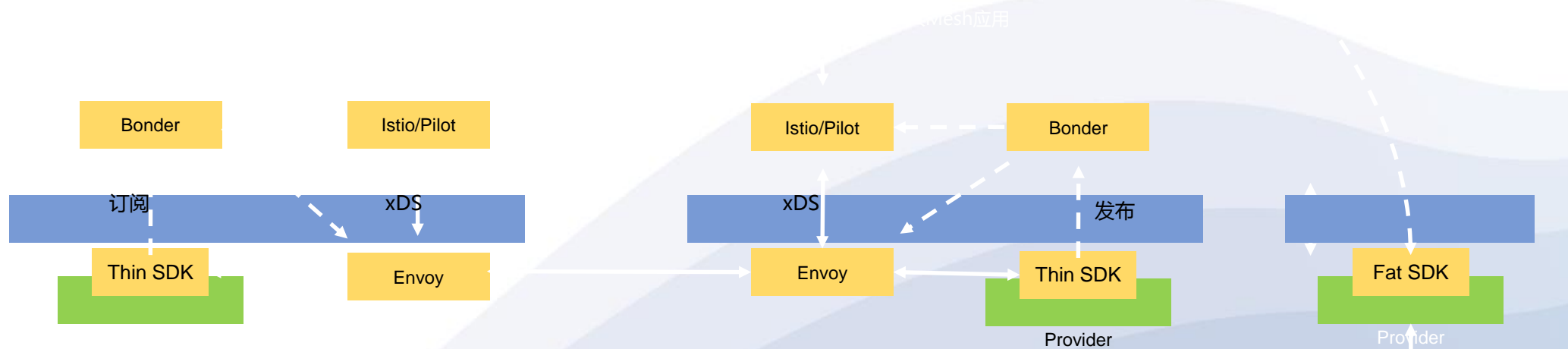
Service Mesh结合Nacos在阿里巴巴落地

- 基于开源Istio，对接Nacos实现服务注册与发现

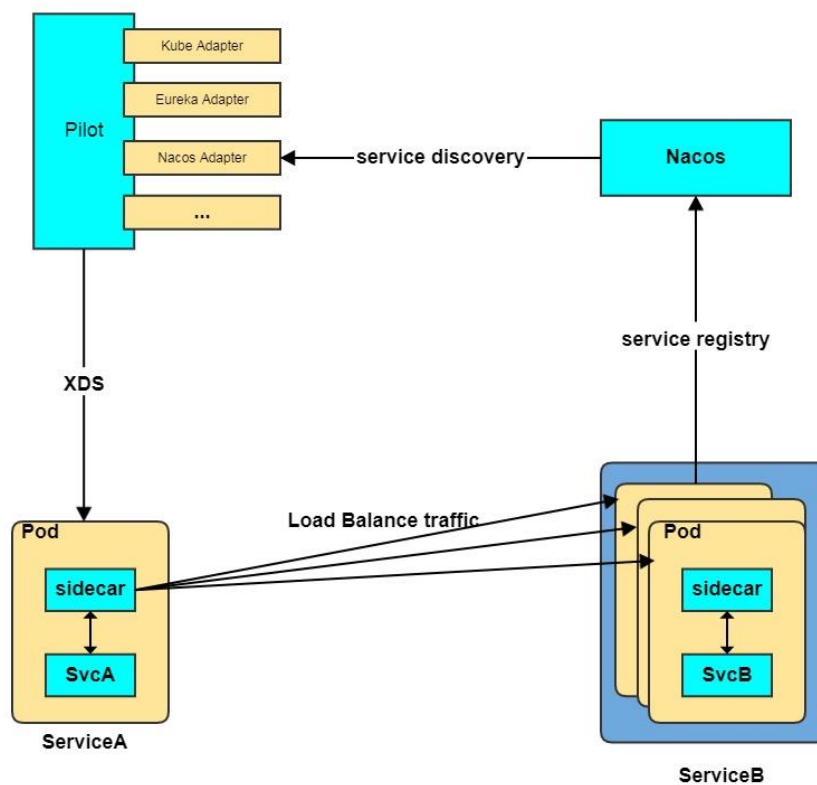
Nacos

在闲鱼落地：

- 解决多技术栈复杂性
- 解决多语言Dubbo客户端功能差异问题



Nacos支持Service Mesh方案



Nacos支持Service Mesh规划

支持Istio服务发现

Nacos

服务发现

Istio/Pilot

xDS

支持Istio服务发现和配置管理

Nacos

服务发现

配置管理

Istio/Pilot

xDS

支持xDS协议

Nacos

Istio/Pilot

xDS

支持SMI

Nacos

SMI

xDS



关注 **ServiceMesher** 微信公众号
获取社区最新信息



关注 金融级分布式架构 微信公众号
获取 **SOFAShark** 最新信息

ServiceMesher 社区是由一群拥有相同价值观和理念的志愿者们共同发起，
于 2018 年 4 月正式成立，致力于成为 Service Mesh 技术在中国的布道者和领航者。

社区官网：<https://www.servicemesh.com>

