



背景介绍

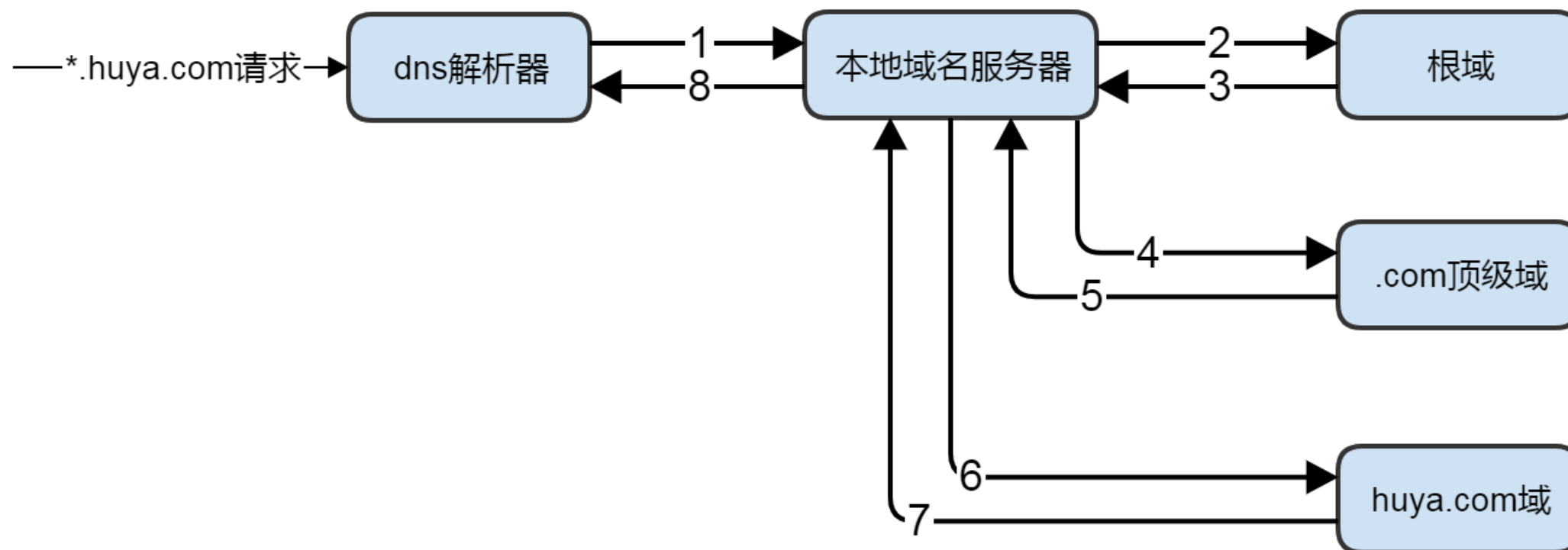
方案设计与对比

高可用

具体实践与落地

规划

A decorative banner at the bottom of the page featuring a dense collection of various icons. The icons include symbols for technology (laptop, smartphone, tablet, camera), business (briefcase, magnifying glass, bar chart, pie chart, dollar sign, yen sign, euro sign), leisure (video game controller, steering wheel, rocket, heart, star, crown, gift, puzzle pieces), and general life (smiley face, speech bubble, key, rocket, star, crown, gift, puzzle pieces). The icons are arranged in a horizontal line, creating a colorful and busy border.



DNS特点

天然分布式

- 分级缓存
- 多级容灾， 流量分散

协议支持面广

- 手机，pc等用户设备
- Sdk
- 服务器



1、依赖公共localDNS，解析不稳定，延迟大



DNS当前现状

2、记录变更生效时间长， 无法及时屏蔽线路和节点异常对业务的影响

- 权威DNS全球各节点数据同步时间不可控，全局生效时间**超过10分钟**
- localDNS缓存过期时间不可控，部分localDNS不遵循TTL时间，缓存时间**超过48小时**

3、内部DNS功能缺失，无法解决**内部服务**调用面临挑战

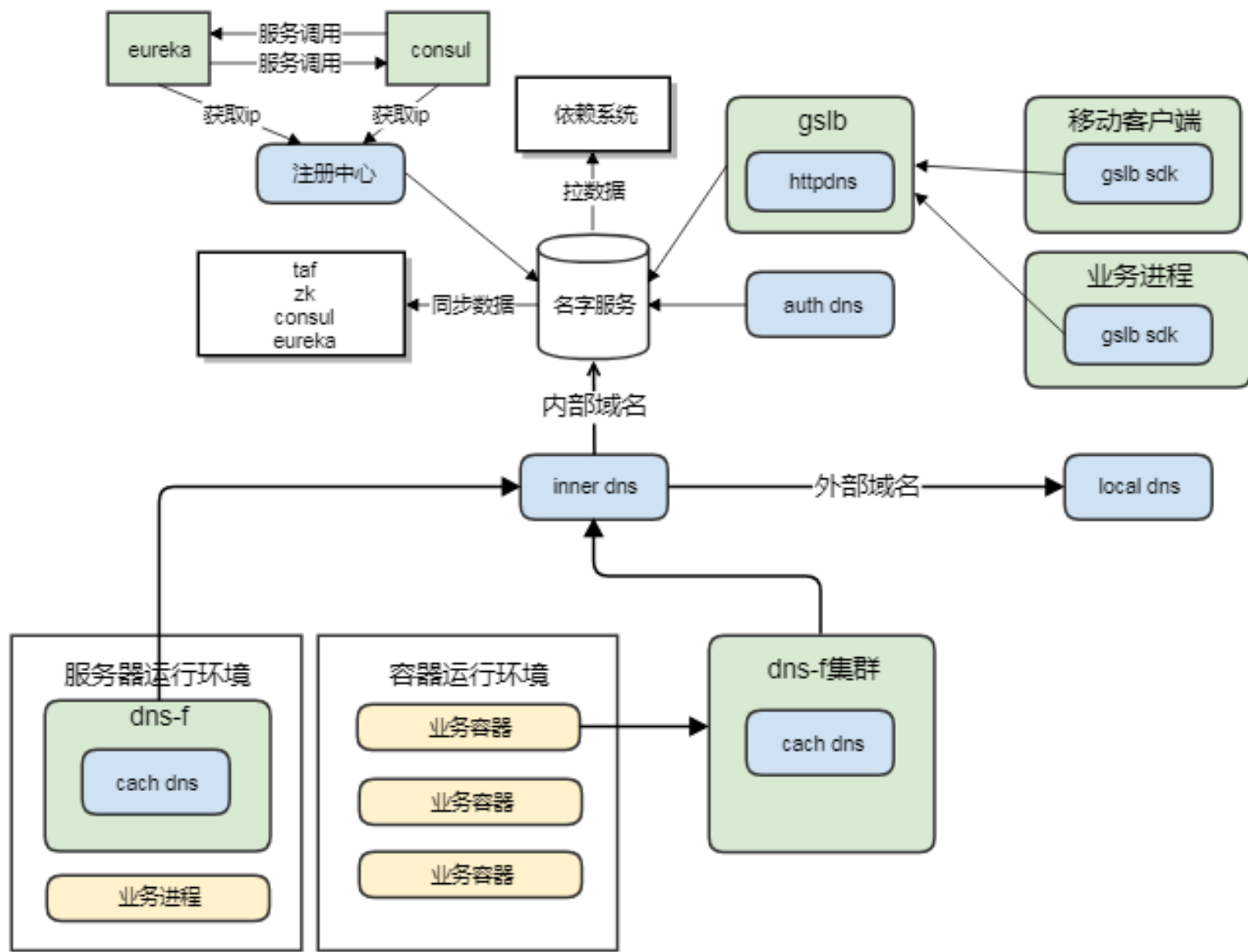
- 时延大，解析不准
- 支持多种调度策略

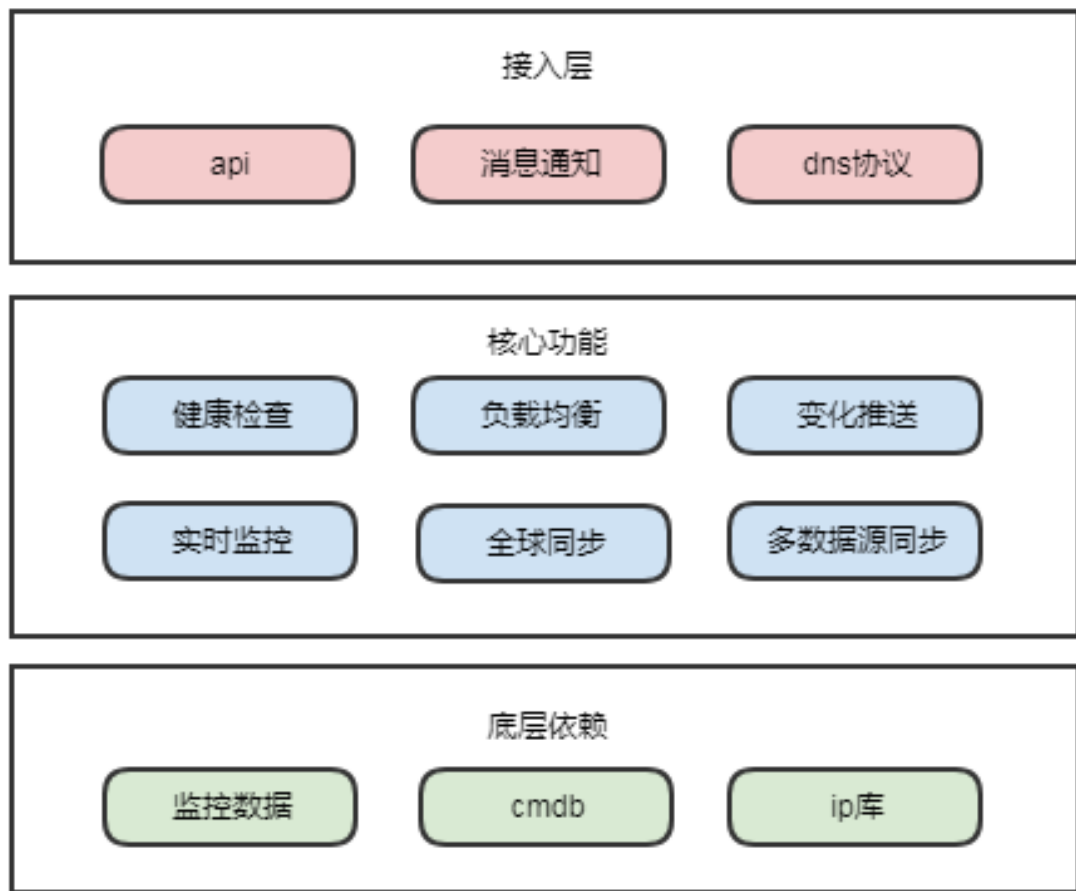
4、无法满足国外业务的快速发展

- 基于DNS的快速扩容方案
- 基于DNS的数据库切换方案

方案设计与对比







Nacos

- 名字服务核心， 提供统一的api， 名字注册， 变化推送， 负载均衡

Nacos-sync

- 集群间数据同步组件

DNS-F

- 客户端组件， 拦截DNS请求， 实现基于DNS的名字服务



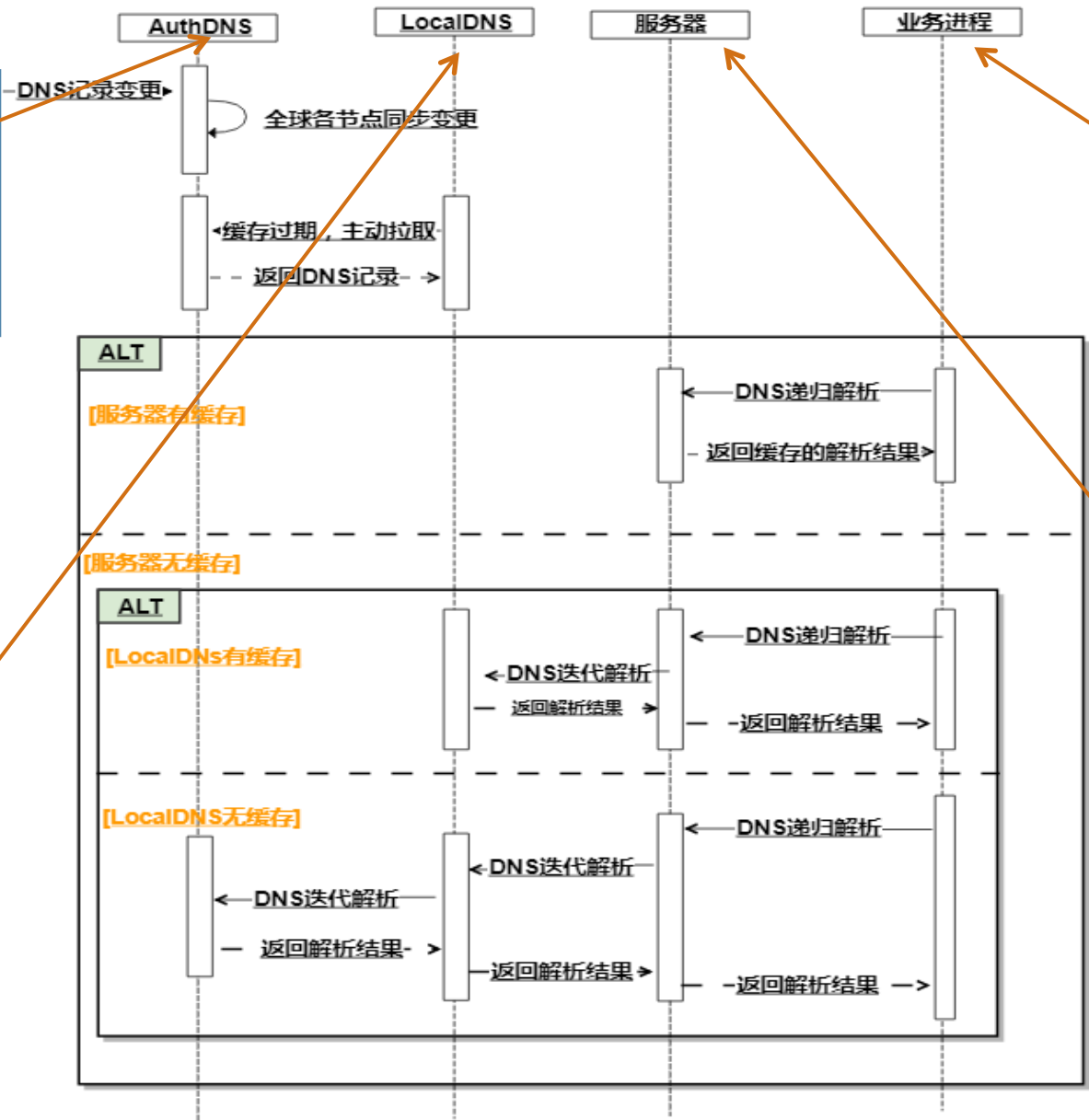
原有dns变更生效流程

影响DNS生效时间因素：
1、跨区域、跨国数据同步慢，不稳定；
2、bind 在数据量比较大的时候，同步比较慢。

影响DNS生效时间因素：
1、根据TTL缓存，过期后才会刷新数据；
2、部分厂商不遵循TTL时间缓存，超过24小时的缓存时间。

影响DNS生效时间因素：
1、应用的DNS缓存，比如java虚拟机、框架层的DNS缓存。

影响DNS生效时间因素：
1、服务器开启nsd做DNS缓存。



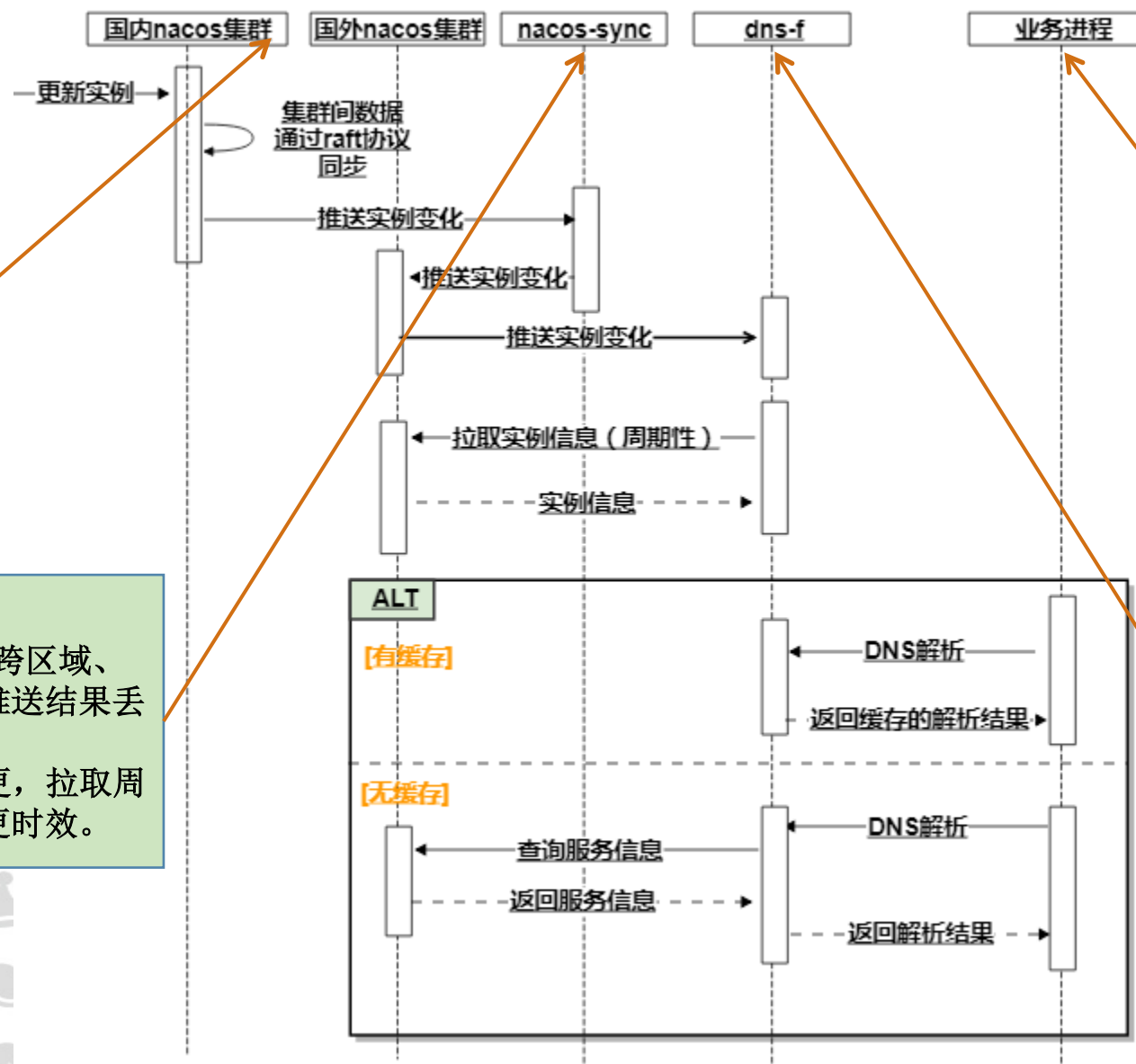
现有dns变更生效流程

影响DNS生效时间因素：
1、集群内通过raft协议同步数据，毫秒级别完成同步。

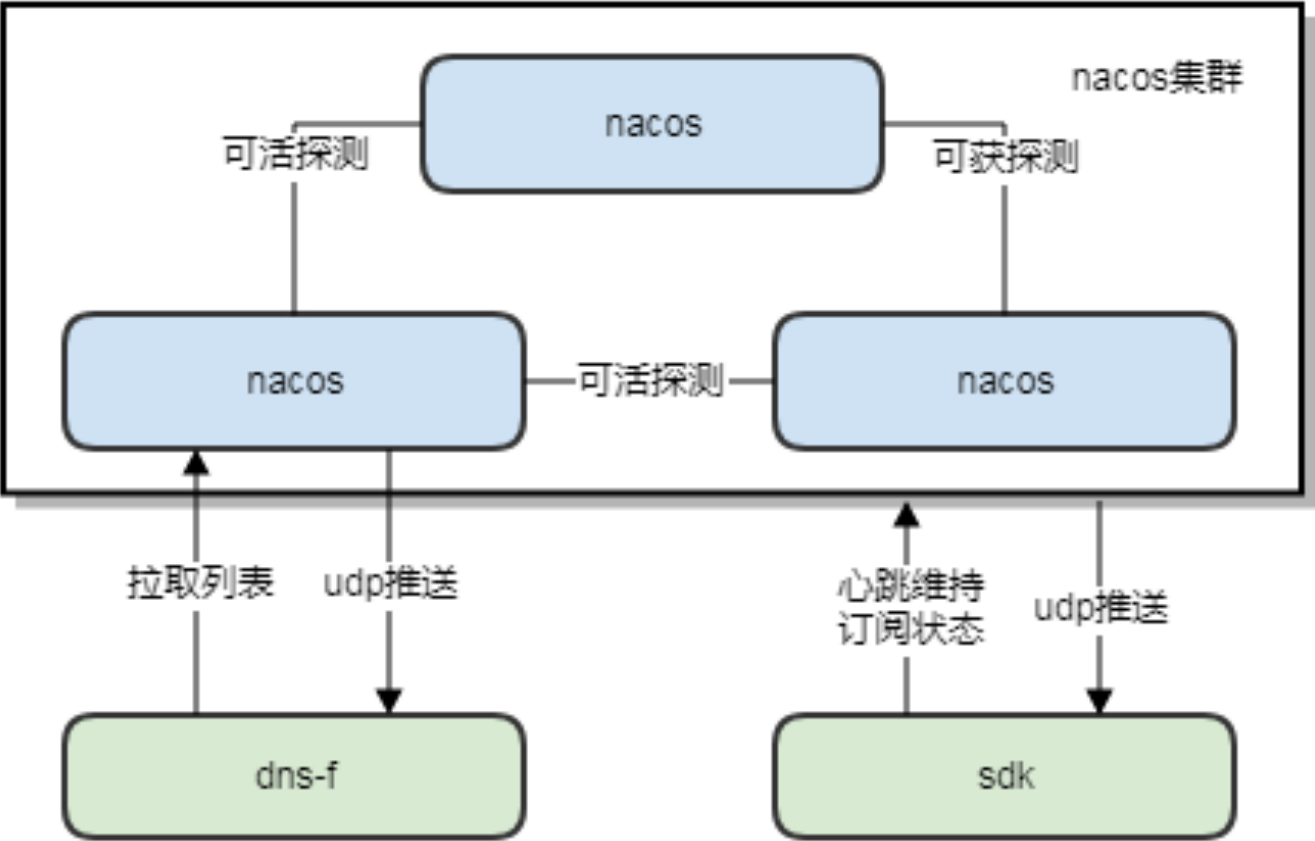
影响DNS生效时间因素：
1、应用禁用DNS缓存

影响DNS生效时间因素：
1、nacos推送变化到nacos-sync，跨区域、跨国网络差的情况下可能会导致推送结果丢失，或者延迟加大。
2、nacos-sync会主动拉取实例变更，拉取周期和监听的服务数量会影响到变更时效。

影响DNS生效时间因素：
1、nacos会将变更推送到dns-f，网络差的情况可能会导致推送结果丢失，或者延迟加大。
2、dns-f会主动拉取实例变更，拉取周期和监听的服务数量会影响到变更时效。



消息推送机制



多集群数据同步

	Nacos-sync	数据库	文件
同步粒度	单个实例	单个操作	单个文件
可用性	集群	单机	单机
同步方式	全量写入	操作重放	增量同步
环形同步	支持	主主同步， 有限制	不支持



Nacos-sync与开源版本的差异

- 通过配置分片任务
- 通过事件合并以及队列控制nacos集群写入量
- 支持从k8s和taf同步数据
- 支持通过endpoint进行跨区域之间数据同步



核心组件设计

DNS-F是基于CoreDNS上开发的，扩展了以下几个组件：

nacos插件： 查询nacos服务信息，监听nacos服务变化，并将服务转化为域名，实现以DNS协议为基础的服务发现；

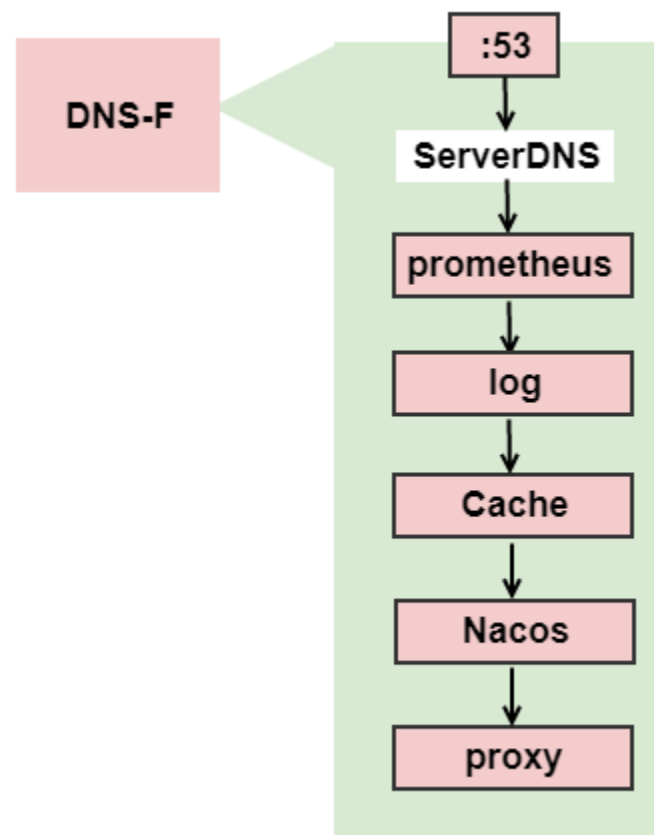
cache插件： 提供域名缓存服务；

Log插件： 将DNS解析日志上报到日志服务；

proxy插件： 代理解析外部域名。

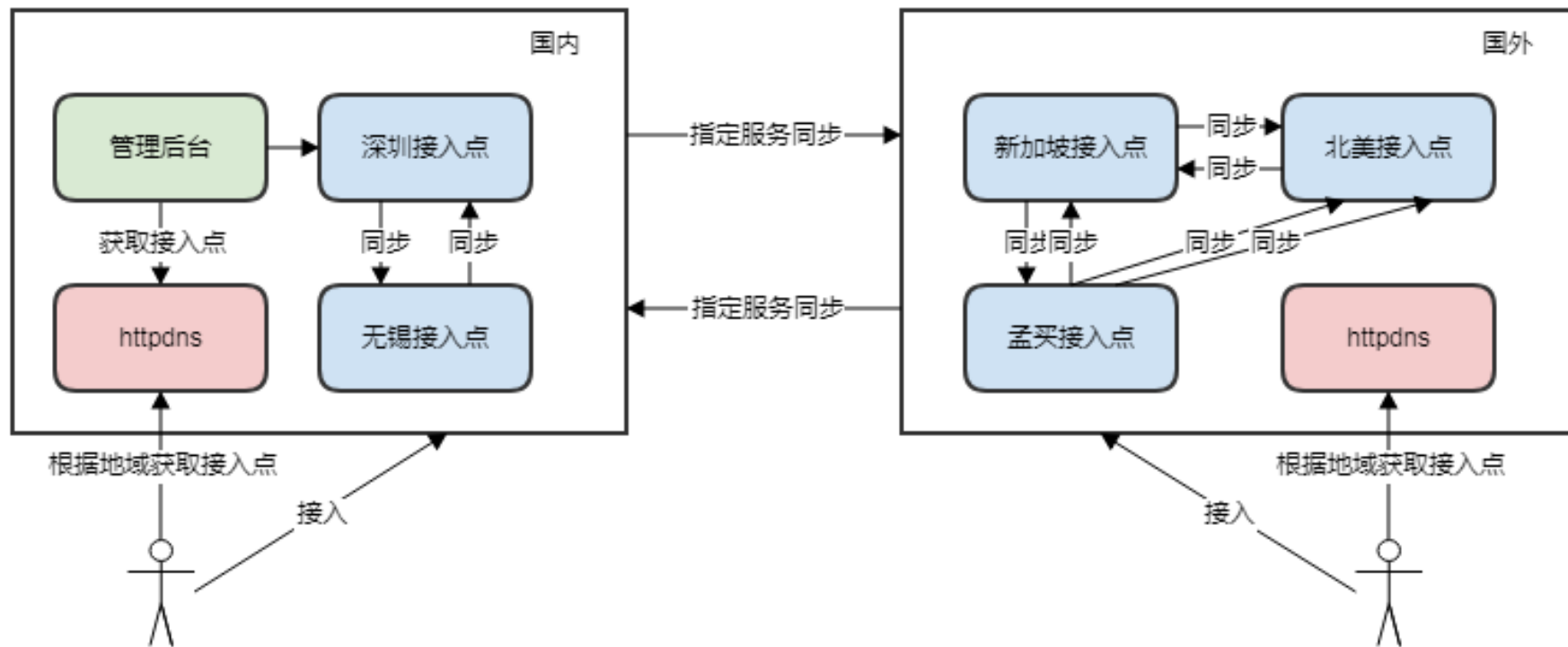
与开源版本差异：

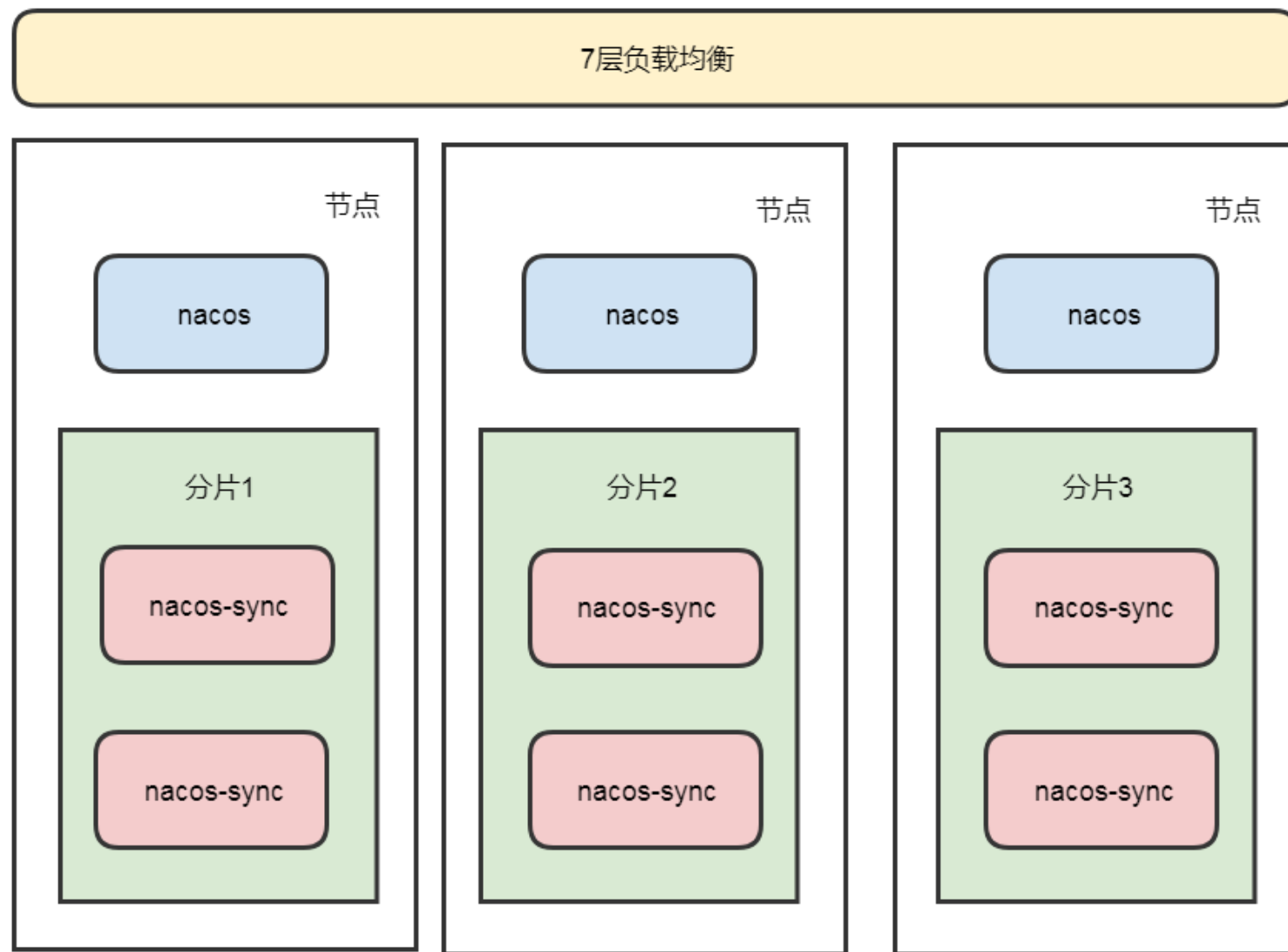
- 缓存功能增强，永不过期，异步刷新缓存
- 高可用保障能力增强，监控和健康检查脚本完善
- 支持集群化部署，解决集群部署的服务变更推送、服务负载均衡等问题



高可用





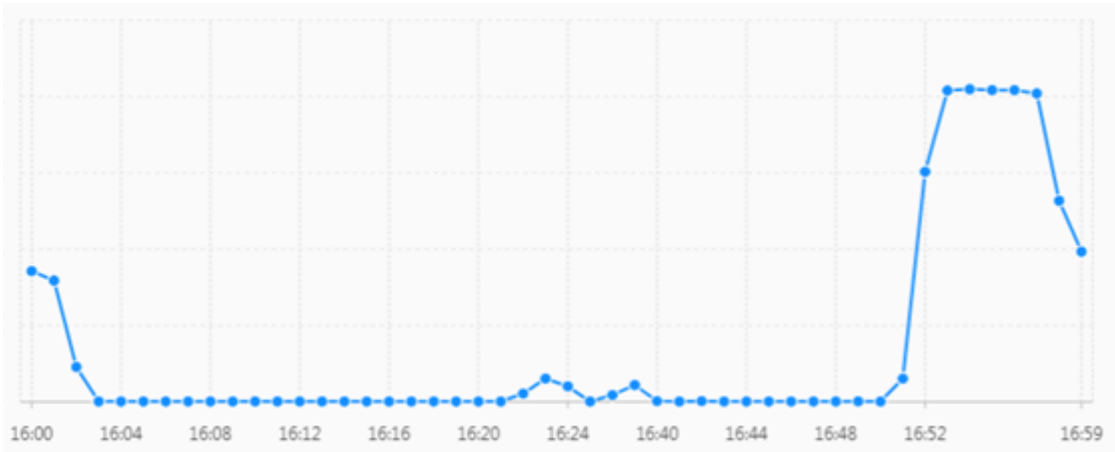


线上演练

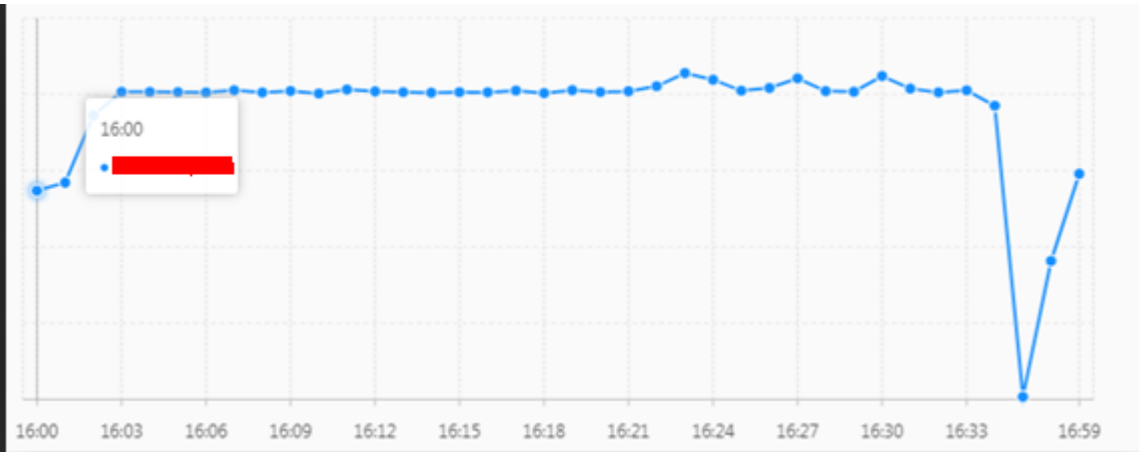
操作记录

时间点	操作
16:01:00	将深圳流量全部切走
16:01:48	日志确认流量已经全部迁移完毕
16:26:12	关闭深圳集群的服务
16:34:52	关闭无锡集群的服务（从这个时刻起，国内两个集群都已经停止服务）
16:48:37	启动深圳集群的服务
16:51:44	切换流量到深圳集群
16:53:22	启动无锡集群的服务
16:59:00	恢复两个集群的流量
17:01:00	数据验证和日志查看，确认正常

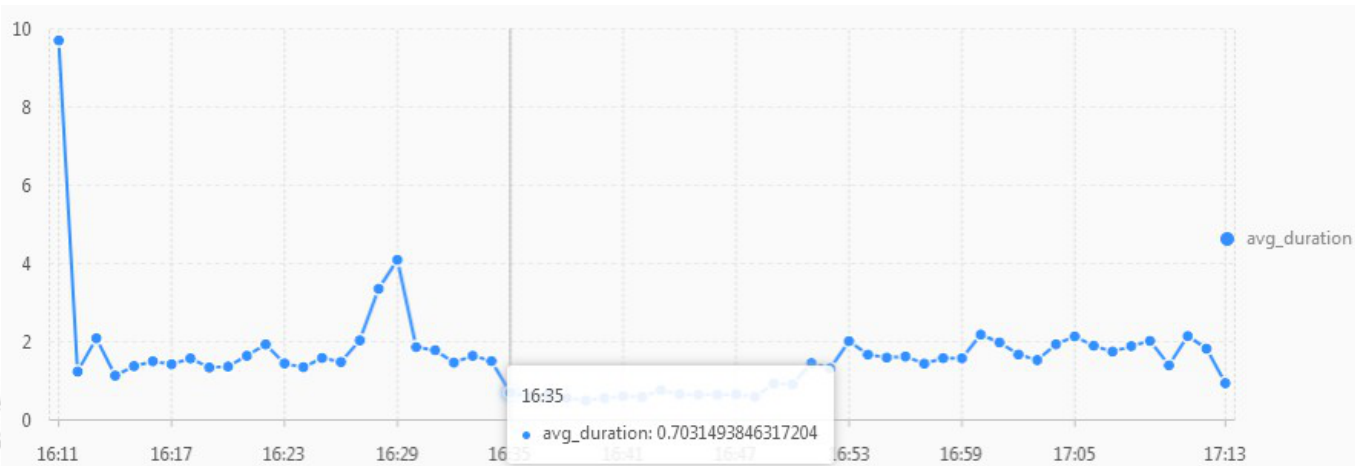
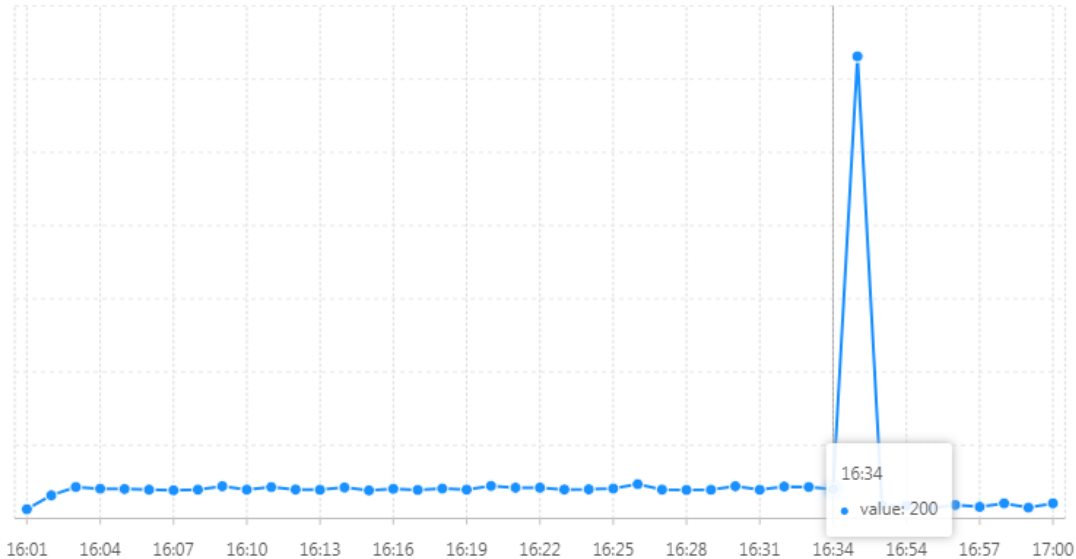
深圳流量



无锡流量



服务端写入影响（写入延迟）



DNS-F解析影响（解析延迟）

双集群可活方案

一致性保证

- 切换后保证元数据和实例数据最终一致
- 集群重新上线， 需要完成数据校验

可用性级别

- 单集群挂掉不会有影响
- 双集群挂掉只影响域名变更， 不影响域名解析



- 大区内部10分钟全量校验
- 大区之间2分钟一次全量校验

- 从其他数据源同步的数据， 通过数据源的时间戳， 增量校验
- 基于api写入日志， 定期校验写入的内容是否已经全部同步

DNS-F高可用

- 1、agent健康状态检测，包括进程存活和是否能正常解析；
- 2、缓存内部域名，nacos集群出问题不影响正常解析；
- 3、提供备用节点。
- 4、resolv.conf配置检测，发现127.0.0.1不在配置中会自动添加；
- 5、限制agent的cpu使用最高为1个cpu；

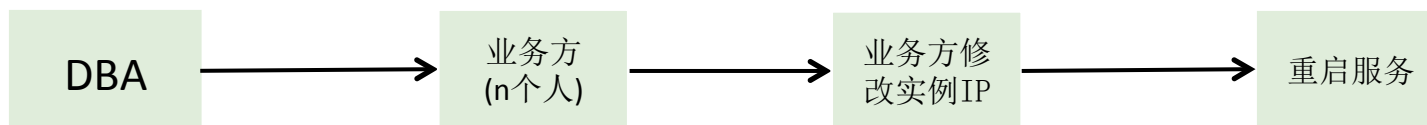


具体实践与落地



背景:

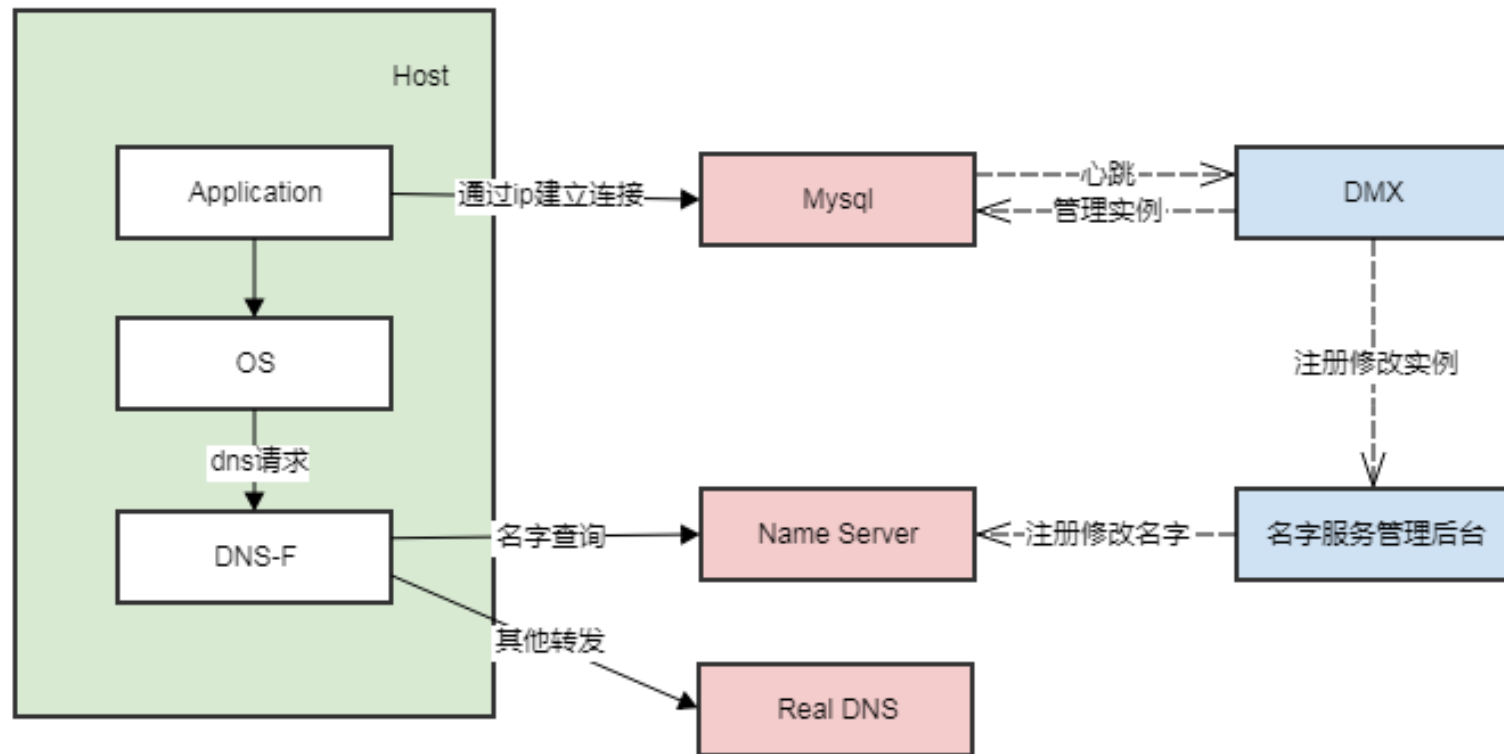
数据库使用ip方式接入，在数据库服务切换时效率低，依赖业务方修改配置，时效不确定，通常需要10分钟以上。



数据库IP切换流程

提升数据库切换效率关键点：

- 通过域名接入，切换时业务方无需参与，无感知实例切换；
- 域名变更能秒级推送到DNS-F，应用快速切换到新实例。



DMX:数据库管理系统，负责管理数据库实例。

最佳实践--数据库域名化改造

落地效果:

- 1、数据库域名变更秒级生效;
- 2、数据库切换10s内完成。



背景:

内部系统之间的调用通过七层负载接入，使用的是外部域名，需要通过公共localDNS解析，存在以下问题:

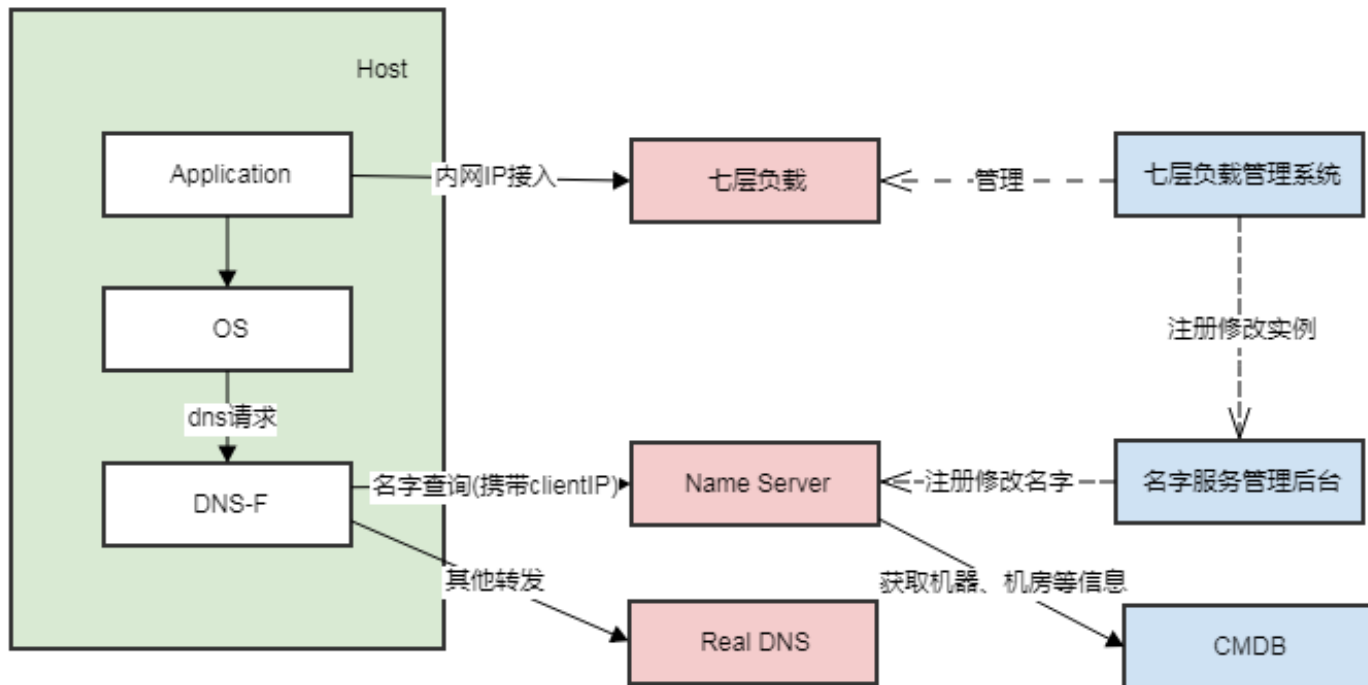
- 1、扩缩容时需要修改DNS记录，生效时间超过10分钟，故障节点影响业务较长时间;
- 2、公共DNS智能解析不准确，影响接入质量;
- 3、不支持定制化的负载均衡策略，比如同机房、同大区优先等策略。



最佳实践--内部调用使用内部域名

提升内部服务调用质量关键点：

- 绕过localDNS，DNS变更记录及时推送到DNS-F，流量快速牵引；
- 与内部系统打通，能根据cmdb信息，支持多种负载均衡策略。



CMDB：配置管理数据库，nacos提供扩展点从CMDB获取主机、机房等信息，并以label方式打标到实例，作为负载均衡的元数据。

落地效果:

- 服务扩缩容在秒级完成, 减少了故障时长;
- 扩展了DNS负载均衡策略, 满足了不同业务场景, 比如有些业务在不同区域有不同接入点, 不能跨区域调用, 原有DNS服务不能满足同区域调度策略, 接入内部域名后, 可根据CMDB信息做同区域调度;
- 业务接入内部域名后延迟下降明显。

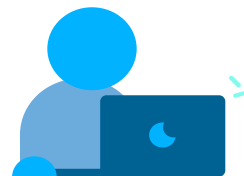


优化后风控服务, 平均延迟下降10ms, 服务超时比例下降25%



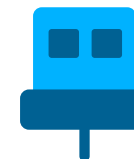
平均解析时间

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%



- 提供基于DNS协议的服务发现能力，消除了异构系统之间服务调用的障碍
- 填补了没有内部域名解析能力的空白
- 解决了内部服务调用面临挑战：时延大，解析不准，不支持多种调度策略，故障牵引慢
- 加速外部域名解析，屏蔽localDNS故障对业务的影响



- 

规划



DNS后续规划

LocalDNS

- 解决公共DNS节点位置影响域名解析准确性问题
- 解决内部使用公共DNS不稳定问题
- 优化内外网解析

AuthDNS

- 精准调度
- 解决全球ns节点同步慢问题



Thanks

