



Saber

分布式服务探索与实践

高性能、透明化的RPC远程服务调用、SOA服务治理方案



郑志升@Bilibili

个人介绍

郑志升@scofield

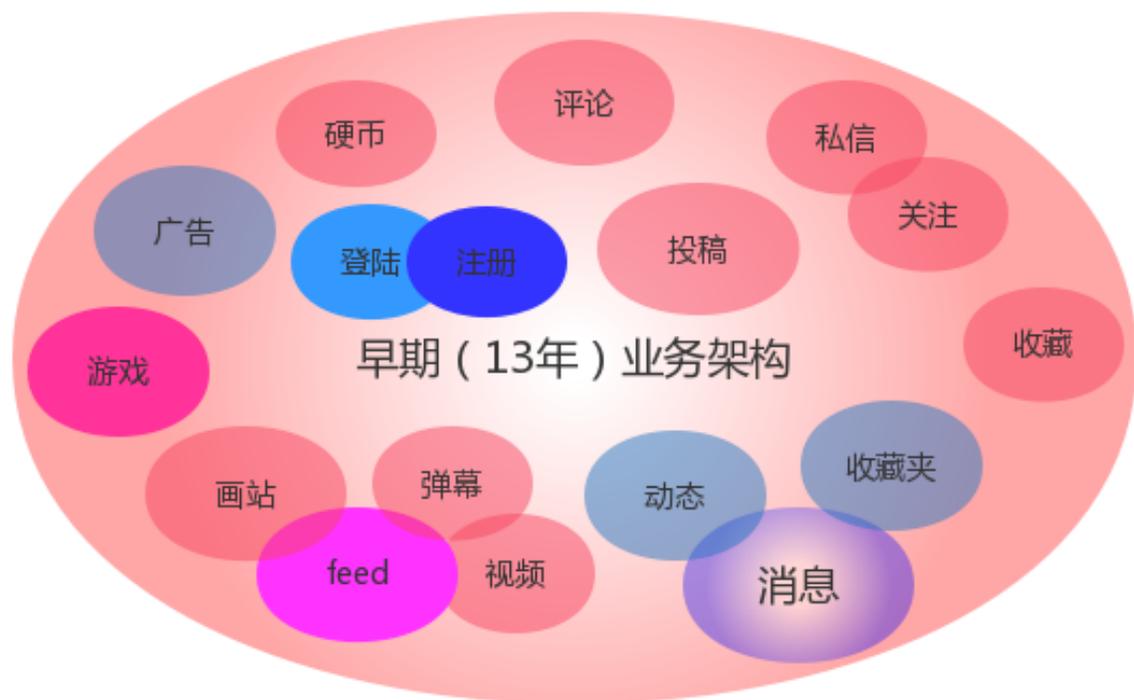
B站研发经理，平台型业务及组件研发

曾就职于UC（现阿里巴巴），游戏事业部

Tags：web架构、分布式服务、羽毛球、骑行、做饭

早期的B站

业务架构



系统架构

代理层 :

Nginx

Web层 :

PHP

基建层 :

Mysql

Memcache

鬼知道经历了什么

* - 2013年

- 单一应用架构 - All in one
- 主站PHP一体化，所有功能部署一起
- 没人能改动代码或追加需求，效率低

2014年

- 垂直应用架构
- 将应用拆分为多个互不相干的几个应用
- 账号拆分、消息拆分、游戏拆分、空间拆分...
- 应用交互，导致服务经常被拖累，核心功能不保障

2015年

- 垂直应用架构 -核心应用拆分
- 服务保障利器，限流、缓存、加机器
- PHP转JAVA化，剥离授权为独立应用
- 主站拆分，投稿、播放器、Feed流、评论

2016年

➢ ? ? ?

拆

拆

拆

3年的历经

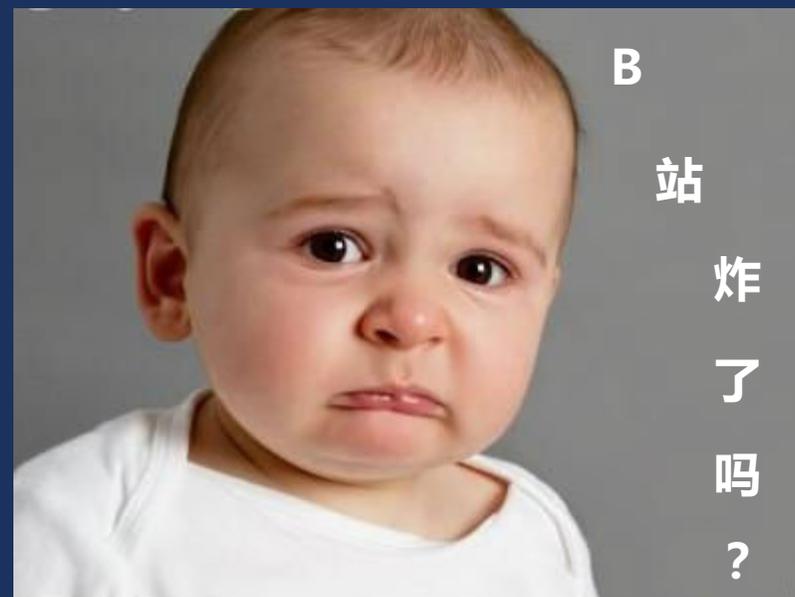
代码复杂度高，业务增长快...

nginx做均衡，上下线不透明也不滑哎

应用依赖严重，局部服务雪崩，扛不住全都等死吧

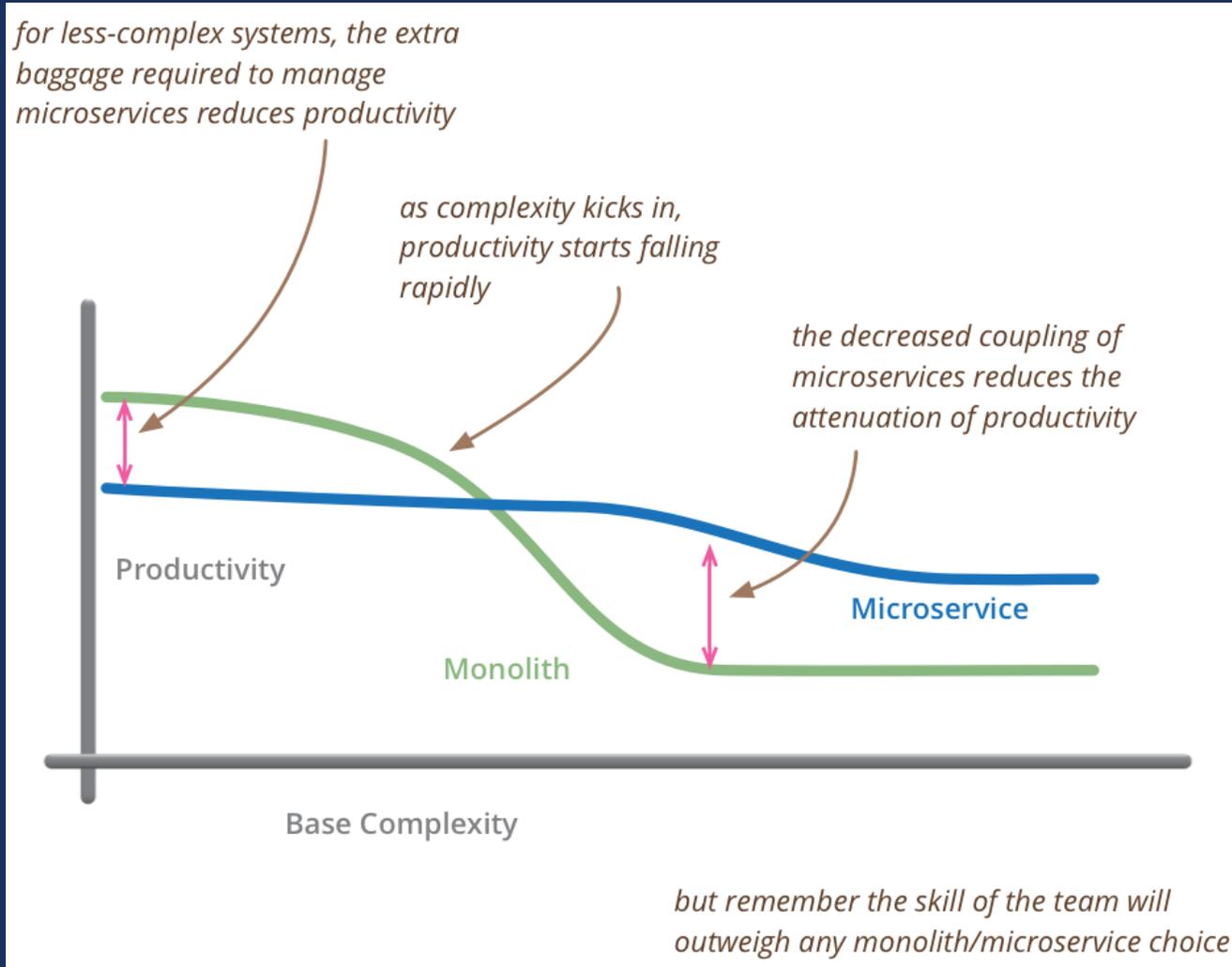
传统通信，http、restful、webservice，性能太差了

...



■ 寻求新方向

↓
micro services



■ 行业解决方案

➤ Google - GRPC

➤ Facebook - Thrift

➤ 阿里 - HSF、dubbo

➤ Twitter - Finagle

■ 为何要自建Saber

➤ 异构场景复杂

跨语言场景 (Java、Php、Python、Go.....)

➤ 高度服务化和可治理

动态限流，流量动态切换、屏蔽、故障转移等.....

➤ 快速接入 / 升级便捷

服务灵活上下线切换，Graceful重启.....

➤ 高性能支持

服务拆分后，请求分解为多个，性能上必须提高

■ 核心攻克

高性能，netty深入优化

服务治理，快速故障转移

跨语言协议，thrift+proxy的巧妙

■ 怎么做到高性能

- 请求响应的异步化
- 0Copy，消息堆外传输
- 串行无锁化，CS处理Codec
- 减少switch，去C端线程池
-

Status	Test Name	Script Name	Modifier	Start Time	Threshold	TPS	MTT	Err Rate	Vusers	
<input type="checkbox"/>	TODAY oauth-750	passport/oauth-appk...	郑志升	2016-02-29 12:15	00:03:00	2,841.9	263.29	0%	750	✘
<input type="checkbox"/>	TODAY oauth-750	passport/oauth-appk...	郑志升	2016-02-29 12:20	00:03:00	2,820.4	265.4	0%	750	✘
<input type="checkbox"/>	TODAY oauth-1000	passport/oauth-appk...	郑志升	2016-02-29 11:17	00:03:00	2,859.8	348.32	0%	1000	✘
<input type="checkbox"/>	oauth-1000	passport/oauth-appk...	郑志升	2016-02-29 11:11	00:03:00	2,824	352.71	0%	1000	✘
<input type="checkbox"/>	TODAY oauth-1500	passport/oauth-appk...	郑志升	2016-02-29 12:04	00:03:00	2,857.2	523.67	0%	1500	✘
<input type="checkbox"/>	TODAY oauth-1500	passport/oauth-appk...	郑志升	2016-02-29 12:08	00:03:00	2,808.8	532.13	0%	1500	✘
<input type="checkbox"/>	TODAY oauth-2000	passport/oauth-appk...	郑志升	2016-02-29 10:34	00:03:00	2,293.9	866.73	0.01%	2000	✘

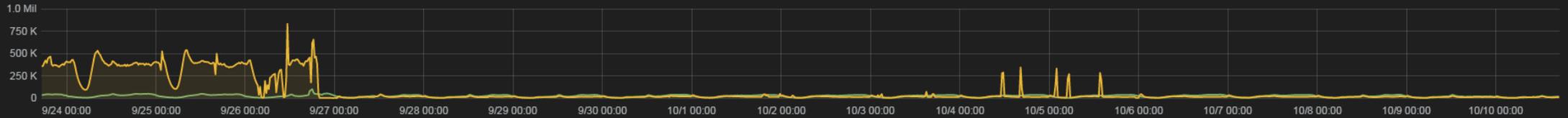
将近8倍性能提升

	TEST	VUSERS	TPS	MIT (ms)	ERRRATE
单节点压测	passport-1000	1000	17300	60	0%
	passport-2000	2000	17200	110	0%
	passport-3000	3000	17500	170	0%
	passport-4000	4000	17000	230	0%
	passport-5000	5000	17000	290	0%

server: aso-rpc-03 group: All api: All client: All

SABER接口访问量

saber接口访问量



aso-rpc-03.aso-shd.com-bilibili-aso-rpc-service-RpcOAuthService.msg-client
 aso-rpc-03.aso-shd.com-bilibili-aso-rpc-thrift-ThriftOAuthService.saber-client

min	max	avg	current	total
2 K	99 K	22 K	17 K	26.678 Mil
714	831 K	76 K	10 K	92.686 Mil

SABER接口访问耗时

saber接口访问耗时



aso-rpc-03.aso-shd.com-bilibili-aso-rpc-service-RpcOau
 aso-rpc-03.aso-shd.com-bilibili-aso-rpc-thrift-ThriftOauth

group: aso host: aso-rpc-01

Processes

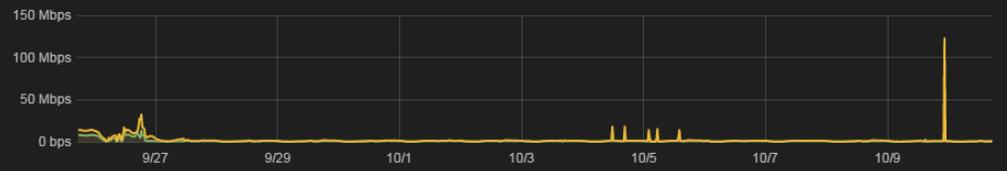
110

Processor load

0.019

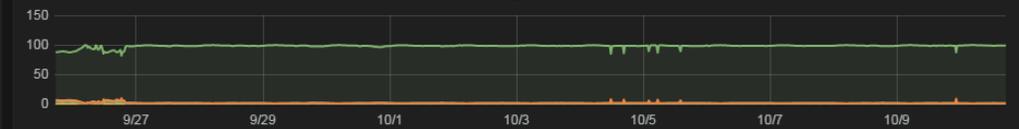
主机基础监控

network traffic



	min	max	avg	current	total
Incoming network traffic on eth0	143 Kbps	13.2 Mbps	1.1 Mbps	607 Kbps	1.0668 Gbps
Outgoing network traffic on eth0	189 Kbps	122.6 Mbps	2.0 Mbps	834 Kbps	1.9288 Gbps

CPU



	min	max	avg	current	total
CPU idle time	82	99	98	99	93.3994 K
CPU iowait time	0	0	0	0	0
CPU system time	0	6	1	0	651
CPU user time	0	9	1	1	1.0439 K

■ 治理服务优化

- 业务分P，服务分组
- 持续心跳，断链重连
- 服务熔断及访问限流
- 流量切换快速故障转移

RPC服务查询

RPC分组

aso-shd

服务分组，根据业务P级做好隔离

服务

com.bilibili.aso. [REDACTED]

查询 恢复 屏蔽

快速熔断服务



服务地址	分组	服务名	是否屏蔽
172. [REDACTED]	aso-shd	com.bilibili.aso. [REDACTED]	否
172. [REDACTED]	aso-shd	com.bilibili.aso. [REDACTED]	否
172. [REDACTED]	aso-shd	com.bilibili.aso. [REDACTED]	否
172. [REDACTED]	aso-shd	com.bilibili.aso. [REDACTED]	否
172. [REDACTED]	aso-shd	com.bilibili.aso. [REDACTED]	否
172. [REDACTED]	aso-shd	com.bilibili.aso. [REDACTED]	否
172. [REDACTED]	aso-shd	com.bilibili.aso. [REDACTED]	否
172. [REDACTED]	aso-shd	com.bilibili.aso. [REDACTED]	否
172. [REDACTED]	aso-shd	com.bilibili.aso. [REDACTED]	否



最大痛点是什么

语言种类多，如何解决异构

Python

Go

Java

Php

...

★ 灵活的协议，支持多平台

■ 解决异构的思路

✘ sdk封装，每个语言都得开发

★ proxy策略，但灵活性差难扩展

■ HTTP

★ 可读性好，易推广，唯一的好处

✘ 头部冗余，浪费流量，且序列化性能差

✘ 短连接，消耗了过多的握手建连时间

■ Protocol Buffer

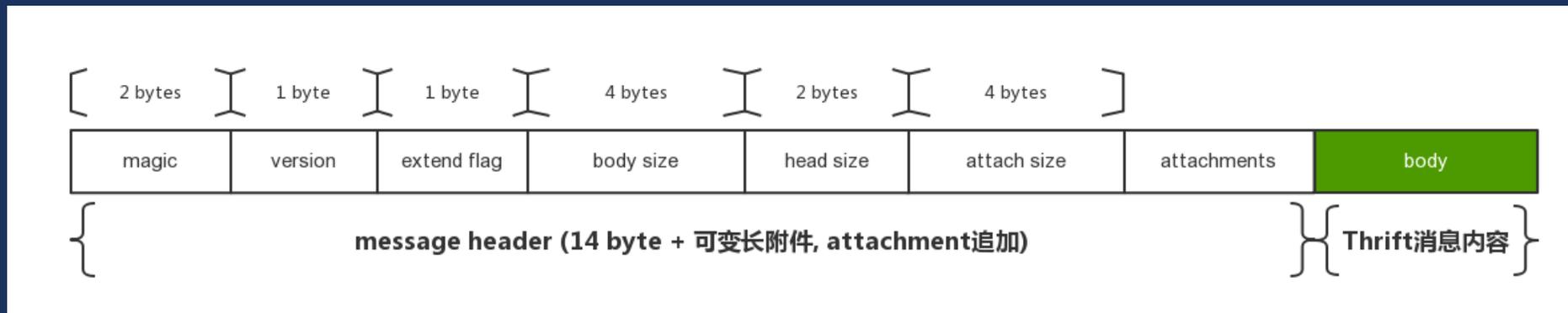
- ★ Google出品，TCP长链
- ★ 序列化性能高，仅支持GRPC
- ✘ 代码侵入性高，多语言支持较差

■ Thrift , 首选

- ★ Facebook出品 , TCP长链
 - ★ 序列化性能高 , 天然RPC通信
 - ★ 多语言支持良好 , 开发效率极高
-

■ 如何解决Saber框架下Thrift协议的原生化

Saber之Thrift协议设计



0 – 15 bit : magic

16 – 23 bit : version , 协议的版本号

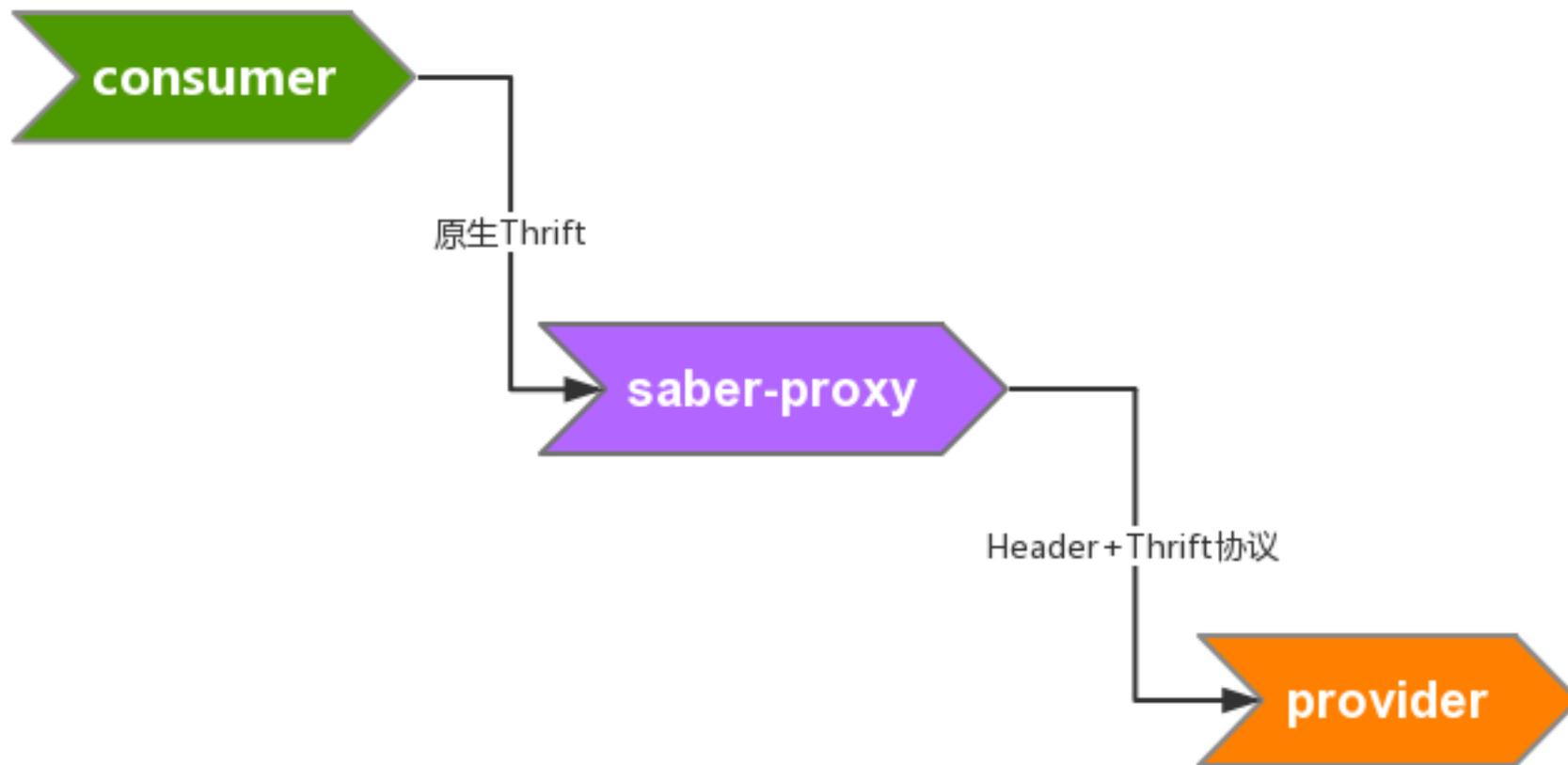
24 – 31 bit : extend flag , 含heartbeat、event类型等

32 – 63 bit : body size

64 – 79 bit : header size

80 – xx bit : attachments , 内容通过k-v存储

Saber协议通信流程



■ 协议巧妙的设计

- Thrift原生化
- Proxy和协议无关
- 中间层损耗低，20%
- Proxy实现HA、Loadbalance

Saber功能介绍

配置管理

- 继承配置
- 可配置可编程

服务匹配

- 服务分组、多版本
- 指定调用
- 只订阅

集群&&容错

- 集群loadbalance
- 容错规则
- 限流
- 降级

多协议

- 不同服务不同协议
- 同一服务多协议
- 本地服务调用

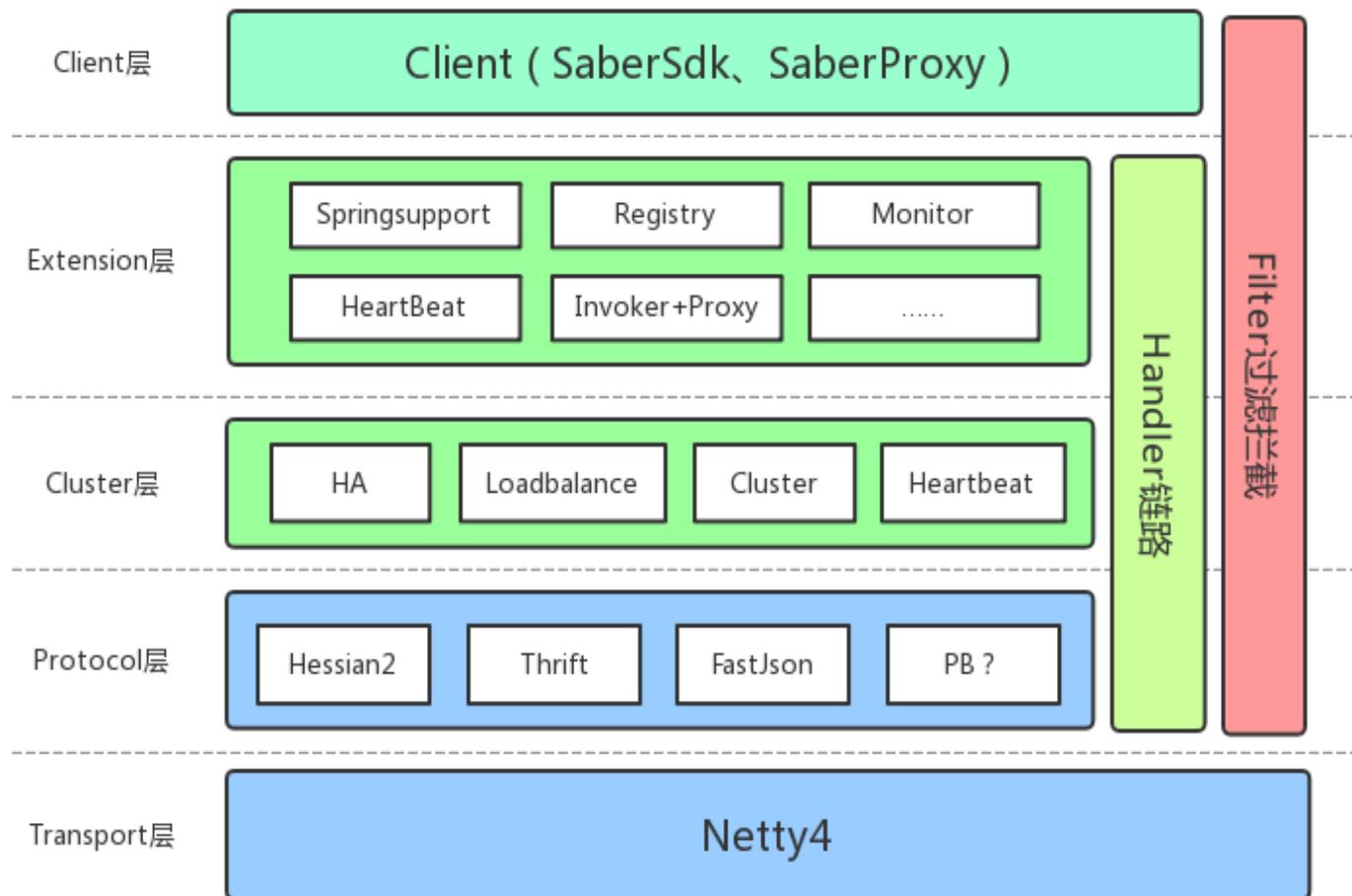
治理管控

- 流量动态切换
- 动态指令（灰度、屏蔽）
- 分组和隔离
- 故障转移

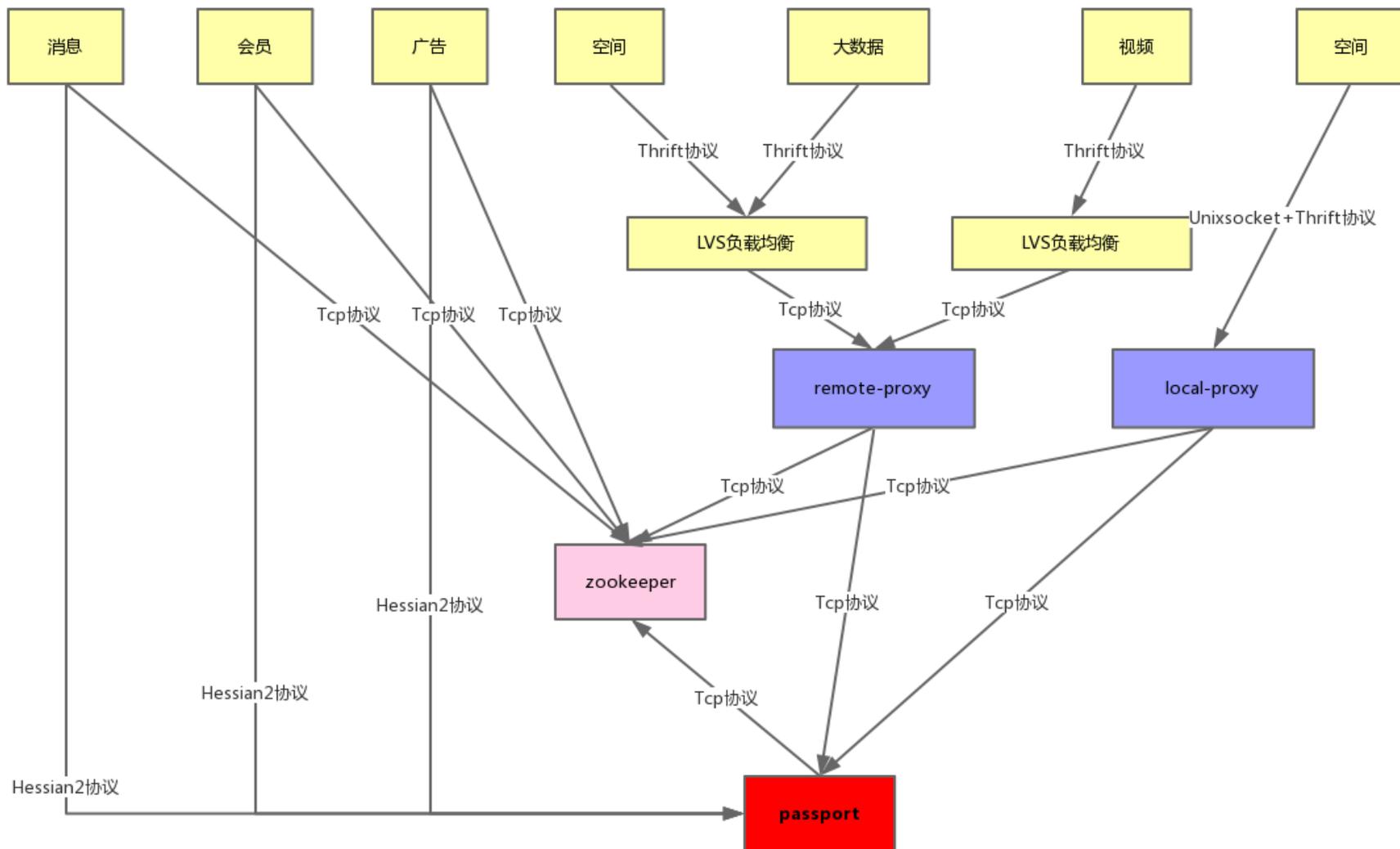
组件套餐

- Saber-proxy
- Mock Server

■ Saber系统架构



线上应用实施



未来的尝试

- 完善监控与报警，如调用链
 - 改善接入，多语言SDK的provider
 - 服务依赖可视化、配置管理及动态下发
 - 拥抱业界，尝试开源，期待更多基友加入
-

总结

- 服务化是痛并快乐的过程，必须得走，一步一脚印！
- 没有一成不变的神作架构，适合自己的才是好的！

谢谢

Thank You

微信：

