酷狗音乐 KUGOU.COM

分布式调用跟踪系统 zipkin

- 基础服务小组—陈晓辰

Zipkin是什么

Zipkin的使用场景

Zipkin的实现

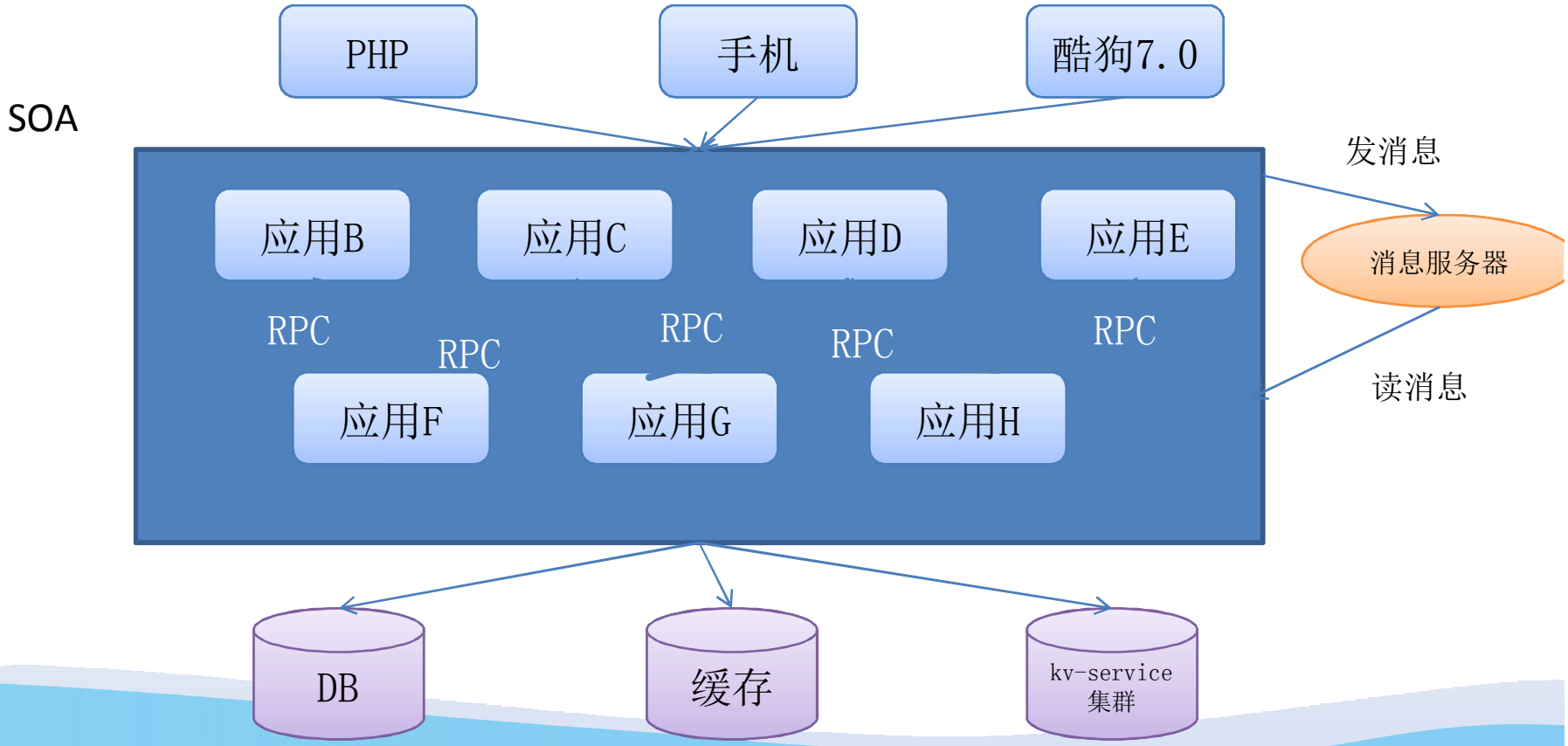
酷狗音乐 KUGOU.COM

背景

- 日趋复杂的分布式系统
  - 服务框架（SOA）
  - 数据层
  - 分布式缓存
  - 分布式存储
  - ……

当前复杂调用链

SOA


PHP
手机
酷狗7.0
应用B
应用C
应用D
应用E
RPC
RPC
RPC
RPC
RPC
应用F
应用G
应用H
发消息
消息服务器
读消息
DB
缓存
kv-service
集群

如何整理跟踪这些后端调用关系？

举个例子

- 设想高速收费站将车辆通行信息记录成日志

[2013-05-01 12:23:34] 鲁A123BC,平度2,S16,济南,￥12
[2013-05-01 12:23:40] 鲁A987DE,平度2,S16,淄博,￥10
[2013-05-01 12:43:15] 鲁A123BC,潍坊1,G20,济南,￥18
[2013-05-01 13:38:29] 鲁A123BC,青州西1,G20,济南,￥10
[2013-05-01 13:38:30] 鲁A567AB,青州西2,G20,潍坊,￥10
[2013-05-01 14:39:27] 鲁A123BC,淄博3,G20,淄博,￥15
[2013-05-01 16:42:58] 鲁A123BC,济南3,G20,济南,￥25
……

举个例子

- 可以得到
  - 收费站的每日总车流量和流量趋势
  - 鲁A123BC在五一期间的行驶路线和费用
  - G20上的车速、路况
  - G20流量过高时，车的来源分布

举个例子

- 高速上行驶的车辆：前端请求
- 高速上的收费站：处理请求的应用
- 关键点：关联日志中记录的车牌号

Zipkin

- 分布式调用链跟踪系统
- Google Dapper 论文的实现
- 核心：调用链，每次请求都生成一个全局唯一的ID（trace_id），通过它将不同系统的“孤立的”请求串在一起，重组成调用链

酷狗音乐 KUGOU.COM

简介

- 覆盖了繁星当前后端主要使用的网络通信
- 数据库：MySQL
- 缓存：memcache
- Thrift RPC

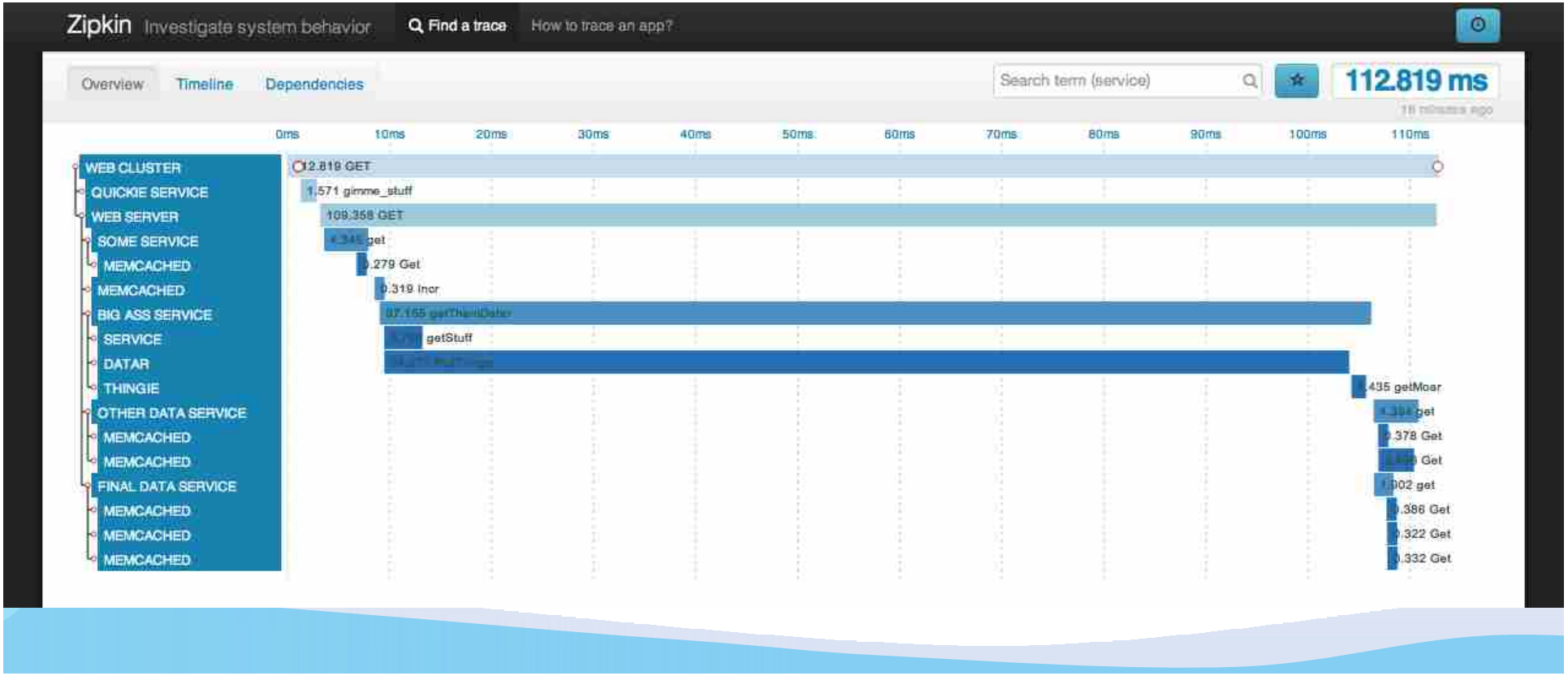
Zipkin是什么

Zipkin的使用场景

Zipkin的实现

酷狗音乐 KUGOU.COM

调用链跟踪



调用链跟踪

上图一行代表一次网络调用，点击其中某一行看该次调用详情

酷狗音乐 KUGOU.COM

调用链跟踪

- 排查前端某页面响应很慢或报错的原因
  - 查看这个页面的调用链，定位瓶颈点、故障点
- 实时跟踪当前客户端的所有请求的调用链
  - 了解每个请求背后的应用间交互过程

调用链跟踪

- 依赖关系
  - 应用直接或间接依赖了哪些服务
  - 各个层次上的依赖的调用指标和错误指标
  - 找出调用链路上的不正常的、多余的依赖调用
- 异常分析
  - 依赖会产生哪些异常
  - 异常时会造成什么影响
  - 把主流程的强依赖转化为弱依赖

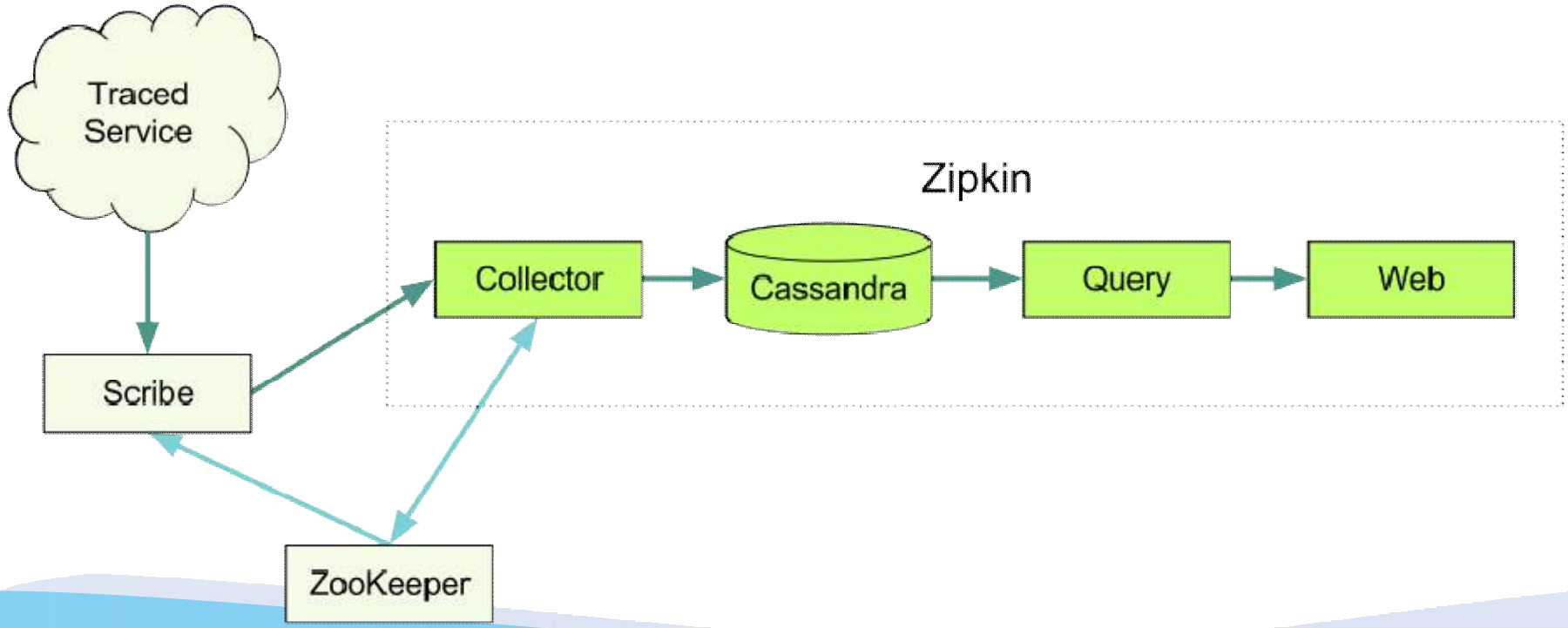
Zipkin是什么

Zipkin的使用场景

Zipkin的实现

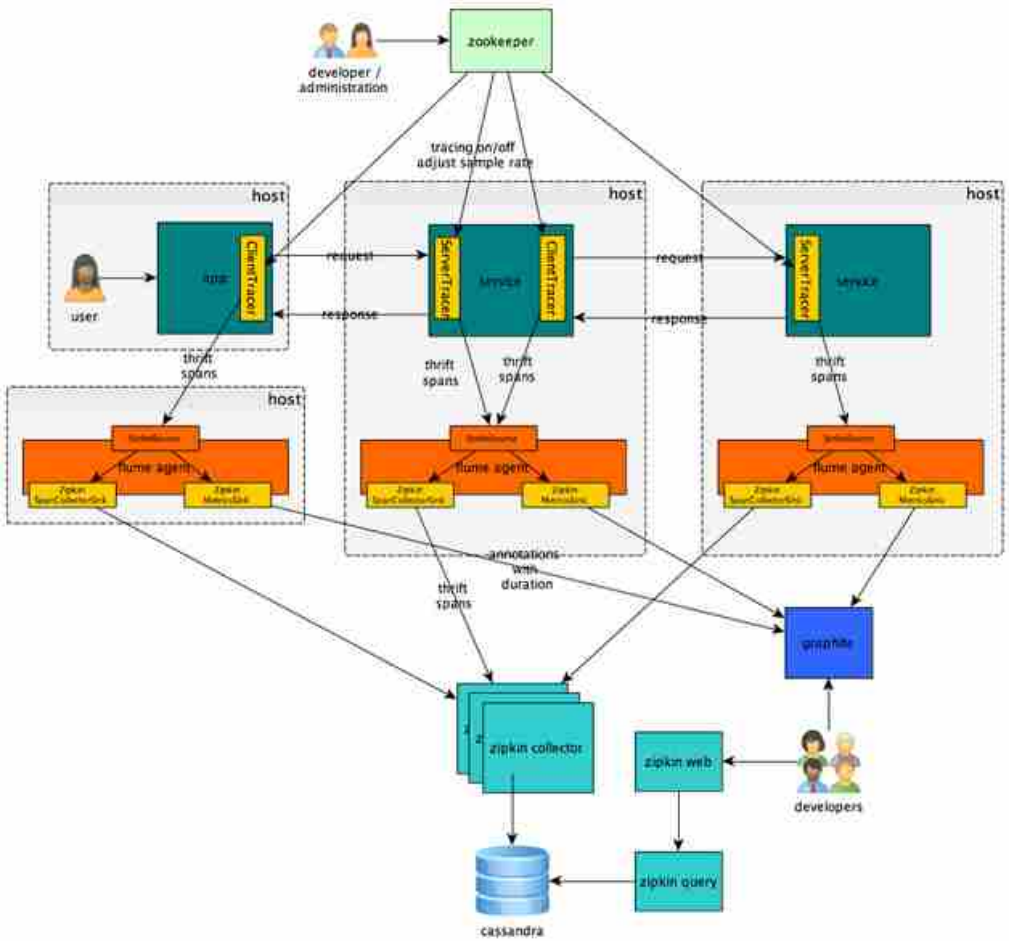
酷狗音乐 KUGOU.COM

架构



酷狗音乐 KUGOU.COM

架构

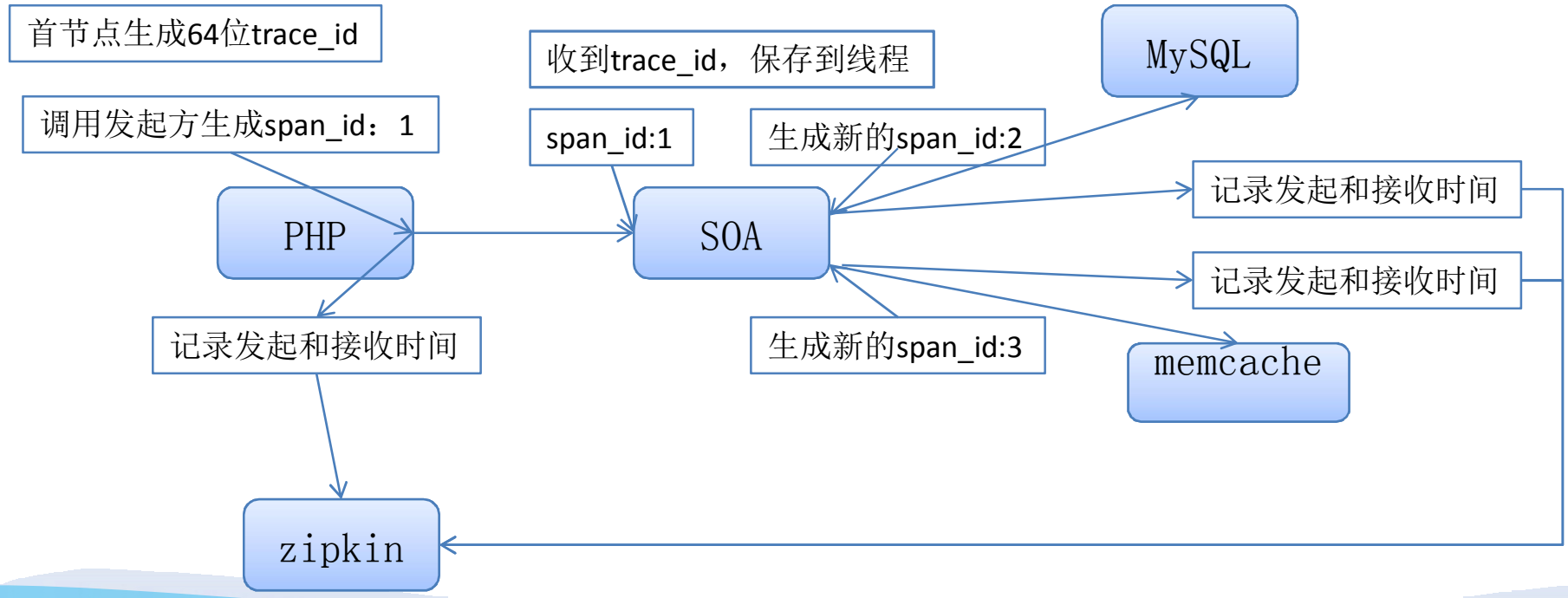


协议

为了能在调用时跟踪到trace_id，在调用时的http header里面传递以下几个参数：
X-B3-TraceId
X-B3-SpanId
X-B3-ParentSpanId
X-B3-Sampled （true或者false，表示该链路需不需要被采样，如果为false，则下面的链路都不会把跟踪信息发送到zipkin）
X-B3-SpanName

酷狗音乐 KUGOU.COM

举例



酷狗音乐 KUGOU.COM

举例

客户端


```
{
    "span": {
        "id": 2132567190233127700,
        "name": "/test/a",
        "trace_id": 1634564891533127700,
        "parent_id": 1634564891533127700,
        "annotations": [
            {
                "host": {
                    "ipv4": 2130706433,
                    "port": 8080,
                    "service_name": "zipkin"
                },
                "timestamp": 1411607696245000,
                "value": "cs"
            },
            {
                "host": {
                    "ipv4": 2130706433,
                    "port": 8080,
                    "service_name": "zipkin"
                },
                "timestamp": 1411607696245000,
                "value": "cr"
            }
        ]
    }
}
```


服务端


```
{
    "span": {
        "id": 2132567190233127700,
        "name": "/test/a",
        "trace_id": 1634564891533127700,
        "parent_id": 1634564891533127700,
        "annotations": [
            {
                "host": {
                    "ipv4": 2130706433,
                    "port": 8080,
                    "service_name": "zipkin"
                },
                "timestamp": 1411607696245000,
                "value": "sr"
            },
            {
                "host": {
                    "ipv4": 2130706433,
                    "port": 8080,
                    "service_name": "zipkin"
                },
                "timestamp": 1411607696245000,
                "value": "ss"
            }
        ]
    }
}
```

采样率

- 每个节点都配置了采样率
- 当前一个请求head中带了X-B3-Sampled，为true时，本节点必采样，并传递该参数。
- 当X-B3-Sampled为false时，认为该条链路都不采样，则不采样，但是依然传递该参数
- 当X-B3-Sampled没有时，认为是没有上文的请求，则在本节点用采样率计算是否采样，然后传递该参数

酷狗音乐 KUGOU.COM

现存在问题

- PHP没有多线程，上报zipkin不能在框架上异步
- 服务器时间必须同步
- 在zipkin上层采用flume日志系统做缓冲，经常发生无法连接的问题。