

# 海量数据推送服务

实时处理架构的演进

SPEAKER

许建辉





#### 促进软件开发领域知识与创新的传播



#### 关注InfoQ官方信息

及时获取QCon软件开发者 大会演讲视频信息



[北京站] 2016年12月2日-3日

咨询热线: 010-89880682



[北京站] 2017年4月16日-18日

咨询热线: 010-64738142

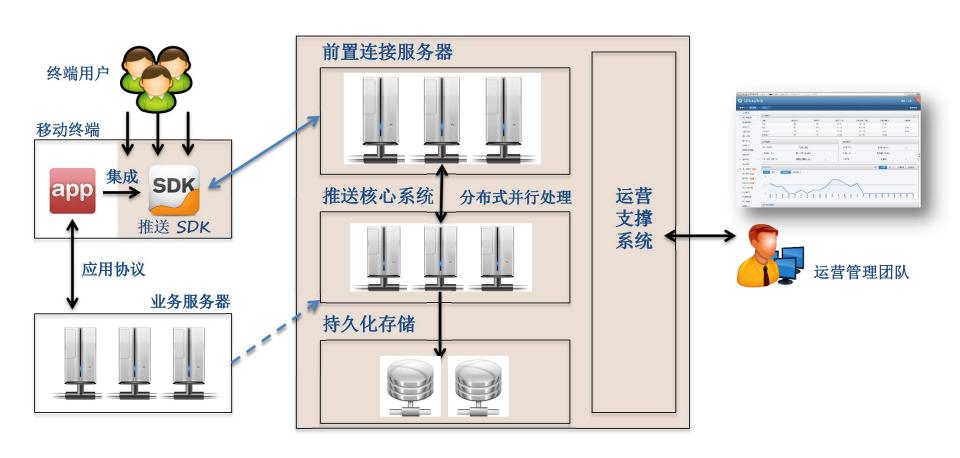
#### 自我介绍

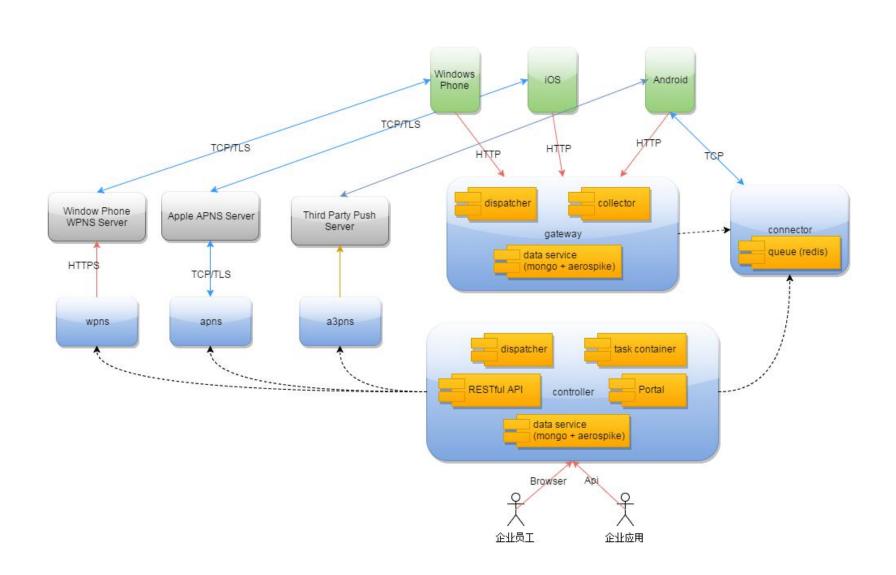
- 近10年开发经验,曾从事财务软件、数字电视、婚恋社交相关软件研发。
- 赢才骏励 原CTO。
- TD 资深研发工程师,推送服务解决方案专家,负责 TD 推送服务后端的研发。
- 2015 年初加入 Push 项目,负责后台系统研发。

### 演讲内容

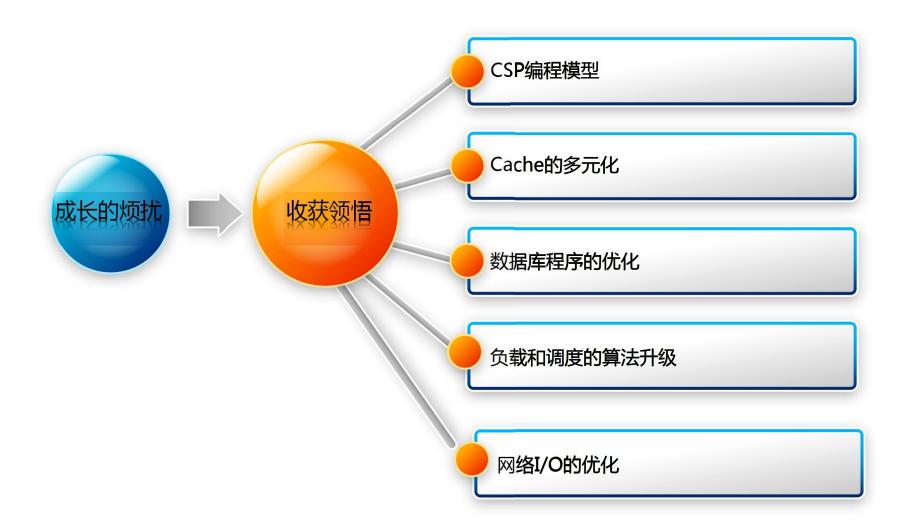
- Td Push 架构
- CSP模型介绍
- 多元化的Cache
- 数据库的性能优化
- 任务分发与调度的演进
- 网络I/O的优化



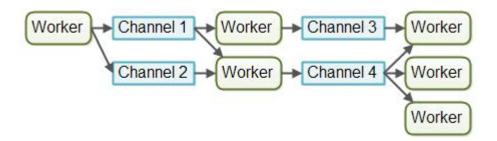








· Communicating Sequential Processes



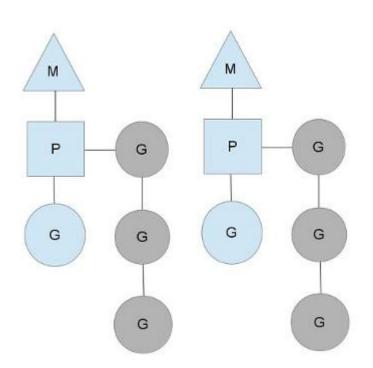
• 模型中,worker之间不直接彼此联系,而是通过不同channel进行消息发布和侦听。消息的发送者和接收者之间通过Channel松耦合,发送者不知道自己消息被哪个接收者消费了,接收者也不知道是哪个发送者发送的消息。

函数

- 函数是一等公民
- 函数式编程
- 利用闭包减少全局变量
- go func(){}

GoRoutine

- 历史悠久,且多种语言均有自己的实现办法
- 内存占用和线程的区别
- 系统调度的区别

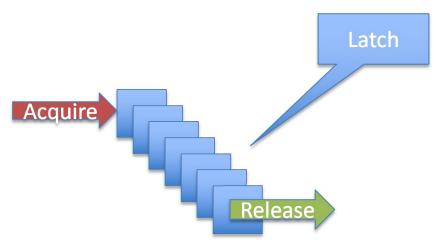


#### Channel

- 在CSP模型中, channel是一个数据传输的纽带
- 在申明channel的时候,可以指定channel的只读、只写、读写权限和buffer大小
- 在对channel进行操作的时候,可以使用阻塞和非阻塞的方式
- 关键字: Select、Range、Default
  - Select: 同时操作多个Channel
  - Range: 遍历Channel
  - Default: 当Select里面存在阻塞,则会执行Default
- Channel的应用
  - 基于Channel返回大量的数据
  - 使用Channel实现超时、心跳等
  - 基于Channel实现Latch,控制并发数
  - 使用Channel实现Recycling memory buffers

#### Channel

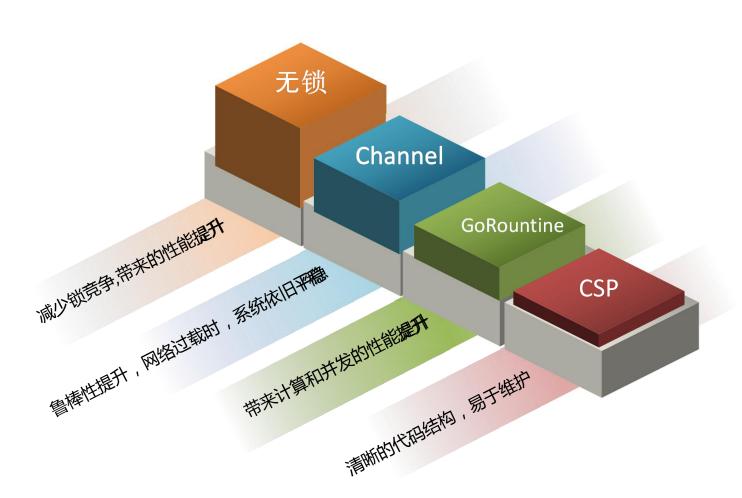
基于Channel实现Latch,控制并发数



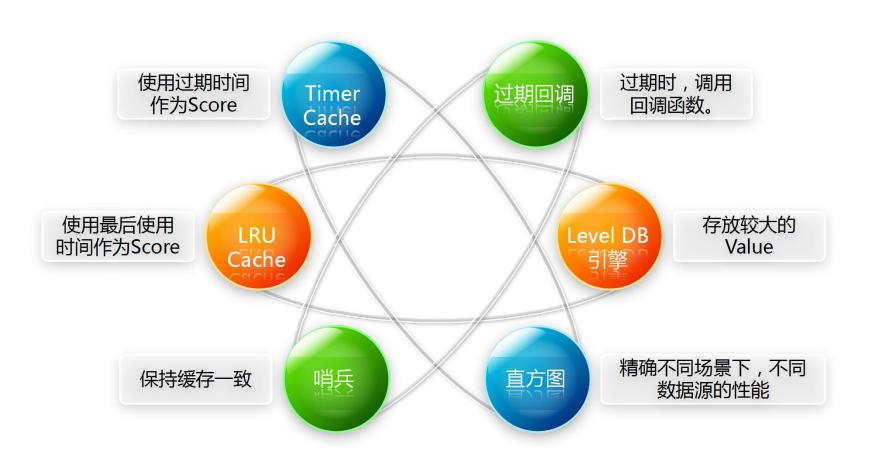
基于Channel实现Recycling memory buffers

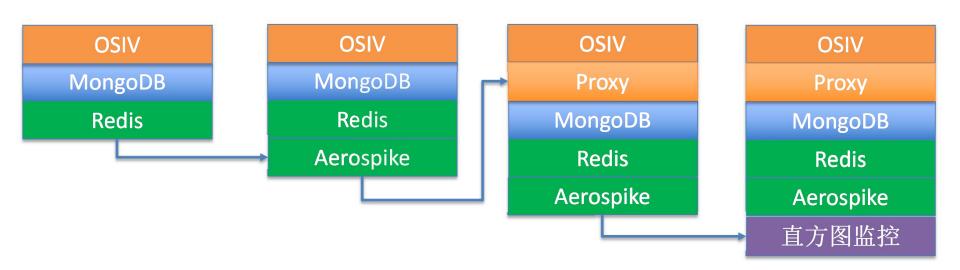


收益



### 多元化的Cache







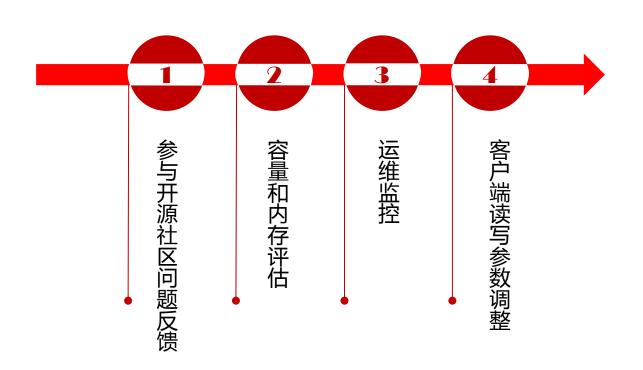
request->	response->
request->	response->
	占用Mongo资源
	等待Mongo连接资源

request->	response->
request->	response->
	占用Mongo资源
	等待Mongo连接资源

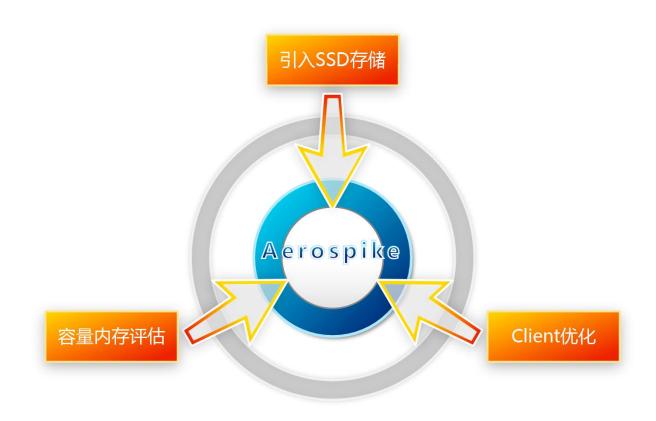


- 沿用了OSIV
- 使用了代理连接
- 代理连接支持自动打开和重复关闭
- 在真正操作数据库的时候,建立连接

Aerospike使用优化



Aerospike使用优化





#### 任务分发与调度的演进

#### Dispatch

当一个终端连接上来的时候,根据Connector当前的负载和权值,进行连接地址的指定。

#### 任务分发

给每个Connector进行全量的消息发送。 对每个当前不在线的设备,需要在shuffle后进行第二次的发送。

#### 健康检查

需要检查当前Connector的负载,并记录。

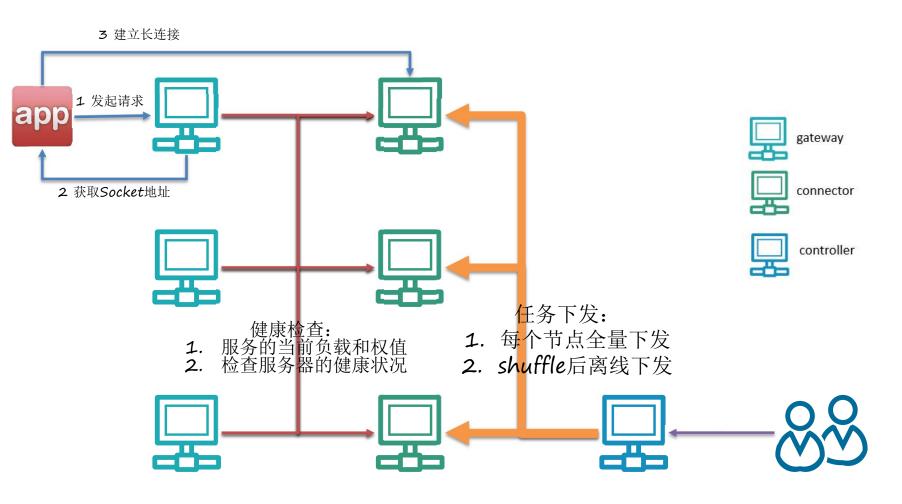






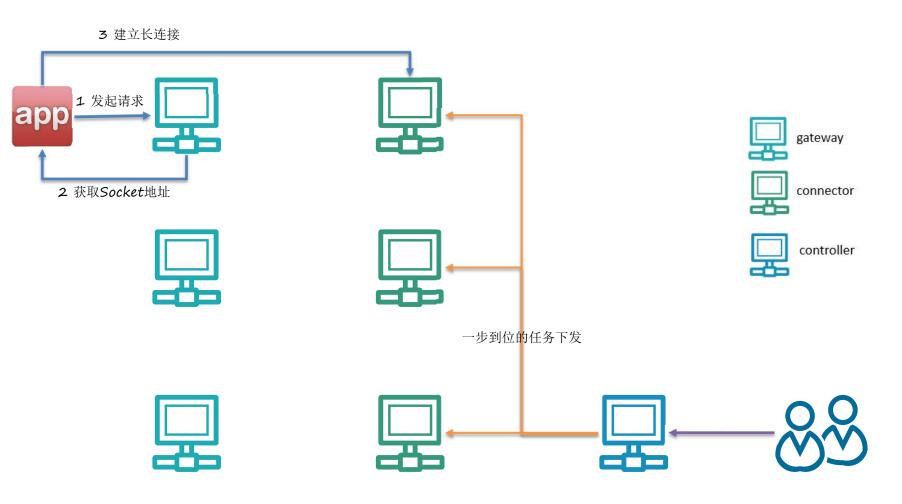
一致性哈希算法

## 任务分发与调度的演进



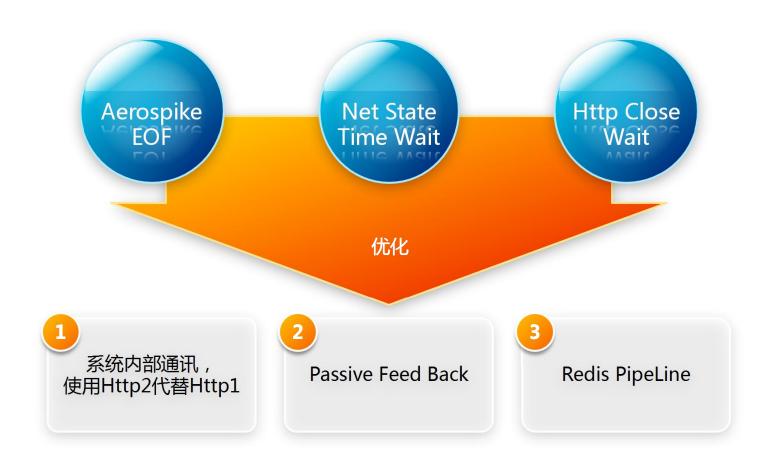
基于权值和健康检查的路由和分发

## 任务分发与调度的演进



基于一致性哈希算法的路由和分发

#### 网络1/0的优化





International Software Development Conference