参考：<http://weibo.com/p/1001643877261186063112>

新兵训练营系列课程——Feed架构介绍

2015年8月26日 14:28 阅读 15595

微博作为当今中国最大的社交媒体平台，每天都有上亿人访问。而Feed作为微博最为核心的功能，每天有高达数十亿的请求，高峰期每秒的请求量上万。如何设计高可用的Feed系统，来应对如此高并发的访问，极具挑战（me:单机肯定不能满足要求，需要使用分布式系统）。

前：负载均衡 反向代理

中：分布式消息中间件 分布式缓存 分布式RPC

后：分布式文件系统 分布式数据库

其他：分布式session 共享 监控 处理突发事件

要求：微服务化 模块化

下面，我们将具体探讨微博Feed系统的架构，希望通过这次讲解，大家能够掌握微博Feed系统架构的基本知识。

大纲

* 初探Feed
  + what？why？how？
* Feed业务体系
  + 核心业务
  + 关联业务
* Feed架构设计
  + Feed存储
  + Feed聚合
* Feed未来发展
  + 业务发展
  + 架构发展

一 初探Feed-What is feed？

从广义上讲，Feed是指为满足用户以某种形式持续获得更新信息的需求而提供的格式标准的信息出口。

1. **PC Feed**



1. **mobile feed**

微博体系中，Feed是指用户通过关注关系，聚合好友最新微博以供自己消费的信息服务，其中也包括分享微博。



**分享**



**倾听**



**看新闻**

****

**围观**



二、Feed业务体系

**1、微博**



1. **page**



**3、Card**



1. **计数器**



1. **未读数、话题、分组、转评赞、广告和趋势**



三、Feed系统架构设计

**1. 如何设计一个高可用的Feed系统？**

* 如何存储用户发表的微博
* 如何聚合关注用户发表的微博

2. Feed存储

**Feed特征**

* 每天发博量大：上亿条
* 发博峰值高：几万/s
* 并发访问量大：几万/s
* 访问热点集中：最近三天占99%

**存储选型**



**存储特点**

* 持久化：MySQL、Redis、HBase
* 并发读QPS：几百万/s：Memcached、Redis
* 最近三天热点数据：Memcached

**实现方案 Memcached+MySQL**

Memcached存储最近三天热点数据、MySQL存储全量数据

3、Feed存储-MySQL

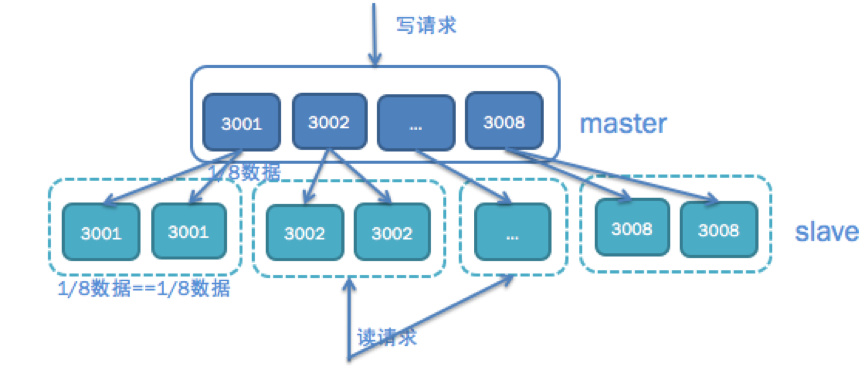
**分库分表**

* 按访问用户维度hash来分库
* 按访问时间维度来分表

**部署形式**

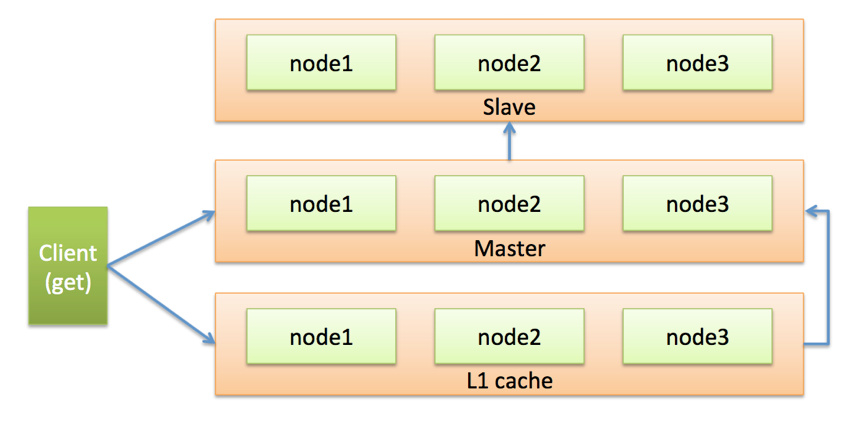
* 多主多从
* 读写分离

**HA**

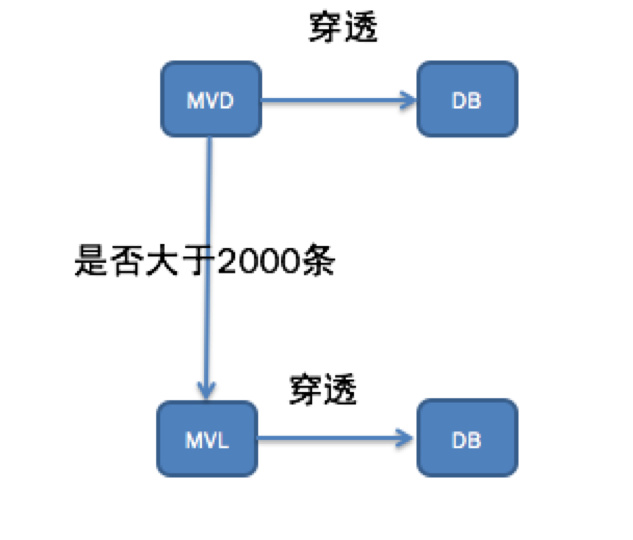
* 主从机制、自动切换
* backup离线灾备
* 

4、Feed存储-Memcached

**高可用性架构**



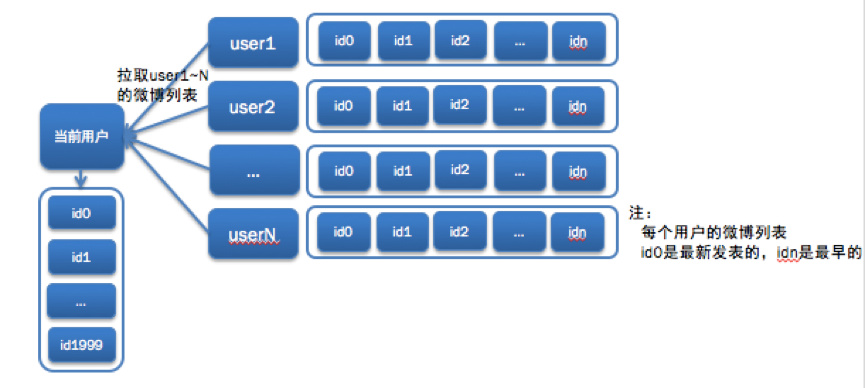
**业务架构**

* mvd：存储用户最近15天50条。
* mvl：存储用户最近200条。
* 

5、Feed聚合

如何聚合关注用户的微博？

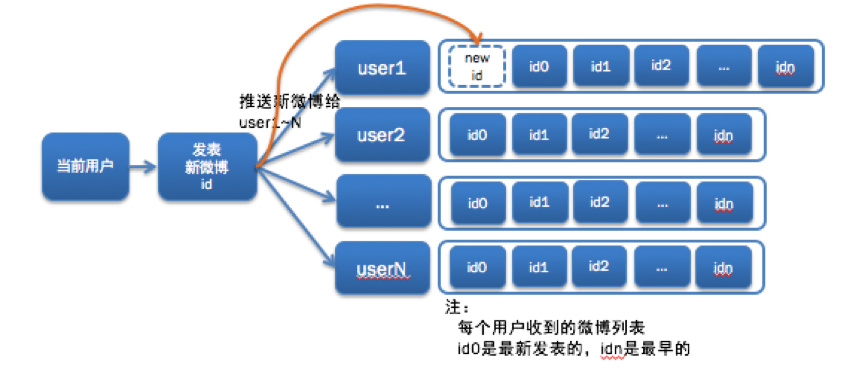
**拉模型**



**优缺点**

* 优点：架构简单、一致性好、写入过程开销小
* 缺点：写少读多，网络开销大，资源读取量大

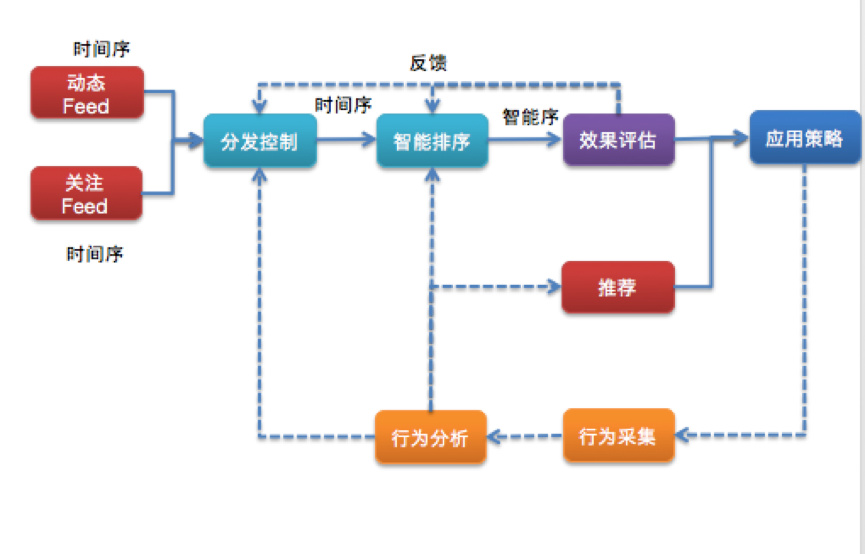
**推模型**



**优缺点**

* 优点：下行网络开销小，资源成本低
* 缺点：写入成本较高，资源写瓶颈高

四 未来



**业务发展**

* 分发控制
* 智能Feed
* 实时反馈

**架构发展**

* 存储架构
* 消息架构
* 可插拔
* cell模型

作业

        设计一个简化的Feed系统。

– 10亿用户，每天发表1亿条微博。

– 平均每秒500次feed请求，平均每个人关注10个用户。

– 平均每秒20000次feed请求，平均每个人关注200个用户。

– 针对上面两种情况，分别设计feed系统，要求给出存储选型和对应的机器数量。

**------------------新兵训练营简介**------------------

微博平台新兵训练营活动是微博平台内部组织的针对新入职同学的团队融入培训课程，目标是团队融入，包括人的融入，氛围融入，技术融入。当前已经进行4期活动，很多学员迅速成长为平台技术骨干。

微博平台是非常注重团队成员融入与成长的团队，在这里有人帮你融入，有人和你一起成长，也欢迎小伙伴们加入微博平台，欢迎私信咨询。

**------------------讲师简介**------------------

胡忠想，微博名[@古月中心相心](http://weibo.com/n/%E5%8F%A4%E6%9C%88%E4%B8%AD%E5%BF%83%E7%9B%B8%E5%BF%83)，目前任职于新浪微博平台研发及大数据部，主要负责微博 Feed 策略相关工作，曾先后参与微博 Feed 存储、微博计数器、微博阅读数等重大业务产品的开发。2012年3月份毕业于北京航空航天大学计算系，同年4月份，加入新浪微博并工作至今。业余爱好户 外，曾徒步过贡嘎、雨崩，攀登过四姑娘三峰。新兵训练营第一期学员。

**------------------下面文字很重要**------------------

**招聘职位**：系统研发工程师Java/C++

**隶属部门**：微博平台研发

**工作描述**：

1. 负责微博核心功能服务的技术支撑，包括：feed，评论，用户关系，开放平台认证授权体系及服务开放策略，微博通讯服务等，为微博主站，移动客户端，第三方应用等提供高效稳定的功能接口及服务。

2. 专注于分布式系统算法和架构，解决高并发、高性能场景下的系统架构问题；处理百亿级/天的服务接口请求，以及百万级/秒的峰值请求。

3. 专注于分布式存储方向，解决海量数据存储过程中的容量及性能瓶颈问题；处理千亿级微博feed、用户关系等在线数据的存储与处理，以及快速的数据增长问题。

4. 致力于服务中间件的开发，提供高内聚的解决方案，以帮助各个核心功能服务简化服务架构，降低系统复杂性，保证服务稳定性。

5. 利用大数据处理技术，为开放平台产品功能及开放策略提供数据分析支持，促进微博开放平台服务更好用户体验。

**职位优势及技术挑战**：

1 良好的技术氛围，定期的技术沙龙和架构分享，促进内部学习成长。

2 完善的技术成长培训体系，新兵训练营活动已经为平台培养多名技术骨干，多名同学成长为业界技术专家。

3 足够的施展空间：十亿级PV、千亿级数据、4w qps以上的队列处理能力、千级以上的服务器规模、百亿级数据HBase存储、千台以上的Docker集群，确保有持续充足的挑战

**任职资格**：

1. 计算机或相关专业，本科及以上学历

2. 有快速学习能力，及较强的解决问题能力，能够在承受压力情况下完成解决问题

3. 掌握Java开发基础，熟悉多线程开发，熟悉Spring框架，理解IOC原理。

4. 有良好的数据结构和算法基础

**简历请发至：qingfeng@staff.weibo.com 或 daoru@staff.weibo.com**