微博系统中”微博评论“的高性能高可用计算架构

1. 业务特点分析

微博的评论对时效性要求不高，但需要可靠的存储，不能丢失。因为评论会作为数据分析来使用。同时也是个人言行的记录。

1. 行为建模和性能估算

考虑到微博是一个看得多发的少的业务，假设平均每天每人发1条微博（只考虑文字微博），则微博每天的发送量约为2.5亿条。 普通的每条微博的评论，平均10条。则评论的次数为2.50亿\*10=25亿。大部分的人发微博集中在早上 8:00~9:00 点，中午 12:00~13:00，晚上 20:00~22:00，假设这几个时间段发微博总量占比为 60%。

评论微博的TPS：25亿\*60%/(4\*3600) = 100K/s

1. 业务特性分析

需要有负载均衡，在热点评论的时候，在高峰的时候，考虑限流。

1. 架构分析

1、微博评论的时候依赖登录状态，登录状态一般都是保存在分布式缓存中的，因此微博评论的时候，将请求发送给任意服务器都可以， 这里选择“轮询”或者“随机”算法。

2、业务服务器数量估算 按照一个服务每秒处理 500 来估算，完成 100K/s 的 TPS，需要 200 台服务器，加上一定的预留量，220 台服务器差不多了。

1. 非热点事件时的高性能计算架构

不考虑服务拆分，同时写评论的话，重点关注复杂均衡和数据分布。

在读取微博的时候，也会同时读取评论，所以写评论不需要涉及缓存

多级负载均衡架构

图形用户界面

描述已自动生成

六、热点事件时的高可用架构：

1、热点事件的评论，这个在业务上实时性要求不高，为了防止浏览评论时刷爆系统，可以限制只看到top 20的，15分钟内的评论；发评论遇到高并发时，只需要尽量少丢失请求即可，考虑使用”漏桶算法”的限流策略，并且设置尽量大的缓存队列。

2、另外，由于热点事件的突发性，也需要做好预防，比如可以用一些算法来监控、预测一些热点事件的评论，发现异常后，及时进行系统的限流、排队甚至熔断的机制，来保证微博系统的正常运行。