

极客时间 架构实战营

模块五作业

姓名： 李天水

学号： G20210607050173

用户行为建模和性能估算

- 用户量
2020.9 月月活 5.11 亿，日活 2.24 亿（参考《微博 2020 用户发展报告》）。
- 关键行为
 - 写评论
 - 读评论
- 性能估算
 - 写评论
大部分的人读写评论的时间应该会 and 发微博、看微博的时间相同，假设每人每天会评论两次，那么在集中时间段，写评论的 TPS 计算如下：
$$2.5 \text{ 亿} * 2 * 60\% / (4 * 3600) = 20\text{K/s}$$
 - 读评论
因为只有前排的少数评论才会被看到，所以假设每条评论的平均查看次数为 50 次，读评论的 QPS 计算如下：
$$2.5 \text{ 亿} * 2 * 50 * 60\% / (4 * 3600) = 1\text{M/s}$$

非热点事件的高性能架构

- 写评论
 - 业务特性分析
写评论是一个典型的写操作，因此不能用缓存，可以用负载均衡。
 - 架构分析
用户量过亿，应该要用多级负载均衡架构，覆盖 DNS→F5→Nginx→网关的多级负载均衡。
 - 架构设计
 - ◆ 负载均衡算法
轮询、随机、最少请求。
 - ◆ 业务服务器数量估算
写评论同样需要内容审核、写入存储系统及缓存，因此同样按照一个服务每秒处理 1000 来估算，完成 20K/s 的 TPS，需要 20 台服务器，考虑到一定的冗余，需要 25 台服务器。
- 读评论
 - 业务特性分析
读评论是一个典型的读场景，可以同时考虑负载均衡及缓存架构。
 - 架构分析
 - ◆ 用户量过亿，应该要用多级负载均衡策略。
 - ◆ 请求量达到 50 亿，应该要用多级缓存架构。
 - 架构设计
 - ◆ 负载均衡算法

轮询、随机、最少请求、哈希。

◆ 业务服务器估算

假设 CDN 能够承载 90% 的用户流量，负责读服务器的业务处理逻辑相对简单，假设单台服务器处理能力是 1000/s，则需要的服务器数量为 100 台，考虑到冗余，需要服务器 120 台。

➤ 整体架构设计

■ 任务分配

双机房、三机房。

■ 任务分解

写评论和读评论拆分到不同服务，避免读评论请求过多导致不能正常相应写评论请求。

热点事件的高可用架构

➤ 写评论

可以考虑采用“漏桶算法”进行限流。

➤ 读评论

延长热点事件相关的原微博及热门评论的缓存有效时间，根据哈希转发请求来提高缓存的命中率。