1. 有3位理发师
2. 每位理发师剪头一次需要30min
3. 每次同一时刻会有不同数量的顾客进店要求剪发
4. 顾客要求苛刻，最多等待时间30min

讨论顾客享受剪头服务的时间变化

不同数量的顾客进店，享受服务器功能的所用时间：

|  |  |
| --- | --- |
| 进店顾客数 | 剪头所用时间 |
| 1 | 30min |
| 2 | 30min |
| 3 | 30min |
| 4 | 1h |
| 5 | 1h |
| 6 | 1h |
| 7 | 1.5h |
| 8 | 1.5h |
| 9 | ... |

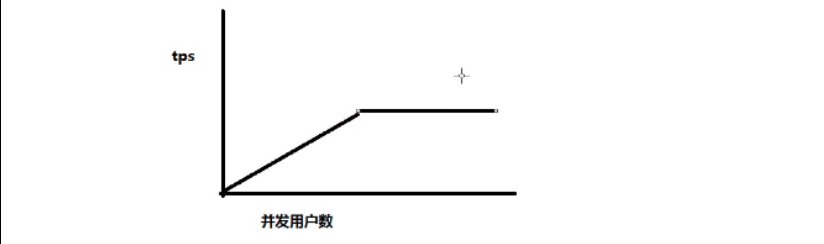
性能指标1：同时进店剪头顾客数-->并发用户数

从进店到完成箭头的总时间（等待时间+执行时间）--响应时间

|  |  |
| --- | --- |
| 同时进店客户数 | 1h内完成的剪头数量 |
| 1 | 1 |
| 2 | 2 |
| 3 | 3 |
| 4 | 4 |
| 5 | 5 |
| 6 | 6 |
| 7 | 6 |
| 8 | 6 |
| 9 | ... |

性能指标2：单位时间内完成事务的数量（事务：剪头）

**单位时间事务数 tps**



总结：

理发店：服务器运行的程序

顾客：从客户端发来的请求

由多个浏览器同时发送到服务器的任务：并发用户

服务程序处理一个浏览器请求的时间：平均事务响应时间，特点：随并发用户的增加而增加

单位时间内服务程序完成客户端请求的数量-->单位事务数随着并发用户的增加而增大，当并发用户数量达到一定量后不再增加