



《手写OS操作系统》小班二期招生，全程直播授课，大牛带你掌握硬核技术！

点此查看

🔍

从所有教程的词条中查询...

全部开发者教程

13. es集成到mxshop\_srvs

14. 倒排索引算法 - 扩展文档

第17周 分布式理论基础、分布式事务解决方案

1. 事务和分布式事务

2. 程序出哪些问题会导致数据不一致?

3. CAP和BASE理论

4. 两\_三阶段提交

5. tcc分布式事务

6. 基于本地消息表的最终一致性

7. 基于可靠消息的最终一致性- 最常用

8. 最大努力通知

9. mq (message queue) 的

bobby · 更新于 2022-11-08

◀ 上一节 7. 负载均衡策略 9. grpc的负载均衡... 下一节 ▶

在分布式系统中，多台服务器同时提供一个服务，并统一到服务配置中心进行管理，消费者通过查询服务配置中心，获取到服务到地址列表，需要选取其中一台来发起RPC远程调用。如何选择，则取决于具体的负载均衡算法，对应于不同的场景，选择的负载均衡算法也不尽相同。负载均衡算法的种类有很多种，常见的负载均衡算法包括轮询法、随机法、源地址哈希法、加权轮询法、加权随机法、最小连接法等，应根据具体的使用场景选取对应的算法。

1. 轮询 (Round Robin) 法

轮询很容易实现，将请求按顺序轮流分配到后台服务器上，均衡的对待每一台服务器，而不关心服务器实际的连接数和当前的系统负载。

2. 随机法

通过系统随机函数，根据后台服务器列表的大小值来随机选取其中一台进行访问。由概率统计理论可以得知，随着调用量的增大，其实际效果越来越接近于平均分配流量到后台的每一台服务器，也就是轮询法的效果。

3. 源地址哈希法

源地址哈希法的思想是根据服务消费者请求客户端的IP地址，通过哈希函数计算得到一个哈希值，将此哈希值和服务器列表的大小进行取模运算，得到的结果便是要访问的服务器地址的序号。采用源地址哈希法进行负载均衡，相同的IP客户端，如果服务器列表不变，将映射到同一个后台服务器进行访问。

4. 加权轮询 (Weight Round Robin) 法

不同的后台服务器可能机器的配置和当前系统的负载并不相同，因此它们的抗压能力也不一样。跟配置高、负载低的机器分配更高的权重，使其能处理更多的请求，而配置低、负载高的机器，则给其分配较低的权重，降低其系统负载，加权轮询很好的处理了这一问题，并将请求按照顺序且根据权重分配给后端。

5. 加权随机 (Weight Random) 法

加权随机法跟加权轮询法类似，根据后台服务器不同的配置和负载情况，配置不同的权重。不同的是，它是按照权重来随机选取服务器的，而非顺序。

6. 最小连接数法

前面我们费尽心思来实现服务消费者请求次数分配的均衡，我们知道这样做是没错的，可以为后端的多台服务器平均分配工作量，最大程度地提高服务器的利用率，但是，实际上，请求次数的均衡并不代表负载的均衡。因此我们需要介绍最小连接数法，最小连接数法比较灵活和智能，由于后台服务器的配置不尽相同，对请求的处理有快有慢，它正是根据后端服务器当前的连接情况，动态的选取其中当前积压连接数最小的那个服务器来处理当前请求，尽可能的提高后台服务器的利用率，将负载合理的分配到每一台服务器。

✎ 意见反馈

📖 收藏教程

🔖 标记书签

 我要提出意见反馈

