

负载均衡SLB

负载均衡（Server Load Balancer）是将访问流量根据转发策略分发到后端多台服务器的流量分发控制服务。负载均衡扩展了应用的服务能力，增强了应用的可用性。

分类

- 四层(传输层)负载均衡:对客户端TCP/IP协议的包转发
- 七层(应用层)负载均衡:Http的应用层负载均衡,Nginx就是一个典型的7层负载均衡SLB

- 中小型的Web应用，比如日PV小于1000万，用Nginx
- DNS轮询
- 大型网站或重要的服务，且服务器比较多时，可以考虑用LVS F5

Nginx的负载均衡算法

- 轮询 每个请求按时间顺序逐一分配到不同的后端服务器;
- ip_hash 每个请求按访问IP的hash结果分配，同一个IP客户端固定访问一个后端服务器。可以保证来自同一ip的请求被打到固定的机器上，可以解决session问题。
- url_hash 按访问url的hash结果来分配请求，使每个url定向到同一个后端服务器。后台服务器为缓存的时候效率。
- fair 这是比上面两个更加智能的负载均衡算法。此种算法可以依据页面大小和加载时间长短智能地进行负载均衡，也就是根据后端服务器的响应时间来分配请求，响应时间短的优先分配。Nginx本身是不支持 fair 的，如果需要使用这种调度算法，必须下载Nginx的 upstream_fair 模块。

<https://www.cnblogs.com/zhaoyanjuan/p/9139390.html>