**Nginx的产生**

没有听过Nginx？那么一定听过它的“同行”Apache吧！Nginx同Apache一样都是一种WEB服务器。基于REST架构风格，以统一资源描述符(Uniform Resources Identifier)URI或者统一资源定位符(Uniform Resources Locator)URL作为沟通依据，通过HTTP协议提供各种网络服务。

然而，这些服务器在设计之初受到当时环境的局限，例如当时的用户规模，网络带宽，产品特点等局限并且各自的定位和发展都不尽相同。这也使得各个WEB服务器有着各自鲜明的特点。

Apache的发展时期很长，而且是毫无争议的世界第一大服务器。它有着很多有点：稳定、开源、跨平台等等。但是由于它出现的时间太长了。它兴起的年代，互联网产业远比不上现在。所以它被设计为一个重量级的。不支持高并发的服务器。在Apache上运行数以万计的并发访问，会导致服务器消耗大量内存。操作系统对其进行进程或线程间的切换也消耗了大量的CPU资源，导致HTTP请求的平均响应速度降低。

这些都决定了Apache不可能成为高性能WEB服务器，轻量级高并发服务器Nginx就应运而生了。

俄罗斯的工程师Igor Sysoev，他在为Rambler Media工作期间，使用C语言开发了Nginx。Nginx作为WEB服务器一直为Rambler Media提供出色而又稳定的服务。

然后呢，Igor Sysoev将Nginx代码开源，并且赋予自由软件许可证。

由于：

* Nginx使用基于事件驱动架构，使得其可以支持数以百万级别的TCP连接
* 高度的模块化和自由软件许可证是的第三方模块层出不穷（这是个开源的时代啊~）
* Nginx是一个跨平台服务器，可以运行在Linux, FreeBSD, Solaris, AIX, Mac OS, Windows等操作系统上
* 这些优秀的设计带来的极大的稳定性。

所以，Nginx火了！

**Nginx的用武之地**

Nginx是一款自由的、开源的、高性能的HTTP服务器和反向代理服务器；同时也是一个IMAP、POP3、SMTP代理服务器；Nginx可以作为一个HTTP服务器进行网站的发布处理，另外nginx可以作为反向代理进行负载均衡的实现。

**关于代理**

说到代理，首先我们要明确一个概念，所谓代理就是一个代表、一个渠道；

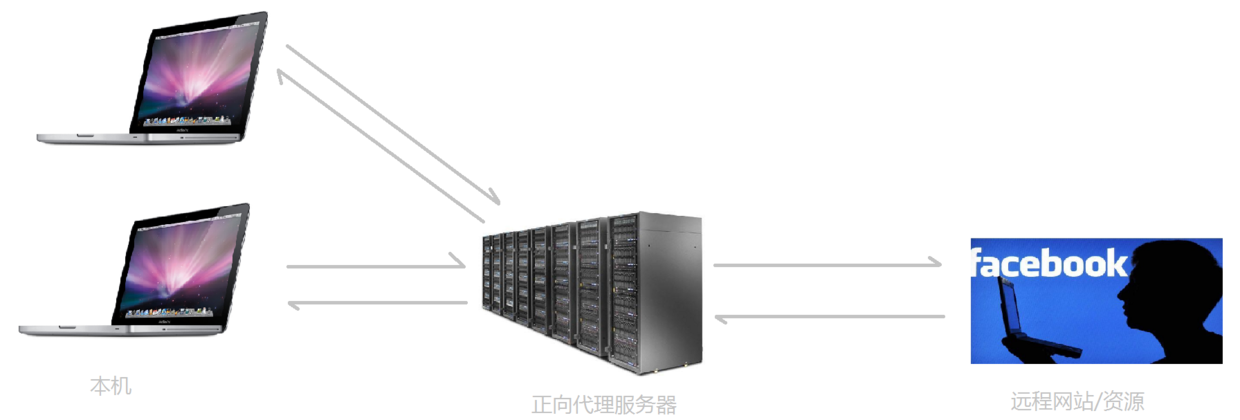
此时就设计到两个角色，一个是被代理角色，一个是目标角色，被代理角色通过这个代理访问目标角色完成一些任务的过程称为代理操作过程；如同生活中的专卖店~客人到adidas专卖店买了一双鞋，这个专卖店就是代理，被代理角色就是adidas厂家，目标角色就是用户。

**正向代理**

说反向代理之前，我们先看看正向代理，正向代理也是大家最常接触的到的代理模式，我们会从两个方面来说关于正向代理的处理模式，分别从软件方面和生活方面来解释一下什么叫正向代理

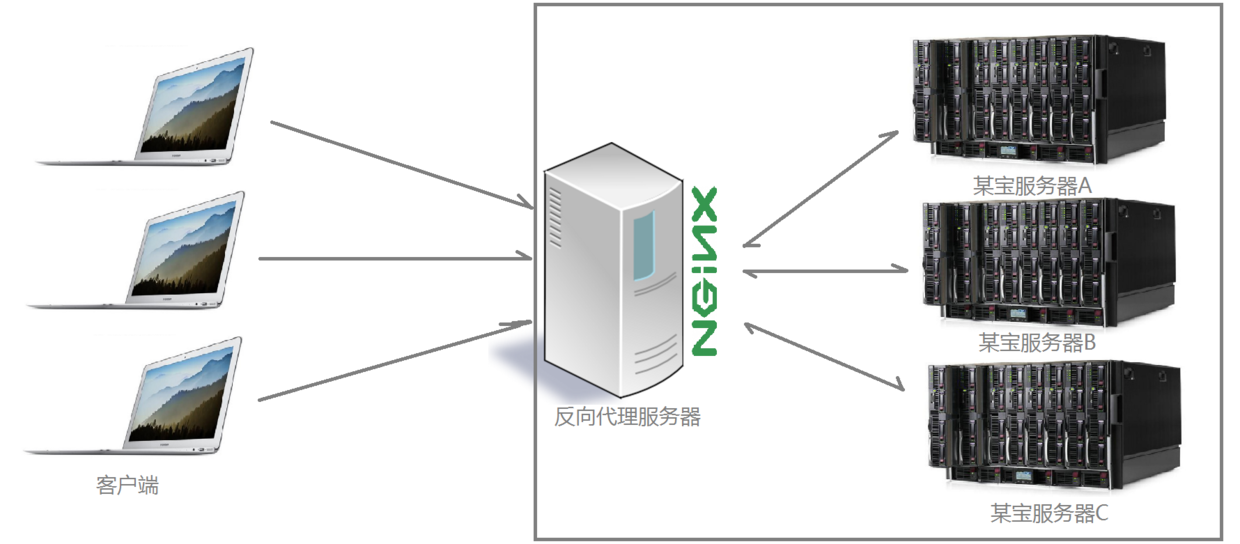
在如今的网络环境下，我们如果由于技术需要要去访问国外的某些网站，此时你会发现位于国外的某网站我们通过浏览器是没有办法访问的，此时大家可能都会用一个操作FQ进行访问，FQ的方式主要是找到一个可以访问国外网站的代理服务器，我们将请求发送给代理服务器，代理服务器去访问国外的网站，然后将访问到的数据传递给我们！

上述这样的代理模式称为正向代理，正向代理最大的特点是客户端非常明确要访问的服务器地址；服务器只清楚请求来自哪个代理服务器，而不清楚来自哪个具体的客户端；正向代理模式屏蔽或者隐藏了真实客户端信息。



**反向代理**

明白了什么是正向代理，我们继续看关于反向代理的处理方式，举例如我大天朝的某宝网站，每天同时连接到网站的访问人数已经爆表，单个服务器远远不能满足人民日益增长的购买欲望了，此时就出现了一个大家耳熟能详的名词：分布式部署；也就是通过部署多台服务器来解决访问人数限制的问题；某宝网站中大部分功能也是直接使用nginx进行反向代理实现的，并且通过封装nginx和其他的组件之后起了个高大上的名字：Tengine，有兴趣的童鞋可以访问Tengine的官网查看具体的信息：http://tengine.taobao.org/。那么反向代理具体是通过什么样的方式实现的分布式的集群操作呢，我们先看一个示意图：

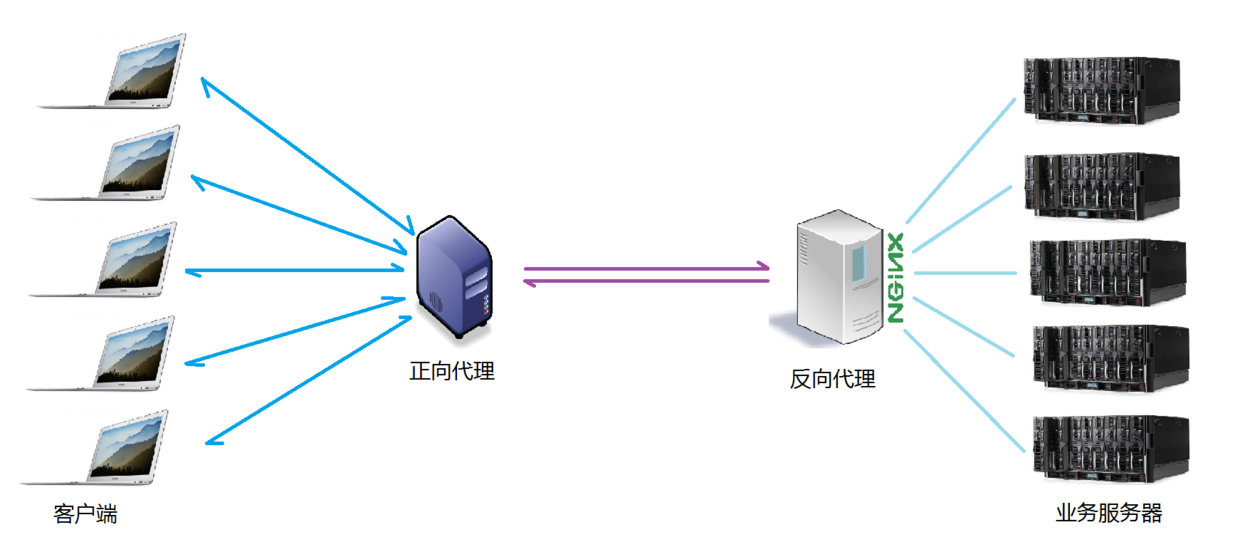


通过上述的图解大家就可以看清楚了，多个客户端给服务器发送的请求，Nginx服务器接收到之后，按照一定的规则分发给了后端的业务处理服务器进行处理了。此时~请求的来源也就是客户端是明确的，但是请求具体由哪台服务器处理的并不明确了，Nginx扮演的就是一个反向代理角色

反向代理，主要用于服务器集群分布式部署的情况下，反向代理隐藏了服务器的信息！

**项目场景**

通常情况下，我们在实际项目操作时，正向代理和反向代理很有可能会存在在一个应用场景中，正向代理代理客户端的请求去访问目标服务器，目标服务器是一个反向单利服务器，反向代理了多台真实的业务处理服务器。具体的拓扑图如下：



**负载均衡**

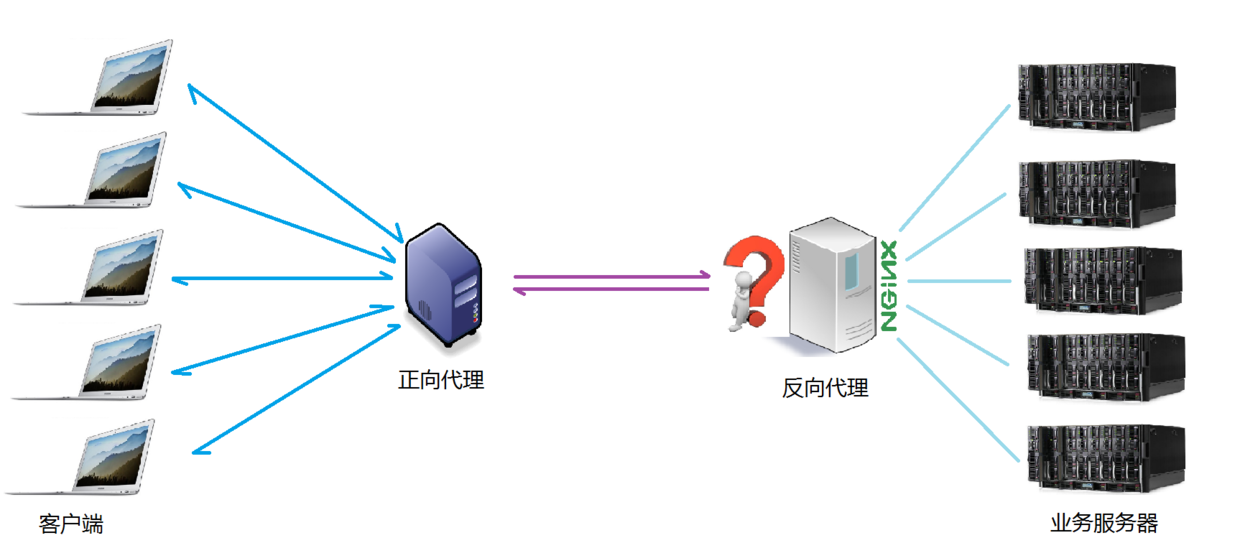
我们已经明确了所谓代理服务器的概念，那么接下来，Nginx扮演了反向代理服务器的角色，它是以依据什么样的规则进行请求分发的呢？不用的项目应用场景，分发的规则是否可以控制呢？

这里提到的客户端发送的、Nginx反向代理服务器接收到的请求数量，就是我们说的负载量。

请求数量按照一定的规则进行分发到不同的服务器处理的规则，就是一种均衡规则。

所以~将服务器接收到的请求按照规则分发的过程，称为负载均衡。

负载均衡在实际项目操作过程中，有硬件负载均衡和软件负载均衡两种，硬件负载均衡也称为硬负载，如F5负载均衡，相对造价昂贵成本较高，但是数据的稳定性安全性等等有非常好的保障，如中国移动中国联通这样的公司才会选择硬负载进行操作；更多的公司考虑到成本原因，会选择使用软件负载均衡，软件负载均衡是利用现有的技术结合主机硬件实现的一种消息队列分发机制。



Nginx支持的负载均衡调度算法方式如下：

1. weight轮询（默认）：接收到的请求按照顺序逐一分配到不同的后端服务器，即使在使用过程中，某一台后端服务器宕机，Nginx会自动将该服务器剔除出队列，请求受理情况不会受到任何影响。 这种方式下，可以给不同的后端服务器设置一个权重值（weight），用于调整不同的服务器上请求的分配率；权重数据越大，被分配到请求的几率越大；该权重值，主要是针对实际工作环境中不同的后端服务器硬件配置进行调整的。
2. ip\_hash：每个请求按照发起客户端的ip的hash结果进行匹配，这样的算法下一个固定ip地址的客户端总会访问到同一个后端服务器，这也在一定程度上解决了集群部署环境下session共享的问题。
3. fair：智能调整调度算法，动态的根据后端服务器的请求处理到响应的时间进行均衡分配，响应时间短处理效率高的服务器分配到请求的概率高，响应时间长处理效率低的服务器分配到的请求少；结合了前两者的优点的一种调度算法。但是需要注意的是Nginx默认不支持fair算法，如果要使用这种调度算法，请安装upstream\_fair模块。
4. url\_hash：按照访问的url的hash结果分配请求，每个请求的url会指向后端固定的某个服务器，可以在nginx作为静态服务器的情况下提高缓存效率。同样要注意Nginx默认不支持这种调度算法，要使用的话需要安装nginx的hash软件包。

**几种常用web服务器对比**

