深入浅出 DDoS 攻击防御应对篇：DDoS 防御方案(1)

谈到 DDoS 防御，首先就是要知道到底遭受了多大的攻击。这个问题看似简单，实际上却有很多不为人知的细节在里面。

防御基础

攻击流量到底多大

谈到 DDoS 防御，首先就是要知道到底遭受了多大的攻击。这个问题看似简单，实际上却有很多不为人知的细节在里面。

以 SYN Flood 为例，为了提高发送效率在服务端产生更多的 SYN 等待队列，攻击程序在填充包头时，IP 首部和 TCP 首部都不填充可选的字段，因此 IP 首部长度恰好是 20 字节，TCP 首部也是 20 字节，共 40 字节。

对于以太网来说，最小的包长度数据段必须达到 46 字节，而攻击报文只有 40 字节，因此，网卡在发送时，会做一些处理，在 TCP 首部的末尾，填充 6 个 0 来满足最小包的长度要求。这个时候，整个数据包的长度为 14 字节的以太网头，20 字节的 IP 头，20 字节的 TCP 头，再加上因为最小包长度要求而填充的 6 个字节的 0，一共是 60 字节。

但这还没有结束。以太网在传输数据时，还有 CRC 检验的要求。网卡会在发送数据之前对数据包进行 CRC 检验，将 4 字节的 CRC 值附加到包头的最后面。这个时候，数据包长度已不再是 40 字节，而是变成 64 字节了，这就是常说的 SYN 小包攻击，数据包结构如下：

|14 字节以太网头部|20 字节 IP 头部|20 字节 TCP|6 字节填充|4 字节检验|

|目的 MAC|源 MAC|协议类型| IP 头 |TCP 头|以太网填充 | CRC 检验 |

到 64 字节时，SYN 数据包已经填充完成，准备开始传输了。攻击数据包很小，远远不够最大传输单元（MTU）的 1500 字节，因此不会被分片。那么这些数据包就像生产流水线上的罐头一样，一个包连着一个包紧密地挤在一起传输吗？事实上不是这样的。

以太网在传输时，还有前导码（preamble）和帧间距（inter-frame gap）。其中前导码占 8 字节（byte），即 64 比特位。前导码前面的 7 字节都是 10101010，1 和 0 间隔而成。但第八个字节就变成了 10101011，当主机监测到连续的两个 1 时，就知道后面开始是数据了。在网络传输时，数据的结构如下：

|8 字节前导码|6 字节目的 MAC 地址|6 字节源 MAC 地址|2 字节上层协议类型|20 字节 IP 头|20 字节 TCP 头|6 字节以太网填充|4 字节 CRC 检验|12 字节帧间距|

也就是说，一个本来只有 40 字节的 SYN 包，在网络上传输时占的带宽，其实是 84 字节。

有了上面的基础，现在可以开始计算攻击流量和网络设备的线速问题了。当只填充 IP 头和 TCP 头的最小 SYN 包跑在以太网络上时，100Mbit 的网络，能支持的最大 PPS（Packet Per Second）是 100×106 / (8 * (64+8+12)) = 148809，1000Mbit 的网络，能支持的最大 PPS 是 1488090。

SYN Flood 防御

前文描述过，SYN Flood 攻击大量消耗服务器的 CPU、内存资源，并占满 SYN 等待队列。相应的，我们修改内核参数即可有效缓解。主要参数如下：

net.ipv4.tcp_syncookies = 1

net.ipv4.tcp_max_syn_backlog = 8192

net.ipv4.tcp_synack_retries = 2

分别为启用 SYN Cookie、设置 SYN 最大队列长度以及设置 SYN+ACK 最大重试次数。

SYN Cookie 的作用是缓解服务器资源压力。启用之前，服务器在接到 SYN 数据包后，立即分配存储空间，并随机化一个数字作为 SYN 号

发送 SYN+ACK 数据包。然后保存连接的状态信息等待客户端确认。启用 SYN Cookie 之后，服务器不再分配存储空间，而且通过基于时间种子的随机数算法设置一个 SYN 号，替代完全随机的 SYN 号。发送完 SYN+ACK 确认报文之后，清空资源不保存任何状态信息。直到服务器接到客户端的最终 ACK 包，通过 Cookie 检验算法鉴定是否与发出去的 SYN+ACK 报文序列号匹配，匹配则通过完成握手，失败则丢弃。当然，前文的高级攻击中有 SYN 混合 ACK 的攻击方法，则是对此种防御方法的反击，其中优劣由双方的硬件配置决定

tcp_max_syn_backlog 则是使用服务器的内存资源，换取更大的等待队列长度，让攻击数据包不至于占满所有连接而导致正常用户无法完成握手。net.ipv4.tcp_synack_retries 是降低服务器 SYN+ACK 报文重试次数，尽快释放等待资源。这三种措施与攻击的三种危害一一对应，完完全全地对症下药。但这些措施也是双刃剑，可能消耗服务器更多的内存资源，甚至影响正常用户建立 TCP 连接，需要评估服务器硬件资源和攻击大小谨慎设置。

除了定制 TCP/IP 协议栈之外，还有一种常见做法是 TCP 首包丢弃方案，利用 TCP 协议的重传机制识别正常用户和攻击报文。当防御设备接到一个 IP 地址的 SYN 报文后，简单比对该 IP 是否存在于白名单中，存在则转发到后端。如不存在于白名单中，检查是否是该 IP 在一定时间段内的首次 SYN 报文，不是则检查是否重传报文，是重传则转发并加入白名单，不是则丢弃并加入黑名单。是首次 SYN 报文则丢弃并

等待一段时间以试图接受该 IP 的 SYN 重传报文，等待超时则判定为攻击报文加入黑名单。

首包丢弃方案对用户体验会略有影响，因为丢弃首包重传会增大业务的响应时间，有鉴于此发展出了一种更优的 TCP Proxy 方案。所有的 SYN 数据报文由清洗设备接受，按照 SYN Cookie 方案处理。和设备成功建立了 TCP 三次握手的 IP 地址被判定为合法用户加入白名单，由设备伪装真实客户端 IP 地址再与真实服务器完成三次握手，随后转发数据。而指定时间内没有和设备完成三次握手的 IP 地址，被判定为恶意 IP 地址屏蔽一定时间。除了 SYN Cookie 结合 TCP Proxy 外，清洗设备还具备多种畸形 TCP 标志位数据包探测的能力，通过对 SYN 报文返回非预期应答测试客户端反应的方式来鉴别正常访问和恶意行为。

清洗设备的硬件具有特殊的网络处理器芯片和特别优化的操作系统、TCP/IP 协议栈，可以处理非常巨大的流量和 SYN 队列。

HTTP Flood 防御

HTTP Flood 攻击防御主要通过缓存的方式进行，尽量由设备的缓存直接返回结果来保护后端业务。大型的互联网企业，会有庞大的 CDN 节点缓存内容。

当高级攻击者穿透缓存时，清洗设备会截获 HTTP 请求做特殊处理。最简单的方法就是对源 IP 的 HTTP 请求频率做统计，高于一定频率的

IP 地址加入黑名单。这种方法过于简单，容易带来误杀，并且无法屏蔽来自代理服务器的攻击，因此逐渐废止，取而代之的是 JavaScript 跳转人机识别方案。

HTTP Flood 是由程序模拟 HTTP 请求，一般来说不会解析服务端返回数据，更不会解析 JS 之类代码。因此当清洗设备截获到 HTTP 请求时，返回一段特殊 JavaScript 代码，正常用户的浏览器会处理并正常跳转不影响使用，而攻击程序会攻击到空处。

DNS Flood 防御

DNS 攻击防御也有类似 HTTP 的防御手段，第一方案是缓存。其次是重发，可以是直接丢弃 DNS 报文导致 UDP 层面的请求重发，可以是返回特殊响应强制要求客户端使用 TCP 协议重发 DNS 查询请求。

特殊的，对于授权域 DNS 的保护，设备会在业务正常时期提取收到的 DNS 域名列表和 ISP DNS IP 列表备用，在攻击时，非此列表的请求一律丢弃，大幅降低性能压力。对于域名，实行同样的域名白名单机制，非白名单中的域名解析请求，做丢弃处理。

慢速连接攻击防御

Slowloris 攻击防御比较简单，主要方案有两个。

第一个是统计每个 TCP 连接的时长并计算单位时间内通过的报文数量即可做精确识别。一个 TCP 连接中，HTTP 报文太少和报文太多都

是不正常的，过少可能是慢速连接攻击，过多可能是使用 HTTP 1.1 协议进行的 HTTP Flood 攻击，在一个 TCP 连接中发送多个 HTTP 请求。

第二个是限制 HTTP 头部传输的最大许可时间。超过指定时间 HTTP Header 还没有传输完成，直接判定源 IP 地址为慢速连接攻击，中断连接并加入黑名单。

企业级防御

互联网企业防御 DDoS 攻击，主要使用上文的基础防御手段，重点在于监控、组织以及流程。

监控需要具备多层监控、纵深防御的概念，从骨干网络、IDC 入口网络的 BPS、PPS、协议分布，负载均衡层的 VIP 新建连接数、并发连接数、BPS、PPS 到主机层的 CPU 状态、TCP 新建连接数状态、TCP 并发连接数状态，到业务层的业务处理量、业务连通性等多个点部署监控系统。即使一个监控点失效，其他监控点也能够及时给出报警信息。多个点信息结合，准确判断被攻击目标和攻击手法。

一旦发现异常，立即启动在虚拟防御组织中的应急流程，防御组织需要囊括到足够全面的人员，至少包含监控部门、运维部门、网络部门、安全部门、客服部门、业务部门等，所有人员都需要 2-3 个备份。流程启动后，除了人工处理，还应该包含一定的自动处理、半自动处理能力。例如自动化的攻击分析，确定攻击类型，自动化、半自动化的

防御策略，在安全人员到位之前，最先发现攻击的部门可以做一些缓解措施。

除了 DDoS 到来之时的流程等工作之外，更多的工作是在攻击到来之前。主要包含 CDN 节点部署、DNS 设置、流程演习等。对于企业来说，具备多个 CDN 节点是 DDoS 防御容量的关键指标。当一个机房承担不住海量数据时，可以通过 DNS 轮询的方式，把流量引导到多个分布节点，使用防御设备分头处理。因此 DNS 的 TTL 值需要设置得足够小，能够快速切换，每个 CDN 节点的各种 VIP 设置也需要准备充分。

在虚拟化时代，各种用户的不同业务共处在相同的物理机平台，遭受 DDoS 攻击的可能性越来越高，而且一个用户被攻击可能牵扯到大量的其他用户，危害被显著放大，因此防御显得尤为重要。阿里云的虚拟化业务，平均每天遭受约 20 起 DDoS 攻击，最大流量达到接近 20Gbit/s，所有这些攻击都在 15 分钟内自动处理完成，让客户远离 DDoS 的威胁，专心发展业务。

总地来说，对 DDoS 防御，主要的工作是幕后积累。台上十分钟，台下十年功，没有充分的资源准备，没有足够的应急演练，没有丰富的处理经验，DDoS 攻击将是所有人的噩梦。