doi:10.3772/j.issn.1002-0470.2014.10.007

基于 KVM 虚拟化的网络环境自动构建技术研究①

张 云②* 唐积强** 闫健恩③* 张兆心*

(*哈尔滨工业大学计算机科学与技术学院 哈尔滨 150001)
(**国家计算机网络应急技术处理协调中心 北京 100029)

摘 要 针对传统网络环境构建存在成本高、效率低、可控性差的问题,进行了虚拟网络环境自动化构建技术研究,提出了一种基于 KVM 虚拟化的网络环境自动构建新技术:首先形式化描述真实网络,根据描述结果利用基于内核的虚拟机(KVM)和 Click 虚拟化生成主机、路由器,然后采用终端路由器路由表洪泛(TRTF)算法计算路由表信息,通过脚本技术对主机与路由器自动设置,形成最终虚拟网络。实验结果表明,利用本文构建算法,构建规模 100 个节点的网络环境仅需 95s。在构建网络中进行了 DDoS 模拟实验,结果表明网络连通性状况良好,构建网络可用于网络安全事件重现。

关键词 网络环境,自动构建,形式化描述语言,KVM 虚拟化

0 引 言

当今网络安全事件频发,给社会带来了极大的危害。为了更好地研究与分析网络安全事件以对其进行有效的防御,需要对其进行重现。由于网络具有不可控性、易变性和不可预测性,同时在对其测评过程中使用的方法、手段和工具可能会影响网络的正常使用甚至引入安全隐患,因此需要依据实际的网络环境,构建一个具备高仿真度的虚拟网络环境,用于网络安全事件的重现。如何构建相应的网络环境是目前网络技术人员所面临的一项重要课题。虽然构建试验床(testbed)可以部分解决此类问题,但是试验床的造价高昂,且对大规模网络试验的支持较差。这方面的工作还包括 Mike 等[1]提出的基于 Emulab 测试平台的虚拟化实验环境构建方法和 Brent[2]提出的基于 PlanetLab 测试平台的虚拟化实验环境构建方法,但这两种方法存在对主机的部署效率相对较低的缺点。中国科学院的王佳宾等[3]提出了一种基于分组自适应的网络环境快速构建方法,但该方法只涉及到对于主机的分组部署,并没有解决自动构建连通性网络的问题。全晓莉等[4]及冯陈伟[5]提出了利用虚拟化技术构建网络试验平台的方法,但其中的 VMWare 虚拟机技术与操作系统虚拟化技术在网络环境构建过程中需要手动操作,对于大规模的网络环境构建也存在效率低下的不足。为了解决以上网络环境构建方法存在的成本高、效率低、可控性差的问题,本文利用基于内核的虚拟机(kernel-based virtual machine, KVM)[6,7]自动虚拟化产生网络中的主机,结合 Click 软件[8,9]构建网络中的路由器,提出了一种基于 KVM 及 Click 软件路由器的网络环境自动构建方案,该方案可根据实际的网络快速自动构建虚拟的高仿真度网络环境。

1 网络元素

在根据实际网络构建用于网络安全事件重现的虚拟网络环境时,需要首先对实际网络中的各个网络元素例如主机、路由器等进行详细的分析与描述,为了便于说明,给出以下定义与描述。

1.1 符号定义

定义 1 真实网络(R-NET):爆发网络安全事件的实际网络,如图 1 所示。

① 国家科技支撑计划(2012BAH45B01),国家自然科学基金(61100189, 61370215, 61370211)和山东省中青年科学家奖励基金(BS2011DX001)资助项目。
② 男,1989 年生,硕士生;研究方向:网络安全;E-mail: hit_zy@126.com
③ 通讯作者,E-mail: yje@hitwh.edu.cn
(收稿日期:2014-05-08)

定义 2　虚拟网络(V-NET):根据 R-NET 利用本文算法构建的虚拟网络环境。

定义 3　H 节点:网络环境中的主机元素。

定义 4　R1 节点:网络中的终端路由器元素,即直接与主机相连的路由器。

定义 5　R2 节点:网络中除终端路由器之外的路由器元素。

定义 6　L1 链路:网络中主机与路由器之间的链路。

定义 7　L2 链路:网络中路由器与路由器之间的链路。

定义 8　网络端口 P:代表主机、路由器中的网卡接口,具有 IP 地址信息,同时与 L1、L2 链路的端点对应。

定义 9　HA(Host Arrary,主机数组):$HA = \{H \mid H \in R - NET\}$。

定义 10　RA(Router Arrary,路由器数组):$RA = \{R \mid R \in R - NET \&\& (R = R1 \parallel R = R2)\}$。

定义 11　LA(Link Arrary,链路关系数组):$LA = \{L \mid L \in R - NET \&\& (L = L1 \parallel L = L2)\}$。

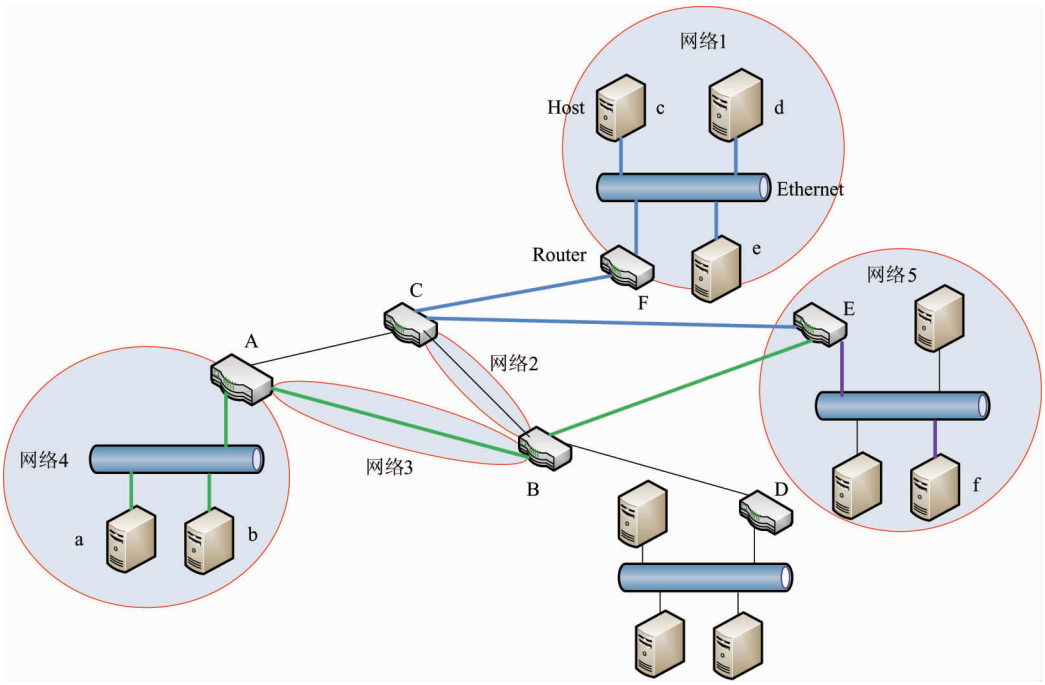

图 1　真实网络 R-NET

由以上定义可知,R-NET 可描述为{HA,RA,LA},即主机,路由器以及链路关系即连接关系的集合。

1.2　元素描述

为了构建 R-NET 对应的虚拟网络环境 V-NET,需要对网络中的主机、路由器、链路关系等元素进行详细的形式化描述。形式化描述如下。

(1)主机

主机元素具有操作系统、处理器、内存等自身特性,而处于网络中的主机同时具备用于与其他网络元素进行沟通的网卡信息,因此,对主机的形式化描述为五元组 H = (H _ De,OS,CPU,MEM,P),其中 H _ De 为主机描述符,唯一表示网络中特定主机,OS 为主机操作系统,CPU 参数指定 CPU 核心数,MEM 指定内存大小,网络端口 P 定义了主机的 IP 地址等信息。

(2)路由器

路由器在具有处理器、内存等自身特性之外,同时具备多个网络端口、用于路由数据包的路由表等属性,因此,对路由器的形式化描述为五元组 R = (R _ De,CPU,MEM,PL,RT),其中 R _ De 为路由器描述符,CPU、MEM 的意义与主机描述相同,PL = (P_1,P_2,P_3,…)为网络端口 P 的列表,分别代表路由器的多个网络端口,RT 定义为路由器中的路由表。

(3)链路关系

在本文中链路关系描述为 L = (L _ De,P_1,P_2),其中 P_1、P_2 分别代表链路两端的网络接口,和与其相连的主机或路由器的形式化描述中的网络端口信息一一对应。当 L 为 L1 时,L. P_1 = H. P,L. P_2 = R1.

P_x,R1.P_x 表示终端路由器 R1 的第 x 个网络接口。

2　基于 KVM 的网络环境自动构建技术

传统网络环境构建是对 R-NET 中的各个元素依次构建,然后进行手动配置最终得到 V-NET。而本研究则在构建过程中,通过对 R-NET 进行解析并形式化描述和基于 KVM 虚拟化的分组构建与自动配置三个过程实现了构建过程的自动化,从而提高了构建效率。

2.1　R-NET 解析

解析是将给定的 R-NET 解析为 1.2 节形式化描述的主机、路由器、链路关系。通过解析,得到通用的 R-NET 的表示形式,为之后主机与路由器元素的虚拟化生成提供输入。在实际构建过程中,R-NET 通常并未包含网络地址信息,因此在解析时,涉及到网络端口 IP 地址分配的问题,依据现实网络环境中网络地址分配规则,需要保证同一局域网网络端口分配属于同一网段的地址,不同局域网网络端口分配不同网段的地址,并保证同一网段地址不被重复分配。本研究在进行网络端口 IP 地址分配时,首先根据链路关系进行划分网络,划分依据如图 1所示,划分结果为多个网络,每个网络包含至少 1 条链路。然后对各个网络分配不同网络地址,对同一网络的网络端口分配该网络地址中的主机地址。网络地址的生成采用选定初始网络地址值并随机步长增长的方式,主机地址则按顺序增长的方式生成。基于网络划分的形式化解析(net partition based formalized analysis,NPFA)算法描述如下:

Input

真实网络 R-NET

Output

形式化描述的主机数组 HA,路由器数组 RA,链路关系数组 LA。

BEGIN

初始化 *HA*、*RA*、*LA* 为空数组
nets←对 R-NET 进行网络划分
for *net in nets* **do**
　net _ ip←生成网络地址
　for 链路 *l in nets* **do**
　　l. P_1←*net _ ip* + +
　　l. P_2←*net _ ip* + +
　　LA. *append*(*l*)
　end for
end for
for 路由器元素 *r in R-NET* **do**
　for 链路 *l in* 与 *r* 相连链路 **do**
　　r. *PL*. *append*(*l*. *P*←*r* 与 *l* 共有的网络端口)
　end for
　RA. *append*(*r*)
end for
for 主机元素 *h in R-NET* **do**
　*l*1←与 *h* 相连的链路
　h. *P*←*l*1. P_1
　HA. *append*(*h*)
end for
END

经过解析,得到形式化描述的主机数组 HA,路由器数组 RA。数组中分别存储每个主机、路由器的网络地址等详细配置信息。

2.2　基于 KVM 的虚拟化技术

在得到真实网络中主机与路由器的形式化描述后,需要根据描述信息生成虚拟网络中的主机与路由器元素。为了降低构建成本,本文构建的虚拟网络环境中主机与路由器元素均以基于内核的虚拟机(KVM)形式出现。

2.2.1　KVM 虚拟化

KVM 是一种用于 Linux 内核中的虚拟化基础设施。利用 KVM 虚拟化技术,可以在一台宿主机上虚拟化得到多台虚拟机,在需要大量主机资源的应用场景中有广泛应用。利用 KVM 进行主机生成主要有直接生成与主机克隆两种方式。直接生成是指根据需要,利用 KVM 进行包括虚拟硬盘生成、主机资源配置、系统安装等在内的所有操作进行虚拟主机的生成,在虚拟机生成时有大量的虚拟硬盘生成、系统安装、读盘等操作,存在严重的时间浪费,生成效率相对较低。主机克隆则是根据已有的模版机克隆生成新的主机,而在克隆过程中,对体积较大的虚拟硬盘的克隆同样存在花费时间长的缺点。

2.2.2　KVM 外置磁盘快照技术

一个虚拟机快照可被看作是虚拟机的在某个指定时间的视图(包括他的操作系统和所有的程序),外置磁盘快照则是对虚拟机磁盘文件的快照,体积很小,并且可以根据它迅速的创建瘦装备虚拟机的实例,磁盘快照可以链式创建,如图 2所示。

图中 Backing file 作为 Overlay(覆盖)-1 的后端文件,Overlay-1同时作为Overlay-1-1、Overlay-1-2、

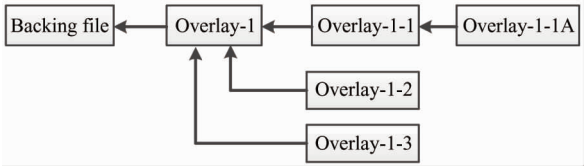

图 2 磁盘快照链

Overlay-1-3 的后端文件,每个快照文件代表了虚拟机的不同时间的视图,利用快照文件和对应的后端文件可进行虚拟机的快速生成。与主机克隆不同之处是,主机克隆需要对模版机体积较大的磁盘文件进行整体复制,而利用外置磁盘快照技术,可以首先将模版机磁盘文件作为后端文件,生成模版机对应的的体积较小的外置磁盘快照。在进行主机克隆生成时,仅需复制多份体积较小的磁盘快照文件,各个磁盘快照共用模版机磁盘文件为后端文件,即可快速生成多个虚拟机,由于克隆时间减少,主机生成速率得到提高。

2.2.3 分组快照虚拟机快速生成技术

网络中存在大量异构元素,例如主机与路由器以及主机之间运行的操作系统也可能各不相同,因此本文提出对主机进行分组的思想,把具有公共属性的主机分为一组,对每一组主机生成一个模版机,对模版机生成磁盘快照,在进行虚拟机生成时,根据解析结果,选择对应的模版机快照即可快速生成虚拟机。

同时,本文利用服务器客户端模式通过由网络环境构建管理程序对多个宿主机进行管理,将虚拟机生成任务平均分配到作为客户端的各个宿主机中,由各个宿主机共同完成虚拟机生成任务,提高了任务的并发性,使得构建效率进一步提高。由于提供了多宿主机的支持,本文网络环境构建技术对规模较大的网络的构建也有了更好的支持。快速虚拟机生成过程如下:

(1) 对主机进行分组并创建模版机,同时生成磁盘快照文件。

(2) 根据解析结果,将虚拟机创建任务平均分配到每个宿主机。

(3) 宿主机对每一个需要生成的虚拟机选择对应的磁盘快照文件。

(4) 宿主机通过克隆磁盘快照文件快速生成虚拟机。

2.3 虚拟路由器配置

经过虚拟化生成得到网络中的主机与路由器,需要对其进行必要的配置,这样才能使得构成的网络能够同真实网络一样正常工作。对生成的主机仅需配置其 IP 地址即可,而本文构建的虚拟网络环境中的路由器是通过在虚拟主机之上安装 Click 软件而得到的软件路由器,为了使软件路由器可以正常工作,需要生成 Click 软件安装时的配置文件,而配置文件的主要内容为路由表信息,因此需要根据待构建的真实网络进行路由表计算。

2.3.1 路由表计算

本文采用基于最短路径的静态路由。传统的基于最短路径的路由算法在计算路由时需要计算各个节点之间的最短路径,计算复杂度高,因而本文提出一种终端路由器路由表洪泛(terminal router routing table flooding,TRTF)传播算法,该算法通过模拟各个终端路由器将本身路由表洪泛方式发送到网络中的各个路由器,每个路由器在收到洪泛消息时将得到的路由表加入到自身路由表中,从而使得计算结束时每个路由器都正确保存了静态路由信息。这种路由表计算算法描述如下:

```
Input
    HA、RA、LA。
Output
    填充了路由表信息的 RA。
BEGIN
    g←根据 HA、RA、LA 生成拓扑图
    for r in RA do
        r. RT ← NULL
        if r is R1 do
            将到达 r 直接相连的主机的路由加入到
            r. RT 中
        end if
    end for
    for r in RA do
        if r is R1 do
            BFS(g,r) //在 BFS 过程中,对于遇到的
            新路由器节点 r_n,执行:
                r_n. RT + = r. RT
        end if
    end for
END
```

计算结束后,RA 中的每个路由器都存储了自身的路由表,根据路由表信息即可得到 Click 软件安装时所需的配置文件。

2.3.2 自动化配置

配置过程自动化的实现过程是:通过在路由器

模版机中放置随系统启动的脚本,在利用 KVM 虚拟化技术根据模版机生成虚拟机并成功启动后,脚本自动向宿主机构建管理程序发起请求,请求自身配置信息,然后宿主机构建管理程序,将请求者对应的配置信息例如 IP 地址信息以及 Click 软件配置文件发送给发起请求的虚拟机,由虚拟机自行进行 IP 地址信息配置、Click 软件的安装,最终得到虚拟网络环境中的主机与路由器。

2.4　自动构建过程

为了使网络环境构建自动化完成,本文提出的基于 KVM 的网络环境自动构建技术的构建过程如图 3 所示。

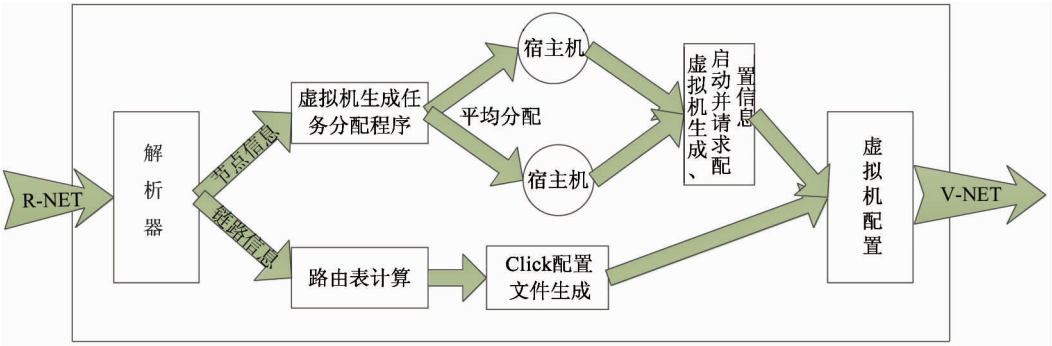

图 3　自动构建过程

解析器根据输入拓扑,将其解析为形式化的主机与路由器描述,根据解析结果,构建管理程序将虚拟机生成任务下发到各个宿主机,由宿主机进行虚拟机的自动生成,同时计算生成 Click 软件路由器所需的配置文件,虚拟机生成并启动成功后,自动向构建管理程序请求自身配置信息,根据得到的配置信息对自身进行配置之后,即可得到虚拟网络环境中的主机与路由器,整个构建过程完毕,输出构建的虚拟网络环境。

3　实　验

为了测试本文算法的可行性、构建效率与正确性,本研究共进行 3 组实验,测试单宿主机生成虚拟机个数,用以测试单台宿主机可生成的虚拟机个数上限与生成时间,实验结果作为构建方案可行性证明;测试不同规模网络在利用本文构建算法进行构建时的时间消耗,用以测试构建效率;在根据图 1 所示的真实网络 R-NET 构建的虚拟网络 V-NET 之上搭建僵尸网络,并模拟 DDoS 安全事件,通过对网络中数据包的统计,测试 V-NET 的连通性,用以对构建算法的正确性做出证明。

3.1　单宿主机生成虚拟机个数测试

本节测试单台宿主机生成虚拟机个数上限与生成时间、内存以及 CPU 占用率之间的关系。试验环境如表 1 所示。

表 1　实验环境配置

配置项目	宿主机	虚拟机
CPU	8 核心、16 线程	1 核心、1 线程
内存	64GB	1GB
操作系统	Ubuntu 13.10	Ubuntu 12.04

虚拟机创建个数与创建时间实验结果如图 4 所示。

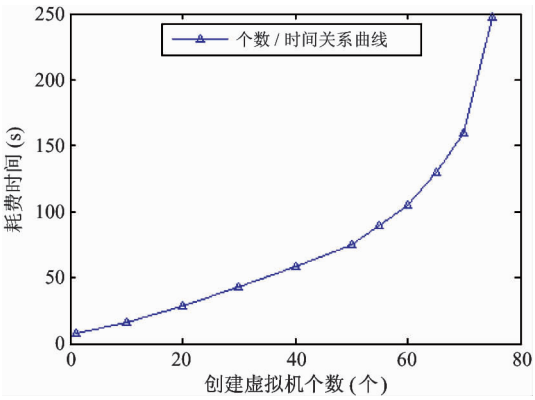

图 4　创建虚拟机个数与创建时间关系图

耗费时间指从开始创建虚拟机到虚拟机全部启动成功所需要的时间。花费时间随着创建的虚拟机个数的增长而变长。由于 KVM 支持内存过载,即虚拟机内存总和可以超过宿主机内存值,因此在实验中,可以创建资源总和比宿主机资源更多的虚拟机。

为了测试虚拟机在不同运行状况下宿主机可支

持的同时运行虚拟机个数上限，分别对虚拟机负载为 CPU 占用 40%、80% 与内存占用 40%、80% 的情况下宿主机的负载状况进行试验。实验结果如图 5、图 6 所示。

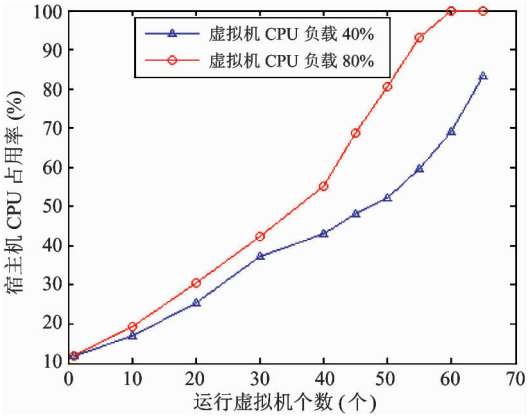

图 5　宿主机 CPU 占用率与虚拟机负载关系图

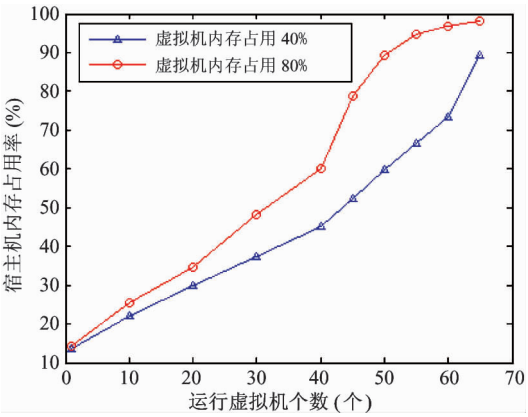

图 6　宿主机内存占用率与虚拟机内存占用率关系图

由实验结果可以看出，随着虚拟机负荷不同，宿主机负荷也有所不同。当虚拟机 CPU 占用率增大时，宿主机 CPU 占用率同时增大；当虚拟机 CPU 占用率达到一定程度时，宿主机 CPU 占用率达到 100%；而对于内存占用率，当虚拟机内存占用总和接近宿主机内存值时，宿主机启用交换内存，此时宿主机内存占用率增加变缓，但宿主机出现明显卡顿现象。实验结果表明，在当前配置条件下，单台宿主机中同时运行虚拟机个数上限为 60 台左右。

3.2　不同网络规模构建时间测试

为了度量不同大小规模的网络环境生成时间，对拓扑规模为 10、50、100 的网络环境进行构建实验。实验利用两台宿主机，宿主机硬件配置与 3.1 节所用宿主机相同。实验结果如图 7 所示。

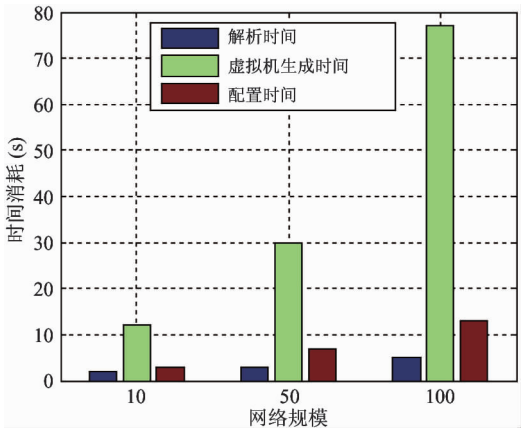

图 7　时间消耗测试结果

实验结果表明，在进行网络环境构建过程中，虚拟机生成所需时间占所有时间比重最大。在试验中，两台宿主机平均分配虚拟机生成任务，由于处于并行工作状态，所需时间比 3.1 节单台宿主机测试时创建相同虚拟机个数所花时间减少约一半。当创建规模为 100 个节点的网络环境时，消耗时间为 95s。

3.3　连通性测试

本节连通性测试通过根据图 1 所示真实网络利用本文构建算法构建得到虚拟网络 V-NET，并利用 V-NET 构建僵尸网络。由僵尸网络组织进行 DDoS 网络安全事件的爆发，攻击方式采用 UDP Flood，DDoS 基本配置如图 8。主机 a、b、c、d、e 对主机 f 进行 DDoS 攻击，攻击频率为 100 数据包/秒，数据包大小为 1KB，攻击持续 5s，在路由器 A、B、C、D、E 以及主机 f 上分别统计经过或收到的数据包数量，统计结果如表 2 所示。

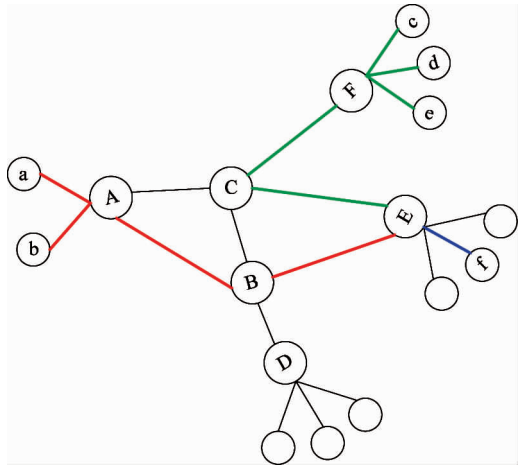

图 8　DDoS 攻击图

表 2　数据包数量统计结果

统计位置	数据包个数	数据包源	数据包目的
路由器 A	1000	a、b	f
路由器 B	1000	a、b	f
路由器 C	1500	c、d、e	f
路由器 D	0		
路由器 E	2500	a、b 、c、d、e	f
路由器 F	1500	c、d、e	f
主机 f	2500	a、b、c、d、e	f

如表 2 所示,主机 a、b 发往主机 f 的数据包路由路径为 F—C—E—f,主机 c、d、e 发往主机 f 的数据包路由路径为 A—B—E—f,路由路径与本文 TRTF 算法得到路径理论值相同,同时在模拟 DDoS 安全事件中,没有发生数据包丢失的现象。因此本文构建的网络环境连通性状况良好,可用于网络安全事件的重现。

4　结　论

本研究通过对网络环境构建问题的研究,提出并实现了网络环境自动构建技术。用此技术可根据实际的网络环境拓扑图,利用虚拟技术自动构建虚拟的网络环境,从而降低了构建成本,由于整个构建过程自动进行,从而减少了配置过程中人工的参与,提高了构建的效率。实验证明,生成规模为 100 的拓扑仅需 95s,从而证实了该技术的可行性,而且连通性测试 DDoS 结果表明自动构建的虚拟网络环境连通性状况良好,整个虚拟网络可用于网络安全事件的模拟重现。

参考文献

[1] Mike H, Robert R, Leigh S, et al. Large-scale virtualization in the Emulab network testbed. In: USENIX 2008 Annual Technical Conference on Annual Technical Conference, Boston, America, 2008. 113-128

[2] Brent C, David C, Timothy R, et al. PlanetLab: an overlay testbed for road-coverage services. *SIGCOMM Computer Communication Review*, 2003, 33(3): 3-12

[3] 王佳宾,连一峰,陈恺. 一种基于分组自适应的网络环境快速构建方法. 中国科学院研究生院学报,2012,29(4):536-542

[4] 全晓莉,周南权,余永辉. 基于虚拟仪器技术的网络实验系统的研究. 计算机工程与设计,2009,32(9):3227-3230

[5] 冯陈伟. 利用 VMware 构建虚拟网络平台. 信息系统工程,2009,(8):78-81

[6] PetrovicD, SchiperA. Implementing virtual machine replication: a case study using Xen and KVM. In: 2012 IEEE 26th International Conference on Advanced Information Networking and Applications (AINA), Fukuoka, Japan, 2012. 73-80

[7] 康辰,朱志祥. 基于云计算技术的网络攻防实验平台. 西安邮电大学学报,2013,18(3):87-91

[8] Rubow E, McGeer R, Mogul J, et al. Chimpp: A Click-based programming and simulation environment for reconfigurable networking hardware. In: 2010 ACM/IEEE Symposium on Architectures for Networking and Communications Systems (ANCS), La Jolla, America, 2010. 1-10

[9] 罗腊咏,贺鹏,关洪涛等. 可编程虚拟路由器关键技术与原型系统. 计算机学报学报,2013,36(7):1349-1363

Research on automatic construction of network environments based on KVM virtualization

Zhang Yun*, Tang Jiqiang**, Yan Jianen*, Zhang Zhaoxin*

(*School of Computer Science and Technology, Harbin Institute of Technology, Harbin 150001)

(**The National Computer Network Emergency Response Technical Team Coordination Center of China, Beijing 100029)

Abstract

The study was conducted to challenge the problems of high-cost, low-efficiency, poor-controllability of the traditional method for construction of network environments, and a novel technique for automatic construction of virtual networks based on the virtualization using a kernel-based virtual machine (KVM). The new technique can be stated below. Firstly, it gives a formalized description of a real network to be structured, and based on the result of the formalized description, it takes advantage of the KVM and the Click virtualization to produce the host and the router in the network. And then, it uses the terminal router routing table flooding (TRTF) algorithm to calculate the routing table, and the host and the router set up themself by running a script. At this time a virtual network is formed. The experimental results indicate that it cost 95s to structure a network having 100 nodes, and the DDoS simulation on the virtual network shows that the network connectivity is in good condition.

Key words: network environment, automatic construction, formal description language, KVM virtualization