文章编号：2095-2163(2020)11-0047-03 中图分类号：TP393 文献标志码：A

基于 IPv6 的网络与信息资产发现关键技术研究

辛 毅，吴 刚，孙洪磊
（哈尔滨工业大学 网络与信息中心，哈尔滨 150001）

摘 要：随着 IPv6 的大规模部署，传统意义的扫描方法已无法完成对资产的快速发现。针对于此，本研究采用无状态扫描技术，收发线程分离，加快扫描速度，采取高速匹配算法，对网络与信息资产的协议指纹进行匹配。同时辅助以管理手段，大大减少了扫描的地址空间，提高了网络与信息资产的发现效率。其次，构建协议指纹库，获取大量通用信息作为发现资产的基础。最后，采用 P2P 方法进行分布式部署，实现数据同步与快速检索，以满足大规模 IPv6 网络下的网络与信息资产的发现。

关键词：IPv6；扫描；资产；协议指纹

Study of key technologies of network and information asset discovery in IPv6 Network

XIN Yi, WU Gang, SUN Honglei
(Network and Information Center, Harbin Institute of Technology, Harbin 150001, China)

【Abstract】 With the large-scale deployment of IPv6, the traditional methods of scanning cannot support the fast discovery of assets. This study uses stateless scanning technology, sending and receiving threads to separate, speed up scanning, and high-speed matching algorithm, by that way the information assets and network protocol fingerprints are matched. The auxiliary management can greatly reduce the scanning address space and improve the discovery efficiency of the network and information assets. Secondly, fingerprint library can provide a large amount of general information as the basis for discovering assets. Finally, this P2P method is used for distributed deployment to achieve data synchronization and fast retrieval to meet the discovery of network and information assets under large-scale IPv6 networks.

【Key words】 IPv6; Scanning; Asset; Protocol fingerprint

0 引 言

近年来，随着信息技术的发展，大量的设备，特别是 ITO 设备都需要接入互联网。由于网络地址分配不均匀，一定程度造成 IPv4 网络地址资源浪费，可供使用的 IPv4 地址已经越来越少，致使网络地址资源即将消耗殆尽。虽然目前可以采用地址转换技术（Network Address Translate：NAT），但仍然不能解决日益增长的需求，IP 地址不足的问题已经成为互联网和通信产业发展的瓶颈。为了从根本上解决 IP 地址空间的不足，需要大力发展下一代互联网协议[1]（Internet Protocol Version 6：IPv6）。IPv6 的 128 位地址格式将以其在 IP 地址数量、移动性、服务质量等方面存在巨大的优势。但是，随着地址数的增加，确定网络与资产进行安全防护变得十分困难。

对于提供网路服务的运营商和企事业单位的网络管理部门来说，详细掌握本单位网络中的全部资产，有效的梳理和管理，针对网络中存在的安全隐患可以精细化排查，合理的配置管辖网络内资源的合理配置，才能够保证网络的安全有效运行。

本文分析了 IPv6 网络下网络与信息资产的特点，针对 IPv6 网络中的网络与信息资产的关键技术进行研究。

1 IPv6 下网络与信息资产特点

随着 IPv6 网络的应用，基于 IPv6 的校园网、私有云及数据中心建设的快速发展，系统规模日益增大、密度不断提升，系统的复杂程度越来越高，运维管理的复杂程度急剧攀升。由于部分企业单位的网络系统，特别是高校校园网，其网络的开放性容易造成网络管理松散，网络资产与信息资产不清等问题。网络中存在大量的服务器、客户主机，其上面部署的网页服务、数据库和中间件等网络基础设施由于缺少必要的管理，导致管理人员无法全面掌握其物理资产和软件资产的情况。更为严重的是网络资产和信息正在向虚拟化迁移，因此在网络上可以很快地新增或部署一台虚拟机，同时由于管理的疏漏，导致服务器上服务开启和端口管控机制也不健全。这些

基金项目：国家重点研发计划项目（2018YFC0830902）；赛尔网络下一代互联网技术创新项目（NGII20170412）。

作者简介：辛 毅（1973-），男，博士，高级工程师，主要研究方向：网络安全、入侵检测与防御。

收稿日期：2020-06-27

资产和数据业务缺少安全检查与访问控制。因此网络中的服务器、客户端主机上的漏洞与脆弱点极易成为攻击者攻入关键业务区的跳板[2-4]。运用多种快速扫描、协议指纹、数据发现、数据获取及数据分析技术,实现基于 IPv6 下校园网的信息资产的快速准确发现,具有重要意义[5-7]。在参考上述研究的基础上,本文同时采取大数据引擎技术,对发现的信息资产进行科学管理,对数据进行重新编排索引,从而实现高效搜索,达到对信息资产的全面掌握与有效管理。

2 IPv6 网络与信息资产关键技术研究

与 IPv4 不同,IPv6 下的网络地址空间巨大,构建高效准确的网络资产的发现与管理,需要从高速扫描、协议与应用指纹库、数据同步与检索等几个部分进行研究。系统架构如图 1 所示。

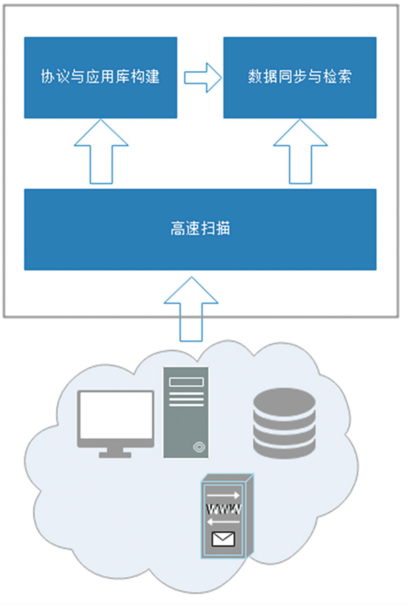

图 1 系统架构图

Fig. 1 System architecture diagram

2.1 高速扫描

为快速地发现网络与信息资产,一方面,对系统的协议栈进行改造,采用无状态扫描技术、收发线程的分离。即发包线程只进行探测包的发送,并对回复的报文进行处理,端口扫描的结果代表目标主机系统对外使用此端口开放某服务。对于使用 TCP 连接的服务探测,需在建立连接的基础上,通过发送探测包,观察回复信息,来判断服务的类型。由于使用半连接端口扫描技术,扫描器将不实际建立连接,在得到端口开放的信息后,系统将自动返回目标主机 RST 报文取消连接,从而达到快速扫描的目的。为了维持 TCP 连接进行服务探测,借鉴 Massscan[8] 的思想,建立统一的结构进行传输控制块 TCB 的记录与管理,系统内使用的方式为 TCB _ ConnectionTable。它数据包时,首先通过解析出的四元组(源 ip、目的 ip、源端口及目的端口)进行与发包前相同逻辑的哈希计算,将得到的结果与 TCB 管理表的掩码进行逻辑与运算,找到该连接在 TCB 管理表中维持了对所有 TCB 的引用,同时负责对其进行统一的创建、管理与销毁。在服务探测过程收到一个的位置。在端口扫描的基础上,针对其开放的特定端口,对 HashTrie 算法[9] 进行了优化。采用基于位图搜索的高速匹配算法,调取协议与应用指纹库进行匹配。该算法主要有位向量 B、F 及校验散列表 M 构成。通过计算散列值,构建位向量表,极大的压缩了内存占用,通过两层过滤的方式,缩小匹配模式集合范围,大大提高了算法的扫描速度。扫描算法见表 1。

表 1 基于位图搜索的高速匹配算法

Tab. 1 High-speed matching algorithm based on bitmap search

```
算法名称: Search
输入: 位向量 B、F,模式 T = T0 T1L… Tn-1、检验链表 M
输出:命中的模式集合
For i = 0 to n - 1Do
h = 0
j = 0
while j < lmax Do
  h = (ah + Ti + j)&(H - 1)
  IFB[h] = 0 Then Break
  IF F[h] = 1 Then
    T = Rnak(F,h)
    Fo reach pattern P in M[t]
      IF Ti…Ti + j = P
        Report (i,P)
      End IF
    End For
  End IF
    j ++
  End While
  End For
```

将相关数据入库,以实现对 IPv6 大规模网络中网络资产特征的快速准确地获取。从而判断出资产的属性。另外一方面,系统针对 IPv6 应用场景进行优化,由于发现资产一般为系统管理员 或者系统安全管理员。因此可以设定特定地址段,并结合 DNS 信息及日志信息进行有针对性地快速扫描,以解决

IPv6 地址空间巨大、扫描时间过长的问题。

2.2　协议与应用指纹库

协议与应用指纹库是判断网络信息资产的基础，此模块对有关协议、操作系统、应用、中间件、第三方架构等特有的信息进行搜集，并对网络中数据进行捕获，采取大数据抽取提取特征并将其入库，作为判断资产的重要指标，主要包括以下内容：

（1）网络协议栈指纹：ACK 序号、TOS、ICMP 地址屏蔽请求、ICMP 错误信息、ISN、FIN 响应、分段重组处理、TTL、最大分片等；

（2）Banner 信息：操作系统 banner 信息、数据库 banner 信息、webserver banner 信息、ftp banner 信息、ssh banner 信息、Cli banner 信息及网络设备、服务器 banner 信息等；

（3）出错信息：系统出错信息、脚本出错信息、数据库出错信息、中间件出错信息，web server 出错信息等；

（4）构架及三方应用信息：路径信息、页面标志、特定文件名等 ；

（5）web 服务器信息：错误返回信息、页面规律信息、页面特征信息、脚本语言信息及出错信息；

（6）后台信息：特定文件信息、特定目录信息、图片 hash 信息等。

2.3　数据同步与检索

为解决 IPv6 下网络与信息资产的发现，采用分布式架构，能高速、高效、准确地进行资产发现。在网络底层采用了 P2P 构架：由于系统没有中央服务器，各个节点之间不存在主从关系，相互平等，共同工作。扫描和信息的融合都在各个节点之间直接进行，不存在单点失效和系统瓶颈。而非中心化的特点，带来了其在可扩展性、健壮性等方面的优势。新的节点可以方便的加入功能覆盖网络，并且随着节点数的不断增加，整体的资源和发现能力也会不断上升。P2P 网络环境下，由于每个节点既是服务器又是客户机，减少了对传统 C/S 结构服务器计算能力、存储能力的要求。由于资源分布在多个节点，更好的实现了整个网络的负载均衡。通过 P2P 协议的调度，系统调度模块将任务进行分解，通过任务调度将分解的任务分配到合适的节点中，对多个节点实现智能的任务分发、负载均衡、进度监测、存储入库、数据分析及统计展现等，以保证系统的效率。数据存储与检索采用 Elastic search[10]，对系统扫描、指纹识别等等非结构化信息进行整合，建立索引，并对数据进行聚合，从而根据不同场景实现实时、动态地生成相关数据。如，在 Struts[11] 新漏洞出现时，快速检索出资产中采取 Strtuts2 构架服务器的详细信息，从而做到快速应急响应。同时通过 elastic search 词法分析器，使系统查询结果更加快速、准确，以提高系统的检索速度。

3　结束语

本文对 IPv6 下资产的发现与管理的关键技术进行了详细论述：

（1）高速扫描采用无状态扫描技术，收发线程的分离为快速地发现网络与信息资产。

（2）协议与应用指纹库，对有关协议、操作系统、应用、第三方架构等特有的信息进行作为判断资产的重要指标。

（3）数据同步与检索，采用 P2P 分布式架构，在端口扫描的基础上，将相关数据入库同对于系统扫描、协议与应用指纹信息等非结构化信息进行整合，建立索引，并对数据进行聚合，从而根据不同场景实现实时、动态地生成相关数据。为 IPv6 网络下进行高速高效、准确地进行资产发现奠定了基础。

参考文献

[1] Network Working Group S. BradnerIP: Next Generation (IPng) White Paper Solicitation RFC1550 https://tools. ietf. org/html/rfc1550.html

[2] ZHANG Y, ZHANG Y, FANG B. A Kind of OpenStack Auto-Deployment Architecture Based on Workflow[J]. Telecommunications Science, 2014, 30(11): 14.

[3] XU H, ZHANG X, WANG C, et al. Patents Analysis of Precision Agriculture Technology and System[J]. Science Focus, 2018, 10(5): 15-33.

[4] 曾诚，何克清，李兵，等. 基于 RGPS 领域资产聚合的按需服务发现方法[J]. 小型微型计算机系统，2012，33(5)：921-928.

[5] 徐前方. 在 IPv4/v6 环境下网络拓扑发现的研究与实现[J]. 计算机工程与设计，2006，27(23)：4507-4509.

[6] 王勇. 软件过程资产库的研究与实现[J]. 计算机应用与软件，2016，33(7)：106-108.

[7] 刘淑霞，王桂玲，赵卓峰. 一种服务网络有序状态分析方法[J]. 小型微型计算机系统，2013，34(12)：2753-2757.

[8] GRAHAM R. MASSCAN: Mass IP port scanner. https://github.com/ robertdavidgraham/masscan.

[9] 张萍，刘燕兵，于静，等. HashTrie：一种空间高效的多模式串匹配算法[J]. 通信学报，2015，36(10)：172-180.

[10] GORMLEY C, TONG Z. Elasticsearch: The definitive guide: A distributed real-time search and analytics engine[M]. " O'Reilly Media, Inc.", 2015.

[11] The Apache Software Foundation, S truts2 Security Bulletins https://cwiki. apache. org/confluence/display/WW/Security + Bulletins