

基于 VxLAN 的网络分组策略研究与分析

孙 兵

(华为技术有限公司, 江苏 南京 210012)

摘 要: IP 网络接入的终端数量和种类越来越多(如企业员工 PC 或 TC 接入、访客手机接入、IP Phone/打印机/物联终端等接入)、承载的业务类型日益丰富,例如企业办公网、生产网、视频监控网、智能楼宇物联网等统一到 IP 网络上承载,员工接入方式也多种多样,如公司分支机构或总部接入、出差远程 VPN 接入等,传统基于 ACL 的网络策略无法应对企业 IP 网络业务场景的变化,面临管理维护复杂度的重要挑战。该文给出一种基于用户逻辑分组(安全组)的策略模型,并全面分析企业应用场景,给出基于 VxLAN 网络的安全组全网同步方案,实现网络策略与网络属性(IP/VLAN/MAC)等无关,大大降低企业 IP 网络策略数量和变更频率,并在实际大型企业的 IT 网络进行应用评估,可以将数以万计的策略数量降低到百计,应用价值高、效果明显,指明了企业 IP 网络策略的演进方向。

关键词: 企业网络;网络分组策略;安全组;访问控制列表(ACL);软件定义网络(SDN)

中图分类号: TP393

文献标识码: A

文章编号: 1673-629X(2021)01-0126-05

doi:10.3969/j.issn.1673-629X.2021.01.023

Research and Analysis of Network Grouping Policy Based on VxLAN

SUN Bing

(Huawei Technology Co., Ltd., Nanjing 210012, China)

Abstract: The number and types of IP network access terminals are increasing, such as PC or TC access for enterprise employees, mobile phone access for visitors, IP Phone/printer/IOT terminal access, etc. The types of services are increasingly diversified, such as office network, production network, video monitoring network, intelligent building IoT, etc., which are all carried on the IP network. The employee accessed methods are also diverse, such as access to branch offices or headquarters of the company, remote VPN access on business trips, etc. Traditional network policy based ACL can't cope with the changes of enterprise IP network scene, so it faces the important challenge of management and maintenance complexity. We propose a network policy model based on user logical grouping (security group) and analyze the enterprise application scenario comprehensively. The security group synchronization scheme based on VxLAN network is presented, and the network policy has nothing to do with network attributes (IP/VLAN/MAC), which greatly reduces the number and change frequency of enterprise IP network policies. It can be applied in the IT network of actual large-scale enterprises, reducing tens of thousands of policy number to hundreds. With high application value and obvious effect, we indicate the evolution direction of enterprise IP network policy.

Key words: enterprise network; network grouping policy; security group; access control list (ACL); software defined network (SDN)

0 引 言

企业网络,终端由原来的有线连接接入、向 WiFi 无线连接演进,终端类型由 PC 机向便携机、PAD 和智能手机扩展,办公地点由固定场所办公到移动办公和在家远程办公拓展,接入终端设备归属由公司资产到个人资产,企业网络由办公网向物联网延伸。企业网络终端接入连接无线化、移动化、物联化,接入终端类型丰富、接入地点多样,导致传统基于 IP 地址 ACL、VLAN 的网络策略控制方式不能适应新的网络场景变

化,因为终端的 IP 地址和 VLAN 会变化,传统的网络策略会随之变化,这样给企业 IT 运维人员的网络策略管理带来很大的复杂度。因此,迫切需要一种更加简单高效、与网络拓扑和 IP/VLAN 等网络自身分配规划无关的策略控制方案。

1 基于 VxLAN 的网络分组策略方案

1.1 方案概述

新一代网络策略控制方案的目标是设计一种能够

满足无论用户身处何地、使用哪个 IP 地址,都可以保证该用户获得相同的网络访问权限的技术解决方案。用户从认证设备接入时,会关联映射一个安全组(又称策略组),用以设置该用户的安全策略;当该用户从其他设备(非认证点)其他地点接入网络,会执行相同的安全策略。该解决方案中的一项关键技术是,需要将用户与安全组的关联映射信息同步到其他设备上。同步方法有如下几种:

(1)SDN 方案的控制器 Controller 全局同步;

(2)通过专用协议在设备之间进行扩散同步(带外方式);

(3)通过用户流量报文将安全组信息携带到其他设备(带内方式)。

上述方案前两种需要额外增加同步机制和协议,第一种方案在 Controller 和网络设备之间增加同步机制,第二种方案在网络设备之间增加机制,方案复杂度相对高,且网络规模大、设备数量多,同步机制本身的性能和网络带宽开销会是问题。第三种方案不需要增加设备间额外的同步机制,仅仅在正常的用户业务流量报文中增加安全组字段,沿途策略点设备进行解析识别和策略执行处理即可。该文设计的技术方案属于上述第三种方案,由 VxLAN 报文头携带安全组通过网络扩散到不同的网元设备中。

文献[1-8]主要介绍传统网络策略的关键技术,仍然基于 IP 五元组来定义网络策略,并结合集中策略管理和 QOS 联动形成增值业务。文献[9]中 IETF 定义了 VxLan 协议报文头部包含的策略组 ID 格式,可作为文中方案的详细实现参考。文献[10-15]重点论述了网络策略集中管理的架构,包括 SDN 方案架构等,解决网络策略跨设备同步的问题。

1.2 方案场景

基于 VxLAN 的网络分组策略方案场景如图 1 所示。

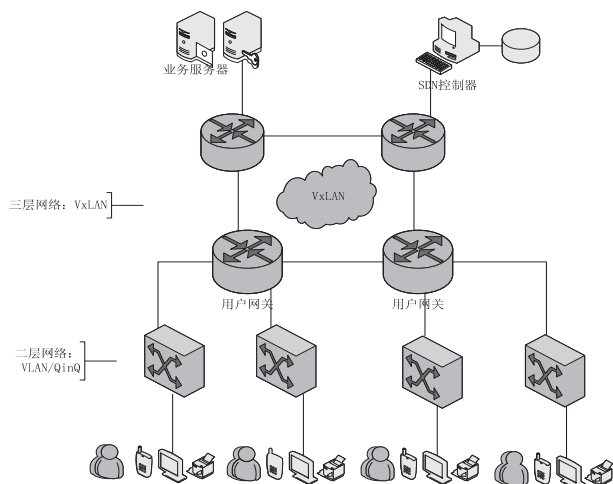


图1 基于 VxLAN 的网络业务策略应用场景区图

(1)用户认证通过网关设备认证上线,网关携带用户 ID 向控制器请求认证结果,控制器分配用户 ID 与安全组 ID 的映射关系,并下发至用户认证网关设备。

(2)用户在网关设备通过认证,用户终端访问报文经由网关设备时,在 VxLAN 头部打上安全组 ID 标识 Group Policy ID,然后转发出去,访问相应的用户业务。

(3)沿途 VxLAN 网关设备会解析报文中 VxLAN 头部携带的安全组 Group Policy ID 信息,根据安全组查询访问策略,实施安全组间的网络访问控制。

该方案可以在企业网络中实现与 IP 地址、MAC 地址、VLAN 等网络属性无关的网络访问策略控制,替代的是使用安全组间的访问策略来控制,例如企业普通员工组不能访问人力资源组的信息资产,只有财务组员工才能访问财务数据库服务器,等等。

安全组策略方案中,能够做到不关心用户终端的 IP 地址/VLAN、接入位置、接入方式(有线/无线/VPN 等),因此网络规划变化(IP 网段或 VLAN 变化)或者接入位置接入方式发生变化(用户出差、在家远程办公等),都不需要更改原先的安全组间策略配置,这相对传统基于 IP 五元组的 ACL 控制方式有着极大的优势。此外安全组方案模型中,用户 ID 与安全组 ID 是多对一映射模型,例如财务组包含 100 个员工、HR 组包括 200 个员工,因此经过汇聚后的组间策略数量规模会大大减少,降低企业 IT 人员的网络访问策略日常管理维护复杂度。

1.3 方案原理

(1)方案架构。

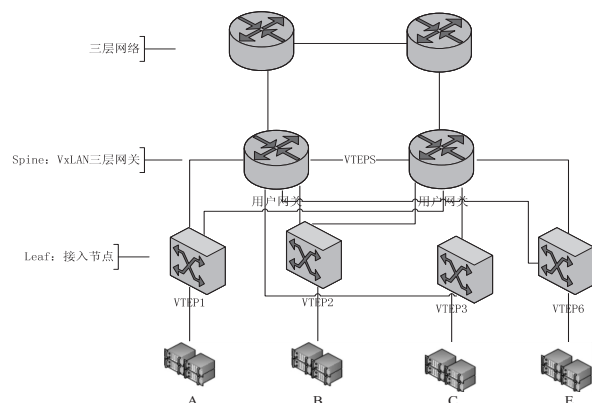


图2 基于 VxLAN 的网络业务策略组网架构图

图2为典型 Spine-Leaf 架构 VxLAN 组网,可使用集中式或分布式网关。Spine 节点的设备使用设备间集群技术增加可靠性。Leaf 设备作为 VxLAN 二层网关,用于传统网络与 VxLAN 网络的二层互通。Spine 设备作为 VxLAN 三层网关,负责不同网段的 VxLAN 网络之间的通信,以及不同网段的 VxLAN 网

络与非 VxLAN 网络之间的通信。

VxLAN 携带用户访问策略,通过 VNI (VxLAN network identifier) 携带安全组/策略组 ID,接入交换机类似数据中心中的 TOR (top of rack) 设备,作为 VxLAN 网关的 VTEP (VxLAN tunnel endpoint),且是 VxLAN 二层网关。汇聚/核心交换机作为 VxLAN 三层网关。

二层广播抑制问题,对于 ARP 的广播请求报文,通过 SDN Controller 代理应答的方式进行处理。流程如下:

(a) SDN Controller 在用户上线认证时缓存其 {IP,MAC} 映射关系,生成 ARP 缓存表。

(b) 二层网关收到 ARP 请求报文,直接上送 SDN Controller 处理。

(c) SDN Controller 向该接入设备下发对应的 ARP 表项,如果 Controller 上 ARP 表项则进行广播

处理。

(d) SDN Controller 也可以选择将 ARP 广播报文修改为单播报文发给被请求设备,这样可保证整个 ARP 请求/应答链路上的设备都能使用 ARP 报文实现其他功能。

(e) ARP 报文上送和下发可使用 SDN OpenFlow 协议。

通过在相同 VNI 之间建立 VxLAN 隧道实现相同安全组内用户互访;不同 VNI 之间的互访原本存在 VxLAN 三层网关上实现,可在三层网关引入安全组间策略,即利用源 VNI 和目的 VNI 之间的组间策略进行网络隔离或放行策略控制。

(2) 方案流程。

基于 VxLAN 的网络分组策略方案流程如图 3 所示。

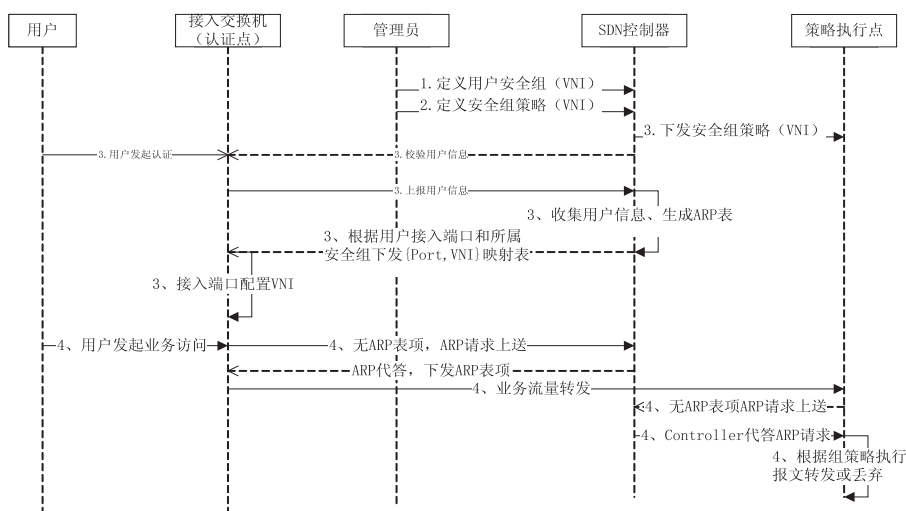


图 3 基于 VxLAN 的网络业务策略方案流程时序图

(a) 控制器根据管理员配置,生成安全组与 VNI 的映射关系,依据安全组间策略生成 VNI 之间的互访策略,并将 VNI 互访策略下发到策略执行点。

(b) 控制器与认证点设备交互,完成对接入用户的认证,依据认证结果将用户划分到对应的安全组。认证点建议使用接入交换机,这样便于直接获取用户终端接入的物理端口。

(c) 控制器根据用户所属安全组,向接入设备下发 {Port, VNI} 映射关系,接入设备依据该映射关系在对应端口上配置 VNI,保证接入用户的所有流量都走该 VNI 标识的 VxLAN 隧道。

(d) 控制器完成所有设备的 ARP 代答,实现对 ARP 的广播抑制。控制器上的 ARP 缓存表在用户认证过程中获取和下发。

(3) 相同网段用户的安全组互访。

相同网段用户的安全组互访策略如图 4 所示。

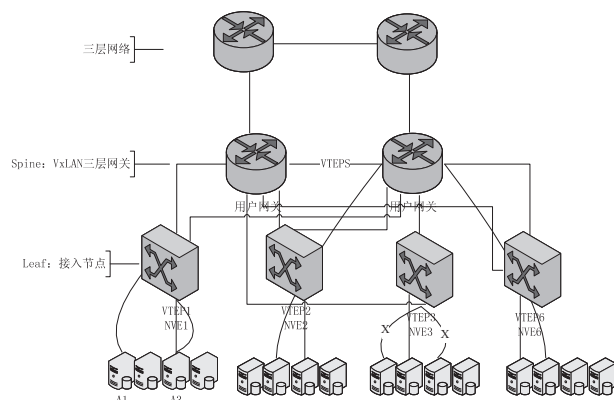


图 4 相同网段用户的安全组互访策略

(a) VxLAN 网络中,不同 VNI 的流量二层不互通,所以在接入设备就保证了不同安全组用户之间的隔离。

(b) Site A 中两个相同网段的用户 A1 和 A3,属于同一安全组,使用相同的 VNI,在 NVE1 设备上实现

互通。

(c) Site B 中 B2 和 Site E 中 E2 也是同一安全组中相同网段的用户, B2 和 E2 使用相同的 VNI, VTEP2 和 VTEP6 之间建立 VxLAN 隧道实现互通, 网关只为该隧道提供三层转发通道。

(d) Site C 中两个不同安全组的用户, 使用不同的 VNI 接入 NVE3, 即使使用同一网段也无法互通。

(4) 不同网段用户的安全组互访。

不同网段用户的安全组互访策略如图 5 所示。

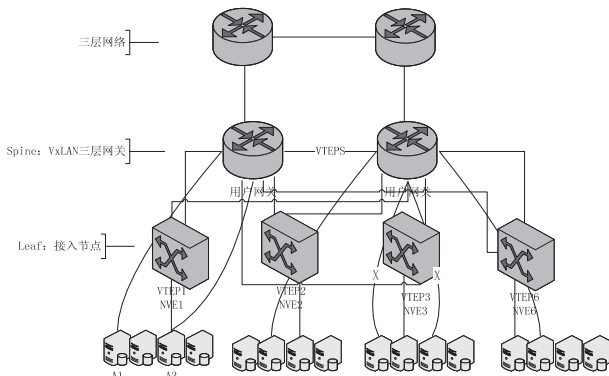


图 5 不同网段用户的安全组互访策略

(a) VxLAN 网络中, 不同 VNI 的流量经过三层网关互通, VxLAN 网络与非 VxLAN 网络也是通过三层网关互通。三层网关上 VxLAN 报文处理流程: 先剥离外层 VxLAN 封装, 依据原始报文走三层流量转发, 重新进行 VxLAN 封装发出。

(b) 在三层网关上执行安全组间策略, 控制器提前向三层网关设备下发 VNI 之间的访问策略, VxLAN 报文在处理流程中, 需判断源 VNI 与目的 VNI 的访问策略, 如允许访问则正常转发, 否则丢弃报文。源 VNI 在报文头部中已携带, 目的 VNI 通过报文内层头部目的 IP 查找获得。

(c) Site A 中两个相同网段的用户 A1 和 A3, 属于同一安全组, 使用相同的 VNI, 在网关上实现互通。

(d) Site B 中 B2 和 Site E 中 E2 分属不同安全组, 网段也不同, 通过三层网关转发实现不同 VNI 之间互通, VNI 之间的策略需预设为允许互通。

(e) Site C 中两个不同安全组的用户, 分属不同安全组, 在三层网关转发时依据 VNI 之间策略, 不允许互通。

(5) 用户移动接入与 QOS 策略。

当用户接入位置发生变化时, 需要重新上线认证, 原端口与 VNI 的绑定关系删除, Controller 为新接入端口下发 {Port, VNI} 映射关系, 新端口绑定 VNI。策略执行点设备不感知用户接入位置的变化。用户所属安全组不变, 其 VNI 不变, 该用户的权限也没有发生变化。Controller 通过 QOS 策略保证 VIP 安全组的优先

级, 为 VIP 安全组对应的 VNI 下发 QOS 策略。VIP 用户在接入 Leaf 节点时, 根据 VNI 的 QOS 策略, 修改内外层 IP 头部的 DSCP, 保证 VIP 用户在整网中优先级。

三层网关除了执行不同 VNI 之间的组间策略, 也可以执行不同 VNI 之间的访问速率控制策略。例如限制 Guest-VNI 用户访问 Internet-VNI 的速率。

1.4 方案效果评估

以某大型企业为例 (50+分支机构, 15 万员工接入, 6000+物理服务器、每服务器运行 20 个虚拟机) 来评估方案应用效果。

该企业原始用户和业务访问策略需求如下:

(1) 企业终端类型分为 PC、IP Phone、TC 云终端、打印机、物联终端; 用户类型分为研发、市场、财务、人力资源、访客;

(2) 用户/终端认证前策略: 只能访问指定的认证服务器、客户端软件下载服务器, 不允许用户/终端间互访;

(3) PC 策略: 允许基于角色 (研发、市场、财务、人力资源) 的服务器业务资源访问, 同一角色内用户间可以互访, 跨角色访问受控 (默认禁止跨角色访问);

(4) TC 云终端策略: 允许 TC 间互访 (TC 的 IP 语音业务)、TC 只能访问服务器桌面云 VM 资源、TC 不能访问其他业务服务器 VM 资源;

(5) IP Phone 策略: 允许 IP Phone 间互访 (IP 语音业务)、IP Phone 不允许访问除了语音业务之外的服务器 VM 业务资源;

(6) 打印机策略: 不允许互访、只能访问 ePrint 打印服务器;

(7) 访客户策略: 只允许访问 Internet, 不能访问服务器所有企业业务资源, 禁止访客间终端互访。

该企业的上述策略需求, 长期以来通过基于 IP 网段/IP 地址的 ACL 技术方案来实施控制, 因为涉及到 50+分支机构, 每个机构又有多个办公楼、多个用户网关, 每个机构又有多种类型终端和用户接入, IP 网段之间的互访策略关系非常复杂, 经与企业 IT 运维人员现场调研核实, 实际有 4 万多条 ACL 网络访问策略数量。分支机构随着办公楼宇基建扩展、人员扩充, 网络规划调整和网络 IP 网段变更时有发生, 按照当前基于 ACL 的网络策略模型及方案, 会要求 IT 运维人员能够在 4 万多条已有的 ACL 策略中进行精准变更, 这个 ACL 策略变更难度和复杂度都非常高, 对日常网络运维是个巨大挑战。IT 运维人员坦言, 实际操作中已有的 ACL 策略根本不敢改变, 害怕对已有业务产生影响, 往往是网络变化时会新增 ACL 策略, 这样长期累积下来就会导致大量实际无用的僵尸 ACL 策略存在, 无人敢问津, 不仅仅网络策略管理维护难, 也存在网络

安全隐患(僵尸 ACL 策略被恶意利用)。

通过与该企业 IT 运维工程师交流,评估应用文中所述分组策略方案,将企业 IT 的终端分为 PC/TC/IPhone/打印机/访客 Guest 几个组,PC 上用户进一步按照研发、市场、财务、人力资源分成四个组,服务器侧 VM 同时匹配终端和用户类型进行相应分组。经过细化分析设计、考虑后续业务扩展,给该企业总共设计 16 个安全策略组,定义组间互访策略(Permit/Deny),并预留后续可能部署的组间访问速率限制,策略总数目会降低在 512 个以内。因为是按照终端/用户/业务类型分组来定义网络访问策略,与 IP 网段/IP 地址无关,因此无论企业分支机构和总部的物理网络/IP 网段如何变化,都不需要变更 512 条以内的既定组间访问策略。该方案的实际应用价值获得该企业 IT 运维工程师的高度认可。

1.5 企业园区网络引入 VxLAN 策略面临的问题

典型园区组网普遍采用三层网络架构,有线无线业务共存,存在总部与分支互访、远程 VPN 用户接入等场景。无线接入的 AP 与接入交换机的互连,SSL VPN 设备与出口路由器的互连,都可归属同一模式:单端口上多用户多业务流量混合承载。

VxLAN 组网中,报文进入 VxLAN 隧道的方式使用二层子接口,进一步通过 VLAN 关联。Spine-Leaf 架构 VxLAN 原本针对数据中心网络场景设计,网络为 VM 提供服务,Leaf 节点下面一般直接连接 VM 或者 Open switch,这些设备可以通过 Port 或 VLAN 来区分,因此使用三种二层子接口方式接入 VLAN 即可满足数据中心网络场景。

园区网络单端口上多用户流量混合,如用 VLAN 识别不同用户,则可以正常接入 VxLAN,但需要接入设备具备基于 VLAN 区分用户的能力,在用户认证过程中与 Controller 交互,为不同安全组用户下发不同功能 VLAN。另外一种更有效的方法是直接根据用户 ID 来映射 VNI,即认证过程中 Controller 维护 {用户 ID,VNI} 的映射关系。

2 结束语

以 IP 网络接入的终端数量类型和承载的业务等发生变化,传统 ACL 网络策略模型和方案面临的重大挑战作为问题输入,研究分析给出网络分组策略方案。用户和策略分组(安全组)映射关系的网络级传播同步机制是方案的关键,系统分析策略分组传播的几种备选方案,给出性能和通过效率高的带内同步方案(基于 VxLAN 的带内同步安全组),详细分析方案架构、安全组映射和同步流程,结合 VxLAN 实际部署给出技术参考方案分析和设计。

参考文献:

- [1] VALENZA F, SPINOSO S, BASILE C, et al. A formal model of network policy analysis[C]//2015 IEEE 1st international forum on research and technologies for society and industry leveraging a better tomorrow (RTSI). Torin, Italy: IEEE, 2015.
- [2] PRAKASH C, ZHANG Y, LEE J, et al. PGA: using graphs to express and automatically reconcile network policies[C]//Proceedings of the 2015 ACM conference on special interest group on data communication. Tokyo, Japan: ACM, 2015: 29-42.
- [3] CUI L, TSO F P, JIA W. Enforcing network policy in heterogeneous network function box environment[J]. Computer Networks, 2018, 138: 108-118.
- [4] TEO L, AHN G J. Towards effective security policy management for heterogeneous network environments[C]//Eighth IEEE international workshop on policies for distributed systems & networks. Bologna, Italy: IEEE, 2007: 241-245.
- [5] LEE SUNG-HYUCK, BANG JONG-HO, JEONG SEONG-H. End-to-end QoS interoperation apparatus and method in heterogeneous network environment: USA, US7944833[P]. 2011-05-17.
- [6] KWAK J Y, KANG S H, SHIN Y Y, et al. Network policy-based virtualization controller in software-defined networks[C]//Sixth international conference on advances in future internet (AFIN 2014). Lisbon, Portugal: IARIA XPS Press, 2014.
- [7] OOWARI T, YOSHINE A, MIZUNO O. B-7-72 the dynamic network control method based on user requests using OpenFlow[C]//IEICE general conference. Tokyo, Japan: CEATEC, 2015.
- [8] GUTIERREZ P A A, MILOUCHEVA I. Automated QoS policy adaptation for heterogeneous access network environments[C]//IEEE 2007 second international conference on systems and networks communications. Cap Eterel, France: IEEE, 2007.
- [9] SMITH M. Draft-smith-VxLAN-group-policy-05[S]. USA: IETF, 2015.
- [10] 董黎刚, 何博翰, 徐倜杰, 等. 面向 SDN 的动态网络策略部署与实现[J]. 电信科学, 2016, 32(10): 137-149.
- [11] 桂勇胜, 陶妍丹, 钟 华. 网络策略控制技术[J]. 现代计算机, 2012(17): 21-24.
- [12] 李庆海, 张德运, 孙朝晖, 等. 层次化网络策略管理关键技术的研究与实现[J]. 西安交通大学学报, 2003, 37(12): 1216-1219.
- [13] 叶 星, 罗兴国, 李 丰. 一种 OpenFlow 网络策略冲突解决方案[J]. 信息工程大学学报, 2018, 19(2): 234-239.
- [14] 白连红, 徐 澍. 基于 SDN 校园网络的用户体验提升策略研究[J]. 软件导刊, 2015, 14(10): 129-130.
- [15] 孙红雨. 制造业企业的计算机网络策略架构的实证研究[J]. 煤炭技术, 2012, 31(3): 277-278.