

基于 SDN 校园网络运维设计研究^①

刘 敏

(湄洲湾职业技术学院现代教育技术中心 福建莆田 351119)

摘 要 [目的] 随着校园网网络应用的不断增长, 网络中的流量监控、故障排查、风险预警等问题面临巨大挑战。[方法] 本文运用SDN技术对南向基础网络自动化部署, 在对北向构建网络管理节点的校园网络基础上, 利用gRPC、ERSPAN技术设计了校园网络运维平台, 对网络中的物理链路、TCP质量、业务进行分析。[结果] 提高校园网络的运维效率。[局限] 在大数据、人工智能等重要技术在网络运维方面的运用还未涉及。[结论] 对后继智能化的网络运维研究具有借鉴意义。

关键词 网络维护; 软件定义网络; gRPC; ERSPAN

Research on Campus Network Operation and Maintenance Design Based on SDN

Liu Min

(Center for Modern Education Technology Meizhouwan Vocational Technology College, Putian, Fujian, 351119, China)

Abstract [Objective] With the continuous growth of campus network applications, network traffic monitoring, troubleshooting, risk warning and other issues are facing great challenges. [Methods] In this paper, SDN technology is used to deploy the basic network automation in the south, and the campus network of network management node is constructed in the north. The campus network operation and maintenance platform is designed by using gRPC and ERSPAN technology. The physical link, TCP quality and business in the network are analyzed. [Results] The operation and maintenance efficiency of the campus network is improved. [Limitations] The application of important technologies such as data and artificial intelligence in network operation and maintenance has not yet been involved. [Conclusions] It is of reference significance to the future intelligent network operation and maintenance research.

Keywords Network Maintenance; Soft Defined Networking; gRPC; ERSPAN

①本课题得到 2019 年福建省中青年教師科研項目“基於 SDN 的網絡運維管理研究與應用”(高校教育信息化專項)的資助(編號: JAT191930)。

1 引言

随着云计算、人工智能、虚拟现实等各种技术在校园网中的应用,网络规模、数据流量快速增长。传统的网络运维方式越来越难以满足校园网络的应用需求。

①网络越来越复杂化,采用传统运维模式,运维成本会越来越高。随着校园网信息化程度的不断深入,各项信息化业务不断上线,面对有不同网络需求的各项业务,使得传统网络运维方式面临前所未有的挑战,运维难度越来越高。

②传统的运维系统大部分是故障驱动,缺少有效的干预手段及事前预测。故障驱动是在运用中出现故障而需要服务时,采用人工登记报修、电话或口头等方式告知运维人员,运维人员在接到通知后,通过电话沟通、远程操作等方法查找故障原因,提出可能的解决方案。

③传统的运维系统是分钟级数据采集,无法实现实时、精准数据采集。传统TCP/IP技术运维采用SNMP协议,从路由器、服务器、交换机等被管设备获取设备运行的数据,SNMP轮询的方式必须要在一定的间隔时间内不断地进行轮询,间隔时间太短增大了网络通信堵塞的风险,一般以分钟级的时间内采集和统计数据,同时MIB II中大量的变量是只读变量,可写变量太少。

④传统的运维系统重点在于网络监控,缺乏网络关联分析。网络监控工具可以将校园网整张网的布局和设备在网络监控平台呈现出来,利用图形化显示当前网络的连通性,直观了解整张网的网络设备,却无法知道网络设备的网络质量是否能满足业务需求。

针对校园网网络中所面临的运维现状,SDN网络架构的基础上,采用ERSPAN对流量

进行实时采集,并进行流量分析,通过推模式主动把设备数据信息上送采集器,从而实现比传统SNMP查询方式更实时、更高效的数据采集性能。

2 相关工作

2.1 SDN (软件定义网络)

SDN技术得到了很多开源组织、创业厂商、设备制造商和运营商的认可和推崇,SDN让网络运作简化的思想是当今众多的网络平台所共同认可的。其核心思想是通过标准化技术实现控制平面与转发平面分离,从而简化网络管理,采用高性能API Gateway提供符合RESTful标准的北向API,向应用层开放使网络的转发功能具有可编程性,南向API支持NetConf、Openflow、BGP-LS、PCEP、SNMP等接口标准,实现对网络流量的灵活化、集中化、细粒度的控制,从而为网络的集中管理和应用创新提供了良好的平台^[1-3]。

2.2 gRPC、ERSPAN

gRPC (Google Remote Procedure Call, Google远程过程调用)^[4]是Google发布的基于HTTP2.0传输层协议承载的高性能开源软件框架,目标是让远程服务调用更加简单、透明,RPC框架负责屏蔽底层的传输方式(TCP或者UDP)、序列化方式(XML/Json/二进制)和通信细节,遵从server/client模型,客户端可以像调用本地函数一样调用server端提供的接口;提供了支持多种编程语言的、对网络设备进行配置和管理的方法,通信双方可以基于该软件框架进行二次开发。

ERSPAN (Encapsulated Remote Switch Port Analyzer, 封装远程端口镜像)^[5]是一种三层远程端口镜像技术,通过复制指定端口、VLAN或CPU的报文,并通过GRE隧道将

复制的报文发送到远程数据监测设备, 使用户可以利用数据监测设备分析这些报文 (称为镜像报文), 以进行网络监控和故障排除。

3 校园网络运维平台研究与设计

3.1 运维平台系统架构 (图 1)

整个系统分为: ①开放的API, 为运维应用以及其他上层应用提供分析能力; ②数据分析平台, 采用Spark、Flink等分布式计算引擎以及AI人工智能模型库完成数据在线/离线分

析任务; ③基于大数据的数据采集器, 分布式部署架构实现数据采集能力的横向扩展以满足不同网络规模的数据采集需求; ④基于SDN的基础网络, 控制器通过NETCONF等方式向设备下发配置, 实现对网络设备的管理, 同时控制器可以根据分析器提供的分析数据, 为网络设备下发配置, 对网络设备的转发行为进行调整, 也可以控制网络设备有选择地对数据进行采样和上报^[6]。

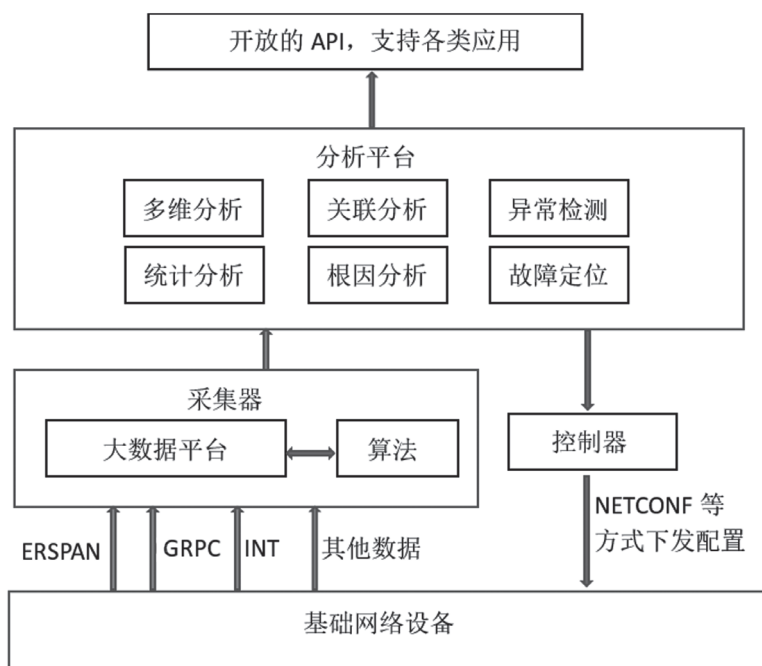


图 1 校园网络的运维系统架构

3.2 基于 EVPN 的 SDN 校园网络运行机制

利用EVPN技术实现SDN在校园网络中的数据控制平面和数据转发平面分离。通过新增5种BGP EVPN消息类型EVPN NLRI (Network Layer Reachability Information, 网络层可达性信息) Ethernet Auto-discovery Route (RT-1)、MAC/IP Advertisement Route (RT-2)、Inclusive

Multicast Ethernet Tag Route (RT-3)、Ethernet Segment Route (RT-4)、IP Prefix Advertisement Route (RT-5)。在数据控制面, 引入RR (路由反射器), 进行核心设备与分支设备协商ibgp evpn邻居, 所有的分支设备都只和RR建立BGP对等体关系, RR发现并接收分支设备发起的RT-2或者RT-3的路由通告后, 连接后形成Client列

表，将从分支设备收到的路由反射给其他所有的分支设备，实现控制面的路由转发。在数据转发面，依靠RT-3建立BUM广播表，在每个分支设备都通告自己的VNI（虚拟网络实例）信息，这样，每个分支设备都有全网的VXLAN信息以及VXLAN和下一跳的关系。分支设备会和那些跟自己有相同的VXLAN的下一跳自动建立VXLAN隧道，并将此VXLAN隧道跟这些相同的VXLAN关联，对每个VXLAN而言，所有这些建立并关联的VXLAN隧道就构成了BUM广播，形成了二层广播域隧道，实现了数据转发。

一跳自动建立VXLAN隧道，并将此VXLAN隧道跟这些相同的VXLAN关联，对每个VXLAN而言，所有这些建立并关联的VXLAN隧道就构成了BUM广播，形成了二层广播域隧道，实现了数据转发。

3.3 校园运维系统数据采集机制（图 2）

gRPC是设备软件平台的能力，一次订

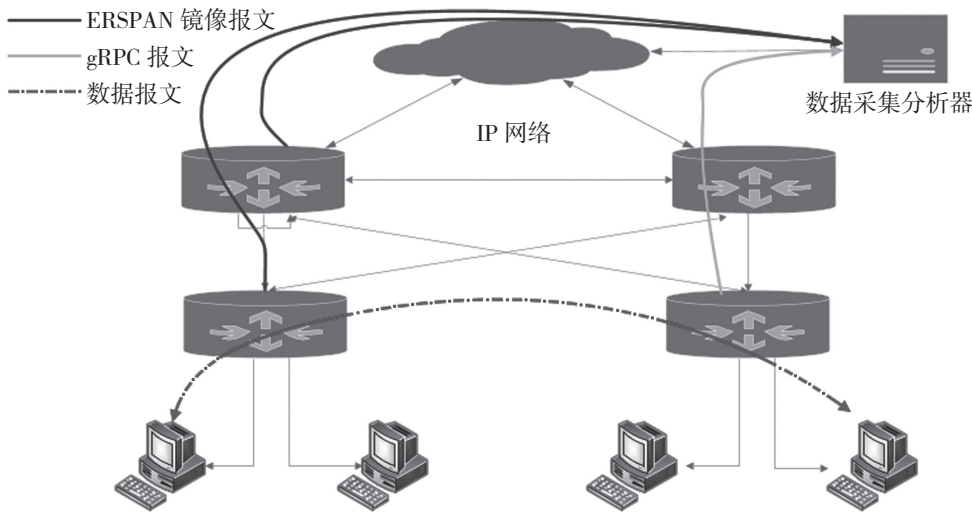


图 2 基础网络数据采集机制

阅，多次推送。采集设备管理、接口管理、IP转发、LLDP、系统日志等业务数据，通过TCP协议打通采集器与设备的数据链路，利用Tls协议对通道加密和双向证书认证，进行安全通信。采集器与设备进行安全通信后，在基于http 2.0协议的基础上，设备指定服务端口等待采集器发起的连接请求，采集器执行相关程序登录到设备并调用proto文件提供的gRPC方法向设备下发配置和发送订阅需要采集的接口流量统计、CPU、告警等数据信息的请求消息，设备以Protocol Buffer编码等形式回复应答消息。

ERSPAN是对原始基于字节流的传输层流量报文进行镜像，可以对TCP报文转发路径上

的下发流匹配规则，将TCP报文镜像到采集器，实现对应用流的流量统计、路径还原、延时计算、应用识别等分析处理。

4 基于 SDN 校园网络运维的实现

4.1 gRPC 远程监控

gRPC采用客户端/服务器模型，使用http 2.0协议传输报文（如图3所示），实现对设备自动读取各种统计信息（CPU、内存、接口等），根据采集器的订阅要求将采集的信息通过gRPC协议上报给采集器，实现更加实时、



图 3 gRPC 数据发送

高效的数据采集功能。

gRPC的工作机制：①服务器通过监听指定服务端口来等待客户端的连接请求。②用户通过执行客户端程序登录到服务器。③客户端调用proto文件提供的gRPC方法发送请求消息。④服务器回复应答消息。

设备数据采集模式：设备作为gRPC服务器，采集器作为gRPC客户端模式。

(1) 公共proto 文件，公共RPC 方法，其内容和含义如下：

```
syntax="proto2";
package grpc_service;
message GetJsonReply { //Get 方法应答结果
    required string result = 1;
}
message SubscribeReply { //订阅结果
    required string result = 1;
}
message ConfigReply { //配置结果
    required string result = 1;
}
message ReportEvent { //订阅事件结果定义
    required string token_id = 1; // token_id
    required string stream_name = 2; //订阅的事件流名称
    required string event_name = 3; //订阅的事件名
    required string json_text = 4; //订阅结果json 字符串
}
message GetReportRequest { //获取事件订阅结果请求
```

```
    required string token_id = 1; //成功后的token_id
}
message SubscribeRequest { //定义事件流名称
    required string stream_name = 1;
}
service GrpcService { //定义gRPC 方法
    rpc SubscribeByStreamName(SubscribeRequest)
    returns (SubscribeReply) {} //订阅事件流
    rpc GetEventReport (GetReportRequest)
    returns (stream ReportEvent) {} //获取事件结果
}
```

(2) 业务模块proto 文件，支持Device、Ifmgr、IPFW、LLDP、Syslog 等多个业务模块的proto 文件，描述具体的业务数据格式。Device 模块数据的RPC 方法，其内容和含义如下：

```
syntax="proto2";
import "grpc_service.proto";
package device;
message DeviceBase { //获取设备基本信息结构定义
    optional string HostName = 1; //设备的名称
    optional string HostOid = 2; //sysoid
    optional uint32 MaxChassisNum = 3; //最大框数
    optional uint32 MaxSlotNum = 4; //最大slot 数
    optional string HostDescription = 5; //设备描述信息
}
message Device Physical Entities { //设
```


备物理实体信息

```
message Entity {
    optional uint32 PhysicalIndex = 1; //实体索引
    optional string VendorType = 2; //vendor 类型
    optional uint32 EntityClass = 3; //实体类型
    optional string SoftwareRev = 4; //软件版本
    optional string Serial Number = 5; //序列号
    optional string Model = 6; //模式
}
```

```
repeated Entity entity = 1;
```

```
}
```

service Device Service { //定义的RPC方法

```
rpc GetJsonDeviceBase ( DeviceBase )
returns ( grpc_service.GetJsonReply ) {} //获取设备基本信息
```

```
rpc GetJsonDevicePhysicalEntities
( DevicePhysicalEntities ) returns ( grpc_service.GetJsonReply ) {} //获取设备实体信息
```

4.2 ERSPAN 流分析

①将镜像报文封装为ERSPAN Type II的GRE报文，GRE报文如图4所示。

```
GRE header for ERSPAN Type II encapsulation (8 octets [34:41])
0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1
+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+
|0|0|0|1|0|0000|00000000|00000| Protocol Type for ERSPAN |
+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+
| Sequence Number (increments per packet per session) |
+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+
```

图4 ERSPAN 的 GRE 报文

②TCP会话传输的数据流大小。

· 请求方向流量：请求方向的FIN-ACK序列号-SYN序列号

· 应答方向流量：应答方向的FIN-ACK序列号-SYN-ACK序列号

③TCP报文在网络设备中进行三层转发时，IP首部的TTL字段会根据逐跳递减原理，先内层匹配内容，再排序内外层报文的TTL，之后再匹配识别完成路径还原；由采集器对每份数据流打时间戳，并根据时间戳信息计算TCP在设备间的逐跳传输延时计算；依据TCP内层报文的五元组信息，加上应用层的主机号和端口信息进行应用识别。

5 实验结果与分析

5.1 gRPC 配置

(1) 配置 Device (gRPC 服务器)。

开启gRPC 功能。

```
<Device> system-view
```

```
[ Device ] grpc enable
```

创建本地用户test，配置该用户的密码，授权用户角色为network-admin，可以使用的服务类型为http 2.0服务。

```
[ Device ] local-user test
```

```
[ Device-luser-manage-test ] password
simple 123456TESTplat&!
```

```
[ Device-luser-manage-test ] authorization-  
attribute user-role network-admin
```

```
[ Device-luser-manage-test ] service-  
type http 2.0
```

```
[ Device-luser-manage-test ] quit
```

(2) 配置 gRPC 客户端。

a. 在 gRPC 客户端安装gRPC 环境。

b. 获取.proto 文件 (该文件中已写入订阅 LLDP 事件的配置), 并通过protocol buffers 编译器生成特定语言 (例如Java、Python、C/C++、Go) 的执行代码。

c. 编写客户端程序, 调用上一步生成的代码。

d. 执行客户端程序, 登录到gRPC 服务器。

5.2 ERSPAN 配置

(1) 配置 Device A。

配置OSPF 协议。

```
[ DeviceA ] ospf 1
```

```
[ DeviceA-ospf-1 ] area 0
```

```
[ DeviceA-ospf-1-area-0.0.0.0 ] network  
10.1.1.0 0.0.0.255
```

```
[ DeviceA-ospf-1-area-0.0.0.0 ] network  
20.1.1.0 0.0.0.255
```

创建本地镜像组1。

```
[ DeviceA ] mirroring-group 1 local
```

配置本地镜像组1的源端口为 GigabitEthernet1/0/1, 目的端口为 GigabitEthernet1/0/2, 镜像报文的目的IP地址为40.1.1.2, 源IP地址为20.1.1.1。

```
[ DeviceA ] mirroring-group 1 mirroring-  
port gigabitethernet 1/0/1 both
```

```
[ DeviceA ] mirroring-group 1 monitor-  
port gigabitethernet 1/0/2 destination-ip  
40.1.1.2 source-ip 20.1.1.1
```

(2) 配置 Device B。

配置OSPF 协议。

```
[ DeviceB ] ospf 1
```

```
[ DeviceB-ospf-1 ] area 0
```

```
[ DeviceB-ospf-1-area-0.0.0.0 ] network  
20.1.1.0 0.0.0.255
```

(3) 验证配置。

显示Device A 上所有镜像组的配置信息。

```
[ DeviceA ] display mirroring-group all  
Mirroring group 1:
```

Type: Local

Status: Active

Mirroring port:

GigabitEthernet1/0/1 Both

Monitor port: GigabitEthernet1/0/2

Encapsulation: Destination IP address
40.1.1.2

Source IP address 20.1.1.1

Destination MAC address 000f-e241-
5e5b

5.3 分析

gRPC采集数据包括网络设备的实时资源信息、RDMA统计信息、RDMA告警信息等, 见表1。

ERSPAN技术采集网络TCP特征报文, 上传数据采集分析器 (如图5所示)。分析器通过TCP流分析技术, 实现如下功能分析: ①分析沿途交换机上报的TCP镜像报文可获得应用流量转发路径; 分析采集TCP报文时间戳可获得应用建立TCP连接时延及沿途交换机转发时延, 定位应用体验差是因为网络慢还是应用本身的问题; 分析TCP报文头可获取应用TCP连接持续时间及流量大小; ②根据TCP流生命周期的交互协议报文, 结合大数据分析算法, 实现TCP连接异常检测: TCP连接异常, TTL会话异常。

表 1 gRPC 采集的相关数据信息

数据大类	数据维度	数据项	备注
资源数据	整机	CPU占用率、内存占用率	采集频率1分钟级
	接口	收/发包数、收/发广播包数、收/发组播包数、收/发单播包数、收/发字节数	采集频率1分钟级
		收/发丢包数、收/发错包数	
RDMA统计信息	状态信息	ingress/egress丢包总量	①每个端口、每个queue ②采集频率1秒级
		ingress/egress buffer统计	
		headroom buffer统计	
RDMA报警信息	故障事件	ingress丢包	采集频率1秒级
		egress丢包	
		headroom buffer超限	
		egress buffer超限	



图 5 ERSpan 采集到的部分数据分析

6 结语

SDN可以通过控制器下发配置使得网络运维更为方便。本文致力于通过gRPC、ERSpan技术持续采集设备数据、日志数据、流量数据、拓扑数据等数据，通过对数据的统计分

析，实时洞察整网状态，结合SDN控制器的网络运行策略，得出网络运行过程中产生的问题，解决快速定位网络故障位置的问题，方便网络运维。本文的设计研究也存在一些不足之处：并未结合人工智能、大数据技术进行研

究,通过模型,利用历史数据训练,预防网络故障的出现;也未对通过网络发现的问题,自动修改SDN控制器的网络配置策略、自动解决网络问题进行研究。

参考文献

- [1] ZHANG Q Y, WANG X W, HUANG M, et al. Software Defined Networking Meets Information Centric Networking: A Survey [J]. IEEE Access, 2018, 6: 39547—39563.
- [2] MCKEOWN N, ANDERSON T, BALAKRISHNAN H, et al. Open Flow: enabling innovation in campus networks [J]. ACM SIGCOMM Computer Communication Review, 2008, 38 (2): 69—74.
- [3] Open Networking Foundation. OpenFlow switch specification, version 1.4.0: wire protocol 0x05: 2013 [S]. 2013.
- [4] Introduction to gRPC. <https://grpc.io/docs/what-is-grpc/introduction/>.
- [5] Cisco Systems' Encapsulated Remote Switch Port Analyzer (ERSPAN) draft-foschiano-erspan-00.txt. <https://tools.ietf.org/html/draft-foschiano-erspan-00>.
- [6] 左青云, 陈鸣, 赵广松, 等. 基于Open Flow 的SDN 技术研究 [J]. 软件学报, 2013, 24 (5): 1078—1097.