

doi:10.3969/j.issn.1673-9833.2014.05.017

一种基于NetMagic的网络拓扑发现算法

廖飞, 龙永新, 钟海

(湖南工业大学 计算机与通信学院, 湖南 株洲 412007)

摘要: 对NetMagic网络实验平台进行研究, 提出了一种基于NetMagic的网络拓扑发现算法。介绍了网络拓扑发现算法原理, 定义探测和响应报文在以太网帧中的封装格式, 在简单拓扑和复杂拓扑情况下对本算法进行仿真实验。仿真实验验证了算法的逻辑可行性和有效性。

关键词: 网络拓扑发现; 互联网技术; NetMagic

中图分类号: TP393

文献标志码: A

文章编号: 1673-9833(2014)05-0084-05

A Netmagic-Based Network Topology Discovery Algorithm

Liao Fei, Long Yongxin, Zhong Hai

(School of Computer and Communication, Hunan University of Technology, Zhuzhou Hunan 412007, China)

Abstract: Investigates the NetMagic network experiment platform, and proposes a network topology discovery algorithm based on NetMagic. Describes the principle of network topology discovery algorithm defines the encapsulation format of detection and response packets in an Ethernet frame, and simulates the algorithm in the cases of simple topology and complex topology. The simulated result verifies the logic feasibility and effectiveness of the algorithm.

Keywords: network topology discovery; internet technology; NetMagic

0 引言

随着新型网络技术研究的爆炸式发展, 现有网络实验平台的局限性越来越突出。一直以来, 路由器和高性能交换机的生产技术是“思科”“华为”“中兴”等国内外大型通信企业的核心技术, 其系统内部结构和处理流程一般不对外公开。商业网络设备厂商因考虑自身利益而导致的封闭性, 已成为网络技术创新和发展的最大障碍。

互联网科学作为一门以实验为基础的科学, 是在真实的网络环境中进行实验并收集数据, 是进行网络技术创新最有效和最具说服力的手段。实验验证是互联网研究的基本方法, 而现有网络和实验平台缺乏灵活的实验机制。因此, 现有互联网技术的演进和新一代互联网技术的研究创新, 都需要一个开放和灵活的实验验证平台。

NetMagic^[1-2]是近年来推出的面向网络技术研究

收稿日期: 2014-07-20

基金项目: 国家自然科学基金资助项目(61170102, 61350011), 湖南省自然科学基金资助项目(11JJ3070), 湖南省自然科学基金重点课题基金资助项目(12JJ2036), 湖南省自然科学基金面上课题基金资助项目(14JJ2115), 湖南省教育厅科研基金资助项目(12A039), 湖南工业大学研究生创新基金资助项目(CX1404)

作者简介: 廖飞(1988-), 男, 湖南常德人, 湖南工业大学硕士生, 主要研究方向为网络嵌入式系统,

E-mail: loveuliaoifei@126.com

通信作者: 龙永新(1966-), 男, 湖南益阳人, 湖南工业大学副教授, 硕士生导师, 博士, 主要从事智能信息处理和数字图像处理方面的教学与研究, E-mail: 1525127621@qq.com

和教学的一种可编程创新实验平台。因其具有灵活性强、可重构等特性,使得基于 NetMagic 的开发和研究日益受到广泛关注。其可适用于很多领域,如网络教学、专用网络建设、网络科研等。NetMagic 实验平台采用的是新型部分可编程交换机体系结构技术,将商用以太网交换芯片和大容量 FPGA (field-programmable gate array) 有机结合,能提供高密度下报文线速交换转发能力和丰富的工作模式配置,有效支持用户自定义报文处理逻辑集成。而软件定义网络 (software define networking, SDN) 是 Emulex 一种新型网络创新架构。其核心技术是将网络设备控制面与数据面分离,从而实现了网络流量的灵活控制,为核心网络及应用的创新提供了良好的平台。

网络拓扑发现是互联网技术研究的核心内容之一。网络拓扑发现是指通过网络中各元素(路由器、交换机、网关和终端主机等)和通信线路确定网络元素之间的互连关系。因此,针对小规模局域网,本文对 NetMagic 网络实验平台进行研究,提出了一种基于 NetMagic 的网络拓扑发现算法。本文以软件定义网络思想的控制和转发分离架构的设计理念为基础,设计算法。通过对两种拓扑结构的仿真实验,验证了本算法的逻辑可行性和有效性。

1 SDN 原理介绍

软件定义网络由于能支持新型网络协议和新型网络体系结构,而得到广泛发展^[3]。在学术界,美国 GENI、Internet2^[4]、欧洲 OFELIA 和日本的 JGN2plus^[5] 先后开展了对 SDN 的研究和部署, IETF (internet engineering task force), ITU (international telecommunication union), ETSI (european telecommunications standards institute) 等标准组织也开始关注 SDN, 讨论 SDN 在各自领域可能的发展场景和架构应用。

图 1 为软件定义网络原理图。图中,网络控制平面与数据转发平面相分离。网络控制逻辑和状态被转移到一个集中控制器上,因此,网络控制平面 (SDN 控制器) 能以一个全局网络的视野去管理和控制网络。数据转发平面 (SDN 交换机) 只专注于报文的转发。SDN 控制器通过配置 SDN 交换机中的规则表来控制交换机的行为,和 SDN 交换机之间通过特定的通信协议传递消息。也就是说,SDN 控制器相当于人的大脑,SDN 交换机类似于人的手脚和肌肉,整个思想活动全由 SDN 控制器控制,SDN 交换机根据 SDN 控制器下发的特定指令执行相应的动作行为,其行为完全是傻瓜式的状态^[6-7]。

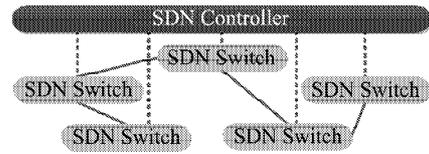


图1 软件定义网络原理图

Fig. 1 Schematic diagram of software defined network

本文所提算法就是采用上述思想。通过 SDN 控制器来定义 SDN 交换机的存储和行为。在网络拓扑发现过程中,只需对 NetMagic 平台内部的 FPGA 编程,将其配置成 SDN 交换机即可(如何通过 FPGA 编程将硬件配置成 SDN 交换机在此不详细讲述)。

2 网络拓扑发现算法设计

网络拓扑发现算法中的探测过程采用主动探测,发现过程工作在数据链路层和网络层之上。主动探测是指探测节点为获取拓扑信息主动向网络中发送路径探测包,并通过收集到的返回信息,分析网络的拓扑结构,最终形成网络拓扑。对于小规模局域网来说,一般是数据链路层的拓扑发现。其主要利用 ARP (address resolution protocol) 协议。本课题组采用自主开发的一套 NMAC (NetMagic access and control protocol) 协议,将探测信息封装在以太网帧中,而后在软硬件之间进行通信。

本文以 SDN 思想的控制和转发分离架构的设计理念为基础来设计算法。以图 2 所示的实验环境为例,对算法进行描述。本实验共有 3 台 PC 机和 4 台 NetMagic 网络实验平台。1 台 PC 机充当 SDN 控制器,主要功能为向 NetMagic 下发拓扑探测报文、处理生成的响应报文和生成网络拓扑结构。另外 2 台 PC 机作为连接在网络中的终端设备,目的在于模拟网络中的真实环境,其是否接入网络环境,不影响整个实验过程。通过硬件逻辑设计,将 4 台 NetMagic 配置成 SDN 交换机,其对控制器下发的拓扑探测报文进行处理,每个交换机的功能和行为完全相同。

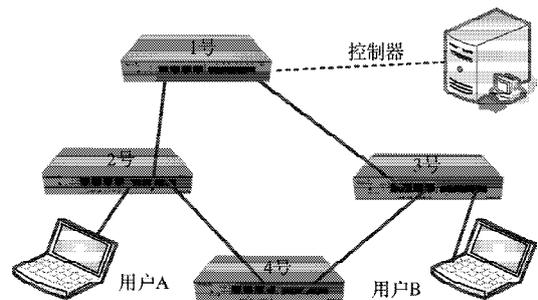


图2 网络拓扑发现实验环境

Fig. 2 The network topology discovery experiment environment

网络运行时, SDN 交换机以报文转发为基础, 内部并不含 CPU, 不运行任何操作系统。对于网络拓扑发现的应用, 还需要一台外部主机对 NetMagic 的 FPGA 进行编程配置。外部主机利用自主开发的 NMAC 协议, 通过以太网报文的形式传到 NetMagicFPGA 的管理配置接口, 然后由 NetMagicFPGA 对这些命令和配置报文进行解析, 改变 NetMagic 的转发行为。

网络拓扑发现算法的具体流程如下:

1) 网络拓扑发现开始。SDN 控制器首先向 1 号交换机发出跳数为 0 的拓扑探测报文, 1 号交换机收到跳数为 0 的探测报文后, 生成响应报文并返回至控制器。

2) 根据返回报文, SDN 控制器配置目的 Mac 地址表, 并将其发送给 1 号交换机。

3) SDN 控制器发送一个跳数为 1 的拓扑探测报文, 1 号交换机发现报文的跳数为 1, 将该报文跳数减 1 并附上自己本机 ID 号和输出端口号后广播。

4) 交换机 2 号、3 号收到跳数为 0 的报文, 生成响应报文并返回至控制器。SDN 控制器收到 2 号、3 号交换机的返回报文, 将配置目的 Mac 地址表发送至 2 号、3 号交换机。

5) SDN 控制器发送一个跳数为 2 的拓扑探测报文, 1 号交换机判断报文跳数为 2, 对报文跳数减 1 后直接进行广播转发; 2 号、3 号交换机收到跳数为 1 的报文, 其发现报文跳数为 1, 将报文跳数减 1 并附上自己本机 ID 号及输出端口号后广播。

6) 交换机 4 号收到跳数为 0 的报文, 将响应报文返回至 SDN 控制器。SDN 控制器收到 4 号交换机的返回报文后, 将配置目的 Mac 地址表发送至 4 号交换机。

7) SDN 控制器再发送一个跳数为 3 的拓扑探测报文。各交换机重复上述工作, 最后控制器发现没有返回的响应报文, 整个网络拓扑发现过程结束。

下面对本算法作如下几点说明:

1) 广播是指向连接交换机的所有活跃端口发送报文(接收到此报文的输入端口除外, 活跃端口指的是已经连通且能够通信的端口)。

2) 根据算法设计, 需要定义 2 类报文, 并将 NMAC 命令进行以太网报文封装。报文定义和封装格式如下。

定义 1 拓扑探测报文是目的 Mac 地址全为 1, Type 类型为 0x09, 协议号为 253 的 NMAC 报文。

探测报文在以太网报文中的封装格式见图 3。



图3 探测报文封装格式

Fig. 3 The detecting packet message encapsulation format

定义 2 响应报文是携带本机 ID 号, Type 类型为 0x10, port 为收到跳数为 0 的输入端口号, Flag 为 SDN 交换机是否连接其它设备标识的 NMAC 报文。

响应报文在以太网报文中的封装格式见图 4。



图4 响应报文封装格式

Fig. 4 The response packet message encapsulation format

在图 3 和图 4 中, 加颜色部分为 NMAC 报文, 规定 NMAC 协议号为 253, 从图中可以看出每个字段的位宽。

3 仿真分析

Jade (Java agent development framework) 是一种人工智能, 其 agent 可以用来模拟交换机之间的逻辑关系。为论证本拓扑发现算法的逻辑可行性和有效

性, 利用 Java 语言构造了不同的拓扑关系进行仿真测试。

3.1 仿真方案设计

仿真方案如下。

1) 交换机的功能有: 记录探测报文从“父节点”(上一个节点)进来的端口号(一般为1号端口); 当接收到跳数为0的探测报文时, 发出响应报文。

2) 初始化。初始化交换机的层数为一个充分大的值(如256)。若收到的探测报文中的 Req_Seq/Ack_Seq 域的值小于自己当前的层数, 则更新交换机的层数, 并记录报文进入的端口号, 该端口将作为自己通往“父节点”的端口。

3) 响应机制。

①当交换机收到的探测报文跳数大于1时, 将跳数减1, 然后向所有活跃端口(除了收到报文的端口)转发报文。

②当交换机收到的探测报文跳数等于1时, 构造新的探测报文, 将本交换机的ID号和Port端口号添加到新的探测报文当中, 并将跳数减1, 然后向所有活跃端口(除了收到报文的端口)转发出报文。

③当交换机收到的探测报文跳数等于0时, 分为3种情况:

情况一 Req_Seq/Ack_Seq 中的值大于自己当前层数, 则层数加1, 丢弃报文; 否则跳转到情况二。

情况二 构造响应报文。将收到的报文添加 Dest_ID 号和 Dest_Port 端口号, 并将自己的ID号和收到报文的端口号分别写入 Src_ID 和 Src_Port 中, 然后将响应报文转发给“父节点”。

情况三 当收到的报文为响应报文, 直接向“父节点”转发, 直到返回给控制器。

3.2 仿真结果与分析

为了验证了所提算法的逻辑可行性和有效性, 本文进行了2组测试实验, 即算法在简单拓扑结构和复杂拓扑结构下的运行情况。简单的网络拓扑结构图如图5所示。通过Java语言编写算法程序, 从各个节点收集拓扑信息, 归类整理, 再调用画图软件绘出拓扑结构图。从图5可以看出, 算法仿真得到的拓扑结构图和本文真实的实验环境完全相同。

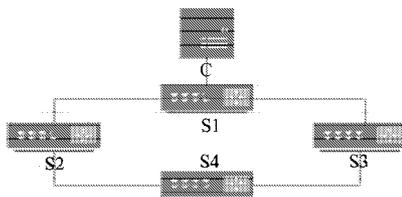


图5 拓扑结构图

Fig. 5 The topology diagram

图6是图5的仿真结果过程图。虚拟机软件VMware平台上有个sniffer可以进行全局抓包, 能够看到控制器与各个交换机之间的报文交换过程。从图中可以看出, 报文交换过程完全与本文算法预期的报文发送过程一样。该仿真结果验证了所提算法的逻辑可行性和有效性。

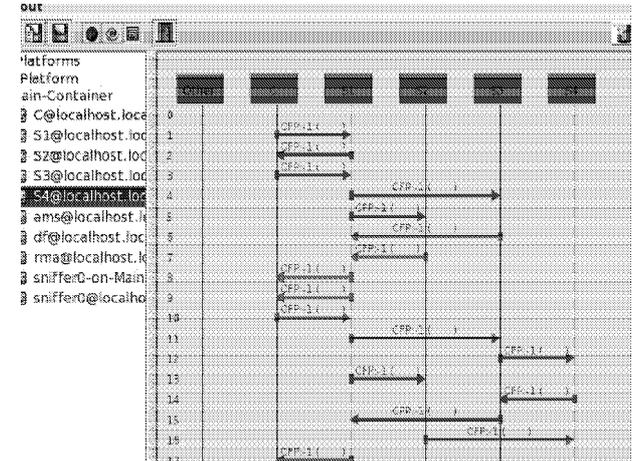


图6 仿真结果过程图

Fig. 6 The simulation results

图7~9是3种复杂情况下的拓扑结构图。从3个拓扑结构中都可以得到本机ID号、本机连父的端口号、父ID号和父连本机的端口号这些关键的拓扑信息, 通过这些信息可以获得全网交换机和控制主机的拓扑图。进一步验证了所提算法的逻辑可行性和有效性。

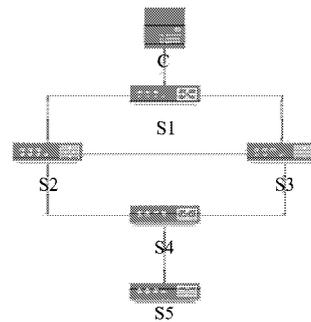


图7 拓扑结构图a

Fig. 7 The topological diagram a

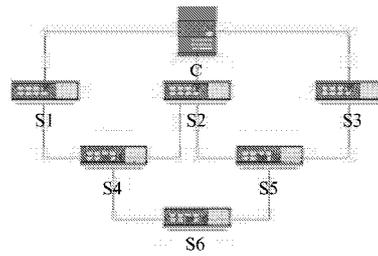


图8 拓扑结构图b

Fig. 8 The topological diagram b

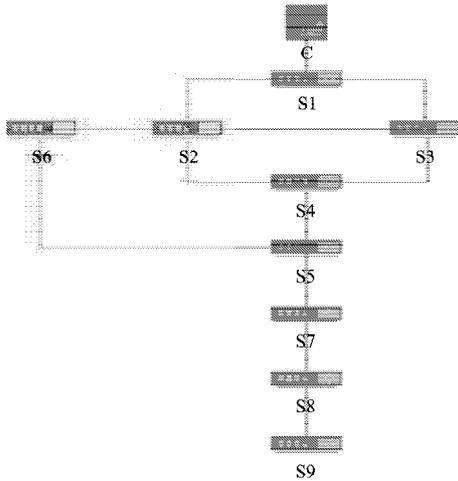


图9 拓扑结构图c

Fig. 9 The topological diagram c

4 结语

本文以SDN思想的控制和转发分离架构的设计理念为基础,提出了一种新型的基于NetMagic的网络拓扑发现算法。先对算法进行了描述,定义了探测和响应报文在以太网帧中的封装格式,并通过软件对算法进行仿真实验,验证了所提算法的逻辑可行性和有效性。实验结果表明:该算法通过本机ID号、本机连父的端口号、父ID号和父连本机的端口号关键四元组拓扑信息,可以主动探测获得全网交换机和控制器主机之间的拓扑结构视图。

参考文献:

[1] Li Tao, Sun Zhigang, Jia Chunbo, et al. Using NetMagic to Observe Fine-Grained Per-Flow Latency Measurements

[C]//Proceedings of the ACM SIGCOMM 2011. Toronto: ACM, 2011: 466-467.

- [2] Naous J, Gibb G, Bolouki S, et al. NetFPGA: Reusable Router Architecture for Experimental Research[C]//Proceedings of the ACM Workshop on Programmable Routers for Extensible Services of Tomorrow (PRESTO'08). Seattle: ACM, 2008: 1-7.
- [3] Nascimento M R, Rothenberg C E, Salvador M R, et al. Virtual Routers as a Service: The RouteFlow Approach Leveraging Software-Defined Networks[C]//Proceedings of the 6th International Conference on Future Internet Technologies(CFI). Seoul: ACM, 2011: 34-37.
- [4] 张 钧, 黄亚楼. 第二代因特网(Internet2)的技术与发展[J]. 计算机应用, 1999, 19(12): 22-24.
Zhang Jun, Huang Yalou. The Technique and Development of Internet2[J]. Computer Applications, 1999, 19(12): 22-24.
- [5] 梁学军, 林昭文, 马 严. 未来互联网实验平台[J]. 计算机学报, 2013, 36(7): 1365-1366.
Liang Junxue, Lin Zhaowen, Ma Yan. Future Internet Experiment Platform[J]. Chinese Journal of Computers, 2013, 36(7): 1365-1366.
- [6] 王志刚, 王汝传, 王绍棣, 等. 网络拓扑发现算法的研究[J]. 通信学报, 2004, 25(8): 36-43.
Wang Zhigang, Wang Ruchuan, Wang Shaodi, et al. Research on Network Topology Discovery Algorithm[J]. Journal on Communications, 2004, 25(8): 36-43.
- [7] 杨国正, 陆余良, 夏 阳. 计算机网络拓扑发现技术研究[J]. 计算机工程与设计, 2006, 27(24): 4710-4712.
Yang Guozheng, Lu Yuliang, Xia Yang. Research on Computer Network Topology Discovery Technology[J]. Computer Engineering and Design, 2006, 27(24): 4710-4712.

(责任编辑: 邓 彬)