

一. OpenFlow 阅读报告

《OpenFlow: Enabling Innovation in Campus Networks》是一篇 2008 年发表于 SIGCOMM^[1] 上的文章，Nick McKeown 教授等人首次提出了 OpenFlow 的概念，并列举了 OpenFlow 的几大应用场景。作者在论文中提出了一种供研究人员在网络中采用实验协议的方法，即 OpenFlow。这是一种标准化网络接口，基于以太网交换机，并且拥有一个内部流表和用来增减流的条目的标准化接口。作者企图将 OpenFlow 应用到应用商的交换机产品中，并将其部署在大学校园中。作者认为 OpenFlow 是一种实用的折衷方案：既允许研究人员以线性速率和高端口密度的统一方式在异构交换机上进行实验，又可以让供应商不公开交换机的内部工作方式。

接下来我将从**可编程网络的必要性**、**OpenFlow 交换机的基本原理**、**使用 OpenFlow 的过程**、**如何部署 OpenFlow** 等方面概述该论文。

1. 可编程网络的必要性

在当今社会，网络承载着现实社会大部分信息传输的功能，任何网络创新对现实世界都会产生相当大的影响。但因其设备和协议的巨大安装基础，以及试验生产流量带来的成本消耗，试验者对网络创新的广泛部署缺乏信心，这导致了网络基础设施的逐步僵化。于是，科学家们转向研究可编程网络，通过使用可编程交换机和路由器，同时处理多个来自互不相连的实验网络的数据包。

虚拟化的可编程网络允许研究人员在开放的可编程虚拟化平台上部署新的协议，测试创新想法的正确性。这降低了新想法的实现障碍，提高网络基础设施的创新速度。

2. OpenFlow 交换机的基本原理

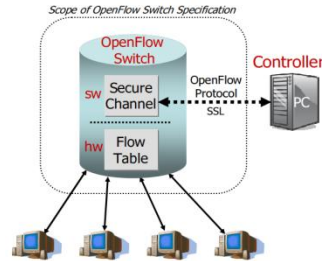
大多数现代以太网交换机和路由器中都包含流表，而 OpenFlow 就是采用流表的思想，在交换机和路由器中运行一组通用函数，并提供一种开放式协议在其中进行编程。在 OpenFlow 交换机和路由器中，存在两种流：一种是普通的日常使用的信号，叫**生产流**；另一种用来给研究人员尝试新的路由协议、安全模型、寻址方案甚至是 IP 的替代方案，叫**研究流**。两种流根据各自数据包遵循的路由和接收的处理选择路径，互不干扰，前者供线路的正常通信，后者用于科学研究。OpenFlow 交换机至少包含 3 个部分：

- 一个**流表**，并且在每个流的入口都有一个与之关联的方法，来告诉交换机如何处理这些流；
- 一个**安全通道**，用来连接交换机和远程控制器并传输命令和数据报；
- **OpenFlow 协议**，为交换机和远程控制器提供开放的标准。

为了区分交换机是否加入 OpenFlow 接口，将交换机分为不支持普通第 2、3 层处理的**专用 OpenFlow 交换机**和支持商用的**通用 OpenFlow 的交换机**。

2.1 专用 OpenFlow 交换机

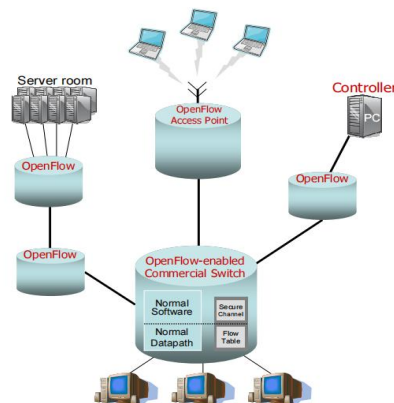
专用 OpenFlow 交换机按远程控制器的定义在端口之间转发数据，包的转发完全由控制器决定（如图）。



在这种情况下，流可以是 TCP 连接或来自不同端口的数据包，每个流入口的操作页较为简单：将此流的数据包转发到给定的端口、封装此流的数据包并转发给控制器。流表中的条目也有 3 个部分：首部，用来定义流；操作部分，用来定义如何转发数据包；统计部分，用来跟踪每个流的数据包数量和字节数。

2.2 通用 OpenFlow 交换机

通用 OpenFlow 交换机将 OpenFlow 作为接口融入日常的交换机中，即在其中添加流表、安全通道和 OpenFlow 协议。通过硬件的复用，将所有流表交给控制器管理，以实现在现有的生产网络中加入可编程虚拟化通道来验证改进的可行性（如图）。



为了保障普通信道的正常通信，必须将 OpenFlow 接口与普通信道分离。作者提出了两种方法：一种是在原先的基础上加入正常处理管道转发流的数据包；另一种是为实验流量和生产流量定义单独的 VLAN 集。

3. 使用 OpenFlow 的过程

在这一部分，作者以研究人员 Amy 在一个 OpenFlow 交换机网络中测试新创立的协议而不改变任何终端主机软件为例，阐述了 OpenFlow 交换机如何测试研究流的同时不影响生产流：在 Amy 的实验中，添加一个流入口，对从她的桌面电脑进入 OpenFlow 交换机网络的流量使用 Amy-OSPF 路由协议，Amy 的电脑与网络中某个特定 OpenFlow 交换机的端口相连，将所有从此端口进入网络的流量定义为一个流，并在这个特定 OpenFlow 交换机的流表中添加流的表项，流表项的行为指示将此流的包全部封装并转发到 Amy-OSPF 控制器。Amy-OSPF 控制器收到包表示网络中产生了新流，它首先为此流选择路由，然后在所选路由经过的所有交换机中添加新流的流表项。该流的后续的包将不再转发给 Amy-OSPF 控制器，而是沿着已经由流表确定好的路由快速的线速转发。

随着实验的进行，也许有人会对 Amy-OSPF 控制器的性能、可靠性和泛化性存在质疑。其实在以太网原型中，使用一个简单的流交换机和中央控制器的情况下就可以足够应付一整个校园的流量需求。因此我们有理由相信 OpenFlow 交换机有相当好的性能。同时，通过使一个控制器变得无状态，允许在多个独立的设备上实现简单的负载平衡，就可以实现泛化性

和冗余性。

后面作者又介绍了生产网络中的实验以及网络管理和访问控制、虚拟网络、移动无线 VOIP 客户端、无 IP 网络、处理数据包等具体实例，此处不加赘述，详见原文。

4. OpenFlow 联盟

文章第 4 部分介绍了 OpenFlow 联盟的宗旨、交换机规范、许可模式等，此处略过。

5. 如何部署 OpenFlow

作者相信 OpenFlow 有着广阔的市场前景，目前也在和交换机、路由器生产商合作，商讨在不改变原有硬件的基础上添加 OpenFlow 特性。并且还在斯坦福大学的计算机科学和电气工程系部署了大型的 OpenFlow 网络，生产流量和实验流量在网络管理员的控制下被隔离在不同的虚拟网络上互不影响。

总的来说，Openflow 的提出将以太网的设计一般化，允许研究人员以统一的方式在异构交换机和路由器上进行实验，而不需要供应商公开其产品的内部工作原理，也不需要研究人员编写针对特定供应商的控制软件。将传统网络设备的数据转发和路由控制两个功能模块相分离，通过集中式控制器以标准化接口对各种网络设备进行管理和配置，从而为网络资源的设计、管理和使用提供更多的可能性，推动了网络的革新与发展。以至于后来基于 OpenFlow 为网络带来的可编程的特性，Nick McKeown 教授和他的团队进一步提出了轰动一时的软件定义网络 SDN，在网络中实现了软硬件的分离以及底层硬件的虚拟化，为网络的发展提供了相当良好的发展平台。

[1]SIGCOMM: 全称是 Special Interest Group on Data Communication，即数据通信专业组，是 ACM 组织在通信网络领域的旗舰型会议，也是目前国际通信网络领域的顶尖会议，由 ACM SIGCOMM 组织举办。

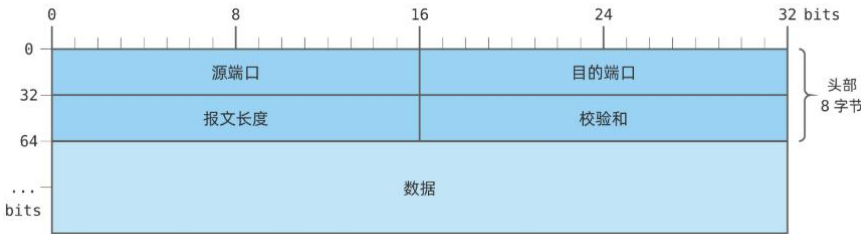
二. RFC 768 阅读报告

在《计算机网络》课程学习中我们知道，Internet 协议支持无连接的传输协议，即用户数据报协议（UDP，User Datagram Protocol）。UDP（即用户数据报协议，下同）为应用程序提供了一种无需建立连接就可以发送封装的 IP 数据包的方法。本篇论文讲述的 RFC 768 就是描述 UDP 的相关数据格式与协议规范。

1. 介绍

UDP 是在一组互连的计算机网络中提供分组交换的计算机通信数据报模式。该协议假设互联网协议被用作底层协议，为应用程序提供了一个最简单的协议机制向其他程序发送消息的过程。UDP 协议是面向报文的，不能保证交付和重复保护。如果需要有序可靠的传输数据，可以使用传输控制协议 TCP。

2. 格式



UDP 数据报格式如上图所示，由首部和数据部分组成。其中首部包含 16bit 的源端口、16bit 的目的端口，16bit 的报文长度和 16bit 的校验和，数据部分长度可变，但不足 32bit 的需要补全。下面将依次介绍各字段含义与规范：

- 源端口：这是一个可选字段，当使用该字段时，它表示发送该数据报的进程的端口，在需要对方回信的情况下表示响应端口。如果不使用该字段，则全部赋为 0。
- 目的端口：指向网络通信地址上下文的端口，在终点交付报文时必须使用。
- 长度：表示这个 UDP 数据报的长度，单位是八进制，即 2 字节。这个长度包括首部和数据。由于首部为 16 字节，因此该长度的最小值为 8。
- 校验和：为了检测 UDP 用户数据报在传输中是否有错，在首部加入了校验和字段。这个校验和是来自 IP 首部、UDP 首部、数据的伪首部的二进制反码求和，需要时还需在结尾补 0 以保证长度为 16bit。

对于 UDP 伪首部，理论上包含源地址、目标地址、协议以及 UDP 长度。在校验和中加入这些信息可以防止数据报在传输中产生差错。该校验和过程与 TCP 中使用的过程相同，此处不加赘述。当目的端口收到 UDP 数据报后，计算出校验和为零，则说明传输无误。

3. 应用层接口

作者在阐述 UDP 数据报接口的时候指出，用户接口必须允许以下操作：

- 创建新的接收端口；
- 接收新创建的端口返回的数据、端口号和端口地址；
- 发送数据并指定源端口和目标端口及其地址；

4. 网络层接口

UDP 数据报必须能够从 IP 数据报的首部中得知源主机和目的主机的 IP 地址以及协议字段。一个可行的 UDP-IP 层间接口需要携带整个 IP 数据报，包括整个 IP 数据报的首部。这个接口能够保障 UDP 数据报传输完整的 IP 数据报。网络层会验证一些字段的可行性，并检验 IP 数据报的校验和。

在文章的最后，作者指出 UDP 协议的主要用途是域名服务器 IDN 和简单文件传输协议 TFTP。并规定了 UDP 协议的序号为 17。

总的来说，UDP 是一个无连接、面向报文的传输协议，没有 TCP 的握手、确认、重传、滑动窗口、拥塞控制等机制，因此在传递数据时速度更快。但也正因如此，UDP 缺少了 TCP 的可靠机制，传输可靠性较低，在网络质量不好的情况下很容易丢包。