

计算机网络课程报告

题 目：SDN驱动下的MPLS探讨

学生姓名：周雨婧

学 号：20141000826

专 业：计算机科学与技术

班 级：191143

指导教师：陈喆

2018年6 月

目录

[第一章 多协议标签交换 3](#_Toc518376754)

[1.1 MPLS概述 3](#_Toc518376755)

[1.1.1 功能 3](#_Toc518376756)

[1.1.2 分层协议 3](#_Toc518376757)

[1.1.3 转发机制 3](#_Toc518376758)

[1.2 MPLS工作原理 5](#_Toc518376759)

[1.2.1 标签 5](#_Toc518376760)

[1.2.2 工作过程 9](#_Toc518376761)

[1.2.3 优势 13](#_Toc518376762)

[第二章 软件定义网络 14](#_Toc518376763)

[2.1 SDN概述 14](#_Toc518376764)

[2.2 SDN架构 15](#_Toc518376765)

[2.2.1 架构特征 15](#_Toc518376766)

[2.2.2 各层功能 16](#_Toc518376767)

[2.3 SDN优势 17](#_Toc518376768)

[2.3.1 简单高效 17](#_Toc518376769)

[2.3.2 低成本 18](#_Toc518376770)

[第三章 基于SDN架构的MPLS 19](#_Toc518376771)

[3.1 结合MPLS部署SDN服务链 19](#_Toc518376772)

[3.1.1 技术趋势 19](#_Toc518376773)

[3.1.2 MPLSDN 19](#_Toc518376774)

[3.2 结合MPLS部署SDN服务链 21](#_Toc518376775)

[3.3 总结 22](#_Toc518376776)

[第四章 总结 23](#_Toc518376777)

# 第一章 多协议标签交换

## 1.1 MPLS概述

### 1.1.1 功能

多协议标签交换（MPLS）是一种用于快速数据包交换和路由的体系，它为网络数据流量提供了目标、路由地址、转发和交换等能力。更特殊的是，它具有管理各种不同形式通信流的机制。

MPLS利用标记进行数据转发，当分组进入网络时，要为其分配固定长度的短的标记并将标记与分组封装在一起，在整个转发过程中，交换节点仅根据标记进行转发。其具有“多协议”特性，对上兼容IPv4、IPv6等多种主流网络层协议，将各种传输技术统一在一个平台之上；对下支持ATM、FR、PPP等链路层多种协议，从而使得多种网络的互连互通成为可能。它的拓展性更可以满足运营商网络的后续发展。

在功能方面，使用MPLS能够实现的基础功能有，打标签，生成标签表，标签交换，路由技术，拓展功能有MPLS VPN、QoS、Multicast等。

### 1.1.2 分层协议

关于MPLS究竟属于分层协议体系结构的哪一层，存在很多争论。由于MPLS头部通常位于分组的第三层和第二层头部之间,它有时被称为2.5层协议，有些人认为，由于IP分组被封装在MPLS头部内，因此MPLS必须在IP“之下”，使它成为第二层协议，而另一些人认为，由于MPLS的控制协议很大一部分是IP使用的同样协议(MPLS使用P路由选择协议和IP編址)因此MPLS必须和P在同一层(即第3层)，严格的分层观点可能很难与现实。通常，标签交换依据标签表一次查找一次转发，基于2.5层转发流量，路由技术则是路由表多次查找一次转发（递归方法），基于3层转发流量。

### 1.1.3 转发机制

在无连接的网络层，一个报文在从源到目的的过程中，每个路由器都会独立的进行转发控制。也就是说，每个路由器都要分析报文的首部，而且每个路由器都要运行网络层路由算法。每个路由器独立的根据报文首部分析的结果和路由算法的结果选择报文的下一跳。

报文首部中包含的信息远远超过用于选择下一跳所需的信息。为一个报文选择下一跳可以被看作是两个功能的组合。首先是将所有的可能收到的报文划分成为一系列的"转发等价类（FECs）"。在MPLS网络中经过相同的LSP，完成相同的转发处理的一些数据分组，这些数据分组具有某些相同的特性，FEC的划分通常依据网络层的目的地址前缀或是主机地址。

其次，将每个等价类映射到一个下一跳上去。映射到同一个等价类中的不同的报文，从转发的观点来看是没有差别的。属于同一个FEC并且是从同一个节点来的所有报文将在同一个通道中传送。（或者，如果使用多路径路由的话，他们将使用这个FEC对应的一套路径。）

在传统的IP转发中，一个路由器在自己的路由表中为两个不同的报文进行路由查找的时候，如果发现他们具有相同的地址前缀X，并且这个X是对每个报文的目的IP地址进行最长匹配的结果，则认为他们属于相同的FEC。当一个报文在网络中传送的过程中，每一跳都重新检查该报文的首部，并且将其分配的一个FEC中。

在MPLS网络中，一个报文被分配到一个FEC的动作只会发生在这个报文进入MPLS网络的时候。报文所属的FEC使用一个定长的值来编码，就是所谓的标签。当一个报文被转发给它的下一跳的时候，标签和报文一起发送，也就是说这个报文转发之前被打上了标签。

在后续的路由器的处理中，不需要再对报文的网络层首部进行分析。报文的标签被当作索引对转发表进行查找，得到下一跳和一个新的标签。老的标签被新的标签替换，报文被转发给它的下一跳。



图1 MPLS体系结构

## 1.2 MPLS工作原理

### 1.2.1 标签

**（1）标签定义**

标签是一个短的、具有固定长度、仅在相邻LSR之间有意义、用来标识和区分转发等价类FEC的标志，被加入到一个报文上的标签表示这个报文所属的转发等价类。

一般来说，一个报文到一个FEC的指定都是基于或者部分基于这个报文的网络层地址。但是，这个标签并不是报文的网络层地址的简单编码。

标签共有4个域：

Label：标签值字段，长度为20bits，用来标识一个FEC。

Exp：3bits，保留，协议中没有明确规定，通常用作CoS。

S：1bit，MPLS支持多重标签。值为1时表示为最底层标签。

TTL：8bits，和IP分组中的TTL意义相同，可以用来防止环路。

http://www.h3c.com.cn/res/200905/04/20090504_758761_image002_606207_30003_0.png

图2 标签的封装结构

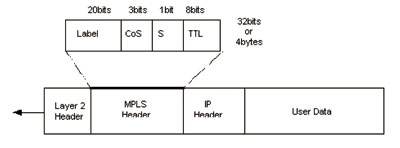


图3 MPLS标签栈头

标记报文指的是被添加了标签的报文，在某些情况下，标签是存在于一个封装的首部，这个首部就是为了标签的目的而添加的。在另外一些情况下，标签存在于已经存在的数据链路层或者网络层首部，只要这里有相应的字段可以完成这个目的。所使用的编码技术对于编码标签的实体和解码标签的实体必须是一致的。

**（2）标签映射**

转发等价类(FEC) 是一个用在多协议标签交换(MPLS)的术语，其用来描述一系列分组拥有相似或同样的特征，以同样的方式被运送。从转发的行为来看，这些分组具有相同的转发属性。FEC的种类分为两种，一种是单目广播分组，其目的地址均与一个IP地址前缀相匹配，另一种FEC是分组的源与目的地址都相同的一组分组。

设A和B为标记交换路由器，他们会约定当A发送一个报文给B的时候，当且仅当这个报文是一个特殊的FEC的一员，A将这个报文加上标签L。也就是说，他们将就在报文从A传送到B时标签L和等价类F的“绑定”达成共识。作为这种共识的结果，L成为A的表示等价类F的“输出标签”，同时成为B的表示等价类F的“输入标签”。

任何不属于从A发送到B的报文，即使它属于转发等价类，也不能用标签L表示。L是一个任意值，它到F的绑定是针对A和B的，是一个本地的操作，而不是针对整个的标记交换路径LSP或者MPLS网络的。

有时候，对B来说，分辨一个收到的报文的标签L是否是由A添加的，是一件非常困难的事情。在这种情况下，Rd必须保证从标签到等价类的绑定是一对一的,也就是说，B绝对不能和A1协定将L绑定到等价类F1，同时又和A2协定将L绑定到F2。除非B可以保证在收到一个标签L的时候，能够区分他们是由A1还是A2发送的。每个LSR都有责任确保他们的解释输入标签是唯一的。

**（3）标签赋值**

mpls标签交换的本质是将网络层的路由和数据链路层的标签进行绑定，并将标签绑定信息在标签交换路由器之间进行转发，这个过程也称为标签赋值。常见的标签赋值有：基于拓扑的控制业务量驱动；基于请求的控制业务量驱动和数据业务量驱动。

第一种是标签赋值对应于正常的路由协议控制业务处理，当lsr处理ospf或者bgp的路由更新时，一方面修改其转发表中的条目，一方面给这些条目分配相应的标签。只要有一条路由存在，网络就预先完成标签赋值，这样转发时就没有标签建立延时。

第二种是标签赋值与基于正常的请求的控制业务量处理相对应，他所对应的协议是rsvp，当lsr处理rsvp时，一方面修改其转发表中的条目，一方面给这些条目分配标签。这种方案要求应用程序事先提出使用标签请求和流规范，以得到标签，它也是根据已存在的路由预先完成标签赋值，没有标签建立延时。

第三种是数据业务量驱动。到达lsr的数据流量“触发“标签赋值和标签分发。此方案中，标签赋值和分发带来的开销是业务流量的正比例函数，存在与标签赋值相关的时延，如果想要将特定的标签分配给特定的网络资源以支持特定的网络程序时，就需要用数据驱动方式。

**（4）标签编码**

MPLS标签封装是基于atm交换机的，atm交换机是利用输入端口和输入vci/vpi作为查找交叉连接表的索引，从交叉连接表里获取输出端口和输出vci/vpi，以完成交换。若通过合适的编码方式将一个或几个标签放入vci/vpi域内，就可以将atm交换机改造为atm－lsr。一般说有三种方式将标签编入信元中：

1. SVC编码。用vci/vpi域对处于栈顶的标签进行编码，将LDP看作SVP，将LDP看作是ATM的信令协议，但是ATM－LSR不能对标签堆栈进行压栈和弹栈操作；
2. SVP编码。用VPI域对栈顶标签进行编码，若是标签堆栈中有第二层标签，则用VCI域对第二层标签进行编码，这种方案较前一种方案的优势是支持ATM的vp交换。
3. SVP多点编码。同样用VPI域对栈顶标签进行编码，用VCI域的一部分对第二层标签进行编码，用VCI域的其他部分指明LSP的入口标识。这种方案可以让ATM交换机支持多点到点的vp。

**（5）标签分发协议**

标签分发协议LDP（Label Distribution Protocol）是 MPLS 体系中的一种控制协议，负责FEC的分类，标签的分配以及LSP的建立和维护等一系列操作。

在 MPLS 网络中，两个标签交换路由器（LSR）必须对在它们之间传输并由它们转发的信息流所携带的标签的含义达成一致。两个相邻LSR就标签含义达成一致的过程需要用到标签分发协议（LDP），LDP就是一个LSR通知另一个LSR关于它已经做出的转发等价类（FEC）和标签之间绑定的过程。



图4 LDP消息格式

根据数据传送的方向，LSR可以分为上游和下游。

上游：以指定的LSR为视角，根据数据传送的方向，所有往本LSR发送MPLS报文的LSR都可以称为上游LSR。

下游：以指定的LSR为视角，根据数据传送的方向，本LSR将MPLS报文发送到的所有下一跳LSR都可以称为下游LSR。

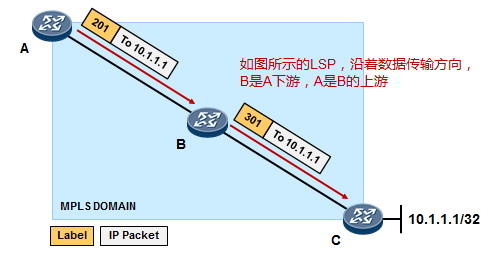


图5 上游与下游

1. 分发方式

根据标签分发的主体的不同，可以分为上游标签分发和下游标签分发，目前使用的是下游标签分发，它又可以分为两种：1）下游按需分发；2）下游主动分发。前一种方式是上游LSR针对某一种FEC向下游LSR提出标签分发申请，后一种是下游LSR主动向上游LSR告知有关标签绑定的信息，而不管上游LSR是否向他提出申请。这两种标签分发方式可以存在于同一网络中，但是不管怎么样，上下游的LSR必须事先对采用何种标签分发方式达成协议。标签分发的过程其实就是为某种FEC建立相应的标签交换路径LSP的过程。标签分发控制模式有两种：独立的标签控制模式和有序的标签控制模式。

1. 分发途径

在显示标签分发模式LDP中，标签分发协议可以在拓扑驱动和业务驱动两种触发方式下工作。

两个LDP对等体间所转发的标签由可以由下游标签分配，也可由上游标签分配，若是由下游LSR进行标签分配，则工作于这种方式下的LSR将该标签用作其交换表的索引，若是由上游LSR进行标签分配，则工作于这种方式下的LSR将该标签用作其交换表查找后的结果。建立标签转发路径LSP可以从入口到出口建立路径，也可以从出口到入口建立路径。

### 1.2.2 工作过程

**（1）基本原理**

MPLS域中的各个路由器使用标记分配协议(LDP)，在边缘路由器之间建立标记交换路径（虚通路）。

一个IP分组到达MPLS域时，入口E-LSR对接收到的IP分组进行分类，将属于不同的FEC的分组映射到不同的标记交换路径LSP。然后为分组加上标记，并转发到下一个标记交换路由器LSR。

LSR对每一个接收到的MPLS分组，利用输入端口号和输入标记查找转发表，找到相应的输出标记和输出端口号。LSR用新的输出标记代替MPLS分组中的旧标记，然后将MPLS分组从输出端口发送出去。

当MPLS分组要离开MPLS域时，出口E-LSR将分组的标记删除，恢复成普通的IP分组转发给IP子网。

控制构件执行路由协议，创建并维护路由表。在路由表的基础上通过LDP建立LSP，并创建和维护LSP对应的转发表。转发构件根据分组的输入端口和所携带的标记去检索转发表，把分组交换到相应输出端口，并用输出标记替代原来的标记。

标记交换路由器LSR由两部分单元组成：一部分是控制单元；另一部分是转发单元，控制单元与转发单元完全分离，可以独立地开发和修改。控制单元执行路由协议，创建并维护路由表。在路由表的基础上通过标记分发协议建立标记交换路径，并创建和维护LSP对应的转发表。转发单元根据分组的输入端口和所携带的标记去检索转发表，把分组交换到相应输出端口，并用输出标记替代原来的标记。

面向过程考虑，MPLS报文转发过程则可以概括为：插入标签，交换标签，弹出标签（弹出一层），去掉标签（全部弹出），包要给谁就发谁的标签，次末跳弹出（Penultimate Hop Poping）。

图6 标签的处理过程

**（2）转发步骤**

典型的MPLS转发过程如下：

1. 所有LSR启用传统路由协议（OSPF、IS-IS等），在LSR中建立IP路由表。
2. 由LDP结合IP路由表来建立LSP。
3. Ingress LER接收IP包，分析IP包头并对应到FEC，然后给IP包加上标记，根据标签转发表中的LSP将已标记的报文送到相应的出接口。
4. LSR收到带有标记的报文，将只分析标记头，不关注标记头之上的部分，根据Label头查找LSP，替换Label，送到相应的出接口。

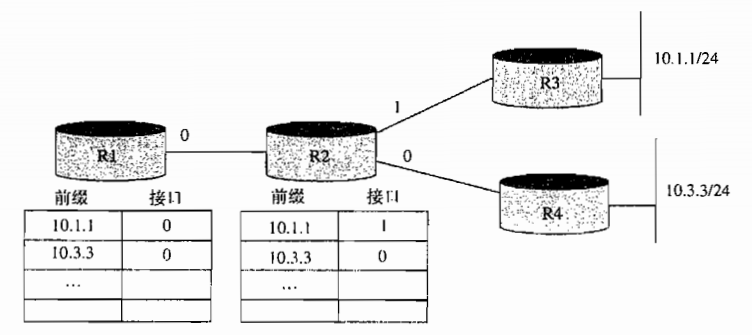
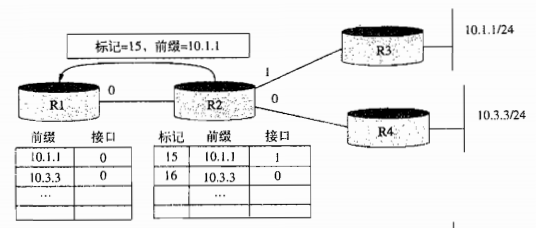


图7 路由表实例

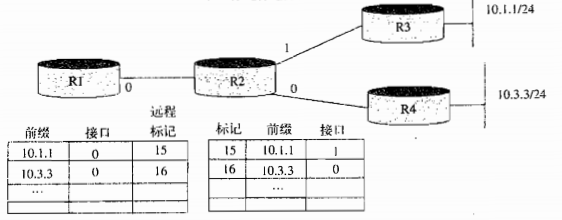
举例说明，如图7网络中的路由表，最右侧2个路由器（R3和R4）中的每一个都连着网络，前缀为10.1.1/24和10.3.3/24。其他2个路由器（R1和R2）有路由表，指出当路由器要转发分组到那两个网络之一时使用哪个发送接口。

当一个路由器能够支持MPLS时,它给路由表中的每个前缀都分配一个标记,并将标记和所表示的前缀通知相邻路由器。此通知的分发由标记分发协议LDP支持。如图7所示,路由器R2给前缀10.1.1分配的标记值是15、给前缀1033分配的标记值是16。这些标记可由分配标记的路由器来选择,可以看作是路由表的索引。

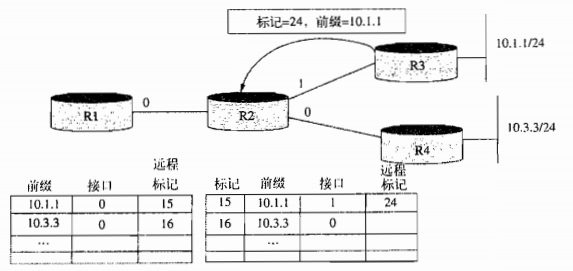
分配标记之后R2将标记绑定通知到相邻节点，在此例中,可以看到R2将标记15和前缀10.1.1之间的绑定通知到R1。事实上,这样的一个通知就相当于R2在说“请将那些发给我的目标前缀为10.1.1的分组全都附着标记15”。R1将标记存在一个表中,旁边是其前缀,表示它是发往那个前缀的任何分组的“远程”或“输出”标记。



a.R2分配标记并将绑定通知R1



b.R1将接收到的标记存在表中



c.R3通知另一个绑定，R2将接收到的标记存在表中

图8 转发过程

在图8-c中，我们看到路由器R3将另一个前缀10.1.1的标记通知R2，并且R2将从R3处得知的“远程”标记放入表中的适当位置。

现在,我们可以来看一个分组在这样的网络中转发时的情况。假设一个目的IP地址为10.1.15的分组从左向右到达R1。在这种情况下，R1叫做标记边路由器(Label Edge Router, LER)，一个LER对到达的IP分组进行完全P查找，然后用它们的标记作为查找的结果。例如，R1发现10.1.1.5与转发表中的前缀10.1.1匹配，并且这条记录中还包含一个输出接口和一个远程标记值。因此R1将远程标记15附着到分组上，然后发出。

当分组到达R2，R2查看分组中的标记。R2中的转发表指示到达的携带标记值15的分组应从接口1发出，并应携带路由器R3通知的标记值24。因此，R2重写或是交换标记，并把它转发到R3。

当标记的这些应用和交换全部进行之后， R2在转发分组时不需要检查IP地址，R3只看输入的标记即可。因此，我们用标记查找代替正常的IP目的地址查找。

### 1.2.3 优势

**（1）精确匹配**

传统IP是基于路由器CPU的，即：数据包到达后，拆3层IP包头，检查目的IP地址和掩码，然后逐一和路由表匹配。虽然IP地址的长度总是相同的,但是IP前缀是变长的，IP目的地址查找算法是最长前缀匹配( longest match)，即找出与将要转发的分组中P地址高比特部分相匹配的最长前缀。总体来说，其寻找路径的过程是基于软件的匹配，普通的ip转发是路由后转发，MPLS交换后转发。

而标记转发机制则是一种精确匹配算法。它根本不需要检查IP地址，只看输入的标记即可。通过使用标记作为数组的索引,其中数组的每个单元是转发表中的一行，就可能实现一个非常简单的精确匹配算法。

MPLS使用定长的标签进行转发，标签交换依据标签表一次查找一次转发，基于2.5层转发流量，即数据包到达后，拆2.5层MPLS包头，找到标签值，然后直接根据标签转发表进行替换并转发，效率更高。并且因为标签是定长的，可以使用asic、np进行处理。由此，基于MPLS的转发方式可以显著提高数据流的转发效率。

MPLS是集成式的IP Over ATM技术，数据包通过虚拟电路来传送，只须在数据链结层执行硬件式交换，可以解决Internet路由的问题，使数据包传送的延迟时间减短，增加网络传输的速度，更适合多媒体信息的传送。因此，MPLS还可以指定数据包传送的先后顺序，其使用标记交换，网络路由器只需要判别标记后即可进行转送处理。

**（2）拓展性强**

MPLS 结合了网络层的灵活连接和可扩展性，以及ATM的标签转发的可靠传输和QoS，提供增值业务的同时不损害效率。

支持多种标准的路由协议、标签生成协议和网络层协议。能有效解决QoS问题，具有标签转发的高性能。

LSP本身就是公网上的隧道，用MPLS来实现VPN有天然的优势，避免了L2TP、GRE等传统VPN在配置和管理上的N平方问题。

VPN 的控制在PE上实现，管理和扩展非常方便。且每个VPN单独构成一个独立的地址空间，即VPN之间可以重用地址，方便的控制VPN 各业务之间的隔离和互通，并且支持流量工程（Traffic Engineering）。

**（3）为IP服务**

MPLS这项技术虽然没有被称为“IP MPLS”，但实际中，却完完全全是为IP服务的。IP是现代网络通讯的基础， MPLS以IP为服务对象，确保该项技术会得到最广泛的吸纳。

MPLS不管是和光集成、和ATM集成、和以太网集成，都脱离不了这个本质。这可以使现有网络结构有充分的过渡空间，也充分支持和现有网络并存。MPLS体系结构中把它当成完成网络结构迁移所使用的技术。

由于充分地依赖于IP，MPLS可以方便的和各种成熟的技术结合，比如BGP、OSPF、IPv6、Multicast，不削弱、不排斥原来技术的应用，而是起到相辅相成的作用，可广泛利用现有成熟协议，是MPLS体系结构足够强大，支撑起新一代网络结构的有力保障。

# 第二章 软件定义网络

## 2.1 SDN概述

软件定义网络（Software-defined networking，SDN），一种新的网络架构。SDN 提出的控制与转发平面分离、网络状态集中控制、支持软件编程等理念并不是什么新鲜事，但是长久以来一直没有非常突破性的进展。

目前 SDN 引起广泛关注得益于网络需求侧翻天覆地的变化：云计算业务（服务器虚拟化技术为代表）成为主流，移动互联网催生的大数据技术日益普及，包括网络在内的资源快速配置、弹性扩容、按需调用需求强烈。传统模式的弊端显现：网络设备硬件、操作系统和网络应用三部分紧耦合在一起，组成一个封闭系统，这三部分相互依赖、每一部分的创新和演进都要求其余部分做出同样的升级。

越来越多的网络新协议和新算法使得网络控制平面变得越来越复杂，但是现在的网络用户却对网络的易用性有更高的要求，希望网络具有更多的可编程能力，从而自动化、智能化网络管理。正如 SDN 的倡导者 Scott Shenker,U.C. Berkeley Professor 所言，网络发展目前还处于“管理复杂性”阶段，这样的架构严重阻碍了网络创新进程的开展。

## 2.2 SDN架构

SDN是一种数据控制分离、软件可编程的新型网络体系架构，其基本架构如下图9所示。SDN采用了集中式的控制平面和分布式的转发平面，两个平面相互分离，控制平面利用控制——转发通信接口对转发平面上的网络设备进行集中式控制。

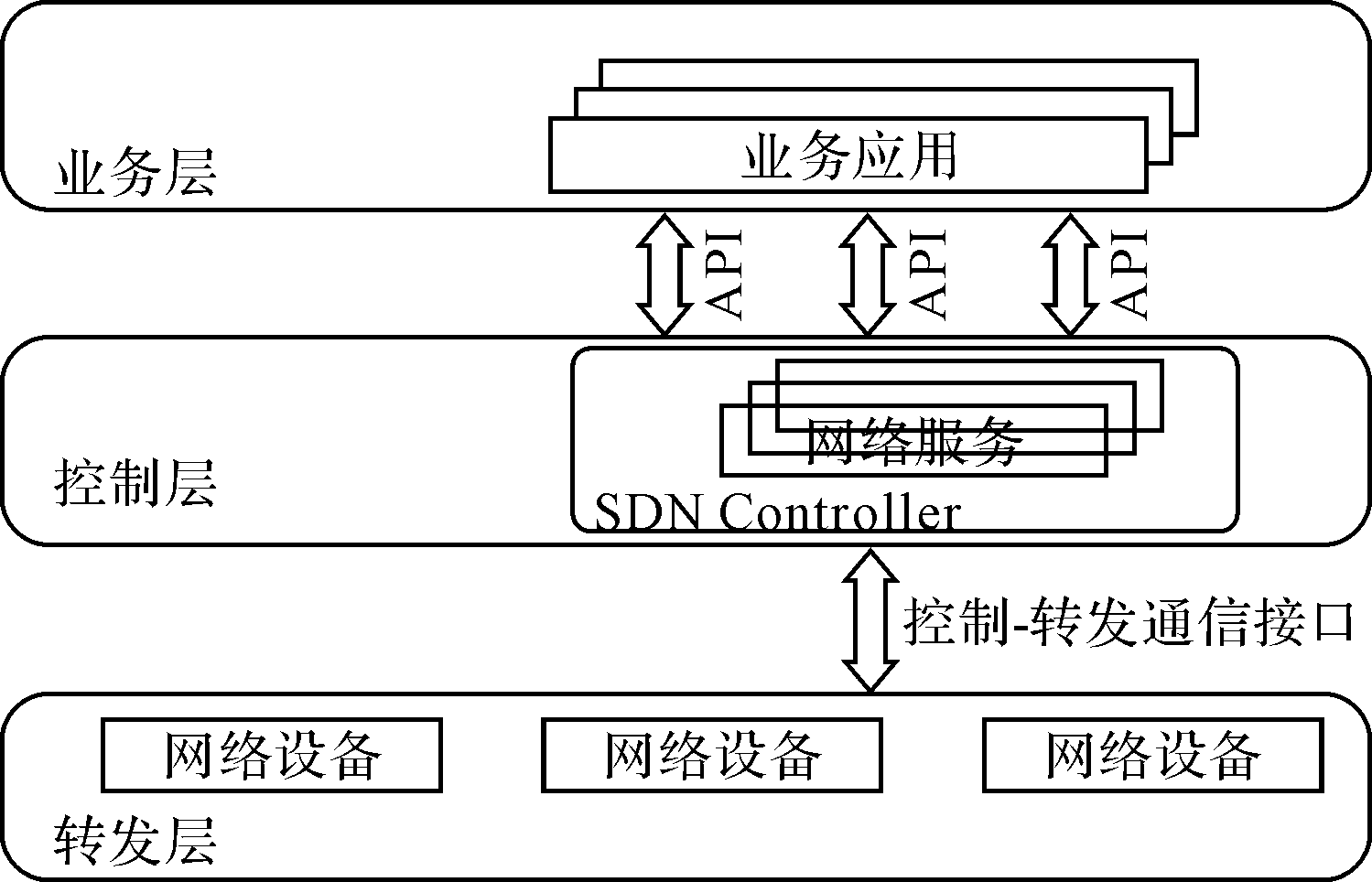


图9 SDN基本架构

### 2.2.1 架构特征

集中控制：逻辑上集中的控制能够支持获得网络资源的全局信息，并根据业务需求进行资源的全局调配和优化，例如流量工程、负载均衡等。

开放接口：通过开放的南向接口和北向接口，能够实现应用和网络的无缝集成，使得应用能告知网络如何运行才能更好地满足应用的需求，比如业务的宽带、时延需求，计费对路由的影响等。

网络虚拟化：通过南向接口的统一和开放，屏蔽了底层物理转发设备的差异，实现了底层网络对上层应用的透明化。

### 2.2.2 各层功能

1. 控制层（Control Layer）

控制层是 SDN 控制器管理网络的基础设施，可以根据需要灵活选择多种控制器。

在这一层中，控制器中包含大量业务逻辑，以获取和维护不同类型的网络信息、状态详细信息、拓扑细节、统计详细信息等。

由于 SDN 控制器是用于管理网络的，所以它必须具有用于现实世界网络使用情况的控制逻辑，网络供应商和开源社区需要在自己的 SDN 控制器中实现自己的服务。这些服务会向上层（应用层）公开自己的API，这使网络管理员可以方便地使用应用程序上的 SDN 控制器的配置、管理和监控网络。

1. 基础设施层（Infrastructure Layer）

基础设施层由各种网络设备构成。它可以是数据中心的一组网络交换机和路由器。控制层负责管理底层物理网络，物理层的实现可以是支持 OpenFlow 的硬件交换机，随着虚拟化技术的完善，SDN 交换机可以是软件形态，例如 Open vSwitch (OVS) 就是一款基于开源技术实现的、能够与服务器虚拟化（Hypervisor）集成，具备交换机的功能，可以实现虚拟化组网。

1. 应用层（Application Layer）

应用层对于开发者来说是开放区域，鼓励开发尽可能多的创新应用。包括网络的可视化：拓扑结构、网络状态、网络统计等；网络自动化相关应用：网络配置管理，网络监控，网络故障排除，网络安全策略等。SDN 应用程序可以为企业和数据中心网络提供各种端到端的解决方案。

1. 南向接口（Southbound interface）

控制层到基础设施层（网络交换机）通讯需要经过南向接口，目前主要的协议是 OpenFlow , NetConf , OVSDB 。 OpenFlow 协议是事实上的国际行业标准,作为一个开放的协议，OpenFlow 突破了传统网络设备厂商各自为政形成的设备能力接口壁垒。

1. 北向接口（Northbound interface）

北向接口：应用层 通过 API 的方式 与 SDN 控制器通讯。与南向接口不同，现在北向接口还缺少业界公认的标准，实现方案思路有的从用户角度出发、有的从运营商角度出发、有的从产品能力角度出发。技术风格上，部分传统的网络设备厂商倾向于在现有的设备上提供编程接口供业务App调用，许多上层应用的开发者也比较倾向于采用 REST API 接口的形式。

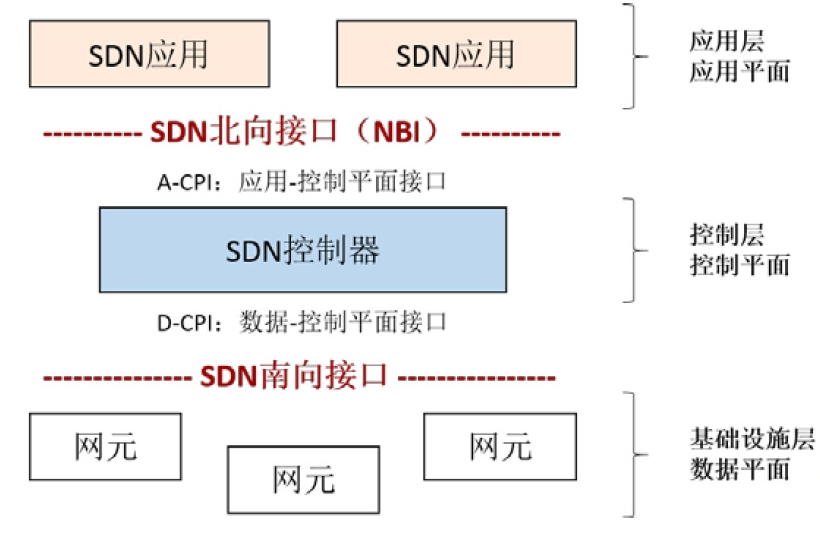


图10 SDN分层结构

## 2.3 SDN优势

### 2.3.1 简单高效

传统网络中的服务链SC与网络拓扑紧密耦合，在SC扩容和变更时，都需要更改底层的网络拓扑，同时重新进行网络设备和中间件的配置，导致基于SC的业务部署比较复杂。传统网络下部署SC—般是借助增强性网关（路由器）将网络流量进行分类之后将需要进行处理的数据包转发给相应的中间件，中间件处理完以后再交给增强型网关，由增强型网关转发给网络中的下一跳路由。从处理方式上就可以看出，服务链上了各个网络功能是彼此独立的，每个网络功能中间件只是对到达的数据包进行相应的功能处理，与服务链上的任何其他的网络功能是没有信息交互的，这就导致了对数据包的重复处理，比如：包头字段的检测。

基于SDN配置服务链的方式很好的解决了传统网络中配置服务链出现的种种问题。

首先，传统网络功能链往往基于手工配置，配置的工作量和复杂度可想而知。而借助SDN中解耦的控制平面提供的北向接口，我们可以通过软件实现功能链的自动配置。

第二，传统的网络功能链依赖于网络拓扑，以至于SC的部署和网络设备的耦合性很大，而基于SDN，通过北向接口可以灵活动态的添加和删除服务链上的服务节点，解耦了网络设备之间的关联。

第三，在传统的SC中，数据包往往要经过多次分类，即解包、封包的过程。而在SDN中，这个过程大大缩减,一般只需分类一次便可以实现数据包的处理，使得整个过程更便捷和高效。

### 2.3.2 低成本

借助软件定义网络 (SDN)，组织能够通过基于策略的工作流程自动化方法加快应用的部署和交付速度，从而大幅度降低 IT 成本。SDN 技术可以大范围提供自动化、按需应用交付和移动功能，为云架构提供支持。这种技术有助于进一步增强数据中心虚拟化的优势，提高资源的灵活性和利用率，并降低基础设施成本和开支。

SDN 之所以能够帮助实现上述业务目标，是因为它能将网络管理和应用服务融合到可扩展的集中式业务协调平台，在整个基础设施范围实现自动调配和自动配置。它还能通过常用的集中式 IT 策略将不同的 IT 组和工作流程联系在一起，借此打造出一个现代化的基础设施，在短短几分钟内实现新应用和新服务的交付，而不必像以前那样耗费几天乃至几周的时间。

SDN 还能为新应用和新业务服务的部署提供速度和灵活性上的保证。灵活、基于策略和可编程，正是思科 SDN 解决方案最独特的特点。此外，思科 SDN 所提供的平台也能胜任当前和未来要求最为苛刻的网络需求。

# 第三章 基于SDN架构的MPLS

## 3.1 结合MPLS部署SDN服务链

### 3.1.1 技术趋势

当前的MPLS技术，存在一些优势的同时也存在缺陷。从优势的角度来看，目前MPLS的优点主要有：服务丰富且成熟；有运营经验；可以进行广泛部署；服务标准化

从存在的问题来看，MPLS网络复杂性高，部署和操作的成本昂贵且效率较低，确定性不高且服务速度有缺陷。

针对MPLS现状，与之有关的新兴技术趋势主要有：

1. MPLSDN
2. 基于云的服务交付和网络功能虚拟化（NfV）
3. IP +光学集成

### 3.1.2 MPLSDN

**（1）演进方向**

MPLSDN简单，快捷，虚拟化，高效，承载了技术演进的未来，其技术演进方向主要有：

1. 拥抱SDN和NFV概念
2. 网络优化的集中控制
3. 基于动态云的基于MPLS网络的服务交付
4. 编程接口和协调

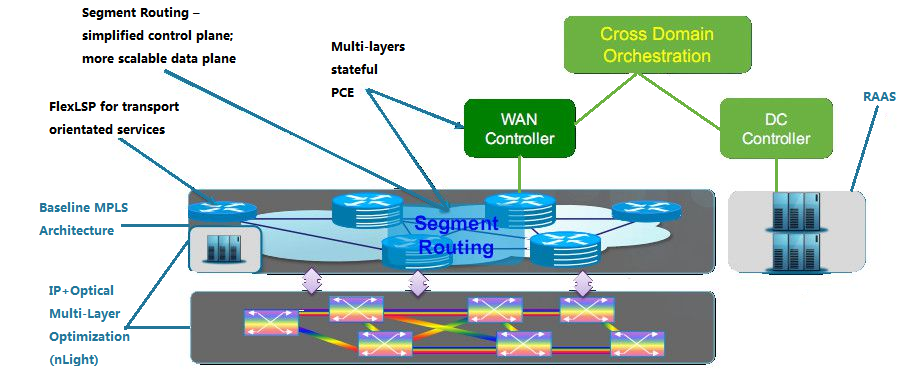


图11 MPLSDN技术体系架构

MPLSDN可以通过技术的继承和平滑演进解决存在的问题。

首先，利用IP / MPLS和光传输中经过验证的概念，在现有的硬件平台上运行。其次，简化的控制平面可以设计更具可扩展性的数据平面。再次，具有用于网络优化和动态服务放置的准入控制和策略引擎的集中控制。

**（2）技术引擎**

MPLSDN是一种新兴的技术引擎，是综合多个技术领域的创新成果，主要有如下几个方向：

1. 多层PCE：IP +光学多层优化（nLight）；WAN控制器。
2. 分段路由
3. FlexLSP
4. IP / MPLS和云集成（RaaS）
5. 跨域协调

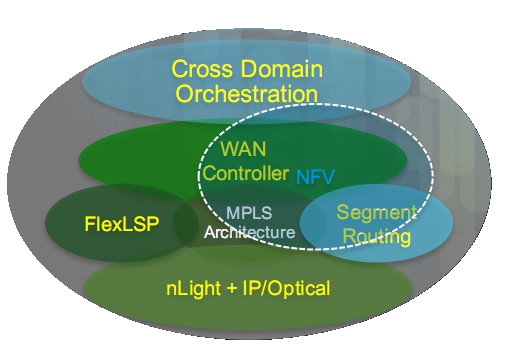


图12 MPLSDN技术引擎

nLight与WAN Controller的结合，是光层与IP层之间的管理和信息交换，使用动态光学控制平面，并使用WAN控制器进行多层优化，是光学和IP域之间的G-MPLS UNI。

FlexLSP是MPLS-TP的一种演进，它将面向传输的服务引入IP/MPLS环境，支持伪线的双向运输定向隧道，可预测路由，保证B/W双向关联LSP，使用MPLS-TP OAM监控LSP状态和驱动保护，编程式VPN服务可以启用NfV。

Segment Routing则是一种更加简化高效的MPLS，无状态MPLS的出现可以通过IGP分发标签，舍去LDP和RSVP，具有简化的优势。路由器的维护规模变小，可以结合WAN控制器进行准入控制和最佳路径确定，并且与现有网络向后兼容。

## 3.2 结合MPLS部署SDN服务链

采用MPLS技术部署服务链，将需要进行相同服务链处理的不同流带上相同的服务链label标签，通过label标签将原本独立的流在服务链上看成一个统一的整体，通过label标签来标识服务链，有效地减少了向数据平面交换机中添加流表的数量。

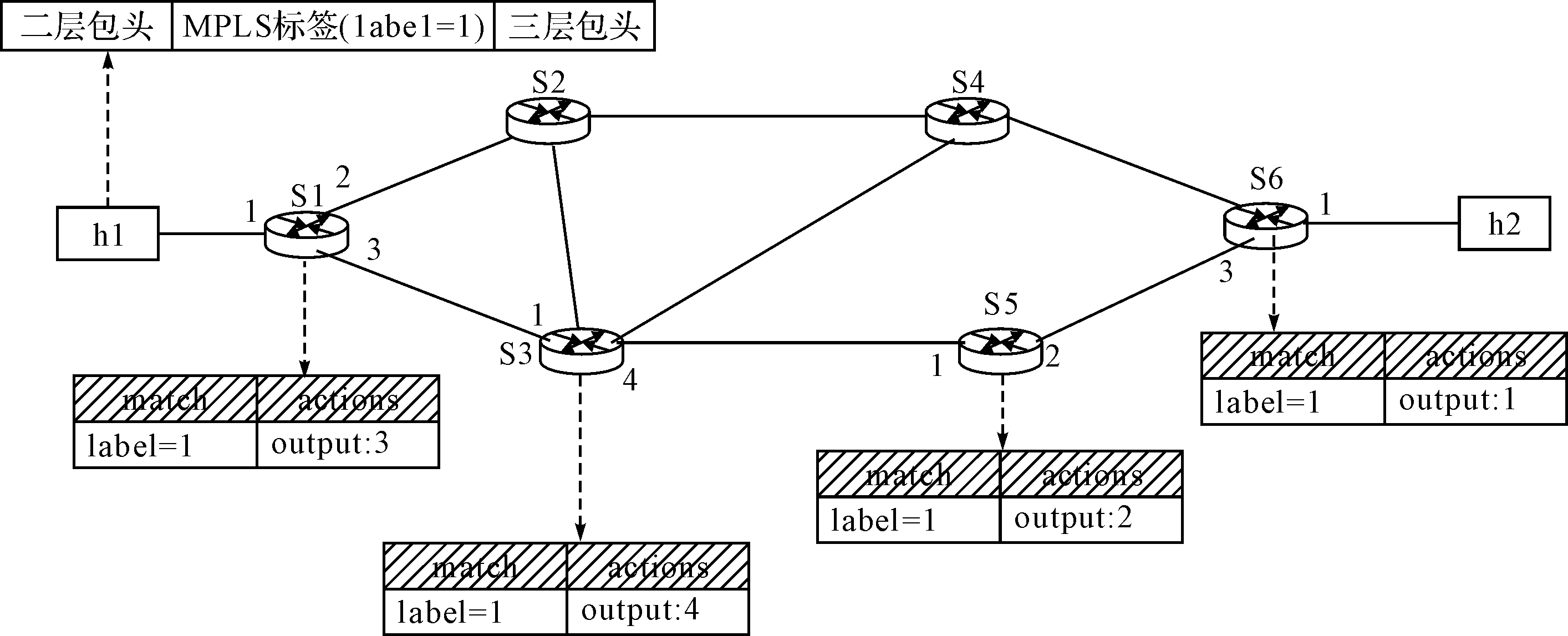


图7 MPLS包头转发流程

采用MPLS技术部署服务链的方式相比传统方式，可以减少向交换机中添加流表的数目，MPLS部署方式针对每条服务链只需对服务链路径上的交换机添加一次对应的流表转发规则，即流的数目的多少对于流表数目没有任何影响。在面对网络规模频繁变动的情况下，若采用基于五元组的部署方式避免不了要频繁对网络中流的传输进行新增或减少的操作。基于MPLS技术部署服务链的方式则很好地应对了此问题，即只要将传输的流正常推送到边缘交换机，借助SDN逻辑层控制器己下发流表的形式，控制正确的流传输到服务链正确的位置。此方法可以轻松地应对频繁变动的网络规模。基于MPLS技术部署服务链方式的灵活性和可扩展性是基于五元组部署服务链所不具备的，因此它也成为解决服务链高效部署的新兴思路。

在故障检测方面，采用MPLS技术部署服务链会为每一段路径附加唯一的标签，在中间件出现意外故障的时候，抓取交换机端口中的数据包解读其中label标签，顺序推理出发生意外故障中间件的具体位置。

在故障恢复方面，由于网络中存在冗余的服务链实例和中间件的信息，在排除故障的过程中，只需将发生故障的中间件从上层维护的中间件集合中删除，然后重新更新一下服务链所对应的中间件集合和与之对应的路径即可。最后使用Ryu控制器向路径上的各个交换机重新下发一遍流表，就可以在有效地解决故障的同时大大压缩故障恢复的时间。

## 3.3 总结

MPLSDN 作为MPLS与SDN技术融合的解决方案，首先能够更快速地业务部署，即实时WAN编排+跨域编排，且使用RaaS的网络功能虚拟化（vCE/vPE）。其次，它具有更紧密的应用可编程，可调度的能力。其路径是分段路由+ WAN协调，需求入场/安置由WAN编排。

再次，MPLSDN具有更高的效率,IP光纤集成和优化（nLight）和统一的MPLS传输,都能使它的效率得到提高。最后，其更加简化，具有更高的扩展性，分段路由的较少的协议可以减少数据平面状态。

采用MPLS技术设计部署服务链，在处理路径相同的流时，只需将流附带上相同的标签进行转发处理，即确保了服务链处理只依据标签进行转发，大大减少了流表的数目，同时基于MPLS的转发方式可以显著提高数据流的转发效率。网络规模的变动对服务链的部署影响很小，配置简单。服务链上的中间件故障定位迅速，故障恢复快，可以灵活的变动和扩展。

# 第四章 总结

总而言之,MPLS是用于解决很多不同的网络互连问题的通用工具，它将通常与虚电路网络相联系的标记交换转发机制,与IP数据报网络的路由选择和控制协议结合起来，产生一类具有两者折中性质的网络。这样就能扩展IP网络的能力,包括使得路由选择的控制更为精确,并且支持一系列VPN服务。

它作为Internet下一代宽带网络核心技术，可以为ISP提供多种大规模IP宽带网络阶必需的功能，例如可预见性、扩展性、可管理胜等。通过MPLS可以建设一个先进的宽带网络，是目前有效实施网络流量工程，更同时支持多种要求特定IPQoS保证的最佳解决方案。

通过这次对MPLS的探究，我通过查阅资料，学习了多协议标记交换的工作原理，应用过程，也讨论了MPLS在技术应用中的优势，收获颇丰。然而，我对于MPLS的了解还远远不够，需要进行进一步的学习。