

[分类浏览](#) [活动](#) [招聘](#) [实验平台](#) [ODL中文社区](#)

搜索



## 热门技术

# Overlay网络与物理网络的关系

SDNLAB君 • 15-01-16 • 1,710 人围观

**编者按：**无论是云计算还是SDN都把注意力集中在数据中心网络的建设上，各种解决方案层出不穷，其中以VMware为代表的软件厂商提出Overlay网络方案后，为数据中心网络的发展提出了新的思路。那么Overlay是如何与物理网络相互依存的？

在以往IT建设中，硬件服务器上运行的是虚拟层的计算，物理网络为了与虚拟服务器对接，需要网络自己进行调整，以便和新的计算层对接（如图1所示）。

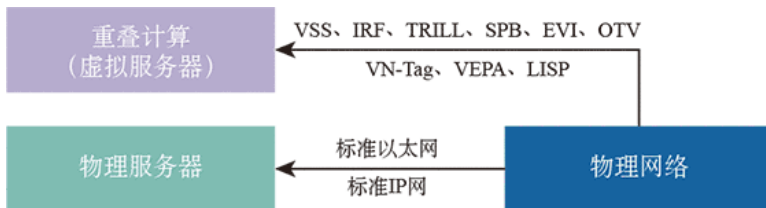


图1 物理网络引入数据中心技术来适配虚拟计算

Overlay是在传统网络上虚拟出一个虚拟网络来，传统网络不需要在做任何适配,这样物理层网络只对应物理层的计算（物理机、虚拟化层管理网），虚拟的网络只对应虚拟计算（虚拟机的业务IP），如图2所示。

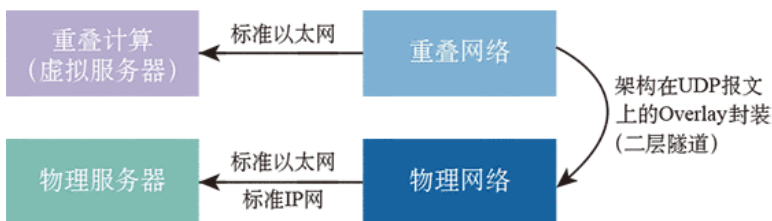


图2 物理网络通过承接虚拟网络来适配虚拟计算

Overlay的技术路线，其实是从架构上对数据中心的建设

## RYU中WSGI学习 笔记

另一篇博文中已经介绍如何使用RYU的RESTAP...

## RYU RESTAPI的使用

使用web拓扑，有时也需要通过API去下发流...

## 广播风暴对比实验部署

环路广播风暴，如果有环路，数据帧将会在...

## OpenFlow网络中的路由服务

所谓OpenFlow网络指的是相互连接的一组Ope...

## OpenDaylight VTN源码及架构...

VTN通过映射机制将虚拟网络资源映射到物理...

[有奖投稿](#) [职位发布](#)

## 最新评论



李呈

你可以去读一读代码，代码里面已经写得很清楚了。@cotyb

模式进行了颠覆，对物理设备的要求降至最低，业务完全定义在层叠网络上。那么，这是否意味着将来数据中心使用Overlay网络不需要硬件支持而只需要软件定义就足够了呢？答案无疑是否定的。

以下讨论Overlay网络与物理网络的依存关系。由于VXLAN ( Virtual eXtensible LAN ) 技术是当前最为主流的Overlay标准，以下VXLAN技术为代表进行具体描述。

## 1. 报文的封装与解封装

VXLAN的核心在于承载于物理网络上的隧道技术，这就意味着要对报文进行封装和解封装，因此需要硬件来加速处理。

在VXLAN网络中，用于建立VXLAN隧道的端点设备称为VTEP ( VXLAN Tunneling End Point , VXLAN隧道终结点 )，封装和解封装在VTEP节点上进行。

在云数据中心，部分业务是不适合进行虚拟化的（如小机服务器，高性能数据库服务器），这些服务器会直接与物理交换机互联，而它们又必须与对应租户/业务的VXLAN网络互通，此时就必须要求与其互联的硬件交换机也能支持VXLAN协议，以接入VXLAN网络。

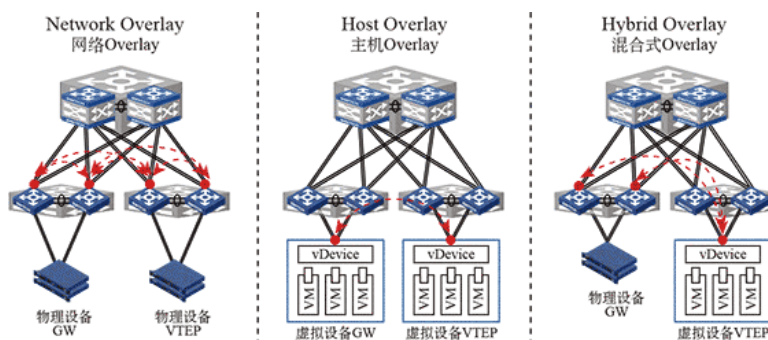


图3 Overlay网络的三种构建模式

考虑到服务器接入的可以是虚拟交换机，也可以是物理交换机，因此存在三种不同的构建模式（如图3所示）：其中网络Overlay方案中，所有终端均采用物理交换机作为VTEP节点；主机Overlay方案中，所有终端均采用虚拟交换机作为VTEP节点；混合Overlay方案中，既有物理交换机接入，又有虚拟交换机接入，且软件VTEP和硬件VTEP之间可以基于标准协议互通。



cotyb

您好,我在用您的最短路径代码做实验,不知道主机如何发起ARP数据包,或者说如何才能让两个主机通信?谢谢你了



moon

你好！我想请问一下，我通过Docker建好2个控制器的平台之后，只要一关了虚拟机，就不能打开\$OC1和\$OC2了这是怎么回事...



AWHANE  
Mark



毛健伟-北邮通信  
^\_^ 是呀，我们是第一批开课的



Dwing  
北邮竟然开了SDN课程。。。



邓晓涛  
好文章:)



毛健伟-北邮通信  
我的博客  
<http://maojianwei.github.io>  
期待与朋友们相互学习、交流 ^\_^



星空夜谈  
@qsdjay 执行cp  
\${ONOS\_ROOT}/tools/  
branding-1.0.0.jar  
\${KARAF\_ROOT}/lib/  
命令



jiameng  
您好！请问是否可以通过OpenVirteX来获取物理链路上的带宽或者

在网络Overlay方案和混合Overlay方案中，都需要有物理交换机设备支持VXLAN协议栈，并能与虚拟交换机构建的VTEP互通。由于在实际组网环境中，服务器种类很多，高吞吐高性能要求的业务一般都采用单独物理服务器甚至小机的硬件环境，而非虚拟化的x86服务器，这就没法使用vSwitch来接入VXLAN网络，只能让支持VXLAN的物理交换机来接入了。

## 2. 组播协议传播

VXLAN网络的MAC表与隧道终端的绑定关系要用组播协议传播，而大规格组播协议离不开物理网络设备的支持。

按照VXLAN的标准，每一个VTEP都需要了解其接入的终端MAC地址，同时还需要知道整网（该VXLAN实例中）其他VTEP下所有的终端MAC地址。只有这样，在本地的VTEP收到报文后需要转发时，才能根据目的MAC查询到需要送到远端的目的VTEP那里。

按照IETF中对VXLAN网络的定义，负责在网络中传播MAC地址和VTEP对应关系的机制，正是依托于物理网络中的组播协议。VTEP将本地的MAC地址表利用组播协议在整个组播中传播，从而使得整网中所有组播成员，也就是其他VTEP都知道本地的MAC地址表。当VTEP下的终端接入情况有所更改，如新增了MAC地址或者减少了MAC地址，也需要利用组播协议通知同一个实例下的所有VTEP。另外，当本地VTEP找不到目的MAC处于哪一个远端VTEP时，也需要发送组播报文来查找目的MAC主机所属的远端VTEP。

traffic状况？另外，我想通过各租户网络的要求（比如根...



微信公众号：搜索"SDNLAB"  
或扫描下面的二维码：



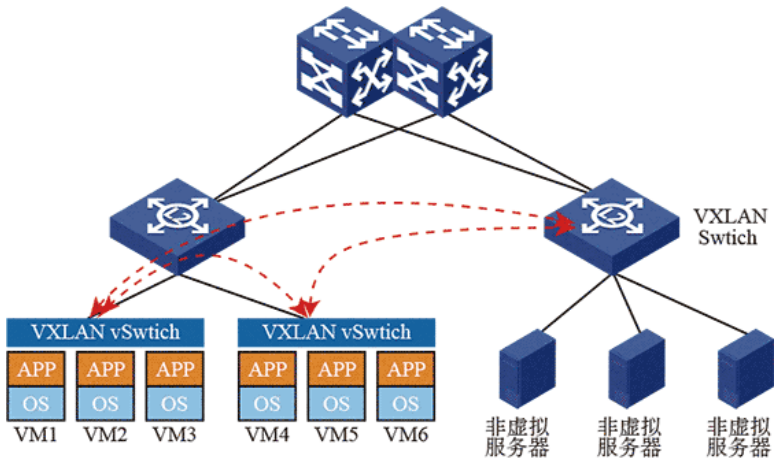


图4 多个VTEP通过物理网络组播传递MAC表信息

如图4所示，多个VTEP需要在整网中传递VTEP下MAC地址信息，逻辑传递路线如绿色虚线所示。由于需要进行逻辑上的Full-Mesh连接，连接逻辑线路会达到N平方量级，因此实际组网中，VXLAN利用了物理网络的组播组，在建立好的组播组中加入VXLAN中所有VTEP成员，传递VTEP变更信息。在多用户多业务情况下，组播组要求与VXLAN数量息息相关。由于VXLAN网络规模的不断拓展（最大可达到16M个VXLAN网络），所需要的组播条目数会不断增加，这实际上对于物理网络承载组播处理能力和规格提出了要求。

由于标准VXLAN架构下使用组播协议，对物理网络组播数规格要求较大，因此H3C VXLAN解决方案基于SDN架构，通过引入全网的SDN Controller来实现VXLAN的管理和维护，使得VTEP之间的信息可以通过Controller来进行反射（如图5所示）。这样，VTEP的MAC地址表映射关系不再通过组播向全网其他VTEP传达，而是统一上报给控制器，由控制器统一下发给需要接受此消息的其他VTEP，由具体的VTEP执行转发机制。

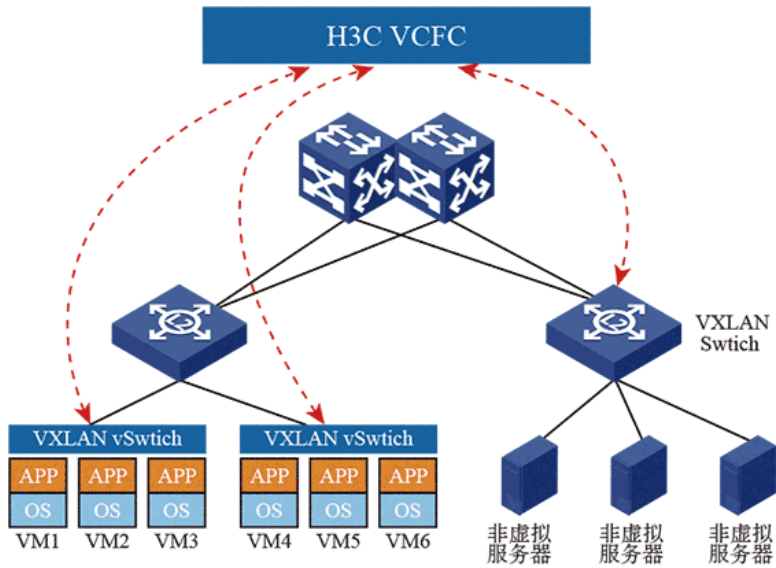


图5 多个VTEP通过SDN Controller传递MAC表信息

可见，在SDN架构下，硬件形态的VTEP需要能够支持集中控制器下发的业务控制信息，同时基于Openflow进行流表转发。而传统硬件交换机不能支持上述特性，必须由新硬件设备来执行和完成的。

3. VXLAN网络互通

在传统L2网络中，报文跨VLAN转发，需要借助三层设备来完成不同VLAN之间的互通问题。VXLAN网络与传统网络、以及VXLAN网络的互通，必须有网络设备的支持。

VXLAN网络框架中定义了两种网关单元。

VXLAN三层网关。用于终结VXLAN网络，将VXLAN报文转换成传统三层报文送至IP网络，适用于VXLAN网络内服务器与远端终端之间的三层互访；同时也用作不同VXLAN网络互通。如图6所示，当服务器访问外部网络时，VXLAN三层网关剥离对应VXLAN报文封装，送入IP网络；当外部终端访问VXLAN内的服务器时，VXLAN根据目的IP地址确定所属VXLAN及所属的VTEP，加上对应的VXLAN报文头封装进入VXLAN网络。VXLAN之间的互访流量与此类似，VXLAN网关剥离VXLAN报文头，并基于目的IP地址确定所属VXLAN及所属的VTEP，重新封装后送入另外的VXLAN网络。



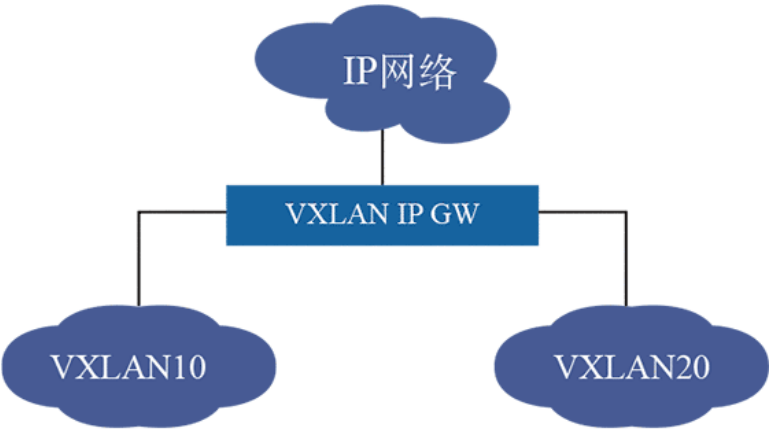


图6 VXLAN网络通过三层网关与传统网络互连

VXLAN二层网关。用于终结VXLAN网络，将VXLAN报文转换成对应的传统二层网络送到传统以太网网络，适用于VXLAN网络内服务器与远端终端或远端服务器的二层互连。如在不同网络中做虚拟机迁移时，当业务需要传统网络中服务器与VXLAN网络中服务器在同一个二层中，此时需要使用VXLAN二层网关打通VXLAN网络和二层网络。如图7所示，VXLAN 10网络中的服务器要和IP网络中VLAN100的业务二层互通，此时就需要通过VXLAN的二层网关进行互连。VXLAN10的报文进入IP网络的流量，剥掉VXLAN的报文头，根据VXLAN的标签查询对应的VLAN网络（此处对应的是VLAN100），并据此在二层报文中加入VLAN的802.1Q报文送入IP网络；相反VLAN100的业务流量进入VXLAN也需要根据VLAN获知对应的VXLAN网络编号，根据目的MAC获知远端VTEP的IP地址，基于以上信息进行VXLAN封装后送入对应的VXLAN网络。

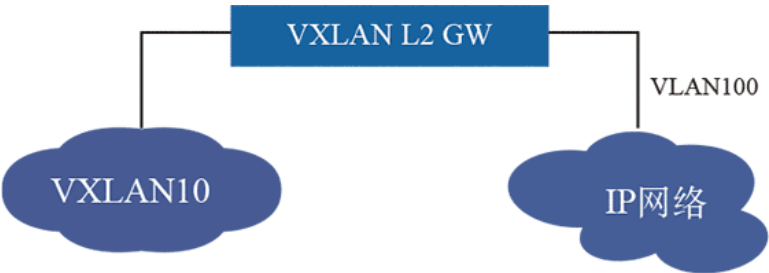


图7 VXLAN网络通过二层网关与传统二层网络互通

可见，无论是二层还是三层网关，均涉及到查表转发、VXLAN报文的解封装和封装操作。从转发效率和执行性能来看，都只能在物理网络设备上实现，并且传统设备无法支持，必须通过新的硬件形式来实现。

## 结束语

Overlay由于其简单、一致的解决问题方法，加上重新定义的网络可以进行软件定义，已经成为数据中心网络最炙手可热的技术方案。然而，它并不是一张完全由软件定义的网络，Overlay网络解决方案必定是一种软硬结合的方案，无论是从接入层VTEP混合组网的组网要求、组播或SDN控制层协议的支持，还是VXLAN网络与传统网络的互通来看，都需要硬件积极的配合和参与，必须构建在坚实和先进的物理网络架构基础上。

转载自：[http://www.h3c.com.cn/About\\_H3C/Company\\_Publication/IP\\_Lh/2014/07/Home/Catalog/201501/852551\\_30008\\_0.htm](http://www.h3c.com.cn/About_H3C/Company_Publication/IP_Lh/2014/07/Home/Catalog/201501/852551_30008_0.htm)

本站声明：本站原创文章可以转载，请注明来自 **SDNLAB**

本文链接：<http://www.sdnlab.com/7241.html>

本文标签：**技术**

分享到：

7 喜欢

## 相关阅读



虚拟化解决方案  
并不仅限于SDN



SDN快速发展  
网络虚拟化准备  
就绪



网络虚拟,云卷云  
舒



企业部署虚拟化  
的七大好处



虚拟化盘  
点:VMware仍是  
主角 Docker异  
军突起



虚拟化：不只  
是“省钱”

上篇：2020年运营商SDN和NFV的投入达...

下篇：SDNLAB每周播报

0条评论

有话不说，憋着多难受啊!

评论

- 友情链接
- [北邮李呈一博客](#) | [蓝点网](#) | [数据小兵](#) | [派克网](#) | [铁匠运维网](#) | [IT那些事](#) | [道普网](#) | [TMTforum](#) | [移动新媒体网](#) | [云时代](#) | [大数据](#) | [极念网](#) | [望星湖畔](#) | [创新网Xilinx中文社区](#) | [智能设备](#) | [泰泽网](#) | [问说网](#) | [云计算](#) | [OpenStack中国社区](#) | [易安在线](#) | [IT资讯](#) | [开源资讯](#)



Copyright ©2014-2015  
SDNLAB.COM All Rights Reserved.  
苏ICP备13052757号-3

关注我们

- RSS订阅
- 微信
- 新浪微博
- 腾讯微博

关于

- [免责声明](#)
- [关于我们](#)
- [加入我们](#)

服务

- [商务合作](#)
- [有奖投稿](#)
- [意见反馈](#)

合作伙伴