

云计算环境下的网络技术研究

徐雷¹, 张云勇¹, 吴俊², 房秉毅¹

(1. 中国联通研究院, 北京 100048; 2. 北京邮电大学 经济管理学院, 北京 100876)

摘 要: 提出了以数据中心为核心的云计算网络技术体系框架, 通过对云计算环境下的网络新需求分析, 对虚拟机本地互访网络、数据中心二层互访网络、跨数据中心二层互访网络、用户接入网络以及 SDN (software defined network) 等新技术进行了详细探讨, 并提出了云计算网络发展建议, 使其为未来云计算网络技术落地做好准备。

关键词: 云计算; 数据中心; 虚拟化; SDN

中图分类号: TN915.81

文献标识码: B

文章编号: 1000-436X(2012)Z1-0216-06

Network technology research under cloud computing environment

XU Lei¹, ZHANG Yun-yong¹, WU Jun², FANG Bing-yi¹

(1. China Unicom Research Institute, Beijing 100048, China;

2. School of Economics and Management, Beijing University of Posts and Telecommunications, Beijing 100876, China)

Abstract: The framework of cloud computing network around data center was proposed. Through the analysis of new demands of the network in the cloud computing environment, new technologies were discussed in detail for the virtual machine local network, data center local network, across data center network, user access network and SDN. Development proposal was provided in order to prepare for the implementation of cloud computing network technology.

Key words: cloud computing; data center; virtualization; SDN

1 引言

近年来, 云计算概念出现并引起产业界高度的重视。云计算是一种通过网络实现对各种 IT 能力进行灵活调用的服务模式。云计算通过分布式计算、虚拟化等技术, 构建用于资源和任务统一管理调度的资源控制层, 将分散的 ICT 资源集中起来形成资源池, 动态按需分配给应用使用。云计算服务的出现大大降低了企业信息化的成本, 提高了信息化系统的资源利用效率, 并推动了云服务市场的快速发展。未来信息通信与消费的世界是一个“终端—网络—业务”三者缺一不可的世界, 网络作为重要组成部分负责连接终端与业务, 云计算业务的发

展使得网络面临新的变革^[1]。

数据中心是云计算业务的基本载体, 基于云计算的信息急剧膨胀使得数据中心网络面临着巨大的压力, 在网络中发生超大规模的数据访问已经成为必然趋势。传统的数据中心网络已不能满足云计算的需求, 云计算向网络技术提出了新的挑战, 也催生了新技术的出现。

本文从以数据中心为核心的云计算网络技术体系框架出发, 通过对云计算带来的网络新需求的分析, 主要对虚拟机本地互访网络、数据中心二层互访网络、跨数据中心二层互访网络以及用户与数据中心之间的网络新技术进行了详细探讨, 并提出了云网络发展建议, 使其为将来云计算网络技术落地做好准备。

收稿日期: 2012-08-25

基金项目: 国家自然科学基金资助项目 (71172134); 国家科技重大专项基金资助项目 (2012ZX03002001)

Foundation Items: The National Natural Science Foundation of China (71172134); The National Science and Technology Major Project (2012ZX03002001)

2 云计算网络技术体系框架

云计算业务主要承载在数据中心内部，数据中心是云计算业务的基本载体，满足云计算的数据中心随着云计算业务的部署逐步发展，经历一系列演进，以承载云计算业务为主要目的云计算数据中心呈现出规模化、虚拟化、绿色化等新特点，云计算将给数据中心带来一系列新的变革。传统数据中心基本上沿用了 IP 网络中层次化汇聚的组网模式，业务服务器、数据库服务器等 IT 设备通过接入以太网交换机（TOR，柜顶交换机）、汇聚交换机、核心交换机，直至出口路由器连接至互联网。由于传统数据中心大多是互联网用户与服务器之间的交互数据，即所谓“南北向”流量，而服务器之间的“东西向”流量较小，因此适合采用这种层层汇聚的层次化组网模式。而云计算的流量特点是横向的流量比重增大，大规模的云计算需要更新数据中心网络架构。

如图 1 所示，云计算网络从业务到终端可分为虚拟机之间、服务器之间、数据中心之间以及用户与数据中心之间 4 个部分，其中虚拟机之间、服务器之间的网络位于数据中心内部。

1) 虚拟机之间

物理服务器内虚拟机之间的网络，也称虚拟机本地互访网络，目前技术主要由软件层面的虚拟交换机（vSwitch）负责进行承载。

2) 服务器之间

物理服务器之间的网络，主要由交换机进行互联，包含纵向与横向的流量。数据中心的分布式计算以横向流量为主，服务器间的虚拟机交换以及横向的迁移需要在二层网络内进行交换，网络面临大二层网络的组网需求。

3) 数据中心之间

数据中心之间互联的网络，同一城域内通过城域网的核心层进行互联，不同城域间的数据中心通过骨干网进行互联。随着云计算数据中心逐步改造，云业务规模不断扩大，某些规模业务的集群组件需要位于不同的数据中心，需要在网络跨数据中心机房搭建二层网络。

4) 用户与数据中心之间

用户与数据中心之间通常通过城域网进行互联，随着云业务的逐步部署，应用逐步集中部署在数据中心内，用户与数据中心之间的流量会逐步增大，承载网络带宽需要增大，网络带宽增大的同时需要更加智能，用于对更加丰富的业务提供不同性能的管道服务。当多个数据中心同时提供服务或者服务从一个数据中心迁移到另一个数据中心时，用户还要面临如何选择数据中心并快速切换的问题。

3 云计算环境下的网络新需求

云计算主要包含 2 种计算模式：一虚多和多虚一。最初的云计算从网格计算、分布式计算演进而来，主要是通过多台服务器完成一台超级服务器才能完成的计算任务，这是通常所说的多虚一。另一种是将一台物理服务器虚拟化出多个虚拟机，实现物理资源的高利用率，称为一虚多。2 种模式都为网络带来新问题，提出新的需求。

1) 横向流量增大带来的新问题

分布式计算和虚拟机迁移带来了服务器之间横向流量的增加，数据中心内部的流量模型从以纵向流量（南北向流量，用户访问服务器流量）为主转变为以横向流量（东西向流量，服务器之间的流量）为主。由于横向流量突发性强，容易产生网络

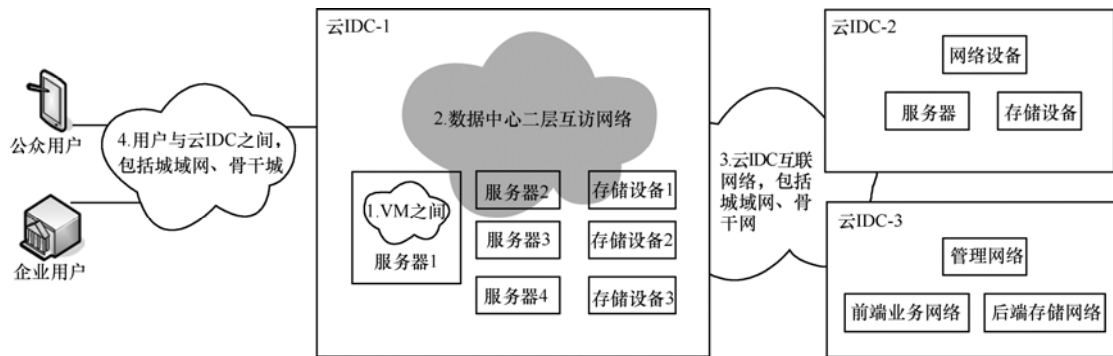


图 1 云计算网络技术体系框架

拥塞,造成分组丢失,给数据中心内部通信的压力增大,对带宽和延迟的要求更高。

2) 虚拟机动态迁移带来的新问题

采用虚拟机动态迁移技术将为数据中心带来更大的灵活性和价值:提高运维效率、节能减排和容灾备份。

但是由于目前虚拟机动态迁移技术要求所迁移的虚拟机维持 IP 地址、MAC 地址不变,在跨域场景下受限于物理路由的可达性,当迁移发生后,虚拟机原有的网络配置将不能与外界正常通信,无法保证业务连接的连续性。因此在跨域场景下首先需要所有的虚拟机位于统一的二层网络下,以满足网络的基本需求,这要求建立统一的二层网络。

3) 虚拟机流量监控带来的新问题

随着服务器被改造成虚拟化平台,接入层从以前的物理端口延伸到服务器内部,为不同虚拟机之间的流量交换提供服务,将虚拟机同网络端口关联起来。目前业界虚拟机之间的数据交换通常由虚拟交换机来完成。vSwitch 实现虚拟机之间通信的流量监管、虚拟交换机端口策略等功能,另外 vSwitch 管理范围被限制在物理服务器网卡之下,无法在整个数据中心提供针对虚拟机的端到端服务。同一台物理服务器上不同虚拟机的流量在离开服务器网卡后仍然混杂在一起,对于上联交换机来说,旧的网管界面无法处理虚拟环境下的多流量共同使用同一个端口的问题,无法做到对同一台物理服务器的虚拟机之间流量进行监控。

4) 网络规模扩大带来的新问题

网络规模扩大是由于服务器和虚拟机数量增长造成的。目前业界采用的方案包括扁平化的二层网络架构和传统三层网络架构。通常采用的二层网络与传统三层组网方案相比,可以有效提高数据中心内部流量的转发效率,并且降低组网成本。但是,网络规模增大后二层网络的可靠性及扩展性差等问题都将成为新的瓶颈。

综合考虑上述新问题,云计算环境下的网络新需求如下:

- 1) 高效无收敛或低收敛的二层网络;
- 2) 具有承载大规模虚拟机的能力;
- 3) 网络架构及设备支持虚拟机的动态迁移;
- 4) 网络设备支持虚拟机的流量监控;
- 5) 跨数据中心的二层网络可达。

4 云计算环境下网络新技术

4.1 虚拟机本地互访网络

目前,业界通过虚拟交换机解决同一台物理服务器内部的虚拟机二层网络互访问题,跨物理服务器的虚拟机互访才交给传统的接入交换机处理。虚拟交换机运行于服务器内部,纯软件实现,简单方便,但也存在 2 大问题。一个是虚拟机之间的流量监控问题,传统的网管系统无法深入服务器内部进行流量监控,造成安全隐患;二是性能问题,虚拟机网络流量越大,vSwitch 就会占用越多的 CPU 资源进行报文转发,降低了服务器支持更多虚拟机的能力。

为解决这 2 大问题,由不同厂商主导,出现了 2 个技术标准体系,包括 IEEE 的 802.1Qbg Edge Virtual Bridging^[2]和 IEEE 的 802.1Br Bridge Port Extension^[3]。

802.1Qbg Edge Virtual Bridging 标准主要包括 VEPA(virtual Ethernet port)和多通道(multichannel) 2 种方案,兼容了 vSwitch 实现,主要由 IBM、HP、H3C 等主流厂商支持。VEPA 不提供虚拟机之间的交换功能,而是将虚拟机报文由外部接入交换机处理,同一台物理服务器之间的报文可以从外部接入交换机重新送回原服务器,实现报文转发。这种方式简化了 vSwitch 功能,使所有 VM 逻辑上对应到外部接入交换机的一个物理网络端口,使外部网络功能延伸到服务器内。多通道技术借用了 Q-in-Q 的 VLAN 标签,实现 vSwitch、VEPA 的混合部署方案,对网卡和交换机链路中的多种类型流量进行细分,使每种类型流量和外部交换机建立单独逻辑通道,实现不同的网络安全、性能及可管理性要求。

802.1Br Bridge Port Extension 是思科提出的新技术,通过引入端口扩展技术(port extender PE,后续简称 PE),为以太网报文增加 TAG 标签,外部端口扩展设备使用 TAG 中的信息来实现报文转发和策略控制。

2 种标准并未成熟,都还在发展中,802.1Br 提供了一整套的网络虚拟化解决方案,但是需要改变硬件设备;802.1Qbg 提出了解决方法,通过修改交换机驱动程序即可实现,最终谁能胜出尚无定论。

4.2 数据中心二层互访网络

数据中心普遍采用树状三层网络架构,分为接

入层、汇聚层和核心层，当 2 个服务器位于树状结构的不同分支时，它们之间的互访要经过核心层转发，多跳才能实现，大大增加了时延，已经不能满足云计算环境下横向流量增加的性能需求。另一方面，在三层网络架构下，如果加入防火墙设备，会在树的一个分支下建立一个 VLAN，当 VLAN 下的某台虚拟机迁移出 VLAN 范围，可能会导致该虚拟机运行异常。解决办法是扁平化数据中心网络结构，由三层结构变为二层结构，即接入层、核心层两级，以提高服务器互访性能，支持虚拟机正常迁移，如图 2 所示。

在二层架构下还有很多问题要处理，虚拟化技术使得一个物理服务器分为多台虚拟机，数据中心内部网络规模成倍的增加。传统网络通过 STP 生成树协议来避免环路，阻塞多路径，浪费带宽，给大规模网络的转发带来了瓶颈。由此产生了控制平面虚拟化和数据平面虚拟化 2 大技术，在避免环路的基础上，充分利用了多路径的带宽。

控制平面虚拟化主要是将网络同一层的多台交换设备横向整合，在逻辑上简化了网络架构，使管理维护和配置简单化，消除了网络环路，保证了网络可靠性和链路的充分利用。其在控制层面由统一逻辑主体完成，转发层面仍是分布式转发结构，典型的技术如思科的 VSS (virtual switching system) /vPC (virtual port-channel)和 H3C 的 IRF(intelligent resilient framework)/IRF2 技术。但是，控制层面统一后，多台交换设备是主备模式工作，无法做到负载均衡，使得部署规模受限于控制节点处理能力，制约了该技术进一步发展。

由于控制层面虚拟化规模受限，使得数据平面虚拟化技术将成为未来的主流。数据平面虚拟化通

过 ISIS 协议进行拓扑路径计算，将整个数据中心网络虚拟成一个大的交换机，通过对原始报文进行封装实现内部转发，避免环路且实现多路径负载分担。目前主要技术包括 IETF 的 TRILL (transparent interconnection of lots of links)^[4]和 IEEE 的 SPB (shortest path bridging)^[5]2 大标准体系。TRILL 需要硬件支持，SPB 靠纯软件实现，其对 CPU 压力也超过了 TRILL。上述 2 个协议在数据中心部署中怎样与其他数据中心协议协同工作的问题都没有解决，2 种标准体系都不成熟，有待发展。

Juniper 的 QFabric^[6]技术将控制平面虚拟化和数据平面虚拟机相结合，从架构上改变了网络形态，使得数据平面所有端口直接互联，之间只有一跳，控制平面由主备控制变为分布式控制。但由于是私有协议，将会制约其发展规模。

4.3 数据中心跨站点二层互访网络

在传统网络中，多数据中心之间主要做扩容和灾备。在云计算环境下，将会出现多站点分布式计算和虚拟机的跨站点动态迁移需求，需要构建跨数据中心的二层网络。多站点二层网络最简单的实现方法就是光纤直连，从国内来看，数据中心基本都没有做到直连，通过城域网或骨干网互联，因此目前技术都是通过 IP 网上打隧道实现二层互联。传统技术包括 VLLoGRE/VPLSoGRE 和 L2TPv3 等。

最新技术是由思科提出的 OTV^[7] (overlay transport virtualization) 私有技术，通过“Mac in IP”方式封装原始以太网报文，使用 MAC 地址路由规则，提供一种叠加 (overlay) 网络，能够在分散的数据中心之间实现二层连接，同时保持这些数据中心的独立性以及 IP 互联的容错性、永续性和负载均衡优势。

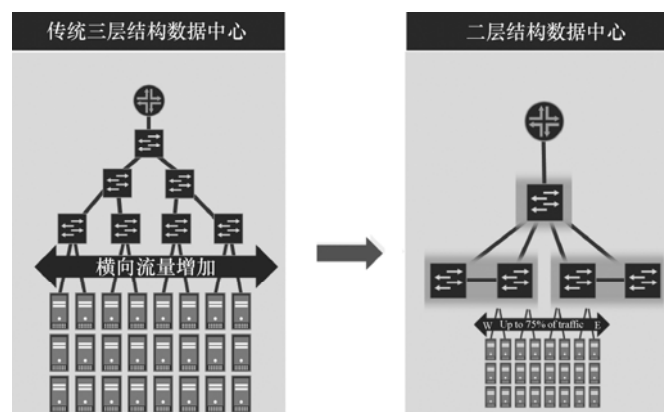


图 2 三层架构转换为二层架构

目前, 国外数据中心互联流量每年递增, 国内需求较新, 需密切关注下一步发展。从技术上讲, 光纤直连最为简单和可靠, OTV 技术较全面和领先, 其他厂商未来可能会出现类似技术。

4.4 用户接入网络

跨数据中心的二层网络的构建, 使得用户接入数据中心时, 可能会面临以下问题: 1) 多站点分布式计算时, 用户是接入数据中心 A 还是数据中心 B; 2) 业务虚拟机从数据中心 A 迁到数据中心 B 后, 用户如何能随之快速切换。在虚拟机跨数据中心迁移时, IP 地址保持不变, 使得多站点选择技术变得更加复杂。

目前的主要技术实现是传统的 DNS 技术和最新的 LISP (locator/ID separation protocol) 路由发现技术。传统的 DNS 技术使得同一个域名解析为不同 IP, 用户通过访问解析出的不同 IP 即可接入不同数据中心。当虚拟机跨数据中心迁移时, 由于 IP 地址不变, 需要在虚拟机外部加一层 NAT 设备, 将其 IP 地址映射为不同的虚拟 IP, 实现用户访问的切换。

思科的 LISP 技术是一个 IPinIP 协议, 是一种名址分离网络技术。通过两层 IP 封装实现虚拟机跨数据中心迁移时保持内层 IP 地址不变, 用户访问仍可精确定位到最新的数据中心并实现最优路由。目前的数据中心多站点选择没有更好的解决方案, 当多数据中心二层互联成为主流时, 可能会促使更新的多站点选择技术的出现。

4.5 SDN 技术

SDN^[8]是一种控制与转发分离并直接可编程

的新的网络架构, 控制层面从传统网络设备紧耦合架构中拆分出来, 形成转发层、控制层和应用层 3 层逻辑架构。控制层可实现全软件编程, 并通过 OpenFlow 等开放协议与转发层通信, 转发层只需关注通用设备即可。SDN 是在 OpenFlow 基础上被提出和关注的最新网络技术, 适用于云计算环境下数据中心内复杂网络, 实现计算资源和网络资源的协调和整合, 方便实现网络虚拟化及虚拟机迁移。

5 云计算网络发展建议

从表 1 可以看出, 云计算环境对端到端网络有新的需求, 目前, 传统网络技术依然可以满足近期需求, 但最新技术正在快速孕育发展。每一个新需求都有多种新技术解决方案实现, 未来何种技术会成为主流无法预测, 但是现在可以为新技术的落地做好准备。数据中心内部传统的三层架构应尽快向二层架构转换, 当前产品都已支持二层架构, 方便未来向 TRILL/SPB 技术演进。业界对数据中心之间的迁移普遍认同度较低, 应尽量避免频繁地在多个数据中心之间发生大规模的数据迁移, 以免由于意外或不可抗力造成传输失败, 导致数据丢失等。但从国外现状看, 数据中心之间数据流量每年递增, 应尽快实现数据中心之间的光纤直连。从现有新技术发展来看, 私有协议引领着技术发展, 一个数据中心内部将可能以单一厂商设备为主, 不同数据中心可以部署不同厂商设备, 它们之间通过公有协议互联。随着 SDN 技术的发展, 设备互联互通可能会得到解决。

表 1 网络类型与技术对比

网络类型	技术需求	最新技术
虚拟机本地互访网络	本地虚拟机互访需要流量监控 需要提高流量转发效率	IEEE 802.1Qbg EVB (Edge Virtual Bridging) IEEE 802.1Br BPE (Bridge Port Extension)
数据中心二层互访网络	数据中心内部横向流量增加需要提高服务器互访 转发效率 需要支持虚拟机迁移 需要支持大规模二层网络	扁平化二层网络 控制平面多虚一: 私有技术 VSS/vPC/IRF/IRF2 数据平面多虚一: IETF TRILL/IEEE 802.1aq (SPB) /Juniper QFabric
跨数据中心二层互访网络	需支持虚拟机跨数据中心迁移	传统通道技术 VLLoGRE/VPLSoGRE 和 L2TPv3 思科 OTV
用户接入网络	需支持多数据中心选择	传统 DNS 技术 思科 LISP

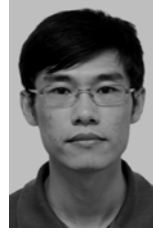
6 结束语

云计算的出现对网络提出了更高的要求，促使了网络新技术的出现。相比计算虚拟化和存储虚拟化来说，网络虚拟化技术更加多样化，更不成熟，随着云计算大规模的落地实施，云计算网络主流技术将逐渐明晰。最新的 SDN 技术的出现，有可能成为云计算网络另一新的发展方向。

参考文献：

- [1] 童晓渝, 吴钢, 张云勇等. 后电信时代[M]. 北京: 人民邮电出版社, 2010.
TONG X Y, WU G, ZHANG Y Y. Post Telecommunications Age[M]. Beijing: Posts and Telecom Press, 2010.
- [2] 802.1Qbg - edge virtual bridging[EB/OL]. <http://www.ieee802.org/1/pages/802.1bg.html>.
- [3] 802.1BR-bridge port extension[EB/OL]. <http://www.ieee802.org/1/pages/802.1br.html>.
- [4] Transparent interconnection of lots of links (trill)[EB/OL]. <http://datatracker.ietf.org/wg/trill/charter/>.
- [5] 802.1aq-shortest path bridging[EB/OL]. <http://www.ieee802.org/1/pages/802.1aq.html>.
- [6] QFabric[EB/OL]. <http://www.juniper.net/us/en/dm/datacenter/>.
- [7] Overlay transport vIrtualization (OTV) [EB/OL]. <http://www.cisco.com/en/US/netsol/ns1153/index.html>.
- [8] Soft defined network[EB/OL]. <https://www.opennetworking.org/>.

作者简介：



徐雷 (1978-), 男, 山东泰安人, 博士, 中国联通研究院工程师, 主要研究方向为云计算、平台、下一代网络研究。



张云勇 (1976-), 男, 江苏盐城人, 博士后, 中国联通研究院平台与云计算研究中心主任, 主要研究方向为下一代开放网络、固定移动融合核心网、移动互联网及业务、公共运算。

吴俊 (1971-), 男, 湖南长沙人, 博士, 北京邮电大学副教授, 主要研究方向为移动互联网应用、云计算商业价值评价等。



房秉毅 (1980-), 男, 山东青岛人, 博士, 中国联通研究院高级工程师, 主要研究方向为云计算、核心网新技术。