



电信运营商数据中心互联技术

唐利莉, 张宁

(中国移动通信集团设计院有限公司网络所, 北京 100080)

摘要:随着云计算和大数据的不断发展,业务对计算和存储资源的需求不断挑战单个数据中心的规模,建设一张数据中心互联网络,成为实现数据中心规模灵活扩展和业务灵活部署的必然要求。通过建设数据中心互联网络,电信运营商可以进一步对IT资源做“云”化整合,以提升资源的利用率以及加快新业务的部署效率,提升核心业务的竞争力。分析了当前具有代表性的数据中心互联方案和技术特点,结合电信运营商的资源整合需求,探讨数据中心互联技术在电信运营商云网络建设中的应用。

关键词:云计算;虚拟机;VDC;SDN;VxLAN

中图分类号: TN915.12

文献标识码: A

doi: 10.11959/j.issn.1000-0801.2015376

Interconnection technology for data center of telecom operator

TANG Lili, ZHANG Ning

Division of Network, China Mobile Group Design Institute Co., Ltd., Beijing 100080, China

Abstract: With the continuous development of cloud computing and big data, service requirement to compute and storage challenges the scale of single data center. Building a data center interconnection network has become a must to achieve the data center scale up and service deployment flexibility. By constructing data center interconnection network, telecom operators can further integrate IT resource as “cloud”, in order to increase the resource utility and fasten new telecom service deployment efficiency, so to enhance the core business competitiveness. The representative solution of data center interconnection and the technology characteristic was analyzed. Combined with the requests of telecom resource integration, several data center interconnection technologies and their deployment in telecom carrier cloud network constructions were analyzed.

Key words: cloud computing, VM, VDC, SDN, VxLAN

1 引言

面对大数据和各类云服务的爆炸式增长,单体数据中心已经无法满足业务发展对计算和存储资源的要求。电信运营商必须进一步考虑将多个数据中心互联,实现更大范围内计算和存储资源的池化。当前,越来越多的数据中心通过部署以 VxLAN(virtual extensible LAN,可扩展虚拟局域

网络)和 SDN(software defined networking,软件定义网络)技术为代表的网络虚拟化技术,成为云化的数据中心;另一方面,传统的非云化数据中心,也通过部署 VxLAN 和 SDN 技术,实现内部碎片化资源的整合和高效利用。

本文在全面分析数据中心网络需求的基础上,着重讨论电信运营商数据中心网络如何进行互联,结合电信运营商的资源整合需求,深入研究各种数据中心互联技术的应用。

2 数据中心业务和互联需求

以计算资源是否虚拟化作为分类标准,数据中心可分为云化数据中心和非云化数据中心。这两种数据中心都需要通过数据中心互联网络实现规模扩展和资源整合;同时由于承载业务的属性不同,其对数据中心互联网络的要求也有不同之处。

2.1 云化数据中心互联需求

云化数据中心通过引入分布式计算与虚拟化等核心技术,提高现有 IT 资源的利用效率,有效地提升单位机柜收入。同时,云化数据中心和 IP 骨干网相结合,将成为运营商网络架构发展的主要特征,成为未来运营商最主要的核心竞争力。云化数据中心的互联需求如下:

- 突破单个数据中心的规模约束,将多个数据中心互联来实现更大规模 IaaS 服务;
- 跨数据中心实现虚拟机热迁移、系统灾备、存储访问等业务;
- 数据中心互联网络需要具备对租户的感知能力,实现基于租户进行流量的按需调整和优化;
- 覆盖多数据中心的统一管理平台,从 IaaS 模式向 PaaS 模式和 SaaS 模式演进。

2.2 非云化数据中心互联需求

非云化数据中心,即未完成计算资源虚拟化改造的数据中心。该类数据中心的业务以出租机架及出口带宽为主体,机架的综合利用率多在 70% 左右,而每个机房的剩余机架无法满足新的租户对于机架数目的要求,所以需要对数据中心做简单改造实现多数据中心互联,盘活数据中心的机架“碎片”,并给业务运营和维护带来附加价值。非云化数据中心的互联需求如下:

- 将多个数据中心整合为基础设施资源池,原来的碎片化硬件资源在重新池化后可以灵活满足新租户的需求;
- 突破单个数据中心规模的约束,实现多局址灵活扩展;
- 把某个租户在多个数据中心的流量,统一调度到指定的出口网关,实现集中化的流量计费和 QoS 控制;
- 通过统一界面下发策略,实现网络设备的自动化配置方式,快速满足租户业务发展需求,简化运维管理流程。

3 DCI 建设方案

未来几年云计算业务将从根本上改变网络流量模型,

网络流向从传统意义上的南北流量,转变成了 IDC 到 IDC 之间或者是资源池之间的横向流量,随着 NFV 技术的引入,这一变化趋势将更加显著。

IDC 之间的流量的承载由 DCI(data center interconnection)负责完成,以满足数据中心之间的通信需求。DCI 网络作为东西向业务的承载网络,将多个数据中心内的资源进行横向打通。DCI 可以分为广域 DCI 和 AZ(available zone,可用区域)/VDC 之间内部不同 DC 之间的 DCI 两类。

3.1 DCI 互联技术简介

DCI 网络使用的互联技术主要包括面向广域网侧的 MPLS(multi-protocol label switching,多协议标签交换) L2/L3 VPN、EVPN(BGP MPLS-based ethernet VPN,基于 BGP MPLS 和以太网的 VPN)以及面向 DC 侧的 VxLAN;相关的流量调优技术包括 MPLS TE 以及 SDN 等。分别简单介绍如下。

SDN 作为一种新的网络架构,成为行业高度关注的热点。转发与控制分离、集中的控制层面、开放的网络编程接口是 SDN 架构的主要特征,开放式网络代表了从应用适应网络能力向网络能力主动适配应用需求这个网络建设理念的改变。

SR(segment routing,分段路由)是由 IETF 推动的支持 SDN 架构的新型路由转发协议,SR 是一种源路由机制,也是以标签交换为基础的,通过对现有的 IGP 进行简单的扩展来实现 TE(traffic engineering,流量工程)、FRR(fast reroute,快速重路由)、MPLS VPN 等功能。在未来的 SDN 架构中,SR 将为网络提供和上层应用快速交互的能力。

VxLAN 是为解决大规模云数据中心的租户隔离需求及通过三层网络实现二层打通的需求而出现的,24 byte 的 VNI(VxLAN network identifier,VxLAN 网络标识)可支持 1 600 万个 VxLAN 段的网络隔离。VxLAN 使用“MAC in UDP”的方法进行封装,数据分组在整个转发过程中保持内部数据的完整,使得原有数据报文内容作为 UDP 报文中的负载在二、三层网络中传送。

EVPN 是 RFC7432 定义的新一代 VPN 协议,可以实现 L2/L3 VPN 业务;与 MPLS L2/L3 VPN 比较,转发平面的选择更为灵活,可以为 MPLS/PBB(provider backbone bridging,骨干网桥接)、VxLAN/NvGRE(利用 GRE 实现网络虚拟化)等;控制平面统一为 MP-BGP,避免了 VPLS(virtual private LAN services,虚拟专用局域网业务)的洪泛问题。

MP-BGP 即 BGP-4 的多协议扩展,实现了对多种网络层协议的支持。采用地址族(address family)来区分不同的



网络层协议,既可支持传统的 IPv4 地址族,又可支持 VPN-IPv4、IPv6 等其他地址族。MP-BGP 可正确处理地址空间重叠的 VPN 路由,完成私网 VPN 路由传递和私网标签分配。

3.2 云化数据中心间的 DCI 建设方案

3.2.1 方案概述

虚拟机热迁移业务对传输时延和传输质量的要求,是云化数据中心 DCI 网络建设的一个主要约束。按照这一约束,需要按照区域内、跨区域两级建设云化数据中心的 DCI 网络。

(1) 区域整合

云化数据中心通过 DCI 网络打通之后,同一个资源池内不同机房/设施的数据中心组成了面向云业务的网络可用区域,即 AZ。AZ 内的云资源可以实现无缝共享,比如虚拟机的热迁移业务可以在 AZ 内实现。为满足热迁移业务,AZ 区域通常在 500 km 范围之内,以保证传输时延和传输质量。

(2) 跨区域整合

通过数据中心互连网络 DCI 将 AZ 之间打通,形成数个大范围区域之间、甚至电信运营商全国的云资源整合。跨区域的整合可以实现系统灾备等非在线业务。

云化数据中心 DCI 网络技术实现方案需要从数据转发平面和控制平面分别考虑。

3.2.2 数据转发平面

云化数据中心互连网络的数据转发平面如图 1 所示。

云化数据中心网络需要承载两类东西向数据流量,分别为 AZ 内 DC 之间的数据流量和 AZ 之间的数据流量。

(1) AZ 内若干个 DC 之间的转发平面

数据连接采用 VxLAN 封装,即数据报文在 AZ 网关上透传。VxLAN 隧道为服务器或 TOR 上端到端建立,通过

VxLAN 隧道实现不同 VM 之间的二层或三层互访,支持虚拟机热迁移。

(2) AZ 之间的转发平面

对于 AZ 之间的数据连接,业界设备厂商都有各自不同的实现方式,比较通用的是采用分段 VPN 的方式。AZ 内部的 VxLAN 在网关上终结,AZ 网关之间根据二层或三层业务需求部署 MPLS L2/L3 VPN 或 EVPN;分段 VPN 方式将租户的业务在 IP 骨干网上做到了可见和可管,是运维的基本条件。AZ 之间的 VPN 业务部署在 DCI 网络的 TE LSP 或 SR 路径上,同时考虑到骨干网对云化业务的可维护性,需要 AZ 网关设备支持对云业务的感知,然后通过广域网流量调优技术实现灵活的流量调度。

3.2.3 控制平面

云化数据中心互连网络的控制平面如图 2 所示。云化数据中心网络存在两套控制平面,分别为 AZ 内 DC 之间的控制面 and AZ 之间的控制平面。

(1) AZ 内 DC 之间的控制平面

通过 MP-BGP 将 SDN 控制器形成联盟,从而实现多个 DC 之间的租户 IP 地址和 MAC 地址信息传递。

(2) AZ 之间的控制平面

使用 MPLS L2/L3VPN,实现跨 AZ 的二层或者三层连接,通过广域网实现二层租户网络或者三层租户网络扩展;采用 MP-BGP 实现。

3.3 非云化数据中心间的 DCI 建设方案

非云数据中心通过 DCI 互联之后,可以省或者区域为单位,将一个省的 IDC 资源组成一个 VDC,然后通过 DCI 网络实现 VDC 之间的互联,从而实现全国性的 IDC 资源整合。非云化数据中心的互联方案如图 3 所示。

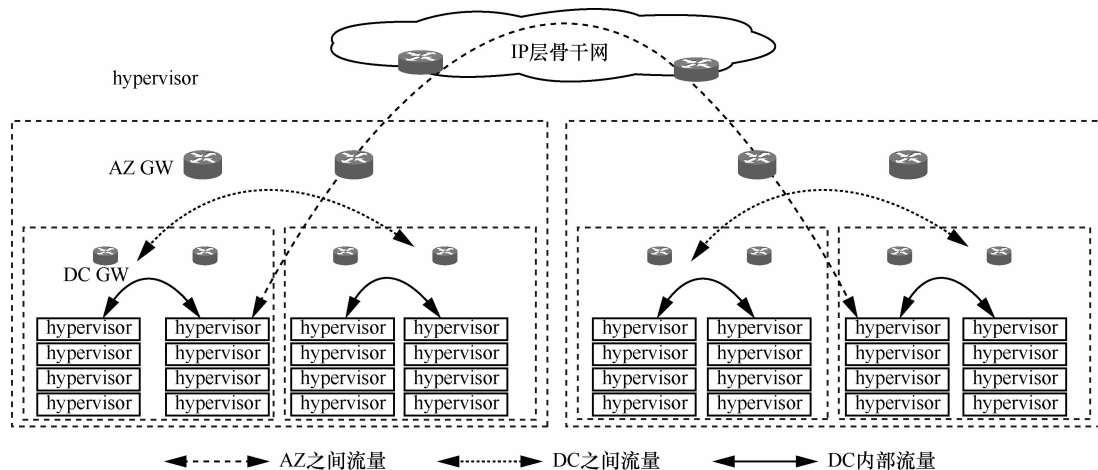


图 1 云化数据中心互连网络的数据转发平面示意

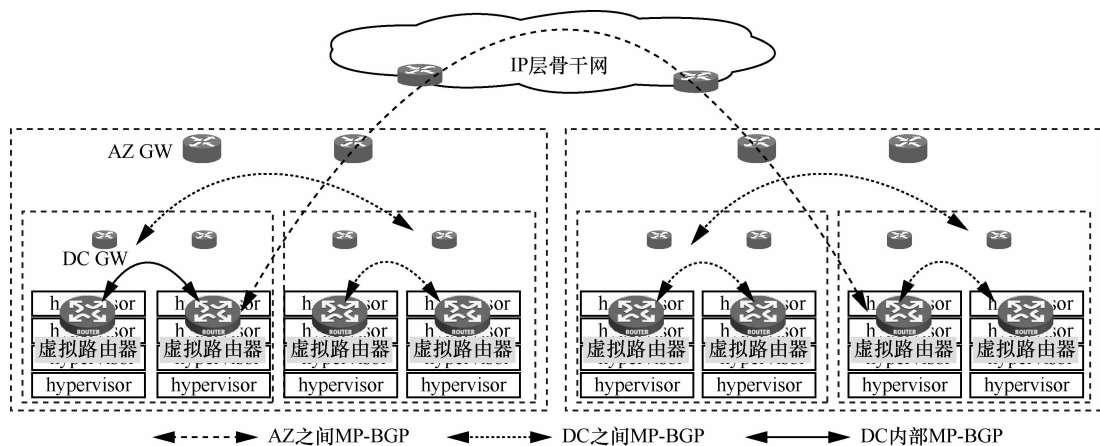


图2 云化数据中心互联网络的控制平面示意

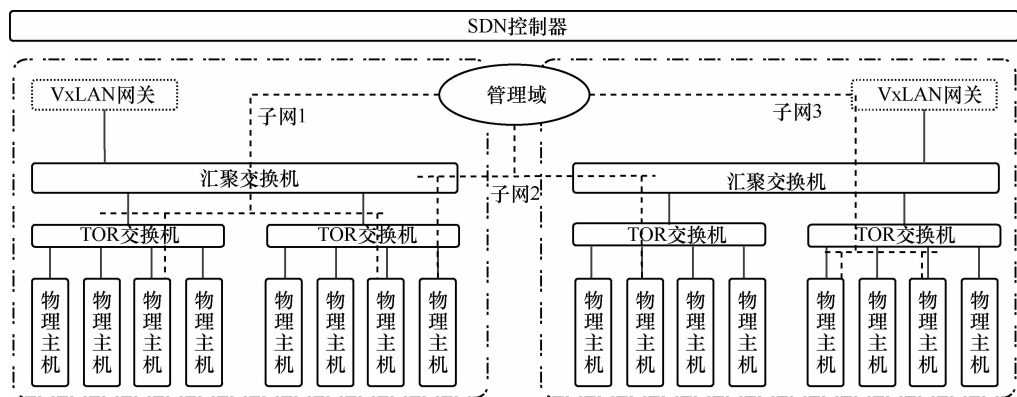


图3 非云化数据中心的互联方案示意

非云化数据中心网络存在两类东西向数据流量,分别为:VDC 内的 DC 之间的数据流量和 VDC 之间的数据流量,技术方案分别如下。

(1)VDC 内部

将多个物理区域分散的 IDC 在网络层面,通过 VxLAN 技术组成一个 VDC,实现区域内原有 IDC 的整合。通过 VxLAN 网关实现 VLAN 与 VxLAN 的映射,非云化的租户通过 VxLAN 的 VNI 进行标识。

(2)VDC 之间

VDC 通过数据中心互联网络 DCI (可采用 MPLS L2/L3 VPN 或 EVPN 技术)实现 VDC 间的互联,从而实现全国 IDC 资源的大整合。

3.4 DCI 的流量调优

电信运营商在完成云化/非云化数据中心的整合之后,更为重要的是如何将跨数据中心的资源在 DCI 网络中进行流量的合理调度。通过 DCI 网络识别业务并进行基于 SDN 的流量调度,真正实现端到端云化/非云化数据中心业务的统一调度与业务的快速响应。

DCI 骨干网流量调优的前提是对云业务的感知。如果骨干网边缘设备不能够识别租户,那么仅作为粗管道是无法针对云业务开展任何运维的。DCI 骨干网边缘设备需要支持的技术包括:对 VxLAN 的支持,对 EVPN 及 MPLS L2/L3 VPN 的支持以及广域网 SDN。

广域网 SDN 技术可以实现骨干网的自动化拓扑计算和更智能的流量调优。SDN 正处在标准化阶段,典型的技术如 SR、PCEP(path computation element protocol,路径计算单元协议)等。在协议完全标准化之前,现阶段也可借鉴 TE 技术,并与 SDN 控制方式相融合,实现现阶段的流量调优。关于广域网 SDN 的控制器,建议由电信运营商建设资源编排器,通过该编排器实现不同网络设备厂商级控制器的整合,实现骨干网的统一管理和资源调度。

3.5 DCI 建网的其他考虑因素

电信运营商在建设 DCI 网络时,除了考虑业务需求、网络技术、运营商自身条件和发展策略外,还需要考虑云计算环境下的其他相关因素,具体分析如下。

· 存储技术和存储网络的发展:不同的计算业务对存



储的要求不同,做 AZ 层面的资源整合时需根据存储网络的覆盖范围,合理划分 AZ 的区域覆盖大小。

- 自动化的管理方式:云计算时代的一个典型特性是虚拟机创建的频繁及其对网络的动态化适配要求。面对这种动态化的网络配置和管理要求以及运营商降低运维成本的实际需求,DCI 骨干网边缘设备需要具备基于 SDN 的自动化配置能力,从而大大节省业务的部署时间,实现与云平台的充分契合。
- 云管理平台的整合:随 DCI 网络的建设实现云管理平台的整合。从 AZ 区域角度出发,建议通过一套云管理平台实现 AZ 下所有虚拟机的管理,实现 AZ 内的一套云管理平台与一套 SDN 及其云网络管理平台对应,从而实现虚拟机的在线迁移。不同的 AZ 之间,可通过层次化云管理平台的方式实现整合,不同 AZ 之间可实现业务的冷备份。

3.6 DCI 网络建设总体思路

DCI 网络的建设,需要考虑云化和非云化两类数据中心的互联组网需求以及 DCI 网络的广域流量调优方案,并综合考虑公有云、私有云和混合云的业务特性,从而形成整体的网络建设方案,实现电信运营商的 IDC 资源整合。

对于电信运营商如何利用自身资源来建设 DCI,本文建议采用循序渐进的策略方案,具体思路如下。

- 分析业务特点和应用需求:根据自身的业务特点和应用需求对各类数据中心互联技术进行分析,评估每种技术方案的成熟度、业务契合度,以选择最适合自身业务发展的 DCI 方案。
- 选择广域网络承载方案:梳理现有 IP 网络的资源情况,包括网络带宽利用率、链路时延、承载业务的类型和特点等,以在使用现有骨干网或新建 DCI 骨干网之间做出选择。当使用现有 IP 骨干网实现 DCI 时,利用现有 IP 骨干网中继实现高速互联,同时升级现有路由器使其支持并部署 DCI 相关的 EVPN、MPLS VPN、SDN、SR 等新技术。
- 合理规划业务类型和承载方案:对数据中心承载的业务进行梳理和分类,并根据云管理平台的管理范围、存储网络的覆盖范围等,对 DCI 所承载的业务类型和承载方案做到合理的规划和设计。
- 先试点再推广,分步建设 DCI 网络:首先就近进行数据中心的区域整合,设立区域整合后的新网关,实现多出口变为统一单出口;然后根据跨区域的互联业

务需求,在该类业务需求量较大的区域之间进行 DCI 网络的试点,同时整合相关云管理平台;最后,在试点成熟的前提下,根据全国的 DCI 互联需求,建设 DCI 网络,并部署 WAN SDN 技术进行流量调优。

4 结束语

云计算技术的快速发展、数据中心的规模部署和 SDN 解决方案的日趋成熟,使数据中心互联成为运营商 IDC 集约化的重要手段。数据中心互联需要考虑的方面非常多,不同运营商基于自身条件会有不同的方案选择。随着数据中心和网络部署经验、云业务运营经验的不断积累,运营商的数据中心互联方案必然越来越成熟,必然给运营商的资源集约化和成本降低带来非常大的价值。

参考文献:

- [1] Eucalyptus System. Cloud computing and open source IT climatology is born[R]. [S.l.: s.n.], 2011.
- [2] 刘承鹏. 云计算及其应用的开源实现研究 [J]. 信息通信, 2015(10).
LIU C P. Research on the realization of open source cloud computing and its applications[J]. Information & Communications, 2015(10).
- [3] Virtual extensible local area network (VxLAN): a framework for overlaying virtualized layer 2 networks over layer 3 Networks: RFC 7348 [S/OL]. [2015-09-28]. <https://tools.ietf.org/html/draft-mahalingam-dutt-dcops-vxlan>.
- [4] BGP MPLS-based ethernet VPN: RFC 7432[S/OL].[2015-09-28]. <https://www.packetizer.com/rfc/rfc7432/>.

[作者简介]



唐利莉(1973-),女,中国移动通信集团设计院有限公司网络所高级工程师、高级咨询设计师,IP 承载网络专家,CMNet 骨干网可研、设计项目总负责人,主持与 CMNet 相关的诸多科研课题。主要研究方向为运营商 IP 承载网络架构演进、运营商 IP 承载网络发展规划及工程建设。



张宁(1982-),男,中国移动通信集团设计院有限公司网络所工程师、高级咨询设计师,CCIE 工程师,主要研究方向为运营商 IP 承载网络架构演进、运营商 IP 承载网络发展规划及工程建设。