

基于 MPLS VPN 和 MSDP 的跨域组播网络设计

陶 骏¹ 匡 磊¹ 徐 旺² 颜云生¹ 万家山¹

(安徽信息工程学院计算机软件与工程系 芜湖 241000)¹

(安徽信息工程学院信息工程系 芜湖 241000)²

摘 要 介绍了组播的基本原理,分析了 MSDP 技术和 MPLS VPN 技术,并针对一般视频服务企业的网络改造提出了 3 种跨域组播的方案,对每种方案的优缺点进行了比较;采用 MSDP 和 MPLS VPN 技术的跨域组播方案可降低网络的流量和提高网络流量的安全性和可靠性。

关键词 MPLS,MSDP,MBGP,流量,时延

中图分类号 TP393.2 文献标识码 A

Multicasting Network System Design of Uniting Domains Based on MPLS VPN and MSDP

TAO Jun¹ KUANG Lei¹ XU Wang² YAN Yun-sheng¹ WAN Jia-shan¹

(Department of Computer Science and Software Engineering, Anhui Institute of Information Technology, Wuhu 241000, China)¹

(Department of Information Engineering, Anhui Institute of Information Technology, Wuhu 241000, China)²

Abstract The article introduced the theory about multicasting, analysed the technology about MSDP and MPLS VPN. As to the network transformation of a normal enterprise on video, three multicasting schemes on uniting domains were proposed, and the priorities and shortcomings about them were compared. The scheme based MSDP and MPLS VPN is the best solution, which not only relieves the traffic of network, but also increases the security and reliability of network.

Keywords MPLS, MSDP, MBGP, Traffic, Delay

随着互联网业务的高速发展,网络视频企业的通讯流量负荷日益加重,企业被迫实时对网络进行扩容,增大了企业负担和经济压力;在运营商(ISP)网络运行高峰期,由于视频业务的可靠性和安全性得不到保障,因此企业网络需要改造,改造的目标是:降低网络的流量负荷,提高网络的安全性和可靠性。一个普通的网络视频企业的现状网络图如图 1 所示。

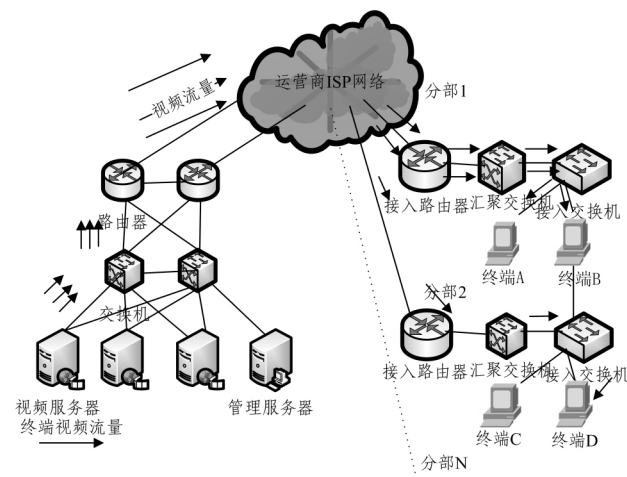


图 1 企业网络现状图

企业网络分为总部和分部,视频服务器都位于总部,用户都位于分部,分部用户向总部视频服务器发起请求,总部视频

服务器根据用户请求下发单播视频流,随着企业业务的发展,下发的视频流越来越多,耗尽了网络的带宽,严重影响了企业的经营与发展,具体的用户和业务流量如表 1 所列。

表 1 用户数目和企业带宽的关系

用户数	总部路由器出口 带宽/Gb/s	某分部路由器 出口带宽/Gb/s	某分部汇聚交换机 上行带宽/Gb/s
10000	20.4	1.8	0.8
15000	30.6	2.4	1.3
30000	58.4	4.6	2.7

网络流量的描述如下:总部路由器到运营商、某分部路由器到运营商和某分部交换机的上行流量为 $c_{总}$, $c_{分}$ 和 $c_{交}$,分部 i 的用户数目为 s_i ,分部的总数目为 n ,分部每个交换机 j 下的用户为 s_j ,每个分部的交换机数目为 m ,每个用户的视频流量为 a , $L_{总}$, $L_{分}$ 和 $L_{交}$ 为总部路由器、分部路由器和分部交换机的常用流量,包括网络管理流量等,近似为常数。

$c_{总}$, $c_{分}$ 和 $c_{交}$ 满足如下公式^[1]:

$$c_{总} = a * \sum_{i=1}^n S_i + L_{总} \quad (1)$$

$$c_{分} = a * \sum_{j=1}^m S_j + L_{分} \quad (2)$$

$$c_{交} = a * S_j + L_{交} \quad (3)$$

由式(1)一(3)可知,当用户数增加时,对应网络流量也将迅速增加。

本文受安徽高校自然科学重点项目(KJ2017A797),安徽高校自然科学重大项目(KJ2017ZD53)资助。

陶 骏(1978—),男,硕士,高级工程师,主要研究方向为网络管理和网络安全;匡 磊(1997—),主要研究方向为网络安全;徐 旺(1990—),硕士,助教,主要研究方向为网络通信;颜云生(1970—),博士,副教授,主要研究方向为信息安全和网络管理;万家山(1988—),硕士,讲师,主要研究方向为数据挖掘。

运营商网络运行高峰时期,企业的视频业务也会受到影响,具体的表现为从总部路由器到分部路由器的时延明显增大,严重影响了用户感知,具体如表 2 所列。

表 2 总部到分部的时延表

时间	00:00—04:00	04:00—08:00	08:00—12:00	12:00—16:00	16:00—20:00	20:00—24:00
时延/ms	30	26	35	48	85	94

采用组播技术降低了网络的流量负荷,使用 MPLS VPN 和 QoS 技术提高了网络的安全性,降低了关键业务的网络时延,使企业的工作效率得到进一步提高,而且节省了企业的投资。

1 组播和 MPLS VPN

组播指在发送者和多个接收者之间实现点对多点的网络连接^[1]。一台发送者给多个接收者同时传输相同的数据,只需复制一份相同数据包,提高了数据传送效率,减少了网络拥塞,解决了单播和广播方式效率低的问题。当网络中的某些用户有特定业务需求时,组播源(即组播信息发送者)仅发送一次信息,相关路由器借助组播路由协议为组播数据包建立树型路由,被传递的数据包在尽可能远的分叉点才开始复制和分发,降低了网络的负荷。

组播协议包括组成员关系协议(路由器和主机之间)和组播路由协议(路由器之间)。组成员关系协议为 IGMP(组管理协议)。组播路由协议分为区域内组播路由协议和区域间组播路由协议^[2],区域内组播路由协议包括 PIM SM(组播稀疏协议模式)、DVMRP(距离向量组播路由协议)等协议,区域间组播路由协议包括 MBGP(多协议边界网关协议)、MSDP(组播路由源发现协议)等。

MPLS VPN(多协议标签交换虚拟网)是指采用 MPLS 技术在宽带 IP 网络上构建企业 IP 虚拟专网^[1,13],实现跨地域、安全和可靠的数据、语音和图像等多业务通信,并结合差别服务、流量工程等相关技术,将公众网可靠的性能、良好的扩展性、丰富的功能与专用网的安全灵活、高效地结合在一起,为用户提供高质量的网络服务。MPLS VPN 的网络在功能上将相关路由器分为 CE(客户路由器)、PE(核心边缘路由器)和 P(核心路由器)。

2 网络实现

采用组播技术能有效地减少网络的流量,经过论证决定在路由器之间采用 PIM SIM 协议,主机和路由器之间的组成员关系采用 IGMP V2 协议^[3],由于公司分部位于全国各地,总部和分部路由器的互联电路涉及多家运营商,因此涉及方案时必须考虑到组播跨域部署的问题,为解决此问题设计了 3 个方案。

方案 1 采用跨域组播协议 MSDP,具体网络拓扑如图 2 所示。总部路由器之间采用 PIMSM 协议,总部路由器和分部路由器充当每个区域的 RP 路由器,路由器和交换机之间、交换机和交换机之间、交换机和主机之间采用 IGMP 协议,视频流量全程采用组播方式传输,此方案的优点是:全程采用了组播发送数据,节省了带宽,对网络的改变小,投资较小;缺点是:由于分部位于全国各地,需要和各地运营商进行协商

配合以调整网络配置,运营商的网络变动后,企业的网络会受到影响,接入运营商网络未采取 QoS(服务质量保障),网络安全性得不到保障^[11-12]。

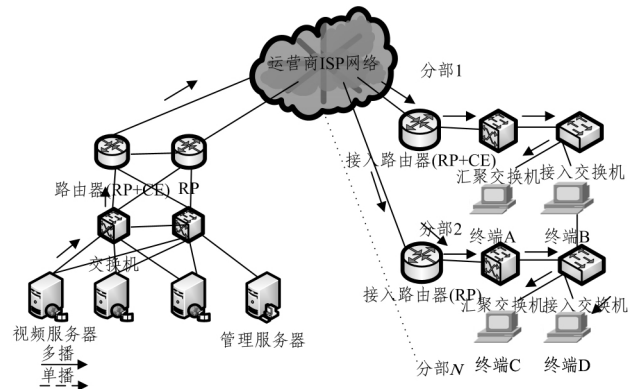


图 2 采用 MSDP 组网的方案

方案 2 采用区域组播的方法,具体的拓扑图如图 3 所示。

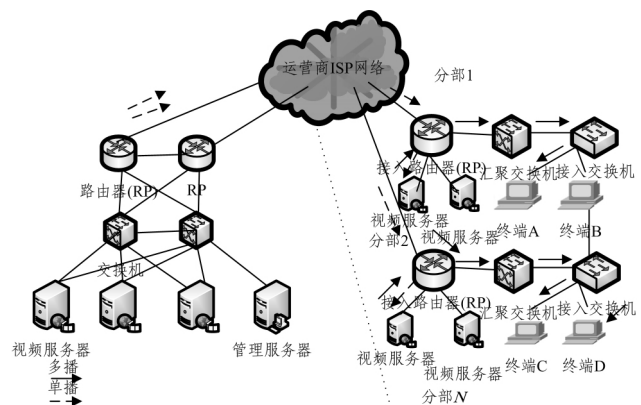


图 3 采用区域组播的方案

方案 2 和方案 1 的区别在于,在方案 2 中:总部和分部路由器没有采用 MSDP 协议,分部增加了视频服务器节点,视频流量的走向是为总部视频服务器以单播的方式发送给分部视频服务器,分部视频服务器再以组播的方式发送到各个节点。此方案的优点是:配置简单,无需通讯运营商配合;分部内采取组播发送数据,节省了带宽;缺点是:需要为每个分部新增视频服务器节点,投资较大;接入运营商网络未采取 QoS。

方案 3 采取 MPLS VPN+MSDP 的方法,具体的网络拓扑图如图 4 所示。

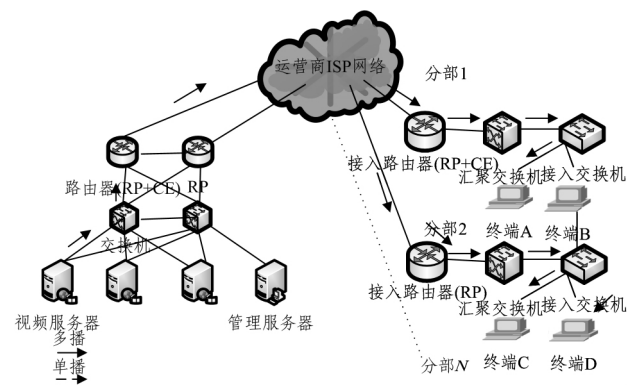


图 4 采用 MPLS VPN+MSDP 的方案

方案 3 是在方案 1 的基础上进行的改进,总部和分部的路由器既充当组播中的 RP(组播集中点)路由器^[4],又充当

MPLS VPN 中的 CE(客户端)路由器,总部和分部的路由器采取 hub and spoke(中心辐射)的方式接入到运营商的网络,只允许总部路由器和分部路由器之间通讯,禁止分部路由器之间相互通讯,对应的运营商网络需要增加 MPLS VPN 配置,并根据公司业务流量的 IP 数据包的 DSCP(差分服务代码点)值做 QoS 保障。此方案的优点是:全程采用组播发送数据,节省带宽;采用 MPLS VPN 发送业务流量,保证了业务流量的安全性和可靠性;缺点是:需要追加投资向运营商购买 MPLS VPN 电路业务,运营商的相关网络配置需要做相应的修改。

3 网络仿真与分析

Waxman 随机图模型广泛应用于路由算法的模拟实验,根据此模型,随机图的节点随机分布在矩形区域内^[12-13],节点 u 与 v 之间存在边的概率为:

$$P_e(u,v) = \beta \exp\left(-\frac{d(u,v)}{\alpha L}\right)$$

其中, $d(u,v)$ 为节点 u 到节点 v 的欧拉距离, L 为节点间最大欧拉距离值,参数 α, β 从区间 $(0,1)$ 间取值, α 越大,距离远的节点间存在边的概率就越大,而 β 越大,网络就越稠密。在模拟实验中取 $\alpha=0.5, \beta=0.3$ 和 $\alpha=0.7, \beta=0.2$ 两种情况,为了更好地比较算法的优越性,作如下的约定。

方案 1 和方案 3 的组播树相同,方案 2 的组播树不同,定义 NMtree 和 KMtree 表示采用方案 3 和方案 2 在同一图中生成的组播树, $C(\text{NMtree})$ 和 $C(\text{KMtree})$ 表示生成两棵树的代价。

目的节点密度 = (目的节点数 + 节点总数) * 100%, 设定每个网络中的节点密度分别为 5%, 20%。

定义 $\delta_i = (C(\text{KMtree}) - C(\text{NMtree})) / C(\text{KMtree})$, $\delta = (\sum_{i=1}^{100} \delta_i) / 500 (0 < i \leq 100)$, δ 为采用方案 3 的优化百分比。

实验在给定节点数目和目的节点密度后,随机生成 100 幅图。比较方案 3 和方案 2 对这些图构造的组播树代价和算法执行时间,实验结果如下。

$\alpha=0.5, \beta=0.3$, 节点密度为 5% 时,结果如图 5 和表 3 所示。

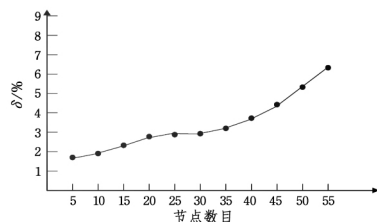


图 5 方案优化百分比示意图 1

表 3 方案比较 1/ms

节点数	10	20	60	100
方案 3	368	780	28432	112025
方案 2	240	470	20043	102350

$\alpha=0.5, \beta=0.3$, 节点密度为 20% 时,结果如图 6 和表 4 所示。

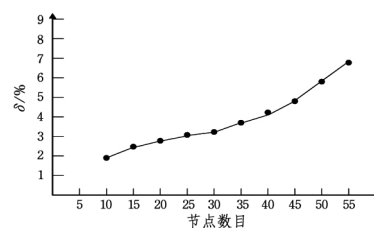


图 6 方案优化百分比示意图 2

表 4 方案比较 2/ms

节点数	10	20	60	100
方案 3	420	842	29872	133012
方案 2	273	491	22056	115764

$\alpha=0.7, \beta=0.2$, 节点密度为 5% 时,结果如图 7 和表 5 所示。

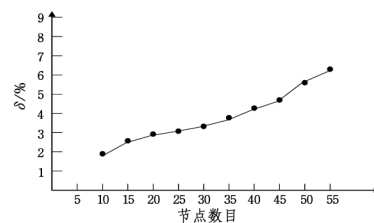


图 7 方案优化百分比示意图 3

表 5 方案比较 3/ms

节点数	10	20	60	100
方案 3	386	801	28771	120011
方案 2	256	469	19982	105538

$\alpha=0.7, \beta=0.2$, 节点密度为 20% 时,结果如图 8 和表 6 所示。

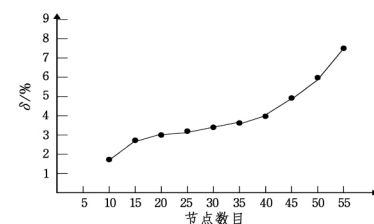


图 8 方案优化百分比示意图 4

表 6 方案比较 4/ms

节点数	10	20	60	100
方案 3	450	1002	31228	123018
方案 2	325	672	23087	108850

如上所示,随着节点数的增多,方案 3 与方案 2 相比能优化的量也越明显,当节点数为 55 时,方案 3 与方案 2 相比能优化 6.97% 左右。随着节点数的增加,方案 3 的执行时间优于方案 2 的执行时间。

通过实验模拟可以发现,网络越复杂越能体现本算法的优越性。通过本算法构造的组播树在最好情况时,相比方案 2 代价可以优化 6.97%,最坏情况不会差于方案 2 构造的组播树。从算法的执行时间来看,方案 3 比方案 2 更优。因此,对于复杂的网络应用方案 3 要优于方案 2。

(下转第 287 页)

Veh. Tech., 2016, 65(1): 204-216.

- [8] WANG D, FU S L, LU K J. Channel coding design to support asynchronous physical layer network coding [C] // Proc 28th IEEE Conference on Global Telecommunications, USA, IEEE, 2009: 3267-3272.
- [9] ZHANG S L, LIEW S C. Channel Coding and Decoding in a Relay System Operated with Physical-Layer Network Coding [J]. IEEE J. Sel. Areas Commun., 2009, 27(5): 788-796.
- [10] WIIBBEN D, LANG Y D. Generalized Sum-Product Algorithm

for Joint Channel Decoding and Physical-Layer Network Coding in Two-Way Relay System [C] // Global Communications Conference, 2010, Florida, USA, 2010: 1-5.

- [11] NAMBOODIRI V, RANJAN B S. Physical Layer Network Coding for Two-Way Relaying with QAM and Latin Squares [C] // Global Communications Conference, IEEE, 2012: 2286-2292.
- [12] QIN J, YANG Z, JIAO J, et al. On Symbol Mapping for FQPSK Modulation Enabled Physical-layer Network Coding [C] // IEEE WCNC, 2013.

(上接第 265 页)

综合投资、安全和可靠性等各种因素后,决定采用方案 3 作为网络实现方案,经过改造后的网络流量如下:

$$c_{\text{总}} = a + L_{\text{总}} \quad (4)$$

$$c_{\text{分}} = a + L_{\text{分}} \quad (5)$$

$$c_{\text{交}} = a * \beta * S_j + L_{\text{交}} \quad (6)$$

其中, β 为组播复制相关系数,相关交换机是组播树的最后一个分叉点时才对组播数据包进行复制, β 就是指交换机是最后一个分叉点的概率^[5-7]。

由上述 3 个公式可以看出,网络的流量已经显著降低。

在采取方案 3 后,总部出口路由器具体的网络流量改造前后对比如表 7 所列。

表 7 改造前后总部出口路由器流量对比

用户数	总部路由器 出口带宽/ Gb/s	某分部路由器 出口带宽/ Gb/s	某分部汇聚 交换机上行带宽/ Gb/s
30000	58.4	4.6	2.7
30000(改造后)	15.3	1.3	0.8

改造后总部出口路由器的流量已经明显降低。由于采取了 MPLS VPN 技术,企业的视频业务得到了 QoS 保障,从总部到分部的时延在运营商的网络高峰期已经正常,具体对比如表 8 所列。

表 8 改造前后总部到某分部的时延对比

时间/ms	00:00— 04:00	04:00— 08:00	08:00— 12:00	12:00— 16:00	16:00— 20:00	20:00— 24:00
时延/ms	30	26	35	48	85	94
时延 (改造后)	27	24	26	28	29	29

采用方案 3 后,企业的流量拥塞和传输时延较大的问题得到了有效的解决。

结束语 随着因特网的飞速发展,视频点播业务成为当前用户的热点业务,组播技术可以有效地减少网络流量^[8],提高网络传输效率,MSDP 协议可以满足跨域组播的实现,MPLS VPN 技术保障了在因特网上传输数据的安全性和可靠性。本文通过 MSDP 组播和 MPLS VPN 技术实现了某企业的跨域组播的网络设计,提高了企业的网络可靠性和安全性,降低了关键业务的网络时延,使企业的工作效率得到进一步提高,而且节省了企业的投资,为类似企业采用组播技术提供了参考方案。

IPv6 和移动网络应用^[9]是不可逆转的大趋势,上述方案暂时还不支持 IPv6 和移动网络^[10],方案还需要进一步优化,要能适用于 IPv6 和移动网络,上述方案还不能有效地控制组播数据包的 UDP 传输时延,这也是下一步的研究方向^[14-16]。

参 考 文 献

- [1] WILLIAMSON B. IP 组播网络的设计与部署 [M]. 北京:人民邮电出版社, 2011: 476-482.
- [2] COMER D E. 用 TCP/IP 进行网络互连(第二卷) [M]. 北京:电子工业出版社, 2009: 149-157.
- [3] 谢希仁. 计算机网络(第六版) [M]. 北京:电子工业出版社, 2013: 169-176.
- [4] 刘莹. Internet 组播体系结构 [M]. 北京:科学出版社, 2008: 100-105.
- [5] 程春玲. 一种支持组播的 SNMPv3 改进模型 [J]. 计算机科学, 2012, 39(4): 89-93.
- [6] ZHAO Q L, TSANG D H K, SAKURAI T. Modeling Nonsaturated IEEE 802.11 DCF Networks Utilizing an Arbitrary Buffer Size [J]. Transactions on Mobile Computing, 2010, 9(8): 1-15.
- [7] PRAKASH G, THANGERAJ P. Throughput Analysis of IEEE 802.11b WLAN Under a Nonsaturated Condition [C] // IEEE International Test Conference, Marriott AUSTIN, 2010: 65-69.
- [8] LIU T H, LIAO W J. Multicast routing in multi-Radio Multi-Channel wireless mesh networks [J]. IEEE Transactions on Wireless Communications, 2010, 9(10): 3031-3039.
- [9] DULAIMI A A, RAWESHIDY A H, COSMAS J, et al. Cognitive mesh networks [J]. IEEE Vehicular Technology Magazine, 2010, 5(3): 54-60.
- [10] PEI T R, ZHAO Z, ZENG W L, et al. A cognitive improved hierarchical AODV routing protocol for cognitive wireless mesh network [J]. Information Technology Journal, 2011, 10(2): 376-384.
- [11] GUO J W, ZHOU X W, MIAO X N, et al. Secure minimum-energy multicast tree based on trust mechanism for cognitive radio networks [J]. Wireless personal Communications, 2011, 62(9): 1-19.
- [12] XIE L M, JIA X H, ZHOU K X. QoS multicast routing in cognitive radio ad hoc networks [J]. International Journal of Communication Systems, 2012, 25(1): 30-46.
- [13] TELATER I E. Capacity of Multi-antenna Gaussian Channels [J]. European on Telecommunications, 1999, 10(6): 585-595.
- [14] ZHENG L Z, TSE D N C. Diversity and Multiplexing: A Fundamental tradeoff in Multiple Antenna Channels [J]. IEEE Trans. on Inform Theory, 2003, 49(5): 1073-1096.
- [15] SHIAH C Y, YEN Y S. Compression of Chinese Document Images by Complex Shape Matching [J]. The Computer Journal, 2013, 56(11): 1292-1304.
- [16] YEN Y S, HSIAO W T, HSU H M, et al. An M-Taiwan experience on mobile WiMAX [J]. International Journal of Autonomous and Adaptive Communications Systems, 2011, 4(3): 252-270.