



IP 城域网业务接入控制层业务备份技术探讨

黄 晋,薛 焱

(中国电信股份有限公司苏州分公司 苏州 215004)

摘 要:越来越多的电信业务通过 IP 网络承载,对 IP 城域网的可靠性提出了极高的要求。业务接入控制层在 IP 城域网中是多业务融合承载的关键设备,在网络中承担着承上启下的功能。根据运营实践和业务测试,分析了业务接入控制层设备业务备份目前的技术现状,提出了在新型承载网络架构中构建高可靠性业务接入控制层的方法和建议。

关键词: IP 城域网;业务接入控制层;业务备份技术;高可靠性

doi: 10.3969/j.issn.1000-0801.2013.09.035

Business Backup Technology Research on Service Access and Control Layer of IP Metropolitan Area Network

Huang Jin, Xue Yao

(Suzhou Branch of China Telecom Co., Ltd., Suzhou 215004, China)

Abstract: As more and more telecom services carrying over the IP network, the IP MAN needs higher reliability requirements. The service access control layer is the key equipment of converging multi-service in IP MAN, and plays a connecting function in the network. Based on the operational practice and operational testing, the current state of technology of the service access control layer equipment business back up was analyzed. The methods and recommendations with building high reliability service access control layer in the new architecture of bearer network were proposed.

Key words: IP metropolitan area network, service access and control layer, business backup technology, high reliability

1 概述

随着电信运营商全业务经营时代的到来,新兴的 IPTV、NGN/3G、大客户专线和 VPN 互联等重要的电信业务均已经由 IP/MPLS 网络承载,面向 IP 化融合业务承载的 IP/MPLS 网络将成为全业务承载网络。业务和用户 IP 网络的故障恢复时间有了更加严格的要求,甚至要求延续以往传统电信网络 50 ms 级的标准。为减少用户明显的故障感知,避免用户对电信网络不信任进而形成客户流失,采用新技术提高业务的可靠性,提升用户体验,已成为运营商在激烈市场角逐中,提升竞争力的关键因素。

在 IP 城域网的层次结构中,业务接入控制层位于网络的二层与三层交汇处,向上接入城域核心,实现三层路由功能,负责数据高速转发;向下接入汇聚层,终结用户二层报文,实现用户的多种业务汇聚和接入控制。它承上启下,是 IP 城域网由简单的互联网接入网络向电信级可运营的城域网转型所需的关键设备,是城域网的枢纽节点。城域网架构如图 1 所示。

2 业务接入控制层运维现状

电信城域网经过多年建设,已经承载了百万级的业务和用户,业务接入控制层设备也从原先的小容量宽带接入

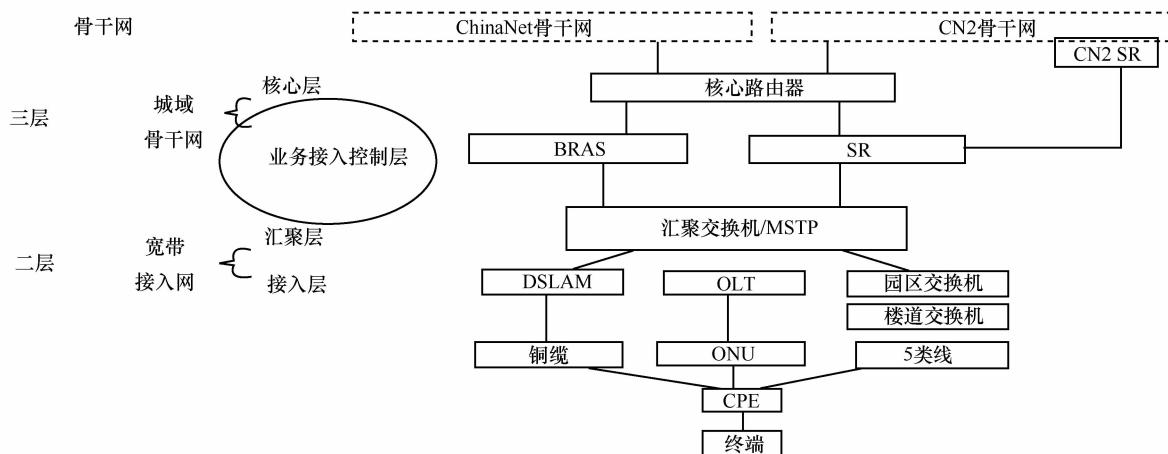


图1 城域网架构

服务器(broadband remote access server, BRAS)、小型三层交换机过渡到了大容量、高密度、集中化的 BRAS 设备和业务路由器(service router, SR)设备。这种集中式大容量设备有利于运营商快速地开展业务,抢占市场,同时降低建设成本,减少运维节点。但是随着宽带用户数的不断增加,每台 BRAS/SR 设备接入的用户数从最初的几千户增加到几万户,并发用户数也已经过万,如一台设备发生故障,就会造成众多用户的网络业务中断,造成大规模用户报障和投诉,严重影响运营商的服务质量和客户满意度。

电信级业务对于 IP 网络的可靠性要求主要由以下 3 个层面实现:设备板卡冗余、链路冗余和网络冗余。主流厂商的设备目前均采用了关键板卡的冗余和热倒换设计,包括主控、矩阵、时钟、电源、风扇等,还采用了各类板卡模块的热插拔设计以及控制和数据相分离的独立总线系统设计,为满足设备级的可靠性要求提供了基础。但只有少数厂商实现了业务板卡的冗余,大部分厂商设备的业务板卡故障后只能通过更换损坏部件来恢复,因此业务恢复时间往往在 1 h 以上。

目前电信网络主要通过增加 BRAS/SR 下行冗余链路的方法,形成接入网设备多链路上联 2 台或多台 BRAS/SR 的拓扑结构,实现链路级冗余。通过加大对主备用链路通断和质量劣化的监控力度(例如监控端口收发光功率变化,进行持续的 ping 检测)可以及时发现链路故障,快速利用冗余链路进行业务调度处理。由于省却了赶赴现场的时间,恢复业务的时间被缩短到数十分钟,不过如果需要调度的业务较多,也会超过 1 h。

链路冗余不能实现业务故障后的自动切换调度:对

PPPoE 方式拨号接入的业务可以实现半自动快速恢复,即当主用链路中断或者主用设备故障后,用户终端主动发起一次新的拨号就可以切换到备用链路或设备上恢复业务,业务恢复时间可以缩短到分钟级别。当切换的用户业务过多时,受限与备用设备的能力(拨号连接速度、地址池地址不足等),冗余也不能完全成功。对于固定 IP 地址接入和 DHCP 方式接入的业务则无法自动快速恢复。

随着客户服务要求的不断提高,随着用户对上网的依赖性不断加强,用户对网络故障造成的业务中断的容忍度不断降低。根据客服部门的统计,一般在 BRAS/SR 发生故障,业务中断后 2~5 min 就有大量用户报障,30 min 后开始有用户第二次催修障碍。综合考虑用户的容忍限度和客服响应能力的极限,一般 BRAS/SR 故障发生后在半小时之内恢复业务,用户反响较小,服务压力不大。而目前的冗余备份技术还不能达到这样的运维和客户服务要求。

3 业务接入控制层备份技术

业务接入控制层设备的备份技术,可以分为设备内备份和设备间备份两种。

设备内备份是将用户的业务通过两条链路接入到同一台 BRAS/SR 的不同端口,当一条链路或端口发生异常后业务能继续通过另外一条链路转发,主要分为同板卡链路备份和跨板卡链路备份两种。除了链路的冗余外,设备也需要将用户的完整接入信息同步到备用端口或板卡上。这种方案无法解决 BRAS/SR 设备发生异常后的业务恢复问题。

设备间的备份是将用户的业务通过两条链路接入到两台不同的 BRAS/SR 设备,当一条链路、一块板卡甚至一



台设备发生异常后,业务可以通过冗余备份继续转发。这种方案可保证在任何异常情况下快速恢复业务。

根据备份时用户业务是否中断的情况,设备间的备份又可分为热备份和冷备份。

热备份的含义是:固定 IP 地址、DHCP 和 PPPoE 等各种方式接入的用户均不中断,PPPoE 用户不需要重新拨号,更重要的是倒换后用户信息不丢失、倒换前后 QoS 调度一致。

冷备份的含义是:DHCP 用户和 PPPoE 用户都掉线,PPPoE 用户需要重新拨号,主要作用是提供一个业务备用通道。

从业务接入控制层的运维现状可知,有必要选择一种有效的备份技术,使得一条链路、一块板卡甚至一台 BRAS/SR 设备出现故障时,能将它承载的业务无缝切换到另一台正常的设备上,既使用户业务体验基本无损伤,又让运营商有足够的时间对故障设备进行维修和更换,从而实现电信级的网络冗余。从和厂商交流以及实际测试的情况看,业界最接近上述目标的备份技术就是设备间的热备份技术。

2011 年末和 2012 年初,笔者组织了几个主流厂商的 BRAS/SR 产品,进行了热备份技术交流和测试。下面就提高业务可靠性的几个主要问题,分析与探讨业务接入控制层热备份的几个关键技术要点。

4 业务接入控制层设备间热备份技术

根据交流和测试的情况,目前的 BRAS/SR 设备间热备份技术在虚拟路由冗余协议(virtual router redundancy protocol,VRRP)基础上,增强了主备选择功能,扩展了用户信息同步备份及业务控制功能。

业务接入控制层设备间的热备份分为两个层面:一是下行二层网络侧,这也是 BRAS/SR 热备份技术的重点和难点所在;二是上行三层网络侧,这一侧发生单条链路故障不会引起设备间的主备切换,而是由路由协议进行快速的路由收敛,保证转发正常,如果是全部链路故障,则引发 VRRP 切换和业务主备切换。

实现 BRAS/SR 设备间备份机制,主要有 4 个环节。

4.1 选择并建立主备控制状态

原来的设计中 BRAS/SR 的热备一般以物理端口为最小备份单元,根据业务环境要求,经过实际测试,选择了以 1 个端口下的 1 个 VLAN 为精细控制单元的热备最小组。主备控制的选择是在 VLAN 逻辑接口上启动扩展的 VRRP,从而协商出这个接口的主备状态,把这个接口下的用户关联到一个 VRRP 保护组,基于保护组状态决定用户

状态。由于具备了这样的功能,会选择 2 台 BRAS/SR 设备互为主备,以充分利用设备的资源。

对于 2 台互为主备的 BRAS/SR 设备,笔者发现如果用标准的 VRRP 仅对节点故障有效,对于链路故障容易产生 2 台 BRAS/SR 同时升为主设备的情况,导致回程流量的部分或全部丢失。只有采用扩展的 VRRP 综合检测 BRAS/SR 与汇聚交换机下行链路的通断状态以及 2 台 BRAS/SR 之间的 VRRP 连通状态,才可以准确选择 BRAS/SR 的主备状态。

此外标准 VRRP 主备倒换,要等到 VRRP hello 报文超时,按照默认规范需要大约 3 s 才能完成主备状态的切换,远远不能满足热备份需求。因此热备技术利用链路层检测协议加快 VRRP 主备状态的检测和倒换,链路检测协议采用 BFD 或以太 OAM,具体检测时间和链路检测协议的设置有关。热备技术采用 BFD+VRRP 综合判断选择业务的主备如图 2 所示。

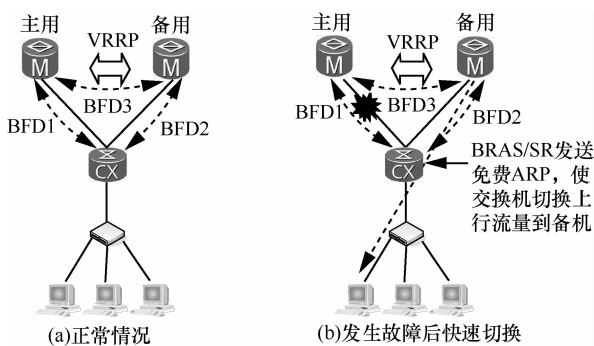


图 2 热备技术采用 BFD+VRRP 综合判断选择业务主备

4.2 进行用户信息备份

一旦选定了设备主备关系,主备设备间就要通过特定的备份协议同步和备份用户信息。

(1) 备份的信息内容

业务接入控制层设备作为业务终结点设备,需要处理大量用户信息,包括以下内容。

- 用户认证、授权信息:如用户的 MAC 地址、session ID、IP 地址、用户名信息、认证信息、Option60 信息等;保证倒换后用户无需重新拨号。
- 计费信息:如计费 ID、流量信息、时长信息;保证倒换后用户话单不丢失。
- 用户的 QoS 信息:如用户优先级信息、QoS 模板信息;保证倒换后用户体验不变。
- 用户物理位置信息:如内外层 VLAN 信息、Option82 信息;保证倒换后用户可正常上线。

这些信息都需要在两台设备间进行同步传递。

(2) 备份的通道

为保证信息传递有效,主备 BRAS/SR 间通常采用 TCP 进行信息备份,传输用户信息数据。经过测试,TCP 连接可以利用 2 台设备的上行链路通过城域网核心网络建立,也可以在 2 台 BRAS/SR 间建立专用链路。

(3) 备份的方式

目前设备上可以实现的备份方式有以下 3 种。

- 实时备份: 单个用户上下线或用户某个信息变化,其用户信息即时备份到备用设备上。
- 批量备份: 首次部署双机热备份,新增备用设备启用后,将会对主用设备所有用户进行批量备份;主用设备故障(整机故障)恢复后,需进行业务回切,此时主用设备没有用户信息,所以备用设备必须批量备份用户信息到主用设备。
- 周期性备份: 正常业务情况下,若一定时间内没有用户信息变化,也需要对用户的计费信息(流量和时长)进行周期性备份。

目前 BRAS/SR 设备备份用户信息的性能号称可以达到每秒备份 1 000 个用户,但是测试中没有测量大量用户实际在线情况下的备份时间。

4.3 发生异常切换时进行流量控制

对于下行到 BRAS/SR 的流量,通过路由控制策略进行流量控制。

由于切换不一定是整台设备一起发生的,如果每个主备接口能够规划一个地址池网段,则通过主发布网段路由、备撤销网段路由的方式控制下行流量的路径。

如果多个主备接口需要共享一个地址池网段(在现网的环境中,地址池中地址数量是较少的,通常是共享地址池),实际选择的方案是主备设备均发布网段路由,主用的路由权值设定为优选,同时在主备设备之间建立一个保护隧道。这样下行到已切换用户的流量,先到达原来的主用设备,通过隧道转发到现在的的主用设备上,再出隧道送达用户。

当网络侧链路故障或主用设备整机故障时,保护隧道 Down,原主用设备会撤销其宣告的地址池网段路由,现在的主用设备宣告的地址段生效,下行流量直接从城域网核心到达现在的主用设备。直到保护隧道重建建立并 Up,原主用设备才会再次发布地址池网段路由。下行流量的控制策略如图 3 所示。

对上行到 BRAS/SR 的流量而言,主备 BRAS/SR 设备使用相同的 MAC 地址与用户建立会话。当业务发生切换,设备从备用状态升为主用时,将向下发送免费 ARP 报文,刷新下

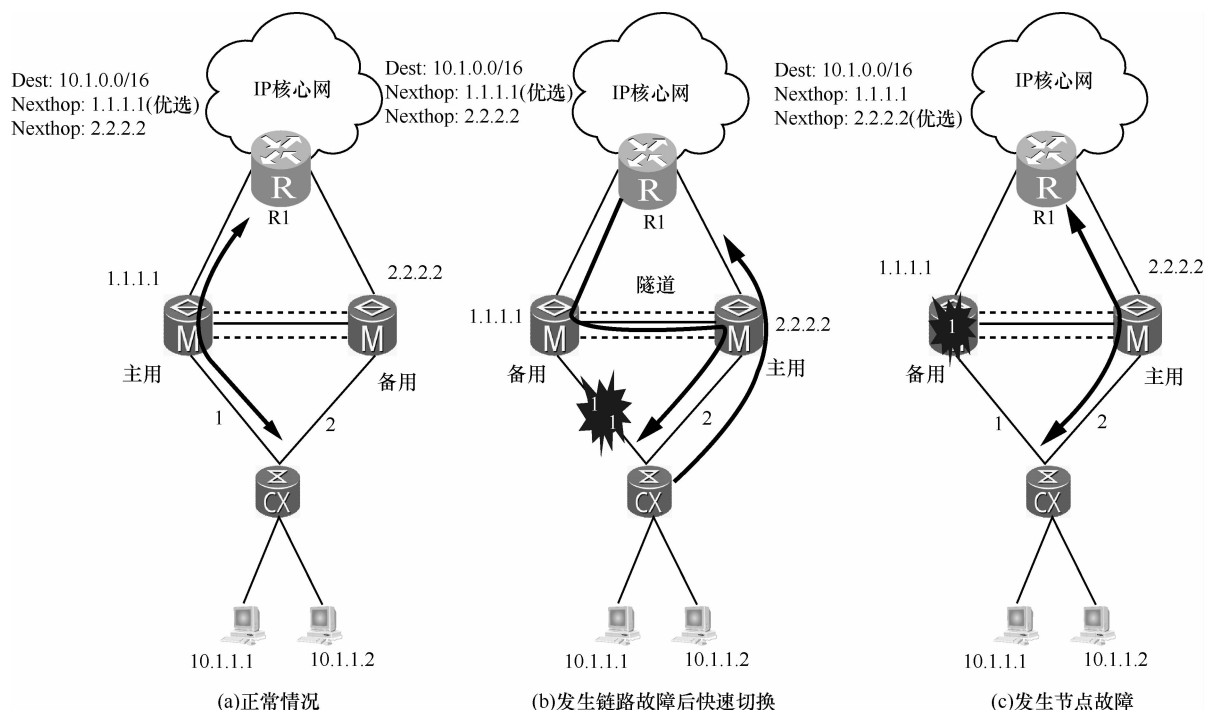


图 3 下行流量的控制策略(路由策略选择+隧道引流)



行交换机的 MAC 转发表,以保证用户流量被正确转发。

另一方面,主用 BRAS/SR 会根据 IP 会话信息对用户报文进行 MAC+IP+VLAN/PVC 的检查以及 QoS 调度,用户感知不到网络发生故障。

4.4 恢复正常时进行回切

当原来的主用链路、板卡、设备恢复正常后,设备经过检测,启动回切操作。首先实施批量同步,然后触发 VRRP 的倒换,最后通过流量控制策略和免费 ARP 发送将用户上下行流量回切。

通过测试发现,在规划为主用的一侧,一定要配置 VRRP 延迟回切,延迟时间设定应较长(10 min 以上);在规划为备的一侧,不配置延迟回切。配置该功能的目的是因为出现单板故障、整机故障时,用户业务切换时间比较长,如果过早回切会造成业务中断;同时防止链路重复通断或分组丢失严重时 BFD 和 VRRP 频繁切换带来的业务质量下降。

5 热备技术的主要难点

通过和厂商的交流以及实际测试,发现目前热备技术还存在一定的技术难题有待克服。

(1) 用户信息备份

目前各厂商的主备设备间用户信息备份均采用了自己的私有协议,因此不具备异种设备之间相互备份的能力。

主备设备之间实时备份及批量备份是否成功,目前也无法进行比较有效的确认。如果出现备份不完整的情况,设备既无法及时发现,更不能给出告警来通知维护人员。因此在主备切换过程中,存在用户信息不全,导致用户业务中断的风险。

当设备或链路发生故障,实施了主备切换后,新的主备设备之间是否还存在正常的实时备份及批量备份操作?从现有技术角度看,当 BRAS/SR 设备下行端口、链路或者板卡故障的时候,两台设备本身没有出现致命性的问题,那么信息同步备份的工作是可以继续的。而当原来的主用设备本身出现了故障,同步备份已经没有了目标,只有等原来的主用设备恢复正常后,才能通过批量备份来将用户信息全部备份过去。

因此,用户信息备份协议必须要首先完善备份状态检查和告警功能,能够实时正确地检查和显示主备设备之间的用户信息备份状态,主动提示不成功的备份。其次要根据设备运行的实际状态,充分考虑物理链路中断、网络分

组丢失、板卡转发异常、设备 CPU 高等多种故障现象,细化备份种类和规则、顺序。

(2) 主备切换后,新上线业务的用户物理信息保持

由于设定为主备关系的 2 台 BRAS/SR 设备在硬件资源分配上不能做到完全一致,比如一台接入网设备,上联主用设备的 2/0/3 端口和备用设备的 4/0/8 端口,就需要采用虚拟端口的技术屏蔽实际端口的不一致。这样当主备切换后,为了保证用户话单的准确性,已经上线的用户在新的主用设备上实际在 4/0/8 端口接入,用户上网信息中仍然保持是在 2/0/3 这个虚拟端口接入;新上线的用户,虽然是在 4/0/8 端口接入,但 BRAS/SR 设备上报给认证计费系统时,要上报成 2/0/3 这个虚拟端口。因此新的主用设备上,存在维持实际端口和虚拟端口的对应关系的必要。

这个对应关系的实时性和准确性,关系到用户业务能否正常使用,所以也需要完善这类关系的准确性检查机制并及时对异常进行告警。

(3) 故障切换和回切时的操作状态呈现

热备技术涉及 BFD 对链路和邻居关系的检查机制、VRRP 组协议的切换机制、用户信息的备份和切换机制、路由和转发切换机制。这 4 者之间的操作顺序和交互关系决定了热备能否成功。作为处在现网运维条件下设备,需要清楚地呈现出这些操作的流程和状态供运维人员确认,需要有明确的日志和告警呈现供运维人员参考。目前的设备上,还无法清晰地展示这些流程和状态,笔者根据实际网络场景,设想了热备的业务切换流程,如图 4 所示。

(4) 不同组网环境下的热备

在城域网的组网环境下,既有接入网设备连接汇聚交换机再上联 BRAS/SR 的,也有接入网设备直连 BRAS/SR 的,还有接入网设备级联后再向上接入。笔者也对不同组网环境下热备能力进行了评估。不同业务组网条件下的热备拓扑如图 5 所示。

在有汇聚交换机的场景下,热备技术能保护汇聚交换机上行链路、BRAS/SR 设备或板卡、端口故障引起的业务中断,不能保障接入网设备上行链路和汇聚交换机设备、板卡、端口故障引起的业务中断。在 OLT 直连 BRAS/SR 的场景下,热备技术能保护 OLT 上行链路、BRAS/SR 设备或板卡、端口故障引起的业务中断,不能保障 OLT 设备、板卡、端口以及下行链路故障引起的业务中断。

此外,热备技术需要在 BRAS/SR 和汇聚交换机、OLT 设备之间启用 BFD 协议,目前老式的汇聚交换机和 OLT

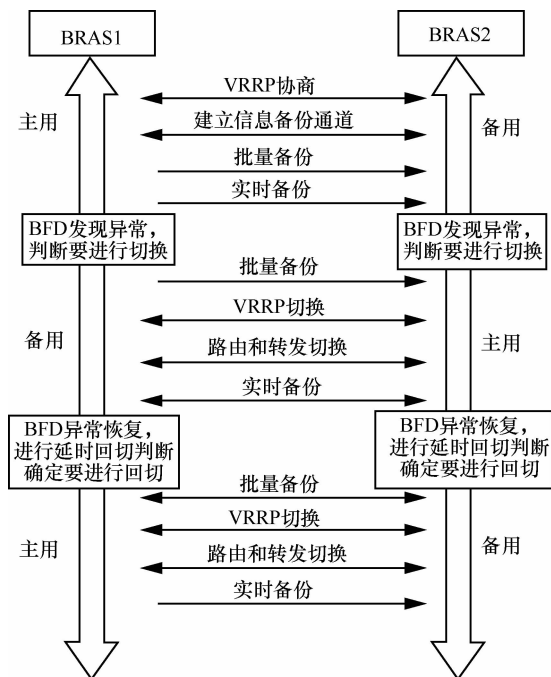


图4 热备切换流程

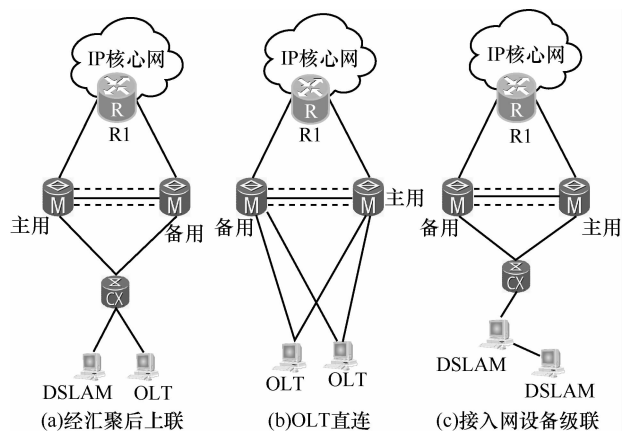


图5 不同业务组网条件下的热备拓扑

不支持 BFD 协议,无法实现完整的热备功能。还有部分老式的汇聚交换机,基于 VLAN 而非 MAC 进行业务转发,更难以实现用户上行流量的主动切换。

(5)实现热备功能的处理单元

目前 IP 网络设备的软肋主要有两个:一是单个 CPU 处理多项重要业务功能和进程,以至于很多设计好的冗余保护功能在 CPU 升高的情况下不能发挥应有的效用,二是内存回收不彻底造成内存泄露。热备功能也是要靠设备软件进行处理的,因此在很大程度上和其他功能一样,凸显出对设备 CPU 及内存的极大依赖。

热备作为 BRAS/SR 设备的重要安全功能,就是要在设备出现故障的时候发挥功效,顺利进行业务切换。因此

这项功能千万不能因为设备其他故障引发的 CPU 占用率过高或内存不够的问题而无法发挥作用。从设备软硬件设计的角度上,应当采用专用 CPU 和专用内存为这项功能服务,才能保证热备功能的正常运转。

(6)目前热备技术不支持的业务功能

目前热备技术仅支持 PPPoE 方式的上网、ITV 单播业务、基于 session 的 DHCP 方式的业务,并不支持固定 IP 地址专线业务、L2TP 方式的 VPDN 业务、IPv4 多播业务、IPv6 业务及 IPv4/IPv6 双栈业务。在全业务融合承载的今天,这些业务的欠缺也意味着这些是热备技术的完善方向。

综上所述,根据现网实际情况,BRAS/SR 可以分别部署热备份、冷备份;其中又可以分为 1:1 和 N:1 等方式。因各厂商 BRAS/SR 备份机制不同,在构建网络 BRAS/SR 高可靠性要求时,暂时还不能采用跨厂商方式部署热备份。建议采用同厂商同型号设备的 1:1 热备。

在多业务 IP/MPLS 承载网络中,网络的高可靠性对于电信级业务而言是非常重要的,尤其是 BRAS/SR 这一层面的设备,保证业务高可靠性更是举足轻重。但是,热备技术还存在厂商互通性差、规划部署复杂等问题,因此需要根据网络的实际情况,逐步试点、不断改进,以满足电信 IP/MPLS 网络多业务承载对于高可靠性的要求。相信通过不断的技术改进,热备技术必将成为构建新的电信级城域网不可缺少的手段。

[作者简介]



黄晋,男,中国电信股份有限公司苏州分公司高级工程师、高级技术经理,从 2003 年起从事苏州电信数据宽带网络的规划、建设和维护管理工作,组织实施了苏州电信城域网、IP 承载网等大型网络的建设、优化项目,完成了 MPLS-VPN、ITV、NGN 等重要业务的网络规划部署工作,在数据通信网络建设维护领域有较为丰富的实践经验。



薛姝,女,中国电信股份有限公司苏州分公司工程师、IP 维护优化班班长,从事 IP 网络的维护与建设工作,参与组织和实施了苏州电信城域网、IP 承载网等大型网络的建设、优化项目。

(收稿日期:2013-06-04)