

融合下一代电信级 IP 网的核心引擎—TB 级路由器

Converged Next – Generation IP Based Telecommunication Network—TSR

曾 斌 (浙江林学院现代教育技术中心 311300)

曾 凯 (浙江林学院信息工程学院 311300)

王宇熙 (浙江林学院信息工程学院 311300)

杨 英 (微创网格信息技术有限公司 310018)

摘要:本文分析了目前电信级核心网络中存在的急需解决的问题,介绍了 T 比特路由器作为融合下一代电信级 IP 网的核心引擎的各项技术优势以及它是如何解决目前电信级核心网络中存在的问题和如何帮助电信运营商建立下一代电信级 IP 网的。

关键词:电信级 IP 网 TB 级核心路由器 PAROLI 光接口 BGP SLA

在目前电信级核心网络中,存在着许多急需解决的问题。与光传输应用的广泛实现,1.6Tbit/s 或更高传输系统的商用化相比,现阶段成熟的 G 比特的路由器已经成为数据通信节点的瓶颈。传统的核心路由器,在日益需求增长和更加复杂的电信级网络中成为了拓展的阻碍。因此新的 T 比特路由器如何解决传统核心网络路由器设备的缺憾,成为 T 比特路由器解决方案最重要的考虑方向。作为解决方案的核心,T 比特路由器的定义及它如何能帮助电信运营商建立下一代电信级的 IP 网,进而协助加速有高附加价值的新型电信业务的开展,突破目前在电信业因传统业务价值的快速下降所造成经营困境。

T 比特路由器 (TSR, Terabit Switching Router) 解决方案正是一个可以在解决以上迫切问题基础上,帮助服务供应商提高业务盈利能力、技术灵活性和运营效率的解决方案。良好的电信级核心路由器——T 比特路由器提供 L2/L3/MPLS 交换引擎必须为分散式网络处理器;能解决日益复杂的路由策略;协助缩短错误恢复的时间;能解决光波段在核心节点短缺的问题;能消除因流量急速变化所需的大量又高速的常备中转线路接口;能在多条负载分担的线路中保持信息传递的顺序,同时保证各种业务相对应的 SLA。

1 T 比特路由器应具备的性能

建立电信级 IP 网的解决办法中,一般一个 T 比特路由器应具备以下特点:

1.1 结构方面

(1) 必须支持多机箱互联,并支持在所有的接口上的单一路由策略;

在多机箱多条负载分担的线路保持信息传递的顺序,并保证各种业务相对的 SLA;

交换矩阵必须在合理的制造成本下,最大化地提升交换矩阵的加速比。

(2) 必须支持中置背板,减少尾纤的维护,加速排查故障及降低维护费用,中置背板的支持可使板卡的尺寸维持在可以维护的范围,进一步地降低维护的费用。

(3) 使用可升级的芯片模块,提高系统的性能,降低能源需求;

可有效地区段化散热系统,热源的隔离则有助于增加整体系统的稳定性;

机箱互连的必须是智能化,并能自动更正人为的连接错误;

机箱的互连不得占用板卡插槽;

支持 40G 板卡,每片板卡必须有单一的 40Gb 以上的背板带宽。

(4) 能将系统逻辑地分开成数个路由系统,并

支持每个子系统的不同路由策略;

子路由系统必须能跨越机箱。

1.2 软件方面

(1) 操作系统软件必须完全模块化,并提供硬件协助的模块通信;

软件模块的地址空间必须被硬件保护。并可经由补丁的方式升级子模块;

软件模块的补丁必须能集群化,以简化打补丁的程序。

(2) 操作系统必须有管理补丁的版本和兼容性;

操作系统的升级决不能造成服务中断或系统重起;

操作系统可经由补丁达到与新版本相同的功能,不必强制升级至新版本。

1.3 其他

(1) 支持 L2/L3/MPLS 二阶转发 (Dual Stage Forwarding);

(2) 支持任意数目的硬件 BGP 处理器;

(3) 支持分散式的 BGP 路由计算,已加速路由收敛的时间;

(4) 提供功能完善的管理功能。

2 以合理成本解决同一节点内中继线路拥塞,简化因流量工程所造成日益复杂的路由策略

现在互联网用户对于信息的来源更加的灵活地选择,使得互联网的流量变得更无法预测,以至于无法有效地利用 BGP 的策略进行流量工程。因此对网管造成而外的工作负担,并使其无法开展其他有利于提高业务收入的工作。

无限制地扩展带宽也不是一个合理的解决办法,因为这样会造成固定投资费用无限制地增加。举例而言,途经 A 城市节点的流量来自于许多不同的广域网线路,而这些线路却不一定接入至 A 城市节点的同一路由器上,在最糟糕的流量模型下,A 城市节点会需要相当于百分之五十的广域网线路带宽作为节点内的中继线路,虽然两台设备之间数十条的尾纤 (Tail Fiber) 不会太昂贵,但线路两端的接口板卡总数将会等于广域网线路的接口板卡总数,因此大幅度地增加了固定

投资费用。而节点内的中继线路通常不需要许多在广域网中才需要的功能,如数以千计的流量整形会话 (Traffic Shaping Sessions) 及数以兆计的缓存 (Packet Buffer), 因此利用高智能接口板卡连接同一节点的两台路由器,不但占用的宝贵的设备插槽,更增加了固定投资费用。而 T 比特核心路由器解决方案将提供如下的方式解决这个问题:

2.1 使用新型的光接口,作为线板机箱之间或线板机箱与交换矩阵的连接

鉴于短距高速光连接的需求,新型的光接口是利用多个 VCSEL 光电转换器所组成的 PAROLI (PARAllel Optical Link) 光接口。一条 PAROLI 光缆由多条光纤组成,其物理带宽可超过 200Gb。一套 PAROLI 光连接系统由多条 PAROLI 光缆所组成,其有效带宽可超过 100 条 STM64 的线路带宽。

2.2 T 比特核心路由器解决方案避免了其他解决方案的弊端

(1) 不使用更高接口密度的路由器,解决了使运营商为追逐更新的科技制造出的更高密度的路由器而不断提高单位成本的弊端。

(2) 在部署 CoS 时不会遇到带宽以及市场成熟度的限制,解决了中继线路不断增加,固定投资费用无法控制的问题。

(3) 解决使用全连接的广域网组网方式导致不具备扩展性能,无须为各个阶段的传输波段资源的短缺感到担心。

3 解决因 IP 业务所造成的光波段在核心节点短缺之现象

如果将光网路看作是 IP 网的交换矩阵,则解决这个需求的方法就可用交换原理中的区段化的方式,利用降低穿透核心节点的波长个数,达到建立具有扩展性和经济性的电信级 IP 网。

传统全连接的方式,在考虑冗余备份的情况下,节点个数的平方和 DWDM 波长数成正比 ($W = N^2 - 1$), 在 DWDM 的波长数受物理参数的限制下 (S/N Ratio, PMD), 必须将平方的参数部分移除,才能有效地加大节点的个数。因此,必须改全连接为多个逻辑星状迂回连接 (如图 1.2 所示,任一节点和核心结点以两个波段连接,两个波段分别位于东环及西环上)。改为多

个逻辑星状迂回连接后,在任一个逻辑星状迂回连接中,仅节点个数和 WDM 波长数成正比(线性),因此,大幅度地提高了光缆环上的节点个数。而星状连接中的 IP 核心节点(T 比特路由器,非光传输交叉机),因为没有穿透的波长,所以有效地解决了这个问题。

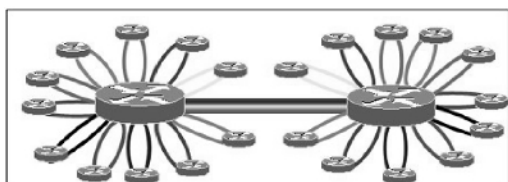


图 1 逻辑图

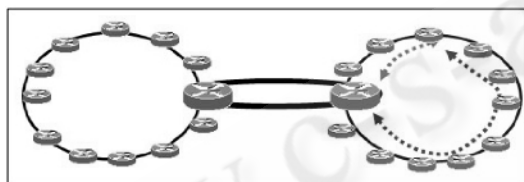


图 2 物理图

4 提高设备的可服务性

4.1 中置背板

中置背板的设计,使机房现场维护人员可以在不接触尾纤(连接于光接口子卡上,前卡)的情况下,轻易地完成接口板卡(后卡)的更换,使得因光纤插错而使服务中断的可能性降至最低。

4.2 可升级的芯片模块

使用高压及高密度的连接器,可以使接口板卡上每个重要功能独立被放置于数个子卡上,如此以来,每个子卡均可独立升级。

4.3 区段化散热系统

中置背板的设计,同时可使板卡面积缩小,并有效地隔离高热的光端口子卡于另一个散热系统中。板卡面积的缩小有助于机房现场维护人员的操作,而热源的隔离则有助于增加整体系统的稳定性。

5 建立有高度扩展性的转发性能及转发机制

5.1 并行处理

良好网路处理器不会因为常用功能的开启而对

其交换能力有明显的性能下降。其工作原理是将常用的功能硬件化,并与其他可程式化的工作同时进行。

5.2 分散式转发

良好的核心路由系统,必须支持分散式的 L2/L3/MPLS 的转发。并且有稳健的硬件支持的 IPC(进程之间的互通信,用于传递两个进程共享的资料,如路由信息),最大化地避免因 IPC 错误所造成的信息不同步。

5.3 二阶段的 L2/L3/MPLS 转发

当一个数据包进入系统后,位于输入端,与光接口相连的网络处理器会将此数据包,根据初步的信息(如 MPLS VPN 所用的 BGP LABEL),转发至相应的输出接口板卡。然后,位于输出端,与背板相连接的网络处理器会利用被指定的路由表进行查找及转发。这样的转发机制被称为二阶段转发。二阶段转发有下列几个优点:

(1) 专用的路由表,如 VPN 的路由表,不必经由 IPC 送至所有的输入板卡,因此降低了因 IPC 错误所造成的路由不同步。同时,因 VPN 的路由表只须放至在输出板卡的硬件转发表内,使系统对于 VPN 的个数及 VPN 路由的条目,可以随着输出板卡的个数无限增加。举例说,当系统中只有一片输出板卡时,系统可以支持一百万条 VPN 路由,当系统有十片输出板卡时,系统便可以支持一千万条 VPN 路由。

(2) 二层的封装信息,不必经由 IPC 送至所有的输入板卡。因此降低了因 IPC 错误所造成的丢包或是误转发。

(3) 不需要因为新的输出端功能,而更改输入端的微指令。这样的转发环境可以使软件的开发加速。

(4) 输入端的 xACL 由输入端的 CAM 记录,输出端的 xACL 则存于输出端的 CAM。这样的设计使 xACL 的运用非常的灵活。

6 在多机箱多条负载分担的线路保持信息传递的顺序,并保证各种业务相对的 SLA

在图 3 中,如要保证到 NETWORK - C 的九条线路都能够保持信息传递的顺序,并保证各种业务相对的 SLA,则这九条线路必须能完全负载分担。也就是说,任何经由这三台路由器到 NETWORK - C 的流量都必须

均匀分流到这九条线路上,避免任何单一的线路阻塞,再配合相应的排队机制和阻塞控制管理,这样才能保证各种业务相对的 SLA(延迟、抖动等)。

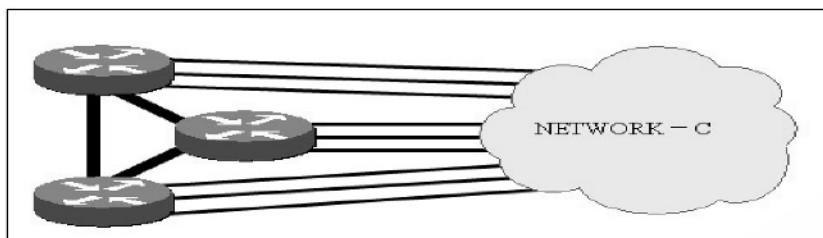


图 3

为了使流量能均匀的分流于不同路由器的线路,图 3 中的三台路由器必须组成一个逻辑子路由系统,此子路由系统将和 NETWORK - C 保持一个 BGP 连接,执行一个 BGP 策略,它的三台路由器将有同样的路由表,而这九条到 NETWORK - C 的线路,将可被视作负载分担的线路。如此一来,这九条线路,将根据输入板卡负载分担的算法进行分流,避免因三个 BGP 策略,三个路由表,而出现负载不均衡的现象。

在九条线路上保证信息传递的顺序则由输入板卡负载分担的算法决定。使用哈希算法同时对源地址及目的地址进行计算,是目前最常见的负载分担算法,但算法中必须允许加入 key 用以避免端化现象(Polarization)。

至于保证各种业务相对应 SLA,则经由上述两个机制实现。任何一个机制的失效,都会造成端化现象,使在不同线路上的相同业务出现不同的服务品质。因此,避免因路由表和负载分担算法而造成的端化现象,是保证各种业务相对应 SLA 的必备条件。

7 降低固定投资成本,使设备的可用年限能大幅提升,改善 IP 业务的投资回报比例

IP 的服务提供商也不断地更新设备以应付业务的需求。但是不断地更新设备却也造了 IP 网络的固定投资成本过高,使得经营 IP 服务地投资回报比例不是甚佳。

新的 T 比特路由器在设计时,必须考虑如何能使系统能根据业务状况平滑升级,进而改善 IP 业务的投资回报比例。以核心路由器而言,需经常更新的零部

件有机箱,电源,接口板卡,交换矩阵和冷却系统。T 比特路由器的优良设计将协助 IP 的服务提供商降低固定投资成本及延长这些零部件的可用年限。

8 以新的半导体制造工艺开发所需的设备,降低对能源的需求,进而降低营运成本

在这个新的世纪中,能源将成为国家是否能持续发展的一个重要因素。直到下一代的能源来源被商用化前,如何能用合理的成本取得能源也成为一企业是否能胜出的重要条件。

电信工业所需的能源主要用于电信设备。利用新的半导体制造工艺开发所需的设备,能降低对能源的需求,进而降低营运成本和提高获利。我们来看这样一个例子:STM64 的光电转换器,如利用传统的砷化镓技术制作其 MUX/DEMUX,其耗电量约为 15W。在改用 CMOS 技术后,耗电量降为 4 - 7W! 以一个有 1000 个 STM64 端口的 Tbit 的路由系统而言,仅仅光电转换器一项,就可节省 11,000W 的能源。新一代的 T 比特路由器将会有大量的高速端口及使用庞大的高速记忆体作为缓存,因此,如使用旧技术制造这些器件,必将使用许多能源,而以新的半导体制造工艺开发的新一代的 T 比特路由器将不会有这样的问题。

选择一个具有良好的扩展性的 T 比特路由器,将能协助电信运营商保护其投资,进而有效地控制固定投资费用。而因减少更换及升级网络设备,使运营商能简化运营维护及降低因服务中断的损失,专注于新型高附加价值电信业务的投产,则能大幅度地减少运营费用,提高营业收入。

参考文献

- 1 张宏科,路由器原理与技术——面向 21 世纪的 IP 网络技术经典著作,国防工业出版社,2003 - 1。
- 2 邱仲潘等,Cisco 路由器从入门到精通,电子工业出版社,2000 - 9。
- 3 蔡康等,下一代网络(NGN)业务及运营,人民邮电出版社,2004 - 8。
- 4 唐纯贞、严建民、鲁碧英,电信网与电信业务,人民邮电出版社,2003 - 8。