

电信级路由器的 MPLS 技术原理及性能测试

孟祥迪^{1,2}, 郭静寰³, 熊木地³

(1. 中国科学院 长春光学精密机械与物理研究所, 吉林 长春 130033;

2. 中国科学院 研究生院, 北京 100039;

3. 大连海事大学 计算机科学与技术学院, 辽宁 大连 116026)

摘要:介绍了基于路由器的多协议标记交换(MPLS)技术出现的相关背景,以及 MPLS 的网络结构、节点体系结构和标记交换的工作过程,分析了目前 MPLS 技术的优点和不足,搭建了 MPLS 网络的测试环境,并利用该组网环境采用国内某通讯公司的路由器进行了性能测试。测试结果较真实地反映了被测路由器的 MPLS 性能。

关键词: 路由器;多协议标记交换;性能测试

中图分类号: TN915.142

文献标识码: A

文章编号: 1673-629X(2006)10-0207-04

Study on Principle and Performance Testing of MPLS
Based on High Level RouterMENG Xiang-di^{1,2}, GUO Jing-huan³, XIONG Mu-di³

(1. Changchun Inst. of Optics, Fine Mechanics and Physics, Chinese Academy of Sciences, Changchun 130033, China;

2. Graduate School of Chinese Academy of Sciences, Beijing 100039, China;

3. Computer Science and Technology College, Dalian Maritime University, Dalian 116026, China)

Abstract: In this paper the technical background of MPLS based on router is presented. Then the MPLS network system, the architecture of node and the process of the label switching are introduced. And the characteristic and shortage of MPLS are analyzed. The system framework for MPLS is introduced. The testing system for the MPLS network is described. Using this system, the performance testing of the MPLS network is conducted with the some domestic routers. The result can really reflect the performance of the MPLS on the router.

Key words: router; multi-protocol label switch; performance test

0 引言

传统的 IP 路由有一些众所周知的局限性:从缩放性问题到支持数据流工程的能力弱以及与大型服务提供商网络中已有的第二层主干贫乏的集成能力。随着因特网的快速增长以及在大多数环境中选择 IP 作为第三层协议,传统 IP 路由的缺陷变得越发凸现出来。由于 Internet 采用面向无链接的 IP 协议,只能提供尽力而为(best-effort)服务,因此无法提供 QoS 保证。当现有 Internet 规模扩充到一定限度后,将在许多方面面临挑战:带宽、路由、网络扩展性、QoS 等^[1]。

ATM 技术的出现为解决 Internet 困境带来了契机,但由于在早期的方案中,所采用的是叠加式(Overlay)模型,使得无论是分组的封装效率、建链的时延、对组播的支持以及对 QoS 的支持都不理想。为解决这些问题,众多

厂商和学者采用集成模型提出了许多新方案,如信元交换路由器(CSR)技术、IP 交换技术、标记交换(Tag Switching)技术、基于路由的集成 IP 交换技术(ARIS)以及虚拟网络交换(VNS)技术,其共同特点是将 IP 路由和 ATM 交换的优点综合起来,这些方案都使用了标记的概念,但都不系统不完整。为了统一标准,IETF 提出了 MPLS 技术^[2]。

1 MPLS 技术

MPLS 技术是将面向非连接的 IP 业务移植到面向连接的标记交换业务之上,实现上将路由选择层面和数据交换层面分离。把网络层(第三层)的转发和数据链路层(第二层)的交换有机地结合起来,在网络层的数据包头与链路层数据包头之间插入固定长度的标签,网络根据标签选择路由并进行转发。这样就解决了路由器中不同长度的路由表在路由判决时所耗费的时间,同时在网络中将网络层的数据包交换转换到链路层实现,使网络能够充分发挥链路层效率和优势。

收稿日期:2005-12-14

作者简介:孟祥迪(1972-),男,山东烟台人,博士研究生,研究方向为高性能路由器设计及相关测试技术。

1.1 MPLS 的网络结构

MPLS 网络由核心部分的标签交换路由器 (Label Switching Router, LSR) 和边缘部分的标签边缘路由器 (Label Edge Router, LER) 组成。LSR 的作用可以看作是 ATM 交换机与传统路由器的结合, 由控制单元和交换单元组成。LER 的作用是分析 IP 包头, 用于决定相应的传送级别和标签交换路径 (LSP)。MPLS 网络组成如图 1 所示。

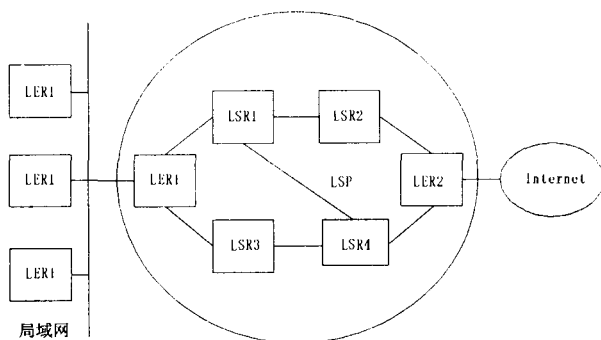


图 1 MPLS 的网络模型

(1) 标记交换边缘路由器 (Edge LSR)。

LER 是位于标记交换网络边缘的含完整的 3 层功能的路由设备, 它检查到来的分组, 在转发给标记交换网络前打上适当的标记, 当分组退出标记交换网络时删去该标记。作为具有完整功能的路由器, 标记边缘路由器也可应用增值的 3 层服务, 如安全、计费 and QoS 分类。LER 的能力不需要特别的硬件, 它作为交换机软件的一个附加特性来实现, 原有的路由器可以通过软件升级成具有 LER 的功能^[3]。

(2) 标记交换路由器 (LSR)。

LSR 是标记交换网络的核心。所谓标记是短的、固定长度的标签, 使 LSR 能用快速的硬件技术来做简单快速的表查询和分组转发。标记可以位于 ATM 信元的 VPI/VCI 域、帧中继的 DLCI、IPv6 的 flowlabel 域或在 2 层和 3 层头信息之间, 这使得标记交换可用于广泛的介质之上, 包括 ATM 连接、以太网等。该设备以标记交换为基础, 对有标记的分组或信元进行交换。LSR 可直接通过对 ATM 交换机的软件升级而实现^[3]。

(3) 标记发布协议 (LDP)。

LDP 与标准的网络层路由协议相结合, 在 MPLS 网络的各设备间发布标记信息。LER 和 LSR 使用标准的路由协议 (如: RIP, OSPF, BGP 等) 建立它们的路由表, 然后相邻的 LSR 和 LER 运用 LDP 协议互相协商标记值, 以达成对具有局部意义标记值的“共识”, 建立标记信息库^[3]。

1.2 MPLS 节点体系结构及标记交换工作过程

MPLS 技术兼有基于第二层交换的分组转发技术和第三层路由技术的优点。与第二层网络 (如 ATM 或帧中继) 相似, MPLS 给分组分配标签, 便于分组能够在基于分组或信元的网络中传输。MPLS 体系结构被分成两个独立的组件: 转发层面 (数据层面) 和控制组件 (控制层面),

如图 2 所示。

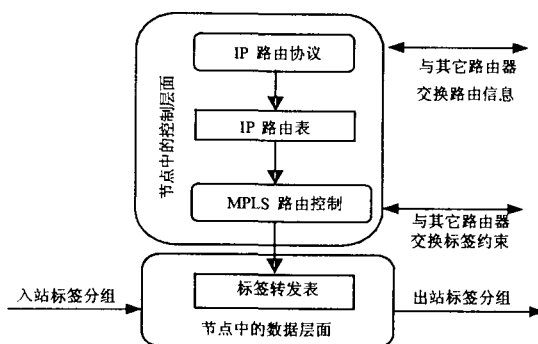


图 2 执行 IP 路由的 MPLS 节点体系结构

转发层面: MPLS 转发层面负责根据所附卷标值来转发资料包。转发层面使用 MPLS 节点维护的卷标转发信息库 (Label Forwarding Information Base, LFIB) 实现对带有卷标的数据包的转发。标签转发的算法不仅利用了包含在 LFIB 中的信息, 还利用了包含在卷标值里的信息。每个 MPLS 节点维护着与 MPLS 转发相关的两张表: 标签信息库 (Label Information Base, LIB) 和 LFIB。LIB 含有本地 MPLS 节点分配的所有的卷标, 以及这些卷标与从邻居 MPLS 节点收到标签之间的映像关系; LFIB 使用 LIB 中所含标签的一个子集来进行实际的数据转发。

控制层面: MPLS 控制层面负责维护 LFIB 表和向 LFIB 表添加条目。所有的 MPLS 节点必须运行 IP 路由选择协议来和网络中的所有其它 MPLS 节点交换 IP 路由选择信息。一般的链路状态型路由选择协议, 如 OSPF 和 IS-IS, 是推荐采用的路由协议^[4]。

基于 MPLS 的路由器上主要有: 帧模式 MPLS 和信元 MPLS 两种。IP 分组在通过 MPLS 主干网时标记交换工作过程主要分为以下 3 个步骤^[4]:

(1) 入口边界 LSR 接收 IP 分组, 将分组归为一个转发等同类 (FEC), 并使用该 FEC 对应的出站卷标栈标记该分组。对于基于目标的单播 IP 路由, FEC 对应于一个目标子网, 分组的分类是在转发表中执行传统的第三层查找。

(2) 主干 LSR 接收被标记的分组, 并使用标签转发表以相同 FEC 对应的出站标签代替输入分组中的入站标签。

(3) 当该特定 FEC 的出口边界 LSR 接收到标记分组后, 删除该标签并对得到的 IP 分组执行传统的第三层查找。

2 MPLS 技术的优点及不足

MPLS 是一种新出现的技术, 旨在解决与当前联网环境中使用的分组转发技术相关的许多问题, 主要是为继承标记交换中继方案和网络层路由技术的基础技术, 改善网络层路由技术的性价比, 提高网络的可缩放性, 并在路由服务方面提供更大的灵活性。但随着对 MPLS 技术研究的不断深入, 也发现了 MPLS 技术存在一些问题。下面

把该技术的优缺点加以介绍。

2.1 MPLS 技术的优点

MPLS 技术是一种在开放的通信网上利用卷标引导数据高速、高效传输的新技术。它的价值在于能够在一个无连接的网络中引入连接模式的特性。其主要优点是减少了网络复杂性,兼容现有各种主流网络技术,能降低 50% 网络成本,在提供 IP 业务时能确保 QoS 和安全性,具有流量工程能力。此外,MPLS 能解决 VPN 扩展问题和维护成本问题。MPLS 技术是下一代最具竞争力的通信网络技术,其技术的优点主要有^[5]:

(1)MPLS 在网络中的分组转发是基于定长标签,由此简化了转发机制,使得转发路由器容量很容易扩展到太比特级;

(2)MPLS 技术充分采用原有的 IP 路由,并在此基础上加以改进,保证了 MPLS 网络路由具有灵活性的特点;

(3)MPLS 抛弃了复杂的 ATM 信令,无缝地将 IP 技术的优点融合到 ATM 的高效硬件转发中;

(4)MPLS 是一种面向连接的传输技术,能够提供有效的 QoS 保证;

(5)MPLS 是一种与链路层无关的技术,它同时支持 X.25、帧中继、ATM、DWDM……,保证了多种网络的互连互通;

(6)MPLS 支持大规模层次化的网络拓扑结构,具有良好的网络扩展性,MPLS 的标签合并机制支持不同数据流的合并传输;

(7)MPLS 支持流量工程、CoS(业务类型)、QoS(服务质量)和大规模的虚拟专用网。

2.2 MPLS 技术不足之处

IP 是无连接的网络,因为 IP 路由使用变长的最长地址前缀匹配做搜索(即搜索匹配前缀最长的一个作为入口),因此无法实现高速转发。而使用 MPLS 标记后,根据固定长度的标签搜索目的地址,能够实现高速转发。但深入研究 MPLS 技术就会发现,MPLS 自身对基本的 IP 网络体系几乎没有增加任何新的内容。除了使用标记做转发外,控制部分还是依赖原来的路由协议。因此简单地认为由于 MPLS 标记比 IP 地址前缀短,使用更短的标记做交换会比用 IP 地址做交换更快,是值得怀疑的。在 MPLS 的最基本层面上,所做的全部工作就是在路由和转发之间增加了新的一层,从而引入更多的间接性。引入新的一层,会在准确关联标记和逐跳转发路径时增加新问题。

在一个 MPLS 网络中,必须依赖于路由协议来既准确地传播可达性信息,又做一些与标记分发相关的工作,因此 MPLS 网络对路由协议的依赖性要高于 IP 网络。但到目前为止,路由系统还不成熟,远未达到令人满意的程

度。如果 MPLS 使用标记交换机制在路由和转发之间引入新一层的间接性,可能会导致 MPLS 网络的故障更具迷惑性,更加难以分析和排除。同时,MPLS-VPN 的可扩展性、流量工程与 TCP 关系的处理以及 MPLS VPN 的安全性和 QoS 等方面都存在相应的不足^[5]。

3 测试环境和结果

MPLS 技术作为新兴的路由交换技术,越来越受到业内的关注。在研究 MPLS 技术的同时,对其相关测试技术的研究也逐渐成为目前网络测试研究的热点。LSR 是 MPLS 网络的基本单元,对 MPLS 的测试也主要集中在对 LSR 的测试上。对 MPLS 的测试主要包括吞吐量测试,即 LSR 对不同大小帧标记分组转发速率的测试,以及对大小帧分组转发时延的测试。LSR 的转发性能测试和传统的路由转发测试基本相通,这里采用了 RFC2455 的建议选择测试方法和度量,对 MPLS 网络进行性能测试^[6,7]。

3.1 MPLS 测试环境的设置

针对上述 MPLS 的特点,采用了如图 3 所示的组网方式,中间的 4 台路由器 RA、RB、RC、RD 组成骨干网,使用 ISIS 协议,机架之间形成全网状 MP-BGP 连接。CE(客户边界)路由器 CE1、CE2 接测试仪,再测试构造相应的数据包,相互收发数据。RC 和 RD 之间存在备份链路 POS12,以构成冗余链路。考虑到 MPLS 协议对各种路由的支持的测试,作为 PE(提供商边界)的路由器 RC 的 vrf 接入采用 BGP 协议,而在 RD 中,为了防止私网路由和公网路由的冲突,vrf 接入采用 Multi-VR 方式的 ISIS 接入。同时考虑到实际组网环境可能会有大量其它的流量与 MPLS 竞争网络资源,在测试环境中增加了 VPWS 和 MPLS-TE 等其它多个业务系统。

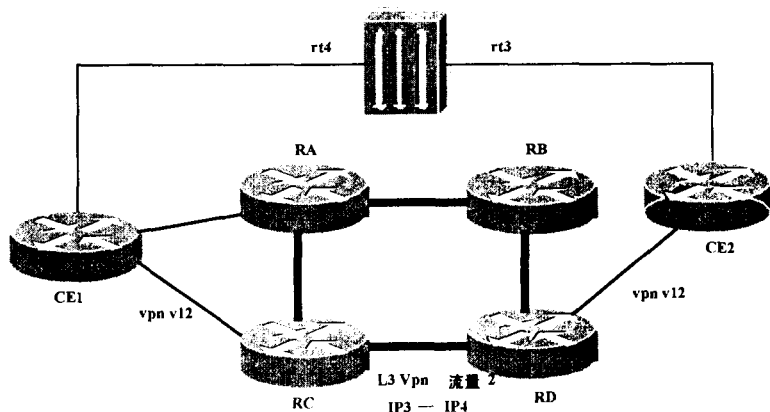


图 3 MPLS 测试组网图

整个测试系统的一个测试接口可以同时生成多个“流”,使用不同的流量生成方案,构成不同的测试连接模型,增加测试系统的实际应用性。如图 3 的组网环境主要是测试目的:

(1)私网路由通告正常:V3 能够导入 V1 和 V2 的私网路由,V1、V2 只能分别导入 V3 的私网路由;

表 1 测试结果

IP Pack Length	Av Latency Limit(μ s)	Max Allowed Lost Pkts	Max Theoretic Rate(pkts/s)	Actual rate (pkts/s)	Throughput (Rx/Tx %)	Tx Load (% of max)	Latency (μ s)		
							Min/	Avg/	Max
48	100000.00	0	145349	88832	100.0000	85.1161	19.42/	108.08/	300.70
64	100000.00	0	122549	88651	100.0000	96.3394	21.60/	133.57/	341.35
128	100000.00	0	75301	75301	100.0000	100.0000	28.71/	156.23/	292.00
256	100000.00	0	42517	42517	100.0000	100.0000	41.25/	233.93/	446.33
512	100000.00	0	22727	22727	100.0000	100.0000	66.86/	402.15/	759.75
768	100000.00	0	15509	15509	100.0000	100.0000	92.45/	570.11/	1073.55
1024	100000.00	0	11770	11770	100.0000	100.0000	117.79/	739.19/	1386.59
1280	100000.00	0	9484	9484	100.0000	100.0000	179.50/	941.46/	1741.92
1500	100000.00	0	8127	8128	100.0000	100.0000	184.11/	1054.66/	1989.89

(2)私网流量通过公网能够正常、稳定、长时间转发。

保证验证路由器的 MPLS 功能系统在复杂组网环境中,并行多个业务系统时,各个业务的运行功能是否正常,48~72 小时运行是否基本可靠,以及冗余链路启用是否正确,同时在测试仪上配置千兆流量对 MPLS 网络的性能进行测试。

3.2 MPLS 测试结果

在测试过程中,使用 RouterTest 配置 MPLS 报文,检验 MPLS 网络的性能,同时为了使环境更加贴近现实环境,测试仪上构建多种业务流量,并在测试过程中检测系统的整体性能。本组网环境对 MPLS/VPN 模块的测试包含了 LDP,MPBGp,BGP,ISIS,RIP 等协议的相关测试,测试了 BGP,RIP,ISIS 的 VPN 接入,MPLP 网络在多业务共存的情况下,可以正常运行,网络性能比较可靠。网络的性能结果如表 1 所示。

从表中可以看出,当报文的长度为小字节时,如 48 或 64,性能会有所差异。在图 3 的组网测试环境中,如果增加一条 MPLS VPN 流量时,超过应有的带宽时,会有随机的丢包现象,不是均匀丢包,在测试仪上看到的现象是两个接收端交替接收到发送端的流量。

(上接第 206 页)

(2)W4006AF 用 CPU 送来的 $\overline{BE}_3 \sim \overline{BE}_0$ 选择 CAS 信号和数据传送类型。用 \overline{BE}_3 选择 $CAS_3, CAS_7, CAS_{11}, CAS_{15}$ 和 $D_{31} \sim D_{24}$; 用 \overline{BE}_2 选择 $CAS_2, CAS_6, CAS_{10}, CAS_{14}$ 和 $D_{23} \sim D_{16}$; 用 \overline{BE}_1 选择 $CAS_1, CAS_5, CAS_9, CAS_{13}$ 和 $D_{15} \sim D_8$; 用 \overline{BE}_0 选择 $CAS_0, CAS_4, CAS_8, CAS_{12}$ 和 $D_7 \sim D_0$, 从而可以实现数据的字节、字或双字类型的传送。

(3)因 W4006AF 控制两个存储体分别采用了相互独立的地址信号、选通信号和读写信号,加上其内部能够满足要求的时序逻辑设计,可以实现对双存储体的交叉访问,构成了并行存储器结构,提高了主存的访问速度。

3 结束语

通过以上分析,对微机主存结构及其工作原理有了较深的掌握。由此也可以想到若采用能够控制多个存储体的 DRAM 控制器,即可构成多体交叉访问的主存结构,使

4 结 论

在分析了 MPLS 网络技术特点基础上,搭建了接近实际环境的测试网络,对 MPLS 性能进行了测试。结果表明,MPLP 网络在多业务共存情况下,运行正常,性能可靠。

参考文献:

- [1] Rosen E, Viswanathan A, Callon R. Multi-protocol label Switching Architecture[Z]. REC 3031, 2001.
- [2] Comer D E. 用 TCP/IP 进行网际互联. 第 1 卷:原理、协议与结构(第 4 版)[M]. 北京:电子工业出版社, 2003.
- [3] Pepelnhak I, Guichard J. MPLS 和 VPN 体系结构[M]. 北京:人民邮电出版社, 2002.
- [4] Aliveyn V. 高级 MPLS 设计与实施[M]. 北京:人民邮电出版社, 2003.
- [5] 曹雄恒. 宽带多协议标记交换(MPLS)网络的原理及关键技术[J]. 计算机应用, 2001, 21(7): 29-31.
- [6] 章 森, 吴建平. 网络互联设备的性能测试: 原理和实践[J]. 小型微型计算机系统, 2004, 25(9): 1587-1591.
- [7] 田 辉. MPLS 测试技术[J]. 电信网技术, 2003, 19(2): 55-57.

主存的并行访问程度更高。若将每个 DRAM 控制器及其控制的存储器作为一个整体,制成电路板。即构成了内存条,由内存条方便地构成主存。

参考文献:

- [1] 窦振中. 单片机外围器件实用手册(存储器分册)[M]. 北京:北京航空航天大学出版社, 1998. 3-13.
- [2] 王 红. 高速缓冲存储器性能分析[J]. 微机发展, 2000, 10(5): 30-32.
- [3] 郇 军. 存储器管理中的虚拟存储技术[J]. 微机发展, 1997, 7(4): 52-54.
- [4] 白中英. 计算机组成原理(第 3 版)[M]. 北京:科学出版社, 2000. 71-104.
- [5] 俸远祯, 阎慧娟, 罗克露. 计算机组成原理[M]. 北京:电子工业出版社, 1996. 150-164.
- [6] 曾家智, 向世清. 微型计算机系统与接口[M]. 成都:电子科技大学出版社, 1992. 7-40.