

基于 NS2 的 MPLS 流量工程仿真研究

李蓬¹, 黄河²

(1. 北京建筑工程学院 电气与信息工程学院, 北京 100044;

2. 北京航空航天大学 软件学院, 北京 100083)

摘要: MPLS 被认为是下一代 Internet 骨干网络的核心技术, 利用 MPLS 可以有效实施流量工程。文中对 MPLS 技术进行了介绍, 针对流量工程的要求, 给出了满足 MPLS 流量工程要求的系统模型。通过对仿真工具 NS2 的扩展, 建立了一个 MPLS 流量工程仿真器, 可以实现对 MPLS 转发技术和约束路由算法进行性能评估。仿真器的核心思想是, 在标记交换路由器 LSR 模型中实现了标记分发、信令协议和路由协议扩展等基本功能; 在标记边缘路由器 LER 模型上实现了对显示路由算法的支持。

关键词: 多协议标记交换; NS2; 流量工程; 边缘交换路由器

中图分类号: TP393

文献标识码: A

文章编号: 1673-629X(2008)09-0053-04

Simulation Research of MPLS Traffic Engineering Based on NS2

LI Peng¹, HUANG He²

(1. School of Electricity and Info. Eng., Beijing Institute of Civil Eng. and Architecture, Beijing 100044, China;

2. College of Software, Beijing University of Aeronautics and Astronautics, Beijing 100083, China)

Abstract: MPLS is regarded as the backbone network technology of next generation Internet network. It can support effectively traffic engineering. In the first, gives a review of MPLS technology. In order to support traffic engineering, the system model supporting MPLS traffic engineering is put forward. And then, a network simulator supporting MPLS traffic engineering is presented by extending the NS2. It can contribute to performance assess of the mechanism of MPLS packet switch and the computation of constraint route. Distribution of the label, the MPLS signaling protocol and the expansion of route protocol are accomplished in the LSR model, and the computation of explicit route is realized in the LER model.

Key words: multiple protocol label switch; network simulator II; traffic engineering; label edge router

0 引言

多协议标记交换 MPLS (Multiple Protocol Switch) 是一种数据包的高速转发技术^[1]。在 MPLS 网络中, 边缘路由器 LER (Label Edge Router)^[2] 按照一定策略将数据包分类形成等效前传类 FEC (Forwarding Equal Class), 然后标记交换路由器 LSR (Label Switch Router) 利用信令协议为 FEC 分配一个定长标记, 标记产生后即形成一条从源端到目的端的标记交换路径 LSP (Label Switch Path)。中间 LSR 仅根据该标记来转发数据包, 不必进行 IP 地址最长匹配查找^[1]。MPLS 将路由和转发机制相分离, 提高了分组的转发

速度^[3], 同时, 通过建立合适的 LSP, MPLS 可以实现对网络流量和资源的合理控制, 有效实施流量工程 TE (Traffic Engineering)^[4]。NS2 (Network Simulator II) 是一种通用的网络测试和仿真平台, 利用该工具可以对网络模型和服务做出性能评价, 但目前它对 MPLS 技术的支持还不完善。文中首先介绍了满足流量工程要求的 LER 体系结构的设计, 然后通过对 NS2 进行扩展, 建立一个 MPLS 流量工程仿真器, 提供对 MPLS 转发技术和约束路由算法进行性能评估的能力。

1 LER 体系结构的设计

LER 是 MPLS 域内的边缘路由器, 它负责为 LSP 请求计算显式路由 ER, 然后通过信令协议沿着该 ER 建立标记交换路径 LSP。由此可见, 在 MPLS 流量工程模型中, 最重要的部分就是 LER 模型。为支持分布实施 MPLS 流量工程, LER 流量工程组件共分四层^[5], 分别是: 路径选择组件、扩展 IGP (Interior Gate-

收稿日期: 2007-12-25

基金项目: 北京市教委科技发展计划面上项目 (KM200610016012)

作者简介: 李蓬 (1975-), 男, 山东东营人, 硕士研究生, 讲师, 主要研究方向为计算机网络及网络安全; 黄河, 博士, 讲师, 主要研究方向为计算机网络性能分析和优化。

way Protocol)消息发布组件、信令组件和包转发组件。如图 1 所示。这几个功能模块与传统的 LSP 路由选择机制共同实现流量工程。

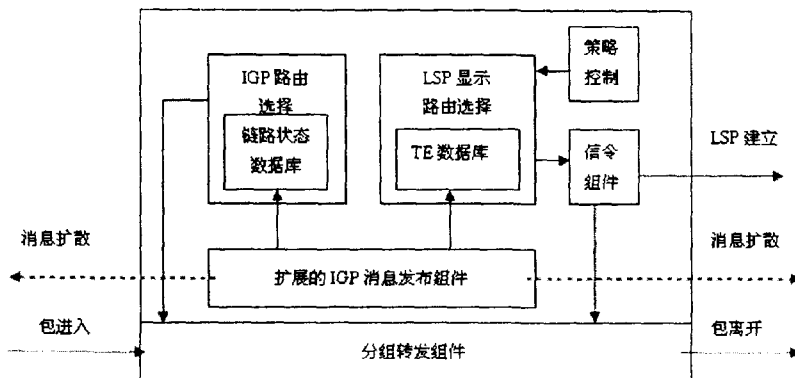


图 1 MPLS 流量工程系统模型

消息发布组件通过对 IGP 消息进行扩展,使原来的链路状态公告 LSA(Link State Advertisement)中包含新增的链路信息,如最大可用带宽、当前保留带宽等。LER 将从消息组件接收到的扩展 LSA 信息和网络拓扑信息存储在其流量工程数据库 TED 中,用于计算 ER。路径选择组件包括策略控制和 LSP 显式路由选择两个模块。其功能是为 LSP 请求计算符合约束条件的显示路由 ER。LER 收到一条 LSP 请求后,利用 TED 中的扩展 LSA 信息,加上必要的附加信息(如管理策略、链路权值等),使用某种 ER 算法,为该请求计算一条显示路由^[6]。

信令组件的作用是使用某种消息机制在 ER 经过的标记交换路由器 LSR 上安装 LSP 状态和进行标记分发。目前,可用于标记分发的 MPLS 信令协议有 RSVP 和 CD-LDP 两种。文中采用的是 CD-LDP。

包转发组件包括三个模块:标记管理、标记映射和包转发模块。采用 MPLS 进行数据包的转发,当信令协议分发标记成功后,一条 LSP 宣告建立。这时,中间 LSR 仅通过查询其 MPLS 转发表即可将数据包发送到相应的接口中,而无需进行 IP 包的最长匹配查找。

LER 的基本工作流程如下:

- 发布扩展 LSA 消息,建立链路状态数据库,生成路由。
- 查询链路权值及全局策略信息,生成符合带宽需求的加权子图。
- 加权子图中计算最短路径作为显式路由。
- 使用信令协议沿着显式路由建立 ER-LSP。
- 更新并发布扩展 LSA 消息。
- 按照 MPLS 转发表转发标记数据包。

2 MPLS 流量工程系统模型仿真

为了研究 MPLS 流量工程在 NS2 中的实现技术,文中以 NS2 为工具,设计和实现包括标记边缘路由器

LER、标记交换路由器 LSR 和一个流量工程策略在内的 MPLS 流量工程模型。LER 具备标记交换,扩展 LSA 消息发布, MPLS 包交换等功能。

2.1 LSR 的设计和实现

在 NS2 中,一个节点包含一个地址分类器 AddressClassifier 和 Agent。AddressClassifier 的作用是将进来的数据包分发到正确的链路或 Agent 上。Agent 用来仿真各层的协议实现。为了使 NS2 支持 MPLS,对类 AddressClassifier 和类 Agent 进行了扩展。如图 2 所示,一个基本 MPLS 节点 LSR 由 MPLSClassifier 和 LDP Agent 两个类构成,完成标记包转发和 LDP 消息发布等基本功能。其中, MPLSClassifier 根据转发表 LIB 进行数据包的转发, LDP Agent(CR-LDP)负责发布 LDP 消息。

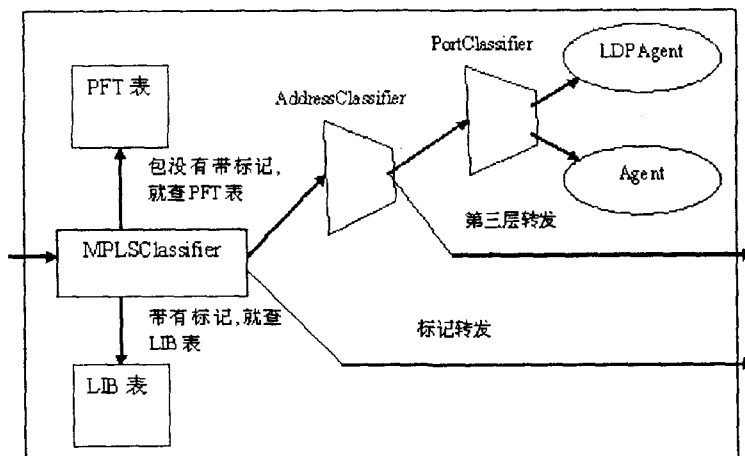


图 2 LSR 的节点结构

为完成数据包的转发,每个 MPLS 节点维护以下三个表:

- 部分转发表 PFT(Partial Forwarding Table):是转发表的一个子集,用于将入口 LSR 的 IP 包映射到 LSP 上。
- 显示路由表 ERB(Explicit Route information Base):包含已建立的显式 LSP(Label Switch Path)的信息,包括 LSP 标识、FEC 以及一个指向 LIB 中入/出标记的指针。
- 标记信息表 LIB(Label Information Base):包含已建立 LSP 的信息,用于标记数据包的转发。它包含输入输出标记和输入输出接口。

2.2 支持显示路由优化的 LER 节点结构

显示路由是利用 MPLS 实施流量工程和进行网

络性能优化的基础,如图3所示。为进行显示路由算法的性能模拟和评估,在基本 MPLS 节点的基础上,对 MPLS 标记边缘路由器 LER 的功能进行了扩展,使它支持扩展 LSA 发布、流分类和显示路由计算。LER 节点的流量工程组件包括显示路由计算、扩展 IGP 消息发布以及 CR-LDP 信令协议、准入控制机制等。其中扩展 IGP 消息发布机制负责通告链路当前可用带宽等流量工程参数。显式路由计算根据扩展 LSA 消息中的链路当前可用带宽和网络拓扑信息,按照某种显式路由算法为 LSP 请求计算显式路由。使用 CR-LDP 作为信令协议,它按照显式路由指定的路径进行标记分发,而准入控制机制则根据信令消息携带的带宽请求值和本网络链路可用带宽的值决定是否接纳该 LSP 请求。CR-LDP 信令协议还将控制资源管理和 MPLS 包的调度,以支持资源分类、区分服务和流量控制。

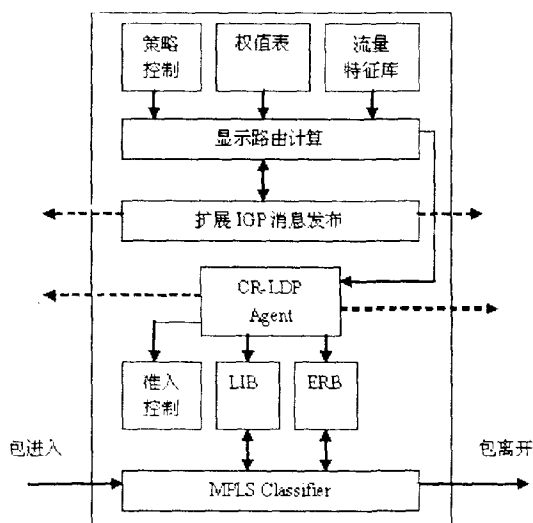


图3 支持显示路由优化的 LER 节点结构

为支持流量工程,MPLS 仿真模型中对路由协议和 CR-LDP 信令协议进行了扩展。其中,对路由协议的扩展分为两方面:

一是对原有最短路径路由算法进行了扩展,仿真模型实现了约束最短路径 CSPF,改进的最小干涉算法 MIRA 以及基于策略和流分类的 PTCR 等多种显式路由算法。二是对链路状态公告 LSA 进行扩展,引入如下流量工程扩展消息,加入到 IGP 的 LSA 中:a)最大链路带宽;b)最大可用带宽;c)当前保留带宽;d)当前使用带宽;e)链路颜色。

对于 LER 的实现,利用前面提到的 LDPAgent 和 MPLSAddressClassifier 这两个类,然后对

类 RoutingModule 进行扩展,创建类 MPLSModule,共同完成 LER 的标记交换功能。在此模型中,每个节点有一个 MPLSAddressClassifier 对象和一个 MPLSModule 对象以及若干 LDPAgent 对象。每一个 LDPAgent 都同另一个 MPLS 节点的一个 LDPAgent 保持会话连接,以便完成消息交换以及标记分发功能等。MPLSModule 除了继承了路由功能外,其他的功能包括定义标记分发机制,维护标记信息,触发标记分发等。

2.3 ER-LSP 的建立

显示路由算法确定的显示标记交换路径 ER-LSP 需要使用 CR-LSP 信令建立。为此在仿真模型中对 LDPAgent 类进行扩展,设计和实现了 CR-LDP 信令协议,完成标记分发过程。协议定义 LDP 消息包头,实现各种 LDP 消息,支持显式路由(Explicit Routing)标记分发和流量工程参数描述。CR-LDP 包头定义支持流量工程的相关参数描述如下:

```
int rc; //描述资源类型
int tr_msgtype; //流量参数类型
char tr[100]; //流参数描述,如频率,权值,
PDR, PBS, CDR, CBS 等
```

其中 tr 描述流的流量工程参数,包括 PDR(Peak Data Rate)、PBS(Peak Burst Size)、CDR(Committed Data Rate)、CBS(Committed Burst Size)、EBR(Excess Burst Size)。rc 用来描述资源类别,可对资源进行约束。CR-LDP 通过 tr 和 rc 可以支持约束路由和显式 LSP 分发,从信令上支持 MPLS 流量工程。

LDPAgent 实现了 CR-LDP 协议定义的五种基本消息:

- 标记通告消息:LDP_NotificationMSG;
- 标记映射消息:LDP_MappingMSG;
- 标记请求消息:LDP_RequestMSG;
- 标记撤销消息:LDP-WithdrawMSG;
- 标记释放消息:LDP_ReleaseMSG。

仿真模型中,利用 CR-LDP 建立 ER-LSP(Explicit Routing-LSP)的合作图如图4所示。

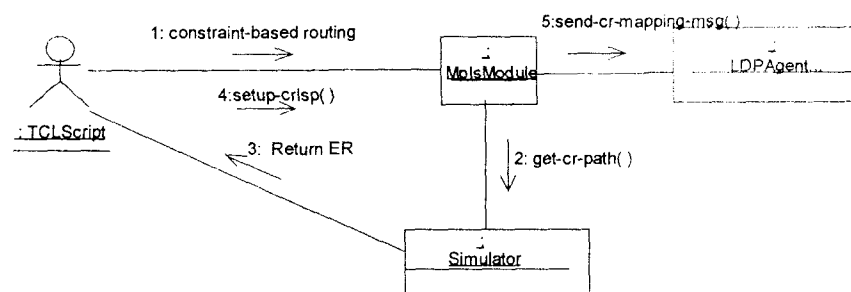


图4 建立 ER-LSP 的合作图

由 TCL 脚本向 MplsModule 类发起 CR(Constraint-based Routing)请求, Simulator 类通过调用显示路由计算模块将一个显示路由 ER(Explicit Route)返回, 然后 TCL 脚本再次发起 CRLSP 建立请求, 最后由 LD-PAgent 根据 CRLSP 请求中指定的显示路由, 使用标记映射消息建立一个显示标记交换路径 ER-LSP。

2.4 包转发机制的设计和实现

MPLS 节点的数据包转发由 MPLS 分类器完成。数据包到达 MPLS 节点后, MPLS 分类器负责执行的操作过程如图 5 所示。

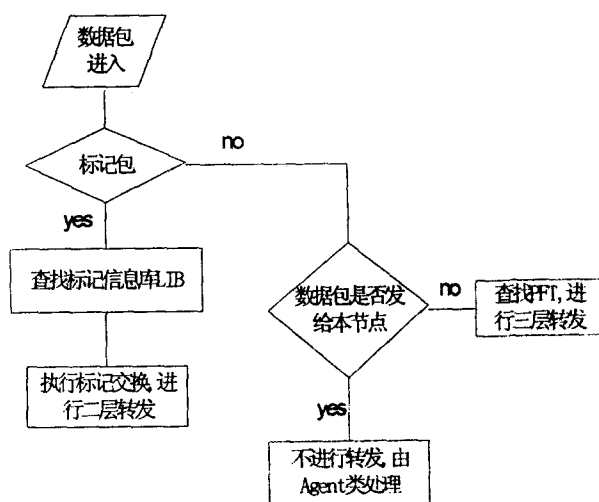


图 5 MPLS 分类器的操作过程

3 结束语

利用 MPLS 实施流量工程可以在现有网络体系结构基本不变的前提下, 合理利用网络资源, 优化网络性能。利用显示路由和故障恢复机制, MPLS 提供了基本的流量工程能力, 这使得大型网络中利用 MPLS 实施性能优化变得十分便利。文中使用 NS2 模拟环境对 MPLS 流量工程进行性能评估和测试。通过对 NS2 的扩展, 实现了一个 MPLS 流量工程仿真器, 可以支持标记分发, 标记交换路径 LSP, 显示路由以及区分服务等。

参考文献:

- [1] Rosen E, Viswanathan A, Callon R. Multiprotocol Label Switching Architecture[S]. IETF RFC 3031, 2001.
- [2] 石晶林, 丁 炜. MPLS 宽带网络互联技术[M]. 北京: 人民邮电出版社, 2001.
- [3] 谢金星, 刑文训. 网络优化[M]. 北京: 清华大学出版社, 2000.
- [4] Awduche D. Requirements for Traffic Engineering Over MPLS[S]. IETF RFC 2702, 1999.
- [5] 黄河, 李伟琴. MPLS 流量工程体系结构优化研究[J]. 北京航空航天大学学报, 2003, 29(3): 221-224.
- [6] Kodialam M, Lakshman T V. Minimum Interference Routing with Applications to MPLS Traffic Engineering[C]//IEEE INFOCOM. 2000. [s. l.]: IEEE Communications Society Press, 2000: 884-893.

(上接第 52 页)

- 2004, 33(3): 61-70.
- [2] 周立柱, 林 玲. 聚焦爬虫技术研究综述[J]. 计算机应用, 2005, 25(9): 1965-1969.
- [3] Cai D, Yu S, Wen J, et al. Extracting Content Structure for Web Pages Based on Visual Representation[C]//In APWeb, 2003. Xi'an: [s. n.], 2003: 406-417.
- [4] Cai D, Yu S, Wen J, et al. VIPS: a Vision-based Page Segmentation Algorithm[R]. Microsoft Research Technical Report, MSR-TR-2003-79, 2003.
- [5] Rahm E, Bernstein P A. A survey of approaches to automatic schema matching[J]. VLDB Journal, 2001, 10(4): 334-350.
- [6] WANG J, WEN J-R, Lochovsky F, et al. Instance-based schema matching for web databases by domain-specific query probing[C]//In VLDB 2004 Conference. Toronto, Canada: [s. n.], 2004.
- [7] WU W, YU C T, Doan A, et al. An interactive clustering-based approach to integrating source query interfaces on the deep web[C]//In SIGMOD 2004 Conference. Paris, France: [s. n.], 2004.
- [8] He B, Chang K C-C, Han J. Automatic complex schema matching across web query interfaces: A correlation mining approach[R]. Technical Report UIUCDCS-R-2003-2388, Dept. of Computer Science, UIUC, 2003.
- [9] Wu W, Doan A, Yu C T. WebIQ: Learning from the Web to Match Deep-Web Query Interfaces[C]//In ICDE 2006. Atlanta, GA, USA: [s. n.], 2006.
- [10] Yu C T, Philip G, Meng W. Distributed top-N query processing with possibly uncooperative local systems[C]//In: Proceedings of the 29th International Conference on Very Large Data Bases. Berlin: [s. n.], 2003: 117-128.
- [11] Zhang Z, He B, Chang K C-C. Light-weight domain-based form assistant: Querying Web Databases On the Fly[C]//In VLDB Conference. Trondheim, Norway: [s. n.], 2005: 97-108.
- [12] He H, Meng W, Yu C, et al. Wise-integrator: An automatic integrator of web search interfaces for e-commerce[C]//In VLDB 2003 Conference. Berlin: [s. n.], 2003.
- [13] Chang K C-C, He B, Zhang Z. Toward large scale integration: Building a metaquerier over databases on the web[C]//In CIDR 2005 Conference. Asilomar, CA, USA: [s. n.], 2005.