

一种基于 OBGP 协议的光网络域间路由的新型机制

王 磊, 张汉一, 郑小平, 李艳和
(清华大学电子工程系, 北京 100084)

摘 要: 本文提出一种采用光边界网关协议(OBGP)实现 WDM 光格状网的域间路由的新型机制; 该机制基于最小跳选择策略建立最优化路径的思想, 同时引进迂回路由 0 策略和/并行信令 0 方式, 实现了动态、分布式的光通道确立过程. 在 OPNet 平台上针对两种网络拓扑进行仿真实验, 结果表明: 该机制能够在网络繁忙时期降低阻塞率, 提高网络性能; 同时发现: 网络规模、单信道波长数以及网络负载等网络参数对于网络性能具有一定的影响.

关键词: 光边界网关协议; 光通道确立; 域间路由; 格状网; 波分复用

中图分类号: TN929 **文献标识码:** A **文章编号:** 0372-2112 (2004) 05-0718-05

An OBGP-Based Mechanism for Inter-Domain Routing in Optical Networks

WANG Lei, ZHANG Han-yi, ZHENG Xiao-ping, LI Yan-he
(Department of Electronic Engineering, Tsinghua University, Beijing 100084, China)

Abstract: The authors proposed a novel mechanism based on Optical Border Gateway Protocol (OBGP) that implements inter-domain routing in optical mesh networks. Based on an idea of minimum-hop optimal path and a bypass routing during the request course of lightpath setup phase, they achieved the procedure of dynamic and distribution lightpath establishment efficiently in optical networks, which results in a lower blocking probability than before. They also made resource reservation and configured hardware with a/parallel signaling mode, which can reduce the end-to-end delay. An experiment of this mechanism using optical network simulation tools has been presented, which has shown that our study can decrease blocking probability on condition of heavy traffics and make network performance better. It also demonstrates that the three factors, such as network scales, number of wavelengths and traffic loads, have influenced on the network performance. The analysis indicates that a value related to the change of blocking probability approximates the product of end-to-end delay and network scales.

Key words: OBGP; lightpath establishment; inter-domain routing; mesh network; WDM

1 引言

Internet 数据业务的快速增长带动具有大容量、高速率的 WDM 光网络的发展. 今天的智能化光网络一方面要求能够在复杂网络拓扑结构下动态、分布式并且快速、可靠的建立光通道, 另一方面必须进一步提高其可管理性和互动能力. 传统的网络是通过运营商来管理和连通不同用户之间的通信业务. 由于网络规模的扩大, 为用户之间的业务建立光通道往往需要经过多个提供光网络资源(即波长)的运营商, 由于每个运营商根据各自的运营策略, 有可能采用不同的协议与信令; 即使使用相同的协议与信令, 运营商出于保护其商业机密的角度, 也不愿意将其内部网络细节告知其他的运营商, 因此需要一种域间路由协议, 来完成各个运营商之间的路由信息的交互, 以实现跨域的路由. Internet 工程任务组(IETF Internet engi-

neering task force) 通过扩展边界网关协议(BGP Border gateway protocol)^[1] 提出的光边界网关协议(OBGP Optical border gateway protocol)^[2], 以及光互联网论坛(OIF Optical internetworking forum) 近期提出的域间路由协议(DDR Domain to domain routing protocol) 和分层路由协议(HRP Hierarchical routing protocol) 是其中广受关注的解决方案.

然而, 现行的 OBGP 协议仅仅针对普通线性网络, 在应用于复杂的格状网络拓扑时, 我们发现在网络业务高峰期会引起较高的阻塞率. 针对上述问题, 本文提出了一种新型机制, 采取在本地路由信息库中增加最优化路径标识并且设计合理的路由、信令策略, 实现光通道的建立. 通过对不同拓扑网络的仿真实验, 我们对原有机制和新型机制作出比较, 证实了新型机制能够提供较好的网络性能, 如在大量的突发性业务产生时降低网络阻塞率, 其代价是具有稍高的端到端时延.

2 路由和信令模型

本文采用图 1 所示的域间模型^[3]. 该模型将光网络和 IP 网络视为不同的自治域(AS Autonomous systems), 但 IP 网络和光网络具有各自独立的控制平面, 两者之间通过 OBGp 域间路由由协议实现路由信息和信令信息的综合传递; 同时, 将光交叉互连节点(OXC)看作部分位于控制平面内的光路由器. 层叠模型具有光网络层拓扑屏蔽特性, 而对等模型的核心路由器需要维护大量的路由信息导致负担过重. 采用域间模型的 OBGp 协议能够有效解决上述问题, 一方面成为 IP 网络和光网络沟通的纽带, 使得两者之间的信息得以交互; 另一方面由于协议本身对于信息有选择的进行传送, 因此能够最大限度的减轻核心路由器的负担.



图 1 域间模型

3 光通道确立机制

目前 OBGp 采用两种机制确立光通道. 一种是基于原有的 BGP2 更新消息(Update message)^[4], 并引入多协议扩展属性^[5]和扩展团体属性^[6]实现光通道的发现、丢失、建立、删除等功能; 另一种是在原有的四种 BGP 消息的基础上, 定义第五种消息类型 BGP2 OBGp 消息(OBGp message)来实现光通道的确立过程^[7]. 第一种机制的特点是采用了光通道可达性阶段和光通道建立阶段两个先后过程来确立光通道. 这种机制对于线性网络或者网络业务不频繁时是有效的, 但对于突发的 IP 业务来说, 网络的链路状态和资源信息不断变化, 两个过程的采用将会加剧网络局部资源的冲突, 增加网络的阻塞率, 降低网络的性能. 第二种机制则存在新的消息类型与原有消息类型之间的协调问题, 实现的时候需要加以必要的考虑. 针对目前这两种机制存在的问题, 本文借鉴第一种机制的思想, 并提出一种新颖的光通道确立机制.

3.1 本地路由信息库(RIB Routing information base)的扩展结构

在本机制中本地 RIB 中的每一个表项由六个域组成, 其中新增加的是最优化路径标识符域(Optimal path ID), 如图 2 所示. 每一个表项记录的是一条由本地 AS 出发的可达性路径. 其中 Destination 域、OXC_IP 域和 NEXT_HOP 域分别显示去往的目的地 AS、本地 AS 和下一跳 AS 编号. AS_PATH 域为一组显式路径列表: 源 AS(本地 AS) y 中间 AS1 y 中间 AS2y, , y 目的地 AS, 它记录了该可达性路径的节点信息. Extended COMM 域主要包含光通道标识符(Lightpath ID), 记录该可达性路径的沿途各段链路及其空闲波长信息. AS_PATH 和 Extended COMM 域的信息分别来自于光通道可达性信息

携带的 AS_PATH 路径属性和 Extended communities 路径属性.



图 2 本地 RIB 的组成结构

新增的最优化路径标识符域标识了该表项是否为去往相关目的地 AS 的最优路径, 使用该标识可以使 OBGp 协议能够充分应用于格状结构的网络中, 根据网络需要在业务建立过程中按照统一的策略寻求最优化的路由选择, 减少建路过程中的不确定因素, 从而提高网络性能. 最优化路径的选择策略可以根据运营商的要求, 如相互利益、运营代价以及技术等, 也可以根据路径因素, 如距离最短、跳数最少、网络拥塞程度最低等来确定. 本文采用基于最小跳的优化选择策略, 其细节将在下面的分析中阐述.

3.1.2 光通道可达性阶段(Lightpath reachability phase)

光通道可达性阶段包括光通道发现(discovery)和丢失(withdrawal)两种情况, 如图 3 所示. 在机制中, 我们将一个 AS 看作一个抽象节点. 为了避免核心路由器负担过重, 更新消息(Update message)采用有选择的发送: 当且仅当一条有向链路的末端 AS 发现空闲波长由无到有(一个或多个)或者由有(一个或多个)到无时, 才向链路始端发送发现更新消息(Reachability discovery message)或者丢失更新消息(Reachability withdrawal message).

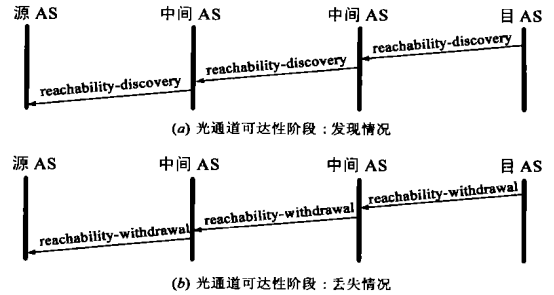


图 3 光通道可达性阶段

3.1.2.1 发现情况 当链路末端 AS 向始端 AS 发送发现更新消息时, 在 AS_PATH 路径属性中添加始端、末端 AS 信息, 在 Extended communities 路径属性中添加链路及其空闲波长信息. 收到发现更新消息的 AS 只对 AS_PATH 路径属性中不含有本地 AS 的消息进行处理. 首先, 根据消息携带的 AS_PATH 和 Extended communities 路径属性在本地 RIB 中添加或更新相应表项. 然后, 在本地 RIB 中对于对应于同一个目的地 AS 的全部表项, 选择其中 AS_PATH 的长度最短(即路径经过的 AS 的跳数最少)的表项, 将其最优化路径标识置为 ON(该表项称作最优化路径表项), 剩余的置为 OFF, 这样就得到了本地 AS 去往所有可达 AS 的最优化路径信息. 最后, 在消息携带的 AS_PATH 和 Extended communities 路径属性中分别增加本地 AS 信息和本地 AS 至上一跳 AS 链路及其空闲波长信息, 并且向相邻可达 AS 转发修改后的消息.

3.1.2.2 丢失情况 当链路末端 AS 向始端 AS 发送丢失更新消息时, 在 Extended communities 路径属性中添加链路信息. 收到丢失更新消息的 AS 在本地 RIB 中删除包含相关链路的表

项.若删除的表项为最优化路径表项,则需要为去往该表项对应的目的地AS选择新的最优化路径表项.最后,向相邻可达AS转发该消息.

313 光通道建立阶段(Lightpath setup phase)

光通道建立阶段包括建立请求过程(Request)和建立预留过程(reservation),如图4所示.

3131 建立成功情况 当AS收到来自本地应用层的业务请求时,查询本地RIB中去往请求目的地的最优化路径表项.根据选定的表项的NEXT_HOP域、Destination域、AS_PATH域和Extended_COMM域构造一个建立请求消息(Setuprequest message),并向下一跳发送.如果需要的话,在本地AS中为去往选定的表项对应的目的地AS选择新的最优化路径表项.

在建立请求过程中,各中间AS同时需要检查本地AS至下一跳AS之间是否存在空闲光通道,如果不存在,则实行迂回路由0策略:查询本地RIB,选定一条新的由本地去往目的地AS的路径,并相应修改AS_PATH和Extended_communities路径属性;向下一跳转发该请求消息.

当目的地AS收到建立请求消息后,构造一个新的建立预留消息(Setupreservation message):复制AS_PATH和Extended_communities路径属性.

在建立预留过程中,目的地AS和各中间AS根据本地Lightpathstatusin标识选定一个可用波长作为上一跳AS至本地AS之间的链路的业务承载波长,并修改Extended_communities路径属性的光通道标识段;在进行硬件配置(如动作光开关等)之前向上一跳转发该预留消息,这体现了一种并行信令的思想,能够有效降低端到端平均时延;为了更新光通道可达性信息,需要发送丢失更新消息,并且设置等待时间参数DWTime(用于目的地AS)或IWTime(用于中间AS),并启动计时器.

当源AS收到建立预留消息后,则建立成功,开始传送数据.中间AS和目的地AS在计时器超时等待前收到数据信息,关闭计时器;如果发现传送数据的链路上存在空闲波长,则向上一跳发送发现更新消息.

31312 建立失败情况 建立失败包括两种情况:请求失败情况(图4(b))和预留失败情况(图4(c)).

在建立请求过程中,若中间AS建立失败(如无去往目的地的空闲光通道或硬件故障等),则构造一个新的建立失败消息(Setupwithdrawal message):复制AS_PATH和Extended_communities路径属性,向上一跳发送该失败消息.

在建立预留过程中,若中间AS预留失败(如硬件故障等),则终止预留过程.等待目的地AS计时器超时,发送建立失败消息.若目的地AS或中间AS计时器超时,发现去往上一跳AS存在空闲光通道,则向上一跳发送发现更新消息.

收到建立失败消息的AS,根据AS_PATH和Extended_communities路径属性恢复本地RIB,并确定最优化路径表项;向上一跳转发该失败消息.

因此若源AS收到建立失败消息,则建立失败.

4 实验环境

新机制可以适用于各种格状网络拓扑结构,我们采用在

OPNet平台对两种网络进行了仿真实验,网络拓扑如图5所示.假设各个AS都支持波长变换功能以及OBGP协议.业务选用均衡模式(Uniform traffic),定时器参数DWTime和IWTime均设定为0.05ms(网络1)或0.08ms(网络2),仿真时间为4ms.实验时,让网络的初始状态尽量为空闲,然后突然产生大量的动态业务请求,由此比较原机制和新机制的网络性能,以及网络规模、单信道波长数以及网络负载等网络参数对于网络性能的影响.

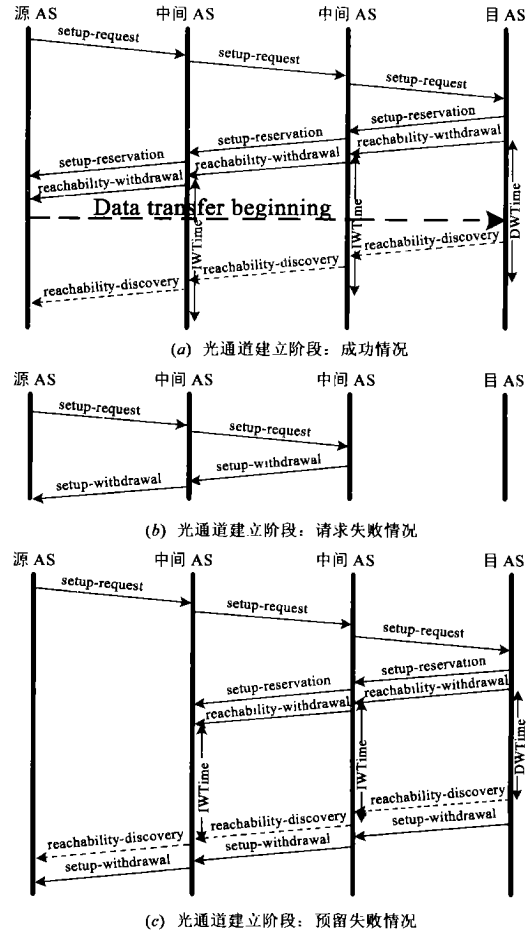


图4 光通道建立阶段

实验中网络性能参数(均为统计量表示)的定义如下:
端到端平均时延:

$$\text{Mean. Delay} = \frac{\sum_{i=1}^{\text{Success_Traffic_Nums}} \text{Delay}_i}{\text{Success_Traffic_Nums}} \quad (1)$$

阻塞率:

$$\text{Blocking_Probability} = \frac{\text{Failure_Traffic_Nums}}{\text{Total_Traffic_Nums}} \quad (2)$$

总业务数:

$$\text{Total_Traffic_Nums} = \text{Success_Traffic_Nums} + \text{Failure_Traffic_Nums} \quad (3)$$

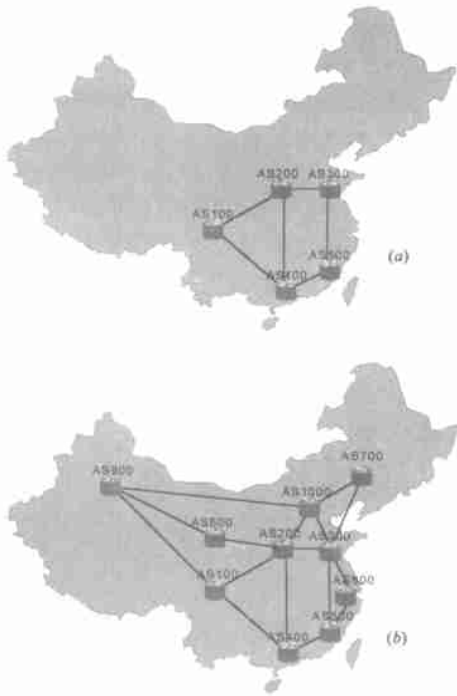


图 5 仿真实验网络拓扑结构(支持 OBGp 的路由器数目:
(a)N= 5;(b)N= 10)

5 实验结果与分析

OBGp 协议的实验结果如图 6, 7, 8 所示, 其中实线表示新机制, 虚线表示原机制. 可以看到在同样的条件下, 新机制的阻塞率比原机制低, 并且网络 2 的阻塞率比网络 1 的阻塞率低. 这说明新机制使网络性能变得优越; 同时, 新机制在规模较大的网络中体现出的性能更佳. 这是因为新机制中(最优路径)策略能够有效利用网络资源, 从而降低阻塞率.

图 6 中业务到达的平均间隔时间为 0. 05ms. 可以看到, 随着单纤波长数变大(从 W= 2 到 W= 8), 网络阻塞率降低, 并且在新机制中网络阻塞率降低得更快. 这是因为在新机制中,

每个 AS 都在本地 RIB 中存储可能多于一个到达相同目的地的不同表项, 而波长资源越丰富, 在同一根光纤中就可以同时存在越多的业务. 因此, 这些不同的路由表项便可以被更加充分的利用.

图 7 为阻塞率随业务繁忙程度的变化, 随着每秒钟平均请求业务的增加, 网络平均阻塞率呈上升趋势. 需要指出的是, 当每秒钟业务的平均请求数小于 $1E+ 4$ (即业务到达的平均时间间隔大于 0. 1ms) 时, 网络阻塞率趋近于零; 而当平均请求数大于 $1E+ 4$ (即时间间隔小于 0. 1ms) 时, 网络阻塞率急剧增大. 我们注意到网络规模 N 的数量级为 $1E+ 1$, 而从图 8 中可以看到, 端到端的平均时延略大于 0. 01ms, 此二者的乘积即为 0. 1ms. 由于实行(并行信令)策略, 当业务到达的平均时间间隔大于端到端时延和网络规模的乘积时, 网络资源相对空闲, 对于多数业务请求都可以承载, 因此阻塞率趋近于零; 当业务到达的平均时间间隔小于该乘积时, 网络资源相对繁忙, 对于业务请求的承载能力下降, 因此越来越多的业务被阻塞.

图 8 显示业务建立的端到端的平均时延几乎不随每秒钟业务的平均请求数的变化而变化, 并且网络规模扩大显而易见的导致时延的增加. 而当每秒钟业务请求数超过 $1E+ 4$ 时, 新机制的端到端时延略高于原机制. 这是因为在业务繁忙时, 业务之间的竞争和冲突是不可避免的, 而新机制提供的(迂回路)策略可以为冲突的业务提供新的路由路径, 这会在一定程度上引起额外的端到端时延.

6 结论

通过在 OPNet 平台上的 WDM 光网络实验, 我们发现在基于 OBGp 协议的域间路由中, 采用最优路径标识符的光通道确立机制, 能够在网络业务繁忙时期降低网络阻塞率, 提高网络性能. 同时发现: 网络规模扩大会导致阻塞率下降, 端到端时延增加; 单纤波长数增多, 网络阻塞率下降; 业务负载低于某个确定的数值时, 阻塞率趋近于零, 而高于该数值时, 阻塞率急剧上升, 并且业务负载对于端到端时延的影响不大. 在实验基础上, 本文分析了产生这些结果的理论依据.

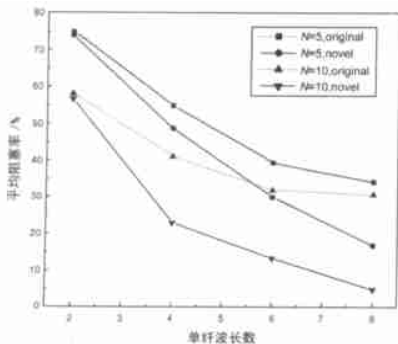


图 6 阻塞率随单纤信道数变化图

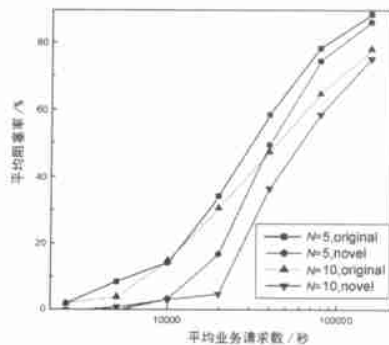


图 7 阻塞率随业务繁忙程度变化图

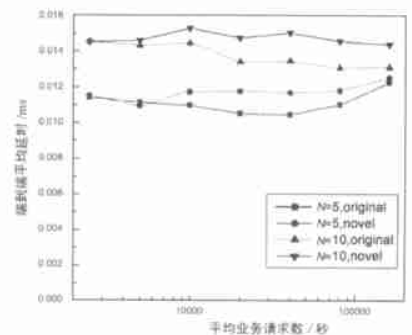


图 8 端到端时延随业务繁忙程度变化图

参考文献:

- [1] REKHTER Y, et al. IETF RFC1771, A border gateway protocol 4 (BGP24) [S]. March 1995.
- [2] BLANCHET M, et al. Iet2draft2paren2obgp201. Optical BGP (OBGP): interAS lightpath provisioning [S]. August 2001.
- [3] RAJAGOPALAN B, et al. IP over optical networks: architectural aspects [J]. IEEE Comm, 2000, 38(9): 94- 108.
- [4] FRANCISCO M J, et al. Interdomain routing in optical networks [A]. Proceedings of SPIE 2Opticomm. 2001 [C]. Denver, CO, USA, 4599, 2001. 120- 129.
- [5] BATES T, et al. IETF RFC2858, Multiprotocol extensions for BGP24 [S]. June 2000.
- [6] RAMACHANDRA S, et al. Iet2draft2ramachandra2bgp2ext2commun2ties208. BGP extended communities attribute [S]. June 2001.
- [7] FRANCISCO M J, et al. End2to2end signaling and routing for optical IP networks [A]. IEEE International Conferenc2ICC. 2002 [C]. New York, NY, USA, 2002. 5. 2870- 2875.

作者简介:



王 磊 男, 1978 年 9 月生于黑龙江哈尔滨, 2001 年 7 月于清华大学电子工程系获学士学位, 现为清华大学电子工程系博士研究生, 研究兴趣有智能化光网络及相关协议、域间路由和信令技术、WDM 光网络软件仿真研究等。

张汉一 男, 1942 年生于黑龙江绥化, 1965 年毕业于清华大学无线电系, 现为清华大学电子工程系教授, 博士生导师, 先后负责承担十余项国家攻关、863 和自然科学基金等任务, 目前的研究领域为 WDM 光通信网络的结构设计与规划、波长路由及动态重构、系统运行及动态传输特性、光网络节点、光网络特性监测、管理、保护、恢复与光网络的安全性、WDM 光网络新型光电子功能器件/ 模块、IP over WDM 等。