

多域光网络中应用混合拓扑聚合的分段保护算法*

张盛峰^{1,2}, 杜书¹

(1. 电子科技大学 宽带光纤传输与通信网技术重点实验室, 成都 611731; 2. 重庆邮电大学 光纤通信重点实验室, 重庆 400065)

摘要: 研究了多域光网络中的路由保护问题; 为了克服多域光网络中可扩展性约束, 提出了一种混合拓扑聚合方法。该方法结合了全连通和生成树拓扑聚合的优点, 在网络中需要存储和发布的链路状态信息与聚合信息反映实际物理拓扑的精确性之间进行了折中; 然后在此混合拓扑聚合方法的基础上, 提出了一种基于查询机制的多域分段保护算法。仿真表明, 相比传统的多域保护算法, 该算法阻塞率低, 可扩展性好。

关键词: 多域光网络; 全连通拓扑聚合; 生成树拓扑聚合; 混合拓扑聚合; 分段保护

中图分类号: TP393 **文献标志码:** A **文章编号:** 1001-3695(2012)09-3369-03

doi:10.3969/j.issn.1001-3695.2012.09.044

Segment-shared protection algorithm based on hybrid topology aggregation for multi-domain optical mesh networks

ZHANG Sheng-feng^{1,2}, DU Shu¹

(1. Key Laboratory of Broadband Optical Fiber Transmission & Communication Networks, University of Electronic Science & Technology of China, Chengdu 611731, China; 2. Key Laboratory of Optical Fiber Communication, Chongqing University of Posts & Telecommunications, Chongqing 400065, China)

Abstract: In order to deal with the survivability problem in multi-domain optical mesh networks, this paper proposed a new hybrid topology aggregation scheme at first. This scheme combined the advantages of the full mesh model and spanning tree model to tradeoff between accuracy in preserving the original topological features and compactness in representation to achieve scaling. Then it proposed an improved segment-shared protection approach called HTASP. Simulation evaluation shows that the proposed scheme is effective in multi-domain path protection with lower blocking probabilities and less inter-domain link update messages.

Key words: multi-domain optical mesh networks; full mesh topology aggregation; spanning tree hybrid topology aggregation; hybrid topology aggregation; segment-shared protection

0 引言

随着光网络单一光纤上承载的数据业务量不断增加, 网络故障造成的损失也愈加严重。因此网络生存性是光网络设计中需要重点考虑的问题。保护机制通过在连接建立初期就给工作通路预先保留保护资源, 一旦工作通路失效, 便立即将业务切换到保护通路上来恢复业务连接, 以实现网络的抗毁性能^[1-4]。随着光网络规模的不断扩大, 网络向着分层多域的层次路由体系结构发展。出于可扩展性和安全性考虑, 每个路由域的拓扑信息首先通过特定的拓扑聚合算法加以抽象才扩散到网络中的其他路由域。这样, 在多域光网络中, 每个路由域内的节点可分成内部节点(interior node)和网关节点(gateway node)。内部节点只获得本域的物理拓扑信息, 而网关节点既具有域内拓扑信息又拥有域间虚逻辑拓扑(virtual topology)信息^[5,6]。传统的应用于单域网络环境下、假定源节点具有全网详细拓扑信息的保护算法^[7,8]已不适合多域光网络。基于多域层次路由, 文献[9]中提出了一种多域扩展共享通路保护算

法(EPSP)。该算法分两步: 首先在多域拓扑聚合(topology aggregation)构成的虚逻辑拓扑上计算两条路径分离的虚通路; 然后将两条虚通路在其经过的域中扩展成完整的物理通路。这样就得到了一对多域端到端的链路分离的工作通路和保护通路。EPSP算法保护恢复时间长, 且无法对多链路故障进行保护。文献[10]中提出了一种本域分段共享保护算法(LSSP)。LSSP算法首先计算一条工作通路, 然后在工作通路所经过的每个域中对该域中的工作路段进行分段共享保护。LSSP相对于EPSP算法有较快的保护恢复时间, 且能对多链路故障(即当网络中每个域中有且仅有一条链路故障)进行保护。但其多域选路采用的是逐域层次路由, 这种方法得出的工作和保护通路往往不是最优的。而且上述保护算法中, 域间虚通路的计算都是基于全连通(full mesh)拓扑聚合的多域虚逻辑拓扑。全连通拓扑聚合方法采用单个域的边界节点对之间的逻辑虚链路来构建汇聚拓扑, 其较准确的逻辑链路信息能进行有效的路由。但是, 全连通拓扑聚合需要洪泛的链路信息是与网关节点的个数的平方成正比的, 因此, 它很可能由于网关

收稿日期: 2012-01-04; **修回日期:** 2012-02-15 **基金项目:** 国家自然科学基金资助项目(61071117, 60972069); 重庆市科委自然科学基金资助项目(CSTC2010BB2413); 重庆市教委资助项目(KJ110527); 重庆邮电大学自然科学基金资助项目(A2009-32, A2009-38)

作者简介: 张盛峰(1977-), 男, 湖北荆州人, 讲师, 博士研究生, 主要研究方向为光网络(zhangsf@cqupt.edu.cn); 杜书(1983-), 男, 江西樟树人, 博士研究生, 主要研究方向为光网络。

节点数量的增加而产生链路状态信息膨胀。而生成树(spanning tree)拓扑聚合方法需要洪泛的链路信息与网关节点数目成线性关系,但其聚合拓扑信息的准确性较全连通次之。本文结合两种拓扑聚合方法的优点,提出了一种混合拓扑聚合方法,并在此基础上提出了一种基于混合拓扑聚合的分段保护算法(HTASP)。

1 问题描述

1.1 混合拓扑聚合方法

拓扑聚合技术是多域光网络路由技术的核心技术之一,它将各路由域的拓扑信息进行压缩,域间交换压缩后的信息,从而使路由协议交换的链路状态信息的数量大大减少,保证网络的可扩展性。典型的拓扑聚合方法有全连通拓扑、对称星型(symmetric star)拓扑、非对称星型(asymmetric star)拓扑以及生成树拓扑^[11],如图1所示。全连通拓扑聚合只保留单域的网关节点,在每一对网关节点之间构造一条虚链路,该虚链路的链路参数根据实际网络拓扑中连接对应两个节点的部分或所有路径的属性来确定,其空间复杂度为 $O(|N|^2)$, N 为单个路由域中网关节点的集合。该聚合方法需要扩散的链路状态信息与网关节点的数目的平方成正比,因此路由消息开销大。当网关节点数远小于内部节点数时,该方法既达到了拓扑压缩目的,又保留了最多的有用信息。但当网关节点与内部节点相当或远大于内部节点时,例如在环形网络中,每个节点都是网关节点,该算法既没有有效压缩信息,又损失了部分信息。极端情况下使用全连通算法甚至会增加冗余信息,与压缩相违背^[12]。

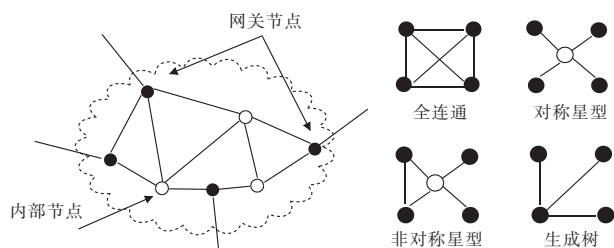


图1 典型的拓扑聚合方法

在采用生成树抽象方法的情况下,带有 $O(|N|^2)$ 个逻辑链路的全连通聚合拓扑将被进一步抽象压缩为带有 $O(|N|)$ 个逻辑链路的生成树聚合拓扑。这样需要洪泛的链路信息与网关节点数目成线性关系。生成树算法可基于节点间最大化带宽,或者基于最小化延迟。但是它的不足之处在于算法可以完全保留必要的凹性(concave)QoS属性信息(如带宽等),从而还原成与原来一样的全连通聚合拓扑;但对于时延等加性(addictive)QoS属性将会损失部分必要信息。

拓扑聚合技术大大减少了网络中需要存储和发布的链路状态信息。但是,由于聚合信息往往不是实际物理拓扑的精确反映,从而导致根据此信息选择的可行路由实际上并不是满足QoS的最优路由。性能良好的拓扑聚合方法试图在两者之间找到最佳的平衡点。为在网络信息准确度和可扩展性之间进行更好的折中,本文结合两种拓扑聚合方法的优点提出了一种混合拓扑聚合方法,其方法描述如下:

a) 当单域网关节点数 $|GN| < 4$ 时,采用全连通拓扑聚合方法。

b) 当单域网关节点数 $|GN| \geq 4$ 时,如果域内链路数 $|IL| < \alpha|GN|(|GN| - 1)$ 时,采用生成树拓扑聚合;否则采用全连通

拓扑聚合方法。其中 α 表示拓扑压缩因子,可根据实际需要适当选取。

1.2 HTASP 基本思想

HTASP 保护算法分两步,即先为业务连接计算工作通路,再计算保护工作路段。参照文献[10]的假设,在网络规划阶段已经充分考虑了域间链路的可靠性(已提供 $1+1$ 的保护),算法仅需要为业务连接在各路由域内的工作路段提供共享保护。首先发起连接请求的源节点向它所在的路由域(称为源域)内距它最近的一个网关节点发送一个查询消息(query message),请求计算域间通路。该网关节点依据它所掌握的基于混合拓扑聚合的多域虚逻辑拓扑信息计算一条源域到达目的节点所在路由域(称为目的域)的最小代价虚通路(loose route)。该虚通路选定一个距目的域最近的源网关节点作为本次连接的出网关节点(该网关节点不一定与接受路由查询的网关节点相同,因为距源节点最近的网关节点并不一定到目的域距离最近或代价最小),并指明了这条跨域通路需要经过的各路由域的网关节点。然后源节点根据查询得到虚通路,找到一条到出网关代价最小的工作路段。随后在虚通路所经过的每个路由域内,由该域的入网关节点(ingress gateway)完成本路由域内的物理工作路段的扩展,扩展后的工作路段指明了其在该域内所需经过的所有节点。最后由目的域的入网关节点完成该网关节点到目的节点工作路段的计算。当整个工作通路计算完成后,在工作通路所经过的每一个域中,为其在这个域中的工作路段计算一条链路分离的共享保护路段。

1.3 多域光网络模型

在本文中,多域光网络由物理拓扑 $G = (\cup_{i=1}^m D_i, L^{\text{INTER}})$ 和逻辑虚拓扑 $G^v = (\cup_{i=1}^m D_i^v, L^{\text{INTER}})$ 表示,两个拓扑都包含有 m 个域。其中: $D_i = (GN_i, IN_i, IL_i, W)$, $D_i^v = (GN_i, VL_i)$, $i = 1, 2, \dots, m$; L^{INTER} 表示连接不同域的网关节点的域间链路; GN_i 为域 i 中的网关节点; IN_i 为域 i 中的内部节点; IL_i 为域 i 中的内链路; W 为每条光纤配置的波长数目; VL_i 为逻辑虚拓扑中域 i 中的虚链路。

在描述算法前,先引入以下算法中需要用到的符号和变量:
 $CR_{s,d}^n$ 表示第 n 次业务请求,此业务请求的源节点为 s ,目的节点为 d 。

$WP_{s,d}^n$ 表示第 n 次业务请求 $CR_{s,d}^n$ 的工作通路。

$BP_{s,d}^n$ 表示第 n 次业务请求 $CR_{s,d}^n$ 的保护通路。

$SW_{v_1,v_2}^{n,i}$ 表示 $WP_{s,d}^n$ 在域 i 中从节点 v_1 到 v_2 之间的工作路段。如果源节点 s 和目的节点 d 都在域 i 中,那么 $SW_{v_1,v_2}^{n,i}$ 正好是工作通路,即 $SW_{v_1,v_2}^{n,i} = WP_{s,d}^n$ 。

$SB_{v_1,v_2}^{n,i}$ 表示 $BP_{s,d}^n$ 在域 i 中从节点 v_1 到 v_2 之间的保护路段。如果源节点 s 和目的节点 d 都在域 i 中,那么 $BP_{s,d}^n$ 正好是保护通路。

cost_e 表示链路 e 的代价。

c_e 表示虚链路的代价, $c_e = 1$,如果虚链路 e 能够扩展为物理链路的集合。

f_e 表示链路 e 上空闲的波长数。

b_e 表示链路 e 上预留留给保护通路的波长数。

$|\Omega|$ 表示集合 Ω 中的元素的个数。

$\phi_{e,i}^f(f, e \in \Omega_i)$: 经过域 i 的内部链路 f ,且其保护路段经过同一域内链路 e 的工作路段集合。

$b_e = \max \{ |\varphi_{e,i}^f|, e \in \Pi_i, \forall f \in \Pi_i, f \neq e \}$ 表示域 i 中链路 e 上已预留给所有保护通路的波长数。

$v_e^* = \max \{ |\varphi_{e,i}^f + \varphi_{e,i}^g|, e \in \Pi_i, \forall (f \in \Pi_i, g \in \text{SW}_{v_1,v_2}^{n,i}), f \neq e, g \neq e \}$ 表示如果保护当前业务请求工作路段的保护路段经过链路 e , 则链路 e 上将预留给域中所有保护路段的波长数。

HTASP 算法描述如下:

输入: $D_i = (\text{GN}_i, \text{IN}_i, \Pi_i, \text{W}_i), i = 1, 2, \dots, m, G^v$, and $\text{CR}_{s,d}^n$

输出: 一条工作通路和一条保护通路; 或者一条工作通路和多个保护路段。

a) 如果 $s, d \in \text{IN}_i (\exists i \in \{1, 2, \dots, m\})$, 则 $v_1, v_2 \leftarrow s, d$, 执行 b); 否则执行 e)。

b) 根据式(1)调整链路代价, 在域 i 用 Dijkstra 算法寻找从节点 v_1 到节点 v_2 的一条最短工作路段 $\text{SW}_{v_1,v_2}^{n,i}$ 。如果该工作路段被找到, 则执行 c); 否则返回 null。

$$\text{cost}_e = \begin{cases} (|\text{W}| + 1 - f_e) / |\text{W}| & \text{if } (f_e > 0) \\ +\infty & \text{otherwise} \end{cases} \quad (1)$$

c) 根据式(2)调整链路代价, 在域 i 中用 Dijkstra 算法寻找一条与工作路段 $\text{SW}_{v_1,v_2}^{n,i}$ 链路分离的共享通路保护路段 $\text{SB}_{v_1,v_2}^{n,i}$ 。如果该保护路段被找到, 则执行 d); 否则返回 null。

$$\text{cost}_e = \begin{cases} (v_e^* + 1 - b_e) / |\text{W}| & \text{if } (e \notin \text{SW}_{v_1,v_2}^{n,i} \text{ and } f_e + b_e \geq v_e^*) \\ +\infty & \text{otherwise} \end{cases} \quad (2)$$

d) 如果 $s, d \in \text{IN}_i (\exists i \in \{1, 2, \dots, m\})$, $(\text{WP}_{s,d}^{n,i}, \text{BP}_{s,d}^{n,i}) \leftarrow (\text{SW}_{v_1,v_2}^{n,i}, \text{SB}_{v_1,v_2}^{n,i})$, 返回已计算出的工作通路和保护通路; 否则记录下该工作和保护路段后, 在工作通路 $\text{WP}_{s,d}^{n,i}$ 中选出邻接域中的工作路段 $\text{SW}_{g_1,g_2}^{n,j} \subset \text{WP}_{s,d}^{n,i} (j \in \{1, 2, \dots, m\})$ 。如果所有的工作路段都迭代完毕, 返回计算出的工作通路和对应的多条保护路段; 否则将 $(v_1, v_2, i \leftarrow g_1, g_2, j)$, 跳回执行 c)。

e) 根据式(3)调整链路 $e (\forall e \in \text{VL} + L^{\text{INTER}})$ 的代价, 寻找从源域网关节点到目的域网关节点的代价最小通路 L 。如果 L 找到, 记录通路 L 上的源域出网关节点 g 和目的域入网关节点 k , 执行 f); 否则返回 null。

$$\text{cost}_e = \begin{cases} 1 & \text{if } (e \in \text{VL}_y \text{ and } c_e = 1 \forall y) \\ (|\text{W}| + 1 - f_e) / |\text{W}| & \text{if } (e \in L^{\text{INTER}} \text{ and } f_e \geq 0) \\ +\infty & \text{otherwise} \end{cases} \quad (3)$$

f) 根据式(1)调整链路代价, 在域 i 中用 Dijkstra 算法寻找一条从源节点 s 到源域出网关节点 g 的最短路。如果找到该工作路段, 记录这条最短路为工作路段 $\text{SW}_{s,g}^{n,i}$, 执行 g); 否则返回 null。

g) 将通路 L 中所有虚链路依次在对应的域中扩展为物理工作路段。根据式(1)调整目的域 t 中链路代价, 用 Dijkstra 算法寻找从目的网关节点 k 到目的节点 d 的代价最小通路。如果该通路找到, 记录该通路为工作路段 $\text{SW}_{k,d}^{n,t}$, 执行 h); 否则返回 null。

h) 将已找到的源域中的工作路段、扩展后的通路 L 、目的域中的工作路段连接成节点 s 到节点 d 的工作通路 $\text{WP}_{s,d}^n$ 。将 $(\text{SW}_{v_1,v_2}^{n,i} \leftarrow \text{SW}_{s,g}^{n,i})$, 跳回执行 c)。

2 仿真和分析

仿真时假设所有业务请求的到达速率服从均值为 β 的泊松分布, 所建业务连接的持续时间服从均值为 $1/\mu$ 的指数分

布, 即全网总负载为 $\beta/\mu = \beta$ 爱尔兰 (Erlang), 仿真时取 $\mu = 1$ 。到达业务请求的源、目的节点在所有节点对之间随机选择, 如果业务连接建立失败则立即丢弃, 即无等待队列。仿真网络拓扑如图 2 所示, 图中的多域光网络包含 5 个域。其中每条域内链路为一根双向光纤, 任意两个域之间的域间链路由一根双向工作光纤和一根双向专用保护光纤组成; 每根光纤支持 16 个波长, 每个业务请求的带宽都是 1 个波长粒度。假设网络中所有节点均具有波长变换能力, 仿真时取拓扑压缩因子 $\alpha = 3.4$ 。与 HTASP 比较的算法是 LSSP, 仿真所得结果是在模拟 10^6 次业务请求后得到的平均值。仿真中比较以下性能: 阻塞率 (blocking probability, BP) 和域间链路更新消息数 (number of inter-domain link update messages, NIDLUM)。

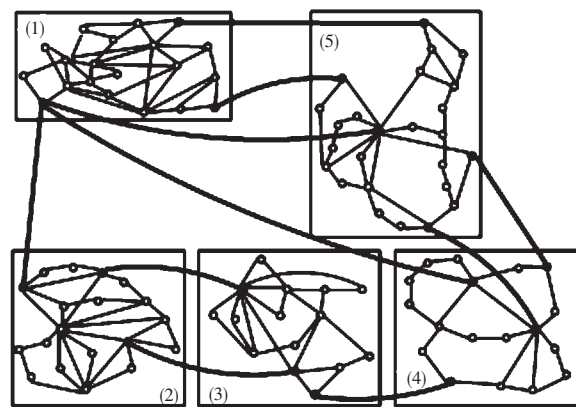


图 2 LARGE-5 网络拓扑^[13]

从图 3 可以看出, 在网络负载变化时 HTASP 的阻塞率比 LSSP 小, 因为 LSSP 在计算多域工作通路时, 在源域选择的出网关大多数情况下并不是到目的域最优的网关, 即在虚拓扑上距目的域最近。所以 LSSP 计算出的工作通路较长, 从而找到的保护通路更长, 进而导致为单个业务请求分配的网络资源 (波长) 就多。

从图 4 可以看出单位时间内域间虚链路更新消息数在网络负载变化时, HTASP 要小于 LSSP。因为 LSSP 采用的拓扑聚合方法为全连通方法, 而 HTASP 采用的是全连通和生成树混合聚合方法。混合聚合方法产生的逻辑虚拓扑上的虚链路的个数要小于全连通方法, 因而为更新网络链路状态而产生的消息数量也要小于 LSSP。

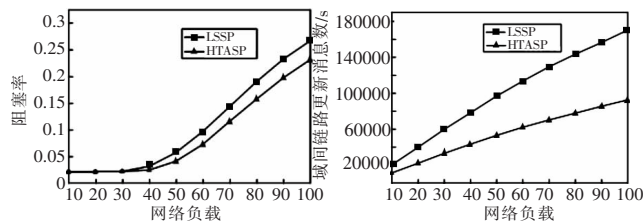


图 3 阻塞率(BP)随网络负载变化

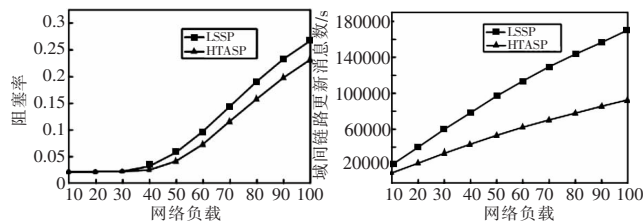


图 4 域间链路更新消息数(NIDLUM)随网络负载变化

3 结束语

网络生存性是光网络设计中需要重点考虑的问题。随着光网络规模的不断扩大, 网络向着分层多域的层次路由体系结构发展。为了克服多域光网络中可扩展性约束, 本文提出了一种混合拓扑聚合方法。该方法结合了全连通和生成树拓扑聚合的优点, 在网络中需要存储和发布的链路状态信息与聚合信息反映实际物理拓扑的精确性之间进行了折中。 (下转第 3375 页)

(上接第3371页)并在此混合拓扑聚合方法的基础上,提出了一种基于查询机制的多域分段保护算法。仿真表明,相比 LSSP 该算法阻塞率低,可扩展性好。

参考文献:

- [1] DOSHI B T, DRAVIDA S, HARSHAVARDHANA P, *et al.* Optical network design and restoration[J]. *Bell Labs Technical Journal*, 1999, 4(1):58-84.
- [2] ZHOU Dong-yun, SUBRAMANIAM S. Survivability in optical networks[J]. *IEEE Network*, 2000, 14(6):16-23.
- [3] AL-KUWAITI M, KYRIAKOPOULOS N, HUSSEIN S. A comparative analysis of network dependability, fault-tolerance, reliability, security, and survivability [J]. *IEEE Communications Surveys & Tutorials*, 2009, 11(2):106-124.
- [4] 赵太飞,李乐民,虞红芳. 基于 k -最短路路由的 mesh 光网络 P 圈构造方法[J]. *计算机应用研究*, 2007, 24(11):278-280.
- [5] LIU Q, KOK M, GHANI N, *et al.* Hierarchical routing in multi-domain optical networks[J]. *Computer Communications*, 2006, 30(1):122-131.
- [6] HAMZA D, BERNARD C, MIKLOS M, *et al.* A survey of survivability in multi-domain optical networks[J]. *Computer Communications*, 2010, 33(8):1005-1012.
- [7] RAMAMURTHY S, SAHASRABUDDHE L, MUKHERJEE B. Survivable WDM mesh networks[J]. *IEEE Journal of Lightwave Technology*, 2003, 21(4):870-883.
- [8] OU Can-hui, ZHANG Jing, ZANG Hui, *et al.* New and improved approaches for shared-path protection in WDM mesh networks[J]. *IEEE Journal of Lightwave Technology*, 2004, 22(5):1223-1232.
- [9] TRUONG D L, THIONGANE B. Dynamic routing for shared path protection in multidomain optical mesh networks[J]. *Journal of Optical Networking*, 2006, 5(1):58-74.
- [10] GUO L. LSSP: a novel local segment-shared protection for multidomain optical mesh networks[J]. *Computer Communications*, 2007, 30(8):1794-1801.
- [11] 杨敏, 向勇, 史美林, 等. 层次网络中的拓扑压缩算法及性能比较[J]. *计算机科学*, 2002, 29(12):44-48.
- [12] JUNG E-S, RANKA S, SAHNI S. Topology aggregation for e-science networks[C]//Proc of the 10th IEEE/ACM International Conference on Cluster, Cloud and Grid Computing. Washington DC: IEEE Computer Society, 2010:530-533.
- [13] TRUONG D L, THIONGANE B. Using topology aggregation for efficient segment shared protection solutions in multi-domain networks[J]. *IEEE Journal on Selected Areas In Communications*, 2007, 25(9):96-107.