

# 基于 STP 协议的物理网络拓扑发现算法

张占国, 刘淑芬, 包 铁, 张欣佳, 王晓燕, 张 婵

(吉林大学计算机科学与技术学院, 长春 130012)

**摘 要:** 拓扑发现是网络管理的重要基础。该文提出一种基于网桥生成树协议 STP 的算法, 利用简单网络管理协议(SNMP)获得各个交换机 MIB 库中的生成树状态信息, 根据生成树协议推导出网络的物理拓扑。和已有方法相比, 该算法不要求各个网桥 FDB 表的信息是完备的, 同时也能很好地发现备份链路和集线器、哑交换机等不支持 SNMP 的设备。实验表明该算法是一个准确、全面的拓扑发现算法。

**关键词:** STP 协议; 物理拓扑发现; 简单网络管理协议; 网络管理

## Physical Topology Discovery Algorithm Based on Spanning Tree Protocol

ZHANG Zhan-guo, LIU Shu-fen, BAO Tie, ZHANG Xin-jia, WANG Xiao-yan, ZHANG Chan

(College of Computer Science and Technology, Jilin University, Changchun 130012)

**【Abstract】** Topology discovery is an important foundation for network management. This paper proposes a new physical topology discovery algorithm relying on Spanning Tree Protocol(STP). The algorithm gets spanning tree status information of each switch by SNMP. According to STP, physical topology of the network is derived. Comparing with other existing algorithms, the algorithm does not require that the FDB information of each bridge is complete. Meanwhile, it can find backup link and equipments that do not support SNMP such as hub and dump switches. Experiments show that the algorithm is an accurate, comprehensive topology discovery algorithm.

**【Key words】** Spanning Tree Protocol(STP); physical topology discovery; SNMP; network management

### 1 概述

随着网络技术的飞速发展, 局域网内设备的连接情况变得十分复杂。找出网络元素之间的实际物理链接是完成许多重要网络管理任务的前提, 例如: 网络管理, 服务器定位, 事件关联等<sup>[1]</sup>。

目前在网络拓扑发现方面主要分为逻辑层网络拓扑发现和物理层网络拓扑发现。逻辑层网络拓扑自动发现算法只能发现网络层(对应OSI<sup>[2]</sup>的第 3 层)设备及它们之间的连接关系, 即路由器到路由器、路由器接口到子网的连接关系。网络层设备拓扑发现的方法比较简单, 遍历网络层设备的路由表即可获得。物理网络拓扑<sup>[2]</sup>指的是一个通信网内部实体的实际物理连接(对应OSI第 2 层)。物理网络拓扑是在网络层拓扑的基础上增加交换机到交换机、交换机到路由器、交换机到主机之间的连接关系。

构建以太网所使用的第 2 层互连设备称为网桥, 它支持 STP(Spanning Tree Protocol)<sup>[3]</sup>协议, 具有地址学习能力。随着网络的发展, 出现了以太网交换机, 它实质上是一种多端口的透明网桥。网桥为了完成数据帧的转发工作, 各自维护自己的生成树状态表(STP)和 MAC 地址转发表(FDB)<sup>[3]</sup>, 并且保存在标准 BRIDGE MIB<sup>[4-6]</sup>库中。

目前, 以太网的物理拓扑发现算法大多数都是根据网桥的 MAC 地址转发表(FDB)来进行的。由于地址转发表中的记录具有动态性和不完整性, 这些算法在执行前必须在网络中注入附加的流量来保证完整性。因此, 这些算法具有一定的局限性。针对这个局限性, 本文提出一种基于 STP 协议的物理拓扑发现算法, 不需要注入附加的流量, 同时可以有效地发现备份链路和集线器、哑交换机等不支持 SNMP 的连接设备。

### 2 相关工作

贝尔实验室 Y.Breitbart 等<sup>[1,7]</sup>和卡奈基·梅隆大学 Bruce Lowekamp<sup>[8]</sup>与郑海等<sup>[9]</sup>分别提出各自基于地址转发表的物理拓扑发现算法。均在网络中注入附加的流量, 增加了网络的负荷, 从而使算法的可操作性受到影响。

### 3 系统模型和相关协议

#### 3.1 系统模型

可以将 IP 网络上的各个网络节点以及它们之间的连接关系组成一个无向图  $N$ , 各网络节点构成图  $N$  中的节点, 各网络设备之间的物理连接构成图  $N$  中的边, 拓扑发现算法就是找出拓扑图中的节点与边之间的关系。如图 1 所示, 无向图的节点对应各网络设备(路由器  $R_1, R_2, R_3$ , 交换机  $S_1, S_2, S_3, S_4, S_5$ ), 无向图的边  $(R_1, R_2), (R_2, S_1), (S_1, S_2) \dots$  对应网络设备的连接关系。

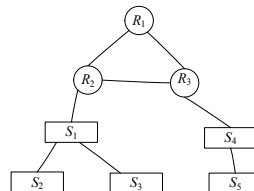


图 1 系统模型的例子

**基金项目:** 国家科技攻关计划基金资助项目(2004BA907A20); 吉林省科技发展计划基金资助重大项目(20040304)

**作者简介:** 张占国(1981—), 男, 硕士研究生, 主研方向: 计算机网络, 网络管理技术; 刘淑芬, 教授、博士生导师; 包 铁、张欣佳、王晓燕, 博士研究生; 张 婵, 硕士研究生

**收稿日期:** 2007-04-02 **E-mail:** zzg\_in\_jlu@163.com

**定义 1** 管理域AD<sup>[1,7]</sup>: 网络中拓扑发现的目标范围称为管理域。

**定义 2** 交换域SD<sup>[1,7]</sup>: 在管理域AD中仅有交换机节点以及交换机之间的连接关系所组成的连通分量。

**定义 3** 令SD中第*i*台交换机为*S<sub>i</sub>*, 令SD中第*i*台交换机的第*j*个端口为*S<sub>ij</sub>*。

**定义 4** 令对于任意端口*S<sub>ij</sub>*, *A<sub>ij</sub>*表示交换机*S<sub>i</sub>*的地址转发表中通过端口*S<sub>ij</sub>*收到的数据帧的源MAC地址集合。

**定义 5** 直接连接: 2个相连节点之间没有其他节点, 称这个连接为直接连接。

**引理** 端口*S<sub>ij</sub>*与*S<sub>kl</sub>*直接连接当且仅当*A<sub>ij</sub> ∩ A<sub>kl</sub> = ∅*, 且*A<sub>ij</sub> ∪ A<sub>kl</sub> = μ*。其中, ∅表示空集合; μ表示交换域中所有交换机的集合<sup>[1,7-8]</sup>。

### 3.2 生成树协议 STP

生成树协议STP是由Digital公司的Dr. Radia Perlman开发的用来确定每次交换信息所经过的转发路径。该协议的目的是实现网桥之间的冗余连接, 同时避免出现网络环路。生成树的主要思想是每个网桥定期发送网桥协议数据单元BPDU<sup>[3]</sup>, 向其他网桥表明自己的存在; 以太网中所有的网桥在收到其他网桥的BPDU帧后运行生成树算法, 以确定根网桥, 决定端口处于转发还是阻塞状态, 形成唯一的生成树, 从而形成拓扑图, 如图2所示。

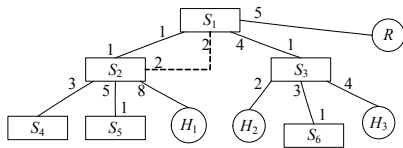


图2 执行STP协议后得到的生成树

BPDU中配置信息包含下列数据: (1)根网桥ID: 具有最小标识符的网桥作为生成树的根网桥, 根网桥总能在它所有端口上转发数据帧。网桥ID是网桥优先级和网桥MAC地址的组合。(2)发送BPDU的网桥ID。(3)网桥到网桥的开销: 从发送网桥到根网桥的最小费用。

表1给出了图2中各交换机在执行STP协议后的MIB库数据。其中, *S<sub>1</sub>*是根网桥, 以\*标记的端口为根端口。

表1 图2中各交换机在执行STP协议后的MIB库数据

交换机	端口号	指派网桥	指派端口	端口状态
<i>S<sub>1</sub></i>	1	<i>S<sub>1</sub></i>	<i>S<sub>11</sub></i>	转发
	2	<i>S<sub>1</sub></i>	<i>S<sub>12</sub></i>	转发
	3	<i>S<sub>1</sub></i>	<i>S<sub>13</sub></i>	转发
	4	<i>S<sub>1</sub></i>	<i>S<sub>14</sub></i>	转发
	5	<i>S<sub>1</sub></i>	<i>S<sub>15</sub></i>	转发
<i>S<sub>2</sub></i>	1*	<i>S<sub>1</sub></i>	<i>S<sub>11</sub></i>	转发
	2	<i>S<sub>1</sub></i>	<i>S<sub>12</sub></i>	阻塞
	3	<i>S<sub>2</sub></i>	<i>S<sub>23</sub></i>	转发
	4	<i>S<sub>2</sub></i>	<i>S<sub>24</sub></i>	阻塞
	5	<i>S<sub>2</sub></i>	<i>S<sub>25</sub></i>	转发
<i>S<sub>3</sub></i>	8	<i>S<sub>2</sub></i>	<i>S<sub>28</sub></i>	转发
	1*	<i>S<sub>1</sub></i>	<i>S<sub>14</sub></i>	转发
	2	<i>S<sub>3</sub></i>	<i>S<sub>32</sub></i>	阻塞
	3	<i>S<sub>3</sub></i>	<i>S<sub>33</sub></i>	转发
<i>S<sub>4</sub></i>	4	<i>S<sub>3</sub></i>	<i>S<sub>34</sub></i>	转发
	1*	<i>S<sub>2</sub></i>	<i>S<sub>23</sub></i>	转发
<i>S<sub>5</sub></i>	1*	<i>S<sub>2</sub></i>	<i>S<sub>24</sub></i>	转发
<i>S<sub>6</sub></i>	1*	<i>S<sub>3</sub></i>	<i>S<sub>33</sub></i>	转发

STP将端口置于转发状态的3条规则: (1)STP选择一个根网桥, STP将根网桥上所有接口都置于转发状态。(2)选择根端口, 每个非根网桥选择一个到达根网桥开销最小的端口作为根端口, 每一个网桥只能有一个根端口(除了根网桥, 根网桥没有根端口), 根端口处于转发状态。(3)选定指派网桥和指派端口。有多个网桥连接到一个以太网段时, 到达根网桥开销最小的网桥的端口处于转发状态。每个网段开销最小的网桥叫指派网桥, 指派网桥路径到网段的端口叫指派端口。指派端口也处于转发状态。

不能转发数据帧的阻塞端口, 包括没有连接物理链路的端口、连接了物理链路但没有工作(上电)的端口或连接了冗余备份链路的端口。

## 4 物理拓扑发现算法

### 4.1 算法的理论基础

算法的出发点在于, 以太网中每台交换机都在Bridge MIB中保存了交换域的生成树的一部分, 利用SNMP获取这些信息, 根据生成树协议的特征, 通过比较推导可以求出整个交换域的拓扑结构。

**定理** 存在直接连接关系的2个转发端口中有且仅有一个根端口<sup>[10]</sup>。

**证明** 采用反证法。

(1)假设存在直连关系的2个转发端口都是根端口。因为2个端口存在于同一物理链路上, 所以2个根端口属于一个网段, 根据根端口的定义, 根端口是负责本网段向根网桥转发数据帧的端口, 这样就存在一个网段通过2个不同的端口有到根网桥的2条路径, 这与生成树协议中2个设备仅有一条路径矛盾。所以存在直连关系的2个转发端口中最多只有一个根端口。

(2)假设存在直连关系的2个转发端口都不是根端口。因为这2个端口存在于同一物理链路上, 所以这2个端口属于一个网段内, 这个网段的数据帧都由这2个端口转发。而这个网段的数据由根端口转发到根网桥, 这与这2个端口都不是根端口矛盾。

由(1),(2)可知, 存在直接连接关系的2个转发端口中有且仅有一个根端口。

**推论 1** 2个非根端口之间若存在直接关系, 则该链路是备份链路, 不属于生成树的边, 即有一个端口的状态为阻塞。

**推论 2** 交换机除了根端口以外的所有其他转发端口的指派网桥都是本交换机。

**推论 3** 若交换机的阻塞端口的指派网桥不等于本交换机, 那么该端口一定用于连接备份链路。

### 4.2 算法使用的判定规则

根据上述生成树的定理以及推论, 在获取MIB中的信息以后, 可以根据下面的规则判定网络设备的连接关系。

**规则 1** 若交换机*S<sub>i</sub>*与*S<sub>k</sub>*满足: *S<sub>i</sub>*的根端口*S<sub>ij</sub>*的指派网桥是*S<sub>k</sub>*, 且*S<sub>k</sub>*的非根端口*S<sub>kl</sub>*的指派端口与*S<sub>ij</sub>*的指派端口相等, 那么*S<sub>ij</sub>*与*S<sub>kl</sub>*直连。

**规则 2** 主机*H*或路由器*R*与交换机*S<sub>i</sub>*的端口*S<sub>ij</sub>*直连, 当且仅当*S<sub>ij</sub>*不是级联端口(用于直连交换机的端口), 且*A<sub>ij</sub>*中包含主机*H*或者路由器*R*的MAC地址。

由于冗余备份链路不属于生成树中的边, 其链路两端的端口没有MAC地址转发记录, 即地址转发表为空。因此基于地址转发表的算法不能发现冗余备份链路。利用生成树端口表中的记录可以发现冗余(阻塞)链路。如图2中<sup>[10]</sup>, *S<sub>2</sub>*的非

根端口 $S_{22}$ 与 $S_1$ 的非根端口 $S_{12}$ 的直连就是冗余备份链路。

**规则 3** 若 $S_i$ 的非根端口 $S_{ij}$ 的指派网桥为 $S_k(k \neq i)$ , 且端口 $S_{ij}$ 的状态是阻塞的, 且 $S_k$ 的非根端口 $S_{ki}$ 的指派端口与 $S_{ij}$ 的指派端口相等, 那么 $S_{ij}$ 与 $S_{ki}$ 直连, 且该链路是备份链路。

许多网络中, 不是所有的设备都与交换机直接连接, 例如, 主机可能通过集线器 Hub 与交换机连接, 而且交换机之间也可能通过不支持 SNMP 的设备相连。而这些设备是不可管理的, 所以这时就无法只靠规则 1 来发现交换机之间的最小生成树。

完整的生成树结构如图 3 所示, 只使用规则 1 得到的是未发现 Hub 的不完整的生成树结构, 如图 4 所示。因此, 必须对规则 1 进行改进。

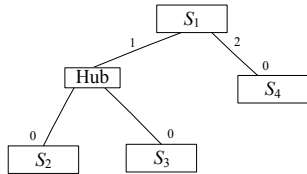


图 3 完整的生成树结构

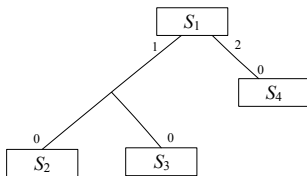


图 4 使用规则 1 未发现 Hub 的生成树结构

**规则 4** 满足规则 1 的 2 个端口 $S_{ij}$ 与 $S_{kl}$ , 不妨假定 $S_{ij}$ 是根端口,  $S_{kl}$ 是非根端口, 若存在其他端口 $S_{xy}(x \neq i)$ 与 $S_{kl}$ 满足规则 1, 那么 $S_i, S_k$ 与 $S_x$ 之间必定存在集线器 Hub 或哑交换机等不支持 SNMP 的设备。

**规则 5** 满足规则 1 的 2 个端口 $S_{ij}$ 与 $S_{kl}$ , 若 $A_{ij} \cap A_{kl} \neq \emptyset$ , 那么 $S_i$ 与 $S_j$ 之间必定存在集线器 Hub 或哑交换机等不支持 SNMP 的设备。

#### 4.3 算法具体描述

- (1) 获取子网内所有的交换机列表。
- (2) 获得每个交换机的生成树端口状态信息表。
- (3) 根据生成树协议确定所搜索的交换域的根网桥 $S_0$ , 将根网桥 $S_0$ 加入到待检测队列(先进先出队列)中。
- (4) 从待检测队列中取出一个交换机 $S_r$ 。
- (5) 从所有交换机列表中查找端口(无论根端口还是非根端口)的指派网桥为 $S_r$ 的交换机 $S_i$ , 将查找到的这些交换机组成一个集合 $\lambda = \{S_i \mid S_i \text{ 的端口 } S_{ij} \text{ 的指派网桥为 } S_r\}$ 。
- (6) 查看集合 $\lambda$ 中是否存在交换机 $S_i$ , 满足: 交换机 $S_i$ 的端口 $S_{ij}$ 的指派端口与交换机 $S_r$ 的非根端口 $S_{ri}$ 的指派端口相同, 并且 $A_{ij} \cap A_{ri} \neq \emptyset$ 。如果存在, 根据规则 5, 这样的交换机 $S_i$ 通过集线器或者哑交换机等不支持 SNMP 的设备与交换机 $S_r$ 相连。
- (7) 查看集合 $\lambda$ 中是否存在多个交换机的端口的指派端口是 $S_r$ 的同一个非根端口。如果存在, 根据规则 4, 这些交换机通过集线器或者哑交换机等不支持 SNMP 的设备与交换机 $S_r$ 相连。
- (8) 集合 $\lambda$ 中不满足条件(6)、条件(7)的的交换机与交换机

$S_r$ 直连。

(9) 对于集合 $\lambda$ 中的每一个交换机 $S_i$ , 如果交换机 $S_i$ 的根端口 $S_{ij}$ 的指派网桥为 $S_r$ , 那么,  $S_i$ 与 $S_r$ 之间的连接为生成树的边, 并将 $S_i$ 加入有待检测队列中; 否则, 交换机 $S_i$ 的非根端口 $S_{ij}$ 的指派网桥为 $S_r$ , 根据规则 3,  $S_i$ 与 $S_r$ 之间的连接为冗余备份链路。

(10) 根据条件(6)~条件(9)的检测结果, 连接交换机 $S_r$ 与集合 $\lambda$ 的每个交换机。

(11) 对于交换机 $S_r$ 不用于连接交换机的任意端口 $S_{ri}$ , 获得它的地址转发表 $A_{ri}$ , 根据规则 2, 其中的主机或路由器的 MAC 所代表的设备与路由器 $S_r$ 直连。如果有多个主机或者路由的 MAC 地址属于 $A_{ri}$ , 那么这些主机或路由器通过集线器或者哑交换机等不支持 SNMP 的设备与交换机 $S_r$ 相连。

(12) 判断待检测队列是否为空, 如果不为空, 转(4)。

## 5 结束语

本文提出了一种基于网桥生成树协议 STP 的算法, 利用 SNMP 获得各个交换机 MIB 库中的交换机端口状态信息, 根据生成树协议推导出网络的物理拓扑。该算法克服了基于 MAC 地址转发表算法的不足, 而且还可以发现交换机之间的冗余备份链路和集线器、哑交换机等不支持 SNMP 的设备。

在接入网中, 经常采用 VLAN 技术, 以此抑制广播数据。虽然各厂商均支持 IEEE802.1Q VLAN 标准, 但是由于各个厂家的实现方法不同, 而且在 MIB-II 中没有规定一个通用信息库来管理这些信息, 因此, 如果存在 VLAN, 物理网络拓扑就无法发现。算法的下一步工作是要解决 VLAN 条件下的物理拓扑发现。

## 参考文献

- [1] Breitbart Y, Garofalakis M. Topology Discovery in Heterogeneous IP Networks[C]//Proc. of INFOCOM'00. Tel Aviv, Israel: [s. n.], 2000.
- [2] Tanenbaum A S. Computer Networks[M]. 3rd ed. Englewood Cliffs: Prentice Hall PTR, 1996.
- [3] IEEE 802.1D-1998 IEEE Standard for Information Technology—Telecommunications and Information Exchange Between Systems—Local and Metropolitan Area Networks—Common Specifications Part 3: Media Access Control(MAC) Bridges[S]. 1998.
- [4] Simple Network Management Protocol[S]. RFC 1157, 1990.
- [5] Rijssinghani A, McCloghrie K. Definitions of Managed Object for Bridges[S]. RFC 1493, 1993.
- [6] Cloghrie M. Management Information Base for Network—Management of TCP/IP-based Internet: MIB-[S]. RFC 1213, 1991.
- [7] Breitbart Y, Garofalakis M, Ja B. Topology Discovery in Heterogeneous IP Networks[J]. IEEE/ACM Trans. on the NetInventory System Networking, 2004, 12(3): 401-404.
- [8] Lowekamp B, David R, Gross T R, et al. Topology Discovery for Large Ethernet Networks[M]. New York: ACM Press, 2000: 237-248.
- [9] 郑海, 张国清. 物理网络拓扑发现算法的研究[J]. 计算机研究与发展, 2002, 39(3): 264-268.
- [10] 李晓红, 张大方. 一种以太网拓扑发现算法[J]. 同济大学学报, 2002, 30(10): 1173-1177.