

文章编号:1001-9081(2005)11-2499-03

IEEE 三种生成树技术的分析与比较

李延冰,马 跃,王 博,万小强

(北京邮电大学 计算机科学与技术学院,北京 100876)

摘 要:简要介绍了 IEEE 的生成树、快速生成树和多生成树协议的工作原理。分别从收敛时间、拓扑变化方式、带宽利用率和资源占用量等几个方面对这三种生成树协议进行了分析与比较。探讨了每种生成树协议的适用环境,为工程应用中生成树技术的选择提供了参考。

关键词:生成树协议;快速生成树协议;多生成树协议

中图分类号: TP393 **文献标识码:** A

Analysis and Comparison of IEEE STPs

LI Yan-bing, MA Yue, WANG Bo, WAN Xiao-qiang

(College of Computer Science and Technology, Beijing University of Posts and Telecommunications, Beijing 100876, China)

Abstract: The basic principles of STP, RSTP and MSTP were introduced. Some analyses and comparisons of the three spanning tree protocols were given in the aspects of convergent time, topology change, available bandwidth and resources taken up. The appropriate environment of each spanning tree protocol was discussed. It can be a reference to select spanning tree protocols in implementing a network project.

Key words: STP (Spanning Tree Protocol); RSTP (Rapid Spanning Tree Protocol); MSTP (Multiple Spanning Tree Protocol)

0 引言

生成树技术是交换网络中一项关键技术。它通过有选择地阻塞端口来消除网络环回,同时也可以实现链路冗余的功能。IEEE 的生成树技术经历了 IEEE 802.1D 生成树协议、IEEE 802.1W 快速生成树协议和 IEEE 802.1S 多生成树协议,性能不断得到改进。

IEEE 802.1D 生成树协议是 IEEE 最早的生成树技术,该生成树技术在早期的交换网络中起了很大作用,达到了消除网络环回、实现冗余链路的目的。但随着交换网络规模的增长,以及用户对网络反应速度要求的提高,这种生成树技术已不能适应大型交换网络的需求。IEEE 802.1W 快速生成树协议在 IEEE 802.1D 生成树协议的基础上进行了改进,大大缩短了收敛时间,同时兼容 IEEE 802.1D 生成树协议。这两种生成树技术都属于单生成树技术,无法解决由于网络规模过大而引起的收敛时间增大的问题,同时会对采用 VLAN 技术的交换网络中的带宽利用率和连通性产生影响。于是 IEEE 又设计了 IEEE 802.1S 多生成树协议,较好地解决了上述问题,同时兼容前两种生成树协议。

笔者结合在交换机上实现这三种生成树协议的经验,对这三种生成树协议进行分析、比较。

1 生成树协议的工作原理

这三种生成树技术都是基于 IEEE 802.1D 生成树协议的基本思想,这一点使得 IEEE 802.1W 快速生成树协议和 IEEE 802.1S 多生成树协议能够很好地兼容早期版本的生成树协议。

1.1 IEEE 802.1D 生成树协议

IEEE 802.1D 生成树协议定义了两种桥协议数据单元(BPDU):配置桥协议数据单元(CONFIG BPDU)和拓扑变化通告桥协议数据单元(TCN BPDU)。它还定义了三种端口角色:根端口、指定端口和替代端口;五种端口状态:阻塞状态、监听状态、学习状态、转发状态和不可用状态。

网桥之间通过交换配置 BPDU 来选举根网桥和确定端口角色;通过传递拓扑变化通告 BPDU 来报告网络拓扑变化。

配置 BPDU 中包含根网桥 ID、根路径开销、指定桥 ID 和指定端口 ID。网桥 ID 由网桥优先级和网桥 MAC 地址组成,端口 ID 由端口优先级和端口号组成,根路径开销是指从本网桥到根网桥所经各端口的端口开销之和。

在选举根网桥时,起初每个网桥都认为是根网桥。如果一个网桥收到的配置 BPDU 中的根网桥 ID 优于自己保存的根网桥 ID(所谓优于是指网桥优先级小或在网桥优先级相等的情况下网桥 MAC 地址小),则该网桥就替换原来的根网桥 ID。通过这种比较,最终具有最优网桥 ID 的网桥被选作根网桥。

选出根网桥后,每个非根网桥都要选出一个根端口,这是通过比较非根网桥的每个端口到根网桥的根路径开销来实现的。具有最小根路径开销的端口被选作根端口;如果两个端口的根路径开销相同,则具有最优端口 ID 的端口被选作根端口。

接着要为每个网段确定一个指定网桥和指定端口。当有多个网桥连接到同一个网段时,具有最小根路径开销的网桥被选作该网段的指定网桥,连接这个网段的网桥端口称为指定端口。当一个指定网桥有多个端口连接到同一个网段,则

收稿日期:2005-06-30;修订日期:2005-08-16

作者简介:李延冰(1980-),男,河北邯郸人,硕士研究生,主要研究方向:计算机网络与通信; 马跃(1960-),男,河南南阳人,副教授,硕士,主要研究方向:计算机网络与通信; 王博(1978-),男,黑龙江哈尔滨人,助理研究员,硕士,主要研究方向:计算机网络与通信; 万小强(1981-),男,湖北公安人,硕士研究生,主要研究方向:计算机网络与通信。

具有最优端口 ID 的端口被选作指定端口。

除了根端口和指定端口外,其余的可用端口被称作替代端口。

端口角色确定之后,根端口和指定端口经过监听状态、学习状态最终进入转发状态;替代端口则直接进入阻塞状态。

当非根网桥检测到网络拓扑变化时,它会向其根端口所连接的网段重复发送拓扑变化通告 BPDU,直到它收到这个网段的指定网桥发送过来的拓扑变化确认(设置配置 BPDU 中的“拓扑变化确认 FLAG”)。然后这个网段的指定网桥重复相同的过程,当根网桥收到这个拓扑变化通告 BPDU 后或者根网桥自己检测到网络拓扑发生了变化,它会在其发送的配置 BPDU 中设置“拓扑变化 FLAG”,每个非根网桥收到这样的配置 BPDU 后会用转发延迟时间来使过滤数据库中的动态地址项快速失效。

1.2 IEEE 802.1W 快速生成树协议

IEEE 802.1W 快速生成树协议在 IEEE 802.1D 生成树协议的配置 BPDU 的基础上定义了快速生成树 BPDU (RST BPDU)。

IEEE 802.1W 快速生成树协议重新定义了端口角色。除了根端口和指定端口外,将 IEEE 802.1D 生成树协议中的替代端口细分为替代端口和备份端口。这两种端口的区别是:替代端口是根端口的冗余端口而备份端口是指定端口的冗余端口。

IEEE 802.1W 快速生成树协议将端口状态定义为丢弃、学习和转发。

IEEE 802.1W 快速生成树协议还定义了边缘端口和点对点链路。边缘端口就是只连了一台主机的端口;点对点链路是指只连接了两个网桥的共享链路。

为了加快收敛速度,IEEE 802.1W 快速生成树协议增加了如下机制:

1) 根端口满足如下条件就可以快速进入转发状态:a) 该端口最近不是备份端口;b) 该网桥除此端口以外的其他端口最近都不是根端口。

2) 连接了点点对点链路的指定端口可以通过和下游网桥握手而快速进入转发状态。握手过程如下:上游网桥欲进入转发状态的指定端口向下游网桥发送的快速生成树 BPDU 中设置了“PROPOSAL FLAG”;下游网桥收到这种快速生成树 BPDU 后,等自己的所有替代端口和备份端口进入阻塞状态、所有的边缘端口和指定端口进入转发状态之后给上游网桥回送一个设置了“AGREEMENT FLAG”的快速生成树 BPDU;上游网桥的指定端口收到这个快速生成树 BPDU 之后直接进入转发状态。

3) 边缘端口只连有一台主机,不会引起网络环回,因此当边缘端口是指定端口时可以立即进入转发状态。

IEEE 802.1W 快速生成树协议只有当非边缘端口进入转发状态时才导致拓扑变化。当网桥得知拓扑变化时先向该网桥所有非边缘端口的指定端口和根端口发送设置了“TOPOLOGY CHANGE FLAG”的快速生成树 BPDU,然后再清除这些端口过滤数据库中的动态地址项,当其他网桥收到设置了“TOPOLOGY CHANGE FLAG”的快速生成树 BPDU 后也做同样动作。

IEEE 802.1W 快速生成树协议可以很好地兼容 IEEE 802.1D 生成树协议,它可以识别 IEEE 802.1D 生成树协议的 BPDU,当运行 IEEE 802.1W 快速生成树协议的网桥的某个

端口收到了 IEEE 802.1D 生成树协议的 BPDU,该端口就立即进入 IEEE 802.1D 生成树协议兼容模式。

1.3 IEEE 802.1S 多生成树协议

IEEE 802.1S 多生成树协议在 IEEE 802.1W 快速生成树协议的基础上定义了多生成树 BPDU (MST BPDU)。

IEEE 802.1S 多生成树协议定义了多生成树域(MST REGION)。多生成树域由一个或多个运行了 IEEE 802.1S 多生成树协议的网桥以及它们之间的网段组成,域内的每个网桥都有相同的多生成树配置 ID、相同的多生成树实例(MSTI)和相同的 VLAN 到多生成树实例的映射。

IEEE 802.1S 多生成树协议定义了内部生成树(IST)、公共生成树(CST)、公共和内部生成树(CIST)以及多生成树实例(MSTI)。它们之间的关系是:内部生成树是多生成树域内的一棵单生成树;公共生成树是连接所有多生成树域的单生成树;内部生成树和公共生成树共同组成公共和内部生成树,它是连接整个交换网络中所有网桥的一棵单生成树;一个多生成树域内可以有多个多生成树实例,每个多生成树实例和一个或多个 VLAN 相对应,彼此之间相互独立。

IEEE 802.1S 多生成树协议定义的端口角色有:根端口,指定端口,主端口,替代端口和备份端口。除主端口外,其他端口角色和 IEEE 802.1W 快速生成树协议定义是一样的。在 IEEE 802.1S 多生成树协议中,如果将整个域看成是一个网桥,那么这个网桥的根端口即是这个域内所有多生成树实例的主端口。

IEEE 802.1S 多生成树协议定义的端口状态和 IEEE 802.1W 快速生成树协议定义的端口状态是一致的。

网桥之间通过交换多生成树 BPDU,选举具有最高优先级的网桥作为公共和内部生成树根(CIST root),每个域内选举到公共和内部生成树根路径最短的网桥作为该域的公共和内部生成树域根(CIST regional root)。最终每个域内生成一棵以公共和内部生成树域根为根的内部生成树;假如把每个域看成是一个网桥,每个域之间生成一棵以公共和内部生成树根为根的公共生成树。公共生成树和内部生成树共同组成公共和内部生成树。每个域内的各多生成树实例单独进行计算,不同的多生成树实例生成的网络拓扑结构可能不同。

IEEE 802.1S 多生成树协议中根端口、指定端口和主端口向转发状态快速迁移的条件和 IEEE 802.1W 快速生成树协议基本相同。

IEEE 802.1S 多生成树协议处理网络拓扑变化时的方式和 IEEE 802.1W 快速生成树协议也基本相同,但进行了一些扩展。公共和内部生成树和每个多生成树实例的拓扑变化信息分别进行传递。多生成树实例中网络拓扑发生变化时不影响公共和内部生成树和其他多生成树实例;公共生成树网络拓扑发生变化时要影响所有生成树;内部生成树网络拓扑发生变化时只影响公共生成树。

IEEE 802.1S 多生成树协议能够很好地兼容 IEEE 802.1D 生成树协议和 IEEE 802.1W 快速生成树协议,它把运行 IEEE 802.1D 生成树协议和 IEEE 802.1W 快速生成树协议的网桥看成是一个单独的域。

2 三种生成树协议的分析与比较

在了解了这三种生成树协议的基本工作原理后,下面分别从收敛时间、拓扑变化方式、带宽利用率和资源占用量这几

方面进行分析与比较。

2.1 收敛时间

当网络拓扑发生变化时,IEEE 802.1D 生成树协议并不能立即察觉,而是要等待 MAX AGE 时间后即所保存的配置消息超时才能察觉,新的配置消息要经过一定的时延 (FORWARD DELAY) 才能传播到整个网络。如果将旧拓扑中处于阻塞状态而在新拓扑中应处于转发状态的端口立即置为转发状态,而旧拓扑中处于转发状态的端口有可能还没有发现自己在新的拓扑中应处于阻塞状态,就会引起网络临时环路。为了防止网络临时环路,端口在从监听状态进入转发状态前要经历一个学习状态。这样当网络拓扑发生变化时,新的网络拓扑要经过 $\text{MAX AGE} + 2 \times \text{FORWARD DELAY}$ 时间才能进入稳定状态。

IEEE 802.1W 快速生成树协议和 IEEE 802.1S 多生成树协议采用了快速收敛的机制。当网络拓扑发生变化时,它们最多等待 $3 \times \text{HELLO TIME}$ 时间就认为所保存的配置消息超时。对于符合 1.2 中所述的条件的根端口和指定端口可以快速进入转发状态,但对于其他端口仍需等待 $2 \times \text{FORWARD DELAY}$ 才可以进入转发状态。因此对于全部使用点对点链路连接起来的交换网络当网络拓扑发生变化时可以快速收敛于稳定状态。

2.2 拓扑变化方式

IEEE 802.1D 生成树协议在网络拓扑发生变化时,首先要将拓扑变化信息一级一级传递给根网桥,然后由根网桥将拓扑变化信息扩散到整个网络中。每个网桥得知网络拓扑发生变化后要使过滤数据库中所有的动态地址项快速失效。

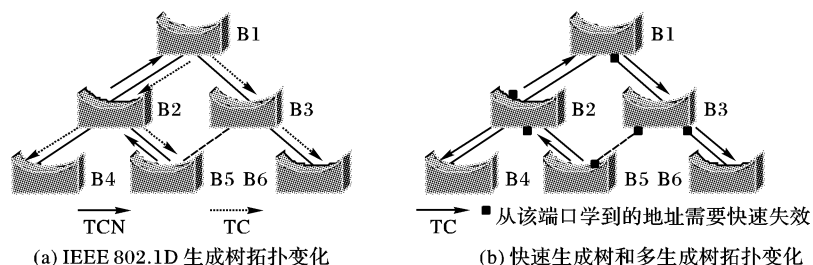


图 1 生成树拓扑变化方式

IEEE 802.1W 快速生成树协议对此作出改进。拓扑变化信息直接由检测到拓扑变化的网桥扩散到全网,而每个得知网络拓扑发生变化的网桥只需使除收到拓扑变化信息的端口之外的所有非边缘端口的指定端口在过滤数据库中所有的动态地址项快速失效,这样避免了网络中出现大量找不到目的的数据帧,但在具体实现当中很多型号的交流芯片并不支持这项特性。

IEEE 802.1S 多生成树协议在支持 IEEE 802.1W 快速生成树协议的这项特性的基础上,把每个多生成树实例中网络拓扑的变化限制在其所在域中,而不扩散至全网,甚至不影响同域中其他多生成树实例。当然这一点也需要交换芯片的支持。

如图 1 所示,IEEE 802.1D 生成树协议在发生拓扑变化时,每个网桥需要将过滤数据库中的所有动态地址项快速失效,而改进后的快速生成树和多生成树协议只需使从部分特定端口学习到的地址项快速失效。

2.3 带宽利用率

现代的大型交换网络中大都采用 IEEE 802.1Q VLAN 技

术,每个数据帧都被加上 VLAN 标签。在连接网桥与网桥的干道链路上可以同时承载不同的 VLAN 的流量。

IEEE 802.1D 生成树协议和 IEEE 802.1W 快速生成树协议都属于单生成树协议,一个端口被阻塞后就不承载任何流量,造成带宽的浪费。当 VLAN 的网络拓扑结构不对称时,甚至会影响某些 VLAN 的连通性。

IEEE 802.1S 多生成树协议能够阻塞一个端口上的某些 VLAN 的流量,而不影响其他 VLAN 的流量。这样可以把网络拓扑结构相同的几个 VLAN 映射到同一个多生成树实例上,既节约了带宽又不会影响连通性,这一特性大部分交换芯片都能支持。

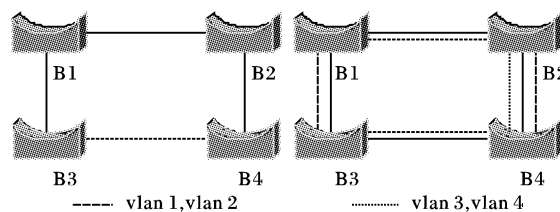


图 2 生成树拓扑结构

如图 2 所示,单生成树将 B3 与 B4 之间的链路阻塞后,该链路不再承载任何流量。而在多生成树拓扑中,vlan1 和 vlan2 所属的生成树实例将 B3 和 B4 之间的链路阻塞,但单该链路还可以承载 vlan3 和 vlan4 的流量。

2.4 资源占用量

IEEE 802.1D 生成树协议和 IEEE 802.1W 快速生成树协议在每个网桥中只维护一棵树,而 IEEE 802.1S 多生成树协议在每个网桥中要维护多棵树,协议规定最多可以维护 64 棵树。

这样,IEEE 802.1S 多生成树协议无论在内存占用量还是在 CPU 占用率上都大大高于单生成树。

3 结语

通过比较分析这三种生成树技术不难发现生成树技术在不断地发展与完善。随着用户对网络反映速度的要求的提高改进了收敛速度;随着 IEEE 802.1Q VLAN 技术在交换网络中的大规模应用,又对 VLAN 技术进行了融合。可以说目前 IEEE 802.1S 多生成树协议在各方面都表现很出色,得到了绝大多数厂家的支持,是目前的主流生成树技术。

但另一方面,我们认为简单灵巧的 IEEE 802.1D 生成树协议和 IEEE 802.1W 快速生成树协议在小型交换网络中还有很好的发展前景,没有必要采用资源占用量较高并且复杂的多生成树协议,而对于像城域网这样大规模的交换网络,IEEE 802.1S 多生成树协议是比较理想的技术。

参考文献:

- [1] ANSI/IEEE Std 802.1D, 1998 Edition, Standard for local and metropolitan area networks: Media Access Control (MAC) Bridges[S], 1998.
- [2] IEEE Std 802.1W-2001, Standard for local and metropolitan area networks: Media Access Control (MAC) Bridges-Amendment 2: Rapid Reconfiguration[S], 2001.
- [3] IEEE Std 802.1S-2002, Standard for local and metropolitan area networks: Virtual Bridged Local Area Networks-Amendment3: Multiple Spanning Trees[S], 2002.