

以太网快速环网保护算法的设计和实现

李 富,程子敬,李 周

(北京卫星信息工程研究所,北京 100086)

摘 要: 在 STP 基础上提出了一种快速的环路保护算法。该算法能够提供毫秒级的环路消除和故障恢复能力且开销小。最后介绍了该算法在硬件上的实现。

关键词: 冗余链路;环路保护;STP

中图分类号: TN915

文献标识码: A

文章编号: 1674-7720(2012)13-0058-03

Research and implementation on a rapid ring protection algorithm

Li Fu, Cheng Zijing, Li Zhou

(Beijing Institute of Satellite Information Engineering, Beijing 100086, China)

Abstract: This paper presents a rapid ring protection algorithm based on STP which can provide capabilities of ms loop elimination, recovery and low overhead. Finally, the paper describes the hardware implementation of this algorithm.

Key words: redundant link; ring protection; STP

交换式以太网已广泛应用到工厂、煤矿、电力等场所。为了提高网络可靠性,网络需要具备冗余链路。交换网络环路为交换网络提供冗余链路,消除了由于单点故障而引起的网络中断,但同时形成数据环路,会引发二层交换网络的广播风暴,导致网络瘫痪^[1]。

为了解决冗余链路引起的问题,环路保护技术应运而生。生成树协议 STP (IEEE 802.1D, Spanning Tree Protocol) 通过阻塞冗余端口进行链路备份,使得网络中断后可在 30 s~60 s 内恢复。为了缩短网络自愈时间,IEEE 又提出了与 STP 兼容的快速生成树协议 RSTP (Rapid Spanning Tree Protocol),其收敛时间为秒级^[2]。

本文在分析 STP 这种主流的环路保护技术的基础上,提出了快速环路保护技术(RRP)。针对环型拓扑,RRP 克服了传统 STP 自愈时间长的缺点,可达到毫秒级的网络自愈时间,而且复杂度低,便于实现。

1 生成树工作原理及其缺点

STP 是将一个存在物理环路的交换网络变成一个没有环路的逻辑树型网络,实现在逻辑上裁剪冗余环路,同时在物理上实现链路备份和路径最优化。STP 通过 Config-BPDU 数据包来构造树型网络,通过 Tcn-BPDU 来通告网络拓扑变化。STP 算法步骤如下:

(1) 选举根节点。拥有最小标识的节点将成为根节

点。选举过程开始时,所有节点都声明自己是根。当节点的一个端口收到高优先级的 Config-BPDU 时,就在该端口保存这些信息,同时向所有端口更新并传播信息。如果收到比自己低优先级的 Config-BPDU,节点就丢弃该信息。根节点的所有端口置为转发状态。

(2) 确定根端口。对每个非根节点,选择一个到根节点路径最短的端口作为此节点的根端口。所有根端口置为转发状态。

(3) 确定指定端口。多个节点连接到同一网段时,代价最小的节点被称为指定节点,取指定节点在此网段上的一个端口作为指定端口。指定端口通过逐个考查与端口相连的网段来确定,选择指定端口的依据首先是路径成本,路径成本低的端口将成为指定端口。所有指定端口置为转发状态。

(4) 等待拓扑变化。若根节点故障,其余各节点的每个端口都收不到 Config-BPDU 数据包,则等待 Config-BPDU 数据包的计时器都超时,都认为自己是根节点,开始重新构建树型网络,重复步骤(1)~(3)。

通过对 STP 工作原理的分析,找到 STP 自愈时间长的几个原因:

(1) 每个节点端口角色确定复杂。节点有根端口、指定端口和阻塞端口 3 种角色。各节点的端口不停地发送

网络与通信 Network and Communication

和接收 Config-BPDU。每个端口根据收到的 BPDU 数据包不断更新端口配置信息,计算出根端口、指定端口和阻塞端口。根端口和指定端口还要经过 2 个 Forward Delay Time 才能进入转发状态。

(2) 根节点没有主动的故障侦测能力,这导致 STP 对拓扑结构的改变响应缓慢。拓扑变化时,发现拓扑变化的节点向根节点方向发送 Tcn-BPDU, 通告过程中存在多次应答,等到根节点收到 Tcn-BPDU 后发送携带拓扑改变标志位的 Config-BPDU,通知其他节点刷新 MAC 地址表,确立新路径。新选出的根端口和指定端口也要经过 2 个 Forward Delay Time 才能进入转发状态。

2 RRP 算法设计及实现

交换机端口对数据的处理无非是丢弃或转发,因此可以将交换机端口状态分为阻塞和转发两种。根交换节点是网络的逻辑中心而非物理中心,为了提高拓扑改变的反应速度,根交换节点需要自发故障检测,而不依赖其他交换节点的故障通告。

2.1 快速环路保护算法 (RRP)

(1) 选择根节点。根节点是环网状态主动检测机制的发起者,也是网络拓扑发生改变后执行操作的决策者。初始时,各交换节点在 hello-timer 的作用下定时从两个级联端口发送环路健康检测报文,即 hello 报文,交换节点收到报文后,进行优先级判断,选择出根交换机, ID 越小,优先级越高。如果收到报文的 ID 比自己的 ID 小,表明自己为传输节点,从另一端口转发收到的报文,自己不再发送 hello 报文;如果收到报文的 ID 比自己的 ID 大,就丢弃此报文;如果收到自己的报文,说明自己为根节点,表明有环,阻塞自己一个端口,定时发送 hello 报文。

(2) 根节点环路检测。根节点定时从两个级联端口发送 hello 报文来检测环路健康状况。若根节点能收到自己的 hello 报文,则表明环路是完整的;如果在 wait-timer 内收不到 hello 报文,就认为环网发生链路故障。

(3) 故障发现。若设备或链路发生故障,与故障链路相连的端口置为阻塞状态。根节点收不到自己发出的 hello 报文,wait-timer 超时,表明环路不完整,出现故障,根节点把之前阻塞的端口打开,同时发送 flush 报文,通知其他传输节点更新地址转发表。传输节点收到根节点的 flush 报文后,刷新地址转发表,重新进行地址学习。

(4) 故障恢复。故障消除后,根节点重新收到自己发出的 hello 报文,表明环路存在,根节点阻塞自己一个级联端口。由于拓扑发生变化,根节点发出 flush 报文,通知其他交换节点刷新地址表,与恢复链路相连的两端口置为转发状态。

(5) 根节点失效检测。若根节点发生故障,传输节点收不到 hello 报文,wait-timer 超时,各传输节点均阻塞端口,开始发送 hello 报文,重新选出一个根节点。

图 1 描述了 RRP 算法的实现过程。图 1(a)中各节

点向两个方向发送 hello 报文。图 1(b)中由于 B1 的 ID 最小,被选为根节点,(B1,P1)阻塞,环路消除。图 1(c)中节点 B3 和 B4 间发生故障,(B3,P1)和 (B4,P0)阻塞,根节点收不到 hello 报文,wait-timer 超时,(B1,P1)置为转发,并向两方向发送 flush 报文。图 1(d)中各节点收到 flush 报文后,刷新地址表,B1 定时发送 hello 报文。图 1(e)中节点 B3 和 B4 间故障修复,(B3,P1)和 (B4,P0)仍保持阻塞,B1 收到自己的 hello 报文,意识到环路的存在,重新阻塞端口 (B1,P1),并向两方向发送 flush 报文。图 1(f)中节点 B3 和 B4 收到根节点的 flush 报文后,把 (B3,P1)和 (B4,P0)置为转发态,根节点定时发送 hello 报文。网络拓扑重新收敛。

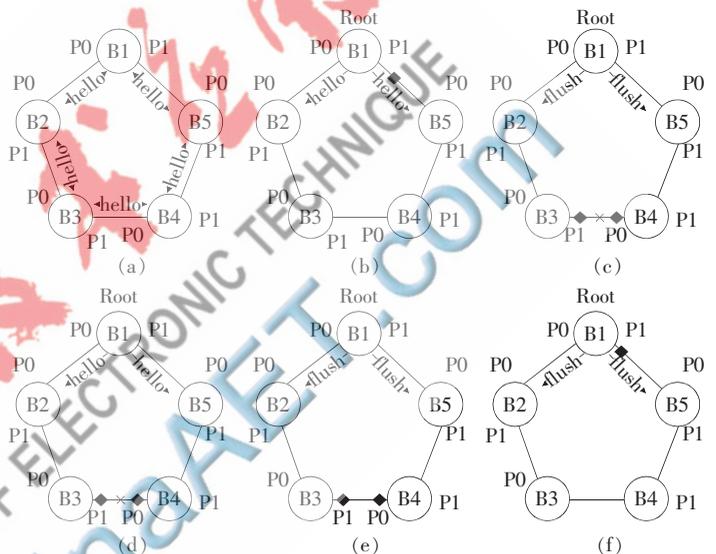


图 1 RRP 算法描述 (◆表示端口阻塞,×表示故障)

2.2 RRP 算法的硬件实现

本文基于稳定拓扑的以太环网保护切换方案定义了 2 种端口状态、9 种节点状态和 8 类事件,使用状态机可以灵活实现状态的转移和事件的处理。

(1) 端口状态

PS0: 阻塞,端口只处理协议控制报文,不接收或转发数据,不进行地址学习;

PS1: 转发,端口接收并转发数据,处理协议控制报文,开始地址学习。

(2) 节点状态

S0: IDLE,环上端口阻塞;

S1: 根节点,环上端口一个转发,一个阻塞,未成环状态,发送 flush 报文;

S2: 根节点,环上端口一个转发,一个阻塞,成环状态,发送 flush 报文;

S3: 非根节点,环上端口转发;

S4: 根节点,环上端口一个转发,一个阻塞,未成环状态;

S5: 根节点,环上端口一个转发,一个阻塞,成环状态;

