

# 基于嵌入式 Linux 的 STP 模块的设计与应用

陈麓<sup>1</sup>, 鹿瑾<sup>1</sup>, 徐曼<sup>2</sup>

(1. 中国电信股份有限公司 徐州分公司 网络维护中心, 江苏 徐州 221000;

2. 中国矿业大学 信息与电气工程学院, 江苏 徐州 221008)

**摘要:** 重点描述了基于嵌入式 Linux 的 STP 模块的开发过程, 深入分析了 Linux 内核的 STP 协议以及 MGMT 帧在内核的处理流程, 并针对内核中 STP 协议存在的问题设计了适用于交换机的 STP 模块, 以及 STP 模块在千兆以太网交换机软件系统中的应用。

**关键词:** STP 协议; 嵌入式 Linux; STA 算法

中图分类号: TP316

文献标识码: A

文章编号: 1674-7720(2010)12-0062-02

## Design and application of STP module based on embedded Linux

CHEN Lu<sup>1</sup>, LU Jin<sup>1</sup>, XU Man<sup>2</sup>

(1. Xuzhou Branch of China Telecom, Xuzhou 221000, China;

2. School of Information and Electronic Engineering, CUMT, Xuzhou 221008, China)

**Abstract:** The essay has mainly described the process of designing STP module based on embedded Linux. By way of analysing STP and the process of MGMT frame in Linux kernel it has designed the STP module of switch to solve the problem of STP in kernel. Furthermore, it has introduced the application of STP module in Gigabit Ethernet network switch.

**Key words:** spanning tree protocol; embedded Linux; spanning tree algorithm

Linux 2.4x 内核中与生成树协议相关的程序由透明网桥模块完成, Linux 的透明网桥协议是以 Linux 内核为数据交换中心, 而在实际工作中, 网络数据的交换与转发都由交换芯片 Marvell 88E6095 完成。因此在不改变 STA 算法的前提下, 移植 Linux 内核透明网桥中与 STP 协议相关部分, 能实现应用于千兆以太网交换机的 STP 模块。

### 1 STP 协议与 STA 算法

生成树协议 STP (Spanning Tree Protocol) 应用于环路网络, 通过一定的算法实现路径冗余, 同时将环路网络修剪成无环路的树型网络, 从而避免报文在环路网络中的增生和无限循环<sup>[1]</sup>。生成树协议的基本原理是通过在交换机之间传递一种特殊的协议报文确定网络的拓扑结构。

STP 协议采用的算法称为生成树算法 STA<sup>[2]</sup> (Spanning Tree Algorithm), 生成树算法很复杂, 但是其过程可以归纳为选择根桥、选项根端口、选择指定端口三个部分。

### 2 基于 Linux 的 STP 模块的设计

#### 2.1 嵌入式 Linux 的 STP 模块整体架构

STP 协议中用于网络拓扑管理的数据流称为 BPDU

(Bridge Protocol Data Unit), 网卡芯片则将 BPDU 封装成 MGMT 帧, MGMT 帧与普通数据包的最大区别在于可以自由进出已经阻塞的端口, 方便网络拓扑的管理。

在 Linux 2.4x 内核中, 驱动程序并不支持类似于 MGMT 帧的多播协议数据流, 为使 MGMT 帧自由进出系统, 需要两个条件: (1) 交换芯片 Marvell 88E6095 必须支持 IEEE 802.1D 协议, 且工作在该模式下; (2) 驱动程序需要开启对多播数据的支持, 即开启 MULTICAST 关键字。

网络数据在 Linux 系统的处理流程如图 1 所示, 以

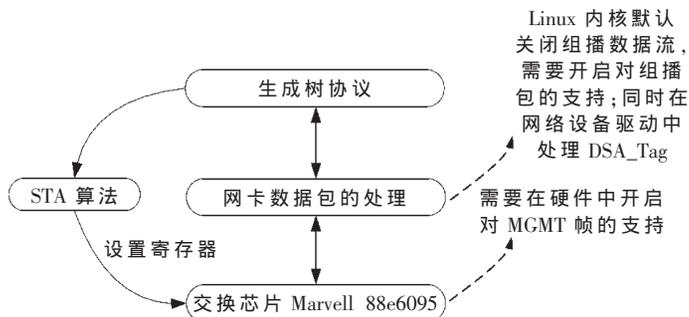


图 1 嵌入式 Linux 的 STP 模块整体架构

# 网络与通信 Network and Communication

以太网数据与 MGMT 帧混合并自由进出系统, Linux 网络驱动将分离出以太网帧由系统正常处理, 而对于 MGMT 帧则需要处理 DSA\_Tag; 当接收时去掉该标记, 而在发送时添加标记。

## 2.2 基于 STP 的网络设备驱动

当硬件设置支持 STP 协议后, 所有网络数据在网卡接收时会添加 DSA\_Tag, 而在发送时去掉相应的 DSA\_Tag, 如图 2 所示<sup>[3]</sup>。

添加 DSA\_Tag 的以太网帧会比正常以太网帧多 4 个字节, 其中 b16~b31 为 STA 算法所用, 其他则应用于 VLAN 模块, 此外 CPU 发送与接收的帧所含的 DSA\_Tag 也不一样, 因此网络驱动也要做出相应改变, 否则由于位数的偏差正常以太网帧会被当成错包丢弃。

## 2.3 生成树算法的实现与分析

以太网交换机实现 STP 模块必须通过 STA 算法计

算出根端口与指定端口, 分析网路拓扑结构, 最终目标是要在网络中生成一颗无环状的树, 以期消除广播风暴以及单播数据帧对网络的影响<sup>[4]</sup>。

STP 协议规定了两种类型的 MGMT 帧, 一种用于交换配置信息的 CONFIG\_BPDU 帧, 另一种则是通知网络拓扑变化的 TCN 帧。

如图 3 所示, 在接收端, Linux 内核接收 MGMT 帧会依次执行以下任务: 更新 CONFIG\_BPDU 帧、执行 STA 收敛算法、设置寄存器; 而在发送端 MGMT 帧由定时器定时发送, 实时监控网络拓扑结构。

## 2.4 内核中与 STP 协议相关的重要数据结构

Linux 内核中, 每个交换机对应一个与网桥有关的数据结构 Struct Net\_Bridge, 包含该交换机全部的端口信息与状态, 并通过 BPDU 包实时更新。而每个交换机端口对应一个 Struct Bridge\_Port, 由 port\_list 串成链表, 存

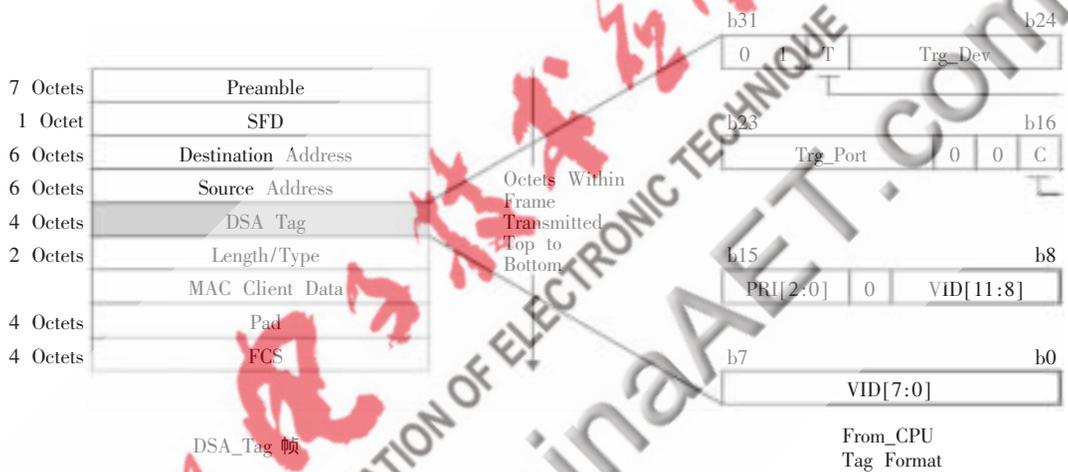


图 2 添加 DSA\_Tag 的以太网帧

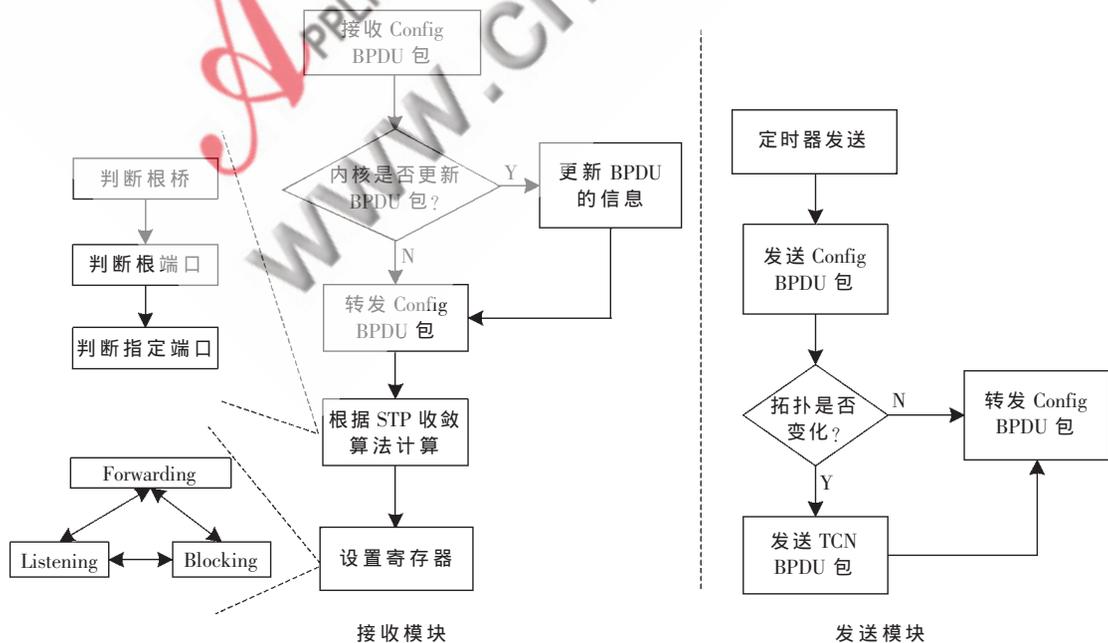


图 3 MGMT 帧在内核中的处理流程

置于 Struct Net\_Bridge 中,如图 4 所示。

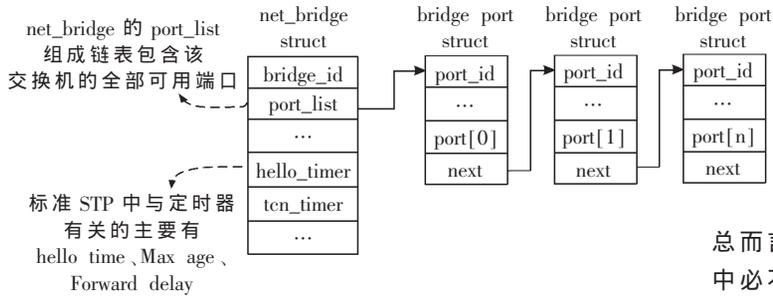


图 4 透明网桥中的重要数据结构

其中 Net\_Bridge 包含与标准 STP 有关的参数,如定时器参数 Hello Time、Max Age、Forwarding Delay 等, Bridge\_Port 则只包含交换机端口信息。

### 3 STP 模块的应用

STP 模块是千兆以太网交换机系统开发中必不可少的一部分,与 VLAN 共同组成了交换机的网络模块。

交换机系统架构如图 5 所示,系统整体上分为应用层、内核层、硬件层。在实际设计中,STP 模块被编译成

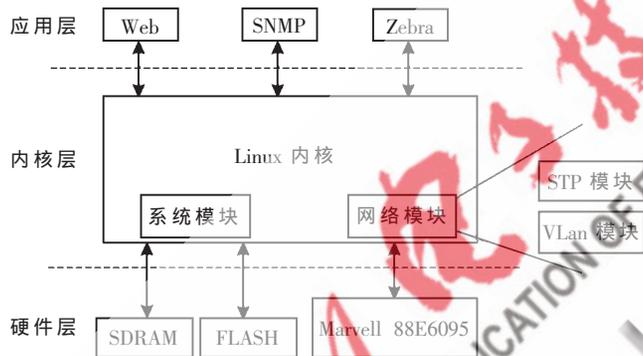


图 5 交换机系统架构

驱动插进系统 Linux 内核中。模块化的管理将更为有效地提高系统稳定性,方便系统的维护与版本的更新。

通过对 Linux2.4x 内核中透明网桥代码分析,利用 Marvell 88E6095 交换芯片,实现了千兆以太网交换机 STP 模块的开发,并提出了嵌入式 Linux 下 STP 模块开发的设计方法与设计思路。由于篇幅所限,文章无法细致分析 STA 算法,且设计时遇到的具体困难也无法详尽说明。

总而言之,STP 模块开发是以太网交换机软件系统开发中必不可少的环节,将极大增加交换机的可靠性与稳定性。且由于 Linux 系统的免费开源等优点,基于 Linux 的 STP 模块的开发在嵌入式领域有广阔的前景。

#### 参考文献

- [1] 李雪梅,孙旭.以太网交换机中 STP 算法的优化设计.微电子学与计算机,2006,23(5):53-55.
- [2] 张劲松,徐志根.快速生成树协议端口角色选择算法的研究[J].西南民族大学学报(自然科学版),2005,31(02):289-293.
- [3] Link Street. 88E6092/88E6095/88E6095F Datasheet. 2006.9.
- [4] 李延冰,马跃,王伟,等.基于生成树的链路层拓扑发现算法[J].计算机工程,2006(18):38-41.

(收稿日期:2010-02-24)

#### 作者简介:

陈麓,男,1980年生,助理工程师,硕士,主要研究方向:通信与信息系统。

鹿瑾,女,1980年生,助理工程师,硕士,主要研究方向:通信与信息系统。

徐曼,女,1986年生,硕士研究生,主要研究方向:通信与信息系统。