Scientific Research

Design of Connection Tracking-Based Packet Filtering Card Based on TCAM

ZHANG Shaobo, LIU Ming

Institute of Science, Information Engineering University, Henan, Zhengzhou, China

Abstract: This paper introduces the technology of firewall based on connection tracking and TCAM briefly, discusses the mechanism of regulation searching and comparing by using TCAM. It focuses on the ways of how to construct a high-speed packet filtering card and how to process the packets passed in and out. Finally, it studies the problem of updating algorithms deeply.

Keywords: connection tracking; firewall; TCAM; packet Filter; updating algorithms

基于TCAM的状态检测包过滤卡设计

张少波，刘 明

信息工程大学理学院，郑州，中国，450001

摘 要: 文章对状态检测防火墙和TCAM技术进行了简要介绍，对使用TCAM进行规则查找比对的工作机理进行了探讨，进而提出了基于TCAM的高速报文过滤卡设计方案，分析了它对进出报文的处理过程。最后，对TCAM的规则更新算法进行了较为深入的研究。

关键词: 状态检测，防火墙，三元内容可寻址存储器，报文过滤，更新算法

1 引言

传统的包过滤防火墙完全基于规则匹配机制，即当接收一个数据包后，检测该数据包头的相关信息与不同的规则依次进行比对，根据匹配结果决定处理情况。这种方法在规则数目较少时性能较好，但当规则数目逐渐变大时，对每个到达的数据包都要从庞大的规则集中找寻相匹配的规则，这显然是低效的。

状态检测技术[1]是近几年才应用的新技术，采用的是一种基于流量的状态检测机制，将属于同一流量的所有包作为一个整体的数据流看待，构成流量状态表。通过状态表中维护的流量信息，避免了对规则表的频繁访问，具有更好性能。

文章所设计的包过滤卡处于安全网关前端，采用PCI 总线与其后的网络处理器交换数据，按照状态检测机制对进出的数据包实施高速过滤，屏蔽不明网络试探和拒绝式服务性攻击数据包的进入，过滤本地子网发出的虚假源 IP 数据包。

2 状态检测过滤机制设计

按照状态检测包过滤的工作原理，我们设计了两张表——规则表和状态表，其中规则表为一个较大的静态规则集合，规则的操作由人工完成；而状态表为紧凑的动态规则集合，对其自身添加的每个表项都设置了固定大小的 TTL 值（时戳），TTL 值随时间进行递减，当其变为 0 时该表项将会被自动清除。

2.1 数据包过滤流程

数据包过滤流程如图 1 所示。

首先，系统接收数据包并存放于缓存中，由报头提取模块提取报头相关信息作比对处理。如果该数据包是打有 SYN 标志的 TCP 包，则根据该数据包的五元组信息与规则表进行匹配。若规则表中不存在相应的匹配表项，或者所匹配的规则表项 r_i 的处置信息域为拒绝，该数据包被丢弃；若 r_i 的处置信息域为通过，则把该数据包的流量信息记录到状态表中，并直接转发该数据包。

如果该数据包不是一个 SYN 包，则直接查找状态表。若在状态表中匹配到了状态表项 s_j，则根据 s_j 的处置信息域进行相应的操作，并刷新该表项的 TTL 值；若没有任何状态表项和该数据包匹配，则进一步把该包与规则表进行匹配，处理流程与 SYN 包类似。

2.2 “TCAM+SRAM”结构

状态包过滤的规则匹配和状态匹配均采用“TCAM+SRAM”结构实现。TCAM[2]是一种专用于进行查表操作的硬件芯片，能够在一个硬件时钟周期内完成关键字的精确匹配查找。TCAM 每一位有 0(zero)、1(one)和 X(Don’t care)等三种可能的值，当执行匹配操

978-1-935068-04-4

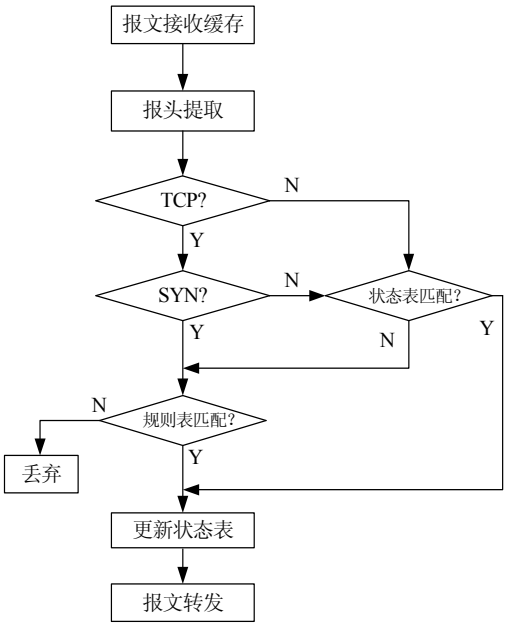

Figure 1. Working procedures of packet filter

图1. 报文过滤流程图

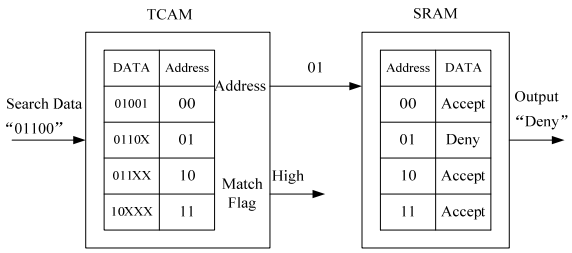

Figure 2. Searching and comparing mechanic by using TCAM

图2. TCAM查表示意图

作时，只有值为 0 或 1 的位会被检查，而被标记为 X 的位会被忽略，因此 TCAM 的匹配方式有精确匹配和范围匹配等两种。

在使用 TCAM 进行规则匹配时，通常需要两类存储器：1 个可以存储规则集的 TCAM 存储器和 1 个存储规则匹配后具体动作的相联存储器（通常为 SRAM），这种无需改动硬件就能改变查找键值与附加信息之间的绑定关系，其工作原理如图 2 所示。

TCAM 是由许多个大小固定的槽组成，根据工作需要，1 条规则可以同时占有 1 个或多个槽，并且每条规则对应一个 Action。当 1 个数据包到来时，包头提取部件从该数据包头提取一个查找匹配的键值，同时将该键值传送至 TCAM 进行查找匹配处理。当在 TCAM 的规则集中匹配到 1 条符合的规则时，该匹配规则就会映射到相联存储器 SRAM 中的地址。系统通过此地址查找出应执行的动作，进而决定数据包的处理策略。

其核心采用 FPGA 实现。考虑 TCAM 占用空间较

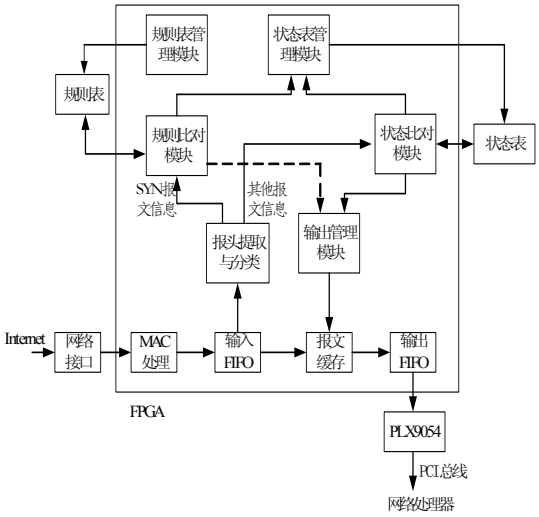

Figure 3. The internal structure ofhigh-speed packet filtering card

图 3. 高速报文过滤卡原理图

3 报文过滤卡设计

3.1 报文过滤卡硬件设计

报文过滤卡设计原理如图 3 所示。

大，我们采用 ALTERA 公司的 APEX 20K1000E 系列芯片，它具有 150 万逻辑门，51840 个逻辑单元，内部可定义的存储单元达到 442368 位，完全可以满足设计需要。

PLX9054 是一种桥接芯片，负责在 PCI 总线和 LOCAL 总线之间快速传递消息，既可以作为两个总线的主控设备去控制总线，也可以作为两个总线的目标设备去响应总线。在本设计中，PLX9054 工作在 C 模式，以 DMA 方式传输报文，并通过 M93C46 加载配置数据。

我们采用外扩 SRAM（静态存储器）的方式来存放规则匹配后的动作，如接收或者拒绝。SRAM 选用日立公司的 HM62W8512，该芯片容量为 4M，具有高密度、高性能和低功耗等特点。

此外，还有一些辅助的外围器件，如 PHY 芯片选用 Micrel 公司的 KS8721BL，使用 MII 和 RMII 接口标准来传输网络报文数据；CY2308 负责提供整个板卡的时钟信号，E2PROM 负责加载 FPGA 的初始配置文件，JTAG 则提供了 FPGA 的调试接口等。

3.2 TCAM 设计

考虑 TCAM 的价格较高，本设计的 TCAM 采用 VHDL 硬件描述语言设计，通过 FPGA 予以实现。在设计中，可借助 QUARTUS II 提供的 altcam 宏模块，选用其 multiple-match 模式实现，如图 4 所示。

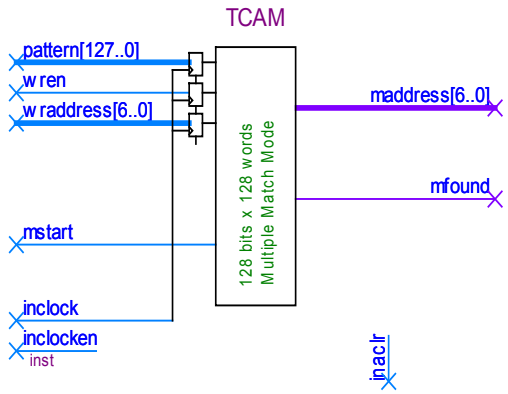

Figure 4 Design of TCAM by using VHDL
图4. 利用altcam宏模块设计TCAM

TCAM 中存放状态表项的五元组信息，配置为 144 位，一条五元组信息占用两个连续地址单元，其中偶地址为高 72 位，奇地址为低 72 位。如地址 0 为高 72 位([143:72])，地址 1 为低 72 位([71:0])。数据以及掩码地址空间均为 0x0000-0xffff (64k×144b)。

IP 报文五元组包括 32 位源 IP 地址、32 位目的 IP 地址、16 位源端口、16 位目的端口及 8 位协议域，共计 104 位。其中源 IP 地址、源端口、协议域(共 56 位)拼上 16 位“1”写入高 72 位单元，目的 IP 地址、目的端口拼上 26 位“1”写入低 72 位单元。72 位数据中只有低 64 位可配置，高 8 位由硬件系统自动置“1”(写入信息时)或置“0”(清除信息时)。对 CAM 的 72 位数据阵列和 MASK 阵列，我们只用其低 64 位，对其进行写操作时自动将[72：64]补全“1”写入，对于 CAM 内部寄存器，因为都只有低位数据有效，写时将最高 8 位再复制作为 72 位数据的最高 8 位。

之后，该五元组作为查表关键词输入 TCAM，由 TCAM 进行高速比对。若比对失败，通过置零 Match Flag 位来告知输入控制模块；若比对成功，通过地址总线访问 SRAM，察看对该数据报文的处理策略，并把该结果返回到输入控制模块。根据查表结果，输入控制模块提取 FIFO1 中的数据报文进行处理。当 TCAM 比对失败或返回的处理结果为“Deny”时，把该数据报文转发至包丢弃池 SRAM3，丢弃该报文；当返回结果为“Accept”时，把该报文转发至 PXL9054，通过 PCI 总线传输至网络处理器，以做进一步的处理。

3.3 报文处理流程

内网发送的数据报文首先由网络处理器进行处理，之后通过 PCI 总线发送至报文过滤卡，经 PLX9054 桥接转换后送至 FIFO2 进行缓存，并发送至报头提取模块提取五元组的数据报文已备查表。

考虑到内网规模一般较小，安全性也远好于外网，我们仅考虑过滤源 IP 为本地网 IP 范围以外的欺骗性数据报文，这样所需的规则并不是太多，所以我们设计的规则比对模块采用逐条比对的方式与规则库 SRAM2 进行比对，并把比对结果返回至输出控制模块，后者从 FIFO2 中提取数据报文，按照与输入控制模块相同的策略处理输出数据报文。

4 表项更新算法

4.1 规则表更新算法

TCAM 的规则更新指的是删除、修改 TCAM 中的规则或向其添加新规则。考虑到规则删除是规则添加的逆过程，二者过程差别不大，而规则修改也比较容易实现，为简单起见，我们在讨论算法时仅考虑规则添加操作。

更新复杂度和查找速度是 TCAM 查找方法的两个重要指标。更新复杂度代表了规则发生变化时的处理能力，而查找速度代表了快速比对的能力。由于在规则表更新时一般不能进行规则查找匹配，而是先将报文缓存，等更新完毕再进行处理，这就要求规则更新速度要尽可能快，否则可能会因缓冲区溢出而造成数据报文丢弃。TCAM 实现简单，查找速度快，复杂度为 $O(1)$，但规则表的更新较为复杂。

4.1.1 TCAM 规则存储方式

在 TCAM 中，长度相同的规则组成一个规则集，多个规则集按照其规则从长到短的顺序从 TCAM 的低地址开始依次存储。若令 R_i 代表长度为 i 的规则集，r_i 代表一条长度为 i 的规则，则当 $i > k$ 时，R_i 中的规则都应该存储在 R_i 中的规则之前（i,k 均在规则长度范围之内）。

由于 TCAM 要求所有的规则按照长度有序存储，所以 TCAM 的规则更新的效率较低。为了提高 TCAM 规则的更新效率，本文在分析传统更新算法的基础上，提出了一种改进的更新算法。

4.1.2 几种传统更新算法[3]

顺序移动法

这是最简单的一种更新算法，如图 5（a）所示。当需要添加一条长度为 k 的新规则 r_k 时，需要把长度小于 k 的所有规则集依次后移一个位置，留出空闲位置后把该规则加入，其效率很低，最差情况下算法复杂度为 $O(N)$（N 当前 TCAM 中保存的规则数目）。

预留空间法

为了尽量避免由于规则插入而造成其它规则的大规模移动，可以为每个长度的规则集预留部分空闲的表项，如图 5（b）所示。当需要加入新的规则 r_k 时，若 R_k 中包含空闲表项，就不需要进行规则移动操作，

978-1-935068-04-4 © 2009 SciRes.

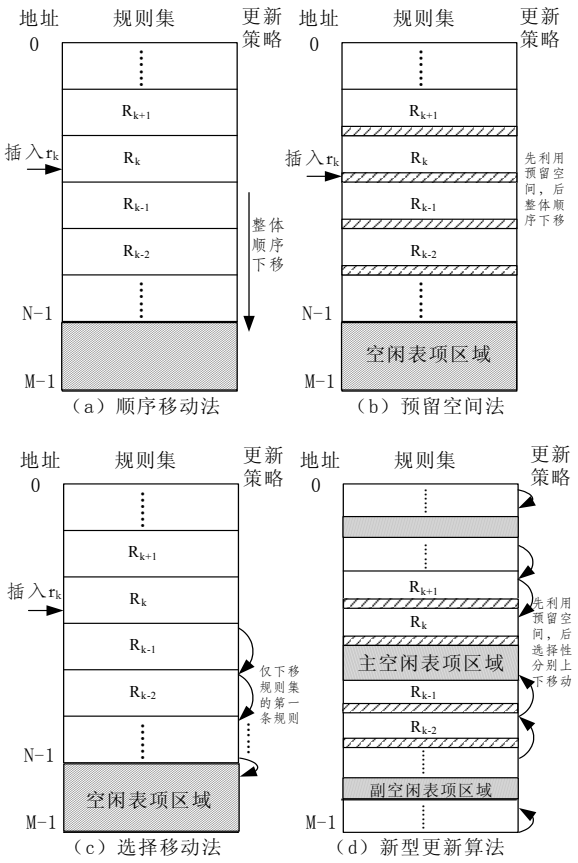

图5. TCAM规则的更新算法

Figure 5. The updating algorithms of TCAM

直接进行添加，否则从相邻的规则集 R_{k+1}、R_{k-1} 中借用空闲表项进行添加，只有在以上规则集均不存在空闲表项时才进行顺序移动。该算法大大降低了规则移动次数，能够提高规则更新的平均效率，但降低了规则存储的效率，且最差情况下的算法复杂度仍为 $O(N)$。

选择移动法

TCAM 要求规则集之间按照长度降序进行排列，但对于每个规则集内部各个规则之间的顺序关系没有严格规定，选择移动法就利用了这个特点，如图 5(c) 所示。

当需要加入新的规则 r_k 时，首先从长度最短的规则集开始，将该规则集的第一项规则移动到空闲表项区域，然后将长度次短的规则集中的第一条规则移动到刚腾 R_{k-1} 出的一条空闲表项处，这样依次类推，直到规则集 R_{k-1} 出现一条空闲表项，在此处添加新规则即可。算法的复杂度为 $O(\omega)$（ω 是规则集的数目），相比前面算法大大降低。

4.1.3 新型更新算法

本设计所采用的新型更新算法综合运用了前述几种算法的思想[4,5]，同时对表项存储进行了两点改进：首先，把空闲表项区域移动到表的中间位置，即把表平均分为上表和下表两部分，这样在表项变动时节省了一半的移动次数；其次，把空闲表项区域进行划分，一半空间作为主空闲区域，仍保留于表的中间位置；另一半空间划分为两个副空闲区域，分别放置于上表和下表的中间位置，即三个空闲区把整个表平均分为四份，以便当表项变动时进一步降低移动次数，如图 5（d）所示。

当添加长度为 k 的规则时，首先利用预留空间法，查看 R_k 及其相邻规则集 R_{k+1}、R_{k-1} 中是否有空闲表项。若不满足要求，则利用选择移动法，把相关规则集的第一条规则向空闲区域逐条移动，最后腾出的一条空闲表项用于添加新规则。考虑到表的主空闲区域较大，因此规定上表中的规则向下移动，而下表中的规则向上移动。该算法是 TCAM 更新平均性能最好的算法，在最坏情况下，该算法的复杂度仅为 $O(\omega/4)$，并已在实际中得到应用。

5 结束语

基于软件的报文匹配过滤方案不能够胜任高速网络的需要，而基于硬件的方法是理想的选择。相比较其它的查询算法而言，TCAM 方案具有更好的性能优势，尤其适用于各种高速的查找操作，在本文的例子中使用了 TCAM 进行规则的查找，经实验证明可以达到比较理想的查找效果。

但是 TCAM 在使用时也存在着一些问题[6]，例如存储量较小且不易扩展、功耗较高以及造价昂贵等，这些缺点都影响着 TCAM 在包过滤等方面的性能的发挥。一些文献对这些问题已经进行了初步的探讨和发展，相信随着不断探索和研究，TCAM 会在更多的领域起着越来越重要的作用。

References (参考文献)

[1] Panigrahy R，Sharma S．Reducing TCAM Power Consumption and Increasitlg Throughput[C].IEEE Hoti,2002.107-112.

[2] Che Hao, Wang Yong, Wang Zhi-jun. A Rule Grouping Technique for Weight-based TCAM Coprocessors [C]. Proceedings of the 11th Symposium on High Performance Interconnects. 2003, 8: 25-28.

[3] Shah D, Gupta P. Fast Updating Algorithms for Tcams [J]. IEEE Micro. Magazine, 2001, 21(1): 36-47.

[4] Zane F, Narlikar G, Basu A．CoolCAMs: Power-Efficient TCAMs for Forwarding Engines[C]．IEEE Infoeonl, 2003.42-52.

[5] Wang Zhi heng, Bai Ying cai.Fast update algorithm for TCAM-based routing lookups[J].Journal of Shanghai Jiaotong University, 2002·E(7): 8-14.