

播存网格中基于 UCL 的硬件过滤系统

孙玉博, 马建国, 朱 敏

(西南科技大学信息工程学院, 绵阳 621010)

摘 要: 针对播存网格, 对节目进行统一内容定位(UCL)标引, 将 UCL 信息封装进 IP 包, 形成用于传输的广播 IP(BIP)包。设计一种基于 UCL 的 BIP 包硬件过滤系统。模拟广播网的数据传输过程, 在源端对节目进行 UCL 标引得到 BIP 包, 以太网终端通过现场可编程门阵列在数据链路层对 BIP 包进行解析和过滤, 并将结果交付用户存储。实验结果验证了该系统的有效性。

关键词: 播存网格; 统一内容定位; 硬件过滤; 内容寻址存储器

Hardware Filtering System Based on UCL in Broadcast-storage Grid

SUN Yu-bo, MA Jian-guo, ZHU Min

(College of Information Engineering, Southwest University of Science and Technology, Mianyang 621010)

【Abstract】 Aiming at broadcast-storage grid, this paper indexes the programs based on Uniform Content Location(UCL), and transmits the Broadcasting IP(BIP) packets comprised of IP packets which is contained of UCL information. It presents a hardware filtering system based on UCL for BIP packet. The transmission process of broadcasting network is simulated, and the BIP packets are formed through UCL indexing in source side. In Ethernet terminal, the BIP packets are resolved and filtered in the data link layer by Field Programmable Gate Array(FPGA). The results are delivered to user to store. Experimental results demonstrate the effectiveness of this system.

【Key words】 broadcast-storage grid; Uniform Content Location(UCL); hardware filtering; Content Addressable Memory(CAM)

1 概述

现代网络存在无尺度分布、数据鸿沟等问题, 文献[1]提出下一代广播和一种新的共享方式——播存网格。播存网格在广播和用户间添加存储环节, 使瞬态数据长期存在并贴近用户, 实现用户对内容的按需服务。为了实现播存网格, 文献[1-2]提出统一内容定位(Uniform Content Location, UCL)的理念及其标引方法。

UCL 是实现播存网格的关键技术, 但其现阶段研究都基于软件平台, 在网络层以上实现 UCL 标引和过滤, 而没有关于其硬件实现方法的研究^[2-3]。软件方法依赖所用软件平台的技术基础, 存在软件瓶颈, 限制了 UCL 技术的推广和应用。因此, 本文提出一种针对 UCL 信息的硬件过滤系统设计方案。

2 UCL 信息标引方法

2.1 UCL 标签

UCL 是网络信息资源的一种属性与内容描述结构, 是基于内容定位的一种新的语义查找方法^[1], 用于精炼地表示语义, 需要对其进行编码。在实验环境中, UCL 按 2 级分类的形式定义, 该 2 级分类的具体名称定义可由应用系统决定。在本实验环境中, 第 1 级分类属于内容提供商; 第 2 级分类属于某个内容提供商的栏目分类。表 1 为 UCL 标签结构定义^[2]。

表 1 UCL 标签结构定义

字段	位数	含义
uVersion	16	版本号, 可细分为多个, 用于表示版本、应用类型等
uGroupID	16	第 1 级分类的标签
uSubID	32	第 2 级分类的标签, 目前不确定该字段是否需要 32 bit, 可以预留一部分空间
uID	32	某个资源的唯一标记, 该标记可以按照日期来生成, 要求是能够唯一标识某一个资源
Reserved	32	保留位, 可用于以后的分类学扩展

2.2 UCL 信息的封装

UCL 数据封装涉及 3 个层次: 基于 IP 层的 BIP 包封装、基于多协议封装(Multi-Protocol Encapsulation, MPE)标准的封装和 TS 流封装, 本文只针对第 1 层 BIP 包进行研究与实现。

BIP(Broadcasting IP)包是一种专用于广播的 IP 包, 并在 IP 包中嵌入了基于统一内容定位的 UCL 代码。它有以下特点: (1)IP 包类型限制。BIP 包是组播类型的 IP 包, 理论上可以跨越所有网段, 其目的是实现数据共享推送的通道。(2)IP 包头嵌入 UCL 信息。UCL 技术选择在 IP option 段嵌入 UCL 信息, 目的是在不影响 IP 包原本结构定义的情况下, 嵌入 UCL 信息。

3 基于 UCL 的硬件过滤系统

UCL 的基础是以库形式存储资源并推送分发资源, 而库以 UCL 标签来管理所存储的资源内容, 每个资源都有自己的 UCL 标签, 终端接收时, UCL 标签就是过滤的关键字段。

(1)资源 UCL 定制。UCL 服务器收集播出的资源内容的 UCL 标签, 形成节目菜单发送给终端, 用户选择后, 将 UCL 标签设定到接收单元, 作为过滤依据。

(2)BIP 数据包过滤。接收单元采用硬件逻辑处理, 对接收到封装有 UCL 标签的 BIP 数据包, 解析获得 UCL 标签, 并与用户定制的 UCL 规则进行匹配, 从而判断应接收或丢弃, 以提高接收效率和实现用户的“按需存储”。图 1 为基于

基金项目: 国家“863”计划基金资助项目(2007AA01Z151); 四川省教育厅科研基金资助项目(08ZA173); 西南科技大学科研基金资助项目(08ZX7117)

作者简介: 孙玉博(1984—), 男, 硕士研究生, 主研方向: 电子设计自动化, 嵌入式系统; 马建国, 教授、博士生导师; 朱 敏, 硕士研究生

收稿日期: 2009-12-07 **E-mail:** hans.sun84@163.com

UCL 的 BIP 包过滤流程。

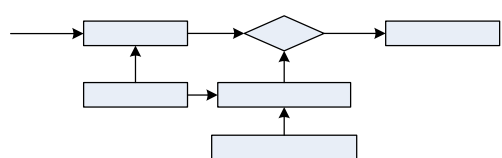


图 1 基于 UCL 的 BIP 包过滤流程

本系统设计面向小型终端，为提高过滤速率和缓解终端系统消耗，设计独立于系统的硬件前端自动过滤系统。发送端封装 UCL 标签形成 BIP 包，通过广播网传输。采用硬件实现 BIP 包的自动接收，在数据链路层解析 UCL 标签，并与用户制定的过滤规则匹配，从而过滤出用户需要的信息。提取信息过滤控制和用户按需存储的目标。用户可以对 UCL 匹配规则进行重新配置，更新用户需求。硬件过滤系统平台见图 2。

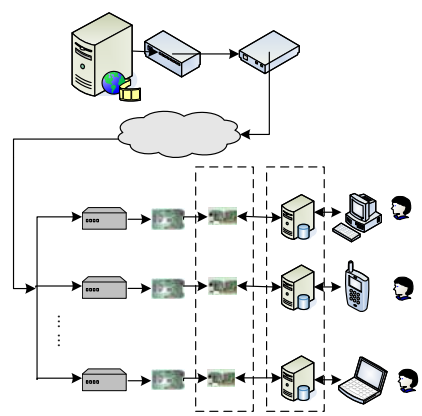


图 2 硬件过滤系统平台

3.1 系统设计

针对硬件信息过滤系统，现场可编程门阵列(Field Programmable Gate Array, FPGA)兼具了 ASIC 的纯硬件高速率特点，且不依赖软件系统，其配置灵活、开发周期短、处理速度快，适合本系统的研究设计。进一步研究发现，UCL 是简单语义描述的代码，本身具有非常强的语义特征，将其封装在数据包里，进行了编码以缩短 UCL 长度。根据对信息的划分等级和类别，要对 UCL 代码的过滤匹配做不同划分。综合 UCL 的特点，本文提出采用 FPGA 构建内容可寻址存储器(Content Addressable Memory, CAM)的方案实现基于 UCL 的 BIP 包硬件过滤。根据实际条件，笔者采用 Altera 公司的 Cyclone2 FPGA 芯片进行设计。UCL 硬件过滤系统结构见图 3。

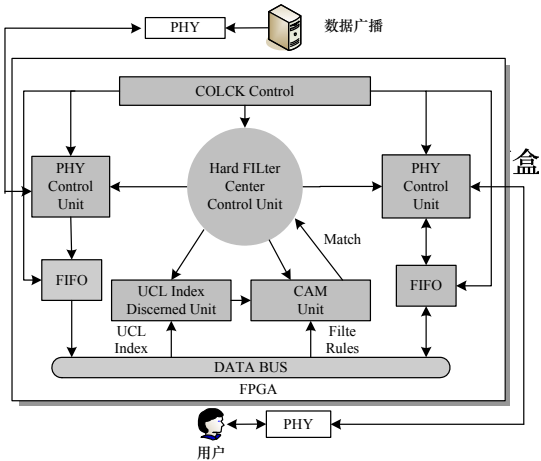


图 3 UCL 硬件过滤系统结构

3.2 CAM过滤规则表

CAM 的设计和实现是整个过滤系统的关键。它是一种专门为快速查找数据地址而设计的存储器，最快能在一个硬件时钟周期内完成关键字的精确匹配查找^[4]。为了提高系统的集成度和速率，系统设计中采用 FPGA 内部构建 CAM 来实现过滤规则表的设计。

针对 UCL 的特点，建立如表 2 所示的 CAM 规则表，本文未把 UCL 版本作为匹配关键字段。BIP 包解析存储配规则含 4 个关键字段，共 112 位，其中，“x”表示无需匹配的字段。

表 2 CAM 规则表

uGroupID	uSubID	uID	Reserved
0001h	00000001h	xxxxxxxh	00000003h
...
...	...	01001000h	xxxxxxxh

用户选择的 UCL 列表

笔者在文献[4-5]的基础上，采用 EP2C35F672 型 FPGA 芯片，基于 M4K 存储器和普通双端口 RAM 设计了一种特殊的快速匹配 CAM。构建一个 2^A Word 深、B bit 宽的 CAM，其电路结构如图 4 所示。wrdatab[B-1..0]为写入匹配项，wrdatab[A-1..0]为写入匹配项地址，两者合并在一起作为双端口 RAM 写地址的输入。

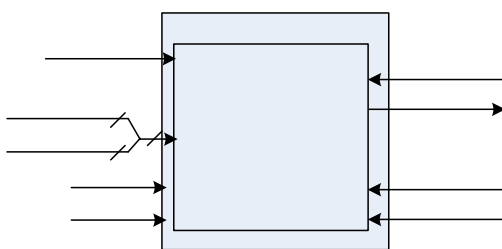


图 4 CAM 电路结构

此快速匹配 CAM 输入匹配项数据和比较查询匹配结果都只需要一个时钟周期，在 EP2C35F672 上最大时钟频率可以达到 210 MHz。16×8 bit CAM 的仿真图如图 5 所示。写入匹配项数据 03h, 15h, 31h, 03h 分别到地址 1h, 3h, 4h, 9h。查询匹配数据 00h, 05h 都没有查询到的情况，输出匹配项地址为 0000h, 03h 为查询到的情况，输出为 0202h(即在地址 1 和地址 9 中查到匹配项数据)，31h 输出为 0010h(即只在地址 3 查到匹配项数据)。

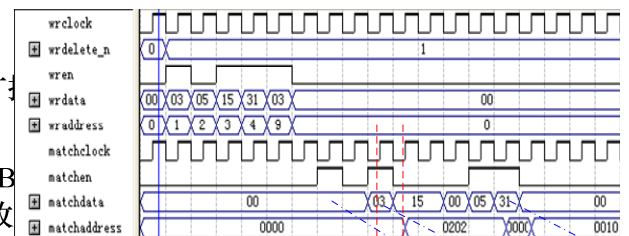


图 5 CAM 的仿真图

上述电路不支持输出匹配信号，且不支持不关心位的比较，为了实现此要求必须添加外围电路。一个 M4K 存储器只能构建一个 16×8 bit 的 CAM。因此，为了设计满足系统要求的 CAM，必须采用并联的方式扩展位宽和深度。

3.3 以太网通信

以太网通信采用 DM9000 网络芯片实现。DM9000 是台湾 DAVICOM 公司推出的一款高速以太网接口芯片，它集成

10/100 Mb/s 物理层接口,内部有 16 KB SRAM 作为接收发送的 FIFO,支持 8 位和 16 位 2 种工作模式。本系统中采用 16 位模式,且为节省硬件成本不使用 EEPROM 进行芯片配置^[6]。

系统中以太网通信设计全部采用 Verilog 硬件描述语言,主要分为以下 3 个部分^[6]:

(1)芯片初始化。激活 PHY,2 次软复位,配置中断状态寄存器,写 6 Byte 以太网地址到物理地址寄存器,清除发送和中断状态,设置中断屏蔽寄存器,配置中断屏蔽寄存器,开始接收数据。

(2)数据接收。读中断状态寄存器,判断接收状态位 PRS 是否为 0,清除 PRS,判断数据帧第 1 个字节是否为 01H,读数据状态和长度判断数据帧是否有错,读取数据。

(3)数据发送。向发送 FIFO 写待发送的数据,写发送数据包长度寄存器,写发送控制寄存器 TCR 请求发送,读中断状态寄存器判断发送状态位 PTS 是否为 0,清除 PTS。

3.4 实验结果分析

实验中采用系统频率 $f=50$ MHz,匹配查询时间为一个时钟周期 $T=20$ ns,以太网芯片采用 16 bit 模式,一个周期读入 $B=16$ bit 数据,一个 UCL 关键字段为 $U=112$ bit,可以算出匹配一个 BIP 包的时间为

$$t = \frac{U}{B} \times T + T = \frac{112}{16} \times T + T = 8T = 0.16 \mu\text{s}$$

判断匹配情况后,将决定是丢弃此 BIP 包,还是继续传送。若继续传送,则可以算出 BIP 包在 FPGA 芯片中的延时为数据包的传输延时。按照网络传输最大数据包为 1 518 Byte 计算,BIP 包的最大延时为 $1\,518T/2=759T$ 。若 $T=20$ ns,则延时为 $15.18 \mu\text{s}$ 。此时,一个 BIP 包的最大延时为 $15.18+$

$0.16=15.34 \mu\text{s}$,则最终可以得到系统过滤 BIP 包的速率为 $1\,518 \text{ Byte}/15.34 \mu\text{s} \approx 791.7 \text{ Mb/s}$,远高于一般以太网和广播网的传输速率,可以满足现阶段以太网和广播网的过滤要求。

4 结束语

本文系统具有一定应用价值,为播存网格和 UCL 信息索引的多层语义挖掘的硬件实现提供了参考。该系统有助于构建跨越网络各层的立体信息理解与过滤机制,实现对网络数据的深度挖掘、分级服务和用户的按需存储,以及在异构的海量数据中获取有意义的信息等。

参考文献

- [1] 李幼平. 营造播存网格——为互联网增添次级结构[J]. 计算机世界报, 2005, (3): 40-43.
- [2] 马建国, 邢玲, 李幼平, 等. 数据广播中的 UCL 标引与传输机制[J]. 电子学报, 2004, 32(10): 1621-1624.
- [3] 薛炳青, 马建国, 李幼平, 等. 播存电视的实验室系统构建[J]. 计算机工程, 2008, 34(19): 280-282.
- [4] An Shengbiao, Gao Shuangxi, Wang Shuhai, et al. CAM Design Based on Virtex Family Devices[C]//Proc. of the 8th International Conference on Electronic Measurement and Instruments. Xi'an, China: [s. n.], 2007.
- [5] 李晨, 王自强, 张东. 基于 RAM 结构的 CAM 的 Verilog HDL 设计[J]. 计算机工程与应用, 2003, 39(27): 157-159.
- [6] Davicom Semiconductor, Inc.. DM9000A Ethernet Controller with General Processor Interface Data Sheet[Z]. 2006.

编辑 陈晖

(上接第 99 页)

了 3 种典型的 UWB 数据率: 53.3 Mb/s 于 CM4 下, 200 Mb/s 于 CM2 下, 480 Mb/s 于 CM1 下。结果显示, 3 种不同的系统参数分别需要 SNR 等于 5.7 dB, 7 dB 和 9.5 dB 以实现文献[1]中要求的 8% 系统误包率性能。与此同时, 如图 4 所示, 本文提出的模拟前端非理想因素估计算法较理想估计情形仅引入了 0.8 dB 的系统性能降低。

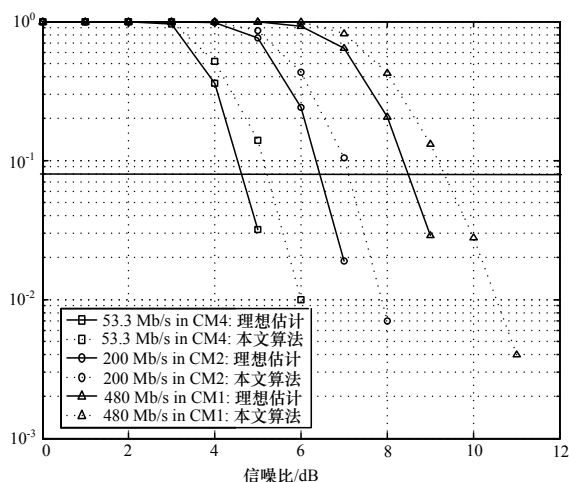


图 4 MB-OFDM UWB 系统误包率

6 结束语

本文针对 MB-OFDM UWB 系统中复杂的模拟前端非理

想因素问题,提出了一种频域联合估计算法。通过利用导频序列,本文提出的联合估计算法可以在存在频率相关性 I/Q 失配的情形下得到准确的 CFO 和 SFO 估计。系统仿真证明,本文提出的联合估计算法能够适用于实际的 MB-OFDM UWB 系统中。

参考文献

- [1] ECMA. ECMA 368-2005 High Rate Ultra Wideband PHY and MAC Standard[S]. 2005.
- [2] Png K B, Peng Xiaoming, Chattong S, et al. Joint Carrier and Sampling Frequency Offset Estimation for MB-OFDM UWB System[C]//Proc. of IEEE Radio and Wireless Symposium. [S. l.]: IEEE Press, 2008: 29-32.
- [3] Valkama M, Renfors M, Koivunen V. Advanced Methods for I/Q Imbalance Compensation in Communication Receivers[J]. IEEE Trans. on Signal Processing, 2001, 49(10): 2335-2344.
- [4] Tang Shigang, Gong Ke, Pan Changyong, et al. Phase Noise Suppression in OFDM systems in Presence of IQ Imbalance[C]//Proc. of IEEE International Conference on Communications, Circuits and Systems. Guilin, China: [s. n.], 2006: 1184-1188.
- [5] Molisch A F, Forester J R, Pendergrass M. Channel Models for Ultrawideband Personal Area Network[J]. IEEE Wireless Communications Magazine, 2003, 10(6): 14-21.

编辑 任吉慧