

# 可重构路由器报文转发引擎设计与实现

陈一骄, 卢泽新, 孙志刚, 李韬, 吕高锋

(国防科学技术大学 计算机学院, 湖南 长沙 410073)

**摘要:** 网络处理功能的时空演化特性要求可重构路由器报文转发引擎除具有基本报文分组交换能力外, 还应具有可重构能力。针对上述需求, 构建了面向可重构路由器的报文转发引擎构件重构模型, 并基于 Pass-Through 模式设计实现了可重构 FPGA 器件与网络处理器相结合的程序/电路构件运行环境。系统实现与应用测试结果表明, 可重构路由器报文转发引擎在保证高吞吐量、低延迟的报文转发处理性能的同时, 可有效支撑多样化业务构件灵活重构与映射。

**关键词:** 路由器; 转发引擎; 可重构; 构件

中图分类号: TP393

文献标识码: A

文章编号: 1000-436X(2012)08-0042-10

## Design and implementation of packet forwarding engine for reconfigurable routers

CHEN Yi-jiao, LU Ze-xin, SUN Zhi-gang, LI Tao, LV Gao-feng

(School of Computer, National University of Defense Technology, Changsha 410073, China)

**Abstract:** The packet forwarding engine should provide not only the basic packet switching function, but also the reconfigurable ability for supporting the network processing functions evolving with time and space. Aiming at the above requirements, a component reconfiguration model of packet forwarding engine for the reconfigurable routers was proposed, and the program/circuit component running environment consisting of FPGA and network processors was designed and implemented based on the Pass-Through mode. The system implementation and application experimental results show that the packet forwarding engine for reconfigurable routers can support flexibly reconfiguration and mapping for the diverse service components, and it can provide high throughput and low-latency packet forwarding ability.

**Key words:** router; forwarding engine; reconfiguration; component

### 1 引言

信息网络技术及应用的迅猛发展对网络路由交换设备(路由器等)提出了巨大挑战, 层出不穷的网络业务需求以及网络带宽和链路速率的飞速增长要求新型网络路由交换设备必须能够提供高

度的可编程性和适应性, 以支持大量多样化用户业务的快速部署和高效实现。然而, 传统路由交换设备体系结构具有刚性和封闭性特点, 对于新型业务的部署通常仅能依赖于设备整体升级和更新换代, 即使是可行的升级和扩展也只能由原提供商实施, 无法充分实现开放和竞争, 网络运营商的投资无法

收稿日期: 2011-09-14; 修回日期: 2012-02-17

基金项目: 国家重点基础研究发展计划(“973”计划)基金资助项目(2009CB320503); 国家高技术研究发展计划(“863”计划)基金资助项目(2008AA01A323, 2008AA01A325, 2009AA01A334)

**Foundation Items:** The National Basic Research Program of China (973 Program)(2009CB320503); The National High Technology Research and Development Program of China (863 Program)(2008AA01A323, 2008AA01A325, 2009AA01A334)

得到有效保护。

此外，随着互联网规模的膨胀及业务模式的演变，对网络路由交换设备的功能需求也在不断演化，网络路由交换设备不应局限于仅提供基本的分组交换能力，而应考虑在数据传输路径上提供较为丰富的功能支撑。如何在网络设备上通过扩展计算、存储等资源，并通过提供开放的标准编程接口（API）有效支持网络设备新的功能原语实现已成为目前的研究热点。

实际上，网络路由交换设备的功能具有时空特性，即路由器的功能随着时间和空间的变化而变化。如图 1 所示，当路由器在  $T_1$  时刻部署在 A 点（如边缘接入网络）和 B 点（如汇聚网络）时需要的功能不同。同样在 B 点，当时间变化到  $T_2$  时，路由器的功能集合也可能变化。例如  $T_1$  时刻路由器需要对网络中爆发的蠕虫病毒进行监测和过滤，而在  $T_2$  时刻蠕虫病毒平息后，并不需要此项功能。同样，在  $T_3$  时刻路由器又部署到 A 点时，其需要的功能又有可能发生变化。

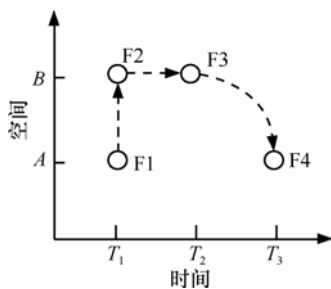


图 1 路由交换设备功能时空特性

基于以上需求和观察，新型可重构路由交换（即可重构路由器）平台<sup>[1]</sup>应运而生，在平台化支撑的构件化处理技术基础上，通过在平台、组件和构件 3 个层次提供开放可重构能力，可有效支持以组合、升级以及替换等方式实现系统功能和性能的演进和扩展。可重构路由器可根据不断变化的功能需求动态加载不同处理构件，不但可以满足各种新出现的互联网处理需求，如视频感知的处理，而且可以降低设备功耗，延长使用周期和减少电子垃圾，更加绿色节能<sup>[2]</sup>。

可重构路由器的设计基于控制与数据平面分离思想，由路由转发平面负责完成数据报文的高速查表、缓冲、交换以及调度等功能，而管理控制平面则主要负责路由计算、设备的管理和配置等功能。

可重构路由器设计与实现中很多开放的问题

有待研究<sup>[3]</sup>。本文聚焦可重构路由器的数据转发平面，着重研究承载各种报文处理任务的转发引擎如何有效支持包括常规转发处理、新的安全、流量管理等业务的重构实现，力图在保证报文处理性能的前提下，通过优化提升可重构路由器数据转发平面的适应性，支撑报文处理业务及网络协议的快速升级、部署及验证。

本文着重研究支撑可重构路由器数据平面处理的<sup>可重构报文转发引擎</sup>，主要贡献包括：在提出构件化报文转发引擎重构模型基础上，通过采用 Pass-Through 模式实现可重构 FPGA 与网络处理器的有机结合，为程序/电路构件的映射实现提供了优化的构件运行环境。可重构报文转发引擎可以灵活实现网络报文处理性能优化，实现多样化程序/电路构件的快速映射与部署，支持多种类型网络接口接入，并可为交换网络的扩展与升级提供有效支撑。

## 2 相关研究

随着网络技术的不断发展，高性能路由器的处理能力和流量也不断提升。然而，商用路由器系统针对性能提升的特定优化却在一定程度上限制了对新型协议以及服务的有效支撑。另一方面，商用路由器系统的封闭性也束缚了第三方软件厂商及增值服务商新业务应用的开发动力。

D.C.LEE 等<sup>[4]</sup>在文献中提出可重构路由器的概念，指出承载报文处理业务的路由器能够在保证处理性能的前提下，具有适应不断变化的网络通信标准和协议的能力。力图能够提高网络报文处理性能，能够更好地适应流量的动态变化。

然而，可重构技术的研究主要集中在可重构计算上，针对网络和路由器的可重构技术研究很少<sup>[1]</sup>。

在实现方式上，可重构路由器系统通常采用通用 CPU、FPGA 以及网络处理器（NP, network processor）等具有可重构能力的技术构建。

基于通用 CPU 实现的路由器又称软件路由器，具有高度可编程性和灵活性，可以提供灵活的业务可重构能力。例如，Click<sup>[5]</sup>是美国麻省理工大学提出的一种由细粒度构件组成的软件路由器，其设计目标是使路由器软件更加灵活并且易于配置和管理。美国华盛顿大学和瑞士苏黎世联邦理工学院提出了一种 Routers Plugins<sup>[6]</sup>，用于在报文分类器匹配特定流时执行，分类器可以安装在任何的“门”间，然而这些“门”在 IP 报文处理路径上是固定的，无

法有效更新和控制路径。PacketShader<sup>[7]</sup>利用 GPU 加速软件路由器报文处理过程,通过应用程序(库程序)控制硬件报文转发过程,可以达到每秒数百万报文的处理速度。然而,这种报文批处理方式也会引入较高的报文处理延迟。

FPGA 技术具有可编程、可擦除、重复使用等特点,可以在不改变外围电路的情况下,通过用户编程实现不同功能。文献[8]中提出了一个基于流的可重构路由器原型结构,并基于 FPGA 实现了可重构硬件处理部件。华盛顿大学可重构研究小组则提出了基于 FPX(field-programmable port extender)<sup>[9]</sup>实现的路由器动态可重构解决方案,其目的是充分利用可重构硬件的高性能和灵活性,开发新的网络应用。斯坦福大学的研究人员则利用网络教学开发平台 NetFPGA<sup>[10]</sup>实现了一个参考可重构路由器,作为课程教学与科研实验的参考设计。SwitchBlade<sup>[11]</sup>基于 FPGA 预置丰富的报文处理功能集合及虚拟数据平面处理,以增强新型网络技术的快速开发和部署能力。然而,完全基于 FPGA 的设计则受限于 FPGA 容量和处理能力,更适用于原型系统验证以及网络实验系统构建等方面。

网络处理器是目前构建商用路由器系统采用最广泛的器件,其试图将精简指令集处理器的低成本、灵活性以及 ASIC 专用网络处理芯片的高性能有效结合,满足网络设备在硬件加速和软件可编程两方面的设计需求。早期,主流网络处理器体系结构通常采用应用特定指令集处理器 ASIP<sup>[12]</sup>方式实现,针对网络任务优化设计的指令集有利于网络业务处理能力的有效提升。然而,ASIP 方式通常要求开发人员使用低层的微码进行编程和优化,要求程序员对网络处理器硬件体系结构有较为深入的了解,对开发人员要求较高。

基于高性能通用多核的网络处理器<sup>[13,14]</sup>作为通用微处理器技术在网络处理领域的延伸,因其通用指令集体系结构带来强大的可编程性和易于使用的设计开发环境近期得到更为广泛地关注。该体系结构大幅降低了网络处理器应用程序设计门槛,却在一定程度上阻碍了系统性能调优。为有效提升网络报文处理性能,基于高性能通用多核的网络处理器通常集成多个硬件加速器,用于加速网络特定业务处理。然而,这种固化的硬件加速器不具备重构能力,难以针对可重构路由器报文处理功能需求演化提供有效支撑。

基于上述相关研究,可重构路由器的设计,尤其是数据转发平面,面临着性能与灵活性(可重构能力)的两难选择。一方面,可重构能力的提升往往会以降低报文处理性能为代价;另一方面,为获得高性能报文转发能力则通常必须对路由器系统进行特定优化设计,而调优过程则意味着路由器功能的固化及灵活性的降低。

### 3 构件重构模型及运行环境

针对可重构路由器在性能和灵活性方面的两难问题,本文在分析路由器报文转发处理工作特点的基础上,提出一种新型异构报文转发引擎构件重构模型。通过将不同业务功能类型优化映射到报文转发引擎的不同处理构件,在保证报文处理性能的前提下,有效支撑可重构路由器系统服务及业务的承载重构能力。

通用路由器报文转发平面处理功能主要由程序功能和电路功能构成。程序功能(又称软件功能)指依赖于处理器指令集实现的逻辑功能,以软件代码形式呈现,适合于控制流较为复杂的非线性任务处理,通常难以达到较高的处理性能。

电路功能指依赖于硬件部件(ASIC 加速芯片、硬件加速器、FPGA 逻辑等)实现的逻辑功能(如 CRC 校验、加解密处理等),这些逻辑功能对处理性能要求较高,较适于线性任务的并行处理,从而可以高效利用硬件在并发处理资源方面的优势实现报文处理性能加速。

可重构路由器路由转发平面功能实现同样依赖于多种程序功能与电路功能的协同,但在针对业务时空性的可重构能力支撑方面提出了更高要求。

#### 3.1 报文转发引擎构件重构模型

针对上述需求,本文提出了一种报文转发引擎重构模型,该模型通过程序/电路构件到构件运行环境的按需映射可有效支撑网络业务应用的演化部署。

如图 2 所示,网络业务应用通常由一组较小的“功能应用”即任务构成,例如报文分类、IP 转发等,这些处理任务可以采用任务图形式表示,图中节点表示报文处理任务,边表示任务间的数据控制通道。

任务图中各任务可由不同构件实现,构件可以有多种实现实例,例如输出调度构件,可以有公平队列、优先级调度和先进先出等多种调度构件实例。构件库用于存储多种按标准接口开发的电路及程序构件实例,需要注意的是,同一构件可以有一

个或多个电路或程序构件实例。为打破传统路由平台封闭性束缚,可重构路由器为程序/电路构件制定了相应的标准化接口,不同构件实例只需满足标准化接口要求,即可纳入构件库管理,从而有效支持第三方用户程序/电路构件的动态加载、安装、升级及卸载功能。

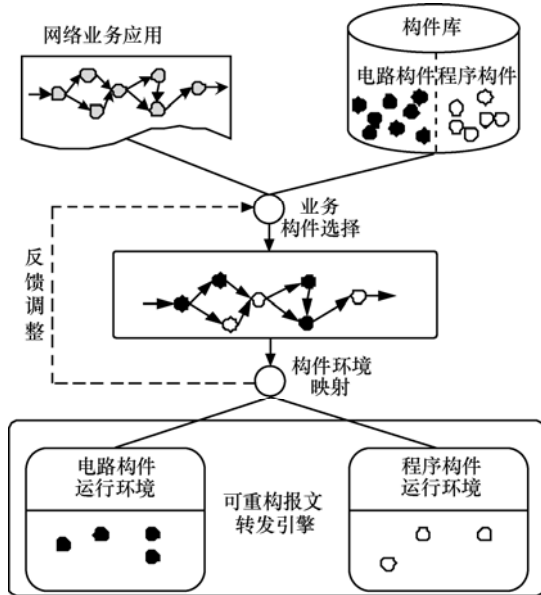


图 2 报文转发引擎构件重构模型

业务构件选择用于根据系统功能性能需求及约束,确定任务图中各节点对应的电路/程序构件实例。构件运行环境为构件实例在可重构路由转发引擎中的运行提供标准接口和运行所需资源。为支持程序及电路功能的动态加载和替换升级,可重构路由器报文转发引擎相应提供 2 种构件运行环境,即程序构件运行环境及电路构件运行环境。可重构路由器报文处理能力生成基于构件实例到构件运行环境的有效映射,构件映射方案必须满足以下约束:

CR1 (资源约束): 用于构件到构件运行环境的映射的逻辑/存储资源不得超过该构件运行环境剩余可用资源;

CR2 (功能约束): 构件到构件运行环境的映射必须匹配构件运行环境功能规范要求,例如流量控制构件不可映射到网络处理器加/解密处理引擎;

CR3 (相关性约束): 对于任意 2 个构件,若它们有直接或间接的数据控制通道连接,则仅可映射到同一运行环境或具有同向数据控制通道的不同运行环境中。

满足上述约束的映射方案称为可行映射,不可

行映射的性能及资源消耗信息则可以用于反馈调整业务构件实例的选择,以实现构件系统的性能迭代调优。

### 3.2 程序/电路构件运行环境

报文转发引擎重构模型中,程序/电路构件运行环境是承载网络处理构件化任务实现的核心,除应为报文转发等业务处理提供资源和性能优化支撑外,还应提供足够的灵活性以支持多样化构件的动态加载、升级和部署,满足可重构路由器功能及性能的时空性要求。

针对上述需求,构件运行环境采用面向网络处理特定业务优化设计的网络处理器和具有快速电路可重构能力的 FPGA 相结合的方式构建。其中,基于冯·诺依曼模型的 NP 线程具有强大的控制流处理能力,较适合非线性构件(带有反馈回路的)实现,而其并发处理能力受限于线程数目,内部固化的硬件加速器则仅限于特定业务构件的加速处理。FPGA 处理引擎(协处理器)具有较强的业务并行处理及可重构能力,适合线性构件实现,而其对构件功能的支撑则受限于内部逻辑及存储资源。

为有效满足多样化程序及电路构件环境在功能、性能以及可重构等方面的需求,将 NP 与 FPGA 有机结合为构件功能提供优化的运行环境。图 3 给出了基于 NP 和 FPGA 器件实现可重构路由器报文转发引擎构件运行环境的 2 种设计选择。其中,程序构件运行环境主要包括 NP 中多核线程池;而电路构件运行环境则主要由 NP 中固化的硬件加速部件(HA, hardware accelerator)和可重构 FPGA 器件等构成。

传统 Look-Aside 模式<sup>[15,16]</sup>将 FPGA 作为从设备旁路挂载在 NP 上,输入报文进入 NP 后被分配到空闲线程处理,对应线程采用 Run-to-Complete 方式负责报文的大部分处理任务。在处理过程中,对于需要加速的特定任务,相应线程调用内部集成的硬件加速器(HA)或者外部协处理器完成。

基于 FPGA 实现的外部协处理器可以根据变化的任务功能进行重构,有效支持多样化业务的演化和部署。然而,Look-Aside 模式下(图 3(a)),实现协处理的 FPGA 与网络处理器间的控制交互及数据传递开销严重阻碍了报文处理性能的提升。虽然网络处理器论坛(NPF)专门针对该模式制定了高速 LA-1 接口标准<sup>[17]</sup>,实际性能测试结果反映 Look-Aside 模式访问延迟等关键性能指标依然难以

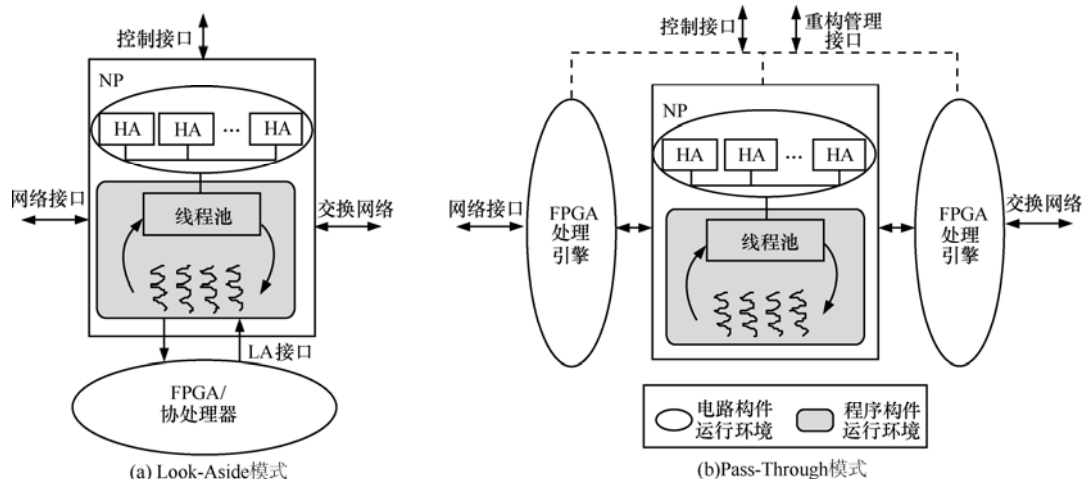


图 3 构件运行环境设计模式

满足高速报文处理需求。本文基于典型的 Netlogic XLR 网络处理器<sup>[18]</sup>, 采用 6 核多线程实现基于 LA-1 接口进行 32bit TCAM 查表报文转发处理, 报文转发率最高仅为 12.42Mbit/s, 与不采用 TCAM 查表的报文转发性能相比, 降低约 26%。

针对传统 Look-Aside 模式存在的问题, 本文提出了一种基于 Pass-Through 模式的可重构构件运行环境。如图 3(b)所示, Pass-Through 模式将 FPGA 处理引擎引入报文处理数据通路, 可有效避免程序构件与可重构电路构件间的控制及数据交互开销, 有效提升网络业务应用处理性能。与 Look-Aside 不同, Pass-Through 模式通过提供重构管理接口, 可支持多样化构件到构件运行环境的动态映射, 实现系统功能及性能按需重构。此外, Pass-Through 模式还具有以下优势。

1) 网络业务构件优化映射: 通过在数据通路引入 FPGA 可重构逻辑处理能力, 扩展了业务构件优化映射空间, 可有效支撑程序/电路构件到相应构件运行环境的优化映射, 支持具有时空性的业务构件动态重构。

2) 支持交换网络拓扑动态配置: 位于输出端 FPGA 处理引擎的集成可以有效支持交换网络拓扑动态可配置。利用 FPGA 处理引擎的逻辑可重构能力, 可以配置线卡交换单元中交换开关的交换方式, 将交换网络配置成冗余或负载均衡等工作模式。此外, 还可以有效实现交换网络吞吐率的动态可调整。

3) 解耦 NP 网络及交换接口: Pass-Through 模式下, FPGA 处理引擎实现了 NP 与网络接口及交换单元接口的解耦, 使系统可灵活支持丰富的网络

接口以及交换接口类型, FPGA 提供的可重构能力则进一步为构件运行环境支持新型网络及交换接口逻辑提供了可行性。

#### 4 系统实现与重构控制机制

可重构路由器报文转发引擎基于 Pass-Through 模型设计实现, 主要由输入/输出控制 FPGA (Altera 公司 Stratix II GX EP2SGX60EF1152C4 器件)、2 片高性能多核网络处理器 XLR732 及其配套外围器件、QDR 接口 TCAM 以及管理配置 EPLD 等部件组成, 为程序/电路构件提供灵活的可重构运行环境。基于上述体系结构实现的单引擎的路由转发能力可达 20Gbit/s, 相应实现框图如图 4 所示。

多核 XLR 网络处理器是报文转发引擎的核心, 基于同时多线程的高性能处理器核为报文转发功能提供了强大的程序构件运行环境。NP 中集成或外置的多样化硬件加速部件 (如加解密、压缩以及查表等), 则可以有效实现功能相对稳定的特定电路构件的有效映射和加速。QDR 接口 TCAM 是 NP 查表加速部件, 用于增强报文转发引擎查表性能。ICF 连接 NP 与网络接口卡, ECF 则实现 XLR 与核心交换单元的连接。I/O 总线控制 EPLD 主要实现 NP 的 I/O 总线 and 局部总线功能, 支持 NP 的 BootFlash 管理、复位、启动以及对其他器件的管理。

##### 4.1 数据/控制通路

报文转发引擎数据通路用于实现报文路由查表、转发决策和 QoS 控制等功能。以输入侧为例, 数据通路对于报文的基本处理流程如下。

网络接口输入报文经过 ICF 的预处理后, 将根

据预设规则分派到特定 NP 上进行处理，该 NP 调度线程池中空闲线程负责输入报文的处理工作；在需要时，处理线程调用相应硬件加速器 HA 对报文处理任务进行加速；报文处理完成后，NP 将报文发送到 ECF 进行报文输出相关处理工作，通过交换单元接口发送到交换矩阵转发。

为支持上述报文处理流程，ICF 预置基本的网络接口逻辑和报文分派逻辑，ECF 则预置输出流量处理逻辑以及交换接口逻辑。除基本预置处理逻辑外，ICF 和 ECF 剩余可用的逻辑及存储资源用于实现可重构电路构件运行环境，为电路构件加载、升级或重构提供相应资源。

可重构路由器报文转发引擎的控制通路主要由吉比特以太网实现，如图 4 所示，XLR 0 和 XLR 1 分别通过 RGMII 接口与 ICF 连接，ICF 通过背板基于吉比特 SerDes 接口连接主/备份控制 CPU，实现主/备份控制 CPU 与 NP 间控制信息交互。

#### 4.2 重构控制机制

根据构件类型的不同，重构管理控制也分为程序重构管理控制和电路重构管理控制。其中，程序重构主要基于 NP 软件代码的更新与加载，实现较简单，而基于 FPGA 的电路重构的控制实现则相对复杂。

可重构报文转发引擎电路重构的实现结构如图 5 所示，由重构代理、配置引擎、配置文件存储器和待配置功能部件（即 ICF/ECF FPGA）4 部分组成。

在本路由转发引擎中，重构代理由网络处理器

XLR 的一个线程实现，主要完成配置命令解析、配置文件的读/写、更新以及启动配置；配置引擎基于 EPLD 器件实现，该器件通过局部总线与 XLR 连接，主要完成 FPGA 的配置控制以及配置文件存储器（flash）的读写控制；配置文件存储器为非易失性 flash，用于存储 FPGA 的重构配置文件。

在电路重构中，用户首先根据应用需求从构件库选择对应的电路构件，针对 FPGA 类型生成并配置比特流文件。重构代理根据用户命令调度选择对应构件的配置比特流文件，通过 FPGA 配置引擎加载该配置文件，从而实现 FPGA 的静态重构。

需要注意的是，在系统加电运行过程中，通过配置控制逻辑主动加载 FPGA 的配置文件，需要中断系统运行，仅可实现 FPGA 静态重构。在需要系统在线动态重构的情况下，可选择采用 Xilinx 或 Altera 等公司支持部分可重构功能的高端 FPGA 实现 ICF 和 ECF，从而在不中断系统运行的前提下实现电路构件的运行更新，提供更高灵活性和重构性能。

在部分动态可重构中，重构管理控制方式与静态重构类似。前期，用户需要首先按照模块化设计方法，将整个设计分解为固定部分和重构部分。然后分别对固定模块和可重构模块进行整体设计综合，根据应用需求生成一个缺省全局配置比特流文件，同时生成所有可重构模块的部分配置比特流文件。在初始配置过程中，重构代理选择加载 FPGA 缺省全局配置文件，在动态部分可重构过程中，可重构代理根据用户命令

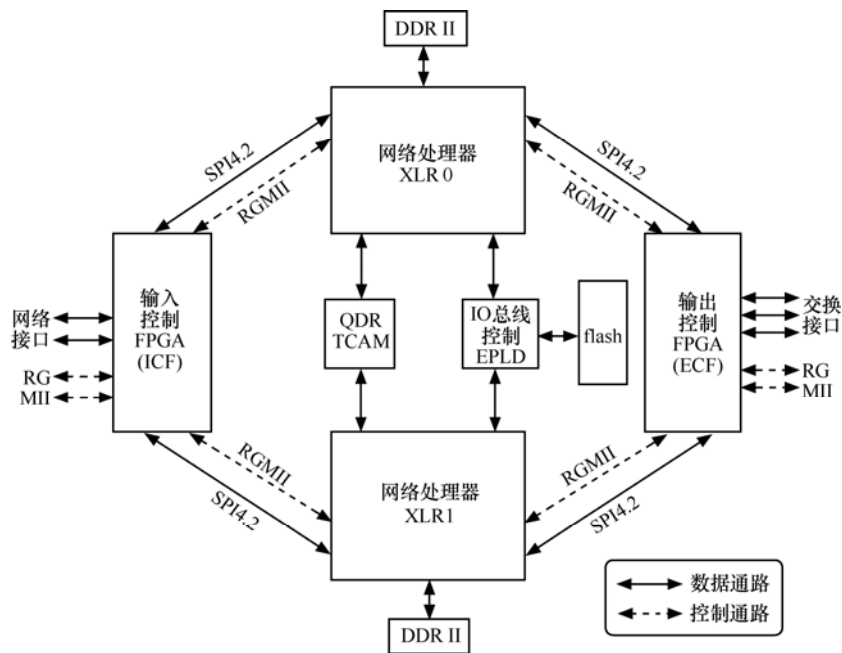


图 4 报文转发引擎实现

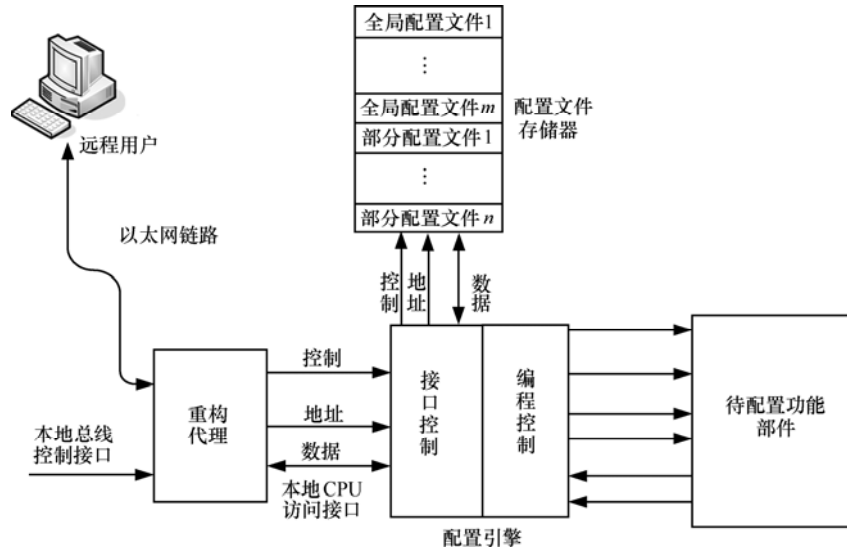


图 5 电路重构控制机制实现

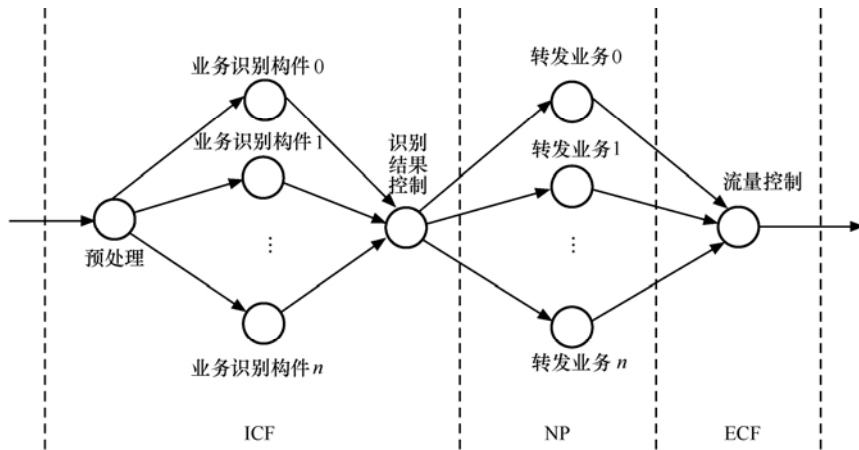


图 6 业务流量识别与控制组件

选择加载对应的部分可重构模块配置文件。重构控制机制具体实现可参见文献[19]。

### 5 应用实例与测试实验测试

#### 5.1 典型应用—业务流量识别与控制组件

业务流量识别与控制组件是 863 “开放式构件化通用路由器研制” 课题中用于支撑电信网、广电网以及互联网三网融合传输业务及其演进的关键部件。业务流量识别与控制组件用于通过深度报文检测 (DPI) 技术识别广电 TS over UDP、电信 RTP over UDP 等多种业务, 并将不同的业务映射到不同优先级上, 实现面向业务的流量管理和服务质量保证。课题基于可重构路由器报文转发引擎构建了支持多种业务识别构件的运行验证环境, 在此基础上完成相应的时序仿真、功能验证和性能评估。

如图 6 所示, 业务流量识别与控制组件主要包

括预处理构件、业务识别构件、识别结果控制构件、转发业务构件和流量控制构件等。

1) 预处理构件: 主要负责从网络接口接收输入报文流、进行报文缓冲、链路层识别及归一化处理等工作。具体包括: 识别以太网或 POS 封装的 MPLS、IPv4 及 IPv6 报文; 剥离链路层头部, IP 报文送识别构件, 非 IP 报文对应的报文处理控制信号直接送识别结果控制模块。

2) 业务识别构件: 进行广电 TS over UDP、电信 RTP over UDP、IPv4、IPv6 等业务协议识别。各构件同时接收报文数据, 并行进行协议业务识别工作, 匹配结果同时送识别结果控制模块。

3) 识别结果控制构件: 负责接收预处理模块输出的报文控制信息及各识别构件的匹配结果信息, 将识别结果与报文统一送 NP 转发处理线程。

4) 转发业务构件: 负责根据输入报文和识别结

果构建查表关键字，进行路由查表，并将查表转发结果与报文一起发送到流量控制构件。

5) 流量控制构件：采用基于令牌桶的流量管理机制，根据预设的流量优先级控制策略和报文识别结果，确定报文的控制处理规则。

业务流量识别与控制组件的预处理、业务识别以及识别结果控制构件映射到 ICF 构件运行环境，转发业务构件分别映射到 NP 程序构件运行环境中多个线程，流量控制构件映射到 ECF 构件运行环境。

本文采用基于 Spirent AX/4000 测试仪，对业务流量识别与控制组件的性能进行了测试。测试仪 TX 端发送测试流量 10Gbit/s 至报文转发引擎，报文转发引擎实现业务流量识别，并按设定流量进行速率控制，转发到测试仪的输入 RX 端口。表 1 给出了对广电业务进行流量识别控制的性能测试结果。

表 1 业务流量识别与控制组件性能

设定流量/(Mbit·s <sup>-1</sup> )	实际流量/(Mbit·s <sup>-1</sup> )	误差/%
2	2.06	3
4	4.11	2.75
8	8.18	2.25
16	16.41	2.56
32	32.81	2.53
64	65.63	2.55
128	131.3	2.58
256	262.53	2.55
512	525.1	2.56
1 024	1 050.19	2.56
2 048	2 100.41	2.56
4 096	4 200.83	2.56
8 192	8 401.75	2.56

由表可知，该功能组件可以有效映射到报文转发引擎程序/电路构件环境，实现不同业务流量高效识别，以及基于设定流量速率的精确流量控制，误差可控制在 3% 以内。

表 2 给出了 ICF 中实现主要业务识别构件的资源占用情况。

5.2 性能比较

基于上述典型应用，本文对可重构路由器报文转发引擎在多种条件下的转发性能进行测试。测试条件包括如下。

1) 单板转发性能测试：主要测试报文转发引擎单板全 10GE 端口全双工报文转发性能；

表 2 业务识别构件资源占用情况

UM 功能	资源名称	使用数量	占用比例
ICF	查找表 LUT	3 110/99 280	3%
	寄存器	639/99 280	<1%
TS over UDP 业务识别构件	存储器位	832/6 727 680	<1%
	查找表 LUT	68/99 280	<1%
RTP over UDP 业务识别构件	寄存器	22/99 280	<1%
	存储器位	0/6 727 680	0%
业务高优先级 识别构件	查找表 LUT	30/99 280	<1%
	寄存器	8/99 280	<1%
	存储器位	0/6 727 680	0%

2) 跨板性能测试：主要测试报文转发引擎板间全双工报文转发性能；

3) 大路由表性能测试：主要测试报文转发引擎在 100K FIB 表项下全双工报文转发性能；

4) 路由振荡条件性能测试：主要测试报文转发引擎在 100K FIB 表项、10K 路由振荡下全双工报文转发性能。

图 7 给出了上述多种测试条件下（测试时间为 30s）报文转发引擎转发延迟性能测试结果。由图可知，报文转发延迟基本保持在微秒级别，与报文长度不存在线性关系。在路由震荡条件下，报文转发延迟相对较大，原因在于：一方面路由表频繁更新会造成转发线程对路由表的访问冲突增多、延迟增大；另一方面，路由表震荡导致的计算开销也会对网络处理器线程处理性能带来负面影响。

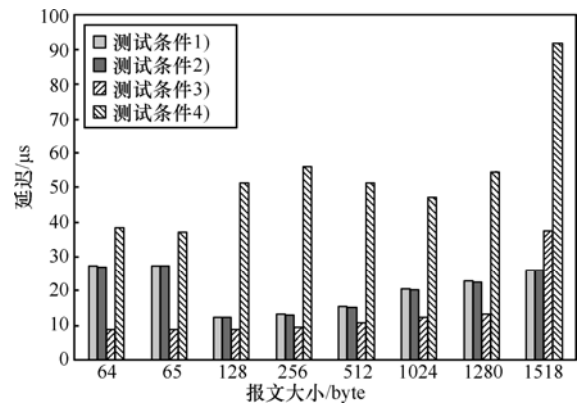


图 7 报文转发引擎性能测试结果

此外,在上述测试条件下测试结果表明,可重构路由器报文转发引擎对于不同报文长度都可获得 100%的转发吞吐率,即保证报文线速转发,由此可见,可重构路由器报文转发引擎能够在提供构件化可重构能力的基础上,有效实现报文线速转发和低延迟处理,实现性能与灵活性的良好折中。

### 5.3 可重构路由器的示范应用

为支持面向三网融合的多种新型网络体制和业务体系的实验和验证,并为各种先进网络架构及网络新技术研究提供有力支撑平台,“新一代高可信网络”项目组于 2011 年在上海构建了一个可重构柔性实验床。本文研制的可重构路由器系统与其他合作单位的研制设备一起,在该实验床上进行了 IPv4、IPv6、广电网络、电信网络等多种网络场景下,分组业务承载能力、服务承载重构能力、网络重构能力以及网络传输性能等方面的测试。测试结果表明,该可重构路由器系统,按照平台化支撑下的构件化实现思想,基于数据/控制平面分离的原则,具有开放性、构件化和可重构等特点,其基于 Pass-Through 模式设计实现的软硬件协同的异构可重构路由转发引擎可以在保证高性能报文处理的前提下,获得高度灵活性,能有效支持新型协议转发模块的快速开发及部署,有效支撑可重构柔性实验床的各种实验。

## 6 结束语

随着多样化网络业务和应用的迅猛发展,路由器处理功能的时空变化特性更趋显著,传统基于网络处理器实现的路由转发引擎难以有效满足程序/电路功能灵活部署和重构。为此,本文提出了可有效支持网络业务重构的路由转发引擎构件重构模型,在此基础上,设计了基于 Pass-Through 模式的程序/电路构件运行环境,通过将可重构 FPGA 器件引入到报文数据通路处理,有效降低程序/电路构件间控制及数据开销,为构件映射提供更广阔的设计选择空间。系统实现和测试结果表明,可重构报文转发引擎可以在满足报文高速处理的前提下,为网络业务提供优化映射环境,满足多样化业务对性能和灵活性的要求。下一步工作主要研究可重构路由

器报文转发引擎对网络虚拟化以及 OpenFlow 技术的支持。

### 参考文献:

- [1] 李玉峰,邱茵,兰巨龙.可重构路由器研究的现状与展望[J].中国工程科学,2008,10(7):82-89.  
LI Y F, QIU H, LAN J L. Status quo and outlook of reconfigurable research[J]. Engineering Science, 2008, 10(7):82-89.
- [2] HU C C, WU C, X W, *et al.* On the design of green reconfigurable router toward energy efficient internet[J]. IEEE Communications Magazine, 2011, 49(6):83-87.
- [3] 龚正虎,傅彬,卢泽新.软件集群路由器体系结构研究[J].国防科学技术大学学报,2006,28(3):40-43.  
GONG Z H, FU B, LU Z X. Research on the architecture of software based cluster routers[J]. Journal of National University of Dependence Technology, 2006, 28(3): 40-43.
- [4] LEE D C, MIDKIFF S F. Reconfigurable routers: a new paradigm for switching device architecture[EB/OL]. <http://www.ccm.ece.vt.edu/papers/>, 1998.
- [5] KOHLER E, MORRIS R, CHEN B, *et al.* The click modular router[A]. Proceedings of the ACM Symposium on Operating Systems Principles (SOSP)[C]. 1999.217-231.
- [6] DAN D, ZUBIN D, GURU P, *et al.* Router plugins: a software architecture for next generation routers[A]. Proc of ACM SIGCOMM[C]. 1998. 229-240.
- [7] HAN S, JANG K, PARK K, *et al.* Packet shader: a GPU-accelerated software router[A]. ACM SIGCOMM[C]. New Delhi, India, 2010. 195-206.
- [8] LEE D, HARPER S, ATHANAS P, *et al.* A stream-based reconfigurable router prototype[A]. IEEE International Conference on Communications[C]. Vancouver, BC, 1999. 581-585.
- [9] LOCKWOOD J, NAUFEL N, TURNER J, *et al.* Reprogrammable network packet processing on the field programmable port extender (FPX)[A]. ACM International Symposium on Field Programmable Gate Arrays (FPGA)[C]. Monterey, CA, USA, 2001. 87-93.
- [10] NAOUS J, GIBB G, BOLOUKI S, *et al.* NetFPGA: reusable router architecture for experimental research[A]. Proceedings of the ACM Workshop on Programmable Routers for Extensible Services of Tomorrow (PRESTO '08)[C]. New York, NY, USA, 2008. 1-7.
- [11] ANWER B, TARIQ M, MOTIWALA M, *et al.* Switch blade: a platform for rapid deployment of network protocols on programmable hardware[A]. Proceedings of the ACM SIGCOMM 2010[C]. New Delhi, India, 2010. 183-194.
- [12] KURT K, SHARAD M, RICHARD A. From ASIC to ASIP: the next design discontinuity[A]. International Conference on Computer

Design[C]. 2002. 84-90.

- [13] LIU D, HUA B, HU X, *et al.* High-performance packet classification algorithm for many-core and multithreaded network processor[A]. Proc of the 6th IEEE International Conference on CASES[C]. 2006.
- [14] FRANKE H, XENIDIS J, BASSO C, *et al.* Introduction to the wire-speed processor and architecture[J]. IBM Journal of Research and Development, 2010, 54(1): 27-37.
- [15] SEDA O, MEMIK G, WILLIAM H, *et al.* Design and analysis of a layer seven network processor accelerator using reconfigurable logic[A]. IEEE Symposium on Field-programmable Custom Computing Machines[C]. Napa Valley, CA, 2002.
- [16] ALBRECHT C, FOAG J, KOCH R, *et al.* DynaCORE – a dynamically reconfigurable coprocessor architecture for network processors[A]. Proc of the Euromicro Conference on Parallel, Distributed and Network-centric Processing[C]. 2006.
- [17] BHUGRA H, CHAPMAN D. Look-Aside (LA-1) Interface Implementation Agreement[S]. 2004.
- [18] NetLogic Corp. XLR 732 datasheet [EB/OL]. <http://www.netlogicmicro.com/Products/MultiCore/XLR.asp>.
- [19] 陈一骄, 卢泽新, 孙志刚. 基于 FPGA 的可重构硬件实现技术研究[J]. 信息工程大学学报, 2009, 10(1):94-97.  
CHEN Y J, LU Z X, SUN Z G. Implementation research of reconfigurable hardware based-on FPGA[J]. Journal of Information Engineering University, 2009, 10(1): 94-97.

#### 作者简介:



**陈一骄** (1972-), 男, 湖南益阳人, 博士, 国防科学技术大学副研究员、硕士生导师, 主要研究方向为高性能可重构路由器和网络处理硬件加速技术。



**卢泽新** (1963-), 男, 湖北黄冈人, 硕士, 国防科学技术大学研究员、硕士生导师, 主要研究方向为计算机网络体系结构、巨型机网络系统和网络协议软件。



**孙志刚** (1973-), 男, 江苏连云港人, 博士, 国防科学技术大学研究员、硕士生导师、主要研究方向为计算机网络与通信、交换技术与调度算法。



**李韬** (1983-), 男, 安徽萧县人, 博士, 国防科学技术大学助理研究员, 主要研究方向为网络处理器、路由与交换技术。



**吕高锋** (1980-), 男, 陕西扶风人, 博士, 国防科学技术大学助理研究员, 主要研究方向为数据中心、高性能路由与交换技术。