

# 基于构件的层次化可重构网络构建及重构方法

刘 强<sup>1)</sup> 汪斌强<sup>1)</sup> 徐 恪<sup>2)</sup>

<sup>1)</sup>(国家数字交换工程技术研究中心 郑州 450002)

<sup>2)</sup>(清华大学计算机科学与技术系 北京 100084)

**摘 要** 下一代网络为了保持网络发展的一致性和连贯性、融合现有底层网络通信基础设施的异构性、能够保证支撑各个领域全面信息化及保证整个网络的可靠可控可信做了很多尝试。作者通过对现有网络及下一代网络发展方向的分析,提出了基于构件的层次化可重构网络体系结构及松耦合、兼容融合、扩展和隔离的构建原则。作者还提出了基于进程代数的重构原则、模型、功能实体合作方法,并提出了具体实现方案,保证了构建和重构过程中的自主性、反应性和主动性。通过搭建实验网络测试,证明了可重构网络具有良好的灵活性和扩展性。

**关键词** 可重构网络;层次化;重构;构件;模型

中图法分类号 TP393 DOI号: 10.3724/SP.J.1016.2010.01557

## Construction and Reconfiguration Scheme of the Hierarchical Reconfiguration Network Based on the Components

LIU Qiang<sup>1)</sup> WANG Bin-Qiang<sup>1)</sup> XU Ke<sup>2)</sup>

<sup>1)</sup>(National Digital Switching System Engineering and Technological R&D Center, Zhengzhou 450002)

<sup>2)</sup>(Department of Computer Science and Technology, Tsinghua University, Beijing 100084)

**Abstract** In order to keep consistency and coherence of network's development, the next-generation network does a lot of attempts to integrate the underlying network heterogeneity of existing communication infrastructure, support information for all areas, and ensure network management, controllability, and reliability. The paper analyzes the development direction of existing network and the next-generation network, and proposes the construction principle of the component based hierarchical reconfigurable networks. The reconfiguration principle, model, function entity cooperate scheme, and the specific implement scheme are introduced so that the autonomy, reactivity and initiative in the construct and reconfiguration process is ensured. Experiments prove the flexibility and scalability of the reconfigurable network.

**Keywords** reconfigurable network; hierarchical; reconfiguration; components; model

## 1 引 言

下一代网络已成为当前研究热点<sup>[1-3]</sup>,学术界对于下一代网络的普遍观点是:新网络的本质应该是

保持网络发展的一致性和连贯性,在融合现有底层网络通信基础设施的异构性的基础上,充分考虑融合未知网络的、能够支撑各个领域全面信息化的、可靠可控可信的信息基础设施。同时近年来国内外各研究机构以业务服务为出发点展开的网络<sup>[4]</sup>、Overlay<sup>[5]</sup>

收稿日期:2010-04-26;最终修改稿收到日期:2010-08-08。本课题得到国家“八六三”高技术研究发展计划项目基金(2008AA01A323)、国家“九七三”重点基础研究发展规划项目基金(2007CB307102)资助。刘 强,男,1982年生,博士研究生,主要研究方向为宽带信息网络和核心路由器。E-mail: lq\_strong@163.com。汪斌强,男,1963年生,博士,教授,博士生导师,主要研究领域为宽带信息网络和核心路由器。徐 恪,男,1974年生,博士,教授,博士生导师,主要研究领域为下一代互联网、交换和路由结构、P2P网络、overlay网络。

等研究也为下一代网络的研究和变革提供了新思路。因此在研究下一代信息网络体系结构时,应该将提供高性能服务作为下一代信息网络的核⼼目标和研究出发点。目前国内由于政策的松动,互联网、广电网、电信网等信息网络融合势在必行。国外各研究机构越来越将计算和通信融合为一个方向。所以新的网络体系结构也要尽可能地将原有信息网络和未知信息基础设施融合进来。

在这些思想和认识的指引下,本文提出了一种以构件作为网络中的最小的物理资源,通过对不同服务的支持,装配或释放相应的构件,来构建网络的新型网络体系——基于构件的层次化的可重构网络。可重构网络面向服务提供基于承载服务,既可以实现控制功能和转发、传送功能松耦合,使得信息网络易于控制和管理,又可以实现业务与网络松耦合,当某种业务出现时,根据其业务特性提供相应服务能力。一旦业务消失,释放对应服务能力和使用的构件。该技术强调使用而不拥有,利于松耦合,并且能够屏蔽资源的异构性,扩展性强。同时通过构件标准化的引入,第三方厂家可以生产出大量标准构件,降低了交换节点的技术门槛,打破了行业垄断。同时可以随时引入新的协议构件构造新型网络,保证了国家的信息安全。

本文第 2 节介绍相关工作;第 3 节给出可重构网络构建原则和四层模型;第 4 节用进程代数的方法描述可重构网络构建和重构过程;第 5 节给出可重构网络具体构建和重构过程,并给出相应实验结果;第 6 节总结全文。

## 2 相关工作

目前人们对下一代网络的研究主要集中在对网络的灵活扩展性、安全性、可控性等方面的研究。在灵活扩展性方面,人们相继开展了主动网<sup>[6]</sup>、可编程网<sup>[7]</sup>、虚拟网络<sup>[8-9]</sup>等方面的研究工作,试图通过节点智能化增强网络的多业务处理能力。其中主动网的研究增强了网络业务适应能力,但它不区分端系统和中间系统,将互联网中所有结点都改造成主动结点,破坏了传统互联网的透明性原则。可编程网通过改变网络中的节点设备的方法来对网络进行改造,但现有研究仅限于框架模型和实现原型这两个方面,缺乏完善、系统的体系架构理论指导。虚拟网络通过聚合多个节点简化了拓扑复杂度,路由系统、流量工程、面向虚拟网嵌入的资源管理等的复杂性

也相应降低。但是现有虚拟网络对网络节点中资源的描述、聚合及分配机制,虚拟物理节点间信息交互的开放通信机制以及多虚拟路由器之间的隔离机制等的研究还有待加强。

在下一代网络的安全性和可信、可控方面的研究主要包括: Intel 研究中心的 Yumerefendi 和 Chase 认为下一代信息网络是可审计的,下一代信息网络的系统行为和状态应该是不可否认和防篡改的。但是可审计网络仅仅关注于网络的几个方面,对全网的规划没有考虑。可信网络<sup>[10]</sup>认为一个可信的网络应该是网络系统的行为及其结果是可以预期的,能够做到行为状态可监测,行为结果可评估,异常行为可控制。但是可信网络研究中许多革命性的成果(如抵御 DDoS 的网络体系结构等)与现有互联网体系 IP 层是不可动摇的基础这个特点相冲突,难以得到实质应用。各种可信机制只能通过补丁的形式在网络中实现,难以形成统一的顶层设计。

为更好地研究下一代网络,国外一些研究机构搭建了实验床。2005 年 8 月, NSF 宣布了 GENI<sup>[11]</sup> 计划。NSF 希望通过 GENI 构建全新的、安全的、能够连接所有设备的互联网,从而促进互联网发展,刺激创新和经济增长。其核⼼目标是改变现有网络和分布系统设计的方式。2008 年 9 月, GPO 发布了 GENI 系统总体框架,并同时发布了 GENI 第 1 阶段的研究概要“GENI Spiral 1”。但 GENI 项目对下一代信息网络的研 究还只是停留在决策层面或者起步阶段,并未形成清晰的理论研究方案,某些研究项目也只是从下一代信息网络的某一个或某几个方面展开研究,缺乏对下一代信息网络体系结构及关键理论与技术的全面性和系统性研究。

源于美国大学校园的 PlanetLab<sup>[12]</sup> 实验床针对未来互联网技术和服 务进行研究、实验的开放式、全球性的实验床平台。PlanetLab 的核⼼思想是资源虚拟应用和分布式重叠网络。其网络节点 (PlanetLab 节点) 都采用修改后的 Linux 操作系统,安装 PlanetLab 软件包。PlanetLab 节点资源被分成许多切片,不同节点上的切片组成一个虚拟网的分片,构成一个通过 Internet 连接的虚拟网。这些节点分属不同的组织、机构,分布在不同的国家,但都由 PLC (PlanetLab Central) 进行管理;节点资源分片构成的虚拟网就可以为用户提供特定的网络服务,整个平台构成一个重叠网实验床。但是它存在缺乏灵活性的局限:机器分片是基于集中容器的资源隔离。一些关键资源是共享的。比如单一内核、网络

协议栈和组件(NAT、Firewall等)、网络资源(IP地址和端口)、VNet(虚拟网络设备)单一地址等等。无法任意选择Linux系统内核/网络协议栈和组件。

通过对现有下一代网络的研究,我们发现现有研究重点仅仅在网络的灵活扩展性、安全性、可控性等某单一方面,很少有计划和方案将这几方面都兼顾到。所以我们在进行下一代网络设计中必须兼顾到面向多样的服务模型和复杂的应用需求,既要能屏蔽现在底层网络通信基础设施的异构性,又必须提供一个统一、开放、安全、可扩展、可管理和高性能的服务平台,以支持快速灵活地开发、集成、部署和定制各种网络应用。

### 3 可重构网络体系结构

#### 3.1 概念

**定义 1.** 重构指系统从一种构形<sup>[13]</sup>向另一种构形的转换。重构又不同于重组,重构指重新构造系统的结构及重新组合系统的功能,需要从系统外引入新的构件,或从系统中移出已有构件,或用一个构件替换另一个构件(物理重构),或保持已有系统构件不变而改变组合方法(逻辑重构)。

**定义 2.** 可重构目标(Reconfiguration Goal, RG). 在需求发生变化时系统中各个功能实体希望通过重构而进入的状态。

可重构目标包括两部分:(1)功能目标,即可重构系统具有何种功能,重构过程结束后各功能实体应完成何种工作,功能实体应达到何种状态;(2)性能目标,即功能实体应在特定的性能指标下合作完成工作,如实体必须在特定性能限制下实现功能目标。在衡量整个可重构网络可重构目标实现的过程中,主要有两个性能目标,即时间和系统资源。

(1)时间(Goal Time, GT). 对实现功能目标的时间限制,即必须在多少时间单位内完成。它与可重构目标RG的关系为

$$GT \cap Goal \Rightarrow (\neg GT \rightarrow \neg RG).$$

(2)系统资源(Goal Resource, GR). 在实现重构目标的过程中各功能实体会消耗一定的系统资源,如CPU、内存、存储等。但由于资源是有限的,所以目标的实现不能以占用或消耗过多的资源为代价。它与可重构目标的关系为

$$GR \cap Goal \Rightarrow (\neg GR \rightarrow \neg RG).$$

为讨论上的方便,谈到资源限制时,是指对同一

种资源的限制;如果系统中需要对多种资源进行限制,那只需增加相应的限制条件即可。

**定义 3.** 可重构网络指具备在业务需求变化时,根据重构目标在网络中引入或移除构件和功能实体,并对其进行相应的配置以在一定资源消耗条件下达到重构目标能力的网络。

#### 3.2 可重构网络体系结构的设计原则

根据可重构网络的定义,结合下一代网络发展的趋势,我们认为在构造满足信息网络发展需求的可重构网络体系结构模型时,一方面既要保持传统互联网开放、简单、灵活和健壮的优点,另一方面还应该在下述设计原则的指导下进行:

(1)松耦合原则。可重构网络中用户业务和网络服务关系是松耦合;构建网络不再依据特定用户业务需求提供网络服务,而是依据网络服务提供能力来进行服务。

由于可重构网络的松耦合性,一种业务在物理实现上就是一个或多个构件通过适当装配而完成。对于整个可重构网络来说就是在业务需要的时候对构件进行相应的装配来组成一个网络。可重构网络将整个网络的最小资源定义为构件,在构件进行标准化后,每个第三方厂家就可以按照构件标准仅仅开发单独的构件,对整个链路设备和交换平台设备不用做过多考虑,大大降低了行业进入门槛,打破了思科等厂家的行业垄断。同时可以通过装配不同的协议构件(非TCP/IP协议构件)组成的网络保证战时的网络安全。

(2)兼容融合原则。可重构网络体现了多种网络系统、多种硬件平台、多种操作系统相互融合的发展趋势,其目标是实现各种开放、异构资源支持下的服务融合、网络互连和优势互补。可重构网络通过可重构网络管理平台,依据网络服务提供能力构建可重构网络,每一种服务可以支撑包括未来新兴的多种特性相似的用户业务,使得可重构网络可良好支持多种现有业务的同时,也可以很方便地支持未来的新型业务,实现了新业务时对网络设备影响最小的目的,使得网络技术和业务技术的独立发展成为可能。

(3)可扩展性原则。可重构网络在兼顾了现有传统网络体系结构的基础上对未知网络也具有较强的可扩展性,对于传统网络中的链路和交换节点进行相应的改造,将其中的各功能单元改造为构件。可重构网络的服务提供能力表现为网络能够提供和去除适应网络服务的资源,即增加和删除个别构件单

元. 因此, 网络需要改造的几率大幅度下降, 网络对于用户业务支撑的鲁棒性显著提高, 可扩展性增强.

(4) 隔离原则. 整个可重构网络由于支持业务的不同而需要构建不同的逻辑承载网络. 为保证可重构网络对每种服务支持的鲁棒性和高效性, 在网络体系结构设计过程中, 选用适宜的技术、方法和机制来分隔各个网络, 尽量保证将不同网络的功能及不同的业务、服务种类从逻辑上进行分离, 从实现上做到解耦.

### 3.3 可重构网络体系结构四层模型

基于 3.2 节中的设计原则, 借鉴了虚拟化网络、可编程网、叠加网、模块化路由器等网络体系结构和路由节点结构的研究思路, 吸纳了近年来软件工程领域所取得的若干研究成果的思想、观点和方法, 我们提出了一种结构分层、面向服务的网络体系结

构——可重构网络体系结构模型. 可重构网络模型具有一定的层次性, 每层和每层之间是松耦合的关系. 下面一层为上面一层提供资源, 上面一层通过命令指导下面一层进行相应的重构. 在设计过程中, 每个层次都有自己独立的管理功能单元, 负责对上层的命令的解析和对本层的管理和控制, 并对下层发送相关的命令. 每一层都是利用网络承载技术, 为上层业务提供承载服务的网络元素的集合. 由于在可重构网络的交换节点内部构件与构件之间也是连接成网的, 所以我们将构件网也作为一层. 可重构网络在逻辑上分为 4 层: 第 1 层为可重构网络管理平台, 第 2 层为嵌入式承载网, 第 3 层为可重构物理网络, 第 4 层为可重构路由平台内部构件系统.

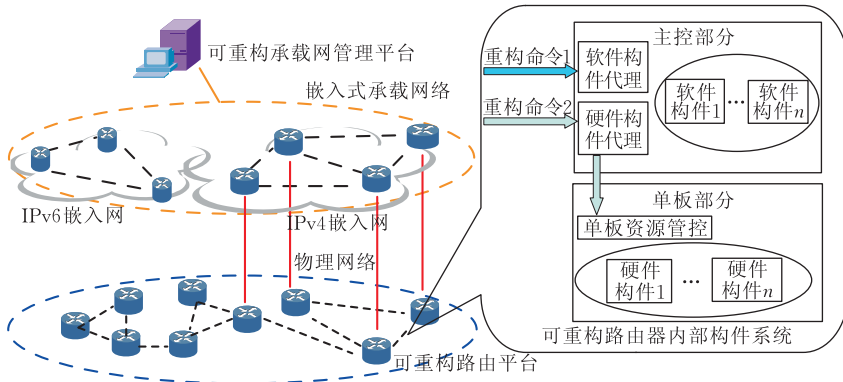


图 1 支持业务重构的可重构网络

在此我们对每层进行定义:

第 1 层可重构网络管理平台主要负责管理整个可重构网络构建和重构方式, 根据用户或网管的命令来确定对整个承载网络应该构建为几个服务提供网络, 每个服务提供网络具体运行何种业务, 每种业务具体需要多少资源来支撑进行规划和目标分解. 可重构网络管理平台主要包括综合管理系统和可重构路由交换平台构件管理系统两部分.

第 2 层嵌入式逻辑承载子网是根据可重构网络管理平台的指令在可重构物理网络的基础上抽象出来的, 具有特定业务特性的虚拟网络. 每个网络和网络之间逻辑上不具有互通性, 但是由于同是在一个物理网络上进行复用可能在具体实现过程中具有一定的互通可能.

第 3 层可重构物理网络在可重构网络物理实现过程中有多少光链路和交换节点设备组成一个大的物理网络, 网络中各光链路和可重构交换节点设备之间如何物理连接.

第 4 层可重构路由平台内部构件网络将原有传统路由器、交换机等交换设备中的各个功能模块(协议栈、路由协议、单板软件等)设计为单独构件. 在接收到上层的重构命令后通过各个构件的灵活装配, 构成支持电信、广电、互联网等网络的可重构路由转发平台. 构件建模方法应被标准化, 以使据此构建的构件模型库是可被第三方扩展的, 达到构件模型库持续可扩充、升级, 满足不断发展的网络服务对新构件的需求. 通过新的构件的加入, 装配成为可满足各种需求的可重构交换平台.

## 4 可重构网络构建及业务重构方法研究

本节我们讨论对可重构网络如何进行重构. 由第 3 节可知可重构网络构建和重构过程主要工作是在可重构管理平台的控制下完成大小两张拓扑网络的构建和重构. 一是根据业务需求搭建可重构嵌入

式逻辑承载子网, 这个网络可能是广电网、电信网、互联网或是一种未知网络, 嵌入式逻辑网络中映射并包含了物理网络中各可重构路由平台节点和光链路等组成的网络设备, 各个网络设备进行互联组成一个拓扑网. 一个是可重构路由平台内部为适应可重构网络业务的差异性和多样性, 满足多样业务需求而选取不同业务构件组成的构件拓扑网. 构件间根据数据流形成相应的拓扑关系. 可重构路由器重构系统将重构目标或需求分解为一系列子目标并指派给这两个拓扑网中的功能实体, 系统目标与功能实体之间形成“目标-指派”关系.

在整个可重构网络构建和重构过程中, 从工程实现的角度看, 其过程为: (1) 产生重构需求, 确定重构目标, (2) 分解重构目标、求解重构结构, (3) 寻求重构功能实体候选集; (4) 选择并调整重构方案; (5) 实现目标. 各功能实体之间是一种“按需设岗, 选择上岗, 调整完成”的模式.

**定义 4.** 可重构模型(Reconfiguration Model, RM)定义为  $M = \langle RG, RTS, PC, RS \rangle$ . 其中:

(1)  $RG$  表示重构目标. 在用户及业务需求的基础上对可重构网络将重构为何种样式具有何种功能进行形式化描述.

(2)  $RTS$  表示重构拓扑结构. 构造出重构过程的目标之间的拓扑结构. 确定完成重构目标的分解, 确定各个子目标的相关性质以及各子目标之间的依赖关系.

(3)  $PC$  表示可重构功能实体候选集合. 即参与重构过程的各功能实体及管理实体的集合. 可重构功能实体候选集合的形式是管理实体根据用户和业务需要对重构目标进行解析, 并对各实体竞争相应地重构合作角色的有关信息而确定的.

(4)  $RS$  表示重构方案. 方案描述了各功能实体间如何进行构建和重构可重构网络的过程, 方案中包括根据网络实际情况进行适当的调整的过程.

#### 4.1 可重构功能实体拓扑结构

在整个可重构网络的重构过程中, 以实现用户及业务需求为目标建立拓扑结构, 即求解重构过程中各功能实体重构方案, 确定完成合作任务所需的角色, 角色的相关性质以及各角色之间的依赖关系. 拓扑结构主要用来规定如何将可重构目标进行逐步细化, 以便将各个子目标分派给参与合作的实体, 形成一个树. 目标细化是在系统的基本构成单位是子目标和功能实体的框架下, 实现并发系统层次化的一种操作. 其基本思想是将系统高层次的基本单位

替换为低层次进程, 然后在该层次上选定一个或几个抽象的未加解释的进程, 并分别用低一层次的具体的活动进行替换, 逐层做下去, 直至得到详尽的系统设计或实现. 在本文是利用进程代数的方法将重构模型形式化. 在细化过程中遵循 3 个性质:

(1) 可重构网络重构系统中细化函数将重构目标细化成子目标系统描述;

(2) 所有的目标都有可能细化;

(3) 在总目标和功能实体进程之间的所有原因和矛盾关系被细化后具有继承性.

**定义 5.** 重构子目标事件结构  $\eta$  是一个四元组  $\langle g, <, \delta, \beta \rangle$ , 其中  $g$  是目标事件集合;  $< \subseteq g \times g$  是非自反的偏序关系, 并满足“有限原因规则”:  $\forall e \in g: \{d \in g \mid d < e\}$  是有限的;  $\delta \subseteq g \times g$  是非自反的有限的矛盾关系, 并满足“矛盾继承规则”:  $\forall d, e, f \in g: d < e \wedge d \delta f \Rightarrow e \delta f$   $\forall e \in g < \subseteq g \times g$ ;  $\beta: g \rightarrow RG$  是目标分解标记函数.

**定义 6.** 重构目标细化函数  $Refine$ . 当且仅当  $\forall a \in RG: refine(a)$  是非空的、有限的和无矛盾的. 设目标事件结构  $\eta \in g$ , 那么  $refine(\eta)$  是一个事件结构, 其定义如下:

$$g_{refine(\eta)} = \{ (e, e') \mid e \in g_\eta, e' \in g_{refine(\beta_\eta(e))} \};$$

$$(d, d') < (e, e')_{refine(\eta)} \text{ 当且仅当 } d < e_\eta$$

$$\text{或者 } d = e \wedge d' < e'_{refine(\beta_\eta(d))}.$$

根据上两个定义可将整个可重构目标进行具体细化.

**定义 7.** 重构子目标集合. 设重构目标  $RG$  及子目标集合  $SG = \{RG_1, RG_2, \dots, RG_n\}$ . 若功能实体实现  $SG$  中的任一目标就相当于实现了目标  $RG$ , 则称  $RG_i$  ( $1 \leq i \leq n$ ) 为  $RG$  的“或”子目标, 记作  $SG_{Or}(RG, RG_i)$ ; 若各功能实体实现了  $SG$  中的所有目标才算实现了目标  $RG$ , 则称  $RG_i$  ( $1 \leq i \leq n$ ) 为  $RG$  的“与”子目标, 记作  $SG_{And}(RG, RG_i)$ . 即

$$(1) SG_{Or}(RG, RG_i) \rightarrow \bigcup_{i=1}^n (reach\ RG_i);$$

$$(2) SG_{And}(RG, RG_i) \rightarrow \bigcap_{i=1}^n (reach\ RG_i).$$

**定义 8.** 重构目标树. 由定义 6~8, 重构目标  $RG$  被分解为集合  $\forall_{1 \leq i \leq n} (SG_{And}(RG, RG_i)) \cup \forall_{1 \leq i \leq n} (SG_{Or}(RG, RG_i))$  的形式, 则根据重构目标及其子目标之间的关系可以构成一颗重构目标树:

$$Tree(RG) = \forall_{1 \leq i \leq n} (SG_{And}(RG, RG_i)) \cup \forall_{1 \leq i \leq n} (SG_{Or}(RG, RG_i)).$$

该树具有如下特征:

(1) 任一结点的子结点可能是“与”子树或“或”子

树,一个子树与子树之间可能有重复任务时,将子节点复制,每个子树各用一个子节点,不能形成环形;

(2) 根结点对应着重构总目标,叶结点表示分解出的子目标,整棵树是对重构总目标分解的求解方案;

(3) 其他结点为树的中间结点,它表示那些还需要进一步分解、要多个子目标组合才能完成的子目标;

(4) 每个目标树节点可以分解为更小的目标树。

在可重构网络中各目标实现过程中的合作不仅仅是一个个子目标独立实现各自内容,各子目标任务在重构期间会出现功能实体之间的资源共享或者竞争以及由此引发的交互和通信等,因此要定义出完整的重构子目标分解结构,还必须在目标树的基础上定义子目标与子目标之间的关系。

**定义 9.** 重构顺序图. 由于可重构网络构建和重构过程中的特殊性,重构目标树结构中的共同祖先是“与”的子结点,同一层的各结点之间的关系为 before, serial, parallel, none. before 即目标  $RG_i$  必须在  $RG_j$  “之前”完成. serial 表示  $RG_i$  和  $RG_j$  之间的顺序“无特定的先后顺序,但不能同时进行”, parallel 表示  $RG_i$  和  $RG_j$  之间必须“并行同时完成”, none 表示  $RG_i$  和  $RG_j$  之间的顺序“无特定要求”。

所以,相对于重构过程中总目标和子目标“与/或”树的重构目标树结构. 树的每一层都对应一个独立的目标分解图,图中结点为对应层的所有(子)目标;目标关系图的构造方法为

(1) 若  $T(RG_i, RG_j) = \text{“none”}$ , 则连线为无向虚弧线;

(2) 若  $T(RG_i, RG_j) = \text{“before”}$ , 则连线为实弧线,连线的方向由  $RG_i$  指向  $RG_j$ ;

(3) 若  $T(RG_i, RG_j) = \text{“parallel”}$ , 则连线为双向实弧线;

(4) 若  $T(RG_i, RG_j) = \text{“serial”}$ , 则连线为实弧线,弧线方向由实际决定。

## 4.2 可重构功能实体候选集合

4.1 节拓扑结构中的重构目标分解成一个个子目标. 子目标中包含了在可重构网络体系结构设计中将要达到的任务以及协调规则. 我们需要将重构子目标落实在每个功能实体上,解决“每个岗位由谁来做的问题”. 从而保证整个可重构网络组网过程中的自动化和可管、可控、可信。

**定义 10.** 可重构网络功能实体模型. 可重构网络功能实体模型定义为 5 元组  $RC(I, C, F, O, S)$ .

其中,  $I$  表示该功能实体的输入接口的集合;  $O$  表示该功能实体的输出接口的集合;  $C$  表示该功能实体的触发条件集,处理对象只有满足其触发条件才会被本功能实体处理,否则不会被处理;  $F$  表示功能实体的功能集合;  $S$  表示该功能实体的功能集合描述信息。

**定义 11.**  $\alpha$  映射. 由定义 10 可知可重构网络功能实体模型  $RC_i$  在  $C_i$  在  $I_i$  输入后输出  $O_i$  完成  $F_i$  功能, 记为  $\alpha(RC_i) = F_i$ . 在可重构网络重构过程中要完成  $F_n$  功能, 则需要对  $F_n$  取逆, 获得相应的构件  $\overline{\alpha(F_i)} = RC_i$ .

**定义 12.** 目标与功能实体体系结构. 可重构网络重构目标规范 RRG, 分解为子目标集合需要一组  $SG = \{RG_1, RG_2, \dots, RG_n\}$ , 即, 需要完成子功能集合  $F = \{F_{RG_1}, F_{RG_2}, \dots, F_{RG_n}\}$ , 对每个子功能取  $\overline{\alpha}$  得到实体功能集合  $RC = \{RC_1, RC_2, \dots, RC_n\}$ .

实质上可能在取  $\overline{\alpha(F_i)} = RC_i$  的过程中会得到多个  $RC$  满足, 所以还存在一个选择重构方案的过程, 因为一旦合作伙伴确定后, 整个可重构网络体系结构也就相应地确定了. 为了减少寻找可重构功能实体候选者的盲目性, 提高效率, 有必要先对重构目标分解进行一定的优化处理, 排除那些不可能达到重构的子目标集合。

## 4.3 可重构方案调整

由 3.2 节, 我们可知在构建可重构网络时, 一旦确定功能实体候选集合后就应该有效地利用有限的物理网络资源构建出数量和种类尽可能多的可重构服务提供网络. 为此, 提出两个可重构网络可重构方案调整的基本原则:

(1) 链路复用优先原则. 即在构建嵌入式网时, 应优先考虑所属物理网络中已占用链路, 以保留尽可能多的空闲可重构端口, 从而使能构建的可重构网络种类尽可能多。

(2) 网络链路的负载均衡原则. 即在嵌入式网络中, 应选用负载强度小的链路, 让物理网络负载更趋均衡, 从而使构建的可重构网络数量尽可能多。

在整个可重构网络中, 我们将物理网和构件网络建模成为图  $G_c = (N_c, E_c)$ , 其中  $N_c$  是底层节点集合而  $E_c$  是物理底层链路集合. 在可重构网络第三层上会存在多个嵌入式逻辑承载网络. 第  $i$  个嵌入式网络的拓扑结构逻辑图  $G_i = (N_i^i, E_i^i)$ ,  $N_i^i$  和  $E_i^i$  分别是逻辑节点和逻辑链路的集合。

在整个可重构网络中对某一业务支撑, 主要取决于整个网络对该业务逻辑承载网络分配的资源上. 每个嵌入式网络的资源分配包括下面两个部分:

(1) 节点资源分配. 每个逻辑节点被映射为相对应的一个底层节点.

$$F_N^i: N_l^i \Rightarrow N_C, f_N^i(m_q) \in N_C, \forall m_q \in N_l^i, \{q | q=1, 2, 3, \dots\}.$$

(2) 链路资源分配: 每条逻辑链路被映射为相对应的两个底层节点之间的底层链路.  $F_L^i: E_l^i \Rightarrow E_C$ , 即  $f_L^i(m_q m_p) \in E_C(f_N^i(m_q), f_N^i(m_p))$ ,  $m_q m_p$  在这里是  $E_C$  所有底层路径的集合,  $E_C(s, t)$  是从节点  $s$  到节点  $t$  的底层路径集合.

为了达到网络链路的负载均衡原则和链路复用优先原则. 定义在  $t$  时刻的底层节点负载  $S_N(t, m)$  为分配给这个底层节点的逻辑节点数量. 定义在  $t$  时刻的底层链路负载  $S_L(t, e_j)$  为对应的底层链路包含了底层链路  $e_j$  的逻辑链路的数量,  $M(e_j)$  为底层链路被复用的次数.

由于在可重构网络提供资源上的限制, 在整个可重构网络的物理层每个底层节点只能对应有限个数的逻辑节点. 因此, 维持整个网络节点均衡的负载能够有效地提高网络资源的利用率. 为了使网络负载的更加平衡, 我们定义了节点负载比极限  $L_N$  和链路负载比极限  $L_L$ .

$$L_N = \lim_{\frac{Num(N_C)}{Num(N_l^i)}} \frac{\max_{q=0} S_N(t, m_q)}{[\sum_{q=0} S_N(t, m_q)] / Num(N_C)},$$

$$L_L = \lim_{\frac{Num(E_C)}{Num(E_l^i)}} \frac{\max_{j=0} S_L(t, e_j)}{[\sum_{j=0} S_L(t, e_j)] / Num(E_C)}.$$

通过调整逻辑节点、逻辑链路和实际物理节点和链路的对应关系, 将负载比极限趋近于 1, 可以达到负载均衡原则, 同时保证  $\sum_{i=1}^{M(t, e_j)} e_i / M(t, e_i)$  最大, 做到链路复用率最大, 从而进一步保证了在可能的资源下可重构网络对业务的支持.

## 5 可重构网络构建及业务重构实现

### 5.1 可重构网络构建及重构

在用户或业务需求改变时可重构网络管理平台管理对可重构网络重构这个总目标进行分解. 并利用第 4 节中的方法, 将每个子目标对应的相关网络资源、功能实体和具体嵌入式网络进行重构. 可重构网络管理平台采用集中方式对可重构网络全网进行映射. 可重构网络管理平台负责接受各业务请求及节点的状况映射, 因此可重构网络管理平台维护基础网的最新信息, 并实时对分解的各个子目标及其关系进行检测. 保证足够的功能实体完成每个子目标.

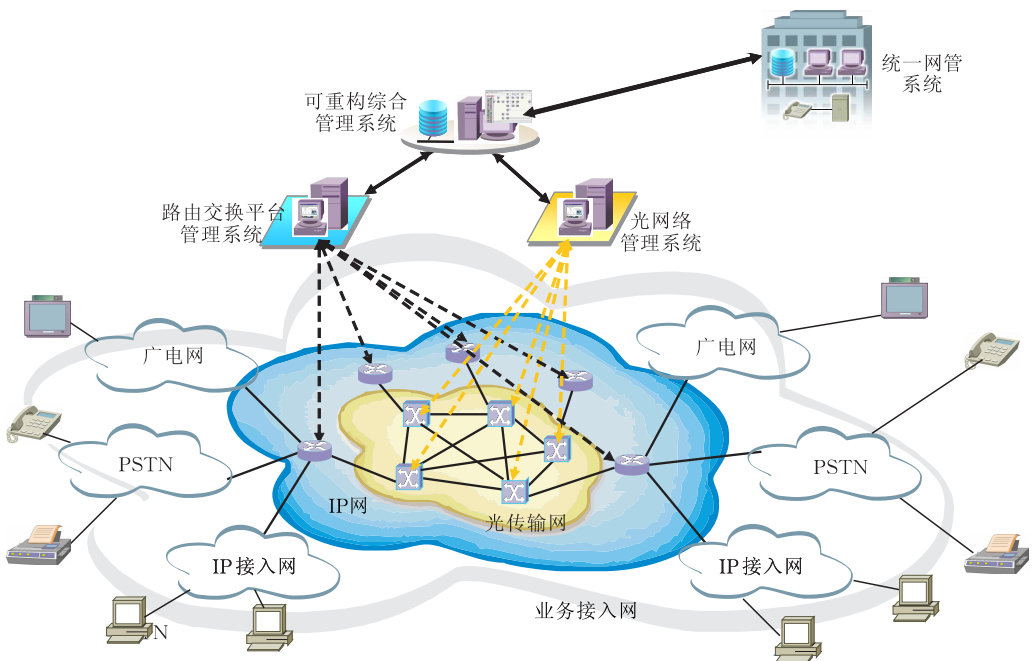


图 2 可重构网络组网方案

在可重构网络中由可重构综合管理系统提供一个北向接口与统一网管系统相连. 由统一网管系统, 将传统网管信息用 SNMP 方式发送给可重构综合管理系统, 再由可重构综合管理系统发送给路由交换平台管理系统, 最后路由交换平台管理系统发送给路由交换设备. 路由交换设备的反馈信息则反向

逐层传递给统一网管系统. 子目标分解后对应的任务和事件(网络构建、网络拆除等)也通过管理系统下达. 图 3、图 4 分别为可重构网络综合管理系统、可重构路由交换平台构件管理系统. 图 5 为可重构网络重构过程中的报文格式.

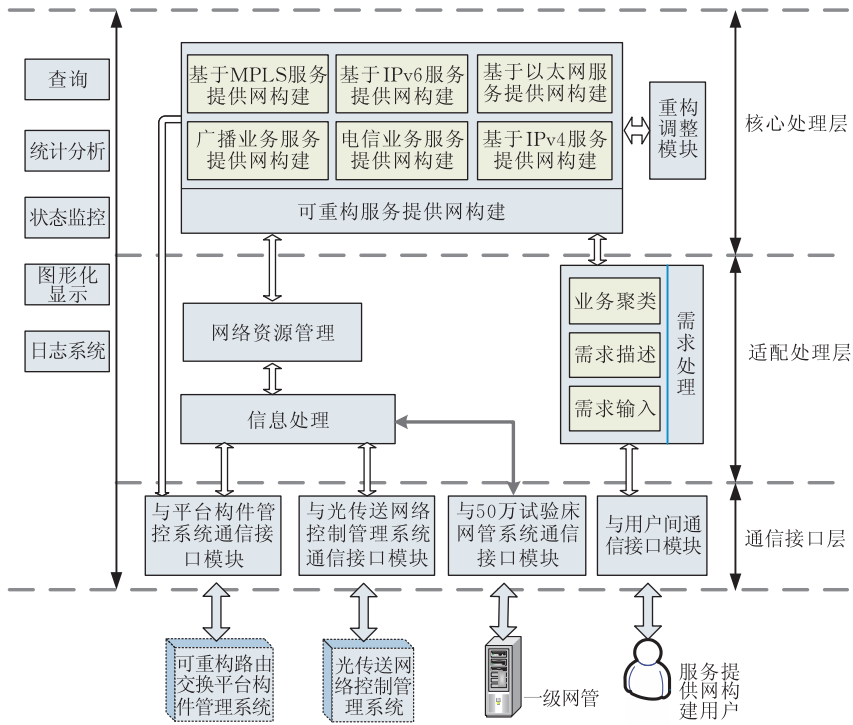


图 3 综合管理系统体系结构

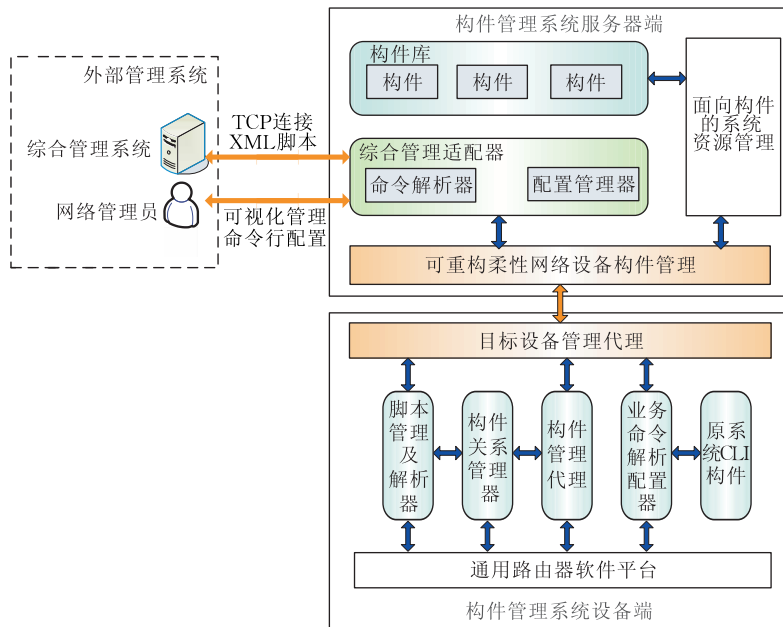


图 4 构件管理系统框架



方向	指令	Type	Value					
			Data (变长)					
N↔N	Hello	1	NodeID	PortID	...	BitField		
N→S	资源感知 (初始化)	2	NodeID	PortCount	PortID	B.W.(2B)	协议	可用协议
			地址(4B/16B)		R.NodeID	R.PortID	Buffer/M(2B)	Delay/μs (2B)
S→N	感知确认	3	NodeID	0=S/F>0				
S→N	网络构建	4	NodeID	LogicID	协议	路由协议	PortCount	PortID
			PathCount	PathID	FromNode	ToNode	SrcIP(4B/16B)	SrcIPMask(~)
			DestIP(4B/16B)		DestIPMask(~)	SrcPortRange(4B)		DestPortRange(4B)
			B.W.(2B)	Buffer(2B)	PortCount,PathCount 二重循环			
N→S	构建反馈	5	NodeID	LogicID	0=S/F>0	PortCount	PortID	PathID
S→N	网络拆除	6	NodeID	LogicID	PortCount	PortID	P.Num=0 全拆	
N→S	拆除反馈	7	NodeID	LogicID	0=S/F>0	PortCount	PortID	
S→N	端口重构	8	NodeID	PortCount	PortID	协议		
N→S	重构反馈	9	NodeID	0=S/F>0	PortCount	PortID		
N→S	统计信息	10	NodeID	PortCount	PortID	Total_Throughput/bps (4B)		DropPackets (4B)
			TotalPackets (4B)		QueueLength (4B)		L.NetCount	LogicID
			PathCount	PathID	Throughput/bps (4B)		Drops (4B)	TotalPackets (4B)
S→N	统计确认	11	NodeID	0=S/F>0	PortCount, L.NetCount, PathCount 三重循环			
N→S	邻接通报	12	NodeID	PortID	Status	Down=0 Up=1 ...		
S→N	邻接反馈	13	NodeID	0=S/F>0				

图 5 可重构报文部分格式

### 5.2 可重构路由交换平台构建及重构

可重构综合管理系统与路由管理平台之间以 Web Service 方式通信,传递构件操作命令、资源感知以及数据统计等信息.路由平台从可重构管理平台获取构件下载、构件加载、构件卸载、资源感知请求指令 4 种信息.可重构管理平台从路由平台获取构件下载/加载/卸载指令的反馈、资源感知信息、端口数据统计、节点邻接状态变更 6 种信息.接收到重构命令后,可重构路由交换平台将功能实体候选集

中的功能实体进行查找确定(附图 1),并将重构命令解析为软件构件重构和硬件构件重构两种命令.前者由路由交换平台主控部分的软件构件代理接收、解析并执行.后者由路由交换平台主控部分的硬件构件代理接收、解析并执行.

可重构路由交换平台业务重构算法如图 6 所示.软件构件的重构发生在主控部分(软件构件重构算法),硬件构件的重构发生在单板部分(硬件构件重构算法).

```

BEGIN
  Receive the reconfiguration command from NCB
  While software
    {Software agent analysis the reconfiguration command
    IF the command is ok THEN
      DO software reconfiguration scheme
      Send the feedback information
    }
  While hardware
    {hardware agent analysis the reconfiguration command
    IF the command is ok THEN
      DO hardware reconfiguration scheme
      Send the feedback information
    }
  END
  
```

图 6 可重构路由交换平台业务重构算法

```

BEGIN
  Receive the reconfiguration command from NCB
  IF command is not ok THEN
    Send warning message
  ELSE
    Download the hardware components scheme description document
  ENDIF
  Get the components information from the description document
  WHILE(the resources of plant is enough)
    {
      Download the hardware components code
      IF (validity checking is not ok) THEN
        Send warning message
      ELSE
        Run the hardware components
        Handover the data to the new components
      ENDIF
    }
  END
  
```

图 7 硬件构件重构算法

```

BEGIN
Receive the reconfiguration command from NCB
IF command is not ok THEN
    Send warning message
ELSE
    Download the software components scheme description
    document
ENDIF
Get the components information from the description
document
WHILE(the resources of plant is enough)
{
    Download the software components code
    IF (validity checking is not ok) THEN
        Send warning message
    ELSE
        Run the software components process
    ENDIF
    Handover the data to the new components
}
END
    
```

图 8 软件构件重构算法

5.3 实验结果

我们按照可重构网络的方法搭建 5 个交换节点的全 mesh 结构网络, 每个交换节点为 CPU AMD3800+2.01GHz、内存 1GB、4 个 USB 网卡的 PC 机器, 可重构网络在可重构网络管理平台的控制下进行重构. 运用直连主机在网络重构后具有 IPv4 和 IPv6 特性的可重构网络中发送 FTP 数据, 在具有广电组播的可重构网络中发送广电数据. 实验结果如图 9~11.

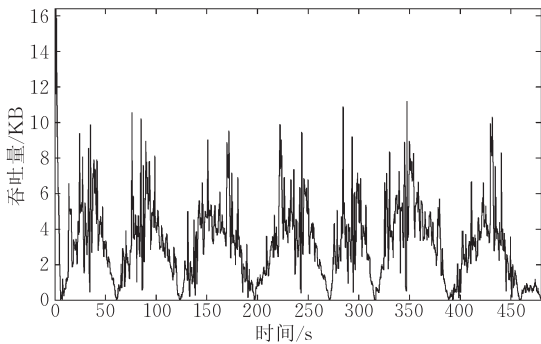


图 9 IPv4 网络中发送小文件 FTP 数据吞吐量示意图

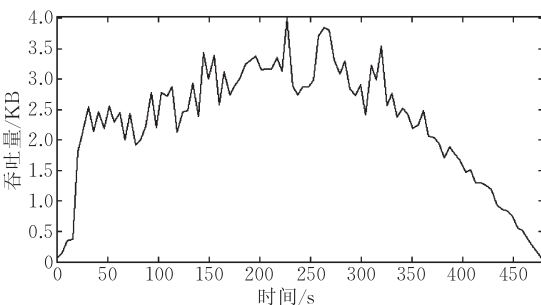


图 10 IPv6 网络中发送大文件 FTP 数据吞吐量示意图

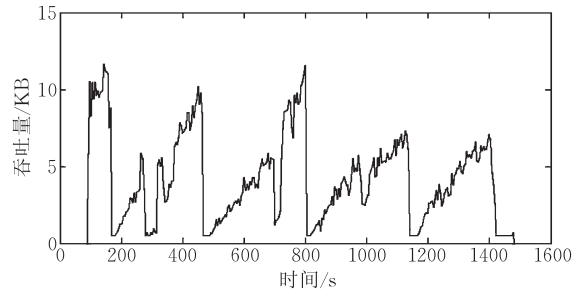


图 11 广电组播网络中发送广电数据吞吐量示意图

在新开发的可重构路由交换平台上加载相应的软件构件和硬件构件进行相应的功能和性能测试. 我们将编译好的软件构件分别下载到不同的运行平台中, 下载时间都控制在 0.5s 以内. 运行启动时间都在 20s 以内, 在任务切换忍受范围之内. 硬件构件主要对多块 FPGA 通过 USB 口下载硬件构件代码实现, 分别对具有不同功能的 FPGA 进行下载硬件构件代码, 通过测试构件功能和性能均能达到设计目标.

表 1 可重构路由平台中软件构件下发和重构时间

功能	大小/MB	运行环境	下载及重构时间/s
BGP 协议(版本 1)	1.09	Vxwork	0.18
BGP 协议(版本 2)	1.85	Linux	0.10
OSPF 协议	1.4	Linux	0.23

表 2 可重构路由平台中硬件构件下发和重构时间

功能	大小/MB	运行环境	下载及重构时间/s
商用芯片的控制和主控通信等功能	1.65	Xilinx 公司 Vertex5 系列	8
从商用芯片出来的数据的数据格式转换	3	Xilinx 公司 Vertex5 系列	8
完成交换排队调度等功能	5.284	Xilinx 公司 Vertex5 系列	23
查表转发程序	8	Xilinx 公司 Vertex5 系列	20

6 总 结

本文基于面向服务提供的思想, 提出以网络承载服务为核心, 给出可重构网络的设计, 并给出了网络构建和重构算法. 它具有以下优点: (1) 以服务为出发点进行业务和资源配置重构; (2) 可扩展性强, 可以适应未来快速发展的业务类型; (3) 可避免为承载特定业务而大规模改造或设计新的核心网络基础设施. 但可重构网络体系架构对未来大规模应用的适应性仍有待组网验证. 因此, 下一步工作将进行可重构网络体系架构的实验网建设, 为大规模工程

部署做了准备。

### 参 考 文 献

[1] Andersen D G, Balakrishnan H, Feamster N. Accountable Internet protocol//Proceeding of the ACM SIGCOMM2008 Conference on Data Communication. Seattle, USA, 2008; 339-350

[2] Mosharaf N M, Chowdhury Kablr, Boutaba Raouf. Network virtualization: State of the art and research challenges. IEEE Communication Magazine, 2009, 47(7): 20-26

[3] Kim Changhoon, Caesar Matthew, Rexford Jennifer. Floodless in SEATTLE: A scalable ethernet architecture for large enterprises//Proceedings of the SIGCOMM 2008. Washington, USA, 2008; 3-14

[4] Amir-Hamed Mohsenian-Rad, Vincent W S Wong, Juri Jatskevich, Robert Schober. Optimal and autonomous incentive-based energy consumption scheduling algorithm for smart grid //Proceedings of the IEEE ISGT2010. Washington, USA, 2010; 1-6

[5] Wang Feng, Xiong Yongqiang, Liu Jiangchuan. mTreebone: A collaborative tree-mesh overlay network for multicast video streaming. IEEE Transactions on Parallel and Distributed Systems, 2010, 21(3): 379-392

[6] Tennenhouse David L, David J. Wetherall towards an active network architecture. ACM SIGCOMM Computer Communication Review, 2007, 37(5): 81-94

[7] Zhang D, Adelman D. An approximate dynamic programming approach to network revenue management with customer choice. Transportation Science, 2009, 43(3): 381-394

[8] Bouabence G, Jelger C, Tschudin C. Virtual network stacks//Proceeding of the ACM Workshop on Programmable Routers for Extensible Services of Tomorrow. Seattle, USA, 2008; 45-50

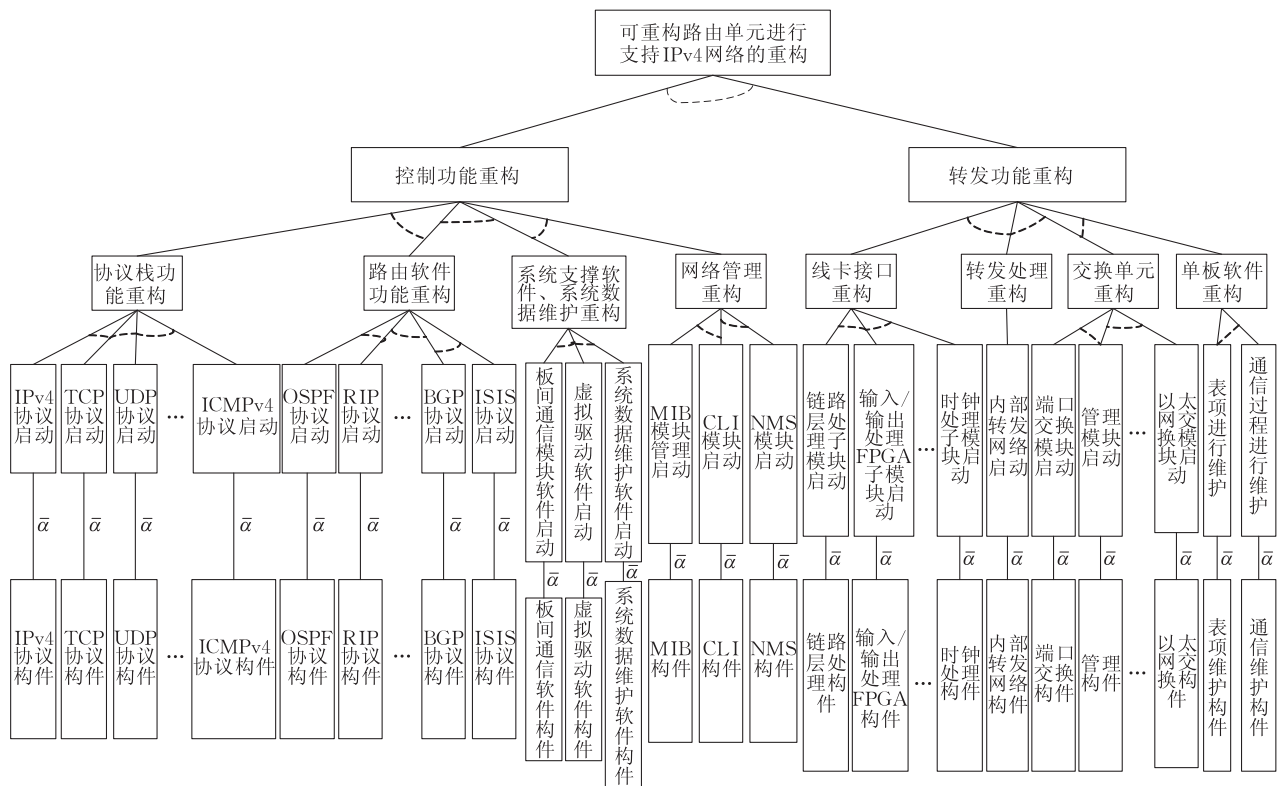
[9] Zhu Y, Ammar M. Algorithms for assigning substrate network resources to virtual network components//Proceedings of the IEEE INFOCOM 2006. Barcelona, 2006; 1-12

[10] Zhang Yi, Sun Zhi-Gang. Virtualization technology for trustworthy network research. Chinese Journal of Computers, 2009, 32(3): 417-423(in Chinese)  
(张怡, 孙志刚. 面向可信网络研究的虚拟化技术. 计算机学报, 2009, 32(3): 417-423)

[11] Jacob J, Partridge C. How can optiRS innovation benefit from GENI? //Proceedings of the 21st Annual Meeting of the IEEE Lasers and Electro-OptiRS Society. Newport Beach, USA, 2008; 545-546

[12] Peterson L, Muir S, Roscoe T, Klingaman A. PlanetLab architecture: An overview. PlanetLab Consortium: Technical Report PDN-06-031, 2006

[13] Wang Cheng-En. On the possibility of manufacturing system reconfiguration. Computer Integrated Manufacturing Systems, 2000, 6(4): 1-5(in Chinese)  
(王成恩. 制造系统的可重构性. 计算机集成制造系统 (CIMS), 2000, 6(4): 1-5)



附图 1 可重构交换平台重构为支持 IPv4 网络目标分解及构件确定



**LIU Qiang**, born in 1982, Ph. D. candidate. His research interests include broadband information network research and core router designing.

**WANG Bing-Qiang**, born in 1963, Ph. D. , professor, Ph.D. supervisor. His research interests include broadband information network research and core router designing.

**XU Ke**, born in 1974, Ph. D. , professor, Ph. D. supervisor. His research interests include next generation Internet, switch and router architecture, P2P and overlay network.

### Background

This research is supported by the National High-Tech-ology Research and Development Program (863 Program) of China project (No. 2008AA01A323) and the National Basic Research Program (973 Program) of China under grant No. 2007CB307102.

In order to keep consistency and coherence of network's development, the next-generation network does a lot of attempts to integrate the underlying network heterogeneity of existing communication infrastructure, support information for all areas, and ensure network management, controllability, and reliability. Based on the research on the exist technology of the next-generation network and service providing-oriented thinking, this project works on the architecture of the reconfiguration network, which is integrated carrying for

all kinds of current network traffic. This project also works on the reconfigurable routing and switching platform. In addition, based on reconfigurable service ability, the reconfiguration network architecture can supports most kinds of future network traffic. An important object of the project is to research into new service providing-oriented network architecture as well as construction technology, which can provide extended service ability for next generation network traffic. Furthermore, this architecture can break away from monopolization of traditional router factory, the third part factory only manufactory the network components as the standard. Achievements of this paper aim to solve the extended service ability providing.