

面向 SPR 的异构路由器聚合交换技术*

管剑波, 苏金树

(国防科技大学 计算机学院, 湖南 长沙 410073)

摘要 互联网的快速发展使得传统的 POP 组网方式面临巨大的挑战, 部署 SPR (Single POP Router) 路由器被认为是解决问题的有效手段。提出了异构路由器的聚合交换技术, 可以将 POP 中的多个异构路由器以集群的方式实现一个规模更大的 SPR 路由器, 具有降低组网成本、保护现有投资、简化网络管理、提高交换效率等优点。给出了聚合交换技术中需要重点研究的几个关键问题以及研究思路。

关键词 单一 POP 路由器; 路由器集群; 聚合交换

中图分类号 TP393 **文献标识码** A

The Aggregating Switching of Heterogeneous Routers for Single POP Applications

GUAN Jian-bo, SU Jin-shu

(College of Computer, National Univ. of Defense Technology, Changsha 410073, China)

Abstract The fast development of Internet challenges the traditional POP network. SPR (Single POP Router) is considered as an effective approach to resolve this problem. A novel clustering technology called aggregating switching is proposed. It can aggregate several heterogeneous routers into one large scale SPR and outperform the traditional POP network in many aspects. Some open problems about aggregating switching are discussed.

Key words single POP router; router cluster; aggregating switching

随着互联网带宽和规模的迅速发展, 网络提供商必须随应用扩展逐渐增加各类路由器, 从而导致服务提供点 POI (Point Of Presence) 的结构日趋复杂, 给网络运营商的部署和管理带来了极大的困难。单一 POP 路由器 SPR (Single POP Router) 是解决这一问题的有效措施, 也是未来高性能路由器发展的方向^[1]。但是已有的 SPR 或解决方案普遍存在更新代价高, 难以扩展等问题。

本文提出采用路由器集群构造 SPR, 以保护运营商已有的投资, 且具有良好的可扩展性。针对异构路由器集群的内部融合问题, 提出了支持交换层面互操作的聚合交换技术, 并给出了聚合交换技术今后研究的几个关键问题。

1 研究背景

提出单一 POP 路由器 (SPR) 的动机是为了解决当前 POP 组网中存在的网络结构日益复杂、管理困难等不足, 而路由器集群方式是实现 SPR 的有效手段。路由器接口的标准化使得构建异构路由器集群成为可能, 而且这种集群在灵活性上具有很大优势, 但是异构路由器的集群仍然存在很多问题值得深入研究, 这也正是本文提出聚合交换技术的背景。

1.1 SPR 的提出

POI (Point Of Presence) 是网络运营商的服务提供点, 它兼有接入网的汇聚、核心网结点和本地网络服务等重要功能, 是互联网的重要组成部分。当前网络运营商的 POP 网络一般部署为一组离散路由器

* 收稿日期: 2005-05-20

基金项目: 国家重点基础研究发展计划项目(2003CB314802), 国家自然科学基金项目(90104001)

作者简介: 管剑波(1977—), 男, 博士生。

的互联 mesh 结构,如图 1 所示。

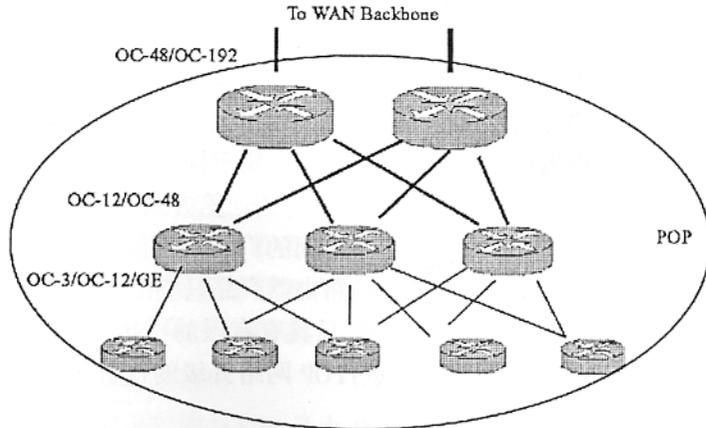


图 1 POP 网络结构

Fig.1 POP network architecture

在这种结构下,POP 网络中的路由器是相互独立的,而且通常是异构的。但是随着网络流量的增长,这种结构面临以下挑战:

- 可扩展性差。由于传统的 IP 路由器只具有单一的容量固定的交换网络,所以当网络流量快速增长使 POP 需要扩展时,网络运营商就需要购买更大规模的路由器来支持,并且需要重新设计和部署 POP 网络,这使得快速而高效地扩展网络变得异常困难。

- 难以保护网络运营商原有投资。为了使 POP 网络能随着流量的增加而不断扩展,网络运营商每 12~18 个月就要升级路由器,这使得设备的生命周期非常短暂,而且不间断地进行新产品的评估、集成和部署既耗时又耗财的。

- 网络复杂度高,难于管理。网络运营商需要对 POP 网络中的每个路由器进行单独管理和配置。然而随着 POP 网络中路由器数量的增加及拓扑结构变得更加复杂,网络管理已经成为网络运营商的噩梦。

- 消耗大量线卡用于 POP 内部路由器间互连,造成资源的浪费。由于 POP 网络中的路由器都是独立运行且通常是异构的,所以它们之间不支持紧耦合的互操作行为。POP 内的路由器之间仍然需要采用普通线卡互连,事实上浪费了线卡上昂贵的复杂逻辑功能部件,例如转发引擎、缓冲区及调度部件、网络接口等等。

正是由于上述情况的存在,近年来单一 POP 路由器 SPR 成为研究的热点。SPR 路由器被认为是解决上述 POP 组网问题的有效途径之一。SPR 路由器具有交换容量大,端口密度高,接口类型多样,可提供多种服务等特点。网络运营商通过部署一台 SPR 路由器可以完成现有 POP 网络的所有功能,即图 1 中所示的整个 POP 网络可视为一台 SPR 路由器。采用 SPR 路由器概念替代现有 POP 组网方式的好处是显而易见的——网络复杂度低,可靠性高,易于扩展且大大简化了网络管理。

目前一些网络设备厂商抛弃了单纯追求背板转发速率的做法,而转向面向 POP 应用的可扩展 SPR 路由器的研制,具有代表性的有 Pluris 公司的 MPR 路由器,Juniper 高速的 T640 路由器以及 CISCO 公司的 CRS-1 路由器等等。由此可见,SPR 路由器具有很大的发展潜力。

1.2 路由器集群技术

从转发交换层面来看,路由器的体系结构可分为集中式、分布式、并行式和集群式四类^[2,3]。各输入端口进来的 IP 分组全部送往中央处理机处理的结构为集中式结构,早期的路由器为这种结构。各线卡就地处理绝大部分 IP 分组的转发,只将特殊 IP 分组送往中央处理机进行处理的为分布式结构,现在的商用高端路由器大部分为这种结构。由于网络链路速度比 CPU 和存储器速度提高快,各线卡处理 IP 分组的能力即将接近极限。近年来,人们开始研究 IP 分组并行转发的问题^[4]。在并行式结构中,高速链路进来的 IP 分组被分派到多个转发器并行处理。分布式路由器和并行式路由器都是紧耦合结构,体系结构相对固定,扩展性受到一定限制。

路由器集群式结构是将多个常规路由器连接起来组成一个单映像路由器。所谓“单映像”是指：从组网的角度来看，一个路由器集群是一台路由器，而不是一个网。可以看出，路由器集群结构非常适合单一 POP 路由器的实现，前述几种 SPR 路由器代表产品也正是采用了集群式的体系结构。它们都是将多个单机架路由器通过互连网络连接起来，构成一个更大规模的可扩展路由器，同时对外保持单一映像。

但是，MPR 路由器、T640 路由器以及 CRS-1 路由器都属于同构式的路由器集群，即组成路由器集群的多个单机架路由器是完全相同的，它们之间通过专有的通信方式互连，不支持与其它种类的路由器集成。造成这种情况的原因是缺乏集群内部路由器之间的广泛的通信接口标准，各厂商只能开发其单机架路由器才支持的专有互连通信协议。所以上述路由器产品只能由各自厂家生产的特定的单机架路由器组成集群，不支持更广泛的异构式路由器集群，只具有有限的可扩展性。在面向单一 POP 应用的环境中，这种同构式的路由器集群具有明显的缺点：POP 网络升级的代价过高，不能延长现有路由器的有效寿命，无法保护运营商已有的投资。

能否将不同厂商生产的异构路由器产品互连，构成更大规模的单一映像的路由器集群呢？本文根据已有的交换网络接口标准，提出了基于开放接口的交换网络聚合技术，以支持异构路由器互连构成路由器集群，然后针对这种聚合交换结构，提出了几个今后仍待继续研究的关键问题。

2 支持 SPR 的聚合交换技术

在图 1 所示的 POP 网络中，多个路由器由于功能和性能的不同，所以往往是异构的，将它们通过集群的方式组成一个规模更大、更易管理且高度灵活可扩展的 SPR 一直是网络运营商追求的目标。这样不仅可以降低 POP 组网成本，大大简化管理，而且可以保护已有投资，减少网络升级的代价。异构路由器的集群技术一直是研究的热点和难点，路由器集群的“单映像”特性主要通过控制层面的软件集群技术来实现，而本文提出的聚合交换技术力图在转发交换层面将多个异构路由器有机地融合在一起，组成一个紧密耦合的更大规模的转发交换平台，为异构路由器集群提供硬件基础的支撑。

2.1 交换网络的 NPSI 接口规范

过去阻碍异构路由器集群的一个障碍是缺乏统一的互连标准，而 NPF (Network Processing Forum) 作为路由器厂商、网络运营商、芯片制造商等多方共同组成的一个标准化组织，于 2002 年发布了 NPSI (NP Streaming Interface) 接口规范，作为未来网络设备中接口单元、网络处理单元 (NPE) 以及交换网络 (Switching Fabric) 之间的接口实现标准^[5]。NPSI 接口在 NPF 推荐的网络交换设备体系结构中的位置如图 2 所示。

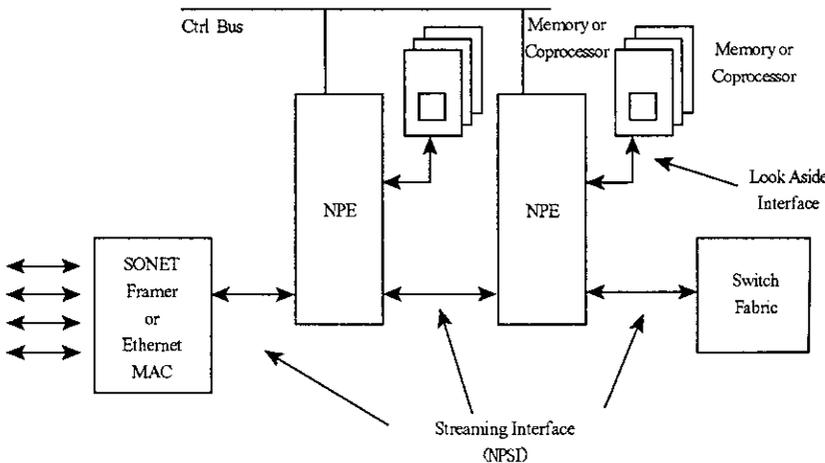


图 2 NPF 体系结构中的 NPSI 接口

Fig.2 NPSI interfaces in NPF architecture

对于交换网络的对外接口,NPSI 规范主要定义了以下内容:

- 物理层通信接口及信号定义
- 数据帧格式及定界、纠错等功能
- 多层次的拥塞信息及流控语义
- 其它可配置参数

可以预见,在不久的将来,不同厂商生产的异构路由器,从高端的核心路由器到低端的边缘路由器,其交换网络对外都将遵循 NPSI 接口规范,实现相同的通信协议,呈现出相同的操作特性,可提供并理解相同的流控信息。这就为多个异构路由器紧密耦合,实现交换网络间的融合在标准上奠定了基础,使得利用 POP 网络中的多个异构路由器通过集群方式实现更大规模的单一 POP 路由器成为可能。NPSI 规范是异构路由器聚合交换技术在标准上的基础。

2.2 聚合交换技术实现 SPR 的物理结构

在物理上,通过聚合交换技术实现的 SPR 是由 POP 中多个异构的路由器按照一定的拓扑结构互连而成的集群结构,如图 3 所示。其具体拓扑可根据 POP 实际应用情况灵活配置,而聚合交换技术侧重于通过标准的互操作行为使各路由器的交换网络协同工作,形成一个有机的整体,从而实现多个异构路由器在转发交换层面的融合。在传统的 POP 组网中,不同路由器之间仍然是通过各自的网络接口线卡进行互连,既浪费了线卡上昂贵的逻辑部件,又阻碍了遵循 NPSI 规范的交换网络间控制信息的互通。为了实现交换网络之间基于 NPSI 的互通和互操作行为,需要引入内部互联卡进行 SPR 内部异构路由器之间的互联。

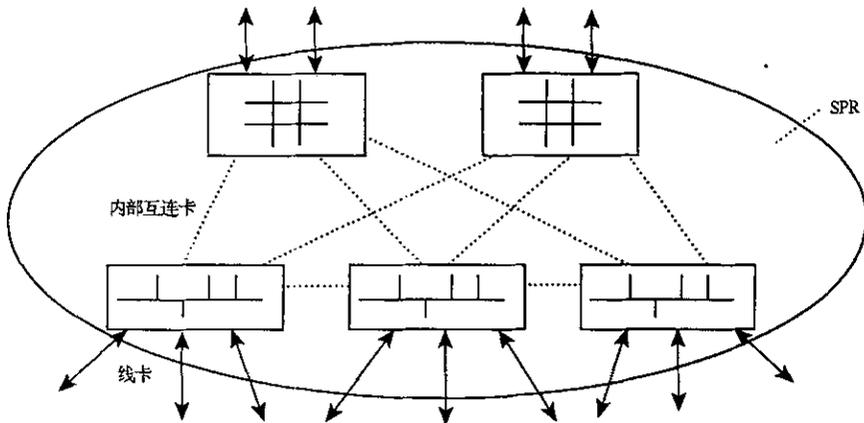


图 3 聚合交换实现 SPR 的物理结构

Fig.3 Physical structure of SPR using aggregating switching

由于各路由器的交换网络对外都遵循 NPSI 规范,所以在上述 SPR 中多个交换网络之间不需要使用昂贵的通用线卡互联,而采用一种简单廉价的内部互联卡实现紧密耦合。在图 3 中,SPR 中采用内部互联卡的连接用虚线表示。内部互联卡可以认为是各厂商为实现交换网络间直接透明互连而设计的一种专用卡,其上仅仅实现交换网络之间传递的 NPSI 控制信息的翻译和转接等功能,物理接口可以采用廉价而成熟的 VSR 技术来实现。图 4 显示了内部互联卡在异构路由器聚合交换中的作用。可以看到在这种集群结构下,SPR 具有很好的可扩展性和灵活性,且功能复杂的通用线卡都用于 SPR 的外部网络接口,避免了资源的浪费,从而降低了 POP 实现的整体代价。

2.3 异构路由器聚合交换实现 SPR 的逻辑模型

从图 3 中可以看出,通过集群方式组成 SPR 的各个路由器的内部结构可能有很大差别,例如高端的核心路由器可能是基于 Crossbar 结构的,而低端的边缘路由器可能是基于共享缓冲结构或者总线互连结构的,但是由于它们对外行为是一致的,所以在逻辑上可以建立一种统一的交换网络模型来表示它们。

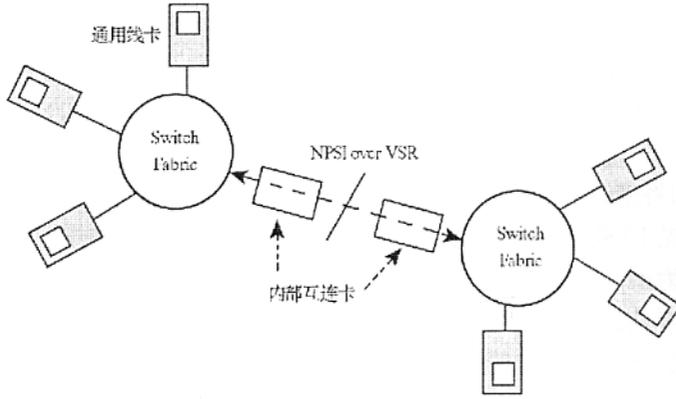


图4 内部互连卡在聚合交换中的作用

Fig.4 Interconnecting cards for aggregating switching

NPSI 规范定义了交换网络之间可交互的多种控制信息,如各端口的拥塞状况、流控信息、调度信息以及其他可自定义的信息。在 SPR 中,这些信息实际上在各个交换网络的调度控制器之间建立了联系。聚合交换的基本思想在于,各交换网络通过交换这些信息,可以及时调整自己的交换行为,达到减少拥塞、屏蔽故障、支持 QoS 等目的,以期获取最高的交换性能和最大的可用性,多个交换网络以及它们之间的内部互联卡共同组成一个更大规模的交换平台。所以在逻辑上,聚合交换将 POP 中多个异构的交换网络融合成一个大规模的交换网络,同时在上层软件协议的支持下,多个异构路由器通过集群的方式融合成一个 SPR 路由器。

在这里给出集群方式组成 SPR 的逻辑模型,其中每个普通路由器被抽象为由一个交换实体和一个路由实体组成。SPR 路由器的交换平面由多个交换实体通过物理互连构成,这里需要重点研究的是聚合交换技术以及分布式交换调度机制及算法,而集群路由器的路由平面由图中的多个路由实体构成,需要研究的关键问题包括分布式路由协议体系结构模型和转发表的同步机制及算法。该模型如图 5 所示。

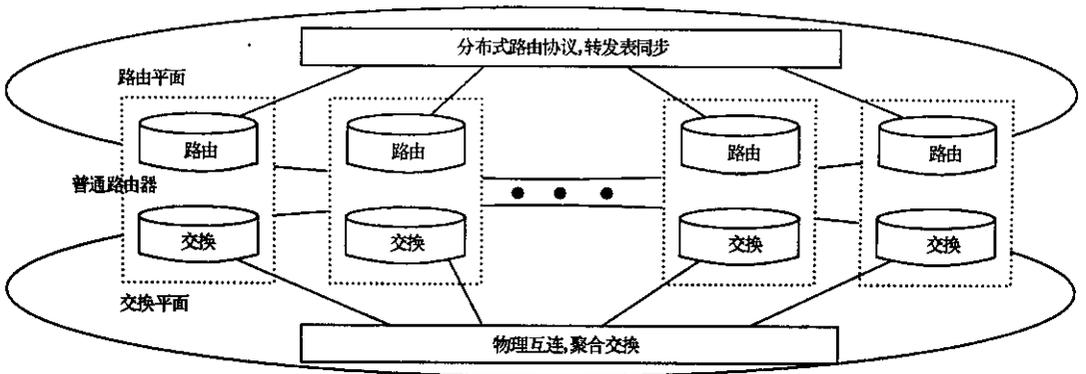


图5 聚合交换实现 SPR 的逻辑模型

Fig.5 Logical model of SPR using aggregating switching

由以上论述可以看出,通过异构路由器聚合交换实现 SPR 具有以下特点:

- 理论上支持任意数量、任何结构路由器的集群,从而具有很好的可扩展性和灵活性。
- 由于各交换网络之间的信息交互以及协同工作,聚合交换实现 SPR 相对传统 POP 组网具有更好的交换效率、更高的可用性,而且使得对 QoS 的支持成为可能。
- SPR 在逻辑上就是一台路由器,从而大大简化了 POP 的管理和配置。

· 网络运营商对网络的升级可以更加灵活和高效,并且保护了已有的投资,延长了现有网络设备的寿命。

3 聚合交换技术研究的几个关键问题

聚合交换技术的提出,使得 POP 网络中多个异构路由器可以通过集群的方式组成一个紧密耦合的 SPR。但是本文给出的仅仅是聚合交换技术的思想和整体结构,在聚合交换技术中,仍然有以下几个关键问题需要在今后进行更加深入的研究:

· 异构交换网络的统一建模问题。用一种统一的模型表示各种异构的交换网络,是实现异构交换网络融合的理论基础,目前还没有看到这方面的研究。一个好的模型应该能够反映各种异构交换网络的规模、容量以及功能等特性,这有利于对路由器集群进行合理的规划。

· 拓扑发现问题。交换网络之间怎样通过交互彼此信息,从而全面了解整个路由器集群的拓扑,以及各个交换网络的端口、速率、交换能力等参数,为自身的交换调度策略提供参考。

· 拥塞控制问题。各个交换网络的拥塞状况如何传递到其他相关交换网络,以及拥塞产生时各交换网络如何反应。

· 支持流量控制和拥塞控制的交换网络调度算法。传统的交换调度算法只专注于单个交换网络的内部阻塞等问题,缺乏对远端反馈信息的支持。而在路由器集群中,各交换网络需要根据交互的远端信息调整自己的交换行为,这有利于异构路由器互连组合成为一个高效率的 SPR。

· 内部互联卡的设计。内部互联卡是实现聚合交换技术的关键,需要在该互联卡上实现 NPSI 接口规范的理解和转译功能,对于实现异构交换网络间的透明直接连接起着重要作用。

4 结 论

提出了面向单一 POP 应用的聚合交换技术,POP 网络中的多个异构路由器可以通过集群的方式组成一个 SPR。与传统的 POP 组网方式相比,聚合交换具有灵活高效、易于扩展、简化 POP 管理、保护现有投资等优点,具有很大的发展潜力。给出的聚合交换技术中需要进一步深入研究的几个关键问题,为下一步的研究指明了方向。

参 考 文 献:

- [1] Chao H J. Next Generation Routers[J]. Proceedings of the IEEE, 2002, 90(9):1518-1558.
- [2] Chan H C B. A Framework for Optimizing the Cost and Performance of Next Generation IP Routers[J]. IEEE Journal on Selected Area in Communication, 17(6), June 1999.
- [3] Aweya J. IP Router Architectures: An Overview[J]. Journal of Systems Architecture, 1999, 46(2000):483-511.
- [4] Iyer S, McKeown N. Making Parallel Packet Switches Practical[A]. IEEE INFOCOM[C], Alaska, April 2001.
- [5] Network Processing Forum Hardware Working Group. Streaming Interface(NPSI) Implementation Agreement[S]. <http://www.npforum.org>, October, 2002.

