摘要

IEEE 目前正在标准化一个新的、全双工、空间重用的环形网络架构—弹性分组环（IEEE P802.17）。该技术集 SDH 光纤环网的可靠性、以太网技术的高效和经济性于一体，是城域网技术发展的重要方向。

本文首先介绍了弹性分组环的特点，在此基础上对弹性分组环的帧结构、拓扑发现机制、保护倒换机制进行了初步的探讨。然后重点研究了弹性分组环公平带宽管理机制，主要分两步进行：第一，提出了一个新的 RPR MAC 参考模型——IA（Ingress Aggregation）参考模型，按功能将其分为五个功能实体并分别对每个功能实体进行具体的定义和设计；第二，基于 IA 参考模型提出了全新的 RPR 带宽分配算法 DBRR（Distributed Bandwidth Reallocated in Rings），该算法很好的解决了当前算法草案中存在的“非平衡业务”情况下持续的带宽震荡现象。最后通过计算机仿真了 DBRR 算法的性能，仿真结果表明 DBRR 算法具有良好的性能。

关键词：弹性分组环　城域网　公平性　环中的分布式带宽重用

ABSTRACT

IEEE is currently standardizing a new full-duplex spatial reuse ring network architecture, called the Resilient Packet Ring (RPR, IEEE P802.17). The reliability of SDH ring, high-efficiency and economy of Ethernet are integrated into this technology. It is an important trend for the evolvement of MANs.

Firstly, this paper introduces the characteristics of RPR. On the basis of which, frame format, topology discovery and protection switching schemes are elementarily discussed. Emphasis of our research is put on fair bandwidth management scheme of RPR. The research work follows two steps: firstly, a new RPR MAC reference model-IA (Ingress Aggregation) is presented. According to functions, it is divided into five functional entities, each of which is defined and designed in detail. Secondly, based on IA, a new algorithm-DBRR (Distributed Bandwidth Reallocated in Rings) is provided, which greatly resolves the phenomena of continual bandwidth oscillations with unbalanced and constant-rate traffic demand scenario in current drafts. Finally, by the means of computer, the performance of DBRR algorithm is simulated, the results show DBRR algorithm has better performance.

Keywords: Resilient Packet Ring MANs Fairness DBRR

创新性声明

本人声明所呈交的论文是我个人在导师的指导下进行的研究工作及所取得的研究成果。尽我所知，除了文中特别加以标注和致谢中所罗列的内容以外，论文中不包含其它人已经发表或撰写过的研究成果；也不包含为获得西安电子科技大学或其它教育机构的学位或证书而使用过的材料。与我一同工作的同志所做的任何贡献均已在论文中做了明确的说明并表示了谢意。

申请学位论文与资料若有不实之处，本人承担一切相关责任。

本人签名：岳鹏　　日期：2003.1.16

关于论文使用授权的说明

本人完全了解西安电子科技大学有关保留和使用学位论文的规定，即：研究生在校攻读学位期间论文工作的知识产权单位属西安电子科技大学。本人保证毕业离校后，发表论文或使用论文工作成果时署名单位仍然为西安电子科技大学。学校有权保留送交论文的复印件，允许查阅和借阅论文；学校可以公布论文的全部或部分内容，可以允许采用影印、缩印、或其它复制手段保存论文。（保密的论文在解密后遵守此规定）

本学位论文属于保密，在2年解密后适用本授权书。

本人签名：岳鹏　　日期：2003.1.16

导师签名：邓智亮　　日期：2003.1.17

第一章 绪论

§1.1 研究的背景和意义

近年来，互联网的业务量以每年翻一番的速率高速增长，远远超过网络中传统的话音业务量。在互联网接入用户数的激增、大量基于 web 的多媒体新业务的产生以及用户终端计算机功能的升级等因素的驱动下，预计互联网业务量将继续以指数式增长，并将成为网络流量中的主要部分。

另一方面，网络的发展也不平衡。在骨干网中，由于光缆的大量铺设和 DWDM（密集波分复用）技术的应用，骨干网络的容量已大大提高，每比特的传输成本也大大降低。与此同时，局域网与接入网的容量也正不断提高，目前十兆与百兆以太网接口已配置到桌面，千兆以太网正成为局域网的主流，10G 以太网技术已经出现，各种接入网技术的应用正使宽带互联网接入千家万户。相比之下，城域网（MAN）技术的发展相对滞后，正成为整个网络通信的瓶颈。目前城域网主要采用的仍是面向同步时分复用的 SDH 技术，它面临着双重压力，即采用新技术的骨干网与局域网容量激增带来的非常迫切的扩容压力和如何高速、有效与低价地承载正成为网络主流的互联网业务。

IP 领域很早就意识到了环形网络结构的价值，并已在这方面作了大量努力，提出了像令牌环和 FDDI（光纤分布数字接口）这样的解决方案，但这些方案却无法满足 IP 流量和带宽增长的要求，也无法满足在拥塞情况下维持高的带宽利用率和转发量、保证结点间业务平衡、从结点或传输媒体故障中迅速恢复、可即插即用等 IP 传输发展的要求。因此，像令牌环和 FDDI 这样的环形网并不适用于城域网。服务提供商和企业需要一种扩展性好、能够健壮地应用在城域网之上、以千兆的速率传输 IP 分组的技术。

正是在这样的背景下，弹性分组环（RPR, Resilient Packet Ring）技术应运而生。2000 年 11 月正式成立了 IEEE's 802.17 弹性分组环工作组(RPRWG)，希望开发一个 RPR (Resilient Packet Rings) MAC 标准，优化在 LAN、MAN 和 WAN 拓扑环上数据包的传输。

§1.2 城域网技术概述

RPR 技术与城域网技术密切相关，正是城域网技术的发展推动了 RPR 技术的产生。因此在介绍 RPR 之前，我们首先简要了解一下几种可能成为主流的城域网技术。

城域网是数据骨干网和长途电话网在城域范围内的延伸和覆盖，它承担着集团用户、商用大楼、智能小区等业务接入和通路出租等纷繁复杂的任务，需要通过各类网关实现话音、数据、图像、多媒体、IP 接入、各种增值业务以及智能业务，并与各运营商的长途网和骨干网实现互通。城域网不仅是连接传统长途网与接入网的桥梁，更是传统电信网与新兴数据网络的交汇点及今后三网融合的基础。

近年来，以 10G SDH 和 DWDM 技术为代表的光纤传输技术有了重大突破，骨干网带宽从 G 比特向 T 比特发展；在企业和居民用户端的网络速率，则随着 G 比特以太网技术进入商业应用而向 G 比特发展。这两个趋势使城域网产生了巨大的带宽压力和多种新的功能需求，主要包括：高带宽、大量的用户结点、灵活的带宽分配、多业务支持和协议无关性、保护和自愈以及便捷的网络管理等。

有需求就有创新，满足上述要求的光城域网技术正蓬勃发展，业界将其统称为多服务提供平台 MSPP （Multiservice Provision Platform）或多服务传送平台 MSTP（Multiservice Transmit Platform）。由于应用技术不同，它的主要发展有以下几个方向：

一、以太网方案

作为城域传送网的最初形态，以太网直连被看作是最简单的城域网方式，其较低的成本以及便捷的开通和运营方式一度受到服务供应商的青睐。

GE、10GE 大容量以太网技术使城域网应用上了一个新的台阶，IEEE 也在 802.3 标准中明确对其进行了定义。这种技术适应了城域网中占据主导地位的 IP 业务增长的需要，且支持附加大带宽、高成本的城域核心网络，可与 TDM 或 DWDM 光纤网络进行无缝连接，满足更大容量组网的需求。但是，该技术也存在一定的不足：

1. 故障恢复时间长：和 SONET/SDH 不同，Ethernet 不能利用环状拓扑实现快速保护机制。Ethernet 通常采用生成树协议（Spanning Tree Protocol）消除回环和拓扑更新，保护过程相对 SONET/SDH 是非常缓慢的（几十秒 vs. <50ms）。即便采用链路会聚（802.1ad），也只能提供链路级的弹性，和 SONET/SDH 相比还是慢（500ms vs. 50ms）。因而不能满足实时性要求很强的业务对时延的要求。
2. 公平（Fairness）性问题：由于带宽共享，Ethernet 不能很好的实现全局（global）的公平性。Ethernet 交换机能够提供链路级的公平性，也就是说 Ethernet 通常为所有输入端口公平的分配输出端口带宽，但这只是局部的，并不能够转化成全局的公平性。

解决以太网城域传送方式缺憾的方法是采用光以太网 RPR 技术（Optical Ethernet RPR）。RPR 是与媒质无关的 MAC（媒质接入控制）协议，它综合了以太网和 SDH 的优点。

RPR 是一个全新的概念，它将交换和传输简化，同时也把交换和传输这两项技术进行了有机的集成，使之成为一个整体。RPR 可承载多种业务，能象路由器一样在环上转发包括 IP 包在内的多种分组，在环上运行的业务可提供单播、组播和广播模式。由 RPR 组成的网络有以下特点：环上的业务是透明的，在 50 毫秒内实现二层保护倒换，具有自动拓扑发现和动态带宽管理功能；环和环上运行的业务具有弹性，可大可小、可多可少；接入环和骨干环可以互相嵌套；双向对称反转环都可用来传送数据、信令和网管帧；可进行动态的结点添加和删除；支持虚拟输出队列（VOQ）和服务质量等级的功能，支持三层（包括 IP 包在内）的存储转发。RPR 帧格式与业务的类型、速率无关。RRP 具体的特点将在下一节详细介绍。

二、SDH 方案

由于具有可靠的业务保护能力，SDH 技术也成为城域网的一种选择。但是令人感到棘手的是：对于固定速率的业务（如传统话音业务），SDH 很容易将其适配到固定容量通道中，但对于可变速率 VBR 业务和任意速率业务，SDH 则显得不够灵活，特别是传送效率不高。

SDH 的市场高占有率以及城域网的巨大增值潜力使 SDH 的倡导者们费尽心思，在原有 SDH 的基础上加入对数据业务层的处理，比如以太网的二层处理、ATM 的统计复用等功能，使其更适合数据业务的传送。

对于以太网业务，其在映射到 VC 之前需要经历处理的过程有：二层交换、协议封装、映射前的处理等。具体如图 1-1 所示

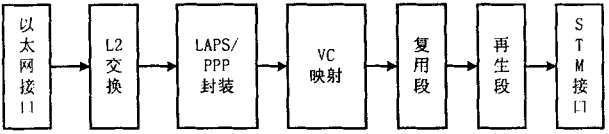

图 1-1 以太网业务在 MSTP 系统中的处理过程

对于 ATM 业务，系统提供统计复用功能，可对多个 ATM 业务流中的非空闲信元进行抽取，复用进一个 ATM 业务流，以提高其在 SDH 线路上的利用率，同时节约了 ATM 交换机的端口数。另外，还可以在 SDH 环路上形成一个 ATM 的虚拟通道环，这样 ATM 的业务层面可以实现环保护。图 1-2 表明了 ATM 业务在 MSTP 系统中的处理过程。

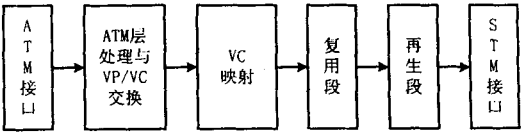

图 1-2 ATM 业务在 MSTP 系统中的处理过程

三、WDM 方案

继在骨干网及长途网络中应用后，波分复用技术也开始在城域网得到应用，特别是其巨大的容量、网络的扩展性以及业务的可扩充性，在城域网中显示出特有的优势。但是 WDM 技术的高成本是城域网环境无法接受的；另外针对城域网用户层业务的多样性和复杂性，城域波分复用技术必须向高效承载多业务方向演进。解决这些矛盾之后，CWDM(粗波分)和 OADM（光分插复用）环网技术将逐渐成为该技术的主导力量。

1. CWDM 城域传输技术

CWDM 技术一般应用于小型城域网或大型城域网的汇聚、接入层，它的波长数目一般为 4 波或 8 波，最多 16 波，波长从 1290nm 到 1610nm（16 波系统）。由于波长间隔较宽，CWDM 系统可以使用非致冷的 DFB（分布反馈）激光器和带宽滤波器，这样既延续了 DWDM 技术的优势，又具备了 DWDM 技术所不具备的低成本，低功耗，小尺寸等优点。它的出现解决了长期困扰城域网建设的性价比问题，而且它最大限度地利用了现有城域光纤基础设施，进而满足了未来小型城域网及大型城域网汇接、接入层业务所需要的带宽。

当然，CWDM 技术也有其不足之处，比如要建设一个 16 波的 CWDM 系统，其带宽范围覆盖了近 400nm 的光纤工作窗口，其中包括 1380nm 的高衰减区，普通的光纤介质根本无法适应，需敷设全波光纤才能满足要求。

2. 城域 OADM 传输技术

城域 OADM 环网技术是在考虑用户信号的可靠性基础上发展起来的。利用该技术，可以实现灵活的波长保护和调度。当前，固定波长的 OADM 在实际工程中已经被采用，波长可调、动态重构的 OADM 产品也即将走向商用。

由于上下波的数目以及要求不同，OADM 又可分为串行、并行、串并结合三种类型。

以上我们简要介绍了三种城域传送技术，可以看到这三种技术各有千秋，总体而言采用 RPR 与以太网技术相结合的光以太网 RPR 技术（Optical Ethernet RPR）无论从性能、复杂度还是组网成本都占很大的优势。光以太网 RPR 技术是否能够成为城域传输技术的主流技术，关键在于 RPR 技术的完善。下一节我们将简要的介绍 RPR 技术，为后几章讨论 RPR 的关键技术做概念上的铺垫。

§1.3 弹性分组环（RPR）概述

1.3.1 RPR 技术介绍

RPR 是一种新型的网络技术，是为了满足基于分组的城域网（MAN）的要求而设计的一种与传输媒质无关的新的 MAC（Media Access Control）层协议标准。RPR 网络是一种环形结构，它由分组交换结点组成，相邻结点通过一对光纤连接。RPR 定义了具有两个接口的媒质接入控制（MAC），系统级接口和物理层接口。RPR 的拓扑结构基于两个反向传输的环，通常外环按逆时针方向传送数据，内环按顺时针方向传送。由于 RPR 技术与媒质无关，因此可以应用于各种物理层技术之上，例如 RPR over SDH、RPR over fiber、RPR over DWDM 等等。RPR 在 IEEE 802 家族中所处位置如图 1-3 所示：

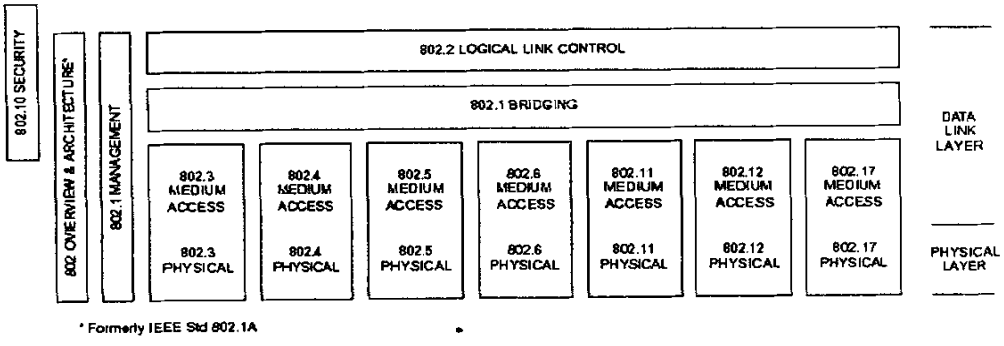

图 1-3 IEEE 802.17 在 IEEE 802 家族中所处的位置

RPR 如何满足城域网的要求反映在制订 RPR 标准的思想中，其目标主要集中在四个方面：弹性、公平性、可伸缩性和效率。

1. 弹性：若物理层检测错误则将该信息通知给 MAC 层。MAC 层根据错误的类型触发不同的保护机制，随后受到影响的 RPR 节点将会启动一个故障操作。这个操作将在故障发生后的 50ms 内完成警报通知并重新定向数据流。
2. 公平性：RPR 把带宽预留与服务质量结合起来管理环上的带宽分配。数据流在进入环路时首先进行分类、缓存，随后根据公平算法计算出的结果进行业务整形。MAC 按照一定的规则对本地业务和转发业务进行调度，保证环中每个节点能够按照事先分配好的权值公平的向环中注入数据。
3. 可伸缩性：RPR 不仅支持当前建议中的物理媒质和线路速率，而且还支持许多早期建议中定义的物理媒质和线路速率，这使得 RPR 组网具有很大的灵活性。除了环的原始带宽外，RPR 还采用标记（例如，RPR 与 MPLS 结合）方案，使得运营商能够具备在一个单环上管理成千上万数据流的能力。
4. 效率：RPR 采用分组交换技术，所以与电路交换相比能更有效的利用带宽。RPR 采用一对互为反向的环路，可根据负载分布的不同动态选择在哪条环路上

传送数据，达到平衡负载的目的。这种设计可以充分利用环的有效带宽，达到很高的带宽利用率。

另外，针对大部分数据业务的实时性不如话音那样强的特点，RPR 采用双环工作方式，和 SDH 相比能够更有效地分配带宽和处理数据，从而降低运营商和用户的成本。传统的 SDH 网是按双光纤环配置的，通常采用 1+1 工作方式，即其中一个环工作，沿一个方向传送信号，另一个环备用。如果工作环断开，备用环就沿反方向传送信号，这种瞬间的倒换是 SDH 的主要特点之一，它能保证话音等实时业务的连接。但是由于互联网业务主要是数据业务，实时业务占的比重较小，网络时延小的要求已经显得不是非常关键。诸如 FTP 和数据备份这样的协议和应用均可以承受一定的时延。对这些应用而言，SDH 时延特性好的特点就大材小用了。而 RPR 在任何时间双环都同时使用，并互为反方向传输，如果一个环出现故障，该环的所有业务全部转到另一个环上，此时工作环由于业务的突然增加有可能发生拥塞，为了解决这样的问题，RPR 引入了 QoS 参数，使高优先级业务先获得它所需的带宽，并且不受断开的影响，但优先级低的业务可能会蒙受时延影响。这使得使用 RPR 的运营商可以对高优先级业务收取较高的费用，对低优先级业务收取较低的费用。RPR 的优先化能力使运营商在城域网内通过以太网运营电信级业务成为可能。它们在提供电信级 QoS 的同时，降低了传送费用，并能提供下一代网络所要求的恢复能力、服务质量和可管理能力。

1.3.2 RPR 技术的特点

虽然 RPR 还处于标准化过程中，但其基本特征已经确定，主要有：双环结构、空间重用、带宽动态分配、统计复用、环上公平接入、保护与恢复、支持 CoS 等。

1. 双环结构。RPR 的典型用法是由传输方向相反的一对光纤组成环形拓扑结构，为区分环网中的两个环，一个称为“内”环，另一个称为“外”环，在环上一个结点有两个方向可以到达另一结点。RPR 环上的每根光纤既传送数据分组，又传送控制分组。控制分组优先级最高。
2. 空间重用技术。RPR 支持空间重用技术（SRP，Spatial Reuse Protocol）[15]。SRP 是一种与媒质无关的 MAC 层协议，可以用于各种物理层技术之上。SRP 协议具有寻址、读取数据分组、带宽管理和控制信息传送的基本功能。SRP 采用目的结点摘除分组的机制，不像 FDDI 那样要经过整个环路最后由源结点摘除。这样，分

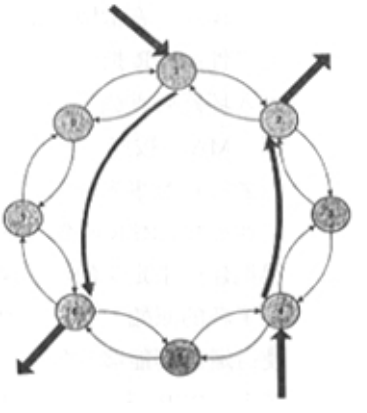

图 1-4 RPR 空间重用

组环可以在环的不同段上互不影响的同时传送数据，充分利用了整个环路的带宽。

图 1-4 所示 RPR 空间重用特性，可以看到结点 1 沿内环向结点 6 传送数据，同时结点 4 也沿内环向结点 2 传送数据。这两个数据流之间互不影响，都可充分利用传输通路的带宽。

3. 带宽动态分配和统计复用。传统的 SDH 环有 50%的环带宽是冗余的。RPR 环不象 SDH 环采用独立的带宽或者预设连接，它一般通过设置各结点的承诺带宽（CB, Committed Bandwidths），通过统计复用，来获得具有高性能处理突发数据的能力。RPR 根据用户需求分配带宽，而不象 SDH 那样分配固定时隙，对于保护，也没有像基于电路的保护那样，为保护预留带宽。
4. 环上公平接入。RPR 的一个目标是结点分布式接入，与 SDH 不同，RPR 不需要众多的结点管理。RPR 采用全球唯一的、永久性的 MAC 地址，加上快速保护和自动重建服务程序，为快速插入和删除结点提供即插即用的机制。
5. 保护与恢复。RPR 采用双环结构传输控制信息和数据信息，同时仍然维持与 SDH APS 类似的保护机制。正在考虑的方法有两种，一种是“环回（Wrapped）”方式，另一种是将数据流“绕开（Steering）”故障点的方式。当采用“环回”方式时，靠近故障点的结点将数据流“环回”到另一个环上（如内环数据转到外环），通过长路径使数据流维持与目的结点之间的连接。“绕开”方式实际上是通过改变数据传送方向，由沿短路径传送变为沿长路径传送，将数据流传送到目的结点，也可以将这两种方式结合起来使用，一旦发生故障，首先采用环回方式，然后根据具体情况，对环中某些结点采用“绕开”方式。
6. 支持 CoS。RPR 在其帧头设置了类型域，支持多种服务等级的 CoS。由于 IP 分组中的 ToS 域可以直接映射到 RPR 的类型域中，因而 RPR 能够有效的支持 IP 突发业务和话音传送业务。RPR 技术为数据业务和话音业务提供了一个 2 层统计复用的平台。无论是话音业务还是数据业务，接入之后都可在同一链路上传送，由于所有业务共享带宽，因而大大提高了带宽利用率。

§1.4 本文主要研究方向

前几节我们简要介绍了几种城域网和RPR技术的特点，无论是城域网还是RPR都有很都多问题尚待解决。

本文重点研究 RPR 的带宽管理机制，主要分两步进行：

第一，提出了一个新的 RPR 参考模型 RPR IA（Ingress Aggregation），以后为

叙述方便称这个新的参考模型为IA。IA参考模型主要是针对城域网的特点提出的，该参考模型与其它建议[3，4，5]提出的参考模型最大的区别在于它首先对输入的业务进行聚合，然后根据带宽分配算法为每个聚合流分配带宽。本文将在第三章对IA模型进行详细讨论。

第二. 在IA参考模型的基础上，提出了一个全新的RPR带宽分配算法DBRR（Distributed Bandwidth Reallocated in Rings）。通过对已提交的带宽分配算法草案的研究，发现目前的算法草案存在一些局限，例如，采用[3，4，5]公平算法在“非平衡业务”的情况下，将会产生严重的链路带宽震荡现象，且这种震荡将是永久的。很显然，持续的链路带宽震荡将严重的影响环网的吞吐率和空间重用效果。另外，研究发现当前的带宽分配算法要经过很多次信息交互，链路才能达到动态的平衡。信息交互的次数与环中相邻结点间的距离成正比，这样就存在一个网络开销问题。前面已经说明RPR技术主要是针对城域网设计、优化的。如果网络开销与网络的规模呈线性增长关系，在网络容量不变的条件下，整个网络的效率将会下降，带宽利用率将会降低。我们将在第四章对这些问题进行详细讨论。

§1.5 本文的内容组织

全文共分五章，第一章为绪论，简要介绍城域网技术、弹性分组环技术以及本文的研究方向。第二章讨论弹性分组环的关键技术，主要对RPR MAC层帧结构、拓扑发现（Topology Discovery）机制和RPR保护倒换机制（Protection Switching）进行初步的研究。第三章主要针对RPR带宽管理机制提出一个新的RPR MAC参考模型IA，在此基础上提出了DBRR算法，并从理论上进行了分析。第四章建立仿真模型，分别从公平性、吞吐率、算法收敛时间以及平均分组传输时延等方面考察采用DBRR算法的RPR网络性能。第五章总结全文，指出RPR研究中有待进一步研究的问题。

第二章 弹性分组环关键技术

§2.1 引言

第一章我们简要讨论了 RPR 技术的特点，本章我们将详细讨论弹性分组环的关键技术。通过对 RPR MAC 层协议草案的研究，我们将弹性分组环的关键技术分为四类，即：RPR MAC 层帧结构、RPR 公平带宽管理机制、RPR 拓扑发现机制和 RPR 保护倒换机制。本章只对 RPR MAC 层帧结构、RPR 拓扑发现机制和 RPR 保护倒换机制进行初步的讨论。RPR 公平带宽管理机制将在下一章重点讨论。

§2.2 RPR MAC 层帧结构

根据[3，4，5]，RPR MAC 帧结构版本有好几种，格式上大同小异。通过比较，我们认为 Gandalf 草案提出的 RPR MAC 层帧结构较为完整。因此本章以 Gandalf 草案为蓝本，讨论 RPR MAC 层帧结构。

RPR MAC 定义了三类帧格式，分别为：

a) RPR 数据帧格式

b) RPR 控制帧格式

   1) 一般的控制帧格式

   2) 公平消息帧格式

c) 扩展的 RPR 帧格式，例如 RPR 帧中携带 IEEE 802.1Q VLAN 标签

图 2-1 RPR 帧的基本格式

图 2-1 表示的是 RPR 帧的基本格式（不含帧定界）。上述的三类帧基本上遵循这样的帧结构，只是局部有所不同。本文先以 RPR 数据帧的帧格式为例，描述 RPR 帧的基本格式。RPR 控制帧和扩展的 RPR 帧的帧格式描述只讨论与 RPR 基本格式不同之处，相同之处不再重复。

2.2.1 RPR 基本的帧格式

为了讨论方便，本文将 RPR 基本的帧格式分为 RPR 帧头（head）域和 RPR 有

效负载（payload）域两部分。

首先描述 RPR 帧头部分。如图 2-2 所示，RPR 帧头中各个域的含义及功能描述如下：

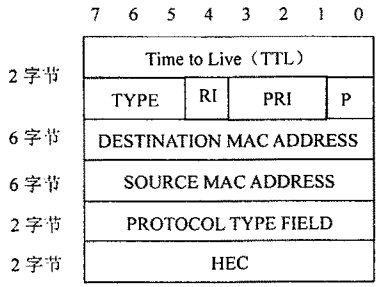

图 2-2 RPR 帧头格式

- ✧ 生存期（TTL）域（8 比特）

  该域是一个长度为 8 比特的跳数计数器。每当环中结点转发一次分组，则转发分组的 TTL 减 1。当 TTL 域的值减到零时，将分组从环上摘除。由于 TTL 域是 8 位的，因此最多允许环上有 256 个结点，即一个分组最多 255 跳，然而考虑到保护倒换的情况，RPR 支持的总结点数为 128。在 RPR 帧注入到环上之前，TTL 至少应当设置为环上总结点数的两倍。其作用是：当环发生保护倒换时，确保分组能够传送到目的结点，而不会被中间结点误摘除。

- ✧ 类型（TYPE）域（3 比特）

  类型域用于指示 RPR 帧的类型，大小为 3 比特。具体定义见表 2-1。

- ✧ 环标识符域（1 比特）

  该比特指示 RPR 结点将分组注入到哪个环上，0 表示外环，1 表示内环。

- ✧ 优先级（PRI）域（3 比特）

  优先级域指示 RPR 分组的优先级等级（0~7），值越大优先级越高。MAC 按优先级的高低向环上注入分组。RPR MAC 一般采用双转移队列——高优先级转移队列（HPTB）和低优先级转移队列（LPTB）。因而一旦分组注入到环上，该分组要么归高优先级分组处理，要么归低优先级分组处理，这取决于分组优先级与预设门限比较的结果。目前建议的草案将预设的优先级判定门限设置为固定值，并且为了保证一致性，环中各结点的优先级门限值相同。

表 2-1 类型值描述

类型值（二进制）	描述
000	保留
001	保留
010	保留
011	保留
100	保留
101	一般控制分组
110	公平消息分组
111	数据分组

- ✧ 校验位（P）域（1 比特）
- ✧ 目的地址域（48 比特）

  RPR 分组的目的地址采用全球唯一的 48 比特 IEEE802.3 地址。
- ✧ 源地址（48 比特）

  RPR 分组的源地址亦采用全球唯一的 48 比特 IEEE802.3 地址。
- ✧ 协议类型域（16 比特）

  协议类型域和以太网帧的类型域基本相同，除了分配给 RPR 的值以外，其它在以太帧类型域中定义的值同样有效。见表 2-2：

表 2-2 协议类型

值（16 进制）	协议类型
0x2007	RPR 控制类型
0x0800	IPv4 类型
0x0806	地址解析协议（ARP）类型
0x8100	VLAN 标签帧类型
其它	待定义

- ✧ 帧头差错校验（HEC）域（16 比特）

  RPR 采用 16 比特 HEC，其生成多项式为：

$$HEC_{16} = x^{16} + x^{12} + x^5 + 1 \qquad (2\text{-}1)$$

图 2-1 中的 RPR 有效载荷域（payload）包括 RPR 分组标签域和 MAC 用户数据域。RPR 帧的标签域是可选项，主要用于 RPR 帧的扩展。MAC 用户数据域用于承载用户数据，长度是可变的。帧校验序列（FCS）采用 32 比特的循环冗余校验（CRC）,其生成多项式为：

$$CRC_{32} = x^{32} + x^{26} + x^{23} + x^{22} + x^{16} + x^{12} + x^{11} + x^{10} + x^8 + x^7 + x^5 + x^4 + x^2 + x^1 + 1 \quad (2\text{-}2)$$

2.2.2 RPR 控制帧格式

本文将控制帧分为一般控制帧和公平消息帧两类。下面分别进行讨论。

1．一般控制帧

如果将 RPR 帧头类型域的值设置成 101（见表 2-2），则说明该分组是一般控制分组。RPR 控制分组可以采用逐跳或者指定（destined）的方式传送到特定结点，如果控制分组只经过一跳就可到达目的结点，那么控制分组不需要任何寻址信息。控制分组的目的地址可置为 0，源地址域设置为发送结点的地址。图 2-3 给出了 RPR 一般控制帧格式。

RPR HEADER（TYPE=0x05）
CONTROL TYPE
RPR CONTROL DATA
FCS

图 2-3 RPR 一般控制帧格式

需要指出的是，当 RPR MAC 发送控制分组的时候，RPR 帧头的 PRI 值为 7，这样 RPR 结点将优先转发和接收 RPR 控制分组。另外，帧头域中协议类型域应置为 0x2007。

表 2-3 控制类型

控制类型	描述
0x01	拓扑发现
0x02	保护消息
0x03	OAM 控制分组
0x04~0xFF	保留

RPR 一般控制帧格式与 RPR 帧的基本格式的最大区别在于 payload 域中增加了控制类型域。该 8 比特域指示控制消息类型。表 2-3 给出了当前已定义的控制类型。图 2-3 中 RPR 控制数据域的长度可变，取决于控制类型。RPR 拓扑发现帧（控制类型为 0x01）格式将在 2.4 节详细讨论，RPR 保护消息帧（控制类型为 0x02）格式将在 2.5 节详细讨论。

2．RPR MAC 公平消息帧格式

将 RPR 帧头类型域的值设置成 110，则表示该控制帧为公平消息帧。由于公平消息帧只发往上游相邻结点，传送本地结点带宽分配算法计算得到的公平消息。因而在公平消息帧格式中，去掉了目的地址域、协议类型域和 HEC，如图 2-4 所示。

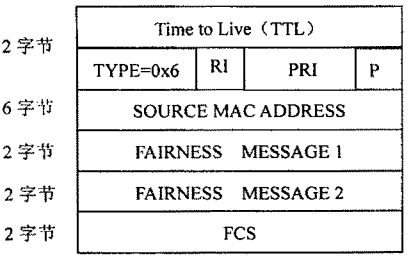

图 2-4 RPR 公平消息帧格式

这样设计的原因是环中公平消息分组相对于其它控制分组而言需要经常在各结点之间交互，因此公平消息帧过长，会引入不小的网络开销。和 RPR 一般控制帧相同，公平消息帧的 PRI 域亦设置为 7，公平消息 1 域携带公平消息。公平消息 2 域预留为今后扩展。

2.2.3 扩展的 RPR 帧格式

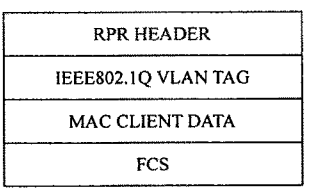

图 2-5 加 VLAN 标签的 RPR 帧格式

我们以加 VLAN 标签的 RPR 帧格式为例，说明 RPR 帧是如何扩展的。如图 2-5 所示，RPR 标头中协议域设置成 0x8100，表示该分组为加 VLAN 标签的分组。IEEE 802.1Q VLAN 标签域携带 VLAN 标签。

以上对 RPR 帧格式进行了初步的讨论。当然在 RPR 帧格式的制定过程中，会加入其它的功能域。尽管 RPR 帧格式并不是本文研究的重点，但是作为 RPR 关键技术之一、RPR 分组的载体，对其进行初步的探讨还是很有必要的。

§2.3 RPR 拓扑发现机制

2.3.1 概述

对于 RPR 而言，拓扑发现是必不可少的。在 RPR 环中，每个结点掌握着环的状态信息。平时，结点没有任何拓扑更新的信息，当环进行初始化、新结点加入、环保护倒换操作时，拓扑发现过程启动，发起结点向环中所有具有逻辑地址的结点发送拓扑发现消息，各结点根据该消息判断发生状态变化的结点和其链路状态。这样在很短的时间内，RPR 环上所有的结点都收集到环的状态信息，包括在环的两个方向上到达其它结点所需的跳数，环上每段链路的状态等。

环上每个结点定时的发送拓扑发现帧以便及时的获取最新的拓扑信息，从而确定结点具体的分布情况。各结点根据自己保存的拓扑图决定源和目的结点之间该如何选路，即确定到达目标结点的最短路径。拓扑发现还能确定环中哪些结点具有环回能力，以便在 RPR 提供保护时决定采用何种保护机制。

2.3.2 拓扑发现帧格式

18 字节	RPR HEADER（TYPE=0x05）
1 字节	CONTROL TYPE（0x01）
2 字节	TOPOLOGY LENGTH
6 字节	ORIGINATOR MAC ADDRESS
2 字节	MAC TYPE
nn 字节	OTHER MAC BINDINGS
4 字节	FCS

图 2-6 RPR 拓扑发现帧格式

RPR 拓扑发现帧属于 RPR 控制帧的一种，其帧格式如图 2-6 所示。

✧ 拓扑长度（TOPOLOGY LENGTH）域（2 字节）

拓扑长度域用来指示拓扑信息的长度，单位是字节。拓扑长度计算公式如下：

$$TOPOLOGY\ LENGTH = ORIGINATOR\ MAC\ ADDRESS + MAC\ TYPE + OTHER\ MAC\ BINDINGS \quad (2\text{-}3)$$

一旦环中结点接收到拓扑发现帧，在绑定其 MAC 地址的同时改变拓扑长度。

✧ 发起结点 MAC 地址（ORIGINATOR MAC ADDRESS）域（6 字节）

发起结点地址域存放发起结点的 MAC 地址。RPR 采用的是 IEEE 802 定义的全球唯一的地址，该域大小为 6 个字节。

✧ MAC 类型域（2 字节）

表 2-4 MAC 类型格式

比特位	含义
0	单转移缓存（0）/双转移缓（1）
1	环标识符（外环 0，内环 1）
2	环回结点（1）/非环回结点（0）
3	环回保护能力（1）
4	绕行保护能力（1）
5~7	公平消息
8	巨帧支持（1）
9~15	保留

MAC 类型域用于指示结点的特征。MAC 类型域中各个比特的含义如表 2-4 所示

✧ 其它结点 MAC 绑定（OTHER MAC BINDINGS）域

每个结点的 MAC 绑定由 MAC 类型域和本结点的 48 位 MAC 地址构成。

2.3.3 RPR 拓扑发现过程

RPR 拓扑发现过程如图 2-7 所示。环中各结点定时向内外环发送拓扑发现帧进行拓扑探测。发起拓扑探测的结点首先在拓扑发现分组中标记发送环标识符，指示该分组将注入到外环还是内环，然后将本地结点的 MAC 地址和 MAC 类型与拓扑发现分组绑定在一起，并在分组中按照（2-3）式设置好长度域。RPR 拓扑发现帧采用广播的方式传送，途中经过的各结点都将其 MAC 地址和 MAC 类型域与拓扑发现分组绑定，并更新长度域，再进行下一跳转发。如果环发生故障，处在环回状态，则离故障点最近的结点在进行 MAC 绑定时将会在拓扑发现帧中指示环处于环回状态，然后环回拓扑发现分组。被环回的拓扑发现分组由于当前传送的环标识符与最初设置不同，因而途中经过的结点的 MAC 地址不会绑定到分组中。

最终拓扑发现帧由发起结点摘除，在接受分组之前，发起结点必需判断拓扑发现帧发送与接收的环标识符是否一致。我们在介绍拓扑发现帧格式时，其中有一个 MAC 类型域，其第 2 比特代表环标识符，见表 2-4。MAC 类型域中这一比特正是用于判断接收的拓扑发现分组和发送环标识符是否匹配。具体判断方法为：当发起结点收到由自己发送的拓扑发现帧之后，判断发送该帧之前设置的 MAC 类型中环标识符与最后一个绑定到该帧的结点的 MAC 类型域中环标识符是否一致。

环中每结点均使用从拓扑发现分组获得的拓扑信息构造拓扑图，只有当连续两次接收到相同的拓扑信息时，拓扑图才会更新。这样设计是为防止网络状态的瞬时变化引起网络拓扑的变化。除了周期性进行拓扑探测外，网络拓扑信息也会在

环中结点接收到保护倒换请求消息或者检测到链路故障时进行更新。需要注意的是拓扑图中只包含可达结点。

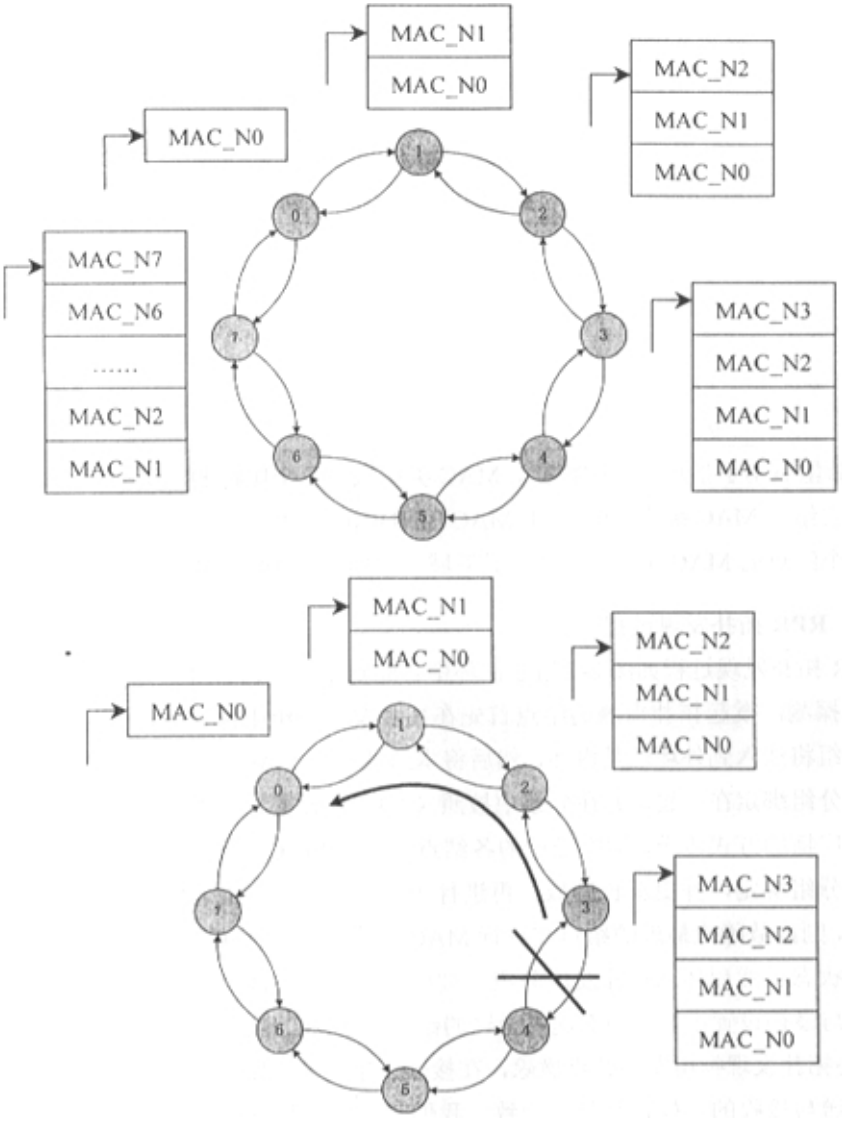

图 2-7 RPR 拓扑发现过程

这里可能会产生混淆，似乎拓扑探测是环保护机制的一部分。实际上，我们可以认为拓扑发现是为 RPR 公平带宽分配机制和 RPR 保护倒换机制服务的。通常，环中各结点独立的发送和接收拓扑发现帧，根据其携带的拓扑信息统计环中的结点数，构造拓扑图。当拓扑探测过程完成后，环中各结点就了解到其它结点的特性，例如哪些结点支持“环回”保护，哪些结点支持“绕行”保护等等。另外，根据拓扑信息，本地结点能够计算它到环中各结点的距离（跳数）。当拓扑信息与

公平带宽分配机制结合后，RPR MAC 不但可以确定本地业务到达各结点的“最优”距离，而且还能知道每业务流传送的具体路径，包括途经的结点和链路。这些信息在进行带宽分配时非常重要。

§ 2.4 RPR 保护倒换机制

弹性是 RPR 技术的特点之一。其目标是当环中链路或结点发生故障，RPR 可以提供 50ms 内的保护。目前采用两种保护机制，“环回”和“绕行”。RPR 保护倒换将支持这两种机制。

在拓扑发现的过程中，每个结点将在拓扑发现分组中指明其支持哪类保护机制。如果环中所有结点都支持“环回”保护，那么 RPR 保护倒换将采用“环回”保护机制，否则将采用“绕行”保护机制。

2.4.1 “环回”保护

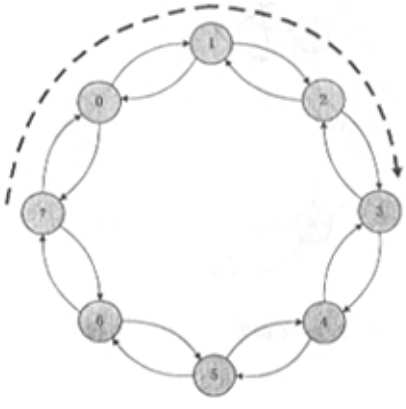

图 2-8a 环路发生故障前数据流

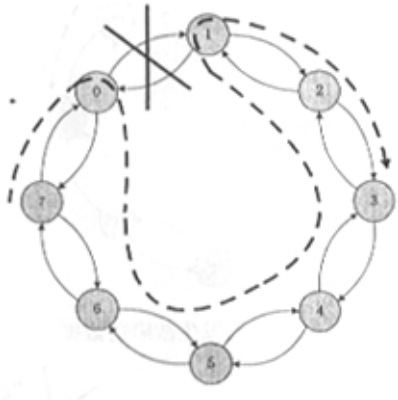

图 2-8b 环路环回保护时数据流

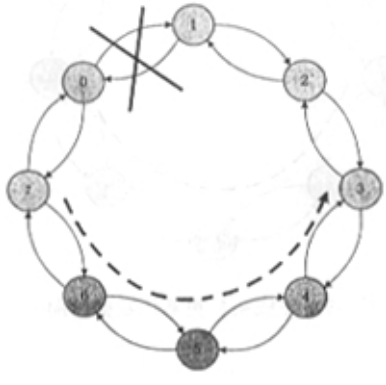

图 2-8c 完成拓扑更新后数据流向

RPR 环是由两个互为反方向传送数据的单光纤环构成。一旦 RPR 检测到环中某结点或链路发生故障，则沿故障链路或故障结点方向传送的业务将会在靠近故障点的结点处“环回”到另一环上，沿相反的方向传送。采用“环回”这样的方式可以重路由业务远离故障点。

下面举例说明。图 2-8a 所示的是链路故障发生之前，结点间数据传送的情况。此时结点 7（N7）沿通路 N7—>N0—>N1—>N2—>N3 向结点 3（N3）发送数据。如果 N0 与 N1 之间的链路发生故障，则 N0 和 N1 将进入环回状态，这时从 N7 发往 N3 的业务将沿着非优化链路 N7—>N0—>N7—>N6—>N5—>N4—>N3—>N2—>N1—>N2—>N3 传送，如图 2-8b 所示。随后当环中各结点新的拓扑图更新完毕之后，N7 发往 N3 的业务将沿着新的优化通路 N7—>N6—>N5—>N4—>N3 传送，如图 2-8c 所示。

2.4.2 “绕行”保护

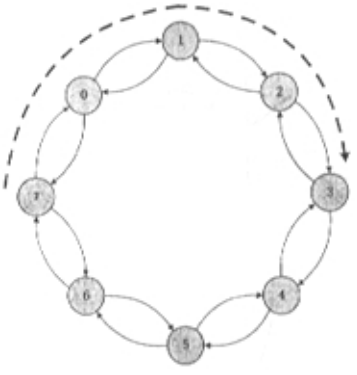

图 2-9a 环路发生故障前数据

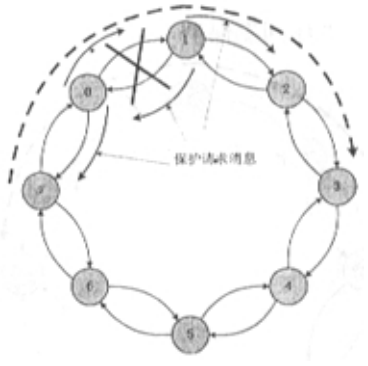

图 2-9b 环路发生故障时数据流向

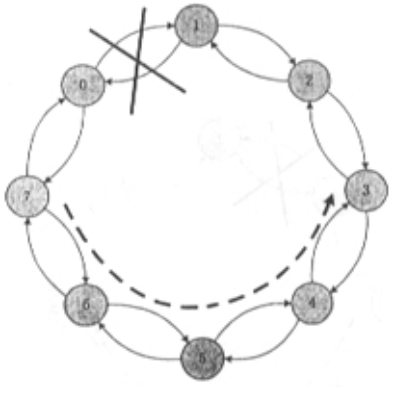

图 2-9c 完成拓扑更新后数据流向

“绕行”保护是 RPR 采用的另一种保护机制，和“环回”保护不同，“绕行”

保护并不将业务从故障点处环回。当 RPR 检测到环中某段链路发生故障时，离故障点最近的结点立即向环中各个结点发送保护请求消息，指示环路发生故障。接收到保护请求消息的结点相应的更新拓扑图。在拓扑图更新后，原先途经故障点的业务将会重新路由，绕开故障点，沿新的通路传送。

图 2-9a 所示的是环路发生故障之前，结点间数据传送的情况。此时 N7 沿通路 N7—>N0—>N1—>N2—>N3 向 N3 发送数据。当 N0 与 N1 之间的环路发生故障，结点 N0 与 N1 立即向环中其它结点发送保护请求消息，指示环路发生故障。如图 2-9b 所示，接收到保护请求消息的结点相应的更新拓扑图。在拓扑图更新后，N7 发往 N3 的业务将沿着新的优化通路 N7—>N6—>N5—>N4—>N3 传送，如图 2-9c 所示。

需要说明的是，若 RPR 采用“绕行”保护机制，在拓扑更新完成之前，已注入环中的分组将会在故障点处被丢弃掉，丢弃分组的数量将取决于拓扑更新的速度。对于 RPR 而言，能够确保在 50ms 内完成拓扑更新，尽管分组的丢弃会影响业务的 QoS，然而 RPR 能够把这样的影响降到最小。正是由于“绕行”保护机制存在分组丢弃现象，因此 RPR 通常采用“环回”和“绕行”相结合的方式进行对数据的保护。RPR 环一旦检测到环路出现故障，先采用“环回”方式，确保已发送的分组不丢失，在拓扑更新完成后，再采用“绕行”方式，指引分组绕过故障点，重路由到达目的结点。

2.4.3 保护倒换消息帧格式

RPR 保护倒换机制的目标是能够从各种各样的链路故障或者链路劣化的情况下，自动的对数据进行保护。要完成这一目标，需要定义一个合适的保护消息帧格式。RPR 保护倒换消息帧格式如图 2-10 所示。

18 字节	RPR HEADER（TYPE=0x05）
1 字节	CONTROL TYPE（0x02）
1 字节	PROTECTION MESSAGE
1 字节	RESERVED
4 字节	FCS

图 2-10 RPR 保护倒换消息帧格式

RPR 标头类型域置为 0x05，表示该分组为 RPR 控制分组。标头中目的地址域设置成全 1，表示广播，控制类型域置为 0x02，表示该控制帧为保护倒换消息帧。保护消息域携带详细的保护信息。保护信息域的格式如表 2-4 所示。

表 2-4 保护消息域格式

比特位	含义
0~3	保护消息请求类型 1101-强迫倒换（FS） 1011-信号故障（SF） 1000-信号劣化（SD） 0110-人工倒换（MS） 0101-等待恢复（WTR） 0000-无请求（IDLE）
4	通路指示符 0-短通路（S） 1-长通路（L）
5~7	状态码 010-保护倒换完成（Wrapped） 000-空闲（IDLE）

✧ 保护请求消息类型

RPR 的保护请求类型按照优先级的高低分为 6 种，分别为强迫倒换（FS）、信号故障（SF）、信号劣化（SD）、人工倒换（MS）、等待恢复（WTR）和无请求（IDLE）。下面分别对这几种保护请求类型做一介绍：

1. 强迫倒换（FS）：强迫倒换由控制台发起。发送 FS 消息的结点将在本结点处环回本地业务和转发业务，与此同时发起结点指引 FS 消息帧到达与之相邻的结点。当相邻结点接收到 FS 消息帧时，立即进入环回状态。FS 主要用于受控方式下，向环中添加新的结点。
2. 信号故障（SF）：信号故障消息是自动发起的。该保护主要是由于媒体“硬故障（例如光纤断裂等）”引起的。在 SONET 系统中，SF 通常由以下原因触发：信号丢失（LOS），帧丢失（LOF），链路误码率（BER）超过 SF 预设的门限等。
3. 信号劣化（SD）：信号劣化消息是自动发起的。该保护主要是由于媒体“软故障”引起的。在 SONET 系统中，SD 通常由以下原因触发：链路误码率或通路误码率超过 SD 预设的门限。
4. 人工倒换（MS）：人工倒换和 FS 一样是由控制台发起，优先级较低。
5. 等待恢复（WTR）：等待恢复消息是自动发起的。当 SF 或 SD 情况消除后，环路检测到当前状态已经满足恢复的门限值，靠近故障点的结点将发起 WTR 消息，通知环中其它结点环路将要恢复正常。这里需要注意的是，

只有在 WTR 定时器超时后，保护状态才结束。这样做是为了防止保护倒换震荡。

6. 空闲（IDLE）：空闲时或者当 WTR 定时器超时后发送，指示环路处于正常状态。

高优先级的保护消息帧可以抢占低优先级保护消息帧的处理，例如当环路正在进行 SD 保护时，环中其它结点发送 SF 消息，则环路立即进行 SF 保护。

✧ 保护消息通路指示符

根据消息发送的方向，保护消息可分为两种类型——长消息类型和短消息类型。所谓“长”与“短”并不是针对消息的长短，而是根据消息传送通路的长短命名的。

✧ 保护状态

RPR 结点在保护倒换的过程中，只有两个状态：一是空闲（idle）状态，该状态表示结点准备好可以随时执行保护倒换；另一个是环回（Wrapped）状态，该状态表示结点已加入到环回保护之中。

2.4.4 工作过程示例

前面我们从 RPR 保护倒换机制、保护消息帧格式以及保护消息类型等方面对 RPR 保护倒换做了初步的探讨。下面将以双环故障为例说明 RPR 保护倒换机制的工作过程。为了叙述方便起见，我们用如下的格式表示 RPR 保护消息：

{ REQUEST_TYPE, SOURCE_ADDRESS, WRAP_STAT, PATH_INDICATOR }

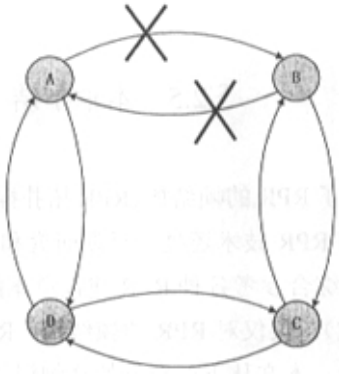

图 2-11 双环故障的情况

如图 2-11 所示，RPR 环由四个结点 A、B、C 和 D 构成，结点 A 和 B 之间环路发生故障（双环）。在故障发生之前，环中所有结点保护状态均处于空闲状态。

1. 结点 A 检测到内环出现信号故障（SF），立即从空闲状态转为环回状态。并沿外环（短路径）向结点 B 发送保护倒换消息帧{SF,A,W,S}，与此同时

沿内环（长路径）向结点 B 发送保护倒换消息帧{SF,A,W,L}。

2. 结点 B 检测到外环出现信号故障（SF），立即从空闲状态转为环回状态。并沿内环（短路径）向结点 A 发送保护倒换消息帧{SF,B,W,S}，与此同时沿外环（长路径）向结点 A 发送保护倒换消息帧{SF,B,W,L}。
3. 环路达到稳态后，环中各结点更新拓扑表，重新选择优化通路传送数据。

当结点 A 与 B 之间的链路恢复正常后，执行以下过程：

1. 结点 A 检测到 SF 已经清除，仍旧保持环回状态，设置 WTR 定时器，沿外环（短路径）向结点 B 发送保护倒换消息帧{WTR,A,W,S}，同时沿内环（长路径）向结点 B 发送保护倒换消息帧{WTR,A,W,L}。
2. 结点 B 检测到 SF 已经清除，尽管此时它很可能已接收到从短路径传来的 WTR 消息帧，但并不马上进入 IDLE 状态，也设置 WTR 定时器，沿内环（短路径）向结点 A 发送保护倒换消息帧{WTR,B,W,S}，与此同时沿外环（长路径）向结点 A 发送保护倒换消息帧{WTR,B,W,L}。
3. 结点 C 和 D 传递长消息帧，并不改变其中的内容。
4. 环路达到稳态
5. 如果结点 A 的 WTR 定时器超时，A 进入空闲状态，然后分别向内、外环发送空闲消息帧{IDLE,A,I,S}和{IDLE,A,I,L}
6. 结点 B 接收到来自短路径的空闲消息后，进入空闲状态。（这里假设结点 A 的 WTR 定时器先超时，如果结点 B 先超时，处理过程类似）
7. 环路达到稳态，环中各结点更新拓扑表，重新选择优化通路传送数据。

§2.5 本章小结

本章我们初步讨论了 RPR 的帧结构、RPR 拓扑探测机制和 RPR 保护倒换机制。需要指出的是由于目前 RPR 技术还处于早期研究和探索阶段，许多关键技术还有待进一步研究。本章在综合参考各种 RPR 建议草案的基础上，按照 RPR 的具体特点和要求（绪论中所述），仅仅对 RPR 的帧结构、RPR 拓扑探测机制和 RPR 保护倒换机制进行初步讨论。本文从下一章开始详细讨论 RPR 技术的重中之重—RPR 公平带宽管理机制。

第三章 弹性分组环公平带宽管理机制的研究

§3.1 前言

RPR 网络是媒质共享的环形拓扑网络，因此带宽管理成为媒质接入控制（MAC）的重要研究课题。对于共享媒质的 RPR 环形网络，每一环段既要承载本地用户的业务又要承载来自上游其它结点用户的业务。因此，除非上游结点 MAC 层对其本地业务的注入速率加以控制，否则本地业务将占用比应得的公平带宽多得多的带宽。这种对带宽不加控制的侵占会引起网络拥塞，进而会剥夺下游用户接入媒质的机会。因此，RPR MAC 层需要采取一定的带宽管理机制，来公平的分配各结点用户业务的带宽。有效的带宽管理不仅可以防止网络拥塞而且可以提高网络带宽利用率。

为了实现环的最大带宽利用率，还必须考虑环路的空间重用特性，即当一个结点的 MAC 层收到并从环上摘除单播数据帧后，空闲的带宽可让给其下游结点的 MAC 用户使用。而只支持单缓存（FIFO）访问的 MAC 通常会发生队头阻塞（Head of Line Blocking）。所谓队头阻塞是指一个准备发向某个未拥塞结点的数据帧可能会等候在因为目的结点拥塞而不能访问环路的 FIFO 队头数据帧的后面。在该队头数据帧从 FIFO 队列中移出之前，它后面的所有数据帧都被阻塞。队头阻塞极大的妨碍了环的空间重用特性，严重降低环路带宽利用率。在下一章我们将通过仿真对队头阻塞现象进行分析。解决队头阻塞问题的一个公认方案是采用虚拟输出队列（VOQ），RPR MAC 的用户可以通过为每一目的结点维护一个 FIFO 来实现 VOQ，采用基于每一个目的地址的多个缓存器，发往不同目的地址的数据帧不会被发往其它目的地址的数据帧所阻塞，因此可以完全避免队头阻塞。而要采用 VOQ，带宽管理必须能够对具有不同目的地址的本地用户业务进行独立的接入控制，必须告诉本地用户每个目的结点的可用带宽。

因此，为了支持最大的环路空间重用以及带宽的公平分配，无论 RPR MAC 用户是否具有支持 VOQ 的能力。RPR MAC 必须能够根据每目的地址对本地业务进行公平的链路带宽分配和有效的媒质接入速率控制。

§3.2 基于输入聚合的 RPR MAC 参考模型

3.2.1 一种新的 RPR MAC 参考模型

公平带宽管理机制是建立在 MAC 层架构之上。这里，首先给出一个新的 RPR MAC 参考模型——基于输入聚合（Ingress Aggregation）的 RPR MAC 参考模型。记为 IA 参考模型，如图 3-1 所示。RPR MAC 内外环的结构是一样的，简单起见，这里只给出外环的结构。

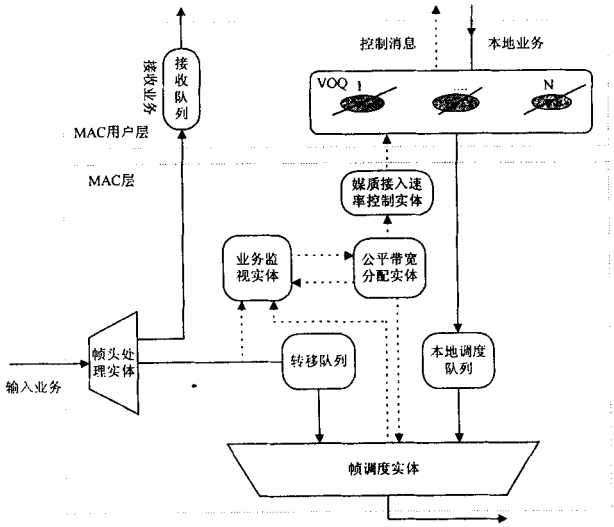

图 3-1 IA 参考模型

IA 参考模型由 5 大功能实体构成，根据功能可分为两类：一类是传送通路部分，包含帧头处理实体和帧调度实体；另一类是带宽管理部分，包含业务监视实体，公平带宽分配实体和媒质接入速率控制实体。下面分别介绍各实体的具体功能：

1. 帧头处理实体

帧头处理实体用于检测 RPR 帧头，判断上游传送来的帧是本地接收帧还是需要转发的中继帧。同时该实体还用于确定转发帧的源结点地址和帧的长度，供业务监视实体使用。另外，帧头处理实体具有循环分组检测和摘除无效帧的能力,即如果帧头处理实体检测到接收帧的生存期（TTL）小于 0 或者在非环回（Wrapped = 0）的状态下，接收帧的源地址与本地地址相同（循环帧）。则帧头处理实体将会把这些无效的帧从网上摘除并丢弃。

2. 帧调度实体

帧调度实体用来解决本地帧、转移帧和 RPR 公平消息帧（RPR-fa-RCM）之间因同时接入而引起的链路竞争问题。调度策略在 3.2.2 节详细介绍。

3. 业务监视实体

业务监视实体用于跟踪测量所有上游结点的业务通过该结点输出链路的速率以及该结点本地业务注入到输出链路的速率。该实体根据测量的结果周期产生 RPR 速率测量信息以更新原先的测量值。

4. 公平带宽分配实体

公平带宽分配实体负责接收来自业务监视实体的速率测量信息。根据消息的内容，采用适当的公平算法按照权值对本地可用带宽进行重新分配，将计算所得的 RPR 公平速率控制消息（RPR_fa_RCM）与来自下游结点的 RPR_fa_RCM 进行比较，选择合适的速率控制消息一方面发往媒质接入速率控制实体，供媒质接入速率控制实体使用该信息对用户业务进行整形，控制每目的业务的接入速率（用户端采用 VOQ 或类似机制）；另一方面将该消息组帧发往帧调度实体，经由帧调度实体调度发往该结点的上游结点。

5. 媒质接入速率控制实体

媒质接入速率控制实体负责控制媒质接入的速率，防止用户端接入的业务超过已分配的链路带宽。在 MAC 侧，当媒质接入速率控制实体接收到来自公平带宽分配实体发来的 RCM，就根据 RCM 中的内容对用户端业务进行整形，动态调整每目的业务的接入速率。

3.2.2 IA 参考模型转移通路的基本设计和操作模式

前面按照功能我们把 IA 参考模型按照功能分为转移通路和带宽管理两部分，并对每部分中各功能实体进行了介绍。这里将详细讨论转移通路的基本设计和操作模式。IEEE 802.17 工作组要求各方提交 RPR MAC 草案时，考虑在满足 RPR 基本功能和特点的同时，尽量使 RPR MAC 简化。因此我们在设计 RPR 转移通路的时候，参照以下目标和要求：

- 转移通路是共享媒质的一部分
- 转移通路应无损（lossless）
- 转移通路中的转移缓存应该优化，尽量的小，其目的是：
  - ✧ 从 RPR MAC 成本的角度考虑，节省内存消耗
  - ✧ 减少分组在转移通路中的时延

图 3-2 给出了 RPR 转移通路的功能模型，可以看出，基于 IA 参考模型的 RPR MAC 层的转移通路由帧头处理实体、转移缓存和帧调度实体构成。转移通路可以看成是媒质的扩展，因此可视为环的一部分。RPR MAC 帧头处理实体完成 RPR 帧的接收和转移操作，转移缓存专门用来避免注入（Inserting）和转发（Forwarding）

帧之间的冲突碰撞。帧注入调度实体采用一定的规则调度注入帧和转发帧。

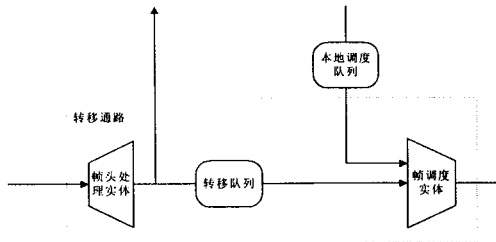

图 3-2 RPR 转移通路模型

1. RPR MAC 接收规则

如图 3-2 所示，当一个帧到达 RPR MAC 的输入端口时，首先在帧头处理实体中进行 RPR MAC 地址适配检测，确定该帧是被接收还是被转发。具体的检测过程如下：

- ✧ 如果接收帧的目的地址与该结点的 MAC 地址匹配，则摘除该帧，判断：
  1) 如果该帧是 RPR-fa-RCM 帧，则发往业务监视实体
  2) 如果该帧是 RPR 数据帧，则发往 RPR 用户端
- ✧ 如果接收帧的目的地址是广播（组播）地址，则判断：
  1) 如果 TTL > 1，摘除该帧并拷贝，原帧发往转移队列，副本发往 RPR 用户端，同时原帧帧头 TTL 域减 1
  2) 如果 TTL = 1，摘除该帧并发往 RPR 用户端
     - ✧ 如果接收帧的源地址与该结点的 MAC 地址匹配，则摘除该帧，并丢弃。
     - ✧ 如果接收帧的 RPR 帧头 CRC 校验出错，则摘除该帧，并丢弃；同时 CRC 出错计数器加 1。当 CRC 出错计数器连续累加超过预设的门限，例如上游链路出现信号劣化（SD），则该结点将发送保护倒换消息，整个环网将进入保护状态，直到其上游链路恢复正常。详细的保护倒换过程参阅 2.4 节。
     - ✧ 如果接收帧的目的地址与该结点的 MAC 地址不匹配，则判断：
       1) 如果 TTL <= 1，摘除该帧，并丢弃；
       2) 如果 TTL > 1，转发该帧，同时 RPR 帧头 TTL 域减 1。

2. RPR MAC 的转移（Transit）规则

这里分为两种情况：一种是采用单转移缓存，另一种是采用双转移缓存。这两种情况的主要区别是：如果采用单转移缓存，则输入的转移帧一般不考虑优先级高低按照先后顺序进行缓存。尽管采用单转移缓存实现起来比较简单且节约成本，

但是其缺点也是显而易见的，由于不同优先级帧进入队列后顺序的进行处理，因此对于某些高优先级业务（例如时延敏感型业务），不能确保端到端时延的要求。而采用双转移缓存，将高优先级帧与低优先级帧分开缓存，且优先调度高优先级缓存中的帧。尽管实现稍微复杂了一些、成本稍微增加了一些，但是其优点是能够确保高优先级业务的服务质量（QoS）。正是由于采用双转移队列可以带来传送性能上的改善，因此本文的转移规则是按照双转移缓存的情况设计的。

转发帧首先按照优先级高低发送到高优先级或者低优先级转移缓存。这里，用 TB_HI 表示高优先级转移缓存，TB_LO 表示低优先级转移缓存。帧调度实体执行以下调度算法：

- ✧ 第一步：判断 TB_HI 是否为空，如果不空，则发送 TB_HI 中的帧，转第一步；否则转第二步
- ✧ 第二步：采用轮寻的策略对两个队列中的帧进行调度，转第三步
- ✧ 第三步：完成发送，转第一步

这里有几点说明：

- ✧ 由于首先调度 TB_HI 中的帧，因而 TB_HI 的长度不应太大，太大不仅浪费而且也没有必要，也不会带来性能的提升。
- ✧ 为了保证帧传输的连续性，帧的传输采用非强占（Nonpreemptive）方式，即当低优先级帧正在发送时，高优先级分组不能强占。

3. RPR MAC 传输（Transmit）规则

为了简化 RPR MAC 层的复杂度，本文将本地传输缓存设置在 RPR 用户层。本地业务在其接入 MAC 层之前，先在用户层排队。在 MAC 层中设计了一个调度队列（Scheduling queue），用于缓存来自 RPR 用户的帧。当满足以下条件时，RPR MAC 发送调度队列中的帧。

- ✧ 如果 TB_HI 和 TB_LO 均为空，则发送调度队列中的帧；
- ✧ 如果 TB_HI 为空，且 TB_LO 不空，则采用轮寻的策略调度两个队列中的帧。

4. RPR MAC 丢弃规则

RPR MAC 遵循以下的丢弃规则：

- ✧ RPR 帧头校验出错
- ✧ 如果 RPR 帧中源地址与接收结点的 MAC 地址匹配，表示该帧又返回发送端，则丢弃该帧
- ✧ TTL 过期，即 TTL < 0

以上 RPR 接收规则、转移规则、传输规则和丢弃规则，是针对 RPR 转移通路的目标和要求进行设计的，其中接收规则和丢弃规则共同实现帧头处理实体的功能，而转移规则和传输规则共同实现帧调度实体的功能。

3.2.3 IA 参考模型公平带宽管理的基本设计

由于媒质是共享媒质，所以带宽管理是 RPR MAC 中的一项必备的功能。因环中每个结点在地理位置上彼此分离，为了保持（maintain）和确保（guarantee）所有结点在竞争带宽资源时能够公平的接入媒质，环中各结点必须执行某种接入策略—带宽管理算法。

执行带宽管理算法的目标是：

- ✧ 确保各结点能够公平接入
- ✧ 在公平接入的前提下，获得最高的带宽（BW）利用率
- ✧ 最大化环的空间重用特性
- ✧ 支持区分服务（Diffserv），环中各结点间相同类型业务可以获得相同的性能，不同类型的业务根据业务类型可能获得较好的或者较差的性能。性能的测度（metrics）包括网络的吞吐率、端到端（end to end）时延和时延抖动等。
- ✧ 支持 RPR 用户具备虚拟输出排队（VOQ）结构或者相似结构以消除单 FIFO 传送不同目的结点的分组时可能产生的队头阻塞（HOL）。

为了实现上述目标，RPR MAC 必须具有动态回收已被系统释放的带宽，且在发送结点之间公平的重新分配这些带宽资源的能力。所以为了满足这一需要，本文设计了以下基本的 RPR MAC 带宽分配功能准则：

- ✧ 计算公平带宽。公平带宽的计算在环中各结点中独立进行。在本文设计的带宽分配算法中，公平带宽的计算是通过本地可分配带宽结合下游结点传来的公平信息一起完成的。
- ✧ 向上游结点发送公平信息。当环中某结点计算完公平带宽之后，将会以消息分组的形式将结果传送给上游结点，供上游结点计算公平带宽之用。
- ✧ 根据计算的公平带宽，警管本地业务的注入速率。环中各结点必需按照已计算的公平带宽的结果严格的控制本地业务的注入速率。
- ✧ 检测拥塞的发生。由于环中业务注入的实时性和系统状态的不确定性，在某些时候很可能会由于环中业务的突然变化导致某些结点产生拥塞。因此，RPR MAC 必需能够尽可能快的检测拥塞是否即将发生，并执行相应的机制避免拥塞的发生。

以上四点功能主要通过下面三个功能实体实现：

1. 业务监视实体，该实体的功能是：
   a) 监视每活动源（per-active source）的链路利用率，目的是检测链路状态是否拥塞。为了保证环路的无损性，即一旦分组注入到环中，除非 TTL 超时或 CRC 校验出错，否则应确保分组在传送的时候不丢失。RPR MAC 采用拥塞避免机制来防止拥塞的发生。拥塞避免机制是一种

主动机制，即它不能在拥塞已经发生的时候才起作用。拥塞避免是通过业务监视实体不断的监视其上游结点以及本地业务的接入速率，及时的向公平带宽分配实体反馈当前的链路状态，由公平带宽分配实体发送公平消息指示上游结点调整其业务接入速率而实现的。

b) 接收来自下游结点的公平消息分组，取出公平信息后传给公平带宽分配实体。

2. 公平带宽分配实体，其主要功能是：

a) 接收来自业务监视实体的公平消息分组；

b) 结合下游结点传送来的公平信息，每隔 T 秒计算一次（加权）公平带宽，将计算出的结果发往媒质接入速率控制实体；

c) 将计算出的公平信息封装成 RPR-fa-RCM 帧发往帧调度实体等待发送。这里需要说明的是，RPR-fa-RCM 帧在发送之前，为了避免与其它分组发生碰撞，首先进入控制消息队列进行缓存（简化起见，图 3-1 中没有画出），再由帧调度实体进行调度。

3. 媒质接入速率控制实体，其主要功能是：

a) 根据公平带宽分配算法计算出的结果，警管本地业务的接入速率。

b) 如果 RPR 用户支持 VOQ，该实体应向 RPR 用户发送速率控制消息（RCM），指示上层实体根据消息内容调整发往不同目的结点业务的速率。

以上我们讨论了 IA 参考模型中实现带宽管理功能的三个功能实体的基本要求和实现方法。下一节我们将通过设计基于 IA 参考模型的分布式环带宽分配——DBRR，进一步讨论 RPR 带宽管理。

§3.3 基于 IA 参考模型的环带宽分配算法——DBRR

3.3.1 GPS 模型

GPS 是一个流体算法。在某一个时刻如果一个连接在系统中有业务等待输出，我们称此时该连接被积压（backlog）。GPS 中每个连接都分配有一个权值，通常情况下，这个权值表示该连接应获得的服务带宽份额。GPS 的服务规则是：在任意时刻 GPS 服务器为被积压的连接按照权值并行服务。令 ϕ_i 表示连接 i 的权值，C 表示输出链路的速率，如果 t 时刻连接 i 被积压，则此时 GPS 服务器为连接 i 服务的速率为：

$$r_i = \frac{\phi_i}{\sum_{k \in B(t)} \phi_k} \cdot C \tag{3-1}$$

其中 B(t)为 t 时刻系统中被积压连接的集合。

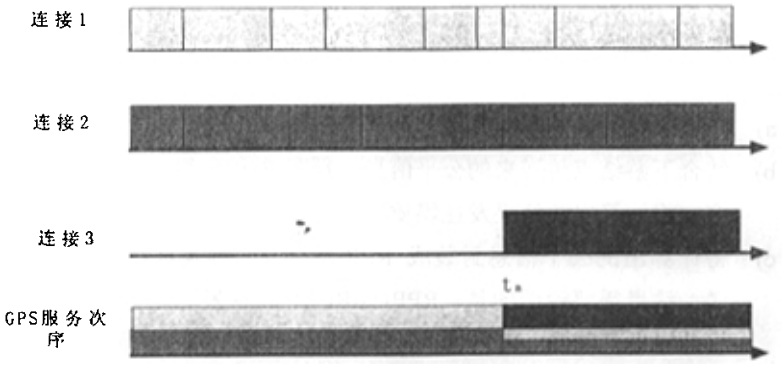

图 3-3 GPS 服务模型

图 3-3 中给出了 GPS 服务器的一个例子，其中 $\phi_1 = \phi_2 = \frac{1}{8}, \phi_3 = \frac{1}{4}$。在 t_a 时刻之前，只有连接 1、连接 2 存在积压分组，它们分享系统带宽，服务速率均为 $\frac{C}{2}$。在 t_a 时刻之后，连接 3 也被积压，于是三个连接的服务速率分别为 $\frac{C}{4}, \frac{C}{4}, \frac{C}{2}$。可以看出，在 GPS 算法中各连接是根据其权值共享系统带宽的。

GPS 算法有三大特点：

第一，只要连接 i 被积压，GPS 服务器就保障以一定的速率为其服务。设 V 为 GPS 系统中复用的连接集合，则连接 i 的服务速率不低于 $\frac{\phi_i}{\sum_{j \in V} \phi_j} C$。

第二，当连接 i 的业务流符合漏桶约束时，由 GPS 服务器构成的网络能够保证连接 i 端到端的时延上限。

第三，GPS 具有最佳的公平性。令 $SW_i(t_1,t_2)$ 为连接 i 在 $[t_1,t_2]$ 期间获得的服务。如果连接 i、连接 j 在 $[t_1,t_2]$ 期间都被积压，则 $\left| \frac{SW_i(t_1,t_2)}{\phi_i} - \frac{SW_j(t_1,t_2)}{\phi_j} \right| = 0$，即连接 i、连接 j 获得的归一化服务量相等。

当 GPS 与漏桶约束控制机制相结合时，能够在相当大的范围内确保网络最坏

情况下吞吐量、公平性和时延的要求。然而目前的建议草案[3，4，5]没有一个能够同时获得在网络吞吐率、公平性和时延三者之间性能上的完美折衷。因此我们设想如果将 GPS 服务规则应用于环的带宽分配上，很可能取得在网络吞吐率、公平性和时延三者之间较好的性能折衷。

GPS 系统一般采用虚时间的方法进行分析。下面我们引入 GPS 系统虚时间（virtual time）的概念。首先给出两个定义：

定义 1：只要系统中有分组在等待服务，输出链路就不间断的输出分组，我们称这样的系统为尽职服务系统，这样的服务为尽职服务（work-conserving）。

定义 2：称系统不间断进行服务的任意一个时间间隔为一个忙期（busy period）。

本文设计的 RPR MAC 系统为尽职服务系统。只要转移队列或本地调度队列中有分组等待服务，则系统就不间断的输出分组。显然，对于尽职服务的系统，忙期只与业务的到达模式和服务速率有关，与系统的排队机制无关。

假设服务器的服务速率为 1，记分组每到达或离开 GPS 服务器一次为一个事件，设 t_j 为第 j 个事件发生的时刻（同时发生的事件先进行任意的排序）。设在一个忙期内第一个到达事件发生的时刻为 t_1=0，当 $j=2,3,....$，在时间间隔 $(t_{j-1},t_j]$ 内系统中处于积压状态的连接个数是不变的。记由这些处于积压状态的连接构成的集合为 B_j。定义 $v(t)$ 表示 GPS 系统的虚时间，考虑在任意一个忙期内，设其开始的时刻为 0，虚时间 $v(t)$ 可表示为[6]：

$$v(0)=0$$

$$v(t_{j-1}+\tau)=v(t_{j-1})+\frac{\tau}{\sum_{k\in B_j}\phi_k}$$

$$\tau\le t_j-t_{j-1},j=2,3,... \qquad (3\text{-}2)$$

式中 ϕ_k 表示连接 k 的权值，B_j 为在时间间隔 $(t_{j-1},t_j]$ 内处于积压状态连接的集合。将（3-2）式进行简单的变形，即：

$$v(t_{j-1}+\tau)-v(t_{j-1})=\frac{\tau}{\sum_{k\in B_j}\phi_k} \qquad (3\text{-}3)$$

（3-3）式表明在时间间隔 $(t_{j-1},t_j]$ 内，虚时间的变化量与处于积压状态的连接获得的服务量相同。由（3-3）式进一步得到虚时间的变化率为：

$$\frac{\partial v(t_j+\tau)}{\partial\tau}=\frac{1}{\sum_{k\in B_j}\phi_k} \qquad (3\text{-}4)$$

由（3-4）式，可以看到虚时间的变化率与事件发生时刻 t_j 无关，故（3-4）式可以写成：

$$\frac{dv(t)}{dt}=\frac{1}{\sum_{k\in B(t)}\phi_k} \tag{3-5}$$

式中，$B(t)$为 t 时刻系统中被积压连接的集合。结合（3-1）式和 C 为 1 的假设条件，有：

$$r_i(t)=\phi_i\cdot\frac{dv(t)}{dt} \tag{3-6}$$

（3-6)式给出了 GPS 服务器为连接 i 服务的速率(也就是为连接 i 分配的带宽)$r_i(t)$与系统虚时间变化率$\frac{dv(t)}{dt}$之间的关系。可以看出，$r_i(t)$与$\frac{dv(t)}{dt}$成正比，一旦确定$\frac{dv(t)}{dt}$则立即可以求得$r_i(t)$。

3.3.2 DBRR：一种基于 IA 参考模型的环带宽分配算法

上一节，我们介绍了 GPS 的概念，可以看出 GPS 的性能是相当理想的，然而 GPS 是一个流体模型，在实际应用中无法实现[6]，因此人们研究了一些调度算法[6，7，8，9，25]去逼近它。但是这些算法也存在一些问题，要么算法复杂度太高[6，8，9]，不适合高速网络，要么尽管算法复杂度降低，但不能保证公平性和端到端时延特性[7]。本文认为现存的基于 GPS 参考模型的调度算法并不适合于环形拓扑。因而很有必要设计一个为环形拓扑优化的算法。

本文设计的算法正是借鉴了 GPS 的基本思想，各个结点独立的计算虚时间，从而确定本结点公平带宽。和[6，7，8，9，25]以分组为最小颗粒计算虚时间或虚时钟（virtual clock）不同，本文提出的 DBRR 算法的关键技术是在 RPR MAC 层中以输入聚合流为最小颗粒构造一个虚时间代理，计算出的虚时间不是用于调度分组的传输而是用于确定每连接的公平速率。通过将计算出的信息分发给环中其它结点，环中各结点可以在本地计算出其在下游结点的公平速率，然后调整本地发往其它目的结点的业务速率为公平速率。

我们先讨论理想情况下 DBRR 算法—单结点的情况。多结点的情况在下一节讨论。下面介绍 DBRR 的几个特点，并由此给出单结点情况计算公平速率的方法。

一、间接带宽控制

间接带宽控制是 DBRR 的特点之一。所谓间接带宽控制是指上游 MAC 速率控制器受下游结点传来的公平信息“控制”。对于发送公平信息的结点，其本地业务和经由其转发的上游结点的业务可以看成是在 GPS 系统中以一定的速率接受服务的不同连接，如图 3-4 所示。

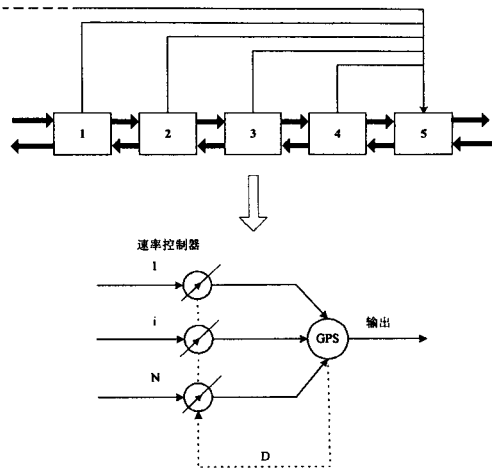

图 3-4 间接带宽控制（单结点情况）

图 3-4 给出了结点 3 处间接带宽控制示意图。其中 D 表示结点 3 的上游结点 i 接收到结点 3 公平消息的时延。环中各结点将根据虚时间的演进信息，动态的、自适应的调整速率控制器的控制接入速率。特别的，考虑在理想情况下，即 $D=0$ 时，当速率控制器连续的接收到来自 GPS 服务器虚时间计算的反馈信息时，简单起见，不妨令链路容量 C=1，每聚合输入流的权值相同均为 ϕ，且 $N\cdot\phi=1$（N 为系统中连接的个数），则根据公式（3-6）可知，由结点 i 发起的流 i 的接入速率控制器的值将由下式给出：

$$r_i(t)=\min(1,\phi\frac{dv(t)}{dt}) \tag{3-7}$$

这里有几点需要说明：第一，在GPS系统中每个积压流 i 的虚时间变化规律 $v_i(t)$ 与系统虚时间的变化规律相同。第二，当系统中无积压流时，此时取 $r_i(t)$ 为 1。

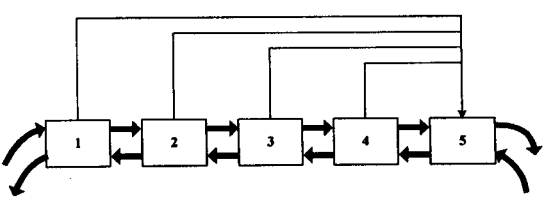

图 3-5 “平行”传送的情况

下面我们举例说明（3-7）式的含义。考虑图 3-5 所示 4 个流“平行”传送的情况。假设系统最初（$t=0^-$）处于空闲状态，从结点 4 观察，此时 $r_i(0^-)=1$（$i=1\sim4$）。在时刻 $t=0^+$时，所有的流立即以尽可能高的速率向环中注入数据。显然如果不加控制很快结点 4 处所有流（包括转发流和本地流）就会产生积压。由（3-5）、（3-7）式可知，$\phi\frac{dv(t)}{dt}$ 为1/4，该值立即反馈，所有结点的速率控制器（结点 1～4）立即设置成 $r_i=1/4$（$i=1\sim4$），随后所有的流将按照这样的接入速率注入业务。假设在之后的某个时刻 t_1，流 4（结点 4 的本地业务）停止向环中注入数据。当流 4 在 GPS 服务器中不再产生积压时，$\phi\frac{dv(t)}{dt}$ 变为1/3，结点速率控制器将设置 $r_i=1/3$（$i=1\sim3$）。根据上面的分析可知，通过监视 GPS 系统虚时间的变化，系统不仅能够公平的为每个结点分配带宽，而且还能够自适应的增加接入速率控制器的值，从而回收未用的带宽。这里有一点需要注意，对于本例的 4 个流而言，接入速率控制器的值一定不会低于1/4——系统最小的公平带宽（1/4），也就是说采用 GPS 服务规则分配带宽可以确保每个结点获得的公平带宽不会低于系统最小的公平带宽，从而避免环中某些结点“恶意”抢占带宽而引起环中某些结点因无法获得带宽而“饿死”。

二. 反馈的公平信息具有延时性和临时会聚性

DBRR 的另一个特点是反馈的公平信息具有延时性和临时会聚性。DBRR 要求环中每个结点收集与其它结点共享的公平信息，这将会引入时延和信息会聚。在前面的讨论中，我们假设虚时间连续不断无延迟的反馈给接入速率控制器。然而实际上，反馈信息必需在一定时间间隔内先进行汇总然后再以公平消息的形式向环中其它结点传送。

我们仍然采用图 3-5 所示的例子，令 D = 0，GPS 服务器每隔 T 秒（其值通常很小）传送一个消息，该消息携带前一个 T 秒期间虚时间演进的信息。如果 GPS 服务器在时间间隔 $[t-T,T]$ 内连续积压（即 GPS 服务器一直忙），则虚时间的演进信息只需简单的时间平均便可得到。如果 GPS 服务器在时间间隔 $[t-T,T]$ 内的某些时间段内空闲，这说明系统还有额外的带宽可供使用，可相应的增加速率控制器的值。为了跟踪在整个时间间隔 $[t-T,T]$ 内可供系统使用的额外带宽的量，从而充分的利用带宽。我们定义 c 为 $[t-T,T]$ 内 GPS 服务器处于忙的时间段与 T 的比值，即：

$$c=\frac{\text{系统处于忙时间段}}{T} \tag{3-8}$$

显然 $c\in[0,1]$，综合上面的结果，我们可以得出考虑延时性和临时会聚性的流 i 接入速率控制器的值：

$$r_i(t)=\min(1,\phi\bullet(\frac{v(t)-v(t-T)}{T})+(1-c)) \tag{3-9}$$

比较（3-7）式和（3-9）式，由于（3-7）式并没有考虑公平信息（虚时间的变化率）反馈的延时性和临时会聚性，并且是以虚时间连续不断的反馈给接入速率控制器为前提得出的结论，故实际应用中并不用（3-7）计算速率控制器的值。（3-9）式给出了考虑延时性和临时会聚性的情况下计算接入速率控制器的值的方法。可以看出确定速率控制器值的关键在于如何计算时间间隔$[t-T,T]$内系统虚时间的平均变化率。我们将在 3.4 节给出一个简单的计算$[t-T,T]$内系统虚时间平均变化率的算法。

最后我们考虑$D>0$的情况。D 表示结点 i 接收到下游结点公平消息的时延，通常为传播（propagation）时延。在这种情况下，结点 i 将在时刻 t 设置时间间隔$[t-T-D,t-D]$内的平均公平接入速率。

3.3.3 DBRR—多点情况

上一节我们讨论了 DBRR 单点情况，给出了计算公平接入速率的方法。这一节我们讨论 DBRR 多点情况。DBRR 多点情况与单点情况的最大区别在于，多点必需考虑带宽细分问题。所谓带宽细分是指系统为每个结点（会聚流）分配的公平带宽需要为该结点发往不同目的结点的微流（micro fluid）进一步分配带宽。多点情况如图 3-6 所示。

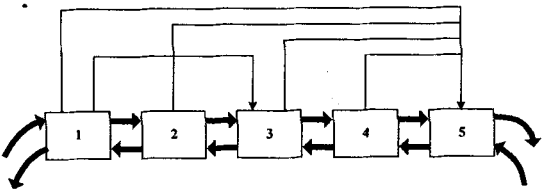

图 3-6 DBRR 多结点情况

图中结点 1 同时向结点 3 和结点 5 发送数据。观察结点 1，系统分配给该结点的公平带宽需要进一步细分，也就是说结点 1 要为 flow（1，5）和 flow（1，3）进一步分配带宽。为了获得多点情况下的网络结点间的公平性以及同一结点内流与流之间的公平性，输入结点必须首先计算出由该结点发起的业务流在整个传送通路上最小的链路公平速率。假设链路容量 $C=1$，每聚合输入流的权值相同，并且结点 1 的两个微流的权值也相同。图 3-6 中各结点均以尽可能高的速率往环中注入数据，由于结点 1 只有一个输入流（两个微流会聚成一个输入流），故最初结点 1 按照公平速率 1 为该输入流分配带宽。由于两个微流的权值相同，结点 1 将分别设置 flow(1,5)和 flow(1,3)接入速率控制器的值为 1/2、1/2。又因为 flow(1,5)途经结点 2~4，结点 1 将会分别接收到来自结点 2~4 反馈的公平信息 1/2、1/3 和 1/4，故结点 1 将设置 flow(1,5)的接入速率控制器的值为 1/4。同样 flow(1,3)途经结点 2，

结点 2 反馈给结点 1 的公平信息为 1/2，所以结点 1 将设置 flow(1,3)接入速率控制器的值为 1/2。以上对带宽细分的讨论可以通过下式求得：

$$r_{i,j}(t)=\min\left(1,\min_{i\le n<j}\left(r_i^n-\sum_{k\ge n,k\ne j}r_{i,k}^n\right)\right) \qquad (3\text{-}10)$$

在解释（3-10）式之前，我们先作如下约定：每段环路（Ring Segment）连接两个结点，我们称上游结点为该环路的拥有者，若将上游结点记为 i，其拥有的环路也记为 i。

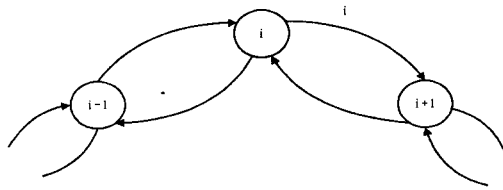

图 3-7 结点与环路的关系

图 3-7 给出结点与环路的关系，结点 i 是环路 i（外环）的拥有者。根据约定，（3-10）式中：$r_{i,j}(t)$ 表示 t 时刻 flow(i,j)的接入速率控制器的值；r_i^n 表示结点 i 发起的业务流在环路 n 上分配的公平速率，r_i^n 可由（3-9）式求得；$r_{i,k}^n$ 表示结点 i 发起的业务流中与流 flow(i,j)共享环路 n 的流 flow(i,k)的接入速率控制器的值。

最后，讨论在多点情况下，DBRR 怎样自适应的实现空闲带宽的再利用。前面对 RPR 多点的讨论中，我们一直没有提及各结点如何再利用空闲带宽。实际上空闲带宽的再利用是自动完成的。

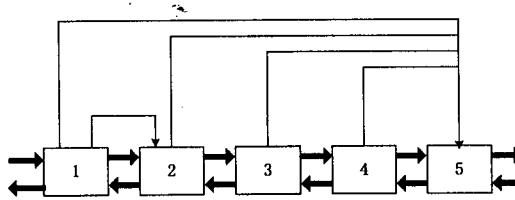

图 3-8 空闲带宽再分配的情况

如图 3-8 所示，设链路容量 $C=1$，接收公平消息的迟延 $D=0$。每聚合输入流的权值相同，且结点 1 的两个微流的权值也相同。所有的流在 $t=0$ 时刻以尽可能高的速率往环中注入数据。环中每个结点每隔 T 秒计算一次公平消息然后向上游结点反馈。最初结点 1 设置本地输入流的公平速率为 1，由于结点 1 的两个微流的权值相同，则结点 1 将分别设置 flow(1,5)和 flow(1,2)接入速率控制器的值为 1/2、

1/2。在第一个 T 秒期间，结点 1 分别接收到来自结点 2、3、4 的公平消息 1/2，1/3 和 1/4，同时计算本地公平速率为 1。因为 flow(1,5)途经结点 2~4，故结点 1 将设置 flow(1,5)接入速率控制器的值为 1/4，流 flow(1,2)仍为 1/2。显然此时结点 1 与结点 2 之间的环路带宽并没有充分利用（1/2+1/4=3/4）。在第二个 T 秒期间，结点 1 分别接收到来自结点 2、3、4 的公平消息 1/4，1/4 和 1/4，并且由（3-9）式计算出本地输入流的公平速率为 1。结合来自结点 2、3、4 的公平消息，由（3-10）式求得流 flow(1,5)的公平接入速率为 1/4，同样由（3-10）式可求得流 flow(1,2)的公平接入速率为 3/4。因而结点 1 将设置流 flow(1,2)的接入速率控制器的值为 3/4，流 flow(1,5)保持不变。通过以上分析，可以看出 DBRR 能够自适应的实现空闲带宽的再利用。

3.3.4 DBRR 公平性分析

在实际的系统中，许多因素导致 DBRR 计算出的公平速率与理想情况下的公平速率之间存在偏差（不公平）。公平消息临时会聚的时间间隔 T 是影响 DBRR 算法公平性的一个最大的因素。本节我们采用一个简单的理论模型研究 T 对系统公平性的影响，该分析方法很容易推广到对诸如传播时延对系统公平性影响的分析上。简化起见，不考虑系统的传播时延。

如图 3-4 所示，假设传播时延为 0，即 D = 0，所有的流均以尽可能高的速率往系统中注入分组，我们考察两个流 i,j 在最坏情况下获得的最大服务偏差，从而确定服务偏差的上界。图 3-4 也可表示为图 3-9 的形式。

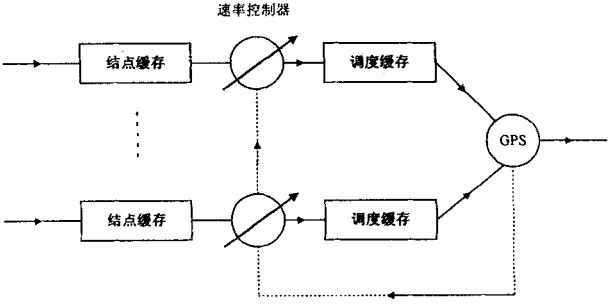

图 3-9 DBRR 单结点模型

如果一个流在其发送结点速率控制器处的缓存非空，我们称该流在结点处被积压，记为 N_backlogged；如果其调度缓存（包括转移队列和本地调度队列）非空，则称该流在调度器处被积压，记为 S_backlogged。进而当 GPS 服务器为流 i 服务

的速率大于结点速率控制器的控制速率时，称该流处于过调（over-throttled）状态。类似的，当 GPS 服务器为流 i 服务的速率小于结点速率控制器的控制速率时，称该流处于欠调（under-throttled）状态。这里需要特别说明的是：由于所有的输入流均尽可能多的往系统中注入数据，因而系统中所有的流均为 N_backlogged 流，所有业务以接入速率控制器规定的速率进入调度缓存。当流 i 处于欠调状态时，调度缓存的队长会不断的增加；反之，当流 i 处于过调状态时，调度缓存会逐渐取空。我们正是通过确定流 i 在时间间隔 T 内最坏情况下获得的过调或欠调的量来确定流 i 与流 j 之间最大的公平速率偏差。

首先推导在时间间隔 T 内流 i 欠调业务量和过调业务量。简化分析，考虑固定长度的分组，即系统为每个分组服务的时间是固定的。设 $v(k)$ 为在时刻 kT 时的虚时间，$c(k)$ 为时间间隔 $[kT,(k+1)T]$ 内系统处于忙状态的时间段与 T 的比值，N 为系统中流的个数。流 i 的欠调业务量可由下面的引理得到。

引理 1 在 DBRR 系统中，一个 N_backlogged 流 i 至多欠调 $(1-\phi_i)CT$，ϕ_i 为流 i 的权值。

证明：对一个 N_backlogged 流 i 而言，当公平速率（GPS 服务速率）减少时，由于结点速率控制器暂时以未调节的速率对该流进行速率控制，意味着此时结点速率控制器的速率比公平速率高。在这种情况下，虚时间 $v(t)$ 在时间间隔 $[kT,(k+1)T]$ 内的平均斜率将减小。对于 N 个流的系统，欠调状态下最坏的情况是在 N 个连续的 T 秒时间间隔内，虚时间的斜率不断减小。因而假设流 i 在时刻 0 时接入系统，U_i 表示截止到 NT 时刻流 i 总的欠调业务量，由式（3-9）有：

$$U_i = C\cdot\left(\sum_{k=1}^{N-1}(\phi_i\cdot(v(k)-v(k-1))-\phi_i\cdot(v(k+1)-v(k)))\right)$$

$$= C\cdot(\phi_i\cdot(v(1)-v(0))-\phi_i\cdot(v(N)-v(N-1)))$$

因为 $\phi_i\cdot T \le \phi_i\cdot(v(k)-v(k-1)) \le T$，故：

$$U_i \le (1-\phi_i)CT$$ 证毕

类似的，下面的引理给出了流 i 过调的业务量。

引理 2：在采用 DBRR 的系统中，一个 N_backlogged 流 i 至多过调 $(1-\phi_i)CT$。

证明：对一个 N_backlogged 流 i 而言，当公平速率（GPS 服务速率）增大时，由于结点速率控制器暂时以未调节的速率对该流进行速率控制，意味着此时结点速率控制器的速率比公平速率低。在这种情况下，虚时间 $v(t)$ 在时间间隔 $[kT,(k+1)T]$ 内的平均斜率将增大。对于 N 个流的系统，过调状态下最坏的情况是在 N 个连续

的 T 秒时间间隔内，虚时间的斜率不断增大。当系统一直处于过调的状态时，GPS 服务器很有可能发生空闲。由式（3-9），在时隙$(k+1)T$ 期间，流 i 获得的总的过调业务量为：$C\bullet(\phi_i\bullet(v(k+1)-v(k))+(1-c(k))\bullet T)$。因而假设流 i 在时刻 0 时接入系统，O_i 表示截至到 NT 时刻流 i 过调的总业务量，有：

$$O_i \le C\bullet\left(\sum_{k=1}^{N-1}((\phi_i\bullet(v(k+1)-v(k))+(1-c(k))\bullet T)-(\phi_i\bullet(v(k)-v(k-1))+(1-c(k-1))\bullet T))\right)$$

$$= C\bullet((\phi_i\bullet(v(N)-v(N-1))+(1-c(N-1))\bullet T)-(\phi_i\bullet(v(1)-v(0))+(1-c(0))\bullet T))$$

因为 $\phi_i\bullet T \le \phi_i\bullet(v(k)-v(k-1))+(1-c(k-1))\bullet T \le T$，故：

$$O_i \le (1-\phi_i)CT$$ 证毕

由引理 1 和引理 2，我们能够很容易推导出采用 DBRR 算法的系统中，当任意两个流 i,j 都尽可能多的往环中注入数据时，最坏情况下流 i,j 之间的最大服务偏差。

定理 1 在采用 DBRR 的系统中，当任意两个流 i,j 都尽可能多的往环中注入数据时，在任意的时间间隔 T 内，其服务偏差不超过 $(1-\phi_i)CT+(1-\phi_j)CT$。

证明：考虑调度器积压的情况，显然 S_backlogged 流将获得不低于系统分配给每个流的公平速率。在欠调情况下，S_backlogged 流的发送速率都限制在公平速率上，两个流接受的服务量相同。因而不公平的现象只会在过调的情况下发生，进一步讲，不公平现象只会在系统过调时新接入的流与已存在的流之间发生。在这种情况下，新接入的流可能获得的额外服务量与它的欠调业务量相同，另一方面，已存在流损失的服务量与它的过调业务量相同。由引理 1 和引理 2 立即得到系统中任意两个流 i,j 之间的最大服务偏差为 $(1-\phi_i)CT+(1-\phi_j)CT$。 证毕

从以上分析可以看出，DBRR 系统中任意两个流之间的服务偏差与公平消息临时会聚的时间间隔 T 密切相关。由定理 1，可以界定系统中任意两个流之间的服务偏差不超过 $(1-\phi_i)CT+(1-\phi_j)CT$。考虑 $T=0$ 这样一种特殊情况，此时系统中任意两个流之间的服务偏差为零。说明当系统不存在公平消息临时会聚时，DBRR 能够达到最佳的公平性。

界定任意两个流之间的服务偏差对设计调度队列（转移队列和本地调度队列）的大小有很大帮助，例如链路速率为 2.5Gbps，流的个数为 64，公平信息会聚时间为 0.5ms 的 RPR 网络，假设由定理 1 可知其结点与结点（流与流）之间的最大服务偏差为 308KB，故调度队列的大小可以设置为 308KB。在实际的应用中，由于要考虑传播时延和业务的突发性，因而实际调度队列的长度将略大于 308KB。

3.3.5 DBRR 算法在实际仿真情况下的具体实现

前几节我们讨论了 DBRR 算法的特点以及在理想情况下具体的实现方案。可以看出为实现 DBRR 算法，关键在于计算虚时间 $v(t)$。下面描述一个用于实际仿真情况的算法。该算法采用每输入（per-ingress）字节计数器来近似的计算时间间隔 T 秒内系统虚时间的演进。

假设所有聚合流的权值相同。记在结点 j 处对第 i 个输入流的流量进行测量的字节计数器为 m_i^j，这里需要注意的是输入流 i 不仅包括转发数据流而且还包括本地的数据流。通过观察在时间间隔 T 内结点 j 处每输入计数器 m_i^j 的变化，可以近似的求得系统虚时间的变化情况，从而确定每输入流的公平接入速率。我们用 ***Fa*** 跟踪虚时间的变化情况，即 $\phi=(\frac{v(t)-v(t-T)}{T})+(1-c)$。***Fa*** 可以通过下面的步骤求得：假设在时间间隔 T 秒内，在结点 j 处测量的数据流有 l 个，则首先对每输入字节计数器的测量值进行排序，即：$m_1^j \le m_2^j \le \cdots \le m_l^j$。下一步对测量值进行归一化处理，即：$n\_m_i^j = \frac{m_i^j}{CT}$。这里需要注意的是，$\sum_{i=1}^{l}\frac{m_i^j}{CT}$ 很有可能大于 1。最后将在结点 j 处将理想的公平速率 $\frac{1}{l}$ 与测量值 $\frac{m_i^j}{CT}$ 进行比较，从而确定 ***Fa*** 的值。计算时间间隔 T 秒内 ***Fa*** 大小的算法由下面的伪代码给出：

伪代码：

```
1   Sort_from_min_to _max (m_i^j ,l) // 每输入流字节计数器的值从小到大进行排列
2   i = 1;
3   for (i = 1,i < = l ,i++)
4   n_m_i^j = m_i^j / CT;                    // 求归一化到达速率
5   if (c < 1)
6       Fa = n_m_i^j + (1- c)                // 回收未用带宽
7   else
8   {
9       i = 1;
10      Fa = 1/l ;                           // 结点j处理想的公平速率。
11      Cnt = l;
12      Reclaim_cap = 1;
13      while (( n_m_i^j < Fa ) && (n_m_i^j >= Fa) && ( i < l ))
        {
14          Cnt --;
15          Reclaim_cap -= n_m_i^j ;
16          Fa = Reclaim_cap/ Cnt ; // 重新分配带宽
17          i = i+1;
```

```
18      }          // 求得Fa
19  }
```

3.3.2 节中定义 c 为 $[t-T,T]$ 内 GPS 服务器处于忙时间段与时间间隔 T 的比值，并给出相应的表达式（3-8），但是在实际应用中，（3-8）式并不适合计算 c，这是因为确定每个流的空闲时间段和忙时间段将是很繁琐很困难的事情。这里我们从 c 的定义出发，采用另一种方式近似的计算 c 的值。考虑到 c 表示系统忙时间段与时间间隔 T 的比值，实际上可以近似的认为结点 i 在时间间隔 T 内服务的总业务量与系统最大可提供服务的业务量 CT 之比，即：

$$c=\frac{\text{结点i在时间间隔}T\text{内服务的总业务量}}{CT} \qquad (3\text{-}11)$$

（3-11）式中，CT 可以立即得到，结点 i 在时间间隔 T 内服务的总业务量可以通过计算 T 期间内结点 i 发送的总比特数得到。

前面我们曾经讲过 RPR 主要是为城域网（MAN）优化而设计的技术。城域网大容量、高速率的特点要求网内每个结点能够高速的处理本地业务和转发业务，因此在设计 RPR MAC 时应尽量简化。这里所谓的简化不仅体现在硬件上—尽可能在不影响网络性能的前提下降低硬件复杂度，同时也体现在软件上—算法的时间复杂度。我们知道一个算法的时间复杂度直接影响到网络设备的处理速度，象城域网这样的高速网络（通常高达 G 比特，以后会更高）要求算法的时间复杂度尽量的低。下面我们将讨论用上述方法实现 DBRR 算法的时间复杂度。对每输入字节计数器 $m_i^j\ (i=1,2\cdots l)$ 进行排队的时间复杂度为 $o(l\log l)$。考虑最坏的情况下算法的计算复杂度，例如由 128 结点构成的环采用最短路径路由，在结点 j 处输入流最多为 64 个，$l\log l$ 为 384，算法中的 while 循环最多进行 64 次迭代。可见 DBRR 算法的计算量很小，能够适用于象城域网这样的高速网络。

最后讨论 DBRR 公平消息帧是如何构造和传送的。DBRR 公平消息帧将按照 2.2 节介绍的公平消息帧结构进行构造。DBRR 公平消息帧的长度为 16 字节，其中公平消息域 1 用来携带本地结点计算的公平带宽信息。公平消息将在下面三种情况之一发生时进行发送：

1. 前面讨论环中各结点每隔 T 秒计算一次公平带宽信息，因而公平消息帧每隔 T 秒发送一次。
2. 当环中某结点发生拥塞时，该结点会立即向其上游结点发送公平消息。
3. 当环中某结点接收到来自其下游结点的公平消息时，该结点将本地计算的公平消息的值与接收到的公平消息的值进行比较，选择较小的值向其上游结点转发。

§3.4 本章小结

本章我们对 RPR 带宽管理机制进行了详细的探讨，首先提出了一个新的 RPR MAC 参考模型——IA 参考模型，并对其中各个功能实体进行定义和设计。IA 参考模型与其它建议提出的参考模型的最大区别在于它首先对输入业务进行聚合，然后根据带宽分配算法为每个聚合流分配带宽。IA 参考模型主要是针对城域网的特点提出的。随后基于 IA 参考模型，进一步提出了一个全新的 RPR 带宽分配算法—DBRR，并从理论上对 DBRR 算法进行了分析和设计，重点分析了 DBRR 间接带宽控制和公平信息具有延迟性和临时会聚性的特点，以及临时会聚的时间间隔 T 对公平性的影响，给出了流与流之间服务偏差的上界。最后根据前面的讨论结果，将 DBRR 实例化，设计了一个用于实际仿真的算法。下一章我们将通过具体的仿真对 DBRR 的性能做进一步考察。

第四章 RPR 网络仿真和性能分析

§4.1 前言

通信系统仿真是借助于计算机对通信系统的模型进行实验，它具有经济、安全、实验周期短的特点。此外，在通信系统的方案设计中使用计算机仿真技术，可以及时的将通信、计算机等行业的最新发展成果运用到自己的系统中去。计算机仿真技术的固有特点使之成为通信系统分析、设计、优化的强有力的工具。

上一章，我们提出了一个新的 RPR MAC 参考模型——IA 参考模型，基于该模型设计了一个全新的公平带宽分配算法（DBRR），并从理论上对 DBRR 算法进行了讨论与分析。本章我们通过计算机仿真再对 DBRR 算法进行讨论，重点考察 DBRR 在网络吞吐率、结点公平性、算法收敛速度（网络达到动态平衡的快慢）以及平均分组传输时延几方面的性能。

本文采用的仿真工具为 OPNET8.0.C。软件运行环境为 Inter PIII 866MHz，操作系统为 Windows2000 操作系统。

本章首先对仿真工具 OPNET8.0.C 做一简单的介绍，然后根据 OPNET 的要求设计基于 IA 参考模型的 RPR MAC 仿真模型。最后给出 DBRR 在网络吞吐率、结点公平性以及算法收敛速度（网络达到动态平衡的快慢）等几方面的性能仿真结果，并与 Gandalf 算法进行比较。

§4.2 OPNET 软件介绍

4.2.1 功能介绍

OPNET Modeler 是一个功能强大的仿真软件包，它不仅支持通用的网络建模，还为特定类型的网络仿真提供支持。OPNET Modeler 可以对许多通信系统进行建模和仿真。下面是它的一些主要应用：

- 基于标准的 LAN 和 WAN 性能模型；
- 互联网络的规划；
- 通信体系和协议的研究开发；
- 分布式传感控制网络；
- 移动分组无线网络；
- 卫星网络。

OPNET Modeler 之所以有以上种种应用，主要取决于其特有的功能：

- 面向对象 OPNET 定义的系统由各个对象组成，而每一个对象拥有可配置的属性族；
- 专用于通信网络和信息系统；
- 体系结构模型 ：对应于实际通信网的结构；
- 自动的运行模型仿真：所有的模型规范都自动用 C 语言编译成可执行的、高效的、离散事件的模型仿真；
- 与应用相关的统计数据：不仅支持系统内部自动生成的统计数据，同时允许用户自定义统计数据；
- 仿真完成后集成的分析工具：可以以图形、动画的形式对输出的统计数据进行分析。

面向对象和离散事件驱动的仿真模型是 OPNET Modeler 的两个最主要的特征；OPNET Modeler 根据不同的功能和层次定义了很多对象，它们构成了系统模型的主要框架。如果说面向对象的思想为 OPNET Modeler 提供了静态的配置，那么离散事件驱动则可以看成为 OPNET Modeler 动态运行仿真提供了依据。

4.2.2 模型规范编辑器

模型规范即研究系统的代表模型，OPNET 允许模型再用，很多网络模型都是基于事先开发出来的低层模型之上而建立的。

OPNET 通过一些获取系统行为特征的编辑器来支持模型规范，由于 OPNET 支持与实际网络系统相似的层次体系结构，故这些编辑器也被划分了层次，具体分为：

- 工程编辑器 开发网络模型，网络模型由子网和结点模型组成；
- 结点编辑器 开发结点模型，结点模型由含有进程模型的模块组成；
- 进程编辑器 开发进程模型，进程模型控制模块行为；
- 链路模型编辑器 开发链路模型
- 分组格式编辑器 分组格式规定了分组内存储信息的结构和顺序；
- ICI 编辑器 创建、编辑、查看接口控制信息（ICI）格式，ICI 用于在进程之间传递控制信息；
- PDF 编辑器 创建、编辑、查看概率密度函数（PDF），PDF 用于控制某些事件，例如源模块中分组产生的频率。
- PATH 编辑器：创建路径的属性
- 天线图编辑器
- 调制曲线编辑器
- 文本编辑器

4.2.3 建模过程

OPNET 的建模过程是分层次进行的，即分别在进程域、结点域和网络域三个空间自下而上（或者自上而下）对其包含的对象进行定义和规范。

1. 进程域(Process Domain)

在进程域内，进程模型用来规定结点域内处理器(processor)和队列(queue)模块的行为；通过它实现子系统，包括通信协议、算法、共享资源（例如磁盘或存储器）、操作系统、排队规则、特定的业务发生器以及统计数据收集等等。

2. 结点域（Node Domain）

OPNET 的结点模型是建立在一种所谓“积木”构架的基础上的。这是用来描述硬件系统的一种很有效的技术，同时它也能很好的描述已定义接口的高层软件实体之间的关系。实际上，国际标准组织（ISO）规定的开放系统互联（OSI）体系网络就是基于这种技术。每个“积木”对应于不同的协议层。与此相似，OPNET 的结点模型也包含了许多处理分组的“积木”，并把它们称为模块（module），每个模块又包含了一组输入和输出，状态存储器，以及根据输入和状态信息如何计算输出的算法。

3. 网络域（Network Domain）

网络模型定义了要仿真的系统，它对系统所含对象进行了高级描述，其中包括它们的位置、互联和配置情况等等。这里，网络的大小和规模可以简单，也可以很复杂，一个网络可能只容纳一个结点，一个子网，或者是许多互联的结点和子网。例如，一个星型拓扑结构的网络包含有一个中心 hub 结点和一些与 hub 通过点到点链路相连的外围结点。

§4.3 RPR 网络仿真模型的设计

4.3.1 RPR 网络的网络（Network）模型设计

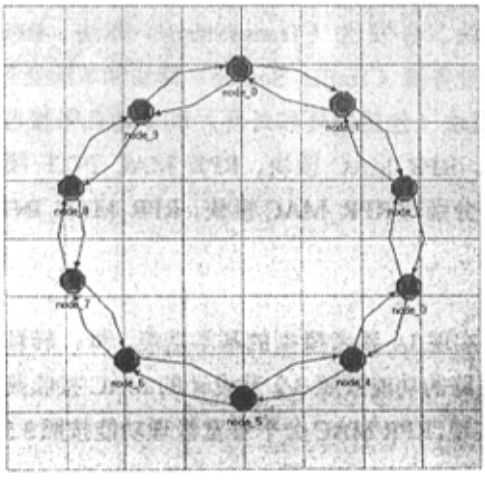

图 4-1 RPR 网络模型

我们在设计 RPR 网络仿真模型时，也遵循 OPNET 建模过程的层次性要求，即分别在进程域、结点域和网络域三个空间自下向上（或自上而下）对其包含的对象进行定义和规范。这里我们采用自上而下的方式设计 RPR 网络的仿真模型。首先进行网络模型的设计。我们设计的 RPR 网络由 10 个 RPR 结点构成，每个 RPR 结点的功能相同，即同时具备传送本地业务和转发上游结点业务的能力，相邻两结点间的距离为 30km，这样整个网络的规模为 300km，基本上达到城域网的规模。结点间采用光纤连接，链路速率为 2.5Gbps（OC-48）。图 4-1 所示为 RPR 网络模型。

4.3.2 RPR 网络的结点（Node）模型设计

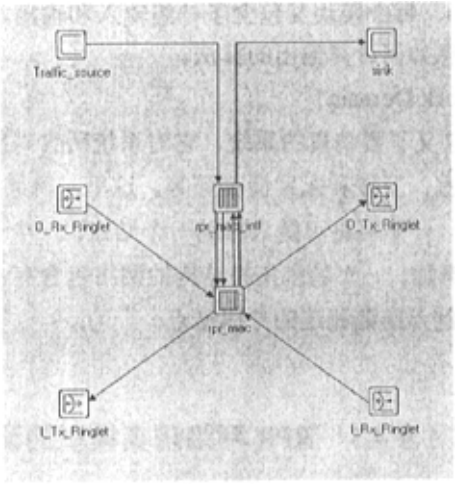

图 4-2 RPR 结点模型

如图 4-2 所示 RPR 结点模型。RPR 结点模型按功能可划分为 6 个模块，即：环接收（Receiving）模块、环发送（Transmitting）模块、RPR_MAC 模块、RPR_MAC_INTF 模块、业务源（Traffic Source）模块和本地业务接收（sink）模块。其中环接收和发送模块（包括内环和外环）和本地业务接收（sink）模块为 OPNET 提供的通用模块。RPR_MAC 模块、RPR_MAC_INTF 模块和业务源模块为自己设计的模块。下面分别对 RPR_MAC 模块、RPR_MAC_INTF 模块和业务源模块做简单描述：

1. RPR_MAC 模块

RPR_MAC 模块主要实现 IA 参考模型的基本功能，即：转移通路和公平带宽管理。RPR MAC 转移通路的功能按照 3.2 节设计的 MAC 接收规则、转移规则、传输规则以及丢弃规则实现。RPR MAC 公平带宽管理功能按照 3.3 节设计的 DBRR 公平带宽分配算法实现。

2. RPR_MAC_INTF 模块

RPR_MAC_INTF 模块主要实现 RPR MAC 与 RPR MAC 用户之间简单的接口功能。此外 RPR_MAC_INTF 模块还实现一些简单的 MAC 用户功能，例如 VOQ。

3. 业务源（Traffic Source）模块

为了简化设计，RPR 结点模型的业务源将根据具体的仿真要求进行设定。在考察 DBRR 算法的公平性和吞吐率时，要求源尽可能多的向环中注入数据，仿真采用预设的恒定比特率业务源，而在考察分组平均传输时延时，采用 Possion 源。

4.3.3 RPR 网络的进程（Process）模型设计

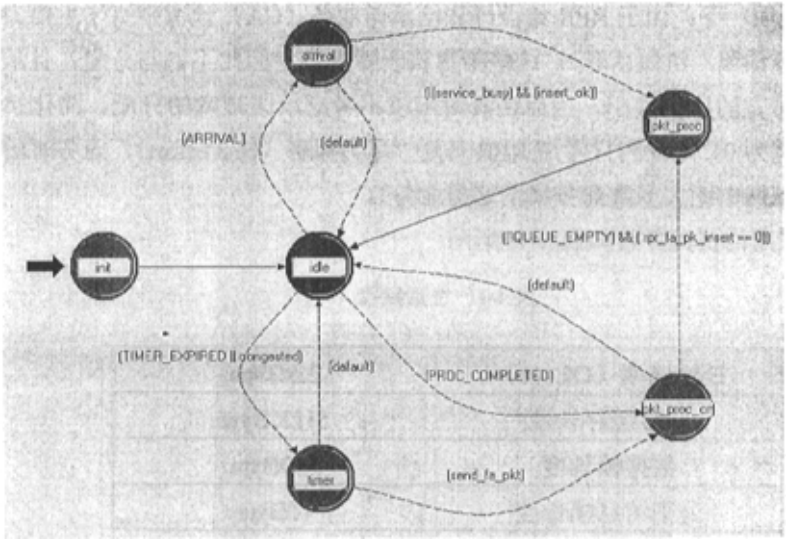

图 4-3 RPR_MAC 模块进程模型

如图 4-3 所示 RPR MAC 模块进程模型。该模型由 6 个进程构成。下面简要介绍每个进程的功能：

1. 初始化（init）：该进程对变量、进程间转移状态、定时器和模块属性等进行初始化。
2. 到达（arrival）：该进程接收到达的帧，包括本地接收帧、转发帧、本地传送帧和控制帧，并进行分类处理（缓存），除了完成 IA 参考模型中帧头处理实体的功能以外，还负责接收来自 RPR MAC 用户的消息，以便 RPR MAC 了解用户的状态。
3. 定时器（timer）：该进程主要是完成公平消息计算和公平带宽分配的功能，即 IA 参考模型中公平带宽分配实体的功能。
4. 分组处理（pkt_proc）：该进程主要完成本地业务和转移业务的调度功能和跟踪

测量所有上游结点的业务通过该结点输出链路的速率以及该结点本地业务注入到输出链路的速率的监视功能。即实现 IA 参考模型中业务监视实体和帧调度实体的功能，另外该进程还进行环路的拥塞检测。

5. 分组处理结束（pkt_proc_cmpl）：该进程主要完成分组的发送。
6. 空闲（idle）：该进程主要用于处理进程间的调度，也就是说，当一个进程执行完毕后，如果满足进入空闲状态的条件，则进程将转移到空闲状态，等待下一次被“唤醒”（通常情况下进程是靠中断“唤醒”的）。

IA 参考模型中还有一个非常重要的功能实体——媒质接入速率控制实体在 RPR_MAC_INTF 模块中实现，另外数据的采集在 sink 模块中完成。

需要说明一下：由于 RPR 结点预先给承诺业务（CA）预留带宽（主要为时延敏感性业务预留，例如话音），这些带宽属于静态带宽因此不参与分配，且承诺带宽占整个带宽的比例很小。因而仿真时可以不考虑承诺带宽的分配。简化起见，令承诺带宽为 0，即我们只考虑 RPR 传送“尽力服务（best effort）”业务时的网络性能（城域网中绝大多数业务属于这类业务）。

本文采用的仿真参数如表 4-1 所示：

表 4-1 仿真参数

链路速率（OC-48）	2.5Gbps
转移缓存长度	512KByte
数据帧长度	500Byte
公平消息帧长度	16Byte
链路传播时延	0.1msec
网络总结点数	10
公平信息会聚收集时间（T）	1mec

§4.4 RPR 网络性能仿真分析

4.4.1 性能指标

衡量一个网络性能的好坏，可以借助于很多性能指标，不同网络侧重点不同。本文根据 RPR 的特点和要求参照以下的性能指标考察采用 DBRR 算法的 RPR 网络性能。

1. 公平性

公平性是指服务器的带宽公平地被系统中复用的各个连接使用。一个“恶意”

的连接不能够依靠多向网络发送分组而“贪婪”的占用系统带宽，影响其它连接的服务。系统带宽资源应该公平地根据各连接的要求按需分配使用，对带宽需求高的连接相应获得的服务也多。

2. 吞吐量

定义单位时间内在网络中成功传送的业务量为吞吐量。假定帧长固定，其长度为 L 比特，且单位时间（秒）内成功传送的帧数为 n，则吞吐量可表示为 $n \cdot L$ bps。通常，用信道传输速率 C 对吞吐量归一化，归一化的吞吐量表示为 S，即：

$$S = \frac{n \cdot L}{C}$$

影响吞吐量的因素有很多，对于弹性分组环而言，有两种情况对吞吐量的影响最大，一种情况是 RPR 结点存在队头阻塞，另一种情况是网络存在业务震荡。下面几节我们将详细讨论这两种情况对吞吐量的影响。

3. 平均传输时延

平均传输时延由 4 部分构成，它包括：处理时延、排队时延、发送时延和传播时延。其中，处理时延指链路层接收到该新产生的业务到分组被送入传送队列之间的时间；排队时延指分组到达传送队列到分组被发送之间的时间；发送时延指分组的第一个比特进入信道到最后一个比特离开信道之间的时间；传播时延指分组的最后一个比特从链路的头结点到达链路的尾结点之间的时间。分组时延主要由排队时延和发送时延构成。本文主要考察分组平均传输时延随网络业务负载变化的规律。为了使仿真结果更清楚明了，仿真中分组平均传输时延只包含分组的排队时延和传输时延。

4. 算法的收敛时间

判断一个算法的好坏，算法的收敛时间也是一个重要的参数。我们这里所说的算法收敛时间是指当网络从一个状态转移到另一个状态，并且达到稳态时需要多长时间。针对公平带宽分配算法，算法的收敛时间特指当环中结点进行带宽重新分配到带宽分配完毕后，网络重新达到稳态的时间。可见算法收敛时间越短越好，这样不但有助于提高网络的吞吐量，而且还有利于减少业务的震荡。

4.4.2 RPR 网络的公平性

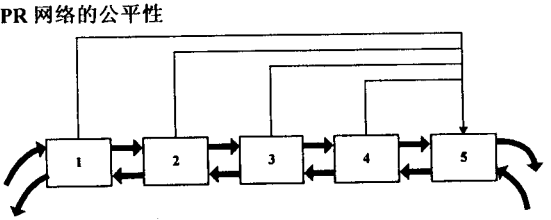

图 4-4 “平行”传送的情况

公平性是 RPR 设计目标之一。前一章在对 DBRR 的讨论和设计中，我们指出 DBRR 具有很好的公平性。下面通过仿真对 DBRR 的公平性进行验证。如图 4-4 所示“平行”传送的情况，假设结点 5 为接收结点，结点 1 到 4 尽可能多的向 5 发送数据，结点 1、2、3、4 分别在 0 秒、1 秒、2 秒和 3 秒时向环注入数据。在第 4 秒时，结点 0 停止发送数据。我们观察在时间段[0,5]秒内，各个流注入速率的变化情况。

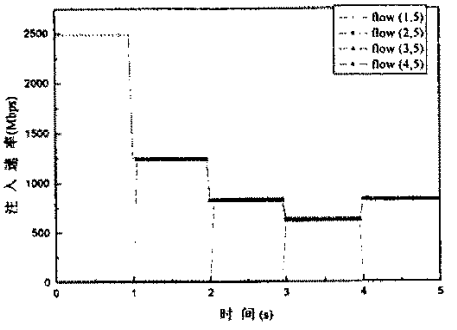

图 4-5 采用 DBRR 算法的公平性

图 4-5 给出了上述情况下的仿真结果，可以看出在时间间隔[0,1]秒内，由于网络中只有 flow(1,5)，故整个环路的带宽全部分给 flow(1,5)使用，flow(1,5)的注入速率大约为 2.5Gbps。在 t=1s 时刻 flow(2,5)接入到网络中，DBRR 立即给 flow(2,5)分配带宽，可以看到 flow(1,5)与 flow(2,5)均以大约 1.2Gbps 速率向环中注入分组。在时刻 t=2s 时，flow(3,5)接入到网络，DBRR 重新为 flow(1,5)、flow(2,5)和 flow(3,5)分配带宽，可以看到这三个流均以大约 800Mbps 速率向环中注入分组。在时刻 t=3s 时，flow(4,5)接入到网络，DBRR 再次重新对带宽进行分配，可以看到在 3 到 4 秒的时间段内，环中每个流均以大约 600Mbps 的速率向环中注入分组。特别的在时刻 t=4s 时，flow(1,5)停止向环中注入速率，可以看到 DBRR 算法立即回收 flow(1,5)释放的带宽，并将回收的带宽重新进行分配，可以看到当带宽分配结束后，环中各个流的注入速率又成为大约 800Mbps。仿真结果表明，我们设计的公平带宽分配算法—DBRR 能够实现 RPR 要求的结点间的公平性。

4.4.3 队头阻塞（HOL）对吞吐量的影响

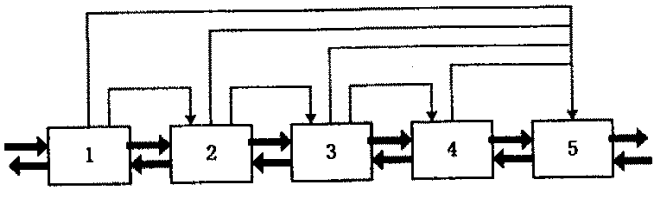

图 4-6 “热线”的情况

本节根据 4.2 节给出的 RPR 网络仿真模型，考察在“热线”情况下，队头阻塞对吞吐量的影响。我们曾经在第三章讨论过队头阻塞形成的原因。通常只支持单缓存访问的 MAC 通常会发生队头阻塞。首先考虑图 4-6 所示“热线”[5]情况下，外环环路带宽分配情况。

由图 4-6，假设结点 5 是接收结点，结点 1 到 4 尽可能多的向 5 发送数据，同时结点 1 到结点 4 也尽可能多的向其下游相邻结点发送数据。

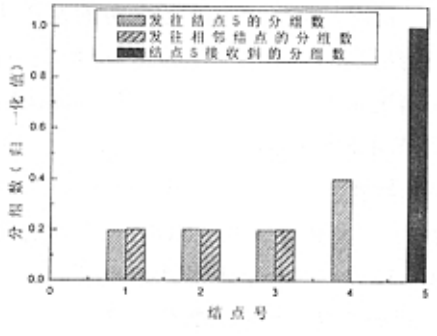

图 4-7 队头阻塞情况

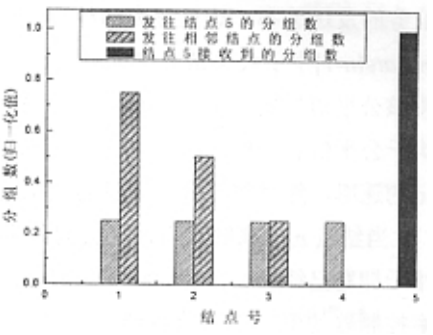

图 4-8 无队头阻塞情况

图 4-7 所示的是只支持单缓存访问的 MAC 得到的仿真结果，可以看到在这种情况下队头阻塞现象非常严重，空间重用的效果大大降低。由仿真结果，当网络达到稳态时，上游结点 4 向 5 发送数据的归一化速率大约为 0.40，该速率通过公平分组传送给结点 4 的上游结点 3，允许其以相同的速率发送分组。因为结点 3 存在队头阻塞，故结点 3 大约只能以公平分组中标记速率的 50%，即 0.20 向结点 6 发送数据，另 50%向结点 5 发送数据。结点 2、1 以类似的方式向结点 5 和相邻的下游结点发送数据。从仿真结果可以看出，除了结点 4 与 5 之间的链路带宽得到充分利用以外，其它结点之间的链路带宽并没有得到充分利用。如结点 0 与 1 之间链路的带宽利用率不到 40%。

图 4-8 是采用 VOQ 得到的仿真结果。我们可以看到结点 1 到 4 几乎以相同的速率向 5 发送数据，同时相邻两站之间的链路带宽也得到充分的利用。充分体现出空间重用的效果。

4.4.4 业务震荡对吞吐量的影响

我们在研究公平带宽管理机制的时候，发现当前提交给 IEEE 802.17 工作组的公平带宽分配算法草案[3，4，5]存在一些局限，其中对网络性能影响最大、最严重的局限是在非平衡业务情况下，如果采用[3，4，5]的带宽分配算法，会产生严

[5] 之所以选择“热线”的情况，因为在这种情况下，队头阻塞现象最严重

重的业务震荡。

以 Gandalf 算法为例，Gandalf 算法采用两个测量参数，*forward_rate* 和 *my_rate*。前者测量所有转移（transit）业务的服务速率；后者测量所有本地业务的服务速率。假设某时刻在结点 *n* 处，*forward_rate*[*n*] + *my_rate*[*n*] > *low_threshold*（预设的拥塞指示门限），表示此时结点 *n* 拥塞。一旦结点 *n* 检测到拥塞，其立刻向上游结点发送公平信息，所发送的公平信息中存放归一化的本地业务服务速率—*myrate*[*n*]。当上游结点 *n*-1 接收到该公平信息后，它就根据接收到的公平信息的内容调整本地业务的发送速率。这里分为两种情况：一种情况是结点 *n*-1 本地业务的发送速率 *myrate*[*n*-1]小于公平信息中标记的速率，则结点 *n*-1 不改变本地的发送速率，直接将该公平消息向其上游结点转发；另一种情况是本地业务的发送速率 *myrate*[*n*-1]大于公平信息中标记的速率，则结点 *n*-1 调节本地业务的发送速率为公平信息中标记的速率，然后将该公平消息向其上游结点转发。

当结点 *n* 拥塞解除后，它会向其上游结点发送内容为空（NULL）公平分组，指示拥塞已经解除，上游结点（例如结点 *n*-1）将以一定的比例不断的增加接入速率控制器的值，尝试让本结点发的更多。只要该结点没有接收到下游结点的拥塞指示消息（公平消息中的值不为空），接入速率控制器的值就不断增加，直到结点 n 拥塞为止。当结点 n 发生拥塞时，其又开始在公平分组中标记发送速率发向它的上游结点，超过标记速率的上游站将会调节它的接入速率控制器的值。这一过程在每个站中同时进行，周而复始。

Gandalf 算法采用这种机制非常简单，当环中结点注入的业务量相差不大时，可以获得较高的吞吐量和结点间的公平性，可是当结点注入的业务量相差较大，即在非平衡业务的情况下，会产生严重的带宽震荡。

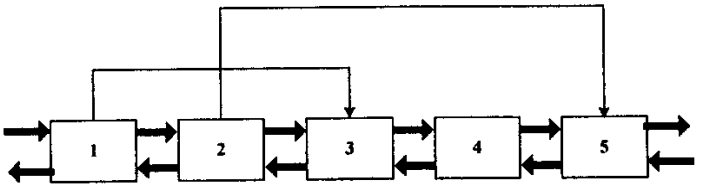

图 4-9 非平衡业务情况

如图 4-9 所示非平衡业务情况，环中有两个流，flow(1,3)和 flow(2,5)，其中 flow(1,3)尽可能的往环中注入数据，flow(2,5)的发送速率恒定为 622Mbps。

图 4-10 是采用 Gandalf 算法得到的仿真结果，我们可以看到 flow(1,3)始终在 622Mbps~1900Mbps 之间震荡。得到这样的结果与 Gandalf 算法采用的机制密切相关。因为 flow(2,5)的速率被限制在 622Mbps，当结点 2 检测到拥塞时，其向上游结点 1 发送公平信息，该公平消息携带结点 2 的本地业务发送速率（622Mbps）。

一旦结点 1 接收到来自结点 2 的公平消息后，立即将它的接入速率控制器的值设置成 622Mbps。当结点 1 以 622Mbps 发送业务时，结点 2 的拥塞将会消除，结点 1 将会不断增加其发送速率，直到结点 2 再一次发生拥塞。显然这样的带宽震荡将降低网络的吞吐量，严重影响网络的性能。在上述非平衡业务情况下，吞吐量损耗大约 15%。图 4-11 是采用 DBRR 算法得到的仿真结果，可以看到 flow(1,3)的带宽震荡几乎为 0，吞吐量几乎没有损耗。

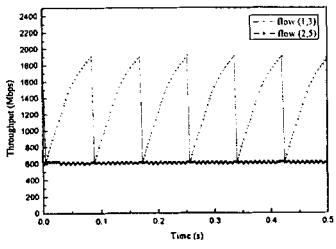

图 4-10 带宽震荡（Gandalf 算法）

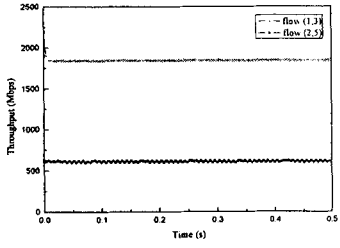

图 4-11 带宽震荡（DBRR 算法）

从以上的分析和仿真结果，我们可以得到以下结论：在非平衡业务情况下，如果在结点发生拥塞时对上游结点所有流（途经该拥塞结点）的发送速率均按照最小测量值进行调整（例如 Gandalf 算法采用拥塞结点的本地发送速率），必然会发生带宽震荡的现象，故更合适的公平带宽分配机制不应将途经拥塞结点的所有流的发送速率与最小的测量值相匹配。由于我们设计的 DBRR 算法并没有采用这种分配带宽的方式，因而很好的避免了在非平衡业务情况下可能出现的带宽震荡现象。

4.4.5 带宽分配算法收敛时间

下面我们研究带宽分配算法的收敛时间。如图 4-4 所示“平行”传送的情况，假设结点 5 为接收结点，结点 1 到 4 尽可能多的向 5 发送数据，结点 1、2、3、4 分别在 0 秒、0.1 秒、0.2 秒和 0.3 秒时向环注入数据。

图 4-12 是采用 DBRR 算法时，算法收敛的情况。可以看到当新的数据流注入到环中时，环路很快趋于稳定而且几乎没有震荡，显示出 DBRR 算法良好的收敛性。图 4-13 是采用 Gandalf 算法时，算法收敛的情况。可以看到 Gandalf 算法的收敛性不如 DBRR 算法，尽管环路最终也能趋于稳定，但是在稳定前震荡较为严重。DBRR 的收敛时间大约 2ms（2T），而 Gandalf 的收敛时间大约为 50ms。除此以外，采用 DBRR 算法的结点在 2ms 内至多发送两个公平分组。而 Gandalf 在收敛之前的 50ms 内要发送大约 500 个公平分组（Gandalf 带宽分配机制采用每 0.1ms 发送一个公平分组）。显然采用 DBRR 算法的网络开销远小于 Gandalf 算法情况。

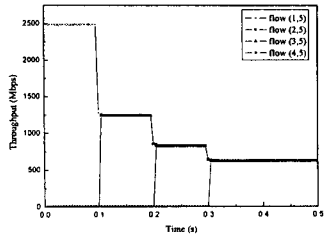

图 4-12 DBRR 算法的收敛性

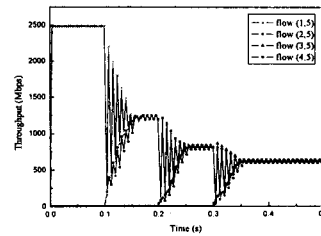

图 4-13 Gandalf 算法的收敛性

4.4.6 总业务量对分组平均传输时延的影响

下面简单讨论总业务量对分组平均传输时延的影响。首先对总业务量进行定义。总业务量是指网络中所有结点在单位时间内要求传送的帧的信息量。本文讨论的是多跳网络，总业务量是指单位时间内新产生的业务量，不包括重传业务和转发业务，即只考虑外部到达网络的业务量。我们对仿真模型做以下要求：

①网络规模不变，仍为 10 个结点，结点间距离为 30km。

②链路速率为 OC-48。

③业务源采用 Possion 源，平均分组长度为 500 字节。

④为了保证环中业务的均匀性，我们令各个结点均以相同的速率（分组到达率）向环中注入分组，并等概的向环中其它各个结点发送数据。

⑤采用最短路径路由。

⑥调度缓存（转移队列和本地发送队列）无限大。

根据[26]定理 2.4、定理 2.5、定理 2.6 以及推论 1 的结论，满足上述条件的环形网络中所有结点的注入流量对一条链路负载的贡献等于一个结点的注入流量对所有同类链路负载的贡献，也就是说从所有的结点看一条链路等于从一个结点看所有的同类链路。因而考察网络总业务量对分组平均传输时延的影响，也就是考察环中任意一个结点的业务注入量或任意一条环路链路利用率对分组平均传输时延的影响。根据上述的结论，通过仿真，考察图 4-1 所示网络中，结点 0 的业务注入量对分组平均传输时延的影响。这里需要说明的是，由于分组在每条链路上的传播时延是固定不变的，故其对分组的平均传输时延影响可忽略不计。

图 4-14 给出具体的仿真结果，可以看到分组平均传输时延随总业务量的增大而不断增大。当结点 0 的业务负载较轻时，如图中 0 到 1200Mbps 期间，平均分组传输时延增长非常缓慢，且很小（< 15 微秒）。当结点 0 的业务负载中等程度时，图中 1200 到 1600Mpbs 期间，平均分组传输时延明显增大，从大约 15 微秒很快增加到大约 50 微秒。当结点 0 的业务负载较重时，图中 1600 到 1800Mbps，分组平

均传输时延急剧增长。这里说明一点，图中 1800Mbps 对应的分组平均传输时延不是实际的测量值，在该点处无法获得平均传输时延的稳定值，即平均传输时延随仿真时间变化而变化，为了能够说明问题，在 1800Mbps 处我们任意设定了一个值。通过仿真我们还发现，当结点 0 的业务负载接近 1800Mbps 时，测得链路利用率几乎达到 100%，即链路接近饱和状态。以上仿真表明，当环路业务负载较轻或者中等时，网络能够获得很低的平均分组传输时延（微秒级）。当环中负载较重时，随着链路逐渐达到饱和，平均传输时延将急剧增长，网络性能急剧恶化。我们将分组平均传输时延增长缓慢的区间称为业务稳定区间，如图中 0 到 1600Mbps 区间。将分组平均传输时延急剧增长的区间称为业务非稳定区间，如图中 1600 到 1800Mbps。显然业务稳定区间越大越好。业务稳定区间的大小与业务源、网络拓扑和路由算法密切相关。在本例设定的条件下，业务稳定区间大约占整个业务负载区间的 90%。

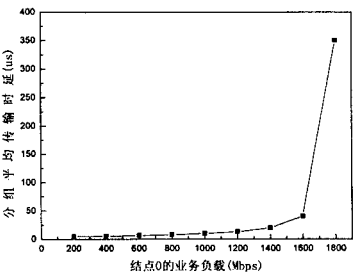

图 4-14 总业务量对分组平均传输时延的影响

§4.4 本章小结

本章通过对结点公平性、队头阻塞和带宽震荡对网络吞吐量的影响、算法收敛时间以及平均分组传输时延的仿真，综合考察了采用 DBRR 算法的 RPR 网络的性能，可以看到 DBRR 具有良好的公平性、吞吐率（避免队头阻塞，消除业务震荡）、较短的算法收敛时间和业务稳定区间内低的平均分组传输时延。仿真结果显示了消除队头阻塞的 RPR 网络的性能要比存在队头阻塞的 RPR 网络好的多。另外，由于目前建议的算法草案存在着非平衡业务情况下带宽震荡的问题，如果不加以解决，将会降低网络的吞吐量，严重影响网络的性能。因而我们在设计带宽分配算法时，从带宽分配机制上很好的避免了这一现象的出现。仿真结果达到了算法的

设计要求，通过与 Gandalf 算法比较，DBRR 算法在非平衡业务情况下几乎没有带宽震荡，吞吐量也几乎没有损耗。仿真结果表明，DBRR 的收敛速度要比 Gandalf 快的多，并且在收敛期间引入的带宽震荡非常小。最后通过仿真研究了 RPR 环形网络中总业务量对平均分组传输时延的影响，仿真结果表明平均分组传输时延随总业务量的增大而不断增大，当业务负载较轻或中等时，平均分组传输时延增长非常缓慢，且时延很小。当业务负载较重时，随着业务负载的增加，平均分组传输时延急剧增长，网络性能急剧恶化。

第五章 总结与展望

§5.1 本文的主要研究成果

弹性分组环（RPR）是城域网（MAN）技术发展的重要方向，已经受到多个国际标准化组织、研究机构及网络设备商的重视。目前RPR技术还处于研究和探索阶段，许多关键技术还有待研究。本文在综合参考各种RPR网络带宽管理标准草案的基础上，研究了RPR网络的带宽管理机制，并提出了实现方案。主要的研究成果有：

一. 提出了一个新的RPR MAC参考模型——IA参考模型。IA参考模型是为RPR优化设计的。根据城域网中数据业务占主体的特点，IA参考模型在为时延敏感（高优先级）业务预留带宽后，剩下的带宽将以每会聚流为最小颗粒统一进行分配，这一点与基于每连接分配带宽的Aladdin协议草案和基于业务优先级分配带宽的Gandalf，Darwin协议草案不同。之所以采用每目的结点数据流先经过会聚再进行带宽分配的策略，主要是针对数据业务对时延要求不高的特点，以及尽量简化MAC复杂度的角度考虑的。我们知道，如果带宽分配采用每连接的方式进行，调度算法将会十分复杂，对于象城域网这样的高速网络，MAC的实现成本非常高。而采用基于业务优先级的方式分配带宽，尽管调度算法实现起来相对每连接分配简单，但是具体为每个优先级分配多少带宽，却无法准确界定。目前基于优先级分配带宽的算法存在非平衡业务情况下带宽震荡以及算法收敛较慢等一些局限。IA参考模型由于采用基于每会聚流分配带宽的机制，在RPR结点处可以采用先来先服务（FIFO）或严格优先级（SP）服务这样简单的调度策略进行调度，因而大大简化了MAC的复杂度。在IA参考模型的设计过程中，本文首先按功能将其分为传送通路和带宽管理两部分，其中传送通路部分由帧头处理实体和帧调度实体构成；带宽管理部分由业务监视实体，公平带宽分配实体和媒质接入速率控制实体构成。然后分别对每个功能实体进行具体的定义和设计。RPR MAC参考模型是设计带宽分配算法的基础，DBRR正是基于IA参考模型设计的。

二. 基于IA参考模型，提出了一个全新的环带宽分配算法（DBRR）。DBRR的基本思想源于GPS服务规则，因而GPS具备的一些优点，DBRR同样具备，例如：GPS能够确保每连接获得的带宽不低于最小公平带宽C/N，其中C为输出链路的带宽，N为GPS系统中连接的个数（假设权值相同），同样DBRR也能确保每个聚合流获得的带宽不低于最小公平带宽C/N。这样就从分配机制上确保了环中各结点间业务的公平性。DBRR是一种分布式算法，环中各结点独立计算公平消息，

结合从下游传来的公平消息的值，动态的调整本地业务的接入速率，从而实现未用带宽的再利用。另外，DBRR 算法解决了[3，4，5]在非平衡业务的情况下带宽震荡问题。从算法收敛时间的角度考察 DBRR 算法，可以看到该算法收敛速度很快，且收敛时几乎不发生带宽震荡。

§5.2 今后有待进一步研究的问题

本文将 RPR 关键技术分为四类，即 RPR MAC 层帧结构、RPR 公平带宽管理机制、RPR 拓扑发现机制和 RPR 保护倒换机制。本文只是对其中的 RPR 公平带宽管理机制进行了深入的研究和探讨，其它三类关键技术只进行了初步的讨论，因而今后很要必要对这三类关键技术做进一步研究。

带宽管理机制是一个很大的课题，涵盖的内容很多，本文仅仅研究了公平带宽分配机制及其相关问题。带宽管理机制中还有另一类很重要的问题——服务质量（QoS）问题，本文没有深入研究。尽管 DBRR 算法具有很多优点，例如能够保证 RPR 结点间的公平性、确保每聚合流获得带宽的下限—最小的公平速率等，这些优点对改善 QoS 是有帮助的，但是 DBRR 并没有对具体业务的带宽需求加以细分（例如区分服务（Diffserv）中定义了 7 类业务，每类业务对 QoS 具有不同的要求）。与 Aladdin、Gandalf 和 Darwin 建议草案一样，DBRR 只是将可用带宽分为两类：一类为承诺带宽，这部分带宽分给对时延和时延抖动要求严格的业务（如话音业务）。承诺带宽是静态的，是运营商对 RPR 结点进行配置时，预先预留好的，占总带宽的比例非常小，不参与公平带宽分配；另一类为剩下的可用带宽，由于预先给高优先级业务分配好了带宽，因而剩下的带宽主要分给对时延和时延抖动要求不高的业务，如尽力服务（best effort）业务。这主要是针对城域网传送业务的特点设计的。前面已经介绍过，目前城域网中的业务绝大多数为对时延和时延抖动要求不高的数据业务，话音等实时性业务所占的比例很小。因而采用这种方式分配带宽，是非常简单的，也是可行的。但是随着诸如远程教学、远程医疗等交互式视频服务的兴起和普及。实时性业务对带宽的需求会越来越大，如果还用静态预留带宽的方式，必然会造成带宽资源的浪费。因而 RPR 的 QoS 保障机制，是今后需要进一步研究的问题。

除此以外，RPR 可以承载不同类型的分组，如 IP 分组、ATM 信元、E1 帧等等。这就存在一个问题，当 RPR 封装 ATM 信元时，由于 ATM 信元长度固定，且较小（53 字节），因而开销很大，约为 30%（22/（22+53））。这么大的开销显然影响带宽利用率。因此采用什么样的方法减少开销也是今后需要进一步研究的问题。

致谢

值此论文结束之时，我要衷心感谢所有关心我、爱护我、帮助过我的人。

首先感谢我的导师刘增基教授、邱智亮副教授。在论文的整个完成过程中，他们给予我悉心的指导，同时他们渊博的学识、严谨的治学态度、勤奋的工作作风和丰富的实践经验都给我留下了深刻的印象。在攻读硕士学位期间，我所取得的成绩和工作上的成果是与他们的谆谆教导分不开的。

感谢综合业务数据网国家重点实验室的徐展琦副教授对我的耐心指导，从他身上我不但得到了许多热心的帮助，也学到了很多科研知识，获得了一些重要的启发。

感谢鲍民权老师、刘焕峰老师在我论文工作过程中给予的指导和帮助。

感谢魏立军、杨帆、游骅、闫江舟、陈震、姚明昕、吕立圣、史琰、苏扬、顾华玺、刘建平、张森、解震春、杨君刚、刘故箐、杨国亮、张途、安丽芳、迟玲等师兄弟姐妹们在工作中给予我的大力支持和帮助。同时，也对重点实验室其他所有老师和同学为我在硕士论文期间提供的良好的工作氛围一并表示衷心的感谢。

感谢张雪山、苏昕、周江在日常生活中的鼓励和支持，让我感受到兄弟般的情意。

深深的感谢家人对我的学习和工作的理解和支持。同时感谢 101 室刘静，在毕设期间给予我的无微不至的关怀和毫无保留的技术帮助。

最后，对所有参加论文评审和对本文提出宝贵意见的专家、教授以及老师们表示衷心的感谢

参考文献

[1] D.Berrsekas, R.Gallager, "Data Networks", second edition, Prentice Hall, 1992.

[2] Mischa Schwartz, "Broadband Integrated Networks", Prentice Hall, 1996.

[3] A. Mekkittikul et al. Alladin Proposal for IEEE Standard 802.17, Draft 1.0, Nov. 2001.

[4] J. Kao et al. Gandalf Proposal for IEEE Standard 802.17, Draft 0.4, Nov. 2001.

[5] J. Kao et al. Darwin Proposal for IEEE Standard 802.17, Draft 1.0, Jan. 2002.

[6] A.k.Parekh, R.G.Gallager, "A Generalized Processor Sharing Approach to Flow Control in Integrated Services Networks: The Single Node-Case", *IEEE/ACM Trans. Networking*, vol.1, No.3, pp.334-357, June, 1993.

[7] S.J.Golestani, "A Self-Clocked Fair Queuing Scheme for Broadband Applications", *Proc. IEEE INFOCOM'94*, pp. 636-646, Apr. 1994.

[8] J.C.R.Benett, HuiZhang, "*W^2FQ*: Worse-case Fair Weighted Fair Queuing", *Proc. IEEE INFOCOM*, pp.120- 128, 1996.

[9] J.C.R Benett, HuiZhang, "Hierarchical Packet Fair Queuing Algorithms", *IEEE Trans. Networking*, pp.557-565, Oct. 1997.

[10]IEEE Standard 802.5-1989, IEEE standard for token ring.

[11]IEEE Standard 802.6-1990, IEEE standard for distributed queue dual bus(DQDB) subnetwork.

[12]F.E.Ross, "Overview of FDDI: The Fiber Distributed Data Interface", *IEEE J. on Selected Areas in Comm.* Vol.7, No.7, Sep. 1989.

[13]IEEE, IEEE Standard 802.17: Resilient Packet ring, http://ieee802.org//17, standard specification in progress.

[14]L. Tassiulas and J. Joung, "Performance measures and scheduling policies in ring networks", *IEEE/ACM Trans. Networking*, 4(5):576-584, Oct. 1995.

[15]D.Tsiang, G. Suwala, "The Cisco SRP MAC Layer Protocol", *IETF* Networking Group, RFC 2892, Aug. 2000.

[16]H.R.van As, "Major Performance Characteristics of the DQDB MAC Protocol",Telecommunications Symposium,1990.ITS'90 Symposium Record, SBT/IEEE 1990

[17]I.Cidon, L.Georgiadis, R.Guerin, "Improved fairness algorithms for rings with spatial reuse",*IEEE/ACM Transactions on Networking*, Vol.5, No.2,1997.

[18]I.Cidon and Y.Ofek, "Metaring – a full-duplex ring with fairness and spatial reuse", *IEEE Transactions on Communications*, 41(1):110-120. January 1993.

[19]M.Shreedhar and G.Varghese, "Efficient fair queuing using deficit round-robin",I*EEE/ACM Transactions on Networking*, 4(5):574-584, October 1995.

[20]L. Georgiadis, R. Guerin, and I. Cidon, "Throughput properties of fair policies in ring networks", *IEEE/ACM Transactions on Networking*, 1(6):718--728, December 1993.

[21]Bob Schiff, "Resilient Packet Rings (RPR)—Delivering Carrier-Class Ethernet in the MAN",Whitepaper ,LANTERN INC.

[22] C. Su, G. de Veciana, and J. Walrand, "Explicit rate flow control for ABR services in ATM networks", *IEEE/ACM Transactions on Networking*, 8(3):350–361, June 2000.

[22]Jay Shuler, "Resilient Packet Transport (RPT) for Metropolitan Area Networks", Whitepaper, Luminous Networks, Inc. September 2001.

[23]Jay Shuler, "Quality of service in Resilient Packet Transport Rings",Whitepaper, Luminous Networks, Inc. September 2001.

[24]韦乐平 编著，《光同步数字传输网》，中国通信学会主编，人民邮电出版社。

[25]杨帆，《分组调度算法及接入允许控制算法的研究》，博士论文，2002 年 6 月。

[26]刘故簪，《大容量可变长分组交换技术理论分析和仿真》，硕士论文，2002 年 12 月。

作者在攻读硕士学位期间的研究成果

学术论文

[1] Peng Yue, Zengji Liu, Jing Liu, "High Performance Fair Bandwidth Allocation Algorithm for Resilient Packet Ring", 已被 AINA2003 国际会议录用。

[2] 岳鹏、徐展琦，"IP 网络传送电路业务的新技术"，《电信快报》，2002 年第 5 期。

[3] 徐展琦、刘增基、胡强，岳鹏，"基于 IP 的 UTRAN 分组传送技术研究"，《电信科学》，2002 年第 3 期。

参加的科研项目

参与 UTRAN 分组传输技术研究。