

# 基于链路共享的分组调度算法研究及仿真

## Research and Simulation of Link-Sharing Based Packet Scheduling Algorithms

张淑清 卢 洁 应启夏

(上海理工大学光电学院, 上海 200093)

**摘 要:** 分组调度管理网络链路带宽, 是实现网络服务质量的核心技术之一。介绍了一种先进的分组调度和资源管理模型——链路共享模型; 说明了 CBQ 和 H-WF<sup>2</sup>Q+ 两种不同的实现算法; 分析了它们用于传输实时流的可行性。使用网络仿真工具 NS2, 对实时流传输的服务质量保证进行了仿真分析。仿真结果表明: 该链路共享模型可以为不同的组织、不同的通信类型提供特定的服务质量保证, 大大地提高了网络带宽的利用率。

**关键词:** 分组调度 链路共享 链路

**中图分类号:** TP301.1

**文献标识码:** A

**Abstract:** Management of network link bandwidth using packet scheduling is one of core technologies for the guaranteed service quality of the network. When an advanced packet scheduling and resource managing model—link sharing model is presented, two different implementation algorithms including CBQ and H-WF<sup>2</sup>Q+ are expounded. Then their feasibility for real-time streaming transmission is analysed. The use of a network simulating tool called NS2 analyses the simulation for the guaranteed service quality. The simulation result shows that this link sharing model can provide different organization and communications types with guaranteed specific service quality, improving utilization of network bandwidth.

**Keywords:** Packet scheduling Link sharing Link Service quality

### 0 引言

在传统的观点上, 互联网与工业控制是两个完全独立的领域。但随着网络技术的发展, 例如服务质量 (QoS) 控制、蓝牙、802.11 无线局域网 WLAN 和 Ad Hoc 等, 以及现代工业对远程实时监控系统的展需求, 现代工业网络越来越多地与 Internet 相融合。但是, Internet 采用分组交换、无连接传输技术, 尽力而为的发送尽可能多的数据, 不能为数据提供时延保证, 并不适用于时延敏感的实时 (real-time) 数据流。实时数据流对发送到网络的每个分组都有严格的时延要求。传统的 Internet 网络提供的平均性能保证已经不能适应这些数据传送要求, 网络的服务质量 QoS 控制 (quality of service) 研究得到快速的发展。

分组调度是实现网络 QoS 控制的核心技术之一。分组调度分配网络链路带宽的使用, 决定等待队列中先服务 (发送) 哪个分组。复杂的分组调度算法需要在充分利用网络资源的前提下, 分级传输用户数据, 使得输入流能够按照预定的方式共享输出链路带宽, 即满足不同的 QoS 要求, 如: 实时流是时延敏感的, 而数据流是丢失敏感的。

本文介绍了一种先进的分组调度模型: 链路共享模型, 对模型的可实现算法 CBQ, WF<sup>2</sup>Q+ 进行了研究, 并使用网络仿真工具 NS2, 对这两种算法实现的性能进行了仿真分析。

### 1 链路共享模型

链路共享方法允许不同流量个体共享一个链路上的带宽, 不同流量个体可以是个体流或汇聚流。分级链路共享允许多机构, 多协议族或不同流量类型以可控的方式共享一个链路。Internet 由多重管理域组成, 很大程度上限制了可用资源, Internet 的资源控制要考虑使用者的本地决策和端到端的性能要求。链路共享方法按照本地需要分配本地链路带宽并分离实时流 (real-time) 和尽力而为型数据流 (best-effort), 可以使用任何一种合适的分组调度算法保证实时流的端到端性能要求。

Sally Floyd 提出了一种分组交换网络中链路共享和资源管理的模型<sup>[1]</sup>, 这种模型建议使用两种调度器: 一般调度器和链路共享调度器。没有拥塞的情况下, 一般调度器可以使用任何调度算法发送分组; 网络中发生拥塞, 链路共享调度器对发送过快的数据流进行限速。链路共享调度器用最小的动作确保了各个数据流在一定时间间隔内得到初始分配的链路带宽。

图 1 所示是一个分级的链路共享结构。本地带宽首先在各机构组织 (agency) 间划分。其中某一个组织使用多种协议, 如 IP 和 IPX 基于不同协议族来划分带宽。分配给某一组织的带宽会在该组织的不同流量类型中共享。对某一特定类型的流使用带宽预留, 就可以在多连接 (connections) 中共享这部分带宽。链路共享和带宽分配可以是静态或动态的。图 1 所示的在多机构和各协议族 (protocol families) 间分配带宽应该是静态的, 应该在接入网络或网络初始化时建立分配关系。在

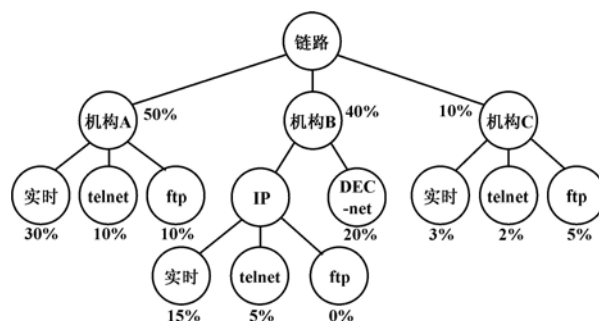


图1 链路共享结构

不同流量类型或端到端的连接类型 (real-time, telnet, ftp) 间分配带宽应该是动态的。一般调度器应该支持包括实时数据流在内的基于优先级的调度算法。链路共享可以充分利用网络资源, 未被占用的链路资源按一定的方式分配给各个流。例如, 图 1 所示, 机构 A 在某一时间段上可以使用 50% 的带宽。如果机构 A 的实时数据流没有足够多的数据需要发送, 那么剩余的带宽会分配给机构 A 的其它子类流。

链路共享的主要目标是:

- ① 每个内部类或子类在一个合适的时间间隔内应该得到基本链路带宽保证, 满足流的基本需求;
- ② 如果所有的子类和内部类已经得到其链路共享带宽保证, “过剩”的带宽不能任意分配, 应遵循合理的分配指导原则。

### 1.1 链路共享指导原则

一个类的分组可以不使用链路共享调度器调度如果它满足下列条件:

- ① 类使用的带宽不多于此类的分配带宽;
- ② 这个类的父类使用的带宽不多于此父类的分配带宽, 并且在链路共享结构中此类的子类带宽要求都得到了保证。

此外所有情况都会引起链路共享调度器的使用。一个类是否应该使用链路共享调度器, 取决于它的父类和它的子类。

## 2 两种实现算法仿真对比研究

Sally Floyd 提出了 CBQ (class-based queueing) 实现算法<sup>[1]</sup>, 链路共享调度器使用基于类的静态优先级的调度算法。Hui Zhang 提出了一种分级基于分组的公平算法 H-PFQ (hierarchical packet fair queueing)<sup>[2]</sup>, 所有的分组公平队列算法 PFQ 都可以实现 H-PFQ, 本文中我们选用 H-WF<sup>2</sup>Q+。

### 2.1 CBQ 算法仿真

CBQ 算法仿真带宽分配如图 2 所示。其中机构 A 的实时流使用 UDP 协议, 干扰流 (interactive) 以指数时间间隔猝发 4 个 1 000 比特长的分组, 机构 B 的实时流使用 UDP 协议, 干扰流以指数时间间隔发送单个 50 比特长的分组。两个机构的 FTP 类可以缓存 20 个分组, 仿真试验中每个 FTP 类有 3 个 TCP 连接。

从图 3、图 4 可以看出, 猝发到达的高优先级类得到了合适的链路带宽分配。基于优先级的 CBQ 调度减少了高优先级类的时延, 同时并没有减少低优先级的吞吐量。平滑到达的高优先级类得到很大一部分的链路带宽。CBQ 算法为高优先级类减少了时延。

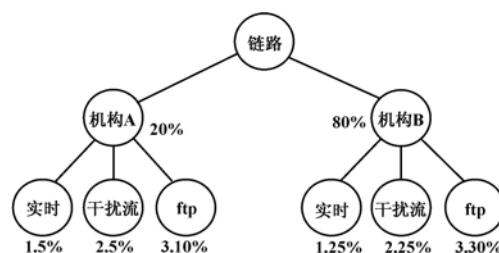


图2 CBQ算法仿真带宽分配

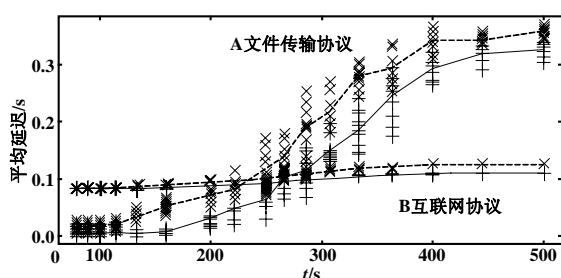


图3 猝发流到达引起的平均到达时延

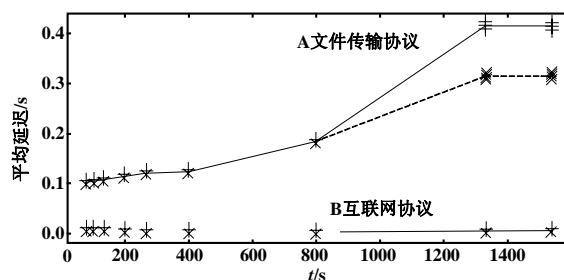


图4 平滑流到达引起的平均到达时延

### 2.2 H-WF<sup>2</sup>Q+算法仿真

H-WF<sup>2</sup>Q+算法计算系统的虚拟时间, 按照分组的虚拟完成时间排序并维持可发送对话集。虚拟时间的计算函数:

$$S_i = \begin{cases} F_i & \text{if } Q_i(\alpha_i^k) \neq 0 \\ \max(F_i, V(\alpha_i^k)) & \text{if } Q_i(\alpha_i^k) = 0 \end{cases} \quad (1)$$

$$F_i = S_i + \frac{L_i^k}{r_i}$$

式中:  $Q_i(t)$ 代表对列  $Q_i$  的第一个分组;  $S_i$ 、 $F_i$ 是分组  $Q_i(t)$  的虚拟开始时间、虚拟结束时间;  $L_i^k$  是分组  $Q_i(t)$  的包长;  $r_i$ 是对话 (session)  $i$  的确保速率。虚拟时间函数:

$$V_i(t_2) = \begin{cases} V_i(t_1) & \text{if } Q_i(\tau) = 0, \forall t_1 < \tau < t_2 \\ \max \{V(t_2), V_i(t_2-)\} & \text{if } Q_i(t_2-) = 0 \wedge Q_i(t_2) \neq 0 \\ V_i(t_1) + \frac{w_i(t_1, t_2)}{\phi_i} \sum_{j=1}^N \phi_j & \text{if } Q_i(\tau) \neq 0, \forall t_1 \leq \tau \leq t_2 \end{cases} \quad (2)$$

该算法确保了每对话连接的传送速率。

图 5 所示算法仿真中的链路带宽分配，有多等级连接，两种流量类型：TCP 和 on-off 源，用 TCP 流代表实时数据。每一等级上都有一个 on/off 源，在 on-off 源干扰情况下观察 TCP 对话流的性能。图 6 所示所有对话连接中 on-off 源激活周期。

图 7 表示每个 TCP 流的带宽分配，平均 100 ms 测量一次所以 TCP 流。100 ms 提供很好的带宽分配测量粒度，相对于路由器中的分组延迟是一个很大的数值。H-WF<sup>2</sup>Q+算法为每个 TCP 流都提供了带宽保证，基本符合网络建立时分配的带宽百分比。

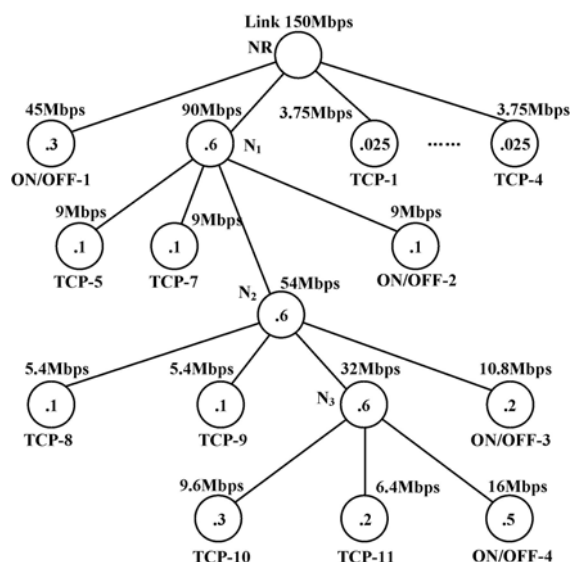


图5 H-WF<sup>2</sup>Q+算法仿真带宽分配

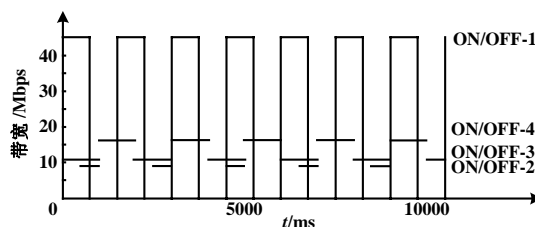


图6 On/Off源的激活周期

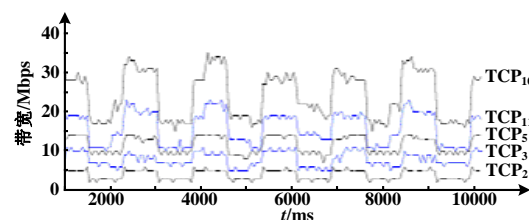


图7 TCP对话流的带宽分配

### 3 结束语

链路共享模型是一种先进的分组调度模型，主要用在广域网的边缘路由器上，使用链路共享模型可以灵活的定义用户类别，为不同的组织，不同的通信类型提供特定的 QoS 保证，在保证实时数据关键应用性能的同时大大的提高了网络带宽的利用率。CBQ 和 H-WF<sup>2</sup>Q+算法是当今具有相当优势的调度算法。

### 参考文献

- 1 Floyd, S., Jacobson, V. Link-sharing and resource management models for packet networks[M]. IEEE/ACM transactions on networking, 1995.
- 2 Bennett, Jon.C.R., Zhang, H. Hierarchical packet fair queueing algorithms[M]. IEEE/ACM transactions on networking, 1997.
- 3 徐雷鸣, 庞 博, 赵 耀. NS 与网络模拟 [M]. 北京: 人民邮电出版社, 2003.
- 4 H. Jonathan Chao, 郭晓雷. 高速网络中的 QoS 控制 (影印版) [M]. 北京: 清华大学出版社, 2004.

收稿日期: 2006-03-24。

第一作者张淑清, 女, 1981 年生, 现为上海理工大学光电学院在读硕士研究生; 主要研究方向为检测技术及自动化装置。