

文章编号: 0529-6579(2000)01-0031-05

一种 MPLS 中 VC 融合模型^{*}

林春光, 李晖晖, 魏颖琪, 张光昭

(中山大学无线电电子学系, 广州 510275)

摘 要: 根据 ATM 中支持 MPLS 的特点, 提出了一种 VC 融合模型, 并采用泊松分组到达 (PPA) 模型, 得出了融合缓冲器在 ON-OFF 模型下的各个主要性能参数的封闭表达式. 计算机仿真结果表明该理论模型具有很高的精确度.

关键词: ATM; MPLS; VC 融合; 泊松分组到达

中图分类号: TN252 **文献标识码:** A

Internet 在近年来的“爆炸性”增长使得 Internet 骨干带宽成为瓶颈^[1], 为了解决这个问题而提出了各种方案, 其中由 IETF 提出的基于集成模型的 MPLS (multi-protocol label switching)^[2] 技术由于有良好的网络扩展能力而倍受关注.

尽管 MPLS 不针对某一种链路层技术, 但是最引人注目的还是使用 ATM 技术作为它的第 2 层转发链路, 这样可以充分发挥 ATM 的优点. 在 ATM 上实现 MPLS 有其特殊的一些特性, 特别是 VPI/VCI 最长只有 28 位, 标签空间有限, 为了解决这个问题, MPLS 提出了流整合操作, 是多点对点的模式, 在 ATM/MPLS 中, 即为 VC 融合. 在融合点, 多个目的地址或目的地网络地址相同的 FEC (forwarding equivalence class-转发等价类), 如果具有相同业务质量, 可以同用一个输出标签, 合为一个 FEC, 从输出端口输出.

VC 融合可以在输入端口实现, 也可以在输出端口实现, 在输入端口实现时, 相当于信元重整成报文再通过交换网络, 相当于帧交换机, 不利于吞吐率的提高、交换的公平性和 QoS 的精确而简单的控制, 并且会产生较大的时延. 在输出端口实现流融合就可以克服以上的问题, 因此一般在输出端口设计一个融合缓冲器来实现 VC 融合的操作. VC 融合有效地减少连接的建立、维护和管理开销以及在 ATM 交换机内部建立连接的数目, 节省标签空间, 使全网路由数量从 $O(N^2)$ 变成 $O(N)$, 从而解决“VC 爆炸”问题.

1 VC 融合的缓冲模型

VC 融合缓冲模型如图 1 所示. VC 融合缓冲模块由输入调度器 SI、重组缓冲器 (RB)、融合调度器 SR、输出缓冲器和输出调度器 SO 等 5 个部分组成. 其中 SI 根据接收信元的

^{*} 基金项目: 广东省自然科学基金资助项目 (960031)

收稿日期: 1999-05-18 作者简介: 林春光 (1970 ~), 男, 博士生.

VPI/VCI来区分不同的FEC,再送到相应的RB_x中;RB_x重整数据分组;SR将RB_x中完整的分组输出到OB中;OB的作用是接收并缓存来自RB_x的数据分组,等待输出;SO将OB中信元按固定时隙,并且转换成同样的VPI/VCI.RB的数量由融合的流数量而决定,一般选用交换机的端口数N.RB和OB采用链表技术组成FIFO的队列结构,全部缓冲器共享一个物理存储器.只要RB_x中有一个完整的分组,RB就向SR发出请求,SR接收到RB_x的请求后,采用先来先服务(FCFS)的顺序服务策略,将某个RB分组串链接到OB队列尾上,由于只涉及到指针操作,因此一个或多个分组从RB_x输出到OB队列的时间非常短,一个信元周期内可以完成许多这样的操作,可以忽略操作时间,对于OB来说,它的信元到达是一个批到达过程.基于以上的调度策略,OB中的信元是按照完整分组顺序排列的,从而消除了信元交织问题,在离开OB时,所有的信元都换成同一个VPI/VCI.

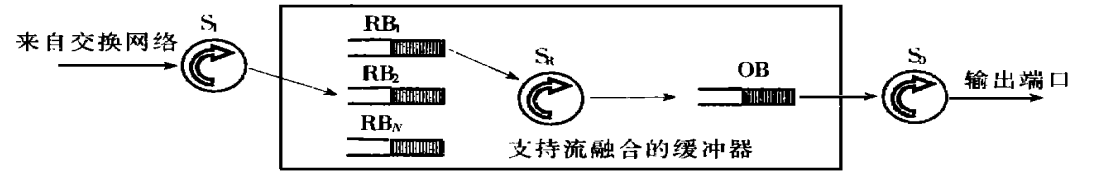


图 1 VC 融合缓冲模型

Fig.1 The buffer model of the VC merging

2 VC 融合缓冲排队模型

2.1 RB 业务模型及排队模型

为了使计算和仿真结果尽量接近实际情况,并且使推导结果比较简单,输入业务采用ON-OFF模型来表示突发业务,输入端口交替处于ON和OFF两种状态,在ON状态,连续有信元到达,在OFF状态则没有信元到达,假设从ON状态转变为OFF状态的概率为a,从OFF状态转变为ON状态的概率为s,则负载 $\rho=s/(a+s)$.处于ON状态和处于OFF状态的时间满足几何分布,即

$$P_{ON} \{ \Delta T_{ON}=k \} = a \cdot (1-a)^k, \quad P_{OFF} \{ \Delta T_{OFF}=m \} = s \cdot (1-s)^m$$

因此

$$\begin{aligned} \overline{\Delta T_{ON}} &= 1/a & \sigma_{ON} &= (1-a)/a^2 \\ \overline{\Delta T_{OFF}} &= 1/s & \sigma_{OFF} &= (1-s)/s^2 \end{aligned} \tag{1}$$

输入端口的ON-OFF到达模型如图2所示.不失一般性,在随机均匀业务模型下,交换机规模为N×N,任一个到达输入端口的信元要交换到某个特定输出端口的概率为1/N,当N很大时,到达某个特定输出端口的概率很小.在设计交换网络时,通过选用适当的加速因子,使得交换网络的吞吐率大于1,可以假设任一个到达输入端口的信元必然可以通过交换网络到达一个特定的端口而不会丢弃,因此可以认为RB的信元到达模型就如图2所示.只要有一个信元在RB中就处于ON状态,当收到分组的最后一个信元,RB由ON状态转变为OFF状态.

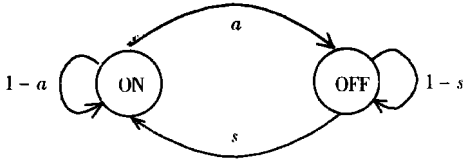


图 2 ON-OFF 到达模型

Fig.2 ON-OFF arrival model

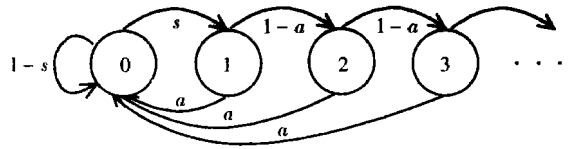


图 3 RB Markov 状态流程图

Fig.3 RB Markov status transmission

根据前面的分析, RB 的 Markov 转移图如图 3 所示。

每个融合模块有 N 个 RB 和一个 OB , 其中每个 RB 对应一个输入端口。

2.2 调度模型和分组丢弃策略

- (1) SR 按 $FCFS$ 的顺序服务策略;
- (2) 假设从 RB 输出到 OB 的时间为 0;
- (3) RB 中最多只有一个分组;
- (4) 不失一般性, 假设只有一个 OB 在工作, 并且输出端口全部带宽由 OB 使用;
- (5) 在每个信元时隙就会输出一个信元, OB 的输出调度是定长的, 记为 D ;
- (6) 如果 RB 中有一个信元丢失, 则丢弃属于这个分组的全部信元。

2.3 OB 业务模型和排队模型

为了得到 OB 的业务到达模型, 考虑 RB 中分组的离开过程. 设 X_i 表示 RB_i 的分组离开

过程, 则 $X = \sum_{i=1}^N X_i$ 为 OB 的到达过程, 其分布就是负二项分布^[3]

$$p^X(x) = \begin{cases} \frac{N+x-1}{x} (p_b)^x (1-p_b)^N, & x=0, 1, 2 \dots \\ 0 & \end{cases}$$

当 N 趋于无穷大时, 上式负二项分布的极限分布是参数为 $\lambda_p = Np_b$ 的泊松分布

$$p^X(x) = \begin{cases} \exp(-\lambda_p) \lambda_p^x / x! & x=0, 1, 2 \dots \\ 0 & \end{cases}$$

即 OB 的分组达到过程可以近似为一个泊松到达过程, 这就是 PPA 简化模型. 排队模型是 $M/G/1$, 其参数为

$$\lambda = \sum_{i=1}^N \lambda_i = N\lambda_0 = Ns / (a+s) \quad (2)$$

由于 OB 的离开是基于信元时隙的, 一个分组的服务时间应该等于该分组长度, 服从几何分布, 其平均服务时间 S 和服务时间方差 σ_s^2 分别为

$$S = 1/a, \quad \sigma_s^2 = (1-a) / a^2 \quad (3)$$

3 理论推导

根据图 3 写出 Markov 链的一步状态转移矩阵 P

$$P=\begin{bmatrix}(1-s) & s & 0 & 0 & \cdots \\ a & 0 & (1-a) & 0 & \cdots \\ a & 0 & 0 & (1-a) & \cdots \\ a & 0 & 0 & 0 & \cdots \\ \vdots & \vdots & \vdots & \vdots & \ddots\end{bmatrix}\tag{4}$$

设 p_i 为稳态时队列长度为 i 时的概率, 设 $x=(p_0, p_1, p_2, \cdots)$, 则通过求解下列方程组

$$\begin{cases} xP=x \\ \sum_{i=0}^{\infty} p_i=1 \end{cases}\tag{5}$$

最后求得它的平均队列长度 $\overline{L_{OB}}$, 报文平均时延 W 为^[4]

$$\overline{L_{OB}}=\frac{\rho}{1-\rho}\left\{1-\frac{\rho}{2}\left[1-\frac{\sigma_S^2}{S^2}\right]\right\}, \quad W=\frac{S}{1-\rho}\left\{1-\frac{\rho}{2}\left[1-\frac{\sigma_S^2}{S^2}\right]\right\}\tag{6}$$

报文在融合缓存中总的时延 T_D 为分别在 RB 中与 OB 中之和

$$T_D=E\left[T_p\right]+W=\frac{4-(2+a) \rho}{2 a(1-\rho)}\tag{7}$$

当缓存器容量为无穷大时, 平均缓存长度为 RB 与 OB 队列长度之和

$$B=N E\left[K\right]+\overline{L_{OB}}=\rho \frac{4-(2+a) \rho}{2 a(1-\rho)}\tag{8}$$

4 模型的计算机仿真

根据前面介绍的调度算法建立计算机仿真模型, 任一到达输入端口的业务模型是相同和相互独立的, 其报文长度满足参数为 $a=0.1$ 的几何分布, 即报文平均长度为 $L=10$, OFF 长度也满足几何分布, 其参数依据负载而定, 输出为每个信元时隙一个信元.

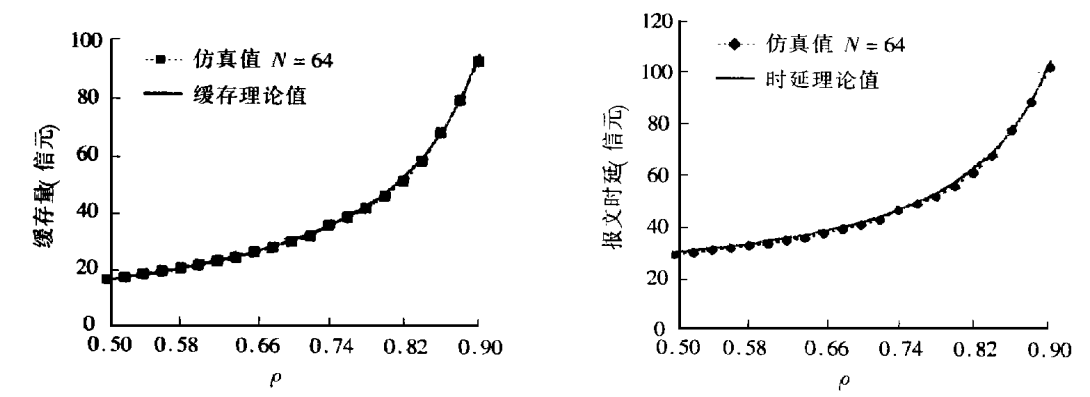


图 4 缓存量和报文时延与负载的关系

Fig.4 Buffer-load and message delay-loadrelation

为了验证理论分析结果, 这里给出报文时延和平均缓存长度的理论分析结果和计算机仿真结果. 图 4 的仿真曲线都是在 $N=64$ 时的仿真结果, 从图 4 中可以看出, 理论计算和仿真结果非常接近, 表明这里所建立的理论模型是对实际情况的很好近似, 可以用来作为分析的工具.

5 结 论

本文首先建立了 *MPLS VC* 融合缓冲器模型, 在分析 *OB* 队列时, 引入了一种新的简化模型(*PPA* 模型), 采用该模型得出了性能参数封闭表达式, 从而避免了复杂的叠代计算, 为在 *VC* 融合条件下实现带宽分配和管理以及性能监控提供一个快速算法. 用计算机模拟仿真得到的结果表明, 在 N 大于 32 时, 理论模型跟仿真结果非常吻合.

参考文献:

- [1] *SCHOEN U, HAMANN J, FUGEL A, et al. Convergence between public switching and the internet[J]. IEEE Comm Mag*, 1998, 50 ~ 65.
- [2] *CALLON R, DOOLAN P, FELDMAN N, et al. A framework for multiprotocol label switching[EB/ OL]. Internet Draft (draft-ietf-mpls-framework-02.txt)*, 1997.
- [3] 林春光. *ATM/ IP 交换技术的研究[D]. 广州: 中山大学无线电电子学系*, 1999, 84 ~ 85.
- [4] *SCHWARTZ M. Broadband integrated networks[M]. Prentice Hall*, 1996, 352.

A Model of VC Merging in Multi-Protocol Label Switching

LIN Chun-guang^{*}, LI Hui-hui, WEI Ying-qi, ZHANG Guang-zhao

Abstract: According to the characteristic of ATM, a model of VC merging in MPLS is proposed. The new PPA (possession packet arrival) model is used to deduce the closed-form of main performance measures with the ON-OFF arrival. It is shown that simulation results are closed to the theoretical model results.

Keywords: ATM; multi-protocol label switching; VC merging; possession packet arrival

^{*} Department of Radio and Electronics Zhongshan University, Guangzhou 510275, China