

一种新的缓存空间动态分配机制 及其分组丢失率分析

程东年,刘增基

(西安电子科技大学综合业务网国家重点实验室,西安 710071)

摘 要: 本文提出了一种适于分组交换机或路由器的缓存空间的动态分配机制:附加缓存块按需分配机制 AB-BA,在给出其排队模型“ $M/M/1/K(2K)$ ”的基础上,导出了其矩阵形式的稳态概率分布和分组丢失概率,并通过理论和仿真两种途径将分组丢失率性能与典型的缓存空间的静态分配系统进行了分析比较.最后分析了附加块动态分配决策机制应考虑的几个重要因素并给出了相应的动态分配算法.

关键词: 分组交换;异步转移模式;服务质量;缓存管理;排队模型

中图分类号: TN913 **文献标识码:** A **文章编号:** 0372-2112 (2001) 05-0634-04

A Novel Type of Dynamic Buffer Allocating Mechanisms and an Analysis of Its Packet Loss Performance

CHENG Dong-nian, LIU Zeng-ji

(National Key Laboratory on Integrated Services Networks, Xidian University, Xi'an 710071, China)

Abstract: This paper proposes a new type of buffer management mechanisms, called “Additional Buffer Block Allocation on demand”(ABBA) scheme, suited for packet switches and routers. With the essential of the proposed mechanism presented, we establish the associated queuing model, denoted by “ $M/M/1/K(2K)$ ”, and then derive the steady-state probability distribution and related packet loss probability. By both analysis and simulations, we evaluate the performance of the proposed ABBA scheme in terms of the packet loss probability and the equivalent buffer size K_e in comparison with typical schemes with static buffer allocation. Finally provided in this paper are some key factors used to determine whether an additional buffer block application is approved or not by a decision-making process, which is described in an algorithm form.

Key words: packet switching; ATM; QoS; buffer management; queuing model

1 引言

现有的缓存空间分配机制^[1~5]大都采用静态方法.由排队论的基本结论^[6]可知,绝大多数情况下顾客等待空间均处于非充满状态,但由于该空间中的未使用部分是“已分配”了的资源,故绝大多数情况下缓冲器均处在非充分利用的状态;此外,就资源利用而言,理想的缓存空间分配方法应是动态的按需分配,这既能大大提高缓冲器的利用率又能很好地适应各业务流到达特性的动态变化,从而更好地确保各业务流的 QoS ^[7,8],尤其是分组丢失率要求.文献^[9]提出了一种根据剩余缓存空间大小动态调整队列门限的机制,但由于不是按需分配,故仍不能使缓存空间得到充分利用.

基于“动态按需分配”的思想,本文提出了一种缓存空间的动态分配方法:基于“附加块”的按需分配机制(ABBA),由于既考虑了对缓存空间的按需动态分配与释放,以适应各业

务流的随机波动所导致的对缓存需求量的动态变化,又将动态分配的容量限制在一定水平上,以减小由于引入动态分配而带来的复杂性,因此,该机制在有效利用缓存空间与确保各业务流分组丢失率性能之间取得了较好的折衷.在描述了该机制工作机理之后,给出了 Poisson 到达、指数服务时间分布条件下其对应的排队模型,导出了其矩阵形式的稳态概率分布和分组丢失概率.在此基础上,分别通过理论和仿真两种途径将分组丢失概率与采用静态缓存分配的“ $M/M/1/K$ ”系统进行了分析和比较,最后,分析了决定附加块动态分配决策过程的几个重要因素并给出了相应的动态分配算法.

2 “缓存空间按需分配”的机理

本文提出的 ABBA 机制建立在“缓存块”的概念上.假设网络结点在连接建立阶段为一用户连接分配的(基本)缓存容量为 K ,称接纳到达分组进入的缓存空间为该连接的“工作

块”:在数据传送阶段,一旦因工作块已被占满而使新到达的分组不能进入系统,便立即申请附加的缓存空间,如果可以,管理者为其分配一片容量仍为 K 的附加缓存空间,称其为该连接的“附加块”;随后,当前和以后到达的分组便只进入其附加块而不再进入原先的工作块.需要指出,一旦申请成功,附加块就变成了新的工作块,而原工作块变为只接受系统服务而不再接纳到达分组的“过渡块”,此后,由于后续到达分组不再进入该过渡块,故它最终会被清空而释放,然后系统才为新的工作块服务.

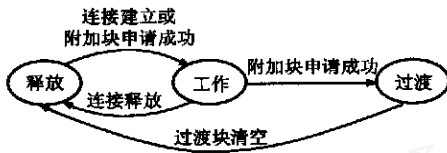


图1 缓存块的状态及演变过程

综上所述,一个缓存块的状态有“工作”、“过渡”和“释放”三种,图1描述了每个缓存块的状态演变过程.

3 “附加块按需分配”机制的排队模型及分析

设一个分组交换机采用上一节给出的附加块动态“分配-释放”的缓存空间管理机制,它以连接为单位分配缓存空间,所分配的工作块或附加块容量均为 K .分组的到达为一个参数为 λ 的 Poisson 过程,系统对其分组的服务时间服从参数为 μ 的负指数分布,采用先来先服务的服务规则.当连接的工作块已被占满且又到达一个新的分组时,便向系统申请一个附加缓存块,其申请成功的概率为 α ,如果申请失败,则丢弃到达分组,而一旦附加块申请成功,新到达及后续到达的分组便均进入新的工作块,过渡块继续接受服务并在清空后被释放.为简化复杂性,限定任一个连接在任何时刻同时拥有的缓存块均不超过两块.

记这一系统为“ $M/M/1/K(2K)$ ”并设其状态为 $(Q(t), q(t))$,其中 $Q(t)$ 和 $q(t)$ 分别表示时刻 t 工作块与附加块的分组队列长度, $q(t) = 0$ 表示在时刻 t 不存在过渡块.假设在处于状态 $(K, 0)$ 时每到达一个新分组便申请一次附加块.由于每次附加块申请成功的概率为 α ,因此从状态 $(K, 0)$ 到 $(1, K)$ 的转移率仅相当于正常情况下的 α 倍,即 $\alpha\lambda$.于是可

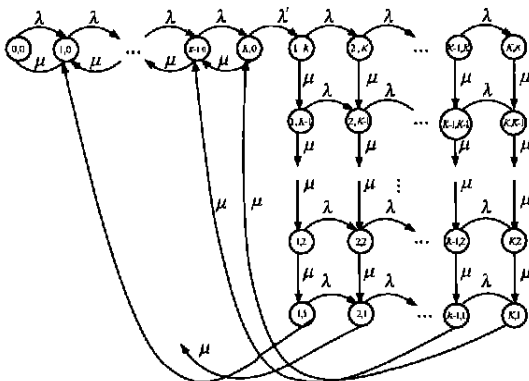


图2 “ $M/M/1/K(2K)$ ”系统的状态转移关系

得到图2所示的该排队系统的状态转移关系.记 $P_{i,j}$ 为系统处于状态 (i, j) 的稳态概率,则该系统的稳态概率方程组为:

$$\begin{cases} P_{0,0} = \mu P_{1,0}, \\ (\lambda + \mu) P_{i,0} = P_{i-1,0} + \mu P_{i+1,0} + \mu P_{i,1}, & i = 1, 2, \dots, K-1 \\ (\lambda + \mu) P_{K,0} = P_{K-1,0} + \mu P_{K,1}, \\ (\lambda + \mu) P_{1,k} = P_{K,0}, \\ (\lambda + \mu) P_{i,k} = P_{i-1,k}, & i = 2, 3, \dots, K-1 \\ \mu P_{K,k} = P_{K-1,k}, \\ (\lambda + \mu) P_{1,j} = \mu P_{1,j+1} & j = 2, 3, \dots, K-1 \\ (\lambda + \mu) P_{i,j} = P_{i-1,j} + \mu P_{i,j+1}, & 2 \leq i, j \leq K-1 \\ \mu P_{K,j} = P_{K-1,j} + \mu P_{K,j+1}, & j = 2, 3, \dots, K-1 \\ (\lambda + \mu) P_{1,1} = \mu P_{1,2}, \\ (\lambda + \mu) P_{i,1} = P_{i-1,1} + \mu P_{i,j+1}, & i = 2, 3, \dots, K-1 \\ \mu P_{K,1} = P_{K-1,1} + \mu P_{K,2}, \end{cases} \quad (1)$$

由图2或式(1)可得到系统的状态转移率矩阵 M ,记 P 为系统的稳态概率分布向量,即

$$P = [P_{0,0}, P_{1,0}, \dots, P_{K,0}, P_{1,K}, P_{2,K}, \dots, P_{K,K}, P_{1,K-1}, P_{2,K-1}, \dots, P_{K,1}]$$

则有

$$\begin{cases} P \cdot M = 0 \\ P \cdot e = 1 \end{cases}$$

其中 e 为其 $(K^2 + K + 1)$ 个元素均为 1 的列向量.假设用方程 $P \cdot e = 1$ 取代方程组 $P \cdot M = 0$ 中的第 i 个方程 ($i = 1, 2, \dots, K^2 + K + 1$) 后,得到的新稳态概率方程组为 $P \cdot M_i = E_i$, E_i 为其第 i 个元素为 1 而其余 $(K^2 + K)$ 个元素均为 0 的行向量,于是有,

$$P = E_i M_i^{-1} \quad (2)$$

系统的分组丢失率可表示为:

$$Pr\{\text{分组丢失}\} = (1 - \alpha) P_{K,0} + \sum_{j=1}^K P_{K,j} \quad (3)$$

不难理解,当附加块申请成功的概率为零时,必有 $P_{i,j} = 0 (0 \leq i, j \leq K)$ 且 $Pr\{\text{分组丢失}\} = P_{K,0}$,此时,该附加块的按需动态分配机制便退化成为采用静态缓存分配的普通“ $M/M/1/K$ ”系统,可见,“ $M/M/1/K$ ”模型仅是“ $M/M/1/K(2K)$ ”系统的一个特例.图3绘出了“ $M/M/1/K(2K)$ ”和“ $M/M/1/K$ ”两个系统分组丢失率的理论曲线.为进一步比较和验证理论分析,还在混合的、Markov 调制的双 Bernoulli 到达模型下,对采用“附加块”机制的“ $NMMBP(2)/D/1/K(K+K)$ ”和不采用该机制的“ $NMMBP(2)/D/1/K$ ”两个系统进行了仿真研究,其分组丢失率也绘制于图3中,由此可看到以下五个重要事实:

(1) 当附加块申请成功的概率 α 不变时, K 的增大能明显减小两个系统的分组丢失率,并且二者的分组丢失率之比也随 K 的增加而减小.但另一方面,大幅度(如成倍)增加 K 使二者的分组丢失率之比降低的幅度始终小于略微增大附加块申请成功概率(比如增大 10%)而带来的效果,这说明,附加块申请成功概率的大小对用户业务流的分组丢失率起着决定性的作用;

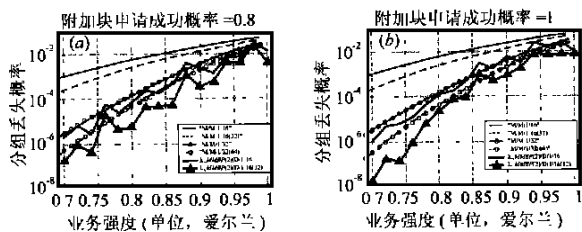


图3 “附加块”机制的分组丢失概率:理论及仿真结果

(2) 当 $\alpha = 1$ 时, “M/M/1/K(2K)”系统的分组丢失率性能比 $\alpha < 1$ 时的情形有着本质上的改善, 这集中表现在此时两个系统的分组丢失率之比有了明显减小, 尤其当 K 较大或系统的业务负荷较小(如小于 0.6)时(未绘出). 其原因在于, 当 $\alpha < 1$ 时, “M/M/1/K(2K)”系统的分组丢失概率必然包括系统稳态概率 $P_{K,0}$, 并且这一概率在数值上构成了该系统在 $\alpha < 1$ 时其分组丢失概率的主导成分; 而当 $\alpha = 1$ 时, 其分组丢失概率便不再包括 $P_{K,0}$, 从而使整个分组丢失概率大大减小. 这再次说明附加块申请成功概率的大小对“M/M/1/K(2K)”系统分组丢失率性能的改善起着决定性的作用, 而缓存块容量 K 的增大对分组丢失性能改善所起的作用却相对较小.

(3) 若在分组丢失概率相同的条件下考察这两个系统, 则“M/M/1/K(2K)”系统是比“M/M/1/K”系统更节省缓存空间容量的机制. 设此时的两个系统分别为“M/M/1/ K_e ”和“M/M/1/K(2K)”, 表 1 中的“等价缓存容量 K_e ”表明两个系统的分组丢失概率在“最小二乘”意义相等条件下, 静态分配系统的等价缓存空间容量. 由表 1 可知, K_e 总是大于 K , 且当 K 较小时, 二者容量之比 K_e/K 会更小, 即“M/M/1/K(2K)”系统节省缓存空间的效果就更明显. 这说明, “M/M/1/K(2K)”系统更适合工作于较小的缓存块. 值得注意的是, 分配较小的缓存块还能使总的剩余缓存空间容量更大, 从而形成各业务流附加块申请成功概率的增大进而使各业务流分组丢失概率进一步减小的良性循环.

表 1 “等价缓存容量 K_e ”与 K 的关系

(α, K)	(0.6, 8)	(0.7, 12)	(0.8, 16)	(0.9, 20)	(1.0, 32)
K_e/K	80 %	81.1 %	82.5 %	83 %	85.8 %

(4) 在 K_e 取值的所有情况下, 两个系统分组丢失概率之比(未绘出)均在系统的业务负荷较轻时(如 $\rho \leq 0.6$)取得较小值, 这就是说, 相对于典型的“M/M/1/K”系统而言, 带附加块动态分配机制的“M/M/1/K(2K)”系统的分组丢失率性能可在系统轻负荷条件下得到更好的改善.

(5) 在 α 取 0.8 和 1 的两种情况下, 分组丢失概率的仿真曲线与相应的理论曲线具有相同的变化趋势, 并且采用“附加块”机制后的仿真结果仍然小于普通的静态机制, 而仿真与理论结果的差别仅在于前者具有更小的绝对数值. 这表明仿真结果验证了理论分析的正确性.

综上所述, 相对于静态缓存分配机制而言, 本文提出的附加块动态分配机制的分组丢失率性能必然得到改善. 通常, 当总的剩余缓存空间较大时, 则只要申请的容量可以满足便为其分配相应的附加块. 但是, 由于附加块的排空需要一定时

间, 且系统中各业务流的“重要程度”不同, 故当系统内总的剩余缓存空间较小时, 对附加块的分配就应考虑更多的因素, 以使系统内各业务流的总体分组丢失率性能最佳. 下面, 就讨论影响附加块的动态分配的几个重要因素并给出相应的附加块动态分配算法.

4 基于“优先级”的附加块动态分配算法

为了实现简单, ABBA 机制采用非剥夺的附加块分配. 由于可能同时存在着多个具有不同丢失性能要求的业务流, 因此, 应当优先考虑“较重要”连接的附加块分配请求. 具体说, 应在为不久的将来也要申请附加块的更高级连接都预留了相应的附加块的基础上, 尽量满足当前请求者对附加块的要求, 而这种附加块的预留正是上述“优先原则”的体现. 为此, 附加块的分配应当考虑连接的丢失优先级、各连接的分队队长及其变化趋势、各高级连接对附加块的未来需求量等因素. 这里, 通过“丢失优先级”、“队长变化趋势因子”、“潜在附加块因子”和“分配偏爱因子”的概念支持基于“优先级”的附加块动态分配.

4.1 连接的“丢失优先级”

为每一个连接(或连接类) C_i 定义一个表征其重要程度的优先数 $i_i (> 0)$, $i_i > j_j$ 表示连接 C_i 的丢失优先级高于 C_j . 由于 i_i 只有与 j_j 相比较时才有意义 ($i_i \sim j_j$), 故优先数 i_i 的值只具有相对意义.

4.2 队长变化趋势、趋势预测及趋势因子

高级连接队列长度的变化趋势是对其未来附加块需求总量估计的重要考虑因素. 总体上, 队长变化可有三种趋势: 减小、维持不变和增加. 显然, 当高级连接的工作块即将被充满且其队列长度呈增加趋势时, 它在不久的将来就很有可能也申请附加块(称为“潜在附加块”). 对队长变化趋势预测得越准确, 则对“潜在附加块”需求量的估计精度也越高, 但另一方面, 相应的预测算法通常也会越复杂. 由于高速网络环境更需要简单而有效的预测算法, 因此, 可采用“指数平滑”预测算法. 记 $q_i(t_r)$ 为 t_r 时刻连接 C_i 工作块中的队列长度, $\tilde{q}_i(t_r)$ 为在 t_{r-1} 时刻得到的对其队长在 t_r 时刻的估计值, $0 < i_i < 1$, 则有:

$$\begin{cases} \tilde{q}_i(t_r) = i_i q_i(t_{r-1}) + (1 - i_i) \tilde{q}_i(t_{r-1}) \\ \tilde{q}_i(t_0) = 0 \end{cases} \quad (4)$$

在实际中, 可选 $i_i = \rho_i / \phi_i$, 其中 ρ_i 为第 i 个业务流(或第 i 类业务流)的分组平均到达率, ϕ_i 为其相应的分配带宽. 式(4)中的 $t_r (r = 1, 2, \dots)$ 是对连接 C_i 队列长度的采样时刻, 相应的采样间隔 T_i 可用分组平均到达间隔 $1/\rho_i$ 作为参照量, 即,

$$T_i = 1/\rho_i, \quad > 0 \quad (5)$$

在得到队长的估值 $\tilde{q}_i(t_r)$ 后, 可进一步对连接 C_i 的队长变化作出估计并用其“队长趋势因子” $i_i(t) \in [0, 1]$ 来表示这一估计, 记 $q_i(t_r) = \tilde{q}_i(t_r) - q_i(t_{r-1})$, 则可将 $i_i(t)$ 简单地定义为:

$$i(t_r) = \begin{cases} 0, & q_i(t_r) < 0 \\ 1/2, & q_i(t_r) = 0 \\ 1, & q_i(t_r) > 0 \end{cases} \quad (6)$$

这样,可用趋势因子的取值 0、1/2、和 1 分别表征队长变化的负增长、零增长或正增长趋势,也可用其表示队长增长的程度高低.下面将看到,趋势因子 $i(t)$ 可用来估算“潜在附加块”的需求量.

4.3 潜在附加块因子、潜在附加块需求量

我们把对任一高级连接“潜在附加块需求量”的估计放在对其附加块申请成功“可能性”大小估计的基础上,并称高级连接 C_j 的这种可能性为它的“潜在附加块因子”,记作 j . 决定 j 的因素主要有三个:连接 C_j 的队长趋势因子 i_j 、分组平均到达率与分配带宽之比 λ_j/ϕ_j 以及其工作块的充满程度 Q_j/K_j ,显然, j 与这三个因素均成正比关系.据此,可将 j 定义成:

$$j = i_j \times \frac{\lambda_j}{\phi_j} \times \frac{Q_j}{K_j} \quad (7)$$

其中 $0 \leq i_j \leq 1$. 一个连接的 i_j 值越接近 1,它在不久的将来申请附加块且成功的可能性也就越大.

此时,可进一步将连接的“潜在附加块需求量”定义为 $j \cdot K_j$,它表征网络结点为其在不久的将来事先预留的缓存空间容量大小.

4.4 偏爱因子

一个连接从申请得到附加块到释放其过渡块的时间长短也可用来指导对其附加块的分配决策.显然,该间隔越短,对剩余缓存空间减小的副作用越小,因而为其分配附加块的理由也越充分.本文可将这种充分性表征为一个新的物理量:“附加块分配偏爱因子”:

$$i = d_i / \bar{d} \quad (8)$$

其中: $d_i = K_i / \phi_i$, $\bar{d} = \frac{1}{N} \sum_{k=1}^N d_k$, d_i 表示连接 C_i 的过渡块被排空所需要的时间; \bar{d} 则为所有连接的过渡块的平均排空时间.显然, i 越小,表明对连接 C_i 过渡块的相对排空时间越短,因而分配机制对它也就越“偏爱”.

4.5 附加块分配决策函数及算法

在引入连接丢失优先级、分组队长变化趋势因子、附加块潜在因子及 i 的基础上,可构造如下的附加块分配决策函数 F_i :

$$F_i = K - \sum_{j > i} j K_j - i K_i \quad (9)$$

K_i 和 K 分别为连接 C_i 的缓存块容量和当前系统中总的剩余缓存空间容量.相应的决策算法可简洁地表示为:

“若 $F_i \geq 0$,则为连接 C_i 分配一个附加块;否则,拒绝本次请求.”

5 结语

为了确保每一业务流分组丢失率性能的前提下更充分

地利用缓存空间资源,本文提出了一种新的缓冲器管理机制:“附加缓存块按需分配”机制 ABBA,该机制所引入的基本缓存空间(工作块)与附加空间(附加块),使得每一业务流(或业务流类)的缓存空间容量具有了一个适度的动态变化范围,从而能较好地适应业务流对缓存需求量的动态变化.本文给出了 ABBA 机制的排队模型“M/M/1/K(2K)”,导出了其矩阵形式的稳态概率分布和分组丢失概率,通过理论和仿真两种方法将其与采用静态缓存分配的系统进行了比较分析,结果表明,ABBA 的分组丢失率性能就优于不采用附加块机制的静态系统;附加块申请成功的概率对 ABBA 的分组丢失率性能起着主导性作用;ABBA 比静态系统更能节省对缓存空间的分配量.最后,研究了附加块动态分配的决策过程应考虑的几个重要因素,给出了相应的分配决策函数和分配算法.目前,基于该附加块动态分配决策过程的模拟研究正在进行之中.

参考文献:

- [1] A K Choudhury, L E Hahne. Space priority management in a shared memory ATM switch [A]. Proc. of IEEE GLOBECOM '93 [C], Houston, 1993:1375 - 1383.
- [2] A Y Lin, J A Silvestre. Priority queueing strategies and buffer allocation protocols for traffic control at an ATM integrated broadband switching systems [J]. IEEE J-SAC, 1991, 9(9): 1524 - 1536.
- [3] D W Petr, V S Forst. Nested threshold cell discarding for ATM overload control: optimization under loss constraints [A]. Proc. of IEEE INFOCOM '91 [C], Bal Harbour, 1991:1403 - 1412.
- [4] L Tassiulas, Y Hung, S S Pauwar. Optimal buffer control during congestion in an ATM network node [A]. Proc. of IEEE INFOCOM '93 [C], San Francisco, 1993:1059 - 1066.
- [5] I Gidon, L Georgiadis, R Guerin et al. Optimal buffer sharing [J]. IEEE J-SAC, 1995, 13(7): 1229 - 1240.
- [6] ITU-T. Traffic control and congestion control in B-ISDN (Rec. I. 371) [S]. Geneva: ITU-T, 1993.
- [7] ATM Forum, ATM Forum Traffic Specification [M], Version 4.0, Apr. 1996.
- [8] R W Wolff. Stochastic Modeling and the Theory of Queues [M]. NJ: Prentice-Hall, Inc., 1989.
- [9] A K Choudhury, L E Hahne. Dynamic queue length thresholds for shared-memory packet switches [J]. IEEE/ACM Trans. on Networking, 1998, 6(2): 130 - 140.

作者简介:



程东年 1982 年于解放军工程技术学院(解放军信息工程大学前身)获工学学士学位,1990 年于国防科技大学获工学硕士学位,现在西安电子科技大学攻读通信与信息系统专业博士学位,副教授.主要研究方向:综合业务网络技术、ATM 的理论及技术、业务量管理及服务质量保证等.