

# 面向多业务通信调度的优先级排队模型研究

柯 鹏<sup>1</sup> 金姗姗<sup>2</sup> 李文翔<sup>3</sup>

(武汉科技大学计算机科学与技术学院 武汉 430072)<sup>1</sup> (武汉大学电子信息学院 武汉 430080)<sup>2</sup>  
(武汉科技大学信息科学与工程学院 武汉 430072)<sup>3</sup>

**摘 要** 互联网业务的快速发展为数据的高效传输提出了挑战,需要一种有效的数据包调度方案来满足各项业务的QoS需求。基于数据包的时延敏感性、长度等属性,采用优先级排队模型研究各种调度策略的性能,指出非抢占式短包优先策略能提供最小的总体时延;在此基础上对时延敏感业务和非时延敏感业务采用不同的策略,设计最优化传输性能的业务调度多优先级排队模型。NS-2平台下的仿真实验证明了所提模型的优越性,并演示了不同流量下的数据包排队特性,以及包交付率、吞吐量、时延等指标的变化规律,从而为性能优化提供了参照。

**关键词** 业务属性,服务质量,数据包调度,优先级排队

**中图分类号** TP393 **文献标识码** A

## Priority-Queuing Model for Multiple Services Scheduling

KE Peng<sup>1</sup> JIN Shan-shan<sup>2</sup> LI Wen-xiang<sup>3</sup>

(College of Computer Science and Technology, Wuhan University of Science and Technology, Wuhan 430072, China)<sup>1</sup>

(College of Electronic Information, Wuhan University, Wuhan 430080, China)<sup>2</sup>

(College of Information Science and Engineering, Wuhan University of Science and Technology, Wuhan 430072, China)<sup>3</sup>

**Abstract** The fast-developing services in Internet pose challenges to the efficient transmission of various types of data, so an effective packet scheduling scheme is needed to meet the QoS constraints of heterogeneous services. In this article, the priority-queuing model was used to study the performances of various strategies based on delay sensitivity and packets length. And the non-preemptive short-packet-first strategy was proved to result in the minimal overall delay. With different strategies for delay-sensitive and non-delay-sensitive services, an optimal priority-queuing model for the scheduling of multiple Internet services was designed based on above conclusions. The results from simulation experiments in NS-2 verified the superiority of the model. The results also demonstrated the features of packets queuing under different traffic scenarios and the law of variation for such performance indices as packet delivery ratio, throughput and average delay, which can be used to design effective measures for performance optimization.

**Keywords** Service properties, Quality of service, Packets scheduling, Priority-queuing

## 1 引言

随着互联网的飞速发展,其所支持的通信业务在类型、长度、QoS需求等方面都具有很大的差异,需要一种针对异构数据包的调度方案,满足各类快速增长的业务量的QoS需求。本文探讨了用于最小化网络的面向多业务调度总体通信时延的优先级队列模型及处理策略。本文的主要工作为:1)针对非时延敏感业务,基于时延敏感性和长度属性提出多种优先级排队策略,并分析了它们的通信时延、队列长度和抢占概率,证明了非抢占式短包优先策略具有最小的总体时延;2)对时延敏感业务采用抢占式策略,对非时延敏感业务采用非抢占式短包优先策略,提出具有最优化传输性能的多业务调度的优先级队列模型。

本文第2节介绍相关的研究现状,指出值得探讨改进的

方向;第3节重点针对非时延敏感业务的性能指标进行最佳策略的分析;第4节进行详细的仿真实验和分析;最后给出结论。

## 2 相关工作

针对给定优先级的数据包的调度,目前广泛采用单服务器优先级队列模型进行研究,对数据包冲突的基本处理方式包括抢占式(Preemptive)和非抢占式(Non-Preemptive)两大类。相关研究工作对抢占式、非抢占式调度策略展开分析,得到等待时延<sup>[1,2,5]</sup>、队列长度<sup>[1,2,4]</sup>、包阻塞概率<sup>[4,8]</sup>及丢包率<sup>[3,6,7,9]</sup>等指标,而这些工作尚未考虑业务的更多属性。抢占式短包优先策略<sup>[10,11]</sup>利用数据包长度的差异性,进一步降低了传输时延。本文在以上这些工作的基础上,基于业务的时延敏感性和数据包长度,探讨所有可能策略的性能,并在实

到稿日期:2012-05-21 返修日期:2012-10-15 本文受国家自然科学基金(61003127)资助。

柯 鹏(1977-),男,讲师,主要研究方向为智能计算、智能网络,E-mail:ke\_peng\_cn@yahoo.com.cn;金姗姗(1985-),女,硕士生,主要研究方向为网络流量工程;李文翔(1979-),男,讲师,主要研究方向为无线传感网。

验的基础上,给出了从中寻找到的最优化的一种策略。

### 3 非时延敏感业务的最优策略选择和分析

#### 3.1 网络业务调度问题描述

网络业务种类很多,它们具有截然不同的特性,对于时延敏感应用,其信息量较小,但对传输可靠性、时延的要求则很高,一般赋以抢占式优先权。要实现整体业务数据通信的时延最优化,在发生调度冲突时,需要考虑的问题包括:1)对多个重要程度相同的非时延敏感业务,如何设定优先级?2)对非时延敏感业务,应该采用抢占式还是非抢占式来处理调度冲突?

#### 3.2 基于数据包长度的各种优先级排队策略

针对上面的2个问题设计排队模型,设系统中存在2种非时延敏感业务,分别是  $App_S$  (如数据消息业务)和  $App_L$  (如多媒体数据业务),其中  $App_S$  的数据包长度  $Len_S$  小于  $App_L$  的数据包长度  $Len_L$ ,上述业务均为 Poisson 输入过程,其数据率分别为  $rate_S = \lambda_S * Len_S$ ,  $rate_L = \lambda_L * Len_L$ ,  $\lambda_S$  和  $\lambda_L$  分别为对应 Poisson 流的到达率。因为数据包的处理时间通常依照一般分布,故将排队模型设定为 M/G/1 型。基于此模型列举出全部的优先级设置策略如下:

1) 不区分优先级(NonPr):  $App_S$  与  $App_L$  同优先级,等同于传统的 FCFS 策略。

2) 抢占式长包优先(PrmLF):  $App_L$  对  $App_S$  具有抢占式优先权。  $App_S$  接受服务时,如果有  $App_L$  数据包到达,则此数据包立即被处理,  $App_S$  数据包回到  $App_S$  的队首等待下一次发送机会。

3) 抢占式短包优先(PrmSF): 与 PrmLF 处理方式类似,只是  $App_S$  与  $App_L$  的相对优先级对调。

4) 非抢占式长包优先(NPrLF):  $App_L$  对  $App_S$  具有非抢占式优先权。当  $App_L$  数据包到达时,它等待正在处理的数据包发送完后,立即获得服务权。

5) 非抢占式短包优先(NPrSF): 与 NPrLF 处理方式类似,只是  $App_S$  与  $App_L$  的相对优先级对调。

在这5个策略中找出具有最小总体时延的一个,即可解决前面提出的2个问题。

#### 3.3 传统多优先级排队模型的相关结论

令  $E(B)$ 、 $E(B^2)$  分别为数据包平均服务时间的数学期望和2阶原点矩,则有流量强度为  $\rho = \lambda * E(B)$ 。新到达数据包的等待时间由2部分组成,一是前面队列中的  $E(L)$  个顾客的服务时间,二是到达时以概率  $\rho$  存在一个正在接收服务的顾客,其残留服务时间为  $E(R)$ ,故等待时间为  $E(W) = E(L)E(B) + \rho E(R)$ ,结合 little 定律  $E(L) = \lambda E(W)$ ,得等待时间  $E(W) = \rho E(R) / (1 - \rho)$ ,此外由文献[12]得残留服务时间为:

$$E(R) = E(B^2) / (2E(B)) \quad (1)$$

由文献[12]中对  $E(W_i)$  及  $E(R_i)$  的分析可以得到,在非抢占式处理策略中,第  $i$  优先级的数据包的平均驻留时间为  $E(W_i)$  与  $E(B_i)$  之和,即:

$$E_{NPr}(S_i) = E(B_i) + \frac{\sum_{j=1}^N \rho_j E(R_j)}{(1 - (\rho_1 + \rho_2 + \dots + \rho_i))(1 - (\rho_1 + \rho_2 + \dots + \rho_{i-1}))} \quad (2)$$

同理由文献[12],得到抢占式处理策略中第  $i$  优先级的数据包的平均驻留时间为:

$$E_{Prm}(S_i) = \frac{E(B_i)}{1 - (\rho_1 + \rho_2 + \dots + \rho_{i-1})} + \frac{\sum_{j=1}^N \rho_j E(R_j)}{(1 - (\rho_1 + \rho_2 + \dots + \rho_i))(1 - (\rho_1 + \rho_2 + \dots + \rho_{i-1}))} \quad (3)$$

#### 3.4 最小总时延策略的选择

定理1 非抢占式短包优先调度策略的总时延小于非抢占式长包优先调度策略的总时延,即

$$Delay_{NPrSF} < Delay_{NPrLF} \quad (4)$$

证明:令  $Len_L = K * Len_S$ ,  $K > 1$ 。第一步求 NPrSF 策略的总时延。设短包优先级为1,长包优先级为2。服务时间的期望分别为:  $E'(B_1) = B_0$ ,  $E'(B_2) = KB_0$ ,二阶矩分别为:  $E'(B_1^2) = Y$ ,  $E'(B_2^2) = K^2 Y$ ,并有  $E'(R_2) = KE'(R_1) = KY / (2B_0)$ 。

由式(1)、式(2)得到短包的总时延:

$$E'(S_1) = \frac{\lambda_S E'(B_1) E'(R_1) + \lambda_L E'(B_2) E'(R_2)}{1 - \lambda_S E'(B_1)} + E'(B_1) = \frac{(\lambda_S + \lambda_L K^2) Y}{2(1 - \lambda_S B_0)} + B_0$$

及长包的总时延:

$$E'(S_2) = \frac{\lambda_S E'(B_1) E'(R_1) + \lambda_L E'(B_2) E'(R_2)}{[1 - \lambda_S E'(B_1)][1 - \lambda_S E'(B_1) - \lambda_L E'(B_2)]} + E'(B_2) = \frac{(\lambda_S + \lambda_L K^2) Y}{2(1 - \lambda_S B_0)(1 - \lambda_S B_0 - \lambda_L KB_0)} + KB_0$$

进而得到总时延:

$$Delay_{NPrSF} = \frac{\lambda_S E'(S_1) + \lambda_L E'(S_2)}{\lambda_S + \lambda_L} = \frac{\lambda_S}{\lambda_S + \lambda_L} \left[ \frac{(\lambda_S + \lambda_L K^2) Y}{2(1 - \lambda_S B_0)} + B_0 \right] + \frac{\lambda_L}{\lambda_S + \lambda_L} \left[ \frac{(\lambda_S + \lambda_L K^2) Y}{2(1 - \lambda_S B_0)(1 - \lambda_S B_0 - \lambda_L KB_0)} + KB_0 \right] \quad (5)$$

第2步 求 NPrLF 策略的总时延。令长包优先级为1,短包优先级为2,服务时间的期望分别为:  $E''(B_1) = KB_0$ ,  $E''(B_2) = B_0$ ,二阶矩分别为:  $E''(B_1^2) = K^2 Y$ ,  $E''(B_2^2) = Y$ ,并有  $E''(R_1) = KE''(R_2) = KY / (B_0^2)$ 。采用类似的方法得到总时延:

$$Delay_{NPrLF} = \frac{\lambda_L E''(S_1) + \lambda_S E''(S_2)}{\lambda_S + \lambda_L} = \frac{\lambda_L}{\lambda_S + \lambda_L} \left[ \frac{(\lambda_S + \lambda_L K^2) Y}{2(1 - \lambda_L KB_0)} + KB_0 \right] + \frac{\lambda_S}{\lambda_S + \lambda_L} \left[ \frac{(\lambda_S + \lambda_L K^2) Y}{2(1 - \lambda_L KB_0)(1 - \lambda_S B_0 - \lambda_L KB_0)} + B_0 \right] \quad (6)$$

第3步 比较两种策略的总时延的差异。

$$Delay_{NPrSF} - Delay_{NPrLF} = \frac{(\lambda_S + \lambda_L K^2)(K - 1)\lambda_S \lambda_L B_0 Y (\lambda_S B_0 + \lambda_L KB_0 - 2)}{2(\lambda_S + \lambda_L)(1 - \lambda_S B_0)(1 - \lambda_L KB_0)(1 - \lambda_S B_0 - \lambda_L KB_0)}$$

在保证足够服务能力的条件下,总流量强度小于1,有  $\lambda_S B_0 + \lambda_L KB_0 < 1 < 2$ ,所以上式小于0。证毕。

定理2 非抢占式长包优先策略的总时延小于抢占式长包优先策略的总时延,非抢占式短包优先策略的总时延小于抢占式短包优先策略的总时延,即:

$$Delay_{NPrLF} < Delay_{PrmLF} \quad (7)$$

$$Delay_{NPrSF} < Delay_{PrmSF} \quad (8)$$

证明:依式(2)、式(4)求出 PrmSF 策略的总时延:

$$Delay_{PrmSF} = \frac{\lambda_S}{\lambda_S + \lambda_L} \left[ \frac{(\lambda_S + \lambda_L K^2)Y}{2(1 - \lambda_S B_0)} + \frac{B_0}{1 - \lambda_S B_0} \right] + \frac{\lambda_L}{\lambda_S + \lambda_L} \left[ \frac{(\lambda_S + \lambda_L K^2)Y}{2(1 - \lambda_S B_0)(1 - \lambda_S B_0 - \lambda_L K B_0)} + \frac{K B_0}{1 - \lambda_S B_0 - \lambda_L K B_0} \right] \quad (9)$$

将式(9)与式(5)相减得:

$$Delay_{NPrSF} - Delay_{PrmSF} = \frac{-1}{\lambda_L + \lambda_S} \left[ \frac{(\lambda_S B_0)^2}{1 - \lambda_S B_0} + \frac{\lambda_L K B_0 (\lambda_S B_0 + \lambda_L K B_0)}{1 - \lambda_S B_0 - \lambda_L K B_0} \right] < 0$$

同理求出 PrmLF 的总时延:

$$Delay_{PrmLF} = \frac{\lambda_L}{\lambda_S + \lambda_L} \left[ \frac{(\lambda_S + \lambda_L K^2)Y}{2(1 - \lambda_L K B_0)} + \frac{K B_0}{1 - \lambda_L K B_0} \right] + \frac{\lambda_S}{\lambda_S + \lambda_L} \left[ \frac{(\lambda_S + \lambda_L K^2)Y}{2(1 - \lambda_L K B_0)(1 - \lambda_S B_0 - \lambda_L K B_0)} + \frac{B_0}{1 - \lambda_S B_0 - \lambda_L K B_0} \right] \quad (10)$$

将式(10)与式(6)相减得:

$$Delay_{NPrLF} - Delay_{PrmLF} = \frac{-1}{\lambda_L + \lambda_S} \left[ \frac{(\lambda_L K B_0)^2}{1 - \lambda_L K B_0} + \frac{\lambda_S B_0 (\lambda_S B_0 + \lambda_L K B_0)}{1 - \lambda_S B_0 - \lambda_L K B_0} \right] < 0$$

**定理 3** 非抢占式短包优先策略相比不区分优先级的策略具有更小的时延,即

$$Delay_{NPrSF} < Delay_{NonPr} \quad (11)$$

证明:对 NonPr 策略,有总到达率  $\lambda = \lambda_S + \lambda_L$ , 总服务率  $\rho = \rho_S + \rho_L = \lambda_S B_0 + \lambda_L K B_0$ 。平均服务时间为  $E(B) = (\lambda_S B_0 + \lambda_L K B_0) / (\lambda_S + \lambda_L)$ , 平均服务时间二阶矩为  $E(B^2) = (\lambda_S Y + \lambda_L K^2 Y) / (\lambda_S + \lambda_L)$ , 由式(1)得残留服务时间为  $E(R) = (\lambda_S Y + \lambda_L K^2 Y) / (2B_0 (\lambda_S + \lambda_L K))$ 。故平均等待时间为:

$$E(W) = \frac{\rho E(R)}{1 - \rho} = \frac{(\lambda_S + \lambda_L K^2)Y}{2(1 - (\lambda_S + \lambda_L K)B_0)}$$

总时延为:

$$Delay_{NonPr} = E(W) + E(B) = \frac{(\lambda_S + \lambda_L K^2)Y}{2(1 - (\lambda_S + \lambda_L K)B_0)} + \frac{\lambda_S B_0 + \lambda_L K B_0}{\lambda_S + \lambda_L} \quad (12)$$

将式(5)与式(12)相减得:

$$Delay_{NPrSF} - Delay_{NonPr} = \frac{\lambda_S \lambda_L B (\lambda_S + \lambda_L K^2) Y (1 - K)}{2(\lambda_S + \lambda_L)(1 - \lambda_S B_0)(1 - (\lambda_S + \lambda_L K)B_0)} < 0$$

综合以上得知,相比其它 4 种策略, NPrSF 在理论上具有最小的总时延。

## 4 仿真与分析

### 4.1 仿真实验方案

通过 NS-2 平台下的仿真实验,比较以上 5 种策略的性能,这 5 种策略都在 CMU-Priqueue 调度算法中实现。仿真场景由图 1 的 4 个移动节点构成,每个节点的速度为 15m/s。节点 N1 配置 3 个无线接口,向 N0 转发来自 N3 的  $App_L$  消息和来自 N2 的  $App_S$  消息,此外,与 N0 之间也有安全警告消息  $App_0$ ,速率为  $rate_0 = 200 \text{ Byte/s}$ ,各业务数据包长度为  $Len_0 = 200 \text{ Byte}$ ,  $Len_S = 200 \text{ Byte}$ ,  $Len_L = 800 \text{ Byte}$ 。  $App_0$  为固定比特率(CBR)类型的时延敏感业务,而  $App_L$  和  $App_S$  为泊松特性的非时延敏感业务。

N3-N1、N2-N1 和 N1-N0 之间的链路带宽分别为 10Mb/s、10Mb/s 和 2Mb/s, N2、N3 和 N1 中的队列长度分别为 50000 包、50000 包和 5000 包,故在表 1 的饱和流量设置方案中,链路 N3-N1 和 N2-N1 具有足够的带宽和处理能力,而链路 N1-N0 的带宽和处理能力可能不足。不考虑信道衰落等因素的影响,可假定所有从 N3 和 N2 发出的数据包都能达到 N1。

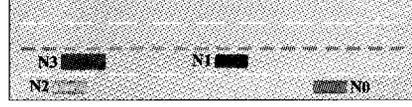


图 1 仿真场景

仿真持续 75s,在前 50s 的数据发送时间内,3 种业务按预定速率持续发送数据,后 25s 用于发送队列中剩余的数据包。预定的总发送流量恒定为 10MByte,因为  $App_0$  流量很小,则有  $50 * (rate_S + rate_L) = 50 * (200 * \lambda_S + 800 * \lambda_L) \approx 10 * 10^6$ ,即  $\lambda_S + 4\lambda_L = 1000$ ,可能有多种组合。设定典型的流量场景如表 1 所列,从 S1 到 S9,发送数据包的总数持续增加。

表 1 非时延敏感业务长短包流量分配

场景	$rate_S$ (Kb/s)	$rate_L$ (Kb/s)	发送流量比	$\lambda_S$	$\lambda_L$
S1	160	1440	1:9	100	225
S2	320	1280	2:8	200	200
S3	480	1120	3:7	300	175
S4	640	960	4:6	400	150
S5	800	800	5:5	500	125
S6	960	640	6:4	600	100
S7	1120	480	7:3	700	75
S8	1280	320	8:2	800	50
S9	1440	160	9:1	900	25

所探讨的性能指标针对 N1 到 N0 的数据通信,包括:分组抵达率(PDR)、以字节为单位的吞吐量(Throughput)、平均端到端时延(Delay)、各优先级队列的等待长度(QueueLen)。

### 4.2 仿真实验结果与分析

#### 1) 对分组抵达率和吞吐量的分析

在仿真中,高优先级分组可很快从 N1 发出,在队列中靠前的低优先级分组可较快地发出,有一定的接收时延,而较晚抵达 N1 的低优先级分组有可能一抵达溢满队列就被丢弃,也可能呆在队列中直到 50s 数据发送停止后才有机会发出。对 PDR 及 Throughput 的仿真结果如图 2—图 4 所示,此外,仿真结果表明  $App_L$  的 PDR 为 100%。

从图 2 可以看出,长包优先策略(PrmLF、NPrLF)的第 2 优先级业务(记为 Pri2)的 PDR 一直接近 100%,其原因在于长包的总数相对较少,预设的队列长度能保证基本不丢包;在 S6 时,短包优先策略(PrmSF、NPrSF)的 PDR 明显下降,其原因为随着  $\lambda_S$  增大,短包数目猛增,其等待队列溢满,产生丢包现象,PDR 和吞吐量受制于最大队列长度。

从图 3 可以看出,对于 NPrSF 策略,其第 3 优先级(记为 Pri3)业务也能得到有效的传输,并在 S5 时存在一个 PDR 极小值点,从 S1 到 S5, Pri2 的短包增多,影响了 Pri3 的长包的传送,导致后者的 PDR 逐渐下降,到 S5 时 PDR 又开始上升,其原因在于  $\lambda_L$  降低,长包数目减少,对应的丢包现象也减少,此队列中来不及发送的包可以保存在队列中,并在 50s 数据发送停止后继续传送;对于 NPrLF 策略,其 Pri3 的短包的 PDR 变化趋势与 NPrSF 相似,在 S4 时存在一个极小值点,但

由于 Pri2 的长包发送时延较大,且短包的数据量较大,因此短包的延误和丢包现象严重些,PDR 都要小些。在 PrmSF 和 PrmLF 策略中,Pri3 业务的 PDR 接近 0,基本被剥夺了传送机会,原因在于抢占式策略导致正在发送的低优先级数据包停止传送,从此数据包开始接收服务到服务被抢占的这一段期间就被浪费,单位时间内成功传输的数据包就很少。

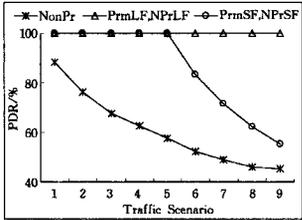


图 2 第 2 优先级业务的分组  
抵达率

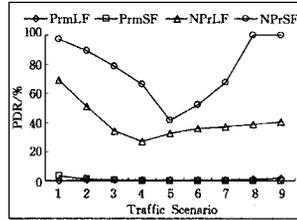


图 3 第 3 优先级业务的分组  
抵达率

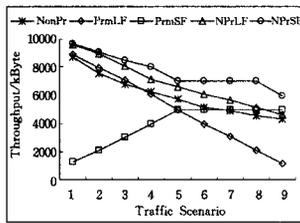


图 4 全部业务的吞吐量

综合 PDR 和数据发送量得到图 4,可以看出 NPrSF 策略能提供最高的吞吐量,总体效能最佳,NPrSF 和 NPrLF 比传统的 FCFS 策略(NonPr)有明显的优势,且吞吐量从 S1 到 S9 逐渐下降,原因在于数据包数目增加,超过了队列容纳能力,导致明显的丢包。PrmSF 和 PrmLF 的流量基本由 Pri2 业务组成,没有 Pri3 业务补充,故其吞吐量曲线变化与图 4 对应的 PDR 规律一致,吞吐量相对小很多。

**结束语** 本文证明了非抢占式短包优先调度策略具有最好的整体性能。将业务类型分为时延敏感和非时延敏感两类,分别采用抢占式策略和非抢占式短包优先策略处理调度冲突,提出了多业务优先级调度模型。通过 NS-2 环境下的饱和和流量仿真实验,证明非抢占式短包优先策略能避免对低优先级业务的抑制情况,提供各业务相对均衡、整体最优的传输效果,可以作为传统 FCFS 策略的取代。

通过对数据包排队长度的分析,可以清楚地体现丢包和阻塞的情况,并提出优化改进的策略。针对数据量增大、排队现象加重、易造成时延增加、PDR 和吞吐量下降的现象,值得进一步研究的内容包括:队列溢满时间预测与动态队列缓冲区分配方案的研究,从而实现队列缓冲区的协调分配,避免运

行中随时可能出现的队列溢满现象。

## 参 考 文 献

- [1] Walraevens J, Steyaert B, Moeneclaey M, et al. A discrete-time HOL priority queue with multiple traffic classes[C]// Proceedings of ICN 2005. Lecture Note in Computer Science 3420. Sydney, Australia, 2005; 620-627
- [2] Kim K, Chae K C. Discrete-time queues with discretionary priorities[J]. European Journal of Operational Research, 2010, 200(2): 473-485
- [3] Zaborovsky V, Zayats O, Mulukha V. Priority queueing with finite buffer size and randomized push-out mechanism[C]// Proceedings of 9th International Conference on Networks. Mennuaires, France, 2010; 316-320
- [4] Awan I, Fretwell R. Analysis of discrete-time queues with space and service priorities for arbitrary arrival processes[C]// Proceedings of 11th International Conference on Parallel and Distributed Systems. Fukuoka, Japan, 2005; 115-119
- [5] Walraevens J, Fiems D, Bruneel H. Performance analysis of priority queueing systems in discrete time[C]// Next Generation Internet. Lecture Notes in Computer Science 5233, 2011; 203-232
- [6] Bouassida M S, Shawky M. A cooperative congestion control approach within VANETs; Formal verification and performance evaluation[J]. EURASIP Journal on Wireless Communications and Networking, 2010, 2010; 712525
- [7] Darus M Y B, Bakar K A. Congestion control framework for disseminating safety messages in Vehicular Ad-Hoc Networks (VANETs)[J]. International Journal of Digital Content Technology and its Applications, 2011, 5(2): 173-180
- [8] Chen Geng-xin, Liu Gui-xiong, Huang Guo-jian. Real-time performance analysis and improvement of NCAP communication mode based on the queuing model with non-preemptive multi-priority [J]. Chinese Journal of Scientific Instrument, 2010, 31(8. supp): 321-325
- [9] 于秦, 毛玉明. 基于 PME 重尾分布服务时间的 M/G/1 模型排队性能研究[J]. 计算机学报, 2005, 28(12): 2103-2108
- [10] Liu Huan-lin, Chen Qian-bin, Pan Ying-jun. Difference length scheduling for asynchronous optical packet switching[C]// Proceedings of Conferences on Optical Transmission, Switching, and Subsystems IV. SPIE 6353, 2006; 63532S
- [11] Li Wen-jie, Liu Bin. Preemptive short-packet-first scheduling in input queueing switches[J]. Acta Electronica Sinica, 2005, 33(4): 577-583
- [12] Adan I, Resing A. Queueing theory[D]. Eindhoven University of Technology, the Netherlands, 2001

(上接第 146 页)

- [2] Corp N. Fermi Whitepaper [EB/OL]. [http://www.nvidia.com/content/PDF/fermi\\_white\\_papers/NVIDIA\\_Fermi\\_Compute\\_Architecture\\_Whitepaper.pdf](http://www.nvidia.com/content/PDF/fermi_white_papers/NVIDIA_Fermi_Compute_Architecture_Whitepaper.pdf), 2010-01-23
- [3] Bellman R E. Dynamic Programming[M]. Princeton University Press, Princeton, N. J., 1957
- [4] Boncz P A, Kersten M L. MIL primitive for querying a fragmented world[J]. The VLDB Journal, 1999, 8(2): 101-119
- [5] Boncz P A, Zukowski M, Nes N. MonetDB/X100; Hyper-pipelining query execution[C]// Böhm K, eds. Proceedings of International Conference on Very Large Data Bases (VLDB) 2005.

Asilomar, CA, USA; Very Large Data Base Endowment, 2005; 255-237

- [6] Harizopoulos S, Shkapenyuk V. QPipe: A Simultaneously Pipelined Relational Query Engine[C]// Özcan F, eds. Proceedings of the ACM SIGMOD International Conference on Management of Data. Baltimore, Maryland, USA; ACM, 2005; 14-16
- [7] Arumugam S, Dobra A, Jermaine C M. The DataPath System; A Data-Centric Analytic Processing Engine for Large Data Warehouses[C]// Elmagarmid A K, eds. Proceedings of the ACM SIGMOD International Conference on Management of Data. Indianapolis, Indiana, USA; ACM, 2010; 6-11
- [8] TPC-H home[OL]. <http://www.tpc.org/tpch/>