

基于轮询方式综合业务总线 LAN 的优先级控制

王超¹, 王立宏², 逯昭义¹

WANG Chao¹, WANG Li-hong², LU Zhao-yi¹

1. 青岛大学 信息工程学院, 山东 青岛 266071

2. 烟台大学 计算机学院, 山东 烟台 264005

1. Information Engineering College, Qingdao University, Qingdao, Shandong 266071, China

2. Computer College, Yantai University, Yantai, Shandong 264005, China

E-mail: chao19821982@163.com

WANG Chao, WANG Li-hong, LU Zhao-yi. Priority control in integrated service bus LAN based on polling model. Computer Engineering and Applications, 2007, 43(36): 132-136.

Abstract: In order to make sure of the QoS (Quality of Service) for the integrated service, such as voice video data, and so on, in the computer networks, it is necessary to solve the problems of priority control. So, under the condition that giving top priority to sending frames with the highest priority and ensuring the sending of lower priority frames (especially the lowest priority frames), "The Improved Polling Model" is first introduced. So that, it expands the use of the polling access model in the integrated service Bus LAN. Simulation results show that the improved protocol increases the QoS for the integrated service Bus LAN.

Key words: polling; priority; integrated service; transmitting polling

摘要: 为了保证语音、视频、数据等综合业务在计算机网络通信中的服务质量 QoS, 需要解决优先级控制问题。为此, 在保证高优先级帧尽可能优先发送, 同时适当兼顾低优先级, 特别是最低优先级帧发送的前提下, 首次提出了“改进型轮询方式”, 从而扩展了轮询介质访问方式在综合业务总线 LAN 中的应用。模拟实验表明, 改进后的协议提高了综合业务总线 LAN 的 QoS。

关键词: 轮询; 优先级; 综合业务; 传递轮询

文章编号: 1002-8331(2007)36-0132-05 **文献标识码:** A **中图分类号:** TP393

1 引言

在综合业务 LAN 中, 优先级问题是一个热门的研究课题^[1-3]。文献[1]是基于访问方式 CSMA/CD 的, 文献[2]是基于令牌总线的, 文献[3]是基于令牌环的。本文探讨基于轮询方式总线型 LAN 的综合业务优先级控制问题。在综合业务网中, 不同的业务对实时性的要求有所不同, 要求网络提供不同质量的服务, 对于那些对实时性要求比较高的业务, 需要网络优先为其提供服务。轮询访问方式是一种比较成熟的介质访问方式, 近年来, 利用轮询方式解决无线 LAN 综合业务优先级问题的研究报告开始增多^[4-5], 文献[4]提出的 QAP(a QoS supportive adaptive polling, 一种支持服务质量的适应性轮询)协议, 实际上是在普通轮询机制的基础上加入了概率算法, 对优先级高的工作站点, 赋予高概率, 访问控制点(AP)依据概率值对无线 LAN 内的各工作站进行轮询访问, 对那些概率值高的站点, 则更有可能被服务。文献[5]利用 UPMA(User-dependent Perfect-scheduling Multiple Access, 根据用户数目妥善安排分组传输的多址接入)协议, 解决了语音、数据两级业务的优先级问题, 网上的站点被分为语音站点和数据站点, 它们各自分别属于实时轮询队列和非实时轮询队列, AP 优先对实时轮询队列内的站点轮询且进行服务, 然后再对非实时轮询队列内的站点轮询且进行

服务, 从而保证不同业务的服务质量。但是, 利用轮询方式来解决总线 LAN 综合业务优先级问题的研究报告比较少, 有关文献对基于轮询方式的总线树型拓扑结构虽有阐述^[6], 但未涉及综合业务。为此, 本文对综合业务总线 LAN 的轮询控制方式进行了深化研究, 提出了一种基于轮询方式总线 LAN 的综合业务优先级控制协议——改进型轮询方式。

2 “改进型轮询方式”的提出

2.1 传递轮询概述

轮询是人们比较熟知的访问方式, 但为了探讨综合业务的优先级机制, 仍需对该访问方式略加阐述。轮询是网上主机按一定顺序逐个询问各工作站有无信息发送, 如有信息发送, 则被询问的工作站立即将信息发给主机; 如无, 则接着询问下一站。轮询分为轮叫轮询和传递轮询。轮叫轮询, 是主机从站 1 开始逐个轮询, 站 1 如有信息, 立即发给主机, 站 1 如无信息, 则向主机发送一种控制帧, 表明无信息可发。随后, 主机询问站 2, 依次类推。从站 1 到站 N 循环访问, 循环发送信息; 传递轮询, 也称集中查询, 主机先向站 N 发出轮询帧, 站 N 在发送的信息帧(如果当时有信息的话)或者控制帧(如果当时无信息的话)之后, 将其相邻站 N-1 的地址附上, 由此站 N-1 得到查询,

依次类推, 能实现一个循环周期的轮询。在下一个循环开始时, 主机再向站 N 发出轮询。传递轮询实为轮询信号传递。如果主机在一次循环轮询完成, 收回轮询信号后, 无停留地把轮询帧发给站 N , 传递轮询与 IEEE802.5 的令牌环协议非常相似。

通过上述轮询协议的运行机理可以看出, 基本的轮询访问方式对不同类型的业务没有区分优先级别。当轮到某一工作站时, 只要有帧即可发送。因此, 不能为不同类型的业务提供不同质量的服务, 不适合综合业务的传输。显然, 将轮询访问方式用在综合业务总线 LAN 上, 需要对原协议进行改进。

2.2 改进型轮询的提出

设网络中存在三种业务——语音、视频、数据, 它们分别对应三种优先级 0 级、1 级、2 级, 优先级依次降低。每个工作站有一个缓冲区, 设定缓冲区的容量为 1, 当到达缓冲区的一帧信息未被服务之前, 本站不再产生新帧, 新产生帧的优先级由工作站自行确定。新协议采用传递轮询方式, 按照站号从高到低的顺序进行轮询。新协议将原来的“轮询”改称为“询问循环”。设定一次“询问循环”包括两轮轮询。如果将一次“询问循环”称为“大轮询”, 一轮轮询称为“小轮询”, 则“大轮询”包括两轮“小轮询”。小轮询时, 对所有工作站只服务同一种优先级别的帧。在一次大轮询中, 第一轮小轮询只让有 0 级帧的工作站发送, 那些有 1 级、2 级帧的工作站在接到轮询帧后并不发送帧, 而是立即把轮询帧传递给下一相邻站; 第二轮小轮询只让有 1 级帧(也可能是 2 级帧)的工作站发送, 究竟是只让有 1 级帧的站发送还是只让有 2 级帧的站发送, 第二轮小轮询开始时由主机随机确定, 按协议主机选取 1 级的概率大于 2 级。当然也可以在第一轮小轮询中让有 1、2 优先级帧的工作站预约, 在按概率确定第二轮小轮询中服务优先级级别时, 控制站将预约的信息作为一个因素予以考虑, 这种情况将待研究取得进展后另行报告。

3 协议运行过程及平均最大滞留时间分析

3.1 参数设定

本文通过平均最大滞留时间来对原协议和改进协议进行性能评价。为了计算平均最大滞留时间, 必须先求出最大滞留时间, 考虑到每轮轮询从 N 号站开始, 因而计算信息帧的最大滞留时间, 只有设定 1 号站为观察站, 且每一轮轮询中 1 号站均有信息帧等待服务, 而且要求该信息帧是在本站前一个信息帧发送结束后, 立即产生的。为分析协议性能, 须首先对协议的运行过程进行阐述, 并设定各类参数如下:

M : 网络开始运行时设定的最大轮询圈数;

m : 当前轮询圈数;

t_m : 第 m 轮的轮询帧经过观察站时, 观察站信息帧的滞留时间。按协议该信息帧是上次发送后立即产生的;

τ : 信息帧时宽;

t_s : 轮询帧在相邻工作站之间的步行时间;

n : 工作站总数;

P : 第二轮小轮询主机选择只对 1 级(或 2 级)帧进行服务的概率, 当选择 1 级时, $P=P_1$; 当选择 2 级时, $P=P_2$ 。 $P_1=1-P_2$;

P' : 轮询帧到达某一工作站时, 工作站有信息帧的概率;

c : 在原协议中, 观察站发送信息帧的总量;

c_{j1} : 在改进协议中, 当观察站前一次发送 1 级(或 2 级)帧, 当前产生的是 j 级帧, 观察站发送 j 级帧的总量, $j=0, 1, 2$;

c_{j2} : 在改进协议中, 当观察站前一次发送 0 级帧, 当前产生的是 j 级帧, 观察站发送 j 级帧的总量, $j=0, 1, 2$;

t_u : 在原协议中, 网络开始运行后, 观察站第 u 次产生的信息帧的

最大滞留时间, $u=1, 2, \dots, c$;

t_{w1} : 在改进协议中, 当观察站前一次发送 1 级(或 2 级)帧, 当前产生的是 j 级帧, 观察站第 w 次产生的 j 级帧的最大滞留时间, $w=1, 2, \dots, c_{j1}$;

t_{v2} : 在改进协议中, 当观察站前一次发送 0 级帧, 当前产生的是 j 级帧, 观察站第 v 次产生的 j 级帧的最大滞留时间, $v=1, 2, \dots, c_{j2}$;

$T_{原}$: 在原协议中, 信息帧的平均最大滞留时间;

T_{j1} : 在改进协议中, 当观察站前一次发送 1 级(或 2 级)帧, 当前产生的是 j 级帧, j 级帧的平均最大滞留时间, $j=0, 1, 2$;

T_{j2} : 在改进协议中, 当观察站前一次发送 0 级帧, 当前产生的是 j 级帧, j 级帧的平均最大滞留时间, $j=0, 1, 2$;

T_j : 在改进协议中, j 级帧的平均最大滞留时间。

3.2 原协议与改进协议 0 级帧的发送过程及平均最大滞留时间分析

对于原协议, 信息帧最大滞留时间为一轮轮询时间, 平均最大滞留时间为多轮轮询时间的平均值(已设定进行 M 次轮询)。发送过程如图 1 所示。由图 1 很容易看出, 原协议的平均最大滞留时间 $T_{原}$ 为:

$$T_{原} = \frac{\sum_{u=1}^c t_u}{c} \quad (1)$$

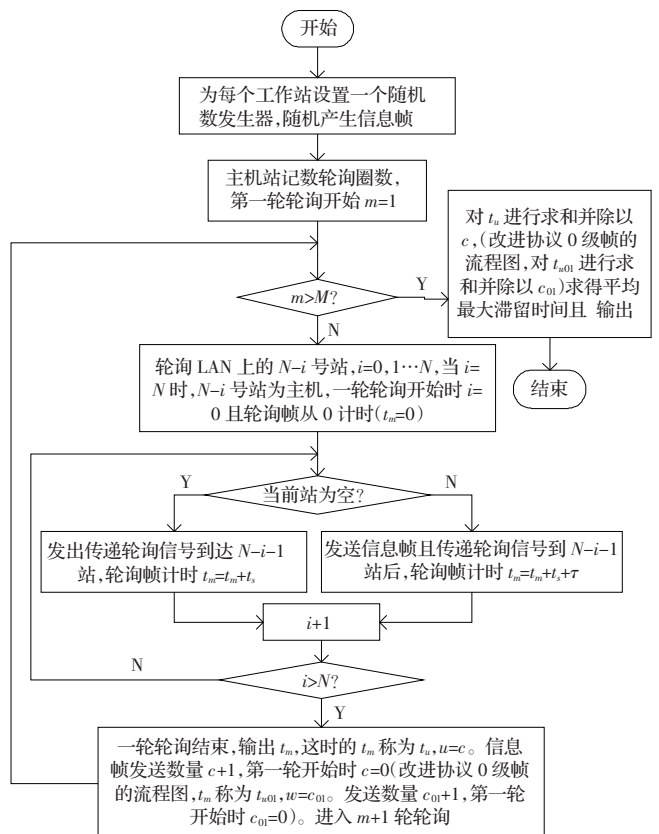


图 1 原协议(或改进协议 0 级帧, 前一次发送 1 级或 2 级帧)流程图

关于改进协议 0 级帧的发送过程与原协议基本一致, 只不过, 由于改进协议将“一次询问循环”改成“两轮小轮询”, 0 级帧的最大滞留时间与观察站(已设定为 1 号站)前一次发送的信息帧类型有关。若前一次发送的为 1 级或 2 级帧, 则当前产生的 0 级帧只要经过一轮小轮询(“第二轮小轮询”)就能得到发送权, 这种情况下的平均最大滞留时间为 T_{j1} , $j=0$; 若前一次发送的为 0 级帧, 则当前产生的 0 级帧需要经过两轮小轮询

(先经过“第二轮小轮询”,再经过“第一轮小轮询”)才能得到发送权,这种情况下的平均最大滞留时间为 $T_{j2}, j=0$;因此,改进协议 0 级帧有两种最大滞留时间 T_{01} 和 T_{02} ,需要对二者再一次求平均值,才能得到平均最大滞留时间 T_0 。这一点在稍后的模拟实验时,给予了充分关注。对于观察站前一次发送 1 级或 2 级帧,当前帧是 0 级帧的情况,发送过程与原协议基本一致,只需将图 1 中的“当前站为空?”和“为每个工作站设置一个随机数发生器,随机产生信息帧”分别改为“当前站有 0 级帧?”和“为每个工作站设置一个随机数发生器,随机产生 0、1、2 级帧,再为主机设置一个随机数发生器,以便在第二轮小轮询中按不同概率确定 1 级或 2 级帧”即可。对于观察站前一次发送 0 级帧,当前帧也是 0 级帧的情况,其发送过程如图 2 所示。参考图 1、2 可得改进协议 0 级帧的平均最大滞留时间 T_0 为:

$$T_{01} = \frac{\sum_{m=1}^{c_{01}} t_{m01}}{c_{01}} \quad T_{02} = \frac{\sum_{v=1}^{c_{02}} t_{v02}}{c_{02}} \quad T_0 = \frac{T_{01} + T_{02}}{2} \quad (2)$$

3.3 改进协议 1、2 级帧的发送过程及平均最大滞留时间分析

上述 0 级帧经过一轮或两轮小轮询必定能够得到发送权,1、2 级帧则不同。1、2 级帧由于受“一次大轮询”中“第二轮小轮询”选择概率 P 的影响,最大滞留时间具有不确定性,并且最大滞留时间也与观察站(已设定为 1 号站)前一次发送的信息帧类型有关。就 1 级帧而言,若前一次发送的为 1 级或 2 级帧,则

当前产生的 1 级帧需经过偶数轮(至少两轮)小轮询才能得到发送权,这时的平均最大滞留时间为 $T_{j1}, j=1$;若前一次发送的为 0 级帧,则当前产生的 1 级帧需经过奇数轮(至少一轮)小轮询才能得到发送权,这时的平均最大滞留时间为 $T_{j2}, j=1$;因此,改进协议 1 级帧也有两种最大滞留时间 T_{11} 和 T_{12} ,需要对二者再一次求平均值,才能得到平均最大滞留时间 T_1 ,这一点在稍后的模拟实验时,同样给予了关注。1 级帧的发送过程分别如图 3、4 所示,由此可得到改进协议 1 级帧的平均最大滞留时间 T_1 为:

$$T_{11} = \frac{\sum_{m=1}^{c_{11}} t_{m11}}{c_{11}} \quad T_{12} = \frac{\sum_{v=1}^{c_{12}} t_{v12}}{c_{12}} \quad T_1 = \frac{T_{11} + T_{12}}{2} \quad (3)$$

2 级帧的发送过程与 1 级帧基本一致,只是必须将图 3 中的①、②号虚线框进行位置上的交换。同样,要将图 4 中②、③号虚线框以及两者之间的虚连线与①号虚线框进行位置上的交换。参考图 3、4 很容易得出改进协议 2 级帧的平均最大滞留时间 T_2 为:

$$T_{21} = \frac{\sum_{m=1}^{c_{21}} t_{m21}}{c_{21}} \quad T_{22} = \frac{\sum_{v=1}^{c_{22}} t_{v22}}{c_{22}} \quad T_2 = \frac{T_{21} + T_{22}}{2} \quad (4)$$

4 模拟实验与结果分析

设定实验在比较理想的状态下进行,无帧的错传、丢失现象,工作站等概率产生各优先级帧,信息帧的长度为定值,相邻

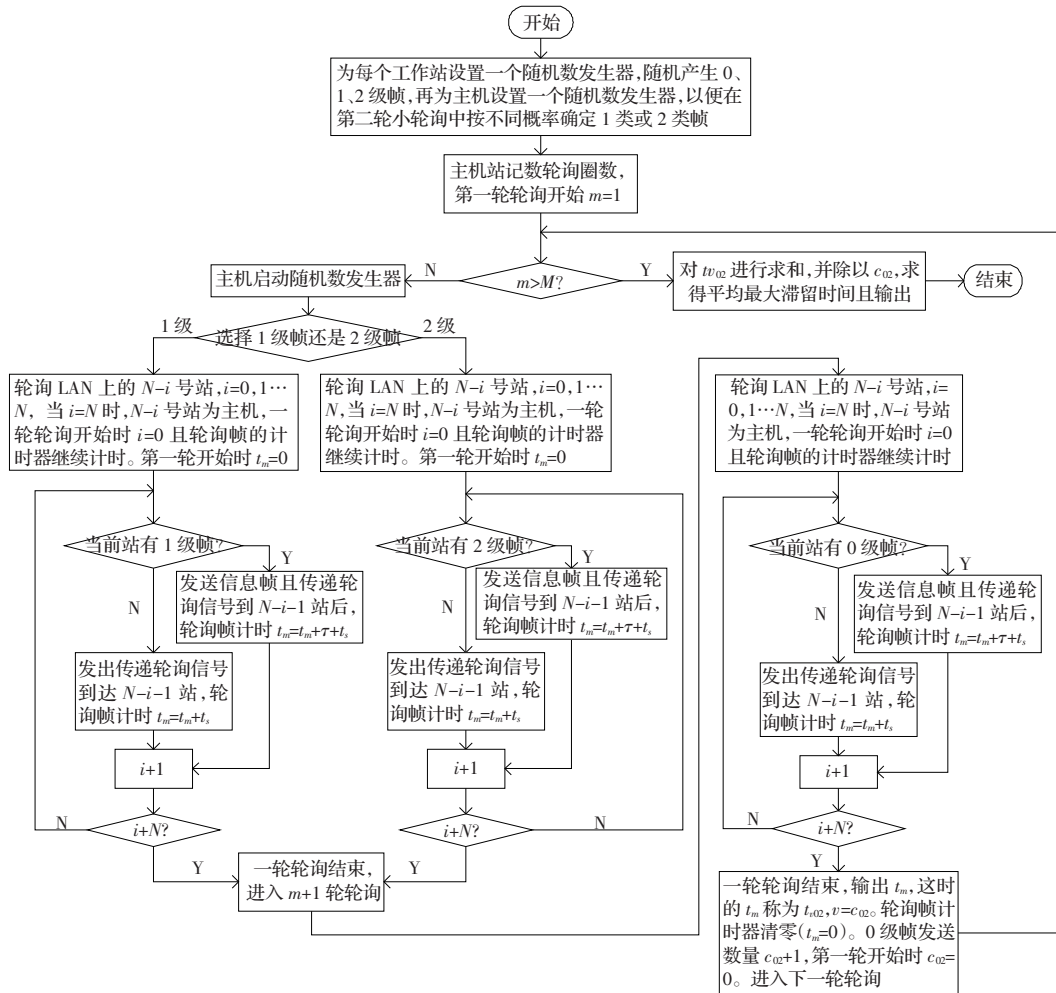


图 2 改进协议 0 级帧流程图(前一次发送 0 级帧)

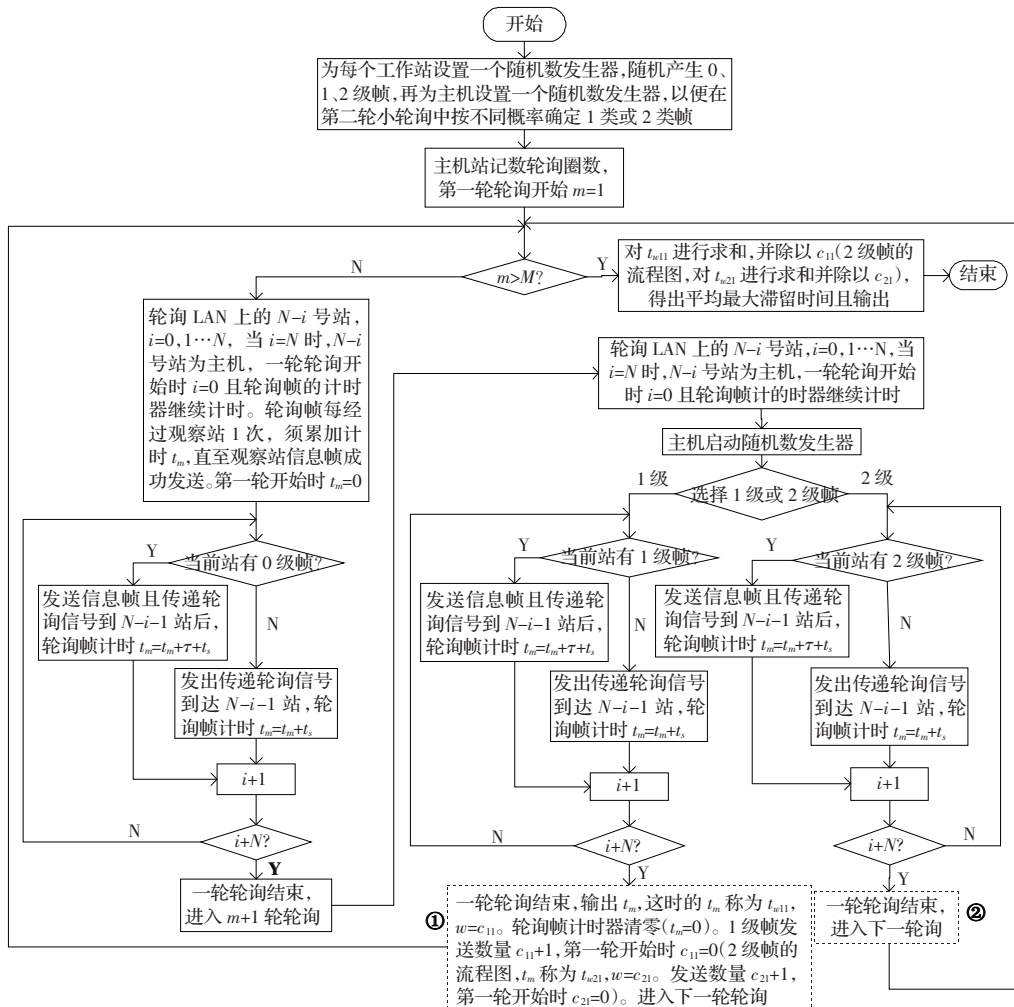


图 3 改进协议 1 级(或 2 级)帧流程图(前一次发送 1 级或 2 级帧)

工作站之间的距离为定长。实验时,令 $M=10\ 000, t_s=1, P'=0.5$ 。按照上述流程,用 C 语言编程,程序约 900 行,有兴趣的读者可以索取,限于篇幅不再赘述。

4.1 信息帧长度 τ 对平均最大滞留时间的影响

为探讨“改进型轮询协议”的性能,需要对信息帧长度问题进行分析。实验已设定信息帧的长度为定值,在此前提下,通过比较改进协议 0 级帧与原协议的平均最大滞留时间以便确定适合改进协议的信息帧长度。设定 $n=50, P_1=0.7$ 。根据流程图 1 和 2,得到实验数据如表 1 所示,实验曲线如图 5 所示。

由图 5 可见,一般情况下($\tau \geq 4$),只要信息帧长度相同,改进协议 0 级帧的性能明显优于原协议,这一点是容易理解的:在原协议的一次循环中,网上凡有信息帧的站都要服务;然而在改进协议的一次大循环(两轮小循环)中,只对 0 级帧和 1 级帧(或 2 级帧)进行服务,对 2 级帧(或 1 级帧)不服务,省出时间在下一大循环的第一轮小轮询中又对 0 级帧服务,因此 0 级帧的滞留时间明显变小。显然,改进型协议适合优先级条件下信息帧的传输。

4.2 站点数 n 对平均最大滞留时间的影响

设定 $\tau=5, P_1=0.7$,改变站点数,依据上述流程图 1、2、3、4 进行实验,对原协议和改进后协议求平均最大滞留时间,得到实验数据如表 2 所示,实验曲线如图 6 所示。

由图 6 可见,改进后协议 0、1、2 优先级帧的滞留时间依次

变大,且改进后协议最高优先级(0 级)帧的滞留时间比原协议要短,较低优先级(1 级)帧的滞留时间适当变长,而最低优先级(2 级)帧的滞留时间并非趋于很大。它表明改进协议对优先级的控制机制是有效的,即改进协议既提高了最高优先级业务的服务质量,同时适当兼顾了低优先级业务特别是最低优先级业务的服务质量,达到了改进协议的预期目的。

4.3 选择概率 P_1 对平均最大滞留时间的影响

设定 $\tau=5, n=20$ 。在流程图 1、2、3、4 的基础上再一次进行实验,以探讨由主机确定的 P_1 值对 0、1、2 优先级帧平均最大滞留时间的影响。实验数据如表 3 所示,实验曲线如图 7 所示。

由图 7 可见,按协议 T_0 几乎不变是容易理解的,而 T_1+T_2 的值随着 P_1 的增加有较大变化,则是由于随着 P_1 增加, T_1 趋于 T_0, T_2 趋于无穷, T_1 的变化趋势平缓,而 T_2 的变化趋势陡急。因此,在不同 P_1 值下, T_1+T_2 并非定值。这说明在网络实际运行中,不能仅为提高较低优先级(1 级)业务的服务质量而无限增加 P_1 值,这样会使得最低优先级(2 级)业务长时间滞留系统甚至永远不能得到服务。(收稿日期:2007 年 7 月)

参考文献:

- [1] 白成林.一种基于 CSMA/CD 的优先级随机分割的综合业务局域网[J].小型微型计算机系统,2004,24(11):1916-1918.
- [2] 吕磊,逯昭义,赵东风.综合业务网中 IEEE802.4 优先级控制准公平性探讨[J].计算机工程与应用,2004,40(7):146-151.

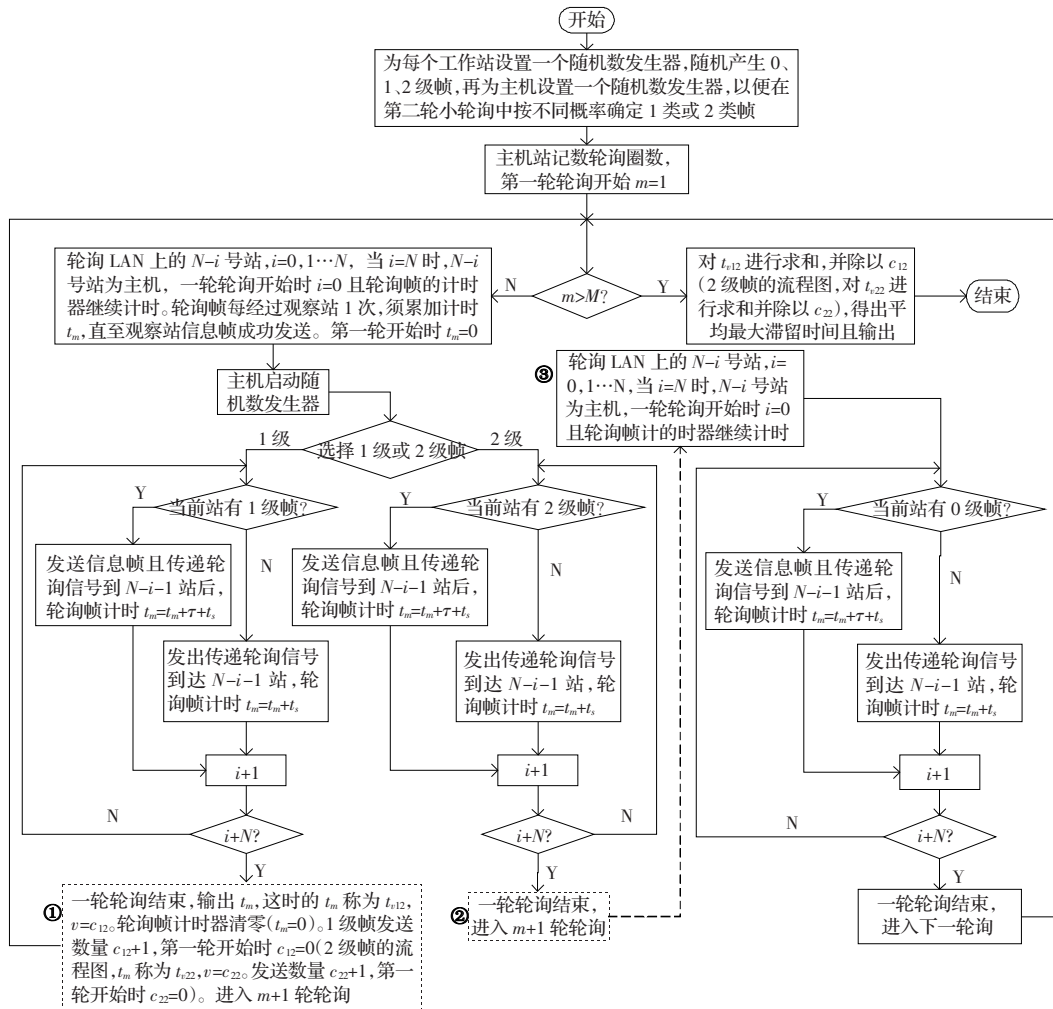


图4 改进协议1级(或2级)帧流程图(前一次发送0级帧)

表1 不同信息帧长度0级帧的平均最大滞留时间

τ	4	5	6	7	8
$T_{原}$	152.20	177.75	203.30	228.85	254.40
T_{01}	90.10	100.12	110.14	120.17	130.19
T_{02}	178.18	197.73	217.27	236.82	256.36
T_0	134.14	148.93	163.71	178.50	193.28

表2 不同 n 值下, 协议改进前后平均最大滞留时间比较

τ	10	20	30	40	50	60
$T_{原}$	37.58	72.54	107.64	142.65	177.75	212.70
T_0	34.48	63.00	91.621	120.25	148.92	177.61
T_1	62.62	127.16	91.883	256.70	321.72	386.71
T_2	122.74	254.14	86.05	518.64	650.47	782.80

表3 不同 p_1 下, 协议改进前后平均最大滞留时间比较

p_1	0.60	0.65	0.70	0.75	0.80	0.85	0.90
T_0	63.52	63.26	63.00	62.68	62.38	61.98	61.58
T_1	147.76	136.38	127.16	118.50	110.08	102.46	96.50
T_2	198.35	223.67	254.14	298.49	365.33	475.98	662.74

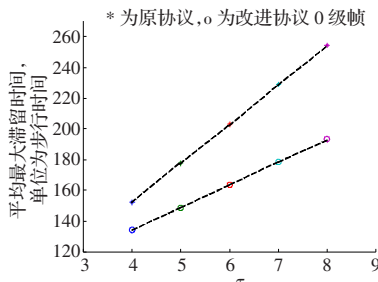


图5 不同信息帧长度 τ 对协议性能的影响

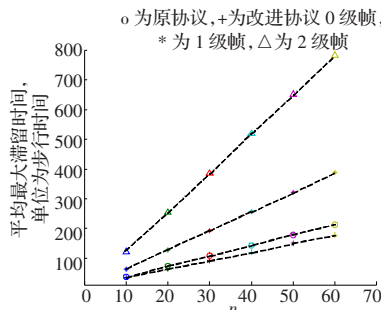


图6 不同 n 值对协议性能的影响

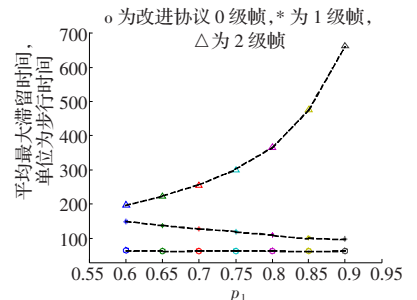


图7 不同 p_1 对协议性能的影响

[3] 逯昭义, 罗秀秀. 综合业务城域网 IEEE802.5 预约级协议改进策略[J]. 电子与信息学报, 2002, 24(11): 1581-1588.
 [4] Lagkas T D, Papadimitriou G I, Pomportsis A S. QAP: a QoS supportive adaptive polling protocol for wireless LANs[J]. Computer Communications, 2006, 29: 618-633.
 [5] 周亚建, 李建东, 吴杰. 支持语音/数据分组并传的 UPMA 多址接入

协议[J]. 电子学报, 2003, 31(8): 1222-1226.
 [6] 顾震, 鲍仁敏, 周明润. Electronic-dog 与链式动态树总线[J]. 计算机工程, 2006, 32(7): 243-252.
 [7] 逯昭义. 计算机网络体系结构-计算机网络原理[M]. 北京: 清华大学出版社, 北方交通大学出版社, 2003: 115-117.