

时间/事件混合触发通讯系统的研究

徐 亮, 张延园, 林 奕, 林 伟

(西北工业大学 计算机学院, 陕西 西安 710129)

摘 要: 时间敏感通讯系统要确保周期任务的实时性, 同时还要能快速响应突发事件. 为了在普通以太网上构建这种系统, 由此提出了一种主从式、时间和事件混合触发的任务调度方法, 使得各个通讯节点能按照调度表分时的占用网络带宽; 此外还提供了一个基于循环队列的事件触发消息缓存一发送模型以简化通讯节点实现. 实验结果表明在系统的时间粒度要求不高时, 利用上述方法构建时间敏感通讯系统具有可行性.

关键词: 实时通讯; 任务调度; 模型设计

中图分类号: TP391

文献标识码: A

文章编号: 1000-7180(2011)04-0068-05

Research of Communication Systems with Time-Triggered and Event-Triggered

XU Liang, ZHANG Yan-yuan, LIN Yi, LIN Wei

(School of Computer Science, Northwestern Polytechnical University, Xi'an 710129, China)

Abstract: Time-sensitive communication systems are supposed not only to ensure real-time performance of periodic tasks, but also to respond quickly to sporadic events. In order to construct such systems based on general Ethernet, this paper proposes a master-slave task scheduling method which mixes time-triggered and event-triggered, enabling different communication nodes occupy the bandwidth according to a schedule. Besides, it provides a store-send model based on round-robin queue for event-triggered messages, with which the implementation of communication nodes can be simplified. Experimental results show that it is feasible to use this method to construct time-sensitive communication systems with low time precision requirement.

Key words: real-time communication; task scheduling; model designing

1 引言

普通以太网具有低价格、高带宽、技术成熟的特点, 已经成为了构建通讯系统的首选, 但由于它采用 CSMA/CD 介质访问控制策略而使得链路存在通讯碰撞, 是典型的非确定性网络, 不能直接用来构建时间敏感通讯系统. 为了改造以太网以满足周期性实时通讯(例如分布并行实时系统^[1])的要求, 时间触发以太网(Time-Triggered Ethernet, TTE)^[2]为以太网底层协议提供一套扩展服务, 基于硬件机制来保证每个通讯节点按照配置表分时确定的占用网络带宽, 保证了通讯的实时性. 然而 TTE 网络的部署

需要专用的硬件设备, 例如 TTE 交换机、板卡、监控器等, 其代价非常高, 限制了其应用的范围. 针对有实时性要求但时间粒度要求不高、希望构建在普通以太网上的通讯应用, 本文提出了一种软件实现方法, 它能通过全局的通讯调度控制, 最大程度确保周期任务不超时, 同时减小系统响应异步事件的延迟时间.

在网络系统中, 时间触发(Time-Triggered, TT)是指通讯任务只有在预先决定的时间点上才能开始运行, 由此可知时间触发机制具有先天的确定性, 只要在网络系统运行前设计好了一套可行的任务调度方案, 这一系统所要求的实时性理论上是可

收稿日期: 2010-06-18; 修回日期: 2010-09-01

基金项目: 国家“八六三”计划重大项目(2009AA01A404)

证明的.然而,时间触发机制给网络系统带来的确定性是以牺牲系统的灵活性为代价的,因为时间触发采取的是静态的、集中的控制方式,在系统运行时,各个通讯节点的任务调度时机是固定的,这样的系统很难随机应变的响应异步事件.事件触发(event-triggered, ET)指系统中任务只有在与之相关的特定事件发生的条件下才开始运行,它适合于应对系统突发的状况,例如系统报警和入侵响应等.时间触发和事件触发各有优缺点而且正好互为补充,因而将两种机制结合起来(ET&TT),使系统在确定性和灵活性上取得平衡这一折衷方案被很多系统^[2-4]所采用.本文提出的调度控制正是这种折衷方案的产物.

2 系统模型

以太网的特点是各个节点竞争共享的传输介质,在发生冲突的时候采用二进制指数退避算法来降低再次冲突的概率.假如能够确保在系统运行的每个时刻,最多只有一个确定的通信节点发送数据,将极大的降低冲突发生的可能.为此我们给网络协议层次中增加一个通讯调度控制层,图1所示,让这个新增协议层来管理本节点的通讯任务,确保各个通讯节点在属于自己的时段内占用网络介质发送数据,避免无序的竞争导致的不可预测性.系统每个节点的通讯调度由一个主控节点全局统筹管理.我们以常见的总线型网络来说明,系统结构如图2所示,其运行机制如下:

① 部署系统时我们根据系统节点数、网络带宽、通讯任务的时间要求等条件,确定调度周期(记为 T)并设计出全系统的任务调度表.调度表存储在主控节点上,是调度信息编码后的二进制串,它的每一项包含一个调度周期 T 的全部信息.

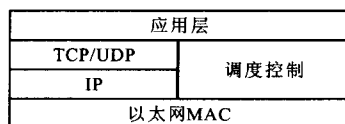


图1 调度控制子层

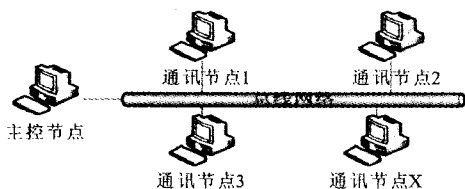


图2 通讯系统结构

② 系统运行时,主控节点读取一个调度表项的

二进制串,通过网络广播给系统所有节点.各个通讯节点接收到控制信息后,先清零本地计时器,以实现系统的时间同步,然后解析调度表,获取本节点在这个调度周期内占有网络的时机和时长.通讯节点将按调度表规定来发送通讯数据,交错确定的访问网络.

以后的每个调度周期内均重复上述过程②.

3 调度方案设计

为了说明方便,我们可认为每个通讯节点有且仅有一个通讯任务基于时间触发发送消息;同时有多个通讯任务基于事件触发发送消息.我们将每个调度周期细分为三种不同的时间片:控制时窗、确定时窗、竞争时窗.任务调度表的实例见图3,其中控制时窗的大小固定,位于每个周期 T 的起始部分,这个时窗被主控节点占用,完成主控节点传输周期调度信息、通讯节点同步计时器和解析调度信息;确定时窗是指分配给某个确定的通讯节点发送时间触发通讯数据的时间片;竞争时窗是指可以被系统所有通讯节点利用发送事件触发通讯数据的时间片.

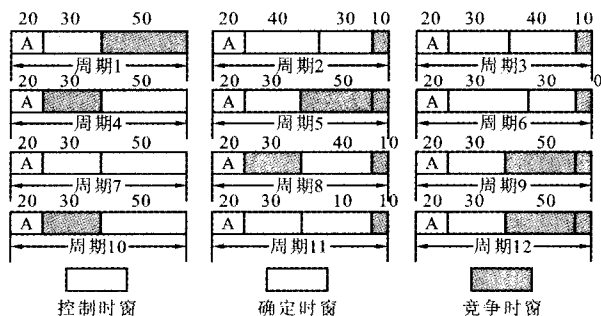


图3 调度表

3.1 调度表的生成

每个周期通讯任务均包含三个要素:执行时间、执行周期、死限.考虑到一个周期任务的执行必须在下个执行周期到来前结束,故可认为死限等于执行周期.生成调度表的关键是根据通讯任务的上述三要素和特定的周期任务调度算法^[5-7]来安排确定时窗,只要确定时窗的大小和位置确定,调度表就确定了.针对本系统的特点,我们采用文献^[7]中提出的方法来构建调度表:

首先,调度周期 T 取值为所有周期任务周期的最大公约数;宏周期即调度表起止时间跨度,取值为所有周期任务周期的最小公倍数.

其次,采用扩展的RM(Rate Monotonic)算法分配周期任务的优先级,即任务的周期越短,优先级越高,周期相同时,任务运行时间越长,优先级越高.

最后,根据任务优先级次序,安排确定时窗在调度表的位置.这一过程的描述较为复杂,详述见文献[7],它同时还给出并证明了调度表有效的充分必要条件.

这是一种硬实时调度方法,对于耗时确定的周期网络通讯,这种方法生成的调度表可以保证所有周期通讯任务的实时性要求.对于本系统,因为操作系统调度导致任务执行时间存在变数,仍存在少数任务丢失死限的可能性.

3.2 调度表编码

每个通讯节点都需要获取属于自己的确定时窗

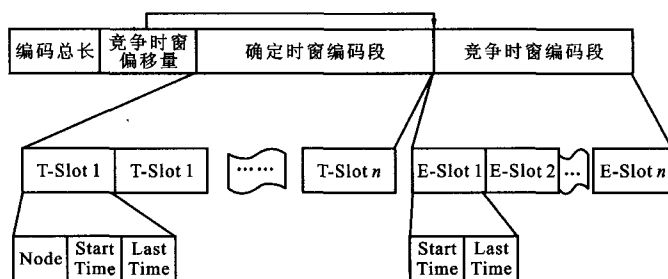


图4 调度表编码

4 通讯节点模型

上一节给出了系统总体的调度控制方法,下面我们关注系统中通讯的主体—通讯节点的模型,在本模型设计时,我们只关心通讯节点发送数据的功能.

4.1 事件触发通讯设计

对于某个通讯节点,假设其事件触发通讯任务有 N 个,分别记为 $T_i (1 \leq i \leq N)$. 为了缓存事件触发消息,我们为每一个事件触发通讯任务配备一个消息循环队列,分别记为 $Q_i (1 \leq i \leq N)$. 队列 Q_i 的大小与 T_i 预期触发频率和每次触发的平均数据量成正比. 当 T_i 被触发开始通讯时,如果循环队列 Q_i 中所有消息均未成功发送过,则任务 T_i 阻塞;否则,任务 T_i 将新消息插入到 Q_i 尾部,需要时覆盖已发送过的消息. 同时系统设有一个只在竞争时窗中执行的任务 T_{send} ,它负责从上述 N 个消息循环队列中读取消息并通过网络接口(如 Socket 等)发送出去——消息队列和网络接口的对应关系是预先确定的. 图5给出了 $N=3$ 时通讯数据发送模型,可知 T_{send} 和通讯任务是典型的单读者—多写者关系,在系统实现时需要同步.

为了控制各个事件触发任务发送消息的延迟时间,发送消息的任务 T_{send} 需要基于某种合理的策略来访问队列. 在此模型中采用“不等量循环优先”访

信息,此外也都需要获取所有竞争时窗信息,所以只编码传输这两个部分. 为使编码后的调度表简短易解析,我们将确定时窗和竞争时窗分开编码. 如图4所示,每个调度表项的编码包含四个部分:编码总长、竞争时窗的偏移量、确定时窗编码段、竞争时窗编码段. 确定时窗编码段由确定时窗信息组成,每个时窗信息包括:所属节点号、起始时刻、持续时长三部分. 竞争时窗编码段的开始位置可以通过竞争时窗偏移量方便地索引到,每个竞争时窗的编码包含两个部分:开始时间和持续时长.

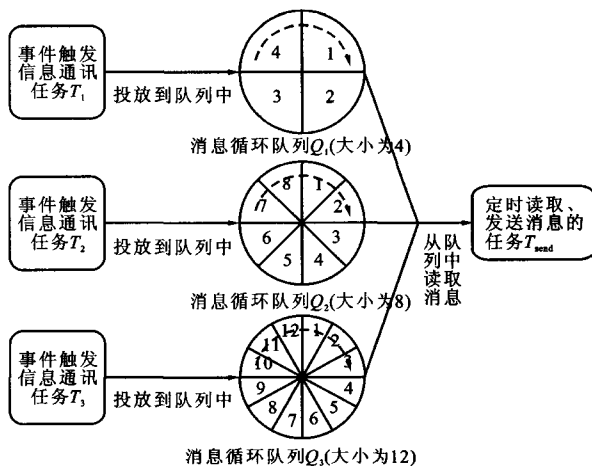


图5 竞争时窗中单个节点调度模型

问策略. “不等量”是指当 T_{send} 读取一个队列消息并发送时,最多可发送的消息个数跟队列的长度成正比. “循环优先”是指 T_{send} 按照 $Q_1, Q_2, \dots, Q_N, Q_1, Q_2, \dots$ 的顺序循环访问队列. 假设当在一个竞争时窗结束时 T_{send} 正在读取 Q_i 中的消息,如果其要求发送的消息个数正好完成,则下一个竞争时窗开始时 T_{send} 将从 Q_{i+1} 开始循环访问队列;否则在下一个竞争时窗开始时查看 Q_i 是否还有没有发送过的消息,并最多补发完上个竞争时窗没有来得及发送的剩余消息后,接着从 Q_{i+1} 开始循环.

由于通讯系统中有多多个通讯节点,不同通讯节点在竞争时窗中发送的消息会争用网络介质,这交

由局域网的介质访问策略来管理。

4.2 通讯节点调度控制

现在考虑通讯节点的实现时,只需要关心 3 个要素:时间触发的调度器;周期通讯任务;读取和发送事件触发消息的任务,即第 4.1 节中 T_{send} 。其中调度器是系统的核心。调度器的时间触发模型描述为

$$M = (Clock, Timer, Action)$$

其中 Clock 为本地计时器,在通讯节点接收到了调度消息的时候清零,考虑到通讯延时时间,Clock 清零时将直接置为一个初始值,而不是置零,其大小 C_0 为通讯延时的均值,通过若干次主控节点和通讯节点收发“消息—响应”的测试得到:

$$C_0 = \frac{1}{2m} \sum_{i=1}^m [T_{接收消息}^{主控节点} - T_{发送消息}^{通讯节点} - (T_{通讯节点}^{接收消息} - T_{通讯节点}^{发送消息})],$$

 m 为测试次数。Timer 记录了本节点占用的所有确定时窗和竞争时窗的开始时刻和持续时长。Action 是任务调度响应程序。在 Timer 记录的时窗开始时,Action 启动通讯任务;在 Timer 记录的确定时窗结束时如果周期任务还未完成则终止正运行的通讯任务,此时周期任务超时;在 Timer 记录的竞争时窗结束时,阻塞 T_{send} 。图 6 给出了调度器的主要控制流程。

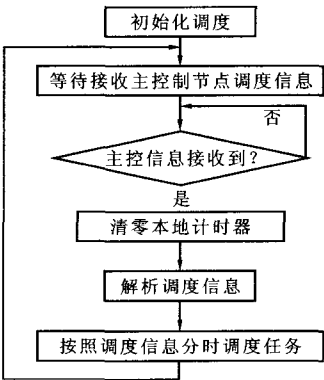


图 6 调度器控制流程

5 实验分析

为了分析上述调度控制方法的性能,我们构建了有 6 个周期任务(包括调度信息传输任务)的系统。任务参数表如表 1 所示。其中 A 为发送调度信息的任务,B—F 为周期任务。任务执行时间取为通讯任务脱离调度器干预的情况下执行 10 次得到的单次用时最大值。生成的调度表如图 3 所示。可知调度表的调度周期为 100 ms,宏周期为 1 200 ms。根据此调度表,得到了系统执行周期个数和周期任务

超时次数的曲线如图 7 所示。可知有少数周期任务错过死限而超时,这跟操作系统任务调度的不确定性有关。通过增大周期任务占用的时窗可以显著的减少超时,但这样会降低系统的带宽利用率,实时性的保证建立在浪费带宽的基础上。此外由于几个方面的原因,这个实验系统的时间精度不能很高:①本系统基于上层软件实现,而一般操作系统线程调度的时间片是毫秒级,这从根本上限制了系统调度的时间粒度。②当系统的时间粒度小时,Clock 初始值的测量误差对系统的影响是很大的,只有时间粒度比较合适时,才可能忽略。③软件计时器的精度一般都在毫秒级别,只有采用特别的机制能得到更精确地计时器。因此,基于本文方法构建的系统在时间粒度要求高的应用中会力不从心,但对于时间粒度要求不高的应用(例如视频音频传输等)和在仿真实验中,它们能满足用户基本需要的。

表 1 任务参数表 ms

任务	A	B	C	D	E	F
执行时间	20	30	50	40	30	40
执行周期	100	200	300	400	400	600

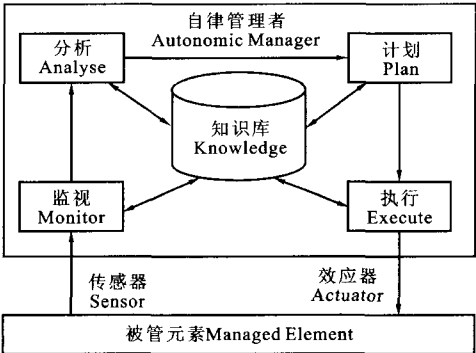


图 7 执行周期和超时次数对照图

6 结束语

本文提出了一个混合时间触发和事件触发的任务调度控制方法以实现时间敏感通讯系统,分析了系统中通讯任务的全局调度、时间表、通讯节点设计实现等关键要素,这些要素确保了周期通讯的实时性,同时也赋予了系统及时响应事件的能力。这种软件实现方法在时间粒度要求不高的实时通讯中,有一定的应用价值。

参考文献:

[1] 胡涛,宋子善. 分布并行实时系统的互连技术研究[J]. 微电子学与计算机,2007,24(4):200—203.

- [2] Hermann Kopetz, Astrit Ademaj, Petr Grillinger, et al. The time-triggered ethernet (TTE) design [C]// Eighth IEEE International Symposium on Object-Oriented Real-Time Distributed Computing, ISORC 2005. Washton; IEEE, 2005(5):22-33.
- [3] 冯晓东, 张争明, 张刚. TTCAN 网络调度平台的设计优化[J]. 微电子学与计算机, 2010, 27(3):67-70.
- [4] 杨仕平, 桑楠, 熊光泽. 基于 Ethernet 技术的安全关键实时网络[J]. 软件学报, 2005, 16(1):121-134.
- [5] 朱智林, 刘晓华, 韩俊刚. TTCAN 周期性任务的优化调度算法[J]. 兰州大学学报: 自然科学版, 2005, 41(4):73-76.
- [6] Qiao Xin, Wang Kunfeng, Sun Yuan, et al. A genetic algorithms based optimization for TTCAN [C]// IEEE International Conference on Vehicular Electronics and

Safety (ICVES) [S. L.]. IEEE, 2007:1-7.

- [7] 周悦, 王智于, 海斌, 等. 扩展单调速率算法及在 FF 调度时间表构建中的应用[J]. 信息与控制, 2001, 30(5):418-421.

作者简介:

徐亮 男, (1986-), 硕士研究生. 研究方向为数据管理、网络软件.

张延国 男, (1956-), 教授. 研究方向为网络存储、软件工程.

林奕 男, (1976-), 副教授. 研究方向为实时网络、存储网络.

林伟 男, (1981-), 博士生. 研究方向为数据存储、网络数据管理.

(上接第 67 页)

4 结束语

文中提出了一种基于可配置共享寄存器堆的多核处理器核间数据交换结构, 详细介绍了该设计的结构及实现机制, 并评估了该结构的性能. 评估结果表明, 使用核间共享寄存器的方式实现核间数据交换可以较明显的提高数据交换的效率.

参考文献:

- [1] 黄国睿, 张平, 魏广博. 多核处理器的关键技术及其发展趋势[J]. 计算机工程与设计, 2009, 30(10):2414-2418.
- [2] Jean Pegga Bissou, Mathieu Dubois, Yvon Savaria, et al. High-speed system bus for a soc network processing platform[C]// IEEE International Conference on Micro-

electronics. [S. L.]. IEEE, 2003:194-197.

- [3] Nozar Tabrizi, Nader Bagherzadeh. A distributed and shared register file for a multiprocessor-on-chip to support real-time applications. System-on-Chip for Real-Time Applications [J]. The 6th International Workshop on system-on-chip for Real-Time Applications, 2006:215-220.

- [4] 陈伟, 卢贵主, 郑灵翔. 二维 DCT 算法及其精简的 VLSI 设计[J]. 微电子学与计算机, 2008(3):107-110.

- [5] 董慧, 卫铭斐, 江丽, 等. 基于 FPGA 的 FFT 处理器的设计与仿真[J]. 微电子学与计算机, 2008(11):117-120.

作者简介:

方颖 女, (1986-), 硕士研究生. 研究方向为大规模集成电路设计与验证.