

## 对低端分布嵌入式系统中通信机制的改进

王 嘉,王 铮,童建林  
(重庆大学 计算机学院,重庆 400044)  
(wangjar@163.com)

**摘 要:**考虑到嵌入式分布系统在通信上的各种局限性,提出了一种新的通信方式。在总线式网络的基础之上,把交换式的概念引入到总线式通信系统中,并借鉴交换机和路由器中的一些优点,在链路层的基础上加入网络层,提高低端总线分布式系统的扩展性能和通信性能。最后,对所提出的方案作出了性能评价。

**关键词:**分布式系统;消息传递;通信;现场总线;以太网

**中图分类号:** TP393.04 **文献标识码:** A

## Improvement on communication of low-end distributed embedded system

WANG Jia, WANG Zheng, TONG Jian-lin  
(College of Computer Science, Chongqing University, Chongqing 400044, China)

**Abstract:** Aiming at many constraints of distributed embedded system into account, a new communication method was presented and analyzed. Based on the bus structure, some ideas from switcher and router were brought into the bus communication system: adding network layer on link layer, to improve the expansibility and performance of low-end bus-based distributed system. The performance of this design was evaluated by simulation tests.

**Key words:** distributed system; message transmission; communication; field bus; Ethernet

### 0 引言

分布式系统就是由多个相互连接的处理资源组成的计算机系统,这些资源可以合作执行一个共同的任务,尽可能少地依赖于集中的程序、数据和硬件等资源。分布式系统和单处理机系统一个最重要的区别是进程间的通信。在单处理机系统中进程间的通信无疑是利用共享存储器。一个典型的例子是生产者/消费者问题,一个进程向共享存储器写入,而另一个进程从该共享存储器中读出,甚至最基本的同步形式,如信号量(信号量自身可变)都要求必须有一个字的内存是共享的。在分布式系统中一般没有共享存储器,进程间的通信必须采用另外的方式,即消息传递方式。消息传递允许两个进程间通过将共享数据物理地拷贝到另一进程的地址空间来进行通信,这通过向包含共享数据的进程传送消息来完成<sup>[1]</sup>。

用于消息传递的互联方式多种多样,最简单的用于分布式系统的互联网络是总线。总线是静态的,并且不要求任何类型的路由。每个处理器和内存模块都连接到总线上,数据的所有事务和移动都发生在这个共享的总线上。Sequent 并行计算机就是用一条总线作为互联网络。虽然这是一个比较简单的互联网络拓扑,但总线限制了模块的带宽,从而产生了通信瓶颈<sup>[2]</sup>。

因为价格方面的原因,总线式的分布系统结构在低端的设备上得到了极为广泛的应用,但是在工厂和过程控制环境下的分布式自动化系统的复杂性与日俱增。特别地,越来越多的设备连结在一起,高速地交换大量的信息,所以提供更高的准确性和其他的附加功能(如远程诊断、维护和管理)十分重要。这也就意味着在通信网络中需要更宽的带宽。为了增加通信支持的扩展性和提高传输速率,人们提出了很多建议,

例如大家熟知的高速网络(如 FDDI)和制造业环境中的异步传输模式(ATM)。但是,这些解决方案都存在两个缺点:首先,它们都很昂贵(至少在几年的时间里),故采用这些先进的技术来连接传感器或致动器一类的简单廉价装置相当不划算;其次,它们与现行工业环境下的应用程序协议和行规(如 CANopen, DeviceNet, PROFIBUS)并不兼容<sup>[3]</sup>。

### 1 国内外现有的嵌入式系统的通信架构

#### 1.1 采用标准的工业用总线方式

国际电工委员会(IEC)制定的总线式协议模型不是按照 ISO 的 OSI 七层协议的规范,而是只由底层协议(包括物理层和数据链路层)和上层协议(应用层和用户层)以及行规组成<sup>[4]</sup>。由于实际功能的需要以及处理速度等原因,中间的网络层直到表示层都是被省略了的,如果要用到其中的功能,就放在链路层或是应用层来处理。

现行广泛应用于分布式控制和自动化系统领域的主要有两种网络模型。

**生产者/消费者模型** 信息按内容来标识,数据源只需将数据发送一次。许多需用此数据的节点通过在网上同时识别这个标识符,可同时从同一生产者取用此消费同一数据。如常见的 CAN<sup>[5]</sup> 协议,采用较为简单实用的短帧结构,其结构如图 1 所示。

**源/目的地模型** 顾名思义,这种模型下数据帧里带有目的主机和源主机的地址。例如巴基斯坦国立科技大学的几名学者提出的一种分布式系统的网络协议 GPCNA (General Purpose Controller Network Architecture)<sup>[6]</sup> 采用 37 个字节的长度,其中头部占 5 个字节,净负荷占剩下的 32 个字节。由于其总线式的连接方式,带来了诸多的不便。其只有一个字节

收稿日期:2004-08-16;修订日期:2004-10-27

作者简介:王嘉(1977-),男,重庆人,硕士研究生,主要研究方向:分布式系统中的网络通信;王铮(1953-),男,上海人,副教授,主要研究方向:操作系统平台、通信协议;童建林(1979-),男,四川眉山人,硕士研究生,主要研究方向:Petri Nets、动态加载。

的 identifier,更是限制了整个网段上可连接设备的个数,从而影响了整个系统的可扩展性。

CAN2.0B 扩展标准的出现解决了这个问题,它把标识位从 11 位扩展至 29 位,这就要求更多的总线访问时间,降低了有效的数据传输速度。而且由于总线争用,即使能够加载上许多的站点,冲突与争用已使其失去了实际意义。

| 起始位 | 控制域          | 仲裁域                                  | 数据域         | 校验域          | 结束位 |
|-----|--------------|--------------------------------------|-------------|--------------|-----|
|     | 11bit<br>标识符 | 远程发送请求<br>(RTR)和 4bit 数<br>据长度码(DLC) | 8Byte<br>数据 | 16bit<br>CRC |     |

(a) CAN帧结构

| 1 Byte                 | 1 Byte            | 1 Byte     | 2 Byte        | 32 Byte |
|------------------------|-------------------|------------|---------------|---------|
| Destination<br>Address | Source<br>Address | Attributes | Header<br>CRC | Payload |

| 1 bit | 1 bit            | 1 bit    | 5 bit             |
|-------|------------------|----------|-------------------|
| Token | Multi or Unicast | Reserved | Length of Payload |

(b) GPCNA帧结构

图1 两种不同的帧结构

## 1.2 采用以太网和嵌入式 TCP/IP 协议

快速以太网和嵌入式的 TCP/IP<sup>[7]</sup> 协议虽然就其功能来说比较完善,可以支持和 Internet 的互联,并且已经有了相关的产品出现,但它们主要用于工作和单位级,而不是低端分布嵌入式系统中以传输短消息为主,故显得过于庞大和造价昂贵。

传统的分布式系统现场总线结构抛弃网络层直到表示层的重要原因就是牵涉到具体的应用,这极大地影响到了整个系统的可扩展性,所以一些分布式系统就使用 IP(或 UDP,它建立在 IP 之上)或嵌入式 TCP/IP 作为基本协议,有以下一些优点:1)协议已经存在,省去了大量的工作;2)该协议已有许多个现成的工具,这也省去了许多工作;3)绝大部分 UNIX 系统中都可发送和接收该信包;4)大多数现有网络支持 IP 与 UDP 信包。而且 IP 与 UDP 适用于大多数现有的 UNIX 系统和网络系统(如 Internet)。这样,编写在 UNIX 系统上运行的客户或者服务器程序较容易,就加快了代码运行和测试的速度。

然而 IP 与 UDP 在性能方面有着自己的缺陷。IP 不是为最终使用者设计的协议。它被设计为一个基础,其上可以建立在不同内部网络的可靠的 TCO 连接。例如,它能协助网关将信包拆分,并以网络允许的最大尺寸在网上传送。尽管基于 LAN 的分布式系统并不需要这个特性,但这些信包消息仍然由发送者填入 IP 信包的头字段中,然后由接收者重组形成一个合法的 IP 信包。IP 信包共有 13 个头字段,但是仅有发送者与目的地的地址和信包长度三个字段是有用的。其他 10 个字段只是随着发送。这 10 个字段中有一个字段是头的校验和,计算校验和将很费时。UDP 协议的情况更糟,它的头字段里有两个字段是校验和<sup>[1]</sup>。这也是在现场总线中一般不使用网络层和传输层的原因之一。

因此,在这里提出了一种解决方案,综合了两种方式的优点,以提高总线式系统的可扩展性。

## 2 对传统现场总线的改进

### 2.1 网络结构

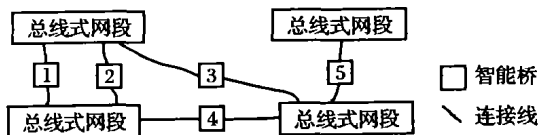


图2 不规则总线网络连接拓扑图

我们考虑把各个不同的总线网段通过一种叫智能桥接器

的设备连接起来,如图2所示。

网段1至4都是连通的,其中1与2之间还有两条线路,来增加网络的冗余性。

### 2.2 工作原理

整个网络系统在物理上是连通的,把每一个总线式网段抽象为一个节点。整个系统加电启动后,各个智能桥就通过广播登录控制信息,以了解整个网络的拓扑状况,并把拓扑结构存入自身的路由表中。为了避免广播风暴,就需智能桥在加电后启动生成树(spanning tree)算法<sup>[8]</sup>,让任意两个有物理通路的节点之间有且仅有一条逻辑路径。

每个网段上可连接的主机的数量由本来总线所能支持的最大数量决定,同一网段里的主机帧地址不允许相同,而不同网段的允许相同。这时每个主机的地点不再仅由帧地址决定,而是由(网段 ID + 帧地址)共同决定。在每个网段里流动的包由桥接器决定是否向外转发,以及向何处转发。由于是嵌入式系统,根据 80/20 的原理,大部分的信息流保持在局部子网之内,这样就把信息限制在一定的范围内,避免大范围内的帧冲突和重传。

信息在传递的过程之中,各个桥因为已经知道整个网络中生成树的结构,会采取第三层交换机组播技术中经常使用的反向路径转发和剪枝技术<sup>[9]</sup>,这样就不会造成消息在整个网络中无休止地循环下去的情况。

允许动态加载或拆卸智能网桥或主机。当智能桥动态加入时,要向邻接的桥发出路由表请求信息,邻接的桥把在路由表中加上此桥后,更改生成树,并向其他桥扩散此信息,以更新它们的路由表。如果是主机动态加入,就向位于同一个网段的其他主机发送类似 ARP Request 的请求,通过其他主机的回应来确定自己的网段 ID 号,有了网段 ID 号之后,就可以顺利通信了。

当有非割边(非关键性的路径,如图中的智能桥1)的线路断开时,智能桥在一定时间内没有收到逻辑相邻的桥传来的消息时,智能桥能够根据网络拓扑结构迅速地选取另一条代价比较小的线路,使整个网在逻辑上尽可能处于连通状态。

如此也就可以在原来的网络框架(如 CAN, GPCNA)基本不变的基础上,增加其扩展性和提高其传输性能。

### 2.3 附加的网络层协议

在不能改变数据链路层的基础之上,我们需要的网络层的包结构如图3。

|            |    |           |    |
|------------|----|-----------|----|
| 0          | 15 | 16        | 31 |
| 16位目标网段ID号 |    | 16位源网段ID号 |    |
| 8位端口号      |    | 类型(优先级等)  | 保留 |
| 净负荷        |    |           |    |
| ...        |    |           |    |
| ...        |    |           |    |
| ...        |    |           |    |

图3 重新定义的包结构

对各个字段含义的详解如下:

目标及源网段 ID 如前所述,总线式网段中 ID 号由第一个登录在此网段上的设备随机生成,然后就可由(网段 ID 号 + 帧地址号)来唯一地确定一台主机,这时其整个网络所拥有的主机的范围在理论上就可由 2032 台(标准 CAN)扩展至以前的  $2^{16}$  倍,使其总数量从理论上扩大 32 768 倍,从而增加了网上可连接的机器数。

虽然与 CAN2.0B 扩展标准的效果等同,但采用此方式与 IP 协议中(子网号 + 主机号)的方式类似,由网段 ID 迅速定位至哪个网段,这样就防止了这个包在整个网络中到处跑动,也可以极大地缩短智能网桥里面路由表的长度。

**端口号** 由于一台主机上可运行多个进程,而一个进程又可能有多个网络需求,所以端口号可确保包准确无误地送到某个应用之中。

**类型** 主要用于对包的优先级的控制。智能桥接器在由于到来的包超过其转发能力而拥塞时,可以对此域进行仲裁,以决定对其中的一些包进行丢弃或优先转发。

**保留** 预留给进一步的扩展,可用于 TTL(生命周期)以及数据分片之用。TTL 是 IP 数据包在网络中存在的时间,如果以秒为单位,则数据包的最大生命期为 255。实际上,该值只以智能网桥计数,每经过一个桥就减 1。当它减到 0 时,网络就丢弃该数据包,并向数据包的源发送一个告警数据包。由于整个拓扑结构存在着环,这一特性旨在防止数据包在网中无限制地循环或漫游。而数据分片主要是考虑到随着技术的不断发展,这个字段可以传送较长的数据的需要。

#### 2.4 性能分析

由理论分析可知,拓扑的变化和处理器的接入位置对性能的影响很大。除此之外,消息长短及其分布对性能也有相当大的影响。

对性能的评价一般用网络延迟和吞吐率作为指标。延迟是以时间单位来度量的。然而,当比较几种设计选择时,绝对值是不重要的。由于很多比较是用网络模拟器进行的,所以延迟可以用模拟器时钟周期来度量。除非另有声明,本节所指的单播消息的延迟是以消息头从源节点注入网络到目的节点接收到最后一个消息单位所经历时间的平均值来度量的<sup>[10]</sup>。单个消息的延迟是不重要的,多数情况下,我们感兴趣的是延迟的平均值。但标准偏差也很重要,因为如果某些消息延迟比平均值大许多的话,实时性就得不到应有的保证。

吞吐率是每单位时间发送的最大信息量,也可以定义为网络可接受的最大流量。这里,流量或者可以接受的流量就是每单位时间发送的信息量。根据使用的是绝对时间还是相对时间,吞吐率可以采用每秒或每个时钟周期的消息数来度量。但是,吞吐率取决于消息的大小和网络规模。因此,吞吐率应进行标准化并需要除以消息大小和网络规模。结果,吞吐率便以每节点每个时钟周期多少位来度量。如果负载最重要的通信其时间使用率达到 100%,那么均匀负载网络的吞吐率就等于网络的容量。

随机有色 Petri Net 是进行网络性能分析比较成熟的工具之一,我们借用它来进行模拟。考虑到平均化和简化,在这里作出严格限制:1)每个网段只连接有 4 个处理器;2)整个网络已经从物理上和逻辑上连成树形拓扑结构;3)由于轻负载下不能体现出交换相对于总线的优点,故着重考虑重负载下的性能;4)消息通过桥接器的时间延迟为一个时间单位;5)波特率、误码率、最坏响应时间、实时报文可调度性等虽然也是衡量指标,因为与本文讨论的主题无关,所以不在本文分析的范围之内。

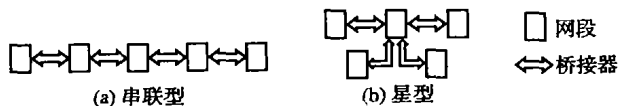


图4 进行评估的两种连接模型

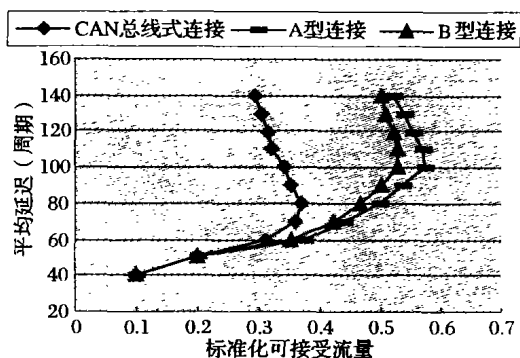
我们用串联型和星型两种最具代表性的拓扑,和原有的总线式连接进行比较,模拟了其在不同的消息字节以及不同的消息延迟下的通信能力,采用 Burton 范式<sup>[10]</sup>(BNF)加以分析。X 轴代表标准化可接受流量,Y 轴代表平均延迟。

从图5可以看出,CAN 总线式连接方式在达到饱和度之后延迟继续上升,吞吐率却在下降。串联型连接达到了最高

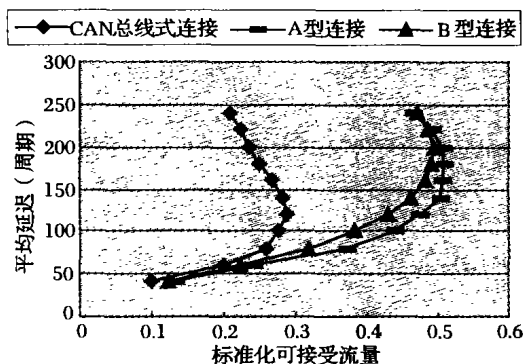
的吞吐率,在其高峰期几乎为总线式的 2 倍。这虽然和理论计算有一定的差距,却也表明了把流量限制于本网内的优势。尽管星型连接网有最小的直径,但它把网段之间的消息交流集中在位于中心的网段之中。这种不平衡直接导致了星型比串联型的最大可持续吞吐率更低。

但出乎意料,星型在很重的网络负荷下有较好的表现。在这种情况下,所以有的桥接器都十分繁忙,且消息极易被阻塞。因此,有最小直径的星型反而有较低的网络延迟和较高的吞吐率。

最后可以肯定的是,在突发传输时网络的最大可持续吞吐率会相当的低,因为桥接器和总线网段会被长消息所占用,而堵塞住了其他消息。



(a) 短消息下的网络性能 (24字节)



(b) 长消息下的网络性能 (128字节)

图5 性能测试结果

#### 2.5 对整个设计的考虑

- 1) 由于是嵌入式系统,应尽可能地把非净负载长度减短;
- 2) 嵌入式系统以短消息为主,故未考虑数据分片的问题;
- 3) 现有网络传输质量的不断提高,已经可以保证数据在链路上的可靠传递,因此没有加上通常意义上的数据校验;
- 4) 桥接器完成了传统的桥和路由器以及交换机的一部分功能,但要求尽量简化,如果太复杂,就突破价廉的限制;
- 5) 众多的网桥可以合成为一个多接口的桥设备,虽然可以进一步提高性能,但是这样就使此桥成为关键性设备,只要它一出故障,就会使整个网络不能有效工作。

### 3 结语

本文对如何在低端分布嵌入式系统中进行高效的消息传递从理论上提出了一个初步的改进措施,使其性能得到了一定的扩展。就其未来的发展而言,还有许多工作要做。

比如可以采用第三层交换机中使用的主干技术(trunking),使两个总线段之间的多条物理链路被定义为主干,主干线路就能同时分担数据,传递消息,实现真正意义的冗余。

拟实验是在一个给定范围的区域( $150 \times 150$ )内,随机散布  $N$  个节点,每个节点有一定的通信半径  $R$ 。通过分别改变  $N$  和  $R$  来比较在不同的失效模型下,不同多路径构造算法的性能。根据文献[9],节点的失效模型可分为两类:I型失效和P型失效。I型失效指的是某些节点随机独立失效,而P型失效指的是在某个半径范围内的节点整体失效。

图3~图6是分别表示了网络内节点数、节点传输半径、I型失效的节点数、P型失效的节点数改变时,对算法容错度的影响。图中M1和M2分别表示不同型值点构造多路径和不同策略构造多路径的算法。

从图中可见,方法一对于I型失效的容忍度比方法二好,尤其是在失效节点比例较高时,这是因为方法一构造的多路径相互之间几乎没有影响,而两种方法对P型失效的容忍度相差不大。当 $R$ 变化或 $N$ 变化时,相比之下,方法一要稍好于方法二。但方法一在具有较好性能的同时也需要更多的开销,因而需要在性能提高和附带开销这两方面进行有效的折中和平衡。

#### 参考文献:

- [1] ACRE J, CLARE L. An integrated architecture for cooperative sensing networks[J]. IEEE Computer, 2000, 33(5): 106-108.
- [2] ESTRIN D, GOVINDAN R, HEIDEMAN J, et al. Next century challenges: Scalable Coordination in Sensor Networks[A]. Proc of the 5th annual ACM/IEEE international conference on mobile computing and networking[C]. Seattle, Washington, 1999. 263-270.
- [3] 李建中, 李金宝, 石胜飞. 传感器网络及其数据管理的概念、问题与进展[J]. 软件学报, 2003, 14(10): 1717-1727.
- [4] 任丰原, 黄海宁, 林闯. 无线传感器网络[J]. 软件学报, 2003, 14(7): 1282-1290.
- [5] PARK VD, CORSON MS. Temporally - ordered routing algorithm (TORA) version 1: Functional specification[S]. draft-ietf-manet-tora-spec-01.txt, 1998.
- [6] BROCH J, JOHNSON DB, MALTZ DA. The dynamic source routing protocol for mobile ad hoc networks[Z]. IETF Internet Draft,

1998.

- [7] Nasipuri A, DAS SR. On-Demand Multipath Routing for Mobile Ad Hoc Networks[A]. 8th International Conference on Computer Communications and Networks (IC3N 99)[C], 1999.
- [8] INTANAGONWIWAT C, GOVINDAN R, ESTRIN D. Directed diffusion: A scalable and robust communication paradigm for sensor networks[A]. Proc Sixth Annual International Conference on Mobile Computing and Network[C], 2000.
- [9] GANESAN D, GOVINDAN R, SHENKER S, et al. Highly-resilient, energy-efficient multipath routing in wireless sensor networks [J]. Mobile Computing and Communications Review, 2001, 4(5).
- [10] DULMAN S, NIEBERG T, WU J, et al. Trade-Off between Traffic Overhead and Reliability in Multipath Routing for Wireless Sensor Networks[Z]. WCNC Workshop, New Orleans, Louisiana, USA, 2003.
- [11] DULMAN S, WU J, HAVINGA P. An energy efficient multipath routing algorithm for wireless sensor networks[A]. IEEE International Symposium on Autonomous Decentralized Systems (ISADS 2003) [C]. Pisa, Italy, 2003.
- [12] YUKSEL M, PRADHAN R, KALYANARAMAN S. Trajectory - Based Forwarding Mechanisms for Ad-Hoc Sensor Networks[A]. IEEE 2nd Upstate New York Workshop on Sensor Networks[C]. Syracuse, NY, 2003.
- [13] PARKINSON B, SPILKER Jr JJ. Global Positioning System: Theory and Application[A]. Progress in Astronautics and Aeronautics[C], 1996, 163.
- [14] NICULESCU D, NATH B. Localized positioning in ad hoc network [A]. Sensor Network Protocols and Applications[C], 2003.
- [15] BRAGINSKY D, ESTRIN D. Rumor routing algorithm for sensor networks[A]. Proceedings of the 1st ACM international workshop on wireless sensor networks and applications[C], 2002.
- [16] WAN P-J, ALZOUBI KM, FRIEDER O. Distributed Construction of Connected Dominating Set in Wireless Ad Hoc Networks[A]. Twenty-First Annual Joint Conference of the IEEE Computer and Communications Societies[C]. 2002, 3. 1597-1604.

(上接第500页)

此外,破坏生成树的结构,使整个链路不仅在物理上,也在逻辑上连通。借鉴路由器中的方法,每个包都可以独立的选择自己的最佳路径,如果最短路径发生堵塞时,连接网段的桥能够自动选取相对最佳的路径供数据的传递。如果能够更加智能化,就可以实现长信息和短信息分别选路。

从长远的眼光来看,随着以太网技术的不断成熟,人们为适应现场总线需要对以太网也作出各种改进,特别是Profibus、Devicenet、Controlnet和Lonworks等公司都在研究通过一种称作管道(tunnel)的简单传递机构,使用Ethernet网络传送报文。与此同时,美国电气工程师协会(IEEE)正着手制定现场装置与Ethernet通信的新标准。该标准能够使网络直接“看到”对象(object)。这些工作都为Ethernet进入工业自动化的现场级打下了基础<sup>[4]</sup>。

在解决了网络的不确定性(Determinism)、环境适应能力、抗干扰能力以及自身的供电能力等问题之后,把以太网运用于工业总线中正成为当前一个比较热门的研究话题。

综上所述,我们的最终目标是在有成本限制的情况下,提高分布式系统互连的可伸缩性、递增可扩充性、可分割性以及可靠性与可修复性。

#### 参考文献:

- [1] TANENBAUM AS. 分布式操作系统[M]. 陆丽娜, 伍卫国, 刘

隆国, 等译. 北京: 电子工业出版社, 1999.

- [2] GALLI DL. 分布式操作系统——原理与实践[M]. 徐良贤, 唐英, 毛家菊, 等译. 北京: 机械工业出版社, 2002. 14-15.
- [3] CENA G, VALENZANO A. FastCAN: A High-Performance Enhanced CAN-Like Network[J]. IEEE Transactions on Industrial Electronics, 2000, 47(4): 951-963.
- [4] 周明. 现场总线控制[M]. 北京: 中国电力出版社, 2002.
- [5] ISO 11898, Road Vehicles-Interchange of Digital Information-Controller Area Network(CAN) for High-Speed Communication(First Edition)[S], 1993.
- [6] YOUNUS B, AHMAD B, BASHIR O, et al. A network protocol for distributed embedded systems[A]. Students Conference, ISCON'02 [C]. IEEE, 2002, 1. 149-153.
- [7] 张满怀. 嵌入式TCP/IP协议研究与实现[D]. 广东工业大学, 2001.
- [8] IEEE Std 802.1D-1990, IEEE standards for local and metropolitan area networks: media access control(MAC) bridges[S], 1990.
- [9] COMER DE. 用TCP/IP进行网际互联——第一卷: 原理、协议与结构[M]. 林瑶, 蒋慧, 杜蔚轩, 等译. 北京: 电子工业出版社, 2001. 233-234.
- [10] DUATO J, YALAMANCHILI S, NI L. 并行计算机互连网络技术[M]. 谢伦国, 张民选, 窦强, 等译. 北京: 电子工业出版社, 2004. 342-343.