



# [12] 发明专利说明书

专利号 ZL 02151524.7

[45] 授权公告日 2005 年 10 月 26 日

[11] 授权公告号 CN 1225101C

[22] 申请日 2002.12.31 [21] 申请号 02151524.7

[71] 专利权人 浪潮电子信息产业股份有限公司

地址 250014 山东省济南市山大路 224 号

[72] 发明人 王守昊 尹宏伟 庄文君

审查员 王红丽

[74] 专利代理机构 济南信达专利事务所有限公司

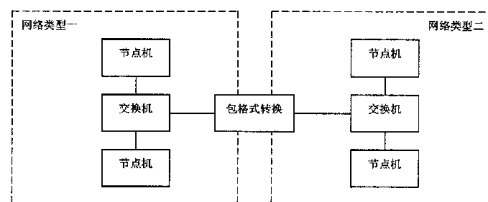
代理人 姜 明

权利要求书 1 页 说明书 5 页 附图 1 页

[54] 发明名称 基于数据包的数值特性的多协议通信方法

## [57] 摘要

本发明提供一种基于数据包的数值特性的多协议通信方法，在该方法中使用数据包中原来某些数值域的数值特性定义没有被覆盖该区域的整个数值域的数值特性，来表示附加的控制或指示等信息，进行多种协议之间通信。本发明的通信方法的优点是在不改变标准的规范和协议定义的数据包格式的情况下，使用已有的数据包实现多协议通信。不需要重新改写数据包的格式，或者重新定义一种标准或格式即可实现不同网络之间的通讯。



1. 基于数据包的数值特性的多协议通信方法, 其特征在于是使用数据包中原来没有定义或没有被保留使用的数值域来表示附加的控制或指示信息和进行多种协议之间通信, 两种不同的网络协议之间数据传送时, 数据包通过第一个网络的第一个适配器发送, 然后由第一个网络的第二个适配器接收, 然后转发到第二个网络的适配器, 具体步骤是:

① Raw 数据包本身由 RWH 和有效负载组成, 当 HCA 希望发送 raw 包到 TCA 时, 通过使用 LRH 指明 TCA, HCA 根据包含在 RWH 和 Raw 包的有效负载中的信息自动生成 V-CRC 值, HCA 接着经过 InfiniBand 发送这个完整的数据包, 最后被指定的 TCA 接收;

② TCA 检验数据包的正确性, 删除 LRH 和 RWH, 将作为结果的 raw 数据包传送到网络适配器, 网络适配器接着将数据包发送到通信网络中。

2. 根据权利要求 1 所述的基于数据包的数值特性的多协议通信方法, 其特征在于第一个网络包括 InfiniBand 网络, 第二个网络包括通信网络。

3. 根据权利要求 2 所述的基于数据包的数值特性的多协议通信方法, 其特征位于 InfiniBand 网络包括 HCA、TCA 和 InfiniBand 交换机, 通信网络包括各种局域网。

4. 根据权利要求 1 所述的基于数据包的数值特性的多协议通信方法, 其特征位于数据区域在 InfiniBand 网络中指的是 RAW 数据报的 RWH 头的 ethertype 区域。

5. 根据权利要求 1 所述的基于数据包的数值特性的多协议通信方法, 其特征位于通过一个大于或等于 1536 (十进制) 或 0x600 (十六进制) 的数值表示特定的数据类型; 通过将该区域设置为小于 1536 的数值来使得 raw 数据报传送没有被 ethertype 定义的类型的数据。

## 基于数据包的数值特性的多协议通信方法

### 1、技术领域

本发明涉及计算机和计算机网络通信领域，尤其是在InfiniBand网络和以太网之间的通信，具体地说是一种基于数据包数值特性的多协议通信方法，尤其是在InfiniBand网络 and 不同类型的以太网之间的通信方法。

### 2、技术背景

在计算机和通信领域，有各种各样的标准，在这些标准的规范和协议中都会定义很多数据包的格式，遵照这种格式与特定的协议之间通信，传送特定的数据，而不能够传送在规范中没有定义的数据类型。尤其是在两种不同的网络之间传送数据的时候，由于没有提供对某种网络协议的支持，而使得这样的网络之间不能够通信。

一个典型的例子就是I/O网络与通信网络之间的数据通信。比如InfiniBand网络和局域网之间的通信。InfiniBand网是一种I/O网络，包括InfiniBand链接和交换机。HCA和TCA通过InfiniBand链接彼此通信，InfiniBand网络之间通过链接和路由器互相通信。目标外设和网络适配器通过TCA发送和接收数据包。数据包也可能通过路由器发送和接收。InfiniBand网提供TCA和HCA之间各种方式的通信。像其他类型的网络一样，InfiniBand网络有物理层，链路层，网络层，传输层和上层协议。像其他类型的包传送网路一样，在InfiniBand网络中，事务被分割成信息，然后被分割成数据包在InfiniBand网上传送。当接收者接收到数据包，将它们重组成信息。

每一种TCA都针对专门的外设比如网络适配器，SCSI适配器等设计。针对以太网适配器设计的TCA使得以太网适配器能够经过InfiniBand网发送和接收数据。

InfiniBand网允许不同的传输服务，包括可靠和不可靠的连接，可靠和不可靠的数据报，raw包支持。在可靠连接和可靠数据报中，通过应答和包序列号确保数据包的顺序。重复的包被丢弃，丢失的包能够被检测到。不可靠的连接和不可靠的数据报中，不产生应答信息，包的顺序不被保证。

在InfiniBand规范中，定义了一种Raw数据报的数据包结构，该数据包中的RWH（RAWHeader）部分有两个2个字节的区域，第一个是Ethertype区

域，也叫长度/类型区域，这一区域的数值设置遵循IEEE 802.3标准，可以设置一个大于或等于1536（十进制）或者0X600（十六进制）的数值，来表示要传送的实际数据的类型，主要是指通信协议类型。这些数值是由IANA（Internet Assigned Numbers Authority）组织来管理和分配的。

Ethertype数值主要定义的是网络层协议。在网络模型中，以太网，令牌环网等局域网协议位于第2层，其上位于第3层的上层协议。这些上层协议包括IPv4/IPv6，ARP，IPX，SNA等。IEEE 802.3 Ethertype数值表示的就是数据包中传送的第3层协议类型。比如，Ethertype数值0X800表示IPv4，0x806表示ARP，0x80D5表示SNA over Ethernet等。第2层通信协议或媒介主要包括以太网（DIX和IEEE 802.3），令牌环网（IEEE 802.5）。其它的还有ATM（RFC2225IETF）、LANE over ATM和令牌总线（IEEE 802.4）。IEEE 802.3以太网，IEEE 802.4 令牌总线，IEEE 802.5令牌环网等都要遵循IEEE 802.2标准，该标准使用LLC（Logic Link Control）头和SNAP（System Network Architecture Protocol）头。这些协议头包括一个提供Ethertype数值的区域。DIX以太网不使用LLC或SNAP头，但是仍然包括提供ethertype数值的区域。

对于大多数的第2层通信，都需要遵循IEEE 802.2 SNAP或者IEEE 802.2 LLC标准。这意味着为了真正在局域网和InfiniBand网之间实现桥接，必须在InfiniBand网络里传送SNAP和LLC头信息。但是为了在InfiniBand数据包中包含以太网信息，InfiniBand标准仅仅指定了IEEE 802.3标准。因为IEEE 802.3标准不传送SNAP和LLC头信息，所以需要这些信息的局域网通信数据包不能够经过InfiniBand网发送。而Ethertype数值也不能够表示第2层数据包。

因此，InfiniBand规范由于使用IEEE 802.3 ethertype限定了非InfiniBand网与InfiniBand网之间通信的数据包只能是上层协议数据包，第2层数据包不能够通过InfiniBand网络传输。

### 3.发明内容

本发明的目的是提供一种基于数据包的数值特性的多协议通信方法。

本发明的基于数据包的数值特性的多协议通信方法，是使用数据包中原来定义ethertype数值域的数值特性以外的没有被覆盖该区域的整个数值域的数值特性，来表示附加的控制或指示等信息，进行多种协议之间通信。

本发明的基于数据包的数值特性的多协议通信方法，是将原来定义ethertype数值域的数值特性以外的没有被覆盖该区域的整个数值域的特性数值设置为小于1536的数值来使得raw数据报传送没有通过ethertype指定的类型

的数据。

本发明的基于数据包的数值特性的多协议通信方法，是在两种不同的网络协议之间进行数据传送，将数据包通过第一个网络的第一个适配器发送，然后由第一个网络的第二个适配器接收，转发到第二个网络的适配器。

在本发明的基于数据包的数值特性的多协议通信方法中，第一个网络包括 InfiniBand 网络，第二个网络包括通信网络，在本发明的基于数据包的数值特性的多协议通信方法中，InfiniBand 网络包括 HCA、TCA 和 InfiniBand 交换机；通信网络包括各种局域网。

本发明的基于数据包的数值特性的多协议通信方法，在 InfiniBand 网络中的数据区域是 RAW 数据报的 RWH 头的 ethertype 区域。

附图 1 是 raw 数据报头；

附图 2 为基于数据包的数值特性的多协议通信方法的两种不同的网络之间通信的结构图结构示意图；

附图 3 是 InfiniBand 规范指定的 raw 数据包格式。

参照说明书附图对本发明的基于数据包的数值特性的多协议通信方法作以下详细地说明。

如附图 1 所示，利用 raw 数据报头的 Ethertype 区域的数值特性，实现多协议之间的通信。raw 数据报头对应于 IEEE 802.3 标准中指定的 ethertype 头，该数据头有两个区域，第二个区域被保留。第一个区域是一个两个字节的区域，通过一个大于或等于 1536（十进制）或 0x600（十六进制）的数值表示特定的数据类型。在本发明所描述的方法中通过将该区域设置为小于 1536 的数值来使得 raw 数据报传送没有通过 ethertype 指定的类型的数据。

附图 2 是本发明提到的两种不同协议的网络之间通信的结构图，数据通过第一种网络发送，经网络服务器对数据处理后转发到第二种网络中。

附图 3 是 InfiniBand 规范指定的 raw 数据包格式。本地路由头(LRIH)表明 RAW 包要被发送到的 TCA 或路由器。Raw 头(RWH)是 raw 数据的实际头，用它表示要传送的数据类型。Raw 包的数据信息包含在 raw 包的有效负载(payload)中。

V-CRC 是循环冗余校验码，通过该值来检验接收到的数据包是否正确。

图 4 展示了数据包在两种不同的网络间之间传送的流程图。

#### 4、实施方式：

一、在 InfiniBand 网络和通信网络之间的通信：

本发明所涉及到的计算机系统包括一个InfiniBand I/O网络，一个HCA和一个TCA，一个连接到TCA和通信网络的网络适配器。一个raw数据报用来在InfiniBand网的HCA和TCA之间传送通信网络信息。通信网络可能是IEEE 802.2 LAN，以太网，或者其他类型的网络。HCA发送raw数据报，TCA接收到以后，经过处理，转发给网络适配器，然后发送到通信网络。此数据报具有一个raw头，对应于IEEE 802.3标准中的ethertype头，该数据头有两个区域，第二个区域被保留。根据IEEE 802.3标准，第一个区域是一个两个字节的区域，通过一个大于或等于1536（十进制）或0x600（十六进制）的数值表示ethertype，这些数值由IANA管理和分配。但是，本发明通过将该区域设置为小于1536的数值来表示的要发送的数据的类型信息。本发明使用已有的数据包结构提供这些要被传送的信息。

当HCA希望发送raw包到TCA时，它使用LRH指明TCA。根据包含在RWH和Raw包的有效负载中的信息，HCA根据数据内容自动生成V-CRC值。Raw数据包本身由RWH和有效负载组成。HCA接着经过InfiniBand发送这个完整的数据包，最后被指定的TCA接收。TCA检验数据包的正确性，删除LRH和RWH，将作为结果的raw数据包传送到网络适配器。网络适配器接着将数据包发送到通信网络中。

InfiniBand交换机仅仅根据LRH转发数据包，因此，Raw包在InfiniBand网络中传输的时候，任何协议头都不会被解码。TCA和网络适配器负责解码raw包。经过InfiniBand网络的时候，数据包由V-CRC提供保护。

## 二、通过ethertype区域指定特定数值从而在数据包中传送的信息：

本发明能够使用ethertype区域来表示各种要被传送数据包的类型信息。Ethertype区域可能设置成0来表示正在传输的raw数据包与DIX帧一致；设置为1来表示正在传输的raw数据包与IEEE 802.2帧一致；可能设置成2来打开多播模式，多播模式允许从网络适配器发送一个包到所有连接在通信网络中的其他网络适配器。ethertype区域还可能设置成其他值来打开VLAN支持，或者在利用千兆以太网络适配器的时候打开巨帧支持。

以下是一个传送与DIX兼容的数据包的例子，LRH指明了网络适配器的TCA。RWH被设置成0。Raw包的有效负载指明了目的以太网地址，源以太网地址，使用的协议和实际的数据。

以下是一个传送与IEEE 802.2兼容的数据包的例子，LRH指明了网络适配器的TCA。RWH被设置成1。Raw包的有效负载指明了目的以太网地址，

源以太网地址，长度，其他一些IEEE 802.2标准需要的数据和实际的数据。

本发明的基于数据包的数值特性的多协议通信方法的优点是在不改变标准的规范和协议定义的数据包格式的情况下，使用已有的数据包实现多协议通信。不需要重新改写数据包的格式，或者重新定义一种标准或格式即可实现不同网络之间的通讯。

RWH 格式



图 1

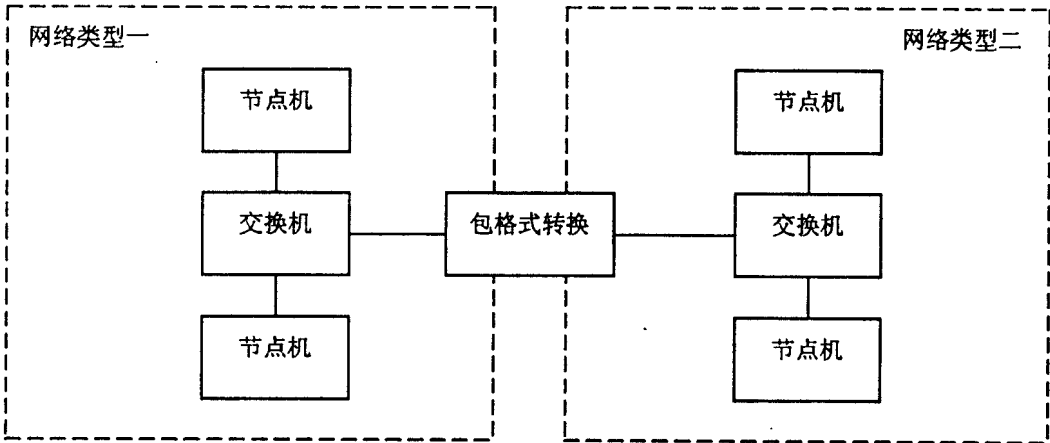


图 2

包顺序

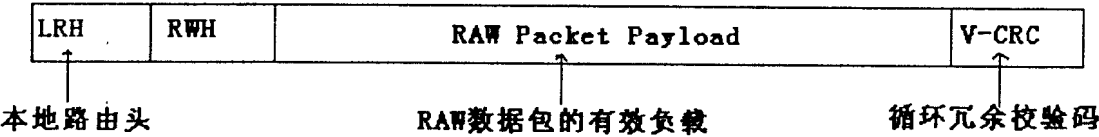


图 3

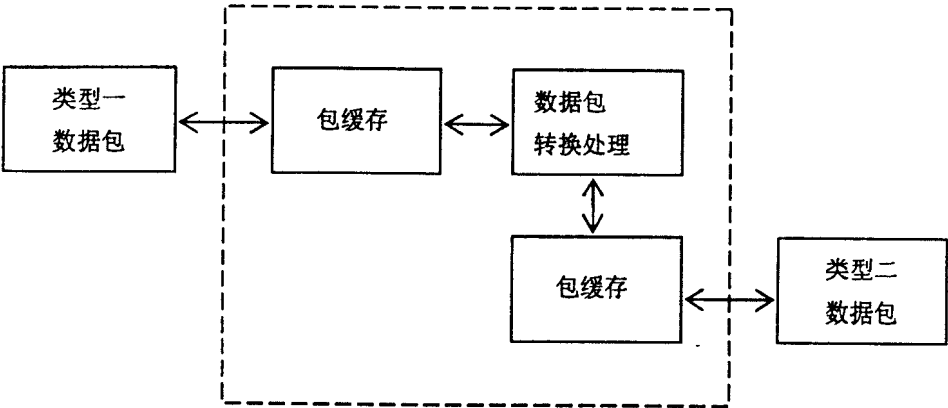


图 4