

(12) 按照专利合作条约所公布的国际申请

(19) 世界知识产权组织
国际局



(43) 国际公布日
2012年9月27日 (27.09.2012)

(10) 国际公布号
WO 2012/126352 A1

- (51) 国际专利分类号:
H04L 12/56 (2006.01) H04L 12/40 (2006.01)
- (21) 国际申请号: PCT/CN2012/072608
- (22) 国际申请日: 2012年3月20日 (20.03.2012)
- (25) 申请语言: 中文
- (26) 公布语言: 中文
- (30) 优先权:
201110070778.6 2011年3月23日 (23.03.2011) CN
- (71) 申请人 (对除美国外的所有指定国): 成都市华为赛门铁克科技有限公司 (CHENGDU HUAWEI SYMANTEC TECHNOLOGIES CO.,LTD.) [CN/CN]; 中国四川省成都市高新西区清水河片区电子科大高新科技园天辰路88号, Sichuan 611731 (CN)。
- (72) 发明人; 及
- (75) 发明人/申请人 (仅对美国): 孟健 (MENG, Jian) [CN/CN]; 中国四川省成都市高新西区清水河片区电子科大高新科技园天辰路88号, Sichuan 611731 (CN)。 王雨晨 (WANG, Yuchen) [CN/CN]; 中国四川省成都市高新西区清水河片区电子科大高新科技园天辰路88号, Sichuan 611731 (CN)。 刘利锋 (LIU, Lifeng) [CN/CN]; 中国四川省成都市高新西区

清水河片区电子科大高新科技园天辰路88号, Sichuan 611731 (CN)。

- (81) 指定国 (除另有指明, 要求每一种可提供的国家保护): AE, AG, AL, AM, AO, AT, AU, AZ, BA, BB, BG, BH, BR, BW, BY, BZ, CA, CH, CL, CN, CO, CR, CU, CZ, DE, DK, DM, DO, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, GT, HN, HR, HU, ID, IL, IN, IS, JP, KE, KG, KM, KN, KP, KR, KZ, LA, LC, LK, LR, LS, LT, LU, LY, MA, MD, ME, MG, MK, MN, MW, MX, MY, MZ, NA, NG, NI, NO, NZ, OM, PE, PG, PH, PL, PT, QA, RO, RS, RU, RW, SC, SD, SE, SG, SK, SL, SM, ST, SV, SY, TH, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, ZA, ZM, ZW。
- (84) 指定国 (除另有指明, 要求每一种可提供的地区保护): ARIPO (BW, GH, GM, KE, LR, LS, MW, MZ, NA, RW, SD, SL, SZ, TZ, UG, ZM, ZW), 欧亚 (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), 欧洲 (AL, AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HR, HU, IE, IS, IT, LT, LU, LV, MC, MK, MT, NL, NO, PL, PT, RO, RS, SE, SI, SK, SM, TR), OAPI (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, ML, MR, NE, SN, TD, TG)。

本国际公布:

- 包括国际检索报告(条约第21条(3))。

(54) Title: METHOD, DEVICE AND SYSTEM FOR TRANSMITTING MESSAGES ON PCIE BUS

(54) 发明名称: 在PCIE总线上传输报文的方法、设备和系统

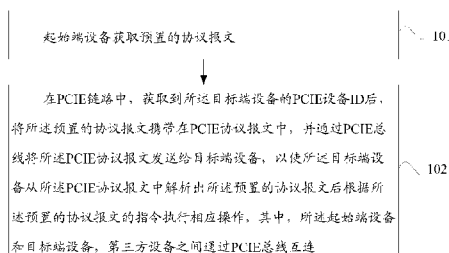


图1 / Fig. 1

101 A START-END DEVICE OBTAINS A PRESET PROTOCOL MESSAGE
 102 IN A PCIE LINK, AFTER OBTAINING THE ID OF THE PCIE DEVICE
 OF THE TARGET-END DEVICE, CARRY THE PRESET PROTOCOL
 MESSAGE IN A PCIE PROTOCOL MESSAGE, AND SEND THE PCIE
 PROTOCOL MESSAGE TO THE TARGET-END DEVICE VIA THE PCIE
 BUS, SO THAT THE TARGET-END DEVICE PARSES THE PRESET
 PROTOCOL MESSAGE FROM THE PCIE PROTOCOL MESSAGE AND
 THEN EXECUTES A CORRESPONDING OPERATION ACCORDING TO
 AN INSTRUCTION IN THE PRESET PROTOCOL MESSAGE; THE
 START-END DEVICE, THE TARGET-END DEVICE AND THE THIRD
 PARTY DEVICE ARE INTERCONNECTED VIA THE PCIE BUS

(57) Abstract: Provided are a method, device and system for transmitting messages on a PCIE bus. The method includes: a protocol message preset by a start-end device processes a preset protocol message and obtains a preset protocol message; after obtaining the ID of the PCIE device of a target-end device, carries the preset protocol message in a PCIE protocol message, and sends the PCIE protocol message to the target-end device via a PCIE bus, so that the target-end device parses the preset protocol message from the PCIE protocol message and then executes a corresponding operation according to an instruction in the preset protocol message; the start-end device, the target-end device and the third party device are interconnected via the PCIE bus. The solution provided in the embodiments of the present invention has high bandwidth for data transmission, quick processing speed, good universality and low cost.

[见续页]



WO 2012/126352 A1



(57) 摘要:

本发明实施例提供一种在 PCIE 总线上传输报文的方法、装置和系统。方法包括：起始端设备预置的协议报文对预置的协议报文进行处理，得到预置的协议报文；获取到所述目标端设备的 PCIE 设备 ID 后，将所述预置的协议报文携带在 PCIE 协议报文中，并通过 PCIE 总线将所述 PCIE 协议报文发送给目标端设备，以使所述目标端设备从所述 PCIE 协议报文中解析出所述预置的协议报文后根据所述预置的协议报文的指令执行相应操作，其中，所述起始端设备和目标端设备，第三方设备之间通过 PCIE 总线互连。本发明实施例提供的方案，数据传输的带宽高、处理速度快，通用性好，成本低。

在PCIE总线上传输报文的方法、设备和系统

本申请要求于 2011 年 3 月 23 日提交中国专利局、申请号为 201110070778.6、发明名称为“在 PCIE 总线上传输报文的方法、设备和系统”的中国专利申请的优先权，其全部内容通过引用结合在本申请中。

5

技术领域

本发明实施例涉及通信技术领域，尤其涉及一种在外部组件高速互连（Peripheral Component Interconnect Express，简称为：PCIE）总线上传输报文的方法、设备和系统。

10

背景技术

PCIE 是最新的总线和接口标准，这个新标准将全面取代现行的外部组件互连（Peripheral Component Interconnection，PCI）标准和加速图形接口（Accelerated Graphics Port，AGP），最终实现总线标准的统一。PCIE 的主要优势是数据传输速率高，目前最高的 16X 2.0 版本可达到 16GB/s，而且还有相当大的发展潜力。PCIE 作为主板级的交互协议，完成应用程序、CPU 和 PCIE 设备（如 NIC，SCSI HBA，FC-HBA 等）间的通信。同时，也可以作为系统级的交互协议。

光纤通道（Fibre Channel，FC）协议通常用于存储区域网络上的服务器和存储设备之间的通信，光纤通道基于 FC 的链路控制和物理层协议，光纤通道的主要特性有热插拔性、高速带宽、远程连接、连接设备数量大等，一般的，业务采用的介质是光纤线缆和铜缆等，然而，FC 设备价格昂贵，但采用铜轴电缆的光纤通道有着铜媒介一样的问题，如传输距离短（30 米，取决于具体的线缆）以及易受电磁干扰（Electromagnetic Interference，EMI）影响等，且目前光纤通道刚开始流行的 8Gb/s 的速率与 PCIE 的最大速率相比有很大差距，不能获得更好的数据传输率。

发明内容

本发明实施例提供一种在 PCIE 总线上传输报文的方法、设备和系统，用以实现在 PCIE 总线上直接传输预置的协议报文；

5 本发明实施例的在外部组件高速互联 PCIE 总线上传输报文的方法，包括：

起始端设备获取预置的协议报文；

在 PCIE 链路中，获取到所述目标端设备的 PCIE 设备 ID 后，将所述预置的协议报文携带在 PCIE 协议报文中，并通过 PCIE 总线将所述 PCIE 协议报
10 文发送给目标端设备，以使所述目标端设备从所述 PCIE 协议报文中解析出所述预置的协议报文后根据所述预置的协议报文的指令执行相应操作，其中，所述起始端设备和目标端设备，第三方设备之间通过 PCIE 总线互连。

本发明实施例的在外部组件高速互联 PCIE 总线上传输报文的装置，其特征在于，包括：

15 封装模块，用于对光纤通道预置的协议报文进行封装，得到预置的协议报文；

收发模块，用于将预置的协议报文携带在 PCIE 协议报文中，并通过 PCIE 总线将所述 PCIE 协议报文发送给目标端设备，以使目标端设备从 PCIE 协议报文中解析出预置的协议报文后根据预置的协议报文的指令执行相应操作；

20 解析模块，用于从 PCIE 协议报文中解析出预置的协议报文；

执行模块，用于根据预置的协议报文的指令执行相应操作。

本发明实施例提供的在 PCIE 总线上传输报文的方法、设备和系统，通过 PCIE 总线使得使用 FC 协议的存储系统进行互连，使得通过 PCIE 总线建立彼此连接的任意设备之间均可以通过预置的协议报文进行相互的操作，可以利用 FC 协议已有的功能优势和管理工具，且数据传输的带宽高、处理速度快，
25 可用于全 PCIE 连接的多主机集群计算系统中，计算节点与存储节点基于 FC 协议实现互连。

附图说明

为了更清楚地说明本发明实施例或现有技术中的技术方案，下面将对实施例或现有技术描述中所需要使用的附图作一简单地介绍，显而易见地，下面描述中的附图是本发明的一些实施例，对于本领域普通技术人员来讲，在不付出创造性劳动的前提下，还可以根据这些附图获得其他的附图。

图 1 为本发明一个实施例提供的在 PCIE 总线上传输报文的方法流程图；

图 2 为本发明一个实施例提供的在 PCIE 总线上传输报文的方法流程图；

图 3 为本发明一个实施例提供的在 PCIE 总线上传输报文的装置结构示意图；

图 4 为本发明一个实施例提供的在 PCIE 总线上传输报文的装置结构示意图；

图 5 为本发明一个实施例提供的在 PCIE 总线上传输报文的系统结构示意图；

图 6 为本发明一个实施例提供的起始端设备的结构示意图；

图 7 为本发明一个实施例提供的第三方设备的结构示意图；

图 8 为本发明一个实施例提供的目标端设备的结构示意图。

具体实施方式

为使本发明实施例的目的、技术方案和优点更加清楚，下面将结合本发明实施例中的附图，对本发明实施例中的技术方案进行清楚、完整地描述，显然，所描述的实施例是本发明一部分实施例，而不是全部的实施例。基于本发明中的实施例，本领域普通技术人员在没有作出创造性劳动前提下所获得的所有其他实施例，都属于本发明保护的范围。

图 1 为本发明一个实施例提供的在 PCIE 总线上传输报文的方法流程图，该方法应用于由通过 PCIE 总线相互连接的设备组成的网络中，所述设备包括起始端设备、目标端设备和第三方设备。

该方法通过将各设备之间交互的一些预定协议的报文，封装在 PCIe 协议报文中，通过 PCIe 总线进行传输。从而提高数据传输带宽和处理速度。下面将以注册过程、或建立虚拟链路过程为例，对本发明实施例提供的 PCIe 总线上传输报文的方法进行介绍，需要说明的是，该方案也同样适用于其他场景中。

如图 1 所示，该方法包括：

步骤 101：起始端设备获取第一预置协议的报文；

本发明实施例可以根据具体的业务类型将预置的协议报文携带在所述 PCIe 协议报文的处理层协议 TLP 包中，获得 FC-TLP 协议报文，其中，FC-TLP 协议报文为 FC-TLP 帧格式，所述预置的协议报文为 FC 协议报文；

也就是说，起始端设备将第一预置协议的报文封装在 PCIe 报文中的 TLP 载荷部分中。

步骤 102：起始端设备获取到所述目标端设备的 PCIe 设备 ID 后，将所述第一预置协议的报文携带在 PCIe 协议报文中，并通过 PCIe 总线将所述 PCIe 协议报文发送给目标端设备，以使所述目标端设备从所述 PCIe 协议报文中解析出所述预置的协议报文后根据所述预置的协议报文的指令执行相应操作，其中，所述起始端设备和目标端设备，第三方设备之间通过 PCIe 总线互连。

其中，本实施例中的第三方设备可以是链路中的特定的网络管理设备。

其中，本实施例中的预置的协议报文为 FC 协议报文或 FCoPIP 协议报文，具体地，本实施例中第一预置协议是指 FC-TLP，第二预置协议是指 FCoPIP；

本实施例中的“携带”也在本领域内也可以称为“封装”，实际上指的是将 FCoPIP 协议报文或 FC 协议报文映射为符合 PCIe 协议报文的格式的一种做法。

其中，可以将预置的协议报文携带在 PCIe 的处理层协议（Transaction Layer Protocol，简称为：TLP）包中。预置的协议报文在 PCIe 协议报文中具体的位置可以如表 1 中所示本发明实施例提供的 PCIe 协议报文的格式示意图。

表 1

帧开始	DLLP 头	TLP 头	预置的协议报文	ECRC	LCRC	帧结束
-----	--------	-------	---------	------	------	-----

其中各字段的定义如下：

帧开始标识表示一个数据帧的开始；

5 DLLP 头，用于链路管理功能，包括物理层的流量控制，主要职责是确保数据包可靠、正确传输，确保数据包的完整性；

TLP 头，用于处理上层应用的读写请求等事务；

TLP 包就是本发明实施例中所描述的 Ethernet-TLP 包；

10 ECRC，指端到端的循环冗余校验码（CRC，Cyclic Redundancy Check），其中，CRC 是数据通信领域中最常用的一种差错校验码；

LCRC，指数据链路层的 CRC，是基于序列标识（ID，IDentity）和 TLP 全部字节的 CRC 字段；

帧结束标识表示一个数据帧的结束。

15 以上 PCIE 数据包中的各个组成部分具体可参见 PCIE 协议，在此不再赘述。

具体来说，起始端设备为了与目标端设备直接进行通信，需要建立起始端设备和目标端设备之前的光纤通道虚拟链路。在此过程中：

20 起始端设备将第一预置协议的注册名称服务器请求信息携带在 PCIE 协议报文中，并通过 PCIE 总线发送给第三方设备，接收第三方设备返回的携带在 PCIE 协议报文中的第一预置协议的第三请求响应信息，所述第三请求响应信息中携带有目标端设备的 PCIE 设备 ID；

25 起始端设备根据从第三请求响应信息中获取到的所述目标端设备的 PCIE 设备 ID，将所述第一预置协议的端口注册请求信息携带在 PCIE 协议报文中，并通过 PCIE 总线发送给目标端设备，并接收目标端设备返回的携带在 PCIE 协议报文中的第一预置协议的第四请求响应信息，从而建立与目标端设备之

间的光纤通道虚拟链路。

在附图 1 所示的起始端设备与目标端设备建立光纤通道虚拟链路的过程之前，还可以采用如下方式进行 PCIE 网络中的注册过程，起始端设备的注册过程具体如下：

起始端设备将第二预置协议的设备发现请求携带在 PCIE 协议报文中，并通过 PCIE 总线发送，接收第三方设备对应返回的携带在 PCIE 协议报文中的第二预置协议的第一请求响应信息，所述第一请求响应信息中携带有第三方设备的 PCIE 设备 ID；

10 起始端设备根据第三方设备的 PCIE 设备 ID，将第二预置协议的虚拟链路实例化请求信息携带在 PCIE 协议报文中，并通过 PCIE 总线发送给第三方设备，所述虚拟链路实例化请求信息携带起始端设备的 PCIE 设备 ID；接收第三方设备返回的携带在 PCIE 协议报文中的第二预置协议的第二请求响应信息，从而在第三方设备中对起始端设备进行注册。

15 类似地，目标端设备也可以采用类似的注册过程，具体如下：

目标端设备将第二预置协议的设备发现请求携带在 PCIE 协议报文中，并通过 PCIE 总线发送，接收第三方设备对应返回的携带在 PCIE 协议报文中的第二预置协议的第一请求响应信息，所述第一请求响应信息中携带有第三方设备的 PCIE 设备 ID；

20 目标端设备根据第三方设备的 PCIE 设备 ID，将第二预置协议的虚拟链路实例化请求信息携带在 PCIE 协议报文中，并通过 PCIE 总线发送给第三方设备，所述虚拟链路实例化请求信息携带目标端设备的 PCIE 设备 ID；接收第三方设备返回的携带在 PCIE 协议报文中的第二预置协议的第二请求响应信息，从而在第三方设备中对目标端设备进行注册。

25 上述将第二预置协议的报文携带在 PCIE 报文中，具体为：

将所述预置的协议报文携带在所述 PCIE 协议报文的处理层协议 TLP 包中，获得 FCoPIP-TLP 协议报文，FCoPIP-TLP 协议报文为 FCoPIP-TLP 帧格式，其中，所述预置的协议报文为 FCoPIP(FC over PCIE initialization protocol)

协议报文。也就是说将所述 FCoPIP-TLP 报文的设备发现请求封装在 PCIE 报文中的 TLP 载荷部分，或者将所述 FCoPIP-TLP 报文的虚拟链路实例化请求信息封装在 PCIE 报文中的 TLP 载荷部分。

需要说明的是，除了附图 1 所示的场景，其中，有多种情况可以使得起始端设备将预置的协议报文进行封装后携带在 PCIE 协议报文中，例如，计算机应用程序或者 CPU 希望通过 PCIE 总线对某一设备进行读写操作，则会将读写请求转换为 FC 读写命令并对其进行封装，然后携带在 PCIE 协议报文中。或者，某一通过 PCIE 总线与计算机相连接的设备接收到计算机通过 PCIE 总线发送的请求报文后需要返回请求响应报文时，该设备会将请求响应报文转换为本发明实施例的协议报文并对其进行封装，然后携带在 PCIE 协议报文中。

本发明实施例提供的在 PCIE 总线上传输报文的方法，使得通过 PCIE 总线建立彼此连接的任意设备之间均可以通过预置的协议报文进行相互的操作，且数据传输的带宽高、处理速度快，可用于全 PCIE 连接的多主机集群计算系统中，计算节点与存储节点基于 FC 协议实现互连。

图 2 是本发明实施例中预置的协议报文进行封装的方法流程图；

本实施例中，获取到所述目标端设备的 PCIE 设备 ID 之前，所述将预置的协议报文携带在 PCIE 协议报文中包括：

首先，步骤 1010，起始端设备，和/或，目标端设备发送设备发现请求，用以获取第三方设备的 PCIE 设备 ID，例如，设备发现请求为 FCoPIP-TLP 报文，设备发现请求对应的 FCoPIP-TLP 报文中 Encapsulated FCoPIP operation 部分可以包含的一些字段的取值如表 1 所示。

表 1

FCoPIP Protocol Code	Subcode	Operation	Contents of Operation
0001h (设备发现)	01h	Discovery Solicitation	PCIE Device ID FC WWN MAX Size
	02h	Discovery Advertisement	Priority

			PCIE Device ID FC WWN MAX Size 等
--	--	--	--

其中，FCoPIP Protocol Code 为 0001h 表示设备发现，Subcode 为 01h 时，表示所述业务类型 Operation 为 Discovery Solicitation，所述业务内容 Contents of Operation 包括 PCIE Device ID，FC 全球唯一名字（world wide name，WWN），MAX Size 等信息；Subcode 为 02h 时，表示业务类型为 Discovery Advertisement，所述业务内容 Contents of Operation 为 Priority（优先级），可
5 选地，此时业务内容还包括 PCIE Device ID，FC WWN，MAX Size 等。其中，MAX Size，表示起始端设备可以接收的报文大小；Priority，表示第三方设备的优先级标识，在 PCIE 链路上也有可能存在多个第三方设备，此时会有多个第三方设备返回第一请求响应消息，这时可以根据第一请求响应信息中的优
10 先级作为参考，从而优选采用第一请求响应信息中优先级标识较高的第三方设备。

具体的，起始端设备需要在 PCIE 链路中，发送设备发现请求信息，其中，所述设备发现请求信息中携带有起始端设备和目标端设备之间通信需要的光纤通道协议参数；本实施例中，可以在 PCIE 链路上通过广播或组播的方式发
15 送设备发现请求信息，并接收第三方设备返回的第一请求响应信息。

步骤 1011，起始端设备，和/或，目标端设备接收所述第三方设备返回的第一请求响应信息，其中，所述第一请求响应信息中携带有所述第三方设备的 PCIE 设备 ID。需要说明的是，本实施例中第一请求响应信息可以是上述的 Discovery Solicitation，表示设备发现请求；
20

当然，本发明实施例也可以通过其他方式获取第三方设备的 PCIE 设备 ID，例如，在 FCoPIP-TLP 协议报文中，通过封装 Subcode 为 02h 的字段，表示 Discovery Advertisement，此时，第三方设备主动向 PCIE 链路上的其他设备公布自己的 PCIE 设备 ID。

需要说明的是，在上述步骤 1010 中起始端设备或目标端设备发送设备发
25 现请求信息时、以及步骤 1011 中第三方设备发送第一请求响应信息时，是通

过如下方式进行发送的:

将 FCoPIP-TLP 报文格式的设备发现请求信息或第一请求响应信息,封装在 PCIE 报文中的 TLP 载荷部分中,然后将封装得到的 PCIE 报文发送给对端设备。

- 5 步骤 1012, 起始端设备, 和/或, 目标端设备获取到第三方设备的 PCIE 设备 ID 后, 根据所述第三方设备的 PCIE 设备 ID 向第三方设备发送虚拟链路实例化请求信息, 并接收 PCIE 链路上第三方设备返回的第二请求响应信息。

其中, 虚拟链路实例化请求信息可以是 Virtual Link Instantiation Request, 例如, 可以是将携带有 FLOGI 消息的 FCoPIP-TLP 协议报文携带在 PCIE
10 协议报文的 TLP 包中;

上述虚拟链路实例化请求信息也被称为虚拟链路初始化请求信息。

第二请求响应信息可以是 Virtual Link Instantiation Reply, 表示该第二请求响应信息是虚拟链路实例化响应, 例如, 可以是将携带有 FLOGI 响应消息的 FCoPIP-TLP 协议报文携带在 PCIE 协议报文的 TLP 包中;

- 15 其中, 本实施例中根据所述第三方设备的 PCIE 设备 ID 向第三方设备发送虚拟链路实例化请求信息, 并接收 PCIE 链路上第三方设备返回的第二请求响应信息, 具体可以包括:

起始端设备在本设备中运行查询命令, 可以查询到本设备的设备 ID, 查询到起始端设备的 PCIE 设备 ID 后, 将所述起始端设备的 PCIE 设备 ID 携
20 带在虚拟链路实例化请求信息中, 在本实施例中, 是将所述起始端设备的 PCIE 设备 ID 携带在 FCoPIP-TLP 协议报文中, 向第三方设备发送虚拟链路实例化请求信息, 并接收 PCIE 链路上第三方设备返回的第二请求响应信息;

本领域技术人员可以了解, 本实施例中也可以是目标端设备查询到第三方设备的 PCIE 设备 ID 后, 根据所述第三方设备的 PCIE 设备 ID 向第三方设
25 备发送虚拟链路实例化请求信息, 并接收 PCIE 链路上第三方设备返回的第二请求响应信息;

需要将所述目标端设备的 PCIE 设备 ID 封装在 FCoPIP-TLP 协议报文中, 向第三方设备发送虚拟链路实例化请求信息, 并接收 PCIE 链路上第三方设备

返回的第二请求响应信息。

需要说明的是，在步骤 1012 中起始端设备或目标端设备发送虚拟链路实例化请求信息时、以及第三方设备对应发送第二请求响应信息时，是通过如下方式进行发送的：

- 5 将 FCoPIP-TLP 报文格式的虚拟链路实例化请求信息或第二请求响应信息，封装在 PCIE 报文中的 TLP 载荷部分中，然后将封装得到的 PCIE 报文发送给对端设备。。

本实施例中，虚拟链路实例化请求信息是 FCoPIP-TLP 帧格式的协议报文，用来实现链路和设备的初始化，虚拟链路实例化请求信息对应的
10 FCoPIP-TLP 报文中 Encapsulated FCoPIP operation 部分可以包含的一些字段的取值如表 2 所示：

表 2

FCoPIP Protocol Code	Subcode	Operation	Contents of Operation
0002h (链路和设备实例化)	01h	Virtual Link Instantiation Request	FC FLOGI 消息 PCIE Device ID
	02h	Virtual Link Instantiation Reply	FC FLOGI 响应消息 PCIE Device ID

其中，FCoPIP Protocol Code 为 0002h 表示链路和设备实例化，Subcode
15 为 01h 时，表示业务类型为虚拟链路实例化请求 Virtual Link Instantiation Request，则所述业务内容 Contents of Operation 包括 FC FLOGI 消息和起始端设备的 PCIE 设备 ID；Subcode 为 02h 表示业务类型为 Virtual Link Instantiation Reply，即虚拟链路实例请求信息的响应，也就是本发明实施例所述的第二请求响应信息，则所述业务内容 Contents of Operation 包括 FC FLOGI
20 响应消息，例如，FC FLOGI 响应消息具体可以是对于 Fabric login 等结构注册信息的响应信息。

步骤 1013，第三方设备返回第二请求响应信息后，为发送所述虚拟链路实例化请求信息的设备（起始端设备，和/或，目标端设备）分配光纤通道端口 ID，并在第三方设备中记录分配的光纤通道端口 ID。

具体地，若发送所述虚拟链路实例化请求信息的设备为起始端设备，则记录为所述起始端设备分配的光纤通道端口 ID，可选地，还可以记录所述起始端设备的光纤通道协议参数等其他信息。若发送所述虚拟链路实例化请求信息的设备为目标端设备，则记录为所述目标端设备分配的光纤通道端口 ID，
 5 可选地，还可以记录所述目标端设备的光纤通道协议参数等其他信息。所述光纤通道端口 ID 是在路由配置中确定的与设备唯一对应的地址标识。在上述步骤 1010~1013 中各交互双方发送报文时，各请求信息或响应信息对应的 FCoPIP-TLP 报文的格式如表 3 所示。

如表 3 中所示本发明实施例提供的 FCoPIP-TLP 协议报文的格式示意图。

10

表 3

Byte	0								1								2								3							
	0	1	2	3	4	5	6	7	0	1	2	3	4	5	6	7	0	1	2	3	4	5	6	7	0	1	2	3	4	5	6	7
0	Type=FCoPIP_TYPE								Version				Reserved																			
1	Encapsulated FCoPIP operation (n words)																															
2																																
3																																
4																																
N																																

其中，Type 域表示预置的协议报文的数据帧的类型，如 FCoPIP_TYPE 表示该协议报文是 FCoPIP-TLP 帧格式的协议报文。Version 表示版本，Reserved 表示预留。Encapsulated FC Frame (n words)表示封装的具体的内容，
 15 Type=FCoPIP_TYPE 时，表示该预置的协议报文类型是 FCoPIP-TLP 协议报文。

表 3 所示的本发明实施例提供的 FCoPIP-TLP 协议报文中 Encapsulated FCoPIP operation 部分所包括的具体的字段、以及字段的排列方式如表 3.1 所示，其中的 FCoPIP Protocol Code，例如，可以参考表 2 中所示，如果取值为
 20 0002h，表示链路和设备实例化；Contents of Operation 可以是表 2 中的 FC FLOGI 消息和 PCIE Device ID 等。

表 3.1

Byte	0								1								2								3							
	0	1	2	3	4	5	6	7	0	1	2	3	4	5	6	7	0	1	2	3	4	5	6	7	0	1	2	3	4	5	6	7
0	FCoPIP Protocol Code																Reserved								Subcode							
1	Contents Length																Reserved															
2	Contents of Operation																															
12																																
16																																
N																																

步骤 1014, 起始端设备发送注册名称服务器请求信息至第三方设备, 以获取目标端设备的 PCIE 设备 ID;

具体地, 起始端设备向第三方设备发送注册名称服务器请求, 用以查询目标端设备的 PCIE 设备 ID, 所述注册名称服务器请求中携带目标端设备的全球唯一名字 (world wide name, WWN)。

起始端设备接收第三方设备对应返回的第三请求响应信息 (注册名称服务器响应), 其中, 第三请求响应信息中携带有所述目标端设备的 PCIE 设备 ID。第三方设备中所存储的各起始端设备和目标端设备的 PCIE 设备 ID, 是第三方设备从每个起始端设备或目标端设备发送的虚拟链路实例化请求信息中得到的, 请参考表 2。

上述注册名称服务器请求和第三请求响应信息为 FC-TLP 报文, 被封装在 PCIE 报文中的 TLP 载荷部分后, 再将封装得到的 PCIE 报文发送给对端设备。

步骤 1015, 起始端设备利用获得的目标端设备的 PCIE 设备 ID, 向目标端设备发送端口注册请求信息, 例如, 可以是发送 PLOGI 信息, 接收到目标端设备返回的第四请求响应信息后, 起始端设备和目标端设备端完成了光纤通道虚拟链路的建立, 并通过该光纤通道虚拟链路进行数据的通信。其中, 第四请求响应信息也可以被看作是光纤通道虚拟链路的建立完成响应信息。其中, 在协议报文传输过程中, 所述端口注册请求信息具体为 FC-TLP 协议报文。

需要说明的是, 在上述步骤 1014 中起始端设备向第三方设备发送注册名称服务器请求、第三方设备对应返回第三请求响应信息, 以及步骤 1015 中起

始端设备向目标端设备发送端口注册请求信息、目标端设备对应返回第四请求响应信息时，是通过如下方式进行发送的：

将 FC-TLP 报文格式的注册名称服务器请求、第三请求响应信息、端口注册请求信息或第四请求响应信息，封装在 PCIE 报文中的 TLP 载荷部分中；

5 再将所述 PCIE 报文发送给对端设备。

如表 4 所示，为一个 FC-TLP 协议报文格式中关于各个字段的定义：

表 4

Byte	0								1								2								3							
	0	1	2	3	4	5	6	7	0	1	2	3	4	5	6	7	0	1	2	3	4	5	6	7	0	1	2	3	4	5	6	7
	0	1	2	3	4	5	6	7	0	1	2	3	4	5	6	7	0	1	2	3	4	5	6	7	0	1	2	3	4	5	6	7
0	Type=FCoP_TYPE																Version				Reserved											
1	Encapsulated FC Frame (n words)																															
2																																
3																																
4																																
N																																

其中，Type 域表示封装的数据帧的类型，FCoP_TYPE 表示该协议报文是 FC-TLP 帧格式的协议报文。Version 表示版本，Reserved 表示预留。Encapsulated FC Frame (n words)表示封装的具体内容。

可选地，通过上述步骤 1010~1015，起始端设备和目标端设备端完成了起始端设备和目标端设备之间光纤通道虚拟链路的建立过程，此后，还可以在所述光纤通道虚拟链路上，发送链路维护信息，具体地：

15 其中上述步骤中，起始端设备接收到目标端设备返回的第四请求响应信息后，起始端设备和/或目标端设备确定当所述链路维护周期到达时，可以发送链路维护信息。链路维护信息中包括起始端设备和目标端设备各自的 PCIE 设备 ID 和光纤通道端口 ID。

需要说明的是，与步骤 1010~1012 中的发现请求信息、以及虚拟链路实例化请求信息相类似，链路维护信息也是 FCoPIP-TLP 报文格式的，是通过如下方式进行发送的：

将 FCoPIP-TLP 报文格式的链路维护信息, 封装在 PCIE 报文中的 TLP 载荷部分中, 然后将封装得到的 PCIE 报文发送给对端设备。

具体示例如下: 其中, 表 5 是起始端设备和目标端设备进行链路维护时发送的 FCoPIP-TLP 报文中 Encapsulated FCoPIP operation 部分可以包含的一些字段的取值。具体各字段的排列方式请参照表 3.1, 在这里不再重复。

表 5

FCoPIP Protocol Code	Subcode	Operation	Contents of Operation
0003h (链路和设备维护)	01h	Keep Alive	FC N_Port ID PCIE Device ID
	02h	Clear Virtual Links	FC N_Port ID PCIE Device ID FC WWN

本发明实施例中, 起始端设备和目标端设备可以在 PCIE 链路上发送链路维护信息, 以便于在 PCIE 链路上的其他设备, 可以获知到该起始端设备和目标端设备的存活状态。

本实施例中, 例如, FCoPIP-TLP 报文格式的链路维护信息中 FCoPIP Protocol Code 为 0003h, 其中, 如上表所示, Subcode 为 01h 时, 表示业务类型为 Keep Alive; 业务内容中携带有 FC N_Port ID 以及 PCIE Device ID; 其中, Subcode 为 02h 时, 表示业务类型为 Clear Virtual Links, 业务内容中携带有 FC N_Port ID, PCIE Device ID, FC WWN; 所述链路维护信息将所述已封装的 FCoPIP 携带在所述 PCIE 协议报文的处理层协议 TLP 包中。

其中, Keep Alive 用于表示发出 Keep Alive 消息的起始端设备或目标端设备还处于存活状态; Clear virtual links 用于第三方设备在 PCIE 链路上探测是否有起始端设备或目标端设备已不存在, 例如设备出现故障, 或设备功能关闭, 设备下电拔掉等, 如第三方设备确定起始端设备或目标端设备已不存在, 则通知与其建立通信链路的另一方设备该链路已不存在, 请求释放资源, 收回分配给起始端设备和/或目标端设备的光纤通道端口 ID。

图 3 为本发明一个实施例提供的在 PCIE 总线上传输报文的装置的结构示意图, 该装置为上述方法的特定执行主体, 因此方法实施例中的内容以引用

的方式带入本装置实施例中，此处不做赘述。

如图 3 所示，该设备包括：

封装模块 201，用于获取预置的协议报文；

收发模块 202，用于将所述预置的协议报文携带在 PCIE 协议报文中，并
5 通过 PCIE 总线将所述 PCIE 协议报文发送给目标端设备，以使所述目标端设备从所述 PCIE 协议报文中解析出所述预置的协议报文后根据所述预置的协议报文的指令执行相应操作；

解析模块 203，用于从接收到的 PCIE 协议报文中解析出预置的协议报文；

执行模块 204，用于根据预置的协议报文的指令执行相应操作。

10 所述封装模块 201 具体用于将预置的协议报文携带在 PCIE 协议报文的处理层协议 TLP 包中。

本发明实施例提供的在 PCIE 总线上传输报文的设备，通过将预置的协议报文携带在 PCIE 协议报文中，并通过 PCIE 总线将携带有预置的协议报文的 PCIE 协议报文传输给目标端设备，实现了在 PCIE 总线上传输预置的协议报
15 文，使得通过 PCIE 总线建立彼此连接的任意设备之间均可以通过预置的协议报文进行相互的操作，且数据传输的带宽高、处理速度快，可用于全 PCIE 连接的多主机集群计算系统中，计算节点与存储节点基于 FC 协议实现互连。

如图 4 所示，为本发明一个实施例提供的在 PCIE 总线上传输报文的装置的结构示意图：

20 封装模块 201，包括：第一封装单元 2011，用于当发送设备发现请求信息至第三方设备时，获取预置的协议报文；

本发明实施例可以根据具体的业务类型可以将所述预置的协议报文封装在所述 PCIE 协议报文的处理层协议 TLP 包中，获得 FC-TLP 协议报文；

25 或将所述预置的协议报文封装在所述 PCIE 协议报文的处理层协议 TLP 包中，获得 FCoPIP-TLP 协议报文；

所述收发模块 202，用于获取所述目标端设备的 PCIE 设备 ID，将所述目标端设备的 PCIE 设备 ID 封装在预置的协议报文中，并将预置的协议报文携带在 PCIE 协议报文中，通过 PCIE 总线将所述 PCIE 协议报文发送给目标

端设备,以使所述目标端设备从所述 PCIE 协议报文中解析出所述预置的协议报文后根据所述预置的协议报文的指令执行相应操作;

当发送设备发现请求信息时,所述收发模块 202 具体用于发送设备发现请求信息至第三方设备,接收所述第三方设备返回的第一请求响应信息,其中,所述第一请求响应中携带有所述第三方设备的 PCIE 设备 ID。

其中,本实施例中的第三方设备可以是链路中的特定的网络管理设备。

本实施例中,所述收发模块 202 发送设备发现请求信息至第三方设备可以通过起始端设备向第三方设备发送,同时,也可以是,通过目标端设备向第三方设备发送设备发现请求信息。

当所述预置的协议报文为虚拟链路实例化请求时,第一封装单元 2011,还用于当所述收发模块 202 向第三方设备发送虚拟链路实例化请求时,将所述起始端设备的 PCIE 设备 ID 封装在 FCoPIP 协议报文中,并将 FCoPIP 协议报文携带在 PCIE 协议报文中,获得 FCoPIP-TLP 协议报文。

收发模块 202 用于向第三方设备发送虚拟链路实例化请求,该请求信息可以是 Virtual Link Instantiation Request;

所述收发模块 202 发送虚拟链路实例化请求至第三方设备可以通过起始端设备向第三方设备发送,同时,还需要,通过目标端设备向第三方设备发送;收发模块 202 接收 PCIE 链路上第三方设备返回的第二请求响应信息;

其中,第二请求响应信息可以是 Virtual Link Instantiation Reply,表示虚拟链路实例化响应,所述第二请求响应信息中携带有所述起始端设备和目标端设备各自的光纤通道协议参数和所述第三方设备分配给起始端设备的光纤通道端口 ID 和目标端设备的光纤通道端口 ID。

进一步的,如果所述预置的协议报文为注册名称服务器 (PLOGINS) 请求时,该封装模块 201 还可以包括:

第二封装单元 2012,用于当发送注册名称服务器请求信息至第三方设备时,对预置的协议报文进行处理,将第三方设备的光纤通道协议参数和光纤通道端口 ID 封装在 FC 协议报文中,并将 FC 协议报文携带在 PCIE 协议报文中以获得 FC-TLP 协议报文;

所述收发模块 202 还用于发送所述注册名称服务器请求到第三方设备，并接收第三方设备返回的第三请求响应信息。其中，所述注册名称服务器请求信息为 FC-TLP 协议报文，所述光纤通道端口 ID 是在路由配置中确定的与设备唯一对应的地址标识。其中，第三请求响应信息中携带有所述目标端设备
5 的 PCIE 设备 ID。

如果发送端口注册请求信息到目标端设备，则第二封装单元 2012 还用于对预置的协议报文进行处理，将目标端设备的 PCIE 设备 ID 封装在 FC 协议报文中，并将 FC 协议报文携带在 PCIE 协议报文中以获得 FC-TLP 协议报文；

所述收发模块 202 用于发送所述端口注册请求信息到所述目标端设备，
10 接收从目标端设备返回的第四请求响应信息，从而建立从起始端设备到目标端设备的光纤通道虚拟链路。

所述设备还包括：

配置模块 401，用于响应接收到的预置的协议报文，为所述起始端设备和目标端设备分别配置各自的光纤通道端口 ID，将所述光纤通道端口 ID 携带于
15 第一请求响应信息中返回至起始端设备和目标端设备；

记录模块 402，用于记录所分配的光纤通道端口 ID 和 PCIE 设备 ID，以及光纤通道协议参数等相关信息，其中，记录的光纤通道协议参数包括起始端设备的光纤通道端口 ID 和目标端设备的光纤通道端口 ID，以及起始端设备和目标端设备各自的光纤通道协议参数。其中，所述光纤通道端口 ID 是在路
20 由配置中确定的与设备唯一对应的地址标识。

本发明实施例提供的在 PCIE 总线上传输报文的设备，通过 PCIE 总线使得使用 FC 协议的存储系统进行互连，使得通过 PCIE 总线建立彼此连接的任意设备之间均可以通过预置的协议报文进行相互的操作，可以利用 FC 协议已有的功能优势和管理工具，且数据传输的带宽高、处理速度快，可用于全 PCIE
25 连接的多主机集群计算系统中，计算节点与存储节点基于 FC 协议实现互连。

图 5 为本发明一个实施例提供的在 PCIE 总线上传输报文的系统结构示意图，如图 5 所示，

该系统包括:

起始端设备 20、目标端设备 30、以及第三方设备 40, 其中, 起始端设备 20、目标端设备 30、以及第三方设备 40 通过 PCIE 总线连接并进行数据交互。

图 6 为本发明实施例提供的起始端设备的结构示意图, 起始端设备 20 包
5 括:

第一发送单元 601, 用于将第一预置协议的注册名称服务器请求信息携带在 PCIE 协议报文中, 并通过 PCIE 总线发送给第三方设备;

第一接收单元 602, 用于接收第三方设备返回的携带在 PCIE 协议报文中的第一预置协议的第三请求响应信息, 第三请求响应信息中携带有目标端设
10 备的 PCIE 设备 ID;

第一发送单元 601, 还用于根据从第三请求响应信息中获取到所述目标端设备的 PCIE 设备 ID, 将所述第一预置协议的端口注册请求信息携带在 PCIE 协议报文中, 并通过 PCIE 总线发送给目标端设备;

第一接收单元 602, 还用于接收目标端设备返回的携带在 PCIE 协议报文中
15 的第一预置协议的第四请求响应信息, 从而建立光纤通道虚拟链路。

可选地, 起始端设备 20 还包括:

第二发送单元 603, 用于在第一发送单元 601 将第一预置协议的注册名称服务器请求信息携带在 PCIE 协议报文中, 并通过 PCIE 总线发送给第三方设备之前, 将第二预置协议的设备发现请求携带在 PCIE 协议报文中, 并通过
20 PCIE 总线发送;

第二接收单元 604, 用于接收第三方设备对应返回的携带在 PCIE 协议报文中的第二预置协议的第一请求响应信息, 所述第一请求响应信息中携带有第三方设备的 PCIE 设备 ID;

第二发送单元 603, 还用于根据第三方设备的 PCIE 设备 ID, 将第二预置
25 协议的虚拟链路实例化请求信息携带在 PCIE 协议报文中, 并通过 PCIE 总线发送给第三方设备, 所述虚拟链路实例化请求信息携带起始端设备的 PCIE 设备 ID;

第二接收单元 604, 还用于接收第三方设备返回的携带在 PCIE 协议报文

中的第二预置协议的第二请求响应信息，从而在第三方设备中对起始端设备进行注册。

可选地，所述第一发送单元 601，还用于在该光纤通道虚拟链路上周期性将第二预置协议的链路维护信息携带在 PCIE 协议报文中，发送给对端设备。

5

图 7 为本发明实施例提供的第三方设备的结构示意图，第三方设备 40 包括：

第三接收单元 701，用于接收起始端设备发送的携带在 PCIE 协议报文中的第一预置协议的注册名称服务器请求信息；

10 第三发送单元 702，用于将第一预置协议的第三请求信息携带在 PCIE 协议报文中，并通过 PCIE 总线发送给所述起始端设备，所述第三请求响应信息中携带有目标端设备的 PCIE 设备 ID。

可选地，第三方设备 40 还包括：

15 第四接收单元 703，用于接收起始端设备或目标端设备发送的携带在 PCIE 协议报文中的第二预置协议的设备发现请求；

第四发送单元 704，用于将第二预置协议的第一请求响应信息携带在 PCIE 协议报文中，并通过 PCIE 总线发送给所述起始端设备或目标端设备。

图 8 为本发明实施例提供的目标端设备的结构示意图，目标端设备 30 包括：

20 第五接收单元 801，用于接收起始端设备发送的携带在 PCIE 协议报文中的第一预置协议的端口注册请求信息；

第五发送单元 802，用于将第一预置协议的第四请求响应信息，携带在 PCIE 协议报文中，并通过 PCIE 总线发送给所述起始端设备，从而建立光纤通道虚拟链路。

25 可选地，目标端设备 30 还包括：

第六发送单元 803，用于在第五接收单元接收端口注册请求信息之前，将第二预置协议的设备发现请求携带在 PCIE 协议报文中，并通过 PCIE 总线发送；

第六接收单元 804，用于接收第三方设备对应返回的携带在 PCIE 协议报文中的第二预置协议的第一请求响应信息，所述第一请求响应信息中携带有第三方设备的 PCIE 设备 ID；

5 第六发送单元 803，还用于根据第三方设备的 PCIE 设备 ID，将第二预置协议的虚拟链路实例化请求信息携带在 PCIE 协议报文中，并通过 PCIE 总线发送给第三方设备，所述虚拟链路实例化请求信息携带目标端设备的 PCIE 设备 ID；

第六接收单元 804，还用于接收第三方设备返回的携带在 PCIE 协议报文中的第二预置协议的第二请求响应信息，从而在第三方设备中对目标端设备进行注册。

这里需要说明的是，起始端设备 20 和目标端设备 30 在现有技术中可以是计算机及其相关外部设备，如计算机、服务器、存储阵列等设备，但并不限制在计算机及其相关外部设备上，其他通过 PCIE 总线连接的特定设备都可以成为起始端设备 20 或者目标端设备 30。

15 本发明实施例可用于全 PCI-E 连接的多主机集群计算系统中，计算节点与存储节点基于 FC 协议实现存储网络的虚拟化。

本发明实施例提供了一种在 PCIE 总线上传输报文的系统，通过将预置的协议报文携带在 PCIE 协议报文中，并通过 PCIE 总线将封装有预置的协议报文的 PCIE 协议报文传输给目标端设备，实现了在 PCIE 总线上传输预置的协议报文，使得通过 PCIE 总线建立彼此连接的任意设备之间均可以通过预置的协议报文进行相互的操作，且数据传输的带宽高、处理速度快，提高了报文的传输速度，降低了设备的成本；同时，通过在 PCIE 总线上传递预置的协议报文，使得通过 PCIE 总线连接的两个对端设备之间可以通过预置的协议报文进行相互的操作，而不需要定义各种私有的协议进行相互的操作，通用性好。

25 另外，采用本方案的存储设备，由于 PCIE 总线沿用了 PCI 总线的原理，但抛弃了并行共享总线，采用了目前业内流行的点对点串行连接，每个设备都有自己的专用连接，不需要向整个总线请求带宽，而且可以把数据传输率提高到一个很高的频率，达到 PCI 所不能提供的高带宽，即所有设备均可通

过 PCIE 进行连接，简化了主板设计上的不必要的麻烦。

本领域普通技术人员可以理解：实现上述方法实施例的全部或部分步骤可以通过程序指令相关的硬件来完成，前述的程序可以存储于一计算机可读存储介质中，该程序在执行时，执行包括上述方法实施例的步骤；而前述的存储介质包括：ROM、RAM、磁碟或者光盘等各种可以存储程序代码的介质。

最后应说明的是：以上实施例仅用以说明本发明的技术方案，而非对其限制；尽管参照前述实施例对本发明进行了详细的说明，本领域的普通技术人员应当理解：其依然可以对前述各实施例所记载的技术方案进行修改，或者对其中部分技术特征进行等同替换；而这些修改或者替换，并不使相应技术方案

10 的本质脱离本发明各实施例技术方案的精神和范围。

权利要求

1、一种在外部组件高速互联 PCIE 总线上传输报文的方法，其特征在于，
包括：

起始端设备将第一预置协议的注册名称服务器请求信息携带在 PCIE 协
议报文中，并通过 PCIE 总线发送给第三方设备，接收第三方设备返回的携带
5 在 PCIE 协议报文中的第一预置协议的第三请求响应信息，所述第三请求响应
信息中携带有目标端设备的 PCIE 设备 ID；

起始端设备根据从第三请求响应信息中获取到的所述目标端设备的 PCIE
设备 ID，将所述第一预置协议的端口注册请求信息携带在 PCIE 协议报文中，
10 并通过 PCIE 总线发送给目标端设备，并接收目标端设备返回的携带在 PCIE
协议报文中的第一预置协议的第四请求响应信息，从而建立与目标端设备之
间的光纤通道虚拟链路，其中，所述起始端设备、目标端设备、第三方设备
之间通过 PCIE 总线互连。

2、如权利要求 1 所述的方法，其特征在于，所述第一预置协议为 FC-TLP。

15 3、如权利要求 2 所述的方法，其特征在于，所述将所述第一预置协议的
注册名称服务器请求信息携带在 PCIE 协议报文中，包括：

将 FC-TLP 协议的注册名称服务器请求信息封装在 PCIE 报文中的 TLP 载
荷部分；

所述将所述第一预置协议的端口注册请求信息携带在 PCIE 协议报文中，
20 包括：

将 FC-TLP 协议的端口注册请求信息封装在 PCIE 报文中的 TLP 载荷部
分。

4、如权利要求 1 所述的方法，其特征在于，所述将第一预置协议的注册
名称服务器请求信息携带在 PCIE 协议报文中，并通过 PCIE 总线发送给第三
25 方设备之前，还包括：

起始端设备将第二预置协议的设备发现请求携带在 PCIE 协议报文中，并
通过 PCIE 总线发送，接收第三方设备对应返回的携带在 PCIE 协议报文中的

第二预置协议的第一请求响应信息，所述第一请求响应信息中携带有第三方设备的 PCIE 设备 ID；

起始端设备根据第三方设备的 PCIE 设备 ID，将第二预置协议的虚拟链路实例化请求信息携带在 PCIE 协议报文中，并通过 PCIE 总线发送给第三方设备，所述虚拟链路实例化请求信息携带起始端设备的 PCIE 设备 ID；接收
5 第三方设备返回的携带在 PCIE 协议报文中的第二预置协议的第二请求响应信息。

5、如权利要求 1 所述的方法，其特征在于，所述起始端设备获取第一预置协议的报文之前，还包括：

10 目标端设备将第二预置协议的设备发现请求携带在 PCIE 协议报文中，并通过 PCIE 总线发送，接收第三方设备对应返回的携带在 PCIE 协议报文中的第二预置协议的第一请求响应信息，所述第一请求响应信息中携带有第三方设备的 PCIE 设备 ID；

15 目标端设备根据第三方设备的 PCIE 设备 ID，将第二预置协议的虚拟链路实例化请求信息携带在 PCIE 协议报文中，并通过 PCIE 总线发送给第三方设备，所述虚拟链路实例化请求信息携带目标端设备的 PCIE 设备 ID；接收第三方设备返回的携带在 PCIE 协议报文中的第二预置协议的第二请求响应信息，从而在第三方设备中对目标端设备进行注册。

20 6、如权利要求 5 所述的方法，其特征在于，所述建立光纤通道虚拟链路后，还包括：

起始端设备和/或目标端设备在该光纤通道虚拟链路上周期性将第二预置协议的链路维护信息携带在 PCIE 协议报文中，发送给对端设备。

7、如权利要求 4 至 6 任一所述的方法，其特征在于，所述第二预置协议为 FCoPIP-TLP。

25 8、如权利要求 7 所述的方法，其特征在于，所述起始端设备将第二预置协议的设备发现请求携带在 PCIE 协议报文中，包括：

将所述 FCoPIP-TLP 报文的设备发现请求封装在 PCIE 报文中的 TLP 载荷部分；

所述将第二预置协议的虚拟链路实例化请求信息携带在 PCIE 协议报文中，包括：

将所述 FCoPIP-TLP 报文的虚拟链路实例化请求信息封装在 PCIE 报文中的 TLP 载荷部分；

5 所述将第二预置协议的链路维护信息携带在 PCIE 协议报文中，包括：

将所述 FCoPIP-TLP 报文的链路维护信息封装在 PCIE 报文中的 TLP 载荷部分。

9、一种起始端设备，其特征在于，包括：

10 第一发送单元，用于将第一预置协议的注册名称服务器请求信息携带在 PCIE 协议报文中，并通过 PCIE 总线发送给第三方设备；

第一接收单元，用于接收第三方设备返回的携带在 PCIE 协议报文中的第一预置协议的第三请求响应信息，第三请求响应信息中携带有目标端设备的 PCIE 设备 ID；

15 第一发送单元，还用于根据从第三请求响应信息中获取到所述目标端设备的 PCIE 设备 ID，将所述第一预置协议的端口注册请求信息携带在 PCIE 协议报文中，并通过 PCIE 总线发送给目标端设备；

第一接收单元，还用于接收目标端设备返回的携带在 PCIE 协议报文中的第一预置协议的第四请求响应信息，从而建立光纤通道虚拟链路。

10、如权利要求 9 所述的设备，其特征在于，还包括：

20 第二发送单元，用于在第一发送单元将第一预置协议的注册名称服务器请求信息携带在 PCIE 协议报文中，并通过 PCIE 总线发送给第三方设备之前，将第二预置协议的设备发现请求携带在 PCIE 协议报文中，并通过 PCIE 总线发送；

25 第二接收单元，用于接收第三方设备对应返回的携带在 PCIE 协议报文中的第二预置协议的第一请求响应信息，所述第一请求响应信息中携带有第三方设备的 PCIE 设备 ID；

第二发送单元，还用于根据第三方设备的 PCIE 设备 ID，将第二预置协议的虚拟链路实例化请求信息携带在 PCIE 协议报文中，并通过 PCIE 总线发

送给第三方设备,所述虚拟链路实例化请求信息携带起始端设备的 PCIE 设备 ID;

第二接收单元,还用于接收第三方设备返回的携带在 PCIE 协议报文中的第二预置协议的第二请求响应信息,从而在第三方设备中对起始端设备进行注册。

11、如权利要求 10 所述的设备,其特征在于,所述第一发送单元,还用于在该光纤通道虚拟链路上周期性将第二预置协议的链路维护信息携带在 PCIE 协议报文中,发送给对端设备。

12、一种第三方设备,其特征在于,包括:

10 第三接收单元,用于接收起始端设备发送的携带在 PCIE 协议报文中的第一预置协议的注册名称服务器请求信息;

第三发送单元,用于将第一预置协议的第三请求信息携带在 PCIE 协议报文中,并通过 PCIE 总线发送给所述起始端设备,所述第三请求响应信息中携带有目标端设备的 PCIE 设备 ID。

15 13、如权利要求 12 所述的设备,其特征在于,还包括:

第四接收单元,用于接收起始端设备或目标端设备发送的携带在 PCIE 协议报文中的第二预置协议的设备发现请求;

第四发送单元,用于将第二预置协议的第一请求响应信息携带在 PCIE 协议报文中,并通过 PCIE 总线发送给所述起始端设备或目标端设备。

20 14、一种目标端设备,其特征在于,包括:

第五接收单元,用于接收起始端设备发送的携带在 PCIE 协议报文中的第一预置协议的端口注册请求信息;

25 第五发送单元,用于将第一预置协议的第四请求响应信息,携带在 PCIE 协议报文中,并通过 PCIE 总线发送给所述起始端设备,从而建立光纤通道虚拟链路。

15、如权利要求 14 所述的设备,其特征在于,还包括:

第六发送单元,用于在第五接收单元接收端口注册请求信息之前,将第二预置协议的设备发现请求携带在 PCIE 协议报文中,并通过 PCIE 总线发送;

第六接收单元,用于接收第三方设备对应返回的携带在 PCIE 协议报文中的第二预置协议的第一请求响应信息,所述第一请求响应信息中携带有第三方设备的 PCIE 设备 ID;

5 第六发送单元,还用于根据第三方设备的 PCIE 设备 ID,将第二预置协议的虚拟链路实例化请求信息携带在 PCIE 协议报文中,并通过 PCIE 总线发送给第三方设备,所述虚拟链路实例化请求信息携带目标端设备的 PCIE 设备 ID;

第六接收单元,还用于接收第三方设备返回的携带在 PCIE 协议报文中的第二预置协议的第二请求响应信息,从而在第三方设备中对目标端设备进行
10 注册。

16、一种在 PCIE 总线传输报文的方法,其特征在于,包括:
获取预置的协议报文;

在 PCIE 链路中,获取到所述目标端设备的 PCIE 设备 ID 后,将所述预置的协议报文携带在 PCIE 协议报文中,并通过 PCIE 总线将所述 PCIE 协议报
15 文发送给目标端设备,以使所述目标端设备从所述 PCIE 协议报文中解析出所述预置的协议报文后根据所述预置的协议报文的指令执行相应操作,其中,所述起始端设备、目标端设备、第三方设备之间通过 PCIE 总线互连。

17、根据权利要求 16 所述的方法,其特征在于,所述将预置的协议报文携带在 PCIE 协议报文中包括:

20 将预置的协议报文携带在 PCIE 协议报文的处理层协议 TLP 包中,获得 FC-TLP 协议报文,其中,所述预置的协议报文为 FC 协议报文;

或者,将预置的协议报文携带在 PCIE 协议报文的处理层协议 TLP 包中,获得 FCoPIP-TLP 协议报文,其中,所述预置的协议报文为 FCoPIP 协议报文。

18. 如权利要求 17 所述的方法,其特征在于,获取到所述目标端设备的
25 PCIE 设备 ID 之前,所述将预置的协议报文携带在 PCIE 协议报文中包括:

在 PCIE 链路中,发送设备发现请求信息,获取第三方设备的 PCIE 设备 ID,其中,设备发现请求信息中包括起始端设备的光纤通道协议参数和目标端设备的光纤通道协议参数,其中,所述设备发现请求信息为 FCoPIP-TLP 协

议报文。

19.如权利要求 18 所述的方法，其特征在于，

发送设备发现请求信息，获取第三方设备的 PCIE 设备 ID，包括：

5 通过广播或组播的方式发送设备发现请求信息于第三方设备，并接收第
三方设备返回的第一请求响应信息，第一请求响应信息中包括第三方设备
PCIE 设备 ID，或，第三方设备主动在 PCIE 链路上公布第三方设备的 PCIE
设备 ID。

20.如权利要求 16-17 中任一项所述的方法，其特征在于，所述将预置的
协议报文携带在 PCIE 协议报文中，还包括：

10 起始端设备发送注册名称服务器请求信息于第三方设备以获取目标端设
备的 PCIE 设备 ID，接收目标端设备返回的第三请求响应信息，其中，注册
名称服务器请求信息是 FC-TLP 协议报文，注册名称服务器请求信息中携带有
目标端设备的光纤通道端口 ID 及目标端设备的光纤通道协议参数，第三请求
响应信息中携带有目标端设备的 PCIE 设备 ID。

15

1/5

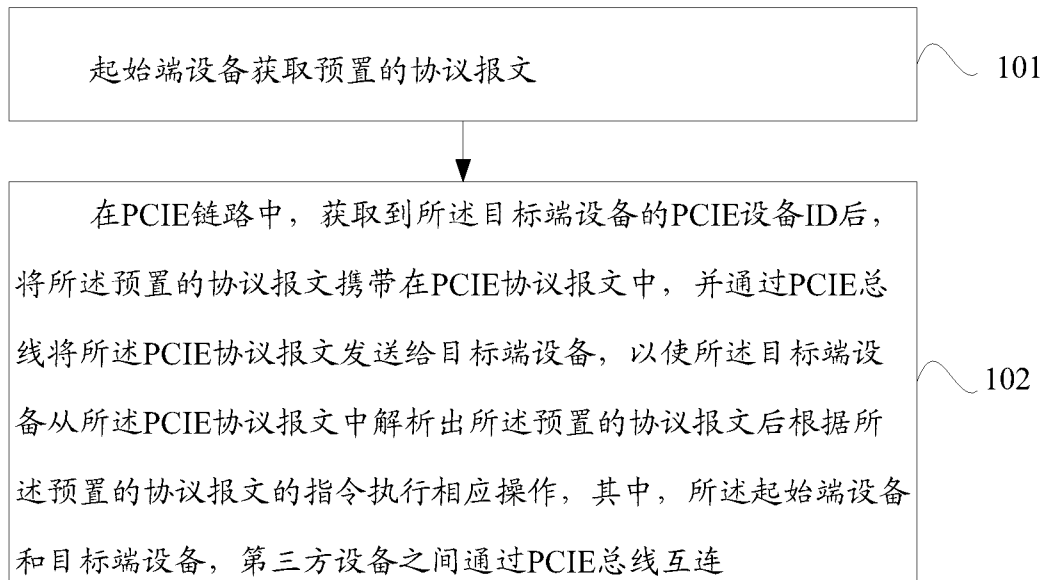


图 1

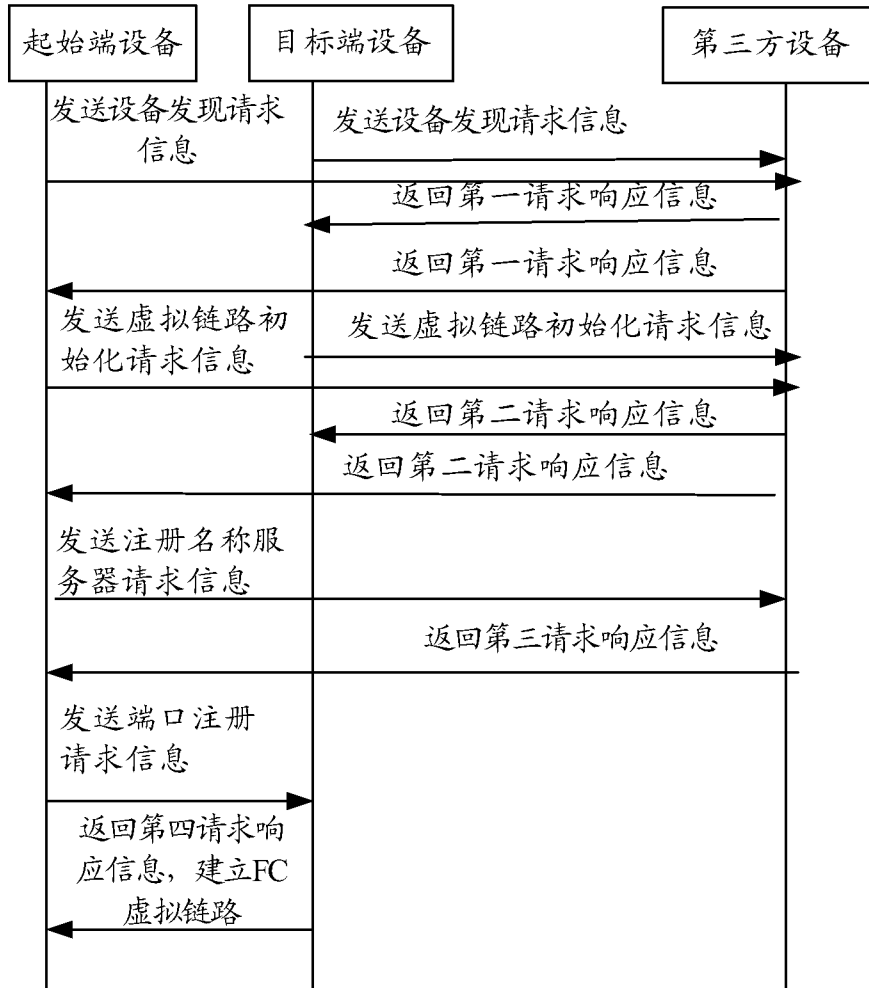


图 2

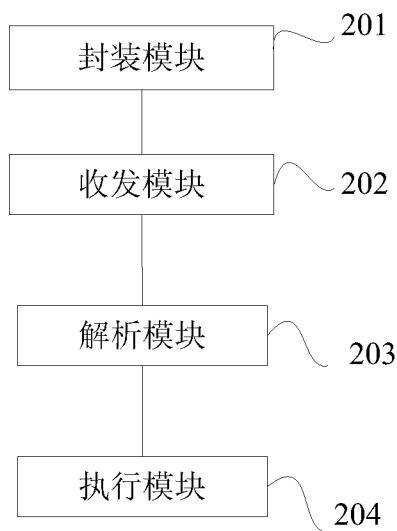


图 3

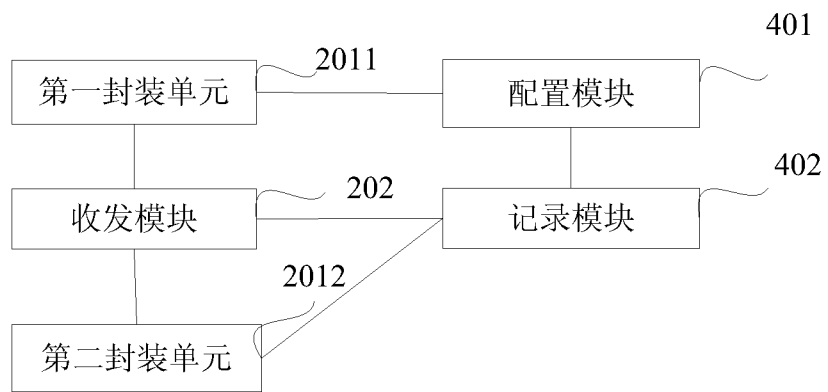


图 4

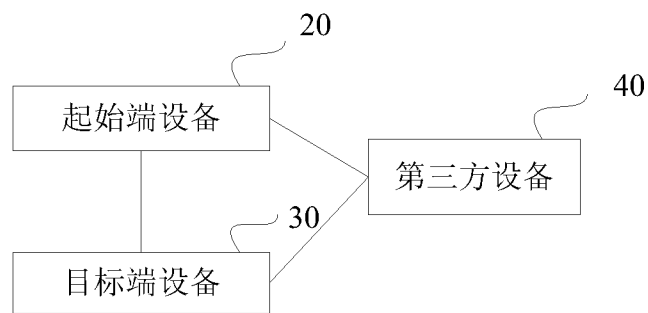


图 5

4/5

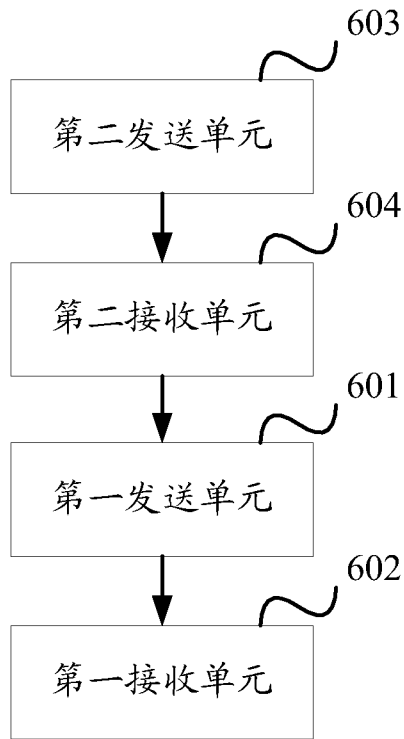


图 6

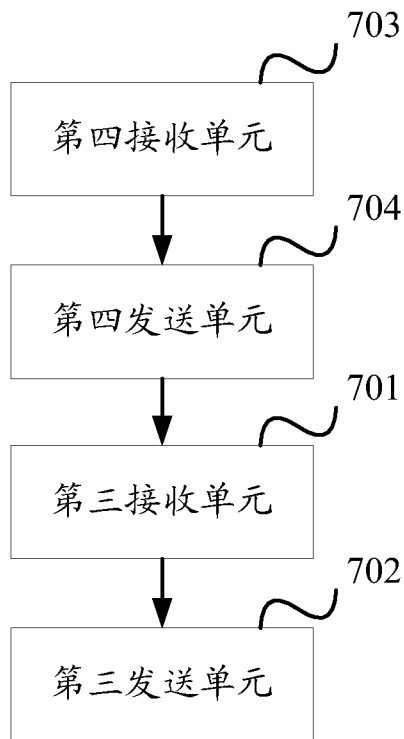


图 7

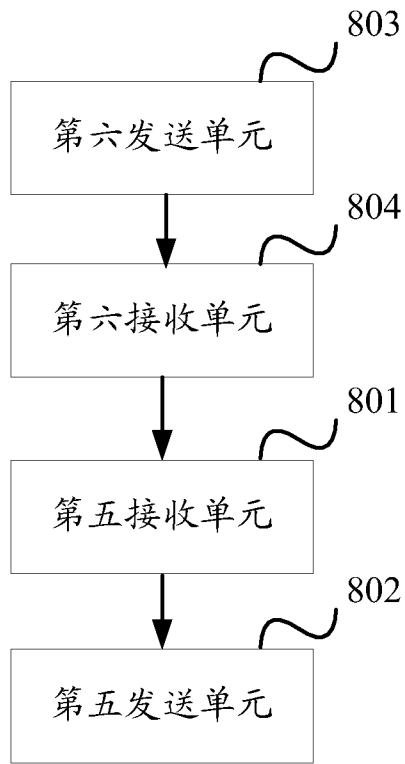


图 8

INTERNATIONAL SEARCH REPORT

International application No.

PCT/CN2012/072608

A. CLASSIFICATION OF SUBJECT MATTER

See the extra sheet

According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)

IPC: H04L

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practicable, search terms used)

CPRSABS, CNTXT, CNKI: PCIE, external bus high speed interconnection, fibre channel, FC, transmit, message, group, carry, encapsulate, device ID, virtual link, FCoPIP, FC-TLP

VEN, GOOGLE: PCIE, fibre channel, FC, transmit, send, packet, message, carry, encapsulate, device, ID, identity, virtual link, FCoPIP, FC-TLP

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
A	US 20070266179 A1 (EMULEX COMMUNICATIONS CORP.), 15 November 2007 (15.11.2007), see the whole document	1-20
A	YANG, A'feng et al., "Development of driver program for high speed data transmission adapter on PCIe port", CHINA MEASUREMENT & TESTING TECHNOLOGY, vol. 34, no. 2, 31 March 2008 (31.03.2008), see the whole document	1-20
A	CN 101588365 A (H3C TECHNOLOGIES CO., LTD.), 25 November 2009 (25.11.2009), see the whole document	1-20
A	CN 101221552 A (INTERNATIONAL BUSINESS MACHINES CORP.), 16 July 2008 (16.07.2008), see the whole document	1-20

Further documents are listed in the continuation of Box C.

See patent family annex.

<p>* Special categories of cited documents:</p> <p>"A" document defining the general state of the art which is not considered to be of particular relevance</p> <p>"E" earlier application or patent but published on or after the international filing date</p> <p>"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)</p> <p>"O" document referring to an oral disclosure, use, exhibition or other means</p> <p>"P" document published prior to the international filing date but later than the priority date claimed</p>	<p>"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention</p> <p>"X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone</p> <p>"Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art</p> <p>"&" document member of the same patent family</p>
---	---

Date of the actual completion of the international search
23 June 2012 (23.06.2012)

Date of mailing of the international search report
05 July 2012 (05.07.2012)

Name and mailing address of the ISA/CN:
State Intellectual Property Office of the P. R. China
No. 6, Xitucheng Road, Jimenqiao
Haidian District, Beijing 100088, China
Facsimile No.: (86-10) 62019451

Authorized officer
ZHENG, Jie
Telephone No.: (86-10) **62412014**

INTERNATIONAL SEARCH REPORT
Information on patent family members

International application No.

PCT/CN2012/072608

Patent Documents referred in the Report	Publication Date	Patent Family	Publication Date
US 20070266179 A1	15.11.2007	US 7743197 B2	22.06.2010
CN 101588365 A	25.11.2009	CN 101588365 B	23.05.2012
CN 101221552 A	16.07.2008	US 2008168255 A1	10.07.2008
		US 7661006 B2	09.02.2010
		CN 101221551 B	21.07.2010

INTERNATIONAL SEARCH REPORT

International application No.

PCT/CN2012/072608

CONTINUATION: A. CLASSIFICATION OF SUBJECT MATTER

According to International Patent Classification (IPC) or to both national classification and IPC

H04L 12/56 (2006.01) i

H04L 12/40 (2006.01) i

国际检索报告

国际申请号
PCT/CN2012/072608

A. 主题的分类		
参见附加页		
按照国际专利分类(IPC)或者同时按照国家分类和 IPC 两种分类		
B. 检索领域		
检索的最低限度文献(标明分类系统和分类号)		
IPC: H04L		
包含在检索领域中的除最低限度文献以外的检索文献		
在国际检索时查阅的电子数据库(数据库的名称, 和使用的检索词(如使用))		
CPRSABS, CNTXT, CNKI: PCIE, 外部总线高速互联, 光纤通道, FC, 传输, 传送, 报文, 分组, 携带, 封装, 设备 ID, 虚拟链路, FCoPIP, FC-TLP		
VEN, GOOGLE: PCIE, fibre channel, FC, transmit, send, packet, message, carry, encapsulate, device, ID, identity, virtual link, FCoPIP, FC-TLP		
C. 相关文件		
类 型*	引用文件, 必要时, 指明相关段落	相关的权利要求
A	US20070266179A1 (EMULEX COMMUNICATIONS CORP.) 15.11 月 2007 (15.11.2007) 参见全文	1-20
A	杨阿锋等, “PCIe 接口高速数据传输卡的驱动程序开发”, 《中国测试技术》, 第 34 卷第 2 期, 31.3 月 2008 (31.03.2008) 参见全文	1-20
A	CN101588365A (杭州华三通信技术有限公司) 25.11 月 2009 (25.11.2009) 参见全文	1-20
A	CN101221552A (国际商业机器公司) 16.7 月 2008 (16.07.2008) 参见全文	1-20
<input type="checkbox"/> 其余文件在 C 栏的续页中列出。 <input checked="" type="checkbox"/> 见同族专利附件。		
* 引用文件的具体类型:		“T” 在申请日或优先权日之后公布, 与申请不相抵触, 但为了理解发明之理论或原理的在后文件
“A” 认为不特别相关的表示了现有技术一般状态的文件		“X” 特别相关的文件, 单独考虑该文件, 认定要求保护的发明不是新颖的或不具有创造性
“E” 在国际申请日的当天或之后公布的在先申请或专利		“Y” 特别相关的文件, 当该文件与另一篇或者多篇该类文件结合并且这种结合对于本领域技术人员为显而易见时, 要求保护的发明不具有创造性
“L” 可能对优先权要求构成怀疑的文件, 或为确定另一篇引用文件的公布日而引用的或者因其他特殊理由而引用的文件(如具体说明的)		“&” 同族专利的文件
“O” 涉及口头公开、使用、展览或其他方式公开的文件		
“P” 公布日先于国际申请日但迟于所要求的优先权日的文件		
国际检索实际完成的日期 23.6 月 2012 (23.06.2012)	国际检索报告邮寄日期 05.7 月 2012 (05.07.2012)	
ISA/CN 的名称和邮寄地址: 中华人民共和国国家知识产权局 中国北京市海淀区蓟门桥西土城路 6 号 100088 传真号: (86-10)62019451	授权官员 郑杰 电话号码: (86-10) 62412014	

国际检索报告
关于同族专利的信息

国际申请号
PCT/CN2012/072608

检索报告中引用的 专利文件	公布日期	同族专利	公布日期
US20070266179A1	15.11.2007	US7743197B2	22.06.2010
CN101588365A	25.11.2009	CN101588365B	23.05.2012
CN101221552A	16.07.2008	US2008168255A1	10.07.2008
		US7661006B2	09.02.2010
		CN101221551B	21.07.2010

续：A. 主题的分类

按照国际专利分类(IPC)或者同时按照国家分类和 IPC 两种分类

H04L 12/56 (2006.01) i

H04L 12/40 (2006.01) i