

(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
2015年4月2日 (02.04.2015)



(10) 国际公布号
WO 2015/043181 A1

- (51) 国际专利分类号:
G06F 3/06 (2006.01)
- (21) 国际申请号: PCT/CN2014/076582
- (22) 国际申请日: 2014年4月30日 (30.04.2014)
- (25) 申请语言: 中文
- (26) 公布语言: 中文
- (30) 优先权:
201310442041.1 2013年9月25日 (25.09.2013) CN
- (71) 申请人: 华为技术有限公司 (HUAWEI TECHNOLOGIES CO., LTD.) [CN/CN]; 中国广东省深圳市龙岗区坂田华为总部办公楼, Guangdong 518129 (CN)。
- (72) 发明人: 李延松 (LI, Yansong); 中国广东省深圳市龙岗区坂田华为总部办公楼, Guangdong 518129 (CN)。 郑玉林 (ZHENG, Yulin); 中国广东省深圳市龙岗区坂田华为总部办公楼, Guangdong 518129 (CN)。
- (74) 代理人: 北京弘权知识产权代理事务所(普通合伙) (CHINABLE IP); 中国北京市朝阳区安定路35号六层35-10-2内620室, Beijing 100029 (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, BN, 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, IR, IS, JP, KE, KG, 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, PA, PE, PG, PH, PL, PT, QA, RO, RS, RU, RW, SA, 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, 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, KM, ML, MR, NE, SN, TD, TG)。

[见续页]

(54) Title: MEMORY EXTENSION SYSTEM AND METHOD

(54) 发明名称: 内存扩展系统及方法

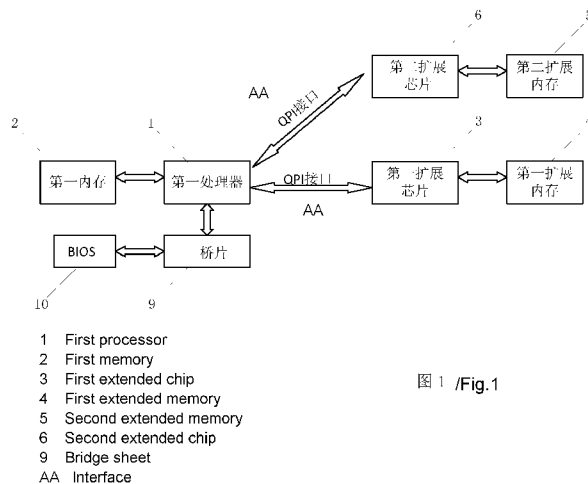


图1 /Fig.1

(57) Abstract: Disclosed are a memory extension system and method. The system comprises: a processor, an extended memory, an extended chip, and multiple processor installation positions, a memory installation position being disposed on each processor installation position. The multiple processor installation positions are inter-connected by using a QuickPath Interconnect (QPI) interface, at least one processor installation position is provided with a processor, and the other at least one processor installation position serves as an extended installation position; the at least one extended installation position is provided with an extended chip; and the extended memory is installed on the memory installation position connected to the extended chip. For the memory extension system, an extended chip is installed on at least one processor installation position to replace a processor, so that an existing processor can access an extended memory of the extended chip by using the extended chip; in this way, a memory capacity of the existing processor is increased without improving a processing capability, thereby solving the problem in the prior art that the processing capability is redundant by adding a processor to extend the memory.

(57) 摘要:

[见续页]



WO 2015/043181 A1

本国际公布:

- 包括国际检索报告(条约第 21 条(3))。

本申请实施例公开了一种内存扩展系统及方法，该系统包括：处理器、扩展内存、扩展芯片和多个处理器安装位，每个处理器安装位上都设置有内存安装位；多个处理器安装位之间通过快速通道互联 QPI 接口相连接，至少一个处理器安装位上安装有处理器，其它至少一个处理器安装位作为扩展安装位；至少一个扩展安装位上安装有扩展芯片；扩展内存安装在与扩展芯片相连接的内存安装位上。该内存扩展系统，在其它处理器位安装扩展芯片替代处理器，使得已有的处理器可以通过扩展芯片访问扩展芯片所带的扩展内存，从而在没有增加处理能力的情况下增加了已有处理器的内存容量，解决了现有技术通过增加处理器来扩展内存导致处理能力冗余的问题。

内存扩展系统及方法

本申请要求于 2013 年 09 月 25 日提交中国专利局、申请号为 201310442041.1、发明名称为“内存扩展系统及方法”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

本申请涉及信息存储技术领域，特别是涉及一种内存扩展系统及方法。

背景技术

对于计算机系统来说，处理器的计算能力、内存容量、输入/输出接口的带宽是三个重要的性能评价指标。为了提高计算能力，系统需要配置高性能的处理器，例如
10 多核多线程的处理器，而且常常将多个处理器互连起来形成 SMP (Symmetrical Multi Processing, 对称多处理器) 或 NUMA (Non Uniform Memory Access, 非均匀性内存访问) 结构，实现业务的并发处理。为了扩大内存容量，处理器通常集成了多个内存控制器，每个控制器可以支持多个内存条，结合上述多处理器技术，可以将内存容量扩大到数百 GB 的规模。为了提高输入/输出接口的带宽，近十年来业界推出了
15 PCI Express (Peripheral Component Interconnect Express, 外设互联标准总线)、RapidIO、Interlaken、10G 以太网等高速接口，速率也随着协议版本的升级在不断提高。

一般来说，计算能力、内存容量和输入/输出接口带宽应该均衡配置，否则容易出现性能瓶颈，影响计算机系统的整体性能。但在某些特殊应用场景下，例如服务器
20 用于不同的业务类型 (Web 服务器、数据库服务器、流媒体服务器等)，对这三个指标的要求有不同的侧重点。对于以控制功能为主的应用场景，处理器的计算能力最重要，相对来说内存容量不是太迫切；而对于虚拟化应用场景来说，由于需要将多种业务部署在同一个处理器的多个内核上，每种业务都要分配一定的内存空间，因此对内存容量的需求相对更为迫切。

25 而目前的服务器一般都是采用通用多路处理器设计，虽然可以支持多种业务，但具体执行何种业务是由客户自行决定的，例如安装不同的应用程序完成不同的业务功能，因此服务器可能在处理能力或内存容量方面存在一定的冗余，增加硬件成本和运

行功耗。举例来说，假设有一个四路服务器，里面最多可以安装四个处理器，每个处理器都集成了内存控制器，可以外接内存条。在上述虚拟化场景下，可能不需要安装所有四个处理器，只安装两个处理器即可满足计算能力的需求，但内存容量需求却可能超过了两个处理器自带内存的容量，如果把另两个处理器也装上去，会增加成本和功耗，而如果不把另两个处理器装上去，很显然内存容量又不够用。

发明内容

本申请中提供了一种内存扩展系统及方法，以解决现有技术中通过增加处理器来扩展内存容量导致处理能力冗余的问题。

为了解决上述技术问题，一方面，本申请提供了一种内存扩展系统，包括：处理器、扩展内存、扩展芯片和多个处理器安装位，其中，每个所述处理器安装位上都设置有与所述处理器安装位相连接的内存安装位；多个所述处理器安装位之间通过快速通道互联 QPI 接口相连接，至少一个处理器安装位上安装有所述处理器，其它至少一个处理器安装位作为扩展安装位；至少一个扩展安装位上安装有所述扩展芯片，所述扩展芯片具有唯一的标识号码；所述扩展内存安装在与所述扩展芯片相连接的内存安装位上，所有所述扩展内存的地址都在所述处理器的地址空间范围内。

在第一方面第一种可能的实现方式中，所述扩展芯片的管脚与所述处理器的管脚兼容。

结合第一方面或第一方面的第一种可能的实现方式，在第一方面第二种可能的实现方式中，所述处理器包括：查询指令生成模块，用于生成标识号码查询指令；查询指令发送模块，用于将所述标识号码查询指令通过所述处理器所在处理器安装位上的 QPI 接口发送；标识号码接收模块，用于通过所述处理器所在处理器安装位上的 QPI 接口接收标识号码；标识号码判断模块，用于判断接收到的所述标识号码是否为扩展芯片的标识号码；配置模块，用于当所述标识号码为扩展芯片的标识号码时，为所述扩展芯片配置内存控制器，在所述处理器的地址空间中分配与所述扩展芯片相连接的扩展内存相对应的内存地址。

结合第一方面的第二种可能的实现方式，在第一方面第三种可能的实现方式中，该处理器还包括：报文生成模块，用于生成用于读取数据或写入数据的、协议为 QPI 协议的报文；报文发送模块，用于将所述报文通过所述处理器所在处理器安装位上的 QPI 接口发送；数据接收模块，用于通过所述处理器所在处理器安装位上的 QPI 接口接收所述扩展芯片根据所述报文读取得到的数据。

结合第一方面的第三种可能的实现方式，在第一方面第四种可能的实现方式中，所述扩展芯片包括：查询指令接收模块，用于通过所述扩展芯片所在扩展安装位上的 QPI 接口接收标识号码查询指令；标识号码获取模块，用于根据所述标识号码查询指令获取所述扩展芯片的标识号码；标识号码发送模块，用于将获取得到的标识号码通过所述扩展芯片所在扩展安装位上的 QPI 接口发送。

结合第一方面的第四种可能的实现方式，在第一方面第五种可能的实现方式中，所述扩展芯片还包括：QPI 接口模块、协议转换模块和内存接口模块，其中，所述 QPI 接口模块与所述扩展芯片所在扩展安装位上的 QPI 接口相连接，用于接收所述处理器发送的、协议为 QPI 协议的报文；所述协议转换模块，用于将 QPI 协议的报文转成内存接口协议后向所述内存接口模块发送；所述内存接口模块与位于所述扩展芯片所在扩展安装位上内存安装位上的扩展内存相连接，用于根据协议转换后的报文在所述扩展内存中读取数据或写入数据，将读取到的数据向所述协议转换模块发送；所述协议转换模块还用于将读取到的数据由内存接口协议转换成 QPI 协议后向所述 QPI 接口模块发送，所述 QPI 接口模块还用于将协议转换后的数据通过所述扩展芯片所在扩展安装位上的 QPI 接口发送。

第二方面，本申请提供了一种内存扩展方法，所述方法包括：处理器生成标识号码查询指令，将所述标识号码查询指令通过所述处理器所在处理器安装位上的快速通道互联 QPI 接口发送；扩展芯片通过所述扩展芯片所在处理器安装位上的 QPI 接口接收查询指令，根据所述标识号码查询指令获取所述扩展芯片的标识号码，将获取得到的标识号码通过所述扩展芯片所在处理器安装位上的 QPI 接口发送；所述处理器通过所述处理器所在处理器安装位上的 QPI 接口接收标识号码，判断所述标识号码是否为扩展芯片的标识号码，当所述标识号码为扩展芯片的标识号码时，为所述扩展芯片配置内存控制器，在所述处理器的地址空间中分配与所述扩展芯片相连接的扩展内存相对应的内存地址。

在第二方面的第一种可能的实现方式中，该方法还包括：所述处理器判断是否向所述处理器所在处理器安装位上的所有 QPI 接口发送标识号码查询指令；如果否，所述处理器向未发送标识号码查询指令的 QPI 接口发送标识号码查询指令。

结合第一方面或第二方面第一种可能的实现方式，在第二方面第二种可能的实现方式中，该方法还包括：所述处理器生成用于读取数据或写入数据的、协议为 QPI 协议的报文，将所述报文通过所述处理器所在处理器安装位上的 QPI 接口发送；所述扩展芯片通过所述扩展芯片所在处理器安装位上的 QPI 接口接收所述报文，将所述报

文转换成内存接口协议后,根据协议转换后的报文在与所述扩展芯片相连接的扩展内存中读取数据或写入数据,将读取到的数据由内存接口协议转换成 QPI 协议后通过所述扩展芯片所在处理器安装位上的 QPI 接口发送;所述处理器通过所述处理器所在处理器安装位上的 QPI 接口接收协议转换后的数据。

5 由以上技术方案可见,本申请提供的该内存扩展系统,在一些处理器安装位上设置有处理器,在其它处理器安装位上设置扩展芯片,处理器与扩展芯片之间可以通过处理器安装位之间的 QPI 接口进行数据交互,每个扩展芯片都可以单独连接有自己的扩展内存,处理器通过扩展芯片可以对与每个扩展芯片连接的扩展内存进行访问。

10 与现有技术相比,当处理能力满足要求但内存容量需求超过处理器自带内存的容量的情况,该内存扩展系统,采用在其它处理器安装位扩展芯片替代处理器,已有的处理器可以通过扩展芯片访问扩展芯片所带的扩展内存,从而可以在没有增加处理能力的情况下增加了已有处理器的内存容量,避免现有技术通过增加处理器来扩展内存容量导致处理能力冗余的问题。

附图说明

15 为了更清楚地说明本发明实施例的技术方案,下面将对实施例中所需要使用的附图作简单地介绍,显而易见地,下面描述中的附图仅仅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性的前提下,还可以根据这些附图获得其他的附图。

图 1 为本申请实施例提供的一种内存扩展系统的结构示意图;

20 图 2 为本申请实施例提供的一种第一处理器的结构示意图;

图 3 为本申请实施例提供的一种扩展芯片的结构示意图;

图 4 为本申请实施例提供的另一种内存扩展系统的结构示意图;

图 5 为本申请实施例提供的另一种第一处理器的结构示意图;

图 6 为本申请实施例提供的另一种扩展芯片的结构示意图;

25 图 7 为本申请实施例提供的一种内存扩展方法的流程示意图;

图 8 为本申请实施例提供的另一种内存扩展方法的流程示意图;

图 9 为本申请实施例提供的又一种内存扩展方法的流程示意图。

具体实施方式

本申请如下实施例提供了一种内存扩展系统及方法,该内存扩展系统在多路处理

器设计的系统中，当处理能力满足要求但内存容量需求超过处理器自带内存的容量的情况，可以在其它处理器安装位上安装扩展芯片以替代处理器，并且在扩展芯片所在的处理器安装位的内存安装位上安装需要的扩展内存，这样已有的处理器就可以通过 QPI 接口与扩展芯片进行通信，实现对扩展芯片连接的扩展内存进行访问，因此可以解决现有多处理器设计的服务器通过增加处理器来扩展内存容量导致处理能力冗余的问题。

为了使本技术领域的人员更好地理解本申请实施例中的技术方案，并使本申请实施例的上述目的、特征和优点能够更加明显易懂，下面结合附图对本申请实施例中技术方案作进一步详细的说明。

一个实施例：

参见图 1，为本申请实施例提供的内存扩展系统的一种结构示意图，图中以三个处理器安装座的多处理器设计的系统为例，该内存扩展系统可以包括：第一处理器 1、第一内存 2、第一扩展芯片 3、第一扩展内存 4、第二扩展内存 5 和第二扩展芯片 6，其中，第一处理器 1、第一扩展芯片 3 和第二扩展芯片 6 分别安装在三个处理器安装位上（图中未示出），三个处理器安装位之间通过 QPI（Quick Path Interconnect，快速通道互联）接口相连接，第一内存 2 为第一处理器 1 的自带内存，第一扩展内存 4 与第一扩展芯片 3 相连接，第二扩展内存 5 与第二扩展芯片 6 相连接。

为了叙述方便，在本申请实施例中，可以将第一处理器 1 所在的处理器安装位称为第一处理器安装位，将第一扩展芯片 3 所在的处理器安装位称为第一扩展安装位，将第二扩展芯片 6 所在的处理器安装位称为第二扩展安装位。第一处理器 1 可以直接焊接在第一处理器安装位，另外，为了方便拆卸或方便对多处理器系统进行组合，第一处理器 1 还可以通过插拔接口安装在第一处理器安装位上，插拔接口可以为相互匹配的管脚和插孔，或者，相互匹配的插卡和插槽。第一扩展芯片 3 和/或第二扩展芯片 6 可以通过焊接的方式固定在相应的扩展安装位上，当然为了方便拆卸或方便对多处理器系统进行组合，也可以通过插拔接口的方式与相应的扩展安装位相连接。三个处理器安装位之间通过 QPI 接口进行数据交互，QPI 接口由 20 对差分发送信号、20 对差分接收信号和两对差分时钟信号组成，信号速率可到 6.4GT/s，双向带宽为 25.6GB/s，此外还支持高速缓冲存储器一致性，可以执行对外围设备的配置访问和对内存的读写操作。

在每个处理器安装位上都设置有一个或多个内存安装位，用于增加内存空间。第一内存 2 设置在第一处理器安装位上的内存安装位上，第一扩展内存 4 设置在第一扩

展安装位上的内存安装位上,第二扩展内存 5 设置在第二扩展安装位上的内存安装位上。处理器安装位与相连接的内存安装位之间通过内存接口相连接,并且处理器安装位与相连接的内存安装位之间交互数据的协议格式为内存接口协议。

每个扩展芯片都有唯一的标识号码,通过该标识号码可以识别出处理器安装位上的是否为扩展芯片。每个扩展芯片都可以连接一个或多个扩展内存,并且所有扩展芯片相连接的扩展内存的地址都在第一处理器 1 的地址空间范围内。另外,为了保证每个扩展芯片都可以安装在处理器安装位上,在本申请实施例中,每个扩展芯片的管脚与处理器的管脚都相兼容。在本申请实施例中,可以将扩展芯片的以下管脚定义与处理器的管脚兼容:

10 1、电源。处理器的电源种类较多,例如核心电压、I/O 电压等,都由单板上的电源模块提供。扩展芯片的核心电压可以与处理器相同,也可以通过扩展芯片内部电路将单板提供的电压转换为自己所需的电压。扩展芯片的 I/O 电压用于给 I/O 接口电路供电,由于扩展芯片也要提供 QPI 接口,因此这部分电路可以直接利用单板提供的电压,管脚位置与处理器保持一致。

15 2、地。地可以为各种电源和信号提供回路,同时为高速信号提供屏蔽功能,扩展芯片可以直接利用处理器已有的地信号。

20 3、时钟。时钟为处理器内部模块、I/O 接口提供参考时钟,处理器内部有锁相环可以将单板提供的外部时钟倍频到更高的频率,还会产生内存接口时钟来同步对内存条的读写访问。扩展芯片可以直接利用处理器的输入参考时钟,并倍频得到自己所需的工作时钟,同时必须接收和产生 QPI 接口时钟与对端处理器通信,还要产生内存接口时钟来同步对内存条的访问,管脚位置、功能和电气特性与处理器保持一致。

 4、复位。复位可以将处理器和单板上的其它器件置为确定的初始状态,然后处理器从第一条指令开始执行。扩展芯片可以直接利用处理器的复位信号,管脚位置、功能和电气特性与处理器保持一致。

25 5、QPI 接口。扩展芯片也要提供 QPI 接口,管脚位置、功能和电气特性与处理器保持一致。

 6、内存接口。扩展芯片也要提供内存接口,包括用于读取内存条信息的 I2C 接口,管脚位置、功能和电气特性与处理器保持一致。

30 图 1 中的 9 为桥片,10 为 BIOS (Basic Input Output System,基本输入输出系统),第一处理器 1 通过 DMI 接口 (Direct Media Interface,直接媒体接口)与桥片 9(一般称为南桥)相连接,通过桥片 9,第一处理器 1 可以访问 BIOS10。通常将可

以访问桥片 9 的第一处理器 1 称为主处理器，将其它处理器称为从处理器（图 1 所示实施例中没有设置从处理器），每次上电复位结束后都是主处理器先启动，执行外围设备的初始化，然后控制从处理器启动。

参见图 2，为本申请实施例提供的第一处理器的一种结构示意图。第一处理器 1
5 可以包括：

查询指令生成模块 21，用于生成标识号码查询指令；

查询指令发送模块 22，用于将标识号码查询指令通过第一处理器安装位上的 QPI
接口发送；

标识号码接收模块 23，用于通过第一处理器安装位上的 QPI 接口接收标识号码；

10 标识号码判断模块 24，用于判断接收到的标识号码是否为扩展芯片的标识号码；

配置模块 25，用于当标识号码为扩展芯片的标识号码时，为扩展芯片配置内存
控制器，在处理器的地址空间中分配与扩展芯片相连接的扩展内存相对应的内存地
址。

相对应图 2 所示的第一处理器，参见图 3，为本申请实施例提供的扩展芯片的一
15 种结构示意图。扩展芯片可以包括：

查询指令接收模块 31，用于通过扩展芯片所在扩展安装位上的 QPI 接口接收标
识号码查询指令；

标识号码获取模块 32，用于根据标识号码查询指令获取扩展芯片的标识号码；

20 标识号码发送模块 33，用于将获取得到的标识号码通过扩展芯片所在扩展安
装位上的 QPI 接口发送。

由此，可以看到，在第一处理器 1 通过对扩展芯片的标识号码进行查询，就可以
对扩展芯片进行初始化。

在本申请实施例中，处理器的数量以 1 个为例。在实际应用中，处理器的数量
根据系统对计算能力的需要可以自由设定，可以为两个或多个，当处理器为多个时，
25 需要处理器能够提供的 QPI 接口更多。

参见图 4，为本申请实施例提供的内存扩展系统的另一种结构示意图。图中，内
存扩展系统还可以包括：第二处理器 7 和第二内存 8，第二处理器 7 所在的处理器安
装位为第二处理器安装位，并且第二内存 8 是第二处理器 7 的自带内存。

30 第一处理器 1 和第二处理器 7 通过 QPI 接口直连，两个处理器都集成有内存控制
器，并且每个处理器都可以外挂一个或多个属于自己的内存，同时两个处理器还支持
彼此内存之间的相互访问。图 4 所示实施例中，第一处理器 1 可以称为主处理器，第

二处理器 2 为从处理器。当系统上电或复位结束后主处理器最先启动，执行 BIOS 中的代码，在完成外围设备的初始化并为从处理器准备好启动代码后才控制其它从处理器启动，然后加载操作系统，开始在所有处理器上运行。

在系统上电后，第一处理器 1 在外围设备初始化完成后以及加载操作系统成功
5 后，就可以对扩展芯片上的扩展内存进行访问。在本申请实施例中，如图 5 所示，第一处理器 1 还可以包括：

报文生成模块 51，用于生成用于读取数据或写入数据的、协议为 QPI 协议的报文；

报文发送模块 52，用于将报文通过第一处理器安装位上的 QPI 接口发送；

10 数据接收模块 53，用于通过第一处理器安装位上的 QPI 接口接收扩展芯片根据所述报文读取得到的数据。

相对应图 5 所示的第一处理器，如图 6 所示，扩展芯片还可以包括：QPI 接口模块 61、协议转换模块 62 和内存接口模块 63。

15 QPI 接口模块 61 与扩展芯片所在扩展安装位上的 QPI 接口相连接，用于接收第一处理器发送的、协议为 QPI 协议的报文。由于处理器安装位之间通过 QPI 接口进行数据交互，所以第一处理器发送的报文为 QPI 协议格式。

20 协议转换模块 62，用于将 QPI 协议的报文转成内存接口协议后向内存接口模块 63 发送。对于扩展芯片和扩展内存，两者之间通过内存接口相连接，即两者之间的交互的数据为内存接口协议格式。所以，协议转换模块 62 在接收到 QPI 接口发送的 QPI 协议格式的报文后，首先根据 QPI 协议对报文进行解析，识别出报文内包含的：
要访问的内存地址、访问类型(读或写)、数据长度和待写入的数据(对于写操作)，然后将识别得到的信息转换成内存接口协议格式后，发送给内存接口模块 63，由内存接口模块 63 发送给扩展内存。另外，当内存接口模块 63 接收到扩展内存发送的数据后，协议转换模块 62 还用于将接收到的数据转换成 QPI 协议格式，然后由 QPI 接口模块
25 61 发送给第一处理器 1。

内存接口模块 63 与位于扩展芯片所在扩展安装位上的内存安装位上的扩展内存相连接，用于根据协议转换后的报文在扩展内存中读取数据或写入数据，将读取到的数据向协议转换模块 62 发送。

30 由以上技术方案可见，本申请实施例提供的该内存扩展系统，在一些处理器安装位上设置有处理器，在其它处理器安装位上设置扩展芯片，处理器与扩展芯片之间可以通过处理器安装位之间的 QPI 接口进行数据交互，每个扩展芯片都可以单独连接有

扩展内存，并且处理器可以对与每个扩展芯片连接的扩展内存进行访问。

与现有技术相比，该内存扩展系统，当处理能力满足要求但内存容量需求超过处理器自带内存的容量的情况，通过在其它处理器位安装扩展芯片替代处理器，通过扩展芯片访问扩展芯片所带的扩展内存，从而可以避免出现现有技术通过增加处理器来扩展内存导致处理能力冗余的问题。

另外，相对于处理器，扩展芯片的作用主要是进行协议转换，扩展芯片的成本和功耗都很小，所以本申请实施例提供的该内存扩展系统用较小的成本和功耗就可以实现增加内存容量。

此外，在本申请实施例中，都是以第一处理器为例进行说明，第一扩展芯片和第二扩展芯片都附属于第一处理器。在本申请其他实施例中，每个扩展芯片都可以与不同处理器之间建立联系，例如：第二扩展芯片既可以附属于第一处理器，也可以附属于第二处理器，同样，第一扩展芯片也是如此，当第二处理器有附属的扩展芯片时，第二处理器的结构可以参见上述实施例中描述的第一处理器的结构，在此不再赘述。这里，扩展芯片与处理器之间的对应关系，可以根据实际使用情况自由设定，例如：在本申请实施例中，第一扩展芯片和第二扩展芯片都属于第一处理器，那么第一处理器的软件设计和控制就需要专门设计，例如第一处理器通过控制是否允许第二处理器检测扩展芯片，进而实现对第二处理器进行控制。

另一实施例：

对应图 1 所示的内存扩展系统，参见图 7，为本申请实施例提供的一种内存扩展方法的流程示意图。所述内存扩展方法可以包括：

S101：处理器生成标识号码查询指令，将标识号码查询指令发送。

处理器将生成标识号码查询指令通过处理器所在处理器安装位上的 QPI 接口发送，用于对与处理器所在处理器安装位相连接的扩展芯片进行确定。

S102：扩展芯片接收标识号码查询指令，根据标识号码查询指令获取扩展芯片的标识号码，将获取得到的标识号码发送。

每个扩展芯片所在的处理器安装位都通过 QPI 接口与处理器所在处理器安装位相连接，所以每个扩展芯片都可以通过所在处理器安装位上的 QPI 接口接收处理器发送的标识号码查询指令。根据标识号码查询指令，扩展芯片可以在查询自身存储的标识号码，并且将查询到的标识号码通过所在处理器安装位上的 QPI 接口返回给处理器。

S103：处理器接收标识号码，判断标识号码是否为扩展芯片的标识号码，当标识

号码为扩展芯片的标识号码时，为扩展芯片配置内存控制器，在处理器的地址空间中分配与扩展芯片相连接的扩展内存相对应的内存地址。

在本申请实施例中，处理器通过向所在处理器安装位上的各个 QPI 接口发送标识号码查询指令，并接收返回的标识号码，通过对标识号码进行识别可以确认与所在处理器安装位相连接的扩展芯片，当识别出扩展芯片后，可以对扩展芯片以及与扩展芯片连接的扩展内存进行初始化操作，包括：为扩展芯片配置的内存控制器，为扩展内存分配地址空间，其中，配置内存控制器可以对与扩展芯片相连接的扩展内存进行数据读取或写入等操作，为扩展内存分配地址空间，可以将扩展内存映射成处理器自己地址空间的一部分。初始化操作完成后，处理器就可以对与扩展芯片相连接的扩展内存进行访问。当然，处理器还可以对自带的内存也进行初始化操作，例如：图 1 中第一处理器 1 的自带内存为第一内存 2，第一处理器 1 可以为第一内存 2 配置内存控制器以及为第一内存 2 分配地址空间。

参见图 8，为本申请实施例提供的内存扩展方法的另一种流程示意图。该内存扩展方法还可以包括：

15 S201：第一处理器生成用于读取数据或写入数据的报文，将报文发送。

第一处理器根据数据读取或写入的需要，在第一处理器内生成读取数据或写入数据的报文，并将报文通过第一处理器安装位上与扩展芯片相连接的 QPI 接口发送出去。由于处理器安装位之间都是通过 QPI 接口进行数据交互，所以生成的读取数据或写入数据的报文的格式为 QPI 协议。

20 S202：扩展芯片接收报文，将报文转换成内存接口协议后，根据协议转换后的报文在扩展内存中读取数据或写入数据，将读取到的数据由内存接口协议转换成 QPI 协议后发送。

扩展芯片通过扩展芯片所在的扩展安装位上的 QPI 接口接收第一处理器 1 发送报文。由于第一处理器发送的报文为 QPI 协议格式，而对于扩展芯片和扩展内存，两者之间通过内存接口相连接，即两者之间的交互的数据为内存接口协议格式，所以，在扩展芯片接收到 QPI 协议格式的报文后，首先根据 QPI 协议对报文进行解析，识别出报文内包含的：要访问的内存地址、访问类型（读或写）、数据长度和待写入的数据（对于写操作），然后将识别得到的信息转换成内存接口协议格式后，发送给扩展内存，在扩展内存中进行数据读取或数据写入。当从扩展内存读取到数据后，扩展芯片接收扩展内存发送的数据，并且将数据转换成 QPI 协议格式，然后由扩展芯片所在的扩展安装位上的 QPI 接口发送给第一处理器 1。

通过上述描述,可见,扩展芯片在数据读取或数据写入过程中,其主要作用是进行协议转换,与现有技术需要增加内存容量采用安装新的处理器而言,安装扩展芯片来增加内存容量的方式,不仅成本大大降低,而且扩展芯片的功耗相对处理器也非常小。

5 S203: 第一处理器接收扩展芯片发送的数据。

第一处理器 1 通过第一处理器安装位上的 QPI 接口接收扩展芯片发送的 QPI 协议格式的数据。

10 由以上技术方案可见,本申请实施例提供的该内存扩展方法,通过查找与处理器相连接的扩展芯片的标识号码,对扩展芯片进行初始化操作,包括:为扩展芯片配置内存控制器、在所述处理器的地址空间中为扩展内存分配相对应的内存地址,向扩展内存加载启动程序。因此,当处理器需要数据读取或写入时,可以通过扩展芯片直接访问与扩展芯片相连接的扩展内存,即以较小的成本和功耗,就可以解决处理能力冗余的问题。

15 另外,对于图 4 所示的多处理器内存扩展系统,第一处理器 1 作为主处理器,第二处理器 7 作为从处理器,在具体的执行过程中,作为主处理器的第一处理器 1 还可以执行图 9 所示的步骤:

S301: 系统复位结束后第一处理器启动;

S302: 通过 QPI 接口查询标识号码;

20 S303: 判断标识号码是否为扩展芯片的标识号码,如果是,执行步骤 S304,如果否,转入步骤 S305;

S304: 为扩展芯片配置内存控制器,为扩展内存分配相应的内存空间;

S305: 判断第一处理器安装位上所有 QPI 接口是否遍历完毕;当第一处理器安装位上所有 QPI 接口遍历完毕时,执行步骤 S306;当第一处理器安装位上所有 QPI 接口未遍历完毕时,转入步骤 S302;

25 S306: 为第二处理器加载启动程序;

S307: 控制第二处理器启动;

S308: 加载操作系统运行。

30 通过上述步骤的配置,可以对与第一处理器相连接的外围设备(指扩展芯片)进行初始化,并且为其它从处理器准备好启动代码,控制其它从处理器开始启动,最后第一处理器加载操作系统,并进行运行。对扩展芯片初始化后,第一处理器就可以对与扩展芯片相连接的扩展内存进行访问。

通过以上的方法实施例的描述，所属领域的技术人员可以清楚地了解到本申请可借助软件加必需的通用硬件平台的方式来实现，当然也可以通过硬件，但很多情况下前者是更佳的实施方式。基于这样的理解，本申请的技术方案本质上或者说对现有技术做出贡献的部分可以以软件产品的形式体现出来，该计算机软件产品存储在一个存储介质中，包括若干指令用以使得一台计算机设备（可以是个人计算机，服务器，或者网络设备等）执行本申请各个实施例所述方法的全部或部分步骤。而前述的存储介质包括：只读存储器（ROM）、随机存取存储器（RAM）、磁碟或者光盘等各种可以存储程序代码的介质。

可以理解的是，本申请可用于众多通用或专用的计算系统环境或配置中。例如：个人计算机、服务器计算机、手持设备或便携式设备、平板型设备、多处理器系统、基于微处理器的系统、置顶盒、可编程的消费电子设备、网络 PC、小型计算机、大型计算机、包括以上任何系统或设备的分布式计算环境等等。

本申请可以在由计算机执行的计算机可执行指令的一般上下文中描述，例如程序模块。一般地，程序模块包括执行特定任务或实现特定抽象数据类型的例程、程序、对象、组件、数据结构等等。也可以在分布式计算环境中实践本申请，在这些分布式计算环境中，由通过通信网络而被连接的远程处理设备来执行任务。在分布式计算环境中，程序模块可以位于包括存储设备在内的本地和远程计算机存储介质中。

需要说明的是，在本文中，诸如第一和第二等之类的关系术语仅仅用来将一个实体或者操作与另一个实体或操作区分开来，而不一定要求或者暗示这些实体或操作之间存在任何这种实际的关系或者顺序。而且，术语“包括”、“包含”或者其任何其它变体意在涵盖非排他性的包含，从而使得包括一系列要素的过程、方法、物品或者设备不仅包括那些要素，而且还包括没有明确列出的其它要素，或者是还包括为这种过程、方法、物品或者设备所固有的要素。在没有更多限制的情况下，由语句“包括一个……”限定的要素，并不排除在包括所述要素的过程、方法、物品或者设备中还存在另外的相同要素。

以上所述仅是本申请的具体实施方式，应当指出，对于本技术领域的普通技术人员来说，在不脱离本申请原理的前提下，还可以做出若干改进和润饰，这些改进和润饰也应视为本申请的保护范围。

权 利 要 求

1、一种内存扩展系统，其特征在于，包括：处理器、扩展内存、扩展芯片和多个处理器安装位，其中，

每个所述处理器安装位上都设置有与所述处理器安装位相连接的内存安装位；

多个所述处理器安装位之间通过快速通道互联 QPI 接口相连接，至少一个处理器安装位上安装有所述处理器，其它至少一个处理器安装位作为扩展安装位；

至少一个扩展安装位上安装有所述扩展芯片，所述扩展芯片具有唯一的标识号码；

所述扩展内存安装在与所述扩展芯片相连接的内存安装位上，所有所述扩展内存的地址都在所述处理器的地址空间范围内。

2、根据权利要求 1 所述的系统，其特征在于，所述扩展芯片的管脚与所述处理器的管脚兼容。

3、根据权利要求 1 或 2 所述的系统，其特征在于，所述处理器包括：

查询指令生成模块，用于生成标识号码查询指令；

查询指令发送模块，用于将所述标识号码查询指令通过所述处理器所在处理器安装位上的 QPI 接口发送；

标识号码接收模块，用于通过所述处理器所在处理器安装位上的 QPI 接口接收标识号码；

标识号码判断模块，用于判断接收到的所述标识号码是否为扩展芯片的标识号码；

配置模块，用于当所述标识号码为扩展芯片的标识号码时，为所述扩展芯片配置内存控制器，在所述处理器的地址空间中分配与所述扩展芯片相连接的扩展内存相对应的内存地址。

4、根据权利要求 3 所述的系统，其特征在于，该处理器还包括：

报文生成模块，用于生成用于读取数据或写入数据的、协议为 QPI 协议的报文；

报文发送模块，用于将所述报文通过所述处理器所在处理器安装位上的 QPI 接口发送；

数据接收模块，用于通过所述处理器所在处理器安装位上的 QPI 接口接收所述扩展芯片根据所述报文读取得到的数据。

5、根据权利要求 1-4 任一项所述的系统，其特征在于，所述扩展芯片包括：

查询指令接收模块，用于通过所述扩展芯片所在扩展安装位上的 QPI 接口接收标识号码查询指令；

标识号码获取模块，用于根据所述标识号码查询指令获取所述扩展芯片的标识号码；

标识号码发送模块，用于将获取得到的标识号码通过所述扩展芯片所在扩展安装位上的 QPI 接口发送。

6、根据权利要求 5 所述的系统，其特征在于，所述扩展芯片还包括：QPI 接口模块、协议转换模块和内存接口模块，其中，

所述 QPI 接口模块与所述扩展芯片所在扩展安装位上的 QPI 接口相连接，用于接收所述处理器发送的、协议为 QPI 协议的报文；

所述协议转换模块，用于将 QPI 协议的报文转成内存接口协议后向所述内存接口模块发送；

所述内存接口模块与位于所述扩展芯片所在扩展安装位上内存安装位上的扩展内存相连接，用于根据协议转换后的报文在所述扩展内存中读取数据或写入数据，将读取到的数据向所述协议转换模块发送；

所述协议转换模块还用于将读取到的数据由内存接口协议转换成 QPI 协议后向所述 QPI 接口模块发送，所述 QPI 接口模块还用于将协议转换后的数据通过所述扩展芯片所在扩展安装位上的 QPI 接口发送。

7、一种内存扩展方法，其特征在于，所述方法包括：

处理器生成标识号码查询指令，将所述标识号码查询指令通过所述处理器所在处理器安装位上的快速通道互联 QPI 接口发送；

扩展芯片通过所述扩展芯片所在处理器安装位上的 QPI 接口接收查询指令，根据所述标识号码查询指令获取所述扩展芯片的标识号码，将获取得到的标识号

码通过所述扩展芯片所在处理器安装位上的 QPI 接口发送；

所述处理器通过所述处理器所在处理器安装位上的 QPI 接口接收标识号码，判断所述标识号码是否为扩展芯片的标识号码，当所述标识号码为扩展芯片的标识号码时，为所述扩展芯片配置内存控制器，在所述处理器的地址空间中分配与所述扩展芯片相连接的扩展内存相对应的内存地址。

8、根据权利要求 7 所述的方法，其特征在于，所述方法还包括：

所述处理器判断是否向所述处理器所在处理器安装位上的所有 QPI 接口发送标识号码查询指令；如果否，所述处理器向未发送标识号码查询指令的 QPI 接口发送标识号码查询指令。

9、根据权利要求 7 或 8 所述的方法，其特征在于，所述方法还包括：

所述处理器生成用于读取数据或写入数据的、协议为 QPI 协议的报文，将所述报文通过所述处理器所在处理器安装位上的 QPI 接口发送；

所述扩展芯片通过所述扩展芯片所在处理器安装位上的 QPI 接口接收所述报文，将所述报文转换成内存接口协议后，根据协议转换后的报文在与所述扩展芯片相连接的扩展内存中读取数据或写入数据，将读取到的数据由内存接口协议转换成 QPI 协议后通过所述扩展芯片所在处理器安装位上的 QPI 接口发送；

所述处理器通过所述处理器所在处理器安装位上的 QPI 接口接收协议转换后的数据。

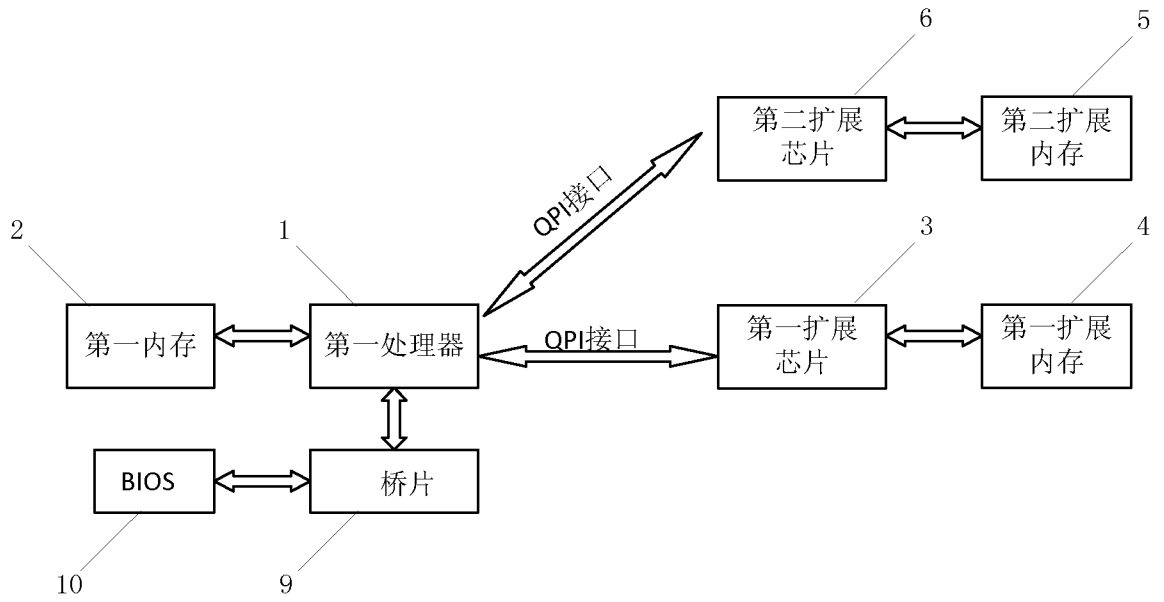


图 1

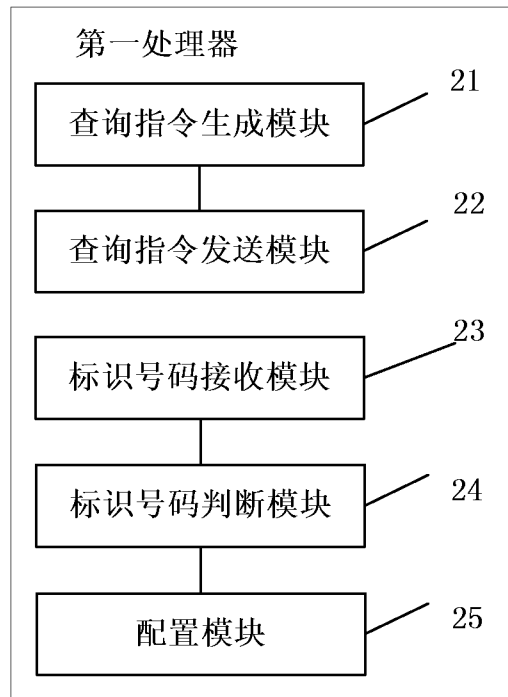


图 2

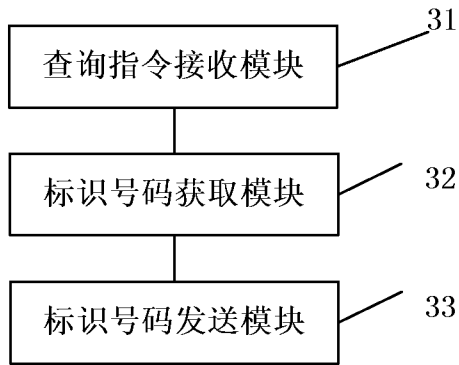


图 3

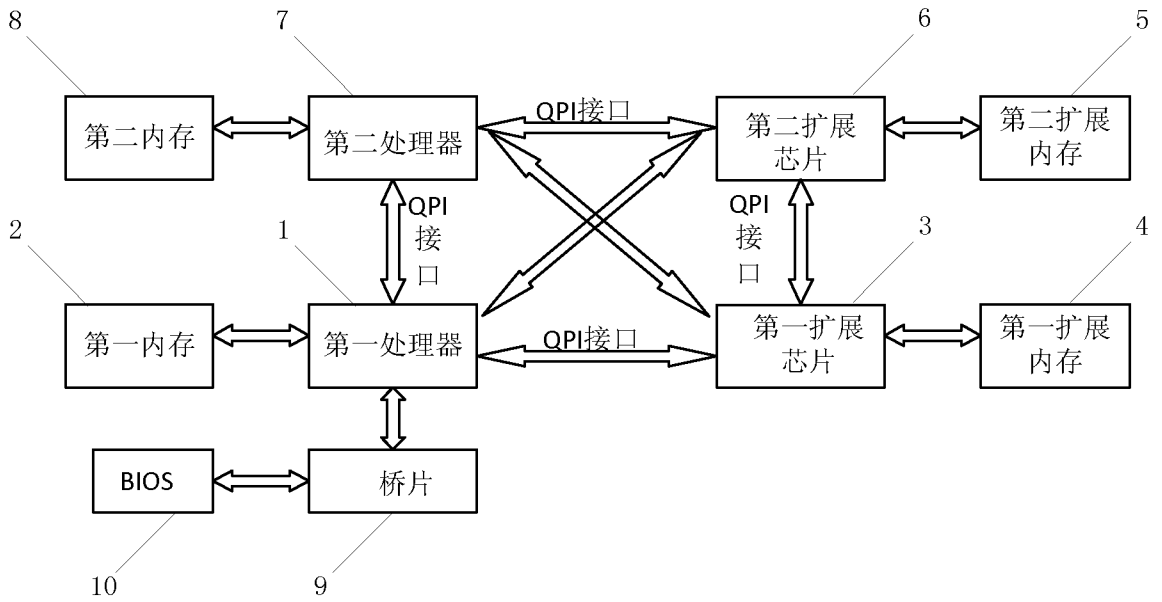


图 4

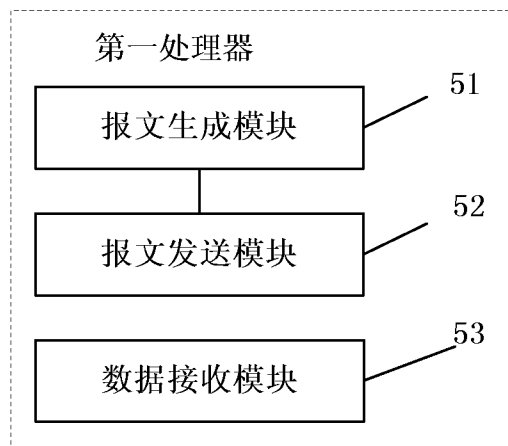


图 5

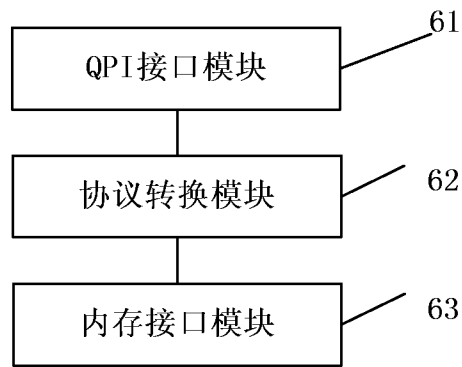


图 6

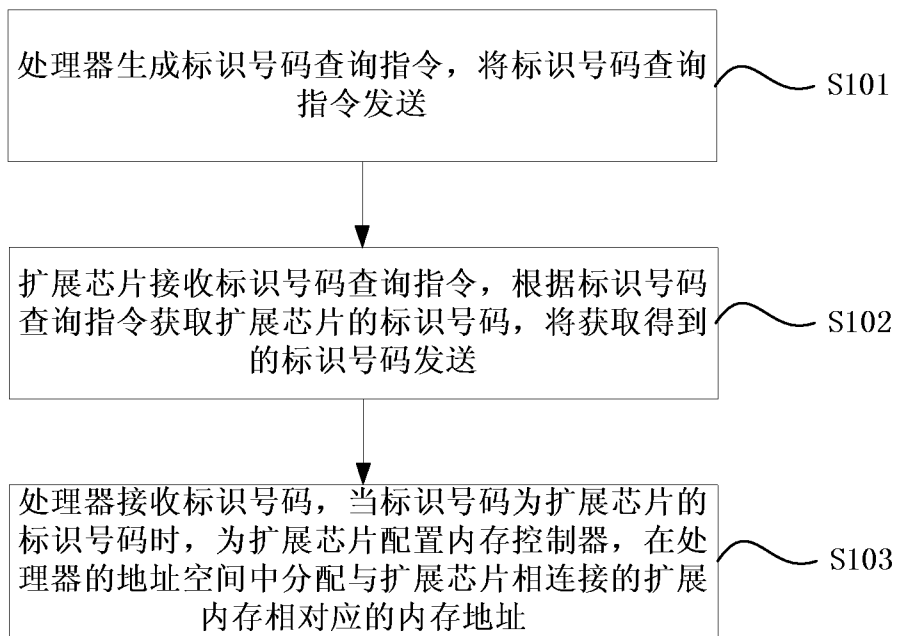


图 7

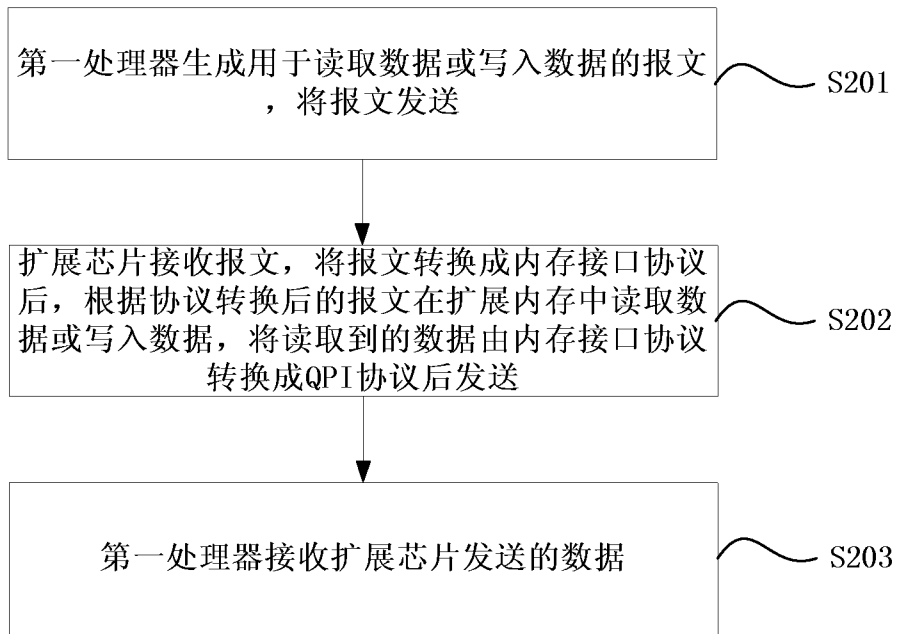


图 8

- 5/5 -

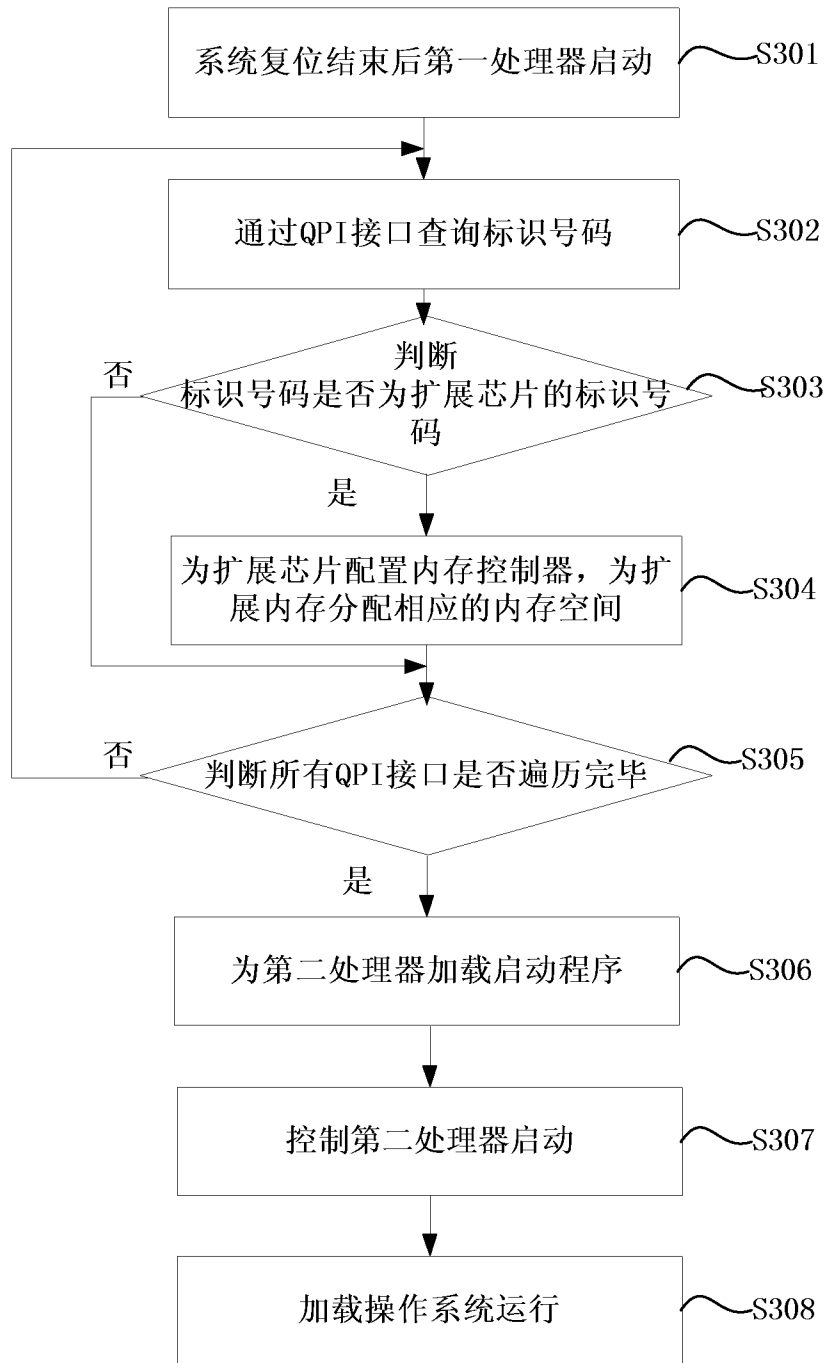


图 9

INTERNATIONAL SEARCH REPORT

International application No.

PCT/CN2014/076582

A. CLASSIFICATION OF SUBJECT MATTER

G06F 3/06 (2006.01) i

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)

G06F

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)

CNABS, VEN, EPODOC, CNKI, GOOGLE: memory, expen+, extent, QPI, quickpath, interconnect+, chip

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
PX	CN 103488436 A (HUAWEI TECHNOLOGIES CO., LTD.) 01 January 2014 (01.01.2014) claims 1 to 9	1-9
Y	CN 102473169 A (INTEL CORP.) 23 May 2012 (23.05.2012) description, paragraphs [0017] to [0025], [0043] and [0064], and figures 1 and 2	1-9
Y	CN 101872308 A (HONGFUJIN PRECISION INDUSTRY SHENZHEN CO., LTD.) 27 October 2010 (27.10.2010) description, paragraphs [0003] to [0010], and figure 1	1-9
A	CN 101872213 A (INSPUR ELECTRONIC INFORMATION INDUSTRY CO., LTD.) 27 October 2010 (27.10.2010) the whole document	1-9
A	WO 2013081579 A1 (INTEL CORPORATION) 06 June 2013 (06.06.2013) the whole document	1-9

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

27 June 2014

Date of mailing of the international search report

30 July 2014

Name and mailing address of the ISA
 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

YANG, Jishuang

Telephone No. (86-10) 62414422

INTERNATIONAL SEARCH REPORT
Information on patent family members

International application No.
PCT/CN2014/076582

Patent Documents referred in the Report	Publication Date	Patent Family	Publication Date
CN 103488436 A	01 January 2014	None	
CN 102473169 A	23 May 2012	US 2011161592 A1	30 June 2011
		JP 2012530327 A	29 November 2012
		WO 2011081840 A2	07 July 2011
		KR 20120026576 A	19 March 2012
		EP 2519892 A2	07 November 2012
		US 2011179311 A1	21 July 2011
CN 101872308 A	27 October 2010	US 2010274999 A1	28 October 2010
CN 101872213 A	27 October 2010	None	
WO 2013081579 A1	06 June 2013	None	

A. 主题的分类 G06F 3/06 (2006.01) i 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) G06F 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)) CNABS, VEN, EPODOC, CNKI, GOOGLE: 内存, 存储, 扩展, 扩充, 扩容, 增容, 快速通道互联, 快速路经互联, 芯片, memory, expen+, exten+, QPI, quickpath, interconnect+, chip		
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
PX	CN 103488436A (华为技术有限公司) 2014年 1月 01日 (2014 - 01 - 01) 权利要求1-9	1-9
Y	CN 102473169A (英特尔公司) 2012年 5月 23日 (2012 - 05 - 23) 说明书第[0017]-[0025]、[0043]、[0064]段, 附图1、2	1-9
Y	CN 101872308A (鸿富锦精密工业深圳有限公司等) 2010年 10月 27日 (2010 - 10 - 27) 说明书第[0003]-[0010]段, 附图1	1-9
A	CN 101872213A (浪潮电子信息产业股份有限公司) 2010年 10月 27日 (2010 - 10 - 27) 全文	1-9
A	WO 2013081579A1 (INTEL CORPORATION等) 2013年 6月 06日 (2013 - 06 - 06) 全文	1-9
<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:		
“A” 认为不特别相关的表示了现有技术一般状态的文件	“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	
“E” 在国际申请日的当天或之后公布的在先申请或专利	“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	
“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)	“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	
“O” 涉及口头公开、使用、展览或其他方式公开的文件	“&” 同族专利的文件	
“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 2014年 6月 27日	国际检索报告邮寄日期 2014年 7月 30日	
ISA/CN的名称和邮寄地址 中华人民共和国国家知识产权局(ISA/CN) 北京市海淀区蓟门桥西土城路6号 100088 中国 传真号 (86-10)62019451	授权官员 杨继爽 电话号码 (86-10)62414422	

国际检索报告
关于同族专利的信息

国际申请号

PCT/CN2014/076582

检索报告引用的专利文件		公布日 (年/月/日)	同族专利		公布日 (年/月/日)
CN	103488436A	2014年 1月 01日	无		
CN	102473169A	2012年 5月 23日	US	2011161592A1	2011年 6月 30日
			JP	2012530327A	2012年 11月 29日
			WO	2011081840A2	2011年 7月 07日
			KR	20120026576A	2012年 3月 19日
			EP	2519892A2	2012年 11月 07日
			US	2011179311A1	2011年 7月 21日
CN	101872308A	2010年 10月 27日	US	2010274999A1	2010年 10月 28日
CN	101872213A	2010年 10月 27日	无		
WO	2013081579A1	2013年 6月 06日	无		