



(12) 发明专利申请

(10) 申请公布号 CN 117178258 A

(43) 申请公布日 2023. 12. 05

(21) 申请号 202180097171.9

(51) Int.Cl.

(22) 申请日 2021.04.22

G06F 9/50 (2006.01)

(85) PCT国际申请进入国家阶段日
2023.10.17

(86) PCT国际申请的申请数据
PCT/JP2021/016313 2021.04.22

(87) PCT国际申请的公布数据
W02022/224409 JA 2022.10.27

(71) 申请人 日本电信电话株式会社
地址 日本东京都

(72) 发明人 横野智也 山部芳朗 石崎晃朗

(74) 专利代理机构 北京三友知识产权代理有限公司 11127

专利代理师 朱丽娟

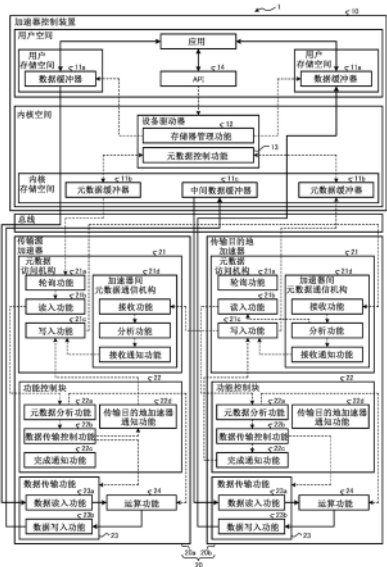
权利要求书2页 说明书10页 附图6页

(54) 发明名称

加速器控制系统、加速器控制方法和加速器
控制程序

(57) 摘要

在加速器 (20) 中, 读入功能 (21b) 从元数据缓冲器 (11b) 取得包含运算处理的对象处理数据的所在和指定加速器 (20) 的运算处理的内容的信息的元数据。运算功能 (24) 按照所取得的元数据中包含的运算处理的对象处理数据的所在和指定加速器 (20) 的运算处理的内容的信息, 对运算处理对象的处理数据进行运算处理。在元数据中包含有传输目的地的加速器 (20b) 的信息的情况下, 写入功能 (21c) 向该加速器 (20b) 传输该控制用数据。在从其他加速器 (20a) 接收到元数据的情况下, 运算功能 (24) 按照元数据中所包含的运算处理的对象的中间数据的所在和指定该运算处理的内容的信息, 进行针对运算处理的对象的中间数据的运算处理。



1. 一种加速器控制系统,其特征在于,具有加速器控制装置以及多个加速器,
所述加速器控制装置包括:

存储部,其存储控制用数据,所述控制用数据包括运算处理的对象数据的所在、要执行
运算处理的加速器的信息以及指定该加速器的运算处理的内容的信息;以及

判定部,其在所述存储部中存储有由所述加速器执行了运算处理的控制用数据的情况
下,判定为该加速器的运算处理完成,

所述加速器包括:

取得部,其从所述存储部取得所述控制用数据;

运算部,其根据所取得的所述控制用数据中包括的所述运算处理的对象数据的所在和
指定加速器的运算处理的内容的信息,对所述运算处理的对象数据进行运算处理;

保存部,其在所述运算处理完成的情况下,在所述控制用数据中不包含传输目的地的
加速器的信息的情况下,将所述控制用数据保存在所述存储部中;

传输部,其在所述控制用数据中包含有传输目的地的加速器的信息的情况下,将该控
制用数据传输给该加速器;以及

第二运算部,其在从其他加速器接收到所述控制用数据的情况下,按照该控制用数据
中包含的所述运算处理的对象数据的所在和指定该运算处理的内容的信息,对所述运算处
理的对象数据进行运算处理。

2. 根据权利要求1所述的加速器控制系统,其特征在于,所述加速器控制系统还具有确
认部,该确认部确认能否接收来自其他加速器的所述控制用数据。

3. 根据权利要求2所述的加速器控制系统,其特征在于,在规定的时间内未从传输目的
地的所述加速器通知能否接收所述控制用数据的情况下,所述传输部重发所述控制用数
据。

4. 根据权利要求1所述的加速器控制系统,其特征在于,在所述控制用数据被传输到传
输目的地的所述加速器的情况下,所述保存部将针对所述运算处理的对象数据的所述运算
处理的结果保存在所述存储部中,

在传输目的地的所述加速器中,所述第二运算部读入存储在所述存储部中的所述运算
处理的结果并作为所述运算处理的对象。

5. 根据权利要求1所述的加速器控制系统,其特征在于,所述加速器控制系统还具有监
视部,该监视部监视所述存储部中是否存在所述控制用数据。

6. 根据权利要求1所述的加速器控制系统,其特征在于,所述判定部监视在所述存储部
中是否存在完成了所述运算处理的控制用数据。

7. 一种加速器控制方法,由加速器控制系统执行,所述加速器控制系统具有加速器控
制装置和多个加速器,其特征在于,

所述加速器控制系统具有存储部,所述存储部存储控制用数据,所述控制用数据包含
运算处理的对象数据的所在、要执行运算处理的加速器的信息以及指定该加速器的运算处
理的内容的信息,

所述加速器控制方法包括如下步骤:

取得步骤,从所述存储部取得所述控制用数据;

运算步骤,按照所取得的所述控制用数据中包含的所述运算处理的对象数据的所在和

指定加速器的动作的信息,对所述运算处理的对象数据进行运算处理;

保存步骤,在所述运算处理完成且所述控制用数据中不包含传输目的地的加速器的信息的情况下,将所述控制用数据保存在所述存储部中;

判定步骤,在所述存储部中保存有完成了运算处理的控制用数据的情况下,判定为运算处理完成;

传输步骤,在所述控制用数据中包含有传输目的地的加速器的信息的情况下,向该加速器传输该控制用数据;以及

第二运算步骤,在从其他加速器接收到所述控制用数据的情况下,按照该控制用数据中包含的所述运算处理的对象数据的所在和指定该运算处理的内容的信息,对所述运算处理的对象数据进行运算处理。

8.一种加速器控制程序,所述加速器控制程序用于使计算机作为权利要求1至6中任一所述的加速器控制系统发挥功能。

加速器控制系统、加速器控制方法和加速器控制程序

技术领域

[0001] 本发明涉及加速器控制系统、加速器控制方法和加速器控制程序。

背景技术

[0002] 以往,已知有在运算处理系统中组装多个运算加速器(加速器)来使处理高速化的技术。由于半导体的微细化技术的极限,处理器的工作频率的提高、运算器的高集成化变得困难,因此通过该技术,专用于特定的处理而将处理转移(offload)到加速器,实现高性能的计算。进而,通过使多台加速器协调工作,实现由多个运算构成的处理的高速化(参照非专利文献1~非专利文献3)。

[0003] 具体地,需要控制将某个加速器的运算结果发送到其他加速器的加速器间通信。因此,在传输源加速器中的运算处理结束后,CPU进行加速器间的数据通信的控制。CPU在向传输目的地加速器发送了请求数据接受的准备的控制信号后,向传输源加速器发送传输开始的控制信号,由此驱动加速器间的数据通信。在加速器间的数据通信之后,传输目的地加速器对数据进行运算处理。

[0004] 现有技术文献

[0005] 非专利文献

[0006] 非专利文献1:J.Yang,D.B.Minturn,F.Hady,“When Poll is Better than Interrupt”,FAST,vol.12,pp.3-3,[online],2012年,[2021年3月29日检索],互联网<URL:<https://www.usenix.org/system/files/conference/fast12/yang.pdf>>

[0007] 非专利文献2:“CUDA Toolkit Documentation v11.2.2”,[online],NVIDIA,[2021年3月29日检索],因特网<URL:<https://docs.nvidia.com/cuda/>>

[0008] 非专利文献3:K.Gulati,S.P.Khatri,“GPU Architecture and the CUDA Programming Model”,Hardware acceleration of EDA algorithms,pp.23-30,[online],Springer,[2021年3月29日检索],因特网<URL:https://link.springer.com/content/pdf/10.1007%2F978-1-4419-0944-2_3.pdf>

发明内容

[0009] 发明所要解决的课题

[0010] 然而,在现有技术中,有时难以使用多个加速器来使实时处理高速化。例如,加速器由CPU控制,因此每当发生加速器间的数据传输的处理时,都会发生传输源加速器・传输目的地加速器的排他控制、处理完成的中断通知、同步控制。这些处理成为开销,延迟时间增大。另外,在使用多个加速器的情况下,除了上述那样的开销以外,还存在由于总线的规格而无法在加速器间直接进行数据传输的情况,因此可扩展性降低。因此,难以在实时处理或各种处理的领域中使用加速器。

[0011] 本发明是鉴于上述情况而完成的,其目的在于使用多个加速器来使实时处理高速化。

[0012] 用于解决课题的手段

[0013] 为了解决上述问题并达到目的,本发明的加速器控制系统具有加速器控制装置以及多个加速器,其特征在于,所述加速器控制装置包括:存储部,其存储控制用数据,所述控制用数据包括运算处理的对象数据的所在、要执行运算处理的加速器的信息以及指定该加速器的运算处理的内容的信息;以及判定部,其在所述存储部中存储有由所述加速器执行了运算处理的控制用数据的情况下,判定为该加速器的运算处理完成,所述加速器包括:取得部,其从所述存储部取得所述控制用数据;运算部,其根据所取得的所述控制用数据中包括的所述运算处理的对象数据的所在和指定加速器的运算处理的内容的信息,对所述运算处理的对象数据进行运算处理;保存部,其在所述运算处理完成的情况下,在所述控制用数据中不包含传输目的地的加速器的信息的情况下,将所述控制用数据保存在所述存储部中;传输部,其在所述控制用数据中包含有传输目的地的加速器的信息的情况下,将该控制用数据传输给该加速器;以及第二运算部,其在从其他加速器接收到所述控制用数据的情况下,按照该控制用数据中包含的所述运算处理的对象数据的所在和指定该运算处理的内容的信息,对所述运算处理的对象数据进行运算处理。

[0014] 发明效果

[0015] 根据本发明,能够使用多个加速器使实时处理高速化。

附图说明

[0016] 图1是例示本实施方式的加速器控制系统的概略结构的示意图。

[0017] 图2是用于说明本实施方式的加速器控制系统的图。

[0018] 图3是例示元数据的数据结构的图。

[0019] 图4是例示元数据缓冲器的结构的图。

[0020] 图5是用于说明数据缓冲器的图。

[0021] 图6是用于说明加速器间通信的图。

[0022] 图7是示出加速器中的加速器控制处理过程的流程图。

[0023] 图8是示出执行加速器控制程序的计算机的一例的图。

具体实施方式

[0024] 以下,参照附图,详细说明本发明的一个实施方式。另外,本发明并不限于该实施方式。另外,在附图的记载中,对同一部分标注同一标号来表示。

[0025] [加速器控制系统的结构]

[0026] 图1是例示本实施方式的加速器控制系统的概略结构的示意图。如图1所示,加速器控制系统1包括加速器控制装置10和多个加速器20。加速器控制装置10和加速器20能够经由总线布线等通信线路相互进行数据通信。

[0027] 加速器控制装置10是使用CPU(中央处理单元)等通用处理器实现的信息处理装置。加速器控制装置10在用户空间中执行作为由用户定义的处理的应用。另外,在内核空间上进行加速器20等物理设备的工作、管理、用户空间处理中的物理设备的抽象化处理。

[0028] 用户存储空间、内核存储空间由RAM(Random Access Memory:随机存取存储器)、闪存(Flash Memory)等半导体存储器元件、或者硬盘、光盘等存储装置实现,作为存储部发

挥功能。使加速器控制装置10工作的处理程序、在处理程序的执行过程中使用的数据等被预先存储在存储部中,或者在每次处理时被临时存储。

[0029] 使用通用处理器实现的控制部通过执行存储于存储器的处理程序,如图1所例示的那样,作为设备驱动器的存储器管理功能12、元数据控制功能13、后述的API14等各功能部发挥功能。此外,控制部也可以具备未图示的其他功能部。

[0030] 在加速器控制装置10中,运算处理的对象数据(下文中也称为处理数据)被设置在用户空间中的用户存储空间中的数据缓冲器11a中。此外,在内核空间中的内核存储空间中的元数据缓冲器11b中设置控制用数据(下文中也称为元数据),该控制用数据包括运算处理的对象数据的所在和指定加速器的运算处理的内容的信息。

[0031] 加速器20是由GPU(Graphics Processing Unit:图形处理单元)、FPGA(Field Programmable Gate Array:现场可编程门阵列)、DPU(Data Processing Unit:数据处理单元)等实现的运算加速器,如图1所例示的那样,作为元数据访问机构21、功能控制块22、数据传输功能23、运算功能24等各功能部发挥功能。

[0032] 加速器20读入处理数据和元数据,实现运算处理的转移。具体地,加速器20通过对元数据缓冲器11b进行自主轮询来监视元数据是否被配置。加速器20在配置有元数据的情况下,读入所配置的元数据并进行分析来驱动加速器20。加速器20将针对处理数据的运算处理的结果存储在数据缓冲器11a中。

[0033] 加速器控制装置10可以与加速器20非同步地进行到下一处理。当运算处理完成时,加速器20在元数据缓冲器11b中设置元数据。因此,加速器控制装置10认为运算处理完成。加速器控制装置10从数据缓冲器11a取得针对处理数据的运算处理的结果,并且进行到下一处理。

[0034] 具体地,加速器控制装置10通过经由安装在设备驱动器中的元数据控制功能13将包括处理数据的位置和运算处理的内容的元数据设置在内核存储空间中的元数据缓冲器11b中,来确定加速器20的工作。

[0035] 加速器控制装置10经由存储器管理功能12在用户存储空间中生成加速器控制装置10和加速器20能够参考的数据缓冲器11a,并且配置处理数据。

[0036] 加速器20具有元数据访问机构21。元数据访问机构21具有对内核存储空间中的元数据缓冲器11b进行轮询的功能、读入元数据的功能以及将元数据写入元数据缓冲器的功能,并且能够自主地读入元数据,而与来自加速器控制装置10的明确控制信号无关。

[0037] 在加载设备驱动器时,通过加速器初始化功能向加速器20通知元数据缓冲器11b的地址等的信息。

[0038] 加速器20从元数据缓冲器11b读入元数据,元数据分析功能22a对数据进行分析,并按照其描述进行工作。加速器20基于元数据中包含的处理数据的地址、长度的信息,驱动数据传输功能23中的数据读入功能23a,从数据缓冲器11a读出处理数据,用于运算功能24中的运算。在运算功能24中,基于元数据中描述的运算信息来执行运算处理。元数据分析功能22a将分析结果传递给运算功能24,设定运算功能24的运算处理内容。

[0039] 在运算处理结束后,从运算功能24将运算处理的结果传输给数据写入功能23b,写入到由数据传输控制功能22b指定的数据缓冲器中。然后,从数据传输功能23向完成通知功能22c发送数据传输结束的通知,通过元数据访问机构21的写入功能将元数据写入到元数

据缓冲器11b,加速器20的工作完成。

[0040] 这样,在加速器控制系统1中,将用于控制数据传输、运算处理的元数据配置于主存储器,加速器20自主地进行轮询和读入。由此,能够减少中断等、加速器控制的开销,进行实时处理。

[0041] 另外,加速器20能够将处理数据直接传输到加速器控制装置10的用户空间,加速器控制装置10和加速器20能够非同步地驱动,因此,延迟和CPU使用率都被削减。

[0042] 此外,由于加速器20可以通过访问加速器控制装置10的主存储器来工作,因此可以扩展到多个加速器20,可扩展性大大提高。

[0043] 另外,在使用多个加速器20进行运算处理的情况下,元数据中还包含进行运算的加速器的信息、加速器中的操作、读入的处理数据的所在等信息。基于这些信息,传输源加速器20a生成传输目的地加速器20b的设定、控制信息(以下称为元数据或加速器间元数据),并且在加速器之间直接执行元数据的数据通信,而不经由加速器控制装置10。

[0044] 在完成运算处理之后,传输源加速器20a将运算处理结果输出到内核存储空间中的中间数据缓冲器11c,并且将元数据发送到传输目的地加速器20b。

[0045] 传输目的地加速器20b根据加速器20b的当前状态判定是否能够接收所发送的元数据。当能够接收元数据时,传输目的地加速器20b向传输源加速器20a通知接收完成。在从传输目的地加速器20b通知接收完成之前,传输源加速器20a处于待机状态,在规定的时间内没有接收到完成的的通知的情况下,再次向传输目的地加速器20b发送元数据。

[0046] 传输目的地加速器20b通过分析接收到的元数据,读入中间数据缓冲器11c的中间数据,执行运算处理。

[0047] 具体地,当多个加速器20协作时,在元数据中记述有助于在多个加速器20中工作的信息。加速器20通过对由元数据分析功能22a读入的元数据进行分析,在运算处理结束后进行加速器间的数据通信。

[0048] 首先,传输源加速器20a在运算处理结束后,向内核存储空间上的中间数据缓冲器11c输出用于传输给传输目的地加速器20b的处理数据的运算结果。

[0049] 之后,传输源加速器20a通过传输目的地加速器通知功能22d,经由总线向传输目的地加速器20b发送包含运算处理结果的位置信息、运算内容、与加速器的协作相关的信息的元数据。

[0050] 传输目的地加速器20b通过加速器间元数据通信机构21d的接收功能接收元数据。元数据被发送到分析功能,并且基于加速器的状态来判断元数据是否被接收。此时,加速器间元数据通信机构21d的接收通知功能生成关于是否接收到的信息,并返回给传输源加速器20a。

[0051] 待机状态的传输源加速器20a通过从传输目的地加速器20b接收到接收完成的的通知,视为传输完成,并转移到下一处理。当在规定的时间内没有接收完成的的通知时,传输源加速器20a再次发送元数据。

[0052] 传输目的地加速器20b如果能够接收被传输的元数据,则基于从加速器间元数据通信机构21d的分析功能读入的元数据,执行运算处理。

[0053] 如上所述,加速器控制系统1在使用多个加速器20的处理中,由于不经由加速器控制装置10而执行处理,因此减少了传输源加速/传输目的地加速的排他处理、处理完成的中

断通知、同步控制等的开销,并且不会降低可扩展性。另外,在通信处理中,加速器控制装置10也能够与各加速器20不同步地进行驱动,因此能够有效地利用各自的运算资源。通过在处理数据的交换中使用主存储器,即使没有加速器控制装置10的控制,也可以从多个加速器20参考数据源。因此,能够与总线的规格无关地进行数据传输,能够构建可扩展且应对广泛的工作量的系统。

[0054] [加速器控制装置]

[0055] 图2是用于说明本实施方式的加速器控制系统的图。如图2所例示,在加速器控制装置10中,API14是能够从应用调用的功能组,提供加速器20的工作所需的功能。例如,API作为数据缓冲器确保功能、数据缓冲器释放功能、元数据生成功能、元数据发送功能、运算完成感知功能、通信用元数据生成功能发挥功能。

[0056] 数据缓冲器确保功能生成用于与加速器20进行数据的收发的存储区域即数据缓冲器11a。数据缓冲器确保功能调用存储器管理功能12,从存储器管理功能12取得数据缓冲器11a的区域信息。

[0057] 数据缓冲器释放功能删除数据缓冲器11a。数据缓冲器释放功能调用存储器管理功能12,通知要释放的数据缓冲器11a的区域信息,释放该区域。

[0058] 元数据生成功能生成被附加了加速器20的控制信息的控制用数据即元数据。元数据发送功能将元数据生成功能生成的元数据发送到元数据控制功能13。运算完成感知功能调用元数据控制功能13,根据元数据控制功能13的返回值,确认任意的加速器20的处理是否结束。通信用元数据生成功能生成用于执行加速器间通信的通信路径信息。

[0059] 设备驱动器存在于内核空间上,提供用于处理特定的设备的存储器管理功能12、元数据控制功能13等功能组。

[0060] 存储器管理功能12被用户空间上的API14调用,进行数据缓冲器11a的区域的确保、释放、管理。

[0061] 元数据缓冲器11b是用于与加速器20进行元数据的发送接收的存储区域,通过加速器初始化功能,根据加速器20的数量而生成。例如,元数据缓冲器11b存储控制用数据,该控制用数据包括运算处理的对象数据的所在、要进行运算处理的加速器20的信息以及指定加速器20的运算处理的内容的信息。

[0062] 元数据缓冲器11b包括用于向加速器20发送元数据的RQ(请求队列)缓冲器和用于从加速器20接收元数据的CQ(完成队列)缓冲器。

[0063] 元数据控制功能13被用户空间上的API14调用,进行元数据缓冲器11b向RQ缓冲器/CQ缓冲器的读入、写入、管理。元数据控制功能13具有轮询功能13a和写入功能13b。

[0064] 轮询功能13a对CQ缓冲器进行轮询,确认运算的完成。也就是说,轮询功能13a作为判定部发挥功能,监视已经由加速器20执行了运算处理的元数据是否存储在元数据缓冲器11b中。当已经由加速器20执行了运算处理的元数据被存储在元数据缓冲器11b中时,轮询功能13a判定为加速器20的运算处理完成。另外,写入功能13b对RQ缓冲器进行写入。

[0065] [加速器]

[0066] 在加速器20中,元数据访问机构21包括轮询功能21a、读入功能21b、写入功能21c和加速器间元数据通信机构21d,作为访问内核空间中的元数据缓冲器11b的功能。

[0067] 轮询功能21a轮询元数据缓冲器11b的RQ缓冲器,确认在RQ缓冲器中是否存在元数

据。即,轮询功能21a作为监视部发挥功能,监视元数据缓冲器11b中是否存在元数据。

[0068] 读入功能21b作为取得部发挥功能,从元数据缓冲器11b取得元数据。即,当轮询功能21a确认在RQ缓冲器中存在元数据时,读入功能21b读入元数据并将元数据传输到功能控制块22。

[0069] 写入功能21c向元数据缓冲器11b的CQ缓冲器写入元数据。即,写入功能21c作为保存部发挥功能,在运算处理完成且元数据中不包含传输目的地的加速器20b的信息的情况下,将元数据存储在元数据缓冲器11b中。

[0070] 功能控制块22控制针对加速器20的各功能部的必要的数据共享、驱动・停止定时。功能控制块22包括元数据分析功能22a、数据传输控制功能22b、完成通知功能22c和传输目的地加速器通知功能22d。

[0071] 元数据分析功能22a从由元数据访问机构21发送的元数据中提取工作、数据缓冲器11a的地址・长度的信息。数据传输控制功能22b提供对数据传输功能23中的数据的读入、写入进行控制的功能。完成通知功能22c向元数据访问机构21传输要写入CQ缓冲器的元数据。

[0072] 运算功能24作为运算部发挥功能,按照所取得的元数据中包含的运算处理的对象处理数据的所在和指定加速器20的运算处理的内容的信息,对运算处理的对象处理数据进行运算处理。

[0073] 具体而言,运算功能24包括输入控制功能、运算电路、输出控制功能。运算电路具有多个独立的运算电路,能够对各个运算电路定义运算。输入控制功能基于从功能控制块22发送的元数据的信息,将从数据传输功能23的数据读入功能23a发送的数据输入到适当的运算电路。

[0074] 输出控制功能将针对从运算电路输出的处理数据的运算处理的结果在适当的定时传输到数据传输功能23的数据写入功能23b。

[0075] 数据传输功能23的数据写入功能23b作为保存部发挥功能,将针对处理数据的运算处理的结果保存在数据缓冲器11a中。

[0076] 接着,图3是例示元数据的数据结构的图。如图3所示,元数据由32位的任务ID、32位的操作、32位的读出数据的长度、64位的读出数据的地址、32位的写入数据的长度、64位的写入数据的地址共计256位的数据构成。

[0077] 另外,图4是例示元数据缓冲器的结构的图。如图4所示,图3所示的元数据被存储在RQ缓冲器/CQ缓冲器中。图4所示的RQ缓冲器的头部(head)信息由元数据控制功能13控制。元数据控制功能13将从API14的元数据发送功能发送的元数据成形为图3所示的结构,并存储在RQ缓冲器中后,将头部(head)的信息偏移到下一个位置。

[0078] 同样地,元数据控制功能13在CQ缓冲器中也使用头部(head)的信息,在被API14的运算完成感知功能调用的情况下,从CQ缓冲器读出元数据,将头部(head)的信息偏移到下一个位置。

[0079] 另外,图5是用于说明数据缓冲器的图。存储器管理功能12通过图5所例示的表来管理数据缓冲器11a的地址和状态。如图5所示,在表中管理数据缓冲器11a的物理地址及其地址的状态。当对加速器20设置元数据时,通过参考关键字(key)取得期望的数据缓冲器11a的物理地址,并且设置读出数据地址和写入数据地址。由此,能够吸收并管理内核空间

和用户空间中的虚拟地址的差异。

[0080] 返回图2的说明。上述的写入功能21c作为传输部发挥功能,在元数据中包含传输目的地的加速器20b的信息的情况下,向该加速器20b传输该元数据。

[0081] 加速器间元数据通信机构21d包括接收功能、分析功能和接收通知功能,并且提供加速器间通信所需的功能。接收功能作为确认部发挥功能,确认能否接收来自其他加速器20a的元数据。即,接收功能生成用于通知能否接收从其他加速器20(传输源加速器20a)发送来的元数据的数据,并发送给传输源加速器20a。

[0082] 功能控制块22的传输目的地加速器通知功能22d提供向传输目的地加速器20b通知元数据的功能。元数据分析功能22a从由元数据访问机构21发送的元数据中提取操作、数据缓冲器的地址・长度的信息。此外,元数据分析功能22a从元数据中提取传输目的地加速器20b的信息。另外,元数据分析功能22a在没有来自传输目的地加速器20b的接收完成的通知的情况下,经由传输目的地加速器通知功能22d和写入功能21c重发元数据。写入功能21c在规定的时间内未从传输目的地的加速器20b通知能否接收元数据的情况下,重发元数据。

[0083] 在传输目的地的加速器20b中,运算功能24作为第二运算部发挥功能,如后所述,在从其他加速器20a接收到元数据的情况下,运算功能24按照元数据中所包含的运算处理的对象的中间数据的所在和指定该运算处理的内容的信息,对运算处理的对象的中间数据进行运算处理。

[0084] 数据传输功能23的数据写入功能23b在向传输目的地的加速器20b传输了元数据的情况下,将针对运算处理对象的处理数据的运算处理的结果的中间数据存储于中间数据缓冲器11c。在传输目的地的加速器20b中,运算功能24读入存储在中间数据缓冲器11c中的运算处理的结果并作为运算处理的对象。

[0085] 这里,图6是用于说明加速器间通信的图。加速器控制系统1使用元数据来实现加速器间通信。如图6的(a)所示,加速器间通信的通信路径以将各加速器20设为顶点 V_0 、 V_1 、 V_2 、 $\cdots$ 的图表来表现。通过将该图表表现为邻接矩阵,作为数据进行处理。例如,如图6的(b)所示,图6的(a)所例示的图表能够表现为邻接矩阵。

[0086] 通过设置成图6的(c)所示的数据结构,可以在加速器20中处理这些信息。在图6的(c)所示的数据结构中,开头的8位表示矩阵的大小,加速器20能够知道矩阵主体的数据的大小。接下来的256位是矩阵主体的数据,包含加速器间的连接信息。在接下来的字段中记述还有用于驱动传输目的地加速器20b的运算内容、处理数据的所在等信息。

[0087] [加速器控制处理]

[0088] 图7是示出加速器20中的加速器控制过程的流程图。图7的流程图例如在由应用指示了开始的定时开始。

[0089] 图7的流程图例如以规定的间隔开始。加速器20通过轮询来监视元数据缓冲器11b的RQ缓冲器(步骤S1)。加速器20确认是否已经从其他加速器20传输了元数据(步骤S2)。

[0090] 在未从其他加速器20传输元数据的情况下(步骤S2,否),加速器20使处理进入步骤S3。另一方面,在从其他加速器20传输了元数据的情况下(步骤S2,是),加速器20(传输目的地加速器20b)向传输源加速器20a发送元数据的接收完成的通知,并且分析元数据(步骤S4),使处理进入步骤S5。

[0091] 在步骤S3中,在RQ缓冲器中未设置有元数据的情况下(步骤S3,否),加速器20使处

理返回到步骤S1。另一方面,在RQ缓冲器中设置有元数据的情况下(步骤S3,是),加速器20读入并分析元数据(步骤S5)。

[0092] 另外,加速器20基于分析结果取得数据缓冲器11a的处理数据(步骤S6),并且执行运算处理(步骤S7)。

[0093] 当不存在作为传输目的地的加速器20时(步骤S8,否),加速器20将经过运算处理的处理数据写入数据缓冲器11a(步骤S9),将元数据发送到CQ缓冲器(步骤S10),完成处理。然后,加速器20将处理返回至步骤S1。

[0094] 另一方面,在存在传输目的地的加速器20的情况下(步骤S8,是),加速器20作为传输源加速器20a,将处理数据的运算处理结果写入中间数据缓冲器11c(步骤S11),并向传输目的地加速器20b发送元数据(步骤S12)。然后,加速器20在从传输目的地加速器20b接收到接收完成的的通知的情况下,视为元数据的传输完成,使处理返回步骤S1。

[0095] [效果]

[0096] 如上所述,在本实施方式的加速器控制系统1中,在加速器控制装置10中,元数据缓冲器11b存储控制用数据,该控制用数据包括运算处理的对象数据的所在、要执行运算处理的加速器20的信息以及指定加速器20的运算处理的内容的信息。当由加速器20执行了运算处理的元数据被存储在元数据缓冲器11b中时,轮询功能13a判定为加速器20的运算处理完成。

[0097] 此外,在加速器20中,读入功能21b从元数据缓冲器11b取得元数据。运算功能24根据所取得的元数据中包含的运算处理的对象处理数据的所在和指定加速器20的运算处理的内容的信息,对运算处理对象的处理数据进行运算处理。在运算处理完成的情况下,且在元数据中不包含传输目的地的加速器的信息的情况下,写入功能21c将元数据储存于元数据缓冲器11b。在元数据中包含传输目的地的加速器20b的信息的情况下,写入功能21c向该加速器20b传输该控制用数据。在从其他加速器20a接收到元数据的情况下,运算功能24按照元数据所包含的运算处理的对象的中间数据的所在和指定该运算处理的内容的信息,进行针对运算处理的对象的中间数据的运算处理。

[0098] 由此,加速器控制系统1在使用了多个加速器20的处理中,不经由加速器控制装置10而执行处理,因此能够减少传输源加速/传输目的地加速的排他处理、处理完成的中断通知、同步控制等的开销,能够不降低可扩展性而高速地进行实时处理。

[0099] 此外,加速器间元数据通信机构21d确认是否能够从其他加速器20a接收元数据。由此,能够可靠地执行加速器间的数据传输。

[0100] 另外,在规定的时间内未从传输目的地的加速器20b通知能否接收元数据的情况下,写入功能21c重发元数据。由此,能够可靠地执行加速器间的数据传输。

[0101] 另外,数据写入功能23b在向传输目的地的加速器20b传输了元数据的情况下,将针对运算处理的对象处理数据的运算处理的结果的中间数据储存于中间数据缓冲器11c。在传输目的地的加速器20b中,运算功能24读入存储在中间数据缓冲器11c中的运算处理的结果并作为运算处理的对象。由此,加速器间的数据传输高效地完成。

[0102] 此外,轮询功能21a监视元数据缓冲器11b中是否存在元数据。因此,加速器20能够自主地执行处理。

[0103] 此外,轮询功能13a监视在元数据缓冲器11b中是否存在经过了运算处理的元数

据。因此,加速器控制装置10可以与加速器20非同步地执行处理。

[0104] [程序]

[0105] 还可以生成以计算机可执行的语言来描述由上述实施方式的加速器控制装置10执行的处理的程序。在一个实施方式中,加速器控制装置10可以通过将执行加速器控制处理的加速器控制程序安装在期望的计算机中作为包软件或在线软件来实现。例如,通过使信息处理装置执行上述加速器控制程序,信息处理装置可以用作加速器控制装置10。在此所说的信息处理装置包括台式或笔记本型的个人计算机。另外,除此之外,信息处理装置还包括智能手机、移动电话机、PHS(Personal Handyphone System:个人手持电话系统)等移动通信终端、以及PDA(Personal Digital Assistant:个人数字助理)等平板终端等。加速器控制装置10的功能可以在云服务器中实现。

[0106] 图8是示出执行加速器控制程序的计算机的示例的图。计算机1000例如具有存储器1010、CPU 1020、硬盘驱动器接口1030、盘驱动器接口1040、串行端口接口1050、视频适配器1060以及网络接口1070。这些各部通过总线1080连接。

[0107] 存储器1010包括ROM(只读存储器)1011和RAM 1012。ROM 1011存储引导程序,例如基本输入输出系统(BIOS)。硬盘驱动器接口1030与硬盘驱动器1031连接。盘驱动器接口1040与盘驱动器1041连接。在盘驱动器1041中例如插入磁盘、光盘等可装卸的存储介质。例如,鼠标1051和键盘1052连接到串行端口接口1050。例如,显示器1061连接到视频适配器1060。

[0108] 这里,硬盘驱动器1031例如存储OS1091、应用程序1092、程序模块1093和程序数据1094。上述实施方式中说明的各信息例如存储于硬盘驱动器1031、存储器1010。

[0109] 另外,加速器控制程序例如作为记述有由计算机1000执行的指令的程序模块1093而存储于硬盘驱动器1031。具体地,描述由上述实施方式中描述的加速器控制装置10执行的处理的程序模块1093被存储在硬盘驱动器1031中。

[0110] 另外,加速器控制程序的信息处理所使用的数据作为程序数据1094例如存储于硬盘驱动器1031。然后,CPU 1020根据需要将硬盘驱动器1031中存储的程序模块1093、程序数据1094读出到RAM 1012,执行上述的各步骤。

[0111] 此外,加速器控制程序所涉及的程序模块1093、程序数据1094不限于存储于硬盘驱动器1031的情况,例如也可以存储于可装卸的存储介质,并经由盘驱动器1041等由CPU 1020读出。或者,加速器控制程序所涉及的程序模块1093、程序数据1094也可以存储于经由LAN、WAN(Wide Area Network:广域网)等网络连接的其他计算机,并经由网络接口1070由CPU 1020读出。

[0112] 以上,对应用了由本发明人完成的发明的实施方式进行了说明,但本发明并不限于构成本实施方式的本发明的公开的一部分的记述以及附图。即,由本领域技术人员等基于本实施方式完成的其他实施方式、实施例以及运用技术等全部包含于本发明的范畴。

[0113] 标号说明

[0114] 10 加速器控制装置

[0115] 11a 数据缓冲器(存储部)

[0116] 11b 元数据缓冲器(存储部)

[0117] 11c 中间数据缓冲器(存储部)

- [0118] 12 存储器管理功能
- [0119] 13 元数据控制功能
- [0120] 13a 轮询功能(判定部)
- [0121] 13b 写入功能
- [0122] 14 API
- [0123] 20 加速器
- [0124] 20a 传输源加速器
- [0125] 20b 传输目的地加速器
- [0126] 21 元数据访问机构
- [0127] 21a 轮询功能(取得部)
- [0128] 21b 读入功能
- [0129] 21c 写入功能(保存部、传输部)
- [0130] 21d 加速器间元数据通信机构(确认部)
- [0131] 22 功能控制块
- [0132] 22a 元数据分析功能
- [0133] 22b 数据传输控制功能
- [0134] 22c 完成通知功能
- [0135] 22d 传输目的地加速器通知功能
- [0136] 23 数据传输功能
- [0137] 23a 数据读入功能
- [0138] 23b 数据写入功能
- [0139] 24 运算功能(运算部、第二传输部)

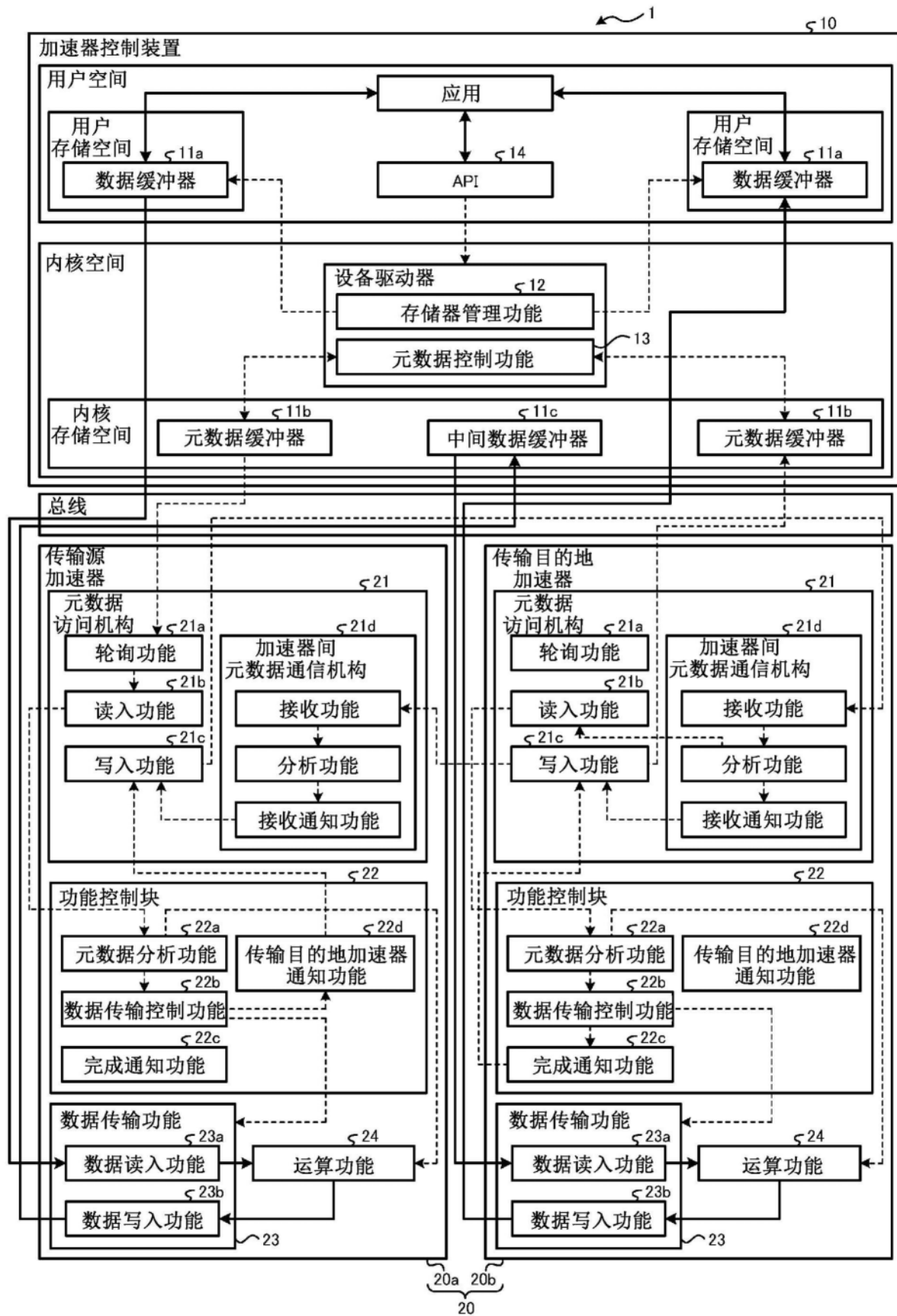


图1

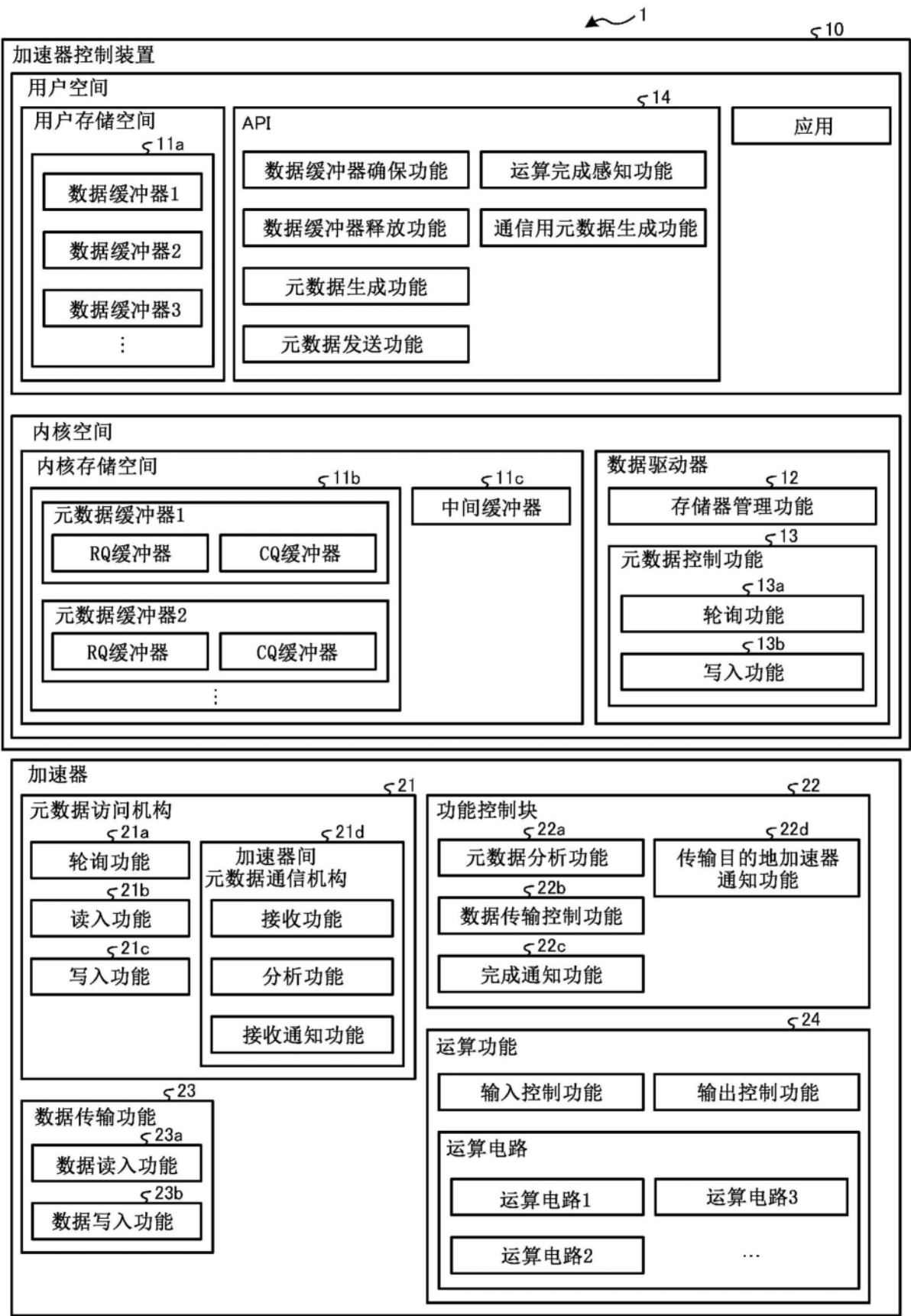


图2



图3

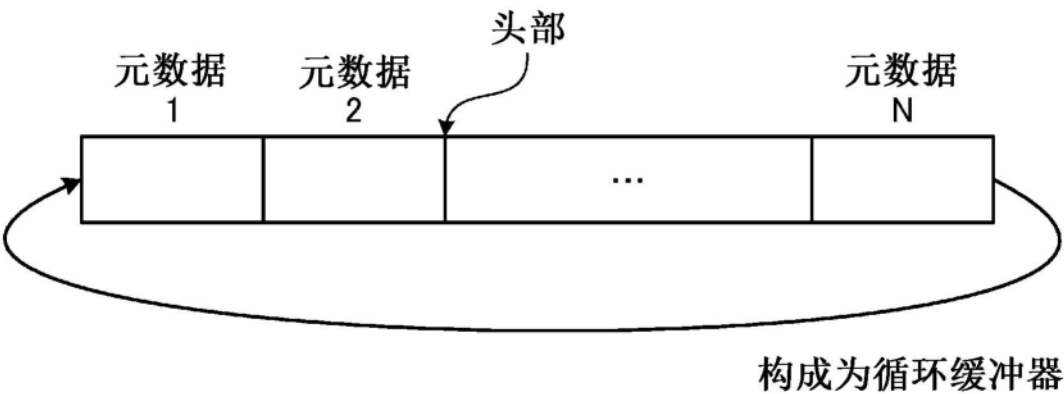


图4

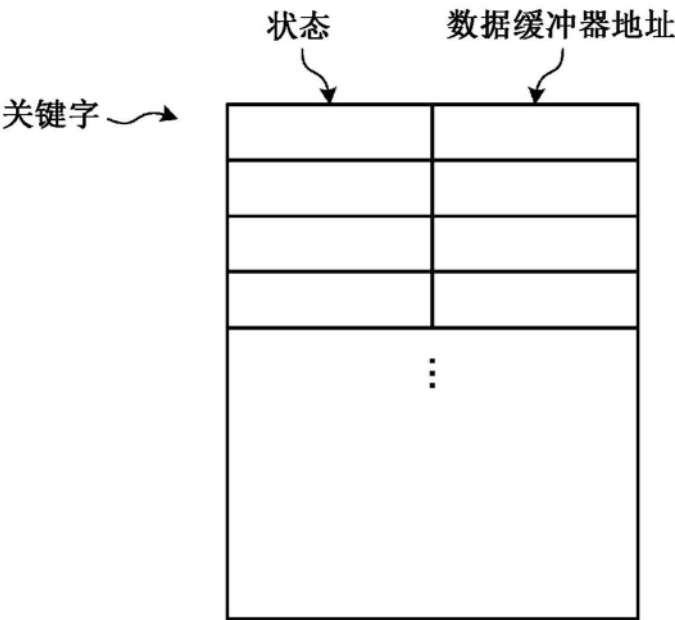


图5

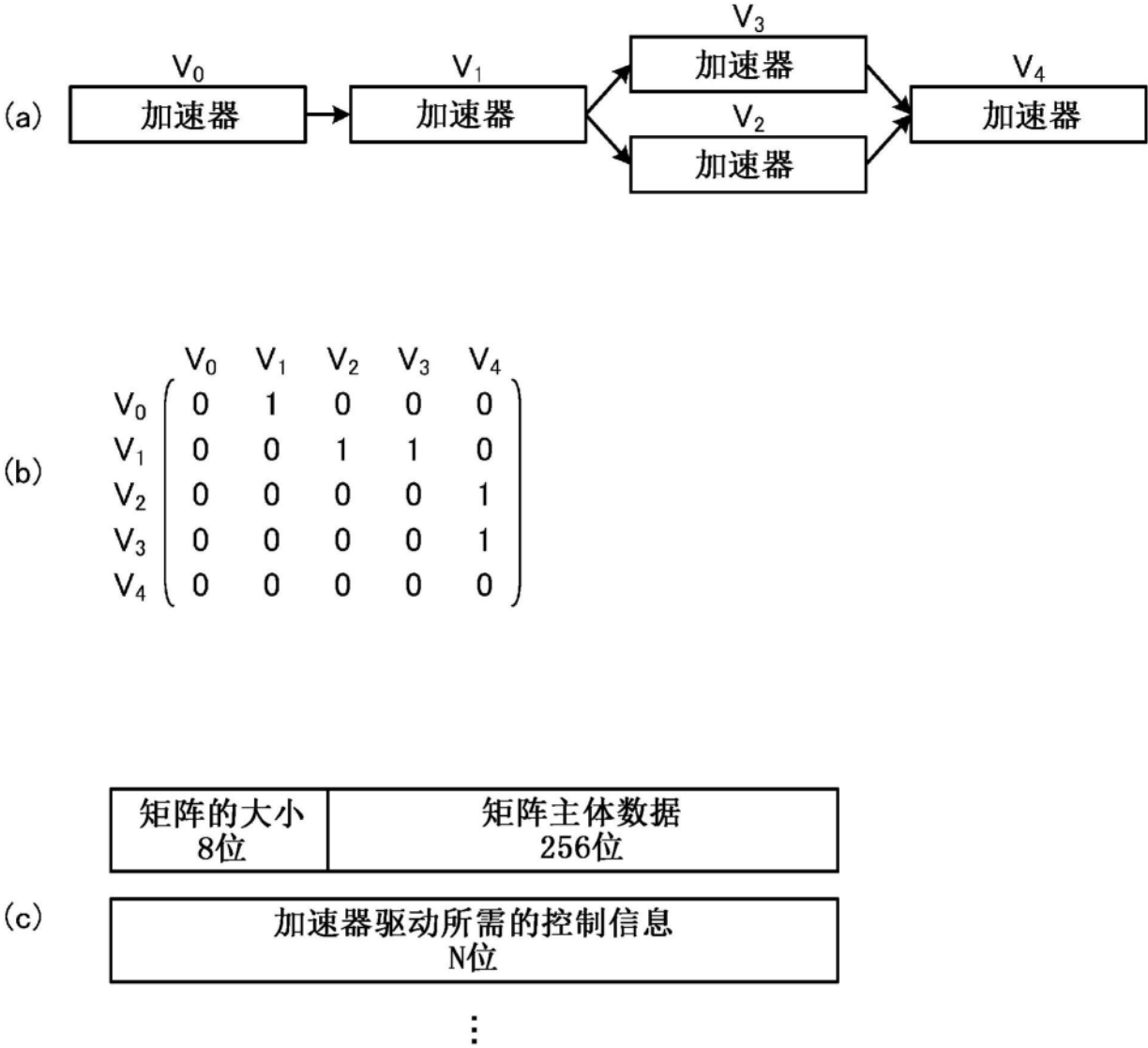


图6

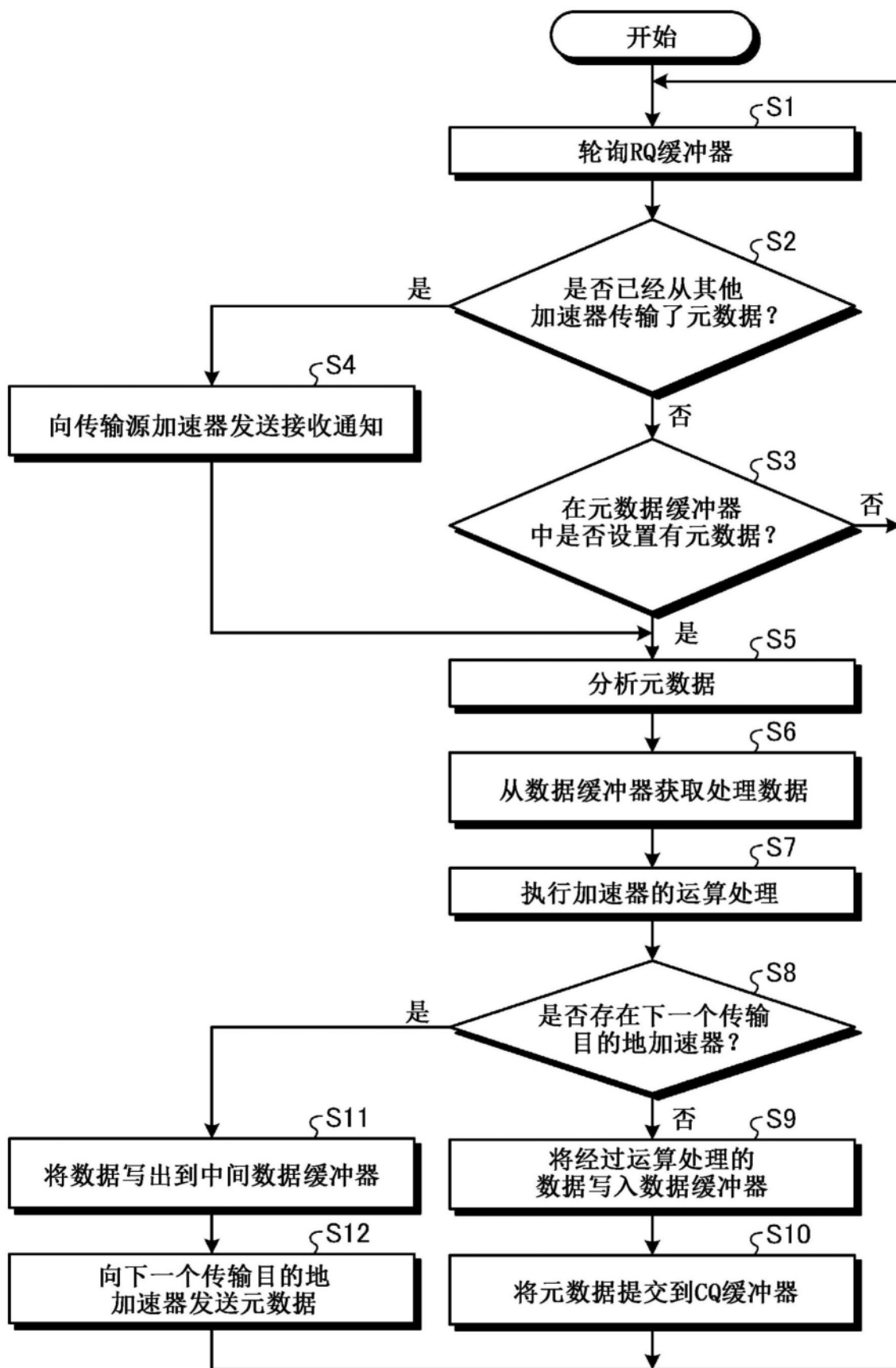


图7

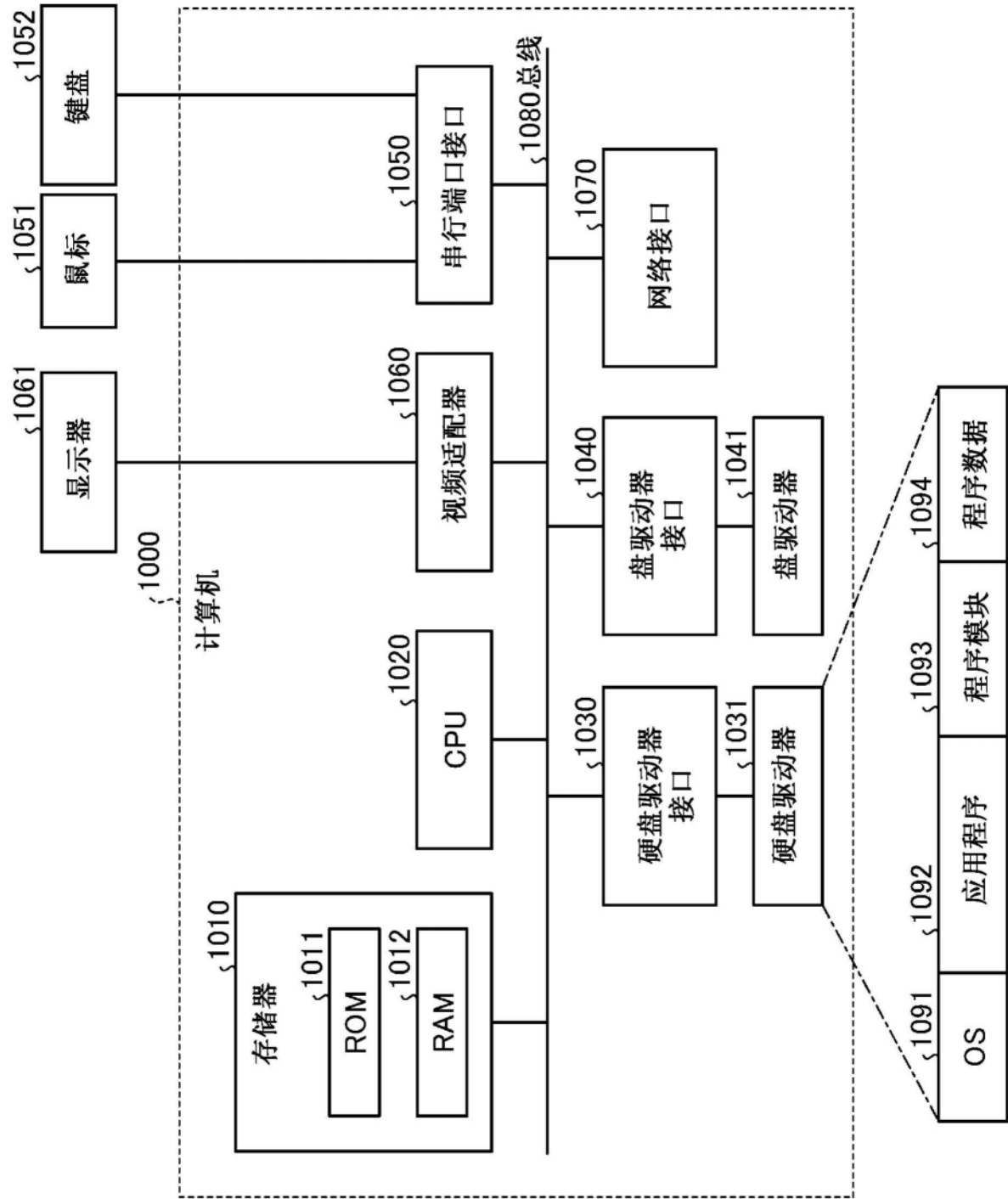


图8