



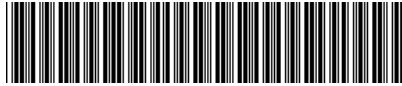
国家知识产权局

250014

山东省济南市历下区经十路 17703 号华特广场 B510 室 济南圣达知
识产权代理有限公司
于凤洋(0531-82961551)

发文日:

2024 年 11 月 20 日



申请号: 202411660753.5

发文序号: 2024112000920810

专利申请受理通知书

根据专利法第 28 条及其实施细则第 43 条、第 44 条的规定, 申请人提出的专利申请已由国家知识产权局受理。现将确定的申请号、申请日等信息通知如下:

申请号: 2024116607535

申请日: 2024 年 11 月 20 日

申请人: 山东省计算中心(国家超级计算济南中心)

发明人: 谭立状, 董鑫, 史慧玲, 张玮

发明创造名称: 基于多级流水线的数据库查询任务 RDMA 卸载方法及系统

经核实, 国家知识产权局确认收到文件如下:

权利要求书 1 份 2 页, 权利要求项数: 10 项

说明书 1 份 8 页

说明书附图 1 份 2 页

说明书摘要 1 份 1 页

发明专利请求书 1 份 5 页

实质审查请求书 文件份数: 1 份

申请方案卷号: 2024710337

提示:

1. 申请人收到专利申请受理通知书之后, 认为其记载的内容与申请人所提交的相应内容不一致时, 可以向国家知识产权局请求更正。

2. 申请人收到专利申请受理通知书之后, 再向国家知识产权局办理各种手续时, 均应当准确、清晰地写明申请号。

审查员: 自动受理

联系电话: 010-62356655

审查部门: 初审及流程管理部



200101
2023.03

纸件申请, 回函请寄: 100088 北京市海淀区蓟门桥西土城路 6 号 国家知识产权局专利局受理处收
电子申请, 应当通过专利业务办理系统以电子文件形式提交相关文件。除另有规定外, 以纸件等其他形式提交的文件视为未提交。

权利要求书

1. 基于多级流水线的数据库查询任务 RDMA 卸载方法，其特征在于，包括：
构建以 RDMA 网络单元为中心，包括计算单元、内存单元和 RDMA 网络单元的分布式计算节点架构；
基于以 RDMA 网络单元为中心的分布式计算节点架构，通过 RDMA 网络单元以及不同计算单元的内存单元，无需 CPU 介入直接完成数据库查询任务数据的转发卸载；
其中，数据库查询任务的计算包括单任务运算模式和多任务融合运算模式，
当下发的任务为多任务融合运算模式时，根据多任务中包含的基本操作选择相应的计算单元组成流水运算，中间产生的计算结果直接转发到该计算单元对应的内存单元后，经由 RDMA 网络单元直接缓存到下一级计算单元的内存单元中，下一级计算单元读取内存单元中的中间结果继续运算，依次类推直至最终结果返回主机节点的 CPU 内存。
2. 如权利要求 1 所述的基于多级流水线的数据库查询任务 RDMA 卸载方法，其特征在于，数据库查询任务的基本操作包括排序操作 P、聚合操作 J、过滤操作 G 以及连接操作 L，基本操作的计算单元包括排序计算单元、聚合计算单元、过滤计算单元及连接计算单元，多任务融合运算模式的基本操作为相应的计算单元组成的流水运算。
3. 如权利要求 1 所述的基于多级流水线的数据库查询任务 RDMA 卸载方法，其特征在于，所述单任务运算模式为单一的排序操作 P、聚合操作 J、过滤操作 G 或者连接操作 L，多任务融合运算模式包括过滤操作 G+排序操作 P、过滤操作 G+连接操作 L、过滤操作 G+排序操作 P+连接操作 L、过滤操作 G+聚合操作 J、排序操作 P+聚合操作 J 以及排序操作 P+连接操作 L。
4. 如权利要求 1 所述的基于多级流水线的数据库查询任务 RDMA 卸载方法，其特征在于，所述数据库查询任务由主机节点 Host 产生，数据库查询任务的下发为：通过调用内核 API 函数，将相关任务类型、目的计算单元的参数通过 PCIE 总线传输到计算单元内，同时 CPU 调用 XRT API 将待处理数据经由 PCIE DMA 缓存到内存单元 M，同时根据任务类型在不同的计算单元部署相应的计算单元，并生成启动信号。
5. 如权利要求 3 所述的基于多级流水线的数据库查询任务 RDMA 卸载方法，其特征在于，当单任务运算模式为连接操作 L 时，计算节点部署连接计算单元，内存单元 M4 分为多块存储区域分别为 M4-1、M4-2、M4-3、M4-4，其中 M4-1 和 M4-2 分别存储内表数据及外表数据。
6. 如权利要求 1 所述的基于多级流水线的数据库查询任务 RDMA 卸载方法，其特征在于，

权利要求书

当数据库下发查询中的单任务运算模式的基本操作任务时，计算系统调度对应的计算单元完成运算，并将结果返回到对应的内存单元，最终返回主机节点的 CPU 内存。

7. 基于多级流水线的数据库查询任务 RDMA 卸载系统，其特征在于，包括：

计算节点架构构建模块，用于构建以 RDMA 网络单元为中心的分布式计算节点架构；

数据库查询任务卸载模块，用于基于以 RDMA 网络单元为中心的分布式计算节点架构，通过 RDMA 网络以及不同计算单元的内存单元，无需 CPU 介入直接完成数据库查询任务数据的转发卸载；

其中，数据库查询任务的计算包括单任务运算模式和多任务融合运算模式，

当下发的任务为多任务融合运算模式时，根据多任务中包含的基本操作选择相应的计算单元组成流水运算，中间产生的计算结果直接转发到该计算单元对应的内存单元后，经由 RDMA 网络直接缓存到下一级计算单元的内存单元中，下一级计算单元读取内存单元中的中间结果继续运算，依次类推直至最终结果返回主机节点的 CPU 内存。

8. 一种计算机程序产品，包括计算机程序，其特征在于，所述计算机程序被处理器执行时实现权利要求 1-6 任一项所述的基于多级流水线的数据库查询任务 RDMA 卸载方法。

9. 一种非暂态计算机可读存储介质，其特征在于，所述非暂态计算机可读存储介质用于存储计算机指令，所述计算机指令被处理器执行时，实现如权利要求 1-6 任一项所述的基于多级流水线的数据库查询任务 RDMA 卸载方法。

10. 一种电子设备，其特征在于，包括：处理器、存储器以及计算机程序；其中，处理器与存储器连接，计算机程序被存储在存储器中，当电子设备运行时，所述处理器执行所述存储器存储的计算机程序，以使电子设备执行实现如权利要求 1-6 任一项所述的基于多级流水线的数据库查询任务 RDMA 卸载方法。

基于多级流水线的数据库查询任务 RDMA 卸载方法及系统

技术领域

[0001] 本公开涉及 RDMA 网络和数据库查询加速技术领域，具体涉及基于多级流水线的数据库查询任务 RDMA 卸载方法及系统。

背景技术

[0002] 本部分的陈述仅仅是提供了与本公开相关的背景技术信息，不必然构成在先技术。

[0003] 近年来，数据库系统已成为融合计算、存储和网络三大基础技术的综合应用平台。当前数据库系统正逐步向将数据密集型和计算密集型操作卸载至专用硬件平台的方向发展，以实现更高效的数据计算。

[0004] 传统的单异构计算节点针对数据库查询操作中的单一操作具有明显计算性能提升，但当一次查询任务包含多个基本操作时，且需要计算节点间流水执行时，由于数据在 CPU 内存与计算节点内存间频繁交互，导致计算加速性能明显下降，即计算节点间相互独立且发生数据交互时必须先回读到 CPU 内存再通过传统的 IP/UDP 网络传到对应的主机节点 CPU 内存，无法实现数据直接转发，导致 CPU 内存与计算节点内存间的频繁数据复制。

发明内容

[0005] 本公开为了解决上述问题，提出了基于多级流水线的数据库查询任务 RDMA 卸载方法及系统，通过将不同的计算节点内的计算单元、内存单元与 RDMA 网络单元组成分布式计算系统，实现数据在计算节点间直接转发，避免流水运算过程中主机节点 CPU 的内存与计算节点间的频繁数据复制，大大提高数据库查询操作的运算效率。

[0006] 根据一些实施例，本公开采用如下技术方案：

基于多级流水线的数据库查询任务 RDMA 卸载方法，包括：

构建以 RDMA 网络单元为中心，包括计算单元、内存单元和 RDMA 网络单元的分布式计算节点架构；

基于以 RDMA 网络单元为中心的分布式计算节点架构，通过 RDMA 网络单元以及不同计算单元的内存单元，无需 CPU 介入直接完成数据库查询任务数据的转发卸载；

其中，数据库查询任务的计算包括单任务运算模式和多任务融合运算模式，当下发的任务为多任务融合运算模式时，根据多任务中包含的基本操作选择相应的计算单元组成流水运算，中间产生的计算结果直接转发到该计算单元对应的内存单元后，经由 RDMA 网络直接缓存到

下一级计算单元的内存单元中，下一级计算单元读取内存单元中的中间结果继续运算，依次类推直至最终结果返回主机节点的 CPU 内存。

[0007] 根据一些实施例，本公开采用如下技术方案：

基于多级流水线的数据库查询任务 RDMA 卸载系统，包括：

计算节点架构构建模块，用于构建以 RDMA 网络单元为中心的分布式计算节点架构；

数据库查询任务卸载模块，用于基于以 RDMA 网络单元为中心的分布式计算节点架构，通过 RDMA 网络以及不同计算单元的内存单元，无需 CPU 介入直接完成数据库查询任务数据的转发卸载；

其中，数据库查询任务的计算包括单任务运算模式和多任务融合运算模式，当下发的任务为多任务融合运算模式时，根据多任务中包含的基本操作选择相应的计算单元组成流水运算，中间产生的计算结果直接转发到该计算单元对应的内存单元后，经由 RDMA 网络直接缓存到下一级计算单元的内存单元中，下一级计算单元读取内存单元中的中间结果继续运算，依次类推直至最终结果返回主机节点的 CPU 内存。

[0008] 根据一些实施例，本公开采用如下技术方案：

一种计算机程序产品，包括计算机程序，所述计算机程序被处理器执行时实现所述的基于多级流水线的数据库查询任务 RDMA 卸载方法。

[0009] 根据一些实施例，本公开采用如下技术方案：

一种非暂态计算机可读存储介质，所述非暂态计算机可读存储介质用于存储计算机指令，所述计算机指令被处理器执行时，实现所述的基于多级流水线的数据库查询任务 RDMA 卸载方法。

[0010] 根据一些实施例，本公开采用如下技术方案：

一种电子设备，包括：处理器、存储器以及计算机程序；其中，处理器与存储器连接，计算机程序被存储在存储器中，当电子设备运行时，所述处理器执行所述存储器存储的计算机程序，以使电子设备执行实现所述的基于多级流水线的数据库查询任务 RDMA 卸载方法。

[0011] 与现有技术相比，本公开的有益效果为：

本公开的基于多级流水线的数据库查询任务 RDMA 卸载方法，针对数据库查询操作包含多个基本操作时，不同的计算节点间通过 CPU 的内存频繁进行数据交换导致查询操作性能下降的问题，提出基于多级流水线的数据库查询任务 RDMA 卸载方法，通过将不同的计算节点内的计算单元、内存单元与 RDMA 网络单元组成分布式计算系统，实现数据在计算节点间直接转

发，避免流水运算过程中主机节点 CPU 的内存与计算节点间的频繁数据复制，大大提高数据库查询操作的运算效率。

[0012] 本公开的基于多级流水线的数据库查询任务 RDMA 卸载方法，基于 RDMA 网络技术将计算单元组成分布式计算系统架构，该网络技术允许数据直接从一个节点的内存复制到另一个节点的内存实现内存扩展与数据高速转发，而无需 CPU 介入。相比于传统计算节点间相互独立且发生数据交互时必须先回读到 CPU 内存再通过传统的 IP/UDP 网络传到对应的主机节点 CPU 内存，本公开方法允许源计算节点通过节点内的 RDMA 网络单元直接访问到目的计算节点的内存单元，实现数据直接转发，避免 CPU 内存与计算节点内存间的频繁数据复制。

[0013] 本公开的基于多级流水线的数据库查询任务 RDMA 卸载方法，以 CPU+FPGA 的异构平台为计算节点，通过 RDMA 网络将不同的计算节点内组成分布式计算系统，实现不同的计算单元之间直接通过内存单元经由 RDMA 网络单元完成数据快速高效的转发，避免了主机节点与计算节点频繁发生数据交换导致整体计算性能下降的问题，大大提高数据库查询操作的运算效率。同时，基于该 RDMA 技术的高带宽、低延时特性使得目的计算节点的内存与源计算节点内存组成内存池，实现内存扩展。

附图说明

[0014] 构成本公开的一部分的说明书附图用来提供对本公开的进一步理解，本公开的示意性实施例及其说明用于解释本公开，并不构成对本公开的不当限定。

[0015] 图 1 为本公开实施例的计算节点单元组成示意图；

图 2 为本公开实施例的以 RDMA 网络为中心的分布式计算节点架构；

图 3 为本公开实施例的多查询任务数据转发卸载示意图；

具体实施方式

下面结合附图与实施例对本公开作进一步说明。

[0016] 应该指出，以下详细说明都是例示性的，旨在对本公开提供进一步的说明。除非另有指明，本文使用的所有技术和科学术语具有与本公开所属技术领域的普通技术人员通常理解的含义。

[0017] 需要注意的是，这里所使用的术语仅是为了描述具体实施方式，而非意图限制根据本公开的示例性实施方式。如在这里所使用的，除非上下文另外明确指出，否则单数形式也意图包括复数形式，此外，还应当理解的是，当在本说明书中使用术语“包含”和/或“包括”时，其指明存在特征、步骤、操作、器件、组件和/或它们的组合。

[0018] 实施例 1

本公开的一种实施例中提供了一种基于多级流水线的数据库查询任务 RDMA 卸载方法,基于 RDMA 网络传输技术,以 CPU+FPGA 的异构平台为计算节点,通过 RDMA 网络将不同的计算节点内组成分布式计算系统,实现不同的计算节点单元之间直接通过内存单元以及 RDMA 网络单元完成数据快速高效的转发,包括:

步骤一:构建以 RDMA 网络单元为中心的分布式计算节点架构;

步骤二:基于以 RDMA 网络单元为中心的分布式计算节点架构,通过 RDMA 网络单元以及不同计算单元的内存单元,无需 CPU 介入直接完成数据库查询任务数据的转发卸载;

其中,数据库查询任务的计算包括单任务运算模式和多任务融合运算模式,当下发的任务为多任务融合运算模式时,根据多任务中包含的基本操作选择相应的计算单元组成流水运算,中间产生的计算结果直接转发到该计算单元对应的内存单元后,经由 RDMA 网络单元直接缓存到下一级计算单元的内存单元中,下一级计算单元读取内存单元中的中间结果继续运算,依次类推直至最终结果返回主机节点的 CPU 内存。

[0019] 作为一种实施例,本公开的基于多级流水线的数据库查询任务 RDMA 卸载方法,允许数据直接从一个节点的内存复制到另一个节点的内存实现内存扩展与数据高速转发,而无需 CPU 介入,允许源计算节点通过节点内的 RDMA 网络单元直接访问到目的计算节点的内存单元,实现数据直接转发,具体的实施过程如下:

步骤 1:构建以 RDMA 网络单元为中心的分布式计算节点架构;

具体地,以 CPU+FPGA 的异构平台为计算节点,通过 RDMA 网络单元将不同的计算节点内组成分布式计算系统,实现不同的计算单元之间直接通过内存单元经由 RDMA 网络单元完成数据快速高效的转发。

[0020] 首先,如图 1 所示,计算节点架构包括主机节点、计算单元、内存单元、RDMA 网络单元以及下一级 RDMA 网络单元,其中,主机节点 Host 产生数据库查询任务,通过调用内核 API 函数,将相关任务类型、目的计算节点等参数通过 PCIE 总线传输到计算节点内,同时 CPU 调用 XRT API 将待处理数据经由 PCIE DMA 缓存到内存单元 M,同时根据任务类型在不同的计算节点部署相应的计算单元,并生成启动信号。

[0021] 其中,所述计算节点包括计算单元、内存单元以及 RDMA 网络单元。所述计算单元与内存单元之间能够实现数据的发送以及接收,内存单元与 RDMA 网络单元之间能够实现通信数据的发送以及接收,计算节点的 RDMA 网络单元与下一级计算单元的下一级 RDMA 网

络单元之间能够实现数据的发送以及接收。

[0022] 其次，数据库查询任务的基本操作包括排序操作 P、聚合操作 J、过滤操作 G 以及连接操作 L，基本操作的计算单元包括排序计算单元、聚合计算单元、过滤计算单元及连接计算单元，该计算系统支持单任务运算模式和多任务融合运算模式，多任务融合运算模式的基本操作为相应的计算节点单元组成的流水运算。

[0023] 进一步地，单任务运算模式为单一的排序操作 P、聚合操作 J、过滤操作 G 或者连接操作 L，多任务融合运算模式包括：

- (1) 排序操作 P+聚合操作 J、
- (2) 排序操作 P+过滤操作 G、
- (3) 排序操作 P+连接操作 L、
- (4) 聚合操作 J+过滤操作 G
- (5) 聚合操作 J+连接操作 L
- (6) 过滤操作 G+连接操作 L
- (7) 排序操作 P+聚合操作 J+过滤操作 G、
- (8) 排序操作 P+聚合操作 J+连接操作 L、
- (9) 排序操作 P+过滤操作 G+连接操作 L、
- (10) 聚合操作 J+过滤操作 G+连接操作 L、
- (11) 排序操作 P+聚合操作 J+过滤操作 G+连接操作 L。

[0024] 在不考虑执行顺序或依赖关系情况下，共计 $C(4, 2)+C(4, 3)+C(4, 4)=6+4+1=11$ 种多任务融合运算基础组合。

[0025] 步骤 2：基于以 RDMA 网络为中心的分布式计算节点架构，通过 RDMA 网络以及不同计算节点单元的内存单元，无需 CPU 介入直接完成数据库查询任务数据的转发卸载；具体地，当数据库下发查询操作中的单任务的基本操作任务时，计算系统调度对应的计算单元完成运算并将结果返回到对应的内存单元最终返回主机节点的 CPU 内存。

[0026] 当下发的是多任务操作融合任务，计算系统需要根据任务中包含的基本操作选择相应的计算单元组成流水运算，中间产生的计算结果并不返回主机节点内存，而是转发到该计算单元对应的内存单元，并经由 RDMA 网络单元直接缓存到下一级计算单元的内存单元中，下一级计算单元读取内存中的中间结果并继续运算，依次类推直至最终结果返回主机内存。通过 RDMA 网络组成的分布式计算系统可有效减少数据在 CPU 内存与计算节点内存之间的复

制次数，避免了主机节点与计算节点频繁发生数据交换导致整体计算性能下降的问题，大大提高数据库查询操作的运算效率。

[0027] 作为一种实施例，数据库下发查询操作的基本操作任务计算由四个计算节点完成，具体过程如下：

S1: 主机节点 Host 产生数据库查询任务，通过调用内核 API 函数，将相关任务类型、目的计算节点等参数通过 PCIE 总线传输到计算节点内，同时 CPU 调用 XRT API 将待处理数据经由 PCIE DMA 缓存到内存单元 M，同时根据任务类型在不同的计算节点部署相应的计算单元，并生成启动信号。

[0028] S2: 当数据库下发查询操作中的单任务的基本操作任务时，计算系统调度对应的计算单元完成运算并将结果返回到对应的内存单元最终返回主机节点的 CPU 内存。

[0029] 其中，若任务只包含单一的连接操作 L 时，计算节点部署连接计算单元，从内存单元 M4 中读取待运算数据。内存单元 M4 分为多块存储区域分别为 M4-1、M4-2、M4-3、M4-4。其中 M4-1 和 M4-2 分别存储内表数据及外表数据。根据连接算法，首先连接计算单元读取内表数据经过哈希地址映射后将内表数据按照哈希地址存储到 M4-3 中，之后连接计算单元再读取 M4-2 中的外表数据元素，假设外表数据元素为 T_o ，该元素经由哈希地址映射后生成哈希地址 H，根据地址 H 从存储区域 M4-3 中读取该地址下的内表元素 T_i ，之后比较 T_o 与 T_i 的值若相等则匹配成功，将合并后的元素 $\{T_o, T_i\}$ 最为连接结果存储到 M4-4 中，依次类推直到遍历外表元素后，得到最终的连接结果均存储在 M4-4 中同时产生结束信号，CPU 检测到计算单元结束信号后从 M4-4 中读取连接结果。

[0030] S3: 假设任务为多任务操作融合任务，计算系统需要根据任务中包含的基本操作选择相应的计算单元组成流水运算，中间产生的计算结果并不返回主机节点内存，而是转发到该计算单元对应的内存单元，并经由 RDMA 网络单元直接缓存到下一级计算单元的内存单元中，下一级计算单元读取内存中的中间结果并继续运算，依次类推直至最终结果返回主机内存。

[0031] 具体地，若包含过滤操作 G+排序操作 P 两种操作时，以部署过滤计算单元的计算节点为主节点，部署排序计算单元的计算节点为从节点开始运算。过滤计算单元作为主节点，首先主机节点将待运算数据缓存到内存单元 M2，之后过滤计算单元基于真值表的过滤架构，将接收的列数据分别与对应列过滤条件进行比较，获得的比较结果即为该行数据的特征，这些结果表征该行各个过滤列是否满足过滤条件。然后，将比较结果作为真值表 RAM 的读取

地址，RAM 的读取值表征当前行数据是否满足过滤条件，若满足过滤条件，输出该行数据缓存到 M2 中，依次类推完成待运算数据的过滤操作，根据任务类型将 M2 中的过滤结果经由 RDMA 网络单元传输到从节点的内存单元 M1 中。之后从节点启动排序计算单元，其内的八个排序子单元采用八通道归并排序算法将待排序数据分成八份，其中八通道归并排序算法采用多级比较逻辑，每级比较逻辑由若干对比较器组成，每对比较器由两个 RAM 和一个比较器组成，其中，RAM 缓存输入数据以及阻塞控制，比较器比较两个 RAM 的输出数据，比较器以前级 RAM 的空标志信号和后级 RAM 的满标志信号作为有效信号，防止输入或输出数据的溢出。经过排序子单元产生的中间排序结果再缓存到 M1 的不同区域中，归并排序子单元读取中间结果使用归并排序进行归并，完成最终排序并将结果缓存到 M1 中。

至此多任务融合型的计算操作已经完成，RDMA 网络单元将最终的运算结果通过链路 E1 返回内存单元 M2 中，主节点通过 PCIE 总线将运算结果数据回读到 CPU 内存中。之后各节点释放计算单元所占用的资源，并拆除 RDMA 网络链路 E1，等待下次运算任务的卸载。

[0032] 如附图 2 所示，针对较为复杂的多任务融合运算，若包含排序操作 P+聚合操作 J+过滤操作 G+连接操作 L，假设存在“排序-过滤-连接”和“排序-聚合-连接”两条流水线，则以排序计算单元的计算节点为主节点，部署排序计算单元和聚合计算单元的计算节点为从节点开始运算，分别按照 M1-E1-M2-E2-M4 和 M1-E4-M3-E3-M4 和两条路径进行流水线任务卸载，最终，RDMA 网络单元将最终的运算结果通过链路 E5 返回内存单元 M1 中，主节点通过 PCIE 总线将运算结果数据回读到 CPU 内存中。

[0033] 如附图 3 所示，单任务运算的查询操作只涉及单一节点，如从 CPU 内存到内存单元到计算单元（如附图 3 中的查询操作 1）；多任务融合运算的查询操作设计两个或以上节点，包含多阶段的查询操作（如附图 3 中的查询操作 2 和查询操作 3）。

[0034] 实施例 2

本公开的一种实施例中提供了一种基于多级流水线的数据库查询任务 RDMA 卸载系统，包括：

计算节点架构构建模块，用于构建以 RDMA 网络为中心的分布式计算节点架构；

数据库查询任务卸载模块，用于基于以 RDMA 网络为中心的分布式计算节点架构，通过 RDMA 网络以及不同计算节点单元的内存单元，无需 CPU 介入直接完成数据库查询任务数据的转发卸载；

其中，数据库查询任务的计算包括单任务运算模式和多任务融合运算模式，

当下发的任务为多任务融合运算模式时，根据多任务中包含的基本操作选择相应的计算节点单元组成流水运算，中间产生的计算结果直接转发到该计算节点单元对应的内存单元后，经由 RDMA 网络直接缓存到下一级计算节点单元的内存单元中，下一级计算节点单元读取内存单元中的中间结果继续运算，依次类推直至最终结果返回主机节点的 CPU 内存。

[0035] 实施例 3

本公开的一种实施例中提供了一种计算机程序产品，包括计算机程序，所述计算机程序被处理器执行时实现所述的基于多级流水线的数据库查询任务 RDMA 卸载方法。

[0036] 实施例 4

本公开的一种实施例中提供了一种非暂态计算机可读存储介质，所述非暂态计算机可读存储介质用于存储计算机指令，所述计算机指令被处理器执行时，实现所述的基于多级流水线的数据库查询任务 RDMA 卸载方法。

[0037] 实施例 5

本公开的一种实施例中提供了一种电子设备，包括：处理器、存储器以及计算机程序；其中，处理器与存储器连接，计算机程序被存储在存储器中，当电子设备运行时，所述处理器执行所述存储器存储的计算机程序，以使电子设备执行实现所述的基于多级流水线的数据库查询任务 RDMA 卸载方法。

[0038] 本公开是参照根据本公开实施例的方法、设备（系统）、和计算机程序产品的流程图和 / 或方框图来描述的。应理解可由计算机程序指令实现流程图和 / 或方框图中的每一流程和 / 或方框、以及流程图和 / 或方框图中的流程和 / 或方框的结合。可提供这些计算机程序指令到通用计算机、专用计算机、嵌入式处理机或其他可编程数据处理设备的处理器以产生一个机器，使得通过计算机或其他可编程数据处理设备的处理器执行的指令产生用于实现在流程图一个流程或多个流程和 / 或方框图一个方框或多个方框中指定的功能的装置。

[0039] 这些计算机程序指令也可装载到计算机或其他可编程数据处理设备上，使得在计算机或其他可编程设备上执行一系列操作步骤以产生计算机实现的处理，从而在计算机或其他可编程设备上执行的指令提供用于实现在流程图一个流程或多个流程和 / 或方框图一个方框或多个方框中指定的功能的步骤。

[0040] 上述虽然结合附图对本公开的具体实施方式进行了描述，但并非对本公开保护范围的限制，所属领域技术人员应该明白，在本公开的技术方案的基础上，本领域技术人员不需要付出创造性劳动即可做出的各种修改或变形仍在本公开的保护范围以内。

说明书附图

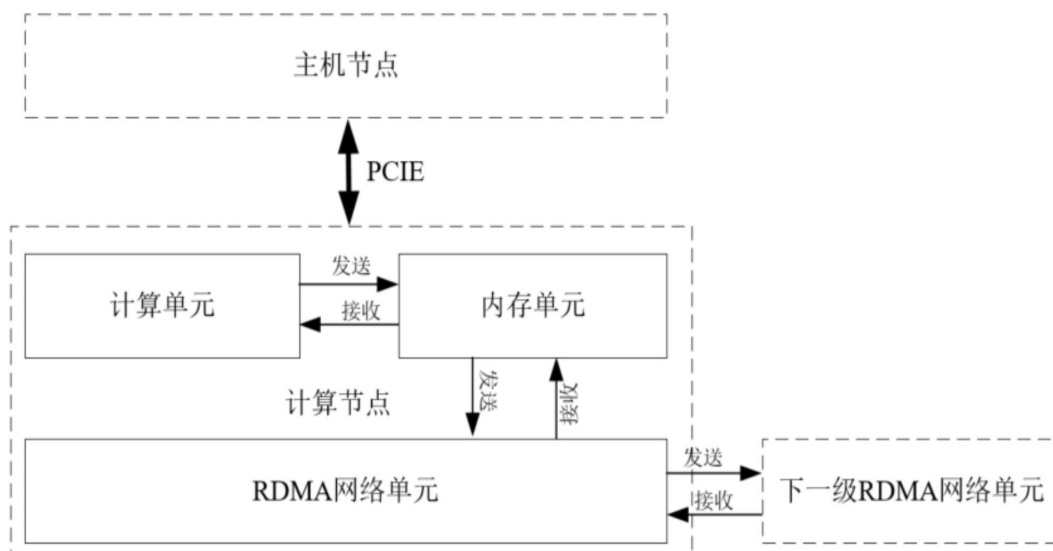


图 1

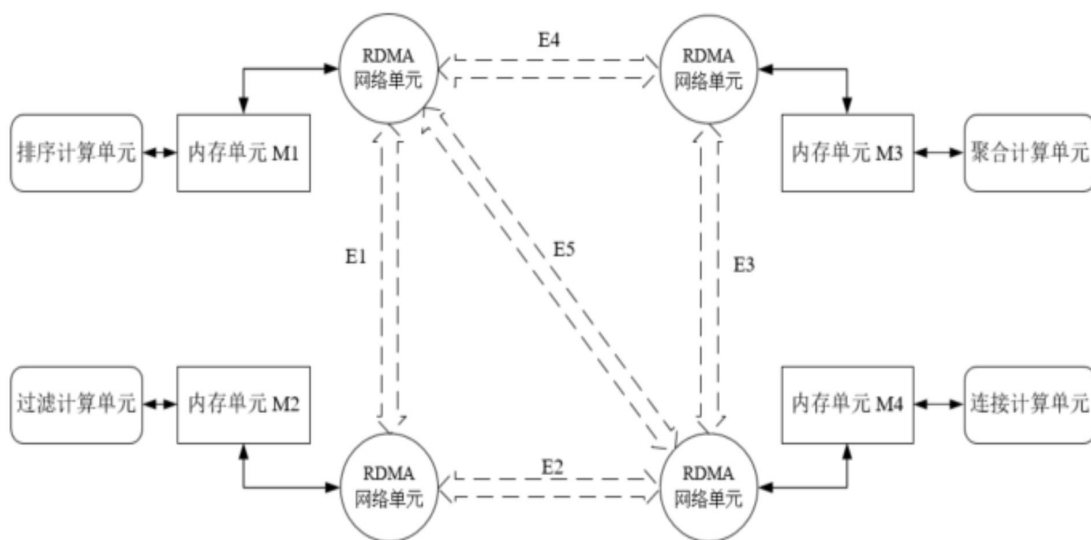


图 2

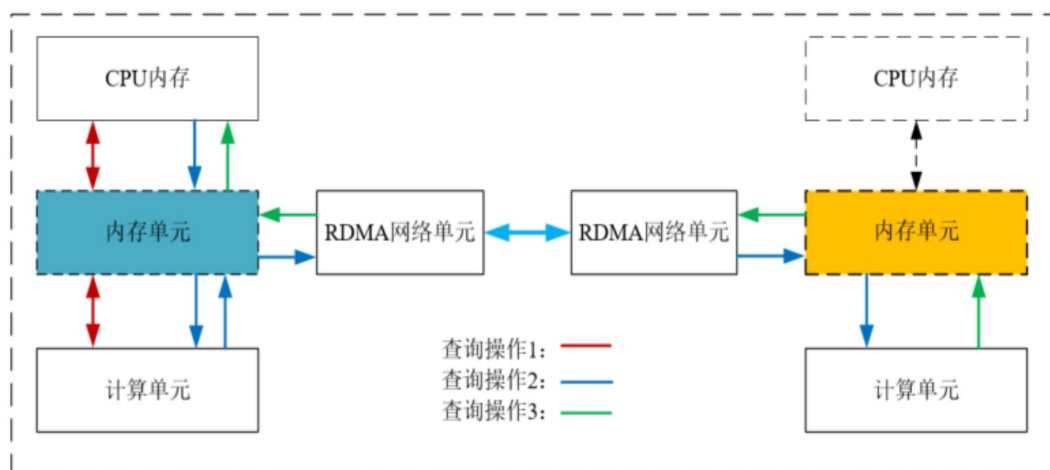


图 3

说明书摘要

本公开提供了基于多级流水线的数据库查询任务 RDMA 卸载方法及系统, 涉及 RDMA 网络和数据库查询加速技术领域, 包括: 基于以 RDMA 网络为中心的分布式计算节点架构, 通过 RDMA 网络以及不同计算节点单元的内存单元直接完成数据库查询任务数据的转发卸载; 其中, 数据库查询任务的计算包括单任务运算模式和多任务融合运算模式, 当下发的任务为多任务融合运算模式时, 根据多任务中包含的基本操作选择相应的计算节点单元组成流水运算, 中间产生的计算结果直接转发到该计算节点单元对应的内存单元后, 经由 RDMA 网络直接缓存到下一级计算节点单元的内存单元中, 下一级计算节点单元读取内存单元中的中间结果继续运算, 依次类推直至最终结果返回主机节点的 CPU 内存。