

(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  
2025年6月5日 (05.06.2025)



(10) 国际公布号  
**WO 2025/112410 A1**

- (51) 国际专利分类号:  
**H04L 67/568** (2022.01)
- (21) 国际申请号: PCT/CN2024/097159
- (22) 国际申请日: 2024年6月4日 (04.06.2024)
- (25) 申请语言: 中文
- (26) 公布语言: 中文
- (30) 优先权:  
202311598492.4 2023年11月27日 (27.11.2023) CN
- (71) 申请人: 海光信息技术股份有限公司 (**HYGON INFORMATION TECHNOLOGY CO., LTD.**) [CN/CN]; 中国天津市滨海新区华苑产业区海泰西路18号北2-204工业孵化-3-8 300392 (CN)。
- (72) 发明人: 程永波 (**CHENG, Yongbo**); 中国天津市滨海新区华苑产业区海泰西路18号北2-204工业孵化-3-8 300392 (CN)。 娄丽芳 (**LOU, Lifang**); 中国天津市滨海新区华苑产业区海泰西路18号北2-204工业孵化-3-8 300392 (CN)。 曹俊 (**CAO, Jun**); 中国天津市滨海新区华苑产业区海泰西路18号北2-204工业孵化-3-8 300392 (CN)。 耿剑波 (**GENG,**

- Jianbo**); 中国天津市滨海新区华苑产业区海泰西路18号北2-204工业孵化-3-8 300392 (CN)。
- (74) 代理人: 北京市柳沈律师事务所 (**LIU, SHEN & ASSOCIATES**); 中国北京市海淀区彩和坊路10号1号楼10层 100080 (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, CV, CZ, DE, DJ, DK, DM, DO, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, GT, HN, HR, HU, ID, IL, IN, IQ, IR, IS, IT, JM, JO, JP, KE, KG, KH, KN, KP, KR, KW, KZ, LA, LC, LK, LR, LS, LU, LY, MA, MD, MG, MK, MN, MU, 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, ST, SV, SY, TH, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, WS, 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, CV, GH, GM, KE, LR, LS, MW, MZ, NA, RW, SC, SD, SL, ST, 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,

(54) **Title:** DATA PROCESSING METHOD AND APPARATUS, AND ELECTRONIC DEVICE AND STORAGE MEDIUM

(54) 发明名称: 数据处理方法和装置、电子设备以及存储介质

- S101 响应于收到第一缓存要将第一数据写回内存的写请求, 缓存一致性节点将所述第一数据写入内存或将所述第一数据丢弃, 其中, 所述写请求携带所述第一数据, 所述缓存一致性节点配置为维护由多个缓存中存储的数据的一致性, 所述多个缓存包括所述第一缓存
- S102 在所述缓存一致性节点将所述第一数据写入内存或将所述第一数据丢弃之后, 所述缓存一致性节点向所述第一缓存返回写响应信号以直接结束对所述写请求的处理操作。

- S101 In response to a write request from a first cache for writing back first data to a memory having been received, a cache coherence node writes the first data into the memory or discards the first data, wherein the write request carries the first data, and the cache coherence node is configured to maintain the consistency of data stored in a plurality of caches, which include the first cache
- S102 After writing the first data into the memory or discarding the first data, the cache coherence node returns a write response signal to the first cache to directly end a processing operation for the write request

图 3A

(57) **Abstract:** Provided in the embodiments of the present disclosure are a data processing method and apparatus, and an electronic device and a storage medium. The data processing method comprises: in response to a write request from a first cache for writing back first data to a memory having been received, a cache coherence node writing the first data into the memory or discarding the first data, wherein the write request carries the first data, and the cache coherence node is configured to maintain the consistency of data stored in a plurality of caches, which include the first cache; and after writing the first data into the memory or discarding the first data, the cache coherence node returning a write response signal to the first cache so as to directly end a processing operation for the write request. The method can reduce the number of ACK messages, lower requirements of protocols in respect of ACK messages, reduce the power consumption, latency and bandwidth occupation generated during ACK message transmission, and optimize the turnover rate of buffers on a master device side. Under the same performance conditions, fewer buffers are required on the master device side, thereby saving on space and reducing the power consumption.

WO 2025/112410 A1

HU, IE, IS, IT, LT, LU, LV, MC, ME, 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)。

本国际公布:

— 包括国际检索报告(条约第21条(3))。

---

**(57)** 摘要: 本公开的实施例提供了数据处理方法和装置、电子设备和存储介质。该数据处理方法包括: 响应于收到第一缓存要将第一数据写回内存的写请求, 缓存一致性节点将第一数据写入内存或将第一数据丢弃, 其中, 写请求携带第一数据, 缓存一致性节点配置为维护由多个缓存中存储的数据的一致性, 多个缓存包括第一缓存; 以及在缓存一致性节点将第一数据写入内存或将第一数据丢弃之后, 缓存一致性节点向第一缓存返回写响应信号以直接结束对写请求的处理操作。该方法能够降低ACK报文的数量, 减少协议对ACK报文的需求, 减少ACK报文传输产生的功耗和延时以及带宽占用, 并且优化了主设备侧缓冲区的周转率, 在相同性能下主设备侧需要的缓冲区数量更少, 节省面积和功耗。

## 数据处理方法和装置、电子设备以及存储介质

本申请要求于 2023 年 11 月 27 日递交的中国专利申请第 202311598492.4 号的优先权，在此全文引用上述中国专利申请公开的内容以作为本申请的一部分。

### 技术领域

本公开的实施例涉及一种数据处理方法、数据处理装置、电子设备和存储介质。

### 背景技术

片上网络 (Network-on-a-Chip, NOC) 系统是一种必要的总线系统，用于实现大规模集成电路中各组件之间的相互通信。在一个系统级芯片 (SOC) 中包含了多种类型的子系统，如 CPU、GPU、内存、IO 设备等，并且每种类型的子系统都可能存在多个实体，这些实体需要彼此通信 (数据信号和控制信号) 协同工作。

在一个 SOC 系统中，多个子系统实体之间需要进行频繁的读写交互。为了确保各子系统之间能够高性能地交互数据，NOC 系统提供了一个大带宽的事务交换网络，用于连接片上各组件。通过这个网络，各个子系统之间可以实现高效的数据传输。当不同的子系统对应的缓存对所缓存的相同地址的数据进行操作时，会产生一致性问题。为了解决这个问题，出现了很多技术来维护数据一致性。

### 发明内容

本公开的至少一个实施例提供了一种数据处理方法，所述数据处理方法包括：响应于收到第一缓存要将第一数据写回内存的写请求，缓存一致性节点将所述第一数据写入内存或将所述第一数据丢弃，其中，所述写请求携带所述第一数据，所述缓存一致性节点配置为维护由多个缓存中存储的数据的一致性，所述多个缓存包括所述第一缓存，以及在所述缓存一致性节点将所述第一数据写入内存或将所述第一数据丢弃之后，所述缓存一致性节点向所

述第一缓存返回写响应信号以直接结束对所述写请求的处理操作。

例如，在本公开的至少一个实施例提供的数据处理方法中，所述响应于收到第一缓存要将第一数据写回内存的写请求，缓存一致性节点将所述第一数据写入内存或将所述第一数据丢弃，包括：响应于收到所述第一缓存要将所述第一数据写回所述内存的所述写请求，所述缓存一致性节点确定所述第一缓存是否与所述缓存一致性节点侧记录的当前独占所述第一数据的缓存相同；以及响应于确定所述第一缓存与所述缓存一致性节点侧记录的当前独占所述第一数据的缓存相同，所述缓存一致性节点将所述第一数据写入所述内存。

10 例如，在本公开的至少一个实施例提供的数据处理方法中，所述在所述缓存一致性节点将所述第一数据写入内存或将所述第一数据丢弃之后，所述缓存一致性节点向所述第一缓存返回写响应信号以直接结束对所述写请求的处理操作，包括：在所述缓存一致性节点将所述第一数据写入内存之后，所述内存向所述缓存一致性节点返回指示所述第一数据已经被写入所述内存的写响应；以及响应于从所述内存收到指示所述第一数据已经被写入所述内存的所述写响应，所述缓存一致性节点向所述第一缓存返回指示所述第一数据已经被写回所述内存的所述写响应信号以直接结束对所述写请求的处理操作。

20 例如，在本公开的至少一个实施例提供的数据处理方法中，所述数据处理方法还包括：在所述缓存一致性节点将所述第一数据写入所述内存之后，所述缓存一致性节点将所述第一缓存的数据状态置为无效态、共享态或独占态，其中，所述无效态表示所述第一缓存没有缓存所述第一数据，所述共享态表示所述第一缓存与其他缓存均缓存有所述第一数据，所述独占态表示仅所述第一缓存中缓存有所述第一数据。

25 例如，在本公开的至少一个实施例提供的数据处理方法中，所述数据处理方法还包括：响应于收到第二缓存对所述第一数据的读请求且所述缓存一致性节点侧记录所述第一数据在所述第一缓存中处于独占状态或共享状态，所述缓存一致性节点向所述第一缓存发起监测信号；响应于所述第一缓存在发出所述写请求之后收到所述监测信号，所述第一缓存响应于所述监测信号向所述缓存一致性节点返回携带所述第一数据的监测响应。

30 例如，在本公开的至少一个实施例提供的数据处理方法中，所述数据处

理方法还包括：响应于收到所述监测响应，所述缓存一致性节点将所述第一数据返回给所述第二缓存；以及在收到所述第一数据之后，所述第二缓存向所述缓存一致性节点返回指示已经收到所述第一数据的报文以结束所述读请求。

5 例如，在本公开的至少一个实施例提供的数据处理方法中，所述响应于收到所述监测响应，所述缓存一致性节点将所述第一数据返回给所述第二缓存，包括：响应于收到所述监测响应，所述缓存一致性节点将所述第一数据返回给所述第二缓存，并且将所述第二缓存的数据状态置为独占态，所述独占态表示仅所述第二缓存中缓存有所述第一数据。

10 例如，在本公开的至少一个实施例提供的数据处理方法中，所述在所述缓存一致性节点将所述第一数据写入内存或将所述第一数据丢弃之后，所述缓存一致性节点向所述第一缓存返回写响应信号以直接结束对所述写请求的处理操作，包括：在收到所述报文之后，所述缓存一致性节点丢弃所述写请求写入所述缓存一致性节点侧的所述第一数据，并且向所述第一缓存返回  
15 所述写响应信号以直接结束对所述写请求的处理操作。

例如，在本公开的至少一个实施例提供的数据处理方法中，所述在收到所述报文之后，所述缓存一致性节点丢弃所述写请求写入所述缓存一致性节点侧的所述第一数据，包括：响应于在收到所述报文之后确定所述写请求写入所述缓存一致性节点侧的所述第一数据为无效数据，所述缓存一致性节点  
20 丢弃所述写请求写入所述缓存一致性节点侧的所述第一数据。

本公开的至少一个实施例还提供了一种电子设备，所述电子设备包括多个缓存和缓存一致性节点，所述多个缓存包括第一缓存，并且所述缓存一致性节点配置为响应于收到所述第一缓存要将第一数据写回内存的写请求，将所述第一数据写入内存或将所述第一数据丢弃，以及在所述缓存一致性节点  
25 将所述第一数据写入内存或将所述第一数据丢弃之后，向所述第一缓存返回写响应信号以直接结束对所述写请求的处理操作，其中，所述写请求携带所述第一数据，所述缓存一致性节点配置为维护由所述多个缓存中存储的数据的一致性。

例如，在本公开的至少一个实施例提供的电子设备中，所述缓存一致性节点还配置为：响应于收到所述第一缓存要将所述第一数据写回所述内存的  
30 所述写请求，确定所述第一缓存是否与所述缓存一致性节点侧记录的当前独

占所述第一数据的缓存相同，以及响应于确定所述第一缓存与所述缓存一致性节点侧记录的当前独占所述第一数据的缓存相同，将所述第一数据写入所述内存。

例如，在本公开的至少一个实施例提供的电子设备中，所述内存配置为：

5 在所述缓存一致性节点将所述第一数据写入内存之后，向所述缓存一致性节点返回指示所述第一数据已经被写入所述内存的写响应，并且所述缓存一致性节点还配置为：响应于从所述内存收到指示所述第一数据已经被写入所述内存的所述写响应，向所述第一缓存返回指示所述第一数据已经被写回所述内存的所述写响应信号以直接结束对所述写请求的处理操作。

10 例如，在本公开的至少一个实施例提供的电子设备中，所述缓存一致性节点还配置为：响应于收到第二缓存对所述第一数据的读请求且所述缓存一致性节点侧记录所述第一数据在所述第一缓存中处于独占状态或共享状态，向所述第一缓存发起监测信号，并且所述第一缓存配置为：响应于所述第一缓存在发出所述写请求之后收到所述监测信号，向所述缓存一致性节点返回  
15 携带所述第一数据的监测响应。

本公开的至少一个实施例还提供了一种数据处理装置，所述数据处理装置包括存储器和处理器，所述存储器配置为存储计算机可执行指令，并且所述处理器配置为执行所述计算机可执行指令，其中，所述计算机可执行指令被所述处理器执行时实现如上任一实施例所述的方法。

20 本公开的至少一个实施例还提供了一种非暂时性存储介质，非暂时性地存储计算机可执行指令，其中，当所述计算机可执行指令由处理器执行时，实现如上任一实施例所述的方法。

#### 附图说明

25 为了更清楚地说明本公开实施例的技术方案，下面将对实施例的附图作简单地介绍，显而易见地，下面描述中的附图仅仅涉及本公开的一些实施例，而非对本公开的限制。

图 1 示出了一种电子设备的示意图。

图 2A 示出了一种电子设备进行数据读取的示意图。

30 图 2B 示出了一种电子设备进行数据写回的示意图。

图 3A 示出了本公开至少一实施例提供的一种数据处理方法的流程示意

图。

图 3B 示出了本公开至少一实施例提供的一种电子设备进行数据写回的示意图。

5 图 4 示出了本公开至少一实施例提供的一种电子设备存在读写冲突时的读写操作示意图。

图 5 示出了本公开至少一实施例提供的一种电子设备的示意图。

图 6 示出了本公开至少一实施例提供的一种数据处理装置的示意图。

图 7 示出了本公开至少一实施例提供的一种非暂时性存储介质的示意图。

10

### 具体实施方式

为使本公开实施例的目的、技术方案和优点更加清楚，下面将结合本公开实施例的附图，对本公开实施例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公开的一部分实施例，而不是全部的实施例。基于所描述的本公开的实施例，本领域普通技术人员在无需创造性劳动的前提下所获得的所有其他实施例，都属于本公开保护的范围。

除非另外定义，本公开使用的技术术语或者科学术语应当为本公开所属领域内具有一般技能的人士所理解的通常意义。本公开中使用的“第一”、“第二”以及类似的词语并不表示任何顺序、数量或者重要性，而只是用来区分不同的组成部分。“包括”或者“包含”等类似的词语意指出现该词前面的元件或者物件涵盖出现在该词后面列举的元件或者物件及其等同，而不排除其他元件或者物件。“连接”或者“相连”等类似的词语并非限定于物理的或者机械的连接，而是可以包括电性的连接，不管是直接的还是间接的。

25 “上”、“下”、“左”、“右”等仅用于表示相对位置关系，当被描述对象的绝对位置改变后，则该相对位置关系也可能相应地改变。

为了保持对本公开实施例的以下说明清楚且简明，本公开省略了部分已知功能和已知部件的详细说明。

在包括多个处理器（或处理器核）和多个高速缓存的计算机系统的运行过程中，相同地址的数据可能会同时存在不同的缓存中。为使得各个缓存中相同地址的数据始终保持一致，产生了基于目录的一致性协议。该协议可以追踪高速缓存中的数据的状态，例如，目录存储结构可以与高速缓存一样，

采用静态随机存储器 (SRAM) 实现。例如, 目录存储结构也可以采用其他形式的存储器实现, 只要能够实现目录存储即可。

图 1 示出了一种电子设备 (或电子系统) 100 的示意图。如图 1 所示, 电子设备 100 可以包括多个高速缓存 (以下也简称“缓存”) 1~n、处理器 (或处理器核) 1~n、互连网络、缓存一致性节点 1~n、内存 1~n 和目录 1~n 等。

例如, 缓存 1~n 中的每个缓存被配置为能够存储至少一个数据信息对应的数据存储信息。例如, 缓存 1~n 可以分别包括一个或多个子缓存。

例如, 缓存一致性节点 1~n 可以分别帮助追踪高速缓存 1~n 中的缓存数据状态, 缓存数据状态例如包括: 缓存 1~n 中只有该缓存数据的单一副本、有该缓存数据的多个副本、或者数据只存在主存中等状态。缓存一致性节点 1~n 会追踪处理器 1~n (例如, 图 1 中的每个处理器可以包括一个或多个子处理器) 的高速缓存 1~n 的状态, 并把追踪的状态信息 (或目录信息) 存储在目录 1~n 中。当缓存一致性节点侦听总线过程中发现一致性事务后, 会查询目录 1~n 中追踪的状态信息并发出相应的探针完成一致性维护。

例如, 目录 1~n 被配置为存储缓存 1~n 中存储的数据信息对应的目录信息以维护缓存一致性。例如, 目录包括多个存储组, 每个存储组可以存储多条目录信息。例如, 当需要向目录中存入目录信息时, 目录根据映射关系将该目录信息放入相应的存储组。

例如, 如图 1 所示, 系统中的互连网络 (例如片上网络 (NOC) 等) 与缓存 1~n 直接耦接, 并且还与缓存一致性节点 1~n 直接耦接。互连网络为信息传输的公共通信干线, 例如, 互连网络可以为由导线等电子元件组成的传输线束。

例如, 如图 1 所示, 缓存一致性节点 1~n 被配置为将目录 1~n 分别连接到互连网络, 缓存一致性节点 1~n 被配置为将内存 1~n 分别连接到互连网络, 从而维护缓存 1~n 中存储的所有数据的一致性。

例如, 在非一致性内存访问架构 (Non Uniform Memory Access, NUMA) 中, 不同的处理器 (或处理器核) 与内存器件的组合从属于不同的节点, 因此基于目录一致性的设计会在每个内存控制器的附近设置一个静态随机存储器 (SRAM) 结构的目录。在 NUMA 架构中, 每个节点都有自己的本地内存。

图 2A 示出了一种电子设备进行数据读取的示意图。图 2B 示出了一种

电子设备进行数据回写的示意图。

例如，在图 2A 中，主设备 0 可以对应于图 1 中的缓存 1~n 中的任何一个，从设备可以对应于图 1 中的缓存一致性节点 1~n 中的任何一个，MEM 可以对应于图 1 中的内存 1~n 中的任何一个。例如，如图 2A 所示，当主设备 0 向从设备发起独占请求 RdE（参见下文的表 2）以获取独占数据，由于系统（例如电子系统）中可能存在多个主设备，所以从设备需要管理多个主设备之间的数据一致性。例如，系统中的从设备侧可以采用基于目录（Directory）的缓存一致性协议，由目录记录数据被哪个主设备缓存，以及该数据当前在对应的主设备中所处的状态。

10 例如，如图 2A 所示，当从设备收到独占请求 RdE，且当前从设备侧对应的目录为例如图 2A 中用 Dir@I 表示的无效态 I（参见下文的表 1，例如，无效态 I 表示当前没有主设备缓存此数据（例如，请求所访问的地址对应的数据））时从设备通过向内存发送读请求 Rdmem 且接收内存向从设备返回的数据响应 rddat，从而直接读取主存（例如，内存）数据，并向作为请求者的主设备 0 返回独占数据响应 DatRspE（参见下文的表 5），并将从设备侧对应的目录置为 E:Master0（例如图 2A 中用 Dir@E:Master0 表示），从而记录主设备 0 缓存此数据且为独占态（E: Exclusive），主设备 0 收到独占数据响应 DatRspE 后向从设备发送报文 ACK 以结束读请求。

主设备处一般缓存较多的数据，且主设备存在溢出的场景，需要从主设备将溢出的数据写回到从设备。例如，如图 2B 所示，当主设备 0 的缓存数据发生溢出时，主设备 0 向从设备发起写回请求（图 2B 中仅以 WBI 为例进行说明，但不限于此写回请求），并将要写回到内存的数据例如通过信号 WrDat 一并发送给从设备，从设备侧确认收到写回数据后向主设备 0 返回写响应 WrRap，主设备 0 收到写响应 WrRap 后向从设备发送报文（或写响应）ACK。例如，如图 2B 所示，从设备收到主设备 0 发送的报文 ACK 后，从设备侧对应的目录状态被从 Dir@E:Master0 置为 Dir@I。例如，从设备侧收到报文 ACK 后通过向内存发送写请求 Rdmem 以将待写回的数据写回内存，且接收内存向从设备返回的数据写响应 WrRap，以结束写回请求。

在现有的缓存一致性协议中，读写事务都需要携带 ACK 报文，ACK 报  
30 文会延长事务在从设备中的生命周期，即占用从设备缓冲区（例如，缓存或其他能够用于实现目录的存储区）的时间变长，降低从设备缓冲区的周转率

和浪费资源，并且 ACK 报文传输需要浪费一定的功耗。

本公开至少一个实施例提供一种数据处理方法。例如，本公开提供一种数据处理方法包括：响应于收到第一缓存要将第一数据写回内存的写请求，缓存一致性节点将第一数据写入内存或将第一数据丢弃，其中，写请求携带第一数据，缓存一致性节点配置为维护由多个缓存中存储的数据的一致性，多个缓存包括第一缓存；以及在缓存一致性节点将第一数据写入内存或将第一数据丢弃之后，缓存一致性节点向第一缓存返回写响应信号以直接结束对写请求的处理操作。

在本公开的上述实施例的数据处理方法中，通过回写事务和目录的比较，进行场景识别，主设备收到写响应后直接结束写事务处理，不需要再向从设备发送 ACK 报文，从而降低 ACK 报文的数量，减少协议对 ACK 报文的需求，减少 ACK 报文传输产生的功耗和延时以及带宽占用，并且优化了主设备侧缓冲区的周转率，在相同性能下主设备侧需要的缓冲区数量更少，节省面积和功耗。

本公开至少一些实施例还提供了一种电子设备，该电子设备包括多个缓存和缓存一致性节点。多个缓存包括第一缓存，缓存一致性节点被配置为响应于收到第一缓存要将第一数据写回内存的写请求，将第一数据写入内存或将第一数据丢弃，以及在缓存一致性节点将第一数据写入内存或将第一数据丢弃之后，向第一缓存返回写响应信号以直接结束对写请求的处理操作，其中，写请求携带第一数据，缓存一致性节点配置为维护由多个缓存中存储的数据的一致性。

本公开上述实施例的电子设备的 technical 效果与上述数据处理方法的 technical 效果相同，因此不再赘述。

上述实施例中，第一缓存可以是多个缓存中的任一缓存，第一数据可以是例如任何被访问的地址对应的数据，即，这里的“第一”（以及“第二”等）仅用于标识作为描述对象的缓存或数据，而非特指某一特定的缓存或数据。

下面将结合具体示例对本公开的各个实施例进行说明。

图 3A 示出了本公开至少一实施例提供的一种数据处理方法的流程示意图。如图 3A 所示，在本公开的一些实施例中，数据处理方法包括以下步骤 S101-S102。

步骤 S101, 响应于收到第一缓存要将第一数据写回内存的写请求, 缓存一致性节点将第一数据写入内存或将第一数据丢弃, 其中, 写请求携带第一数据, 缓存一致性节点配置为维护由多个缓存中存储的数据的一致性, 多个缓存包括第一缓存。

5 步骤 S102, 在缓存一致性节点将第一数据写入内存或将第一数据丢弃之后, 缓存一致性节点向第一缓存返回写响应信号以直接结束对写请求的处理操作。

图 3B 示出了本公开至少一实施例提供的一种电子设备进行数据回写的示意图。例如, 如图 3B 所示, 响应于收到第一缓存 (或 “主设备 0”) 要将第一数据写回内存的写请求, 缓存一致性节点 (或 “从设备”) 将第一数据写入内存 MEM。例如, 写请求携带第一数据, 缓存一致性节点配置为维护由多个缓存中存储的数据的一致性, 该多个缓存包括第一缓存。例如, 在缓存一致性节点将第一数据写入内存 MEM, 缓存一致性节点向第一缓存返回写响应信号 WrRsp 以直接结束对写请求的处理操作。

15 图 4 示出了本公开至少一实施例提供的一种电子设备存在读写冲突时的读写操作示意图。例如, 如图 4 所示, 响应于收到第一缓存 (或 “主设备 0”) 要将第一数据写回内存的写请求, 缓存一致性节点 (或 “从设备”) 将第一数据丢弃。例如, 写请求携带第一数据, 缓存一致性节点配置为维护由多个缓存中存储的数据的一致性, 该多个缓存包括第一缓存。例如, 在缓存一致性节点将第一数据丢弃之后, 缓存一致性节点向第一缓存返回写响应信号 WrRsp 以直接结束对写请求的处理操作。

需要说明的是, 虽然本公开的例如图 3B 和图 4 中示出了具体的第一缓存 (例如, 主设备 0)、第二缓存 (例如, 主设备 1)、缓存一致性节点 (例如, 从设备) 和内存 (例如, MEM) 及其相关的具体操作, 但图中示出的内容仅为示例性的, 以方便对本公开的实施例进行说明, 但本公开的实施例不限于此, 在本公开的具体示例的基础上能够想到的任何其他操作均包括在本公开的范围內。

以下结合表 1-9 对本公开的实施例的缓存一致性协议的一些规定进行示例性描述。

30 例如, 本公开的实施例的协议规定的主设备 (或缓存) 均有多种状态, 如表 1 中所示:

表 1

缓存状态	说明	主存数据有效	其他缓存可能的状态
M	独占此数据且数据已被修改	否	I
E	独占此数据且数据未被修改	是	I
S	与其他缓存共享此数据	是	S
I	未缓存此数据	是	M/E/S/I

例如，在表 1 中，缓存状态可以为第一缓存的状态，主存可以是内存 MEM。例如，第一缓存可以包括 M 状态、E 状态、S 状态和 I 状态。M 状态表示第一缓存独占相应的数据，且该数据已被修改；E 状态表示第一缓存独占相应的数据，且该数据未被修改；S 状态表示第一缓存与其他主设备共享相应的数据；I 状态表示第一缓存未缓存相应的数据。例如，当第一缓存为 M 状态时，由于第一缓存已经修改了相应的数据，因此内存 MEM 中对应的数据已经无效，此时其他缓存均为未缓存有此修改后的数据的 I 状态；当第一缓存为 E 状态时，其他缓存均为未缓存有此数据的 I 状态；当第一缓存为 S 状态时，存在其他缓存处于共享此数据的 S 状态；当第一缓存为 I 状态时，其他缓存可以处于 M 状态、E 状态、S 状态和 I 状态中的任何一种状态。

需要说明的是，在本公开的实施例中，主设备（例如，第一缓存）的状态不限以上列出的四种状态，主设备可以根据不同需要而具有其他不同的状态。

例如，表 2 为本公开的协议规定的主设备可以发出的请求类型，以及对应的缓存（Cache）状态转换和从设备响应报文。

表 2

请求类型	发起请求时允许的缓存状态	请求结束时缓存状态	从设备返回给主设备的响应	说明
RdE	I/S/E	M/E	DatRspM/ DatRspE	获取独占数据副本
RdD	I	M/E/S	DatRspM/ DatRspE/ DatRspS	获取非 I 态的数据副本，即需要缓存此数据副本
RdS	I	S	DatRspS	获取共享数据副本
RdI	I	I	DatRspI	获取数据快照，不缓

				存此数据副本
RdEE	I/S	E	RspE	获取数据独占权限，不获取数据
WBI	M	I	WrRsp	将已缓存的数据副本写回主存，本缓存置为无效 I 态
WBS	M	I/S	WrRsp	将已缓存的数据副本写回主存，本缓存置为共享 S 态
WBE	M	I/S/E	WrRsp	将已缓存的数据副本写回主存，本缓存置为独占 E 态

下面以主设备为第一缓存时为例进行说明。

例如，在表 2 中，请求类型可以为第一缓存向从设备发起的请求类型。例如，第一缓存可以发起的读请求类型可以包括 RdE（获取独占数据副本）、RdD（获取非 I 状态的数据副本，即需要缓存此数据副本）、RdS（获取共享数据副本）、RdI（获取数据快照，不缓存此数据副本）和 RdEE（获取数据独占权限，不获取数据）。本公开的实施例不限于此，第一缓存也可以根据需要发起其他不同的读请求类型。

例如，当第一缓存向从设备发起的读请求类型为 RdE 时，第一缓存发起请求时的缓存状态（或缓存数据状态）可以为 I 状态、S 状态或 E 状态；当 RdE 请求结束时，第一缓存的缓存状态可以为 E 状态（在此情况下，从设备向第一缓存返回响应 DatRspE（参见下文的表 6），使得从设备获取独占态数据副本的响应）或 M 状态（在此情况下，从设备返回响应 DatRspM（参见下文的文表 6），使得从设备获取修改态数据副本的响应）。例如，当第一缓存向从设备发起的读请求类型为 RdD 时，第一缓存发起请求时的缓存状态可以为 I 状态；当 RdD 请求结束时，第一缓存的缓存状态可以为 S 状态（在此情况下，从设备向第一缓存返回响应 DatRspS（参见下文的表 6），使得从设备获取共享态数据副本的响应）、E 状态（在此情况下，从设备向第一缓存返回响应 DatRspE）或 M 状态（在此情况下，从设备向第一缓存返回响应 DatRspM）。例如，当第一缓存向从设备发起的读请求类型为 RdS 时，第一缓存发起请求时的缓存状态可以为 I 状态；当 RdS 请求结束时，第一缓存的

缓存状态可以为 S 状态(在此情况下,从设备向第一缓存返回响应 DatRspS)。例如,当第一缓存向从设备发起的读请求类型为 RdI 时,第一缓存发起请求时的缓存状态可以为 I 状态;当 RdI 请求结束时,第一缓存的缓存状态可以为 I 状态(在此情况下,从设备向第一缓存返回响应 DatRspI(参见下文的表 6),使得从设备获取数据副本的响应)。例如,当第一缓存向从设备发起的读请求类型为 RdEE 时,第一缓存发起请求时的缓存状态可以为 I 状态或 S 状态;当 RdEE 请求结束时,第一缓存的缓存状态可以为 E 状态(在此情况下,从设备返回响应 RspE(参见下文的表 6),使得从设备获取独占权限的响应)。

10 例如,在表 2 中,第一缓存可以发起的写请求类型可以包括 WBI(将已缓存的数据副本写回主存,本缓存置为无效 I 态)、WBS(将已缓存的数据副本写回主存,本缓存置为共享 S 态)和 WBE(将已缓存的数据副本写回主存,本缓存置为独占 E 态)。本公开的实施例不限于此,第一缓存也可以根据需要发起其他不同的写请求类型。

15 例如,当第一缓存向从设备发起的写请求类型为 WBI 时,第一缓存发起请求时的缓存状态可以为 M 状态;当写请求 WBI 结束时,第一缓存的缓存状态可以为 I 状态(在此情况下,从设备向第一缓存返回响应 WrRsp)。例如,当第一缓存向从设备发起的写请求类型为 WBS 时,第一缓存发起请求时的缓存状态可以为 M 状态;当写请求 WBS 结束时,第一缓存的缓存状态可以为 I 状态(在此情况下,从设备向第一缓存返回响应 WrRsp)或 S 状态(在此情况下,从设备向第一缓存返回响应 WrRsp)。例如,当第一缓存向从设备发起的写请求类型为 WBE 时,第一缓存发起请求时的缓存状态可以为 M 状态;当写请求 WBE 结束时,第一缓存的缓存状态可以为 I 状态(在此情况下,从设备向第一缓存返回响应 WrRsp)、S 状态(在此情况下,从设备向第一缓存返回响应 WrRsp)或 E 状态(在此情况下,从设备向第一缓存返回响应 WrRsp)。

25 例如,在图 3B 和图 4 中,第一缓存写请求不限于 WBI,写请求也可以是表 2 中示出的 WBS 或 WBE。例如,第一缓存的缓存数据溢出时,第一缓存可以向从设备发起写请求 WBI,响应于收到第一缓存要将第一数据写回内存 MEM 的写请求 WBI,缓存一致性节点将第一数据写入内存 MEM(如图 3B 所示)或将第一数据丢弃(如图 4 所示)。例如,在缓存一致性节点将第

一数据写入内存 MEM 或将第一数据丢弃之后，缓存一致性节点向第一缓存返回写响应信号 WrRsp 以直接结束对写请求的处理操作。例如，第一缓存的缓存数据溢出时，第一缓存可以向从设备发起写请求 WBS，响应于收到第一缓存要将第一数据写回内存 MEM 的写请求 WBS，缓存一致性节点将第一数据写入内存 MEM（如图 3B 所示）或将第一数据丢弃（如图 4 所示）。  
5 例如，在缓存一致性节点将第一数据写入内存 MEM 或将第一数据丢弃之后，缓存一致性节点向第一缓存返回写响应信号 WrRsp 以直接结束对写请求的处理操作。例如，第一缓存的缓存数据溢出时，第一缓存可以向从设备发起写请求 WBE，响应于收到第一缓存要将第一数据写回内存 MEM 的写请求  
10 WBE，缓存一致性节点将第一数据写入内存 MEM（如图 3B 所示）或将第一数据丢弃（如图 4 所示）。例如，在缓存一致性节点将第一数据写入内存 MEM 或将第一数据丢弃之后，缓存一致性节点向第一缓存返回写响应信号 WrRsp 以直接结束对写请求的处理操作。

例如，响应于收到第一缓存要将第一数据写回内存 MEM 的写请求，缓存一致性节点确定第一缓存是否与缓存一致性节点侧记录的当前独占第一数据的缓存相同，以及响应于确定第一缓存与缓存一致性节点侧记录的当前独占第一数据的缓存相同，缓存一致性节点将第一数据写入内存 MEM。  
15

例如，如图 3B 所示，响应于收到主设备 0 要将第一数据写回内存 MEM 的写请求，从设备确定主设备 0 是否与从设备侧记录的当前独占第一数据的缓存相同。例如，如图 3B 中所示，响应于确定主设备 0 与从设备侧记录的当前独占第一数据的缓存相同（例如，从设备侧记录了 Src:Master0==Dir@E:Master0），从设备通过写内存操作 Wrmem 将第一数据写回内存 MEM，内存 MEM 收到写内存操作 Wrmem 后向缓存一致性节点返回写响应 WrRsp 以向从设备指示内存 MEM 已经成功收到第一数据。例如，  
20 响应于确定主设备 0 与从设备侧记录的当前独占第一数据的缓存不相同（例如，从设备侧记录了 Src:Master0!=Dir@E:Master0），则判断从设备没有将第一数据写入内存 MEM 的权限，并且从设备不执行将第一数据写入内存 MEM 的操作。

例如，在缓存一致性节点将第一数据写入内存之后，内存向缓存一致性节点返回指示第一数据已经被写入内存的写响应，以及响应于从内存收到指示第一数据已经被写入内存的写响应，缓存一致性节点向第一缓存返回指示  
30

第一数据已经被写回内存的写响应信号以直接结束对写请求的处理操作。

例如，如图 3B 所示，在从设备将第一数据写入内存 MEM 之后，内存 MEM 向从设备返回指示第一数据已经被写入内存 MEM 的写响应 WrRsp，

以及响应于从内存 MEM 收到指示第一数据已经被写入内存 MEM 的写响应  
5 WrRsp，从设备向主设备 0 返回指示第一数据已经被写回内存 MEM 的写响应信号 WrRsp 以直接结束对写请求的处理操作。例如，此时主设备 0 不需要再向从设备发送报文就直接结束对写请求的处理操作。

例如，在缓存一致性节点将第一数据写入内存之后，缓存一致性节点将  
10 第一缓存的数据状态置为无效态、共享态或独占态。例如，无效态表示第一缓存没有缓存第一数据，共享态表示第一缓存与其他缓存均缓存有第一数据，独占态表示仅第一缓存中缓存有第一数据。

例如，在从设备将第一数据写回内存 MEM 之后，从设备将多个缓存中的  
15 第一缓存（例如，主设备 0）的数据状态为无效态（I 状态）、共享态（S 状态）或独占态（例如，E 状态）。例如，无效态表示第一缓存没有缓存第一数据，共享态表示第一缓存与其他缓存均缓存有第一数据，独占态表示仅第一缓存中缓存有第一数据。

例如，如以上表 2 所示，如果第一缓存向从设备发起写请求 WBI，则当  
20 写请求 WBI 结束时，第一缓存的缓存状态可以为 I 状态。例如，如果第一缓存向从设备发起写请求 WBS，则当写请求 WBS 结束时，第一缓存的缓存状态可以为 I 状态（例如，第一缓存将数据写回内存 MEM 后，第一缓存中缓存的该数据被其他数据替换）或 S 状态（例如，第一缓存将数据写回内存 MEM 后，第一缓存中缓存的该数据未被其他数据替换）。例如，如果第一缓存向从设备发起写请求 WBE，则当写请求 WBE 结束时，第一缓存的缓存状态  
25 可以为 I 状态（例如，第一缓存将数据写回内存 MEM 后，第一缓存中缓存的该数据被其他数据替换）、S 状态（例如，第一缓存写回内存 MEM 的数据被其他缓存共享）或 E 状态（例如，第一缓存写回内存 MEM 的数据未被其他缓存共享，并且第一缓存将数据写回内存 MEM 后，第一缓存中缓存的该数据未被其他数据替换）。

例如，如图 4 所示，响应于收到第二缓存（或“主设备 1”）对第一数据  
30 的读请求且缓存一致性节点侧（例如，从设备对应的目录）记录第一数据在第一缓存（或“主设备 0”）中处于独占状态或共享状态，缓存一致性节点向

第一缓存发起监测信号（例如，ExpI）；响应于第一缓存在发出写请求之后收到监测信号，第一缓存响应于监测信号向缓存一致性节点返回携带第一数据的监测响应（例如，ExpRspIDat）。

例如，表 3 为本公开的协议规定的监测信号类型及其说明，检测信号 5 Exp\*用于监测主设备处的缓存状态和数据，并将被监测的主设备置为相应的状态。

表 3

监测信号类型	说明
ExpI	监测目标缓存，获取最新数据副本，并将目标缓存数据置为无效状态。
ExpS	监测目标缓存，获取最新数据副本，并将目标缓存数据置为无效态或共享状态。
ExpE	监测目标缓存，获取最新数据副本，不改变目标缓存数据状态。

表 4 为本公开的协议规定的监测响应类型，即主设备被检测信号 Exp\* 10 监测时可能向从设备返回的监测响应。

表 4

监测响应类型	说明
ExpRspI	无数据监测响应，被监测缓存置为 I 态。
ExpRspS	无数据监测响应，被监测缓存置为 S 态。
ExpRspE	无数据监测响应，被监测缓存置为 E/M 态。
ExpRspIDat	有数据监测响应，被监测缓存置为 I 态。
ExpRspSDat	有数据监测响应，被监测缓存置为 S 态。
ExpRspEDat	有数据监测响应，被监测缓存置为 E/M 态。

表 5 为本公开的协议规定的非监测类响应事务类型及其说明。

表 5

响应类型	说明
RspE	使得从设备获取独占权限的响应

DatRspI	使得从设备获取数据副本的响应
DatRspS	使得从设备获取共享态数据副本的响应
DatRspE	使得从设备获取独占态数据副本的响应
DatRspM	使得从设备获取修改态数据副本的响应
WrRsp	写响应
ACK	读确认

表6为本公开的协议规定的监测事务在主设备侧的状态转移表以及监测响应动作，定义了主设备侧缓存被监测后可能转换到的目标状态，以及发送相应的监测响应报文。

5

表6

监测信号	当前缓存状态	处理监测信号后的缓存状态	监测响应
ExpI	I/S/E/M	I	ExpRspI/ExpRspIDat
ExpS	I	I	ExpRspI
	S	S	ExpRspS
	E/M	S	ExpRspS/ExpRspSDat
ExpE	I	I	ExpRspI
	S	S	ExpRspS
	E/M	E/M	ExpRspE/ExpRspEDat

例如，如表6中所示，当第一缓存收到检测信号 ExpI 后，如果第一缓存的当前缓存状态为 I、S、E 或 M，则处理监测信号后第一缓存的状态为 I，第一缓存向缓存一致性节点返回检测响应 ExpRspI（例如，在第一缓存的当前状态为 I 时）或 ExpRspIDat（例如，在第一缓存的当前状态为 S、E 或 M 时）。

例如，当第一缓存收到检测信号 ExpS 后，如果第一缓存的当前缓存状态为 I，则处理监测信号后第一缓存的状态为 I，第一缓存向缓存一致性节点返回检测响应 ExpRspI；当第一缓存收到检测信号 ExpS 后，如果第一缓存的当前缓存状态为 S，则处理监测信号后第一缓存的状态为 S，第一缓存向缓存一致性节点返回检测响应 ExpRspS；当第一缓存收到检测信号 ExpS 后，

如果第一缓存的当前缓存状态为 E 或 M, 则处理监测信号后第一缓存的状态为 S, 第一缓存向缓存一致性节点返回检测响应 ExpRspS (例如, 在第一缓存执行写请求 WBS 或 WBE 时) 或 ExpRspSDat (例如, 在第一缓存执行写请求 WBI 时)。

- 5 例如, 当第一缓存收到检测信号 ExpE 后, 如果第一缓存的当前缓存状态为 I, 则处理监测信号后第一缓存的状态为 I, 第一缓存向缓存一致性节点返回检测响应 ExpRspI; 当第一缓存收到检测信号 ExpE 后, 如果第一缓存的当前缓存状态为 S, 则处理监测信号后第一缓存的状态为 S, 第一缓存向缓存一致性节点返回检测响应 ExpRspS; 当第一缓存收到检测信号 ExpE 后, 10 如果第一缓存的当前缓存状态为 E 或 M, 则处理监测信号后第一缓存的状态为 E 或 M, 第一缓存向缓存一致性节点返回检测响应 ExpRspE (例如, 在第一缓存执行写请求 WBS 或 WBE 时) 或 ExpRspEDat (例如, 例如, 在第一缓存执行写请求 WBI 时)。

表 7 为本公开的协议规定的监测事务在主设备侧处理且存在冲突的情况, 即监测事务到达主设备时, 主设备当前已经对相同地址的缓存行发起了读事务 (或读请求) Rd\*或写事务 (或写请求) WB\*。

表 7

监测信号类型	当前缓存状态	处理监测命令后的缓存状态	冲突请求类型	监测响应
ExpI	I/S/E	I	Rd*	ExpRspI/ ExpRspIDat
	M	I	WB*	ExpRspIDat
ExpS	I/S/E	I	Rd*	ExpRspI/ ExpRspIDat
	M	I	WB*	ExpRspIDat
ExpE	I/S/E	I	Rd*	ExpRspI/ ExpRspIDat
	M	I	WB*	ExpRspIDat

例如, 如表 7 中所示, 当缓存一致性节点向第一缓存发送的检测信号为

ExpI, 且监测事务到达第一缓存时, 第一缓存已经对相同地址的缓存行发起了读事务 Rd\*时, 如果第一缓存当前状态为 I、S 或 E, 则处理监测信号后第一缓存的状态为 I, 第一缓存向缓存一致性节点返回检测响应 ExpRspI (例如, 在主设备的状态为 I 时) 或 ExpRspIDat (例如, 在第一缓存的状态为 S 或 E 时); 当缓存一致性节点向第一缓存发送的检测信号为 ExpI, 且监测事务到达第一缓存时, 第一缓存已经对相同地址的缓存行发起了写事务 WB\*时, 如果第一缓存当前状态为 M, 则第一缓存向缓存一致性节点返回检测响应 ExpRspIDat, 处理监测信号后第一缓存的状态为 I。

例如, 如表 7 中所示, 当缓存一致性节点向第一缓存发送的检测信号为 ExpS, 且监测事务到达第一缓存时, 第一缓存已经对相同地址的缓存行发起了读事务 Rd\*时, 如果第一缓存当前状态为 I、S 或 E, 则处理监测信号后第一缓存的状态为 I, 第一缓存向缓存一致性节点返回检测响应 ExpRspI (例如, 在第一缓存的状态为 I 时) 或 ExpRspIDat (例如, 在第一缓存的状态为 S 或 E 时); 当缓存一致性节点向第一缓存发送的检测信号为 ExpS, 且监测事务到达第一缓存时, 第一缓存已经对相同地址的缓存行发起了读事务 WB\*时, 如果第一缓存当前状态为 M, 第一缓存向缓存一致性节点返回检测响应 ExpRspIDat, 处理监测信号 ExpS 后第一缓存的状态为 I。

例如, 如表 7 中所示, 当缓存一致性节点向第一缓存发送的检测信号为 ExpE, 且监测事务到达第一缓存时, 第一缓存已经对相同地址的缓存行发起了读事务 Rd\*时, 如果第一缓存当前状态为 I、S 或 E, 则处理监测信号后第一缓存的状态为 I, 第一缓存向缓存一致性节点返回检测响应 ExpRspI (例如, 在第一缓存的状态为 I 时) 或 ExpRspIDat (例如, 在第一缓存的状态为 S 或 E 时); 当缓存一致性节点向第一缓存发送的检测信号为 ExpE, 且监测事务到达第一缓存时, 第一缓存已经对相同地址的缓存行发起了读事务 WB\*时, 如果第一缓存当前状态为 M, 第一缓存向缓存一致性节点返回检测响应 ExpRspIDat, 处理监测信号 ExpE 后第一缓存的状态为 I。

表 8 为本公开的协议定义的从设备收到主设备读写请求后, 可能需要根据目录状态进行的状态转移 (包括触发监测命令, 向主设备返回响应命令, 或对内存数据进行读写的处理)。

表 8

请求类型	目录状态	触发的监测	触发的响应	内存数据
------	------	-------	-------	------

		信号	信号	
RdE	I			读取内存数据
	S/E	ExpI		
RdD	I/S			读取内存数据
	E	ExpS		
RdS	I/S			读取内存数据
	E	ExpS		
RdI	I/S			读取内存数据
	E	ExpS		
RdEE	I		RspE	
	S/E	ExpE		
WBI	I/S/E		WrRsp	独占且指向发起 WBI 的 Master 则数据写入内存，否则丢弃数据
WBS	I/S/E		WrRsp	
WBE	I/S/E		WrRsp	

例如，如表 8 中所示，响应于收到第二缓存对第一数据的读请求为 RdE 且第一数据在缓存一致性节点对应的目录中处于 I 状态，缓存一致性节点直接从内存中读取第一数据，并将第一数据返回给第二缓存。例如，响应于收到第二缓存对第一数据的读请求 RdE 且缓存一致性节点侧（例如，从设备对应的目录）记录第一数据在缓存一致性节点对应的目录中处于 S 状态或独占状态 E，缓存一致性节点向第一缓存发起监测信号 ExpI（监测目标缓存，获取最新数据副本，并将其缓存数据置为无效状态）。

例如，响应于收到监测响应，缓存一致性节点将第一数据返回给第二缓存，以及在收到第一数据之后，第二缓存向缓存一致性节点返回指示已经收到第一数据的报文以结束读请求。例如，如图 4 所示，响应于收到监测响应 ExpRspIDat，从设备响应于监测响应 ExpRspIDat 而（例如，通过响应信号 DatRspM）将第一数据返回给主设备 1，并且将主设备 1 在从设备侧的数据

状态置为独占态（例如，从设备侧记录 Dir@E:Master1），独占态（或 Dir@E:Master1）表示仅主设备 1 中缓存有第一数据。例如，在收到第一数据之后，主设备 1 向从设备返回指示已经收到第一数据的报文 ACK 以结束读请求 RdE。

5 例如，在收到报文之后，缓存一致性节点丢弃写请求写入缓存一致性节点侧的第一数据，并且向第一缓存返回写响应信号以直接结束对写请求的处理操作。例如，如图 4 所示，在从设备收到主设备 1 的报文 ACK 之后，主设备之前发出的读请求 RdE 处理完成，从设备侧的处于阻塞状态（在从设备处理主设备 1 发起的读请求后处于阻塞状态（即，由于从设备对相同地址事务同一时刻只能处理一个，因此在主设备 1 的 RdE 处理完成之前主设备 0 所发起的写请求 WB\*一直处于阻塞状态））的主设备 0 之前发出的写请求 WB\*被唤醒继续处理，此时由于从设备侧记录了主设备 1 已经获取了第一数据的独占状态，而写请求由主设备 0 发起，因此主设备 0 的写事务中的数据已经是过期的无效数据，所以从设备判断此写事务或写操作由于冲突应为无效数据（如图 4 中所示，Src:Master0!=Dir@E:Master1，即，请求回写的主设备 0 不是从设备指向的独占第一数据的主设备 1），直接丢弃由主设备 0 已经写入从设备的数据。例如，从设备确定写请求写入从设备侧的第一数据为无效数据，从设备丢弃写请求写入从设备侧的第一数据，并且向主设备 0 返回写响应信号 WrRsp 以直接结束对写请求的处理操作。例如，从设备收到写响应信号 WrRsp 后直接结束对应写事务处理，不需要再向从设备发送报文 ACK。

例如，响应于收到第二缓存对第一数据的读请求为 RdD、RdS 或 RdI 且第一数据在缓存一致性节点对应的目录中处于 I 状态或 S 状态，缓存一致性节点可以直接从内存中读取第一数据，并将第一数据返回给第二缓存。例如，响应于收到第二缓存对第一数据的读请求 RdD、RdS 或 RdI 且缓存一致性节点侧（例如，从设备对应的目录）记录第一数据在缓存一致性节点对应的目录中处于独占状态 E，缓存一致性节点向第一缓存发起监测信号 ExpS（监测目标缓存，获取最新数据副本，并将其缓存数据置为无效或共享状态）。

30 例如，响应于收到监测响应，缓存一致性节点将第一数据返回给第二缓存，以及在收到第一数据之后，第二缓存向缓存一致性节点返回指示已经收

到第一数据的报文以结束读请求。例如，如图 4 所示，响应于收到监测响应 ExpRspIDat，从设备响应于监测响应 ExpRspIDat 而（例如，通过响应信号 DatRspM）将第一数据返回给主设备 1，并且将主设备 1 在从设备侧的数据状态置为独占态（例如，从设备侧记录 Dir@E:Master1），独占态（或 Dir@E:Master1）表示仅主设备 1 中缓存有第一数据。例如，在收到第一数据之后，主设备 1 向从设备返回指示已经收到第一数据的报文 ACK 以结束读请求（例如，RdD 或 RdS）。

例如，在收到报文之后，缓存一致性节点丢弃写请求写入缓存一致性节点侧的第一数据，并且向第一缓存返回写响应信号以直接结束对写请求的处理操作。例如，如图 4 所示，在从设备收到主设备 1 的报文 ACK 之后，主设备之前发出的读请求 RdD、RdS 或 RdI 处理完成，从设备侧的处于阻塞状态的主设备 0 之前发出的写请求 WB\*被唤醒继续处理，此时由于从设备侧记录了主设备 1 已经获取了第一数据的独占状态或共享状态，而写请求由主设备 0 发起，因此主设备 0 的写事务中的数据已经是过期的无效数据，所以从设备判断此写事务或写操作由于冲突应为无效数据，直接丢弃由主设备 0 已经写入从设备的数据。例如，从设备确定写请求写入从设备侧的第一数据为无效数据，从设备丢弃写请求写入从设备侧的第一数据，并且向主设备 0 返回写响应信号 WrRsp 以直接结束对写请求的处理操作。例如，从设备收到写响应信号 WrRsp 后直接结束对应写事务处理，不需要再发送报文 ACK。

例如，响应于收到第二缓存对第一数据的读请求为 RdEE 且第一数据在缓存一致性节点对应的目录中处于 I 状态，缓存一致性节点向第二缓存返回响应 RspE（无数据监测响应，被监测缓存置为 E 态或 M 态）以结束读请求 RdEE。例如，响应于收到第二缓存对第一数据的读请求 RdEE 且缓存一致性节点侧（例如，从设备对应的目录）记录第一数据在缓存一致性节点对应的目录中处于 S 状态或独占状态 E，缓存一致性节点向第一缓存发起监测信号 ExpE（监测目标缓存，获取最新数据副本，不改变其缓存数据状态）。

例如，响应于收到监测响应，缓存一致性节点将第一数据返回给第二缓存，以及在收到第一数据之后，第二缓存向缓存一致性节点返回指示已经收到第一数据的报文以结束读请求。例如，如图 4 所示，响应于收到监测响应 ExpRspIDat，从设备响应于监测响应 ExpRspIDat 而（例如，通过响应信号 DatRspM）将第一数据返回给主设备 1。例如，在收到第一数据之后，主设

备 1 向从设备返回指示已经收到第一数据的报文 ACK 以结束读请求 RdEE。

例如，在收到报文之后，缓存一致性节点丢弃写请求写入缓存一致性节点侧的第一数据，并且向第一缓存返回写响应信号以直接结束对写请求的处理操作。例如，如图 4 所示，在从设备收到主设备 1 的报文 ACK 之后，主设备之前发出的读请求 RdEE 处理完成，从设备侧的处于阻塞状态的主设备 0 之前发出的写请求 WB\*被唤醒继续处理，此时由于从设备侧记录了主设备 0 已经不是独占第一数据的状态，而写请求由主设备 0 发起，因此主设备 0 的写事务中的数据已经是过期的无效数据，所以从设备判断此写事务或写操作由于冲突应为无效数据，直接丢弃由主设备 0 已经写入从设备的数据。例如，从设备确定写请求写入从设备侧的第一数据为无效数据，从设备丢弃写请求写入从设备侧的第一数据，并且向主设备 0 返回写响应信号 WrRsp 以直接结束对写请求的处理操作。例如，从设备收到写响应信号 WrRsp 后直接结束对应写事务处理，不需要再发送报文 ACK。

例如，响应于收到第二缓存向缓存一致性节点发送对数据的写请求（例如，WBI、WBS 或 WBE），无论此时目录处于什么状态（例如，I、S 或 E），均不会不会触发检测信号，缓存一致性节点向第二缓存返回响应 WrRsp。例如，如果此时缓存一致性节点侧记录第二缓存需要写入的数据处于独占状态且指向发起写请求 WBI 的第二缓存，则将该数据写入内存，否则丢弃数据。

表 9 为本公开的协议规定的从设备从主设备接收响应后的状态转换表。

表 9

主设备向从设备发送的响应信号	从设备收到的请求类型	从设备向主设备发送的响应	内存数据	请求事务表项
ExpRspI	*		读取内存数据	WaitACK
ExpRspS	*		读取内存数据	WaitACK
ExpRspE	*		读取内存数据	WaitACK
ExpRspIDat	RdI	DatRspI		WaitACK
	RdS	DatRspS		WaitACK
	RdD/RdE	DatRspE/DatRspM		WaitACK
ExpRspSDat	RdS/RdD	DatRspS		WaitACK

ExpRspEDat	RdI	DatRspI		WaitACK
ACK	Rd*			结束

例如，如表 9 中所示，如果第一缓存向缓存一致性节点发送响应信号 ExpRspI、ExpRspS 或 ExpRspE，且缓存一致性节点收到第二缓存的任何类型的读请求（例如，缓存一致性节点响应于收到第二缓存的读请求，创建读请求事务处理表项，记录当前读请求事务相关的状态），则在收到第一缓存向缓存一致性节点发送的响应信号 ExpRspI、ExpRspS 或 ExpRspE 后，缓存一致性节点需要读取内存数据。例如，在缓存一致性节点读取内存数据之后，将该数据发送给第二缓存，并且将与第二缓存的读请求对应的读请求事务处理表项的状态置为等待确认状态（例如，WaitACK），等待第二缓存发送确认已经成功接收数据的确认信号（ACK）。

例如，如表 9 中所示，如果第一缓存向缓存一致性节点发送响应信号 ExpRspIDat，且缓存一致性节点收到第二缓存的请求类型为读请求 RdI（例如，在缓存一致性节点响应于收到第二缓存的读请求 RdI，创建读请求事务处理表项，记录当前读请求事务相关的状态），则在收到第一缓存向缓存一致性节点发送的响应信号 ExpRspIDat 后，缓存一致性节点向第二缓存发送携带数据的响应信号 DatRspI，并且将与第二缓存的读请求 RdI 对应的读请求事务处理表项的状态置为等待确认状态（例如，WaitACK），等待第二缓存发送确认已经成功接收数据的确认信号（ACK）。

例如，如表 9 中所示，如果第一缓存向缓存一致性节点发送响应信号 ExpRspIDat，且缓存一致性节点收到第二缓存的请求类型为读请求 RdS（例如，在缓存一致性节点响应于收到第二缓存的读请求 RdS 后，创建读请求事务处理表项，记录当前读请求事务相关的状态），则在收到第一缓存向缓存一致性节点发送的响应信号 ExpRspIDat 后，缓存一致性节点向第二缓存发送携带数据的响应信号 DatRspS，并且将与第二缓存的读请求 RdS 对应的读请求事务处理表项的状态置为等待确认状态（例如，WaitACK），等待第二缓存发送确认已经成功接收数据的确认信号（ACK）。

例如，如表 9 中所示，如果第一缓存向缓存一致性节点发送响应信号 ExpRspIDat，且缓存一致性节点收到第二缓存的请求类型为读请求 RdD（例如，在缓存一致性节点响应于收到第二缓存的读请求 RdD 后，创建读请求

事务处理表项，记录当前读请求事务相关的状态)，则在收到第一缓存向缓存一致性节点发送的响应信号 ExpRspIDat 后，缓存一致性节点向第二缓存发送携带数据的响应信号 DatRspE (记录第二缓存独占此数据) 或 DatRspM (记录第二缓存独占此被修改后的数据)，并且将与第二缓存的读请求 RdD 对应的读请求事务处理表项的状态置为等待确认状态 (例如，WaitACK)，等待第二缓存发送确认已经成功接收数据的确认信号 (ACK)。

例如，如表 9 中所示，如果第一缓存向缓存一致性节点发送响应信号 ExpRspIDat，且缓存一致性节点收到第二缓存的请求类型为读请求 RdE (例如，在缓存一致性节点响应于收到第二缓存的读请求 RdE 后，创建读请求事务处理表项，记录当前读请求事务相关的状态)，则在收到第一缓存向缓存一致性节点发送的响应信号 ExpRspIDat 后，缓存一致性节点向第二缓存发送携带数据的响应信号 DatRspE (记录第二缓存独占此数据) 或 DatRspM (记录第二缓存独占此被修改后的数据)，并且将与第二缓存的读请求 RdE 对应的读请求事务处理表项的状态置为等待确认状态 (例如，WaitACK)，等待第二缓存发送确认已经成功接收数据的确认信号 (ACK)。

例如，如表 9 中所示，如果第一缓存向缓存一致性节点发送响应信号 ExpRspSDat，且缓存一致性节点收到第二缓存的请求类型为读请求 RdS 或 RdD (例如，在缓存一致性节点响应于收到第二缓存的读请求 RdS 或 RdD 后，创建读请求事务处理表项，记录当前读请求事务相关的状态)，则在收到第一缓存向缓存一致性节点发送的响应信号 ExpRspSDat 后，缓存一致性节点向第二缓存发送携带数据的响应信号 DatRspS (记录第二缓存共享此数据)，并且将与第二缓存的读请求 RdS 或 RdD 对应的读请求事务处理表项的状态置为等待确认状态 (例如，WaitACK)，等待第二缓存发送确认已经成功接收数据的确认信号 (ACK)。

例如，如表 9 中所示，如果第一缓存向缓存一致性节点发送响应信号 ExpRspEDat，且缓存一致性节点收到第二缓存的请求类型为读请求 RdI (例如，在缓存一致性节点响应于收到第二缓存的读请求 RdI 后，创建读请求事务处理表项，记录当前读请求事务相关的状态)，则在收到第一缓存向缓存一致性节点发送的响应信号 ExpRspEDat 后，缓存一致性节点向第二缓存发送携带数据的响应信号 DatRspI (记录第二缓存获取此数据的副本)，并且将与第二缓存的读请求 RdI 对应的读请求事务处理表项的状态置为等待确认状

态（例如，WaitACK），等待第二缓存发送确认已经成功接收数据的确认信号（ACK）。

例如，如表 9 中所示，如果第一缓存向缓存一致性节点发送响应信号 ACK，且缓存一致性节点收到第二缓存的请求类型为读请求 Rd\*（例如，在  
5 缓存一致性节点响应于收到第二缓存的读请求 Rd\*后，创建读请求事务处理表项或写请求事务处理表项，记录当前读请求事务相关的状态），则在收到第一缓存向缓存一致性节点发送的响应信号 ACK 后，缓存一致性节点将读请求事务处理表项的状态置为结束状态，以确认已经成功完成数据的读取。

在本公开的实施例的数据处理方法中，通过回写事务和目录的比较，进  
10 行场景识别，主设备收到写响应后直接结束写事务处理，不需要再向从设备发送 ACK 报文，从而降低 ACK 报文的数量，减少协议对 ACK 报文的需求，减少 ACK 报文传输产生的功耗和延时以及带宽占用，并且优化了主设备侧缓冲区的周转率，在相同性能下主设备侧需要的缓冲区数量更少，节省面积和功耗。

15 图 5 示出了本公开至少一实施例提供的一种电子设备 60 的示意图。例如，该电子设备 60 包括缓存一致性节点 605、第一缓存 606 和第二缓存 607。例如，缓存一致性节点 605 配置为响应于收到第一缓存 606 要将第一数据写回内存的写请求，将第一数据写入内存或将第一数据丢弃，以及在缓存一致性节点 605 将第一数据写入内存或将第一数据丢弃之后，向第一缓存 606 返  
20 回写响应信号以直接结束对写请求的处理操作。例如，该写请求携带第一数据，缓存一致性节点 605 配置为维护由多个缓存中存储的数据的一致性。本公开的电子设备 60 包括的缓存不限于图 5 中示出的两个缓存（第一缓存 606 和第二缓存 607），电子设备 60 的数量可以根据需要设置为任何数量。

例如，缓存一致性节点 605 还配置为：响应于收到第一缓存 606 要将第  
25 一数据写回内存的写请求，确定第一缓存 606 是否与缓存一致性节点 605 侧记录的当前独占第一数据的缓存相同，以及响应于确定第一缓存 606 与缓存一致性节点 605 侧记录的当前独占第一数据的缓存相同，将第一数据写入内存。

例如，内存配置为：在缓存一致性节点 605 将第一数据写入内存之后，  
30 向缓存一致性节点 605 返回指示第一数据已经被写入内存的写响应。例如，缓存一致性节点 605 还配置为：响应于从内存收到指示第一数据已经被写入

内存的写响应，向第一缓存 606 返回指示第一数据已经被写回内存的写响应信号以直接结束对写请求的处理操作。

例如，缓存一致性节点 605 还配置为：在缓存一致性节点 605 将第一数据写入内存之后，将第一缓存 606 的数据状态置为无效态、共享态或独占态。

5 例如，无效态表示第一缓存 606 没有缓存第一数据，共享态表示第一缓存 606 与其他缓存均缓存有第一数据，独占态表示仅第一缓存 606 中缓存有第一数据。

例如，缓存一致性节点 605 还配置为：响应于收到第二缓存 607 对第一数据的读请求且缓存一致性节点 605 侧记录第一数据在第一缓存 606 中处于  
10 独占状态或共享状态，向第一缓存 606 发起监测信号。例如，第一缓存 606 配置为：响应于第一缓存 606 在发出写请求之后收到监测信号，向缓存一致性节点 605 返回携带第一数据的监测响应。

例如，缓存一致性节点 605 还配置为：响应于收到监测响应，将第一数据返回给第二缓存 607。例如，第二缓存 607 还配置为：在收到第一数据之后，  
15 向缓存一致性节点 605 返回指示已经收到第一数据的报文以结束读请求。

例如，缓存一致性节点 605 还配置为：响应于收到监测响应，将第一数据返回给第二缓存 607，并且将第二缓存 607 的数据状态置为独占态。例如，独占态表示仅第二缓存 607 中缓存有第一数据。

例如，缓存一致性节点 605 还配置为：在收到报文之后，丢弃写请求写入缓存一致性节点 605 侧的第一数据，并且向第一缓存 606 返回写响应信号  
20 以直接结束对写请求的处理操作。

例如，缓存一致性节点 605 还配置为：响应于在收到报文之后确定写请求写入缓存一致性节点 605 侧的第一数据为无效数据，丢弃写请求写入缓存一致性节点 605 侧的第一数据。

25 在本公开的实施例的电子设备 60 中，通过回写事务和目录的比较，进行场景识别，主设备收到写响应后直接结束写事务处理，不需要再向从设备发送 ACK 报文，从而降低 ACK 报文的数量，减少协议对 ACK 报文的需求，减少 ACK 报文传输产生的功耗和延时以及带宽占用，并且优化了主设备侧缓冲区的周转率，在相同性能下主设备侧需要的缓冲区数量更少，节省面积  
30 和功耗。

本公开至少一些实施例还提供了一种数据处理装置，该数据处理装置包

括存储器和处理器。例如，存储器配置为存储计算机可执行指令，处理器配置为执行计算机可执行指令。例如，计算机可执行指令被处理器执行时实现本公开至少一实施例提供的数据处理方法。

5 图 6 示出了本公开至少一实施例提供的一种数据处理装置 500 的示意图。

如图 6 所示，根据本公开实施例的电子设备 500 包括处理器 501 以及存储器 502，处理器 501 和存储器 502 可以通过总线 503 进行互连。

10 处理器 501 可以根据存储在存储器 502 中的程序或代码执行各种动作和处理。具体地，处理器 501 可以是一种集成电路芯片，具有信号的处理能力。例如，上述处理器 501 可以是通用处理器、数字信号处理器 (DSP)、专用集成电路 (ASIC)、现成可编程门阵列 (FPGA) 或者其他可编程逻辑器件、分立门或者晶体管逻辑器件、分立硬件组件，可以实现或者执行本公开实施例中公开的各种方法和步骤。通用处理器可以是微处理器或者该处理器也可以是任何常规的处理器等，可以是 X86 架构或者是 ARM 架构等。

15 存储器 502 用于非暂时性存储计算机可执行指令，处理器 501 用于运行计算机可执行指令。当计算机可执行指令在被处理器 501 执行时实现本公开至少一实施例提供的数据处理方法。

20 例如，存储器 502 可以是易失性存储器或非易失性存储器，或可包括易失性和非易失性存储器两者。非易失性存储器可以是只读存储器 (ROM)、可编程只读存储器 (PROM)、可擦除可编程只读存储器 (EPROM)、电可擦除可编程只读存储器 (EEPROM) 或闪存。易失性存储器可以是随机存取存储器 (RAM)，其用作外部高速缓存。通过示例性但不是限制性说明，许多形式的 RAM 可用，例如静态随机存取存储器 (SRAM)、动态随机存取存储器 (DRAM)、同步动态随机存取存储器 (SDRAM)、双倍数据速率同步动态随机存取存储器 (DDRSDRAM)、增强型同步动态随机存取存储器 (ESDRAM)、同步连接动态随机存取存储器 (SLDRAM) 和直接内存总线随机存取存储器 (DRRAM)。应注意，本文描述的方法的存储器旨在包括但不限于这些和任意其它适合类型的存储器。

30 在本公开的实施例的数据处理装置 500 中，通过回写事务和目录的比较，进行场景识别，主设备收到写响应后直接结束写事务处理，不需要再向从设备发送 ACK 报文，从而降低 ACK 报文的数量，减少协议对 ACK 报文的需

求，减少 ACK 报文传输产生的功耗和延时以及带宽占用，并且优化了主设备侧缓冲区的周转率，在相同性能下主设备侧需要的缓冲区数量更少，节省面积和功耗。

本公开的至少一个实施例还提供了一种非暂时性存储介质，非暂时性地存储计算机可执行指令。例如，当计算机可执行指令由处理器执行时，实现本公开至少一实施例提供的数据处理方法。

图 7 是本公开一些实施例提供的一种非暂时性存储介质的示意图。如图 7 所示，非暂时性存储介质 600 可以非暂时性地存储计算机可执行指令 610，计算机可执行指令 610 在被计算机执行时实现本公开任一实施例提供的数据处理方法。

类似地，本公开实施例中的非暂时性存储介质可以是易失性存储器或非易失性存储器，或可包括易失性和非易失性存储器两者。应注意，本文描述的方法的存储器旨在包括但不限于这些和任意其它适合类型的存储器。

上述非暂时性存储介质的技术效果与上述数据处理方法的技术效果相同，此处不再赘述。

需要说明的是，附图中的流程图和框图，图示了按照本公开的各种实施例的系统、方法和计算机程序产品的可能实现的体系架构、功能和操作。在这点上，流程图或框图中的每个方框可以代表一个模块、程序段、或代码的一部分，所述模块、程序段、或代码的一部分包含至少一个用于实现规定的逻辑功能的可执行指令。也应当注意，在有些作为替换的实现中，方框中所标注的功能也可以以不同于附图中所标注的顺序发生。例如，两个接连地表示的方框实际上可以基本并行地执行，它们有时也可以按相反的顺序执行，这依所涉及的功能而定。也要注意的，框图和/或流程图中的每个方框、以及框图和/或流程图中的方框的组合，可以用执行规定的功能或操作的专用的基于硬件的系统来实现，或者可以用专用硬件与计算机指令的组合来实现。

一般而言，本公开的各种示例实施例可以在硬件或专用电路、软件、固件、逻辑，或其任何组合中实施。某些方面可以在硬件中实施，而其他方面可以在可以由控制器、微处理器或其他计算设备执行的固件或软件中实施。当本公开的实施例的各方面被图示或描述为框图、流程图或使用某些其他图形表示时，将理解此处描述的方框、装置、系统、技术或方法可以作为非限制性的示例在硬件、软件、固件、专用电路或逻辑、通用硬件或控制器或其

他计算设备，或其某些组合中实施。

对于本公开，还有以下几点需要说明：

(1) 本公开实施例的附图只涉及到与本公开实施例涉及到的结构，其他结构可参考通常设计。

5 (2) 为了清晰起见，在用于描述本公开的实施例的附图中，层或结构的厚度和尺寸被放大。可以理解，当诸如层、膜、区域或基板之类的元件被称作位于另一元件“上”或“下”时，该元件可以“直接”位于另一元件“上”或“下”，或者可以存在中间元件。

10 (3) 在不冲突的情况下，本公开的实施例及实施例中的特征可以相互组合以得到新的实施例。

以上所述仅为本公开的具体实施方式，但本公开的保护范围并不局限于此，本公开的保护范围应以所述权利要求的保护范围为准。

## 权利要求书

### 1.一种数据处理方法，包括：

响应于收到第一缓存要将第一数据写回内存的写请求，缓存一致性节点将所述第一数据写入内存或将所述第一数据丢弃，其中，所述写请求携带所述第一数据，所述缓存一致性节点配置为维护由多个缓存中存储的数据的一致性，所述多个缓存包括所述第一缓存，以及

在所述缓存一致性节点将所述第一数据写入内存或将所述第一数据丢弃之后，所述缓存一致性节点向所述第一缓存返回写响应信号以直接结束对所述写请求的处理操作。

2.根据权利要求1所述的数据处理方法，所述响应于收到第一缓存要将第一数据写回内存的写请求，缓存一致性节点将所述第一数据写入内存或将所述第一数据丢弃，包括：

响应于收到所述第一缓存要将所述第一数据写回所述内存的所述写请求，所述缓存一致性节点确定所述第一缓存是否与所述缓存一致性节点侧记录的当前独占所述第一数据的缓存相同，以及

响应于确定所述第一缓存与所述缓存一致性节点侧记录的当前独占所述第一数据的缓存相同，所述缓存一致性节点将所述第一数据写入所述内存。

3.根据权利要求1或2所述的数据处理方法，所述在所述缓存一致性节点将所述第一数据写入内存或将所述第一数据丢弃之后，所述缓存一致性节点向所述第一缓存返回写响应信号以直接结束对所述写请求的处理操作，包括：

在所述缓存一致性节点将所述第一数据写入内存之后，所述内存向所述缓存一致性节点返回指示所述第一数据已经被写入所述内存的写响应，以及

响应于从所述内存收到指示所述第一数据已经被写入所述内存的所述写响应，所述缓存一致性节点向所述第一缓存返回指示所述第一数据已经被写回所述内存的所述写响应信号以直接结束对所述写请求的处理操作。

4.根据权利要求 1-3 中的任一项所述的数据处理方法，还包括：

在所述缓存一致性节点将所述第一数据写入所述内存之后，所述缓存一致性节点将所述第一缓存的数据状态置为无效态、共享态或独占态，

其中，所述无效态表示所述第一缓存没有缓存所述第一数据，所述共享态表示所述第一缓存与其他缓存均缓存有所述第一数据，所述独占态表示仅所述第一缓存中缓存有所述第一数据。

5.根据权利要求 1-4 中的任一项所述的数据处理方法，还包括：

响应于收到第二缓存对所述第一数据的读请求且所述缓存一致性节点侧记录所述第一数据在所述第一缓存中处于独占状态或共享状态，所述缓存一致性节点向所述第一缓存发起监测信号；

响应于所述第一缓存在发出所述写请求之后收到所述监测信号，所述第一缓存响应于所述监测信号向所述缓存一致性节点返回携带所述第一数据的监测响应。

6.根据权利要求 5 所述的数据处理方法，还包括：

响应于收到所述监测响应，所述缓存一致性节点将所述第一数据返回给所述第二缓存，以及

在收到所述第一数据之后，所述第二缓存向所述缓存一致性节点返回指示已经收到所述第一数据的报文以结束所述读请求。

7.根据权利要求 6 所述的数据处理方法，所述响应于收到所述监测响应，所述缓存一致性节点将所述第一数据返回给所述第二缓存，包括：

响应于收到所述监测响应，所述缓存一致性节点将所述第一数据返回给所述第二缓存，并且将所述第二缓存的数据状态置为独占态，所述独占态表示仅所述第二缓存中缓存有所述第一数据。

8.根据权利要求 6 所述的数据处理方法，所述在所述缓存一致性节点将所述第一数据写入内存或将所述第一数据丢弃之后，所述缓存一致性节点向所述第一缓存返回写响应信号以直接结束对所述写请求的处理操作，包括：

在收到所述报文之后，所述缓存一致性节点丢弃所述写请求写入所述缓

存一致性节点侧的所述第一数据，并且向所述第一缓存返回所述写响应信号以直接结束对所述写请求的处理操作。

9.根据权利要求 8 所述的数据处理方法，所述在收到所述报文之后，所述缓存一致性节点丢弃所述写请求写入所述缓存一致性节点侧的所述第一数据，包括：

响应于在收到所述报文之后确定所述写请求写入所述缓存一致性节点侧的所述第一数据为无效数据，所述缓存一致性节点丢弃所述写请求写入所述缓存一致性节点侧的所述第一数据。

10. 一种电子设备，包括：

多个缓存，包括第一缓存；以及

缓存一致性节点，配置为响应于收到所述第一缓存要将第一数据写回内存的写请求，将所述第一数据写入内存或将所述第一数据丢弃，以及在所述缓存一致性节点将所述第一数据写入内存或将所述第一数据丢弃之后，向所述第一缓存返回写响应信号以直接结束对所述写请求的处理操作，其中，所述写请求携带所述第一数据，所述缓存一致性节点配置为维护由所述多个缓存中存储的数据的一致性。

11.根据权利要求 10 所述的电子设备，其中，所述缓存一致性节点还配置为：

响应于收到所述第一缓存要将所述第一数据写回所述内存的所述写请求，确定所述第一缓存是否与所述缓存一致性节点侧记录的当前独占所述第一数据的缓存相同，以及

响应于确定所述第一缓存与所述缓存一致性节点侧记录的当前独占所述第一数据的缓存相同，将所述第一数据写入所述内存。

12.根据权利要求 10 或 11 所述的电子设备，其中，所述内存配置为：在所述缓存一致性节点将所述第一数据写入内存之后，向所述缓存一致性节点返回指示所述第一数据已经被写入所述内存的写响应，并且

所述缓存一致性节点还配置为：响应于从所述内存收到指示所述第一数

据已经被写入所述内存的所述写响应，向所述第一缓存返回指示所述第一数据已经被写回所述内存的所述写响应信号以直接结束对所述写请求的处理操作。

13.根据权利要求 10-12 中的任一项所述的电子设备，其中，所述缓存一致性节点还配置为：响应于收到第二缓存对所述第一数据的读请求且所述缓存一致性节点侧记录所述第一数据在所述第一缓存中处于独占状态或共享状态，向所述第一缓存发起监测信号，并且

所述第一缓存配置为：响应于所述第一缓存在发出所述写请求之后收到所述监测信号，向所述缓存一致性节点返回携带所述第一数据的监测响应。

14. 一种数据处理装置，包括：

存储器，配置为存储计算机可执行指令；以及

处理器，配置为执行所述计算机可执行指令，

其中，所述计算机可执行指令被所述处理器执行时实现根据权利要求 1-9 中任一项所述的方法。

15. 一种非暂时性存储介质，非暂时性地存储计算机可执行指令，其中，当所述计算机可执行指令由处理器执行时，实现根据权利要求 1-9 中任一项所述的方法。

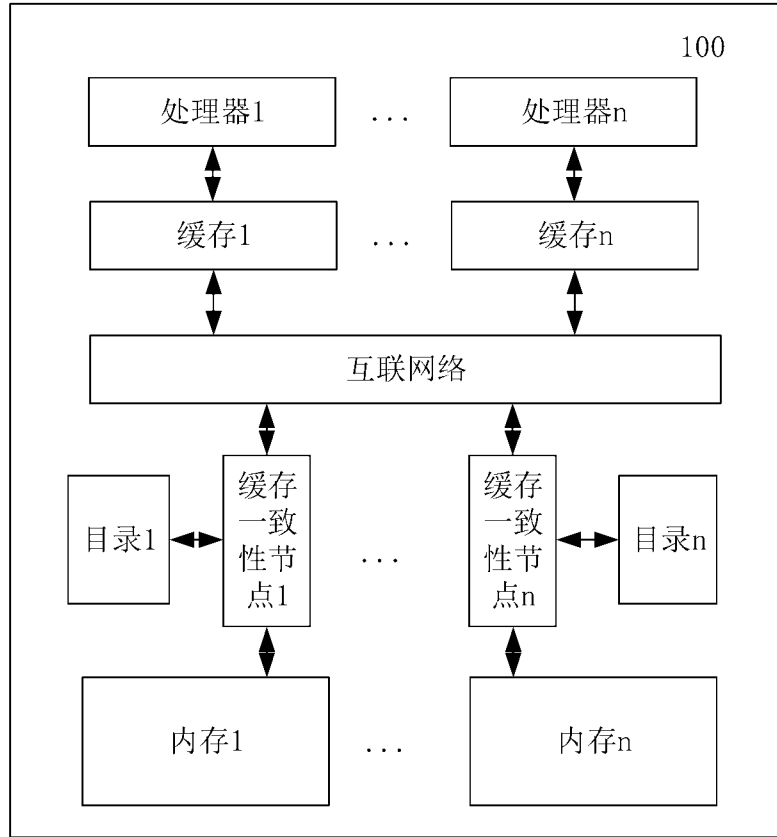


图 1

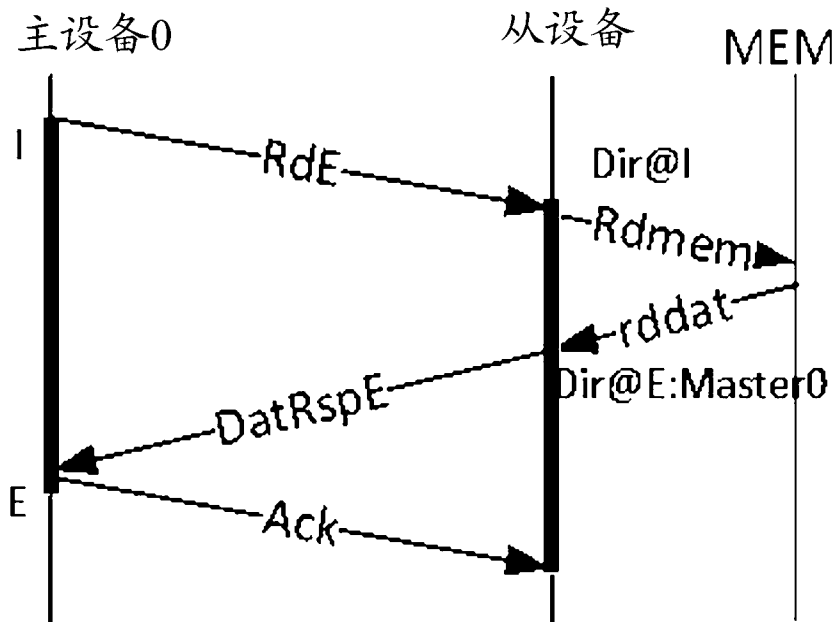


图 2A

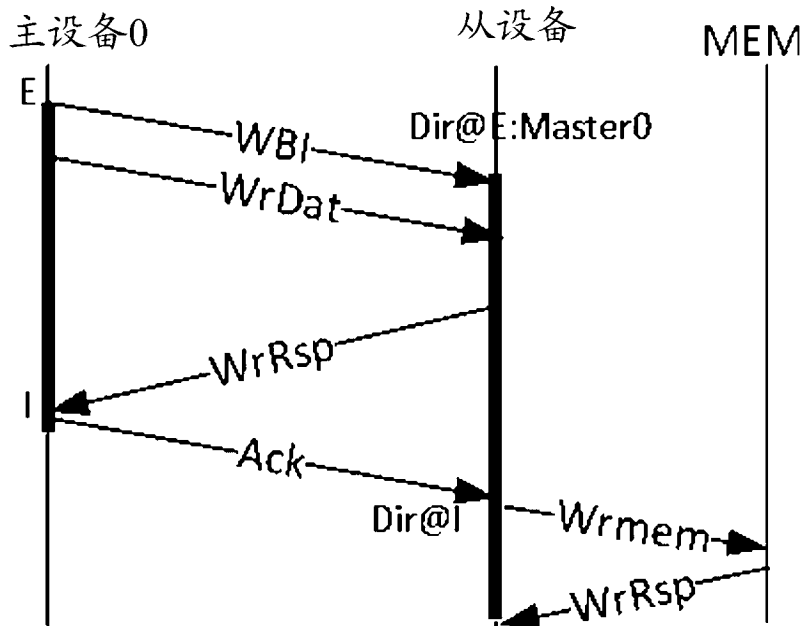


图 2B

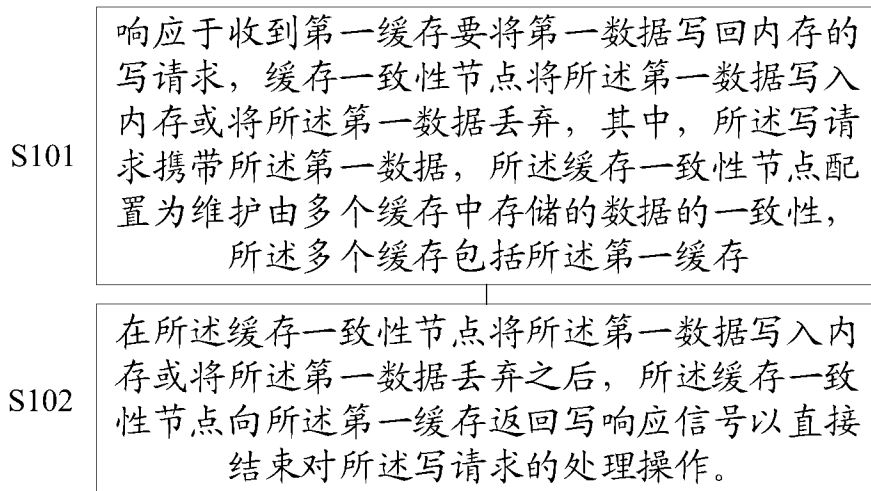


图 3A

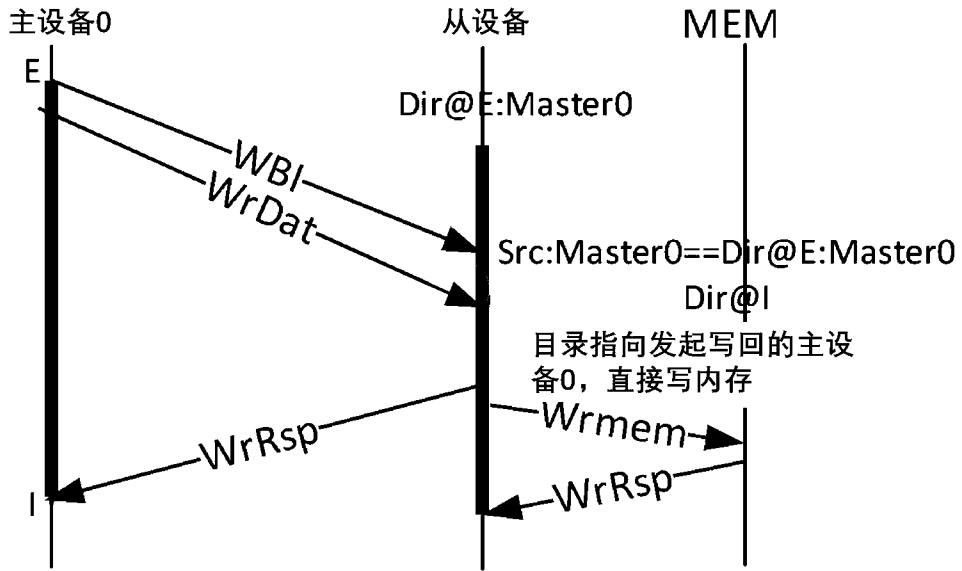


图 3B

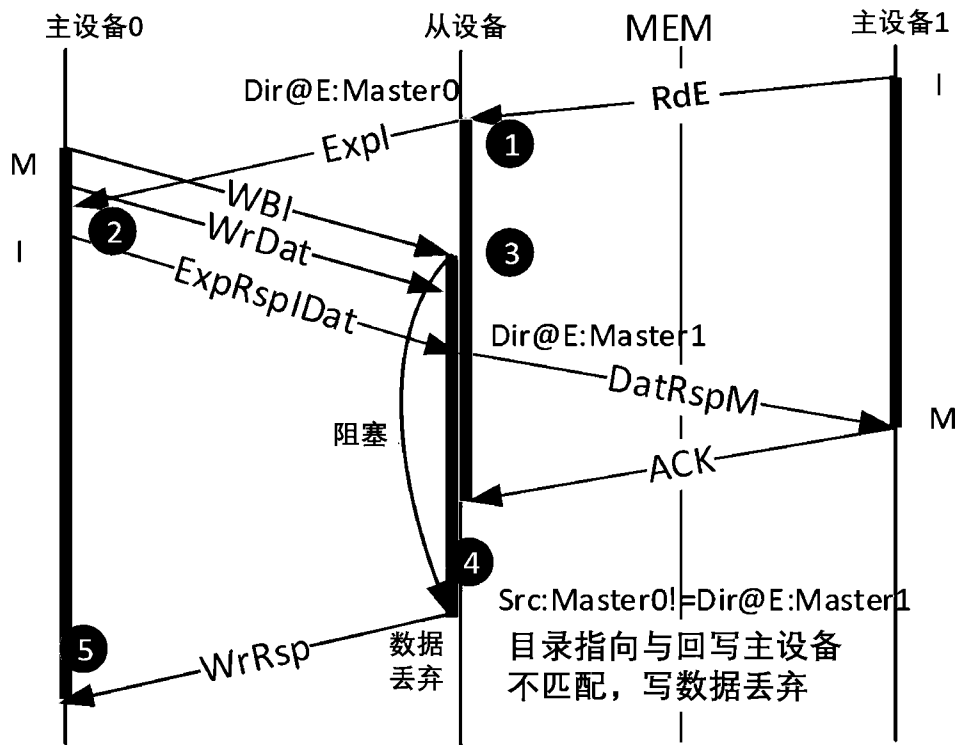


图 4



图 5

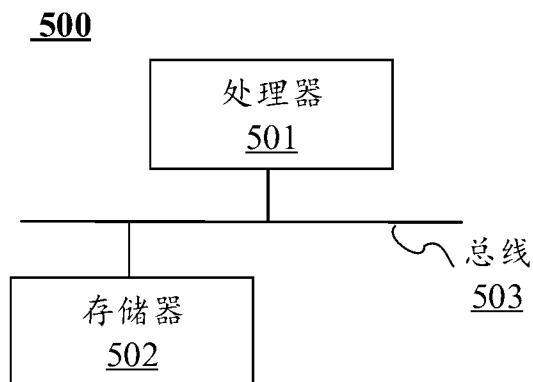


图 6



图 7

## INTERNATIONAL SEARCH REPORT

International application No.

PCT/CN2024/097159

<b>A. CLASSIFICATION OF SUBJECT MATTER</b>		
H04L67/568(2022.01)i		
According to International Patent Classification (IPC) or to both national classification and IPC		
<b>B. FIELDS SEARCHED</b>		
Minimum documentation searched (classification system followed by classification symbols)		
IPC:H04L,H04W		
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)		
CNTXT, CNABS, CNKI, ENTXT, VEN: 片上网络, 缓存, 一致, 写, 请求, 独占, 专用, E态, 丢弃, 忽略, 拒绝, 拦截, 阻止, network-on-a-chip, NOC, cache, consisten+, writ+, request, WrDat, Exclusive, E-state, discard+, ignor+, reject+, block+, intercept+		
<b>C. DOCUMENTS CONSIDERED TO BE RELEVANT</b>		
Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
PX	CN 117560422 A (HYGON INFORMATION TECHNOLOGY CO., LTD.) 13 February 2024 (2024-02-13) claims 1-15	1-15
X	CN 114116531 A (SUZHOU INSPUR INTELLIGENT TECHNOLOGY CO., LTD.) 01 March 2022 (2022-03-01) claims 1-4 and 9-10, and description, paragraphs [0020]-[0032]	1-15
A	CN 104991868 A (INSPUR (BEIJING) ELECTRONIC INFORMATION INDUSTRY CO., LTD.) 21 October 2015 (2015-10-21) entire document	1-15
A	CN 116089116 A (CHENGDU HAIGUANG INTEGRATED CIRCUIT DESIGN CO., LTD.) 09 May 2023 (2023-05-09) entire document	1-15
A	US 2018157590 A1 (ADVANCED RISC MACHINES LTD.) 07 June 2018 (2018-06-07) entire document	1-15
<input type="checkbox"/> Further documents are listed in the continuation of Box C. <input checked="" type="checkbox"/> See patent family annex.		
* Special categories of cited documents: "A" document defining the general state of the art which is not considered to be of particular relevance "D" document cited by the applicant in the international application "E" earlier application or patent but published on or after the international filing date "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) "O" document referring to an oral disclosure, use, exhibition or other means "P" document published prior to the international filing date but later than the priority date claimed "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 "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 "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 "&" document member of the same patent family		
Date of the actual completion of the international search		Date of mailing of the international search report
07 August 2024		07 September 2024
Name and mailing address of the ISA/CN		Authorized officer
China National Intellectual Property Administration (ISA/CN) China No. 6, Xitucheng Road, Jimenqiao, Haidian District, Beijing 100088		Telephone No.

**INTERNATIONAL SEARCH REPORT**  
**Information on patent family members**

International application No.

**PCT/CN2024/097159**

Patent document cited in search report			Publication date (day/month/year)	Patent family member(s)			Publication date (day/month/year)
CN	117560422	A	13 February 2024	None			
CN	114116531	A	01 March 2022	None			
CN	104991868	A	21 October 2015	None			
CN	116089116	A	09 May 2023	None			
US	2018157590	A1	07 June 2018	GB	201620464	D0	18 January 2017
				GB	2557254	A	20 June 2018
				GB	2557254	B	12 February 2020
				US	10664399	B2	26 May 2020

A. 主题的分类 H04L67/568(2022.01)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) IPC:H04L,H04W 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)) CNTXT,CNABS,CNKI,ENTXT,VEN:片上网络,缓存,一致,写,请求,独占,专用,E态,丢弃,忽略,拒绝,拦截,阻止,network-on-a-chip,NOC,cache,consisten+,writ+,request,WrDat,Exclusive,E-state,discard+,ignor+,reject+,block+,intercept+		
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 117560422 A (海光信息技术股份有限公司) 2024年2月13日 (2024 - 02 - 13) 权利要求1-15	1-15
X	CN 114116531 A (苏州浪潮智能科技有限公司) 2022年3月1日 (2022 - 03 - 01) 权利要求1-4、9-10及说明书第[0020]-[0032]段	1-15
A	CN 104991868 A (浪潮(北京)电子信息产业有限公司) 2015年10月21日 (2015 - 10 - 21) 全文	1-15
A	CN 116089116 A (成都海光集成电路设计有限公司) 2023年5月9日 (2023 - 05 - 09) 全文	1-15
A	US 2018157590 A1 (ADVANCED RISC MACH LTD) 2018年6月7日 (2018 - 06 - 07) 全文	1-15
<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 “D” 申请人在国际申请中引证的文件 “E” 在国际申请日的当天或之后公布的在先申请或专利 “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) “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 “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 “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 “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 “&” 同族专利的文件		
国际检索实际完成的日期 2024年8月7日	国际检索报告邮寄日期 2024年9月7日	
ISA/CN的名称和邮寄地址 中国国家知识产权局 中国北京市海淀区蓟门桥西土城路6号 100088	授权官员 蔡文慧 电话号码 (+86) 010-62411053	

国际检索报告  
关于同族专利的信息

国际申请号

PCT/CN2024/097159

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	117560422	A	2024年2月13日	无			
CN	114116531	A	2022年3月1日	无			
CN	104991868	A	2015年10月21日	无			
CN	116089116	A	2023年5月9日	无			
US	2018157590	A1	2018年6月7日	GB	201620464	D0	2017年1月18日
				GB	2557254	A	2018年6月20日
				GB	2557254	B	2020年2月12日
				US	10664399	B2	2020年5月26日