



(12) 发明专利申请

(10) 申请公布号 CN 104508640 A

(43) 申请公布日 2015. 04. 08

(21) 申请号 201380041056. 5

(74) 专利代理机构 北京三友知识产权代理有限公司 11127

(22) 申请日 2013. 04. 16

代理人 李辉 马建军

(30) 优先权数据

JP2012-183338 2012. 08. 22 JP

(51) Int. Cl.

G06F 12/08(2006. 01)

(85) PCT国际申请进入国家阶段日

2015. 02. 02

(86) PCT国际申请的申请数据

PCT/JP2013/061244 2013. 04. 16

(87) PCT国际申请的公布数据

W02014/030387 JA 2014. 02. 27

(71) 申请人 三菱电机株式会社

地址 日本东京都

(72) 发明人 田中沙织 贵岛淳子 内藤正博

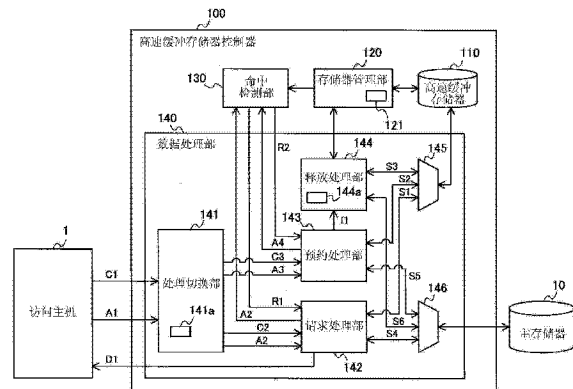
权利要求书5页 说明书24页 附图26页

(54) 发明名称

高速缓冲存储器控制器和高速缓冲存储器控制方法

(57) 摘要

高速缓冲存储器控制器(100)与主存储器(10)和访问主机(1)连接,其中,主存储器(10)具有存储第1程序的命令区域和存储由第1程序中包含的特定命令利用的数据区域,访问主机(1)执行第1程序中包含的命令,高速缓冲存储器控制器(100)具有:高速缓冲存储器(110),其存储主存储器(10)的一部分数据;以及数据处理部(140),其按照包含特定命令的起始地址的传输预约信息,在访问主机(1)执行特定命令之前,根据从访问主机(1)正在执行的命令的地址起到特定命令的起始地址为止的剩余命令步数,计算访问间隔,按照该访问间隔,将由特定命令利用的数据从主存储器(10)传输到高速缓冲存储器(110)。



1. 一种高速缓冲存储器控制器,其与主存储器和访问主机连接,其中,所述主存储器具有存储第 1 程序的命令区域和存储由该第 1 程序中包含的特定命令利用的数据的数据区域,所述访问主机执行所述第 1 程序中包含的命令,其特征在于,所述高速缓冲存储器控制器具有:

高速缓冲存储器,其存储所述主存储器的一部分数据;以及

数据处理部,其按照包含所述特定命令的起始地址的传输预约信息,在所述访问主机执行所述特定命令之前,根据从所述访问主机正在执行的命令的地址起到所述特定命令的起始地址为止的剩余命令步数,计算访问间隔,按照所述访问间隔,将由所述特定命令利用的数据从所述主存储器传输到所述高速缓冲存储器。

2. 根据权利要求 1 所述的高速缓冲存储器控制器,其特征在于,

所述第 1 程序包含具有所述特定命令的起始地址的第 1 传输预约命令,

所述访问主机在执行所述特定命令之前先执行所述第 1 传输预约命令,由此,所述数据处理部从所述访问主机取得所述传输预约信息,在所述访问主机执行所述特定命令之前,将由所述特定命令利用的数据从所述主存储器传输到所述高速缓冲存储器。

3. 根据权利要求 2 所述的高速缓冲存储器控制器,其特征在于,

每当所述访问主机正在执行的命令的地址被更新时,所述数据处理部计算所述访问间隔。

4. 根据权利要求 2 或 3 所述的高速缓冲存储器控制器,其特征在于,

所述数据处理部从所述高速缓冲存储器释放按照所述第 1 传输预约命令存储在所述高速缓冲存储器的数据中的、没有被所述访问主机访问的时间最长且所述访问主机曾访问过 1 次以上的数据。

5. 根据权利要求 2 ~ 4 中的任意一项所述的高速缓冲存储器控制器,其特征在于,

所述高速缓冲存储器控制器还具有程序转换部,该程序转换部将包含表示所述第 1 传输预约命令的代码的第 2 程序转换成所述第 1 程序。

6. 根据权利要求 2 ~ 4 中的任意一项所述的高速缓冲存储器控制器,其特征在于,

所述高速缓冲存储器控制器还具有程序转换部,该程序转换部将包含表示所述特定命令的代码的第 2 程序转换成所述第 1 程序,

所述程序转换部分析表示所述特定命令的代码,生成所述第 1 传输预约命令。

7. 根据权利要求 1 所述的高速缓冲存储器控制器,其特征在于,

所述数据处理部通过分析所述第 1 程序,生成所述传输预约信息,在所述访问主机执行所述特定命令之前,按照所述生成的传输预约信息,将由所述特定命令利用的数据从所述主存储器传输到所述高速缓冲存储器。

8. 根据权利要求 2 ~ 7 中的任意一项所述的高速缓冲存储器控制器,其特征在于,

所述数据处理部根据所述访问间隔,决定从所述访问主机接收到的多个所述传输预约信息中最优先的传输预约信息。

9. 根据权利要求 2 ~ 7 中的任意一项所述的高速缓冲存储器控制器,其特征在于,

所述数据处理部根据所述访问间隔和访问经过时间,决定从所述访问主机接收到的多个所述传输预约信息中最优先的传输预约信息,其中,所述访问经过时间是从所述访问主机向在所述主存储器中预先设定的多个连续区域的各个连续区域进行访问的经过时

间。

10. 根据权利要求 8 或 9 所述的高速缓冲存储器控制器,其特征在于,
所述数据处理部将从所述访问主机接收到的多个所述传输预约信息中的所述访问间隔最小的传输预约信息决定为最优先的传输预约信息。

11. 根据权利要求 8 或 9 所述的高速缓冲存储器控制器,其特征在于,
在对从所述访问主机接收到的多个所述传输预约信息计算出的所述访问间隔全部为同等程度的情况下,所述数据处理部将对所述访问经过时间最长的连续区域中存储的数据进行访问的传输预约信息决定为最优先的传输预约信息。

12. 根据权利要求 8 或 9 所述的高速缓冲存储器控制器,其特征在于,
在对从所述访问主机接收到的多个所述传输预约信息计算出的所述访问间隔为预先设定的第 1 阈值以下,而且在参照存储有所述要传输的数据的连续区域的命令组的起始地址中,最小的起始地址与第二小的起始地址的差分为预先设定的第 2 阈值以上的情况下,所述数据处理部将参照存储有所述要传输的数据的连续区域的命令组的起始地址最小的传输预约信息决定为最优先的传输预约信息。

13. 根据权利要求 8 或 9 所述的高速缓冲存储器控制器,其特征在于,
在对从所述访问主机接收到的多个所述传输预约信息计算出的所述访问间隔为第 1 阈值以上的情况下,所述数据处理部不决定最优先的传输预约命令而使处理等待,直到对从所述访问主机接收到的多个所述传输预约信息再次计算出的访问间隔成为所述第 1 阈值以下为止。

14. 根据权利要求 8 ~ 13 中的任意一项所述的高速缓冲存储器控制器,其特征在于,
每当完成按照已决定的所述最优先的传输预约信息传输的连续区域中的预先设定的传输单位尺寸的传输时,所述数据处理部再次决定最优先的传输预约信息。

15. 根据权利要求 8 ~ 14 中的任意一项所述的高速缓冲存储器控制器,其特征在于,
每当所述访问主机正在执行的命令的地址前进时,所述数据处理部决定所述最优先的传输预约信息。

16. 根据权利要求 8 ~ 15 中的任意一项所述的高速缓冲存储器控制器,其特征在于,
每当由已决定的所述最优先的传输预约信息表示的、连续区域中存储的全部数据的传输完成时,所述数据处理部再次决定最优先的传输预约信息。

17. 根据权利要求 2 ~ 16 中的任意一项所述的高速缓冲存储器控制器,其特征在于,
所述数据处理部根据从所述访问主机接收到的命令指令中附加的地址来判断是所述传输预约信息的情况,按照所述传输预约信息,在所述访问主机执行所述特定命令之前,将由所述特定命令利用的数据从所述主存储器传输到所述高速缓冲存储器。

18. 根据权利要求 2 ~ 17 中的任意一项所述的高速缓冲存储器控制器,其特征在于,
在所述高速缓冲存储器中存储有按照所述传输预约信息传输的连续区域的数据的情况下,所述数据处理部不从所述高速缓冲存储器释放该数据。

19. 一种高速缓冲存储器控制方法,使用高速缓冲存储器从主存储器向访问主机提供由特定命令利用的数据,其中,所述主存储器具有存储第 1 程序的命令区域和存储由该第 1 程序中包含的所述特定命令利用的数据的数据区域,所述访问主机执行所述第 1 程序中包含的命令,其特征在于,所述高速缓冲存储器控制方法具有:

传输步骤,按照包含所述特定命令的起始地址的传输预约信息,在所述访问主机执行所述特定命令之前,根据从所述访问主机正在执行的命令的地址起到所述特定命令的起始地址为止的剩余命令步数,计算访问间隔,按照所述访问间隔,将由所述特定命令利用的数据从所述主存储器传输到所述高速缓冲存储器;以及

提供步骤,在所述访问主机执行所述特定命令时,从所述高速缓冲存储器向所述访问主机提供由所述特定命令利用的数据。

20. 根据权利要求 19 所述的高速缓冲存储器控制方法,其特征在于,

所述第 1 程序包含具有所述特定命令的起始地址的第 1 传输预约命令,

所述访问主机在执行所述特定命令之前先执行所述第 1 传输预约命令,由此,在所述传输步骤中,从所述访问主机取得所述传输预约信息,在所述访问主机执行所述特定命令之前,将由所述特定命令利用的数据从所述主存储器传输到所述高速缓冲存储器。

21. 根据权利要求 20 所述的高速缓冲存储器控制方法,其特征在于,

在所述传输步骤中,每当所述访问主机正在执行的命令的地址被更新时,计算所述访问间隔。

22. 根据权利要求 20 或 21 所述的高速缓冲存储器控制方法,其特征在于,

所述高速缓冲存储器控制方法还具有释放步骤,在该释放步骤中,从所述高速缓冲存储器释放按照所述第 1 传输预约命令存储在所述高速缓冲存储器的数据中的、没有被所述访问主机访问的时间最长且所述访问主机曾访问过 1 次以上的数据。

23. 根据权利要求 20 ~ 22 中的任意一项所述的高速缓冲存储器控制方法,其特征在于,

所述高速缓冲存储器控制方法还具有程序转换步骤,在该程序转换步骤中将包含表示所述第 1 传输预约命令的代码的第 2 程序转换成所述第 1 程序。

24. 根据权利要求 20 ~ 22 中的任意一项所述的高速缓冲存储器控制方法,其特征在于,

所述高速缓冲存储器控制方法还具有程序转换步骤,在该程序转换步骤中将包含表示所述特定命令的代码的第 2 程序转换成所述第 1 程序,

在所述程序转换步骤中,分析表示所述特定命令的代码,生成所述第 1 传输预约命令。

25. 根据权利要求 19 所述的高速缓冲存储器控制方法,其特征在于,

在所述传输步骤中,通过分析所述第 1 程序,生成所述传输预约信息,在所述访问主机执行所述特定命令之前,按照所述生成的传输预约信息,将由所述特定命令利用的数据从所述主存储器传输到所述高速缓冲存储器。

26. 根据权利要求 20 ~ 25 中的任意一项所述的高速缓冲存储器控制方法,其特征在于,

在所述传输步骤中,根据所述访问间隔,决定从所述访问主机接收到的多个所述传输预约信息中最优先的传输预约信息。

27. 根据权利要求 20 ~ 25 中的任意一项所述的高速缓冲存储器控制方法,其特征在于,

在所述传输步骤中,根据所述访问间隔和访问经过时间,决定从所述访问主机接收到的多个所述传输预约信息中最优先的传输预约信息,其中,所述访问经过时间是没有从所

述访问主机向在所述主存储器中预先设定的多个连续区域的各个连续区域进行访问的经过时间。

28. 根据权利要求 26 或 27 所述的高速缓冲存储器控制方法,其特征在于,

在所述传输步骤中,将从所述访问主机接收到的多个所述传输预约信息中的所述访问间隔最小的传输预约命令决定为最优先的传输预约信息。

29. 根据权利要求 26 或 27 所述的高速缓冲存储器控制方法,其特征在于,

在所述传输步骤中,在对从所述访问主机接收到的多个所述传输预约信息计算出的所述访问间隔全部为同等程度的情况下,将对所述访问经过时间最长的连续区域中存储的数据进行访问的传输预约信息决定为最优先的传输预约信息。

30. 根据权利要求 26 或 27 所述的高速缓冲存储器控制方法,其特征在于,

在前記传输步骤中,在对从所述访问主机接收到的多个所述传输预约信息计算出的所述访问间隔为预先设定的第 1 阈值以下,而且在参照存储有所述要传输的数据的连续区域的命令组的起始地址中,最小的起始地址与第二小的起始地址的差分为预先设定的第 2 阈值以上的情况下,将参照存储有所述要传输的数据的连续区域的命令组的起始地址最小的传输预约信息决定为最优先的传输预约信息。

31. 根据权利要求 26 或 27 所述的高速缓冲存储器控制方法,其特征在于,

在所述传输步骤中,在对从所述访问主机接收到的多个所述传输预约信息计算出的所述访问间隔为第 1 阈值以上的情况下,不决定最优先的传输预约信息而使处理等待,直到对从所述访问主机接收到的多个所述传输预约信息再次计算出的访问间隔成为所述第 1 阈值以下为止。

32. 根据权利要求 26 ~ 31 中的任意一项所述的高速缓冲存储器控制方法,其特征在于,

在所述传输步骤中,每当完成按照已决定的所述最优先的传输预约信息传输的连续区域中的预先设定的传输单位尺寸的传输时,再次决定最优先的传输预约信息。

33. 根据权利要求 26 ~ 32 中的任意一项所述的高速缓冲存储器控制方法,其特征在于,

在所述传输步骤中,每当所述访问主机正在执行的命令的地址前进时,决定最优先的传输预约信息。

34. 根据权利要求 26 ~ 33 中的任意一项所述的高速缓冲存储器控制方法,其特征在于,

在所述传输步骤中,在由已决定的所述最优先的传输预约信息表示的、连续区域中存储的全部数据的传输完成时,再次决定最优先的传输预约信息。

35. 根据权利要求 20 ~ 34 中的任意一项所述的高速缓冲存储器控制方法,其特征在于,

在所述传输步骤中,根据从所述访问主机接收到的命令指令中附加的地址来判断是所述传输预约信息的情况,按照所述传输预约信息,在所述访问主机执行所述特定命令之前,将由所述特定命令利用的数据从所述主存储器传输到所述高速缓冲存储器。

36. 根据权利要求 20 ~ 35 中的任意一项所述的高速缓冲存储器控制方法,其特征在于,

在所述传输步骤中,在所述高速缓冲存储器中存储有按照所述传输预约信息传输的连续区域的数据的情况下,不从所述高速缓冲存储器释放该数据。

高速缓冲存储器控制器和高速缓冲存储器控制方法

技术领域

[0001] 本发明涉及高速缓冲存储器控制器和高速缓冲存储器控制方法。

背景技术

[0002] 近年来,由设备处理的计算机程序和视频等的数量不断增加,设备中搭载的硬盘和主存储器也大容量化。主存储器分成命令区域和数据区域。在命令区域存储有程序等命令,在数据区域存储有这些命令在处理中使用的视频等数据。主存储器的工作频率比 CPU 等访问主机的工作频率低,因此,通常使用能够高速访问的高速缓冲存储器(cache memory)。访问主机通过访问高速缓冲存储器,能够进行更高速的数据读写。

[0003] 但是,高速缓冲存储器的每单位面积的容量小且昂贵,因此,在多数情况下,难以将主存储器整体置换成高速缓冲存储器。因此,采用将主存储器的一部分数据传输到高速缓冲存储器的方法。从主存储器向高速缓冲存储器的传输是以作为高速缓冲存储器的管理单位的高速缓冲行单位进行的。在高速缓冲存储器存储有所需的数据且能够可靠地访问的情况下,访问主机能够高速地进行数据读写。将该情况称作高速缓冲命中(cache hit)。

[0004] 与此相对,将在访问主机进行访问时高速缓冲存储器没有存储被请求访问的地址的数据的情况称作高速缓冲未命中(cache miss)。在该情况下,需要将请求访问的数据从主存储器传输到高速缓冲存储器。由此,导致由于产生程序的等待时间而引起的低速化和功耗的增加。因此,期望的是,预先从主存储器预读访问主机需要的数据并将其传输到高速缓冲存储器,由此,提高能够可靠地访问数据的概率(高速缓冲命中率)。

[0005] 作为数据的预读方法,在专利文献 1 记载有如下的信息处理装置:将来自访问主机的命令存储到缓冲器,基于过去的中断命令历史预读数据并将其存储到高速缓冲存储器。由此,在再次执行访问主机过去曾执行过的中断命令的情况下成为高速缓冲命中,能够高速地向中断例程分支以及高速地从中断例程或子例程返回。

[0006] 现有技术文献

[0007] 专利文献

[0008] 专利文献 1:日本特许第 4739380 号公报

发明内容

[0009] 发明要解决的问题

[0010] 但是,在专利文献 1 所述的信息处理装置中,能够预读的数据只是具有过去曾执行过的历史的中断命令。因此,不能预读访问主机尚未执行的中断命令和成为与过去执行时不同的分支目的地的数据。因此,存在产生高速缓冲未命中的问题。

[0011] 因此,本发明的目的在于,即使对于访问主机未访问过的命令和数据,也能够可靠地实现高速缓冲命中。

[0012] 用于解决问题的手段

[0013] 本发明的一个方式的高速缓冲存储器控制器与主存储器和访问主机连接,其中,

所述主存储器具有存储第 1 程序的命令区域和存储由该第 1 程序中包含的特定命令利用的数据的数据区域,所述访问主机执行所述第 1 程序中包含的命令,其特征在于,所述高速缓冲存储器控制器具有:高速缓冲存储器,其存储所述主存储器的一部分数据;以及数据处理部,其按照包含所述特定命令的起始地址的传输预约信息,在所述访问主机执行所述特定命令之前,根据从所述访问主机正在执行的命令的地址起到所述特定命令的起始地址为止的剩余命令步数,计算访问间隔,按照所述访问间隔,将由所述特定命令利用的数据从所述主存储器传输到所述高速缓冲存储器。

[0014] 本发明的一个方式的高速缓冲存储器控制方法使用高速缓冲存储器,从主存储器向访问主机提供由特定命令利用的数据,其中,所述主存储器具有存储第 1 程序的命令区域和存储由该第 1 程序中包含的所述特定命令利用的数据的数据区域,所述访问主机执行所述第 1 程序中包含的命令,其特征在于,所述高速缓冲存储器控制方法具有:传输步骤,按照包含所述特定命令的起始地址的传输预约信息,在所述访问主机执行所述特定命令之前,根据从所述访问主机正在执行的命令的地址起到所述特定命令的起始地址为止的剩余命令步数,计算访问间隔,按照所述访问间隔,将由所述特定命令利用的数据从所述主存储器传输到所述高速缓冲存储器;以及提供步骤,在所述访问主机执行所述特定命令时,从所述高速缓冲存储器向所述访问主机提供由所述特定命令利用的数据。

[0015] 发明效果

[0016] 根据本发明的一个方式,即使对于访问主机未访问过的命令和数据,也能够可靠地实现高速缓冲命中。

附图说明

[0017] 图 1 是概略地示出实施方式 1 的高速缓冲存储器控制器的结构的框图。

[0018] 图 2 是示出用于使实施方式 1 的高速缓冲存储器控制器工作的传输预约函数的概略图。

[0019] 图 3 是示出实施方式 1 的第 2 程序应用传输预约函数的例子的概略图。

[0020] 图 4 是示出实施方式 1 的将第 2 程序编译成第 1 程序的处理的流程图。

[0021] 图 5 是示出实施方式 1 的编译器的输入与输出之间的关系概略图。

[0022] 图 6 是示出实施方式 1 的编译器编译第 2 程序而生成的第 1 程序的一例的概略图。

[0023] 图 7 是示出实施方式 1 的第 1 程序向主存储器的配置例的概略图。

[0024] 图 8 是示出实施方式 1 的数据处理部的处理切换部进行的处理的流程图。

[0025] 图 9 是示出实施方式 1 的处理切换部中的处理的时序图的一例的概略图。

[0026] 图 10 是示出实施方式 1 的数据处理部的请求处理部进行的处理的流程图。

[0027] 图 11 是示出实施方式 1 的数据处理部的预约处理部进行的处理的流程图。

[0028] 图 12 的 (a) ~ (c) 是示出实施方式 1 的预约处理部进行的数据传输的推移的概略图。

[0029] 图 13 是示出实施方式 1 的预约处理部进行的处理的时序图的一例的概略图。

[0030] 图 14 是示出实施方式 1 的数据处理部的释放处理部进行的处理的流程图。

[0031] 图 15 是示出实施方式 1 的第 2 程序的变形例的概略图。

[0032] 图 16 是概略地示出实施方式 2 的高速缓冲存储器控制器的结构的框图。

- [0033] 图 17 是实施方式 2 的预约处理部存储的表。
- [0034] 图 18 是示出实施方式 2 的访问管理信息的一例的概略图。
- [0035] 图 19 是示出实施方式 2 的数据处理部的优先级决定部进行的处理的流程图。
- [0036] 图 20 是示出实施方式 2 的数据处理部的优先级决定部决定优先级的、第 1 程序中包含的 2 个传输预约信息的第 1 例子的图。
- [0037] 图 21 是示出实施方式 2 的数据处理部的优先级决定部决定优先级的、第 1 程序中包含的 2 个传输预约信息的第 2 例子的图。
- [0038] 图 22 是示出实施方式 2 的数据处理部的优先级决定部决定优先级的、第 1 程序中包含的 2 个传输预约信息的第 3 例子的图。
- [0039] 图 23 是示出实施方式 2 的数据处理部的优先级决定部决定优先级的、第 1 程序中包含的 2 个传输预约信息的第 4 例子的图。
- [0040] 图 24 是示出实施方式 2 的数据处理部的优先级决定部决定优先级的、第 1 程序中包含的 2 个传输预约信息的第 5 例子的图。
- [0041] 图 25 是示出实施方式 2 的数据处理部的优先级决定部中的处理的时序图的一例的概略图。
- [0042] 图 26 是示出实施方式 2 的数据处理部的预约处理部进行的处理的流程图。
- [0043] 图 27 是示出实施方式 2 的数据处理部的优先级决定部中的处理的时序图的一例的概略图。
- [0044] 图 28 是概略地示出实施方式 2 的高速缓冲存储器控制器的变形例的框图。
- [0045] 图 29 是示出实施方式 2 的传输预约管理信息的变形例的概略图。
- [0046] 图 30 是示出实施方式 2 的运行地址管理信息的一例的概略图。

具体实施方式

[0047] 实施方式 1

[0048] 图 1 是概略地示出实施方式 1 的高速缓冲存储器控制器 100 的结构框图。高速缓冲存储器控制器 100 具有高速缓冲存储器 110、存储器管理部 120、命中检测部 130、数据处理部 140。

[0049] 在图 1 中,简单地示出访问主机 1、高速缓冲存储器控制器 100、主存储器 10 之间的连接关系。高速缓冲存储器控制器 100 根据来自访问主机 1 的命令指令 C1,进行对后述的高速缓冲存储器 110 或主存储器 10 中存储的数据的访问。在此,命令指令 C1 是来自访问主机 1 的对主存储器 10 上的地址的访问请求。例如,如果命令指令 C1 为读请求,则从访问主机 1 向高速缓冲存储器控制器 100 输入命令指令 C1 和表示主存储器 10 上的地址的命令地址 A1。然后,高速缓冲存储器控制器 100 将与命令指令 C1 和命令地址 A1 对应的读数据 D1 输出到访问主机 1。

[0050] 此外,图 1 示出访问主机 1 构成为 1 个的高速缓冲存储器控制器 100,但是,也可以是多个访问主机 1 共享高速缓冲存储器控制器 100。此外,访问主机 1 例如由 CPU 等控制部构成,按照主存储器 10 中存储的计算机程序(第 1 程序)的过程执行命令。

[0051] 主存储器 10 具有命令区域和数据区域。在命令区域存储有访问主机 1 执行的命令,在数据区域存储有访问主机 1 在处理中利用的数据。例如,在本实施方式中,在命令区

域存储有第 1 程序,在数据区域存储有由第 1 程序中包含的命令利用的数据。

[0052] 高速缓冲存储器 110 存储主存储器 10 中存储的一部分数据。高速缓冲存储器 110 例如由 SRAM(Static Random Access Memory :静态随机访问存储器) 那样的半导体存储器构成,能够比主存储器 10 更高速地访问数据。例如,高速缓冲存储器 110 按 64 字节进行分割,并将该分割后的单位称作高速缓冲行。而且,在高速缓冲行存储有主存储器 10 的连续 64 字节的数据。

[0053] 存储器管理部 120 进行高速缓冲存储器 110 的管理。例如,存储器管理部 120 具有作为管理信息存储部的标签存储器 121,使用该标签存储器 121 来进行高速缓冲存储器 110 的管理。

[0054] 作为管理信息,标签存储器 121 存储有 :高速缓冲存储器 110 的各高速缓冲行中存储的数据在主存储器 10 中的地址信息 Ta ;状态标记 Fs,其是表示在各高速缓冲行中是否存在数据的状态识别信息 ;以及访问标记 Fa,其是表示访问主机 1 是否访问过各高速缓冲行的访问识别信息。

[0055] 在高速缓冲存储器 110 的各高速缓冲行中存在数据的情况下,状态标记 Fs 示为“有效”,在不存在数据的情况下,状态标记 Fs 示为“无效”。

[0056] 在访问主机 1 访问过高速缓冲存储器 110 的各高速缓冲行的情况下,访问标记 Fa 示为“有效”,在访问主机 1 未访问过的情况下,访问标记 Fa 示为“无效”。存储器管理部 120 例如根据调查是否在规定的滞留时间访问过的 LRU(Least Recently Used :最近最少使用)方式,在规定的时刻,例如在定时器(未图示)计测出经过了预先设定的时间的时刻,将访问标记 Fa 复位。由此,存储器管理部 120 能够求出最近未访问过的高速缓冲行。

[0057] 命中检测部 130 判断在高速缓冲存储器 110 中是否存储有被请求访问的主存储器 10 上的地址的数据。然后,在高速缓冲存储器 110 中存储有那样的数据的情况下,命中检测部 130 将表示高速缓冲命中的命中检测结果发送给数据处理部 140,在高速缓冲存储器 110 中没有存储那样的数据的情况下,将表示高速缓冲未命中的命中检测结果发送给数据处理部 140。命中检测部 130 参照标签存储器 121 内的状态标记 Fs 示为“有效”的高速缓冲行的地址信息 Ta,判断在高速缓冲存储器 110 中是否存储有那样的数据。在存储器管理部 120 上存储有与被请求访问的地址一致的地址信息 Ta 的情况下,成为高速缓冲命中,在没有存储那样的地址信息 Ta 的情况下,成为高速缓冲未命中。

[0058] 数据处理部 140 将主存储器 10 中存储的数据传输到高速缓冲存储器 110。例如,在本实施方式中,数据处理部 140 按照包含由特定命令利用的数据存储在主存储器 10 中的地址的传输预约信息,在访问主机 1 执行第 1 程序中包含的该特定命令之前,将由该特定命令利用的数据从主存储器 10 传输到高速缓冲存储器 110。此外,数据处理部 140 根据来自访问主机 1 的请求,从高速缓冲存储器 110 或主存储器 10 读出数据。此外,数据处理部 140 根据来自访问主机 1 的请求,向高速缓冲存储器 110 或主存储器 10 写入数据。数据处理部 140 具有处理切换部 141、请求处理部 142、预约处理部 143、释放处理部 144、高速缓冲存储器访问仲裁部 145 以及主存储器访问仲裁部 146。

[0059] 处理切换部 141 分析来自访问主机 1 的命令指令 C1,将数据的输出目的地在请求处理部 142 和预约处理部 143 之间切换。例如,在命令指令 C1 表示读或写的情况下,处理切换部 141 将命令指令 C1 设为请求指令 C2 并将命令地址 A1 设为请求地址 A2 而分别发送

给请求处理部 142。此外,处理切换部 141 在命令指令 C1 表示读的情况下,将其命令地址 A1 设为运行地址 A3 而使存储器(运行地址存储部)141a 进行存储。在此,运行地址 A3 表示访问主机 1 执行中的地址。另一方面,在命令指令 C1 表示出数据传输的预约的传输预约信息的情况下,处理切换部 141 将命令指令 C1 设为传输预约指令 C3,将其与存储器 141a 中存储的运行地址 A3 一起发送给预约处理部 143。

[0060] 请求处理部 142 根据从处理切换部 141 输入的请求指令 C2 和请求地址 A2 以及从命中检测部 130 输入的命中检测结果 R1,对高速缓冲存储器 110 或主存储器 10 上的数据进行读或写。例如,请求处理部 142 在从处理切换部 141 接收到请求地址 A2 时,将该请求地址 A2 发送给命中检测部 130。然后,请求处理部 142 作为其响应从命中检测部 130 取得请求地址 A2 的命中检测结果 R1。在从处理切换部 141 输入的请求指令 C2 表示读的情况下,请求处理部 142 将从高速缓冲存储器 110 或主存储器 10 读到的读数据 D1 输出到访问主机 1。

[0061] 此外,请求处理部 142 与高速缓冲存储器访问仲裁部 145 之间的信息交换是经由信号 S1 进行的。此外,请求处理部 142 与主存储器访问仲裁部 146 之间的信息交换是经由信号 S4 进行的。

[0062] 预约处理部 143 根据从处理切换部 141 输入的运行地址 A3 和传输预约指令 C3 以及从命中检测部 130 输入的命中检测结果 R2,从主存储器 10 向高速缓冲存储器 110 传输数据。例如,预约处理部 143 在从处理切换部 141 接收到传输预约指令 C3 时,确定要进行传输的数据存储在主存储器 10 中的地址(传输预约地址),将确定的地址 A4 发送给命中检测部 130。然后,预约处理部 143 作为其响应从命中检测部 130 取得地址 A4 的命中检测结果 R2。然后,预约处理部 143 在命中检测结果 R2 为高速缓冲未命中的情况下,将该地址的数据从主存储器 10 传输到高速缓冲存储器 110。此外,预约处理部 143 在基于传输预约指令 C3 开始传输时,将表示存储目的地的高速缓冲行的预约区域信息 I1 发送给释放处理部 144。

[0063] 此外,预约处理部 143 与高速缓冲存储器访问仲裁部 145 之间的信息交换是经由信号 S2 进行的。此外,预约处理部 143 与主存储器访问仲裁部 146 之间的信息交换是经由信号 S5 进行的。

[0064] 释放处理部 144 在经由存储器管理部 120 判断为高速缓冲存储器 110 的空闲容量变少的情况下,选择要释放的高速缓冲行。例如,释放处理部 144 监视存储器管理部 120 的标签存储器 121 的状态标记 F_s ,在示为“无效”的状态标记 F_s 例如变成 T 个那样的预先设定的个数(阈值)以下的情况下,判断为高速缓冲存储器 110 的空闲容量变少。在此,释放处理部 144 根据标签存储器 121 的访问标记 F_a 和来自预约处理部 143 的预约区域信息 I1,选择要释放的高速缓冲行。释放处理部 144 在选择出要释放的高速缓冲行时,向高速缓冲存储器访问仲裁部 145 发送释放信息,该释放信息表示将选择出的高速缓冲行的数据写回到主存储器 10。

[0065] 释放处理部 144 在选择要释放的高速缓冲行时,参照来自预约处理部 143 的预约区域信息 I1。例如,释放处理部 144 根据预约区域信息 I1,监视由预约处理部 143 存储有数据的高速缓冲行的访问标记 F_a ,将访问主机 1 是否对这些高速缓冲行中存储的数据进行 1 次以上的访问而使访问标记 F_a 成为有效作为预约区域访问标记 F_{ra} ,记录到存储器(访

问历史信息存储部)144a。在访问主机 1 没有对这些高速缓冲行的数据进行过任何访问的情况下,释放处理部 144 不将该高速缓冲行作为释放对象。此外,在高速缓冲行被释放后,释放处理部 144 将该预约区域访问标记 Fra 复位。

[0066] 此外,释放处理部 144 与高速缓冲存储器访问仲裁部 145 之间的信息交换是经由信号 S3 进行的。此外,释放处理部 144 与主存储器访问仲裁部 146 之间的信息交换是经由信号 S6 进行的。

[0067] 高速缓冲存储器访问仲裁部 145 基于预先设定的优选顺序,根据从请求处理部 142、预约处理部 143、释放处理部 144 输入的信号 S1 ~ S3,对访问高速缓冲存储器 110 的顺序进行仲裁。然后,高速缓冲存储器访问仲裁部 145 按照仲裁出的顺序,将信号 S1 ~ S3 发送给高速缓冲存储器 110。例如,假定该优选顺序从高到低依次是请求处理部 142、释放处理部 144 以及预约处理部 143。

[0068] 因此,在从请求处理部 142、预约处理部 143 以及释放处理部 144 中的任意 2 个以上同时向高速缓冲存储器访问仲裁部 145 输入了信号的情况下,停止基于优选顺序较低的信号的对高速缓冲存储器 110 的访问,最优先地进行基于优选顺序较高的信号的对高速缓冲存储器 110 的访问。在基于优选顺序较高的信号的访问结束之后,开始基于同时输入的信号中优选顺序较低的信号的访问。

[0069] 主存储器访问仲裁部 146 基于预先设定的优选顺序,根据从请求处理部 142、预约处理部 143、释放处理部 144 输入的信号 S4 ~ S6,对访问主存储器 10 的顺序进行仲裁。例如,与高速缓冲存储器访问仲裁部 145 的顺序同样地,假定该优选顺序从高到低依次是请求处理部 142、释放处理部 144 以及预约处理部 143。

[0070] 因此,在从请求处理部 142、预约处理部 143 以及释放处理部 144 中的任意 2 个以上同时向主存储器访问仲裁部 146 输入了信号的情况下,停止基于优选顺序较低的信号的对主存储器 10 的访问,最优先地进行基于优选顺序较高的信号的对主存储器 10 的访问。在基于优选顺序较高的信号的访问结束之后,开始基于同时输入的信号中优选顺序较低的信号的访问。

[0071] 此时,在从请求处理部 142 和预约处理部 143 同时对相同地址的数据请求访问的情况下,在高速缓冲存储器 110 上没有存储该地址的数据时,主存储器访问仲裁部 146 仅将来自请求处理部 142 的访问请求设为有效。主存储器访问仲裁部 146 向预约处理部 143 输出表示该地址的数据的传输已完成的预约传输完成标记 Ftf。

[0072] 图 2 是示出用于使高速缓冲存储器控制器 100 进行工作的传输预约函数 160 的概略图。传输预约函数 160 是表示将特定的数据从主存储器 10 传输到高速缓冲存储器 110 的传输预约命令的代码。如图 2 所示,传输预约函数 160 定义在执行该函数后访问主机 1 参照的连续区域的起始地址 MM_ADDR、该连续区域的尺寸 H*V、以及作为参照该连续区域的命令组的函数的起始地址 PROC。连续区域是主存储器 10 上的命令区域或数据区域的地址连续的一部分区域。此外,在本实施方式中,起始地址 MM_ADDR、尺寸 H*V 以及起始地址 PROC 是传输预约信息。

[0073] 访问主机 1 执行以例如汇编语言等访问主机 1 能够执行的形式记述的第 1 程序。另一方面,图 2 所示的传输预约函数 160 包含于以例如 C 语言这样的高级语言等访问主机 1 不能执行的形式记述的第 2 程序。第 1 程序是通过编译第 2 程序而生成的。

[0074] 图 3 是示出第 2 程序应用图 2 所示的传输预约函数 160 的例子的概略图。图 3 所示的第 2 程序 170 包含：函数 173，其由访问主机 1 参照配置在主存储器 10 上的连续区域中的数据中的数据的命令构成；以及函数 174，其由参照配置在别的连续区域中的数据中的数据的命令构成。在程序 170 中，在执行函数 173、174 之前，对传输预约函数 171 和传输预约函数 172 进行处理。

[0075] 第 2 程序被记述成在访问主机 1 进行参照连续区域的数据的处理之前，对传输预约函数 160 进行处理，因此，能够以访问主机 1 可执行的形式，在第 1 程序中插入用于将希望可靠地实现高速缓冲命中的连续区域的数据从主存储器 10 传输到高速缓冲存储器 110 的指令。

[0076] 图 4 是示出将图 3 所示的第 2 程序 170 编译（转换）成第 1 程序的处理的流程图。如图 5 所示，图 4 所示的流程是通过将第 2 程序输入到编译器 Cgc 而执行的。编译器 Cgc 是具有按照通常使用的规格对图 2 所示的传输预约函数 160 进行编译的功能的程序转换部。此外，假定第 2 程序 170 存储在未图示的存储部。

[0077] 编译器 Cgc 在开始编译时，判定第 2 程序中的作为编译对象的源代码是否是图 2 所示的传输预约函数 160（步骤 S10）。然后，在源代码是传输预约函数 160 的情况下（S10：是），处理进入到步骤 S11，在源代码不是传输预约函数 160 的情况下（S10：否），处理进入到步骤 S12。

[0078] 在步骤 S11 中，编译器 Cgc 将源代码编译成与第 1 程序对应的指令。另一方面，在步骤 S12 中，编译器 Cgc 按照通常使用的规格对源代码进行编译。

[0079] 然后，编译器 Cgc 判定作为编译对象的源代码是否是第 2 程序的末尾（S13）。在源代码不是末尾的情况下（S13：否），编译器 Cgc 将编译对象更新成下一个源代码，处理进入到步骤 S10。在源代码是末尾的情况下（S13：是），编译器 Cgc 结束流程。

[0080] 图 6 是示出编译器 Cgc 编译图 3 所示的第 2 程序 170 而得到的第 1 程序的一例的概略图。在图 6 所示的第 1 程序 180 中，左列以 16 进制数表记地址，右列以 16 进制数表记命令语句。命令的形式取决于编译器 Cgc，因此，在此省略与命令语句形式相关的说明，对各命令表示的动作进行说明。

[0081] 图 6 所示的命令 181a ~ 181c 是基于图 3 所示的传输预约函数 171 而生成的传输预约命令。命令 181a 是将图 3 所示的传输预约函数 171 中记述的连续区域的起始地址 MM_ADDR1 通知给处理切换部 141 的命令。命令 181b 是将图 3 所示的传输预约函数 171 中记述的连续区域的尺寸 H1*V1 通知给处理切换部 141 的命令。命令 181c 是将图 3 所示的传输预约函数 171 中记述的起始地址 PROC1 通知给处理切换部 141 的命令。此外，命令 182a ~ 182c 是基于图 3 中的传输预约函数 172 而生成的传输预约命令。命令 182a 是将图 3 所示的传输预约函数 172 中记述的连续区域的起始地址 MM_ADDR2 通知给处理切换部 141 的命令。命令 182b 是将图 3 所示的传输预约函数 172 中记述的连续区域的尺寸 H2*V2 通知给处理切换部 141 的命令。命令 182c 是将图 3 所示的传输预约函数 172 中记述的起始地址 PROC2 通知给处理切换部 141 的命令。

[0082] 如图 6 所示，在第 1 程序 180 中，传输预约命令 181a ~ 181c、182a ~ 182c 被记载在执行利用主存储器 10 的数据进行的处理 P1、P2 的命令组 1831、1832、1833、•••、1841、1842、1843、••• 之前，因此，访问主机 1 在执行这些命令组之前先执行传输预约命

令 181a ~ 181c、182a ~ 182c。因此,访问主机 1 能够在执行这些命令组之前,按照传输预约命令 181a ~ 181c、182a ~ 182c,将表示传输预约信息的命令指令 C1 发送给处理切换部 141。由此,在访问主机 1 执行这些命令组之前,将由这些命令组利用的数据从主存储器 10 传输到高速缓冲存储器 110。

[0083] 图 7 是示出图 6 所示的第 1 程序 180 向主存储器 10 的配置例的概略图。在图 7 中,将图 6 所示的第 1 程序 180 存储在主存储器 10 上的命令区域 190 中。特别地,将编译图 3 所示的传输预约函数 171 而得到的命令 181a ~ 181c 配置在区域 191p 中,将编译图 3 所示的传输预约函数 172 而得到的命令 182a ~ 182c 配置在区域 192p 中。

[0084] 访问主机 1 执行配置在区域 191p 中的命令 181a(参照图 6),由此,将在图 3 所示的函数 173 中参照的连续区域 197d 的起始地址通知给预约处理部 143。接下来,访问主机 1 执行配置在区域 191p 中的命令 181b(参照图 6),由此,将连续区域 197d 的尺寸通知给预约处理部 143。此外,访问主机 1 执行配置在区域 191p 中的命令 181c(参照图 6),由此,将存储有与图 3 所示的函数 173 对应的指令的区域 193p 的起始地址通知给预约处理部 143。

[0085] 同样,访问主机 1 执行配置在区域 192p 中的命令 182a ~ 182c(参照图 6),由此,将在图 3 所示的函数 174 中参照的连续区域 198d 的起始地址和尺寸以及存储有与图 3 所示的函数 174 对应的指令的区域 194p 的起始地址通知给预约处理部 143。

[0086] 接下来,使用流程图,对高速缓冲存储器控制器 100 中的处理的流程进行说明。

[0087] 图 8 是示出数据处理部 140 的处理切换部 141 进行的处理的流程图。处理切换部 141 在从访问主机 1 输入命令指令 C1 和命令地址 A1 时开始处理。

[0088] 首先,处理切换部 141 判定输入的命令地址 A1 是否是图 7 所示的命令区域 190 中包含的地址 (S20)。在命令地址 A1 是命令区域 190 的地址的情况下 (S20:是),处理进入到步骤 S21,在命令地址 A1 不是命令区域 190 的地址的情况下 (S20:否),处理进入到步骤 S22。

[0089] 在步骤 S21 中,处理切换部 141 将输入的命令地址 A1 作为运行地址 A3 存储到存储器 141a。此外,在已经存储有运行地址 A3 的情况下,处理切换部 141 更新其值。另一方面,在命令地址 A1 不是命令区域 190 的地址的情况下,不更新运行地址 A3。

[0090] 在步骤 S22 中,处理切换部 141 将存储器 141a 中存储的运行地址 A3 发送给预约处理部 143。

[0091] 接下来,处理切换部 141 判定从访问主机 1 输入的命令指令 C1 是否是读和写中的任意一个 (S23)。在命令指令 C1 是读和写中的任意一个的情况下 (S23:是),处理进入到步骤 S24,在命令指令 C1 既不是读也不是写的情况下 (S23:否),处理进入到步骤 S25。

[0092] 在步骤 S24 中,处理切换部 141 将输入的命令指令 C1 和命令地址 A1 分别作为请求指令 C2 和请求地址 A2 发送给请求处理部 142。另一方面,在步骤 S25 中,处理切换部 141 将命令指令 C1 作为传输预约指令 C3 发送给预约处理部 143。传输预约指令 C3 例如是图 6 所示的命令 181a ~ 181c 和命令 182a ~ 182c。

[0093] 接下来,处理切换部 141 判定在数据处理部 140 是否存在来自访问主机 1 的命令指令 C1(是否存在未处理的命令指令 C1) (S26)。在存在未处理的命令指令 C1 的情况下 (S26:是),处理进入到步骤 S20,在不存在未处理的命令指令 C1 的情况下 (S26:否),处理切换部 141 结束流程。

[0094] 图 9 是示出处理切换部 141 中的处理的时序图的一例的概略图。在图 9 中, 示出从访问主机 1 向处理切换部 141 输入命令指令 C1 的时序、向请求处理部 142 输出数据的时序以及向预约处理部 143 输出数据的时序。

[0095] 处理切换部 141 在时刻 t10 从访问主机 1 输入表示读请求的命令指令 C1 和表示程序区域的命令地址 A1 时, 在时刻 t11, 将它们分别作为请求指令 C2 和请求地址 A2 输出到请求处理部 142。此外, 处理切换部 141 将运行地址 A3 的值更新成表示命令地址 A2 的值并发送给预约处理部 143。

[0096] 接下来, 处理切换部 141 在时刻 t12 从访问主机 1 输入表示写请求的命令指令 C1 和表示数据区域的命令地址 A1 时, 在时刻 t13, 将它们分别作为请求指令 C2 和请求地址 A2 发送给请求处理部 142。此外, 由于命令地址 A1 表示数据区域, 因此, 处理切换部 141 不更新运行地址 A3 的值, 而将存储器 141a 中记录的运行地址 A3 发送给预约处理部 143。

[0097] 处理切换部 141 在时刻 t13 从访问主机 1 输入表示数据的传输预约的命令指令 C1 (TRI) 和表示传输预约的命令地址 A1 (TRIA) 时, 不进行向请求处理部 142 的输出, 在时刻 t14, 将存储器 141a 中记录的运行地址 A3 发送给预约处理部 143, 并且, 将命令指令 C1 作为传输预约指令 C3 发送给预约处理部 143。

[0098] 在时刻 t14 从访问主机 1 输入表示读请求的命令指令 C1 和表示程序区域的命令地址 A1 时, 在时刻 t15, 将它们分别作为请求指令 C2 和请求地址 A2 发送给请求处理部 142。此外, 处理切换部 141 将运行地址 A3 的值更新成由命令地址 A1 表示的值发送给预约处理部 143。

[0099] 如上所述, 处理切换部 141 根据从访问主机 1 输入的命令指令 C1 和命令地址 A1, 将输出目的地在请求处理部 142 或预约处理部 143 之间切换。

[0100] 图 10 是示出数据处理部 140 的请求处理部 142 进行的处理的流程图。请求处理部 142 在被输入来自处理切换部 141 的请求指令 C2 和请求地址 A2 以及来自命中检测部 130 的表示高速缓冲命中或高速缓冲未命中的命中检测结果 R1 时, 开始处理。

[0101] 首先, 请求处理部 142 根据来自命中检测部 130 的命中检测结果 R1, 判定高速缓冲存储器 110 中是否存在由请求指令 C2 请求的数据 (S30)。在高速缓冲存储器 110 中不存在由请求指令 C2 请求的数据的情况下 (S30: 否), 换言之, 在命中结果 R1 表示高速缓冲未命中的情况下, 处理进入到步骤 S31。另一方面, 在高速缓冲存储器 110 中存在由请求指令 C2 请求的数据的情况下 (S30: 是), 换言之, 在命中结果 R1 表示高速缓冲命中的情况下, 处理进入到步骤 S32。

[0102] 在步骤 S31 中, 请求处理部 142 为了从主存储器 10 向高速缓冲存储器 110 传输数据, 将进行数据传输的命令发送给主存储器访问仲裁部 146。假定该命令包含请求地址 A2。然后, 主存储器访问仲裁部 146 进行将由请求地址 A2 表示的地址中存储的数据从主存储器 10 传输到高速缓冲存储器 110 的处理。例如, 主存储器访问仲裁部 146 从主存储器 10 读出由请求地址 A2 表示的地址中存储的数据, 将该数据发送给请求处理部 142。请求处理部 142 将发送来的数据发送给高速缓冲存储器访问仲裁部 145, 使其写入到高速缓冲存储器 110。此外, 在向高速缓冲存储器 110 写入数据的情况下, 存储器管理部 120 将写入到标签存储器 121 中的数据的地址作为地址信息 Ta 存储, 并且, 将对应的状态标记 Fs 更新成表示存在数据。

[0103] 在步骤 S32 中,请求处理部 142 进行向高速缓冲存储器 110 的访问。在请求指令 C2 表示写命令的情况下,为了将请求指令 C2 表示的数据写到高速缓冲行,请求处理部 142 向高速缓冲存储器访问仲裁部 145 发送请求指令 C2 和请求地址 A2。另一方面,在请求指令 C2 表示读命令的情况下,从高速缓冲行读取请求指令 C2 表示的数据,请求处理部 142 向高速缓冲存储器访问仲裁部 145 发送作为读对象的请求地址 A2。将这样取得的数据作为读数据 D1 发送给访问主机 1。此外,在对高速缓冲存储器 110 进行了访问的情况下,存储器管理部 120 将标签存储器 121 中存储的访问标记 Fa 更新成表示已访问。

[0104] 在此,在请求指令 C2 表示读请求且命中检测结果 R1 表示高速缓冲未命中的情况下,请求处理部 142 可以将经由主存储器访问仲裁部 146 从主存储器 10 读出的数据,经由高速缓冲存储器访问仲裁部 145 传输到高速缓冲存储器 110,并且,作为读数据输出到访问主机 1。在该情况下,不进行图 10 的步骤 S32 的处理。但是,存储器管理部 120 将标签存储器 121 中存储的访问标记 Fa 更新成表示已访问。

[0105] 图 11 是示出数据处理部 140 的预约处理部 143 进行的处理的流程图。预约处理部 143 在从处理切换部 141 输入运行地址 A3、存储有要传输的数据的连续区域的起始地址 MM_ADDR、存储有要传输的数据的连续区域的尺寸 H*V 以及参照存储有要传输的数据的连续区域的命令组的起始地址 PROC,且从命中检测部 130 输入表示高速缓冲命中或高速缓冲未命中的命中检测结果 R2 时,开始处理。

[0106] 首先,预约处理部 143 计算访问主存储器 10 的间隔(以下,称作访问间隔 Da)(S40)。在访问间隔 Da 的计算中,使用从运行地址 A3 到参照存储有要传输的数据的连续区域的命令组的起始地址 PROC 的命令步数 Ds。命令步数 Ds 可以按照下述(1)式所示,根据从作为开始预约处理时刻的命令地址的运行地址 A3 到参照存储有要传输的数据的连续区域的命令组的起始地址 PROC 的差分来计算。

$$[0107] \quad Ds = (\text{命令组的起始地址 PROC}) - (\text{运行地址 A3}) \quad (1)$$

[0108] 然后,预约处理部 143 可以按照下述(2)式所示,将命令步数 Ds 除以从处理切换部 141 输入的存储有要传输的数据的连续区域的尺寸 H*V 中的未完成传输的剩余的连续区域的尺寸 Rs,由此计算每单位尺寸的传输所需的命令步数 Dspu。预约处理部 143 将该每单位尺寸的传输所需的命令步数 Dspu 作为访问间隔 Da 使用。

$$[0109] \quad Dspu = (\text{命令步数 Ds}) \div (\text{剩余的连续区域的尺寸 Rs}) \quad (2)$$

[0110] 接下来,预约处理部 143 判定在高速缓冲存储器 110 中是否存在要传输的连续区域的数据中作为接下来要传输的对象的数据(S41)。在此,作为接下来要传输的对象的数据是要传输的连续区域的数据中接下来要传输的传输单位长的数据。然后,在高速缓冲存储器 110 中不存在作为接下来要传输的对象的数据的情况下(S41:否),换言之,在作为接下来要传输的对象的数据的命中检测结果 R2 表示高速缓冲未命中的情况下,需要从主存储器 10 向高速缓冲存储器 110 传输该数据,因此,预约处理部 143 使处理进入到步骤 S42。另一方面,在高速缓冲存储器 110 中存在作为接下来要传输的对象的数据的情况下(S41:是),换言之,在作为接下来要传输的对象的数据的命中检测结果 R2 表示高速缓冲命中的情况下,不需要从主存储器 10 向高速缓冲存储器 110 传输该数据,因此,预约处理部 143 不传输该数据而使处理进入到步骤 S43。

[0111] 在步骤 S42 中,预约处理部 143 以在步骤 S40 中计算出的访问间隔 Da,将从主存储

器 10 向高速缓冲存储器 110 进行数据传输的命令发送给主存储器访问仲裁部 146。例如，在访问间隔 D_a 为“8”的情况下，需要在运行地址 A_3 前进 8 步之前传输传输单位长的数据。因此，预约处理部 143 在该 8 步中的任意一个例如该 8 步中最初的步，将传输命令发送给主存储器访问仲裁部 146。接收到这样的传输命令的主存储器访问仲裁部 146 从主存储器 10 读出作为传输对象的数据并发送给预约处理部 143。预约处理部 143 将发送来的数据发送给高速缓冲存储器访问仲裁部 145，使其写入到高速缓冲存储器 110。此外，在向高速缓冲存储器 110 写入数据的情况下，存储器管理部 120 将写入到标签存储器 121 中的数据的地址作为地址信息 T_a 存储，并且，将对应的状态标记 F_s 更新成表示存在数据。

[0112] 接下来，预约处理部 143 更新传输完成尺寸，该传输完成尺寸表示已经传输的数据的合计尺寸 (S_{43})。

[0113] 接下来，预约处理部 143 判定传输完成尺寸是否为从处理切换部 141 输入的存储有要传输的数据的连续区域的尺寸 $H*V$ 以上 (S_{44})。在传输完成尺寸小于存储有要传输的数据的连续区域的尺寸 $H*V$ 的情况下 (S_{44} :否)，处理进入到步骤 S_{45} ，在为存储有要传输的数据的连续区域的尺寸 $H*V$ 以上的情况下 (S_{44} :是)，结束流程。

[0114] 在步骤 S_{45} 中，预约处理部 143 判断是否已从处理切换部 141 取得更新后的运行地址 A_3 。在已取得运行地址 A_3 的情况下 (S_{45} :是)，处理进入到步骤 S_{40} ，在未取得运行地址 A_3 的情况下 (S_{45} :否)，处理进入到步骤 S_{41} 。

[0115] 如上所述，预约处理部 143 基于传输预约指令 C_3 ，将所需的数据传输到高速缓冲存储器 110，由此，即使针对访问主机 1 未访问过的主存储器 10 上的区域，也能够可靠地提高高速缓冲命中率。

[0116] 图 12 的 (a) ~ (c) 是示出预约处理部 143 进行的数据传输的推移的概略图。图 12 的 (a) 示出运行地址 A_3 的推移，图 12 的 (b) 示出存储有要传输的数据的连续区域的剩余尺寸的推移，图 12 的 (c) 示出访问间隔 D_a 的推移。

[0117] 如图 12 的 (c) 所示，时刻 t_0 的访问间隔 D_{a0} 是根据向预约处理部 143 输入传输预约指令 C_3 的时刻而计算出的。

[0118] 如图 12 的 (a) 所示，在时刻 t_1 ，在运行地址 A_3 的值被更新时，根据存储有要在该时刻传输的数据的连续区域的剩余尺寸（参照图 12 的 (b)）和从运行地址 A_3 到参照连续区域的命令组的起始地址 PROC 的剩余步数（参照图 12 的 (a)），预约处理部 143 再次计算访问间隔 D_{a1} （参照图 12 的 (c)），调节访问主存储器 10 的间隔。

[0119] 这样，预约处理部 143 在运行地址 A_3 被更新时，再次计算访问间隔 D_a ，以该访问间隔 D_a 进行连续区域的传输。然后，进行传输处理，使得在运行地址 A_3 成为参照连续区域的命令组的起始地址 PROC 的时刻 t_n 之前，存储有要传输的数据的连续区域的剩余尺寸为“0”即传输完成。在时刻 t_n ，运行地址 A_3 成为参照连续区域的命令组的起始地址 PROC，因而访问间隔 D_{an} 成为“0”。

[0120] 图 13 是示出预约处理部 143 进行的处理的时序图的一例的概略图。图 13 示出从处理切换部 141 输入运行地址 A_3 和存储有要传输的数据的连续区域的尺寸 $H*V$ 的时序、切换存储有要传输的数据的连续区域中剩余的连续区域的尺寸的时序、计算访问间隔 D_a 的时序以及访问主存储器 10 的时序。

[0121] 在时刻 t_0 ，在从处理切换部 141 输入运行地址 A_3 和传输预约指令 C_3 时，预约处

理部 143 计算访问间隔 Da0。预约处理部 143 以计算出的访问间隔 Da0 进行对主存储器 10 的访问,直到来自处理切换部 141 的运行地址 A3 被更新的时刻 t1 为止。

[0122] 在时刻 t1,在从处理切换部 141 输入的运行地址 A3 被更新时,预约处理部 143 计算访问间隔 Da1,以计算出的访问间隔 Da1 进行对主存储器 10 的访问,直到运行地址 A3 被更新的时刻 t2 为止。

[0123] 同样,在时刻 t2,在从处理切换部 141 输入的运行地址 A3 被更新时,预约处理部 143 计算访问间隔 Da2,以计算出的访问间隔 Da2 进行对主存储器 10 的访问,直到运行地址 A3 被更新为止。

[0124] 以后,预约处理部 143 每当从处理切换部 141 输入的运行地址 A3 被更新时,以计算出的访问间隔 Da 访问主存储器 10,在运行地址 A3 成为由传输预约指令 C3 通知的参照连续区域的命令组的起始地址 PROC 之前完成传输。

[0125] 如上所述,参照运行地址 A3 调节访问主存储器 10 的间隔,在执行参照连续区域的命令组之前完成传输,由此,能够高效地使用高速缓冲存储器 110。

[0126] 图 14 是示出数据处理部 140 的释放处理部 144 进行的处理的流程图。释放处理部 144 始终监视在存储器管理部 120 的标签存储器 121 中,表示无效(没有存储数据)的状态标记 Fs 是否为预先设定的个数例如 T 个以下。

[0127] 释放处理部 144 判定状态标记 Fs 是否为 T 个以下(S50)。在状态标记 Fs 多于 T 个的情况下(S50:否),释放处理部 144 进行等待,继续监视存储器管理部 120 的状态标记 Fs。另一方面,在状态标记 Fs 为 T 个以下的情况下(S50:是),处理进入到步骤 S51。

[0128] 在步骤 S51 中,释放处理部 144 选择高速缓冲存储器 110 上的各高速缓冲行中作为释放候选的高速缓冲行。选择作为释放候选的高速缓冲行的方法例如应用选择未被参照的时间最长的高速缓冲行的 LRU 方式。

[0129] 然后,释放处理部 144 判定在步骤 S51 中选择出的高速缓冲行是否是释放对象(S52)。在高速缓冲存储器 110 的各高速缓冲行中,存储有由传输预约指令 C3 传输的数据且从未被访问主机 1 访问的高速缓冲行不是释放对象。换言之,成为释放对象的是,作为释放候选选择出的高速缓冲行中通过请求指令 C2 传输到高速缓冲存储器 110 的高速缓冲行、以及作为释放候选选择出的高速缓冲行中由预约区域访问标记 Fra 表示访问标记 Fa 有效的高速缓冲行。因此,释放处理部 144 根据来自预约处理部 143 的预约区域信息 I1,监视预约处理部 143 存储数据的高速缓冲行的访问标记 Fa,将访问标记 Fa 是否 1 次以上成为有效作为预约区域访问标记 Fra 记录到存储器 144a。

[0130] 然后,在步骤 S51 中选择出的高速缓冲行不是释放对象的情况下(S52:否),处理进入到步骤 S51,释放处理部 144 再次根据 LRU 方式选择作为释放候选的高速缓冲行。另一方面,在步骤 S51 中选择出的高速缓冲行是释放对象的情况下(S52:是),处理进入到步骤 S53。

[0131] 在步骤 S53 中,释放处理部 144 向高速缓冲存储器访问仲裁部 145 发送命令,该命令将作为释放对象的高速缓冲行中存储的数据写回到主存储器 10(S53)。接收到这样的命令的高速缓冲存储器访问仲裁部 145 读出作为释放对象的高速缓冲行中存储的数据,将该数据发送给释放处理部 144。释放处理部 144 将该数据发送给主存储器访问仲裁部 146,使其将该数据写入主存储器 10。

[0132] 在高速缓冲行从高速缓冲存储器 110 向主存储器 10 的写回完成时,处理进入到步骤 S50,释放处理部 144 继续监视标签存储器 121 的状态标记 Fs。

[0133] 如上所述,不将存储有通过传输预约指令 C3 传输的数据且访问主机 1 未访问过的高速缓冲行作为释放对象,由此,将访问主机参照的数据可靠地存储到高速缓冲存储器 110。

[0134] 在以上记载的实施方式 1 中说明了如下例子:在图 3 所示的第 2 程序 170 记述传输预约函数 171、172,使用由编译器 Cgc 编译这些传输预约函数 171、172 而生成的第 1 程序 180(参照图 6),使高速缓冲存储器控制器 100 进行工作,但是,不限于这样的例子。在第 1 程序中记述有传输预约指令即可。因此,也可以使用从图 15 所示的第 2 程序 270 生成具有图 6 所示的传输预约指令 181a ~ 181c、182a ~ 182c 的第 1 程序 180 的编译器。换言之,只要编译器通过分析第 2 程序中包含的表示利用主存储器 10 的数据区域中存储的数据的命令的代码,生成传输预约命令即可。由此,即使在第 2 程序 270 没有有意识地记述传输预约函数 160,也能够适当地可靠地记述传输预约指令,能够达到期望的目的。

[0135] 此外,在以上记载的实施方式 1 中,对从访问主机 1 发出传输预约指令的例子进行了说明,但是,不限于这样的例子。例如,也可以在访问主机 1 执行的第 1 程序中不包含传输预约命令。在这样的情况下,只要预约处理部 143 分析高速缓冲存储器 110 内存储的主存储器 10 上的规定的程序区域中存储的第 1 程序,生成将在之后执行的处理中参照的数据从主存储器 10 传输到高速缓冲存储器 110 的传输预约信息即可。由此,能够使用通用的编译器从第 2 程序生成第 1 程序,可靠地将访问主机 1 参照的数据存储到高速缓冲存储器 110。

[0136] 实施方式 2

[0137] 根据图 16 ~ 图 30,对实施方式 2 进行说明。

[0138] 图 16 是概略地示出实施方式 2 的高速缓冲存储器控制器 200 的结构的框图。高速缓冲存储器控制器 200 具有高速缓冲存储器 110、存储器管理部 120、命中检测部 130 以及数据处理部 240。

[0139] 在图 16 中,简单地示出访问主机 1、高速缓冲存储器控制器 200 以及主存储器 10 的连接关系。

[0140] 主存储器 10 按照被称作段 (bank) 的一定容量的容量而汇总地管理。段内被分成命令区域和数据区域。此外,针对主存储器 10,通过指定行 (Row) 地址和列 (Column) 地址,能够对特定的连续区域进行访问。

[0141] 高速缓冲存储器 110、存储器管理部 120 以及命中检测部 130 的功能与实施方式 1 相同,由于已经说明,因此在此省略说明。

[0142] 数据处理部 240 将主存储器 10 中存储的数据传输到高速缓冲存储器 110。在本实施方式中,每当接收到包含由特定命令利用的数据存储在主存储器 10 中的地址的传输预约信息时,数据处理部 240 存储接收到的传输预约信息。然后,数据处理部 240 在存储有多个传输预约信息的情况下,对它们进行仲裁。数据处理部 240 按照通过仲裁判定为优先级高的传输预约信息,在访问主机 1 执行第 1 程序中包含的特定命令之前,将由该特定命令利用的数据从主存储器 10 传输到高速缓冲存储器 110。此外,数据处理部 240 根据来自访问主机 1 的请求,将数据写入高速缓冲存储器 110 或主存储器 10。数据处理部 240 具有处理切换部 141、请求处理部 142、预约处理部 243、释放处理部 144、高速缓冲存储器访问仲裁部

145、主存储器访问仲裁部 146、优先级决定部 247 以及访问管理部 248。

[0143] 处理切换部 141、请求处理部 142、释放处理部 144、高速缓冲存储器访问仲裁部 145 以及主存储器访问仲裁部 146 的功能与实施方式 1 相同,由于已经说明,因而在此省略说明。

[0144] 优先级决定部 247 从处理切换部 141 接收运行地址 A3 和传输预约指令 C3。然后,优先级决定部 247 将由接收到的传输预约指令 C3 表示的传输预约信息存储到存储器 247a(传输预约存储部)。优先级决定部 247 按照每个已存储的传输预约信息计算访问间隔 Da。访问间隔 Da 的计算方法与实施方式 1 相同。优先级决定部 247 根据计算出的访问间隔 Da 和从访问管理部 248 得到的访问经过时间 R5,决定最优先使预约处理部 243 进行处理的传输预约信息。优先级决定部 247 将由传输预约信息表示的、存储有要传输的数据的连续区域的起始地址 Am 作为地址 A5 发送给访问管理部 248,由访问管理部 248 接收访问经过时间 R5 作为其响应。此外,优先级决定部 247 将已决定的最优先的传输预约信息的对主存储器 10 上的连续区域的访问间隔 Da 以及存储有要传输的数据的连续区域的起始地址 Am 发送给预约处理部 243。在来自预约处理部 243 的预约传输状况信号 V1 表示传输完成,作为传输预约信息存储的存储有要传输的数据的连续区域的尺寸 H*V 全部被传输完时,优先级决定部 247 从存储器 247a 删除该传输预约信息。

[0145] 图 17 是示出优先级决定部 247 的存储器 247a 中存储的传输预约管理信息 201 的概略图。传输预约管理信息 201 具有到达顺序栏 201a、参照命令组起始地址栏 201b、连续区域起始地址栏 201c、传输剩余尺寸栏 201d 以及传输状况栏 201e。

[0146] 到达顺序栏 201a 存储表示传输预约信息的到达顺序的信息。

[0147] 参照命令组起始地址栏 201b 存储传输预约指令 C3 中包含的、作为参照存储有要传输的数据的连续区域的命令组的函数的起始地址 PROC。

[0148] 连续区域起始地址栏 201c 存储基于传输预约指令 C3 传输的数据存储在主存储器 10 中的连续区域的起始地址。此外,连续区域起始地址栏 201c 的初始值是传输预约指令 C3 中包含的、访问主机 1 参照的连续区域的起始地址 MM_ADDR。

[0149] 传输剩余尺寸栏 201d 存储基于传输预约指令 C3 传输的数据的剩余尺寸。此外,传输剩余尺寸栏 201d 的初始值是传输预约指令 C3 中包含的、存储有要传输的数据的连续区域的尺寸 H*V。

[0150] 传输状况栏 201e 存储传输状况信息,该传输状况信息表示是否正在根据与参照命令组起始地址栏 201b 和连续区域起始地址栏 201c 中存储的信息对应的传输预约信息传输数据。例如,在本实施方式中,如果该栏为“1”,则表示处于传输中,如果为“0”,则表示处于传输等待中。

[0151] 此外,优先级决定部 247 按照接收到传输预约指令 C3 的顺序,将由传输预约指令 C3 表示的访问主机 1 参照的连续区域的起始地址 MM_ADDR、存储有要传输的数据的连续区域的尺寸 H*V 以及作为参照存储有要传输的数据的连续区域的命令组的函数的起始地址 PROC 作为传输预约信息,存储到传输预约管理信息 201。优先级决定部 247 在从预约处理部 243 接收到预约传输状况信号 V1 时,更新存储器 247a 中存储的内容。

[0152] 返回到图 16,访问管理部 248 监视来自处理切换部 141 的请求地址 A2,利用定时器(未图示)计测从前次访问起经过的时间作为访问经过时间 Td,使存储器 248a(访问经

过时间存储部) 存储计测出的访问经过时间 T_d 。访问管理部 248 将属于主存储器 10 上的各段的、由按行地址和列地址分割的多个地址构成的连续区域作为 1 个单位, 每当从处理切换部 141 向请求处理部 142 发送请求地址 A_2 时, 确定请求地址 A_2 所属的连续区域, 将确定出的连续区域的访问经过时间 T_d 复位。此外, 在从优先级决定部 247 发送来地址 A_5 的情况下, 访问管理部 248 确定地址 A_5 所属的连续区域, 将确定出的连续区域的访问经过时间 T_d 作为响应 R_5 , 从存储器 248a 读出并将其发送给优先级决定部 247。

[0153] 图 18 是示出访问管理部 248 的存储器 248a 中存储的访问管理信息 202 的概略图。

[0154] 访问管理信息 202 具有段序号栏 202a、行地址栏 202b、列地址栏 202c 以及访问经过时间栏 202d。

[0155] 段序号栏 202a 存储用于识别主存储器 10 的段的段序号。

[0156] 行地址栏 202b 存储主存储器 10 的段中形成的连续区域的行地址的范围。

[0157] 列地址栏 202c 存储主存储器 10 的段中形成的连续区域的列地址的范围。

[0158] 访问经过时间栏 202d 存储访问经过时间 T_d , 该访问经过时间 T_d 表示在由段序号栏 202a 确定的段的、由行地址栏 202b 和列地址栏 202c 确定的连续区域中, 自前次访问起的经过时间。关于未曾被访问主机 1 访问过的连续区域, 访问经过时间 T_d 是自高速缓冲存储器控制器 200 启动起的经过时间。

[0159] 返回到图 16, 预约处理部 243 根据从命中检测部 130 发送的命中检测结果 R_2 、从优先级决定部 247 发送的存储有要传输的数据的连续区域的起始地址 A_m 以及对主存储器 10 的访问间隔 D_a , 对主存储器 10 进行访问, 向高速缓冲存储器 110 传输作为预先设定的传输单位的 1 行数据。在向高速缓冲存储器 110 传输 1 行数据完成时, 预约处理部 243 将预约传输状况信号 V_1 更新成表示传输已完成的值, 由此, 向优先级决定部 247 通知 1 行数据的传输已完成。例如, 预约处理部 243 将从优先级决定部 247 接收到的存储有要传输的数据的连续区域的起始地址 A_m 作为地址 A_4 发送给命中检测部 130。然后, 预约处理部 243 从命中检测部 130 取得地址 A_4 的命中检测结果 R_2 作为其响应。然后, 预约处理部 243 在命中检测结果 R_2 为高速缓冲未命中的情况下, 将该地址的数据从主存储器 10 传输到高速缓冲存储器 110。此外, 在存储有要传输的数据的连续区域的起始地址 A_m 的传输开始时, 预约处理部 243 将表示存储目的地的高速缓冲行的预约区域信息 I_1 发送给释放处理部 144。最后, 在从存储有要传输的数据的连续区域的起始地址 A_m 起完成 1 行数据的传输时, 预约处理部 243 将预约传输状况信号 V_1 更新成表示传输已完成, 向优先级决定部 247 进行通知。预约传输状况信号 V_1 例如为 1 比特的信号, 在传输已完成而能够接收下一个传输预约的情况下设为 $H(1)$, 相反, 在不能接收其它传输预约的情况下设为 $L(0)$ 。

[0160] 接下来, 使用流程图, 对优先级决定部 247 中的处理流程进行说明。

[0161] 图 19 是示出优先级决定部 247 对多个传输预约信息进行仲裁时的处理的流程图。优先级决定部 247 在从处理切换部 141 接收到包含运行地址 A_3 、存储有要传输的数据的连续区域的起始地址 MM_ADDR 、存储有要传输的数据的连续区域的尺寸 $H*V$ 以及参照存储有要传输的数据的连续区域的命令组的起始地址 $PROC$ 的传输预约指令 C_3 时, 开始处理。

[0162] 首先, 优先级决定部 247 判定预约处理部 243 是否正在处理 (S60)。在来自预约处理部 243 的预约传输状况信号 V_1 表示不能接收传输预约 ($V_1 = L$) 的情况下 (S60 : 否), 等待到传输完成。在来自预约处理部 243 的预约传输状况信号 V_1 表示能够接收传输预约

(V1 = H) 的情况下 (S60 :是), 优先级决定部 247 使处理进入到步骤 S61。

[0163] 接下来, 在步骤 S61 中, 优先级决定部 247 判定存储器 247a 中存储的传输预约管理信息 201 的传输状况栏 201e 是否有表示传输中的传输预约信息。在已存储的传输预约信息的任意 1 个中, 在传输预约管理信息 201 的传输状况栏 201e 存储有表示传输中的“1”的情况下 (S61 :是), 优先级决定部 247 使处理进入到步骤 S62。另一方面, 在已存储的全部传输预约信息的传输状况栏 201e 存储有表示传输等待中的“0”的情况下 (S61 :否), 优先级决定部 247 使处理进入到步骤 S63。

[0164] 在步骤 S62 中, 优先级决定部 247 更新传输预约管理信息 201 的传输状况栏 201e 中存储有表示传输中的“1”的记录。具体而言, 优先级决定部 247 从传输预约管理信息 201 的对应记录的传输剩余尺寸栏 201d 的值减去作为传输单位的高速缓冲存储器 10 的 1 行的尺寸 (在此, 例如为“1”), 由此进行更新。此外, 优先级决定部 247 在传输剩余尺寸栏 201d 的值变为“0”的情况下, 删除该传输预约信息 (记录)。另一方面, 在从传输剩余尺寸栏 201d 的值减去 1 行的尺寸之后传输剩余尺寸栏 201d 的值仍为“1”以上的情况下, 优先级决定部 247 将连续区域起始地址栏 201c 中存储的连续区域的起始地址加上 1 行。换言之, 优先级决定部 247 将连续区域的起始地址更新成 1 行后的地址。此外, 优先级决定部 247 将对应记录的传输状况栏 201e 更新成表示等待中的“0”, 使处理进入到步骤 S63。

[0165] 在步骤 S63 中, 优先级决定部 247 判定存储器 247a 中存储的、传输等待中的传输预约信息的个数。在传输预约信息的个数为“0”个的情况下, 优先级决定部 247 结束流程。在传输预约信息的个数为“1”个的情况下, 优先级决定部 247 使处理进入到步骤 S64。此外, 在传输预约信息的个数为“2”个以上的情况下, 优先级决定部 247 使处理进入到步骤 S65。

[0166] 在步骤 S64 中, 优先级决定部 247 将从处理切换部 141 接收到的 1 个传输预约信息, 换言之存储器 247a 中存储的唯一的传输预约信息决定为最优先传输预约信息。然后, 处理进入到步骤 S69。

[0167] 在步骤 S65 中, 优先级决定部 247 计算存储器 247a 中存储的各传输预约信息的访问间隔 D_a 。然后, 处理进入到步骤 S66。访问间隔 D_a 按照实施方式 1 的 (2) 式来计算。

[0168] 接下来, 在步骤 S66 中, 优先级决定部 247 判定在步骤 S65 中计算出的各传输预约信息的访问间隔 D_a 是否为同等程度。此时, 在各传输预约信息的访问间隔 D_a 为预先设定的容许误差范围内的情况下, 优先级决定部 247 判定为同等程度 (S66 :是), 使处理进入到步骤 S67。在不为同等程度的情况下 (S66 :否), 优先级决定部 247 使处理进入到步骤 S68。访问间隔 D_a 的容许误差为预先设定的值, 例如“ ± 10 ”。

[0169] 在步骤 S67 中, 优先级决定部 247 根据来自访问管理部 248 的、针对各传输预约信息的访问经过时间 R_5 , 决定最优先传输预约信息。例如, 优先级决定部 247 将存储器 247a 中存储的传输预约管理信息 201 的连续区域起始地址栏 201c 中存储的地址 A5 发送给访问管理部 248, 作为其响应而取得访问经过时间 R_5 。然后, 优先级决定部 247 将访问经过时间 R_5 最大的值的传输预约信息决定为最优先传输预约信息。然后, 处理进入到步骤 S69。

[0170] 在步骤 S68 中, 优先级决定部 247 将在步骤 S65 中计算出的全部传输预约信息的访问间隔 D_a 中访问间隔 D_a 最小的传输预约信息决定为最优先传输预约信息。然后, 处理进入到步骤 S69。

[0171] 在步骤 S69 中, 优先级决定部 247 根据在步骤 S64、S67 或 S68 中决定的最优先传

输预约信息,将传输预约管理信息 201 的对应的连续区域起始地址栏 201c 中存储的地址作为存储有要传输的数据的连续区域的起始地址 A_m 。然后,优先级决定部 247 将存储有要传输的数据的连续区域的起始地址 A_m 和访问间隔 D_a 发送给预约处理部 243。然后,处理进入到步骤 S70。

[0172] 在步骤 S70 中,将存储器 247a 中存储的传输预约管理信息 201 中的、在步骤 S69 中发送给预约处理部 243 的传输预约信息的传输状况从“0(传输等待中)”变更成“1(传输中)”。然后,处理进入到步骤 S60。

[0173] 如上所述,优先级决定部 247 基于根据运行地址 A_3 计算出的访问间隔 D_a ,根据存储器 247a 中存储的全部传输预约信息,决定最优先传输预约信息,由此,即使对于更加紧急的传输预约信息,也能够访问主机 1 执行由起始地址 PROC 表示的命令组之前,完成所需的数据传输,能够防止高速缓冲存储器 110 的浪费。

[0174] 此外,优先级决定部 247 基于根据运行地址 A_3 计算出的访问间隔 D_a 和来自访问管理部 248 的访问经过时间 R_5 ,根据存储器 247a 中存储的全部传输预约信息,决定最优先传输预约信息,由此,即使针对从高速缓冲存储器 110 释放的可能性高的主存储器 10 上的连续区域,也能够提高高速缓冲命中率。

[0175] 图 20 ~ 图 24 分别是示出优先级决定部 247 从处理切换部 141 接收的 2 个传输预约信息、存储有要传输的数据的连续区域的尺寸以及参照存储有要传输的数据的连续区域的命令组的起始地址之间的关系的关系的概略图。图 20 ~ 图 24 的纵轴表示存储有要传输的数据的连续区域的尺寸,其横轴表示时间。优先级决定部 247 在时间“0”开始决定从处理切换部 141 接收到的传输预约信息 #1 和传输预约信息 #2 的优先级的处理。横轴记载的 $t_{\#1}$ 和 $t_{\#2}$ 分别表示由从处理切换部 141 接收到的传输预约信息 #1 和传输预约信息 #2 表示的、参照存储有要传输的数据的连续区域的命令组的起始地址 PROC1 以及起始地址 PROC2 分别被执行的时序。此外,图中的 $H_1 \times V_1$ 和 $H_2 \times V_2$ 分别表示由从处理切换部 141 接收到的传输预约信息 #1 和传输预约信息 #2 表示的、存储有要传输的数据的连续区域的尺寸。

[0176] 图 20 和图 21 分别是在图 19 的步骤 S66 中判断为全部传输预约信息的访问间隔 D_a 不是同等程度的例子。

[0177] 图 20 为示出如下情况的一例:与传输预约信息 #2 相比,在传输预约信息 #1 中,存储有要传输的数据的连续区域的尺寸较大,到参照存储有要传输的数据的连续区域的命令组的起始地址为止的命令步数较少,换言之,到开始执行参照存储有要传输的数据的连续区域的命令组为止的时间较短。在该情况下,关于在图 19 的步骤 S65 中计算出的访问间隔 D_a ,与传输预约信息 #2 的值相比,传输预约信息 #1 的值超过容许误差而变小。因此,在图 19 的步骤 S66 中,判定为传输预约信息的访问间隔 D_a 不是同等程度(S66:否),将访问间隔 D_a 较小的传输预约信息 #1 决定为最优先传输预约信息。

[0178] 图 21 与图 20 相反,是优先级决定部 247 将传输预约信息 #2 决定为最优先传输预约信息的例子。在图 21 中,与传输预约信息 #1 相比,在传输预约信息 #2 中,存储有要传输的数据的连续区域的尺寸较大,到参照存储有要传输的数据的连续区域的命令组的起始地址为止的命令步数也较大,换言之,到开始执行参照存储有要传输的数据的连续区域的命令组为止的时间较长。因此,在传输预约信息 #2 中,虽然到开始执行参照存储有要传输的数据的连续区域的命令组为止的时间较长,但是参照存储有要传输的数据的连续区域的命

令组的起始地址 PROC1 与 PROC2 的差分较小。因此,关于在图 19 的步骤 S65 中计算出的访问间隔 D_a ,与传输预约信息 #1 的值相比,传输预约信息 #2 的值超过容许误差而变小。因此,在图 19 的步骤 S66 中,判定为传输预约信息的访问间隔 D_a 不是同等程度 (S66:否),将访问间隔 D_a 较小的传输预约信息 #2 决定为最优先传输预约信息。

[0179] 如上所述,优先级决定部 247 将访问间隔 D_a 最小的传输预约信息决定为最优先传输预约信息,将计算出的访问间隔 D_a 和由最优先传输预约信息表示的、存储有要传输的数据的连续区域的起始地址 A_m 发送给预约处理部 243。由此,不需要预约处理部 243 计算访问间隔 D_a ,能够使预约处理的开始提前,因而能够提高从主存储器 10 向高速缓冲存储器 110 的数据传输效率。

[0180] 图 22 ~ 图 24 分别是在图 19 的步骤 S66 中判断为全部传输预约信息的访问间隔 D_a 是同等程度的例子。

[0181] 在图 22 中,在传输预约信息 #1 和传输预约信息 #2 双方中,相对于存储有要传输的数据的连续区域的尺寸,到参照存储有要传输的数据的连续区域的命令组的起始地址为止的命令步数较少。因此,在图 19 的步骤 S65 中计算出的访问间隔 D_a 的值均同等程度地变小,在步骤 S66 中,判断为任何传输预约信息的访问间隔 D_a 均为同等程度。

[0182] 在图 23 中,在传输预约信息 #1 中,虽然存储有要传输的数据的连续区域的尺寸非常小,但是到参照存储有要传输的数据的连续区域的命令组的起始地址为止的命令步数也较少。另一方面,在传输预约信息 #2 中,虽然存储有要传输的数据的连续区域的尺寸较大,但是到参照存储有要传输的数据的连续区域的命令组的起始地址为止的命令步数也较大。因此,在图 19 的步骤 S65 中计算出的访问间隔 D_a 的值均同等程度地变小,在步骤 S66 中,判断为任何传输预约信息的访问间隔 D_a 均为同等程度。

[0183] 在图 24 中,在传输预约信息 #1 中,虽然存储有要传输的数据的连续区域的尺寸较大,但是到参照存储有要传输的数据的连续区域的命令组的起始地址为止的命令步数也非常大。此外,在传输预约信息 #2 中,也是虽然存储有要传输的数据的连续区域的尺寸非常大,但是到参照存储有要传输的数据的连续区域的命令组的起始地址为止的命令步数也非常大。因此,在图 19 的步骤 S65 中计算出的访问间隔 D_a 的值均同等程度地变大,在步骤 S66 中,判断为任何传输预约信息的访问间隔 D_a 均为同等程度。

[0184] 在图 22 ~ 图 24 中的任意一个情况下,在相对于传输预约信息 #1,传输预约信息 #2 的访问经过时间 R_5 较短的情况下,在图 19 的步骤 S67 中,将访问经过时刻 R_5 较长的传输预约信息 #1 决定为最优先传输预约信息。相反,在相对于传输预约信息 #1,传输预约信息 #2 的访问经过时间 R_5 较长的情况下,在图 19 的步骤 S67 中,将访问经过时间 R_5 较长的传输预约信息 #2 决定为最优先传输预约信息。

[0185] 如上所述,通过将访问经过时间 R_5 最大的传输预约信息决定为最优先传输预约信息,由此,即使针对优先级决定部 247 计算出的访问间隔 D_a 为同等程度的多个传输预约信息,也能够防止从高速缓冲存储器 110 释放的可能性高的连续区域的释放,能够使高速缓冲存储器控制器 200 的处理高速化。

[0186] 如图 23 所示,在计算出的访问间隔 D_a 均较小,传输预约信息 #1 和传输预约信息 #2 中包含的、参照存储有要传输的数据的连续区域的命令组的起始地址 PROC1 与 PROC2 的差分较大的情况下,优先级决定部 247 将访问经过时刻 R_5 最大的传输预约信息决定为最优先

先传输预约信息,但是不限于此。例如,也可以是,在访问间隔 D_a 为预先设定的第 1 阈值以下,并且参照存储有要传输的数据的连续区域的命令组的起始地址 PROC1 与 PROC2 的差分为预先设定的第 2 阈值以上的情况下,优先级决定部 247 将参照存储有要传输的数据的连续区域的命令组的起始地址最小的传输预约信息决定为最优先传输预约信息。由此,高速缓冲存储器 110 不会不需要地存储数据,能够防止高速缓冲存储器 110 的浪费。

[0187] 如图 24 所示,当与传输预约信息 #1 相比,在传输预约信息 #2 中,存储有要传输的数据的连续区域的尺寸和到参照存储有要传输的数据的连续区域的命令组的起始地址为止的命令步数均较大,在图 19 的步骤 S65 中计算出的访问间隔 D_a 的值均同等程度地大的情况下,优先级决定部 247 将访问经过时间 R_5 最大的传输预约信息决定为最优先传输预约信息,但是不限于此。例如,也可以是,在计算出的访问间隔 D_a 为预先设定的第 1 阈值以上的情况下,优先级决定部 247 等待到访问间隔 D_a 变为第 1 阈值以下或输入下一个传输预约信息为止。由此,在接收到访问间隔 D_a 比之前接收到的传输预约信息小的传输预约信息的情况下,优先级决定部 247 也能够优先处理访问间隔 D_a 较小的传输预约信息,能够高效地使用高速缓冲存储器 110。

[0188] 以上对优先级决定部 247 的存储器 247a 中存储的传输预约信息为 2 个的情况下的例子进行了说明,但是,实际上不限于此。在存储器 247a 中存储有 3 个以上的传输预约信息的情况下,通过与上述同样地根据访问间隔 D_a 和访问经过时间来决定最优先传输预约信息,也能够防止高速缓冲存储器 110 的浪费。

[0189] 图 25 是示出优先级决定部 247 进行的处理的时序图的一例的概略图。图 25 示出从处理切换部 141 输入运行地址 A3 和命令指令 C3 的时序、决定最优先传输预约信息的时序、以及根据优先级决定部 247 从处理切换部 141 接收到的多个传输预约信息切换要传输的数据中的剩余数据的尺寸的时序。

[0190] 在时刻 t_{20} ,在从处理切换部 141 输入运行地址 A3(1fc00fff) 和传输预约指令 C3(TRI#1) 时,优先级决定部 247 基于传输预约指令 C3,使存储器 247a 存储传输预约信息 (TR#1)。然后,优先级决定部 247 决定最优先传输预约信息。在时刻 t_{20} 的阶段中,优先级决定部 247 接收到的传输预约信息只有 TR#1,因此,优先级决定部 247 将 TR#1 决定为最优先传输预约信息。然后,优先级决定部 247 将该连续区域的起始地址 A_m 和访问间隔 D_a 输出到预约处理部 243,开始预约处理。预约处理部 243 完成由 TR#1 指定的存储有要传输的数据的连续区域的尺寸 $H_1 \times V_1$ 中的 1 行的传输,优先级决定部 247 使处理等待,直到预约传输状况信号 V_1 变成能够接收传输预约状态 ($V_1 = H$) 的时刻 t_{24} 为止。

[0191] 在时刻 $t_{21} \sim t_{23}$,从处理切换部 141 向优先级决定部 247 输入的运行地址 A3 和传输预约指令 C3 被更新。优先级决定部 247 根据在时刻 t_{21} 输入的传输预约指令 C3 (TRI#2),使存储器 247a 存储传输预约信息 (TR#2),此外,根据在时刻 t_{23} 输入的传输预约指令 C3 (TRI#3),使存储器 247a 存储传输预约信息 (TR#3)。但是,由于来自预约处理部 243 的预约传输状况信号 V_1 为不能接收传输预约状态 ($V_1 = L$),因而优先级决定部 247 使处理等待。

[0192] 接下来,在时刻 t_{24} ,来自预约处理部 243 的预约传输状况信号 V_1 变成能够接收传输预约状态 ($V_1 = H$),因而优先级决定部 247 开始最优先传输预约信息的决定处理。

[0193] 在时刻 t_{25} ,优先级决定部 247 决定最优先传输预约信息为 TR#3,将该连续区域的

起始地址 A_m 和访问间隔 D_a 输出到预约处理部 243, 开始预约处理。

[0194] 在时刻 t_{27} , TR#3 中的 1 行的连续区域的传输完成, 来自预约处理部 243 的预约传输状况信号 V_1 成为能够接收传输预约状态 ($V_1 = H$), 因此, 与时刻 t_{24} 同样地, 优先级决定部 247 开始最优先传输预约信息的决定处理。在此, 将 TR#2 决定为最优先传输预约信息。

[0195] 以后, 同样地, 每当在各个时刻来自预约处理部 243 的预约传输状况信号 V_1 变成能够接收传输预约状态 ($V_1 = H$) 时, 优先级决定部 247 决定最优先传输预约信息, 向预约处理部 243 输出最优先传输预约信息的连续区域的起始地址 A_m 和访问间隔 D_a 。而且, 如果来自预约处理部 243 的预约传输状况信号 V_1 变成不能接收传输预约状态 ($V_1 = L$), 则优先级决定部 247 使处理等待。

[0196] 最后, 在时刻 t_{31} , 在优先级决定部 247 的存储器 247a 存储的最后的传输预约信息 (TR#2) 的传输完成时, 来自预约处理部 243 的预约传输状况信号 V_1 保持能够接收传输预约状态 ($V_1 = H$), 优先级决定部 247 使处理等待, 直到从处理切换部 141 输入传输预约指令 C_3 为止。

[0197] 如上所述, 优先级决定部 247 每当按照传输预约信息完成作为预先设定的传输单位尺寸的 1 行的传输时, 根据存储器 247a 中存储的全部传输预约信息决定最优先传输预约信息, 在执行参照各连续区域的命令组之前完成传输, 由此, 能够提高对访问主机 1 的数据传输效率。

[0198] 接下来, 使用流程图, 对数据处理部 240 的预约处理部 243 进行的处理流程进行说明。

[0199] 图 26 是示出预约处理部 243 基于传输预约指令 C_3 传输数据时的处理的流程图。预约处理部 243 在从命中检测部 130 输入表示高速缓冲命中或高速缓冲未命中的命中检测结果 R_2 并从优先级决定部 247 输入存储有要传输的数据的连续区域的起始地址 A_m 和对主存储器的访问间隔 D_a 时, 开始处理。

[0200] 首先, 预约处理部 243 判定在高速缓冲存储器 110 中是否存在与起始地址 A_m 对应的数据 (S_{80})。在此, 与起始地址 A_m 对应的数据是从起始地址 A_m 起存储的传输单位长的数据。预约处理部 243 根据来自命中检测部 130 的命中检测结果 R_2 进行该判定。然后, 在高速缓冲存储器 110 中不存在与起始地址 A_m 对应的数据的情况下 (S_{80} : 否), 需要将该数据从主存储器 10 传输到高速缓冲存储器, 因此, 预约处理部 243 使处理进入到步骤 S_{81} 。另一方面, 在高速缓冲存储器 110 中存在与起始地址 A_m 对应的数据的情况下 (S_{80} : 是), 不需要将该数据从主存储器 10 传输到高速缓冲存储器 110, 因此, 预约处理部 243 不传输该数据而使处理进入到步骤 S_{82} 。

[0201] 在步骤 S_{81} 中, 预约处理部 243 按照从优先级决定部 247 输入的访问间隔 D_d , 将进行从主存储器 10 到高速缓冲存储器 110 的数据传输的命令发送给主存储器访问仲裁部 146。接收到进行传输的命令的主存储器访问仲裁部 146 从主存储器 10 读出作为传输对象的数据并发送给预约处理部 243。预约处理部 243 将发送来的数据发送给高速缓冲存储器仲裁部 145, 使其写入到高速缓冲存储器 110。此外, 此后的存储器管理部 120 的处理与实施方式 1 相同, 因而省略说明。

[0202] 接下来, 预约处理部 243 判定传输完成尺寸是否为作为传输单位的 1 行的尺寸以上 (S_{82})。在传输完成尺寸小于 1 行尺寸的情况下 (S_{82} : 否), 处理返回到步骤 S_{80} , 在传输

完成尺寸为 1 行尺寸以上的情况下 (S82 :是), 处理进入到 S83。

[0203] 在步骤 S83 中, 预约处理部 243 将预约传输状况信号 V1 设定成能够接收传输预约状态 (V1 = H) 而输出到优先级决定部 247, 结束预约处理, 该能够接收传输预约状态表示已完成从存储有要传输的数据的连续区域的起始地址 Am 起 1 行的传输。

[0204] 如上所述, 预约处理部 243 逐行传输从优先级决定部 247 输入的传输预约信息中的各行, 由此, 能够高效地使用高速缓冲存储器 110。

[0205] 此外, 在以上记载的预约处理部 243 的处理流程中, 对逐行传输由传输预约指令 C3 指定的存储有要传输的数据的连续区域中的各行的例子进行了说明, 但不限于这样的例子。例如, 也可以是, 每当从处理切换部 141 向优先级决定部 247 输入运行地址 A3 时, 优先级决定部 247 决定最优先传输预约信息并输出到预约处理部 243, 预约处理部 243 据此进行传输预约处理。在这样的情况下, 需要考虑每当对主存储器 10 的行 (Row) 地址不同的区域进行访问时产生的预充电时间而计算访问间隔 Da。由此, 即使针对访问间隔 Da 较小的多个传输预约信息, 也能够进行从主存储器 10 向高速缓冲存储器 110 的传输, 能够提高高速缓冲命中率。

[0206] 在此, 对考虑到预充电时间的访问间隔 Da 的计算方法的一例进行说明。

[0207] 首先, 优先级决定部 247 判定是否需要考虑预充电时间。在主存储器 10 上的 Row 地址不同的情况下, 产生预充电时间。因此, 优先级决定部 247 将由前次决定的最优先传输预约信息传输的数据的连续区域的地址存储到存储器 247a, 通过进行 Row 地址的比较, 判定作为前次传输对象的传输预约信息的 Row 地址与作为本次传输对象的传输预约信息的 Row 地址是否不同。在 Row 地址不同的情况下, 优先级决定部 247 考虑预充电时间。由于访问间隔 Da 是每传输单位尺寸的传输所需的命令步数, 因而将预充电时间 Tpri (Cycle) 转换成命令步数。预充电时间 Tpri 向命令步数的转换按下述 (3) 式来进行。

[0208]

$$\text{变换后预充电 } Spri = \frac{1 \text{ 命令步骤耗费的周期数 } Tos}{\text{预充电时间 } Tpri} \quad (3)$$

[0209] 在此, 在式 (3) 中, 优先级决定部 247 预先保持定时器, 对执行计测出的 1 命令步骤耗费的周期数 Tos 进行计测, 将其值用于预充电时间 Tpri 的转换。此外, 执行 1 命令步骤耗费的周期数 Tos 是从前次命令起到本次命令为止耗费的周期数或到本次命令为止的执行 1 命令耗费的周期数的平均值。

[0210] 然后, 使用利用 (3) 式计算出的变换后预充电 Spri, 利用 (4) 式计算考虑到预充电时间的访问间隔 Dap。

$$Dap = (\text{访问间隔 } Da) - (\text{变换后预充电 } Spri) \quad (4)$$

[0212] 此外, 预约处理部 243 也可以将由传输预约信息表示的、存储有要传输的数据的连续区域的尺寸 H*V 作为传输单位。在该情况下, 从优先级决定部 247 向预约处理部 243 发送存储有要传输的数据的连续区域的尺寸 H*V, 在传输完成后, 将预约传输状况信号 V1 设为能够接收传输预约状态 (V1 = H)。由此, 能够减少决定最优先传输预约信息的处理, 能够缩短决定最优先传输预约信息所需的计算时间, 提高从主存储器 10 向高速缓冲存储器 110 的数据传输效率。

[0213] 图 27 是示出将预约处理部 243 的传输单位设为由传输预约信息表示的、存储有要

传输的数据的连续区域的尺寸 $H*V$ 的情况下, 优先级决定部 247 进行的处理的时序图的一例的概略图。图 27 示出从处理切换部 141 输入运行地址 A3 和命令指令 C3 的时序、决定最优先传输预约信息的时序、以及根据优先级决定部 247 从处理切换部 141 接收到的多个传输预约信息切换存储有要传输的数据的连续区域中剩余的连续区域的尺寸的时序。

[0214] 在时刻 t_{30} , 在从处理切换部 141 输入运行地址 A3 (1fc00fff) 和传输预约指令 C3 (TRI#1) 时, 优先级决定部 247 基于传输预约指令 C3, 使存储器 247a 存储传输预约信息 (TR#1)。然后, 优先级决定部 247 决定最优先传输预约信息。在时刻 t_{30} 的阶段中, 由于优先级决定部 247 接收到的传输预约信息只有 TR#1, 因而优先级决定部 247 将 TR#1 决定为最优先传输预约信息。然后, 优先级决定部 247 将由 TR#1 表示的存储有要传输的数据的连续区域的起始地址 PROC1、存储有要传输的数据的连续区域的尺寸 $H1*V1$ 以及计算出的访问间隔 D_a 输出到预约处理部 243, 开始预约处理。

[0215] 在时刻 $t_{31} \sim t_{35}$, 从处理切换部 141 向优先级决定部 247 输入的运行地址 A3 和传输预约指令 C3 被更新。优先级决定部 247 根据在时刻 t_{31} 输入的传输预约指令 C3 (TRI#2), 使存储器 247a 存储传输预约信息 (TR#2), 此外, 根据在时刻 t_{33} 输入的传输预约指令 C3 (TRI#3), 使存储器 247a 存储传输预约信息 (TR#3)。但是, 由于来自预约处理部 243 的预约传输状况信号 V1 为不能接收传输预约状态 ($V1 = L$), 因而优先级决定部 247 使处理等待。

[0216] 在时刻 t_{35} , 从预约处理部 243 输入的预约传输状况信号 V1 变成能够接收传输预约状态 ($V1 = H$), 因而优先级决定部 247 决定下一个最优先传输预约信息。

[0217] 在时刻 t_{36} , 优先级决定部 247 将 TR#3 决定为最优先传输预约信息。然后, 优先级决定部 247 将由决定为最优先传输预约信息的 TR#3 表示的存储有要传输的数据的连续区域的起始地址 PROC3、存储有要传输的数据的连续区域的尺寸 $H3*V3$ 以及计算出的访问间隔 D_a 输出到预约处理部 243, 开始预约处理。然后, 优先级决定部 247 使处理等待, 直到从从预约处理部 243 输入的预约传输状况信号 V1 变成能够接收传输预约状态 ($V1 = H$) 为止。

[0218] 以后, 与上述同样地, 在时刻 t_{37} , 在预约传输状况信号 V1 变成能够接收传输预约状态 ($V1 = H$) 时, 优先级决定部 247 再次决定最优先传输预约信息, 将由作为已决定的最优先传输预约信息的传输预约信息 #2 表示的、存储有要传输的数据的连续区域的起始地址 PROC2、存储有要传输的数据的连续区域的尺寸 $H2*V2$ 以及计算出的访问间隔 D_a 输出到预约处理部 243。然后, 使处理等待, 直到从预约处理部 243 输入的预约传输状况信号 V1 变成能够接收传输预约状态 ($V1 = H$) 为止。

[0219] 在以上记载的访问管理部 248 中, 将从前次访问起经过的时间作为访问经过时间 T_d 存储到存储器 248a, 但是不限于此。例如, 也可以是, 在从前次访问起经过的时间经过了按照高速缓冲存储器 110 的各高速缓冲行中决定释放候选的 LRU 方式设定的时间的情况下, 访问管理部 248 将对应的访问经过时间 T_d 复位。由此, 能够优先将如下的传输预约信息输出到预约处理部 243: 该传输预约信息传输从高速缓冲存储器 110 释放的可能性更高的高速缓冲行中存储的主存储器 10 上的连续区域中存储的数据。因此, 能够减少释放处理部 144 进行释放处理的次数, 使高速缓冲存储器控制器 200 的处理高速化。

[0220] 在以上记载的高速缓冲存储器控制器 200 中, 对访问主机 1 为 1 个的情况进行了

说明,但是不限于这样的例子。高速缓冲存储器控制器 200 可以连接多个访问主机 1。图 28 是示出高速缓冲存储器控制器 300 与访问主机 1#1 和访问主机 1#2 这 2 台访问主机连接的例子的概略图。

[0221] 在与高速缓冲存储器控制器 300 连接的访问主机 1 存在多个的情况下,处理切换部 341 在存储器 341a 内存储每个访问主机 1 的运行地址 A3。因此,在从连接的访问主机 1#1 或 1#2 输入的命令地址 A1#1 或 A1#2 为图 7 所示的命令区域 190 中包含的地址的情况下,处理切换部 341 将其作为存储器 341a 内的、输入命令地址的访问主机 1#1 或 1#2 的运行地址 A3 进行存储,或者,在已经存储有该访问主机 1#1 或 1#2 的运行地址 A3 的情况下,更新其值。然后,处理切换部 341 将作为运行地址 A3 存储的命令地址和访问主机编号 Mn 输出到优先级决定部 347,其中,访问主机编号 Mn 表示该访问主机 1#1 或 1#2,是按照每个访问主机而预先设定的。此外,在从访问主机 1#1 或 1#2 输入的命令指令 C1#1 或 C1#2 既不是读也不是写的情况下,将命令指令 C1#1 或 C1#2 作为传输预约指令 C3 输出到优先级决定部 347。

[0222] 接下来,优先级决定部 347 从处理切换部 341 接收传输预约指令 C3、运行地址 A3 以及访问主机编号 Mn。然后,优先级决定部 347 使存储器 347a(传输预约存储部)存储接收到的访问主机编号 Mn、由接收到的传输预约指令 C3 表示的访问主机 1 参照的连续区域的起始地址 MM_ADDR、该连续区域的尺寸 H*V 以及作为参照该连续区域的命令组的函数的起始地址 PROC。此外,优先级决定部 347 使存储器 347b(运行地址存储部)存储接收到的访问主机编号 Mn 和接收到的运行地址 A3。优先级决定部 347 基于存储器 347b 中存储的各访问主机 1 的运行地址计算访问间隔 Da,决定最优先传输预约信息。

[0223] 图 29 是示出优先级决定部 347 的存储器 347a 中存储的传输预约管理信息 301 的概略图。传输预约管理信息 301 具有到达顺序栏 301a、访问主机编号栏 301f、参照命令组起始地址栏 301b、连续区域起始地址栏 301c、传输剩余尺寸栏 301d 以及传输状况栏 301e。此外,图 29 中的到达顺序栏 301a、参照命令组起始地址栏 301b、连续区域起始地址栏 301c、传输剩余尺寸栏 301d 以及传输状况栏 301e 与图 17 中的到达顺序栏 201a、参照命令组起始地址栏 201b、连续区域起始地址栏 201c、传输剩余尺寸栏 201d 以及传输状况栏 201e 相同,因而省略说明。

[0224] 访问主机编号栏 301f 存储从处理切换部 341 发送的访问主机编号 Mn。

[0225] 图 30 是示出优先级决定部 347 的存储器 347b 中存储的运行地址管理信息 303 的概略图。运行地址管理信息具有访问主机编号栏 303a 和运行地址栏 303b。

[0226] 访问主机编号栏 303a 存储从处理切换部 341 发送的访问主机编号 Mn。

[0227] 运行地址栏 303b 存储从处理切换部 341 发送的运行地址 A3。

[0228] 如上所述,能够决定从与高速缓冲存储器控制器 300 连接的多个从访问主机 1 输入的多个传输预约信息的优先级,根据各访问主机 1 的程序的运行状况,从主存储器 10 向高速缓冲存储器 110 传输数据,由此,能够设计出高速缓冲存储器控制器 300 与多个访问主机 1 连接的系统。因此,能够扩大要构筑的系统的规模。

[0229] 在以上记载的高速缓冲存储器控制器 100 ~ 300 中,指定了主存储器 10 上的数据区域作为传输预约信息表示的存储有要传输的数据的连续区域,但是不限于此。存储有要传输的数据的连续区域也可以是主存储器 10 上的命令区域。这样的情况下,只要由传输预

约信息指定的参照存储有要传输的数据的连续区域的命令组的起始地址 PROC,是执行要传输的属于主存储器 10 上的命令区域的连续区域的命令之前的命令的地址即可。由此,例如,即使在访问主机 1 执行的命令是分支命令的情况下,也能够执行它们之前将分支目的地的命令组传输到高速缓冲存储器 110,能够防止访问主机 1 的工作速度下降。

[0230] 以上记载的处理切换部 141、341 通过分析来自自己连接的访问主机 1 的命令指令 C1 来判断是读还是写,从而切换处理,但是不限于此。例如也可以是,在从访问主机 1 输入的命令指令 C1 是读和写中的任意一个且这样的命令指令 C1 中附加有优先级决定部 247、347 的存储器 247a、347a 的地址的情况下,处理切换部 141、341 通过对输入的地址进行解码来进行切换处理。在该情况下,在命令指令 C1 中附加的地址不表示主存储器 10 上的地址而表示优先级决定部 247、347 的存储器 247a、347a 的地址的情况下,处理切换部 141、341 可以判断为是传输预约信息。由此,作为访问主机 1 可以使用通用的 CPU 等硬件,在访问主机 1 与高速缓冲存储器控制器 100 ~ 300 的连接中,例如可以使用 AMBA AXI (Advanced eXtensible Interface :高级可扩展接口) 等通用总线,能够提高高速缓冲存储器控制器 100 ~ 300 的通用性。

[0231] 在以上记载的预约处理部 143、243 中,在命中检测结果 R2 为高速缓冲未命中的情况下,将其地址的数据从主存储器 10 传输到高速缓冲存储器 110,将表示存储目的地的高速缓冲行的预约区域信息 I1 发送给释放处理部 144,但是不限于此。例如也可以是,即使在命中检测结果 R2 为高速缓冲命中的情况下,也将表示存储目的地的高速缓冲行的预约区域信息 I1 发送给释放处理部 144。在该情况下,释放处理部 144 在接收到预约区域信息 I1 时,将标签存储器 121 中存储的、预约区域信息 I1 表示的高速缓冲行的访问标记 Fa 设定成表示访问主机 1 没有访问的“无效”,并且,将预约区域访问标记 Fra 设定成表示访问标记 Fa 未曾变成有效。由此,即使在访问主机 1 执行由传输预约信息表示的、参照存储有要传输的数据的连续区域的命令组之前耗费时间的情况下,也能够可靠地实现高速缓冲命中而不释放该高速缓冲行。

[0232] 标号说明

[0233] 1 访问主机,10 主存储器,100、200、300 高速缓冲存储器控制器,110 高速缓冲存储器,120 存储器管理部,130 命中检测部,140、240、340 数据处理部,141、341 处理切换部,142 请求处理部,143、243 预约处理部,144 释放处理部,145 高速缓冲存储器访问仲裁部,146 主存储器访问仲裁部,247、347 优先级决定部,248 访问管理部。

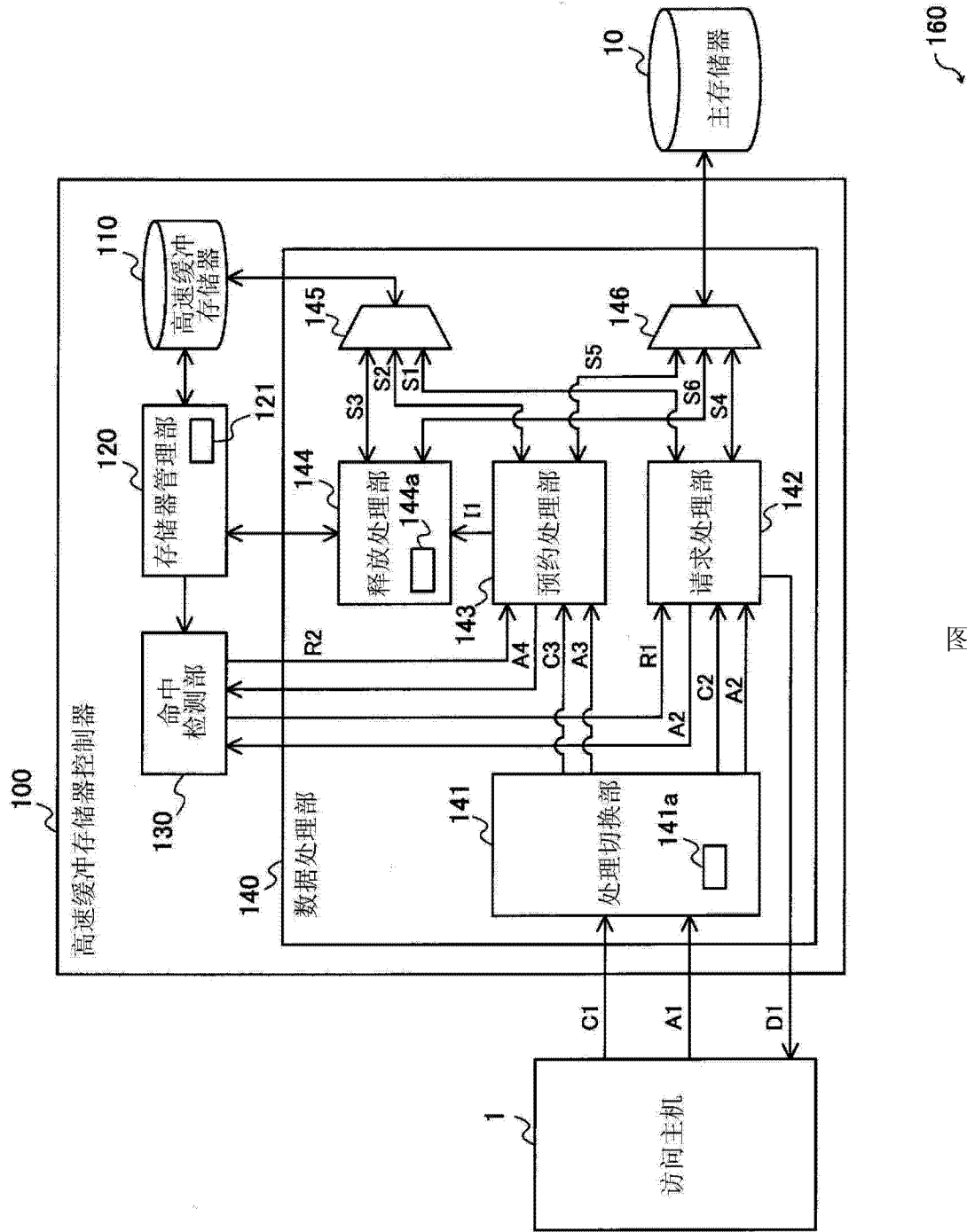


图 1

图 2
cache_reserve(MM_ADDR, H*V, PROC);

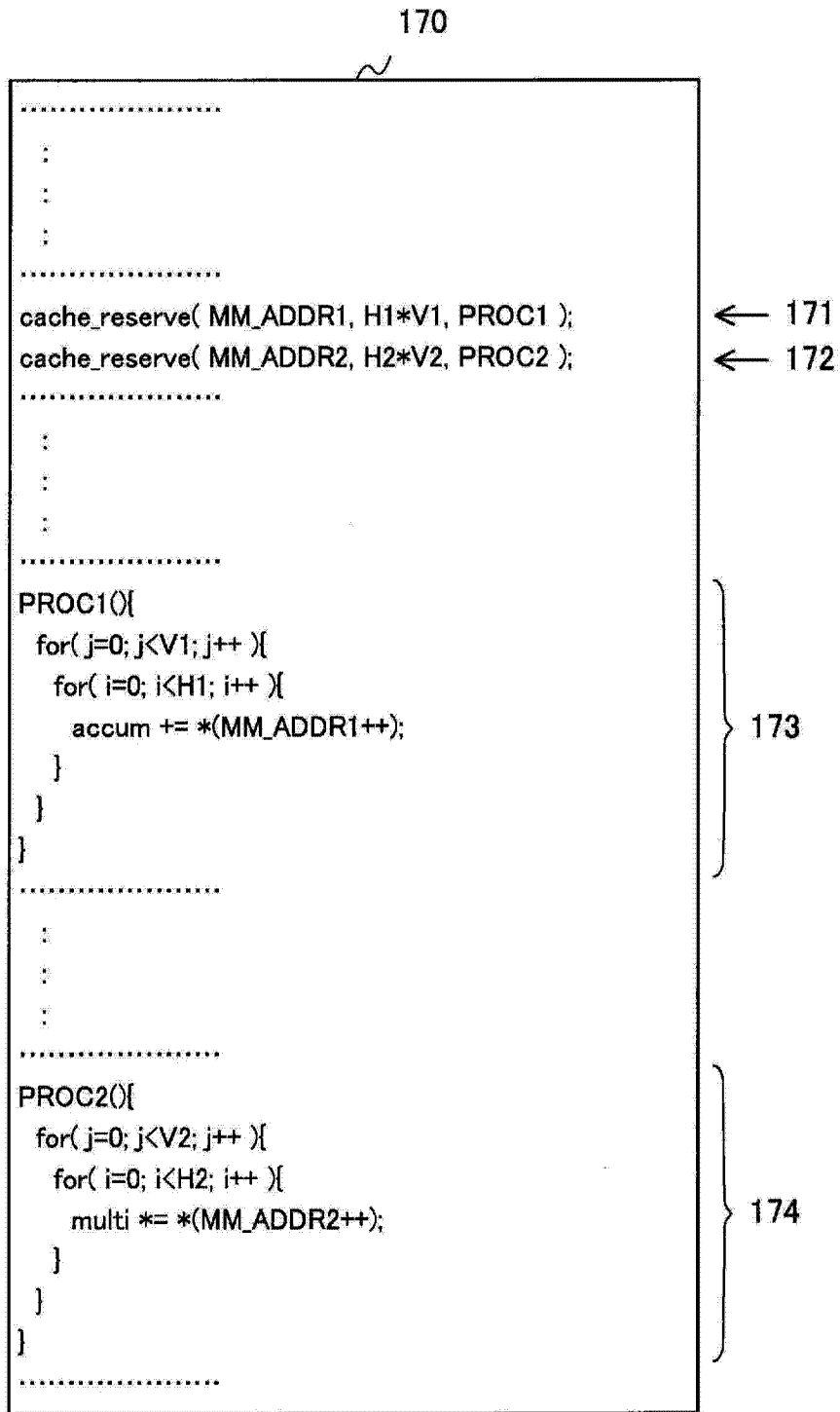


图 3

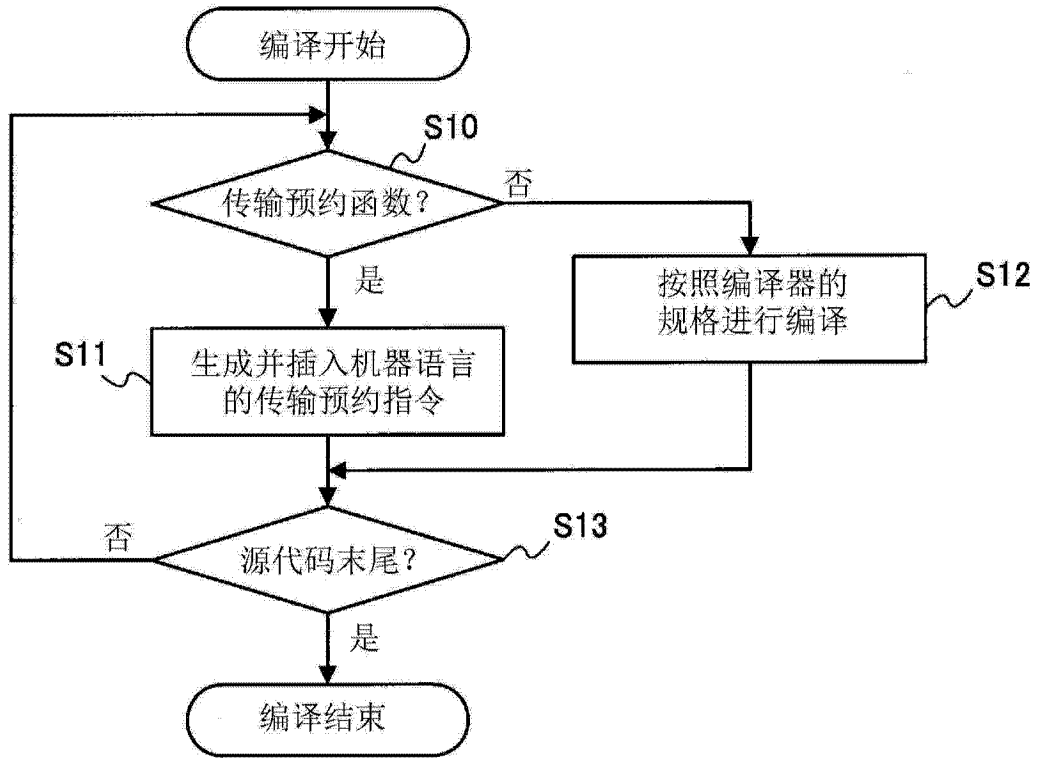


图 4

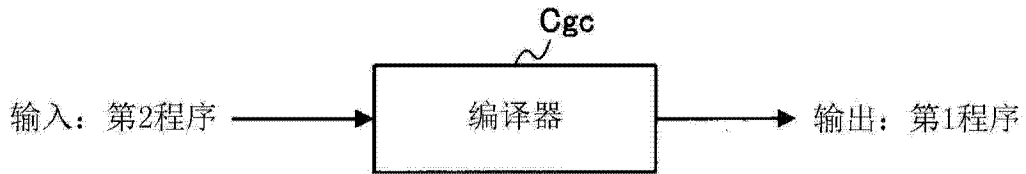


图 5

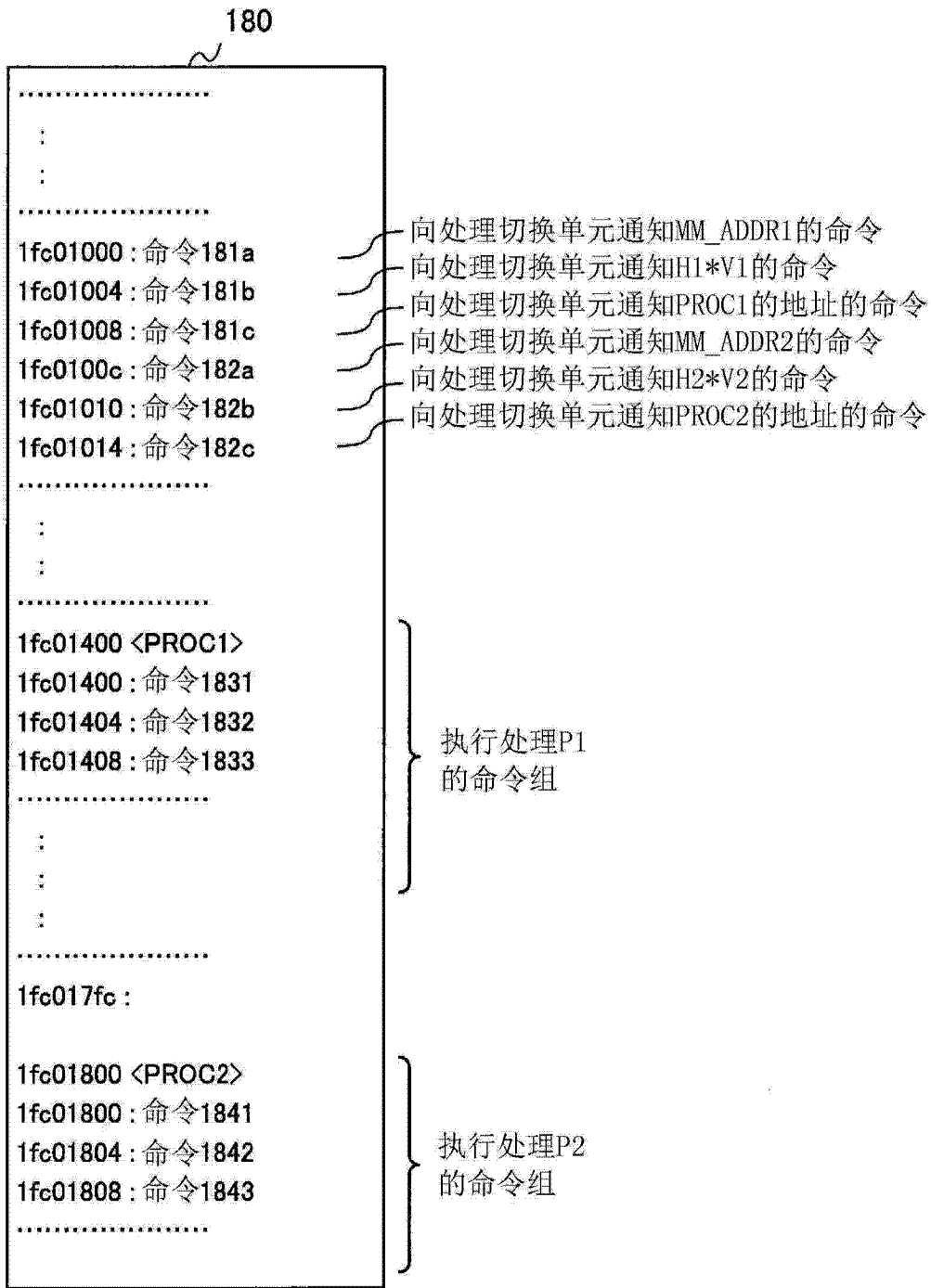


图 6

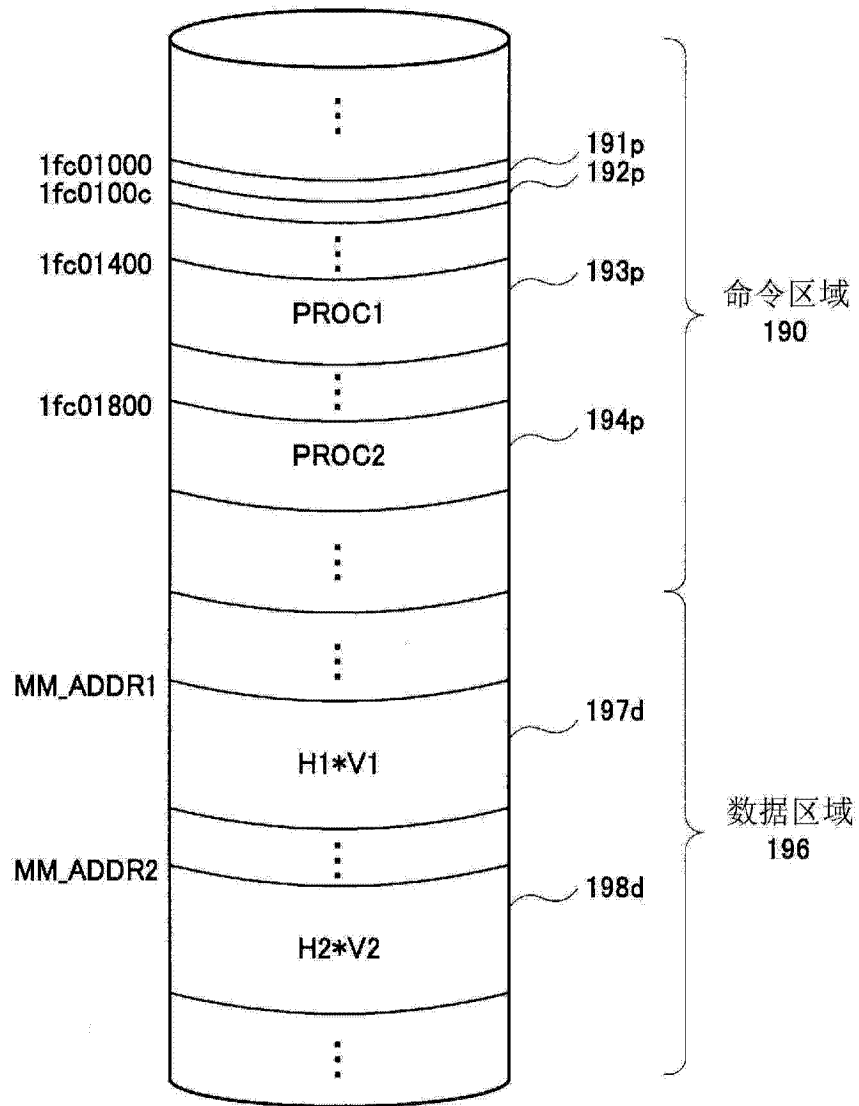


图 7

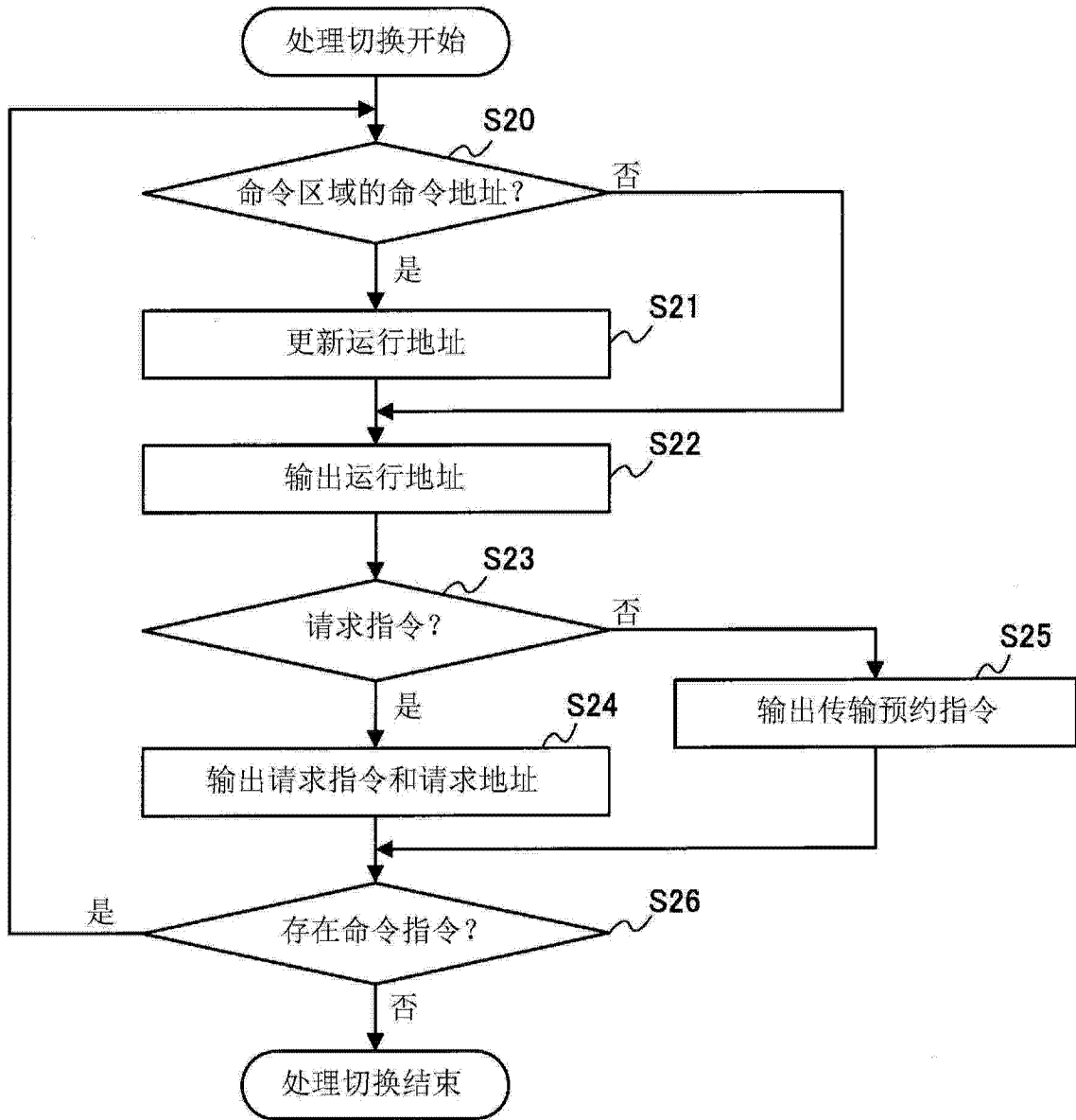


图 8

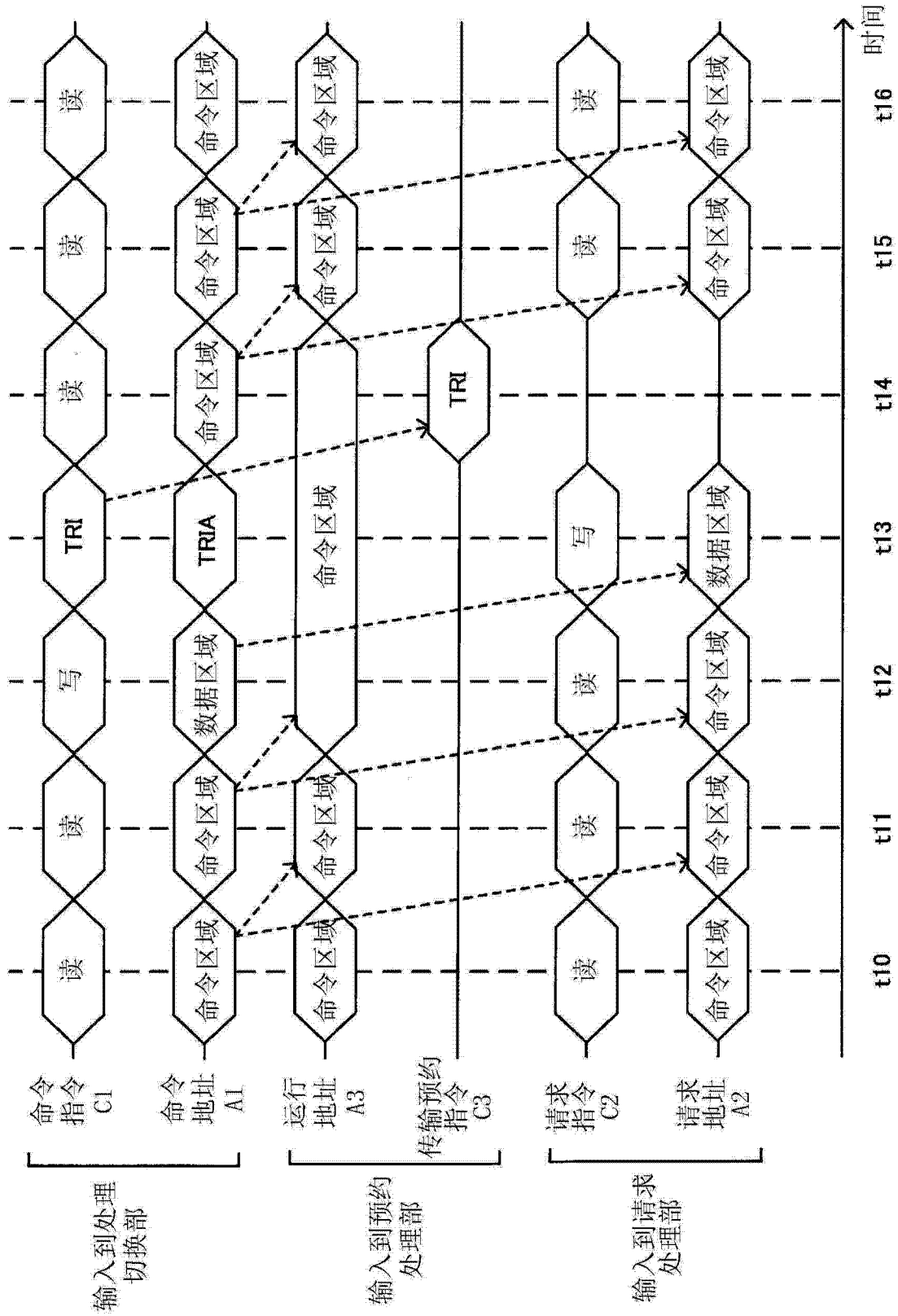


图 9

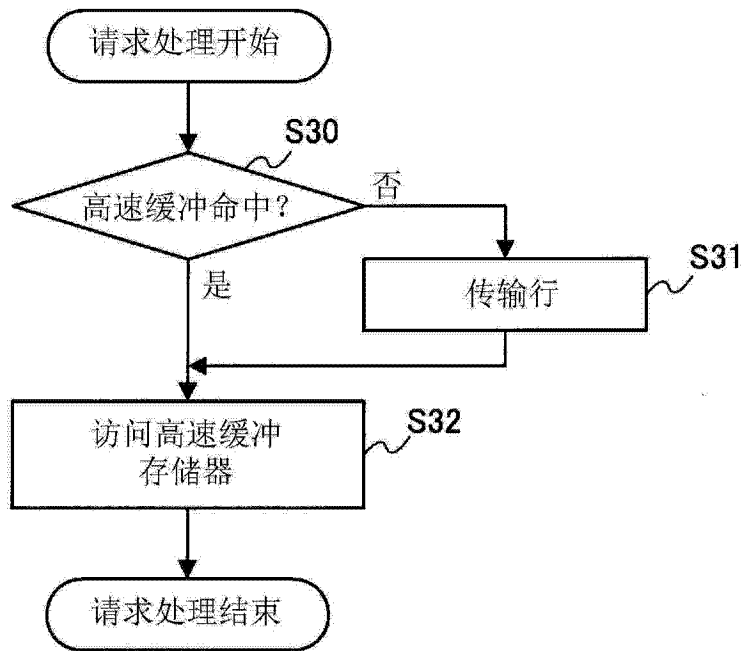


图 10

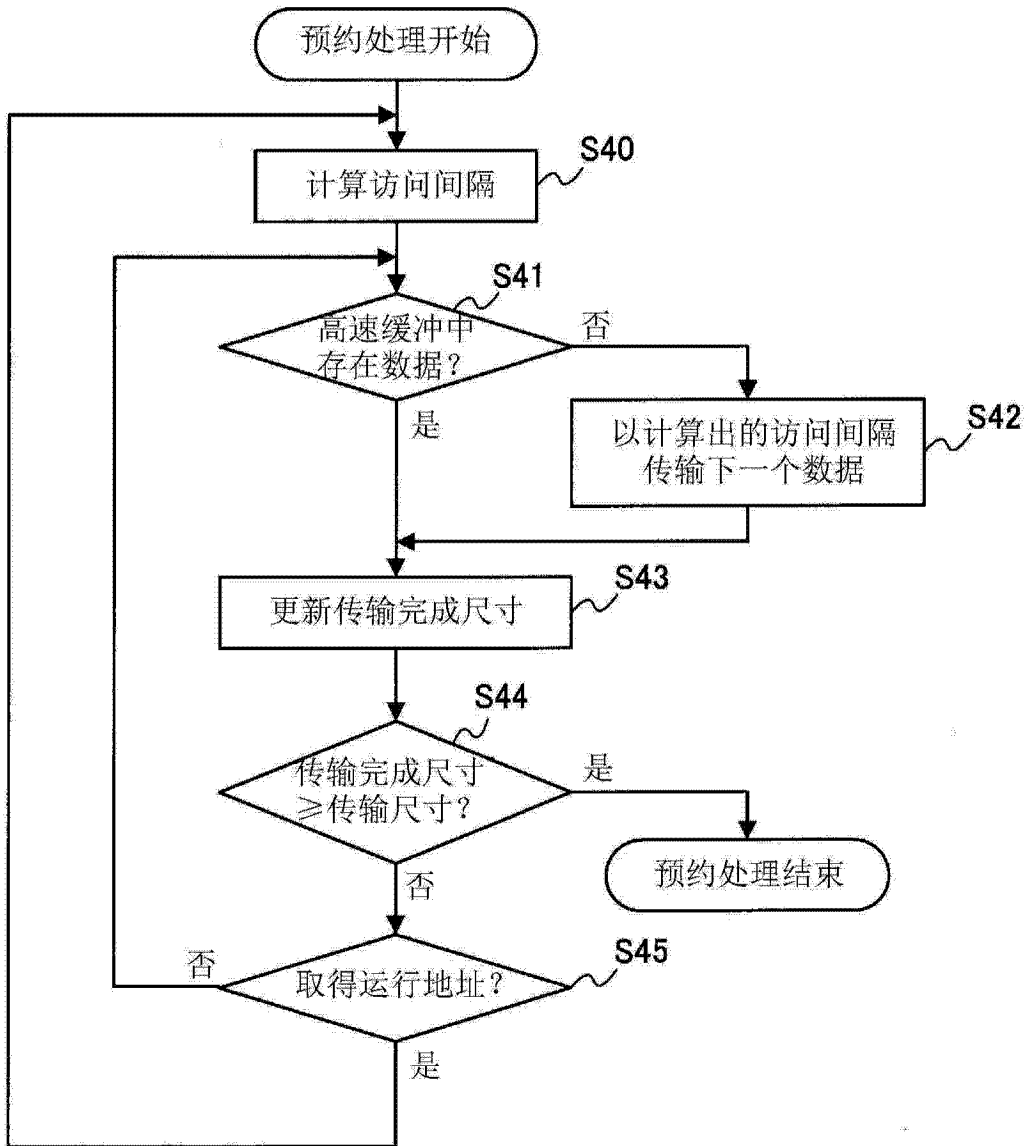


图 11

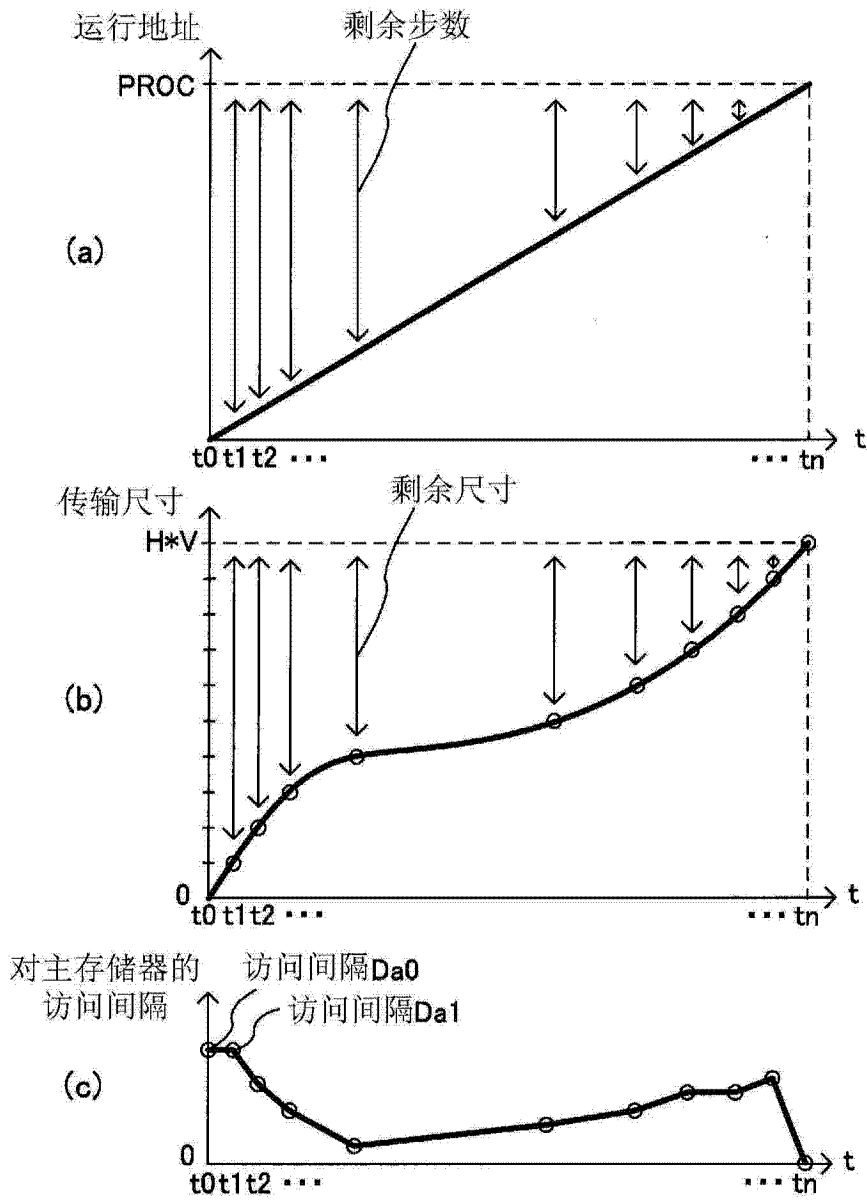


图 12

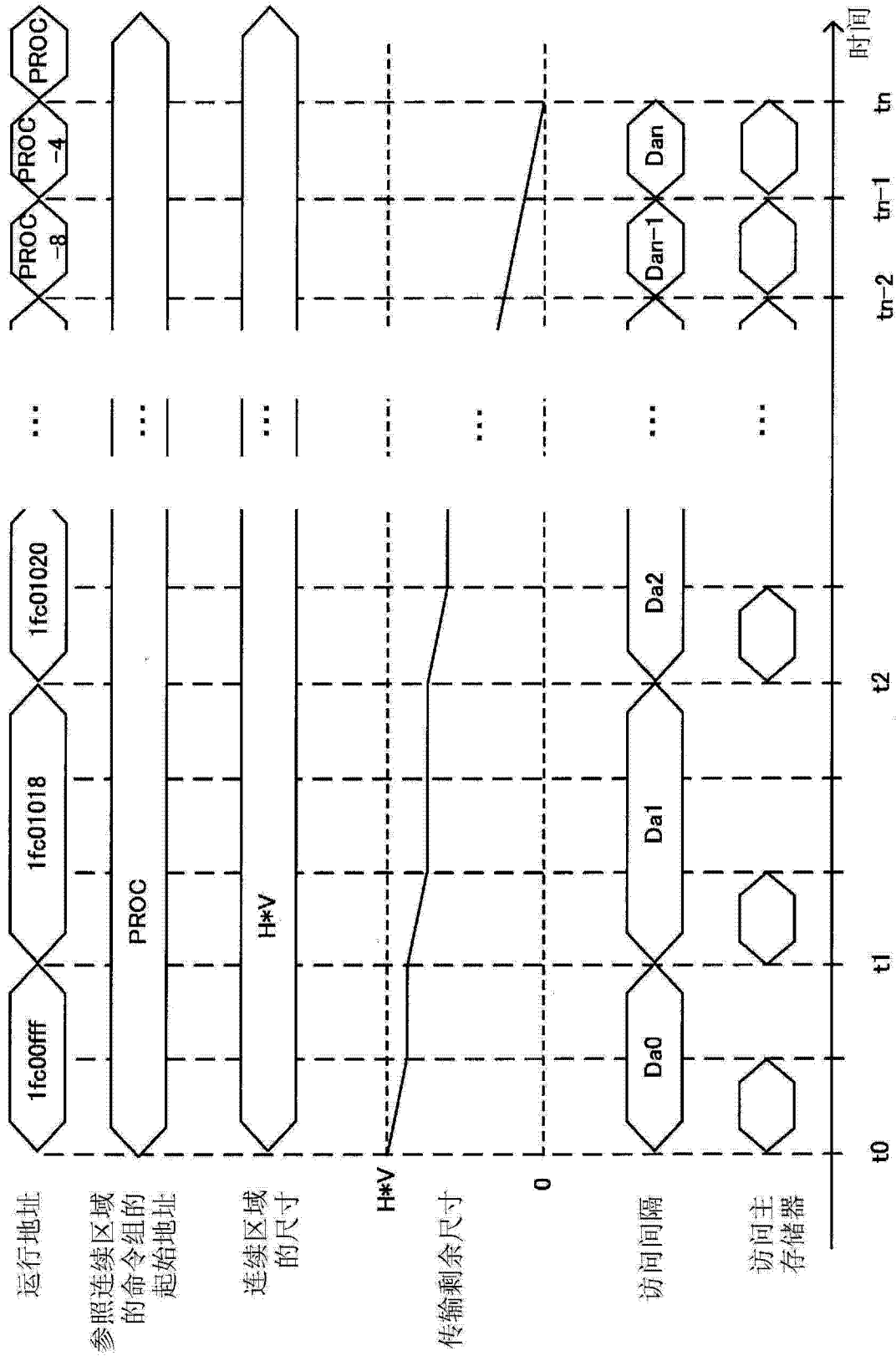


图 13

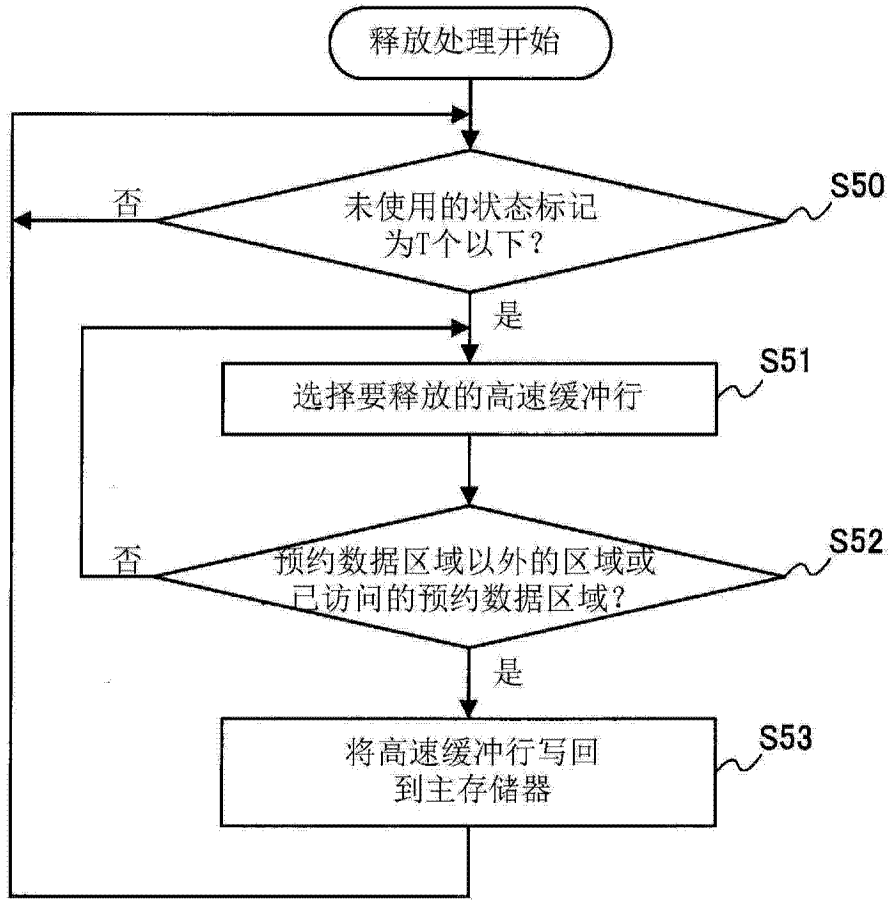


图 14

270

```
.....  
:  
:  
:  
.....  
:  
:  
:  
.....  
PROC1()  
  for(j=0; j<V1; j++)  
    for(i=0; i<H1; i++)  
      accum += *(MM_ADDR1++);  
    }  
  }  
}  
.....  
:  
:  
:  
.....  
PROC2()  
  for(j=0; j<V2; j++)  
    for(i=0; i<H2; i++)  
      multi *= *(MM_ADDR2++);  
    }  
  }  
}  
.....
```

图 15

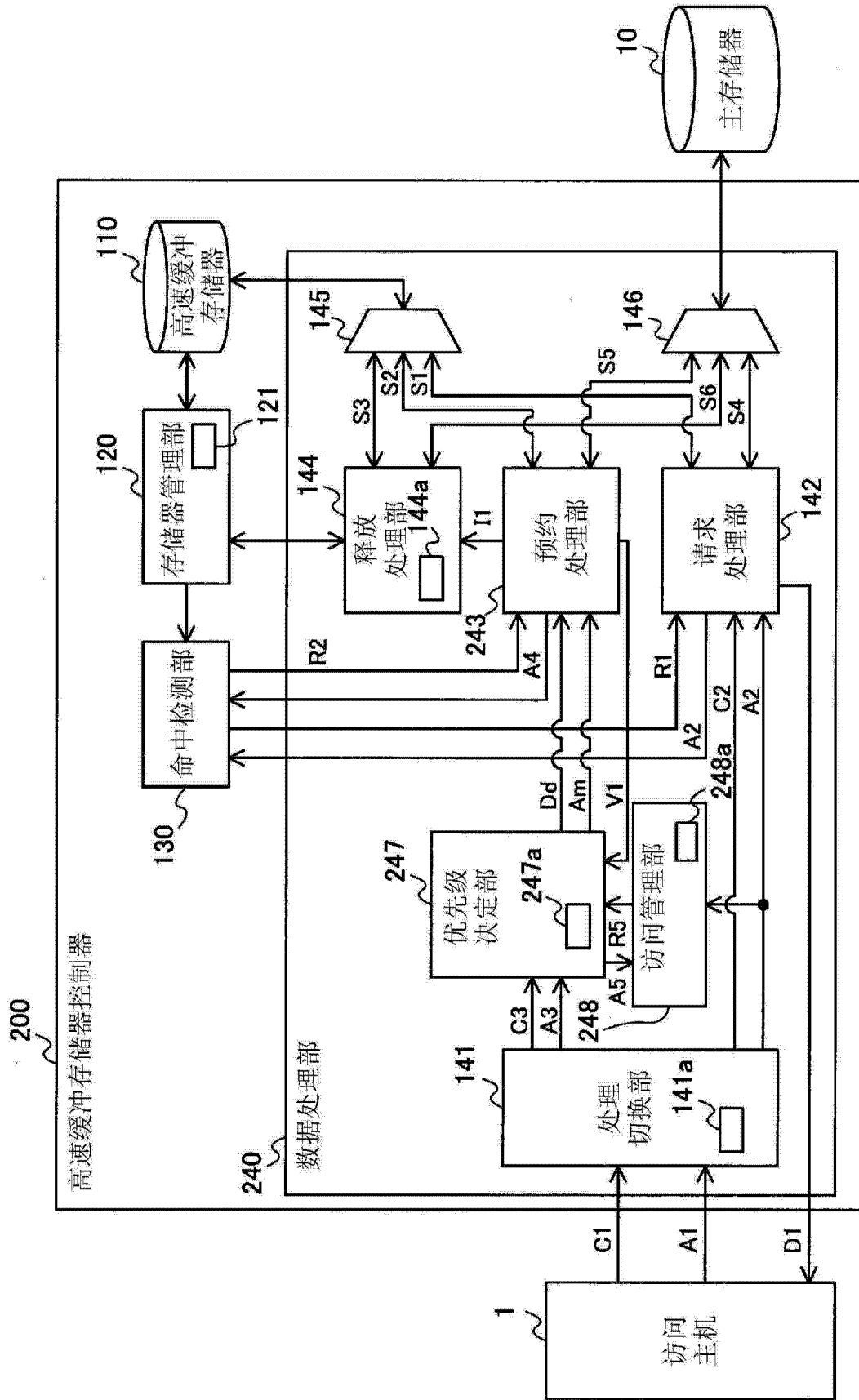


图 16

201

201a	201b	201c	201d	201e
到达顺序	参照命令组 起始地址	连续区域 起始地址	传输剩余尺寸	传输状况
1	PROC1	MM_ADDR1	H1*V1	1(传输中)
2	PROC2	MM_ADDR2	H2*V2	0(传输等待中)
3	PROC3	MM_ADDR3	H3*V3	0(传输等待中)
∴	∴	∴	∴	∴

图 17

202

202a 段序号	202b 行地址	202c 列地址	202d 访问经过时间Td
1	RAS#1	CAS#1	200
1	RAS#2	CAS#1	1062
1	RAS#3	CAS#1	896
⋮	⋮	⋮	⋮

图 18

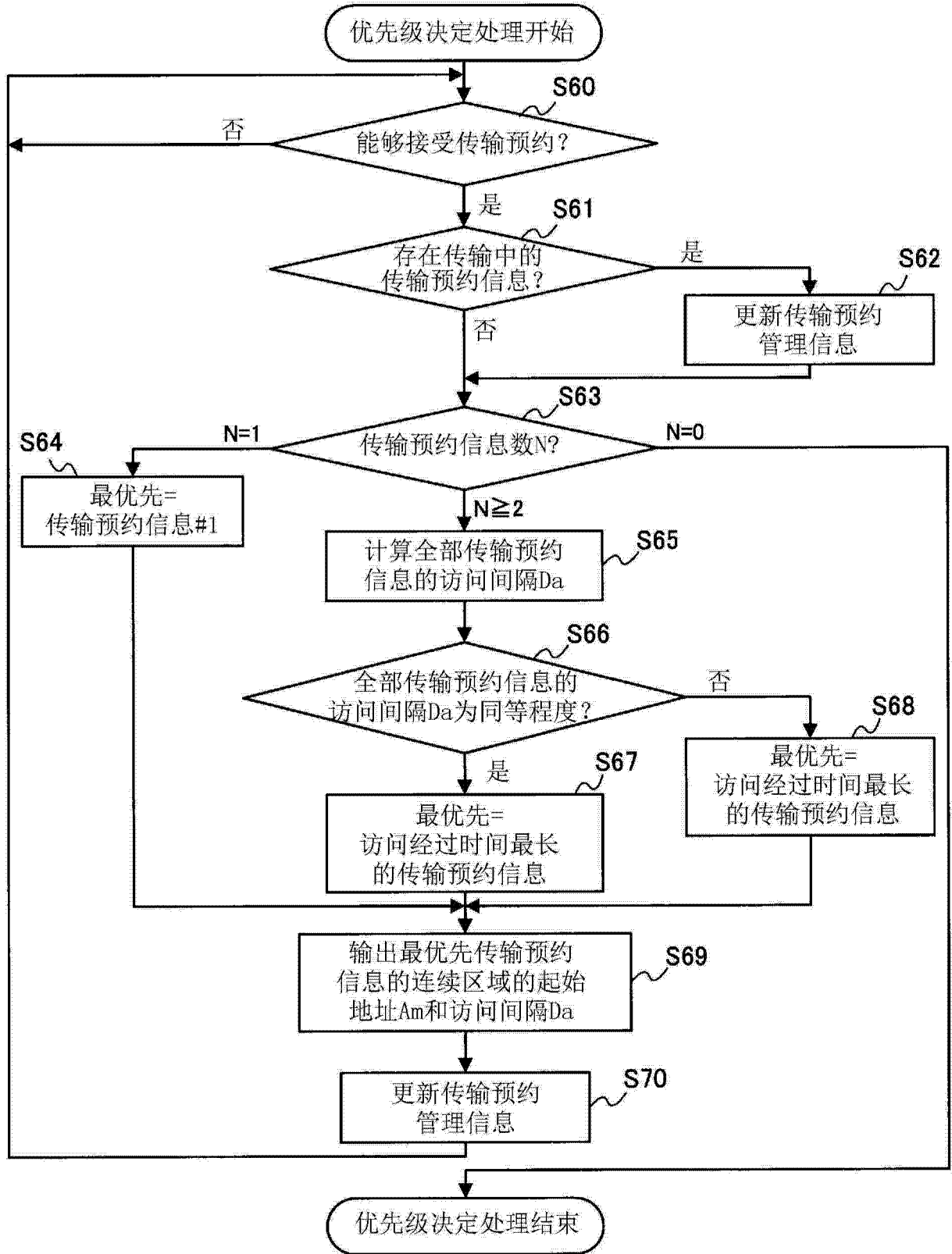


图 19

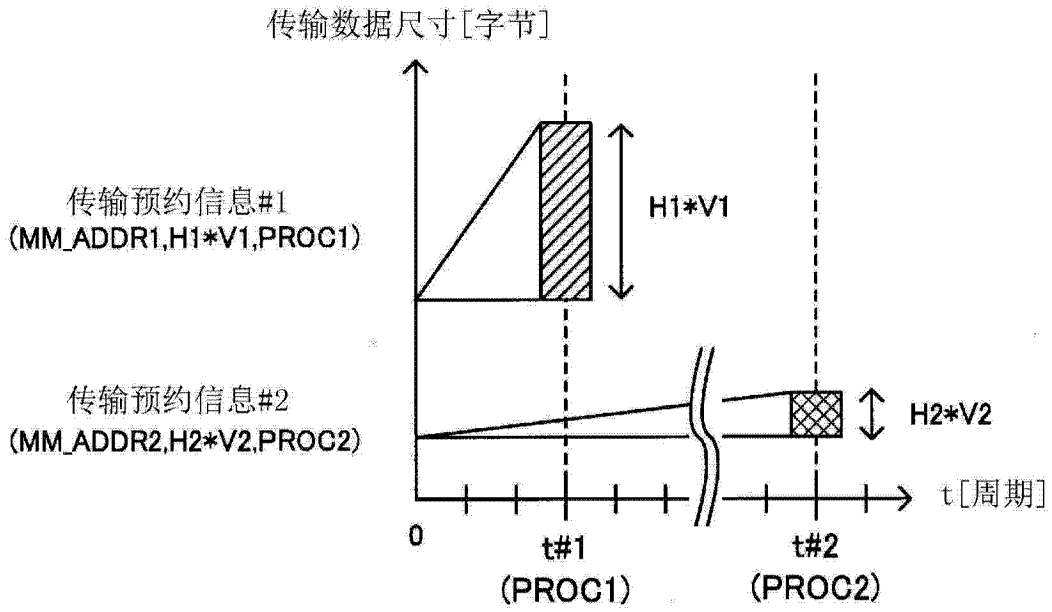


图 20

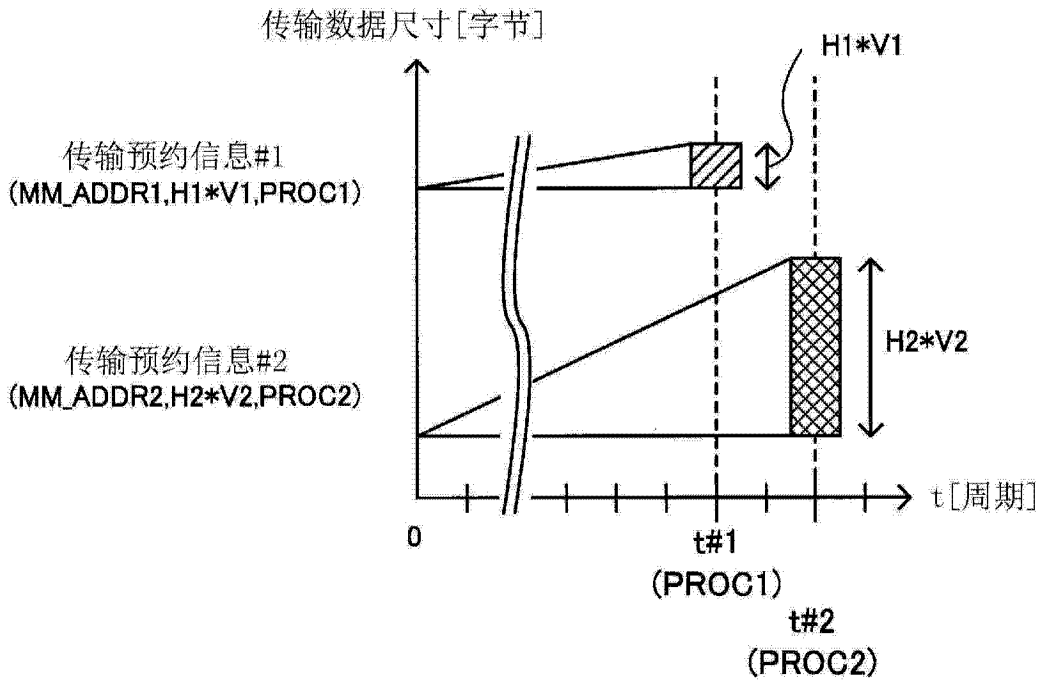


图 21

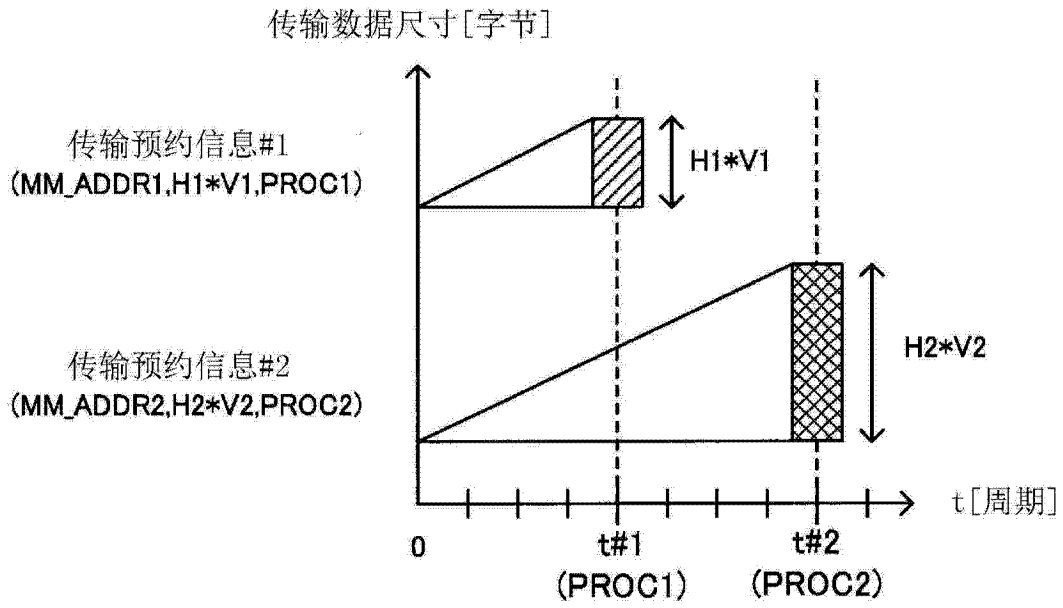


图 22

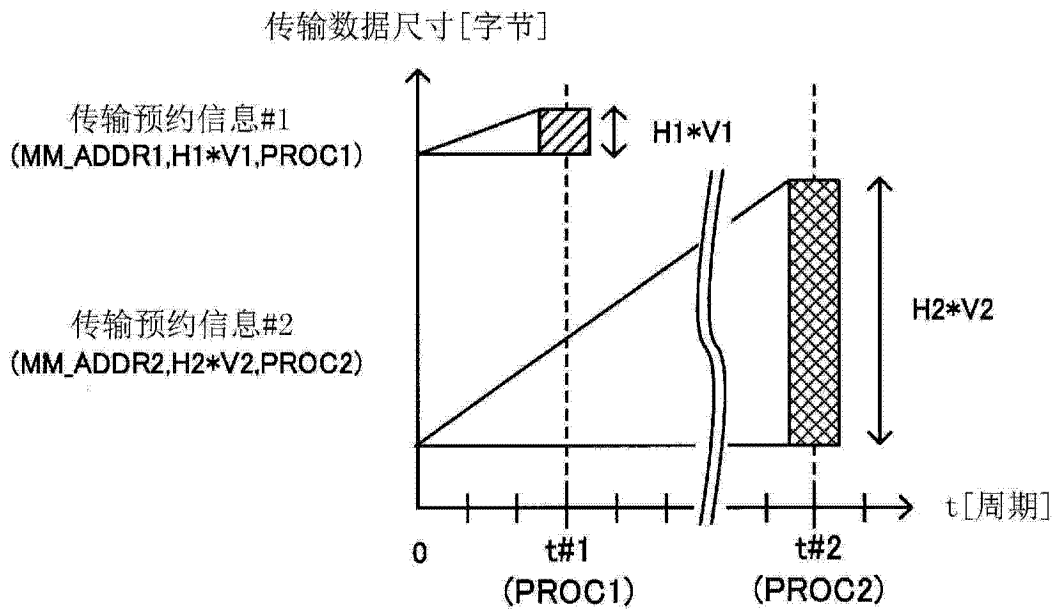


图 23

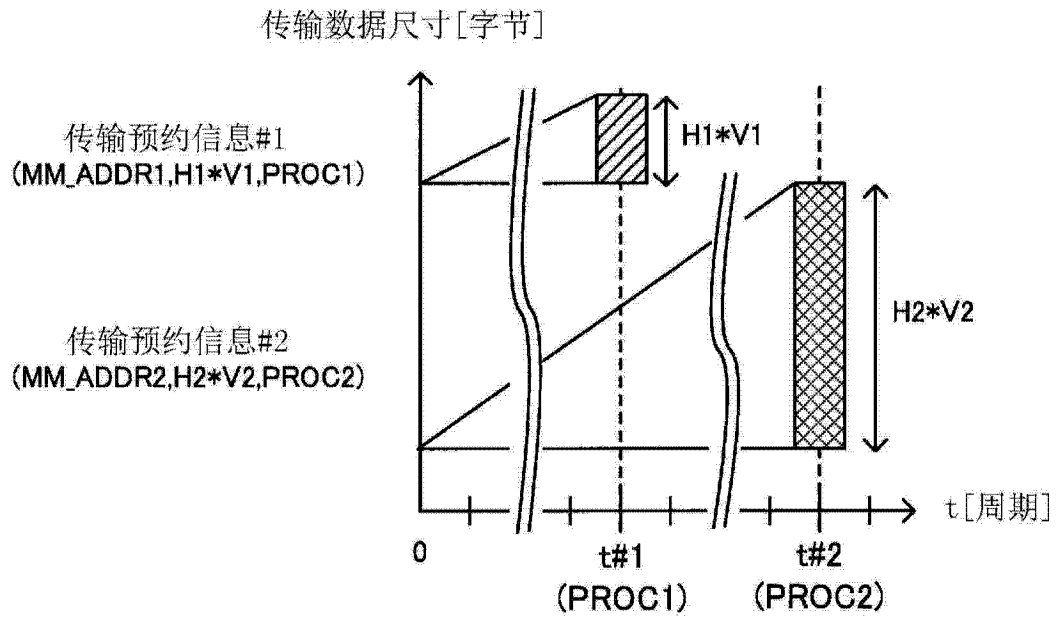


图 24

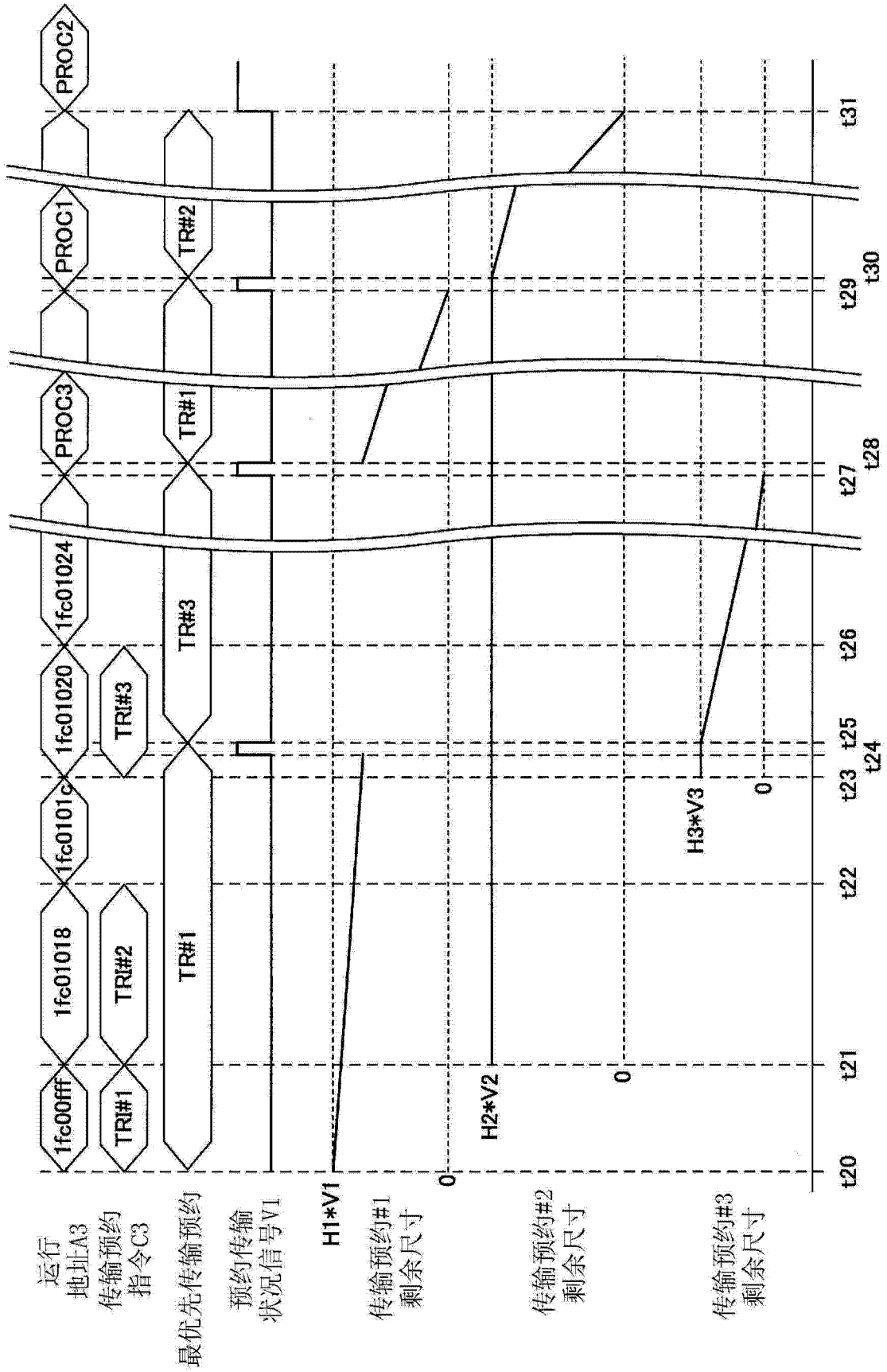


图 25

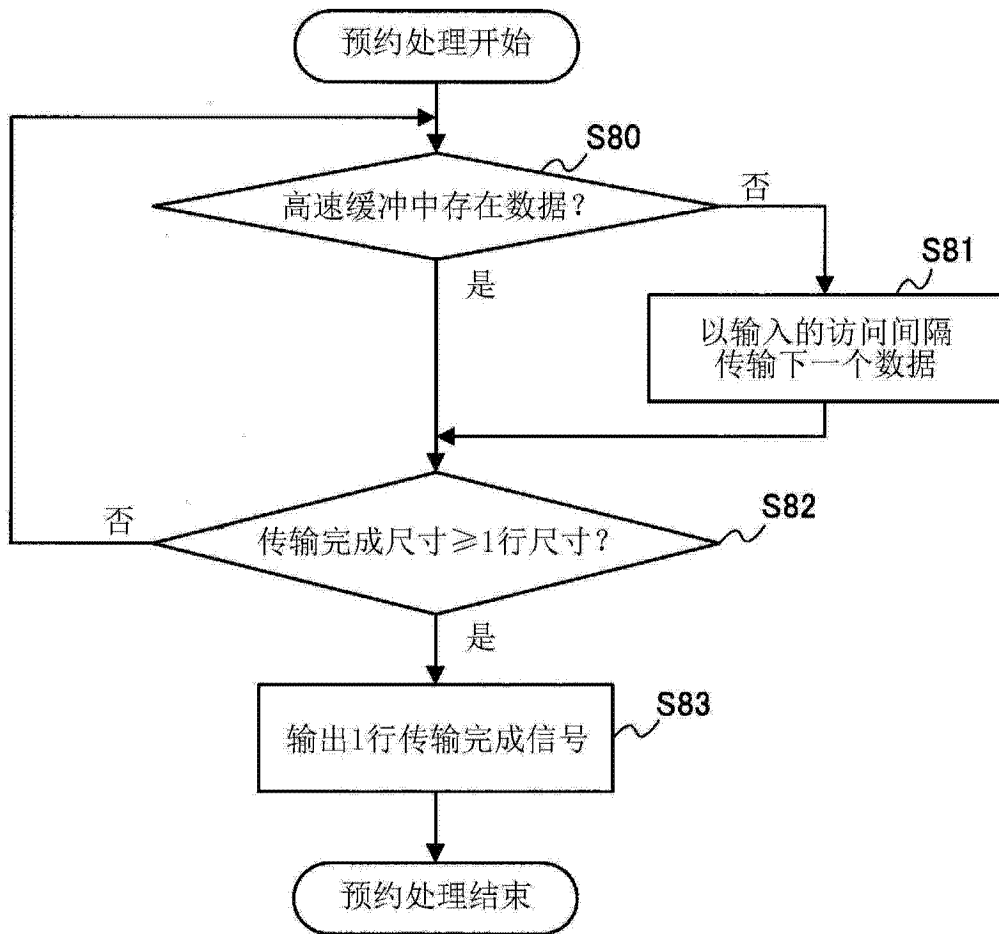


图 26

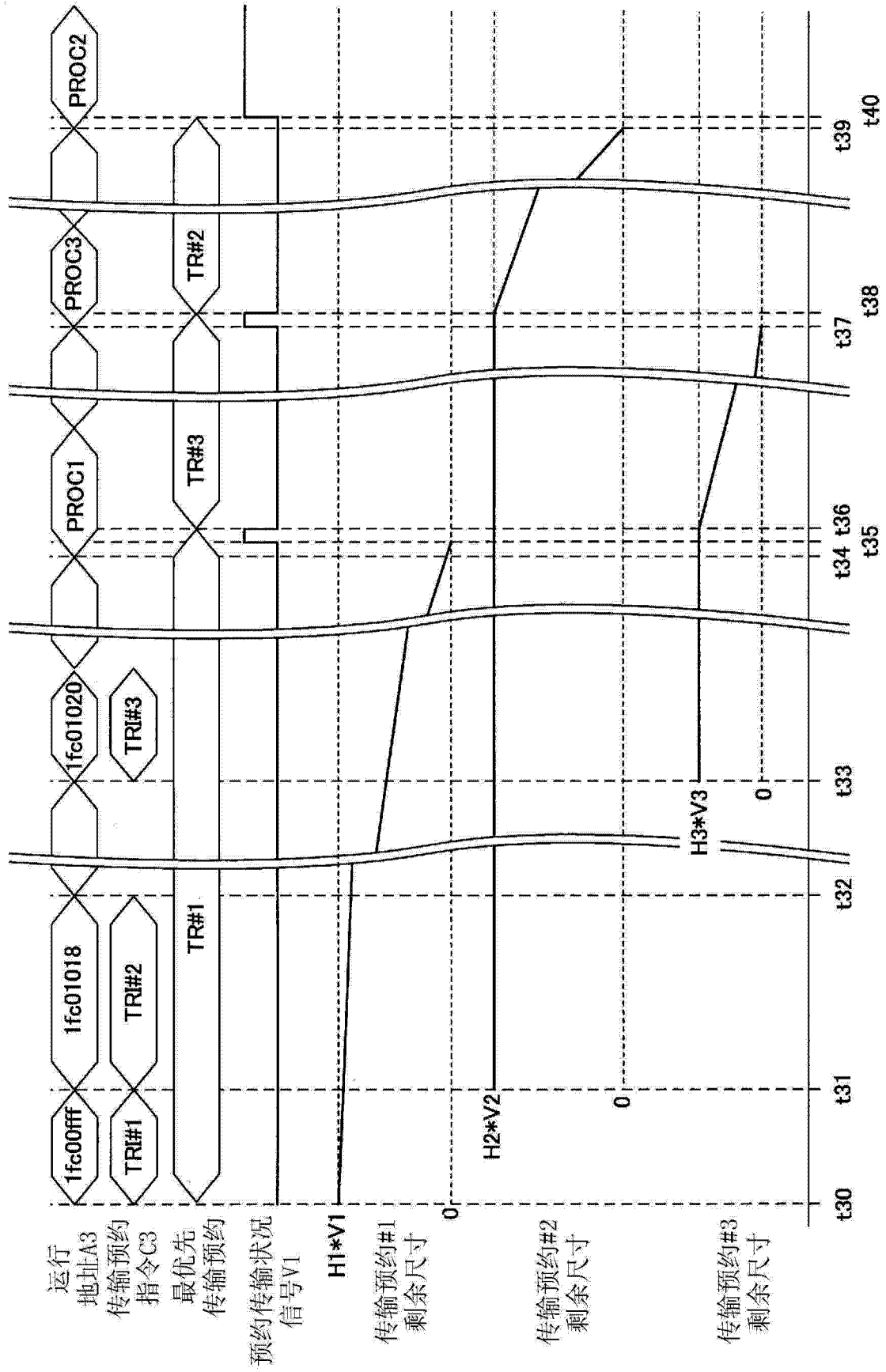


图 27

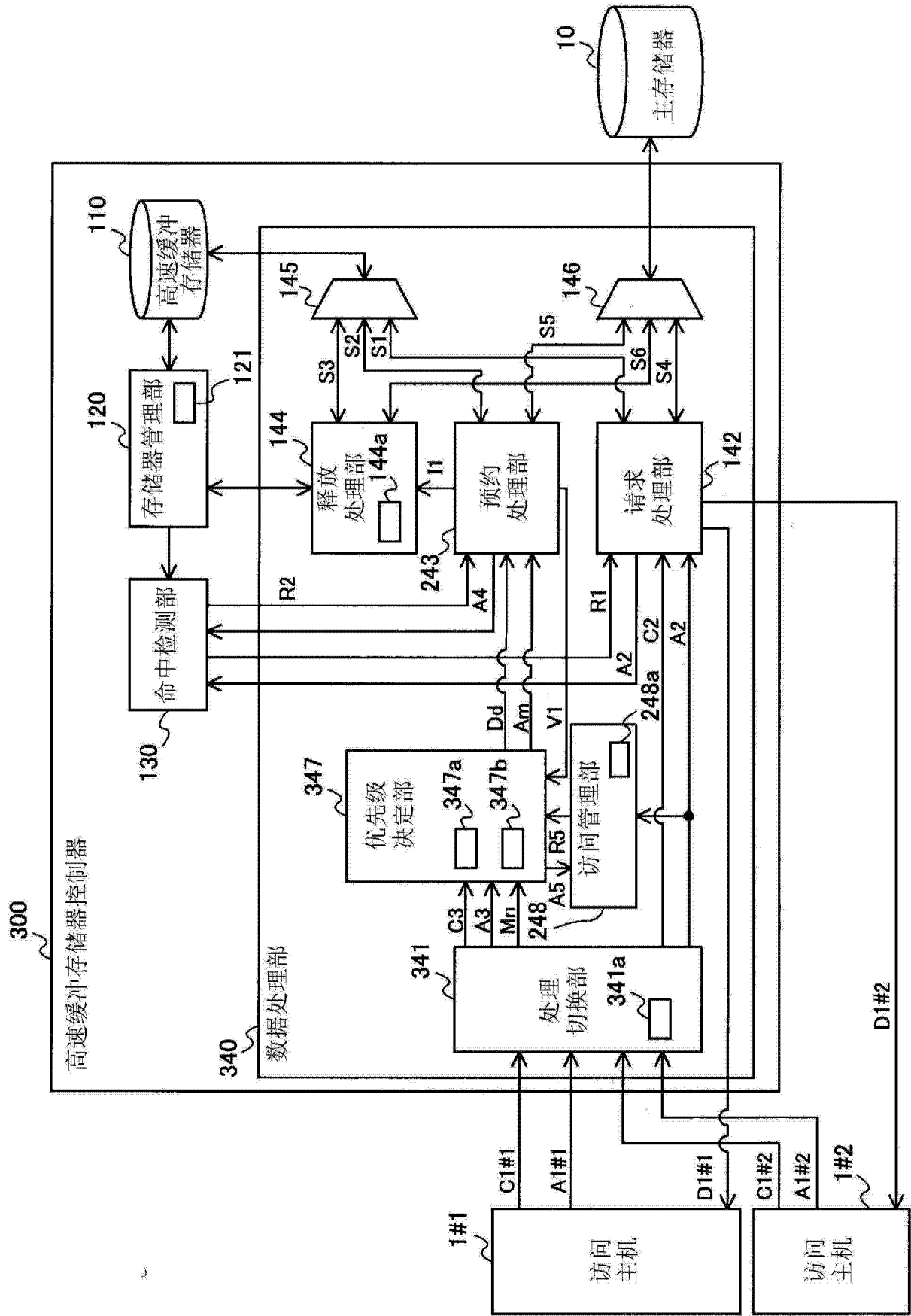


图 28

301

301a	301f	301b	301c	301d	301e
到达顺序	访问主机 编号	参照命令组 起始地址	连续区域 起始地址	传输剩余尺寸	传输状况
1	1	PROC1	MM_ADDR1	H1*V1	1(传输中)
2	1	PROC2	MM_ADDR2	H2*V2	0(传输等待中)
3	2	PROC3	MM_ADDR3	H3*V3	0(传输等待中)
⋮	⋮	⋮	⋮	⋮	⋮

图 29

The diagram shows a table with two columns and four rows. The first column is labeled '303a' and contains '访问主机编号' (Access Host Number), '1', '2', and a vertical ellipsis. The second column is labeled '303b' and contains '运行地址' (Execution Address), '1fc00ffff', '0fc0118', and a vertical ellipsis. A bracket labeled '303' encompasses the entire table.

访问主机编号	运行地址
1	1fc00ffff
2	0fc0118
⋮	⋮

图 30