



(12) 发明专利申请

(10) 申请公布号 CN 114830100 A

(43) 申请公布日 2022. 07. 29

(21) 申请号 202080088074.9

(22) 申请日 2020.11.20

(30) 优先权数据

16/718,162 2019.12.17 US

(85) PCT国际申请进入国家阶段日

2022.06.17

(86) PCT国际申请的申请数据

PCT/US2020/061672 2020.11.20

(87) PCT国际申请的公布数据

W02021/126471 EN 2021.06.24

(71) 申请人 超威半导体公司

地址 美国加利福尼亚州

(72) 发明人 保罗·莫耶

(74) 专利代理机构 上海胜康律师事务所 31263

专利代理师 樊英如 张静

(51) Int.Cl.

G06F 12/0862 (2006.01)

G06F 12/0897 (2006.01)

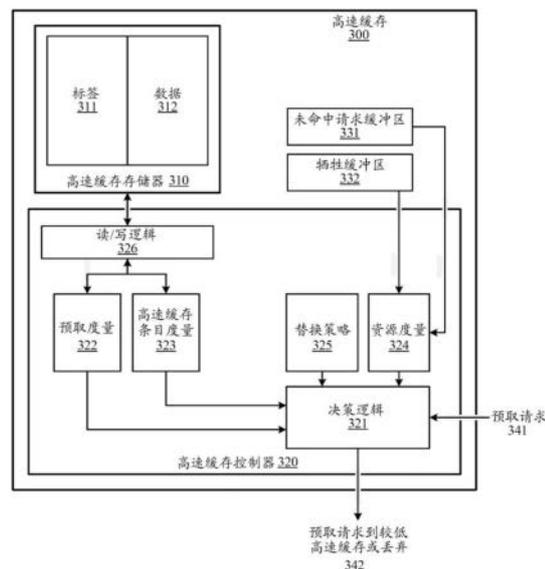
权利要求书4页 说明书12页 附图6页

(54) 发明名称

预取级别降级

(57) 摘要

一种方法包括:记录目标高速缓存的第一高速缓存性能度量集;对于在所述目标高速缓存处接收的多个预取请求中的每个预取请求,基于所述第一高速缓存性能度量集确定所述预取请求相对于所述目标高速缓存的阈值优先级级别的相对优先级;对于所述多个预取请求中的每个低优先级预取请求,响应于确定所述低优先级预取请求的优先级小于所述目标高速缓存的所述阈值优先级级别,将所述低优先级预取请求重定向到第一较低级别高速缓存;以及对于所述多个预取请求中的每个高优先级预取请求,响应于确定所述高优先级预取请求的优先级大于所述目标高速缓存的所述阈值优先级级别,根据所述高优先级预取请求将预取数据存储在所述目标高速缓存中。



1. 一种方法,其包括:

记录目标高速缓存的第一高速缓存性能度量集;

对于在所述目标高速缓存处接收的多个预取请求中的每个预取请求,基于所述第一高速缓存性能度量集确定所述预取请求相对于所述目标高速缓存的阈值优先级级别的相对优先级;

对于所述多个预取请求的第一子集中优先级不超过所述阈值优先级级别的每个低优先级预取请求,响应于确定所述低优先级预取请求的所述优先级不超过所述阈值优先级级别,将所述低优先级预取请求重定向到第一较低级别高速缓存;以及

对于所述多个预取请求的第二子集中优先级超过所述阈值优先级级别的每个高优先级预取请求,响应于确定所述高优先级预取请求的所述优先级超过所述阈值优先级级别,根据所述高优先级预取请求将预取数据存储存储在所述目标高速缓存中。

2. 如权利要求1所述的方法,其还包括,对于所述低优先级预取请求中的每一者:

从所述目标高速缓存的高速缓存层次结构中选择另一高速缓存作为所述第一较低级别高速缓存,其中所述第一较低级别高速缓存具有比所述目标高速缓存大的容量;以及

根据所述低优先级预取请求,将预取数据存储存储在所述第一较低级别高速缓存中。

3. 如权利要求1所述的方法,其还包括:

对于所述第一子集的一个或多个预取请求,基于所述第一较低级别高速缓存的第二高速缓存性能度量集,将所述一个或多个预取请求从所述第一较低级别高速缓存重定向到第二较低级别高速缓存,其中所述第二较低级别高速缓存具有比所述第一较低级别高速缓存高的容量。

4. 如权利要求1所述的方法,其还包括对于所述第一子集中的一个或多个预取请求:

将所述一个或多个预取请求重定向到所述目标高速缓存的所述高速缓存层次结构中的最低级别高速缓存;以及

响应于确定所述一个或多个预取请求的所述优先级小于所述最低级别高速缓存的阈值优先级级别,丢弃所述一个或多个预取请求。

5. 如权利要求1所述的方法,其还包括:

对于所述多个预取请求中的每个预取请求,基于预取准确性度量来确定所述预取请求的优先级,其中所述预取准确性度量是对于所述目标高速缓存的预取条目集,基于未使用的预取条目与已使用的预取条目的比例来确定的。

6. 如权利要求1所述的方法,其还包括:

对于所述多个预取请求中的每个预取请求,基于所述预取请求的来源来确定所述预取请求的优先级,其中所述来源包括硬件预取器和用户应用程序中的一者。

7. 如权利要求1所述的方法,其还包括基于以下项确定所述目标高速缓存的所述阈值优先级级别:

高速缓存替换策略,以及

对于所述目标高速缓存中的多个高速缓存条目中的每个高速缓存条目,所述高速缓存条目的访问频率和与所述高速缓存条目相关联的操作类型。

8. 如权利要求1所述的方法,其中所述第一高速缓存性能度量集包括所述目标高速缓存的牺牲缓冲区的牺牲缓冲区占用率度量和所述目标高速缓存的未命中请求缓冲区的未

命中请求缓冲区占用率度量。

9. 一种计算装置,其包括:

监视电路,所述监视电路被配置为记录目标高速缓存的第一高速缓存性能度量集;

第一决策逻辑电路,所述第一决策逻辑电路与所述监视电路耦合并且被配置为:

对于在所述目标高速缓存处接收的多个预取请求中的每个预取请求,基于所述第一高速缓存性能度量集确定所述预取请求相对于所述目标高速缓存的阈值优先级级别的相对优先级;

对于所述多个预取请求的第一子集中优先级不超过所述阈值优先级级别的每个低优先级预取请求,响应于确定所述低优先级预取请求的所述优先级不超过所述阈值优先级级别,将所述低优先级预取请求重定向到第一较低级别高速缓存;并且

对于所述多个预取请求的第二子集中优先级超过所述阈值优先级级别的每个高优先级预取请求,响应于确定所述高优先级预取请求的所述优先级超过所述阈值优先级级别,根据所述高优先级预取请求将预取数据存储在上述目标高速缓存中。

10. 如权利要求9所述的计算装置,其中:

所述第一决策逻辑电路被进一步配置为对于所述低优先级预取请求中的每一者,从所述目标高速缓存的高速缓存层次结构中选择另一高速缓存作为所述第一较低级别高速缓存,

所述第一较低级别高速缓存具有比所述目标高速缓存大的容量;并且

所述第一较低级别高速缓存被配置为对于所述低优先级预取请求中的每一者,根据所述低优先级预取请求来存储预取数据。

11. 如权利要求9所述的计算装置,其还包括:

第二决策逻辑电路,所述第二决策逻辑电路在所述第一较低级别高速缓存中;以及

第二较低级别高速缓存,所述第二较低级别高速缓存与所述第二决策逻辑电路耦合并且具有比所述第一较低级别高速缓存高的容量,其中:

所述第二决策逻辑电路被配置为对于所述第一子集中的一个或多个预取请求,基于所述第一较低级别高速缓存的第二高速缓存性能度量集,将所述一个或多个预取请求从所述第一较低级别高速缓存重定向到所述第二较低级别高速缓存。

12. 如权利要求9所述的计算装置,其还包括:

最低级别高速缓存,所述最低级别高速缓存在所述目标高速缓存的所述高速缓存层次结构中;以及

第二决策逻辑电路,所述第二决策逻辑电路被配置为对于所述第一子集中的一个或多个预取请求,将所述一个或多个预取请求重定向到所述最低级别高速缓存;以及

第三决策逻辑电路,所述第三决策逻辑电路在所述最低级别高速缓存中,所述第三决策逻辑电路被配置为响应于确定所述一个或多个预取请求的所述优先级小于所述最低级别高速缓存的阈值优先级级别,丢弃所述一个或多个预取请求。

13. 如权利要求9所述的计算装置,其还包括:

预取度量模块,所述预取度量模块与所述第一决策逻辑耦合并且被配置为对于所述目标高速缓存的预取条目集,基于未使用的预取条目与已使用的预取条目的比例来确定预取准确性度量,

其中所述第一决策逻辑被进一步配置为对于所述多个预取请求中的每个预取请求,基于所述预取准确性度量来确定所述预取请求的优先级。

14. 如权利要求9所述的计算装置,其还包括:

硬件预取器;以及

处理器,所述处理器被配置为基于应用程序的执行指令生成所述多个预取请求中的一者或多者,其中所述决策逻辑被进一步配置为对于所述多个预取请求中的每个预取请求,基于所述预取请求的来源来确定所述预取请求的优先级,其中所述来源包括所述硬件预取器和所述处理器中的一者。

15. 如权利要求9所述的计算装置,其还包括:

高速缓存条目度量模块,所述高速缓存条目度量模块被配置为记录高速缓存条目度量,所述高速缓存条目度量对于所述目标高速缓存中的多个高速缓存条目中的每个高速缓存条目,包括所述高速缓存条目的访问频率和与所述高速缓存条目相关联的操作类型,其中所述决策逻辑被进一步配置为基于高速缓存替换策略和所述高速缓存条目度量来确定所述阈值优先级级别。

16. 一种计算系统,其包括:

处理单元,所述处理单元被配置为执行应用程序;

多个高速缓存,所述多个高速缓存在与所述处理单元耦合的高速缓存层次结构中;以及

高速缓存控制器,所述高速缓存控制器与所述多个高速缓存耦合,所述高速缓存控制器包括:

监视电路,所述监视电路被配置为记录目标高速缓存的第一高速缓存性能度量集,以及

决策逻辑电路,所述决策逻辑电路与所述监视电路耦合并且被配置为:

对于在所述目标高速缓存处接收的多个预取请求中的每个预取请求,基于所述第一高速缓存性能度量集确定所述预取请求相对于所述目标高速缓存的阈值优先级级别的相对优先级;

对于所述多个预取请求的第一子集中优先级不超过所述阈值优先级级别的每个低优先级预取请求,响应于确定所述低优先级预取请求的所述优先级不超过所述阈值优先级级别,将所述低优先级预取请求重定向到第一较低级别高速缓存;并且

对于所述多个预取请求的第二子集中优先级超过所述阈值优先级级别的每个高优先级预取请求,响应于确定所述高优先级预取请求的所述优先级超过所述阈值优先级级别,根据所述高优先级预取请求将预取数据存储在与所述目标高速缓存中。

17. 如权利要求16所述的计算系统,其还包括:

存储器控制器,所述存储器控制器与所述高速缓存控制器耦合,其中所述决策逻辑电路被进一步配置为:

对于所述第一子集中的一个或多个预取请求,将所述一个或多个预取请求重定向到所述目标高速缓存的所述高速缓存层次结构中的最低级别高速缓存,

响应于确定所述一个或多个预取请求的所述优先级小于所述最低级别高速缓存的阈值优先级级别,丢弃所述一个或多个预取请求,以及

将所述预取请求的指示传输到存储器控制器;并且

所述存储器控制器被配置为响应于所述指示,初始化对包含所述预取数据的所述存储器的访问。

18. 如权利要求16所述的计算系统,其还包括:

硬件预取器,所述硬件预取器与所述高速缓存层次结构耦合并且被配置为基于以下项为所述应用程序生成所述多个预取请求中的一者或多者:

为所述应用程序执行分支预测,以及

基于所述应用程序的先前存储器访问模式来预测存储器访问。

19. 如权利要求16所述的计算系统,其中:

所述处理单元被进一步配置为根据所述应用程序的指令生成所述多个预取请求中的一者或多者。

20. 如权利要求16所述的计算系统,其还包括:

多个高速缓存控制器,所述多个高速缓存控制器包括所述高速缓存控制器,其中所述多个高速缓存控制器中的每一者被配置为:

控制所述高速缓存层次结构中的所述多个高速缓存中的一者,并且

将所述第一子集中的所述低优先级预取请求中的一者或多者重定向到所述高速缓存层次结构中具有比所述多个高速缓存中的相关联一个高速缓存高的容量的另一高速缓存。

预取级别降级

背景技术

[0001] 现代计算系统中的处理器通常可能比存储处理器使用的指令或其他数据的主存储器运行得更快。因此,在许多情况下,结合主存储器使用更小且更快的高速缓存,以提供对指令或数据的快速访问。在实际需要数据之前,当处理器请求将数据存储到高速缓存中时,将数据预取到高速缓存中发生。然后,当需要数据时,可从高速缓存中检索数据,而不会导致从主存储器请求数据出现额外延迟。

[0002] 由于大多数程序是按顺序执行的或表现出其他规则的执行模式,因此可按程序顺序或根据存储器访问流中的其他识别模式来获取指令或其他数据。然而,预取不正确的数据或在不适当的时间预取数据可能会降低预取实现方式带来的整体益处。

附图说明

[0003] 在附图的图示中通过举例的方式而非限制的方式示出本公开。

[0004] 图1示出根据实施方案的计算系统。

[0005] 图2示出根据实施方案的计算系统中的存储器层次结构。

[0006] 图3示出根据实施方案的高速缓存的部件。

[0007] 图4示出根据实施方案的存储在高速缓存标签中的信息。

[0008] 图5示出根据实施方案的高速缓存层次结构中的预取的降级。

[0009] 图6是根据实施方案的示出预取过程的流程图。

具体实施方式

[0010] 以下描述阐述了许多具体细节,诸如具体系统、部件、方法等的示例,以便提供对实施方案的良好理解。然而,本领域的技术人员将明白,可在没有这些具体细节的情况下实践至少一些实施方案。在其他情况下,未详细描述公知的部件或方法,或以简单的框图格式呈现公知的部件或方法,以避免对实施方案不必要的混淆。因此,阐述的具体细节仅仅是示例性的。特定实现方式可与这些示例性细节不同,并且仍被认为在实施方案的范围内。

[0011] 在包括多个级别的高速缓存(例如,L1、L2和L3)的计算系统中,数据或指令的预取由硬件预取器或软件(诸如用户应用程序)以高速缓存级别中的特定一个为目标。例如,包括多个级别的高速缓存的计算系统还每个高速缓存级别包括硬件预取器,所述硬件预取器监视存储器访问流并确定从主存储器中提取哪些数据到其相关联高速缓存级别或较低级别(较高编号)高速缓存。此外,可根据指令(例如,如在x86指令集中提供)生成以给定高速缓存级别为目标的预取;由编译器使用启发式方法生成此类指令,以预测在运行时预取哪些项目。因此,硬件预取机制和软件预取机制两者都以特定级别的高速缓存为目标,所述特定级别的高速缓存是在不考虑目标高速缓存中资源的可用性的情况下选择的。

[0012] 在某些情况下,目标级别的高速缓存的资源被过度利用,并且预取更适合以较低高速缓存级别为目标。此外,预取到被硬件预取器或软件预取器当成目标的高速缓存级别并不始终导致所消耗的低级别高速缓存容量的量出现最低延迟;在某些情况下,预取到较

低级别(即较高编号)高速缓存可能会具有更好的容量/延迟影响,尤其是在大量预取被确定为不准确或执行得太早的情况下。由于对目标高速缓存级别的高速缓存行容量和资源可用性的压力增加,发送到目标不当高速缓存级别的预取可能会导致预取的高速缓存行或该级别的其他高速缓存行(例如,由于预取提前或不准确而被逐出的高速缓存行)的延迟增加。

[0013] 在一个实施方案中,高速缓存层次结构中的每个级别都包括高速缓存控制器,所述高速缓存控制器具有用于将预取降级到较低(即,更高编号和更高容量)高速缓存级别的逻辑。例如,当满足某些条件时,初始以L2高速缓存为目标的预取将降级到L3高速缓存,这指示预取应被赋予低于作为初始目标的L2高速缓存中的现有数据的优先级。

[0014] 在一个实施方案中,如果目标高速缓存的未命中请求缓冲区和/或牺牲缓冲区已满或接近满,或者特定高速缓存索引的未命中数超过阈值数(在高速缓存未命中由高速缓存标记本身跟踪的实现方式中),则目标高速缓存级别的高速缓存控制器将预取降级到较低高速缓存级别。

[0015] 在一个实施方案中,高速缓存控制器基于按需操作对先前预取的数据或指令的使用来跟踪预取使用度量。在基于这些预取度量确定之前的预取不准确(即,在被需要之前从高速缓存中逐出)或不及时(即,预取信息被需要的太迟)时,高速缓存控制器降低传入该目标高速缓存级别的部分或全部预取的优先级。因此,较低优先级预取不导致目标高速缓存中较高优先级高速缓存行的容量逐出。

[0016] 在一个实施方案中,高速缓存控制器根据高速缓存替换策略识别高优先级高速缓存行。例如,高优先级高速缓存行经常被重用,或者被比其他操作更关键的操作(诸如指令提取、转换后备缓冲区(TLB)提取、关键路径中的加载/存储等)重用。如果高优先级高速缓存行的比例超过阈值,则将预取降级到下一个较低级别的高速缓存允许高优先级“热”高速缓存行在目标高速缓存中保持不受干扰。

[0017] 图1示出实现预取降级机制的计算系统100的实施方案。通常,计算系统100体现为许多不同类型的装置中的任一种,包括但不限于膝上型计算机或台式计算机、移动电话、服务器、网络开关或路由器等。计算系统100包括可通过总线101彼此通信的多个硬件资源,包括部件102-108。在计算系统100中,部件102-108中的每一者都能够直接通过总线101或者经由其他部件102-108中的一者或多者与任何其他部件102-108通信。计算系统100中的部件101-108包含在单个物理外壳内,诸如膝上型或台式计算机机架或移动电话外壳。在替代实施方案中,计算系统100的一些部件体现为外部外围设备,使得整个计算系统100不驻留在单个物理外壳内。

[0018] 计算系统100还包括用于从用户接收信息或向用户提供信息的用户接口装置。具体地,计算系统100包括输入装置102,诸如键盘、鼠标、触摸屏或用于从用户接收信息的其他设备。计算系统100经由诸如监测器、发光二极管(LED)显示器、液晶显示器或其他输出装置的显示器105向用户显示信息。

[0019] 计算系统100另外包括用于通过有线或无线网络传输和接收数据的网络适配器107。计算系统100还包括一个或多个外围设备108。外围设备108可包括大容量存储装置、位置检测装置、传感器、输入装置、或由计算系统100使用的其他类型的装置。存储器系统106包括计算系统100使用的存储器装置,诸如随机存取存储器(RAM)模块、只读存储器(ROM)模

块、硬盘和其他非暂时性计算机可读介质。

[0020] 计算系统100包括处理单元104。在一个实施方案中,处理单元104包括驻留在公共集成电路衬底上的多个处理核心。处理单元104接收并执行存储在主存储器106中的指令109。指令109的至少一部分定义了包括可由处理单元104执行的指令的应用程序。

[0021] 计算系统100的一些实施方案可包括比图1所示的实施方案更少或更多的部件。例如,在没有任何显示器105或输入装置102的情况下实现某些实施方案。其他实施方案具有多于一个特定部件;例如,计算系统100的实施方案可具有多个处理单元104、总线101、网络适配器107、存储器系统106等。

[0022] 图2示出了根据实施方案的处理单元的高速缓存层次结构。处理单元104包括高速缓存层次结构,所述高速缓存层次结构包括L1高速缓存201、L2高速缓存202和L3高速缓存203。其他装置(诸如处理器核心230)经由高速缓存控制器211-213与这些高速缓存201-203交互,所述高速缓存控制器分别控制高速缓存201-203。处理器核心230通过执行指令109来运行操作系统231和用户应用程序232。在高速缓存层次结构中,最高L1高速缓存201是层次结构中速度最快且容量最小的高速缓存。连续较低高速缓存L2 202和L3 203越来越慢(即,延迟越来越高)和/或容量越来越大。

[0023] 硬件预取器221-223分别与高速缓存级别201-203相关联,并且针对它们的相关联高速缓存级别或低于其相关联高速缓存级别的高速缓存级别生成预取请求。所述预取请求通过向目标高速缓存加载数据或指令来支持应用程序232的执行,所述数据或指令在需要之前将由应用程序232使用。因此,硬件预取器221-223通过对应用程序232执行分支预测和/或通过基于应用程序232的先前存储器访问模式预测应用程序232的未来存储器访问来确定要预取那些数据或指令。预取请求也由处理器核心230执行应用程序232的指令而生成。例如,应用程序232指令可包括将某些数据或指令预取到特定指定级别的高速缓存的显式指令。

[0024] 图3示出根据实施方案的高速缓存300中的电路部件。高速缓存201-203中的每一者都以与高速缓存300类似的方式包括类似的部件和功能。高速缓存300包括存储器310,所述存储器存储高速缓存行阵列,每个高速缓存行将标签311中的一个或多个与高速缓存行中的数据312的一部分相关联。标签311包括关于其相关联高速缓存行中的数据的信息,诸如数据是否来自预取、预取的来源(例如,硬件预取器、应用程序等)、数据的访问频率、使用数据的操作类型等。高速缓存控制器320包括用于在存储器310中读取和写入标签311和数据312的读/写逻辑326。

[0025] 高速缓存300包含监视电路,所述监视电路包括预取度量322模块、高速缓存条目度量323模块和资源度量324模块,所述模块记录高速缓存300的性能度量。预取度量模块322基于标签311中的信息来测量指示预取准确性和及时性的度量。在一个实施方案中,当预取请求被高速缓存300接受时,控制器320更新与高速缓存行相关联的标签(例如,通过断言位),所述标签指示高速缓存行包含尚未使用的预取数据。当较高级别高速缓存或处理器230随后需要预取数据时,更新标签(例如,通过清除位)以反映需要数据的事实。随着时间的推移,预取度量模块322跟踪接收到需要请求的已使用预取高速缓存行与在需要之前被从高速缓存300逐出的未使用预取高速缓存行的比例。高比例的未使用预取指示由于分支预测错误或其他因素,预取不准确。

[0026] 在一个实施方案中,原始预取请求的来源也在标签311中被跟踪;例如,标签311指示预取请求是来自目标高速缓存的硬件预取器、较高级别高速缓存的硬件预取器还是来自执行应用程序指令的处理器230。在一个实施方案中,将线程标识符或其他标识应用程序的信息添加到标签,以标识发起预取请求的特定线程或进程。在一个实施方案中,系统100包括不同类型的硬件预取器(其基于观察不同类型的模式生成预取请求),所述不同类型的硬件预取器也被作为不同的预取来源进行跟踪。然后,控制器320能够独立地跟踪源自不同预取来源中的每一个的预取的预取准确性和及时性。

[0027] 预取度量模块322还跟踪最终需要预取数据的时间,但在较低高速缓存级别而不是作为初始目标的高速缓存级别。这往往指示数据相对于以相同高速缓存为目标的其他数据而言过早预取。在这种情况下,预取初始以较低级别高速缓存为目标的计算成本较低。因此,当预取不准确或不及时时,将预取请求降级到下一个较低高速缓存级别,以避免来自该高速缓存级别的不太可能被需要的预取数据污染作为初始目标的高速缓存。

[0028] 高速缓存条目度量模块323记录存储器310中的条目(例如,高速缓存行)的度量,诸如每个高速缓存行的访问频率、与高速缓存行相关联的操作时间等。在一个实施方案中,所述度量记录在标签311中。高速缓存条目度量用于确定高速缓存条目的优先级。相对于被需要的可能性较低的预取,被频繁访问或被较高优先级操作(例如,指令提取、TLB提取、关键路径中的加载/存储等)需要的高速缓存行被赋予较高优先级。

[0029] 资源度量模块324监视高速缓存300的未命中请求缓冲区331和牺牲缓冲区332,以指示高速缓存资源过度利用,诸如高速缓存未命中流量高。未命中请求缓冲区331存储从高速缓存中丢失的行,直到它们可被传送到高速缓存存储器310中,而牺牲缓冲区332存储由于高速缓存未命中而从高速缓存存储器310逐出的行。因此,当高速缓存300正经历高未命中率时,未命中请求缓冲区331和牺牲缓冲区332中的空间需要增加。发生这种情况时,高速缓存资源被过度利用;因此,较低优先级预取被降级到层次结构中较低级别、较高容量的高速缓存。

[0030] 决策逻辑321基于由监视电路跟踪的高速缓存性能度量来确定预取请求341是在高速缓存300处被接受还是被降级到较低级别高速缓存。预取请求341在高速缓存控制器320中的决策逻辑321处被接收,并且响应于接收到预取请求341,决策逻辑321相对于高速缓存存储器310中的现有条目的优先级确定预取的优先级。在一个实施方案中,现有条目包括当前在高速缓存存储器310中的条目,以及指定放置在高速缓存中的条目(例如,存在于未命中请求缓冲区中但尚未在高速缓存存储器310中的条目)。在一个实施方案中,相对优先级是预取的优先级级别与高速缓存存储器310中的一个或多个现有条目的优先级级别之间的差。换句话说,预取请求的相对优先级指示预取请求的优先级是高于还是低于目标高速缓存的阈值优先级级别。在一个实施方案中,高速缓存的阈值优先级级别是基于现有高速缓存行的最低优先级级别来确定的,所述高速缓存行是供预取逐出的候选高速缓存行。如果传入预取的优先级不大于任何现有高速缓存条目的优先级,则预取将降级到下一个较低级别的高速缓存。

[0031] 决策逻辑321基于高速缓存替换策略325和模块322-324跟踪的各种度量来确定预取请求341的优先级和现有高速缓存行的阈值优先级级别。因此,决策逻辑321确定哪些高速缓存行是最重要的并且应被保存在高速缓存300中。

[0032] 替换策略325定义了一组规则,用于识别在将新高速缓存行写入高速缓存存储器310时要逐出的最低优先级高速缓存行。例如,最不经常使用(LFU)替换策略指定最不经常使用高速缓存行以便在更经常使用高速缓存行被逐出之前从高速缓存300中逐出,而最近最少使用(LRU)替换策略在逐出最近更经常使用高速缓存行之前逐出最近最少使用高速缓存行。在一个实施方案中,高速缓存实现重新参考间隔预测(RRIP)替换策略,所述RRIP替换策略预测哪些高速缓存行可能在不久的将来被重用。

[0033] 因此,决策逻辑321基于替换策略确定存储器310中现有高速缓存行的优先级级别。在一个实施方案中,更可能被重用的高速缓存行被分配更高的优先级。如果传入预取请求的优先级级别不大于任何现有高速缓存行的优先级级别,则预取请求被降级到下一个较低级别的高速缓存以避免逐出任何较高优先级的现有高速缓存行。在替代实施方案中,决策逻辑321使用除高速缓存替换策略之外的机制来确定现有高速缓存行的相对优先级,并作为确定是降级还是接受传入预取请求的基础。

[0034] 除了重用的频率或新近度之外,现有高速缓存行的优先级还基于重用高速缓存行的操作类型来确定。在一个实施方案中,需要高速缓存行的操作的类型被记录在高速缓存行的标签中。当接收到预取请求341时,决策逻辑321将高优先级分配给由高优先级操作使用的高速缓存行。例如,与例如不在关键路径中的操作所使用的高速缓存行相比,应用程序232的关键路径中的转换后备缓冲区(TLB)遍历器或加载/存储操作所使用的高速缓存行被赋予了更高的优先级级别。因此,传入预取被降级,以避免逐出此类高优先级操作所使用的高速缓存行。

[0035] 除了高速缓存条目度量323之外,决策逻辑321还基于资源度量324确定是否将传入预取请求341降级。在一个实施方案中,当未命中请求缓冲区331和/或牺牲缓冲区332各自已满或已填满超过占用率阈值时,决策逻辑321将所有预取降级到下一个较低级别高速缓存。在替代实施方案中,决策逻辑321接受较高优先级预取请求的子集,而不是将所有预取降级。

[0036] 决策逻辑321基于预取度量322确定预取请求341的优先级级别。预取度量322指示先前预取是否准确和及时。如果先前预取不准确或不及时,则决策逻辑321将较低优先级分配给传入预取请求341。在一个实施方案中,针对预取的每个来源(例如,硬件预取器、执行应用程序指令的处理器等)分别跟踪预取准确性和及时性,使得由一个来源发出的不准确或不及时的预取请求不影响从不同来源发出的预取请求的优先级。决策逻辑321将较高优先级分配给源自先前已生成更准确和更及时预取的来源的预取请求,同时将较低优先级分配给源自自己生成不准确和/或不及时预取请求的来源的预取请求。在一个实施方案中,决策逻辑321将较高优先级分配给对将由高优先级操作(例如,关键路径中的操作等)使用的数据或指令的预取请求。

[0037] 对于在高速缓存300处接收的每个预取请求(诸如预取请求341),决策逻辑321通过将预取请求的优先级与现有高速缓存行的优先级进行比较来确定预取请求的相对优先级。如果预取请求的优先级低于已在高速缓存存储器310中的任何高速缓存行,则决策逻辑321通过将预取请求342的副本重定向到较低高速缓存级别来将预取请求341降级到较低高速缓存级别。如果高速缓存300已经是高速缓存层次结构中的最低级别高速缓存,则预取请求342被丢弃而不是被重定向到较低级别高速缓存。

[0038] 在一个实施方案中,决策逻辑321默认将预取请求342重定向到层次结构中的下一个较低级别高速缓存(例如,L2高速缓存将低优先级预取降级到L3高速缓存)。在替代实施方案中,决策逻辑321选择多个较低高速缓存级别中的任一者来接收降级的预取请求。在较低级别高速缓存处接收到降级的预取请求342后,较低级别高速缓存中的另一决策逻辑类似地基于其自身的预取度量、高速缓存条目度量和资源度量来确定是接受预取请求342还是将请求342再次降级到下一个较低级别高速缓存。

[0039] 对于由决策逻辑321确定的优先级高于接收高速缓存300的阈值优先级(例如,来自最低优先级高速缓存行)的每个预取请求341,决策逻辑321通过逐出最低优先级高速缓存行并将预取数据存储到预取请求341中指定的其存储器310中来接受预取请求341。为了跟踪传入预取的准确性和及时性,在标签311中设置位,指示数据来自预取。如果针对每个预取来源跟踪预取准确性和及时性,则预取请求341的来源也被记录在标签311中。

[0040] 图4示出根据实施方案的针对高速缓存存储器310中的每个高速缓存行存储在标签311中的信息。标签311包括预取指示401、预取来源402、访问频率403和操作类型404等。预取指示401被实现为单个位,当相关联高速缓存行包含预取的数据时所述位被断言,否则被取消断言。当高速缓存行中的数据是预取数据时,预取来源402指示预取请求的来源,并且可包括诸如线程标识符、装置标识符(例如,用于硬件预取器)和/或用于源自另一高速缓存级别的硬件预取器的预取的高速缓存级别标识符的信息。预取来源402还指示预取请求是否从更高的高速缓存级别降级。访问频率403指示在对高速缓存或高速缓存索引的访问期间,需要相关联高速缓存行中的数据的频率。操作类型404指示需要相关联高速缓存行中的数据的一个或多个操作的类型。标签401-404在高速缓存数据被访问(即,被写入或需要)时由高速缓存控制器320更新,并由决策逻辑321使用来确定传入预取请求和现有高速缓存行的优先级,如先前所述。

[0041] 图5示出根据实施方案的对在高速缓存层次结构中的不同级别接收的多个预取请求进行操作的预取降级机制。图5示出处理器核心230和高速缓存层次结构,所述高速缓存层次结构包括L1高速缓存201、L2高速缓存202和L3高速缓存203以及它们相应的高速缓存控制器211-213和预取器221-223。

[0042] 由硬件预取器221在L1高速缓存级别生成第一预取请求501。预取请求501以L1高速缓存201为目标并在高速缓存控制器211处被接收。高速缓存控制器211中的决策逻辑确定预取请求501具有比其现有高速缓存行中的至少一个高的优先级,因此逐出最低优先级高速缓存行以接受预取数据。

[0043] 作为处理器核心230执行应用程序232的预取指令的结果,从处理器核心230发出预取请求502。在L1高速缓存级别201,高速缓存控制器211中的决策逻辑确定预取请求502的优先级小于高速缓存201中最低优先级高速缓存行的优先级,因此将预取502降级到L2高速缓存级别202。可替代地,如果L1高速缓存201的资源由于高未命中率或其他原因而被过度利用,则可将预取502降级到L2高速缓存级别202。L2高速缓存控制器212中的决策逻辑确定预取请求502具有比L2高速缓存202中的最低优先级高速缓存行相对更高的优先级,并且预取请求502在L2高速缓存202中被接受。

[0044] 在一个实施方案中,特定缓存级别的硬件预取器能够生成以层次结构中较低高速缓存级别为目标的预取请求。因此,L1预取器221生成指向L2高速缓存级别202的预取请求

503。高速缓存控制器212中的决策逻辑确定预取请求503具有比L2高速缓存202中的任何现有高速缓存行低的优先级。作为响应,决策逻辑将预取请求503降级到下一个较低高速缓存级别L3 203。在L3高速缓存级别203,高速缓存控制器213中的决策逻辑基于其自身的高速缓存性能度量确定预取请求503的优先级也低于其任何现有高速缓存行。由于L3高速缓存203是层次结构中的最低高速缓存级别,因此预取请求503被丢弃。

[0045] 在一个实施方案中,L3高速缓存控制器213另外将预取请求503被丢弃的指示504传输到丢弃的预取503的数据所在的主存储器106的存储器控制器520。响应于接收到指示504,在预期对所尝试的预取数据存在紧急的需要请求时,存储器控制器520准备读取由丢弃的预取503指定的预取数据。例如,存储器控制器520通过打开包含数据的存储器页面来初始化对包含数据的存储器的访问,使得在需要所述数据时可以较低的延迟读取所述数据。

[0046] 图6是根据实施方案的示出预取过程600的流程图。预取过程600由计算系统100中的部件执行,所述部件包括高速缓存201-203(在图3中表示为高速缓存300)、高速缓存控制器211-213(即,高速缓存控制器320)、处理器核心230、存储器控制器520等。在框601处,高速缓存控制器320更新标签311并记录高速缓存性能度量,诸如预取度量322、高速缓存条目度量323和资源度量324。所述度量被记录在标签311和/或高速缓存控制器320中的寄存器和计数器中。

[0047] 在框603处,高速缓存300的高速缓存控制器320接收预取请求341。预取请求341是根据应用程序232中的显式预取指令,从高速缓存300的硬件预取器、更高级别高速缓存的硬件预取器或处理器230接收的。

[0048] 在框605处,决策逻辑321基于预取准确性度量322中的一者或多者确定预取请求341的优先级,所述预取准确性度量包括未使用的预取条目与已使用的预取条目的比例。未使用的预取条目包括在需要其之前从高速缓存300逐出的预取数据,而已使用的预取条目包括来自高速缓存300的需要的预取数据。决策逻辑321为预取请求341分配较高优先级,对应于已使用的预取条目的较高比例,这指示预取是准确的和及时的。在一个实施方案中,对于预取请求的每个来源,诸如相同级别或更高级别高速缓存的硬件预取器(包括降级的预取)、应用程序等,独立地跟踪预取准确性和及时性。

[0049] 在框607处,决策逻辑321基于高速缓存条目度量323和替换策略325确定高速缓存300的阈值优先级。在一个实施方案中,高速缓存条目的优先级级别对应于条目的访问频率、条目的访问最近度、与条目相关联的操作类型和/或替换策略325中定义的其他因素而增加。

[0050] 在框609处,如果高速缓存300的资源未被过度利用,则过程600在框611处继续。在框611处,决策逻辑321基于高速缓存性能度量确定预取请求341的相对优先级。在一个实施方案中,相对优先级是预取的优先级级别与目标高速缓存级别的阈值优先级之间的差。如果预取优先级高于高速缓存阈值优先级(例如,高速缓存300中的最低优先级条目),则在框613处逐出最低优先级高速缓存条目,并将用于预取请求341的预取数据存储存在高速缓存存储器中。

[0051] 在框601处更新标签311,并记录更新的高速缓存性能度量。例如,由于写入了包含预取数据的新高速缓存行,因此在新高速缓存行的标签中断言位,以指示数据是预取数据。

作为另一示例,如果在框613处逐出的数据是未使用的预取数据,则针对最初请求对逐出的数据进行预取的来源更新预取准确性度量(例如,已使用的预取条目与未使用的预取条目的比率)。此外,如果自上次更新601以来,需要任何先前预取的数据或逐出任何未使用的预取数据,则更新预取准确性度量。通过框603-615的操作,接受由高速缓存控制器接收的作为高优先级预取请求(即,具有比高速缓存阈值优先级高的优先级)的预取请求子集,并使预取数据存储存在高速缓存存储器300中。

[0052] 在框609处,如果高速缓存资源未被过度利用,则过程600在框617处继续。高速缓存资源的过度利用由高速缓存性能度量指示,所述高速缓存性能度量包括牺牲缓冲区占用率度量和未命中请求缓冲区占用率度量。牺牲缓冲区占用率度量表示目标高速缓存的牺牲缓冲区中已使用容量的量。未命中请求缓冲区占用率度量表示目标高速缓存的未命中请求缓冲区中已使用容量的量。在一个实施方案中,当牺牲缓冲区占用率度量或未命中请求缓冲区占用率度量超过相应阈值时,则认为高速缓存资源被过度利用,使得过程600从框609继续到框617。

[0053] 当在框611处,预取优先级不高于高速缓存阈值优先级时,也到达框617。当已在高速缓存存储器310中的每个高速缓存行具有比预取请求341更高的优先级时,这是真的。在这种情况下,不会逐出任何现有的较高优先级高速缓存行,而是将相对较低优先级预取降级。

[0054] 在框617处,如果目标高速缓存300的高速缓存层次结构中存在一个或多个较低高速缓存级别,则决策逻辑321在框619处选择较低高速缓存级别中的一者以接收降级的预取请求342。在一个实施方案中,决策逻辑321自动选择下一个较低高速缓存级别(例如,L1高速缓存201中的决策逻辑选择L2高速缓存202来接收降级的预取请求)。在替代实施方案中,决策逻辑不一定选择紧邻的下一个较低高速缓存(例如,L1高速缓存201选择L3高速缓存203来接收降级的预取请求)。在框621处,决策逻辑321通过将预取请求342重定向到所选择的较低级别高速缓存来将所述预取请求降级。通过框609和617-621的操作,在高速缓存资源被过度利用时,以高速缓存300为目标的预取请求被降级到较低级别高速缓存。

[0055] 在替代实施方案中,在框607处,使用高速缓存资源利用率级别来设置高速缓存的阈值优先级级别,然后过程600从框607继续到框611。框607处的决策逻辑321与高速缓存资源的利用率增加成比例地增加现有高速缓存条目的阈值优先级级别,从而在资源利用率高时限制经由框611接受到高速缓存300中的预取的数量。

[0056] 在框603处,在较低级别高速缓存处接收降级的预取请求342。较低级别高速缓存类似地对接收的预取请求执行过程600,并基于根据其自身的高速缓存性能度量确定的优先级来接受或降级预取请求。也就是说,如果先前降级的预取请求的优先级高于较低级别高速缓存的高速缓存阈值优先级,则根据低优先级预取请求在较低级别高速缓存中接受预取数据,如框615处所提供的。

[0057] 低优先级预取请求(即,优先级低于高速缓存阈值优先级)再次降级到下一个较低高速缓存级别。因此,预取可通过连续较低的高速缓存级别多次降级,直到它被接受或丢弃。

[0058] 在框617处,如果层次结构中不存在较低级别的高速缓存(即,预取请求341初始以最低高速缓存级别为目标或降级到最低高速缓存级别),则过程600在框623处继续。在框

623处,在最低高速缓存级别(例如,L3高速缓存203)丢弃342预取请求,并且不基于所述请求来预取数据。在框625处,丢弃预取请求342的决策逻辑321向存储器控制器520指示丢弃的预取。存储器控制器520通过例如打开存储器页面以开始访问来准备读取在丢弃的预取请求中指定的数据,如在627处所提供的。过程600从框627返回到框601,以更新标签和高速缓存性能度量。

[0059] 在高速缓存300的给定级别,预取过程600针对在高速缓存300处接收的多个预取请求中的每一者重复。因此,优先级低于高速缓存阈值优先级的预取请求子集被降级到一个或多个较低高速缓存级别,并可能在最低高速缓存级别被丢弃。在高速缓存300处接受优先级高于高速缓存阈值优先级的预取请求子集,并根据每个接受请求将数据预取到高速缓存。在一个实施方案中,如果高速缓存300是层次结构中的最低高速缓存级别,则在高速缓存300被过度利用时接收的预取请求被降级到较低级别高速缓存或被丢弃。替代地,使用高速缓存资源利用率度量来确定高速缓存阈值优先级级别。

[0060] 一种方法包括:记录目标高速缓存的第一高速缓存性能度量集;对于在所述目标高速缓存处接收的多个预取请求中的每个预取请求,基于所述第一高速缓存性能度量集确定所述预取请求相对于所述目标高速缓存的阈值优先级级别的相对优先级;对于所述多个预取请求的第一子集中的每个低优先级预取请求,响应于确定所述低优先级预取请求的优先级小于所述目标高速缓存的所述阈值优先级级别,将所述低优先级预取请求重定向到第一较低级别高速缓存;以及对于所述多个预取请求的第二子集中的每个高优先级预取请求,响应于确定所述高优先级预取请求的优先级大于所述目标高速缓存的所述阈值优先级级别,根据所述高优先级预期请求将预取数据存储到所述目标高速缓存中。

[0061] 所述方法还包括:对于所述低优先级预取请求中的每一者,从所述目标高速缓存的高速缓存层次结构中选择另一高速缓存作为所述第一较低级别高速缓存,其中所述第一较低级别高速缓存具有大于所述目标高速缓存的容量;以及根据所述低优先级预取请求,将预取数据存储到所述第一较低级别高速缓存中。

[0062] 所述方法还包括对于所述第一子集的一个或多个预取请求,基于所述第一较低级别高速缓存的第二高速缓存性能度量集,将所述一个或多个预取请求从所述第一较低级别高速缓存重定向到第二较低级别高速缓存,其中所述第二较低级别高速缓存具有比所述第一较低级别高速缓存高的容量。

[0063] 所述方法还包括:对于所述第一子集的一个或多个预取请求,将所述一个或多个预取请求重定向到所述目标高速缓存的所述高速缓存层次结构中的最低级别高速缓存;以及响应于确定所述一个或多个预取请求的所述优先级小于所述最低级别高速缓存的阈值优先级级别,丢弃所述一个或多个预取请求。

[0064] 所述方法还包括,对于所述多个预取请求中的每个预取请求,基于预取准确性度量来确定所述预取请求的优先级。所述预取准确性度量是对于所述目标高速缓存的预取条目集,基于未使用的预取条目与已使用的预取条目的比例来确定的。

[0065] 所述方法还包括对于所述多个预取请求中的每个预取请求,基于所述预取请求的来源来确定所述预取请求的优先级。所述来源包括硬件预取器和用户应用程序中的一者。

[0066] 所述方法还包括基于高速缓存替换策略,对于所述目标高速缓存中的多个高速缓存条目中的每个高速缓存条目、所述高速缓存条目的访问频率和与所述高速缓存条目相关

联的操作类型,确定所述目标高速缓存的所述阈值优先级级别。

[0067] 在所述方法中,所述第一高速缓存性能度量集包括所述目标高速缓存的牺牲缓冲区的牺牲缓冲区占用率度量和所述目标高速缓存的未命中请求缓冲区的未命中请求缓冲区占用率度量。

[0068] 一种计算装置包括:监视电路,所述监视电路用于记录目标高速缓存的第一高速缓存性能度量集;以及第一决策逻辑电路,所述第一决策逻辑电路与所述监视电路耦合。所述第一决策逻辑电路对于在所述目标高速缓存处接收的多个预取请求中的每个预取请求,基于所述第一高速缓存性能度量集确定所述预取请求相对于所述目标高速缓存的阈值优先级级别的相对优先级;对于所述多个预取请求的第一子集中的每个低优先级预取请求,响应于确定所述低优先级预取请求的优先级小于所述目标高速缓存的所述阈值优先级级别,将所述低优先级预取请求重定向到第一较低级别高速缓存;以及对于所述多个预取请求的第二子集中的每个高优先级预取请求,响应于确定所述高优先级预取请求的所述优先级大于所述目标高速缓存的所述阈值优先级级别,根据所述高优先级预期请求将预取数据存储存储在所述目标高速缓存中。

[0069] 在所述计算装置中,所述第一决策逻辑电路进一步对于所述低优先级预取请求中的每一者,从所述目标高速缓存的高速缓存层次结构中选择另一高速缓存作为所述第一较低级别高速缓存。所述第一较低级别高速缓存具有大于所述目标高速缓存的容量。所述第一较低级别高速缓存对于所述低优先级预取请求中的每一者,根据所述低优先级预取请求来存储预取数据。

[0070] 所述计算装置还包括:第二决策逻辑电路,所述第二决策逻辑电路在所述第一较低级别高速缓存中;以及第二较低级别高速缓存,所述第二较低级别高速缓存与所述第二决策逻辑电路耦合并且具有比所述第一较低级别高速缓存高的容量。所述第二决策逻辑电路对于所述第一子集中的一个或多个预取请求,基于所述第一较低级别高速缓存的第二高速缓存性能度量集,将所述一个或多个预取请求从所述第一较低级别高速缓存重定向到所述第二较低级别高速缓存。

[0071] 所述计算装置还包括:最低级别高速缓存,所述最低级别高速缓存在所述目标高速缓存的所述高速缓存层次结构中;以及第二决策逻辑电路,所述第二决策逻辑电路对于所述第一子集中的一个或多个预取请求,将所述一个或多个预取请求重定向到所述最低级别高速缓存;以及第三决策逻辑电路,所述第三决策逻辑电路在所述最低级别高速缓存中,所述第三决策逻辑电路响应于确定所述一个或多个预取请求的所述优先级小于所述最低级别高速缓存的阈值优先级级别,丢弃所述一个或多个预取请求。

[0072] 所述计算装置还包括预取度量模块,所述预取度量模块与所述第一决策逻辑耦合,所述预取度量模块对于所述目标高速缓存的预取条目集,基于未使用的预取条目与已使用的预取条目的比例来确定预取准确性度量。所述第一决策逻辑进一步对于所述多个预取请求中的每个预取请求,基于所述预取准确性度量来确定所述预取请求的优先级。

[0073] 所述计算装置还包括硬件预取器和处理器,所述处理器基于应用程序的执行指令生成所述多个预取请求中的一个或多个。所述决策逻辑进一步对于所述多个预取请求中的每个预取请求,基于所述预取请求的来源来确定所述预取请求的优先级。所述来源包括所述硬件预取器和所述处理器中的一者。

[0074] 所述计算装置还包括高速缓存条目度量模块,所述高速缓存条目度量模块记录高速缓存条目度量,所述高速缓存条目度量对于所述目标高速缓存中的多个高速缓存条目中的每个高速缓存条目,包括所述高速缓存条目的访问频率和与所述高速缓存条目相关联的操作类型。所述决策逻辑进一步基于高速缓存替换策略和所述高速缓存条目度量确定所述阈值优先级级别。

[0075] 一种计算系统包括:处理单元,所述处理单元执行应用程序;多个高速缓存,所述多个高速缓存在与所述处理单元耦合的高速缓存层次结构中;以及高速缓存控制器,所述高速缓存控制器与所述多个高速缓存耦合。所述高速缓存控制器包括监视电路,所述监视电路记录目标高速缓存的第一高速缓存性能度量集;以及决策逻辑电路,所述决策逻辑电路与所述监视电路耦合。所述决策逻辑电路对于在所述目标高速缓存处接收的多个预取请求中的每个预取请求,基于所述第一高速缓存性能度量集确定所述预取请求相对于所述目标高速缓存的阈值优先级级别的相对优先级;对于所述多个预取请求的第一子集中的每个低优先级预取请求,响应于确定所述低优先级预取请求的优先级小于所述目标高速缓存的所述阈值优先级级别,将所述低优先级预取请求重定向到第一较低级别高速缓存;以及对于所述多个预取请求的第二子集中的每个高优先级预取请求,响应于确定所述高优先级预取请求的所述优先级大于所述目标高速缓存的所述阈值优先级级别,根据所述高优先级预取请求将预取数据存储于所述目标高速缓存中。

[0076] 所述计算系统还包括存储器控制器,所述存储器控制器与所述高速缓存控制器耦合。所述决策逻辑电路进一步:对于所述第一子集中的一个或多个预取请求,将所述一个或多个预取请求重定向到所述目标高速缓存的所述高速缓存层次结构中的最低级别高速缓存;响应于确定所述一个或多个预取请求的所述优先级小于所述最低级别高速缓存的阈值优先级级别,丢弃所述一个或多个预取请求;以及将所述预取请求的指示传输到存储器控制器。所述存储器控制器响应于所述指示,初始化对包含所述预取数据的所述存储器的访问。

[0077] 所述计算系统还包括硬件预取器,所述硬件预取器与所述高速缓存层次结构耦合以基于以下项为所述应用程序生成所述多个预取请求中的一者或多者:为所述应用程序执行分支预测;以及基于所述应用程序的先前存储器访问模式预测存储器访问。

[0078] 在所述计算系统中,所述处理单元进一步根据所述应用程序的指令生成所述多个预取请求中的一者或多者。

[0079] 所述计算系统还包括多个高速缓存控制器,所述多个高速缓存控制器包括所述高速缓存控制器。所述多个高速缓存控制器中的每个高速缓存控制器控制所述高速缓存层次结构中的所述多个高速缓存中的一个,并将所述第一子集中的所述低优先级预取请求中的一者或多者重定向到所述高速缓存层次结构中具有比所述多个高速缓存中的相关联一个高速缓存高的容量的另一高速缓存。

[0080] 如本文所用,术语“耦合到”可意味着直接或通过一个或多个中间部件间接地耦合。通过本文所述的各种总线提供的任何信号可与其他信号进行时分复用,并且可通过一根或多根公共总线提供。另外,电路部件或框之间的互连可显示为总线或单根信号线。总线中的每一个替代地可以是一根或多根单信号线,并且单信号线中的每一个替代地可是总线。

[0081] 某些实施方案可实现为计算机程序产品,其可包括存储在非暂时性计算机可读介质上的指令。这些指令可用于对通用或专用处理器进行编程以执行所描述的操作。计算机可读介质包括用于以机器(例如,计算机)可读的形式(例如,软件、处理应用)存储或传输信息的任何机制。非暂时性计算机可读存储介质可包括但不限于磁存储介质(例如,软盘);光存储介质(例如,CD-ROM);磁光存储介质;只读存储器(ROM);随机存取存储器(RAM);可擦可编程存储器(例如,EPROM和EEPROM);闪存或另一种类型的适合存储电子指令的介质。

[0082] 另外,一些实施方案可在分布式计算环境中实践,其中计算机可读介质存储在多于一个计算机系统上和/或由多于一个计算机系统执行。另外,可在连接计算机系统的传输介质间拉取或推送计算机系统之间传递的信息。

[0083] 通常,表示计算系统100的数据结构和/或其承载在计算机可读存储介质上的部分可以是数据库或其他数据结构,其可以由程序读取并直接或间接地用于制造包括计算系统100的硬件。例如,数据结构可以是采用高级设计语言(HDL)(诸如Verilog或VHDL)的硬件功能的行为级描述或寄存器传送级(RTL)描述。所述描述可由合成工具读取,所述合成工具可合成所述描述以产生包括来自合成库的门的列表的网表。所述网表包括一组门,其还表示包括计算系统100的硬件的功能性。然后,可放置并且路由该网表以产生描述要应用于掩模的几何形状的数据集。所述掩模然后可用于各种半导体制造步骤中,以产生对应于计算系统100的一个或多个半导体电路。替代地,计算机可读存储介质上的数据库可以是网表(带有或没有合成库)或数据集(根据需要),或图形数据系统(GDS) II数据。

[0084] 尽管以特定顺序显示和描述了本文方法的操作,但是可更改每种方法的操作顺序,以便可以相反的顺序执行某些操作或者以使得可至少部分地与其他操作同时执行某些操作。在另一实施方案中,不同操作的指令或子操作可以是间歇和/或交替的方式。

[0085] 在以上说明书中,已参考其具体示例性实施方案描述了实施方案。然而,将明显的是:在不脱离如在所附权利要求中阐述的实施方案的更广范围的情况下,可对其做出各种修改和改变。因此,说明书和附图被认为是说明性的意义而不是限制性的意义。

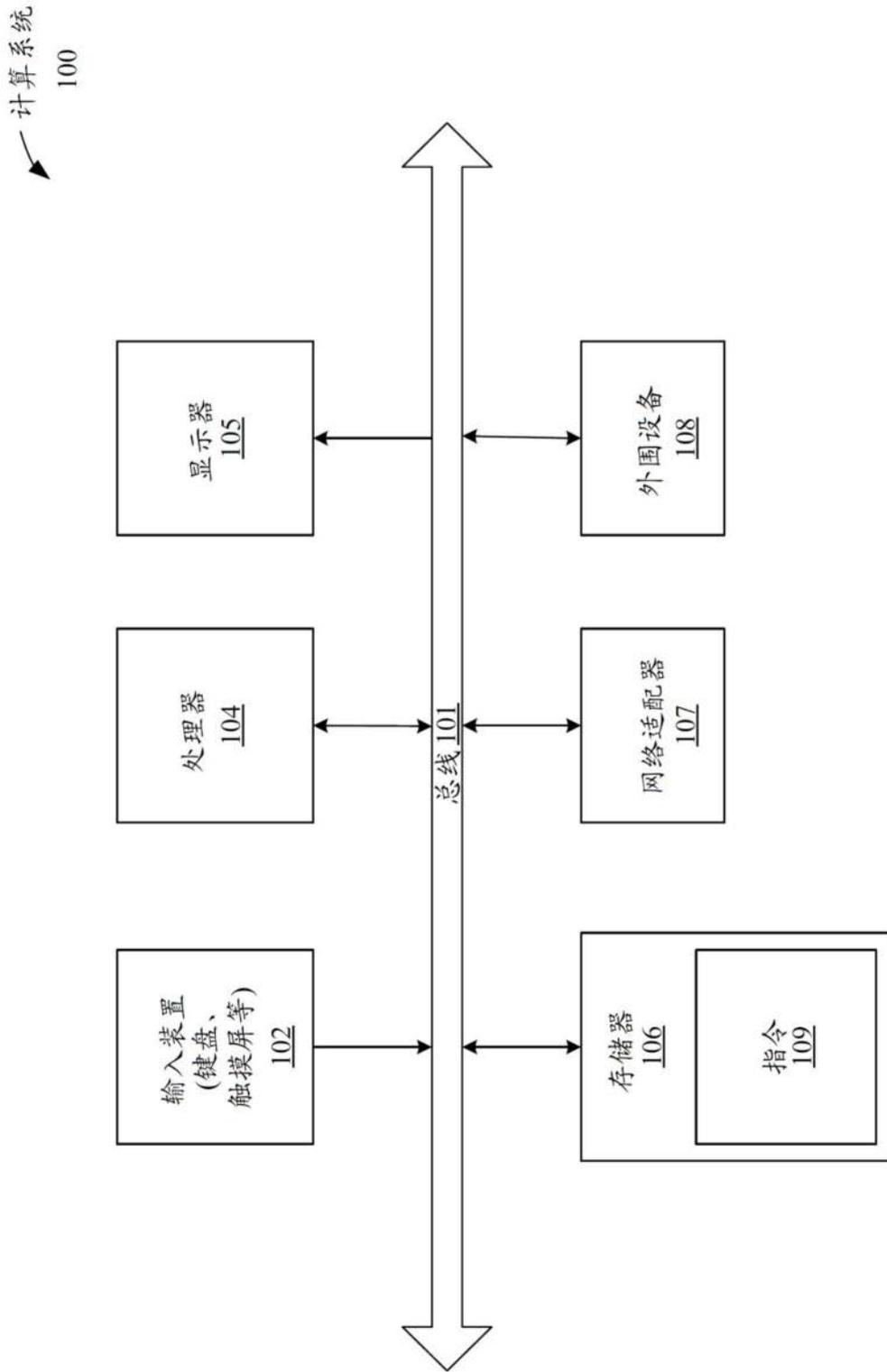


图1

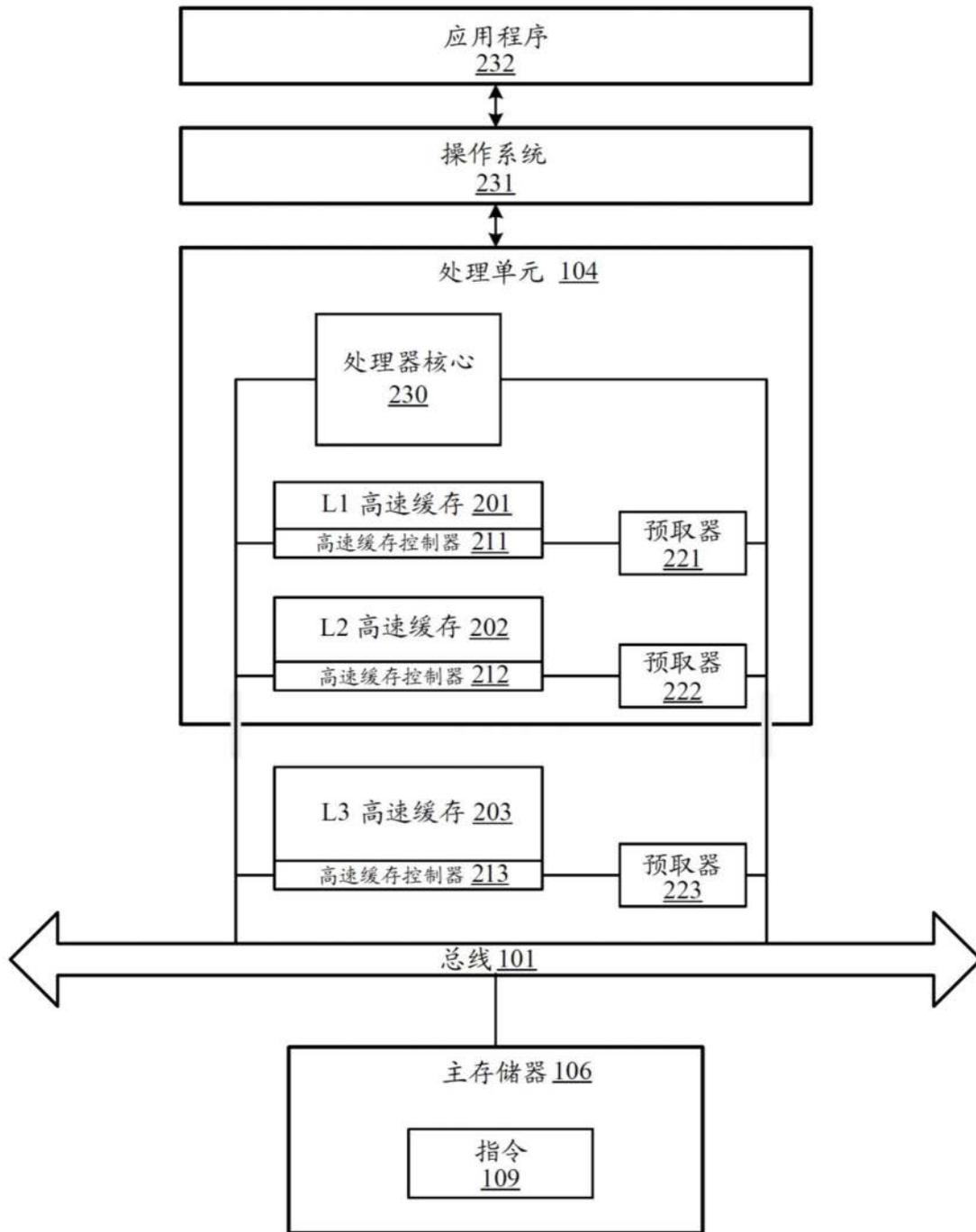


图2

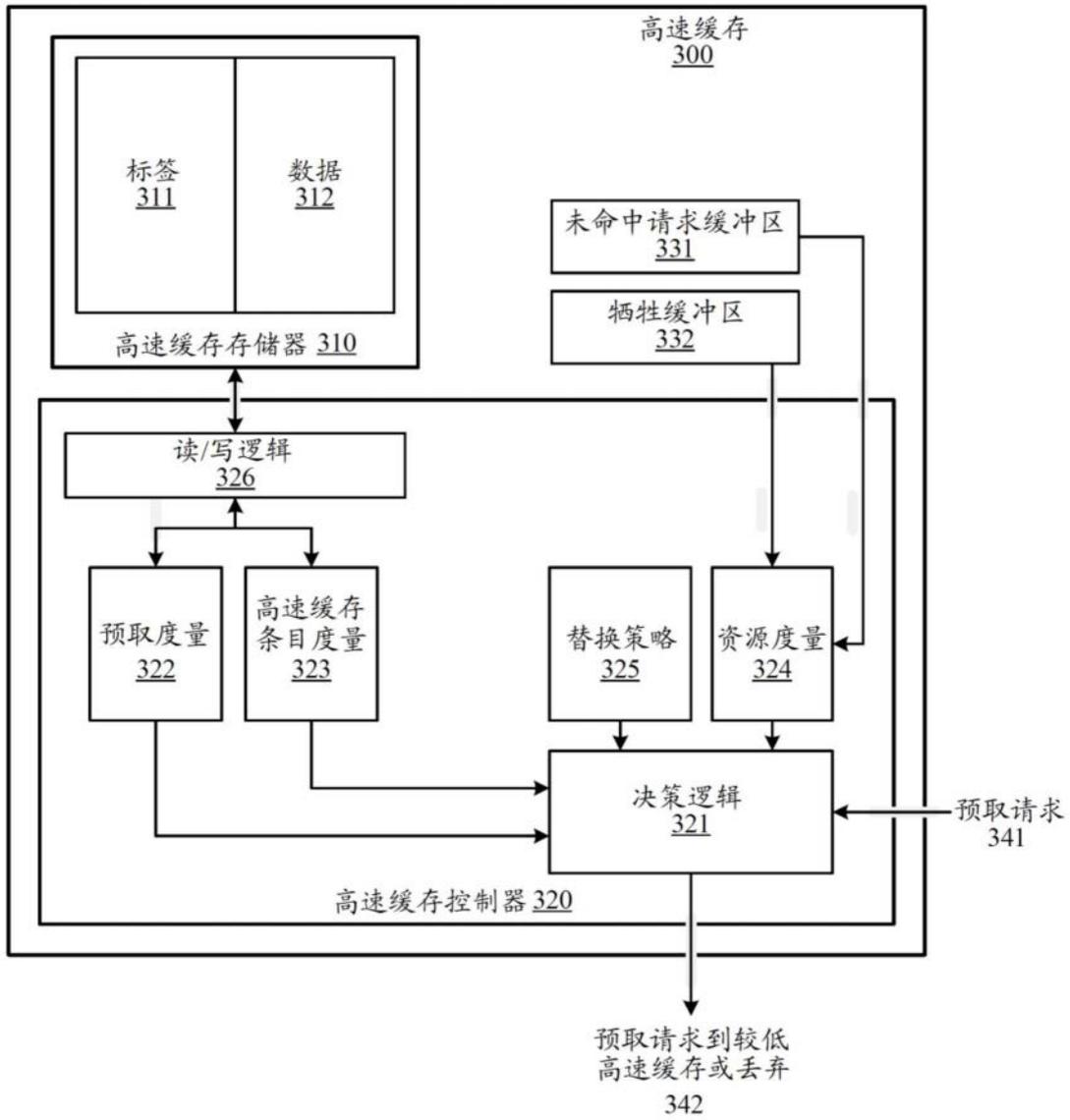


图3

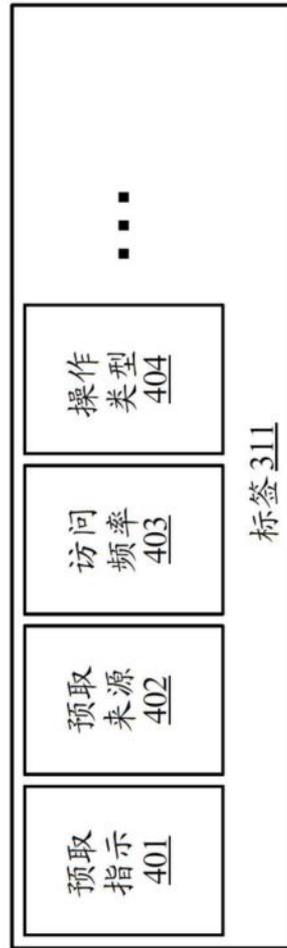


图4

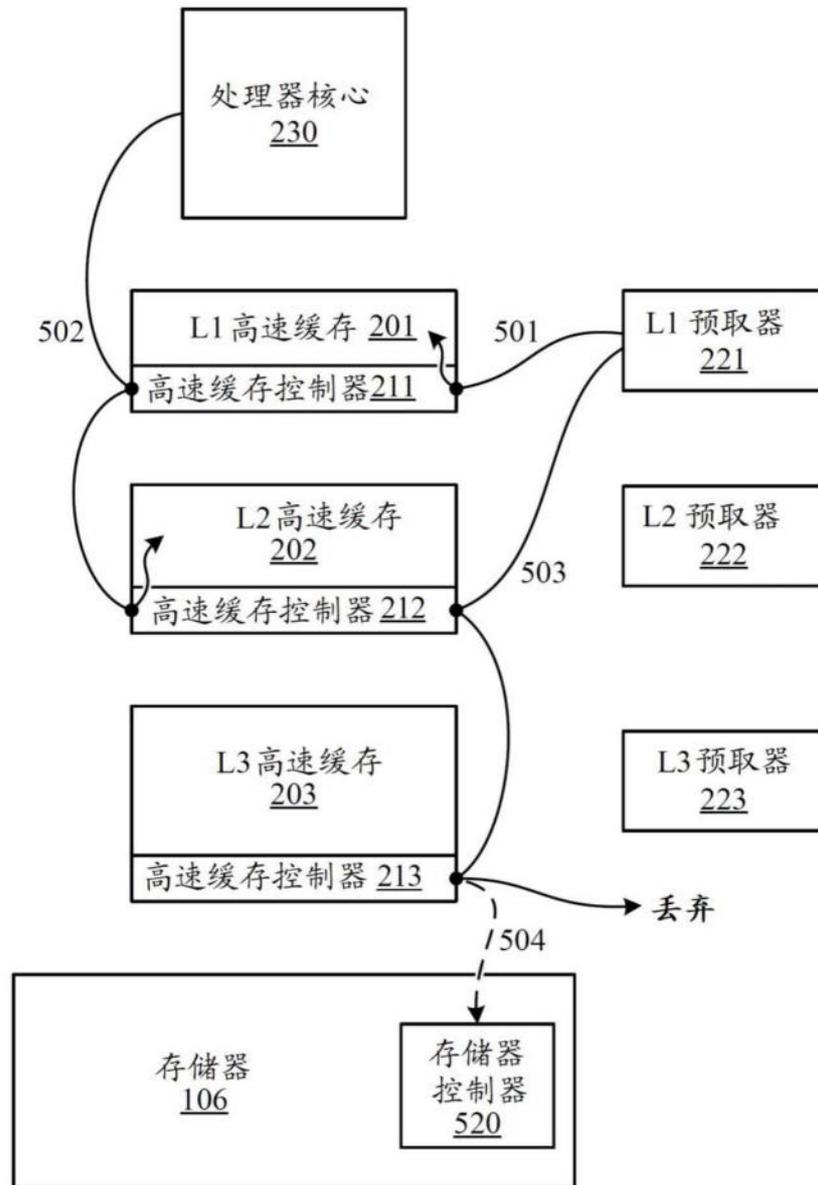


图5

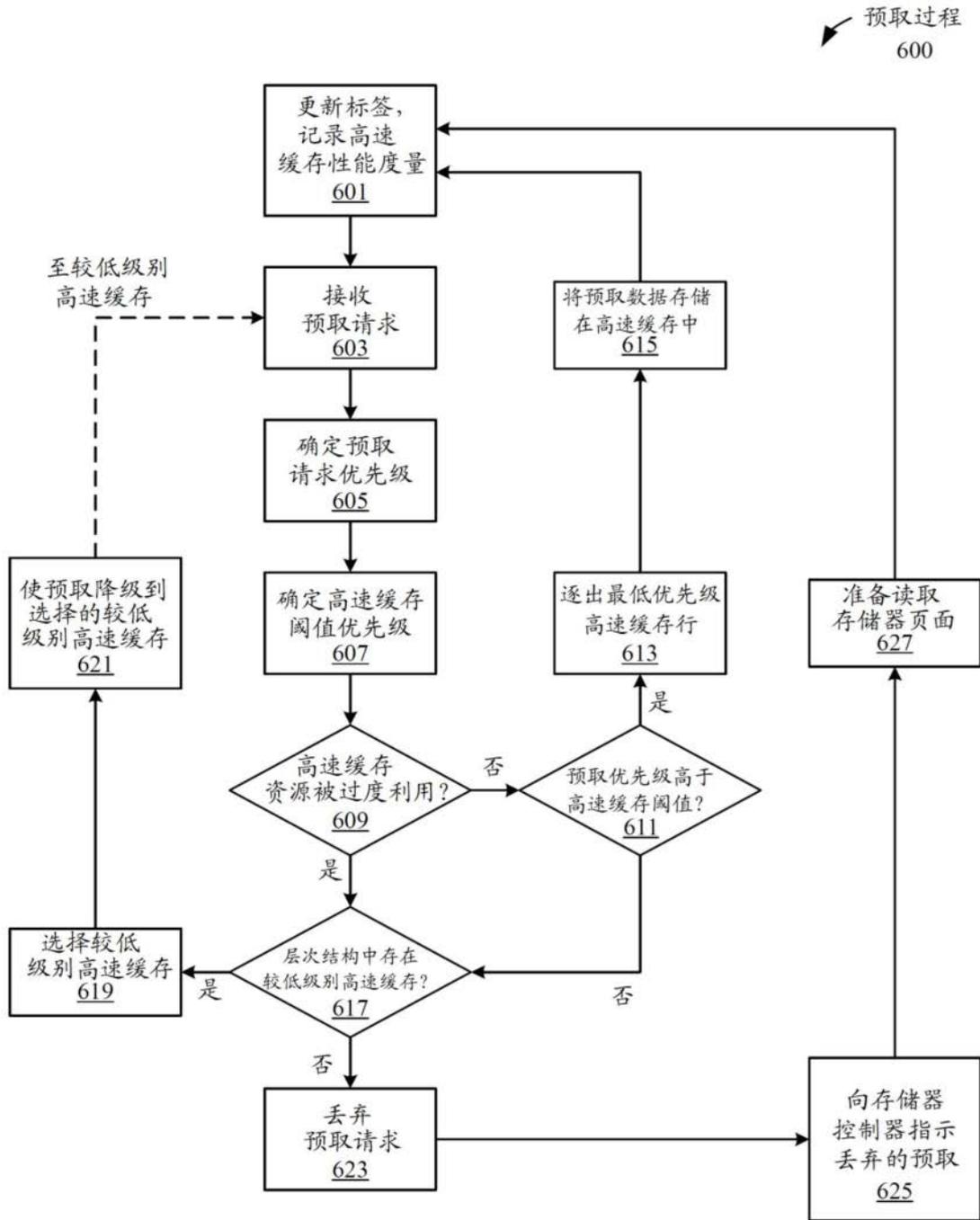


图6