



[12] 发 明 专 利 说 明 书

[21] ZL 专利号 99101097.3

[45] 授权公告日 2003 年 10 月 29 日

[11] 授权公告号 CN 1126046C

[22] 申请日 1999. 1. 15 [21] 申请号 99101097. 3

[30] 优先权

[32] 1998. 2. 17 [33] US [31] 024318

[71] 专利权人 国际商业机器公司

地址 美国纽约州

[72] 发明人 R · K · 阿里米里 J · S · 多德森

J · D · 刘易斯

审查员 哈雅坤

[74] 专利代理机构 中国专利代理(香港)有限公司

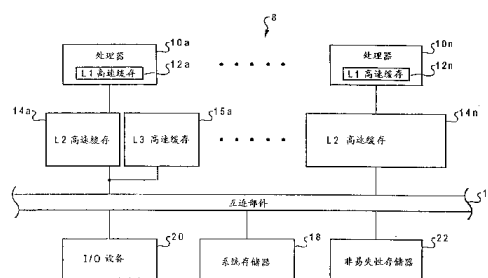
代理人 王 勇 王忠忠

权利要求书 3 页 说明书 22 页 附图 6 页

[54] 发明名称 包括多层次分级存储器体系的数据
处理系统的高速缓存相关性协议

[57] 摘要

本文介绍一种数据处理系统和方法。数据处理系统包括多个高速缓存和多个处理器，它们至少被划分成第一、第二两个群集，各自有至少一个高层次高速缓存和至少一个低层次高速缓存。第一群集的高层次高速缓存的第一个数据项是与表示某特定地址的一个地址标记关联存储的。第一群集的高层次高速缓存的相关性指示符被设置到第一状态。在第二群集的高层次高速缓存中，第二数据项是与表示该特定地址的一个地址标记关联存储的。



1. 一种保持数据处理系统中高速缓存相关性的方法，其中的数据处理系统包括互连部件、多个高速缓存和多个处理器，该多个高速缓存和多个处理器至少被划分成第一、第二两个群集，所述第一和第二群集中每个群集与所述互连部件连接、并且至少有一个高层次高速缓存和至少有一个低层次高速缓存，该方法包括：

在所述第一群集的第一高层次高速缓存中，将第一数据项与表示某地址的某地址标记关联存储，并将第一相关性指示符设置到第一状态，该第一状态表示所述地址标记有效、所述第一数据项无效；

10 在所述第二群集的第二高层次高速缓存中，当所述第一高层次高速缓存正以所述第一状态存储所述第一数据项时，将第二数据项与表示所述地址的某地址标记关联存储，并将第二相关性指示符设置到所述第一状态，该第一状态表示所述地址标记有效、所述第二数据项无效；

15 响应所述第一和所述第二高层次高速缓存二者之一，检测与由所述数据处理系统中另一设备启动的所述地址关联的数据传输，该数据传输包括第三数据项，在检测所述第一和所述第二高层次高速缓存的一个中用所述第三数据项替换所述第一和第二数据项中相应的一个，并将所述第一和第二相关性指示符中相应的一个设置到第二状态，其中，所述检测所述第一和所述第二高层次高速缓存中的一个是在其群集中，响应一个请求，通过对另一个高速缓存共享干预，决定提供所述第三数据项。

2. 权利要求 1 的方法，其特征在于，替换所述第一和第二数据项中相应的一个包括响应检测通过读取请求启动的数据传输替换所述第一和第二数据项中相应的一个。

3. 权利要求 1 的方法，其特征在于，所述数据处理系统进一步包括一个系统存储器，还在于，替换所述第一和第二数据项中相应的一个包括响应检测包含所述第三数据项的数据传输替换所述第一和第二数据项中相应的一个，其中所述第三数据项相对系统存储器中的某对应数据项进行修改。

4. 权利要求 3 的方法，其特征在于，所述数据传输是由所述第一群集中第三高层次高速缓存将所述第三数据项回写到所述系统存

储器。

5. 权利要求 2 的方法，其特征在于，进一步包括以下步骤：

在所述相关性指示符设置在所述第二状态期间，响应在所述第一
5 高层次高速缓存收到要求所述第三数据项的请求时，提供所述第三数
据项；和

将所述第一高层次高速缓存中的相关性指示符更新到第三状态。

6. 权利要求 1 的方法，其特征在于，将所述第一高层次高速缓存中的相关性指示符设置到第一状态包含，响应所述多个处理器的其
10 中之一表示要向所述地址标记表示的所述地址存储，将所述第一高层次高速缓存中的所述相关性指示符设置到所述第一状态。

7. 权利要求 1 的方法，其特征在于，进一步包含，将所述第一群集的较低层次高速缓存中的某相关性指示符设置到所述第一状态，其中该相关性指示符与表示所述地址的地址标记关联。

15 8. 一种数据处理系统，它包含：

一个互连部件；

一个与所述互连部件相连的系统存储器；

与所述互连部件相连的第一群集和第二群集，其中：

20 所述第一群集和所述第二群集各自至少有一个处理器、一个高层次高速缓存、一个低层次高速缓存，其中所述第一群集的第一高层次高速缓存和所述第二群集的第二高层次高速缓存各自同时包括：

数据存储器；

标记存储器，用于存储表示特定地址的地址标记，该地址标记与所述数据存储器中的某数据项关联；

25 相关性指示符，设置到第一状态，该状态表示所述地址标记有效而所述关联数据项无效；

高速缓存控制器，响应与由所述数据处理系统中另一设备启动的所述特定地址关联的数据传输的检测，该数据传输包括第二数据项，该控制器用所述第二数据项替换所述数据项，并将所述相关性指示符
30 设置到第二状态，其中，包含所述高速缓存控制器的高层次高速缓存存在其群集中响应一个请求，通过共享干预，提供所述第二数据项。

9. 权利要求 8 的数据处理系统，其特征在于，所述数据传输是

响应一个读取请求而启动的读取数据传输。

10. 权利要求 9 的数据处理系统，其特征在于，在所述数据传输中的第二数据项相对所述系统存储器中的某对应数据项进行修改。

11. 权利要求 10 的数据处理系统，其特征在于，所述数据传输
5 包含由第三高层次高速缓存将所述第二数据项回写到所述系统存储器。

12. 权利要求 9 的数据处理系统，其特征在于，所述第一高层次高速缓存的所述高速缓存控制器，在所述相关性指示符设置到所述第二状态期间，响应在所述第一高层次高速缓存收到要求所述第二数据
10 项的请求，提供所述第二数据项，并且将所述第一高层次高速缓存中的所述相关性指示符更新到第三状态。

13. 权利要求 8 的数据处理系统，其特征在于，所述第一高层次高速缓存的所述高速缓存控制器，响应所述数据处理系统中某个处理器表示要向所述地址标记表示的所述地址存储，将所述相关性指示符
15 设置到所述第一状态。

14. 权利要求 8 的数据处理系统，其特征在于，在所述第一高层次高速缓存的所述相关性指示符被设置到第一状态期间，所述第一群集的某较低层次高速缓存中的某相关性指示符被设置到所述第一状态，其中在所述第一群集的所述较低层次高速缓存中的该相关性指示
20 符与表示所述地址的地址标记关联。

15. 权利要求 8 的数据处理系统，其特征在于，所述第一群集包括高速缓存互连和多个高层次高速缓存，该多个高层次高速缓存包括所述第一高层次高速缓存，并且每个都与所述高速缓存互连相连。

16. 权利要求 8 的数据处理系统，其特征在于，所述第一群集包
25 括多个处理器和多个高层次高速缓存，其中每个所述高层次高速缓存只能被所述多个处理器中的相应一个直接访问。

包括多层次分级存储器体系的数据处理系统的
高速缓存相关性协议

5 技术领域

本发明总体涉及一种数据处理方法和系统，特别是一种在多处理器数据处理系统中保持高速缓存相关性的方法和系统。更具体来说，本发明涉及一种用于含有多层次分级存储器体系的数据处理系统的改进高速缓存相关性协议。

10 背景技术

在常规的对称多处理器（SMP）数据处理系统中，所有处理器都基本相同，就是说，所有处理器都使用相同的指令集和通讯协议，都具有类似的硬件结构，一般都配备类似的分级存储器体系。例如，某常规 SMP 数据处理系统中可能含有一个系统存储器，多个处理单元—
15 每个处理单元包括一个处理器和一级或多级高速缓存，以及一个将各处理单元互相连接并与系统存储器连接的系统总线。要获得 SMP 数据处理系统中的有效执行结果，重要的是保持一个相干的分级存储器体系，即对所有处理器提供单一内存内容视图。

要保持相干的分级存储器体系，就要选择使用某种存储相关性协议，例如 MESI 协议。在 MESI 协议中，至少所有高层次（高速缓冲）
20 存储器的每个相关区组（coherency granule）（例如高速缓冲存储器行或区段）都相联地存储着一个表示相关性状态的指示项。各相关区组具有下述 4 种状态之一：修改（M）、排他（E）、共享（S）、无效（I）。在高速缓存目录中用两个二进制位来表示状态。修改状态表示，相关区组（coherency granule）只有存储在被修改相关区组的高速缓存中，并且修改相关区组的值尚未写入系统存储器时才是有效的。如果相关区组被指示处于排他状态，则该相关区组仅驻留在
25 分级存储器体系中所有同一层次高速缓存中含有排他状态的相关区组的高速缓存中。不过，排他状态的数据是与系统存储器一致的。如果相关区组在高速缓存目录中被指示处于共享状态，则该相关区组驻留在相联高速缓存中，同时还驻留在分级存储器体系中至少另一个同一层次的高速缓存中，并且相关区组的所有拷贝都与系统存储器一
30

致。最后一个是无效状态，它表示与相关区组相联的数据和地址标记都无效。

各相关区组（例如高速缓冲存储器行）被设置到何种状态，既依赖于高速缓冲存储器行的上一个状态，也依赖于请求处理器寻求的存储器访问的类型。因此，要保持多处理器数据处理系统中的存储相关性，就要求处理器在系统总线上传播表明其对存储单元读写意向的消息。例如，如果某处理器希望向某存储单元写数据，则处理器首先必须向所有其它处理单元通报其要向该存储单元写数据的意向，在收到所有其它处理单元的允许信息后，才能执行写操作。由请求处理器接收的允许消息，表示对该存储单元内容的所有其它高速缓存的拷贝都已经被无效，由此保证其它处理器不会访问过时的本地数据。这种消息交换称作交叉无效（XI）。

本发明认为，尽管高速缓存项目的交叉无效有助于保持 SMP 数据处理系统中的存储相关性，但是远程处理器对高速缓存项目的无效却会因为本地高速缓存命中率的降低，而对数据处理系统的效能产生负面影响。因此，即使配备了大容量的本地高速缓存，当某个处理单元检索曾经在本地高速缓存中驻留的、来自另一个处理单元中远程高速缓存或来自系统存储器的数据时，也可能会引起长时间的访问延滞。以上表明，需要提供某种方法和系统，它要能既保持 SMP 数据处理系统中的存储相关性，同时又减少由于高速缓存项目的交叉无效而对系统效能带来的损失。

本发明还认为，常规 MESI 协议并不支持有效地检索远程高速缓存中存储的数据。某些已知的多处理器数据处理系统支持所谓的修改干预（modified intervention），而修改干预允许存储着处于修改状态的数据的远程高速缓存响应读请求而提供修改数据（modified data）。尽管如此，MESI 协议并没有规定当被请求数据处于共享状态时将数据提供到另一个处理单元的责任。因此，尽管被请求数据可以以相对短的访问延滞存储在多个远程处理单元中，但在常规多处理器数据处理系统中，共享数据总是必须从系统存储器检索。因此，需要提供一种支持共享干预（shared intervention）的改进高速缓存相关性协议。

在支持修改干预的常规多处理器数据处理系统中，当修改数据每

次通过共享总线被提供到另一个高速缓存时，修改数据就被回写到系统存储器。本发明认为，尽管这种过程确实能保持系统存储器与高速缓存之间相关性，它却为执行一次本来无必要的写操作而要耗费系统存储器有限带宽的一部分。例如，如果某修改高速缓冲存储器行
5 (modified cache line)被第一个高速缓存因响应读请求而提供到第二个高速缓存，第二个高速缓存中的数据有可能被再次修改，因此要求另一次写向系统存储器。如果在第二个高速缓存发出读请求后至该高速缓冲存储器行随后被修改的这段时间里，没有其它对该高速缓冲存储器行的请求发出，则伴随修改干预而由第一个高速缓存向系统存
10 储器的写操作纯属多余。

减少不必要的向系统存储器的回写操作的一种可能方法是，即使请求高速缓存只提出一次读请求，也将源高速缓存中的数据标记为无效，将请求高速缓存中的数据标记为修改。这样，被请求数据就不必由于修改干预而被回写系统存储器。然而，这种保持高速缓存相关性的
15 的方法并不允许在处理器之间共享更新数据，而如果在高速缓存之间传送数据又会增加总线通信量。因此需要提供一种改进高速缓存相关性协议，它要既支持高效率地将数据回写系统存储器又允许高速缓存之间共享更新数据。

发明内容

20 因此，本发明的一个目的是，提供一种改进的数据处理方法和系统。

本发明的另一个目的是，提供一种在多处理器数据处理系统中保持高速缓存相关性的方法和系统。

25 本发明的再一个目的是，提供一种用于含有多层次分级存储器体系的数据处理系统的高速缓存的改进相关性协议。

以下说明上述目的的实现方法。根据本发明一个方面，提供一种数据处理系统，它包含：

一个互连部件；

一个与所述互连部件相连的系统存储器；

30 与所述互连部件相连的第一群集和第二群集，其中：

所述第一群集和所述第二群集各自至少有一个处理器、一个高层次高速缓存、一个低层次高速缓存，其中所述第一群集的第一高层次

高速缓存和所述第二群集的第二高层次高速缓存各自同时包括：

数据存储器；

标记存储器，用于存储表示特定地址的地址标记，该地址标记与所述数据存储器中的某数据项关联；

- 5 相关性指示符，设置到第一状态，该状态表示所述地址标记有效而所述关联数据项无效；

高速缓存控制器，响应与由所述数据处理系统中另一设备启动的所述特定地址关联的数据传输的检测，该数据传输包括第二数据项，该控制器用所述第二数据项替换所述数据项，并将所述相关性指示符
10 设置到第二状态，其中，包含所述高速缓存控制器的高层次高速缓存存在其群集中响应一个请求，通过共享干预，提供所述第二数据项。

根据本发明的另一发明，提供一种保持数据处理系统中高速缓存相关性的方法，其中的数据处理系统包括互连部件、多个高速缓存和多个处理器，该多个高速缓存和多个处理器至少被划分成第一、第二
15 两个群集，所述第一和第二群集中每个群集与所述互连部件连接、并且至少有一个高层次高速缓存和至少有一个低层次高速缓存，该方法包括：

在所述第一群集的第一高层次高速缓存中，将第一数据项与表示某地址的某地址标记关联存储，
20 并将第一相关性指示符设置到第一状态，该第一状态表示所述地址标记有效、所述第一数据项无效；

在所述第二群集的第二高层次高速缓存中，当所述第一高层次高速缓存正以所述第一状态存储所述第一数据项时，将第二数据项与表示所述地址的某地址标记关联存储，并将第二相关性指示符设置到所述
25 第一状态，该第一状态表示所述地址标记有效、所述第二数据项无效；响应所述第一和所述第二高层次高速缓存二者之一，检测与由所述数据处理系统中另一设备启动的所述地址关联的数据传输，该数据传输包括第三数据项，在检测所述第一和所述第二高层次高速缓存的一个中用所述第三数据项替换所述第一和第二数据项中相应的一
30 个，并将所述第一和第二相关性指示符中相应的一个设置到第二状态，其中，所述检测所述第一和所述第二高层次高速缓存中的一个是在其群集中，响应一个请求，通过对另一个高速缓存共享干预，决定

提供所述第三数据项。

与此类似，在第二群集的高层次高速缓存中，第二数据项与表示该特定地址的一个地址标记关联存储。此外，第二群集的高层次高速缓存中的相关性指示符被设置到第一状态。这样，数据处理系统执行的
5 的相关性协议，允许将第一、第二两个群集的高层次高速缓存中的相关性指示符都设置到第一状态。

本发明的上述及其它目的、特点和优点显见于以下的详细说明中。

附图说明

10 本说明书后附的权利要求中陈述了本发明特有的新颖特点。然而，欲透彻理解本发明本身以及它的一种优选使用方式及其进一步的目的和优点，则要参阅以下结合各附图对示例性实施例的说明。附图简介如下：

15 图 1 表示按照本发明的多处理器数据处理系统的一个示例性实施例。

图 2 是表示按照本发明的高速缓存的一个示例性实施例的框图。

图 3 是表示 HR-MESI 高速缓存相关性协议的一个示例性实施例的状态图。

20 图 4 是表示 HRT-MESI 高速缓存相关性协议的一个示例性实施例的状态图。

图 5 是表示用派生 H 状态来保持 N 层高速缓存相关性历史数据的状态图。

25 图 6 表示包括了一个能很好地应用本发明的高速缓存相关性协议的多层次分级存储器体系的多处理器数据处理系统的一个示例性实施例。

具体实施方式

现在参见各附图，特别是图 1，该图用高度抽象的框图表示了按照本发明的一个多处理器数据处理系统。如图所示，数据处理系统 8
30 中包括许多处理器 10a ~ 10n，每个处理器最好是 IBM 公司生产的 PowerPC™ 系列的某种处理器。除了常规寄存器、指令流逻辑和用来执行程序指令的执行单元之外，处理器 10a ~ 10n 中每个处理器还包括一个相联的第一层次的 (L1) 机载高速缓存 12a ~ 12n，用于暂时存储

相联处理器可能要访问的指令和数据。图1中将L1高速缓存12a~12n一律表示为既存储指令又存储数据(本文以后将指令和数据统称为数据)的高速缓存,不过本领域的熟练人员明白,L1高速缓存12a~12n的每个高速缓存也都可以选择用分开的指令高速缓存和数据高速缓存来实现。

5 为了尽量降低数据访问延滞,数据处理系统8还包括另外一层或多层高速缓存,例如第二层次的(L2)高速缓存14a~14n,用于分级传递数据到L1高速缓存。换言之,第二层次的(L2)高速缓存14a~14n起着系统存储器与L1高速缓存12a~12n之间中间存储器的作用,与L1高速缓存12a~12n相比,它一般能存储更多的数据量,但访问延滞的时间更长。例如,L2高速缓存14a~14n的存储容量为256或512KB,而L1高速缓存12a~12n的存储容量为64或128KB。处理器10a进一步由第三层次的(L3)后备高速缓存15a支持。后备高速缓存与L2高速缓存14a并联到互连部件16,存储容量最好等于或大于L2高速缓存14a。

10 如图所示,数据处理系统8进一步包括输入输出设备20、系统存储器18和非易失性存储器22,它们都与互连部件16相连。输入输出设备20包括有常规的外部设备,例如显示器、键盘和图形指针,它们都通过常规适配器而与互连部件16相连。非易失性存储器22存储着操作系统和其它软件,它们要随数据处理系统的通电而被写入易失性系统存储器18中。当然,本领域的熟练人员明白,数据处理系统8可以包含未在图1中表示的许多其它部件,例如用于连接网络或附属设备的串行与并行端口,控制对系统存储器18访问的存储控制器,等等。

25 互连部件16中含有一个或多个总线,或者一个交叉点开关,它作为一个管道,用于L2高速缓存14a~14n、L3高速缓存15a、系统存储器18、输入/输出设备20和非易失性存储器18之间的通讯事务。互连部件16上典型的通讯事务包括一个表示事务源的源标记、一个指定事务的接收者的目的地标记、以及地址和/或数据。与互连部件16相连的各设备最好都窥探在互连部件16上的所有通讯事务。

30 现在参见图2,该图是表示按照本发明的L2高速缓存14的一个示例性实施例的更详细的框图。在该示例性实施例中,L2高速缓存

14 是一个使用 32 位地址的 8 分路 (way) 组关联高速缓存。与此相应, L2 高速缓存 14 的数据阵列 34 中包含许多同余类 (congruence classes), 每个同余类含有 8 个用于存储高速缓冲存储器行的分路。如常规的组关联高速缓存中的一样, 利用存储单元地址内的变址位 (例如 32 位地址中的 20~26 位), 就能将系统存储器 18 的存储单元映射到数据阵列 34 内的特定同余类。

在数据阵列 34 中存储的高速缓冲存储器行被记录在高速缓存目录 32 中。对应数据阵列 34 中的每个分路, 该目录都有一个目录项。每个目录项中包含标记域 40、相关性状态域 42、最近最少使用 (LRU—least recently used) 域 44 和蕴含域 46。标记域 40 通过存入高速缓冲存储器行的系统存储器地址中的标记位 (例如 0~19 位), 表明哪个高速缓冲存储器行被存储在数据阵列 34 的对应分路中。相关性状态域 42 利用预定的位组合, 来表示存储在数据阵列 34 的对应分路中的数据的相关性状态 (下文将结合图 3 详细讨论)。LRU 域 44 表示的是, 数据阵列 34 的对应分路, 相对于其同余类的其它分路, 在最近什么时间被访问过, 由此进而指示, 当发生高速缓存失中 (cache miss) 时, 应当将哪个高速缓冲存储器行从该同余类中抛弃。最后是蕴含域 46, 它表示在数据阵列 34 的对应分路中存储的高速缓冲存储器行, 是否也存储在关联的 L1 高速缓存 12 中。

还是参见图 2, L2 高速缓存 14 进一步包括高速缓存控制器 36, 它管理着数据阵列 34 内部数据的存储和检索, 以及根据从相关的 L1 高速缓存 12 接收的信号和在互连部件 16 上窥探到的事务, 对高速缓存目录 32 进行更新。如图所示, 高速缓存控制器 36 中含有一个读队列 50 和一个写队列 52, 高速缓存控制器 32 根据队列的内容, 执行对高速缓存目录 32 的更新以及对数据阵列 34 的访问。例如, 如果从关联的 L1 高速缓存 12 接收到一个读请求, 高速缓存控制器 36 就将该读请求放置到读队列 50 内的一个单元上。高速缓存控制器 36 为该读请求所作的服务是, 将所请求的数据提供给关联的 L1 高速缓存 12, 然后将该读请求从读队列 50 中删除。另一个例子是, 高速缓存控制器 36 窥探到由 L2 高速缓存 14a~14n 中的另一个高速缓存启动的一个事务, 这表明某远程处理器 10 打算修改其对某高速缓冲存储器行的本地拷贝的内容。当窥探到该事务时, 高速缓存控制器 36 就在读

队列 50 中放置一个要求读高速缓存目录 32 的请求，目的是确定该高速缓冲存储器行是否驻留在数据阵列 34 中。如果是，高速缓存控制器 36 就在互连部件 16 上放置一个适当的应答，必要的话，还在写队列 52 中插入一个目录写请求—如果目录写请求得到服务，就更新该高速缓冲存储器行相关的相关性状态域。尽管图 2 所示的实施例只使用了一个读队列和一个写队列，应当明白，高速缓存控制器 36 所采用的队列的数目只是一个设计选择问题，而且高速缓存控制器 36 还可以采用独立的队列分别用于对高速缓存目录的访问和对数据阵列的访问。

10 L3 高速缓存 15a 与图 2 所示的 L2 高速缓存 14 的结构几乎相同，只有很少的例外。具体来说，L3 高速缓存 15a 的高速缓存控制器并不是通过高速缓存间链路与 L2 高速缓存 14a 相连，而是与 L2 高速缓存 14a 并联到互连部件 16，使得 L3 高速缓存 15a 的高速缓存控制器能窥探到互连部件 16 上由 L2 高速缓存 14a 发出的所有事务。

15 参见图 3。该图表示了根据本发明的 HR-MESI 高速缓存相关性协议的一个示例性实施例的状态图。所示 HR-MESI 协议，最好由分级存储器体系中最低层次的联机 (in-line) 高速缓存 (例如 L2 高速缓存 14a ~ 14n)，以及稍作变化，由较低层次的后备高速缓存 (例如 L3 高速缓存 15a) 来执行。较高层次的高速缓存最好执行诸如 MESI 协议 H-MESI 协议、或 R-MESI 协议之类的协议子集。本文的引证文献—美国专利申请系列号 09/024610 描述了 H-MESI 协议；美国专利申请系列号 08/839,548 描述了 R-MESI 协议。然而应当明白，数据处理系统 8 的另选方案的实施例中，可以以增加高速缓存间通信量为代价，在分级存储器体系中任何层次的高速缓存上执行 HR-MESI 协议。

25 如图 3 所示，HR-MESI 高速缓存相关性协议包括 MESI 协议的 4 个常规状态：修改 (M)、排他 (E)、共享 (S)、无效 (I)，图中分别用引用号 80、82、84、86 表示。此外，本发明的 HR-MESI 高速缓存相关性协议还包括空悬 (H) 状态 90 和 R 状态 92。空悬状态表示，在关联的标记域 40 中存储的地址标记是有效的，但是在数据阵列 34 的对应分路中存储的数据项 (例如高速缓冲存储器行或高速缓存区段) 却是无效的。R 状态表示，在存储某数据项的拷贝的多个高速缓存中，哪个高速缓存在离当前最近的时间通过互连部件 16 上的某个

事务接收了该共享数据项。

在一个最佳实施例中，在系统通电时，各 L2 高速缓存目录 32 的各目录项的相关性状态域 42 被初始化到 I 状态 86，表示标记域 40 以及在数据阵列 34 的对应分路中存储的数据项均为无效。与此类似，
5 L1 和 L3 高速缓存的目录项的初始化状态也是无效状态。然后，在 L2 高速缓存 14a~14n 的其中之一、或在处于无效状态 86 的 L3 高速缓存 15a 中存储的高速缓冲存储器行(或高速缓存区段)的相关性状态，可以被更新到 M 状态 80、E 状态 82 或 R 状态 92 的其中某一个状态，具体状态取决于处理器 10a~10n 提出的存储器请求的类型以及分级
10 存储器体系对这些请求的响应。

例如，如果处理器 10a 根据某个加载指令而提出读请求，L1 高速缓存 12a 首先判断被请求数据是否驻留在 L1 高速缓存 12a。如果在 L1 高速缓存 12a 中命中该数据，L1 高速缓存 12a 就将该被请求数据提供给处理器 10a。然而如果在 L1 高速缓存 12a 中没有命中该数据，
15 L1 高速缓存 12a 就通过高速缓存间连路，将该读请求传递到 L2 高速缓存 14a。如果在 L2 高速缓存 14a 中命中该数据，L2 高速缓存 14a 就将该被请求数据提供给 L1 高速缓存 12a，后者将该被请求数据结合适当的相关性状态存储起来，并将该被请求数据传递到处理器 10a。

然而，如果该读请求在 L1 高速缓存 12a 和 L2 高速缓存 14a 中都没有命中，则 L2 高速缓存 14a 的高速缓存控制器 36 就将读请求作为一个事务传送到互连部件 16 上，L2 高速缓存 14b~14n 及 L3 高速缓存 15a 都能窥探到该事务。根据对互连部件 16 上的该读请求的窥探，L2 高速缓存 14b~14n 中每个的高速缓存控制器 36 就要判断，被请求数据是驻留在其数据阵列 34 中还是驻留在 L1 高速缓存 12b~12n 中
20 一个关联的高速缓存中。L3 高速缓存 15a 的高速缓存控制器 36 类似地判断读请求是否在 L3 高速缓存目录中命中。如果 L2 高速缓存 14b~14n、L1 高速缓存 12b~12n 及 L3 高速缓存 15a 都没有存储被请求数据，则 L2 高速缓存 14b~14n 及 L3 高速缓存 15a 每个都返回一个空应答 (null response) 给 L2 高速缓存 14a，后者然后就向系统存储器 18 请求该数据。当被请求数据从系统存储器 18 返回到 L2 高速
25 缓存 14a，高速缓存控制器 36 就将被请求数据传递到 L1 高速缓存 12a，在其数据阵列 34 中存储该请求数据，将与存储被请求数据的分

路相关的相关性状态域 42 从 I 状态 86 更新为 E 状态 82—如引用号 100 所示。如常规的 MESI 协议一样，E 状态 82 表示相关的高速缓冲存储器行有效，并且不驻留在分级存储器体系中第二层次的任何其它高速缓存中。L3 高速缓存 15a 类似地提取被请求数据并将被请求数据以 E 状态 82 储存起来。

如果由处理器 10a 请求的数据没有驻留在 L1 高速缓存 12a、L2 高速缓存 14a、或 L3 高速缓存 15a 中，但却存储在例如 L1 高速缓存 12n 中，处于 M 状态 80，则 L2 高速缓存 14n 的高速缓存控制器 36 根据读请求产生一个修改应答，指示 L1 高速缓存 12n 将该被请求数据压入 L2 高速缓存 14n。以后，L2 高速缓存 14n 将在互连部件 16 上提供该被请求数据。根据读请求，L1 高速缓存 12n 和 L2 高速缓存 14n 中的被请求数据的相关性状态被更新为 S 状态 84—如引用号 102 所示。根据本发明，L2 高速缓存 14a 中存储被请求数据的分路的相关性状态从 I 状态 86 转换到 R 状态 92—如引用号 104 所示。如上所述，R 状态 92 表示，其关联的数据是最晚被 L2 高速缓存 14a 引用的，如果窥探到互连部件 16 上有对该数据的请求，则 L2 高速缓存 14a 要提供该数据。当 L3 高速缓存 15a 窥探到该被请求数据从系统存储器 18 返回时，L3 高速缓存 15a 提取被请求数据，将其以 S 状态 84 存储起来。图中引用号 105 所示的 I 状态 86 与 S 状态 84 之间转换是用虚线表示的，虚线表示这种转换仅适用于 L3 高速缓存 15a。本领域的熟练人员由以上的例子会明白，在不支持修改干预的其它选择方案的本发明实施例中也能作同样的状态转换，然而，在这些其它选择方案的实施例中，被请求数据是从系统存储器 18 获得的，其代价是增加了访问延滞。

L2 高速缓存 14a 在收到因其读请求而产生的共享干预应答或共享应答时，类似地进行一次由 I 状态 86 到 R 状态 92 的转换。如果 L2 高速缓存 14b~14n 之一，例如 L2 高速缓存 14n 存储着 R 状态 92 或者 E 状态 82 的被请求数据，则 L2 高速缓存 14a 会收到一个共享干预应答（以及被请求数据）。根据读请求，L2 高速缓存 14n 的高速缓存控制器 36 将其对被请求数据的拷贝的相关性状态由 R 状态 92 更新到 S 状态 84—如引用号 106 所示。如果 L2 高速缓存 14b~14n 都没有存储着 R 状态 92 的被请求数据，但至少有一个存储着 S 状态 84 的被请

求数据, 则存储着 S 状态 84 的被请求数据的 L2 高速缓存 14 用一个共享应答来响应该读请求, 将其对被请求数据的拷贝保留在 S 状态 84。如果只收到一个共享应答, L2 高速缓存 14a 就从系统存储器 18 检索被请求数据, 并将被请求数据以 R 状态 92 存储起来。在这两种情况, L3 高速缓存 15a 采样被请求数据, 并将其以 S 状态 84 存储。

如果 L2 高速缓存 14a 在互连部件 16 上发出一个“意欲修改的读”请求, 表示处理器 10a 需要排他使用某存储单元, 目的是对其修改, 如上所述, 该请求数据可以从 L3 高速缓存 15a、某个远程 L2 高速缓存 14 或者系统存储器 18 获得。然而当被请求高速缓冲存储器行被获得时, L1 高速缓存 12a 和 L2 高速缓存 14a 以 M 状态将被请求高速缓冲存储器行存储起来—如引用号 107 所示。此外, 由于“意欲修改的读”请求事务指出, 其它对被请求高速缓冲存储器行的拷贝将变得过时无用, 所以 L3 高速缓存 15a 和远程 L1 和 L2 高速缓存必须指示, 它们对被请求高速缓冲存储器行的拷贝是无效的。在 L1 高速缓存 12b ~ 12n 中, 所有对被请求高速缓冲存储器行的拷贝都被简单地标记为无效。然而, 在 L3 高速缓存 15a 和 L2 高速缓存 14b ~ 14n 中存储的对被请求高速缓冲存储器行的拷贝的相关性状态, 却不像在使用交叉无效 (X1) 的常规多处理器数据处理系统中的那样被更新到 I 状态 86。相反, 根据本发明的一个重要方面, 存储着对被请求高速缓冲存储器行的拷贝的 L3 高速缓存 15a 和 L2 高速缓存 14b ~ 14n 中的各高速缓存, 要将其拷贝关联的相关性状态域 42, 从 R 状态 92、S 状态 84、M 状态 80 或 E 状态 82 中的任何状态更新到 H 状态 90—分别如引用号 108、110、112、114 所示。如上所述, H 状态 90 表示, 标记域 40 中的标记保持有效, 而数据阵列 34 内关联的高速缓冲存储器行是无效的。L2 和 L3 高速缓存目录中的目录项, 根据窥探到的其它要求使数据无效的事务—包括删除 (kills, 即明确地无效掉特定数据块的事务)、清仓 (flushes, 即无效掉特定数据块并将修改数据拷贝到系统存储器的事务) 放弃 (dclaims, 即响应某个因存储而被修改的本地高速缓冲存储器行的拷贝, 无效掉各远程高速缓存中标记为共享的该高速缓冲存储器行的拷贝的事务), 也类似地被更新到 H 状态 90。

如参考号 116 和 118 所示, 高速缓存目录项根据高速缓存接收的

事务的类型,可以由H状态90分别转换到E状态82或M状态80。例如,处于H状态90的L2高速缓存14a的某目录项,在提出读请求的处理器10a(在L1高速缓存12a和L2高速缓存14a中都没有命中后)收到L3高速缓存15a和L2高速缓存14b~14n发来的空应答时,转换到E状态82。E状态82表示,从系统存储器18检索的数据,将只被存储到L2高速缓存14a~14n中的L2高速缓存14a中。另一方面,如果处理器10a提出意愿,要向L1高速缓存12a的某个处于H状态90的分路存储数据,则L1高速缓存12a向L2高速缓存14a指出该意愿,后者然后在互连部件16上发出一个“意欲修改的读”事务。

如上所述,当窥探到“意欲修改的读”事务时,在L3高速缓存15a和L2高速缓存14b~14n中存储的被请求高速缓冲存储器行的拷贝被更新到H状态90,在L1高速缓存12b~12n中存储的被请求高速缓冲存储器行的拷贝被标记为无效。一旦获得被请求高速缓冲存储器行,L2高速缓存14a和L1高速缓存12a中的该高速缓冲存储器行的相关性状态就被设置到M状态80,表示该高速缓冲存储器行有效,但与系统存储器18不相干。

H状态90的高速缓存目录项也可以被更新到S状态84—如引用号120所示。如上所述,L2高速缓存14a~14n每个都能窥探一切在互连16上发出的事务。如果L2高速缓存14a~14n其中之一,例如L2高速缓存14a,窥探到由L2高速缓存14b~14n中的另一个发出的事务含有L2高速缓存14a存储的H状态90的数据的更新(即有效)拷贝,则L2高速缓存14a的高速缓存控制器36从互连部件16提取该数据,将该窥探数据存储到数据阵列34,并将其关联的相关性状态域42由H状态90更新到S状态84。这样,本发明允许将高速缓存目录项,在相关的处理器10没有提出数据请求、或高速缓存没有在互连部件16上启动事务的情况下,由H状态90更新到S状态84。当然,在上述情形中,如果需要应答来保持相关性,则L2高速缓存14a也向被窥探的事务提供应答。例如,如果被窥探的事务是个读请求,则L2高速缓存14a必须提供一个共享应答,表示其要提取被请求数据的意愿,使得提出请求的L2高速缓存以R状态92而不是E状态82存储被请求数据。能按这种方式窥探、以刷新与某有效地址标记关联的无效数据的互连部件16上的事务包括读事务、写事务、因高速缓冲

存储器行的丢弃而向系统存储器 18 的数据回写，等等。

H 状态 90 的高速缓存目录项，也可以根据许多不同的请求/响应情况，被更新到 R 状态 92。例如，如果处理器 10a 提出的读请求在 L1 高速缓存 12a 中没有命中，导致在 L2 高速缓存 14a 命中一个标记，并且如果该匹配标记相关的相关性状态域 42 处于 H 状态 90，则 L2 高速缓存 14a 的高速缓存控制器 36 在互连部件 16 上发出一个读事务。如果从 L3 高速缓存 15a 和 L2 高速缓存 14b~14n 收到对该读事务的共享干预应答、共享应答或修改应答，则 L2 高速缓存 14a 的高速缓存控制器 36 将被请求数据相关的相关性状态域 42 由 H 状态 90 更新到 R 状态 92—如引用号 122 所示。

更重要的是，高速缓存目录项，在相关的处理器没有请求数据的情况下，也能被更新到 R 状态 92。例如，在由处理器 10a 提出“意欲修改的读”请求后，L1 高速缓存 12a 和 L2 高速缓存 14a 以 M 状态 80 存储被请求数据，同时将该高速缓冲存储器行在 L3 高速缓存 15a 和 L2 高速缓存 14b~14n 中的拷贝的相关性状态设置到 H 状态 90。如果例如因为读漏失 (read miss)，随后从 L2 高速缓存 14a 选择被请求数据作为替代，则 L2 高速缓存 14a 的高速缓存控制器 36 在互连部件 16 上发出一个重定位写 (deallocation write) 事务，以便将被请求数据存储在系统存储器 18。当 L3 高速缓存 15a 的高速缓存控制器 36 窥探到这个写事务时，就提取数据，将数据存储在它的 L3 数据阵列，并将相关的相关性状态域由 H 状态 90 更新到 R 状态 92，这样，L3 高速缓存 15a 以后就能在收到读或“意欲修改的读”请求时作为数据源来提供该数据。

下面的表 I，总结了图 3 所示的 HR-MESI 高速缓存相关性协议的示例性实施例中所能进行的状态转换。

表 I

状态 转换	原因	注
I → E	处理器读得到空应答	“处理器读”指相关 (即本地) 处理器的读请求
I → R	(1) 处理器读得到共享干预 (ShI) 应答;	

	(2) 处理器读得到共享应答; (3) 处理器读得到修改应答	
I → S	处理器读得到除了空以外的应答	仅指 L3 后备高速缓存; L2 高速缓存转换到 R, 如上所述
I → M	处理器“意欲修改的读” (rwitm)	
E → S	窥探读	提供干预数据
E → M	处理器“意欲修改的读”	
E → H	窥探数据无效	窥探数据无效包括“意欲修改的读”、放弃、清仓、删除等
R → S	窥探读	提供干预数据
R → M	处理器“意欲修改的读”	在互连部件上发出放弃
R → H	窥探数据无效	
S → M	处理器“意欲修改的读”	在互连部件上发出放弃
S → H	窥探数据无效	
M → S	窥探读	如果支持修改干预, 提供数据
M → H	窥探数据无效	如果窥探事务是“意欲修改的读”, 则如果支持修改干预, 就提供数据
H → E	处理器读得到空应答	
H → R	(1) 处理器读得到共享干预应答; (2) 处理器读得到共享应答; (3) 处理器读得到修改应答; (4) 窥探本地高速缓存控制器发出的读	
H → S	窥探远程高速缓存控制器发出的读或写	
H → M	处理器“意欲修改的读”	

根据本发明, 以上描述的 HR-MESI 高速缓存相关性协议可以扩展

到包括支持其它功能的其它状态。例如，现在参见图 4，图中表示了 HRT-MESI 高速缓存相关性协议的一个示例性实施例的状态图。如果该 HRT-MESI 高速缓存相关性协议被执行，它最好用于分级存储器体系中最底层次的联机高速缓存（在图 1 所示的实施例中，包括 L2 高速缓存 14a ~ 14n）。最好所有较低层次的后备高速缓存例如 L3 高速缓存 15a 都执行 HR-MESI 协议，而较高层次的高速缓存最好执行 HR-MESI 协议的子集，例如上述 H-MESI 协议、R-MESI 协议、MESI 协议之一。除了以上描述的 HR-MESI 协议内的状态外，图 4 所示的 HRT-MESI 高速缓存相关性协议还包括一个标记（T）状态 94。美国专利申请系列号 09/024,393 详细描述了该种状态，本文引用其全部描述作为参考。T 状态 94 表示某关联高速缓冲存储器行已经被不是本地处理器的某处理器修改，但是尚未回写系统存储器（即与系统存储器不一致）。这样，本地处理器对待 T 状态 94 就像对待 S 状态 84，对于互连上的窥探来说，T 状态 94 就像 M 状态 80 一样被对待。

如图 4 中引用号 130 和 134 所示，高速缓冲存储器行在读请求收到修改应答时可以从 I 状态 86 转换到 H 状态 90 或 T 状态 94。例如，如果 L2 高速缓存 14a 对 L2 高速缓存 14n 存储的 M 状态 80 的高速缓冲存储器行提出读请求，L2 高速缓存 14n 就通过上述的修改干预将该高速缓冲存储器行提供给 L2 高速缓存 14a。如以前一样，L2 高速缓存 14n 中被请求高速缓冲存储器行的拷贝被更新到 S 状态 84—如引用号 131 所示。另一方面，L2 高速缓存 14a 将被请求高速缓冲存储器行以 T 状态 94 存储。然后，根据 L2 高速缓存 14b 对该修改高速缓冲存储器行的读请求，L2 高速缓存 14a 通过修改干预来提供被请求高速缓冲存储器行并将其对修改高速缓冲存储器行的拷贝的相关性状态更新到 S 状态 84—如引用号 132 所示。这个过程可以再进行多次，使得由最近一次访问修改高速缓冲存储器行的 L2 高速缓存 14 以 T 状态存储该高速缓冲存储器行，而 L2 高速缓存 14a ~ 14n 中其它 L2 高速缓存以 S 状态 84 存储修改高速缓冲存储器行。这样，L2 高速缓存 14a ~ 14n 其中之一被标记表示为暂时负责通过修改干预向其它高速缓存提供该高速缓冲存储器行并将该高速缓冲存储器行回写到系统存储器 18。

重要的是，当存储着某 T 状态 94 的高速缓冲存储器行的 L2 高速

缓存 14 通过修改干预向互连部件 16 提供数据时，系统存储器 18 的窥探机构（例如存储控制器）并不提取和存储该修改数据，相反却不理睬该事务，以释放系统存储器带宽。所以，只有在需要时，处于 T 状态的修改高速缓冲存储器行才会被回写系统存储器 18，例如某 L2 高速缓冲存储器行抛弃（castout）时或者窥探到一个要求将修改高速缓冲存储器行推入系统存储器 18 的事务时。当然，T 状态 94 的修改高速缓冲存储器行也可能永远不会被回写到系统存储器 18。例如设想这种情形，即处理器 10a 发出指定了 L2 高速缓存 14n 中存储的 T 状态 94 的某高速缓冲存储器行的“意欲修改的读”请求。根据该“意欲修改的读”请求，存储着 S 状态 84 的指定高速缓冲存储器行的 L2 高速缓存 14n 或其它 L2 高速缓存 14，分别都将其对指定高速缓冲存储器行的拷贝更新到 H 状态 90—如引用号 135 和 110 所示。由于通过转换到 H 状态 90，高速缓冲存储器行数据被无效，因而就无须将指定高速缓冲存储器行回写到系统存储器 18。

如引用号 138 所示，T 状态 94 的高速缓冲存储器行，也可以根据本地处理器的“意欲修改的读”请求，转换到 M 状态 80。因为本地 L2 高速缓存 14 中存储被请求高速缓冲存储器行，所以就不在互连部件 16 上发出要求数据的请求事务。相反，本地 L2 在互连部件 16 上发出一个放弃事务，指示其它 L2 高速缓存 14 将它们对该被请求高速缓冲存储器行的拷贝无效，方法是将关联的相关性状态更新到 H 状态 90。

下面的表 II，总结了图 4 所示的 HRT-MESI 高速缓存相关性协议的示例性实施例中所能进行的状态转换。

表 II

状态转换	原因	注释
I → E	cpu 读得到空应答	
I → R	(1) cpu 读得到共享干预 (ShI) 应答; (2) cpu 读得到共享应答	
I → T	cpu 读得到修改应答	
I → M	cpu “意欲修改的读” (rwitm)	
E → S	窥探读	提供干预数据

E → M	cpu “意欲修改的读”	
E → H	窥探数据无效	窥探数据无效包括“意欲修改的读”、放弃、清仓、删除等。如果窥探事务是“意欲修改的读”，则应答共享干预并提供数据
R → S	窥探读	提供干预数据
R → M	cpu “意欲修改的读”	在互连部件上发出放弃
R → H	窥探数据无效	如果是“意欲修改的读”，则应答共享干预并提供数据
S → M	cpu “意欲修改的读”	在互连部件上发出放弃
S → H	窥探数据无效	
M → S	窥探读	如果支持修改干预，则提供数据
M → H	窥探数据无效	如果窥探事务是“意欲修改的读”，则如果支持修改干预，提供数据
T → S	窥探读	应答修改并提供干预数据
T → M	cpu “意欲修改的读”	在互连部件上发出放弃
T → H	窥探数据无效	
H → E	cpu 读得到空应答	
H → R	(1) cpu 读得到共享干预应答； (2) cpu 读得到共享应答； (3) 本地高速缓存控制器发出的窥探写	
H → S	远程高速缓存控制器发出的窥探读或写	
H → T	cpu 读得到修改应答	
H → M	cpu “意欲修改的读”	

进一步增强包括有 H 状态 90 的高速缓存相关性协议的方法是，增加一个或多个表示相关性状态历史数据的派生 H 状态 90。例如，上述的 HR-MESI 和 HRT-MESI 相关性协议可以被修改为增加一个 H_r 状态。与 H 状态 90 一样， H_r 状态也表示关联高速缓冲存储器行无效、关联地址标记有效。 H_r 状态还表示，在其数据无效之前，关联高速缓冲存储器行是以排他状态（即 M 状态 80 或 E 状态 82）或者以能提供干预数据的某个相关性状态（例如 R 状态 92 或 T 状态 94）存储的。例如设想这种情形，即 L2 高速缓存 14a 以 M 状态 80、E 状态 82、R 状态 92 或 T 状态 94 之一存储某高速缓冲存储器行。如果 L2 高速缓存 14n 在互连部件 16 上发出一个对该高速缓冲存储器行的“意欲修改的读”请求，则 L2 高速缓存 14a 就发出一个适当的应答，并将该该高速缓冲存储器行的相关性状态更新到 H_r 状态（而不是 H 状态 90）。然后，当 L2 高速缓存 14n 在互连部件 16 上发出一个将该高速缓冲存储器行回写系统存储器 18 的回写事务（例如高速缓冲存储器行重定位或窥探压入（snoop push））时，L2 高速缓存 14a 就提取和存储该高速缓冲存储器行数据，并将该高速缓冲存储器行的相关性状态更新到 R 状态 92。如此地使用 H_r 状态允许在修改高速缓冲存储器行的重定位后共享干预的保持，由此减少数据访问延滞。

其次，假设 L2 高速缓存 14a 存储着某高速缓冲存储器行的拷贝，其状态是 H_r 状态；L2 高速缓存 14n 存储着同一高速缓冲存储器行的拷贝，其状态是允许 L2 高速缓存 14n 提供干预数据的某个状态。如果 L2 高速缓存 14b 在互连部件 16 上发出一个对该高速缓冲存储器行的“意欲修改的读”请求，L2 高速缓存 14n 就发出一个适当的应答，通过干预来提供被请求高速缓冲存储器行，然后将其对被请求高速缓冲存储器行的拷贝的相关性状态更新到 H_r 状态。根据窥探的含有更新高速缓冲存储器行的干预应答，L2 高速缓存 14a 将其对被请求高速缓冲存储器行的拷贝的相关性状态从 H_r 状态更新到 H 状态 90。另外也可以如图 5 所示，通过在顶级派生状态 H_{r1} 与 H 状态 90 之间插入一个或多个附加派生状态（ H_{r2} 或 H_{r3} ）来保留附加层次的历史数据。这样，当窥探到 L2 高速缓存 14n 发出的干预应答时，L2 高速缓存 14a 不是将状态转换到 H 状态 90，而是由 H_{r1} 状态转换到 H_{r2} 状态。然后，当窥探到含有相关高速缓冲存储器行的回写事务时，L2 高速缓存 14a 会将

其对该高速缓冲存储器行的拷贝的相关性状态升级到 H_{R1} 状态。这样，就可以实现 N 层次的相关性状态历史数据，以增加至少有一个高速缓存能通过干预而提供一个被请求高速缓冲存储器行的可能性。

如图 5 中虚线所示，可以将高速缓存相关性协议改变成包括有 N 层次的相关性状态历史数据的变体。具体来说，当窥探到相关高速缓冲存储器行的回写时，较低层次的历史状态 (H_{R2} 、 H_{R3} ... H_{RN}) 中选定的某些状态或所有状态被直接更新到 S 状态 84，这是因为，以顶层 H_{R1} 状态存储该高速缓冲存储器行的高速缓存将能够通过共享干预来提供 R 状态 92 的该高速缓冲存储器行。

增强包括有至少一个派生 H 状态的高速缓存相关性协议的另一个方法是，支持多层次分级存储器体系内各互连部件上某特定高速缓冲存储器行的一个顶层派生 H 状态 (H_R 或 H_{R1})。现在参见图 6，图中表示了一个实现这种增强的多处理器数据处理系统的典型实施例。如图所示，数据处理系统 150 包括两个对称的群集 152 和 154。群集 152 含有处理器 160a 和处理器 160b，除了处理电路之外，各处理器还包括一个机载 $L1$ 高速缓存 162 和 $L2$ 高速缓存 164。 $L2$ 高速缓存 164a 和 164b 各通过高速缓存间连路 166a 与共享 $L3$ 高速缓存 170a 相连，后者又通过互连部件 180 与系统存储器 182 相连。与此类似，群集 154 含有处理器 160c 和 160d，除了处理电路之外，各处理器还包括一个 $L1$ 高速缓存 162 和 $L2$ 高速缓存 164。 $L2$ 高速缓存 164c 和 164d 各通过高速缓存间连路 166b 与共享 $L3$ 高速缓存 170b 相连，后者又通过互连部件 180 与系统存储器 182 相连。

表 III 中总结了表示在多层次分级存储器体系中使用派生 H 状态的效能优点的例示处理情况。如表 III 第 1 行所示， $L2$ 高速缓存 164a ~ 164h 和 $L3$ 高速缓存 170a ~ 170b 开始操作时，它们各自对某特定高速缓冲存储器行的相关性指示符被设置为 I 状态 86。下一步，如第 2 行所示，根据处理器 160a 提出的读请求， $L2$ 高速缓存 164a 在高速缓存间连路 166a 上发出一个读请求。根据 $L2$ 高速缓存 164b ~ 164d 的空应答， $L3$ 高速缓存 170a 将该读请求放置到互连部件 180 上。系统存储器 182 响应读请求，提供被 $L3$ 高速缓存 170a 与 $L2$ 高速缓存 164a 二者以 E 状态 82 存储的该被请求高速缓冲存储器行。如果处理器 160d 指出有向该同一高速缓冲存储器行存储的意向，则 $L2$ 高速缓存 164d

在高速缓存间连路 166a 上发出一个“意欲修改的读”(RWITM)请求。根据窥探的该 RWITM 请求, L2 高速缓存 164a 作出一个共享干预应答, 提供该高速缓冲存储器行, 并将相关性状态指示符更新到 H_r 状态一如上所述。如表 III 所示, L2 高速缓存 164d 以 M 状态 80 存储被请求高速缓冲存储器行。L3 高速缓存 170a 也类似地转换到 M 状态 80, 以表示上游高速缓存存储着该高速缓冲存储器行的修改拷贝。重要的是, L3 高速缓存 170a 不必在互连部件 180 上放置 RWITM 请求, 这是因为 L3 高速缓存 170a 以排他状态存储着该高速缓冲存储器行。

如表 III 第 4 行所示, 如果处理器 160f 指出有向该高速缓冲存储器行存储的意向, 则 L2 高速缓存 164f 在高速缓存间连路 166b 上发出一个 RWITM 请求。L3 高速缓存 170b 于是将该 RWITM 请求放置到互连部件 180 上。根据窥探的该 RWITM 请求, L3 高速缓存 170a 将该请求放置到高速缓存间连路 166a 上, 因为它“知道”至少有一个上游高速缓存中存储着被请求高速缓冲存储器行的修改拷贝。根据窥探的该 RWITM 请求, L2 高速缓存 164d 应答修改, 通过修改干预提供被请求数据, 并将其对该高速缓冲存储器行的拷贝的相关性状态更新到 H_r 状态。L2 高速缓存 164a 和 L3 高速缓存 170d 也更新它们对该高速缓冲存储器行的拷贝的相关性状态—此例中分别更新到 H 状态 90 和 H_r 状态。L3 高速缓存 170b 和 L2 高速缓存 164f 都以 M 状态 80 存储着被请求高速缓冲存储器行。

然后, 根据 L2 高速缓存 164h 在高速缓存间连路 166b 上发出的 RWITM 请求, L2 高速缓存 164f 应答修改, 提供被请求高速缓存数据, 并将其对被请求高速缓冲存储器行的拷贝的相关性状态更新到 H_r 状态。收到被请求高速缓冲存储器行后, L2 高速缓存 164f 将该高速缓冲存储器行以 M 状态 80 存储。其余高速缓存中存储的该高速缓冲存储器行的相关性状态保持不变。这样, 如表 III 第 5 行所示, 某个高速缓存组有一个高速缓存能保存一个 H_r 状态的特定高速缓冲存储器行, 其中, 高速缓存组的定义是, 其包括的高速缓存 (1) 位于分级存储器体系中的相同层次; (2) 相连到共同的互连部件或高速缓存间连路。

现在继续说明本例。如果 L2 高速缓存 164g 然后在高速缓存间连路 166b 上发出读请求, 则 L2 高速缓存 164h 应答修改, 通过修改干

- 预提供高速缓冲存储器行，并将其对该高速缓冲存储器行的拷贝的相关性状态更新到 S 状态 84。当窥探到该被请求高速缓冲存储器行时，L2 高速缓存 164g 就将该高速缓冲存储器行以 T 状态 94 存储。L3 高速缓存 170b 不必在互连部件 180 上发出任何事务，这是因为，它存储着 M 状态 80 的该高速缓冲存储器行。最后，如表 III 第 7 行所示，当 L2 高速缓存 164g 重定位高速缓冲存储器行时，L2 高速缓存 164g 将关联的相关性状态更新到某个中间状态（以“X”表示），L2 高速缓存 164f 提取该高速缓冲存储器行，并将其相关性状态由 H_r 状态更新到 R 状态 92。
- 10 L3 高速缓存 170b 可以被配置为能以多种方式响应 L2 高速缓存 164g 对高速缓冲存储器行的重定位。首先，L3 高速缓存 170b 可以只简单地刷新其对高速缓冲存储器行的拷贝，而不将修改高速缓冲存储器行写到系统存储器 182。另外，如表 III 第 7 行所示，L3 高速缓存 170b 也可以既用该修改数据更新其对高速缓冲存储器行的拷贝，又通过互
- 15 连部件 180 将修改高速缓冲存储器行写到系统存储器 182，由此使修改高速缓冲存储器行对其它 L3 高速缓存透明。在第二种实现方式中，L3 高速缓存 170a 以共享应答来响应重定位写，以表示它要提取该高速缓冲存储器行数据。然后，在提取该高速缓冲存储器行数据时，L3 高速缓存 170a 并将其对该高速缓冲存储器行的拷贝的相关性状态由
- 20 H_r 状态更新到 R 状态 92，从而有利地保持了通过共享干预在分级存储器体系的互连层次上提供该高速缓冲存储器行的能力。如表中所示，L3 高速缓存 170b 将其对该高速缓冲存储器行的拷贝的相关性状态更新到 S 状态 84。

表 III

动作	L2 高速缓存 164					L3 高速缓存 170	
	a	d	f	g	h	a	b
初始状态	I	I	I	I	I	I	I
L2 164a 读	E	I	I	I	I	E	I
L2 164d RWITM	H _r	M	I	I	I	M	I
L2 164f RWITM	H	H _r	M	I	I	H _r	M
L2 164h RWITM	H	H _r	H _r	I	M	H _r	M

L2 164g 读	H	H _R	H _R	T	S	H _R	M
L2 164g 重定位	H	H _R	R	X	S	R	S

如上所述，本发明提供一种用于保持多处理器数据处理系统中高速缓存相关性的改进方法和装置。本发明提供的改进高速缓存相关性协议能在本地处理器不明确发出读或写请求的情况下，将与有效地址标记关联存储在某高速缓存中的无效数据项自动更新为有效数据项。以这种方式，因远程处理器的操作而无效的数据能在本地处理器访问之前被刷新，由此实质性地减少数据访问延滞，这是因为不需要从远程高速缓存或系统存储器中检索数据。实质性地减少数据访问延滞的另一个原因是，本发明的高速缓存相关性协议支持共享干预，即不从系统存储器而是从远程高速缓存提供共享状态下存储的数据。

根据本发明的另一个方面，高速缓存相关性协议可以扩展成包括一个标记（T）状态，通过分配向系统存储器回写修改数据的责任，提高系统存储器带宽的使用效率。根据本发明的高速缓存相关性协议还可以通过蕴含一个或多个派生状态，保持相关性状态历史数据。这种派生状态能根据第一个高速缓存对含有数据的某高速缓冲存储器行的重定位，将通过干预而提供数据的责任由第一个高速缓存转移到第二个高速缓存。这些派生状态在包括有多个互连部件或高速缓存间连路的多层次分级存储器体系中特别有用，这是因为，与各个部件互连或高速缓存间连路相连的一个高速缓存被赋予通过干预而提供数据的责任。

以上结合示例性实施例具体地表示和叙述了本发明，然而本领域的熟练人员明白，在不偏离本发明的实质和范围的前提下，对实施例的形式和内容可以有各种改变。例如，可以修改 HR-MESI 和 HRT-MESI 高速缓存相关性协议，删除其中的 I 状态 86，因为该状态仅仅用于在通电时初始化目录项，其它状态永远不会转换到这个状态。如果将 I 状态 86 删除，在通电时各 L2 目录项的相关性状态域就被初始化为 H 状态 90，各 L2 目录项的标记域就被初始化成一个至少在同一同余类中是唯一的标记值。

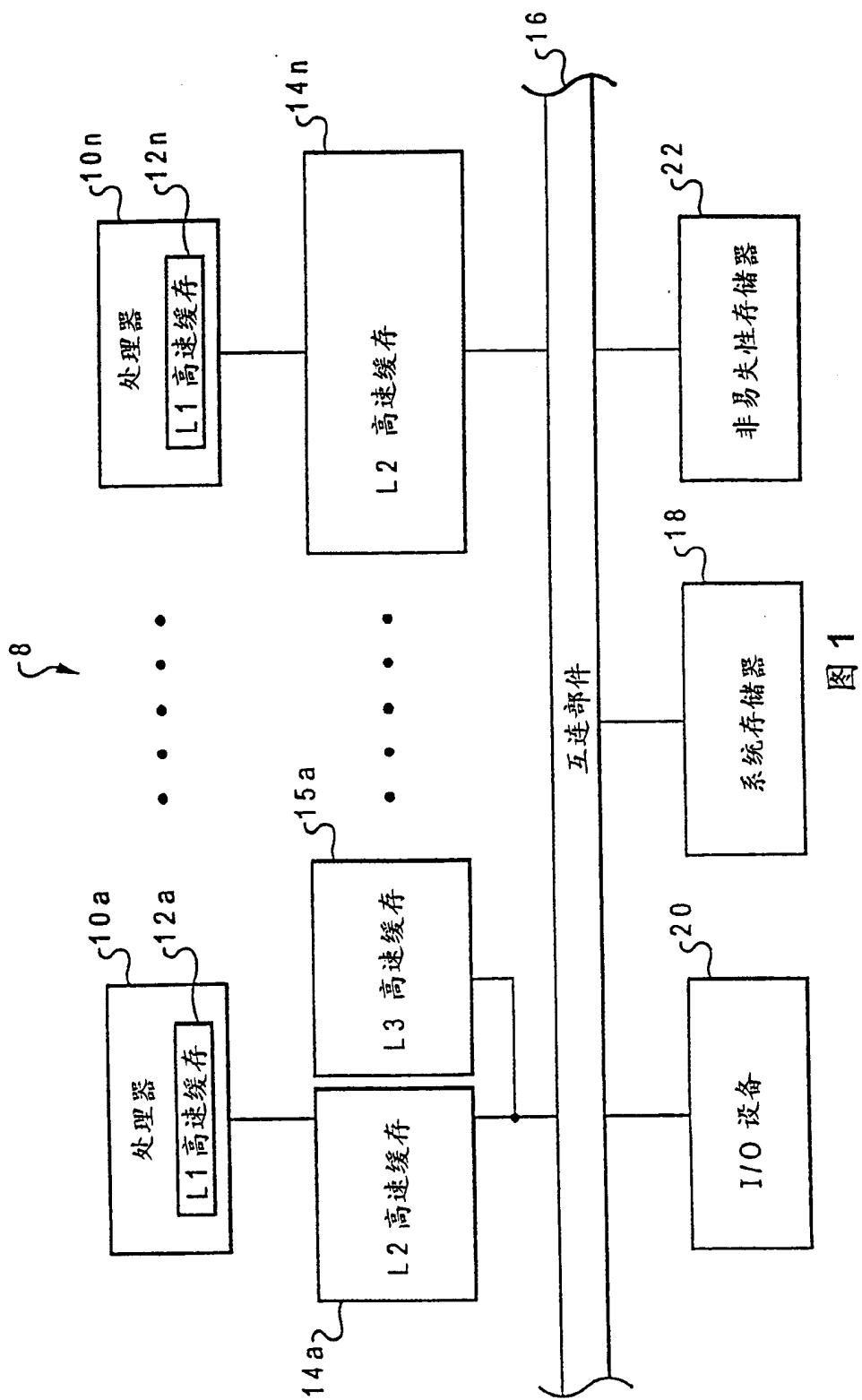


图 1

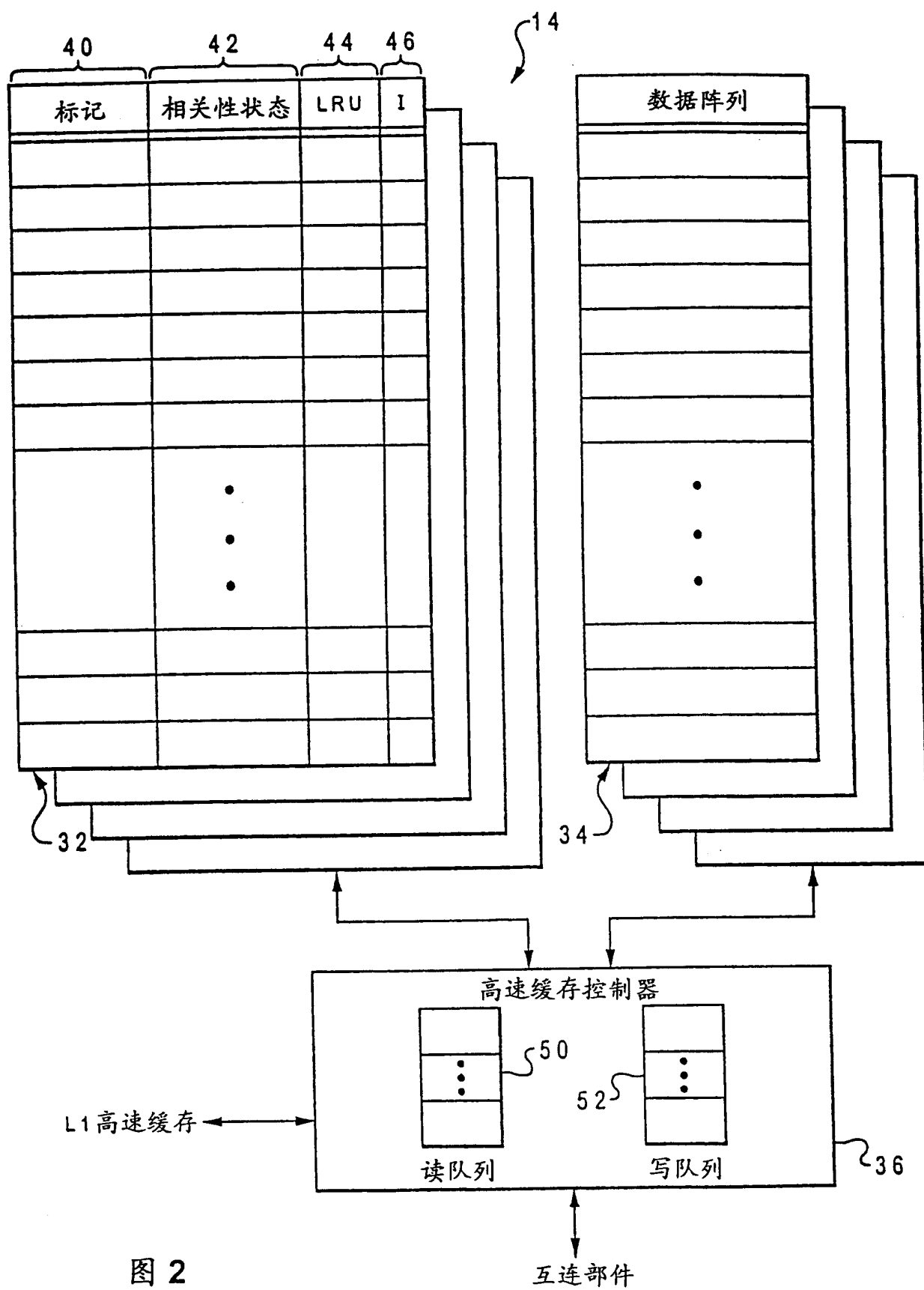


图 2

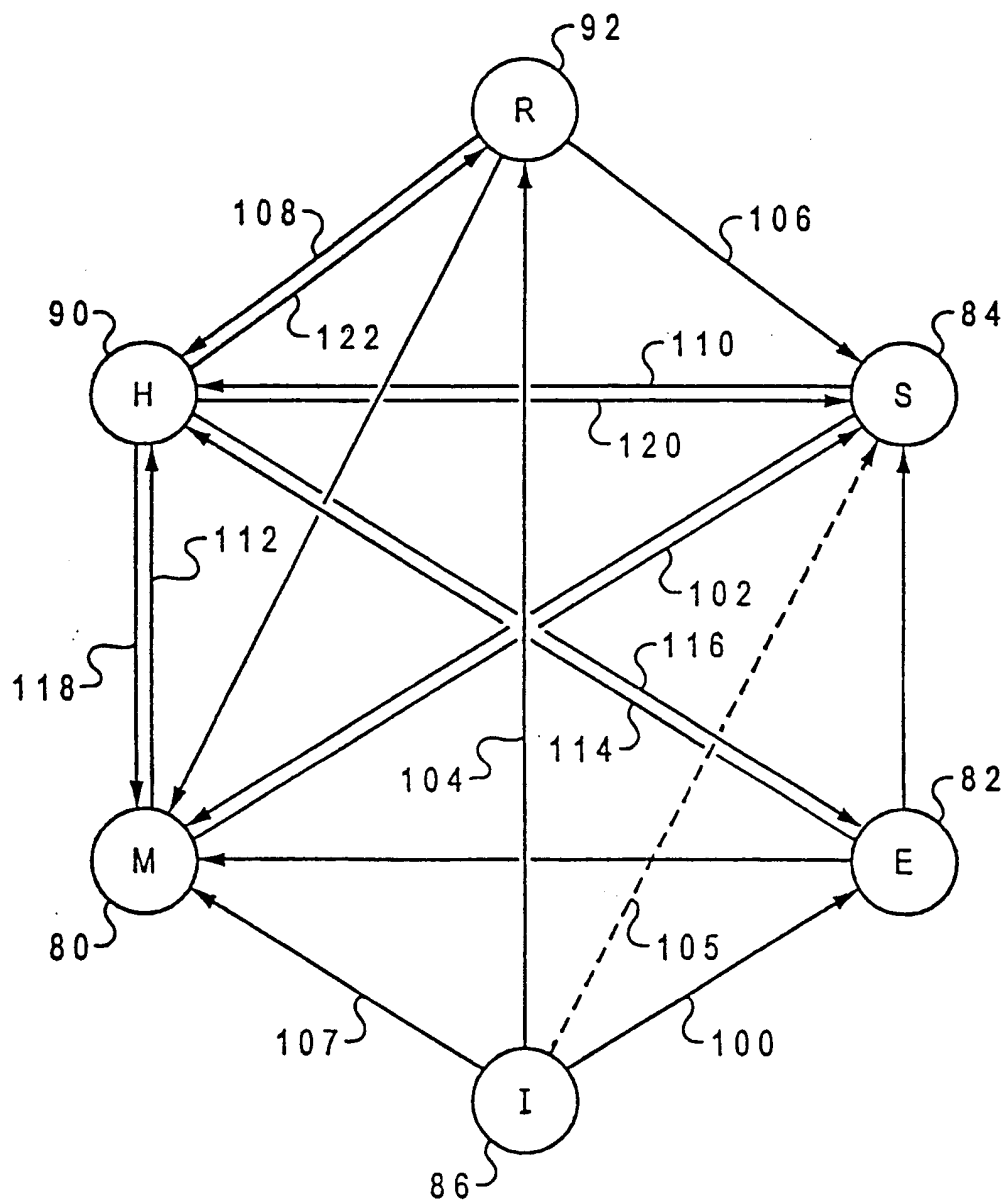


图 3

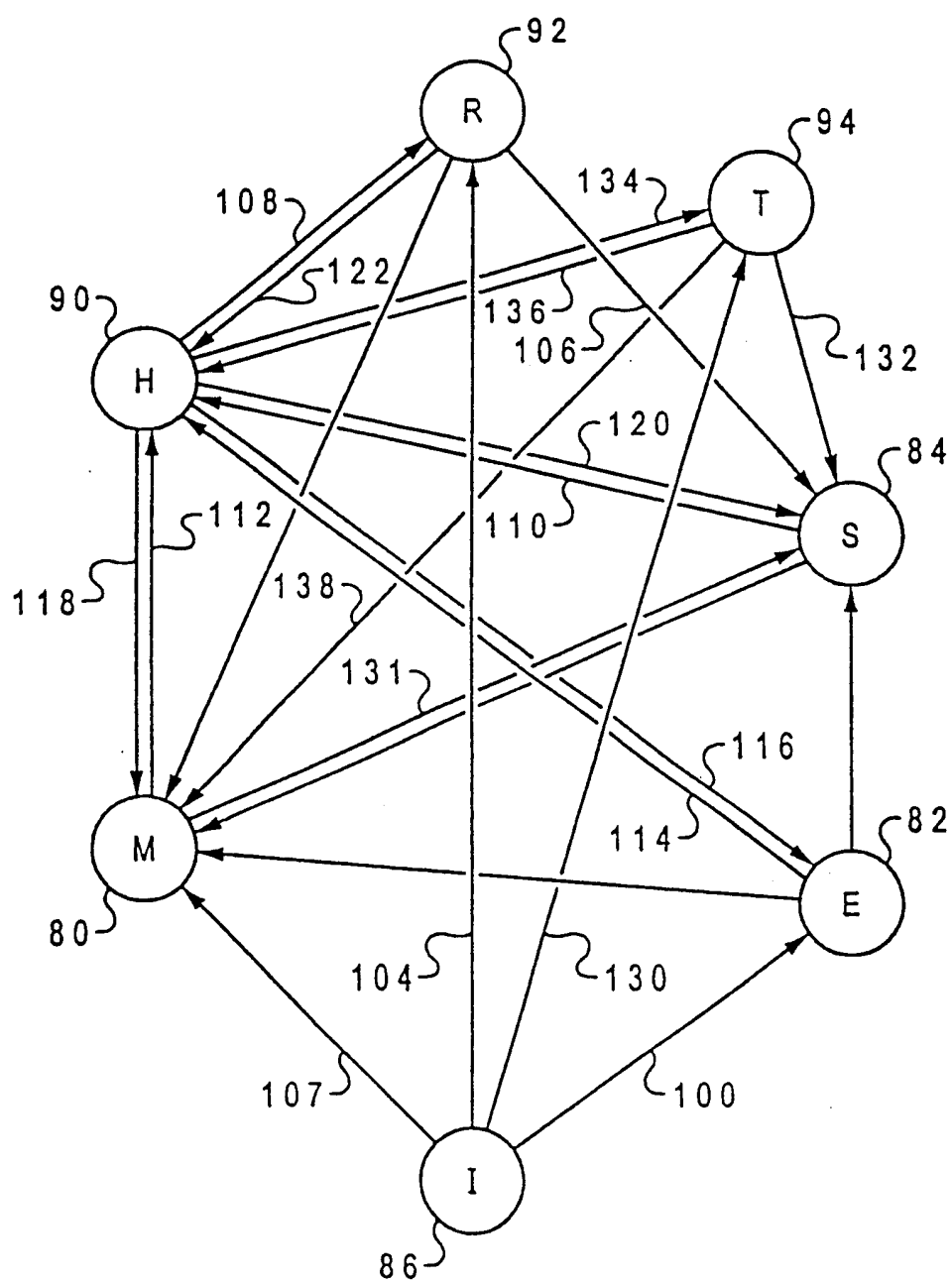


图 4

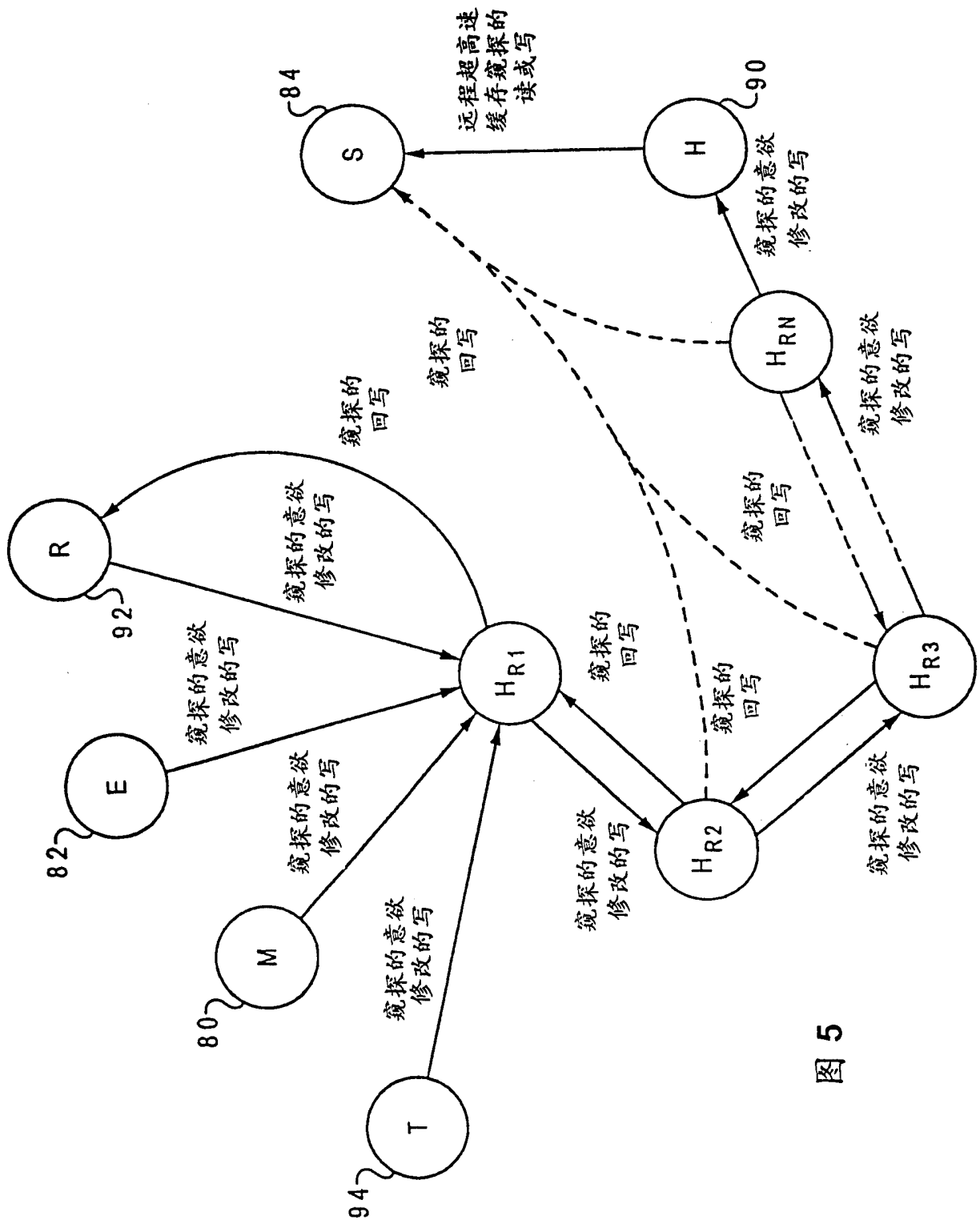


图 5

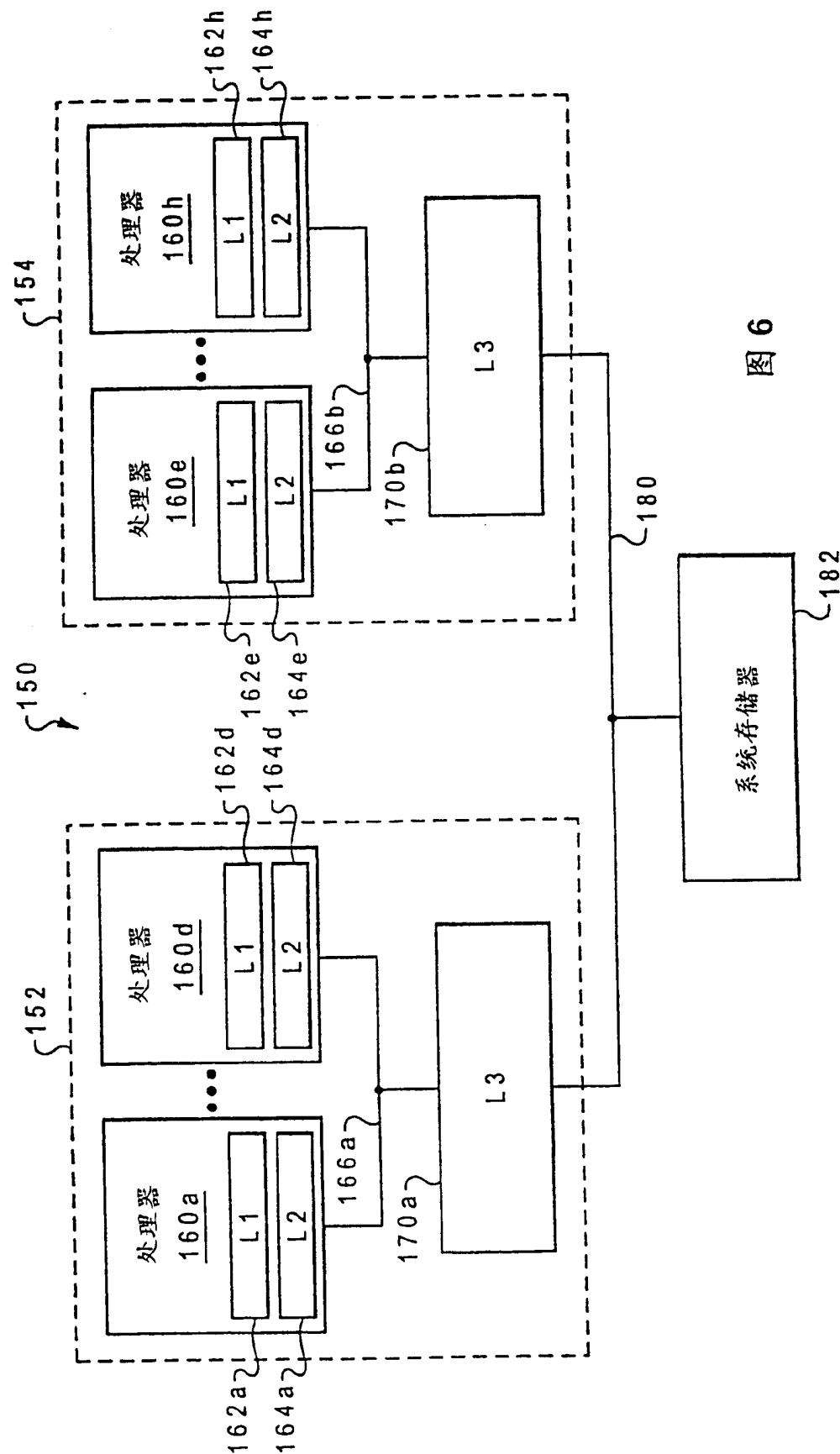


图 6