

[19] 中华人民共和国国家知识产权局

[51] Int. Cl.

G06F 11/34 (2006.01)

G06F 9/38 (2006.01)



[12] 发明专利申请公布说明书

[21] 申请号 200480001642.8

[43] 公开日 2007年6月20日

[11] 公开号 CN 1985242A

[22] 申请日 2004.4.14

[21] 申请号 200480001642.8

[30] 优先权

[32] 2003.4.23 [33] US [31] 10/422,025

[86] 国际申请 PCT/GB2004/001586 2004.4.14

[87] 国际公布 WO2004/095282 英 2004.11.4

[85] 进入国家阶段日期 2005.6.21

[71] 申请人 国际商业机器公司

地址 美国纽约

[72] 发明人 威廉·约瑟夫·阿姆斯特朗

迈克尔·斯蒂芬·佛洛伊德

拉里·斯科特·利特纳

罗纳德·尼克·卡拉

巴拉拉姆·辛哈洛伊

[74] 专利代理机构 中国国际贸易促进委员会专利商
标事务所

代理人 李德山

权利要求书4页 说明书8页 附图4页

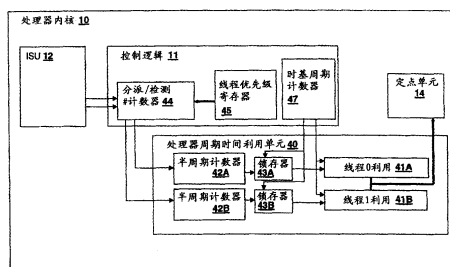
[54] 发明名称

确定同时多线程(SMT)处理器中每线程处理器资源利用的核算方法和逻辑

[57] 摘要

用于确定同时多线程(SMT)处理器中的每线程处理器资源利用的核算方法和逻辑提供了用于核算程序和程序内的线程的处理器资源利用的机制。通过检测针对处理器内活跃的多个线程的指令分派来确定相对资源利用,所述多个线程可能包含仍然占用处理器资源的空闲线程。如果针对所有线程分派指令或没有对线程分派指令,则同样地对所有线程核算处理器周期。可选地,如果没有线程处于分派状态,则可以使用先前状态或根据线程的优先级的比值来进行核算。如果只有一个线程正进行分派,则对该线程核算上整个处理器周期。如果多个线程正进行分派,但并不是所有线程正进行分派(在支持2个以上线程的处理器中),则将处理器周期平均记帐到进行分派的线程上。可以针对线程检测多个

分派,并且可以根据其分数利用更新针对每个线程确定的分数资源利用和计数器。



- 1.一种核算处理器中的处理器时间利用的方法，包括：
以周期性间隔确定所述处理器内所有执行线程的相对资源利用；
和
响应所述确定，根据所述相对资源利用更新多个处理器资源利用计数，其中每个所述计数均与多个线程中的特定一个相关。
- 2.如权利要求1所述的方法，其中所述周期性间隔是每个处理器时钟周期，并且所述确定确定每个线程是否处于特定周期状态。
- 3.如权利要求2所述的方法，其中所述确定还确定资源的数量，其中针对该资源的每个线程处于特定周期状态。
- 4.如权利要求3所述的方法，其中所述更新根据所述确定的由所述相关线程使用的资源的数量，分数更新所述多个所述处理器资源利用计数的每一个。
- 5.如权利要求2或3所述的方法，其中所述特定周期状态是指示针对线程的指令分派的状态。
- 6.如权利要求2到 5中任何一个所述的方法，其中响应所述确定确定所述多个线程的一或多个处于所述特定周期状态，所述更新同等更新与所述一或多个线程相关的所述多个所述处理器资源利用计数中的每一个。
- 7.如权利要求2到 6中任何一个所述的方法，其中响应所述确定确定所述多个线程中没有线程处于所述特定周期状态，所述更新同等更新所述多个所述处理器资源利用计数中的每一个。
- 8.如权利要求2到 7中任何一个所述的方法，其中响应所述确定进一步确定所述多个线程中没有线程处于所述特定周期状态，所述更新根据由所述周期性间隔的先前一个间隔确定的所述相关线程使用的资源的数量，分数更新所述多个所述处理器资源利用计数中的每一个。
- 9.如权利要求2到 8中任何一个所述的方法，其中响应所述确定确定所述多个线程中没有线程处于所述特定周期状态，所述更新根据

所述相关线程的优先级分数更新所述处理器资源利用计数中的每一个。

10.如权利要求3到9中任何一个所述的方法,其中所述处理器具有单线程操作模式和多线程操作模式,并且该方法还包括:

确定所述处理器是否处于所述单线程模式;和

响应确定所述处理器处于所述单线程模式,执行所述更新,使得每个处理器周期归于所述处理器内执行的单线程。

11.一种支持多个线程的并行执行的处理器,所述处理器包括:

由所述多个线程使用的多个资源;

连接到所述多个资源的指令控制单元,其中所述指令控制单元以周期性间隔控制所述处理器使用的资源的数量;

多个处理器资源利用周期计数器,每个周期计数器均与所述多个线程中的特定一个相关;和

连接到所述指令控制单元和所述处理器资源利用周期计数器的逻辑电路,其中所述逻辑电路根据所述相关线程的相对资源利用的确定,更新所述处理器资源利用周期计数器中的每一个。

12.如权利要求11所述的处理器,其中所述周期性间隔是每个处理器时钟周期,并且所述逻辑电路确定每个线程是否处于特定周期状态。

13.如权利要求12所述的处理器,其中所述控制逻辑还确定资源的数量,其中针对该资源的每个线程处于特定周期状态。

14.如权利要求13所述的处理器,其中所述控制逻辑根据所述确定的所述相关线程使用的资源的数量,分数更新所述多个所述处理器资源利用计数中的每一个。

15.如权利要求12、13或14所述的处理器,其中所述特定周期状态是指示针对线程的指令分派的状态。

16.如权利要求12到15中任何一个所述的处理器,其中所述控制逻辑响应确定所述多个线程的一或多个处于所述特定周期状态,同等更新与所述一或多个线程相关的所述多个所述处理器资源利用计数中的每一个。

17.如权利要求12到16中任何一个所述的处理器,其中所述控制逻辑响应确定所述多个线程中没有线程处于所述特定周期状态,同等更新所述多个所述处理器资源利用计数中的每一个。

18.如权利要求12到17中任何一个所述的处理器,其中所述控制逻辑响应确定所述多个线程中没有线程处于所述特定周期状态,根据确定的在所述周期性间隔的先前一个间隔中所述相关线程使用的资源的数量,分数更新所述多个所述处理器资源利用计数中的每一个。

19.如权利要求12到18中任何一个所述的处理器,其中所述控制逻辑响应确定所述多个线程中没有线程处于所述特定周期状态,根据所述相关线程的优先级,分数更新所述多个所述处理器资源利用计数中的每一个。

20.如权利要求12到19中任何一个所述的处理器,其中所述处理器具有单线程操作模式和多线程操作模式,所述控制逻辑确定所述处理器是否处于所述单线程模式,并且响应确定所述处理器处于所述单线程模式,在每个周期更新与所述处理器内执行的单线程相关的所述处理器资源利用周期计数器。

21.一种支持多个线程的并行执行的处理器,所述处理器包括:
由所述多个线程使用的多个资源;

连接到所述多个资源的指令控制单元,其中所述指令控制单元在每个时钟周期控制处理器内指令的分派;

多个处理器资源利用周期计数器,每个周期计数器均与所述多个线程中的特定一个相关;和

连接到所述指令控制单元和所述处理器资源利用周期计数器的逻辑电路,其中所述逻辑电路在所述指令控制单元没有正针对所述线程分派指令或正针对所有所述线程分派指令的情况下,同等更新所述处理器资源利用周期计数器中的每一个,并且在所述指令控制单元正只针对第一线程分派指令的情况下,更新与所述第一线程相关的所述处理器资源利用周期计数器。

22.如权利要求21所述的处理器,其中所述处理器具有单线程操作

模式和多线程操作模式，所述控制逻辑确定所述处理器是否处于所述单线程模式，并且响应确定所述处理器处于所述单线程模式，在每个周期更新与所述处理器内执行的单线程相关的所述处理器资源利用周期计数器。

确定同时多线程(SMT)处理器中 每线程处理器资源利用的核算方法和逻辑

技术领域

本发明涉及处理器和计算机系统，尤其涉及同时多线程 (SMT) 处理器。本发明也涉及处理器利用核算系统。

背景技术

今天的高速处理器包含同时执行指令，推测执行和加载指令，以及同时操作处理器内的各种资源的能力。具体地，已经发现最好是管理处理器内一或多个线程的执行，使得不止一个执行线程可以利用处理器，并且与通常由单线程利用资源的情况相比，更加有效地利用资源。

现有的处理器设计解决了通过从一个线程的执行到另一个线程的执行的硬件状态切换来管理多个线程的问题。这种处理器被称作硬件多线程 (HMT) 处理器，并且其能够提供一个或其它线程的执行之间的硬件切换。HMT 处理器通过允许硬件切换执行到另一个线程克服了等待停止线程的限制。通过在每个线程遇到例如高速缓存命中失误的停止条件时为另一个线程分配执行时间片，能够非同时地进行两个线程的执行。

由于多个线程可以同时利用处理器资源，同时多线程 (SMT) 处理器提供了对处理器资源的更加高效的利用。在 SMT 处理器中并行执行多个线程，使得多个处理器执行单元，例如浮点单元，定点指令单元，加载/存储单元和其它单元能够同时执行多个线程中的一个 (或更多，取决于执行单元的能力) 的任务。也可以逐个线程地分配存储器和寄存器资源，使得避免 HMT 的完全内部状态切换。

当例如某些处理器运行被用于研究和开发活动并且这些活动允许以不同方式为税务目的而不是其它用途将硬件转化为资本时，对外

理器时间利用的核算对于计算机服务销售的管理以及内部成本核算管理是需要的。服务器和处理器时间可以"按需"或按使用分别划分和告知给多个用户。另外,处理器时间可以被硬件所有者或出租人使用,以及转包给为服务付费的实体。因此,处理器执行时间的精确核算在计算机体系结构和软件模型中是必要的。并且,处理器时间利用,尤其是逐个线程的处理器时间利用被用于形成程序分布以进行优化。

在单线程处理系统中,核算通常是直接的。可以为完整任务运行提供处理器周期利用的计数,或甚至简单的"挂钟"时间测量,因为即使执行多个程序内的多个线程,然而它们不被同时执行,而是顺序执行。维护周期时间的记录,直到任务完成,并且为核算目的提供合计。测量的时间直接与处理器资源利用相关。

在 HMT 机器中,任务的类似之处在于,在处理器硬件内,每次只执行一个线程,并且多个计数器被用来跟踪每个线程的处理器时间利用,同时,只运行其相关线程为活跃的计数器。然而在 SMT 处理器中,两个或更多线程可能同时在单处理器内核内执行,并且通过简单的执行计数或时间测量不能容易地确定每个线程对资源的利用。

因此期望提供一种能够核算 SMT 处理器中的处理器时间利用的方法和设备。还期望提供一种核算这种处理器内执行的线程中的 SMT 处理器利用内的资源利用的方法。

发明内容

通过具有处理器利用计数器的处理器和用于核算处理器时间利用的方法来实现核算同时多线程 (SMT) 处理器内的处理器时间利用的目的。

处理器引入多个处理器利用周期计数器,处理器内执行的每个线程都有一个处理器利用周期计数器与之相关,以提供对应于每个线程的处理器资源利用的计数。相对资源利用是逐个线程的检测,并且被用来更新周期计数器。通过检测指示线程的活跃指令处理的特定周期状态的存在,可以进行资源利用检测。通过可以在每个时钟周期进行的周期性采样来检测周期状态,并且根据检测的周期状态递增(或可

选地, 递减)计数器。检测的状态被选择为指示每个线程的相对处理器资源利用。

检测的周期状态可以是针对线程的指令分派。没有线程处于特定周期状态的周期可以由每个线程平均负担, 或者可以根据所有线程的最近分派状态来负担该周期。可选地, 可以根据提供相对线程资源利用的指示的排序线程优先级来负担非指示周期。其中一个线程处于特定周期状态并且另一个线程不处于特定周期状态的周期可以完全由处于特定周期状态的线程负担, 从而规定活跃线程优先于空闲线程负担费用。并且, 可以根据指示每个线程在指定周期中的使用的资源数量的分数值来负担周期。

处理器可以能够处于 SMT和单线程 (ST)操作模式, 并且核算可以有选择地响应选定模式: 从而由执行于 ST模式的单线程负担所有周期。

通过以下如附图所示的更加具体的关于本发明优选实施例的说明, 可以理解本发明的上述和其它目的, 特征和优点。

附图说明

在所附权利要求书中提出了被认为是本发明的特征的新颖特性。然而参照下列结合附图对图解实施例进行的详细描述可以更好地理解本发明自身, 最优使用模式, 其它目的和优点, 其中类似附图标记指示类似的部分, 并且:

图1是根据本发明的一个实施例的系统的模块图。

图2是根据本发明的一个实施例的处理器内核的模块图。

图3是图2的处理器内核10内的功能单元的详细的模块图。

图4是描述根据本发明实施例的方法的流程图。

具体实施方式

参照附图并且具体参照图1, 其中描述了一个基于本发明最优实施例的系统的模块图。系统包含处理器组5, 处理器组5可以通过桥接器37连接到其它处理器组以形成超标量处理器。处理器组5连接到 L3高速缓存单元36, 系统局部存储器38和各种外设34, 以及 2个服务处

处理器 34A和34B。服务处理器提供故障监视，启动支持和测试能力给处理器组5，并且可以具有其自身的到其它处理器组以及连接所有处理器 30A -D的互连路径。

处理器组5内是多个处理器30A-D，其通常制造在单个单元中，并且包含连接到 L2高速缓存32和存储器控制器4的多个处理器内核 10A和10B。内核 10A和10B为通用处理功能提供指令执行和数据值的操作。桥接器37以及系统内的其它桥接器提供宽总线上与其它处理器组的通信，并且总线35提供处理器30A-D，桥接器37，外设34， L3高速缓存36和系统局部存储器38的连接。其它全局系统存储器可以外部连接到桥接器37以便由所有处理器组对称访问。

处理器内核 10A和10B是能够并行执行多个线程的同时多线程 (SMT)处理器。处理器内核 10A和10B还支持单线程操作模式，以便在程序执行条件规定单线程操作时，例如当必须在已知时间前完成高优先级程序执行时，或当多线程处理器中的一个线程已知为空闲时，高效地执行单线程。与单线程的全时执行相比，多线程引入某些低效率因素，而是总体上存在系统效率优势，因为线程经常因等待其它任务完成而空闲。因此，单线程和多线程模式之间的转变提供了适应一或多个上述条件方面的优势，并且本发明的实施例以和提供响应这种转变的处理器时间核算的处理器相符合的方式提供处理器时间核算。

现在参照图2，描述具有与处理器内核 10A和10B相同的特征的处理器内核10的细节。总线接口单元23将处理器内核10连接到其它SMT处理器和外设，并且连接用于存储数据值的 L1 Dcache 22，用于存储程序指令的 L1 Icache 20，和高速缓存接口单元21到外部存储器，处理器和其它设备。L1 Icache 20结合指令获取单元 IFU 16提供指令流的加载，指令获取单元 IFU 16预取指令，并且可以包含推测加载和分支预测能力。指令定序器单元 (ISU) 12控制提交到例如执行通用操作的定点单元 (FXU) 14和执行浮点操作的浮点单元 (FPU) 15的各种内部单元的指令的顺序。全局完成表 (GCT) 13通过标签跟踪 ISU 12提交的指令，直到该指令寻的的具体执行单元指示指令已经完成执行。

定点单元14和浮点单元15被连接到各种资源，例如通用寄存器 (GPR) 18A，浮点寄存器 (FPR) 18B，条件寄存器 (CR) 18C，重命名缓冲区 18D，计数寄存器/链接寄存器 (CTR/LR) 18E和异常寄存器 (XER) 18F。GPR 18A和 FPR 18B对加载存储单元 (LSU) 19从 L1 Dcache 22加载和存储的数据值提供数据值存储。CR 18C存储条件分支信息，重命名缓冲区18D (可以包括与各个内部执行单元相关的若干重命名单元)为执行单元提供操作数和结果存储。XER 18F存储分支和定点异常信息，CTR/LR 18E存储程序分支执行的分支链接信息和计数信息。GPR 18A，FPR 18B，CR 18C，重命名缓冲区18D，CTR/LR 18E和XER 18F是包含某些固定(经结构设计)的寄存器的资源，这些寄存器存储程序执行期间的信息，并且必须被提供为针对每个执行线程的固定设置，上述资源内的其它非结构设计寄存器是空闲的，以用于重命名用途。控制逻辑11被连接到处理器内核10内的各个执行单元和资源，并且被用来提供对执行单元和资源的全面控制。SCOM/XSCOM 接口单元25提供到外部服务处理器 34A-B的连接。

现在参照图3描述处理器内核10内的功能模块的细节。ISU 12指示控制逻辑11何时针对用处理器内核10执行的具体线程分派指令。控制逻辑11根据 ISU 12在分派检测电路44所检测的指定周期内是否为与指定计数器相关的线程 (针对半周期计数器42A的线程0和针对半周期计数器42B的线程1)分派了指令，有选择地更新半周期计数器42A和42B。半周期计数器 42A和42B的输出被连接到时基周期计数器47的锁存器 43A和43B锁存。在时基周期计数器47的每个第8处理器周期上，提供锁存器 43A和43B的输出以递增线程利用计数器 41A和/或41B，从而每8个处理器周期(时基周期计数器47活跃)提供线程利用计数器 41A和/或41B的更新 (将排除其中时基不活跃的特殊操作由活跃线程来负担的情况)。在递增之后，时基周期计数器47复位锁存器43A和43B。上述动作在线程利用计数器 41A和/或41B的输出处提供作为处理器的时基计数器的分数的计数，并且能够被直接用于按照"利用时间"值计算机相对线程资源利用。

根据本发明的可选实施例,控制逻辑11和 ISU 12 (或处理器内核10内的其它执行单元)能够利用其它具体周期状态在每个周期中触发半周期计数器42A和/或 42B的更新,而是在本发明的示例性说明中,选择指令分派,因为指令分派提供测量使用处理器内核10的每个处理活动的开始,而不是这种活动的进度或完成(取决于代码效率)的机制。例如,低效编码的程序可能产生许多不正确的分支预测和相应的预取和刷新,从而用掉关键处理器资源并且产生较少的完成。因此,利用指令完成而不是指令分派会产生对执行低效代码的线程的较低处理器时间利用测量。因此,将指令分派用作更新半周期计数器 41A和 41B的特定触发器事件提供了没有"报答"低效代码的处理器资源利用估计。

虽然现有核算模型已经能够基本上为特定线程分配所有处理器内核10时间,然而由于在指定处理器中缺乏线程的同时执行,在 SMT 处理器中,逐个线程的处理器时间利用的"记帐"必须被分布在可以同时执行于处理器内核10内的两个或更多线程上。当 ISU 12不针对使用处理器内核10的任何线程分派指令时,仍然有必要核算所有线程对处理器内核10的利用。线程可能没有开始分派指令的形式的新工作,但是仍然"占用"处理器内核10,或可能在不同于指令分派的执行阶段活跃。例如,一个线程可能正导致长的预取序列,而另一个线程可能正执行冗长的浮点操作。根据上述例子,处理器资源和执行单元正被两个线程主动利用,同时只有少量指令完成会与处理区间相关。对完成是优选的另一个特定指令状态是指令获取,它会提供类似于指令分派的指示,但是提供更加偏向 IFU 16和 L1 Icache 20利用的利用计数,而不是包含 FXU 14, FPU 15及其相关资源的执行单元利用。

因此,这里描述的示例性实施例提供了即使在指定周期没有出现分派 (或在可选实施例中,由控制逻辑11检测的其它特定周期状态或资源利用)时,仍然能够根据执行单元和资源利用估计核算所有处理器内核10周期的机制。在没有出现针对任何线程的指令分派的周期中,以针对处理器内核10内执行的每个线程(包含尚未从处理器内核10中

卸载并且仍然占用处理器资源的空闲线程)均分周期值的方式核算周期。但是,在控制逻辑11的可选实施例或可选行为中,控制逻辑11可以记忆最近分派状态,并且可以用此更新半周期计数器42A和/或42B。另一个可选方式是控制逻辑11利用处理器内核10中设置的相对线程优先级分数分割用来更新半周期计数器42A和/或42B的时钟周期值,以便适当分配其中没有出现针对任何线程的分派的周期。

类似地,当在指定周期针对不止一个线程分派指令时,以针对被分派指令的每个线程均分周期值的方式核算周期。或者再次地,控制逻辑11可以利用处理器内核10中设置的相对线程优先级分数分割用来更新半周期计数器42A和/或42B的时钟周期值,可以便适当分配其中没有出现针对任何线程的分派的周期。

因此,在均分方法(不是前面作为可选方式列出的优先级或先前周期模式)中,在支持一个或2个线程的执行的处理器内核中,对于其中没有出现分派的周期,或其中出现针对每个线程的指令分派的周期,对每个线程核算 $1/2$ 值。在只出现针对一个线程的指令分派的周期中,将整个周期记帐到出现针对其的指令分派的线程上。控制逻辑11检测以下4个潜在状态之一:没有线程分派,线程0分派,线程1分派,或两个线程均分派,并且根据上述周期记帐分布更新线程0半周期计数器42A和线程1半周期计数器42B。

由于上述最低分布值被用于 $1/2$ 周期增量的分配,因此对于半周期用增量(或减量)1,或对于完全周期分配用增量(或减量)2更新半周期计数器42A和42B。实际上,用基于芯片面积,功率和定时原因的可能指令分派率的 $1/8$ (或由于因相同原因被选择为处理器周期时间的 $1/8$ 的时基周期计数器47的速率)更新每个线程利用计数器41A和41B,所以不形成"经结构设计"的线程利用计数器的部分的4位半周期计数器42A和42B提供每次溢出的有效8周期利用计数(由于值2会被分配给半周期计数器42A和42B的总增量)。半周期计数器42A和42B溢出到当核算例程(程序)检索处理器周期时间利用信息时则定点单元14读取的线程利用计数器41A和41B的部分。半周期计数器42A和42B被

实现为将完全周期分配的两个子计数(sub-count)相加或将半周期分配的一个子计数相加的加法器。

在本发明的可选实施例中，处理器周期时间利用单元40使用加法递增(或使用减法递减)线程利用计数器41A和41B。由检测正对指定线程进行多个分派的分派检测电路44来确定相加的值，并且用一个值(已经相加其上)递增相关的子计数器，该值与周期针对该线程分派的指令的数量与针对所有线程分派的指令的总数的比值成比例。例如，当处理器内核10正执行2个线程并且在指定周期中针对第一线程分派3个指令并且针对第二线程分派一个指令时，值 $3/4$ 被分配给第一线程的线程利用计数器，并且 $1/4$ 被分配给第二线程的线程利用计数器。如果在指定周期没有进行分派，则仍然在线程之间平均分配该周期。

现在参照图4，在流程图中描述了根据本发明实施例的方法。所描述的方法用于具有SMT和ST模式、并且能够在SMT模式中同时执行多个线程的SMT处理器内的核算。如果处理器处于ST模式(判决50)，则每个周期用1周期值更新针对执行线程的线程周期计数器(步骤51)。如果处理器处于SMT模式(判决50)，并且如果只有一个线程正进行分派(判决52)，则用1周期值更新相关线程周期计数器(步骤53)。如果所有线程正进行分派(判决54)，则根据由线程数量确定的周期分数，或根据针对每个线程提交的分派的数量(或使用的资源)，同等更新所有线程的线程周期计数器(步骤55)。如果没有线程进行分派，则能够同等地，或与最后分派采样成比例的，或与其优先级成比例地更新线程周期计数器(步骤57)。最终，如果没有线程进行分派，则能够同样地同等地，或与最后分派采样成比例的，或与其优先级成比例地更新线程周期计数器(步骤58)。

虽然前面已经具体参照本发明的优选实施例示出和描述了本发明，然而本领域的技术人员会理解，在不偏离本发明的实质和范围的前提下，可以在形式和细节方面进行上述和其它变化。

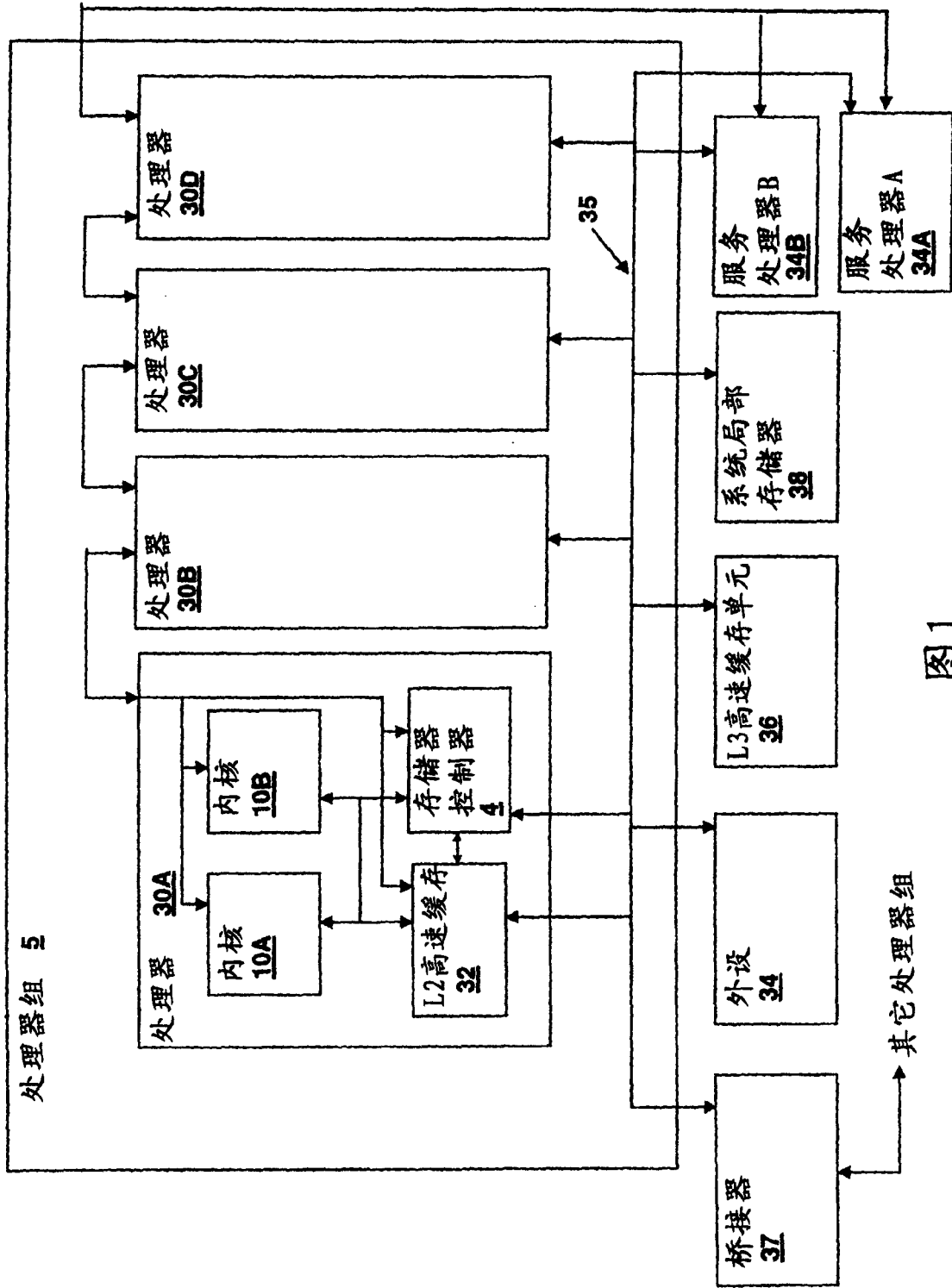


图1

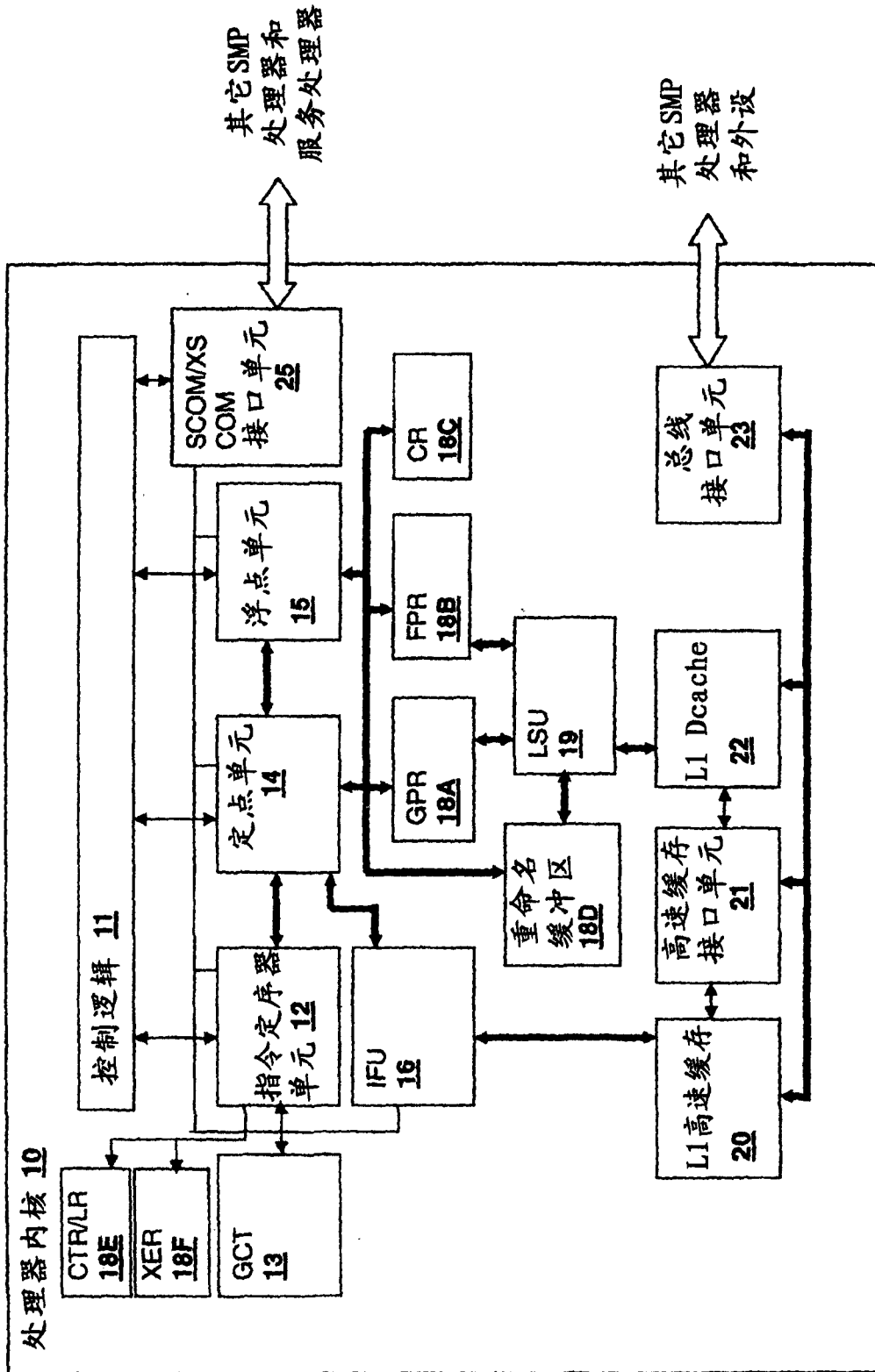


图2

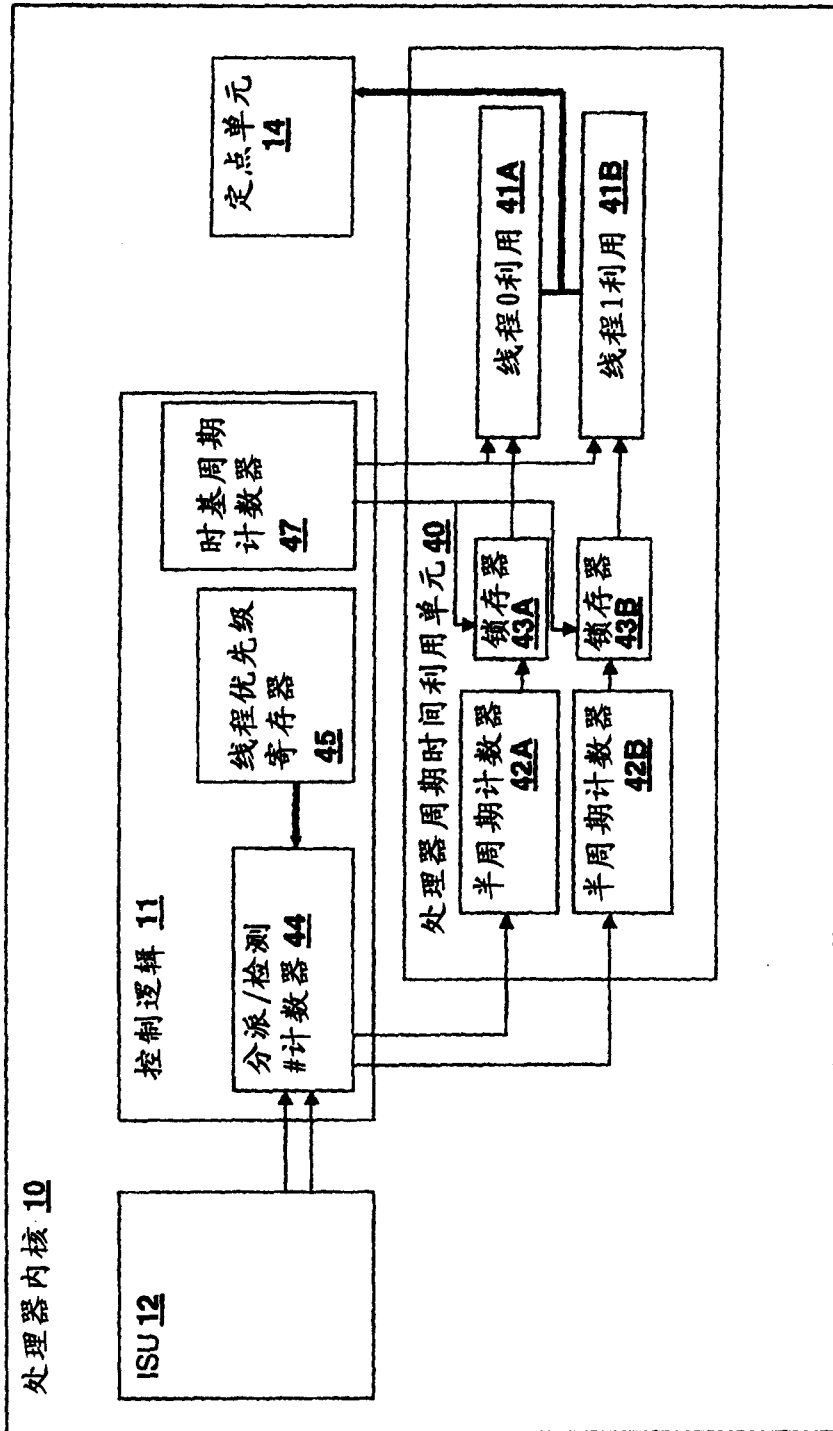


图3

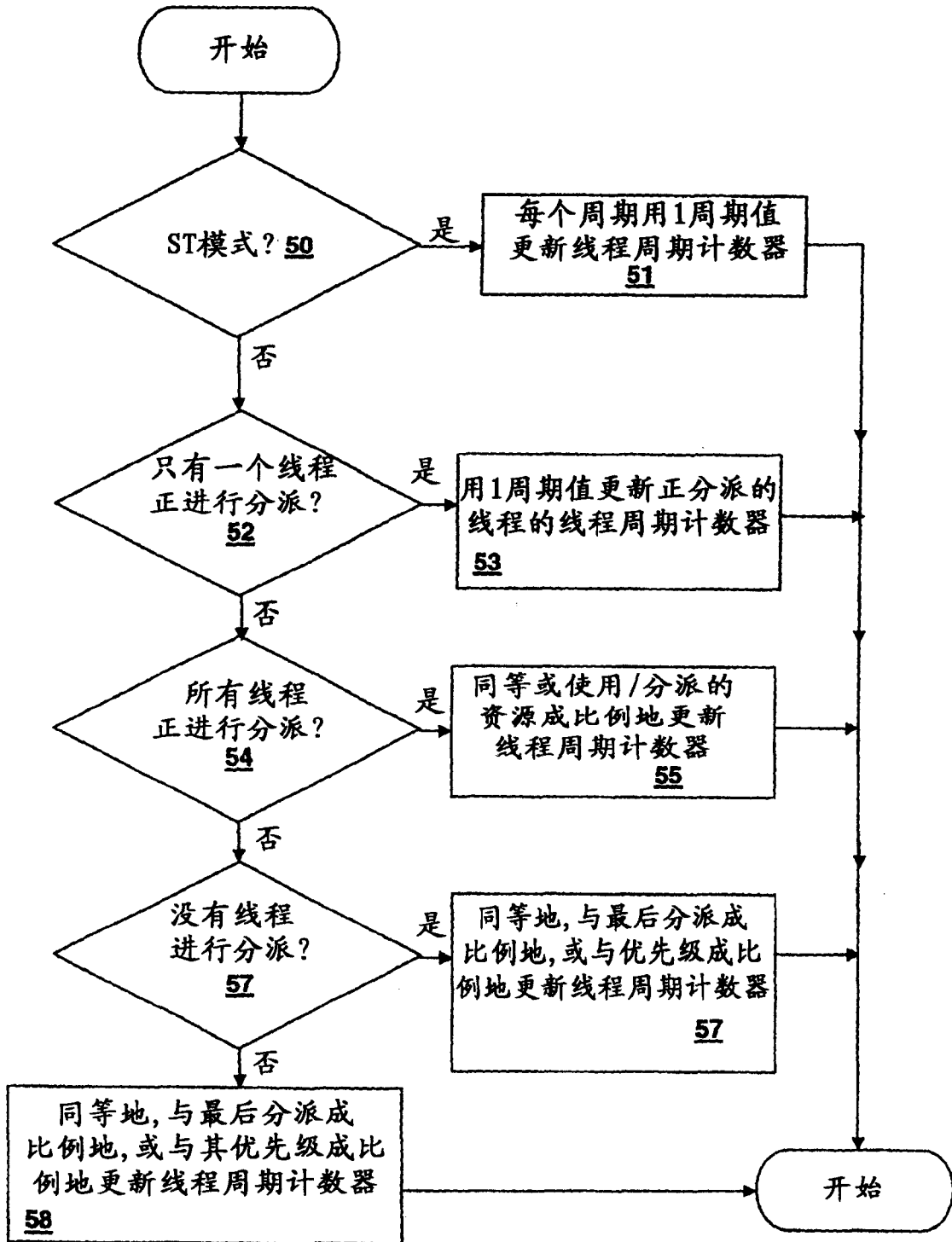


图 4