



(19)中華民國智慧財產局

(12)發明說明書公開本

(11)公開編號：TW 201730761 A

(43)公開日：中華民國 106 (2017) 年 09 月 01 日

(21)申請案號：105137288

(22)申請日：中華民國 105 (2016) 年 11 月 15 日

(51)Int. Cl. : **G06F9/46 (2006.01)**

(30)優先權：2015/12/21 美國

14/976,360

(71)申請人：英特爾股份有限公司 (美國) INTEL CORPORATION (US)
美國

(72)發明人：賽瑞 菲利浦 THIERRY, PHILIPPE (FR)；歐德亞麥德維爾 艾蒙斯特阿法 OULD-AHMED-VALL, ELMOUSTAPHA (MR)

(74)代理人：林志剛

申請實體審查：無 申請專利範圍項數：20 項 圖式數：12 共 76 頁

(54)名稱

高度平行應用的排程

SCHEDULING HIGHLY PARALLEL APPLICATIONS

(57)摘要

本文揭示用於藉由高性能計算系統來執行之高度平行應用排程的系統與方法。一種實例處理系統可包含：一控制暫存器；以及一處理核心，其通訊地耦合至該控制暫存器。該處理核心可經組態以接收具體指定應用的預期運行時間以及叢集節點之節點請求數目的節點分配請求；鑑於該節點分配請求以及複數個節點上的電流負載，決定要分配給該應用的節點之實際數目，其中要分配給該應用的該節點實際數目將叢集負載標準最佳化；以及使用該控制暫存器，向該應用通知該節點實際數目。

Systems and methods for scheduling highly-parallel applications executed by high-performance computing systems. An example processing system may comprise: a control register and a processing core, communicatively coupled to the control register. The processing core may be configured to receive a node allocation request specifying an expected running time of an application and a requested number of nodes of a cluster of nodes; determine, in view of the node allocation request and a current load on the plurality of nodes, an actual number of nodes to be allocated to the application, wherein the actual number of nodes to be allocated to the application optimizes a cluster load criterion; and notify, using the control register, the application of the actual number of nodes.

指定代表圖：

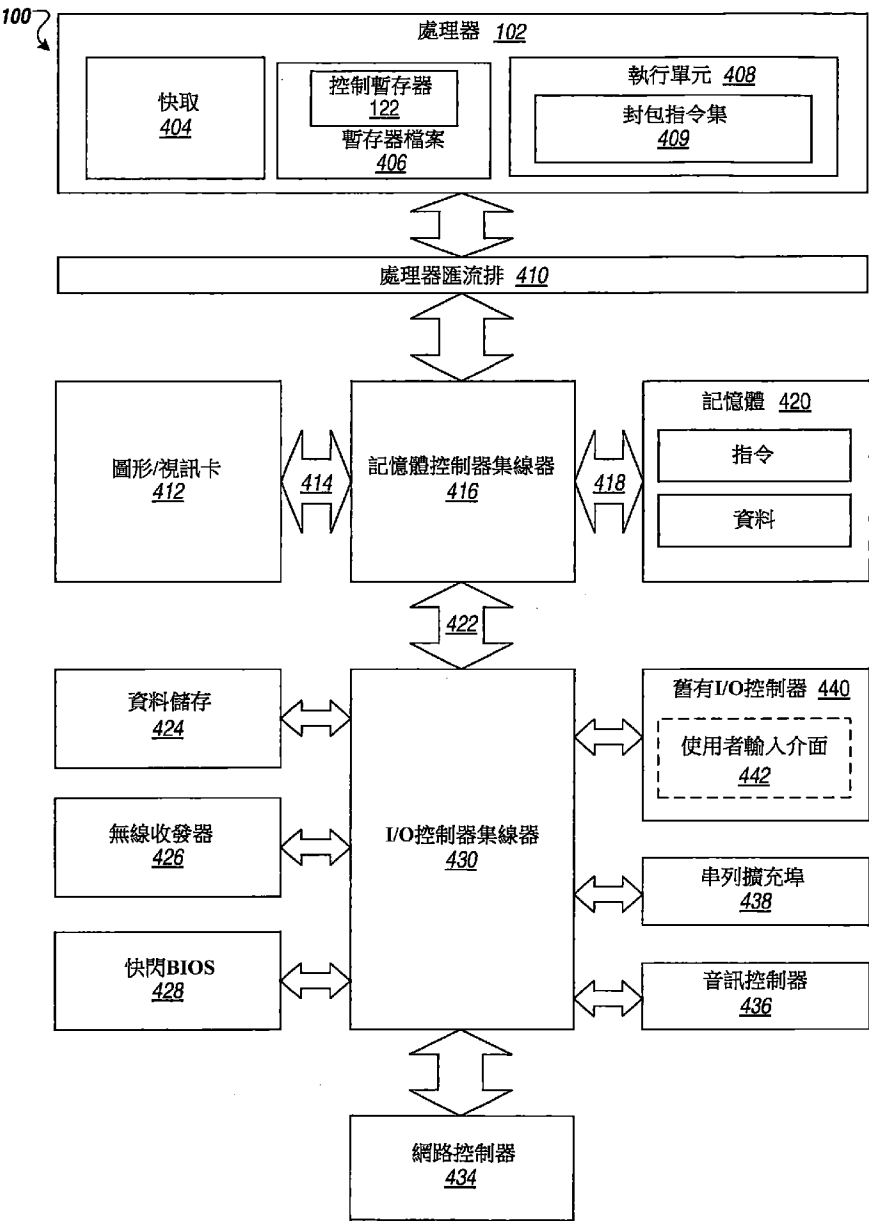


圖 3

符號簡單說明：

- 100 . . . 處理系統
- 102 . . . 處理器
- 122 . . . 控制暫存器
- 404 . . . 級別 1(L1) 內部快取
- 406 . . . 暫存器檔案
- 408 . . . 執行單元
- 409 . . . 封包指令集
- 410 . . . 處理器匯流排
- 412 . . . 圖形視訊卡
- 416 . . . 記憶體控制器集線器
- 420 . . . 記憶體
- 422 . . . 系統輸入/輸出
- 424 . . . 資料儲存器
- 426 . . . 無線收發器
- 428 . . . 快閃基本輸出入系統
- 430 . . . 輸入/輸出控制器集線器
- 434 . . . 網路控制器
- 436 . . . 音訊控制器
- 438 . . . 串列擴充埠
- 442 . . . 使用者輸入面

發明摘要

※申請案號：105137288

※申請日：105 年 11 月 15 日

※IPC 分類：G06F 9/46 (2006.01)

【發明名稱】(中文/英文)

高度平行應用的排程

Scheduling highly parallel applications

【中文】

本文揭示用於藉由高性能計算系統來執行之高度平行應用排程的系統與方法。一種實例處理系統可包含：一控制暫存器；以及一處理核心，其通訊地耦合至該控制暫存器。該處理核心可經組態以接收具體指定應用的預期運行時間以及叢集節點之節點請求數目的節點分配請求；鑑於該節點分配請求以及複數個節點上的電流負載，決定要分配給該應用的節點之實際數目，其中要分配給該應用的該節點實際數目將叢集負載標準最佳化；以及使用該控制暫存器，向該應用通知該節點實際數目。

【 英文 】

Systems and methods for scheduling highly-parallel applications executed by high-performance computing systems. An example processing system may comprise: a control register and a processing core, communicatively coupled to the control register. The processing core may be configured to receive a node allocation request specifying an expected running time of an application and a requested number of nodes of a cluster of nodes; determine, in view of the node allocation request and a current load on the plurality of nodes, an actual number of nodes to be allocated to the application, wherein the actual number of nodes to be allocated to the application optimizes a cluster load criterion; and notify, using the control register, the application of the actual number of nodes.

【代表圖】

【本案指定代表圖】：第(3)圖。

【本代表圖之符號簡單說明】：

- 100：處理系統
- 102：處理器
- 122：控制暫存器
- 404：級別 1(L1)內部快取
- 406：暫存器檔案
- 408：執行單元
- 409：封包指令集
- 410：處理器匯流排
- 412：圖形視訊卡
- 416：記憶體控制器集線器
- 420：記憶體
- 422：系統輸入/輸出
- 424：資料儲存器
- 426：無線收發器
- 428：快閃基本輸出入系統
- 430：輸入/輸出控制器集線器
- 434：網路控制器
- 436：音訊控制器
- 438：串列擴充埠
- 442：使用者輸入面

【本案若有化學式時，請揭示最能顯示發明特徵的化學式】：
無

發明專利說明書

(本說明書格式、順序，請勿任意更動)

【發明名稱】(中文/英文)

高度平行應用的排程

Scheduling highly parallel applications

【技術領域】

本揭露通常關於高性能計算機（HPC）系統，且更具體地，關於用於高度平行應用之排程的指令集架構支持。

【先前技術】

高性能計算（HPC）系統可包括組成一或多個計算叢集的數個計算節點。計算節點可由控制節點管理。

【圖式簡單說明】

本揭露以舉例的方式而不是以限制的方式來繪示，且當結合圖式來考慮時，可能可參考以下詳細說明而更令人完整理解。

圖 1 示意地繪示根據本揭露一或多項態樣來操作之實例高性能計算系統的方塊圖。

圖 2A 至圖 2B 示意地繪示根據本揭露一或多項態樣的多個叢集組態。

圖 3 係根據本揭露一或多項態樣來操作之實例處理系統的方塊圖。

圖 4 描繪根據本揭露一或多項態樣之用於決定在節點的叢集中之應用組態的實例方法 400 的流程圖。

圖 5 繪示根據本揭露一或多項態樣的封包資料類型。

圖 6A 繪示根據本揭露一或多項態樣之處理器微架構的元件。

圖 6B 繪示根據本揭露一或多項態樣之處理器微架構的元件。

圖 7 係根據本揭露一或多項態樣來操作之處理器的方塊圖。

圖 8 係根據本揭露一或多項態樣來操作之處理器的方塊圖。

圖 9 係根據本揭露一或多項態樣來操作之系統單晶片的方塊圖。

圖 10 係根據本揭露一或多項態樣來操作之計算系統的方塊圖。

圖 11 繪示用於根據本揭露一或多項態樣來操作之計算系統的另一實例方塊圖。

圖 12 繪示實例形式為根據本揭露一或多項態樣來操作之計算系統之機器的圖示。

【發明內容及實施方式】

在本文中說明用於高度平行應用之排程的系統與方法。

高性能計算（HPC）系統可包括組成一或多個計算叢

集的多個計算節點。存在於專用節點上或存在於其中一個計算節點上的批次排程器可用來將在計算節點上執行的計算工作（亦稱為應用）排程。可請求一或多個計算節點，以用於執行已知的應用，且可指定期待的運行時間。排程器可利用這些與其他的執行參數，以將應用指派到用於執行的計算節點。

排程器可應用將計算節點之使用量程度最大化的排程策略，以便能夠增加整個系統成本效率。不過，在一般的實施方案中，計算節點的平均使用量稀有地超過其等容量的 80%，因為可用於及時在已知時刻排程的節點數目可能不一定等於欲排程之一或多個應用之節點請求數目的總和。換言之，在可用節點數目與分配至一或多個應用之節點數目之間的差，可小於將等待執行之另一應用排程所必要的節點數目。

計算節點的未充分利用不利地影響整個系統成本效率，因為甚至非主動計算節點消耗用於其操作與冷卻的電能。以 80% 容量利用消耗 20 與 40 兆瓦（MW）電能之間的一叢集計算節點，意味著每小時將損失 4MW 至 8MW 的電能。

本揭露透過提供用於批次排程方法的指令集架構（ISA）支持來解決上述與其他缺陷，該批次排程方法導致應用程式在運行時間上動態地增加或減少應用足跡（亦即，計算節點的數目）。根據本揭露的一或多項態樣，執行批次排程器的處理器可實施控制暫存器，以用於儲存反

映目前應用組態之記憶體資料結構的位址（例如，指派至應用之節點的識別器）。處理器可進一步實施數個指令，以操縱控制暫存器。此等指令可包括用於讀取組態的負載組態（LDCONFIG）、用於儲存組態的儲存組態（STCONFIG）、以及用於請求應用組態資料結構之鎖定的請求組態所有權（REQCONFIG）。

根據本揭露一或多項態樣來操作的批次排程器可經程式化，以決定將特定叢集負載標準最佳化的叢集組態。在例示性實例中，批次排程器可接收來自正等待執行或正由高性能計算系統執行之應用的節點分配請求。節點分配請求可指定期待的應用以及用於執行應用之節點的請求數目。

回應於接收節點分配請求，批次排程器可決定將選出之叢集負載標準最佳化的叢集組態。批次排程器隨後可利用處理器控制暫存器，以通知該應用，指派到該應用之節點的實際數目，如以下在本文中更詳細的說明。

以上參考方法與系統的多項態樣在本文中以實例方式而非以限制方式詳細說明如下。

現在參考圖 1，實例高性能計算系統 1000 可包括主要節點 110 以及由網路 135 互連的複數個計算節點 120A 至 120N。主要節點 110 可執行批次排程器 140，其可例如經由訊息傳遞介面（MPI）訊息 150 通訊至節點 120A 至 120N 的各者。

主要節點 110 可實施輸入工作佇列 160，以用於將等

待排程的應用 170A 至 170K 佇列。各應用可藉由各別應用識別器、請求用來執行應用之計算節點的數目、以及應用的期待運行時間來說明。

可將批次排程器 140 程式化，以決定將特定叢集負載標準最佳化的叢集組態。如圖 2A 示意地繪示，批次排程器 140 可產生應用 170A 至 170K 的複數個組合，且選擇在任何已知時間、將仍然未指派至任何應用之節點數目最小化的組合。或者，批次排程器 140 可選擇在已知時段內、將仍然未使用之總處理容量最小化的組合。圖 2 示意地繪示在特定時段內、將欲由複數個計算節點 120A 至 120N 執行之複數個應用 170A 至 170K 排程的實例。如圖 2A 所示，習知排程演算法可產生具有明確程度之未使用處理容量的排程（標為方塊 180）。

因為一些應用（例如，客戶端-伺服器應用）可能具有可輕易地適用於節點之動態變化數目的架構，所以批次排程器 140 則可藉由應用排程方法來改善叢集效率，該等排程方法導致該等應用在運行時間動態地增加或減少應用足跡（亦即，計算節點的數目）。相較於習知方法，此等排程方法可產生更有效率的排程，其進一步將未使用之處理容量的程度節點，如圖 2B 示意地繪示。

圖 3 示意地繪示根據本揭露一或多項態樣來操作之實例處理系統的方塊圖。如圖 3 示意地繪示，處理系統 100 可包括一或多個處理器 102，其提供用於批次排程方法的 ISA 支持，該批次排程方法指示該等應用根據本揭露的一

或多項態樣，動態地增加或減少在運行時間的應用足跡。

在圖 3 的說明性實例中，處理器 102 包括經組態以施行特定指令集之指令的一或多個執行單元 408。處理器 102 耦合至處理器匯流排 410，該處理器匯流排傳送資料訊號於處理器 102 與系統 100 的其他組件之間。系統 100 的其他元件（例如，圖形視訊卡 412、記憶體控制器集線器 416、記憶體 412、I/O 控制器集線器 430、無線收發器 426、快閃 BIOS 428、網路控制器 434、音訊控制器 436、串列擴充埠 438、舊有 I/O 控制器 440 等等）可施行熟悉該技術者眾所皆知的其等習知功能。

各處理器 102 可包含複數個暫存器，其包括通用暫存器與專用暫存器。在特定實施方案中，處理系統 100 也可包括在圖 3 中沒有顯示的多個其他組件。處理系統 100 之多個組件的更詳細說明在本文中呈現如下。

處理器 102 可包含控制暫存器 122，該控制暫存器可用來儲存反映目前應用組態之記憶體資料結構的位址。在說明性實例中，應用組態資料結構可將欲指派用於執行應用的應用識別器與節點數目儲存。在說明性實例中，應用組態資料結構最初可由應用產生，其可記錄節點的請求數目。批次排程器最終可修改指派至該應用的節點數目，以試圖將叢集負載標準最佳化。

或者，應用組態資料結構可由批次排程器產生，其可記錄指派至該應用之節點的數目。應用組態資料結構的產生器可使用 STCONFIG 指令，將資料結構的位址儲存在控

制暫存器 122 中。該應用可使用 LDCONFIG 指令來讀取控制暫存器 122 的內容，且可擷取資料結構的位址，以儲存用於應用之節點的現有組態。

在特定實施方案中，處理器 102 可進一步實施用於在修改資料結構之前、請求應用組態資料結構之鎖定的請求組態所有權（REQCONFIG）。在說明性實例中，在自獲得鎖定起之預定逾時期滿時，可鬆開鎖定。或者，藉由執行具有指示鎖定鬆開之參數的 REQCONFIG 命令，可明顯地鬆開鎖定。

在特定實施方案中，批次排程器可決定，在預定時段內，應用無法鬆開應用組態資料結構之先前得到的鎖定。批次排程器可能認為該應用是無反應的，且可啟動適當的恢復程序（例如，藉由結束無反應的應用）。

如本文中的上文所提及，可將根據本揭露之一或多項態樣來操作的批次排程器程式化，以決定將特定叢集負載標準最佳化的叢集組態。在說明性實例中，批次排程器可從等待執行或正由高性能計算系統執行的應用，接收節點分配請求。節點分配請求可具體指定期待的應用，以及用於執行應用之節點的請求數目。在特定實施方案中，節點分配請求可進一步具體指定所需要之節點的最小數目，等待偏好參數則指示額外的等待時間在分配較少數目的節點上是否優於所請求的應用優先級、及/或各種其他的應用參數。

回應於接收節點分配請求，批次排程器可決定欲分配

至應用之節點的實際數目。如本文中的上文所提及，可選擇節點的實際數目，以將特定叢集的負載標準最佳化。在說明性實例中，可如下地決定叢集負載標準：

$$L = \sum_{i=1, \dots, n} (n_i * t_i)$$

其中，L 標示欲由批次排程器最小化的叢集未使用負載，且 n_i 係為在時段 t_i 仍維持未使用的節點數目。

在說明性實例中，批次排程器可產生欲由叢集執行之複數個應用的複數個應用組合，決定用於各產生組合之叢集負載標準的值，並且選擇與叢集負載標準之最佳（最大或最小）值有關聯的組合。

在各種說明性實例中，批次排程器可決定滿足額外限制的叢集組態，該等額外限制係從節點分配請求的各種參數獲得，諸如欲分配到應用之節點的最小數目，及／或指示額外的等待時間在分配較少數目的節點上是否優於請求的等待偏好參數。

回應於決定將叢集負載標準最佳化的叢集組態，批次排程器可通知該應用已經指派到該應用之節點的實際數目，如在本文中之上文的更詳細說明。

圖 4 描繪根據本揭露一或多項態樣之用於決定在節點叢集中之應用組態之實例方法 400 的流程圖。方法 400 可藉由處理邏輯來施行，該邏輯可包含硬體（例如，電路、專用邏輯、可程式化邏輯、微碼等等）、軟體（諸如由功能性單元施行的操作）、韌體或其組合。在特定實施方案

中，方法 400 係由欲由圖 4 之實例處理系統 100 所執行的批次排程器所施行。

參考圖 4，方法 400 開始於方塊 410，在此，實施該方法的處理器可接收節點分配請求。節點分配請求可具體指定所期待的應用以及用於執行應用之節點的請求數目。在特定實施方案中，節點分配請求可進一步具體指定多個額外的應用執行參數，如在本文中之上文的更詳細說明。

在方塊 420，根據節點分配請求以及在複數個節點上的現有負載，處理器可決定用於請求應用的應用組態，包括欲分配至應用的節點數目。可選定應用組態，以將特定的叢集負載標準最佳化，如在本文中之上文的更詳細說明。

在方塊 430，處理器可將應用組態（包括節點數目）儲存在應用組態資料結構中，如在本文中之上文的更詳細說明。

在方塊 440，處理器可將應用組態資料結構的位址儲存在控制暫存器中，如在本文中之上文的更詳細說明。

在方塊 450，處理器可導致應用根據所決定的應用組態來執行。

本揭露的實施方案不限於電腦系統。本揭露的替代實施方案可使用於其他裝置中，諸如手持裝置與嵌入式應用。手持裝置的一些實例包括蜂巢式電話、網際網路協定裝置、數位相機、個人數位助理（PDA）、以及手持個人電腦（PC）。嵌入式應用可包括微控制器、數位訊號處理

器（DSP）、系統單晶片、網路電腦（NetPC）、機上盒、網路集線器、廣域網路（WAN）切換器、或可根據至少一項實施方案來施行一或多個指令的任何其他系統。

再度參考圖 3，處理器 102 包括實施用以執行至少一項指令之演算法的一或多個執行單元 408。一項實施方案可說明於單一處理器桌面或伺服器系統的情境中，但是替代的實施方案卻可包括在多處理器系統中。系統 100 是「集線器」系統架構的實例。處理系統 100 包括用以處理資料訊號的處理器 102。舉一個說明性實例，處理器 102 包括複雜指令集計算（CISC）微處理器、精簡指令集計算（RISC）微處理器、超長指令字組（VLIW）微處理器、實施指令集組合的處理器、或任何其他處理器裝置，諸如例如數位訊號處理器。處理器 102 耦合至處理器匯流排 410，該處理器匯流排傳送資料訊號於處理器 102 與系統 100 的其他組件之間。系統 100 的元件（例如，圖形加速器 412、記憶體控制器集線器 416、記憶體 420、I/O 控制器集線器 430、無線收發器 426、快閃 BIOS 428、網路控制器 434、音訊控制器 436、串列擴充埠 438、I/O 控制器 430 等等）施行熟悉該技術者所熟知的其等習知功能。

在一項實施方案中，處理器 102 包括級別 1（L1）內部快取 404。依據該架構，處理器 102 可具有單一內部快取或多層內部快取。其他實施方案包括依據特別之實施方案與需求之內部與外部快取兩者的組合。這些快取中的一或多者可設定為具有關聯性且可允許位址之鎖定。就其本

身而言，它可包括一項實施方案，由此，快取控制器實施鎖定與解鎖快取方式的原理，如本文中所說明。暫存器檔案 406 係用以儲存不同類型的資料於多種暫存器中，包括整數暫存器、浮動點暫存器、向量暫存器、組合暫存器、陰影暫存器、檢查點暫存器、狀態暫存器、以及指令指標暫存器。

執行單元 408，包括用以施行整數與浮動點操作的邏輯，亦存在於處理器 102 中。在一項實施方案中，處理器 102 包括用以儲存微碼的微碼（ucode）ROM，當執行該微碼時，其係用以施行用於特定巨型指令的演算法，或處理複雜的情節。在此，微碼可潛在地更新，以處理用於處理器 102 的邏輯錯誤/固定。就一項實施方案而言，執行單元 408 包括用以處理封包指令集 409 的邏輯。藉由將封包指令集 409 包括在通用處理器 102 的指令集中，連同用以執行指令的相關電路，由許多多媒體應用所使用的操作則可使用在通用處理器 102 中的封包資料來施行。因此，藉由使用用於施行操作於封包資料上之處理器資料匯流排的全寬度，可更有效率地加速且執行許多多媒體應用。這潛在地刪除傳送更小單元的資料通過處理器之資料匯流排以施行一或多個操作（一次一個資料元件）的需要。

執行單元 408 的替代實施方案亦可使用於微控制器、嵌入式處理器、圖形裝置、DSP、以及其他類型的邏輯電路。系統 100 包括記憶體 420。記憶體 420 包括動態隨機存取記憶體（DRAM）裝置、靜態隨機存取記憶體

(SRAM) 裝置、快閃記憶體裝置、或其他記憶體裝置。記憶體 420 儲存由欲由處理器 120 執行之資料訊號所代表的指令及／或資料。

記憶體控制器集線器 (MCH) 416 耦合至處理器匯流排 410 與記憶體 420。處理器 102 可經由處理器匯流排 410 通訊至 MCH 416。MCH 416 提供高頻寬記憶體路徑 418 至記憶體 420，以用於指令與資料之儲存且用於圖形命令、資料與紋理之儲存。MCH 416 用以引導資料訊號於處理器 102、記憶體 402、及系統 100 中的其他組件之間，且用以橋接在處理器匯流排 410、記憶體 420、及系統 I/O 422 之間的資料訊號。在一些實施方案中，系統邏輯晶片 416 可提供用於耦合至圖形控制器 412 的圖形埠。MCH 416 係經由記憶體介面 418 耦合至記憶體 420。圖形卡 412 係經由加速圖形埠 (AGP) 互連 414 而耦合至 MCH 416。也可在記憶體控制器集線器 416 中實施快取，以提供比記憶體 420 更快的反應。此快取可設定為具有關聯性且可允許位址之鎖定，如在本文中的說明。

系統 100 可使用專用的集線器介面匯流排 422，以將 MCH 416 耦合至 I/O 控制器集線器 (ICH) 430。ICH 430 經由局部 I/O 匯流排提供直接連接至一些 I/O 裝置。局部的 I/O 匯流排係為用於連接週邊裝置至記憶體 420、晶片組、及處理器 102 的高速 I/O 匯流排。一些實例係為音訊控制器、韌體集線器 (快閃 BIOS) 428、無線收發器 426、資料儲存器 424、含有使用者輸入與鍵盤介面的舊

有 I/O 控制器、串列擴充埠（諸如通用串列匯流排（USB））、及網路控制器 434。資料儲存裝置 424 可包含硬碟驅動器、軟碟驅動器、CD-ROM 裝置、快閃記憶體裝置、或其他大量儲存裝置。

圖 5 繪示根據本揭露一項實施方案之處理器暫存器中之多種封包資料類型的呈現。圖 5 繪示用於 128 位元寬運算元之封包位元組 510、封包字組 520、以及封包雙字組（dword）530 的資料類型。此實例的封包位元組格式 510 係為 128 位元長且含有十六個封包位元組的資料元件。一位元組在此定義為 8 位元的資料。用於各位元組資料元件的資訊係儲存在用於位元組 0 的位元 7 至位元 0、用於位元組 1 的位元 15 至位元 8、用於位元組 2 的位元 23 至位元 16、以及最後用於位元組 15 的位元 120 至位元 127。因此，全部的有效位元均使用於暫存器中。此儲存排列增加處理器的儲存效率。並且，由於存取十六個資料元件，所以現在可在十六個資料元件上平行施行一個操作。

通常來說，一資料元件係為儲存在具有相同長度之其他資料元件的單一暫存器或記憶體位置的各別資料片。在與 SSEx 科技有關的封包資料序列中，儲存在 XMM 暫存器中之資料元件的數目是 128 位元除以各別資料元件的長度（以位元為單位）。同樣地，在與 MMX 及 SSE 科技有關的封包資料序列中，儲存在 MMX 暫存器中之資料元件的數目是 64 位元除以各別資料元件的長度（以位元為單位）。雖然在圖 5 中繪示的資料類型是 128 位元長，但是

實施方案亦可以 64 位元寬或其他尺寸的運算元來操作。此實例的封包字組格式 520 是 128 位元長且含有八個封包字組資料元件。各封包字組含有十六位元的資訊。圖 5 的封包雙字組格式 530 是 128 位元長且含有四個封包雙字組資料元件。各封包雙字組資料元件含有三十二位元的資訊。封包四字組是 128 位元長且含有兩個封包四字組的資料元件。

圖 6A 係為繪示根據本揭露至少一項實施方案之依序管線與暫存器更名階段、亂序發出/執行管線的方塊圖。圖 6B 係為繪示根據本揭露至少一項實施方案之包括在處理器中之依序架構核心與暫存器更名邏輯、亂序發出/執行邏輯的方塊圖。在圖 6A 中的實線盒繪示依序管線，而虛線盒繪示暫存器更名、亂序發出/執行管線。相同地，在圖 6B 中的實線盒繪示依序架構邏輯，而虛線盒繪示暫存器更名與亂序發出/執行邏輯。

在圖 6A 中，處理器管線 600 包括提取階段 602、長度解碼階段 604、解碼階段 606、分配階段 608、更名階段 610、排程（亦稱為配送或發出）階段 612、暫存器讀取/記憶體讀取階段 614、執行階段 616、回寫/記憶體寫入階段 618、異常處理階段 622、及提交階段 624。

在圖 6B 中，箭頭標示兩或更多個單元之間的耦合，且箭頭的方向指示在那些單元之間資料流的方向。圖 6B 顯示包括耦合至執行引擎單元 650 之前端單元 630 的處理器核心 690，且兩者均耦合至記憶體單元 670。

核心 690 可以是精簡指令集計算（RISC）核心、複雜指令集計算（CISC）核心、超長指令字組（VLIW）核心、或混成或替代核心類型。作為仍另一個選項，核心 690 可以是專用核心，諸如，例如，網路或通訊核心、壓縮引擎、圖形核心、或類似物。

前端單元 630 包括耦合至指令快取單元 634 的分支預測單元 632，該指令快取單元耦合至指令轉譯後備緩衝器（TLB）636，該指令轉譯後備緩衝器耦合至指令提取單元 638，該指令提取單元耦合至解碼單元 640。解碼單元或解碼器可解碼指令，且產生一或多個微操作、微碼進入點、微指令、其他指令、或其他控制訊號，以作為輸出，其係從原始指令解碼，或另外反映原始指令，或從原始指令取得。解碼器可使用許多不同的機構來實施。適當機構的實例包括但不限於查找表、硬體實施方案、可程式化邏輯陣列（PLA）、微碼唯讀記憶體（ROM）等等。指令快取單元 634 進一步耦合至記憶體單元 670 中的級別 2（L2）快取單元 676。解碼單元 640 耦合至在執行引擎單元 650 中的更名/分配器單元 652。

執行引擎單元 650 包括更名/分配器單元 652，該更名/分配器單元耦合至退休單元 654 及一組一或多個排程器單元 656。排程器單元 656 代表任何數目的不同排程器，其包括保留站、中央指令窗等等。排程器單元 656 係耦合至實體暫存器檔案單元 658。實體暫存器檔案單元 658 之各者代表一或多個實體暫存器檔案，該等檔案中的不同檔

案儲存一或多個不同資料類型，諸如純量整數、純量浮動點、封包整數、封包浮動點、向量整數、向量浮動點、狀態（例如，係為欲執行之下一指令位址的指令指示器）等等。實體暫存器檔案單元 658 與退休單元 654 重疊，以繪示可實施暫存器更名與亂序執行的許多方式（例如，使用重排序緩衝器及退休暫存器檔案，使用未來檔案、歷史緩衝器、及退休暫存器檔案；使用暫存器地圖及大批的暫存器；等等）。通常來說，從處理器外面或從程式設計員的觀點可看見架構性暫存器。暫存器不限於任何已知特別類型的電路。多種不同類型的暫存器是合適的，只要它們能夠儲存且提供在本文中所說明的資料。合適暫存器的實例包括但不限於專屬的實體暫存器、使用暫存器更名的動態分配實體暫存器、專屬與動態分配實體暫存器之組合等等。退休單元 654 與實體暫存器檔案單元 658 係耦合至執行叢集 660。執行叢集 660 包括一組一或多個執行單元 162 以及一組一或多個記憶體存取單元 664。執行單元 662 可施行多種操作（例如，移位、加法、減法、乘法）並在多種類型的資料上施行（例如，純量浮動點、封包整數、封包浮動點、向量整數、向量浮動點）。雖然一些實施方案可包括專門用於具體函數或函數組的一些執行單元，但是其他實施例方案可包括一個執行單元或者全部施行全部函數的數個執行單元。排程器單元 656、實體暫存器檔案單元 658、及執行叢集 660 係顯示為可能複數個，因為特定實施方案產生用於特定類型資料/操作的分開管

線（例如，純量整數管線、純量浮動點/封包整數/封包浮動點/向量整數/向量浮動點管線、及／或記憶體存取管線，各具有它們自己的排程器單元、實體暫存器檔案單元、及／或執行叢集一且在分開記憶體存取管線的情形中，實施特定實施方案，其中此管線的執行叢集具有記憶體存取單元 664）。亦應該理解的，在使用分開管線之處，一或多個這些管線可亂序發出/執行且剩下的依序。

該組記憶體存取單元 664 耦合至記憶體單元 670，其包括資料 TLB 單元 672，該資料 TLB 單元則耦合至資料快取單元 674，該資料快取單元則耦合至級別 2（L2）快取單元 676。在一項例示性實施方案中，記憶體存取單元 664 可包括載入單元、儲存位址單元、以及儲存資料單元，其中各者耦合至記憶體單元 670 中的資料 TLB 單元 672。L2 快取單元 676 耦合至一或多個其他級別的快取且最終耦合至主要記憶體。L2 快取單元 676 可設定為具有關聯性且可允許位址之鎖定，如本文中的說明。

舉例來說，暫存器更名、亂序發出/執行核心架構可實施管線 500 如下：1）指令提取 638 施行提取與長度解碼階段 602 與 604；2）解碼單元 640 施行解碼階段 606；3）更名/分配器單元 652 施行分配階段 608 與更名階段 610；4）排程器單元 656 施行排程階段 612；5）實體暫存器檔案單元 658 與記憶體單元 670 施行暫存器讀取/記憶體讀取階段 614；執行叢集 660 施行執行階段 616；6）記憶體單元 670 與實體暫存器檔案單元 658 施行回寫/記

記憶體寫入階段 618；7) 多種單元可涉及於異常處理階段 622；以及 8) 退休單元 654 與實體暫存器檔案單元 658 施行提交階段 624。

核心 690 可支持一或多個指令集（例如，x86 指令集（具有已添加較新版本的一些擴充程式）；Sunnyvale, CA 的 MIPS Technologies 的 MIPS 指令集；Sunnyvale, CA 的 ARM Holdings 的 ARM 指令集（具有額外擴充程式，諸如 NEON））。

雖然暫存器更名係在亂序執行的情境中說明，但是應該理解暫存器更名可使用於依序架構中。雖然處理器的繪示實施方案也包括分開的指令與資料快取單元 634/674 與共享的 L2 快取單元 676，但是替代性實施方案可具有用於指令與資料兩者的單一內部快取，諸如例如級別 1（L1）內部快取或多數級別的內部快取。在一些實施方案中，該系統可包括內部快取以及在核心及／或處理器外部的快取之組合。或者，全部的快取均可在核心及／或處理器外面。

圖 7 係為用於處理器 700 之微架構的一方塊圖，該處理器包括用以根據本揭露一項實施方案來施行指令的邏輯電路。在一些實施方案中，可實施根據一項實施方案的指令，以在具有位元組、字組、雙字組、四字組等等之大小以及資料形式（諸如單一與雙重準確整數與浮動點資料形式）的資料元件上操作。在一項實施方案中，依序前端 701 是處理器 700 的一部份，該處理器提取欲執行的指令

且準備稍後於處理器管線中使用。前端 701 可包括數個單元。在一項實施方案中，指令預取器 726 從記憶體提取指令，且將其等饋送到指令解碼器 728，該指令解碼器接著將其等解碼或解譯。例如，在一項實施方案中，解碼器將所收到的指令解碼成機器可執行的一或多個操作，該等操作稱為「微指令」或「微操作」（也稱為微 op 或 uops）。在其他實施方案中，解碼器將指令剖析成操作代碼（opcode）及對應的資料及控制欄位，其係由微架構所使用以施行根據一項實施方案的操作。在一項實施方案中，軌跡快取 730 進行經解碼的 uops 且在微操作（uop）佇列 734 中將其等組合成程式排序序列或軌跡以用於執行。當軌跡快取 730 遇到複雜的指令時，微碼 ROM 732 則提供完成該操作所必要的 uops。

一些指令轉換成單一個微操作，而其他則需要數個微操作來完成全部操作。在一項實施方案中，假如需要大於 4 個的微操作來完成指令，解碼器 728 存取微碼 ROM 732 以進行指令。就一項實施方案而言，可將指令解碼成少數個微操作，以用於在指令解碼器 728 上處理。在另一實施方案中，萬一需要一些微操作來完成操作，則可將指令儲存在微碼 ROM 732 內。軌跡快取 730 意指進入點可程式化邏輯陣列（PLA），以決定用於讀取微碼序列的正確微指令指向器，以從微碼 ROM 732 完成根據一項實施方案的一或多個指令。在微碼 ROM 732 結束定序用於一指令的微操作之後，機器的前端 701 則恢復從軌跡快取 730 提

取微操作。

亂序執行引擎 703 係為準備用來執行指令之處。亂序執行邏輯具有一些緩衝器，以使指令流平滑且重新排序，以當它們使管線下降且排程以用來執行時將性能最佳化。分配器邏輯將各 uop 為了執行所需要的機器緩衝器與資源分配。暫存器更名邏輯更名暫存器檔案中之入口上的邏輯暫存器。在指令排程器（記憶體排程器、快排程器 702、慢/一般浮動點排程器 704、以及簡單浮動點排程器 706）之前，分配器也分配一入口以用於在兩個 uop 佇列（一個用於記憶體操作且一個用於非記憶體操作）之一個中的各 uop。Uop 排程器 702、704、706 決定 uop 何時準備執行，其依據它們相依的輸入暫存器運算元來源之準備度以及 uops 用以完成其等操作所需要之執行資源的可用性。一項實施方案的快排程器 702 可在主要時鐘循環的各半循環上排程，而其他排程器則可每逢一主要的處理器時鐘循環排程一次。排程器仲裁配送口，以將用於執行的 uops 排程。

暫存器檔案 708、710 位於排程器 702、704、706 與在執行方塊 711 中的執行單元 712、714、716、718、720、722、724 之間。有各別用於整數與浮動點操作的分開暫存器檔案 708、710。一項實施方案的各暫存器檔案 708、710 也包括旁通網路，該旁通網路可將尚未寫入於暫存器檔案內的剛完成結果旁通或前傳到新的相依 uops。整數暫存器檔案 708 與浮動點暫存器檔案 710 也能夠彼此

通訊資料。就一項實施方案而言，整數暫存器檔案 708 分成兩個分開的暫存器檔案，一個暫存器檔案用於低階的 32 位元資料，且第二暫存器檔案用於高階的 32 位元資料。因為浮動點指令一般具有寬度從 64 至 128 位元的運算元，所以一項實施方案的浮動點暫存器檔案 710 具有 128 位元寬的入口。

執行方塊 711 含有實際上執行指令的執行單元 712、714、716、718、720、722、724。此區段包括暫存器檔案 708、710，該等暫存器檔案儲存微指令必須執行的整數與浮動點資料運算元值。一項實施方案的處理器 700 包含許多執行單元：位址產生單元（AGU）712、AGU 714、快 ALU 716、快 ALU 718、慢 ALU 720、浮動點 ALU 722、浮動點移動單元 724。就一項實施方案而言，浮動點執行方塊 722、724 執行浮動點、MMX、SIMD、及 SSE、或其他操作。一項實施方案的浮動點 ALU 722 包括 64 位元乘以 64 位元的浮動點除法器，以執行除法、平方根、以及剩下的微操作。就本揭露的實施方案而言，涉及浮動點值的指令可以浮動點硬體處理。在一項實施方案中，ALU 操作前進到高速 ALU 執行單元 716、718。一項實施方案的快 ALU 716、718 可執行具有半個時脈週期之有效延遲的快操作。就一項實施方案而言，因為慢 ALU 720 包括用於長延遲類型操作（諸如乘法、移位、旗標邏輯、以及分支處理）的整數執行硬體，所以大多數的複雜整數操作則前進到慢 ALU 720。記憶體載入/儲存操作係由 AGU

712、714 所執行。就一項實施方案而言，在 64 位元資料運算元上施行整數運算的情境中，將整數 ALU 716、718、720 說明。在替代性實施方案中，可實施 ALU 716、718、720，以支持包括 16、32、128、756 等等的多種資料位元。類似地，可實施浮動點單元 722、724，以支持具有多種寬度位元的運算元範圍。就一項實施方案而言，在結合 SIMD 與多媒體指令之 128 位元寬的封包資料運算元上，可操作浮動點單元 722、724。

在一項實施方案中，在上層（parent）載入已經結束執行之前，uops 排程器 702、704、706 配送相依的操作。因為 uops 在處理器 700 中經臆測性排程與執行，所以處理器 700 也包括用以處理記憶體遺漏的邏輯。假如資料載入在資料快取中遺漏，在已經離開具有暫時不正確資料之排程器之管線中的飛行中，則會有相依的操作。重播機構追蹤且重新執行使用不正確資料的指令。應該將相依的操作重播，且促使獨立的操作完成。處理器之一項實施方案的排程器與重播機構也經設計，以捕獲用於文字串比較操作的指令序列。

術語「暫存器」可意指使用當作用以識別運算元之一部份指令的板上處理器儲存位置。換言之，暫存器可以是可從處理器外面（從程式設計員的觀點）使用的那些。不過，一實施方案的暫存器在意義上應該不限於特別類型的電路。更確切地說，一實施方案的暫存器能夠儲存且提供資料，且實施在本文中說明的功能。藉由在使用任何數目

之不同技術之處理器內的電路（諸如專屬實體暫存器、使用暫存器更名的動態分配實體暫存器、專屬與動態分配實體暫存器之組合等等），可實施在本文中說明的暫存器。在一項實施方案中，整數暫存器儲存三十二位元的整數資料。一項實施方案的暫存器檔案也含有用於封包資料的八個多媒體 SIMD 暫存器。就下文的討論而言，暫存器被理解為經設計以保留封包資料的資料暫存器，諸如在以 MMX™ 科技（來自 Santa Clara, California 的 Intel Corporation）所致能之微處理器中的 64 位元寬 MMX 暫存器（在一些情形中，也稱為「mm」暫存器）。可用於整數與浮動點形式兩者中的這些 MMX 暫存器，可以伴隨 SIMD 與 SSE 指令的封包資料元件來操作。類似地，與 SSE2、SSE3、SSE4、或以上（通常稱為「SSEx」）科技有關的 128 位元寬 XMM 暫存器，也可用來保留此等封包資料運算元。在一項實施方案中，在儲存封包資料與整數資料時，暫存器不需要區別該等兩資料類型。在一項實施方案中，在相同暫存器檔案或不同暫存器檔案中，含有整數與浮動點。更者，在一項實施方案中，浮動點與整數資料可儲存在不同暫存器或相同暫存器中。

圖 8 係為根據本揭露實施方案之具有積體記憶體控制器與圖形之單核心處理器與多核心處理器 800 的方塊圖。在圖 8 中的實線盒繪示具有單核心 802A、系統媒介 810、一組一或多個匯流排控制器單元 816 的處理器 800，而虛線盒的添加繪示具有數個核心 802A-N、系統媒介單元

810 中的一組一或多個積體記憶體控制器單元 814、以及積體圖形邏輯 808 的替代處理器 800。

記憶體階層包括在核心內的一或多個級別快取、一組一或多個共享快取單元 806、以及耦合至該組積體記憶體控制器單元 814 的外部記憶體（未圖示）。該組共享快取單元 806 可包括一或多個中級別快取，諸如級別 2（L2）、級別 3（L3）、級別 4（L4）、或其他級別快取、最後級別快取（LLC）、及／或其組合。此組共享快取單元 806 可設定為具有關聯性且可允許位址之鎖定，如本文中的說明。雖然在一項實施方案中，環狀互連單元 812 將積體圖形邏輯 808、該組共享快取單元 806、以及系統媒介單元 810 互連，但是替代實施方案則可使用任何數目的著名技術來將此等單元互連。

在一些實施方案中，核心 802A-N 中的一或多個能夠多線程化。

系統媒介 810 包括協調與操作核心 802A-N 的那些組件。系統媒介單元 810 可例如包括功率控制單元（PCU）與顯示單元。PCU 可以是或可包括用來調節核心 802A-N 與積體圖形邏輯 808 之功率狀態所需要的邏輯與組件。顯示單元係用於驅動一或多個外部連接顯示器。

在架構及／或指令集方面，核心 802A-N 可以是均質或異質的。例如，核心 802A-N 中的一些可以是依序的，而其他者則是亂序的。如另一實例，核心 802A-N 中的兩個或更多個能夠執行相同指令集，而其他者則能夠執行那

指令集或不同指令集的子集。如進一步的實例，核心可以是不同的架構。

處理器可包括一或多個不同的通用處理器，諸如 Core™i3、i5、i7、2Duo 與 Quad、Xeon™、Itanium™、Atom™、XScale™或 StrongARM™處理器，其可從 Santa Clara, Calif.的 Intel Corporation 得到。例如，一個核心可能是 Corei7™核心，而該處理器的另一個核心可能是 Atom™核心。或者，處理器可來自另一公司，諸如 ARM Holdings, Ltd.、MIPS 等等。處理器可以是專用處理器，諸如例如網路或通訊處理器、壓縮引擎、圖形處理器、共用處理器、嵌入式處理器、或類似物。處理器可在一或多個晶片上實施。處理器 800 可以是一或多個基板的一部份及/或可使用一些製程技術之任一者而在一或多個基板上實施，諸如例如雙極互補金氧半導體（BiCMOS）、互補金氧半導體（CMOS）、或 N 型金氧半導體（NMOS）。

現在參考圖 9，其顯示根據本揭露之一實施方案之 SoC 900 的方塊圖。在圖 9 中的類似元件擁有相同的參考號碼。同樣地，虛線盒係為在更高級 SoC 上的特徵。在圖 9 中，互連單元 902 係耦合至：應用處理器 910，該處理器包括一組一或多個核心 902A-N 及共享快取單元 906；系統媒介單元 910；匯流排控制器單元 916；積體記憶體控制器單元 914；可包括積體圖形邏輯 908 的一組一或多個媒體處理器 920、用於提供靜止及/或攝影機功能的影像處理器 924、用於提供硬體音訊加速的音訊處理器 926、

以及用於提供視訊編碼/解碼加速的視訊處理器 928；靜態隨機存取記憶體（SRAM）單元 930；直接記憶體存取（DMA）單元 932；以及用於耦合至一或多個外部顯示器的顯示單元 940。

可以許多不同系統類型來實施實施方案。現在參考圖 10，其顯示根據一些實施方案之多處理器系統 1000 的方塊圖。如圖 10 所示，多處理器系統 1000 係為點對點互連系統，且包括經由點對點互連 1050 而耦合的第一處理器 1070 與第二處理器 1080。如圖 10 所示，處理器 1070 與 1080 之各者可能是多核心處理器，其包括第一與第二處理器核心（亦即，處理器核心 1074a 與 1074b 以及處理器核心 1084a 與 1084b），雖然許多更多核心潛在地存在於處理器中。處理器各者可包括根據本發明之實施方案的混合寫入模式邏輯。

雖然以兩個處理器 1070、1080 顯示，但是要理解，本揭露的範圍不會如此受到限制。在其他實施方案中，一或多個額外處理器可存在於已知的處理器中。

處理器 1070 與 1080 係顯示各別包括積體記憶體控制器單元 8102 與 8102。處理器 1070 也包括點對點（P-P）介面 1076 與 1078 以作為其匯流排控制器單元的一部份；類似地，第二處理器 1080 包括點對點介面 1086 與 1088。處理器 1070、1080 可經由使用點對點介面電路 1078、1088 的點對點（P-P）介面 1050 來交換資訊。如圖 10 所示，IMC 1072 與 IMC 1082 將處理器耦合至各別

記憶體（亦即，記憶體 1032 與記憶體 1034），其可以是局部附著到各別處理器之主要記憶體的部份。

處理器 1070、1080 各者可經由使用點對點介面電路 1076、1094、1086、1098 的各別點對點介面 1052、1054 而與晶片組 1090 交換資訊。晶片組 1090 也可經由高性能圖形介面 1039 而與高性能圖形電路 1038 交換資訊。

共享快取（未圖示）可包括在處理器中或兩處理器外，但仍經由點對點互連而與處理器連接，使得假如將處理器置於低功率模式，任一或兩處理器的區域快取資訊可能可儲存在共享快取中。

晶片組 1090 可經由介面 1096 耦合至第一匯流排 1016。在一項實施方案中，第一匯流排 1016 可以是週邊組件互連（PCI）匯流排，或者譬如 PCI 快速匯流排或另一第三代 I/O 互連匯流排的匯流排，雖然本揭露的範圍不會如此受到限制。

如在圖 10 中所示，多種 I/O 裝置 1014 可連同匯流排橋接器 1018 耦合至第一匯流排 1016，該匯流排橋接器將第一匯流排 1016 耦合至第二匯流排 1020。在一項實施方案中，第二匯流排 1020 可以是低接腳數（LPC）匯流排。多種裝置可耦合至第二匯流排 1020，例如包括鍵盤及／或滑鼠 1022、通訊裝置 1027 及儲存單元 1028（諸如磁碟驅動裝置或在一項實施方案中可包括指令／代碼及資料 1030 的其他大量儲存裝置）。更者，音訊 I/O 1024 可耦合至第二匯流排 1020。注意，其他架構是可能的。例

如，一系統可實施多點匯流排或其他此架構，而非圖 10 的點對點架構。

接著回到圖 11，其描繪根據本揭露實施方案之系統單晶片（SoC）設計的實施方案。如說明性實例，SOC 1100 包括在使用者設備（UE）中。在一項實施方案中，UE 意指由終端用戶用來通訊的任何裝置，諸如手持電話、智慧型手機、平板、超薄筆電、具有寬頻配接器的筆電、或任何其他類似的通訊裝置。UE 可連接到基地台或節點，其在本質上對應在 GSM 網路中的行動台（MS）。

在此，SOC 1100 包括 2 核心—1106 與 1107。與以上討論類似，核心 1106 與 1107 可符合指令集架構，諸如具有 Intel® Architecture Core™的處理器、Advanced Micro Devices, Inc.（AMD）處理器、MIPS 式的處理器、ARM 式的處理器設計、或其用戶、以及其等執照持有人或採用者。核心 1106 與 1107 係耦合至快取控制 1108，該快取控制與匯流排介面單元 1109 及 L2 快取 1110 有關聯，以與系統 1100 的其他部件通訊。互連 1111 包括晶片上互連，諸如 IOSF、AMBA、或上文所討論的其他互連，其可實施所說明揭露的一或多項態樣。

互連 1111 提供通訊通道至其他組件，諸如用以與 SIM 卡介面接合的用戶識別模組（SIM）1130、用以保留啟動碼以用於由核心 1106 與 1107 執行以啟始且啟動 SOC 1100 的啟動 ROM 1135、用以與外部記憶體（例如，DRAM 1160）介面接合的 SDRAM 控制器 1140、用以與持

續或非揮發性記憶體（例如，快閃 1165）介面接合的快閃控制器 1145、用以與週邊介面接合的週邊控制 1150（例如，序列週邊介面）、用以顯示與接收輸入（例如，觸控致能輸入）的視訊編解碼器 1120 與視訊介面 1125、用以施行與圖形相關之計算的圖形處理器（GPU）1115 等等。這些介面的任一者可結合在本文中所說明之實施方案的態樣。

此外，系統繪示用於通訊的週邊，諸如藍芽模組 1170、數據機 1175（例如，3G、4G、長期演進技術（LTE）、長期演進技術升級版等等）、全球定位系統（GPS）1180、無線上網（Wi-Fi）1185、紫蜂（Zigbee）（未圖示）、以及 Z-波（未圖示）。注意，如上文所述，UE 包括用於通訊的無線電。結果，這些週邊通訊模組可能不會全部包括在內。不過，在 UE 中，應該包括用於外部通訊之某種型式的無線電。

圖 12 繪示實例形式為計算系統 1200 之機器的示意代表圖，在該計算系統內，可將用於導致機器施行本文中所討論方法之任一者或多者的一組指令執行。在替代性實施方案中，機器可連接（例如，網路化）到在 LAN、內部網路、外部網路、或網際網路中的其他機器。該機器可在客戶端-伺服器網路環境中以伺服器或客戶裝置的身份或在對等式（或分佈式）網路環境中以對等機器來操作。該機器可以是個人電腦（PC）、平板電腦、機上盒（STB）、個人數位助理（PDA）、遊戲機、蜂巢式電話、數位相

機、手持個人電腦、網路設備、伺服器、網路路由器、交換器或橋接器、微控制器、數位訊號處理器（DSP）、系統單晶片、網路電腦（NetPC）、網路集線器、廣域網路（WAN）交換器、或能夠執行指定那機器所進行之動作之一組指令的任何機器（逐次或另外）。再者，雖然就系統架構 100 而言，只繪示單一個機器，但是術語「機器」也應包括各別或聯合執行一組（或多數組）指令以施行本文中所討論方法之任一者或多者的任何機器集合。實施方案不限於電腦系統。

計算系統 1200 包括處理裝置 1202、主要記憶體 1204（例如，唯讀記憶體（ROM）、快閃記憶體、動態隨機存取記憶體（DRAM）（諸如同步 DRAM（SDRAM）或 DRAM（RDRAM）等等））、靜態記憶體 1206（例如，快閃記憶體、靜態隨機存取記憶體（SRAM）等等）、以及資料儲存裝置 1216，其經由匯流排 1208 彼此通訊。

處理裝置 1202 代表一或多個通用處理裝置，諸如微處理器、中央處理單元、或類似物。更特別地，處理裝置可以是複雜的指令集計算（CISC）微處理器、精簡指令集電腦（RISC）微處理器、超長指令字組（VLIW）微處理器、或實施其他指令集的處理器、或實施指令集之組合的處理器。處理裝置 1202 也可以是一或多個專用處理裝置，諸如特殊應用積體電路（ASIC）、場可程式化閘陣列（FPGA）、數位訊號處理器（DSP）、網路處理器、或類似物。在一項實施方案中，處理裝置 1202 可包括一或

多個處理核心。處理裝置 1202 經組態以執行處理邏輯 1226，以用來施行本文中所討論的操作。在一項實施方案中，處理裝置 1202 可以是圖 1A 之系統架構 100 或圖 1B 之 SoC 190 的一部份。或者，計算系統 1200 可包括本文中所說明的其他組件。應理解，核心可支持多線程（執行兩或更多個平行組操作或線程），並且可以多種方式如此進行，包括分時間片多線程化、同時線程化（其中，單一的實體核心提供邏輯核心給實體核心正同時多線程化的各線程）或其組合（例如，之後譬如在 Intel® Hyperthreading 科技中的分時間片提取與解碼與同時多線程化）。

計算系統 1200 可代表以 PENTIUM III™、PENTIUM 4™、Celeron™、Xeon™、Itanium™、Xscale™、StrongARM™、Core™、Core2™、Atom™、及／或 Intel® Architecture Core™為基礎的處理系統，諸如可購自 Santa Clara, California 之 Intel Corporation 的 i3、i5、i7 微處理器，雖然也可使用其他系統（包括具有其他微處理器、工程化工作站、機上盒以及類似物的個人電腦）。不過，要理解，其他低功率處理器，諸如可購自 Sunnyvale, CA 的 Advanced Micro Devices, Inc. (AMD)、來自 Sunnyvale, CA 之 MIPS Technologies, Inc. 的 MIPS 式設計、ARM Holdings, Ltd. 或其用戶或其等執照持有人或採用者所授權的 ARM 式設計，則反而可能出現在其他實施方案中，諸如 Apple A5/A6 處理器、Qualcomm Snapdragon 處理器、

或 TI OMAP 處理器。在一項實施方案中，處理裝置 1202 執行可購自 Redmond, Washington 之 Microsoft Corporation 的 WINDOWS™ 操作系統版本，雖然也可使用其他操作系統（例如，OS X、UNIX、Linux、Android、iOS、Symbian）、嵌入式軟體、及／或圖形使用者介面。因此，本揭露的實施方案不限於硬體電路與軟體的任何具體組合。可在單一處理器桌面或伺服器系統的情境中說明一個實施方案，但是替代的實施方案則可包括在多處理器系統中。計算系統 1200 可以是「集線器」系統架構的實例。

計算系統 1200 可進一步包括可通信地耦合至網路 1218 的網路介面裝置 1222。計算系統 1200 也可包括顯示裝置 1210（例如，液晶顯示器（LCD）或陰極射線管（CRT））、文數輸入裝置 1212（例如，鍵盤）、游標控制裝置 1214（例如，滑鼠）、訊號產生裝置 1220（例如，喇叭）、或其他週邊裝置。更者，計算系統 1200 可包括圖形處理單元（未繪示）、視訊處理單元（未繪示）以及音訊處理單元（未繪示）。在另一實施方案中，計算系統 1200 可包括晶片組（未繪示），其意指一組積體電路或晶片，其經設計以與處理裝置 1202 一起操作，且控制處理裝置 1202 與外部裝置之間的通訊。例如，晶片組可能是主機板上的一組晶片，其連結處理裝置 1202 到非常高速的裝置，諸如主要記憶體 1204 與圖形控制器，以及連結處理裝置 1202 到週邊裝置之較低速度的週邊匯流

排，諸如通用序列匯流排（USB）、週邊組件互連（PCI）、或工業標準架構（ISA）匯流排。

資料儲存裝置 1216 可包括電腦可讀取儲存媒體 1224，在該媒體上係為經儲存的指令 1226，該等指令實施本文中所說明功能之方法的任一者或多者。在藉由計算系統 1200 來執行的期間內，指令 1226 也可完全或至少部份地存在於主要記憶體 1204 內，以作為指令 1226，及／或存在於處理裝置 1202 內，以作為處理邏輯 1226；主要記憶體 1204 與處理裝置 1202 也可組成電腦可讀取儲存媒體。

電腦可讀取儲存媒體 1224 也可用來儲存指令 1226，以利用處理裝置 1202（諸如相關於圖 1A 所說明的）、及／或含有呼叫以上應用程式之方法的軟體程式庫。雖然在一實例實施方案中，電腦可讀取儲存媒體 1224 顯示為單一媒體，但是術語「電腦可讀取儲存媒體」應該包括儲存一或多組指令的單一個媒體或數個媒體（例如，集中式或分布式資料庫、及／或相關聯的快取與伺服器）。術語「電腦可讀取儲存媒體」也應包括能夠儲存、編碼或承載供機器執行之一組指令且導致該機器施行本實施方案之方法之任一者或多者的任何媒體。術語「電腦可讀取儲存媒體」因此應包括但不限於固態記憶體以及光學與磁性媒體。

雖然本揭露相關於有限數目的實施方案被說明，但是所屬技術領域中具有通常知識者將理解源自那的許多修改

與變化。附加申請專利範圍旨在涵蓋在本揭露之真實精神與範圍內的所有此等修改與變化。

在本文中的說明中，陳述了許多具體細節，諸如具體類型的處理器與系統組態、具體硬體結構、具體結構性與微結構性細節、具體暫存器組態、具體指令類型、具體系統組件、具體測量/高度、具體處理器管線階段與操作等等之實例，以便提供本揭露的完整理解。不過，所屬技術領域中具有通常知識者將明瞭，不需要應用這些具體細節來實行本揭露。在其他情況中，著名的組件或方法，諸如具體與替代的處理器架構、用於所說明演算法的具體邏輯電路/代碼、具體韌體代碼、具體互連操作、具體邏輯組態、具體製造技術與材料、具體編譯器實施方案、用代碼之演算法的具體表達、具體功率下降與閘控技術/邏輯、以及電腦系統的其他具體操作性細節，則不會予以詳細說明，以便避免非必要地混淆本揭露。

該等實施方案係參考在具體積體電路中的混合穿線來說明，諸如在計算平台或微處理器中。該等實施方案也可應用到其他類型的積體電路與可程式化邏輯裝置。例如，所揭露的實施方案不限於桌上型電腦系統或可攜式電腦，諸如 Intel® Ultrabooks™ 電腦。而且也可使用在其他裝置中，諸如手持裝置、平板、其他薄筆電、系統單晶片（SOC）裝置、以及嵌入式應用。手持裝置的一些實例包括蜂巢式電話、網際網路協定裝置、數位相機、個人數位助理（PDA）、以及手持個人電腦（PC）。嵌入式應用一

般而言包括微控制器、數位訊號處理器（DSP）、系統單晶片、網路電腦（NetPC）、機上盒、網路集線器、廣域網路（WAN）切換器、或可施行以下教導之功能與操作的任何其他系統。據說，該系統可能是任何種類的電腦或嵌入式系統。所揭示的實施方案尤其可使用於低端裝置，像穿戴式裝置（例如，手錶）、電子植入物、感測與控制基礎結構裝置、控制器、監視控制與資料採集（SCADA）系統、或類似物。更者，在本文中所說明的設備、方法、及系統不限於實體計算裝置，但也可相關於用於能量保存與效率的軟體最佳化。如在下面的說明中將變得顯而易見的，在本文中所說明之方法、設備及系統的實施方案（不管是否參照硬體、韌體、軟體、或其組合）對於與性能考量平衡的「綠色科技」未來而言是至關重要的。

雖然在本文中的實施方案是參考處理器來說明的，但是其它實施方案則可應用於其它類型的積體電路與邏輯裝置。本揭露之實施方案的類似技術與教學可應用於可受益於較高管線流通量與改進性能的其它類型電路或半導體裝置。本揭露之實施方案的教學可應用於施行資料操縱的任何處理器或機器。不過，本揭露不限於施行 512 位元、256 位元、128 位元、64 位元、32 位元、或 16 位元資料操作的處理器或機器，且可應用於施行資料之操縱或管理的任何處理器與機器。此外，在本文中的說明提供實例，且附圖顯示用於繪示之目的多種實例。不過，這些實例不應以限制意義詮釋，因為它們僅旨在提供本揭露之實施方

案的實例，而不是提供本揭露之實施方案之全部可能實施方案的完整列表。

雖然以下的實例說明在執行單元與邏輯電路之情境中的指令處理與分佈，但是本揭露的其他實施方案則可透過儲存在機器可讀取、有形媒體上的資料或指令來實現，當由機器施行時，其導致機器施行與本揭露之至少一個實施方案一致的功能。在一項實施方案中，與本揭露之實施方案相關聯的功能以機器可執行指令實施。該等指令可用來導致以該等指令程式化的通用或專用處理器施行本揭露之步驟。本揭露的實施方案可能以電腦程式產品或軟體來提供，其可能包括具有指令儲存於上的機器或電腦可讀取媒體，該等指令可用來程式化電腦（或其他電子裝置）以施行根據本揭露之實施方案的一或多個操作。或者，藉由含有用於施行該等操作之固定功能邏輯的具體硬體組件，或藉由經程式化之電腦組件與固定功能硬體組件的任何組合，可施行本揭露之實施方案的操作。

用來程式化邏輯以施行本揭露之實施方案的指令可儲存在系統中的記憶體內，諸如 DRAM、快取、快閃記憶體、或其他儲存器。再者，該等指令可經由網路或透過其他電腦可讀取媒體來分佈。因此，機器可讀取媒體可包括用於將形式為可由機器（例如，電腦）讀取的資訊儲存或傳送的任何機構，但不限於軟式磁片、光碟、唯讀光碟記憶體（CD-ROMs）、及磁光碟、唯讀記憶體（ROMs）、隨機存取記憶體（RAM）、可抹除可程式化唯讀記憶體

(EPROM) 、電性可抹除可程式化唯讀記憶體 (EEPROM) 、磁性或光學卡、快閃記憶體、或有形機器可讀取儲存器，其係使用於在網際網路上、經由電性、光學、聲音或其他形式傳播訊號（例如，載波、紅外線訊號、數位訊號等等）的資訊傳送。因此，電腦可讀取媒體包括任何類型的有形機器可讀取媒體，其適合用於將形式為可由機器（例如，電腦）讀取的電子指令或資訊儲存或傳送。

一項設計可能歷經多種階段，從生產，到模擬，到製造。代表設計的資料可能以一些方式來代表設計。首先，如在模擬方面有用的，可使用硬體描述語言或另一功能性描述語言來代表硬體。此外，可在設計過程的一些階段上產生具有邏輯及／或電晶體閘極的電路級別模型。更者，在某階段，大部分的設計會達到代表在硬體模型中之多種裝置之實體布局的資料級別。在使用習知半導體製造技術的情形中，就使用以產生積體電路的遮罩而言，代表硬體模型的資料可以是指定不同遮罩層上之多種特徵存在或不存在的資料。在該設計的任何呈現中，該資料可以任何形式的機械可讀取媒體儲存。記憶體或磁性或光學儲存器（諸如碟片），其係可以用以儲存經由光波或電波來發射之資訊的機械可讀取媒體，該光波或電波係經調變或以別的方式產生以發送此資訊。當將指示或承載該代碼或設計的電載波發射時，在施行電訊號之複製、緩衝、或重新發射的方面來說，可進行新的複製。因此，通訊提供者或

網路提供者則可儲存在有形、機械可讀取媒體上，至少暫時地，在一物體上，諸如編碼成載波的資訊上，以實施本揭露實施方案的技術。

在本文中所使用的模組意指硬體、軟體、及／或韌體的任何組合。舉個實例，模組包括與非暫態媒體有關聯的硬體（諸如微控制器），其用以儲存經調適以由微控制器執行的代碼。因此，在一項實施方案中，提到一模組，意指硬體，其係經具體組態以確認及／或執行在非暫態媒體上所保留的代碼。更者，在另一項實施方案中，模組之使用意指包括該代碼的非暫態媒體，其經具體調適以由微控制器所執行，以施行預定的操作。且如可推斷的，在仍另一項實施方案中，術語模組（在本實例中）可意指微控制器與非暫態媒體之組合。經常地，繪示為分開的模組邊界通常會改變且潛在地重疊。例如，第一與第二模組可共享硬體、軟體、韌體、或其組合，同時潛在地保留一些獨立硬體、軟體、或韌體。在一項實施方案中，術語邏輯之使用包括硬體，諸如電晶體、暫存器，或其他硬體，諸如可程式化邏輯裝置。

在一項實施方案中，用詞「經組態以」之使用，意指將一設備、硬體、邏輯、或元件排列、裝配、製造、提供銷售、進口及／或設計，以施行經指定或經決定的任務。在本實例中，沒正在操作的設備或其元件仍「經組態以」施行一指定任務，假如其經設計、耦合、及／或互連以施行該指定任務。舉個純粹為說明性的實例，在操作期間

內，邏輯閘可提供 0 或 1。但是「經組態以」提供致能訊號到時鐘的邏輯閘不包括可提供 1 或 0 的每一潛在邏輯閘。反之，邏輯閘係為以在操作期間 1 或 0 輸出用以致能時鐘的某種方式而耦合者。再一次注意，術語「經組態以」之使用不規定操作，但反而聚焦於設備、硬體、及／或元件的潛在狀態，其中在潛在狀態中，該設備、硬體、及／或元件經設計以當設備、硬體、及／或元件正被操作時施行特別的任務。

更者，在一項實施方案中，用詞「以」、「能夠/能夠以」、「及／或」「可操作以」之使用，意指某設備、邏輯、硬體、及／或元件以能夠以指定方式使用該設備、邏輯、硬體、及／或元件的此方式來設計。如上文所提及的，在一項實施方案中，以、能夠以、或可操作以的使用意指一設備、邏輯、硬體、及／或元件的潛在狀態，其中該設備、邏輯、硬體、及／或元件不被操作，但卻以能夠以指定方式來使用設備的此種方式被設計。

在本文中所使用的值包括數目、狀態、邏輯的狀態、或二進制邏輯的狀態之任何已知表示。邏輯位準、邏輯值、或邏輯的值之使用也經常稱為 1 (1's) 以及 0 (0's)，其簡單地表示二進制的邏輯狀態。例如，1 意指高的邏輯位準且 0 意指低的邏輯位準。在一項實施方案中，儲存單元，諸如電晶體或快閃單元，其可能能夠保持單一的邏輯的值或數個邏輯的值。不過，在電腦系統中之值的其他表示則早已被使用。例如，十進位的十也可用二

進位的值 1010 及十六進位的字 A 表示。十六進位的值也可用字首表示，諸如「0x」。因此，一值包括能夠保持在電腦系統中之資訊的任何表示。

更者，狀態可由值或部份的值表示。舉個實例，第一值，諸如邏輯的一，可表示預設或最初的狀態，而第二值，諸如邏輯的零，可表示非預設的狀態。此外，在一項實施方案中，術語重設與設定分別意指預設與更新的值或狀態。例如，預設的值潛在地包括高的邏輯的值，亦即重設，而更新的值潛在地包括低的邏輯的值，亦即設定。注意，值的任何組合可用來表示狀態的任何數目。

上文陳述之方法、硬體、軟體、韌體或代碼的實施方案，其可經由儲存在機器可存取、機器可讀取、電腦可存取、或電腦可讀取媒體上、可由處理元件來執行的指令或代碼所實施。非暫態機器可存取/可讀取媒體包括提供（亦即，儲存及／或發送）形式為可由機器（諸如電腦或電子系統）讀取之資訊的任何機構。例如，非暫態機器可存取媒體包括隨機存取記憶體（RAM），諸如靜態 RAM（SRAM）、或動態 RAM（DRAM）；唯讀記憶體（ROM）；磁性或光學儲存媒體；快閃記憶體裝置；電性儲存裝置；光學儲存裝置；聲音儲存裝置；其他形式的儲存裝置，其用於保持從暫態（傳播）訊號（例如，載波、紅外線訊號、數位訊號）收到的資訊；等等，其係與可自其接收資訊的非暫態媒體有所區別。

在整個本說明書中，提及「一項實施方案」或「一實

施方案」意味著，結合該實施方案來說明的特定特徵、結構、或特色，係包括在本揭露的至少一項實施方案中。因此，用詞「在一項實施方案中」或「在一實施方案中」之出現在整個本說明書的各處，不一定全部意指相同的實施方案。再者，在一或多項的實施方案中，特定特徵、結構、或特色可以任何適當的方式組合。

在本說明書中，實施方式已經參考具體實例的實施方案來產生。不過，顯然，可對其進行許多修改與改變而不脫離附加申請專利範圍中所陳述之揭露的更廣泛精神與範圍。因此，說明書與圖式則被視為說明性的意義而非限制性的意義。再者，實施方案與其他例示性語言的上述使用不一定意指相同實施方案或相同實例，但卻可意指不同且明顯的實施方案，以及潛在地相同的實施方案。

以下實例屬於進一步的實施方案。

實例 1 係為一種處理系統，其包含：控制暫存器；以及處理核心，其通訊地耦合至該控制暫存器；其中該處理核心經組態以：接收具體指定一應用的預期運行時間以及叢集節點之請求數目節點的節點分配請求；鑑於該節點分配請求以及該複數個節點上的電流負載，決定要分配給該應用的節點之實際數目，其中要分配給該應用的該實際數目節點將叢集負載標準最佳化；以及使用該控制暫存器，向該應用通知該節點之實際數目。

實例 2 係為實例 1 之處理系統，其中向該應用通知該實際數目的節點包含：將該節點的實際數目儲存在用來儲

存與該應用相關聯之節點之組態的記憶體資料結構中；以及將該記憶體資料結構的位址儲存在該控制暫存器中。

實例 3 係為實例 2 之處理系統，其中將該記憶體資料結構的該位址儲存在該控制暫存器中，包含獲得該控制暫存器之一鎖定。

實例 4 係為實例 1 之處理系統，其中該處理核心進一步組態以：導致該應用使用該節點的實際數目來執行。

實例 5 係為實例 1 之處理系統，其中該叢集負載標準反映在特定時段內保持未指派給任何目前運行之應用的數個節點。

實例 6 係為實例 1 之處理系統，其中該叢集負載標準反映在特定時段內保持未指派給任何運行之應用的數個節點中，乘以該時段。

實例 7 係為實例 1 之處理系統，其中決定該節點的實際數目進一步包含：產生將要排程之應用的複數個組合；決定用於各產生組合之該叢集負載標準的一值；以及選擇與該叢集負載標準之一最佳值有關聯的一組合。

實例 8 係為實例 1 之處理系統，其中該處理系統係以系統單晶片（SoC）實施。

實例 9 係為一種方法，其包含：藉由一處理裝置，接收具體指定應用的預期運行時間以及叢集節點之節點請求數目的節點分配請求；鑑於該節點分配請求以及該複數個節點上的電流負載，決定要分配給該應用的節點之實際數目，其中要分配給該應用的該節點之實際數目將叢集負載

標準最佳化；以及使用該處理系統的控制暫存器，向該應用通知該節點之實際數目。

實例 10 係為實例 9 之方法，其中向該應用通知該節點之實際數目包含：將該節點之實際數目儲存在用來儲存與該應用相關聯之節點之組態的記憶體資料結構中；以及將該記憶體資料結構的位址儲存在該控制暫存器中。

實例 11 係為實例 10 之方法，其中將該記憶體資料結構的該位址儲存在該控制暫存器中，包含獲得該控制暫存器之鎖定。

實例 12 係為實例 9 之方法，進一步包含：導致該應用使用該節點之實際數目來執行。

實例 13 係為實例 9 之方法，其中該叢集負載標準反映在特定時段內保持未指派給任何目前運行之應用的數個節點。

實例 14 係為實例 9 之方法，其中該叢集負載標準反映在特定時段內保持未指派給任何運行之應用的數個節點中，乘以該時段。

實例 15 係為實例 9 之方法，其中決定該等實際數目的節點進一步包含：產生將要排程之應用的複數個組合；決定用於各產生組合之該叢集負載標準的值；以及選擇與該叢集負載標準之最佳值有關聯的組合。

實例 16 係為一種設備，其包含：一記憶體；以及耦合至該記憶體的一處理裝置，該處理裝置用以施行實例 9 至實例 15 任一項的方法。

實例 17 係為一種包含可執行指令的非暫態電腦可讀取儲存媒體，該等可執行指令在由處理系統執行時導致該處理系統執行操作，該等操作包含：接收具體指定應用的預期運行時間以及叢集節點之節點請求數目的節點分配請求；鑑於該節點分配請求以及該複數個節點上的電流負載，決定要分配給該應用的節點之實際數目，其中要分配給該應用的該節點之實際數目將叢集負載標準最佳化；以及使用該處理系統的控制暫存器，向該應用通知該節點之實際數目。

實例 18 係為實例 17 的非暫態電腦可讀取儲存媒體，其中向該應用通知該節點之實際數目包含：將該節點之實際數目儲存在用來儲存與該應用相關聯之節點之組態的記憶體資料結構中；以及將該記憶體資料結構的位址儲存在該控制暫存器中。

實例 19 係為實例 18 的非暫態電腦可讀取儲存媒體，其中將該記憶體資料結構的該位址儲存在該控制暫存器中，包含獲得該控制暫存器之鎖定。

實例 20 係為實例 17 的非暫態電腦可讀取儲存媒體，進一步包含可執行指令，其導致該處理系統：導致該應用使用該節點之實際數目來執行。

實例 21 係為實例 17 的非暫態電腦可讀取儲存媒體，其中該叢集負載標準反映在特定時段內保持未指派給任何目前運行之應用的數個節點。

實例 22 係為實例 17 的非暫態電腦可讀取儲存媒體，

其中該叢集負載標準反映在特定時段內保持未指派給任何運行之應用的數個節點中，乘以該時段。

實例 23 係為實例 17 的非暫態電腦可讀取儲存媒體，其中決定該等實際數目的節點進一步包含：產生將要排程之應用的複數個組合；決定用於各產生組合之該叢集負載標準的值；以及選擇與該叢集負載標準之最佳值有關聯的組合。

實例 24 係為一種高性能計算系統，其包含：控制節點；以及計算叢集，其包含複數個計算節點；其中該控制節點經組態以：接收具體指定應用的預期運行時間以及該計算叢集之計算節點的請求數目之計算節點分配請求；鑑於該計算節點分配請求以及該等複數個計算節點上的一電流負載，決定要分配給該應用的計算節點之實際數目，其中要分配給該應用的該計算節點之實際數目將叢集負載標準最佳化；以及向該應用通知該計算節點之實際數目。

實例 25 係為實例 24 之高性能計算系統，其中該控制節點進一步經組態以：導致該應用使用該節點之實際數目來執行。

實例 26 係為實例 24 之高性能計算系統，其中向該應用通知該節點之實際數目包含：將該等節點之實際數目儲存在用來儲存與該應用相關聯之節點之組態的記憶體資料結構中；以及將該記憶體資料結構的位址儲存在該控制暫存器中。

實例 27 係為實例 24 之高性能計算系統，其中將該記

憶體資料結構的該位址儲存在該控制暫存器中，包含獲得該控制暫存器之鎖定。

實例 28 係為實例 24 之高性能計算系統，其中該叢集負載標準反映在特定時段內保持未指派給任何運行之應用的數個節點。

實例 29 係為實例 24 之高性能計算系統，其中該叢集負載標準反映在特定時段內保持未指派給任何運行之應用的數個節點中，乘以該時段。

實例 30 係為實例 24 之高性能計算系統，其中決定該等實際數目的節點進一步包含：產生將要排程之應用的複數個組合；決定用於各產生組合之該叢集負載標準的值；以及選擇與該叢集負載標準之最佳值有關聯的組合。

詳細說明的某些部分是根據電腦記憶體內之資料位元上之操作的演算法與符號代表來呈現。這些演算法說明與代表係為所屬資料處理技術領域中具有通常知識者用以將其等工作之主旨最有效地傳遞給其他所屬技術領域中具有通常知識者所使用的方法。演算法在此且通常地令人認為是導致希望結果的自相容的操作順序。該等操作係為需要物理量之物理操作的那些。通常，雖然並非必要地，這些量的形式為能夠經儲存、傳送、組合、比較以及另外操縱的電或磁訊號。有時，主要為了普通用法之因素，已經證明將這些訊號稱為位元、值、元件、符號、字元、項、號碼、或類似物是方便的。在本文中所說明的方塊可以是硬體、軟體、韌體或其組合。

不過，應該記住，所有這些與類似的術語將與適當的物理量相關聯，且僅僅是應用於這些量的方便標記。除非另有特別陳述（如從以上討論明顯可見的），應該明白，在整個說明中，利用術語（諸如「定義」、「接收」、「決定」、「發出」、「連結」、「關聯」、「得到」、「鑑別」、「禁止」、「執行」、「請求」、「通訊」、「監視」、「計算」、或類似物）的討論，意指計算系統或類似電子計算裝置的動作與過程，其將表示為計算系統的暫存器與記憶體內之物理（例如，電子）量的資料操縱且轉換為類似地表示為計算系統記憶體或暫存器或其他此等資訊儲存、傳輸或顯示裝置內之物理量的其他資料。

在本文中，字「實例」或「例示性」用來意味著充當實例、示例、或說明。在本文中說明為「實例」或「例示性」的任何態樣或設計不一定詮釋為比其他態樣或設計較佳或有利。相反的，字「實例」或「例示性」的使用旨在以具體方式呈現概念。如在本申請案中使用的，術語「或」旨在意味著包容性的「或」而不是排他性的「或」。也就是，除非另有規定，或從上下文清楚理解，「X 包括 A 或 B」旨在意味著自然包含性排列的任一者。也就是，假如 X 包括 A；X 包括 B；或 X 包括 A 與 B 兩者，那麼在以上示例的任一者之下，則滿足「X 包括 A 或 B」。此外，在本申請案中與附加申請專利範圍中使用的冠詞「一（a 與 an）」通常詮釋為意味著「一或多個」，除非另有規定或從上下文清楚是針對單數形式。更者，在

全文中，術語「一實施方案」或「一個實施方案」之使用並非旨在意味著相同實施方案或諸實施方案，除非如此說明。同樣地，在本文中使用的術語「第一」、「第二」、「第三」、「第四」等等意味著區分不同元件的標記，且可能不一定具有根據它們數字名稱的序數意義。

【符號說明】

1000：高性能計算系統

100：處理系統

102：處理器

110：主要節點

120A~120N：計算節點

122：控制暫存器

135：網路

140：批次排程器

150：訊息傳遞介面訊息

160：輸入工作佇列

170A~170K：應用

180：方塊

400：方法

402：記憶體

404：級別 1（L1）內部快取

406：暫存器檔案

408：執行單元

- 409：封包指令集
- 410：方塊
- 410：處理器匯流排
- 412：圖形視訊卡
- 412：圖形加速器
- 412：圖形控制器
- 414：加速圖形埠互連
- 416：記憶體控制器集線器
- 418：高頻寬記憶體路徑
- 418：記憶體介面
- 420：記憶體
- 420：方塊
- 422：系統輸入/輸出
- 422：集線器介面匯流排
- 424：資料儲存器
- 426：無線收發器
- 428：快閃基本輸出入系統
- 430：方塊
- 430：輸入/輸出控制器集線器
- 434：網路控制器
- 436：音訊控制器
- 438：串列擴充埠
- 440：方塊
- 440：舊有輸入/輸出控制器

- 450：方塊
- 510：封包位元組
- 520：封包字組
- 530：封包雙字組
- 600：處理器管線
- 602：提取階段
- 604：長度解碼階段
- 606：解碼階段
- 608：分配階段
- 610：更名階段
- 612：排程階段
- 614：暫存器讀取/記憶體讀取階段
- 616：執行階段
- 618：回寫/記憶體寫入階段
- 622：異常處理階段
- 624：提交階段
- 630：前端單元
- 632：分支預測單元
- 634：指令快取單元
- 636：指令轉譯後備緩衝器
- 638：指令提取單元
- 640：解碼單元
- 650：執行引擎單元
- 652：更名/分配器單元

- 654：退休單元
- 656：排程器單元
- 658：實體暫存器檔案單元
- 660：執行叢集
- 662：執行單元
- 664：記憶體存取單元
- 670：記憶體單元
- 672：資料轉譯後備緩衝器單元
- 674：資料快取單元
- 676：級別 2 快取單元
- 690：處理器核心
- 700：處理器
- 701：依序前端
- 702：快排程器
- 703：亂序執行引擎
- 704：慢/一般浮動點排程器
- 706：簡單浮動點排程器
- 708：暫存器檔案
- 710：暫存器檔案
- 711：執行方塊
- 712：執行單元
- 714：執行單元
- 716：執行單元
- 718：執行單元

- 720：執行單元
- 722：浮動點單元
- 724：浮動點單元
- 726：指令預取器
- 728：指令解碼器
- 730：軌跡快取
- 732：微碼唯讀記憶體
- 734：微操作佇列
- 800：處理器
- 802A~802N：核心
- 806：共享快取單元
- 808：積體圖形邏輯
- 810：系統媒介單元
- 812：環狀互連單元
- 814：積體記憶體控制器單元
- 816：匯流排控制器單元
- 900：系統單晶片
- 902：互連單元
- 902A~902N：核心
- 906：共享快取單元
- 908：積體圖形邏輯
- 910：應用處理器
- 910：系統媒介單元
- 914：積體記憶體控制器單元

- 916：匯流排控制器單元
- 920：媒體處理器
- 924：影像處理器
- 926：音訊處理器
- 928：視訊處理器
- 930：靜態隨機存取記憶體單元
- 932：直接記憶體存取單元
- 940：顯示單元
- 1000：多處理器系統
- 1014：輸入/輸出裝置
- 1016：第一匯流排
- 1018：匯流排橋接器
- 1020：第二匯流排
- 1022：鍵盤及/或滑鼠
- 1024：音訊輸入/輸出
- 1027：通訊裝置
- 1028：儲存單元
- 1030：指令/代碼及資料
- 1032：記憶體
- 1034：記憶體
- 1038：高性能圖形電路
- 1039：高性能圖形介面
- 1050：點對點互連
- 1052：點對點介面

- 1054：點對點介面
- 1070：第一處理器
- 1072：積體記憶體控制器單元
- 1074a：處理器核心
- 1074b：處理器核心
- 1076：點對點介面電路
- 1078：點對點介面
- 1080：第二處理器
- 1082：積體記憶體控制器單元
- 1084a：處理器核心
- 1084b：處理器核心
- 1086：點對點介面電路
- 1088：點對點介面
- 1090：晶片組
- 1094：點對點介面電路
- 1096：介面
- 1098：點對點介面電路
- 1100：系統單晶片
- 1106：核心
- 1107：核心
- 1108：快取控制
- 1109：匯流排介面單元
- 1110：L2 快取
- 1111：互連

- 1115：圖形處理器
- 1120：視訊編解碼器
- 1125：視訊介面
- 1130：用戶識別模組
- 1135：啟動 ROM
- 1140：SDRAM 控制器
- 1145：快閃控制器
- 1150：週邊控制
- 1160：動態隨機存取記憶體
- 1165：快閃
- 1170：藍芽模組
- 1175：數據機
- 1180：全球定位系統
- 1185：無線上網
- 1200：計算系統
- 1202：處理裝置
- 1204：主要記憶體
- 1206：靜態記憶體
- 1208：匯流排
- 1210：顯示裝置
- 1212：文數輸入裝置
- 1214：游標控制裝置
- 1216：資料儲存裝置
- 1218：網路

1220：訊號產生裝置

1222：網路介面裝置

1224：電腦可讀取儲存媒體

1226：處理邏輯

申請專利範圍

1. 一種處理系統，其包含：

控制暫存器；以及

處理核心，其通訊地耦合至該控制暫存器；

其中該處理核心用以：

接收具體指定一應用的預期運行時間以及叢集節點之請求數目節點的節點分配請求；

鑑於該節點分配請求以及該複數個節點上的電流負載，決定要分配給該應用的節點實際數目，其中要分配給該應用的該節點實際數目將叢集負載標準最佳化；以及

使用該控制暫存器，向該應用通知該節點實際數目。

2. 如申請專利範圍第 1 項之處理系統，其中向該應用通知該節點實際數目包含：

將該節點實際數目儲存在用來儲存與該應用相關聯之節點之組態的記憶體資料結構中；以及

將該記憶體資料結構的位址儲存在該控制暫存器中。

3. 如申請專利範圍第 2 項之處理系統，其中將該記憶體資料結構的該位址儲存在該控制暫存器中，包含獲得該控制暫存器之鎖定。

4. 如申請專利範圍第 1 項之處理系統，其中該處理核心進一步用以：

使用該節點實際數目來執行該應用。

5. 如申請專利範圍第 1 項之處理系統，其中該叢集負載標準反映在特定時段內保持未指派給任何目前運行中

之應用的數個節點。

6. 如申請專利範圍第 1 項之處理系統，其中該叢集負載標準反映在特定時段內保持未指派給任何運行中之應用的數個節點乘以該時段。

7. 如申請專利範圍第 1 項之處理系統，其中決定該節點實際數目進一步包含：

產生將要排程之應用的複數個組合；

決定用於產生的組合各者之該叢集負載標準的一值；

以及

選擇與該叢集負載標準之一最佳值有關聯的一組合。

8. 一種包含可執行指令的非暫態電腦可讀取儲存媒體，該等可執行指令在由處理系統執行時使該處理系統執行操作，該等操作包含：

接收具體指定應用的預期運行時間以及叢集節點之節點請求數目的節點分配請求；

鑑於該節點分配請求以及該複數個節點上的電流負載，決定要分配給該應用的節點實際數目，其中要分配給該應用的該節點實際數目將叢集負載標準最佳化；以及

使用該處理系統的控制暫存器，向該應用通知該節點之實際數目。

9. 如申請專利範圍第 8 項之非暫態電腦可讀取儲存媒體，其中向該應用通知該節點實際數目包含：

將該節點實際數目儲存在用來儲存與該應用相關聯之節點組態的記憶體資料結構中；以及

將該記憶體資料結構的位址儲存在該控制暫存器中。

10. 如申請專利範圍第 9 項之非暫態電腦可讀取儲存媒體，其中將該記憶體資料結構的該位址儲存在該控制暫存器中，包含獲得該控制暫存器之鎖定。

11. 如申請專利範圍第 8 項之非暫態電腦可讀取儲存媒體，進一步包含可執行指令，其使該處理系統用以：
使用該節點實際數目來執行該應用。

12. 如申請專利範圍第 8 項之非暫態電腦可讀取儲存媒體，其中該叢集負載標準反映在特定時段內保持未指派給任何目前運行中之應用的數個節點。

13. 如申請專利範圍第 8 項之非暫態電腦可讀取儲存媒體，其中該叢集負載標準反映在特定時段內保持未指派給任何運行中之應用的數個節點乘以該時段。

14. 如申請專利範圍第 8 項之非暫態電腦可讀取儲存媒體，其中決定該節點實際數目進一步包含：

產生將要排程之應用的複數個組合；

決定用於產生之組合各者之該叢集負載標準的值；以及

選擇與該叢集負載標準之最佳值有關聯的組合。

15. 一種高性能計算系統，其包含：

控制節點；以及

計算叢集，其包含複數個計算節點；

其中該控制節點用以：

接收具體指定應用的預期運行時間以及該計算叢集之

計算節點的請求數目之計算節點分配請求；

鑑於該計算節點分配請求以及該等複數個節點上的電流負載，決定要分配給該應用的計算節點實際數目，其中分配給該應用的該計算節點實際數目將叢集負載標準最佳化；以及

向該應用通知該計算節點實際數目。

16. 如申請專利範圍第 15 項之高性能計算系統，其中該控制節點進一步用以：

使用該節點實際數目來執行該應用。

17. 如申請專利範圍第 15 項之高性能計算系統，其中向該應用通知該節點實際數目包含：

將該等節點實際數目儲存在用來儲存與該應用相關聯之節點組態的記憶體資料結構中；以及

將該記憶體資料結構的位址儲存在該控制暫存器中。

18. 如申請專利範圍第 17 項之高性能計算系統，其中將該記憶體資料結構的該位址儲存在該控制暫存器中，包含獲得該控制暫存器之鎖定。

19. 如申請專利範圍第 15 項之電腦可讀取儲存媒體，其中該叢集負載標準反映在特定時段內保持未指派給任何運行中之應用的數個節點。

20. 如申請專利範圍第 15 項之高性能計算系統，其中決定該等節點之實際數目進一步包含：

產生將要排程之應用的複數個組合；

決定用於產生組合各者之該叢集負載標準的值；以及

選擇與該叢集負載標準之最佳值有關聯的組合。

圖式

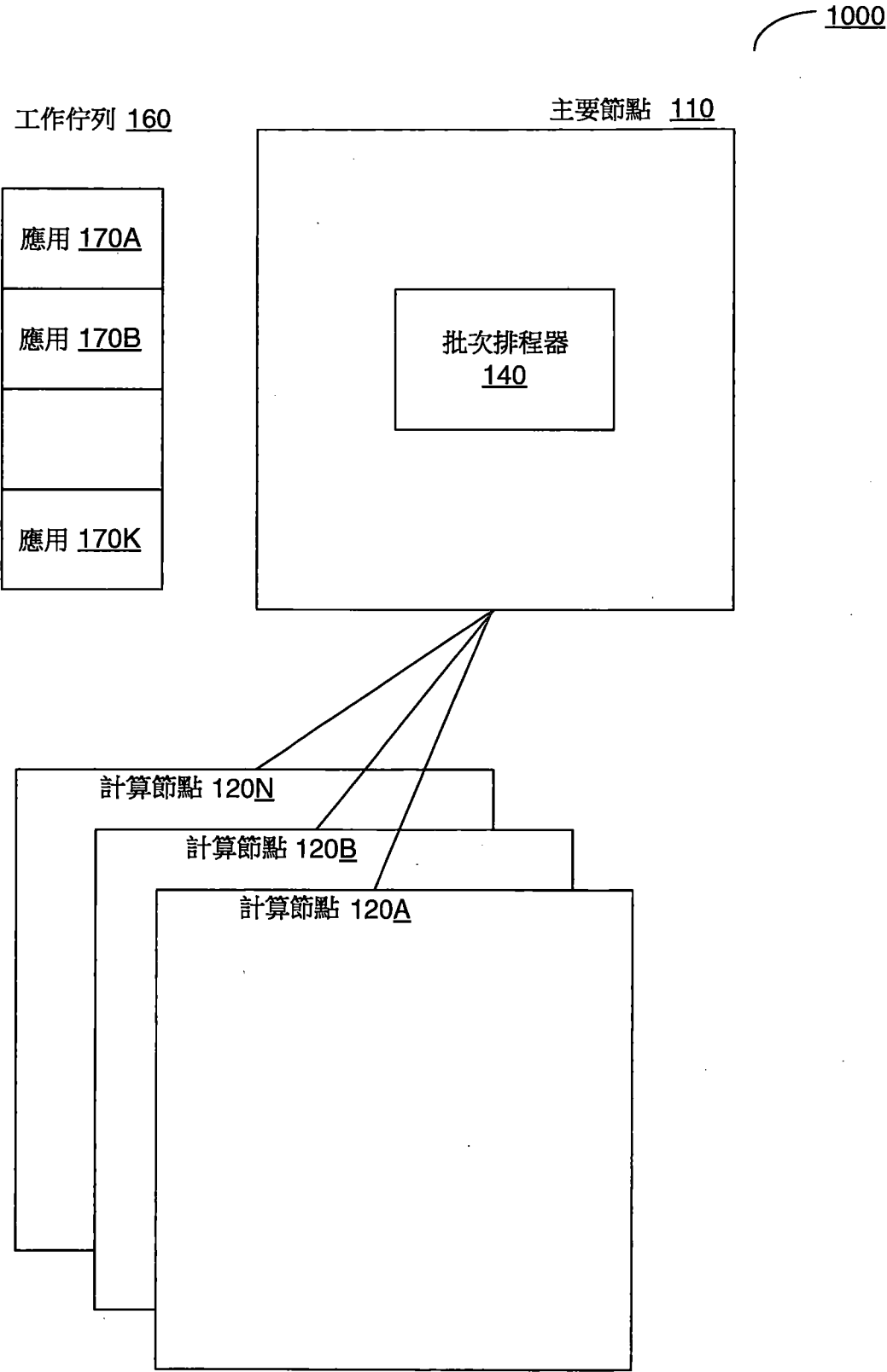


圖 1

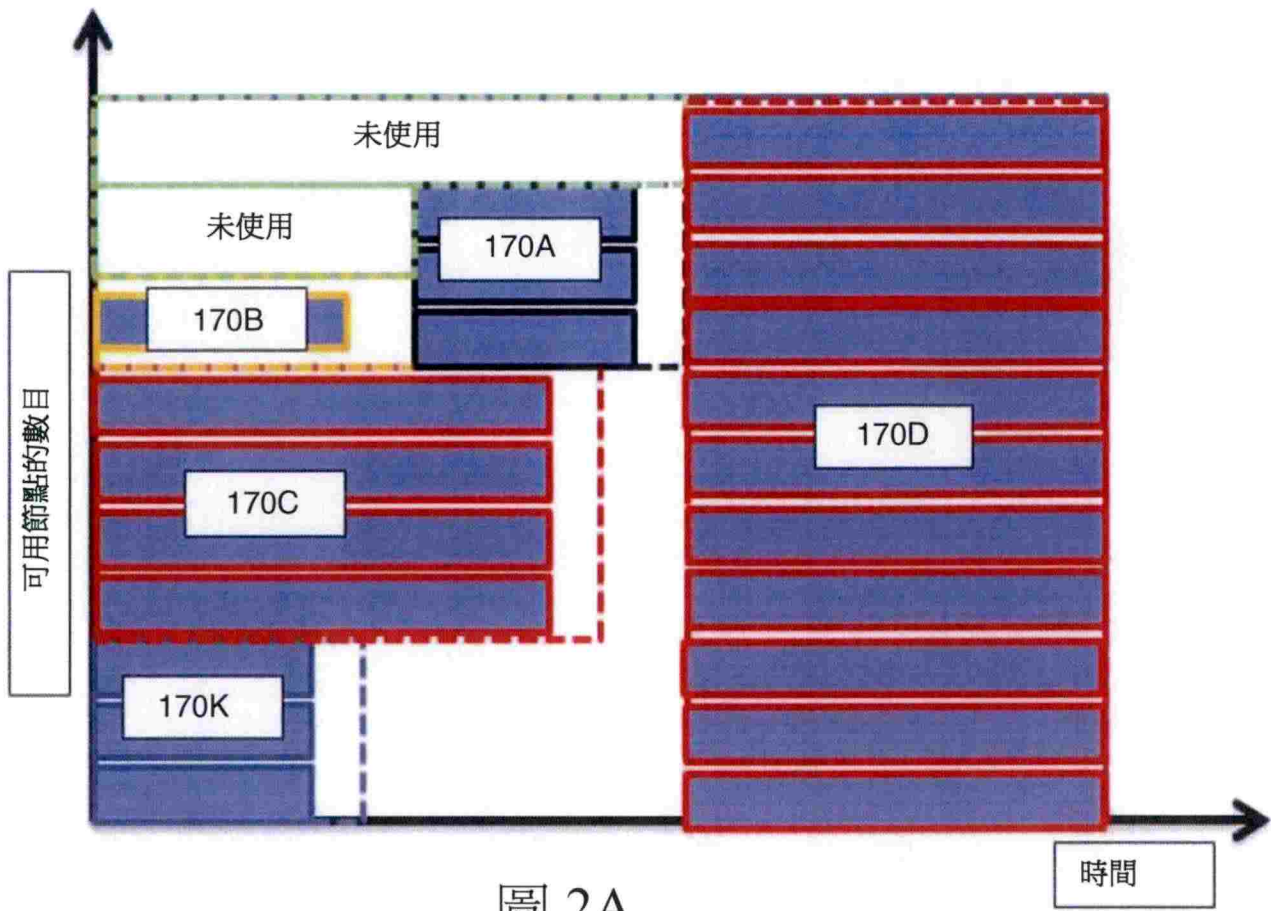


圖 2A

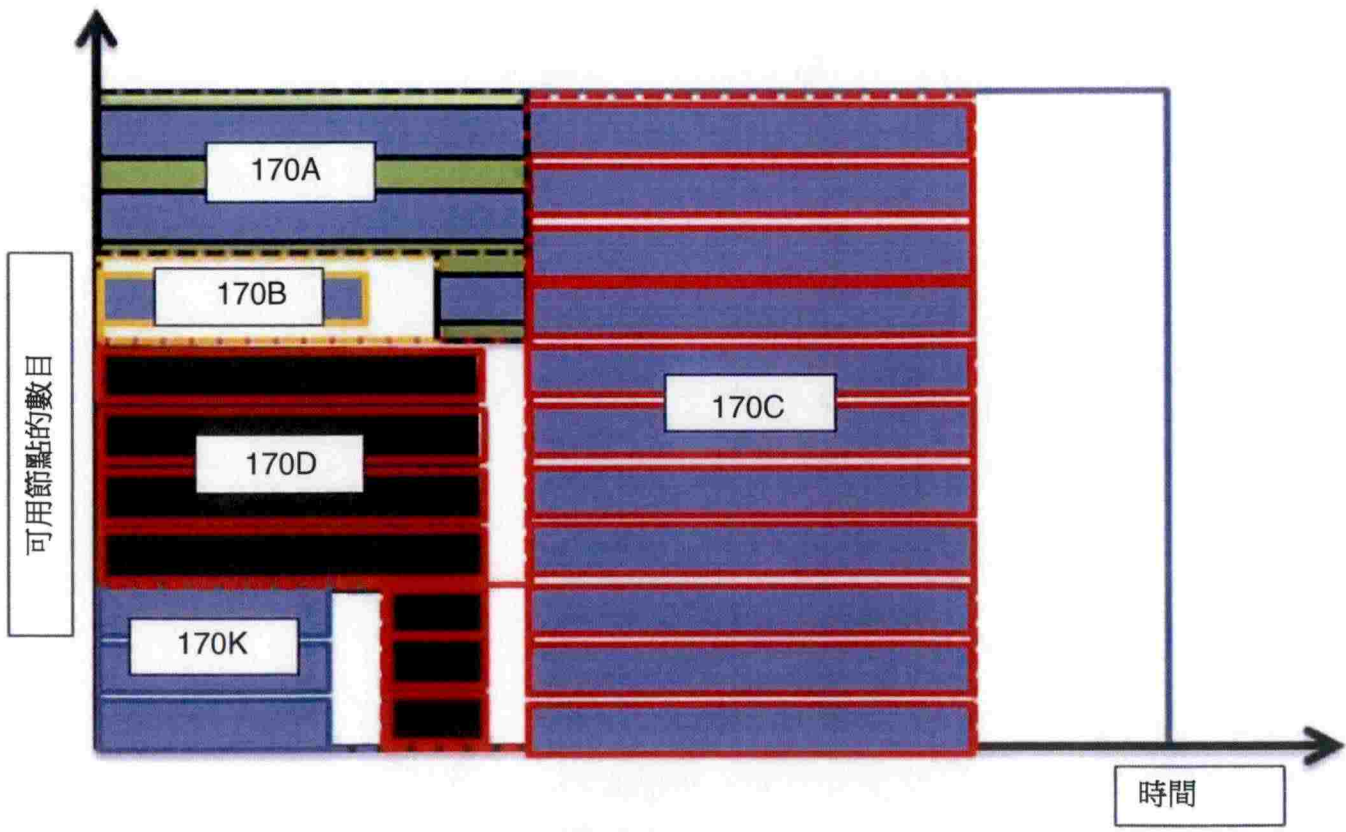


圖 2B

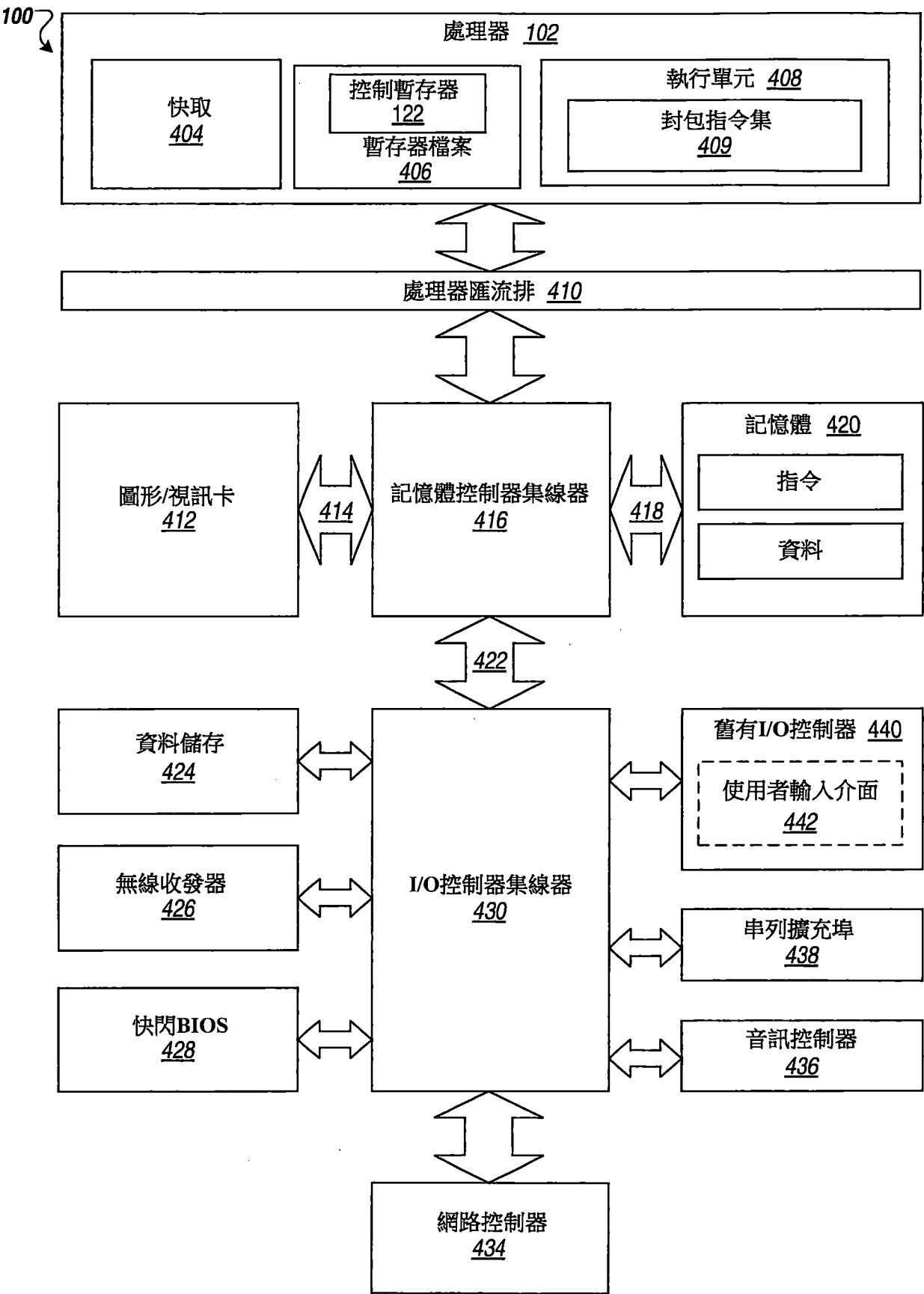


圖 3

400 ↘

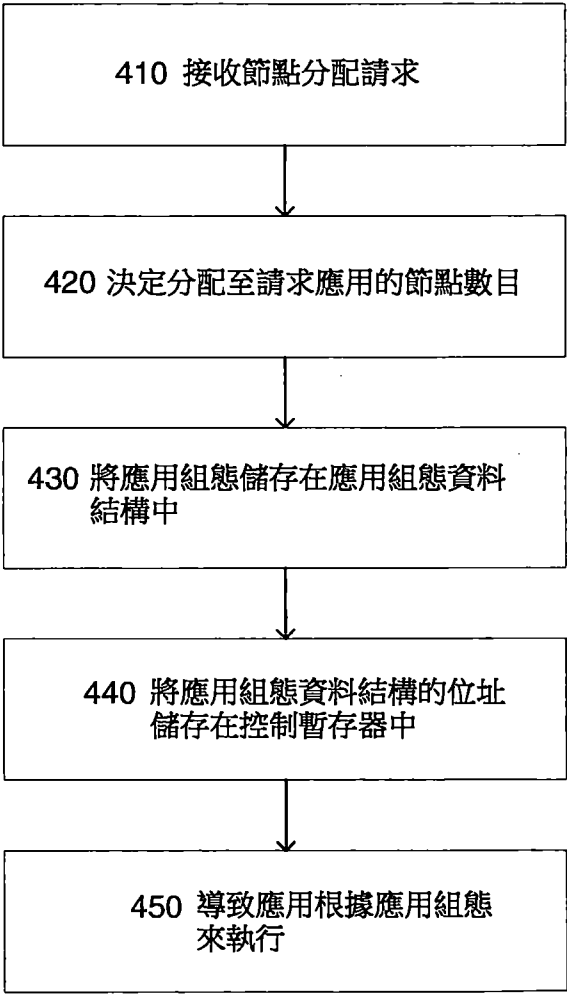


圖 4

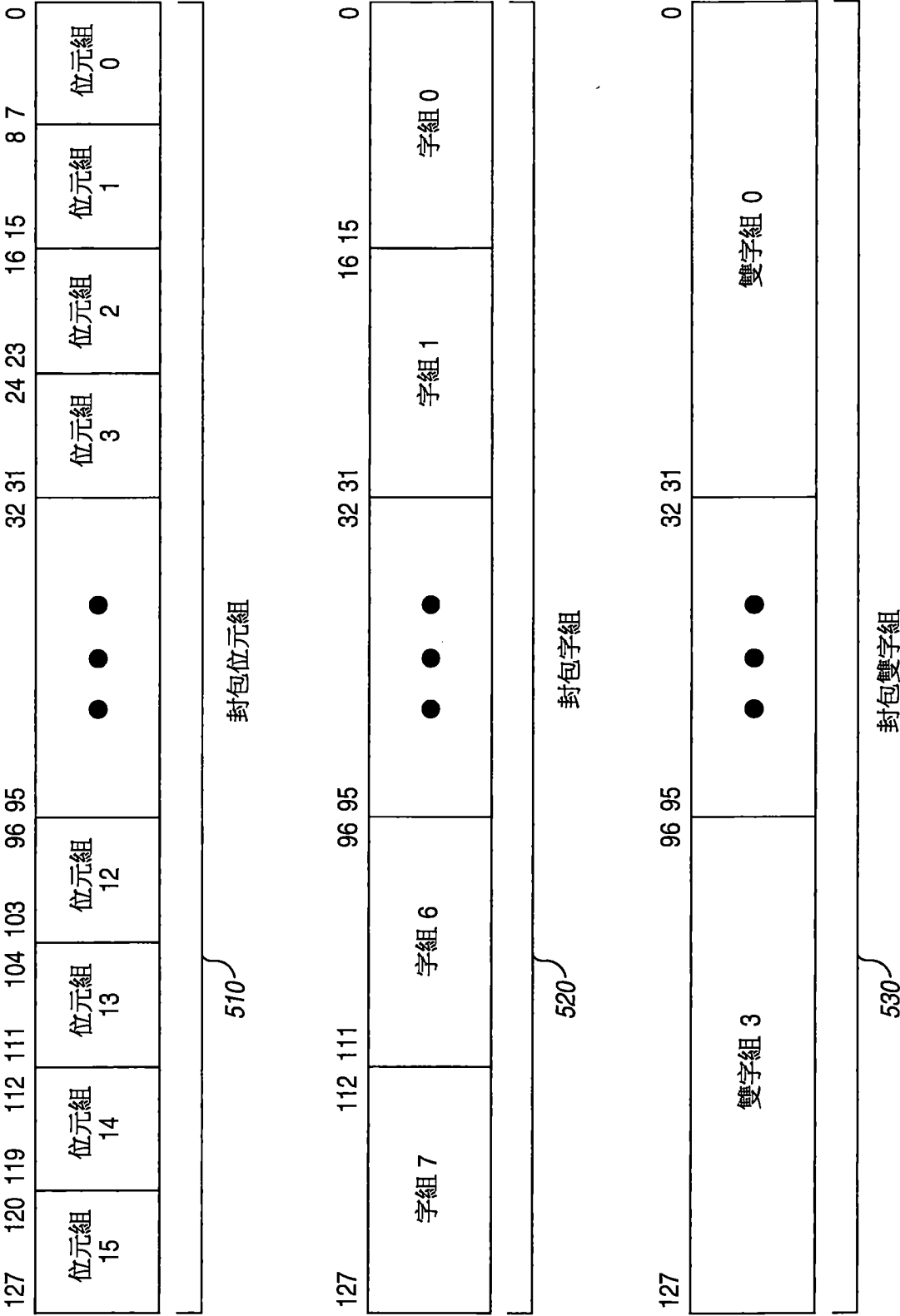


圖 5

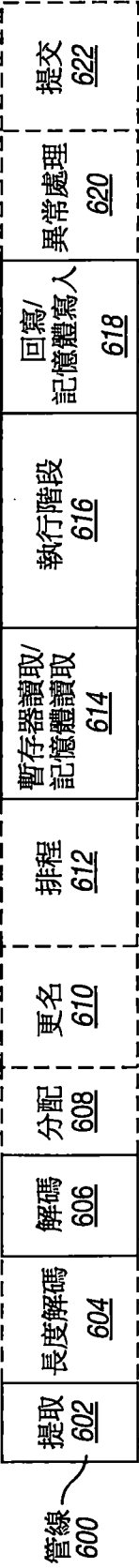


圖 6A

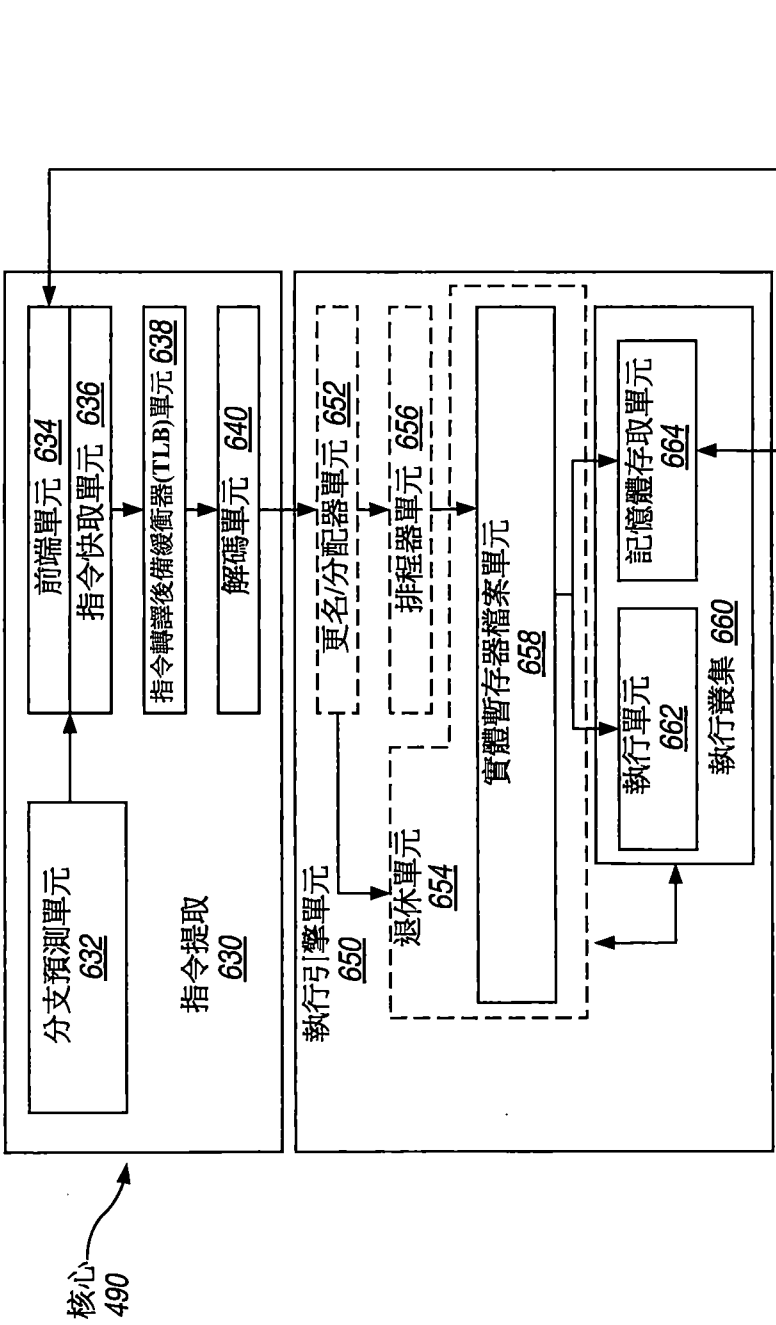


圖 6B

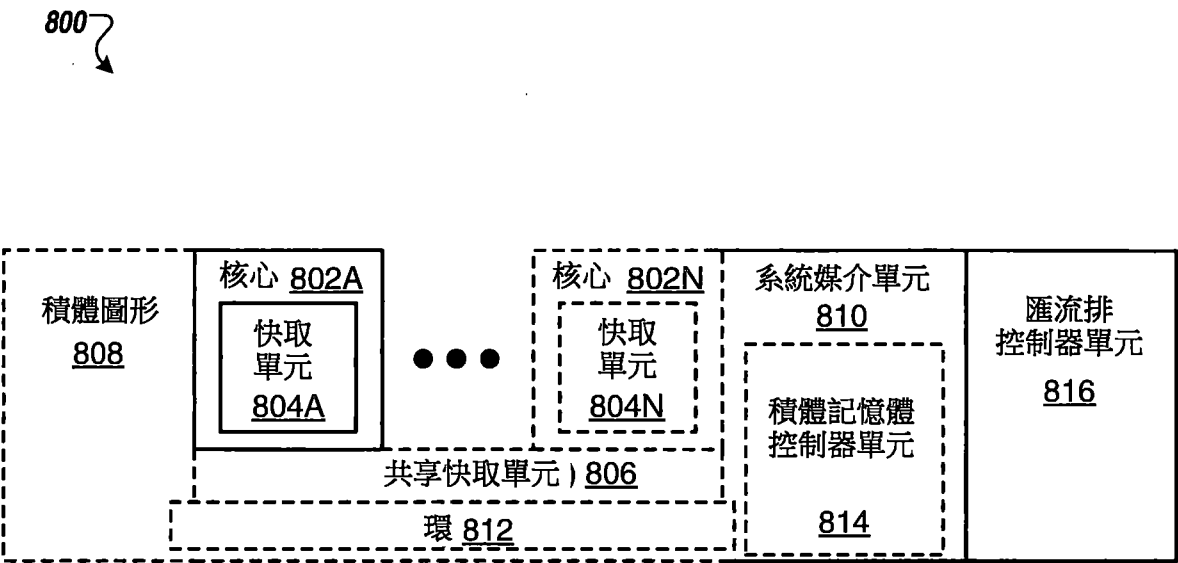


圖 8

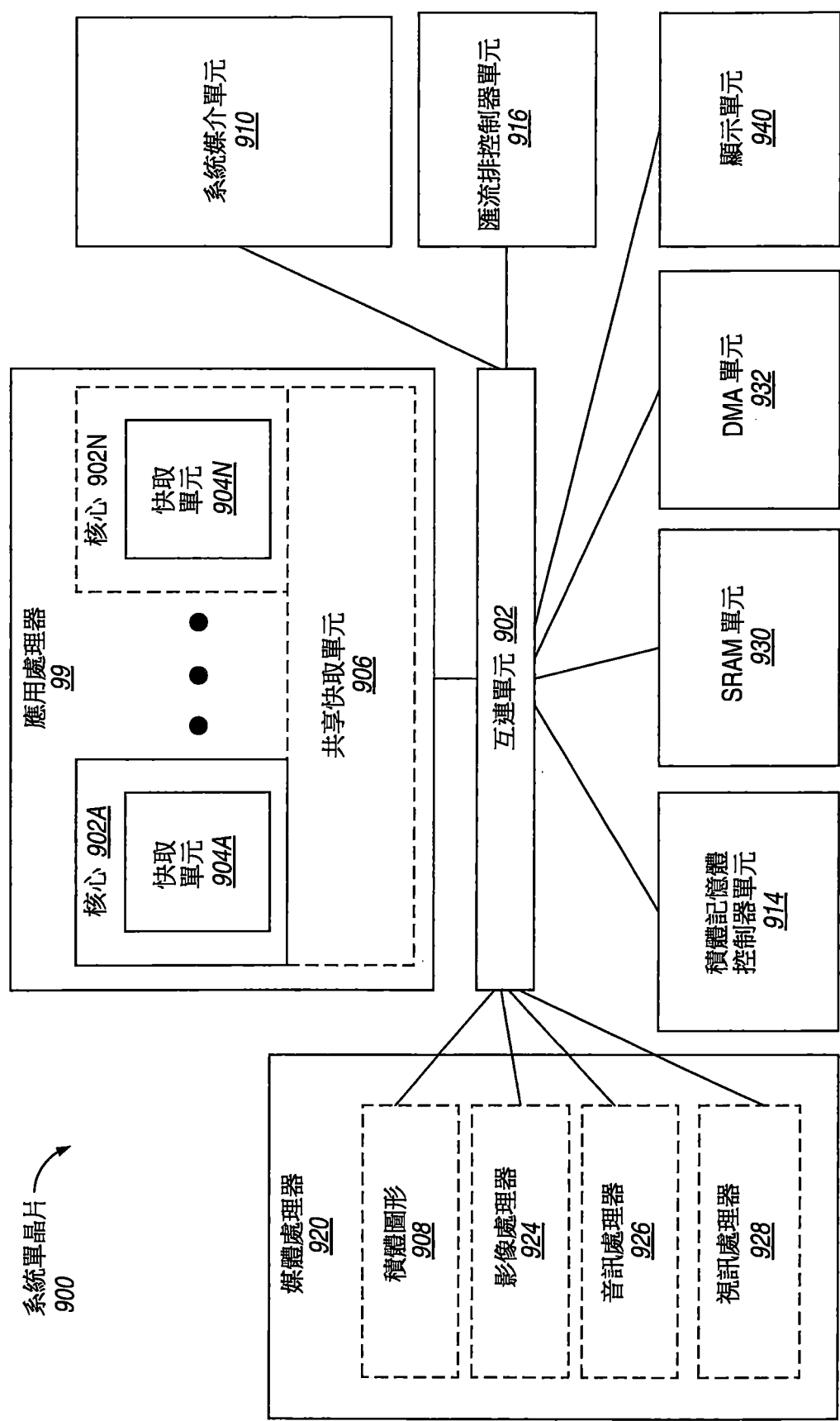


圖 9

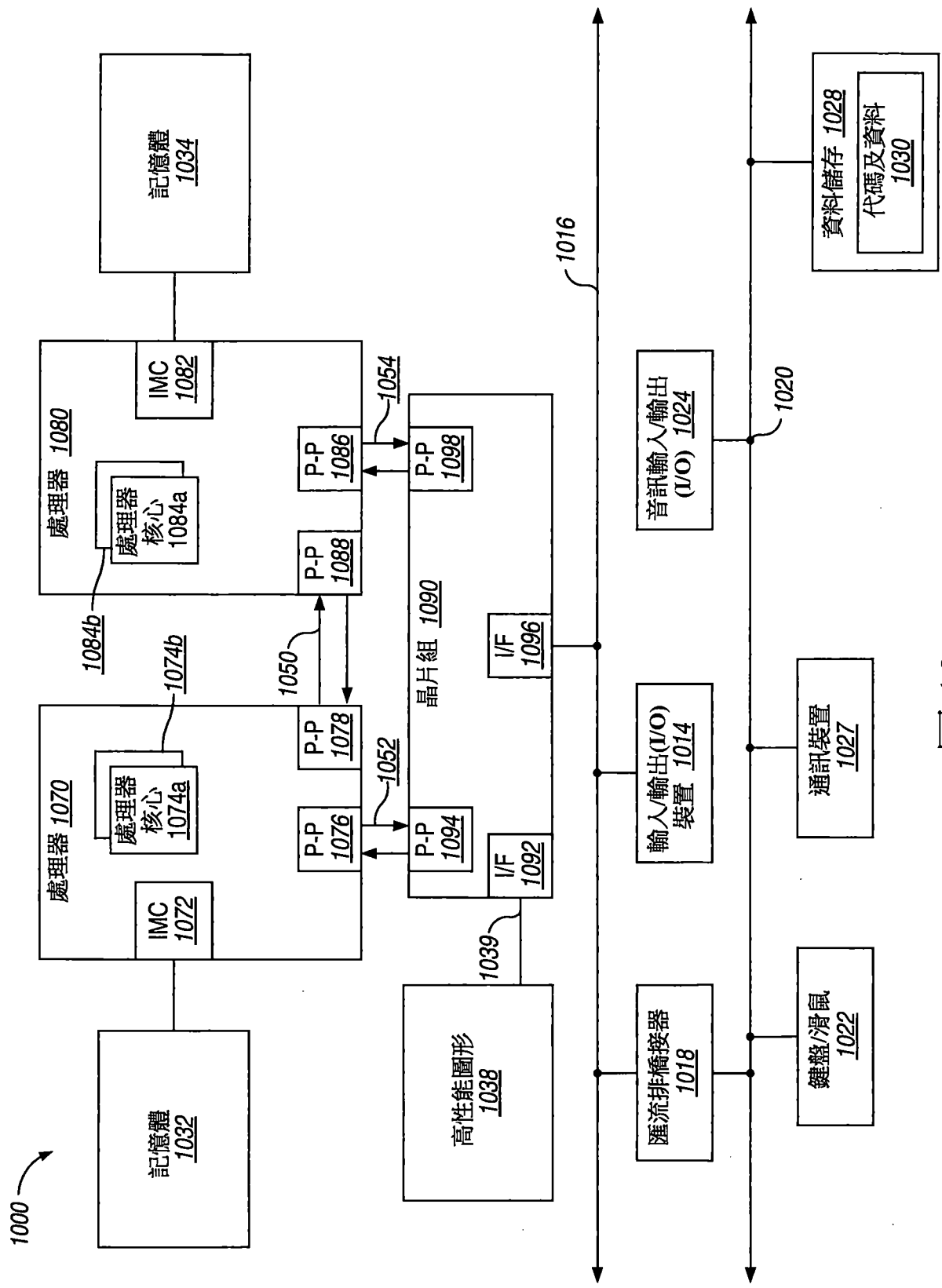


圖 10

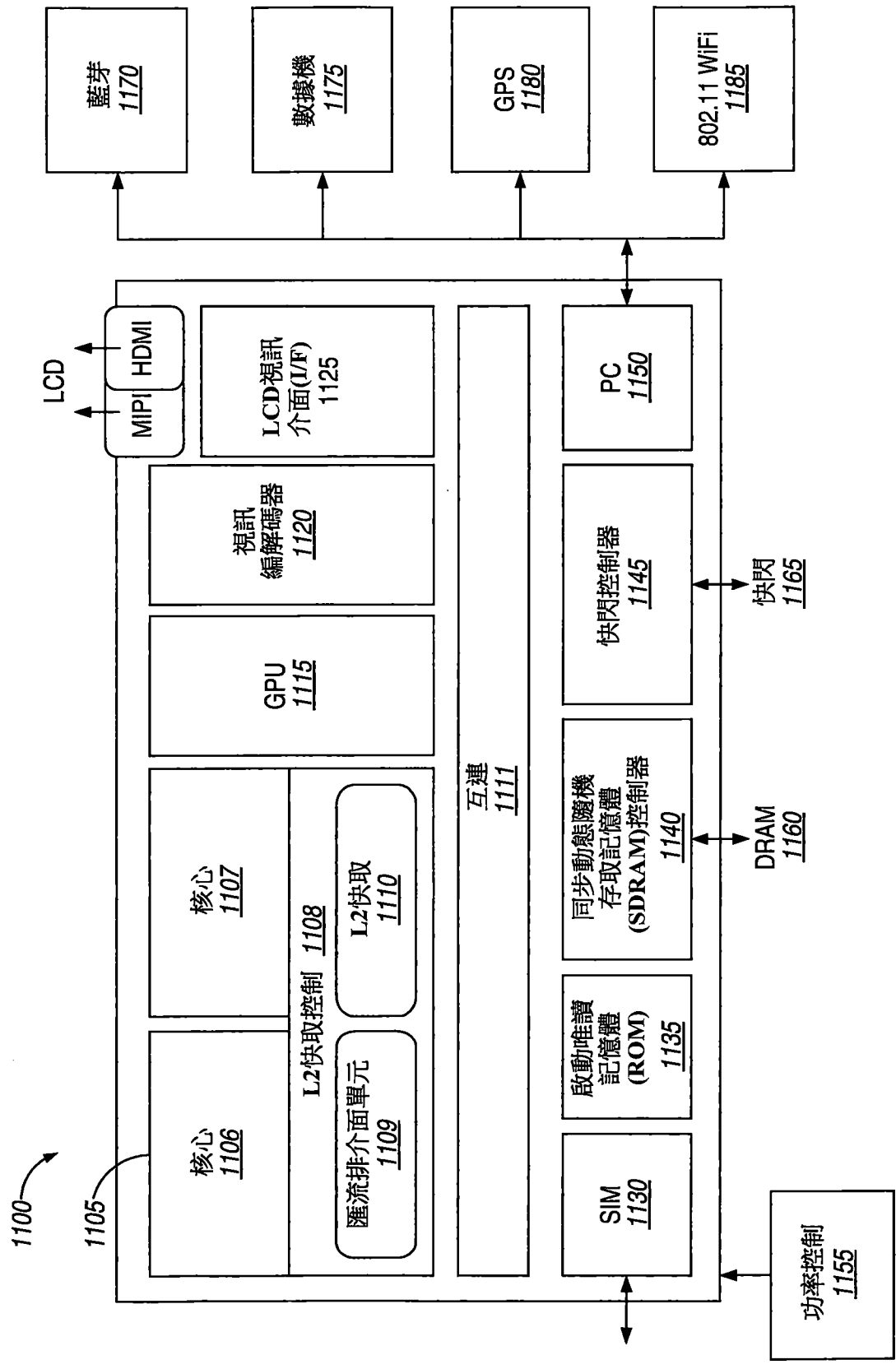


圖 11

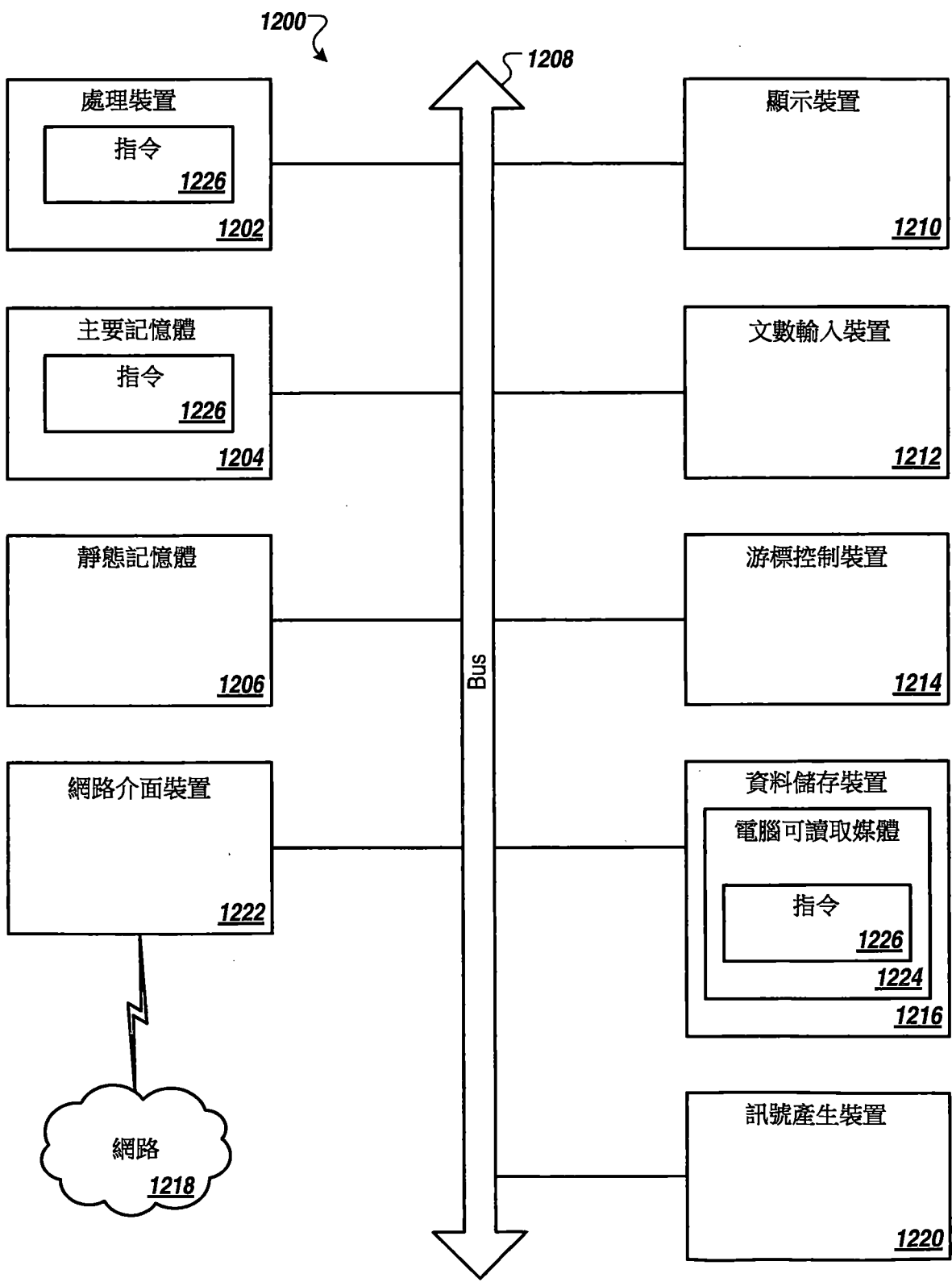


圖 12