



(19)中華民國智慧財產局

(12)發明說明書公告本

(11)證書號數：TW I536269 B

(45)公告日：中華民國 105 (2016) 年 06 月 01 日

(21)申請案號：101113633

(22)申請日：中華民國 101 (2012) 年 04 月 17 日

(51)Int. Cl. : G06F9/50 (2006.01)

G06F9/28 (2006.01)

(30)優先權：2011/06/08 美國

13/155,387

(71)申請人：微軟技術授權有限責任公司 (美國) MICROSOFT TECHNOLOGY LICENSING, LLC
(US)

美國

(72)發明人：奧斯海姆傑瑞德 AASHEIM, JERED (US)

(74)代理人：蔡坤財；李世章

(56)參考文獻：

TW I262443

TW 200517963A

TW 200820082A

US 7007183B2

US 7539841B2

審查人員：劉季涵

申請專利範圍項數：20 項 圖式數：4 共 34 頁

(54)名稱

去耦合異質計算之作業系統

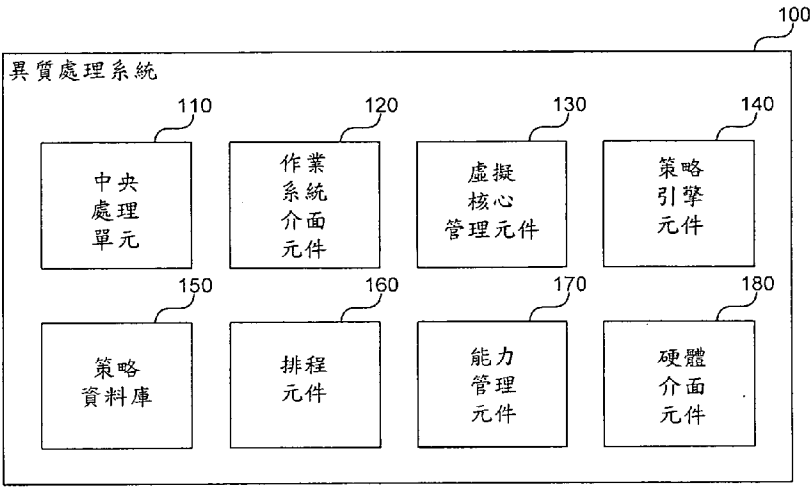
OPERATING SYSTEM DECOUPLED HETEROGENEOUS COMPUTING

(57)摘要

本文描述一種異質處理系統，該異質處理系統提供在無作業系統的知曉或涉入的情況下自主控制跨越大核心與小核心之作業系統線程排程之軟體超管理器，以改良能源效率或達成其他處理目標。系統向作業系統提供有限集合之虛擬化計算核心，系統向作業系統排程用於執行的線程。隨後，超管理器智慧控制實體指派及選擇哪個(些)核心執行每一線程以管理能源使用或其他處理要求。藉由使用軟體超管理器以抽取基礎大小電腦架構，核心之間之效能與功率作業差異對作業系統保持不透明狀態。固有間接性亦使具有新能力之硬體之發佈與作業系統發佈排程去耦合。

A heterogeneous processing system is described herein that provides a software hypervisor to autonomously control operating system thread scheduling across big and little cores without the operating system's awareness or involvement to improve energy efficiency or meet other processing goals. The system presents a finite set of virtualized compute cores to the operating system to which the system schedules threads for execution. Subsequently, the hypervisor intelligently controls the physical assignment and selection of which core(s) execute each thread to manage energy use or other processing requirements. By using a software hypervisor to abstract the underlying big and little computer architecture, the performance and power operating differences between the cores remain opaque to the operating system. The inherent indirection also decouples the release of hardware with new capabilities from the operating system release schedule.

指定代表圖：



第1圖

- 符號簡單說明：
- 110 . . . 中央處理單元
 - 120 . . . 作業系統介面元件
 - 130 . . . 虛擬核心管理元件
 - 140 . . . 策略引擎元件
 - 150 . . . 策略資料庫
 - 160 . . . 排程元件
 - 170 . . . 能力管理元件
 - 180 . . . 硬體介面元件

發明專利說明書

(本說明書格式、順序，請勿任意更動，※記號部分請勿填寫；惟已有申請案號者請填寫)

※ 申請案號：101113633

※ 申請日期：101 年 4 月 17 日

※IPC 分類：

G06F9/50 (2006.01)

G06F9/28 (2006.01)

一、發明名稱：(中文/英文)

去耦合異質計算之作業系統/OPERATING SYSTEM DECOUPLED
HETEROGENEOUS COMPUTING

二、中文發明摘要：

本文描述一種異質處理系統，該異質處理系統提供在無作業系統的知曉或涉入的情況下自主控制跨越大核心與小核心之作業系統線程排程之軟體超管理器，以改良能源效率或達成其他處理目標。系統向作業系統提供有限集合之虛擬化計算核心，系統向作業系統排程用於執行的線程。隨後，超管理器智慧控制實體指派及選擇哪個(些)核心執行每一線程以管理能源使用或其他處理要求。藉由使用軟體超管理器以抽取基礎大小電腦架構，核心之間之效能與功率作業差異對作業系統保持不透明狀態。固有間接性亦使具有新能力之硬體之發佈與作業系統發佈排程去耦合。

三、英文發明摘要：

A heterogeneous processing system is described herein that provides a software hypervisor to autonomously control operating system thread scheduling across big and little cores without the operating system's awareness or involvement to improve energy efficiency or meet other processing goals. The system presents a finite set of virtualized compute cores to the operating

system to which the system schedules threads for execution. Subsequently, the hypervisor intelligently controls the physical assignment and selection of which core(s) execute each thread to manage energy use or other processing requirements. By using a software hypervisor to abstract the underlying big and little computer architecture, the performance and power operating differences between the cores remain opaque to the operating system. The inherent indirection also decouples the release of hardware with new capabilities from the operating system release schedule.

四、指定代表圖：

(一)本案指定代表圖為：第（ 1 ）圖。

(二)本代表圖之元件符號簡單說明：

- 110 中央處理單元
- 120 作業系統介面元件
- 130 虛擬核心管理元件
- 140 策略引擎元件
- 150 策略資料庫
- 160 排程元件
- 170 能力管理元件
- 180 硬體介面元件

五、本案若有化學式時，請揭示最能顯示發明特徵的化學式：

無

六、發明說明：

【發明所屬之技術領域】

本發明係關於去耦合異質計算之作業系統。

【先前技術】

能源效率日益成為從行動電話到資料處理中心之重要區分者。消費者樂意為可持續較久的移動裝置之使用感受付費，且亦渴望從此等相同裝置得到增加之效能。另一方面，資料處理中心繼續擴大計算能力但面臨可有效冷卻之熱限制。此外，公眾日益意識到能源使用及能源使用造成的環境影響。因此在多種計算系統中有效利用能源係主要優先設計目標。

此等技術上對立之目的（使用更少功率而達到更多效能）引起了產業實驗，該實驗具有把「大」計算核心與「小」計算核心在單個系統或矽晶片中緊密耦合的異質設計，在本文中稱為異質核心或異質處理。大核心設計為在較大功率包絡（power envelope）中提供高效能，同時小核心設計為在較小功率包絡中提供較低效能。傳統觀點認為作業系統之排程程式將隨後視工作負載而有選擇地在大核心或小核心上排程線程。在一天中的至少某些時候，作業系統可能能夠完全關掉大核心且依靠功率吸取之小核心。

大核心及小核心可能或可能不共享相同指令組或特徵結構。舉例而言，小核心可能包括較少指令組或其他差

異，該等差異涉及作業系統所做之進一步決策以在可相容核心上排程程序。一個傳統實例為包括中央處理單元（CPU）及圖形處理單元（GPU）之系統，且該系統在 GPU 閑置或利用不足時允許 GPU 用於計算任務。

現存之解決方法取決於修改作業系統之內核以「啓發」作業系統大小核心之存在、大小核心各別之效能及功率特徵，及作業系統可監視該系統中之何設備（例如，CPU 效能計數器、快取未命中或命中計數器、匯流排活動計數器，等等）以決定在何（等）核心上排程特定線程。此方法具有若干缺點：1）此方法涉及修改所有受支援之作業系統之內核；2）此方法要求已修改內核瞭解跨越潛在不同架構（例如，支援多個不同實施）之大或小設計之差異；及 3）此方法緊密耦合作業系統內核之發佈排程與基礎電腦架構。電腦架構之變化則涉及在核心可支援新市售核心之前等待下一個已排程作業系統發佈（亦即，大概若干年或更久）（或反之亦然）。

【發明內容】

本文描述一種異質處理系統，該異質處理系統提供在無作業系統的知曉或涉入的情況下自主控制跨越大核心與小核心之作業系統線程排程之軟體超管理器，以改良能源效率或達成其他處理目標。該系統向作業系統提供有限集合之虛擬化計算核心，該系統向作業系統排程用於執行的線程。隨後，在表面之下，超管理器智慧控制

實體指派及選擇哪個（些）核心執行每一線程以管理能源使用或其他處理要求。藉由使用軟體超管理器以抽取基礎大小電腦架構，核心之間之效能與功率作業差異對作業系統保持不透明狀態。固有間接性亦將具有新能力之硬體之發佈與作業系統發佈排程去耦合。硬體供應商可發佈已更新之超管理器，且允許新硬體與供應商選擇之任何版本之作業系統一起工作。

本文提供此發明內容以用簡化形式介紹在下文實施方式中進一步描述之概念選擇。本發明內容不欲識別所主張標的之關鍵特徵或基本特徵，亦不欲用以限制所主張標的之範疇。

【實施方式】

本文描述一種異質處理系統，該異質處理系統提供在無作業系統的知曉或涉入的情況下自主控制跨越大核心與小核心之作業系統線程排程之軟體超管理器以改良能源效率或達成其他處理目標。該系統向作業系統提供有限集合之虛擬化計算核心，該系統向作業系統排程用於執行的線程。隨後，在表面之下，超管理器智慧控制實體指派及選擇哪個（些）核心執行每一線程以管理能源使用或其他處理要求。藉由使用軟體超管理器以抽取基礎大小電腦架構、核心之間之效能與功率作業差異對作業系統保持不透明狀態。固有間接性亦使具有新能力之硬體之發佈與作業系統發佈排程去耦合。硬體供應商可

發佈已更新之超管理器，且允許新硬體與供應商選擇之任何版本之作業系統一起工作。

超管理器實施緊密耦合至基礎電腦架構及使用可用之系統反饋（例如，中央處理單元利用率、匯流排或快取活動，諸如此類）以自主分配適當核心用於所請求之工作負載。此方法允許基礎電腦架構頻繁改變以與軟體超管理器合作且從上述作業系統中去耦合此演變。異質處理系統提供簡單、過程消除功率管理而不用修改作業系統內核本身。因此，異質處理系統允許更多快速硬體革新，且允許現有資料處理中心及其他設施當下得益於可用異質處理硬體。

異質計算係行業中之新興領域，異質計算系統之目標為根據在該系統中可用之不同類型之計算核心（例如，CPU、GPU、加速器，等等）最佳化工作負載之執行。最佳化可針對效能、功率、延遲或其他目標。當異質處理系統可適用於此等更普遍之情況時，異質處理系統在具有相同功能等效但不同效能或功率作業特性之核心之系統上亦可達成。通常，此等系統具有一或更多個大核心及一或更多個小核心。大核心通常具有深管線、無序執行、大快取記憶體、高時鐘速率，且使用較高洩漏製程（例如，40G）製造。小核心通常具有較短管線、較小快取記憶體、較低時鐘速率、各種功率水平，且使用低洩漏製程（例如，40LP）製造。

在某些實施例中，大小核心可能具有架構等效、微架

構等效、全域中斷控制器、同調及虛擬化。架構等效可能包括相同指令組架構 (ISA)、單指令多資料 (SIMD)、浮點 (FP)、協處理器有效性及 ISA 之擴充部分。微架構等效可能包括不同效能但相同可配置特徵 (例如, 快取記憶體線長度)。全域中斷控制器提供管理、處理及向所有核心轉發中斷之能力。同調意味著所有核心可按需要用轉發從其他核心存取 (快取) 資料。虛擬化係用以從或向核心切換或遷移工作負載。

在某些實施例中, 異質處理系統也許能夠處理核心中之次要差異。舉例而言, 不支援串流單指令、多資料 (SIMD) 擴充 (SSE) (現在存在於四迭代中: SSE1、SSE2、SSE3 及 SSE4) 之小核心仍可處理其他基於英特爾 x86 之軟體代碼。超管理器可偵測在指令串流中不受支援之指令, 及激活適當核心以向該核心分配此指令串流。其他指令串流可在任何核心上可靠作業。在某些諸如使用僅少數不受支援之指令之情況下, 超管理器可包括某種模擬水平以在可用之指令組上模擬不受支援之指令。舉例而言, 諸如矢量數學之操作可常常使用標準數學指令以較低效率分解及實施。

軟體超管理器在作業系統 (OS) 初始化之前在裝置啟動過程中安裝超管理器本身。在完成特定硬體配置 (亦即, 配置記憶體、初始化虛擬化設備, 等等) 以後, 超管理器則經由策略配置安裝在計算裝置中之大小處理核心。舉例而言, 若裝置係行動電話, 則策略可指示超管

理器以最小量可用效能啟動作業系統且最佳化電池壽命；超管理器將隨後排程作業系統線程至一或更多個小核心。或者，若裝置係資料處理中心布列茲（blade），則策略可指示以最大量可用效能啟動超管理器且犧牲能源效率；超管理器將隨後排程作業系統線程至可用之大核心以及可能還視可用熱預算排程至小核心。在完成初始化以後，軟體超管理器載入作業系統啟動管理器，作業系統啟動管理器繼而載入作業系統。

在執行中，異質處理系統向作業系統提供虛擬化核心集合。作業特性及核心之間之差異對作業系統係不透明的，且藉由軟體超管理器根據已定義之作業策略秘密地管理。作業策略可在系統初始化期間設定或在執行期間動態設定。

超管理器使用作業策略連同可用之系統設備（例如，CPU 效能計數器、快取未命中或命中計數器、匯流排活動計數器，等等）以決定排程作業系統線程至哪些核心。超管理器將使用此資訊以瞭解 CPU 核心之利用、一段時間內之趨勢、位置資訊及輸入或輸出（I/O）模式。由此資訊，超管理器可適當動態及可預測地遷移作業系統線程跨越大小核心。另外，超管理器亦可以根據該系統實施代表作業系統控制動態頻率及電壓調節（DFVS）。

此為超管理器可控制之可用作業策略之示例：最小功率（MiPo）、最大效能（MaPe）、最小功率、按需效能（MiPoD）及最大效能、閑置時斷電（MaPeI）。此等示

例之每一個在下文段落中描述。此外，儘管如此，按任何特定實施之選擇可實施更高級作業策略。

最小功率（MiPo）排程線程至最小核心集合。此舉通常將意味著超管理器排程線程至小核心，且當需要時使用 DVFS 以爲核心控制功率及效能作業點。當需要時可供電及排程額外小核心。

最大效能（MaPe）排程線程至最大核心集合。此舉通常將意味著超管理器從大核心開始排程線程至全部可用之核心及當需要時使用 DVFS 以爲核心控制功率及效能工作點。只要可用熱預算允許，亦可供電和排程小核心。

最小功率、按需效能（MiPoD）一般以最低可用功率狀態（例如，在一或更多個小核心上）作業，但按工作負載需求提升效能。此舉通常被認爲是「渦輪」或「提升」作業模式，且藉由動態分派及排程至大核心實現。一旦完成工作負載，系統回到最小功率狀態（例如，在小核心上）。

最大效能、在閑置時斷電（MaPeI）一般以最大可用之效能狀態（例如，在一或更多個大核心上）作業，但一旦達到閑置閾值時默許至較低功率狀態。在此情況下該閑置閾值不是典型的接近零之 CPU 利用率，但可任意地定義在某種每秒整數運算百萬次指令（DMIPS）或由策略所定義之 CPU 利用率百分比。當將要閑置時，超管理器動態分派及排程至小核心，且使不用的大核心進入備用或休眠狀態。策略及/或將來工作負載決定系統何時

回到最大可用效能狀態（例如，在大核心上）。

第 1 圖係方塊圖，該圖圖示了在一個實施例中異質處理系統之元件。系統 100 包括一或更多個中央處理單元 110、作業系統介面元件 120、虛擬核心管理元件 130、策略引擎元件 140、策略資料庫 150、排程元件 160、能力管理元件 170 及硬體介面元件 180。此等元件之每一個在本文中更詳細地描述。以下元件可能在軟體超管理器中實施，軟體超管理器位於作業系統及計算裝置之硬體資源之間。

一或更多個中央處理單元 110 包括一或更多個處理核心，該等處理核心具有異質處理能力及功率輪廓。通常，每一個 CPU 複合物位於單個矽芯晶粒上，且 CPU 複合物之每一個核心共享矽芯晶粒。硬體可在各種類型裝置之各種套件中實施。舉例而言，更新的行動裝置及甚至某些新近桌面處理器包括在相同晶片上之 CPU 及 GPU 用以在兩個且較低的功率使用之間有效通訊。每一 CPU 複合物可包括一或更多個大小核心。或者或另外，一個 CPU 複合物可包括所有大核心同時另一 CPU 複合物包括所有小核心。此處使用之 CPU 複合物應用於 GPU 及其他可執行軟體指令之硬體。

作業系統介面元件 120 在超管理器及作業系統之間通訊以接收用以傳遞至硬體資源及接收來自硬體資源之輸出之指令。作業系統可排程線程、提供指標至指令串流（例如，程式計數器（PC））、寫入傳送指令至硬體之記

憶體區域，諸如此類。作業系統通常與計算裝置上之硬體直接互動。然而，超管理器出於各種目的在作業系統及硬體之間插入間接層。超管理器往往用以提供虛擬化以便多個作業系統可同時在同一硬體上執行。超管理器亦可用以向作業系統提供不同於安裝在計算裝置上之實際硬體之虛擬硬體。在異質處理系統 100 的情況下，此舉可包括使大小核心對作業系統看起來相同。系統 100 甚至可以向作業系統提供不同於實際存在於裝置中之數目的核心。

虛擬核心管理元件 130 管理一或更多個虛擬核心，超管理器將該等虛擬核心提供給作業系統。虛擬核心作為 CPU 核心出現至作業系統，但可能在特徵方面與在計算裝置中之可用實體硬體不同。舉例而言，虛擬核心向作業系統隱瞞在處理或功率能力上之差異，以便不設計為與異質大小核心工作之作業系統可以為作業系統所設計之方式作業。在此情況下，超管理器提供任何所需之專用程式以利用異質計算環境，以便不需修改作業系統。

策略引擎元件 140 管理一或更多個策略用以根據可用之一或更多個中央處理單元排程作業系統線程及向作業系統提供虛擬核心。策略引擎元件 140 可包括特定於某個超管理器實施之硬編碼策略或可包括管理員可配置策略，該等管理員可配置策略可經修改以適合特定安裝目標。策略可決定首先排程哪些核心、權衡功率使用與目標處理、如何關閉及喚起核心以節省功率、虛擬核心如

何向作業系統提供，諸如此類。

策略資料庫 150 儲存一或更多個策略於可在啓動及執行時由超管理器存取之儲存設備。策略資料庫 150 可包括一或更多個檔案、檔案系統、硬碟、資料庫，或用以跨越系統 100 之執行會話存留資料的其他儲存設備。在某些實施例中，管理員執行設置步驟，該設置步驟使系統 100 經由配置階段以儲存策略之初始設置供超管理器使用。

排程元件 160 從作業系統排程接受為線程之一或更多個指令串流至安裝在計算裝置中中央處理單元中之一或更多者。排程元件從作業系統接收虛擬核心識別符，該虛擬核心識別符識別虛擬核心，作業系統請求排程線程至該虛擬核心。排程元件 160 檢查排程請求且決定實體核心，在該實體核心上排程執行該等線程。舉例而言，元件 160 可決定功率還是處理對線程更相關，及排程適當小或大核心作為回應。在某些情況下，元件 160 可避免排程線程至某些核心以允許彼等核心斷電以節省功率。

能力管理元件 170 視情況管理大小處理核心之間之一或更多個差異。在某些情況下，系統 100 可能僅僅作業在大小核心共享相同能力之處理單元上，且不需要容量管理元件 170。在其他情況下，系統 100 處理可用處理核心之間之次要或主要差異。舉例而言，系統 100 可能等待不受某些核心支援之指令及將相應線程排程到支援

彼等指令之核心上。在更複雜實施中，元件 170 可能在小核心上虛擬化或模擬大核心能力（或反之亦然）以滿足功率目標或其他輪廓目標。

硬體介面元件 180 在超管理器與中央處理單元之間通訊以排程軟體指令以在可用之實體核心上執行。硬體介面元件 180 可包括用以存取真實硬體之真實記憶體位址或其他設備，該等真實硬體或其他設備對其他元件尤其向藉由超管理器管理之客作業系統隱瞞。

實施異質處理系統之計算裝置可包括中央處理單元、記憶體、輸入裝置（例如，鍵盤及指向裝置）、輸出裝置（例如，顯示裝置）及儲存裝置（例如，磁碟或其他非揮發性儲存媒體）。記憶體及儲存裝置係電腦可讀取儲存媒體，該等儲存媒體可用實施或實現系統的電腦可執行指令（例如，軟體）編碼。此外，資料結構及訊息結構可經由諸如在通訊鏈路上之訊號之資料傳輸媒體儲存或傳輸。可使用各種通訊鏈路，諸如網際網路、區域網路、廣域網路、點對點撥號連接、蜂巢式電話網路，等等。

系統之實施例可在各種作業環境中實施，該等作業環境包括個人電腦、伺服器電腦、手持或膝上型裝置、多處理器系統、基於微處理器之系統、可程式化消費性電子設備、數位攝像機、網路 PC、小型電腦、大型電腦、包括任何上述系統或裝置之分佈式計算環境、視訊轉換器、晶片上系統（SOCs），等等。電腦系統可為蜂巢式電話、個人數位助理、智慧電話、個人電腦、可程式化

消費性電子設備、數位攝像機，等等。

系統可在由一或更多個電腦或其他裝置執行之諸如程式模組之電腦可執行指令的大體上下文中描述。大體而言，程式模組包括常式、程式、物件、元件、資料結構，等等，該等程式模組執行特定任務或實施特定抽象資料類型。通常，程式模組之功能性可按需在各種實施例中組合或分配。

第 2 圖係流程圖，該圖圖示了在一個實施例中處理異質處理系統在異質處理核心及作業系統之間使用超管理器初始化具有異質處理核心之計算裝置。

從方塊 210 開始，系統接收啟動要求以初始化計算裝置。舉例而言，基本輸入或輸出系統 (BIOS)、可擴充韌體介面 (EFI)、啟動裝入程式或其他初始裝置軟體可載入且調用實施異質計算系統之超管理器。在某些情況下，管理員先前已執行安裝階段以在計算裝置上安裝超管理器，儘管系統亦可支援網路啟動及通常提供給計算裝置之其他非安裝方案。

在方塊 220 中繼續，系統枚舉計算裝置之兩個或兩個以上實體處理核心。在某些實施例中，至少兩個核心提供不同效能及功率使用特徵。然而，系統亦可在不存在不對稱的情況下使用。舉例而言，使用用於功率管理的軟體超管理器仍然可適用在具有晶粒上之 N 個實體 CPU 但僅 K 個可作業之情況中，該作業基於諸如周圍溫度、外殼形狀因數、可用功率之成本等等外部因素。在啟動

時，超管理器可使用此「策略」資訊以向作業系統報告 K 個核心之虛擬化集合，且此舉可隨每一個啟動循環變化。超管理器將在對稱核心之情況下執行相同任務。系統可調用 BIOS 或其他基礎層以決定計算裝置安裝了多少及何種處理器，且可執行 CPUID 或其他類似指令以決定有關處理器之處理能力之資訊。在某些實施例中，系統可包括可擴充性介面，超管理器製造商或第三方可經由該可擴充性介面實施或添加驅動程式或其他超管理器擴充，以對該超管理器添加新處理硬體支援而不用必須更新該超管理器本身。

在方塊 230 中繼續，系統決定每一個已枚舉處理核心之能力。該等能力可包括藉由每一個核心提供之一或更多個功率輪廓、藉由每一個核心支援之一或更多個指令組、每一個核心之效能特徵，諸如此類。該系統可利用核心本身之資訊介面（諸如先前提到之 CPUID 指令）或由驅動程式或其他擴充對超管理器所提供之資訊，以決定每一個核心之能力。系統使用已決定能力以分配與核心相容之線程至每個核心，且以符合已接收之策略及處理目標之方式執行排程。

在方塊 240 中繼續，系統識別一或更多個作業系統，超管理器將為該等作業系統管理已枚舉之實體核心的存取及排程。系統可存取硬碟、快閃記憶體，或計算裝置之其他儲存裝置以決定在初始化超管理器以後調用哪個作業系統。超管理器可設計為具有關於各種作業系統之

資訊，且可包括可擴充性以便不用更新超管理器本身就可支援新作業系統。每一個作業系統及作業系統版本可具有不同排程語義或其他細微差別，超管理器處理該等各種排程語義或其他細微差別以允許作業系統在虛擬化處理來源上正確執行。在某些情況下，超管理器可受要求以允許多個作業系統共享已枚舉之實體處理核心，且策略可指示如何處理該共享。

在方塊 250 中繼續，系統存取規定一或更多個目標之超管理器策略資訊，該等目標係為了在已枚舉之實體處理核心上排程作業系統線程。目標可包括效能目標、功率使用目標或其他方向以決定在哪個或哪些核心上執行作業系統線程。策略可能儲存在與計算裝置關聯之儲存裝置中、硬編碼至超管理器實施中，諸如此類。超管理器可經由向管理員提供之管理介面接收策略的更新。

在方塊 260 中繼續，系統建立一或更多個虛擬核心以暴露於已識別之作業系統，其中每一個虛擬核心使作業系統與實體處理核心間之已決定之能力差異分離。舉例而言，異質處理系統可將兩個或兩個以上大小核心作為單一類型之統一虛擬核心向作業系統提供。在從作業系統接收排程請求以在虛擬核心上執行線程後，系統根據已存取之超管理器策略決定選擇哪個實體核心工作。超管理器策略可規定超管理器提供與實體核心數目不同之虛擬核心數目，諸如當以使至少某些較高功率需求之核心無間隙地斷電以使用較低功率需求之核心為目標時。

或者，系統仍然可使作業系統察覺之核心斷電但在作業系統選擇使用該等核心或使用若干不能獨自由較低功率核心滿足之核心時激活該等核心。

在方塊 270 中繼續，系統調用已識別之作業系統且向作業系統提供經建立之虛擬核心，同時使已識別之作業系統與已枚舉之實體處理核心分離。調用作業系統可包括調用作業系統裝入程式及提供超管理器以代替通常的 BIOS 或在作業系統之下之其他層。作業系統仿佛直接在實體硬體上執行一般作業，但超管理器位於作業系統與實體硬體之間以執行排程邏輯，而如本文所描述作業系統並不知曉。在方塊 270 以後，此等步驟結束。

第 3 圖係流程圖，該圖圖示了在一個實施例中處理異質處理系統以經由管理異質處理核心之超管理器排程一或更多個作業系統線程。

從方塊 310 開始，系統從作業系統接收線程排程請求以在由超管理器提供之已識別之虛擬核心上執行線程指令，其中虛擬核心使作業系統與在計算裝置上可存取之兩個或兩個以上實體處理核心之間之一或更多個能力差異分離。作業系統通常包括用於每一個已偵測之處理核心之閒置迴路，在該閒置迴路中，作業系統可排程及置放作業系統想要在該核心上執行之任何指令。作業系統可時間切片多個應用程式線程以在特定處理核心上執行。無論作業系統選擇以執行線程之特定虛擬核心為何，超管理器皆可依照一或更多個超管理器策略選擇任

何特定實體處理核心以執行線程。

在方塊 320 中繼續，系統決定已接收之排程請求之處理需要。舉例而言，系統可決定藉由已排程之線程所使用之特定指令組（例如，是否請求一或更多個指令組擴充、協處理器，或其他能力）、線程之效能要求、線程是否適於較低功率使用之較慢執行、線程是否可推遲直到額外處理來源可用，諸如此類。系統可使用關於特定作業系統或經由策略接收之指令之特定認識以決定對特定線程之處理需要。舉例而言，系統可識別與作業系統之內部作業、應用程式線程等等相關之線程，且根據策略處理每個線程。

在方塊 330 中繼續，系統存取規定用以操作裝置之一或更多個目標之排程策略。舉例而言，策略可請求對功率使用、效能或兩者之混合之最佳化。策略可儲存在與裝置關聯之資料庫中或硬編碼至超管理器之特定實施中。舉例而言，系統可提供超管理器之低功率使用版本，該低功率使用版本對小處理核心有利直到作業系統排程執行高效能任務的線程。在那時，系統可在大核心上排程高效能任務，然後指令大核心在任務完成以後休眠。

在方塊 340 中繼續，該系統選擇實體處理核心，在該核心上執行與已接收之排程請求關聯之線程，其中根據已存取之排程策略作出該選擇。系統可具有多個不同能力及效能或功率特徵之可用核心，系統可在該等核心之上排程線程。根據系統對核心之選擇，計算裝置將使用

或多或少的功率，且將在或多或少時間內結束線程之執行。工作或排程策略將允許該系統以促進用以管理計算裝置之效能、功率或其他特徵之一或更多個目標的方式做出選擇。行動裝置可優選較低功率使用，而高效能伺服器可能優選較高效能。在某些情況下，策略可能根據當日時間（例如，高峰對非高峰之電耗成本）或其他考慮而不同，以便策略一段時間內或根據某些條件變化。

在方塊 350 中繼續，系統視情況處理線程與所選擇之實體處理核心之間的任何能力差異。舉例而言，若線程包括在所選核心上不可用之指令，則系統可模擬該指令或以所選核心支援之一或更多個等效指令替換該指令。管理能力差異使系統之複雜性顯著增加，且超管理器實施者可選擇任何特定實施將支援之處理核心之間的能力差異程度（若存在差異）。

在方塊 360 中繼續，系統排程線程以在所選擇之實體處理核心上執行線程。系統亦處理任何輸出，且將輸出返回給作業系統，使得輸出似乎來自虛擬核心，作業系統分配線程至虛擬核心。因此，作業系統始終不察覺由超管理器管理之核心類型及數目，且如同作業系統在沒有超管理器及異質處理核心之系統中正常使用一樣使用虛擬核心集合。在方塊 360 以後，此等步驟結束。

第 4 圖係方塊圖，該圖圖示了在一個實施例中異質處理系統之作業環境。客作業系統 410 看到由軟體超管理器 430 提供之一或更多個虛擬處理核心 420，軟體超管

理器 430 實施本文所描述之異質處理系統。軟體超管理器 430 管理異質處理硬體 440，異質處理硬體 440 包括全域中斷控制器 450、一或更多個大核心 460 及一或更多個小核心 470。應注意儘管為說明簡單僅僅圖示兩種核心類型（大的及小的），但系統可以任何數目之不同核心作業。舉例而言，某些處理器套件可包括若干處理儲存器，該等處理儲存器之功率使用及效能逐步減低。軟體超管理器 430 管理異質處理硬體 440 以使客作業系統 410 與任何特別認識或有效使用硬體 440 所需要之處理分離。因此，軟體超管理器 430 允許諸如客作業系統 410 之舊有未改性作業系統，以利用較新的異質處理硬體 440。超管理器 430 可經修改以跟上硬體 440 之變化，同時作業系統 410 仍繼續正常（儘管功率或效能特徵愈來愈佳）作業。

在某些實施例中，異質處理系統在線程已經執行以後從一個實體處理核心遷移線程至另一個實體處理核心。在某些情況下，一或更多個線程可能在超管理器決定減少功率消耗、增強效能或實行其他策略目標後已經執行。核心可共享快取儲存器或其他設備以便超管理器不用影響線程之存取資料就可遷移線程至另一核心。因此，超管理器可中斷線程之執行、移動線程之指令串流至不同實體處理核心，且在目標核心上重新開始執行。

在某些實施例中，異質處理系統在選擇不同核心前應用處理器電壓及頻繁修改以減少功率或增強效能到達閾

值。舉例而言，系統可在大核心上開始執行特定線程，繼而藉由減少大核心之作業電壓按比例縮減大核心之功率使用，且最終可遷移該大核心之工作至小核心。此舉允許該系統逐步減低功率使用以管理熱包絡或滿足策略所規定之其他計算目標。

在某些實施例中，異質處理系統允許某些處理任務遷移至雲端計算設備。系統可向雲端計算設備僅作為另一處理核心提供，任務可排程至該另一處理核心。對適當任務而言，系統可能得以從計算裝置完全卸載任務且隨後返回任務之輸出至客作業系統。此舉可允許系統在計算裝置上進入較低功率狀態或從高峰電耗資料處理中心轉換工作至較低電耗資料處理中心。

在某些實施例中，異質處理系統處理競態條件且應用軟體鎖定範式以管理作業系統期望。在很多情況下，作業系統之排程線程根據相依性或在特定線程之間缺少關聯性。軟體可利用由作業系統提供之鎖定、互斥、信號，或其他同步原語以允許軟體代碼在多個同時執行之線程之環境中正確作業。異質計算系統確保滿足作業系統對於線程安全性及其他同步之保證，且可能引入輔助鎖定或以確保沒有新競態條件或其他問題被引入之方式決定線程排程。

在某些實施例中，異質處理系統包括硬體超管理器。儘管軟體超管理器已用於本文之實例中，彼等一般技術者將認識到對用以實施計算任務之硬體或軟體之選擇常

常可轉變成滿足效能或其他目標之實施細節。因此，系統可使用硬體超管理器實施且可製造某些處理單元以將系統包括在處理單元本身之中。

根據前述內容，應瞭解本異質處理系統之特定實施例出於說明之目的已於本文中描述，但可做各種修改而不背離本發明之精神及範疇。因此，本發明除附加申請專利範圍之外不受限制。

【圖式簡單說明】

第 1 圖係方塊圖，該圖圖示了在一個實施例中異質處理系統之元件。

第 2 圖係流程圖，該圖圖示了在一個實施例中處理異質處理系統在異質處理核心與作業系統之間使用超管理器以初始化具有異質處理核心之計算裝置。

第 3 圖係流程圖，該圖圖示了在一個實施例中處理異質處理系統以經由管理異質處理系統核心之超管理器排程一或更多個作業系統線程。

第 4 圖係方塊圖，該圖圖示了在一個實施例中異質處理系統之作業環境。

【主要元件符號說明】

100	異質處理系統	110	中央處理單元
120	作業系統介面元件	130	虛擬核心管理元件
140	策略引擎元件	150	策略資料庫
160	排程元件	170	能力管理元件

180	硬體介面元件	210	接收啟動請求
220	枚舉實體核心	230	決定核心能力
240	識別作業系統	250	存取超管理器策略
260	建立虛擬核心	270	調用作業系統
310	接收線程排程請求	320	決定線程處理需要
330	存取排程策略	340	選擇實體處理核心
350	管理核心能力差異	360	排程線程以在已選 核心上執行
410	客作業系統	420	虛擬處理核心
430	軟體超管理器	440	異質處理硬體
450	全域中斷控制器	460	大核心
470	小核心		

七、申請專利範圍：

1. 一種經由一超管理器排程一或更多個作業系統線程之電腦實施方法，該超管理器管理異質處理核心，該方法包含以下步驟：

從一作業系統接收一線程排程請求以在由一超管理器所提供之一已識別虛擬核心上執行一線程之指令，其中該虛擬核心使該作業系統與在一計算裝置可存取之實體處理核心之間的一或更多個能力差異分離；

決定該已接收之排程請求之一或更多個處理需要；

存取一排程策略，該排程策略規定操作該計算裝置之一或更多個目標；

選擇一實體處理核心，在該實體處理核心上執行與該已接收之排程請求關聯之該線程，其中該選擇係根據該存取排程策略所做；及

排程該線程以在該已選擇之實體處理核心上執行。

2. 如請求項 1 所述之方法，其中接收一線程排程請求之步驟包含以下步驟：不管該作業系統選擇以執行一線程之該特定虛擬核心為如何，一超管理器依照一或更多個超管理器策略選擇任何特定實體處理核心以執行該線程。

3. 如請求項 1 所述之方法，其中決定處理需要之步驟包含以下步驟：決定由該已排程線程使用之一特定指令組、

可用之協處理器、指令組擴充或構架擴充。

4. 如請求項 1 所述之方法，其中決定處理需要之步驟包含以下步驟：決定該線程之一或更多個效能要求。
5. 如請求項 1 所述之方法，其中決定處理需要之步驟包含以下步驟：決定該線程是否適於在較低功率使用下之較慢執行。
6. 如請求項 1 所述之方法，其中決定處理需要之步驟包含以下步驟：決定該線程是否可被推遲直到額外處理資源可用。
7. 如請求項 1 所述之方法，其中決定處理需要之步驟包含以下步驟：存取關於一特定作業系統之特定認識以決定一特定線程之處理需要。
8. 如請求項 1 所述之方法，其中存取該排程策略之步驟包含以下步驟：決定該策略請求功率使用之最佳化且規定排程線程至可用之較低功率處理核心。
9. 如請求項 1 所述之方法，其中存取該排程策略之步驟包含以下步驟：決定該策略請求效能之最佳化且規定排程線程至可用之較高效能處理核心。

- 10.如請求項 1 所述之方法，其中存取該排程策略之步驟包含以下步驟：決定該計算裝置係具有有限可用功率之一行動裝置或具有熱限制之一裝置且經由在核心上排程線程之該等核心的選擇規定最佳化功率使用或管理熱量。
- 11.如請求項 1 所述之方法，其中選擇該實體處理核心之步驟包含以下步驟：從多個不同能力及效能或功率特徵之可用核心中選擇一個核心，該系統可在該核心上排程該線程。
- 12.如請求項 1 所述之方法，其中選擇該實體處理核心之步驟包含以下步驟：以促進用以管理該計算裝置之效能及功率使用之一或更多個目標之一方式做出該選擇。
- 13.如請求項 1 所述之方法，該方法進一步包含以下步驟：處理該線程及該已選擇之實體處理核心之間的一或更多個能力差異。
- 14.如請求項 1 所述之方法，其中排程該線程以執行之步驟進一步包含以下步驟：處理輸出及提供該輸出返回至該作業系統，使該輸出對該作業系統看來似乎來自該虛擬核心，該作業系統分配該線程至該虛擬核心。

15. 一種經由一超管理器用以提供去耦合異質計算之作業系統之電腦系統，該系統包含：

- 一或更多個處理複合物，該一或更多個處理複合物包括一或更多個處理核心，該一或更多個處理核心具有異質處理能力及功率輪廓；
- 一作業系統介面元件，該作業系統介面元件在一超管理器與一作業系統之間通訊以接收指令，該等指令用以傳遞至硬體資源及用以接收來自該等硬體資源之輸出以使得不同核心對該作業系統看起來相同；
- 一虛擬核心管理元件，該虛擬核心管理元件管理該超管理器向該作業系統提供之一或更多個虛擬核心；
- 一策略引擎元件，該策略引擎元件管理一或更多策略用以根據該可用之一或更多個處理複合物排程作業系統線程及向該作業系統提供虛擬核心；
- 一排程元件，該排程元件排程從該作業系統接收作為線程之一或更多個指令串流至安裝在該計算裝置中之該等處理複合物之一或多個處理複合物；及
- 一硬體介面元件，該硬體介面元件在該超管理器與處理複合物之間通訊以排程軟體指令以在可用之實體核心上執行。

16. 如請求項 15 所述之系統，其中該虛擬核心管理元件提供一虛擬核心，該虛擬核心對該作業系統出現為一 CPU 核心，但與可用之實體硬體特徵不同，以允許該作業系

統與硬體一起工作，該作業系統未設計為與該硬體一起工作。

17.如請求項 15 所述之系統，其中該排程元件從該作業系統接收一虛擬核心識別，該虛擬核心識別識別該虛擬核心，該作業系統請求排程該線程至該虛擬核心。

18.如請求項 15 所述之系統，其中該排程元件檢查該排程請求且決定一實體核心，在該實體核心上排程執行該線程。

19.如請求項 15 所述之系統，該系統進一步包含：一能力管理元件，該能力管理元件管理處理核心之間的一或多個差異。

20.一種電腦可讀取儲存媒體，該電腦可讀取儲存媒體包含指令，該等指令用以控制一電腦系統以使用該等核心與一作業系統之間的一超管理器初始化具有異質處理核心之一計算裝置，其中該等指令在執行時致使一處理器執行動作，該等動作包含：

接收一啟動請求以初始化一計算裝置；

枚舉該計算裝置可存取之兩個或兩個以上處理核心；

決定每一已枚舉處理核心之能力，其中該等能力包括由每一核心所提供之一或多個功率輪廓及每一核心之效能

特徵；

識別一或更多個作業系統，該超管理器將為該等作業系統

管理該等已枚舉實體核心之存取及排程；

存取超管理器策略資訊，該策略資訊規定用以在該等已枚

舉實體處理核心上排程作業系統線程之一或更多個目標；

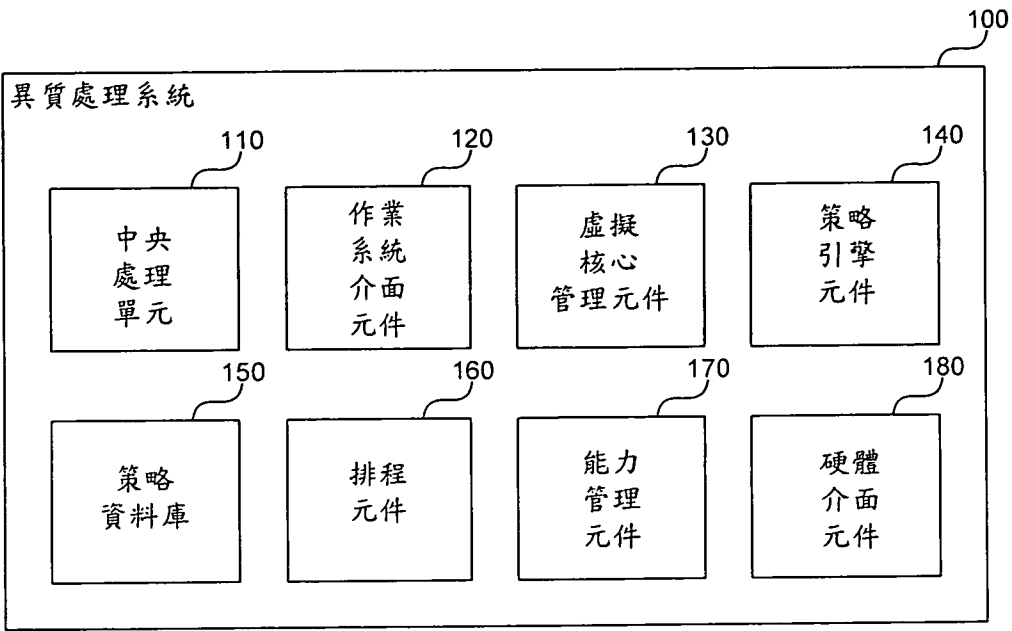
建立一或更多個虛擬核心以暴露於該已識別作業系統，其

中每一虛擬核心使該作業系統與該等實體處理核心間之已決定之能力差異分離；及

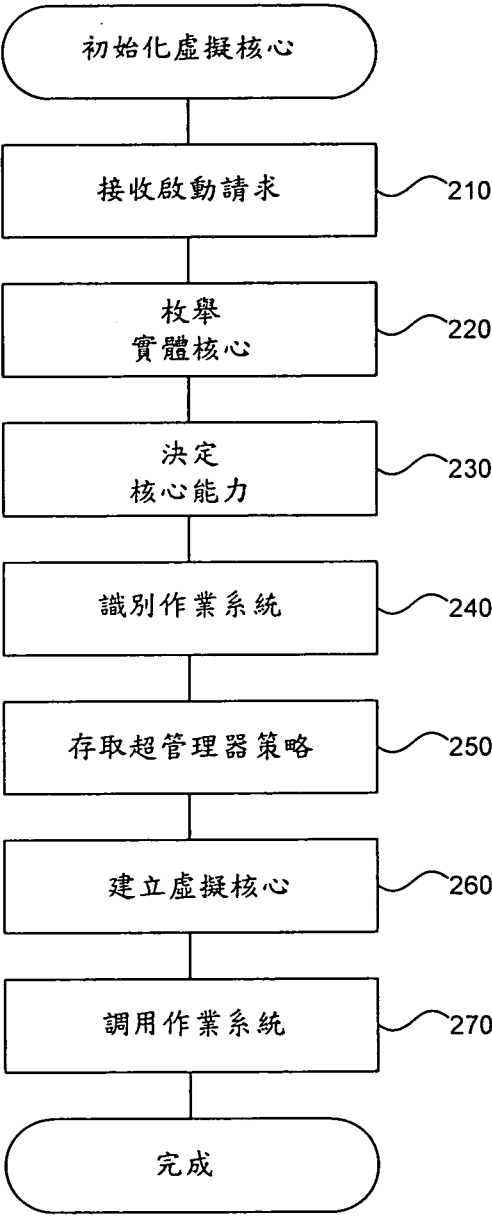
調用該已識別作業系統及向該作業系統提供該等被建立之

虛擬核心同時使該等已識別作業系統與該等已枚舉實體處理核心分離。

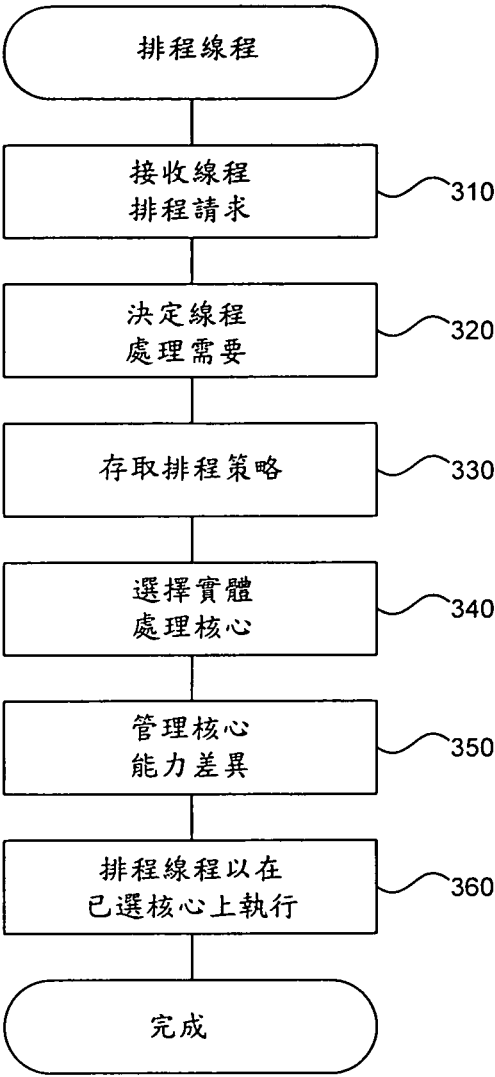
八、圖式：



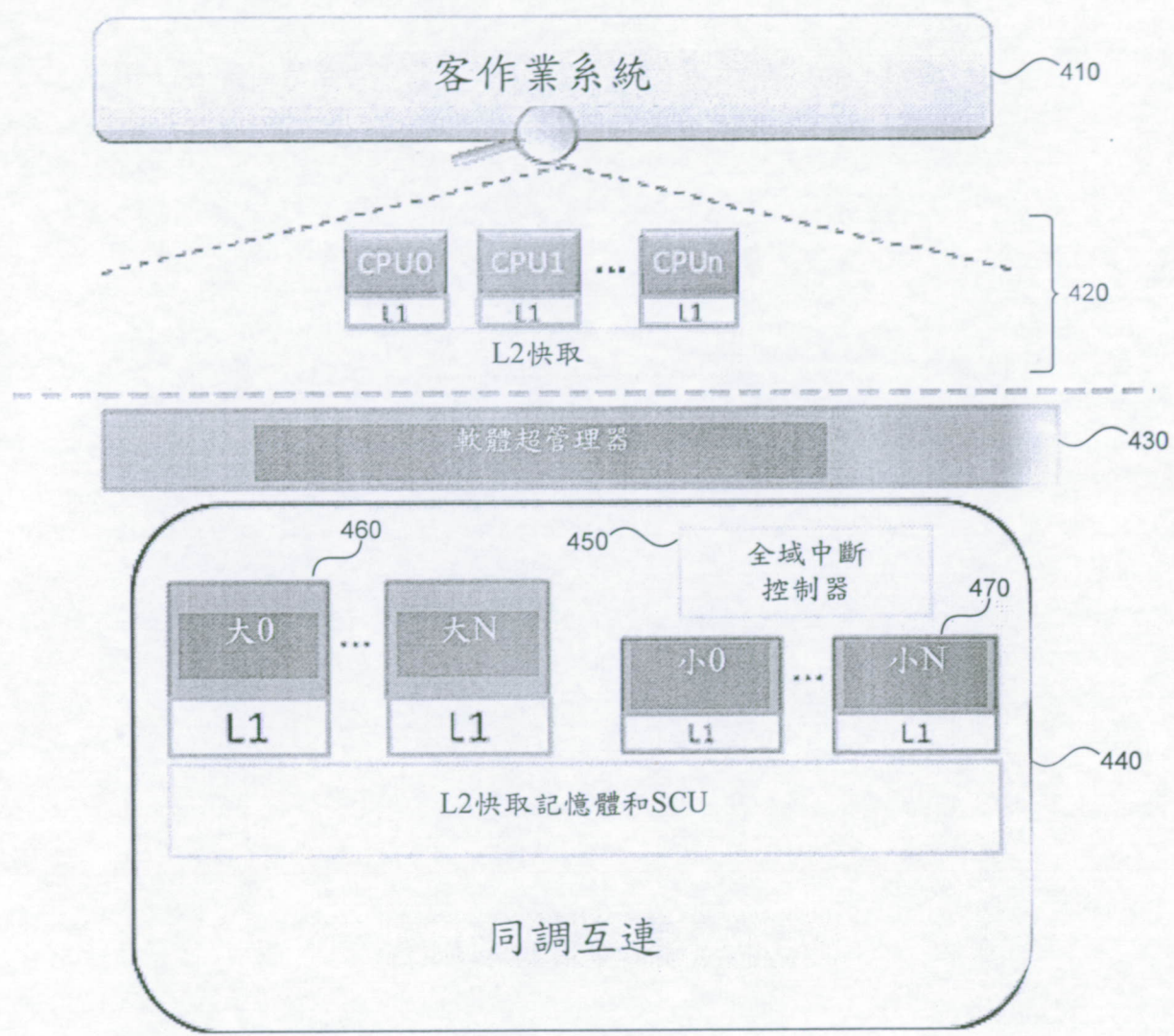
第1圖



第2圖



第3圖



第4圖