四、聲明事項：

☐ 主張專利法第二十二條第二項☐第一款或☐第二款規定之事實，其事實發生日期為：　年　月　日。

☑ 申請前已向下列國家（地區）申請專利：

【格式請依：受理國家（地區）、申請日、申請案號　順序註記】

☑ 有主張專利法第二十七條第一項國際優先權：

1. 美國；2004年05月13日；10/845,550

2.

☐ 無主張專利法第二十七條第一項國際優先權：

1.

2.

☐ 主張專利法第二十九條第一項國內優先權：

【格式請依：申請日、申請案號　順序註記】

☐ 主張專利法第三十條生物材料：

☐ 須寄存生物材料者：

國內生物材料　【格式請依：寄存機構、日期、號碼　順序註記】

國外生物材料　【格式請依：寄存國家、機構、日期、號碼　順序註記】

☐ 不須寄存生物材料者：

所屬技術領域中具有通常知識者易於獲得時，不須寄存。

九、發明說明：

【發明所屬之技術領域】

本發明大體上係關於一種改良之資料處理系統，且尤其係關於一種用於處理資料之方法及裝置。更特定言之，本發明係關於一種用於在一邏輯空間分割資料處理系統中管理追蹤資料之方法、裝置及電腦指令。

【先前技術】

愈來愈多之大型對稱多重處理器資料處理系統，諸如可自國際商用機器公司購得之IBM eServer P690、可自惠普公司購得之DHP9000 Superdome Enterprise Server及可自SUN微系統公司購得之Sunfire 15K server，不再用作單一大型資料處理系統。實情為，該等類型之資料處理系統正被空間分割且用作較小之系統。此等系統亦被稱作邏輯空間分割(LPAR)資料處理系統。一資料處理系統內之一邏輯空間分割之功能性允許一單一作業系統的多個複本或多個不同種類之作業系統同時運作於一單一資料處理系統平臺上。一作業系統鏡像運作於一空間分割中，該空間分割被指派予平臺資源之一非重疊子集。此等平臺可指派資源包括具有中斷管理區域之一或多個不同架構之處理器、系統記憶體區域及輸入/輸出(I/O)配接器匯流排插槽。該等空間分割之資源由作業系統鏡像之平臺韌體表示。

運作於一平臺內之每一不同作業系統或一作業系統之鏡像相互間受保護，以便一邏輯空間分割上之軟體錯誤不會影響任一其它空間分割之正確運作。藉由分配一組由每一

作業系統鏡像直接管理之不相交平臺資源且藉由提供用於確保各種鏡像不可控制未分配至該鏡像之任何資源之機制來提供此保護。此外，防止由作業系統之已分配資源控制之軟體錯誤影響任何其它鏡像之資源。

因此，作業系統之每一鏡像或每一不同之作業系統直接控制平臺內一組不同之可分配資源。相對於一邏輯空間分割資料處理系統中之硬體資源，此等資源係不相交地共用於各種空間分割中。此等資源可包括(例如)輸入/輸出(I/O)配接器、記憶體DIMM、非揮發性隨機存取記憶體(NVRAM)及硬碟驅動器。一LPAR資料處理系統內之每一空間分割可反覆啟動及關閉，而不必動力循環整個資料處理系統。

當一邏輯空間分割資料處理系統經歷一故障時，需要有關處理及系統狀態之資料幫助識別及分析該故障。在目前之邏輯空間分割資料處理系統中，由於目前系統設計之原因，診斷一故障所需之某些資料無法獲得。舉例而言，平臺韌體包括一允許在該韌體中追蹤代碼路徑之追蹤功能。用於邏輯空間分割資料處理系統中之平臺韌體之一實例為一超級監督器，其可自國際商用機器公司購得。

藉由目前所使用之追蹤功能，當每一空間分割產生平臺韌體呼叫時，展示平臺固件中所採用之代碼路徑的追蹤資訊及關鍵資料值寫入一追蹤緩衝器。當一空間分割遇到一錯誤且此錯誤路徑連同關鍵資料值均被追蹤時，此追蹤資訊尤其關鍵。

目前所有邏輯空間分割模式資料處理系統平臺支持一超級監督器追蹤功能，其用於在超級監督器執行期間將超級監督器代碼執行追蹤點資料寫入一位於超級監督器空間中之追蹤緩衝器中。此超級監督器追蹤資料對於系統故障事件領域中之有效故障分析係關鍵的。

所有超級監督器呼叫追蹤點使用相同之追蹤緩衝器且藉由呼叫一包括被指派為空間分割0之空間分割管理器之空間分割號碼來識別。因為該超級監督器就記憶體分配而言係靜態的，所以此緩衝器具有固定長度且對於平臺之最大數目之空間分割必須足夠大。當需要追蹤資料時，其可使用一作業系統指令(諸如進階互動式執行(AIX)作業系統指令"fetchdog")自任一空間分割中擷取。此指令載入一特定互動式執行(AIX)作業系統設備驅動，該驅動使一超級監督器呼叫將追蹤資料複製至空間分割之空間中。

因為此指令收集整個超級監督器追蹤緩衝器之內容，所以該命令允許來自包括空間分割管理器之所有空間分割之資料由一單一空間分割擷取，使邏輯空間分割資料處理系統曝露於安全弱點。此等安全弱點包括雙向隱藏儲存通道、由另一空間分割分析空間分割運作及超級監督器運作之分析。

雙向隱藏儲存通道可建立於空間分割之間。因為追蹤緩衝器被組成一循環緩衝器，所以通信之所有跡象由系統週期性移除。隨著更多資料寫入該緩衝器中，舊資料最終被覆寫。通道將資料儲存一足夠長之時期以成為以非常高之

資料速率轉移之資料的有效構件。

建立一雙向隱藏儲存通道中發生之運作包括使傳輸空間分割使用編碼輸入參數產生一超級監督器呼叫。此等參數藉由超級監督器追蹤功能寫入該超級追蹤緩衝器中。接收空間分割產生連續呼叫以使用"fetchdbh"指令擷取追蹤資料或使用呼叫"h_hypervisor_debug()"直接呼叫該超級監督器以擷取該追蹤資料。該接收空間分割可由空間分割過濾追蹤資料，且當呼叫偵測到關鍵模式為已知呼叫時完成隱藏通道通信路徑。另外，該接收空間分割藉由使用編碼輸入參數產生超級監督器呼叫亦可充當傳輸器。此處理可用於在空間分割之間建立全雙向隱藏"插口"。

一額外安全弱點為藉由另一空間分割分析一空間分割之運作。可藉由分析空間分割進行此類型之分析，使得每一超級監督器呼叫且讀取追蹤緩衝器以建立每一超級監督器呼叫之特徵。分析空間分割產生連續呼叫以擷取追蹤資料。然後，藉由使用由讀取該追蹤緩衝器所獲得之資料的簡單分析，該空間分割可推斷所有其它空間分割處於何狀態且開發一種攻擊方法。

另一安全弱點為平臺韌體運作之分析，諸如一超級監督器之該等運作。藉由監視自每一追蹤點傳回之值及尤其是該空間分割管理器之值，有可能開發出能夠起動有效空間分割間攻擊以及意欲使整個機器崩潰之攻擊方法。

因此，有利地應存在一用於消去與追蹤資料相關之安全弱點之經改良方法、裝置及電腦指令。

【發明內容】

本發明提供一種在一邏輯空間分割資料處理系統中用於管理追蹤資料之方法、裝置及電腦指令。自該邏輯空間分割資料處理系統之複數個空間分割內之一呼叫空間分割接收一追蹤資料之呼叫。識別一與該呼叫空間分割相關聯之緩衝器中之追蹤資料以形成經識別之追蹤資料。僅傳回呼叫空間分割之經識別追蹤資料。複數個空間分割內之其它空間分割之追蹤資料不被傳回至呼叫空間分割。

【實施方式】

現參看圖式，且詳言之參看圖1，其描述了可實施本發明之一資料處理系統的一方塊圖。資料處理系統100可為一對稱型多重處理器(SMP)系統，其包括連接至系統匯流排106之複數個處理器101、102、103及104。舉例而言，資料處理系統100可為在一網路內充當伺服器之IBM eServer，一位於紐約Armonk之國際商用機器公司(International Business Machines Corporation)之產品。或者，可採用一單一處理器系統。記憶體控制器/快取記憶體108亦連接至系統匯流排106，其向複數個本地記憶體160至163提供一介面。I/O匯流排橋接器110連接至系統匯流排106且向I/O匯流排112提供一介面。如圖所描繪，記憶體控制器/快取記憶體108及I/O匯流排橋接器110可整合到一起。

資料處理系統100係一邏輯空間分割(LPAR)資料處理系統。因此，資料處理系統100可具有同時運作的多個不同

種類之作業系統(或一單一作業系統之多個實例)。此等多個作業系統中之每一個可具有於其內執行之任何數目之軟體程式。資料處理系統100被邏輯空間分割以便不同之PCI I/O配接器120-121、128-129及136、圖形配接器148及硬碟配接器149可指派給不同之邏輯空間分割。在此情況下，圖形配接器148為一顯示設備(未圖示)提供一連接，而硬碟配接器149提供一連接以控制硬碟150。

因此，舉例而言，假定資料處理系統100劃分為三個邏輯空間分割P1、P2及P3。PCI I/O配接器120-121、128-129、136中之每一個、圖形配接器148、硬碟配接器149、主機處理器101至104中之每一個及來自本地記憶體160至163之記憶體被指派給該等三個空間分割中之每一個。在此等實例中，記憶體160至163可採取雙同軸(dual in-line)記憶體模組(DIMMs)之形式。DIMMs通常不以每一DIMM為基礎指派給空間分割。實情為，一空間分割將獲得平臺所見之全部記憶體之一部分。舉例而言，處理器101、來自本地記憶體160至163之記憶體的某些部分及I/O配接器120、128及129可指派給邏輯空間分割P1；處理器102至103、來自本地記憶體160至163之記憶體的某些部分及PCI I/O配接器121及136可指派給邏輯空間分割P2；且處理器104、來自本機記憶體160至163之記憶體的某些部分及圖形配接器148及硬碟配接器149可指派給邏輯空間分割P3。

在資料處理系統100內執行之每一作業系統指派給不同之邏輯空間分割。因此，在資料處理系統100內執行之每

一作業系統僅可存取處於其邏輯空間分割內之該等I/O單元。因此，舉例而言，進階互動式執行(AIX)作業系統之一實例可在空間分割P1內執行，該AIX作業系統之第二實例(鏡像)可在空間分割P2內執行，且一Linux或OS/400作業系統可在邏輯空間分割P3內運作。

連接至I/O匯流排112之周邊元件互連(PCI)主機橋接器114向PCI本地匯流排115提供一介面。多個PCI輸入/輸出配接器120至121可經由PCI至PCI橋接器116、PCI匯流排118、PCI匯流排119、I/O插槽170及I/O插槽171連接至PCI匯流排115。PCI至PCI橋接器116向PCI匯流排118及PCI匯流排119提供一介面。PCI I/O配接器120及121分別置放於I/O插槽170及171中。典型之PCI匯流排實施將支持4個與8個之間的I/O配接器(意即附加連接器之擴充槽)。每一PCI I/O配接器120至121在資料處理系統100與輸入/輸出設備諸如(例如)資料處理系統100之用戶端的其它網路電腦之間提供一介面。

一額外PCI主機橋接器122為一額外PCI匯流排123提供一介面。PCI匯流排123連接至複數個PCI I/O配接器128至129。PCI I/O配接器128至129可經由PCI至PCI橋接器124、PCI匯流排126、PCI匯流排127、I/O插槽172及I/O插槽173連接至PCI匯流排123。PCI至PCI橋接器124向PCI匯流排126及PCI匯流排127提供一介面。PCI I/O配接器128及129分別置放於I/O插槽172及173中。以此方式，額外I/O設備諸如(例如)數據機或網路配接器可經由PCI I/O配接器

128至129中之每一個支撐。以此方式，資料處理系統100允許連接至多個網路電腦。

一插入I/O插槽174之記憶體映射圖形配接器148可經由PCI匯流排144、PCI至PCI橋接器142、PCI匯流排141及PCI主機橋接器140連接至I/O匯流排112。硬碟配接器149可置放於連接至PCI匯流排145之I/O插槽175中。接著，此匯流排連接至PCI至PCI橋接器142，PCI至PCI橋接器142藉由PCI匯流排141連接至PCI主機橋接器140。

一PCI主機橋接器130為一PCI匯流排131提供一介面以連接至I/O匯流排112。PCI I/O配接器136連接至I/O插槽176，I/O插槽176藉由PCI匯流排133連接至PCI至PCI橋接器132。PCI至PCI橋接器132連接至PCI匯流排131。此PCI匯流排亦將PCI主機橋接器130連接至服務處理器郵箱介面與ISA匯流排存取通過邏輯194及PCI至PCI橋接器132。服務處理器郵箱介面與ISA匯流排存取通過邏輯194將PCI存取目標指向PCI/ISA橋接器193。NVRAM儲存器192連接至ISA匯流排196。服務處理器135經由其本地PCI匯流排195耦接於服務處理器郵箱介面與ISA匯流排存取通過邏輯194。服務處理器135亦經由複數個JTAG/I^2C匯流排134連接至處理器101至104。JTAG/I^2C匯流排134係JTAG/掃描匯流排(見IEEE 1149.1)與Phillips I^2C匯流排之組合。然而，或者，JTAG/I^2C匯流排134可由單個Phillips I^2C匯流排或單個JTAG/掃描匯流排替代。主機處理器101、102、103及104之所有SP-ATTN訊號共同連接至一服務處理器之中斷

輸入訊號。該服務處理器135具有其自身之本地記憶體191，且可存取硬體OP面板190。

當資料處理系統100最初加電時，服務處理器135使用JTAG/I^2C匯流排134以詢問系統(主機)處理器101-104、記憶體控制器/快取記憶體108及I/O橋接器110。此步驟完成時，服務處理器135具有資料處理系統100之清單及拓撲協定(topology understanding)。服務處理器135亦藉由詢問主機處理器101至104、記憶體控制器/快取記憶體108及I/O橋接器110對所發現之所有元件執行內建式自我測試常用程式(BIST)、基本保證測試(BAT)及記憶體測試。在BIST、BAT及記憶體測試期間所偵測到之故障的任何錯誤資訊由服務處理器135收集及報告。

若在BIST、BAT及記憶體測試期間去除有故障元件之後系統資源之有意義/有效組態仍可行，則允許資料處理系統100繼續將可執行代碼載入本地(主機)記憶體160至163。然後服務器處理器135釋放主機處理器101至104用於執行載入本地記憶體160至163中之代碼。當主機處理器101至104在資料處理系統100內執行來自個別作業系統之代碼時，服務處理器135進入一監視及報告錯誤模式。服務處理器135所監視之項目類型包括(例如)冷卻風扇速度及運作、熱感應器、電源調節器及由處理器101至104所報告之可恢復與不可恢復之錯誤、本機記憶體160至163及I/O橋接器110。

服務處理器135負責保存及報告與資料處理系統100中之

所有監視項目相關之錯誤資訊。服務處理器135亦基於錯誤類型及所界定之臨限值進行運作。舉例而言，服務處理器135可注意一處理器之快取記憶體上之連續可恢復錯誤，且判定此係一硬體故障之預兆。基於此判定，服務處理器135可在當前運作之會話及將來之初始程式載入(IPL)期間標記取消組態之資源。IPL有時亦被稱作"啟動程式(boot)"或"引導程式(bootstrap)"。

資料處理系統100可使用各種市售電腦系統實施。舉例而言，資料處理系統100可使用可自國際商用機器公司購得之IBM eServer iSeries Model 840系統實施。此系統可使用亦可自國際商用機器公司購得之OS/400作業系統支持邏輯空間分割。

此項技術中之該等一般技術者將瞭解圖1中所描繪之硬體可變化。舉例而言，除了所描繪之硬體之外或替代所描繪之硬體亦可使用其它周邊設備，諸如光碟驅動器及其類似物。所描繪之實例並非意謂關於本發明之架構限制。

現在參看圖2，描繪了可實施本發明之一例示性邏輯空間分割平臺之一方塊圖。邏輯空間分割平臺200中之硬體可實施為(例如)圖1中之資料處理系統100。邏輯空間分割平臺200包括空間分割之硬體230、作業系統202、204、206、208及空間分割管理韌體210。作業系統202、204、206及208可為單一作業系統之多個複本或在邏輯空間分割平臺200上同時運作之多個不同種類之作業系統。此等作業系統可使用OS/400實施，OS/400設計用以與一空間分割

管理韌體(諸如超級監督器)建立介面。OS/400僅用作此等說明性實施例中之一實例。當然，可視特定實施而定使用其它類型之作業系統，諸如AIX及linux。作業系統202、204、206及208位於空間分割203、205、207及209中。超級監督器軟體為一可用以實施空間分割管理韌體210之軟體的實例，且可自國際商用機器公司購得。韌體為儲存於一記憶體晶片中之"軟體"，該記憶體芯片在沒有電功率之情況下固持其內容，諸如(例如)唯讀記憶體(ROM)、可程式化ROM(PROM)，可擦可程式化ROM(EPROM)、電可擦可程式化ROM(EEPROM)及非揮發性隨機存取記憶體(非揮發性RAM)。

另外，此等部分亦包括空間分割韌體211、213、215及217。空間分割韌體211、213、215及217可使用初始引導程式代碼、IEEE-1275標準開放韌體及可自國際商用機器公司購得之運作時間抽象軟體(RTAS)實施。當空間分割203、205、207及209實體化時，一引導程式代碼之複本藉由平臺韌體210載入至空間分割203、205、207及209上。此後，控制轉移至引導程式代碼，然後引導程式代碼載入開放韌體及RTAS。然後與該等空間分割相關聯或被指派給該等空間分割之處理器被分派至空間分割之記憶體以執行空間分割韌體。

空間分割硬體230包括複數個處理器232至238、複數個系統記憶體單元240至246、複數個輸入/輸出(I/O)配接器248至262及一儲存單元270。處理器232至238、記憶體單

元240至246、NVRAM儲存器298及I/O配接器248至262中之每一個可指派給邏輯空間分割平臺200內之多個空間分割中的一個，該等空間分割中之每一個對應於作業系統202、204、206及208中的一個。

空間分割管理韌體210對空間分割203、205、207及209執行多個功能及服務以創建並增強邏輯空間分割平臺200之空間分割。空間分割管理韌體210為一實施與下層硬體一致之虛擬機的韌體。因此，空間分割管理韌體210藉由虛擬化邏輯空間分割平臺200之所有硬體資源允許同時執行獨立之OS鏡像202、204、206及208。

服務處理器290可用於提供各種服務，諸如空間分割中平臺錯誤之處理。此等服務亦可充當服務代理以將錯誤報告傳回至廠商，諸如國際商用機器公司。可經由一硬體管理控制臺諸如硬體管理控制臺280來控制不同空間分割之運作。硬體管理控制臺280為一獨立資料處理系統，系統管理員可自其上執行包括將資源再分配至不同空間分割之各種功能。

本發明提供一種以一減少安全弱點之方式管理追蹤資料之方法、裝置及電腦指令。詳言之，防止雙向隱藏儲存通道、由另一空間分割分析空間分割運作及平臺韌體運作之分析。本發明之機制採用一過濾器以選擇資料，該資料被傳回以回應來自追蹤資料之空間分割的呼叫。此過濾器僅傳回與呼叫空間分割相關聯之追蹤資料。不傳回其它空間分割之其它資料。藉由所選之空間分割諸如一服務空間分

割，彌補了安全弱點。

現在參看圖3A及3B，其展示了用以處理追蹤資料之目前可利用之邏輯空間分割資料處理系統中之組件的圖式。在此說明性實例中，空間分割300含有圖3A中之作業系統302及RTAS 304。空間分割306含有作業系統308及RTAS 310。在此等所描繪之實施例中，可存在總計255個空間分割。所有此等空間分割經由平臺韌體諸如超級監督器312管理。

如同代碼路徑318，空間分割管理器314將追蹤資訊寫入超級監督器追蹤緩衝器316中。當此元件產生對超級監督器312之呼叫時，寫入此追蹤資訊。空間分割管理器314係超級監督器312內之組件。此組件用於管理空間分割且包括諸如開始及終止空間分割之功能。由空間分割管理器314產生之呼叫儲存於超級監督器追蹤緩衝器316內。

以一相似之方式，當呼叫由空間分割306產生對超級監督器312之呼叫時，超級監督器代碼路徑320由儲存於超級監督器追蹤緩衝器316中之此等呼叫之追蹤資料形成。以一相似之方式，當呼叫超級監督器312接收到由空間分割306產生之呼叫時，形成超級監督器代碼路徑320。用於此路徑之追蹤資料亦儲存於超級監督器追蹤緩衝器316中。

當每一空間分割啟動時，載入作業系統且開始執行。當一作業系統諸如作業系統302需要平臺資源時，該作業系統向RTAS 304產生RTAS呼叫，該RTAS 304又向超級監督器312產生超級監督器呼叫。當一超級監督器呼叫執行

時，執行特定"追蹤點"，其中追蹤資料寫入超級監督器追蹤緩衝器316中，不與空間分割資料隔離。換言之，所有空間分割之所有呼叫之所有追蹤資料均置放於此緩衝器中。

不僅執行代表空間分割之超級監督器呼叫之超級監督器資料被寫入單一緩衝器，而且空間分割管理器314(其為一超級監督器裝備)亦將其追蹤資料寫入相同之追蹤緩衝器。另外，機器檢驗中斷處理器322使用相同之追蹤裝備且將資訊儲存於超級監督器追蹤緩衝器316中。

在圖3B中之擷取追蹤資料之正常除錯(debug)運作中，空間分割300建立資料緩衝器324且向"h_hypervisor_debug()"常用程式產生一呼叫。此呼叫之執行導致原超級監督器追蹤資料被複製至空間分割空間中用於分析。在此特定實例中，該資料被複製至資料緩衝器324。當此呼叫發生時，當前對資料轉移不加以限制。此類型之非限制資料複本亦允許在兩個空間分割之間建立一隱藏儲存通道。

因此，資料緩衝器324中之資料可藉由來自空間分割306之一系列超級監督器呼叫傳遞至超級監督器312所提供之追蹤裝備中之超級監督器追蹤緩衝器316中。位於超級監督器追蹤緩衝器316中之空間分割306之此資料可使用"h_hypervisor_debug()"超級監督器呼叫由空間分割300擷取。本文所描述之此呼叫及其它特定呼叫為超級監督器產品中目前可利用之呼叫。以此方式，可建立一隱藏儲存通

道。另外，一空間分割諸如空間分割306可自超級監督器追蹤緩衝器316擷取由空間分割管理器314及機器檢驗中斷處理器322產生之追蹤資料。此資訊允許分析空間分割之運作及超級監督器之運作。

現在參看圖4，根據本發明之一較佳實施例描繪了一用於管理追蹤資料以消去安全弱點之組態的一圖式。在此說明性實例中，存在空間分割400及空間分割402且經由超級監督器404進行管理。空間分割400含有作業系統406及RTAS 408，而空間分割402含有作業系統410及RTAS 412。在此實例中，作業系統400亦包括資料緩衝器414且作業系統402含有資料緩衝器416。當空間分割400或空間分割402向超級監督器404產生呼叫時，藉由儲存於超級監督器追蹤緩衝器422中之追蹤資料產生超級監督器代碼路徑418及超級監督器代碼路徑420。

此外，空間分割管理器424可產生呼叫，其中追蹤資料儲存於超級監督器追蹤緩衝器422中。機器檢驗中斷處理器426亦將資料儲存於超級監督器追蹤緩衝器422中。在此說明性實例中，過濾器428亦存在於超級監督器404中。於此等實例中，當超級監督器404接收到追蹤資料之呼叫時，此過濾器用以限制傳回之資料。

在說明性實例中，空間分割之間的隱藏通道經由應用過濾器428中之簡單過濾演算法來消去。在此等說明性實例中，此過濾器用於"h_hypervisor_debug()"呼叫中。此常用程式呼叫read_trace ()常用程式430以實際上轉移追蹤資

料。

本發明之機制藉由增加過濾器428來修改此常用程式以選擇回應一追蹤資料呼叫之待傳回資料。過濾器428防止非限制資料自超級監督器追蹤緩衝器422轉移至一空間分割，諸如空間分割400。在所描繪之實例中，當輸入至一過濾演算法或處理時，過濾器428使用自空間分割之呼叫所識別之當前空間分割數目及追蹤緩衝器內識別該資料屬於何空間分割之資料段作為過濾演算法或程序的輸入。

超級監督器追蹤緩衝器422中之資料排列作為記錄。在該等說明性實例中，當追蹤資料由來自一空間分割之呼叫產生時，創建超級監督器追蹤緩衝器422中之每一記錄。在記錄之資料段中識別引起在記錄中產生追蹤資料之呼叫的空間分割。在過濾器428中本發明之過濾機制比較呼叫空間分割之識別及與超級監督器追蹤緩衝器422中之記錄相關聯之空間分割的識別。

此過濾器僅將呼叫空間分割之資料傳遞至該空間分割，除非呼叫空間分割被識別為"服務空間分割"之情況。在此等說明性實例中，該服務空間分割為一給定特殊許可以執行服務功能(例如代碼更新)之空間分割。因為平臺管理員必須使用硬體管理控制臺指派服務空間分割，所以假定此空間分割安全且此空間分割不接收隱藏資料。

以此方式，防止了空間分割之間的隱藏通道。藉由限制過濾器428之資料轉移，消去了由另一空間分割分析空間分割運作及超級監督運作之分析。

現在參看圖5，根據本發明之一較佳實施例描繪了用於傳回追蹤資料以回應空間分割中追蹤資料之一呼叫的程序之流程圖。圖5中所說明之程序可在一過濾器(諸如位於圖4之read_trace常用程式430中之過濾器428)中實施。

作為對自一空間分割接收追蹤資料之呼叫或請求之回應，開始該程序。此請求包括呼叫空間分割諸如一空間分割數目或一位址之識別，及一資料緩衝器之識別。做出關於該呼叫空間分割是否為一服務空間分割之判定(步驟500)。在此等說明性實例中之服務空間分割不考慮潛在之安全風險。若該呼叫空間分割不為一服務空間分割，則該程序轉至追蹤緩衝器中下一未處理之追蹤資料記錄(步驟502)。在此等說明性實例中，該資料被組織為追蹤緩衝器中之多個記錄。

然後，使當前空間分割數目與追蹤資料記錄之資料段進行比較(步驟504)。追蹤記錄中之資料段包括一產生超級監督器呼叫之空間分割之識別，該超級監督器呼叫導致追蹤記錄中之資料的產生。此後，做出關於該追蹤資料記錄是否屬於當前空間分割之判定(步驟506)。若該追蹤資料記錄屬於當前空間分割，則該追蹤資料記錄被複製至資料緩衝器(步驟508)。然後，做出關於在追蹤緩衝器中是否存在更多未處理之追蹤資料之判定(步驟510)。若在追蹤緩衝器中不存在更多未處理之追蹤資料，則該程序結束。

返回步驟500，若該呼叫空間分割為一服務空間分割，則高達資料緩衝器容量之數目的追蹤資料記錄自追蹤緩衝

器轉移至資料緩衝器(步驟512)，其後結束該程序。在步驟506中，若追蹤資料記錄不屬於當前空間分割，則如上文所描述該程序繼續進行至步驟510。在步驟510中，若追蹤緩衝器中存在更多未處理之追蹤資料記錄，則如上文所描述該處理繼續進行至步驟502。

然後轉至圖6，根據本發明之一較佳實施例描繪了用於轉移追蹤資料之代碼的圖式。圖6中所說明之代碼可作為平臺韌體(諸如圖4中之超級監督器404)中之呼叫實施。詳言之，所說明之代碼可在一資料傳輸常用程式(諸如圖4中之read_trace常用程式430)中實施。

在此等說明性實例中，h_hypervisor_debug()常用程式執行初始處理，然後呼叫read_trace()常用程式600以轉移追蹤資料。本發明之過濾機制位於此常用程式中。在此實例中，read_trace()常用程式600為上文圖5中所描繪之程序的實例示實施。本發明之過濾機制在read_trace()常用程式600之區域602中實施。

因此，本發明提供一種用於保護追蹤資料之經改良的方法、裝置及電腦指令。本發明之機制自空間分割接收追蹤資料之請求，且自追蹤緩衝器僅傳回請求空間分割之資料。以此方式，可經由使用本發明之過濾機制消去空間分割之間的隱藏通道。因此，在藉由平臺韌體處理追蹤資料時減少或消去了安全弱點。

重要的是，應注意雖然已經在全功能資料處理系統之情況下描述了本發明，但是此項技術中之該等一般技術者將

瞭解本發明之程序能夠以電腦可讀取媒體之指令形式及各種形式分佈，且不論實際用於進行該分佈之特定類型之訊號承載媒體本發明可同等地應用。電腦可讀取媒體之實例包括可記錄型媒體，諸如軟碟、硬碟驅動器、RAM、CD-ROM、DVD-ROM及傳輸型媒體，諸如數位及類比通信鏈路、使用諸如(例如)射頻及光波傳輸之傳輸形式的有線或無線通信鏈路。該電腦可讀取媒體可採用編碼格式，在特定資料處理系統中該編碼格式被解碼用於實際應用。

本發明之描述僅出於說明及描述之目的，且並非意欲無遺漏的或將本發明限制於所揭示之形式。熟習此項技術者將易瞭解諸多修改及變化。本發明所選及所描述之實施例是為了最好地說明本發明之原理及實際應用，且使此項技術中之其它一般技術者能夠理解本發明及具有適於預期之特殊使用的各種修改之各種實施例。

【圖式簡單說明】

圖1係可實施本發明之一資料處理系統的一方塊圖；

圖2係可實施本發明之一例示性邏輯空間分割平臺的一方塊圖；

圖3A及3B係用以處理追蹤資料之目前可利用之邏輯空間分割資料處理系統中之組件的圖式；

圖4係根據本發明之一較佳實施例用於管理追蹤資料以消去安全弱點之組態的一圖式；

圖5係根據本發明之一較佳實施例用於傳回追蹤資料以回應一空間分割中追蹤資料之呼叫的程序之流程圖；及

圖6係根據本發明之一較佳實施例用於轉移追蹤資料之代碼的圖式。

【主要元件符號說明】

100	資料處理系統
101、102、103、104	處理器
106	系統匯流排
108	記憶體控制器/快取記憶體
110	I/O匯流排橋接器
112	I/O匯流排
114	周邊元件互連(PCI)主機橋接器
115	PCI本地匯流排
116、124、132、142	PCI至PCI橋接器
118、119、123、126、127、131、133、141、144、145	PCI匯流排
120、121、128、129、136	PCI I/O配接器
122、130、140	PCI主機橋接器
134	JTAG/I^2C匯流排
135	服務處理器
148	圖形配接器

149	硬碟配接器
150	硬碟
160、161、162、163	本地記憶體
170、171、172、173、174、175、176	I/O插槽
190	硬體OP面板
191	本地記憶體
192	NVRAM儲存器
193	PCI/ISA橋接器
194	服務處理器郵箱介面與ISA匯流排存取通過邏輯
195	本地PCI匯流排
196	ISA匯流排
200	邏輯空間分割平臺
202、204、206、208	作業系統
203、205、207、209	空間分割
210	空間分割管理韌體/平臺韌體
211、213、215、217	空間分割韌體
230	空間分割硬體

232、234、236、238	處理器
240、242、244、246	記憶體單元
248、250、252、254、256、258、260、262	輸入/輸出(I/O)配接器
270	儲存單元
280	硬體管理控制臺
290	服務處理器
298	NVRAM儲存器
300、306	空間分割
302、308	作業系統
304、310	RTAS
312	超級監督器
314	空間分割管理器
316	超級監督器追蹤緩衝器
318	代碼路徑
320	超級監督器代碼路徑
322	機器檢驗中斷處理器
324	資料緩衝器
400、402	空間分割
404	超級監督器
406、410	作業系統

408、412	RTAS
414、416	資料緩衝器
418、420	超級監督器代碼路徑
422	超級監督器追蹤緩衝器
424	空間分割管理器
426	機器檢驗中斷處理器
428	過濾器
430	read_trace()常用程式
600	read_trace()常用程式

五、中文發明摘要：

本發明揭示一種在一邏輯空間分割資料處理系統中用於管理追蹤資料之方法、裝置及電腦指令。自該邏輯空間分割資料處理系統之複數個空間分割內的一呼叫空間分割接收一對該追蹤資料之呼叫。識別與該呼叫空間分割相關聯之一緩衝器中之該追蹤資料以形成經識別之追蹤資料。僅傳回該呼叫空間分割之該經識別之追蹤資料。該等複數個空間分割內之其它空間分割之該追蹤資料不被傳回至該呼叫空間分割。

六、英文發明摘要：

十一、圖式：

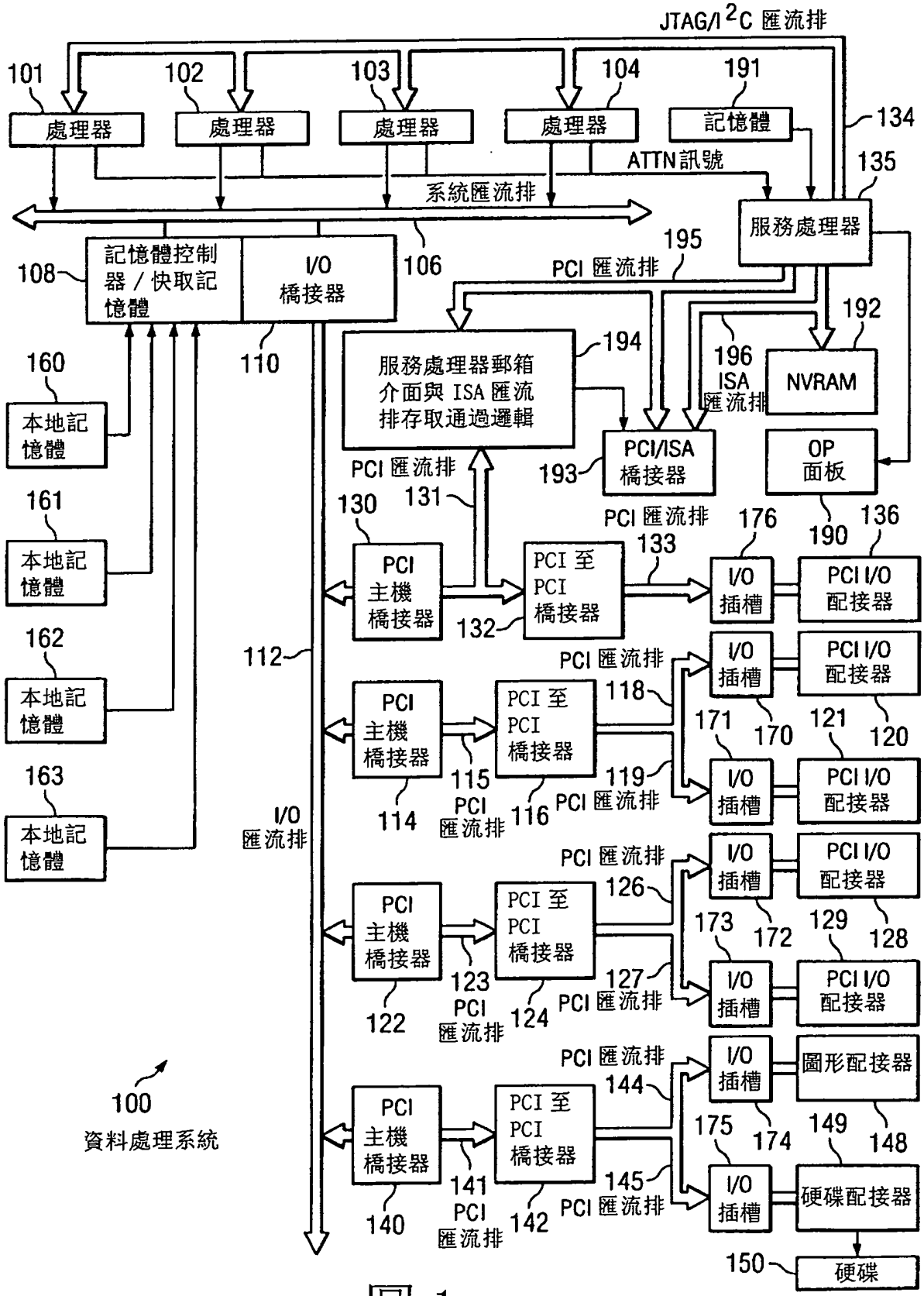
JTAG/I2C 匯流排
101
102
103
104
191
處理器
處理器
處理器
處理器
記憶體
134
ATTN訊號
135
系統匯流排
服務處理器
195
108
記憶體控制器／快取記憶體
I/O 橋接器
106
PCI 匯流排
192
160
110
服務處理器郵箱介面與 ISA 匯流排存取通過邏輯
194
196
ISA 匯流排
NVRAM
本地記憶體
PCI/ISA 橋接器
193
OP 面板
PCI 匯流排
161
130
131
PCI 匯流排
176
190
136
本地記憶體
PCI 主機橋接器
PCI 至 PCI 橋接器
133
I/O 插槽
PCI I/O 配接器
132
162
112
PCI 匯流排
118
171
170
121
120
本地記憶體
PCI 主機橋接器
PCI 至 PCI 橋接器
I/O 插槽
PCI I/O 配接器
115
119
163
114
PCI 匯流排
116
PCI 匯流排
I/O 匯流排
PCI 匯流排
126
173
172
129
128
本地記憶體
PCI 主機橋接器
PCI 至 PCI 橋接器
I/O 插槽
PCI I/O 配接器
123
127
122
PCI 匯流排
124
PCI 匯流排
PCI 匯流排
144
175
174
149
148
100
資料處理系統
PCI 主機橋接器
PCI 至 PCI 橋接器
I/O 插槽
圖形配接器
141
145
140
PCI 匯流排
142
PCI 匯流排
I/O 插槽
硬碟配接器
150
硬碟

圖 1

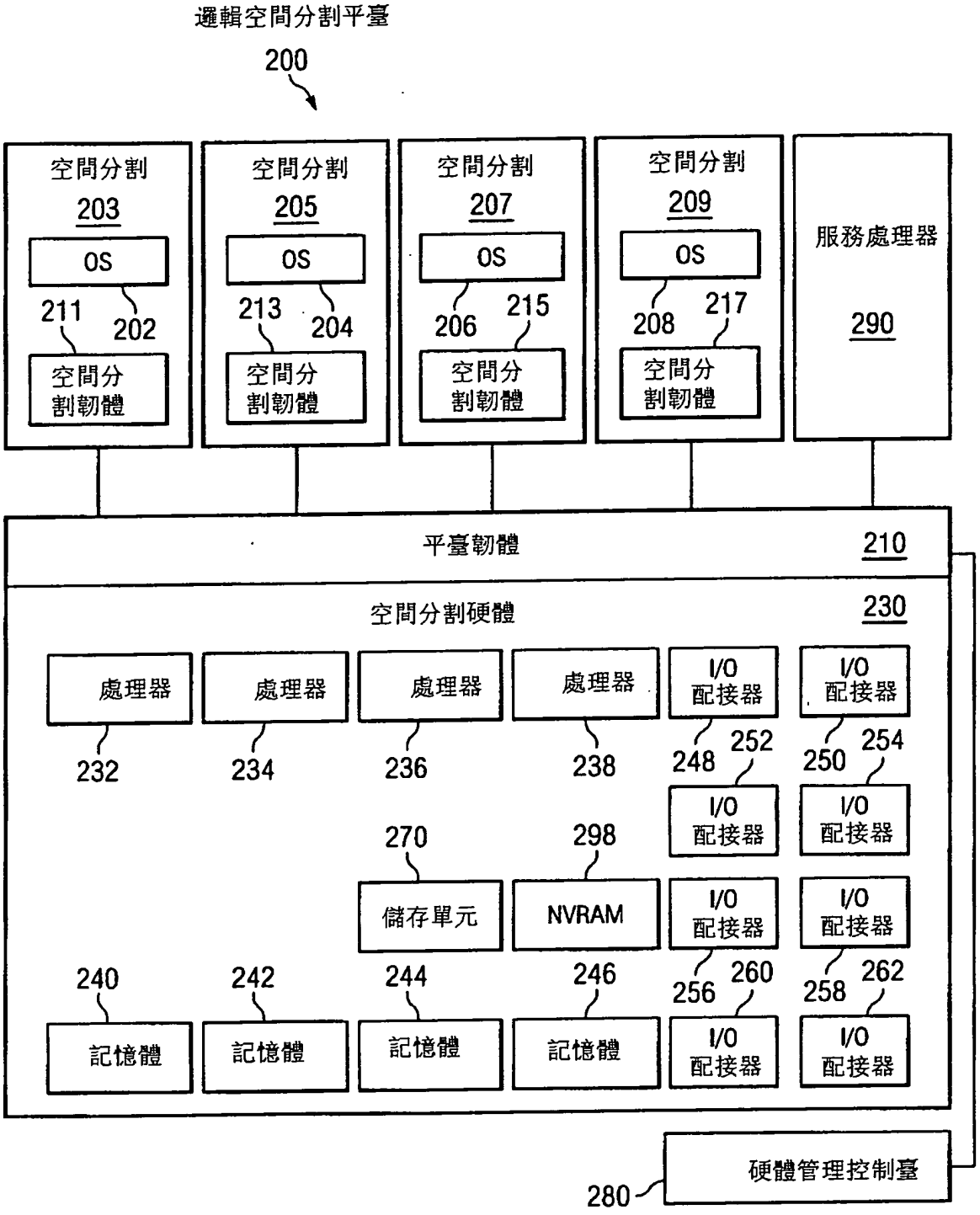
邏輯空間分割平臺
200
空間分割
203
OS
211
202
空間分割韌體
空間分割
205
OS
213
204
空間分割韌體
空間分割
207
OS
206
215
空間分割韌體
空間分割
209
OS
208
217
空間分割韌體
服務處理器
290
平臺韌體
210
空間分割硬體
230
處理器
232
處理器
234
處理器
236
處理器
238
I/O 配接器
248
252
I/O 配接器
I/O 配接器
250
254
I/O 配接器
270
儲存單元
298
NVRAM
I/O 配接器
I/O 配接器
240
記憶體
242
記憶體
244
記憶體
246
記憶體
256
260
I/O 配接器
258
262
I/O 配接器
硬體管理控制臺
280

圖 2

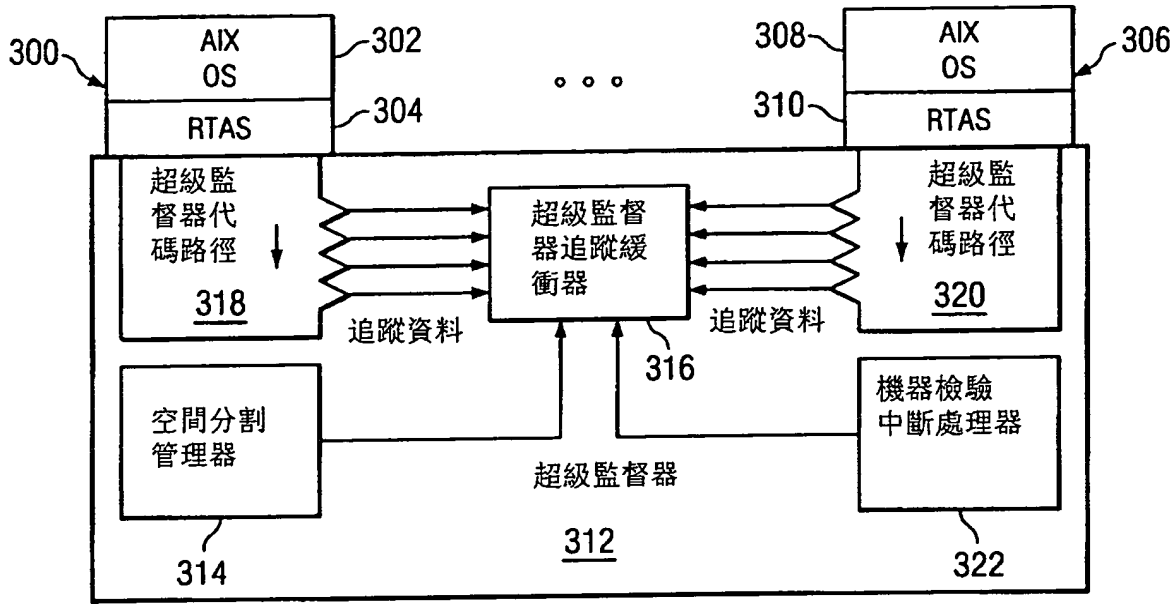

圖 3A

(先前技術)

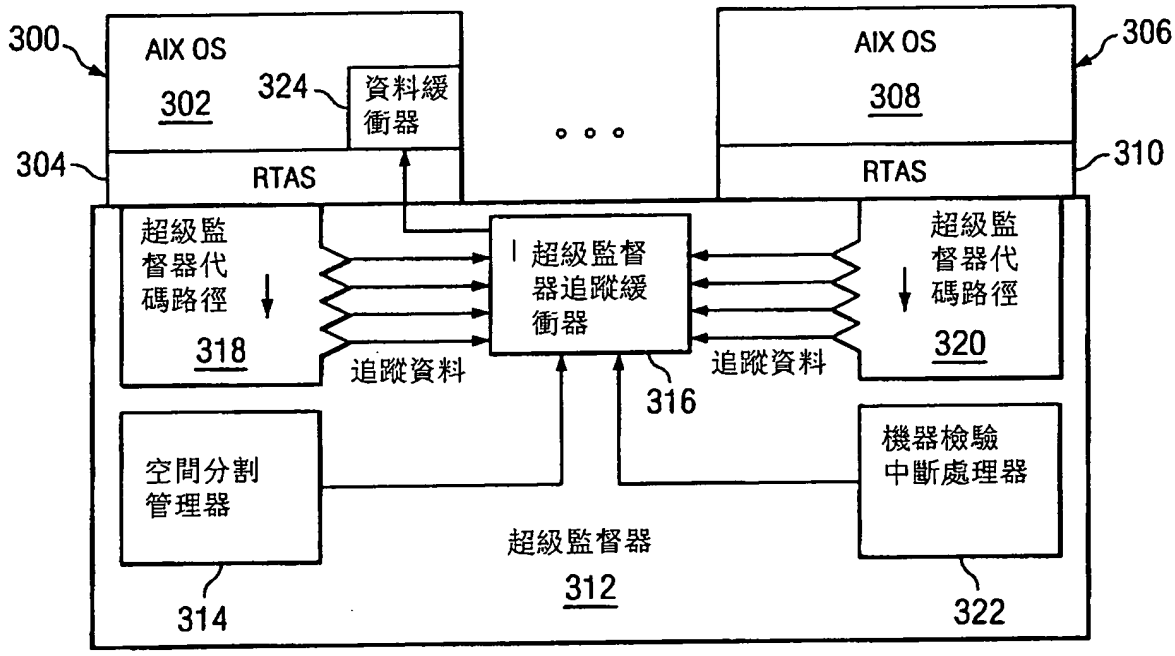

圖 3B

(先前技術)

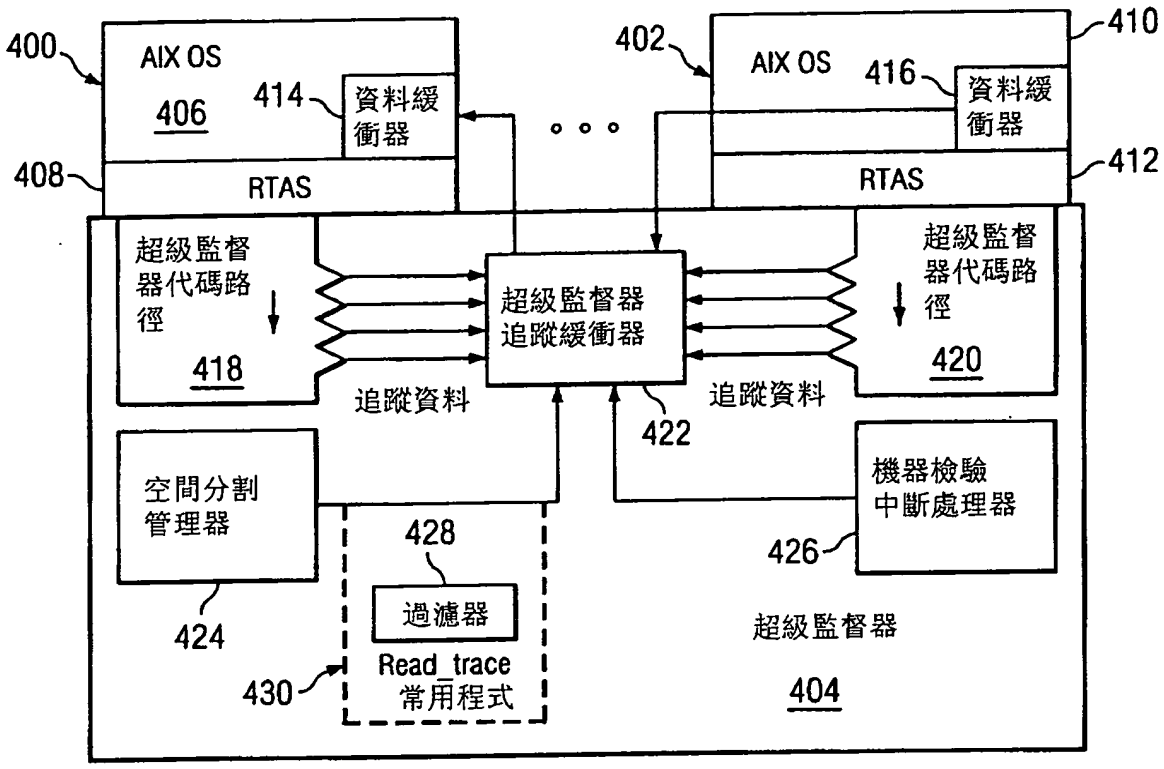
400
AIX OS
406
414
資料緩衝器
408
RTAS
402
AIX OS
416
資料緩衝器
410
412
RTAS
超級監督器代碼路徑
418
追蹤資料
超級監督器追蹤緩衝器
422
追蹤資料
超級監督器代碼路徑
420
空間分割管理器
424
428
過濾器
Read_trace
常用程式
430
機器檢驗中斷處理器
426
超級監督器
404

圖 4

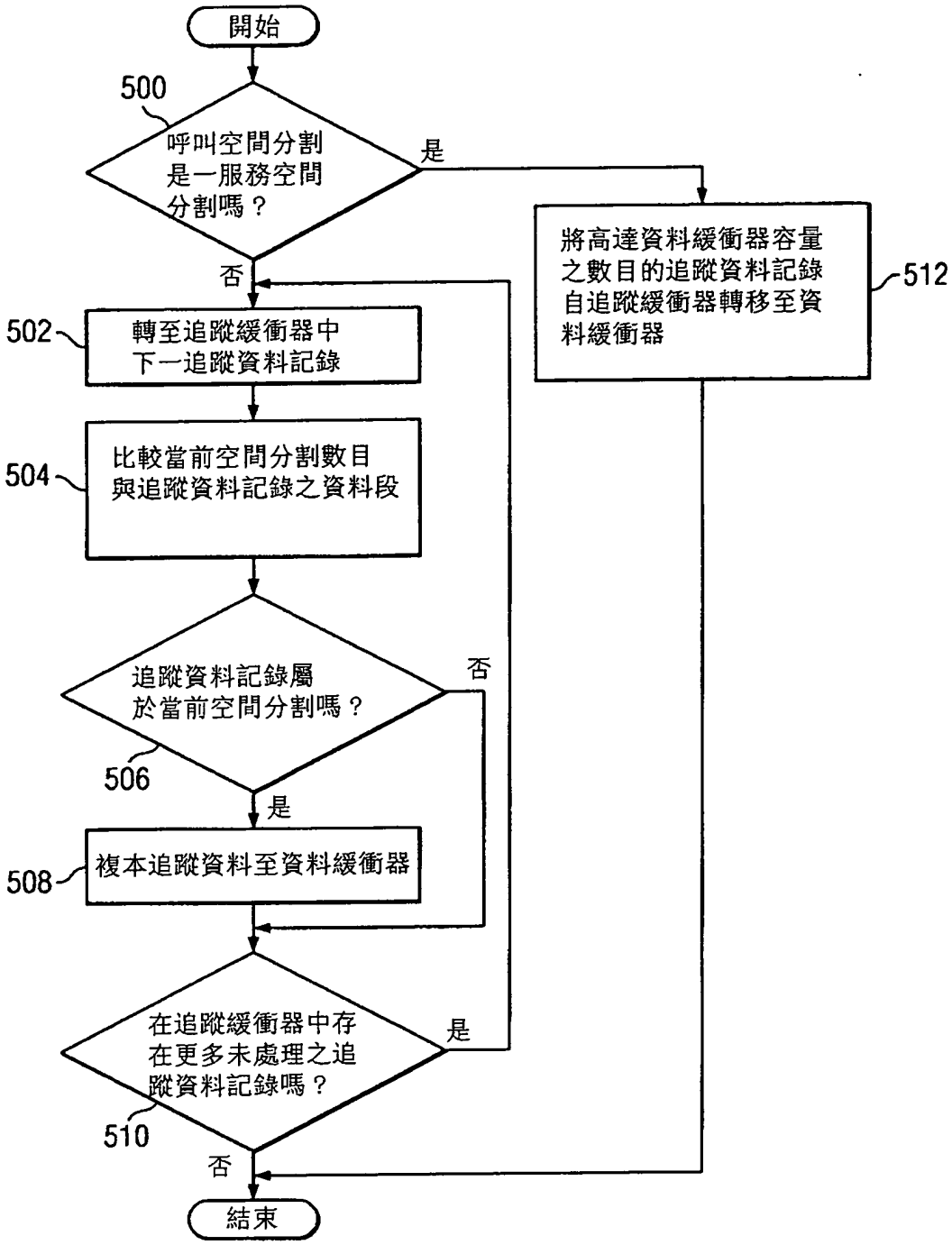
開始
500
呼叫空間分割是一服務空間分割嗎？
是
將高達資料緩衝器容量之數目的追蹤資料記錄自追蹤緩衝器轉移至資料緩衝器
512
否
502
轉至追蹤緩衝器中下一追蹤資料記錄
504
比較當前空間分割數目與追蹤資料記錄之資料段
追蹤資料記錄屬於當前空間分割嗎？
否
506
是
508
複本追蹤資料至資料緩衝器
在追蹤緩衝器中存在更多未處理之追蹤資料記錄嗎？
是
510
否
結束

圖 5

```
++++++++++++++++++++++++++++++++++++++++++++++++++++++++++
int64_t
read_trace (int64_taddr, struct io_struct*ios)
{
   /*A byte count of zero indicates that all trace data */
   /* has been transferred.
   if (trace_save_area.byte_count = = 0) {
      ios->resid = ios->xfer_length;
   }
   else if (trace_save_area.byte_count>ios->xfer_length) {
   /*There is more data to copy out than provided by   */
   /*space in the user buffer.    Thus, copy in the max */
   /*amount of data the buffer will contain.          */

  If service partition {                                          }
     Memcpy64 (void*) addr,                                       }
                 (void*) trace_save_area.current_data_ptr,        }
                  ios->xfer_length) ;                             }
                                                                  }
}else {                                                           } 602
     while  (trace data available) {                              }
        if ( trace data record belongs to current partition)      }
                 copy record to user buffer                       }
     }                                                            }
}                                                                 }

   trace_save_area.byte_count-=ios->xfer_length;
     trace_save_area.current_data_ptr =
          (void*)   ((uint64_t) trace_save_area.current_data_ptr+
          ios->xfer_length) ;
    /*There is no residual data.*/
    ios->resid = 0 ;
   }
   else{
   /*The user buffer provided can hold all remaining trace*/
   /*data.    Counters are zeroed to reflect this.          */
     memcpy64 ((void*)  addr,
          (void*) trace_save_area.current_data_ptr,
              trace_save_area.byte_count) ;
     ios->resid = ios->xfer_length - trace_save_area.byte_cou
     /* All data has been transferred so set the byte count*/
     /* to zero.
     trace_save_area.byte_count = 0 ;
}
   /*There is currently no failure reason for this routine,       *
   /*however, a return code will be passed given the future       *
   /*possibility of memcpy adding a return code.      */
```

600

圖 6

七、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：

100	資料處理系統
101、102、103、104	處理器
106	系統匯流排
108	記憶體控制器/快取記憶體
110	I/O匯流排橋接器
112	I/O匯流排
114	周邊元件互連(PCI)主機橋接器
115	PCI本地匯流排
116、124、132、142	PCI至PCI橋接器
118、119、123、126、127、131、133、141、144、145	PCI匯流排
120、121、128、129、136	PCI I/O配接器
122、130、140P	CI主機橋接器
134	JTAG/I^2C匯流排
135	服務處理器
148	圖形配接器
149	硬碟配接器

150	硬碟
160、161、162、163	本地記憶體
170、171、172、173、174、175、176	I/O插槽
190	硬體OP面板
191	本地記憶體
192	NVRAM儲存器
193	PCI/ISA橋接器
194	服務處理器郵箱介面與ISA匯流排存取通過邏輯
195	本地PCI匯流排
196	ISA匯流排

八、本案若有化學式時，請揭示最能顯示發明特徵的化學式：

（無）

I326820

97年4月30日修(更)正替換頁

發明專利說明書

中文說明書替換頁(97年4月)

公告本

（本說明書格式、順序及粗體字，請勿任意更動，※記號部分請勿填寫）

※ 申請案號：094114173

※ 申請日期：94. 05. 03　　※IPC 分類：G06F11/00(2006.01)

一、發明名稱：（中文/英文）

在一邏輯空間分割資料處理系統中用於管理追蹤資料的方法，相關系統及記錄相關電腦程式的電腦可讀取媒體

A METHOD, A SYSTEM AND A COMPUTER READABLE MEDIUM WITH A COMPUTER PROGRAM IN A LOGICAL PARTITIONED DATA PROCESSING SYSTEM FOR MANAGING TRACE DATE

二、申 請 人：（共 1 人）

姓名或名稱：（中文/英文）

美商萬國商業機器公司

INTERNATIONAL BUSINESS MACHINES CORPORATION

代表人：（中文/英文）

傑拉德　羅森賽

ROSENTHAL, GERALD

住居所或營業所地址：（中文/英文）

美國紐約州阿蒙市新果園路

NEW ORCHARD ROAD, ARMONK, NY 10504, U.S.A.

國　籍：（中文/英文）

美國　U.S.A.

三、發明人：（共 2 人）

姓　名：（中文/英文）

1.高登　D　麥尹托希

MCINTOSH, GORDON D.

2.蓋瑞　李　盧茲克

RUZEK, GARY LEE

國　籍：（中文/英文）

1.-2.均美國　U.S.A.

97年4月30日修(更)正替換頁

十、申請專利範圍：

1. 一種在一邏輯空間分割資料處理系統中用於管理追蹤資料之方法，該方法包含：

自該邏輯空間分割資料處理系統中之複數個空間分割內之一呼叫空間分割，接收一對該追蹤資料之呼叫；

識別與該呼叫空間分割相關聯之一緩衝器中之該追蹤資料，以形成經識別之追蹤資料；及

僅傳回該呼叫空間分割之該經識別之追蹤資料，其中該等複數個空間分割內之其它空間分割之該追蹤資料不被傳回至該呼叫空間分割。

2. 如請求項1之方法，其中該緩衝器係一循環緩衝器。

3. 如請求項1之方法，其進一步包含：

作為對該呼叫空間分割為一服務空間分割之回應，將由該呼叫所請求之所有該等追蹤資料傳回至一服務空間分割。

4. 如請求項1之方法，其中該傳回步驟將該經識別之追蹤資料傳回至該呼叫空間分割中之一資料緩衝器。

5. 如請求項1之方法，其中該傳回步驟包含：

識別該呼叫空間分割之一資料緩衝器中存在之一自由空間的量；及

僅傳回大小適合於該資料緩衝器之該呼叫空間分割的該經識別之追蹤資料的一部分。

6. 如請求項1之方法，其中藉由平臺韌體來執行該接收步驟、該識別步驟及該傳回步驟。

7. 如請求項1之方法，其中該呼叫包括一該呼叫空間分割之識別及一與該呼叫空間分割相關聯之一資料緩衝器之識別。

8. 一種用於管理追蹤資料之邏輯空間分割資料處理系統，該資料處理系統包含：

接收構件，其用於自該邏輯空間分割資料處理系統之複數個空間分割內之一呼叫空間分割，接收一對該追蹤資料之呼叫；

識別構件，其用於識別與該呼叫空間分割相關聯之一緩衝器中之該追蹤資料，以形成經識別之追蹤資料；及

傳回構件，其用於僅傳回該呼叫空間分割之該經識別之追蹤資料，其中該等複數個空間分割內之其它空間分割之該追蹤資料不被傳回至該呼叫空間分割。

9. 如請求項8之資料處理系統，其中該緩衝器係一循環緩衝器。

10. 如請求項8之資料處理系統，其中該傳回構件係一第一傳回構件且其進一步包含：

第二傳回構件，其用於作為對該呼叫空間分割為一服務空間分割之回應，將由該呼叫所請求之所有該等追蹤資料傳回至一服務空間分割。

11. 如請求項8之資料處理系統，其中該傳回構件傳回該經識別之追蹤資料至該呼叫空間分割中之一資料緩衝器。

12. 如請求項8之資料處理系統，其中該傳回構件包含：

用於識別該呼叫空間分割之一資料緩衝器中存在之一

自由空間的量之構件；及

用於僅傳回大小適合於該資料緩衝器之該呼叫空間分割之該經識別追蹤資料的一部分之構件。

13. 如請求項8之資料處理系統，其中該接收構件、該識別構件及該傳回構件均由平臺韌體來執行。

14. 如請求項8之資料處理系統，其中該呼叫包括一該呼叫空間分割之識別及一與該呼叫空間分割相關聯之一資料緩衝器之識別。

15. 一種記錄可以執行在一邏輯空間分割資料處理系統中用於管理追蹤資料之電腦程式的一電腦可讀取媒體，該電腦程式包含：

第一指令，其用於自該邏輯空間分割資料處理系統之複數個空間分割內之一呼叫空間分割，接收一對該追蹤資料之呼叫；

第二指令，其用於識別與該呼叫空間分割相關聯之一緩衝器中之追蹤資料，以形成經識別之追蹤資料；及

第三指令，其用於僅傳回該呼叫空間分割之該經識別之追蹤資料，其中該等複數個空間分割內之其它空間分割之該追蹤資料不被傳回至該呼叫空間分割。

16. 如請求項15之電腦可讀取媒體，其中該緩衝器係一循環緩衝器。

17. 如請求項15之電腦可讀取媒體，其進一步包含：

第四指令，其作為對該呼叫空間分割為一服務空間分割之回應，用於傳回該呼叫所請求之所有該等追蹤資料

至一服務空間分割。

18. 如請求項15之電腦可讀取媒體，其中該等第三指令傳回該經識別之追蹤資料至該呼叫空間分割中之一資料緩衝器。

19. 如請求項15之電腦可讀取媒體，其中該等第三指令包含：

第一子指令，其用於識別該呼叫空間分割之一資料緩衝器中存在之一自由空間的量；及

第二子指令，其用於僅傳回大小適合於該資料緩衝器之該呼叫空間分割之該經識別追蹤資料的一部分。

20. 如請求項15之電腦可讀取媒體，其中該等第一指令、該等第二指令及該等第三指令均由平臺韌體來執行。

21. 如請求項15之電腦可讀取媒體，其中該呼叫包括一該呼叫空間分割之識別及一與該呼叫空間分割相關聯之一資料緩衝器之識別。

22. 一種用於管理追蹤資料之邏輯空間分割資料處理系統，其包含：

一匯流排系統；

一連接至該匯流排系統之記憶體，其中該記憶體包括一組指令；及

一連接至該匯流排系統之處理單元，其中該處理單元執行一組指令以自該邏輯空間分割，資料處理系統中複數個空間分割內之一呼叫空間分割接收一對該追蹤資料之呼叫；識別與該呼叫空間分割相關聯之一緩衝器中之

該追蹤資料，以形成經識別之追蹤資料；及僅傳回該呼叫空間分割之該經識別之追蹤資料，其中該等複數個空間分割內之其它空間分割之該追蹤資料不被傳回至該呼叫空間分割。