



(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 I475379 B

(45)公告日：中華民國 104 (2015) 年 03 月 01 日

(21)申請案號：100121300

(22)申請日：中華民國 100 (2011) 年 06 月 17 日

(51)Int. Cl. : G06F11/16 (2006.01)

G06F12/06 (2006.01)

G06F3/06 (2006.01)

(30)優先權：2010/06/18 美國

61/356,443

2011/06/16 美國

61/497,525

(71)申請人：L S I 公司(美國) LSI CORPORATION (US)

美國

(72)發明人：卡尼帕 提姆帝 勞倫斯 CANEPA, TIMOTHY LAWRENCE (US)；阿姆達爾 卡

爾頓 基恩 AMDAHL, CARLTON GENE (US)

(74)代理人：陳長文

(56)參考文獻：

TW I233551

TW I274353

TW I285304

TW I286690

US 6304980B1

US 6442709B1

US 6654831B1

審查人員：何旭智

申請專利範圍項數：19 項 圖式數：5 共 111 頁

(54)名稱

提供可擴展儲存之方法、裝置及系統

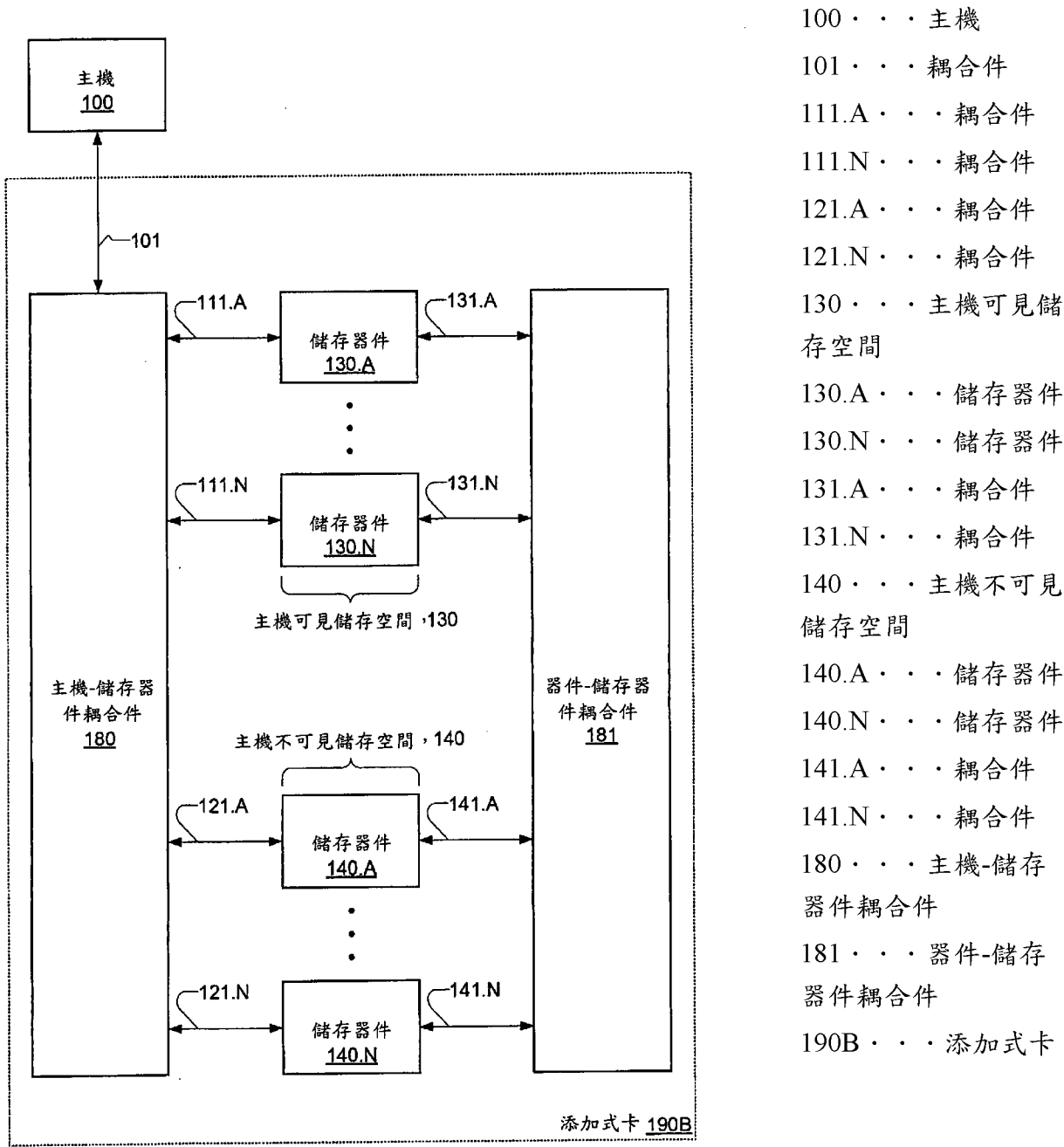
METHOD, APPARATUS AND SYSTEM FOR PROVIDING SCALABLE STORAGE

(57)摘要

使用可擴展儲存器件之技術將複數個主機可存取儲存器件表示為單個邏輯介面，從而在概念上彙總由該等器件實施之儲存空間。該等器件之一初級代理程式使用一主機介面協定自主機接受儲存請求，在內部處理該等請求及/或使用一同級間協定將該等請求作為子請求轉發至該等儲存器件之次級代理程式。該等次級代理程式接受及處理該等子請求，並將該等子請求中之每一者之子狀態資訊報告給該初級代理程式及/或該主機。該初級代理程式視情況將該等子狀態累積成一總狀態以提供至該主機。該等代理程式之間的同級間通信視情況用於在主機存取及/或故障恢復期間傳遞冗餘資訊。各種故障恢復技術重新分配儲存空間、重新指派初級/次級/可組態代理程式、經由冗餘資訊恢復資料或其任一組合。

Techniques using scalable storage devices represent a plurality of host-accessible storage devices as a single logical interface, conceptually aggregating storage implemented by the devices. A primary agent of the devices accepts storage requests from the host using a host-interface protocol, processing the requests internally and/or forwarding the requests as sub-requests to secondary agents of the storage devices using a peer-to-peer protocol. The secondary agents accept and process the sub-requests, and report sub-status information for each of the sub-requests to the primary agent and/or the host. The primary agent optionally accumulates the sub-statuses into an overall status for providing to the host. Peer-to-peer communication between the agents is optionally used to communicate redundancy information during host accesses and/or

failure recoveries. Various failure recovery techniques reallocate storage, reassign primary/secondary/configurable agents, recover data via redundancy information, or any combination thereof.



- 100 . . . 主機
- 101 . . . 耦合件
- 111.A . . . 耦合件
- 111.N . . . 耦合件
- 121.A . . . 耦合件
- 121.N . . . 耦合件
- 130 . . . 主機可見儲存空間
- 130.A . . . 儲存器件
- 130.N . . . 儲存器件
- 131.A . . . 耦合件
- 131.N . . . 耦合件
- 140 . . . 主機不可見儲存空間
- 140.A . . . 儲存器件
- 140.N . . . 儲存器件
- 141.A . . . 耦合件
- 141.N . . . 耦合件
- 180 . . . 主機-儲存器件耦合件
- 181 . . . 器件-儲存器件耦合件
- 190B . . . 添加式卡

圖 1B

發明專利說明書

中文說明書替換頁(103年8月)

(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：100121300

※ 申請日：100年6月17日

※IPC 分類：G06F 11/16 (2006.01)

G06F 12/06 (2006.01)

G06F 3/06 (2006.01)

一、發明名稱：(中文/英文)

提供可擴展儲存之方法、裝置及系統

METHOD, APPARATUS AND SYSTEM FOR PROVIDING SCALABLE
STORAGE

二、中文發明摘要：

使用可擴展儲存器件之技術將複數個主機可存取儲存器件表示為單個邏輯介面，從而在概念上彙總由該等器件實施之儲存空間。該等器件之一初級代理程式使用一主機介面協定自主機接受儲存請求，在內部處理該等請求及/或使用一同級間協定將該等請求作為子請求轉發至該等儲存器件之次級代理程式。該等次級代理程式接受及處理該等子請求，並將該等子請求中之每一者之子狀態資訊報告給該初級代理程式及/或該主機。該初級代理程式視情況將該等子狀態累積成一總狀態以提供至該主機。該等代理程式之間的同級間通信視情況用於在主機存取及/或故障恢復期間傳遞冗餘資訊。各種故障恢復技術重新分配儲存空間、重新指派初級/次級/可組態代理程式、經由冗餘資訊恢復資料或其任一組合。

103年 8月 19日修(受)正替換頁

三、英文發明摘要：

Techniques using scalable storage devices represent a plurality of host-accessible storage devices as a single logical interface, conceptually aggregating storage implemented by the devices. A primary agent of the devices accepts storage requests from the host using a host-interface protocol, processing the requests internally and/or forwarding the requests as sub-requests to secondary agents of the storage devices using a peer-to-peer protocol. The secondary agents accept and process the sub-requests, and report sub-status information for each of the sub-requests to the primary agent and/or the host. The primary agent optionally accumulates the sub-statuses into an overall status for providing to the host. Peer-to-peer communication between the agents is optionally used to communicate redundancy information during host accesses and/or failure recoveries. Various failure recovery techniques reallocate storage, reassign primary/secondary/configurable agents, recover data via redundancy information, or any combination thereof.

四、指定代表圖：

(一)本案指定代表圖為：第(1B)圖。

(二)本代表圖之元件符號簡單說明：

100	主機
101	耦合件
111.A	耦合件
111.N	耦合件
121.A	耦合件
121.N	耦合件
130	主機可見儲存空間
130.A	儲存器件
130.N	儲存器件
131.A	耦合件
131.N	耦合件
140	主機不可見儲存空間
140.A	儲存器件
140.N	儲存器件
141.A	耦合件
141.N	耦合件
180	主機-儲存器件耦合件
181	器件-儲存器件耦合件
190B	添加式卡

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】

領域：需要儲存器件之存取之進步以提供效能、效率及使用效用之改良。

【先前技術】

相關領域：除非明確識別為公眾習知或眾所周知，否則本文提及之出於上下文、定義或比較目的而包含之技術及概念不應視為承認此等技術及概念係先前眾所周知或另外作為先前技術之一部分。本文所引用之包含專利、專利申請案及公開案之所有參考(若存在)皆出於各種目的以全文引用之方式明確地或不明確地併入本文中。

【發明內容】

本發明可以多種方式實施，包含作為一製程、一製造物件、一裝置、一系統、一物質組成及諸如一電腦可讀儲存媒體之一電腦可讀媒體(例如，一光學/磁性大容量儲存器件中之媒體，諸如一碟，或具有諸如快閃儲存器之非揮發性儲存器之一積體電路)或其中在光學或電子通信鏈路上發送程式指令之一電腦網路。在此說明書中，此等實施方案或本發明可採取之任何其他形式可稱為技術。[實施方式]提供本發明之實現在上文所識別領域中之效能、效率及使用效用之改良的一或多個實施例之一說明。[實施方式]包含一導論以促進對[實施方式]之剩餘部分之更快速理解。該導論包含根據本文所述概念之一或多個系統、方法、製造物件及電腦可讀媒體之實例性實施例。如在總論

中更詳細地論述，本發明涵蓋在所發佈之申請專利範圍之範疇內的所有可能修改及變化。

【實施方式】

結合圖解說明本發明之選定細節之附圖在下文中提供本發明之一或多個實施例之一詳細說明。本發明係與該等實施例一起闡述。本文之實施例應理解為僅係實例性的，本發明明確地不限於或不受限於任何或所有本文實施例，且本發明涵蓋各種替代、修改及等效物。為避免說明之單調性，可將各種文書標籤(包含但不限於：第一、最後、某一(些)、各種、進一步、其他、特定、選擇、某一(些)及顯著的)應用於單獨組之實施例；如本文所使用，此等標籤明確地不意欲表達品質或任一形式之偏好或偏見，而僅在該等單獨組間便利地加以區分。所揭示過程之某些操作之次序在本發明之範疇內可變更。儘管多個實施例起到闡述製程、方法及/或程式指令特徵中之變化之作用，但亦涵蓋根據一預定準則或一動態判定之準則執行分別對應於複數個該多個實施例之複數個操作模式中之一者之靜態及/或動態選擇之其他實施例。在下列說明中列舉各種具體細節以提供對本發明之一透徹理解。出於實例性目的提供該等細節，且可在不具有該等細節中之某些或全部之情況下根據申請權利範圍來實行本發明。出於簡要之目的，並未詳細闡述關於本發明之技術領域所習知之技術材料，以免不必要地使本發明難以理解。

導論

本文僅出於促進更快速地理解[實施方式]之目的而包含此導論；本發明不限於該導論中所提供之概念(若存在，包含明確實例)，乃因任何導論之段落係整個主題之示意圖，且不意欲係窮盡性或約束性說明。舉例而言，下文之導論提供由空間所限之概略資訊及對僅某些實施例之組織。貫穿該說明書均衡地存在諸多其他實施例，包含申請專利範圍將涉及之彼等實施例。

縮寫字

本文所界定之各種速記縮寫或縮寫字之至少某些係關於本文中使用的某些元件。

縮寫字	說明
AHCI	先進主機控制器介面
CF	緊密快閃
DIF	資料完整性欄位
DIX	資料完整性擴展
DMA	直接記憶體存取
eNVMHCI	企業非揮發性記憶體主機控制器介面
eSATA	外部串列先進技術附件
IC	積體電路
IDE	積體驅動電子器件
IO	輸入/輸出
JBOD	簡單磁碟綁定
LBA	邏輯區塊位址
MMC	多媒體卡
MsgD	具有資料有效負載之訊息請求
PC	個人電腦
PCIe	高速周邊組件互連(高速PCI)
PDA	個人數位助理

縮寫字	說明
RAID	廉價/獨立磁碟冗餘陣列
RMW	讀取-修改-寫入
ROM	唯讀記憶體
SAS	串列附接小電腦系統介面(串列SCSI)
SATA	串列先進技術附件(串列ATA)
SCSI	小電腦系統介面
SD	安全數位
SSD	固態磁碟/驅動器
USB	通用串列匯流排

使用可擴展儲存器件之技術將複數個主機可存取儲存器件表示為單個邏輯介面，從而在概念上彙總由該等器件實施之儲存空間。該等器件之一初級代理程式使用一主機介面協定自主機接受儲存請求，在內部處理該等請求及/或使用一同級間協定將該等請求作為子請求轉發至儲存器件之次級代理程式。該等次級代理程式接受及處理該等子請求，並將該等子請求中之每一者之子狀態資訊報告給該初級代理程式及/或主機。該初級代理程式視情況將該等子狀態累積成一總狀態以提供至該主機。該初級代理程式將可用儲存空間報告給主機，該可用儲存空間包含由初級代理程式實施之儲存空間以及由次級代理程式實施及分配以彙總之儲存空間。次級代理程式將零可用儲存空間或在該彙總分配之後剩餘的儲存空間報告給主機。

該等代理程式之間的同級間通信視情況用於在主機存取及/或故障恢復期間傳遞冗餘資訊。各種故障恢復技術重新分配儲存空間、重新指派初級/次級/可組態代理程式、

經由冗餘資訊恢復資料或其任一組合。

PCIe介面具有與一系統中之一主機之一固有單數關係。每一PCIe介面表示對一器件之一功能(該功能係儲存、網路連結或某一其他計算系統能力)之單個入口。該系統中之每一PCIe器件表示該系統中之一各別相異實體。然而，透過同級間訊務，將數個個別PCIe器件一起編組以使得該等器件能夠對主機顯現為單個PCIe器件。藉由利用PCIe同級間能力，藉由經由同級間通信將複數個PCIe儲存器件一起接合及/或編組以使得該等經接合/編組器件對主機顯現為單個儲存子系統來實現一可擴展儲存架構。在某些實施例中，僅基於LBA在該等PCIe儲存器件間散佈請求。該唯LBA散佈不要求該群組中之任一特定器件之儲存空間具有為該特定器件路由請求之功能，因此作為單個故障點排除該儲存空間。

在某些實施例及/或使用情境中，該接合/編組使得效能及/或容量增加，同時對主機顯現為單個邏輯介面。在某些實施例及/或使用情境中，經由基於LBA之請求路由排除所有單個儲存空間故障點。在某些實施例及/或使用情境中，該接合/編組跨越若干個儲存器件散佈請求負載，因此對一主機透明地擴展容量及效能。該容量及效能擴展對主機係透明的意指主機不需要控制且不要求觀察該接合/編組而非容量及延遲改變。在某些實施例及/或使用情境中，該接合/編組提供用於一對稱多處理器儲存器複合體。

在某些實施例及/或使用情境中，該接合/編組使得一主機能夠作為單個邏輯器件存取散佈於複數個實體器件上之資源之一彙總。舉例而言，該接合/編組使得一主機能夠作為單個邏輯儲存器件存取散佈於複數個實體儲存器件上之複數個儲存範圍之一彙總。在作為該請求之部分轉發如主機所提供之一LBA(及視情況，一長度)之後，保留該LBA(及選用長度)資訊。保留該LBA(及選用長度)資訊並在一目標器件上進行轉譯使得維持準確的完整性元資料(諸如DIV/DIX參考完整性)及/或實現器件之獨立(例如，並行)操作。

在某些實施例及/或使用情境中，該接合/編組實現相對於儲存容量之透明冗餘及/或恢復，諸如各種RAID、鏡像及失效切換實施方案。在各種實施例中，主機訊務主要係資料訊務，而同級間訊務主要係控制訊務。在各種實施例中，同級間訊務包含資料(例如，在一RAID恢復操作期間或在累積同位資訊之輸送期間)。

在各種實施例及/或使用情境中，該接合/編組啟用具有耦合至SSD之一主機之系統，其中耦合至該主機之一或多個PCIe鏈路(各自具有一或多個巷道)之帶寬係跨越複數個SSD、複數個快閃控制器(諸如用於SSD中)及/或複數個快閃晶片散佈(對主機透明地)。舉例而言，1個主機耦合式PCIe鏈路(例如，具有8個巷道)係跨越4個SSD或(另一選擇係)8個快閃控制器(每一SSD或快閃控制器具有一個PCIe鏈路，一個PCIe鏈路具有(例如)少於8個巷道)散佈。針對另

一實例，4個主機耦合式PCIe鏈路之帶寬係跨越64個SSD或(另一選擇係)256個快閃控制器散佈。針對另一實例，1個主機耦合式PCIe鏈路之帶寬係跨越32個快閃晶片或(另一選擇係)跨越64個快閃晶片散佈。針對另一實例，4個主機耦合式PCIe鏈路之帶寬係跨越128個快閃晶片或(另一選擇係)跨越256個快閃晶片散佈。

在各種實施例中，一初級代理程式判定主機至次級代理程式儲存映射(例如，哪些主機LBA範圍對應於哪些次級代理程式)。舉例而言，主機LBA跨越一預定大小(例如，64 KB)之實體驅動器等量化。針對另一實例，根據一RAID實施方案(例如，RAID 0、RAID 1或RAID 5)主機LBA等量化。針對再一實例，等量化係根據工作負載動態改變(例如，針對一第一工作負載使用一第一等量化，且針對一第二工作負載使用一第二等量化)。

在某些實施例中，經由一匯流排拓撲耦合一群組之PCIe儲存器件(可自一或多個主機存取)以實現彼此之間的同級間通信。該群組之PCIe儲存器件經組態以使得該群組中之僅一個PCIe器件接收一特定主機邏輯連接之請求。在某些實施例中，存在對該群組之複數個連接，實現不同主機邏輯連接以供應由該群組中之不同PCIe儲存器件接收之請求。

該群組之PCIe儲存器件之組態相依於如何使該等器件曝露至主機。若該群組中之所有器件直接曝露至主機，例如經由一透明開關，則除了實現主機邏輯連接之器件外的所

有器件將零儲存容量報告回至該(等)主機，而實現主機邏輯連接之該等器件針對一特定主機邏輯連接報告該群組之一經組態容量。該經組態之容量相依於如何跨越主機邏輯連接分割該群組之儲存空間，以及該儲存空間如何組態或組織(諸如JBOD或RAID)。若該等器件中之某些器件在例如一橋接器或一不透明開關後面自該主機隱藏，則不由不曝露至該主機之該等器件執行該零容量報告。一主機邏輯連接僅將請求直接發送至單個PCIe器件介面，然後該單個PCIe器件介面將請求分配及協調至該群組中之其他器件。

在進一步實施例中，實現各別主機邏輯連接之兩個或兩個以上器件報告在該等器件及/或其他器件之儲存容量內重疊之儲存容量。舉例而言，實現一第一主機邏輯連接之一第一器件報告等於該第一器件之總儲存容量與一第二器件之總儲存容量之一彙總儲存容量。實現一第二主機邏輯連接之一第三器件亦報告該彙總儲存容量。經由該第一主機邏輯連接接收或經由該第二主機邏輯連接接收之主機請求參考第一器件及第二器件之相同實體儲存空間(視情況及/或可選地使用不同LBA)。

由一儲存器件實施之儲存空間不需要整個地直接曝露(或不曝露)至一主機。舉例而言，一特定儲存器件實施部分地分配至一群組之儲存器件(例如，不直接曝露至一主機)且部分地分配以供直接曝露至主機之儲存空間。該特定儲存器件報告對應於直接曝露至主機之該部分分配之一儲存空間量，而非將零容量報告給主機。

在某些實施例中，由駐存於該(等)主機上及/或在一或多個選項ROM中之軟體執行該群組之器件之初始組態。該初始組態之一實例係哪些器件欲啟用主機邏輯連接而哪些不欲啟用。該初始組態之另一實例係將多少儲存空間分配至一群組之器件對直接曝露至一主機。

在某些實施例中，該群組之器件之組態係在一系統操作時執行，諸如在插入一「備用磁碟(hot spare)」時或在自該群組動態地(例如，回應於一故障)添加或移除一器件時。舉例而言，在插入一備用磁碟器件時，涉及一主機及一初級代理程式上之一器件驅動器之一或多者之一發現過程辨識所插入之器件，並相應地組態所插入之器件(諸如替換一故障器件)。該器件之插入係經由(舉例而言)來自所插入器件及/或一「即插即用」事件之一宣告來指示。

在一主機發佈一儲存空間讀取/寫入請求(諸如一eNVMHCI或AHCI相容請求)時，接收該請求之器件使用該請求之一LBA(及視情況，長度)以將該請求轉發至實施所請求之儲存空間之該群組之器件。端視該請求之長度及該群組之組織(例如，JBOD或RAID)，將該請求(例如)作為複數個子請求路由至該群組之多於一個器件。

根據各種實施例，在該群組中之一特定器件接收一子請求時，該特定器件將該LBA(及視情況，任一供應之長度)轉譯至一本端LBA(及長度)，處理該子請求且然後將該子請求之資料直接傳送至/自主機記憶體。主機記憶體之一實例係一元件，該元件經啟用以經由執行將資料儲存至記

憶體在一特定位址處之一指令而保持由該主機寫入之資料，且經啟用以經由執行自記憶體在該特定位址處讀取資料之一指令使該主機讀取所寫入之資料。該特定器件視情況及/或可選地獲得將資料自以下各項中之一或多者傳送至之一主機記憶體位址：(a)自(所轉發之)子請求，(b)自主機駐存轉譯結構(諸如分散/集中列表)，及(c)藉由將代理請求發佈至充當對主機之一控制介面之一器件。在某些實施例中，代理請求之發佈僅係針對IO虛擬化執行。

在每一子請求之資料傳送完成時，該傳送器件將子狀態報告給充當對主機之控制介面之器件。充當對主機之控制介面之該器件隨後在一請求中所涉及之所有器件皆已報告子狀態時將狀態公佈至主機。

類似地處理非資料命令(舉例而言，一SATA識別器件命令)。在充當對一主機之一控制介面之一特定器件接收一非資料命令時，該特定器件詢問該群組中之其他器件(若必要)，且然後將該詢問之摘要及/或彙總結果報告回至該主機。

在各種實施例中，一主機將複數個器件(諸如各自具有一x1或x2之PCIe連接之器件)之一彙總視為涵蓋所有該等器件之儲存空間之單個邏輯器件(例如，該等器件係虛擬化為該單個邏輯器件)。在各種實施例中，該主機將任何或所有該等器件視為具有總儲存空間之部分，除資料訊務外亦實現分離控制訊務。

藉由(例如)該主機知曉之器件中之一特定者處理經由一

提交佇列(諸如，用於並行活動串流之複數個同時作用中之提交佇列中之一個提交佇列)來自主機之一請求。視情況及/或選擇性地將該提交佇列中之一個別請求作為一子請求轉發(經由來自該特定器件之同級間請求)至實施該請求所參考之儲存空間中之全部或部分儲存空間之器件。舉例而言，該儲存空間藉由個別LBA或LBA群組在該等器件間等量化。在某些實施例及/或使用情境中，在LBA群組中之等量化經由較大資料塊至個別器件之彙總實現效能裨益。在具有複數個提交佇列之某些實施例中，主機通知該等器件之一特定者新的提交佇列，且該特定器件將該等新提交佇列之服務分配至該等器件之其他提交佇列。

該等個別器件處理所轉發之子請求，包含至/自該主機獨立傳送之資料。某些協定(諸如某些PCIe相容協定)使用經由每一提交佇列之一各別完成佇列提供之「完成」。針對一提交佇列中之每一條目，將一對應完成通知發送回至該各別完成佇列。該等器件中之任何一或多者經視情況及/或可選地啟用以針對一既定請求「彙總」完成資訊及更新完成佇列。在各種實施例中，每一完成佇列指派一個器件以啟用原子性。在某些實施例中，所指派器件係處理對應提交佇列中之實體之相同器件，而在其他實施例中，所指派器件不同於處理對應提交佇列中之實體之器件。

實例性實施例

在結束該導論至實施方式後，接下來是實例性實施例之一集合，包含至少某些明確地列舉為「EC(實例組合)」之

實例性實施例，提供根據本文所述概念之各種實施例類型之額外說明；此等實例不意欲係彼此排外性、窮盡性或約束性的；且本發明不限於此等實例性實施例，而是涵蓋在所發佈申請專利範圍之範疇內的所有可能的修改及變化。

EC1) 一種系統，其包括：

用於接受來自一主機之一請求以存取一位址處之儲存空間之構件；

用於至少部分地基於該位址判定該請求對應於複數個儲存器件中之哪一或多者之構件；

用於至少部分地基於該位址判定對應於該等所判定儲存器件之一各別子請求之構件；

用於將該等子請求發送至該等對應之所判定儲存器件之構件；

用於自該等所判定儲存器件接受各別子狀態之構件；

用於至少部分地基於該等各別子狀態中之每一者判定一總狀態之構件；

用於將該總狀態提供至該主機之構件；

其中該用於接受該請求之構件與一主機介面協定相容，且該用於發送該等子請求之構件及該用於接受該等子狀態之構件與可獨立於該主機操作之一同級間協定相容；且其中該等儲存器件中之至少一者實施可經由該等子請求中之一者存取之儲存空間，且該等儲存器件中之該至少一者停用經由該主機介面協定對該所實施儲存空間之存取。

EC2) 如 EC1 之系統，其中該用於接受一請求之構件、該用於發送該等子請求之構件、該用於接受各別子狀態之構件及該用於提供該總狀態之構件中之兩個或兩個以上可在一共用實體介面上操作。

EC3) 如 EC2 之系統，其中一高速周邊組件互連(PCIe)拓撲與一無限帶寬拓撲中之至少一者構成該共用實體介面。

EC4) 如 EC2 之系統，其中一高速周邊組件互連(PCIe)介面與一無限帶寬介面中之至少一者構成該共用實體介面。

EC5) 如 EC1 之系統，其進一步包括該等儲存器件之特定一者，其包括該用於接受該請求之構件、該用於判定該等儲存器件之構件、該用於判定該等子請求之構件、該用於發送該等子請求之構件、該用於接受該等子狀態之構件、該用於判定該總狀態之構件及該用於提供該總狀態之構件。

EC6) 如 EC1 之系統，其中該等儲存器件中之兩個或兩個以上經啟用以經由該主機介面協定與該主機通信。

EC7) 如 EC6 之系統，其中該等儲存器件中之該兩個或兩個以上皆係該等儲存器件。

EC8) 如 EC1 之系統，其中該位址包括一開始位置及一長度。

EC9) 如 EC1 之系統，其中該位址係一邏輯區塊位址(LBA)。

EC10) 如 EC1 之系統，其中該等構件係在單個積體電路(IC)中一起實施。

EC11) 如 EC1 之系統，其中該等構件係在單個添加式卡中一起實施。

EC12) 如 EC1 之系統，其中該等構件係包括於一固態磁碟(SSD)中。

EC13) 如 EC1 之系統，其中該主機介面協定與以下各項中之一或多者相容

- 一通用串列匯流排(USB)介面標準，
- 一緊密快閃(CF)介面標準，
- 一多媒體卡(MMC)介面標準，
- 一安全數位(SD)介面標準，
- 一記憶棒介面標準，
- 一xD圖片卡介面標準，
- 一積體驅動電子器件(IDE)介面標準，
- 一串列先進技術附件(SATA)介面標準，
- 一外部 SATA(eSATA)介面標準，
- 一小電腦系統介面(SCSI)介面標準，
- 一串列附接小電腦系統介面(SAS)介面標準，
- 一光纖通道介面標準，
- 一乙太網路介面標準，及
- 一高速周邊組件互連(PCIe)介面標準。

EC14) 如 EC1 之系統，其進一步包括該主機之所有或任何部分。

EC15) 如 EC14 之系統，其中該主機包括以下各項中之一或多者

- 一 電腦，
- 一 工作站電腦，
- 一 伺服器電腦，
- 一 儲存伺服器，
- 一 個人電腦(PC)，
- 一 膝上型電腦，
- 一 筆記型電腦，
- 一 迷你筆記型電腦，
- 一 個人數位助理(PDA)，
- 一 媒體播放器，
- 一 媒體錄製器，
- 一 數位相機，
- 一 蜂巢式手機，
- 一 無線手持電話，及
- 一 電子遊戲。

EC16) 如EC1之系統，其進一步包括該等儲存器件之所有或任何部分。

EC17) 如EC1之系統，其進一步包括在該等儲存器件中之一或更多者中包括之快閃記憶體。

EC18) 一種電腦可讀媒體，其具有儲存於其中之一組指令，該組指令在由一處理元件執行時致使該處理元件執行包括以下各項之操作：

管理接受來自一主機之一請求以存取在一位址處之儲存空間；

管理至少部分地基於該位址判定該請求對應於複數個儲存器件中之哪一或多者；

管理至少部分地基於該位址判定對應於該等所判定儲存器件中之每一者之一各別子請求；

管理將該等子請求發送至該等對應判定之儲存器件；

管理自該等所判定之儲存器件中之每一者接受各別子狀態；

管理至少部分地基於該等各別子狀態中之每一者判定一總狀態；

管理將該總狀態提供至該主機；

其中該接受該請求係經由一主機介面協定，且該發送該等子請求及該接受該等子狀態係經由可獨立於該主機操作之一同級間協定；且

其中該等儲存器件中之至少一者實施可經由該等子請求中之一者存取之儲存空間，且該等儲存器件中之該至少一者停用經由該主機介面協定對該所實施儲存空間之存取。

EC19) 如 EC18 之電腦可讀媒體，其中該主機介面及該等同級間協定與一高速周邊組件互連(PCIe)通道相容。

EC20) 一種方法，其包括：

接受來自一主機之一請求以存取在一位址處之儲存空間；

至少部分地基於該位址判定該請求對應於複數個儲存器件中之哪一或多者；

至少部分地基於該位址判定對應於該等所判定儲存器件中

之每一者之一各別子請求；

將該等子請求發送至該等對應之所判定儲存器件；

自該等所判定儲存器件中之每一者接受各別子狀態；

至少部分地基於該等各別子狀態中之每一者判定一總狀態；

將該總狀態提供至該主機；

其中該接受該請求係與一主機介面協定相容，且該發送該等子請求及該接受該等子狀態係與可獨立於該主機操作之一同級間協定相容；且

其中該等儲存器件中之至少一者實施可經由該等子請求中之一者存取之儲存空間，且該等儲存器件中之該至少一者停用經由該主機介面協定對該所實施儲存空間之存取。

EC21) 一種系統，其包括：

主機介面邏輯硬體，其經啟用以接受來自一主機之一請求以存取在一位址處之儲存空間；

控制邏輯硬體，其經啟用以

至少部分地基於該位址判定該請求對應於複數個儲存器件中之哪一或多者及對應於該等所判定儲存器件中之每一者之一各別子請求，

將該等子請求發送至該等對應之所判定儲存器件，

自該等所判定之儲存器件中之每一者接受各別子狀態，及至少部分地基於該等各別子狀態中之每一者判定一總狀態；

其中該主機介面邏輯硬體進一步經啟用以將該總狀態提供至該主機；

其中該接受該請求係與一主機介面協定相容，且該發送該等子請求及該接受該等子狀態係與可獨立於該主機操作之一同級間協定相容；且

其中該等儲存器件中之至少一者實施可經由該等子請求中之一者存取之儲存空間，且該等儲存器件中之該至少一者停用經由該主機介面協定對該所實施儲存空間之存取。

EC22) 一種方法，其包括：

一第一儲存器件自複數個主機請求佇列中之一第一者讀取一第一請求；

一第二儲存器件自該等主機請求佇列中之一第二者讀取一第二請求，該第二主機請求佇列與該第一主機請求佇列相異，該第二儲存器件之該讀取獨立於該第一儲存器件之該讀取；

該第一儲存器件將該第一請求之至少一部分轉發至該第二儲存器件；

該第二儲存器件將該第二請求之至少一部分轉發至該第一儲存器件，該第二儲存器件之該轉發獨立於該第一儲存器件之該轉發；

該第一儲存器件根據該第二請求之至少該部分存取該第一儲存器件之第一儲存器，並至少部分地基於該第一儲存器件之該存取將一第一子狀態發送至該第二儲存器件；及

該第二儲存器件根據該第一請求之至少該部分存取該第二儲存器件之第二儲存器，並至少部分地基於該第二儲存器件之該存取將一第二子狀態發送至該第一儲存器件。

EC23) 一種系統，其包括：

用於使一第一儲存器件自複數個主機請求佇列中之一第一者讀取一第一請求之構件；

用於使一第二儲存器件自該等主機請求佇列之一第二者讀取一第二請求之構件，該第二主機請求佇列與該第一主機請求佇列相異，該第二儲存器件之該讀取獨立於該第一儲存器件之該讀取；

用於使該第一儲存器件將該第一請求之至少一部分轉發至該第二儲存器件之構件；

用於使該第二儲存器件將該第二請求之至少一部分轉發至該第一儲存器件之構件，該第二儲存器件之該轉發獨立於該第一儲存器件之該轉發；

用於使該第一儲存器件根據該第二請求之至少該部分存取該第一儲存器件之第一儲存器並至少部分地基於該第一儲存器件之該存取將一第一子狀態發送至該第二儲存器件之構件；及

用於使該第二儲存器件根據該第一請求之至少該部分存取該第二儲存器件之第二儲存器並至少部分地基於該第二儲存器件之該存取將一第二子狀態發送至該第一儲存器件之構件。

EC24) 一種系統，其包括：

- 一 第一儲存器件之主機介面邏輯硬體，其經啟用以自複數個主機請求佇列中之一第一者讀取一第一請求；
- 一 第二儲存器件之主機介面邏輯硬體，其經啟用以自該等主機請求佇列之一第二者讀取一第二請求，該第二主機請求佇列與該第一主機請求佇列相異，該第二儲存器件之該讀取獨立於該第一儲存器件之該讀取；
- 該第一儲存器件之請求轉發邏輯硬體，其經啟用以將該第一請求之至少一部分轉發至該第二儲存器件；
- 該第二儲存器件之請求轉發邏輯硬體，其經啟用以將該第二請求之至少一部分轉發至該第一儲存器件，該第二儲存器件之該轉發獨立於該第一儲存器件之該轉發；
- 該第一儲存器件之儲存介面邏輯硬體，其經啟用以根據該第二請求之至少該部分存取該第一儲存器件之第一儲存器；
- 該第一儲存器件之子狀態產生邏輯硬體，其經啟用以至少部分地基於該第一儲存器件之該存取判定一第一子狀態；
- 該第一儲存器件之子狀態轉發邏輯硬體，其經啟用以將該第一子狀態轉發至該第二儲存器件；
- 該第二儲存器件之儲存介面邏輯硬體，其經啟用以根據該第一請求之至少該部分存取該第二儲存器件之第二儲存器；
- 該第二儲存器件之子狀態產生邏輯硬體，其經啟用以至少部分地基於該第二儲存器件之該存取判定一第二子狀態；

態；及

該第二儲存器件之子狀態轉發邏輯硬體，其經啟用以將該第二子狀態轉發至該第一儲存器件。

EC25) 一種方法，其包括：

一第一儲存器件自複數個主機請求佇列中之一第一者讀取一第一請求；

一第二儲存器件自該等主機請求佇列中之一第二者讀取一第二請求，該第二主機請求佇列與該第一主機請求佇列相異，該第二儲存器件之該讀取獨立於該第一儲存器件之該讀取；

該第一儲存器件將該第一請求之至少一部分轉發至該第二儲存器件；

該第二儲存器件將該第二請求之至少一部分轉發至該第一儲存器件，該第二儲存器件之該轉發獨立於該第一儲存器件之該轉發；

該第一儲存器件根據該第二請求之至少該部分存取第一儲存器，並至少部分地基於該第一儲存器件之該存取將一第一狀態寫入至複數個主機狀態佇列中之一第一者；及
該第二儲存器件根據該第一請求之至少該部分存取第二儲存器並至少部分地基於該第二儲存器件之該存取將一第二狀態寫入至該等主機狀態佇列之一第二者，該第二儲存器件之該存取獨立於該第一儲存器件之該存取，該第二儲存器與該第一儲存器相異。

EC26) 一種系統，其包括：

用於使一第一儲存器件自複數個主機請求佇列中之一第一者讀取一第一請求之構件；

用於使一第二儲存器件自該等主機請求佇列之一第二者讀取一第二請求之構件，該第二主機請求佇列與該第一主機請求佇列相異，該第二儲存器件之該讀取獨立於該第一儲存器件之該讀取；

用於使該第一儲存器件將該第一請求之至少一部分轉發至該第二儲存器件之構件；

用於使該第二儲存器件將該第二請求之至少一部分轉發至該第一儲存器件之構件，該第二儲存器件之該轉發獨立於該第一儲存器件之該轉發；

用於使該第一儲存器件根據該第二請求之至少該部分存取第一儲存器並至少部分地基於該第一儲存器件之該存取將一第一狀態寫入至複數個主機狀態佇列中之一第一者之構件；及

用於使該第二儲存器件根據該第一請求之至少該部分存取第二儲存器並至少部分地基於該第二儲存器件之該存取將一第二狀態寫入至該等主機狀態佇列之一第二者之構件，該第二儲存器件之該存取獨立於該第一儲存器件之該存取，該第二儲存器與該第一儲存器相異。

EC27) 一種系統，其包括：

一第一儲存器件之主機介面邏輯硬體，其經啟用以自複數個主機請求佇列中之一第一者讀取一第一請求；

一第二儲存器件之主機介面邏輯硬體，其經啟用以自該等

主機請求佇列之一第二者讀取一第二請求，該第二主機請求佇列與該第一主機請求佇列相異，該第二儲存器件之該讀取獨立於該第一儲存器件之該讀取；

該第一儲存器件之請求轉發邏輯硬體，其經啟用以將該第一請求之至少一部分轉發至該第二儲存器件；

該第二儲存器件之請求轉發邏輯硬體，其經啟用以將該第二請求之至少一部分轉發至該第一儲存器件，該第二儲存器件之該轉發獨立於該第一儲存器件之該轉發；

該第一儲存器件之儲存介面邏輯硬體，其經啟用以根據該第二請求之至少該部分存取第一儲存器；

該第一儲存器件之狀態產生邏輯硬體，其經啟用以至少部分地基於該第一儲存器件之該存取判定一第一狀態，且進一步經啟用以將該第一狀態提供至該第一儲存器件之該主機介面邏輯硬體以寫入至複數個主機狀態佇列中之一第一者；

該第二儲存器件之儲存介面邏輯硬體，其經啟用以根據該第一請求之至少該部分存取第二儲存器；及

該第二儲存器件之狀態產生邏輯硬體，其經啟用以至少部分地基於該第二儲存器件之該存取判定一第二狀態，且進一步經啟用以將該第二狀態提供至該第二儲存器件之該主機介面邏輯硬體以寫入至該等主機狀態佇列之一第二者，該第二儲存器件之該存取獨立於該第一儲存器件之該存取，該第二儲存器與該第一儲存器相異。

EC28) 一種系統，其包括：

用於使得一主機能夠作為單個邏輯介面存取散佈於複數個實體儲存器件上之複數個儲存範圍之一彙總之構件，該等實體儲存器件中之一者經啟用以操作為一初級代理程式且該等實體儲存器件中之一或多者經啟用以操作為次級代理程式；

用於在該初級代理程式與該等次級代理程式中之至少一者之間傳遞同級間訊務之構件；

其中該初級代理程式經啟用以管理該等次級代理程式之至少一部分供用於冗餘資料儲存；且

其中該同級間訊務包括冗餘訊務以實施該冗餘資料儲存。

EC29) 如 EC28 之系統，其中該冗餘訊務之至少一部分包括累積同位資訊。

EC30) 如 EC29 之系統，其中至少該部分係自該初級代理程式及/或該等次級代理程式中之一者轉發。

EC31) 如 EC29 之系統，其中至少該部分係轉發至實施對應於該累積同位資訊之累積同位儲存之該等次級代理程式中之一者。

EC32) 如 EC28 之系統，其中該冗餘訊務之至少一部分係由該初級代理程式提供至該等次級代理程式中之一或多者。

EC33) 如 EC28 之系統，其中該冗餘訊務之至少一部分係由該等次級代理程式中之一或多者提供至該初級代理程式。

EC34) 如 EC28 之系統，其中該冗餘訊務之至少一部分

係在複數個該等次級代理程式之間。

EC35) 如 EC28 之系統，其中該冗餘訊務包括控制資訊、未轉換資料及已轉換資料。

EC36) 如 EC35 之系統，其中該控制資訊包括將針對一特定儲存位址發生多少冗餘更新之一指示。

EC37) 如 EC36 之系統，其中該初級代理程式及該等次級代理程式中之至少一者經啟用以使用該指示來快取與該等冗餘更新有關之資訊直至已針對該特定儲存位址發生該等冗餘更新為止。

EC38) 如 EC35 之系統，其中該未轉換資料包括來自該主機之寫入資料之一或多個部分。

EC39) 如 EC38 之系統，其中寫入資料之該等部分係至少部分地基於一鏡像操作來判定。

EC40) 如 EC35 之系統，其中該經轉換資料包括以下各項中之一或多者

至少部分地基於來自該主機之寫入資料之同位資料，

至少部分地基於來自該主機之寫入資料之互斥「或」

(XOR) 資料，及

至少部分地基於廉價/獨立磁碟冗餘陣列(RAID)實施方案之冗餘資訊。

EC41) 如 EC28 之系統，其中該冗餘訊務包括讀取-修改-寫入(RMW)操作控制及/或資料冗餘訊務。

EC42) 如 EC41 之系統，其中該 RMW 操作資料冗餘訊務係來自該主機之寫入資料之所有或任何部分之一複本。

EC43) 如 EC41 之系統，其中該 RMW 操作資料冗餘訊務係至少部分地基於來自該主機之寫入資料之一部分之經轉換資料。

EC1) 如 EC41 之系統，其中該 RMW 操作控制冗餘訊務係將針對一特定位址出現多少冗餘更新之一指示。

EC45) 如 EC28 之系統，其中該冗餘訊務包括資料恢復操作控制及/或資料冗餘訊務。

EC46) 如 EC28 之系統，其中該冗餘資料儲存係根據一或多個廉價/獨立磁碟冗餘陣列(RAID)技術。

EC47) 如 EC28 之系統，其中該初級代理程式經啟用以自該主機接受一請求以存取儲存空間、將該請求作為一或多個子請求轉發至該等次級代理程式之所有或任何部分及自該等子請求所轉發至之該等次級代理程式接受與該等子請求相關聯之子狀態。

EC48) 如 EC47 之系統，其中該等子請求係至少部分地基於該冗餘資料儲存所根據之一冗餘技術來判定。

EC49) 如 EC48 之系統，其中該冗餘技術係根據一廉價/獨立磁碟冗餘陣列(RAID)實施方案。

EC50) 如 EC47 之系統，其中該等子請求之至少一部分係至少部分地基於跨越該等子請求所轉發至之該等實體器件之一等量化。

EC51) 如 EC50 之系統，其中該等量化係根據一廉價/獨立磁碟冗餘陣列(RAID)實施方案。

EC52) 如 EC50 之系統，其中該等量化係根據工作負載

動態地改變。

EC53) 如 EC28 之系統，其中該等次級代理程式中之每一者經啟用以自該初級代理程式接受一或多個子請求、轉譯該等所接受之子請求之主機上下文定址資訊以存取本端儲存器及至少部分地基於該本端存取而將各別子狀態提供至該初級代理程式。

EC54) 如 EC28 之系統，其中該等次級代理程式中之每一者經啟用以自該初級代理程式接受一或多個子請求、轉譯該等所接受子請求之主機上下文定址資訊以存取本端儲存器、存取該本端儲存器及藉助該主機傳送資料，該資料係該存取之讀取/寫入資料。

EC55) 如 EC28 之系統，其進一步包括用於在該主機與該初級代理程式之間傳遞主機訊務之構件。

EC56) 如 EC55 之系統，其中該用於傳遞同級間訊務之構件係經由與該用於傳遞主機訊務之構件之邏輯通道相異之至少一個邏輯通道。

EC57) 如 EC55 之系統，其中該用於傳遞同級間訊務之構件係經由與該用於傳遞主機訊務之構件之實體通道相異之至少一個實體通道。

EC58) 如 EC55 之系統，其中該用於傳遞同級間訊務之構件及該用於傳遞主機訊務之構件中的一或多者與一高速周邊組件互連(PCIe)標準相容。

EC59) 如 EC55 之系統，其中該用於傳遞同級間訊務之構件與該用於傳遞主機訊務之構件中的一或多者與一無限

帶寬標準相容。

EC60) 如 EC55 之系統，其中該用於傳遞同級間訊務之構件包括一不透明開關之所有或任何部分。

EC61) 如 EC55 之系統，其中該用於傳遞主機訊務之構件包括一透明開關之所有或任何部分。

EC62) 一種方法，其包括：

使得一主機能夠作為單個邏輯介面存取散佈於複數個實體儲存器件上之複數個儲存範圍之一彙總，該等實體儲存器件中之一者經啟用以操作為一初級代理程式且該等實體儲存器件中之一或多者經啟用以操作為次級代理程式；

在該初級代理程式與該等次級代理程式中之至少一者之間傳遞同級間訊務；

其中該初級代理程式經啟用以管理該等次級代理程式之至少一部分供用於冗餘資料儲存；且

其中該同級間訊務包括冗餘訊務以實施該冗餘資料儲存。

EC63) 一種系統，其包括：

主機介面邏輯硬體；

複數個實體儲存器件，其經啟用以至少部分地經由該主機介面邏輯硬體與一主機選擇性地通信，該等實體儲存器件中之至少一者經啟用以操作為一初級代理程式且該等實體儲存器件中之一或多者經啟用以操作為各別次級代理程式；

同級間通信邏輯硬體，其實現該初級代理程式與該等次級

代理程式中之至少一者之間的同級間訊務；

其中該等實體儲存器件實施散佈於該等實體儲存器件上之複數個儲存範圍之一彙總；

其中該初級代理程式經啟用以管理該等次級代理程式之至少一部分以實施冗餘資料儲存；且

其中該同級間訊務包括冗餘訊務以實施該冗餘資料儲存。

EC64) 一種系統，其包括：

用於經由一同級間協定在複數個儲存代理程式之間介接之構件；

用於偵測該等儲存代理程式之一出故障者之一故障之構件，該故障係該出故障之儲存代理程式不再實施一特定儲存空間部分；及

用於分配經受再分配之儲存空間以實施該特定儲存空間部分之構件，該經受再分配之儲存空間包括由該等儲存代理程式之無故障者實施之儲存空間之任一組合。

EC65) 如 EC64 之系統，其中該等儲存代理程式包括次級代理程式。

EC66) 如 EC64 之系統，其中該等儲存代理程式包括初級代理程式。

EC67) 如 EC64 之系統，其中該等構件係包括於一初級代理程式中，且該等儲存代理程式包括次級代理程式。

EC68) 如 EC64 之系統，其中該等儲存代理程式中之每一者實施實體儲存空間之至少一各別部分。

EC69) 如 EC68 之系統，其中該等各別部分包括一或多

個非揮發性記憶體。

EC70) 如 EC64 之系統，其進一步包括本端儲存器且其中該經受分配之儲存空間進一步包括該本端儲存器。

EC71) 如 EC64 之系統，其中該故障係一部分故障，且該出故障之儲存代理程式在該部分故障之後繼續實施至少某些儲存空間。

EC72) 如 EC71 之系統，其中該經受分配之儲存空間進一步包括該至少某些儲存空間。

EC73) 如 EC64 之系統，其中該用於偵測之構件係至少部分地經由該等儲存代理程式中之一者實施。

EC74) 如 EC64 之系統，其中該用於偵測之構件係至少部分地經由一器件驅動器實施。

EC75) 如 EC64 之系統，其進一步包括用於經由一或多個冗餘技術恢復資料之構件。

EC76) 如 EC75 之系統，其中該等冗餘技術包括一鏡像技術。

EC77) 如 EC75 之系統，其中該等冗餘技術包括一廉價/獨立磁碟冗餘陣列(RAID)技術。

EC78) 一種方法，其包括：
經由一同級間協定在複數個儲存代理程式之間介接；
偵測該等儲存代理程式之一出故障者之一故障，該故障係該出故障之儲存代理程式不再實施一特定儲存空間部分；及
分配經受再分配之儲存空間以實施該特定儲存空間部分，

該經受再分配之儲存空間包括由該等儲存代理程式之無故障者實施之儲存空間之任一組合。

EC79) 一種系統，其包括：

同級間通信邏輯硬體，其實現複數個儲存代理程式之間的同級間協定通信；

故障偵測邏輯硬體，其經啟用以偵測該等儲存代理程式之一出故障者之一故障，該故障係該出故障之儲存代理程式不再實施一特定儲存空間部分；及

一處理器，其經啟用以執行在由該處理器執行時致使該處理器執行包括分配經受再分配之儲存空間以實施該特定儲存空間部分之操作的指令，該經受再分配之儲存空間包括由該等儲存代理程式之無故障者實施之儲存空間之任一組合。

EC80) 一種具有儲存於其中之一組指令之電腦可讀媒體，該組指令在由一處理元件執行時致使該處理元件執行包括以下之操作：

管理經由一同級間協定在複數個儲存代理程式之間介接；

管理偵測該等儲存代理程式之一出故障者之一故障，該故障係該出故障之儲存代理程式不再實施一特定儲存空間部分；及

管理分配經受再分配之儲存空間以實施該特定儲存空間部分，該經受再分配之儲存空間包括由該等儲存代理程式之無故障者實施之儲存空間之任一組合。

EC81) 一種系統，其包括：

用於偵測複數個儲存代理程式之一出故障者之一故障之構件，該故障係該出故障之儲存代理不再使得一主機能夠作為單個邏輯介面存取散佈於該等儲存代理程式上之複數個儲存範圍之一彙總；

用於識別一取代儲存代理程式以取代該出故障之儲存代理程式之構件；及

用於組態該取代儲存代理程式以提供該單個邏輯介面之構件。

EC82) 如 EC81 之系統，其中該取代儲存代理程式係一備用磁碟。

EC83) 如 EC81 之系統，其中該取代儲存代理程式係該等儲存代理程式之不同於該出故障之儲存代理程式之一者。

EC84) 如 EC83 之系統，其中該取代儲存代理程式係一初級代理程式。

EC85) 如 EC83 之系統，其中該取代儲存代理程式係選擇性地操作為一初級代理程式或一次級代理程式之一可組態代理程式。

EC86) 如 EC81 之系統，其中該用於偵測之構件係至少部分地經由該等儲存代理程式中之一者實施。

EC87) 如 EC81 之系統，其中該用於偵測之構件係至少部分地經由在該主機上執行之一器件驅動器實施。

EC88) 如 EC81 之系統，其進一步包括用於經由一主機介面協定介接至該主機之構件，其包括自一提交佇列中之

一條目讀取一儲存請求，該用於介接之構件係根據該提供該單個邏輯介面。

EC89) 如 EC88 之系統，其進一步包括用於至少部分地基於該儲存請求之位址資訊判定一或多個儲存子請求之構件。

EC90) 如 EC89 之系統，其進一步包括用於將該等子請求轉發至該等儲存代理程式中之次級代理程式之構件。

EC91) 如 EC90 之系統，其進一步包括用於接收分別對應於該等子請求之子狀態資訊並至少部分地基於該所接收子狀態資訊將總狀態返回至該主機。

EC92) 如 EC81 之系統，其進一步包括用於在該等儲存代理程式間傳遞同級間訊務之構件。

EC93) 一種方法，其包括：

偵測複數個儲存代理程式之一出故障者之一故障，該故障係該出故障之儲存代理程式不再使得一主機能夠作為單個邏輯介面存取散佈於該等儲存代理程式上之複數個儲存範圍之一彙總；

識別一取代儲存代理程式以取代該出故障之儲存代理程式；及

組態該取代儲存代理程式以提供該單個邏輯介面。

EC94) 一種系統，其包括：

故障偵測邏輯硬體，其經啟用以偵測複數個儲存代理程式之一出故障者之一故障，該故障係該出故障之儲存代理程式不再使得一主機能夠作為單個邏輯介面存取散佈於

該等儲存代理程式上之複數個儲存範圍之一彙總；及
一處理器，其經啟用以執行在由該處理器執行時致使該處理器執行包括以下各項之操作之指令

識別一取代儲存代理程式以取代該出故障之儲存代理程式，及

組態該取代儲存代理程式以提供該單個邏輯介面。

EC95) 一種具有儲存於其中之一組指令之電腦可讀媒體，該組指令在由一處理元件執行時致使該處理元件執行包括以下各項之操作：

管理複數個儲存代理程式之一出故障者之一故障，該故障係該出故障之儲存代理程式不再使得一主機能夠作為單個邏輯介面存取散佈於該等儲存代理程式上之複數個儲存範圍之一彙總；

管理識別一取代儲存代理程式以取代該出故障之儲存代理程式；及

管理組態該取代儲存代理程式以提供該單個邏輯介面。

EC96) 一種系統，其包括：

用於介接至大容量儲存器之儲存介面構件；

用於經由一主機介面協定介接至一主機之主機介面構件，該主機介面構件包括用於自一提交佇列中之一條目讀取一儲存請求之構件；

用於經由一同級間協定介接至一初級代理程式之初級代理程式介面構件；

用於判定至少一個位址以經由該儲存介面構件存取該大容

量儲存器之位址判定構件，該用於判定之構件至少部分地基於包括於該儲存請求中之位址資訊操作；
用於將一零儲存容量報告回至該主機且進一步用於將根據該大容量儲存器之一儲存容量之一儲存容量報告回至該初級代理程式之儲存容量報告構件；且
其中該主機介面構件及該初級代理程式介面構件係至少部分地經由同一實體通道實施。

EC97) 如 EC96 之系統，其進一步包括用於至少部分地基於該位址資訊判定將所累計同位資料轉發至之至少一個目的地之目的地判定構件。

EC98) 如 EC97 之系統，其中該至少一個目的地可經由該同級間協定到達。

EC99) 如 EC98 之系統，其中該至少一個目的地係該初級代理程式。

EC100) 如 EC98 之系統，其中該至少一個目的地係一次級代理程式。

EC101) 如 EC96 之系統，其中該位址判定構件根據一或多個冗餘技術操作。

EC102) 如 EC101 之系統，其中該等冗餘技術包括一或多個鏡像技術及 / 或一或多個廉價 / 獨立磁碟冗餘陣列 (RAID) 技術。

EC103) 如 EC96 之系統，其中該主機介面構件進一步包括用於在該主機與該大容量儲存器之間傳送資料之構件。

EC104) 如 EC96 之系統，其中該同一實體通道與一高速

周邊組件互連(PCIe)標準相容。

EC105) 如 EC96 之系統，其中該同一實體通道與一無限帶寬標準相容。

EC106) 如 EC96 之系統，其中該大容量儲存器包括一或多個非揮發性記憶體。

EC107) 一種方法，其包括：

介接至大容量儲存器；

經由一主機介面協定介接至一主機，該介接至該主機包括自一提交佇列中之一條目讀取一儲存請求；

經由一同級間協定介接至一初級代理程式；

判定至少一個位址以經由該介接至大容量儲存器來存取該大容量儲存器，該判定可至少部分地基於包括於該儲存請求中之位址資訊來操作；

將一零儲存容量報告回至該主機，且將根據該大容量儲存器之一儲存容量之一儲存容量報告回至該初級代理程式；且

其中該介接至該主機及該介接至該初級代理程式係至少部分地經由同一實體通道實施。

EC108) 一種系統，其包括：

一大容量儲存器硬體介面，其經啟用以與大容量儲存器介接；

主機介接邏輯硬體，其實現與該主機介接、與一主機介面協定相容及經啟用以自一提交佇列中之一條目讀取一儲存請求；

同級間通信邏輯硬體，其實現經由一同級間協定與一初級代理程式介接；

存取位址判定邏輯硬體，其經啟用以判定至少一個位址以經由該大容量儲存器硬體介面存取該大容量儲存器，該判定可至少部分地基於包括於該儲存請求中之位址資訊操作；

儲存容量報告邏輯硬體，其經啟用以將一零儲存容量報告回至該主機，且進一步經啟用以將根據該大容量儲存器之一儲存容量之一儲存容量報告回至該初級代理程式；且

其中該與該主機介接及該與該初級代理程式介接係至少部分地經由同一實體通道實施。

EC109) 一種具有儲存於其中之一組指令之電腦可讀媒體，該組指令在由一處理元件執行時致使該處理元件執行包括以下各項之操作：

管理介接至大容量儲存器；

管理經由一主機介面協定介接至一主機，該介接至該主機包括自一提交佇列中之一條目讀取一儲存請求；

管理經由一同級間協定介接至一初級代理程式；

管理判定至少一個位址以經由該介接至大容量儲存器來存取該大容量儲存器，該判定可至少部分地基於包括於該儲存請求中之位址資訊來操作；

管理將一零儲存容量報告回至該主機，及將根據該大容量儲存器之一儲存容量之一儲存容量報告回至該初級代理

程式；且

其中該介接至該主機及該介接至該初級代理程式係至少部分地經由同一實體通道實施。

EC110) 一種系統，其包括：

用於介接至大容量儲存器之儲存介面構件；

用於經由一主機介面協定介接至一主機之主機介面構件，該主機介面構件包括用於在該主機與該大容量儲存器之間傳送資料之構件；

用於經由一同級間協定介接至一初級代理程式之初級代理程式介面構件，該初級代理程式介面構件包括用於自該初級代理程式接收一經轉發之儲存請求；

用於判定至少一個位址以存取該大容量儲存器之位址判定構件，該用於判定之構件至少部分地基於包括於該所轉發儲存請求中之位址資訊操作；

其中該所轉發儲存請求對應於自一提交佇列獲得之一請求，該請求及該所轉發請求係指相同的一或多個位址範圍；且

其中該主機介面構件及該初級代理程式介面構件係至少部分地經由同一實體通道實施。

EC111) 如EC110之系統，其進一步包括用於至少部分地基於該位址資訊判定欲將一累積同位資料轉發至之至少一個目的地。

EC112) 如EC111之系統，其中該至少一個目的地可經由該同級間協定到達。

EC113) 如 EC112 之系統，其中該至少一個目的地係該初級代理程式。

EC114) 如 EC112 之系統，其中該至少一個目的地係一次級代理程式。

EC115) 如 EC110 之系統，其中該位址判定構件根據一或多個冗餘技術操作。

EC116) 如 EC115 之系統，其中該等冗餘技術包括一或多個鏡像技術及 / 或一或多個廉價 / 獨立磁碟冗餘陣列 (RAID) 技術。

EC117) 如 EC110 之系統，其進一步包括用於將一零儲存容量報告回至該主機及進一步用於將根據該大容量儲存器之一儲存容量之一儲存容量報告回至該初級代理程式之儲存容量報告構件。

EC118) 如 EC110 之系統，其中該同一實體通道與一高速周邊組件互連 (PCIe) 標準相容。

EC119) 如 EC110 之系統，其中該同一實體通道與一無限帶寬標準相容。

EC120) 如 EC110 之系統，其中該大容量儲存器包括一或多個非揮發性記憶體。

EC121) 一種方法，其包括：

介接至大容量儲存器；

經由一主機介面協定介接至一主機，該介接至該主機包括

在該主機與該大容量儲存器之間傳送資料；

經由一同級間協定介接至一初級代理程式，該介接至該初

級代理程式包括自該初級代理程式接收一所轉發儲存請求；

判定至少一個位址以經由該介接至大容量儲存器來存取該大容量儲存器，該判定可至少部分地基於包括於該儲存請求中之位址資訊來操作；

其中該所轉發儲存請求對應於自一提交佇列獲得之一請求，該請求及該所轉發請求係指相同的一或多個位址範圍；且

其中該介接至該主機及該介接至該初級代理程式係至少部分地經由同一實體通道實施。

EC122) 一種系統，其包括：

一大容量儲存器硬體介面，其經啟用以與大容量儲存器介接；

主機介面邏輯硬體，其實現與該主機介接、與一主機介面協定相容及經啟用以至少部分地經由該大容量儲存器硬體介面在該主機與該大容量儲存器之間傳送資料；

同級間通信邏輯硬體，其實現經由一同級間協定與一初級代理程式介接；

儲存請求接受邏輯硬體，其實現至少部分地經由該同級間通信邏輯硬體自該初級代理程式接收一所轉發儲存請求；

存取位址判定邏輯硬體，其經啟用以判定至少一個位址以經由該大容量儲存器硬體介面存取該大容量儲存器，該判定可至少部分地基於包括於該儲存請求中之位址資訊

操作；

其中該所轉發儲存請求對應於自一提交佇列獲得之一請求，該請求及該所轉發請求係指相同的一或多個位址範圍；且

其中該與該主機介接及該與該初級代理程式介接係至少部分地經由同一實體通道實施。

EC123) 一種具有儲存於其中之一組指令之電腦可讀媒體，該組指令在由一處理元件執行時致使該處理元件執行包括以下各項之操作：

管理介接至大容量儲存器；

管理經由一主機介面協定介接至一主機，該介接至該主機包括在該主機與該大容量儲存器之間傳送資料；

管理經由一同級間協定介接至一初級代理程式，該介接至該初級代理程式包括自該初級代理程式接收一所轉發儲存請求；

管理判定至少一個位址以經由該介接至大容量儲存器來存取該大容量儲存器，該判定可至少部分地基於包括於該儲存請求中之位址資訊來操作；

其中該所轉發儲存請求對應於自一提交佇列獲得之一請求，該請求及該所轉發請求係指相同的一或多個位址範圍；且

其中該介接至該主機及該介接至該初級代理程式係至少部分地經由同一實體通道實施。

EC124) 一種系統，其包括：

用於經由一主機介面協定介接至一主機之主機介面構件，
該主機介面構件包括用於自一提交佇列中之一條目讀取
一儲存請求之構件；

用於經由一同級間協定介接至一次級代理程式之次級代理
程式介面構件，該次級代理程式介面構件包括用於將一
儲存子請求轉發至該次級代理程式之構件；

用於至少部分地基於包括於該儲存請求中之位址資訊判定
該子請求之子請求產生構件；

用於將一總儲存容量報告回至該主機之儲存容量報告構
件，該總儲存容量至少部分基於該次級代理程式之一儲
存容量；且

其中該主機介面構件及該次級代理程式介面構件係至少部
分地經由同一實體通道實施。

EC125) 如 EC124 之系統，其進一步包括用於介接至大
容量儲存器之儲存介面構件。

EC126) 如 EC125 之系統，其中該大容量儲存器包括一
或多個非揮發性記憶體。

EC127) 一種方法，其包括：

經由一主機介面協定與一主機介接，該主機介接包括自一
提交佇列中之一條目讀取一儲存請求；

經由一同級間協定與一次級代理程式介接，該次級代理程
式介接包括將一儲存子請求轉發至該次級代理程式；

至少部分地基於包括於該儲存請求中之位址資訊判定該子
請求；

將一總儲存容量報告回至該主機，該總儲存容量至少部分地基於該次級代理程式之一儲存容量；及
其中該與該主機介接及該與該次級代理程式介接係至少部分地經由同一實體通道實施。

EC128) 一種系統，其包括：

主機介接邏輯硬體，其實現與該主機介接、與一主機介面協定相容及經啟用以自一提交佇列中之一條目讀取一儲存請求；

同級間通信邏輯硬體，其實現經由一同級間協定與一次級代理程式介接；

請求轉發邏輯硬體，其經啟用以至少部分地經由該同級間通信邏輯硬體將一儲存子請求轉發至該次級代理程式；
子請求判定邏輯硬體，其經啟用以至少部分地基於包括於該儲存請求中之位址資訊判定該子請求；

儲存容量報告邏輯硬體，其經啟用以將一總儲存容量報告回至該主機，該總儲存容量至少部分地基於該次級代理程式之一儲存容量；且

其中該與該主機介接及該與該次級代理程式介接係至少部分地經由同一實體通道實施。

EC129) 一種具有儲存於其中之一組指令之電腦可讀媒體，該組指令在由一處理元件執行時致使該處理元件執行包括以下之操作：

管理經由一主機介面協定與一主機介接，該主機介接包括自一提交佇列中之一條目讀取一儲存請求；

管理經由一同級間協定與一次級代理程式介接，該次級代理程式介接包括將一儲存子請求轉發至該次級代理程式；

管理至少部分地基於包括於該儲存請求中之位址資訊判定該子請求；

管理將一總儲存容量報告回至該主機，該總儲存容量至少部分地基於該次級代理程式之一儲存容量；且

其中該與該主機介接及該與該次級代理程式介接係至少部分地經由同一實體通道實施。

EC130) 一種系統，其包括：

用於介接至大容量儲存器之儲存介面構件；

用於經由一主機介面協定介接至一主機之主機介面構件，該主機介面構件包括用於在該主機與該大容量儲存器之間傳送資料之構件之主機介面構件；

用於經由一同級間協定介接至一初級代理程式之初級代理程式介面構件，該初級代理程式介面構件包括用於自該初級代理程式接收一所轉發儲存子請求之構件；

用於將一零儲存容量報告回至該主機且進一步用於將根據該大容量儲存器之一儲存容量之一儲存容量報告回至該初級代理程式之儲存容量報告構件；且

其中該主機介面構件及該初級代理程式介面構件係至少部分地經由同一實體通道實施。

EC131) 如 EC130 之系統，其中該大容量儲存器包括一或多個非揮發性記憶體。

EC132) 一種方法，其包括：

介接至大容量儲存器；

經由一主機介面協定與一主機介接，該主機介接包括在該主機與該大容量儲存器之間傳送資料；

經由一同級間協定與一初級代理程式介接，該初級代理程式介接包括自該初級代理程式接收一所轉發儲存子請求；

將一零儲存容量報告回至該主機，且進一步將根據該大容量儲存器之一儲存容量之一儲存容量報告回至該初級代理程式；且

其中該與該主機介接及該與該初級代理程式介接係至少部分地經由同一實體通道實施。

EC133) 一種系統，其包括：

一大容量儲存器硬體介面，其經啟用以與大容量儲存器介接；

主機介面邏輯硬體，其實現與該主機介接、與一主機介面協定相容及經啟用以至少部分地經由該大容量儲存器硬體介面在該主機與該大容量儲存器之間傳送資料；

同級間通信邏輯硬體，其實現經由一同級間協定與一初級代理程式介接；

子請求接收邏輯硬體，其經啟用以至少部分地經由該同級間通信邏輯硬體自該初級代理程式接收一所轉發儲存子請求；

儲存容量報告邏輯硬體，其經啟用以將一零儲存容量報告

回至該主機，且進一步經啟用以將根據該大容量儲存器之一儲存容量之一儲存容量報告回至該初級代理程式；且其中該與該主機介接及該與該初級代理程式介接係至少部分地經由同一實體通道實施。

EC134) 一種具有儲存於其中之一組指令之電腦可讀媒體，該組指令在由一處理元件執行時致使該處理元件執行包括以下各項之操作：

管理介接至大容量儲存器；

管理經由一主機介面協定與一主機介接，該主機介接包括在該主機與該大容量儲存器之間傳送資料；

管理經由一同級間協定與一初級代理程式介接，該初級代理程式介接包括自該初級代理程式接收一所轉發儲存子請求；

管理將一零儲存容量報告回至該主機，且進一步將根據該大容量儲存器之一儲存容量之一儲存容量報告回至該初級代理程式；且

其中該與該主機介接及該與該初級代理程式介接係至少部分地經由同一實體通道實施。

EC135) 一種系統，其包括：

用於介接至大容量儲存器之儲存介面構件；

用於經由一主機介面協定介接至一主機之主機介面構件，該主機介面構件包括用於在該主機與該大容量儲存器之間傳送資料之構件；

用於經由一同級間協定介接至一初級代理程式之初級代理

程式介面構件，該初級代理程式介面構件包括用於自該初級代理程式接收一子請求之構件及用於將一子狀態發送至該初級代理程式之構件，該子狀態至少部分地基於根據該子請求存取該大容量儲存器之結果；

用於將一零儲存容量報告回至該主機且進一步用於將根據該大容量儲存器之一儲存容量之一儲存容量報告回至該初級代理程式之儲存容量報告構件；且

其中該主機介面構件及該初級代理程式介面構件係至少部分地經由同一實體通道實施。

EC136) 如 EC135 之系統，其中該初級代理程式介面構件進一步包括用於與該初級代理程式傳遞冗餘資訊之構件。

EC137) 如 EC136 之系統，其中該冗餘資訊與一或多個鏡像技術及/或一或多個廉價/獨立磁碟冗餘陣列(RAID)技術相容。

EC138) 如 EC135 之系統，其中該同一實體通道與一高速周邊組件互連(PCIe)標準相容。

EC139) 如 EC135 之系統，其中該同一實體通道與一無限帶寬標準相容。

EC140) 如 EC135 之系統，其中該大容量儲存器包括一或多個非揮發性記憶體。

EC141) 一種方法，其包括：

介接至大容量儲存器；

經由一主機介面協定與一主機介接，該與該主機介接包括

在該主機與該大容量儲存之間傳送資料；

經由一同級間協定與一初級代理程式介接，該與該初級代理程式介接包括自該初級代理程式接收一子請求及將一子狀態發送至該初級代理程式，該子狀態至少部分地基於根據該子請求存取該大容量儲存之結果；

將一零儲存容量報告回至該主機，且進一步將根據該大容量儲存器之一儲存容量之一儲存容量報告回至該初級代理程式；且

其中該與該主機介接及該與該初級代理程式介接係至少部分地經由同一實體通道實施。

EC142) 一種系統，其包括：

一大容量儲存器硬體介面，其經啟用以與大容量儲存器介接；

主機介面邏輯硬體，其實現與該主機介接、與一主機介面協定相容及經啟用以至少部分地經由該大容量儲存器硬體介面在該主機與該大容量儲存器之間傳送資料；

同級間通信邏輯硬體，其實現經由一同級間協定與一初級代理程式介接；

子請求接收邏輯硬體，其經啟用以至少部分地經由該同級間通信邏輯硬體自該初級代理程式接收一所轉發儲存子請求；

子狀態判定邏輯硬體，其經啟用以至少部分地基於根據該子請求及經由該大容量儲存器硬體介面存取該大容量儲存器來判定一子狀態；

子狀態轉發邏輯硬體，其經啟用以至少部分地經由該同級間通信邏輯硬體將該子狀態轉發至該初級代理程式；

儲存容量報告邏輯硬體，其經啟用以將一零儲存容量報告回至該主機，且進一步經啟用以將根據該大容量儲存器之一儲存容量之一儲存容量報告回至該初級代理程式；且其中該與該主機介接及該與該初級代理程式介接係至少部分地經由同一實體通道實施。

EC143) 一種具有儲存於其中之一組指令之電腦可讀媒體，該組指令在由一處理元件執行時致使該處理元件執行包括以下各項之操作：

管理介接至大容量儲存器；

管理經由一主機介面協定與一主機介接，該與該主機介接包括在該主機與該大容量儲存器之間傳送資料；

管理經由一同級間協定與一初級代理程式介接，該與該初級代理程式介接包括自該初級代理程式接收一子請求及將一子狀態發送至該初級代理程式，該子狀態至少部分地基於根據該子請求存取該大容量儲存器之結果；

管理將一零儲存容量報告回至該主機，且進一步管理將根據該大容量儲存器之一儲存容量之一儲存容量報告回至該初級代理程式；且

其中該與該主機介接及該與該初級代理程式介接係至少部分地經由同一實體通道實施。

EC144) 一種方法，其包括：

接受來自一主機之一請求以存取在一位址處之儲存空間；

至少部分地基於該位址判定該請求對應於複數個儲存器件中之哪一或多者；

至少部分地基於該位址判定對應於該等所判定儲存器件中之每一者之一各別子請求；

將該等子請求發送至該等對應之所判定儲存器件；

自該等所判定儲存器件中之每一者接受各別子狀態；

至少部分地基於該等各別子狀態中之每一者判定一總狀態；

將該總狀態提供至該主機；

其中該接受該請求、該判定該等儲存器件、該判定及該發送該等子請求、該接受該等子狀態及該判定及該提供該總狀態係由該等儲存器件中之一特定者執行；

其中該接受該請求係經由一主機介面協定，且該發送該等子請求及該接受該等子狀態係經由可獨立於該主機操作之一同級間協定；且

其中該等儲存器件中之至少一者實施可經由該等子請求中之一者存取之儲存空間，且該等儲存器件中之該至少一者停用經由該主機介面協定對該所實施儲存空間之存取。

EC145) 如 EC144 之方法，其中該主機介面及該等同級間協定與一高速周邊組件互連(PCIe)通道相容。

EC146) 如 EC144 之方法，其中該接受該請求包括讀取主機記憶體中所保持之一提交佇列中之一條目。

EC147) 如 EC144 之方法，其中該提供該總狀態包括將

一條目寫入主機記憶體中所保持之一完成佇列中。

EC148) 如 EC144 之方法，其進一步包括由該等所判定儲存器件中之每一者獨立地執行該對應子請求之一資料傳送部分。

EC149) 如 EC148 之方法，其中該獨立執行包括存取主機記憶體。

EC150) 一種方法，其包括：

一第一儲存器件自複數個主機請求佇列中之一第一者讀取一第一請求；

一第二儲存器件自該等主機請求佇列中之一第二者讀取一第二請求，該第二主機請求佇列與該第一主機請求佇列相異，該第二儲存器件之該讀取獨立於該第一儲存器件之該讀取；

該第一儲存器件將該第一請求之至少一部分轉發至該第二儲存器件；

該第二儲存器件將該第二請求之至少一部分轉發至該第一儲存器件，該第二儲存器件之該轉發獨立於該第一儲存器件之該轉發；

該第一儲存器件根據該第二請求之至少該部分存取第一儲存空間並至少部分地基於該第一儲存器件進行之該存取將一第一狀態返回至該第二儲存器件；及

該第二儲存器件根據該第一請求之至少該部分存取第二儲存空間並至少部分地基於該第二儲存器件進行之該存取將一第二狀態返回至該第一儲存器件，該第二儲存器件

之該存取獨立於該第一儲存器件之該存取，該第二儲存空間與該第一儲存空間相異。

EC151) 如 EC150 之方法，其中該等主機請求佇列可由一主機存取，且進一步包括該第二儲存器件至少部分地基於該第二請求將一第三狀態返回至該主機且該第一儲存器件至少部分地基於該第一請求將一第四狀態返回至該主機。

EC152) 如 EC150 之方法，其中該第二請求之至少該部分係至少一第一部分，且進一步包括該第二儲存器件根據該第二請求之至少一第二部分存取第三儲存空間，該第二請求之該第一部分及該第二部分係該第二請求之相異部分。

EC153) 如 EC150 之方法，其中該等主機請求佇列可由一主機存取且該第二請求之至少該部分係至少一第一部分，且進一步包括該第二儲存器件根據該第二請求之至少一第二部分存取第三儲存空間且該第二儲存器件至少部分地基於對該第三儲存空間之該存取將一第三狀態返回至該主機，該第二請求之該第一部分及該第二部分係該第二請求之相異部分。

EC154) 一種方法，其包括：

將對複數個實體儲存器件之單個邏輯介面提供至一主機，

該等實體儲存器件中之一者操作為一初級代理程式且該

等實體儲存器件中之一或多者操作為次級代理程式；

其中該提供包括

該初級代理程式自該主機接受一請求以存取儲存空間、將該請求作為一或多個子請求轉發至該等次級代理程式之所有或任何部分、自該等子請求所轉發至之該等次級代理程式接受與該等子請求相關聯之子狀態及至少部分地基於該等子狀態將一總狀態提供至該主機，及

該等次級代理程式之該所有或任何部分接受該等子請求，將該等子請求之主機上下文定址資訊轉譯成本端上下文定址資訊，至少部分地基於該本端上下文定址資訊而存取本端儲存空間，且將該等子狀態提供至該初級代理程式；

其中在由該主機針對儲存容量查詢時該初級代理程式至少部分地基於由該初級代理程式實施之儲存空間及由該等次級代理程式實施之儲存空間來報告該儲存容量；且

其中在由該主機針對儲存容量查詢時該等次級代理程式中之每一者至少部分地基於該初級代理程式所報告之該儲存容量報告一各別儲存容量。

EC155) 如 EC154 之方法，其中由該初級代理程式所報告之該儲存容量根據由該等次級代理程式之一特定一者所實施之該儲存空間之至少一部分而增加，且進一步包括該特定次級代理程式在該特定次級代理程式之一儲存容量根據由該特定第二代理程式所實施之該儲存空間之至少一部分而減少時將該特定次級代理程式之一儲存容量報告給該主機。

EC156) 如 EC154 之方法，其中若由該主機針對儲存容

量查詢時該等次級代理程式之特定一者在由該初級代理程式所報告之該儲存容量中計及由該特定次級代理程式所實施之全部儲存空間，則報告回零儲存容量。

EC157) 如 EC154 之方法，其中該初級代理程式經啟用以經由一主機介面協定與該主機通信。

EC158) 如 EC154 之方法，其中該初級代理程式及該等次級代理程式經啟用以經由一同級間協定彼此通信。

EC159) 如 EC154 之方法，其中該初級代理程式經啟用以經由同一實體鏈路與該主機及該等次級代理程式中之至少一者通信。

EC160) 一種系統，其包括：

主機介面邏輯，其經啟用以經由一主機介面協定介接至一主機及經由一同級間協定介接至一次級代理程式，該介接至該主機包括自一提交佇列中之一條目讀取一儲存請求，該介接至該次級代理程式包括將一儲存子請求轉發至該次級代理程式，該介接至該主機及該介接至該次級代理程式係經由同一實體通道；

子請求產生邏輯，其經啟用以至少部分地基於包括於該儲存請求中之位址資訊判定該子請求；及

儲存容量報告邏輯，其經啟用以將一總儲存容量報告回至該主機，該總儲存容量至少部分地基於該次級代理程式之一儲存容量。

EC161) 如 EC160 之系統，其中該實體通道與一高速周邊組件互連(PCIe)通道及一無限帶寬通道中之至少一者相

容。

EC162) 如 EC160 之系統，其進一步包括該主機、該次級代理程式及將該主機耦合至該主機介面邏輯及/或該次級代理程式之一開關中之一或多者。

EC163) 如 EC160 之系統，其進一步包括儲存介面邏輯及大容量儲存器，該儲存介面邏輯經啟用以介接至該大容量儲存器，且其中該總儲存容量進一步係根據該大容量儲存器之一儲存容量。

EC164) 如 EC160 之系統，其中該介接至該主機進一步包括將一總狀態寫入至一完成佇列中之一條目，且該介接至該次級代理程式進一步包括經由該總狀態對該完成佇列之寫入將一子狀態作為該總狀態之至少部分自該次級代理程式轉發至該主機；且進一步包括經啟用以至少部分地基於該子狀態判定該總狀態之子狀態累積邏輯。

EC165) 如 EC160 之系統，其中系統係包括於一固態磁碟(SSD)中。

EC166) 一種系統，其包括：

主機介面邏輯，其經啟用以經由一主機介面協定介接至一主機及經由一同級間協定介接至一次級代理程式，該介接至該主機包括自一提交佇列中之一條目讀取一儲存請求，該介接至該次級代理程式包括將一儲存子請求轉發至該次級代理程式，該介接至該主機及該介接至該次級代理程式係經由同一實體通道；

子請求產生邏輯，其經啟用以至少部分地基於包括於該儲

存請求中之位址資訊判定該子請求；及
 儲存容量報告邏輯，其經啟用以將一總儲存容量報告回至該主機，該總儲存容量至少部分地基於該次級代理程式之一儲存容量。

EC167) 如 EC166 之系統，其中該實體通道與一高速周邊組件互連(PCIe)通道及一無限帶寬通道中之至少一者相容。

EC168) 如 EC166 之系統，其進一步包括該主機、該初級代理程式、將該主機耦合至該主機介面邏輯及/或該初級代理程式之一開關及該大容量儲存器之所有或任何部分中之一或多者。

EC169) 如 EC166 之系統，其中該大容量儲存器包括複數個快閃儲存器件。

EC170) 如 EC166 之系統，其中該系統係包括於一固態磁碟中。

可擴展儲存系統

圖 1A 圖解說明一種用於可擴展儲存器件之技術之一實施例之選定結構細節，該等可擴展儲存器件包含一主機、具有操作為各別初級代理程式之一或多個儲存器件之主機可見儲存空間及具有操作為各別次級代理程式之一或多個儲存器件之主機不可見儲存空間。主機 100 經由主機-儲存器器件耦合件 180 耦合至主機可見儲存空間 110 及主機不可見儲存空間 120。該主機-儲存器器件耦合件以及該主機可見及主機不可見儲存空間視情況實施為一可插式模組(圖解說

明為添加式卡 190A)及/或耦合件 101視情況實施為一纜線。在某些實施例中，該添加式卡之所有或任何部分係實施為一SSD。該主機可見儲存空間具有一或多個儲存器件(圖解說明為儲存器件 110.A及儲存器件 110.N)。類似地，該主機不可見儲存空間具有一或多個儲存器件(圖解說明為儲存器件 120.A及儲存器件 120.N)。在各種實施例中，該等儲存器件之任何一或多者係實體儲存器件，諸如SSD。

耦合件 101、111.A、111.N、121.A及 121.N實現主機、主機可見儲存空間與主機不可見儲存空間之間的請求、狀態及資料傳送。該等耦合件其中之一或多者實現經由一主機介面協定之傳送(諸如主機充當一主裝置且該主機可見儲存空間之該等元件中之一者操作為一從裝置)。該等耦合件其中之一或多者實現經由一同級間協定之傳送(諸如該主機可見儲存空間之該等元件中之一者操作為一初級代理程式且該主機不可見儲存空間之該等元件中之一者或該主機可見儲存空間之其他元件中之一者操作為一次級代理程式)。在各種實施例中，該等耦合件之一或多者與一介面標準相容(諸如PCIe或無限帶寬)。在各種實施例中，主機-儲存器件耦合件 180係經由一或多個PCIe及/或無限帶寬開關實施。在某些實施例中，主機-儲存器件耦合件 180與主機 100整合以作為一主機計算複合體之所有或任何部分。

該等主機可見儲存空間元件實施儲存空間，且在初始化期間，該等元件經組態以實現至少某些所實施儲存空間之

主機存取，因此提供對主機「可見」之儲存空間，諸如經由一主機介面協定。該等主機不可見儲存空間元件亦實施儲存空間，但在初始化期間，該等元件經組態以停用至少某些所實施儲存空間之主機存取，因此提供對主機「不可見」之儲存空間。然而，對主機不可見之儲存空間可由主機經由主機可見儲存空間之元件(諸如經由一同級間協定)間接地存取。

虛線箭頭 151 概念性地圖解說明主機 100 與儲存器件 110.A 之間的資訊傳送，且代表主機與主機可見儲存空間 110 之該等元件中之任一者之間的資訊傳送。該等資訊傳送包含源自主機之欲存取儲存空間之一請求、關於該請求之一狀態回應及關於該請求之一資料傳送中之一或多者。該請求包含一位址(諸如一 LBA)及一長度(諸如以位元組或 LBA 配量為單位)中之一或多者。自主機至儲存器件之該資訊傳送係經由耦合件 101、主機-儲存器件耦合件 180 及然後耦合件 111.A 來傳遞，且針對自儲存器件至主機之資訊傳送係反向的。

虛線箭頭 152 概念性地圖解說明在儲存器件 110.A 與儲存器件 120.A 之間的資訊傳送，且表示在主機可見儲存空間 110 之該等元件中之任一者與主機不可見儲存空間 120 之該等元件中之任一者之間的同級間資訊傳送。該等資訊傳送包含自儲存器件 110.A (充當一初級代理程式)至主機不可見儲存空間 120 (充當一次級代理程式)之該等元件中之任一者之一子請求、關於該子請求之一子狀態回應及關於該子請

求之一資料傳送中之一或多者。該子請求包含一位址(諸如一LBA)及一長度(諸如以位元組或LBA配量為單位)中之一或多者。自初級代理程式至次級代理程式之資訊傳送係經由耦合件111.A、主機儲存器件耦合件180及然後耦合件121.A來傳遞，且針對自次級代理程式至初級代理程式之資訊傳送係反向的。

虛線箭頭153概念性地圖解說明在儲存器件110.A與儲存器件110.N之間的資訊傳送，且表示在主機可見儲存空間110之該等元件中之任一者之間的同級間資訊傳送。該等資訊傳送包含自儲存器件110.A(充當一初級代理程式)至主機可見儲存空間110(充當一次級代理程式)之其他元件中之任一者之一子請求、關於該子請求之一子狀態回應及關於該子請求之一資料傳送中之一或多者。自初級代理程式至次級代理程式之資訊傳送係經由耦合件111.A、主機-儲存器件耦合件180及然後耦合件111.N來傳遞，且針對自次級代理程式至初級代理程式之資訊傳送係反向的。

虛線箭頭154概念性地圖解說明在主機100與儲存器件120.A之間的資訊傳送，且表示在主機與主機不可見儲存空間120之該等元件中之任一者之間的資訊傳送。該等資訊傳送包含關於該請求之一總狀態回應及關於該請求之一資料傳送中之一或多者。儘管在圖式中未清晰地圖解說明，但自主機至儲存器件之資訊傳送係經由耦合件101、主機-儲存器件耦合件180及然後耦合件121.A來傳遞，且針對自儲存器件至主機之資訊傳送係反向的。

在某些實施例中，源自一主機之請求係經由可經由耦合件101存取之主機記憶體中之一或多個提交佇列(例如，由一初級代理程式經由一或多個DMA操作讀取一提交佇列中之條目)。在某些實施例中，提供至一主機之狀態係經由可存取至初級代理程式及/或次級代理程式之主機記憶體中之一或多個完成佇列(例如，由一初級代理程式/次級代理程式經由一或多個DMA操作將條目寫入一完成佇列中)。在某些實施例中，請求及/或狀態係至少部分地經由一或多個門鈴暫存器及/或標尾暫存器傳遞，視情況結合一或多個提交及/或完成佇列。在某些實施例中，至少部分地經由一主機所起始之經程式化IO操作將請求及/或狀態傳遞至初級代理程式。在某些實施例中，至少部分地經由一主機所起始之經程式化IO操作將提交及/或完成佇列描述符傳遞至初級代理程式。

在某些實施例中，接受來自一主機之一請求包含自一提交佇列讀取一或多個條目。舉例而言，一主機將一特定條目寫入主機記憶體中所保持之一提交佇列中，相應地修改一標尾暫存器，及設定一對應門鈴暫存器(諸如在一初級代理程式中實施及/或可由一初級代理程式存取之一門鈴暫存器)。回應於該門鈴暫存器之設定，該初級代理程式存取該標尾暫存器並讀取該特定提交佇列條目(諸如經由來自該主機記憶體之一DMA讀取)。在某些實施例中，將狀態提供至一主機包含將一或多個條目寫入一完成佇列。舉例而言，一初級代理程式將一特定條目寫入在主機記憶

體中保持之一完成佇列中(諸如經由一DMA寫入至該主機記憶體)、相應地修改一標尾暫存器及通知主機一已完成操作(例如,藉由設定一對應門鈴暫存器或提供一中斷)。回應於該門鈴暫存器之該設定,主機存取標尾暫存器並讀取該特定完成佇列條目。

在某些實施例中,使用與自其判定該等子請求之一請求不同的路由、定址及/或囊封將子請求發送至次級代理程式,且在與該請求相同之一耦合件上輸送該等子請求。舉例而言,在使用PCIe耦合件用於輸送任何或所有請求、子請求及子狀態訊務之一實例性系統中,作為一讀取請求之一完成將一請求提供至一初級代理程式,同時在一MsgD請求中將一子請求轉發至一次級代理程式。在某些實施例中,一子請求經重封裝以包含除自其判定該子請求之一請求外的資訊。舉例而言,一經重封裝之子請求包含諸如流動控制、增強型路由資訊之額外資訊或關於請求內容之其他額外資訊。在某些實施例中,增強型路由資訊使得經由一或多個耦合件將子請求發送至遠端次級代理程式。

在某些實施例中,使用與該等子狀態對應之一請求不同的路由、定址及/或囊封將子狀態發送至初級代理程式,且在與該請求相同之一耦合件上輸送該等子狀態。舉例而言,在使用PCIe耦合件輸送任何或所有請求、子請求及子狀態訊務之一實例性系統中,作為一讀取請求之完成將一請求提供至一初級代理程式,同時由一次級代理程式將一子狀態在一MsgD請求中返回至該初級代理程式。在某些

實施例中，一子狀態經重封裝以包含除該子狀態及/或該子狀態對應之一請求外的資訊。舉例而言，一經重封裝之子狀態包含諸如流動控制、增強型路由資訊之額外資訊或關於該請求內容之其他額外資訊。在某些實施例中，增強型路由資訊使得經由一或多個耦合件將子狀態自遠端次級代理程式發送至初級代理程式。

作為一特定實例，充當一主裝置之主機100經由一主機介面協定將存取儲存空間之一請求(概念性地由虛線箭頭151圖解說明)提供至儲存器件110.A。作為回應，充當一從裝置之儲存器件110.A回應於該請求。儲存器件110.A經由主機介面協定接受該請求(亦由虛線箭頭151概念性地圖解說明)，且然後判定關於該請求之儲存空間係由儲存器件110.A及/或由主機不可見儲存空間120之該等元件之一或多者實施還是主機可見儲存空間110之其他元件中之任一者實施。儲存器件110.A然後在內部處理該請求(若該儲存空間之至少某些係由儲存器件110.A實施)及/或充當一初級代理程式經由同級間協定將該請求作為一或多個同級間子請求轉發至儲存器件120.A(由虛線箭頭152概念性地圖解說明)。充當一次級代理程式之儲存器件120.A經由該同級間協定接受該子請求(亦由虛線箭頭152概念性地圖解說明)，且然後在內部處理該子請求。儲存器件120.A然後將對應於該子請求之子狀態返回至儲存器件110.A(亦由虛線箭頭152概念性地圖解說明)。儲存器件110.A然後判定該請求之一總狀態並將該總狀態提供至主機(由虛線箭頭151

概念性地圖解說明)。另一選擇係，儲存器件120.A將子狀態直接返回至主機100(由虛線箭頭154概念性地圖解說明)，而不是儲存器件120.A將子狀態返回至儲存器件110.A(以隨後轉發至主機100)。

繼續該特定實例，用於讀取之資料類似於狀態地流動。儲存器件120.A將對應於該子請求之資料返回至儲存器件110.A(由虛線箭頭152概念性地圖解說明)，且然後儲存器件110.A將該資料提供至主機(由虛線箭頭151概念性地圖解說明)。另一選擇係，儲存器件120.A將該資料直接返回至主機100(由虛線箭頭154概念性地圖解說明)，而不是儲存器件120.A將該資料返回至儲存器件110.A(以然後轉發至主機100)。一特定請求之狀態及資料不限於同一流動。舉例而言，針對一個請求，狀態及資料兩者皆經由一初級代理程式(例如，儲存器件110.A)自一次級器件(例如，儲存器件120.A)流動至主機100。針對另一請求，狀態經由初級代理程式自該次級代理程式流動至主機，且資料直接自次級代理程式流動至主機。用於寫入之資料經由初級代理程式自主機流動至次級代理程式(由虛線箭頭151及152概念性地圖解說明)，或另一選擇係，直接自主機至次級代理程式(由虛線箭頭154概念性地圖解說明)。

其他實例包含充當一或多個初級代理程式之主機可見儲存空間110之該等元件之任何一或多者，及主機不可見儲存空間120之該等元件之任何一或多者，或充當一或多個次級代理程式之主機可見儲存空間110之其他元件之任何

一或多者。

圖 1B 圖解說明一種用於可擴展儲存器件之技術之另一實施例之選定結構細節，該等可擴展儲存器件包含一主機、具有操作為各別初級代理程式之一或多個儲存器件之主機可見儲存空間及具有操作為各別次級代理程式之一或多個儲存器件之主機不可見儲存空間。該技術使用在操作及結構上與圖 1A 之類似識別之元件一致之某些元件(主機 100、主機-儲存器件耦合件 180 及耦合件 101、111.A、111.N、121.A 及 121.N)。主機可見儲存空間 130 類似於圖 1A 之主機可見儲存空間 110，但其中之儲存器件(儲存器件 130.A 及儲存器件 130.N)各自經啟用以經由一額外耦合件(分別地，耦合件 131.A 及 131.N)通信。主機不可見儲存空間 140 類似於圖 1A 之主機不可見儲存空間 120，但其中之儲存器件(儲存器件 140.A 及儲存器件 140.N)經啟用以經由一額外耦合件(分別地，耦合件 141.A 及 141.N)通信。器件-儲存器件耦合件 181 耦合至耦合件 131.A、131.N、141.A 及 141.N，且使得主機可見儲存空間 130 與主機不可見儲存空間 140 之該等元件中之任一者之間的額外通信帶寬有利於(舉例而言)同級間通信。舉例而言，器件-儲存器件耦合件 181 實現儲存器件 130.A 與主機不可見儲存空間 140 之該等元件中之任一者之間的額外通信帶寬。針對另一實例，器件-儲存器件耦合件 181 實現儲存器件 130.A 與儲存器件 130.N 之間的額外通信帶寬。

主機-儲存器件耦合件與器件-儲存器件耦合件以及主機

可見儲存空間及主機不可見儲存空間係視情況實施為一可插式模組(圖解說明為添加式卡190B)及/或耦合件101視情況實施為一纜線。在某些實施例中，該添加式卡之所有或任何部分係實施為一SSD。在各種實施例中，該等儲存器件之任何一或多者係SSD。

在某些實施例及/或使用情境中，由器件-儲存器件耦合件181提供之額外通信帶寬實現經由同級間通信(諸如RAID 5)實現之儲存冗餘技術，而對主機與儲存器件之間的通信可用之帶寬幾乎或完全不影響。該同級間通信包含未轉換資料(諸如在一RAID鏡像操作期間)、已轉換資料(諸如在一RAID同位產生及/或寫入期間之XOR資料)及控制資訊之通信之一或多者。

在某些實施例中，藉由以一等量化的方式跨越複數個儲存器件(諸如正交於(非RAID)資料散佈之方向)散佈RAID資料來啟用RAID冗餘。在某些情境中，在讀取一主機請求時，自該等儲存器件之兩個或兩個以上讀取資訊。在一主機請求係一寫入時，維持RAID冗餘資訊視情況包含執行RMW操作或另一選擇係相對於該等儲存器件中之兩個或兩個以上之資料複製操作。

在一初級代理程式或一次級代理程式作為一主機寫入請求之一結果擷取資料時，該代理程式將一同級間請求(諸如經由器件-儲存器件耦合件181)連同特定RAID資料一起傳遞至另一代理程式(一初級代理程式或一次級代理程式)。在某些境況(諸如鏡像RAID)中，該特定RAID資料係

針對該寫入請求之資料之一複本。在某些境況(諸如 RAID 5)中，基於針對該寫入請求之資料及對應於該寫入請求之一 LBA 之先前資料(例如，將該先前資料用於一 RMW 操作中，諸如具有針對該寫入請求之資料之一 XOR)計算該特定 RAID 資料。

舉例而言，接收針對該寫入請求之資料之代理程式藉由用對應於該寫入請求之 LBA 之先前資料互斥「或」針對該寫入請求之資料來計算一 RAID 差異。該代理程式將 RAID 差異發送至儲存與該 LBA 相關聯之 RAID 冗餘之代理程式中之一者。儲存 RAID 冗餘之該代理程式然後執行一 RMW 以使用 RAID 差異更新該 RAID 冗餘。若儲存該 RAID 冗餘之代理程式已被告知期望對 RAID 冗餘之多個更新，則儲存該 RAID 冗餘之該代理程式視情況及/或選擇性地組合該等 RMW 操作供用於更大效率。

在某些實施例及/或使用情境中，一初級代理程式告知另一代理程式(另一初級代理程式或一次級代理程式)將針對特定 LBA 存在多少 RAID 更新。在各種實施例中，所告知之代理程式經啟用以快取關於該等 RAID 更新之資訊直至已執行針對該特定 LBA 之所有更新為止。在各種實施例中，RAID 更新之數目係隱含地或明確地作為一子請求之部分傳遞。在完成該等 RAID 更新時，所告知之代理程式將一子狀態發送至該初級代理程式。

在各種實施例中，完全獨立地或部分獨立地執行關於單個主機請求之複數個資料傳送。舉例而言，在某些 RAID

RMW情境中，寫入資料之傳送獨立於先前累積之同位資料之傳送。在各種情境中，由一初級代理程式及複數個次級代理程式中之任一者自主機記憶體讀取寫入資料，由該等初級代理程式及次級代理程式中之任一者自一儲存器件讀取該所累積同位資料，及然後將新同位資料寫入至(例如)自其讀取所累積之同位資料之儲存器件。

在各種實施例中，耦合件131.A、131.N、141.A及141.N中之一或多者與圖1A之耦合件111.A、111.N、121.A及121.N一致或大致類似。在某些實施例中，器件-儲存耦合件器件-儲存器件耦合件181在實施上與主機-儲存器件耦合件180一致或大致類似。在某些實施例中，將主機-儲存器件耦合件180及器件-儲存器件耦合件181組合成單個元件。在某些實施例中，耦合件131.A、131.N、141.A及141.N中之任何一或多者耦合至主機-儲存器件耦合件180之一變化形式之額外部分而非耦合至器件-儲存器件耦合件181。

在各種實施例中，諸如與圖1A或圖1B中之任一者有關之某些實施例，主機可見儲存空間具有操作為各別初級代理程式之一或多個儲存器件及操作為次級代理程式之零個或零個以上儲存器件，而主機不可見儲存空間具有操作為各別次級代理程式之零個或零個以上儲存器件。

在各種實施例中，一儲存器件對應於一實體儲存器件，且該實體儲存器件經啟用以根據一或多個操作模式實施主機可見儲存空間與主機不可見儲存空間之任一組合。在一

第一操作情境中，一特定實體器件經操作以實施僅主機可見儲存空間。在一第二操作情境中，該特定實體器件經操作以實施僅主機不可見儲存空間。在一第三操作情境中，該特定實體器件經操作以實施主機可見儲存空間與主機不可見儲存空間之一組合。在某些情況(諸如回應於工作負載之一改變及/或一故障)中，操作係逐操作情境地動態改變。

舉例而言，參照圖1A，操作如圖解說明地開始，其中主機可見儲存空間110具有儲存器件110.A及儲存器件110.N，且主機不可見儲存空間120具有儲存器件120.A及儲存器件120.N。繼續該實例，儲存器件110.A、110.N、120.A及120.N中之每一者對應於各別實體儲存器件。回應於工作負載之一改變，儲存器件110.N自操作為主機可見儲存空間之一元件動態地(重)組態至操作為主機不可見儲存空間之一元件。在該動態重組態之後，主機可見儲存空間110具有儲存器件110.A，且主機不可見儲存空間120具有儲存器件110.N、儲存器件120.A及儲存器件120.N(未圖解說明)。

針對另一實例，參照圖1B，操作如圖解說明地開始，其中主機可見儲存空間130具有儲存器件130.A及儲存器件130.N，且主機不可見儲存空間140具有儲存器件140.A及儲存器件140.N。繼續該實例，儲存器件130.A、130.N、140.A及140.N中之每一者對應於各別實體儲存器件。回應於儲存器件130.A之一故障，儲存器件140.A經動態(重)組

態以操作為主機可見儲存空間之一元件以及繼續操作為主機不可見儲存空間之一元件。舉例而言，由儲存器件140.A實施之儲存空間在該故障之前係在用作取代儲存器件130.A所實施之儲存空間之所有或任何部分之間分配，且繼續用作主機不可見儲存空間。在該動態重組態之後，主機可見儲存空間130具有由儲存器件140.A及儲存器件130.N實施之儲存空間之一部分，且主機不可見儲存空間140具有儲存器件140.A(少於用於取代儲存器件130.A之部分)及儲存器件140.N(未圖解說明)。

在各種實施例中，主機可見儲存空間係可經由一透明開關存取至一主機之儲存空間，及/或主機不可見儲存空間係經由一不透明開關自主機隱藏之儲存空間。該主機可見儲存空間及該主機不可見儲存空間係啟用用於同級間通信，例如經由該等透明及/或不透明開關之所有或任何部分。在某些實施例中，一整合開關提供分別對應於前述透明開關及不透明開關之一透明部分及一不透明部分。舉例而言，參照圖1B，主機100、耦合件101、主機-儲存器件耦合件180及器件-儲存器件耦合件181之所有或任何部分之一或多者實施一整合開關。該整合開關參照主機可見儲存空間130之該等元件之一或多者操作為一透明開關，且亦參照主機不可見儲存空間140之該等元件之一或多者同時操作為一不透明開關。

可擴展儲存系統操作

圖2圖解說明一種用於可擴展儲存器件之技術之一實施

例之選定處理細節。所圖解說明的係由充當一初級代理程式之一儲存器件(諸如圖1A之主機可見儲存空間110之該等元件之一或多者)執行為「初級代理程式動作」209之動作，以及由充當一次級代理程式之一儲存器件(諸如圖1A之主機不可見儲存空間120之該等元件之一或多者及/或主機可見儲存空間110之其他元件之任何一或多者)實施為「次級代理程式動作」219之動作。

首先闡述其中由單個對應代理程式滿足用以存取儲存空間之一請求之一情境，及然後闡述其中由多於一個對應代理程式滿足單個請求之情境。一主機(諸如圖1A之主機100)將用以存取儲存空間之一請求提供至充當一初級代理程式之一儲存器件(概念性地由「開始」201表示)。該初級代理程式自主機接受該請求，諸如藉由讀取主機記憶體中所保持之一提交佇列條目，視情況提供與該主機之信號交換以指示該請求已被無錯誤地接受(「自主機接受請求」202)。然後初級代理程式檢查該請求以判定該請求是否係針對由該初級代理程式實施之儲存空間(「本端？」203)。若是(「是」，203Y)，則流程進行，其中初級代理程式處理與該請求一起提供之位址資訊以將該位址資訊自與該請求相關聯之一(主機)上下文轉譯至與該初級代理程式實施之對應儲存空間相關聯之一(初級代理程式)上下文(「轉譯至本端LBA」204)。針對一實例，該請求指定LBA 100，且該初級代理程式將LBA 100轉譯至LBA 0，對應於由該初級代理程式實施之(本端)儲存空間之一開始位置。然後

該初級代理程式執行該請求(「本端地處理」205)，並將狀態提供至主機(「將狀態提供至主機」，206)，諸如藉由寫入保持於該主機記憶體中之完成佇列條目，因此完成該請求之處理(「結束」299)。在某些實施例中，該轉譯包括在該等LBA之一函數與本端儲存位址之間的一映射，類似於一頁表。

若該請求並非針對由初級代理程式實施之儲存空間(「否」，203N)，則在該初級代理程式內之流程進行至將該請求作為一子請求發送至充當一次級代理程式之一儲存器件(「轉發至次級代理程式」207)。該初級代理程式自該請求判定一子請求，且由虛線「子請求」207R圖解說明子請求自初級代理程式至次級代理程式之通信。在該初級代理程式內，流程然後進行至等待次級代理程式完成該子請求(「等待完成」208)。在該次級代理程式已完成該子請求(具有或不具有錯誤)時，次級代理程式將子狀態供應至初級代理程式(由虛線「子狀態」214S概念性地圖解說明)。然後初級代理程式接受來自次級代理程式之子狀態，視情況提供與次級代理程式之信號交換以指示已無錯誤地接受該子狀態(「接受來自次級代理程式之子狀態」209A)。然後初級代理程式進行至將該子狀態作為狀態提供至主機(「將狀態提供至主機」206)，且然後完成該請求之處理(「結束」299)。

在次級代理程式內，流程藉由接受來自初級代理程式之子請求而開始，視情況提供與初級代理程式之信號交換以

指示已無錯誤地接受該子請求(「接受來自初級代理程式之子請求」211)。該子請求包含由主機在該子請求所對應之請求中所提供之位址資訊(諸如一LBA)。次級代理程式處理與該子請求一起提供之位址資訊以將該位址資訊自與該請求相關聯之一(主機)上下文轉譯至與該次級代理程式所實施之對應儲存空間相關聯之一(次級代理程式)上下文(「轉譯至本端LBA」212)。舉例而言，一請求指定LBA 200，子請求指定LBA 200，且次級代理程式將(主機)LBA 200轉譯至(本端)LBA 0，對應於該次級代理程式所實施之(本端)儲存空間之一開始位置。繼續該實例，另一請求指定LBA 201，一對應子請求指定LBA 201，且該次級代理程式將(主機)LBA 201轉譯至(本端)LBA 1，對應於該次級代理程式所實施之(本端)儲存空間之下一位置。接下來在圖3之說明中闡述其他實例。在轉譯該子請求之後，流程則進行至其中次級代理程式執行該請求(「本端地處理」213)且然後將子狀態返回至初級代理程式(「將子狀態提供至初級代理程式」214)，如虛線「子狀態」214S所圖解說明。

在其中由多於一個回應代理程式滿足用以存取儲存空間之單個請求之一情境中，將該單個請求概念性地分成複數個子請求。該等子請求之零個、一個或多個係由初級代理程式滿足，且該等子請求之零個、一個或多個係由零個、一個或多個次級代理程式滿足。舉例而言，針對LBA 400之一請求(具有兩個LBA之一長度)係由針對LBA 400之一

次級代理程式之一子請求與針對LBA 401之另一次級代理程式之另一子請求滿足。每一次級代理程式將該各別子請求之LBA獨立地轉譯成對應於該各別次級代理程式之本端儲存空間之一LBA。

該初級代理程式基於該請求之位址及長度資訊判定哪些次級代理程式(若存在)欲接收子請求。因此在圖2中，判定該請求是否係針對由初級代理程式實施之儲存空間(「本端？」203)係針對複數個子請求中之每一者概念性地判定。針對該等子請求中之某些子請求，結果係該等子請求係本端的，而針對該等子請求之其他，結果係其他子請求並非本端的，且係作為子請求轉發至一或多個次級代理程式(「轉發至次級代理程式」207及「子請求」207R)。在某些情境中，單個次級代理程式接收單個子請求，而在其他情境中，該單個次級代理程式接收複數個子請求。然後針對該複數個子請求中之每一者執行在「次級代理程式動作」219中表示之流程(在各種實施例中，在串列及/或並行處理之任一組合中)。在某些情境中，複數個次級代理程式中之每一者接收一或多個子請求。然後由該複數個次級代理程式中之每一者獨立地執行「次級代理程式動作」219中表示之流程。

該初級代理程式判定如何作為在「轉發至次級代理程式」207中執行之處理之部分將單個請求映射成對一或多個次級代理程式之複數個子請求，諸如經由一或多個函數及/或映射表(例如，在圖1A之儲存器件110.A中實施)。在

將複數個子請求發送至一或多個次級代理程式時，該初級代理程式藉由針對獨立於其他子請求之子請求中之每一者之等待完成(「等待完成」208)及接受對應子狀態(「接受來自次級代理程式之子狀態」209A)來概念性地累積狀態資訊。在已接收一特定請求之針對所有該等子請求之子狀態時，該初級代理程式產生一總狀態以作為在「將狀態提供至主機」206中執行之處理之部分提供至該主機。

在某些實施例中，次級代理程式接收一請求之整體作為一致的子請求(或另一選擇係，讀取來自一提交佇列中之一條目之請求)。該等次級代理程式中之每一者然後判定各別次級代理程式欲執行該請求之哪一部分(若存在)，並視情況判定欲將累積之同位資料(若存在)發送至哪一代理程式。舉例而言，一請求指定具有200個LBA之一長度之LBA 500。一初級代理程式將一致的子請求發送至第一及第二次級代理程式(或另一選擇係，指向至該第一次級代理程式及第二次級代理程式之單個子請求)。該(等)子請求指定具有200個LBA之一長度之LBA 500。該第一次級代理程式(獨立於主機及第二次級代理程式)判定該子請求之LBA 600至699對應於第一次級代理程式欲執行之該子請求之一部分。該第二次級代理程式(獨立於主機及第二次級代理程式)判定該子請求之LBA 500至599對應於該第二次級代理程式欲執行之該子請求之一部分。相依於一特定冗餘方案，第一及第二次級代理程式判定所累積之同位資訊係該等代理程式之一適合者。

在某些實施例及/或使用情境中，由一次級代理程式將狀態直接提供至主機，而不是透過一初級代理程式轉發(由圖1A之虛線箭頭154概念性地圖解說明)。舉例而言，狀態之直接提供出現在將一請求整體之子請求轉發至單個次級代理程式時。然後該單個次級代理程式將狀態直接返回至主機(舉例而言，由該單個次級代理程式將一條目寫入至保持於主機記憶體中之一完成佇列)。應注意，主機不能區分寫入該條目之單個次級代理程式與寫入該條目之初級代理程式之動作，且因此即使該單個次級代理程式寫入該條目，該單個次級代理程式滿足該請求亦對主機透明。在某些實施例中，一次級代理程式(而不是一初級代理程式)累積對應於一特定請求之複數個子請求之子狀態並將所累積之子狀態作為一總狀態提供至主機。

在某些實施例中，一或多個儲存器件至少相對於來自一或多個主機之某些請求同時作為一初級代理程式及作為一次級代理程式操作。舉例而言，一系統具有兩個儲存器件。該等儲存器件之第一者相對於該等儲存器件之操作為一第二代理程式之第二者操作為一初級代理程式，且該第二儲存器件相對於操作為一次級代理程式之該第一儲存器件操作為一初級代理程式。更具體而言，該第一儲存器件接受來自該等主機中之一特定者之一請求並將該請求作為一子請求發送至第二儲存器件。該第二儲存器件自該特定主機接受另一請求並將該另一請求作為一子請求轉發至第一儲存器件。在某些實施例及/或使用情境中，該等請求

係來自同一主機請求佇列。在某些實施例及/或使用情境(諸如其中該等請求係來自相異主機)中，該等請求係來自相異主機請求佇列。

根據各種實施例，請求、狀態及資料傳送中之一或多者係經由(舉例而言)可存取至主機(諸如主機記憶體)(或作為其部分)之記憶體與可存取至該等儲存器件(諸如實施於一儲存器件中之一緩衝器)中之一者(或其部分)之間的DMA。因此，請求、狀態及資料傳送之前述流動中之一或多者係藉由一或多個分散/集中清單與根據該等分散/集中清單之後續DMA傳送之通信。另外，請求、狀態及資料傳送之該等流動之一或多者係藉由完成及/或狀態佇列資訊與至/自該等佇列之條目之後續DMA傳送之通信。在各種實施例中，該請求、該狀態及該等資料傳送中之任一者係直接執行及/或由該等儲存器件之一或多者中之一或多個處理器管理。

定址

圖3圖解說明一種用於可擴展儲存器件之技術之一實施例中之主機、初級代理程式及次級代理程式定址之選定細節。該圖圖解說明在由一主機(諸如圖1A之主機100)查看時儲存位址空間(主機位址空間310)之實例性窗口。該主機儲存位址空間具有主機LBA範圍1至5 311至315。該圖進一步圖解說明在由一初級代理程式(諸如圖1A之主機可見儲存空間110之該等元件中之任一者)查看時儲存位址空間(初級代理程式位址空間320)之實例性窗口。該初級代理程式

儲存位址空間具有初級代理程式LBA範圍1至3 321至323。該圖進一步圖解說明在由兩個次級代理程式(次級代理程式A及B，諸如圖1A之主機不可見儲存空間120之該等元件之任何兩者及/或主機可見儲存空間110之其他元件中之任一者)查看時儲存位址空間(次級代理程式位址空間330)之實例性窗口。次級代理程式A儲存位址空間具有次級代理程式A LBA範圍1至3 331A、332A及333A。次級代理程式B儲存位址空間具有次級代理程式B LBA範圍1至2 331B及332B。儘管未繪製至任一特定比例，但在該圖式中各種位址範圍元件之垂直高度變化，概念性地表示在該等範圍中之每一者中存在多少位置之變化形式。

在一第一實例中，單個主機位址範圍(主機LBA範圍1 311)對應於次級代理程式A之單個位址範圍(次級代理程式A LBA範圍1 331A)。一初級代理程式(諸如圖1A之儲存器件110.A)判定對主機LBA範圍1 311內之一位址之一主機請求對應於並未由初級代理程式實施之儲存空間(諸如經由與圖2之「本端？」203及「否」203N相關聯之處理)，且該請求對應於由次級代理程式A實施之儲存空間。該初級代理程式然後將一子請求轉發至次級代理程式A。次級代理程式A然後轉譯該子請求(諸如與圖2之「轉譯至本端LBA」212相關聯之處理)，及執行對可經由次級代理程式A LBA範圍1 331A存取之儲存空間之一本端存取(諸如經由圖2之「本端地處理」213進行處理)。

在一第二實例中，單個主機位址範圍(主機LBA範圍2

312)對應於次級代理程式B之單個位址範圍(次級代理程式B LBA範圍1 331B)。該初級代理程式判定對主機LBA範圍2 312內之一位址之一請求對應於由次級代理程式B實施之儲存空間，且然後將一子請求轉發至次級代理程式B。次級代理程式B然後轉譯該子請求並執行對可經由次級代理程式B LBA範圍1 331B存取之儲存空間之一本端存取。

在一第三實例中，單個主機位址範圍(主機LBA範圍3 313)對應於該主機代理程式之單個位址範圍(初級代理程式LBA範圍1 321)。該初級代理程式判定在主機LBA範圍3 313內之一位址之一請求對應於由初級代理程式實施之儲存空間，轉譯該請求(諸如經由與「轉譯至本端LBA」204相關聯之處理)及執行對可經由初級代理程式LBA範圍1 321存取之儲存空間之一本端存取。

在一第四實例中，單個主機位址範圍(主機LBA範圍4 314)對應於次級代理程式A之兩個位址範圍(次級代理程式A LBA範圍2 332A及3 333A)及次級代理程式B之單個位址範圍(次級代理程式B LBA範圍2 332B)。該初級代理程式判定對主機LBA範圍4 314之LBA整體之一請求(例如，基於一開始LBA及欲存取之連續LBA之一數目)對應於由次級代理程式A及次級代理程式B實施之儲存空間。然後該初級代理程式將兩個子請求轉發至次級代理程式A及將單個子請求轉發至次級代理程式B。次級代理程式A轉譯該兩個子請求並執行對可經由次級代理程式A LBA範圍2 332A及3 333A存取之儲存空間之本端存取(系列地或整個地或

部分並行地)。次級代理程式B轉譯該單個子請求並執行對可經由次級代理程式B LBA範圍2 332B存取之儲存空間之一本端存取。

在一第五實例中，單個主機位址範圍(主機LBA範圍5 315)對應於該初級代理程式之兩個位址範圍(初級代理程式LBA範圍2 322及3 323)。該初級代理程式判定對主機LBA範圍5 315之LBA整體之一請求(例如，基於一開始LBA及欲存取之連續LBA之一數目)對應於由初級代理程式所實施之儲存空間。該初級代理程式將該請求轉譯成兩個本端LBA範圍，並執行對可經由初級代理程式LBA範圍2 322及3 323存取之儲存空間之本端存取。

在某些實施例中，一初級代理程式經啟用以將一主機請求之每一子請求單獨地轉發至一或多個次級代理程式之特定一者。在其他實施例中，一初級代理程式經啟用以將一主機請求之兩個或兩個以上作為單個組合子請求轉發至複數個次級代理程式之特定一者。舉例而言，在各種實施例中，一初級代理程式判定一或多個次級代理程式以處理一主機請求之一或多個子請求並將具有一相同LBA範圍及一相同長度之主機請求轉發至該一或多個次級代理程式。該一或多個次級代理程式經啟用以隨之解譯該LBA範圍及該長度以判定該一或多個次級代理程式中之每一者欲處理及回應於之主機請求之各別部分。在其他實施例中，一初級代理程式處理一LBA範圍及一主機請求之一長度並將該LBA範圍及該長度之一經處理版本發送至一或多個次級代

理程式之每一者。舉例而言，若資料在64 KB之邊界上之N個代理程式(N-1個次級代理程式及該初級代理程式)間等量化，該初級代理程式將該LBA範圍除以 $N \times 64 \text{ KB}$ 以將子請求分配至該N-1個次級代理程式。該N-1個次級代理程式中之每一者接收具有一已除降之開始LBA、對應於資料在該N個代理程式間等量化的長度之一部分及視情況及/或可選地自除降該LBA所得之一剩餘部分之一各別子請求；因此該等次級代理程式之一第一者及/或最後一者視情況處理小於64 KB之一傳送。

可擴展儲存器件

圖4圖解說明經啟用以操作為一初級代理程式(具體而言，圖1A之儲存器件110.A)之一可擴展儲存器件之一實施例之選定結構細節。該儲存器件包含對一PCIe匯流排之一介面(PCIe介面401)。該PCIe介面包含一DMA單元以實現直接在主機記憶體與PCIe介面之緩衝器之間的可選及/或選擇性傳送，舉例而言以讀取/寫入完成/狀態佇列條目及/或直接地傳送資料自/至主機記憶體。該PCIe介面耦合至一轉譯區塊(LBA，長度轉譯邏輯402)。該轉譯區塊經啟用以將主機儲存位址資訊(例如，LBA及長度)轉譯成本端儲存位址資訊，諸如參照圖2之「轉譯至本端LBA」204所闡述。該轉譯區塊耦合至一大容量儲存器介面(儲存器介面邏輯403)，該大容量儲存器介面又耦合至一或多個大容量儲存器單元(一起圖解說明為大容量儲存器404)。該等大容量儲存器單元包含一快閃儲存器單元(快閃404F)、一磁碟

儲存器單元(磁性件404M)、一光碟儲存器單元(光學件404O)及/或任一類型之非揮發性儲存器單元中之任何一或多者。在各種實施例中，該快閃儲存器單元包括一或多個快閃記憶體晶片及/或晶粒，諸如NAND快閃或NOR快閃。

儲存器件110.A包含一控制區塊，其經啟用以執行關於作為視情況耦合至PCIe介面401、LBA、長度轉譯邏輯402及儲存器介面邏輯403中之一或多者之一初級代理程式(子請求產生及子狀態累積405)之操作之各種操作。該等操作包含(舉例而言)實施及/或管理與圖2之「初級代理程式動作」209之該等元件之任何一或多者相關聯之處理之所有或任何部分。在某些實施例中，子請求產生及子狀態累積405實施一或多個函數及/或映射表，該一或多個函數及/或映射表追蹤如何將主機位址分配至各種次級器件(諸如由圖2之「轉發至次級器件」207所使用之映射表)。該等操作進一步包含回應於主機針對儲存容量之詢問，好似一或多個次級代理程式所實施之儲存容量(諸如圖5之儲存器件120.A)係由儲存器件110.A實施。

圖5圖解說明經啟用以操作為一次級代理程式(諸如圖1A之儲存器件120.A)之一可擴展儲存器件之一實施例之選定結構細節。該儲存器件包含在操作及結構上與圖4之類似識別元件(PCIe介面401、LBA、長度轉譯邏輯402、儲存介面邏輯403及具有元件快閃404F、磁性件404M及404O之大容量儲存器404)一致之某些元件。

儲存器件120.A包含一控制區塊，其經啟用以執行關於

作為視情況耦合至PCIe介面401、LBA、長度轉譯邏輯402及儲存介面邏輯403中之一或多者之一次級代理程式(子請求接受及子狀態產生501)之操作之各種操作。該等操作包含(舉例而言)實施及/或管理與圖2之「次級代理程式動作」219之該等元件之任何一或多者相關聯之處理之所有或任何部分。該等操作進一步包含回應於主機針對儲存容量之詢問，好似不存在由儲存器件120.A實施之儲存空間，或另一選擇係，好似用於由一主機經由一初級代理程式(而非直接由主機)存取而分配之儲存空間並未由該儲存器件實施。相反地，由儲存器件120.A實施之儲存空間係由一初級代理程式報告，諸如圖4之儲存器件110.A。

在某些實施例中(未圖解說明)，一可組態儲存器件包含對應於在圖4之儲存器件110.A中圖解說明之所有該等元件之功能的元件，且進一步包含對應於圖5之子請求接受及子狀態產生501之功能的元件。一實例性可組態儲存器件包含PCIe介面401、LBA、長度轉譯邏輯402、儲存介面邏輯403及大容量儲存器404以及子請求產生及子狀態累積405與子請求接受及子狀態產生501。該可組態儲存器件經啟用以(舉例而言)回應於可程式化之所包含模式狀態選擇性地操作為一初級代理程式或一次級代理程式。該模式狀態可藉由各種技術來程式化及/或替換，諸如由管理軟體使用特殊(例如，賣方專用)命令經由所包含之CPIe介面、經由諸如非揮發性記憶體之一初始元件(例如，快閃記憶體或一選項ROM)或經由對可組態儲存器件及/或與該可組

態儲存器件連通之另一儲存器件之一事件之偵測。在某些實施例及/或使用情境中，一可組態儲存器件實現操作模式之一動態改變，諸如一次級代理程式動態地取代一出故障之初級代理程式，因此消除儲存空間作為單個故障點。

故障情境

在各種實施例及/或使用情境中啟用數個故障恢復技術，其中一初級代理程式以及一或多個次級代理程式對一主機顯現為單個邏輯介面。該等恢復技術對主機完全透明，或僅對在該主機上執行之一器件驅動器可見。

在一次級代理程式完全地或部分地出故障時，可應用一第一故障恢復技術，以使得出故障的次級代理程式不再實施一特定儲存空間部分。該初級代理程式變得知曉該故障(例如，該初級代理程式偵測該故障，或該出故障的次級代理程式將該故障報告給該初級代理程式)。作為回應，該初級代理程式經由自該初級代理程式及/或自該等次級代理程式之零個或零個以上之儲存空間(視情況及/或可選擇地包含該出故障之次級代理程式)(重)分配儲存空間以實施該特定儲存空間部分。該初級代理程式及該等次級代理程式然後根據參照圖 1A、圖 1B 及圖 2 至圖 5 闡述之前述實施例及/或情境中之任一者操作。

在某些情況中，諸如在一次級代理程式出故障使得該次級代理程式不能向該重新分配貢獻儲存空間時，該重新分配排除該出故障的次級代理程式。該次級代理程式出於各種原因在各種情境中不能貢獻儲存空間，諸如不具有可操

作的足夠自由儲存空間、不能正確地與初級代理程式及/或主機通信、實體移除或總故障。在其他情況中，諸如在一次級代理程式部分地出故障且繼續實施至少某些儲存空間時，該重新分配視情況包含該部分出故障之第二代理程式之儲存空間。在替代之相關實施例及/或使用情境中，在該主機上執行之一器件驅動器變得知曉該故障(而非該初級代理程式)並通知該初級代理程式該故障及/或指示該初級代理程式執行一重新分配。

在一初級代理程式整個地或部分地出故障時可應用一第二故障恢復技術，以使得該初級代理程式不再實施一特定儲存空間部分。該第二故障恢復技術類似於第一故障恢復技術，但該重新分配係參照不再由整個地或部分地出故障之初級代理程式實施之儲存空間。如在第一故障恢復技術中，欲正確地實施儲存空間之初級代理程式之故障可由該初級代理程式、一次級代理程式或一主機上之一器件驅動器中之一或多者偵測。

在一初級代理程式不能正確地充當一初級代理程式時可應用一第三故障恢復技術，例如與一或多個次級代理程式及/或一主機傳遞資訊(例如，子請求及/或子狀態)。一監測代理程式變得知曉該故障且作為回應識別另一初級代理程式或次級代理程式以充當一取代初級代理程式。在某些實施例中，該取代初級代理程式係一備用磁碟。該監測代理程式之一實例係包含該出故障之初級代理程式之一儲存器件之另一初級(或次級)代理程式。該監測代理程式之另

一實例係一主機上之一器件驅動器，諸如實現在該主機上執行之一OS及/或應用程式與在該故障之前的出故障之初級代理程式之間的通信之一器件驅動器。若該出故障之初級代理程式故障亦導致該出故障之初級代理程式不再實施一特定儲存空間部分，則該取代初級代理程式重新分配儲存空間以實施該特定儲存空間部分，諸如以該第二故障恢復技術。該重新分配係經由對以下各項中之任何一或多者之儲存空間之分配：該出故障之初級代理程式(若任何儲存空間保持在出故障之初級代理程式中可操作)、該取代初級代理程式及/或零個或零個以上該等次級代理程式。該取代初級代理程式充當在該故障之前的出故障之初級代理程式，且根據前述實施例及/或參照圖1A、圖1B及圖2至圖5所述之情境中之任一者來操作。

一第四故障恢復技術係在具有一或多個可組態代理程式之一系統中啟用，且可在一初級代理程式不能正確地充當一初級代理程式時應用。該可組態代理程式之一實例係可操作為一初級代理程式或一次級代理程式之一可組態儲存器件，諸如基於概念性地包含圖4及圖5之所有元件之一可組態儲存器件。類似於該第三故障恢復技術，回應於一監測代理程式變得知曉正確地充當一初級代理程式之一初級代理程式之故障，該監測代理程式識別該等可組態代理程式中之一者以充當一取代初級代理程式。該監測代理程式請求所識別之可組態代理程式經組態(例如，經由程式化模式資訊)以操作為一取代初級代理程式。所識別之可組

態元件動態地回應於該組態改變請求，並開始充當一初級代理程式以執行為一取代初級代理程式。類似於第三故障恢復技術，若該出故障之初級代理程式故障亦導致該出故障之初級代理程式不再實施一特定儲存空間部分，則該取代初級代理程式重新分配儲存空間以實施該特定儲存空間部分，諸如以第二故障恢復技術。該取代初級代理程式充當在該故障之前的出故障之初級代理程式，且根據前述實施例及/或參照圖1A、圖1B及圖2至圖5所述之情境中之任一者來操作。

在前述故障恢復技術中之任何一或多者中，在一故障之後，視情況經由冗餘資訊盡可能地恢復資料，該資料恢復係視情況經由一主機側耦合件(例如，圖1A或圖1B之主機-儲存器件耦合件180)與一器件側耦合件(例如，圖1B之器件-儲存器件耦合件181)之任一組合。如本文其他處所闡述，該冗餘資訊係根據(例如)各種RAID及/或鏡像實施方案。

在前述故障恢復技術中之任何一或多者中，該重新分配視情況減少可用之自由空間，例如如向主機所呈現。

在前述故障恢復技術中之任何一或多者中，該重新分配之任何一或多者係至少部分地由執行指令之一處理器執行及/或管理。

在前述故障恢復技術中之任何一或多者中，一初級代理程式或一次級代理程式之一故障之辨識係經由以下各項中之一或多者：一代理程式之一故障之一具體偵測、由一代

理程式接收一活動訊號指示之一故障及一故障已發生之一較高位準指示。該具體偵測包含另一代理程式已出故障之一初級代理程式、一次級代理程式或一器件驅動器判定，例如經由判定一所返回狀態係不合適的、一協定已被違反或一請求已超時。

在其中一透明開關實現一主機與主機可見儲存空間之間的通信之某些實施例中，該主機可見儲存空間之一元件之一故障對(例如)在該主機上執行之一器件驅動器可見。該器件驅動器視情況參與自該故障恢復。在其中在一主機與主機不可見儲存空間之間具有一不透明開關之某些實施例中，該主機不可見儲存空間之一元件之一故障對該主機不可見。自該故障恢復係視情況對主機不可見。

在各種實施例中，由圖1A及圖1B之主機可見儲存空間110、主機不可見儲存空間120、主機可見儲存空間130及主機不可見儲存空間140以及其中之單元及次單元執行之操作之所有或任何部分係由執行儲存於一視情況有形之電腦可讀媒體中之指令之一或多個處理器(諸如韌體)管理及/或實施。舉例而言，子請求產生及子狀態累積405及/或子請求接受及子狀態產生501係由執行儲存於一電腦可讀媒體中之指令之一處理器(諸如韌體)管理及/或實施。繼續該實例，該電腦可讀媒體係儲存於該可擴展儲存器件中之快閃記憶體之一部分(諸如快閃404F)。在各種實施例中，由圖1A及圖1B之主機可見儲存空間110、主機不可見儲存空間120、主機可見儲存空間130及主機不可見儲存空間140

之該等元件中之任一者實施之操作之所有或任何部分以及其中的單元及次單元係由一或多個狀態機管理及/或實施。舉例而言，由子請求產生及子狀態累積405及/或子請求接受及子狀態產生501執行之操作之所有或任何部分係由至少部分地由邏輯閘電路實施之一或多個狀態機管理及/或實施。

在各種實施例中，具有一或多個可擴展儲存器件(經啟用以操作為初級代理程式及/或次級代理程式)之一可擴展儲存系統之所有或任何部分係在一設備架之一或多個抽屜及/或擱板中(諸如在一資料中心中)實施。該可擴展儲存系統視情況經啟用以與一或多個主機一起操作，諸如在該設備架之其他抽屜及/或擱板中實施之計算複合體。

實例性實施技術

在某些實施例中，由一可擴展儲存器件執行之操作或包含於一可擴展儲存器件中之區塊之全部或部分之各種組合(諸如圖4之儲存器件110.A或圖5之儲存器件120.A)及一處理器、微處理器、系統單晶片、專用積體電路、硬體加速器或提供前述操作或區塊之全部或部分之其他電路之部分係由與一電腦系統之處理相容之一規範指定。該規範係根據各種說明，諸如硬體說明語言、電路說明、網路連線表說明、遮罩說明或佈局說明。實例性說明包含：Verilog、VHDL、SPICE、SPICE變體(諸如PSpice)、IBIS、LEF、DEF、GDS-II、OASIS或其他說明。在各種實施例中，該處理包含解譯、編譯、模擬及合成之任一組合以產生、驗

證或指定適合用於包含於一或多個積體電路上之邏輯及/或電路。根據各種實施例，每一積體電路可根據各種技術來設計及/或製造。該等技術包含一可程式化技術(諸如一場可程式化閘極陣列積體電路或遮罩可程式化閘極陣列積體電路)、一半定製技術(諸如一完全或部分地基於胞之積體電路)及一完全定製技術(諸如大致特定之一積體電路)、其任一組合或與積體電路之設計及/或製造相容之任一其他技術。

在某些實施例中，由具有儲存於其中之一組指令之一電腦可讀媒體闡述之操作之全部或部分之各種組合係藉由一或多個程式指令之執行及/或解譯、藉由一或多個源語言及/或指令碼語言陳述式之解譯及/或編譯或藉由執行藉由編譯、轉譯及/或解譯在程式化及/或指令語言陳述式中表達之資訊而產生之二進制指令來執行。該等陳述式係與任一標準程式化或指令語言(諸如C、C++、Fortran、Pascal、Ada、Java、VBscript及Shell)相容。該等程式指令、語言陳述式或二進制指令中之一或多者係視情況儲存於一或多個電腦可讀儲存媒體元件上。在各種實施例中，該等程式指令之某些、所有或各種部分係實現為一或多個功能、常式、次常式、內嵌常式、程序、巨集或其部分。

總論

已僅出於便於比較該文字與圖式之目的而在說明中做出某些選擇，且除非存在對相反內容之一指示，否則不應將該等選擇本身視為傳達針對所述實施例之結構或操作之額

外資訊。該等選擇之實例包含：用於圖式編號之設計之特定組織或指派及用於識別及引用該等實施例之特徵及元件之元件識別符(例如圖說文字或數字指示符)之特定組織或指派。

措詞「包含(includes或including)」係明確地意欲被視為闡述開放範疇之邏輯組之抽象詞，且除非明確地後跟有措詞「在...內」否則不意欲表達實體內含。

儘管已出於清晰說明及理解之目的在某些細節上闡述前述實施例，但本發明並不限於所提供之細節。存在本發明之諸多實施例。該等所揭示實施例僅係實例性而非約束性的。

應瞭解，構造、配置及使用之諸多變化形式可能與此說明一致，且在所發佈之專利之申請專利範圍之範疇內。舉例而言，互連及功能單元位元寬度、時鐘速度及所使用技術類型係根據每一組件區塊中之各種實施例可變的。賦予互連及邏輯之名稱僅係實例性的，且不應被視為限制所闡述之該等概念。流程圖及流程圖過程、動作及功能元件之次序及配置根據各種實施例係可變的。而且，除非明確地提及為相反，否則所指定之值範圍、所使用之最大及最小值或其他特定規範(諸如快閃記憶體技術類型；及暫存區與緩衝器中之條目或級之數目)僅係所述實施例之彼等，期望追蹤實施技術中之改良及改變且不應被視為限制。

此項技術中習知之在功能上等價之技術可替代所闡述之彼等技術來採用以實施各種組件、子系統、操作、功能、

常式、次常式、內嵌常式、程序、巨集或其部分。亦應瞭解，諸多實施例之功能態樣可在硬體(亦即，通常為專用電路)或軟體(亦即，經由某一方式之經程式化控制器或處理器)中選擇性地實現，作為實施例相依之設計約束及較快處理之技術趨勢(促進先前在硬體中之功能遷移至軟體中)及較高整合密度(促進先前在軟體中之功能遷移至硬體中)。各種實施例中之特定變化形式包含但不限於：分割差異；不同的形狀因數及組態；不同作業系統及其他系統軟體之使用；不同介面標準、網路協定或通信鏈路之使用；及在根據一特定應用程式之獨特工程及商務約束實施本文所述概念時將期望之其他變化形式。

已藉助超出所述實施例之諸多態樣之一最小實施方案所要求之細節及環境上下文來闡述該等實施例。熟習此項技術者將瞭解到，某些實施例在不變更該等剩餘元件間的基本合作之情況下省略所揭示之組件及特徵。因此，應瞭解，所揭示之多數細節並不要求實施所闡述之該等實施例之各種態樣。針對剩餘元件與先前技術可區分之程度而言，所省略之組件及特徵並非限定本文所闡述之概念。

所有此等設計之變化形式係在所述實施例傳達之教示上之非實質性改變。亦應瞭解，本文所述實施例對其他計算及網路連結應用程式具有廣泛的可應用性，且不限於所述實施例之特定應用或產業。因此本發明應視為包含涵蓋於所發佈專利之申請專利範圍之範疇內之所有可能修改及變化形式。

【圖式簡單說明】

圖 1A 圖解說明一種用於可擴展儲存器件之技術之一實施例之選定結構細節，該等可擴展儲存器件包含一主機、具有操作為各別初級代理程式之一或多個儲存器件之主機可見儲存空間及具有操作為各別次級代理程式之一或多個儲存器件之主機不可見儲存空間；

圖 1B 圖解說明一種用於可擴展儲存器件之技術之另一實施例之選定結構細節，該等可擴展儲存器件包含一主機、具有操作為各別初級代理程式之一或多個儲存器件之主機可見儲存空間及具有操作為各別次級代理程式之一或多個儲存器件之主機不可見儲存空間；

圖 2 圖解說明一種用於可擴展儲存器件之技術之一實施例之選定處理細節，包含由一初級代理程式及一或多個次級代理程式執行之動作；

圖 3 圖解說明一種用於可擴展儲存器件之技術之一實施例中之主機、初級代理程式及次級代理程式定址之選定細節；

圖 4 圖解說明經啟用以操作為一初級代理程式之一可擴展儲存器件之一實施例之選定結構細節；及

圖 5 圖解說明經啟用以操作為一次級代理程式之一可擴展儲存器件之一實施例之選定結構細節。

【主要元件符號說明】

100	主機
101	耦合件
110	主機可見儲存空間

110.A	儲存器件
110.N	儲存器件
111.A	耦合件
111.N	耦合件
120	主機不可見儲存空間
120.A	儲存器件
120.N	儲存器件
121.A	耦合件
121.N	耦合件
130	主機可見儲存空間
130.A	儲存器件
130.N	儲存器件
131.A	耦合件
131.N	耦合件
140	主機不可見儲存空間
140.A	儲存器件
140.N	儲存器件
141.A	耦合件
141.N	耦合件
151	虛線箭頭
152	虛線箭頭
153	虛線箭頭
154	虛線箭頭
180	主機-儲存器件耦合件

181	器 件 - 儲 存 器 件 耦 合 件
190A	添 加 式 卡
190B	添 加 式 卡
201	開 始
202	自 主 機 接 受 請 求
203	本 端 ？
203N	否
203Y	是
204	轉 譯 至 本 端 邏 輯 區 塊 位 址
205	本 端 地 處 理
206	將 狀 態 提 供 至 主 機
207	轉 發 至 次 級 代 理 程 式
207R	子 請 求
208	等 待 完 成
209A	接 受 來 自 次 級 代 理 程 式 之 子 狀 態
209	初 級 代 理 程 式 動 作
211	接 受 來 自 初 級 代 理 程 式 之 子 請 求
212	轉 譯 至 本 端 邏 輯 區 塊 位 址
213	本 端 地 處 理
214	將 子 狀 態 提 供 至 初 級 代 理 程 式
214S	子 狀 態
219	次 級 代 理 程 式 動 作
299	結 束
310	主 機 位 址 空 間

311	主機邏輯區塊位址範圍 1
312	主機邏輯區塊位址範圍 2
313	主機邏輯區塊位址範圍 3
314	主機邏輯區塊位址範圍 4
315	主機邏輯區塊位址範圍 5
320	初級代理程式位址空間
321	初級代理程式邏輯區塊位址範圍 1
322	初級代理程式邏輯區塊位址範圍 2
323	初級代理程式邏輯區塊位址範圍 3
330	次級代理程式位址空間
331A	次級代理程式 A 邏輯區塊位址範圍 1
331B	次級代理程式 B 邏輯區塊位址範圍 1
332A	次級代理程式 A 邏輯區塊位址範圍 3
332B	次級代理程式 B 邏輯區塊位址範圍 2
333A	次級代理程式 A 邏輯區塊位址範圍 2
401	高速周邊組件互連介面
402	長度轉移邏輯
403	儲存器介面邏輯
404	大容量儲存器
404F	快閃
404M	磁性件
404O	光學件
405	子請求產生及子狀態累積
501	子請求接受及子狀態產生

七、申請專利範圍：

102年8月19日修正本

1. 一種提供可擴展儲存之方法，其包括：

將一第一儲存器件操作為一初級代理程式及將至少一第二儲存器件操作為一或多個次級代理程式，該等初級代理程式及次級代理程式中之每一儲存器件具有至少一個各別儲存範圍；

該初級代理程式提供用於與一主機之通信之一邏輯介面及使得該主機能夠作為單一邏輯器件存取散佈於該等初級代理程式及次級代理程式之該等儲存器件上之該等儲存範圍之一彙總，與該主機之該等初級代理程式通信包括由該初級代理程式自該主機接受之儲存空間存取請求及由該初級代理程式發送至該主機之總狀態；

該初級代理程式將接受為一或多個子請求之每一請求轉發至該等次級代理程式中之所有或任何部分，自該等子請求所轉發至之該等次級代理程式接受與該等子請求相關聯之子狀態，且公式化至少基於該等子狀態而發送之每一總狀態；及

該等次級代理程式之該等所有或任何部分接受該等子請求，將該等子請求之主機上下文定址資訊轉譯成本端上下文定址資訊，至少部分地基於該本端上下文定址資訊存取本端儲存空間，且將該等子狀態提供至該初級代理程式。

2. 如請求項1之方法，其進一步包括：

該初級代理程式至少部分地基於該等初級代理程式及

次級代理程式之該等儲存範圍判定該單一邏輯器件之一彙總儲存容量；

回應於該主機針對儲存容量查詢該單一邏輯器件，該初級代理程式報告該彙總儲存容量；及

回應於該主機針對儲存容量查詢該等次級代理程式，該等次級代理程式中之每一者報告一各別儲存容量，該各別儲存容量排除在該彙總儲存容量中表示之該次級代理程式之任一儲存範圍之表示。

3. 如請求項1之方法，其進一步包括：

回應於該主機針對儲存容量查詢該等次級代理程式中之特定一者，該特定次級代理程式報告回一特定儲存容量，該特定儲存容量排除在由該初級代理程式所判定之一彙總儲存容量中表示之由該特定次級代理程式所實施之儲存空間之彼等部分之表示。

4. 如請求項1之方法，其進一步包括：

根據一主機介面協定執行與該主機之該等初級代理程式通信中之至少某些通信。

5. 如請求項1之方法，其進一步包括：

執行與該主機之該等初級代理程式通信中之至少某些通信及經由共用同一實體鏈路與該等次級代理程式中之至少一者之至少某些初級代理程式通信；且

其中與該等次級代理程式中之至少一者之該等初級代理程式通信包括在該初級代理程式與該等次級代理程式之間交換的該等子請求中之至少某些子請求及該等子狀

態中之至少某些子狀態。

6. 如請求項1之方法，其進一步包括：

根據一同級間協定執行與該等次級代理程式之至少某些初級代理程式通信；且

其中與該等次級代理程式之該等初級代理程式通信包括在該初級代理程式與該等次級代理程式之間交換的至少該等子請求及該等子狀態。

7. 如請求項6之方法，其中：

與該等次級代理程式之該等初級代理程式通信進一步包括經由該同級間協定交換冗餘資訊。

8. 一種提供可擴展儲存之裝置，其包括：

用於將一第一儲存器件操作為一初級代理程式及將至少一第二儲存器件操作為一或多個次級代理程式之構件，該等初級代理程式及次級代理程式中之每一儲存器件具有至少一個各別儲存範圍；

用於使該初級代理程式提供用於與一主機之通信之一邏輯介面且使得該主機能夠作為單一邏輯器件存取散佈於該等初級代理程式及次級代理程式之該等儲存器件上之該等儲存範圍之一彙總之構件，與該主機之該等初級代理程式通信包括由該初級代理程式自該主機接受之儲存空間存取請求及由該初級代理程式發送至該主機之總狀態；

用於使該初級代理程式將接受為一或多個子請求之每一請求轉發至該等次級代理程式之所有或任何部分、自

該等子請求所轉發至之該等次級代理程式接受與該等子請求相關聯之子狀態且至少基於該等子狀態公式化所發送之每一總狀態之構件；及

用於使該等次級代理程式之該等所有或任何部分接受該等子請求、將該等子請求之主機上下文定址資訊轉譯成本端上下文定址資訊、至少部分地基於該本端上下文定址資訊存取本端儲存空間且將該等子狀態提供至該初級代理程式之構件。

9. 如請求項8之裝置，其進一步包括：

用於使該初級代理程式至少部分地基於該等初級代理程式及次級代理程式之該等儲存範圍判定該單一邏輯器件之一彙總儲存容量之構件；

用於使該初級代理程式回應於該主機針對儲存容量查詢該單一邏輯器件而報告該彙總之儲存容量之構件；及

用於使該等次級代理程式中之每一者回應於該主機針對儲存容量查詢該等次級代理程式而報告一各別儲存容量之構件，該各別儲存容量排除在該彙總之儲存容量中表示之該第二次級代理程式之任一儲存範圍之表示。

10. 如請求項8之裝置，其進一步包括：

用於使該特定次級代理程式回應於該主機針對儲存容量查詢該等次級代理程式中之特定一者而報告一特定儲存容量之構件，該特定儲存容量排除在由該初級代理程式所判定之一彙總儲存容量中表示之由該特定次級代理程式所實施之儲存空間之彼等部分之表示。

11. 如請求項8之裝置，其進一步包括：

用於根據一主機介面協定執行與該主機之該等初級代理程式通信中之至少某些通信之構件。

12. 如請求項8之裝置，其進一步包括：

用於執行與該主機之該等初級代理程式通信中之至少某些通信及經由共用同一實體鏈路與該等次級代理程式中之至少一者之至少某些初級代理程式通信之構件；且
其中與該等次級代理程式中之至少一者之該等初級代理程式通信包括在該初級代理程式與該等次級代理程式之間交換的該等子請求中之至少某些子請求及該等子狀態中之至少某些子狀態。

13. 如請求項8之裝置，其進一步包括：

用於根據一同級間協定執行與該等次級代理程式之至少某些初級代理程式通信之構件；且

其中與該等次級代理程式之該等初級代理程式通信包括在該初級代理程式與該等次級代理程式之間交換的至少該等子請求及該等子狀態。

14. 如請求項13之裝置，其中：

與該等次級代理程式之該等初級代理程式通信進一步包括經由該同級間協定交換之冗餘資訊。

15. 一種提供可擴展儲存之系統，其包括：

一儲存子系統，其具有包括複數個實體儲存器件之實體組件及具有複數個埠之一實體開關部分；

一主機，其經由一專用點對點鏈路耦合至該複數個埠

中之一專用埠；

其中每一實體儲存器件包括：

至少一個儲存範圍，

至少一個埠，其經啟用以經由一各別點對點鏈路耦合至該複數個埠中之一各別埠，及

可組態代理程式邏輯，其使得該實體儲存器件能夠操作為該儲存子系統之一或多個初級代理程式及一或多個次級代理程式中之選定一者，每一初級代理程式經啟用以自該主機接受儲存空間存取請求及產生子請求且累積子狀態，每一次級代理程式經啟用以接受該等子請求中之至少一者且產生該等子狀態中之至少一者；且

其中該儲存子系統經啟用以操作一或多個邏輯儲存器件，每一邏輯儲存器件使得該主機能夠經由單個邏輯介面存取散佈於對應於該等初級代理程式中之一者及該等次級代理程式中之至少一者之該等儲存器件上之該等儲存範圍之一彙總。

16. 如請求項15之系統，其進一步包括：

其中具有複數個埠之該實體開關部分係具有第一複數個埠之一第一實體開關部分且該等實體組件進一步包括具有第二複數個埠之一第二實體開關部分；且

其中每一實體儲存器件進一步包括經啟用以經由一各別點對點鏈路耦合至該第二複數個埠中之一各別埠之至少一個埠。

17. 如請求項16之系統，其進一步包括：

其中該第二實體開關部分在對經由該第一實體開關部分之通信可用之帶寬具有不重要影響之情況下實現經由同級間通信而實施之儲存冗餘技術。

18. 如請求項17之系統：

其中該同級間通信係經由該第二實體開關部分且包括控制資訊、未轉換之冗餘資料及已轉換之冗餘資料中之一或多者。

19. 如請求項15之系統，其進一步包括：

其中該單個邏輯介面實現欲經由該等初級代理程式與該等次級代理程式之間的主機閉塞式同級間通信而實施之鏡像、等量化、RAID同位及失效切換中之一或多者。

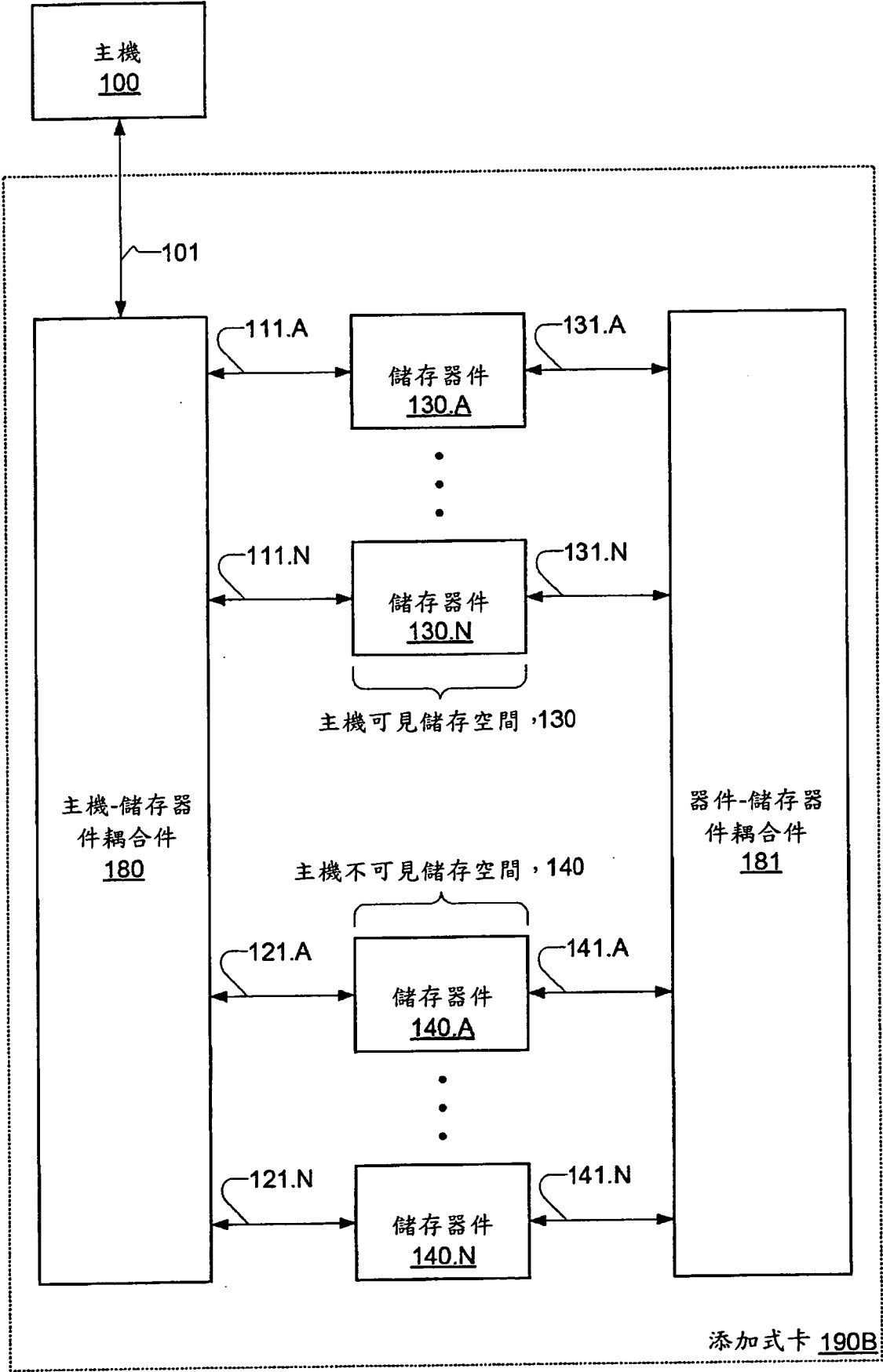


圖 1B

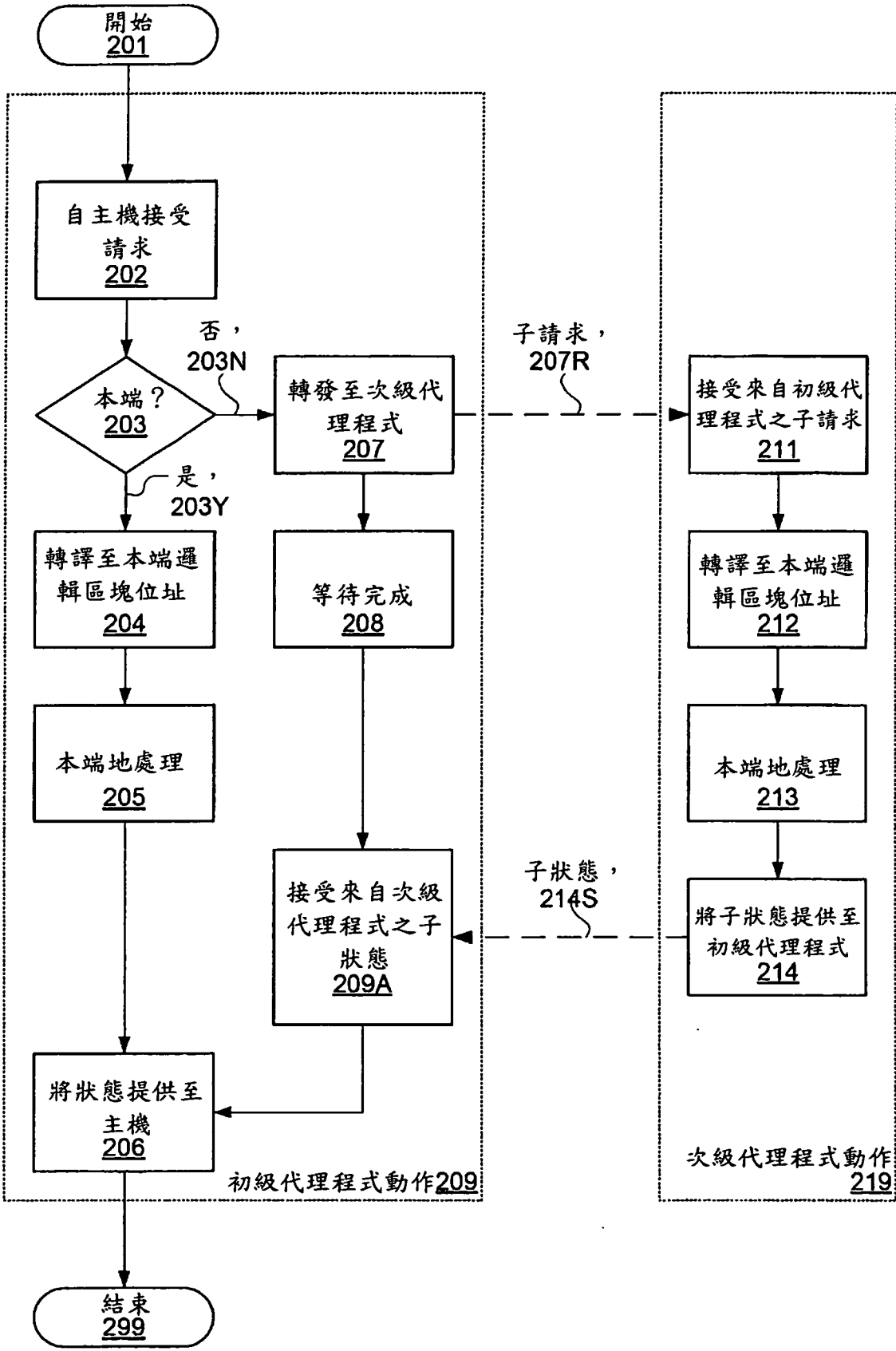


圖 2

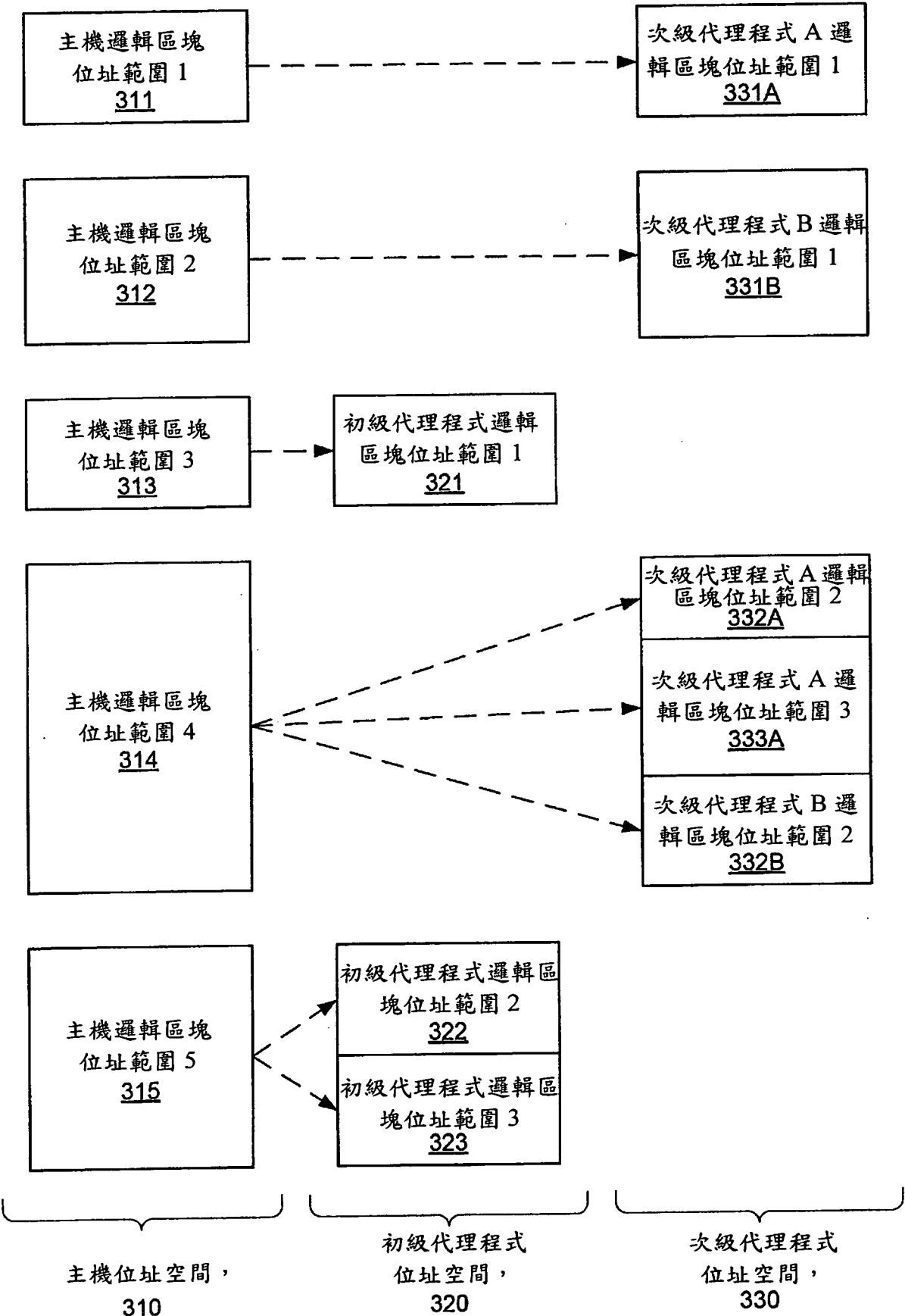


圖 3

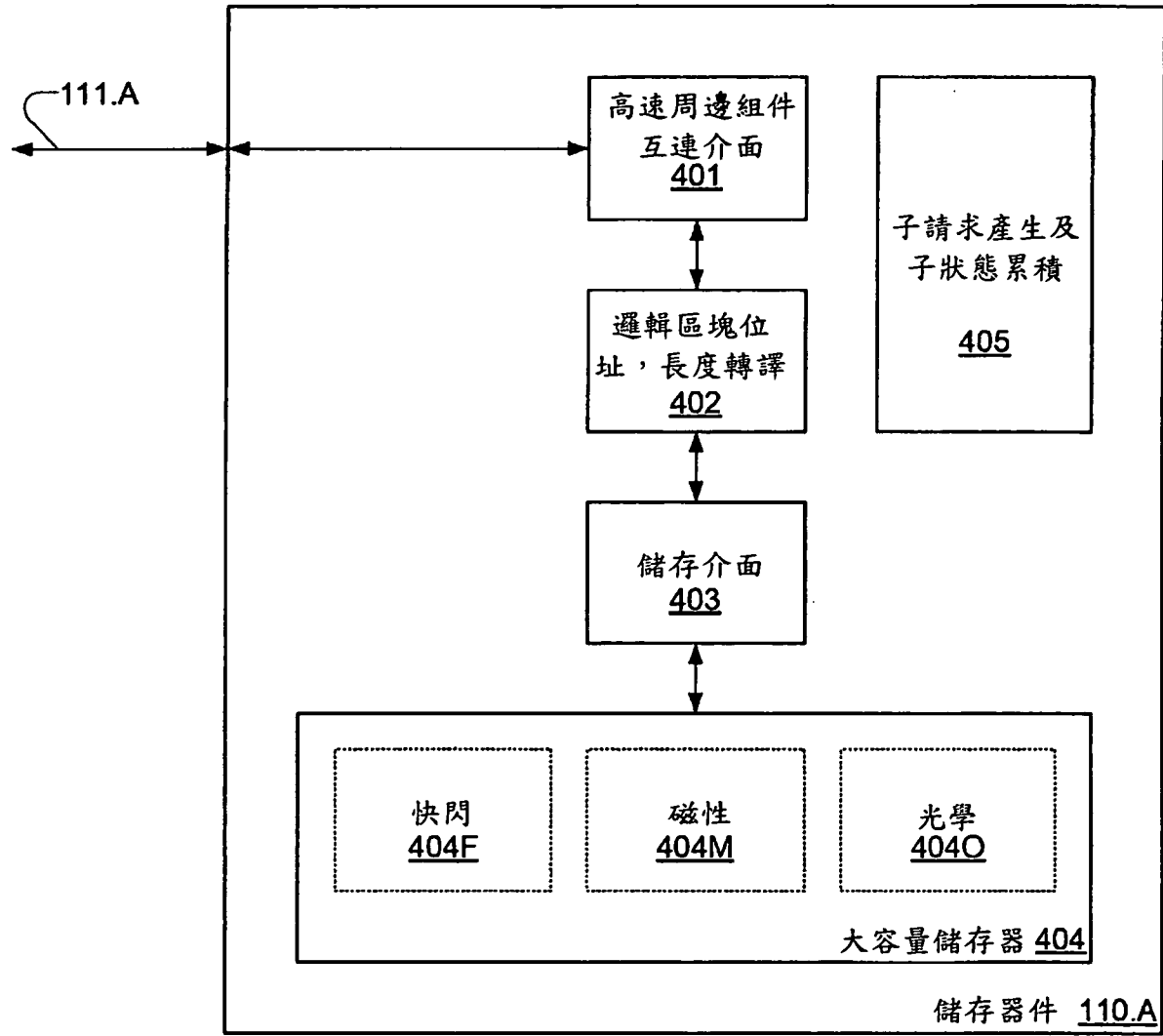


圖 4

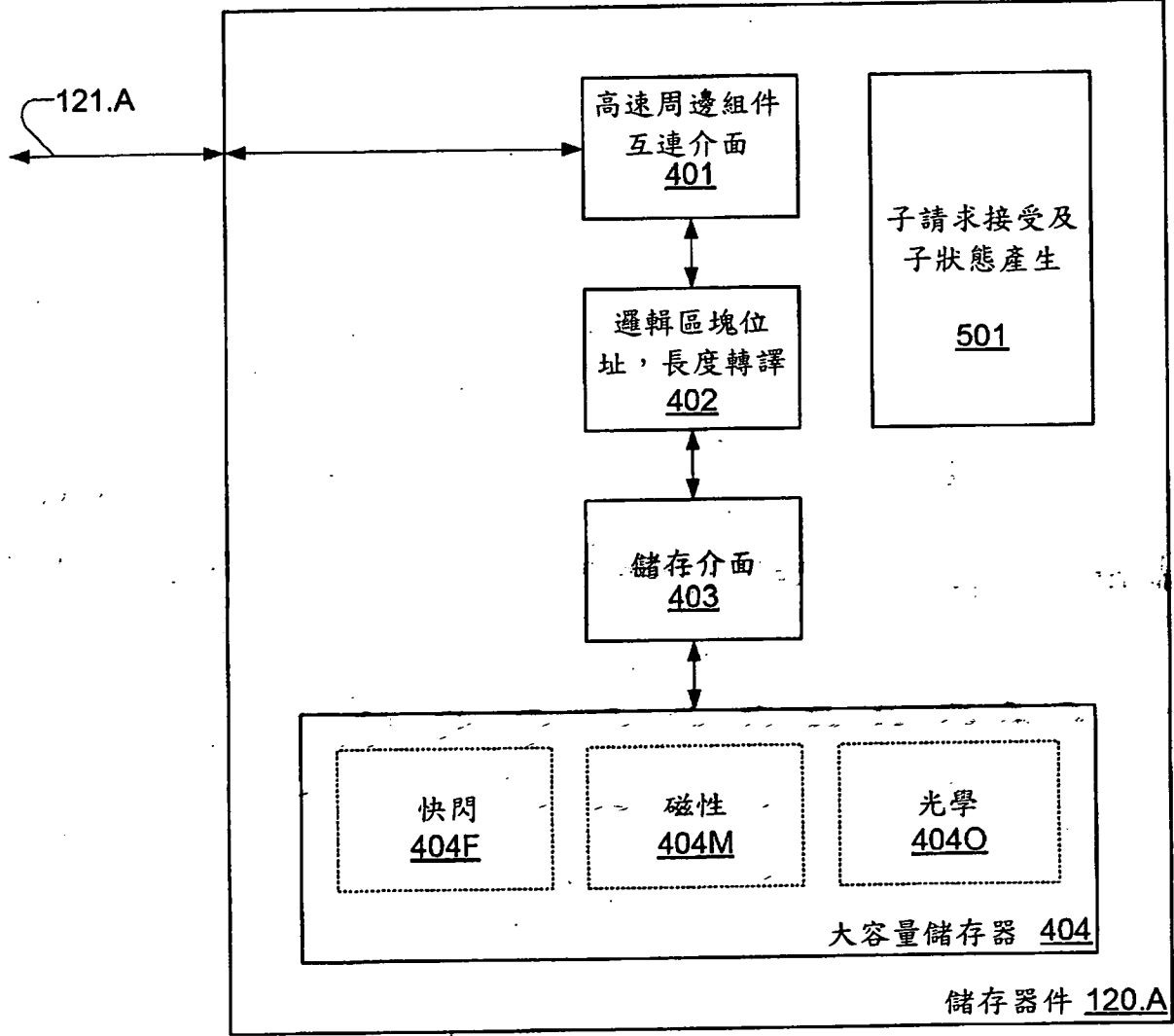


圖 5