

101年8月28日修正替換頁

案號：95104022
101年08月28日修正-替換頁

發明專利說明書

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：95104022

※ 申請日期：95年2月7日

※IPC 分類：G06F15/177, G06F15/82
(2006.01)

一、發明名稱：(中文/英文)

元資料重製與復原之方法，系統與電腦可讀取媒體
METHOD, SYSTEM AND COMPUTER READABLE STORAGE
MEDIUM FOR METADATA REPLICATION AND
RESTORATION

二、申請人：(共1人)

姓名或名稱：(中文/英文)(簽章)

萬國商業機器公司

INTERNATIONAL BUSINESS MACHINES CORPORATION

指定 為應受送達人

代表人：(中文/英文)(簽章)

琳奈 D 安德森/ANDERSON, LYNNE D.

住居所或營業所地址：(中文/英文)

美國紐約州 10504 亞芒克市新奧爾察德路
New Orchard Road, Armonk, NY 10504, U.S.A.

國籍：(中文/英文) 美國/US

三、發明人：(共4人)

姓名：(中文/英文)

1. 肯尼士 偉恩 包德 / BOYD, KENNETH WAYNE
2. 肯尼士 費爾克羅夫 戴 三世 / DAY, KENNETH FAIRCLOUGH III
3. 馬克 艾德華 迪恩 / DEAN, MARK EDWARD
4. 約翰 傑 沃夫根 / WOLFGANG, JOHN JAY

國籍：(中文/英文)

1-4 皆為美國 / US

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：

美國 US；西元 2005 年 02 月 09 日；11/054,976

無主張專利法第二十七條第一項國際優先權：

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。

五、中文發明摘要：

本案提供一方法、系統與製成品，其中資料與元資料 (metadata) 係產生於一計算 (computational) 裝置中，其中所產生資料係儲存於耦合至計算裝置之一第一實體儲存裝置中，且其中所產生的元資料係短暫地儲存於計算裝置中。資料與元資料被重製 (replicated) 至一第二實體儲存裝置。第二實體儲存裝置中的所重製的資料與所重製的元資料被用以自計算裝置與第一實體儲存裝置之至少一失效 (failure) 中恢復 (recover)。

六、英文發明摘要：

Provided are a method, system, and article of manufacture, wherein data and metadata are generated in a computational device, wherein the generated data is stored in a first physical storage device coupled to the computational device, and wherein the generated metadata is stored transiently in the computational device. The data and the metadata are replicated to a second physical storage device. The replicated data and the replicated metadata in the second physical storage device are used to recover from a failure of at least one of the computational device and the first physical storage device.

七、指定代表圖：

(一)本案指定代表圖為：圖一。

(二)本代表圖之元件符號簡單說明：

100 計算裝置

102 主要儲存裝置

104 重製儲存裝置

106a 使用者應用程式

106b 使用者應用程式

108 應用程式元資料

110 重製器

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

九、發明說明：

【發明所屬之技術領域】

此揭露內容關於重製與復原 (restoring) 元資料之一方法、系統、與製成品。

【先前技術】

資訊科技系統，包含儲存系統在內，會需要保護免於據點災難 (site disasters) 或停機 (outages)。進一步來說，資訊科技系統會需要資料遷移 (migration)、資料備份、或資料複製 (duplication) 的功能。災難或停機恢復、資料遷移、資料備份、與資料複製可包含自一儲存系統鏡像 (mirroring) 或複製資料至另一儲存系統。舉例而言，於某些資訊科技系統中，資料可自一主要儲存系統重製至一次要儲存系統。若主要儲存系統不可用，則次要儲存系統中的所重製資料可用來替代主要儲存系統中的不可用資料。

恢復時間目標 (recovery time objective) 是資料可用性應復原的時限。舉例而言，若一股票交易的一交易履行程式應在主要儲存系統變成不可用之三十秒內工作並執行，則此交易履行程式的恢復時間目標為三十秒。

恢復點目標 (recovery point objective) 是資料須復原至的時間點，以重新開始處理。舉例而言，如果一股票交易的一交易履行程式失效，則若所有資料沒有同步備份

至次要儲存，則某些資料會遺失。恢復點目標是交易履行程式與任何遺失的資料應復原至的時間點。恢復時間目標係一段時間，而恢復點目標係一時間點。

【發明內容】

所提供為一方法、系統與製成品，其中資料與元資料 (metadata) 係產生於一計算 (computational) 裝置中，其中所產生的資料係儲存於耦合至計算裝置之一第一實體儲存裝置中，且其中所產生的元資料係短暫地儲存於計算裝置中。資料與元資料被重製 (replicated) 至一第二實體儲存裝置。第二實體儲存裝置中的所重製的資料與所重製元資料被用以自計算裝置與第一實體儲存裝置至少其中之一的一失效 (failure) 中恢復 (recover)。

某些實施例中，計算裝置係一第一計算裝置，其中資料與元資料係由第一計算裝置中執行的一第一應用程式所產生，且其中自失效中恢復更包含復原 (restoring) 所重製的元資料至一第二計算裝置，並容許一第二應用程式使用第二實體儲存裝置中的所重製的資料，此第二應用程式為第一應用程式之複本 (copy)，其中第二應用程式可執行於第二計算裝置中。自第一應用程式之一失效點後執行第二應用程式，係藉由使用所復原的元資料與所重製的資料。

某些其他實施例中，資料與元資料重製至第二實體儲

存裝置更包含自第一實體儲存裝置複製資料至第二實體儲存裝置，以及自計算裝置之記憶體複製元資料至第二實體儲存裝置。

其他實施例中，元資料包含自計算裝置至另一計算裝置之通訊用信息 (messages for communication)。又其他實施例中，元資料包含產生資料與元資料之一執行中應用程式的狀態資訊。

進一步實施例中，元資料並非儲存於第一實體儲存裝置中。

又進一步實施例中，相較於重製資料但不重製元資料的一第二系統中自失效恢復的一第二時段，於重製資料與元資料的一第一系統中自失效恢復的一第一時段是歷時較短。

再進一步實施例中，元資料係一第一元資料，選自計算裝置產生之一第二元資料，其中選自第二元資料之第一元資料之一數量，係基於一恢復時間之目標。

其他實施例中，資料與元資料係以一致的方式重製至第二實體儲存裝置。

【實施方式】

以下說明中，請參照本文及例示實施例所附的圖式，整體觀之。應瞭解的是仍有其他的實施例，可做其他結構性或運作上的改變。

在某些應用程式中，像是資料庫應用程式，可讓元資料短暫地儲存在快取或內部記憶體中。雖然應用程式的資料大部分會位在一實體儲存裝置中，並不是所有資料一直都保存在實體儲存中。即使實體儲存裝置為資料保護被重製，在一系統失效之情形中，短暫地儲存在快取或內部記憶體中的元資料可能遺失。結果，可能無法一致地恢復一失效應用程式之版本，且可能花大量時間來完成，其中此大量時間可能大於一恢復時間之目標。

某些情況中，恢復時間之目標會是重要的，且失效的應用程式會必須在恢復時間目標所定義的期間內開始運作。某些實施例將短暫地儲存在快取與內部記憶體中的元資料進行重製，以相較沒有重製暫態儲存之元資料的情況，減少恢復時間。舉例而言，因為恢復資料庫之一致版本所需的所有資料與元資料已重製，使得在一資料庫之失效發生的某些實施例中，恢復資料庫可能快很多。

圖一例示根據某些實施例之一計算環境的一方塊圖。計算裝置 100 耦合至一主要儲存裝置 102 以及一重製儲存裝置 104。計算裝置 100 可為任何適合的裝置，包含

目前技藝已知者，像是一個人電腦、一工作站 (workstation)、一伺服器、一主機 (mainframe)、一手持式電腦、一掌上型個人電腦、一電話裝置、一網路設備、一刀鋒式電腦 (blade computer)、一儲存伺服器等等。主要儲存裝置 102 與重製儲存裝置 104 可包含任何合適的資料儲存庫 (repository) 系統，包含目前技藝已知者，像是一直接存取儲存裝置 (Direct Access Storage Device, DASD)、一或更多磁碟冗餘陣列階層 (RAID rank)、集束磁碟 (Just a bunch of disks, JBOD) 等等。某些實施例中，主要儲存裝置 102 與重製儲存裝置 104 可為儲存控制器。

雖然圖一顯示計算裝置 100 直接耦合至主要儲存裝置 102 與重製儲存裝置 104，其他實施例中，計算裝置 100、主要儲存裝置 102、與重製儲存裝置 104 可透過一合適網路耦合，包含目前技藝已知者，像是一儲存區域網路 (Storage area network, SAN)、一區域網路、網際網路、一內部網路 (intranet)、一廣域網路 (Wide area network, WAN)、一無線網路等等。

計算裝置 100 可包含一執行中的使用者應用程式 106a、應用程式元資料 108 和一重製器 (replicator) 110。使用者應用程式 106a 是任何合適的應用程式，包含目前技藝已知者，像是一資料庫應用程式、一試算表 (spreadsheet) 應用程式、一文字處理應用程式等等。某些實施例中，使用者應用程式 106a 於計算裝置 100 中

執行。應用程式元資料 108 係與執行中應用程式 106a 相關的元資料。元資料係關於資料的資料。舉例而言，當使用者應用程式 106a 執行於計算裝置 100 中時，與使用者應用程式 106a 相關的應用程式資料 114 可儲存於主要儲存裝置 102。應用程式元資料 108 可為應用程式資料 114 之元資料，其中當使用者應用程式 106a 執行於計算裝置 100 中時，應用程式元資料 108 短暫地儲存於計算裝置 100 中。舉例而言，使用者應用程式 106a 可為一資料庫應用程式，應用程式資料 114 可為對應此資料庫應用程式之資料庫，而應用程式元資料 108 可為對應執行中資料庫應用程式與資料庫之元資料。

重製器 (replicator) 110 可為任何合適的應用程式，實施於硬體、軟體或韌體中，其中重製器 110 複製應用程式資料 114 至重製儲存裝置 104。某些實施例中，所複製的應用程式資料儲存於資料元件「應用程式資料之複本 (the copy of application data)」116 中。此外，重製器 110 亦複製應用程式元資料 108 之所選部分至重製儲存裝置 104，供儲存於資料元件「所選應用程式元資料」118 中。

某些實施例中，計算裝置 100 係一第一計算裝置，一第二計算裝置 120 可耦合至重製儲存裝置 104。某些實施例中，一復原器 (restorer) 112 可提供於第二計算裝置中，其中復原器 112 可為任何合適的應用程式，實施於硬體、軟體、或韌體中。如果計算裝置 100 發生失效的情形時，

包含第二計算裝置 120 與重製儲存裝置 104 之系統可藉由使用應用程式資料之複本 116 與應用程式元資料之複本 118，而自失效恢復，以從使用者應用程式 106a 失效的時點，重新啟動為使用者應用程式 106a 之一複本之一使用者應用程式 106b。

復原器 112 可關聯於一較高階應用程式（未圖示），藉使用於重製儲存裝置 104 中的所選的應用程式元資料 118，於一恢復點目標（recovery point objective）內啟動使用者應用程式 106b。既然使用者應用程式 106b 係使用者應用程式 106a 之複本，重新啟動使用者應用程式 106b 可實際上繼續失效使用者應用程式 106a 之執行。因此，於第一計算裝置 100 之一失效的情形中，復原器 112 復原所選應用程式元資料 118 至第二計算裝置 120。某些實施例中，與復原器 112 相關聯的一較高階應用程式（未圖示），可藉由啟動為使用者應用程式 106a 之複本的使用者應用程式 106b，以及所選應用程式元資料 118，而繼續使用者應用程式 106a 之執行。

因此，圖一例示一實施例，其中一第一計算裝置 100 中的重製器 110 重製對應一執行中使用者應用程式 106a 的應用程式元資料 108 與應用程式資料 114，於一重製儲存裝置 104 中。於第一計算裝置 100 之一失效的情形中，相關聯於一較高階應用程式之第二計算裝置 120 中的重製器 112，可使用儲存於所選應用程式元資料 118 中的應用程式元資料，以及儲存於應用程式資料之複本 116 中所

重製的應用程式資料，來在使用者應用程式 106a 之失效的一恢復點目標內，重新啟動使用者應用程式 106b，其中使用者應用程式 106b 係使用者應用程式 106a 之一複本。

圖二例示一方塊圖，顯示根據某些實施例，實施於計算裝置 100 中的範例應用程式元資料 200。某些實施例中，範例應用程式元資料 200 可對應圖一之應用程式元資料 108。

範例應用程式元資料 200 可包含快取中的元資料 202、主記憶體中的元資料 204、信息 206、執行使用者應用程式 106a 的狀態資訊 208、執行使用者應用程式 106a 的內部記憶體狀態 210、中間計算資料(intermediate calculated data) 212、以及沒有儲存在主要儲存裝置 102 中的其他資訊 214。圖二中包含範例應用程式元資料 200 的元件係為範例，而出現在一範例組件中的資訊可出現在其他範例組件。舉例而言，主記憶體中元資料 204 之部分亦可出現在快取中的元資料 202。範例應用程式元資料 200 的組件亦可與圖二所示不同範疇。

範例應用程式元資料 200 之部分可儲存於計算裝置 100 之一快取或主記憶體中，其中儲存於快取或主記憶體的部分可為暫態，即暫時的，且沒有存在主要儲存裝置 102 中。舉例而言，某些實施例中，於計算裝置之快取中的元

資料 202 及計算裝置 100 之主記憶體中的元資料 204，可以不儲存在主要儲存裝置 102 中。元資料 202、204 可於使用者應用程式 106a 之執行期間產生。

於使用者應用程式 106a 之執行期間，某些訊息 206 可被產生。舉例而言，指示執行中使用者應用程式 106a 之狀態之一信息可產生於計算裝置 100 中，供傳送至另一計算裝置。如此產生的信息 206 可短暫地儲存於計算裝置 100 中，且可不儲存於主要儲存裝置 102 中。

某些實施例中，執行中使用者應用程式 106a 之狀態資訊 208 與執行中使用者應用程式 106a 之內部記憶體狀態 210，可短暫地儲存於計算裝置 100。某些其他實施例中，可由執行中使用者應用程式 106a 執行一系列計算，而一系列計算期間產生之中間值 (intermediate values) 可短暫地儲存於中間計算資料 212 中，而非主要儲存裝置 102 中。某些實施例中，只有一系列計算之最後結果可儲存於主要儲存裝置 102 中。某些實施例中，未儲存於主要儲存裝置 102 中的其他資訊 214，即未儲存於應用程式資料 114 的資訊，亦可包含範例應用程式元資料 200。

圖二例示一實施例，其中範例應用程式元資料 200 可以複數個形式短暫地，即暫時地儲存於計算裝置 100 中，其中使用者應用程式 106a 執行期間，範例應用程式元資料 200 未儲存於主要儲存裝置 102 中。

圖三例示一方塊圖，根據某些實施例供重製應用程式元資料 108。

執行中使用者應用程式 106a 產生（參照編號 300）應用程式資料 114，且亦會產生（參照編號 302）應用程式元資料 108。應用程式資料 114 可被寫入（參照編號 304）至主要儲存裝置 102。

某些實施例中，重製器 110 讀取（參照編號 306）應用程式資料 114，並控制（參照編號 308）應用程式資料 114 之寫入至重製儲存裝置 104，以產生應用程式資料 116 之複製。重製器 110 可於應用程式資料 114 由使用者應用程式 106a 產生時，讀取應用程式資料 114，或自主要儲存系統 102 讀取應用程式資料 114。某些替代實施例中，自主主要儲存裝置 102 複製應用程式資料 114 至重製儲存裝置 104，可由異於重製器 110 之一應用程式執行。

某些實施例中，重製器 110 自計算裝置 100 讀取（參照編號 310）應用程式元資料 108，並寫入（參照編號 312）應用程式元資料 108 之所選部分至所選應用程式元資料 118 中的重製儲存裝置 104。

所選應用程式元資料 118 結合應用程式資料 116 之複

本，可適於在計算裝置 100 或主要儲存裝置 102 之一失效事件中，於恢復時間目標內，供重新啟動執行中使用者應用程式 106a。某些實施例中，復原器 112 可與一較高階應用程式相關，於第一計算裝置 100 之一失效事件中，啟動第二計算裝置 120 中的使用者應用程式 106b（使用者應用程式 106b 係使用者應用程式 106a 之一複本）。某些實施例中，若一第一恢復時間目標係比一第二恢復時間目標較短的一段時間，則會須選擇一較大量的應用程式元資料 108，供儲存於所選應用程式元資料 118 中。某些實施例中，應用程式資料 114 和應用程式元資料 108 以一致的方式重製至重製儲存裝置 104。為避免重製被非同步地執行的情形中，某些實施例會需要一致的重製。

圖三例示一實施例，其中重製器 110 儲存應用程式資料 114 與應用程式元資料 108 之所選部分於重製儲存裝置 104 中。

圖四例示一方塊圖，根據某些實施例，供藉復原應用程式元資料 108，於一恢復時間目標內繼續使用者應用程式 106a 之執行。某些實施例中，於第一計算裝置 100 之一失效事件中，應用程式元資料 108 由復原器 112 復原至第二計算裝置 120。

於無法獲得儲存在主要儲存裝置 102 之應用程式資料 114、或使用者應用程式 106a 的一失效之事件中，復

原器 112，即復原應用程式，建立與重製儲存裝置 104 的通訊（參照編號 400）。復原器 112 自重製儲存裝置 104 讀取（參照編號 402）應用程式資料 116 之複本以及所選元資料 118 的複本。應用程式資料 116 之複本以及所選元資料 118 的複本，可能已由重製器 110 在無法獲得應用程式資料 114 或使用者應用程式 106a 的失效前儲存。

復原器 112 復原（參照編號 404）所選的元資料 118 至一計算裝置，像是第二計算裝置 120。舉例而言，已復原所選的元資料 406，可儲存於第二計算裝置 120 中。可能在第二計算裝置 120 內或外的一較高階應用程式 407，可重新啟動（參照編號 408）使用者應用程式 106b（於圖四標為執行中使用者應用程式 410）。重新啟動的執行中使用者應用程式 410 用已復原所選的元資料 406 及應用程式資料 116 之複本來執行。所選應用程式元資料 118 之復原可容許已重新啟動之執行中使用者應用程式 410（可對應使用者應用程式 106a）工作，並以比沒有復原所選元資料 118 較短之一等待期間運作。

因此，圖四例示某些實施例，其中所選元資料 118 之復原容許已復原之執行中應用程式 410 工作，並於小於一恢復時間目標的一時間內運作。而圖四顯示計算裝置 120 中的復原器 112，某些替代實施例中，復原器 112 亦可在計算裝置 100 中。

圖五例示根據某些實施例，重製與復原元資料之操作 (operations)。某些實施例中，重製實施於第一計算裝置 100 中，而復原實施於第二計算裝置 120 中。

控制始於方塊 500，其中一重製應用程式，像是重製器 110，讀取應用程式資料 114 與應用程式元資料 108。應用程式資料 114 與應用程式元資料 108 可能已由使用者應用程式 106a 之執行所產生。

重製應用程式 110 控制 (於方塊 502) 應用程式資料 114 之寫入至重製儲存裝置 104。平行地，重製應用程式選擇 (於方塊 504) 應用程式元資料 108 之至少某部分，其中在無法獲得儲存於主要儲存裝置 102 之應用程式資料 114、或使用者應用程式 106a 的一失效之情形中，應用程式元資料之所選部分可用以滿足恢復時間目標。

在選擇 (於方塊 504) 應用程式元資料之至少某部分後，重製應用程式 110 將所選的應用程式元資料 118 寫入 (於方塊 506) 至重製儲存裝置 104。控制自方塊 502、506 回到方塊 500，而重製應用程式 110 讀取 (於方塊 500) 進一步產生的應用程式資料 114 與應用程式元資料 108。因此，當執行中使用者應用程式 106a 致使已更新應用程式資料 114 與已更新應用程式元資料 108 之產生時，重製應用程式 110 可重複地繼續更新重製儲存裝置 104。

控制亦可從方塊 500 前進至方塊 508，其中主要儲存系統 102 無法讓執行中使用者應用程式 106a 獲得應用程式資料 114，或執行中使用者應用程式 106a 失效。執行中使用者應用程式 106a 無法獲得應用程式資料 114 可有複數個原因，包含磁碟或主要儲存裝置 102 中的其他單元之失效。無法從主要儲存裝置 102 獲得應用程式資料 114 可導致執行中使用者應用程式 106a 在執行上沒有進一步進行。使用者應用程式 106a 亦可因第一計算裝置 100 之一失效而失效。

一復原應用程式，像是復原器 112，建立(於方塊 510)與重製儲存裝置 104 之通訊。復原應用程式 112 讀取(於方塊 512)所選應用程式元資料 118 之複本。

復原應用程式 112 結合較高階應用程式 407，復原(於方塊 514)執行中使用者應用程式 106a (經由使用者應用程式 106a 之複本，即經由使用者應用程式 106b)，用已復原所選元資料 408 和應用程式資料 116 之複本來繼續執行。所選元資料 408 之復原，使得使用者應用程式 106a 繼續執行具有較短的等待期間，相較於對應使用者應用程式 106a 之應用程式元資料沒有被重製到重製儲存裝置 104 的情況而言。已繼續執行的使用者應用程式 106a (即 106b) 期望應用程式資料在重製儲存裝置 104 中，則復原應用程式 112 不須讀取應用程式資料 116 之複本。

因此，圖六例示某些實施例，其中一應用程式可重製應用程式元資料 108 至一重製儲存裝置 104。在一系統失效之事件中，重製應用程式元資料可用來以比沒有實施應用程式元資料之重製的一情況相較之下較短的一時間，自一失效中恢復。

某些實施例中，復原器 112 和重製器 110 可在同一或不同計算裝置上。舉例而言，某些實施例中，重製器 110 可在第一計算裝置 100 中，而復原器 112 在第二計算裝置 120 中，如圖一所示。這樣的情形中，復原器 112 會復原重製元資料 118 至第二計算裝置 120，以供自一失效中恢復。

某些實施例中，重製器 110 和復原器 112 可藉創造一個保存重製之所有相關元資料的一容器類別 (container class) 來實施。可加強此容器類別，以在一資料更新之事件中，自動地重製所有元資料至重製儲存裝置 104。於某些其他實施例中，一使用者可選擇元資料哪個部分要重製，以及可能重製之對應元資料類別。

另一種方式是，某些實施例中，可實作一應用程式設計介面，供讀取與寫入元資料至內部記憶體。應用程式設計介面可重製所有所需元資料至重製儲存裝置 104，無論一資料更新何時實施。某些實作中，可創造計算裝置 100 之內部記憶體中的一記憶體磁碟 (ramdisk)，來儲存所有

元資料。一記憶體磁碟是記憶體的一部分，被分配來用作一分割區 (partition)，即記憶體包含記憶體磁碟的部分可以類似一硬碟之方式儲存資料。存取此記憶體磁碟記憶體可經由一記憶體磁碟裝置驅動器，其亦同步地重製所有至記憶體磁碟之寫入到重製儲存裝置 104。

某些實施例中，要更快自一失效恢復，提供至重製儲存裝置 104 的一高速資料鏈結 (link)，而重製儲存裝置 104 保持接近繼續執行使用者應用程式 106a 的第二計算裝置 120。

某些實施例中，重製儲存裝置 104 結合元資料與磁碟資料更新兩者於單一的一致性群組 (consistency group)。因此，重製資料與重製元資料可包含所有所需的資料，以於一恢復時間目標內致能 (enable) 一系統。資料和元資料以一致的方式重製至重製儲存裝置 104，特別是在非同步重製架構中。

某些實施例在資料重製方案中容許重製非僅實體 (physical) 儲存的實體 (entities)。資料重製工業中可重製實體磁碟卷 (disk volumes) 至一次要儲存點的產品，係可能有一個低的已重製資料恢復點目標，但卻無手段來控制恢復時間目標。某些實施例提供與實體儲存一起重製非實體儲存之能力。舉例而言，一伺服器之內部記憶體或快取的重製，或發生在二伺服器間的一資料線上的資料傳

輸被重製。某些實施例亦維持實體卷 (physical volume) 與非實體卷間的資料一致性。

某些實施例提供一重製方案，可提供已重製實體儲存與已重製內部快取記憶體或狀態兩者，並期望有了額外已重製元資料，可以較少時間自重製儲存裝置恢復。資料重製方案之恢復時間目標可因此於某些實施例來達成。

額外實施例細節

所說明技術可實作為一方法、裝置或製成品，包括軟體、韌體、微碼 (micro-code)、硬體及/或任何其結合。詞彙「製成品」於文中係用以指程式指令、程式碼、及/或實作為電路之邏輯 (例如一積體電路晶片、可程式化閘陣列[Programmable Gate Array, PGA]、應用特定積體電路[ASIC]等等)、及/或一電腦可讀取媒體 (例如磁儲存媒體，像是一硬碟、軟碟、磁帶)、光儲存 (例如 CD-ROM、DVD-ROM、光碟等等)、揮發性與非揮發性記憶體裝置 (例如電子可抹除可程式化唯讀記憶體[Electrically Erasable Programmable Read Only Memory, EEPROM]、唯讀記憶體、可程式化唯讀記憶體、隨機存取記憶體、動態隨機存取記憶體、靜態隨機存取記憶體、快閃、韌體、可程式化邏輯等等)。電腦可讀取媒體中的程式碼可由一機器存取與執行，像是一處理器。某些實施例中，實施例實作於其中之程式碼可進一步透過一傳輸媒體或經一網路自一檔案伺服器存取。這樣的情形中，程式碼實作於其中之製成品可包含一傳輸媒體，像是一網路傳輸線、無線傳輸媒

體、經過太空傳播的信號、無線電波、紅外線信號等等。當然，熟此技藝者會瞭解到，仍有許多不悖離實施例之範疇的修正，且製成品可包含任何技藝中已知資訊承載媒體。舉例而言，製成品可包含儲存有指令之一儲存媒體，而當該指令由一機器執行時，會實行前述之操作。

圖六例示一系統 600 之一方塊圖，某些實施例可實作於其中。某些實施例中，計算裝置 100、120 可根據系統 600 實作。系統 600 可包含一電路 602，其於某些實施例中可包含一處理器 604。系統 600 亦可包含一記憶體 606（例如一揮發性記憶體裝置），以及儲存器 608。系統 600 之某些元件可能會或不會在計算裝置 100、120 中找到。儲存 608 可包含一非揮發性記憶體裝置（例如電子可抹除可程式化唯讀記憶體、唯讀記憶體、可程式化唯讀記憶體、隨機存取記憶體、動態隨機存取記憶體、靜態隨機存取記憶體、快閃、韌體、可程式化邏輯等等）、磁碟機、光碟機、磁帶機等等。儲存 608 可包含一內部儲存裝置、一附接儲存裝置及/或一網路可存取儲存裝置。系統 600 可包含一程式邏輯 610，程式邏輯 610 包含程式碼 612，程式碼 612 可被載入記憶體 606 並由處理器 604 或電路 602 執行。某些實施例中，包含程式碼 612 之程式邏輯 610 可儲存於儲存器 608。

某些實施例會針對一方法，供部署（deploying）一個人之計算指令或自動化處理整合電腦可讀取程式碼於一計算系統中，其中程式碼結合計算系統可執行所述實施例

之操作。

至少某些圖五的操作可順序地亦可平行地執行。另些實施例中，某些操作會以不同次序執行、修改、或移除。

此外，為例示之目的，許多軟體或硬體組件以分離模組說明。這樣的組件可整合為更少數目的組件或分成更多數目的組件。另外，描述為由一特定組件執行之某些操作亦可由其他組件執行。

圖一至六顯示或提及之資料結構與組件被說明為有特定資訊類型。另些實施例中，資料結構和組件可有不同結構，並有較圖中顯示或提及的更少、更多或不同之欄位（fields）或不同功能。因此，實施例之前述說明已為例示與說明之目的提出。其不欲為窮舉性或限制實施例於所揭露確切形式。鑑於以上教示，可有許多修改與變化。

【圖式簡單說明】

現在參照至圖式，其中類似的參照編號代表相對應的部分：

圖一顯示根據一實施例的計算環境之一方塊圖；

圖二顯示根據一實施例的範例應用程式元資料之一方塊圖；

圖三顯示根據一實施例供重製元資料之一方塊圖；

圖四顯示根據一實施例供復原元資料之一方塊圖；

圖五顯示根據一實施例供重製與復原元資料之操作；以及

圖六顯示一實施例實作於其中之系統。

【主要元件符號說明】

100 計算裝置	102 主要儲存裝置
104 重製儲存裝置	106a 使用者應用程式
106b 使用者應用程式	108 應用程式元資料
110 重製器	112 復原器
114 應用程式資料	116 應用程式資料之複本
118 所選應用程式元資料	120 計算裝置
200 範例應用程式元資料	202 元資料
204 元資料	206 信息
208 狀態資訊	210 內部記憶體狀態
212 中間計算資料	214 其他資訊
406 已復原所選元資料	407 更高階應用程式
410 已重新啟動執行中的使用者應用程式	
600 系統	602 電路
604 處理器	606 記憶體
608 儲存器	610 程式邏輯
612 程式碼	

十、申請專利範圍：

1. 一種元資料重製與復原之方法，包含：

創造一容器類別以實作一重製應用程式以及一復原應用程式，其中該重製應用程式執行於一第一計算裝置，而該復原應用程式執行於一第二計算裝置，而該容器類別保存將被重製之所有相關元資料；

於該第一計算（computational）裝置中藉由執行一第一應用程式產生資料與第一元資料（metadata），其中該產生的資料係儲存於耦合至該第一計算裝置之一第一實體儲存裝置中，於該第一計算裝置中暫時儲存該產生的元資料至該第一計算裝置之一記憶體磁碟（ramdisk），該記憶體磁碟是該第一計算裝置之一內部記憶體的一部分，分配該記憶體磁碟為一分割區（partition），以及經由一記憶體磁碟裝置驅動器存取該記憶體磁碟記憶體；

其中為了產生該第一元資料，該執行該第一應用程式之步驟包含：

- (a) 儲存執行該第一應用程式的狀態資訊於該第一元資料；
- (b) 儲存自該第一計算裝置至另一計算裝置之通訊用信息於該第一元資料；
- (c) 儲存執行該第一應用程式的內部記憶體狀態於該第一元資料；
- (d) 儲存暫時儲存藉由執行該第一應用程式所執行一系列計算而產生之中間值之中間計算資料（intermediate calculated data）於該第一元資料，以及暫時儲存該第一元資料之部份於該第一計算裝置，而不儲存該

第一元資料之該部份於該第一實體儲存裝置；

藉由於該第一計算裝置執行該重製應用程式，重製（replicating）該資料與該第一元資料之至少一部分至一第二實體儲存裝置，該步驟係藉由：

- (i) 自該第一實體儲存裝置複製該資料至該第二實體儲存裝置；以及
- (ii) 基於一恢復時間之目標，從該第一元資料中選擇第二元資料，該恢復時間之目標係定義為從在該第一計算裝置上執行該第一應用程式之任何失效中恢復之時間，其中該第二元資料係為該第一元資料之部份且在數量上小於該第一元資料，以及藉由經由該記憶體磁碟同步地重製所有至該記憶體磁碟之寫入至該第二實體儲存裝置，將該第二元資料從該第一計算裝置之該記憶體磁碟複製到該第二實體儲存裝置，其中該第二實體儲存裝置結合該資料以及該第二元資料為一單一的一致性群組，而該資料以及該第二元資料係以一致的方式重製至該第二實體儲存裝置；以及

在由該恢復時間之目標所定義之一段時間中，自執行該第一應用程式之失效（failure）中恢復（recovering），該步驟係藉由：

- (i) 藉由於該第二計算裝置執行該復原應用程式，

復原 (restoring) 該重製的第二元資料至該第二計算裝置，其中該第二計算裝置係不同於該第一計算裝置；

- (ii) 容許一第二應用程式使用該第二實體儲存裝置中該重製的資料，該第二應用程式為執行於該第一計算裝置之該第一應用程式之一複本 (copy)；以及
- (iii) 自該第一應用程式之一失效點後於該第二計算裝置執行該第二應用程式，係藉由使用該復原的第二元資料與該重製的資料，該復原的第二元資料包含執行該第一應用程式的狀態資訊，其中該第二應用程式在由該恢復時間之目標所定義之該段時間開始執行。

2. 如請求項 1 所述之方法，其中該元資料並非儲存於該第一實體儲存裝置中。

3. 如請求項 1 所述之方法，其中相較於重製該資料但不重製該元資料的一第二系統中自該失效恢復的一第二時段間，於重製該資料與該元資料的一第一系統中自該失效恢復的一第一時段是歷時較短。

4. 如請求項 1 所述之方法，其中該資料與該元資料係以一致的方式重製至該第二實體儲存裝置。

5. 如請求項 1 所述之方法，其中：

- (i) 該第一計算裝置包含：(a) 一重製應用程式，以及

- (b) 一使用者程式係為該第一應用程式；以及
- (ii) 該第二計算裝置包含：(a) 一復原應用程式，以及
- (b) 該使用者程式之一複本，其中該使用者程式之該複本係為該第二應用程式。

6. 一種元資料重製與復原之系統，包含：

記憶體；

處理器，耦合至該記憶體，其中該處理器可執行一操作，該操作包含：

創造一容器類別以實作一重製應用程式以及一復原應用程式，其中該重製應用程式執行於一第一計算裝置，而該復原應用程式執行於一第二計算裝置，而該容器類別保存將被重製之所有相關元資料；

於該第一計算 (computational) 裝置中藉由執行一第一應用程式產生資料與第一元資料 (metadata)，其中該產生的資料係儲存於耦合至該第一計算裝置之一第一實體儲存裝置中，於該第一計算裝置中暫時儲存該產生的元資料至該第一計算裝置之一記憶體磁碟 (ramdisk)，該記憶體磁碟是該第一計算裝置之一內部記憶體的一部分，分配該記憶體磁碟為一分割區 (partition)，以及經由一記憶體磁碟裝置驅動器存取該記憶體磁碟記憶體；

其中為了產生該第一元資料，該執行該第一應用程式之步驟包含：

- (a) 儲存執行該第一應用程式的狀態資訊於該第一元資料；
- (b) 儲存自該第一計算裝置至另一計算裝置之通訊用信息於該第一元資料；

- (c) 儲存執行該第一應用程式的內部記憶體狀態於該第一元資料；
- (d) 儲存暫時儲存藉由執行該第一應用程式所執行一系列計算而產生之中間值之中間計算資料 (intermediate calculated data) 於該第一元資料，以及暫時儲存該第一元資料之部份於該第一計算裝置，而不儲存該第一元資料之該部份於該第一實體儲存裝置；

藉由於該第一計算裝置執行該重製應用程式，重製 (replicating) 該資料與該第一元資料之至少一部分至一第二實體儲存裝置，該步驟係藉由：

- (i) 自該第一實體儲存裝置複製該資料至該第二實體儲存裝置；以及
- (ii) 基於一恢復時間之目標，從該第一元資料中選擇第二元資料，該恢復時間之目標係定義為從在該第一計算裝置上執行該第一應用程式之任何失效中恢復之時間，其中該第二元資料係為該第一元資料之部份且在數量上小於該第一元資料，以及藉由經由該記憶體磁碟同步地重製所有至該記憶體磁碟之寫入至該第二實體儲存裝置，將該第二元資料從該第一計算裝置之該記憶體磁碟複製到該第二實體儲存裝置，其中該第二實體儲存裝置結合該資料以及該第二元資料為一單一的一致性群組，而該資料以及該第二元資料係以一致的方式重製至該第二實體儲存裝

置；以及

在由該恢復時間之目標所定義之一段時間中，自執行該第一應用程式之失效 (failure) 中恢復 (recovering)，該步驟係藉由：

- (i) 藉由於該第二計算裝置執行該復原應用程式，復原 (restoring) 該重製的第二元資料至該第二計算裝置，其中該第二計算裝置係不同於該第一計算裝置；
- (ii) 容許一第二應用程式使用該第二實體儲存裝置中該重製的資料，該第二應用程式為執行於該第一計算裝置之該第一應用程式之一複本 (copy)；以及
- (iii) 自該第一應用程式之一失效點後於該第二計算裝置執行該第二應用程式，係藉由使用該復原的第二元資料與該重製的資料，該復原的第二元資料包含執行該第一應用程式的狀態資訊，其中該第二應用程式在由該恢復時間之目標所定義之該段時間開始執行。

7. 如請求項 6 所述之系統，其中該元資料並非儲存於該第一實體儲存裝置中。

8. 如請求項 6 所述之系統，其中相較於重製該資料但不重製該元資料的一第二系統中自該失效恢復的一第二時段間，於重製該資料與該元資料的一第一系統中自該失效恢復的一第一時段是歷時較短。

9. 如請求項 6 所述之系統，其中該資料與該元資料係以一致的方式重製至該第二實體儲存裝置。

10. 如請求項 6 所述之系統，其中：

- (i) 該第一計算裝置包含：(a) 一重製應用程式，以及 (b) 一使用者程式係為該第一應用程式；以及
- (ii) 該第二計算裝置包含：(a) 一復原應用程式，以及 (b) 該使用者程式之一複本，其中該使用者程式之該複本係為該第二應用程式。

11. 一種電腦可讀取媒體，其中儲存該電腦可讀取媒體之碼當由一處理器執行時，可產生一操作，該操作包含：

創造一容器類別以實作一重製應用程式以及一復原應用程式，其中該重製應用程式執行於一第一計算裝置，而該復原應用程式執行於一第二計算裝置，而該容器類別保存將被重製之所有相關元資料；

於該第一計算 (computational) 裝置中藉由執行一第一應用程式產生資料與第一元資料 (metadata)，其中該產生的資料係儲存於耦合至該第一計算裝置之一第一實體儲存裝置中，於該第一計算裝置中暫時儲存該產生的元資料至該第一計算裝置之一記憶體磁碟 (ramdisk)，該記憶體磁碟是該第一計算裝置之一內部記憶體的一部分，分配該記憶體磁碟為一分割區 (partition)，以及經由一記憶體磁碟裝置驅動器存取該記憶體磁碟記憶體；

其中為了產生該第一元資料，該執行該第一應用程式之步驟包含：

- (a) 儲存執行該第一應用程式的狀態資訊

於該第一元資料；

- (b) 儲存自該第一計算裝置至另一計算裝置之通訊用信息於該第一元資料；
- (c) 儲存執行該第一應用程式的內部記憶體狀態於該第一元資料；
- (d) 儲存暫時儲存藉由執行該第一應用程式所執行一系列計算而產生之中間值之中間計算資料（intermediate calculated data）於該第一元資料，以及暫時儲存該第一元資料之部份於該第一計算裝置，而不儲存該第一元資料之該部份於該第一實體儲存裝置；

藉由於該第一計算裝置執行該重製應用程式，重製（replicating）該資料與該第一元資料之至少一部分至一第二實體儲存裝置，該步驟係藉由：

- (i) 自該第一實體儲存裝置複製該資料至該第二實體儲存裝置；以及
- (ii) 基於一恢復時間之目標，從該第一元資料中選擇第二元資料，該恢復時間之目標係定義為從在該第一計算裝置上執行該第一應用程式之任何失效中恢復之時間，其中該第二元資料係為該第一元資料之部份且在數量上小於該第一元資料，以及藉由經由該記憶體磁碟同步地重製所有至該記憶體磁碟之寫入至該第二實體儲存裝置，將該第二元資料從該第一計

算裝置之該記憶體磁碟複製到該第二實體儲存裝置，其中該第二實體儲存裝置結合該資料以及該第二元資料為一單一的一致性群組，而該資料以及該第二元資料係以一致的方式重製至該第二實體儲存裝置；以及

在由該恢復時間之目標所定義之一段時間中，自執行該第一應用程式之失效 (failure) 中恢復 (recovering)，該步驟係藉由：

- (i) 藉由於該第二計算裝置執行該復原應用程式，復原 (restoring) 該重製的第二元資料至該第二計算裝置，其中該第二計算裝置係不同於該第一計算裝置；
- (ii) 容許一第二應用程式使用該第二實體儲存裝置中該重製的資料，該第二應用程式為執行於該第一計算裝置之該第一應用程式之一複本 (copy)；以及
- (iii) 自該第一應用程式之一失效點後於該第二計算裝置執行該第二應用程式，係藉由使用該復原的第二元資料與該重製的資料，該復原的第二元資料包含執行該第一應用程式的狀態資訊，其中該第二應用程式在由該恢復時間之目標所定義之該段時間開始執行。

12. 如請求項 11 所述之電腦可讀取媒體，其中該元資料並非儲存於該第一實體儲存裝置中。

13. 如請求項 11 所述之電腦可讀取媒體，其中相較於重製該資料但不重製該元資料的一第二系統中自該失效恢復的一第二時段間，於重製該資料與該元資料的一第一系統中自該失效恢復的一第一時段是歷時較短。

14. 如請求項 11 所述之電腦可讀取媒體，其中該資料與該元資料係以一致的方式重製至該第二實體儲存裝置。

15. 如請求項 11 所述之電腦可讀取媒體，其中：

- (i) 該第一計算裝置包含：
 - (a) 一重製應用程式，以及
 - (b) 一使用者程式係為該第一應用程式；以及
- (ii) 該第二計算裝置包含：
 - (a) 一復原應用程式，以及
 - (b) 該使用者程式之一複本，其中該使用者程式之該複本係為該第二應用程式。

16. 一種部署 (deploying) 計算基礎建設 (infrastructure) 之方法，包含整合電腦可讀取碼於一計算系統中，其中該碼結合該計算系統可用於執行下列步驟：

創造一容器類別以實作一重製應用程式以及一復原應用程式，其中該重製應用程式執行於一第一計算裝置，而該復原應用程式執行於一第二計算裝置，而該容器類別保存將被重製之所有相關元資料；

於該第一計算 (computational) 裝置中藉由執行一第一應用程式產生資料與第一元資料 (metadata)，其中該產生的資料係儲存於耦合至該第一計算裝置之一第一實體儲存裝置中，於該第一計算裝置中暫時儲存該產生的元資料至該第一計算裝置之一記憶體磁碟 (ramdisk)，該記憶體

磁碟是該第一計算裝置之一內部記憶體的一部分，分配該記憶體磁碟為一分割區（partition），以及經由一記憶體磁碟裝置驅動器存取該記憶體磁碟記憶體；

其中為了產生該第一元資料，該執行該第一應用程式之步驟包含：

- (a) 儲存執行該第一應用程式的狀態資訊於該第一元資料；
- (b) 儲存自該第一計算裝置至另一計算裝置之通訊用信息於該第一元資料；
- (c) 儲存執行該第一應用程式的內部記憶體狀態於該第一元資料；
- (d) 儲存暫時儲存藉由執行該第一應用程式所執行一系列計算而產生之中間值之中間計算資料（intermediate calculated data）於該第一元資料，以及暫時儲存該第一元資料之部份於該第一計算裝置，而不儲存該第一元資料之該部份於該第一實體儲存裝置；

藉由於該第一計算裝置執行該重製應用程式，重製（replicating）該資料與該第一元資料之至少一部分至一第二實體儲存裝置，該步驟係藉由：

- (i) 自該第一實體儲存裝置複製該資料至該第二實體儲存裝置；以及
- (ii) 基於一恢復時間之目標，從該第一元資料中選擇第二元資料，該恢復時間之目標係定義為從在該第一計算裝置上執行該第一應用

程式之任何失效中恢復之時間，其中該第二元資料係為該第一元資料之部份且在數量上小於該第一元資料，以及藉由經由該記憶體磁碟同步地重製所有至該記憶體磁碟之寫入至該第二實體儲存裝置，將該第二元資料從該第一計算裝置之該記憶體磁碟複製到該第二實體儲存裝置，其中該第二實體儲存裝置結合該資料以及該第二元資料為一單一的一致性群組，而該資料以及該第二元資料係以一致的方式重製至該第二實體儲存裝置；以及

在由該恢復時間之目標所定義之一段時間中，自執行該第一應用程式之失效 (failure) 中恢復 (recovering)，該步驟係藉由：

- (i) 藉由於該第二計算裝置執行該復原應用程式，復原 (restoring) 該重製的第二元資料至該第二計算裝置，其中該第二計算裝置係不同於該第一計算裝置；
- (ii) 容許一第二應用程式使用該第二實體儲存裝置中該重製的資料，該第二應用程式為執行於該第一計算裝置之該第一應用程式之一複本 (copy)；以及
- (iii) 自該第一應用程式之一失效點後於該第二計算裝置執行該第二應用程式，係藉由使用該復原

的第二元資料與該重製的資料，該復原的第二元資料包含執行該第一應用程式的狀態資訊，其中該第二應用程式在由該恢復時間之目標所定義之該段時間開始執行。

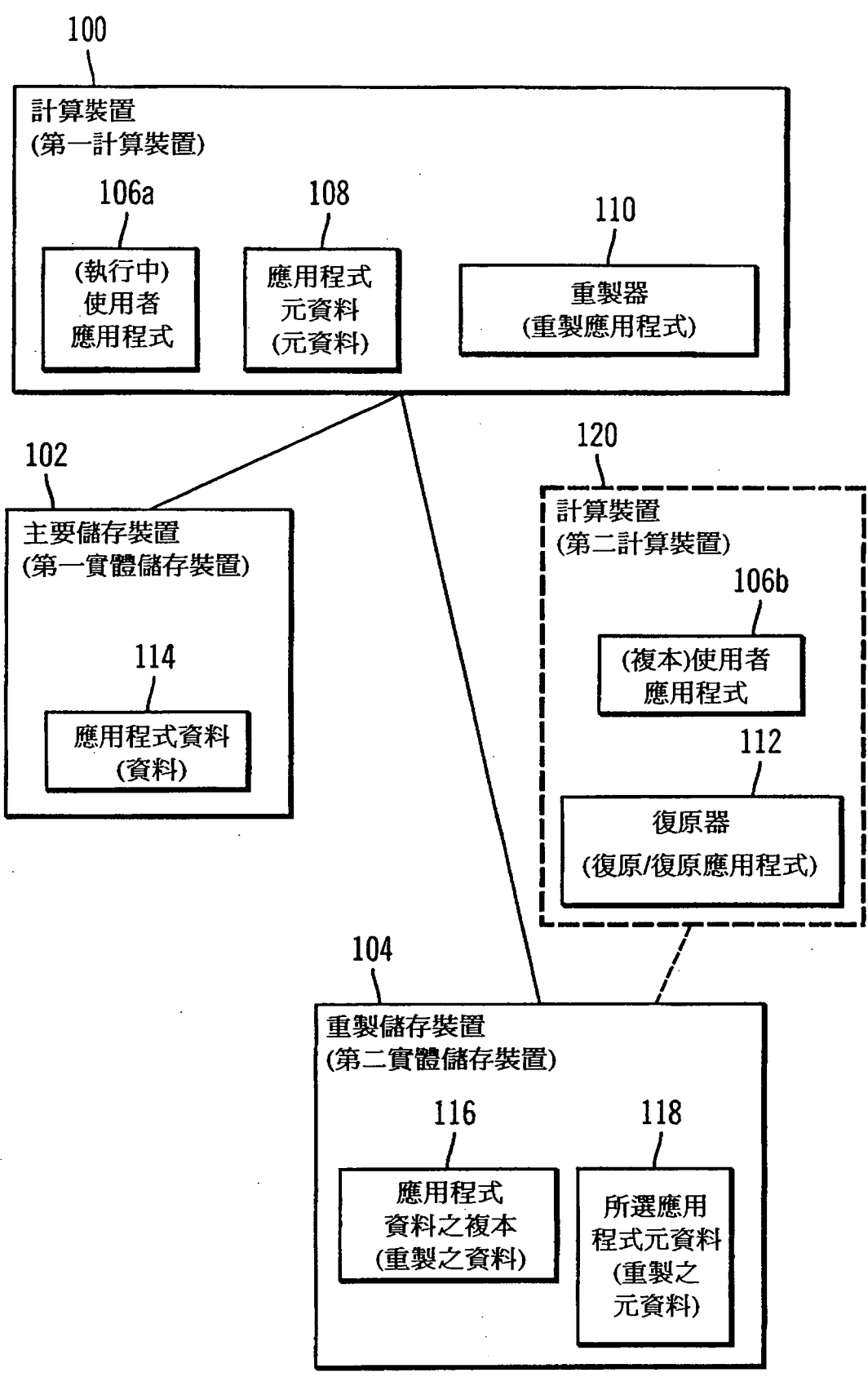
17. 如請求項 16 所述之方法，其中該元資料並非儲存於該第一實體儲存裝置中。

18. 如請求項 16 所述之方法，其中相較於重製該資料但不重製該元資料的一第二系統中自該失效恢復的一第二時段間，於重製該資料與該元資料的一第一系統中自該失效恢復的一第一時段是歷時較短。

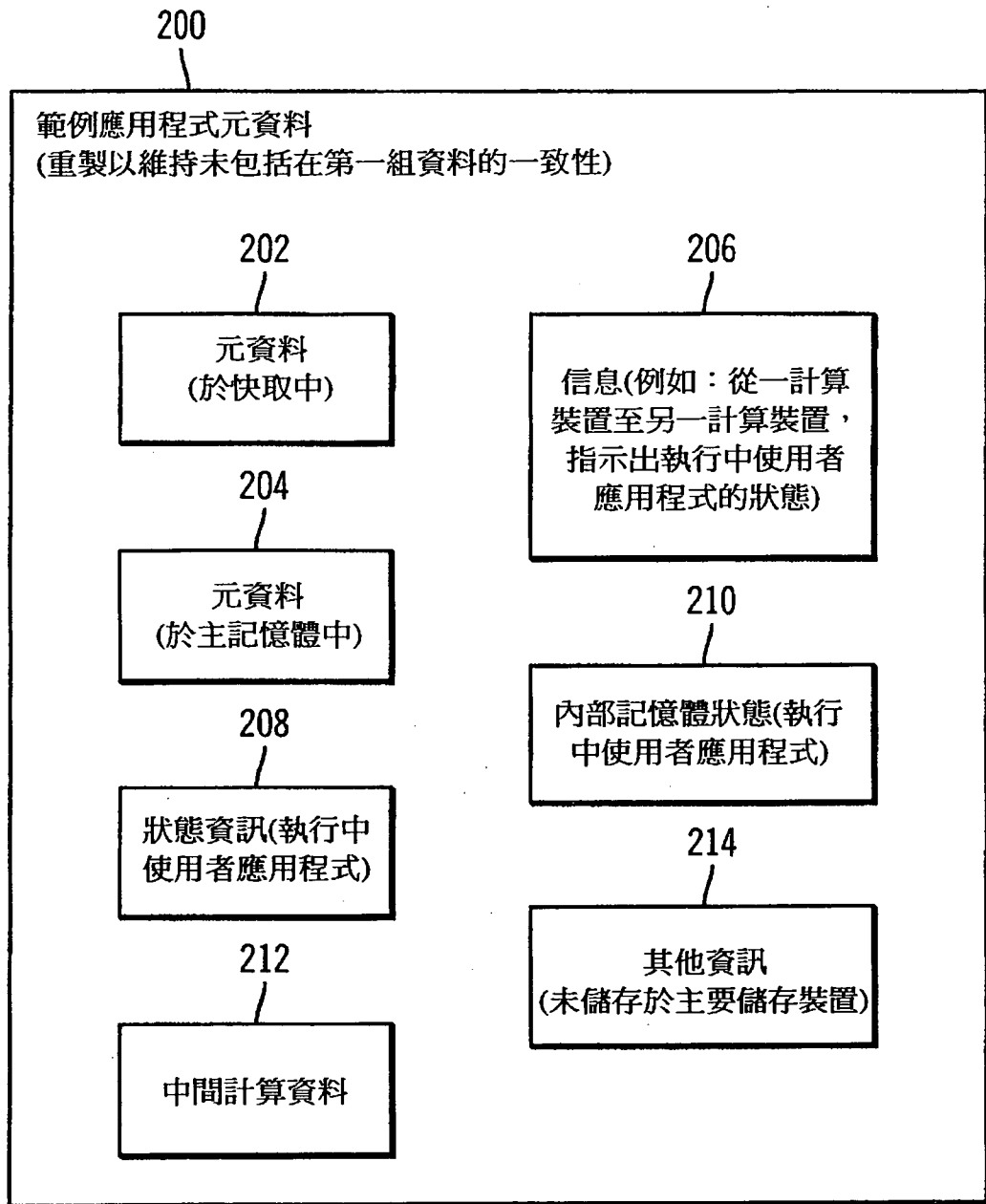
19. 如請求項 16 所述之方法，其中該資料與該元資料係以一致的方式重製至該第二實體儲存裝置。

20. 如請求項 16 所述之方法，其中：

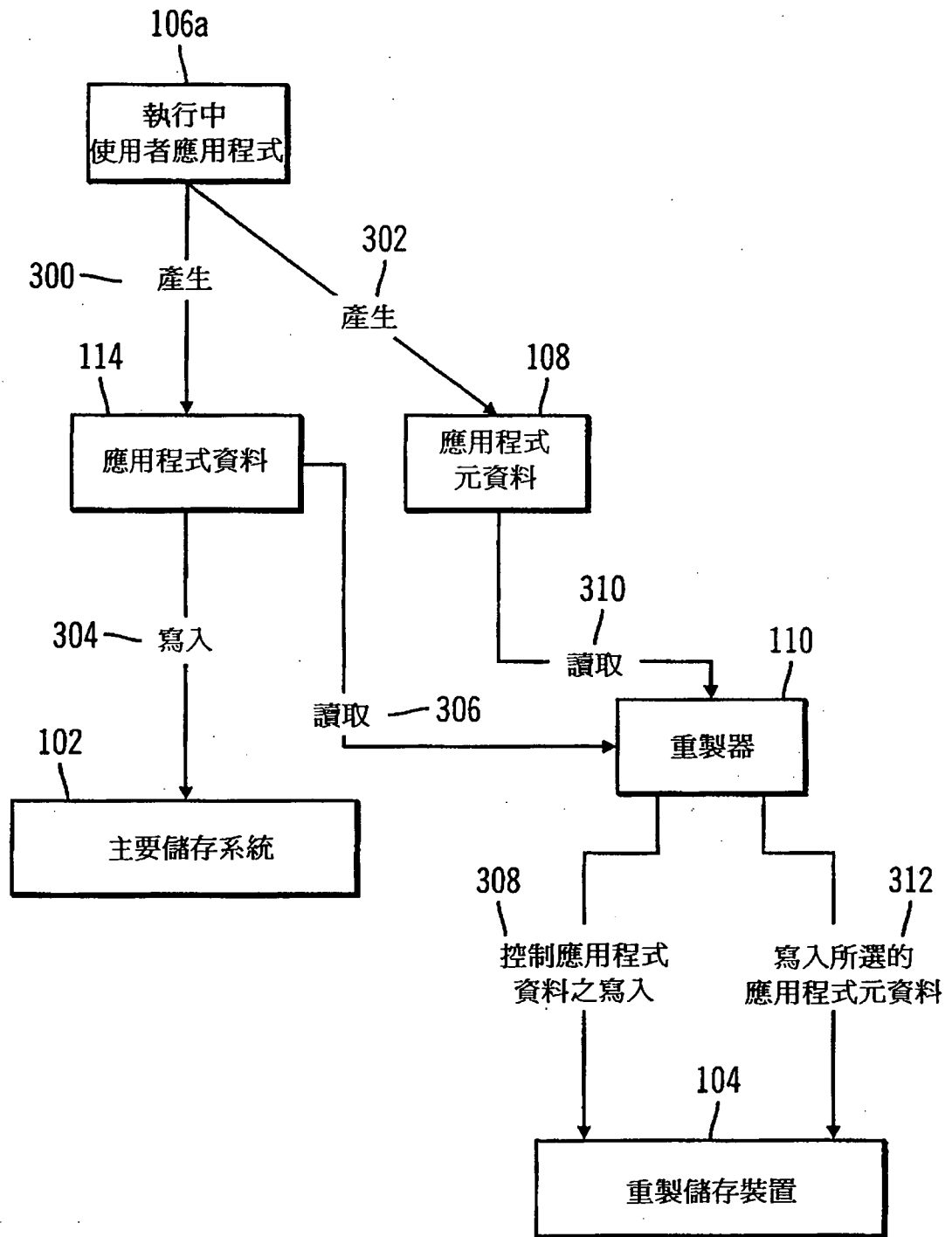
- (i) 該第一計算裝置包含：(a) 一重製應用程式，以及 (b) 一使用者程式係為該第一應用程式；以及
- (ii) 該第二計算裝置包含：(a) 一復原應用程式，以及 (b) 該使用者程式之一複本，其中該使用者程式之該複本係為該第二應用程式。



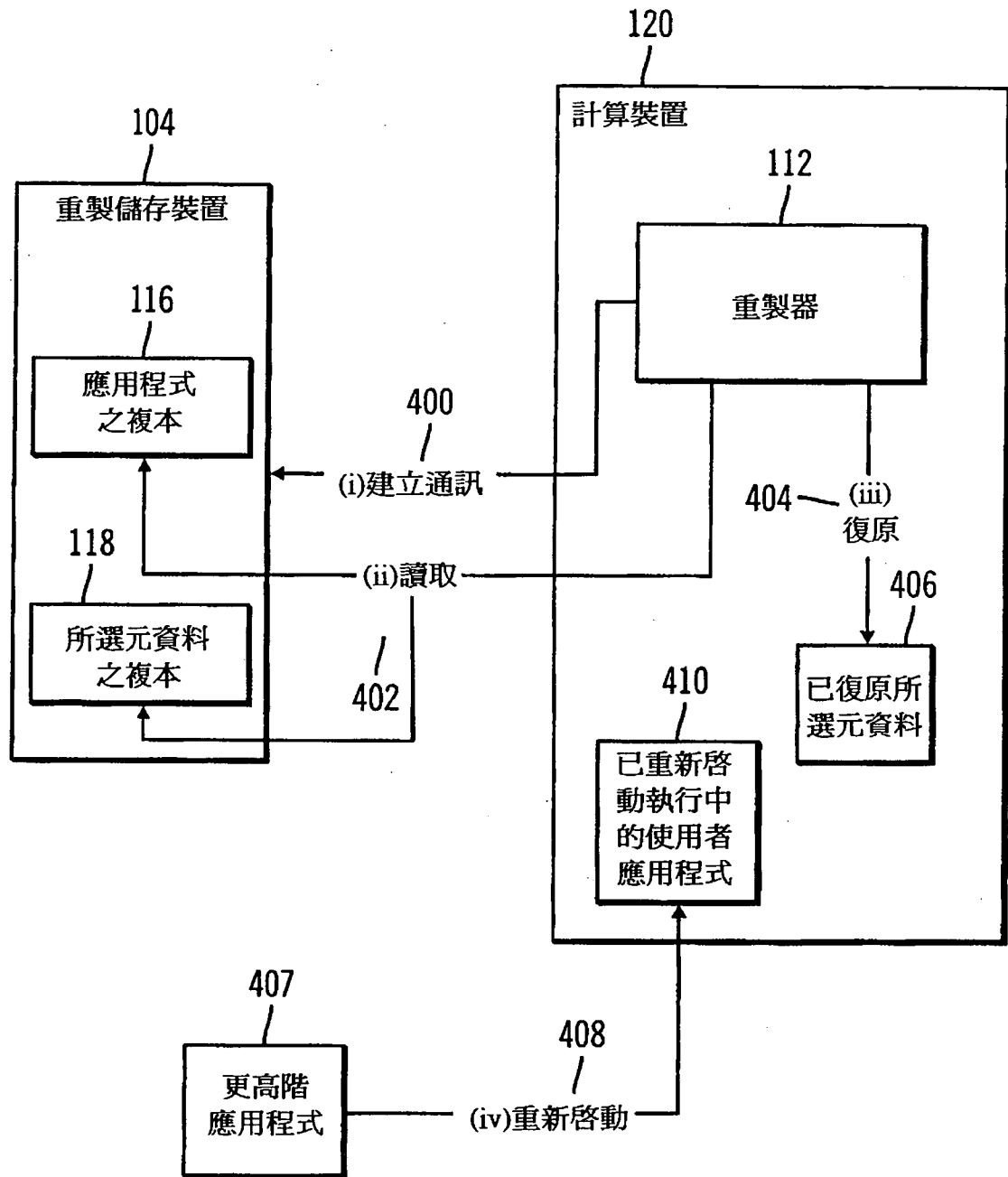
圖一



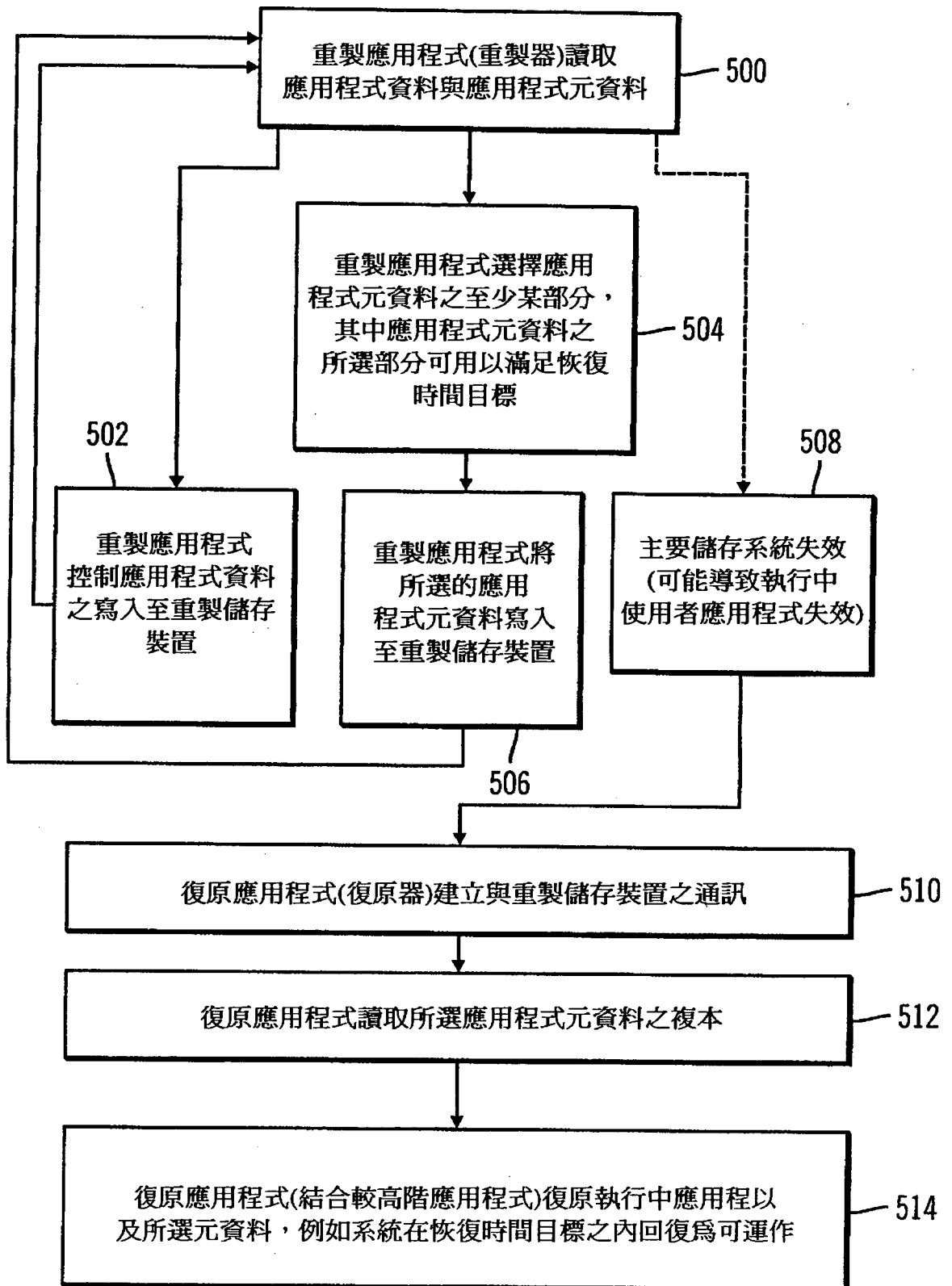
圖二



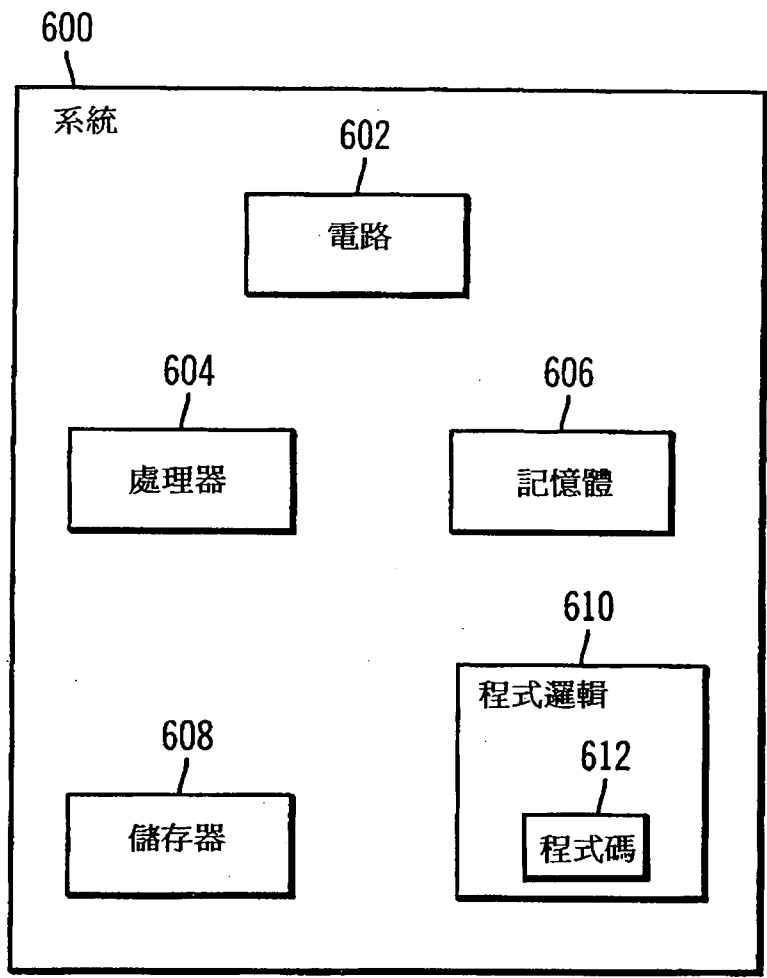
圖三



圖四



圖五



圖六