

380223

申請日期	87. 6. 22
案 號	87110004
類 別	G06F 13/00

A4
C4

(以上各欄由本局填註)

380223

發明專利說明書

一、發明 名稱	中 文	於非同步客戶/伺服器計算系統中針對一組伺服器執行工作 負載管理之資料處理裝置、方法及電腦程式產品
	英 文	"DATA PROCESSING APPARATUS, METHOD AND COMPUTER PROGRAM PRODUCT FOR CARRYING OUT WORKLOAD MANAGEMENT WITH RESPECT TO A GROUP OF SERVERS IN AN ASYNCHRONOUS CLIENT/SERVER COMPUTING SYSTEM"
二、發明 人	姓 名	1. 安德魯 大衛 詹姆斯 班克斯 2. 李察 史考特 麥斯威爾 3. 丹尼爾 納爾 米爾沃茲
	國 籍	均英國
三、申請人	住、居所	1. 英國漢普夏郡若希區米契瑪希市新路藍鈴屋 2. 英國漢普夏郡若希區懷特納波克洛斯路8號 3. 英國漢普夏郡南阿波頓市豐德萊巷203號
	姓 名 (名稱)	美商萬國商業機器公司
三、申請人	國 籍	美國
	住、居所 (事務所)	美國紐約州阿蒙市
三、申請人	代 表 人 姓 名	費羅普

裝

訂

線

(由本局填寫)

承辦人代碼：
大 類：
I P C分類：

A6

B6

本案已向：

國(地區) 申請專利，申請日期： 案號： ，☐有 ☐無主張優先權
英國 1998年4月14日 9807683.9 ☐有 ☒無主張優先權

有關微生物已寄存於： ，寄存日期： ，寄存號碼：

(請先閱讀背面之注意事項再填寫本頁各欄)

裝

訂

線

五、發明說明(1)

發明範圍

(本發明與電腦系統有關，特別是客戶/伺服器(也稱為“分散式”)計算系統，其中一計算裝置(“客戶端”)要求另一計算裝置(“伺服器端”)執行部份的客戶端工作。)

發明背景

客戶/伺服器計算過去幾年在資訊科技世界中變得越來越重要。這種型態的分散式計算允許一機器指派其一些工作給，例如可能較適合執行該工作的，另一機器。例如，伺服器可以是一執行資料庫程式管理著大量資料儲存的強力電腦，而客戶僅是在其一區域程式中由資料庫要求資訊使用的桌上個人電腦(PC)。

訊息佇列(也稱為訊息代理)資料處理技術在現今的客戶/伺服器電腦網路中已變得越來越普遍。此技術允許一客戶電腦系統與一伺服器電腦系統聯繫，即使此二系統在作業系統上、資料格式上與通訊協定上彼此非常不同。再者，由於此技術非同步之特性，客戶可以送給伺服器一訊息，而伺服器可以將此訊息儲存在一佇列中並在稍後時處理與回應該訊息。這與要求客戶與伺服器即時交談(例如，在客戶繼續其他工作之前，客戶將等待伺服器的回應)的同步客戶/伺服器模式相當不同。

訊息佇列與已上市的訊息佇列產品在1994年McGraw-Hill出版B.Blakeley, H.Harris & R.Lewis, 所著的 "Messaging and Queuing Using the MQI" 中，與可由IBM公司買到的下列出版品："An Introduction to

(請先閱讀背面之注意事項再填寫本頁)

裝

訂

五、發明說明(2)

Messaging and Queuing"(IBM文件號碼GC33-0805-00)與"MQSeries - Message Queue Interface Technical Reference"(IBM文件號碼SC33-0850-01)中做了描述。IBM與MQSeries是IBM公司的商標。IBM的MQSeries訊息軟體產品提供處理訊息的支援，在工作的邏輯單元中依據一訊息協定同時發送訊息以確保即使在系統事件或通訊失敗中訊息之發送。MQSeries產品藉由確認訊息安全地儲存到接收系統以前，不做發送系統中所儲存訊息的最後刪除，與使用複雜的回復設備以確保訊息之送達。在確認成功儲存所傳送的訊息之前，發送系統中所儲存訊息的刪除與接收系統中所儲存訊息的插入保持“存疑”，且可在失敗事件中做些微回復。此訊息傳輸協定與相關處理概念及回復設備在國際專利申請WO 95/10805與US專利5465328中做了描述，並列在此以爲參考。

在一非同步客戶/伺服器網路中，諸如使用訊息與佇列軟體架構者，提供每一伺服器皆可滿足一客戶要求的一群伺服器變得越來越平常。如此，客戶的要求可以更快地被服務，因爲若一伺服器正在忙則要求可以被送到另一伺服器。基本上，客戶端的工作負載管理單元使用每個伺服器的可使用資料爲基礎以決定哪個伺服器應被指派負責一特定要求。若這些可使用資料未完全可用，工作負載管理單元等到接收到這些資料後，使用整組資料以決定一伺服器。然後客戶的要求帶著工作負載管理單元所選擇伺服器

(請先閱讀背面之注意事項再填寫本頁)

裝

訂

五、發明說明(3)

的附加指示，被置入傳輸佇列等候傳送到所選的伺服器上。此客戶也將有關佇列中的要求，包括所選擇伺服器的標示之一些資訊寫入到區域儲存中，以便發生傳輸問題致伺服器未能接收到所傳輸的客戶要求時，該要求可以由客戶再送一次到伺服器。

在此技術領域的目前狀態是不便利的，因為工作負載管理單元必須涉及複雜的軟體部份，它持續地檢查每個伺服器的伺服器可使用資料的狀態，以判斷何時由伺服器完整地收到所有這些資料。此複雜的軟體規劃增加了系統整體的大量額外成本。

再者，若尚未收到所有的伺服器可使用資料，一要求不能傳送到該組伺服器，因此暫停住客戶端的資料傳輸處理。這是非常不利的，特別是在一高傳輸速度非常重要的環境中。

發明概要

首先，本發明提供一種資料處理裝置以實施一非同步客戶/伺服器計算系統中一組伺服器個別的工作負載管理，有：一接收來自客戶應用程式工作要求的裝置；一裝置，用以查閱有關該組伺服器的可使用資料，及用以在該組伺服器中以所查閱得之可使用資料為基礎，選擇一特定伺服器來滿足所收到的工作要求，其中為了由該組伺服器中選擇一伺服器，所查閱得之可使用資料皆係使用收到工作要求當時的目前狀態，而不論是否收到該組伺服器中每一伺服器之可使用資料；及一裝置，用以將所收到的工作要求

(請先閱讀背面之注意事項再填寫本頁)

裝

訂

五、發明說明(4)

置入一發訊佇列，在其中工作要求等待發送到該組伺服器中之該特定伺服器。

最好，該裝置進一步有：一裝置以判斷在接收到工作要求後可使用資料是否有改變；及一裝置以在可使用資料有改變時查閱改變的可使用資料，並用以在該組伺服器中以改變後之可使用資料為基礎，選擇一伺服器來滿足所收到的工作要求；及一裝置用以將發訊佇列中的工作要求更正，以在根據改變後之可使用資料為基礎所選擇伺服器不同於該特定伺服器時，使該工作要求被送到該組伺服器中一與該特定伺服器不同之伺服器。

最好再進一步，改變的可使用資料包括了在最初收到工作要求時，該組伺服器中沒有可使用資料伺服器新增加的可使用資料。

或者，改變的可使用資料包括了在最初收到工作要求時，該組伺服器中有可使用資料伺服器變更過的可使用資料。再者，本發明提供一種實施上述工作負載管理裝置功能的方法。

第三，本發明提供一種儲存在電腦的可讀取儲存媒介上的電腦程式產品，當其在電腦上執行時用以實施上述功能。

如此，以本發明，工作負載管理單元軟體可以大為簡化，因為不需要持續地檢查是否已收到所有的伺服器可使用資料。代替的是，一簡單的軟體架構，可用在所有客戶要求都被相同地處理，換句話說，當客戶發出要求就用伺服器可使用資料的目前狀態做決定，並該要求送到傳輸佇

(請先閱讀背面之注意事項再填寫本頁)

裝

訂

五、發明說明(5)

列。而且，其目標也記錄在客戶要求的連續儲存中當成相同運作的一部份。

在要求由傳輸佇列被送到所選擇的伺服器之前應會有新的可使用資料可用，應可以選擇新的伺服器並將佇列中的要求更新以送到不同的伺服器。但，在要求由傳輸佇列被傳送之前若無新的可使用資料可用，則要求會被送到所選擇的伺服器而不管是否收到該組伺服器中每一可用伺服器的可使用資料。如此，增加的好處是工作負載管理單元不用忍受必須等待每一可用伺服器的可使用資料造成的傳輸延誤。本發明使用在工作要求剛收到當時可用的可使用資料做成一初始決定，以選擇該組伺服器中的一個伺服器。只有在非常狀況中新的可使用資料才會出現，做成目標的新記錄。

圖示概述

本發明的一實施例將配合下列圖示說明：

圖1是顯示本發明較佳實施例運作的基本客戶/伺服器環境；

圖2是說明根據本發明一較佳實施例的發訊與佇列單元所採行步驟的流程圖；及

圖3是說明根據本發明一較佳實施例發訊與佇列單元的工作負載管理單元所查閱可使用資料的表列。

發明詳述

在圖1中一客戶資料處理裝置11與一群12伺服器資料處理裝置(121, 122, 123)透過傳輸網路13連接在一起。

(請先閱讀背面之注意事項再填寫本頁)

表

訂

五、發明說明(6)

客戶11與伺服器群12正執行著發訊與佇列軟體產品，如IBM的MQSeries，以便在客戶與伺服器間進行非同步傳輸。也就是，當客戶送一工作要求給伺服器時，伺服器不需要是可運作的。該要求將被置入一佇列中而伺服器可在稍後其可運作時取出佇列中的要求。發訊與佇列產品也執行任何必要的傳輸協定交談以便讓客戶11與伺服器群12在傳輸網路13上對談。

在較佳實施例中，客戶與伺服器群位於傳輸網路13上連接的個別處理單元，但本發明並不限於此架構。客戶與伺服器群可位於相同資料處理單元。

客戶11包含一客戶應用程式111與一發訊與佇列軟體系統112，它包括一工作負載管理單元112a與一傳輸佇列112b。客戶應用程式111是發出工作要求給伺服器的消耗軟體部份。當伺服器進行所要求的工作後，結果可以送回到客戶應用程式111。傳輸佇列112b是發訊與佇列軟體架構的一部份，並用來將所收到來自客戶應用程式111的工作要求透過傳輸網路13傳送到伺服器群12之前予以佇列起來。工作負載管理單元112a，也是發訊與佇列軟體的一部份，是負責為所收到來自客戶應用程式111的每一工作要求，由伺服器群12中選擇伺服器121，122，123之一的軟體元件。

工作負載管理單元112a由傳輸網路13上接收有關每一伺服器121，122，123的資料(以下稱為“可使用資料”)。此資料以每個伺服器接受客戶的新工作要求的就緒

(請先閱讀背面之注意事項再填寫本頁)

訂

五、發明說明(7)

程度來指出每個伺服器的目前可用性。工作負載管理單元以此可使用資料來決定伺服器121，122，123中的哪一個應被選來滿足特定的工作要求。

現合併圖2中的流程圖描述發訊與佇列系統112的運作。

在步驟21，發訊與佇列系統112由客戶應用程式111接收一工作要求。在步驟22，工作負載管理單元112a查閱一可使用資料表列(見圖3)以決定伺服器群12中伺服器的可使用性來處理所收到的工作要求。

在圖3中，剛收到工作要求當時當成時點1。在時點1，可使用表列指出伺服器1(121)可用以接收工作要求(如時點1與伺服器1的交集方塊中字母A所代表“可使用”)。在時點1，可使用表列指出伺服器2(122)不可用以接收工作要求(如時點1與伺服器2的交集方塊中字母NA所代表“不可使用”)。在時點1，應注意圖3的表列中並無登錄伺服器3(123)，因此表示伺服器3尚未提供其可使用資料給客戶11中的發訊與佇列軟體112。

在時點1所給的可使用資料的狀態，工作負載管理單元112a選擇(步驟23)伺服器1(121)指派給該工作要求，因為伺服器1可使用，伺服器2不可使用而伺服器3不知是否可使用。然後工作要求被置入(步驟24)傳輸佇列112b，且一控制指示被加到工作要求以指出該工作要求將被送到伺服器1(121)。然後該工作要求在佇列112b中等到伺服器1(121)指示其已就緒可接收要求(伺服器1可能與客戶11在不同的時區且在時點1不可運作)。

五、發明說明(8)

時點1稍後的時點2，客戶11接收到更新的伺服器可使用資料，且由發訊與佇列軟體112在步驟25中做決定。如此，控制迴路回到步驟22，在其中工作負載管理單元112a再次查閱可使用資料並做決定：伺服器1仍然可用(如時點2與伺服器1的交集方塊中字母A所代表“可使用”)；伺服器2現已可用(如時點2與伺服器2的交集方塊中字母A所代表“可使用”)；而伺服器3不可使用(如時點2與伺服器3的交集方塊中字母NA所代表“不可使用”)。因此，在時點1與時點2間的時間中，新的可使用資料由伺服器群12被送到客戶11，伺服器2改變其可使用性，而伺服器3提供了之前未提供的可使用資料。

因此，在步驟23中，工作負載管理單元112a以此更新過的可使用資料選擇三個伺服器中的一個。任何工作負載平衡規劃方法都可以使用，但最有效率的將會是若兩個伺服器都可使用，且其中之一在流程圖的前一回次中已被選用，則仍使用該伺服器以避免花費處理能力修正佇列112b中的工作要求。這也節省了區域儲存的資料寫入運作，當每個佇列中的工作要求所選擇的伺服器必須登入到區域儲存以在傳輸失敗時作回復的情形下。因為，在此情形中，步驟24將不會更新佇列中的工作要求。

然後，在時點2稍後的時點3中，工作要求仍在佇列112b中，且已收到新的伺服器可使用資料，這些資料的接收在步驟25被發訊與佇列軟體112所注意到。因此，控制迴路回到步驟22，在其中再次查閱了圖3中可使用表列

(請先閱讀背面之注意事項再填寫本頁)

裝

訂

五、發明說明(9)

以決定依據新收到的伺服器可使用資料是否需要改變所選擇的伺服器。如圖3中所示，伺服器1現不可使用，伺服器2仍然可用而伺服器3現為可用。因此，步驟23中，工作負載管理單元112a選擇伺服器2或伺服器3兩者之一。再次，任何通用的工作負載平衡演算法都可以用來選擇此二伺服器應用何者。根據“round-robin”方法，選擇了伺服器2，因為它排在伺服器1之後(而伺服器3會是在排列中接著的回合被選用的)。

然後，在步驟24，傳輸佇列112b中的工作要求被更新，以使該要求現在包含的目標為伺服器2而非伺服器1。這也涉及執行區域儲存的寫入以提供一資料回復時的必要記錄。

在此運作當中的某些點上，所選擇的伺服器會由佇列112b抓取工作要求，且該要求將透過網路13傳輸到伺服器2，如此結束圖2中一特定工作要求的循環流程。

應注意到本發明運作的基本假設是伺服器可使用資料並不經常改變，且在步驟25非常少必要採用YES分支而回到步驟22再查閱圖3的表列。因此，在客戶應用程式發出要求後很早就選擇了初始的伺服器。這使工作要求在使用目前可用的伺服器可使用資料時盡快被置入傳輸佇列112b，以便要求可被快速而有效地處理以傳輸到伺服器群。只有伺服器可使用資料的改變需要修正初始的伺服器選擇。

(請先閱讀背面之注意事項再填寫本頁)

裝

訂

四、中文發明摘要(發明之名稱:於非同步客戶/伺服器計算系統中針對一組伺服器執行工作負載管理之資料處理裝置、方法及電腦程式產品)

一種於非同步客戶/伺服器計算系統中針對一組伺服器執行工作負載管理之資料處理裝置，具有：一用以接收來自客戶端應用程式的工作要求之裝置；一用以查閱有關該組伺服器的可使用資料之裝置，並用以在該組伺服器中以所查閱得之可使用資料為基礎，選擇一特定伺服器來滿足所收到的工作要求，其中，為了由該組伺服器中選擇一伺服器，所查閱得之可使用資料皆係使用收到工作要求當時的目前狀態，而不論是否收到該組伺服器中每一伺服器之可使用資料；以及一用以將所收到的工作要求置入一傳輸佇列之裝置，在後佇列中工作要求等待傳輸到該組伺服器中之該特定伺服器。

英文發明摘要(發明之名稱: DATA PROCESSING APPARATUS, METHOD AND COMPUTER PROGRAM PRODUCT FOR CARRYING OUT WORKLOAD MANAGEMENT WITH RESPECT TO A GROUP OF SERVERS IN AN ASYNCHRONOUS CLIENT/SERVER COMPUTING SYSTEM)

A data processing apparatus for carrying out workload management with respect to a group of servers in an asynchronous client/server computing system, has: a means for receiving a work request from a client application; a means for consulting availability data concerning the group of servers and for selecting a particular server in the group to be used to satisfy the received work request on the basis of the consulted availability data, where the consulted availability data is always used in its present state when the work request is received, irrespective of whether availability data has been received for each server in the group, in order to select a server from the group; and a means for placing the received work request in a transmission queue where it awaits transmission to the particular server in the group.

六、申請專利範圍

1. 一種於一非同步客戶/伺服器計算系統中針對一組伺服器執行工作負載管理之資料處理裝置，該裝置包含：
 - 一接收來自客戶應用程式工作要求的裝置；
 - 一用以查閱有關該組伺服器的可使用資料之裝置，並用以在該組伺服器中以所查閱得之可使用資料為基礎，選擇一特定伺服器來滿足所收到的工作要求，其中為由該組伺服器中選擇一伺服器，所查閱得之可使用資料皆係使用收到工作要求當時的目前狀態，而不論是否收到該組伺服器中每一伺服器之可使用資料；以及
 - 一用以將所收到的工作要求置入一發訊佇列之裝置，在該佇列中工作要求等待發送到該組伺服器中之該特定伺服器。
2. 如申請專利範圍第1項中的裝置，進一步包括：
 - 一裝置以判斷在接收到工作要求後可使用資料是否有改變；及
 - 一裝置以在可使用資料有改變時查閱改變的可使用資料，並用以在該組伺服器中以改變後之可使用資料為基礎，選擇一伺服器來滿足所收到的工作要求；及
 - 一裝置用以將發訊佇列中的工作要求更正，以在根據改變後之可使用資料為基礎所選擇伺服器不同於該特定伺服器時，使該工作要求被送到該組伺服器中一與該特定伺服器不同之伺服器。
3. 如申請專利範圍第2項中的裝置，其中改變的可使用資料包括了在最初收到工作要求時，該組伺服器中沒有可

六、申請專利範圍

使用資料的伺服器新增加的可使用資料。

4. 如申請專利範圍第2項中的裝置，其中改變的可使用資料包括了在最初收到工作要求時，該組伺服器中有可使用資料伺服器變更過的可使用資料。
5. 一種於一非同步客戶/伺服器計算系統中針對一組伺服器執行工作負載管理之資料處理方法，該方法包含步驟：

接收一來自一客戶應用程式之工作要求；

查閱有關該組伺服器的可使用資料，且在該組伺服器中以所查閱得之可使用資料為基礎，選擇一特定伺服器來滿足所收到的工作要求，其中為了由該組伺服器中選擇一伺服器，所查閱得之可使用資料皆係使用收到工作要求當時的目前狀態，而不論是否收到該組伺服器中每一伺服器之可使用資料；以及

將所收到的工作要求置入一發訊佇列，在其中工作要求等待發送到該組伺服器中之該特定伺服器。

6. 如申請專利範圍第5項中的方法，進一步包括步驟：

判斷在接收到工作要求後可使用資料是否有改變；及

在可使用資料有改變時查閱改變的可使用資料，並用以在該組伺服器中以改變後之可使用資料為基礎，選擇一伺服器來滿足所收到的工作要求；及

將發訊佇列中的工作要求更正，以在根據改變後之可使用資料為基礎所選擇伺服器不同於該特定伺服器時，使該工作要求被送到該組伺服器中一與該特定伺服器不同之伺服器。

(請先閱讀背面之注意事項再填寫本頁)

張

訂

六、申請專利範圍

7. 如申請專利範圍第6項中的方法，其中改變的可使用資料包括了在最初收到工作要求時，該組伺服器中沒有可使用資料的伺服器新增加的可使用資料。
8. 如申請專利範圍第6項中的方法，其中改變的可使用資料包括了在最初收到工作要求時，該組伺服器中有可使用資料伺服器變更過的可使用資料。
9. 一種儲存在一電腦的可讀取儲存媒介上的電腦程式產品，當其在該電腦上執行時，用以於一非同步客戶/伺服器計算系統中針對一組伺服器執行工作負載管理之資料處理方法，該方法包含步驟：

接收一來自一客戶應用程式的工作要求；

查閱有關該組伺服器的可使用資料，且在該組伺服器中以所查閱得之可使用資料為基礎，選擇一特定伺服器來滿足所收到的工作要求，其中為了由該組伺服器中選擇一伺服器，所查閱得之可使用資料皆係使用收到工作要求當時的目前狀態，而不論是否收到該組伺服器中每一伺服器之可使用資料；以及

將所收到的工作要求置入一發訊佇列，在其中工作要求等待發送到該組伺服器中之該特定伺服器。

10. 如申請專利範圍第9項中的產品，進一步包括步驟：

判斷在接收到工作要求後可使用資料是否有改變；及

在可使用資料有改變時查閱改變的可使用資料，並用以在該組伺服器中以改變後之可使用資料為基礎，選擇一伺服器來滿足所收到的工作要求；及

(請先閱讀背面之注意事項再填寫本頁)

張

訂

六、申請專利範圍

將發訊佇列中的工作要求更正，以在根據改變後之可使用資料為基礎所選擇伺服器不同於該特定伺服器時，使該工作要求被送到該組伺服器中一與該特定伺服器不同之伺服器。

11. 如申請專利範圍第10項中的產品，其中改變的可使用資料包括了在最初收到工作要求時，該組伺服器中沒有可使用資料的伺服器新增加的可使用資料。
12. 如申請專利範圍第10項中的產品，其中改變的可使用資料包括了在最初收到工作要求時，該組伺服器中有可使用資料伺服器變更過的可使用資料。

(請先閱讀背面之注意事項再填寫本頁)

裝

訂

380223

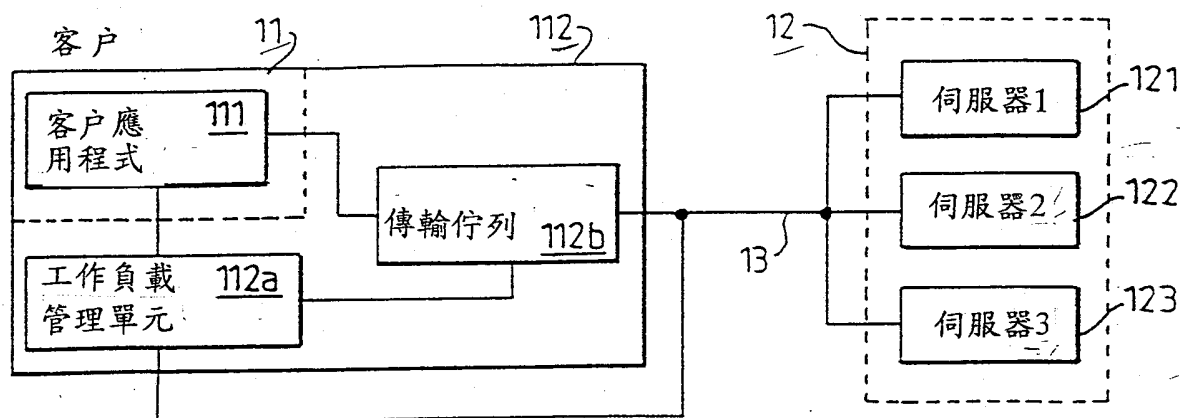


圖 1

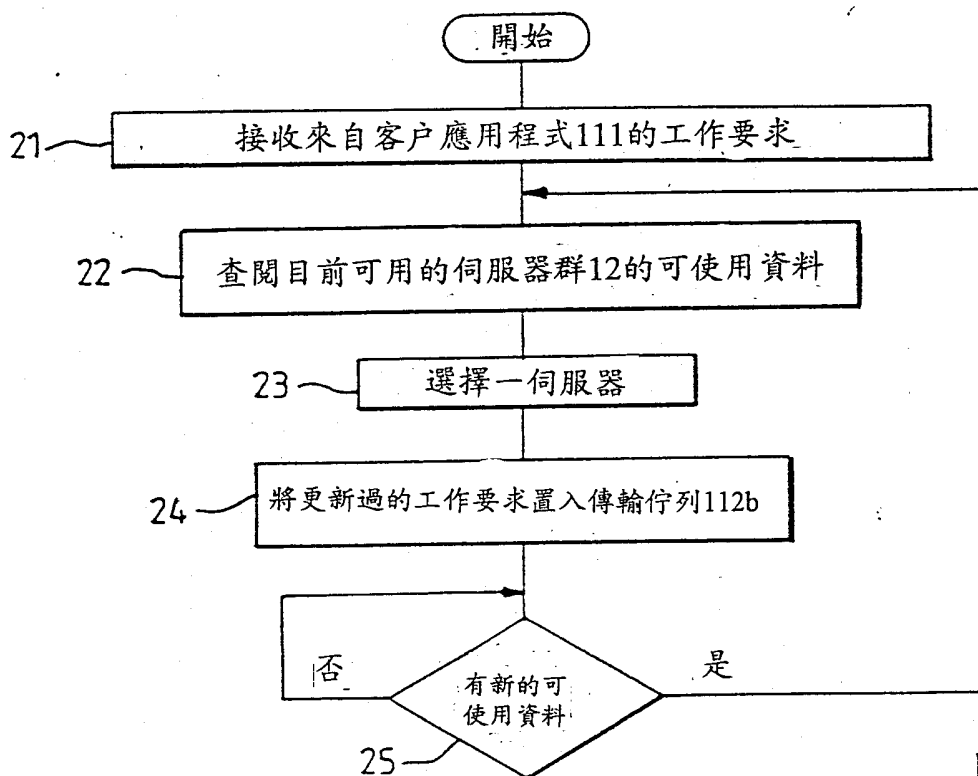


圖 2

時點 \ 伺服器	伺服器1	伺服器2	伺服器3
1	A	NA	
2	A	A	NA
3	NA	A	A

圖 3