

# 發明專利說明書

(填寫本書件時請先行詳閱申請書後之申請須知，作※記號部分請勿填寫)

※ 申請案號：C2103629 ※IPC分類：H04L 12/44

※ 申請日期：92-2-21

## 壹、發明名稱

(中文) 多重星形網路上管理通信流的反向多工  
(英文) INVERSE MULTIPLEXING OF MANAGED TRAFFIC FLOWS OVER A MULTI-STAR NETWORK

## 貳、發明人 (共 3 人)

發明人 1 (如發明人超過一人，請填說明書發明人續頁)

姓名：(中文) 傑若得·利比薩伊  
(英文) LEBIZAY, Gerald

住居所地址：(中文) 美國紐澤西州 07940 麥狄遜, 遠景街 11-B 號  
(英文) 11-B Prospect Street, Madison, NJ 07940, U.S.A.

國籍：(中文) 比利時 (英文) Belguim

## 參、申請人 (共 1 人)

申請人 1 (如發明人超過一人，請填說明書申請人續頁)

姓名或名稱：(中文) 英特爾公司  
(英文) Intel Corporation

住居所或營業所地址：(中文) 美國, 加州 95052-8119 聖塔克拉拉 58119 郵政信箱, 米生大學大道 2200 號  
(英文) 2200 Mission College Boulevard., P.O. Box 58119, Santa Clara, California 91367, U.S.A.

國籍：(中文) 美國 (英文) U.S.A.

代表人：(中文) 大衛·賽門  
(英文) Simon, David

☐ 續發明人或申請人續頁 (發明人或申請人欄位不敷使用時，請註記並使用續頁)

發明人 2

姓名：(中文) 大衛 W·季石

(英文) GISH, David W.

住居所地址：(中文) 美國,紐澤西州 07457-1511 河溪市,史托佛得區 6 號

(英文) 6 Stratford Place, Riverdale, NJ 07457-1511, U.S.A.

國籍：(中文) 美國

(英文) U.S.A.

發明人 3

姓名：(中文) 亨利 M·米薛爾

(英文) MITCHEL, Henry M.

住居所地址：(中文) 美國,紐澤西州 07470-2614 偉恩,阿姆絡西路 29 號

(英文) 29 Almroth Drive, Wayne, NJ 07470-2614, U.S.A.

國籍：(中文) 美國

(英文) U.S.A.

## 捌、聲明事項

☐ 本案係符合專利法第二十條第一項第一款但書或第二款但書規定之期間，其日期為：\_\_\_\_\_

☐ 本案已向下列國家（地區）申請專利，申請日期及案號資料如下：

【格式請依：申請國家（地區）；申請日期；申請案號 順序註記】

1. \_\_\_\_\_
2. \_\_\_\_\_
3. \_\_\_\_\_

☒ 主張專利法第二十四條第一項優先權：

【格式請依：受理國家（地區）；日期；案號 順序註記】

1. 美國；2002.2.21；10/080,329
2. \_\_\_\_\_
3. \_\_\_\_\_
4. \_\_\_\_\_
5. \_\_\_\_\_
6. \_\_\_\_\_
7. \_\_\_\_\_
8. \_\_\_\_\_
9. \_\_\_\_\_
10. \_\_\_\_\_

☐ 主張專利法第二十五條之一第一項優先權：

【格式請依：申請日；申請案號 順序註記】

1. \_\_\_\_\_
2. \_\_\_\_\_
3. \_\_\_\_\_

☐ 主張專利法第二十六條微生物：

☐ 國內微生物 【格式請依：寄存機構；日期；號碼 順序註記】

1. \_\_\_\_\_
2. \_\_\_\_\_
3. \_\_\_\_\_

☐ 國外微生物 【格式請依：寄存國名；機構；日期；號碼 順序註記】

1. \_\_\_\_\_
2. \_\_\_\_\_
3. \_\_\_\_\_

☐ 熟習該項技術者易於獲得，不須寄存。

## 玖、發明說明

### 【發明所屬之技術領域】

本發明係關於電腦系統，且特定關於其運用諸如無限頻帶 (InfiniBand) 架構的封包交換式組織 (packet-switching fabric) 之系統。

### 【先前技術】

於現今世代之電腦，中央處理單元 (CPU, central processing unit) 係藉著諸如週邊構件介面 (PCI, Peripheral Component Interface) 匯流排或工業標準架構 (ISA, Industry Standard Architecture) 匯流排之一種共享式 (shared) 並行匯流排而連接至系統記憶體與週邊裝置。類似而言，欲伺服連接與連結 (link) 至諸如一遠端儲存器與網路裝置之其他伺服器相關系統的現行伺服器 (server) 係取決於並行匯流排技術。取決於一共享式匯流排輸入/輸出 (I/O, input/output) 架構之伺服器設計係可給予例如 512 MB/秒之潛在的頻寬，其為共享於連接至匯流排的裝置。隨著另外的連接係加至伺服器，每個連接的潛在性能係減低，且 I/O 競爭係增高。

隨著資料路徑寬度增大，且時脈速度係變得更快，共享式並行匯流排係變得成本太高且過於複雜以跟得上系統需求。因應於此，電腦工業係正努力以開發下一代的匯流排標準。諸多提議的新標準係具有某些共同處。其提議為處理所運用於 PCI 之共享式匯流排技術而移至一種點對點 (point-to-point) 交換連接。因此，電腦工業係移向快速、封包化、串行輸入/輸出匯流排架構，於其，電腦主機與週邊

係由其普遍稱為一交換組織之一交換網路所連結。諸多的此種型式之架構係已經提出，且第一個下一代的標準係已為適當。無限頻帶架構係由一群工業領導者發起之協會所促進。

無限頻帶架構係一種 I/O 下部結構(infrastructure)技術，其簡化且加速伺服器-至-伺服器之連接以及對於諸如遠端儲存及網路裝置之其他的伺服器相關系統之連結。無限頻帶組織係於一資料中心之伺服器與遠端網路及儲存裝置之間的連接之中央網路。無限頻帶架構係亦透過多階層之冗餘性而構成高度可利用的資料中心。藉著經由多個連結而連接節點，無限頻帶系統係持續以執行，即使若有一個連結為失效。針對提高的可靠度，於一組織之多個開關係提供冗餘的路徑，其允許當萬一於開關之間的一個連結為失效時之透過該組織的資料之無接縫式(seamless)重新路由(re-routing)。完全冗餘的組織係可構成以供最高階層的可靠度，且若整個組織為失效時而仍可持續以執行。

對於任何網路之一個共同的問題係如何透過網路而提供服務品質(QoS, Quality of Service)。欲提供 QoS，網路通信係必須為差異化。某些通信係應為處理於一種方式，而其他的通信則為以另一種方式。

舉例而言，一實體係可與一網路服務提供者(即:其提供網路之實體)而設定一服務階層協定(SLA, Service Level Agreement)，其載明該實體之通信係恆為具有可利用的某個頻寬(例如:每秒鐘為 10 百萬位元或者 10 Mbps)與等待時間(

例如:小於 1 毫秒(ms))。接著，每當一封包係偵測為來自或為去向該實體時，該封包係應收到特定的處理(handling)。若對於實體之整體流係目前為小於 10 Mbps，則該封包係應為通過而無須為棄置且具有小於 1 ms 之等待時間。此型式之流(flow)係稱為藉由確保遞送(AF, Assured Forwarding)所處理。當目前的流為大於 10 Mbps 而到達的封包係將以不同方式處理，或許是如為最佳效力(BE, Best Effort)通信(參閱下文)。

作為另一個實例，一路由器(router)係可作設定以辨識某些型式之流為即時(real-time)流。即時流係特徵為其想法在於，若封包係未於及時到達，則其將可能同樣為根本未到達。舉例而言，於電話會話中之聲音資料的一封包係必須在當其為需要時而可利用於接收器，否則其為無用。太遲者，其係無法運用而且將作棄置。是以，屬於一聲音會話的即時通信(一串流之封包)係應由其習稱為快速遞送(EF, Expedited Forwarding)之一類的行為所處理。以此方式所處理之一封包係將為極快地遞送(具有低的等待時間(latency))。可希望的是，於等待時間之變化(習稱為顫動(jitter))係亦將為低。如同為一折衷者，於該串流之封包係可為純然棄置，若其整體的頻寬係超過某一個臨限值。此外，適用於該等封包之一 SLA 係可能對於購買者為昂貴，因為提供此種服務係要求的是，路由器具有特徵而使得其為昂貴以建立。

一第三個實例係針對其並非為任何 SLA 所適用之通信

，稱為最佳效力(BE)通信。此種型式之通信係典型為可見於網際網路上。最佳效力之封包係可為針對任何理由而棄置，且不具有關於等待時間或顫動之任何特別的保證。

因此，運用多重開關架構之邊際效應的一者係需要判斷哪些封包為進行於某一開關，因為任何的開關均可到達至任何目的。諸多可能的演算法係可實施至其跨於多重無限頻帶開關子網路之整體資料。是以，存在對於其為簡單、快速、適宜、且強健的一種演算法之需求。

### 【實施方式】

本發明係存在於其運用多個無限頻帶開關子網路而連接於一系統內的多個實體(通常為於一系統內的多個板)之環境。子網路之數目係不重要，只要其為二或更多個。第一圖係說明一種無限頻帶多重星形組織，其具有高達 32 個板 110、111、112、113、與 114 且為互連至三個無限頻帶開關裝置 120、121、與 122。第二圖係說明一種無限頻帶多重星形組織 220，其為連接至一來源節點 A/板 A 210 與一目的節點 B/板 B 230。一多重星形開關組織係一組的並聯星形，各個星形係由其為連接至每個節點之單一個開關所組成，該等開關係並未為互連於任何方式。

參考第三圖，實體(板或節點) 301、302、與 303 係提供對於透過系統的各個封包流之不同階層的服務。節點 301、302、與 303 係由一個或多個電子式處理器所構成，其共同執行本發明之作用。經由連結，節點 301、302、與 303 係藉著透過開關 310、320、與 330 以送出封包而連通於彼

此。因為各個節點 301、302、與 303 係連接至多個開關 310、320、與 330，各個節點 301、302、與 303 係可經由多個路徑(明確而言，每個開關為一或多個路徑)而送出封包至另一個節點 301、302、與 303。

連結(link)係介於節點與開關(或是一般之節點至節點、或開關至開關)之間的連接。其係能夠於任一時間傳送不超過某量的資訊，習稱為頻寬。舉例而言，頻寬係可測量於位元/秒。一般而言，不具有任何限制於開關(N)之數目，除了其須為大於 1。

一流(flow)係一組的封包，其共用一組的特性。典型而言，該等特性係包括:封包之來源與目的位址、以及其協定型式與可能為其優先性與分類。重要的是，於一流之所有封包係維持其送出之某一順序，較佳為以相同的順序而到達其目的。若其為脫序而到達，其係可為重新排序，或依序放回，然而，最終係需要大量工作以重新排序封包。因此，一個良好設計係企圖保持於一流之所有的封包均為依序通過網路，使得其為依序到達遠端而不需要重新排序。

舉例而言，若二個人係經由網際網路(IP, Internet)通話法而通訊，則其涉及四個流:來自各側至另一者之資料串流(stream)(其載有聲音)、與來自各側至另一者之一控制串流(藉其，各側係告知另一者該資料為如何通過)。各個資料串流係由一序列的封包而組成，各者含有僅為一個短暫的會話，或許僅為 5 微秒之長。因此，需有 200 個封包以承載有各個單一秒的電話會話。將為最佳的是，若封包係以其

送出之相同順序而到達。若實際上有一些封包係變換，系統係將仍為運作，然而，較佳為若是其均以進入時之相同順序而流出。

管理通信係由系統操作者尋求確保取得通過系統於某個邊限內之流組而組成。舉例而言，操作者係可能付出以確定自一個特定來源網路至一個特定目的網路之所有通信均為取得通過而無損失(無掉落的封包)，且具有限定的等待時間與顫動(等待時間係一封包欲跨越系統所耗費的時間量，而顫動係於各個封包之等待時間的平均變化)。因為管理通信係一收益(revenue)來源，其為值得建立一系統以強化其需求。

存有二種管理通信流:其頻寬係在流起始之前而作交涉者，習稱為頻寬交涉式管理通信(BNMT, Bandwidth-Negotiated Managed Traffic);及，其頻寬係未在時間前而作交涉者，習稱為未交涉式管理通信(UMT, Un-negotiated Managed Traffic)。

頻寬交涉式管理通信(BNMT)包括:關於服務之流，其具有已知或易於估計的頻寬，諸如於 IP 之聲音(VoIP, Voice over IP)，於其，符合一個特定的定義檔(profile)之一資料流係可能約為 64 Kbps。符合一個不同的定義檔之一 VoIP 資料流係可僅為 20 Kbps。重點在於事實為，將發生之流係建立為超前時間(藉由其設立一 VoIP 頻道之二者所建立)，且該頻道之頻寬係在資料流開始前而為已知。

另一種 BNMT 流係一整體者(aggregate)，對於其，操作

者係不知道其確實的頻寬需求，但其為足夠重要以預留頻寬。此種通信之一個實例係於系統本身之內部控制通信。種種的節點係必須通訊介於其本身之間的控制資訊，且此種通信係極高的優先性—其必須以最小的等待時間而作處理。對於此種 BNNT，操作者係估計所需要的頻寬量且在時間前而預留。

未交涉式管理通信(UMT)包括其非為 BNMT 之所有的管理通信。舉例而言，對於自某一個來源網路至某一個目的網路之所有通信係可能存有一協定以能夠流通為具有最小的等待時間與損失，高達至一預先設定極限之頻寬，之後，超過於其者係均將視為未管理的通信。於此情形，整體係作管理，但是並不知道於該整體中之各個流將消耗多少的頻寬。

所有其他的通信係未管理通信(UT, un-managed traffic)。網際網路通信之大多數者係未管理通信(UT)。未管理通信係經常稱為最佳效力者(BE)，因為其為期望於系統部分之所有效力者。一網路係應該盡其全力使得封包取得通過。然而，若其為棄置、重新排序、或握住於數分鐘，將係可接受。

本發明包括：管理通信流為如何實施。更為明確而言，任何節點係如何選取一封包的管理通信(BNMT 或 UMT)將進行至任何其他節點之路徑。該節點係必須選取自 1 至 N 之開關，封包係將送出至該開關。

自入口至組織輸入之封包的處理係顯示於第四圖。如

圖所繪者係透過單一個節點而行進至開關 420、421、422 之封包流。於該節點上的單一個分類器 410 係分類所有進入的通信而成爲一特定的流。該流係定義一組的參數，其包括：目的節點(A-Z)、藉其而將到達目的節點之路徑(開關 1 至 N，即 420、421、422)、與分類(管理或未管理)。個別的封包係根據該分類而完結於佇列 430、440。

對於每個開關之每個目的節點係存有一個單獨組的佇列 430，以供管理通信。爲了簡化，該組的佇列係顯示爲單一方塊，且標示爲 SSMT 以針對特定開關管理通信(Switch-Specific Managed-Traffic)佇列。

每個開關係維持多個佇列，其針對二個理由。主要理由係在於，多個佇列係提供一推進(steering)作用。在分類之後，開關係已經識別。若封包係放回至相同的佇列，開關資訊係將必須爲隨著封包而承載。此種設計係實際爲進行於某些實施。然而，將封包分離爲單獨的佇列係另一種方式以載有資訊。

第二個理由係防止由於開關之線路頭端阻斷(head-of-line blocking)。此發生情形係應爲一罕見的狀況(開關係不應極爲經常阻塞，但是其將發生，而單獨的佇列係防止此狀況)。此線路頭端阻斷係將發生，若僅存有單一個佇列，且舉例而言，若其預定爲通過開關 1 (310)之一個封包係在其預定爲通過開關 2 (320)之一個封包的前方，且開關 1 (310)係重載而開關 2 (320)爲否。預定爲通過開關 1 (310)之封包係將無法爲脫離佇列，因爲開關 1 (310)係將爲不可利

用，且將阻斷其預定為於開關 2 (320)之封包，即使當其預定為於開關 2 (320)之封包可行進(若其可脫離佇列)。

供判斷一流應為採取哪個路徑之方法係納入一開關之選取，基於一種通信技術(traffic-engineering)演算法。該判斷係嵌入於針對該流之節點的一路由表 411。一路由表 411係載明各個流在其離開節點後所應該行進之處。本發明係載明的是，通信技術演算法係亦編碼該流所應行進通過之開關 420、421、422。

再者，於初始的通信技術判斷為作成之相同時間或稍後，一備用路徑係決定。備用路徑係預先計算，使得若原始路徑失效時，該流係可切換至備用者而不會棄置任何封包。

對於 BNMT，通信技術演算法係可運作於已知的資訊。因為該流頻寬預留係已經為先前建立，負載平衡演算法係置放該流而為隨其所願。舉例而言，此置放係可為於其具有目前最小量的管理通信之開關。或者其可能為，通信技術演算法係寧願一起置放某些類似型式(以其偏好之任何分類)之所有流於相同開關。

對於 UMT 係存有二個選項。於第一個選項，通信技術演算法係運用於其為同意之整體頻寬，即：當該流係初始建立時，其將歸屬於一管理分類，對其係存有一預先交涉的整體頻寬。通信技術演算法係可選取以置放對於該整體之所有通信於一特定開關與於一不同開關之其備用者。以此種方式運作，負載平衡器係本質為分類 UMT 通信而成為一

整體的 BNMT 分類，且接著置放該整體流於每當其應行進時。

供處理一 UMT 流之第二個選項係指定其至一特定開關，且接著度量其於一段時間(即:決定其平均頻寬)。接著，通信技術演算法係可判斷實際以置放該流之處，如同關於 BNMT 所為。

因此，當一管理流係起始，通信技術演算法係指定該流至一特定開關 420、421、422。同時，通信技術演算法係指明一個備用開關，以於主要開關為失效時而承載通信。通信技術演算法係置放主要與備用開關資訊於針對該流之路由表 411。該資訊係接著決定該流為載於其之開關，以確保對於一特定流之封包為並未錯誤排列。

為了通信技術演算法以指定該流至一特定開關，其需要知道(或者估計)該流所將耗費之頻寬。具有三個選項以運用此資訊:當 BNMT 流係起始時，其個別的頻寬需求係已知。故，其可考慮作為該頻寬為已知之一個整體的部分者。或者，其係可為初始指定者，其頻寬係測量，且接著該測量值係運用以修正(若必要時)初始指定者。

未管理之所有通信係藉著未管理之定義，且流至其標示為 UT 之佇列，針對未管理通信佇列 440，每個目的節點為一個佇列。

針對各個開關，於各個節點係存有一開關輸入調度處理 450，自該節點之佇列 430 而取得封包，且將其送至一特定開關 420、421、422。於各個節點，存有針對各個開關之

一個該種處理。此係針對各個開關 420、421、422 之處理，其掌控以選取哪個佇列 430 應為接著作服務。

多個佇列 430 係亦為維持對於同一個開關之每個目的節點，藉以允許介於透過各個單一開關的目的流之間的隨機化。此配置係一種最佳化，其為論述於下文。

當多個節點係同時送出資料至相同的目的節點，該目的節點係將暫時成為過載。此狀況係將發生在特別是當某些處理為碰巧產生介於其送出至該單一個目的節點之不同的輸入節點之間的同步化。該同步化係可發生為如同下述之方式：傳輸控制協定(TCP, Transmission Control Protocol)係透過一路由器而成為同步化，若一末端棄置(tail-drop)策略係運用以供整頓該等輸入佇列而非為一隨機初端棄置(RED, Random Early Drop)策略。

同理，多個節點係可能為偶然一次送出所有資料至相同的開關。雖然平均負載係將為低，於該開關上的瞬間負載係將為相當高而且引起問題。

欲平滑處理該尖峰(spike)之一個方式係引入隨機元件於封包之調度。較佳而言，隨機元件係無法偶然同步化於諸個來源節點，或者其將使得狀況為更糟而非為更佳。

假設某人欲執行隨機化以平滑處理通信，則其可進行此舉為藉著分割該等佇列以使得各個目的節點具有其本身單獨組的管理佇列 430 與其本身的未管理佇列 440，如於第四圖所示。

一隨機化器作用 460、461、462、與 463 係加在各組的

佇列 430、440 之後，以隨機判斷哪個目的節點佇列 430、440 將接著作服務。

隨機化器 460、461、462、與 463 係在正常佇列調度作用之後而運作，舉例而言，使得管理佇列 430 係仍保有其 BNMT 流對(vs.)UMT 流之行爲。

當針對一特定開關之開關輸入調度處理 450、451、與 452 係需要擷取封包，則引取(pull)自其優先佇列 470、471、與 472 (即:管理對(vs.)未管理):

1)首先，引取自針對管理佇列之隨機化器 460、461、462。

a.若存有可用的封包於管理佇列 430，管理佇列隨機化器 460、461、462 係透過標準佇列引取機構而引取。舉例而言，頻寬交涉管理通信(BNMT)佇列係將首先引取，接著爲該未交涉管理通信(UMT)佇列。注意，隨機化器係必須詢問(poll)所有 BNMT 佇列且在詢問任何 UMT 佇列之前而選自可利用者。

2)若於管理佇列隨機化器 460、461、462 不具有可利用的封包，開關輸入調度處理 450、451、452 係引取自針對未管理佇列之隨機化器 463。若存有可利用的封包，則自該等可利用者而隨機選出一個佇列，且返回該佇列之頭端封包。

3)此外，若尚未具有對於該開關之可利用的封包，開關輸入調度處理 450、451、452 係再次詢問。

第五圖說明一個流程圖，針對實施一種用於多重星形

無限頻帶網路上的管理通信封包之反向多工處理的方法。

於節點上的單一個分類器 410 係分類所有進入的通信而成為一特定流(於步驟 500)。該流係定義一組參數，其包括:目的節點(A-Z)、藉其而將到達目的節點之路徑(開關 1 至 N，即: 420、421、422)、與分類(管理 BNMT、UMT 或未管理)。

該分類器係運用一種通信技術演算法，以判斷哪個路徑為一流所應採取者、與哪個開關為該流所應通過者。該判斷係可嵌入於針對該流之節點的一路由表(於步驟 510)。一路由表係載明各個流在其離開該節點後而應行進之處。一個備用路徑係亦為決定，使得若原始的路徑為失效時，該流係可切換至備用路徑(於步驟 520)。個別的封包係根據分類而送至佇列 430 (於步驟 530)。

一開關輸入調度處理 450 係自該節點之佇列而取得封包，且將其送至一特定開關。於各個節點，存有對於各個開關之一個該種處理。對於各個開關，此處理係掌控選取哪個佇列為接著應作服務。

一隨機化器作用係開關輸入調度處理之部分者。隨機化器作用係隨機判斷哪個目的節點為接著將作服務(於步驟 540)。當針對一特定開關之開關輸入調度處理 450、451、452 係需要擷取一封包(於步驟 550)，則引取自其優先佇列 470、471、472 (即:管理對(vs.)未管理)，其引取自用於管理佇列之隨機化器 460、461、462。若於管理佇列 430 係存有可用的封包，管理佇列隨機化器 460、461、462 係透過標準佇列引取機構而引取(於步驟 560)。若於管理佇列隨機化

器 460、461、462 係不具有任何可用的封包，開關輸入調度處理 450、451、452 係引取自未管理佇列之隨機化器 463。若其存有可用的封包，則自該等可用者而隨機選取一佇列且返回該佇列之頭端封包(於步驟 570)。此外，尚未存有對於該開關之可用的封包，且開關輸入調度處理 450、451、452 係再次詢問。由該開關輸入調度處理所選取的封包係送至特定開關(於步驟 580)。來源節點係傳送該等封包為分類後的流，透過一開關組織至目的節點。

於目的節點，透過開關組織而傳送之封包係由開關輸出處理 601、602、603 而排序，且送出至意欲的輸出佇列。參閱第六圖。

本發明係藉著多工處理其為跨於一多重星形無限頻帶開關網路之管理與未管理的通信流而利用可用的頻寬。欲提供服務之品質，網路通信係差異化。一開關組織之運用係提供多個階層之冗餘性。一隨機化作用係運用以平滑處理該差異化的通信流。

儘管以上敘述係參照本發明之特定實施例，將瞭解的是，諸多的修改係可作出而不偏離其精神。伴隨之申請專利範圍係意欲以涵蓋該等修改，如同歸屬於本發明之範疇與精神。因此，目前揭示的實施例係在各方面均視為說明性質而非為限制性質，由隨附的申請專利範圍而非前述的說明所指出之本發明的範疇以及成為於申請專利範圍之等效性的意義與範圍內之所有變化係因此意欲為包含於其。

【圖式簡單說明】

### (一) 圖式部分

第一圖說明根據本發明的一個實施例之一種無限頻帶多重星形組織，其具有互連至三個無限頻帶開關裝置之高達 32 個板；

第二圖說明根據本發明的一個實施例之一種無限頻帶多重星形組織，其為連接至一來源節點 A/板 A 與一目的節點 B/板 B；

第三圖說明根據本發明的一個實施例之連接至多個開關的多個節點；

第四圖說明根據本發明的一個實施例之自入口至組織輸入的封包處理；

第五圖說明根據本發明的一個實施例之一個流程圖，用以實施於多重星形無限頻帶網路上之管理通信流的反向多工之一種方法；及

第六圖說明根據本發明的一個實施例，其傳送自來源節點 A 而透過一開關組織至一組織輸出之封包係由一開關輸出處理而作用於目的節點 B。

### (二) 元件代表符號

|         |               |
|---------|---------------|
| 100-114 | 板             |
| 120-122 | 無限頻帶(IBA)開關裝置 |
| 210     | 來源節點 A/板 A    |
| 220     | 多重星形組織        |
| 230     | 目的節點 B/板 B    |
| 300     | 入口/出口         |

|             |                         |
|-------------|-------------------------|
| 301-303     | 實體(板或節點)                |
| 310、320、330 | 開關                      |
| 410         | 分類器                     |
| 411         | 路由表                     |
| 420-422     | 開關                      |
| 430、440     | 佇列                      |
| 450-452     | 開關輸入調度處理                |
| 460-462     | 管理佇列隨機化器                |
| 463         | 未管理佇列隨機化器               |
| 470-472     | 優先佇列                    |
| 500-580     | 第五圖之流程圖的步驟              |
| 601-603     | 開關輸出處理                  |
| 610         | 序列檢查處理                  |
| 620         | 特定來源節點之未管理通信的緩衝器(SNSUT) |

## 肆、中文發明摘要

一種用以反向多工處理於一多重星形開關網路上的管理通信流之方法係包括一來源節點分類器。運用一種通信技術演算法，來源節點分類器係基於流參數而分類進入的通信，嵌入流參數於針對一流之一節點的一路由表，置放來自分類後的流之封包於特定開關管理通信佇列(SSMT)與一來源節點未管理通信佇列。針對一開關之一來源節點開關輸入調度處理係選取來自預定用於該開關之 SSMT 的所有管理封包，然後選取來自來源節點未管理通信佇列之單一個未管理封包。來源節點係透過一開關組織以傳送其如為分類後的流之封包至目的節點。於目的節點，透過開關組織而傳送之封包係由一開關輸出處理而排序，且送出至意欲的輸出佇列。

## 伍、英文發明摘要

A method for inverse multiplexing of managed traffic flows over a multi-star switch network includes a source node classifier. The source node classifier, using a traffic-engineering algorithm, classifies incoming traffic based on flow parameters, embeds the flow parameters in a routing table in a node for a flow, places packets from classified flows into Switch-Specific Managed-Traffic Queues (SSMT) and a source node unmanaged traffic queue. A source node switch input scheduler process for a switch selects all managed packets from the SSMT destined for the switch, then selects a single unmanaged packet from the source node unmanaged traffic queue. The source node transmits the packets as classified flows through a switch fabric to the destination node. At the destination node packets transmitted through the switch fabric are sorted by a Switch Output Process and sent to intended output queues.

## 拾、申請專利範圍

1.一種多重星形交換網路上的管理通信流之反向多工方法，包含：

基於流參數而分類進入的通信以形成分類後的流，且嵌入該流參數於針對一流之一來源節點的一路由表；

置放來自該分類後的流之封包於特定開關管理通信佇列(SSMT)與一未管理通信佇列；

選取以使得對於一開關之一開關輸入調度處理為選取其可得自 SSMT 之所有管理封包，然後選取來自未管理通信佇列之單一個未管理的封包，以形成所選取的封包；

透過一開關組織，傳送來自該來源節點之分類後的流之選取的封包至一目的節點；

由目的節點而接收分類後的流，其中，一開關輸出處理係排序其接收自一組織輸出之分類後的流，且送出該分類後的流至一意欲的輸出佇列。

2.如申請專利範圍第 1 項之方法，其中該流參數係定義該來源節點、一路徑、一備份路徑、一選取開關、一備份開關、目的節點、以及該流是否為管理或未管理。

3.如申請專利範圍第 2 項之方法，其中，運用一種通信技術演算法之一來源節點分類器係基於流參數而分類進入的通信，嵌入該流參數於針對該流之來源節點的路由表，且置放來自分類後的流之封包於 SSMT 與未管理通信佇列。

4.如申請專利範圍第 3 項之方法，其中該通信技術演算

法係基於該流所需之頻寬而分類。

5.如申請專利範圍第 4 項之方法，其中，一頻寬交握式管理通信(BNMT)流係具有一已知的頻寬需求。

6.如申請專利範圍第 4 項之方法，其中，具有一未知的頻寬需求之一未交握式管理通信(UMT)流係視為一整體之部分者，針對於其，頻寬係已知或為初始置放於針對一選取開關之一佇列，UMT 頻寬係作測量，且測量後的頻寬係運用以修正初始之置放。

7.如申請專利範圍第 2 項之方法，其中該選取係包括藉著開關輸入調度處理以移動來自 SSMT 與未管理通信佇列之分類後的流之所選取的封包至選取的開關。

8.如申請專利範圍第 1 項之方法，其中該藉著開關輸入調度處理之選取係包括一隨機化作用。

9.如申請專利範圍第 1 項之方法，其中該來源節點係具有針對開關之開關輸入調度處理。

10.一種程式碼儲存裝置，包含：

一機器可讀取之儲存媒體；及

機器可讀取之程式碼，其為儲存於該機器可讀取之儲存媒體，且具有指令以基於流參數而分類進入的通信以形成分類後的流，且嵌入該流參數於針對一流之一來源節點的一路由表，

置放來自該分類後的流之封包於特定開關管理通信佇列(SSMT)與一未管理通信佇列，

選取以使得對於一開關之一開關輸入調度處理為先選

取其可得自 SSMT 之所有的管理封包，然後選取來自未管理通信佇列之單一個未管理的封包，以形成所選取的封包

，

透過一開關組織以傳送來自該來源節點之分類後的流之選取的封包至一目的節點，

接收該分類後的流於目的節點，其中，一開關輸出處理係排序其接收自一組織輸出之分類後的流，且送出該分類後的流至一意欲的輸出佇列。

11.如申請專利範圍第 10 項之程式碼儲存裝置，其中該流參數係定義該來源節點、一路徑、一備份路徑、一選取開關、一備份開關、目的節點、以及該流是否為管理或未管理。

12.如申請專利範圍第 10 項之程式碼儲存裝置，其中，運用一種通信技術演算法之一來源節點分類器係基於流參數而分類該進入的通信，嵌入該流參數於針對該流之來源節點的路由表，且置放來自該分類後的流之封包於 SSMT 與未管理通信佇列。

13.如申請專利範圍第 12 項之程式碼儲存裝置，其中該通信技術演算法係基於該流所需之頻寬而分類。

14.如申請專利範圍第 13 項之程式碼儲存裝置，其中，一頻寬交握式管理通信(BNMT)流係具有一已知的頻寬需求。

15.如申請專利範圍第 13 項之程式碼儲存裝置，其中，其具有一未知的頻寬需求之一未交握式管理通信(UMT)流係

視為一整體之部分者，針對於其，頻寬係已知或者為初始置放於針對一選取開關之一佇列，UMT 頻寬係作測量，且測量後的頻寬係運用以修正初始之置放。

16.如申請專利範圍第 11 項之程式碼儲存裝置，其中該開關輸入調度處理係移動來自 SSMT 與未管理通信佇列之封包至選取的開關。

17.如申請專利範圍第 10 項之程式碼儲存裝置，其中該開關輸入調度處理係包括一隨機化作用。

18.如申請專利範圍第 10 項之程式碼儲存裝置，其中該來源節點係具有針對開關之一開關輸入調度處理。

19.一種多重星形交換網路，包含：

一開關組織；

一輸入裝置，其為連接至該開關組織，且具有一路由表、複數個特定開關管理通信佇列(SSMT)、一未管理通信佇列、與一開關輸入調度器，其中，該輸入裝置係基於流參數而分類進入的通信以形成分類後的流，嵌入該流參數於針對一流之路由表，置放來自分類後的流之封包於 SSMT 與未管理通信佇列，選取以使得對於一開關之開關輸入調度器為選取其可得自 SSMT 之所有的管理封包且然後選取來自未管理通信佇列之單一個未管理的封包以形成所選取的封包，透過該開關組織以傳送來自輸入裝置之分類後的流之選取的封包至一輸出裝置；及

該輸出裝置，其為連接至該開關組織，含有一開關輸出處理，其排序所接收自一組織輸出之分類後的流，且送

出該分類後的流至一意欲的輸出佇列。

20.如申請專利範圍第 19 項之多重星形交換網路，其中該流參數係包括該輸入裝置、一路徑、一備份路徑、一選取開關、一備份開關、輸出裝置、以及該流是否為管理或未管理。

21.如申請專利範圍第 20 項之多重星形交換網路，其中該輸入裝置運用一種通信技術演算法以基於流參數而分類該進入的通信，嵌入該流參數於針對該流之路由表，且置放來自分類後的流之封包於 SSMT 與未管理通信佇列。

22.如申請專利範圍第 21 項之多重星形交換網路，其中該通信技術演算法係基於該流所需之頻寬而分類。

23.如申請專利範圍第 22 項之多重星形交換網路，其中，一頻寬交握式管理通信(BNMT)流係具有一已知的頻寬需求。

24.如申請專利範圍第 22 項之多重星形交換網路，其中，具有一未知的頻寬需求之一未交握式管理通信(UMT)流係視為一整體之部分者，針對於其，頻寬係已知或者為初始置放於針對一選取開關之一佇列，UMT 頻寬係作測量，且測量後的頻寬係運用以修正初始之置放。

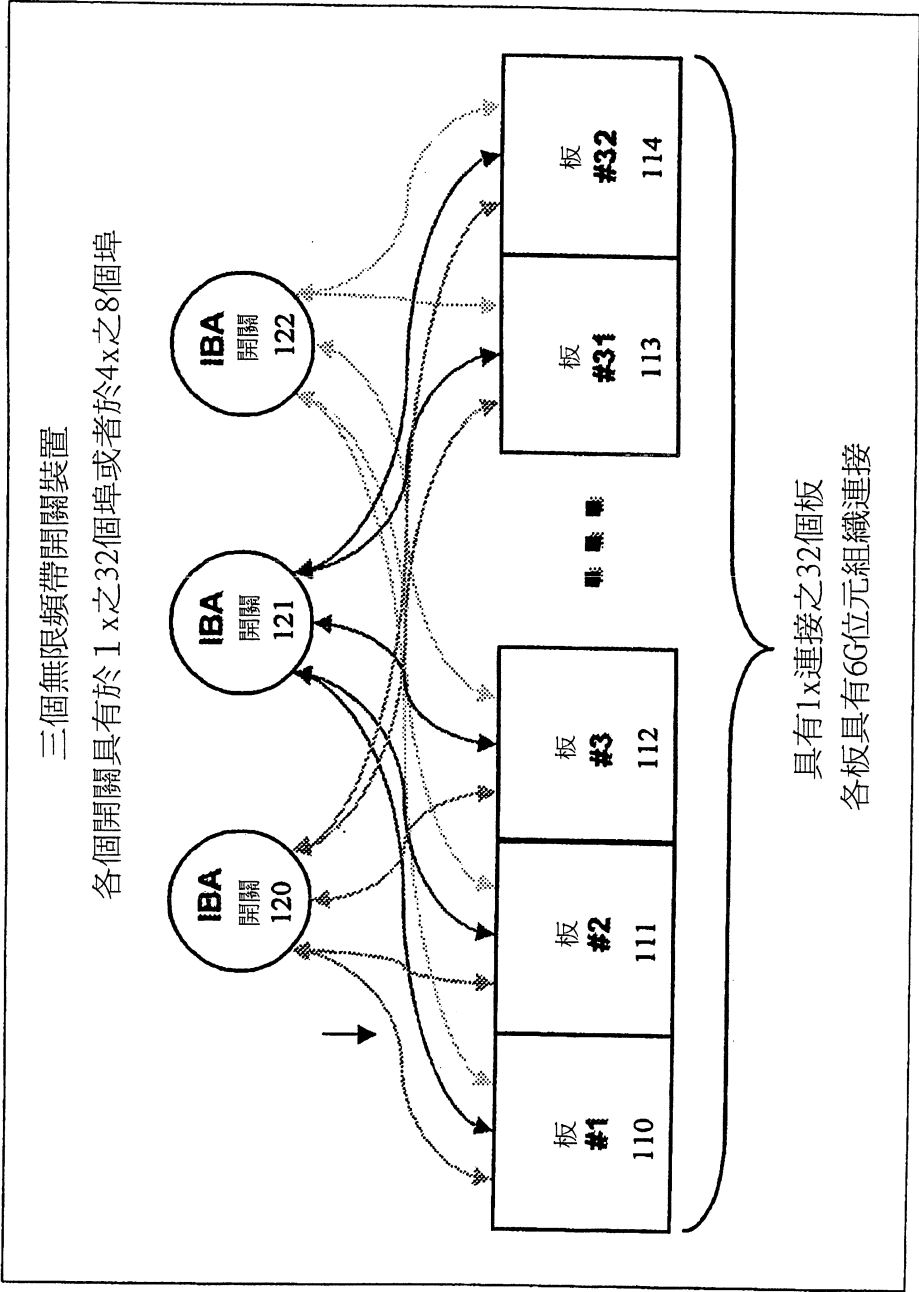
25.如申請專利範圍第 20 項之多重星形交換網路，其中該開關輸入調度器係移動來自 SSMT 與未管理通信佇列之封包至選取的開關。

26.如申請專利範圍第 19 項之多重星形交換網路，其中該開關輸入調度器係包括一隨機化作用。

27.如申請專利範圍第 19 項之多重星形交換網路，其中該輸入裝置係具有針對開關之一開關輸入調度器。

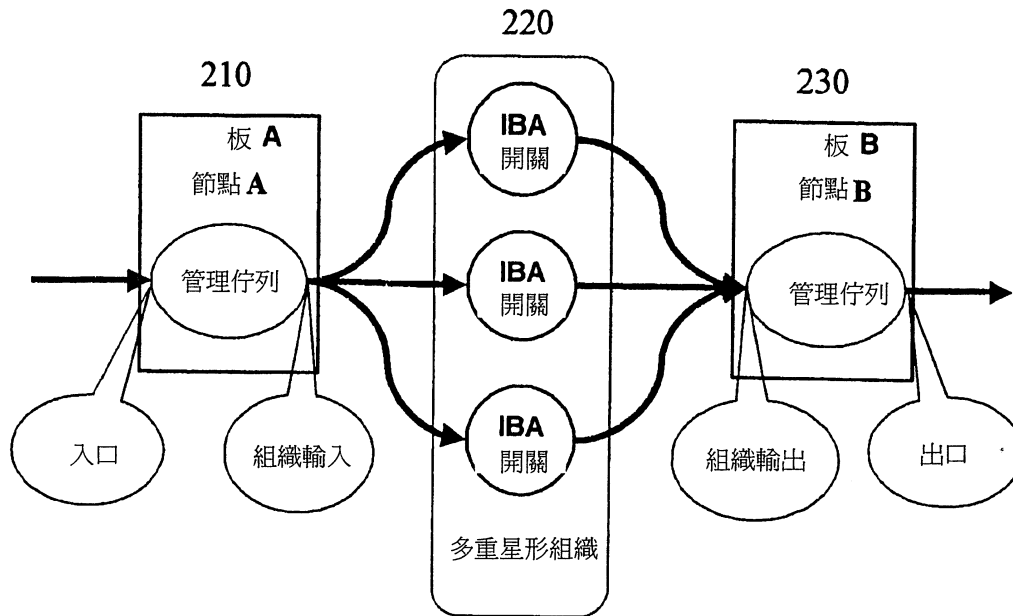
## 拾壹、圖式

如次頁

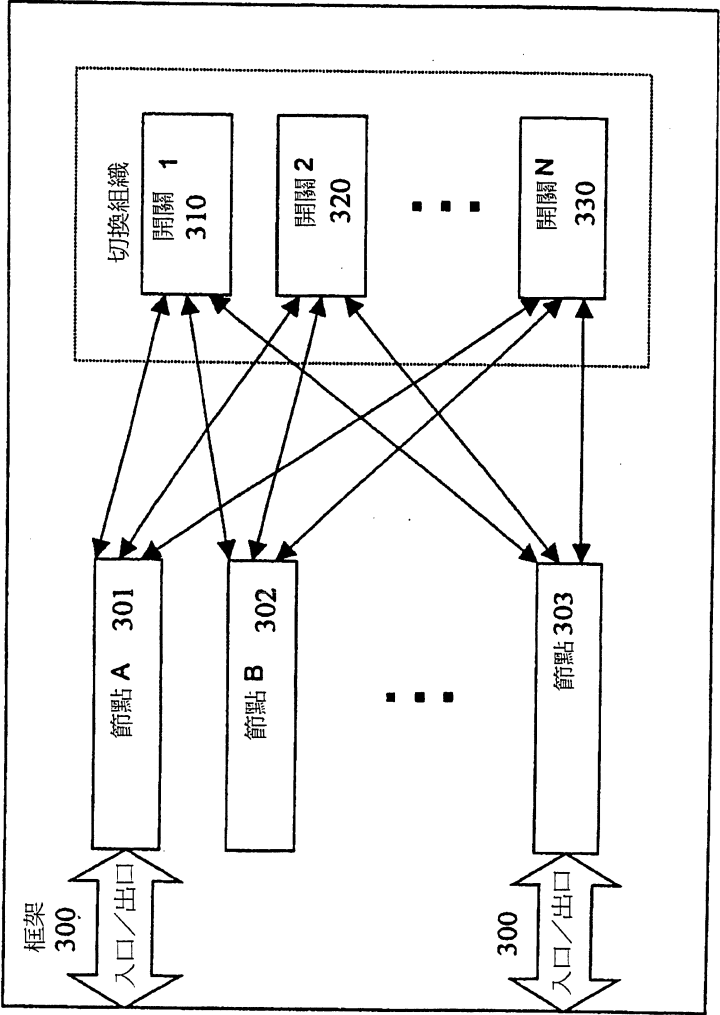


第一圖

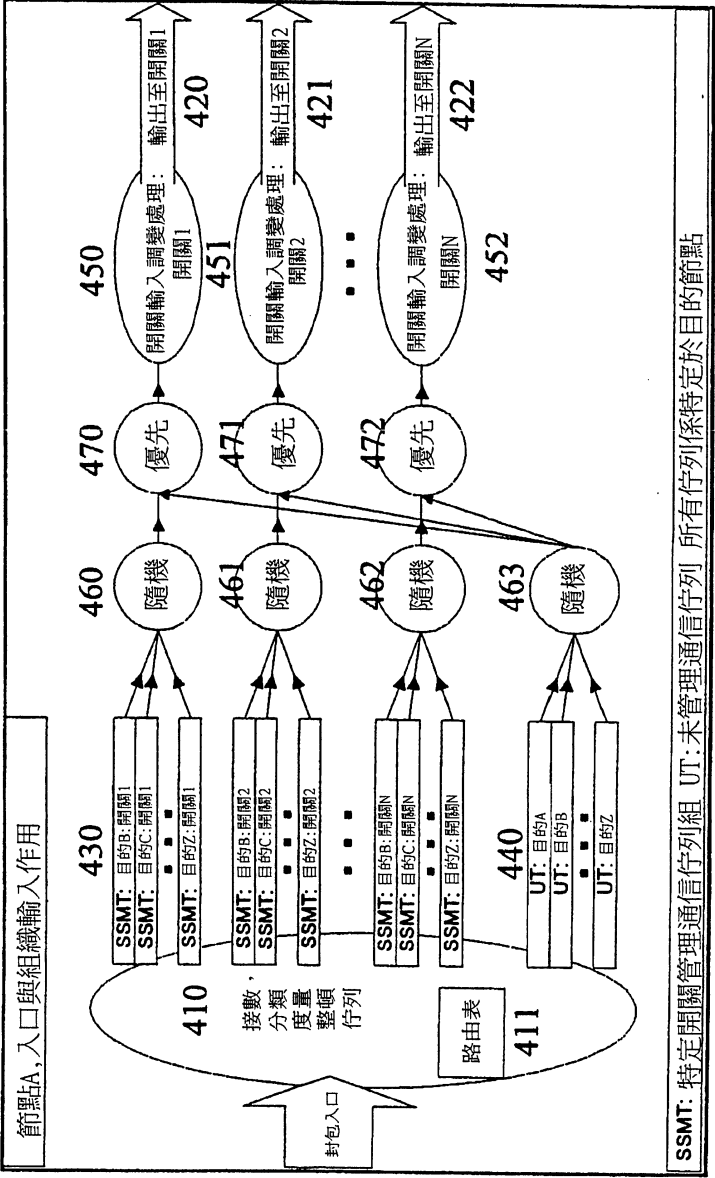
管理佇列點用辭



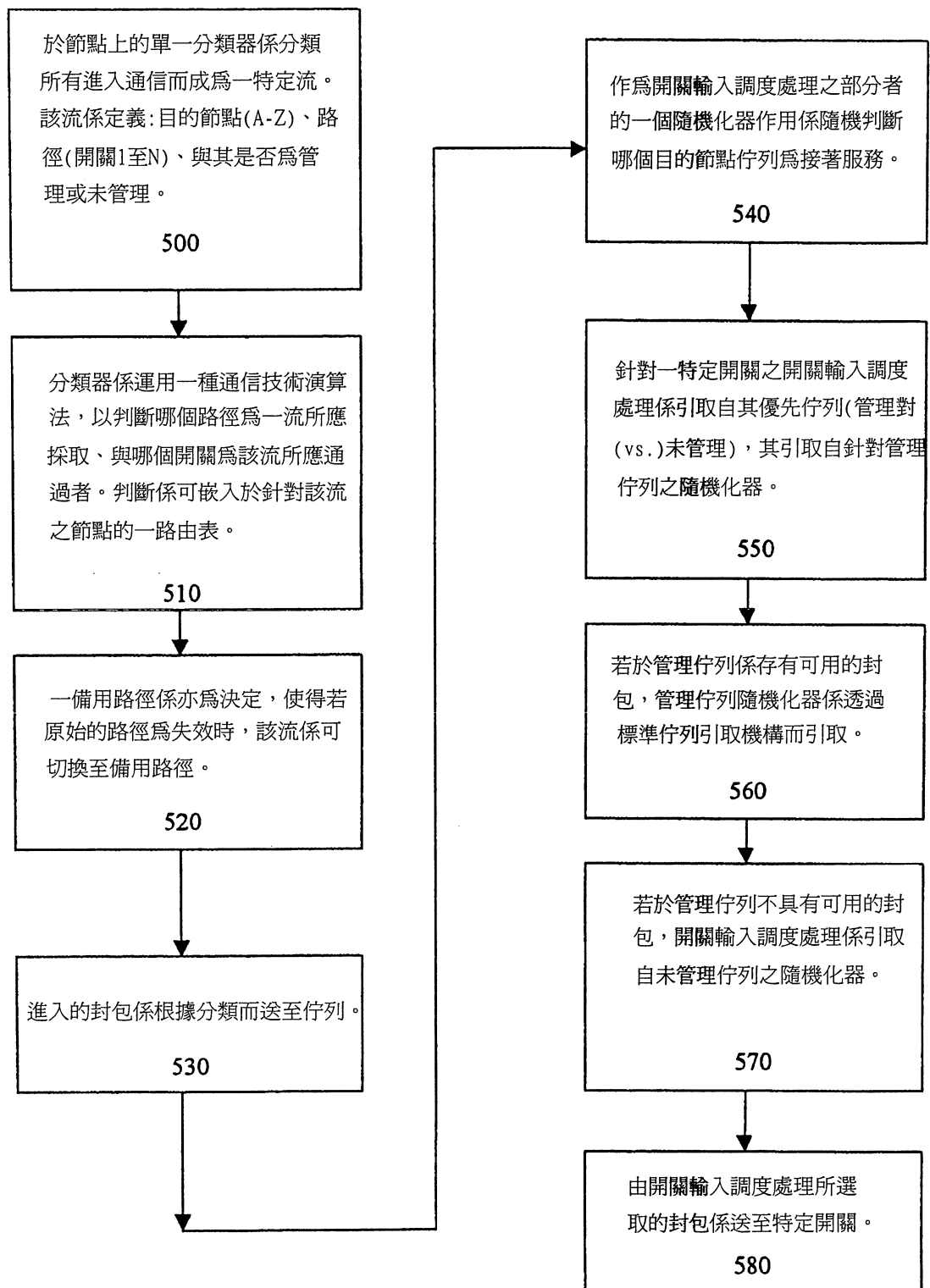
第二圖



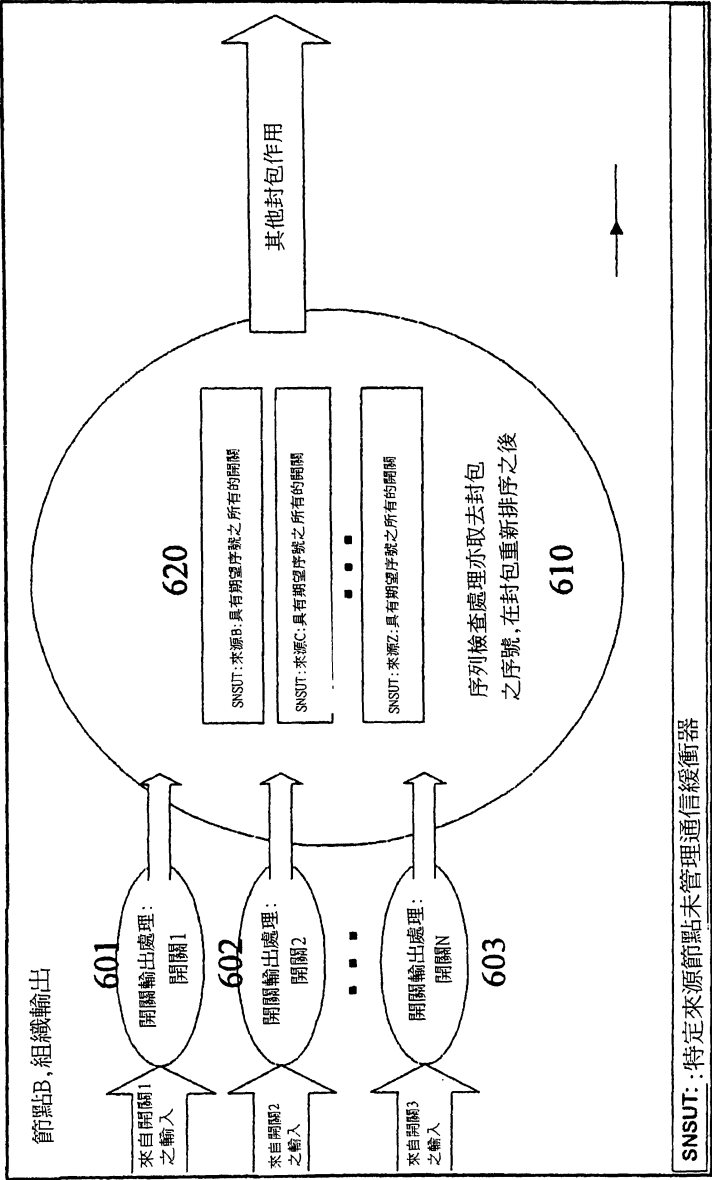
第三圖



第四圖



第五圖



第六圖

陸、(一)、本案指定代表圖爲：第 1 圖

(二)、本代表圖之元件代表符號簡單說明：

100-114 板

120-122 無限頻帶(IBA)開關裝置

柒、本案若有化學式時，請揭示最能顯示發明特徵的化學式：

無