(19)日本国特許庁(JP)　(12)特　許　公　報(B2)　(11)特許番号

特許第6392745号
(P6392745)

(45)発行日　平成30年9月19日(2018.9.19)　(24)登録日　平成30年8月31日(2018.8.31)

(51)Int.Cl.　　F I
*H04L 12/931 (2013.01)*　　H04L 12/931
*G06F 15/173 (2006.01)*　　G06F 15/173　685S

請求項の数 24　(全 23 頁)

(21)出願番号　特願2015-512707(P2015-512707)
(86)(22)出願日　平成25年5月10日(2013.5.10)
(65)公表番号　特表2015-522991(P2015-522991A)
(43)公表日　平成27年8月6日(2015.8.6)
(86)国際出願番号　PCT/US2013/040508
(87)国際公開番号　WO2013/173181
(87)国際公開日　平成25年11月21日(2013.11.21)
審査請求日　平成28年4月14日(2016.4.14)
(31)優先権主張番号　13/470,847
(32)優先日　平成24年5月14日(2012.5.14)
(33)優先権主張国　米国(US)

(73)特許権者　591016172
アドバンスト・マイクロ・ディバイシズ・インコーポレイテッド
ＡＤＶＡＮＣＥＤ　ＭＩＣＲＯ　ＤＥＶＩＣＥＳ　ＩＮＣＯＲＰＯＲＡＴＥＤ
アメリカ合衆国　９５０５４　カリフォルニア州、　サンタ　クララ、オーガスティン　ドライブ　２４８５
(74)代理人　100108833
弁理士　早川　裕司
(74)代理人　100111615
弁理士　佐野　良太
(74)代理人　100162156
弁理士　村雨　圭介

最終頁に続く

(54)【発明の名称】サーバノード相互接続デバイス及びサーバノード相互接続方法

(57)【特許請求の範囲】
【請求項１】
複数の電子デバイスと通信するスイッチファブリックと、
前記スイッチファブリックと通信するランデブーメモリと、
前記スイッチファブリックと１つ以上のネットワークインターフェースカード（ＮＩＣｓ）との間に接続された入出力（Ｉ／Ｏ）プロセッサと、を備え、
前記複数の電子デバイスのうち第１の電子デバイスのメモリからデータが出力可能な状態にあるとの判定に応じて、且つ、前記ランデブーメモリに前記データ用の位置が割り当てられたことに応じて、前記データが前記第１の電子デバイスから前記ランデブーメモリに転送され、
前記Ｉ／Ｏプロセッサは、前記ＮＩＣｓに接続されたネットワーク上のリモート電子デバイスへの前記ランデブーメモリからの前記ＮＩＣｓを介したデータの転送に関連する動作を実行する、
集約デバイス。
【請求項２】
前記第1の電子デバイスは、前記スイッチファブリックに接続したサーバノードである、請求項１に記載の集約デバイス。
【請求項３】
前記第１の電子デバイスは、前記集約デバイスと通信するリモートコンピュータである、請求項１に記載の集約デバイス。

【請求項４】
前記データは、前記ランデブーメモリから前記複数の電子デバイスのうち第２の電子デバイスに前記データが出力可能な状態にあるとの判定に応じて、且つ、前記第２の電子デバイスのメモリに前記データ用の位置が割り当てられたことに応じて、前記ランデブーメモリから前記第２の電子デバイスに転送される、請求項１に記載の集約デバイス。
【請求項５】
前記データが第２の電子デバイスに提供可能になるまでに、前記ランデブーメモリにおいて前記データ用の位置を割り当てるランデブー管理モジュールをさらに含む、請求項１に記載の集約デバイス。
【請求項６】
前記Ｉ／Ｏプロセッサは、出力可能な状態の前記データが受信される位置であって、前記ランデブーメモリに割り当てられた位置を識別する、請求項１に記載の集約デバイス。
【請求項７】
前記複数の電子デバイスの間でデータを送信するスイッチングコントロールプレーンを管理するコントロールプレーンプロセッサをさらに含む、請求項１に記載の集約デバイス。
【請求項８】
前記ランデブーメモリはＴＳＶメモリを含む、請求項１に記載の集約デバイス。
【請求項９】
前記ランデブーメモリはオンダイメモリを含む、請求項１に記載の集約デバイス。
【請求項１０】
複数のサーバノードと、
前記複数のサーバノードに接続した集約デバイスとを備え、
前記集約デバイスは、
前記複数のサーバノードと通信するスイッチファブリックと、
前記複数のサーバノードのうち第１のサーバノードのローカルメモリからデータが出力可能な状態にあるとの判定と、ランデブーメモリに前記データ用の位置が割り当てられたことに応じた、前記第１のサーバノードのローカルメモリを用いたデータの転送に関与するランデブーメモリと、
前記スイッチファブリックと１つ以上のネットワークインターフェースカード（ＮＩＣｓ）との間に接続された入出力（Ｉ／Ｏ）プロセッサであって、前記ＮＩＣｓに接続されたネットワーク上のリモート電子デバイスへの前記ランデブーメモリからの前記ＮＩＣｓを介したデータの転送に関連する動作を実行するＩ／Ｏプロセッサと、
を含む、サーバ集約システム。
【請求項１１】
前記データは、前記ランデブーメモリから宛先電子デバイスに前記データが出力可能な状態にあるとの判定と、前記データを受信するための位置が前記宛先電子デバイスに割り当てられたことに応じて、前記ランデブーメモリから前記宛先電子デバイスに転送される、請求項１０に記載のサーバ集約システム。
【請求項１２】
前記宛先電子デバイスは、前記複数のサーバノードのうち第２のサーバノード、又は、前記集約デバイスと通信するリモートコンピュータである、請求項１１に記載のサーバ集約システム。
【請求項１３】
前記データが宛先電子デバイスに提供可能になるまでに、前記ランデブーメモリにおいて前記データ用の位置を割り当てるランデブー管理モジュールをさらに含む、請求項１０に記載のサーバ集約システム。
【請求項１４】
前記Ｉ／Ｏプロセッサは、出力可能な状態の前記データが受信される位置であって、前記ランデブーメモリに割り当てられた位置を識別する、請求項１０に記載のサーバ集約シ

ステム。
【請求項１５】
前記第１のサーバノードの前記ローカルメモリと前記ランデブーメモリとの間で前記データの転送を確立するコントロールプレーンプロセッサをさらに含む、請求項１０に記載のサーバ集約システム。
【請求項１６】
複数の電子デバイスと集約システムとの間で通信する方法であって、
第２の電子デバイスに送信するデータを、前記複数の電子デバイスのうち第１の電子デバイスのローカルメモリに提供することと、
前記第１の電子デバイスの前記ローカルメモリから前記データを受信するための前記集約システムのランデブーメモリを提供することと、
前記第１の電子デバイスの前記ローカルメモリに前記データが伝送用として存在するとの前記集約システムによる判定と、前記第１の電子デバイスの前記ローカルメモリから受信される前記データを記憶するための位置が前記ランデブーメモリに割り当てられているとの判定とに応じて、前記データを、前記第１の電子デバイスの前記ローカルメモリから前記ランデブーメモリに転送することと、
前記データを前記ランデブーメモリからリモートコンピュータである前記第２の電子デバイスのローカルメモリに転送することであって、前記データを、前記集約システムのＩ／Ｏプロセッサと１つ以上のネットワークインターフェースカード（ＮＩＣｓ）を介して、前記ＮＩＣｓ及び前記第２の電子デバイスに接続されたネットワークを経由して前記第２の電子デバイスに転送することを含む、ことと、
を含む方法。
【請求項１７】
前記第１の電子デバイスの前記ローカルメモリにおいて前記データが伝送用に使用可能であるという前記第１の電子デバイスからの通知を生成することと、
前記ランデブーメモリにて使用可能なバッファを識別することと、
前記第１の電子デバイスの前記ローカルメモリから前記データを受信するために前記使用可能なバッファを割り当てることと、
前記データを、前記第１の電子デバイスの前記ローカルメモリから前記ランデブーメモリの前記使用可能なバッファに出力することと、
をさらに含む請求項１６に記載の方法。
【請求項１８】
前記第２の電子デバイスのローカルメモリが前記データを受信可能であるとの判定を行うことと、
前記判定に基づいて、前記データを、前記ランデブーメモリから前記第２の電子デバイスの前記ローカルメモリに転送することと、
をさらに含む請求項１６に記載の方法。
【請求項１９】
前記ランデブーメモリにおいて前記データが伝送用に使用可能であるという前記ランデブーメモリからの通知を生成することと、
前記第２の電子デバイスの前記ローカルメモリにて使用可能なバッファを識別することと、
前記第２の電子デバイスの前記ローカルメモリにて前記使用可能なバッファを割り当てることと、
前記割り当てに基づいて、前記データを、前記ランデブーメモリから前記第２のデバイスの前記ローカルメモリの前記使用可能なバッファに転送することと、
をさらに含む請求項１８に記載の方法。
【請求項２０】
前記第１の電子デバイスはサーバノードを含む、請求項１６に記載の方法。
【請求項２１】

前記第２の電子デバイスはリモートコンピュータである、請求項１６に記載の方法。
【請求項２２】
複数の電子デバイスと集約システムとの間で通信する方法であって、
第２の電子デバイスに送信するデータを、前記複数の電子デバイスのうち第１の電子デバイスのローカルメモリに提供することと、
前記集約システムの仮想ネットワークインターフェースカード（ｖＮＩＣ）に対して、前記第１の電子デバイスの前記ローカルメモリにおいて前記データが伝送用として使用可能であると通知することと、
前記第１の電子デバイスの前記ローカルメモリに伝送用として前記データがあるとの判定と、前記集約システムのランデブーメモリに前記データ用の位置が割り当てられたこととに応じて、前記データを、前記第１の電子デバイスの前記ローカルメモリから前記ランデブーメモリに転送することと、
前記第２の電子デバイスにおける少なくとも一つの使用可能なバッファの通知を生成することと、
前記データを受信するために前記第２の電子デバイスにおける少なくとも一つの使用可能なバッファを割り当てることと、
前記データを前記ランデブーメモリからリモートコンピュータである前記第２の電子デバイスにおける少なくとも一つの使用可能なバッファに転送することであって、前記データを、前記集約システムのＩ／Ｏプロセッサと１つ以上のネットワークインターフェースカード（ＮＩＣｓ）を介して、前記ＮＩＣｓ及び前記第２の電子デバイスに接続されたネットワークを経由して前記第２の電子デバイスに転送することを含む、ことと、
を含む方法。
【請求項２３】
前記電子デバイスの少なくとも一つはサーバノードを含む、請求項２２に記載の方法。
【請求項２４】
コンピュータ可読プログラムコードを記憶するコンピュータ可読記憶媒体であって、
前記コンピュータ可読プログラムコードは、
集約システムを介して第２の電子デバイスにデータを送信するために、第1の電子デバイスのローカルメモリに前記データを提供するように構成されたコンピュータ可読プログラムコードと、
前記集約システムの仮想ネットワークインターフェースカード（ｖＮＩＣ）に対して、前記第１の電子デバイスの前記ローカルメモリにおいて前記データが伝送用として使用可能であると通知するように構成されたコンピュータ可読プログラムコードと、
前記第１の電子デバイスの前記ローカルメモリに前記データが伝送用として存在するとの判定と、前記ランデブーメモリに前記データ用の位置が割り当てられたこととに応じて、前記データを、前記第１の電子デバイスの前記ローカルメモリから前記ランデブーメモリに転送するように構成されたコンピュータ可読プログラムコードと、
前記第２の電子デバイスにおける少なくとも一つの使用可能なバッファの通知を生成するように構成されたコンピュータ可読プログラムコードと、
前記データを受信するために前記第２の電子デバイスにおける前記少なくとも一つの使用可能なバッファを割り当てるように構成されたコンピュータ可読プログラムコードと、
前記ｖＮＩＣが、前記データを、前記ランデブーメモリからリモートコンピュータである前記第２の電子デバイスにおける前記少なくとも一つの使用可能なバッファに転送することであって、前記データを、前記集約システムのＩ／Ｏプロセッサと１つ以上のネットワークインターフェースカード（ＮＩＣｓ）を介して、前記ＮＩＣｓ及び前記第２の電子デバイスに接続されたネットワークを経由して前記第２の電子デバイスに転送することを含む、ことを行うように構成されたコンピュータ可読プログラムコードと、
を含むコンピュータ可読記憶媒体。
【発明の詳細な説明】
【技術分野】

【０００１】
本発明は、概して、ネットワークスイッチアーキテクチャに関し、より詳細には、スイッチファブリックを用いてサーバノードを相互接続するシステム及び方法に関する。
【背景技術】
【０００２】
データセンタは、概して、ビジネス及び組織をサポートするインターネット及びイントラネットサービスを提供する、一元管理型の設備である。典型的なデータセンタは、様々なタイプの電子機器（例えば、コンピュータ、サーバ（例えば、電子メールサーバ、プロキシサーバ及びＤＮＳサーバ）、ネットワークスイッチ、ルーター、データ記憶デバイス、及び、その他の関連する構成要素）を格納している。所与のデータセンタは、スイッチ及びルーターを含むスイッチングアーキテクチャを介して相互に及び外部デバイスと通信する数百又は数千の相互接続されたサーバノードを有することが可能である。また、従来のデータセンタは、仮想化のための構成をとることにより、サーバノードに対して、ネットワークインタフェイスカード（ＮＩＣ）、ハードディスクドライブ又はその他のハードウェアの共有を可能にすることもできる。このようにする際には、通信データセンタアーキテクチャを介したノード間通信を容易にするために、複雑なスイッチファブリックが必要である。
【発明の概要】
【課題を解決するための手段】
【０００３】
一態様によれば、スイッチファブリックとランデブーメモリとを含む集約デバイスが提供される。スイッチファブリックは、複数の電子デバイスと通信する。ランデブーメモリは、スイッチファブリックと通信する。データは、複数の電子デバイスのうち第１の電子デバイスのメモリから前記データが出力可能な状態にあるという判定と、前記データ用のランデブーメモリに割り当てられた位置とに応じて、前記第１の電子デバイスから前記ランデブーメモリに転送される。
【０００４】
別の態様によれば、複数のサーバノードと、前記複数のサーバノードに接続した集約デバイスとを含むサーバ集約システムが提供される。集約デバイスは、前記複数のサーバノードのうち第１のサーバノードのローカルメモリを用いたデータ転送に関与するランデブーメモリを含む。このデータ転送は、前記第１のサーバノードのメモリからデータが出力可能であるという判定と、前記ランデブーメモリに割り当てられた前記データ用の位置とに応じて行われる。
【０００５】
別の態様によれば、複数の電子デバイスと集約システムとの間で通信する方法が提供される。この方法は、複数の電子デバイスのうち第２の電子デバイスに送信するデータを、複数の電子デバイスのうち第１の電子デバイスのローカルメモリに提供することと、前記第１の電子デバイスのローカルメモリからデータを受信するために、ランデブーメモリを集約システムに設けることと、伝送用のデータが前記第１の電子デバイスのローカルメモリに存在するとの集約システムによる判定と、前記第１の電子デバイスのローカルメモリから受信したデータを記憶するための位置がランデブーメモリに割り当てられたとの判定とに応じて、第１の電子デバイスのローカルメモリからランデブーメモリにデータを転送することと、を含む。
【０００６】
別の態様によれば、複数の電子デバイスと集約システムとの間で通信する方法が提供される。この方法は、複数の電子デバイスのうち第２の電子デバイスに送信するデータを、複数の電子デバイスのうち第１の電子デバイスのローカルメモリに提供することと、第１の電子デバイスのローカルメモリにおいてデータが伝送用に使用可能であることを、集約システムの仮想ネットワークインターフェースカード（ｖＮＩＣ）に通知することと、第２の電子デバイスにおける少なくとも一つの使用可能なバッファの通知を生成することと

、データを受信するために前記少なくとも一つの使用可能なバッファを割り当てることと、データを、第２の電子デバイスにおける前記少なくとも一つの使用可能なバッファにｖＮＩＣによって出力することと、伝送用のデータが前記第１の電子デバイスのローカルメモリに存在するとの判定と、前記第２の電子デバイスにて前記少なくとも一つの使用可能なバッファを割り当てたとの判定とに応じて、第１の電子デバイスのローカルメモリからランデブーメモリにデータを転送することと、を含む。

【０００７】

別の態様によれば、コンピュータ可読プログラムコードを有するコンピュータ可読記憶媒体を含むコンピュータプログラム製品であって、前記記憶媒体を用いてプログラムコードが具体化されるコンピュータプログラム製品が提供される。コンピュータ可読プログラムコードは、集約システムを介して第２の電子デバイスに送信するためのデータを、第１の電子デバイスのローカルメモリに提供するように構成されたコンピュータ可読プログラムコードと、第１の電子デバイスのローカルメモリにおいてデータが伝送用に使用可能であることを、集約システムの仮想ネットワークインターフェースカード（ｖＮＩＣ）に通知するように構成されたコンピュータ可読プログラムコードと、第２の電子デバイスにおける少なくとも一つの使用可能なバッファの通知を生成するように構成されたコンピュータ可読プログラムコードと、データを受信するために前記少なくとも一つの使用可能なバッファを割り当てるように構成されたコンピュータ可読プログラムコードと、第２の電子デバイスにおける少なくとも一つの使用可能なバッファに対して、ｖＮＩＣによってデータを出力するように構成されたコンピュータ可読プログラムコードと、伝送用のデータが前記第１の電子デバイスのローカルメモリに存在するとの判定と、前記第２の電子デバイスにて前記少なくとも一つの使用可能なバッファを割り当てたとの判定とに応じて、第１の電子デバイスのローカルメモリからランデブーメモリにデータを転送するように構成されたコンピュータ可読プログラムコードと、を含む。

【０００８】

本発明の上記の利点及び更なる利点は、添付の図面に関連した後述の記載を参照することによって、より良く理解し得る。これらの図面において同様に表された数字は、同様の構成要素及び特徴を示している。図面は、必ずしも縮尺が揃っていないが、本発明の原理を例示することに重点が置かれている。

【図面の簡単な説明】

【０００９】

【図１】従来のデータセンタネットワーク階層のブロック図である。

【図２】従来のインターコネクトファブリックスイッチに結合した複数のサーバノードを含むデータセンタのブロック図である。

【図３】一実施形態における、ランデブーメモリを有するサーバ集約システムと通信している複数のサーバノードを含む計算インフラストラクチャーのブロック図である。

【図４】一実施形態における、図３のランデブー管理装置のブロック図である。

【図５】一実施形態における、図３のネットワークプロセッサのブロック図である。

【図６】一実施形態における、サーバノードと、ランデブーメモリを有するサーバ集約システムとの間で電子通信を行う方法のフロー図である。

【図７】一実施形態における、ソースサーバノードと、ランデブーメモリを有するサーバ集約システムとの間で電子通信する方法のフロー図である。

【図８】一実施形態における、サーバ集約システムと、宛先サーバノードとの間で通信を確立する方法のフロー図である。

【図９】他の実施形態における、ランデブースイッチング用に構成されたサーバ集約システムに結合した第１及び第２のサーバノード間で通信を確立する方法のフロー図である。

【発明を実施するための形態】

【００１０】

以下の記載では具体的な詳細が記載されるが、本発明にかかるシステム及び方法が少なくともいくつかの詳細無しで実行できることは、当業者にとって認識し得るはずである。

場合によっては、本発明が曖昧にならないように、既知の特徴又は工程は、詳細に記載されていない。

【００１１】

図１は、従来のデータセンタネットワーク階層１０のブロック図である。図１では、複数のサーバノード１２－１～１２－Ｎ（Ｎは、１より大きい整数）は、例えば、ラック１８－１のイーサネット（登録商標）ローカルエリアネットワーク（ＬＡＮ）又は関連するデータネットワークなどのＬＡＮを介して、ラックスイッチ１４と通信することができる。ラック１８－１は、データセンタネットワーク階層１０における他の一つ以上のラック１８－２，１８－Ｎ（Ｎは、１より大きい整数）とともに、クラスタの一部として構成されてもよい。各クラスタは、集約スイッチ２２－１，２２－２，２２－Ｎ（Ｎは、１より大きい整数）を含むことができ、これらは、イーサネット（登録商標）又は他のネットワーク接続を介して、コアルータ２４に接続されている。ユーザコンピュータ３２（例えば、ラップトップ、スマートフォン又はその他の電子デバイス）は、ネットワーク２６（例えば、ＩＰネットワーク、インターネットなど）を介して、サーバノード１２－１～１２－Ｎ（概して、１２）と通信することができる。データセンタスイッチに関連する規模、経費及び電力の削減を求める一方で、データセンタの拡張性の増大を求めるという、止むことのない要求が存在する。

【００１２】

図２は、従来のインターコネクトファブリックスイッチ１２０に結合した複数のサーバノード１１２－１～１１２－Ｎ（Ｎは、１より大きい整数）を含むデータセンタ２０のブロック図である。インターコネクトファブリックスイッチ１２０は、単一のシャーシのもとでの集約スイッチ及びラックスイッチの一元管理機能によって、スイッチの数を削減することができ、従来のサーバラックにおいて使用される従来のイーサネット（登録商標）スイッチングの要求の多くを置き換えることができる。

【００１３】

サーバノード１１２－１～１１２－Ｎ（概して、１１２）は、プロセッサクラスタとして、又は、他の周知の装置として構成することができる。サーバノード１１２は、単一のソケットサーバ、又は、共用インフラストラクチャーを共有する低消費電力プロセッサ１０２を含むことができる。サーバノードプロセッサ１０２は、一つ以上のマイクロプロセッサ、中央演算処理装置（ＣＰＵ）、画像処理装置（ＧＰＵ）、デジタルシグナルプロセッサ（ＤＳＰ）、特定用途向け集積回路（ＡＳＩＣ）、メモリコントローラ、マルチコアプロセッサ、及び、その他のタイプのデータ処理デバイスだけでなく、これらのデバイスや他のデバイスの一部及び／又は組合せを含むことができる。

【００１４】

また、サーバノード１１２は、ローカルメモリ１０４と、Ｉ／Ｏロジック１０６とを含む。ローカルメモリ１０４は、不揮発性又は揮発性メモリ（例えば、一つ以上のチャネルのダイナミックランダムアクセスメモリ（ＤＲＡＭ）又はスタティックＲＡＭ（ＳＲＡＭ））を含むことができる。Ｉ／Ｏロジック１０６は、コンピュータのＩ／Ｏ機能（例えば、サーバノード１１２とインターコネクトファブリックスイッチ１２０との間のデータ転送）を管理するために、サウスブリッジ等を含むＩ／Ｏコントローラを有するように構成されてもよい。Ｉ／Ｏロジック１０６は、インターコネクトファブリックスイッチ１２０と電子通信を確立するために、イーサネット（登録商標）、ＰＣＩｅ又はその他のネットワークコネクター１１４を含んでいてもよい。

【００１５】

インターコネクトファブリックスイッチ１２０は、複数の入出力ポートと、可変長又は固定長のフレーム、データパケット、セル等を入出力ポート間でルーティングし得るクロスバー１２４とを含み、サーバノード１１２、例えばＮＩＣ１４２又はハードドライブ１４４等の共有デバイス、及び／又は、例えばユーザコンピュータ１５２等の外部電子デバイス間での通信を容易にすることができるものである。ポートは、仮想ポート、物理ポート又はこれらの組合せを含んでいてもよい。ポートは、単方向又は双方向通信用の構成で

あってもよい。

【００１６】

クロスバー１２４は、行と列の配列に組織化され得る。データ伝送の間には、所与の行のいくつかの入力ポートは、ある列における出力ポートを奪い合う場合がある。入力ポートは、入力バッファ１２６と通信することができる。入力バッファは、クロスバー１２４が、使用可能な出力ポートにデータを送信することができるまで、サーバノード１１２から受信したデータを一時的に記憶するものである。出力ポートは、出力バッファ１２８を含む。出力バッファは、ネットワークバス１４０（例えば、イーサネット（登録商標）バス、ＰＣＩｅバス等）を介してパケットを宛先に送信するために、所望の出力ポートが使用可能になるまで、一つ以上の入力ポートから受信したデータを一時的に記憶する。

【００１７】

しかしながら、インターコネクトファブリックスイッチ１２０は、輻輳に関連する問題（例えば、ヘッドライン（ＨＯＬ）ブロッキング）が生じやすい。この問題は、スイッチ１２０の複数の入力ポートが同一の出力ポートを奪い合う場合に発生するものである。また、スケーリングに関連する課題も発生するが、これは、スイッチ１２０でのＩ／Ｏポートの追加によって、共有のリソースを奪い合うというリスクが増加するからである。関連する課題は、サーバノード１１２と、インターコネクトファブリックスイッチ１２０との間でバンド幅を充分に使用しないことである。例えば、ＨＯＬブロッキングは、特に、入力ポートで待機しているパケット、セル等が、ファーストイン・ファーストアウト（ＦＩＦＯ）のキューで記憶されている場合に、クロスバー１２４でかなりの量のバンド幅を消費する可能性がある。また、入力及び出力バッファ１２６，１２８でのキューはすぐに満杯になり、その結果、望ましくない待ち時間、ジッター又はパケット喪失のみならず、オーバーラン及び／又はアンダーラン条件が生じ、これらは、バンド幅の不充分な使用及び性能の課題を生じる。

【００１８】

フロー制御技術は、容易に使用可能であり、データセンタ２０でのネットワーク輻輳を緩和することができる。例えば、複雑なスケジューリング技術を適用して、クロスバー１２４を介したデータトラヒックを管理することができる。しかしながら、そのような技術は、概して、実現するには高価である。

【００１９】

他の周知である技術は、イーサネット（登録商標）をベースにした再送信を実現することである。しかしながら、イーサネット（登録商標）をベースにした再送信は、貴重なバンド幅を消費する。

【００２０】

他のフロー制御メカニズムを上流のデバイスに適用して、パケットの通過を停止させるように当該デバイスに要求することができる。そのような手法は、バッファオーバーフローを低減させるのに効果的であるものの、ネットワーク輻輳を完全に緩和しない。

【００２１】

他の手法は、一時記憶用の外部メモリデバイスと、パケットのキューイング（ｑｕｅｕｉｎｇ）とに依るものである。しかしながら、この手法は、インターコネクトファブリックスイッチ１２０に対して、さらなるピン、及び、すぐにでも使用可能なメモリデバイスへの「配線」又はコネクターを必要とし、その結果として、拡張性の制限と、バンド幅に関連した問題とが、さらなる電力消費を求める関連要求とともに生じる。

【００２２】

本発明の概念は、従来のデータセンタ相互接続スイッチに付随する輻輳に関連した課題を、ランデブーデバイス（ｒｅｎｄｅｚｖｏｕｓ　ｄｅｖｉｃｅ）を含むサーバ集約システムを導入することによって低減又は取り除くものである。ランデブーデバイスは、制御された方法で、高バンド幅スイッチファブリックを介して、複数のサーバノード又は関連する電子デバイスのローカルメモリと通信する。ランデブーデバイスは、出力可能なデータを有する伝送デバイスと、当該データの受信用に使用可能なメモリバッファを有する受

信デバイスとの間で通信し、伝送デバイス及び受信デバイスの間で効率的なデータ伝送経路を提供する。例えば、データが、伝送デバイスのローカルメモリから明示的に転送され得るのは、データが出力可能であるとサーバ集約システムが判定した後であって、データ記憶用のランデブーデバイスのメモリ記憶位置に特定のメモリ記憶位置が割り当てられた後である。これらの条件が真であると判定された場合、すなわち、送信するデータを伝送デバイスが有していて、使用可能なバッファ（空きバッファ）を受信デバイスが有している場合に、伝送デバイスから受信デバイスへの効率的で待ち時間の短いデータ移動が生じる。

【００２３】

図２に記載された従来のネットワークスイッチ環境では、データは、クロスバースイッチキュー１２６，１２８が使用可能な空間を有しているか否かに関係なく、キュー１２６，１２８に出力される。キュー１２６，１２８が使用可能な空間を有していない場合には、データは通常破棄され、複雑でバンド幅を消費する再送信過程が実行される。

【００２４】

一実施形態では、サーバ集約システムは、ソースサーバノード等からのデータをいつ受信するのかを判定して、「ランデブー（ｒｅｎｄｅｚｖｏｕｓ）」する位置を確立することができる。この位置は、宛先サーバノード等のメモリ記憶位置と通信して、宛先サーバノードにデータをいつ送信するのかを提供することができる。事前に割り当てられた位置は、サーバ集約システムのあらゆる入力ポートからデータを着信するために、ランデブーメモリ及び／又は宛先サーバノードメモリに提供される。これを行う場合、ソースサーバノードメモリからランデブーメモリに、又は、ランデブーメモリから宛先サーバノードメモリにデータを移動させるために、解放されているデータバッファが割り当てられる。ソースサーバノードから送信されることが意図されるデータ用のランデブーメモリにおいて使用可能な位置が存在しない場合には、ランデブーメモリに記憶位置の空きができるまで、データはランデブーメモリに転送されない。また、宛先サーバノードが、データを受信する宛先サーバノードのローカルメモリに使用可能な記憶位置があることをサーバ集約システムに通知するまで、受信データをランデブーメモリに記憶することができる。

【００２５】

他の実施形態では、リモートコンピュータは、外部ネットワーク上のサーバ集約システムへのデータを、外部ネットワークとサーバ集約システムとの間に結合されたＮＩＣ又は関連インターフェースに提供する。ここでは、ＮＩＣは、（例えば、ＮＩＣの受信リングでの記述子により識別される）ランデブーメモリの受信バッファにデータを転送する。使用可能なバッファが不十分である場合には、データは、脱落又はフィルタリングされる。そうでなければ、受信バッファは使用可能であり、データはランデブーメモリに提供され、ランデブーメモリと通信している処理複合体に通知される。次いで、メモリバッファは検査され、メモリバッファをどこにルーティングするべきか、そしてスイッチングファブリックへの処理複合体接続における記述子上のどこに置くべきかが判定される。この時点で、データは、あたかもそれがイングレスｖＮＩＣにあるのと同様な方法で移動される。ここでは、ランデブーメモリ管理装置は、宛先サーバノードによって、例えばｖＮＩＣを介して、宛先サーバノードメモリに空き記憶位置があるかどうか通知される。ランデブーメモリ管理装置は、受信バッファが宛先サーバノードメモリに割り当てられていない限り、宛先サーバノードメモリへの伝送用のデータをフェッチすることはしない。

【００２６】

従って、本発明の概念の特徴は、データ転送におけるランデブーメモリの入力及び／又は出力領域、すなわち、ソースサーバノードと通信するランデブーメモリの一端と、宛先サーバノードと通信するランデブーメモリのもう一端とで、フロー制御を生じさせることができるということである。また、スイッチファブリックにより実行されるバッファリング（ｂｕｆｆｅｒｉｎｇ）は、ほとんど又は全く存在しないことから、スイッチファブリックのイングレスポートからエグレスポートへの処理待ち時間が低い。例えば、データトランザクションをスケジューリングすることにより、投機的データ移動が原因となる輻輳

又はその他のトラヒックに起因するブロッキングのリスクが低減するから、待ち時間を改善することができる。

【００２７】

本発明の概念の他の特徴は、バンド幅を、サーバ集約システム内及びサーバノード間で効率的に分配可能なことである。これは、出力することになるデータをソースサーバノードのローカルメモリが有しているかどうか、ランデブーメモリがそのデータを受信できるかどうか、及び／又は、データのメモリ間交換を行っているサーバノードでのローカルメモリが、データを受信するのに充分な空間を有しているかどうかを、サーバ集約システムが実際のデータ転送に先立って判定することができるからである。

【００２８】

図３は、本実施形態における、ランデブーメモリ（Ｒ－Ｖメモリ）３０８を有するサーバ集約システム３００と通信している複数のサーバノード３１２－１～３１２－Ｎを含む計算インフラストラクチャー３０のブロック図である。計算インフラストラクチャー３０は、大規模なデータセンタ、クラウド等を含むことができる。計算インフラストラクチャーは、図２を参照して記載されるデータセンタ２０と同様の仮想化した構成をとることができる。従って、仮想化関連の構成要素に関する詳細は、簡潔にするために図３では省略される。

【００２９】

サーバ集約システム３００は、スイッチファブリック３０２と、ランデブー管理モジュール３０４と、Ｉ／Ｏプロセッサ３０６と、ランデブーメモリ３０８と、コントロールプレーンプロセッサ３１０とを含む。サーバ集約システム３００は、仮想化した構成とすることができる。

【００３０】

サーバノード３１２－１～３１２－Ｎ（概して、３１２）、又は、マイクロサーバ及び／若しくは少なくとも一つのプロセッサを有するその他の電子デバイスは、コネクター３１４、好ましくはＰＣＩｅバス、又は、その他のネットワークコネクターを介して、サーバ集約システム３００と通信することができる。各コネクター３１４は、一つ以上のサーバノード３１２とサーバ集約システム３００のと間のデータ経路を提供することができる。サーバ集約システム３００及びサーバノード３１２は、同一のマルチプロセッシングユニット（例えばチップ、計算デバイス又はラック）に一緒に設置されてもよい。他の実施形態では、サーバノード３１２は、一つ以上のユニット上に形成され、サーバ集約システム３００は、独立ユニット上（例えば、チップ上）に形成される。

【００３１】

スイッチファブリック３０２は、複数の入力ポート及び出力ポートを含んでおり、データ又はその他の電子情報を、その入力ポートと通信するサーバノード３１２と、そのエグレスポートと通信するランデブーメモリ３０８との間で移動させる構成とすることができる。スイッチファブリック３０２は、イングレスポートとエグレスポートとの間でデータを移動させるスイッチング構成（例えば、クロスバー）を含むことができ、このことは当業者には周知である。従来のスイッチファブリックとは異なり、スイッチファブリック３０２では、従来のキューイング技術が必要ないことから、バッファリングがほとんど又は全く必要ない。スイッチファブリック３０２及びランデブーメモリ３０８は、相互接続することができ、それらの間でのＩ／Ｏバンド幅通信チャネルを有しており、従って、ランデブーメモリ３０８にデータ用として割り当てられた充分な空間がある限り、いくつものサーバノード３１２からのデータも受信することができる。

【００３２】

ランデブー管理モジュール３０４は、コントロールプレーンプロセッサ３１０及びＩ／Ｏプロセッサ３０６とともに、サーバノード３１２におけるメモリ１０４－１～１０４－Ｎ（概して、１０４）の一つ以上とランデブーメモリ３０８との間のデータパケットの明示的な転送を監視する。ランデブー管理モジュール３０４は、ランデブーメモリ３０８にバッファ空間を割り当てて、宛先サーバノード又は外部計算デバイスが、ソースサーバノ

ードから送信されたデータをランデブーメモリ３０８から読み出すことができるまで、それを「駐車（ｐａｒｋｉｎｇ）」させておくことができる。ランデブー管理モジュール３０４は、記述子等を使用して、ランデブーメモリ３０８から宛先メモリにデータが伝送される方法を制御することができる。ランデブー管理モジュール３０４は、ランデブーメモリ３０８の受信バッファの空きを監視し、受信バッファがポスト（ｐｏｓｔ）されるまで待機する。換言すれば、ランデブー管理モジュール３０４は、転送操作における一方の当事者（すなわち、ソースサーバノード又はランデブーメモリ）が伝送用データを有すること、転送操作における他方の当事者（すなわち、宛先サーバ又はランデブーメモリ）がデータを受信するのに充分な空間を有していることを保証することができる。

【００３３】

ランデブーメモリ３０８は、複数のネットワークスイッチのリンクと通信する場合には、バンド幅が充分に高い（例えば、１０Ｇｂ以上）貫通ビアシリコン（ＴＳＶ）、ＳＲＡＭ又はオンダイ（ｏｎ－ｄｉｅ）メモリを含むことができる。ランデブーメモリ３０８は、例えば、メモリのブロックに組織化された解放された複数のデータバッファを含むように構成され得るものであって、これらのデータバッファは、データをローカルメモリ１０４からランデブーメモリ３０８に移動させる場合に、一つ以上のｖＮＩＣ３３４に割り当てられ得る。

【００３４】

Ｉ／Ｏプロセッサ３０６は、ランデブーメモリ３０８と、一つ以上の電子デバイス（例えば、サーバノード３１２及び／又はリモート計算デバイス３５２）との間で転送されたデータを処理する。Ｉ／Ｏプロセッサ３０６は、マルチプレクサ、並びに、ランデブーメモリ３０８へのデータ転送及びランデブーメモリ３０８からのデータ転送を実行するその他のロジックを含むことができるが、転送の実行は、ランデブー管理モジュール３０４、コントロールプレーンプロセッサ３１０又はそれらの組合せを用いて形成された通信に従ってなされる。このように、Ｉ／Ｏプロセッサ３０６は、サーバ集約システム３００へ及びサーバ集約システム３００から（例えば、二つ以上のサーバノード３１２の間、又はサーバノード３１２とリモート計算デバイス３５２との間）データを移動させる集結地（ｓｔａｇｉｎｇ　ａｒｅａ）としての役割を果たすことができる。

【００３５】

Ｉ／Ｏプロセッサ３０６は、使用可能なバッファを示す受信記述子をポストすることによって、ソースサーバノード３１２からの伝送可能状態にあると判定された特定のデータが一時的に（例えば、宛先サーバノード３１２がデータを読み出すまで）記憶されることになるランデブーメモリ３０８におけるバッファを識別することができる。Ｉ／Ｏプロセッサ３０６は、ランデブーメモリ３０８からバッファ記述子を処理するｖＮＩＣ３３４へのデータ移動のために、例えばランデブーメモリ３０８の一つ以上のチャネル用のバッファ記述子を保持することができる。Ｉ／Ｏプロセッサ３０６は、受信バッファのリソースを受信記述子に追加する。

【００３６】

コントロールプレーンプロセッサ３１０は、ネットワークルーティングプロトコルを処理することによって、サーバノード３１２とサーバ集約システム３００との間のデータ送信用のスイッチングコントロールプレーン（図示省略）を管理し、サーバ集約システム３００等で受信されるデータパケット、フレーム、セル等の転送に関与する。コントロールプレーンプロセッサ３１０の他の機能は、データレディ（ｄａｔａ　ｒｅａｄｙ）通知、受信バッファの空き通知、バッファ解放等の生成を含む。コントロールプレーンプロセッサ３１０は、サーバ集約システム３００内のブロック、ポート及びノードの間の通信を提供することができ、ポート間のデータ移動のためにスイッチファブリック３０２と通信する。コントロールプレーンプロセッサ３１０は、ランデブーメモリ３０８及び／又は一つ以上のサーバノードメモリ１０４への書き込み及び／又は読み出しを行うことができる。コントロールプレーンは、使用可能な受信バッファの数を問い合わせるためのデータ可用メッセージ（例えば、送信可能メッセージ）及びバッファ解放メッセージを送信するよう

に構成され得る。

【００３７】

コントロールプレーンプロセッサ３１０は、ゼロ又は一つ以上のｖＮＩＣ３３４を、好ましくはダイレクトメモリアクセス（ＤＭＡ）エージェント（ａｇｅｎｔ）（図示省略）、又は、スイッチファブリック３０２における関連するエージェントと組み合わせて実現及び管理することができる。様々なタイプの通信が、コントロールプレーンプロセッサ３１０の管理下で、コントロールプレーン上で行われてもよい。例えば、このような通信は、サーバノードインターフェースドライバーから書き込まれ、バッファ記述子メモリ記憶位置の状態の変化を指示する伝送／受信記述子のドアベル（ｄｏｏｒ　ｂｅｌｌ）通知であってもよい。他の例では、コントロールプレーンは、伝送を終えたバッファを、さらなる伝送のために又は受信バッファとして再利用させるなどのバッファ解放メッセージを管理することができる。コントロールプレーンの他の機能は、データレディ通知の提供、バッファ空き通知の受信、及び、バッファ解放等を含むことができる。

【００３８】

図４は、一実施形態における、図３のランデブー管理モジュール３０４のブロック図である。ランデブー管理モジュール３０４は、メモリ割り当てモジュール４０２、通知モジュール４０４、データ配信モジュール４０６、割り込み生成モジュール４０８及び／又はタイマ４１０を含むことができる。ランデブー管理モジュール３０４は、同一デバイス（例えば、チップ、ラック等）の一部として示されている。代替として、ランデブー管理モジュール３０４のいくつかの構成要素は、図３に示された計算インフラストラクチャー３０内の別の位置に物理的に配置することができる。

【００３９】

メモリ割り当てモジュール４０２は、ソースサーバノード３１２－１からデータをフェッチするときにｖＮＩＣ３３４が使用するバッファ空間を、ランデブーメモリ３０８内に割り当てる。より詳細には、メモリ割り当てモジュール４０２は、ランデブーメモリ３０８のブロックを示すインデックスを提供することができる。このインデックスは、ランデブーメモリの空きデータバッファブロックを示すバッファ記述子を含む。

【００４０】

通知モジュール４０４は、ランデブーメモリ３０８にデータが入力された場合に、バッファ割り当てに関連した通知を生成する。例えば、通知モジュール４０４は、バッファを要求するｖＮＩＣ３３４に応答して、受信バッファ割り当てをｖＮＩＣ３３４に通知することができる。通知モジュール４０４は、サーバノード３１２からの伝送用にデータが使用可能であることを示す通知を、例えばｖＮＩＣ３３４から受信することができる。通知モジュール４０４は、ランデブーメモリ３０８から宛先デバイス用のｖＮＩＣ３３４への記述子及び／又はデータをフェッチした後に新規データが使用可能であることを、ｖＮＩＣ３３４に通知することができる。

【００４１】

データ配信モジュール４０６は、割り当て用のランデブーメモリ３０８からｖＮＩＣ３３４への記述子、データ等をフェッチすることができる。このｖＮＩＣ３３４は、次いで、データを宛先ローカルメモリに提供する。

【００４２】

割り込み生成モジュール４０８は、新規データがランデブーメモリ３０８にある場合に、割り込み信号をＩ／Ｏプロセッサ３０６に出力する。これにより、Ｉ／Ｏプロセッサ３０６は、データをランデブーメモリからプル（ｐｕｌｌ）し、当該データを提供する。また、割り込み生成モジュール４０８は、伝送記述子の解放後（例えば、ランデブーメモリ３０８から宛先ローカルメモリ１０４－Ｎにデータを移動させた後）に、割り込み信号をネットワークプロセッサに出力することができる。

【００４３】

タイマ４１０は、受信バッファが、ランデブーメモリ３０８にてｖＮＩＣ３３４への割り当て用に使用可能でない（空きがない）場合に起動し得る。タイマ４１０は、データを

ランデブーメモリ３０８に転送することが不可能であることをサーバ集約システム３００の構成要素に示し、これにより、フロー制御のレベルを提供する。

【００４４】

図５は、一実施形態における、図３のＩ／Ｏプロセッサ３０６のブロック図である。Ｉ／Ｏプロセッサ３０６は、プロセッサ５０２、記述子処理モジュール５０４及び／又はスケジューラ５０８を含む。Ｉ／Ｏプロセッサは、メモリ（図示省略）を含んでもよい。Ｉ／Ｏプロセッサ３０６は、同一のデバイス（例えばチップ、ラック等）の一部として示されているが、Ｉ／Ｏプロセッサ３０６の構成要素のいくつか又はすべては、図３に示された計算インフラストラクチャー３０の他の構成要素にあってもよい。

【００４５】

プロセッサ５０２は、データ出力に関連する構成要素（例えば、ＮＩＣ３４２）とのインターフェースをとるのに使用されるプログラムのプログラムコードを実行することが可能である。プロセッサ５０２は、ルーティングの決定を行い、受信バッファを、受信記述子リングから宛先伝送記述子リングに移動させることができる。

【００４６】

記述子処理モジュール５０４は、ソースサーバノード３１２－１からデータを受信するランデブーメモリ３０８の空き記憶位置を示す受信記述子を、生成及び／又はポストすることができる。受信記述子は、ランデブーメモリ３０８において、バッファが常駐する位置、バッファのサイズ、複数のセグメント等を示すことができる。Ｉ／Ｏプロセッサ３０６は、現在解放されている（データを受信するのに空いている）データバッファの記述子を、ランデブーメモリ３０８からＩ／Ｏプロセッサ３０６へ受け渡すことによって、ランデブーメモリ３０８内の使用可能なデータバッファに関する情報を受信することができる。また、記述子処理モジュール５０４は、ランデブーメモリ３０８から宛先サーバノード３１２－Ｎのローカルメモリ１０４－Ｎにデータを転送する命令、分散収集リスト等を含む伝送記述子を生成することができる。

【００４７】

スケジューラ５０８は、例えば、ＮＩＣ３４２若しくはＰＣＩｅコネクター３１４を介してサーバ集約システム３００が新規データを受信するとの通知、又は、宛先デバイスへの伝送用としてランデブーメモリ３０８内で新規データが使用可能であるとの通知を受信した場合に、起動する。スケジューラ５０８は、ランデブーメモリと、一つ以上のサーバノード３１２及び／又はリモート計算デバイス３５２との間のデータの転送を、例えばラウンドロビンの順、ファーストインファーストアウトの順又は本発明の技術分野で周知の他の順で調整することができる。スケジューラは、所定の方針に従ってデータを転送するように調整することができる。例えば、方針は、フロー又はポートを優先度付きでタグ付けするメカニズムを含むことができる。Ｉ／Ｏプロセッサ３０６は、あるレベルのパケット検査及び分類を実行してフローを差別化することができる。厳格な優先度スケジューラ５０８を実現することができる。代替方法として、スケジューラ５０８は、最小バンド幅割り当てを有する異なるフローを提供することができる。スケジューラ５０８は、サーバ集約システム３００のｖＮＩＣを管理することにより、ある宛先用のデータを選択することができる。

【００４８】

図６は、一実施形態における、サーバノードと、ランデブーメモリを有するサーバ集約システムとの間で電子通信を行う方法６００のフロー図である。方法６００は、図３の一つ以上のサーバノード３１２のメモリ１０４及び／又はサーバ集約システム３００に記憶された命令に従うことができる。従って、図２～５が参照される。方法６００のいくつか又はすべてを、ＡＳＩＣ、システムオンチップ（ＳＯＣ）又は関連するデバイスにて、オンダイで実行することができる。

【００４９】

ブロック６０２では、ソースサーバノード３１２－１は、セル、パケット、フレーム又は他のローカルメモリ１０４－１内のデータのユニットが、宛先（例えば、他のサーバノ

ード３１２－Ｎ又はリモート計算デバイス３５２）へのＮＩＣ３４２を介した出力に使用可能であると告知する。ソースサーバノード３１２－１は、伝送用のデータがメモリ１０４－１内にあって伝送用に使用可能であることを、ｖＮＩＣ３３４に通知することができる。サーバノード３１２は、データを処理するためにメモリ１０４－１内で定義されているリングバッファ等を維持して、データ伝送又は受信処理を実行することができる。

【００５０】

ひし型の判断部６０４では、ランデブーメモリ３０８が、ソースサーバノード３１２－１のメモリ１０４－１内の伝送用データを受信することができるかどうかの判定が行われる。この判定は、ランデブー管理モジュール３０４によって行うことができる。ランデブー管理モジュール３０４は、ソースサーバノード３１２－１から直接的にデータを伝送することができるバッファ空間が、ランデブーメモリ３０８内に空いている（使用可能である）かどうかを確立することができる。ランデブーメモリ３０８がデータを受信することができない場合には、次にブロック６０８において、ランデブー管理モジュール３０４は、受信バッファがポストされるまで待機し得る。一方、ランデブーメモリ３０８がデータを受信することができる場合には、ブロック６０６にて、データ、記述子等を、ソースサーバノード３１２－１のメモリ１０４－１からランデブーメモリ３０８に転送することができる。

【００５１】

ひし形の判断部６１０では、Ｉ／Ｏプロセッサ３０６は、ソースサーバノード３１２－１からランデブーメモリ３０８に転送されるデータが、宛先サーバノード（例えば、サーバノード３１２－Ｎ）に転送されることになるかどうかを判定する。この判定は、宛先アドレスに基づいて行われてもよいし、伝送記述子又は宛先ポートを識別するパケットに関するメタ情報をプリペンド（ｐｒｅｐｅｎｄ）することによって行われてもよい。データを宛先サーバノードに転送しないという判定結果であれば、次にブロック６１２において、Ｉ／Ｏプロセッサ３０６がデータを消費することができ、例えば、データは／プロセッサ３０６において終了する。代替方法として、Ｉ／Ｏプロセッサ３０６は、データを、データ内容に応じて（例えば、宛先アドレスに基づいて）ＮＩＣ３４２に転送する。例えば、データパケットは、スイッチファブリック３０２を介してフロー管理に関連付けられることが可能であり、Ｉ／Ｏプロセッサ３０６の構成要素、データのルーティングを制御する処理複合体と相互作用する。そうでなければ、ひし型の判断部６１４では、宛先サーバノード３１２－Ｎが、ランデブーメモリ３０８から伝送可能な状態のデータを受信するローカルメモリ１０４－Ｎに空き空間を有しているかどうかという判定がなされる。この判定は、コントロールプレーンプロセッサ３１０が行うことができ、この場合には、仮想ＮＩＣ３３４が、宛先ローカルメモリ１０４－Ｎにおいてバッファ空間が空いているかどうかを確定することができる。判断部６１４にて「いいえ」の場合には、ブロック６１６において、ランデブー管理モジュール３０４は、宛先サーバノードＣＰＵ１０２によって受信バッファがポストされるまで、宛先ローカルメモリ１０４－Ｎへのデータ送信を待機することができる。判断部６１４にて「はい」の場合には、ブロック６１８において、ランデブーメモリ３０８にてデータが転送用として使用可能であって、且つ、ランデブーメモリ３０８用に受信バッファが割り当てられた、との判定がなされた場合に、ランデブーメモリ３０８から宛先サーバノードメモリ１０４－Ｎにデータを転送することができる。

【００５２】

図７は、一実施形態における、サーバノードと、ランデブーメモリを有するサーバ集約システムとの間で電子通信を行う方法７００のフロー図である。方法７００は、図３の一つ以上のサーバノード３１２のメモリ１０４及び／又はサーバ集約システム３００に記憶された命令に従うことができる。従って、図２～５が参照される。方法７００のいくつか又はすべては、ＡＳＩＣ、システムオンチップ（ＳＯＣ）又は関連する集積回路にて、オンダイで実行することができる。

【００５３】

先ず、サーバノード３１２－１は、セル、パケット、フレーム又は他のデータのユニッ

トがローカルメモリ１０４－１にあって、宛先（例えば、他のサーバノード３１２－Ｎ）又はＮＩＣ３４２を介してリモート計算デバイス３５２に出力されるかどうかの判定を行う。ローカルメモリ１０４－１は、データを処理するためのキュー、リングバッファ、リンクされたリスト等を含む。

【００５４】

ブロック７０２では、サーバノード３１２－１のＣＰＵ１０２は、サーバノード３１２－１から出力されるローカルメモリ１０４－１内のデータに関する記述子又は関連情報を生成することができる。記述子は、ランデブーメモリ３０８に転送されるデータに関連したコマンド、分散収集リスト等を含むことができる。関連する記述子情報は、データが常駐するメモリ１０４－１における記憶位置、データが伝送される宛先アドレス、移動させるデータのバイト数、及び／又は、ＣＰＵ１０２と他のサーバ集約システム３００におけるｖＮＩＣ３３４との間で通信を確立するための関連情報を識別することができる。

【００５５】

ブロック７０４では、第１のサーバノード３１２－１のＣＰＵ１０２は、新規記述子がメモリ１０４－１にあって伝送に使用可能であるということを、ｖＮＩＣ３３４に通知する。ＣＰＵ１０２は、メールボックス書き込み又は関連イベントの通知を、通信経路３１４を介してｖＮＩＣ３３４に送信することによって、伝送用データの使用可能性をｖＮＩＣ３３４に通知することができる。

【００５６】

ブロック７０６では、Ｉ／Ｏプロセッサ３０６は、ランデブーメモリ３０８内の使用可能なデータバッファ（空きデータバッファ）を示す受信記述子をポストする。受信記述子は、バッファのアドレス、長さ又は関連情報を含むことができる。

【００５７】

ブロック７０８では、Ｉ／Ｏプロセッサ３０６は、ｖＮＩＣ３３４用の一つ以上のランデブーメモリバッファの使用可能性（空き状況）を、ランデブー管理モジュール３０４に通知する。

【００５８】

ブロック７１０では、コントロールプレーンプロセッサ３１０は、ランデブーメモリ３０８の一つ以上のバッファ記憶位置にｖＮＩＣ３３４用の空きを要求するように、ランデブー管理モジュール３０４にメッセージを送信する。

【００５９】

ブロック７１２では、ランデブー管理モジュール３０４は、ランデブーメモリ３０８に対して、ｖＮＩＣ３３４用の一つ以上の受信バッファを割り当てて、ｖＮＩＣ３３４が、サーバノードメモリ１０４－１からランデブーメモリ３０８にｖＮＩＣ３３４がデータを転送、コピー又は移動させることができるようにする。ランデブー管理モジュール３０４は、バッファ割り当てに関する通知をコントロールプレーンプロセッサ３１０に送信する。現在空いている受信バッファがない場合には、ランデブー管理モジュール３０４は、受信バッファが空くまで待機することができる。記述子及び／又はデータは、この待ちの期間にフェッチされない。一実施形態では、入力記述子は、待ち時間を削減するためにプリフェッチ（ｐｒｅｆｅｔｃｈ）され得る。これは、入力記述子が、転送待ち時間を削減するために多くのメモリ資源を消費するということがないためである。

【００６０】

ブロック７１４では、ｖＮＩＣ３３４は、データをフェッチするのに使用されるサーバノードメモリ１０４－１から伝送記述子をフェッチする。ｖＮＩＣ３３４は、一つ以上の記述子をフェッチする、又は、フェッチするための一連の記述子に従うことができる。ｖＮＩＣ３３４は、サーバノードメモリ１０４－１からの記述子情報（例えば、アドレス）に従ってデータをフェッチし、データを、スイッチファブリック３０２を介して、ランデブー管理モジュール３０４によって割り当てられた空き（使用可能な）ランデブーメモリ３０８に移動させる。例えば、コントロールプレーンプロセッサ３１０は、ＰＣＩｅコネクター３１４を介してメモリ１０４－１に読み出しを発令し、データをフェッチしてラン

デブーメモリ３０８に移動させることができる。

【００６１】

ブロック７１６では、コントロールプレーンプロセッサ３１０は、ランデブー管理モジュール３０４に対して、フェッチされたデータがランデブーメモリ３０８において使用可能であるとの通知を送信することができる。ランデブー管理モジュール３０４は、Ｉ／Ｏプロセッサ３０６に伝送される割り込みを生成することができる。これにより、Ｉ／Ｏプロセッサ３０６は、データをランデブーメモリ３０８からプル（ｐｕｌｌ）し、当該データを提供することができる。

【００６２】

ブロック７１８では、Ｉ／Ｏプロセッサ３０６は、ランデブーメモリ３０８内のデータを処理して、例えば、データを消費するかどうか、ＮＩＣ３４２に当該データを転送するかどうか、又は、計算インフラストラクチャー３０の一部である他のサーバノード３１２のＣＰＵ１０２に当該データを転送するかどうかを判定することができる。

【００６３】

示されてはいないが、コントロールプレーンプロセッサ３１０は、サーバノード３１２－１で生成された伝送記述子を解放することができる。コントロールプレーンプロセッサ３１０は、任意に、サーバノードＣＰＵ１０２に割り込むことができる。Ｉ／Ｏプロセッサ３０６は、例えば、スケジューラ５０８に従って、ランデブーメモリ３０８内のデータの送信を制御することができる。このように、Ｉ／Ｏプロセッサ３０６は、その宛先（例えば、宛先サーバノード３１２－Ｎ）へのデータの送信を保証することができ、これは図８に記載されるとおりである。

【００６４】

図８は、一実施形態における、ランデブースイッチングを行う構成のサーバ集約システムに結合した第１及び第２のサーバノード間で通信を確立する方法８００のフロー図である。方法８００は、サーバ集約システムのネットワークプロセッサによる図７のブロック７１８での判定に応じて実行することができる。ネットワークプロセッサは、ランデブーメモリ３０８内の使用可能なフェッチされたデータを、他のサーバノード、又は、ＣＰＵを有する関連する電子デバイス（例えば、図３に示されたサーバノード３１２－Ｎ）に転送する。方法８００は、図３のサーバノード３１２のメモリ１０４及び／又はサーバ集約システム３００に記憶された命令により制御することができる。この場合には、図２～５及び図７が参照される。

【００６５】

ブロック８０２では、Ｉ／Ｏプロセッサ３０６の記述子処理モジュール５０４は、図７に記載される方法７００に従ってランデブーメモリ３０８に移動されるデータに関する記述子又は関連する情報を提供することができる。記述子は、宛先サーバノード３１２－Ｎのローカルメモリ１０４－Ｎにデータを転送する命令、分散収集リスト等を含むことができる。関連する記述子情報は、データが常駐するメモリ１０４－１における記憶位置、データが伝送される宛先アドレス、移動させるデータのバイト数、及び／又は、Ｉ／Ｏプロセッサ３０６と、ローカルメモリ１０４－Ｎへのデータの転送に関与するｖＮＩＣ３３４との間で通信を確立する他の関連情報を識別することができる。

【００６６】

ブロック８０４では、Ｉ／Ｏプロセッサ３０６は、ランデブーメモリ３０８内のデータが伝送用に使用可能であることを、ランデブー管理モジュール３０４に通知する。

【００６７】

ブロック８０６では、宛先サーバノード３１２－ＮのＣＰＵ１０２は、宛先サーバノード３１２－Ｎのメモリ１０４－Ｎ内の使用可能なデータバッファ（空きデータバッファ）を示す受信記述子をポストする。受信記述子は、バッファのアドレス、長さ又は関連情報を含むことができる。

【００６８】

ブロック８０８では、宛先サーバノード３１２－ＮのＣＰＵ１０２は、コントロールプ

レーンプロセッサ３１０に対して、メモリ１０４－Ｎにおけるランデブー管理モジュール３０４用の一つ以上のホストメモリバッファの使用可能性（空き状況）を通知する。

【００６９】

ブロック８１０では、宛先サーバノード３１２－Ｎ内のＣＰＵ１０２は、コントロールプレーンプロセッサ３１０に要求を送信して、宛先ノードメモリ１０４－Ｎにおける一つ以上のバッファ記憶位置を空きにするようにランデブー管理モジュール３０４に要求する。

【００７０】

ブロック８１２では、コントロールプレーンプロセッサ３１０は、ランデブー管理モジュール３０４用に宛先ノードメモリ１０４－Ｎの受信バッファを割り当てて、ランデブー管理モジュール３０４が、ランデブーメモリ３０８から宛先ノードメモリ１０４－Ｎへのデータの転送、コピー等を行えるようにする。コントロールプレーンプロセッサ３１０は、バッファ割り当てに関する通知を、ランデブー管理モジュール３０４に送信することができる。現在空いている受信バッファがない場合には、ランデブー管理モジュール３０４は、宛先サーバノードメモリ１０４－Ｎの受信バッファが空くまで待機することができる。一実施形態では、記述子及び／又はデータは、待ち期間の間フェッチされない。

【００７１】

ブロック８１４では、ランデブー管理モジュール３０４は、ランデブーメモリ３０８からのデータをフェッチするのに使用される伝送記述子をフェッチする。ランデブー管理モジュール３０４は、ランデブーメモリ３０８からの記述子情報（例えば、アドレス）に従ってデータをフェッチし、データを、スイッチファブリック３０２を介して、ｖｎｉｃ６３４に移動させる。ランデブー管理モジュール３０４は、コントロールプレーンプロセッサ３１０に対して、宛先サーバノード３１２－Ｎ用のランデブーメモリ３０８からｖＮＩＣ３３４への記述子及び／又はデータをフェッチした後に新規データが使用可能であることを、通知することができる。

【００７２】

ブロック８１６では、ｖＮＩＣ／ＣＰは、受信バッファ用の記述子をフェッチ及び処理し、データを宛先サーバノードメモリ１０４－Ｎに移動させる。コントロールプレーンプロセッサ３１０は、メモリ１０４－Ｎに移動したデータを提供するために、宛先サーバノード３１２－ＮのＣＰＵ１０２に出力される割り込みを生成することができる。ランデブー管理モジュール３０４は、伝送記述子を解放することができ、サーバノードＩ／Ｏプロセッサ３０６に任意に割り込みすることができる。

【００７３】

本明細書に記載の方法の実施形態により、データが宛先に送信できる状態にあるとサーバ集約システム３００が判定した場合には、要求に応じてデータを転送することができる。ランデブーメモリ３０８がデータを受信できる状態にない場合には、データは宛先に伝送されない。これは、仮に、データがソースサーバノード３１２－１において転送できる状態にあってもそうである。同様に、宛先サーバノード３１２－Ｎがデータを受信できる状態にない場合には、データは、ランデブーメモリ３０８から伝送されない。このように、従来のキューイング技術を必要とすることなく、バッファ空間がないことに起因してデータが失われるということはない。むしろ、データの移動は記述子の使用可能性と結びついている。記述子が確立される場合には、サーバ集約システム３００は、移動させることになるデータが存在するとの判定を行うことができる。

【００７４】

図９は、一実施形態における、ランデブースイッチングを行う構成のサーバ集約システムに結合した第１及び第２のサーバノード間で電子通信する方法９００のフロー図である。方法９００は、一つ以上のサーバノード３１２のメモリ１０４及び／又は上記のサーバ集約システム３００に記憶された命令により制御することができる。方法９００の一つ以上の構成要素は、上述したものと同様であってもよい。図２～５が参照されるが、方法９００は、ランデブーメモリ３０８無しに実行される。特に、ランデブー管理モジュール３

０４及びコントロールプレーンプロセッサ３１０は、方法９００に関与することができ、この方法では、データ転送が、二つのサーバノード２１２－１，２１２－Ｎの間で、ランデブーメモリ３０８の介在無しに行われる。

【００７５】

ブロック９０２では、記述子は、第１のサーバノード３１２－１で生成される。

【００７６】

ブロック９０４では、ｖＮＩＣ３３４は、データが伝送用に使用可能であることを通知される。

【００７７】

ブロック９０６では、受信記述子は、第２のサーバノード３１２－Ｎの使用可能なバッファ（空きバッファ）でポストされる。

【００７８】

ブロック９０８では、宛先サーバノード３１２－Ｎは、コントロールプレーンプロセッサ３１０に対して、ランデブー管理モジュール３０４用のメモリ１０４－Ｎの一つ以上のホストメモリバッファの使用可能性（空き状況）を通知する。

【００７９】

ブロック９１０では、宛先メモリ１０４－Ｎのバッファの要求がなされる。

【００８０】

ブロック９１２では、受信バッファが割り当てられる。バッファは、ｖＮＩＣ３３４用に割り当てることができ、これにより、コントロールプレーンプロセッサ３１０のｖＮＩＣ３３４が、サーバノードメモリ１０４－１から宛先ノードメモリ１０４－Ｎにデータを転送、コピー又は移動させることができる。

【００８１】

ブロック９１４では、ｖＮＩＣ３３４は、受信バッファ用の記述子をフェッチ及び処理して、データを宛先サーバノードメモリ１０４－Ｎに移動させる。

【００８２】

当業者が認識し得るように、本発明の態様は、システム、方法又はコンピュータプログラム製造物として具体化されてもよい。従って、本発明の態様は、完全にハードウェアの実施形態、完全にソフトウェアの実施形態（ファームウェア、常駐ソフトウェア、マイクロコード等を含む）、又は、ソフトウェア及びハードウェアの態様を組み合わせた実施形態であって、本明細書では「回路」、「モジュール」又は「システム」と称される形態であってもよい。さらには、本発明の態様は、コンピュータプログラム製品の形態であって、一つ以上のコンピュータ可読媒体で具体化されてもよい。コンピュータ可読媒体は、コンピュータ可読プログラムコードを有していてもよい。

【００８３】

一つ以上のコンピュータ可読媒体のあらゆる組合せが利用できる。コンピュータ可読媒体は、コンピュータ可読信号媒体、又は、コンピュータ可読記憶媒体であってもよい。コンピュータ可読記憶媒体は、例えば、電気的、磁気的、光学的、電磁気的、赤外線であってもよく、半導体システム、装置若しくはデバイスであってもよく、これらのあらゆる好適な組合せであってよいが、これらに限定されるものではない。コンピュータで可読記憶媒体のさらに特定の例（非包括的な列挙）は、以下のものを含む。すなわち、一つ以上の配線を有する電気的接続、ポータブルコンピュータディスケット、ハードディスク、ランダムアクセスメモリ（ＲＡＭ）、リードオンリーメモリ（ＲＯＭ）、イレーザブルプログラマブルリードオンリーメモリ（ＥＰＲＯＭ若しくはフラッシュメモリ）、光ファイバー、ポータブルコンパクトディスクリードオンリーメモリ（ＣＤ－ＲＯＭ）、光記憶デバイス、磁気記憶デバイス、又は、これらのあらゆる好適な組合せである。本文献の文脈においては、コンピュータ可読記憶媒体は、あらゆる有形の媒体であって、命令実行システム、装置若しくはデバイスにより使用されるプログラム、又は、これらに接続して使用するプログラムを含む又は記憶するものであってもよい。

【００８４】

コンピュータ可読信号媒体は、その内部で具体化されるコンピュータ可読プログラムコード（例えば、ベースバンド又は搬送波の一部として伝搬されるデータ信号）を含んでいてもよい。このような伝搬信号は、あらゆる多様な形態（例えば、電磁的、光学的又はこれらのあらゆる好適な組合せの形態）をとってもよいが、これらには限定されない。コンピュータ可読信号媒体は、コンピュータで読み取り可能なあらゆる媒体であって、コンピュータ可読記憶媒体ではなく、命令実行システム、装置若しくはデバイスによって使用されるプログラム、又は、これらと接続して使用するプログラムを、通信、伝搬又は移動することが可能な媒体であってもよい。コンピュータ可読媒体上で具体化されるプログラムコードは、あらゆる適切な媒体、例えば、限定はされないが、無線、有線、光ファイバーケーブル、ＲＦ等又は前述のあらゆる好適な組合せを用いて伝送することができる。

【００８５】

本発明の態様の動作を実行するコンピュータプログラムコードは、Ｊａｖａ（登録商標）、Ｓｍａｌｌｔａｌｋ、Ｃ＋＋等のオブジェクト指向プログラミング言語及び「Ｃ」プログラミング言語等の従来の手続型言語、又は、同様のプログラミング言語を含む一つ以上のプログラミング言語のあらゆる組合せで記述されていてもよい。プログラムコードは、全部がユーザのコンピュータ上で実行されてもよいし、一部がユーザのコンピュータ上で実行されてもよい。また、プログラムコードは、スタンドアロンのソフトウェアパッケージとして、一部がユーザのコンピュータ上又はリモートコンピュータ上で実行されてもよいし、全部がリモートコンピュータ上又はサーバ上で実行されてもよい。後者のシナリオでは、リモートコンピュータは、ローカルエリアネットワーク（ＬＡＮ）又はワイドエリアネットワーク（ＷＡＮ）を含むあらゆるタイプのネットワークを介してユーザのコンピュータに接続されていてもよいし、（例えば、インターネットサービスプロバイダを用いたインターネットを介して）外部のコンピュータと接続されていてもよい。

【００８６】

本発明の態様は、本発明の実施形態の方法、装置(システム)及びコンピュータプログラム製品のフローチャート並びに／又はブロック図を参照しつつ記載されている。フローチャート図及び／又はブロック図の各ブロックと、フローチャート図及び／又はブロック図におけるブロックの組合せとは、コンピュータプログラムの命令により実現可能であることは理解されるであろう。コンピュータプログラムのこれらの命令は、汎用コンピュータ、専用コンピュータ、又は、機械を製造する他のプログラマブルデータ処理装置のプロセッサに提供され、これらのコンピュータ又は他のプログラマブルデータ処理装置のプロセッサにおいて実行され、フローチャート及び／又はブロック図のブロックにおいて特定された機能／作用を実現する手段を生成するようにしてもよい。

【００８７】

また、コンピュータプログラムのこれらの命令は、コンピュータ可読媒体に記憶されてもよく、この媒体が、コンピュータ、他のプログラマブルデータ処理装置、又は、特定の方法で機能する他のデバイスに命令して、コンピュータ可読媒体に記憶された命令によって製品が製造され、その製品に、フローチャート及び／又はブロック図のブロックにおいて特定された機能／作用を実現する命令が含まれるようにしてもよい。さらに、コンピュータプログラムの命令は、コンピュータ、他のプログラマブルデータ処理装置、又は、他のデバイスにロードされて、コンピュータ、他のプログラマブル装置、又は、他のデバイス上で一連の処理ステップを実行させ、コンピュータで実現されるプロセスを生成させ、コンピュータ又は他のプログラマブル装置上で実行されるこれらの命令が、フローチャート及び／又はブロック図のブロックにおいて特定される機能／作用を実現するプロセスを提供するようにしてもよい。

【００８８】

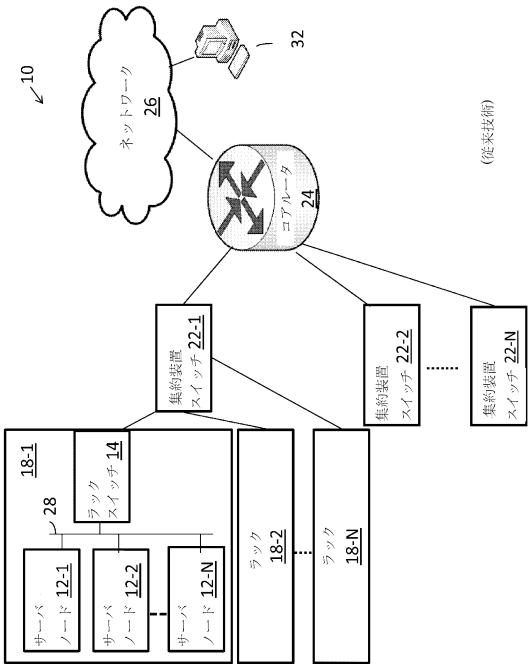
図におけるフローチャート及びブロック図は、本発明の様々な実施形態によるシステム、方法及びコンピュータプログラム製品の実施可能なアーキテクチャ、機能性及び動作を例示する。この点で、フローチャート又はブロック図における各ブロックは、モジュール、セグメント又はコードの一部を表してもよく、これらは、特定された論理機能を実現す

る一つ以上の実行可能な命令を含む。いくつかの代替の実施物では、ブロックで指摘される機能は、図において指摘される順序に従うことなく生じてもよいことも留意されたい。例えば、連続して示された二つのブロックは、実際には、実質的に同時に実行されてもよいし、ブロックは、場合によっては逆順に実行されてもよく、これは関係する機能による。ブロック図及び／又はフローチャートの各ブロックと、ブロック図及び／又はフローチャートにおけるブロックの組合せとは、専用のハードウェア及びコンピュータ命令の特定された機能若しくは作用又は組合せを実行する、専用のハードウェアをベースにしたシステムによって実現されてもよいことも留意されるであろう。

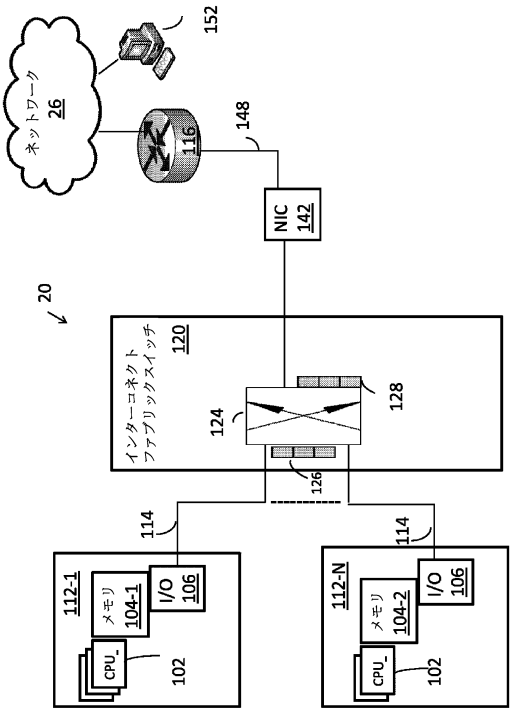
【００８９】

特定の実施形態を参照して、本発明を示し記載してきたが、形態及び詳細の様々な変更が、本発明の精神及び範囲を逸脱することなく、当業者により行われてもよいことは理解されるべきである。

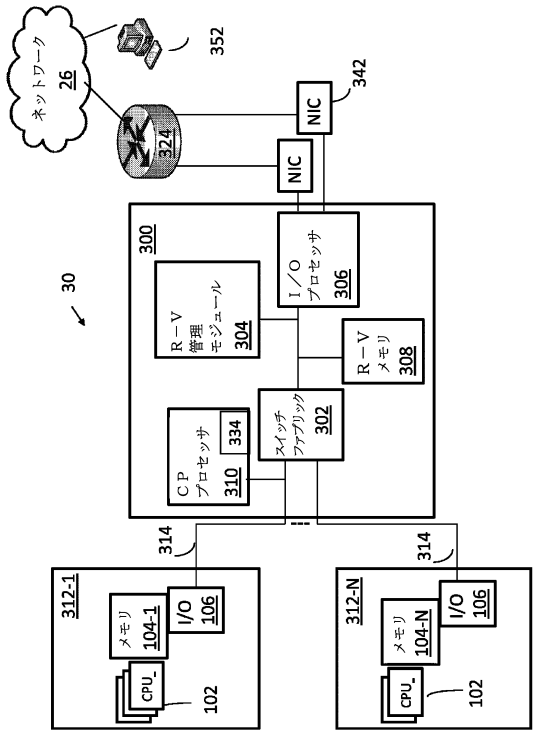
【図１】



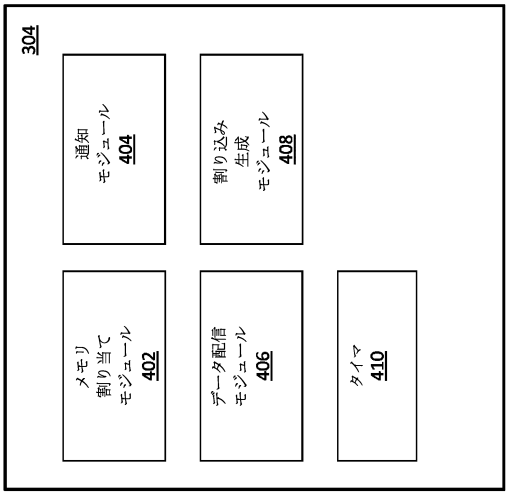

【図２】



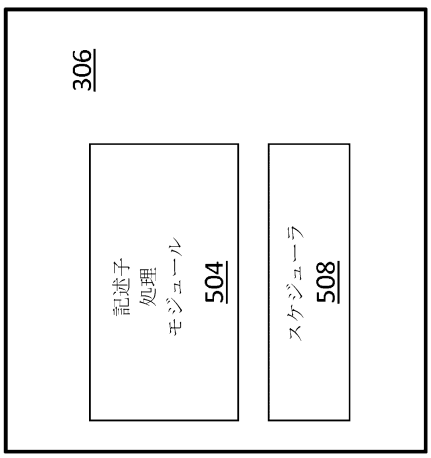
【図３】


ネットワーク
26
352
324
NIC
342
NIC
30
300
R－V
管理
モジュール
304
I／O
プロセッサ
306
R－V
メモリ
308
CP
プロセッサ
310
334
スイッチ
ファブリック
302
314
314
312-1
メモリ
104-1
I/O
106
CPU
102
312-N
メモリ
104-N
I/O
106
CPU
102

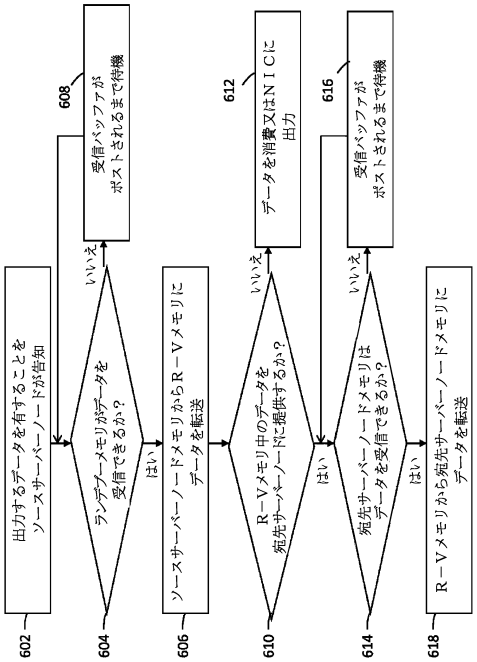
【図４】


304
通知
モジュール
404
割り込み
生成
モジュール
408
メモリ
割り当て
モジュール
402
データ配信
モジュール
406
タイマ
410

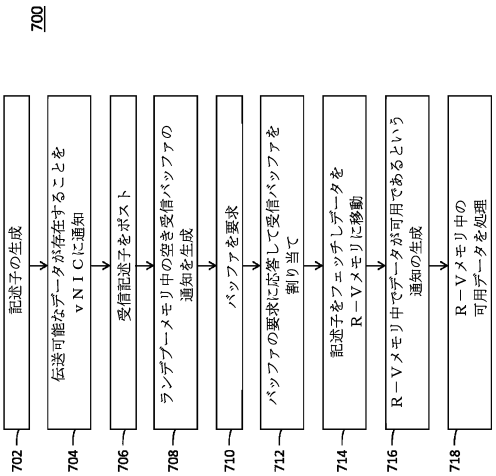
【図５】


306
記述子
処理
モジュール
504
スケジューラ
508

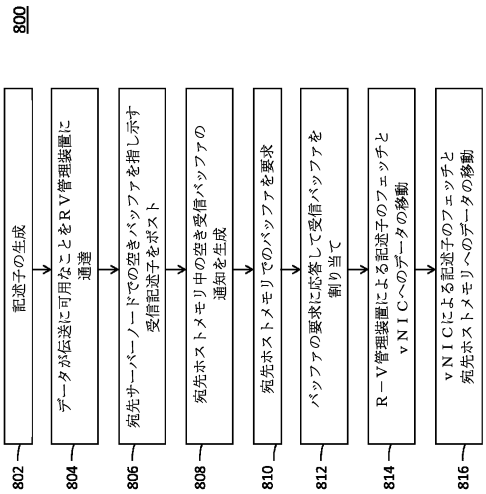
【図６】


出力するデータを有することを
ソースサーバーノードが告知
602
ランデブーメモリがデータを
受信できるか？
604
いいえ
受信バッファが
ポストされるまで待機
608
はい
ソースサーバーノードからR－Vメモリに
データを転送
606
R－Vメモリ中のデータを
宛先サーバーノードに提供するか？
610
いいえ
データを消費又はＮＩＣに
出力
612
はい
宛先サーバーノードメモリは
データを受信できるか？
614
いいえ
受信バッファが
ポストされるまで待機
616
はい
R－Vメモリから宛先サーバーノードメモリに
データを転送
618

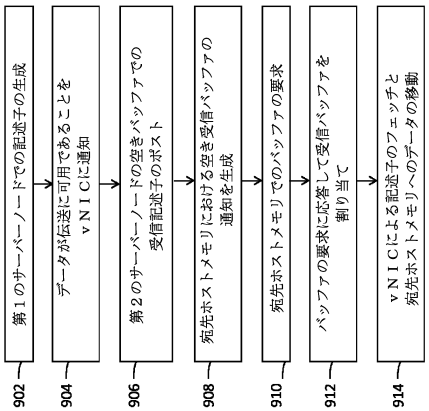
【図７】


700
記述子の生成
伝送可能なデータが存在することをvNICに通知
受信記述子をポスト
ランデブーメモリ中の空き受信バッファの通知を生成
バッファを要求
バッファの要求に応答して受信バッファを割り当て
記述子をフェッチしデータをR－Vメモリに移動
R－Vメモリ中でデータが可用であるという通知の生成
R－Vメモリ中の可用データを処理
702
704
706
708
710
712
714
716
718

【図８】


800
記述子の生成
データが伝送に可用なことをRV管理装置に通達
宛先サーバーノードでの空きバッファを指し示す受信記述子をポスト
宛先ホストメモリ中の空き受信バッファの通知を生成
宛先ホストメモリでのバッファを要求
バッファの要求に応答して受信バッファを割り当て
R－V管理装置による記述子のフェッチとvNICへのデータの移動
vNICによる記述子のフェッチと宛先ホストメモリへのデータの移動
802
804
806
808
810
812
814
816

【図９】


第１のサーバーノードでの記述子の生成
データが伝送に可用であることをvNICに通知
第２のサーバーノードの空きバッファでの受信記述子のポスト
宛先ホストメモリにおける空き受信バッファの通知を生成
宛先ホストメモリでのバッファの要求
バッファの要求に応答して受信バッファを割り当て
vNICによる記述子のフェッチと宛先ホストメモリへのデータの移動
902
904
906
908
910
912
914

フロントページの続き

(72)発明者　マーク　ハメル
アメリカ合衆国　０２０３８　マサチューセッツ州、フランクリン、スチュワート　ストリート　６８
(72)発明者　デイビッド　メイヒュー
アメリカ合衆国　０１５３２　マサチューセッツ州、ノースボロー、プレザント　ストリート　１５９
(72)発明者　マイケル　オズボーン
アメリカ合衆国　０３０４９　ニューハンプシャー州、ホリス、ブラック　オーク　ドライブ　５０

審査官　衣鳩　文彦

(56)参考文献　特開平１０－２２９４２９（ＪＰ，Ａ）
特開２００５－２０４０８９（ＪＰ，Ａ）
米国特許出願公開第２００５／００４１５１０（ＵＳ，Ａ１）
米国特許出願公開第２０１１／０２０２７０１（ＵＳ，Ａ１）
米国特許出願公開第２００３／０２０８６３１（ＵＳ，Ａ１）
特開２００９－２８２９１７（ＪＰ，Ａ）
特開平０３－０８２２４４（ＪＰ，Ａ）

(58)調査した分野(Int.Cl.，ＤＢ名)
Ｈ０４Ｌ　１２／９３１
Ｇ０６Ｆ　１５／１７３