

(19) 日本国特許庁(JP)

(12) 特許公報(B2)

(11) 特許番号

特許第5834178号  
(P5834178)

(45) 発行日 平成27年12月16日(2015.12.16)

(24) 登録日 平成27年11月13日(2015.11.13)

(51) Int.Cl.		F I	
HO 4 L 12/40	(2006.01)	HO 4 L 12/40	B
HO 4 L 12/701	(2013.01)	HO 4 L 12/701	
HO 4 L 12/911	(2013.01)	HO 4 L 12/911	

請求項の数 11 (全 27 頁)

(21) 出願番号	特願2014-527396 (P2014-527396)	(73) 特許権者	314012076
(86) (22) 出願日	平成25年10月31日(2013.10.31)		パナソニックIPマネジメント株式会社
(86) 国際出願番号	PCT/JP2013/006457		大阪府大阪市中央区城見2丁目1番61号
(87) 国際公開番号	W02014/073188	(74) 代理人	100101683
(87) 国際公開日	平成26年5月15日(2014.5.15)		弁理士 奥田 誠司
審査請求日	平成26年6月9日(2014.6.9)	(74) 代理人	100155000
(31) 優先権主張番号	特願2012-246474 (P2012-246474)		弁理士 喜多 修市
(32) 優先日	平成24年11月8日(2012.11.8)	(74) 代理人	100180529
(33) 優先権主張国	日本国(JP)		弁理士 梶谷 美道
		(74) 代理人	100125922
			弁理士 三宅 章子
		(74) 代理人	100135703
			弁理士 岡部 英隆
		(74) 代理人	100188813
			弁理士 川喜田 徹

最終頁に続く

(54) 【発明の名称】 半導体回路のバスシステム

(57) 【特許請求の範囲】

【請求項1】

第1伝送レートを有する第1バス、および前記第1伝送レートよりも速い第2伝送レートを有する第2バスと、

データを送信する第1ノードと、

前記第1ノードを前記第1バスに接続するバスインタフェースと、

前記第1バスおよび前記第2バスを接続するルータと、

前記第2バスに接続され、前記データを受信する第2ノードと

を備えた半導体回路のバスシステムであって、

前記第1バスは、前記バスインタフェースから前記ルータへ至る伝送経路を複数有する分散型バスであり、

前記ルータは、

所定の基準に従って前記第1バスの各伝送経路に送信可能なデータ送信量を割り当て、各伝送経路の送信量に関する前記情報を前記バスインタフェースに通知する割当部と、

前記第1バスの各伝送経路を流れる前記データを受け取って前記第2バスに転送するルータ処理部と、

前記第2バスに流れる前記データの流量を制御する第2制御部と

を備え、

前記バスインタフェースは、

前記第1ノードから受け取った前記データを、前記第1バスの各伝送経路に転送する

10

20

転送処理部と、

前記ルータから通知された、各伝送経路の送信量に関する情報に基づいて、前記第1バスの各伝送経路を流れるデータの流量を制御する第1制御部と

を備えた、半導体回路のバスシステム。

【請求項2】

前記第2制御部は、前記第1バスの各伝送経路のスループット、および前記第2バスのスループットを利用して、前記第2バスに流れる前記データの流量を制御する、請求項1に記載のバスシステム。

【請求項3】

前記バスインタフェースの転送処理部は、前記第1ノードから受け取ったデータを、パケット化して前記第1バスの各伝送経路に転送し、

前記所定の基準として、前記第1バスの各伝送経路の送信可能なパケット数の上限値の情報が予め定められており、

前記ルータの割当部は、前記上限値の情報に従って、前記各伝送経路に送信可能なデータ送信量を割り当てる、請求項1または2に記載のバスシステム。

【請求項4】

前記バスインタフェースの転送処理部は、前記第1ノードから受け取ったデータを、パケット化して前記第1バスの各伝送経路に転送し、

前記ルータの割当部は、前記所定の基準として、前記第1バスの各伝送経路のスループットに従って、前記各伝送経路に送信可能なデータ送信量を割り当てる、請求項1または2に記載のバスシステム。

【請求項5】

前記バスインタフェースの転送処理部は、前記第1ノードから受け取ったデータを、パケット化して前記第1バスの各伝送経路に転送し、

前記ルータの割当部は、前記パケット数の上限値の情報に加えて、さらに前記第1バスの各伝送経路のスループットの情報を前記所定の基準として用いて、前記各伝送経路に送信可能なデータ送信量を割り当てる、請求項3に記載のバスシステム。

【請求項6】

前記ルータの割当部は、前記第1バスの全ての伝送経路のスループットに対する各伝送経路のスループットの割合に従って、前記各伝送経路に送信可能なデータ送信量を割り当てる、請求項4または5に記載のバスシステム。

【請求項7】

前記第2ノードは、前記第1ノードにデータを送信することが可能であり、

前記ルータ処理部は、前記第2ノードから受け取ったデータを、前記第1バスを介して前記第1ノードに転送する、請求項1に記載のバスシステム。

【請求項8】

前記ルータ処理部は、前記第2ノードから受け取ったデータを、前記第1バスの複数の伝送経路を介して前記第1ノードに転送する、請求項7に記載のバスシステム。

【請求項9】

第1伝送レートを有する第1バス、および前記第1伝送レートよりも速い第2伝送レートを有する第2バスと、

各々がデータを送信する複数の第1ノードと、

前記複数の第1ノードの各々を前記第1バスに接続する複数のバスインタフェースと、

前記第1バスおよび前記第2バスを接続するルータと、

前記第2バスに接続され、前記データを受信する、少なくとも1つの第2ノードと

を備えた半導体回路のバスシステムであって、

前記第1バスは、前記バスインタフェースから前記ルータへ至る伝送経路を複数有する分散型バスであり、

前記ルータは、

所定の基準に従って前記第1バスの各伝送経路に送信可能なデータ送信量を割り当て

10

20

30

40

50

、各伝送経路の送信量に関する前記情報を前記バスインタフェースに通知する割当部と、  
前記第 1 バスの各伝送経路を流れる前記データを受け取って前記第 2 バスに転送する  
ルータ処理部と、

前記第 2 バスに流れる前記データの流量を制御する第 2 制御部であって、前記複数の  
第 1 ノードから受け取った複数のデータを、第 1 ノード毎に保証すべき伝送レートを確保  
できるように集約して、前記第 2 ノードに伝送する、第 2 制御部と

を備え、

各バスインタフェースは、

接続された第 1 ノードから受け取ったデータを、前記第 1 バスの各伝送経路に転送す  
る転送処理部と、

10

保証されるべき伝送レートに基づいて制限された伝送帯域、および、前記ルータから  
通知された、各伝送経路の送信量に関する情報に基づいて、前記第 1 バスの各伝送経路を  
流れるデータの流量を制御する第 1 制御部と

を備えた、半導体回路のバスシステム。

【請求項 10】

保証されるべき伝送レートに基づいて制限された前記伝送帯域は、前記ルータ、または  
前記ルータとは異なるルータによって制限された伝送帯域である、請求項 9 に記載のバス  
システム。

【請求項 11】

第 1 伝送レートを有する第 1 バス、前記第 1 伝送レートよりも遅い第 2 伝送レートを有  
する第 2 バス、および前記第 2 バスよりも速い第 3 バスと、

20

データを送信する第 1 ノードと、

前記第 1 ノードを前記第 1 バスに接続するバスインタフェースと、

前記第 1 バスおよび前記第 2 バスを接続する第 1 ルータと、

前記第 2 バスおよび前記第 3 バスを接続する第 2 ルータと、

前記第 3 バスに接続され、前記データを受信する第 2 ノードと

を備えた半導体回路のバスシステムであって、

前記第 1 バスは、前記バスインタフェースから前記ルータへ至る伝送経路を複数有する  
分散型バスであり、

前記第 1 ルータは、

30

所定の基準に従って前記第 1 バスの各伝送経路に送信可能なデータ送信量を割り当て  
、各伝送経路の送信量に関する前記情報を前記バスインタフェースに通知する割当部と、

前記第 1 バスの各伝送経路を流れる前記データを受け取って前記第 2 バスに転送する  
ルータ処理部と、

前記第 2 バスに流れる前記データの流量を制御する第 2 制御部と

を備え、

前記第 2 ルータは、

前記第 2 バスの各伝送経路を流れる前記データを受け取って前記第 3 バスに転送する  
ルータ処理部と、

前記第 3 バスに流れる前記データの流量を制御する第 2 制御部と

40

を備え、

前記バスインタフェースは、

前記第 1 ノードから受け取ったデータを、前記第 1 バスの各伝送経路に転送する転送  
処理部と、

前記第 1 ルータから通知された、各伝送経路の送信量に関する情報に基づいて、前記  
第 1 バスの各伝送経路を流れるデータの流量を制御する第 1 制御部と

を備えた、半導体回路のバスシステム。

【発明の詳細な説明】

【技術分野】

【0001】

50

本開示は、半導体回路のバスシステムに関する。

【背景技術】

【0002】

図1(a)は、集中型のバス制御の例を示す。集中型のバス制御を行う従来の集積回路では、主に複数のバスマスタとメモリとの間は1つのバスによって接続され、アービターによって各バスマスタからメモリへのアクセスが調停される。しかし、集積回路の高機能化、マルチコア化により、回路の規模が増大し、トラヒックも複雑に変動しながらバスを流れており、集中型のバス制御による集積回路の設計が困難になってきている。

【0003】

その一方で、近年、並列計算機での接続技術や、ATM(Asynchronous Transfer Mode)網などのネットワーク制御の技術を取り入れた分散型のバスを有する半導体集積回路の開発が進んでいる。図1(b)は、分散型のバス制御の例を示す。分散型のバスを有する半導体集積回路は、複数のルータが複数のバスで接続されて構成されている。近年、図1(b)に示すような分散型のバスを用いることにより、大規模化した集積回路内のトラヒックを、複数のバスに分散して伝送するネットワークオンチップ(Network on Chip)の取り組みがある。

【0004】

分散型のバス制御を行うNoC上でも、バスのトラヒックは増加する傾向にあり、バスの伝送性能の向上が求められている。バスの伝送性能を維持するためには、トラヒックの増加にあわせて、伝送路を並列化して高速化する方法が考えられる。

【0005】

特許文献1は、イーサネット(登録商標)規格のネットワーク(通信網)において複数の通信端末が伝送路を共有しながら通信を行う、一般的な通信システムを開示する。特許文献1は、伝送路を並列化して高速化する方法として、必要に応じて、1つの送信端末が複数の伝送路を使用して広帯域化を図る通信方式を提案している。

【0006】

図2は、特許文献1に記載された通信システムの構成を示している。この通信システムでは、伝送路よりも左側に記載された送信端末と、伝送路よりも右側に記載された受信端末とが、複数の伝送路を用いてより高速な通信を行う。

【0007】

従来技術では、伝送路毎の通信品質を往復伝搬遅延時間(RTT)として計測する。RTTが大きい場合は、伝送路は混雑していると判断し、送信レートを下げる。逆に、RTTが小さい場合は、伝送路は空いていると判断し、送信レートを上げる。従って、従来技術では伝送路が混雑して破綻するまで伝送を行い、輻輳制御を行うため、予め想定した伝送品質を維持してデータを伝送するのは難しい(特許文献1、非特許文献1)。

【0008】

イーサネット(登録商標)回線を用いるような一般のネットワークでは、ルータのバッファでのパケットの溢れなどの破綻が生じた場合には、ルータが自動的にパケットを廃棄する。一方、NoCでは、一般のネットワークと異なり、ルータはパケットを廃棄しない。その理由は、NoCでは、ネットワークに接続されるバスマスタ等が予め想定され、ネットワーク上を流れるトラヒックデータのデータ量を見込むことができるため、パケットを廃棄しなくてもよいよう、NoCを設計できるからである。

【0009】

したがって、一般のネットワークに適用されてきた技術を、NoCにそのまま適用することはできない。バスの伝送性能を維持するために、これまでの技術を用いて伝送路を並列化して高速化することは困難である。

【先行技術文献】

【特許文献】

【0010】

【特許文献1】特開2009-49742号公報

10

20

30

40

50

## 【非特許文献】

## 【0011】

【非特許文献1】マルチパス環境における高速高信頼トランスポートプロトコルの性能評価、電子情報通信学会信学技法、2005.2

## 【発明の概要】

## 【発明が解決しようとする課題】

## 【0012】

NOCにおいても、バスの伝送性能を維持するために、データの伝送効率をより高めることが求められている。

## 【0013】

本願の、限定的ではない例示的なある実施形態は、たとえば、集積回路上の複数のノードがバスで接続され、分散型のバス制御が行われるNOCにおいて、効率的なパケット伝送を実現する技術を提供する。

## 【課題を解決するための手段】

## 【0014】

上記課題を解決するために、本発明の一態様は、第1伝送レートを有する第1バス、および前記第1伝送レートよりも速い第2伝送レートを有する第2バスと、データを送信する第1ノードと、前記第1ノードを前記第1バスに接続するバスインタフェースと、前記第1バスおよび前記第2バスを接続するルータと、前記第2バスに接続され、前記データを受信する第2ノードとを備えた半導体回路のバスシステムであって、前記第1バスは、前記バスインタフェースから前記ルータへ至る伝送経路を複数有する分散型バスであり、前記バスインタフェースは、前記第1ノードから受け取ったデータを、前記第1バスの各伝送経路に転送する転送処理部と、前記ルータから通知された、各伝送経路の送信量に関する情報に基づいて、前記第1バスの各伝送経路を流れるデータの流量を制御する第1制御部とを備え、前記ルータは、所定の基準に従って前記第1バスの各伝送経路に送信可能なデータ送信量を割り当て、各伝送経路の送信量に関する前記情報を前記バスインタフェースに通知する割当部と、前記第1バスの各伝送経路を流れる前記データを受け取って前記第2バスに転送するルータ処理部と、前記第2バスに流れる前記データの流量を制御する第2制御部とを備えた、半導体回路のバスシステムを含む。

## 【0015】

上述の一般的かつ特定の態様は、システム、方法およびコンピュータプログラムを用いて実装され、またはシステム、方法およびコンピュータプログラムの組み合わせを用いて実現され得る。

## 【発明の効果】

## 【0016】

本発明の一態様にかかる半導体回路のバスシステムによれば、半導体システムのデータ伝送において、伝送速度が異なるバスを相互接続していても、バスの混雑を回避し、バス動作周波数を抑制したバスの広帯域化が可能になる。

## 【図面の簡単な説明】

## 【0017】

【図1】(a)は集中型のバス制御の例を示す図であり、(b)は分散型のバス制御の例を示す図である。

【図2】従来の文献に記載された通信システムの構成を示す図である。

【図3】例示的な実施形態による伝送システムの構成例を示す図である。

【図4】パケット600の伝送フォーマットの一例と、パケット600を複数のフリットに分割した例とを示す図である。

【図5】例示的な実施形態によるルータ600の構成を示す図である。

【図6】ルータ600の動作手順を示すフローチャートである。

【図7】ローカルバスとシステムバスとの関係を具体的に示す図である。

【図8】システムバス10に接続されているローカルバスの構成を示す図である。

10

20

30

40

50

【図 9】システムバス上のメモリとルータにおいてトラヒックの干渉が生じている例を模式的に示す図である。

【図 10】例示的な実施形態による半導体回路のバスシステム 100 の構成を示す図である。

【図 11】往路での並列伝送時におけるNIC 803 の動作手順を示すフローチャートである。

【図 12】往路での並列伝送時におけるNIC 803 とルータGW 807 との動作シーケンスを示す図である。

【図 13】ローカルバスの伝送経路 1 ~ 3 を模式的に示す図である。

【図 14】図 13 に示す伝送経路 1 ~ 3 の各伝送量の割当規則を示す図である。

10

【図 15】往路での並列伝送時における、ルータGW の処理手順を示すフローチャートである。

【図 16】往路での並列伝送時におけるNIC とルータGW との動作シーケンスを示す図である。

【図 17】復路における伝送帯域のギャップに対応するための処理の手順を示す図である。

【図 18】3 つの伝送経路 4 ~ 6 を用いてデータを伝送する態様を示す図である。

【図 19】システムバスとローカルバスで生じる伝送帯域のギャップに対応する処理の手順を示すフローチャートである。

【図 20】プロセッサと伝送経路との対応付けを示す図である。

20

【図 21】3 つのバスマスタであるプロセッサから伝送されるトラヒックと保証レートとを示す図である。

【図 22】クラスに応じた調停順位を示す図である。

【図 23】往路における保証レートを越えた伝送処理の手順を示すフローチャートである。

【図 24】データが、主経路と、空き帯域のある伝送経路とを用いて伝送される様子を示す図である。

【図 25】( i ) は主経路の帯域の割当状況を示す図であり、( i i ) は空き帯域が存在する、主経路以外の伝送経路の帯域の割当状況を示す図である。

【図 26】複数のSoC を相互接続する場合に適用する場合の例を示す図である。

30

【発明を実施するための形態】

【0018】

本願発明者らは、上述した従来技術の問題を検討した。

【0019】

集積回路上の複数のノードがバスで接続され、分散型のバス制御が行われるNoCのバスの伝送性能を維持するためには、トラヒックの増加にあわせて、バス動作周波数を高めていくか、バス幅を大きくする必要がある。しかしながら、バス動作周波数を高めると消費電力が高くなり、バス幅を大きくすると配線の混雑を引き起こしやすくなる。

【0020】

これらの課題を解決するためのアプローチとして、機能ブロック（たとえば、映像、音声、ペリフェラルの各ブロック）に応じて、バス幅やバス動作周波数を変更する構成が採用されている。例えば、機能ブロック単位で構成されるローカルバスと、機能ブロック間を相互接続するシステムバスとを有する半導体バスシステムが構築されている。システムバスを設けることにより、機能ブロック間やバスマスタ間でメモリ上のデータを共有することが可能となる。このような構成において、ローカルバスのバス動作周波数はシステムバスのバス動作周波数よりも低く設定される。つまり、ローカルバスの伝送速度が相対的に低くなり、システムバスの動作速度が相対的に高くなる。これにより、NoCのバス伝送性能の維持を図っている。

40

【0021】

しかしながら、このような半導体バスシステムでは、ローカルバスとシステムバスとが

50

接続されるノードにおいて、伝送帯域のギャップに起因してデータ処理の停滞が生じやすい。また、システムバスではローカルバス間のトラヒック干渉が生じやすく、伝送可能なトラヒック量が時々刻々と変化する。以上の原因で、ローカルバスとシステムバスを接続するルータや、ローカルバス、システムバスにおいて、混雑が発生しやすくなる。よって、伝送性能が異なるバスを介してトラヒックを伝送する際には、伝送性能（スループットや遅延）を維持することが困難である。

【0022】

そこで本開示は、複数の送信ノードで伝送路を共有しつつ、伝送速度が異なるバスが相互接続されたネットワークモデルに対して、予め想定した伝送品質を維持しながら、複数の伝送路を活用したバスの広帯域化を実現する技術を提供する。

10

【0023】

本発明の一態様の概要は以下のとおりである。

【0024】

本発明の一態様であるバスシステムは、第1伝送レートを有する第1バス、および前記第1伝送レートよりも速い第2伝送レートを有する第2バスと、データを送信する第1ノードと、前記第1ノードを前記第1バスに接続するバスインタフェースと、前記第1バスおよび前記第2バスを接続するルータと、前記第2バスに接続され、前記データを受信する第2ノードとを備えた半導体回路のバスシステムであって、前記第1バスは、前記バスインタフェースから前記ルータへ至る伝送経路を複数有する分散型バスであり、前記ルータは、所定の基準に従って前記第1バスの各伝送経路に送信可能なデータ送信量を割り当て、各伝送経路の送信量に関する前記情報を前記バスインタフェースに通知する割当部と、前記第1バスの各伝送経路を流れる前記データを受け取って前記第2バスに転送するルータ処理部と、前記第2バスに流れる前記データの流量を制御する第2制御部とを備え、前記バスインタフェースは、前記第1ノードから受け取ったデータを、前記第1バスの各伝送経路に転送する転送処理部と、前記ルータから通知された、各伝送経路の送信量に関する情報に基づいて、前記第1バスの各伝送経路を流れるデータの流量を制御する第1制御部とを備えている。

20

【0025】

例示的なある実施形態において、前記第2制御部は、前記第1バスの各伝送経路のスループット、および前記第2バスのスループットを利用して、前記第2バスに流れる前記データの流量を制御する。

30

【0026】

例示的なある実施形態において、前記バスインタフェースの転送処理部は、前記第1ノードから受け取ったデータを、パケット化して前記第1バスの各伝送経路に転送し、前記所定の基準として、前記第1バスの各伝送経路の送信可能なパケット数の上限値の情報が予め定められており、前記ルータの割当部は、前記上限値の情報に従って、前記各伝送経路に送信可能なデータ送信量を割り当てる。

【0027】

例示的なある実施形態において、前記バスインタフェースの転送処理部は、前記第1ノードから受け取ったデータを、パケット化して前記第1バスの各伝送経路に転送し、前記ルータの割当部は、前記所定の基準として、前記第1バスの各伝送経路のスループットに従って、前記各伝送経路に送信可能なデータ送信量を割り当てる。

40

【0028】

例示的なある実施形態において、前記バスインタフェースの転送処理部は、前記第1ノードから受け取ったデータを、パケット化して前記第1バスの各伝送経路に転送し、前記ルータの割当部は、前記パケット数の上限値の情報に加えて、さらに前記第1バスの各伝送経路のスループットの情報を前記所定の基準として用いて、前記各伝送経路に送信可能なデータ送信量を割り当てる。

【0029】

例示的なある実施形態において、前記ルータの割当部は、前記第1バスの全ての伝送経

50

路のスループットに対する各伝送経路のスループットの割合に従って、前記各伝送経路に送信可能なデータ送信量を割り当てる。

【0030】

例示的なある実施形態において、前記第2ノードは、前記第1ノードにデータを送信することが可能であり、前記ルータ処理部は、前記第2ノードから受け取ったデータを、前記第1バスを介して前記第1ノードに転送する。

【0031】

例示的なある実施形態において、前記ルータ処理部は、前記第2ノードから受け取ったデータを、前記第1バスの複数の伝送経路を介して前記第1ノードに転送する。

【0032】

本発明の他の一態様であるバスシステムは、第1伝送レートを有する第1バス、および前記第1伝送レートよりも速い第2伝送レートを有する第2バスと、各々がデータを送信する複数の第1ノードと、前記複数の第1ノードの各々を前記第1バスに接続する複数のバスインタフェースと、前記第1バスおよび前記第2バスを接続するルータと、前記第2バスに接続され、前記データを受信する、少なくとも1つの第2ノードとを備えた半導体回路のバスシステムであって、前記第1バスは、前記バスインタフェースから前記ルータへ至る伝送経路を複数有する分散型バスであり、前記ルータは、所定の基準に従って前記第1バスの各伝送経路に送信可能なデータ送信量を割り当て、各伝送経路の送信量に関する前記情報を前記バスインタフェースに通知する割当部と、前記第1バスの各伝送経路を流れる前記データを受け取って前記第2バスに転送するルータ処理部と、前記第2バスに流れる前記データの流量を制御する第2制御部であって、前記複数の第1ノードから受け取った複数のデータを、第1ノード毎に保証すべき伝送レートを確保できるように集約して、前記第2ノードに伝送する、第2制御部とを備え、各バスインタフェースは、接続された第1ノードから受け取ったデータを、前記第1バスの各伝送経路に転送する転送処理部と、保証されるべき伝送レートに基づいて制限された伝送帯域、および、前記ルータから通知された、各伝送経路の送信量に関する情報に基づいて、前記第1バスの各伝送経路を流れるデータの流量を制御する第1制御部とを備えている。

【0033】

例示的なある実施形態において、保証されるべき伝送レートに基づいて制限された前記伝送帯域は、前記ルータ、または前記ルータとは異なるルータによって制限された伝送帯域である。

【0034】

本発明のさらに他の一態様であるバスシステムは、第1伝送レートを有する第1バス、前記第1伝送レートよりも遅い第2伝送レートを有する第2バス、および前記第2バスよりも速い第3バスと、データを送信する第1ノードと、前記第1ノードを前記第1バスに接続するバスインタフェースと、前記第1バスおよび前記第2バスを接続する第1ルータと、前記第2バスおよび前記第3バスを接続する第2ルータと、前記第3バスに接続され、前記データを受信する第2ノードとを備えた半導体回路のバスシステムであって、前記第1バスは、前記バスインタフェースから前記ルータへ至る伝送経路を複数有する分散型バスであり、前記第1ルータは、所定の基準に従って前記第1バスの各伝送経路に送信可能なデータ送信量を割り当て、各伝送経路の送信量に関する前記情報を前記バスインタフェースに通知する割当部と、前記第1バスの各伝送経路を流れる前記データを受け取って前記第2バスに転送するルータ処理部と、前記第2バスに流れる前記データの流量を制御する第2制御部とを備え、前記第2ルータは、前記第2バスの各伝送経路を流れる前記データを受け取って前記第3バスに転送するルータ処理部と、前記第3バスに流れる前記データの流量を制御する第2制御部とを備え、前記バスインタフェースは、前記第1ノードから受け取ったデータを、前記第1バスの各伝送経路に転送する転送処理部と、前記第1ルータから通知された、各伝送経路の送信量に関する情報に基づいて、前記第1バスの各伝送経路を流れるデータの流量を制御する第1制御部とを備え、ている。

【0035】



以下、添付の図面を参照しながら、本開示による伝送装置および伝送方法の実施形態を説明する。

【0036】

なお、以下の説明では、「バスマスタ」は、たとえばプロセッサや画像処理などの演算処理を行うノードである。バスマスタはトラヒックを送信する役割を有しているため、「送信ノード」と呼ぶことがある。また、「スレーブ」は、たとえばメモリ（メモリコントローラも含む）やI/Oである。スレーブはトラヒックを受信する役割を有しているため、「受信ノード」または「宛先ノード」と呼ぶことがある。本実施形態では、スレーブはメモリであるとして説明する。

【0037】

< 1 . システム構成 >

図3は、本実施形態による伝送システムの構成例を示す。図示されるシステムは、送信ノードとして8個のバスマスタ（BM）101を有し、受信ノードとして8個のメモリ105を有している。これら複数のバスマスタ101および複数のメモリ105は、12個のルータ600を介してバスで多段に接続されている。図3に示されるネットワーク構成は、いわゆるメッシュ網と呼ばれる。

【0038】

このように、図3に示す伝送システムは、多段接続網（Multistage Interconnection Network：MIN）を構成している。

【0039】

本実施形態におけるバスマスタ101は、例えばプロセッサやDSPなどの、バスを用いてデータの転送制御を行うことができるデバイスである。メモリ105は、例えばDRAMやSRAM等の半導体メモリである。なお、本開示の実施形態としては、メモリ105を、プロセッサ、DSP、I/Oなど、メモリ以外のノードを接続する形態で構成してもよい。ルータ600は、例えば半導体回路であり、複数のバスマスタ101および複数のメモリ105の間で伝送されるデータを中継する機能を備えている。

【0040】

バスマスタ101からメモリ105へのデータ伝送は、パケット交換方式により行われる。各バスマスタは、図3（a）に示すように、送信すべきパケットをフリットと呼ばれる最小単位に分割してから隣接のルータに送信する。1つのパケットを分割した複数のフリットのうち、最初に送信されるフリットは、ヘッダフリットと呼ばれる。ヘッダフリットには、パケットの先頭であることを示すフラグ情報や、パケットの送信先のアドレス情報などが記述されている。本実施形態におけるパケットおよびフリットのデータ構造の詳細は後述する。

【0041】

図3に示す多段接続網を構成する各ルータは、2入力2出力のクロスバスイッチを備えている。各ルータは、図3（b）に示すように、クロスバスイッチを切り替えることによって入力と出力の組合せを変更することができる。このため、トラヒックの流れを2つの伝送経路（伝送経路1、2）の間で切り替えることができる。出力先の伝送経路が異なっていれば、ルータは2つのトラヒックを各伝送経路に同時に出力することが可能である。このような多段接続網によれば、各ルータでのスイッチの切り替えによって、全バスマスタと全メモリ間で、必ず1つ以上の伝送経路を構築することができる。

【0042】

メッシュ網では、複数の送信ノードで伝送路を共有しつつ、必要に応じて、1つの送信ノードが複数の伝送路を使用して、広帯域化を図る通信方式を適用することが可能である。なお、本実施形態においてはメッシュ網を例に挙げて説明するが、これは一例である。集積回路の構成がその他のトポロジー（例えば、バタフライ網）であっても本開示を適用することが可能である。

【0043】

< 2 . パケット、フリットの構成 >

10

20

30

40

50

次に、本実施形態におけるパケットおよびフリットの構造を説明する。なお、本実施形態では、パケットまたはフリットを例として説明しているが、伝送に際してパケット化することは設計事項に過ぎない。パケットまたはフリットを一般化して、単に「データ」と呼んでもよい。

【0044】

図4は、パケット600の伝送フォーマットの一例と、パケット600を複数のフリットに分割した例とを示している。パケット600は、ヘッダフィールド601、データフィールド602、および制御コードフィールド603を有している。

【0045】

ヘッダフィールド601には、たとえば、送信先のアドレス、送信元のアドレス、パケットの送信されてからの経過時間を示す情報（経時情報）が記述される。経時情報の記述方法は任意であり、送信されてから経過した時間を判別できる値であればどのような記述方法であっても良い。例えば、送信された時刻が直接記載されてもよいし、送信されてからの経過時刻、また、現在までに通過したルータの数（ホップ数）などが記載されても良い。各ルータはパケット600を中継する度に、ホップ数をインクリメントしてもよいし、予め定められた最大ホップ数を順にデクリメントさせてもよい。

10

【0046】

ヘッダフィールド601には、上記以外の情報が記述されていてもよい。ヘッダフィールド601のデータのうち、送信先のアドレスおよび送信元のアドレスに基づいて、パケット600の中継処理と、受信側での受信処理とが行われる。

20

【0047】

データフィールド602には、例えば、映像データや音声データなどが記述される。制御コードフィールド603には、例えば、予め定められたパケット600の終了コードが記述される。ルータは、終了コードを検出することによってパケット600の末尾を判別できる。制御コードフィールド603には、終了コード以外の情報が格納されていてもよい。

【0048】

送信側のバスマスタ101は、前述のように、パケット600を、フリットと呼ばれる小さなパケット単位に分解して伝送する。1フリットのサイズは、バスを使って1サイクルで伝送可能なサイズであり、バス幅に応じて決定される。パケット600は、ヘッダフリット604、複数のデータフリット605、およびテイルフリット606に分割される。

30

【0049】

ヘッダフリット604には、上記のヘッダフィールド601に格納されたフラグ情報や送信先アドレス情報などが含まれる。

【0050】

ヘッダフリット604に続く各フリット（データフリット605、テイルフリット606）には、送信先を特定するアドレス情報は格納されていない。その理由は、ヘッダフリット604に続く各フリットは、ヘッダフリット604と同じ宛先に送られるためである。ヘッダフリット604によって宛先が決まり、そのトラヒックのフリットを出力する出力バッファ606が決まると、後に続くフリットは、ヘッダフリット604と同じ出力バッファ606を利用してヘッダフリット604が示す宛先に伝送される。

40

【0051】

テイルフリット606には、そのフリットがパケットを構成する最後のフリットであることを示すフラグ情報（制御コードフィールド603に格納されている終了コード）が付与されている。また、ヘッダフリット604とテイルフリット606以外の複数のフリット605は、主にデータを伝送するフリット（データフリット）であり、パケット600のデータフィールド602に対応する。

【0052】

受信側のメモリ105は、テイルフリット606に記述されたフラグ情報（終了コード

50

)を検出すると、その終了コードに基づき、伝送された複数のフリットをパケットへ再構築する。

【 0 0 5 3 】

1パケットのサイズは、例えば128バイトであり、1フリットのサイズは、例えば32ビットまたは64ビットに設定される。ただし、1パケットのサイズおよび1フリットのサイズはアプリケーションによって異なり得るため、上記のサイズは一例に過ぎない。1フリットの長さは、送信先のアドレスや送信元のアドレスなどの制御データを記述できる長さを基準としてもよい。

【 0 0 5 4 】

後述するように、各ルータは、送られてきたフリットを蓄えるためのバッファを備えている。フリットは、一旦そのバッファに蓄えられた後、スイッチの切り替えによって宛先のメモリへと続くルータ、または宛先のメモリへ送信される。

10

【 0 0 5 5 】

< 3 . ルータの構成および動作 >

< 3 . 1 . ルータの概略構成 >

図5は、ルータ600の構成を示す。図6は、ルータ600の動作手順を示すフローチャートである。

【 0 0 5 6 】

ルータ600は、複数の入力ポート613および出力ポート615を備えている。前段のルータ(群)600aから入力ポート613を経由してフリットを受け取り、出力ポート615を介して、後段のルータ(群)600bにフリットを送出する。

20

【 0 0 5 7 】

ルータ600は、入力バッファ203と、スイッチ204と、ルーティング処理部205と、バッファアロケータ部206と、スイッチアロケータ部207とを有している。以下、図6の各ステップを説明しながら、ルータ600の各構成要素の機能または動作を説明する。

【 0 0 5 8 】

まずルータ600は、入力ポート613において隣接ルータ(群)600aからのフリットを受信し、受信したフリットを入力バッファ203に格納する(S301)。ルーティング処理部205はフリットの転送先のノードを決定する(S302)。次に、バッファアロケータ部206は、隣接する後段のルータ(群)600bのどの入力バッファにフリットを格納するかを決定する(S303)。スイッチアロケータ部207は、入力バッファ203と転送先のルータのバッファ203とを対応付けてスイッチを設定する(S304)。スイッチアロケータ部207がスイッチを接続することにより、出力ポート615を介してフリットが転送先のルータ202のバッファへ送信される(S305)。

30

【 0 0 5 9 】

なお、図5の隣接ルータ(群)600aおよび600bに代えて、それぞれバスマスタおよび/またはメモリが設けられていてもよい。

【 0 0 6 0 】

図7および図8は、ローカルバスとシステムバスとの関係を具体的に示している。

40

【 0 0 6 1 】

図7に例示されるように、半導体バスシステムは、複数のローカルバス(ローカルバス1~3)を備えている。そして基幹となるシステムバス10が、それらローカルバスを相互に接続している。換言すると、1つのシステムバス10が、ローカルバス1~3を集約している。

【 0 0 6 2 】

図7はまた、システムバス10においてトラヒックの干渉が生じている例を模式的に示している。システムバス10がローカルバス1~3を集約しているため、システムバス10では、各ローカルバスからメモリへ向かうトラヒックが干渉することがある。図7に示される二種類の破線は、それぞれトラヒックを表している。共通のメモリへアクセスする

50

トラヒック同士がシステムバス10上で干渉するため、システムバス10では、伝送可能なトラヒック量は時間的に変動する。

【0063】

図8は、システムバス10に接続されているローカルバスの構成を示している。図3に関連して説明したように、ローカルバスは、3つのバスマスタ(BM1~BM3)がバスを共有できるよう、メッシュ型で構成されている。バスインタフェース(I/F)とも呼ばれるNIC(Network Interface Controller)は、バスマスタとルータとを接続するために設けられており、バスマスタから出力されるデータをパケット化してルータに送信し、ルータから受け取ったパケットを脱パケット化し、得られたデータをバスマスタに送信する。

10

【0064】

また、NICは並列的にデータを伝送するかどうかの判断と、伝送経路の伝送量を制御する。図8には、空き経路を活用したトラヒックの伝送例が模式的に示されている。図7の例では、バスマスタBM2から出力されたデータのトラヒックが3つの伝送経路で並列的に伝送されている。図8に示される三本の破線の各々は、別個のトラヒックを表している。なお、バスマスタBM2とNICとの間の伝送経路は3つのトラヒックを同時に伝送できる帯域が確保されているとする。

【0065】

ローカルバスは、ルータGWを介してシステムバス10と接続されている。ルータGWは、システムバスの伝送速度とローカルバスの伝送速度との違いを吸収するために、比較的大きい容量のバッファを備えている。ルータGWは速度変換ルータとも呼ばれることがある。

20

【0066】

各バスマスタには、主経路が予め定義されている。主経路とは、バスマスタが、システムバスとローカルバスとを相互に接続するルータGW(ゲートウェイ)までトラヒックを送信する際に、主として利用する伝送経路である。図8に示されるように、バスマスタBM1に関してはバスマスタBM1からルータGWまでの最短経路が主経路として定義されている。図の例では、全てのバスマスタの主経路は最短経路で定義されている。主経路は設計者が自由に設定することが可能である。

【0067】

並列伝送に使用する伝送経路の数や、ローカルバスやシステムバスでのトラヒック干渉などにより、伝送できるトラヒック量は時間的に変動する。特定のバスマスタがあまりに大量のトラヒックを送信すると、他のバスマスタが伝送できるトラヒック量が大幅に減少し、必要な処理ができなくなる。そこで、バスマスタ毎の伝送可能なデータ量(伝送量)が定義されている。すなわち、各バスマスタには予め「パケットの最大送信数」が定義されている。バスマスタは、パケットの最大送信数を上限とする数のパケットをネットワーク上に送信することができる。パケットの最大送信数が定義されることにより、1回の送信処理によって連続して送信可能なパケット数を制限することができる。

30

【0068】

パケットを送信したバスマスタは、送信したパケットに対する応答パケットをスレーブ(宛先のノード)から受信した場合には、受信した応答パケットの分だけパケットをさらに送信することができる。従って、伝送路上にはそのバスマスタから送信されたパケットが、最大でも当該最大送信数分しか伝送されないことになる。これにより、バスの伝送性能(スループットや遅延)を維持、保証することが可能となり、あるバスマスタが想定以上のパケットを送信することによるバスの混雑を回避することが可能となる。

40

【0069】

主経路以外の経路も利用してパケットを伝送するかどうかの判断や伝送量の判断に関しては、ルータGWが、各伝送経路からの伝送要求や伝送量に基づき一括して行えばよい。または、そのような判断を各NICが分散的に行ってもよい。そのときは、バスマスタに接続された各NIC間で、各伝送経路からの伝送要求や、伝送量に関する情報を交換すれ

50

ばよい。本実施形態においては、後に図 1 3 および図 1 4 を参照しながら、ルータ GW が一括して判断を行う例を説明する。

【 0 0 7 0 】

図 9 は、システムバス上のメモリとルータにおいてトラヒックの干渉が生じている例を模式的に示す。図 9 は、図 7 に示すトラヒックの干渉がルータに関しても、メモリに関しても発生している様子を示す。トラヒックの干渉は、各ロ カルバスから伝送されてきたトラヒック（破線）が交わっていること、または共通するメモリに送られていることによって示されている。

【 0 0 7 1 】

上述のとおり、システムバスのルータやメモリで伝送できるトラヒック量は時間的に変動する。そこで、伝送されるトラヒック（データ）の重要度の決定が行われる。ローカルバス 3 に存在するバスマスタは、処理優先度、データ種別、締切時刻などにより、伝送されるデータの重要度を決定する。この重要度に応じて、伝送を中継するルータやメモリにおける処理の優先順位が決定される。ルータおよびメモリで処理すべきトラヒック量に応じて処理時間が変化する。この結果、伝送できるトラヒック量は時間的に変動する。この状況はローカルバス上においても同様に発生し得る。たとえば、ローカルバス内にあるメモリやルータにより、ローカルバスで伝送できるトラヒック量は時間的に変動する。

【 0 0 7 2 】

図 1 0 は、本実施形態による半導体回路のバスシステム 1 0 0 の構成を示す。バスシステム 1 0 0 は、バスマスタと、N I C 8 0 3 と、アクセスルータと、ルータ G W 8 0 7 と、スレーブとを有している。アクセスルータと、ルータ G W 8 0 7 との間には、ローカルバスが設けられている。

【 0 0 7 3 】

N I C 8 0 3 は、バスマスタをローカルバスに接続するために設けられている。N I C 8 0 3 は、転送処理部 8 0 1 と送信流量制御部 8 0 2 とを有する。

【 0 0 7 4 】

転送処理部 8 0 1 は、バスマスタからのデータをパケット化し、スレーブからのパケットを脱パケット化する。

【 0 0 7 5 】

送信流量制御部 8 0 2 は、伝送経路の伝送量を制御する。この制御は、ローカルバス内で並列的にデータを伝送する際にも行われる（図 8 ）。

【 0 0 7 6 】

いま、N I C 8 0 3 に接続されるルータを「アクセスルータ」と定義し、ローカルバスとシステムバスとを接続するルータを「ゲートウェイ」（G W ）と定義する。以下では「ルータ G W 」と記述する。

【 0 0 7 7 】

ルータ G W はパケットの転送機能を備えたルータ処理部 8 0 4 と、システムバスへのパケットの送信量を制御する送信流量制御部 8 0 5 と、ローカルバスの各伝送経路の伝送量を決定する受信量割当部 8 0 6 とを有する。上述した機能以外の機能に関しては、以下で詳細に説明する。なお、ルータ処理部 8 0 4 は、図 5 で示したルータの構成（ 2 0 3 ~ 2 0 7 ）に対応する。また、本実施形態では、ルータ G W を除いては、アクセスルータも含めたルータは図 5 に示す一般的な構成を備えたとする。

【 0 0 7 8 】

図の例では、バスマスタからルータ G W を結ぶバスをローカルバスとし、ルータ G W からスレーブを結ぶバスをシステムバスとしている。一般的に、システムバスはローカルバスよりもバス動作周波数を高く設定する。

【 0 0 7 9 】

ここでは、説明を単純化するためにシステムバスは 1 本のみで記しているが、ローカルバスと同様に、システムバスが複数本存在して、同一のスレーブに伝送される形態であってもよい。つまり、M 対 N 型のトポロジでバスを構築してもよい。例えば、システムバ

10

20

30

40

50

スにある複数の伝送経路の最大送信数の算出方法をローカルバスの最大送信数の算出方法と同様にすれば実現は可能である。

【 0 0 8 0 】

さらに、図 1 0 では 1 つのスレーブが示されているが、複数のスレーブが設けられていてもよい。各スレーブに複数本のシステムバスが接続されていてもよい。たとえばスレーブであるメモリが複数存在してもよい。

【 0 0 8 1 】

なお、図 1 0 では、ルータ G W 8 0 7 とスレーブとがシステムバスによって直結されているように記載されている。しかしながら、スレーブもまた N I C を介してシステムバスに接続され得る。このとき、メモリコントローラがスレーブに対応する。メモリコントローラは N I C を介してシステムバスに接続されている。メモリへのアクセスはメモリコントローラを介して行われる。

10

【 0 0 8 2 】

図 1 1 は、往路での並列伝送時における N I C 8 0 3 の動作手順を示すフローチャートである。図 1 2 は、往路での並列伝送時における N I C 8 0 3 とルータ G W 8 0 7 との動作シーケンスを示す。

【 0 0 8 3 】

いま、ユースケース（例えば、実行されるアプリケーションの種類）が切り替わるタイミングに注目する。このタイミングで、たとえば異なるバスマスタが新たにデータの送信を開始する。新たに動作するバスマスタに接続された N I C 8 0 3 の送信流量制御部 8 0 2 は、ルータ G W 8 0 7 の受信量割当部 8 0 6 に対して各伝送経路に関するパケットの最大送信数を要求する（ S 9 0 1 ）。この動作により、各伝送経路において最大に送信できるパケット数を規定する。本実施形態では、この規定値はユースケース毎に予め設計段階で決められているとする。

20

【 0 0 8 4 】

ルータ G W 8 0 7 において、ルータ処理部 8 0 4 がパケットの最大送信数の要求を受け取ると、受信量割当部 8 0 6 は、ユースケースに応じた各バスマスタに対応した伝送経路毎の最大送信数を、自らが管理するテーブルから読み込む（ S 9 0 2 ）。管理している情報の例としては、図 1 3 および図 1 4 を例として挙げるができる。詳細は図 1 3 および図 1 4 を参照しながら説明する。

30

【 0 0 8 5 】

受信量割当部 8 0 6 は、各バスマスタに対応した伝送経路毎の通信量に関する情報、より具体的には最大送信数を特定する情報を、ルータ処理部 8 0 4 を介して N I C 8 0 3 の送信流量制御部 8 0 2 へ送信する（ S 9 0 3 ）。これにより N I C 8 0 3 は、伝送経路毎の最大送信数を特定することができる。 N I C 8 0 3 の送信流量制御部 8 0 2 は、ルータ G W 8 0 7 から通知された伝送経路毎の最大送信数の情報にしたがって、伝送経路の伝送量を制御する。

【 0 0 8 6 】

上述の動作の後に N I C 8 0 3 はバスマスタにデータの送信を許可する。その結果、バスマスタはデータ送信を開始する。以上の手順により、ユースケースに適したパケット送信数で各バスマスタはトラヒック伝送できるようになり、ローカルバス、システムバス、ルータ G W 8 0 7 などでの混雑を抑制することが可能となる。なお、システムバスには複数のルータ G W が接続され得る。

40

【 0 0 8 7 】

図 1 3 は、ローカルバスの伝送経路 1 ~ 3 を模式的に示し、図 1 4 は当該伝送経路 1 ~ 3 の各伝送量の割当規則を示す。

【 0 0 8 8 】

ルータ G W 8 0 7 の受信量割当部 8 0 6 で管理する最大送信数の例に関して説明する。図 1 3 および図 1 4 の例は、各バスマスタの最短経路が、各バスマスタが優先的に利用する主経路（例：バスマスタ B M 1 の主経路は経路 1 ）として設定されている。ユースケー

50

ス毎に、ルータGW807では各バスマスタの伝送経路毎に最大送信数が割り当てられている。

【0089】

例えば、BM1に関しては、5個の packets が経路1への packets の最大送信数として割り当てられている。BM2に関しては、3、5、4個の packets がそれぞれ経路1、2、3への packets の最大送信数として割り当てられている。一方、BM3に関しては、いずれの経路にも packets の最大送信数は割り当てられていない。なお、ゲートウェイで集中的に最大送信数を管理する方法を説明したが、各バスマスタが備えるNICで分散的に管理をしてもよい。

【0090】

図15は、往路での並列伝送時における、ルータGWの処理手順を示すフローチャートである。より具体的には、図15に示すフローチャートは、ルータGWの受信量割当部806によって実行される。また図16は、往路での並列伝送時におけるNICとルータGWとの動作シーケンスを示す。

【0091】

1つのバスマスタから特定のスレーブ(たとえばメモリ)へ、ローカルバスの複数の伝送経路を用いて packets を伝送する場合(往路)に、ローカルバスにおける伝送経路毎の伝送量の制御方法を説明する。

【0092】

ルータGW807の受信量割当部806は、ルータGW807からスレーブへの伝送部分であるシステムバスでのスループット(単位時間当たりの送信 packets 数)を計測する(S1201)。なお、スループットの計測は、ルータGW807で行ってもよいし、受信側であるスレーブで行ってもよい。また、実装を簡略化するためにスループットの値の計測は行わずに、システムバスで設計時に規定する最大送信数を、スループットとして固定値で与えてもよい。図16に示した動作シーケンスでは、システムバスへ転送している単位時間当たりの packets 数を「TS」で表している。

【0093】

次に、受信量割当部806は、ローカルバスの各伝送経路でのスループットの計測を行う(S1202)。ローカルバスのスループットの計測についても、バスマスタ側のNICで行ってもよいし、受信側であるゲートウェイで行ってもよい。図16では、ルータGW807の受信量割当部806において、ローカルバスの各伝送経路のスループットTH1~TH3を計測している。

【0094】

受信量割当部806は、システムバスでのスループットとローカルバスの伝送経路毎のスループットとに基づいて、ローカルバスの各伝送経路で転送可能な packets 数を算出する。具体的には、受信量割当部806は、システムバスでのスループットに対してローカルバスの各伝送経路でのスループットの比率を掛けて、ローカルバスの各伝送経路での packets の最大送信数を求める(S1203)。例えば、図16の例では、ローカルバスの1つの伝送経路での packets の最大送信数PL1を求める。図16のステップS1203に示されるように、 $PL1 = TH1 * TS / (TH1 + TH2 + TH3)$ で求めることができる。

【0095】

受信量割当部806は、各伝送経路で伝送される packets 数がローカルバスの各伝送経路で転送処理できる最大送信数以下に制限されるよう、バスマスタからルータGW807に伝送された packets に応答して、バスマスタに応答 packets を送信する応答処理を行う(S1204)。この応答処理により、バスマスタからの過剰な packets の伝送を抑制できる。

【0096】

以上の手順により、ローカルバスの各伝送経路の最大送信数を固定的に、設計時に決めておくよりは、動的に決定することで、ローカルバスやシステムバスでのトラヒック変動

10

20

30

40

50

に強くなる。これにより、ゲートウェイでの混雑の回避や、ローカルバスやシステムバスでの伝送効率（スループットとレイテンシ）を高めることが可能となる。なお、システムバスとローカルバスのスループットの計測は、独立かつ、定期的に行ってもよい。

【 0 0 9 7 】

図 1 7 および図 1 8 は、復路に対するシステムバスとローカルバスで生じる伝送帯域のギャップに対応する処理を示す。

【 0 0 9 8 】

いま、システムバスの伝送速度は相対的に大きく、ローカルバスの伝送速度は相対的に小さいという条件が与えられたとする。この条件では、システムバスとローカルバスとの間で伝送帯域のギャップが発生する。具体的には、この条件ではシステムバスからローカルにデータを伝送する復路において伝送帯域は狭くなるため、ローカルバスの入口で混雑が発生する。この混雑は、特定の伝送経路のみを用いてデータが伝送されることに起因している。

【 0 0 9 9 】

そこで、ローカルバスにおいて複数の伝送経路を用いてデータを伝送することで、ローカルバスのバス動作周波数を上げずに、ローカルバスの伝送速度をシステムバスの伝送速度まで高めることができる。以下では、復路において、伝送帯域のギャップをなくす目的で、復路の伝送路を複数用いる場合を説明する。

【 0 1 0 0 】

図 1 7 は、復路における伝送帯域のギャップに対応するための処理の手順を示す。図 1 8 は、3つの伝送経路 4 ~ 6 を用いてデータを伝送する様子を示す。本来 1 本であった伝送経路を 3 本に分けるため、この処理を本明細書では分割処理とも呼ぶことがある。

【 0 1 0 1 】

ユースケースの変化時点に、各バスマスタに接続された N I C 8 0 3 の送信流量制御部 8 0 2 は、ゲートウェイ G W 8 0 7 の受信量割当部 8 0 6 に対して、復路に関しては並列伝送を行うか否かの問い合わせを行う（ S 1 3 0 1 ）。この問い合わせは、並列伝送の実施の有無と、並列伝送を行う場合において使用する伝送経路に関するものである。

【 0 1 0 2 】

復路にて並列伝送を行うかどうかは、設計時にバスマスタが要求する伝送品質から決定しておく。また、復路において並列伝送を行う場合には、バスマスタ間でのトラヒック干渉が生じないように並列伝送に用いる伝送経路を決定しておく。なお、ゲートウェイ G W 8 0 7 で集中的に並列伝送を管理する方法を説明したが、各バスマスタが備える N I C で分散的に管理をしてもよい。

【 0 1 0 3 】

図 1 8 の例では、復路においてバスマスタ B M 2 のみが並列伝送を行い、他のバスマスタはデータ伝送を行わないとする。すると、バスマスタ B M 2 が伝送経路を 3 つ利用して並列伝送を行うことが可能である。

【 0 1 0 4 】

復路での伝送時には、並列伝送を行うバスマスタに対しては、ゲートウェイ G W 8 0 7 の送信流量制御部 8 0 5 は設計時に予め決めた伝送経路数にデータを分割して伝送を行う（ S 1 3 0 2 ）。図 1 8 の例では 3 つの伝送経路 4 ~ 6 で並列伝送を行うため、データを 3 分割している。実装を簡略化するために、ゲートウェイ G W 8 0 7 はシステムバスから受信したパケットを、対象となる伝送経路に対してラウンドロビンで均等になるように振り分けて伝送してもよい。

【 0 1 0 5 】

なお、図 1 8 に示す経路番号は、図 1 3 および図 1 4 に示す経路番号と異なっている。これは、往路と復路とが独立して伝送経路を決定できることを意味している。往路に伝送経路 1 で送信されたトラヒック、またはそのトラヒックに起因して発生した別のトラヒックが、復路でも同じ伝送経路 1 で伝送されるとは限られないことを意味している。

【 0 1 0 6 】

10

20

30

40

50



バスマスタBM2側のNIC803の送信流量制御部802は、分割されたデータを統合してバスマスタBM2に渡す(S1303)。図18の例では、復路での伝送経路の数である3つ分のトラヒックがNICで統合されて、バスマスタBM2へ伝送されている。NIC803の送信流量制御部802は、各伝送経路への振分け前のパケットの順序とNIC803で統合した後のパケットの順序が一致するようにパケットの順序制御をNICの送信流量制御部802で行う。例えば、送信流量制御部802は、分割した順に番号を割り当てて管理する。

【0107】

以上の手順により、復路において、ローカルバスとシステムバス間の伝送帯域のギャップの課題に対して、同一のローカルバス内のバスマスタ間でのトラヒック干渉を抑えながら、ローカルバス内での伝送を高速化することが可能となり、伝送帯域のギャップも抑えることができる。

10

【0108】

図19は、システムバスとローカルバスで生じる伝送帯域のギャップに対応する処理の手順を示すフローチャートである。図19は、プロセッサのデータを伝送する処理を示している。図20は、プロセッサと伝送経路との対応付けを示す図である。

【0109】

ローカルバスの各伝送経路と、プロセッサとの対応付けをしやすいようにするために、バスマスタから発行されたメモリのリード要求に対して、NIC803の送信流量制御部802は、送信元であるバスマスタを区別できる情報をパケットに付与する(S1401)。図20の例では、各バスマスタBMには、予めそのバスマスタBMに割り当てられた主経路が割り当てられている。これにより、伝送経路が独立され、トラヒック干渉を防ぐことが可能になる。

20

【0110】

各バスマスタからのトラヒックを、ゲートウェイGW807のルータ処理部804は保証レートの観点から集約し、送信流量制御部805は、システムバスで要求される伝送品質で、ゲートウェイGW807からデータを伝送する(S1402)。ゲートウェイGW807では、各バスマスタからの伝送量が保証レート(プロセッサが必要とするスループット)を確保できるように、送信流量制御部805が帯域制御を行う。実現方法として、例えば、固定レートで伝送レートを制御する方法でよい。ローカルバスで予め帯域制御を行うことで、システムバスでは、保証レートが高い1つの送信ノードからトラヒックが伝送されているように扱えるため、システムバスでのルータの性能保証のための実装が簡略化できる。

30

【0111】

たとえば図21は、3つのバスマスタであるプロセッサから伝送されるトラヒックと保証レートとを示す。

【0112】

図21に示すように、複数のバスマスタであるプロセッサからトラヒックが伝送される場合、各バスマスタにとって保証されるべき保証レートを超過してトラヒックが伝送される場合がある。この場合、全てのバスマスタの保証レートを保証するために、ローカルバスからシステムバスにデータを受け渡すルータGWの直前に配置されたルータ(ルータR1、R2、R3)は、各バスマスタから出力されるトラヒックに割り当てる帯域を制限する。図の例では、バスマスタBM1に対してはB1の伝送帯域が割り当てられ、バスマスタBM2に対してはB2の伝送帯域が割り当てられ、バスマスタBM3に対してはB3の伝送帯域が割り当てられる。帯域制限は、ルータGWによって設定された最大送信数に基づく帯域制限に加えて課される。なお、帯域制限に関しては、前述したようにルータGWにて一括で行ってもよい。

40

【0113】

そして、各プロセッサから送信されたトラヒックは、プロセッサクラスの1つのトラヒックとして集約されて、受信ノードに伝送される。図21に示されるように、このトラヒ

50

ックの伝送帯域は、 $B1 + B2 + B3$ である。

【0114】

なお、トラヒックには伝送の優先度に相当する「クラス」が設定され得る。たとえば、遅延の条件が最も厳しい「低遅延クラス」、その次に遅延の条件が厳しい「プロセッサクラス」、そして、上記2つのクラスよりも遅延の条件が緩やかな「グラフィックスクラス」である。たとえば図21のルータGWには、各クラスのデータを保持するためのバッファが別個設けられている。これらのバッファは、たとえば送信流量制御部805に設けられる。

【0115】

図21においては、各バスマスタに付随したNICが示されていないが、これは単に記載を省略しているに過ぎない。帯域制限されることにより、各NICからのパケットの出力は制限される。また、図21においては、各バスマスタからルータGWに至るローカルバスは特に分散型バスとして明示されていない。しかしながら、これは便宜的な記載である。バスマスタから、たとえば帯域を制限するルータRn (n:整数)までの伝送経路は分散型バスであってもよい。

10

【0116】

図の例では、要求品質の違いでルータのバッファを用意すればよい。要求品質に応じて、3つのクラス(「低遅延」、「プロセッサ」、「グラフィックス」)が用意されている。ルータは、「低遅延」、「プロセッサ」、「グラフィックス」の順に優先的にデータを送信する。

20

【0117】

なお、保証レートを超えたトラヒックは、全てのバスマスタの保証レート分のトラヒックを伝送していない期間でトラヒックを伝送する。これにより、システムバス側のルータはプロセッサの数(バスマスタの数)だけ、ルータのバッファを備える必要はなくなるため、ルータの実装を簡略化することが可能となる。

【0118】

図22は、クラスに応じた調停順位を示す。保証レート以下の範囲内においては、上述の規則に従い、「低遅延」、「プロセッサ」、「グラフィックス」の調停順位にしたがってデータが伝送される(調停順位(1)~(3))。そして保証レートを超過した場合には、予め定められた上限(閾値)未満までであれば、「プロセッサ」、「グラフィックス」の順でデータが伝送される(調停順位(4)、(5))。予め定められた上限(閾値)以上であれば、さらに空き帯域がある場合にデータが伝送される(調停順位(6))。この調停順位に従えば、プロセッサクラスのデータに関しては、調停順位に相違が生じるものの、保証レートには上限がないと言える。上記の説明では、プロセッサクラスには保証レートの上限がないとしたが、同様にグラフィックスクラスに対しても、この調停方法を適用してもよい。

30

【0119】

再び図19を参照する。

【0120】

メモリコントローラがリード要求を解釈し、要求に対応するデータをメモリから読み出し、往路のシステムバスと往路のローカルバスを介してバスマスタへデータを送信する(S1403)。

40

【0121】

ルータGW807の送信流量制御部805は、システムバスを介してメモリから伝送されてきたデータを、ローカルバスの伝送経路に振り分けて、そのデータのリード要求を発行したバスマスタ(プロセッサ)へ伝送する(S1404)。

【0122】

メモリからバスマスタに伝送する復路の伝送では、メモリから伝送されるトラヒックは中継するルータでブロックされることなく、伝送する必要がある。このため、本実施形態では、復路の伝送経路は、往路の伝送経路の逆を辿るよう設定される。これは、たとえば

50

ゲートウェイ 807 の送信流量制御部 805 が、各バスマスタからの要求をそれぞれ識別可能な情報（たとえばパケット ID）を保持しておき、そのパケット ID とローカルバスの経路とを対応付けて情報を保持しておくことにより、実現される。リード要求に応答して読み出され、メモリから送信されたデータにも同じパケット ID が付加されることにより、ゲートウェイ 807 の送信流量制御部 805 は復路のデータを、往路と同じ経路で、そのデータのリード要求を発行したバスマスタに伝送することができる。復路の伝送はプロセッサ毎に独立して行われる。これにより、復路でのトラフィック干渉を抑制することができる（図 20）。

#### 【0123】

本実施形態では、具体例としてプロセッサを対象に説明したが、グラフィックスや他の同様な特性を備えたバスマスタであっても同様に実現することができる。他の図の例の説明に関しても、用途はプロセッサに限定されない。また、伝送経路は、図で示したように直線的な経路を取る必要は無い。予め経路を静的に決定する方法や動的に経路を決定する方法を用いて任意の伝送経路で実現してもよい。同様に、他の図の例の説明に関しても、任意の伝送経路で実現してもよい。上述の例では、読み込み処理を例にして示したが、書き込みの場合も同様に実現することができる。

#### 【0124】

図 23 は、往路における保証レートを超えた伝送処理の手順を示す。

#### 【0125】

保証レートを超えた伝送とは、例えば、保証レート以内のパケット伝送は主経路のみを使い、保証レートを超えたパケット伝送は、迂回路も用いて伝送することを示す。保証レートとは、バスマスタの必要最小限の伝送帯域を表し、保証レートを超えて伝送することで、バスマスタでの処理を円滑にし、パケット伝送の効率（スループットや遅延）もより一層高めることができる。

#### 【0126】

バスマスタ側の NIC 803 は、ゲートウェイ 807 に保証レートを超えた伝送の要求を発行する（S1501）。

#### 【0127】

ゲートウェイ 807 の受信量割当部 806 は、バスマスタからゲートウェイ 807 までの各伝送経路の伝送量を計測する（S1502）。

#### 【0128】

受信量割当部 806 は、計測した伝送量に基づいて空き帯域を明らかにし、空き帯域に応じて主経路以外にも使用できる伝送経路を抽出する（S1503）。例えば受信量割当部 806 は、50% の空き帯域があれば主経路以外にも伝送できる経路であると認定する。但し、主経路は 50% の空き帯域がなくても伝送できるとする。

#### 【0129】

ゲートウェイ 807 の受信量割当部 806 は、使用するローカルバスの伝送経路（主経路と空き帯域のある伝送経路）を決定し、ルータ処理部 804 を介して対象となる NIC に通知する（S1504）。

#### 【0130】

バスマスタ側の NIC 803 の送信流量制御部 802 は、決定された伝送経路に基づき各伝送経路で伝送できる伝送レートの上限を設定する（S1505）。

#### 【0131】

送信流量制御部 802 は、主経路には保証レート内のトラフィックを伝送し、他の空き帯域のある伝送路には保証レートを超えたトラフィックを伝送する（S1506）。

#### 【0132】

図 24 は、データが、主経路と、空き帯域のある伝送経路とを用いて伝送される様子を示す。伝送対象は、バスマスタ BM1 から伝送される、プロセッサクラスの要求である。データの一部は、保証レート B1 で主経路で伝送される。データの残りの一部は、保証レート B1 では伝送しきれずに残されたデータであり、空き帯域が存在する他の経路で伝送

10

20

30

40

50

される。

【 0 1 3 3 】

図 2 5 の ( i ) は主経路の帯域の割当状況を示し、図 2 5 の ( i i ) は、空き帯域が存在する、主経路以外の伝送経路の帯域の割当状況を示す。図 2 5 の ( i ) に示されるように、主経路で伝送する保証レート内のトラヒックに対しては、予め設計時に伝送帯域が確保される。一方、図 2 5 の ( i i ) に示されるように、主経路以外の伝送経路は、主経路として伝送されている / 伝送される可能性のある伝送帯域 ( 設計時に確保されている帯域 ) と、主経路ではないが伝送している伝送量の両方を、伝送経路が伝送できる最大の帯域から減算して空き帯域を求める。なお、保証レートを越えたトラヒックに対しては、各バスマスタが各伝送経路の帯域を予約 / 解放の手続きをゲートウェイとバスマスタ間で行う方法で実現してもよい。これにより、空き帯域の計測処理は不要になり、ハード実装は簡略化できる。

10

【 0 1 3 4 】

主経路で伝送する保証レート内のトラヒックは高優先度で、主経路以外で伝送される保証レートを越えたトラヒックは、同じ伝送路で伝送される保証レート内のトラヒックよりも低優先度で、バスマスタからゲートウェイまでを中継するルータにおいて、転送処理されるとする。従って、保証レートを越えたトラヒックは、保証レート内のトラヒックよりも低い優先度で処理されてもよい。このようなトラヒックは、例えば、プロセッサを対象としたアプリケーションで、高優先度に対応付けられたメディア処理のアプリケーションは保証レート内の伝送に対応付け、低優先度に対応付けられた W e b やメールのアプリケーションは保証レートを越えた伝送に対応付ける。このようにプロセッサの処理ではアプリケーションにより要求性能が大きく異なる。

20

【 0 1 3 5 】

以上の構成により、ローカルバスのバス動作周波数を低く抑えながら、バスの混雑を回避しながら、システムバスと同様な高速なトラヒック伝送を実現することができる。

【 0 1 3 6 】

本開示は、チップ内のバス配線だけではなく、チップ間のバス配線に対しても適用が可能である。図 2 6 は、複数の S o C を相互接続する場合に適用する場合の例を示す。S o C 間をつなぐ配線は、プリント基板上に実装されるため、チップ内で配線される場合に比べて、高速にデータを伝送することが難しい。このため、S o C 間をつなぐ配線部分 Z に対して伝送経路の並列化を図ることで、バス動作周波数を大きく上げずにバス伝送の高速化が可能となる。

30

【 0 1 3 7 】

また、図 2 6 のように、速度差がある 3 つ以上のバス ( S o C 1 内のバス、S o C 2 内のバス、S o C 1 と S o C 2 とを接続するバス ) で構成される場合でも適用してもよい。

【 0 1 3 8 】

なお図 2 6 においても、N I C は記載されていないが、単に記載を省略したに過ぎない。各 S o C は、図 1 0 に示すバスマスタからルータ G W に相当する構成を有している。そして各 S o C のルータ G W 同士が、S o C 間を接続するバスを介して接続されている。

【 産業上の利用可能性 】

40

【 0 1 3 9 】

本開示にかかる半導体回路のバスシステムは、システム L S I 上のデータ伝送に利用可能である。

【 符号の説明 】

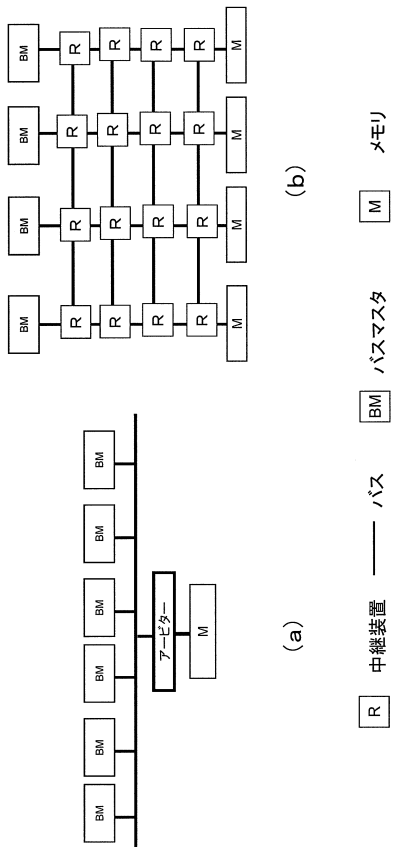
【 0 1 4 0 】

- 1 0 0 半導体回路のバスシステム
- 8 0 1 転送処理部
- 8 0 2 送信流量制御部
- 8 0 3 ネットワークインタフェースコントローラ ( N I C )
- 8 0 4 ルータ処理部

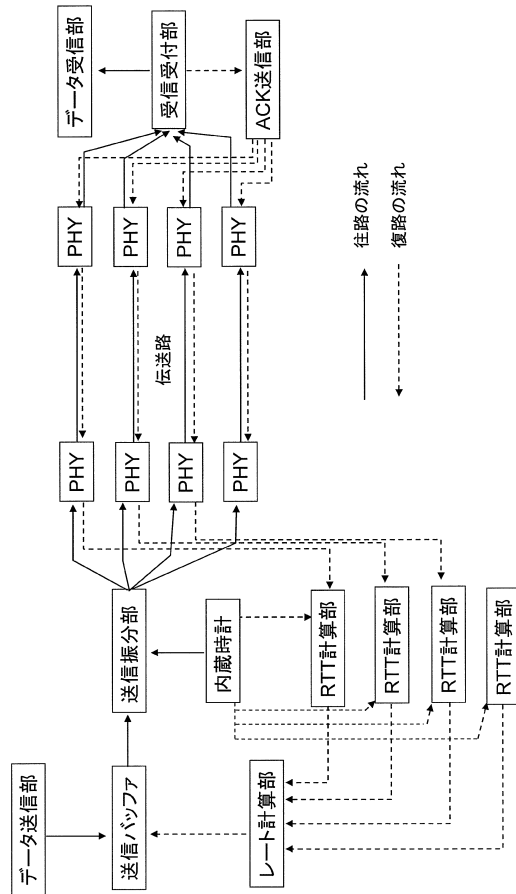
50

- 805 送信流量制御部
- 806 受信料割当部
- 807 ルータGW

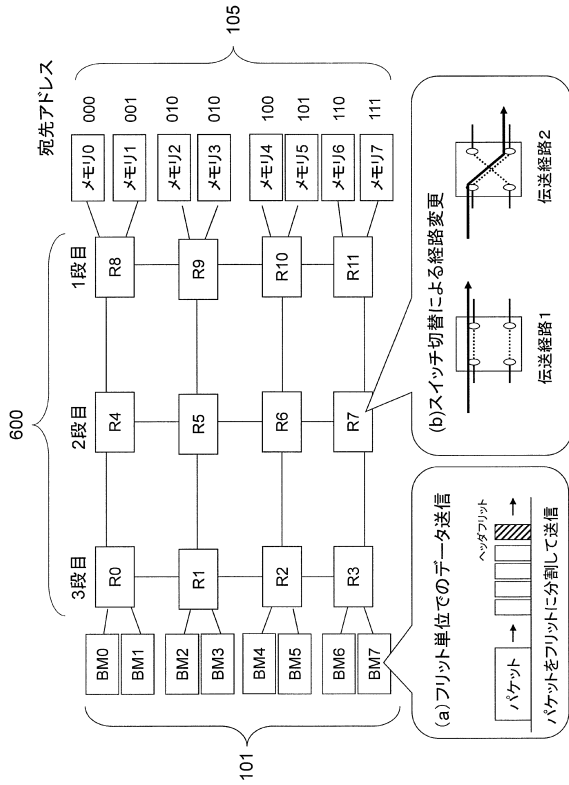
【図1】



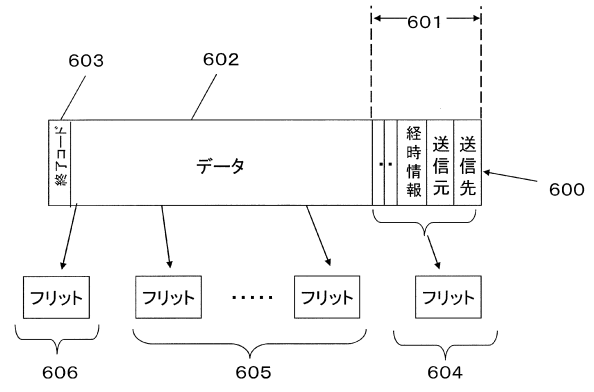
【図2】



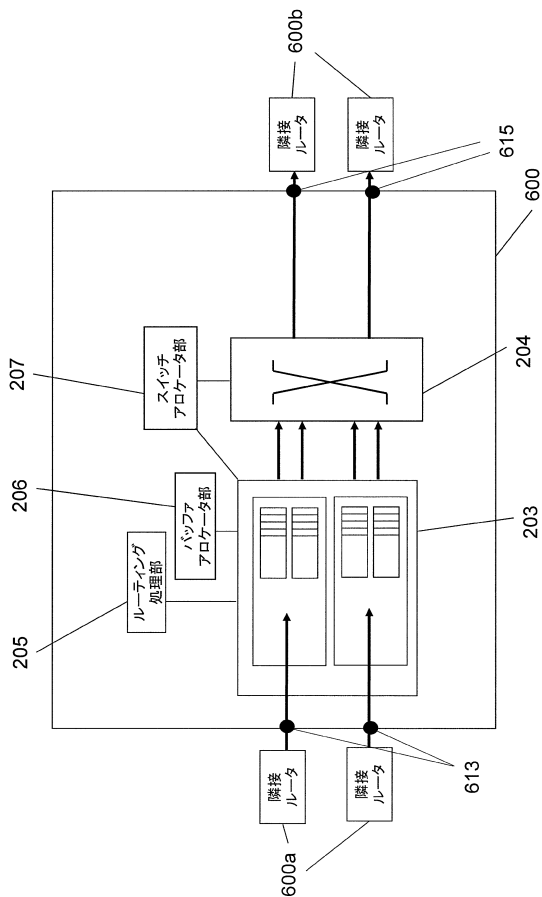
【図3】



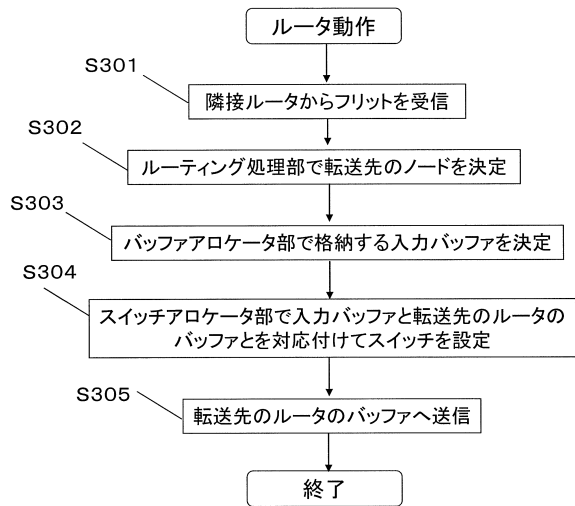
【図4】



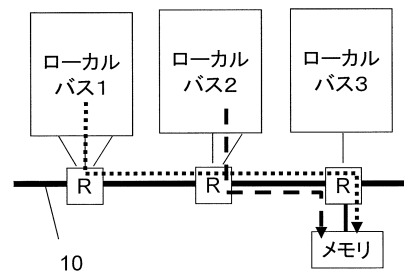
【図5】



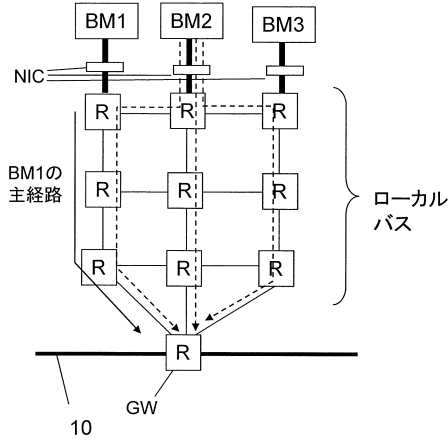
【図6】



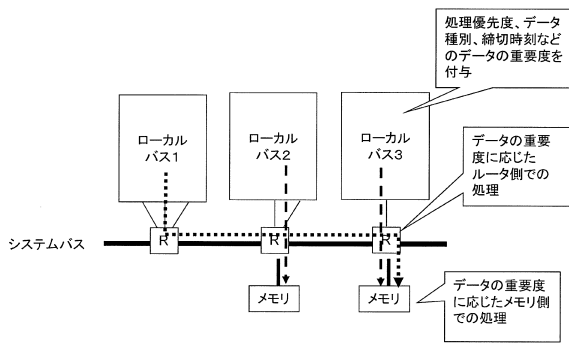
【図7】



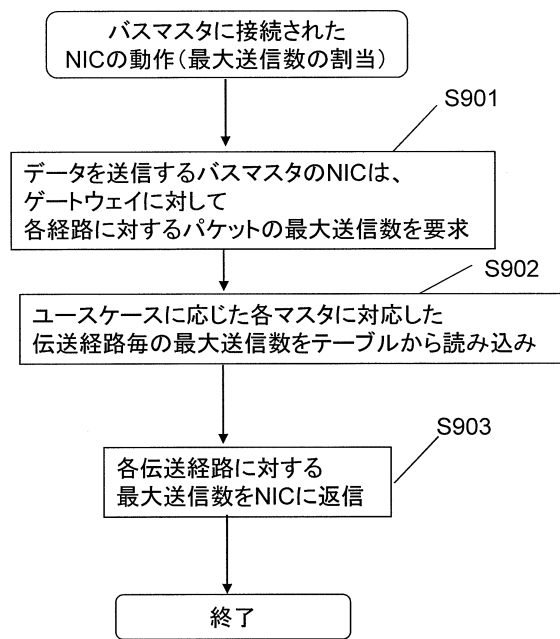
【図8】



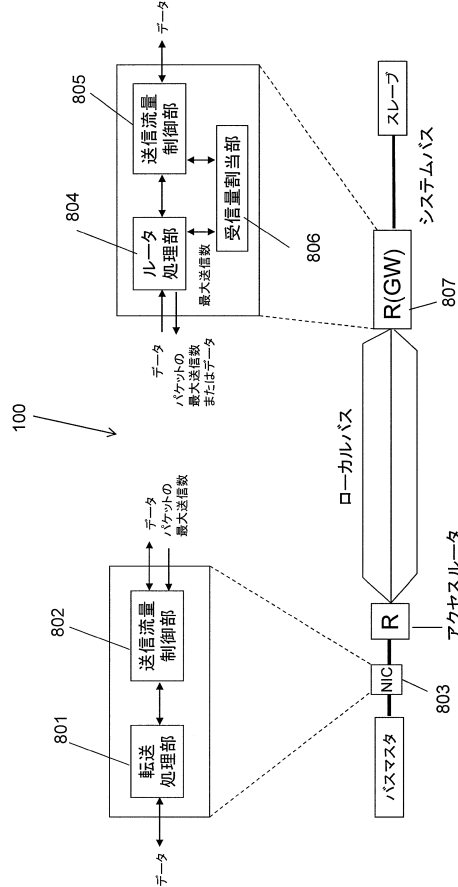
【図9】



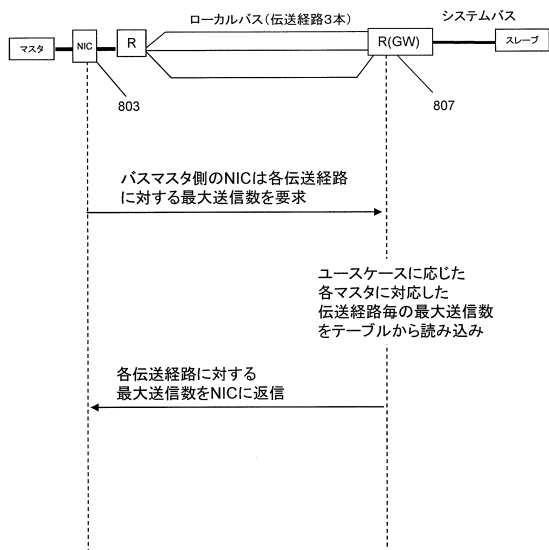
【図11】



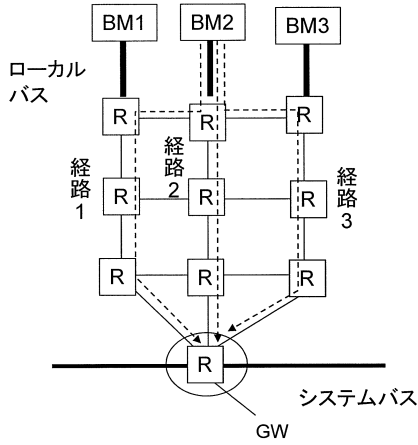
【図10】



【図12】



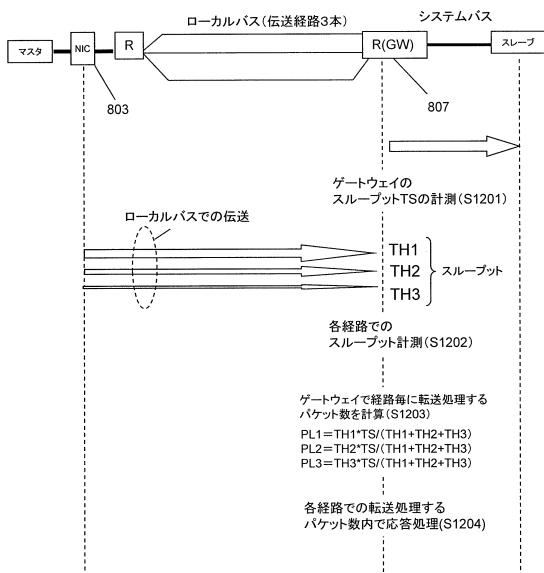
【図13】



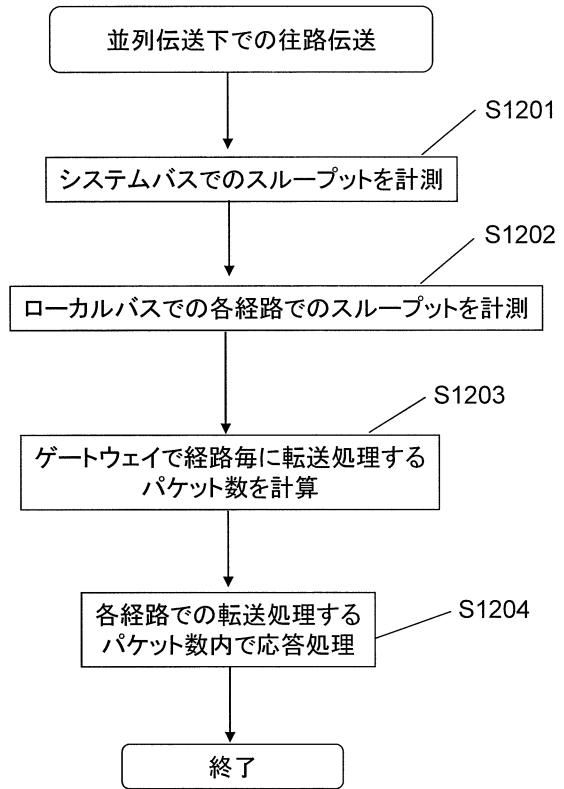
【図14】

	経路1	経路2	経路3
BM1	5(主経路)	0	0
BM2	3	5(主経路)	4
BM3	0	0	0(主経路)

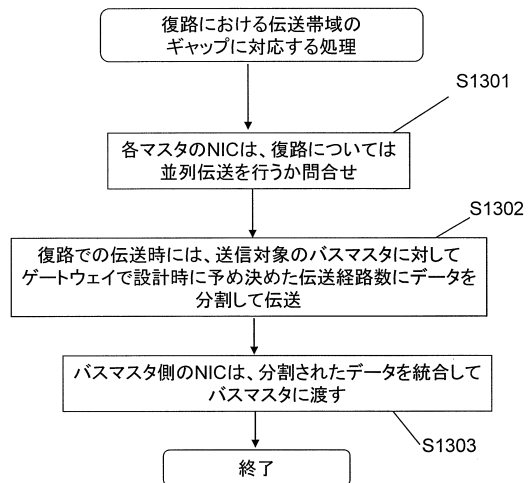
【図16】



【図15】

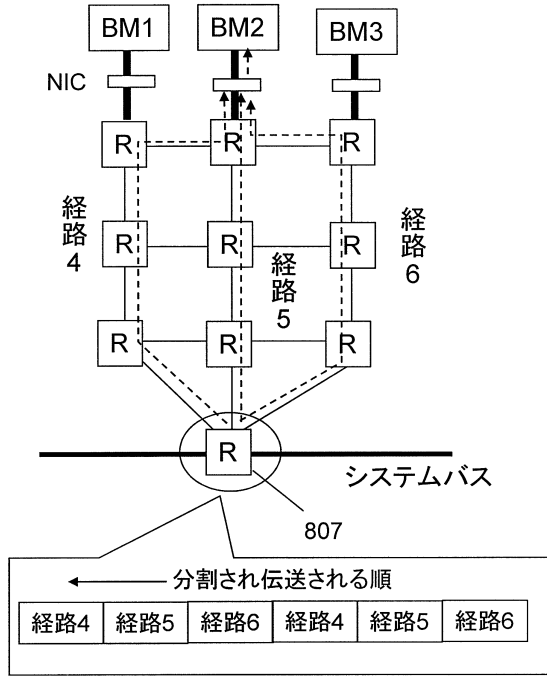


【図17】

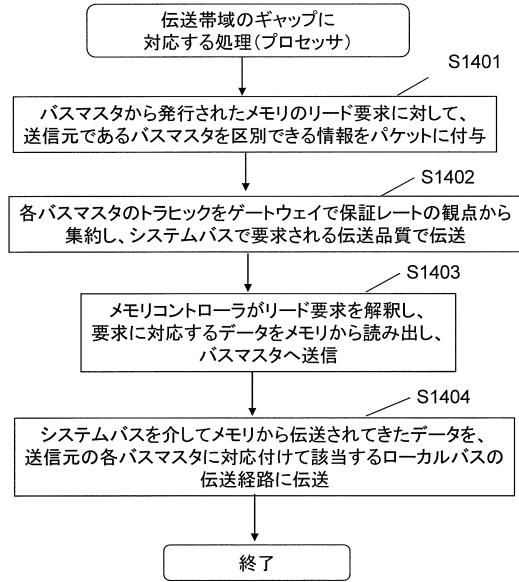




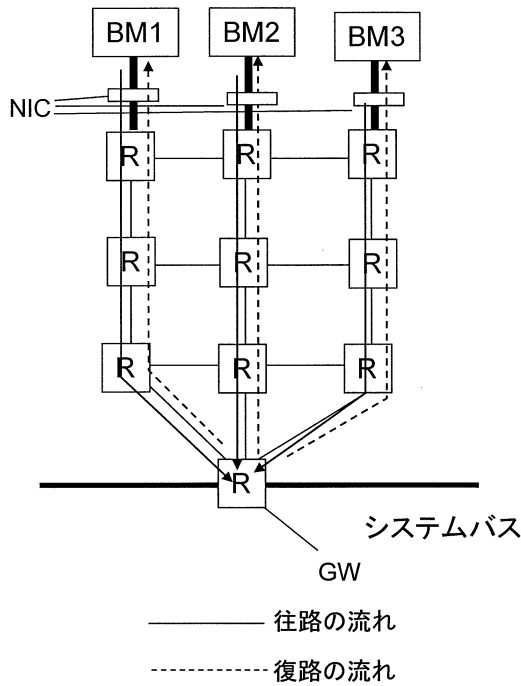
【図18】



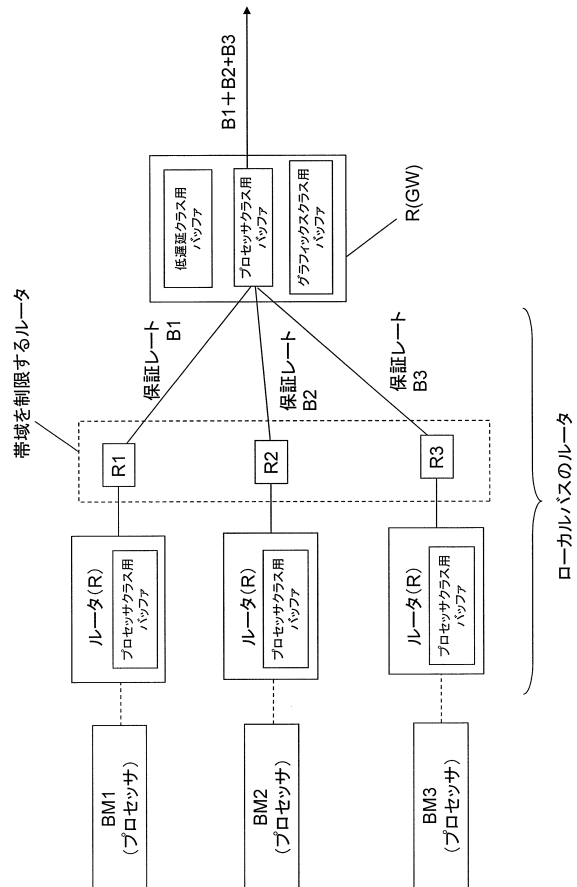
【図19】



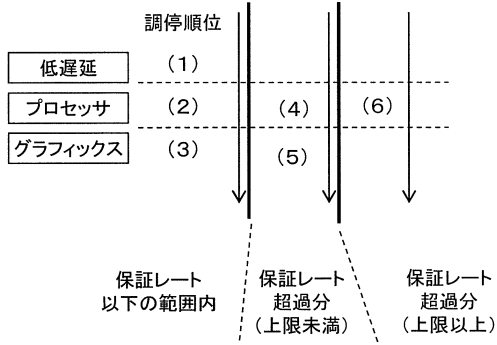
【図20】



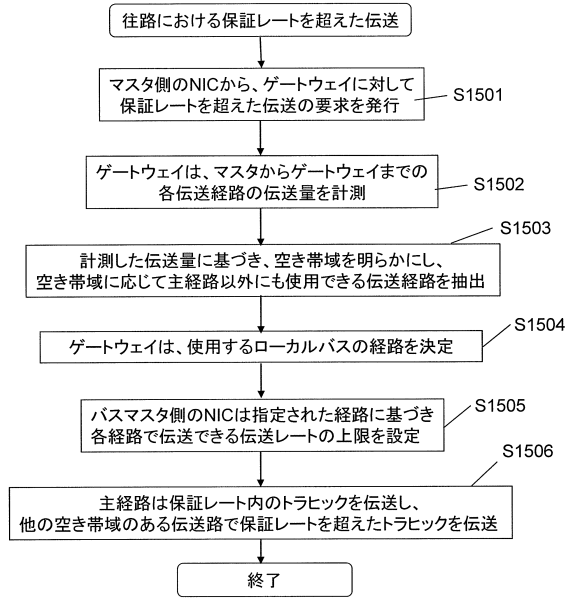
【図21】



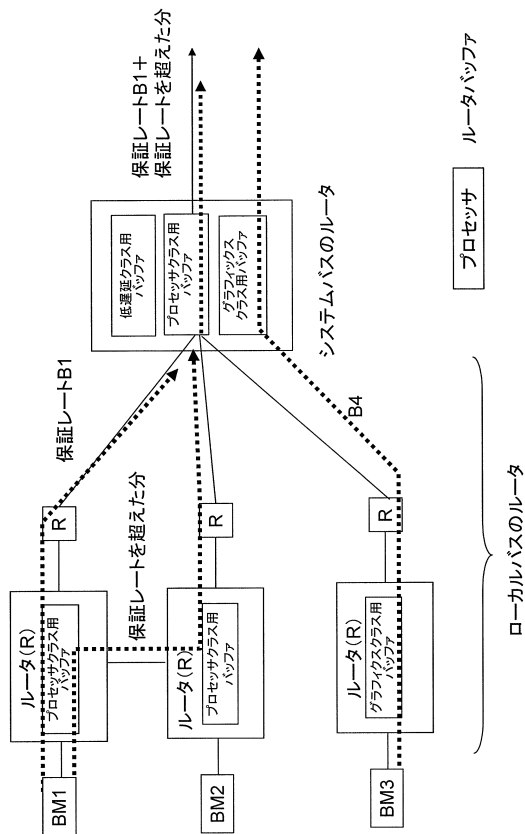
【図 2 2】



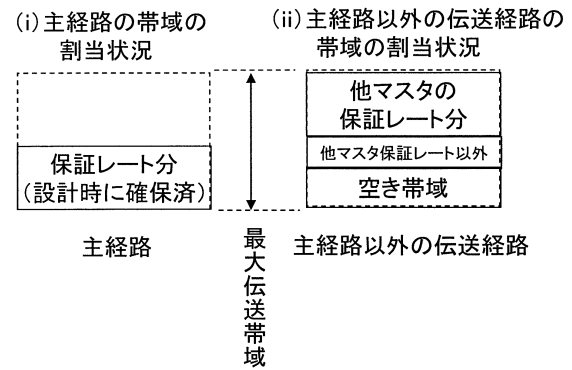
【図 2 3】



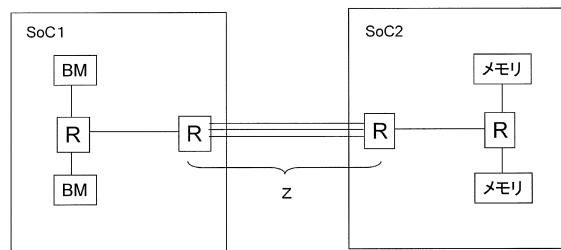
【図 2 4】



【図 2 5】



【図 2 6】



## フロントページの続き

- (72)発明者 山口 孝雄  
大阪府門真市大字門真1006番地 パナソニック株式会社内
- (72)発明者 吉田 篤  
大阪府門真市大字門真1006番地 パナソニック株式会社内
- (72)発明者 石井 友規  
大阪府門真市大字門真1006番地 パナソニック株式会社内
- (72)発明者 得津 覚  
大阪府門真市大字門真1006番地 パナソニック株式会社内

審査官 安藤 一道

- (56)参考文献 特開2009-118489(JP,A)  
国際公開第2012/132263(WO,A1)  
国際公開第2011/004566(WO,A1)  
吉田篤、他2名、分散型バスにおける伝送負荷均一化方式の提案、情報処理学会研究報告、情報処理学会、2011年7月20日、Vol.2011-ARC-196 No.7, pp.1-8

(58)調査した分野(Int.Cl., DB名)

H04L 12/40  
H04L 12/701  
H04L 12/911