(19)日本国特許庁(JP)　(12)公開特許公報(A)　(11)特許出願公開番号
特開2019-197319
(P2019-197319A)
(43)公開日　令和1年11月14日(2019.11.14)

(51)Int.Cl.		F I			テーマコード（参考）
G06F 9/50	(2006.01)	G06F	9/50	150D	5B042
G06F 11/34	(2006.01)	G06F	11/34	133	
G06F 13/14	(2006.01)	G06F	11/34	119	
		G06F	13/14	310F	

審査請求 未請求　請求項の数 8　OL　（全 15 頁）

(21)出願番号　特願2018-90015(P2018-90015)
(22)出願日　平成30年5月8日(2018.5.8)

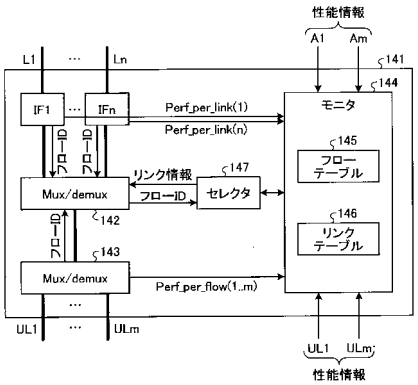
(71)出願人　000005223
富士通株式会社
神奈川県川崎市中原区上小田中４丁目１番１号
(74)代理人　110002147
特許業務法人酒井国際特許事務所
(72)発明者　三吉　貴史
神奈川県川崎市中原区上小田中４丁目１番１号　富士通株式会社内

Ｆターム(参考)　5B042 JJ20　MA14　MC33

(54)【発明の名称】情報処理装置、情報処理方法および情報処理プログラム

(57)【要約】
【課題】システム全体の性能を最大化するようにフローを振り分けることができる情報処理装置、情報処理方法および情報処理プログラムを提供する。
【解決手段】情報処理装置は、複数のリンクを経由してＣＰＵと接続されるオフロード回路を備える。情報処理装置のオフロード回路は、論理回路と、収集部と、選択部とを備える。論理回路は、アプリケーションの処理を演算する。収集部は、アプリケーションの処理に対応するフローごとのリンクの性能情報を示す値と、リンクごとの使用可能な性能情報の最大値とを収集する。選択部は、フローごとのリンクの性能情報を示す値に基づいて、要する性能情報を満たしていないフローを判定する。また、選択部は、リンクごとの使用可能な性能情報の最大値と、リンクごとの現在使用されている性能情報の値とに基づいて、フローの振り分け先のリンクを選択して振り分ける。
【選択図】図４

制御回路の機能構成の一例を示すブロック図



【特許請求の範囲】
【請求項１】
複数のリンクを経由してＣＰＵと接続されるオフロード回路において、
アプリケーションの処理を演算する論理回路と、
前記アプリケーションの処理に対応するフローごとの前記リンクの性能情報を示す値と、前記リンクごとの使用可能な性能情報の最大値とを収集する収集部と、
前記フローごとの前記リンクの性能情報を示す値に基づいて、要する性能情報を満たしていないフローを判定し、前記リンクごとの使用可能な性能情報の最大値と、前記リンクごとの現在使用されている性能情報の値とに基づいて、前記フローの振り分け先のリンクを選択して振り分ける選択部と、
を備えることを特徴とする情報処理装置。
【請求項２】
前記収集部は、前記アプリケーションの処理に対応するフローごとの前記リンクの使用帯域を示す値と、前記リンクごとの使用可能な帯域の最大値とを収集し、
前記選択部は、前記フローごとの前記リンクの使用帯域を示す値に基づいて、要する帯域を満たしていないフローを判定し、前記リンクごとの使用可能な帯域の最大値と、前記リンクごとの現在使用されている帯域の値とに基づいて、前記フローの振り分け先のリンクを選択して振り分ける、
ことを特徴とする請求項１に記載の情報処理装置。
【請求項３】
前記収集部は、前記アプリケーションの処理、または、前記論理回路の性能が性能要件を満たしていない場合、対応する前記フローの優先度を上げる、
ことを特徴とする請求項２に記載の情報処理装置。
【請求項４】
前記選択部は、前記リンクごとの使用可能な帯域の最大値と、前記リンクごとの現在使用されている帯域の値とに基づいて、前記フローの振り分け先のリンクを選択し、選択した前記振り分け先のリンクに該フローを振り分けると、前記振り分け先のリンクの使用可能な帯域の最大値を超える場合、前記振り分け先のリンクを使用する最も優先度が低いフローを、振り分け元のリンクに振り分ける、
ことを特徴とする請求項２または３に記載の情報処理装置。
【請求項５】
前記収集部は、所定時間ごとに、前記リンクの使用帯域を示す値と、前記リンクごとの使用可能な帯域の最大値とを収集し、
前記選択部は、前記所定時間ごとに、前記要する帯域を満たしていないフローを判定する、
ことを特徴とする請求項２～４のいずれか１つに記載の情報処理装置。
【請求項６】
前記収集部は、前記アプリケーションの処理に対応するフローごとの前記リンクのレイテンシを示す値と、前記リンクごとのレイテンシの最大値とを収集し、
前記選択部は、前記フローごとの前記リンクのレイテンシを示す値に基づいて、要するレイテンシを満たしていないフローを判定し、前記リンクごとのレイテンシの最大値と、前記リンクごとの現在のレイテンシの値とに基づいて、前記フローの振り分け先のリンクを選択して振り分ける、
ことを特徴とする請求項１に記載の情報処理装置。
【請求項７】
複数のリンクを経由してＣＰＵと接続され、アプリケーションの処理を演算する論理回路を備えるオフロード回路において、
前記アプリケーションの処理に対応するフローごとの前記リンクの性能情報を示す値と、前記リンクごとの使用可能な性能情報の最大値とを収集し、
前記フローごとの前記リンクの性能情報を示す値に基づいて、要する性能情報を満たし

ていないフローを判定し、前記リンクごとの使用可能な性能情報の最大値と、前記リンクごとの現在使用されている性能情報の値とに基づいて、前記フローの振り分け先のリンクを選択して振り分ける、
ことを特徴とする情報処理方法。
【請求項８】
複数のリンクを経由してＣＰＵと接続され、アプリケーションの処理を演算する論理回路を備えるオフロード回路において、
前記アプリケーションの処理に対応するフローごとの前記リンクの性能情報を示す値と、前記リンクごとの使用可能な性能情報の最大値とを収集し、
前記フローごとの前記リンクの性能情報を示す値に基づいて、要する性能情報を満たしていないフローを判定し、前記リンクごとの使用可能な性能情報の最大値と、前記リンクごとの現在使用されている性能情報の値とに基づいて、前記フローの振り分け先のリンクを選択して振り分ける、
処理をコンピュータに実行させることを特徴とする情報処理プログラム。
【発明の詳細な説明】
【技術分野】
【０００１】
本発明は、情報処理装置、情報処理方法および情報処理プログラムに関する。
【背景技術】
【０００２】
近年、ＣＰＵ（Central Processing Unit）で動作するアプリケーションを、高速化のためにＦＰＧＡ（Field Programmable Gate Array）等のハードウェアにオフロードして動作させることが提案されている。オフロードしたアプリケーションの一部は、ユーザロジック（オフロード回路）と呼ばれ、１つのＦＰＧＡに複数のユーザロジックを搭載する場合がある。複数のユーザロジックは、それぞれＣＰＵで動作するアプリケーションに応じて、任意の組み合わせで動作することが想定されている。ＣＰＵとＦＰＧＡとの間は、例えば、ＰＣＩ Ｅｘｐｒｅｓｓ（以下、ＰＣＩｅともいう。）等のＩ／Ｏバス経由での接続と、ＯｐｅｎＣＡＰＩ（Open Coherent Accelerator Processor Interface）等のメモリコヒーレントバス経由での接続とがある。
【０００３】
Ｉ／Ｏバス経由の接続は、アクセス遅延が大きく、ＣＰＵでキャッシュできないが、帯域は中から広帯域であり、ＦＰＧＡ側からのＤＭＡ（Direct Memory Access）転送に向く。一方、メモリコヒーレントバス経由の接続は、アクセス遅延が小さく、ＣＰＵでキャッシュできるが、帯域は小から中帯域であり、ＦＰＧＡ側からのメモリアクセスが可能である。ＣＰＵとＦＰＧＡとの間の接続（リンク）は、Ｉ／Ｏバス経由の接続と、メモリコヒーレントバス経由の接続との両方が用意される場合がある。この場合、ユーザロジックは、どのリンクを使用するかを事前に固定的に選択する。
【先行技術文献】
【特許文献】
【０００４】
【特許文献１】特開２０１４－１７０３６３号公報
【特許文献２】特表２００８－５４６０７２号公報
【発明の概要】
【発明が解決しようとする課題】
【０００５】
しかしながら、ユーザロジックは、他のユーザロジックがどのリンクを使用するのかが運用時まで判らない。このため、例えば、ユーザロジック、コマンド種別およびアドレス範囲の組み合わせを表すフローを単位として各リンクに振り分ける場合、フローが要求する帯域に対してリンクの帯域が不足する等により、アプリケーションの性能要求を満たさない場合がある。

【０００６】
一つの側面では、システム全体の性能を最大化するようにフローを振り分けることができる情報処理装置、情報処理方法および情報処理プログラムを提供することにある。
【課題を解決するための手段】
【０００７】
一つの態様では、情報処理装置は、複数のリンクを経由してＣＰＵと接続されるオフロード回路を備える。情報処理装置のオフロード回路は、論理回路と、収集部と、選択部とを備える。論理回路は、アプリケーションの処理を演算する。収集部は、前記アプリケーションの処理に対応するフローごとの前記リンクの性能情報を示す値と、前記リンクごとの使用可能な性能情報の最大値とを収集する。選択部は、前記フローごとの前記リンクの性能情報を示す値に基づいて、要する性能情報を満たしていないフローを判定する。また、選択部は、前記リンクごとの使用可能な性能情報の最大値と、前記リンクごとの現在使用されている性能情報の値とに基づいて、前記フローの振り分け先のリンクを選択して振り分ける。
【発明の効果】
【０００８】
システム全体の性能を最大化するようにフローを振り分けることができる。
【図面の簡単な説明】
【０００９】
【図１】図１は、複数のリンクでＣＰＵとＦＰＧＡを接続する場合の一例を示す図である。
【図２】図２は、実施例の情報処理装置のハードウェア構成の一例を示すブロック図である。
【図３】図３は、実施例の情報処理装置の機能構成の一例を示すブロック図である。
【図４】図４は、制御回路の機能構成の一例を示すブロック図である。
【図５】図５は、フローテーブルの一例を示す図である。
【図６】図６は、リンクテーブルの一例を示す図である。
【図７】図７は、実施例の経路制御処理の一例を示すフローチャートである。
【図８】図８は、実施例の切替処理の一例を示すフローチャートである。
【発明を実施するための形態】
【００１０】
以下、図面に基づいて、本願の開示する情報処理装置、情報処理方法および情報処理プログラムの実施例を詳細に説明する。なお、本実施例により、開示技術が限定されるものではない。また、以下の実施例は、矛盾しない範囲で適宜組みあわせてもよい。
【実施例】
【００１１】
まず、図１を用いてリンクの帯域が不足する場合について説明する。図１は、複数のリンクでＣＰＵとＦＰＧＡを接続する場合の一例を示す図である。図１の例では、ＣＰＵ１０にはメモリ１１が接続され、ＣＰＵ１０とＦＰＧＡ１２との間のリンクとして、ＰＣＩｅ経由の接続と、ＯｐｅｎＣＡＰＩ経由の接続との両方が用意される。ユーザロジックＵＬ２０１～ＵＬ２０３は、Ｍｕｘ／Ｄｅｍｕｘ１３を介して、ＰＣＩｅおよびＯｐｅｎＣＡＰＩと接続される。
【００１２】
ここで、ＰＣＩｅは、Ｇｅｎ４　ｘ１６ｌａｎｅ　ｘ２ｓｌｏｔであり、帯域が５１．２ＧＢ／ｓ、アクセス遅延が１μｓ～であるとする。また、ＯｐｅｎＣＡＰＩは、帯域が２５．６ＧＢ／ｓ、アクセス遅延が１００ｎｓ～であるとする。ユーザロジックの要求性能は、ユーザロジックＵＬ２０１は、必要帯域１０ＧＢ／ｓ、許容アクセス遅延２００ｎｓである。ユーザロジックＵＬ２０２は、必要帯域４０ＧＢ／ｓ、許容アクセス遅延２μｓである。ユーザロジックＵＬ２０３は、必要帯域２０ＧＢ／ｓ、許容アクセス遅延１μｓである。

【００１３】
ユーザロジックＵＬ２０１～ＵＬ２０３は、事前にどのリンクを使用するかは固定選択であるとし、運用時まで他のユーザロジックがどれだけリンクを使用するか不明であるとする。この場合、図１の例では、ユーザロジックＵＬ２０１は、ＯｐｅｎＣＡＰＩに割り振り、ユーザロジックＵＬ２０２は、ＰＣＩｅに割り振る。ところが、ユーザロジックＵＬ２０３は、ＯｐｅｎＣＡＰＩとＰＣＩｅとのうち、どちらに割り振っても、必要帯域がリンクの帯域を上回ることになる。つまり、図１は、フローが要求する帯域に対してリンクの帯域が不足し、アプリケーションの性能要求を満たさない場合の例である。これに対し、動的にユーザロジックＵＬ２０３の通信を２つの性質の異なるリンクに割り振ることができれば、システム全体としての性能を向上させることができる。
【００１４】
図２は、実施例の情報処理装置のハードウェア構成の一例を示すブロック図である。図２に示すように、情報処理装置１００は、通信部１１０と、表示部１１１と、操作部１１２と、ＨＤＤ（Hard Disk Drive）１１３と、メモリ１２０と、ＣＰＵ１３０と、ＦＰＧＡ１４０とを有する。なお、通信部１１０、表示部１１１、操作部１１２、ＨＤＤ１１３、ＣＰＵ１３０およびＦＰＧＡ１４０は、バス１１４を介して相互に接続される。また、ＣＰＵ１３０は、メモリ１２０およびＦＰＧＡ１４０と接続される。なお、情報処理装置１００は、図２に示す機能部以外にも既知のコンピュータが有する各種の機能部、例えば各種の入力デバイスや音声出力デバイス等の機能部を有することとしてもかまわない。
【００１５】
通信部１１０は、例えば、ＮＩＣ（Network Interface Card）等によって実現される。通信部１１０は、図示しないネットワークを介して他の情報処理装置と有線または無線で接続され、他の情報処理装置との間で情報の通信を司る通信インタフェースである。
【００１６】
表示部１１１は、各種情報を表示するための表示デバイスである。表示部１１１は、例えば、表示デバイスとして液晶ディスプレイ等によって実現される。表示部１１１は、ＣＰＵ１３０から図示しない表示制御部を介して入力された表示画面等の各種画面を表示する。
【００１７】
操作部１１２は、情報処理装置１００のユーザから各種操作を受け付ける入力デバイスである。操作部１１２は、例えば、入力デバイスとして、キーボードやマウス等によって実現される。操作部１１２は、ユーザによって入力された操作を操作情報としてＣＰＵ１３０に出力する。なお、操作部１１２は、入力デバイスとして、タッチパネル等によって実現されるようにしてもよく、表示部１１１の表示デバイスと、操作部１１２の入力デバイスとは、一体化されるようにしてもよい。
【００１８】
ＨＤＤ１１３は、補助記憶装置であり、ＣＰＵ１３０で動作するＯＳ（Operating System）や各種データを記憶する。なお、ＨＤＤ１１３は、ハードディスクドライブの他にもフラッシュメモリ等の半導体メモリ素子を用いたＳＳＤ（Solid State Drive）や光ディスク等の記憶装置によって実現されてもよい。
【００１９】
メモリ１２０は、主記憶装置であり、例えば、各種のＳＤＲＡＭ（Synchronous Dynamic Random Access Memory）等のようなＲＡＭ（Random Access Memory）等の半導体メモリ素子等の記憶装置によって実現される。メモリ１２０は、ＣＰＵ１３０での処理に用いる情報を記憶する。なお、メモリ１２０は、バス１１４に接続してもよいし、ＦＰＧＡ１４０と直接接続してもよい。
【００２０】
ＣＰＵ１３０は、ＨＤＤ１１３等の記憶部に記憶されているプログラムに従って、メモリ１２０等のＲＡＭを作業領域として各種処理を実行する。すなわち、ＣＰＵ１３０は、ＯＳやＶＭ（Virtual Machine）によって制御され、各種処理を実行する。

【００２１】
ＦＰＧＡ１４０は、ＣＰＵ１３０で動作するアプリケーションをオフロードするユーザロジックを動作させる。ＦＰＧＡ１４０は、ＣＰＵ１３０とバス１１４（例えば、ＰＣＩｅ。）経由、および、メモリコヒーレントバス（例えば、ＯｐｅｎＣＡＰＩ。）経由で接続される。すなわち、ＦＰＧＡ１４０は、複数のリンクを経由してＣＰＵ１３０と接続されるオフロード回路を構成する。なお、複数のリンクは、異なる種類のバスでなく、同じ種類のバスであってもよい。
【００２２】
図３は、実施例の情報処理装置の機能構成の一例を示すブロック図である。図３に示すように、情報処理装置１００では、ＣＰＵ１３０でＯＳ／ＶＭ１３１が動作し、さらに、ＯＳ／ＶＭ１３１上でアプリケーションＡ１～Ａｍが動作している。また、ＦＰＧＡ１４０は、制御回路１４１と複数のユーザロジックＵＬ１～ＵＬｍとを有する。制御回路１４１は、リンクＬ１～ＬｎでＣＰＵ１３０と接続される。リンクＬ１～Ｌｎは、バス１１４およびメモリコヒーレントバスに対応する。ユーザロジックＵＬ１～ＵＬｍは、アプリケーションＡ１～Ａｍと対応付けられており、対応するアプリケーションＡ１～Ａｍとの間で、制御回路１４１およびリンクＬ１～Ｌｎを介して通信を行う。つまり、ユーザロジックＵＬ１～ＵＬｍは、アプリケーションＡ１～Ａｍの処理を演算する論理回路である。また、アプリケーションＡ１～Ａｍ、および、ユーザロジックＵＬ１～ＵＬｍは、使用するリンクＬ１～Ｌｎのいずれかを用いて、それぞれ性能情報を制御回路１４１に出力する。なお、性能情報は、例えば、性能要件を満たす場合を「１」とし、性能要件を満たさない場合を「０」とする。
【００２３】
図４は、制御回路の機能構成の一例を示すブロック図である。図４に示すように、制御回路１４１は、インタフェースＩＦ１～ＩＦｎと、Ｍｕｘ／ｄｅｍｕｘ１４２と、Ｍｕｘ／ｄｅｍｕｘ１４３と、モニタ１４４と、セレクタ１４７とを有する。また、モニタ１４４は、フローテーブル１４５と、リンクテーブル１４６とを有する。
【００２４】
インタフェースＩＦ１～ＩＦｎは、リンクＬ１～Ｌｎに対応するインタフェースである。インタフェースＩＦ１～ＩＦｎは、リンクごとの使用帯域を表すデータであるPerf_per_link(1)～(n)を、それぞれモニタ１４４に出力する。また、インタフェースＩＦ１～ＩＦｎは、フローを識別するフローＩＤ（Identifier）をＭｕｘ／ｄｅｍｕｘ１４２に出力する。なお、フローは、ユーザロジックＵＬ１～ＵＬｍ、コマンド種別（Ｒｅａｄ／Ｗｒｉｔｅ）およびアドレス範囲の組み合わせに基づくデータの流れを表し、フローＩＤを用いて各フローを識別する。
【００２５】
Ｍｕｘ／ｄｅｍｕｘ１４２は、マルチプレクサおよびデマルチプレクサである。Ｍｕｘ／ｄｅｍｕｘ１４２は、インタフェースＩＦ１～ＩＦｎ側から入力されるフローを多重化し、Ｍｕｘ／ｄｅｍｕｘ１４３に出力する。また、Ｍｕｘ／ｄｅｍｕｘ１４２は、Ｍｕｘ／ｄｅｍｕｘ１４３から入力される多重化されたフローを複数のフローに戻し、対応するインタフェースＩＦ１～ＩＦｎに出力する。Ｍｕｘ／ｄｅｍｕｘ１４２は、インタフェースＩＦ１～ＩＦｎ側、および、Ｍｕｘ／ｄｅｍｕｘ１４３側から入力された各フローのフローＩＤをセレクタ１４７に出力する。Ｍｕｘ／ｄｅｍｕｘ１４２には、セレクタ１４７からフローＩＤに応じたリンク情報が入力される。Ｍｕｘ／ｄｅｍｕｘ１４２は、リンク情報に応じたインタフェースＩＦ１～ＩＦｎに、対応するフローを出力するようにフローの経路を切り替える。
【００２６】
Ｍｕｘ／ｄｅｍｕｘ１４３は、マルチプレクサおよびデマルチプレクサである。Ｍｕｘ／ｄｅｍｕｘ１４３は、ユーザロジックＵＬ１～ＵＬｍ側から入力されるパケットからフローを抽出して多重化し、Ｍｕｘ／ｄｅｍｕｘ１４２に出力する。また、Ｍｕｘ／ｄｅｍｕｘ１４３は、Ｍｕｘ／ｄｅｍｕｘ１４２から入力される多重化されたフローを複数のフ

ローに戻し、対応するユーザロジックＵＬ１～ＵＬｍに出力する。Ｍｕｘ／ｄｅｍｕｘ１４３は、各フローに対応するフローＩＤをＭｕｘ／ｄｅｍｕｘ１４２に出力する。また、Ｍｕｘ／ｄｅｍｕｘ１４３は、抽出したフローに関するフロー情報と、フローごとの使用帯域とを含むデータであるPerf_per_flow(1...m)をモニタ１４４に出力する。

【００２７】

モニタ１４４は、リンクＬ１～Ｌｎの使用帯域と、各フローの使用帯域と、性能情報とを収集し、フローテーブル１４５に記憶する。ここで、図５および図６を用いて、フローテーブル１４５およびリンクテーブル１４６について説明する。

【００２８】

フローテーブル１４５は、フローごとにリンク情報（経路情報）、優先度、性能情報等を対応付けて記憶する。図５は、フローテーブルの一例を示す図である。図５に示すように、フローテーブル１４５は、「フローＩＤ」、「フロー情報」、「リンク情報」、「優先度」、「性能情報（現在値）」といった項目を有する。また、「フロー情報」は、「ユーザロジック」、「コマンド種別」、「アドレス範囲」といった項目を有する。また、「性能情報（現在値）」は、「アプリ性能」、「ＵＬ性能」、「使用帯域」といった項目を有する。

【００２９】

「フローＩＤ」は、フローを識別する識別子である。「ユーザロジック」は、当該フローに対応するユーザロジックを識別する識別子である。「コマンド種別」は、Ｒｅａｄ／Ｗｒｉｔｅの別を示す情報である。「コマンド種別」は、Ｒｅａｄ／Ｗｒｉｔｅのどちらも対象である場合はＲ／Ｗとする。「アドレス範囲」は、当該フローに対応するアドレス範囲を示す情報である。「リンク情報」は、当該フローが現在割り振られているリンクを示す情報である。

【００３０】

「優先度」は、各フローの振り分けの優先度を示し、例えば、「０～２５５」の２５６段階で表すことができる。この場合、数値が大きいほど優先度が高いとしている。「アプリ性能」は、アプリケーションＡ１～Ａｍから収集した性能情報である。「アプリ性能」は、ＯＫであれば性能要件を満たし、ＮＧであれば性能要件を満たさないことを示す。ここで、「アプリ性能」が性能要件を満たすとは、例えば、アプリケーションからユーザロジックに対してリクエストを送信してレスポンスが返ってくるまでの応答時間やスループット等が条件を満たす場合である。

【００３１】

「ＵＬ性能」は、ユーザロジックＵＬ１～ＵＬｍから収集した性能情報である。「ＵＬ性能」は、ＯＫであれば性能要件を満たし、ＮＧであれば性能要件を満たさないことを示す。ここで、「ＵＬ性能」が性能要件を満たすとは、例えば、ユーザロジックからアプリケーションにレスポンスを送信してＡＣＫが返ってくるまでの応答時間や、アプリケーションからの単位時間あたりのリクエスト回数等が条件を満たす場合である。

【００３２】

「アプリ性能」および「ＵＬ性能」では、ＯＫは、モニタ１４４が収集した性能情報「１」に対応し、ＮＧは、モニタ１４４が収集した性能情報「０」に対応する。性能情報は、例えば、モニタ１４４内に、アプリケーションＡ１～Ａｍ、および、ユーザロジックＵＬ１～ＵＬｍに対応するレジスタを設け、それぞれがレジスタに「０」または「１」を書き込み、モニタ１４４が定期的にレジスタを参照することで収集できる。「使用帯域」は、当該フローが使用しているリンクＬ１～Ｌｎの帯域を示す情報である。

【００３３】

リンクテーブル１４６は、リンクＬ１～Ｌｎの性能情報を記憶する。図６は、リンクテーブルの一例を示す図である。図６に示すように、リンクテーブル１４６は、「リンクＩＤ」、「性能情報」といった項目を有する。また、「性能情報」は、「最大帯域」、「最小遅延」といった項目を有する。

【００３４】

「リンクＩＤ」は、リンクＬ１～Ｌｎを識別する識別子である。「最大帯域」は、当該リンクが収容可能な最大の帯域を示す情報である。「最小遅延」は、当該リンクにおける最小の遅延時間を示す情報である。

【００３５】

図４の説明に戻って、モニタ１４４は、使用帯域や性能情報の収集とともに、各フローの優先度を制御する。モニタ１４４は、電源が投入されると、初期状態を設定する。モニタ１４４は、初期状態として、フローテーブル１４５の優先度欄を全て「１」に設定する。モニタ１４４は、アプリケーションＡ１～Ａm、および、ユーザロジックＵＬ１～ＵＬmの動作が開始されると、フローテーブル１４５のフロー情報、リンク情報および性能情報を更新する。なお、フローＩＤは、例えば、アプリケーションの起動時に設定されたフローＩＤをアプリケーションから取得する。また、フローＩＤは、例えば、アプリケーションの終了時にアプリケーションからの指示により削除される。

【００３６】

つまり、モニタ１４４は、Ｍｕｘ／ｄｅｍｕｘ１４３から入力されるフロー情報に基づいて、フローテーブル１４５のフロー情報を更新する。モニタ１４４は、リンクテーブル１４６を参照し、インタフェースＩＦ１～ＩＦｎから入力されるリンクごとの使用帯域に基づいて、使用可能な帯域が多いリンクから順にフローを割り振って、フローテーブル１４５のリンク情報を更新する。モニタ１４４は、Ｍｕｘ／ｄｅｍｕｘ１４３から入力されるフローごとの使用帯域に基づいて、フローテーブル１４５のフローごとの使用帯域を更新する。モニタ１４４は、アプリケーションＡ１～Ａm、および、ユーザロジックＵＬ１～ＵＬmに対応するレジスタを参照し、フローテーブル１４５のアプリ性能およびＵＬ性能を更新する。

【００３７】

モニタ１４４は、フローテーブル１４５のアプリ性能およびＵＬ性能がＮＧであるフローがある場合、当該フローの優先度をインクリメントする。次に、モニタ１４４は、セレクタ１４７に対して、切替処理の実行を指示する。

【００３８】

モニタ１４４は、切替処理の終了後、フローテーブル１４５を参照し、全フローにおいて、アプリ性能またはＵＬ性能がＮＧであれば、全フローの優先度をデクリメントする。また、モニタ１４４は、フローテーブル１４５を参照し、全フローにおいて、アプリ性能およびＵＬ性能がＯＫであれば、全フローの優先度をデクリメントし、一定時間の待機後、再びアプリ性能およびＵＬ性能を更新して優先度の制御を繰り返す。

【００３９】

言い換えると、モニタ１４４は、アプリケーションの処理に対応するフローごとのリンクの性能情報を示す値と、リンクごとの使用可能な性能情報の最大値とを収集する収集部の一例である。また、モニタ１４４は、アプリケーションの処理に対応するフローごとのリンクの使用帯域を示す値と、リンクごとの使用可能な帯域の最大値とを収集する。また、モニタ１４４は、アプリケーションの処理、または、論理回路の性能が性能要件を満たしていない場合、対応するフローの優先度を上げる。また、モニタ１４４は、所定時間ごとに、リンクの使用帯域を示す値と、リンクごとの使用可能な帯域の最大値とを収集する。

【００４０】

セレクタ１４７は、Ｍｕｘ／ｄｅｍｕｘ１４２からフローＩＤが入力されると、フローテーブル１４５を参照し、入力されたフローＩＤに対応するリンク情報をＭｕｘ／ｄｅｍｕｘ１４２に出力する。すなわち、セレクタ１４７は、フローテーブル１４５を参照し、リンクＬ１～Ｌｎを流れるパケットを、適切なユーザロジックＵＬ１～ＵＬmに接続する。

【００４１】

また、セレクタ１４７は、モニタ１４４から切替処理の実行を指示されると、フローとリンクとの対応を切り替える切替処理を実行する。セレクタ１４７は、フローテーブル１

４５を参照し、未判定のフローがあるか否かを判定する。セレクタ１４７は、未判定のフローがないと判定した場合には、切替処理を終了する。

【００４２】

セレクタ１４７は、未判定のフローがあると判定した場合には、未判定のフローのうち、優先度が最も高いフローを対象フローとして選択する。セレクタ１４７は、選択した対象フローについて、アプリ性能およびＵＬ性能がＯＫであるか否かを判定する。セレクタ１４７は、アプリ性能およびＵＬ性能がＯＫであると判定した場合には、次のフローの判定に進む。

【００４３】

セレクタ１４７は、アプリ性能およびＵＬ性能がＯＫでないと判定した場合には、フローテーブル１４５およびリンクテーブル１４６を参照し、最も使用可能な帯域が大きいリンクを選択する。セレクタ１４７は、選択したリンクが対象フローのリンクと同じであるか否かを判定する。セレクタ１４７は、選択したリンクが対象フローのリンクと同じであると判定した場合には、次のフローの判定に進む。

【００４４】

セレクタ１４７は、選択したリンクが対象フローのリンクと同じでないと判定した場合には、対象フローに選択したリンクを設定する。セレクタ１４７は、選択したリンクの使用帯域が最大帯域に収まるか否かを判定する。セレクタ１４７は、選択したリンクの使用帯域が最大帯域に収まると判定した場合には、切替処理を終了する。

【００４５】

セレクタ１４７は、選択したリンクの使用帯域が最大帯域に収まらないと判定した場合には、フローテーブル１４５を参照し、選択したリンクを使用する最も優先度が低いフローを選択する。セレクタ１４７は、選択した最も優先度が低いフローのリンクに、対象フローの元のリンクを設定し、切替処理を終了する。すなわち、セレクタ１４７は、対象フローの元のリンクと、最も優先度が低いフローのリンクとを入れ替える。

【００４６】

言い換えると、セレクタ１４７は、フローごとのリンクの性能情報を示す値に基づいて、要する性能情報を満たしていないフローを判定する。セレクタ１４７は、リンクごとの使用可能な性能情報の最大値と、リンクごとの現在使用されている性能情報の値とに基づいて、フローの振り分け先のリンクを選択して振り分ける。つまり、セレクタ１４７は、選択部の一例である。すなわち、セレクタ１４７は、フローごとのリンクの使用帯域を示す値に基づいて、要する帯域を満たしていないフローを判定する。セレクタ１４７は、リンクごとの使用可能な帯域の最大値と、リンクごとの現在使用されている帯域の値とに基づいて、フローの振り分け先のリンクを選択して振り分ける。

【００４７】

また、セレクタ１４７は、リンクごとの使用可能な帯域の最大値と、リンクごとの現在使用されている帯域の値とに基づいて、フローの振り分け先のリンクを選択する。セレクタ１４７は、選択した振り分け先のリンクに該フローを振り分けると、振り分け先のリンクの使用可能な帯域の最大値を超える場合、振り分け先のリンクを使用する最も優先度が低いフローを、振り分け元のリンクに振り分ける。また、セレクタ１４７は、所定時間ごとに、要する帯域を満たしていないフローを判定する。

【００４８】

次に、実施例の情報処理装置１００の動作について説明する。図７は、実施例の経路制御処理の一例を示すフローチャートである。

【００４９】

制御回路１４１のモニタ１４４は、電源が投入されると、初期状態を設定する（ステップＳ１）。モニタ１４４は、アプリケーションＡ１～Ａｍ、および、ユーザロジックＵＬ１～ＵＬｍの動作が開始されると、フローテーブル１４５のフロー情報、リンク情報および性能情報を更新する。モニタ１４４は、動作開始後、一定時間待機する（ステップＳ２）。

【００５０】
その後、モニタ１４４は、アプリケーションＡ１～Ａｍ、および、ユーザロジックＵＬ１～ＵＬｍに対応するレジスタを参照し、フローテーブル１４５のアプリ性能およびＵＬ性能を更新する（ステップＳ３）。
【００５１】
モニタ１４４は、フローテーブル１４５のアプリ性能およびＵＬ性能がＮＧであるフローの優先度をインクリメントする（ステップＳ４）。次に、モニタ１４４は、セレクタ１４７に対して、切替処理の実行を指示する（ステップＳ５）。
【００５２】
ここで、図８を用いて切替処理について説明する。図８は、実施例の切替処理の一例を示すフローチャートである。
【００５３】
セレクタ１４７は、モニタ１４４から切替処理の実行を指示されると、フローとリンクとの対応を切り替える切替処理を実行する。セレクタ１４７は、フローテーブル１４５を参照し、未判定のフローがあるか否かを判定する（ステップＳ５１）。セレクタ１４７は、未判定のフローがないと判定した場合には（ステップＳ５１：否定）、切替処理を終了し、経路制御処理に戻る。
【００５４】
セレクタ１４７は、未判定のフローがあると判定した場合には（ステップＳ５１：肯定）、未判定のフローのうち、優先度が最も高いフローを対象フローとして選択する（ステップＳ５２）。セレクタ１４７は、選択した対象フローについて、アプリ性能およびＵＬ性能がＯＫであるか否かを判定する（ステップＳ５３）。セレクタ１４７は、アプリ性能およびＵＬ性能がＯＫであると判定した場合には（ステップＳ５３：肯定）、次のフローの判定を行うため、ステップＳ５１に戻る。
【００５５】
セレクタ１４７は、アプリ性能およびＵＬ性能がＯＫでないと判定した場合には（ステップＳ５３：否定）、フローテーブル１４５およびリンクテーブル１４６を参照し、最も使用可能な帯域が大きいリンクを選択する（ステップＳ５４）。セレクタ１４７は、選択したリンクが対象フローのリンクと同じであるか否かを判定する（ステップＳ５５）。セレクタ１４７は、選択したリンクが対象フローのリンクと同じであると判定した場合には（ステップＳ５５：肯定）、次のフローの判定を行うため、ステップＳ５１に戻る。
【００５６】
セレクタ１４７は、選択したリンクが対象フローのリンクと同じでないと判定した場合には（ステップＳ５５：否定）、対象フローに選択したリンクを設定する（ステップＳ５６）。セレクタ１４７は、選択したリンクの使用帯域が最大帯域に収まるか否かを判定する（ステップＳ５７）。セレクタ１４７は、選択したリンクの使用帯域が最大帯域に収まると判定した場合には（ステップＳ５７：肯定）、切替処理を終了し、経路制御処理に戻る。
【００５７】
セレクタ１４７は、選択したリンクの使用帯域が最大帯域に収まらないと判定した場合には（ステップＳ５７：否定）、フローテーブル１４５を参照し、選択したリンクを使用する最も優先度が低いフローを選択する（ステップＳ５８）。セレクタ１４７は、選択した最も優先度が低いフローのリンクに、対象フローの元のリンクを設定し（ステップＳ５９）、切替処理を終了して経路制御処理に戻る。これにより、セレクタ１４７は、優先度の高いフローからリンクに振り分けることができる。
【００５８】
図７の説明に戻って、モニタ１４４は、切替処理の終了後、フローテーブル１４５を参照し、全フローにおいて、アプリ性能またはＵＬ性能がＮＧであれば、全フローの優先度をデクリメントする（ステップＳ６）。モニタ１４４は、フローテーブル１４５を参照し、全フローにおいて、アプリ性能およびＵＬ性能がＯＫであれば、全フローの優先度をデ

クリメントし（ステップＳ７）、ステップＳ２に戻る。これにより、制御回路１４１は、システム全体の性能を最大化するようにフローを振り分けることができる。また、制御回路１４１は、ユーザロジックに依存することなく、システム全体の性能を最大化する通信パターンを自動的に決定できる。また、制御回路１４１は、ユーザロジックのＦＰＧＡ接続方式に依存する部分を削減することができる。すなわち、情報処理装置１００では、他のシステムへのユーザロジックの移植が容易となる。また、情報処理装置１００では、インタフェース設計を共通化できるので、開発工数を削減できる。

【００５９】

なお、上記実施例では、リンクの性能情報として使用帯域を用いたが、これに限定されない。例えば、リンクの性能情報としてレイテンシを用いてもよい。この場合、制御回路１４１は、Ｍｕｘ／ｄｅｍｕｘ１４２にレイテンシチェッカを接続し、定期的に計測用パケットを各リンクに送信して取得したレイテンシ情報に基づいて、各フローを各リンクに振り分ける。すなわち、モニタ１４４は、アプリケーションの処理に対応するフローごとのリンクのレイテンシを示す値と、リンクごとのレイテンシの最大値とを収集する。セレクタ１４７は、フローごとのリンクのレイテンシを示す値に基づいて、要するレイテンシを満たしていないフローを判定する。セレクタ１４７は、リンクごとのレイテンシの最大値と、リンクごとの現在のレイテンシの値とに基づいて、フローの振り分け先のリンクを選択して振り分ける。

【００６０】

このように、情報処理装置１００は、複数のリンクを経由してＣＰＵと接続されるオフロード回路を備える。オフロード回路は、ＦＰＧＡ１４０として、制御回路１４１と、ユーザロジックである論理回路とを備える。論理回路は、アプリケーションの処理を演算する。制御回路１４１は、アプリケーションの処理に対応するフローごとのリンクの性能情報を示す値と、リンクごとの使用可能な性能情報の最大値とを収集する。また、制御回路１４１は、フローごとのリンクの性能情報を示す値に基づいて、要する性能情報を満たしていないフローを判定する。また、制御回路１４１は、リンクごとの使用可能な性能情報の最大値と、リンクごとの現在使用されている性能情報の値とに基づいて、フローの振り分け先のリンクを選択して振り分ける。その結果、制御回路１４１は、システム全体の性能を最大化するようにフローを振り分けることができる。

【００６１】

また、制御回路１４１は、アプリケーションの処理に対応するフローごとのリンクの使用帯域を示す値と、リンクごとの使用可能な帯域の最大値とを収集する。また、制御回路１４１は、フローごとのリンクの使用帯域を示す値に基づいて、要する帯域を満たしていないフローを判定する。また、制御回路１４１は、リンクごとの使用可能な帯域の最大値と、リンクごとの現在使用されている帯域の値とに基づいて、フローの振り分け先のリンクを選択して振り分ける。その結果、制御回路１４１は、リンクの使用帯域に基づいてシステム全体の性能を最大化するようにフローを振り分けることができる。

【００６２】

また、制御回路１４１は、アプリケーションの処理、または、論理回路の性能が性能要件を満たしていない場合、対応するフローの優先度を上げる。その結果、制御回路１４１は、優先度に応じてフローを振り分けることができる。

【００６３】

また、制御回路１４１は、リンクごとの使用可能な帯域の最大値と、リンクごとの現在使用されている帯域の値とに基づいて、フローの振り分け先のリンクを選択する。制御回路１４１は、選択した振り分け先のリンクに該フローを振り分けると、振り分け先のリンクの使用可能な帯域の最大値を超える場合、振り分け先のリンクを使用する最も優先度が低いフローを、振り分け元のリンクに振り分ける。その結果、制御回路１４１は、優先度の高いフローと優先度の低いフローとが使用するリンクを入れ替えることができる。

【００６４】

また、制御回路１４１は、所定時間ごとに、リンクの使用帯域を示す値と、リンクごと

の使用可能な帯域の最大値とを収集する。また、制御回路１４１は、所定時間ごとに、要する帯域を満たしていないフローを判定する。その結果、制御回路１４１は、動的にフローを振り分けることができる。

【００６５】

また、制御回路１４１は、アプリケーションの処理に対応するフローごとのリンクのレイテンシを示す値と、リンクごとのレイテンシの最大値とを収集する。また、制御回路１４１は、フローごとのリンクのレイテンシを示す値に基づいて、要するレイテンシを満たしていないフローを判定する。制御回路１４１は、リンクごとのレイテンシの最大値と、リンクごとの現在のレイテンシの値とに基づいて、フローの振り分け先のリンクを選択して振り分ける。その結果、制御回路１４１は、リンクのレイテンシに基づいてシステム全体の性能を最大化するようにフローを振り分けることができる。

【００６６】

また、図示した各部の各構成要素は、必ずしも物理的に図示の如く構成されていることを要しない。すなわち、各部の分散・統合の具体的形態は図示のものに限られず、その全部または一部を、各種の負荷や使用状況等に応じて、任意の単位で機能的または物理的に分散・統合して構成することができる。例えば、Ｍｕｘ／ｄｅｍｕｘ１４２とＭｕｘ／ｄｅｍｕｘ１４３とを統合してクロスバースイッチとしてもよい。また、図示した各処理は、上記の順番に限定されるものでなく、処理内容を矛盾させない範囲において、同時に実施してもよく、順序を入れ替えて実施してもよい。

【００６７】

さらに、制御回路１４１で行われる各種処理機能は、ＣＰＵ（またはＭＰＵ、ＭＣＵ（Micro Controller Unit）等のマイクロ・コンピュータ）上で、その全部または任意の一部を実行するようにしてもよい。また、各種処理機能は、ＣＰＵ（またはＭＰＵ、ＭＣＵ等のマイクロ・コンピュータ）で解析実行されるプログラム上、またはワイヤードロジックによるハードウェア上で、その全部または任意の一部を実行するようにしてもよいことは言うまでもない。

【００６８】

なお、上記実施例で説明した制御回路１４１は、プログラムを読み込んで実行することで、図４、図７、図８等で説明した処理と同様の機能を実行することができる。例えば、制御回路１４１は、モニタ１４４、セレクタ１４７と同様の処理を実行するプロセスを実行することで、上記実施例と同様の処理を実行することができる。

【００６９】

これらのプログラムは、インターネットなどのネットワークを介して配布することができる。また、これらのプログラムは、ハードディスク、フレキシブルディスク（ＦＤ）、ＣＤ－ＲＯＭ、ＭＯ、ＤＶＤなどのコンピュータで読み取り可能な記録媒体に記録され、コンピュータによって記録媒体から読み出されることによって実行することができる。

【符号の説明】

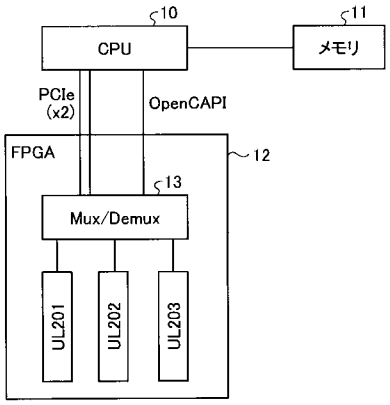
【００７０】

１００　情報処理装置
１１０　通信部
１１１　表示部
１１２　操作部
１１３　ＨＤＤ
１１４　バス
１２０　メモリ
１３０　ＣＰＵ
１３１　ＯＳ／ＶＭ
１４０　ＦＰＧＡ
１４１　制御回路
１４２，１４３　Ｍｕｘ／ｄｅｍｕｘ

１４４　モニタ
１４５　フローテーブル
１４６　リンクテーブル
１４７　セレクタ
Ａ１～Ａｍ　アプリケーション
ＩＦ１～ＩＦｎ　インタフェース
Ｌ１～Ｌｎ　リンク
ＵＬ１～ＵＬｍ　ユーザロジック

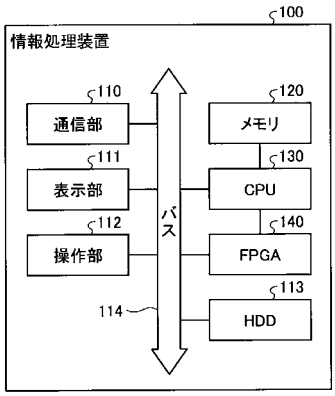
【図１】

複数のリンクでCPUとFPGAを接続する場合の一例を示す図



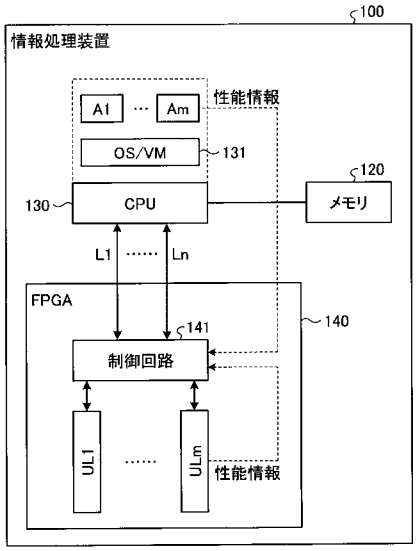
【図２】

実施例の情報処理装置のハードウェア構成の一例を示すブロック図



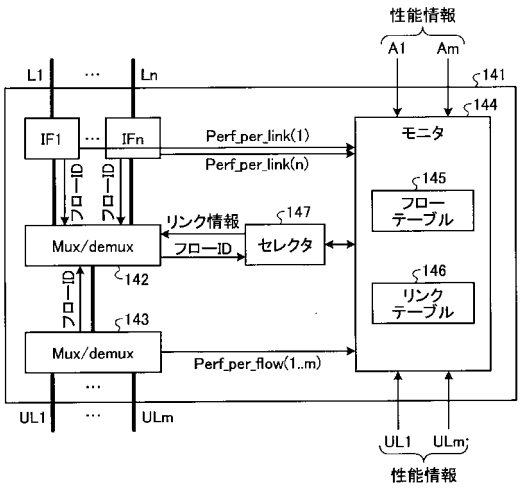
【図３】

実施例の情報処理装置の機能構成の一例を示すブロック図



【図４】

制御回路の機能構成の一例を示すブロック図



【図５】

フローテーブルの一例を示す図

145

フローID	フロー情報			リンク情報	優先度	性能情報(現在値)		
	ユーザロジック	コマンド種別	アドレス範囲		0:低～255:高	アプリ性能	UL性能	使用帯域
1	UL1	Read	All	L1	10	OK	NG	10GB/s
2	UL1	Write	All	L1	1	OK	OK	1GB/s
3	UL2	R/W	All	L1	11	NG	OK	20GB/s
4	UL3	R/W	0x0000-0xFFFF	L2	12	OK	OK	30GB/s
5	UL3	R/W	0x10000-0x1FFF	L2	10	OK	OK	10GB/s
…	…	…	…	…	…	…	…	…

【図６】

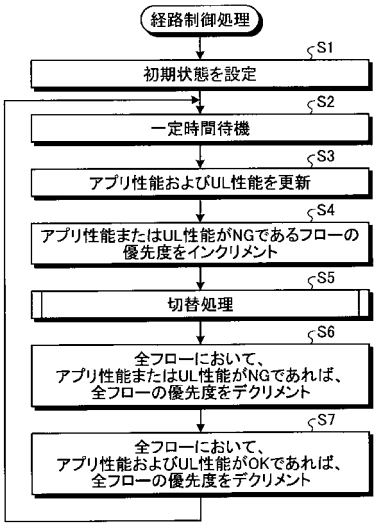
リンクテーブルの一例を示す図

146

リンクID	性能情報		
	最大帯域	最小遅延	…
L1	50GB/s	1000ns	…
L2	20GB/s	300ns	…
L3	10GB/s	100ns	…
…	…	…	…

【図７】

実施例の経路制御処理の一例を示すフローチャート



【図８】

実施例の切替処理の一例を示すフローチャート