

(19)日本国特許庁(JP)

(12)特許公報(B2)

(11)特許番号
特許第7206485号
(P7206485)

(45)発行日 令和5年1月18日(2023.1.18)

(24)登録日 令和5年1月10日(2023.1.10)

(51)国際特許分類

F I

G 0 6 F 13/38 (2006.01)

G 0 6 F 13/38 3 4 0 A

G 0 6 F 13/36 (2006.01)

G 0 6 F 13/36 3 1 0 E

請求項の数 20 (全22頁)

(21)出願番号	特願2018-553804(P2018-553804)	(73)特許権者	514315159
(86)(22)出願日	平成29年11月22日(2017.11.22)		株式会社ソシオネクスト
(86)国際出願番号	PCT/JP2017/041935		神奈川県横浜市港北区新横浜2丁目10番23
(87)国際公開番号	WO2018/101136	(74)代理人	110002918
(87)国際公開日	平成30年6月7日(2018.6.7)		弁理士法人扶桑国際特許事務所
審査請求日	令和2年10月14日(2020.10.14)	(72)発明者	後藤 誠司
(31)優先権主張番号	特願2016-232340(P2016-232340)		神奈川県横浜市港北区新横浜二丁目10番23 株式会社ソシオネクスト内
(32)優先日	平成28年11月30日(2016.11.30)	(72)発明者	仁茂田 永一
(33)優先権主張国・地域又は機関	日本国(JP)		神奈川県横浜市港北区新横浜二丁目10番23 株式会社ソシオネクスト内
前置審査		(72)発明者	岡本 諭
			神奈川県横浜市港北区新横浜二丁目10番23 株式会社ソシオネクスト内
			最終頁に続く

(54)【発明の名称】 情報処理システム、半導体集積回路及び情報処理方法

(57)【特許請求の範囲】

【請求項1】

複数の半導体集積回路のそれぞれが用いる各メモリのアドレスが定義された第1のメモリマップ情報に基づいて、前記複数の半導体集積回路の1つであるデータ送信先に接続された第1のメモリの第1のアドレスを指定し、前記第1のアドレスを、前記データ送信先で参照され、前記第1のメモリマップ情報とは異なる第2のメモリマップ情報において定義される前記第1のメモリの第2のアドレスに変換し、前記第2のアドレスと送信データとを、前記データ送信先からの応答を待たずに連続的な送信が可能なバスインタフェースを用いて出力する第1の半導体集積回路と、

前記第2のアドレスと前記送信データとを、前記バスインタフェースを用いて前記データ送信先に転送するスイッチと、

前記データ送信先であり、前記バスインタフェースを用いて、前記第2のアドレスと前記送信データとを受信し、前記第2のアドレスに対応する前記第1のメモリの受信バッファ領域に前記送信データを書き込む第2の半導体集積回路と、

を有し、

前記第1のメモリマップ情報と前記第2のメモリマップ情報は、互いに共通のフォーマットで定義された第1のアドレス空間と、第2のアドレス空間とを有し、

前記第1のメモリマップ情報の前記第1のアドレス空間には、前記第1の半導体集積回路に接続された第2のメモリのアドレスが定義されており、

前記第1のメモリマップ情報の前記第2のアドレス空間には、前記複数の半導体集積回

路のうち、前記第 1 の半導体集積回路以外の半導体集積回路に接続されたメモリのアドレスが定義されており、

前記第 2 のメモリマップ情報の前記第 1 のアドレス空間には、前記第 1 のメモリのアドレスが定義されており、

前記第 2 のメモリマップ情報の前記第 2 のアドレス空間には、前記複数の半導体集積回路のうち、前記第 2 の半導体集積回路以外の半導体集積回路に接続されたメモリのアドレスが定義されている、

ことを特徴とする情報処理システム。

【請求項 2】

前記第 1 のメモリマップ情報には、前記複数の半導体集積回路のそれぞれにおいて、送信が終了したことを示す割り込み設定値が設定される各記憶回路のアドレスがさらに定義されており、

前記第 1 の半導体集積回路は、前記第 1 のメモリマップ情報に基づいて、前記データ送信先において前記割り込み設定値が設定される第 1 の記憶回路の第 3 のアドレスを指定し、前記第 3 のアドレスを、前記第 2 のメモリマップ情報において定義される前記第 1 の記憶回路の第 4 のアドレスに変換し、前記第 4 のアドレスと前記割り込み設定値とを、前記バスインタフェースを用いて出力し、

前記スイッチは、前記第 4 のアドレスと前記割り込み設定値とを、前記バスインタフェースを用いて前記データ送信先に転送し、

前記第 2 の半導体集積回路は、前記バスインタフェースを用いて、前記第 4 のアドレスと前記割り込み設定値とを受信し、前記第 4 のアドレスに対応する前記第 1 の記憶回路に前記割り込み設定値を書き込み、前記割り込み設定値に基づいて、前記送信データの受信の終了を判定する、

ことを特徴とする請求項 1 に記載の情報処理システム。

【請求項 3】

前記スイッチは、第 1 のスイッチと、前記第 1 のスイッチと前記バスインタフェースを用いて情報の送受信を行う第 2 のスイッチとを含み、

前記第 2 のアドレス、前記第 4 のアドレス、前記送信データ及び前記割り込み設定値は、前記第 1 のスイッチと前記第 2 のスイッチとを介して前記第 2 の半導体集積回路に転送される、

ことを特徴とする請求項 2 に記載の情報処理システム。

【請求項 4】

前記スイッチは、前記第 2 のアドレス及び前記第 4 のアドレスを、前記複数の半導体集積回路のうち、第 3 の半導体集積回路が参照する第 3 のメモリマップ情報において定義される、前記第 3 の半導体集積回路に接続された第 3 のメモリの第 5 のアドレスと、前記第 3 の半導体集積回路において前記割り込み設定値が設定される第 2 の記憶回路の第 6 のアドレスに変換する、ことを特徴とする請求項 2 または 3 に記載の情報処理システム。

【請求項 5】

前記第 2 の半導体集積回路は、第 2 の制御回路と、前記割り込み設定値に基づいて、前記送信データの受信が終了したことを示す割り込み信号を前記第 2 の制御回路に通知する割り込み制御回路と、を有することを特徴とする請求項 2 乃至 4 の何れか一項に記載の情報処理システム。

【請求項 6】

前記第 1 のメモリマップ情報は、さらに、

前記第 1 の半導体集積回路の前記割り込み設定値が設定される記憶回路のアドレスを定義する第 3 のアドレス空間と、

前記第 1 の半導体集積回路以外の半導体集積回路の割り込み設定値が設定される記憶回路のアドレスを定義する第 4 のアドレス空間と、

を含むことを特徴とする請求項 2 乃至 5 の何れか一項に記載の情報処理システム。

【請求項 7】

10

20

30

40

50

前記受信バッファ領域には、前記複数の半導体集積回路のそれぞれに対応したバッファ領域が、それぞれ異なるアドレスで指定されるように設定されている、ことを特徴とする請求項 1 乃至 6 の何れか一項に記載の情報処理システム。

【請求項 8】

前記第 1 の半導体集積回路は、前記第 1 のアドレスを指定する第 1 の制御回路と、前記第 1 のアドレスを前記第 2 のアドレスに変換する第 1 のアドレス変換回路と、前記バスインタフェースとを有する、ことを特徴とする請求項 1 乃至 7 の何れか一項に記載の情報処理システム。

【請求項 9】

前記データ送信先からの応答を待たずに連続的な送信が可能な前記バスインタフェースは、P C I e インタフェースであることを特徴とする請求項 1 乃至 8 の何れか一項に記載の情報処理システム。

【請求項 10】

前記複数の半導体集積回路のうちの 2 つ以上には、前記複数の半導体集積回路のそれぞれが用いる各メモリよりも容量が大きい第 4 のメモリが接続されている、請求項 1 乃至 9 の何れか一項に記載の情報処理システム。

【請求項 11】

前記複数の半導体集積回路のうちの 2 つ以上には、特定機能向け L S I が接続されている、請求項 1 乃至 10 の何れか一項に記載の情報処理システム。

【請求項 12】

複数の半導体集積回路のそれぞれが用いる各メモリのアドレスが定義された第 1 のメモリマップ情報に基づいて、前記複数の半導体集積回路の 1 つであるデータ送信先に接続された第 1 のメモリの第 1 のアドレスを指定する制御回路と、

前記第 1 のアドレスを、前記データ送信先で参照され、前記第 1 のメモリマップ情報とは異なる第 2 のメモリマップ情報において定義される前記第 1 のメモリの第 2 のアドレスに変換するアドレス変換回路と、

前記第 2 のアドレスと送信データとを出力する、前記データ送信先からの応答を待たずに連続的な送信が可能な第 1 のバスインタフェースと、

を有し、

前記第 1 のメモリマップ情報と前記第 2 のメモリマップ情報は、互いに共通のフォーマットで定義された第 1 のアドレス空間と、第 2 のアドレス空間とを有し、

前記第 1 のメモリマップ情報の前記第 1 のアドレス空間には、自身が用いる第 2 のメモリのアドレスが定義されており、

前記第 1 のメモリマップ情報の前記第 2 のアドレス空間には、前記複数の半導体集積回路のうち、自身以外の半導体集積回路に接続されたメモリのアドレスが定義されており、

前記第 2 のメモリマップ情報の前記第 1 のアドレス空間には、前記第 1 のメモリのアドレスが定義されており、

前記第 2 のメモリマップ情報の前記第 2 のアドレス空間には、前記複数の半導体集積回路のうち、前記データ送信先以外の半導体集積回路に接続されたメモリのアドレスが定義されている、ことを特徴とする半導体集積回路。

【請求項 13】

前記第 1 のメモリマップ情報には、前記複数の半導体集積回路のそれぞれにおいて、送信が終了したことを示す割り込み設定値が設定される各記憶回路のアドレスがさらに定義されており、

前記制御回路は、前記第 1 のメモリマップ情報に基づいて、前記データ送信先において前記割り込み設定値が設定される第 1 の記憶回路の第 3 のアドレスを指定し、

前記アドレス変換回路は、前記第 3 のアドレスを、前記第 2 のメモリマップ情報において定義される前記第 1 の記憶回路の第 4 のアドレスに変換し、

前記第 1 のバスインタフェースは、前記第 4 のアドレスと前記割り込み設定値とを出力する、ことを特徴とする請求項 12 に記載の半導体集積回路。

10

20

30

40

50

【請求項 14】

前記第1のバスインタフェースは、他の半導体集積回路から、前記第1のメモリマップ情報において定義される前記第2のメモリの第5のアドレスと、書き込みデータとを受信し、前記第5のアドレスに対応する前記第2のメモリの受信バッファ領域に前記書き込みデータを書き込む、ことを特徴とする請求項13に記載の半導体集積回路。

【請求項 15】

前記割り込み設定値が設定される第2の記憶回路をさらに有し、

前記第1のメモリマップ情報は、さらに、

前記第2の記憶回路のアドレスを定義する第3のアドレス空間と、

前記第1の記憶回路のアドレスを定義する第4のアドレス空間と、

を含む、ことを特徴とする請求項14に記載の半導体集積回路。

10

【請求項 16】

前記第1のバスインタフェースは、前記他の半導体集積回路から出力された、前記第1のメモリマップ情報において定義される前記第2の記憶回路の第6のアドレスと、前記割り込み設定値とを受信し、前記第6のアドレスに対応する前記第2の記憶回路に前記割り込み設定値を書き込み、

前記制御回路は、前記割り込み設定値に基づいて、前記書き込みデータの受信の終了を判定する、ことを特徴とする請求項15に記載の半導体集積回路。

【請求項 17】

前記割り込み設定値に基づいて、前記書き込みデータの受信が終了したことを示す割り込み信号を前記制御回路に通知する割り込み制御回路をさらに有する、ことを特徴とする請求項16に記載の半導体集積回路。

20

【請求項 18】

前記複数の半導体集積回路のそれぞれが用いる各メモリよりも容量が大きい第3のメモリ、または、特定機能向けLSIと接続可能な第2のバスインタフェースを含む、ことを特徴とする請求項12乃至17の何れか一項に記載の半導体集積回路。

【請求項 19】

第1の半導体集積回路が、

複数の半導体集積回路のそれぞれが用いる各メモリのアドレスが定義された第1のメモリマップ情報に基づいて、前記複数の半導体集積回路の1つであるデータ送信先に接続された第1のメモリの第1のアドレスを指定し、前記第1のアドレスを、前記データ送信先で参照され、前記第1のメモリマップ情報とは異なる第2のメモリマップ情報において定義される前記第1のメモリの第2のアドレスに変換し、前記第2のアドレスと送信データとを、前記データ送信先からの応答を待たずに連続的な送信が可能なバスインタフェースを用いて出力し、

30

スイッチが、

前記第2のアドレスと前記送信データとを、前記バスインタフェースを用いて前記データ送信先に転送し、

前記データ送信先である第2の半導体集積回路が、

前記バスインタフェースを用いて、前記第2のアドレスと前記送信データとを受信し、前記第2のアドレスに対応する前記第1のメモリの受信バッファ領域に前記送信データを書き込み、

40

前記第1のメモリマップ情報と前記第2のメモリマップ情報は、互いに共通のフォーマットで定義された第1のアドレス空間と、第2のアドレス空間とを有し、

前記第1のメモリマップ情報の前記第1のアドレス空間には、前記第1の半導体集積回路に接続された第2のメモリのアドレスが定義されており、

前記第1のメモリマップ情報の前記第2のアドレス空間には、前記複数の半導体集積回路のうち、前記第1の半導体集積回路以外の半導体集積回路に接続されたメモリのアドレスが定義されており、

前記第2のメモリマップ情報の前記第1のアドレス空間には、前記第1のメモリのアド

50

レスが定義されており、

前記第2のメモリマップ情報の前記第2のアドレス空間には、前記複数の半導体集積回路のうち、前記第2の半導体集積回路以外の半導体集積回路に接続されたメモリのアドレスが定義されている、

ことを特徴とする情報処理方法。

【請求項20】

前記第1のメモリマップ情報には、前記複数の半導体集積回路のそれぞれにおいて、送信が終了したことを示す割り込み設定値が設定される各記憶回路のアドレスがさらに定義されており、

前記第1の半導体集積回路は、前記第1のメモリマップ情報に基づいて、前記データ送信先において前記割り込み設定値が設定される第1の記憶回路の第3のアドレスを指定し、前記第3のアドレスを、前記第2のメモリマップ情報において定義される前記第1の記憶回路の第4のアドレスに変換し、前記第4のアドレスと前記割り込み設定値とを、前記バスインタフェースを用いて出力し、

前記スイッチは、前記第4のアドレスと前記割り込み設定値とを、前記バスインタフェースを用いて前記データ送信先に転送し、

前記第2の半導体集積回路は、前記バスインタフェースを用いて、前記第4のアドレスと前記割り込み設定値とを受信し、前記第4のアドレスに対応する前記第1の記憶回路に前記割り込み設定値を書き込み、前記割り込み設定値に基づいて、前記送信データの受信の終了を判定する、

ことを特徴とする請求項19に記載の情報処理方法。

【発明の詳細な説明】

【技術分野】

【0001】

本発明は、情報処理システム、半導体集積回路及び情報処理方法に関する。

【背景技術】

【0002】

情報処理システムの複雑化や大規模化に伴い、複数のSoC (System on Chip) を含む情報処理システムが提案されている。複数のSoCを含む情報処理システムの1つとして、複数のCPU (Central Processing Unit) をネットワークで接続し、各CPUによる並列処理を実現するCPUクラスタが提案されている。

【0003】

なお、従来、イーサネット（登録商標）スイッチを用いて複数のサーバを接続する際に、イーサネットスイッチのハードウェアを削減するために、PCIe (Peripheral Component Interconnect express) を用いる技術があった。また、PCIeスイッチを用いて、複数のプロセッサを接続し、あるプロセッサが他のプロセッサに接続されるメモリにアクセスする技術があった。また、データセンタ内のサーバ間通信に、イーサネットの代わりにPCIeを用いることが提案されている。

【先行技術文献】

【特許文献】

【0004】

【文献】米国特許第7480303号明細書

米国特許出願公開第2015/0067229号明細書

国際公開第2013/136522号

【発明の概要】

【発明が解決しようとする課題】

【0005】

高速な処理が可能なCPUクラスタを実現するために、クラスタ構成を密結合クラスタとすることが考えられる。密結合クラスタでは、共有メモリやプロセス間通信が用いられ、各CPUが連携処理を行う。複数のCPUを接続するネットワークについては、比較的

10

20

30

40

50

低コストで高速通信が可能なイーサネットを用いることが考えられる。しかし、イーサネットを用いる場合、送信元は送信先からの応答信号の受信を待ってからでないといの送信を行えないなどのプロトコルオーバーヘッドなどが存在し、密結合クラスタの高速化を制限してしまう問題があった。

【課題を解決するための手段】

【0006】

発明の一観点によれば、複数の半導体集積回路のそれぞれが用いる各メモリのアドレスが定義された第1のメモリマップ情報に基づいて、前記複数の半導体集積回路の1つであるデータ送信先に接続された第1のメモリの第1のアドレスを指定し、前記第1のアドレスを、前記データ送信先で参照される第2のメモリマップ情報において定義される前記第1のメモリの第2のアドレスに変換し、前記第2のアドレスと送信データとを、前記データ送信先からの応答を待たずに連続的な送信が可能なバスインタフェースを用いて出力する第1の半導体集積回路と、前記第2のアドレスと前記送信データとを、前記バスインタフェースを用いて前記データ送信先に転送するスイッチと、前記データ送信先であり、前記バスインタフェースを用いて、前記第2のアドレスと前記送信データとを受信し、前記第2のアドレスに対応する前記第1のメモリの受信バッファ領域に前記送信データを書き込む第2の半導体集積回路と、を有する情報処理システムが提供される。

10

【0007】

また、発明の一観点によれば、複数の半導体集積回路のそれぞれが用いる各メモリのアドレスが定義された第1のメモリマップ情報に基づいて、前記複数の半導体集積回路の1つであるデータ送信先に接続された第1のメモリの第1のアドレスを指定する制御回路と、前記第1のアドレスを、前記データ送信先で参照される第2のメモリマップ情報において定義される前記第1のメモリの第2のアドレスに変換するアドレス変換回路と、前記第2のアドレスと送信データとを出力する、前記データ送信先からの応答を待たずに連続的な送信が可能な第1のバスインタフェースと、を有する半導体集積回路が提供される。

20

【0008】

また、発明の一観点によれば、第1の半導体集積回路が、複数の半導体集積回路のそれぞれが用いる各メモリのアドレスが定義された第1のメモリマップ情報に基づいて、前記複数の半導体集積回路の1つであるデータ送信先に接続された第1のメモリの第1のアドレスを指定し、前記第1のアドレスを、前記データ送信先で参照される第2のメモリマップ情報において定義される前記第1のメモリの第2のアドレスに変換し、前記第2のアドレスと送信データとを、前記データ送信先からの応答を待たずに連続的な送信が可能なバスインタフェースを用いて出力し、スイッチが、前記第2のアドレスと前記送信データとを、前記バスインタフェースを用いて前記データ送信先に転送し、前記データ送信先である第2の半導体集積回路が、前記バスインタフェースを用いて、前記第2のアドレスと前記送信データとを受信し、前記第2のアドレスに対応する前記第1のメモリの受信バッファ領域に前記送信データを書き込む、情報処理方法が提供される。

30

【発明の効果】

【0009】

情報処理システムの処理を高速化できる。

40

本発明の上記及び他の目的、特徴及び利点は本発明の例として好ましい実施の形態を表す添付の図面と関連した以下の説明により明らかになるであろう。

【図面の簡単な説明】

【0010】

【図1】第1の実施の形態の情報処理システムの一例を示す図である。

【図2】第2の実施の形態の情報処理システムの一例を示す図である。

【図3】メモリマップ情報の一例を示す図である。

【図4】各SOCのDRAM上の共用領域の一例を示す図である。

【図5】受信バッファ領域に格納されるパケットデータの一例を示す図である。

【図6】情報処理システムによる送信制御の一例の流れを示すフローチャートである。

50

【図 7】情報処理システムによる受信制御（読み出し制御）の一例の流れを示すフローチャートである。

【図 8】第 3 の実施の形態の情報処理システムの一例を示す図である。

【図 9】第 4 の実施の形態の情報処理システムの一例を示す図である。

【発明を実施するための形態】

【0011】

以下、発明を実施するための形態を、図面を参照しつつ説明する。

（第 1 の実施の形態）

図 1 は、第 1 の実施の形態の情報処理システムの一例を示す図である。

【0012】

情報処理システム 10 は、半導体集積回路 11、12、スイッチ 13、メモリ 14、15 を有する。

半導体集積回路 11、12、スイッチ 13 は、たとえば、それぞれ 1 チップの SoC である。図 1 では、説明を簡単にするために、2 つの半導体集積回路 11、12 と 1 つのスイッチ 13 が図示されているが、この数に限定されるものではない。3 つ以上の半導体集積回路と 2 つ以上のスイッチを有する情報処理システムの例については後述する。

【0013】

半導体集積回路 11 は、制御回路 11a、割り込み制御回路 11b、アドレス変換回路 11c、PCIe インタフェース（I/F）11d、システムバス 11e を有する。また、半導体集積回路 11 には、メモリ 14 が接続されている。

【0014】

制御回路 11a は、システムバス 11e を介して半導体集積回路 11 の各部を制御して、種々のアプリケーション処理を実行する。制御回路 11a は、たとえば、マイクロコントローラ、CPU、複数の CPU コアを有する CPU などである。

【0015】

割り込み制御回路 11b は、図示しない記憶回路（以下割り込み設定レジスタと呼ぶ）に設定される割り込み設定値に基づいて、割り込みの発生を制御回路 11a に通知する。

アドレス変換回路 11c は、複数の半導体集積回路 11、12 のそれぞれが用いるメモリ 14、15 のアドレスなどが定義されたメモリマップ情報 11m、12m に基づいて、アドレス変換を行う。

【0016】

PCIe インタフェース 11d は、送信先からの応答を待たずに連続的な送信が可能なバスインタフェースの一例であり、PCIe バスを介してスイッチ 13 と情報の送受信を行う。

【0017】

半導体集積回路 12 も、半導体集積回路 11 と同様に、制御回路 12a、割り込み制御回路 12b、アドレス変換回路 12c、PCIe インタフェース 12d、システムバス 12e を有する。

【0018】

スイッチ 13 は、PCIe インタフェース 13a、13b と、PCIe インタフェース 13a、13b を接続するシステムバス 13c を有する。

PCIe インタフェース 13a は、半導体集積回路 11 の PCIe インタフェース 11d と情報の送受信を行う。

【0019】

PCIe インタフェース 13b は、半導体集積回路 12 の PCIe インタフェース 12d と情報の送受信を行う。

図 1 に示すような情報処理システム 10 において、PCIe インタフェース 11d、12d、13a、13b が用いられる場合、PCIe インタフェース 11d、12d には、半導体集積回路 11、12 がエンドポイントとなるような設定がなされる。また、PCIe インタフェース 13a、13b には、スイッチ 13 がルートコンプレックスになるよう

10

20

30

40

50

な設定がなされる。

【0020】

なお、P C I e インタフェース 1 1 d , 1 2 d , 1 3 a , 1 3 b の代わりに、送信先からの応答を待たずに連続的な送信が可能なバスインタフェースの他の例である、U S B (Universal Serial Bus) インタフェースなどを用いてもよい。

【0021】

メモリ 1 4 , 1 5 はそれぞれ、半導体集積回路 1 1 , 1 2 に接続されたメモリであり、制御回路 1 1 a , 1 2 a が実行するプログラムや各種データ（受信データなども含む）を記憶する。メモリ 1 4 , 1 5 は、たとえば、D R A M (Dynamic Random Access Memory) などの揮発性の半導体メモリでもよいし、フラッシュメモリなどの揮発性のストレージでもよい。なお、メモリ 1 4 , 1 5 は、半導体集積回路 1 1 , 1 2 に含まれていてもよい。また、メモリ 1 4 , 1 5 は、それぞれ複数あってもよい。

10

【0022】

以下、半導体集積回路 1 1 をデータの送信元、半導体集積回路 1 2 をデータの送信先として、情報処理システム 1 0 の動作の一例を説明する。半導体集積回路 1 1 がデータの送信先、半導体集積回路 1 2 がデータの送信元である場合も同様の処理が行われる。

【0023】

なお、以下では半導体集積回路 1 1 を S O C 1 、半導体集積回路 1 2 を S O C 2 と呼ぶ場合もある。

制御回路 1 1 a は、半導体集積回路 1 2 へのデータの送信の際、図 1 に示すような S O C 1 のメモリマップ情報 1 1 m に基づいて、アドレスの指定を行う。

20

【0024】

メモリマップ情報 1 1 m には、自 S O C (つまり S O C 1) のメモリ 1 4 のメモリ空間のアドレスと、他 S O C (つまり S O C 2) のメモリ 1 5 のメモリ空間のアドレスが定義されている。さらに、メモリマップ情報 1 1 m には、自 S O C の割り込み設定レジスタ空間のアドレスと、他 S O C の割り込み設定レジスタ空間のアドレスが定義されている。割り込み設定レジスタ（図示せず）は、たとえば、割り込み制御回路 1 1 b , 1 2 b 内に設けられている。

【0025】

なお、半導体集積回路 1 2 が参照するメモリマップ情報 1 2 m においても、メモリマップ情報 1 1 m と同様に、自 S O C (つまり S O C 2) のメモリ 1 5 のメモリ空間のアドレスと、他 S O C (つまり S O C 1) のメモリ空間のアドレスが定義されている。さらに、メモリマップ情報 1 2 m には、自 S O C の割り込み設定レジスタ空間のアドレスと、他 S O C の割り込み設定レジスタ空間のアドレスが定義されている。

30

【0026】

制御回路 1 1 a は、半導体集積回路 1 2 へのデータの送信の際、そのデータ（送信データ）を書き込むメモリ 1 5 の、メモリマップ情報 1 1 m で定義されているアドレスを指定する。図 1 では、制御回路 1 1 a が、メモリマップ情報 1 1 m で定義されている他 S O C メモリ空間のアドレス a 1 を指定した例が示されている。

【0027】

アドレス変換回路 1 1 c は、制御回路 1 1 a がアドレス a 1 を指定した場合、たとえば、図示しない変換テーブルなどによって、S O C 2 のメモリマップ情報 1 2 m で定義されている S O C 2 のメモリ空間のアドレス a 3 に変換する。P C I e インタフェース 1 1 d は、変換したアドレス a 3 と、送信データとを出力（送信）する。アドレス a 3 と送信データは、スイッチ 1 3 により、半導体集積回路 1 2 に転送される。

40

【0028】

半導体集積回路 1 2 は、P C I e インタフェース 1 2 d を用いて、アドレス a 3 と送信データとを受信し、そのアドレスに対応するメモリ 1 5 の受信バッファ領域に送信データを書き込む。たとえば、アドレス変換回路 1 2 c は、図示しない変換テーブルなどによって、アドレス a 3 を、メモリ 1 5 の実際のアドレスに変換する。これにより、そのアドレ

50

スから始まるメモリ 15 の受信バッファ領域に送信データが書き込まれる。

【0029】

その後、半導体集積回路 11 の制御回路 11a は、半導体集積回路 12 の割り込み設定レジスタの、メモリマップ情報 11m で定義されているアドレスを指定する。図 1 では、制御回路 11a が、メモリマップ情報 11m で定義されている他 SOC レジスタ空間のアドレス a2 を指定した例が示されている。

【0030】

アドレス変換回路 11c は、制御回路 11a がアドレス a2 を指定した場合、たとえば、図示しない変換テーブルなどによって、SOC2 のメモリマップ情報 12m で定義されている SOC2 の割り込み設定レジスタ空間のアドレス a4 に変換する。PCIe インタフェース 11d は、変換したアドレス a4 と、割り込みの発生を示す割り込み設定値（たとえば、“1”）とを出力（送信）する。アドレス a4 と割り込み設定値とは、スイッチ 13 により、半導体集積回路 12 に転送される。

【0031】

半導体集積回路 12 は、PCIe インタフェース 12d を用いて、アドレス a4 と割り込み設定値とを受信し、そのアドレスに対応する割り込み設定レジスタに割り込み設定値を書き込む。たとえば、アドレス変換回路 12c は、図示しない変換テーブルなどによって、アドレス a4 を、割り込み設定レジスタの実際のアドレスに変換する。そのアドレスで指定される割り込み設定レジスタに割り込み設定値が書き込まれる。

【0032】

割り込み設定レジスタに、割り込みの発生を示す割り込み設定値が書き込まれると、割り込み制御回路 12b は、制御回路 12a に、割り込みが発生した旨を通知する。これにより、制御回路 12a は、送信データの受信（書き込み）の終了を判定し、メモリ 15 の受信バッファ領域に書き込まれた送信データを読み出す。そして、制御回路 12a は、たとえば、読み出した送信データを用いて、種々のアプリケーション処理を実行する。

【0033】

以上のように、第 1 の実施の形態の情報処理システム 10 において、半導体集積回路 11 は、メモリマップ情報 11m に基づいて、複数の半導体集積回路 11, 12 の 1 つであるデータ送信先（半導体集積回路 12）が用いるメモリ 15 のアドレス a1 を指定する。そして、半導体集積回路 11 は、そのアドレスを、データ送信先で参照されるメモリマップ情報 12m において定義されたメモリ 15 のアドレス a3 に変換する。そして、半導体集積回路 11 は、アドレス a3 と送信データとを、PCIe インタフェース 11d を用いて出力する。スイッチ 13 は、アドレス a3 と送信データとを、PCIe インタフェース 13a, 13b を用いてデータ送信先に転送する。データ転送先である半導体集積回路 12 は、PCIe インタフェースを用いて、アドレス a3 と送信データとを受信し、アドレス a3 に対応するメモリ 15 の受信バッファ領域に送信データの書き込みを行う。

【0034】

これによって、半導体集積回路 11 は、半導体集積回路 12 が用いるメモリ 15 のアドレスが定義されたメモリマップ情報 11m に基づいて、直接そのアドレスを指定した送信（書き込み）処理が可能となる。また、情報処理システム 10 では、PCIe インタフェース 11d, 12d, 13a, 13b が用いられているため、半導体集積回路 11 は、データ送信先からの応答を待つことなく、データ送信を連続的に行える。以上のことから、情報処理システム 10 の処理を高速化できる。

【0035】

また、スーパーコンピュータに用いられるような高価な高速ネットワークを用いなくても済み、比較的安価に密結合、高並列・高性能の CPU クラスタを実現できるようになる。また、消費電力の大きい高性能な CPU を使用しなくて済むため、情報処理システム 10 の消費電力を削減できる。

【0036】

また、半導体集積回路 11 は、半導体集積回路 12 の割り込み設定レジスタのアドレス

10

20

30

40

50

が定義されたメモリマップ情報 1 1 m に基づいて、直接そのアドレスを指定して割り込み設定値の設定を行う。そして、半導体集積回路 1 2 は、割り込みの発生を示す割り込み設定値が設定されると、受信（書き込み）が終了したと判定し、送信データの読み出しを行う。これによって、パケット受信を契機にデータ読み込み処理を行うようなイーサネットプロトコルを擬似的に再現できる。

【0037】

（第2の実施の形態）

図2は、第2の実施の形態の情報処理システムの一例を示す図である。

情報処理システム20は、半導体集積回路21a0, 21a1, 21an, 21an+1を含む複数の半導体集積回路と、DRAM22a0, 22a1, 22an, 22an+1を含む複数のDRAMを有する。さらに、2段のスイッチ23, 24を有する。

10

【0038】

半導体集積回路21a0は、CPU21b0、ICU（Interrupt Controller Unit）21c0、MMU（Memory Management Unit）21d0、PCIeインタフェース21e0、システムバス21f0を有する。また、半導体集積回路21a0には、DRAM22a0が接続されている。

【0039】

CPU21b0は、図1に示した制御回路11aの一例であり、システムバス21f0を介して半導体集積回路21a0の各部を制御する。

ICU21c0は、図1に示した割り込み制御回路11bの一例であり、割り込みの発生をCPU21b0に通知する。

20

【0040】

MMU21d0は、図1に示したアドレス変換回路11cの機能を有し、複数のDRAMのアドレスなどが定義されたメモリマップ情報に基づいて、アドレス変換を行う。

PCIeインタフェース21e0は、送信先からの応答を待たずに連続的な送信が可能なバスインタフェースの一例であり、スイッチ23と情報の送受信を行う。

【0041】

半導体集積回路21a1も、半導体集積回路21a0と同様に、CPU21b1、ICU21c1、MMU21d1、PCIeインタフェース21e1、システムバス21f1を有する。図示を省略しているが、その他の半導体集積回路も同様の要素を有する。

30

【0042】

スイッチ23は、PCIeインタフェース23a0, 23a1, ..., 23an, 23an+1、MMU23b0, 23b1, ..., 23bn, 23bn+1、システムバス23cを有する。スイッチ24は、PCIeインタフェース24a1, 24a2を含む複数のPCIeインタフェース、MMU24b1, 24b2を含む複数のMMU、システムバス24cを有する。

【0043】

スイッチ23において、PCIeインタフェース23a0~23an+1は、MMU23b0~23bn+1を介してシステムバス23cに接続されている。スイッチ24において、PCIeインタフェース24a1, 24a2もMMU24b1, 24b2を介してシステムバス24cに接続されている。

40

【0044】

PCIeインタフェース23a0は、半導体集積回路21a0のPCIeインタフェース21e0と情報の送受信を行う。PCIeインタフェース23a1は、半導体集積回路21a1のPCIeインタフェース21e1と情報の送受信を行う。PCIeインタフェース23anは、半導体集積回路21anのPCIeインタフェース（図示せず）と情報の送受信を行う。PCIeインタフェース23an+1は、スイッチ24のPCIeインタフェース24a1と情報の送受信を行う。PCIeインタフェース24a2は、半導体集積回路21an+1のPCIeインタフェース（図示せず）と情報の送受信を行う。

【0045】

50

図2に示すような情報処理システム20では、P C I eインタフェース21e0, 21e1には、半導体集積回路21a0, 21a1がエンドポイントとなるような設定がなされる。また、P C I eインタフェース23a0~23an, 24a1, 24a2には、スイッチ23, 24がルートコンプレックスになるような設定がなされる。また、P C I eインタフェース23an+1には、スイッチ23がエンドポイントとなるような設定がなされる。

【0046】

MMU23b0は、図示しない変換テーブルなどによって、P C I eインタフェース23a0が受信したアドレスを、別のアドレスに変換する機能を有する。MMU23b1は、図示しない変換テーブルなどによって、P C I eインタフェース23a1が受信したアドレスを、別のアドレスに変換する機能を有する。MMU23bnは、図示しない変換テーブルなどによって、P C I eインタフェース23anが受信したアドレスを、別のアドレスに変換する機能を有する。MMU23bn+1は、図示しない変換テーブルなどによって、P C I eインタフェース23an+1が受信したアドレスや、P C I eインタフェース23an+1が送信するアドレスを、別のアドレスに変換する機能を有する。MMU24b1は、図示しない変換テーブルなどによって、P C I eインタフェース23anが受信したアドレスを、別のアドレスに変換する機能を有する。MMU24b2は、図示しない変換テーブルなどによって、P C I eインタフェース24a2が受信したアドレスを、別のアドレスに変換する機能を有する。

【0047】

上記の情報処理システム20では、スイッチ23に、n+1個の半導体集積回路21a0~21anが接続されている。また、スイッチ24にも半導体集積回路21an+1を含む複数の半導体集積回路が接続されている。

【0048】

このため、半導体集積回路21a0がD R A M 22an+1にデータの書き込みを行う場合、半導体集積回路21a0が出力する2種類のアドレス、送信データ及び割り込み設定値は、スイッチ23, 24を介して半導体集積回路21an+1に転送される。

【0049】

たとえば、物理的な制約などによって、1つのスイッチ23に接続可能な半導体集積回路の最大数が決まっている場合、図2に示すように、スイッチ24を追加することにより、多数の半導体集積回路を含む情報処理システム20を実現できる。なお、図1では、2つのスイッチ23, 24を有する例が示されているが、スイッチは3つ以上設けられていてもよい。

【0050】

また、上記の情報処理システム20では、スイッチ23, 24にもMMU23b0~23bn+1, 24b1, 24b2が設けられている。

このようなMMU23b0~23bn+1, 24b1, 24b2を設けることで、通信の接続状況を変更することができる。

【0051】

たとえば、半導体集積回路21a0が、データ送信先（書き込み先）のアドレスとして、半導体集積回路21a1が参照するメモリマップ情報により定義されているD R A M 22a1のアドレスを出力する場合を考える。このとき、スイッチ23のMMU23b0は、そのアドレスを、たとえば、半導体集積回路21anが参照するメモリマップ情報により定義されているD R A M 22anのアドレスに変換すると、送信データをD R A M 22anに書き込ませることができる。MMU23b0は、割り込み設定レジスタのアドレスについても同様の変換を行うことができる。

【0052】

これにより、たとえば、半導体集積回路21a1とスイッチ23との接続がなくなった場合にも、他の半導体集積回路を変更せずに、データ送信先を変更できる。つまり通信の接続状況を変更できる。

【0053】

上記のような情報処理システム20において、半導体集積回路21a0～21an+1を含むN個の半導体集積回路は、たとえば、以下に示すようなメモリマップ情報に基づいて、書き込み先のアドレス指定を行う。メモリマップ情報は、たとえば、半導体集積回路21a0～21an+1を含むN個の半導体集積回路が用いるメモリ（DRAM22a0）、または、N個の半導体集積回路のそれぞれの中に設けられたROM（Read Only Memory）などに格納されていてもよい。

【0054】

なお、以下では、半導体集積回路21a0～21an+1を含むN個の半導体集積回路をSOC[0]～SOC[N]と表記する。たとえば、半導体集積回路21a0を、SOC[0]、半導体集積回路21a1を、SOC[1]と表記する。

10

【0055】

図3は、メモリマップ情報の一例を示す図である。

SOC[0]用のメモリマップ情報30a0には、ICUレジスタ空間30b0、ローカルDRAM空間30c0, 30d0, 30e0、PCIeバス空間30f0, 30g0のアドレスが定義されている。

【0056】

ICUレジスタ空間30b0のアドレスは、SOC[0]のICU21c0が割り込みの発生をCPU21b0に通知するか否かを決定するために用いる割り込み設定値を格納する割り込み設定レジスタのアドレスに対応している。割り込み設定レジスタは、たとえば、ICU21c0に設けられている。

20

【0057】

ローカルDRAM空間30c0～30e0のアドレスは、DRAM22a0のアドレスに対応している。図3の例では、ローカルDRAM空間30c0～30e0は3つある。ローカルDRAM空間30c0は、アドレス0x0000_8000_0000～0x0000_FFFF_FFFFで定義される。ローカルDRAM空間30d0は、アドレス0x0008_8000_0000～0x000F_FFFF_FFFFで定義される。DRAM空間30e0は、アドレス0x0088_0000_0000～0x008F_FFFF_FFFFで定義される。ローカルDRAM空間の数は3つに限定されない。

【0058】

PCIeバス空間30f0のアドレスは、PCIeバスで接続される他のSOC[1], SOC[2], ..., SOC[N]の割り込み設定レジスタのアドレスに対応している。図3の例では、PCIeバス空間30f0は、アドレス0x0006_0000_0000～0x0006_FFFF_FFFFで定義される。PCIeバス空間30f0のアドレスは、SOC[1]～SOC[N]用のメモリマップ情報30a1, 30a2, ..., 30aNで定義されるICUレジスタ空間30b1, 30b2～30bNのアドレスに対応している。

30

【0059】

PCIeバス空間30g0のアドレスは、PCIeバスで接続される他のSOC[1], SOC[2], ..., SOC[N]が用いるメモリ（DRAM22a1など）のアドレスに対応している。図3の例では、PCIeバス空間30g0は、アドレス0x0040_0000_0000～0x007F_FFFF_FFFFで定義される。SOC[1]～SOC[N]が用いるメモリのアドレスは、メモリマップ情報30a1～30aNではローカルDRAM空間30c1, 30d1, 30e1, 30c2, 30d2, 30e2, ..., 30cN, 30dN, 30eNのアドレスとして定義されている。PCIeバス空間30g0のアドレスは、メモリマップ情報30a1～30aNのローカルDRAM空間30e1～30eNのアドレスと対応付けられている。

40

【0060】

図3の例では、各メモリマップ情報30a0～30aNで定義されるローカルDRAM空間30e0～30eNは、他のSOCから送信される書き込みデータが書き込まれる受

50

信バッファ領域として機能する。つまり、ローカルDRAM空間30e0～30eNは、他のSOCからアクセス可能な共用領域である。

【0061】

なお、SOC[1]～SOC[N]用のメモリマップ情報30a1～30aNにおいても、PCIeバス空間が定義されているが、図示が省略されている。また、メモリマップ情報30a0～30aNには、他の周辺機器についてのメモリ空間のアドレスが定義されているともよい。

【0062】

図4は、各SOCのDRAM上の共用領域の一例を示す図である。図4では、8つのSOC[0]～SOC[7]が用いるDRAM(DRAM22a0～22an+1など)上の共用領域の一例が示されている。

【0063】

SOC[0]が用いるDRAM22a0上の共用領域には、SOC[1]～SOC[7]用の受信バッファ領域が順に設定されている。SOC[1]のDRAM22a1上の共用領域には、SOC[0]，SOC[2]～SOC[7]用の受信バッファ領域が順に設定されている。SOC[7]のDRAM上の共用領域には、SOC[0]～SOC[6]用の受信バッファ領域が順に設定されている。

【0064】

SOC[n]のDRAM22an上の共用領域のSOC[m]用の受信バッファ領域には、SOC[m]が送信する送信データが書き込まれる。各共用領域における受信バッファ領域は、SOC[0]～SOC[7]の受信バッファのそれぞれは、異なるアドレスで指定されるように設定されている。つまり、各共用領域における受信バッファの割り当ては固定であり、同一バッファに複数のSOCが書き込みを行うことはない。このため、他のSOCの送信データによる上書きを防止できる。

【0065】

なお、図4の例では、各共用領域は7分割されているが、N=63の場合には、63分割されるようにしてもよい。ただ、ソフトウェア実装の観点から扱いやすい、8分割また64分割としてもよく、DRAM容量に余裕がある場合、余る1領域は不使用状態で放置しておいてもよい。

【0066】

図5は、受信バッファ領域に格納されるパケットデータの一例を示す図である。図5には、図4に示したSOC[n]のDRAM上の共用領域のSOC[m]用の受信バッファ領域に格納されるパケットデータの一例が示されている。

【0067】

受信バッファ領域は、たとえば、それぞれ64KiBのpacket_buf[0]，packet_buf[1]，…，packet_buf[i]，…，packet_buf[last]に分割されている。

【0068】

packet_buf[0]～packet_buf[last]のそれぞれには、以下に示す情報を含むパケットデータが格納される。

パケットデータは、validフィールド、lengthフィールド、padding[2]フィールド、data[65526]フィールドを含む。

【0069】

validフィールドは、当該パケットデータが有効かつ、SOC[n]による受信処理待ちであるか否かを示すbool型の値を含む。validフィールドには、当該パケットデータが有効かつ、SOC[n]による受信処理待ちである場合には1、当該パケットデータが無効の場合には、0が格納される。

【0070】

lengthフィールドは、data[]フィールドに格納されたイーサネットフレームの有効サイズ[bytes]を示す32ビットの整数型の値を含む。

10

20

30

40

50

`padding [2]` フィールドは、`data []` フィールドに格納されたイーサネットフレームの IP (Internet Protocol) ヘッダの先頭を 4 byte アラインするためのパディングを示す 8 ビットの整数型の値を含む。

【0071】

`data [65526]` フィールドは、イーサネットフレームデータを示す 8 ビットの整数型の値を含む。なお、65526 は、`packet_buf [i]` のサイズである 64 KiB から逆算して得られるバイト数である。この結果、最大の MTU (Maximum Transmission Unit) は 65512 バイトとなる。

【0072】

以下、第 2 の実施の形態の情報処理システム 20 の動作の一例を、フローチャートを用いて説明する。

図 6 は、情報処理システムによる送信制御の一例の流れを示すフローチャートである。

【0073】

以下では、SOC [0] が、SOC [1] のメモリ (DRAM22a1) にデータの書き込みを行う場合を例にして説明する。

まず、SOC [0] の CPU21b0 は、書き込み先のアドレスとして、SOC [1] 用のメモリマップ情報 30a1 のローカル DRAM 空間 30e1 のアドレスに対応付けられている PCIe バス空間 30g0 のアドレスを指定する (ステップ S10)。

【0074】

なお、CPU21b0 は、図 5 に示した `packet_buf [0] ~ packet [last]` が昇順にアクセスされるようにアドレスを指定する。`packet [last]` がアクセスされた後は、`packet_buf [0]` がアクセスされるようにアドレスが指定される。

【0075】

図 3 に示したような PCIe バス空間 30g0 のアドレスは、MMU21d0 によって、送信先の SOC [1] が参照するメモリマップ情報 30a1 のローカル DRAM 空間 30e1 のアドレスに変換される (ステップ S11)。

【0076】

次に、CPU21b0 は、書き込み先の `valid` が 0 であることを確認する (ステップ S12)。

たとえば、CPU21b0 は、MMU21d0 が変換したアドレスに対する読み出し命令を発行する。そして、CPU21b0 は、そのアドレスに対応した SOC [1] の DRAM22a1 の共用領域にある SOC [0] 用の受信バッファ領域の `packet_buf [i]` の `valid` の値を、スイッチ 23 を介して読み出し、確認する。

【0077】

なお、図示を省略しているが、CPU21b0 は、`valid` の値が 1 の場合 (受信バッファ領域が `full` である場合)、たとえば、データ送信を一定期間中断し、一定期間後も `valid` の値が 1 である場合には、送信制御を終了する。

【0078】

次に、CPU21b0 は、書き込み先の `length`、`data []` (図 5 参照) にデータを設定する (データ送信) (ステップ S13)。

CPU21b0 は、`length`、`data []` として受信バッファ領域に書き込むデータを、MMU21d0 が変換したアドレスとともに、PCIe インタフェースに出力させる。アドレスとデータは、スイッチ 23 によって SOC [1] に転送される。アドレスは、SOC [1] の MMU21d1 にて、DRAM22a1 の実際のアドレスに変換され、SOC [0] 用の受信バッファ領域にデータが書き込まれる。

【0079】

その後、CPU21b0 は、`valid` を 1 に設定する (ステップ S14)。`valid` の設定は、ステップ S13 の処理と同様に行われる。

最後に、CPU21b0 は、受信先である SOC [1] の CPU21b1 に送信 (書き

10

20

30

40

50

込み)の終了を知らせるために、SOC [1]の割り込み設定レジスタに割り込み設定値を設定する(ステップS15)。

【0080】

たとえば、CPU21b0は、SOC [1]用のメモリマップ情報30a1のICUレジスタ空間30b1のアドレスに対応付けられているPCIeバス空間30f0のアドレスを指定する。指定されたアドレスは、MMU21d0によって、SOC [1]用のメモリマップ情報30a1のICUレジスタ空間30b1のアドレスに変換される。そして、そのアドレスと割り込みの発生を示す割り込み設定値(たとえば、“1”)が、CPU21b0の制御のもと、PCIeインタフェース21e0から出力される。アドレスと割り込み設定値は、スイッチ23によって、SOC [1]に転送される。アドレスは、SOC [1]のMMU21d1にて、割り込み設定レジスタの実際のアドレスに変換され、割り込み設定レジスタに割り込み設定値が書き込まれる。

10

【0081】

図7は、情報処理システムによる受信制御(読み出し制御)の一例の流れを示すフローチャートである。

以下では、SOC [1]が、受信したデータを読み出す制御について説明する。

【0082】

SOC [1]のCPU21b1は、ICU21c1から割り込みの発生を通知されると受信制御(読み出し制御)を開始する。

CPU21b1は、まず、DRAM22a1の共用領域の各SOCの受信バッファ領域の読み出し位置のvalidが0であるか否かを判定する(ステップS20)。

20

【0083】

validが0である場合には、図5に示したように、パケットデータは無効であるので、CPU21b1は受信制御を終了する。

validが1である場合、CPU21b1は、ステップS21の処理を行う。ステップS21では、図5に示したようなdata[]フィールドに格納されたイーサネットフレームデータを、たとえば、OS(Operating System)のソケットバッファ(sk__buffと表記されている)構造体にコピーする読み出し処理が行われる。

【0084】

その後、CPU21b1は、validを0に設定し(ステップS22)、ソケットバッファ構造体をTCP(Transmission Control Protocol)/IPスタックに引き渡す(ステップS23)。

30

【0085】

その後、ステップS20の処理に戻り、次の読み出し位置のvalidが0であるか否かが判定され、同様の処理が繰り返される。

以上のように、SOC [0]~SOC [N]のそれぞれは、他のSOCのメモリ(DRAM22a1など)のアドレスが定義されたメモリマップ情報30a0~30aNに基づいて、直接そのアドレスを指定した送信及び書き込み処理が可能である。また、情報処理システム20では、PCIeインタフェース21e0, 21e1, 23a0~23an+1, 24a1, 24a2が用いられている。そのため、SOC [0]~SOC [N]のそれぞれは、データ送信先からの応答を待つことなく、データ送信を連続的に行える。以上のことから、情報処理システム20の処理を高速化できる。

40

【0086】

また、スーパーコンピュータに用いられるような高価な高速ネットワークを用いなくても済み、比較的安価に密結合、高並列・高性能のCPUクラスタを実現できるようになる。また、消費電力の大きい高性能なCPUを使用しなくて済むため、情報処理システム20の消費電力を削減できる。

【0087】

また、SOC [0]~SOC [N]のそれぞれは、他のSOCの割り込み設定レジスタのアドレスが定義されたメモリマップ情報30a0~30aNに基づいて、直接そのアド

50

レスを指定して割り込み設定値の設定を行う。そして、データ送信先のSOCは、割り込みの発生を示す割り込み設定値が設定されると、受信（書き込み）が終了したと判定し、送信データの読み出しを行う。これによって、パケット受信を契機にデータ読み込み処理を行うようなイーサネットプロトコルを擬似的に再現できる。

【0088】

また、第2の実施の形態の情報処理システム20は、スイッチ23, 24にも、アドレス変換回路として機能するMMU23b0~23bn+1, 24b1, 24b2を有している。これにより、SOC[0]~SOC[N]の構成を変えずに、書き込み先の変更など、通信の接続状況を変更することができる。

【0089】

また、図2のようにスイッチ23, 24を複数段設けることで、1つのスイッチに接続できるSOCの数に物理的な制約がある場合でも、より多くのSOCを含む情報処理システム20を実現することができる。

【0090】

（第3の実施の形態）

図8は、第3の実施の形態の情報処理システムの一例を示す図である。図8において、図2に示した第2の実施の形態の情報処理システム20と同じ要素については同一符号が付されている。また、以下では、半導体集積回路41a0~41an+1を含むN個の半導体集積回路をSOC[0]~SOC[N]と表記する。たとえば、半導体集積回路41a0を、SOC[0]、半導体集積回路41a1を、SOC[1]と表記する。

【0091】

第3の実施の形態の情報処理システム40において、SOC[0]~SOC[N]のうちの2つ以上には、DRAMとは異なるメモリが接続されている。図8の例では、SOC[0]にはメモリ42a0が接続され、SOC[n]にはメモリ42anが接続され、SOC[n+1]にはメモリ42an+1が接続されている。

【0092】

メモリ42a0~42an+1のそれぞれは、DRAM22a0~22an+1のそれぞれよりも、容量が大きい情報記憶装置である。メモリ42a0~42an+1は、たとえば、SSD（Solid State Drive）メモリである。

【0093】

メモリ42a0と接続するためのインタフェースの一例として、SOC[0]は、PCIeインタフェース41b0を有する。PCIeインタフェース41b0は、MMU41c0を介して、バス21f0に接続される。すなわち、メモリ42a0は、スイッチ23がPCIeインタフェース21e0とMMU21d0を介してバス21f0に接続されるのと同様に、バス21f0に接続される。MMU41c0は、MMU21d0と同様の機能を有しているが、PCIeインタフェース41b0がメモリ42a0に接続される場合であって、たとえば、メモリ42a0へのアクセス（ライト動作またはリード動作）がSOC[0]内のCPU21b0からのアクセスに限定される場合は、その機能が無効となってもよく、また、MMU41c0自体がなくてもよい。

【0094】

メモリ42a0は、CPU21b0が出力するコマンドによって、ライト動作またはリード動作を行う。

なお、MMU41c0が、MMU21d0と同様の機能を有している場合、スイッチ23がPCIeインタフェース41b0に接続され、メモリ42a0がPCIeインタフェース21e0に接続されていてもよい。

【0095】

図示を省略しているがSOC[n]、SOC[n+1]なども、SOC[0]と同様の回路構成である。

上記のような情報処理システム40によれば、個々のSOCが扱えるデータサイズを増やせるため、より高効率な分散処理が可能となる。また、情報処理システム40は、高速

10

20

30

40

50

処理が可能な第2の実施の形態の情報処理システム20をベースとしているため、その利点を生かして、大規模なデータを高速に処理できる。

【0096】

(第4の実施の形態)

図9は、第4の実施の形態の情報処理システムの一例を示す図である。図9において、図8に示した第3の実施の形態の情報処理システム40と同じ要素については同一符号が付されている。

【0097】

第4の実施の形態の情報処理システム50において、SOC[0]～SOC[N]のうちの2つ以上には、特定機能向けLSI (Large Scale Integrated circuit) が接続されている。図9の例では、SOC[0]には特定機能向けLSI51a0が接続され、SOC[n]には特定機能向けLSI51anが接続され、SOC[n+1]には特定機能向けLSI51an+1が接続されている。

10

【0098】

特定機能向けLSI51a0～51an+1のそれぞれは、たとえば、映像データのエンコード処理など、負荷が比較的重いタスクを行うLSIであり、たとえば、ASSP (Application Specific Standard Product) である。

【0099】

特定機能向けLSI51a0は、たとえば、インタフェースの一例であるPCIeインタフェース41b0と、MMU41c0を介して、バス21f0に接続される。MMU41c0は、MMU21d0と同様の機能を有しているが、PCIeインタフェース41b0が特定機能向けLSI51a0に接続される場合であって、たとえば、特定機能向けLSI51a0の動作がSOC[0]内のCPU21b0が出力する制御信号に基づく動作に限定される場合は、その機能が無効となってもよく、また、MMU41c0自体がなくてもよい。

20

【0100】

なお、MMU41c0が、MMU21d0と同様の機能を有している場合、スイッチ23がPCIeインタフェース41b0に接続され、特定機能向けLSI51a0がPCIeインタフェース21e0に接続されていてもよい。

【0101】

特定機能向けLSI51a0は、CPU21b0が出力する制御信号に基づいて動作する。特定機能向けLSI51a0が、たとえば、HEVC (High Efficiency Video Coding) エンコーダである場合には、特定機能向けLSI51a0は、動画データデータをSOC[0]から受け、その動画データを圧縮したHEVCデータを生成して出力する。

30

【0102】

特定機能向けLSI51an, 51an+1なども同様にSOC[n], SOC[n+1]に接続され、SOC[n], SOC[n+1]のCPUが出力する制御信号に基づいて動作する。

【0103】

上記のような情報処理システム50によれば、個々のSOCに接続された各特定機能向けLSIが並列処理を行うことで、映像データの解析やメタデータの抽出、統合などの負荷が比較的重いタスクをより高効率に処理可能となる。

40

【0104】

なお、上記第3の実施の形態の情報処理システム40と第4の実施の形態の情報処理システム50は、互いに組み合わせることもできる。すなわち、SOC[0]～SOC[N]のうち、あるSOCにはSSDメモリなどのメモリが接続され、他のSOCには特定機能向けLSIが接続されてもよい。

【0105】

また、SOC[0]～SOC[n+1]において、PCIeインタフェースなどのインタフェースをさらに追加して、SSDメモリなどのメモリと、特定機能向けLSIの両方

50

が接続されるようにしてもよい。

【0106】

上記については単に本発明の原理を示すものである。さらに、多数の変形、変更が当業者にとって可能であり、本発明は上記に示し、説明した正確な構成及び応用例に限定されるものではなく、対応する全ての変形例及び均等物は、添付の請求項及びその均等物による本発明の範囲とみなされる。

【符号の説明】

【0107】

10 情報処理システム

11, 12 半導体集積回路

10

11a, 12a 制御回路

11b, 12b 割り込み制御回路

11c, 12c アドレス変換回路

11d, 12d, 13a, 13b PCIEインタフェース

11e, 12e, 13c システムバス

11m, 12m メモリマップ情報

a1～a4 アドレス

13 スイッチ

14, 15 メモリ

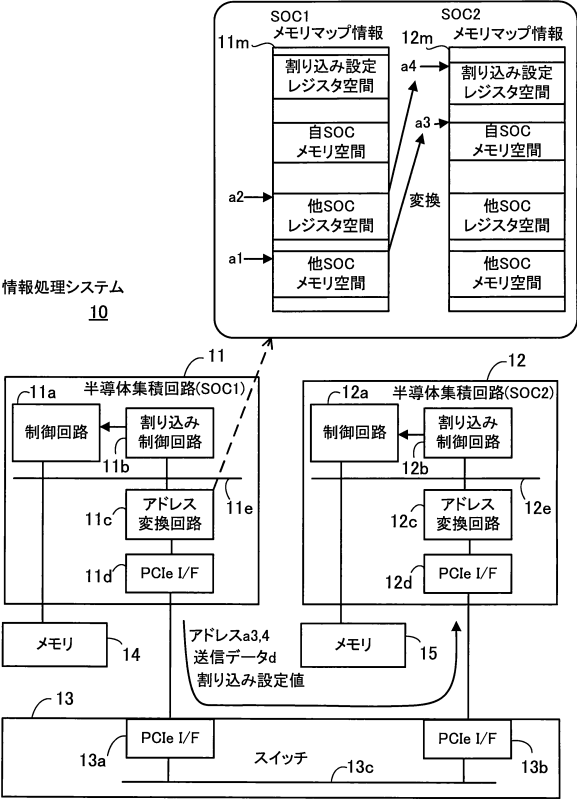
20

30

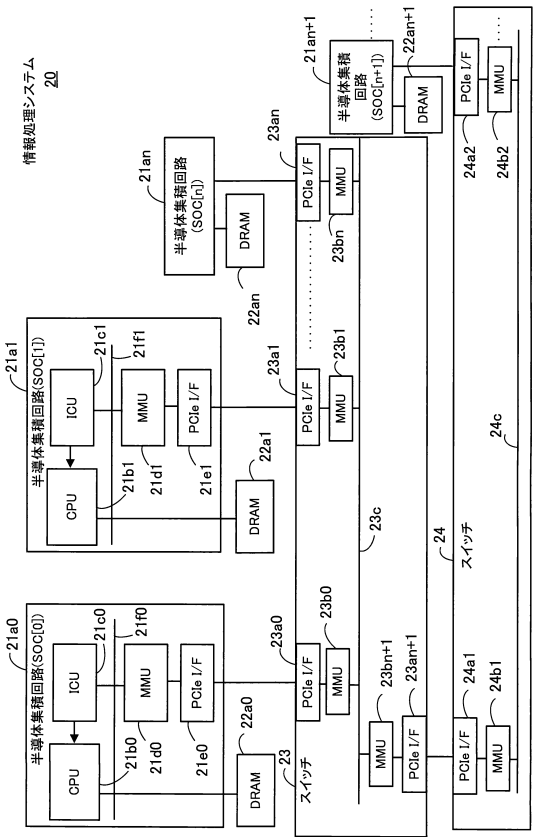
40

50

【図面】
【図 1】



【図 2】



10

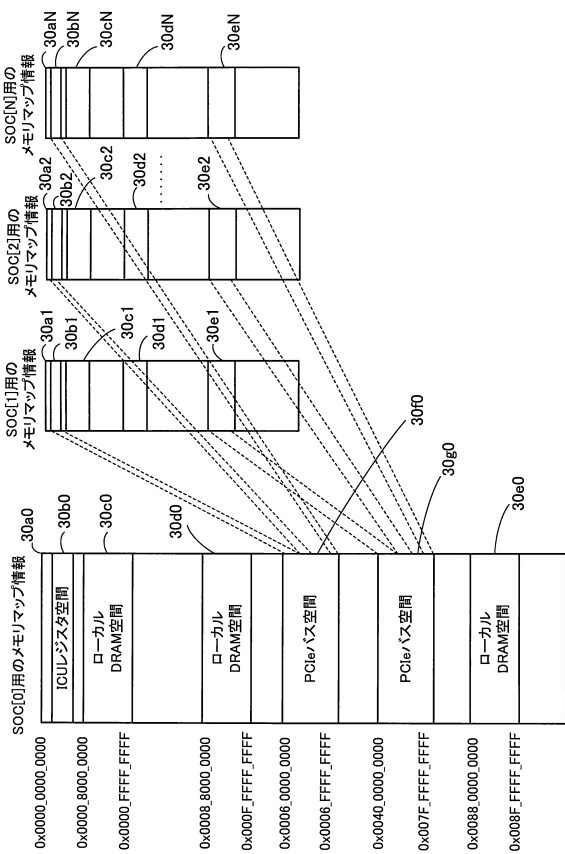
20

30

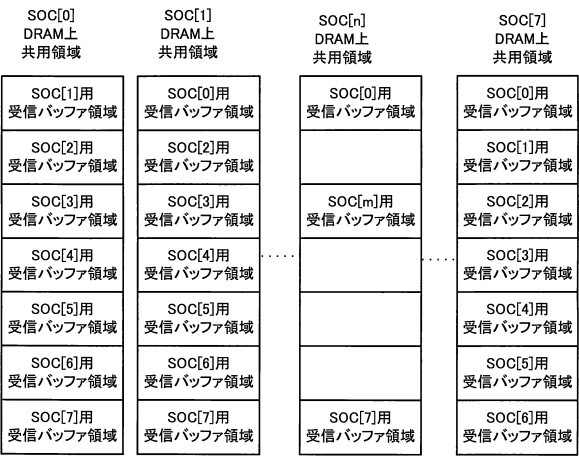
40

50

【図 3】



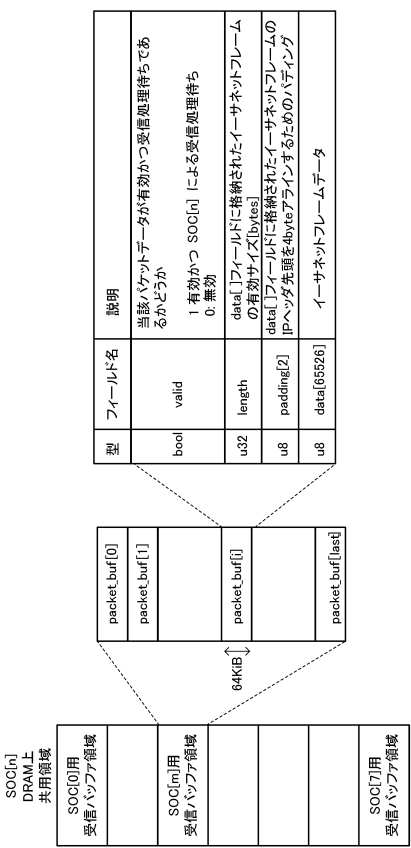
【図 4】



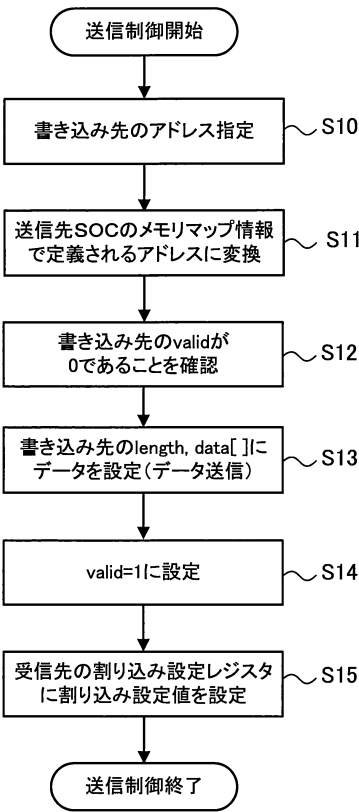
10

20

【図 5】



【図 6】

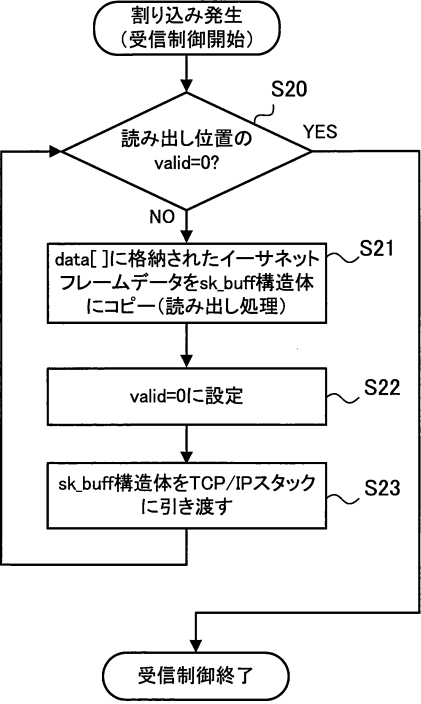


30

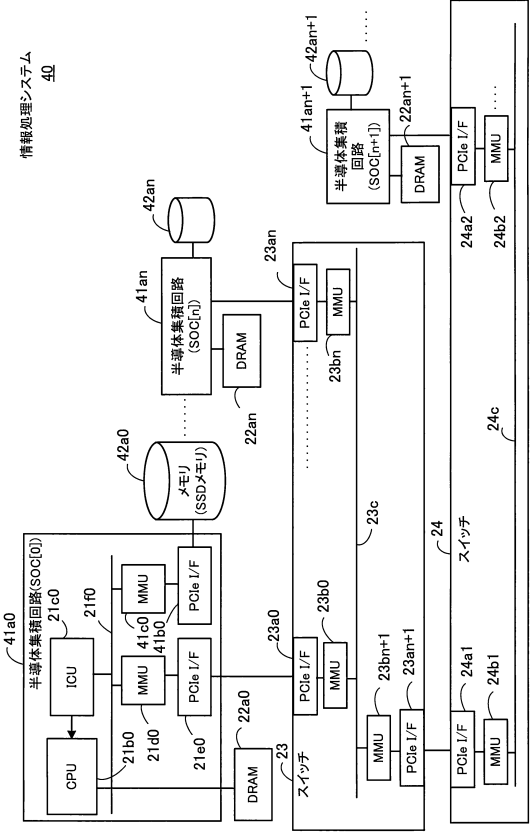
40

50

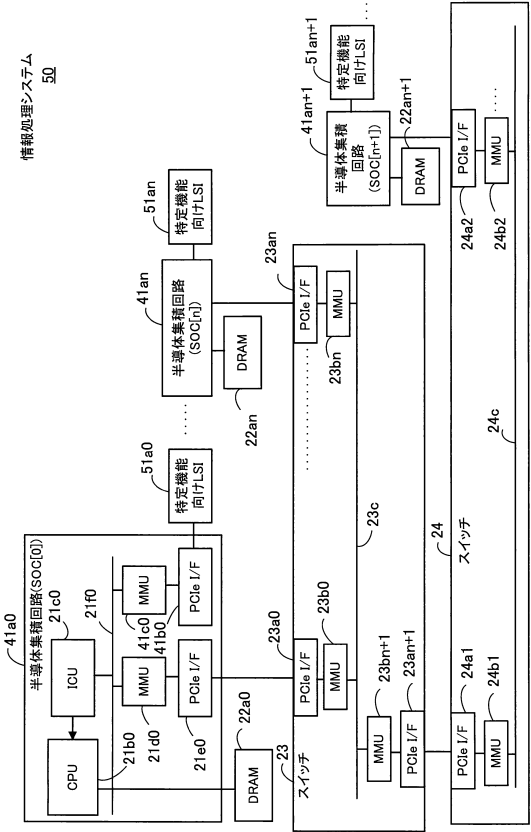
【図 7】



【図 8】



【図 9】



フロントページの続き

審査官 田名網 忠雄

- (56)参考文献 国際公開第2007/105373 (WO, A1)
特開2012-128717 (JP, A)
米国特許出願公開第2015/0067229 (US, A1)
欧州特許出願公開第02515241 (EP, A2)
- (58)調査した分野 (Int.Cl., DB名)
G06F 13/36-13/38