

(12) 特許協力条約に基づいて公開された国際出願

(19) 世界知的所有権機関
国際事務局

(43) 国際公開日
2021年6月10日(10.06.2021)



(10) 国際公開番号

WO 2021/111490 A1

(51) 国際特許分類:

G06N 3/08 (2006.01)

(21) 国際出願番号:

PCT/JP2019/046966

(22) 国際出願日:

2019年12月2日(02.12.2019)

(25) 国際出願の言語:

日本語

(26) 国際公開の言語:

日本語

(71) 出願人: 日本電信電話株式会社 (NIPPON TELEGRAPH AND TELEPHONE CORPORATION) [JP/JP]; 〒1008116 東京都千代田区大手町一丁目5番1号 Tokyo (JP).

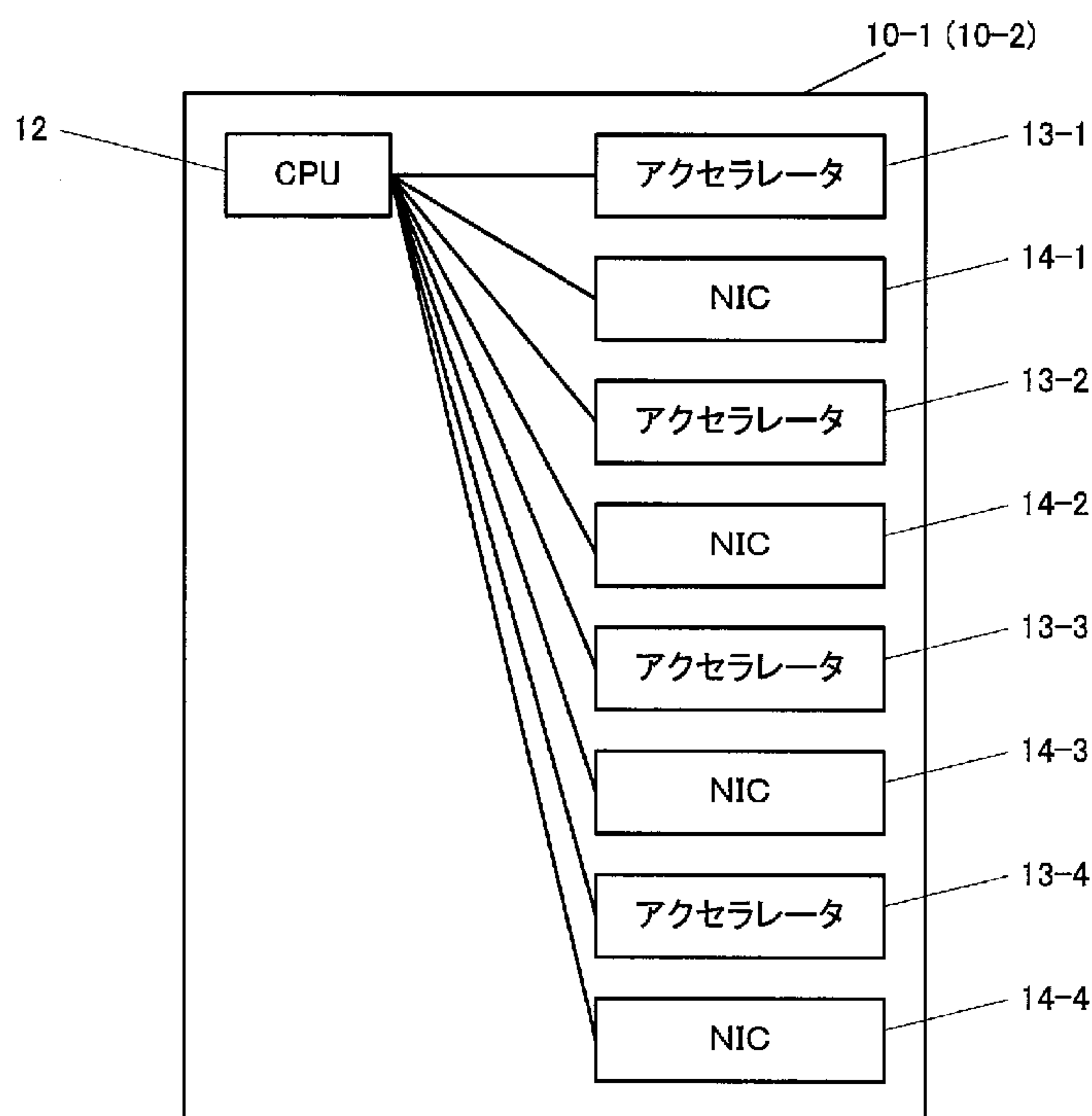
(72) 発明者: 田仲 顕至(TANAKA, Kenji); 〒1808585 東京都武蔵野市緑町3丁目9-11 NTT 知的財産センタ内 Tokyo (JP). 有川 勇輝

(ARIKAWA, Yuki); 〒1808585 東京都武蔵野市緑町3丁目9-11 NTT 知的財産センタ内 Tokyo (JP). 伊藤 猛(ITO, Tsuyoshi); 〒1808585 東京都武蔵野市緑町3丁目9-11 NTT 知的財産センタ内 Tokyo (JP). 寺田 和彦(TERADA, Kazuhiko); 〒1808585 東京都武蔵野市緑町3丁目9-11 NTT 知的財産センタ内 Tokyo (JP). 坂本 健(SAKAMOTO, Takeshi); 〒1808585 東京都武蔵野市緑町3丁目9-11 NTT 知的財産センタ内 Tokyo (JP).

(74) 代理人: 山川 茂樹, 外(YAMAKAWA, Shigeki et al.); 〒1006104 東京都千代田区永田町2丁目11番1号 山王パークタワー4階 山川国際特許事務所内 Tokyo (JP).

(54) Title: DISTRIBUTED PROCESS CALCULATOR AND DISTRIBUTED DEEP LEARNING SYSTEM

(54) 発明の名称: 分散処理用計算機および分散深層学習システム



13-1, 13-2, 13-3, 13-4 Accelerator

(57) Abstract: In order to perform distributed deep learning in a shorter time period by using a calculator provided with a plurality of accelerators, a distributed process calculator (10-1, 10-2) is provided with: a plurality of accelerators (13-1 to 13-4) each of which is assigned with a part of a neural network, and is configured to derive a learning result on the basis of input data and to update a parameter value included in the part of the neural network by using the learning result; a plurality of network interface circuits (14-1 to 14-4) each configured to transmit/receive information concerning learning

(81) 指定国(表示のない限り、全ての種類の国内保護が可能): AE, AG, AL, AM, AO, AT, AU, AZ, BA, BB, BG, BH, BN, BR, BW, BY, BZ, CA, CH, CL, CN, CO, CR, CU, CZ, DE, DJ, DK, DM, DO, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, GT, HN, HR, HU, ID, IL, IN, IR, IS, JO, JP, KE, KG, KH, KN, KP, KR, KW, KZ, LA, LC, LK, LR, LS, LU, LY, MA, MD, ME, MG, MK, MN, MW, MX, MY, MZ, NA, NG, NI, NO, NZ, OM, PA, PE, PG, PH, PL, PT, QA, RO, RS, RU, RW, SA, SC, SD, SE, SG, SK, SL, SM, ST, SV, SY, TH, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, ZA, ZM, ZW.

(84) 指定国(表示のない限り、全ての種類の広域保護が可能): ARIPO (BW, GH, GM, KE, LR, LS, MW, MZ, NA, RW, SD, SL, ST, SZ, TZ, UG, ZM, ZW), ユーラシア (AM, AZ, BY, KG, KZ, RU, TJ, TM), ヨーロッパ (AL, AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HR, HU, IE, IS, IT, LT, LU, LV, MC, MK, MT, NL, NO, PL, PT, RO, RS, SE, SI, SK, SM, TR), OAPI (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, KM, ML, MR, NE, SN, TD, TG).

添付公開書類:

一 国際調査報告 (条約第21条(3))

including the learning result via a network; and an arithmetic processing device (12) configured to control the plurality of accelerators and the plurality of network interface circuits so as to cause each of the plurality of accelerators to derive a learning result on the basis of input data and to cause the plurality of network interface circuits to transmit/receive in parallel information concerning learning including a plurality of the learning results derived by the plurality of accelerators.

(57) 要約: 複数のアクセラレータを備えた計算機を用いてより短い時間で分散深層学習を行うために、分散処理用計算機(10-1, 10-2)は、それぞれにニューラルネットワークの一部が割り付けられ、入力データに基づいて学習結果を導出するとともに、前記学習結果を用いて前記ニューラルネットワークの一部に含まれるパラメータの値をそれぞれ更新するように構成された、複数のアクセラレータ(13-1~13-4)と、それぞれがネットワークを介して前記学習結果を含む学習に関する情報を送受信するように構成された複数のネットワークインターフェース回路(14-1~14-4)と、前記複数のアクセラレータおよび前記複数のネットワークインターフェース回路を制御して、前記複数のアクセラレータのそれぞれに入力データに基づいて学習結果を導出させるとともに、前記複数のネットワークインターフェース回路に、前記複数のアクセラレータにそれぞれ導出させた前記学習結果を含む学習に関する情報を並列に送受信させるように構成された、演算処理装置(12)とを備える。

明 細 書

発明の名称：分散処理用計算機および分散深層学習システム

技術分野

[0001] 本発明は、分散処理用計算機および分散深層学習システムに関し、特に、演算処理装置（セントラルプロセッシングユニット：CPU）と、複数のアクセラレータと、ネットワークインターフェース回路とを有する分散処理用計算機、および複数の分散処理用計算機を用いた分散深層学習システムに関する。

背景技術

[0002] 多層のニューラルネットワークによる深層学習では、順伝播（Feed forward）と逆伝播（Back propagation）とを交互に行うことで、入力データに適合したモデルを学習する。順伝播および逆伝播を効率的に行うために、近年ではGPU（Graphics Processing Unit）などのアクセラレータが用いられる。

[0003] さらに、近年では、データサイズとモデルサイズとが増大したために、これらを分散処理する方法が提案されている。

[0004] たとえば、深層学習を1台の計算機で処理しようとした場合には、入力データサイズが増大したことによって、ストレージやI/Oのボトルネックが発生してしまう。そこで、複数台の計算機の間でデータを分散させて学習を行う「データ並列分散深層学習」が提案されている。

[0005] また、モデルサイズが増大すると、1つのアクセラレータのデバイスのメモリにモデルが記録しきれなくなり、主記憶にモデルを記憶する必要が生じる。すると、アクセラレータにおいてモデルデータの読み書きにより遅延が発生してしまう。そこで、複数台のアクセラレータを用い、モデルを分散させて学習を行う「パイプライン並列分散深層学習」が提案されている。

[0006] 並列分散深層学習においては、上述した複数の計算機による分散処理（データ並列）と計算機内での分散処理（パイプライン並列）とを併用すること

もできる。

[0007] ところで、データ並列分散深層学習では、複数の計算機のそれぞれは、分割されたデータに基づいて計算機毎に順伝播および逆伝播を行って、個別にモデル（レプリカ）の重みデータを導出し、その個別に導出された逆伝播後の重みデータを通信により全計算機間で共有する。この重みデータの共有は、「A l l r e d u c e」という集団通信である。このA l l r e d u c e処理は、計算機毎に得られた逆伝播後の重みデータを合計（R e d u c e）し、全計算機に分配（B r o a d c a s t）する処理を含む。各計算機は、A l l r e d u c e処理後の共有された重みデータ（多くの場合においては、計算機のそれぞれにおいて導出された重みデータの平均）をもって各パラメータの値を更新する。このA l l r e d u c eは、データ並列分散深層学習において重要な役割を持つ反面、ボトルネックになることが知られている。

[0008] 一方、パイプライン並列分散深層学習では、多層ニューラルネットの各層をそれぞれひとつのアクセラレータに割り付ける。各アクセラレータは、担当する層の処理が終了次第、自らよりも下層を担当しているアクセラレータに計算結果を転送し、次のデータの処理にとりかかる。このような処理はパイプライン処理と呼ばれる。

[0009] 図12は、上述した並列分散深層学習に用いられる計算機の構成例を示している。

図12に示すように、計算機100は、CPU120と、複数のアクセラレータ130-1～130-4と、ネットワークインターフェースカード（NIC）140とを備えている。ここで複数のアクセラレータ130-1～130-4は、それぞれニューラルネットワークの複数の層にそれぞれ対応している。

[0010] このような計算機100では、学習用に入力されたデータに基づいて、アクセラレータ130-1～130-4がそれぞれ順伝播処理と逆伝播処理を行い、最終的にニューラルネットワークのパラメータの値を算出する。図

13に示すように、この例において計算機100は、複数のアクセラレータ130-1~130-4によって4つのデータをまとめて処理している。これらまとめて処理される複数のデータは「ミニバッチ」と呼ばれる。すなわち、計算機100は、4つのデータを含むミニバッチが入力されるごとに、各データに対してパイプライン並列処理により順次、順伝播処理を行い、逆伝播処理を行い、データごとに導出された学習結果の平均を算出している。

[0011] なお、図13等において、「 F_{ij} 」は、ニューラルネットワークの*i*番目の層に対応するアクセラレータ（*i*番目のアクセラレータ）による、ミニバッチに含まれる*j*番目のデータに基づく順伝播処理を表し、「 B_{ik} 」は、*i*番目のアクセラレータによる*k*番目のデータに対する出力結果に基づく逆伝播処理を表している。例えば、図13において、「 F_{11} 」は、ニューラルネットワークの第1層が割り付けられたアクセラレータ1による、ミニバッチに含まれる1番目のデータに基づく順伝播処理を表し、「 F_{21} 」は、順伝播処理 F_{11} の結果に基づいてニューラルネットワークの第2層が割り付けられたアクセラレータ2が行う順伝播処理を表す。同様に、「 B_{41} 」は、ニューラルネットワークの第4層（出力層）が割り付けられたアクセラレータ4による順伝播処理 F_{41} の結果に基づいて、同じアクセラレータ4が行う逆伝播処理を表し、「 B_{31} 」は、逆伝播処理 B_{41} の結果に基づき、ニューラルネットワークの第3層が割り付けられたアクセラレータ3が行う逆伝播処理を表している。

[0012] 一般的なパイプライン並列では、ミニバッチに含まれるデータごとに導出された学習結果の平均を算出した後、この学習結果の平均を用いて各パラメータを更新させて、次のミニバッチの処理を開始するが、パイプライン並列とデータ並列とを併用する場合には、ミニバッチによる学習結果の平均処理とパラメータの更新処理との間に、ネットワークを介して通信可能に接続された他の計算機との間でAll reduce処理が実行される。

先行技術文献

非特許文献

- [0013] 非特許文献1 : Huang, Yanping, et al. "Gpipe: Efficient training of giant neural networks using pipeline parallelism." arXiv preprint arXiv:1811.06965 (2018)

発明の概要

発明が解決しようとする課題

- [0014] しかしながら、アクセラレータの数に対してN I Cの数が少ない場合には、学習結果が一挙にN I Cに到来するために輻輳が発生し、A l l r e d u c e処理が深刻なボトルネックとなってしまう。
- [0015] そこで、本発明は、複数のアクセラレータを備えた計算機を用いてより短い時間で分散深層学習を行うことができる計算機、分散深層学習システムおよび方法を提供することを目的とする。

課題を解決するための手段

- [0016] 上記の目的を達成するために、本発明に係る分散処理用計算機（1 0 - 1 , 1 0 - 2）は、それぞれにニューラルネットワークの一部分が割り付けられ、入力データに基づいて学習結果を導出するとともに、前記学習結果を用いて前記ニューラルネットワークの一部分に含まれるパラメータの値をそれぞれ更新するように構成された、複数のアクセラレータ（1 3 - 1 ~ 1 3 - 4）と、それぞれがネットワークを介して前記学習結果を含む学習に関する情報を送受信するように構成された複数のネットワークインターフェース回路（1 4 - 1 ~ 1 4 - 4）と、前記複数のアクセラレータおよび前記複数のネットワークインターフェース回路を制御して、前記複数のアクセラレータのそれぞれに入力データに基づいて学習結果を導出させるとともに、前記複数のネットワークインターフェース回路に、前記複数のアクセラレータにそれぞれ導出させた前記学習結果を含む学習に関する情報を並列に送受信させるように構成された、演算処理装置（1 2）とを備える。
- [0017] また、本発明に係る分散深層学習システムは、ネットワークを介して互いに通信可能に接続された複数の計算機（1 0 - 1 , 1 0 - 2）を備え、前記複数の計算機のそれぞれは、上記の分散処理用計算機であり、前記複数の計

算機のうちの一の計算機の前記複数のネットワークインターフェース回路（ $14-1 \sim 14-4$ ）は、前記一の計算機の複数のアクセラレータがそれぞれ導出した複数の学習結果と前記複数の計算機の中の他の計算機の複数のアクセラレータがそれぞれ導出した複数の学習結果とを並列に送受信し、前記一の計算機の前記複数のアクセラレータ（ $13-1 \sim 13-4$ ）は、前記複数のアクセラレータが導出した前記学習結果と、前記複数のネットワークインターフェース回路を介して前記他の計算機から受信した前記複数の学習結果とに基づいて前記パラメータの値を更新する。

発明の効果

[0018] 本発明によれば、複数のアクセラレータがそれぞれ導出した学習結果を含む学習に関する情報を複数のネットワークインターフェース回路によって並列に送受信するので、より短い時間で分散深層学習を行うことができる。

図面の簡単な説明

[0019] [図1]図1は、本発明の第1の実施の形態に係る分散深層学習システムの概要を示す図である。

[図2]図2は、本発明の第1の実施の形態に係る分散深層学習システムを構成する分散処理用計算機の構成を示す図である。

[図3]図3は、ニューラルネットワークと分散処理用計算機との関係を説明するための図である。

[図4]図4は、本発明の第1の実施の形態に係る分散深層学習システムを構成する分散処理用計算機の動作を説明するための図である。

[図5]図5は、本発明の第1の実施の形態の変形例に係る分散処理用計算機の構成を示す図である。

[図6]図6は、本発明の第1の実施の形態の変形例に係る分散処理用計算機の動作を説明するための図である。

[図7]図7は、本発明の第2の実施の形態に係る分散処理用計算機の動作の一例を説明するための図である。

[図8A]図8Aは、本発明の第3の実施の形態に係る分散処理用計算機の構成

を示す図である。

[図8B]図8Bは、本発明の第3の実施の形態に係る分散処理用計算機において用いられるFPGA NICの構成を概念的に示す図である。

[図9]図9は、本発明の第3の実施の形態に係る分散処理用計算機の動作の一例を説明するための図である。

[図10A]図10Aは、本発明の第4の実施の形態に係る分散処理用計算機の構成を示す図である。

[図10B]図10Bは、本発明の第4の実施の形態に係る分散処理用計算機において用いられるFPGA NICの構成を概念的に示す図である。

[図11]図11は、本発明の第4の実施の形態に係る分散処理用計算機の動作の一例を説明するための図である。

[図12]図12は、従来の分散処理用計算機の構成を示す図である。

[図13]図13は、従来の分散処理用計算機の動作の一例を説明するための図である。

発明を実施するための形態

[0020] 以下、本発明の実施の形態について図面を参照して説明する。

[0021] [第1の実施の形態]

[分散深層学習システムおよび分散処理用計算機の構成]

本発明の第1の実施の形態に係る分散深層学習システムは、互いに通信可能に接続された複数の分散処理用計算機の間でデータを分散させて学習を行うデータ並列深層学習システムである。図1に示すように、分散深層学習システム1は、通信ネットワーク50を介して互いに通信可能に接続された2台の分散処理用計算機10-1, 10-2から構成されている。

[0022] 分散処理用計算機10-1, 10-2は、互いに同一のハードウェア構成を有している。図2に示すように、分散処理用計算機10-1, 10-2は、それぞれ、複数のアクセラレータ13-1~13-4と、複数のネットワークインターフェースカード（以下、「NIC」という。）14-1~14-4と、これらアクセラレータ13-1~13-4およびNIC14-1~

14-4を制御する中央演算処理装置（以下、「CPU」という。）12を備えている。

[0023] ここで、アクセラレータ13-1～13-4としては、例えば、GPUや、FPGA（Field-Programmable Gate Array）、DSP（Digital Signal Processor）などを用いることができる。

また、NIC14-1～14-4は、それぞれ分散処理用計算機10-1，10-2を通信ネットワーク50に接続するためのカード型の拡張装置である。

なお、図1においては、アクセラレータ13-1～13-4とNIC14-1～14-4とは、CPU12を介して接続されているが、これは、PCI Express（PCIe）を用いた接続の例である。この他に、CPUを介さずにPCIeスイッチを介してアクセラレータ13-1～13-4とNIC14-1～14-4とを接続してもよい。

[0024] 分散処理用計算機10-1，10-2は、それぞれ、複数の中間層を有する深層ニューラルネットワークをエミュレートする装置である。より具体的には、図3に示すように、深層ニューラルネットワークDNNはレイヤー（層）単位でいくつかに分割されて、複数の層が、分散処理用計算機10-1，10-2それぞれのアクセラレータ13-1～13-4にそれぞれ割り付けられている。本実施の形態においては、分割された各層が本発明における「ニューラルネットワークの一部」に相当する。なお、簡単のため、図3においては、分散処理用計算機10-1，10-2のCPU12は図示されていない。

[0025] 複数のNIC14-1～14-4は、それぞれ、複数のアクセラレータ13-1～13-4のいずれかと対応づけられている。本実施の形態においては、アクセラレータ13-1～13-4と同数のNIC14-1～14-4が設けられ、NIC14-1～14-4とアクセラレータ13-1～13-4とは、互いに1対1に対応づけられている。その結果、深層ニューラルネ

ットワークDNNの各層は、アクセラレータとNICとの組に割り付けられている。

[0026] [分散深層学習システムおよび分散処理用計算機の動作]

本実施の形態に係る分散深層学習システムにおいては、データ並列処理のため、分散処理用計算機10-1, 10-2には、それぞれ異なる入力データが与えられる。

分散処理用計算機10-1, 10-2のそれぞれにおいては、複数の入力データを含むミニバッチ単位で学習が行われる。本実施の形態においては、分散処理用計算機10-1, 10-2のそれぞれに、4つの入力データを含むミニバッチが与えられるものとする。

[0027] 簡単のため、分散処理用計算機10-1に着目して、分散処理用計算機10-1の動作を、図4を参照して説明する。

[0028] まず、分散処理用計算機10-1にミニバッチを構成する4つの入力データが入力されると、CPU12によって制御されたアクセラレータ13-1～13-4は、順次、順伝播処理と逆伝播処理とを行う。より具体的には、アクセラレータ13-1～13-4は、まず、入力データのそれぞれに基づいて、順次、順伝播処理をパイプライン並列によって行う（図4「順伝播処理」フェーズ）。順伝播処理は、深層ニューラルネットワークDNNの入力層から順に、入力データに基づく各ニューロンの出力と次の層のニューロンとの間の重みとの積和演算を行う処理であり、層が割り付けられたアクセラレータ13-1～13-4によって実行される。本発明における「パラメータ」の一つとして、この重みが挙げられる。

[0029] ミニバッチを構成する入力データのすべてに基づく順伝播処理が終了すると、アクセラレータ13-1～13-4は、順次、逆伝播処理をパイプライン並列によって行う（図4「逆伝播処理」フェーズ）。逆伝播処理は、例えば、順伝播処理によって得られた出力の誤差を出力層から入力層に向かって伝播させて、各重みの勾配を求める処理である。

[0030] このようにアクセラレータ13-1～13-4は、それぞれ、層ごとに4

つの入力データのそれぞれに基づく各重みの勾配を導出すると、層ごとに各重みの勾配の平均を算出し、この平均をミニバッチに基づく学習結果とする（図4「バッチ平均」フェーズ。以下、ミニバッチを単位として、各入力データより導出された学習結果の平均を算出する処理を「ミニバッチ平均」または単に「バッチ平均」ということがある。）。

なお、各重みの勾配の平均を算出する処理は、アクセラレータ13-1～13-4の間で並列に行われる。

以上のような順伝播処理、逆伝播処理、ミニバッチ平均処理は、他の分散処理用計算機10-2においても異なる入力データを含むミニバッチに対して行われ、各重みの勾配が求められる。

[0031] 分散処理用計算機10-1, 10-2のそれぞれにおいて、アクセラレータ13-1～13-4のそれぞれが学習結果、すなわち層ごとの各重みの勾配の平均を算出すると、各分散処理用計算機のNIC14-1～14-4は、それぞれ、図4に示すように、自装置のアクセラレータ13-1～13-4が導出した各層の学習結果を他の分散処理用計算機に送信するとともに、他の分散処理用計算機のアクセラレータ13-1～13-4が導出した各層の学習結果をそれぞれ受信して、自装置のアクセラレータ13-1～13-4が導出した各層の学習結果と足し合わせて、これを他の計算機に送信する。

より具体的には、複数の分散処理用計算機10-1, 10-2のそれぞれにおいて、CIN14-1～14-4は、通信ネットワーク50を介して他の分散処理用計算機10-2の対応するNIC、すなわち深層ニューラルネットワークDNNの層ごとに互いに対応づけられたNICと通信を行い、層ごとに導出された学習に関する情報、すなわち各層の各重みの平均を送受信することによって、分散処理用計算機10-1, 10-2においてそれぞれ導出された学習結果、すなわち、各重みの平均を共有する（図4「通信」フェーズ）。このようにNIC14-1～14-4によって学習結果を共有する処理は、データ並列分散深層学習におけるAll reduce処理に他なら

ない。

[0032] 複数のアクセラレータ 13-1～13-4 のそれぞれは、NIC 14-1～14-4 のそれぞれを介して共有された学習結果、すなわち、各重みの勾配を用いて、深層ニューラルネットワーク DNN の各層に含まれる各重みをそれぞれ更新する（図4、「更新」フェーズ）。

[0033] 各重みが更新されたら、分散処理用計算機 10-1，10-2 にはそれぞれ、新たなミニバッチが与えられ、この新たなミニバッチに含まれる入力データに基づいて上述した処理を行って各重みを更新する。このような学習処理を、入力データの数に応じて所定の回数、または各重みの勾配が所定の値よりも小さくなるまで繰り返す。

[0034] [第1の実施の形態の効果]

本実施の形態に係る分散深層学習システム1に用いられる分散処理用計算機10-1，10-2が、それぞれ、複数のNIC 14-1～14-4を備えているので、分散処理用計算機10-1，10-2間の通信をデバイスレベルで並列して実行することができ、オーバヘッドが少なくなる。したがって、分散処理用計算機10-1，10-2間で学習結果を共有する際に輻輳が発生することを回避することができ、深層ニューラルネットワークのより高速な学習が可能となる。

[0035] なお、本実施の形態においては、分散深層学習システム1が2台の分散処理用計算機10-1，10-2から構成される場合を例に説明したが、分散処理用計算機は2台に限定されることはなく、3台以上の分散処理用計算機から構成するようにしてもよい。この場合、複数の分散処理用計算機を含む通信ネットワーク50は、リング型ネットワークを構成してもよいし、スター型ネットワークを構成してもよい。

[0036] また、本実施の形態においては、NIC 14-1～14-4とアクセラレータ13-1～13-4とは互いに1対1に対応づけられていると説明したが、本発明においては、NIC 14-1～14-4とアクセラレータ13-1～13-4とは必ずしも同数でなくてもよい。すなわち、アクセラレータ

とN I Cの比率は、輻輳が発生しない程度とすればよく、例えば、複数のN I Cのひとつに複数のアクセラレータを対応づけてもよい。

[0037] また、本実施の形態においては、深層ニューラルネットワークD N Nが4層として、分散処理用計算機1 0 - 1, 1 0 - 2がそれぞれ4個のアクセラレータを備える態様を例に説明したが、アクセラレータの数は、深層ニューラルネットワークD N Nを構成する層の数やそれらの分割の仕方に応じて決められる。

[0038] また、分散処理用計算機1 0 - 1, 1 0 - 2のそれぞれに与えられるミニバッチは、4つの入力データを含む場合を例にパイプライン並列処理を説明したが、ミニバッチのサイズは、任意に定めることができ、4に限定されない。

[0039] [第1の実施の形態の変形例]

上述した第1の実施の形態の変形例に係る分散処理用計算機2 0 - 1 (2 0 - 2) は、図5に示すように、C P U 1 2および複数のアクセラレータ1 3 - 1 ~ 1 3 - 4を含むことは、上述した第1の実施の形態に係る分散処理用計算機1 0 - 1 (1 0 - 2) と共通するが、第1の実施の形態に係る分散処理用計算機1 0 - 1 (1 0 - 2) におけるN I C 1 4 - 1 ~ 1 4 - 4 に代えて、F P G A N I C 2 4 - 1 ~ 2 4 - 4を備えている。これらF P G A

N I C 2 4 - 1 ~ 2 4 - 4には、それぞれA l l r e d u c e処理を実行する回路が搭載されている。ここで「A l l r e d u c eを実行する回路」とは、自装置の対応するアクセラレータ1 3 - 1 ~ 1 3 - 4から受け取った学習結果を他の分散処理用計算機の対応するF P G A N I Cに送信するとともに、すべての計算機のF P G A N I Cを足し合わせ、その結果をすべてのF P G A N I Cに分配する機能を実現する回路を言う。

[0040] このような分散処理用計算機2 0 - 1 (2 0 - 2) の動作は、図6に示すように、第1の実施の形態に係る分散処理用計算機1 0 - 1 (1 0 - 2) と同じく、順伝播処理フェーズ、逆伝播処理フェーズ、バッチ平均フェーズ、通信フェーズ、更新フェーズの順となる。しかし、第1の実施の形態に係る

分散処理用計算機 10-1 (10-2) においては、*All reduce* 処理の足し算はアクセラレータ 13-1 ~ 13-4 内で実行されるため、計算カーネルと通信カーネルとが交互に実行されることとなるのに対し、この変形例に係る分散処理用計算機 20-1 (20-2) では、この変形例に係る分散処理用計算機 20-1 (20-2) では、*All reduce* 処理が *FPGA NIC* 24-1 ~ 24-4 においてそれぞれ行われるので、*All reduce* の実行を命じるだけで良い。

[0041] したがって、第 1 の実施の形態に係る分散処理用計算機 10-1 (10-2) においては、計算カーネルと通信カーネルが交互に実行された場合オーバーヘッドが発生するが、第 1 の実施の形態に係る分散処理用計算機 10-1 (10-2) においては、カーネル切り替えのオーバーヘッドが発生しないため、通信時間がより短くなる。

[0042] [第 2 の実施の形態]

第 2 の実施の形態に係る分散処理用計算機は、図 5 に示した第 1 の実施の形態の変形例に係る分散処理用計算機 20-1 (20-2) と同じく、CPU と、複数のアクセラレータと、複数の *FPGA NIC* とを備えている。複数の *FPGA NIC* には、それぞれ、上述した第 1 の実施の形態の変形例と同様に、*All reduce* 処理を実行する回路が搭載されている。

[0043] 図 7 を参照して、第 2 の実施の形態に係る分散処理用計算機の動作を説明する。

例えば、第 1 の実施の形態に係る分散処理用計算機 10-1 (10-2) や、その変形例に係る分散処理用計算機 20-1 (20-2) においては、それぞれ図 3 や図 6 に示すように、ミニバッチに含まれる入力データのそれぞれに基づいて順伝播処理と逆伝播処理とを行い、各データの学習結果をミニバッチ平均した後に、他の分散処理用計算機との通信を伴う *All reduce* 処理を行うのに対し、この第 2 の実施の形態に係る分散処理用計算機においては、図 7 に示すように、入力データそれぞれに基づく逆伝播計算が終了するたびに、すぐに計算機間での通信を伴う *All reduce* 処理を

実行し、その後ミニバッチ平均処理を行い、パラメータの値（重み）を更新する。これを換言するならば、第2の実施の形態に係る分散処理用計算機においては、逆伝播処理とA l l r e d u c e処理とをオーバーラップさせ、入力データごとにパイプライン処理される。したがって、アクセラレータは、それぞれ、入力データのそれぞれに基づいて導出され、かつ、A l l r e d u c e処理が行われた、複数の学習結果に対して、それらの平均を算出する。

[0044] 第1の実施の形態に係る分散処理用計算機10-1（10-2）においては、計算カーネルと通信カーネルとが交互に実行された場合にオーバヘッドが発生する。さらに、第1の実施の形態に係る分散処理用計算機10-1（10-2）においては、アクセラレータ13-1～13-4がミニバッチ平均した後に、データがN I C 14-1～14-4にそれぞれ送られてA l l r e d u c e処理が実行され、その後に再度アクセラレータに送られ更新されるので、データの移動時間が陽に存在する。

これに対し、第2の実施の形態に係る分散処理用計算機によれば、カーネル切り替えのオーバヘッドが発生しないため、通信時間がより短くなる。また、第2の実施の形態に係る分散処理用計算機では、アクセラレータからF P G A N I Cにデータを転送する時間が逆伝播処理に隠蔽されるため、学習がより短時間になる。

[0045] [第3の実施の形態]

第3の実施の形態に係る分散処理用計算機は、図8Aに示すように、C P U 12と、複数のアクセラレータ13-1～13-4と、複数のF P G A N I C 34-1～34-4とを備えている。複数のF P G A N I C 34-1～34-4には、それぞれ、図8Bに示すように、N I Cとともに、順伝播処理回路と、逆伝播処理回路の一部または全てが搭載されている。

[0046] 仮にF P G A N I C 34-1～34-4が、それぞれ、N I Cとともに順伝播処理回路と逆伝播処理回路の全てを実装していた場合（以下、「ケース1」という。）は、図9に示すように、順伝播処理フェーズと逆伝播処理

フェーズの双方において、アクセラレータ 13-1 ~ 13-4 と F P G A N I C 34-1 ~ 34-4 とが、それぞれミニバッチに含まれる異なる入力データを処理することができる。

[0047] また、F P G A N I C 34-1 ~ 34-4 が、それぞれ、N I C とともに順伝播処理回路と逆伝播処理の一部とを搭載している場合（以下、「ケース 2」という。）は、順伝播処理フェーズでは、アクセラレータ 13-1 ~ 13-4 と F P G A N I C 34-1 ~ 34-4 とが、それぞれミニバッチに含まれる異なる入力データを処理し、逆伝播処理フェーズでは、アクセラレータ 13-1 ~ 13-4 がそれぞれ逆伝播処理の途中まで計算した後、対応する F P G A N I C 34-1 ~ 34-4 に計算を引き継ぎ、アクセラレータ 13-1 ~ 13-4 または F P G A N I C 34-1 ~ 34-4 でミニバッチ平均を行い、A l l r e d u c e 処理のための通信を行う。

[0048] また、F P G A N I C 34-1 ~ 34-4 が、それぞれ、N I C とともに順伝播処理回路と逆伝播処理回路のうち順伝播処理回路のみを持つ場合（以下、「ケース 3」という。）は、順伝播処理フェーズでは、アクセラレータ 13-1 ~ 13-4 と F P G A N I C 34-1 ~ 34-4 とが、それぞれミニバッチに含まれる異なるデータを処理し、逆伝播処理フェーズでは、逆伝播に必要な途中計算をアクセラレータ 13-1 ~ 13-4 は保持せず、F P G A N I C 34-1 ~ 34-4 がそれぞれ都度再計算を行って逆伝播処理を実行する。

[0049] 上記のケース 1 ~ 3 のいずれの場合も、F P G A N I C 34-1 ~ 34-4 は、従来のノイマン型アクセラレータとは異なり、処理をハードウェア実装するので、より高速に学習を行うことができる。

[0050] また、上記ケース 1 の場合には、順伝播処理と逆伝播処理とをアクセラレータ 13-1 ~ 13-4 と F P G A N I C のデータのスループットが倍になる。

また、上記ケース 2 の場合には、順伝播のスループットが倍になり、逆伝播も F P G A が実行する分だけ高速になる。また、ミニバッチ平均を F P G

Aで実行する場合にはデータ移動が軽減されるので、その分のオーバーヘッドが削減される。

また、上記ケース3の場合には、順伝播のスループットが倍になり、アクセラレータは途中計算を廃棄できるので、アクセラレータのデバイスメモリをより有効活用できるようになる。

[0051] [第4の実施の形態]

第4の実施の形態に係る分散処理用計算機40-1(40-2)は、図10Aに示すように、CPU12と、複数のアクセラレータ13-1~13-4と、複数のFPGA NIC44-1~44-4とを備えている。

FPGA NIC44-1~44-4は、図10Bに示すように、それぞれ、NICとともに順伝播処理回路を備えている。

[0052] 第4の実施の形態に係る分散処理用計算機40-1(40-2)は、図11に示すように、順伝播処理フェーズでは、FPGA NIC44-1~44-4に搭載された順伝播処理回路によって、ミニバッチに含まれる複数の入力データに基づいて順次、順伝播処理を行い、順伝播の処理中に逆伝播を開始し、また、逆伝播の処理中に通信を開始する。

[0053] より具体的には、深層ニューラルネットワークDNNの出力層である第4層が割り付けられた(図3参照)FPGA NIC44-4がミニバッチに含まれる入力データに基づく順伝播処理F41~F44を順次終了させ次第、アクセラレータ13-4が逆伝播処理B41~B44を順次行う。これによって、順伝播と逆伝播とをオーバーラップさせる。さらに、FPGA NIC44-1~44-4が、それぞれ、順伝播処理に引き続いて、他の分散処理用計算機との間で通信を順次開始する。これによって、逆伝播と通信とをオーバーラップさせる。

[0054] このように順伝播と逆伝播とのオーバーラップ、および逆伝播と通信とのオーバーラップにより、処理時間をさらに短縮することができる。

[0055] 以上、説明したように、本発明の実施の形態に係る分散処理用計算機からなる深層学習システムによれば、複数のアクセラレータがそれぞれ導出した

学習結果を含む学習に関する情報を複数のネットワークインターフェース回路によって並列に送受信するので、より短い時間で分散深層学習を行うことができる。

産業上の利用可能性

[0056] 本発明は、深層学習システムに利用することができる。

符号の説明

[0057] 1…分散深層学習システム、10-1, 10-2, 20-1, 30-1, 40-1…分散処理用計算機、12…CPU、13-1～13-4…アクセラレータ、14-1～14-4…NIC、24-1～24-4, 34-1～34-4, 44-1～44-4…FPGA NIC。

請求の範囲

- [請求項1] それぞれにニューラルネットワークの一部が割り付けられ、入力データに基づいて学習結果を導出するとともに、前記学習結果を用いて前記ニューラルネットワークの一部に含まれるパラメータの値をそれぞれ更新するように構成された、複数のアクセラレータと、
- それぞれが前記学習結果を含む学習に関する情報を送受信するように構成された複数のネットワークインターフェース回路と、
- 前記複数のアクセラレータおよび前記複数のネットワークインターフェース回路を制御して、前記複数のアクセラレータのそれぞれに入力データに基づいて学習結果を導出させるとともに、前記複数のネットワークインターフェース回路に、前記複数のアクセラレータにそれぞれ導出させた前記学習結果を含む学習に関する情報を並列に送受信させるように構成された、演算処理装置と
- を備える分散処理用計算機。
- [請求項2] 請求項1に記載された分散処理用計算機において、
- 前記複数のアクセラレータには、前記ニューラルネットワークの一部として、前記ニューラルネットワークを構成する複数の層がそれぞれ割り付けられ、
- 前記複数のアクセラレータは、前記複数の層ごとに順伝播処理と逆伝播処理とを実行する、
- 分散処理用計算機。
- [請求項3] 請求項1または2に記載された分散処理用計算機において、
- 前記複数のネットワークインターフェース回路は、
- 前記複数のアクセラレータが前記学習結果を導出すると、前記学習結果を含む学習に関する情報を並列に送受信する、分散処理用計算機。
- [請求項4] 請求項3に記載された分散処理用計算機において、
- 前記複数のアクセラレータは、

複数の入力データを含むミニバッチに対して、前記複数の入力データのそれぞれに基づく順伝播処理と逆伝播処理とをパイプラインで行って前記複数の入力データのそれぞれに基づいて複数の学習結果を導出し、導出された前記複数の学習結果の平均を前記ミニバッチに基づく学習結果として算出し、

前記複数のネットワークインターフェース回路は、

前記複数のアクセラレータが前記複数の入力データに基づいて複数の学習結果を算出すると、前記複数の学習結果を含む学習に関する情報を並列に送受信する、分散処理用計算機。

[請求項5]

請求項3に記載された分散処理用計算機において、

前記複数のアクセラレータは、

複数の入力データを含むミニバッチに対して、前記複数の入力データのそれぞれに基づく順伝播処理と逆伝播処理とをパイプラインで行って前記複数の入力データのそれぞれに基づいて複数の学習結果を導出し、

前記複数のネットワークインターフェース回路は、

前記複数のアクセラレータが前記ミニバッチに含まれる前記複数の入力データに基づく複数の学習結果を導出するごとに前記学習結果を含む学習に関する情報を並列に送受信する、分散処理用計算機。

[請求項6]

請求項3に記載された分散処理用計算機において、

前記複数のアクセラレータは、

複数の入力データを含むミニバッチに対して、前記複数の入力データのそれぞれに基づく順伝播処理と逆伝播処理とをパイプラインで行って前記複数の入力データのそれぞれに基づいて複数の学習結果を導出し、

前記複数のネットワークインターフェース回路は、

前記複数のアクセラレータが前記ミニバッチに含まれる前記複数の入力データに基づく複数の学習結果を導出するごとに、前記逆伝播処

理とオーバーラップさせて前記学習結果を含む学習に関する情報を順次に送受信する、分散処理用計算機。

[請求項7] 請求項1～6のいずれか一項に記載された分散処理用計算機において、

前記複数のネットワークインターフェース回路は、前記複数のアクセラレータと同数設けられている、分散処理用計算機。

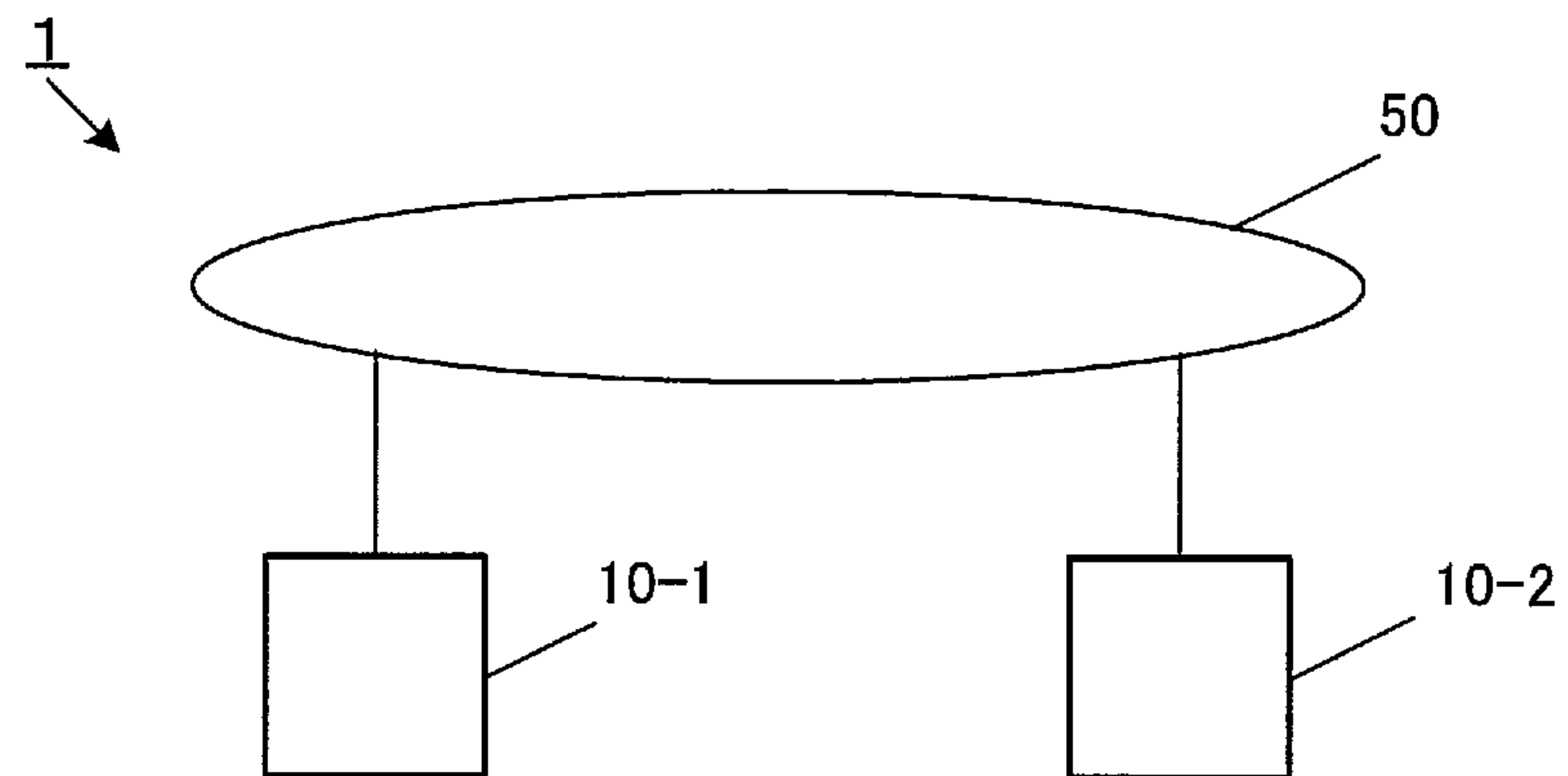
[請求項8] 互いに通信可能に接続された複数の計算機を備え、

前記複数の計算機のそれぞれは、請求項1～7のいずれか一項に記載された分散処理用計算機であり、

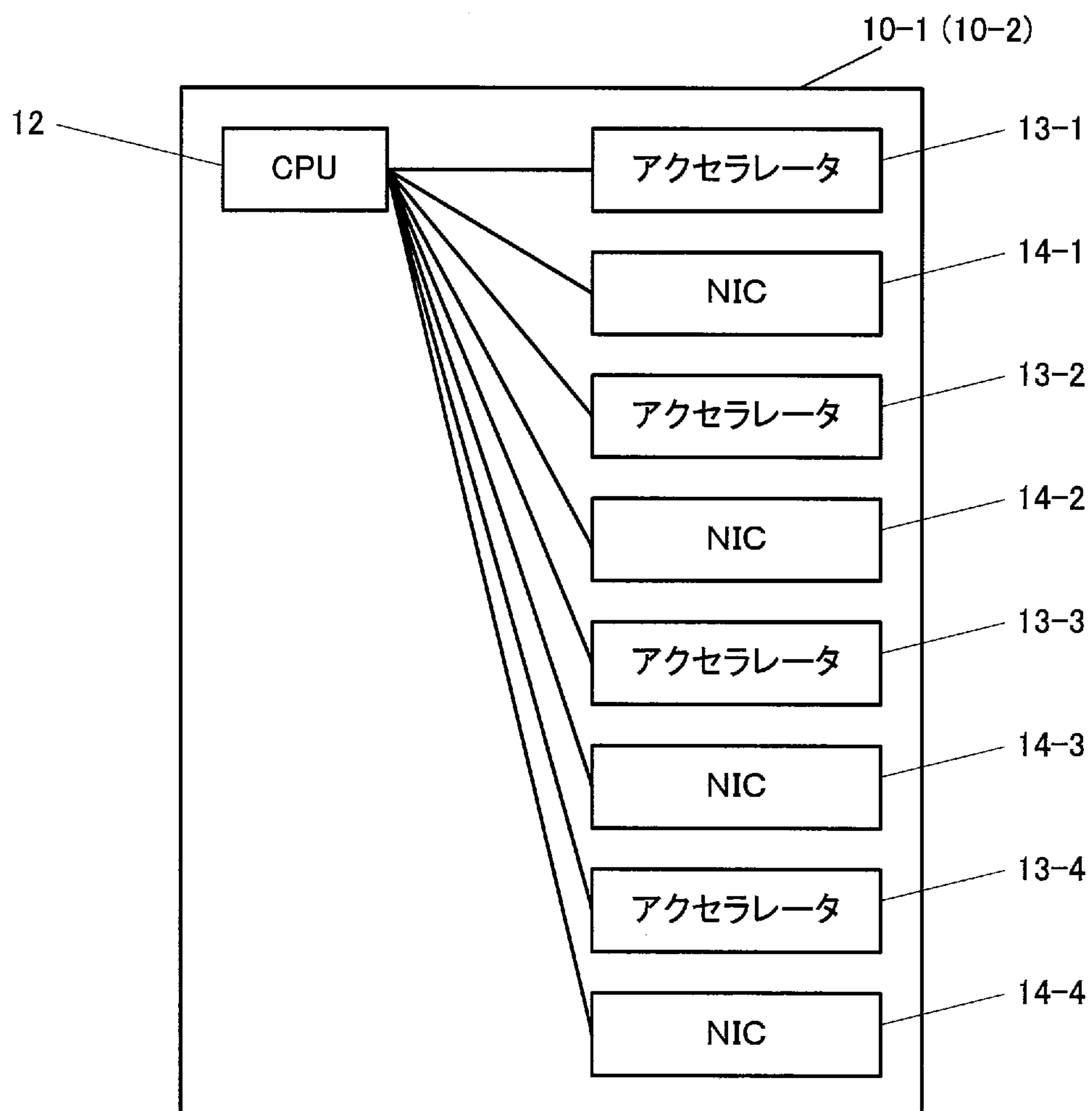
前記複数の計算機のうちの一の計算機の前記複数のネットワークインターフェース回路は、前記一の計算機の複数のアクセラレータがそれぞれ導出した複数の学習結果と前記複数の計算機のうちの他の計算機の複数のアクセラレータがそれぞれ導出した複数の学習結果とを並列に送受信し、

前記一の計算機の前記複数のアクセラレータは、前記複数のアクセラレータが導出した前記学習結果と、前記複数のネットワークインターフェース回路を介して前記他の計算機から受信した前記複数の学習結果とに基づいて前記パラメータの値を更新する、分散深層学習システム。

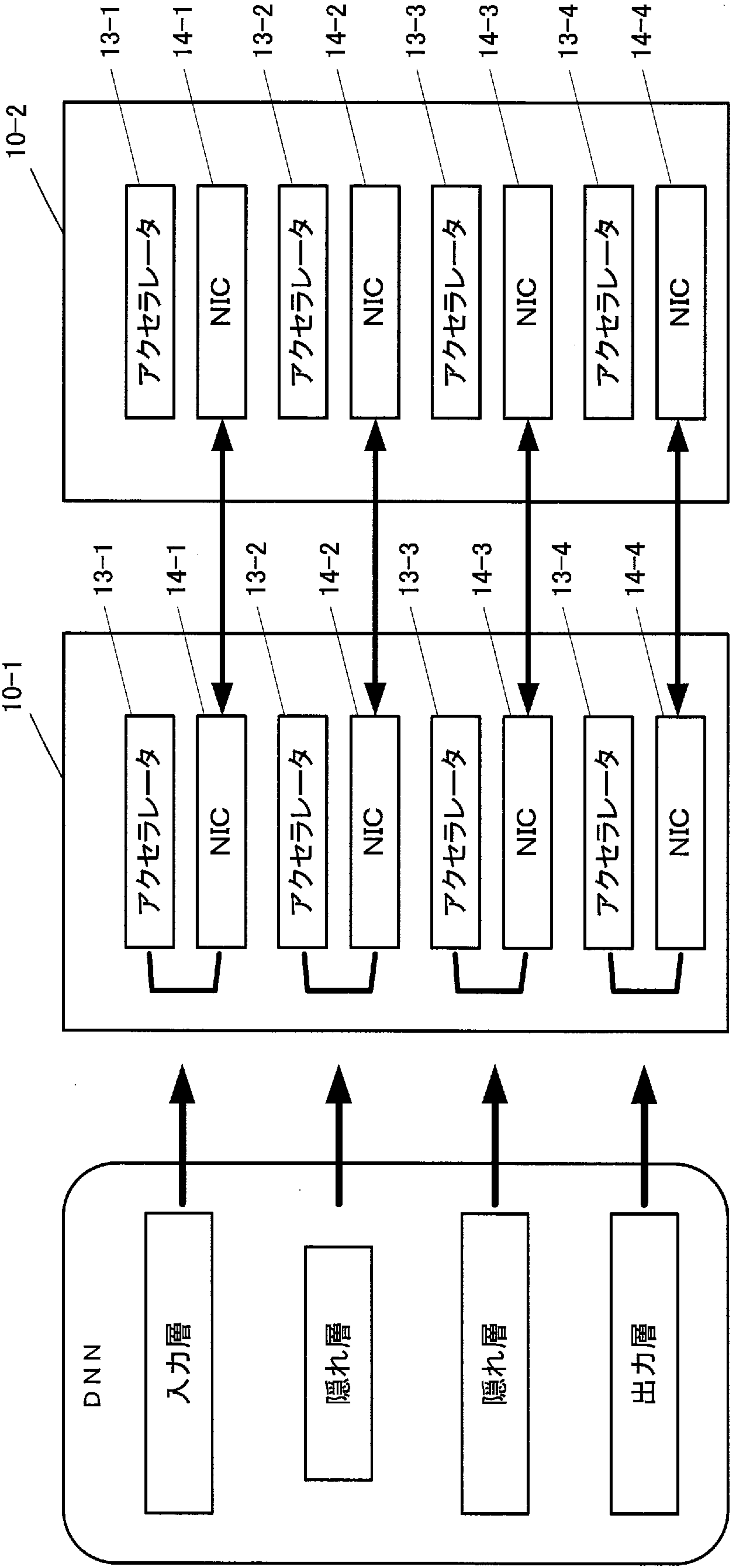
[図1]



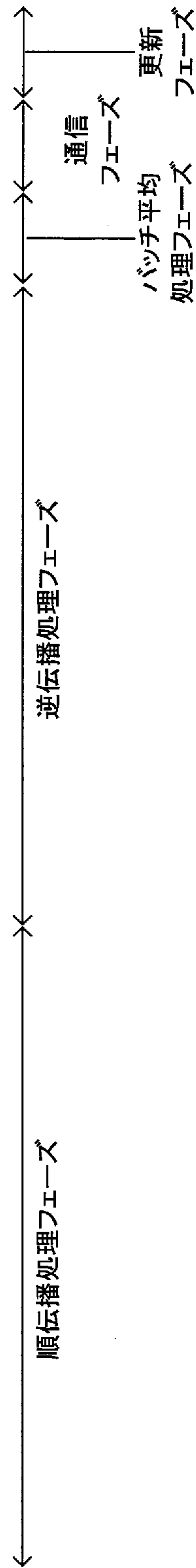
[図2]



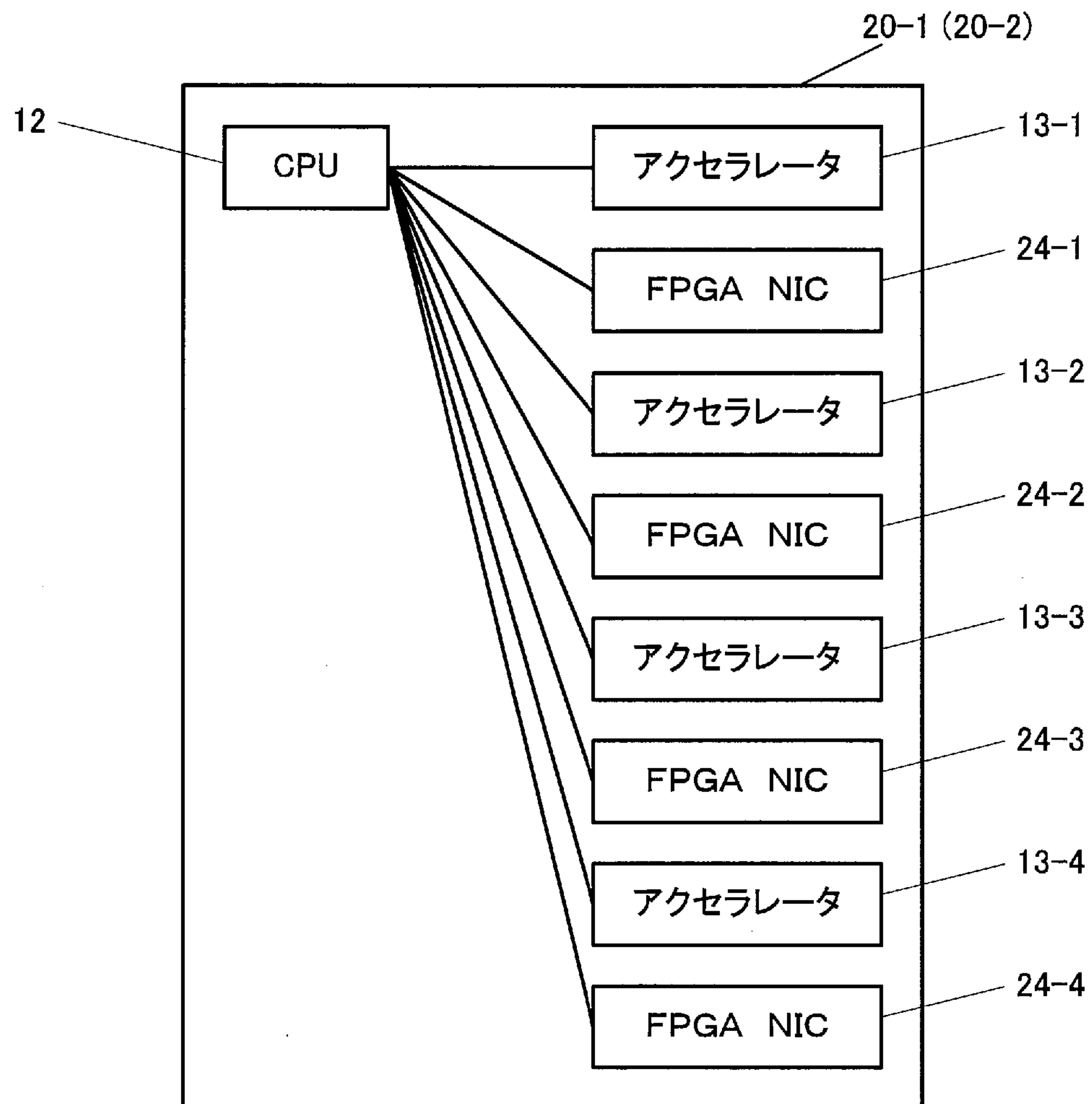
[図3]



[圖4]

[illegible]

[図5]



【図6】

アクセラレータ 1	F 11	F 12	F 13	F 14								B 11	B 12	B 13	B 14	バッチ平均		更新
F P G A N I C 1																	第 1 層通信	
アクセラレータ 2		F 21	F 22	F 23	F 24							B 21	B 22	B 23	B 24	バッチ平均		更新
F P G A N I C 2																	第 2 層通信	
アクセラレータ 3			F 31	F 32	F 33	F 34					B 31	B 32	B 33	B 34		バッチ平均		更新
F P G A N I C 3																	第 3 層通信	
アクセラレータ 4				F 41	F 42	F 43	F 44	B 41	B 42	B 43	B 44					バッチ平均		更新
F P G A N I C 4																	第 4 層通信	

←

順伝播処理フェーズ

→

逆伝播処理フェーズ

←

通信フェーズ

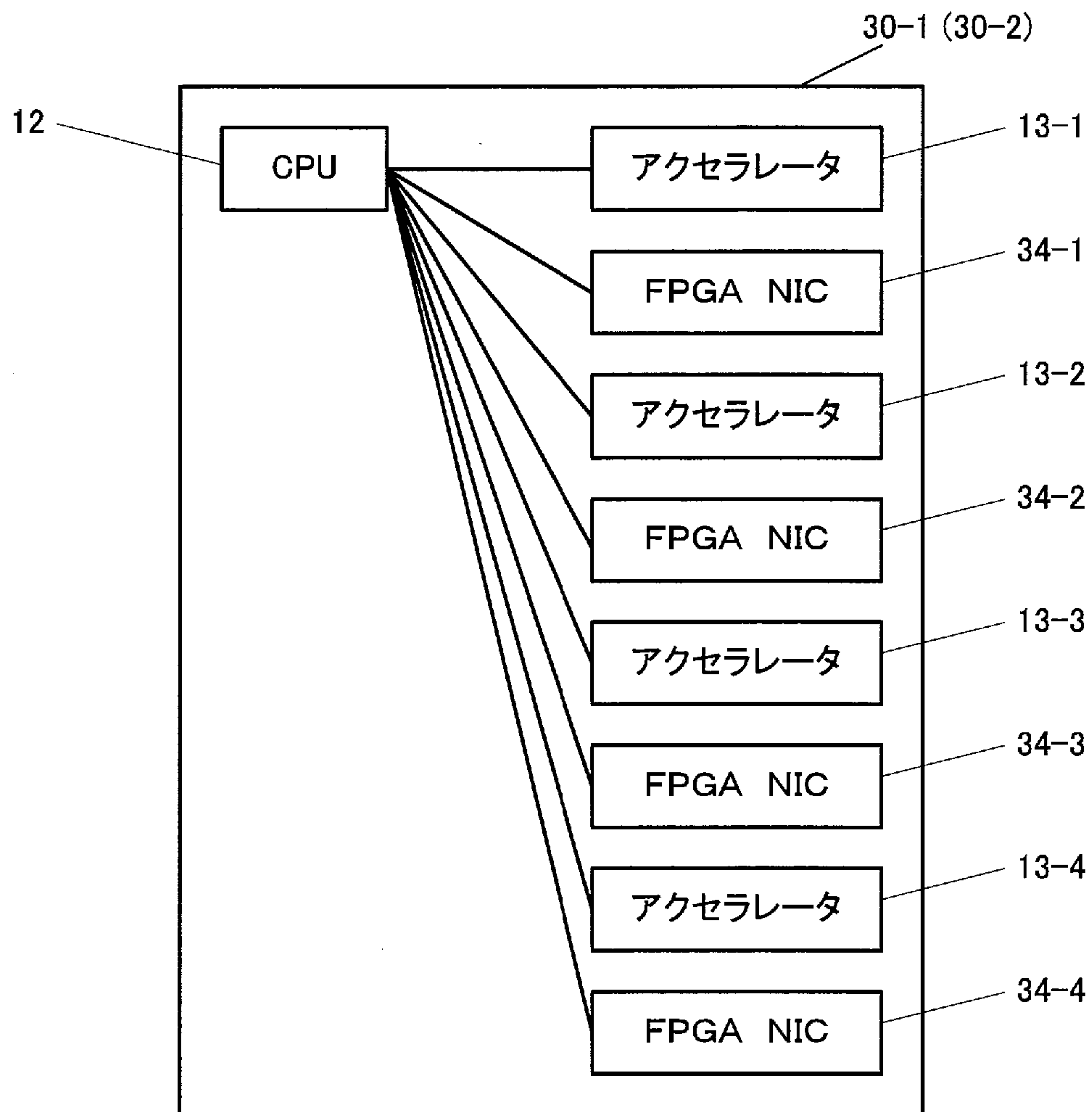
バッチ平均
処理フェーズ

更新
フェーズ

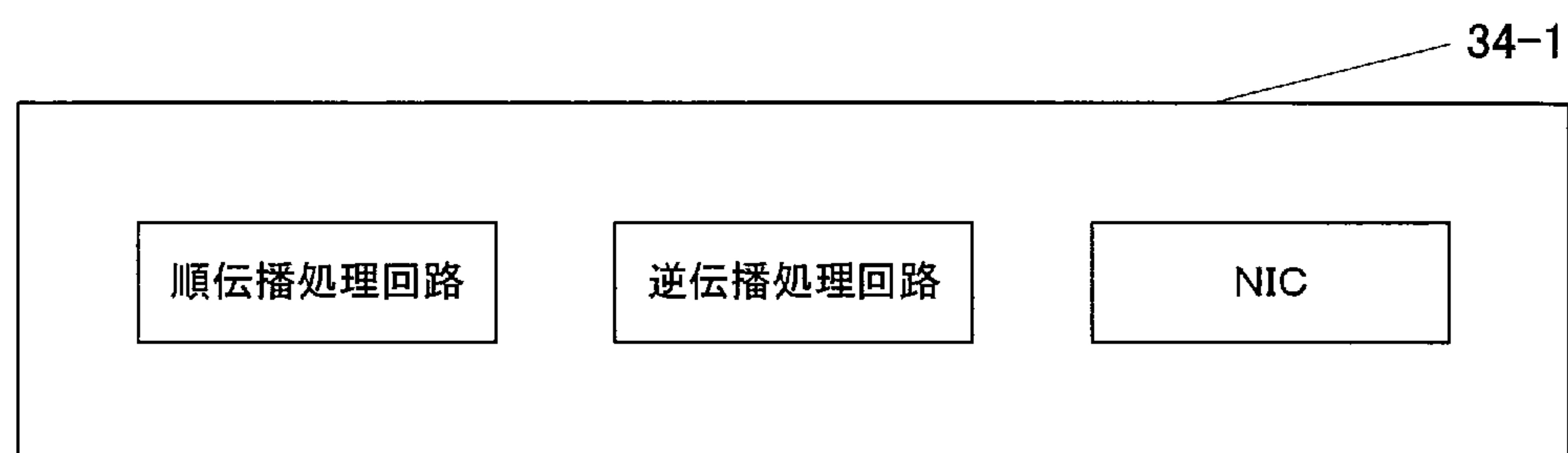
[図7]

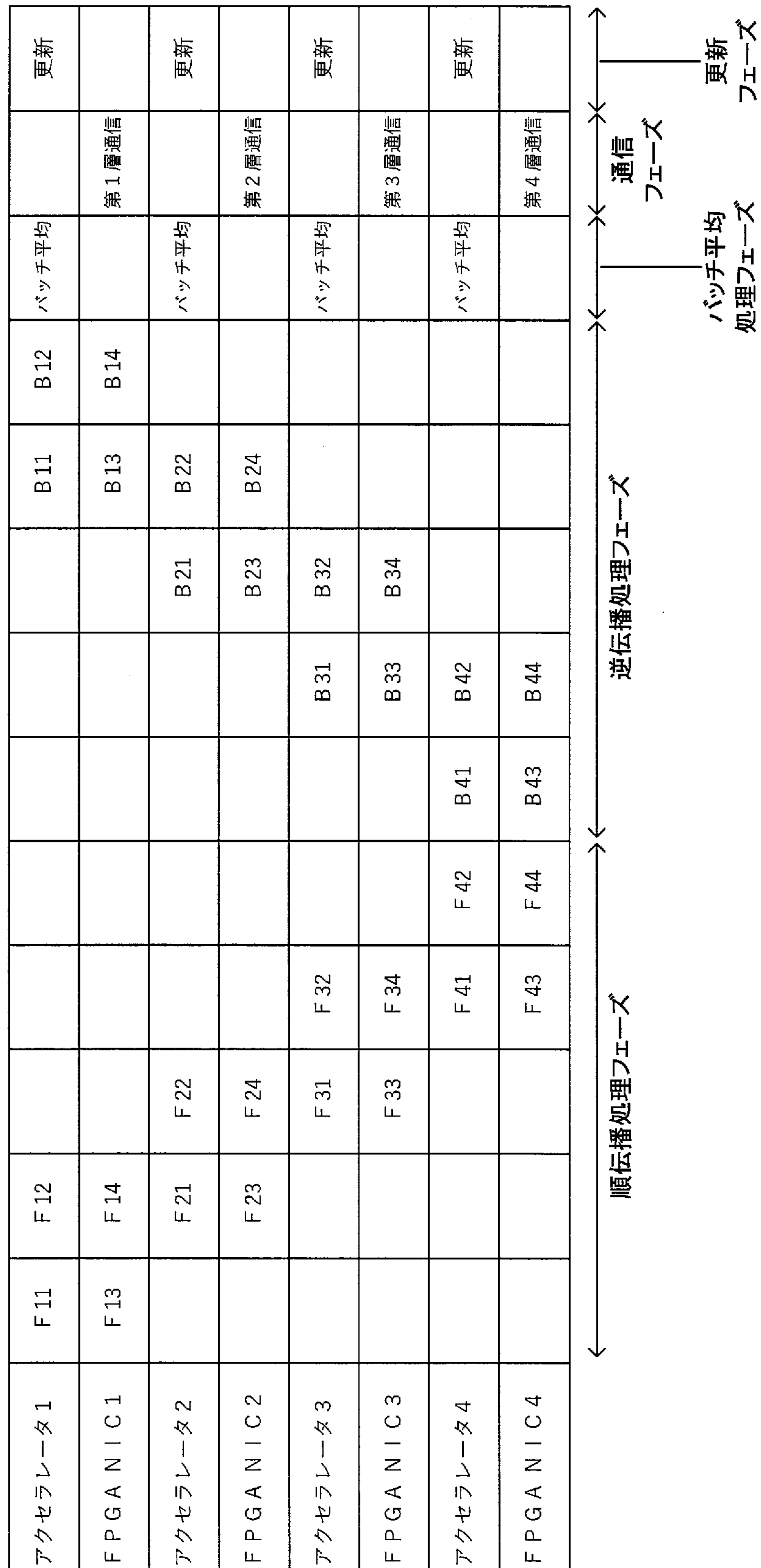
アクセラレータ 1	F 11	F 12	F 13	F 14								B 11	B 12	B 13	B 14		バッチ平均	更新
F P G A N I C 1													第 1 層通信	第 1 層通信	第 1 層通信			
アクセラレータ 2		F 21	F 22	F 23	F 24							B 21	B 22	B 23	B 24		バッチ平均	更新
F P G A N I C 2													第 2 層通信	第 2 層通信	第 2 層通信			
アクセラレータ 3			F 31	F 32	F 33	F 34					B 31	B 32	B 33	B 34			バッチ平均	更新
F P G A N I C 3												第 3 層通信	第 3 層通信	第 3 層通信	第 3 層通信			
アクセラレータ 4				F 41	F 42	F 43	F 44	B 41	B 42	B 43	B 44						バッチ平均	更新
F P G A N I C 4									第 4 層通信	第 4 層通信	第 4 層通信		第 4 層通信	第 4 層通信				

[図8A]

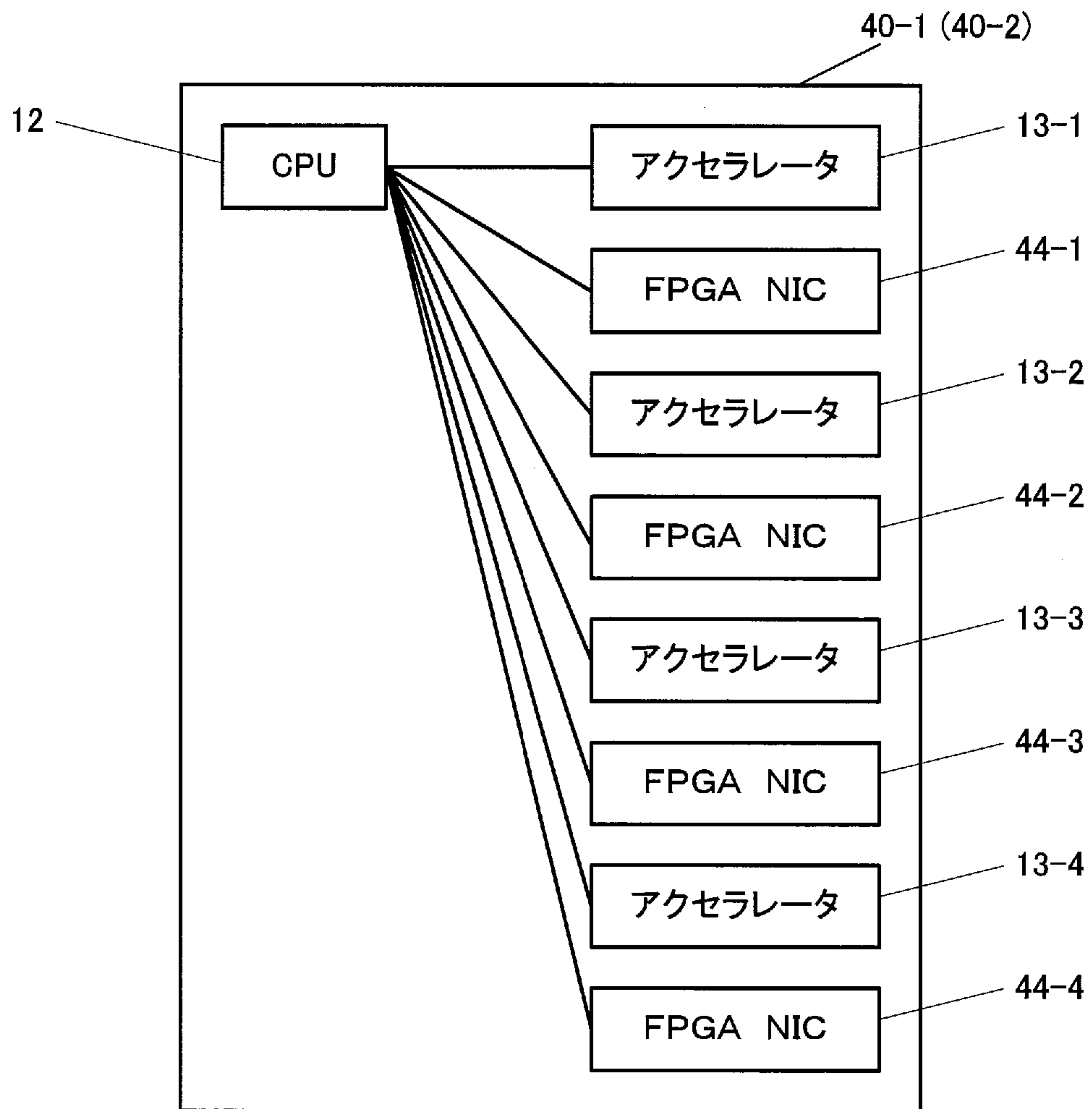


[図8B]

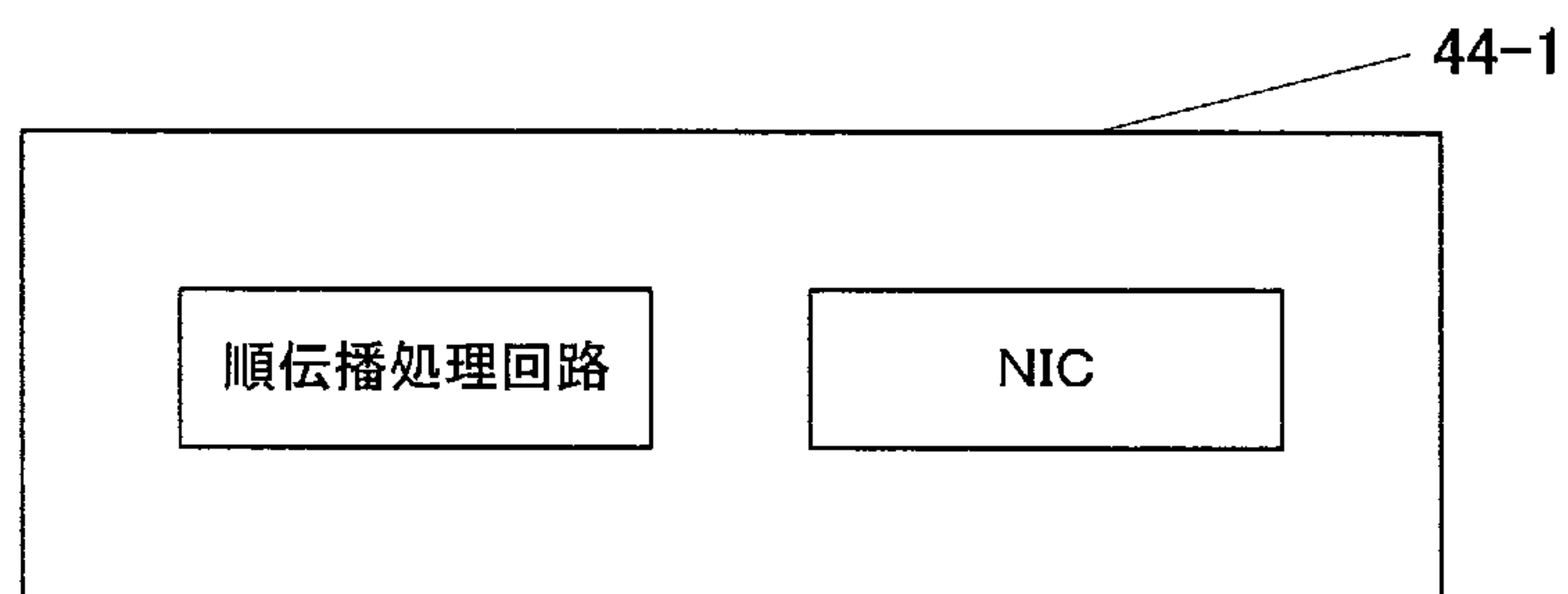




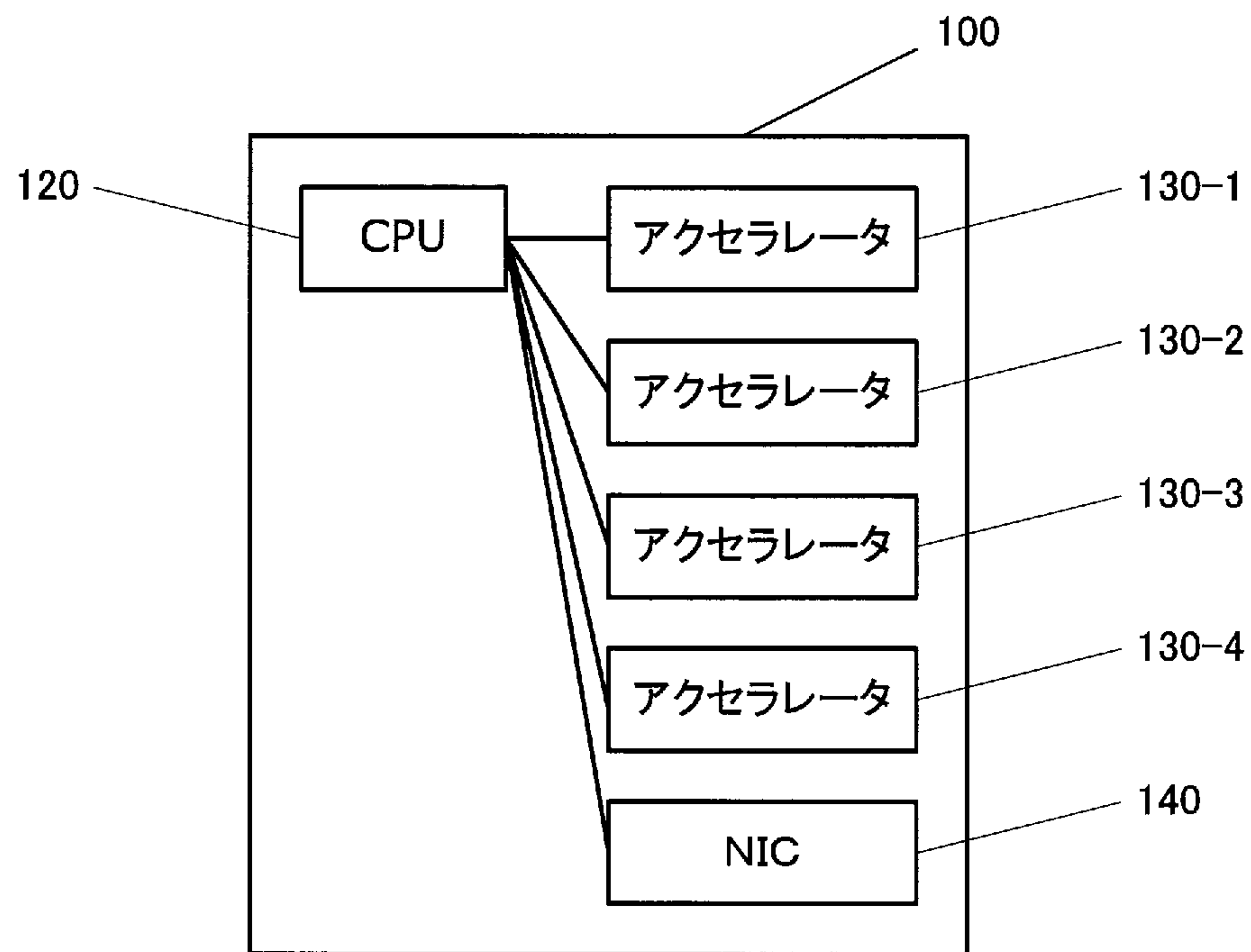
[図10A]



[図10B]



[図12]



[図13]

アクセラレータ 1	F 11	F 12	F 13	F 14														B 11	B 12	B 13	B 14	バッチ平均					更新
アクセラレータ 2		F 21	F 22	F 23	F 24													B 21	B 23	B 24		バッチ平均					更新
アクセラレータ 3			F 31	F 32	F 33	F 34												B 31	B 32	B 34		バッチ平均					更新
アクセラレータ 4				F 41	F 42	F 43	F 44	B 41	B 42	B 43	B 44											バッチ平均					更新
N I C																								第 1 層通信	第 2 層通信	第 3 層通信	第 4 層通信

INTERNATIONAL SEARCH REPORT

International application No.

PCT/JP2019/046966

A. CLASSIFICATION OF SUBJECT MATTER Int. Cl. G06N3/08(2006.01) i FI: G06N3/05 According to International Patent Classification (IPC) or to both national classification and IPC														
B. FIELDS SEARCHED Minimum documentation searched (classification system followed by classification symbols) Int. Cl. G06N3/00-99/00 Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched Published examined utility model applications of Japan 1922-1996 Published unexamined utility model applications of Japan 1971-2020 Registered utility model specifications of Japan 1996-2020 Published registered utility model applications of Japan 1994-2020 Electronic data base consulted during the international search (name of data base and, where practicable, search terms used)														
C. DOCUMENTS CONSIDERED TO BE RELEVANT														
Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.												
Y A	本田巧、外 8 名, 大規模 GPU クラスタにおける ResNet-50/ImageNet 学習の高速化, 情報処理学会研究報告ハイパフォーマンスコンピューティング (HPC), 17 July 2019, vol. 1.2019-HPC-170, no. 6, pp. 1-7, in particular「3. DNN 学習」-「5.2 学習率の推移と精度」non-official translation (HONDA, Takumi et al. Acceleration of ResNet-50/ImageNet Learning in Large-scale GPU Clusters. IPSJ SIG Technical Report. High Performance Computing (HPC). "3. DNN Learning" to "5.2 Learning Ratios Estimates and Accuracy".)	1-4, 7-8 5-6												
Y	HUANG, Yanping et al., GPipe: Efficient Training of Giant Neural Networks using Pipeline Parallelism, arXiv [online], 16 November 2018 [retrieved on: 05 August 2020], internet: <URL: https://arxiv.org/abs/1811.06965v1> in particular, [3.2 Algorithm]	1-4, 7-8												
☐ Further documents are listed in the continuation of Box C. ☐ See patent family annex.														
<table><tr><td>* Special categories of cited documents:</td><td>"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention</td></tr><tr><td>"A" document defining the general state of the art which is not considered to be of particular relevance</td><td>"X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone</td></tr><tr><td>"E" earlier application or patent but published on or after the international filing date</td><td>"Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art</td></tr><tr><td>"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)</td><td>"&" document member of the same patent family</td></tr><tr><td>"O" document referring to an oral disclosure, use, exhibition or other means</td><td></td></tr><tr><td>"P" document published prior to the international filing date but later than the priority date claimed</td><td></td></tr></table>			* Special categories of cited documents:	"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention	"A" document defining the general state of the art which is not considered to be of particular relevance	"X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone	"E" earlier application or patent but published on or after the international filing date	"Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art	"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)	"&" document member of the same patent family	"O" document referring to an oral disclosure, use, exhibition or other means		"P" document published prior to the international filing date but later than the priority date claimed	
* Special categories of cited documents:	"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention													
"A" document defining the general state of the art which is not considered to be of particular relevance	"X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone													
"E" earlier application or patent but published on or after the international filing date	"Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art													
"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)	"&" document member of the same patent family													
"O" document referring to an oral disclosure, use, exhibition or other means														
"P" document published prior to the international filing date but later than the priority date claimed														
Date of the actual completion of the international search 05.08.2020		Date of mailing of the international search report 18.08.2020												
Name and mailing address of the ISA/ Japan Patent Office 3-4-3, Kasumigaseki, Chiyoda-ku, Tokyo 100-8915, Japan		Authorized officer Telephone No.												

国際調査報告		国際出願番号 PCT/JP2019/046966
A. 発明の属する分野の分類（国際特許分類（IPC）） G06N 3/08(2006.01)i FI: G06N3/08		
B. 調査を行った分野 調査を行った最小限資料（国際特許分類（IPC）） G06N3/00-99/00 最小限資料以外の資料で調査を行った分野に含まれるもの 日本国実用新案公報 1922 - 1996年 日本国公開実用新案公報 1971 - 2020年 日本国実用新案登録公報 1996 - 2020年 日本国登録実用新案公報 1994 - 2020年 国際調査で使用した電子データベース（データベースの名称、調査に使用した用語）		
C. 関連すると認められる文献		
引用文献の カテゴリー*	引用文献名 及び一部の箇所が関連するときは、その関連する箇所の表示	関連する 請求項の番号
Y	本田巧、外8名、大規模GPUクラスタにおけるResNet-50/ImageNet学習の高速化、情報処理学会研究報告 ハイパフォーマンスコМПユーティング (HPC), 2019.07.17, Vol.2019-HPC-170 No.6, pp.1-7 特に「3. DNN学習」－「5.2 学習率の推移と精度」	1-4, 7-8
A		5-6
Y	HUANG, Yanping, et al., GPipe: Efficient Training of Giant Neural Networks using Pipeline Parallelism, arXiv [online], 2018.11.16 [検索日: 2020.08.05], インターネット: <URL: https://arxiv.org/abs/1811.06965v1> 特に「3.2 Algorithm」	1-4, 7-8
☐ C欄の続きにも文献が列挙されている。 ☐ パテントファミリーに関する別紙を参照。		
* 引用文献のカテゴリー “A” 特に関連のある文献ではなく、一般的技術水準を示すもの “E” 国際出願日前の出願または特許であるが、国際出願日以後に公表されたもの “L” 優先権主張に疑義を提起する文献又は他の文献の発行日若しくは他の特別な理由を確立するために引用する文献（理由を付す） “O” 口頭による開示、使用、展示等に言及する文献 “P” 国際出願日前で、かつ優先権の主張の基礎となる出願の日の後に公表された文献		“T” 国際出願日又は優先日後に公表された文献であって出願と抵触するものではなく、発明の原理又は理論の理解のために引用するもの “X” 特に関連のある文献であって、当該文献のみで発明の新規性又は進歩性がないと考えられるもの “Y” 特に関連のある文献であって、当該文献と他の1以上の文献との、当業者にとって自明である組合せによって進歩性がないと考えられるもの “&” 同一パテントファミリー文献
国際調査を完了した日 05.08.2020		国際調査報告の発送日 18.08.2020
名称及びあて先 日本国特許庁 (ISA/JP) 〒100-8915 日本国 東京都千代田区霞が関三丁目4番3号		権限のある職員（特許庁審査官） 杉浦 孝光 5B 5287 電話番号 03-3581-1101 内線 3545