

(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
2019年12月12日(12.12.2019)



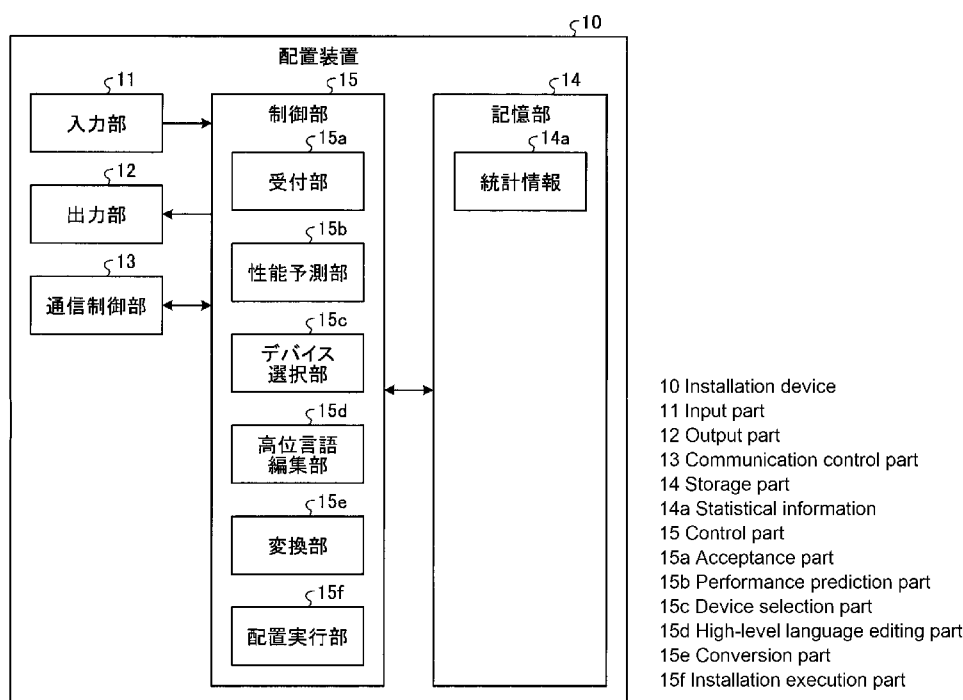
(10) 国際公開番号

WO 2019/235271 A1

- (51) 国際特許分類:
G06F 9/50 (2006.01)
- (21) 国際出願番号: PCT/JP2019/020775
- (22) 国際出願日: 2019年5月24日(24.05.2019)
- (25) 国際出願の言語: 日本語
- (26) 国際公開の言語: 日本語
- (30) 優先権データ:
特願 2018-108254 2018年6月6日(06.06.2018) JP
- (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: 武井 勇樹 (TAKELI, 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). 西山 聡史 (NISHIYAMA, Satoshi); 〒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: 特許業務法人酒井国際特許事務所 (SAKAI INTERNATIONAL PATENT OFFICE); 〒1000013 東京都千代田区霞が関3丁目8番1号 虎の門三井ビルディング Tokyo (JP).
- (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,

(54) Title: INSTALLATION DEVICE AND INSTALLATION METHOD

(54) 発明の名称: 配置装置および配置方法



(57) Abstract: In the present invention, a storage part (14) stores statistical information (14a) which includes: resource consumption amounts of each of multiple types of hardware components constituting candidate function installation destinations; and performance information which represents the performance of each of said hardware components. An acceptance part (15a) accepts an input of: description content of the function being described in high-level languages which are hardware description languages corresponding to the multiple types of hardware components; and performance re-

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, 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))

quirements which represent the requested performance. For each of the hardware components, a performance prediction part (15b) uses the description content and a prescribed algorithm to compute a predicted performance and a predicted resource consumption amount. A device selection part (15c) selects, as an installation destination, the hardware component for which the computed predicted performance and the performance information satisfy the performance requirements and the total of the predicted resource consumption amount and the resource consumption amount is less than or equal to a resource capacity.

(57) 要約: 記憶部 (14) が、機能の配置先の候補である複数種のハードウェアのそれぞれのリソース消費量と性能を表す性能情報とを含む統計情報 (14 a) を記憶し、受付部 (15 a) が、複数種のハードウェアに対応したハードウェア記述言語である高位言語による機能の記述内容と、要求される性能を表す性能要件との入力を受け付け、性能予測部 (15 b) が、ハードウェアごとに、記述内容と所定のアルゴリズムとを用いて、予測される性能である予測性能と、予測されるリソース消費量である予測リソース消費量とを算出し、デバイス選択部 (15 c) が、算出された予測性能および性能情報が性能要件を満たし、かつ予測リソース消費量とリソース消費量との合計値がリソース容量以下であるハードウェアを、配置先として選択する。

明 細 書

発明の名称： 配置装置および配置方法

技術分野

[0001] 本発明は、配置装置および配置方法に関する。

背景技術

[0002] キャリアネットワークにおいて、トラフィックの増加に伴って、サーバのCPU (Central Processing Unit) のパケット処理負荷が増加している。そこで、サーバのCPU負荷をオフロードするため、CPUとは異なるハードウェア種別のアクセラレータをサーバと組み合わせて、ネットワーク機能（以下、単に機能とも記す。）の処理を分散させる技術が知られている（非特許文献1 参照）。このように、ハードウェア（以下、デバイスとも記す。）を組み合わせて、システム全体の性能の向上を図る技術が検討されている。

[0003] その場合に、アクセラレータに使用するハードウェアプロセッサによって、システムのアーキテクチャの特徴やネットワーク機能の実現手法が異なる。そこで、これらの違いを隠蔽して機能開発を効率化するために、抽象度を高めて複数種のデバイスに対応したハードウェア記述言語であるP4等の高位言語が提案されている（非特許文献2～4 参照）。

[0004] また、アクセラレータを1つのネットワーク機能で占有するとリソースに無駄が生じるため、アクセラレータには複数の機能を配置することが望ましい。そして、システム全体の性能向上を図るためには、複数種のハードウェアを組み合わせたシステム構成が必要である。

先行技術文献

非特許文献

[0005] 非特許文献1：J.F.Zazo, et al., “A PCIe DMA engine to support the virtualization of 40 Gbps FPGA-accelerated network appliances”, in Reconfigurable Computing and FPGAs, pp.1-6, 2015年

非特許文献2：「P4/CによるSmartNICファームウェア開発の取組み」、[onlin

e]、2017年、富士通、[2018年5月11日検索]、インターネット<URL:https://www.okinawaopenlabs.com/ood/2017/wp-content/uploads/sites/4/2017/12/fujitsu_3.pdf>

非特許文献3：]P.Bosshart, et al., “P4: Programming Protocol-Independent Packet Processors”, SIGCOMM CCR, vol.44, no.3, pp.87-95, 2014年

非特許文献4：Henning Stubbe, “P4 Compiler & Interpreter: A Survey”, Proc. Future Internet (FI) Innov. Internet Technol. Mobile Commun. (IITM). Vol.47, pp.47-52, 2017年

発明の概要

発明が解決しようとする課題

[0006] しかしながら、従来、複数種のハードウェアを用いてネットワーク機能を実現する場合に、ハードウェアごとに異なるシミュレーションや動作確認を実行することにより、処理性能や消費リソース容量を確認していた。そのため、シミュレーションや動作確認に長時間を要し、迅速にネットワーク機能の配置を行うことができなかった。その結果、リソースに無駄が生じる場合があった。

[0007] 本発明は、上記に鑑みてなされたものであって、複数種のハードウェアへのネットワーク機能の配置を迅速に行うことを目的とする。

課題を解決するための手段

[0008] 上述した課題を解決し、目的を達成するために、本発明に係る配置装置は、機能の配置先の候補である複数種のハードウェアのそれぞれのリソース消費量と性能を表す性能情報とを含む統計情報を記憶する記憶部と、複数種のハードウェアに対応したハードウェア記述言語である高位言語による機能の記述内容と、要求される性能を表す性能要件との入力を受け付ける受付部と、ハードウェアごとに、前記記述内容と所定のアルゴリズムとを用いて、予測される性能である予測性能と、予測されるリソース消費量である予測リソース消費量とを算出する予測部と、算出された前記予測性能および前記性能

情報が前記性能要件を満たし、前記予測リソース消費量と前記リソース消費量との合計値がリソース容量以下であるハードウェアを、配置先として選択する選択部と、を備えることを特徴とする。

発明の効果

[0009] 本発明によれば、複数種のハードウェアへのネットワーク機能の配置を迅速に行うことができる。

図面の簡単な説明

[0010] [図1]図1は、本実施形態に係る配置装置の対象システムを説明するための説明図である。

[図2]図2は、配置装置の概略構成を例示する模式図である。

[図3]図3は、統計情報のデータ構成を例示する図である。

[図4]図4は、配置装置の処理を説明するための説明図である。

[図5]図5は、配置処理手順を例示するフローチャートである。

[図6]図6は、配置プログラムを実行するコンピュータを例示する図である。

発明を実施するための形態

[0011] 以下、図面を参照して、本発明の一実施形態を詳細に説明する。なお、この実施形態により本発明が限定されるものではない。また、図面の記載において、同一部分には同一の符号を付して示している。

[0012] [対象システム]

図1は、本実施形態に係る配置装置の対象システムを説明するための説明図である。図1に示すように、配置装置による配置処理の対象となるシステムは、サーバにアクセラレータが組み込まれて構成される。アクセラレータは、NP (Network Processor) やFPGA (Field Programmable Gate Array) 等のサーバのCPUとは種別の異なるハードウェアで実現される。

[0013] このシステムにおいて、アクセラレータが通信ネットワークの他のネットワーク装置から受信したパケットのトラヒックが、パケットに対して要求される処理に応じて、アクセラレータまたはサーバのCPUに振り分けられる。アクセラレータには、サーバのCPUの負荷要因となるネットワーク機能

が配置される。これにより、サーバのCPU負荷をアクセラレータにオフロードすることが可能となる。

[0014] 本実施形態の配置装置は、このようなシステムにおいて、各ネットワーク機能をCPU、またはアクセラレータのいずれのハードウェアに配置するかを決定する配置処理を実行する。

[0015] [配置装置の構成]

図2は、配置装置10の概略構成を例示する模式図である。図2に例示するように、本実施形態の配置装置10は、ワークステーションやパソコン等の汎用コンピュータで実現され、入力部11、出力部12、通信制御部13、記憶部14、および制御部15を備える。

[0016] 入力部11は、キーボードやマウス等の入力デバイスを用いて実現され、操作者による入力操作に対応して、制御部15に対して処理開始などの各種指示情報を入力する。出力部12は、液晶ディスプレイなどの表示装置、プリンター等の印刷装置、情報通信装置等によって実現される。通信制御部13は、LAN (Local Area Network) やインターネットなどの電気通信回線を介したネットワーク装置等の外部の装置との通信を制御する。例えば、通信制御部13は、図1に例示した配置処理対象のシステムから、CPU、NP、またはFPGA等のハードウェアについての性能の測定値等を受信する。

[0017] 記憶部14は、RAM (Random Access Memory)、フラッシュメモリ (Flash Memory) 等の半導体メモリ素子、または、ハードディスク、光ディスク等の記憶装置によって実現される。なお、記憶部14は、LANやインターネット等の電気通信回線を介して制御部15と通信する構成でもよい。

[0018] 本実施形態において、記憶部14には、統計情報14aが記憶されている。この統計情報14aは、機能の配置先の候補である複数種のハードウェアのそれぞれのリソース消費量と性能を表す性能情報とを含む情報である。

[0019] ここで、図3は、統計情報14aのデータ構成を例示する図である。図3に例示するように、統計情報14aは、デバイスごとのリソース消費量と性

能情報とを含む。図3には、デバイスとして、CPU、FPGA1、FPGA2等が例示されている。

[0020] リソース消費量は、各デバイスの使用率(%)等で表される実績値である。図3に示す例では、リソース消費量には、さらに、デバイスごとに異なる詳細な情報が含まれる。例えば、図3に示すCPUのリソース消費量には、コア数、コア数使用量、メモリ数、メモリ使用量等が含まれ、FPGA1のリソース消費量には、FF(フリップフロップ)数、FF使用量、LUT(ルックアップテーブル)数、LUT使用量等が含まれている。

[0021] 性能情報は、スループット(Gbps)、遅延(ms)等を含む実績値である。ここで、スループットとは、各デバイスが一定時間内に処理できるデータ量である。また、遅延は、例えば、パケットを送信してから、受信した側が送信側にACKパケットを送信し、送信側でそれを受信するまでの時間で表される。

[0022] また、統計情報14aは、ステータスと、使用機能とを含む。ステータスは、各デバイスが使用可能か否かを示す情報であり、「稼働」あるいは「故障」で表される。配置装置10は、後述する配置処理において、ステータスが「稼働」であるデバイスを対象に、ネットワーク機能の配置先として選択する。

[0023] また、使用機能は、各デバイスに配置されている機能と、各機能によるリソース消費量の実績値とで表される。配置装置10は、後述する配置処理において、高位言語のソースコードを、ネットワーク機能の配置先として選択したデバイスのソースコードに変換し、変換したソースコードを用いてその機能のリソース消費量を測定し、この測定値を実績値として用いて、統計情報14aを更新する。なお、上記の「リソース消費量」は、各デバイスに配置されている全ての「使用機能」のリソース消費量の合計値に相当する。

[0024] 図2の説明に戻る。制御部15は、CPU等で実現され、メモリに記憶された処理プログラムを実行し、図2に示すように、受付部15a、性能予測部15b、デバイス選択部15c、高位言語編集部15d、変換部15eお

よび配置実行部 15 f として機能する。

[0025] 受付部 15 a は、複数種のハードウェアに対応したハードウェア記述言語である高位言語による機能の記述内容と、要求される性能を表す性能要件との入力を受け付ける。具体的には、受付部 15 a は、入力部 11 または通信制御部 13 を介して、配置処理の対象とされる機能の高位言語で記述されたソースコードと、スループットや遅延等の性能についての要求条件である性能要件の入力を受け付ける。

[0026] 性能予測部 15 b は、予測部として機能する。すなわち、性能予測部 15 b は、ハードウェアごとに、高位言語の記述内容と所定のアルゴリズムとを用いて、予測される処理性能である予測性能と、予測されるリソース消費量である予測リソース消費量とを算出する。

[0027] また、性能予測部 15 b は、高位言語の記述内容と、この記述内容に対応する機能について測定されたリソース消費量および性能情報とを用いた学習により、アルゴリズムを生成する。

[0028] 具体的には、性能予測部 15 b は、機能の高位言語のソースコードと、この機能についての実績値としての統計情報 14 a のリソース消費量および性能情報とを用いて、高位言語のソースコードと各ハードウェアのリソース消費量および性能情報との関係を学習する。これにより、性能予測部 15 b は、高位言語のソースコードと各ハードウェアのリソース消費量および性能情報との関係を算出するアルゴリズムを生成する。性能予測部 15 b は、生成したアルゴリズムを記憶部 14 に記憶させる。

[0029] そして、性能予測部 15 b は、ハードウェアごとに、生成したアルゴリズムを適用して、配置処理の対象として入力された機能の高位言語のソースコードから、予測される性能すなわち予測性能と、予測されるリソース消費量すなわち予測リソース消費量とを算出する。

[0030] デバイス選択部 15 c は、選択部として機能する。すなわち、デバイス選択部 15 c は、算出された予測性能および性能情報が性能要件を満たし、かつ予測リソース消費量とリソース消費量との合計値がリソース容量以下であ

るハードウェアを、配置先として選択する。

[0031] 具体的には、デバイス選択部 15 c は、ハードウェアごとに、入力された性能要件と算出された予測性能とハードウェアについての統計情報 14 a の性能情報とを比較する。そして、デバイス選択部 15 c は、予測性能、性能情報の双方が性能要件を満たす場合に、このハードウェアを配置先の候補とする。

[0032] 次に、デバイス選択部 15 c は、算出されたリソース予測リソース消費量と統計情報 14 a のリソース消費量との合計値を算出する。そして、デバイス選択部 15 c は、算出した合計値がこのハードウェアのリソース容量以下である場合に、このハードウェアを機能の配置先として選択する。

[0033] ここで、図 4 は、配置装置 10 の処理を説明するための説明図である。図 4 に示す例では、デバイス選択部 15 c は、選択機能部と比較機能部とで構成されている。そして、選択機能部が、性能予測部 15 b に対して、機能の配置先として適切か否かを判定する処理のターゲットのデバイスを順に指定している。この場合には、まず、選択機能部は、予め入力部 11 または通信制御部 13 を介して入力され記憶部 14 に記憶されている初期配置情報を参照し、初期配置情報で指定されているデバイスを、最初のターゲットのデバイスとして、デバイス選択部 15 c に通知する。

[0034] また、デバイス選択部 15 c の選択機能部は、受付部 15 a として機能して、配置処理の対象として入力された機能の高位言語のソースコードを受け付けて、性能予測部 15 b に転送する。

[0035] 性能予測部 15 b は、指定されたデバイスについて、高位言語のソースコードとアルゴリズムとを用いて、予測性能と予測リソース消費量とを算出し、予測結果としてデバイス選択部 15 c の比較機能部に転送する。このアルゴリズムは、性能予測部 15 b が、予め機能の高位言語のソースコードと、この機能についての実績値としての統計情報 14 a のリソース消費量および性能情報とを用いて生成したものである。

[0036] デバイス選択部 15 c の比較機能部は、受付部 15 a として機能して、性

能要件を受け付ける。また、比較機能部は、入力された性能要件と、予測結果と、このデバイスについての統計情報 14 a の性能情報およびリソース消費量とを比較する。

[0037] そして、比較機能部は、算出された予測性能および性能情報が性能要件を満たし、かつ予測リソース消費量とデバイスのリソース消費量との合計値がこのデバイスのリソース容量の範囲内である場合に、判定結果 OK として、このデバイスを配置先として選択する。

[0038] なお、比較機能部は、判定結果 OK ではない場合には、判定結果 NG として、ターゲットのデバイスの変更を選択機能部に指示する。このようにして、選択機能部が、性能予測部 15 b に対して、配置処理の配置先のターゲットのデバイスを順に指定する。ターゲットのデバイスの指定順は特に限定されないが、例えば、リソース消費量が少ない順、すなわち、デバイスのリソース容量からデバイスのリソース消費量を減算したリソース許容量が大きい順に指定されればよい。

[0039] 図 3 に示した統計情報 14 a において、入力された性能要件が遅延 5 m s、スループット 10 G b p s である場合に、ターゲットのデバイスが CPU の場合には、遅延 100 m s、スループット 1 G b p s であり、性能要件を満たさない。したがって、比較機能部は、判定結果 NG として、選択機能部にターゲットのデバイスの変更を指示する。

[0040] そこで、選択機能部は、ターゲットのデバイスを F P G A 1 に変更する。その場合に、統計情報 14 a において、F P G A 1 の遅延 1 m s、スループット 40 G b p s であり、性能要件を満たす。したがって、選択機能部は、予測性能が性能要件を満たし、かつ予測リソース消費量とリソース消費量との合計値がデバイスのリソース容量以下であれば、F P G A 1 を機能の配置先として決定する。

[0041] 例えば、F P G A 1 のリソース消費量が 10 %、予測リソース消費量が 20 %、またリソース容量が 100 % であれば、予測リソース消費量とリソース消費量との合計はリソース容量以下であるため、配置先として決定される

。なお、この場合には、配置後のFPGA1のリソース消費量は30%になる。また、この場合には、CPUのリソース消費量は変化しない。

[0042] なお、デバイス選択部15cの比較機能部は、さらに統計情報14aのステータスを参照し、ターゲットのデバイスが故障していないことを確認したうえで、配置先として決定する。

[0043] このように、配置装置10では、配置対象の機能の高位言語によるソースコードを用いて、機能によるリソース消費量と性能情報を予測することができる。そのため、シミュレーションや実機による動作確認を実行しなくても、機能の配置先のハードウェアを迅速かつ適切に選択することが可能となる。

[0044] 図2の説明に戻る。高位言語編集部15dは、編集部として機能する。すなわち、高位言語編集部15dは、ハードウェアのリソース容量からリソース消費量を減算した値であるリソース許容量が予測リソース消費量より小さい場合に、高位言語による記述内容を分割する。また、高位言語編集部15dは、リソース許容量が予測リソース消費量より大きい場合に、該リソース許容量の範囲で複数の機能の高位言語による記述内容を結合する。

[0045] 変換部15eは、機能の高位言語によるソースコードをハードウェアに対応したソースコードに変換する。また、配置実行部15fは、ハードウェアに対応したソースコードを用いて、統計情報を測定する。

[0046] 具体的には、高位言語編集部15dは、例えば図4に示すように、比較機能部による判定結果がNGの場合に、配置処理の対象の機能の高位言語のソースコードを用いて分割あるいは他の機能のソースコードと結合する等の編集を行う。また、高位言語編集部15dは、編集後のソースコードをデバイス選択部15cに引き渡す。その場合に、デバイス選択部15cは、編集後のソースコードで記述される機能を配置処理の対象に変更して、後続の処理を行う。

[0047] また、上記の配置処理により機能の配置先のデバイスが決定した場合に、変換部15eが、デバイス選択部15cから受け取った機能の高位言語によ

るソースコードを、決定したデバイスのソースコードに変換し、配置実行部 15 f に引き渡す。

[0048] 配置実行部 15 f は、変換部 15 e が変換したソースコードを用いて、機能のシミュレーションまたは実機における動作確認を実行する。そして、配置実行部 15 f は、その機能のリソース消費量を測定し、この測定値を実績値として用いて、統計情報 14 a を更新する。すなわち、配置後の機能のリソース消費量を用いて、統計情報 14 a の「使用機能」および「リソース消費量」を更新する。

[0049] この更新された統計情報 14 a を用いて、性能予測部 15 b が性能予測に用いるアルゴリズムを生成し記憶部 14 のアルゴリズムを更新する。これにより、アルゴリズムの精度が向上する。

[0050] [配置処理]

次に、図 5 を参照して、本実施形態に係る配置装置による配置処理について説明する。図 5 は、配置処理手順を例示するフローチャートである。図 5 のフローチャートは、例えば、ユーザが開始を指示したタイミングで開始される。

[0051] 最初に、配置装置は、各デバイスから、デバイスの種別と、遅延、スループット等の性能情報を取得して（ステップ S 1）、統計情報 14 a として記憶部 14 に記憶する（ステップ S 2）。また、配置装置は、各デバイスのリソース消費量を取得する（ステップ S 3）。

[0052] 次に、受付部 15 a が、配置処理対象の機能の高位言語によるソースコードと、性能要件との入力を受け付ける（ステップ S 4）。本実施形態では、受付部 15 a は、性能要件として、遅延要件およびスループット要件の入力を受け付ける。また、受付部 15 a は、配置処理のターゲットとするデバイスの初期値である初期配置情報との入力を受け付けて、統計情報 14 a のデバイスを識別する登録番号 i ($i = 0, 1, \dots$) についての初期化处理 ($i = 0$) を行う（ステップ S 4 ~ S 5）。

[0053] 次に、デバイス選択部 15 c の選択機能部が、初期配置情報が入力されて

いることを確認し（ステップS 6）、入力されている場合に（ステップS 6、Y e s）、初期配置情報のデバイスが配置先として存在することを確認する（ステップS 7）。選択機能部は、初期配置情報のデバイスが配置先として存在する場合に（ステップS 7、Y e s）、このデバイスを配置処理のターゲットのデバイスとして選択する（ステップS 8）。

[0054] 一方、選択機能部は、初期配置情報が入力されていない場合（ステップS 6、N o）、または初期配置情報のデバイスが配置先として存在しない場合には（ステップS 7、N o）、統計情報1 4 aの登録番号iに基づいて配置処理のターゲットのデバイスを選択する（ステップS 9）。

[0055] 選択機能部は、選択したデバイスを指定する情報と、高位言語のソースコードとを性能予測部1 5 bに転送する（ステップS 1 0）。その場合に、性能予測部1 5 bは、予測リソース消費量と予測性能とを算出し（ステップS 2 1）、算出した予測結果をデバイス選択部1 5 cの比較機能部に出力する（ステップS 2 2）。本実施形態では、予測性能として、性能要件に対応して、遅延およびスループットの予測値が比較機能部に出力される。

[0056] 比較機能部は、性能要件が入力されていることを確認する。比較機能部は、まず、遅延要件が入力されていることを確認し（ステップS 3 1）、入力されている場合に（ステップS 3 1、Y e s）、予想結果および統計情報1 4 aの遅延が遅延要件を満たすか否かを確認する（ステップS 3 2）。比較機能部は、予想結果および統計情報1 4 aの遅延が遅延要件を満たす場合に（ステップS 3 2、Y e s）、ステップS 3 3に処理を進め、満たさない場合に（ステップS 3 2、N o）、ステップS 3 6に処理を進める。

[0057] ステップS 3 3の処理では、比較機能部が、スループット要件が入力されていることを確認し、入力されている場合に（ステップS 3 3、Y e s）、予想結果および統計情報1 4 aのスループットがスループット要件を満たすか否かを確認する（ステップS 3 4）。比較機能部は、予想結果および統計情報1 4 aのスループットがスループット要件を満たす場合に（ステップS 3 4、Y e s）、ステップS 3 5に処理を進め、満たさない場合に（ステッ

プS 3 4、N o)、ステップS 3 6に処理を進める。

[0058] ステップS 3 5の処理では、比較機能部が、予測結果の予測リソース消費量と統計情報1 4 aのリソース消費量との合計値がこのデバイスのリソース容量以下の許容範囲か否かを確認する。比較機能部は、合計値が許容範囲ではない場合に(ステップS 3 5、N o)、ステップS 3 6に処理を進め、許容範囲である場合に(ステップS 3 5、Y e s)、このデバイスを機能の配置先として選択し、ステップS 3 7に処理を進める。

[0059] ステップS 3 6の処理では、比較機能部が、統計情報1 4 aのデバイスを識別する登録番号iのインクリメントを行って、ステップS 9に処理を戻す。

[0060] ステップS 3 7の処理では、比較機能部が、変換部1 5 eに、機能の配置先として選択したデバイスと高位言語のソースコードとを通知して、コンパイラ処理の実行を指示する(ステップS 3 7)。

[0061] 変換部1 5 eは、高位言語のソースコードを配置先のデバイスのソースコードに変換するコンパイラ処理を実行し、配置実行部1 5 fに引き渡す。配置実行部1 5 fは、コンパイラ処理された機能をデバイスに配置し(ステップS 4 1)、配置された機能のリソース消費量を測定し(ステップS 4 2)、この測定値を実績値として用いて統計情報1 4 aを更新する。これにより、一連の配置処理が終了する。

[0062] 以上、説明したように、本実施形態の配置装置1 0では、記憶部1 4が、機能の配置先の候補である複数種のハードウェアのそれぞれのリソース消費量と性能を表す性能情報とを含む統計情報1 4 aを記憶する。また、受付部1 5 aが、複数種のハードウェアに対応したハードウェア記述言語である高位言語による機能の記述内容と、要求される性能を表す性能要件との入力を受け付ける。また、性能予測部1 5 bが、ハードウェアごとに、前記記述内容と所定のアルゴリズムとを用いて、予測される性能である予測性能と、予測されるリソース消費量である予測リソース消費量とを算出する。また、デバイス選択部1 5 cが、算出された予測性能および統計情報1 4 aの性能情

報が性能要件を満たし、かつ算出された予測リソース消費量と統計情報 1 4 a のリソース消費量との合計値がリソース容量以下であるハードウェアを、配置先として選択する。

[0063] これにより、配置装置 1 0 では、配置対象の機能の高位言語によるソースコードを用いて機能によるリソース消費量と性能情報を予測することができる。そのため、シミュレーションや実機による動作確認を実行しなくても、機能の配置先のハードウェアを迅速かつ適切に選択することが可能となる。したがって、配置装置 1 0 の配置処理によれば、複数種のハードウェアへのネットワーク機能の配置を迅速に行うことが可能となる。

[0064] また、性能予測部 1 5 b は、高位言語による記述内容と、該記述内容に対応して測定されたリソース消費量および性能情報とを用いた学習により、アルゴリズムを生成する。これにより、性能予測に用いられるアルゴリズムの精度が向上する。

[0065] また、高位言語編集部 1 5 d は、ハードウェアのリソース容量からリソース消費量を減算した値であるリソース許容量が予測リソース消費量より小さい場合に、高位言語による記述内容を分割し、リソース許容量が予測リソース消費量より大きい場合に、該リソース許容量の範囲で複数の機能の高位言語による記述内容を結合する。これにより、さらに適切にネットワーク機能の配置を行うことが可能となる。

[0066] [プログラム]

上記実施形態に係る配置装置 1 0 が実行する処理をコンピュータが実行可能な言語で記述したプログラムを作成することもできる。一実施形態として、配置装置 1 0 は、パッケージソフトウェアやオンラインソフトウェアとして上記の配置処理を実行する配置プログラムを所望のコンピュータにインストールさせることによって実装できる。例えば、上記の配置プログラムを情報処理装置に実行させることにより、情報処理装置を配置装置 1 0 として機能させることができる。ここで言う情報処理装置には、デスクトップ型またはノート型のパーソナルコンピュータが含まれる。また、その他にも、情報

処理装置にはスマートフォン、携帯電話機やPHS (Personal Handyphone System) 等の移動体通信端末、さらには、PDA (Personal Digital Assistants) 等のスレート端末等がその範疇に含まれる。

[0067] また、配置装置10は、ユーザが使用する端末装置をクライアントとし、当該クライアントに上記の配置処理に関するサービスを提供するサーバ装置として実装することもできる。例えば、配置装置10は、性能要件と機能の高位言語によるソースコードとを入力とし、機能の配置先のハードウェアを出力する配置処理サービスを提供するサーバ装置として実装される。この場合、配置装置10は、Webサーバとして実装することとしてもよいし、アウトソーシングによって上記の配置処理に関するサービスを提供するクラウドとして実装することとしてもかまわない。以下に、配置装置10と同様の機能を実現する配置プログラムを実行するコンピュータの一例を説明する。

[0068] 図6は、配置プログラムを実行するコンピュータの一例を示す図である。コンピュータ1000は、例えば、メモリ1010と、CPU1020と、ハードディスクドライブインタフェース1030と、ディスクドライブインタフェース1040と、シリアルポートインタフェース1050と、ビデオアダプタ1060と、ネットワークインタフェース1070とを有する。これらの各部は、バス1080によって接続される。

[0069] メモリ1010は、ROM (Read Only Memory) 1011およびRAM 1012を含む。ROM 1011は、例えば、BIOS (Basic Input Output System) 等のブートプログラムを記憶する。ハードディスクドライブインタフェース1030は、ハードディスクドライブ1031に接続される。ディスクドライブインタフェース1040は、ディスクドライブ1041に接続される。ディスクドライブ1041には、例えば、磁気ディスクや光ディスク等の着脱可能な記憶媒体が挿入される。シリアルポートインタフェース1050には、例えば、マウス1051およびキーボード1052が接続される。ビデオアダプタ1060には、例えば、ディスプレイ1061が接続される。

- [0070] ここで、ハードディスクドライブ１０３１は、例えば、ＯＳ１０９１、アプリケーションプログラム１０９２、プログラムモジュール１０９３およびプログラムデータ１０９４を記憶する。上記実施形態で説明した各情報は、例えばハードディスクドライブ１０３１やメモリ１０１０に記憶される。
- [0071] また、配置プログラムは、例えば、コンピュータ１０００によって実行される指令が記述されたプログラムモジュール１０９３として、ハードディスクドライブ１０３１に記憶される。具体的には、上記実施形態で説明した配置装置１０が実行する各処理が記述されたプログラムモジュール１０９３が、ハードディスクドライブ１０３１に記憶される。
- [0072] また、配置プログラムによる情報処理に用いられるデータは、プログラムデータ１０９４として、例えば、ハードディスクドライブ１０３１に記憶される。そして、ＣＰＵ１０２０が、ハードディスクドライブ１０３１に記憶されたプログラムモジュール１０９３やプログラムデータ１０９４を必要に応じてＲＡＭ１０１２に読み出して、上述した各手順を実行する。
- [0073] なお、配置プログラムに係るプログラムモジュール１０９３やプログラムデータ１０９４は、ハードディスクドライブ１０３１に記憶される場合に限られず、例えば、着脱可能な記憶媒体に記憶されて、ディスクドライブ１０４１等を介してＣＰＵ１０２０によって読み出されてもよい。あるいは、配置プログラムに係るプログラムモジュール１０９３やプログラムデータ１０９４は、ＬＡＮ（Local Area Network）やＷＡＮ（Wide Area Network）等のネットワークを介して接続された他のコンピュータに記憶され、ネットワークインタフェース１０７０を介してＣＰＵ１０２０によって読み出されてもよい。
- [0074] 以上、本発明者によってなされた発明を適用した実施形態について説明したが、本実施形態による本発明の開示の一部をなす記述および図面により本発明は限定されることはない。すなわち、本実施形態に基づいて当業者等によりなされる他の実施形態、実施例および運用技術等は全て本発明の範疇に含まれる。

符号の説明

- [0075] 1 0 配置装置
- 1 1 入力部
- 1 2 出力部
- 1 3 通信制御部
- 1 4 記憶部
- 1 4 a 統計情報
- 1 5 制御部
- 1 5 a 受付部
- 1 5 b 性能予測部
- 1 5 c デバイス選択部
- 1 5 d 高位言語編集部
- 1 5 e 変換部
- 1 5 f 配置実行部

請求の範囲

- [請求項1] 機能の配置先の候補である複数種のハードウェアのそれぞれのリソース消費量と性能を表す性能情報とを含む統計情報を記憶する記憶部と、
- 複数種のハードウェアに対応したハードウェア記述言語である高位言語による機能の記述内容と、要求される性能を表す性能要件との入力を受け付ける受付部と、
- ハードウェアごとに、前記記述内容と所定のアルゴリズムとを用いて、予測される性能である予測性能と、予測されるリソース消費量である予測リソース消費量とを算出する予測部と、
- 算出された前記予測性能および前記性能情報が前記性能要件を満たし、かつ前記予測リソース消費量と前記リソース消費量との合計値がリソース容量以下であるハードウェアを、配置先として選択する選択部と、
- を備えることを特徴とする配置装置。
- [請求項2] 前記予測部は、さらに前記記述内容と、該記述内容に対応する機能について測定された前記リソース消費量および前記性能情報とを用いた学習により、前記アルゴリズムを生成することを特徴とする請求項1に記載の配置装置。
- [請求項3] ハードウェアのリソース容量から前記リソース消費量を減算した値であるリソース許容量が前記予測リソース消費量より小さい場合に、前記記述内容を分割し、前記リソース許容量が前記予測リソース消費量より大きい場合に、該リソース許容量の範囲で複数の機能の前記記述内容を結合する編集部を、さらに備えることを特徴とする請求項1または2に記載の配置装置。
- [請求項4] 機能の配置先の候補である複数種のハードウェアのそれぞれのリソース消費量と性能を表す性能情報とを含む統計情報を記憶する記憶部を備えた配置装置において実行される配置方法であって、

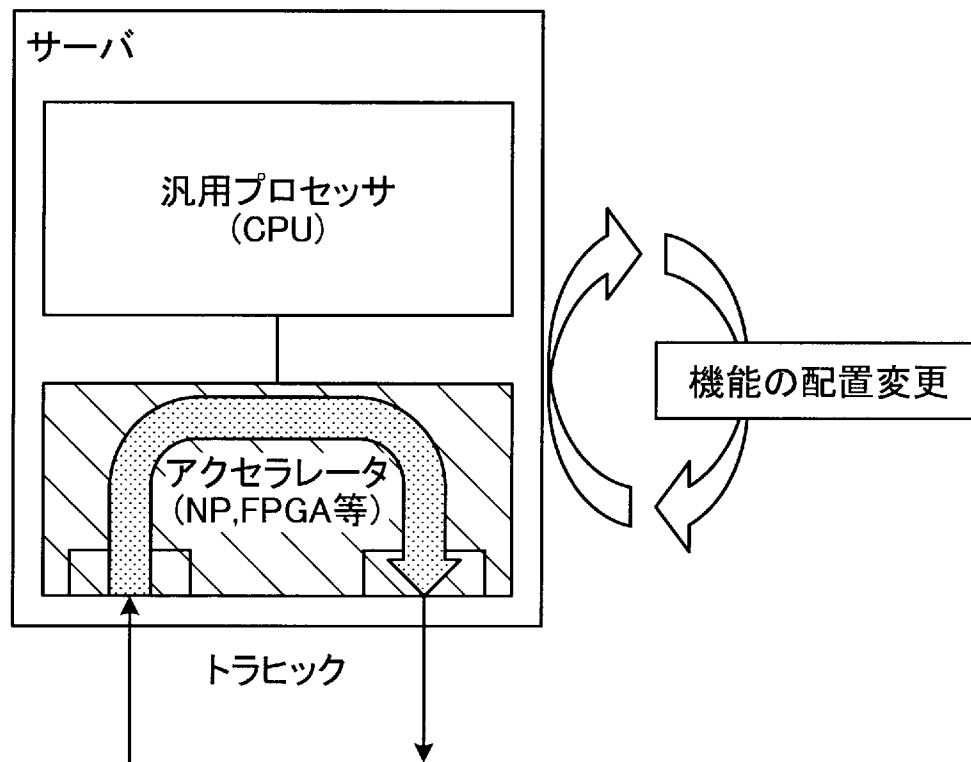
複数種のハードウェアに対応したハードウェア記述言語である高位言語による機能の記述内容と、要求される性能を表す性能要件との入力を受け付ける受付工程と、

ハードウェアごとに、前記記述内容と所定のアルゴリズムとを用いて、予測される性能である予測性能と、予測されるリソース消費量である予測リソース消費量とを算出する予測工程と、

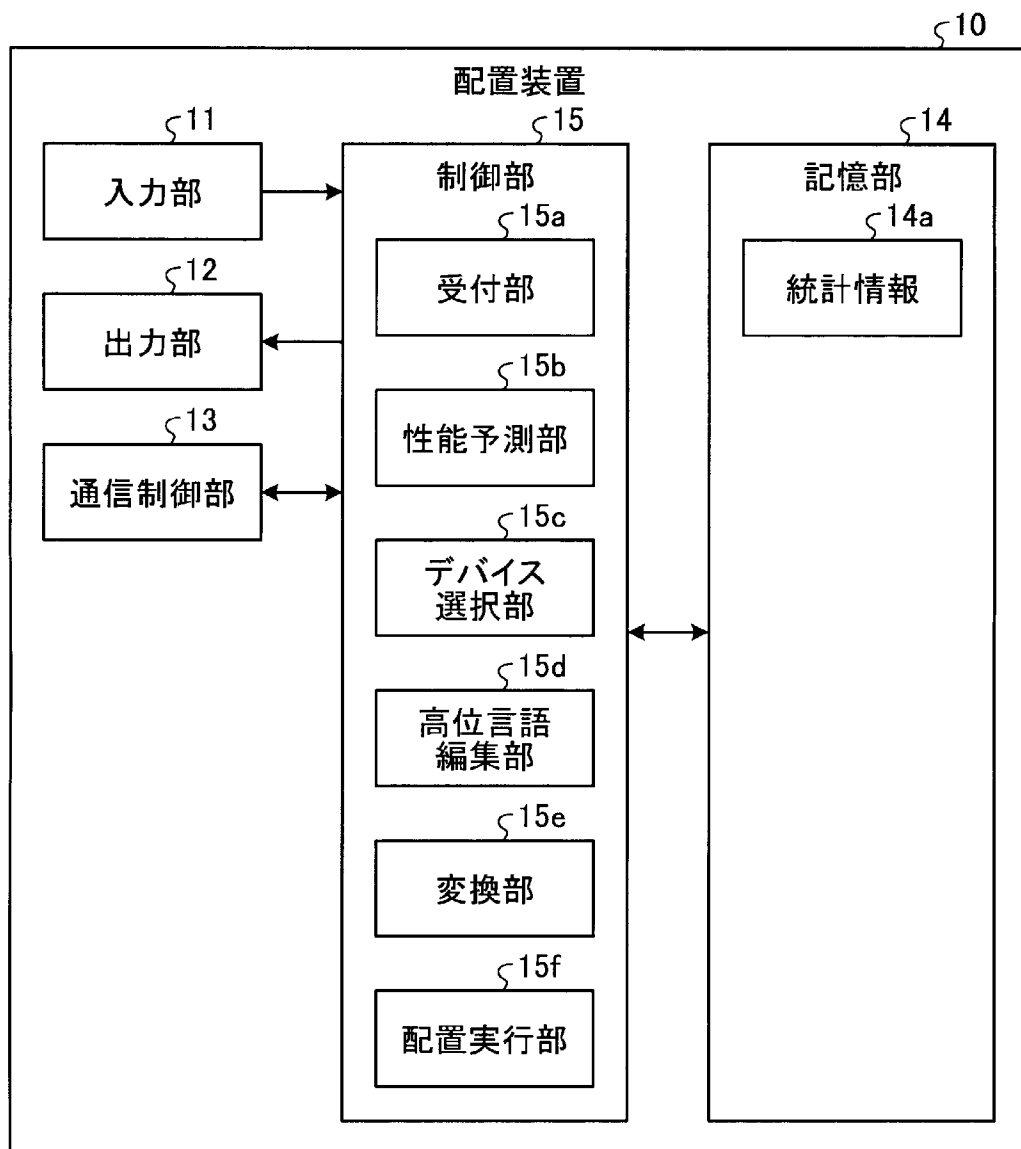
算出された前記予測性能および前記性能情報が前記性能要件を満たし、かつ前記予測リソース消費量と前記リソース消費量との合計値がリソース容量以下であるハードウェアを、配置先として選択する選択工程と、

を含んだことを特徴とする配置方法。

[図1]



[図2]



[図3]

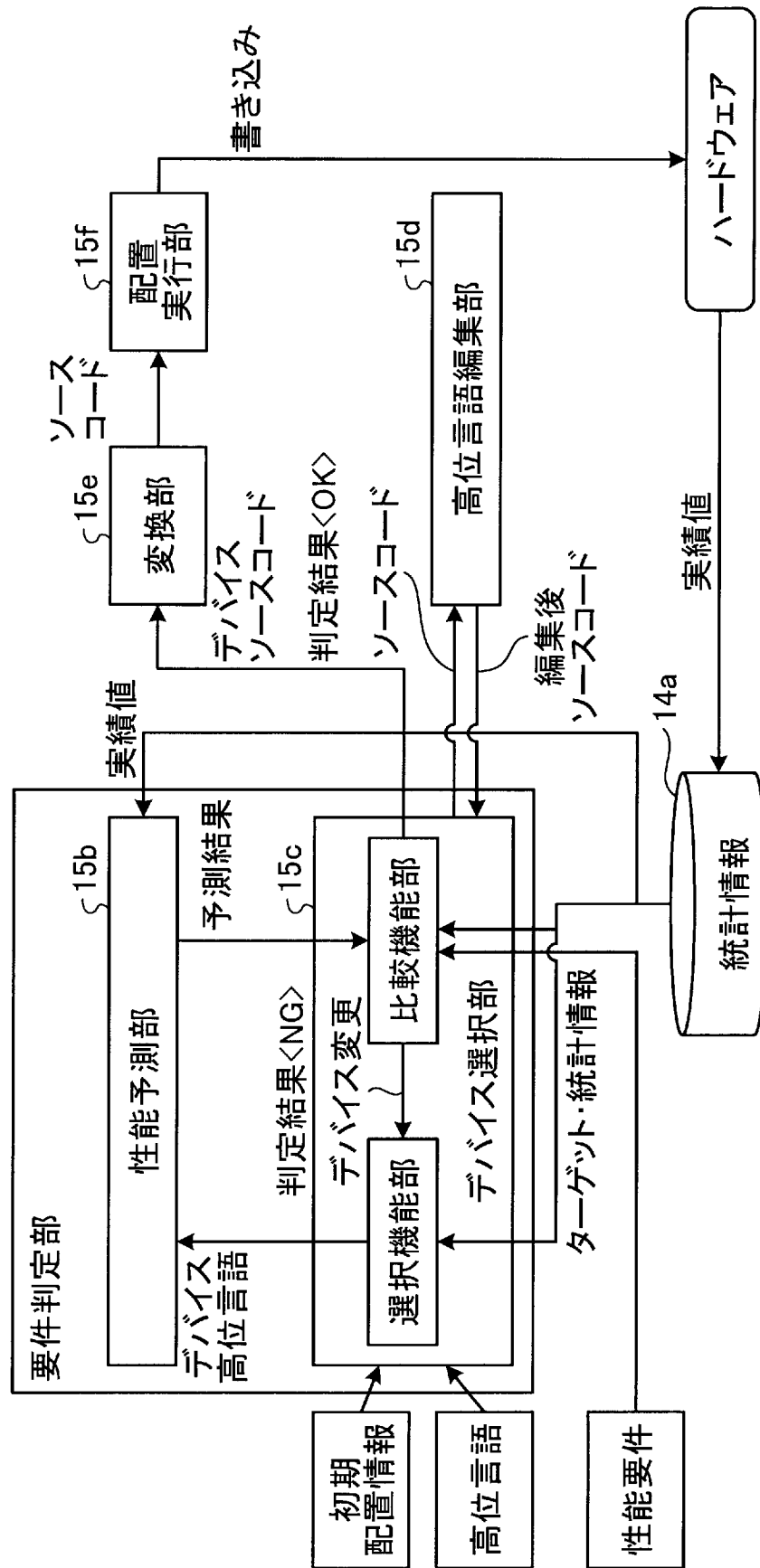
14a

デバイス	CPU	FPGA1	FPGA2	...
リソース消費量	80%	10%	0%	...
スループット	1Gbps	40Gbps	-	...
遅延	100ms	1ms	-	...
...	...	...	...	...
ステータス	稼動	稼動	故障	...
使用機能	機能A(30%) 機能B(50%)	機能C(10%)		

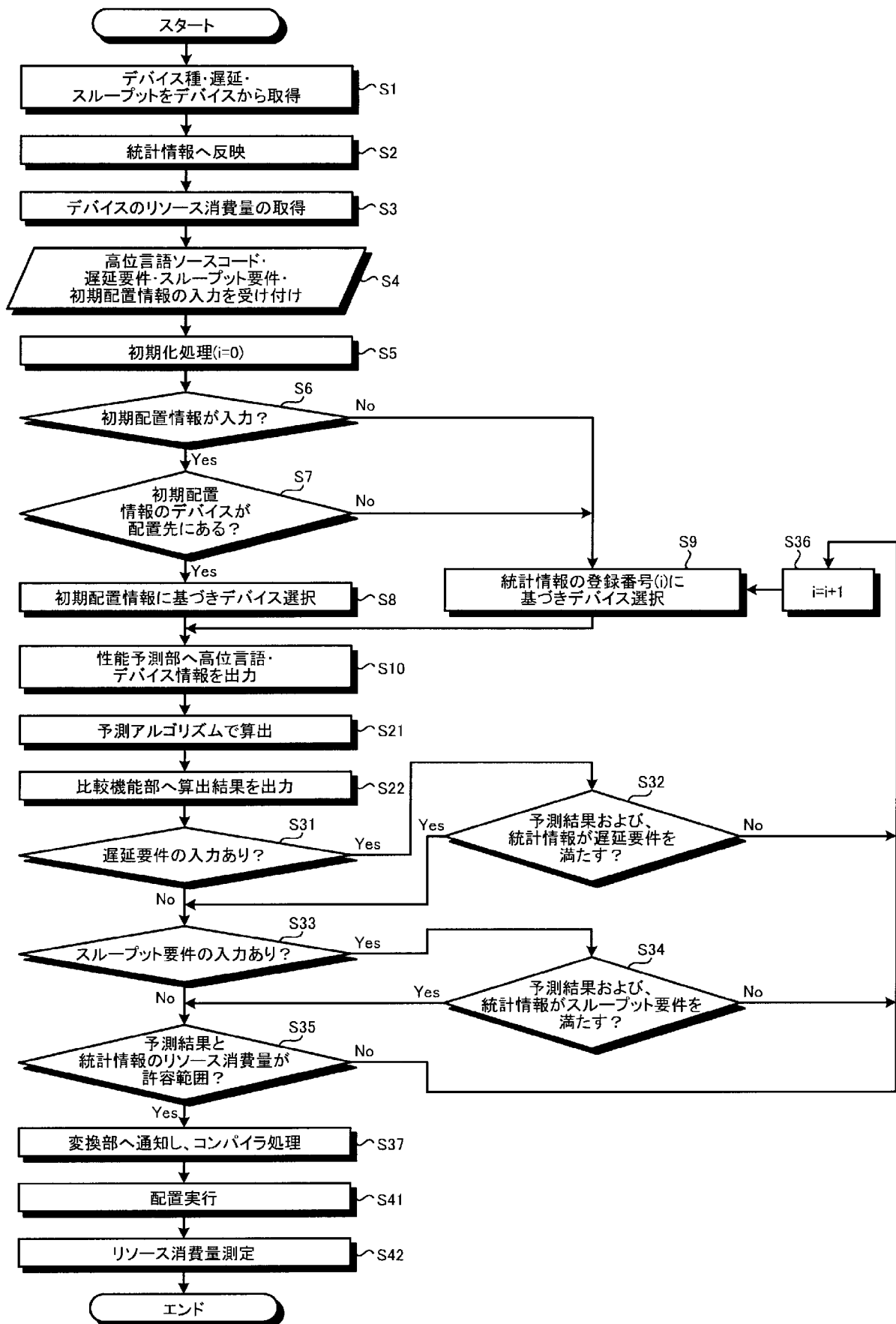
デバイス	CPU	FPGA1
コア数	16	FF数
コア数 使用量	10	FF使用量
メモリ数	1TB	LUT数
メモリ 使用量	800GB	LUT 使用量
...	...	...

リソース消費量詳細

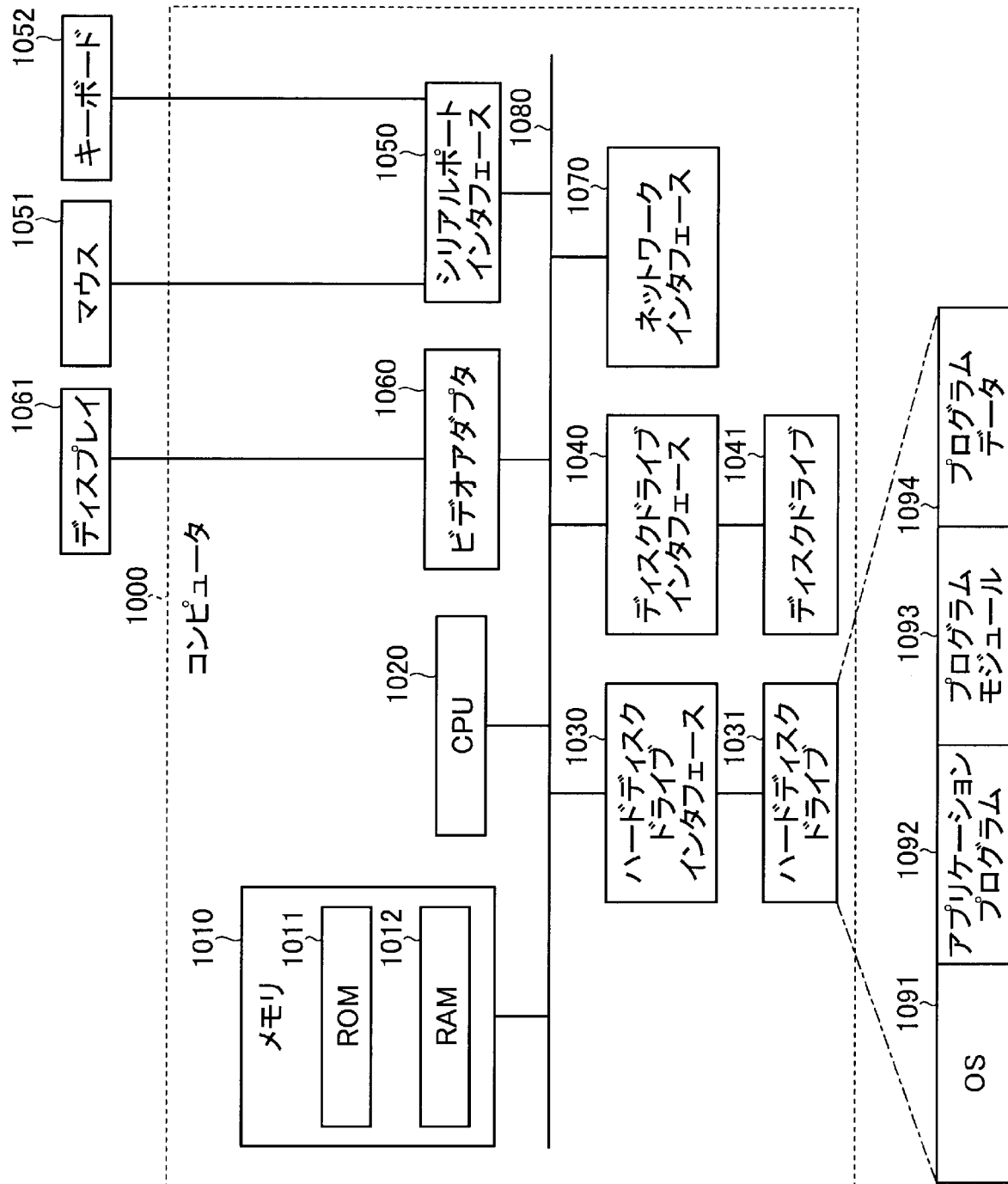
[図4]



[図5]



[図6]



INTERNATIONAL SEARCH REPORT

International application No.

PCT/JP2019/020775

A. CLASSIFICATION OF SUBJECT MATTER

Int.Cl. G06F9/50 (2006.01) i

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. G06F9/50

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 1971-2019

Registered utility model specifications of Japan 1996-2019

Published registered utility model applications of Japan 1994-2019

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.
A	JP 2017-129951 A (CANON INC.) 27 July 2017 (Family: none)	1-4
A	JP 2004-240530 A (MATSUSHITA ELECTRIC INDUSTRIAL CO., LTD.) 26 August 2004 (Family: none)	1-4
A	US 8826205 B2 (FREESCALE SEMICONDUCTOR, INC.) 02 September 2014 & CN 103780259 A	1-4

☐ Further documents are listed in the continuation of Box C. ☐ See patent family annex.

* Special categories of cited documents:	"I" 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
23 July 2019 (23.07.2019)

Date of mailing of the international search report
06 August 2019 (06.08.2019)

Name and mailing address of the ISA/
Japan Patent Office
3-4-3, Kasumigaseki, Chiyoda-ku,
Tokyo 100-8915, Japan

Authorized officer

Telephone No.

A. 発明の属する分野の分類（国際特許分類（IPC））

Int.Cl. G06F9/50(2006.01)i

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））

Int.Cl. G06F9/50

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	1971-2019年
日本国実用新案登録公報	1996-2019年
日本国登録実用新案公報	1994-2019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
A	JP 2017-129951 A（キヤノン株式会社）2017.07.27 （ファミリーなし）	1-4
A	JP 2004-240530 A（松下電器産業株式会社）2004.08.26 （ファミリーなし）	1-4
A	US 8826205 B2（FREESCALE SEMICONDUCTOR, INC.）2014.09.02 & CN 103780259 A	1-4

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矛盾するものではなく、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

23.07.2019

国際調査報告の発送日

06.08.2019

国際調査機関の名称及びあて先

日本国特許庁（ISA/J P）

郵便番号100-8915

東京都千代田区霞が関三丁目4番3号

特許庁審査官（権限のある職員）

大桃 由紀雄

5 B

5583

電話番号 03-3581-1101 内線 3545