

(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
2012年8月2日(02.08.2012)

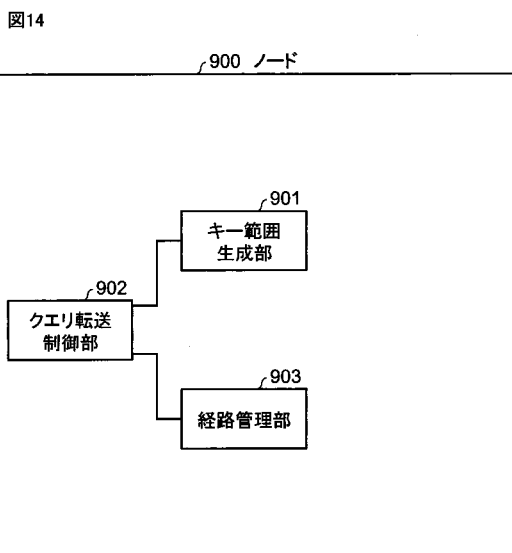


(10) 国際公開番号
WO 2012/102102 A1

- (51) 国際特許分類:
G06F 15/00 (2006.01) G06F 13/00 (2006.01)
 - (21) 国際出願番号: PCT/JP2012/050647
 - (22) 国際出願日: 2012年1月6日(06.01.2012)
 - (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 2011-012946 2011年1月25日(25.01.2011) JP
 - (71) 出願人(米国を除く全ての指定国について): 日本電気株式会社(NEC CORPORATION) [JP/JP]; 〒1088001 東京都港区芝五丁目7番1号 Tokyo (JP).
 - (72) 発明者; および
 - (75) 発明者/出願人(米国についてのみ): 佐藤 正 (SATO, Tadashi) [JP/JP]; 〒1088001 東京都港区芝五丁目7番1号日本電気株式会社内 Tokyo (JP).
 - (74) 代理人: 下坂 直樹 (SHIMOSAKA, Naoki); 〒1088001 東京都港区芝五丁目7番1号日本電気株式会社内 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, BR, BW, BY, BZ, CA, CH, CL, CN, CO, CR, CU, CZ, DE, DK, DM, DO, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, GT, HN, HR, HU, ID, IL, IN, IS, JP, KE, KG, KM, KN, KP, KR, KZ, LA, LC, LK, LR, LS, LT, LU, LY, MA, MD, ME, MG, MK, MN, MW, MX, MY, MZ, NA, NG, NI, NO, NZ, OM, PE, PG, PH, PL, PT, QA, RO, RS, RU, RW, 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, SZ, TZ, UG, ZM, ZW), ユーラシア (AM, AZ, BY, KG, KZ, MD, 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, ML, MR, NE, SN, TD, TG).
- 添付公開書類:
— 国際調査報告(条約第21条(3))

(54) Title: INFORMATION PROCESSING DEVICE, INFORMATION PROCESSING SYSTEM, INFORMATION PROCESSING METHOD AND INFORMATION PROCESSING PROGRAM

(54) 発明の名称: 情報処理装置、情報処理システム、情報処理方法および情報処理プログラム



900 NODE
 901 KEY RANGE GENERATION UNIT
 902 QUERY FORWARDING CONTROL UNIT
 903 PATH MANAGEMENT UNIT

(57) Abstract: The present invention reduces redundant query processing when processing large volumes of queries. An information processing device: receives a query specifying a certain value range; generates keys indicating values contained in the query; generates a key range group which contains at least one key range, which is a set of prescribed keys; stores a key indicating the bounds of the keys to be managed by each device; receives a key range group; specifies, for each key range contained in the received key range group, one key which is an address key on the basis of the magnitude relation between the value indicated by a prescribed key contained in the key range and the values indicated by the stored keys; appends the keys contained in the key range to a key range group, which is a set of keys associated with a node identifier of a device identified on the basis of the value of the specified address key; and forwards each key range group to the various devices identified by the node identifiers associated with each key range group.

(57) 要約:

[続葉有]

WO 2012/102102 A1

本発明は、大量のクエリを処理するときに冗長なクエリ処理を低減する。 情報処理装置は、ある値の範囲を特定するクエリを受け取り、クエリに含まれる値をそれぞれ示すキーを生成し、所定のキーの集合であるキー範囲を少なくとも一つ含むキー範囲群を生成し、各装置が管理すべきキーの境界を示すキーを記憶し、キー範囲群を受け取り、受け取ったキー範囲群に含まれるキー範囲毎に、キー範囲に含まれる所定のキーが示す値と記憶する各キーがそれぞれ示す値との大小関係に基づいて一のキーである宛先キーを特定し、キー範囲に含まれるキーを、特定した宛先キーの値に基づいて識別される装置のノード識別子に関連付けられるキーの集合であるキー範囲グループに追加し、各キー範囲グループを、各キー範囲グループに関連付けられているノード識別子で識別される装置へそれぞれ転送する。

明細書

発明の名称

情報処理装置、情報処理システム、情報処理方法および情報処理プログラム
5 ム

技術分野

本発明は、構造化P2P(Peer-to-Peer)におけるクエリの転送方法に関する。

背景技術

- 10 構造化P2P(Peer-to-Peer)では、システムを構成するノードは、外部システムから入力されるクエリを分担して処理する。ノードは、CPU(Central Processing Unit)やメモリから構成される装置のメモリ上に配備されるプログラムである。該プログラムをCPUが読み込んで実行することで、このCPUを備えるコンピュータは、データやクエリを転送したり、メモリ上に保持したりするノードとして機能する。クエリからハッシュ関数などによって生成される値(キー)によって、該クエリが到達すべきノードが決定される。
- 15 構造化P2Pではキーがとりうる値の空間であるキー空間が定義され、各ノードにキー空間の一部が範囲(キー範囲)としてそれぞれ割り当てられる。ノードに割り当てられるキーは、該ノードの割り当てキーと呼ばれ、割り当てられるキーの範囲は割り当てキー範囲と呼ばれる。割り当てキー範囲のうち、最小の値を示すキーは最小割り当てキーと呼ばれ、割り当てキー範囲のうち、最大の値を示すキーは最大割り当てキーと呼ばれる。
- 20 外部システムから入力されたクエリは、0以上のノードを経由して、該クエリから生成したキーが割り当てられたノードに送信される。クエリから複数のキーが生成される可能性があり、その場合、各キーが割り当てられたノードの全てに該クエリが送信される。経由されるノードの特定は、ルーティングプロトコルに依存する。あるキーが割り当てられたノードにクエリが送信されるとき、該キーは、該ノードの宛先キーと呼ばれる。
- 25 外部システムからデータが入力されると、ノードは、該データからキーを生成し、生成したキーが割り当てられたノードへ、データを転送する。よって、同じキーが生成されるデータとクエリとが、そのキーが割り当てられた同一のノードに送信される。各ノードがデータとクエリとの完全一致を判定する場合には、データを入力としてハッシュ関数により得られる値をキーとし、クエリで指定されるデータを入力としてハッシュ関数により得

られる値をキーとする構成であればよい。これにより、完全一致するデータとクエリとが同一のノードに到達することとなる。

外部システムから入力されるクエリを、そのクエリを処理すべきノードにルーティングする方法が開示されている(例えば非特許文献1参照)。

- 5 非特許文献1に記載された技術は、例えば以下のように動作する。P2Pシステムの各ノードは、多次元の属性のそれぞれの値を示す属性値を、空間充填曲線を使って一次元のデータを示すキーにマッピングする。そしてクエリが入力されると、各ノードは、該クエリからある値を示すキーの集合であるキー範囲を生成する。次に、ノードは、該クエリが各キー範囲の最小キーが割り当てられたノードに送信されるように、該クエリの
- 10 転送先のノードを決定する。そして該ノードは、その転送先のノードに該クエリを送信する。クエリは、0以上のノードを経由して、該キーが割り当てられたノードに送信される。クエリの転送先のノードは、クエリを受信し、受信したクエリを処理する。その処理の際、該ノードに割り当てられたキー範囲の最大値よりもクエリのキー範囲の最大値の方が大きい場合、該ノードはサクセッサ(successor)に受信したクエリを転送する。ノード
- 15 のサクセッサとは、該ノードの割り当てキーが示すそれぞれの値の最大値に1を加えた値を示すキーが割り当てられたノードである。ノードがそのサクセッサにクエリを転送する処理は、クエリから生成されたキー範囲の最大値が該ノードに割り当てられたキー範囲の最大値よりも小さくなるまで繰り返される。

[先行技術文献]

- 20 [非特許文献]

[非特許文献1] Prasanna Ganesan et al., "One torus to rule them all: Multi-dimensional queries in p2p systems," In WebDB '04: Proceedings of the 7th International Workshop on the Web and Databases, pages 19-24, New York, NY, USA, 2004. ACM Press.

- 25 [非特許文献2] Christian Bohm et al., "XZ-Ordering: A Space-Filling Curve for Objects with Spatial Extension," Proceedings of the 6th International Symposium on Advances in Spatial Databases, pages 75-90, July 20-23, 1999.

[非特許文献3] Ion Stoica et al., "Chord: A Scalable Peer-to-peer Lookup Service for Internet Applications," In Proceedings of the 2001 Conference

onApplications, Technologies, Architectures, and Protocols For Computer Communications. SIGCOMM '01. ACM, New York, NY, 149-160.

発明の概要

発明が解決しようとする課題

5 非特許文献1に記載された技術は、クエリが中継されるノードの割当キー範囲を考慮しないでクエリをいくつかのキー範囲に分割し、各キー範囲を送信する。よって、クエリが中継されるノードにおいて同一のクエリが同じノードを複数回経由する可能性がある。

本発明の目的の一例は、大量のクエリを転送するときに冗長なクエリ転送を低減する情報処理装置、情報処理システム、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である。

課題を解決するための手段

本発明の一態様における情報処理装置は、ある値の範囲を特定するクエリを受け取り、当該クエリに含まれる値をそれぞれ示すキーを生成し、所定のキーの集合であるキー範囲を少なくとも一つ含むキー範囲群を生成するキー範囲生成部と、各装置が管理すべきキーの境界を示すキーを記憶する経路管理部と、キー範囲群を受け取り、当該キー範囲群に含まれるキー範囲毎に、キー範囲に含まれる所定のキーが示す値と前記経路管理部が記憶する各キーがそれぞれ示す値との大小関係に基づいて一のキーである宛先キーを特定し、当該キー範囲に含まれるキーを、前記宛先キーの値に基づいて識別される装置のノード識別子に関連付けられるキーの集合であるキー範囲グループに追加する、クエリ転送制御部と、を備え、前記クエリ転送制御部は、各キー範囲グループを、当該各キー範囲グループに関連付けられているノード識別子で識別される装置へそれぞれ転送する。

本発明の一態様における情報処理システムは、少なくとも一つの情報処理装置を備え、当該情報処理装置は、ある値の範囲を特定するクエリを受け取り、当該クエリに含まれる値をそれぞれ示すキーを生成し、所定のキーの集合であるキー範囲を少なくとも一つ含むキー範囲群を生成するキー範囲生成部と、各情報処理装置が管理すべきキーの境界を示すキーを記憶する経路管理部と、キー範囲群を受け取り、当該キー範囲群に含まれるキー範囲毎に、キー範囲に含まれる所定のキーが示す値と前記経路

管理部が記憶する各キーがそれぞれ示す値との大小関係に基づいて一のキーである宛先キーを特定し、当該キー範囲に含まれるキーを、前記宛先キーの値に基づいて識別される情報処理装置のノード識別子に関連付けられるキーの集合であるキー範囲グループに追加する、クエリ転送制御部と、を備え、前記クエリ転送制御部は、各キー
5 範囲グループを、当該各キー範囲グループに関連付けられているノード識別子で識別される情報処理装置へそれぞれ転送する。

本発明の一態様における情報処理方法は、ある値の範囲を特定するクエリを受け取り、当該クエリに含まれる値をそれぞれ示すキーを生成し、所定のキーの集合であるキー範囲を少なくとも一つ含むキー範囲群を生成し、各装置が管理すべきキーの境界を示すキーを経路管理部に記憶し、キー範囲群を受け取り、当該キー範囲群に含まれるキー範囲毎に、キー範囲に含まれる所定のキーが示す値と前記経路管理部に記憶される各キーがそれぞれ示す値との大小関係に基づいて一のキーである宛先キーを特定し、当該キー範囲に含まれるキーを、前記宛先キーの値に基づいて識別される装置のノード識別子に関連付けられるキーの集合であるキー範囲グループに追加し、各
10 キー範囲グループを、当該各キー範囲グループに関連付けられているノード識別子で識別される装置へそれぞれ転送する。

本発明の一態様における情報処理プログラムは、コンピュータに、ある値の範囲を特定するクエリを受け取り、当該クエリに含まれる値をそれぞれ示すキーを生成し、所定のキーの集合であるキー範囲を少なくとも一つ含むキー範囲群を生成する処理と、各装置が管理すべきキーの境界を示すキーを経路管理部に記憶する処理と、キー範囲群を受け取り、当該キー範囲群に含まれるキー範囲毎に、キー範囲に含まれる所定のキーが示す値と前記経路管理部に記憶される各キーがそれぞれ示す値との大小関係に基づいて一のキーである宛先キーを特定する処理と、当該キー範囲に含まれるキーを、前記宛先キーの値に基づいて識別される装置のノード識別子に関連付けられる
15 キーの集合であるキー範囲グループに追加する処理と、各キー範囲グループを、当該各キー範囲グループに関連付けられているノード識別子で識別される装置へそれぞれ転送する処理と、を実行させる。

本発明の一態様における記録媒体は、コンピュータに、ある値の範囲を特定するクエリを受け取り、当該クエリに含まれる値をそれぞれ示すキーを生成し、所定のキーの集

合であるキー範囲を少なくとも一つ含むキー範囲群を生成する処理と、各装置が管理すべきキーの境界を示すキーを経路管理部に記憶する処理と、キー範囲群を受け取り、当該キー範囲群に含まれるキー範囲毎に、キー範囲に含まれる所定のキーが示す値と前記経路管理部に記憶される各キーがそれぞれ示す値との大小関係に基づいて

5 一のキーである宛先キーを特定する処理と、当該キー範囲に含まれるキーを、前記宛先キーの値に基づいて識別される装置のノード識別子に関連付けられるキーの集合であるキー範囲グループに追加する処理と、各キー範囲グループを、当該各キー範囲グループに関連付けられているノード識別子で識別される装置へそれぞれ転送する処理と、を実行させるための情報処理プログラムを格納する。

10 発明の効果

本発明の効果の一例は、大量のクエリを処理するときに冗長なクエリ処理を低減することができることである。

図面の簡単な説明

[図1] 図1は、本発明の第一の実施の形態における情報処理システムの構成を示すブロック図である。

[図2] 図2は、二次元の属性を示す例である。

[図3] 図3は、情報処理システムが、空間充填曲線を使って、多次元の属性を一次元の情報を示すキーにマッピングする手順の一例を示す図である。

[図4] 図4は、本発明の第一の実施の形態におけるノードの構成を示すブロック図である。

[図5] 図5は、経路管理部が記憶する情報の一例である。

[図6] 図6は、最小割当キー保持部が保持する最小割当キー表の例を示す図である。

[図7] 図7は、ノードが、自ノードの割当キー範囲を記憶する割当キー範囲記憶部106を備える構成を示すブロック図である。

[図8] 図8は、クエリ転送制御部が生成するメッセージの例を示す図である。

[図9] 図9は、ノードが備えるプログラム記憶装置を示すブロック図である。

[図10] 図10は、本発明の第一の実施の形態におけるノードのキー範囲生成部の動作の概要の一例を示すフローチャートである。

[図11]図11は、本発明の第一の実施の形態におけるノードのクエリ転送制御部の動作の概要の一例を示すフローチャートである。

[図12]図12は、本発明の第一の実施の形態におけるノードのクエリ転送制御部による対の生成処理の動作の概要の一例を示すフローチャートである。

5 [図13]図13は、第一の実施の形態におけるノードの動作の概要を示す具体例である。

[図14]図14は、本発明の第二の実施の形態におけるノードの構成を示すブロック図である。

[図15]図15は、ノードが備えるプログラム記憶装置を示すブロック図である。

10 発明を実施するための形態

本発明を実施するための形態について図面を参照して詳細に説明する。なお、各図面および明細書記載の各実施の形態において、同様の機能を備える構成要素には同様の符号が与えられている。

[第一の実施の形態]

15 図1は、本発明の第一の実施の形態における情報処理システム10の構成を示すブロック図である。

図1を参照すると、本発明の第一の実施の形態における情報処理システム10は、複数のノード100aないしノード100rを備える。ここで、ノード100aないしノード100rは、以下代表してノード100とも記載される。各ノード100は、少なくとも一つの他のノード
20 100と通信可能に接続されている。

情報処理システム10は、空間充填曲線を使って多次元の属性を、ある値を示すキーにマッピングする。図2は、二次元の属性を示す例である。二次元の属性のそれぞれの属性値の組み合わせが、あるキー803にマッピングされる。例えば、図2において、
25 二次元の属性からなる空間である属性空間804の少なくとも一部を示すクエリエリア805に含まれるある属性値の組み合わせ (x_a, y_a) が、あるキー803にマッピングされている。

図1において、情報処理システム10は、ノード100aからノード100rまでを備える。そして、それぞれのノード100に対して一次元のキー空間801の一部のキー範囲802が割り当てられている。

本明細書において、データは、複数の属性を含んでもよい。例えば、データは、位置情報として「緯度」属性と「経度」属性とを含んでもよい。属性は、属性の種別を識別する情報である属性名と、属性名に対する値(属性値)を含む。例えば、「緯度」属性は、属性名「緯度」と、属性値「北緯35度40分」とをそれぞれ含んでもよい。

- 5 本明細書において、クエリは、属性値の範囲を指定する情報であってもよい。例えば、クエリは、北緯35度30分から北緯35度50分まで、東経139度40分から東経139度50分までといったように各属性値の範囲を指定する情報であってもよい。「北緯35度40分、東経139度46分」というデータは、前述のクエリで指定される範囲に含まれる。よって該データは、前述のクエリ「北緯35度30分から北緯35度50分まで、東経139度40分から東経139度50分まで」に合致する。

次に、情報処理システム10が、空間充填曲線を使って、多次元の属性を一次元の情報を示すキーにマッピングする手順について説明する。

図3は、情報処理システム10が、空間充填曲線を使って、多次元の属性を一次元の情報を示すキーにマッピングする手順の一例を示す図である。

- 15 本例は、属性xと属性yとを含む二次元の属性が一次元の情報を示すキーにマッピングされる例である。まず、属性xの属性値範囲と属性yの属性値範囲とから、属性xサブキー範囲[011, 101]と属性yサブキー範囲[00, 01]が生成される。ここで、サブキーとは、キーを生成するために使用される値である。図3の例では、属性xのサブキーは、3ビットである。また属性yのサブキーは、2ビットである。本例において、属性xサブキー範囲に含まれるサブキー集合{011, 100, 101}と属性yサブキー範囲に含まれるサブキー集合{00, 01}の全組み合わせについて、各ビットを交互に並べたビット列がキーである。例えば、「011」と「00」の各ビットが交互に並べられると、「00101」が生成される。以上の処理により、属性xの属性値範囲と属性yの属性値範囲から、2×3個(=6個)のキーが生成される。本明細書では、サブキーの組み合わせである、「011」と「00」と、からキーを生成することを、「(011, 00)からキーを生成する」と表記する。上記のキー生成方法は例示であって、キー生成方法は、サブキー集合に含まれるビットの組み合わせに基づいてキーが生成される方法であればよい。

図4は、本発明の第一の実施の形態におけるノード100の構成を示すブロック図である。

図4を参照すると、本発明の第一の実施の形態におけるノード100は、キー範囲生成部101と、クエリ転送制御部102と、経路管理部103と、クエリ処理部104と、最小割当キー保持部105と、を備える。

5 キー範囲生成部101は、クエリを図示しない外部システムから受信する。そしてキー範囲生成部101は、該クエリから一または複数のキー範囲を含むキー範囲群のデータを生成する。キー範囲とは、キーがとりうる値の空間であるキー空間の一部の範囲である。あるキー範囲に含まれる各キーの値は、連続する値を示す。

10 キー範囲生成部101がクエリからキー範囲群を生成する方法として、非特許文献2が開示する手段が挙げられる。例えば、クエリとして、緯度の範囲と経度の範囲で表されるエリアが指定される場合、非特許文献2が開示する方法では、以下のように処理が行われる。まず、キー範囲生成部101は、緯度と経度で表すことができる全エリアに対して、そのエリアがクエリに完全に包含されるかどうかを検証する。包含されていなければ、キー範囲生成部101は、そのエリアを四等分する。四等分された各エリアに対して、一つのエリアがクエリとオーバーラップする場合に、キー範囲生成部101は、その
15 エリアをさらに四等分する。キー範囲生成部101は、以上のような処理を再帰的に行う。以上の処理が完了すると、キー範囲生成部101は、クエリに完全に包含されるエリアの集合と、クエリとオーバーラップしないエリアの集合と、を求めることができる。キー範囲生成部101は、クエリに完全に包含されるエリアに基づいて一つのキー範囲を生成することができる。非特許文献2が開示する手段は、分割されたエリアから連続した
20 キーが生成されるように、複数の属性の属性値が空間充填曲線に基づいて一次元のキーに変換されていることが前提となる。

キー範囲生成部101は、生成した各キー範囲について、キー範囲を、自ノード100の割当キー範囲とオーバーラップするキー範囲と、オーバーラップしないキー範囲とに分割する。例えばキー範囲生成部101は、自ノード100の割当キー範囲の最小割当
25 キー k に基づいて、 k 以上の値を示すキーのキー範囲と、 k 未満の値を示すキーのキー範囲とに分割する。つまり、キー範囲生成部101は、キー範囲 $[k_{min}, k_{max}]$ が自ノード100に割り当てられたキー範囲(割当キー範囲)の最小割当キー k とその k より小さい値を示すキー $k-1$ とを含む場合に、以下の処理を行う。すなわちキー範囲生成部101は、キー範囲 $[k_{min}, k_{max}]$ を、二つのキー範囲 $[k_{min}, k-1]$ と $[k, k_{max}]$

x]とに分割する。

キー範囲生成部101は、自ノード100が受信したクエリと、生成したキー範囲を含むキー範囲群と、をクエリ転送制御部102に渡す。

クエリ処理部104は、後述のクエリ転送制御部102からクエリを受け取り、受け取ったクエリを処理する。クエリ処理部104が行う処理は、情報処理システム10の応用方法によって異なる。例えば、情報処理システム10は、Publish/Subscribe型のシステムとして利用されうる。

Publish/Subscribe型のシステムでは、アプリケーションが、あらかじめデータの条件(データの条件とは、これまでの説明のクエリに相当する)をシステムに登録する。データの条件に合致するイベントが発生すると、そのイベントは、データの条件に登録したアプリケーションに通知される。イベントがシステムに入力されることは、Publishと呼ばれる。またデータの条件がシステムに登録されることは、Subscribeと呼ばれる。

本実施の形態において、例えばあらかじめ、ユーザがデータの条件(データ条件)を情報処理システム10に登録しておく。そしてデータ条件にマッチするデータが情報処理システム10に入力されたら、該データを、データ条件に登録した外部システムに通知する。この場合、データ条件とは、クエリに相当する。よってクエリ処理部104は、クエリをメモリ上に記憶し、データが入力されたときにクエリとデータとの合致判定をする機能手段を備える。その後、データがクエリ処理部104に入力されたら、クエリ処理部104は、メモリ上に保持しているクエリとの合致判定を行う。クエリ処理部104は、合致したと判定した場合に、所定の通知先にその旨を通知する。

経路管理部103は、キー(宛先キー)に対して、該キーが割り当てられたノードに到達するために、次の転送先であるノードのアドレス(ノードアドレス)を算出する。例えば、経路管理部103は、Chord(例えば、非特許文献3参照)などのルーティングプロトコルに基づいて、ノードアドレスを算出する。Chordは、宛先キーに最も近いキーが割り当てられたノードを次の転送先として特定するプロトコルである。

経路管理部103は、他のノード100との経路を示す情報(例えば、Chordであれば、Finger tableに相当する)を記憶する。図5は、経路管理部103が記憶する情報の一例である。この情報はフィンガーテーブル825とも呼ばれる。図5において、「start」は、自ノード100に対応付けられているキーが示す値に所定数だけ(例えば 2^i 、ただ

しは0以上の整数)加算した値を示すキーである。また、「int.」は、「start」のキーに対応するキー範囲を示す情報である。例えばキーの値がnであるキーに対応付けられているノードにおいて「start」は、 $n+2i$ (ただしiは0以上の整数)である。そしてこのキー「 $n+2i$ 」に対応付けられている「int.」は、 $n+2i \leq k < 2i+1$ を満たすキーkの
5 集合、すなわち、キーkのキー範囲である。「succ.」は、int. で示されている各キーに対応するノードのうちstartに対応付けられているキーの値に最も近いキーに対応付けられている他のノード100を示す識別子である。

キーはいずれかのノード100を識別できる情報であってもよい。また、キーは、ノード100を識別できる識別子、またはノードアドレスであってもよい。キーは、あるノードを
10 識別できる識別子またはノードアドレスに一意に対応する情報であってもよい。

図5において、「int.」は、「succ.」で識別されるノードに対応付けられるキーの集合を示す情報であってもよい。また、「succ.」に含まれるノードの識別子は、ノードのノードアドレスであってもよい。

各ノード100は、この経路を示す情報を生成するために、自ノード100のノードアドレスと自ノードに割り当てられているキーとを他のノード100に通知する。経路管理部103は、通知される情報を元に他ノード100との経路を示す情報を生成し、その情報を記憶する。
15

最小割当キー保持部105は、他のノード100の最小割当キーを保持する。最小割当キー保持部105が保持する最小割当キー表の例が図6に示されている。図6に示される例は、ノード100bが保持する最小割当キー表806、ノード100jが保持する最小割当キー表807およびノード100qが保持する最小割当キー表808、である。図6を参照すると、ノード100bの最小割当キー保持部105は、ノードアドレスc、d、f、jに対して、それらのノードの最小割当キーが、それぞれ、 $k_c - \min$ 、 $k_d - \min$ 、 $k_f - \min$ 、 $k_j - \min$ であることを記憶している。
20

経路管理部103が他ノード100との経路を示す情報を生成するときに、ノード100がノードアドレスを別のノード100に通知する処理を行う。該ノード100の最小割当キーは、その通知メッセージに含まれてもよい。最小割当キー保持部105は、経路管理部103が受け取った通知メッセージに含まれるノードアドレスと、最小割当キーとを対応付けて記憶する。これによって、経路管理部103が他ノード100との経路を示す情
25

報を構築する際に、最小割当キー保持部105は、最小割当キー表を記憶することができる。

クエリ転送制御部102は、キー範囲生成部101、あるいは、他のノード100のクエリ転送制御部102からクエリとキー範囲群を受信する。そして、クエリ転送制御部102は、自ノードの割当キー範囲が、キー範囲群のいずれかとオーバーラップするか否かを検証する。すなわち、クエリ転送制御部102は、キー範囲群に含まれるキーの中に自ノードの割当キー範囲にも含まれるキーがあるか否かを検証する。

自ノードの割当キー範囲は、クエリ転送制御部102に記憶されていてもよい。あるいは、図7に示されるように、ノード100が、自ノードの割当キー範囲を記憶する割当キー範囲記憶部106を備えていてもよい。

クエリ転送制御部102は、自ノードの割当キー範囲が、キー範囲群のいずれかとオーバーラップすると判定した場合に、受信したクエリをクエリ処理部104に渡す。そしてクエリ転送制御部102は、キー範囲群から、割当キー範囲に完全に包含されるキー範囲を削除する。例えば、自ノードの割当キー範囲が $[k1, k2]$ であり、キー範囲群のうちの一つのキー範囲が $[kmin, kmax]$ であると仮定する。このとき、例えば $k1 \leq kmin < kmax \leq k2$ の場合が、『割当キー範囲に完全に包含されるキー範囲』の例に該当し、このキー範囲 $[kmin, kmax]$ は削除される。

またクエリ転送制御部102は、キー範囲群のうち、自ノードの割当キー範囲とオーバーラップするキー範囲を二つのキー範囲に分割する。そしてクエリ転送制御部102は、分割したキー範囲のうち、オーバーラップする部分を削除する。例えば自ノードの割当キー範囲に含まれるキーの中で最大値を示すキーである最大割当キーが k であると仮定する。そしてキー範囲 $[kmin, kmax]$ について、 $kmin < k \leq kmax$ なる関係が満たされるとき、クエリ転送制御部102は、キー範囲 $[kmin, kmax]$ を二つのキー範囲 $[kmin, k]$ と $[k+1, kmax]$ に分割する。またクエリ転送制御部102は、キー範囲 $[kmin, k]$ を削除する。

クエリ転送制御部102は、一以上のキー範囲を含む集合であるキー範囲グループとそのキー範囲グループの送付先であるノードのノードアドレスとを対応付けた対を生成する。クエリ転送制御部102は、キー範囲群を構成する各キー範囲 $[kmin, kmax]$ について以降に示す処理を行う。

まずクエリ転送制御部102は、経路管理部103を参照し、 k_{max} に基づいて特定されるノードアドレス“addr”を取得する。次にクエリ転送制御部102は、ノードアドレス“addr”に対応付けられている最小割当キー k を最小割当キー保持部105から読み出す。

- 5 そしてクエリ転送制御部102は、キー範囲 $[k_{min}, k_{max}]$ のうち、読み出した最小割当キー k が示す値以上であり、 k_{max} が示す値以下である値を示すキー範囲 $[k, k_{max}]$ のキーを特定する。そしてクエリ転送制御部102は、この特定したキーをノードアドレス“addr”に送信するキー範囲グループに追加する。そしてクエリ転送制御部102は、残りのキー範囲 $[k_{min}, k-1]$ を新たに $[k_{min}, k_{max}]$ とみなして前述の処理を繰り返し行う。

- 10 一方、クエリ転送制御部102は、読み出した最小割当キー k が示す値よりも k_{min} が示す値のほうが小さいと判定した場合に、キー範囲 $[k_{min}, k_{max}]$ をノードアドレス“addr”に送信するキー範囲グループに含まれると特定する。もし、“addr”に対応するキー範囲グループがない場合、クエリ転送制御部102は、新たにキー範囲グループを生成し、そのキー範囲グループを“addr”に送信するキー範囲グループと特定する。

- 15 クエリ転送制御部102は、特定されたキー範囲グループ毎に、各キー範囲グループに対応付けられているノードアドレス“addr”で特定されるノード100へ、そのキー範囲グループを送信する。具体的には、クエリ転送制御部102は、生成した対毎に図8に示す構造のメッセージを生成する。このメッセージ809は、クエリ810と、キー範囲グループに含まれる各キー範囲811と、各キー範囲の最小キー812と最大キー813とを含む。最小キー812は、そのキー範囲に含まれるキーの中で最小値を示すキーである。また、最大キー813は、そのキー範囲に含まれるキーの中で最大値を示すキーである。

- 20 キー範囲生成部101と、クエリ転送制御部102と、経路管理部103と、クエリ処理部104とは、例えば、プログラム(情報処理プログラム)によって動作するCPUによって実現される。キー範囲生成部101と、クエリ転送制御部102と、経路管理部103と、クエリ処理部104とが、同一のCPUによって実現されてもよい。この場合、プログラムは、例えば、ノード100が備えるプログラム記憶装置107(図9参照)に記憶される。CPU108は、そのプログラムを読み込み、そのプログラムにしたがって、キー範囲生成

部101と、クエリ転送制御部102と、経路管理部103と、クエリ処理部104として動作すればよい。

また、前述のプログラムのコードを記録した記録媒体(または記憶媒体)が、ノード100に供給され、ノード100が記録媒体に格納されたプログラムのコードを読み出し実行してもよい。すなわち、本発明は、第一の実施の形態におけるノード100が実行するためのソフトウェア(情報処理プログラム)を一時的に記憶するまたは非一時的に記憶する記録媒体110も含む。

最小割当キー保持部105は、例えば、ハードディスク109のような記憶装置によって実現される。

10 次に、本発明の第一の実施の形態におけるノード100の動作の一例について説明する。図10は、本発明の第一の実施の形態におけるノード100のキー範囲生成部101の動作の概要の一例を示すフローチャートである。

キー範囲生成部101は、図示しない外部システムからクエリを受け取る(ステップA1)。そして、キー範囲生成部101は、入力されたクエリからキー範囲1、キー範囲2、およびキー範囲3を含むキー範囲群を生成する(ステップA2)。

キー範囲生成部101は、生成した各キー範囲について以下の処理を行う。あるキー範囲が $[kmin, kmax]$ であると仮定する。キー範囲 $[kmin, kmax]$ が自ノード100の割当キー範囲の最小割当キー k と、この k より小さい値を示すキー $k-1$ と、を含む場合、キー範囲生成部101は、以下の処理を実行する。すなわちキー範囲生成部101は、キー範囲 $[kmin, kmax]$ を k に基づいて二つに分割する(ステップA3)。具体的には、キー範囲生成部101は、キー範囲 $[kmin, kmax]$ を二つのキー範囲 $[kmin, k-1]$ と $[k, kmax]$ に分割する。

キー範囲生成部101は、すべてのキー範囲について処理が行われたか否か判定する(ステップA4)。すべてのキー範囲について、処理が行われていない場合(ステップA4のNo)、キー範囲生成部101の処理は、ステップA3に戻る。すべてのキー範囲について、処理が行われた場合(ステップA4のYes)、キー範囲生成部101は次の処理を行う。すなわちキー範囲生成部101は、自ノード100が受信したクエリと、生成したキー範囲群とをクエリ転送制御部102に渡す(ステップA5)。

図11は、本発明の第一の実施の形態におけるノード100のクエリ転送制御部102

の動作の概要の一例を示すフローチャートである。

クエリ転送制御部102は、キー範囲生成部101、または他のノード100のクエリ転送制御部102から少なくともクエリとキー範囲群とを含む情報を受け取る(ステップB1)。そしてクエリ転送制御部102は、受け取ったキー範囲群に含まれるいずれかのキー範囲が自ノード100の割当キー範囲とオーバーラップするか否か検証する(ステップB2)。

クエリ転送制御部102は、受け取ったキー範囲群に含まれるいずれかのキー範囲が自ノード100の割当キー範囲とオーバーラップすると判定した場合(ステップB2のYes)、クエリをクエリ処理部104に渡す(ステップB3)。そして、クエリ転送制御部102は、受け取ったキー範囲群に含まれる各キー範囲から、自ノード100の割当キー範囲に完全に包含されるキー範囲を削除する(ステップB4)。また、クエリ転送制御部102は、受け取ったキー範囲群に含まれる各キー範囲のうち、自ノード100の割当キー範囲とオーバーラップするキー範囲 $[kmin, kmax]$ を二つのキー範囲 $[kmin, k]$ と $[k+1, kmax]$ に分割する。そして、クエリ転送制御部102は、キー範囲 $[kmin, k]$ を削除する(ステップB5)。ここで、 k は自ノード100の最大割当キーである。

クエリ転送制御部102は、ステップB5の処理後、またはステップB2において受け取ったキー範囲群に含まれるいずれかのキー範囲が自ノード100の割当キー範囲とオーバーラップしないと判定した場合(ステップB2のNo)、以下の処理を行う。すなわちクエリ転送制御部102は、キー範囲グループとノードアドレスとを含む0以上の対を生成する(ステップB6)。ステップB6におけるクエリ転送制御部102の処理は、後で説明される。クエリ転送制御部102は、生成した各対について、クエリと、対の一方であるキー範囲グループとに基づいて図8に示される構造のメッセージを生成する。そしてクエリ転送制御部102は、生成したメッセージを対のもう一方であるノードアドレスで特定されるノードに送信する(ステップB7)。

図12は、本発明の第一の実施の形態におけるノード100のクエリ転送制御部102による前述ステップB6における処理の動作の概要の一例を示すフローチャートである。

クエリ転送制御部102は、これまでのステップにより生成されている各キー範囲 $[kmin, kmax]$ についてそれぞれ以下の処理を行う(ステップB601)。

クエリ転送制御部102は、キー範囲に含まれるキーの中で最大値を示すキー k_{max} を経路管理部103に渡し、経路管理部103からその k_{max} に対応するノードアドレス“addr”を受け取る(ステップB602)。

次に、クエリ転送制御部102は、ノードアドレス“addr”に対応付けられている最小割
5 当キー k を最小割当キー保持部105から読み出す(ステップB603)。

クエリ転送制御部102は、 k_{min} が示す値よりも k が示す値が大きいか否か判定する(ステップB604)。クエリ転送制御部102は、 $k_{min} < k$ であると判定した場合(ステップB604のYes)、キー範囲 $[k_{min}, k_{max}]$ を二つのキー範囲 $[k_{min}, k-1]$ と $[k, k_{max}]$ に分割する(ステップB605)。

10 次に、クエリ転送制御部102は、キー範囲 $[k, k_{max}]$ をノードアドレス“addr”に対応するキー範囲グループに追加する。もし“addr”に対応するキー範囲グループがない場合、クエリ転送制御部102は、新たにキー範囲グループを生成し、該キー範囲グループを“addr”と関連付ける(ステップB606)。以降、クエリ転送制御部102は、キー範囲 $[k_{min}, k-1]$ を $[k_{min}, k_{max}]$ と見なす(ステップB607)。そして、クエリ転
15 送制御部102の処理は、ステップB602へ移行する。

一方、クエリ転送制御部102は、 $k_{min} < k$ でないと判定した場合(ステップB604のNo)、キー範囲 $[k_{min}, k_{max}]$ をノードアドレス“addr”に対応するキー範囲グループに追加する。もし“addr”に対応するキー範囲グループがない場合、クエリ転送制御部102は、新たにキー範囲グループを生成し、該キー範囲グループを“addr”と関連付け
20 る(ステップB608)。

クエリ転送制御部102は、ステップB602-B608における処理をキー範囲毎に行う(ステップB609)。

図13は、第一の実施の形態におけるノード100bの動作の概要を示す具体例である。本例では、ノード100bが、キー範囲1(814)、キー範囲2(815)およびキー範囲
25 3(818)に対応するクエリのサブスクリプトを受け、それぞれ転送を行う。

キー範囲1(814)は、 $[k_{1min}, k_{1max}]$ と表される。そしてキー範囲1(814)は、ノード100qの割当キー範囲 $[k_{qmin}, k_{qmax}]$ に対して、 $k_{qmin} < k_{1min} < k_{1max} < k_{qmax}$ という関係を満たす。また、キー範囲2(815)は、 $[k_{2min}, k_{2max}]$ と表される。そしてキー範囲2(815)は、ノード100mの割当キー範囲 $[k$

$m - \min, km - \max$]および、ノード100nの割当キー範囲 $[kn - \min, kn - \max]$ に対して、次の関係をそれぞれ満たす。その関係とはすなわち、 $km - \min < k_{2\min} < km - \max < kn - \min < k_{2\max} < kn - \max$ という関係である。また、キー範囲3(818)は、 $[k_{3\min}, k_{3\max}]$ と表される。またキー範囲3は、ノード100nの割当

5 キー範囲 $[kn - \min, kn - \max]$ に対して、 $kn - \min < k_{3\min} < k_{3\max} \leq kn - \max$ という関係を満たす。

ノード100bのキー範囲生成部101は、受信したクエリからキー範囲群{キー範囲1, キー範囲2, キー範囲3}を生成する。ノード100bの最小割当キー $kb - \min$ と $kb - \min - 1$ の両方を含むキー範囲は、キー範囲群の中にない。したがって、キー範囲生成

10 部101によるキー範囲の分割はない。その後、ノード100bのキー範囲生成部101は、クエリ転送制御部102へ該クエリと、生成したキー範囲群{キー範囲1, キー範囲2, キー範囲3}とを渡す。

ノード100bのクエリ転送制御部102は、キー範囲生成部101からクエリとキー範囲群{キー範囲1, キー範囲2, キー範囲3}とを受け取る。ノード100bの割当キー範囲

15 $[kb - \min, kb - \max]$ とオーバーラップするキー範囲がキー範囲群には含まれていない。したがって、クエリは、クエリ処理部104に渡されない。

次にクエリ転送制御部102は、受け取ったキー範囲群に基づいてキー範囲グループとそのキー範囲グループを送信するノード100を識別できるノードアドレスとを対応付けた対を求める。ノード100bの最小割当キー保持部105は、例えば図6に示される

20 最小割当キー表806を保持していると仮定する。

また経路管理部103は、Chordのルーティングアルゴリズムにしたがって、 i のキーに対するノードアドレスを算出すると仮定する。この場合、経路管理部103は、自ノードから $2i$

(i は0以上の整数)ホップだけ離れたノードのノードアドレスを有する。例えば、ノード1

25 00bの経路管理部103は、ノード100c、ノード100e、ノード100fおよびノード100jのノードアドレスを管理する。また経路管理部103は、自身が管理するノードアドレスに対応付けてそれぞれに対応するキーの集合を記憶する。例えば、ノード100bの経路管理部103は、ノード100jのノードアドレスに対応付けて、ノード100jからノード100pまでの、およびノード100aの、それぞれの割当キー範囲に含まれるキーを記憶する。

ここで、キー範囲1について、 k_{1max} が、ノード100qの割当キー範囲に含まれている。また経路管理部103は、ノード100qの割当キー範囲に含まれるキーをノード100jのノードアドレスと対応付けて記憶している。よって、クエリ転送制御部102は、キー範囲1について経路管理部103を参照することにより、 k_{1max} に基づいて、ノード
5 100jのノードアドレスを取得する。また、クエリ転送制御部102は、ノード100jのノードアドレスに基づいて、最小割当キー保持部105からノード100jの最小割当キー k_{j-min} を取得する。

ここで、 $k_{1min} < k_{j-min}$ が成り立たない。よってクエリ転送制御部102は、キー範囲1 [k_{1min}, k_{1max}] をノード100jのノードアドレスに対応するキー範囲グループjに追加する。キー範囲2およびキー範囲3は、キー範囲1の場合と同様に、グループjに追加される。

次に、クエリ転送制御部102は、クエリとキー範囲グループj {キー範囲1, キー範囲2, キー範囲3} をノード100jのノードアドレスへ送信する。

ノード100jのクエリ転送制御部102は、ノード100bからクエリとキー範囲群 {キー範囲1, キー範囲2, キー範囲3} とを受信する。ノード100jの割当キー範囲 [k_{j-min}, k_{j-max}] とオーバーラップするキー範囲がキー範囲群には含まれていない。したがって、クエリは、クエリ処理部104に渡されない。

次にクエリ転送制御部102は、受け取ったキー範囲群に基づいてキー範囲グループとそのキー範囲グループを送信するノード100を識別できるノードアドレスとを対応付けた対を求める。ノード100jの最小割当キー保持部105は、例えば図6に示される最小割当キー表807を保持していると仮定する。

またノード100jの経路管理部103も、Chordのルーティングアルゴリズムにしたがって、一のキーに対するノードアドレスを算出すると仮定する。例えば、ノード100jの経路管理部103は、ノード100k、ノード100q、ノード100nおよびノード100bのノードアドレスを管理する。また経路管理部103は、自身が管理するノードアドレスに対応付けてそれぞれに対応するキーの集合を記憶する。例えば、ノード100jの経路管理部103は、ノード100qのノードアドレスに対応付けて、ノード100qおよびノード100mのそれぞれの割当キー範囲に含まれるキーを記憶する。

ここで、キー範囲1について、 k_{1max} が、ノード100qの割当キー範囲に含まれて

いる。また経路管理部103は、ノード100qの割当キー範囲に含まれるキーをノード100qのノードアドレスと対応付けて記憶している。よって、クエリ転送制御部102は、キー範囲1について経路管理部103を参照することにより、 k_{1max} に基づいて、ノード100qのノードアドレスを取得する。また、クエリ転送制御部102は、ノード100qのノードアドレスに基づいて、最小割当キー保持部105からノード100qの最小割当キー k_{q-min} を取得する。

ここで、 $k_{1min} < k_{q-min}$ が成り立たない。よってクエリ転送制御部102は、キー範囲1 [k_{1min} , k_{1max}] をノード100qのノードアドレスに対応するキー範囲グループqに追加する。

10 キー範囲2について、 k_{2max} が、ノード100nの割当キー範囲に含まれている。また経路管理部103は、ノード100nの割当キー範囲に含まれるキーをノード100nのノードアドレスと対応付けて記憶している。よって、クエリ転送制御部102は、キー範囲2について経路管理部103を参照することにより、 k_{2max} に基づいて、ノード100nのノードアドレスを取得する。また、クエリ転送制御部102は、ノード100nのノードアドレスに基づいて、最小割当キー保持部105からノード100nの最小割当キー k_{n-min} を取得する。

ここで、 $k_{1min} < k_{n-min}$ が成り立つ。よってクエリ転送制御部102は、キー範囲2 [k_{1min} , k_{1max}] をキー範囲2-1(816) [k_{2min} , k_{nmin-1}] とキー範囲2-2(817) [k_{n-min} , k_{2max}] とに分割する。そしてクエリ転送制御部102は、キー範囲2-2 [k_{n-min} , k_{2max}] をノード100nのノードアドレスに対応するキー範囲グループnに追加する。そして、クエリ転送制御部102は、キー範囲2-1 [k_{2min} , k_{nmin-1}] について、上述の処理と同様の処理を行う。

25 キー範囲2-1について、 k_{nmin-1} が、ノード100mの割当キー範囲に含まれている。また経路管理部103は、ノード100mの割当キー範囲に含まれるキーをノード100qのノードアドレスと対応付けて記憶している。よって、クエリ転送制御部102は、キー範囲2-1について経路管理部103を参照することにより、 k_{nmin-1} に基づいて、ノード100qのノードアドレスを取得する。また、クエリ転送制御部102は、ノード100qのノードアドレスに基づいて、最小割当キー保持部105からノード100qの最小割当キー k_{q-min} を取得する。

ここで、 $k_{2\min} < k_q - \min$ が成り立たない。よってクエリ転送制御部102は、キー範囲2-1 [$k_{2\min}, k_n - \min - 1$] をノード100qのノードアドレスに対応するキー範囲グループqに追加する。

同様に、クエリ転送制御部102は、キー範囲3をノード100nのノードアドレスに対応するキー範囲グループnに追加する。クエリ転送制御部102が、キー範囲3をキー範囲グループnに追加する過程については、上述の例と類似している。したがって記述は省略される。

次に、クエリ転送制御部102は、クエリとキー範囲グループq {キー範囲1, キー範囲2-1} をノード100qのノードアドレスへ送信する。また、クエリ転送制御部102は、クエリとキー範囲グループn {キー範囲2-2, キー範囲3} をノード100nのノードアドレスへ送信する。

ノード100qのクエリ転送制御部102は、ノード100jからクエリとキー範囲群 {キー範囲1, キー範囲2-1} を受信する。キー範囲1は、ノード100qの割当キー範囲に完全に包含される。したがってクエリ転送制御部102は、クエリをクエリ処理部104に渡し、キー範囲1をキー範囲群から削除する。そしてクエリ転送制御部102は、キー範囲群 {キー範囲2-1} からキー範囲グループとノードアドレスの対を求める。以降の、キー範囲グループとノードアドレスとの対を求める処理についての説明は、省略される。

このような処理を行うことにより、最終的に、ノード100q、100m、100nのクエリ処理部104に所望のクエリが入力されるように、適切にクエリが転送される。

本実施の形態における情報処理システム10は、クエリから形成されるキー範囲群のうち、次の転送先のノードが同じであるキー範囲群をキー範囲グループにまとめる。そして情報処理システム10は、そのキー範囲グループを該ノードへ転送する。例えば図13を参照すると、ノード100bは、キー範囲1、キー範囲2およびキー範囲3を次の転送先ノードであるノード100jへまとめて転送する。一方非特許文献1に開示された技術は、ノード100bからノード100jへ、キー範囲毎に転送を行う。したがって、非特許文献1の技術によれば、ノード100bからノード100jへの通信が3回行われることになる。一方、本実施の形態の情報処理システム10において、ノード間を転送するクエリは、同じノードを二回以上経由しない。

情報処理システム10は、ステップB607において、キー範囲 [$k_{\min}, k-1$] を [k_{\min}

n, kmax]と見なし、処理がステップB602へ戻る。この処理はキー範囲毎に繰り返される。よって、各ノードアドレスに高々一つのキー範囲グループが関連付けられる。一つの装置には高々一つのキー範囲グループとクエリとが送信される。つまり、ノード間を転送するクエリは、同じノードを二回以上経由しない。

5 このような冗長性の低減により、本実施の形態の情報処理システム10は、各ノードにおいて、クエリを転送するための処理負荷を低減させることができる。

また、本実施の形態における情報処理システム10は、クエリから生成された各キー範囲について、キー範囲に含まれる各キーから求められる転送先ノードが同じになるように、キー範囲を分割する。そして情報処理システム10は、分割されたキー範囲毎に
10 キー範囲グループにまとめる。図13を参照すると、ノード100jは、キー範囲1、キー範囲2およびキー範囲3を受信する。そしてノード100jは、キー範囲2を、転送先がノード100qであるキー範囲2-1と、転送先がノード100nであるキー範囲2-2とに分割する。そして、ノード100jのクエリ転送制御部102は、キー範囲2-2と、同じく転送先がノード100nであるキー範囲3とを、キー範囲グループnとして一つのグループとして扱う。
15 そしてクエリ転送制御部102は、そのキー範囲グループnをノード100nに転送する。一方、非特許文献1に開示された技術は、ノード100jからノード100qへのクエリの転送が二回行われる。さらに非特許文献1に開示された技術は、ノード100nに二回、同一のクエリが到達する。したがって、本実施の形態における情報処理システム10は、ノード100nに対して、キー範囲グループnが別経路にて二重に通知されることはなくなる。
20

このような冗長性の低減により、本実施の形態の情報処理システム10は、各ノードにおいて、クエリを転送するための処理負荷を低減することができる。

本実施の形態の情報処理システム10は、クエリ処理時の冗長な処理を低減できる。そして、本実施の形態の情報処理システム10は、大量のクエリを処理する場合、例え
25 ば非特許文献1に開示された技術と比べて、より高いパフォーマンスを得ることができる。

[第二の実施の形態]

図14は、本発明の第二の実施の形態におけるノード900の構成を示すブロック図である。

図14を参照すると、本発明の第二の実施の形態におけるノード900は、キー範囲生成部901と、クエリ転送制御部902と、経路管理部903と、を備える。ノード900は、データを処理するためのクエリを転送する。

キー範囲生成部901は、ある値の範囲を特定するクエリを受け取る。そしてキー範囲生成部901は、受け取ったクエリに含まれる値をそれぞれ示すキーを生成する。キーは少なくともある値を示す情報である。そしてキー範囲生成部901は、所定のキーの集合であるキー範囲を少なくとも一つ含むキー範囲群を生成する。

経路管理部903は、各装置が管理すべきキーの境界を示すキーを記憶する。

クエリ転送制御部902は、他のノード900またはキー範囲生成部901からキー範囲群を受け取る。そしてクエリ転送制御部902は、受け取ったキー範囲群に含まれるキー範囲毎に、以下の処理を行う。すなわちクエリ転送制御部902は、キー範囲に含まれる所定のキーを経路管理部903に渡す。経路管理部903はクエリ転送制御部902から受け取るキーが示す値と記憶している各キーがそれぞれ示す値との大小関係に基づいて一のキーである宛先キーを特定する。クエリ転送制御部902は、経路管理部903が特定する宛先キーを経路管理部903から受け取る。そしてクエリ転送制御部902は、対象のキー範囲に含まれるキーを、経路管理部903から受け取った宛先キーの値に基づいて識別される装置のノード識別子に関連付けられるキーの集合であるキー範囲グループに追加する。

またクエリ転送制御部902は、各キー範囲グループを、その各キー範囲グループに関連付けられているノード識別子で識別される装置へそれぞれ転送する。

キー範囲生成部901と、クエリ転送制御部902と、経路管理部903とは、例えば、プログラム(情報処理プログラム)によって動作するCPUによって実現される。キー範囲生成部901と、クエリ転送制御部902と、経路管理部903とが、同一のCPUによって実現されてもよい。この場合、プログラムは、例えば、ノード900が備えるプログラム記憶装置907(図15参照)に記憶される。CPU908は、そのプログラムを読み込み、そのプログラムにしたがって、キー範囲生成部901と、クエリ転送制御部902と、経路管理部903として動作すればよい。

また、前述のプログラムのコードを記録した記録媒体(または記憶媒体)が、ノード900に供給され、ノード900が記録媒体に格納されたプログラムのコードを読み出し実行

してもよい。すなわち、本発明は、第一の実施の形態におけるノード900が実行するためのソフトウェア(情報処理プログラム)を一時的に記憶するまたは非一時的に記憶する記録媒体910も含む。

5 経路管理部903は、例えば、ハードディスク909のような記憶装置によっても実現される。

このような構成により、本実施の形態におけるノード900は、大量のクエリを処理するときに冗長なクエリ処理を低減することができる。

10 以上、これまで述べてきた各実施の形態は、本発明の好適な実施形態であり、上記実施の形態のみに本発明の範囲を限定するものではない。各実施の形態は、本発明の要旨を逸脱しない範囲において種々の変更を施した形態での実施が可能である。

本発明の各実施の形態において、ノード100およびノード900は、情報処理装置またはその一部に備えられる。したがって、一つのノードが一つの情報処理装置から構成されてもよい。あるいは、一つのノードに、情報処理装置に含まれるメモリ、CPUおよびハードディスクなどの資源の少なくとも一部が割り当てられるように構成されてもよい。
15 つまり、複数のノードが一つの情報処理装置から構成されてもよい。

上記の各実施の形態の一部または全部は、以下の付記のようにも記載されうるが、以下には限られない。

(付記1)

20 ある値の範囲を特定するクエリを受け取り、当該クエリに含まれる値をそれぞれ示すキーを生成し、所定のキーの集合であるキー範囲を少なくとも一つ含むキー範囲群を生成するキー範囲生成部と、

各装置が管理すべきキーの境界を示すキーを記憶する経路管理部と、

キー範囲群を受け取り、当該キー範囲群に含まれるキー範囲毎に、

25 キー範囲に含まれる所定のキーが示す値と前記経路管理部が記憶する各キーがそれぞれ示す値との大小関係に基づいて一のキーである宛先キーを特定し、

当該キー範囲に含まれるキーを、前記宛先キーの値に基づいて識別される装置のノード識別子に関連付けられるキーの集合であるキー範囲グループに追加する、クエリ転送制御部と、

を備え、

前記クエリ転送制御部は、各キー範囲グループを、当該各キー範囲グループに関連付けられているノード識別子で識別される装置へそれぞれ転送する、情報処理装置。

(付記2)

付記1に記載の情報処理装置であって、

- 5 ノード識別子とキーとを対応付けて記憶する最小割当キー保持部を備え、

前記クエリ転送制御部は、受け取ったキー範囲群に含まれるキー範囲毎に、

前記宛先キーに基づいて識別される装置のノード識別子に対応付けられている一のキーを前記最小割当キー保持部から読み出し、

- 10 当該キー範囲に含まれるキーのうち前記一のキーが示す値と前記所定のキーが示す値との間の値を示すキーを、前記ノード識別子に関連付けられる一の集合であるキー範囲グループに追加する、情報処理装置。

(付記3)

付記2に記載の情報処理装置であって、

- 15 前記所定のキーは当該所定のキーを含むキー範囲に含まれるキーの中で最大値を示すキーである、情報処理装置。

(付記4)

付記3に記載の情報処理装置であって、

- 20 前記最小割当キー保持部は、ノード識別子と当該ノード識別子で識別される装置が管理すべきキーの中で最小値を示すキーである最小割当キーとを対応付けて記憶し、前記クエリ転送制御部は、受け取ったキー範囲群に含まれるキー範囲毎に、

前記キー範囲に含まれるキーのうち前記最小割当キーの値以上であり、前記所定のキーが示す値以下の値を示すキーを、前記ノード識別子に関連付けられる一の集合であるキー範囲グループに追加する、情報処理装置。

(付記5)

- 25 付記4に記載の情報処理装置であって、

前記クエリ転送制御部は、受け取ったキー範囲群に含まれるキー範囲毎に、前記キー範囲に含まれるキーであって、前記最小割当キーの値未満、または、前記所定のキーが示す値を超える値を示すキーを新たなキー範囲として特定する、情報処理装置。

(付記6)

付記1ないし5のいずれか1項に記載の情報処理装置であって、
前記経路管理部が記憶するキーは、対応する装置のノード識別子である、情報処理装置。

(付記7)

- 5 付記1ないし6のいずれか1項に記載の情報処理装置であって、
あるキーの集合を示す割当キー範囲を記憶する割当キー範囲記憶部と、
クエリを処理するクエリ処理部と、を備え、
前記クエリ転送制御部は、受け取ったキー範囲群に含まれるキー範囲毎に、
キー範囲を、前記割当キー範囲に重複するキーを含む第一のキー範囲と重複しないキーを含む第二のキー範囲とに分割し、前記第一のキー範囲を削除し、
10 前記クエリ処理部は、前記クエリ転送制御部が第一のキー範囲を削除した際にクエリを処理する、情報処理装置。

(付記8)

- 少なくとも一つの情報処理装置を備え、
15 当該情報処理装置は、
ある値の範囲を特定するクエリを受け取り、当該クエリに含まれる値をそれぞれ示すキーを生成し、所定のキーの集合であるキー範囲を少なくとも一つ含むキー範囲群を生成するキー範囲生成部と、
各情報処理装置が管理すべきキーの境界を示すキーを記憶する経路管理部と、
20 キー範囲群を受け取り、当該キー範囲群に含まれるキー範囲毎に、
キー範囲に含まれる所定のキーが示す値と前記経路管理部が記憶する各キーがそれぞれ示す値との大小関係に基づいて一のキーである宛先キーを特定し、
当該キー範囲に含まれるキーを、前記宛先キーの値に基づいて識別される情報処理装置のノード識別子に関連付けられるキーの集合であるキー範囲グループに追加
25 する、クエリ転送制御部と、
を備え、
前記クエリ転送制御部は、各キー範囲グループを、当該各キー範囲グループに関連付けられているノード識別子で識別される情報処理装置へそれぞれ転送する、情報処理システム。

(付記9)

付記8に記載の情報処理システムであって、

前記情報処理装置は、

ノード識別子とキーとを対応付けて記憶する最小割当キー保持部を備え、

5 前記クエリ転送制御部は、受け取ったキー範囲群に含まれるキー範囲毎に、

前記宛先キーに基づいて識別される装置のノード識別子に対応付けられている一のキーを前記最小割当キー保持部から読み出し、

当該キー範囲に含まれるキーのうち前記一のキーが示す値と前記所定のキーが示す値との間の値を示すキーを、前記ノード識別子に関連付けられる一の集合であるキ

10 ー範囲グループに追加する、情報処理システム。

(付記10)

付記9に記載の情報処理システムであって、前記所定のキーは当該所定のキーを含むキー範囲に含まれるキーの中で最大値を示すキーである、情報処理システム。

(付記11)

15 付記10に記載の情報処理システムであって、

前記最小割当キー保持部は、ノード識別子と当該ノード識別子で識別される装置が管理すべきキーの中で最小値を示すキーである最小割当キーとを対応付けて記憶し、

前記クエリ転送制御部は、受け取ったキー範囲群に含まれるキー範囲毎に、

20 前記キー範囲に含まれるキーのうち前記最小割当キーの値以上であり、前記所定のキーが示す値以下の値を示すキーを、前記ノード識別子に関連付けられる一の集合であるキー範囲グループに追加する、情報処理システム。

(付記12)

付記11に記載の情報処理システムであって、

25 前記クエリ転送制御部は、受け取ったキー範囲群に含まれるキー範囲毎に、前記キー範囲に含まれるキーであって、前記最小割当キーの値未満、または、前記所定のキーが示す値を超える値を示すキーを新たなキー範囲として特定する、情報処理システム。

(付記13)

付記8ないし12のいずれか1項に記載の情報処理システムであって、

前記経路管理部が記憶するキーは、対応する装置のノード識別子である、情報処理システム。

(付記14)

付記8ないし13のいずれか1項に記載の情報処理システムであって、

5 前記情報処理装置は、

あるキーの集合を示す割当キー範囲を記憶する割当キー範囲記憶部と、

クエリを処理するクエリ処理部と、を備え、

前記クエリ転送制御部は、受け取ったキー範囲群に含まれるキー範囲毎に、

キー範囲を、前記割当キー範囲に重複するキーを含む第一のキー範囲と重複しな

10 いキーを含む第二のキー範囲とに分割し、前記第一のキー範囲を削除し、

前記クエリ処理部は、前記クエリ転送制御部が第一のキー範囲を削除した際にクエリを処理する、情報処理システム。

(付記15)

ある値の範囲を特定するクエリを受け取り、当該クエリに含まれる値をそれぞれ示す

15 キーを生成し、所定のキーの集合であるキー範囲を少なくとも一つ含むキー範囲群を生成し、

各装置が管理すべきキーの境界を示すキーを経路管理部に記憶し、

キー範囲群を受け取り、当該キー範囲群に含まれるキー範囲毎に、

キー範囲に含まれる所定のキーが示す値と前記経路管理部に記憶される各キー

20 がそれぞれ示す値との大小関係に基づいて一のキーである宛先キーを特定し、

当該キー範囲に含まれるキーを、前記宛先キーの値に基づいて識別される装置のノード識別子に関連付けられるキーの集合であるキー範囲グループに追加し、

各キー範囲グループを、当該各キー範囲グループに関連付けられているノード識別子で識別される装置へそれぞれ転送する、情報処理方法。

25 (付記16)

付記15に記載の情報処理方法であって、

ノード識別子とキーとを対応付けて最小割当キー保持部に記憶し、

前記受け取ったキー範囲群に含まれるキー範囲毎に、

前記宛先キーに基づいて識別される装置のノード識別子に対応付けられている一

のキーを前記最小割当キー保持部から読み出し、

当該キー範囲に含まれるキーのうち前記一のキーが示す値と前記所定のキーが示す値との間の値を示すキーを、前記ノード識別子に関連付けられる一の集合であるキー範囲グループに追加する、情報処理方法。

5 (付記17)

付記16に記載の情報処理方法であって、

前記所定のキーは当該所定のキーを含むキー範囲に含まれるキーの中で最大値を示すキーである、情報処理方法。

(付記18)

10 付記17に記載の情報処理方法であって、

前記最小割当キー保持部に、ノード識別子と当該ノード識別子で識別される装置が管理すべきキーの中で最小値を示すキーである最小割当キーとを対応付けて記憶し、前記受け取ったキー範囲群に含まれるキー範囲毎に、

15 前記キー範囲に含まれるキーのうち前記最小割当キーの値以上であり、前記所定のキーが示す値以下の値を示すキーを、前記ノード識別子に関連付けられる一の集合であるキー範囲グループに追加する、情報処理方法。

(付記19)

付記18に記載の情報処理方法であって、

20 前記受け取ったキー範囲群に含まれるキー範囲毎に、前記キー範囲に含まれるキーであって、前記最小割当キーの値未満、または、前記所定のキーが示す値を超える値を示すキーを新たなキー範囲として特定する、情報処理方法。

(付記20)

付記15ないし19のいずれか1項に記載の情報処理方法であって、

25 前記経路管理部に記憶されるキーは、対応する装置のノード識別子である、情報処理方法。

(付記21)

付記15ないし20のいずれか1項に記載の情報処理方法であって、

あるキーの集合を示す割当キー範囲を割当キー範囲記憶部に記憶し、

前記受け取ったキー範囲群に含まれるキー範囲毎に、

キー範囲を、前記割当キー範囲に重複するキーを含む第一のキー範囲と重複しないキーを含む第二のキー範囲とに分割し、前記第一のキー範囲を削除し、

前記クエリ転送制御部が第一のキー範囲を削除した際にクエリを処理する、情報処理方法。

5 (付記22)

コンピュータに、

ある値の範囲を特定するクエリを受け取り、当該クエリに含まれる値をそれぞれ示すキーを生成し、所定のキーの集合であるキー範囲を少なくとも一つ含むキー範囲群を生成する処理と、

10 各装置が管理すべきキーの境界を示すキーを経路管理部に記憶する処理と、

キー範囲群を受け取り、当該キー範囲群に含まれるキー範囲毎に、

キー範囲に含まれる所定のキーが示す値と前記経路管理部に記憶される各キーがそれぞれ示す値との大小関係に基づいて一のキーである宛先キーを特定する処理と、

15 当該キー範囲に含まれるキーを、前記宛先キーの値に基づいて識別される装置のノード識別子に関連付けられるキーの集合であるキー範囲グループに追加する処理と、各キー範囲グループを、当該各キー範囲グループに関連付けられているノード識別子で識別される装置へそれぞれ転送する処理と、を実行させるための情報処理プログラム。

20 (付記23)

付記22に記載の情報処理プログラムであって、

前記コンピュータに、

ノード識別子とキーとを対応付けて最小割当キー保持部に記憶する処理と、

前記受け取ったキー範囲群に含まれるキー範囲毎に、

25 前記宛先キーに基づいて識別される装置のノード識別子に対応付けられている一のキーを前記最小割当キー保持部から読み出す処理と、

当該キー範囲に含まれるキーのうち前記一のキーが示す値と前記所定のキーが示す値との間の値を示すキーを、前記ノード識別子に関連付けられる一の集合であるキー範囲グループに追加する処理と、を実行させるための情報処理プログラム。

(付記24)

付記23に記載の情報処理プログラムであって、

前記所定のキーは当該所定のキーを含むキー範囲に含まれるキーの中で最大値を示すキーである、情報処理プログラム。

5 (付記25)

付記24に記載の情報処理プログラムであって、

前記コンピュータに、

前記最小割当キー保持部に、ノード識別子と当該ノード識別子で識別される装置が管理すべきキーの中で最小値を示すキーである最小割当キーとを対応付けて記憶する処理と、

10

前記受け取ったキー範囲群に含まれるキー範囲毎に、

前記キー範囲に含まれるキーのうち前記最小割当キーの値以上であり、前記所定のキーが示す値以下の値を示すキーを、前記ノード識別子に関連付けられる一の集合であるキー範囲グループに追加する処理と、を実行させるための情報処理プログラム。

15

(付記26)

付記25に記載の情報処理プログラムであって、

前記コンピュータに、

前記受け取ったキー範囲群に含まれるキー範囲毎に、前記キー範囲に含まれるキーであって、前記最小割当キーの値未満、または、前記所定のキーが示す値を超える値を示すキーを新たなキー範囲として特定する処理を実行させるための情報処理プログラム。

20

(付記27)

付記22ないし26のいずれか1項に記載の情報処理プログラムであって、

前記経路管理部に記憶されるキーは、対応する装置のノード識別子である、情報処理プログラム。

25

(付記28)

付記22ないし27のいずれか1項に記載の情報処理プログラムであって、

前記コンピュータに、

あるキーの集合を示す割当キー範囲を割当キー範囲記憶部に記憶する処理と、
前記受け取ったキー範囲群に含まれるキー範囲毎に、

- キー範囲を、前記割当キー範囲に重複するキーを含む第一のキー範囲と重複しないキーを含む第二のキー範囲とに分割し、前記第一のキー範囲を削除する処理と、
5 前記クエリ転送制御部が第一のキー範囲を削除した際にクエリを処理する処理と、を
実行させるための情報処理プログラム。

(付記29)

コンピュータに、

- ある値の範囲を特定するクエリを受け取り、当該クエリに含まれる値をそれぞれ示す
10 キーを生成し、所定のキーの集合であるキー範囲を少なくとも一つ含むキー範囲群を
生成する処理と、

各装置が管理すべきキーの境界を示すキーを経路管理部に記憶する処理と、

キー範囲群を受け取り、当該キー範囲群に含まれるキー範囲毎に、

- キー範囲に含まれる所定のキーが示す値と前記経路管理部に記憶される各キー
15 がそれぞれ示す値との大小関係に基づいて一のキーである宛先キーを特定する処理
と、

- 当該キー範囲に含まれるキーを、前記宛先キーの値に基づいて識別される装置の
ノード識別子に関連付けられるキーの集合であるキー範囲グループに追加する処理と、
各キー範囲グループを、当該各キー範囲グループに関連付けられているノード識別子
20 で識別される装置へそれぞれ転送する処理と、を実行させるための情報処理プログラ
ムを格納した記録媒体。

この出願は、2011年1月25日に出願された日本出願特願2011-012946を基
礎とする優先権を主張し、その開示の全てをここに取り込む。

産業上の利用可能性

- 25 本発明の情報処理システムは、RFID(Radio Frequency IDentification)やセ
ンサなどのデバイスから大量に発生するデータ(イベント)を、それを要求するアプリケー
ションに対して配信するPublish/Subscribe型のシステムに応用されうる。

符号の説明

10 情報処理システム

- 100、900 ノード
- 101、901 キー範囲生成部
- 102、902 クエリ転送制御部
- 104 クエリ処理部
- 5 103、903 経路管理部
- 105 最小割当キー保持部
- 106 割当キー範囲記憶部
- 107、907 プログラム記憶装置
- 108、908 CPU
- 10 109、909 ハードディスク
- 110、910 記録媒体
- 801 キー空間
- 802 キー範囲
- 803 キー
- 15 804 属性空間
- 805 クエリエリア
- 806、807、808 最小割当キー表
- 809 メッセージ
- 810 クエリ
- 20 811 キー範囲
- 812 最小キー
- 813 最大キー
- 814 キー範囲1
- 815 キー範囲2
- 25 816 キー範囲2-1
- 817 キー範囲2-2
- 818 キー範囲3
- 825 フィンガーテーブル

請求の範囲

[請求項1]

- ある値の範囲を特定するクエリを受け取り、当該クエリに含まれる値をそれぞれ示すキーを生成し、所定のキーの集合であるキー範囲を少なくとも一つ含むキー範囲群を
- 5 生成するキー範囲生成手段と、
- 各装置が管理すべきキーの境界を示すキーを記憶する経路管理手段と、
- キー範囲群を受け取り、当該キー範囲群に含まれるキー範囲毎に、
- キー範囲に含まれる所定のキーが示す値と前記経路管理手段が記憶する各キーがそれぞれ示す値との大小関係に基づいて一のキーである宛先キーを特定し、
- 10 当該キー範囲に含まれるキーを、前記宛先キーの値に基づいて識別される装置のノード識別子に関連付けられるキーの集合であるキー範囲グループに追加する、クエリ転送制御手段と、
- を備え、
- 前記クエリ転送制御手段は、各キー範囲グループを、当該各キー範囲グループに関
- 15 連付けられているノード識別子で識別される装置へそれぞれ転送する、情報処理装置。

[請求項2]

- 請求項1に記載の情報処理装置であって、
- ノード識別子とキーとを対応付けて記憶する最小割当キー保持手段を備え、
- 20 前記クエリ転送制御手段は、受け取ったキー範囲群に含まれるキー範囲毎に、
- 前記宛先キーに基づいて識別される装置のノード識別子に対応付けられている一のキーを前記最小割当キー保持手段から読み出し、
- 当該キー範囲に含まれるキーのうち前記一のキーが示す値と前記所定のキーが示す値との間の値を示すキーを、前記ノード識別子に関連付けられる一の集合であるキ
- 25 ー範囲グループに追加する、情報処理装置。

[請求項3]

- 請求項2に記載の情報処理装置であって、
- 前記所定のキーは当該所定のキーを含むキー範囲に含まれるキーの中で最大値を示すキーである、情報処理装置。

30 [請求項4]

請求項3に記載の情報処理装置であって、

前記最小割当キー保持手段は、ノード識別子と当該ノード識別子で識別される装置が管理すべきキーの中で最小値を示すキーである最小割当キーとを対応付けて記憶し、

5 前記クエリ転送制御手段は、受け取ったキー範囲群に含まれるキー範囲毎に、

前記キー範囲に含まれるキーのうち前記最小割当キーの値以上であり、前記所定のキーが示す値以下の値を示すキーを、前記ノード識別子に関連付けられる一の集合であるキー範囲グループに追加する、情報処理装置。

[請求項5]

10 請求項4に記載の情報処理装置であって、

前記クエリ転送制御手段は、受け取ったキー範囲群に含まれるキー範囲毎に、前記キー範囲に含まれるキーであって、前記最小割当キーの値未満、または、前記所定のキーが示す値を超える値を示すキーを新たなキー範囲として特定する、情報処理装置。

15 [請求項6]

請求項1ないし5のいずれか1項に記載の情報処理装置であって、

前記経路管理手段が記憶するキーは、対応する装置のノード識別子である、情報処理装置。

[請求項7]

20 請求項1ないし6のいずれか1項に記載の情報処理装置であって、

あるキーの集合を示す割当キー範囲を記憶する割当キー範囲記憶手段と、

クエリを処理するクエリ処理手段と、を備え、

前記クエリ転送制御手段は、受け取ったキー範囲群に含まれるキー範囲毎に、

25 キー範囲を、前記割当キー範囲に重複するキーを含む第一のキー範囲と重複しないキーを含む第二のキー範囲とに分割し、前記第一のキー範囲を削除し、

前記クエリ処理手段は、前記クエリ転送制御手段が第一のキー範囲を削除した際にクエリを処理する、情報処理装置。

[請求項8]

少なくとも一つの情報処理装置を備え、

30 当該情報処理装置は、

ある値の範囲を特定するクエリを受け取り、当該クエリに含まれる値をそれぞれ示すキーを生成し、所定のキーの集合であるキー範囲を少なくとも一つ含むキー範囲群を生成するキー範囲生成手段と、

各情報処理装置が管理すべきキーの境界を示すキーを記憶する経路管理手段と、

5 キー範囲群を受け取り、当該キー範囲群に含まれるキー範囲毎に、

キー範囲に含まれる所定のキーが示す値と前記経路管理手段が記憶する各キーがそれぞれ示す値との大小関係に基づいて一のキーである宛先キーを特定し、

当該キー範囲に含まれるキーを、前記宛先キーの値に基づいて識別される情報処理装置のノード識別子に関連付けられるキーの集合であるキー範囲グループに追加

10 する、クエリ転送制御手段と、

を備え、

前記クエリ転送制御手段は、各キー範囲グループを、当該各キー範囲グループに関連付けられているノード識別子で識別される情報処理装置へそれぞれ転送する、情報処理システム。

15 [請求項9]

ある値の範囲を特定するクエリを受け取り、当該クエリに含まれる値をそれぞれ示すキーを生成し、所定のキーの集合であるキー範囲を少なくとも一つ含むキー範囲群を生成し、

各装置が管理すべきキーの境界を示すキーを経路管理手段に記憶し、

20 キー範囲群を受け取り、当該キー範囲群に含まれるキー範囲毎に、

キー範囲に含まれる所定のキーが示す値と前記経路管理手段に記憶される各キーがそれぞれ示す値との大小関係に基づいて一のキーである宛先キーを特定し、

当該キー範囲に含まれるキーを、前記宛先キーの値に基づいて識別される装置のノード識別子に関連付けられるキーの集合であるキー範囲グループに追加し、

25 各キー範囲グループを、当該各キー範囲グループに関連付けられているノード識別子で識別される装置へそれぞれ転送する、情報処理方法。

[請求項10]

コンピュータに、

ある値の範囲を特定するクエリを受け取り、当該クエリに含まれる値をそれぞれ示す
30 キーを生成し、所定のキーの集合であるキー範囲を少なくとも一つ含むキー範囲群を

生成する処理と、

各装置が管理すべきキーの境界を示すキーを経路管理手段に記憶する処理と、

キー範囲群を受け取り、当該キー範囲群に含まれるキー範囲毎に、

- 5 キー範囲に含まれる所定のキーが示す値と前記経路管理手段に記憶される各キーがそれぞれ示す値との大小関係に基づいて一のキーである宛先キーを特定する処理と、

当該キー範囲に含まれるキーを、前記宛先キーの値に基づいて識別される装置のノード識別子に関連付けられるキーの集合であるキー範囲グループに追加する処理と、

- 10 各キー範囲グループを、当該各キー範囲グループに関連付けられているノード識別子で識別される装置へそれぞれ転送する処理と、を実行させるための情報処理プログラムを格納する記録媒体。

図1

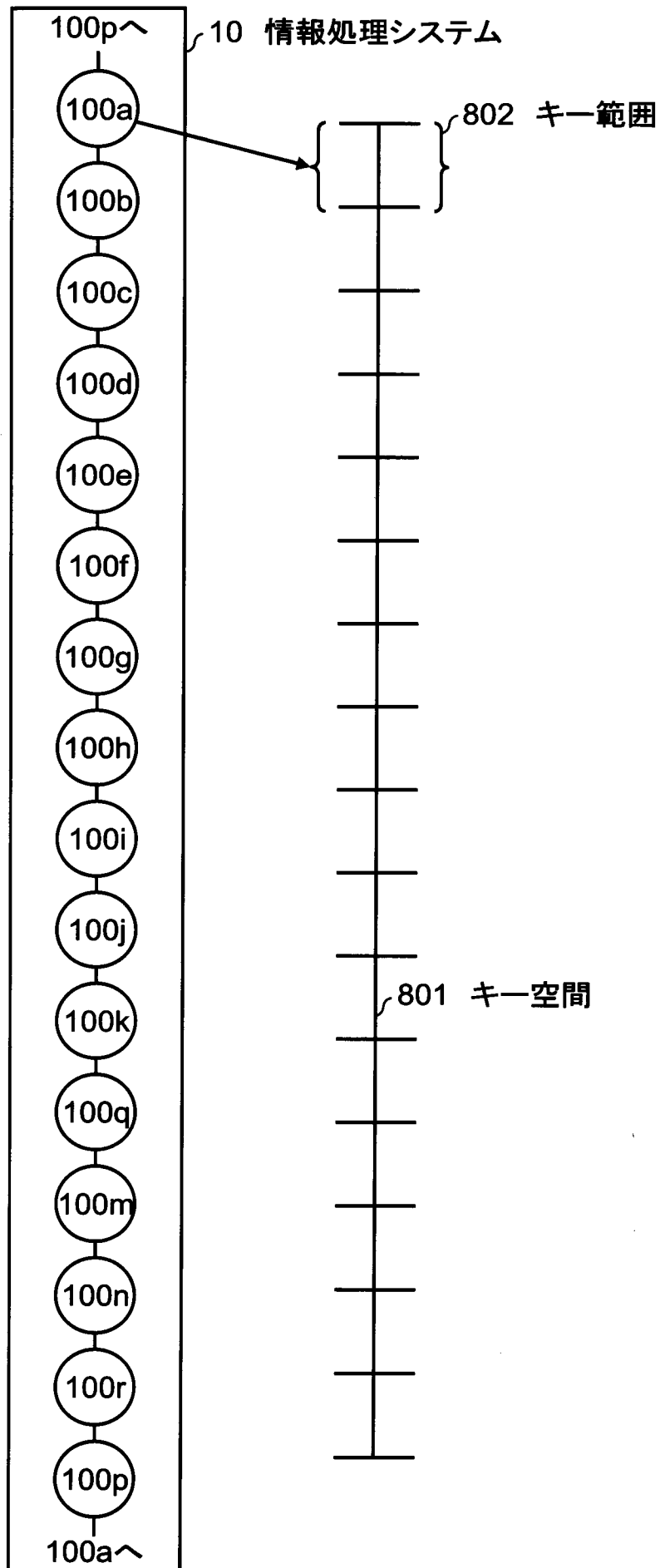


図2

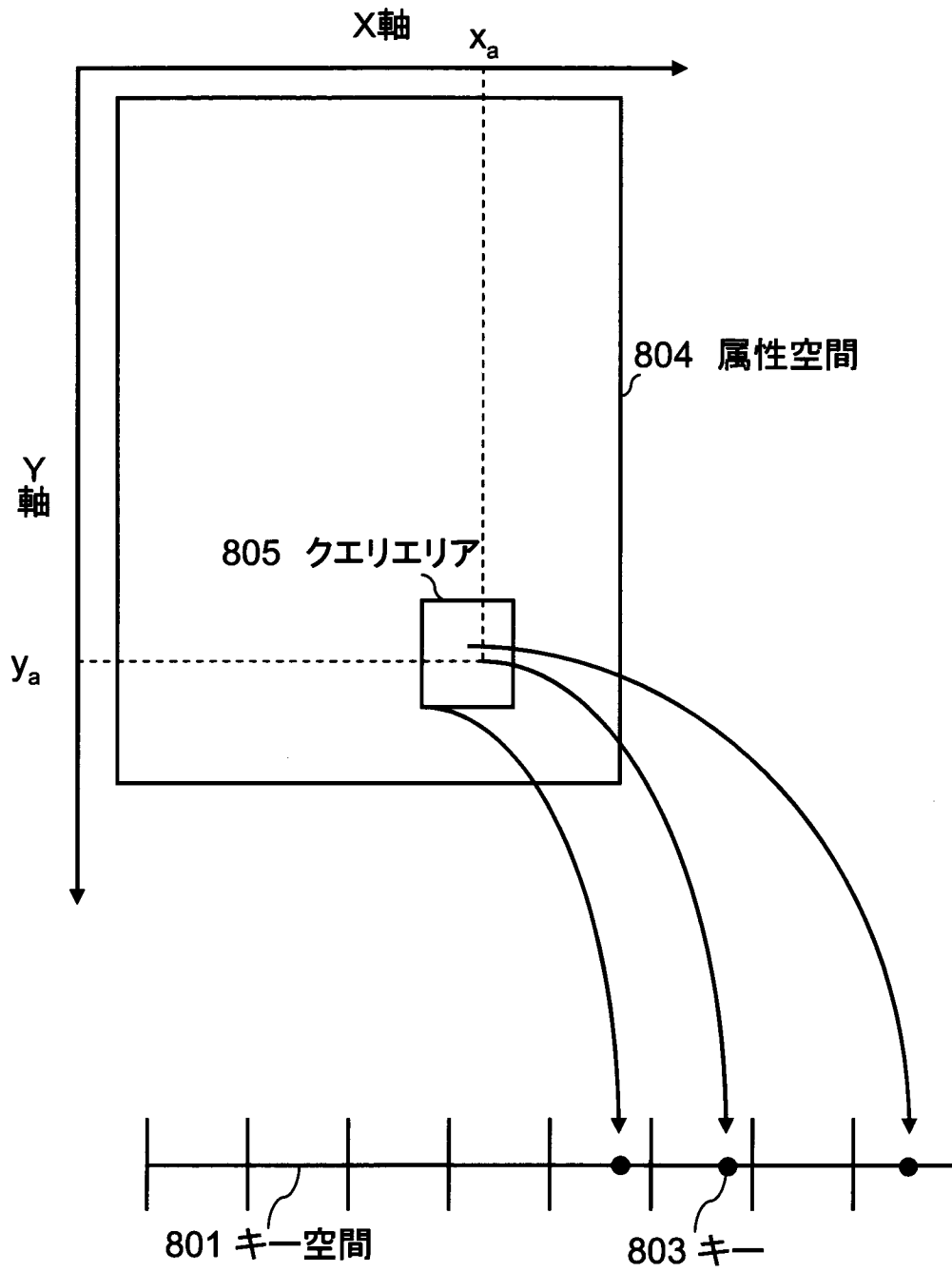


図3

X \ Y	000	001	010	011	100	101	110	111
00	00000	00001	00100	00101	10000	10001	10100	10101
01	00010	00011	00110	00111	10010	10011	10110	10111
10	01000	01001	01100	01101	11000	11001	11100	11101
11	01010	01011	01110	01111	11010	11011	11110	11111

図4

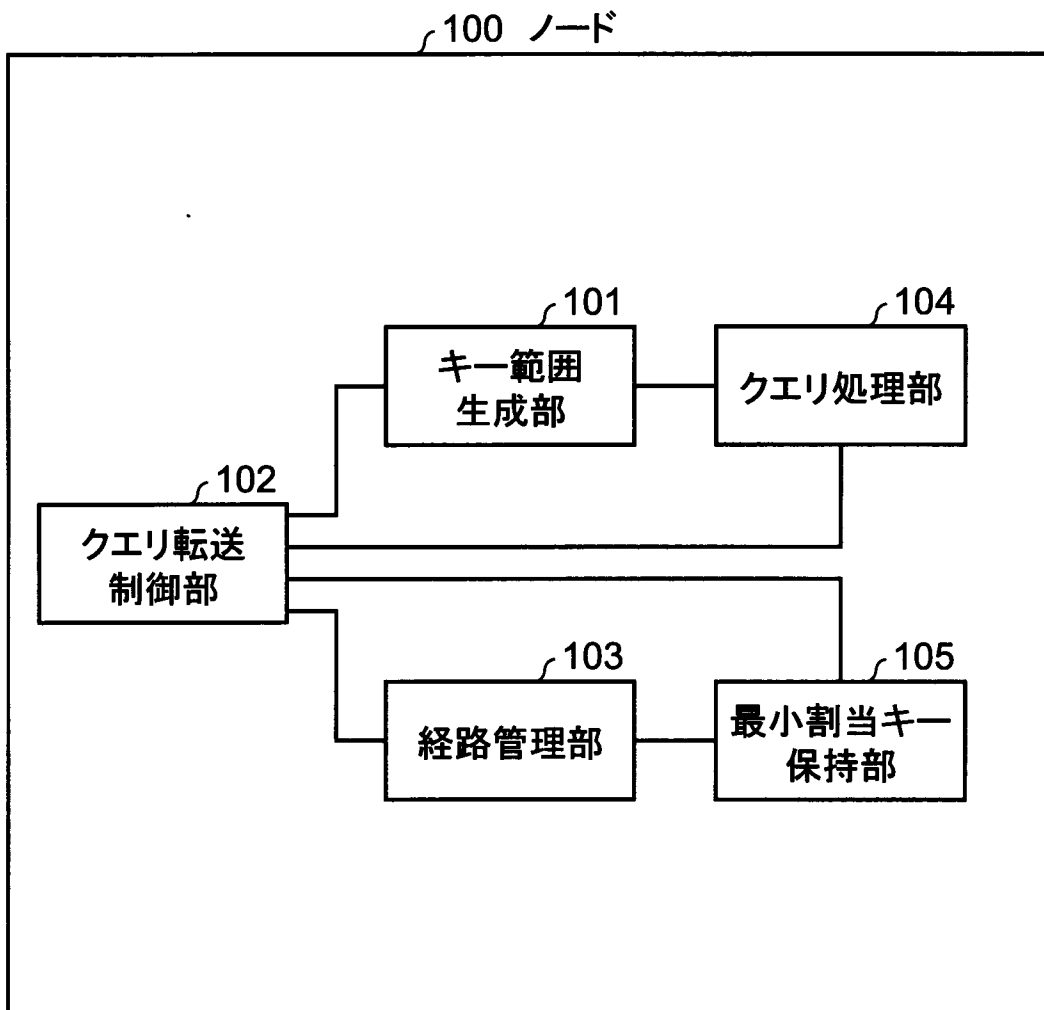


図5

825 フィンガーテーブル

start	int.	succ.
1	[1,2)	1
2	[2,4)	3
4	[4,0)	0

図6

806 最小割当キー表

ノードアドレス	最小割当キー
c	k_{c-min}
d	k_{d-min}
f	k_{f-min}
j	k_{j-min}

807 最小割当キー表

ノードアドレス	最小割当キー
k	k_{k-min}
q	k_{q-min}
n	k_{n-min}

808 最小割当キー表

ノードアドレス	最小割当キー
m	k_{m-min}
n	k_{n-min}
p	k_{p-min}

図7

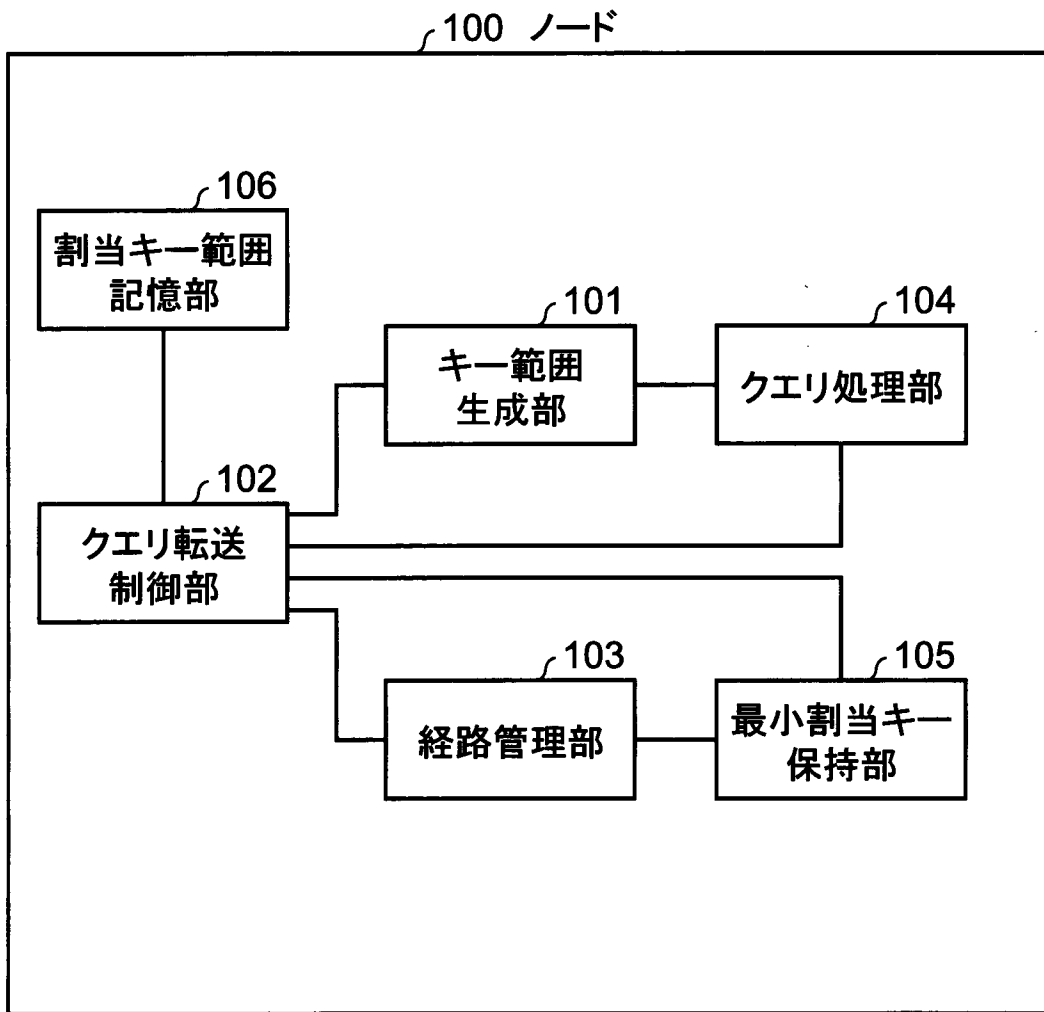


図8

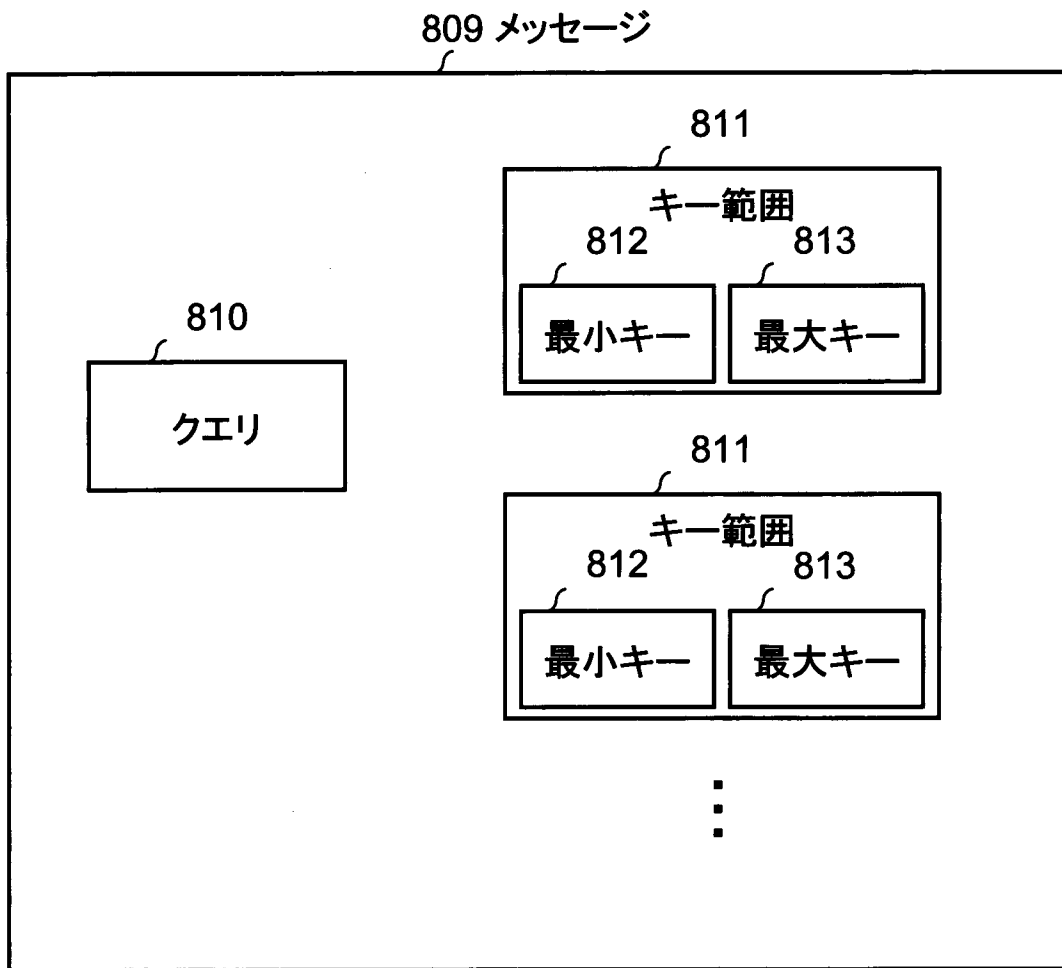


図9

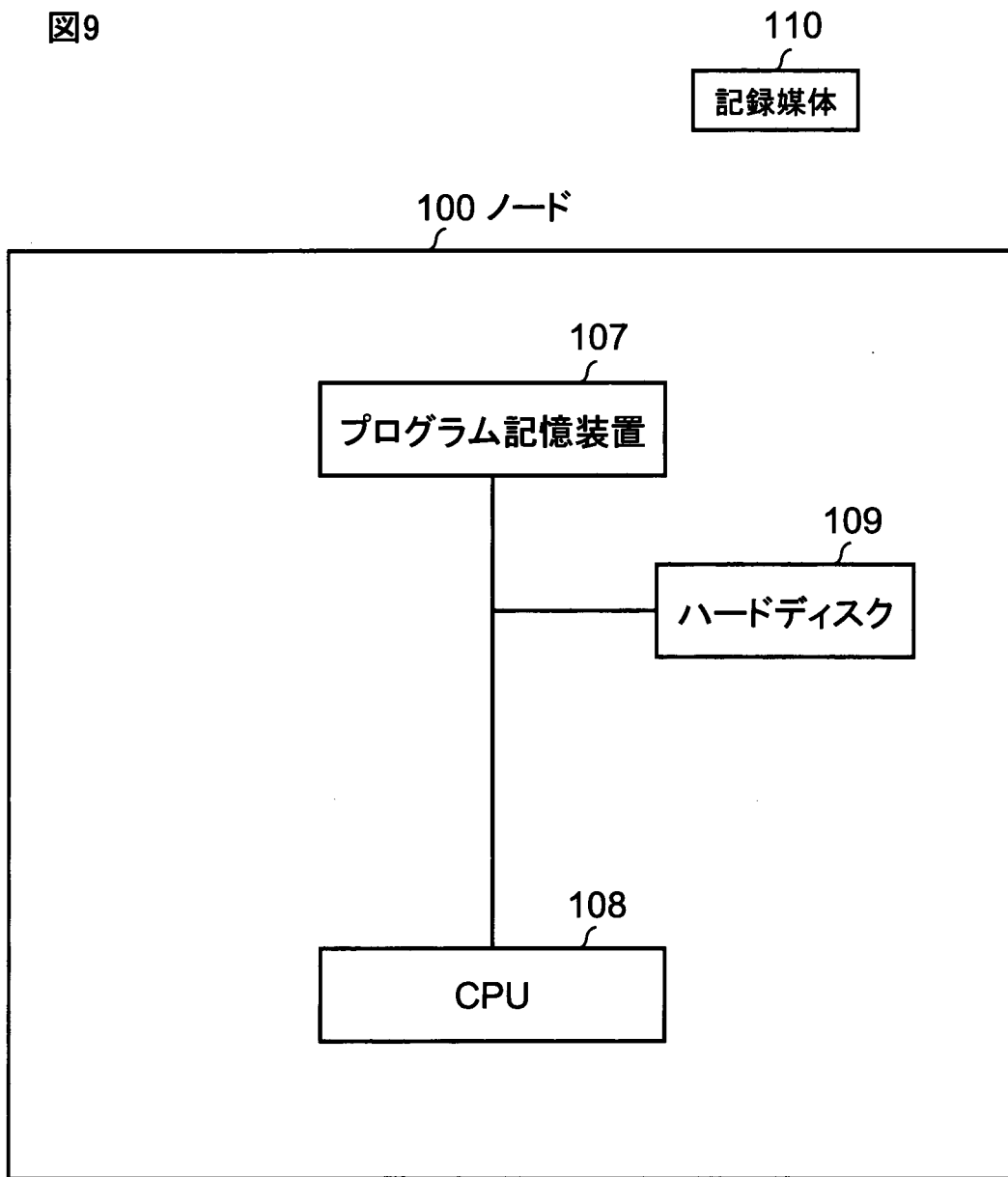


図10

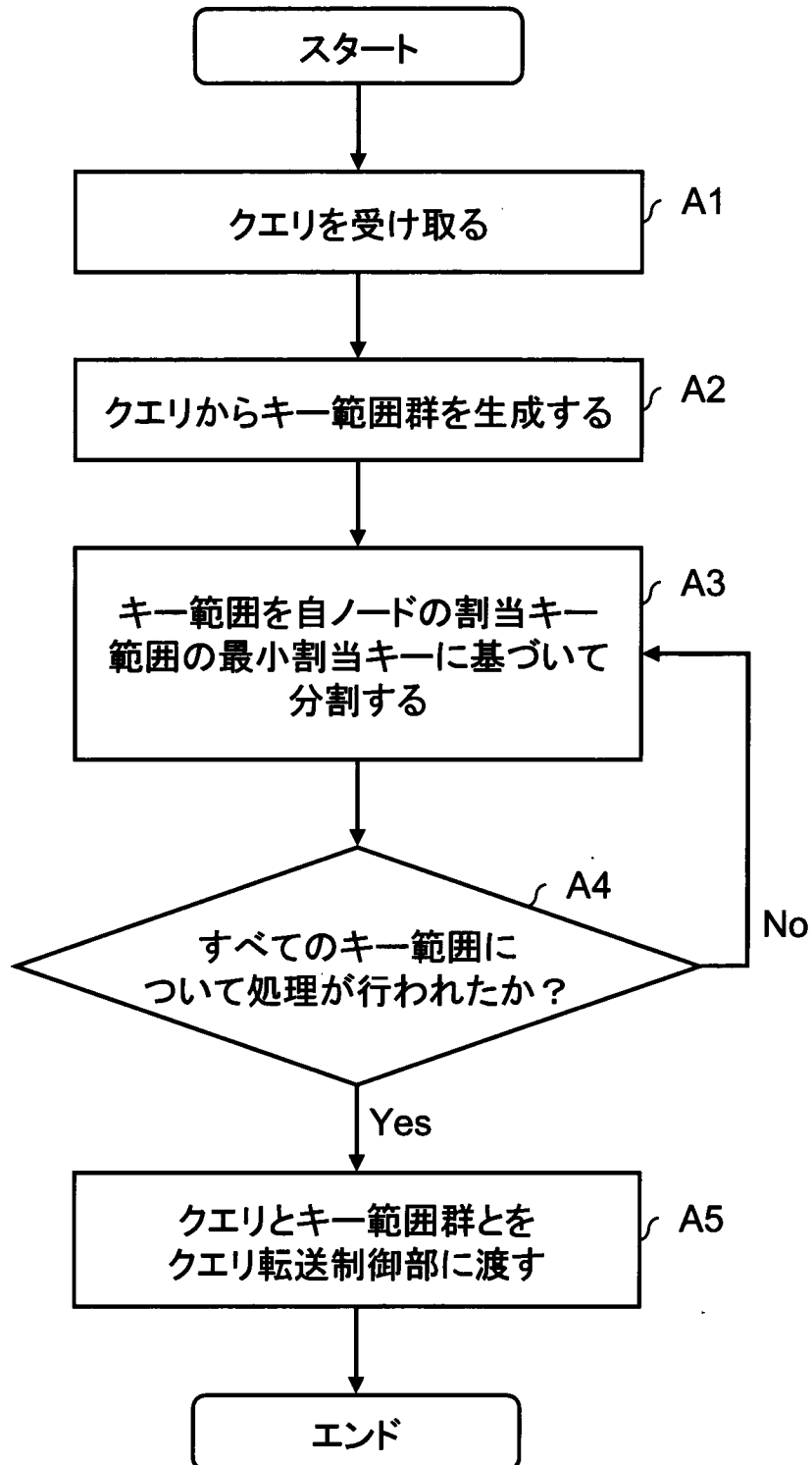


図11

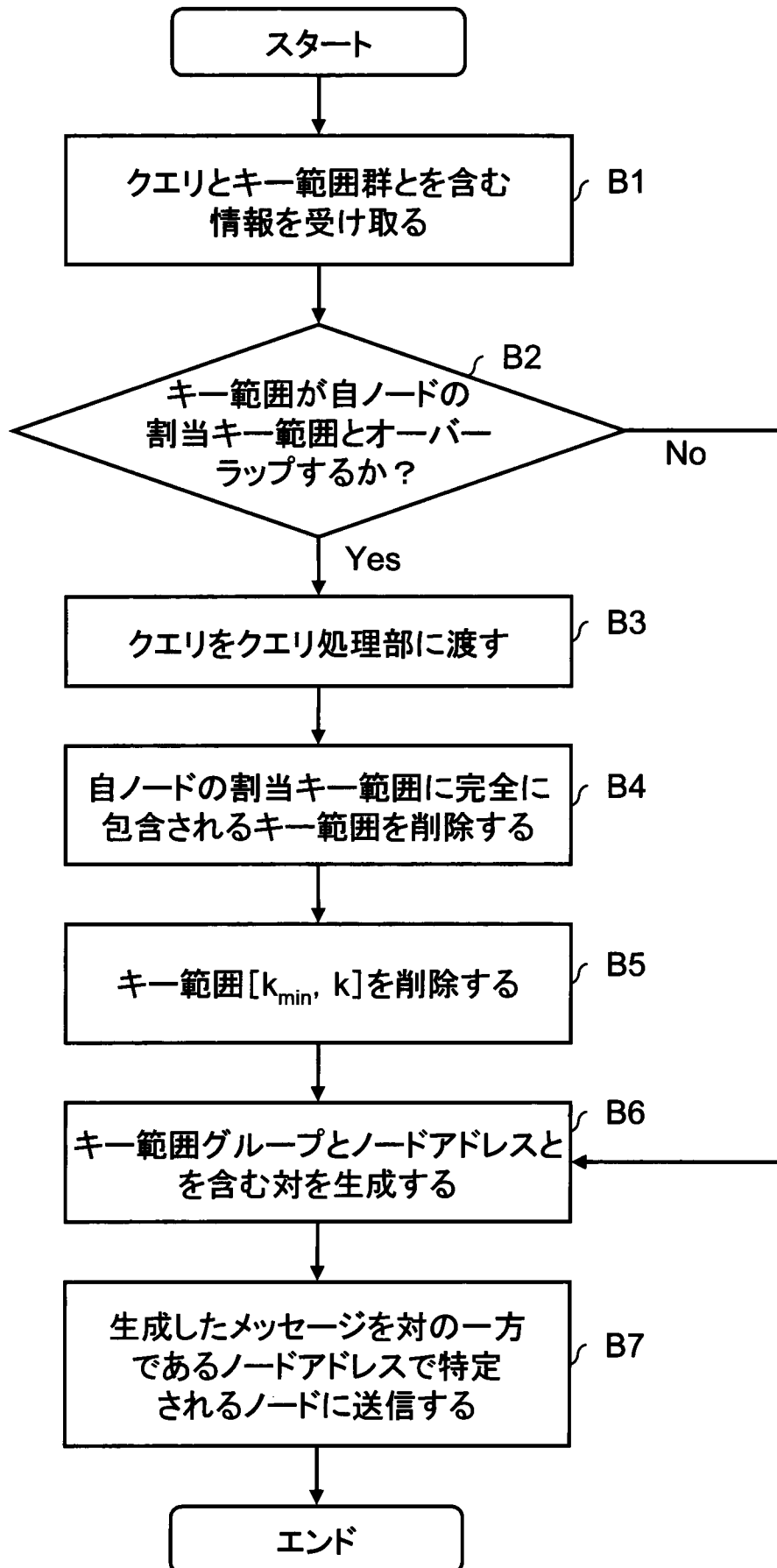


図12

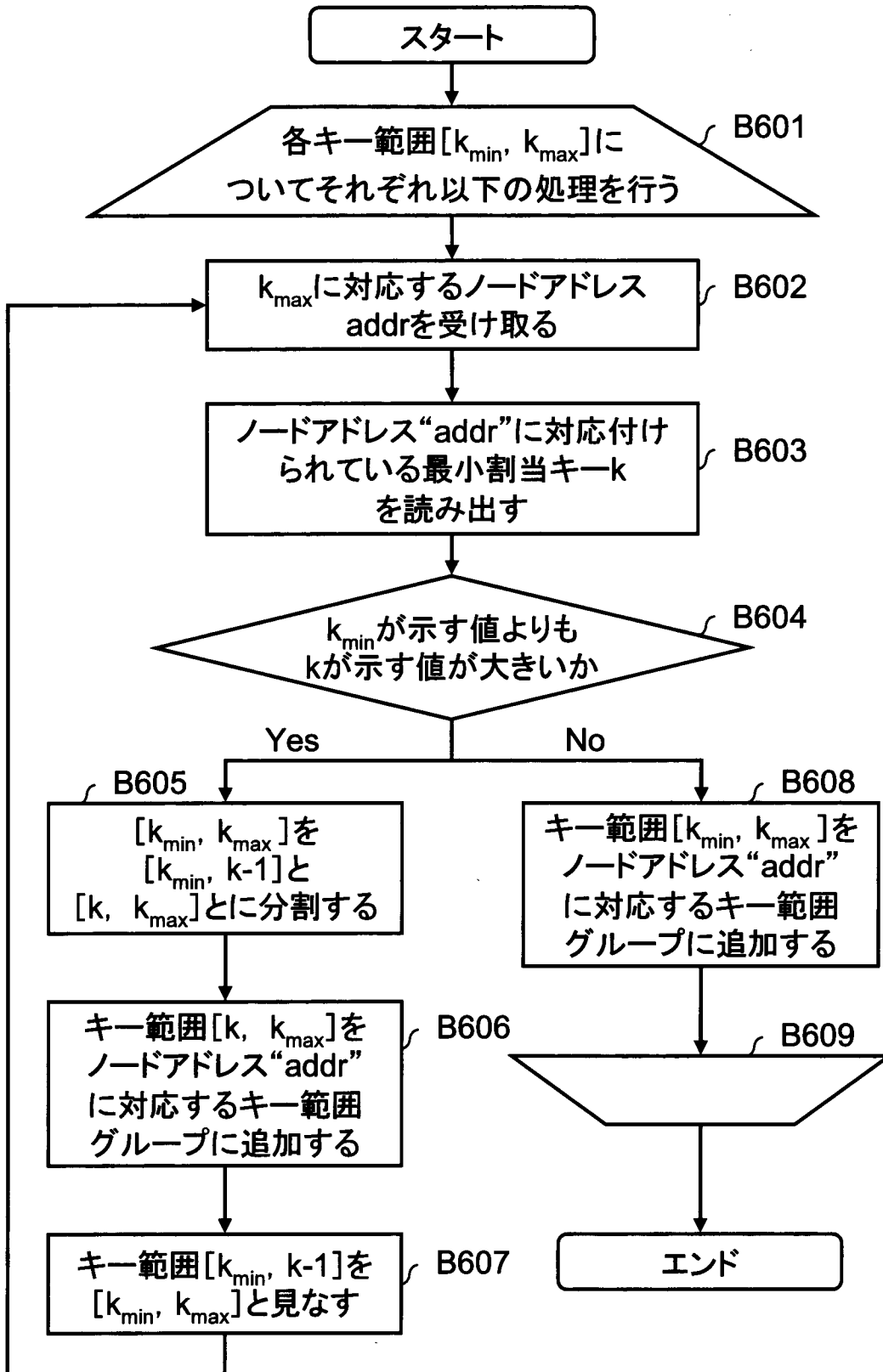


図13

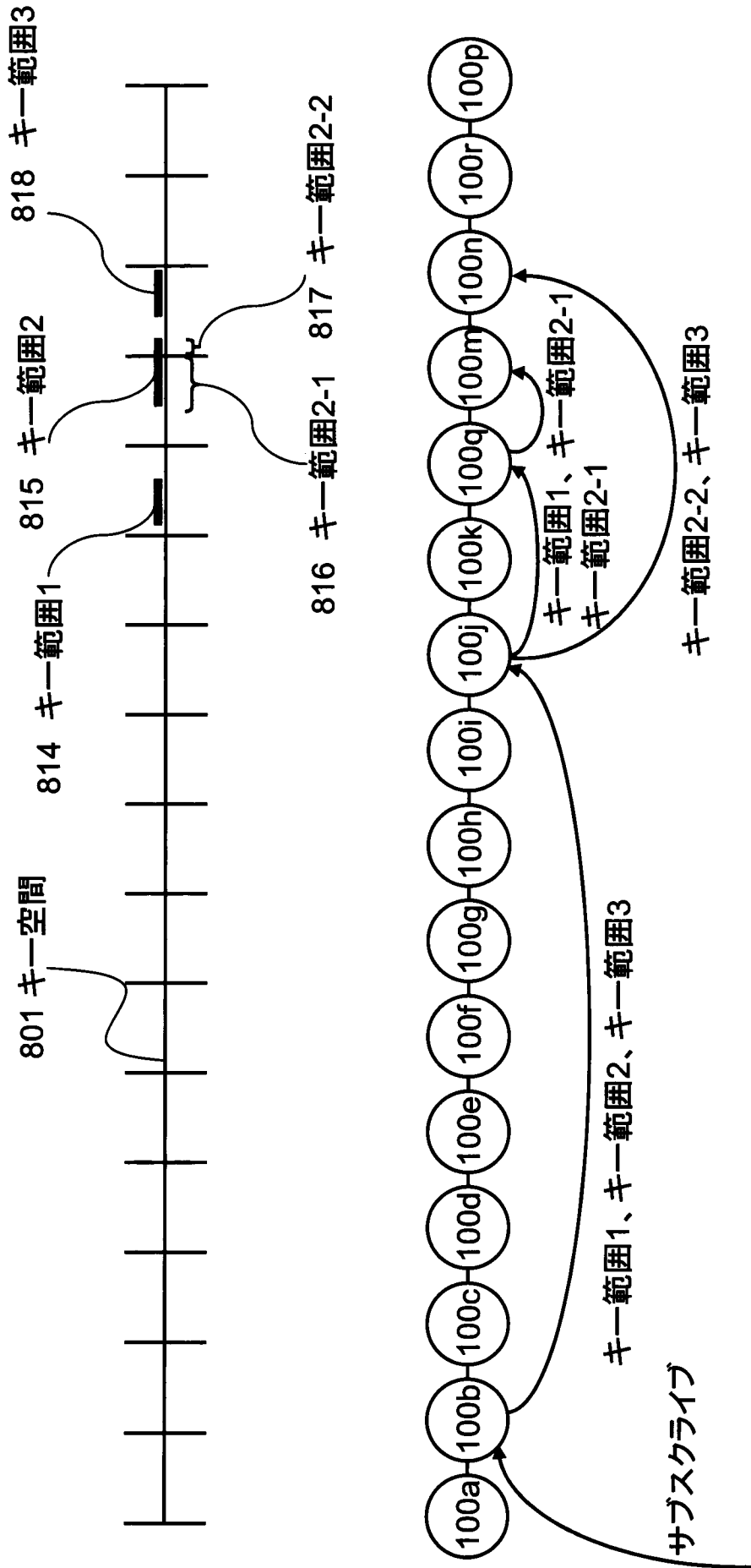


図14

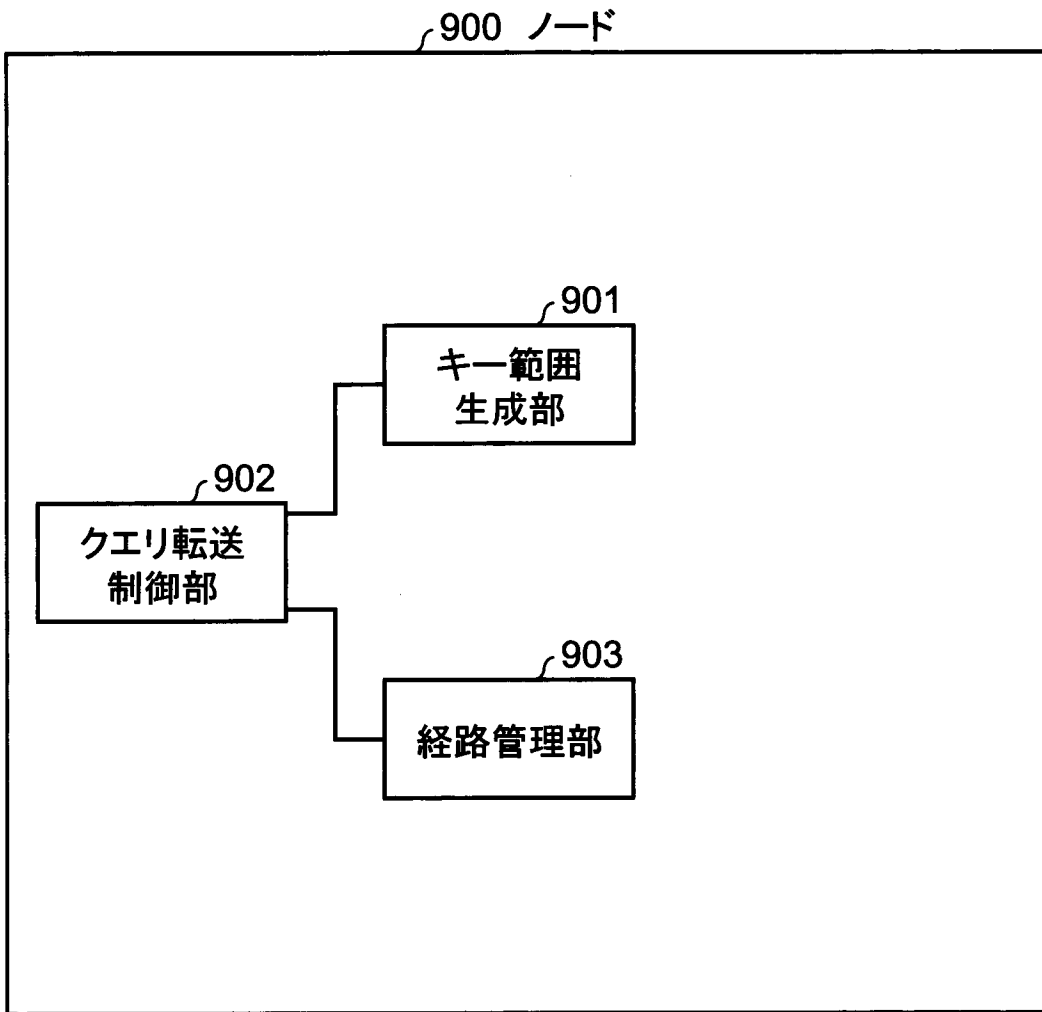
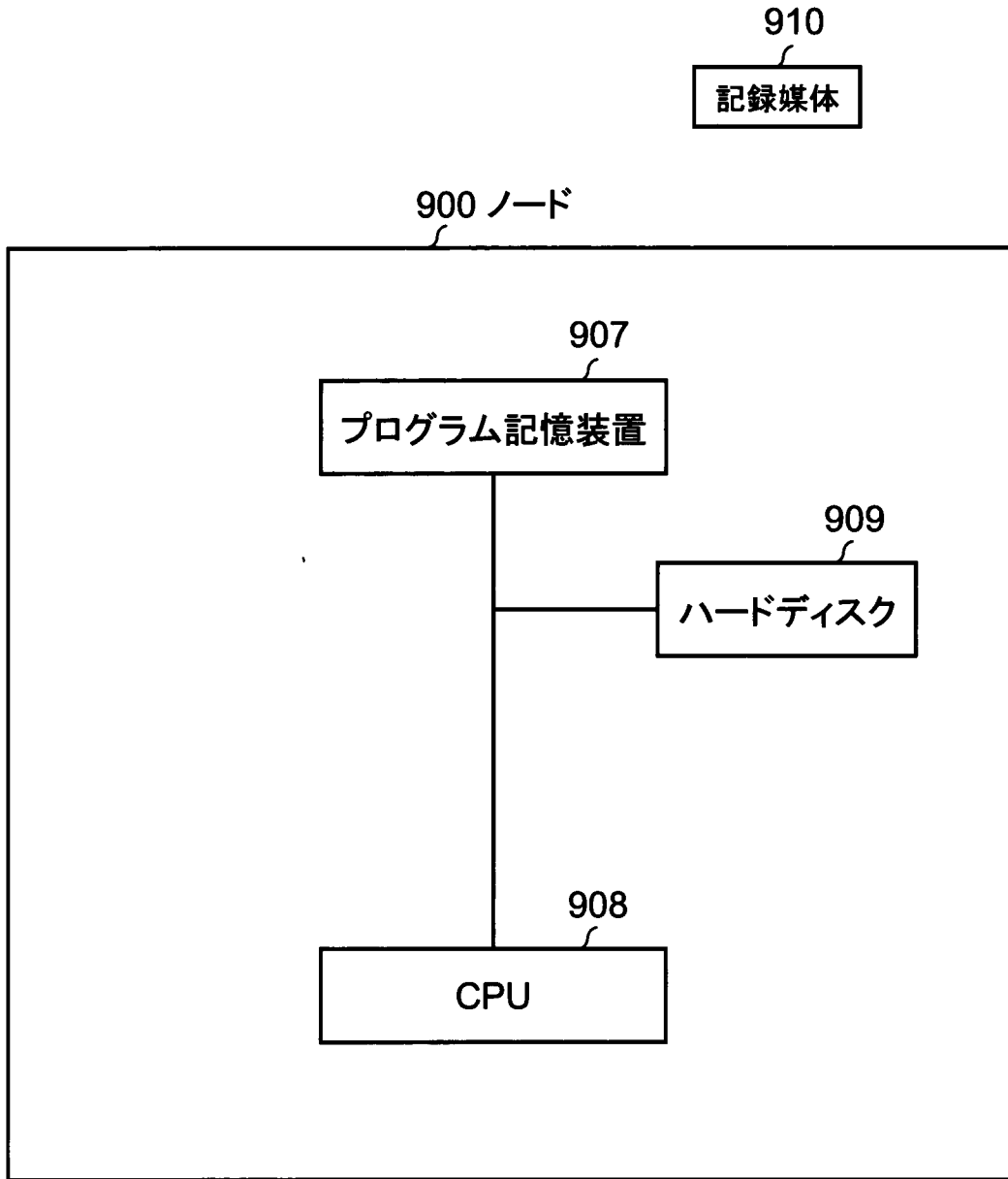


図15



INTERNATIONAL SEARCH REPORT

International application No.

PCT/JP2012/050647

A. CLASSIFICATION OF SUBJECT MATTER

G06F15/00(2006.01) i, G06F13/00(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)

G06F15/00, G06F13/00

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Jitsuyo Shinan Koho	1922-1996	Jitsuyo Shinan Toroku Koho	1996-2012
Kokai Jitsuyo Shinan Koho	1971-2012	Toroku Jitsuyo Shinan Koho	1994-2012

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 2010-271798 A (Nippon Telegraph and Telephone Corp.), 02 December 2010 (02.12.2010), paragraphs [0025] to [0064]; fig. 3 to 9 (Family: none)	1-10
A	JP 2010-538552 A (Huawei Technologies Co., Ltd.), 09 December 2010 (09.12.2010), paragraphs [0035] to [0108]; fig. 1 to 6 & EP 2187576 A1 & WO 2009/039796 A1 & CN 101399746 A & KR 10-2010-0039385 A	1-10

Further documents are listed in the continuation of Box C.

See patent family annex.

* Special categories of cited documents:

“A” document defining the general state of the art which is not considered to be of particular relevance

“E” earlier application or patent but published on or after the international filing date

“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)

“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

“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

“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

“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

“&” document member of the same patent family

Date of the actual completion of the international search
04 April, 2012 (04.04.12)

Date of mailing of the international search report
17 April, 2012 (17.04.12)

Name and mailing address of the ISA/
Japanese Patent Office

Authorized officer

Facsimile No.

Telephone No.

A. 発明の属する分野の分類 (国際特許分類 (IPC))
 Int.Cl. G06F15/00(2006.01)i, G06F13/00(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. G06F15/00, G06F13/00

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 1971-2012年
 日本国実用新案登録公報 1996-2012年
 日本国登録実用新案公報 1994-2012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 2010-271798 A (日本電信電話株式会社) 2010. 12. 02, 段落【0025】 - 【0064】, 第 3-9 図 (ファミリーなし)	1-10
A	JP 2010-538552 A (華為技術有限公司) 2010. 12. 09, 段落【0035】 - 【0108】, 第 1-6 図 & EP 2187576 A1 & WO 2009/039796 A1 & CN 101399746 A & KR 10-2010-0039385 A	1-10

☐ 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	の日の後に公表された文献
「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	「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
「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	「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
「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)	「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
「O」口頭による開示、使用、展示等に言及する文献	「&」同一パテントファミリー文献
「P」国際出願日前で、かつ優先権の主張の基礎となる出願	

国際調査を完了した日
 0 4 . 0 4 . 2 0 1 2

国際調査報告の発送日
 1 7 . 0 4 . 2 0 1 2

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 1 0 0 - 8 9 1 5
 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)	5 B	9 3 6 6
漆原 孝治		
電話番号 0 3 - 3 5 8 1 - 1 1 0 1 内線 3 5 4 5		