

(19) 世界知的所有権機関
国際事務局



(43) 国際公開日
2005年2月10日 (10.02.2005)

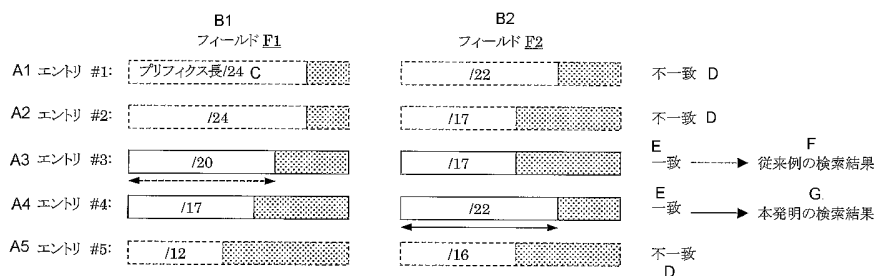
PCT

(10) 国際公開番号
WO 2005/013566 A1

- (51) 国際特許分類: H04L 12/56, G06F 17/30 (UZAWA, Yuichi) [JP/JP]; 〒211-8588 神奈川県 川崎市 中原区上小田中4丁目1番1号 富士通株式会社内 Kanagawa (JP). 大場 康弘 (OOBA, Yasuhiro) [JP/JP]; 〒211-8588 神奈川県 川崎市 中原区上小田中4丁目1番1号 富士通株式会社内 Kanagawa (JP).
- (21) 国際出願番号: PCT/JP2003/009740
- (22) 国際出願日: 2003年7月31日 (31.07.2003)
- (25) 国際出願の言語: 日本語 (74) 代理人: 茂泉 修司 (MOIZUMI, Shuji); 〒108-0074 東京都 港区 高輪3丁目25番27-809号 Tokyo (JP).
- (26) 国際公開の言語: 日本語 (81) 指定国 (国内): JP, US.
- (71) 出願人 (米国を除く全ての指定国について): 富士通株式会社 (FUJITSU LIMITED) [JP/JP]; 〒211-8588 神奈川県 川崎市 中原区上小田中4丁目1番1号 Kanagawa (JP). 添付公開書類:
— 国際調査報告書
- (72) 発明者; および 2文字コード及び他の略語については、定期発行される各PCTガゼットの巻頭に掲載されている「コードと略語のガイダンスノート」を参照。
- (75) 発明者/出願人 (米国についてのみ): 鷗澤 裕一

(54) Title: DATA SEARCH METHOD AND DEVICE

(54) 発明の名称: データ検索方法及び装置



A1...ENTRY #1
A2...ENTRY #2
A3...ENTRY #3
A4...ENTRY #4
A5...ENTRY #5
B1...FIELD F1
B2...FIELD F2
C...PREFIX LENGTH/24
D...NOT MATCHED
E...MATCHED
F...CONVENTIONAL SEARCH RESULT
G...SEARCH RESULT OF PRESENT INVENTION

(57) Abstract: Among entries divided into a plurality of fields and stored in a database, an entry in which all the non-mask bits in all the search object fields are matched with the bit data of the corresponding search key is obtained as an intermediate search result. From the prefix length of each search object field of the intermediate search result, the longest prefix length is obtained for each search object field. In the intermediate search result, an entry having the longest prefix length in the search object field of the highest priority is made a search result.

(57) 要約: データベースにおいて複数フィールドに分割して格納されたエントリの中から、全ての検索対象フィールドにおいて全ての非マスクビットと、これらに対応する検索キーのビットデータとが一致するエントリを中間検索結果とし、該中間検索結果の各検索対象フィールドのプリフィクス長の中から検索対象フィールド毎の

[続葉有]

WO 2005/013566 A1



最長プリフィクス長を求め、該中間検索結果の内、優先度が最も高い検索対象フィールドにおいて該最長プリフィクス長を有するエントリを検索結果とする。

明 細 書

データ検索方法及び装置

5 技術分野

本発明はデータ検索方法及び装置に関し、特にデータベースの中から検索キーに一致するデータを検索するデータ検索方法及び装置に関するものである。

背景技術

10 ネットワークの中継点となるルータ又はスイッチにおいては、ネットワーク内を流れるパケット（例えば IP データグラム）の宛先に応じた経路選択処理、すなわちルーティング処理を行なう必要がある。

例えば、図 6 に示したルータ 101 はネットワークの中継点となっており、このルータ 101 にはネットワーク回線 102～105 が接続されている。また、ネットワーク回線 103～105 の先にはサーバ 130、140、及び 150 がそれぞれ接続されている。

15 図示の如く、ルータ 101 がネットワーク回線 102 を介して受信するパケット 110 のヘッダには、レイヤー 3 に属する宛先 IP アドレス 111 及び送信元 IP アドレス 112 や、レイヤー 4 に属する宛先ポート番号 113 及び送信元ポート番号 114 の各情報が含まれている。

20 宛先 IP アドレス 111 からはパケット 110 を送信すべき宛先のネットワークを識別することが出来る。また、送信元 IP アドレス 112 からはパケット 110 の送信者を識別することが出来るため、識別された送信者との間で締結されているネットワークの利用に関する契約の内容（回線容量等）を把握することも出来る。さらに、宛先ポート番号 113 及び送信元ポート番号 114 からはアプリケーションの種類などを判別することが出来る。

25 従って、ルータ 101 はこれらの情報に基づき、パケット 110 を適切なネットワーク回線 103, 104, 105 に出力するように、レイヤー 3 又はレイヤー 4 のルーティング処理を行なう。

例えば、ネットワーク回線 103 及び 104 の回線容量（帯域）について、それぞ

れの実線の太さで表される如く、ネットワーク回線 103 が大容量であり、ネットワーク回線 104 が小容量であるような場合、パケット 110 の送信者との回線容量に関する契約内容に応じてルーティング処理を行うことが可能となる。

また、アプリケーション毎にサーバ 130、140、及び 150 をそれぞれ対応させておくことにより、レイヤー3 のルーティング処理で負荷分散を行うこともできる。

ルータ 101 においてルーティング処理を適切に行えば、柔軟なネットワークが構築可能となる。

一般に、上記のルーティング処理はソフトウェア及びハードウェアによる宛先のデータ検索処理によって実現されている。特に、高速・大容量のルーティング処理を行なうためのデータ検索装置としては、従来より連想メモリ (Content Addressable/Associative Memory : CAM) デバイスが主として用いられている。

なお、CAM デバイスは、上記のルータ 101 の様にレイヤー3 又はレイヤー4 でのルーティング処理を行うルータだけでなく、レイヤー2 スイッチ (図示せず) の様に MAC アドレスによるスイッチングにも用いられている。

CAM デバイスを用いたルーティング処理の一例として、IP アドレスの検索処理について図 7 を参照して以下に説明する。

図 7 は、例えば図 6 のルータ 101 内において、ルーティング処理に関連した部分の構成例を示したものである。

CAM デバイス 200 は、検索対象のエントリ群を格納したデータベース 203 を有するものであり、ネットワークプロセッサ 210 は、ネットワークの処理に特化したプロセッサである。また、コンテキスト RAM (Context-RAM 又は Associated-RAM) 220 は、CAM デバイス 200 の各エントリに対応した処理内容等を格納したメモリであり、通常 SRAM、DRAM などのメモリデバイスで構成されるものである。

CAM デバイス 200、ネットワークプロセッサ 210、及びコンテキスト RAM 220 は、入力インタフェース 211、出力バス 201 及び 221 でそれぞれ相互に接続されている。

動作において、ネットワークプロセッサ 210 は受信したパケット 110 からヘッダ情報を抽出して、検索したい IP アドレスまたはポート番号等を検索キー 212 として入力インタフェース 211 を介して CAM デバイス 200 に入力する。CAM デバイ

ス 200 内部のデータベース 203 には、予めネットワーク管理者によって入力された IP アドレス情報がエン트리群として収容されており、CAM デバイス 200 は、検索キー 212 とデータベース 203 の各エン트리とを比較する。

ここで、CAM デバイス 200 による比較動作は各エントリの各ビットに対して同時に
5 行われる。検索した結果、適切なエントリが有った（ヒットした）場合、その中で“最も”適切なエントリの物理アドレス（同図の例ではアドレス[2]）をヒットアドレス 202 として出力バス 201 を介してコンテキスト RAM220 に与える。

コンテキスト RAM220 は、例えば、同図のヒットアドレス 202 として得られたアドレス[2]に対応して、対処すべき処理内容や出力ポート（配布先）のデータを出力
10 バス 221 を介してネットワークプロセッサ 210 に検索結果 222 として与える。

ネットワークプロセッサ 210 は、検索結果 222 に応じてパケット 110 を処理したり、再度検索を行ったりする。

上記の CAM デバイス 200 内における従来の IP アドレス検索処理について、図 8 を参照してより詳細に説明する。

15 図 8 は、図 7 の CAM デバイス 200 において IP アドレス検索処理に関連した部分をより詳細に示したものである。従って、図 7 のデータベース 203 に予めネットワーク管理者によって入力されたエン트리群の内、説明の便宜上、上位 4 つのエントリ #1～#4 のみが示されている。また、各エントリ #1～#4 毎に一致線 321～324 が設けられており、それぞれプライオリティ・エンコーダ 330 に接続されている。

20 各エントリ #1～#4 は、IP アドレスとサブネットマスクが一对となってデータベース 203 に収容されたものである。同図の各エントリ #1～#4 の上段にはそれぞれ、IP アドレスとして“A1”、“A2”、“A3”、及び“A4”が示されており、下段にはそれぞれのサブネットマスクが連続したビット“1”及び“0”で示されている。この場合、連続したビット“1”はプリフィクス部分を示しており、連続したビット
25 “0”はマスク部分を示している。

エントリ #1～#4 は、図示の如くサブネットマスクのプリフィクス部分（ビット“1”）の長さ、すなわちプリフィクス長（prefix length）が長い順に並べられている。また、各エントリ #1～#4 の右側に示された[1]～[4]は、それぞれの物理アドレスを示している。

IP アドレスは右側の下位ビットほどネットワークの末端を現し、地図上の住所と同様に、より詳細な内容である下位ビットまで一致している検索結果ほど適切であるといえることができる。すなわち、IP アドレスの検索結果としては、プリフィクス長が最も長いアドレスが最も適切であり、このような検索結果はロングスト・プリフィクス・マッチ (Longest Prefix Match) と呼ばれるものである。エントリをプリフィクス長に従って並べて収容するのはこのためである。

運用中の CAM デバイス 200 の動作においては、ネットワークプロセッサ 210 などから検索キー 212 として例えば IP アドレス “A-KEY” が入力される。CAM デバイス 200 は全エントリに対して同時に各ビットについて比較を行なう。但し、サブネットマスクが “0” のビットに関しては比較を行なわない (ドントケア条件)。

比較したビットが 1 つでも「不一致」である場合、一致線 321~324 に不一致をあらわす信号が出力されて、不一致エントリとして識別される。同図の例では、不一致信号 331 及び 333 がそれぞれ一致線 321 及び 323 に出力されているため、エントリ #1 及び #3 が不一致エントリとして識別されている。

比較したビットが全て「一致」している場合は、エントリ #2 及び #4 の様に一致エントリとして識別される。

このように複数ヒットした場合は、従来よりプライオリティエンコーダ 330 は、プリフィクス長の長い方、すなわちエントリ #2 の物理アドレス [2] をヒットアドレス 202 として出力する。

このように、検索キー 212 のデータビットと、エントリ #1~#4 のデータビットとを比較する場合、マスクビットによるマスク処理を行う CAM デバイスは 3 値のデータを扱うことからターナリ CAM デバイスと呼ばれ、一方、検索キー 212 のデータビットとエントリ #1~#4 のデータビットとの比較のみを行ない、マスク処理を行なわない CAM デバイスは 2 値のデータを扱うことからバイナリ CAM デバイスと呼ばれている。

図 9 は、ターナリ CAM デバイスにおける 1 ビット分に相当するターナリ CAM セル 600 の機能及び検索結果の真理値表 610 を示したものである。

同図 (1) において、ターナリ CAM セル 600 は、データビット 601 及びマスクビット 602 を保持しており、さらに、入力された検索キー 607 とデータビット 601

とを比較する比較処理部 603 と、この比較処理部 603 が出力する比較結果に対しマスクビット 602 に応じてマスク処理を施すマスク処理部 604 を有している。

データベースの 1 つのエントリは共通の一致線 606 に接続された複数のターナリ CAM セル 600 によって構成されている。

- 5 動作において、ターナリ CAM セル 600 は、比較処理部 603 によって検索キー 607 とデータビット 601 とが比較された結果に対して、マスク処理部 604 によってマスク処理を施した結果を、最終的な検索結果 605 として一致線 606 に出力する。

- すなわち、比較処理部 603 は、検索キー 607 とデータビット 601 の値が異なれば“0”、等しければ“1”を出力し、マスク処理部 604 は、マスクビット 602 の値が“1”のときは比較処理部 603 の出力をそのまま検索結果 605 として出力し、マスクビット 602 の値が“0”のときは比較処理部 603 の出力に関わらず“1”を検索結果 605 として強制的に出力する。同図（2）は、このような検索結果を真理値表の形で示したものである。
- 10

なお、ここで動作が相等しければ論理反転させても良い。

- 15 データビット 601 及びマスクビット 602 を保持する機能としては、例えばフリップフロップ、キャパシタ、ヒステリシスを有する高誘電膜あるいは磁性膜を用いられる。

- 上記の図 7 及び図 8 に示した CAM デバイス 200 がターナリ CAM デバイスである場合、各エントリ #1～#4 は、エントリのビット数に応じたターナリ CAM セル 600 によって構成されることになる。この場合、各一致線 321～324 は一致線 606 に対応している。
- 20

- 上記の図 8 においては、特定の項目、すなわち、IP アドレスのみを検索する場合を例にとって説明したため、各エントリ #1～#4 には IP アドレスのみが收容されているように示したが、実際の CAM デバイス 200 内のデータベース 203 には、図 10（1）に示す如く、エントリ #1～#4 には例えばフィールド F1～F5 のように複数フィールドに分割されたデータが格納されている。
- 25

すなわち、エントリ #1～#4 のフィールド F1 には、それぞれ送信元 IP アドレスとして“SA1”～“SA4”が格納されており、同様に、フィールド F2 には宛先 IP アドレスとして“DA1”～“DA4”が格納されている。さらに、フィールド F3 には、

それぞれ送信元ポート番号として“SP1”～“SP4”が格納されており、フィールド F4 には宛先ポート番号として“DP1”～“DP4”が格納されている。

また、フィールド F5 には、その他の情報として、例えば、エントリの有効/無効を示すタグのように、外部から通知される管理用の情報が格納されている。

5 なお、フィールド F1 及び F2 は各エントリ #1～#4 について上下 2 段で示されているが、これは、フィールド F1 及び F2 には通常サブネットマスクが適用されるため、それぞれのプリフィクス部分 431, 432 及びマスク部分 441, 442 を下段に示したものである。フィールド F3～F5 については、通常はサブネットマスクを適用しないため 1 段で示されている。

10 同図 (2) は、同図 (1) のフィールド F1～F5 の内、特定のフィールドのみを検索対象に指定したい場合に用いられるグローバルマスクパターン 450 を示したものである。

 このようなグローバルマスクパターン 450 を用いれば、同図の例では、マスク部分 451 及び 453 が検索対象外となり、検索対象部分 452 に対応して、宛先 IP
15 アドレスを示すフィールド F2 のみが検索対象になる。

 従来より提案されている CAM デバイスとしては、連想メモリアレイを複数のフィールドから構成し、検索対象となるフィールドを指示することにより、指示されたフィールドには検索対象データを、その他のフィールドにマスク信号を与えて検索を行なうものがある (例えば、特許文献 1 参照)。

20 また、IP アドレスとプリフィクス長との組み合わせにより表現された経路情報とともに別途当該プリフィクス長の情報を格納したもの (例えば、特許文献 2 参照) や、データ選択の優先順序を外部から任意に設定可能としたものもある (例えば、特許文献 3 参照)。

 さらには、低消費電力によるデータ検索を可能とし、より集積度の高いメモリ
25 幅拡張機能を有するもの (例えば、特許文献 4 参照) や、可変幅のデータを検索するための可変幅の CAM デバイスもある (例えば、特許文献 5 参照)。

< 特許文献 1 >

特開昭 58-200494 号公報 (特許請求の範囲、図 1)

< 特許文献 2 >

特開 2002-305539 号公報 (要約、図 1)

<特許文献 3>

特開平 11-102589 号公報 (要約、図 1)

<特許文献 4>

5 特開平 5-189978 号公報 (要約、図 1)

<特許文献 5>

特開 2001-160292 号公報 (要約、図 2A)

10 CAM デバイスを用いた検索を数十ビットの幅でしか行えない状況では、例えば、フロー識別などの処理を行なうために複数回 CAM 検索を繰り返す必要があったが、近年では、多くの CAM デバイスが数百ビットを超える多ビット幅検索に対応しており、例えば、IPv4 によるフロー識別を例にとると、宛先 IP アドレス (32 ビット)、送信元 IP アドレス (32 ビット)、宛先ポート番号 (16 ビット)、及び送信元ポート番号 (16 ビット) を 1 つのエントリとして合計 96 ビット以上を一度に検索することが可能になっている。

15 しかしながら、従来の技術においては、上述の如く、複数のエントリがヒットした場合にプライオリティエンコーダ 330 が出力する最終検索結果は、エントリの格納順によって決まってしまう。

別の最終検索結果を得るためには、エントリの格納順を変更する必要があるが、これには、データベース 203 の書き換えが必要になる。このようなデータベース
20 203 のメンテナンスを行なうためには、ルータ 101 のルーティング処理を停止する必要があり、ネットワークのダウン時間を増大させることになる。

従って、ネットワークの常時運用を実現するためには、エントリの格納順は容易に変更することが出来ない。

一方、近年における仮想閉域網 (Virtual Private Network : VPN) の利用の拡大に伴い、顧客毎の柔軟な対応がネットワークの運用において重要になって来
25 ている。例えば、ネットワークの障害時において、特定の顧客の回線を優先させる必要がある場合、当初の格納順が低位のエントリを優先させることが望まれる。

また、運用中のサーバ等の負荷が集中しないように、適切な負荷分散を行なう場合にも、当初の格納順が低位のエントリを優先させることが望まれる。

従って、本発明は、データベースの中から検索キーに一致するデータを検索するデータ検索方法及び装置に関し、エントリの格納順に拘束されることなく、所望の優先度に基づく検索結果を得ることを目的とする。

5 発明の開示

上記の目的を達成するため、本発明に係るデータ検索方法は、データベースにおいて複数フィールドに分割して格納されたエントリの中から、全ての検索対象フィールドにおいて全ての非マスクビットと、これらに対応する検索キーのデータビットとが一致するエントリを中間検索結果とする第1ステップと、該中間検索結果の各検索対象フィールドのプリフィクス長の中から該検索対象フィールド毎の最長プリフィクス長を求める第2ステップと、該中間検索結果の内、優先度が最も高い検索対象フィールドにおいて該最長プリフィクス長を有するエントリを検索結果とする第3ステップと、を有することを特徴としている。

本発明の原理を図1を用いて以下に説明する。同図は、データベースに格納されたエントリ#1～#5の各々について、例えば複数のフィールドF1及びF2を検索対象とする検索キー（図示せず）が入力された際の一致／不一致の状態を模式的に示したものである。

各エントリ#1～#5の点線又は実線で囲まれた各フィールドF1及びF2内の白地部分において、例えば“/24”のように示された数値は、プリフィクス長（サブネットマスクされていないプリフィクス部分の長さ）を示すビット数である。また、白地部分に隣接した網掛部分は、マスク部分（サブネットマスクされた部分）を示したものである。

すなわち、同図のエントリ#1のフィールドF1のプリフィクス長は24ビットであり、フィールドF2のプリフィクス長は22ビットであることを示している。同様に、エントリ#2～#5のフィールドF1のプリフィクス長はそれぞれ、24ビット、20ビット、17ビット、12ビットであり、フィールドF2のプリフィクス長はそれぞれ、17ビット、17ビット、22ビット、16ビットである。

また、点線で示したエントリ#1, #2, 及び#5は、プリフィクス部分の少なくとも1ビットが対応する検索キーのデータビットに一致しないために、「不一致」であ

ると判定されている。

他方、実線で示したエントリ 503 及び 504 は、プリフィクス部分の全ビット（非マスクビット）が対応する検索キーのビットデータに「一致」しており、中間検索結果になっている。

5 ここで、従来の連想メモリの検索方法では、エントリ#1～#5 の格納順を優先順位として検索結果を出力するため、上記の如く抽出されたエントリ#3 及び#4 の内、より上位に格納されたエントリ#3 が検索結果として出力されることになる。

10 このような従来例とは異なり、本発明においては、上記の如く抽出されたエントリ#3 及び#4 についてプリフィクス長レジスタに保持された各検索対象フィールド F1 及び F2 のプリフィクス長から、フィールド F1 及び F2 における最長プリフィクス長を求める。

15 すなわち、同図において、フィールド F1 における最長プリフィクス長は、エントリ#3 で求めた 20 ビットであり、フィールド F2 の最長プリフィクス長は、エントリ#4 で求めた 22 ビットである。本発明ではさらに、優先度が最も高い検索対象フィールドにおいて最長プリフィクス長を有するエントリを最終検索結果とする。

従って、優先度が最も高い検索対象フィールドがフィールド F2 であれば、上記の如く抽出されたエントリ#3 及び#4 の内、より下位に格納されたエントリ#4 を最終検索結果とすることが可能になる。

20 このように、優先度を適切に設定することにより、従来のようにエントリ#1～#5 の格納順に拘束されることなく、所望の優先度に基づく検索結果が得られるようになる。

25 上記の優先度が最も高い検索対象フィールドは、該第 3 ステップにおいて求められた該最長プリフィクス長同士の内、最長の検索対象フィールドであってもよい。

すなわち、上記の優先度の指標として最長プリフィクス長を用いれば、該優先度が最も高い検索対象フィールドは、該第 3 ステップにおいて求められた該最長プリフィクス長同士の内、最長になった検索対象フィールドである。

図 1 の場合に、フィールド F1 の最長プリフィクス長である 20 ビットよりも、

フィールド F2 の最長プリフィクス長である 22 ビットの方が長いため、優先度が最も高い検索対象フィールドはフィールド F2 となり、このフィールド F2 において最長プリフィクス長 (22 ビット) を有するエン트리 #4 が検索結果として出力される。

5 このように、エントリの格納順に拘らず、より適切な検索結果を得ることが可能になる。

また、検索開始前に各フィールドに該優先度を設定してもよい。

すなわち、上記の優先度を検索開始前に各フィールドに設定する場合、例えば
10 図 1 におけるフィールド F2 の優先度をフィールド F1 よりも高く設定しておけば、優先度が最も高い検索対象フィールドはフィールド F2 となり、このフィールド F2 において最長プリフィクス長 (22 ビット) を有するエン트리 #4 が検索結果として出力される。

このように、エントリの格納順に拘らず、検索開始前に各フィールドに設定した優先度に基づいた検索結果を得ることが可能になる。

15 この場合の優先度は、検索を行う度に変更可能であればよい。

すなわち、上記の優先度を固定的に設定するのではなく、検索を行う度に変更できるようにすれば、より柔軟に所望の優先度に基づいた検索結果を得ることが可能になる。

20 また、上記の優先度として所定の値が設定されたフィールドを検索対象外にしてもよい。

すなわち、フィールドを検索対象外とする所定の値を定めておくことにより、該優先度として所定の値が設定されたフィールドを検索対象外にすることができる。これにより、最終検索結果を判定するためだけでなく、フィールドを検索対象とするか否かの判断材料としても上記の優先度を流用することが出来る。

25 また、上記の目的を達成するため、本発明に係るデータ検索方法を実現する本発明に係るデータ検索装置は、エント리를複数フィールドに分割して格納したデータベースと、各エントリのフィールド毎のプリフィクス長を保持するプリフィクス長保持部と、全ての検索対象フィールドにおいてエントリの全ての非マスクビットと対応する検索キーのビットデータとが一致するエント리를中間検索結果

とし、該プリフィクス長保持部に保持された該中間検索結果の各検索対象フィールドのプリフィクス長から検索対象フィールド毎の最長プリフィクス長を求め、該中間検索結果の内、優先度が最も高い検索対象フィールドにおいて該最長プリフィクス長を有するエントリを検索結果とする演算部と、を備えたことを特徴として
5

上記の優先度が最も高い検索対象フィールドは、該演算部によって求められた該最長プリフィクス長同士の内、最長の検索対象フィールドであってもよい。

また、検索開始前に各フィールドに該優先度を設定してもよく、この場合の優先度は、検索を行う度に変更可能であればよい。

10 さらに、上記の優先度として所定の値が設定されたフィールドを検索対象外にしてもよい。

図面の簡単な説明

15 図1は、本発明に係るデータ検索方法及び装置の原理を示したブロック図である。

図2は、本発明に係るデータ検索方法及び装置の実施例(1)及び(2)に共通したエントリの状態例を示したブロック図である。

図3は、本発明に係るデータ検索方法及び装置の実施例(1)を示したブロック図である。

20 図4は、本発明に係るデータ検索方法及び装置の実施例(2)を示したブロック図である。

図5は、本発明に係るデータ検索方法及び装置の実施例(2)の変形例を示したブロック図である。

25 図6は、一般的なルーティング処理の概要を説明するためのブロック図である。

図7は、図6におけるルータ101内のルーティング処理に関連した部分の構成例を示したブロック図である。

図8は、図7におけるCAMデバイス200の詳細な構成を示したブロック図である。

図9は、一般的なCAMデバイスに用いられるターナリCAMセルの構成例及び検索結果の真理値表を示した図である。

図10は、一般的なCAMデバイス内部のエントリ構成例及びグローバルマスクパターンを示したブロック図である。

5	符号の説明	
	#1～#5	エントリ
	F1～F5	フィールド
	R1～Rn, 931, 932,	優先度レジスタ
	101	ルータ
	102～105	ネットワーク回線
10	110	パケット
	111	宛先IPアドレス
	112	送信元IPアドレス
	113	宛先ポート番号
	114	送信元ポート番号
	130, 140, 150	サーバ
	200	CAMデバイス
	201, 221	出力バス
	202	ヒットアドレス
	203	データベース
15	210	ネットワークプロセッサ
	211	入力インタフェース
	212, 607	検索キー
	220	コンテキストRAM
	222, 605	検索結果
	321～324, 606, 711～715	一致線
	330	プライオリティエンコーダ
	431, 432	プリフィクス部分
	441, 442, 451, 453	マスク部分
	450	グローバルマスクパターン
20	452	検索対象部分
	600	ターナリCAMセル
	601	データビット
	602	マスクビット
	603	比較処理部
	604	マスク処理部
	801, 802, 901, 902	レジスタ

図中、同一符号は同一又は相当部分を示す。

25

発明を実施するための最良の形態

本発明の実施例として、フィールド毎の最長プリフィクス長同士を比較する場合の実施例(1)と、予めフィールド毎の優先度を設定しておく場合の実施例(2)の2種類について以下に説明する。

まず、上記の実施例(1)及び(2)に共通したエントリ#1～#5の状態を図2に示す。同図は、図1と同様に、データベースに格納されたエントリ#1～#5の各々について、例えば複数のフィールドF1及びF2を検索対象とする検索キー(図示せず)が入力された際の一致/不一致の状態を模式的に示したものである。

- 5 図2には、図1に加え、一致線711～715がそれぞれエントリ#1～#5に対応して示されており、エントリ#1、#2、及び#5については、それぞれ一致線711、712、及び715に対して、下向きの矢印で不一致信号が出力されている(例えば、エントリ#5における不一致信号716)。

10 このように、不一致信号が出力されたエントリ#1、#2、及び#5は、「不一致」であるが、不一致信号が出力されていないエントリ#3及び#4は「一致」であるとして中間検索結果が得られる。

実施例(1)

15 図2の状態において、エントリ#1～#5の格納順に拘らず、検索対象フィールド毎の最長プリフィクス長同士を比較することによって、最終検索結果を得る方法について図3を用いて説明する。

同図に示す如く、エントリ#1～#5の各フィールドF1及びF2のプリフィクス長を保持するレジスタ801及び802を設けておく。すなわち、レジスタ801には、エントリ#1～#5のフィールドF1のプリフィクス長である“24”、“24”、“20”、“17”、及び“12”がそれぞれ保持されている。同様に、レジスタ802には、エン
20 トリ#1～#5のフィールドF2のプリフィクス長である“22”、“17”、“17”、“22”、及び“16”がそれぞれ保持されている。

なお、上記のレジスタ801及び802に対する各エントリ#1～#5のプリフィクス長の値は検索運用時前に予め設定されておく必要があるが、ネットワーク管理者がエントリ作成時にプリフィクス長を入力しておけばよい。或いは、マスクビット
25 トに設定された値をCAMデバイス自身がカウントしてレジスタ801及び802に自動的に入力するようにしてもよい。

同図において、中間検索結果として得られているエントリ#3及び#4について、点線矢印811～814に示す如く、それぞれレジスタ801及び802を参照し、フィールドF1については、エントリ#3のプリフィクス長“20”及びエントリ#4のプリ

フィクス長“17”が得られ、フィールド F2 については、エントリ#3 のプリフィクス長“17”及びエントリ#4 のプリフィクス長“22”が得られる。

これらのプリフィクス長について、点線矢印 815 及び 816 に示す如く各フィールド F1 及び F2 において最大値 821 及び 822 を求めると、フィールド F1 における
5 最大値 821 は、エントリ#3 のプリフィクス長“20”であり、フィールド F2 における最大値 822 は、エントリ#4 のプリフィクス長“22”である。

次に、点線矢印 817 に示す如く各最大値 821 及び 822 同士を比較し、最大の値を求めると、最大値 822 の“22”であることから、フィールド F2 を優先し、CAM 検索結果としては、フィールド F2 のプリフィクス長が最長となるエントリ#4 が
10 最終検索結果になる。

実施例(2)

図 2 の状態において、エントリ#1~#5 の格納順に拘らず、予めフィールド毎の優先度を設定しておくことによって、最終検索結果を得る方法について図 4 を用いて説明する。

15 図 4 に示したレジスタ 901 及び 902 は、図 3 に示したレジスタ 801 及び 802 に相当するものであり、実施例(1)の場合と同様に、エントリ#1~#5 の各フィールド F1 及び F2 のプリフィクス長を保持している。

また、レジスタ 901 及び 902 における最大値 921 及び 922 も、図 3 に示したレジスタ 801 及び 802 における最大値 821 及び 822 と同様にして求めたものである。

20 図 4 は、図 3 と異なり、フィールド F1 及び F2 に対し、それぞれ優先度レジスタ 931 及び 932 が設けられている。なお、優先度レジスタ 931 及び 932 に設定する値は、検索を行なう度に変更可能にしておく。

図示の例では、最高の優先度を示す“1”がフィールド F1 の優先度レジスタ 931 に設定され、フィールド F2 の優先度レジスタ 932 には“2”が設定された状態が
25 示されている。

この場合、点線矢印 933 に示すフィールド間の優先度の比較は、優先度レジスタ 931 及び 932 に基づいて行なわれ、より優先度が高いフィールド F1 において、プリフィクス長が最長となるエントリ#3 が最終検索結果になる。

図 5 は、実施例(2)の変形例として、優先度レジスタをグローバルマスクパター

ンとして使用する場合は示したものである。例えば、 $n < 255$ として、各フィールド $F_1 \sim F_n$ に対する優先度レジスタ $R_1 \sim R_n$ が設けられているものとする。

この場合、非選択を示す優先度である“255”が優先度レジスタ R_n に設定されたフィールド F_n を検索対象から外すことが可能になる。なお、フィールド数が n となっているため非選択を表す値は n より大きくとる必要がある。

以上説明したように、本発明に係るデータ検索方法及び装置によれば、データベースにおいて複数フィールドに分割して格納されたエントリの中から、全ての検索対象フィールドにおいて全ての非マスクビットと、これらに対応する検索キーのビットデータとが一致するエントリを中間検索結果とし、該中間検索結果の各検索対象フィールドのプリフィクス長の中から検索対象フィールド毎の最長プリフィクス長を求め、該中間検索結果の内、優先度が最も高い検索対象フィールドにおいて該最長プリフィクス長を有するエントリを検索結果とするように構成したので、エントリの格納順に拘束されることなく、より適切な検索結果を得ることが可能になる。

15

請求の範囲

1. データベースにおいて複数フィールドに分割して格納されたエントリの中から、全ての検索対象フィールドにおいて全ての非マスクビットと、これらに対応する検索キーのデータビットとが一致するエントリを中間検索結果とする第1ステップと、
5
該中間検索結果の各検索対象フィールドのプリフィクス長の中から該検索対象フィールド毎の最長プリフィクス長を求める第2ステップと、
該中間検索結果の内、優先度が最も高い検索対象フィールドにおいて該最長プリフィクス長を有するエントリを検索結果とする第3ステップと、
10
を有することを特徴とするデータ検索方法。
2. 請求の範囲1において、
該優先度が最も高い検索対象フィールドが、該第3ステップにおいて求められた該最長プリフィクス長同士の内、最長の検索対象フィールドであることを特徴とするデータ検索方法。
15
3. 請求の範囲1において、
検索開始前に各フィールドに該優先度を設定することを特徴とするデータ検索方法。
4. 請求の範囲3において、
20
該優先度は、検索を行う度に変更可能であることを特徴とするデータ検索方法。
5. 請求の範囲3において、
該優先度として所定の値が設定されたフィールドを検索対象外にすることを特徴とするデータ検索方法。
- 25
6. エントリを複数フィールドに分割して格納したデータベースと、
各エントリのフィールド毎のプリフィクス長を保持するプリフィクス長保持部と、
全ての検索対象フィールドにおいてエントリの全ての非マスクビットと対応する検索キーのビットデータとが一致するエントリを中間検索結果とし、該プリフ

イクス長保持部に保持された該中間検索結果の各検索対象フィールドのプリフィクス長から検索対象フィールド毎の最長プリフィクス長を求め、該中間検索結果の内、優先度が最も高い検索対象フィールドにおいて該最長プリフィクス長を有するエントリを検索結果とする演算部と、

5 備えたことを特徴とするデータ検索装置。

7. 請求の範囲6において、

該優先度が最も高い検索対象フィールドが、該演算部によって求められた該最長プリフィクス長同士の内、最長の検索対象フィールドであることを特徴とするデータ検索装置。

10 8. 請求の範囲6において、

検索開始前に各フィールドに該優先度を設定したことを特徴とするデータ検索装置。

9. 請求の範囲8において、

該優先度は、検索を行う度に変更可能であることを特徴とするデータ検索装置

15 。

10. 請求の範囲8において、

該優先度として所定の値が設定されたフィールドを検索対象外にすることを特徴とするデータ検索装置。

図1

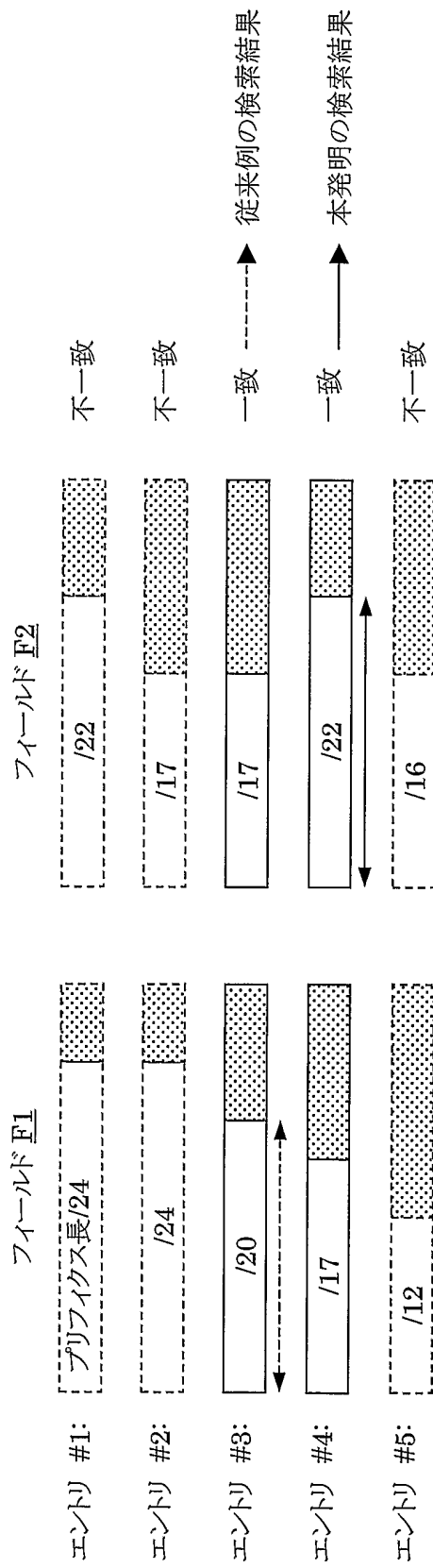


図2

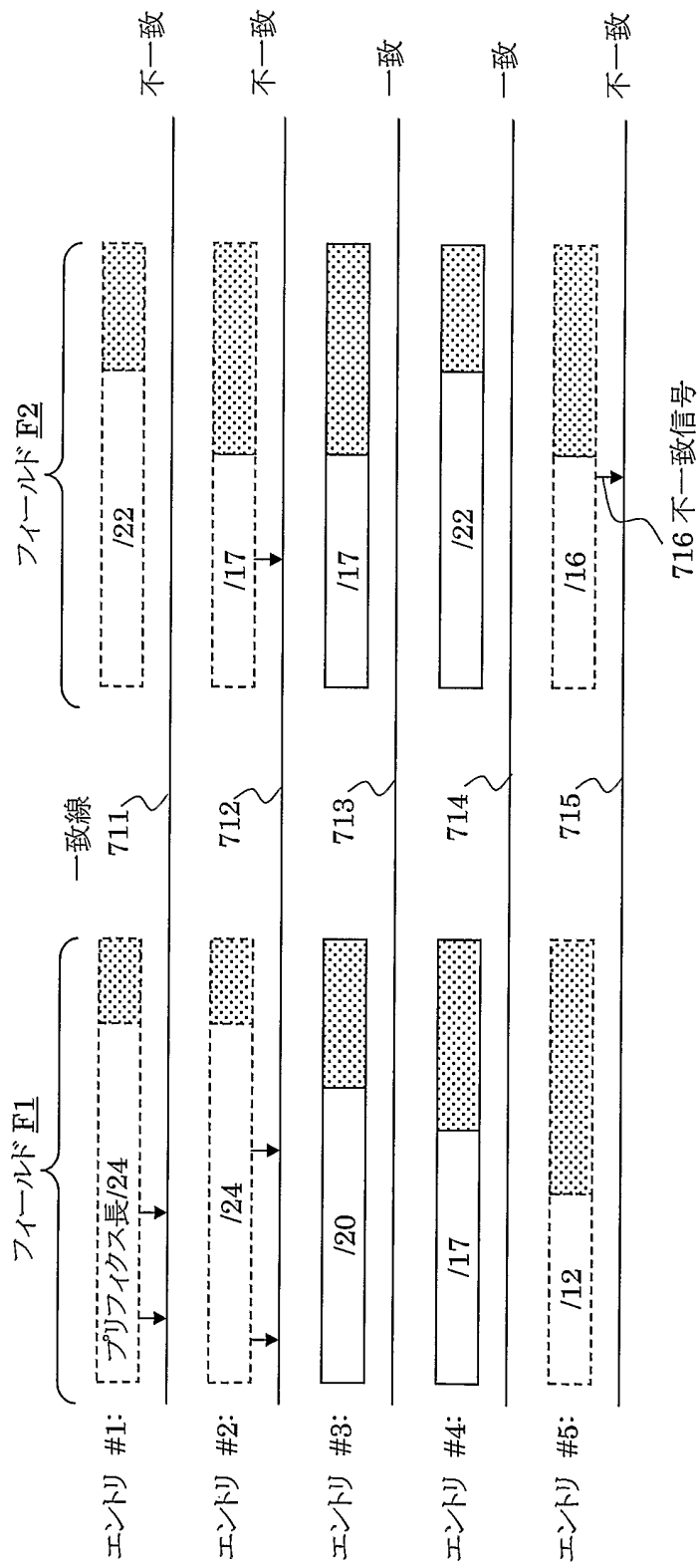


図3

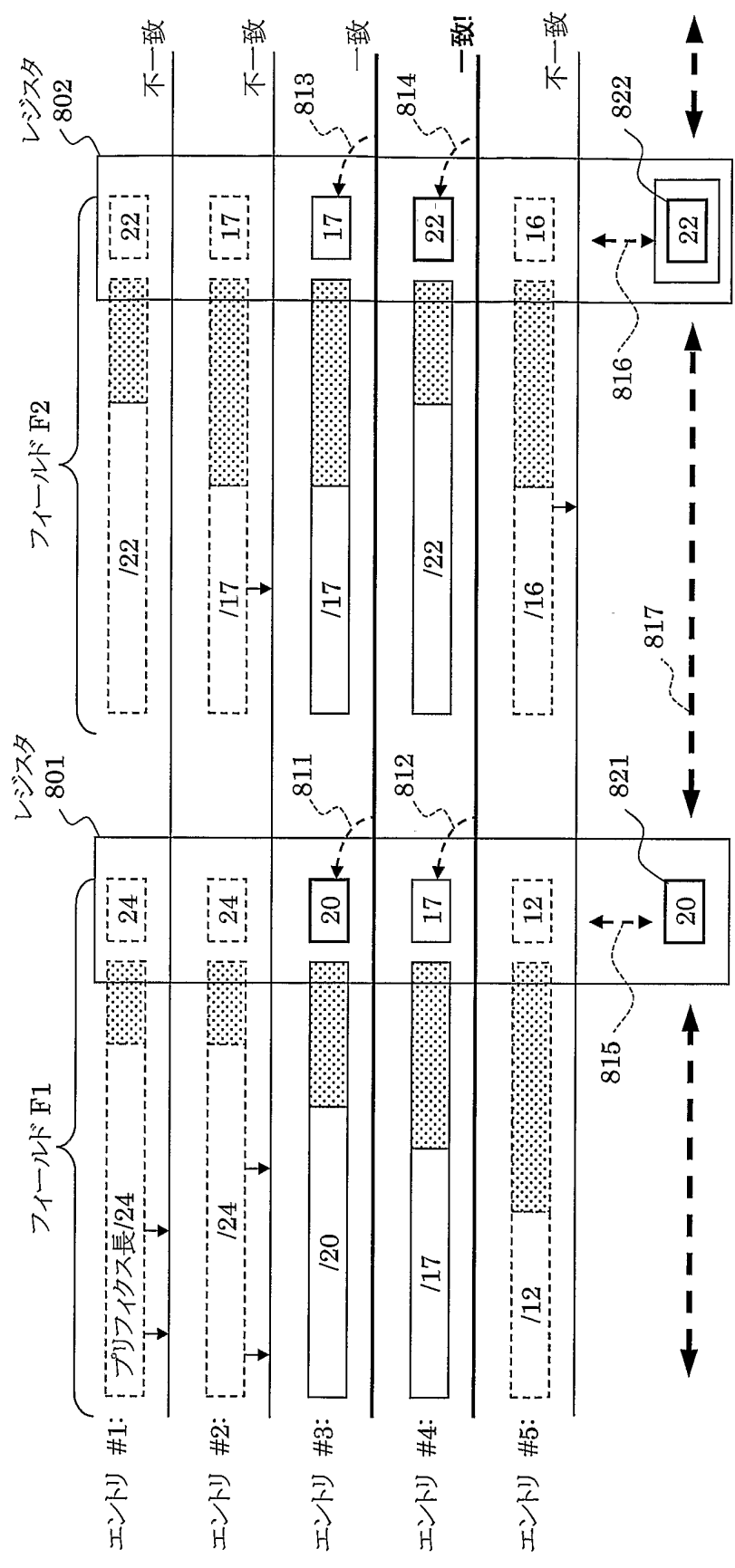


図4

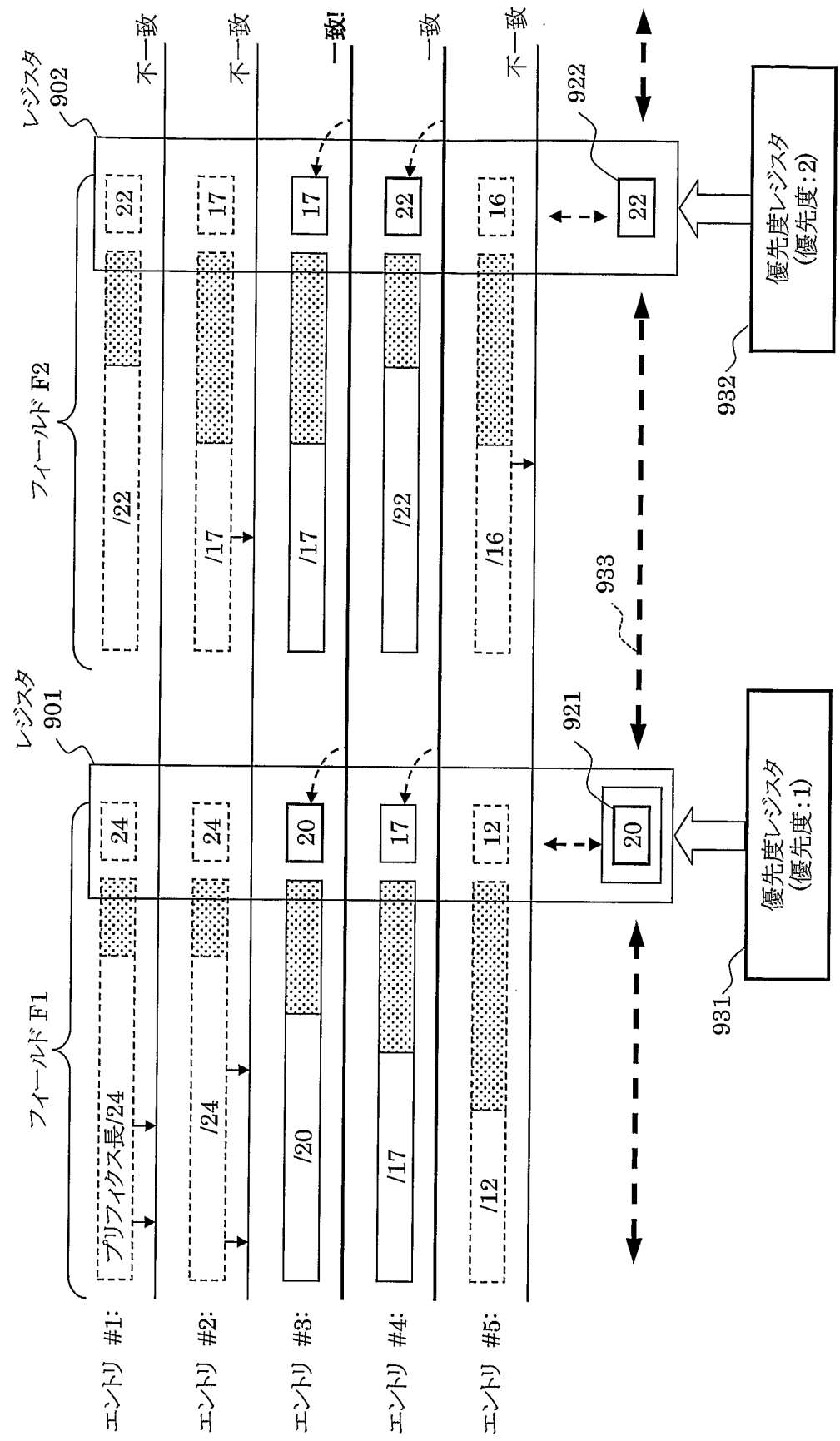


図5

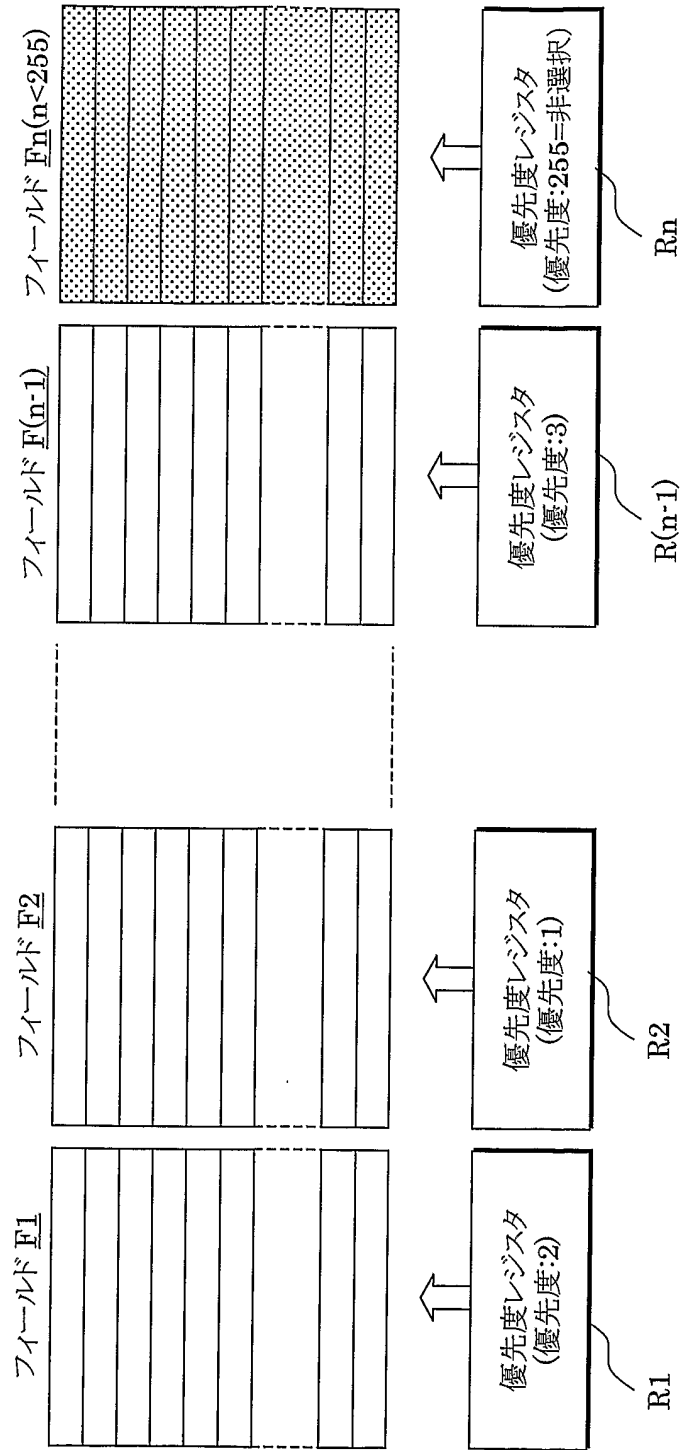


図6

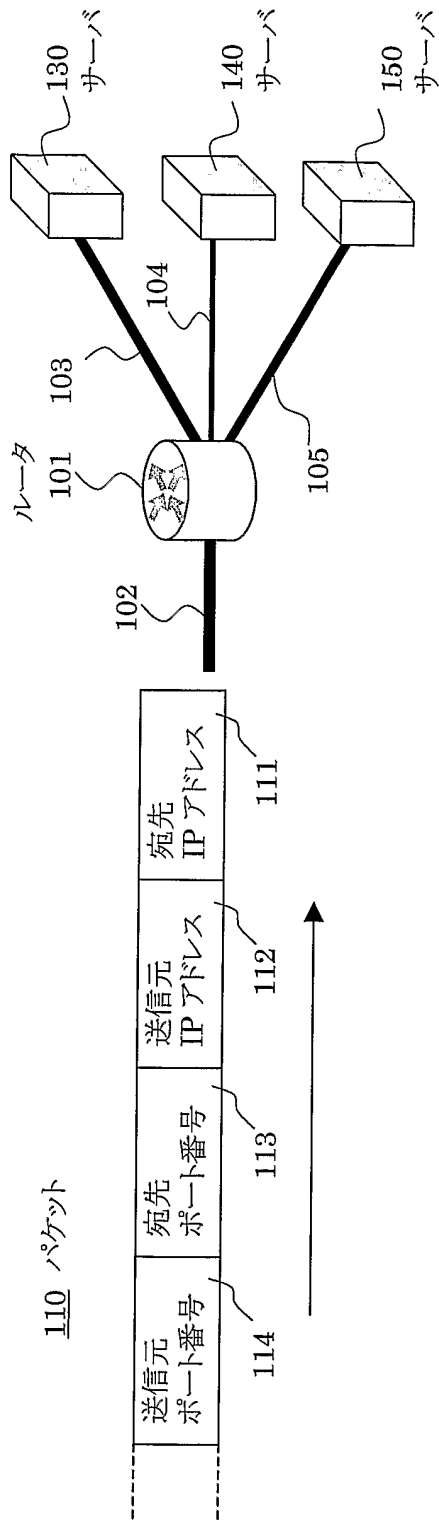
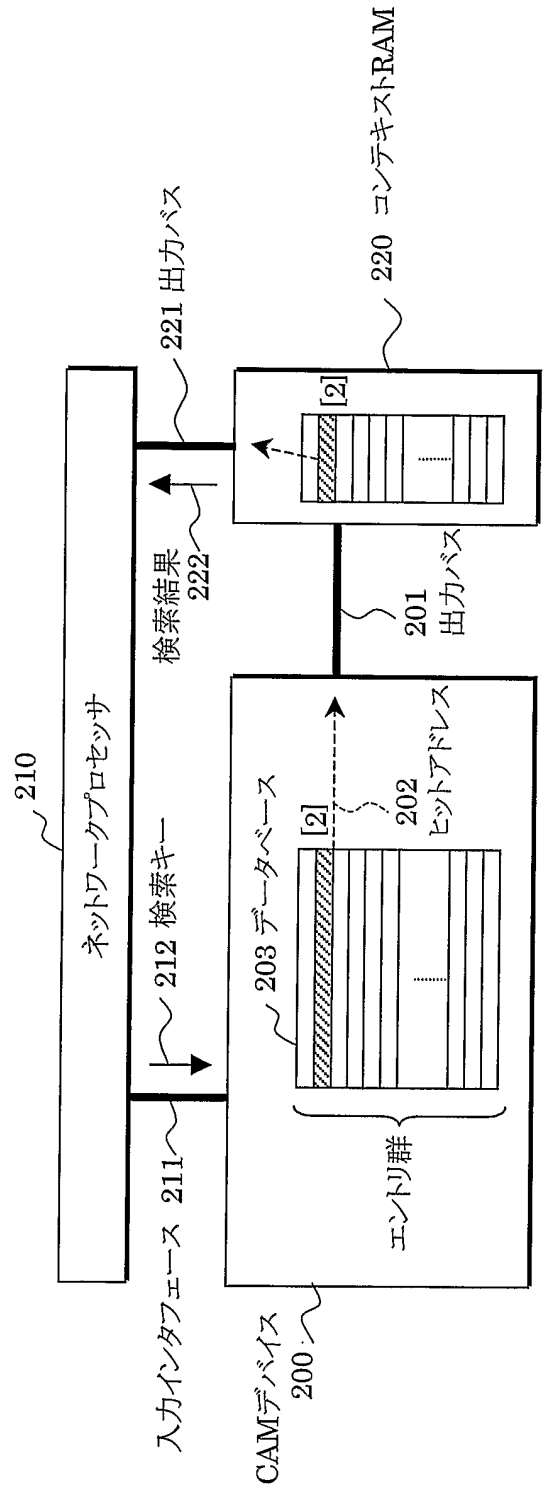


図7



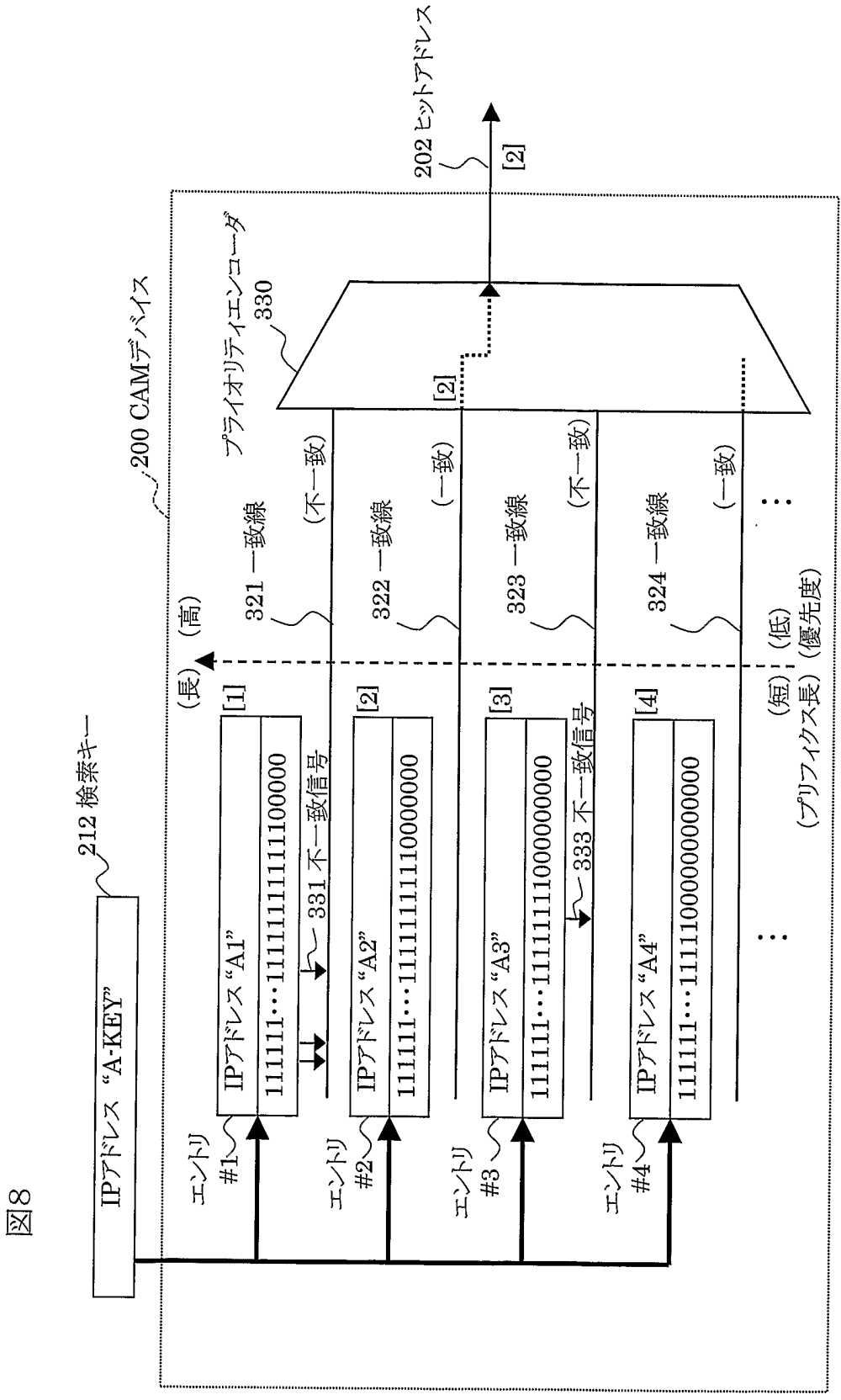
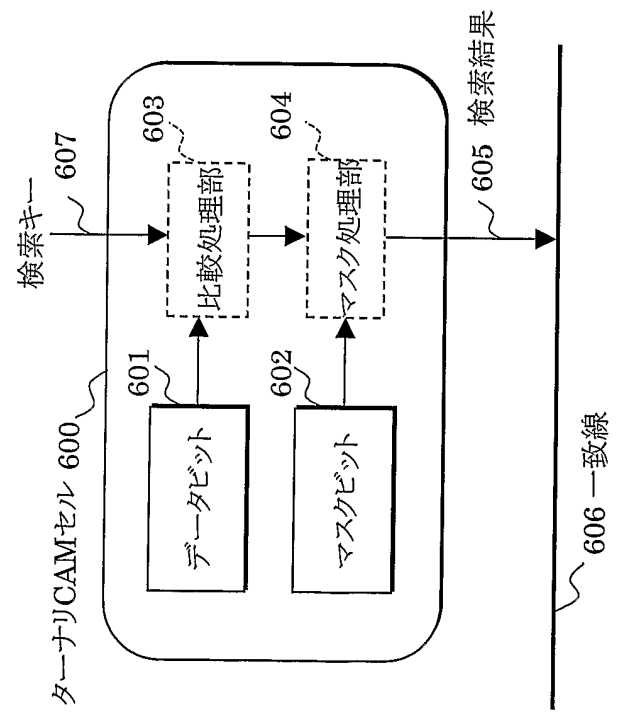


図9

(1)

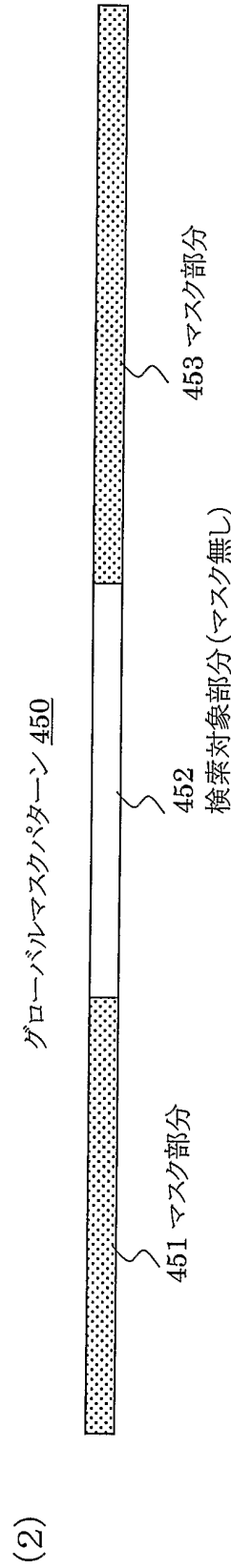
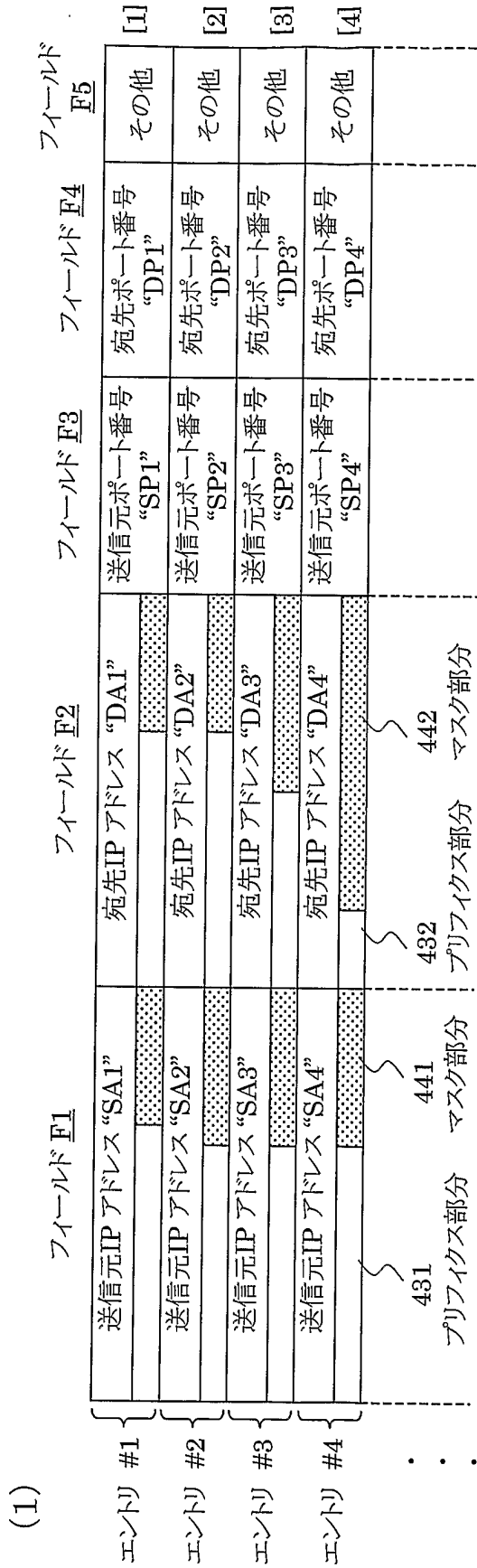


(2)

検索結果の真理値表 610

	データ "1"	データ "0"
キー"1"	マスク"1"	"0"
	マスク"0"	"1"
キー"0"	マスク"1"	"1"
	マスク"0"	"1"

図10



INTERNATIONAL SEARCH REPORT

International application No. PCT/JP03/09740

A. CLASSIFICATION OF SUBJECT MATTER
Int.Cl⁷ H04L12/56, G06F17/30

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.Cl⁷ H04L12/56, G06F17/30

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

Jitsuyo Shinan Koho	1926-1996	Toroku Jitsuyo Shinan Koho	1994-2003
Kokai Jitsuyo Shinan Koho	1971-2003	Jitsuyo Shinan Toroku Koho	1996-2003

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 2000-349811 A (NEC Corp.), 15 December, 2000 (15.12.00), Fig. 3 (Family: none)	1-10
A	JP 11-284658 A (Kawasaki Steel Corp.), 15 October, 1999 (15.10.99), Fig. 4 (Family: none)	1-10
A	JP 11-261647 A (NEC Corp.), 24 September, 1999 (24.09.99), Fig. 3 (Family: none)	1-10

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

<p>* Special categories of cited documents:</p> <p>"A" document defining the general state of the art which is not considered to be of particular relevance</p> <p>"E" earlier document but published on or after the international filing date</p> <p>"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)</p> <p>"O" document referring to an oral disclosure, use, exhibition or other means</p> <p>"P" document published prior to the international filing date but later than the priority date claimed</p>	<p>"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</p> <p>"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</p> <p>"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</p> <p>"&" document member of the same patent family</p>
--	---

Date of the actual completion of the international search 04 November, 2003 (04.11.03)	Date of mailing of the international search report 18 November, 2003 (18.11.03)
---	--

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. Cl⁷ H04L12/56, G06F17/30

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. Cl⁷ H04L12/56, G06F17/30

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 1926-1996年
- 日本国公開実用新案公報 1971-2003年
- 日本国登録実用新案公報 1994-2003年
- 日本国実用新案登録公報 1996-2003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 請求の範囲の番号
A	JP 2000-349811 A (日本電気株式会社)、2000.12.15、図3 (ファミリー無し)	1~10
A	JP 11-284658 A (川崎製鉄株式会社)、1999.10.15、図4 (ファミリー無し)	1~10
A	JP 11-261647 A (日本電気株式会社)、1999.09.24、図3 (ファミリー無し)	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


- 「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 04.11.03

国際調査報告の発送日 18.11.03

国際調査機関の名称及びあて先
 日本国特許庁 (ISA/JP)
 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)
 石井 研一  5X 8124
 電話番号 03-3581-1101 内線 3556