

(19) 日本国特許庁(JP)

(12) 特 許 公 報(B2)

(11) 特許番号

特許第4685299号  
(P4685299)

(45) 発行日 平成23年5月18日(2011.5.18)

(24) 登録日 平成23年2月18日(2011.2.18)

(51) Int.Cl.

F I

H04L 12/56 (2006.01)

H04L 12/56 260A

請求項の数 30 (全 23 頁)

(21) 出願番号 特願2001-500493 (P2001-500493)  
 (86) (22) 出願日 平成12年6月1日(2000.6.1)  
 (65) 公表番号 特表2003-501881 (P2003-501881A)  
 (43) 公表日 平成15年1月14日(2003.1.14)  
 (86) 国際出願番号 PCT/US2000/015417  
 (87) 国際公開番号 W02000/074312  
 (87) 国際公開日 平成12年12月7日(2000.12.7)  
 審査請求日 平成19年5月21日(2007.5.21)  
 (31) 優先権主張番号 60/137,153  
 (32) 優先日 平成11年6月1日(1999.6.1)  
 (33) 優先権主張国 米国(US)  
 (31) 優先権主張番号 09/412,815  
 (32) 優先日 平成11年10月5日(1999.10.5)  
 (33) 優先権主張国 米国(US)

(73) 特許権者 501438485  
 ヤフー！ インコーポレイテッド  
 アメリカ合衆国 カリフォルニア州 94  
 089 サニーヴェイル ファースト ア  
 ヴェニュー 701  
 (74) 代理人 100082005  
 弁理士 熊倉 禎男  
 (74) 代理人 100067013  
 弁理士 大塚 文昭  
 (74) 代理人 100086771  
 弁理士 西島 孝喜  
 (74) 代理人 100109070  
 弁理士 須田 洋之  
 (74) 代理人 100109335  
 弁理士 上杉 浩

最終頁に続く

(54) 【発明の名称】 マルチキャストする方法および装置

(57) 【特許請求の範囲】

【請求項 1】

データネットワークにおけるノードにおいてデータを処理する処理エージェントであって、該データネットワークは、複数のノードと、該複数のノードの少なくとも一部であって、マルチキャストグループを形成する少なくとも一部とを接続し、該処理エージェントは、

該マルチキャストグループに関連した複数のエントリを格納するように作用可能なデータストアと、

該マルチキャストグループのメンバーに該複数のエントリのそれぞれを広めるように作用可能であるロジックと、

を備え、

前記処理エージェントが前記マルチキャストグループにおいてランデブーノードとして指定されたことを指示するように作用可能なロジックをさらに備え、前記ランデブーノードとしての指定は、前記処理エージェントが前記マルチキャストグループのメンバーに前記複数のエントリのそれぞれを広めることを指示するものである  
 処理エージェント。

【請求項 2】

前記マルチキャストグループのメンバーでない外部ノードから、クエリを実行する要求を受信するように作用可能であるロジックであって、該クエリが整合基準を特定する、ロジックと、

前記データストア内の前記複数のエントリに対して該クエリを実行するように作用可能であるロジックと、

前記外部ノードに、該整合基準を満たす一つ以上のエントリを広めるように作用可能であるロジックと

をさらに備える、請求項 1 に記載の処理エージェント。

【請求項 3】

第 1 のノードからの要求であって、第 1 のエントリを追加する要求に応答して、前記データストア内の前記複数のエントリに該第 1 のエントリを追加するように作用可能であるロジックをさらに備える、請求項 1 に記載の処理エージェント。

【請求項 4】

前記広めるように作用可能であるロジックは、前記第 1 のノードからの要求であって、前記複数のエントリに前記第 1 のエントリを追加する要求に応答して、前記マルチキャストグループを形成する該複数のノードに、該第 1 のエントリを自動的に広めるようにさらに作用可能である、請求項 3 に記載の処理エージェント。

【請求項 5】

第 1 のノードからの要求であって、前記データストア内の前記複数のエントリのうちの第 1 のエントリを破棄する要求に応答して、該第 1 のエントリを除去するロジックをさらに備える、請求項 1 に記載の処理エージェント。

【請求項 6】

前記マルチキャストグループを形成する前記複数のノードに、前記第 1 のエントリが破棄されたことを指示するように作用可能であるロジックをさらに備え、該指示が、前記第 1 のノードからの要求であって、該第 1 のエントリを破棄する要求に応答する、請求項 5 に記載の処理エージェント。

【請求項 7】

前記複数のエントリの内の特定のエントリを発行したソースが、前記マルチキャストグループのメンバーでない、請求項 1 に記載の処理エージェント。

【請求項 8】

前記複数のエントリの内の特定のエントリを発行したソースが、前記マルチキャストグループのメンバーである、請求項 1 に記載の処理エージェント。

【請求項 9】

前記複数のエントリのそれぞれが、他のエントリに対する自身の送達優先順位を特定する優先順位に関連している、請求項 1 に記載の処理エージェント。

【請求項 10】

データネットワークにおける選択されたノードに結合された処理エージェントの動作方法であって、該データネットワークは、複数のノードと、該複数のノードの少なくとも一部であって、該選択されたノードを含んでおり、マルチキャストグループを形成する少なくとも一部とを接続し、該方法は、

該マルチキャストグループに関連した複数のエントリを該処理エージェントにおいて格納するステップと、

該マルチキャストグループのメンバーに該複数のエントリのそれぞれを広めるステップと

を包含し、

前記選択されたノードは前記マルチキャストグループでランデブーノードとして指定され、

前記ランデブーノードとしての指定は、前記選択されたノードが前記マルチキャストグループのメンバーに前記複数のエントリのそれぞれを広めることを指示するものである方法。

【請求項 11】

前記マルチキャストグループのメンバーでないノードからの要求であって、前記処理エージェントにおいて格納された前記複数のエントリに対してクエリを実行する要求を受信

10

20

30

40

50

するステップであって、該クエリが整合基準を特定する、ステップと、

該マルチキャストグループのメンバーでない該ノードに、該整合基準を満たす一つ以上のエントリを広めるステップと

をさらに包含する、請求項 1 0 に記載の方法。

【請求項 1 2】

第 1 のノードからの要求であって、第 1 のエントリを追加する要求に応答して、前記処理エージェントにおいて格納された前記複数のエントリに該第 1 のエントリを追加するステップをさらに包含する、請求項 1 0 に記載の方法。

【請求項 1 3】

前記第 1 のノードからの要求であって、前記第 1 のエントリを追加する要求に応答して、前記マルチキャストグループを形成する前記複数のノードに、該第 1 のエントリを自動的に広めるステップをさらに包含する、請求項 1 2 に記載の方法。

10

【請求項 1 4】

第 1 のノードからの要求であって、前記処理エージェントにおいて格納された前記複数のエントリのうちの第 1 のエントリを破棄する要求に応答して、該第 1 のエントリを除去するステップをさらに包含する、請求項 1 0 に記載の方法。

【請求項 1 5】

前記マルチキャストグループを形成する前記複数のノードに、前記処理エージェントにおいて格納された前記複数のエントリのうちの前記第 1 のエントリが破棄されたことを指示するステップをさらに包含し、該指示するステップが、前記第 1 のノードからの要求であって、該第 1 のエントリを破棄する要求に応答して実行される、請求項 1 4 に記載の方法。

20

【請求項 1 6】

前記複数のエントリの内の特定のエントリを発行したソースが、前記マルチキャストグループのメンバーでない、請求項 1 0 に記載の方法。

【請求項 1 7】

前記複数のエントリの内の特定のエントリを発行したソースが、前記マルチキャストグループのメンバーである、請求項 1 0 に記載の方法。

【請求項 1 8】

前記複数のエントリのそれぞれが、他のエントリに対する自身の送達優先順位を特定する優先順位に関連している、請求項 1 0 に記載の方法。

30

【請求項 1 9】

特定のノードから、前記処理エージェントにおいて格納された前記複数のエントリに対してクエリを実行する要求を受信するステップであって、該クエリは、一つ以上のエントリを発行したソースを特定する、ステップと、

該特定のノードに、第 1 のノードに対する変更を非同期的に通知するステップと

をさらに包含し、該特定のノードに非同期的に通知するステップが、該第 1 のエントリを発行したソースが該第 1 のエントリを変更したことに応答して、実行される、請求項 1 0 に記載の方法。

【請求項 2 0】

40

データを伝送するデータネットワークであって、該データネットワークは、複数のノードと、該複数のノードの少なくとも一部であって、マルチキャストグループを形成する少なくとも一部とを接続し、該データネットワークは、複数の処理エージェントを備え、

該複数の処理エージェントのそれぞれは、

該マルチキャストグループに関連した複数のエントリを格納するように作用可能なデータストアと、

該マルチキャストグループのメンバーに該複数のエントリのそれぞれを広めるように作用可能であるロジックと、

前記複数の処理エージェントの特定の処理エージェントが前記マルチキャストグループにおいてランデブーノードとして指定されたことを指示するように作用可能なロジックで

50

あって、前記ランデブーノードとしての指定は、前記特定の処理エージェントが前記マルチキャストグループのメンバーに前記複数のエントリのそれぞれを広めることを指示するものであるロジックと、

を備える、  
データネットワーク。

【請求項 2 1】

該マルチキャストグループ内の該複数のノードのそれぞれにおいて、グループ状態を含むデータストアを維持するステップと、

前記ランデブーノードにおいて状態アップデートを受信するステップと、

該状態アップデートを用いて、該ランデブーノードにおいて、該データストア内の該グループ状態をアップデートするステップと、

該ランデブーノードから、該マルチキャストグループ内の他のノードに該状態アップデートを伝搬するステップと、

該マルチキャストグループ内の他のノードにおいて、該データストア内の該グループ状態をアップデートするステップと

を包含する、請求項 1 0 に記載の方法。

【請求項 2 2】

前記複数のノードのうちの 1 つである接続ノードが、前記マルチキャストグループに加えられ、特定のデータストアを有しており、前記方法が、

前記グループ状態を該接続ノードに伝搬するステップと、

該グループ状態を用いて、該接続ノードにおいて該特定のデータストアをアップデートするステップと

をさらに包含する、請求項 2 1 に記載の方法。

【請求項 2 3】

前記グループ状態を伝搬するステップが、該グループ状態を、前記マルチキャストグループ内の特定のノードから前記接続ノードに伝搬するステップを包含する、請求項 2 2 に記載の方法。

【請求項 2 4】

前記グループ状態を伝搬するステップが、該グループ状態を、前記マルチキャストグループ内の特定のノードであって、前記接続ノードの最も近傍にある特定のノードから、該接続ノードに伝搬するステップを包含する、請求項 2 3 に記載の方法。

【請求項 2 5】

前記グループ状態を伝搬するステップが、該グループ状態を、前記マルチキャストグループ内の特定のノードから前記接続ノードに伝搬するステップを包含し、該特定のノードが、ネットワークルーティングプロトコルから決定される、請求項 2 3 に記載の方法。

【請求項 2 6】

前記状態アップデートを伝搬するステップが、信頼できるプロトコルを用いて該状態アップデートを伝搬するステップを包含する、請求項 2 1 に記載の方法。

【請求項 2 7】

前記ランデブーノードにおける前記データストアが特定の複数のエントリを有し、該特定の複数のエントリのそれぞれのエントリが、該エントリを発行したソースを識別する、請求項 2 1 に記載の方法。

【請求項 2 8】

信頼できるプロトコルを使用して前記複数のエントリの中の特定のエントリを広めるように作用可能なロジックをさらに備える請求項 1 に記載の処理エージェント。

【請求項 2 9】

状態アップデートを第 2 の処理エージェントから受信するように作用可能なロジックと

前記データストアにおいて前記状態アップデートのための情報を格納するように作用可能なロジックと、

10

20

30

40

50

前記マルチキャストグループにおいて1つ以上のノードに前記状態アップデートを伝搬するように作用可能なロジックと、をさらに備える請求項1に記載の処理エージェント。

【請求項30】

前記状態アップデートは前記ランデブーノード以外の前記マルチキャストグループ内の特定のノードから受信される請求項21に記載の方法。

【発明の詳細な説明】

【0001】

(関連出願の相互参照)

本願は、1999年6月1日出願した同時係属中の米国仮特許出願第60/137,153号の優先権を主張する。本願は、1999年6月1日出願した、タイトル「PERFORMING MULTICAST COMMUNICATION IN COMPUTER NETWORKS BY USING OVERLAY ROUTING」の米国特許出願第09/323,869号(以下、「McCanne'869」と参照)および、1999年8月27日出願した、タイトル「SYSTEM FOR BANDWIDTH ALLOCATION IN A COMPUTER NETWORK」の米国特許出願第09/384,865号(以下、「McCanne'865」と参照)に関する。これらの出願の各々を、本願において詳細に記載されているように、参考として援用する。

【0002】

(発明の分野)

本発明は、一般に、コンピュータネットワークの分野に関し、より詳細には、コンピュータネットワークで使用するためのマルチポイント転送プロトコルに関する。

【0003】

(発明の背景)

インターネットの大衆化が進むにつれて、インターネットを介して、マルチメディア情報などの情報の「マルチキャスト」を備えることが望ましい。マルチキャストは、データネットワーク上のホストからそのデータネットワーク上の選択された複数のホストに情報を伝送するプロセスである。この「選択された複数」は、しばしば、「マルチキャストグループ」と呼ばれる。

【0004】

データのユニキャスト(unicast)送達インターネットの基本構成要素として多大な成功を収めている一方で、マルチキャストはかなり複雑であることが示され、マルチキャストが広いエリアにわたって展開することを妨げる多くの技術的な障壁が残っている。例えば、ドメイン間のマルチキャストのルーティングは、まだうまく実現されておらず、現在の形態のマルチキャストはインターネットの全体にわたって広く展開できないかもしれないと考えられる多くの理由が存在する。他方、1つの管理されたネットワークドメインに限定されている場合、マルチキャストは、設定および管理がかなり容易になり、あるアプリケーションに対して、許容される性能を提供し得る。

【0005】

現在のマルチキャスト技術に関連する1つの課題は、1つの管理されたネットワークドメインにおいてさえ、グループメンバーが行き来した時に、情報が全ての現在のグループメンバーに確実に送達されることを保証する送達機構はないことである。さらに、一般的には、マルチキャストグループの全体にわたって情報を効率的にルーティングすること保証する送達機構はない。そのような送達機構がないために、マルチキャストの使用は、信頼できる送達と効率的なルーティングとが必要とされないアプリケーションにおける使用に大きく制限されている。

【0006】

(発明の要旨)

本発明は、データネットワーク内のマルチポイントインフラストラクチャトランスポート(MINT)プロトコルをインプリメントするための方法および装置を提供する。MIN

10

20

30

40

50

Tプロトコルは、データネットワークの単一ノードと他の全てのインフラストラクチャ、ならびに、特定のグループに属する、データネットワークのエンドホストノードとの間に、信頼性情報送達機構を提供する。本発明は、スパース(sparse)モードPIM状すなわちコアベースツリー(CBT)のIPマルチキャストルーティングプロトコル、または、他のマルチキャストプロトコルを用いて形成されたグループ(このマルチキャストグループは、関連したランデブーポイントまたはノードを有する)で使用するのに適している。そのようなプロトコルの例は、McCannet'869に記載され、そこで、オーバーレイマルチキャストネットワーク(OMN)の記載で開示される。

【0007】

本発明のある実施形態は、データネットワーク内でデータを配信するための方法を提供する。データネットワークは複数のノードを接続し、その複数のノードの少なくとも一部がマルチキャストグループを形成する。マルチキャストグループのノードの1つは、ランデブーノードとして指定される。その方法は、マルチキャストグループにおけるノードの各々でグループ状態を含むデータ格納部を保持するステップを包含する。ランデブーノードで受信された状態アップデートを使用して、ランデブーノードのデータ格納部内のグループ状態をアップデートする。状態アップデートは、信頼性プロトコルを用いて、ランデブーノードからマルチキャストグループ内の他のノードに伝搬される。最終的に、マルチキャストグループ内の他のノードでのデータ格納部内のグループ状態をアップデートする。

【0008】

本発明の別の実施形態において、データネットワーク内のあるノードでデータを処理するための処理エージェントが提供される。データネットワークは複数のノードを接続し、その複数のノードの少なくとも一部がマルチキャストグループを形成する。マルチキャストグループのノードの1つは、ランデブーノードとして指定される。処理エージェントは、状態メモリと、プロトコルプロセッサとを含む。プロトコルプロセッサは、データネットワーク内の選択されたノードに結合されたロジックを有し、かつ、信頼性プロトコルを用いたデータチャネルを介してデータネットワーク内の他の処理エージェントでデータを伝送し、受信するためのロジックを有する。プロトコルプロセッサはまた状態メモリに結合され、データを、それぞれ、状態メモリに格納し、状態メモリから取り出すためのロジックを有する。

【0009】

(特定の実施形態の説明)

ある実施形態において、本発明は、データネットワーク内のMINTプロトコルをインプリメントし、データネットワーク内のセnderノードとマルチキャストグループのメンバー、インフラストラクチャおよび/またはデータネットワーク内のエンドホストとの間の信頼性情報送達機構を提供するための方法および装置を提供する。

【0010】

MINTを用いて、セnderは、データネットワーク内に、かつ、データネットワークにわたって発行された名前の付けられた値をマルチキャストグループに関連付け、それにより、他のグループメンバーおよびネットワークエンティティが、分散状態のこの「データベース」をクエリする。「ミント(min t)」とよばれるデータベース内の各タプルは、そのオーナ(そのマルチキャストのセnder)、名前およびマルチキャストグループによって、識別される。このミントは、アクティブグループ参加者でネットワークの全ての部分に確実に広められる。好ましくは、ミントは、ソースからそのグループのアクティブレシーバのセットまでのパスに沿って入るルータにのみ流れる。この結果、MINT情報の効率的なルーティングを生じ、これは、効率的なルーティングおよび配信にかかわらず情報をネットワーク全体に流すことによって動作する、従来のシステムに対して利点がある。エンドホストは、それぞれのオーナによって発行された全ての既知のミントを発見し、かつ/または、列挙するためのマルチキャストシステムにクエリし得る。今度は、ミント値は、名前/オーナを参照してクエリし得、クエリを行なうエージェントは、オーナが値を変更する場合に、非同期的に通知され得る。

10

20

30

40

50

## 【 0 0 1 1 】

ある実施形態において、特定のミントが、システム特有の機能のために保存される。システム特有の機能とは、例えば、グループをアプリケーションのタイプにマッピングするか、または、グループの属性を説明する機能である。それにより、グループは、ネットワークの異なる部分で局地的に規定されたトラフィッククラスにマッピングされ得る。例えば、伝送されたデータストリームが、アプリケーションレベル処理および/またはトラフィック管理を必要とする場合、具体的な「セットアップミント」は、必須の情報を提供し、データの伝送を進める。

## 【 0 0 1 2 】

別の実施形態において、情報ソースは、MINTプロトコルを使用して、グループ内に入れられたデータストリームに注釈をつけたミントを発行し得る。該当するグループに対するマルチキャストツリーの各ノードに位置付けられた、専用パケット転送エンジンは、ストリームの注釈に基づいて、受信したデータストリームを処理する。例えば、そのパケット転送エンジンは、ストリームの注釈に基づいて、データストリームにネットワーク帯域を割り当て得る。

## 【 0 0 1 3 】

図1Aは、本発明で使用するのに適したコンピュータシステム1を示す。コンピュータシステム1は、ディスプレイスクリーン5を有するディスプレイ3を含む。キャビネット7は、ディスクドライブ、CDROMドライブ、ディスプレイアダプタ、ネットワークカード、ランダムアクセスメモリ(RAM)、中央処理ユニット(CPU)および他の構成要素、サブシステム、デバイスなどの標準的なコンピュータ構成要素(図示せず)を収容する。ボタン13を有するマウス11およびキーボード9などのユーザ入力デバイスが示される。他のユーザ入力デバイス(トラックボール、タッチスクリーン、ディジタイザタブレット等)が使用されてもよい。一般に、コンピュータシステムは、本発明で使用するのに適したコンピュータシステムのあるタイプ(例えば、デスクトップコンピュータ)の例示にすぎない。コンピュータは、多くの異なるハードウェア構成要素で構成され得、多くの寸法およびスタイル(例えば、ラップトップ、パームトップ、ペントップ、サーバ、ワークステーション、メインフレーム)で作製され得る。本明細書で記載される処理を実行するのに適した任意のハードウェアプラットフォームは、本発明で使用するのに適している。

## 【 0 0 1 4 】

図1Bは、コンピュータ内に一般に見出され得るサブシステム(例えば、コンピュータ1のキャビネット7)を図示する。図1Bにおいて、サブシステムは内部バス22と直接的にインターフェースをとる。サブシステムは、入力/出力(I/O)コントローラ24、システムランダムアクセスメモリ(RAM)26、中央処理ユニット(CPU)28、ディスプレイアダプタ30、シリアルポート40、固定ディスク42、ネットワークインターフェースアダプタ44を含む。バス22の使用により、サブシステムの各々は、サブシステム間でのデータ転送、そして、最も重要にはCPUへのデータ転送が可能になる。外部デバイスは、バス22を介して、バス上のサブシステムとインターフェースをとることによって、CPUまたは他のサブシステムと連絡し得る。モニタ46はディスプレイアダプタ30を介してバスに接続する。マウスなどの相対ポインティングデバイス(RPD)48はシリアルポート40を介して接続する。キーボード50などのいくつかのデバイスは、メインデータバスを用いることなく直接的な手段(例えば、割り込みコントローラおよび関連レジスタ(図示せず)によって、CPUと連絡し得る。

## 【 0 0 1 5 】

図1Aに示される外部の物理的構成と同様に、多くのサブシステム構成が可能である。図1Bは、ある適切な構成の例示にすぎない。図1Bに示されるものとは異なる、サブシステム、構成要素またはデバイスを追加してもよい。図1Bに示されるサブシステムの全てを用いることなく、適切なコンピュータシステムを達成することができる。例えば、CDROMドライブ、グラフィックアクセラレータなどの他のサブシステムは、本発明のシス

10

20

30

40

50

テムの性能に影響を与えることなく、その構成において含まれ得る。

【 0 0 1 6 】

図 2 は、本発明に従って構築されたデータネットワーク 2 0 0 の一部を示す。ネットワーク 2 0 0 は、ルーティングノード（またはネイティブルータ）2 0 2、2 0 4、2 0 6、2 0 8、2 1 0 および 2 1 2 を含む。ノードは、2 1 4、2 1 6 および 2 1 8 に示されるリンクのような双方向リンクによって相互接続されている。ネットワーク 2 0 0 には、情報ソース 2 2 0 および 2 2 2 が含まれる。ネットワーク 2 0 0 にはまた、さらなる双方向リンクを介してルーティングノードに接続されたクライアント 2 2 4、2 2 6、2 2 8 および 2 3 0 が示されている。ネットワーク 2 0 0 は、ルーティングノード 2 0 4 および 2 1 0 が、それぞれ、実際には、ドメイン X およびドメイン Y のような他のネットワークドメインの一部であり得ることを示す。従って、ネットワーク 2 0 0 は、本発明の単なる 1 つの実施形態を示す。当業者には、本発明の他の実施形態が他のタイプのネットワークアーキテクチャまたはトポロジーに用いられ得ることが明らかである。

10

【 0 0 1 7 】

ネットワーク 2 0 0 の各ノードには M I N T 処理エージェント 2 3 2、2 3 4、2 3 6、2 3 8、2 4 0 および 2 4 2 が接続されている。M I N T 処理エージェントは、ルーティングノードの外部にあるように示されているが、M I N T 処理エージェントは、各ルーティングノード内に組み込まれていてもよい。M I N T 処理エージェントは、M I N T プロトコルをインプリメントするために、それら M I N T 処理エージェントに関連付けられたノードを介して情報を受信し、かつ伝送する。ネットワーク 2 0 0 は、例えば、スパースモード P I M またはコア型ツリーのような I P マルチキャストルーティングプロトコルにおけるマルチキャストグループを形成することができる。このマルチキャストグループは、関連付けられたランデブーポイントまたはノードを有する。

20

【 0 0 1 8 】

図 3 は、本発明に従って構築された M I N T 処理エージェント 3 0 0 の例示的な実施形態を示す。M I N T 処理エージェント 3 0 0 は、図 2 に示される M I N T 処理エージェント 2 3 2 のような M I N T 処理エージェントを示す。M I N T 処理エージェント 3 0 0 は、M I N T プロトコルをインプリメントするためにデータネットワーク 2 0 0 において用いられ得る。M I N T 処理エージェントは、M I N T プロトコルモジュール（P M）3 0 2、データ格納部 3 0 4 およびオプションのパケット転送エンジン 3 0 6 を備える。

30

【 0 0 1 9 】

M I N T - P M 3 0 2 は、リンク 3 0 8 を介してデータネットワークのルーティングノードと接続する。M I N T - P M は、リンク 3 0 8 を用いてルーティングノードと通信し、M I N T チャネルを形成する。この M I N T チャネルによって、データネットワーク内の M I N T 処理エージェントが互いに通信することができる。例えば、M I N T チャネルを用いて M I N T 処理エージェント間、および / または M I N T 処理エージェントとクライアント、情報ソースおよびデータネットワークにおける任意の他のエンドホストとの間で情報を伝送し受信する。データ格納部 3 0 4 は M I N T - P M に接続しており、分散された状態のデータベースを形成するミント（m i n t）情報を格納する。

【 0 0 2 0 】

オプションのパケット転送エンジン 3 0 6 は、M I N T 処理エージェントを用いて M c C a n n e ' 8 6 5 に記載されるようにミント情報に基づいてトラフィックストリームを調整する場合に用いられ得る。パケット転送エンジン 3 0 6 は、ネットワーク 2 0 0 上で伝送されたデータパケット 3 1 0 を受信し、受信されたデータパケットを処理して、ネットワーク上で伝送するための出力データストリーム 3 1 2 を形成する。パケット転送エンジン 3 0 6 は M I N T - P M 3 0 2 およびデータ格納部 3 0 4 に接続し、パケット転送エンジン 3 0 6 がどのようにして受信されたデータパケットを処理するかを決定するために用いられる情報を交換する。例えば、パケット転送エンジン 3 0 6 がデータ格納部 3 0 4 から取り出されたミント情報を用いて、受信されたデータパケット 3 1 0 用にデータネットワーク上での帯域幅割り当てを決定する。別の実施例では、パケット転送エンジン 3 0

40

50



6 がデータ格納部 304 から取り出されたミント情報を用いて、ミント情報に含まれる優先情報に基づいて、出力データストリーム 312 内のパケットをスケジューリングする。

【0021】

MINT 処理エージェント 300 の別の実施形態では、パケット転送エンジン 306 は MINT 処理エージェントから省略されており、付随するネイティブルータ内にあると考えられる。このような実施形態では、MINT 処理エージェントは、データネットワーク内でミントを処理し伝送するために用いられるが、ネットワーク内で伝送されたデータパケットについては何ら処理しない。従って、MINT 処理エージェントは、MINT を処理するタスク、およびデータネットワーク内で MINT の確実な送達を提供するタスクに制限される。

10

【0022】

(MINT プロトコル)

MINT プロトコルは、グループ指向型の確実な情報送達メカニズムをデータネットワーク内のノードのサブセットに提供する。これらノードは、対応するグループを支援するマルチキャストルーティングツリーに広がっている。さらに、エンドホストソースは、MINT 命令を該当するグループのランデブーポイントへと向けることによって、ネットワーク内にデータを発行し得る。MINT プロトコルは、あるメカニズムを提供する。これにより発行された 1 組の値は、メンバーが行き来する際に、広がるツリーにあるアクティブノードに関連付けられた全ての MINT 処理エージェントで維持される。さらなる MINT プロトコルの機能は、任意のネットワーククライアントまたは管理エージェントによるクエリを提供し、発行された最新の 1 組の値を入手する。

20

【0023】

MINT チャンネルは、各アクティブマルチキャストグループに関連付けられる。MINT チャンネルは、複数の MINT 命令を含む MINT アクセスプロトコルに従う TCP を用いた確実なコントロール接続であり得る。アプリケーションは、MINT 命令をランデブーポイントに向けることによって、「ミント」と呼ばれる名前を付けられたデータタブルを MINT チャンネル内へと発行し、次いで、そのランデブーポイントにある MINT - PM は、このようなミントの各々が、そのグループについての分散型ツリーに付随するルーティングノードに関連付けられたすべての MINT 処理エージェントへ伝搬されることを保証する。これによりエッジアプリケーションは、ネットワーク内へ状態を発行し、アプリケーションレベルの処理エージェント（すなわち、プラグイン）と通信することができる。このアプリケーションレベルの処理エージェントは、ネットワーク内にあってもよいし、または MINT 処理エージェントの一部として設けられてもよい。例えば、パケット転送エンジンは、MINT チャンネルを介してエッジアプリケーションと通信し得るアプリケーションレベルの処理エージェントを実行して、ネットワークの帯域幅をエッジアプリケーションに割り当て得る。MINT プロトコルはまた、サブツリーが行き来する際に、動的状態で広がったツリーのパスに沿ってミントを広めるために、明確な通信抽出化を提供する。ルータまたはノードが、ブランチを所与のグループの広がったツリーに接続するときは常に、そのグループのすべてのミントがその新たに作成されたブランチに沿って流れるかまたは伝搬される。その結果、状態は、新たに接続されたブランチに沿って全ての MINT 処理エージェントへと確実に伝搬される。

30

40

【0024】

(データモデル)

MINT プロトコルによって仮定されるデータモデルは、名前タブルまたはミントの固定データ格納部である。オリジンノード（またはオーナー）がネットワーク内にミントを発行し得るか、または MINT アクセス命令を用いて固定データ格納部にその添付を破棄し得る。ノードが失敗するか、またはネットワークから切断される場合には、（該当するグループについて）マルチキャストルーティングツリーの対応するレッグがティヤードウンされると、その発行されたバインディングのすべてをその関連付けられたデータ格納部から削除する。

50

## 【 0 0 2 5 】

ミントが固定であるため、M I N T 処理エージェントは、ネットワーク内に発行されたすべてのミントを維持するためにリソースを使いきってしまう可能性がある。この場合、ミント発行プロセスは失敗する。この失敗をエンドクライアントに通知するために、特別な予備エラーミントがそのグループに添付され、全ての既存のミントよりも高い優先順位を有する。

## 【 0 0 2 6 】

静的優先順位がミントに割り当てられ得る。分散型ツリーのレッグが行き来する際に、この割り当てによって、M I N T 処理エージェント間でミントが伝播されると、ミントの相対的オーダリングが制御される。各ミントは、構造上の階層名によって名前が付けられ、それによってプレフィックス、一般表現、または他の再検討技術によってミントの分類を再検討するためのリッチメカニズムを提供する。M I N T 値が任意の2値データであるのに対して、M I N T データ名は、テキストストリングとして表される。

## 【 0 0 2 7 】

図4は、M I N T 処理エージェント300のデータ格納部304で生じ得る例示的なM I N T 情報400を示す。M I N T 情報400は、任意のタイプの情報を表すミントパラメータのタプルを含む。例えば、ミントパラメータは、データネットワークにおいて伝送される情報を示し得る。M I N T 情報400の特定の実施形態を図4に示すが、当業者には、本発明の範囲を逸脱することなく、他のタイプのデータを表すM I N T 情報の他の実施形態が可能であることが明らかである。

## 【 0 0 2 8 】

M I N T 情報400は、グループ401、オリジン402、名前404、値406および優先順位408を有するミントを含む。データネットワークにおけるノードは1つ以上のマルチキャストグループと関連付けられているので、M I N T 情報は、1つ以上のマルチキャストグループと関連付けられたミントパラメータを含み得る。図4に示されるように、オリジンは、いくつかの名前の付けられた値に関連付けられ得る。M I N T 情報400の特定の実施形態において、名前と値との対によって映像データストリームが説明される。当業者には、名前と値との対が任意のタイプの情報を表し得ることが明らかである。

## 【 0 0 2 9 】

( 名前空間 )

M I N T データ格納部にインデックス付けされた名前は、名前空間を自然に形成する。各グループは、自律名前空間に関連付けられる(すなわち、各グループのミントは全ての他のグループと完全に独立である)。リッチかつ効率的なクエリをこれらの名前空間にわたって支援するために、名前は構造化されているが、なおかつ簡単な形式で表される。特に、名前は階層名前空間を形成し、ちょうどU n i x (登録商標)ファイルシステムが、ディレクトリ名を階層へと並べ替え、「 / 」区切り記号を用いて、ツリー型階層中のパスの相対的な要素を示すのと同様に、階層の境界は、「 / 」区切り記号によって示される。

## 【 0 0 3 0 】

階層名前空間の表現は、マッチングするクエリを既存の名前空間に対して実行させることができる。例えば、広告掲載サービスを構築するために、放送システムは、プレフィックス「 / a d / i n f o 」という形の一連のミントとして、広告情報を発行し得る。従って、クライアントは、「 g l o b b i n g 」マッチ(例えば、「 / a d / i n f o / 」)を用いて、このプレフィックスの形のどの名前が存在するかを調べるためにデータベースにクエリしたいと望むかもしれない。同様に、ネットワークエージェントは、「 / a d / i n f o / 」にマッチする任意のミントが作成されるか、消去されるか、または修正される場合にイベントコールバックが起こり得るように、この情報が変更する場合は常に通知されることを望むかもしれない。

## 【 0 0 3 1 】

( プロトコルモジュール )

ネットワークにおける各ミント処理エージェントは、M I N T プロトコルモジュール ( M

10

20

30

40

50

INT - PM)を含む。このMINTプロトコルモジュールは、データ格納部を維持しており、そのエージェントに既知であるすべての発行されたミントのオリジン、グループおよび名前によってインデックス付けされている。ミントの発行元は、オリジンまたはオーナーと呼ばれる。オリジンがデータネットワーク内のどこかに位置され得る間に、そのグループのランデブーポイントを介して特定のグループ用のミントを発行する必要がある。

#### 【0032】

本発明の1実施形態において、各ミントは、不規則に発行され、リフレッシュプロセスはない。このことが可能であるのは、(その元となるルーティングプロトコルと関連する)MINTプロトコルによってグループ状態は、グループに広がるツリーを介して一貫して維持され得るためである。例えば、いくつかのルーティングプロトコルにおいて失敗が生じると、ルーティングシステムは、ティアードダウンし、次いで必要に応じてグループを再確立することによって応答する。その結果、該当するグループに添付されたミントすべてが、グループが再構築される際に、MINTプロトコルによって伝搬される。従って、グループ状態が一貫して管理され、リフレッシュ/タイムアウトプロセスを必要としない。

#### 【0033】

任意の1つのノードがネットワークへ送ることができるミント状態の量は、構築可能なパラメータによって制限される。任意の所与のグループに関連付けられた各タブルに格納されるデータは以下の要素を含む。

#### 【0034】

##### 【表1】

グループ	マルチキャストグループの名前
オリジン	このエントリを発行したノードの主IPアドレス
名前	このエントリのストリングに基づく名前。各タブルは、そのオリジンに対して一意に定まる名前を有する。
値	名前を付けられたデータの値。これは、任意の2値データであり得、この意味はこのプロトコル仕様の範囲外である。
優先順位 (pri)	他の発行された値と比較した、このアイテムの送達優先順位を表す小さな整数(より高い優先順位値であれば、より低い優先順位値よりも前に送信される傾向にある)

各ノードのデータ格納部は、証明可能に、各グループ用にランデブーポイントで維持されるデータ格納部へと集められるので、シーケンス番号またはタイムスタンプを必要としない。

#### 【0035】

プロトコルメッセージには3つのタイプがある。

- ・発行メッセージ：このメッセージによりMINTが作成され、該当するグループが及ぶ放送ツリーにわたって伝搬されて維持される。
- ・破棄メッセージ：このメッセージによりオリジンに代わってミントバインディングが明示的にティアードダウンされる。
- ・クエリメッセージ：このメッセージによりMINTデータ格納部が名前および値情報を求めてクエリされ得る。

#### 【0036】

これらのメッセージは、ピアMINT - PMと交換され、以下の抽出形式を有する。

【 0 0 3 7 】

【 表 2 】

発行する(G、 名前、値、優先 順位)	グループG用のMINTチャンネルを介してネットワーク内へ名前の付けられた値を発行する。データ値が、Gのためのマルチキャストルーティングツリーに広がるネットワーク内のすべてのノードに確実に送達される。各ノードは、クエリコールによって名前別の値にアクセスすることができる。優先順位は、他の発行された値と比較した、このアイテムの送達優先順位を表す小さな整数である(より高い優先順位値であれば、より低い優先順位値よりも前に送信される傾向にある)。このバインディングは、オリジンが、破棄コマンドによってバインディングを終了するかまたは破棄するまで維持される。値は、状態量が支援可能な制限を越える場合には、うまく発行されないかもしれない。この場合、発行コールが行われた後、エラーが数回非同期に送達される。
破棄(G、名前)	グループG用のMINTチャンネルを介してネットワーク内へ以前に発行された名前の付けられた値を破棄する。
クエリ__名前 (G)	グループG用のネットワーク内へ発行されたデータバインディングの全ての既知の名前(および各オリジン)を返す。
クエリ(G、オリ ジン、名前)	グループG用のネットワーク内へ発行されたMINTの値をクエリする。このキーとなるものは、名前およびオリジンである。値を返す。

10

20

(発行メッセージ)

本発明の信頼性は、グループメンバーシッププロトコルと提携する新規の送達メカニズムに基づく。MINTプロトコルはグループ単位ごとに実行されるので、該当するグループを一般的に「グループG」とよぶ。MINT PMが発行されたミントを受信するピアは、一般的に、ランデブーポイントに戻る最も短い反転パスルートである。パスは元のルーティングプロセスに依存するので、これは常に当てはまるわけではないかもしれない。このピアは「ペアレント」ピアと呼ばれ得る。ちょうどよい時の任意の地点で、グループGの広がったツリーにおける変化とともに生じ得るルーティングの変化の結果として、発行されたミントは、異なるペアレントから受信され得る。これらのミントのすべては、MINT PMに関連したデータ格納部において保持される。

30

【 0 0 3 8 】

グループGなどのグループ内のピアに関連するMINT PMは、ユニキャストの形態でMINTチャンネルを介してお互いにミントを通信する。MINTチャンネルは、信頼性のある接続、例えば、元のルータのピアリング関係に適合する信頼性TCP接続である。

40

【 0 0 3 9 】

MINT PMがそのペアレントのピアに関連するMINT PMから新たなミントを受信する場合、ミントをそのデータ格納部に入力し、ミントのコピーを、Gに対して現在のマルチキャストの広がったツリーに関する他のピアに関連するMINT PMに転送する。例えば、MINT PMは、そのペアレントのピアから発行されたミントを受信し、そのデータ格納部を更新し、次いで、ミントを、ペアレントとして、他のMINT PMに発行する。この動作は、広がったツリーへの任意の変更に対して自動的に実行されることに留意されたい。目標とすることは、グループGに対する広がったツリーに関するピアに関連するすべてのMINT PMが、それぞれのペアレントのデータ格納部に格納された

50

全てのミントを確実に配信することを不変に保持することである。ここで、最終的なペアレントは、そのグループに対するランデブーポイントと関連するMINT PMである。

【0040】

MINT PMがすでにそのテーブル内にあるミントを（そのペアレントから）受信する場合、データの値が異なっているかどうかを知るためにチェックする。もし、異なっていないのであれば、（ネットワーク管理プロトコルを介してアクセス可能な）エラーカウンタをインクリメントする。なぜなら、そのピアは冗長なアップデータを送信しないことが知られているべきであるからである。その値が異なる場合、MINT PMはそのデータ格納部を更新し、その変化を（ペアレントのピアとしてミントに再発行することによって）、Gに対する現在のマルチキャストの広がったツリーに関するチャイルドのピアのそれぞれに伝搬する。結局、MINT - PMは、ペアレントであったかのように、別の発行コマンドをピアMINT PMに発行する。

10

【0041】

MINT PMがグループGに対するペアレントではないピアからミントを受信する場合、そのピアのためのシャドウテーブルにミントアップデートを記録する。そのピアが後にGのペアレントになる場合、このシャドウテーブルは実際のデータ格納部（および変更テーブルが通常のミント到達、変更または削除として取り扱われる間に生じる任意の差異）になる。

【0042】

グループGノードがグラフト（graft）メッセージを受信する場合、要求ノードはグループGにグラフトされ、グループGに関連する全てのミントは要求ノードに関連するMINT PMに送信される。そのミントは、（タプル内の優先順位フィールドにしたがった）静的順位の順番に送信される。すべての既知のミントの集合は、後のミントの到達および他の状態変更に対して自動的に形成される必要がある。ノードがグループG内の別のノードからブーン（prune）メッセージを受信する場合、何かをする必要はなく、ダウンストリームピアがグループGに対するすべてのミントを忘れると推定しなければならない。

20

【0043】

MINT PMがグループGに対するマルチキャストの広がったツリーに関しないピアからミントを受信する場合、アップデートは無視され、エラーカウンタをインクリメントする。ピアがグループに以前に接続されていない場合、ピアはグループGに対するミントを送信することができないので、これはエラー状態である。

30

【0044】

（破棄メッセージ）

グループGに対するミントが（ランデブーポイントに送信された破棄メッセージを介して）オリジンノードによって削除される場合、ランデブーポイントのMINT PMは、データ格納部から対応するミントを取り除き、Gに対する現在のマルチキャストの広がったツリーに関するチャイルドのピアのそれぞれに破棄メッセージを伝搬する。

【0045】

MINT PMはペアレントのピアからのミントに対する破棄メッセージを受信する場合、（オーナおよび名前によってインデックス付けされた）データ格納部を調べる。同じオーナおよび名前を有するタプルが存在する場合、データ格納部から対応するミントを取り除き、Gに対する現在のマルチキャストの広がったツリーに関するチャイルドのピアのそれぞれに破棄メッセージを伝搬する。その名前およびオーナを有するミントが全く存在しない場合、エラーカウンタはエラー状態を示すようにインクリメントされる。破棄メッセージを非ペアレントピアから受信する場合、シャドウテーブルを更新し、非ペアレントピアがペアレントになる場合に使用する。ミントの破棄に関連する任意の事象は、ミントがデータ格納部から除去される場合、ディスパッチされる。

40

【0046】

（トランザクションの実施例）

50

以下の説明は、本発明にしたがったMINTプロトコルを用いた本トランザクションの実施例を示す。このトランザクションの実施例は、ネットワークグループがルーティングプロトコルによって形成され得ると想定する。ルーティングプロトコルは、ランデブーポイント(RP)を使用して、グループに対するアンカーとして機能する。一般的なネットワーク動作の間、各ルーティングノードは、特定のグループおよびその関連RPに直接または間接にアクセスし得る。

【0047】

図5は、本発明によるデータネットワーク200および関連するMINT処理エージェントを示す。ネットワーク200において、グループに対するRPとしてルーティングノード206を有するマルチキャストグループAが形成される。グループAは、ルーティングノード202、206、208および204を含む。ルーティングノード202、206、208および204は、これらのノードのダウストリームに位置する他のグループAメンバーに情報をルーティングする。本発明は、単一のドメイン内に形成されるグループで、または、グループA内のドメインXのノード204を内包することによって示されるように、複数のドメイン内のノードから形成されるグループで使用されるのに適している。

10

【0048】

図6は、本発明にしたがって、図5のネットワーク200の動作の方法600に対するブロック図を示す。ブロック602において、方法600は、図5に示されるように、ノード206に位置付けられるグループAに対するRPを有するグループAの形成から開始する。

20

【0049】

ブロック604において、情報ソース220は、グループAにミントを発行する。例えば、ある実施形態において、情報ソース220は、グループA、ノード202に発行コマンド(ミント情報を含む)を伝送する。元のルーティングプロトコルの一部として、ノード202は、ノード206がグループAに対するRPであることに気づく。結果として、ノード202は、最終的に受信されるRPの方に発行命令をルーティングする。別の実施形態において、情報ソース202は、ディレクトリサービスを用いて、例えば、グループAに対するRPの位置を決定するために、ネットワークにクエリし得る。RPの位置を知ると、情報ソースは、発行コマンドを直接RPに伝送し得る。トランザクションパス502は、ソース220からRPまでの発行コマンドのルーティングを示す。

30

【0050】

ブロック606において、RPは、トランザクションパス504によって示されるように、MINT処理エージェント236に転送される発行コマンドを受信する。

【0051】

ブロック608において、MINT処理エージェント236は、そのデータ格納部を新たなミント情報で更新する。この動作は、MINT PM302がMINTチャネル308を介して発行ミントを受信した場合に発生し、そのミント情報を用いて、関連するデータ格納部304を更新する。

40

【0052】

ブロック610において、更新されたミント情報を、グループAの他のMINT処理エージェント、すなわち、エージェント232、234および240に伝搬する。これを達成するために、RPに関連するMINT PMは、新たなミント情報を、MINTチャネルを介して、グループA内の他のMINT処理エージェントに配信する。MINT処理エージェント236は新たなMINT情報をグループAに発行し、元のルーティングプロトコルによって決定されるように、そのアップデートはグループAのルーティングに続く。例えば、トランザクションパス506は、MINT処理エージェント236からの新たに発行されたミント情報の、グループA内の他のMINT処理エージェントへのルーティングを示す。上述するようなミント配信規則にしたがって、一貫性が保持される。結果として、情報ソース220によって発行されたミント情報が、グループA内のMINT処理エー

50

ジェントのすべてに確実に配信され得る。

【 0 0 5 3 】

図 7 は、グループ A のメンバーシップへの変化が示される、図 5 のネットワーク 2 0 0 を示す。例えば、グループ A に対するメンバーシップが、線 7 5 0 によって指定されるグループから線 7 5 2 によって指定されるグループに増加し、ノード 2 1 2 を含むことに留意されたい。

【 0 0 5 4 】

図 8 は、本発明に従って、図 7 のネットワークの操作の方法 8 0 0 に関するブロック図を示す。方法 8 0 0 において、クライアント 2 2 6 はグループ A に接続することを要求し、結果として、グループ A はノード 2 1 2 および M I N T 処理エージェント 2 4 2 を含むように拡張される。方法 8 0 0 は、ミント情報が新しいグループメンバーと適合するように、どのように確実に伝搬するかを説明する。

【 0 0 5 5 】

ブロック 8 0 2 において、グループ A の現在のメンバーシップは、図 7 の 7 5 0 に示されるようにノード 2 0 2、2 0 4、2 0 6 および 2 0 8 を含む。ブロック 8 0 4 では、クライアント 2 2 8 は、グループ A に接続するためにノード 2 1 2 に要求を伝送する。クライアント 2 2 8 は、グループ A にわたって現在マルチキャストされている情報を受信しようとし得る。ブロック 8 0 6 では、ノード 2 1 2 は、接続要求をノード 2 0 6 へ転送する。このノード 2 0 6 は既にグループ A のメンバーである。接続要求はトランザクションパス 7 0 2 により示される。ノード 2 1 2 がノード 2 1 2 からの接続要求を受信した場合、ノード 2 1 2 はグループ A の広がったツリー ( s p a n n i n g t r e e ) に含まれ、その結果グループ A メンバーシップは 7 5 2 で示される。ノード 2 1 2 と関連する M I N T 処理エージェント 2 4 2 はまた、グループ A のメンバーであり得る。

【 0 0 5 6 】

ブロック 8 0 8 では、ノード 2 0 6 は M I N T 処理エージェント 2 3 6 にノード 2 1 2 ( およびその関連した M I T 処理エージェント 2 4 2 ) がグループ A に接続したことを通知する。これはトランザクションパス 7 0 4 で示される。ブロック 8 1 0 では、M I N T 処理エージェント 2 3 6 は、グループ A に関するミントを、その M I N T データ格納部から新しく追加された M I N T 処理エージェント 2 4 2 へ伝搬する。このミントは、エージェント 2 3 6 の M I N T - P M がエージェント 2 4 2 の M I N T - P M にミント情報を発行すると、M I N T チャンネルにわたって伝搬する。これは、トランザクションパス 7 0 6 で示される。ブロック 8 1 2 では、M I N T 処理エージェント 2 4 2 は、新しい M I N T 情報を有するそのデータ格納部をアップデートし、その結果グループ A のすべての M I N T 処理エージェントは同一のグループ A の M I N T 情報を有する。M I N T エージェント 2 3 6 は M I N T エージェント 2 4 2 のペアレントピア ( p a r e n t p e e r ) であることに留意されたい。M I N T 処理エージェント 2 4 2 が他の新たに添付された M I N T 処理エージェントに対するペアレントピアであったならば、それらの他の M I N T 処理エージェントに新たなミント情報を再発行する。

【 0 0 5 7 】

上記の例では、一つのノードのみがグループ A に加えられ、次いで、これはミント情報で引き続いてアップデートされる。数個のノードがグループに追加される場合においては、ミント伝搬は接続要求の結果として確立されたルートに従い得る。例えば、一つの実施形態では、ミント情報は ( 接続要求と比較して ) 逆方向に伝搬する。これは、接続要求を受信したグループのノードからホップバイホップ ( h o p b y h o p ) で始まって、接続リクエストに戻る。逆のホップバイホップルートの各 M I N T 処理エージェントは、グループの広がったツリーの新しいブランチと関連したすべての M I N T 処理エージェントがアップデートされるまで、アップデートされる。

【 0 0 5 8 】

各データ格納部内でグループ A に関する同一のミント情報を有するすべての M I N T 処理エージェントを用いると、ソース 2 2 0 はグループ A にアップデートされたミント情報を

10

20

30

40

50

発行しようとする。ブロック 8 1 4 では、ソース 2 2 0 は発行コマンドを、ノード 2 0 2 を介して R P へ伝送する。これは、トランザクションパス 7 0 8 で示される。先のように、ソースは複数の方法のうちの一つを使用して発行コマンドをグループ A の R P に伝送し得る。

【 0 0 5 9 】

ブロック 8 1 6 では、R P はソース 2 2 0 から発行コマンドを受信する。ブロック 8 1 8 では、R P は M I N T 処理エージェント 2 3 6 に発行コマンド（パス 7 0 8 ）を通知し、そして M I N T 処理エージェント 2 3 6 はミントを受信し、新しい発行コマンドのミントに基づいたそのデータ格納部をアップデートする。ブロック 8 2 0 では、M I N T 処理エージェント 2 3 6 は、（ペアレントとして）新しいミント情報をグループ A に関連したすべてのピア M I N T 処理エージェント（チャイルドピア）へ伝搬する。これが起こり得る一つの様式は、トランザクションパス 7 1 0 で示されるように、M I N T 処理エージェント 2 3 6 がグループ A の他のメンバーに M I N T チャネルにわたって発行コマンドを発行する場合である。結果として、新しいミント情報はノード 2 0 2、2 0 4、2 0 8 および 2 1 2 に確実に伝搬される。これらのノードはグループ A のすべての部分であり、エージェント 2 3 6 に対するチャイルドピアである。

【 0 0 6 0 】

この例では、エージェント 2 3 6 によって発行された新しいミント情報は、図 7 に示されるように、チャイルドピアに到達するように一つのホップを伝搬する必要があるだけである。しかし、チャイルドピアはグループ A の他のノードに（ペアレントとして）ミント情報を再発行し得ることが当業者には明らかである。従って、グループ A の広がったツリーが相互に接続したノードの大きなブランチを含むならば、新しいミント情報はグループ A に関連したすべてのノード（および、M I N T 処理エージェント）に、広がったツリーの下に（ペアレントからチャイルドまで）ホップバイホップで伝搬する。

【 0 0 6 1 】

ブロック 8 2 2 では、M I N T 処理エージェント 2 3 2、2 3 4、2 4 0 および 2 4 2 はすべて、新しいミント情報を受信し、そしてその新しいミント情報を有するそれらの関連したデータ格納部をアップデートする。従って、本発明によれば、新しく発行されたミント情報は、グループ A の広がったツリーのアクティブなノードに関連したすべての M I N T 処理エージェントに確実に配信される。

【 0 0 6 2 】

方法 8 0 0 はまた、本発明に従って、ミント情報がどのようにクエリされ得るかを示す。

【 0 0 6 3 】

ブロック 8 2 4 では、クライアント 2 2 8 はグループ A に関連したミント情報をクエリすることを望む。クライアント 2 2 8 はクエリ命令を目的のグループとしてグループ A を特定するノード 2 1 2 に伝送する。用いられるクエリのタイプは、グループ A のネットワーク中に発行されたデータバインディングのすべての既知の名前（および、それぞれのオリジン）を返す。例えば、上記の名前ベースのクエリ命令（`query - name (A)`）は、この情報を返す。

【 0 0 6 4 】

ブロック 8 2 6 では、M I N T 処理エージェント 2 4 2 はクエリ命令を受信する。これは、トランザクションパス 7 1 2 で示される。

【 0 0 6 5 】

ブロック 8 2 8 では、M I N T 処理エージェント 2 4 2 は、トランザクションパス 7 1 4 により示されるように、クエリの結果をクライアント 2 2 8 に伝送することにより、要求されたミント情報に回答する。これは、エージェント 2 4 2 での M I N T - P M がその関連したミントデータ格納部から要求された情報を取り出し、M I N T チャネルにわたる結果をクライアント 2 2 8 に伝送する場合に発生する。ブロック 8 3 0 では、クライアント 2 2 8 は要求されたミント情報を受信し、そして結果として、クライアント 2 2 8 は返されたミント情報を使用してグループ A の状態を判定し得るか、またはグループ A において



伝送されたデータストリームを受信するアクションをとり得る。

【0066】

図9は、クライアント228がグループAへのその接続を終結し、そのMINTバインディングを破棄する方法900のブロック図を示す。この方法900は、図10および図11を参照して説明される。

【0067】

図10および11は、ネットワーク200および方法900の操作時に起こる関連のトランザクションを示す。

【0068】

図9を再度参照すると、ブロック902では、クライアント228はグループAに添付され、これは、図10に示されるように、ノードおよびMINT処理エージェントから構成される。ブロック904では、クライアント228は、クライアント228からノード212までリープ要求(leave request)を伝送することにより、グループAからの終結を要求する。このリープ要求のルートは1002で示される。

【0069】

ブロック906では、ノード212はエージェント242に、クライアント228はグループAからそのメンバーシップを終結していることを通知し、従ってノード212はグループAから取り除かれる。ブロック908では、ノード212がグループAから取り除かれるので、エージェント242はグループAに関連するミントを破棄する。しかし、ノード212が他のグループのメンバーである場合は、それらの他のグループに関連するミントはエージェント242により維持されることに注目すべきである。他の実施形態では、エージェント242は、以下のセクションで説明するように、本発明の他の局面に従って、グループを離れた後にミントを維持し得る。

【0070】

ブロック910では、ノード212は、それが最終的に受信されるRP(ノード206)の方へ、リープ要求を伝搬する。RPはエージェント236にトランザクションパス1004で示されるような(クライアント228による)リープ要求を通知する。ブロック912では、エージェント236は、それがグループのRPと関連するために、それをグループAのミントの格納部を維持する。グループAが存在する限り、エージェント236は、ミントを他のグループメンバーに伝搬することをエージェントが要求する場合において、ミントのそのデータ格納部を維持する。

【0071】

ブロック914では、RP(ノード206)はクライアント228からのリープ要求を処理し、結果としてノード212はグループAから取り除かれる。これが起こった後、生じたグループは図11のグループAで示されるように、ノード202、206および208を含む。

【0072】

ここで、図11を参照すると、グループAから取り除かれたノード212では、方法900は以下に説明する追加のトランザクションで継続する。

【0073】

ブロック916では、情報ソース220は、グループAに関する新たなミントを発行する。ノード202は発行コマンドを受信し、そしてそれをRPの方にルート付けする。上記のように、情報ソースはRPの位置を見つけて、発行コマンドを直接RPに発行することができる。あるいは、潜在的なグループルーティングプロトコルの結果として、ノード202はRPの位置を知ることができ、そしてそれゆえに、RPの方へ発行コマンドをルート付けする。このトランザクションはトランザクションパス1102に示される。

【0074】

ブロック918では、トランザクションパス1104で示されるように、RPは発行コマンドを受信し、発行されたミントをMINT処理エージェント236に転送する。ブロック920では、MINT処理エージェント236は、新たなミント情報によりそのデータ

10

20

30

40

50

格納部をアップデートする。ブロック 9 2 2 では、M I N T 処理エージェントは新たなミント情報をグループ A の他の M I N T 処理エージェント（すなわち、エージェント 2 3 2、2 3 4 および 2 4 0）に伝搬する。これはトランザクションパス 1 1 0 6 で示される。ミントの伝搬は、エージェント 2 3 6 が新たなミント情報を有する発行コマンドをグループ A の他のノードに発行するとき起こる。

【 0 0 7 5 】

グループ A へのその添付を終結するクライアント 2 2 8 の結果として、そして結果的に、ノード 2 1 2 はグループ A の広がったツリーから取り除かれ、M I N T 処理エージェント 2 4 2 は、もはやグループ A に関する新たなミント情報でアップデートされない。しかし、M I N T プロトコルは、グループ A のアクティブなメンバーである M I N T 処理エージェントのためのミントデータ格納部を確実にアップデートし続ける。ノード 2 1 2 が将来、グループ A を接続することを要求すれば、アップデートされたミントは再度ノード 2 1 2 へ伝搬され、その後 M I N T 処理エージェント 2 4 2 に伝搬される。

【 0 0 7 6 】

別の実施形態では、M I N T プロトコルは過剰なルーティング変動と関連する問題を克服するように作動する。過剰なルーティング変動の間、特定のノードが反復して離れ、次いでグループに再接続する場合に、それらのノードに関連したデータ格納部中のミント情報は繰り返し破棄され、再配置（*repopulate*）される。これは、データネットワーク上のミント情報の過剰なトランザクションをもたらす。この問題を回避するために、M I N T プロトコルに対する拡張は、過剰なルーティング変化の結果として、データ格納部の破棄および再配置を回避する。

【 0 0 7 7 】

一つの実施形態では、M I N T ダイジェスト（*digest*）は、データ格納部中のミントにわたって計算される。M I N T ダイジェストは、データ格納部中のすべてのミントを表してもよいし、あるいはデータ格納部中のミントの選択された部分を表してもよい。ノードがグループを離れたときにミント情報を破棄する代わりに、そのノードに関連したミント情報はその関連の M I N T ダイジェストとともにデータ格納部に保存される。ノードがグループに再接続すると、それはその M I N T ダイジェストをグループに伝送する。ノードにおける M I N T ダイジェストがグループの現在の M I N T ダイジェストとは異なると、ノードはミント情報の新たなコピーとともにアップデートされる。次いで、ノードはそのミントデータ格納部およびその関連のダイジェストをアップデートする。ノードからの M I N T ダイジェストがグループからの M I N T ダイジェストと適合すると、ミント情報の新たなコピーをノードに伝送する必要はない。それゆえ、拡張された M I N T プロトコルは、ネットワーク中のミント情報の過剰な伝送を回避する。

【 0 0 7 8 】

別の実施形態では、データ格納部のリソースが古いミント情報を格納するために用いられることがないように、時間パラメータが使用される。ノードがグループを離れると、そのノードに関連した M I N T 処理エージェントは時間パラメータを使用して、データ格納部にミント情報をどれほど長く保存するかを決定する。この時間パラメータ値はネットワーク管理者により決定され得る。時間パラメータにより規定された期間中、データ格納部およびその関連の M I N T ダイジェストを保存することにより、ミント情報の過剰な伝達は、上記のように防止され得る。しかし、ノードがグループを離れ、時間パラメータにより規定された期間が満了すると、ミントデータ格納部はそのグループのミントから取り除かれ得、それによりデータ格納物のリソースを無くす（*free up*）。それゆえ、M I N T 処理エージェントはデータ格納部を保存して、ネットワークフラッピング間の冗長なミント伝達を防止する。そして選択された期間の満了後、データ格納部を消去して、さらなるミントを格納する価値あるリソースを解放する。

【 0 0 7 9 】

当業者に明らかなように、M I N T プロトコルをインプリメントするための上記の方法および装置における改変は、本発明の範囲から逸脱することなく可能である。従って、本明

10

20

30

40

50

細書中の開示および記載は、提示であり、限定されるものではなく、特許請求の範囲に示される本発明の範囲内である。

【図面の簡単な説明】

【図 1 A】 図 1 A は、本発明を使用するのに適したコンピュータを示す。

【図 1 B】 図 1 B は、図 1 A のコンピュータのサブシステムを示す。

【図 2】 図 2 は、本発明にしたがって構成されたデータネットワークを示す。

【図 3】 図 3 は、本発明にしたがって構成された M I N T 処理エージェント 3 0 0 のブロック図を示す。

【図 4】 図 4 は、例示的な M I N T 情報を示す。

【図 5】 図 5 は、図 2 のデータネットワークに規定されたマルチキャストグループを示す。

10

【図 6】 図 6 は、図 5 に示されるネットワークを介して M I N T 情報を伝搬する方法のブロック図を示す。

【図 7】 図 7 は、接続ノードを含むように改変された改変マルチキャストグループを示す。

【図 8】 図 8 は、図 7 の改変マルチキャストグループを介して M I N T 情報を伝搬する方法のブロック図を示す。

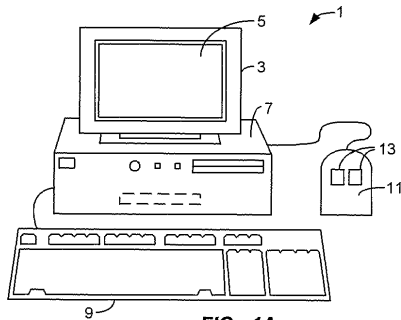
【図 9】 図 9 は、マルチキャストグループ内のノードを解放し、その関連する M I N T バインディングを終端させる方法のブロック図を示す。

【図 1 0】 図 1 0 は、図 9 に示される方法の動作の間のトランザクションステップをさらに示すための図である。

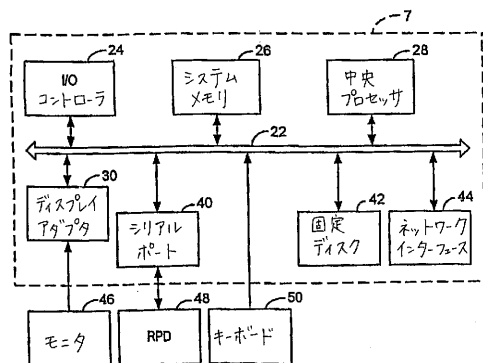
20

【図 1 1】 図 1 1 は、図 9 に示される方法の動作の間のトランザクションステップをさらに示すための図である。

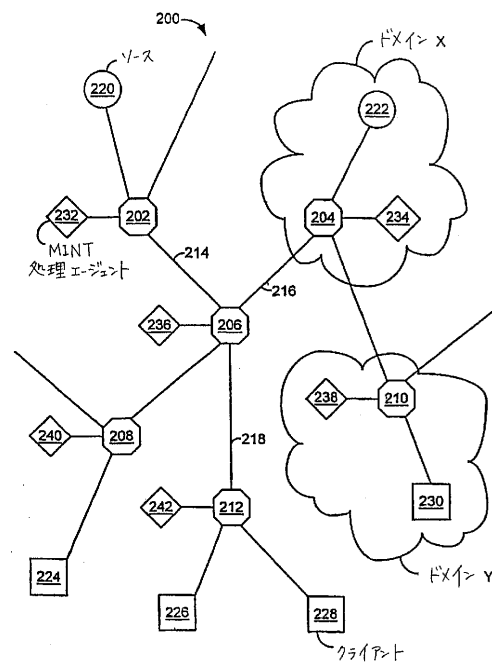
【図 1 A】



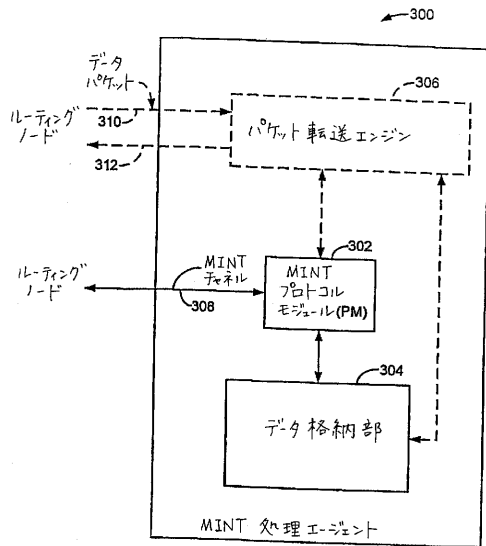
【図 1 B】



【図 2】



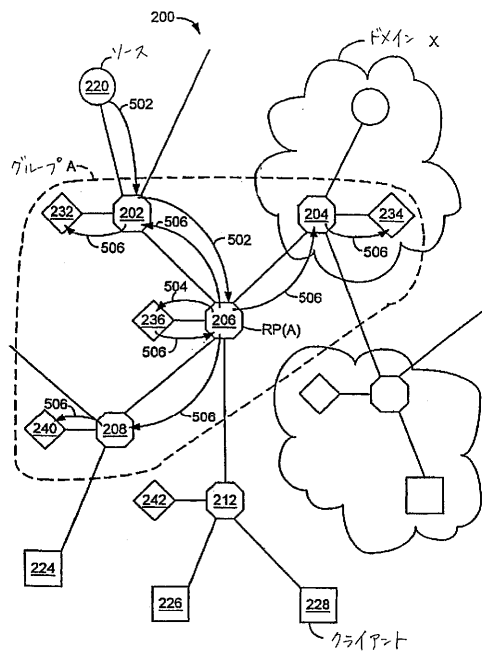
【図 3】



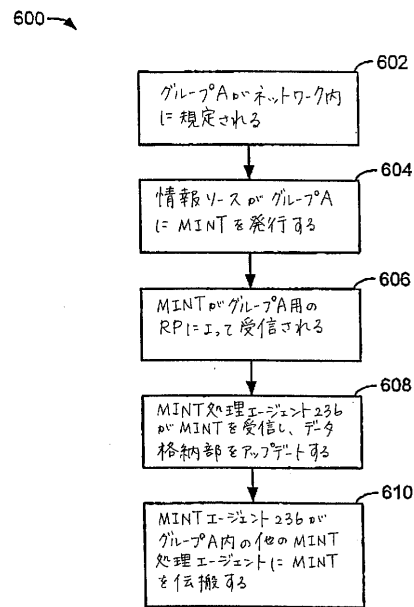
【図 4】

優先順位	値	名前	オリジン (オナ)	グループ
1	NETSHOW	MEDIATYPE	10.1.1.8	G
3	THOMPSON	AUTHOR		
2	FEAR AND LOATHING	TITLE		
3	G	RATING		
1	MEDIATYPE	MEDIATYPE	10.11.77.132	
3	OBIS SPORTS	AUTHOR		
2	PENGUINS V DUCKS	TITLE		
5	2:30:00	DURATION		

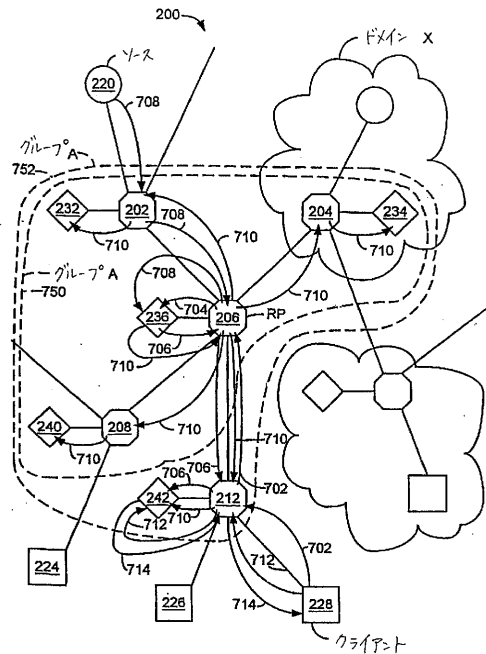
【図 5】



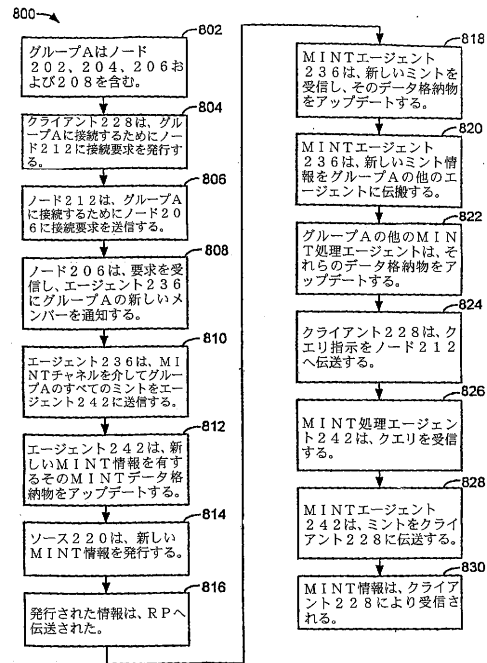
【図 6】



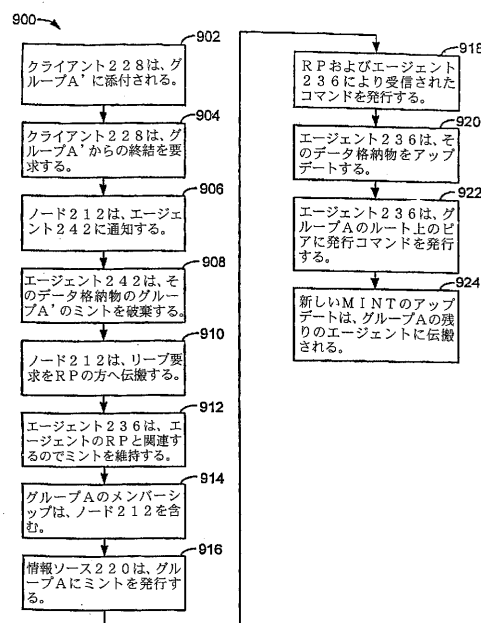
【 図 7 】



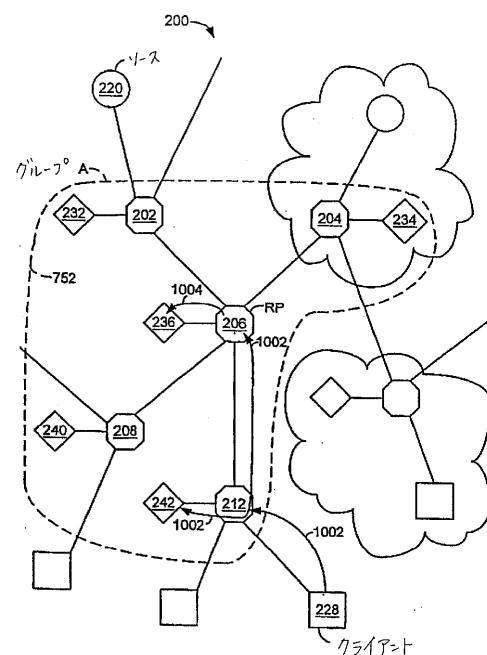
【 図 8 】



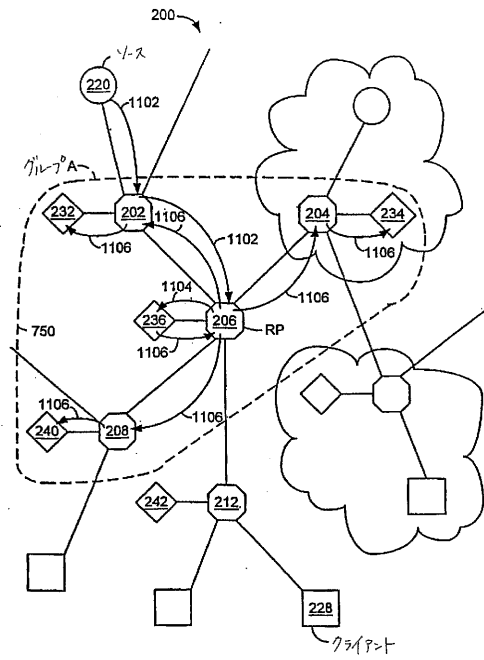
【圖 9】



【 図 1 0 】



【図 11】



---

フロントページの続き

(74)代理人 100120525

弁理士 近藤 直樹

(72)発明者 マッカンネ, スティーブン

アメリカ合衆国 カリフォルニア 94707, バークレイ, サン ロレンゾ アベニュー  
1571

(72)発明者 スワン, アンドリュー

アメリカ合衆国 カリフォルニア 94702, バークレイ, シーダー ストリート 135  
0

審査官 松崎 孝大

(56)参考文献 特開平3-273727(JP,A)

特開平9-200269(JP,A)

特開平10-93573(JP,A)

特開2001-24700(JP,A)

尾上 裕子、藤井 敬三、徳田 英幸、マルチキャストサーバにおける動的QOS制御、情報処理学会研究報告 Vol.94 No.56, 社団法人情報処理学会, 1994年 7月 8日, p.1-6

大出 富康, 1997 パソコンTCP/IP最前線, Software Design, 株式会社 技術評論社, 1997年 6月18日, p.42-53

(58)調査した分野(Int.Cl., DB名)

H04L 12/56