

(19) 日本国特許庁(JP)

(12) 特 許 公 報(B2)

(11) 特許番号

特許第6172296号
(P6172296)

(45) 発行日 平成29年8月2日(2017.8.2)

(24) 登録日 平成29年7月14日(2017.7.14)

(51) Int.Cl.

F I

H O 4 L 12/46 (2006.01)

H O 4 L 12/46 E

H O 4 L 29/08 (2006.01)

H O 4 L 13/00 3 O 7 Z

G O 6 F 13/00 (2006.01)

H O 4 L 12/46 2 O O X

G O 6 F 13/00 5 2 O A

請求項の数 10 (全 24 頁)

(21) 出願番号 特願2015-557635 (P2015-557635)
 (86) (22) 出願日 平成26年1月16日(2014.1.16)
 (86) 国際出願番号 PCT/JP2014/050687
 (87) 国際公開番号 W02015/107653
 (87) 国際公開日 平成27年7月23日(2015.7.23)
 審査請求日 平成28年6月23日(2016.6.23)

(73) 特許権者 000005223
 富士通株式会社
 神奈川県川崎市中原区上小田中4丁目1番
 1号
 (74) 代理人 100074099
 弁理士 大菅 義之
 (74) 代理人 100133570
 弁理士 ▲徳▼永 民雄
 (72) 発明者 佐沢 真一
 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
 審査官 森谷 哲朗

最終頁に続く

(54) 【発明の名称】 通信装置、通信方法、および、通信プログラム

(57) 【特許請求の範囲】

【請求項 1】

データを前記データの宛先に送信した回数を、前記データを識別する識別子に対応付けて記憶する記憶部と、

前記宛先へ送信する対象のデータである対象データの送信回数が閾値に達すると、前記対象データの代替として前記宛先に送信する代替データを、前記対象データと前記対象データの保持を要求する保持要求の組み合わせに設定し、前記対象データに対する保持要求の送信後は、前記対象データの識別子を前記代替データに設定する設定部と、

前記代替データを前記宛先に送信する送信部

を備えることを特徴とする通信装置。

10

【請求項 2】

前記設定部は、前記対象データの送信回数が前記閾値未満の間、前記代替データを、前記対象データと前記対象データの保持を禁止する情報の組み合わせに設定する

ことを特徴とする請求項 1 に記載の通信装置。

【請求項 3】

前記通信装置に向けて送信されたデータを受信する受信部と、

受信データのうちで、前記保持要求が付加されたデータを保持する受信キャッシュと、

前記通信装置にデータを送信する装置である送信元から受信するデータのトラフィックレートを算出する算出部と、

前記送信元が前記保持要求を付加するまでに前記通信装置に同じデータを送信できる回

20

数の上限値の指標として前記送信元が使用する指標値を、前記トラフィックレートが低いほど小さな値になるように決定する決定部を備え、

前記送信部は、前記指標値を含むパケットを前記送信元に送信することを特徴とする請求項 1 または 2 に記載の通信装置。

【請求項 4】

第 1 の通信装置と第 2 の通信装置を含むネットワークにおいて、

前記第 1 の通信装置は、

前記第 2 の通信装置に送信するデータを識別する識別子を生成し、

前記第 2 の通信装置へ送信する対象のデータである対象データの送信回数が閾値に達すると、前記対象データの代替として前記第 2 の通信装置に送信する代替データを、前記対象データと前記対象データの保持を要求する保持要求との組み合わせに設定し、

前記第 2 の通信装置は、前記保持要求と共に受信した前記対象データを前記第 2 の通信装置が保持する受信キャッシュに保持し、

前記第 1 の通信装置は、前記保持要求の送信後は、前記対象データの識別子を前記代替データに設定し、

前記第 2 の通信装置は、前記識別子を受信すると、前記識別子に対応付けられたデータを、前記第 1 の通信装置から受信したデータとして扱う

ことを特徴とする通信方法。

【請求項 5】

前記第 1 の通信装置は、前記第 2 の通信装置への前記対象データの送信回数が前記閾値未満の間、前記代替データを、前記対象データと前記対象データの保持を禁止する禁止情報の組み合わせに設定し、

前記第 2 の通信装置は、前記禁止情報と共に受信した対象データを、前記第 1 の通信装置からの受信データとして処理する際に、前記受信キャッシュに格納しない

ことを特徴とする請求項 4 に記載の通信方法。

【請求項 6】

前記第 2 の通信装置は、

前記第 1 の通信装置から受信するデータのトラフィックレートを算出し、

前記閾値の値を、前記トラフィックレートが低いほど小さな値になるように決定し、

前記第 1 の通信装置に、前記閾値として決定した値を通知し、

前記第 1 の通信装置は、前記第 2 の通信装置から通知された値を前記閾値として用いて、前記保持要求を付してデータを前記第 2 の通信装置に送信するかを判定する

ことを特徴とする請求項 4 または 5 に記載の通信方法。

【請求項 7】

前記第 1 の通信装置は、第 1 の閾値を用いて前記第 2 の通信装置に送信するデータに前記保持要求を含めるかを判定し、

第 3 の通信装置は、第 2 の閾値を用いて前記第 2 の通信装置に送信するデータに前記保持要求を含めるかを判定し、

前記第 2 の通信装置は、

前記第 1 の通信装置から受信するデータのトラフィックレートである第 1 のレートと、前記第 3 の通信装置から受信するデータのトラフィックレートである第 2 のレートを算出し、

前記第 2 のレートが前記第 1 のレートより高い場合、前記第 2 の閾値の値が前記第 1 の閾値よりも大きくなるように、前記第 1 および第 2 の閾値の更新値を決定し、

前記第 1 の通信装置は、前記第 2 の通信装置から通知された前記第 1 の閾値の更新値を用いて、前記第 2 の通信装置に送信するデータに前記保持要求を含めるかを判定し、

前記第 3 の通信装置は、前記第 2 の通信装置から通知された前記第 2 の閾値の更新値を用いて、前記第 2 の通信装置に送信するデータに前記保持要求を含めるかを判定する

ことを特徴とする請求項 4 または 5 に記載の通信方法。

【請求項 8】

データを前記データの宛先に送信した回数を、前記データを識別する識別子に対応付けて記憶し、

前記宛先へ送信する対象のデータである対象データの送信回数が閾値に達すると、前記対象データの代替として前記宛先に送信する代替データを、前記対象データと前記対象データの保持を要求する保持要求の組み合わせに設定し、

前記対象データに対する保持要求の送信後は、前記対象データの識別子を前記代替データに設定し、

前記代替データを前記宛先に送信する

処理を通信装置に行わせることを特徴とする通信プログラム。

【請求項 9】

前記対象データの送信回数が前記閾値未満の間、前記代替データを、前記対象データと前記対象データの保持を禁止する情報の組み合わせに設定する

処理を、前記通信装置に行わせることを特徴とする請求項 8 に記載の通信プログラム。

【請求項 10】

前記通信装置に向けて送信されたデータを受信し、

前記通信装置にデータを送信する装置である送信元から受信するデータのトラフィックレートを算出し、

前記送信元が前記保持要求を付加するまでに前記通信装置に同じデータを送信できる回数の上限値の指標として前記送信元が使用する指標値を、前記トラフィックレートが低いほど小さな値になるように決定し、

前記指標値を含むパケットを前記送信元に送信する

処理を、前記通信装置に行わせることを特徴とする請求項 8 または 9 に記載の通信プログラム。

【発明の詳細な説明】

【技術分野】

【0001】

本発明は、複数の通信装置の間で行われる通信に関する。

【背景技術】

【0002】

近年、通信の高速化が求められてきており、通信速度を高速化するために様々な試みが行われてきている。例えば、データの送信量を減らすために、送信対象のデータに対して重複除去が行われることがある。重複除去では、受信側の通信装置が、過去に送信側の通信装置から受信したデータを、識別子に対応付けてキャッシュに保存しており、送信側の通信装置も過去に送信したデータの識別子を記憶している。送信側の通信装置は、送信しようとするデータがすでに送信したことがあるデータであるかを、データの識別子を用いて判定し、すでに送信したことがあるデータについては、データの代わりに、そのデータに対応付けられた識別子を受信側の通信装置に送信する。受信側の通信装置は、受信した識別子に対応付けられたデータをキャッシュから読み出して、読み出したデータを送信側の通信装置から送信されたデータとして扱う。

【0003】

重複するデータの処理に関連して、同一のイメージデータを処理する場合、プリンタドライバはその旨を示す情報をデータに付加してプリンタに転送し、プリンタがイメージデータをキャッシュする方法が提案されている（例えば、特許文献 1）。

【先行技術文献】

【特許文献】

【0004】

【特許文献 1】特開 2002 - 91731 号公報

【発明の概要】

【発明が解決しようとする課題】

【0005】

10

20

30

40

50

キャッシュを用いて送信データ中の重複を除去する場合、受信側の通信装置が備えているキャッシュのサイズが大きいほど多くのデータがキャッシュに格納されるため、重複したデータの送信を防ぎやすくなる。しかし、受信側の通信装置が備えるキャッシュの容量は有限であるため、受信側の装置は、キャッシュの空き容量がなくなると、最終アクセスからの経過時間が長いデータから順に削除することになる。このため、受信側でのデータの削除により、効率的に重複除去ができない場合がある。例えば、送信対象のデータ中に、データAが複数回含まれていたとしても、先に送信されるデータAと後に送信されるデータAの間に、重複していないデータが大量に送信される場合がある。このとき、後から送信されるデータAの送信前に受信側の通信装置のキャッシュの空き容量がなくなり、データAが受信側の通信装置で削除されると、送信側の通信装置は、データAを受信側の通信装置に再度送信することになる。

10

【0006】

本発明は、1つの側面として、重複部分を含むデータの転送を効率的に防止し、通信を高速化することを目的とする。

【課題を解決するための手段】

【0007】

実施形態に係る通信装置は、記憶部、設定部、送信部を備える。記憶部は、データを前記データの宛先に送信した回数を、前記データを識別する識別子に対応付けて記憶する。設定部は、前記宛先へ送信する対象のデータである対象データの送信回数が閾値に達すると、前記対象データの代替として前記宛先に送信する代替データを、前記対象データと前記対象データの保持を要求する保持要求の組み合わせに設定する。設定部は、前記対象データに対する保持要求の送信後は、前記対象データの識別子を前記代替データに設定する。送信部は、前記代替データを前記宛先に送信する。

20

【発明の効果】

【0008】

重複部分を含むデータの転送が効率的に防止され、通信が高速化される。

【図面の簡単な説明】

【0009】

【図1】実施形態に係る通信の例を説明する図である。

【図2】実施形態にかかる通信装置が用いられるネットワークの例を示す。

30

【図3】通信装置の構成の例を示す図である。

【図4】通信装置のハードウェア構成の例を示す図である。

【図5】重複検知部の処理の例を説明する図である。

【図6】送信データテーブルの例を示す図である

【図7】送信処理の例を説明する図である。

【図8】受信キャッシュでのデータの保存方法の例とインデックス値の決定方法の例を示す図である。

【図9】通信処理の例を説明する図である。

【図10】データを送信する通信装置の処理の例を説明するフローチャートである。

【図11】データを受信する通信装置の処理の例を説明するフローチャートである。

40

【図12】第2の実施形態にかかる通信装置の構成の例を示す図である。

【図13】閾値の決定方法の例を説明する図である。

【図14】データを受信する通信装置の処理の例を説明するフローチャートである。

【発明を実施するための形態】

【0010】

図1は、実施形態に係る通信の例を示す。以下、送信側の通信装置を通信装置10a、受信側の通信装置を通信装置10bとする。さらに、通信装置10bが備えている受信キャッシュを、受信キャッシュ32bとする。

【0011】

通信装置10aは、通信装置10bにデータを送信するとき、送信対象のデータを一意

50

に識別できる識別子を生成する。通信装置 10 a は、送信対象のデータを通信装置 10 b に送信するときに、そのデータを受信側の通信装置に送信した回数を、識別子に対応付けて記憶するものとする。図 1 の例では、通信装置 10 a と通信装置 10 b の通信が開始されたところであり、通信装置 10 a から、まだデータの送信が行われていないものとする。通信装置 10 a は、予め閾値 N を記憶しており、同じデータの送信回数が N 回未満の間は、送信するデータのキャッシュへの保持を禁止する情報を付して、通信装置 10 b にデータを送信する。以下、送信するデータのキャッシュへの保持を禁止する情報のことを「禁止情報」と記載することがある。通信装置 10 b は、禁止情報と共に受信したデータは受信キャッシュ 32 b に格納しない。

【0012】

一方、同じデータの送信回数が閾値 N と一致すると、通信装置 10 a は、送信するデータのキャッシュへの保持を要求する情報を付して、通信装置 10 b にデータを送信する。以下、送信するデータのキャッシュへの保持を要求する情報のことを、「保持要求」と記載することがある。通信装置 10 b は、保持要求と共に受信したデータを、そのデータの識別子に対応付けて受信キャッシュ 32 b に格納する。

【0013】

同じデータの送信回数が閾値 N を越えると、通信装置 10 a は、送信するデータの代わりに、そのデータの識別子を通信装置 10 b に通知する。通信装置 10 b は、通知された識別子に対応付けられたデータを受信キャッシュ 32 b から読み出し、読み出したデータを、通信装置 10 a から受信したデータとして扱う。

【0014】

例えば、時刻 T0 から通信装置 10 a と通信装置 10 b が通信を開始した場合を例として説明する。ここでは、閾値 N は 2 に設定されているものとする。通信装置 10 a は、時刻 T0 の時点で、データ A を通信装置 10 b に送信するために、データ A の通信装置 10 b への送信回数を閾値 N と比較する。この時点では、データ A は通信装置 10 b に送信されていないので、今回の送信が 1 回目である。通信装置 10 a は、今回の送信での送信回数が閾値 N に達していないため、図 1 の矢印 に示すように、データ A と共にデータ A についての禁止情報を、通信装置 10 b に送信する。すると、通信装置 10 b は、データ A を受信しても、データ A を受信キャッシュ 32 b に格納しない。このため、受信キャッシュ 32 b __0 に示すように、受信キャッシュ 32 b にデータが格納されない。その後、通信装置 10 a は、データ B、データ C、データ D、データ E についてもデータ A と同様に、禁止情報と共に、通信装置 10 b に送信する。このため、通信装置 10 b は、データ E の受信までは、通信装置 10 a から受信したいずれのデータも受信キャッシュ 32 b に格納しない。

【0015】

次に、通信装置 10 a は、再度データ A を通信装置 10 b に送信するための処理を行う。通信装置 10 a から通信装置 10 b へのデータ A の送信回数は、今回の送信で 2 回になる。通信装置 10 a は、送信回数が閾値 N と同じ値になっているので、図 1 の矢印 に示すように、データ A とともに、データ A についての保持要求を通信装置 10 b に送信する。通信装置 10 b は、データ A を受信すると受信キャッシュ 32 b に格納する。このため、時刻 T1 で行われる送信処理により、通信装置 10 b は受信キャッシュ 32 b __1 を保持する。

【0016】

その後、通信装置 10 a は、データ F を初めて送信するので、禁止情報と共にデータ F を、通信装置 10 b に送信する。すると、通信装置 10 b はデータ F を格納しないため、受信キャッシュ 32 b に保持されているデータは変更されない。

【0017】

時刻 T2 に、通信装置 10 a は、通信装置 10 b に向けたデータ B の 2 回目の送信処理を行う。このとき、通信装置 10 a は、データ B の送信回数が閾値 N と同じ値になるので、図 1 の矢印 に示すように、データ B とデータ B の保持要求を通信装置 10 b に送信す

10

20

30

40

50

る。通信装置 10 b はデータ B を受信キャッシュ 32 b に格納するため、時刻 T2 で行われる送信処理により、通信装置 10 b は、受信キャッシュ 32 b __ 1 を受信キャッシュ 32 b __ 2 に更新する。

【0018】

その後、通信装置 10 a は、データ G、データ H を禁止情報と共に通信装置 10 b に送信する。通信装置 10 b はデータ G、データ H を格納しないため、受信キャッシュ 32 b __ 2 に保持されているデータは変更されない。

【0019】

時刻 T3 に通信装置 10 a は、データ A を 3 回目に送信しようとする。このとき、データ A の送信回数は、閾値 N を上回っているので、通信装置 10 a は、データ A がすでに通信装置 10 b に格納されていると判定する。そこで、通信装置 10 a は、矢印 に示すように、データ A の識別子を通信装置 10 b に送信する。通信装置 10 b は、データ A の識別子を受信すると、受信した識別子をキーとして受信キャッシュ 32 b __ 2 を検索し、データ A を取得する。通信装置 10 b は、取得したデータ A を、通信装置 10 a からの送信データとして扱う。

【0020】

時刻 T4 に通信装置 10 a は、データ B の 3 回目の送信処理を行う。このときも、時刻 T3 でのデータ A の送信処理と同様に、送信回数が閾値 N を上回っているので、通信装置 10 a は、データ B の識別子を通信装置 10 b に送信する。通信装置 10 b は、データ B の識別子をキーとして受信キャッシュ 32 b __ 2 を検索し、データ B を取得する。通信装置 10 b は、取得したデータ B を、通信装置 10 a からの送信データとして扱う。

【0021】

このように、実施形態にかかる方法では、閾値以上の回数送信されるデータは受信側の通信装置 10 が備える受信キャッシュ 32 に格納されるが、送信回数が閾値未満のデータは受信側の通信装置 10 に格納されない。このため、閾値以上の回数送信されるデータ同士の間、閾値未満の回数しか送信されないデータが受信キャッシュ 32 の容量以上に含まれている場合でも、受信側の通信装置 10 は、重複して送信されてくるデータを受信キャッシュ 32 に記憶できる。図 1 の例では、時刻 T1 でのデータ A の送信から時刻 T3 でのデータ A の送信までの間に、受信キャッシュ 32 b の容量分だけ、データ A 以外のデータが通信装置 10 a から通信装置 10 b に送信される。しかし、時刻 T1 ~ T3 に送信されるデータは、データ B 以外は閾値以下の回数しか送信されていないことから受信キャッシュ 32 b への格納対象にならない。このため、受信キャッシュ 32 b での容量不足が発生せず、時刻 T3 にデータ A が再送されるときにも、データ A が受信キャッシュ 32 b に保持される。このため、重複除去が効率的に行われる。

【0022】

換言すると、実施形態にかかる送信側の通信装置 10 では、データの送信回数を、重複の発生頻度の指標として用いることにより、受信側の通信装置 10 に、重複して受信する可能性が比較的高いデータを選択的に記憶させている。従って、通信装置 10 は、重複部分を含むデータの転送を効率的に防止できる。

【0023】

< ネットワークの例と装置構成 >

図 2 は、実施形態にかかる通信装置 10 が用いられるネットワークの例を示す。図 2 は、Wide Area Network (W A N) 5 をはさんで、エリア 4 a 中の装置とエリア 4 b 中の装置が通信する場合のネットワークの例を示す。エリア 4 a、4 b には通信を行う任意の種類の装置が任意の数だけ含まれる。図 2 の例では、エリア 4 a にサーバ 1 a ~ 1 c、タブレット 2 a、携帯電話 3 a が含まれ、エリア 4 b にサーバ 1 d ~ 1 f、タブレット 2 b、携帯電話 3 b が含まれている場合を示している。エリア 4 a に含まれている装置は、通信装置 10 a を介して W A N 5 にアクセスし、エリア 4 b に含まれている装置は、通信装置 10 b を介して W A N 5 にアクセスするものとする。ここで、通信装置 10 a および通信装置 10 b は、W A N 高速化装置として動作することができる。図 2 の例では、エリア 4

10

20

30

40

50

aに含まれている装置が通信装置10aと通信装置10bを介してエリア4bに含まれている装置にデータを送信している。

【0024】

図3は、通信装置10の構成の例を示す。通信装置10は、送信部11、受信部12、送受信部13、接続管理部14、キャッシュ管理部15、データ処理部16、重複検知部20、記憶部30を備える。重複検知部20は、分割部21、識別子生成部22、および、設定部23を有する。記憶部30は、送信データテーブル31、受信キャッシュ32を有する。

【0025】

送信部11は、通信装置10がWAN5に含まれている装置や、他の通信装置10にパケットを送信するとき使用される。送信部11は、設定部23などから入力されたパケットを、宛先に向けて送信する。受信部12は、WAN5に含まれている装置や、他の通信装置10からパケットを受信する。受信部12は、受信したパケットを、キャッシュ管理部15に出力する。送受信部13は、通信装置10とWAN5を介さずに通信する装置との間で通信装置10がパケットを送受信するとき使用される。ここで、WAN5を介さずに通信装置10と通信する装置は、通信装置10と通信可能なエリア4に含まれているサーバ1などの装置である。送受信部13は、エリア4に含まれている装置から受信したパケットを、データ処理部16に出力する。

【0026】

接続管理部14は、通信装置10同士の接続や、通信装置10が通信可能なエリア4に含まれている装置との間の接続状態に関する情報を保持する。キャッシュ管理部15は、受信したパケットに保持要求が含まれている場合、保持要求と共に取得したデータを受信キャッシュ32に格納し、さらに、受信パケット中のデータをデータ処理部16に出力する。一方、受信パケットに禁止情報が含まれている場合、キャッシュ管理部15は、禁止情報と共に取得したデータをデータ処理部16に出力するが、禁止情報と共に取得したデータを受信キャッシュ32には格納しない。受信部12から入力されたパケット中のデータが、特定のデータを表わす識別子である場合、キャッシュ管理部15は、受信キャッシュ32から識別子に対応付けられたデータを読み出し、読み出したデータをデータ処理部16に出力する。なお、受信パケットが、データの受信側の通信装置10で保持されている受信キャッシュ32の情報を通知する制御パケットである場合、キャッシュ管理部15は、入力された制御パケットを設定部23に出力する。

【0027】

データ処理部16は、キャッシュ管理部15から取得したデータを、送信元の通信装置10から送られてきたデータとして処理する。キャッシュ管理部15から取得したデータが通信装置10自身に宛てたデータである場合、データ処理部16は、アプリケーション等を用いて、入力されたデータを処理する。一方、キャッシュ管理部15から入力されたデータが、通信装置10と通信可能なエリア4に含まれている装置に宛てられたものである場合、データ処理部16は、データを宛先の装置へ転送するための処理を行う。

【0028】

重複検知部20は、データを他の通信装置10に送信する通信装置10で動作する。分割部21は、送信対象とするファイルなどを予め決められたデータ長に分割することにより、個々の送信パケットに含めることができる大きさのデータを生成する。以下、個々の送信パケットに含めることができる大きさのデータのことを「チャンク」と記載する。チャンクは、データの重複が発生しているかの判定や、送信回数を特定するときの対象とされるデータのブロックである。分割部21は、生成したチャンクを識別子生成部22に出力する。識別子生成部22は、各チャンクを一意的に識別するための識別子を生成する。識別子の生成方法は、任意であるが、例えば、ハッシュ関数を識別子の生成に用いることができる。識別子生成部22は、生成した識別子を設定部23に出力する。

【0029】

設定部23は、入力された識別子をキーとして、送信データテーブル31を検索する。

送信データテーブル 31 の例については後で示すが、送信データテーブル 31 は、チャンクを受信側の通信装置 10 に送信した回数を、そのチャンクを識別する識別子に対応付けている。送信データテーブル 31 は、キーとして使用する識別子が送信データテーブル 31 に含まれていない場合、処理対象の識別子を、送信回数 = 1 という情報に対応付けて送信データテーブル 31 に記録する。一方、識別子に対応付けられた送信回数を送信データテーブル 31 から得られた場合、設定部 23 は、得られた回数に 1 を加えて、閾値 N と比較する。ここで、閾値 N は、保持要求を有するパケットの送信を含めて、受信側の通信装置 10 に処理対象のチャンクを送信する回数を示す。送信回数が閾値 N よりも小さい間は、設定部 23 は、チャンクとそのチャンクに対する禁止情報を、送信パケットに含める。一方、送信回数が閾値 N に一致すると、設定部 23 は、処理対象のチャンクに保持要求を含めて送信パケットを生成する。送信回数が閾値 N を超えると、設定部 23 は、識別子を送信パケットに含めて受信側の通信装置 10 に送信する。なお、受信側の通信装置 10 が個々のチャンクを保持する領域を示すインデックス値を用いて、送信側の通信装置 10 に受信側の受信キャッシュ 32 の状況を通知しているときには、設定部 23 は、インデックス値を識別子として用いても良い。インデックス値を用いた通信の例については、後述する。

【0030】

図 4 は、通信装置 10 のハードウェア構成の例を示す。通信装置 10 は、プロセッサ 101、メモリ 102、バス 105、外部記憶装置 106、ネットワーク接続装置 109 を備える。さらにオプションとして、通信装置 10 は、入力装置 103、出力装置 104、媒体駆動装置 107 を備えても良い。通信装置 10 は、例えば、コンピュータなどで実現されることがある。

【0031】

プロセッサ 101 は、Central Processing Unit (CPU) を含む任意の処理回路とすることができる。プロセッサ 101 は、接続管理部 14、キャッシュ管理部 15、データ処理部 16、および、重複検知部 20 として動作する。なお、プロセッサ 101 は、例えば、外部記憶装置 106 に記憶されたプログラムを実行することができる。メモリ 102 は、記憶部 30 として動作し、送信データテーブル 31 を記憶する他、受信キャッシュ 32 を実現する。さらに、メモリ 102 は、プロセッサ 101 の動作により得られたデータや、プロセッサ 101 の処理に用いられるデータも、適宜、記憶する。ネットワーク接続装置 109 は、他の装置との通信に使用され、送信部 11、受信部 12、送受信部 13 として動作する。

【0032】

入力装置 103 は、例えば、ボタン、キーボードやマウスとして実現され、出力装置 104 は、ディスプレイなどとして実現される。バス 105 は、プロセッサ 101、メモリ 102、入力装置 103、出力装置 104、外部記憶装置 106、媒体駆動装置 107、ネットワーク接続装置 109 の間を相互にデータの受け渡しが行えるように接続する。外部記憶装置 106 は、プログラムやデータなどを格納し、格納している情報を、適宜、プロセッサ 101 などに提供する。媒体駆動装置 107 は、メモリ 102 や外部記憶装置 106 のデータを可搬記憶媒体 108 に出力することができ、また、可搬記憶媒体 108 からプログラムやデータ等を読み出すことができる。ここで、可搬記憶媒体 108 は、フロッピーディスク、Magneto-Optical (MO) ディスク、Compact Disc Recordable (CD-R) や Digital Versatile Disk Recordable (DVD-R) を含む、持ち運びが可能な任意の記憶媒体とすることができる。

【0033】

< 第 1 の実施形態 >

以下、図 2 に示すネットワークにおいて、エリア 4a 中の装置からエリア 4b 中の装置にデータが送信される場合を例として、通信装置 10a と通信装置 10b で行われる処理の例を説明する。以下、理解し易くするために、エリア 4a 中の装置とエリア 4b 中の装置の間の通信の開始により、通信装置 10a と通信装置 10b の通信が開始された場合を

例として説明する。以下の例でも閾値Nは2である場合を例とするが、閾値Nの値は実装に応じて任意の値に設定され得る。

【0034】

また、以下の説明では、いずれの装置による動作であるかを分り易くするために、符号の後に、その通信装置10の符号の末尾のアルファベットを付すものとする。例えば、設定部23aは、通信装置10aに備えられている設定部23であり、受信キャッシュ32bは、通信装置10bに備えられた受信キャッシュ32である。

【0035】

(1) 通信装置10aは、エリア4aに含まれている装置から、送受信部13aを介してデータを受信する。送受信部13aは、受信したパケットをデータ処理部16aに出力する。データ処理部16aは、入力されたパケットから受信データと送信元や宛先の情報などを取得し、通信装置10b宛てのデータとする。データ処理部16aは、通信装置10b宛てのデータを分割部21aに出力する。

【0036】

図5は、重複検知部20aでの処理の例を説明する図である。分割部21aは、入力されたデータを予め決められたデータ長のチャンクに分割する。図5の例では、分割部21aは、入力されたデータをX個のチャンクに分割している。分割部21aは、個々のチャンクを、識別子生成部22aに出力する。なお、分割部21aに入力されたデータの大きさが、チャンクとして許容されるデータの大きさ以下であれば、分割部21aは、データを分割せずに識別子生成部22aに出力する。

【0037】

図5の例では、識別子生成部22aは、予め保持しているハッシュ関数を用いて、チャンクからハッシュ値を生成し、得られたハッシュ値を個々のチャンクの識別子とする。なお、識別子生成部22aで使用されるハッシュ関数は、個々のチャンクを一意に識別できるようなハッシュ値を生成可能な任意の関数であるものとする。例えば、Secure Hash Algorithm-1 (SHA1) 等のハッシュ関数が識別子生成部22aで使用されることがある。

【0038】

(2) 識別子生成部22aは、個々のチャンクとハッシュ値を対応付けて、設定部23aに出力する。設定部23aは、ハッシュ値を用いて、送信データテーブル31aの検索や送信データテーブル31aの更新処理を行う。

【0039】

図6は、送信データテーブル31の例を示す。図6に示す送信データテーブル31は、ハッシュ値、送信回数、インデックス(index)を対応付けている。ハッシュ値は、識別子生成部22aで生成された値である。設定部23aは、ハッシュ値で識別されるチャンクの各々について、通信装置10bに送信した回数を記録している。インデックスは、受信側の通信装置10bが保持する受信キャッシュ32bにおいて、ハッシュ値で特定されるチャンクが保持されている領域の特定に使用される値である。設定部23aは、保持要求を付して送信するチャンクの大きさを用いて、インデックスを計算し、得られた値を送信データテーブル31aに記録する。

【0040】

図7のC1は、データAであるチャンクがHash1として、設定部23aに入力された場合に行われる送信処理の例を示す。なお、図7の場合では、図を簡略化するために、データAがチャンクとして許容される大きさ以下のデータであった場合の例を示している。識別子生成部22aから設定部23aに、Hash1というハッシュ値と、データAが入力されたとする。今までに通信装置10aから通信装置10bにデータの送信が行われていないので、設定部23aは、送信データテーブル31aにHash1のチャンクのエントリを生成し、Hash1のチャンクの送信回数を1に設定する。新たにエントリを生成する時点では、設定部23aは、インデックスの値を未定とする。このため、Hash1について、図6の送信データテーブル31a__1の1番目のエントリに示すデータが、

10

20

30

40

50

送信データテーブル 3 1 a に記録される。

【 0 0 4 1 】

設定部 2 3 a は、送信データテーブル 3 1 a 中のエントリの更新処理を行った後に、H a s h 1 のチャンクの送信回数と閾値 N を比較する。ここでは、H a s h 1 の送信回数が 1 であるので閾値 N よりも送信回数は小さい。そこで、設定部 2 3 a は、H a s h 1 のチャンク（データ A）、禁止情報、データ A の最終的な宛先などの情報をペイロードとした通信装置 1 0 a 宛てのパケット P 1 を生成する。

【 0 0 4 2 】

ここで、禁止情報は、ペイロード中のデータを受信キャッシュ 3 2 b に格納しないことを通信装置 1 0 b 中のキャッシュ管理部 1 5 b に通知できる情報であれば、任意の形式を取ることができる。以下の説明では、パケット中のヘッダの直後の 2 ビットを、キャッシュ管理部 1 5 b がペイロード中のデータを受信キャッシュ 3 2 b に格納するかを判定するために使用するフラグとして使用する場合を例として説明する。以下、禁止情報では、ヘッダの直後の 2 ビットの値が 0 0 であり、保持要求では、ヘッダの直後の 2 ビットの値が 0 1 であるものとする。設定部 2 3 a は、生成したパケット P 1 を、送信部 1 1 a に出力する。送信部 1 1 a は、図 7 の C 1 に示すように、パケット P 1 を通信装置 1 0 b に送信する。

【 0 0 4 3 】

（ 3 ）通信装置 1 0 b の受信部 1 2 b は、通信装置 1 0 a から送信されたパケット P 1 を受信すると、パケット P 1 をキャッシュ管理部 1 5 b に出力する。キャッシュ管理部 1 5 b は、パケット中のヘッダの直後の 2 ビットの値が 0 0 であることから、禁止情報が付されたデータを取得したと判定する。すると、キャッシュ管理部 1 5 b は、パケット P 1 からデータ A と最終的な宛先の情報などを抽出して、データ処理部 1 6 b に出力する。このとき、図 7 の C 1 に示すように、キャッシュ管理部 1 5 b は、抽出したデータ A を受信キャッシュ 3 2 b に格納しない。データ処理部 1 6 b は、キャッシュ管理部 1 5 b から取得した情報に基づいて、データ A をエリア 4 b 中の装置に送信するためのパケットを生成する。データ処理部 1 6 b は、送受信部 1 3 b を介して、生成したパケットを宛先の装置に送信する。

【 0 0 4 4 】

（ 4 ）その後もエリア 4 a 中の装置からエリア 4 b 中の装置へのデータの送信が続けられ、送信対象のチャンクに同様の処理が行われたとする。送信処理に伴い、送信データテーブル 3 1 a の更新も行われる。このため、例えば、H a s h 1 ~ H a s h 6 のチャンクが通信装置 1 0 a から通信装置 1 0 b に 1 回ずつ送信されると、送信データテーブル 3 1 a は、図 6 に示す送信データテーブル 3 1 a _ 1 に示すようになる。

【 0 0 4 5 】

（ 5 ）図 7 の C 2 に示すように、通信装置 1 0 a に再度データ A が送信されたとする。すると、通信装置 1 0 a の分割部 2 1 a は、手順（ 1 ）で述べたようにデータ A を識別子生成部 2 2 a に出力する。識別子生成部 2 2 a は、記憶しているハッシュ関数を用いてデータ A に対するハッシュ値（H a s h 1）を計算する。識別子生成部 2 2 a は、H a s h 1 というハッシュ値と共に、データ A を設定部 2 3 a に出力する。

【 0 0 4 6 】

（ 6 ）設定部 2 3 a は、識別子生成部 2 2 a から入力された H a s h 1 というハッシュ値をキーとして、送信データテーブル 3 1 a _ 1（図 6）を検索する。すると、1 番目のエントリの情報がヒットする。そこで、設定部 2 3 a は、送信回数を 1 つインクリメントした値を、閾値 N と比較する。ここでは、送信データテーブル 3 1 a _ 1 での H a s h 1 に対応付けられた送信回数の値は 1 であり、今回の処理により、設定部 2 3 a は、H a s h 1 で識別されるチャンクを通信装置 1 0 b に送信するのは 2 回目になる。閾値 N は 2 であるので、設定部 2 3 a は、送信回数を 1 つインクリメントした値と閾値 N が一致したと判定する。そこで、設定部 2 3 a は、通信装置 1 0 b に H a s h 1 で識別されるチャンクを受信キャッシュ 3 2 b に保持することを要求する。

10

20

30

40

50

【 0 0 4 7 】

設定部 2 3 a は、これまでに通信装置 1 0 b に対して保持要求を出したデータの大きさの合計値を、通信装置 1 0 b の受信キャッシュ 3 2 b において、H a s h 1 のデータが保持される領域のインデックス値に設定する。ここでは、設定部 2 3 a はこれまでに保持要求を付して通信装置 1 0 b に送信したデータは無いものとする。すると、設定部 2 3 a は、H a s h 1 のインデックスを 0 に設定する。このため、2 回目に通信装置 1 0 a が受信したデータ A の処理により、送信データテーブル 3 1 a 中の H a s h 1 に対するエントリ中の情報は、送信データテーブル 3 1 a _ 2 の 1 番目のエントリに示す通りに更新される。このとき、設定部 2 3 a は、保持要求を出したデータのデータ長を、保持要求を出したデータのインデックスに加えた値を、次に保持要求を付して送信するデータに対するインデックスとして記憶する。例えば、H a s h 1 のデータの大きさが 1 0 0 0 バイトである場合、設定部 2 3 a は、次に通信装置 1 0 b に対して保持することを要求するデータのインデックス値が 1 0 0 0 になることを記憶する。

10

【 0 0 4 8 】

さらに、設定部 2 3 a は、通信装置 1 0 b に H a s h 1 で識別されるチャンクを受信キャッシュ 3 2 b に保持することを要求するために、データ A と保持要求を含むパケット P 2 を生成する。設定部 2 3 a は、生成したパケット P 2 を送信部 1 1 a に出力する。送信部 1 1 a は、図 7 の C 2 に示すように、パケット P 2 を通信装置 1 0 b に送信する。

【 0 0 4 9 】

(7) 通信装置 1 0 b のキャッシュ管理部 1 5 b は、手順 (3) と同様に、受信部 1 2 b を介してパケット P 2 を取得する。キャッシュ管理部 1 5 b は、パケット中のヘッダの直後の 2 ビットの値が 0 1 であることから、保持要求が付されたデータを取得したと判定する。すると、キャッシュ管理部 1 5 b は、パケット P 2 からデータ A と最終的な宛先の情報などを抽出して、データ処理部 1 6 b に出力する。さらに、図 7 の C 2 に示すように、キャッシュ管理部 1 5 b は、抽出したデータ A を受信キャッシュ 3 2 b に格納する。データ処理部 1 6 b は、キャッシュ管理部 1 5 b から取得した情報に基づいて、手順 (3) と同様に処理する。

20

【 0 0 5 0 】

(8) その後、データ B を通信装置 1 0 a がエリア 4 a 中の装置から 2 回目に受信したとする。ここで、データ B のハッシュ値は、H a s h 2 であるとする。2 回目にデータ B を受信すると、通信装置 1 0 a は、手順 (5)、(6) と同様の処理を用いて、データ B についての保持要求とデータ B を通信装置 1 0 b に送信する。すなわち、H a s h 2 のデータを 2 回目に送信対象として取得すると、設定部 2 3 a は、H a s h 2 のデータに保持要求を付して通信装置 1 0 b に送信する。このとき、設定部 2 3 a は、H a s h 2 のインデックス値を、手順 (6) での計算に従って、1 0 0 0 に設定する。このため、H a s h 2 についての送信データテーブル 3 1 a の情報は、図 6 の送信データテーブル 3 1 a _ 2 の 2 番目のエントリに示すとおり更新される。ここで、H a s h 2 のデータのデータ長が 1 2 0 0 バイトであるとする、設定部 2 3 a は、次に保持要求を出すデータについてのインデックス値を 2 2 0 0 に設定する。

30

【 0 0 5 1 】

(9) H a s h 2 のチャンクを含むパケットを受信した通信装置 1 0 b は、手順 (7) と同様の処理を行う。このため、H a s h 2 のチャンクは、H a s h 2 のチャンクの最終的な宛先に転送されると共に、受信キャッシュ 3 2 b に格納される。図 8 に受信キャッシュ 3 2 でのデータの保存方法の例とインデックス値の決定方法の例を示す。図 8 の左側のテーブルは、受信キャッシュ 3 2 b へのデータの格納状況をキャッシュ管理部 1 5 b が管理するために保持する情報の例である。右側の図は、受信キャッシュ 3 2 b でのデータの格納例を示す。キャッシュ管理部 1 5 b は、受信キャッシュ 3 2 b に格納されている個々のチャンクについて、ハッシュ値、容量 (データ長)、そのチャンクが格納されている領域の先頭の位置を特定するためのインデックス値を記憶する。

40

【 0 0 5 2 】

50

手順(7)で述べた処理により、H a s h 1で識別されるデータAがインデックス0から1000バイト分の領域に保持されているので、キャッシュ管理部15bは、インデックス1000から始まる領域にH a s h 2で識別されるチャンクを格納する。ここで、H a s h 2で識別されるチャンクはデータBであるとする、H a s h 2のチャンクは、図8の右側に示すように、受信キャッシュ32bにおいてデータAが格納された領域の続きに格納される。また、キャッシュ管理部15bは、H a s h 2のチャンクの格納状況を、図8の左側のテーブルの2番目のエントリに示すように記憶する。

【0053】

(10) その後も、エリア4a中の装置からエリア4b中の装置へのデータの送信が続けられ、H a s h 3 ~ H a s h 5のチャンクについても、閾値Nと同数の送信が行われたとする。すると、通信装置10aは、H a s h 3 ~ H a s h 5のチャンクの各々についても、手順(5)、(6)、(8)等で述べた処理を行うことにより、受信キャッシュ32bに格納することを、通信装置10bに要求する。なお、H a s h 3 ~ H a s h 5のチャンクの送信の合間に、送信回数が閾値Nに達していない別のチャンク(H a s h 7、H a s h 8)の送信を行う場合、通信装置10aは、手順(3)と同様の処理を行う。このため、H a s h 3 ~ H a s h 5のチャンクの2回目の送信と、H a s h 6 ~ H a s h 8のチャンクの1回目の送信を終えた段階では、通信装置10aが保持する送信データテーブル31aは、送信データテーブル31a_2(図6)に示す通りになる。

【0054】

(11) 一方、通信装置10bでは、H a s h 3 ~ H a s h 5のチャンクを含むパケットを受信すると、手順(9)で述べたように、受信したパケットに含まれているチャンクを受信キャッシュ32bに格納する。さらに、H a s h 3 ~ H a s h 5のチャンクのそれぞれについて、ハッシュ値とインデックスを対応付ける。このため、H a s h 3 ~ H a s h 5のチャンクの処理により、受信キャッシュ32bには、図8の右側に示すようにデータが格納され、キャッシュ管理部15bは、図8の左側に示すような情報を保持する。

【0055】

(12) 図7のC3に示すように、さらに、エリア4a中の装置から通信装置10aにデータAが送信されたとする。すると、通信装置10aの分割部21aは、手順(1)で述べたようにデータAを識別子生成部22aに出力する。識別子生成部22aは、記憶しているハッシュ関数を用いてデータAに対するハッシュ値(H a s h 1)を計算する。識別子生成部22aは、H a s h 1というハッシュ値と共に、データAを設定部23aに出力する。

【0056】

(13) 設定部23aは、識別子生成部22aから入力されたハッシュ値H a s h 1をキーとして、送信データテーブル31a_2(図6)を検索する。すると、1番目のエントリから、H a s h 1のチャンクが既に2階送信されていることを特定する。そこで、設定部23aは、送信回数を1つインクリメントした値を、閾値Nと比較する。ここでは、H a s h 1で識別されるチャンクを通信装置10bに送信するのは3回目になるのに対し、閾値Nは2であるため、設定部23aは、H a s h 1のチャンクについては、送信回数が閾値Nを越えると判定する。そこで、設定部23aは、通信装置10bにH a s h 1で識別されるチャンクの代わりに、そのチャンクを識別する値を含めたパケットを生成する。ここで、チャンクの識別に使用される値は、通信装置10bにおいてチャンクを一意に識別できる任意の情報である。例えば、インデックス値と、チャンクのデータ長がチャンクの識別子としてパケットに含められても良い。この場合、設定部23aは、H a s h 1に対応付けられたインデックスである0と、データ長が1000バイトであることをデータとして含めたパケットP3を送信部11aに出力する。このとき、設定部23aは、送信するデータがチャンクを識別するための情報であることを示す情報をパケットP3に含めるものとする。以下の説明では、送信するデータがチャンクを識別するための情報であることを示す情報として、設定部23aは、パケット中のヘッダの直後の2ビットの値を11に設定するものとする。図7のC3に示すように、送信部11aは、パケットP3を

10

20

30

40

50

通信装置 10 b に送信する。なお、設定部 23 a は、送信データテーブル 31 a の Hash 1 のチャンクについての送信回数を更新する。

【0057】

(14) 通信装置 10 b のキャッシュ管理部 15 b は、手順 (3) と同様に、受信部 12 b を介してパケット P3 を取得する。キャッシュ管理部 15 b は、パケット中のヘッダの直後の 2 ビットの値が 11 であることから、受信キャッシュ 32 b に格納済みのチャンクに対応付けられたインデックス値とデータ長が通知された判定する。すると、キャッシュ管理部 15 b は、パケット P3 からインデックス値やデータ長を抽出する。キャッシュ管理部 15 b は、抽出したインデックス値で表わされる受信キャッシュ 32 b の領域から、抽出したデータ長の情報を、通信装置 10 a から受信したデータとして取得する。キャッシュ管理部 15 b は取得したデータをデータ処理部 16 b に出力する。データ処理部 16 b は、キャッシュ管理部 15 b から取得した情報に基づいて、手順 (3) と同様に処理する。

【0058】

図 9 は、通信処理の手順の例を説明する図である。手順 (1) ~ (14) で説明した処理が行われることにより、エリア 4 a 中の装置からエリア 4 b 中の装置へのデータの送信の際に、通信装置 10 a と通信装置 10 b との間で効率的に重複排除処理が行われる。すなわち、通信装置 10 a は、エリア 4 a 中の装置から、エリア 4 b 中の装置への送信データを受信すると、受信したデータを適宜、チャンクに分割して、個々のチャンクについての送信回数の判定を行う。このとき、通信装置 10 a 中の設定部 23 a は、送信データテーブル 31 a のデータを使用して、送信の回数と閾値の関係を判定する。送信回数が N 回以下のチャンクについては、図 9 中の矢印 A に示すように、通信装置 10 a は、実データを含むパケットを通信装置 10 b に送信する。このとき、送信回数が N 回未満の間、設定部 23 a は、通信装置 10 b に送信するパケット中に、禁止情報を含めておくことにより、通信装置 10 b でのデータの格納を抑制する。また、送信回数が N 回になると、保持要求を含めることにより、設定部 23 a は、送信するチャンクを通信装置 10 b に保持させる。送信回数が N + 1 回以上になると、図 9 中の矢印 B に示すように、設定部 23 a は、実データの代わりにチャンクを一意に特定できる識別子 (ID データ) を含むパケットを通信装置 10 b に送信する。すなわち、設定部 23 a は、同じデータの送信回数と閾値 N との比較結果に応じて、送信対象のデータの代替として送信する代替のデータを生成し、生成した代替データを受信側の通信装置 10 に送信するデータとしている。さらに、通信装置 10 a では、矢印 C に示すように、データの送信状況を送信データテーブル 31 a に保存する。

【0059】

禁止情報を含むパケットを受信した場合、通信装置 10 b は、図 9 の矢印 D に示すように、チャンクの実データを転送処理に用いるが、受信キャッシュ 32 b に受信したデータを格納しない。保持要求を含むパケットを受信した場合、通信装置 10 b は、図 9 の矢印 E に示すように、チャンクの実データを転送処理に用い、さらに、実データを受信キャッシュ 32 b に格納する。実データの代わりにチャンクの識別子を含むパケットを受信した場合、通信装置 10 b 中のキャッシュ管理部 15 b は、図 9 の矢印 F に示すように、識別子で識別されるチャンクの実データを受信キャッシュ 32 b から読み出すことによりデータを復元する。通信装置 10 b は、復元されたデータを転送処理に用いる。なお、チャンクの識別に使用される識別子は、図 6 ~ 図 8 を参照しながら説明したように、インデックス値とチャンクのデータ長とすることができが、その他の情報を用いても良い。例えば、チャンクのハッシュ値が識別子として用いられても良いものとする。

【0060】

図 10 は、データを送信する通信装置 10 での処理の例を説明するフローチャートである。なお、図 10 は一例であり、例えば、ステップ S5 と S7 の判断の順序が互いに変更されるなどの変更が加えられても良い。送信側の通信装置 10 の送受信部 13 は、送信データを取得する (ステップ S1)。送受信部 13 からデータを入力されると、分割部 21

は、データを所定の長さ以下のチャンクに分割する（ステップS2）。識別子生成部22は、分割部21で生成された各チャンクについて、ハッシュ関数を用いてハッシュ値を計算する（ステップS3）。識別子生成部22は、ハッシュ値とチャンクのデータを設定部23に出力する。設定部23は、入力されたチャンクについて、送信データテーブル31を用いて、今回の送信が同じ通信装置10への何回目の送信であるかを特定する（ステップS4）。設定部23は、今回の送信を行ったときの送信回数が閾値未満であるかを判定する（ステップS5）。今回の送信を行ったときの送信回数が閾値未満の場合、設定部23は、実データと禁止情報を受信側の通信装置10に送信する（ステップS5で閾値未満、ステップS6）。今回の送信を行ったときの送信回数が閾値未満ではない場合、設定部23は、今回の送信を行ったときの送信回数が閾値を超えているかを判定する（ステップS5でNo、ステップS7）。送信回数が閾値と一致している場合、設定部23は、実データと保持要求を、受信側の通信装置10に送信する（ステップS7で閾値と一致、ステップS8）。送信回数が閾値を超えている場合、設定部23は、実データを含めずに、チャンクの識別子のみを、受信側の通信装置10に送信する（ステップS7で閾値より大、ステップS9）。

10

【0061】

図11は、データを受信する通信装置10での処理の例を説明するフローチャートである。図11は一例であり、例えば、ステップS23では、禁止情報を含むかを判定する代わりに、保持要求を含むかを判定するように変形されても良い。受信部12は、送信側の通信装置10からパケットを受信すると、データをキャッシュ管理部15に出力する（ステップS21）。キャッシュ管理部15は、パケット中の予め決められている領域の値を特定することにより、パケット中に実データが含まれているかを判定する（ステップS22）。パケット中に実データが含まれている場合、キャッシュ管理部15は、パケット中に禁止情報が含まれているかを判定する（ステップS22でYes、ステップS23）。パケット中に禁止情報が含まれていない場合、キャッシュ管理部15は、実データを受信キャッシュ32に書き込んだ後で、実データの転送処理を行う（ステップS23でNo、ステップS24、S25）。一方、パケット中に禁止情報が含まれている場合、キャッシュ管理部15は、実データを受信キャッシュ32に書き込まずに、実データの転送処理を行う（ステップS23でYes、ステップS25）。一方、受信したパケットに実データが含まれていない場合、キャッシュ管理部15は、受信パケットからデータの識別子を取得する（ステップS22でNo、ステップS26）。キャッシュ管理部15は、取得した識別子に対応付けられたデータを受信キャッシュ32から読み込む（ステップS27）。キャッシュ管理部15は、読み込んだデータをデータ処理部16に出力する。データ処理部16は、キャッシュ管理部15から入力されたデータの転送処理を行う（ステップS28）。

20

30

【0062】

このように、実施形態にかかる方法では、送信回数が閾値未満の間は、送信データが受信側の通信装置10において格納されない。このため、閾値以上の回数、送信されるデータ同士の間、閾値未満の回数しか送信されないデータが含まれている場合でも、受信側の通信装置10は、閾値以上の回数だけ送信される可能性が高いデータを、選択的に受信キャッシュ32に記憶することができる。従って、通信装置10を用いたシステムでは、重複部分を含むデータの転送が効率的に防止される。

40

【0063】

ネットワーク中で送受信されるトラフィックの傾向として、重複して送受信される回数の多いデータは、一部のデータに集中する傾向があることが報告されている。（例えば、A. Anand, et al., "Redundancy in Network Traffic: Findings and Implications," in SIGMETRICS/Performance 2009, Seattle, WA USA）。例えば、1回送信されたデータを全て記憶することができる受信キャッシュを用いて重複排除を行った場合に、送信対象とするデータの量を、重複排除を行わない場合に比べてA%削減できるとする。同じデータについて、重複回数の多い順にチャンクを並べたときの上位20%に重複するチャンク

50

だけを重複排除の対象とした場合に削減できる量は、概ね、 $0.8 \times A\%$ 程度であるということも報告されている。すなわち、重複除去に使用するデータ数を5分の1に削減しても、重複除去の性能の低下は20%程度に過ぎない。従って、第1の実施形態で示すように、最初の数回は禁止情報を付した実データを受信側の通信装置10に送信することによって受信キャッシュ32への格納を妨げたとしても、送受信されるデータ量の削減に対して、ほとんど影響を与えないといえる。

【0064】

さらに、実施形態にかかる方法では、受信キャッシュ32に格納されるデータは、受信側の通信装置10に少なくとも閾値と同じ回数は送信されるチャンクである。このため、全ての受信データが受信キャッシュに格納される場合に比べて、通信装置10中の受信キャッシュ32の容量を小さくすることができる。

【0065】

なお、以上の説明では、分かり易くするために通信装置10aが送信側、通信装置10bが受信側の場合を例として説明したが、いずれの通信装置10も、送信側と受信側の両方の装置として動作することができる。さらに、通信装置10は、複数の通信装置10にデータの送信をすることができ、また、複数の通信装置10からデータを受信することもできる。このため、複数の通信装置10との間で通信が行われる場合、キャッシュ管理部15は、送信側の通信装置10と受信側の通信装置10の組合せごとに分けて、受信キャッシュ32の情報を管理する。従って、キャッシュ管理部15は、通信装置10aから通信装置10bに送信されたインデックス0に格納されたデータと、通信装置10cから通信装置10bに送信されたインデックス0に格納されたデータを別個に管理できる。同様に、設定部23も、宛先の通信装置10ごとに、別の送信データテーブル31を保持しており、受信側の通信装置10にあわせた送信データテーブル31を用いて、送信パケットに含める情報を設定しているものとする。

【0066】

< 第2の実施形態 >

第2の実施形態では、トラフィックレートに応じて、通信装置50が閾値Nを変更する場合の例を説明する。以下、ネットワーク上を単位時間あたりに流れるデータ量を「トラフィックレート」と記載することがある。また、以下の説明では、特定の経路を1秒に流れるデータ量をbps (Bits Per Second) で示すものとする。

【0067】

図12は、第2の実施形態にかかる通信装置50の構成の例を示す図である。通信装置50は、閾値設定部40を備え、さらに、送信部11、受信部12、送受信部13、接続管理部14、キャッシュ管理部15、データ処理部16、重複検知部20、記憶部30を備える。送信部11、受信部12、送受信部13、接続管理部14、キャッシュ管理部15、データ処理部16、重複検知部20、記憶部30は、第1の実施形態と同様に動作する。

【0068】

閾値設定部40は、閾値決定部41とトラフィックレート算出部42を有する。トラフィックレート算出部42は、接続管理部14から、通信装置50自身が受信側として通信している通信装置50の数と、それらの通信装置50の識別子を取得する。閾値決定部41は、送信側の通信装置50の各々について、その装置から受信するデータのトラフィックレートを算出する。トラフィックレート算出部42は、得られた結果を閾値決定部41に出力する。

【0069】

閾値決定部41は、受信キャッシュ32の空き容量を求め、送信元の装置ごとに、保持要求を付してデータを送信するときの送信回数(閾値N)を計算する。このとき、閾値決定部41は、受信データのトラフィックレートが低いほど、送信側の装置で使用される閾値Nが小さくなるように、閾値Nを決定する。このため、トラフィックレートが低いほど、同じデータが繰り返し送受信される回数が小さくなり、通信経路を有効活用できるよう

になる。一方、トラフィックレートが高い場合は、閾値Nが比較的大きく設定されるので、同じデータが繰り返して送受信されることはあるが、受信側の通信装置50において受信キャッシュ32に格納するデータを、送信回数が比較的大きいデータに限定できる。

【0070】

図13は、閾値の決定方法の例を説明する図である。図13では、通信装置50aが通信装置50b～50eからデータを受信しており、通信装置50aが通信装置50b～50eの各々で使用される閾値Nを計算する場合を例として説明する。なお、図13では、通信装置50a～50eの各々が送信側と受信側のいずれとして動作しているときの閾値が計算されているかを見やすくするために、便宜上、通信装置50aを受信装置、通信装置50b～50eを送信装置と表記している。

10

【0071】

通信装置50a中の接続管理部14aは、通信装置50b～50eとの間で、通信装置50aが受信側として通信していることを、トラフィックレート算出部42aに通知する。トラフィックレート算出部42aは、通信装置50b～50eの各々から受信するデータのトラフィックレートを算出する。例えば、トラフィックレート算出部42aは、送信元の装置別に所定の時間中に受信したパケットの数を計数することにより、トラフィックレートを計算できる。また、トラフィックレート算出部42aは、送信元の装置別に、所定のデータの量を受信するまでにかかった時間から、トラフィックレートを計算することができる。図13の例では、トラフィックレート算出部42aは、通信装置50aが通信装置50bとの間では50Mbps、通信装置50cとの間では20Mbps、通信装置50dおよび50eとの間では15Mbpsで通信していると計算したとする。トラフィックレート算出部42aは、各通信装置50について得られたトラフィックレートの値を、閾値決定部41aに出力する。

20

【0072】

閾値決定部41aは、受信キャッシュ32aの空き容量を計算する。閾値決定部41aは、閾値Nを、トラフィックレートの割合に正比例するように決定することにより、トラフィックレートの低い通信で優先的に受信キャッシュ32aを使用できるようにする。すなわち、以下の式を用いて、閾値Nを決定する。

$$N = M \times R \times C \quad \cdots (1)$$

ここで、Mは、受信キャッシュ32aの空き容量、Cは定数である。Rは、通信装置50aが受信しているデータのトラフィックレートの合計値のうち、閾値Nを計算する対象とする通信装置50との間のトラフィックレートが占める割合である。乗数Cは、閾値Nが1～20程度の予め決められた値の範囲に収まるように設定される。閾値決定部41aは、通信装置50aにデータを送信してきている全ての通信装置50b～50eに対して同じ定数Cを用いて、閾値Nを計算する。なお、計算で得られた閾値Nが整数でない場合、閾値決定部41aは、得られた計算値より小さい整数の最大値を、閾値Nに設定するものとする。なお、得られた計算値より小さい整数の最大値が1より小さくなる場合、閾値決定部41aは、閾値Nを1に設定する。

30

【0073】

例えば、受信キャッシュ32aの空き容量が1Mバイトであるとする、式(1)中のMは 1×10^9 となる。ここで、閾値決定部41aが使用する定数Cが 1×10^{-8} であるとする。通信装置50aが通信装置50b～50eとの間で行っている通信では、トラフィックレートの合計値は、 $50 \text{ Mbps} + 20 \text{ Mbps} + 15 \text{ Mbps} + 15 \text{ Mbps} = 100 \text{ Mbps}$ である。そこで、トラフィックレートの割合Rは、通信装置50bでは $50 / 100$ 、通信装置50cでは $20 / 100$ であり、通信装置50dと50eでは $15 / 100$ である。この場合、閾値決定部41aは、式(1)より、閾値Nを以下のように計算する。

40

$$N(\text{通信装置 } 50b) = 1 \times 10^9 \times (50 / 100) \times 1 \times 10^{-8} = 5$$

$$N(\text{通信装置 } 50c) = 1 \times 10^9 \times (20 / 100) \times 1 \times 10^{-8} = 2$$

$$N(\text{通信装置 } 50d、50e)$$

50

$$= 1 \times 10^9 \times (15 / 100) \times 1 \times 10^{-8} = 1.5$$

従って、閾値決定部 41a は、通信装置 50b との間の閾値 N を 5、通信装置 50c との間の閾値 N を 2 とする。さらに、通信装置 50d と通信装置 50e については、計算値が 1.5 であるので、閾値 N を 1 に設定する。

【0074】

閾値決定部 41a は、得られた値を、送信側の通信装置 50 に通知するための制御パケットを生成する。制御パケットには、宛先のアドレスを指定するためのヘッダ、閾値 N の値を通知することを示す情報、および、得られた閾値 N が含まれる。閾値決定部 41a は、生成した制御パケットを、送信部 11a を介して送信先の通信装置 50 に送信する。

【0075】

通信装置 50b の受信部 12b は、通信装置 50a から受信した制御パケットを設定部 23b に出力する。設定部 23b は、記憶している閾値 N を、制御パケットに格納されている値に更新する。設定部 23b は、閾値 N を更新すると、以後は、更新後の値を用いて、通信装置 50a に送信するデータを設定する。すなわち、新たに設定された閾値 N よりも送信回数が小さい間、設定部 23b は、実データに禁止情報を付したものを、送信対象とし、新たな閾値 N と同じ回数の送信では、実データと保持情報を通信装置 50a への送信対象とする。通信装置 50c ~ 通信装置 50e も、通信装置 50a から制御パケットを受信すると、同様の処理を行う。なお、送信側の通信装置 50 は、同じデータの送信回数が閾値 N になると保持要求と共にデータを送信するので、閾値 N は、送信側の通信装置 50 にとって、保持要求を付さずに同じデータを送信できる回数の上限値の指標であるともいえる。このため、第 2 の実施形態では、受信側の通信装置 50 が送信側の通信装置 50 に対して、同じデータを保持要求を付さずに送信できる回数の上限値の指標を決定しているともいえる。

【0076】

図 14 は、データを受信する通信装置 50 の処理の例を説明するフローチャートである。図 14 は、閾値 N の再計算が所定の周期ごとに行われる場合の例を示している。この場合、各通信装置 50 はタイマを保持しており、受信側の通信装置 50 が所定の周期ごとに閾値 N の計算処理を行うものとする。図 14 の例では、変数 n、m、定数 Y を使用する。定数 Y は、閾値 N を計算する通信装置 50 に対してデータを送信している通信装置 50 の数である。変数 n は、トラフィックレートの計算を行った経路の数を計数するために使用される変数である。また、変数 m は、閾値 N の計算を行った通信装置 50 の数を計数するために使用される。なお、図 14 でも、データの送信側の通信装置 50 を分かりやすくするために、フローチャート中では、送信側の通信装置 50 のことを、便宜上、送信装置と記載している。

【0077】

トラフィックレート算出部 42 は、タイマを初期化する（ステップ S41）。トラフィックレート算出部 42 は、変数 n を 1 に設定すると共に、接続管理部 14 から通信装置 50 にデータを送信している送信元の通信装置 50 の数（Y）を取得する（ステップ S42）。トラフィックレート算出部 42 は、n 番目の送信側の通信装置 50 との間の通信のトラフィックレートを計算する（ステップ S43）。トラフィックレート算出部 42 は、n を 1 つインクリメントして、n が Y 以下であるかを判定する（ステップ S44、S45）。変数 n が Y 以下である場合、トラフィックレート算出部 42 は、ステップ S43 以降の処理を繰り返す（ステップ S45 で Yes）。一方、変数 n が Y を超えると、閾値決定部 41 は、変数 m を 1 に設定する（ステップ S45 で No、ステップ S46）。閾値決定部 41 は、m 番目の送信側の通信装置 50 との間で使用する閾値 N を計算する（ステップ S47）。閾値決定部 41 は、変数 m を 1 つインクリメントして、m が Y 以下であるかを判定する（ステップ S48、S49）。変数 m が Y 以下である場合、閾値決定部 41 は、ステップ S47 以降の処理を繰り返す（ステップ S49 で Yes）。一方、変数 m が Y を超えると、トラフィックレート算出部 42 は、タイマが満了したかを判定する（ステップ S49 で No、ステップ S50）。タイマが満了していない場合、トラフィックレート算出

10

20

30

40

50

部 4 2 は、タイマが満了するまで待機する（ステップ S 5 0 で N o）。タイマが満了すると、トラフィックレート算出部 4 2 は、ステップ S 4 1 以降の処理を繰り返す（ステップ S 5 0 で Y e s）。

【 0 0 7 8 】

第 2 の実施形態によると、通信装置 5 0 は、受信キャッシュ 3 2 の空き領域に応じて閾値 N の値を変更することができる。さらに、送信側の個々の通信装置 5 0 で使用される閾値 N の計算は、定期的に行われるので、受信側の通信装置 5 0 と通信する送信側の通信装置 5 0 の数が変動した場合にも、変動に応じて閾値 N が変更できる。従って、通信装置 5 0 で行われている通信の状況に応じて、自律的に閾値 N の値が変更されるため、状況に適した通信の設定が行われる。

10

【 0 0 7 9 】

< その他 >

なお、実施形態は上記に限られるものではなく、様々に変形可能である。以下にその例をいくつか述べる。

【 0 0 8 0 】

送信回数が閾値を超えたときにチャンクの識別のために送信される情報は、受信側の通信装置 1 0 においてチャンクを一意に識別できる任意の情報であり、例えば、ハッシュ値であっても良い。ハッシュ値が識別子として使用される場合、送信側の通信装置 1 0 に含まれる設定部 2 3 は、送信しようとするチャンクについてハッシュ値を、チャンクの代わりにパケットに含めて受信側の通信装置 1 0 に送信する。この場合、受信側の通信装置 1 0 は、通知されたハッシュ値に対応付けられたデータを受信したものとして処理を行う。

20

【 0 0 8 1 】

識別子としてインデックス値とチャンクの長さが用いられる場合でも、受信側の通信装置 1 0 に備えられている受信キャッシュ 3 2 の空き容量がなくなることもあり得る。そこで、受信側の通信装置 1 0 は、受信キャッシュ 3 2 の空き容量がなくなったことによりチャンクを削除する場合は、削除するチャンクを送信側の通信装置 1 0 に通知する。この処理により、送信側の通信装置 1 0 と受信側の通信装置 1 0 の間で同じチャンクに対するインデックス値を一致させることができる。なお、通信装置 5 0 同士の通信においても、インデックス値を用いて重複しているデータを受信側の通信装置 5 0 に通知する場合は、送信側と受信側の通信装置 5 0 で同様の処理が行われる。

30

【 0 0 8 2 】

以上の説明で記載した送信データテーブル 3 1 や受信キャッシュ 3 2 に記憶される情報要素や情報の形式は、一例に過ぎない。送信データテーブル 3 1 や受信キャッシュ 3 2 に記憶される情報要素や情報の記憶の形式等は、実装に応じて任意に変更され得る。

【 0 0 8 3 】

第 1 および第 2 の実施形態のいずれにおいても、受信側の装置が保持する受信キャッシュ 3 2 の空き領域が小さいほど、保持情報を含めたデータを送信するかを決定するときに使用する閾値の値が大きくなるように調整されるように、変形することができる。例えば、第 2 の実施形態の場合では、閾値決定部 4 1 は、閾値 N を計算する際に、受信キャッシュ 3 2 の空き容量が小さいほど、定数 C の値を大きくすることができる。このとき、閾値決定部 4 1 は、受信キャッシュ 3 2 の空き容量と定数 C の関係を予め記憶しているものとし、閾値 N を計算するときの受信キャッシュ 3 2 の空き容量に合わせて、定数 C を選択できるものとする。第 1 の実施形態では、設定部 2 3 に設定される閾値 N は、受信キャッシュ 3 2 のうちの使用可能な容量が小さいほど、閾値 N が小さくなるように設定されるものとする。

40

【 0 0 8 4 】

第 2 の実施形態において、受信側の装置が通信装置 5 0 であれば、送信側の装置は通信装置 5 0 と通信装置 1 0 のいずれでも、同様に閾値の変更処理が可能である。なお、この場合、通信装置 1 0 も、データの受信側の通信装置 5 0 から受信した制御パケットについては、送信側の通信装置 5 0 と同様の処理を行うものとする。

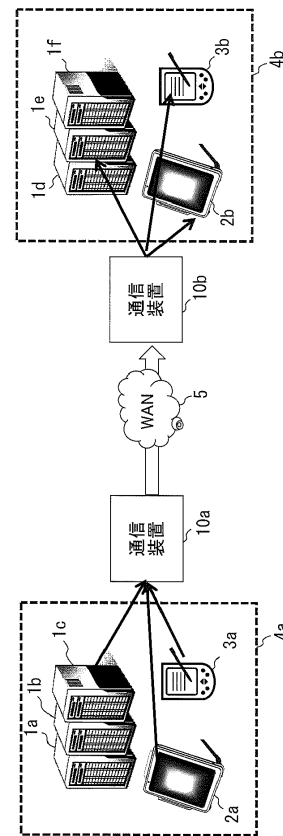
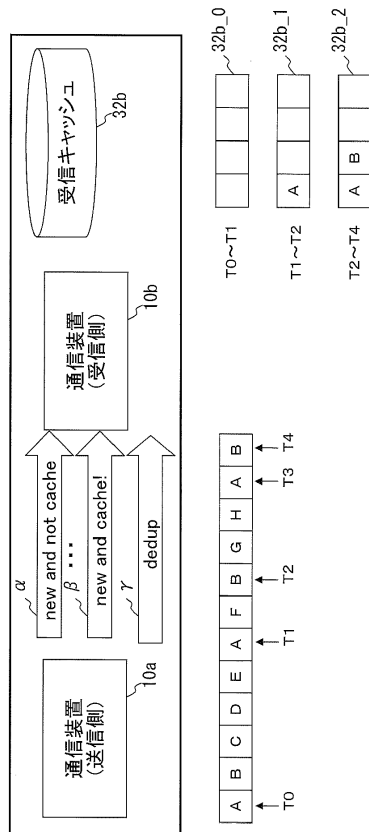
50

【符号の説明】

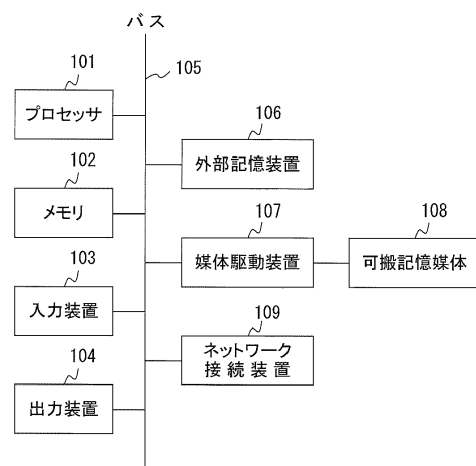
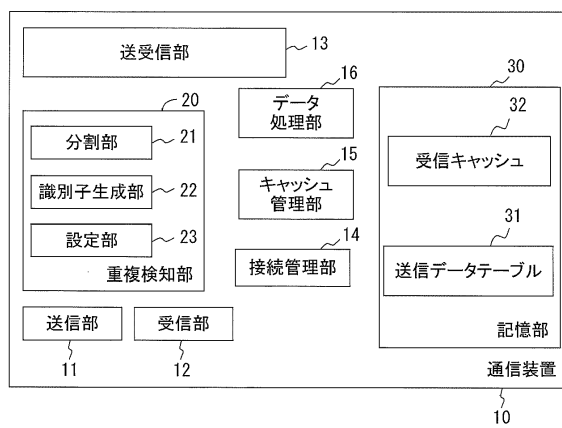
【0085】

1	サーバ	
2	タブレット	
3	携帯電話	
4	エリア	
5	W A N	
10、50	通信装置	
11	送信部	
12	受信部	10
13	送受信部	
14	接続管理部	
15	キャッシュ管理部	
16	データ処理部	
20	重複検知部	
21	分割部	
22	識別子生成部	
23	設定部	
30	記憶部	
31	送信データテーブル	20
32	受信キャッシュ	
40	閾値設定部	
41	閾値決定部	
42	トラフィックレート算出部	
101	プロセッサ	
102	メモリ	
103	入力装置	
104	出力装置	
105	バス	
106	外部記憶装置	30
107	媒体駆動装置	
108	可搬記憶媒体	
109	ネットワーク接続装置	

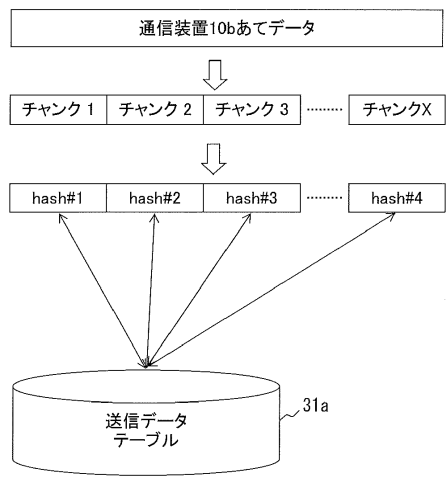
【 図 2 】



【圖 4】



【図 5】

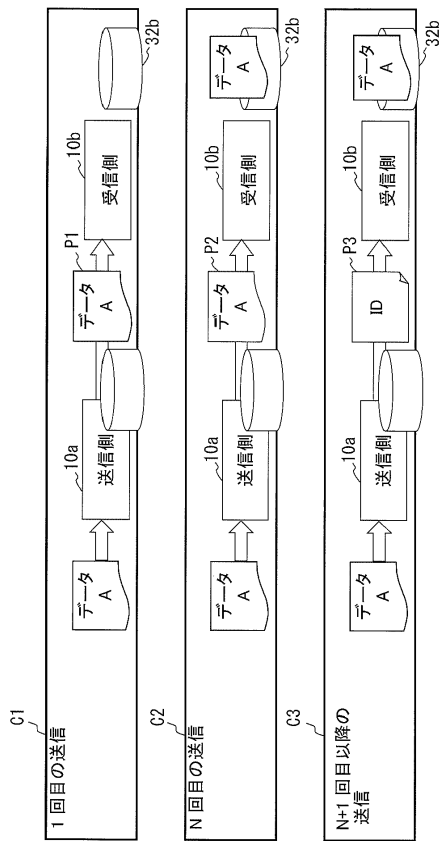


【図 6】

31a_1		
ハッシュ値	送信回数	index
Hash1	1	
Hash2	1	
Hash3	1	
Hash4	1	
Hash5	1	
Hash6	1	

31a_2		
ハッシュ値	送信回数	index
Hash1	2	0
Hash2	2	1000
Hash3	2	2200
Hash4	2	3000
Hash5	2	4500
Hash6	1	
Hash7	1	
Hash8	1	

【図 7】

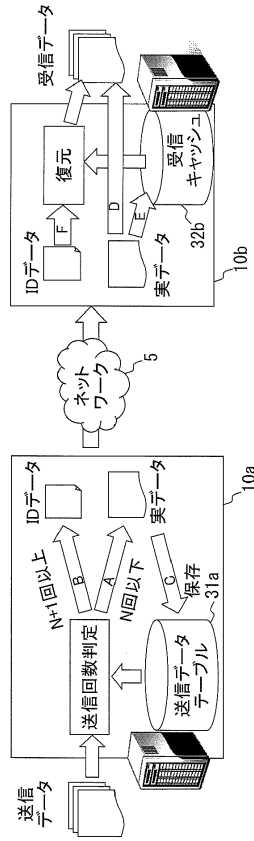


【図 8】

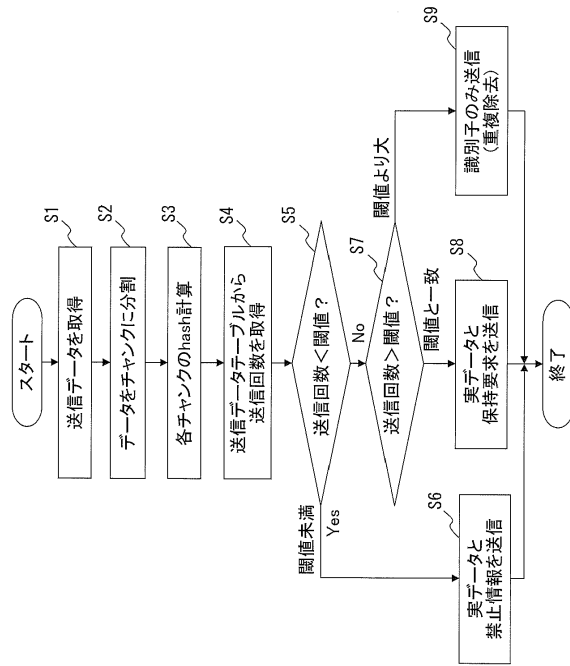
データA (Hash1)	データB (Hash2)	データC (Hash3)	データD (Hash4)	データE (Hash5)

ハッシュ値	容量	index
Hash1	1000	0
Hash2	1200	1000
Hash3	800	2200
Hash4	1500	3000
Hash5	1000	4500

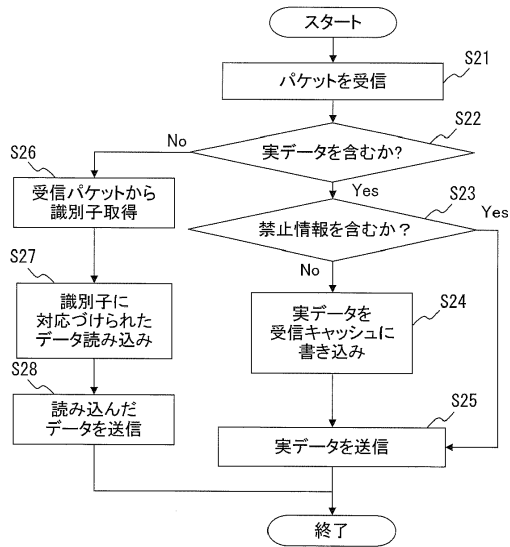
【図 9】



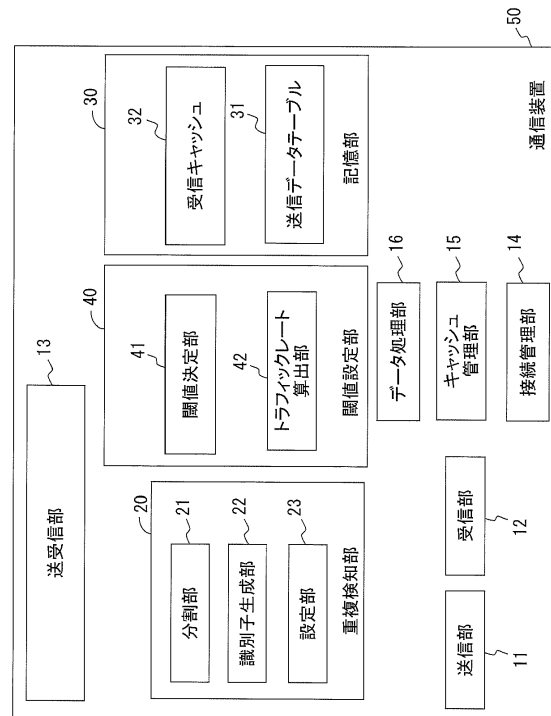
【図 10】



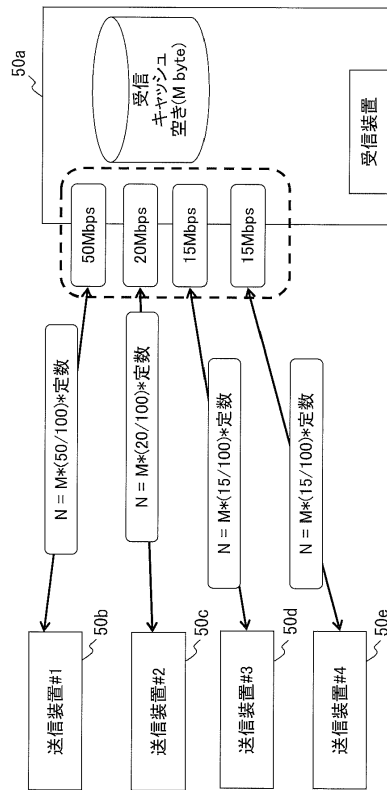
【図 11】



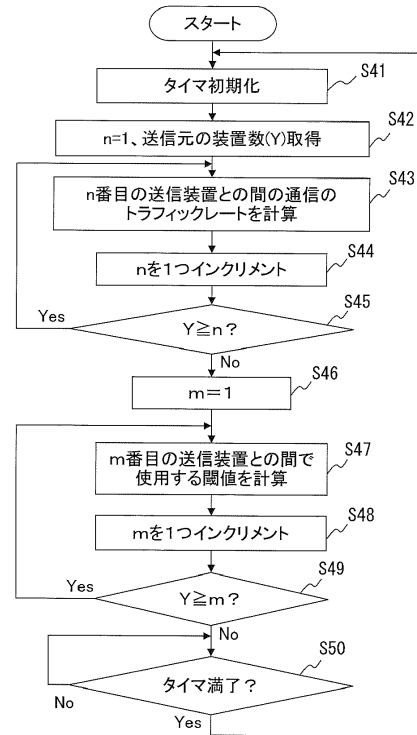
【図 12】



【図 13】



【図 14】



フロントページの続き

(56)参考文献 特開2009-205201(JP,A)
特開2009-187393(JP,A)
特表2003-524968(JP,A)
特開2012-79044(JP,A)

(58)調査した分野(Int.Cl., DB名)
H04L 12/46
G06F 13/00
H04L 29/08