

(19) 日本国特許庁(JP)

(12) 特 許 公 報(B2)

(11) 特許番号

特許第6730437号
(P6730437)

(45) 発行日 令和2年7月29日(2020.7.29)

(24) 登録日 令和2年7月6日(2020.7.6)

(51) Int.Cl. F I
G06F 13/00 (2006.01) G06F 13/00 520B

請求項の数 7 (全 40 頁)

(21) 出願番号	特願2018-535971 (P2018-535971)	(73) 特許権者	000005108 株式会社日立製作所 東京都千代田区丸の内一丁目6番6号
(86) (22) 出願日	平成28年8月24日(2016.8.24)	(74) 代理人	110001678 特許業務法人藤央特許事務所
(86) 国際出願番号	PCT/JP2016/074632	(72) 発明者	谷田 直輝 東京都千代田区丸の内一丁目6番6号 株 株式会社日立製作所内
(87) 国際公開番号	W02018/037495	(72) 発明者	明石 大 東京都千代田区丸の内一丁目6番6号 株 株式会社日立製作所内
(87) 国際公開日	平成30年3月1日(2018.3.1)	(72) 発明者	藤嶋 堅三郎 東京都千代田区丸の内一丁目6番6号 株 株式会社日立製作所内
審査請求日	平成30年9月6日(2018.9.6)		

最終頁に続く

(54) 【発明の名称】 転送装置及びネットワークシステム

(57) 【特許請求の範囲】

【請求項1】

端末間で送受信されるデータを転送する転送装置であって、
プロセッサ、前記プロセッサに接続される主記憶装置、前記プロセッサに接続される副記憶装置、及び前記プロセッサに接続されるネットワークインタフェースを備え、

前記端末間のデータの転送処理を制御する転送制御部を有し、

ネットワークを介して他の転送装置と接続され、

前記副記憶装置は、データを分割することによって生成されるデータブロックであって、前記端末間で過去に送信されたデータブロックを管理するデータブロック情報を格納し、

前記主記憶装置は、前記データブロック情報に格納される前記データブロックを検索するための検索インデックスと、前記データブロックを一時的に保存するバッファとを格納し、

前記検索インデックスは、前記データブロックから算出された比較値及び前記データブロック情報に格納される前記データブロックへアクセスするためのポインタ値の組から構成される要素を複数含み、

前記データブロック情報は、前記ポインタ値及び前記データブロックから構成されるエントリを複数含み、

前記転送制御部は、

前記端末からデータを受信した場合、前記受信したデータから前記データブロックを生

成するための分割処理を実行し、

前記分割処理によって生成された第1データブロックから、前記検索インデックスを参照するための検索値及び参照値を算出し、

前記参照値に基づいて、前記検索インデックスに含まれる複数の要素の中から参照する少なくとも一つの要素を特定し、

前記特定された少なくとも一つの要素に含まれる前記比較値と前記検索値とを比較することによって、前記第1データブロックが前記データブロック情報に格納されているか否かを判定し、

前記第1データブロックが前記データブロック情報に格納されていると判定された場合、前記第1データブロックの識別情報を前記他の転送装置に送信し、前記バッファに前記データブロックを保存し、前記他の転送装置から当該第1データブロックの転送完了を通知する応答を受信した後、前記バッファから前記第1データブロックを削除し、

前記第1データブロックが前記データブロック情報に格納されていないと判定された場合、前記第1データブロックの識別情報及び前記第1データブロックを前記他の転送装置に送信し、前記バッファに前記データブロックを保存し、前記他の転送装置から当該第1データブロックの転送完了を通知する応答を受信した後、前記データブロック情報に前記第1データブロックを格納し、前記バッファから前記第1データブロックを削除し、

前記第1データブロックの識別情報を送信した前記他の転送装置から、前記第1データブロックの送信を要求するデータブロック要求を受信した場合、前記バッファに保存される前記第1データブロックを前記他の転送装置に送信することを特徴とする転送装置。

【請求項2】

請求項1に記載の転送装置であって、

前記比較値及び前記検索値は、ハッシュ値であり、

前記転送制御部は、第1ハッシュ関数及び第2ハッシュ関数を保持し、

前記第1ハッシュ関数及び前記第1データブロックを用いて前記参照値を算出し、

前記第2ハッシュ関数及び前記第1データブロックを用いて前記検索値を算出し、

前記検索値及び前記参照値を前記第1データブロックの識別情報として送信し、

前記検索値が前記比較値と一致する場合、前記第1データブロックが前記データブロック情報に格納されると判定することを特徴とする転送装置。

【請求項3】

請求項2に記載の転送装置であって、

前記転送制御部は、

前記第1データブロックの識別情報及び前記第1データブロックを送信した前記他の転送装置から当該第1データブロックの格納完了を通知する応答を受信した場合、前記検索値及び前記参照値を用いて前記ポインタ値を算出し、

前記参照値に基づいて前記検索インデックスを参照して、値を設定する要素を一つ選択し、

前記選択された要素に、前記検索値及び前記算出されたポインタ値を設定し、

前記データブロック情報にエントリを追加し、

前記追加されたエントリに、前記算出されたポインタ値及び前記第1データブロックを格納することを特徴とする転送装置。

【請求項4】

端末間で送受信されるデータを転送する複数の転送装置を備えるネットワークシステムであって、

前記複数の転送装置の各々は、

プロセッサ、前記プロセッサに接続される主記憶装置、前記プロセッサに接続される副記憶装置、及び前記プロセッサに接続されるネットワークインタフェースを有し、

前記端末間のデータの転送処理を制御する転送制御部を有し、

前記副記憶装置は、データを分割することによって生成されるデータブロックであって、前記端末間で過去に送信されたデータブロックを管理するデータブロック情報を格納し

10

20

30

40

50

前記主記憶装置は、前記データブロック情報に格納される前記データブロックを検索するための検索インデックスと、前記データブロックを一時的に保存するバッファとを格納し、

前記検索インデックスは、前記データブロックから算出された比較値及び前記データブロック情報に格納される前記データブロックへアクセスするためのポインタ値の組から構成される要素を複数含み、

前記データブロック情報は、前記ポインタ値及び前記データブロックから構成されるエントリを複数含み、

前記複数の転送装置は、第1検索インデックス及び第1データブロック情報を保持する第1転送装置、並びに、第2検索インデックス及び第2データブロック情報を保持する第2転送装置を含み、

前記第1転送装置は、

前記第1転送装置に接続する端末から送信データを受信した場合、前記送信データから前記データブロックを生成するための分割処理を実行し、

前記分割処理によって生成された第1データブロックから、前記検索インデックスを参照するための検索値と参照値を算出し、

前記参照値に基づいて、前記第1検索インデックスに含まれる複数の要素の中から参照する少なくとも一つの要素を特定し、

前記特定された少なくとも一つの要素に含まれる前記比較値と前記検索値とを比較することによって、前記第1データブロックが前記第1データブロック情報に格納されているか否かを判定し、

前記第1データブロックが前記第1データブロック情報に格納されていると判定された場合、前記検索値及び前記参照値を前記第1データブロックの識別情報として前記第2転送装置に送信し、前記バッファに前記データブロックを保存し、前記第2転送装置から当該第1データブロックの転送完了を通知する応答を受信した後、前記バッファから前記第1データブロックを削除し、

前記第1データブロックが前記第1データブロック情報に格納されていないと判定された場合、前記第1データブロックの識別情報及び前記第1データブロックを前記第2転送装置に送信し、前記バッファに前記データブロックを保存し、前記第2転送装置から当該第1データブロックの転送完了を通知する応答を受信した後、前記データブロック情報に前記第1データブロックを格納し、前記バッファから前記第1データブロックを削除し、

前記第1データブロックの識別情報を送信した前記第2転送装置から、前記第1データブロックの送信を要求するデータブロック要求を受信した場合、前記バッファに保存される前記第1データブロックを前記第2転送装置に送信し、

前記第2転送装置は、

前記第1データブロックの識別情報を受信した場合、前記参照値に基づいて、前記第2検索インデックスに含まれる複数の要素の中から参照する少なくとも一つの要素を特定し

前記特定された少なくとも一つの要素に含まれる前記比較値と前記第1データブロックの識別情報に含まれる前記検索値とを比較することによって、前記第1データブロックが前記第2データブロック情報に格納されているか否かを判定し、

前記第1データブロックが前記第2データブロック情報に格納されていると判定された場合、前記第2データブロック情報から前記第1データブロックを読み出して、前記第1データブロックを前記第2転送装置に接続される端末に送信し、前記第1データブロックの転送完了を通知する応答を前記第1転送装置に送信し、

前記第1データブロックが前記第2データブロック情報に格納されていないと判定された場合、前記データブロック要求を前記第1転送装置に送信することを特徴とするネットワークシステム。

【請求項5】

10

20

30

40

50

請求項 4 に記載のネットワークシステムであって、
前記第 1 転送装置は、
第 1 ハッシュ関数及び第 2 ハッシュ関数を保持し、
前記第 1 ハッシュ関数及び前記第 1 データブロックを用いて前記参照値を算出し、
前記第 2 ハッシュ関数及び前記第 1 データブロックを用いて前記検索値を算出し、
前記検索値が前記比較値と一致する場合、前記第 1 データブロックが前記データブロッ
ク情報に格納されると判定することを特徴とするネットワークシステム。

【請求項 6】

請求項 5 に記載のネットワークシステムであって、
前記第 1 転送装置は、
前記第 1 データブロックの識別情報及び前記第 1 データブロックを送信した前記第 2 転
送装置から当該第 1 データブロックの格納完了を通知する応答を受信した場合、前記検索
値及び前記参照値を用いて前記ポインタ値を算出し、
前記参照値に基づいて前記検索インデックスを参照して、値を設定する要素を一つ選択
し、
前記選択された要素に、前記検索値及び前記算出されたポインタ値を設定し、
前記データブロック情報にエントリを追加し、
前記追加されたエントリに、前記算出されたポインタ値及び前記第 1 データブロックを
格納することを特徴とするネットワークシステム。

10

【請求項 7】

請求項 5 に記載のネットワークシステムであって、
前記第 2 転送装置は、
前記第 1 データブロック及び前記第 1 データブロックの識別情報を受信した場合、前記
検索値及び前記参照値を用いて前記ポインタ値を算出し、
前記参照値に基づいて前記検索インデックスを参照して、値を設定する要素を選択し、
前記選択された要素に、前記検索値及び前記算出されたポインタ値を設定し、
前記データブロック情報にエントリを追加し、
前記追加されたエントリに、前記算出されたポインタ値及び前記第 1 データブロックを
格納することを特徴とするネットワークシステム。

20

【発明の詳細な説明】

【技術分野】

【0001】

本発明は、データの転送制御に関する。

【背景技術】

【0002】

クラウドコンピューティングの普及によって、ネットワークを介してデータを利用する機会が増えている。

【0003】

ネットワーク帯域が制限されている場合、データへのアクセス時間がかかり、利用者の快適性が損なわれるという問題がある。また、同一のデータに対して繰り返し、アクセスが行われることによって、システム全体の通信量が増加し、ネットワーク帯域を圧迫するという問題もある。

40

【0004】

本技術分野の背景技術として、特許文献 1 に記載の技術が知られている。特許文献 1 には、通信データを転送する際に、`persistent segment store` に保存されているセグメントを `reference` に置き換えて転送することが記載されている。

【0005】

また、特許文献 2 には、「ネットワークの利用度が利用度閾値に達したまたはこれを超過したときに、暗号化ハッシュ関数を適用して通信トラフィックにハッシュ・ダイジェス

50

トを組み込み、データの完全性点検およびデータの重複排除を提供し、これによって、送信側ネットワーク・デバイスと受信側ネットワーク・デバイスとの間で伝送されるデータの量を低減する」ことが開示されている。

【先行技術文献】

【特許文献】

【0006】

【特許文献1】米国特許出願公開第2013/0246508号明細書

【特許文献2】特表2016-517551号公報

【発明の概要】

【発明が解決しようとする課題】

【0007】

しかし、persistent segment storeが大きい場合、主記憶装置に格納できないため、二次記憶装置に保存されるpersistent segment storeからセグメントを検索する必要がある。一般に、二次記憶装置はアクセス速度が遅いため、通信データの転送速度が遅くなる。

【課題を解決するための手段】

【0008】

本願において開示される発明の代表的な一例を示せば以下の通りである。すなわち、端末間で送受信されるデータを転送する転送装置であって、プロセッサ、前記プロセッサに接続される主記憶装置、前記プロセッサに接続される副記憶装置、及び前記プロセッサに接続されるネットワークインタフェースを備え、前記端末間のデータの転送処理を制御する転送制御部を有し、ネットワークを介して他の転送装置と接続され、前記副記憶装置は、データを分割することによって生成されるデータブロックであって、前記端末間で過去に送信されたデータブロックを管理するデータブロック情報を格納し、前記主記憶装置は、前記データブロック情報に格納される前記データブロックを検索するための検索インデックスと、前記データブロックを一時的に保存するバッファとを格納し、前記検索インデックスは、前記データブロックから算出された比較値及び前記データブロック情報に格納される前記データブロックへアクセスするためのポインタ値の組から構成される要素を複数含み、前記データブロック情報は、前記ポインタ値及び前記データブロックから構成されるエントリを複数含み、前記転送制御部は、前記端末からデータを受信した場合、前記受信したデータから前記データブロックを生成するための分割処理を実行し、前記分割処理によって生成された第1データブロックから、前記検索インデックスを参照するための検索値及び参照値を算出し、前記参照値に基づいて、前記検索インデックスに含まれる複数の要素の中から参照する少なくとも一つの要素を特定し、前記特定された少なくとも一つの要素に含まれる前記比較値と前記検索値とを比較することによって、前記第1データブロックが前記データブロック情報に格納されているか否かを判定し、前記第1データブロックが前記データブロック情報に格納されていると判定された場合、前記第1データブロックの識別情報を前記他の転送装置に送信し、前記バッファに前記データブロックを保存し、前記他の転送装置から当該第1データブロックの転送完了を通知する応答を受信した後、前記バッファから前記第1データブロックを削除し、前記第1データブロックが前記データブロック情報に格納されていないと判定された場合、前記第1データブロックの識別情報及び前記第1データブロックを前記他の転送装置に送信し、前記バッファに前記データブロックを保存し、前記他の転送装置から当該第1データブロックの転送完了を通知する応答を受信した後、前記データブロック情報に前記第1データブロックを格納し、前記バッファから前記第1データブロックを削除し、前記第1データブロックの識別情報を送信した前記他の転送装置から、前記第1データブロックの送信を要求するデータブロック要求を受信した場合、前記バッファに保存される前記第1データブロックを前記他の転送装置に送信することを特徴とする。

【発明の効果】

【0009】

10

20

30

40

50

本発明によれば、検索インデックスを用いることによって、データブロックの検索を高速に行うことができる。前述した以外の課題、構成及び効果は、以下の実施例の説明によって明らかにされる。

【図面の簡単な説明】

【0010】

【図1】実施例1のネットワークシステムの構成例を示す図である。

【図2】実施例1の転送装置のハードウェア構成及びソフトウェア構成の一例を示す図である。

【図3】実施例1の転送装置のソフトウェア構成の論理的な接続関係を示す図である。

【図4】実施例1の転送装置がLANを介して送受信するTCPパケットのフォーマットの一例を示す図である。 10

【図5】実施例1の転送装置がWANを介して送受信するTCPパケットのフォーマットの一例を示す図である。

【図6】実施例1のフロー情報の一例を示す図である。

【図7】実施例1の送信検索インデックスの一例を示す図である。

【図8】実施例1の送信データブロック情報の一例を示す図である。

【図9A】実施例1のデータ転送処理の流れを説明するシーケンス図である。

【図9B】実施例1のデータ転送処理の流れを説明するシーケンス図である。

【図9C】実施例1のデータ転送処理の流れを説明するシーケンス図である。

【図10】実施例1の転送制御部がLANからデータを受信した場合に実行する処理を説明するフローチャートである。 20

【図11】実施例1の転送制御部がデータ分割部から応答を受け取った場合に実行する処理を説明するフローチャートである。

【図12】実施例1の転送制御部がタイマに設定された時間が経過した場合に実行する処理を説明するフローチャートである。

【図13】実施例1の転送制御部が他の転送装置から応答を受信した場合に実行する処理を説明するフローチャートである。

【図14】実施例1のデータ分割部が実行する処理を説明するフローチャートである。

【図15】実施例1の重複排除部が転送制御部によって呼び出された場合に実行する処理を説明するフローチャートである。 30

【図16】実施例1の重複排除部が転送制御部から更新指示を受け付けた場合に実行する処理を説明するフローチャートである。

【図17】実施例1の転送制御部がWANからデータを受信した場合に実行する処理を説明するフローチャートである。

【図18】実施例1の重複復元部が転送制御部によって呼び出された場合に実行する処理を説明するフローチャートである。

【図19】実施例1の重複排除部が実行する処理の具体例を示す図である。

【図20】実施例2の転送装置のハードウェア構成及びソフトウェア構成の一例を示す図である。

【図21】実施例2の転送装置のソフトウェア構成の論理的な接続関係を示す図である。 40

【図22】実施例2のフロー情報の一例を示す図である。

【図23】実施例2のデータの分割方法の概念を示す図である。

【図24】実施例2の副検索インデックスの一例を示す図である。

【図25】実施例2の転送制御部がLANからデータを受信した場合に実行する処理を説明するフローチャートである。

【図26】実施例2の重複排除部が転送制御部から副データブロックの検出指示を受け付けた場合に実行する処理を説明するフローチャートである。

【図27】実施例2の重複排除部が転送制御部によって呼び出された場合に実行する処理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】 50

【0011】

以下、添付図面を参照して本発明の実施例を説明する。各図において共通の構成については同一の参照符号が付されている。

【実施例1】

【0012】

実施例1では、本発明の基本的な一例として、転送装置が、二次記憶装置に格納されるデータブロック情報からデータブロックを高速に検索する転送制御方法を説明する。これによって、データの高速な転送を実現できる。本実施例では、TCP/IP通信におけるデータ転送を一例として示す。ただし、本発明は、データ転送に用いる通信プロトコルに限定されない。

10

【0013】

図1は、実施例1のネットワークシステムの構成例を示す図である。

【0014】

ネットワークシステムは、複数の端末100、複数の転送装置110、及び各装置を接続するネットワークから構成される。なお、本発明は、ネットワークシステムに含まれる装置及びネットワークの数には限定されない。

【0015】

端末100及び転送装置110は、LAN(Local Area Network)140を介して互いに接続される。また、複数の転送装置110は、WAN(Wide Area Network)を介して互いに接続される。各ネットワークの接続方式は、無線及び有線のいずれもよい。LAN140及びWAN120には、中継装置130が一つ以上含まれる。

20

【0016】

なお、各装置を接続するネットワークの種別に限定されず、専用回線及びインターネット等のネットワークでもよい。また、端末100及び転送装置110は、直接接続されてもよい。複数の転送装置110は、直接接続されてもよい。

【0017】

図1では、二つの端末100-1、100-2が、LAN140-1、140-2、転送装置110-1、110-2、及びWAN120を介して、互いに通信を行う。

【0018】

端末100は、通信を行う装置である。端末100は、物理計算機及びプロキシ装置等が考えられる。端末100は、図示しない演算装置、主記憶装置、及びネットワークインタフェースを有する。なお、端末100は、仮想計算機を用いて実現してもよい。

30

【0019】

中継装置130は、LAN140及びWAN120等のネットワークを構成する装置であり、スイッチ、ルータ、及びゲートウェイ装置等が考えられる。中継装置130は、図示しない演算装置、主記憶装置、及びネットワークインタフェースを有する。

【0020】

転送装置110は、端末100間のTCP通信の終端及び中継を行う装置であり、端末100間で送受信されるデータを転送する。本実施例の転送装置110は、フロー単位で通信を制御するものとする。

40

【0021】

端末100-1と端末100-2との間のTCP通信150において、転送装置110-1は、TCP通信150を終端し、LAN140-1を介して端末100-1との間でTCP通信151を行う。なお、TCP通信151では、転送装置110-1は、TCP通信150において端末100-2が用いるIPアドレス及びポート番号を、転送装置110-1のIPアドレス及びポート番号として用いる。同様に、TCP通信150において、転送装置110-2は、TCP通信150を終端し、LAN140-2を介して端末100-2とTCP通信153を行う。TCP通信153では、転送装置110-2は、TCP通信150において端末100-1が用いるIPアドレス及びポート番号を、転送

50

装置 110 - 2 の IP アドレス及びポート番号として用いる。

【 0022 】

また、転送装置 110 - 1 及び転送装置 110 - 2 は、WAN120 を介して、TCP 通信 152 を行う。TCP 通信 152 では、転送装置 110 - 1 は、端末 100 - 1 の IP アドレス及びポート番号を用い、また、転送装置 110 - 2 は、端末 100 - 2 の IP アドレス及びポート番号を用いる。

【 0023 】

前述のような TCP 通信によって、端末 100 - 1 及び端末 100 - 2 は、端末 100 間で直接通信する TCP 通信 150 を行っている場合と同様の動作を行う。

【 0024 】

また、本実施例の転送装置 110 は、重複したデータの転送を排除するように通信を制御する。具体的には、転送装置 110 は、送信したデータ及び受信したデータを格納する情報を管理する。転送元の端末 100 から過去に送信されたデータと同一のデータを受信した場合、転送装置 110 は当該データの識別情報を送信する。データの識別情報を受信した場合、転送装置 110 は、当該データの識別情報に一致するデータを情報から読み出し、転送先の端末 100 に読み出されたデータを送信する。

【 0025 】

前述したようなデータの転送制御を行うことによって、WAN120 へのデータ転送量が削減される。したがって、通信コストの削減及び端末 100 間の通信速度の高速化を実現できる。

【 0026 】

以下の説明では、ネットワークシステム内で過去に送信されたデータと同一のデータの転送を排除する処理を「重複排除」とも記載する。また、以下の説明では、ネットワークシステム内で過去に送信されたデータを転送装置 110 が管理する情報から取得する処理を「重複復元」とも記載する。

【 0027 】

図 2 は、実施例 1 の転送装置 110 のハードウェア構成及びソフトウェア構成の一例を示す図である。図 3 は、実施例 1 の転送装置 110 のソフトウェア構成の論理的な接続関係を示す図である。

【 0028 】

転送装置 110 は、演算装置 201、主記憶装置 202、二次記憶装置 203、及び複数のネットワークインタフェース 204 を有する。転送装置 110 が有するハードウェアは、システムバス 205 を介して互いに接続され、システムバス 205 を介して互いに通信を行う。

【 0029 】

なお、転送装置 110 が有する各ハードウェアの数は二つ以上でもよい。また、各ハードウェアは、複数のシステムバス 205 を介して、互いに接続されてもよい。また、各ハードウェアは、直接接続されてもよい。

【 0030 】

なお、前述したハードウェアは、仮想的なハードウェアを用いて実現してもよい。例えば、一つのネットワークインタフェース 204 上に複数の論理的なネットワークインタフェースを構成し、論理的なネットワークインタフェースを用いることもできる。

【 0031 】

なお、転送装置 110 は、入出力装置を有してもよい。入出力装置には、キーボード、マウス、タッチパネル、及びディスプレイ等が含まれる。

【 0032 】

演算装置 201 は、主記憶装置 202 に格納されるプログラムを実行する。演算装置 201 は、CPU (Central Processing Unit) 及び GPU (Graphics Processing Unit) 等が考えられる。演算装置 201 がプログラムを実行することによって転送装置 110 が有する機能を実現できる。以下の説明

10

20

30

40

50

では、機能部（モジュール）を主語に処理を説明する場合、演算装置 201 が当該機能部を実現するプログラムを実行していることを示す。

【0033】

主記憶装置 202 は、演算装置 201 が実行するプログラム及び当該プログラムによって使用される情報を格納する。主記憶装置 202 は、また、プログラムが使用するワークエリア及びバッファ等の一時記憶領域を含む。主記憶装置 202 は、メモリ等が考えられる。

【0034】

二次記憶装置 203 は、永続的にデータを格納する。また、二次記憶装置 203 は、主記憶装置 202 に格納できないデータを格納する一時記憶領域を含む。二次記憶装置 203 は、HDD (Hard Disk Drive) 及び SSD (Solid State Drive) 等が考えられる。

10

【0035】

主記憶装置 202 に格納されるプログラム及び情報は、二次記憶装置 203 に格納されてもよい。この場合、演算装置 201 は、二次記憶装置 203 から、プログラム及び情報を読み出し、主記憶装置 202 にプログラム及び情報をロードする。

【0036】

ネットワークインタフェース 204 は、ネットワークと接続するためのインタフェースである。本実施例では、転送装置 110 は、ネットワークインタフェース 204 - 1 を介して LAN 140 と接続し、また、ネットワークインタフェース 204 - 2 を介して WAN 120 と接続する。

20

【0037】

ここで、主記憶装置 202 及び二次記憶装置 203 に格納されるプログラム及び情報に吹いて説明する。

【0038】

主記憶装置 202 は、プログラムとして、重複排除部 211、重複復元部 212、転送制御部 213、プロトコル処理部 214、及びデータ分割部 215 を実現するプログラムを格納する。主記憶装置 202 は、情報として、フロー情報 217、送信検索インデックス 218、及び受信検索インデックス 219 を格納する。また、主記憶装置 202 は、一時記憶領域としてデータブロックバッファ 216 を含む。

30

【0039】

二次記憶装置 203 は、送信データブロック情報 220 及び受信データブロック情報 221 を格納する。

【0040】

フロー情報 217 は、フローを管理するための情報である。フロー情報 217 の詳細は、図 6 を用いて説明する。

【0041】

送信データブロック情報 220 は、重複排除を行う場合に使用されるデータブロックを格納する。受信データブロック情報 221 は、重複復元を行う場合に使用されるデータブロックを格納する。送信データブロック情報 220 及び受信データブロック情報 221 の詳細は、図 8 を用いて説明する。

40

【0042】

送信検索インデックス 218 は、送信データブロック情報 220 に格納されるデータブロックを検索する場合に使用されるインデックスである。受信検索インデックス 219 は、受信データブロック情報 221 に格納されるデータブロックを検索する場合に使用されるインデックスである。送信検索インデックス 218 及び受信検索インデックス 219 の詳細は、図 7 を用いて説明する。

【0043】

データブロックバッファ 216 は、データブロックを一時的に格納する記憶領域である。

50

【0044】

転送制御部213は、転送装置110全体を制御する。転送制御部213は、内部変数として更新フラグを保持する。転送制御部213は、ネットワークインタフェース204-1を介してLAN140からデータを受信した場合、重複排除部211及びデータ分割部215等と連携して、重複排除を行い、ネットワークインタフェース204-2を介してWAN120にデータを転送する。転送制御部213は、ネットワークインタフェース204-2を介してWAN120からデータを受信した場合、重複復元部212等と連携して、重複復元を行い、ネットワークインタフェース204-1を介してLAN140にデータを転送する。転送制御部213が実行する処理の詳細は後述する。

【0045】

プロトコル処理部214は、ネットワークインタフェース204を管理し、所定の通信プロトコルに基づく通信処理を実行する。本実施例の通信プロトコルは、TCP/IPプロトコルである。また、プロトコル処理部214は、アプリケーション等に対して通信を行うためのAPI(Application Programming Interface)を提供する。例えば、プロトコル処理部214は、送信アドレス及び送信ポート、並びに受信アドレス及び受信ポートを転送制御部213に通知する。また、プロトコル処理部214は、ネットワークインタフェース204を介してデータを受信した場合、転送制御部213にデータを出力し、転送制御部213からデータを受信した場合、ネットワークインタフェース204を介してデータをネットワークに送信する。

【0046】

なお、プロトコル処理部214は、OS(Operating System)に含まれるTCP/IP機能として実現されてもよいし、専用のプログラムとして実現されてもよい。

【0047】

データ分割部215は、所定の分割方法(プロトコル)に基づいて、データを分割することによって複数のデータブロックを生成する。データ分割部215が実行する処理の詳細は後述する。

【0048】

重複排除部211は、送信検索インデックス218を用いて重複排除を行う。また、重複排除部211は、送信検索インデックス218及び送信データブロック情報220を管理する。重複排除部211が実行する処理の詳細は後述する。

【0049】

重複復元部212は、受信検索インデックス219を用いて重複復元を行う。また、重複復元部212は、受信検索インデックス219及び受信データブロック情報221を管理する。重複復元部212が実行する処理の詳細は後述する。

【0050】

なお、重複排除部211は、第1ハッシュ関数及び第2ハッシュ関数を保持する。ハッシュ関数は、例えば、md5及びsha1等が考えられる。

【0051】

なお、前述した各機能部は、複数の機能部を一つの機能部に集約してもよいし、また、一つの機能部を複数の機能部に分割してもよい。例えば、重複排除部211、重複復元部212、及びデータ分割部215を、転送制御部213に集約してもよい。

【0052】

なお、前述した各機能部は、ソフトウェアとして実現されているが、これらの機能の一部又は全てを演算装置201及びネットワークインタフェース204等のハードウェアを用いて実現されてもよい。

【0053】

本実施例では、送信検索インデックス218及び受信検索インデックス219を主記憶装置202に格納することによって、データブロックの検索処理を高速に行うことができる。また、後述するように、送信検索インデックス218及び受信検索インデックス21

10

20

30

40

50

9のデータ構造を工夫することによって、検索処理の高速化及び検索精度の向上を実現できる。

【0054】

次に、図4及び図5を用いてTCP通信に使用されるTCPパケットについて説明する。

【0055】

図4は、実施例1の転送装置110がLAN140を介して送受信するTCPパケット400のフォーマットの一例を示す図である。

【0056】

TCPパケット400は、MACヘッダ410、IPヘッダ420、IPオプションヘッダ430、TCPヘッダ440、TCPオプションヘッダ450、及びTCPペイロード460を含む。

10

【0057】

MACヘッダ410は、DMAC411、SMAC412、TPID413、TCI414、及びType415を含む。DMAC411は、宛先MACアドレスを表す。SMAC412は、送信元MACアドレスを表す。TPID413は、タグ付きフレームであること及びタグの種類を表す。TCI414は、タグの情報を表す。Type415は、MACフレームタイプを表す。

【0058】

TCI414は、さらに、PCP416、CFI417、及びVID418を含む。PCP416は優先度を表す。CFI417は、MACアドレスが正規フォーマットであるかを表す。VID418は、VLANのIDを表す。なお、VLANが使用されていないネットワークの場合、TPID413及びTCI414は存在しない。この場合、転送装置110は、VID418が「0」であるものとして処理を行う。

20

【0059】

IPヘッダ420は、IPlength421、protocol422、SIP423、及びDIP424を含む。IPlength421は、MACヘッダ410を除くパケット長を表す。protocol422は、プロトコル番号を表す。SIP423は、送信元IPアドレスを表す。DIP424は、宛先IPアドレスを表す。

【0060】

IPオプションヘッダ430は、0又は複数のIPオプション情報431を含む。

30

【0061】

TCPヘッダ440は、src.port441、dst.port442、SEQ443、ACK444、flag445、tcphlen446、及びwin_size447を含む。src.port441は、送信元ポート番号を表す。dst.port442は、宛先ポート番号を表す。SEQ443は、送信シーケンス番号を表す。ACK444は、受信シーケンス番号を表す。flag445は、TCPフラグ番号を表す。tcphlen446は、TCPのヘッダ長を表す。win_size447は、対向装置へ通知する広告ウィンドウサイズを表す。

【0062】

TCPオプションヘッダ450は、0又は複数のオプションを含む。例えば、optionkind451、optionlength452、及びオプション情報453等のオプションが含まれる。optionkind451は、オプションの種別を表す。optionlength452は、オプション長を表す。オプション情報453は、オプションの種類に応じた情報を表す。

40

【0063】

図5は、実施例1の転送装置110がWAN120を介して送受信するTCPパケット500のフォーマットの一例を示す図である。

【0064】

TCPパケット500は、TCPパケット400と同様にMACヘッダ410、IPヘ

50

ッダ420、IPオプションヘッダ430、TCPヘッダ440、及びTCPオプションヘッダ450を含む。さらに、TCPパケット500は、符号化されたデータ部分に対応するDedupヘッダ510及びDedupペイロード520の組を一つ以上含む。

【0065】

TCPパケット500に含まれるMACヘッダ410、IPヘッダ420、TCPヘッダ440、及びTCPオプションヘッダ450は、TCPパケット400に含まれるものと同一である。TCPパケット500に含まれるIPオプションヘッダ430には、オプション情報の一つとしてDedup501を含む。Dedup501は、転送装置110によって重複排除が行われたパケットであることを表す。

【0066】

Dedupヘッダ510は、dedup.type511、dedup.data__len512、dedup.hash1 513、及びdedup.hash2 514を含む。Dedupヘッダ510に含まれる各情報のデータサイズは、一例であって、必要に応じてデータサイズを変更できる。

【0067】

dedup.type511は、Dedupペイロード520の種類を表す。dedup.data__len512は、Dedupペイロード520の長さを表す。dedup.hash1 513及びdedup.hash2 514は、データブロックを検索するために用いるハッシュ値を表す。本実施例では、dedup.hash1 513及びdedup.hash2 514が、データブロックの識別情報として用いられる。

【0068】

なお、dedup.type511には「1」、「2」、「3」のいずれかが設定される。「1」は、データブロックの識別情報のみを含むTCPパケット500であることを示す。「2」は、データブロック及びデータブロックの識別情報を含むTCPパケット500であることを示す。「3」は、分割されなかったデータを含むTCPパケット500であることを示す。

【0069】

Dedupペイロード520は、Dedup.type511の値に応じて、データブロック又はLAN140から受信したデータ等が格納される。なお、データブロックの識別情報を送信する場合、空のDedupペイロード520が付与されてもよいし、Dedupペイロード520が付与されなくてもよい。

【0070】

次に、図6、図7、及び図8を用いて、転送装置110が管理する情報等について説明する。

【0071】

図6は、実施例1のフロー情報217の一例を示す図である。

【0072】

フロー情報217は、フローを管理する情報を格納するエントリを含む。エントリは、フローID601、VLAN602、LAN-IP603、WAN-IP604、LAN-PORT605、WAN-PORT606、バッファ607、及びタイマ608から構成される。

【0073】

フローID601は、フローを一意に識別するための識別情報である。VLAN602は、フローが属するVLANの識別情報である。LAN-IP603及びLAN-PORT605は、LAN140を介して接続される端末100のIPアドレス及びポート番号である。WAN-IP604及びWAN-PORT606は、WAN120を介して接続される端末100のIPアドレス及びポート番号である。バッファ607は、LAN140から受信したデータをWAN120に転送する場合に、一時的にデータを格納するための領域である。タイマ608は、データ転送に用いるタイマ値である。

【0074】

10

20

30

40

50

図7は、実施例1の送信検索インデックス218の一例を示す図である。なお、受信検索インデックス219のデータ構造は、送信検索インデックス218と同一であるため、説明を省略する。

【0075】

送信検索インデックス218は、2次元配列のデータ構造である。図7では、N行4列の2次元配列となっている。なお、行及び列の数は一例であって、任意の値に設定できる。

【0076】

2次元配列の一つの要素701には、ハッシュ値X及びポインタ値Pが含まれる。ハッシュ値Xは、ハッシュ関数にデータブロックを入力することによって算出されるハッシュ値である。ハッシュ値Xは、データブロックが送信データブロック情報220に格納されているか否かを判定する場合に使用される。ポインタ値は、送信データブロック情報220内のデータブロックをアクセスするための値である。ポインタ値には、後述するファイル名801の値が設定される。

【0077】

後述するように、重複排除部211は、二つのハッシュ関数を用いて、送信検索インデックス218を参照する。

【0078】

図8は、実施例1の送信データブロック情報220の一例を示す図である。なお、受信データブロック情報221のデータ構造は、送信データブロック情報220と同一であるため、説明を省略する。

【0079】

送信データブロック情報220は、データブロックを格納するエントリを含む。エントリは、ファイル名801及びデータブロック802から構成される。本実施例では、個々のデータブロックがファイルとして管理される。

【0080】

ファイル名801は、データブロックを管理するファイルの識別情報である。本実施例では、ファイルの識別情報に数値を用いる。データブロック802は、データブロックそのものである。

【0081】

次に、図9A、図9B、及び図9Cを用いて、データ転送処理の一連の流れについて説明する。以下では、一例として、端末100-1から端末100-2へデータを転送する場合のデータ転送処理について説明する。なお、端末100-2から端末100-2へデータを転送する場合も同様の処理となる。

【0082】

図9A、図9B、及び図9Cは、実施例1のデータ転送処理の流れを説明するシーケンス図である。なお、TCP通信における確認応答等のTCPレイヤの処理については省略している。

【0083】

図9Aは、転送装置110-1及び転送装置110-2が、端末100-1から送信されるデータ(データブロック)を保持している場合のデータ転送処理の流れを示す。

【0084】

端末100-1が端末100-2へデータを送信する(ステップS901)。

【0085】

転送装置110-1は、LAN140-1を介して、端末100-1からデータを受信した場合、当該データをデータブロックに分割する(ステップS902)。転送装置110-1は、データブロックのハッシュ値に基づいて送信検索インデックス218を検索し、データブロックが送信データブロック情報220に格納されているか否かを判定する(ステップS903)。

【0086】

10

20

30

40

50

データブロックは送信データブロック情報 220 に格納されているため、転送装置 110 - 1 は、WAN 120 を介して、端末 100 - 2 にデータブロックの識別情報を送信する (ステップ S904)。このとき、転送装置 110 - 1 は、データブロックバッファ 216 にデータブロックを格納する。

【0087】

転送装置 110 - 2 は、データブロックの識別情報を受信した場合、データブロックの識別情報に基づいて受信検索インデックス 219 を検索し、データブロックが受信データブロック情報 221 に格納されているか否かを判定する (ステップ S905)。

【0088】

データブロックは受信データブロック情報 221 に格納されているため、転送装置 110 - 2 は、WAN 120 を介して、転送装置 110 - 1 に復元確認通知を送信する (ステップ S906)。なお、復元確認通知は、データブロックの転送が完了した旨を示す通知である。

【0089】

また、転送装置 110 - 2 は、データブロックの識別情報に基づいて受信データブロック情報 221 からデータブロックを取得し、LAN 140 - 2 を介して、端末 100 - 2 に取得されたデータブロックを端末 100 - 1 から送信されたデータとして送信する (ステップ S907)。

【0090】

転送装置 110 - 1 は、WAN 120 を介して、復元確認通知を受信した場合、データブロックバッファ 216 に格納されるデータブロックを削除する (ステップ S908)。

【0091】

なお、端末 100 - 1 から受信したデータは複数のデータブロックに分割される。したがって、転送装置 110 - 1 は、Dedup ヘッダ 510 及び Dedup ペイロード 520 の組を一つ以上含む TCP パケット 500 の送信を一回以上行う。また、転送装置 110 - 2 は、データブロックを送信する場合、複数のデータブロックを含む一つの TCP パケット 400 を送信してもよいし、一つのデータブロックを含む TCP パケット 400 を複数回送信してもよい。

【0092】

図 9B は、転送装置 110 - 1 が、端末 100 - 1 から送信されるデータ (データブロック) を保持していない場合のデータ転送処理の流れを示す。

【0093】

端末 100 - 1 が端末 100 - 2 へデータを送信する (ステップ S911)。

【0094】

転送装置 110 - 1 は、LAN 140 - 1 を介して、端末 100 - 1 からデータを受信した場合、当該データをデータブロックに分割する (ステップ S912)。転送装置 110 - 1 は、データブロックのハッシュ値に基づいて送信検索インデックス 218 を検索し、データブロックが送信データブロック情報 220 に格納されているか否かを判定する (ステップ S913)。

【0095】

データブロックは送信データブロック情報 220 に格納されていないため、転送装置 110 - 1 は、WAN 120 を介して、転送装置 110 - 2 にデータブロックを送信する (ステップ S914)。

【0096】

転送装置 110 - 2 は、データブロックを受信した場合、受信データブロック情報 221 に当該データブロックを格納する (ステップ S915)。転送装置 110 - 2 は、WAN 120 を介して、転送装置 110 - 1 に復元確認通知を送信する (ステップ S916)。また、転送装置 110 - 2 は、LAN 140 - 2 を介して、端末 100 - 2 に受信したデータブロックを端末 100 - 1 から送信されたデータとして送信する (ステップ S917)。

10

20

30

40

50

【 0 0 9 7 】

転送装置 1 1 0 - 1 は、WAN 1 2 0 を介して、復元確認通知を受信した場合、データブロックバッファ 2 1 6 に格納されるデータブロックを送信データブロック情報 2 2 0 に格納する（ステップ S 9 1 8 ）。

【 0 0 9 8 】

図 9 C は、転送装置 1 1 0 - 1 のみが、端末 1 0 0 - 1 から送信されるデータ（データブロック）を保持している場合のデータ転送処理の流れを示す。

【 0 0 9 9 】

ステップ S 9 2 1 からステップ S 9 2 4 の処理は、ステップ S 9 0 1 からステップ S 9 0 4 の処理と同一である。

10

【 0 1 0 0 】

転送装置 1 1 0 - 2 は、データブロックの識別情報を受信した場合、データブロックの識別情報に基づいて受信検索インデックス 2 1 9 を検索し、データブロックが受信データブロック情報 2 2 1 に格納されているか否かを判定する（ステップ S 9 2 5 ）。

【 0 1 0 1 】

データブロックは受信データブロック情報 2 2 1 に格納されていないため、転送装置 1 1 0 - 2 は、WAN 1 2 0 を介して、データブロック要求を送信する（ステップ S 9 2 6 ）。

【 0 1 0 2 】

転送装置 1 1 0 - 1 は、WAN 1 2 0 を介して、データブロック要求を受信した場合、WAN 1 2 0 を介して、転送装置 1 1 0 - 2 にデータブロックバッファ 2 1 6 に格納されるデータブロックを送信する（ステップ S 9 2 7 ）。

20

【 0 1 0 3 】

転送装置 1 1 0 - 2 は、WAN 1 2 0 を介して、データブロックを受信した場合、受信データブロック情報 2 2 1 に当該データブロックを格納する（ステップ S 9 2 8 ）。

転送装置 1 1 0 - 2 は、WAN 1 2 0 を介して、転送装置 1 1 0 - 1 に復元確認通知を送信する（ステップ S 9 2 9 ）。

また、転送装置 1 1 0 - 2 は、LAN 1 4 0 - 2 を介して、端末 1 0 0 - 2 に受信したデータブロックを端末 1 0 0 - 1 から送信されたデータとして送信する（ステップ S 9 3 0 ）。

【 0 1 0 4 】

転送装置 1 1 0 - 1 は、WAN 1 2 0 を介して、復元確認通知を受信した場合、データブロックバッファ 2 1 6 に格納されるデータブロックを削除する（ステップ S 9 3 1 ）。

30

【 0 1 0 5 】

転送装置 1 1 0 - 1 は、復元確認通知を受信するまで、データブロックバッファ 2 1 6 にデータブロックを格納するため、転送装置 1 1 0 - 2 からのデータブロック要求を受信した場合、即座に転送装置 1 1 0 - 2 にデータブロックを送信できる。

【 0 1 0 6 】

また、図 9 B に示すように、転送装置 1 1 0 - 1 は、復元確認通知を受信した後に、送信データブロック情報 2 2 0 にデータブロックを格納する。送信データブロック情報 2 2 0 へのデータブロックの格納順番が逆の場合、転送装置 1 1 0 - 2 にデータブロックが登録されていない可能性がある。そのため、図 9 C に示すようなデータ転送処理が発生する可能性がある。一方、図 9 B に示すような処理手順にすることによって、図 9 C に示すようなデータ転送処理が発生する可能性を低くすることができる。図 9 C に示すデータ転送処理では、転送装置 1 1 0 間の通信回数が増えることから、通信の遅延が発生する。したがって、図 9 C に示すようなデータ転送処理の発生を防止することによって、ネットワークシステム全体の通信速度、通信品質等を向上させることができる。

40

【 0 1 0 7 】

次に、図 1 0 から図 1 8 を用いて、転送装置 1 1 0 の各機能部が実行する処理の詳細を説明する。

【 0 1 0 8 】

50

図10は、実施例1の転送制御部213がLAN140からデータを受信した場合に実行する処理を説明するフローチャートである。

【0109】

転送制御部213は、プロトコル処理部214から、LAN140から受信したデータを受け取った場合(ステップS1001)、フロー情報217にデータを格納する(ステップS1002)。

【0110】

具体的には、転送制御部213は、プロトコル処理部214から受け取ったIPアドレス、ポート番号、及びVLANのID等に基づいてフロー情報217を参照し、フローに対応するエントリを特定する。転送制御部213は、特定されたエントリのバッファ607に、受け取ったデータ(TCPペイロード460)を格納する。なお、バッファ607にデータが格納されている場合、転送制御部213は、当該データの後方に受け取ったデータを追加する。

10

【0111】

次に、転送制御部213は、データ分割部215を呼び出す(ステップS1003)。このとき、転送制御部213は、受け取ったデータをデータ分割部215に出力する。その後、転送制御部213は、処理を終了する。

【0112】

図11は、実施例1の転送制御部213がデータ分割部215から応答を受け取った場合に実行する処理を説明するフローチャートである。

20

【0113】

転送制御部213は、データ分割部215から受け取った応答が分割不可通知であるか否かを判定する(ステップS1101)。

【0114】

データ分割部215から受け取った応答が分割不可通知であると判定された場合、転送制御部213は、ステップS1106に進む。

【0115】

データ分割部215から受け取った応答が分割不可通知ではないと判定された場合、すなわち、データ分割部215からデータブロックを受け取った場合、転送制御部213は、ステップS1002において特定されたエントリのバッファ607に格納されるデータから当該データブロックを削除する(ステップS1102)。

30

【0116】

具体的には、転送制御部213は、ステップS1002において特定されたエントリのバッファ607に格納されるデータからデータブロックに一致するデータ列を削除する。

【0117】

次に、転送制御部213は、重複排除部211を呼び出す(ステップS1103)。転送制御部213は、重複排除部211から応答があるまで待ち状態に移行する。このとき、転送制御部213は、データブロックを重複排除部211に入力する。

【0118】

転送制御部213は、一定時間経過した後、重複排除部211から符号化されたデータを取得したか否かを判定する(ステップS1104)。

40

【0119】

重複排除部211から符号化されたデータを取得していないと判定された場合、転送制御部213は、待ち状態に移行し、一定時間経過した後にステップS1104に戻る。

【0120】

重複排除部211から符号化されたデータを取得したと判定された場合、転送制御部213は、WAN120を介して、符号化されたデータを送信し、また、データブロックバッファ216にデータブロックを格納する(ステップS1105)。

【0121】

具体的には、転送制御部213は、符号化されたデータをプロトコル処理部214に出

50

力する。なお、符号化されたデータとは、一つ以上のDedupヘッダ510及びDedupペイロード520の組を表す。

【0122】

ステップS1101の判定結果がYESの場合、又は、ステップS1105の処理が実行された後、転送制御部213は、ステップS1002において特定されたエントリのバッファ607にデータが残っているか否かを判定する(ステップS1106)。

【0123】

ステップS1002において特定されたエントリのバッファ607にデータが残っていると判定された場合、転送制御部213は、当該エントリのタイマ608に任意の時間を設定する(ステップS1107)。その後、転送制御部213は、処理を終了する。

10

【0124】

ステップS1002において特定されたエントリのバッファ607にデータが残っていないと判定された場合、転送制御部213は、当該エントリのタイマ608に「0」を設定する(ステップS1108)。その後、転送制御部213は、処理を終了する。

【0125】

図12は、実施例1の転送制御部213がタイマ608に設定された時間が経過した場合に実行する処理を説明するフローチャートである。

【0126】

転送制御部213は、各フローについてタイマをカウントし、タイマ608に設定された時間が経過したエントリを検出した場合、以下の処理を開始する。以下では、タイマ608に設定された時間が経過したエントリを対象エントリと記載する。

20

【0127】

転送制御部213は、対象エントリのバッファ607に格納されるデータの符号化を重複排除部211に指示する(ステップS1201)。転送制御部213は、重複排除部211から符号化されたデータを取得するまで待ち状態に移行する。なお、当該指示には、対象エントリのバッファ607に格納されるデータが含まれる。

【0128】

重複排除部211は、当該指示を受け付けた場合、dedup.type511に「3」を設定し、dedup.data_len512にバッファ607に格納されるデータの長さを設定し、dedup.hash1_513及びdedup.hash2_514にダミーの値として「0」を設定することによってDedupヘッダ510を生成する。また、重複排除部211は、Dedupペイロード520に対象エントリのバッファ607に格納されるデータを設定する。重複排除部211は、Dedupヘッダ510及びDedupペイロード520を転送制御部213に出力する。

30

【0129】

次に、転送制御部213は、重複排除部211から符号化されたデータを取得した場合、WAN120を介して、符号化されたデータを送信する(ステップS1202)。

【0130】

具体的には、転送制御部213は、符号化されたデータをプロトコル処理部214に出力する。これによって、Dedupヘッダ510及びDedupペイロード520を含むTCPパケット500が送信される。

40

【0131】

次に、転送制御部213は、対象エントリのタイマ608に「0」を設定する(ステップS1203)。その後、転送制御部213は、処理を終了する。

【0132】

このように、タイマ608に設定された時間が経過した場合に、転送装置110がバッファ607に格納されるデータを送信することによって、受信したデータが転送されない状態を防止できる。

【0133】

図13は、実施例1の転送制御部213が他の転送装置110から応答を受信した場合

50

に実行する処理を説明するフローチャートである。

【0134】

転送制御部213は、プロトコル処理部214から、WAN120から受信した他の転送装置110の応答を受け取った場合(ステップS1301)、受け取った応答がデータブロック要求であるか否かを判定する(ステップS1302)。

【0135】

受け取った応答がデータブロック要求であると判定された場合、転送制御部213は、データブロックバッファ216からデータブロックを読み出し、重複排除部211にデータブロックの符号化を指示する(ステップS1303)。転送制御部213は、重複排除部211から符号化されたデータブロックを取得するまで待ち状態に移行する。なお、当該指示には、データブロックが含まれる。

10

【0136】

重複排除部211は、当該指示を受け付けた場合、第1ハッシュ関数及び第2ハッシュ関数を用いて、変数 及び変数 を算出する。変数 及び変数 の算出方法は、図15を用いて説明する。重複排除部211は、`dedup.type511`に「2」を設定し、`dedup.data_len512`にデータブロックの長さを設定し、`dedup.hash1513`に変数 を設定し、`dedup.hash2514`に変数 を設定することによってDedupヘッダ510を生成する。また、重複排除部211は、Dedupペイロード520にデータブロックを設定する。重複排除部211は、Dedupヘッダ510及びDedupペイロード520を転送制御部213に出力する。

20

【0137】

次に、転送制御部213は、重複排除部211から符号化されたデータブロックを取得した場合、WAN120を介して、符号化されたデータブロックを送信する(ステップS1304)。その後、転送制御部213は、処理を終了する。

【0138】

具体的には、転送制御部213は、符号化されたデータブロックをプロトコル処理部214に出力する。これによって、Dedupヘッダ510及びDedupペイロード520を含むTCPパケット500が送信される。

【0139】

ステップS1302において、受け取った応答がデータブロック要求でないと判定された場合、転送制御部213は、更新フラグがONであるか否かを判定する(ステップS1305)。

30

【0140】

更新フラグがOFFであると判定された場合、転送制御部213は、ステップS1307に進む。

【0141】

更新フラグがONであると判定された場合、転送制御部213は、重複排除部211に更新指示を出力する(ステップS1306)。このとき、転送制御部213は、更新フラグをOFFに設定する。当該更新指示にはデータブロックが含まれる。重複排除部211が、更新指示を受け付けた場合に実行する処理は図16を用いて説明する。

40

【0142】

ステップS1305の判定結果がNOである場合、又は、ステップS1306の処理が実行された後、転送制御部213は、データブロックバッファ216からデータブロックを削除する(ステップS1307)。その後、転送制御部213は、処理を終了する。

【0143】

図14は、実施例1のデータ分割部215が実行する処理を説明するフローチャートである。

【0144】

ここでは、データ分割部215が、Content Defined Chunkingとして知られるアルゴリズムに基づいて、データを可変長のデータブロックに分割する

50

方法について説明する。Content Defined Chunkingでは、計算機が、固定長のウィンドウ（例えば、48バイト）をデータの先頭から1バイトずつスライドし、ウィンドウに含まれるデータのハッシュ値を用いて分割位置を決定する。

【0145】

データ分割部215は、転送制御部213から呼び出された場合、処理を開始する。まず、データ分割部215は、転送制御部213からLAN140から受信したデータを取得する（ステップS1401）。

【0146】

データ分割部215は、データの先頭に所定のサイズのウィンドウを配置する（ステップS1402）。データ分割部215は、ウィンドウに含まれるデータのハッシュ値を算出する（ステップS1403）。ハッシュ値を算出するためのハッシュ関数は、例えば、ローリングハッシュとして知られるハッシュ関数を用いる。

10

【0147】

次に、データ分割部215は、ハッシュ値が所定の条件を満たすか否かを判定する（ステップS1404）。

【0148】

例えば、2進法表記におけるハッシュ値の下10桁が全て1であるか否かが判定される。ハッシュ値が入力されたデータに応じて適切に散らばる場合、2進法表記におけるハッシュ値の下10桁が全て1になる確率は1024分の1である。このとき、データブロックの平均サイズは約1024バイトとなる。

20

【0149】

ハッシュ値が所定の条件を満たすと判定された場合、データ分割部215は、ウィンドウに含まれるデータを出力するデータブロックに決定し、転送制御部213に当該データブロックを応答として出力する（ステップS1405）。その後、転送制御部213は処理を終了する。

【0150】

ステップS1404において、ハッシュ値が所定の条件を満たさないと判定された場合、データ分割部215は、ウィンドウの移動が可能であるか否かを判定する（ステップS1406）。

【0151】

具体的には、データ分割部215は、ウィンドウの終端がLAN140から受信したデータの終端と一致するか否かが判定される。ウィンドウの終端がLAN140から受信したデータの終端と一致しない場合、データ分割部215は、ウィンドウの移動が可能であると判定する。

30

【0152】

ウィンドウの移動が不可能であると判定された場合、データ分割部215は、転送制御部213に分割不可通知を応答として出力する（ステップS1407）。その後、転送制御部213は処理を終了する。

【0153】

ウィンドウの移動が可能であると判定された場合、データ分割部215は、ウィンドウをデータの後方にスライドさせ（ステップS1408）、その後、ステップS1403に戻る。例えば、データ分割部215は、ウィンドウを1バイト分だけ後方にスライドさせる。

40

【0154】

図15は、実施例1の重複排除部211が転送制御部213によって呼び出された場合に実行する処理を説明するフローチャートである。

【0155】

重複排除部211は、転送制御部213からデータブロックを取得する（ステップS1501）。重複排除部211は、第1ハッシュ関数及び第2ハッシュ関数を用いて、変数及び変数を算出する（ステップS1502）。

50

【0156】

具体的には、重複排除部211は、第1ハッシュ関数にデータブロックを入力することによって算出された第1ハッシュ値を2次元配列の行数Nで除算し、その余りを変数として算出する。また、重複排除部211は、第2ハッシュ関数にデータブロックを入力することによって算出される第2ハッシュ値を変数として算出する。

【0157】

次に、重複排除部211は、送信検索インデックス218を参照し(ステップS1503)、ハッシュ値が変数に一致する要素701が存在するか否かを判定する(ステップS1504)。

【0158】

具体的には、重複排除部211は、行目の行に含まれる要素701を参照する要素701群として特定する。重複排除部211は、行目に含まれる各要素701に格納されるハッシュ値Xと、変数とを比較する。

【0159】

ハッシュ値Xが変数に一致する要素701が存在しないと判定された場合、重複排除部211は、ステップS1508に進む。

【0160】

ハッシュ値Xが変数に一致する要素701が存在すると判定された場合、重複排除部211は、送信データブロック情報220からデータブロックを読み出す(ステップS1505)。

【0161】

具体的には、重複排除部211は、ハッシュ値Xが変数に一致する要素701に含まれるポインタ値Pを取得し、送信データブロック情報220を参照して、当該ポインタ値Pがファイル名801に一致するエントリを検索する。重複排除部211は、検索されたエントリのデータブロック802からデータブロックを読み出す。

【0162】

次に、重複排除部211は、転送制御部213から取得したデータブロックと、送信データブロック情報220から読み出したデータブロックとが一致するか否かを判定する(ステップS1506)。

【0163】

二つのデータブロックが一致しないと判定された場合、重複排除部211は、ステップS1508に進む。

【0164】

二つのデータブロックが一致すると判定された場合、重複排除部211は、転送制御部213に、符号化されたデータブロックの識別情報を出力する(ステップS1507)。その後、重複排除部211は、処理を終了する。

【0165】

具体的には、重複排除部211は、`dedup.type511`に「1」を設定し、`dedup.data_len512`に「0」を設定し、`dedup.hash1513`に変数を設定し、`dedup.hash2514`に変数を設定することによってDedupヘッダ510を生成する。重複排除部211は、Dedupヘッダ510及びDedupペイロード520を転送制御部213に出力する。なお、Dedupペイロード520は空であるものとする。

【0166】

Dedupヘッダ510のサイズが28バイト、データブロックのサイズの平均値が4キロバイトである場合、データブロックの識別情報のみを送信する通信は、データブロックを送信する通信より通信量が100分の1以下となる。なお、前述の比較は、TCPヘッダ440等を除いたものである。

【0167】

ステップS1504の判定結果がNOである場合、又は、ステップS1506の判定結

10

20

30

40

50

果がNOである場合、重複排除部211は、転送制御部213の更新フラグをONに設定する(ステップS1508)。このとき、重複排除部211は、変数 及び変数 を一時的に保持する。

【0168】

次に、重複排除部211は、転送制御部213に、符号化されたデータブロックの識別情報及び符号化されたデータブロックを出力する(ステップS1509)。その後、重複排除部211は、処理を終了する。

【0169】

具体的には、重複排除部211は、`dedup.type511`に「2」を設定し、`dedup.data_len512`にデータブロックのサイズを設定し、`dedup.hash1513`に変数 を設定し、`dedup.hash2514`に変数 を設定することによってDedupヘッダ510を生成する。また、重複排除部211は、Dedupペイロード520にデータブロックを設定する。重複排除部211は、Dedupヘッダ510及びDedupペイロード520を転送制御部213に出力する。

【0170】

変数 は、行数Nで割った余りであるため、0からN-1までの値となる。したがって、任意のデータブロックの変数 は、他のデータブロックを用いて算出された変数 とN分の1の確率で一致する。送信検索インデックス218のデータサイズは行数Nの大きさに依存するため、行数Nの大きさには制限がある。一方、変数 は、第2ハッシュ値をそのまま入力したものであり、Nより十分大きな値を設定できる。

【0171】

例えば、変数 のサイズを20ビット、変数 のサイズを128ビット、ポインタ値Pのサイズを64ビットとした場合、送信検索インデックス218のサイズは、式(1)に示すように約100MBとなる。変数 のサイズを128ビット、変数 のサイズを20ビット、ポインタ値Pのサイズを64ビットとした場合、送信検索インデックス218のサイズは、式(2)に示すように約 1.2×10^{16} YBとなる。

【0172】

【数1】

$$2^{20} \times 4 \times 192 \text{ ビット} \approx 100\text{MB} \quad \dots(1)$$

【0173】

【数2】

$$2^{128} \times 4 \times 84 \text{ ビット} \approx 1.2 \times 10^{16} \text{ YB} \quad \dots(2)$$

【0174】

変数 のサイズを大きくすると、ハッシュ値の衝突が発生する確率が非常に低くなる。したがって、変数 とハッシュ値Xとが一致する場合、受信したデータブロックと、送信データブロック情報220から読み出したデータブロックとは、高い確率で一致する。この場合、変数 及びハッシュ値Xの比較のみを行えばよく、データブロック同士を比較する必要はない。そのため、ステップS1506の省略できる。

【0175】

ステップS1506の処理を省略した場合、二次記憶装置203へのアクセスする必要がないため、重複排除部211は、さらに高速に、データブロックを検索することができる。したがって、処理速度を重視する場合には、変数 のサイズを大きくし、ステップS1506の処理を行わないように設定すればよい。

【0176】

図16は、実施例1の重複排除部211が転送制御部213から更新指示を受け付けた場合に実行する処理を説明するフローチャートである。

【0177】

重複排除部211は、転送制御部213から更新指示を受け取った場合(ステップS1

10

20

30

40

50

601)、送信検索インデックス218及び送信データブロック情報220を更新する(ステップS1602)。その後、重複排除部211は、処理を終了する。具体的には、以下のような処理が実行される。

【0178】

重複排除部211は、ポインタ値Pに設定する値を式(3)を用いて算出する。なお、mは、第2ハッシュ関数の出力値の最大ビット数である。

【0179】

【数3】

$$\alpha \times 2^m + \beta \quad \dots(3)$$

10

【0180】

式(3)に示すように変数 α 及び変数 β に基づいて一意に定まる値をポインタ値に用いることによって、異なるデータブロックのポインタ値が同一となる確率を非常に小さくできる。

【0181】

重複排除部211は、送信検索インデックス218の i 行目を参照し、 i 行目に含まれる要素701の中に空の要素701が存在するか否かを判定する。空の要素701とは、ハッシュ値X及びポインタ値Pが設定されていない要素701を示す。

【0182】

空の要素701が存在する場合、重複排除部211は、当該要素701のハッシュ値Xに変数 α を設定し、ポインタ値Pに式(3)を用いて算出された値を設定する。また、重複排除部211は、送信データブロック情報220にエントリを追加し、追加されたエントリのファイル名801に式(3)を用いて算出された値を設定し、データブロック802に更新指示に含まれるデータブロックを格納する。

20

【0183】

空の要素701が存在しない場合、重複排除部211は、 i 行目に含まれる要素701の中から値を設定する要素701を一つ選択する。要素701を選択する方法は様々考えられる。例えば、ランダムに選択する方法、及びLRU(Least Recently Used)アルゴリズムに基づいて選択する方法等が考えられる。

【0184】

重複排除部211は、送信データブロック情報220を参照し、ファイル名801が選択された要素701のポインタ値Pに一致するエントリを検索し、検索されたエントリを削除する。重複排除部211は、選択された要素701のハッシュ値Xに変数 α を設定し、ポインタ値Pに式(3)を用いて算出された値を設定する。また、重複排除部211は、送信データブロック情報220にエントリを追加し、追加されたエントリのファイル名801に式(3)を用いて算出された値を設定し、データブロック802にデータブロックを格納する。以上がステップS1602の処理の説明である。

30

【0185】

図17は、実施例1の転送制御部213がWAN120からデータを受信した場合に実行する処理を説明するフローチャートである。

40

【0186】

転送制御部213は、プロトコル処理部214から、WAN120から受信したデータ(TCPパケット500)を受け取った場合(ステップS1701)、重複復元部212を呼び出す(ステップS1702)。このとき、転送制御部213は、受け取ったデータを重複復元部212に出力する。転送制御部213は、重複復元部212から応答を受け付けるまで待ち状態に移行する。

【0187】

転送制御部213は、重複復元部212からの応答に基づいて、データブロックを取得する必要があるか否かを判定する(ステップS1703)。

【0188】

50

具体的には、転送制御部 2 1 3 は、重複復元部 2 1 2 からの応答が「該当データ無し」であるか否かを判定する。重複復元部 2 1 2 からの応答が「該当データ無し」である場合、転送制御部 2 1 3 は、データブロックを取得する必要があると判定する。

【 0 1 8 9 】

データブロックを取得する必要があると判定された場合、転送制御部 2 1 3 は、LAN 1 4 0 を介して、復元されたデータを送信する（ステップ S 1 7 0 4 ）。

【 0 1 9 0 】

具体的には、転送制御部 2 1 3 は、復元されたデータをプロトコル処理部 2 1 4 に出力する。これによって、Dedup ペイロード 5 2 0 に設定されたデータが TCP ペイロード 4 6 0 に設定された TCP パケット 4 0 0 が送信される。

10

【 0 1 9 1 】

次に、転送制御部 2 1 3 は、WAN 1 2 0 を介して、復元確認通知を送信する（ステップ S 1 7 0 5 ）。その後、転送制御部 2 1 3 は、処理を終了する。

【 0 1 9 2 】

具体的には、転送制御部 2 1 3 は、プロトコル処理部 2 1 4 に復元確認通知の送信を指示する。なお、復元確認通知を送信する場合には、TCP パケット 4 0 0 及び TCP パケット 5 0 0 のいずれを用いてもよい。

【 0 1 9 3 】

ステップ S 1 7 0 3 において、データブロックを取得する必要があると判定された場合、転送制御部 2 1 3 は、WAN 1 2 0 を介して、データブロック要求を送信する（ステップ S 1 7 0 6 ）。その後、転送制御部 2 1 3 は、処理を終了する。

20

【 0 1 9 4 】

具体的には、転送制御部 2 1 3 は、プロトコル処理部 2 1 4 にデータブロック要求の送信を指示する。なお、データブロック要求を送信する場合には、TCP パケット 4 0 0 及び TCP パケット 5 0 0 のいずれを用いてもよい。

【 0 1 9 5 】

図 1 8 は、実施例 1 の重複復元部 2 1 2 が転送制御部 2 1 3 によって呼び出された場合に実行する処理を説明するフローチャートである。

【 0 1 9 6 】

重複復元部 2 1 2 は、転送制御部 2 1 3 からデータを取得し（ステップ S 1 8 0 1 ）、Dedup ヘッダ 5 1 0 に含まれる dedup . type 5 1 1 の値が「2」であるか否かを判定する（ステップ S 1 8 0 2 ）。

30

【 0 1 9 7 】

dedup . type 5 1 1 の値が「2」ではないと判定された場合、重複復元部 2 1 2 は、dedup . type 5 1 1 の値が「3」であるか否かを判定する（ステップ S 1 8 0 3 ）。

【 0 1 9 8 】

dedup . type 5 1 1 の値が「1」であると判定された場合、重複復元部 2 1 2 は、受信検索インデックス 2 1 9 を参照し、ハッシュ値 X が dedup . hash 2 5 1 4 の値に一致する要素 7 0 1 が存在するか否かを判定する（ステップ S 1 8 0 4 ）。

40

【 0 1 9 9 】

具体的には、重複復元部 2 1 2 は、(dedup . hash 1 5 1 3) 行目の行に含まれる要素 7 0 1 を参照する要素 7 0 1 群として特定する。重複復元部 2 1 2 は、(dedup . hash 1 5 1 3) 行目に含まれる各要素 7 0 1 に格納されるハッシュ値 X と、dedup . hash 2 5 1 4 の値とを比較する。

【 0 2 0 0 】

ハッシュ値 X が dedup . hash 2 5 1 4 の値に一致する要素 7 0 1 が存在すると判定された場合、重複復元部 2 1 2 は、受信データブロック情報 2 2 1 からデータブロックを読み出す（ステップ S 1 8 0 5 ）。

【 0 2 0 1 】

50

具体的には、重複復元部 212 は、ハッシュ値 X が `dedup.hash2 514` の値に一致する要素 701 からポインタ値 P を取得し、受信データブロック情報 221 を参照して、当該ポインタ値 P がファイル名 801 に一致するエントリを検索する。重複復元部 212 は、検索されたエントリのデータブロック 802 からデータブロックを読み出す。

【0202】

次に、重複復元部 212 は、転送制御部 213 に読み出されたブロックデータを応答として出力する (ステップ S1806)。その後、重複復元部 212 は、処理を終了する。

【0203】

ステップ S1804 において、ハッシュ値が `dedup.hash2 514` の値に一致する要素 701 が存在しないと判定された場合、重複復元部 212 は、転送制御部 213 に「該当データ無し」を応答として出力する (ステップ S1808)。その後、重複復元部 212 は、処理を終了する。

【0204】

ステップ S1803 において、`dedup.type511` の値が「3」と判定された場合、重複復元部 212 は、転送制御部 213 に、`Dedup` ペイロード 520 に設定されたデータを応答として出力する (ステップ S1807)。その後、重複復元部 212 は、処理を終了する。

【0205】

ステップ S1802 において、`dedup.type511` の値が「2」と判定された場合、重複復元部 212 は、受信検索インデックス 219 及び受信データブロック情報 221 を更新する (ステップ S1809)。具体的には、以下のような処理が実行される。

【0206】

重複復元部 212 は、ポインタ値 P に設定する値を式 (4) を用いて算出する。

【0207】

【数 4】

$$(\text{dedup.hash1 } 513) \times 2^m + (\text{dedup.hash2 } 514) \dots(4)$$

【0208】

式 (4) に示すように `dedup.hash1 513` 及び `dedup.hash2 514` に基づいて一意に定まる値をポインタ値に用いることによって、異なるデータブロックのポインタ値が同一となる確率を非常に小さくできる。

【0209】

重複復元部 212 は、受信検索インデックス 219 の (`dedup.hash1 513`) 行目を参照し、(`dedup.hash1 513`) 行目に含まれる要素 701 の中に空の要素 701 が存在するか否かを判定する。

【0210】

空の要素 701 が存在する場合、重複復元部 212 は、当該要素 701 のハッシュ値 X に `dedup.hash2 514` の値を設定し、ポインタ値 P に式 (4) を用いて算出された値を設定する。また、重複復元部 212 は、受信データブロック情報 221 にエントリを追加し、追加されたエントリのファイル名 801 に式 (4) を用いて算出された値を設定し、データブロック 802 にデータブロックを格納する。

【0211】

空の要素 701 が存在しない場合、重複復元部 212 は、(`dedup.hash1 513`) 行目に含まれる要素 701 の中から、値を設定する要素 701 を一つ選択する。要素 701 を選択する方法は重複排除部 211 が使用する方法と同一の方法とする。重複排除部 211 及び重複復元部 212 が同一の方法を用いることによって、送信データブロック情報 220 及び受信データブロック情報 221 の内容が一致する可能性が高くなる。これによって、データブロック要求を送信する回数を低減できる。

10

20

30

40

50

【 0 2 1 2 】

重複復元部 2 1 2 は、受信データブロック情報 2 2 1 を参照し、ファイル名 8 0 1 が選択された要素 7 0 1 のポインタ値 P に一致するエントリを検索し、検索されたエントリを削除する。重複復元部 2 1 2 は、選択された要素 7 0 1 のハッシュ値 X に `dedup.hash2_514` の値を設定し、ポインタ値 P に式 (4) を用いて算出された値を設定する。また、重複復元部 2 1 2 は、受信データブロック情報 2 2 1 にエントリを追加し、追加されたエントリのファイル名 8 0 1 に式 (4) を用いて算出された値を設定し、データブロック 8 0 2 にデータブロックを格納する。以上がステップ S 1 8 0 9 の処理の説明である。

【 0 2 1 3 】

次に、重複復元部 2 1 2 は、転送制御部 2 1 3 に、Dedup ペイロード 5 2 0 に設定されたデータブロックを応答として出力する (ステップ S 1 8 1 0)。その後、重複復元部 2 1 2 は、処理を終了する。

【 0 2 1 4 】

次に、図 1 9 を用いて、重複排除部 2 1 1 が実行する具体的な処理内容について説明する。

【 0 2 1 5 】

図 1 9 は、実施例 1 の重複排除部 2 1 1 が実行する処理の具体例を示す図である。ここで、送信検索インデックス 2 1 8 及び送信データブロック情報 2 2 0 は、それぞれ、図 1 9 に示すエントリが格納されるものとする。なお、送信検索インデックスの左側の数字は行数を表す。

【 0 2 1 6 】

まず、重複排除部 2 1 1 がデータブロック「`abc...z`」を受信した場合を例に説明する。なお、当該データブロックの変数 `は「6423」、変数 は「0x11121239」であるものとする。`

【 0 2 1 7 】

ステップ S 1 5 0 3 において、重複排除部 2 1 1 は、6 4 2 3 行目に含まれる四つの要素 7 0 1 のハッシュ値 X と、変数 `とを比較する。このとき、2 列目の要素 7 0 1 のハッシュ値 X と変数 とが一致するため、ステップ S 1 5 0 4 の判定結果は YES となる。`

【 0 2 1 8 】

ステップ S 1 5 0 5 では、重複排除部 2 1 1 は、2 列目の要素 7 0 1 のポインタ値 P の値「`p2`」を用いて、送信データブロック情報 2 2 0 からデータブロック「`abc...z`」を読み出す。

【 0 2 1 9 】

転送制御部 2 1 3 から取得したデータブロックと、送信データブロック情報 2 2 0 から読み出したデータブロックとは一致するため、ステップ S 1 5 0 6 の判定結果は YES となる。

【 0 2 2 0 】

したがって、重複排除部 2 1 1 は、符号化されたデータブロックの識別情報を出力する (ステップ S 1 5 0 7)。なお、Dedup ヘッド 5 1 0 の `dedup.hash1_513` には「`6242`」が設定され、`dedup.hash2_514` には「`0x11121239`」が設定される。

【 0 2 2 1 】

次に、重複排除部 2 1 1 がデータブロック「`012...9`」を受信した場合を例に説明する。なお、当該データブロックの変数 `は「312」、変数 は「0x121fd9」であるものとする。`

【 0 2 2 2 】

ステップ S 1 5 0 3 において、重複排除部 2 1 1 は、3 1 2 行目に含まれる四つの要素 7 0 1 のハッシュ値 X と、変数 `とを比較する。このとき、変数 に一致するハッシュ値 X を含む要素 7 0 1 は存在しないため、ステップ S 1 5 0 4 の判定結果は NO となる。`

10

20

30

40

50

【0223】

したがって、重複排除部211は、更新フラグをONに設定し(ステップS1508)、符号化されたデータブロックの識別情報及びデータブロックを出力する(ステップS1509)。なお、Dedupヘッダ510のdedup.hash1 513には「312」が設定され、dedup.hash2 514には「0x121f12d9」が設定される。また、Dedupペイロード520にはデータブロック「012...9」が設定される。

【0224】

重複排除部211は、転送制御部213から更新指示を受け取った場合、データブロック「012...9」から変数 及び変数 を算出し、また、式(3)を用いてポインタ値Pに設定する値を算出する。

10

【0225】

重複排除部211は、送信検索インデックス218の312行目を参照する。312行目には空の要素701が存在しないため、重複排除部211は、312行目に含まれる要素701の中から一つの要素を選択する。ここでは、2列目の要素701が選択されたものとする。

【0226】

重複排除部211は、送信データブロック情報220を参照して、ファイル名801がp6であるエントリを検索し、検索されたエントリを削除する。

【0227】

重複排除部211は、312行目の2列目の要素701のハッシュ値Xに変数 を設定し、また、ポインタ値Pに式(3)を用いて算出された値を設定する。また、重複排除部211は、送信データブロック情報220にエントリを追加し、追加されたエントリのファイル名801に式(3)を用いて算出された値を設定し、データブロック802にデータブロック「012...9」を設定する。

20

【0228】

二つのデータブロックを例に、送信検索インデックス218及び送信データブロック情報220の使用方法について説明した。なお、受信検索インデックス219及び受信データブロック情報221の使用方法も同様である。

【0229】

実施例1によれば、転送装置110は、送信検索インデックス218及び受信検索インデックス219を用いることによって、二次記憶装置203に格納される送信データブロック情報220及び受信データブロック情報221にアクセスすることなくデータブロックを保持しているか否かを判定できる。

30

【0230】

特に、dedup.hash2 514(変数)のサイズを大きくすることによって、ハッシュ値の衝突を回避することができるため、転送装置110は、送信検索インデックス218及び受信検索インデックス219のみを用いてデータブロックを検索することができる。

【0231】

また、送信検索インデックス218及び受信検索インデックス219は、送信データブロック情報220及び受信データブロック情報221よりサイズを小さくできるため、主記憶装置202に送信検索インデックス218及び受信検索インデックス219を格納することができる。

40

【0232】

したがって、転送装置110は、送信検索インデックス218及び受信検索インデックス219を用いて、データブロックの重複排除及び重複復元を高速に行うことができる。これによって、データ転送処理の高速化を実現することができる。

【実施例2】

【0233】

50

実施例 2 では、データブロックの分割方法が実施例 1 と異なる。以下、実施例 1 との差異を中心に実施例 2 について説明する。なお、実施例 1 と同一の構成、同一の処理には、同一の符号を付し、説明を省略する。

【 0 2 3 4 】

実施例 2 のネットワークシステムの構成は、実施例 1 のネットワークシステムの構成と同一であるため説明を省略する。

【 0 2 3 5 】

図 2 0 は、実施例 2 の転送装置 1 1 0 のハードウェア構成及びソフトウェア構成の一例を示す図である。図 2 1 は、実施例 2 の転送装置 1 1 0 のソフトウェア構成の論理的な接続関係を示す図である。

10

【 0 2 3 6 】

実施例 2 の転送装置 1 1 0 のハードウェア構成は、実施例 1 の転送装置 1 1 0 のハードウェア構成と同一であるため説明を省略する。

【 0 2 3 7 】

実施例 2 の転送装置 1 1 0 の主記憶装置 2 0 2 は、副検索インデックス 2 0 0 1 を格納する。副検索インデックス 2 0 0 1 は、送信データブロック情報 2 2 0 に格納されるデータブロックを検索する場合に使用されるインデックスである。副検索インデックス 2 0 0 1 の詳細は、図 2 4 を用いて説明する。

【 0 2 3 8 】

実施例 2 のフロー情報 2 1 7 のデータ構造は、実施例 1 のフロー情報 2 1 7 のデータ構造と異なる。実施例 2 のフロー情報 2 1 7 の詳細は、図 2 2 を用いて説明する。

20

【 0 2 3 9 】

実施例 2 のデータ分割部 2 1 5 は、実施例 1 とは異なる方法を用いてデータブロックを生成する。データ分割部 2 1 5 が生成するデータブロックについて、図 2 3 を用いて説明する。

【 0 2 4 0 】

実施例 2 の転送制御部 2 1 3 が実行する処理は、実施例 1 の転送制御部 2 1 3 が実行する処理と一部異なる。異なる処理については、図 2 5 を用いて説明する。実施例 2 の重複排除部 2 1 1 が実行する処理は、実施例 1 の重複排除部 2 1 1 が実行する処理と一部異なる。異なる処理については、図 2 6 及び図 2 7 を用いて説明する。また、実施例 2 の重複排除部 2 1 1 は、第 1 ハッシュ関数及び第 2 ハッシュ関数の他に、第 3 ハッシュ関数を保持する。

30

【 0 2 4 1 】

その他の情報及び機能部が実行する処理は、実施例 1 と同一である。

【 0 2 4 2 】

図 2 2 は、実施例 2 のフロー情報 2 1 7 の一例を示す図である。

【 0 2 4 3 】

フロー情報 2 1 7 に含まれるエントリは、head 2 2 0 1 を含む。head 2 2 0 1 は、バッファ 6 0 7 に格納されるデータの先頭がデータブロックの先頭であるか否かを示す情報であり、初期値は「NO」である。エントリに含まれるその他のカラムは実施例 1 と同一である。

40

【 0 2 4 4 】

図 2 3 は、実施例 2 のデータの分割方法の概念を示す図である。

【 0 2 4 5 】

実施例 2 では、LAN 1 4 0 を介して受信したデータは、固定長のデータブロックに分割される。図 2 3 では、i 番目のデータブロックの識別情報を M_i と記載している。

【 0 2 4 6 】

さらに、実施例 2 では、一つのデータブロックが、固定長の副データブロックに分割される。転送装置 1 1 0 は、副データブロックの場所を検索するためのインデックスとして、副検索インデックス 2 0 0 1 を保持する。

50

【0247】

図24は、実施例2の副検索インデックス2001の一例を示す図である。

【0248】

副検索インデックス2001は、2次元配列のデータ構造である。図24では、N行4列の2次元配列となっている。なお、行及び列の数は一例であって、任意の値に設定できる。

【0249】

2次元配列の一つの要素2401には、ハッシュ値XX及び添字Iが含まれる。ハッシュ値XXは、第3ハッシュ関数にデータブロックを入力することによって算出されるハッシュ値である。ハッシュ値XXは、データブロックが送信データブロック情報220に格納されているか否かを判定する場合に使用される。添字Iは、データブロックの先頭から何番目の副データブロックであることを示す値である。

10

【0250】

図25は、実施例2の転送制御部213がLAN140からデータを受信した場合に実行する処理を説明するフローチャートである。

【0251】

転送制御部213は、ステップS1002の処理が完了した後、特定されたエントリのバッファ607に格納されるデータの先頭がデータブロックの先頭であるか否かを判定する(ステップS2501)。

【0252】

具体的には、転送制御部213は、特定されたエントリのhead2201が「YES」であるか否かを判定する。head2201が「YES」である場合、転送制御部213は、特定されたエントリのバッファ607に格納されるデータの先頭がデータブロックの先頭であると判定する。

20

【0253】

特定されたエントリのバッファ607に格納されるデータの先頭がデータブロックの先頭であると判定された場合、転送制御部213は、データ分割部215を呼び出す(ステップS1003)。このとき、転送制御部213は、受け取ったデータをデータ分割部215に出力する。その後、転送制御部213は、処理を終了する。

【0254】

特定されたエントリのバッファ607に格納されるデータの先頭がデータブロックの先頭ではないと判定された場合、転送制御部213は、特定されたエントリのバッファ607に格納されるデータのサイズが比較ブロックサイズ以上であるか否かを判定する(ステップS2502)。なお、比較ブロックサイズは、予め設定されているものとする。

30

【0255】

特定されたエントリのバッファ607に格納されるデータのサイズが比較ブロックサイズより小さいと判定された場合、転送制御部213は、特定されたエントリのバッファ607にデータが残っているか否かを判定する(ステップS2503)。

【0256】

特定されたエントリのバッファ607にデータが残っていると判定された場合、転送制御部213は、当該エントリのタイマ608に任意の時間を設定する(ステップS2504)。その後、転送制御部213は、処理を終了する。

40

【0257】

特定されたエントリのバッファ607にデータが残っていないと判定された場合、転送制御部213は、当該エントリのタイマ608に「0」を設定する(ステップS2505)。その後、転送制御部213は、処理を終了する。

【0258】

ステップS2502において、特定されたエントリのバッファ607に格納されるデータのサイズが比較ブロックサイズ以上であると判定された場合、転送制御部213は、重複排除部211に副データブロックの検出指示を出力する(ステップS2506)。転送

50

制御部 213 は、重複排除部 211 から応答を受け取るまで待ち状態に移行する。なお、副データブロックの検索処理は、図 26 を用いて説明する。

【0259】

転送制御部 213 は、重複排除部 211 から受け取った応答に基づいて、特定されたエントリのバッファ 607 に格納されるデータに副データブロックが存在するか否かを判定する（ステップ S2507）。

【0260】

具体的には、転送制御部 213 は、重複排除部 211 からの応答が「該当データ無し」であるか否かを判定する。重複排除部 211 からの応答が「該当データ無し」である場合、転送制御部 213 は、特定されたエントリのバッファ 607 に格納されるデータに副データブロックが存在すると判定する。

10

【0261】

特定されたエントリのバッファ 607 に格納されるデータに副データブロックが存在すると判定された場合、転送制御部 213 は、副データブロックを含むデータブロックを特定し、特定されたエントリのバッファ 607 に格納されるデータの先頭から当該データブロックの手前までのデータを削除し、符号化されたデータを出力するとともに、特定されたエントリの head 2201 を「YES」に更新する（ステップ S2508）。その後、転送制御部 213 は、ステップ S1003 に進む。具体的には、以下のような処理が実行される。

【0262】

20

転送制御部 213 は、重複排除部 211 を介して、検索された副データブロックの識別情報（添字 I）に基づいて、特定されたエントリのバッファ 607 に格納されるデータにおける、データブロックの先頭を特定する。

【0263】

転送制御部 213 は、重複排除部 211 にデータブロックの符号化を指示する。転送制御部 213 は、重複排除部 211 から符号化されたデータを取得するまで待ち状態に移行する。当該指示には、特定されたエントリのバッファ 607 に格納されるデータが含まれる。以下では、バッファ 607 に格納されるデータの先頭から当該データブロックの手前までのデータを対象データと記載する。

【0264】

30

重複排除部 211 は、当該指示を受け付けた場合、dedup.type 511 に「3」を設定し、dedup.data_len 512 に対象データの長さを設定し、dedup.hash1 513 及び dedup.hash2 514 にダミーの値として「0」を設定することによって Dedup ヘッダ 510 を生成する。また、重複排除部 211 は、Dedup ペイロード 520 に対象データを設定する。以上がステップ S2508 の処理の説明である。

【0265】

特定されたエントリのバッファ 607 に格納されるデータに副データブロックが存在しないと判定された場合、転送制御部 213 は、特定されたエントリのバッファ 607 に格納されるデータから比較ブロックサイズ分のデータを削除し、符号化されたデータを出力する（ステップ S2509）。

40

【0266】

ステップ S2509 の処理は、ステップ S2508 の処理と同様の処理である。ただし、Dedup ペイロード 520 にはバッファ 607 から削除されたデータが含まれ、また、特定されたエントリの head 2201 は更新されない。

【0267】

比較ブロックサイズは、副データブロックのサイズより大きい値が設定される。比較ブロックサイズが大きい場合、処理の負荷が軽くなるため、データの分割を高速に行うことができる。一方、比較ブロックサイズが大きい場合、重複排除されるデータの割合が小さくなる。

50

【0268】

図26は、実施例2の重複排除部211が転送制御部213から副データブロックの検出指示を受け付けた場合に実行する処理を説明するフローチャートである。

【0269】

重複排除部211は、転送制御部213から副データブロックの検出指示を受け付けた場合、特定されたエントリのバッファ607に格納されるデータを取得する(ステップS2601)。

【0270】

重複排除部211は、変数Sに「1」を設定し(ステップS2602)、また、データの先頭に所定のサイズのウィンドウを配置する(ステップS2603)。データ分割部215は、ウィンドウに含まれるデータのハッシュ値を算出する(ステップS2604)。

10

【0271】

具体的には、データ分割部215は、変数と同様の手順に基づいてハッシュ値SAを算出し、第3ハッシュ関数にデータを入力することによってハッシュ値SCを算出する。なお、ウィンドウのサイズは、副データブロックのサイズと同一のものとする。

【0272】

次に、重複排除部211は、副検索インデックス2001を参照し(ステップS2605)、ハッシュ値XXがハッシュ値SCに一致する要素2401が存在するか否かを判定する(ステップS2606)。

【0273】

具体的には、重複排除部211は、SA行目の行に含まれる各要素2401に格納されるハッシュ値XXと、ハッシュ値SCとを比較する。

20

【0274】

ハッシュ値XXがハッシュ値SCに一致する要素2401が存在すると判定された場合、重複排除部211は、転送制御部213に、当該要素2401に含まれる添字Iを応答として出力する(ステップS2607)。その後、重複排除部211は処理を終了する。

【0275】

ハッシュ値XXがハッシュ値SCに一致する要素2401が存在しないと判定された場合、重複排除部211は、変数Sが閾値以上であるか否かを判定する(ステップS2608)。閾値は任意に設定できる。

30

【0276】

変数Sが閾値以上であると判定された場合、重複排除部211は、転送制御部213に、「該当データ無し」を応答として出力する(ステップS2609)。

【0277】

変数Sが閾値より小さいと判定された場合、重複排除部211は、変数Sに「1」を加算し(ステップS2610)、また、ウィンドウをデータの後方にスライドさせる(ステップS2611)。その後、重複排除部211は、ステップS2604に戻る。例えば、重複排除部211は、ウィンドウを1バイト分だけ後方にスライドさせる。

【0278】

図27は、実施例2の重複排除部211が転送制御部213によって呼び出された場合に実行する処理を説明するフローチャートである。

40

【0279】

実施例2では、重複排除部211は、ステップS1508の処理が完了した後、副検索インデックス2001を更新する(ステップS2701)。その後、重複排除部211は、ステップS1509に進む。具体的には、以下のような処理が実行される。

【0280】

重複排除部211は、データブロックに含まれる副データブロックの中から、対象の副データブロックを選択し、当該副データブロックを第1ハッシュ関数及び第2ハッシュ関数に入力し、ハッシュ値Sa及びハッシュ値Scを算出する。

【0281】

50

重複排除部 211 は、副検索インデックス 2001 を参照し、S a 行目に含まれる要素 2401 に空の要素 2401 が存在するか否かを判定する。

【0282】

S a 行目に含まれる要素 2401 に空の要素 2401 が存在する場合、重複排除部 211 は、空きの要素 2401 のハッシュ値 X X にハッシュ値 S c を格納し、添字 I にデータブロックにおける対象の副データブロックの位置を示す値を格納する。

【0283】

S a 行目に含まれる要素 2401 に空の要素 2401 が存在しない場合、重複排除部 211 は、所定のアルゴリズムに基づいて、S a 行目に含まれる要素 2401 の中から値を設定する要素 2401 を一つ選択する。重複排除部 211 は、選択された要素 2401 のハッシュ値 X X にハッシュ値 S c を格納し、添字 I にデータブロックにおける対象の副データブロックの位置を示す値を格納する。以上がステップ S 2701 の処理の説明である。

【0284】

実施例 2 によれば、実施例 1 と同様にデータ転送処理の高速化を実現することができる。また、実施例 2 によれば、転送装置 110 は、比較ブロックサイズの間隔をあげながら副データブロックを用いて、データブロックの境界をバイト単位で調べることができる。データ全体に対してデータブロックの境界を特定する処理が転送処理の速度を律速する場合、比較ブロックサイズのサイズを大きくすることによって、データブロックの境界を特定する処理を間引くことができる。

【0285】

また、データブロックサイズと独立したパラメータである比較ブロックサイズを用いることによって、バイト単位のデータブロックの境界の特定処理の間引き方を調整できる。データブロックサイズを小さくした場合、データブロックが一致する可能性が高まるため重複排除の効果を高めることができ、比較ブロックサイズを大きくした場合、バイト単位のデータブロックの境界の特定処理が間引かれるため処理を高速化できる。

【0286】

請求の範囲に記載した以外の発明の観点の代表的なものとして、次のものがあげられる。

(1) 端末間で送受信されるデータを転送する複数の転送装置を備えるネットワークシステムであって、

前記複数の転送装置の各々は、

前記端末間のデータの転送処理を制御する転送制御部を有し、

データを分割することによって生成されるデータブロックであって、前記端末間で過去に送信されたデータブロックを管理するデータブロック情報と、前記データブロック情報に格納される前記データブロックを検索するための識別情報を管理する検索インデックスと、を保持し、

前記複数の転送装置は、第 1 検索インデックス及び第 1 データブロック情報を保持する第 1 転送装置、並びに、第 2 検索インデックス及び第 2 データブロック情報を保持する第 2 転送装置を含み、

前記第 1 転送装置は、

前記第 1 転送装置に接続する端末から送信データを受信した場合、前記送信データから前記データブロックを生成するための分割処理を実行し、

前記分割処理によって生成された第 1 データブロックから、前記検索インデックスを参照するための第 1 識別情報を算出し、

前記第 1 識別情報に基づいて前記第 1 検索インデックスを参照して、前記第 1 データブロックが前記第 1 データブロック情報に格納されているか否かを判定し、

前記第 1 データブロックが前記第 1 データブロック情報に格納されていないと判定された場合、前記第 1 識別情報及び前記第 1 データブロックを前記第 2 転送装置に送信し、

前記第 2 転送装置は、

10

20

30

40

50

前記第1識別情報及び前記第1データブロックを受信した場合、前記第2検索インデックスに前記第1識別情報を格納し、前記第2データブロック情報に前記第1データブロックを格納し、

前記第1転送装置に応答を送信し、

前記第1転送装置は、前記第2転送装置からの応答を受信した場合、前記第1検索インデックスに前記第1識別情報を格納し、前記第1データブロック情報に前記第1データブロックを格納することを特徴とするネットワークシステム。

【0287】

なお、本発明は上記した実施例に限定されるものではなく、様々な変形例が含まれる。また、例えば、上記した実施例は本発明を分かりやすく説明するために構成を詳細に説明したものであり、必ずしも説明した全ての構成を備えるものに限定されるものではない。また、実施例の構成の一部について、他の構成に追加、削除、置換することが可能である。

10

【0288】

また、上記の各構成、機能、処理部、処理手段等は、それらの一部又は全部を、例えば集積回路で設計する等によりハードウェアで実現してもよい。また、本発明は、実施例の機能を実現するソフトウェアのプログラムコードによっても実現できる。この場合、プログラムコードを記録した記憶媒体をコンピュータに提供し、そのコンピュータが備えるCPUが記憶媒体に格納されたプログラムコードを読み出す。この場合、記憶媒体から読み出されたプログラムコード自体が前述した実施例の機能を実現することになり、そのプログラムコード自体、及びそれを記憶した記憶媒体は本発明を構成することになる。このようなプログラムコードを供給するための記憶媒体としては、例えば、フレキシブルディスク、CD-ROM、DVD-ROM、ハードディスク、SSD(Solid State Drive)、光ディスク、光磁気ディスク、CD-R、磁気テープ、不揮発性のメモリカード、ROMなどが用いられる。

20

【0289】

また、本実施例に記載の機能を実現するプログラムコードは、例えば、アセンブラ、C/C++、perl、Shell、PHP、Java等の広範囲のプログラム又はスクリプト言語で実装できる。

【0290】

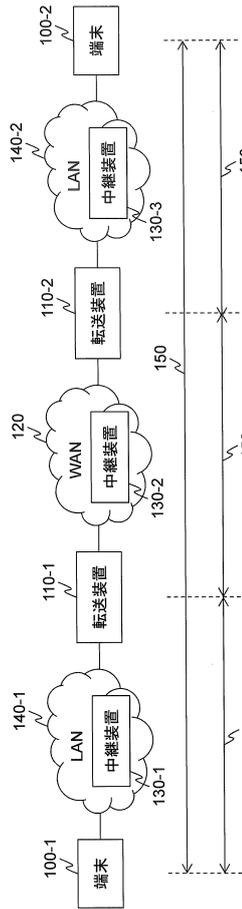
さらに、実施例の機能を実現するソフトウェアのプログラムコードを、ネットワークを介して配信することによって、それをコンピュータのハードディスクやメモリ等の記憶手段又はCD-RW、CD-R等の記憶媒体に格納し、コンピュータが備えるCPUが当該記憶手段や当該記憶媒体に格納されたプログラムコードを読み出して実行するようにしてもよい。

30

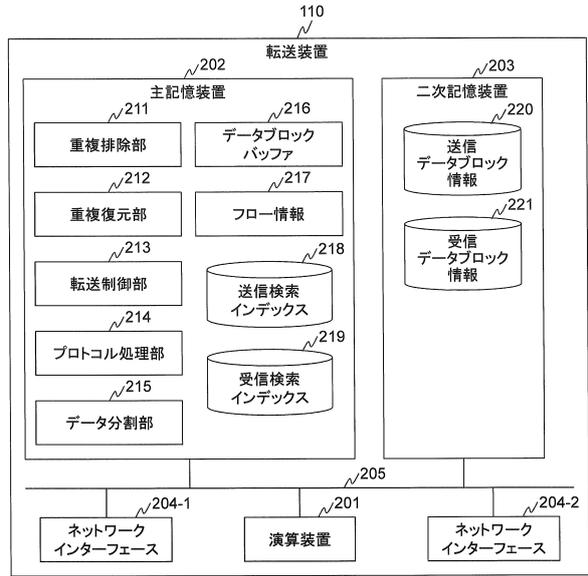
【0291】

上述の実施例において、制御線や情報線は、説明上必要と考えられるものを示しており、製品上必ずしも全ての制御線や情報線を示しているとは限らない。全ての構成が相互に接続されていてもよい。

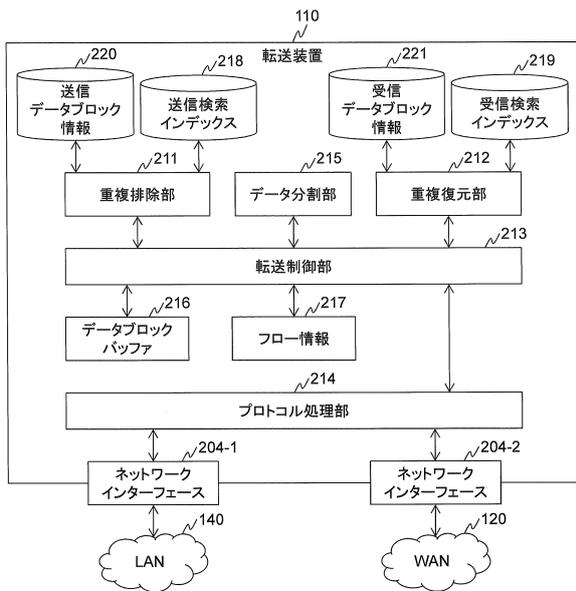
【図1】



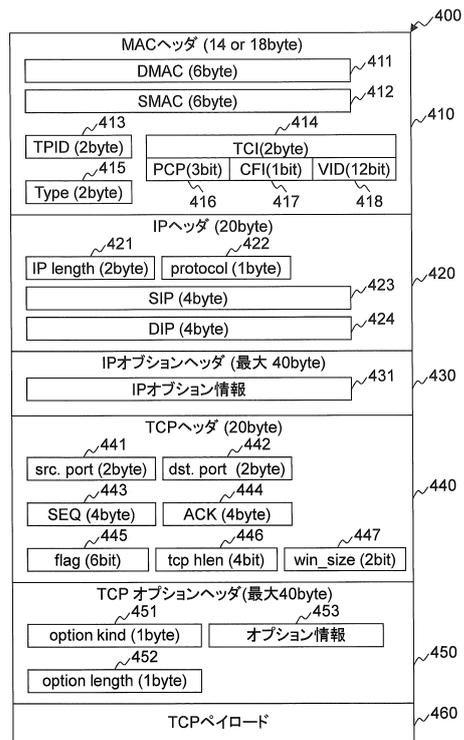
【図2】



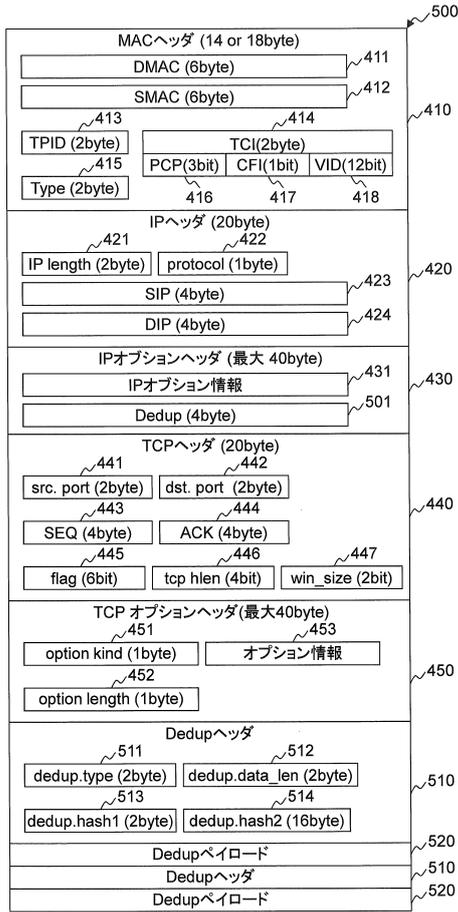
【図3】



【図4】



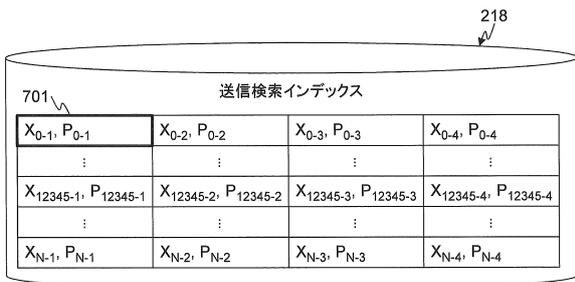
【図5】



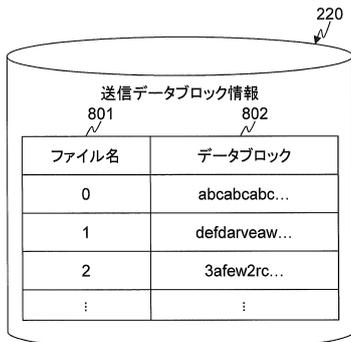
【図6】

ポートID	VLAN	LAN-IP	WAN-IP	LAN-PORT	WAN-PORT	バッファ	タイム
1	1	192.168.0.1	10.0.0.1	80	12345	abcabca...	150
2	10	192.169.0.2	123.123.123.123	8080	20842	xyzxyz...	100
3	10	192.170.0.3	111.111.111.111	80	45728	-	0

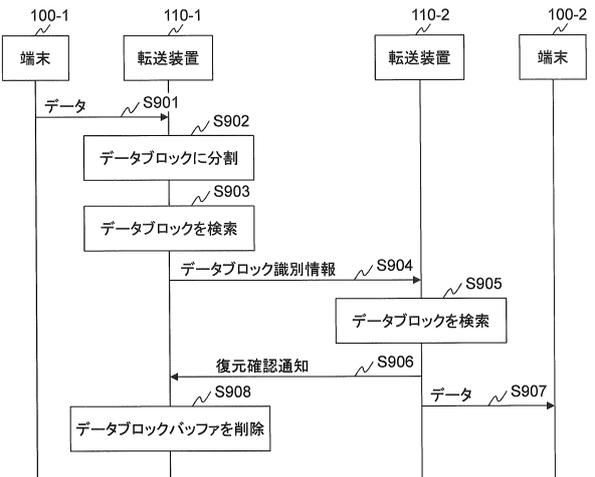
【図7】



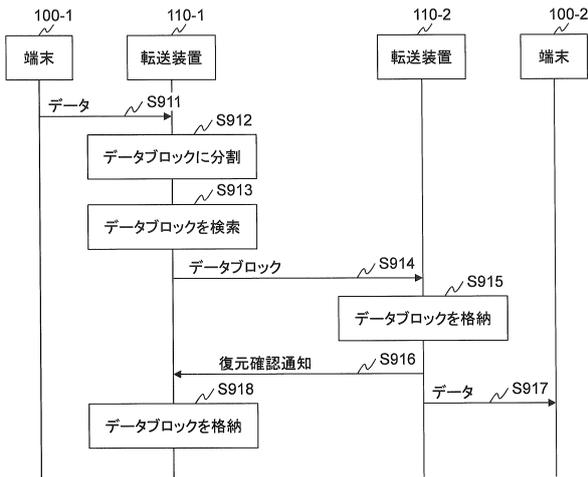
【図8】



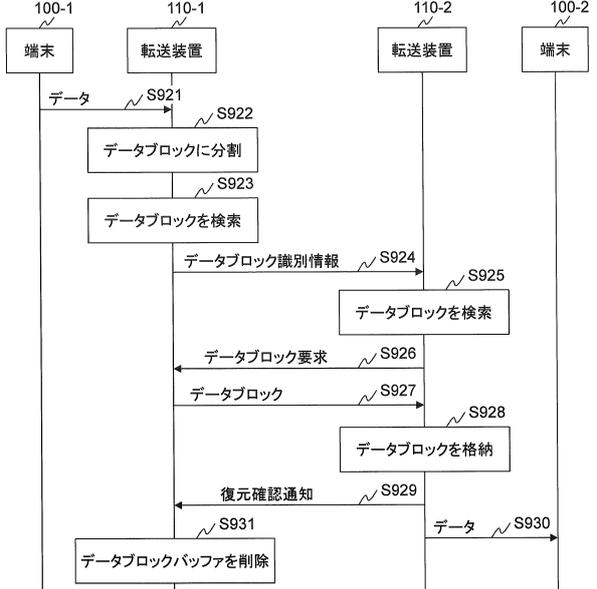
【図9A】



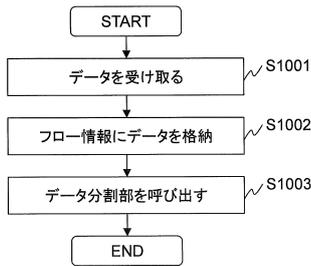
【図9B】



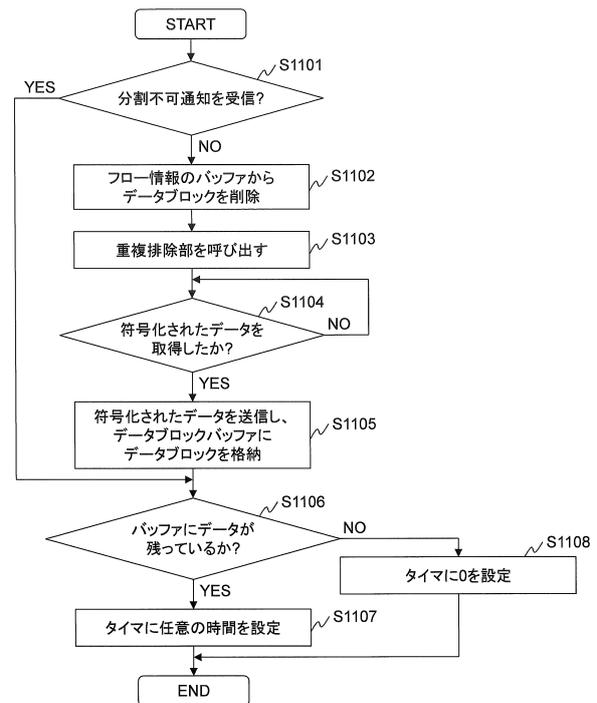
【図9C】



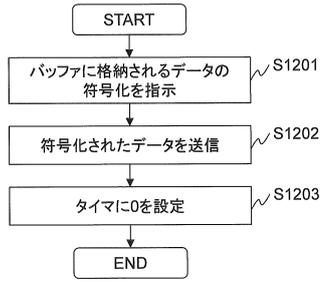
【図10】



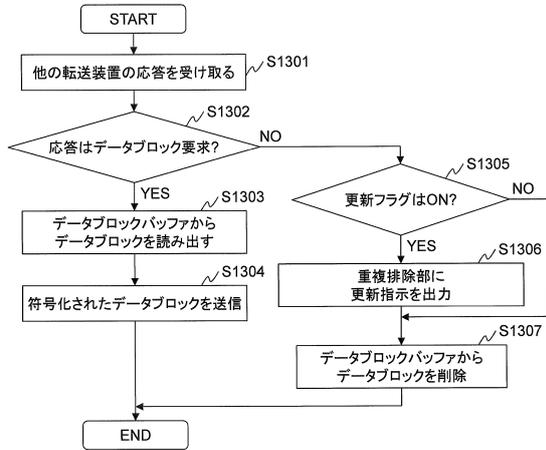
【図11】



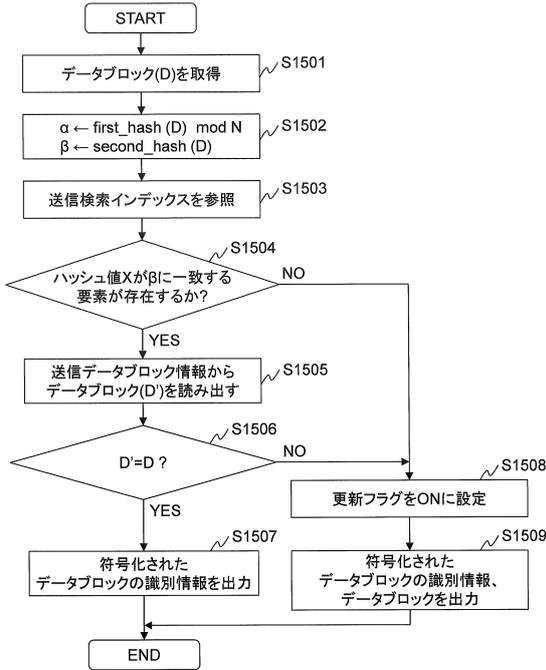
【図12】



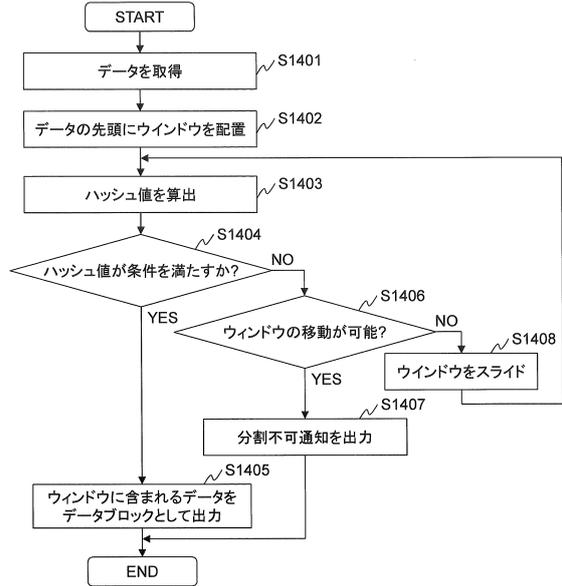
【図13】



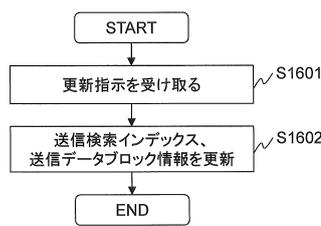
【図15】



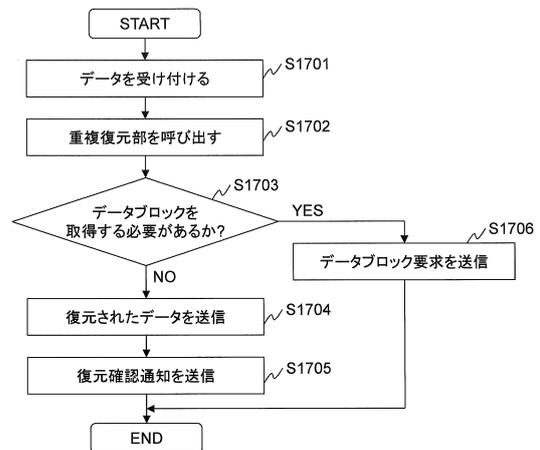
【図14】



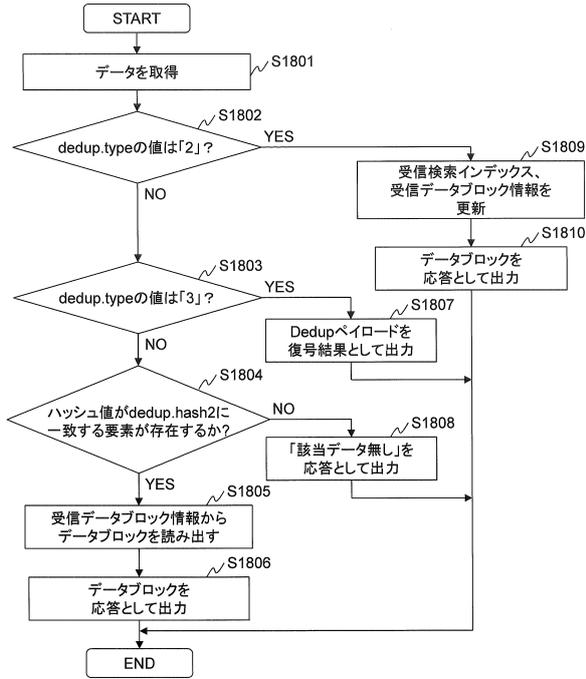
【図16】



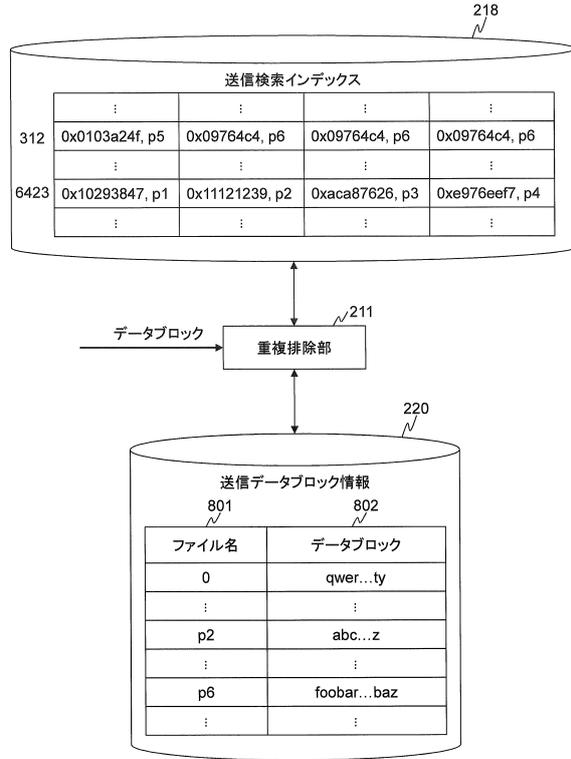
【図17】



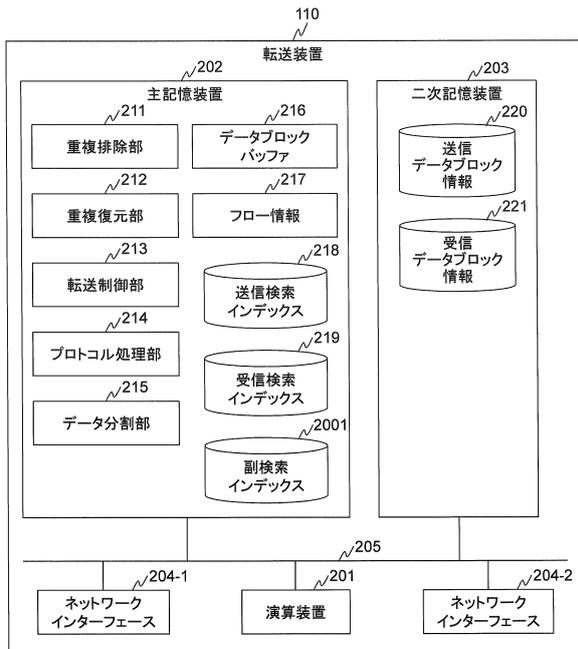
【図18】



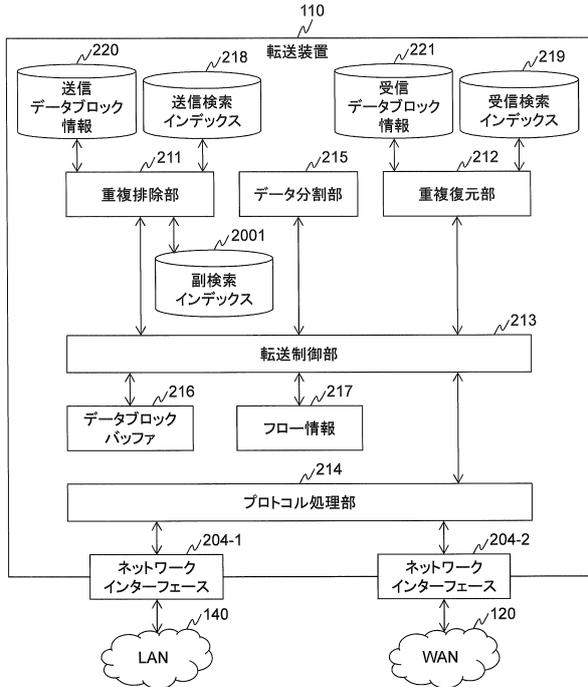
【図19】



【図20】



【図21】



【図22】

217	head	2201
608	タイム	150
607	パッパ	abcabca...
606	WAN-PORT	12345
605	LAN-PORT	80
604	WAN-IP	10.0.0.1
603	LAN-IP	192.168.0.1
602	VLAN	1
601	プロID	1
		2
		3
		100
		20842
		45728
		0
		NO
		YES
		NO

【図23】

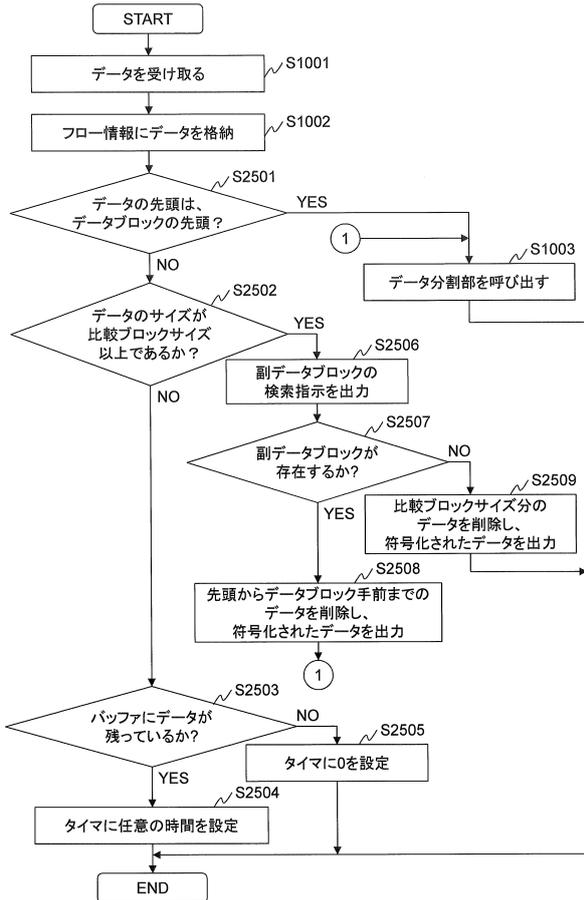
データ											
データブロック M1						データブロック M2					
S1	S2	S3	...	S16	S17	S18	S19	...	S32

【図24】

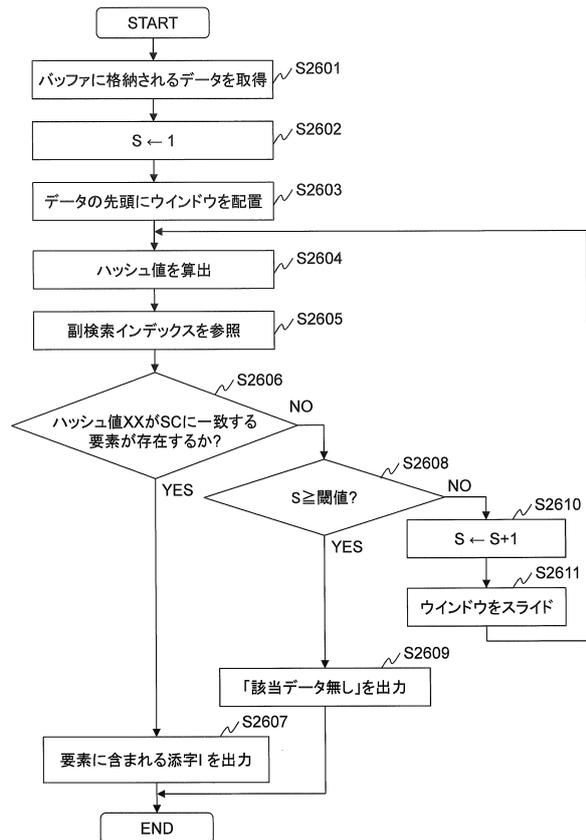
2001

副検索インデックス			
XX ₀₋₁ , I ₀₋₁	XX ₀₋₂ , I ₀₋₂	XX ₀₋₃ , I ₀₋₃	XX ₀₋₄ , I ₀₋₄
:	:	:	:
XX ₄₅₆₋₁ , I ₄₅₆₋₁	XX ₄₅₆₋₂ , I ₄₅₆₋₂	XX ₄₅₆₋₃ , I ₄₅₆₋₃	XX ₄₅₆₋₄ , I ₄₅₆₋₄
:	:	:	:
XX _{N-1} , I _{N-1}	XX _{N-2} , I _{N-2}	XX _{N-3} , I _{N-3}	XX _{N-4} , I _{N-4}

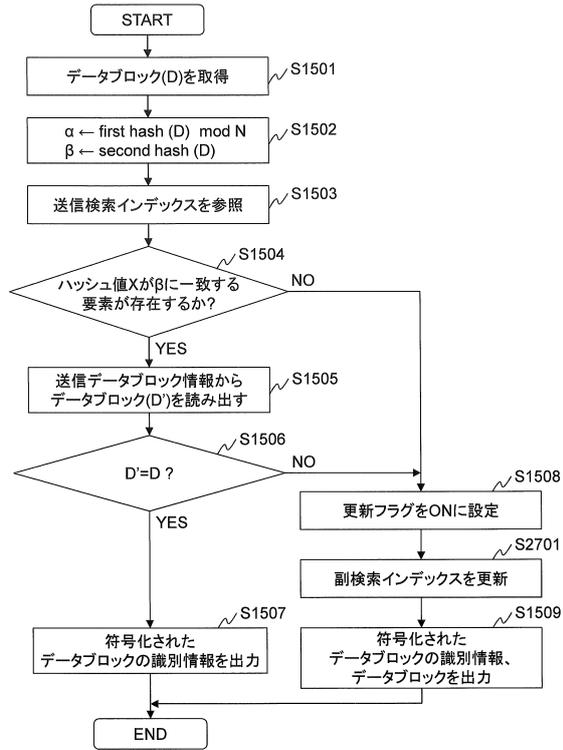
【図25】



【図26】



【図27】



フロントページの続き

- (72)発明者 對馬 雄次
東京都千代田区丸の内一丁目6番6号 株式会社日立製作所内
- (72)発明者 宮崎 悠樹
東京都千代田区丸の内一丁目6番6号 株式会社日立製作所内

審査官 小林 義晴

- (56)参考文献 国際公開第2016/121024(WO, A1)
特開2004-096692(JP, A)
特開平05-151047(JP, A)
特開2012-113350(JP, A)
特表2006-505215(JP, A)

- (58)調査した分野(Int.Cl., DB名)
G06F 13/00