

(19) 日本国特許庁 (JP)

(12) 特 許 公 報 (B2)

(11) 特許番号

特許第4829896号
(P4829896)

(45) 発行日 平成23年12月7日 (2011. 12. 7)

(24) 登録日 平成23年9月22日 (2011. 9. 22)

(51) Int. Cl.

F I

H O 4 L 29/08 (2006. 01)

H O 4 L 13/00 3 O 7 Z

H O 4 L 12/56 (2006. 01)

H O 4 L 12/56 3 O O D

請求項の数 16 (全 17 頁)

(21) 出願番号 特願2007-546070 (P2007-546070)
 (86) (22) 出願日 平成17年12月15日 (2005. 12. 15)
 (65) 公表番号 特表2008-524891 (P2008-524891A)
 (43) 公表日 平成20年7月10日 (2008. 7. 10)
 (86) 国際出願番号 PCT/EP2005/056832
 (87) 国際公開番号 W02006/064047
 (87) 国際公開日 平成18年6月22日 (2006. 6. 22)
 審査請求日 平成20年9月10日 (2008. 9. 10)
 (31) 優先権主張番号 11/014, 552
 (32) 優先日 平成16年12月16日 (2004. 12. 16)
 (33) 優先権主張国 米国 (US)

(73) 特許権者 390009531
 インターナショナル・ビジネス・マシーンズ・コーポレーション
 INTERNATIONAL BUSIN
 ESS MACHINES CORPO
 RATION
 アメリカ合衆国10504 ニューヨーク
 州 アーモンク ニュー オーチャード
 ロード
 (74) 代理人 100108501
 弁理士 上野 剛史
 (74) 代理人 100112690
 弁理士 太佐 種一
 (74) 代理人 100091568
 弁理士 市位 嘉宏

最終頁に続く

(54) 【発明の名称】 データ破壊を避けることによる改善されたネットワーク性能のための方法、システム及び物品

(57) 【特許請求の範囲】

【請求項 1】

データ処理システム・ネットワークにおいて通信を処理するためのインターネット・プロトコル (IP) パケットを生成するための方法であって、

送信されるべき IP パケットの IP 宛先が当該 IP パケットの IP ソース・アドレスと同じサブネット上にあるかないかを、IP パケットのセンダによって決定するステップと、

IP パケットの IP 宛先は当該 IP パケットの IP ソース・アドレスと同じサブネット上にないときには、前記パケットの転送のラウンド・トリップ時間を計算したかどうかを、前記センダによって決定するステップと、

前記パケットの転送のラウンド・トリップ時間が計算されるときには、前記ラウンド・トリップ時間にセットされたリアセンブリ・タイム・オプションズ・フィールドを備える前記 IP パケットを生成するステップと、

前記パケットの転送のラウンド・トリップ時間が計算されないときには、IP パケットが生成された時間がセットされるタイム・スタンプ・リアセンブリ・タイム・オプションズ・フィールドを備える前記 IP パケットを生成するステップと

を含む、前記方法。

【請求項 2】

IP パケットの IP 宛先は当該 IP パケットの IP ソース・アドレスと同じサブネット上にあるときには、前記サブネットの送信速度に対応する時間にセットされたリアセンブ

リ・タイマ・オプションズ・フィールドを備える I P データ・パケットを生成するステップをさらに含む、請求項 1 に記載の方法。

【請求項 3】

前記 I P パケットの前記センダが前記 I P パケットのレシーバと同じサブネット上にあるときには、当該センダが前記ネットワークの既知高速リンクを介して前記レシーバに接続されている、請求項 1 に記載の方法。

【請求項 4】

前記ラウンド・トリップ時間が前記センダから前記 I P パケットのレシーバへ送信される前記 I P パケットについてのラウンド・トリップ時間である、請求項 1 に記載の方法。

【請求項 5】

前記ラウンド・トリップ時間が T C P 層により計算されたラウンド・トリップ時間である、請求項 1 に記載の方法。

【請求項 6】

前記タイム・スタンプ・リアセンブリ・タイマ・オプションズ・フィールドを備える前記 I P パケットが生成される場合に、前記タイム・スタンプ・リアセンブリ・タイマ・オプションズ・フィールドが I P パケットが生成された時間を明示するタイム・スタンプを含む、請求項 1 に記載の方法。

【請求項 7】

リアセンブリ・タイマ・オプションズ・フィールドを備える前記 I P パケットを生成するステップは、前記 I P パケットのヘッダにおいてリアセンブリ時間を明示するオプションズ・フィールドを設定するステップを含む、請求項 1 に記載の方法。

【請求項 8】

タイム・スタンプ・リアセンブリ・タイマ・オプションズ・フィールドを備える前記 I P パケットを生成するステップは、前記 I P パケットのヘッダにおいて前記 I P パケットが生成された時間を明示するタイム・スタンプを備えるオプションズ・フィールドを設定するステップを含む、請求項 1 に記載の方法。

【請求項 9】

データ処理システムであって、

送信されるべき I P パケットの I P 宛先が当該 I P パケットの I P ソース・アドレスと同じサブネット上にあるかないかを、I P パケットのセンダによって決定するための手段と、

I P パケットの I P 宛先は当該 I P パケットの I P ソース・アドレスと同じサブネット上にないときには、前記パケットの転送のラウンド・トリップ時間を計算したかどうかを、前記センダによって決定するための手段と、

前記パケットの転送のラウンド・トリップ時間が計算されるときには、前記ラウンド・トリップ時間にセットされたリアセンブリ・タイマ・オプションズ・フィールドを備える前記 I P パケットを生成するための手段と、

前記パケットの転送のラウンド・トリップ時間が計算されないときには、I P パケットが生成された時間がセットされるタイム・スタンプ・リアセンブリ・タイマ・オプションズ・フィールドを備える前記 I P パケットを生成するための手段と

を含む、前記データ処理システム。

【請求項 10】

I P パケットの I P 宛先が当該 I P パケットの I P ソース・アドレスと同じサブネット上にあるときには、前記サブネットの送信速度に対応する時間にセットされたリアセンブリ・タイマ・オプションズ・フィールドを備える I P データ・パケットを生成するための手段をさらに含む、請求項 9 に記載のデータ処理システム。

【請求項 11】

データ処理システム・ネットワークを介する通信のための方法であって、

送信されるべき特定のデータ・パケットについてカスタマイズされたリアセンブリ時間を決定するステップであって、前記決定するステップは、前記カスタマイズされたリアセ

10

20

30

40

50

ンブリ時間を、もし前記データ・パケットのセンダが前記ネットワークの既知高速リンクを介して前記データ・パケットのレシーバに接続されているならば第１時間に、もし前記データ・パケットのセンダが前記ネットワークの既知高速リンクを介して前記レシーバに接続されていなければ第２時間に、セットすることを含み、前記センダが前記ネットワークの前記データ・パケットの前記レシーバと同じサブネットの中にあるときには前記データ・パケットの前記センダは前記ネットワークの既知高速リンクを介して前記レシーバに接続されている、前記決定するステップと、

前記カスタマイズされたリアセンブリ時間を前記データ・パケットのフィールドに挿入することを含む、前記データ・パケットを生成するステップと、

前記データ・パケットを送信するステップと

10

を含み、

前記データ・パケットのレシーバは、前記カスタマイズされたリアセンブリ時間を利用して、フラグメントの受信後に前記カスタマイズされたリアセンブリ時間により示されるより長く分割されたままになっている前記パケットの前記フラグメントを捨てる、前記方法。

【請求項 1 2】

カスタマイズされたリアセンブリ時間を決定する前記ステップは、カスタマイズされたリアセンブリ時間をセンダから前記レシーバに送信されたデータ・パケットについて計算されたラウンド・トリップ時間にセットすることを含む、請求項 1 1 に記載の方法。

【請求項 1 3】

20

前記カスタマイズされたリアセンブリ時間は前記データ・パケットが生成された時間を明示するタイム・スタンプであり、前記データ・パケットの前記レシーバは、前記カスタマイズされたリアセンブリ時間を利用して前記データ・フラグメントの到着時間に基いて前記データ・フラグメントの横断時間を計算し、前記レシーバは前記データ・フラグメントについて前記横断時間に基いてリアセンブリ時間をセットする、請求項 1 1 又は 1 2 に記載の方法。

【請求項 1 4】

データ処理システムであって、

送信されるべき特定のデータ・パケットについてカスタマイズされたリアセンブリ時間を決定するための手段であって、前記決定は、前記カスタマイズされたリアセンブリ時間を、もし前記データ・パケットのセンダがネットワークの既知高速リンクを介して前記データ・パケットのレシーバに接続されているならば第１時間に、もし前記データ・パケットのセンダが前記ネットワークの既知高速リンクを介して前記レシーバに接続されていなければ第２時間に、セットすることを含み、前記センダが前記ネットワークの前記データ・パケットの前記レシーバと同じサブネットの中にあるときには前記データ・パケットの前記センダは前記ネットワークの既知高速リンクを介して前記レシーバに接続されている、前記決定するための手段と、

30

前記カスタマイズされたリアセンブリ時間を前記データ・パケットのフィールドに挿入することを含む、前記データ・パケットを生成するための手段と、

前記データ・パケットを送信するための手段と

40

を含み、

前記データ・パケットのレシーバは、前記カスタマイズされたリアセンブリ時間を利用して、フラグメントの受信後に前記カスタマイズされたリアセンブリ時間により示されるより長く分割されたままになっている前記パケットの前記フラグメントを捨てる、前記データ処理システム。

【請求項 1 5】

前記決定するための手段は、カスタマイズされたリアセンブリ時間をセンダから前記レシーバに送信されたデータ・パケットについて計算されたラウンド・トリップ時間にセットすることを含む、請求項 1 4 に記載のデータ処理システム。

【請求項 1 6】

50

前記カスタマイズされたリアセンブリ時間は前記データ・パケットが生成された時間を明示するタイム・スタンプであり、前記データ・パケットの前記レシーバは、前記カスタマイズされたリアセンブリ時間を利用して前記データ・フラグメントの到着時間に基いて前記データ・フラグメントの横断時間を計算し、前記レシーバは前記データ・フラグメントについて前記横断時間に基いてリアセンブリ時間をセットする、請求項14又は15に記載のデータ処理システム。

【発明の詳細な説明】

【技術分野】

【0001】

本発明は、一般的にはデータ処理ネットワーク・システムにおけるデータ転送に関し、特にインターネット及び類似ネットワークを介するデータ・ブロックの転送に関する。更に特に、本発明は改善されたインターネット・プロトコル（IP）ネットワーク通信に関する。

【背景技術】

【0002】

コンピュータ・ネットワークはエンティティー間でデータを移送するための相互接続された通信媒体の地理的に分散した集合体である。エンティティーは、通信媒体を介してネットワーク・メッセージのソーシング（すなわち、送信）及び受信の少なくとも一方を行うホスト又はエンド・ステーションのような任意の装置から成り得る。多くのタイプのコンピュータ・ネットワークが利用可能であり、そのタイプはローカル・エリア・ネットワーク（LAN）からワイド・エリア・ネットワーク（WAN）に及ぶ。エンド・ステーションは、パーソナル・コンピュータ又はワークステーションを含み得るものであって、通例、所定のプロトコルに従ってデータの、フレーム又はパケットのような、離散的メッセージを交換することによって通信する。この文脈において、通信プロトコル・スタックはステーション同士がどの様に相互作用するかを定義する規則の集合から成る。

【0003】

インターネットは、データ（テキスト、コード、画像、ビデオ、オーディオ、又は混合）及びソフトウェアの送信及び配布のための重要なコンピュータ・ネットワークになっている。インターネット通信アーキテクチャ・プロトコル・スタックの主要なプロトコルは、ネットワーク層（第3層）におけるインターネット・プロトコル（IP）とトランスポート層（第4層）における伝送制御プロトコル（TCP）とである。TCP/IPという用語はインターネット・アーキテクチャを指すために一般的に使われており、これはインターネット及びイントラネット技術において広く実行される標準的通信プロトコルになっており、クライアント、サーバ、及びそれらを結合させる通信システムの間での広範な異質性を可能にしている。IPは、ネットワーク・レベルにおいて“データグラム”配信サービスを提供する。TCPは、2つのIPホスト間でデータ・ストリームの信頼できる順次配信を提供する接続指向のトランスポート・レベル・サービスを築く。TCP/IP伝送における信頼性は一般に3つのイベント、すなわちデータ・ロス、データ破壊、及びデータの再配列（reordering）、により損われる。

【0004】

データ・ロスはTCP/IPではタイム・アウト・メカニズムによって管理される。TCPは、レシーバからの送信されたセグメントについての肯定応答（ACK）の受信の遅れを測定するタイマ（再送タイマ）を維持する。推定された時間間隔（再送タイム・アウト（retransmission time-out（RTO））内にACKが到着しないときには、対応するセグメントは失われたと見なされ、再送される。更に、TCPは、伝統的に、パケット・ロスはネットワーク輻輳の兆候であるという前提に基いているので、TCPは、“スロー・スタート”を入力することによってその伝送速度を後退させ、これによりその輻輳ウィンドウを1セグメントまで徹底的に減少させる。

【0005】

TCPは、セグメントがレシーバに到着するに連れてセグメントに関してチェックサム

10

20

30

40

50

を実行することによってデータ破壊を管理する。チェックサム・アルゴリズムはTCPヘッダ及びデータの全ての16ビット・ワードの1の補数の和の16ビットの1の補数である。TCPセグメントは、パケット・データに関してチェックサムを計算し、この2バイト値をTCPヘッダにロードする。TCPヘッダのチェックサム・フィールドは、IPヘッダからの情報を含む1つの12バイト擬似ヘッダも含む。レシーバは、受け取られたデータに関するチェックサムを計算し(TCPヘッダ内の2バイト・チェックサム・フィールドを除く)、それがヘッダ内のチェックサム値と一致することを確認する。

【0006】

TCPは、入ってくるパケットをそれらが順に再配列されるまで待ち行列に入れるリアセンブリ・キューを維持することによってデータの再配列又はセグメントの順番が狂った到着を管理する。この待ち行列内のデータが順番に到着したときに限って、それはユーザの受信バッファに移され、ここでそれはユーザから見える。レシーバが、受け取られたパケットのシーケンス番号の中に“穴”を見つけたとき、それは、自分が受け取るその後の全ての“順番が狂った”パケットのために重複肯定応答(duplicate acknowledgment(DACK))を生成する。抜けているパケットが受け取られるまで、より高いシーケンス番号を有する各々の受け取られたデータ・パケットは“順番が狂っている”と見なされ、DACKを生成させる。

【0007】

図10は、IPヘッダ部分110とペイロード/データ部分150とを含むIPパケット100の略ブロック図である。IPヘッダ110は、IPヘッダのフォーマットを示すバージョン・フィールド102と、インターネット・ヘッダの長さを示すインターネット・ヘッダ長さ(Internet header length(IHL))フィールド104と、所望のサービス品質のパラメータの表示を提供するサービスのタイプ(type of service(TOS))フィールド106とを含む。IP全長フィールド108はIPヘッダ及びペイロード/データを含むIPパケットの長さを明示し、IP識別フィールド110はパケットのフラグメントを組み立てるのに役立つように送信エンティティにより割り当てられた識別値を明示する。

【0008】

IPヘッダは更に、モア・フラグメント(more fragment(MF))フラグ112と、IPパケット内でのそのフラグメントの配置を明示するIPフラグメント・オフセット・フィールド114と、そのパケットがネットワーク内で存続することを許される最大時間を示すタイム・ツー・ライブ(time to live(TTL))フィールド116とを含む。プロトコル・フィールド118はパケットのペイロード/データ部分150で使われる次レベル・プロトコルを示し、ヘッダ・チェックサム・フィールド120はIPヘッダだけについてのチェックサムを提供する。IPヘッダは更に、送信エンティティのIPソース・アドレスを含むソース・アドレス・フィールド122と、受信エンティティのIP宛先アドレスを含む宛先アドレス・フィールド124と、オプションズ・フィールド126(以下、オプションともいう)とパディング・フィールド128とを含む。

【0009】

ソース及び宛先エンティティに関連付けられたLAN標準規格が異種のものであるならば(例えば、イーサネット及びトークン・リング)、IPデータグラム(以降、パケットと称される)の分割が必要になることが良くある。その様な場合、ネットワークのルータ及びスイッチはパケットのフォーマットを、それが宛先エンティティにより受信され得るように、変更しなければならないかもしれない。例えば、もしパケットが大きなパケット・サイズを許すネットワークから生じて、パケットをより小さなサイズに限定する1つ以上のリンク又はローカル・ネットワークを通過するならば、それらのネットワーク同士を接続するスイッチはIPパケットを分割しなければならない。TCP/IPネットワーク環境の文脈においては、分割とリアセンブリの手続きは良く知られていて、『インターネット・プロトコル、コメントについての要求』(Internet Proto

10

20

30

40

50

col, Request for Comments (RFC) 791, by Information Sciences Institute University of Southern California (1981)) に詳しく記載されている。RFC 791 によれば、IP 分割は、IP パケットを、後に再組み立てされ得る任意の数のフラグメントに分割する。

【0010】

IP パケットを分割するために、ソース又は中間のシステム（例えば、スイッチ）は、2 つ以上の新しい IP フラグメントを生成し、オリジナル・パケットからの IP ヘッダ・フィールドの一部の内容をフラグメントの IP ヘッダの各々にコピーする。フラグメントの受信エンティティは、異なるパケットのフラグメントが混ぜ合わされないことを確実にするために IP 識別フィールド 110 の内容（すなわち、パケット識別子（ID））を使用する。すなわち、識別フィールド 110 は、1 つのパケットのフラグメントを他のものから区別するために使われる。IP フラグメント・オフセット・フィールド 114 は、フラグメントのオリジナル・パケットにおける位置を受信エンティティに知らせる。各フラグメントのフラグメント・オフセット・フィールド及び IP 全長フィールド 108 の内容は、オリジナル・パケットのそのフラグメントによりカバーされる部分を決定する。MF フラグ 112 は（例えば、リセット時に）最後のフラグメントを指し示す。完全な IP パケットの起源ホストは、IP 識別フィールド 110 を、そのパケットがネットワークにおいてアクティブである間ソースノ宛先アドレス・ペア及びプロトコル（例えば、TCP、UDP）のためにユニークである値にセットする。完全なパケットの起源ホストは、また、MF フラグ 112 を例えばゼロにセットし、IP フラグメント・オフセット・フィールド 114 をゼロにセットする。

【0011】

IP 識別フィールド 110 は 2 バイトのフィールドであり、これは 65535 に達した後にはラップ・アラウンドしなければならない（すなわち、番号付けを 1 から再開しなければならない）。毎秒数千個の IP パケットを生成する高速ネットワークでは、フィールド 110 の IP 識別子（IP-ID）は 1 秒あたりに幾度もラップ・アラウンドし得る。例えば、ギガビット・イーサネットでは、1 秒間に 80,000 個のパケットが生成され得、このことは IP-ID のラップ・アラウンドが 1 秒以内に起こり得ることを意味する。ネットワークがもっと高速になるに連れて、このラップ・アラウンドはもっと頻繁に起こるようになる。例えば、10 ギガビット・イーサネットでは、ラップ・アラウンドは数ミリ秒で起こり得る。

【0012】

その結果として、もしラップ・アラウンドされた IP-ID がオリジナル IP-ID により識別される別の IP パケットのフラグメントとリアセンブリされるならば、IP-ID ラップ・アラウンドはネットワークにおけるデータ破壊の原因となり得る。破壊を検出するためにヘッダ・チェックサム 120 を利用するインターネット・チェックサム・アルゴリズムは余り強力ではないので、TCP 又は UDP のような上層プロトコルはその様な破壊を検出し得ないかもしれない。この問題は、IP では、リアセンブリ・タイマの使用により対処されている（RFC 731 を参照）。IP フラグメント・リアセンブリは、リアセンブリ・タイマを使って、もし識別されたパケットの全フラグメントがリアセンブリ・タイマ期間内に受信されなければ、フラグメントを捨てる。IP フラグメント・リアセンブリの多くのインプリメンテーションがリアセンブリ・タイマのために通例 30 秒を使用する。非常に高速のネットワークでは、IP-ID はこの期間内に幾度もラップ・アラウンドし、フラグメントがラップ・アラウンドされた IP-ID の間違ったフラグメントと不適切に組み合わされる可能性を高める。

【発明の開示】

【発明が解決しようとする課題】

【0013】

高速ネットワークにおける IP-ID ラップ・アラウンドに対する解決策は、リアセン

10

20

30

40

50

ブリ・タイマを非常に低い値にセットし、これにより未解決の重複IP-IDの数を減らすことであった。しかし、これは、高速及び低速の両方のネットワークの多彩な速度を有するネットワーク環境においては性能低下の原因となる。なぜならば、IPパケットが単にまだ到着していないだけのときに低速リンクに沿ってフラグメントが間違っ捨てられてしまうからである。低速ネットワーク接続においてIPフラグメントを捨てることは、不必要な再送が上の方の層から要求されるという結果をもたらす。可変速度ネットワークにおけるIP-IDラップ・アラウンドに起因するデータ破壊問題に対して、公知解決策より優れた性能を提供する解決策を提供することが望ましいということが分かる。

【課題を解決するための手段】

【0014】

10

本発明により、データ処理システム・ネットワークを介しての通信のための改良された方法、システム及び製造物品が開示される。本発明の1つの好ましい方法は、送信されるべき特定のデータ・パケットのためにカスタマイズされたリアセンブリ時間を決定するステップと、そのカスタマイズされたリアセンブリ時間をそのデータ・パケットのフィールドに挿入することを含むそのデータ・パケットを生成するステップと、そのデータ・パケットを送信するステップとを含んでおり、そのデータ・パケットのレシーバは、そのカスタマイズされたリアセンブリ時間を利用して、そのフラグメントの受信後にそのカスタマイズされたリアセンブリ時間により示されるより長く分割されたままになっているそのパケットのフラグメントを捨てる。

【0015】

20

データ処理システム・ネットワークを介する通信のための本発明の他の好ましい方法は、データ・フラグメントを含むデータ・パケットのためのカスタマイズされたリアセンブリ時間を示すオプションを含むデータ・フラグメントをネットワークを介して受信するステップと、そのデータ・フラグメントの受信後、もしそのデータ・フラグメントがカスタマイズされたリアセンブリ時間により示されるより長くデータ・パケットから分割されたままになっていたらそのデータ・フラグメントを捨てるステップと、を含む。

【0016】

本発明の全ての目的、特徴、及び利点は、以下の詳細な記述において明らかになるう。

【0017】

本発明は、図面と関連して以下の記述において好ましい実施態様に関して説明され、図面では同様の数は同じ又は類似の要素を表す。

30

【0018】

本発明の代表的実施態様についての以下の詳細な記述においては、添付図面が参照され、それは本書の一部をなしており、また、そこには本発明が実施され得る特定の代表的実施態様が例として示されている。

【発明を実施するための最良の形態】

【0019】

本発明の代表的実施態様についての以下の詳細な記述においては、本発明が実施され得る特定の代表的実施態様が、当業者が本発明を実施することを可能にするために十分に詳しく記述されており、そして、他の実施態様が利用され得ること、また、本発明の範囲から逸脱すること無く論理に関する変更、アーキテクチャに関する変更、プログラムに関する変更、機械に関する変更、電気に関する変更、又は他の変更が行われ得ることが理解されなければならない。従って、以下の詳細な記述は限定的意味に解されるべきではなく、本発明の範囲は添付されている請求項のみによって定義される。特に、好ましい実施態様が以下でTCP/IPネットワーク環境に関して記述されるけれども、本発明がTCP/IPインプリメンテーションに限定されないことが理解されるであろう。

40

【0020】

ここで図面を参照すると、特に図1を参照すると、本発明の好ましい実施態様が実現され得るデータ処理システム・ネットワークが描かれている。データ処理システム・ネットワーク101は、インターネット107のような少なくとも1つのネットワークを介して

50

少なくとも1つのクライアント・システム105に結合された少なくとも1つのサーバ・システム103を含む。サーバ103とクライアント105との間のデータ転送は、TCP/IP仕様、及びファイル転送プロトコル(FTP)、ハイパーテキスト転送プロトコル(HTTP)、又は何らかの類似通信プロトコルに従う。単一のサーバ103と単一のクライアント・システム105とだけが示されているけれども、データ処理システム・ネットワーク101が、インターネット107を含む1つ以上の接続及びネットワークにより相互に接続された任意の数のサーバ及びクライアント・システム(図示されていない)を含み得ることが理解されるであろう。例証の目的上、サーバ103及びクライアント・システム105はTCP/IPを用いて通信するが、本発明はその様には限定されなくて、代替の実施態様では他の適切なプロトコルも採用することができる。

10

【0021】

図2は、本発明の好ましい実施態様で利用され得る、サーバ-クライアント・システムのための代表的なソフトウェア・アーキテクチャの描写である。サーバ103及びクライアント105は、各々、ソフトウェア・アーキテクチャ200で構築されている。最低のレベルでは、ユーザと他のソフトウェアとに高レベル機能を提供するためにオペレーティング・システム205が利用される。このオペレーティング・システムは、通例、BIOS(Basic Input Output System(基本入出力システム))を含む。通信ソフトウェア210は、オペレーティング・システム機能を直接起動することによって、又は間接的に、オペレーティング・システムを迂回してネットワークでの通信のためにハードウェアにアクセスすることによって、物理的通信リンクを介してインターネットのようなネットワークへの外部ポートを通しての通信を提供する。アプリケーション・プログラミング・インターフェース215は、システムのユーザ(個人であってもソフトウェア・ルーチンであっても)が、具体的な機能がどの様に実行されるかを懸念せずに標準的な矛盾しないインターフェースを用いてシステム機能を起動することを可能にする。インターネット・ソフトウェア220は、インターネット機能をコンピュータに持たせるために利用しうる幾つかの標準的な市販のパッケージのうちの任意の1つを表す。アプリケーション・ソフトウェア225は、通信ポートを通してデータに反応してユーザが求める所望の機能を提供するように設計された任意の数のソフトウェア・アプリケーションを表す。このレベルのアプリケーションは、インターネットのユーザによりアクセスされ得るデータ、ビデオ、グラフィックス、写真又はテキストを処理するために必要なものを含むことができる。

20

30

【0022】

図3に示されているように、TCP/IP及び類似プロトコルは、アプリケーション層310、トランスポート層312、ネットワーク層314、リンク層316及び物理層318を含むネットワークのための4層通信アーキテクチャにより利用される。各層は、次のように、種々の通信タスクを処理する責任を有する。

【0023】

物理層318は、一般に、ネットワーク内の種々の物理的装置間での通信のためにデータをビットに変換する。物理層で採用されるプロトコルの例は、イーサネット・カード、トークン・リング・カード、無線ネットワーク・カードなどのようなネットワーク・インターフェース・カードからのプロトコルを含むが、それらに限定はされない。電気電子技術者協会(IEEE)は、物理層で採用され得る種々のプロトコル、すなわちIEEE802.3標準規格(イーサネット)、IEEE802.5標準規格(トークン・リング)、及びIEEE802.11標準規格(無線イーサネット)のようなプロトコル、を公表している。

40

【0024】

リンク層316は、一般に、ネットワークを介して伝達されるデータを取り扱う。リンク層316(データ・リンク層又はネットワーク・インターフェース層とも称される)は、普通、デバイス・ドライバを含んでいて、また、例えば媒体アクセス制御(MAC)のようなプロトコルを含むことができる。連携して、物理層及びリンク層は、使用されるネ

50

ットワーク媒体と物理的にインターフェースするハードウェア細目を取り扱う。

【 0 0 2 5 】

ネットワーク層 3 1 4 (インターネット層とも称される) は、ネットワークの周囲でのデータのパケットの移動を取り扱う。例えば、ネットワーク層は、ネットワークを介して転送されるデータの種々のパケットのルーティングを処理する。TCP / IP スイートにおけるネットワーク層は、IP (インターネット・プロトコル)、ICMP (インターネット・コントロール・メッセージ・プロトコル)、及び IGM P (インターネット・グループ管理プロトコル) を含む、幾つかのプロトコルから構成される。

【 0 0 2 6 】

トランスポート層 3 1 2 は、2つのホスト・コンピュータ間でのデータ転送を容易にするネットワーク層 3 1 4 及びアプリケーション層 3 1 0 の間のインターフェースを提供する。トランスポート層は、アプリケーションからそれに渡されたデータを下のネットワーク層のために適切なサイズのチャンクに分割すること、受け取られたパケットについて肯定応答をすること、送られたパケットについて他端が肯定応答をしたことを確かめるためにタイムアウトをセットすること、などのことに関係する。TCP / IP プロトコル・スイートには明白に異なる2つのトランスポート・プロトコル、すなわち TCP (伝送制御プロトコル) 及び UDP (ユーザ・データグラム・プロトコル)、がある。TCP は、抜けの検出及び再送サービスを含む、データが2つのホスト間で適切に伝送されることを確実にする信頼性サービスを提供する。逆に、UDP は、データグラムと呼ばれるデータの packets を、そのデータが適切に転送されることを保証するメカニズムを全く提供すること無く、1つのホストから他方に単に送ることによって、遥かに簡単なサービスをアプリケーション層に提供する。UDP を用いるときには、アプリケーション層は信頼性機能を実行しなければならない。

【 0 0 2 7 】

アプリケーション層 3 1 0 は、具体的アプリケーションの細目を取り扱う。殆ど全てのインプリメンテーションが提供するありふれた TCP / IP アプリケーションが多数あり、それは (1) リモート・ログインのためのテルネット、(2) FTP、すなわちファイル転送プロトコル、(3) SMTP、すなわち電子メールのためのシンプル・メール・トランスファ・プロトコル、(4) SNMP、すなわちシンプル・ネットワーク・マネジメント・プロトコル、及び (5) ポスト・オフィス・プロトコル v 3 (POP 3) を含む。

【 0 0 2 8 】

インターネットのようなネットワークはルータにより相互に接続され、それは各々2つ以上のネットワークを相互に結合させる。代表的なルータは、入力接続及び出力接続を有する特別目的ハードウェア・ボックスと、イーサネット、トークン・リング、ポイント・ツー・ポイント・リンクなどの様々なタイプの物理的ネットワークの接続を可能にする専用のハードウェア及び埋め込み型ソフトウェアのうちの少なくとも一方とを含む。図4は、ルータ 4 3 6 によりトークン・リング・ネットワーク 4 3 4 に結合されたイーサネット・ネットワーク 4 3 2 を含むインターネット 4 0 0 を示す。図4は通信する2つのホストを示すだけであるが、イーサネット・ネットワーク上の任意のホストがその上の任意のホストと、又はトークン・リング・ネットワーク上の任意のホストと、またその逆に、通信することができる。

【 0 0 2 9 】

図4に示されているように、ルータ 4 3 6 は、ネットワーク層モジュール 4 3 8 (この場合には IP モジュール) と、ホスト・ネットワークとの接続のための適切なネットワーク・ドライバ、すなわちイーサネット・ドライバ 4 4 0 及びトークン・リング・ドライバ 4 4 2 とを含む。アプリケーション層において、ネットワークは FTP クライアント 4 2 0 と FTP サーバ 4 2 2 とを含み、トランスポート層においてはネットワークは TCP クライアント 4 2 4 と TCP サーバ 4 2 6 とを含み、ネットワーク層においてはネットワークは IP クライアント 4 2 8 と IP サーバ 4 3 0 とを含む。殆どのネットワーク・アプリケーションは、一端がクライアントで他方の側がサーバであるように設計される。サーバ

は、種々のクライアントに何らかのタイプのサービスを、この場合にはサーバ・ホスト上のファイルへのアクセスを、提供する。各層は、同じ層に存するそのピアと通信するための1つ以上のプロトコルを有する。これらの通信プロトコルは、アプリケーション層のFTPプロトコル444と、トランスポート層のTCPプロトコル446と、ネットワーク層のIPプロトコル448と、リンク層のイーサネット・プロトコル450及びトークン・リング・プロトコル454とを含む。アプリケーション層にとってはユーザ・プロセスを処理するのは普通のことであるが、下の方の3つの層（トランスポート、ネットワーク及びリンク）は、UNIX（登録商標）又はWindows（登録商標）オペレーティング・システムのようなオペレーティング・システムのカーネルにおいて実現される。例えば、ネットワーク・インターフェース層の目的は通信媒体（イーサネット、トークン・リングなど）の細目を取り扱うことであり、アプリケーション層の目的は1つの特定のユーザ・アプリケーション（FTP、テルネットなど）を取り扱うことである。

10

【0030】

アプリケーション層とトランスポート層とは、エンド・ツー・エンドのプロトコル（FTPプロトコル444、TCPプロトコル446）を使用する。ネットワーク層は、2つのエンド・システムとその間の全ての中間システム（明瞭性を目的として、ここでは1つの中間システムだけが示されている）とで使用されるホップ・ツー・ホップ（hop-to-hop）のプロトコルを提供する。例えば、ルータ436のIPモジュール438は、IPプロトコル448によって2つのホストに接続されている。ルータに接続された種々のタイプのホスト・ネットワークに特有の、それらのネットワークとルータとの間のリンク層における通信を取り扱うリンク層プロトコルもある。従って、イーサネット・プロトコル450はルータ436内のイーサネット・ドライバ440とイーサネット・ネットワーク432上のホストのイーサネット・ドライバ452との間の通信を取り扱うために使用され、トークン・リング・プロトコル454はルータ436のトークン・リング・ドライバ442とトークン・リング・ネットワーク434上のホストのトークン・リング・ドライバ456との間の通信を取り扱うために使用される。

20

【0031】

アプリケーションがTCP/IPを用いてデータを送るとき、データは、ビットのストリームとしてネットワークに送り出されるまで各層を通り、プロトコル・スタックに沿って送られる。図5に示されているように、プロトコル・スタック500の各層は、自分が受け取ったデータにヘッダをプリペンドすることによって（且つ、時にはトレーラ情報を付け加えることによって）そのデータに情報を付け加える。例えば、アプリケーション層においてはアプリケーション・データ584を形成するべくアプリケーション・ヘッダ580がユーザ・データ582にプリペンドされる。トランスポート層においてはトランスポート・プロトコル・ヘッダがアプリケーション・データにプリペンドされる。図5の場合、トランスポート層はTCPであり、従ってTCPヘッダ586がアプリケーション・データ584にプリペンドされ、これにより、ネットワーク層IPに送られるTCPフレーム588を形成する。TCPヘッダ586は20バイトを含む。同様に、ネットワーク層において、ネットワーク層ヘッダがトランスポート層データにプリペンドされる。TCP/IPの場合、IPデータグラム592を形成するべくIPヘッダ590がTCPフレーム588にプリペンドされる。IPヘッダ590も20バイトを含む。最後に、リンク層において、データのフレームを形成するべく、ネットワーク層から受け取られたデータにイーサネット・ヘッダ594のような媒体ヘッダが付加される。媒体がイーサネットであるような場合には、媒体トレーラもデータの末端に添付される。例えば、図5においては、イーサネット・フレーム598を形成するべくイーサネット・トレーラ96がイーサネット・ヘッダ594及びIPデータグラム592に添付される。イーサネット・フレームは、ネットワークを横切って流れるオリジナル・アプリケーション・メッセージ・データに対応するビットのストリームを含む。ヘッダの下の数（14, 20, 20, 4）はバイト単位でのヘッダの代表的サイズであり、例えば、イーサネット・ヘッダ94は14バイトを含む。フレームのサイズは、データ・パケットを転送するために使用されるタイプ

30

40

50

のネットワークの最大送信単位 (MTU) によって制限される。例えば、イーサネット・ネットワークのMTUは1500バイトである。ネットワーク層は分割 (データグラムをより小さな断片に分解すること) を自動的に実行するので、各フラグメントはネットワークのMTUより小さい。

【0032】

図1及び4に戻って参照して、ヘテロジニアス・ネットワーク環境で使用される本発明の好ましい実施態様に設けられたダイナミックIPパケット・リアセンブリ・タイマの動作が説明される。本発明の種々の実施態様において、IPフラグメントを捨てるためのリアセンブリ・タイマは、各IPパケットについて個別に動的に計算される。これらの実施態様は、IPネットワーク・システムにおけるIP-IDラップ・アラウンドに起因するデータ破壊を避ける上で大幅な改善を提供する。

10

【0033】

第1の実施態様では、IPパケット・センダ (サーバ103又はクライアント105) は、生成されるIPパケットについてリアセンブリ・タイマ値を明示するIPパケット・ヘッダ100内の新しいIPオプション126を明示する。本発明の好ましい実施態様に従って、この新しいオプションはリアセンブリ・タイマ・オプションと称される。本発明の好ましい実施態様によるIPヘッダのためのリアセンブリ・タイマ・オプションのためのフォーマットが図6に示されている。

【0034】

図6に見られるように、リアセンブリ・タイマ・オプション602は、そのリアセンブリ・タイマ・オプションについてIPで認識されるオプション126のオプション番号を明示する1バイト・フィールド604を含む。フィールド604は、ペイロード長さをバイト数で明示する1バイト・フィールドである。この例では、リアセンブリ・タイマ・オプション602へのペイロードはリアセンブリ時間の数値であり、それは4バイト・フィールドとして明示される。フィールド608は、このIPパケットのためにレシーバのリアセンブリ・タイマに割り当てられるべきリアセンブリ時間を示す。センダは、ネットワーク、センダ及びレシーバ、及び送られるIPパケットに関連するいろいろな要素に基いてリアセンブリ時間608のための値を決定する。

20

【0035】

リアセンブリ時間608をセットするための好ましい実施態様では、センダはリアセンブリ時間をセットすることについて2つの考慮事項を有する。第1の考慮事項として、センダは、IPパケットのための宛先アドレスがセンダと同じサブネット上にあるか否かを判定する。センダとレシーバとが同じサブネット上にあるならば、センダはリアセンブリ時間608を、センダとレシーバとを接続するサブネットの速度に基かせることができる。例えば、毎秒80,000個のパケットを生成することのできるギガビット・イーサネット・ネットワークでは、リアセンブリ・タイマは1秒の半分 (0.5秒) に送られ得、10ギガビット・イーサネット・ネットワークのためには、リアセンブリ・タイマはこの時間の十分の一 (すなわち、50ミリ秒) にセットされ得る。

30

【0036】

第2の考慮事項として、センダは、再送タイマ計算として用いられるTCPにより計算されるラウンド・トリップ時間 (RTT) 値を検討することができる。この値を用いることにより、センダは、正常なネットワーク状況下で期待されるラウンド・トリップ時間の推定値をレシーバに提供し、これによりRTTより長くかかるフラグメントが適切に再送されるべきことを提案することができる。このアルゴリズムを使って、センダはリアセンブリ時間608をRTTに、又はRTTプラスある程度のデルタ (例えば、 $RTT * 1.5$) に等しくセットすることができる。リアセンブリ時間608をセットするときにセンダが任意の数の要素及びアルゴリズムを考慮し得ること、また、いずれか1つの考慮事項だけ又はそれらの若しくは他の考慮事項の組み合わせが使用され得ることが理解されるべきである。

40

【0037】

50

本発明の代わりの好ましい実施態様では、適切なリアセンブリ・タイマを決定するためにセンダとレシーバとは同期化タイミング・アルゴリズムを調整する。この好ましい実施態様は、ネットワーク・プロトコルがラウンド・トリップ時間の推定値を有さない環境において又は宛先が同じサブネット上に無い場合に、使用され得る。

【 0 0 3 8 】

図 7 は、本発明の好ましい実施態様による、“タイム・スタンプ・リアセンブリ・タイマ”と称される新しい IP オプション 1 2 6 のためのデータ・フォーマットを示す。タイム・スタンプ・リアセンブリ・タイマ・オプション 7 0 2 は、IP パケットの IP ヘッダ 1 0 0 の中のオプション 1 2 6 である。フィールド 7 0 4 は、タイム・スタンプ・リアセンブリ・タイマ・オプションについて IP で認識されるオプション番号を示す。フィールド 7 0 6 は、オプションのペイロードの長さを示す 1 バイト・フィールドである。この場合、オプションは 4 バイト・ペイロードを示す。フィールド 7 0 8 は現在時刻の数値を明示するが、それは 4 バイト・フィールドとして明示される。タイム・スタンプ 7 0 8 は、IP ヘッダ 1 0 0 が生成されるときに送信サーバによってセットされる。

【 0 0 3 9 】

センダは、始めにネットワーク・タイム・プロトコルを用いてその時間を同期化し、現在時刻をタイム・スタンプ 7 0 8 として送る。IP ヘッダ 1 0 0 のオプション 1 2 6 に含まれているタイム・スタンプ・リアセンブリ・タイマ・オプション 7 0 2 を受け取ると、レシーバは、現在時刻とタイム・スタンプ 7 0 8 との差を計算することによってそのフラグメントの横断時間を計算する。レシーバは、その後、IP パケットのフラグメントについてリアセンブリ・タイマを決定するために、計算された横断時間に“デルタ”期間を加える。例えば、フラグメントの横断時間を計算した後、レシーバは、リアセンブリ・タイマを IP パケットの計算された横断時間の 1 . 5 倍にセットするべく横断時間の半分に等しい量の時間を加えることができる。この計算されたリアセンブリ・タイマは、普通は、受け取られたパケットが IP パケットの第 1 フラグメントであったならば、又は計算されたリアセンブリ・タイマが受信システムにおいて現在使用されているタイマより大きいと判定されたならば、レシーバによって使用される。リアセンブリ時間をセットするためのこのアルゴリズムは、望まれる粒状度 (g r a n u l a r i t y) に応じて、IP パケットの第 1 フラグメントだけに関して又は全てのフラグメントに関して使用され得る。リアセンブリ・タイマ及びタイム・スタンプ・リアセンブリ・タイマ・オプションが利用可能でないか又は受け取られないネットワーク環境においては、レシーバはリアセンブリ・タイマの代わりに前もってプログラムされた値を使用することができる。

【 0 0 4 0 】

ここで図 8 を参照すると、本発明の好ましい実施態様によって、リアセンブリ時間を含む IP パケットを生成するためのプロセスの流れ図が示されている。プロセスは、IP パケットを生成するように送信サーバが指令されたときにステップ 8 0 2 から開始する。判定ブロック 8 0 4 で、送信サーバにより、IP 宛先が IP ソース・アドレスと同じサブネット上にあるか否かが判定される。もし肯定ならば、サブネットは既知の高速スループットを有し、IP フラグメントのために非常に短い再送タイマがセットされるべきことを示唆する。プロセスはステップ 8 0 6 に進み、ここで送信サーバは、サーバとクライアントとの間の既知サブネットの速度に対応する非常に短い時間にセットされたリアセンブリ時間 6 0 8 を有するリアセンブリ・タイマ・オプション 6 0 2 をオプション 1 2 6 に含む IP パケットを生成する。その後、サーバはステップ 8 0 8 で示されているように IP パケット (またはフラグメント) を送信する (プロセスは終了する) 。

【 0 0 4 1 】

判定ブロック 8 0 4 に戻って、もし IP 宛先がセンダと同じサブネット上になければ、プロセスは判定ブロック 8 1 0 に進み、ここで送信サーバは、トランスポート層がデータ通信についてのラウンド・トリップ時間を計算したか否かを判定する。例えば、TCP は RTT を計算して、この計算結果を TCP ヘッダに含める。もし肯定ならば、プロセスはステップ 8 1 2 に進み、ここで送信サーバは、トランスポート層のために計算されたラウ

ンド・トリップ時間にセットされたリアセンブリ・タイマ・オプション 602 を含む IP パケットを生成する。TCP において、RTT にセットされる。その後、プロセスはステップ 808 で終了し、このとき送信サーバは IP パケットを送信する。

【0042】

判定ブロック 810 に戻って、もしトランスポート層がラウンド・トリップ時間を計算していなかったならば、又はその様な計算がトランスポート層によってサポートされないならば、プロセスはステップ 814 に進み、ここで送信サーバはタイム・スタンプ・リアセンブリ・タイマ・オプション 702 を含む IP パケットを生成する。センダは、IP パケットが生成されるときに現在時刻を表す数値でタイム・スタンプ 708 をセットする。その後、ステップ 808 に見られるように、送信サーバが IP パケットをクライアントに送信した後にプロセスは終了する。いまや認められるであろうように、これらのヘッダ・オプションの各々は別々に又は組み合わされて使用され得る。

【0043】

ここで図 9 を参照すると、本発明の好ましい実施態様による、IP フラグメントを受け取ったことに応答してリアセンブリ・タイマをセットするためのプロセスの流れ図が示されている。プロセスは、ネットワーク上でクライアントにより IP フラグメントが受け取られたときにステップ 902 から始まる。プロセスは判定ブロック 904 に進み、ここで受信クライアントは、IP フラグメントがリアセンブリ・タイマ・オプション 602 を含むか否か判定する。肯定ならば、ステップ 906 に見られるように、受信クライアントは、受け取られた IP フラグメントについてのリアセンブリ時間をオプションのペイロード・フィールド 708 においてセットされている値にセットする。もし IP フラグメントがリアセンブリ・タイマ・オプション 602 を含まなければ、プロセスは判定ブロック 908 に進み、ここで受信クライアントは IP フラグメントがタイム・スタンプ・リアセンブリ・タイマ・オプション 702 を含むか否か判定する。肯定ならば、ステップ 910 に見られるように、レシーバは、ペイロードに含まれているタイム・スタンプ 708 をレシーバにおける現在時刻から引くことによって IP フラグメントの横断時間を計算する。レシーバは、その後、ステップ 912 に見られるように、計算された横断時間に基いてリアセンブリ・タイマをセットする。判定ブロック 908 に戻って、もしタイム・スタンプ・リアセンブリ・オプションがセットされていないければ、プロセスはブロック 914 で終了し、ここで受信クライアントは、リアセンブリ・タイマを、ユーザにより予め決定された又は IP フラグメントを捨てて再送を要求するために最適の時間としてシステムにより計算された格納されている値に、セットする。

【0044】

上記の方法が一般的には 1 つ以上のプロセッサ（図示されていない）上で動作するソフトウェアで実行されること、また、そのソフトウェアが磁気又は光学コンピュータ・ディスクのような任意の適切なデータ・キャリア（これも図示されていない）に担持されるコンピュータ・プログラム・エレメントとして提供され得ることが当業者に理解されるであろう。データ伝送のためのチャネルも同様にあらゆる種類の記憶媒体、並びに、有線又は無線信号媒体のような信号伝達媒体を含むことができる。

【0045】

従って、本発明はコンピュータ・システムに使用されるコンピュータ・プログラムとして適切に具体化され得る。この様なインプリメンテーションは一連のコンピュータ可読命令を含むことができ、それは、例えばディスク、CD-ROM、ROM、若しくはハード・ディスクなどのコンピュータ可読媒体のような有形の媒体に固定され、又は、光学又はアナログ通信回線を含むがこれらに限定はされない有形媒体に沿って又はマイクロウェーブ、赤外線若しくはその他の伝送技術を含むがこれらに限定はされない無線技術を用いてモデム若しくは他のインターフェース装置を介してコンピュータ・システムに伝送可能である。その一連のコンピュータ可読命令は前記の機能の全部又は一部を具体化する。

【0046】

当業者は、その様なコンピュータ可読命令が多くのコンピュータ・アーキテクチャ又はオペレーティング・システムに用いられる多数のプログラミング言語で書かれ得ることを理解するであろう。更に、その様な命令は、半導体、磁気、若しくは光学を含むがこれらに限定はされない現在又は将来の任意のメモリ技術を用いて格納され得、或いは、光学、赤外線、又はマイクロウェーブを含むがこれらに限定はされない現在又は将来の任意の通信技術を用いて伝送され得る。その様なコンピュータ・プログラムが、例えばコンピュータ・システムのシステムROM若しくは固定ディスクに事前ロードされたシュリンク・ラップされたソフトウェアなど、印刷された若しくは電子式の取扱説明書が添付されている取り外し可能な媒体として頒布され得、又は例えばインターネット若しくはワールド・ワイド・ウェブなどのネットワークを介してサーバ若しくは電子掲示板から頒布され得ることが意図されている。

10

【0047】

本発明は好ましい実施態様に関して具体的に示され記述されたが、本発明の範囲から逸脱せずにその形及び詳細に関して種々の変更をなし得ることが当業者に理解されるであろう。例えば、本発明は、コンピュータ・プログラミング・ソフトウェア、ファームウェア若しくはハードウェアの任意の組み合わせを用いて実施され得る。本発明を実行し或いは本発明の装置を構築する準備段階として、本発明によるコンピュータ・プログラミング・コード（ソフトウェア若しくはファームウェア）は、一般的には、固定（ハード）ディスク、ディスク、光ディスク、磁気テープ、ROM、PROMなどの半導体メモリのような1つ以上の機械可読記憶媒体に格納され、これにより本発明の製造物品を生じさせる。そのコンピュータ・プログラミング・コードを含む製造物品は、そのコードを記憶装置から直接実行することにより、コードを記憶装置からハード・ディスク、RAMなどのような他の記憶装置にコピーすることにより、又はコードをリモート実行するために伝送することにより、使用される。本発明の方法形は、本発明によるコードを含む1つ以上の機械可読記憶装置を適切な標準的コンピュータ・ハードウェアと組み合わせてその中に含まれているコードを実行することによって実行され得る。本発明を実行するための装置は、本発明により符号化されたコンピュータ・プログラムへのネットワーク・アクセスを含むか若しくは有する1つ以上のコンピュータ及び記憶システムであり得る。

20

【図面の簡単な説明】

【0048】

30

【図1】本発明の好ましい実施態様を実現し得るデータ処理システム・ネットワークを示す。

【図2】本発明の好ましい実施態様に利用され得る、サーバ・クライアント・システムのための代表的なソフトウェア・アーキテクチャの描写である。

【図3】本発明の好ましい実施態様に利用され得る、ネットワークのための4層通信アーキテクチャである。

【図4】本発明の好ましい実施態様に利用され得る、ルータにより相互に接続されたインターネットのようなネットワークのブロック図を示す。

【図5】本発明の好ましい実施態様に利用され得る、データ・パケットの、TCP/IPプロトコル・スタックを通過するときのフォーマットを示す。

40

【図6】本発明の好ましい実施態様による、IPヘッダのためのリアセンブリ・タイマ・オプションのためのフォーマットを示す。

【図7】本発明の好ましい実施態様による、IPヘッダのためのタイム・スタンプ・リアセンブリ・タイマ・オプションのためのフォーマットを示す。

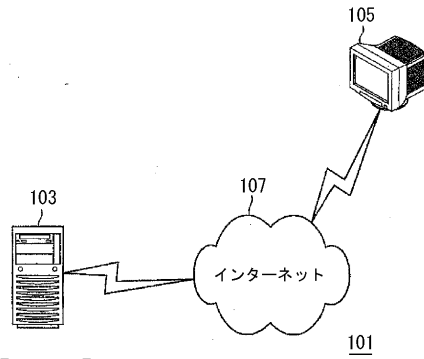
【図8】本発明の好ましい実施態様による、リアセンブリ時間を含むIPパケットを生成するためのプロセスの流れ図を示す。

【図9】本発明の好ましい実施態様による、IPフラグメントを受信したことに応答してリアセンブリ・タイマをセットするためのプロセスの流れ図を示す。

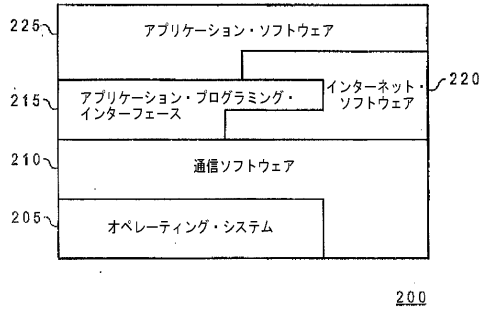
【図10】IPヘッダ部分及びペイロード/データ部分を含むIPパケット・データ・フォーマットの略ブロック図である。

50

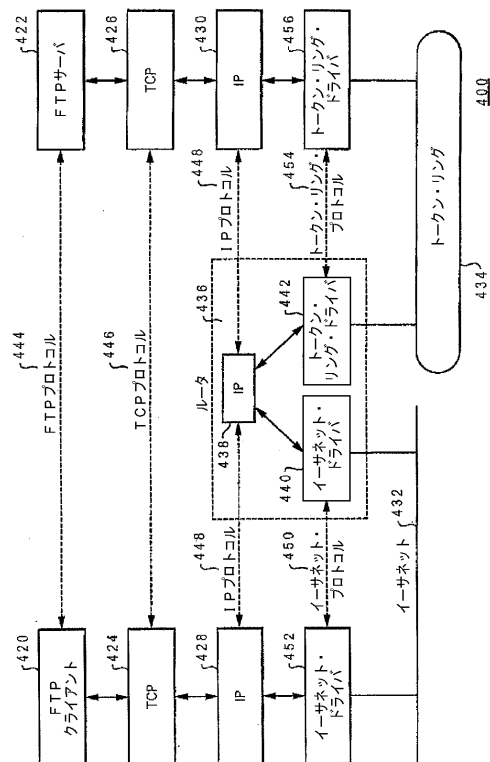
【図 1】



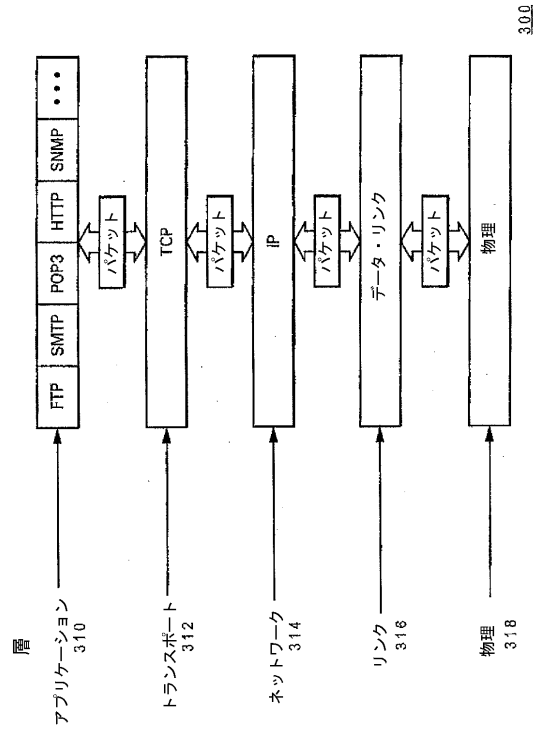
【図 2】



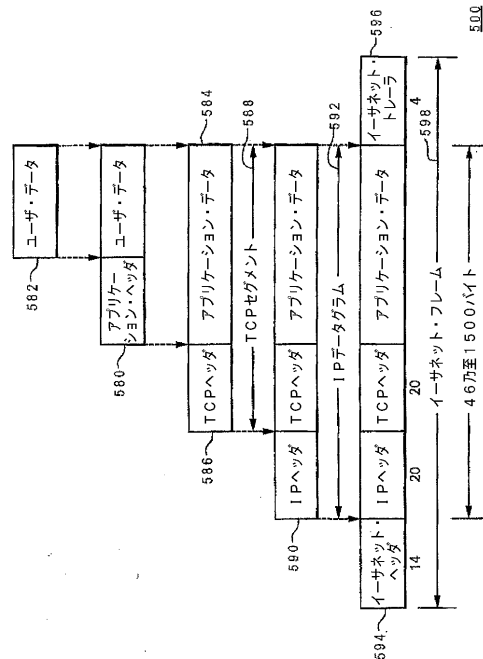
【図 4】



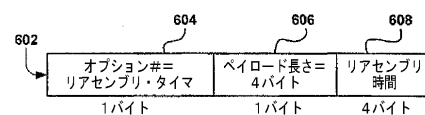
【図 3】



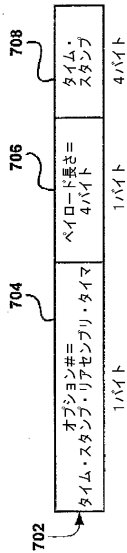
【図 5】



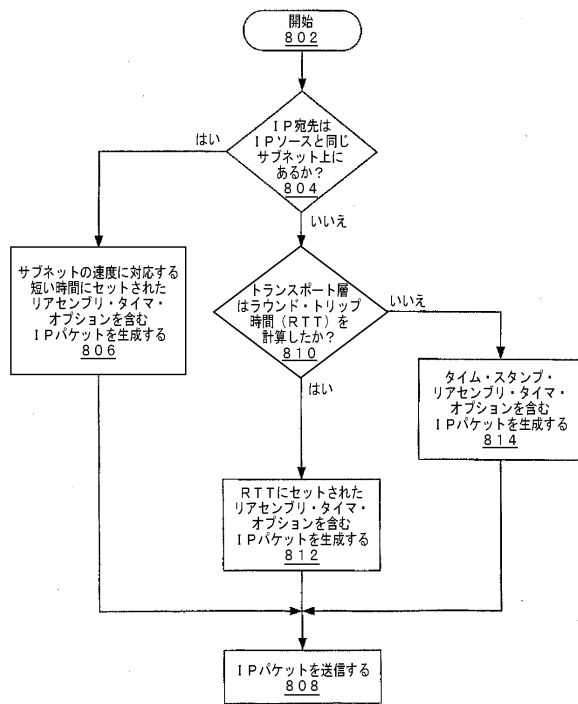
【図 6】



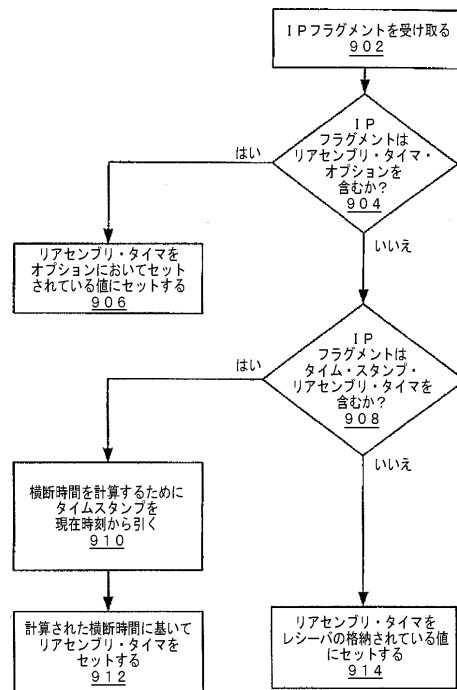
【図 7】



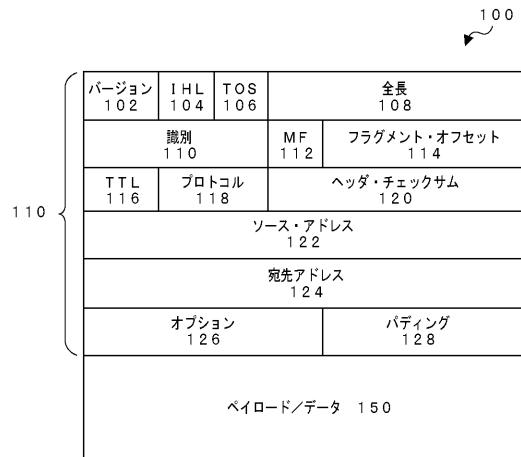
【図 8】



【図 9】



【図 10】



フロントページの続き

(74)代理人 100086243

弁理士 坂口 博

(72)発明者 マンジュナサ、シャンカー

アメリカ合衆国 7 8 7 5 8 テキサス州 オースティン ストーンホロウ・ドライブ 1 1 9 1
5

(72)発明者 ヴァラブハネニ、ヴァス

アメリカ合衆国 7 8 7 5 0 テキサス州 オースティン トップリッジ・ドライブ 9 1 0 3

(72)発明者 ヴェンカットスブラ、ヴェンカット

アメリカ合衆国 7 8 7 1 7 テキサス州 オースティン ブライアー・クリーク・トレイル 8
6 0 7

(72)発明者 ヤングマン、リチャード、ペリー

アメリカ合衆国 7 8 6 1 3 テキサス州 シーダー・パーク オーククレスト・ドライブ 4 0
9

審査官 森谷 哲朗

(56)参考文献 特開平 0 8 - 0 3 2 6 2 3 (J P , A)

特開 2 0 0 4 - 3 1 2 0 6 2 (J P , A)

特開平 0 2 - 0 0 2 7 5 7 (J P , A)

特開 2 0 0 3 - 0 6 0 6 9 9 (J P , A)

特開平 1 1 - 1 4 5 9 7 5 (J P , A)

特開平 0 9 - 1 9 1 3 1 4 (J P , A)

特開 2 0 0 4 - 2 0 1 1 9 6 (J P , A)

(58)調査した分野(Int.Cl. , D B 名)

H04L 29/08

H04L 12/56