

(19) 日本国特許庁(JP)

(12) 特許公報(B2)

(11) 特許番号

特許第3613863号

(P3613863)

(45) 発行日 平成17年1月26日(2005.1.26)

(24) 登録日 平成16年11月12日(2004.11.12)

(51) Int. Cl.⁷

F I

H04L 29/04

H04L 13/00 303Z

G06F 13/00

G06F 13/00 357Z

請求項の数 13 (全 32 頁)

(21) 出願番号	特願平7-333028	(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
(22) 出願日	平成7年12月21日(1995.12.21)	(74) 代理人	100075096 弁理士 作田 康夫
(65) 公開番号	特開平9-93304	(72) 発明者	北井 克佳 東京都国分寺市東恋ヶ窪1丁目280番地 株式会社日立製作所中央研究 所内
(43) 公開日	平成9年4月4日(1997.4.4)	(72) 発明者	吉沢 聡 東京都国分寺市東恋ヶ窪1丁目280番地 株式会社日立製作所中央研究 所内
審査請求日	平成14年12月5日(2002.12.5)		
(31) 優先権主張番号	特願平7-182429		
(32) 優先日	平成7年7月19日(1995.7.19)		
(33) 優先権主張国	日本国(JP)		

最終頁に続く

(54) 【発明の名称】 ネットワーク接続システム及び並列ネットワーク接続方法

(57) 【特許請求の範囲】

【請求項1】

複数のネットワーク・インタフェースを持つサーバ計算機から、ネットワークに接続されたネットワーク接続手段を介して、クライアント計算機に対してデータ通信を行うネットワーク接続システムにおいて、

前記ネットワーク接続手段として、

前記ネットワークと前記サーバ計算機との間に設けられ、前記サーバ計算機の複数のネットワーク・インタフェースにそれぞれ対応した帯域予約可能な複数の通信路により前記サーバ計算機と前記ネットワークとの接続を行うLANスイッチと、

前記クライアント計算機から要求されるサービス品質(QoS)に応じて、前記複数の帯域予約可能な通信路の一つを選択する選択手段を備え、

前記選択手段は、前記サーバ計算機内のルーティング・テーブルを用い、そのルーティング・テーブル中の、前記クライアント計算機が接続されるネットワークのアドレスと、前記複数の通信路に対応した前記サーバ計算機のネットワーク・インタフェースのアドレスを用いて、パケットを送出する前記ネットワーク接続手段のネットワーク・アドレスを求めることにより、通信路を選択することを特徴とするネットワーク接続システム。

【請求項2】

請求項1記載において、前記LANスイッチは、ATMスイッチで構成されることを特徴とするネットワークネットワーク接続システム。

【請求項3】

10

20

請求項 1 記載において、前記サーバ計算機内に、前記ネットワーク・インタフェースごとに仮想チャネルの帯域及び動的な負荷の統計情報を保持する Q O S 管理テーブルを設け、通信路の確立要求時にこのテーブル Q O S 管理テーブルを参照することを特徴とするネットワーク接続システム。

【請求項 4】

請求項 1 または 3 記載において、前記サービス品質 (Q O S) は、通常のデータ通信である B E (Best Effort)、バースト転送向けのデータ通信である G B (Guaranteed Burst)、ビデオや音声などのストリーム・データの通信である G S (Guaranteed Stream) の 3 種類からなるサービス・クラス、ピークバンド幅 (M b p s)、平均バンド幅 (M b p s) の 3 項目で構成することを特徴とするネットワーク接続システム。

10

【請求項 5】

複数のネットワーク・インタフェースを備え、他の計算機とネットワーク通信を行う計算機において、

前記計算機に、パケットの最終宛先ネットワーク・アドレスと、前記ネットワーク・インタフェースのアドレスと、両者のアドレスから求められる、パケットの次の送付先であるゲートウェイのネットワーク・アドレスとを各エントリに保持するルーティング・テーブルを備え、

少なくとも 1 つのネットワーク・インタフェースを選択し、選択されたネットワーク・インタフェースから上記ゲートウェイにパケットを送付することを特徴とする計算機。

【請求項 6】

20

請求項 5 記載において、前記計算機は、前記ネットワーク・インタフェースのそれぞれの使用状況を保持するサービス品質 (Q O S) 保持手段を備え、前記ネットワーク・インタフェースを選択するときには、前記ルーティング・テーブルのエントリの中から、最終宛先ネットワーク・アドレスが前記他の計算機のネットワーク・アドレスと等しい少なくとも 1 つのエントリを選択し、選択されたルーティング・テーブルのエントリのネットワーク・インタフェースのアドレスで示されるネットワーク・インタフェースの上記 Q O S 保持手段を比較した結果に基づいて、少なくとも 1 つのネットワーク・インタフェースを選択することを特徴とする計算機。

【請求項 7】

請求項 6 記載において、前記他の計算機が、前記計算機に対して通信路を開設することを要求する時に、前記他の計算機から指定された、通信路のサービス・クラスと平均通信量とピーク通信量からなる Q O S パラメタの指定に対し、前記計算機は、前記ルーティング・テーブルのエントリの中から、最終宛先ネットワーク・アドレスが前記他の計算機のネットワーク・アドレスと等しい少なくとも 1 つのエントリを選択し、選択されたルーティング・テーブルのエントリのネットワーク・インタフェースのアドレスで示されるネットワーク・インタフェースの前記 Q O S 保持手段と前記 Q O S パラメタを比較し、前記 Q O S パラメタで指定される条件を満たすことができるように、少なくとも 1 つのネットワーク・インタフェースを選択することを特徴とする計算機。

30

【請求項 8】

請求項 6 記載において、前記他の計算機と前記計算機が通信路を開設するとき、通信路の開設時に送信する同期 (S Y N) パケットに付随させて、前記計算機が選択した少なくとも 1 つの前記ネットワーク・インタフェースのネットワーク・インタフェースのアドレスを前記他の計算機に送付することを特徴とする計算機。

40

【請求項 9】

請求項 6 記載において、前記計算機は、前記複数のネットワーク・インタフェースに接続される複数の第 1 の通信路を第 2 の通信路に束ねて公衆網へ送付する L A N スイッチを備え、前記 L A N スイッチは、第 1 の通信路と第 2 の通信路とのスイッチングを指定するスイッチング保持手段を備え、前記計算機が選択した第 1 の通信路が、前記他の計算機が指定した第 1 の通信路とは異なった場合には、前記第 1 の計算機が、前記スイッチング保持手段を、前記第 1 の計算機が選択した第 1 の通信路になるように、前記 L A N スイッチに

50

指示することを特徴とする計算機。

【請求項 10】

複数のネットワーク・インタフェースを備えた第1の計算機と、ネットワークにゲートウェイ計算機を介して接続された第2の計算機を備え、ネットワーク通信を行う際に、前記第1の計算機に、パケットの最終宛先ネットワーク・アドレスと、前記ネットワーク・インタフェースのアドレスと、両者のアドレスから求められる、パケットの次の送付先であるゲートウェイ計算機のネットワーク・アドレスとを各エントリに保持するルーティング・テーブルを備え、前記第1の計算機は、少なくとも1つのネットワーク・インタフェースを選択し、選択されたネットワーク・インタフェースから上記ゲートウェイにパケットを送付することを特徴とする並列ネットワーク接続方法。

10

【請求項 11】

請求項 10 記載において、前記第2の計算機が、前記第1の計算機に対して通信路の開設を要求したとき、前記第1の計算機はルーティング・テーブルのエントリの中から、最終宛先ネットワーク・アドレスが第2の計算機のネットワーク・アドレスと等しい少なくとも1つのエントリを選択し、選択したルーティング・テーブルのエントリのネットワーク・インタフェース・アドレスで示されるネットワーク・インタフェースのQOS保持手段と前記第2の計算機が指定するQOSパラメタを比較し、QOSパラメタで指定される条件を満たすことができるように、少なくとも1つのネットワーク・インタフェースを選択し、そのネットワーク・インタフェースを通して前記第1の計算機はゲートウェイにパケ

20

【請求項 12】

請求項 10 記載において、前記第1の計算機は前記第2の計算機に対して、通信路の開設時に送信する同期(SYN)パケットに付随させて、選択した全てのネットワーク・インタフェースのネットワーク・アドレスを教えることを特徴とする並列ネットワーク接続方法。

【請求項 13】

請求項 10 記載において、帯域の大きいネットワーク・インタフェースを備えた第2の計算機の場合に、前記第1の計算機は複数のネットワーク・インタフェースを選択し、前記第2の計算機に対して、通信路の開設時に送信する同期(SYN)パケットに付随させて、選択した複数のネットワーク・インタフェースのネットワーク・アドレスを教え、もって前記選択した複数のネットワーク・インタフェース並列に用いてデータ通信を行なうことを特徴とする並列ネットワーク接続方法。

30

【発明の詳細な説明】

【0001】

【発明の属する技術分野】

本発明は、複数のネットワーク・インタフェースを備えた計算機のネットワーク接続システム及び並列ネットワーク接続方法に関し、特に、ATMネットワークのような多様なサービス品質(QOS)に応えることが要求されるネットワーク接続システムに係わる。

【0002】

【従来の技術】

図15に、複数のネットワーク・インタフェースを備えたサーバにおける従来のネットワーク接続を示す。複数のネットワーク・インタフェースを備えたサーバとしては、例えば、Auspe社(NFSサーバ)や、Maximum Strategy社のストレージサーバが知られている。

【0003】

図15において、500はサーバ、501、502、503はネットワーク・インタフェース、505はルーティング・テーブル、510、520、530はそれぞれネットワーク・インタフェース501、502、503に接続されるネットワーク、511ないし513、521ないし523、531ないし533は、それぞれネットワーク510、52

50

0、530に接続されるクライアント(ワークステーションやPCなど)を示す。540は公衆網570に接続されるネットワーク、541は、ネットワーク540に接続されるクライアントを示す。尚、サーバに備えられるネットワーク・インタフェースは、実体としてはインタフェース用のカード及び関連するソフトウェアなどで構成される。

【0004】

サーバ500は、ネットワーク510に接続されたゲートウェイ550とPBX(Private Branch Exchange:構内交換機)560を介して、公衆網570と接続される。クライアント511、512、513にはそれぞれ、ネットワーク510のネットワーク・アドレスnet1.1、net1.2、net1.3が割り当てられ、ネットワーク・インタフェース501には、ネットワーク510のネットワーク・アドレスnet1.11が割り当てられている。クライアント511、512、513は、net1.11を使ってのみ、つまりネットワーク510を通してのみサーバ500とデータ通信を行うことができる。同様に、クライアント521、522、523はネットワーク520を通してサーバ500とデータ通信を行い、クライアント531、532、533はネットワーク530を通してサーバ500とデータ通信を行う。公衆網570に接続されたクライアント541は、PBX560、ゲートウェイ550、ネットワーク510を経由してサーバ500とデータ通信を行う。LAN1~LAN3に接続されているクライアントは、あらかじめ接続されているネットワーク510、520、530を通じてのみ、データ通信を行なうことができる。

【0005】

サーバ500からクライアント511、512、513、クライアント521、522、523、またはクライアント531、532、533へデータを転送するときには、サーバ500のオペレーティング・システムが管理するルーティング・テーブル505の指定にしたがって、ネットワーク・インタフェース501、502、503(ネットワーク510、520、530)のいずれかが一つが使われる。一方、サーバ500からクライアント541へデータを転送するときには、ルーティング・テーブル505の指定にしたがって、ネットワーク510に接続されているゲートウェイ550が選択され、ゲートウェイ550はサーバ500から送られてきたパケットをネットワーク540に送る。この場合には、ネットワーク・インタフェース501のみが使用されることになる。

【0006】

図2に従来のルーティング・テーブルの構成を示す。図2において、161ないし165はルーティング・テーブルのエントリの項目を示す。各々のエントリは、宛先アドレスをハッシュ関数180で変換した値で定まるエントリ100ないし150をヘッダとする線型リストに接続される。宛先アドレスをキーにして線型リストをたどり、宛先アドレス161がキーと一致するエントリを見つける。最初に見つかったエントリのネットワーク・アドレスであるゲートウェイ・アドレス162からパケットを送出する。例えば、図15では、サーバ500からは、ネットワーク・インタフェース501を介し、ゲートウェイ550に接続されているネットワーク510を使ってクライアント541と通信することができる。(尚、サーバ500のネットワーク・インタフェース501、502、503を介して、ネットワーク510、520、530に接続されている各クライアントと通信を行う場合には、ゲートウェイ・アドレス162の代わりにネットワーク・アドレスnet1.11、net2.11、net3.11を直接指定することになる。)

ルーティング・テーブルの詳細は、Douglas E. Comer他著の"Internetworking with TCP/IP, Volume I, II"(Prentice Hall)や、S. J. Leffler他著の"The Design and Implementation of the 4.3BSD UNIX Operating system"(Addison Wesley)に示されている。このように、ルーティング・テーブルの役割は、パケットの宛先アドレスをキーにして、次にこのパケットを送出するゲートウェイなどのネットワーク・アドレスを明らかにする点にある。

10

20

30

40

50

【0007】

さらに、ルーティング・テーブル・エントリの中のネットワーク・インタフェース情報テーブルのエントリへのポインタ163によって、ゲートウェイに到達するネットワーク・インタフェースに関する情報171ないし176を得ることができる。ネットワーク・インタフェースに関する情報として、ネットワーク・インタフェースの最大転送長(MTU: Maximum Transmission Unit)172、このネットワーク・インタフェースを通過した入力・出力パケット数173、174などがある。

【0008】

【発明が解決しようとする課題】

上記従来技術によると、クライアントとサーバのデータ通信では、サーバのネットワーク・インタフェースが複数個存在するにもかかわらず、クライアントが接続されているただ一つのネットワーク・インタフェースしか使用できないため、ネットワーク・インタフェースの負荷状況に応じて動的にネットワーク・インタフェースを選択し、負荷を分散させることができないという問題があった。これは、クライアントの多様なサービス品質(QOS)に応ずることができないという問題につながる。例えば、昨今、音声や画像やデータなどマルチメディア・データを通信するニーズが高まっているが、容量がかなり大きくなるため、負荷が一つのネットワーク・インタフェースに集中してしまい、クライアントが要求するサービス品質(QOS)を満たすことが困難であった。

【0009】

さらに、上記従来技術によると、クライアントのネットワーク・インタフェースのバンド幅の方が、サーバの各ネットワーク・インタフェースのバンド幅よりも大きい場合には、サーバからクライアントへの転送に使用されるネットワーク・インタフェースがただ一つに限られるため、クライアントのネットワーク・インタフェースのバンド幅を十分に活用できないという問題があった。

【0010】

本発明の目的は、上記課題を解決し、複数のネットワーク・インタフェースを備えたサーバにおいて、クライアントが要求するサービス品質(QOS)に応じたネットワーク・インタフェースを用いて通信を行ったり、複数本のネットワーク・インタフェースを並列に用いて通信を行うことができるネットワーク接続システム及び並列ネットワーク接続方法を提供することにある。

【0011】

また、本発明の他の目的は、ATMネットワークのような帯域を予約できるネットワークを用いて、複数のネットワーク・インタフェースを有するサーバに接続し、コネクション設立時にクライアントが要求するサービス品質(QOS)を満たすネットワーク・インタフェースを使用することができるネットワーク接続システム及び並列ネットワーク接続方法を提供することにある。

【0012】

【課題を解決するための手段】

上記目的を達成するために、本発明は、第1の計算機(サーバ)のルーティング・テーブルの各エントリにネットワーク・インタフェースのアドレスであるネットワーク・インタフェース・アドレスを保持する。

【0013】

また、各ネットワーク・インタフェースの使用状況を保持するサービス品質(QOS)保持手段を備える。更に、第2の計算機(クライアント)が第1の計算機と通信路を開設するときに指定するQOSパラメタを指定する手段を備える。具体的には、第2の計算機(クライアント)が、第1の計算機(サーバ)に対して通信路の開設を要求したとき、第1の計算機はルーティング・テーブルのエントリの中から、最終宛先ネットワーク・アドレスが第2の計算機のネットワーク・アドレスと等しい少なくとも1つのエントリを選択し、選択したルーティング・テーブルのエントリのネットワーク・インタフェース・アドレスで示されるネットワーク・インタフェースのQOS保持手段と第2の計算機が指定する

10

20

30

40

50

QOSパラメタを比較し、QOSパラメタで指定される条件を満たすことができるように、少なくとも1つのネットワーク・インタフェースを選択し、そのネットワーク・インタフェースを通して第1の計算機はゲートウェイにパケットを送付する。また、第1の計算機は第2の計算機に対して、通信路の開設時に送信する同期(SYN)パケットに付随させて、選択した全てのネットワーク・インタフェースのネットワーク・アドレスを教える。

【0014】

また、帯域の大きいネットワーク・インタフェースを備えた第2の計算機の場合には、第1の計算機の複数のネットワーク・インタフェースを並列に用いてデータ通信を行えるように構成する。

10

【0015】

クライアントが複数のネットワーク・インタフェースを備えたサーバとデータ通信を行うとき、クライアントが要求するサービス品質(QOS: Quality of Service)を満たすことができるように、各ネットワーク・インタフェースの動的な負荷の大きさなどに応じて、サーバはネットワーク・インタフェースを選択できることが望ましい。

【0016】

本発明では、まず、サーバが同一のクライアントに対して、異なるネットワーク・インタフェースを用いてデータ通信を行えるようにするために、ATM(Asynchronous Transfer Mode)スイッチのようなLANスイッチの複数の帯域予約可能な仮想チャネルを用いることとし、従来のルーティング・テーブル・エントリにネットワーク・インタフェースのアドレスを追加する。クライアントのネットワーク・アドレスとネットワーク・インタフェース・アドレスの両者によって、パケットを送出するゲートウェイのネットワーク・アドレスを求めることによって、異なる経路を選択できるようにする。

20

【0017】

さらに、(a)サーバに、ネットワーク・インタフェースごとに仮想チャネルの帯域や動的な負荷の統計情報を保持するQOS管理テーブルを設け、データ通信路の確立要求時にこのテーブル・エントリを参照し、(b)従来の3ウェイハンドシェイクを拡張して、コネクション確立時において、サーバがクライアントにSYN(同期)/ACK(応答)を返すときに、SYNのパラメタとしてサーバが選択したコネクションのネットワーク・インタフェース・アドレスを添付し、クライアントはSYNで指定されたネットワーク・インタフェース・アドレスに対してACKを返すことによって、サーバが選択したネットワーク・インタフェースを用いてコネクションを確立し、以降のデータ通信を行えるようにする。また、帯域の大きいネットワーク・インタフェースを備えたクライアントの場合には、サーバの複数のネットワーク・インタフェースを並列に用いてデータ通信を行えるようにするために、3ウェイハンドシェイクによるコネクション確立時において、サーバがクライアントにSYN/ACKを返すときに、SYNのパラメタとしてサーバが選択したネットワーク・インタフェース数と各々のネットワーク・アドレスを添付し、クライアントはSYNで指定された全てのネットワーク・アドレスに対してACKを返し、サーバが選択したネットワーク・インタフェースを用いてコネクションを確立し、以降のデータ通信を行えるようにする。さらに、プロトコル層とアプリケーション層の間において、パケットの分割・統合を行うことによって、アプリケーション・プログラムの変更なしに、並列通信を実現することができるようにする。

30

40

【0018】

【発明の実施の形態】

(1)実施形態例1

以下、本発明の詳細な実施形態例を説明する。

【0019】

図3は、本発明の実施形態例1に係るネットワーク接続システムの全体構成図を示す。図

50

3において、3000は本発明が適用されるサーバ計算機、3050は例えばATM (Asynchronous Transfer Mode) スイッチのようなLANスイッチ、3060はPBX (Private Branch Exchange: 構内交換機)、3070は公衆網、3100、3200、3300、3400はLAN (Local Area Network)、3101、3201、3301、3401はクライアント計算機、3102、3202は本発明が適用されるゲートウェイ計算機を示す。

【0020】

サーバ3000は、ワークステーションなどの単一の計算機または並列計算機で構成され、ネットワーク・インタフェース3001、3002、3003、3004を通して、それぞれ通信路3011、3012、3013、3014に接続され、これらの通信路はLANスイッチ3050と接続される。また、ネットワーク・インタフェース3005、3006を通して、通信路3015、3016に接続され、これらの通信路はそれぞれLAN3300、3400と接続される。LANスイッチ3050は、ATMスイッチやファイバーチャネルのようなスイッチで構成され、通信路3074を通してPBX3060と接続され、PBX3060は通信路3073を通して公衆網3070に接続される。ここで、通信路3011、3012、3013、3014、3074、3073は、例えばATMのように、それぞれ、1つないしは複数の帯域予約可能な仮想チャネルから構成され、LANスイッチ3050やPBX3060は、仮想チャネル間のスイッチングを行う。LANスイッチとサーバとの間に複数の仮想チャネルを有していることがポイントとなっている。クライアント3101、3201、3301、3401は、PCやワークステーションなどの計算機で構成され、それぞれ、通信路3110、3210、3310、3410を通してLAN3100、3200、3300、3400と接続され、ゲートウェイ3102は、ネットワーク接続するための計算機であり、通信路3111、3071、3081を通して、それぞれLAN3100、公衆網3070、公衆網3080と接続され、ゲートウェイ3202も同じくネットワーク接続するための計算機であり、通信路3211、3072、3082を通して、それぞれLAN3200、公衆網3070、公衆網3080と接続される。尚、ネットワーク接続手段としては、ゲートウェイの他に、ルータやスイッチなども含まれる。

【0021】

本実施形態例では、LANスイッチ3050と公衆網3070を含むネットワークには、ネットワーク・アドレスとしてnet1.*が割り当てられ、公衆網3080を含むネットワークにはnet2.*、LAN1、LAN2、LAN3、LAN4にはそれぞれnet11.*、net12.*、net13.*、net14.*が割り当てられている。通信路3011ないし3016が接続されるサーバ3000のネットワーク・インタフェース3001ないし3006には、それぞれネットワーク・アドレスnet1.1、net1.2、net1.3、net1.4、net13.1、net14.1が割り当てられ、クライアント3101、3201、3301、3401には、それぞれnet11.10、net12.10、net13.10、net14.10、ゲートウェイ3102の各ネットワーク・インタフェースには、net11.9、net1.11、net2.11、ゲートウェイ3202の各ネットワーク・インタフェースには、net12.9、net1.12、net2.12が割り当てられている。

【0022】

図3において、サーバ3000はクライアント3101とデータ通信を行う場合、複数の経路を選択することができる。サーバ3000は、ネットワーク・インタフェース3001、3002、3003、3004のいずれを使用してもLANスイッチ3050、PBX3060、公衆網3070、ゲートウェイ3102、LAN3100を経て、クライアント3101と通信ができる。また、公衆網3070からクライアント3101までの他の経路として、ゲートウェイ3202、公衆網3080、ゲートウェイ3102、LAN3100を経る経路がある。

【0023】

10

20

30

40

50

いま、クライアント3101がサーバ3000とTCP/IPプロトコルを用いてコネクション指向のデータ通信を行う場合を用いて、本発明の実施形態例1を説明する。まず、図16にプログラム記述の実施形態例を示す。本プログラムは、W. D. Stevens著の"UNIX network programming" (Prentice Hall)にも記されているsocketを用いたプログラムを拡張している。9001から9015がサーバ3000で実行されるプログラムを示し、9050から9061がクライアント3101で実行されるプログラムを示す。サーバ3000は、socketの生成(9002)、socketのアドレス付け(9006)を行った後、listen()コールによって、任意のクライアント(9004)からのコネクション設立要求を待つ(9007)。クライアント3101は、socketを生成(9053)した後、サーバ3000のネットワーク・アドレスのひとつであるnet1.1(9051)を指定して(9055)、connect()コールによってサーバとのコネクション設立を要求する(9058)。

10

【0024】

サーバとのコネクション設立を要求するとき、通信路のQOSを指定する(9057)。QOSは、9052に示すように、サービス・クラス、ピークバンド幅(Mbps)、平均バンド幅(Mbps)の3項目で構成される。サービス・クラスにはBest Effort (BE)、GB (Guaranteed Burst)、GS (Guaranteed Stream)の3種類がある。BEは通常のコネクションで、通信路の込み具合に応じてデータ・スループットが変化する。GBはバースト転送向けのデータ通信で、バースト転送時のスループットをできる限り保証する。バースト転送開始前に帯域の予約を増減させることもできる。GSはビデオや音声などのストリーム・データの通信で、あらかじめ確保されたバンド幅を保証する。サービス・クラス指定をオフにすることもできる。

20

【0025】

クライアント3101がconnect()コールによって通信路の開設を要求し、サーバ3000がクライアントの要求を受け入れると、accept()コールによってクライアントとサーバ間の通信路が確立する(9009)。サーバは新たに確立した通信路で用いるソケットの記述子newfdを割り当て(9009)、子プロセスを生成し(9010)、子プロセスとクライアントとの間でデータ通信が行われる(9013)。親プロセスは、他のクライアントからの要求を受け付けるべく待ち状態に戻る(9015、9008)。クライアント3101も通信路が確立すると(9058)、サーバとの間でデータ通信を行う(9059)。

30

【0026】

次に図4、図1、図9、図10、図11、図12を用いて、TCP/IPプロトコル層におけるクライアント3101とサーバ3000との間のコネクションの確立方法の実施形態例を示す。

【0027】

まず、図4を用いてコネクション確立方法の処理の概略を示す。図4において、3500、3501、3502、3503、3550、3551、3552は、TCPプロトコルの状態を示し、3570、3571、3572はクライアント3101とサーバ3000との間でコネクションを確立する3ウェイ・ハンドシェイクを示す。オリジナルの3ウェイ・ハンドシェイクに関しては、Douglas E. Comer他著の"Internet networking with TCP/IP, Volume I, II" (Prentice Hall)に詳しく記されている。

40

【0028】

CLOSED状態(3500)のサーバ3000は、図16に示したプログラム中のlisten()コールが実行されると(9007)LISTEN状態になり、クライアントからのコネクション開設要求を待つ(3501)。CLOSED状態3550のクライアント3101は、図16に示したプログラム中のconnect()コールが実行される

50

と(9058)、サーバ3000のネットワーク・インタフェースのひとつであるネットワーク・アドレスnet1.1とのデータ通信を制御するために必要な制御ブロックTCB(Transmission Control Block)を作成する(3560)。次に、net1.1に対してコネクション開設を要求するために(3561)、SYN(同期)フラグを立てたTCPパケットをサーバ3000へ送る。このとき、クライアント3101は、コネクションのQOS(サービス品質)をSYNフラグ付TCPパケットのパラメタとしてサーバ3000へ送付する(3570)。QOSは、図16の9052に示したように、サービスクラス、ピークバンド幅(Mbps)、平均バンド幅(Mbps)の3項目で構成される。クライアント3101は、SYNフラグ付TCPパケットを送ると、SYN SENT状態になる(3551)。

10

【0029】

LISTEN状態のサーバ3000は、QOS(サービス品質)付きのTCPパケットを受け取ると、クライアント3101のQOSを満たすネットワーク・インタフェースを探す。ネットワーク・インタフェースを選択する手順の実施形態例は後述する。いま、図3におけるネットワーク・インタフェース3003、ネットワーク・アドレスnet1.3のインタフェースを選択した場合について述べる(3510)。サーバ3000は、ネットワーク・アドレスnet1.3に対応するTCB(Transmission Control Block)を作成する(3511)。TCBには、クライアント3101のネットワーク・アドレス、サーバ3000の選択されたネットワーク・インタフェースのネットワーク・アドレスnet1.3、通信プロトコル名TCPが保持されている。サーバ3000は、TCBを作成した後、クライアント3101に対して、クライアントからのSYNに対するACK(応答)フラグの付いたTCPパケットと、SYNフラグの付いたTCPパケットをクライアント3101に送り(3571)、このTCBをSYN RECD状態にする(3502)。SYNフラグには、パラメタとしてネットワーク・アドレスnet1.3を添付する。クライアント3101は、パラメタとしてネットワーク・アドレスnet1.3付随したSYNフラグ付きTCPパケットを受け取ると、net1.1用に作成したTCBを削除し(3562)、net1.3に対応したTCBを作成する(3563)。次に、クライアント3101はnet1.3に対して、ACKフラグ付TCPパケットを送り(3572)、ESTABLISHED状態となる(3552)。サーバ3000もACKを受け取るとESTABLISHED状態となり(3503)、以降、クライアントとサーバ間でSEND/RECEIVEによる通信を行うことができる(3573)。

20

30

【0030】

以上図4を用いて示した本発明のコネクション確立方法の詳細を、図1、図9、図10、図11、図12を用いて説明する。図9、図10、図11は、LISTEN状態(3501)でクライアントからのコネクション設立要求を待っている状態以降のサーバ3000の処理のフローチャート図、図12はクライアント3101の処理のフローチャート図、図1は、ネットワーク・インタフェースを選択するとき用いるデータ構造を示す。

【0031】

図9において、LISTEN状態(3501)のサーバ3000は、クライアント3101からQOS指定付きのSYNフラグの付いたTCPパケットを受信すると(3570)、クライアント3101のQOS要求を満たし、サーバ3000のネットワーク・インタフェースの負荷を分散できるネットワーク・インタフェースを選択するため(4010)、以下の処理を行う。

40

【0032】

図1を一緒に用いて、ネットワーク・インタフェースを選択する手順について説明する。図1は、本発明の実施形態例1にかかるルーティング・テーブル、ネットワーク・インタフェース情報テーブル、QOS管理テーブルを示す。サーバ3000はネットワーク・インタフェースを選択するとき、これらのテーブルを参照する。

【0033】

50

図1において、10ないし60はルーティング・テーブルの各エントリの並びのヘッダである。クライアント3101のネットワーク・アドレス(サーバ3000から見た場合には宛先アドレス)net11.10をハッシュ関数で変換し、ルーティング・テーブル・エントリをたどる。ルーティング・テーブルの各エントリは、最終的にパケットを送り届ける先であるクライアントの宛先アドレス70、宛先アドレスへ到達できるサーバ3000のネットワーク・インタフェースのネットワーク・アドレス71、宛先アドレス70へ到達するために次ぎにパケットを送出するゲートウェイのネットワーク・アドレス72、ゲートウェイが接続されているネットワーク・インタフェースの情報とネットワーク・インタフェースのQOS情報を保持するネットワーク・インタフェース情報テーブルへのポインタ73、宛先が等しいルーティング・テーブル・エントリを高速にサーチするためのポインタ74、ルーティング・テーブル・エントリのリストを作成するためのポインタ75、その他76から構成される。ネットワーク・インタフェース情報テーブルの各エントリは、ネットワーク・インタフェースの名前80、当該ネットワーク・インタフェースが処理できる最大パケット長(MTU: Maximum Transmission Unit)81、当該ネットワーク・インタフェースを通して送受信されたパケット数の累積値を保持する受信パケット数82と送信パケット数83、ネットワーク・インタフェース情報テーブルのエントリの線型リストを作成するためのポインタ84、当該ネットワーク・インタフェースのQOS情報や動的負荷情報を保持するQOS管理テーブル・エントリへのポインタ85、その他86から構成される。QOS管理テーブルの各エントリは、当該ネットワーク・インタフェースの最大バンド幅(メガビット/秒)90、バーチャル・チャンネル(VC)の本数91、各バーチャル・チャンネルに割り当てられているバンド幅(帯域幅)(メガビット/秒)92、各バーチャル・チャンネルのコネクションが設立されているか否かを示すフラグ93、サービスクラスがGBかGSで、コネクションが設立されているバーチャル・チャンネル本数94、サービスクラスがGBかGSでコネクションが設立され、予約済になっているバンド幅の合計値(メガビット/秒)95、最新1分間のピーク転送量(メガビット/秒)96、最新1分間の平均転送量(メガビット/秒)97などからなる。

【0034】

図13を用いて、図3のネットワーク接続構成における、サーバ3000のルーティング・テーブル全体の実施形態例を示す。図13において、3800がルーティング・テーブルの全体、3801ないし3814の各々がルーティング・テーブル・エントリの中の宛先アドレス(70)、ネットワーク・インタフェースのネットワーク・アドレス(71)、ゲートウェイのネットワーク・アドレス(72)を示す。

【0035】

クライアント3101を含むLAN3100のネットワーク・アドレスは、宛先アドレスnet11.* (3801ないし3808)で示し、LAN3200、3300、3400のそれぞれのネットワーク・アドレスは、宛先アドレスnet12.* (3809ないし3812)、net13.* (3813)、net14.* (3814)で示している。エントリ3801は、net11.*へのルーティングは、サーバ3000のネットワーク・インタフェース3001(ネットワーク・アドレスはnet1.1)を用いて、ゲートウェイ3102(ネットワーク・アドレスnet1.11)経由で可能であることを示している。同様に、エントリ3802、3803、3804は、それぞれネットワーク・インタフェース3002、3003、3004(ネットワーク・アドレスはnet1.2, net1.3, net1.4)を用いて、ゲートウェイ3102経由でnet11.*へのルーティングが可能であることを示し、エントリ3805、3806、3807、3808は、それぞれネットワーク・インタフェース3001、3002、3003、3004(ネットワーク・アドレスはnet1.1, net1.2, net1.3, net1.4)を用いて、ゲートウェイ3202(ネットワーク・アドレスnet1.12)経由でnet11.*へのルーティングが可能であることを示す。以上のように、ルーティング・テーブルにあらかじめnet11.*へ到達可能なネットワーク・イン

10

20

30

40

50

タフェースとゲートウェイ・アドレスを設定しておき、コネクション設立時に、クライアント3101が要求するQOSやサーバ3000のネットワーク・インタフェースの負荷やLANスイッチ3050の負荷に応じてデータ通信路（経路）を選択する。同様に、エントリ3809、3810、3811、3812は、それぞれネットワーク・インタフェース3001、3002、3003、3004（ネットワーク・アドレスはnet1.1、net1.2、net1.3、net1.4）を用いて、ゲートウェイ3202（ネットワーク・アドレスnet1.12）経由でnet12.*へのルーティングが可能であることを示す。

【0036】

図3によると、ゲートウェイ3102経由でもnet12.*へのルーティングは可能であるが、図13のルーティング・テーブルの設定によると、ゲートウェイ3102経由のルーティングは行わせていない。エントリ3813、3814のネットワーク・インタフェース・アドレスが未指定なのは、net13.*とnet14.*へのルーティングは、サーバ3000のネットワーク・インタフェースの数が1つずつしかないため、指定する必要がないためである。なお、従来のルーティング・テーブルでは、ネットワーク・インタフェースのネットワーク・アドレスのカラムが、ルーティング・テーブルには無かった。

【0037】

コネクションが確立されると、テーブル3820に、データ通信を行うクライアント（宛先アドレス）とサーバのネットワークインタフェースのネットワーク・アドレス（送信元アドレス）を登録する。図13のテーブル3820の場合には、クライアント3101とサーバ3000のネットワーク・インタフェース3002、3003、3004の間でコネクションが確立されている（3821、3822、3833）。

【0038】

図9の説明に戻る。サーバ3000は、クライアント3101のネットワーク・アドレスnet11.10をハッシュ関数にかけ、ルーティング・テーブル・エントリのリストのヘッダ40を求める（4020）。ヘッダ40からルーティング・テーブル・エントリをたどり、そのエントリの宛先アドレス70が、クライアント3101のネットワーク・アドレスnet11.10もしくはクライアント3101のサブネットワーク・アドレスnet11と等しいか否かを調べる（4030）。もし、アドレスが不一致の場合には、ルーティング・テーブル・エントリの次エントリへのポインタ75をたどり、次のエントリへ進み（4035）、クライアント3101のネットワーク・アドレスnet11.10もしくはクライアント3101のサブネットワーク・アドレスnet11と等しい宛先アドレス70のルーティング・テーブル・エントリが見つかるまで繰り返す（4030、4035）。

【0039】

アドレスが等しいルーティング・テーブル・エントリが見つかった場合について説明する。

【0040】

(A) クライアント3101が要求するサービス・クラスがGB (Guaranteed Burst) かGS (Guaranteed Stream) の場合には、ルーティング・テーブル・エントリ中のポインタ73から、ネットワーク・インタフェース情報テーブル・エントリを調べ（4040）、その中のQOS管理テーブル・エントリへのポインタ85の値をワークメモリ領域に記録する（4050）。ルーティング・テーブル・エントリ中のポインタ74を用いて、同じ宛先アドレスの全てのルーティング・テーブル・エントリを求め（4060）、そのネットワーク・インタフェース情報テーブル・エントリの中のQOS管理テーブル・エントリへのポインタ85の値をワークメモリ領域に記録する（4050）。本実施形態例の場合には、図13のルーティング・テーブル3800によると、エントリ3801ないし3808の8つのエントリが選択される。

【0041】

10

20

30

40

50

図10に移る。次に、4050で記録したQOS管理テーブル・エントリへのポインタをたどり、QOS管理テーブル・エントリ中の記録を調べ、次の2条件を満たすサーバ3000のネットワーク・インタフェースを選択する(4070)。

【0042】

1) クライアント3101が要求する平均バンド幅が、最大バンド幅(90)から予約済みバンド幅の合計(95)を引いた値よりも小さいこと、

2) クライアント3101が要求するピーク・バンド幅が、未予約のバーチャル・チャンネル(93)のバンド幅の中で、最大のものよりも小さいこと、の2条件である。

【0043】

もし該当するネットワーク・インタフェースがない場合には、最近一分間のピーク転送量(96)が最小のネットワーク・インタフェースを選択する(4080)。図10や図4には示していないが、クライアント3101の要求を満たすことができない場合には、サーバ3100はSYN SENT状態のクライアント3101に対して、クライアント3101の要求を満たすことができない旨を、ACKフラグ付TCPパケットのパラメタとして伝え、クライアント3101とQOSの協議を行い、クライアント3101がQOSを指定し直して、再びSYNフラグ付TCPパケットをサーバ3000に送り(3570)、新しいQOSに基づいてサーバ3000がネットワーク・インタフェースを探す(3510)、というサーバ3000とクライアント3101との間の協議が行われてもよい。

10

【0044】

再び図10へ戻って、4070において、もし該当するネットワーク・インタフェースがある場合には、条件を満たすネットワーク・インタフェースの中から、予約済みバンド幅が最小のインタフェースを選択する(4085)。選択したネットワーク・インタフェースのQOS管理テーブル・エントリの予約済みバンド幅(95)に、クライアント3101が要求している平均バンド幅を加える。

20

【0045】

選択したネットワーク・インタフェースのアドレスを含むルーティング・テーブル・エントリを以下では用いる。

【0046】

(B) 4030に戻り、クライアント3101が要求するサービス・クラスがBE(Best Effort)が未指定の場合には、最初に見つかったルーティング・テーブル・エントリを用いる。

30

【0047】

選択したルーティング・テーブル・エントリから、ネットワーク・インタフェースのネットワーク・アドレス(71)を求める(4200)。TCPプロトコルを用いてEnd-to-Endのデータ通信を行うのに必要なTCB(Transmission Control Block)を作成する(4210)。TCBに記す送信元アドレスには4200で求めたネットワーク・アドレスを指定し、宛先アドレスにはクライアント3101のネットワーク・アドレスを指定する。TCBを作成すると、クライアント3101にACKフラグ付TCPパケットと、ネットワーク・インタフェースのネットワーク・アドレスをパラメタとして付加したSYNフラグ付TCPパケットを送り(4220)、SYN RECD状態になる(3502)。

40

【0048】

図11に移る。SYN RECD状態において、クライアント3101からACKフラグ付TCPパケットを受け取ると(4250)、接続確立済み通信路の管理テーブル(図13におけるテーブル3820)に、当該ネットワーク・インタフェース・アドレスとクライアント3101のネットワーク・アドレスを登録し(4260)、ESTABLISHED状態になる(3503)。以降、クライアントとサーバ間で、QOSに従ったデータ通信を行う。

【0049】

50

図12を用いて、クライアント3101の詳細な処理手順を示す。CLOSED状態(3550)のクライアントでは、アプリケーション・プログラムが、サーバ3000の複数のネットワーク・インタフェースの中の一つのネットワーク・アドレスnet1.A(9051)とQOS(9057)を指定して、connect()コールを実行する(9058)。クライアント上のOSは、connect()コールを契機にして通信路開設要求を受ける(4300)。OSはまず、net1.Aとの通信のために必要なTCB(Transmission Control Block)を作成し、connect()コールで指定されたQOS値をパラメタとする、SYNフラグ付TCPパケットを、サーバ3000のネットワーク・インタフェースの中の一つであるnet1.A宛てに送る(4310)。SYNフラグ付TCPパケットを送ると、クライアント3101はSYN SENT状態になり、サーバ3000から送られてくるACKフラグ付のTCPパケットを待つ(3551)。ACKおよび、SYNフラグ付TCPパケットを受け取ると、SYNのパラメタである、サーバ3000のネットワーク・インタフェースのアドレスのひとつであるネットワーク・アドレスnet1.Bがnet1.Aと等しいかどうかを調べる(4320)。もし、等しい場合には、net1.A宛てでサーバ3000へACKフラグ付TCPパケットを送る(4345)。もし、不一致の場合は、サーバ3000がネットワーク・インタフェースを変更したものとみなし、4310で作成したnet1.Aとの通信のためのTCBを削除し、新たにnet1.Bとの通信のためのTCBを作成する(4330)。TCBを作成後、net1.B宛てでサーバ3000へACKフラグ付TCPパケットを送る(4340)。ACKを返すと、次に、クライアントの中のコネクション確立済み通信路管理テーブルに、サーバ3000のネットワーク・インタフェースのネットワーク・アドレスnet1.Bと、クライアント3101のネットワーク・アドレスnet11.10を登録し(4350)、ESTABLISHED状態になる(3552)。以降、クライアントとサーバ間で、QOSに従ったデータ通信を行う。

【0050】

次に、図7、図8を用いて、コネクション確立方法におけるLANスイッチ3050の動作を説明する。

【0051】

図7において、3011、3012、3013、3014、3074は、図3と同じく通信路を表す。3011a、3011bは通信路3011に設定されるバーチャル・チャンネルを示す。図7では、図面の都合上、2本のバーチャル・チャンネルを記しているだけであるが、通常はもっと多くの本数のバーチャル・チャンネルが存在する。また、バーチャル・チャンネルはあらかじめ設定される場合もあるが、動的に生成されてもよい。3012a、3012bは通信路3012に設定されるバーチャル・チャンネル、3013a、3013bは通信路3013に設定されるバーチャル・チャンネル、3014a、3014bは通信路3014に設定されるバーチャル・チャンネル、3074a、3074b、3074c、3074d、3074e、3074fは通信路3074に設定されるバーチャル・チャンネルである。通信路3074において、バーチャル・チャンネル3074a(VC=11)、3074b(VC=12)、3074c(VC=13)、3074d(VC=14)は、通信路3071のバーチャル・チャンネルに接続され、バーチャル・チャンネル3074e(VC=21)と3074f(VC=22)は、通信路3072のバーチャル・チャンネルに接続される。

【0052】

図7において、5000は、サーバ3000の各ネットワーク・インタフェースのネットワーク・アドレスとLANスイッチ3050のポート番号の変換テーブル、5010(a)はLANスイッチ3050の各ポートとバーチャル・チャンネルの組が、相互にどのように接続されるかを示すスイッチング・テーブルである。変換テーブル5000によって、ネットワーク・アドレスをLANスイッチ3050の物理アドレスに変換し、パケットをネットワーク・アドレスで指定されるポートへ送出することができる。スイッチング・テーブル5010は、コネクションが設立されるときに設定されるテーブルである。図7に

において、5100、5200、5300、5400、5500は、LANスイッチ3050のポート#0、#1、#2、#3、#4のQOS管理テーブルのエントリを示す。各エントリのフォーマットは、図1の90ないし97と同じである。サーバ3000が、LANスイッチ3050のポート#0のQOSも管理することによって、サーバ3000がクライアントとしてコネクションを設立するときや、サーバ3000がゲートウェイとしてルーティングを行う場合、図3における公衆網3070に接続される複数のゲートウェイの中で、どのゲートウェイへルーティングすると、クライアントのQOSを満たすことができるかを判断することができる。つまり、図9、図10、図11を用いて述べた実施形態例では、ネットワーク・インタフェース3001、3002、3003、3004のQOS管理テーブル・エントリ5110、5120、5130、5140から、クライアント3101とのコネクション設立時の経路を決定したが、さらにLANスイッチ3050のポート#0のQOS管理テーブル・エントリ5100も一緒に用いて、LANスイッチ3050のポート#0のバーチャル・チャンネルと、ポート#1、2、3のバーチャル・チャンネルのスイッチングの設定も考慮して、サーバ3000とLANスイッチ3050との経路を選択することができる。その判断のための処理手順は、図9、図10、図11と同様である。

【0053】

図8を用いて、コネクション確立方法におけるLANスイッチ3050の動作を説明する。クライアント3101からサーバ3000のnet1.1宛のSYNフラグ付TCPパケットが、LANスイッチ3050のポート#0、VC=11(3074a)へルーティングされて届いた場合(5500)について説明する。LANスイッチ3050は、変換テーブル5000を調べ、net1.1宛のSYNフラグ付TCPパケットの送出先ポート#が1であることを求める(5510)。次に、ポート#1のバーチャル・チャンネルの中からVC=1(3011a)を選択し、(ポート#0、VC=11)と(ポート#1、VC=1)が相互にスイッチングされるように、アドレス・スイッチング・テーブル5010(a)に登録する(5520)。今、サーバ3000がネットワーク・インタフェースとして通信路3013(ネットワーク・アドレスnet1.3)を選択した場合、サーバ3000は、LANスイッチ3050に対して、アドレス・スイッチング・テーブル5010に登録されている(ポート#0、VC=11)と(ポート#1、VC=1)の対応を記したエントリを削除させ、さらに(ポート#0、VC=11)と(ポート#3、VC=1)の対応を記したエントリを追加させる(5530)。LANスイッチ3050は、アドレス・スイッチング・テーブル5010(a)の内容を変更する(5540)。変更の結果、テーブル5010(a)は、テーブル5010(b)に示すとおり更新される。この結果、サーバ3000は通信路3013を使って、クライアント3101へ、ACKフラグ付TCPパケットと、ネットワーク・アドレスnet1.3がパラメタとして付随したSYNフラグ付TCPパケットを、送ることができる(5550)。つまり、サーバ3000が(ポート#3、VC=1)へ送り出したパケットは、クライアント3101がSYNフラグ付TCPパケットをサーバ3000に送るときに使用した(ポート#0、VC=11)へスイッチングされるため、クライアント3101へ届くことができるだけでなく、クライアント3010からLANスイッチ3050までに確立された経路をそのまま使用することができる。クライアント3101は、サーバ3000に対して、(ポート#0、VC=11)からネットワーク・アドレスnet1.3に対応する(ポート#3、VC=1)へ、ACKフラグ付TCPパケットを送出し(5560)、コネクションが確立される(5570)。

【0054】

以上のように、上記の本発明の実施形態例1によると、クライアントが要求するQOSとサーバの負荷に応じて、サーバは複数のネットワーク・インタフェースの中から、条件に見合ったネットワーク・インタフェースを選択し、データ通信を行うことができる。

【0055】

また、上記の実施形態例1によると、クライアント3101は、コネクション確立時には

10

20

30

40

50

サーバ3000の全てのネットワーク・インタフェースのネットワーク・アドレスを知らなくとも、クライアントが要求するQOSに応じたネットワーク・インタフェースを用いてコネクションを確立でき、データ通信を行うことができる。上記の実施形態例1によると、サーバ3000からの指示にしたがって、LANスイッチ3050の中のスイッチング・テーブルを更新できるため、サーバ3000の公衆網3070に接続されるネットワーク・アドレスは、ネットワーク・インタフェース3001のアドレスをただ一つ割り当てておくだけでも、異なる経路を用いて、クライアントとデータ通信を行うことができる。

【0056】

また、上記の実施形態例1ではネットワーク・インタフェースを選択する手順のひとつを示したに過ぎないが、QOS管理テーブルの情報の使い方によって、種々の選択手順を考えることができる。

【0057】

(2) 実施形態例2

次に、本発明の実施形態例2である、クライアント3101とサーバ3000との間の並列コネクションの確立方法について、図5、図1、図18、図19、図20、図21を用いて説明する。

【0058】

まず、図5を用いて、おおまかな処理の流れを示す。

【0059】

図5において、3500、3501、3502、3503、3550、3551、3552は、TCPプロトコルの状態を示し、3590、3591、3592、3593、3594はクライアント3101とサーバ3000との間でコネクションを確立する3ウェイ・ハンドシェイクを示す。

【0060】

CLOSED状態3500のサーバ3000は、図11に示したプログラム中のlisten()コールが実行されると(9007)LISTEN状態になり、クライアントからのコネクション開設要求を待つ(3501)。CLOSED状態3550のクライアント3101は、図16に示したプログラム中のconnect()コールが実行されると(9058)、サーバ3000のネットワーク・インタフェースのひとつであるネットワーク・アドレスnet1.1とのデータ通信を制御するために必要な制御ブロックTCB(Transmission Control Block)を作成する(3580)。次に、net1.1に対してコネクション開設を要求するために(3581)、SYNフラグを立てたTCPパケットをサーバ3000へ送る。このとき、クライアント3101は、コネクションのQOS(サービス品質)をSYNフラグ付TCPパケットのパラメタとしてサーバ3000へ送付する(3590)。QOSは、図16の9052に示したように、サービスクラス、ピークバンド幅(Mbps)、平均バンド幅(Mbps)の3項目で構成される。クライアント3101は、SYNフラグ付TCPパケットを送ると、SYN SENT状態になる(3551)。

【0061】

LISTEN状態のサーバ3000は、QOS(サービス品質)付きのTCPパケットを受け取ると、クライアント3101のQOSを満たすネットワーク・インタフェースを探す。ひとつのネットワーク・インタフェースだけではクライアント3101が要求するQOSを満たすことができない場合には、複数のネットワーク・インタフェースを選択し、それらの合計値でクライアント3101が要求するQOSを満たすようにする。複数のネットワーク・インタフェースを選択する手順の実施形態例は後述する。いま、図3におけるネットワーク・インタフェース3002、3003、3004の3個(ネットワーク・アドレスnet1.2, net1.3, net1.4)を選択した場合について述べる(3520)。サーバ3000は、ネットワーク・アドレスnet1.2, net1.3, net1.4に対応する3個のTCB(Transmission Contr

10

20

30

40

50

o l B l o c k) を作成する (3 5 2 1) 。サーバ 3 0 0 0 は、3 個の T C B を作成した後、クライアント 3 1 0 1 に対して、クライアントからの S Y N に対する A C K フラグの付いた T C P パケットと、S Y N フラグの付いた T C P パケットをクライアント 3 1 0 1 に送り (3 5 9 1) 、これらの T C B を S Y N R E C V D 状態にする (3 5 0 2) 。S Y N フラグには、パラメタとしてネットワーク・アドレスの個数 3 、ネットワーク・アドレス *net 1 . 2* , *net 1 . 3* , *net 1 . 4* 、および各経路に割り当てられた Q O S (平均バンド幅) を添付する。クライアント 3 1 0 1 は、パラメタが付随した S Y N フラグ付き T C P パケットを受け取ると、*net 1 . 1* 用に作成した T C B を削除し (3 5 8 2) 、*net 1 . 2* , *net 1 . 3* , *net 1 . 4* に対応した 3 個の T C B を作成する (3 5 8 3) 。次に、クライアント 3 1 0 1 は *net 1 . 2* , *net 1 . 3* , *net 1 . 4* に対して、それぞれ A C K フラグ付 T C P パケットを送り (3 5 9 2 、 3 5 9 3 、 3 5 9 4) 、 E S T A B L I S H E D 状態となる (3 5 5 2) 。サーバ 3 0 0 0 も 3 個の A C K を受け取ると E S T A B L I S H E D 状態となり (3 5 0 3) 、3 本のコネクションが確立される。以降、クライアントとサーバ間では、3 本のコネクションを用いて、S E N D / R E C E I V E による並列通信を行うことができる (3 5 9 5) 。

【 0 0 6 2 】

並列通信のための分割・統合の方法の実施形態例を図 1 7 を用いて説明する。図 1 7 では、サーバ 3 0 0 0 からクライアント 3 1 0 1 へデータを送る場合について示す。6 0 3 0 、 6 0 3 1 、 6 0 3 2 、 6 0 3 3 はサーバ 3 0 0 0 内のバッファ、6 1 3 0 、 6 1 3 1 、 6 1 3 2 、 6 1 3 3 はクライアント 3 1 0 1 内のバッファ、6 0 1 2 、 6 0 1 3 、 6 0 1 4 はサーバ 3 0 0 0 からネットワーク 6 0 7 0 に接続される通信路 (図 3 における通信路 3 0 1 2 、 3 0 1 3 、 3 0 1 4 に相当する) 、6 1 1 0 はクライアント 3 1 0 1 からネットワーク 6 0 7 0 に接続される通信路 (図 3 における通信路 3 1 1 0 に相当する) を示す。今、アプリケーションからデータを S E N D すると、そのデータは T C P プロトコル層のバッファ 6 0 3 0 に一時的に保持される。6 0 3 0 内のデータは、アプリケーションが指定するセグメント長ごとに 3 つに分割し、バッファ 6 0 3 1 、 6 0 3 2 、 6 0 3 3 に、3 つの通信路に割り当てられたバンド幅に比例した個数ずつ分配する。例えば、バンド幅の比率が 2 : 1 : 1 の場合には、6 0 3 1 、 6 0 3 1 、 6 0 3 2 、 6 0 3 3 の順序で振り分け、各バッファは一時的にデータを保持する。T C P 層では、サーバ 3 0 0 0 のバッファ 6 0 3 1 、 6 0 3 2 、 6 0 3 3 と、クライアント 3 1 0 1 のバッファ 6 1 3 1 、 6 1 3 2 、 6 1 3 3 を 1 対 1 で対応させ、それぞれが独立したコネクションとして、互いに独立にサーバ 3 0 0 0 からクライアント 3 1 0 1 へデータを送る (S E N D する) 。つまり、バッファ 6 0 3 1 、 6 0 3 2 、 6 0 3 3 のデータは、それぞれ通信路 6 0 1 2 、 6 0 1 3 、 6 0 1 4 から通信路 6 1 1 0 へ送られる。クライアント 3 1 0 1 内のバッファ 6 1 3 1 、 6 1 3 2 、 6 1 3 3 へ送られてきたデータは、ひとつのデータに統合し、バッファ 6 1 3 0 へ送り、クライアント 3 1 0 1 上のアプリケーションがデータを受け取る。以上のように、T C P 層とアプリケーション層との間で、データの分割と統合を行うため、アプリケーション・プログラムには並列通信を隠蔽できる。そのため、シングル通信と同じプログラムを用いて並列通信を実現し、クライアントが要求する Q O S を満たすことができる。

【 0 0 6 3 】

また、分割と統合のアルゴリズムは、次のようであってもよい。バッファ 6 0 3 0 内のデータをセグメント長ごとに分割するとき、各セグメントにシリアル番号を付加する。3 つの通信路の負荷状態は、動的に変化するため、バッファ 6 0 3 0 からバッファ 6 0 3 1 、 6 0 3 2 、 6 0 3 3 へセグメントを分配するときには、バッファ 6 0 3 1 、 6 0 3 2 、 6 0 3 3 の中で未送付のセグメント数が最少のバッファを選択し、未送付のセグメント数が等しければ、先頭のセグメントのシリアル番号が一番大きなバッファを選択することによって、最も S E N D 処理が進行している通信路に付加を加えることができるため、動的に負荷を分散させ、より高速な通信が可能になる。

【 0 0 6 4 】

以上図5を用いて示した本発明の実施形態例2の詳細を、図1、図18、図19、図20、図21を用いて説明する。図18、図19、図20は、LISTEN状態(3501)でクライアントからの接続設立要求を待っている状態以降のサーバ3000の処理のフローチャート図、図21はクライアント3101の処理のフローチャート図を示す。なお、図18は、図9と同じフローチャート図のため、説明を省略する。

【0065】

図19において、図18の4050で記録したQOS管理テーブル・エントリへのポイントをたどり、QOS管理テーブル・エントリ中の記録を調べ、次の2条件を満たすサーバ3000のネットワーク・インタフェースを選択する。

【0066】

1) クライアント3101が要求する平均バンド幅が、最大バンド幅(90)から予約済みバンド幅の合計(95)を引いた値よりも小さいこと、
2) クライアント3101が要求するピーク・バンド幅が、未予約のバーチャル・チャンネル(93)のバンド幅の中で、最大のものよりも小さいこと(4070)、の2条件である。

【0067】

もし該当するネットワーク・インタフェースがない場合には、複数のネットワーク・インタフェースを選択し、各インタフェースの合計値がクライアント3101が要求するQOSを満たすようにする。つまり、4050で記録したQOS管理テーブル・エントリへのポイントをたどり、QOS管理テーブル・エントリ中の記録を調べ、次の2条件を満たすサーバ3000のネットワーク・インタフェースを選択する。

【0068】

1) クライアント3101が要求する平均バンド幅が、選択した複数個のネットワーク・インタフェースの最大バンド幅(90)の合計から、選択した複数個のネットワーク・インタフェースの予約済みバンド幅の合計(95)の合計を引いた値よりも小さいこと、
2) 選択した複数個のネットワーク・インタフェースの未予約のバーチャル・チャンネル(93)の中でバンド幅が最大のバーチャル・チャンネルを1本ずつ選び、それらの合計よりもクライアント3101が要求するピーク・バンド幅が小さいこと(4100)、の2条件である。

【0069】

複数個のネットワーク・インタフェースを用いても、クライアント3101の要求を満たすことができない場合には、複数個のネットワーク・インタフェースは用いずに、最近一分間のピーク転送量(96)が最小のネットワーク・インタフェースを選択する(4110)。図19や図5には示していないが、クライアント3101の要求を満たすことができない場合には、サーバ3100はSYNSENT状態のクライアント3101に対して、クライアント3101の要求を満たすことができない旨を、ACKフラグ付TCPパケットのパラメタとして伝え、クライアント3101とQOSの協議を行い、クライアント3101がQOSを指定し直して、再びSYNフラグ付TCPパケットをサーバ3000に送り(3590)、新しいQOSに基づいてサーバ3000がネットワーク・インタフェースを探す(3520)、というサーバ3000とクライアント3101との間の協議が行われてもよい。

【0070】

再び図19へ戻って、4070において、もし該当するネットワーク・インタフェースがある場合には、条件を満たすネットワーク・インタフェースの中から、予約済みバンド幅が最小のインタフェースを選択する(4085)。以上、4085、4100、4110のいずれかで選択した全てのネットワーク・インタフェースのQOS管理テーブル・エントリの予約済みバンド幅(95)に、各ネットワーク・インタフェースに分割して割り当てたクライアント3101が要求している平均バンド幅の各ネットワーク・インタフェース担当分の値を加える(4120)。選択したネットワーク・インタフェースのアドレスを含むルーティング・テーブル・エントリを以下では用いる。

10

20

30

40

50

【 0 0 7 1 】

図 2 0 に移る。選択したルーティング・テーブル・エントリから、全てのネットワーク・インタフェースのネットワーク・アドレス (7 1) を求める (4 5 0 0)。次に、TCP プロトコルを用いて End - to - End のデータ通信を行うのに必要な T C B (T r a n s m i s s i o n c o n t r o l B l o c k) を全ての選択したネットワーク・インタフェースに対応して作成する (4 5 1 0)。各 T C B に記す送信元アドレスには 4 5 0 0 で求めたネットワーク・アドレスを 1 つずつ指定し、宛先アドレスにはクライアント 3 1 0 1 のネットワーク・アドレス n e t 1 1 . 1 0 を指定する。T C B を作成すると、クライアント 3 1 0 1 に A C K フラグ付 T C P パケットと、選択したネットワーク・インタフェースの個数、全てのネットワーク・アドレス、各ネットワーク・インタフェースに割り当てた平均バンド幅をパラメタとして付加した S Y N フラグ付 T C P パケットを送り (4 5 2 0)、S Y N R E C V D 状態になる (3 5 0 2)。サーバ 3 0 0 0 は、A C K フラグ付 T C P パケットと S Y N フラグ付 T C P パケットを、サーバ 3 0 0 0 がクライアント 3 1 0 1 から S Y N フラグ付 T C P パケットを受け取ったのと同じネットワーク・インタフェースを用いて、クライアント 3 1 0 1 に送付する。S Y N R E C V D 状態において、クライアント 3 1 0 1 から、4 5 2 0 の S Y N フラグ付 T C P パケットに添付した全てのネットワーク・アドレスに対応するネットワーク・インタフェースにおいて、A C K フラグ付 T C P パケットを受け取ると (4 5 3 0)、コネクション確立済み通信路の管理テーブルに、当該全てのネットワーク・インタフェース・アドレスとクライアント 3 1 0 1 のネットワーク・アドレスを登録し (4 5 4 0)、E S T A B L I S H E D 状態になる (3 5 0 3)。以降、クライアントとサーバ間で、Q O S に従った並列データ通信を行う。

10

20

【 0 0 7 2 】

図 2 1 を用いて、クライアント 3 1 0 1 の詳細な処理手順を示す。S Y N S E N T 状態 3 5 5 1 になるまでは、図 1 2 と同じフローチャート図のため省略する。

【 0 0 7 3 】

S Y N S E N T 状態のクライアント 3 1 0 1 は、サーバ 3 0 0 0 から送られてくる A C K フラグ付の T C P パケットを待つ (3 5 5 1)。A C K および、S Y N フラグ付 T C P パケットを受け取ると、S Y N のパラメタである、選択されたネットワーク・インタフェースの個数を調べる (4 6 0 0)。0 個の場合には n e t 1 . A 宛てでサーバ 3 0 0 0 へ A C K フラグ付 T C P パケットを送る (4 6 4 0)。1 個の場合には、パラメタで指定されているネットワーク・アドレスが、n e t 1 . A と等しいかどうかを調べる (4 6 1 0)。等しい場合には、n e t 1 . A 宛てでサーバ 3 0 0 0 へ A C K フラグ付 T C P パケットを送る (4 6 4 0)。不一致の場合、もしくは、4 6 0 0 において S Y N のパラメタの値が 2 個以上の場合には、サーバ 3 0 0 0 がネットワーク・インタフェースを変更したものとみなし、4 3 1 0 で作成した n e t 1 . A との通信のための T C B を削除し、新たに S Y N のパラメタで指定されている全てのネットワーク・インタフェースと通信を行えるように、各々のネットワーク・アドレスに対応する T C B を作成する (4 6 2 0)。A T M のような帯域予約可能なネットワークの場合には、クライアント 3 1 0 1 はサーバから知らされた 3 個の Q O S (平均バンド幅) にしたがって、3 本のバーチャル・チャネルを選択して、3 本のコネクションの確立を図る。次に、S Y N のパラメタで指定されていたネットワーク・アドレスの各々宛てに、A C K フラグ付 T C P パケットを送る (4 6 3 0)。A T M のようなネットワークの場合には、3 本のバーチャル・チャネルのそれぞれから A C K フラグ付 T C P パケットを送る。A C K を返すと、次に、クライアントの中のコネクション確立済み通信路管理テーブルに、S Y N のパラメタで指定されていたサーバ 3 0 0 0 の全てのネットワーク・アドレスと、クライアント 3 1 0 1 のネットワーク・アドレス n e t 1 1 . 1 0 を登録し (4 6 5 0)、E S T A B L I S H E D 状態になる (3 5 5 2)。以降、クライアントとサーバ間で、Q O S に従った並列データ通信を行う。

30

40

【 0 0 7 4 】

この時のサーバ 3 1 0 0 とクライアント 3 1 0 1 のコネクション確立済み通信路管理テ

50

ブルを、それぞれ図 1 3 と図 1 4 に示す。図 1 3 において、3 8 2 0 がコネクション確立済み通信路管理テーブルを示し、3 8 2 1、3 8 2 2、3 8 2 3 の各エントリによって、3 本のコネクションが確立されていることを示す。同様に、図 1 4 において、3 9 1 0 がコネクション確立済み通信路管理テーブルを示し、3 9 1 1、3 9 1 2、3 9 1 3 の各エントリによって、3 本のコネクションが確立されていることを示す。このようにして、E n d - t o - E n d の確立済みの通信路が管理される。

【 0 0 7 5 】

次に、図 2 2、図 2 3 を用いて、並列コネクション確立方法における L A N スイッチ 3 0 5 0 の動作を説明する。図 2 2 は、アドレス・スイッチング・テーブルを除いて図 7 と同じであるため、図の説明は省略する。図 2 2 において、5 0 1 1 (a) は、クライアント 3 1 0 1 がサーバ 3 0 0 0 に対して、S Y N フラグ付 T C P パケットを送出した時のアドレス・スイッチング・テーブルを示し、5 0 1 1 (b) は、クライアント 3 1 0 1 がサーバ 3 0 0 0 に対して、A C K フラグ付 T C P パケットを送出してコネクションが確立したあとのアドレス・スイッチング・テーブルを示す。

【 0 0 7 6 】

図 2 3 を用いて、L A N スイッチ 3 0 5 0 の動作を説明する。クライアント 3 1 0 1 からサーバ 3 0 0 0 の n e t 1 . 1 宛の S Y N フラグ付 T C P パケットが、L A N スイッチ 3 0 5 0 のポート # 0、V C = 1 1 (3 0 7 4 a) ヘルディングされて届いた場合 (5 7 0 0) について説明する。L A N スイッチ 3 0 5 0 は、変換テーブル 5 0 0 0 を調べ、n e t 1 . 1 宛の S Y N フラグ付 T C P パケットの送出先ポート # が 1 であることを求める (5 7 1 0)。次に、ポート # 1 のバーチャル・チャンネルの中から V C = 1 (3 0 1 1 a) を選択し、(ポート # 0、V C = 1 1) と (ポート # 1、V C = 1) が相互にスイッチングされるように、アドレス・スイッチング・テーブル 5 0 1 1 (a) に登録する (5 7 2 0)。今、サーバ 3 0 0 0 がクライアント 3 1 0 1 のコネクション設立要求に対して、Q O S 管理テーブル・エントリ 5 1 1 0、5 1 2 0、5 1 3 0、5 1 4 0 を検索し、ネットワーク・インタフェースとして通信路 3 0 1 2、3 0 1 3、3 0 1 4 (ネットワーク・アドレス n e t 1 . 2, n e t 1 . 3 n e t 1 . 4) の 3 本を選択した場合について述べる。サーバ 3 0 0 0 は、まず、3 本の通信路を使用できるようにするために、L A N スイッチ 3 0 5 0 に対して (ポート # 2、V C = 1)、(ポート # 3、V C = 1)、(ポート # 4、V C = 1) の 3 つのバーチャル・チャンネルをリザーブする (5 7 3 0)。次に、サーバ 3 0 0 0 は、クライアント 3 1 0 1 に対して、L A N スイッチ 3 0 5 0 の (ポート # 1、V C = 1) から (ポート # 0、V C = 1 1) を経由して、A C K フラグ付 T C P パケットと、パラメタの付いた S Y N フラグ付 T C P パケットを返す。S Y N フラグには、選択したネットワーク・インタフェースのネットワーク・アドレスと各ネットワーク・インタフェースが使用できる平均バンド幅を Q O S として付加する (5 7 4 0)。その後、L A N スイッチ 3 0 5 0 に対して、(ポート # 1、V C = 1) と (ポート # 0、V C = 1 1) とのスイッチングを指定するエントリを削除するように要求し、この経路を他者が使用できるように解放する (5 7 5 0)。

【 0 0 7 7 】

クライアント 3 1 0 1 は、S Y N フラグ付 T C P パケットを受け、選択されたネットワーク・インタフェースのネットワーク・アドレスと、各ネットワーク・インタフェースが使用できる平均バンド幅を知ると、各ネットワーク・インタフェースに対して、A C K を返す。A C K を返す過程で、平均バンド幅の条件を満たすように、クライアント 3 1 0 1 から L A N スイッチ 3 0 5 0 に到るまでの経路のバーチャル・チャンネルを新たに確保する。

今、L A N スイッチ 3 0 5 0 へは、(ポート # 0、V C = 1 2)、(ポート # 0、V C = 1 3)、(ポート # 0、V C = 1 4) の 3 つの経路で到達したとする (5 7 6 0)。L A N スイッチ 3 0 5 0 は、3 つの A C K フラグ付 T C P パケットの宛先 n e t 1 . 2, n e t 1 . 3, n e t 1 . 4 とポート番号変換テーブル 5 0 0 0 より、ネットワーク・インタフェースへ到達するポート番号 # 2、# 3、# 4 を得る。L A N スイッチ 3 0 5 0 は、5 7 3 0 で予約されている (ポート # 2、V C = 1)、(ポート # 3、V C = 1)、

10

20

30

40

50

(ポート# 4、VC = 1)と、それぞれ(ポート# 0、VC = 12)、(ポート# 0、VC = 13)、(ポート# 0、VC = 14)をスイッチングするように、アドレス・スイッチング・テーブルのエントリを登録する(5770)。その結果、アドレス・スイッチング・テーブルは5011(b)に示すように更新され(5780)、3本のコネクションが確立されて(5790)、並列のSEND/RECEIVEによるデータ通信が行われる。

【0078】

以上のように、上記の本発明の実施形態例2によると、クライアントが要求するQOSとサーバの負荷に応じて、サーバは複数のネットワーク・インタフェースの中から、条件に見合ったネットワーク・インタフェースを複数個選択し、並列データ通信を行える。

10

【0079】

また、上記の実施形態例2によると、クライアント3101は、コネクション確立時にはサーバ3000の全てのネットワーク・インタフェースのネットワーク・アドレスを知らなくとも、クライアントが要求するQOSに応じたネットワーク・インタフェースを用いてコネクションを確立でき、並列データ通信を行うことができる。

【0080】

また、上記の実施形態例2ではネットワーク・インタフェースを選択する手順のひとつを示したに過ぎないが、QOS管理テーブルの情報の使い方によって、種々の選択手順を考えることができる。

【0081】

20

(3) 実施形態例3

上記の実施形態例2では、サーバ3000のネットワーク・インタフェース3001、3002、3003、3004には、それぞれ固有のネットワーク・アドレスが割り当てられていたが、図6と図24に示す実施形態例によれば、単一のネットワーク・アドレスをサーバ3000に割り当てただけで、ネットワーク・インタフェース3001、3002、3003、3004を用いて並列通信を行うことができる。

【0082】

図24において、3000はサーバ、3050はLANスイッチ、3011、3012、3013、3014、3074はそれぞれポート#1、#2、#3、#4、#0の通信路を表し、図3の同じ番号と対応する。7000は並列通信フラグ検出回路、7010はパケット・スイッチング回路、7100はパケットのヘッダ、7110は送信元ネットワーク・アドレス、7120は宛先ネットワーク・アドレス、7130は並列通信フラグ、7135はサーバ3000のポート番号を示す。並列通信フラグ7130とポート番号7135はパケットのオプションフィールドに追加した項目である。ネットワーク・アドレスnet1.1は、ポート#1ないし4の中で、ポート#1にのみ割り当てられている。

30

【0083】

図6において、3500、3501、3502、3503、3550、3551、3552は、TCPプロトコルの状態を示し、3690、3691、3692、3693、3694はクライアント3101とサーバ3000との間でコネクションを確立する3ウェイ・ハンドシェイクを示す。図6を用いて、第2の並列コネクション確立方法の処理の流れを示す。サーバ3000がLISTEN状態(3501)になるまでと、クライアント3101がSYN SENT状態(3561)になるまでは、図5と同じであるため説明を省略する。

40

【0084】

LISTEN状態(3501)のサーバ3000は、QOS(サービス品質)付きのTCPパケットをクライアント3101から受け取ると(3690)、クライアント3101のQOSを満たすネットワーク・インタフェースを探す。ひとつのネットワーク・インタフェースだけではクライアント3101が要求するQOSを満たすことができない場合には、複数のネットワーク・インタフェースを選択し、それらの合計値でクライアント3101が要求するQOSを満たすようにする。いま、図3におけるネットワーク・インタフ

50

エース3002、3003、3004の3個(ポート番号2、3、4)を選択した場合について述べる(3620)。サーバ3000は、ポート番号2、3、4に対応する3個のTCB(Transmission Control Block)を作成する(3621)。サーバ3000は、3個のTCBを作成した後、クライアント3101に対して、クライアントからのSYNに対するACKフラグの付いたTCPパケットと、SYNフラグの付いたTCPパケットをクライアント3101に送り(3691)、このTCBをSYN RECV状態にする(3502)。SYNフラグには、パラメタとしてポート番号の個数3、ポート番号2、3、4)を添付する。クライアント3101は、パラメタが付随したSYNフラグ付きTCPパケットを受け取ると、net1.1用に作成したTCBを削除し(3682)、net1.1かつポート番号2、net1.1かつポート番号3、net1.1かつポート番号4に対応した3個のTCBを作成する(3683)。次に、クライアント3101はnet1.1かつポート番号2、net1.1かつポート番号3、net1.1かつポート番号4に対して、それぞれACKフラグ付TCPパケットを送り(3692、3693、3694)、ESTABLISHED状態となる(3552)。サーバ3000も3個のACKを受け取るとESTABLISHED状態となり(3503)、3本のコネクションが確立される。以降、クライアントとサーバ間では、3本のコネクションを用いて、SEND/RECEIVEによる並列通信を行うことができる(3695)。

【0085】

図24に、“net1.1かつポート番号2”を表現できるパケットのヘッダ形式を示す。図24において、複数のポートを並列に用いて通信を行う場合には、パケットのヘッダのオプション・フィールドのなかの並列通信フラグ7130を“1”にセットし、同じくオプション・フィールドのなかのポート番号7135に“2”を設定する。LANスイッチ3050の中の並列通信フラグ検出回路は、パケットのヘッダのなかの並列通信フラグが1であることを検出すると、このパケットをパケット・スイッチング回路7010へ送り、パケットのヘッダのなかのポート番号を検出して、指定されたポート番号のネットワーク・インタフェースを用いて、サーバ3000にパケットを送る。以上のように、クライアント3101からサーバ3000に対して送り出すパケットに関しては、ネットワーク・インタフェース3001、3002、3003、3004を用いて並列通信を行うことができる。

【0086】

一方、サーバ3000からクライアント3101に対して送り出すパケットに関しては、図1で示したルーティング・テーブル・エントリの中のネットワーク・インタフェースのネットワーク・アドレスを、ポート番号として設定することによって、単一のネットワーク・アドレスをサーバ3000に割り当てるだけで、ネットワーク・インタフェース3001、3002、3003、3004を用いて並列通信を行うことができる。

【0087】

以上のように、パケットのヘッダ形式を拡張することによって、単一のネットワーク・アドレスをサーバ3000に割り当てるだけで、クライアント3101とサーバ3000は、ネットワーク・インタフェース3001、3002、3003、3004を用いて並列通信を行うことができる。

【0088】

上記(1)ないし(3)の実施形態例では、TCP/IPプロトコル層におけるクライアントとサーバ間のコネクション確立方式について述べ、ルーティング・テーブルにネットワーク・インタフェースのネットワーク・アドレスのカラム(71)を追加することによって、サーバはクライアントのQOS要求を満たすネットワーク・インタフェースを選択できるようになった。

【0089】

しかし、図25、図26に示す実施形態例によれば、図25に示すコネクション管理テーブルを用いることによって、クライアントのQOS要求を満たすだけでなく、サーバが

10

20

30

40

50

自らのネットワーク・インタフェース間の通信処理の負荷を分散させるように、ネットワーク・インタフェースを選択することができる。さらに、次のパケット送出先であるゲートウェイ・アドレスは従来通りルーティング・テーブルを用いて求めるが、パケットを送出するネットワーク・インタフェースは、図25に示すコネクション管理テーブルとLANスイッチに接続されているネットワーク・インタフェース・アドレスを記した登録テーブルを用いることによって、クライアントがサーバに接続を要求したネットワーク・インタフェースを選択できるようになる。その結果、クライアントがサーバに接続を要求したネットワーク・インタフェースを使った双方向通信が可能になる。

【0090】

図25は、図3のサーバ3000上にあるコネクション管理テーブルを示す。図25において、8050, 8051, 8052はコネクション管理テーブルの各エントリを示し、サーバ3000がクライアントとの間で開設しているコネクションに関する情報を保持する。コネクション管理テーブル・エントリ8050において、8001は通信で使用しているプロトコル名、8002はクライアントがサーバに対して接続を要求したサーバのネットワーク・インタフェースのネットワーク・アドレス、8003はサーバがコネクションを識別するためのポート番号、8004はクライアントのネットワーク・アドレス、8005はクライアント側のポート番号、8006は図4に示したようなCLOSED、LISTEN、ESTABLISHEDのようなコネクションの状態を示し、8006は次のコネクション管理テーブル・エントリへのポインタを示す。コネクション管理テーブルのエントリは、コネクション確立時に生成される。8040は、コネクション管理テーブルの先頭のエントリへのポインタを示す。

【0091】

クライアントとサーバの間で通信を行うときのコネクション識別子(ファイル記述子)8010ないし8012に対応して、コネクション管理テーブルの当該エントリへのポインタ8015ないし8017がある。ポインタ8015ないし8017は、各々のコネクション管理テーブルの各エントリ8050ないし8052を指す。

【0092】

8020は、LANスイッチに接続されているネットワーク・インタフェース・アドレスの登録テーブルを示す。エントリ8021ないし8024に登録されているネットワーク・インタフェース・アドレスのネットワーク・インタフェース3001ないし3004はいずれも、図3に示すようにLANスイッチ3050に接続される。それゆえ、8021ないし8024のいずれのネットワーク・インタフェースからパケットを送出しても、同じネットワークにパケットをルーティングできる。したがって、ネットワーク・インタフェース3001ないし3004の負荷を分散させるようにネットワーク・インタフェースを選択すれば、高いトータル・スループットを得ることができる。以下、詳細な実施形態例を記す。

【0093】

図26のフローチャートは、クライアント3101とサーバ3000の間の通信をTCP/IPプロトコル層の上位で実現する、本発明の並列ネットワーク接続方式の1実施形態例を示す。

【0094】

クライアント3101は、コネクション開設要求をサーバ3000に送る(8111)。サーバ3000は、コネクション開設要求を待ち、クライアント3101からのコネクション開設要求を受け付ける(8101)。クライアント3101は、QOSパラメタをサーバ3000に送る(8113)。このとき、QOSパラメタとして、複数のネットワーク・インタフェースを使用する場合にはネットワーク・インタフェースの数を指定したり、コネクションに対するピーク・バンド幅や平均バンド幅を指定する。サーバ3000はクライアント3101からQOSパラメタを受け取ると、図3に示すLANスイッチ3050からクライアント3101へルーティング可能かどうかを調べ、ルーティング可能ならば、図25に示したコネクション管理テーブルのエントリをたどって、ESTABLISHED

10

20

30

40

50

S H E D状態のコネクション数が最小のネットワーク・インタフェースを選択する(8103)。LANスイッチ3050からクライアント3101へルーティング可能かどうかは、図2に示した従来のルーティング・テーブル100ないし150を、クライアント3101のネットワーク・アドレスをキーにして検索し、得られたゲートウェイ・アドレスが、LANスイッチ3050のネットワーク・アドレスと接続されるならば、ルーティング可能であると判断できる。

【0095】

LANスイッチ3050からクライアント3101へルーティング可能な場合について、図25を用いてネットワーク・インタフェースの選択手順を詳細に説明する。コネクション管理テーブルの先頭エントリへのポインタ8040から、コネクション管理テーブル・エントリ8050をアクセスする。コネクションの状態8006がE S T A B L I S H E Dであるコネクション数をローカル・アドレス8002が等しいネットワーク・インタフェースごとに数える。ローカル・アドレス8002はサーバ3000のネットワーク・インタフェース・アドレスの中のいずれかと等しい。そこで、LANスイッチに接続されているネットワーク・インタフェース・アドレスの登録テーブル8020に登録されているネットワーク・インタフェースの中から、エントリ数が最小のネットワーク・インタフェースを求める。また、実施形態例(1)と同様に、図1に示すネットワーク・インタフェース情報テーブル中の次エントリへのポインタ84を使って各エントリをたどり、QOS管理テーブル・エントリへのポインタで示されるQOS管理テーブル・エントリに記されている90ないし97の統計情報にしたがって、負荷が最小のネットワーク・インタフェースを選択してもよい。例えば、バーチャル・チャンネル数94が最小のネットワーク・インタフェースを選択したり、最近1分間の平均転送量97が最小のネットワーク・インタフェースを選択してもよい。

【0096】

また、QOSパラメタとして、1より大きなネットワーク・インタフェースの数が記されていた場合には、コネクション数が少ないエントリから順々に指定された数だけ選ぶ。

【0097】

図26のフローチャートに戻る。サーバ3000は選択したネットワーク・インタフェースのネットワーク・アドレスと通信のためのポート番号をクライアント3101に送る(8105)。クライアント3101はネットワーク・アドレスとポート番号を受け取ると(8115)、通信用のs o c k e tを生成して(8117)、受信したサーバ3000のネットワーク・インタフェースに対してコネクションの開設要求を送る(8119)。クライアントが、サーバ3000が指定したネットワーク・インタフェースとコネクションを開設し直すことによって、サーバの各ネットワーク・インタフェースの負荷を分散できる。サーバ3000はクライアント3101からのコネクション再開要求を受付け(8107)ると、通信処理の負荷分散を図るために子プロセスをf o r kする。特に並列計算機のように各ノードにネットワーク・インタフェースが存在する場合には、選択したネットワーク・インタフェースに対応する各ノードに子プロセスをf o r kする(8108)。その後、クライアントとサーバ間でデータを送受信する(8109、8121)。サーバがクライアントへデータを送信する時には、当該コネクションに対応するサーバのコネクション識別子8010を用いて、コネクション管理テーブル・エントリへのポインタ8015からコネクション管理テーブル・エントリ8050を求めて、そのエントリ中のローカル・アドレス8002で示されるネットワーク・インタフェースからパケットを送出する(8109)。その結果、クライアントがサーバに接続を要求したネットワーク・インタフェースを使った双方向通信が可能になる。

【0098】

以上のように、LANスイッチに接続されているネットワーク・インタフェース・アドレスの登録テーブルを追加し、さらにコネクション管理テーブルのローカル・アドレスを用いることによって、従来のルーティング・テーブルを変更することなしに、クライアントがサーバに接続を要求したネットワーク・インタフェースを使った双方向通信を行える。

10

20

30

40

50

また、コネクション管理テーブルのエントリ数を数えることによって、各ネットワーク・インタフェースの負荷を判断し、サーバが主体となってクライアントとの通信で使用するネットワーク・インタフェースを選択することによって、通信処理の負荷分散を図ることができるようになり、ネットワーク処理のシステム・スループットを向上できる。

【0099】

【発明の効果】

以上のように、本発明によれば、複数のネットワーク・インタフェースを備えたサーバにおいて、クライアントが要求するQOSに応じたネットワーク・インタフェースを用いて通信を行ったり、複数本のネットワーク・インタフェースを用いて高速かつ高効率な並列通信を行うことができる。

10

【図面の簡単な説明】

【図1】本発明に係わるルーティング・テーブルとそのエントリ、ネットワーク・インタフェース情報テーブルのエントリ、QOS（サービス品質）管理テーブルのエントリの実施の形態例を示す図である。

【図2】従来のルーティング・テーブルとそのエントリ、ネットワーク・インタフェース情報テーブルのエントリを示す図である。

【図3】本発明のネットワーク接続システムの全体構成図である。

【図4】本発明のコネクション確立方法の実施の形態例を示す図である。

【図5】本発明の並列コネクション確立方法の実施の形態例を示す図である。

【図6】本発明の第2の並列コネクション確立方法の実施の形態例を示す図である。

20

【図7】本発明のコネクション確立方法におけるLANスイッチの動作の実施の形態例を説明するブロック図である。

【図8】本発明のコネクション確立方法におけるLANスイッチの動作の実施の形態例を説明するフローチャート図である。

【図9】本発明のコネクション確立方法におけるサーバの動作の実施の形態例を説明するフローチャート図である。

【図10】本発明のコネクション確立方法におけるサーバの動作の実施の形態例を説明するフローチャート図である。

【図11】本発明のコネクション確立方法におけるサーバの動作の実施の形態例を説明するフローチャート図である。

30

【図12】本発明のコネクション確立方法におけるクライアントの動作の実施の形態例を説明するフローチャート図である。

【図13】本発明のネットワーク接続方法におけるサーバのルーティング・テーブルの実施の形態例を説明する図である。

【図14】本発明のネットワーク接続方法におけるクライアントのルーティング・テーブルを説明する図である。

【図15】従来のサーバにおけるネットワーク接続を示す図である。

【図16】アプリケーション・プログラムにおけるクライアントとサーバのプログラムの実施の形態例を示す図である。

【図17】並列通信のためのデータの分割・統合方法の実施の形態例を示す図である。

40

【図18】本発明の並列コネクション確立方法におけるサーバの動作の実施の形態例を説明するフローチャート図である。

【図19】本発明の並列コネクション確立方法におけるサーバの動作の実施の形態例を説明するフローチャート図である。

【図20】本発明の並列コネクション確立方法におけるサーバの動作の実施の形態例を説明するフローチャート図である。

【図21】本発明の並列コネクション確立方法におけるクライアントの動作の実施の形態例を説明するフローチャート図である。

【図22】本発明の並列コネクション確立方法におけるLANスイッチの動作の実施の形態例を説明するブロック図である。

50

【図23】本発明の並列コネクション確立方法におけるLANスイッチの動作の実施の形態例を説明するフローチャート図である。

【図24】本発明の第2の並列コネクション確立方法におけるLANスイッチのブロック図と、パケットのヘッダ形式の実施の形態例を示す図である。

【図25】本発明の並列コネクション確立方法におけるサーバのルーティング方式を説明する図である。

【図26】本発明の第4の並列コネクション確立方法の実施の形態例を示す図である。

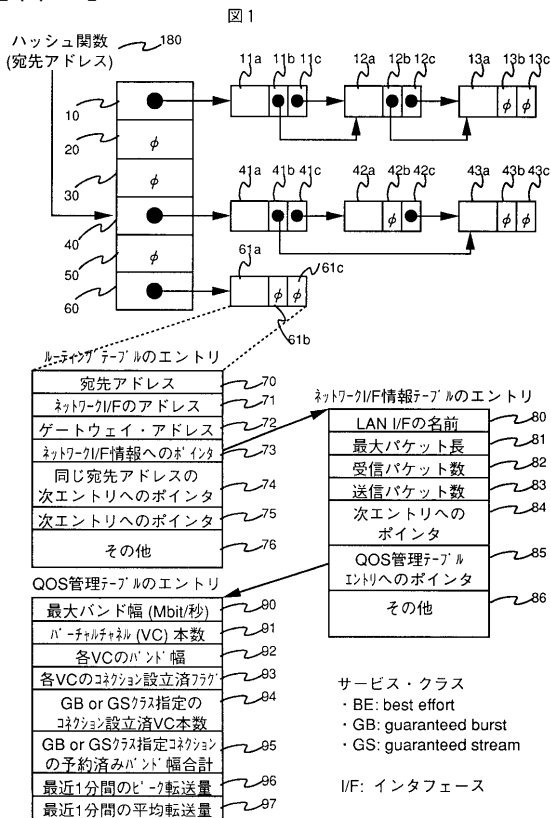
【符号の説明】

71：ネットワーク・インタフェースのネットワーク・アドレス、74：同じ宛先アドレスの次エントリへのポインタ、85：QOS管理テーブル・エントリへのポインタ、90：当該ネットワーク・インタフェースの最大バンド幅、92：各バーチャル・チャンネルのバンド幅、95：既に予約されているバンド幅の合計、96：最近1分間のピーク転送量、97：最近1分間の平均転送量、3000：サーバ、3011ないし3014：転送路、3050：LANスイッチ、3070：PBX、3101：クライアント1、3102：ゲートウェイ、3570：QOSパラメタ付きSYN（同期）パケット、3571：ACKパケット+ネットワーク・アドレス付きSYN（同期）パケット、3572：ACKパケット、3800：サーバのルーティング・テーブル、3900：クライアント1のルーティング・テーブル、9001ないし9015：サーバのプログラム例、9050ないし9061：クライアントのプログラム例、6030ないし6033：サーバのバッファ、6130ないし6133：クライアント1のバッファ。

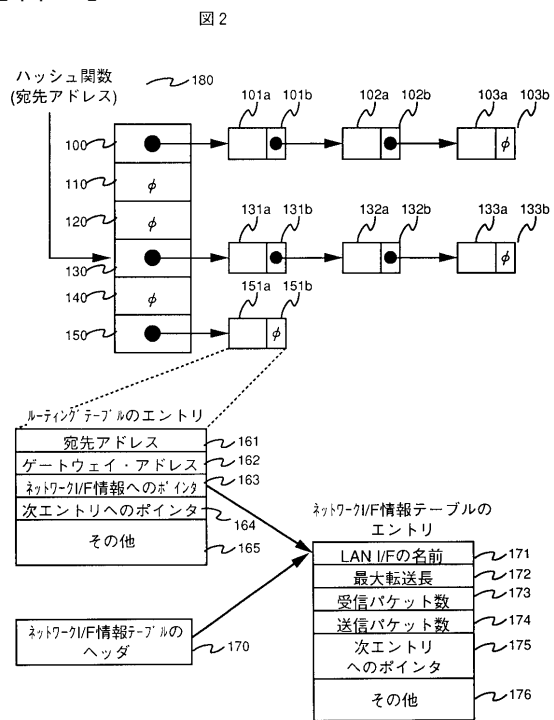
10

20

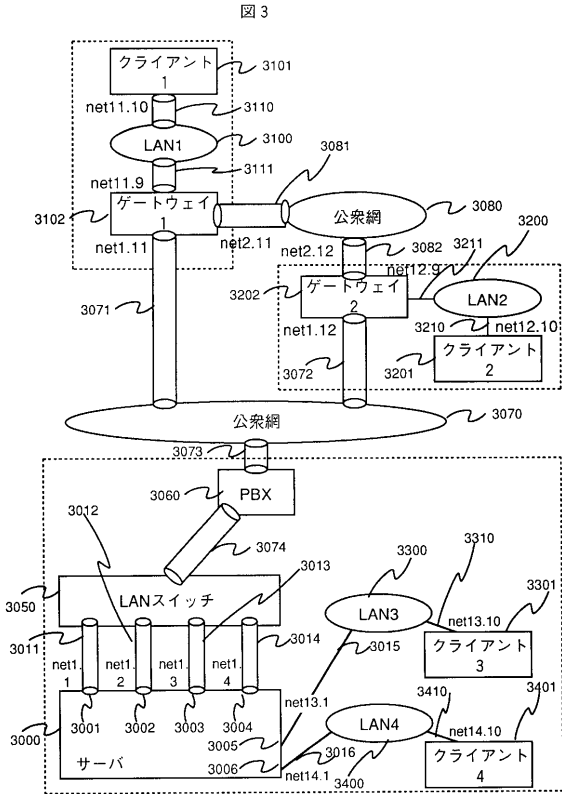
【図1】



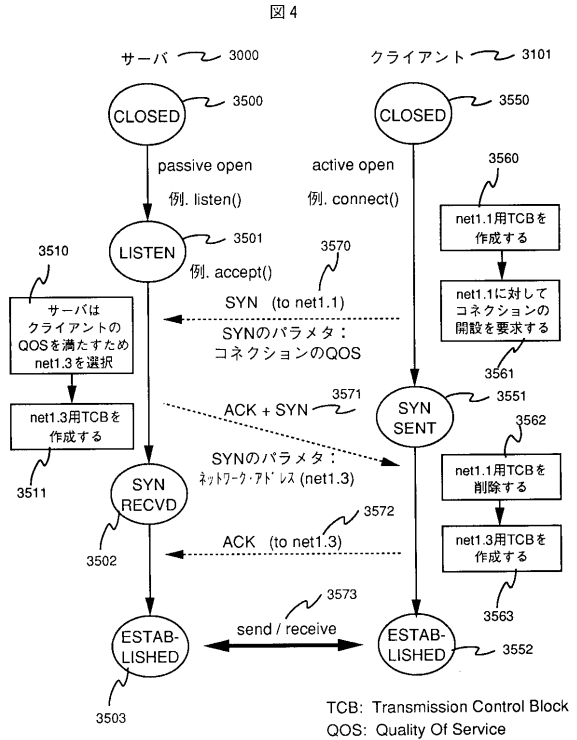
【図2】



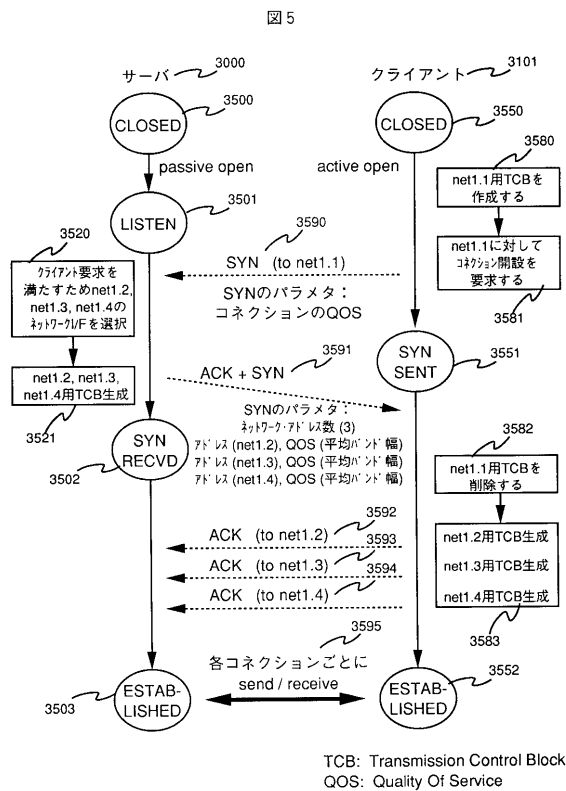
【 図 3 】



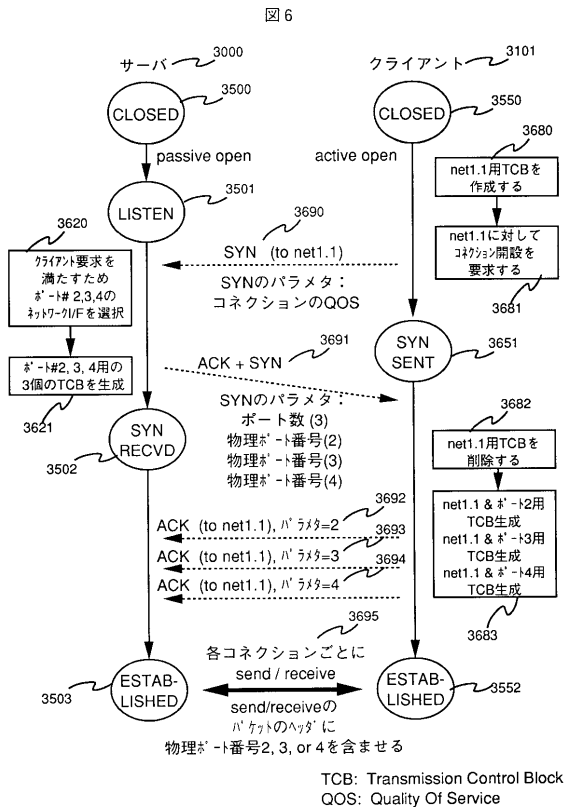
【 図 4 】



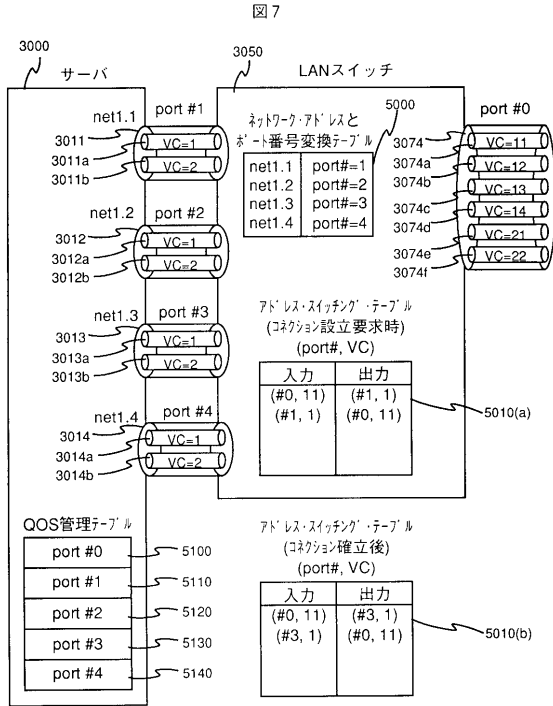
【 図 5 】



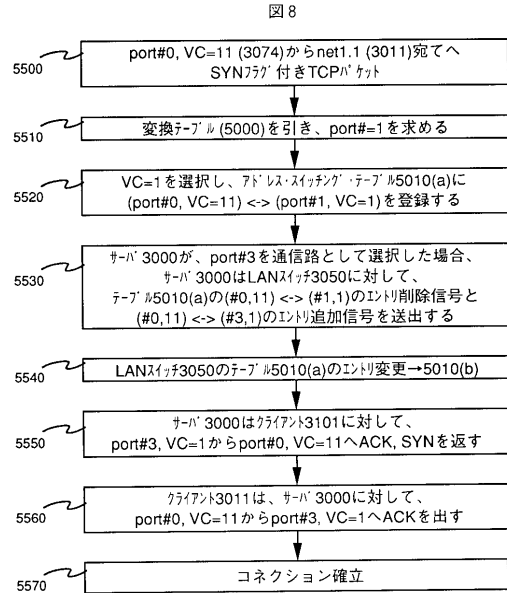
【 図 6 】



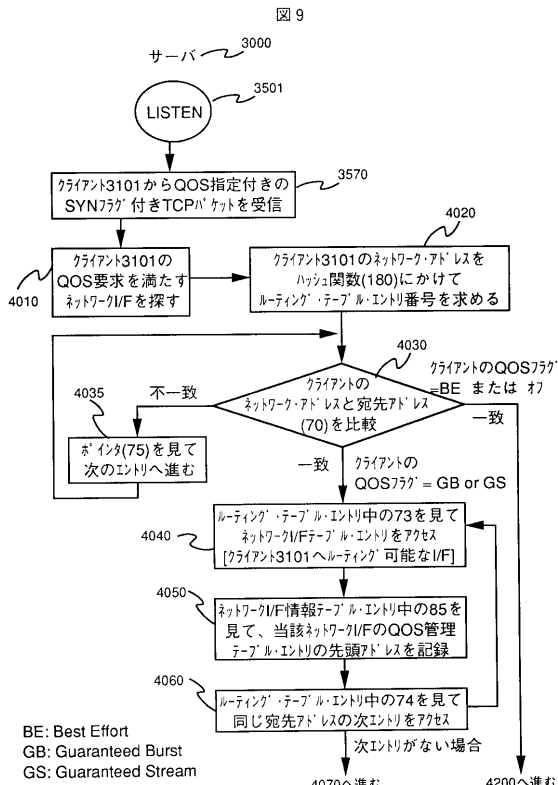
【 図 7 】



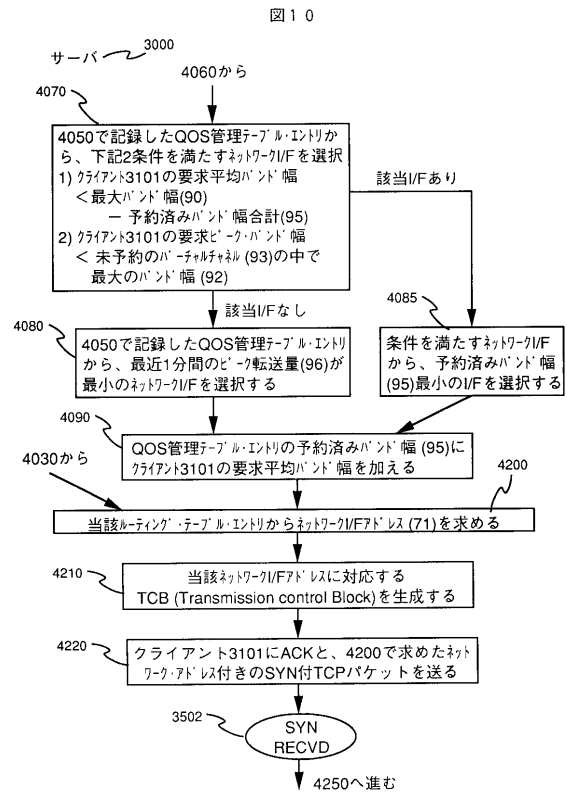
【 図 8 】



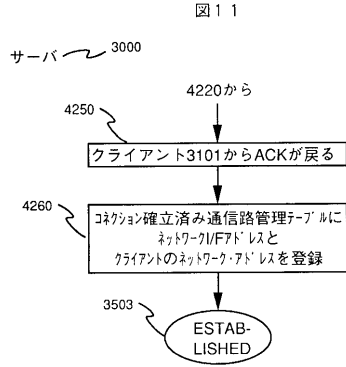
【 図 9 】



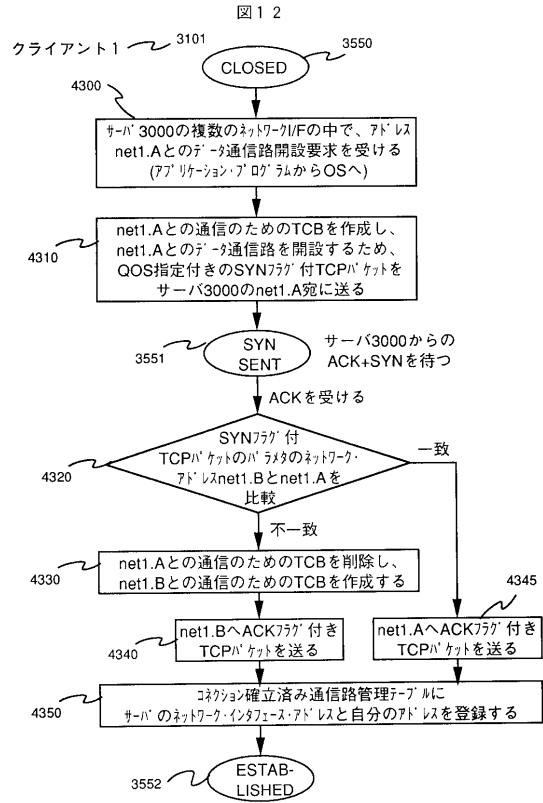
【 図 10 】



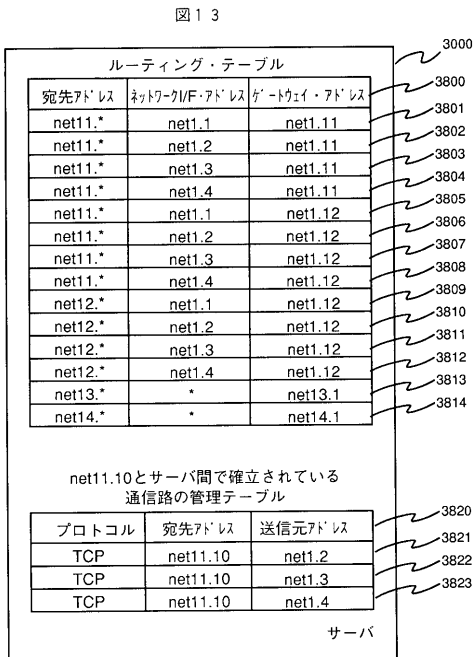
【 図 1 1 】



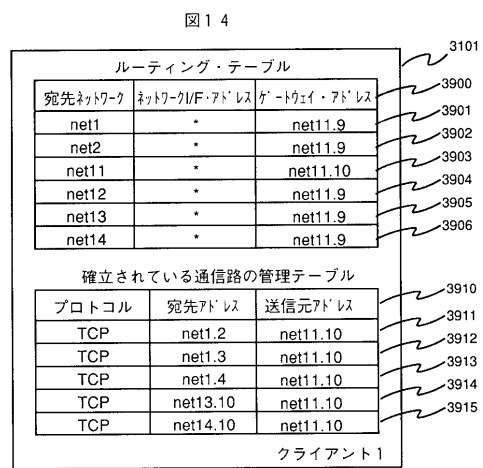
【 図 1 2 】



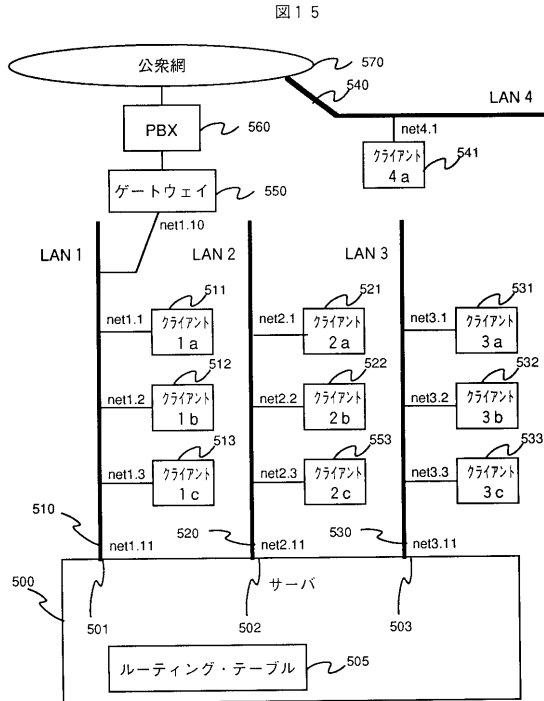
【 図 1 3 】



【 図 1 4 】



【 図 1 5 】



【 図 1 6 】

図 1 6

サーバ 3 0 0 0 のプログラム例

```

#define SERV_TCP_PORT 6001 ~ 9001
fd = socket(AF_INET, SOCK_STREAM, 0); ~ 9002

serv_addr.sin_family = AF_INET; ~ 9003
serv_addr.sin_addr.s_addr = htonl(INADDR_ANY); ~ 9004
serv_addr.sin_port = htons(SERV_TCP_PORT); ~ 9005

bind(fd, (struct sockaddr *) &serv_addr, ~ 9006
      sizeof(serv_addr));

listen (fd, 5); ~ 9007

for ( ; ; ) { ~ 9008
  newfd = accept(fd, (struct sockaddr *) ~ 9009
                &cli_addr, &cli_len);
  childpid = fork(); ~ 9010
  if (childpid == 0) { /* child process */ ~ 9012
    close(fd); ~ 9011
    send and receive data to and from CLIENT; ~ 9013
    exit(0); ~ 9014
  }
  close(newfd); /* parent process */ ~ 9015
}

クライアント 3 1 0 1 のプログラム例
#define SERV_TCP_PORT 6001 ~ 9050
#define SERV_HOST_ADDR net1.1 ~ 9051
#define QOS [SERVICE_CLASS, ~ 9052
           PEAK_BANDWIDTH, AVERAGE_BANDWIDTH];

fd = socket(AF_INET, SOCK_STREAM, 0); ~ 9053

serv_addr.sin_family = AF_INET; ~ 9054
serv_addr.sin_addr.s_addr = htonl(SERV_HOST_ADDR); ~ 9055
serv_addr.sin_port = htons(SERV_TCP_PORT); ~ 9056
serv_addr.qos = QOS; ~ 9057

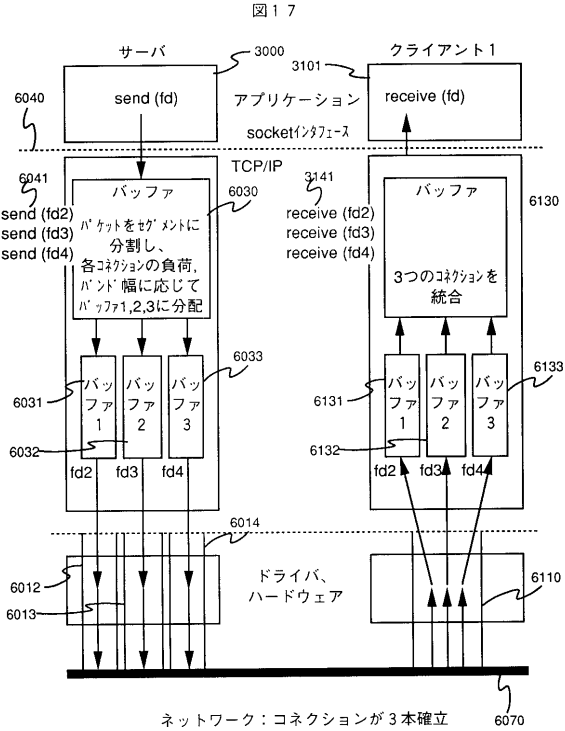
connect(fd, (struct sockaddr *) &serv_addr, ~ 9058
        sizeof(serv_addr));

send and receive data to and from SERVER; ~ 9059

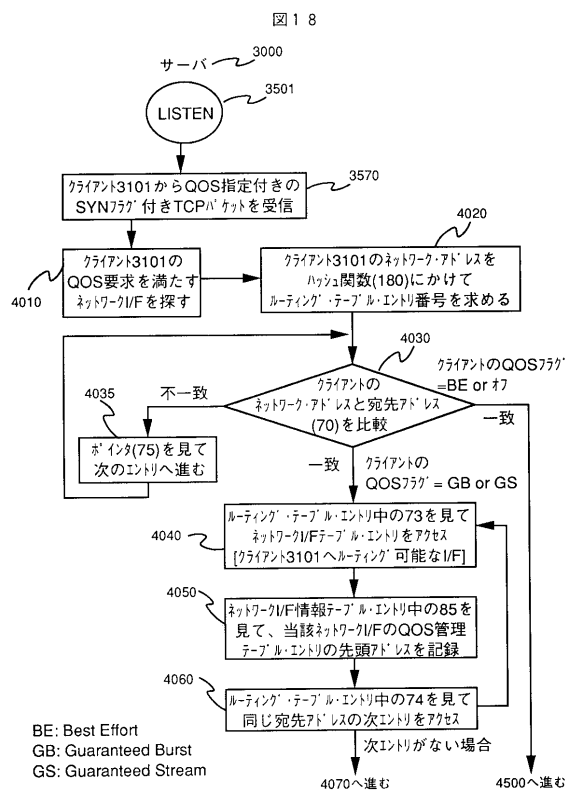
close(fd); ~ 9060
exit(0); ~ 9061

```

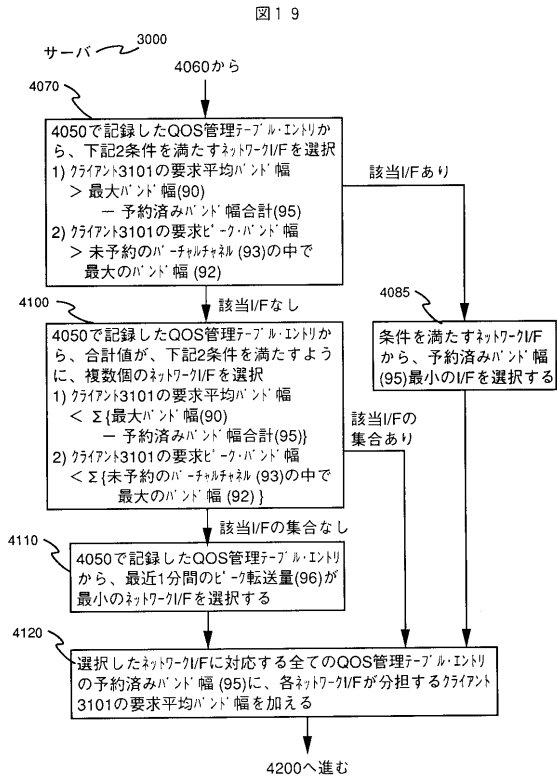
【 図 1 7 】



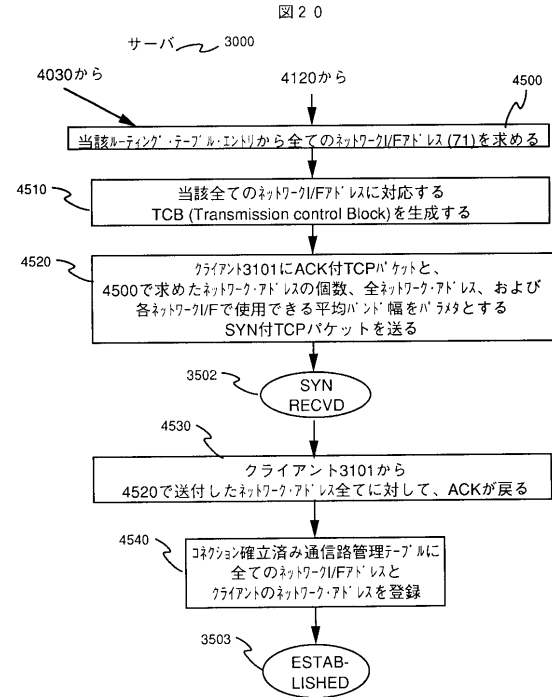
【 図 1 8 】



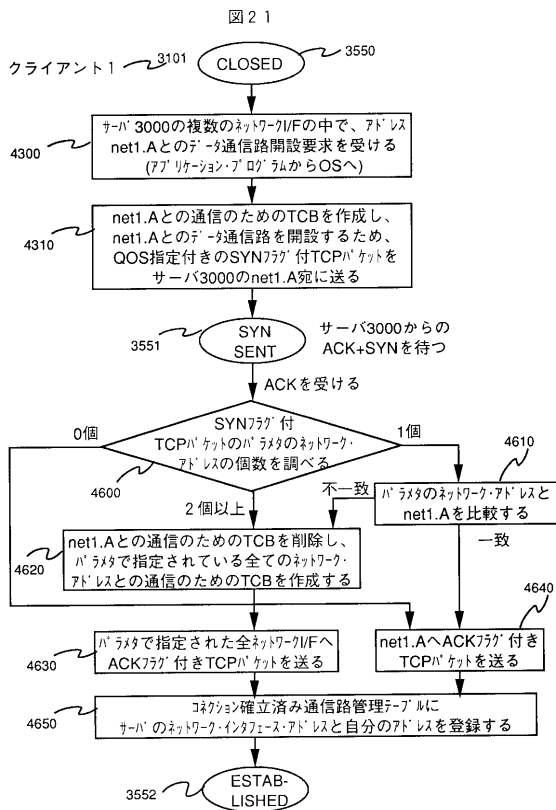
【図19】



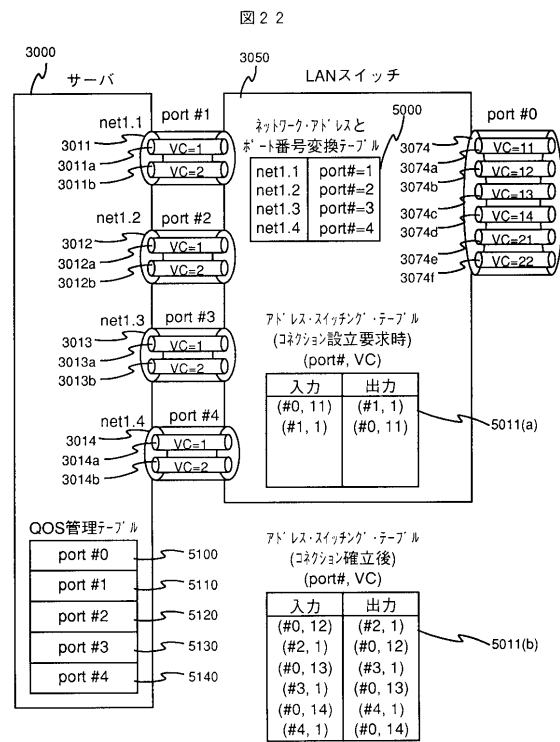
【図20】



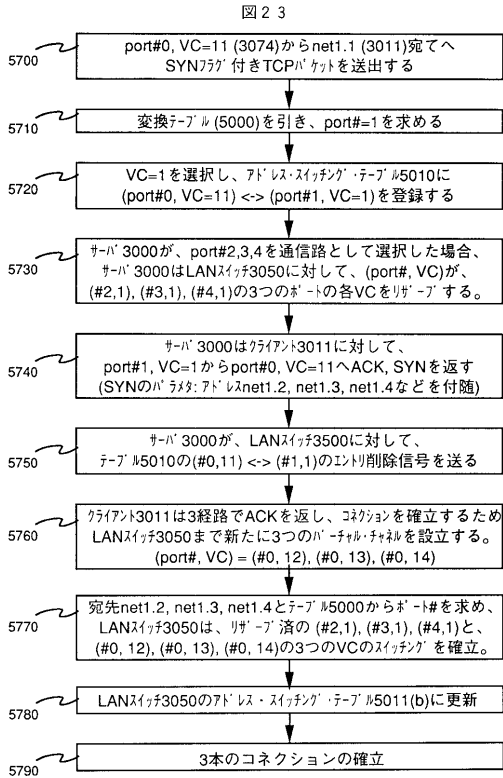
【図21】



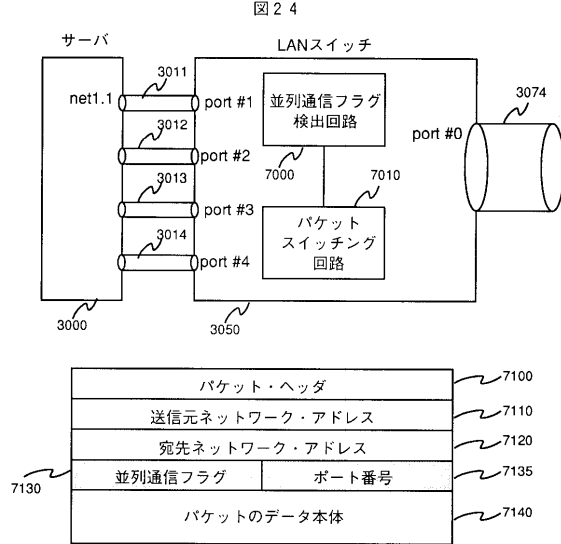
【図22】



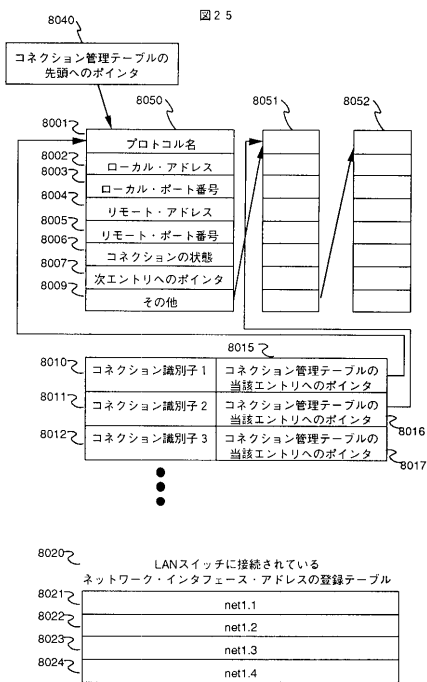
【 図 2 3 】



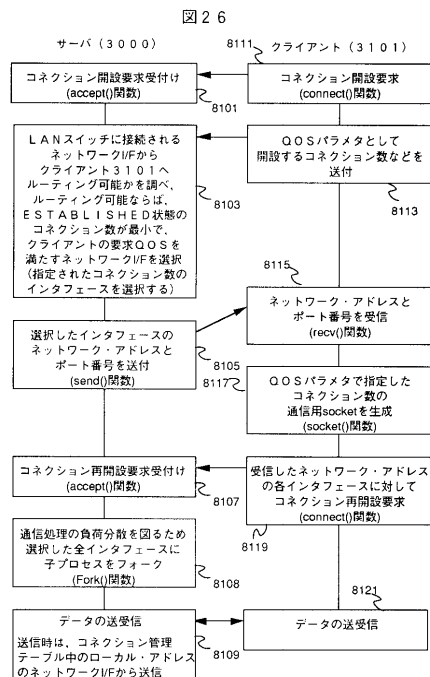
【 図 2 4 】



【 図 2 5 】



【 図 2 6 】



フロントページの続き

- (72)発明者 鍵政 豊彦
東京都国分寺市東恋ヶ窪1丁目280番地 株式会社日立製作所中央研究所内
- (72)発明者 野田 文雄
東京都国分寺市東恋ヶ窪1丁目280番地 株式会社日立製作所中央研究所内
- (72)発明者 増岡 義政
東京都国分寺市東恋ヶ窪1丁目280番地 株式会社日立製作所中央研究所内
- (72)発明者 高本 良史
東京都国分寺市東恋ヶ窪1丁目280番地 株式会社日立製作所中央研究所内

審査官 矢頭 尚之

- (56)参考文献 特開平07-170271(JP,A)
特開平07-170270(JP,A)
特開平07-074774(JP,A)

- (58)調査した分野(Int.Cl.⁷, DB名)
H04L 29/04
G06F 13/00 357
H04L 12/28