

(19) 日本国特許庁 (JP)

(12) 特 許 公 報 (B2)

(11) 特許番号

特許第5074872号
(P5074872)

(45) 発行日 平成24年11月14日 (2012.11.14)

(24) 登録日 平成24年8月31日 (2012.8.31)

(51) Int.Cl.

G 0 6 F 13/00 (2006.01)

F I

G 0 6 F 13/00 3 5 3 C

請求項の数 13 (全 20 頁)

(21) 出願番号	特願2007-248185 (P2007-248185)	(73) 特許権者	000001007
(22) 出願日	平成19年9月25日 (2007.9.25)		キヤノン株式会社
(65) 公開番号	特開2009-80584 (P2009-80584A)		東京都大田区下丸子3丁目30番2号
(43) 公開日	平成21年4月16日 (2009.4.16)	(74) 代理人	100076428
審査請求日	平成22年8月25日 (2010.8.25)		弁理士 大塚 康德
		(74) 代理人	100112508
			弁理士 高柳 司郎
		(74) 代理人	100115071
			弁理士 大塚 康弘
		(74) 代理人	100116894
			弁理士 木村 秀二
		(74) 代理人	100130409
			弁理士 下山 治
		(74) 代理人	100134175
			弁理士 永川 行光

最終頁に続く

(54) 【発明の名称】 プロトコル処理装置及び制御方法

(57) 【特許請求の範囲】

【請求項 1】

プロセッサによるデータ通信のプロトコル処理をオフロードするプロトコル処理装置であって、

前記プロセッサからの、パケット送受信においてパケットデータを格納するためのメモリアドレスを含む転送情報とパケット送受信に関わるタイマ制御のためのタイマ情報とを含む送受信要求を保持する保持手段と、

前記保持手段に保持された送受信要求に含まれる転送情報にしたがって、メモリとデータリンク層との間のデータ転送を実行する転送手段と、を備え、

前記転送手段は、前記データ転送が完了した後に、前記送受信要求に含まれるタイマ情報にしたがってタイマを制御し、前記タイマの満了時には当該タイマの種別に基づいて前記プロセッサに割り込み処理を実行させるか否かを判定し、前記プロセッサに割り込みを実行させると判定された場合は、割り込み信号を前記プロセッサへ出力することを特徴とするプロトコル処理装置。

【請求項 2】

前記タイマ情報が、タイムアウト値と、送受信のコネクション識別子と、タイマの種別と、タイマの起動を示す場合、

前記転送手段は、前記タイマ情報にしたがってタイマを起動し、前記タイムアウト値による当該タイマのタイマ満了時に、前記コネクション識別子と前記タイマの種別を前記プロセッサに通知することを特徴とする請求項 1 に記載のプロトコル処理装置。

10

20

【請求項 3】

前記タイマ情報が、送受信のコネクション識別子と、タイマの種別と、タイマの停止を示す場合、

前記転送手段は、前記コネクション識別子と前記タイマの種別によって特定されるタイマを停止することを特徴とする請求項 1 に記載のプロトコル処理装置。

【請求項 4】

前記タイマ情報は、更にタイマ満了時の通知先のプロセッサを示す通知先情報を含み、

前記転送手段は、前記タイマ情報にしたがって起動したタイマの満了時に、前記通知先情報によって示されるプロセッサに対して前記コネクション識別子と前記タイマの種別を通知することを特徴とする請求項 2 に記載のプロトコル処理装置。

10

【請求項 5】

前記タイマの種別は、再送タイマ、遅延 A C K タイマ、持続タイマのいずれかであることを特徴とする請求項 2 乃至 4 のいずれか 1 項に記載のプロトコル処理装置。

【請求項 6】

前記送受信要求の前記保持手段における格納アドレスを登録する登録手段を更に備え、

前記転送手段は、前記登録手段に登録された格納アドレスにしたがって前記保持手段から送受信要求を読み出し、読み出した送受信要求にしたがって、データ転送及びタイマ制御を実行することを特徴とする請求項 1 に記載のプロトコル処理装置。

【請求項 7】

プロセッサによるデータ通信のプロトコル処理をオフロードするプロトコル処理装置であって、

20

前記プロセッサからの送受信要求に応じて、メモリとデータリンク層との間のデータ転送を実行する第 1 の転送手段と、

タイマ値とともに、コネクション識別子とタイマの種別を含むネットワークタイマ情報をレジスタに保持することにより、前記送受信要求の処理に関連したタイマを起動するタイマ手段と、

前記タイマ手段で起動したタイマの計数を管理し、タイマ満了となったタイマに対応するレジスタに保持されている前記ネットワークタイマ情報に基づいて送信すべきパケットを決定し、該パケットを送信するべくメモリとデータリンク層との間のデータ転送を実行する第 2 の転送手段とを備え、

30

前記第 2 の転送手段は、前記タイマの種別に基づいて前記プロセッサに割り込み処理を実行させるか否かを判定し、前記プロセッサに割り込みを実行させると判定された場合は、割り込み信号を前記プロセッサへ出力することを特徴とするプロトコル処理装置。

【請求項 8】

前記第 2 の転送手段は、ネットワークタイマ情報に対応するディスクリプタを登録したメモリを有し、タイマ満了となったタイマのネットワークタイマ情報に対応するディスクリプタを前記メモリから読み出し、読み出したディスクリプタにしたがってデータ転送を行うことを特徴とする請求項 7 に記載のプロトコル処理装置。

【請求項 9】

前記第 2 の転送手段は、前記タイマの種別が T C P / I P 通信における A C K タイマを示す場合、自端末が受信したシーケンス番号を相手端末に通知する A C K パケットを送出させることを特徴とする請求項 7 に記載のプロトコル処理装置。

40

【請求項 10】

前記第 2 の転送手段は、前記タイマの種別が T C P / I P 通信における持続タイマを示す場合、プロービングパケットを構成するパケットを送出させることを特徴とする請求項 7 に記載のプロトコル処理装置。

【請求項 11】

前記第 2 の転送手段は、送信すべきデータの格納アドレスと転送サイズが記述された送信要求のアドレスを、前記ネットワークタイマ情報に基づいて特定し、特定されたアドレスに格納された送信要求にしたがってパケットを送信するべく、メモリとデータリンク層

50

との間のデータ転送を実行することを特徴とする請求項 7 に記載の Protokol 処理装置。

【請求項 1 2】

プロセッサによるデータ通信の Protokol 処理をオフロードする Protokol 処理装置の制御方法であって、

保持手段が、前記プロセッサからの、パケット送受信においてパケットデータを格納するためのメモリアドレスを含む転送情報とパケット送受信に関わるタイマ制御のためのタイマ情報とを含む送受信要求を保持する保持工程と、

転送手段が、前記保持工程で保持された送受信要求に含まれる転送情報にしたがって、メモリとデータリンク層との間のデータ転送を実行する転送工程と、を有し、

前記転送工程では、前記データ転送が完了した後に、前記送受信要求に含まれるタイマ情報にしたがってタイマを制御し、前記タイマの満了時には当該タイマの種別に基づいて前記プロセッサに割り込み処理を実行させるか否かを判定し、プロセッサに割り込みを実行させると判定された場合は、割り込み信号を前記プロセッサへ出力することを特徴とする Protokol 処理装置の制御方法。

【請求項 1 3】

プロセッサによるデータ通信の Protokol 処理をオフロードする Protokol 処理装置の制御方法であって、

第 1 の転送手段が、前記プロセッサからの送受信要求に応じて、メモリとデータリンク層との間のデータ転送を実行する第 1 の転送工程と、

タイマ制御手段が、タイマ値とともに、コネクション識別子とタイマの種別を含むネットワークタイマ情報をレジスタに保持することにより、前記送受信要求の処理に関連したタイマを起動するタイマ制御工程と、

第 2 の転送手段が、前記タイマ制御工程で起動したタイマの計数を管理し、タイマ満了となったタイマに対応するレジスタに保持されている前記ネットワークタイマ情報に基づいて送信すべきパケットを決定し、該パケットを送信するべくメモリとデータリンク層との間のデータ転送を実行する第 2 の転送工程とを備え、

前記第 2 の転送工程では、前記タイマの種別に基づいて前記プロセッサに割り込み処理を実行させるか否かを判定し、前記プロセッサに割り込みを実行させると判定された場合は、割り込み信号を前記プロセッサへ出力することを特徴とする Protokol 処理装置の制御方法。

【発明の詳細な説明】

【技術分野】

【0001】

本発明は、ネットワーク通信の分野に属し、より詳細にはプロセッサからパケット送受信処理をオフロードする Protokol 処理装置及び制御方法に関する。

【背景技術】

【0002】

ギガビットイーサネット（登録商標）などの普及により汎用 PC のみならず、組込み機器においても Protokol 処理を高速に実行できることが要求されている。ところで、ギガビットイーサネット（登録商標）の Full-wire の速度を達成するためには動作周波数 3 GHz 程度のプロセッサが必要になると言われている。これは、今日の組込み機器が一般的に搭載するプロセッサ能力を遥かに超えている。そこで、TOE（TCP/IP オフロードエンジン）といった Protokol 処理に特化した補助的デバイスをシステムに付加し、広帯域なネットワーク通信を実現することが一般化しつつある。

【0003】

さて、TCP/IP の Protokol 処理において重要となるのは、Protokol 処理と密接な関係を持つタイマ処理である（特許文献 1）。従来の Protokol 処理装置においては、ソフトウェアタイマによる実装がよく用いられている。ソフトウェアタイマは、書き換え可能な記憶装置にタイマ値を記憶させ、一定間隔でプロセッサなどが記憶装置上のタイマ値をアップ、またはダウンカウントする。そして、タイマ値の規定値への到達をタイマ満

10

20

30

40

50

了と見なし、タイマ満了時の処理を開始する手法である。ソフトウェアタイマは記憶装置の容量に比例して容易にタイマ数を増やすことが可能であるが、記憶装置上の全てのタイマ値を一定間隔でアップ、またはダウンカウントするため、プロセッサの負荷が増大する。

【 0 0 0 4 】

また、常時計時しているフリーランカウンタと記憶装置上のタイマ値を比較することでソフトウェアタイマが実現できる。このソフトウェアタイマでは、設定時のフリーランカウンタ値に規定値を加算した値をタイマ値として、一定間隔でフリーランカウンタとタイマ値を比較し、一致した場合にタイマが満了したと見なす手法である。この場合も、個別にタイマ値を更新する必要は無いものの、一定間隔で全てのタイマのタイマ値を読み出してフリーランカウンタと比較する必要があるため、タイマ数の増加がプロセッサ負荷の増大につながる。

10

【 0 0 0 5 】

さらに、TCP/IPのプロトコル処理においては、持続タイマ、遅延ACKタイマ、再送タイマ、キープアライブタイマ、2MSLタイマなど様々なタイマがコネクション毎に必要である。その結果、数十から数百のコネクションを処理する場合、その数倍のタイマが必要となるため、タイマを管理するためのプロセッサの負荷増大は計り知れないものとなる。

【 0 0 0 6 】

上述したソフトウェアタイマの欠点を補うために、ハードウェアタイマ(タイマコントローラ)を用いる場合がある。これは、メモリ領域とは別にタイムカウント専用のレジスタを有し、タイマのアップカウント、またはダウンカウントを専用ハードウェアにより実現する手法である。この専用ハードウェアは、タイマ値が規定値に到達すると割り込み信号を出力してプロセッサにタイマの満了を通知する。プロセッサはこの割り込み信号を受信すると通常の処理を一時中断し、レジスタや変数などを退避させた後に対応する割り込み処理を開始する。プロセッサはタイマ値のアップカウント、またはダウンカウントを行う必要がなくなるため、ソフトウェアタイマと比較してプロセッサの負荷を抑えることが可能になっている。すなわち、この手法によれば、記憶装置に対するトラヒックの増加や、プロセッサに対する処理負荷の増大を押さえることが可能である。

20

【 0 0 0 7 】

タイマを使用した具体的なプロトコル処理として、以下のような処理が存在する。

30

【 0 0 0 8 】

パケット送信後に再送タイマを起動し、再送タイマが満了した場合、送信したパケットがネットワーク途中で喪失したと判断し、パケットの再送信を行う。再送タイマが満了する前にACKパケットが到着した場合、再送タイマを停止する。一方、この再送タイマが満了になる以前にACK応答パケットを受信しなければ、それ以前に受信したACK応答パケットのシーケンス番号に戻ってTCPパケットが再度送信される。

【 0 0 0 9 】

パケットを受信すると遅延ACKタイマを起動し、異なるパケットを受信した場合は遅延ACKタイマを停止する。遅延ACKタイマが満了した場合は、パケット送信元へACKパケットを送信する。遅延ACKタイマはパケットを受信した場合にセットされ、ACK応答パケットを送出する時間を待機するのに用いられる。この待機により、複数のパケットに対してACK応答パケットを一つにまとめることができ、ACK応答パケットによるトラヒックとACKカウント処理を削減する効果が得られる。

40

【 0 0 1 0 】

持続タイマは通信相手先が0サイズの受信ウィンドウを公告してきた場合にセットされる。この場合、通信相手先は何も受信出来ない状態であるので自局側はパケットを送信出来ない状態が継続されることになる。このとき、相手先の受信ウィンドウの状態を調べるために、ペイロードが1バイトサイズのプロービングパケットと呼ばれるパケットを相手先に送信し、通信相手先のACKパケットの到着を待つ。このプロービングパケットの送

50

出タイミングを決定するために持続タイマは用いられる。

【0011】

これらのタイマは接続時や、コネクションクローズ時に使用されるタイマとは異なり、TCPのコネクションが確立されている期間にわたって何度も使用される。そのため、割り込みを使用したタイマの場合、データ転送と並行してこれらのタイマに起因した割り込み処理が幾度も行われることになる。

【0012】

以上述べたタイマ以外にもコネクション確立タイマ、キープアライブタイマ、FIN_WAIT_2タイマ、2MSLタイマなどがコネクションの確立時や切断時に使用されている。さらにTCP/IP通信を実現するためにこれらのタイマ値を基に多くのタイマ変数が用いられている。従って、通信システムにおけるタイマ処理のためには十分な記憶領域の確保と、タイマ処理能力をもったプロセッサを用いる必要がある。

10

【0013】

パケット送受信処理がプロセッサからオフロードされているプロトコル処理装置においては、パケット送受信処理はプロセッサからの要求によって開始され、パケット送受信処理の完了は割り込み信号によってプロセッサへ通知される。パケット送受信処理完了が通知されたプロセッサは、完了したパケット送受信処理の内容に応じてタイマの起動・停止を行っている。割り込み処理に関しては、割り込みディスクリプタテーブルによってタイプ別に割り込みを処理することが可能な割り込みコントローラが特許文献2に記載されている。

20

【特許文献1】特開昭59-17757号公報

【特許文献2】特開2005-202964号公報

【発明の開示】

【発明が解決しようとする課題】

【0014】

以上述べたように上記従来例に見られるプロトコル処理装置では、TCP/IPでは持続タイマ、遅延ACKタイマ、再送タイマ、キープアライブタイマ、2MSLタイマなどのタイマセットがコネクション毎に必要である。数十から数百のコネクションを使用する場合、コネクションの数だけタイマセットが用いられるので、プロセッサのタイマ管理に関する負荷が大きくなってしまいう課題があった。

30

【0015】

また、タイマセットの中でも遅延ACKタイマ、再送タイマ、持続タイマはパケットの送受信中に、頻繁にタイマの更新を行う必要がある場合があり、タイマ管理負荷の増大の一つの要因となっている。更に、ハードウェアタイマを用いた場合でも割り込み処理によるタイマの起動・停止がプロセッサの負荷となっている。複数コネクションが並列に動作し、タイマの起動・停止が頻発した場合、プロセッサの負荷が大きくなり、他処理に割り当て可能なプロセッサのリソースが減少してしまう。

【0016】

本発明は、上記の課題に鑑みてなされたものであり、その目的は、プロトコル処理に含まれるタイマに関わる処理をプロセッサからオフロードし、プロセッサの負荷を軽減するプロトコル処理装置を提供することにある。

40

【課題を解決するための手段】

【0017】

上記の目的を達成するための本発明の一態様によるプロトコル処理装置は以下の構成を備える。すなわち、

プロセッサによるデータ通信のプロトコル処理をオフロードするプロトコル処理装置であって、

前記プロセッサからの、パケット送受信においてパケットデータを格納するためのメモリアドレスを含む転送情報とパケット送受信に関わるタイマ制御のためのタイマ情報とを含む送受信要求を保持する保持手段と、

50

前記保持手段に保持された送受信要求に含まれる転送情報にしたがって、メモリとデータリンク層との間のデータ転送を実行する転送手段と、を備え、

前記転送手段は、前記データ転送が完了した後に、前記送受信要求に含まれるタイマ情報にしたがってタイマを制御し、前記タイマの満了時には当該タイマの種別に基づいて前記プロセッサに割り込み処理を実行させるか否かを判定し、プロセッサに割り込みを実行させると判定された場合は、割り込み信号を前記プロセッサへ出力する。

【 0 0 1 8 】

また、上記の目的を達成するための本発明の他の態様によるプロトコル処理装置は以下の構成を備える。すなわち、

プロセッサによるデータ通信のプロトコル処理をオフロードするプロトコル処理装置であって、

前記プロセッサからの送受信要求に応じて、メモリとデータリンク層との間のデータの転送処理を実行する第 1 の転送手段と、

タイマ値とともに、コネクション識別子とタイマの種別を含むネットワークタイマ情報をレジスタに保持することにより、前記送受信要求の処理に関連したタイマを起動するタイマ手段と、

前記タイマ手段で起動したタイマの計数を管理し、タイマ満了となったタイマに対応するレジスタに保持されている前記ネットワークタイマ情報に基づいて送信すべきパケットを決定し、該パケットを送信するべくメモリとデータリンク層との間のデータの転送処理を実行する第 2 の転送手段とを備える。

【発明の効果】

【 0 0 1 9 】

本発明によれば、プロトコル処理に含まれるタイマに関わる処理をプロセッサからオフロードするので、プロセッサの負荷を軽減するプロトコル処理装置が提供される。

【発明を実施するための最良の形態】

【 0 0 2 0 】

以下、添付の図面を参照して、本発明の好適な実施形態を説明する。

【 0 0 2 1 】

[第 1 実施形態]

図 1 は、第 1 実施形態に係るプロトコル処理装置を含むシステムの構成例を説明する図である。

【 0 0 2 2 】

図 1 において、101 はプロセッサであり、アプリケーションや OS、デバイスドライバなどの一般的なソフトウェアを実行する。102 はデータ転送部であり、プロセッサからのパケット送受信要求を受けて主記憶部 105 と MAC / PHY 106 の間でパケットの送受信処理を行う。103 はタイマであり、再送タイマ、遅延 ACK タイマ、持続タイマとして使用される。タイマ 103 は複数の時間を同時に計測することが可能である。タイマ 103 の各計測には、

- ・再送タイマ、遅延 ACK タイマ、持続タイマのいずれとして動作しているかを示すタイマ種別と、
 - ・タイムアウト値と、
 - ・タイマが使用されるコネクションを識別するためのコネクション識別子とが設定される。
- タイマ 103 は、起動されると、設定されたタイムアウト値を計測し、タイマ満了時に割り込み信号をアサートする。

【 0 0 2 3 】

104 は割り込みコントローラであり、タイマ 103 や MAC / PHY 106、データ転送部 102 から出力される割り込み信号をプロセッサ 101 へ通知する。105 は主記憶部であり、プロセッサ 101 が実行するプログラムとその作業領域のほか、データ転送部 102 が利用するデータ、作業領域およびプロセッサ 101 とデータ転送部 102 との通信用領域などが置かれる。106 はネットワークの MAC (データリンク層) や PHY

(物理層)であり、ネットワークから図示されないスイッチ、ハブ、ルータ等を介して他のネットワーク機器と接続される。107はバスであり、プロセッサ101、データ転送部102、タイマ103、割り込みコントローラ104、主記憶部105、MAC/PHY106を接続する。

【0024】

続いて、図1に示したプロトコル処理装置の動作について説明する。なお、プロトコル処理装置のタイマ103としては、再送タイマ、遅延ACKタイマ、持続タイマなどが存在するが、本実施形態においては再送タイマを用いて説明を行う。図2は、第1実施形態によるデータ通信処理を示すフローチャートである。

【0025】

まず、パケット送信時の動作について説明する。プロセッサ101上で動作するユーザアプリケーションプログラムからTCPによるストリーム送信が要求されると、プロセッサ101は主記憶部105上にあるストリームデータからパケットを生成する(図2のステップS101)。パケット生成後、プロセッサ101はデータ転送部102に対するパケット送受信要求を主記憶部105へ書き込む(ステップS102)。

【0026】

本実施形態のパケット送受信要求は、図3の(a)で示すように、転送情報とタイマ情報とで構成される。転送情報は、

- ・データ転送部102にパケット送信処理を行わせるか、パケット受信処理を行わせるかを示す要求処理、

- ・パケット送信処理要求時には送信するパケットの主記憶部105上における先頭アドレスを、パケット受信処理要求時には受信するパケットを格納する主記憶部105上における先頭アドレスを示す転送先/転送元アドレス、

- ・転送するデータ量を示す転送サイズ、

を有する。また、タイマ情報は、

- ・パケット送受信処理を行うコネクションを識別するためのコネクション識別子と、

- ・パケット送受信処理後に操作するタイマの種別を示すタイマ種別と、

- ・タイムアウト値とタイマに対して起動・停止を行うかを示すタイマ操作と、

- ・パケット送受信処理後に複数タイマを操作する場合に次のタイマ情報のアドレスを示す次タイマ情報アドレスとを有する。

【0027】

本実施形態においてはタイマ103にタイムアウト値を設定する例を示しているが、タイマ種別によってタイムアウト値を固定できる場合のタイマ情報のデータ構成例は、タイムアウト値が不要となるので、図3の(d)のようになる。例えば、ストリーム送信が要求されるコネクションの識別子が0であり、送信すべきパケットの先頭アドレスが0x1000_0000、転送サイズが1024バイトであったとする。この場合、プロセッサ101は図3の(b)に示すパケット送受信要求(送信要求)を主記憶部105へ書き込む。こうして、主記憶部105は、プロセッサ101からの、パケット送受信のための転送情報と、タイマ制御(タイマ操作)のためのタイマ情報とを含むパケット送受信要求を保持する保持部として機能する。

【0028】

パケット送受信要求を主記憶部105へ書き込んだ後、プロセッサ101はパケット送受信要求の主記憶部105上の先頭アドレスをデータ転送部102へ通知する(ステップS103)。

【0029】

データ転送部102は通知されたアドレスを用いて、主記憶部105からパケット送受信要求を読み出す(ステップS111)。そして、データ転送部102は、読み出したパケット送受信要求から、当該要求が送信要求か受信要求かを判定する(ステップS112)。ここでは、送信要求であるので、データ転送部102は、読み出したパケット送受信要求に従って転送先/転送元アドレスで指定されたアドレスから転送サイズで指定された

10

20

30

40

50

データ量のデータを読み出し、MAC / PHY 106へ転送する（ステップS113）。
こうして、データ転送部102は、主記憶部105に保持された送受信要求に含まれる転送情報にしたがって、メモリとデータリンク層との間のデータの転送処理を実行する。MAC / PHY 106へパケットを転送した後、データ転送部102は、パケット送受信要求に含まれているタイマ情報に基づいてコネクション識別子とタイマ種別、タイムアウト値をタイマ103に設定し、時間計測を起動する（ステップS115）。すなわち、本実施形態のデータ転送部102は、データの転送処理が完了した後に、当該送受信要求に含まれているタイマ情報にしたがってタイマを操作するタイマ制御部としても機能するものである。図3の（b）に示される送受信要求の場合、タイマ種別 = 再送タイマ、タイマ操作 = 起動であるので、再送タイマが起動される。

10

【0030】

タイマ103は設定されたタイマの計測を開始する（ステップS121）。そして、タイムアウト値まで時間を計測すると、タイマ満了の割り込みを出力する（ステップS122、S123）。

【0031】

タイマ103から出力された割り込み信号は、割り込みコントローラ104経由で、プロセッサ101に通知される。割り込みを通知されたプロセッサ101は、タイマ103から当該満了した時間計測に関して設定されているコネクション識別子とタイマ種別を読み出す。すなわち、タイムアップを通知したタイマから、そのタイマに設定されたコネクション識別子とタイマ種別をプロセッサ101が読み出す。そして、読み出したコネクション識別子により処理を行うコネクションを特定し、タイマ種別により行う割り込み処理を特定する（ステップS104）。例えば、上記の例では、タイマ種別が再送タイマとなっているので、プロセッサ101は、コネクション識別子で示されるコネクションに対して、パケットの再転送処理を開始する。

20

【0032】

次に、パケット受信時の動作について説明する。MAC / PHY 106がネットワークからパケットを受信すると、割り込み信号をアサートし、割り込みコントローラ104がプロセッサ101へ通知する。割り込みを通知されたプロセッサ101は、受信されたパケットが格納されるMAC / PHY 106のバッファからIPヘッダ及びTCPヘッダを読み出す。プロセッサ101はIPヘッダ内の送信元 / 受信先IPアドレスからコネクション識別子を判断する。また、TCPヘッダ内のACKフラグが立ち、かつ、確認応答番号がその時点で最後に送信したデータのシーケンス番号 + 1であった場合、プロセッサ101は、図3の（c）に示すパケット送受信要求（受信要求）を主記憶部105へ書き込む。パケット送受信要求を主記憶部105へ書き込んだ後、プロセッサ101はパケット送受信要求の主記憶部105上の先頭アドレスをデータ転送部102へ通知する。なお、以上のプロセッサ101の動作についてはフローチャートによる図示は省略した。

30

【0033】

データ転送部102は、通知された先頭アドレスからパケット送受信要求を読み出す（ステップS111）。そして、読み出したパケット送受信要求（ここでは受信要求）に従ってMAC / PHY 106のバッファから転送サイズで指定されたサイズ分のデータを読み出し、主記憶部105上の転送先 / 転送元アドレスで指定されたアドレスへ書き込む。この処理は、図2のステップS112、S114に示されている。図3の（c）において、タイマ操作 = 停止となっているので、主記憶部105への書き込みが終了した後、データ転送部102はタイマ情報のコネクション識別子とタイマ種別から特定されるタイマ103の時間計測を停止する（ステップS116）。このように、受信処理においても、主記憶部105が、プロセッサ101からの、パケット送受信のための転送情報と、タイマ制御のためのタイマ情報とを含む送受信要求を保持する。そして、データ転送部102は、主記憶部105に保持された送受信要求に含まれる転送情報にしたがって、データリンク層のバッファから主記憶部105へのデータの転送処理を実行する。そして、データ転送部102は、データの転送処理が完了した後に、送受信要求に含まれるタイマ情報にし

40

50

たがってタイマを制御（操作）する（上記の例では、特定のタイマの停止を行う）タイマ制御部としても機能する。

【 0 0 3 4 】

1つのパケット送受信要求において複数のタイマを起動・停止する場合は、タイマ情報内の次タイマ情報には次タイマ情報の開始アドレスが格納され、操作すべきタイマがリスト化される。データ転送部 1 0 2 は、ステップ S 1 1 5 またはステップ S 1 1 6 において、パケット送受信処理終了後、リスト化されたタイマ情報を手繰り、順にタイマ操作を行う。

【 0 0 3 5 】

以上のように、第 1 実施形態によれば、プロトコル処理と密接な関係を持つタイマの起動・停止の操作がプロセッサからオフロードされることにより、プロセッサの負荷を軽減することが可能となる。そして、プロセッサのリソースを他プロトコル処理に割り当てることが可能になるため、プロトコル処理の高速化が可能になる。

【 0 0 3 6 】

[第 2 実施形態]

図 4 は、第 2 実施形態に係るプロトコル処理装置を含むシステムの構成例を示す図である。

【 0 0 3 7 】

図 4 において、送信処理用プロセッサ 2 0 1 は T C P / I P 送信処理用のプロセッサであり、受信処理用プロセッサ 2 0 8 は T C P / I P 受信処理用のプロセッサである。2 0 2 はデータ転送部であり、送信処理用プロセッサ 2 0 1、受信処理用プロセッサ 2 0 8 からのパケット送受信要求を受けて主記憶部 2 0 5 と M A C / P H Y 2 0 6 間でパケットの転送を行う。2 1 0 はパケット送受信要求格納部であり、データ転送部 2 0 2 がパケット送受信処理中であるために、送信処理用プロセッサ 2 0 1 や受信処理用プロセッサ 2 0 8 からのパケット送受信要求を受け取れない場合のためにパケット送受信要求を格納する。より具体的には、主記憶部 2 0 5 上の、パケットの送受信要求が格納されている先頭アドレスがパケット送受信要求格納部 2 1 0 に格納される。なお、送信処理用プロセッサ 2 0 1 がデータ転送部 2 0 2 の状態をチェックし、パケット送受信要求を受け取れないと判定した場合に、パケット送受信要求の格納アドレスをパケット送受信要求格納部 2 1 0 に格納するように構成しても良い。

【 0 0 3 8 】

2 0 3 はタイマであり、再送タイマ、遅延 A C K タイマ、持続タイマ等として使用される。タイマ 2 0 3 は複数の時間を同時に計測することが可能である。タイマ 2 0 3 の各計測には、

- ・再送タイマ、遅延 A C K タイマ、持続タイマのいずれとして動作しているかを示すタイマ種別と、
- ・タイムアウト値と、
- ・タイマが使用されるコネクションを識別するためのコネクション識別子と、
- ・満了時の割り込み通知先を示すタイマ満了時通知先とが設定される。

第 2 実施形態では、起動されたタイマ 2 0 3 は設定されたタイムアウト値を計測する。そして、タイマ 2 0 3 は、タイマ満了時には送信処理用割り込みコントローラ 2 0 4 や受信処理用割り込みコントローラ 2 0 9 との協働により、設定された割り込み通知先へ割り込み信号をアサートする。

【 0 0 3 9 】

2 0 4、2 0 9 はそれぞれ送信処理用割り込みコントローラ、受信処理用割り込みコントローラであり、タイマ 2 0 3 の割り込み信号や、M A C / P H Y 2 0 6 からの割り込み信号を送信処理用プロセッサ 2 0 1 や受信処理用プロセッサ 2 0 8 へ通知する。2 0 5 は主記憶部であり、送信処理用プロセッサ 2 0 1 及び受信処理用プロセッサ 2 0 8 が実行するプログラムとその作業領域が配置される。更に、主記憶部 2 0 5 には、データ転送部 2 0 2 の利用するデータや作業領域、送信処理用プロセッサ 2 0 1 及び受信処理用プロセッサ

10

20

30

40

50

サ 2 0 8 とデータ転送部 2 0 2 との間の通信用領域などが置かれる。2 0 6 はネットワークの M A C (データリンク層) や P H Y (物理層) であり、ネットワークから図示されないスイッチ、ハブ、ルータ等を介して他のネットワーク機器と接続される。2 0 7 はバスであり、上述した各部を接続する。

【 0 0 4 0 】

続いて、第 2 実施形態におけるプロトコル処理装置の動作について説明する。プロトコル処理装置が実行するプロトコル処理において再送タイマ、遅延 A C K タイマ、持続タイマなどのタイマが存在するが、本実施形態においては再送タイマを用いて説明する。なお、第 2 実施形態におけるプロトコル処理装置の動作を示すフローチャートとして、図 2 を流用する。

10

【 0 0 4 1 】

送信処理用プロセッサ 2 0 1 上で動作するユーザアプリケーションプログラムから T C P によるストリーム送信が要求されると、送信処理用プロセッサ 2 0 1 は主記憶部 2 0 5 上にあるストリームデータからパケットを生成する (ステップ S 1 0 1) 。パケット生成後、送信処理用プロセッサ 2 0 1 は主記憶部 2 0 5 上にデータ転送部 2 0 2 に対するパケット送受信要求を書き込む (ステップ S 1 0 2) 。

【 0 0 4 2 】

パケット送受信要求は図 5 の (a) に示すように、

- ・データ転送部 2 0 2 にパケット送信処理を行わせるか、パケット受信処理を行わせるかを示す要求処理と、
- ・パケット送信処理要求時には送信するパケットの主記憶部 2 0 5 上の先頭アドレスを示し、パケット受信処理要求時には受信するパケットを格納する主記憶部 2 0 5 上の先頭アドレスを示す転送先 / 転送元アドレスと、
- ・転送するデータ量を示す転送サイズ、

の各情報 (転送情報) に加えて、

- ・パケット送受信処理を行うコネクションを識別するためのコネクション識別子と、
- ・パケット送受信処理後に操作するタイマの種別を示すタイマ種別と、
- ・タイマのタイムアウト値と、
- ・タイマに対して起動・停止を行うかを示すタイマ操作とタイマ満了時の割り込み通知先を示すタイマ満了時通知先と、
- ・パケット送受信処理後に複数タイマを操作する場合に次のタイマ情報のアドレスを示す次タイマ情報アドレスと、

を含むタイマ情報で構成される。

20

30

【 0 0 4 3 】

なお、本実施形態において、タイマ 2 0 3 にタイムアウト値を設定しているが、タイマ種別によってタイムアウト値を固定する場合のタイマ情報は図 5 の (d) のようになる。例えば、ストリーム送信が要求されるコネクションの識別子が 0 であり、送信すべきパケットの先頭アドレスが 0 x 1 0 0 0 _ 0 0 0 0 、転送サイズが 1 0 2 4 バイトであったとする。この場合、送信処理用プロセッサ 2 0 1 は図 5 の (b) にあるパケット送受信要求を主記憶部 2 0 5 へ書き込む。すなわち、主記憶部 2 0 5 は、送信処理用プロセッサ 2 0 1 から、パケット送受信のための転送情報と、タイマ操作のためのタイマ情報とを含む送受信要求を保持する保持部として機能する。

40

【 0 0 4 4 】

続いて、送信処理用プロセッサ 2 0 1 はパケット送受信要求の先頭アドレスをデータ転送部 2 0 2 に通知する (ステップ S 1 0 3) 。第 2 実施形態では、データ転送部 2 0 2 は、まず、パケット送受信要求格納部 2 1 0 に格納し、データ転送部 2 0 2 がパケット送受信要求格納部 2 1 0 に格納されたアドレスを読み出す、という構成を有する。パケット送受信要求格納部 2 1 0 は、主記憶部 2 0 5 に保持された送受信要求の格納アドレスを登録する登録部として機能する。

【 0 0 4 5 】

50

データ転送部 202 は、パケット送受信要求格納部 210 にアドレスが格納されている場合、パケット送受信要求格納部 210 へ最初に格納されたアドレスを読み出す。すなわち、パケット送受信要求格納部 210 に存在するアドレスのうち、最も最初に格納されたアドレスから順に読み出される。そして、データ転送部 202 は、そのアドレスにしたがって主記憶部 205 からパケット送受信要求を読み出す（ステップ S111）。なお、読み出されたアドレスは、パケット送受信要求格納部 210 から消去される。消去されるタイミングとしては、パケット送受信要求格納部 210 からアドレスを読み出した後から次にパケット送受信要求格納部 210 からアドレスが読み出されるまでの間であればよい。本実施形態では、読み出しと同時に消去が行われる。第 1 実施形態では、プロセッサからデータ転送部へ直接に先頭アドレスが通知されたが、第 2 実施形態では、パケット送受信要求格納部 210 を介してプロセッサからデータ転送部へ先頭アドレスが通知される。そして、データ転送部 202 は、読み出したパケット送受信要求に従って、転送先 / 転送元アドレスで指定されたアドレスから転送サイズで指定されたデータ量のデータを読み出し、MAC / PHY 206 へ転送する（ステップ S112、S113）。MAC / PHY 206 へパケットを転送した後、データ転送部 202 は、タイマ情報に基づいてコネクション識別子とタイマ種別、タイムアウト値、タイマ満了時通知先をタイマ 203 に設定し、時間計測を起動する（ステップ S115）。このように、データ転送部 202 は、登録部としてのパケット送受信要求格納部 210 に登録された格納アドレスにしたがって主記憶部 205 から送受信要求を読み出し、データの転送処理及びタイマ操作を実行する。

【0046】

タイマ 203 は設定されたタイムアウト値までの時間を計測すると、タイマ満了時通知先に指定されたプロセッサ（ここでは送信処理用プロセッサ 201）に送信処理用割り込みコントローラ 204 経由で割り込みを通知する（ステップ S121～S122）。すなわち、上述したように、タイマ情報はタイマ満了時の通知先のプロセッサを示す通知先情報を含む。そして、タイマ 203 は、タイマの満了時に、対応するタイマ情報に含まれている通知先情報が示すプロセッサに対して当該満了の通知（割り込み通知）を行う。割り込みを通知された送信処理用プロセッサ 201 は、タイマ 203 から満了した時間計測に設定されたコネクション識別子とタイマ種別を読み出す。そして、送信処理用プロセッサ 201 は、読み出したコネクション識別子により処理を行うコネクションを、タイマ種別により行う処理を特定し、実行する（ステップ S104）。図 5 の（b）に示されるパケット送受信要求が処理された場合、送信処理用プロセッサ 201 はコネクション識別子 0 で示されるコネクションに対して、パケットの再転送処理を開始する。

【0047】

次に、データの受信処理について説明する。MAC / PHY 206 がネットワークからパケットを受信すると、割り込み信号をアサートする。受信処理用割り込みコントローラ 209 は、この割り込み信号を受信処理用プロセッサ 208 へ通知する。割り込みを通知された受信処理用プロセッサ 208 は、受信されたパケットが格納される MAC / PHY 206 のバッファから IP ヘッダ及び TCP ヘッダを読み出す。受信処理用プロセッサ 208 は IP ヘッダ内の送信元 / 受信先 IP アドレスからコネクション識別子を判断する。また、受信処理用プロセッサ 208 は、TCP ヘッダ内の ACK フラグが立ち、かつ、確認応答番号がその時点で最後に送信したデータのシーケンス番号 + 1 であった場合、図 5 の（c）にあるパケット送受信要求を主記憶部 205 へ書き込む。

【0048】

パケット送受信要求を主記憶部 205 へ書き込んだ後、受信処理用プロセッサ 208 はパケット送受信要求の主記憶部 205 上の先頭アドレスをパケット送受信要求格納部 210 へ格納する。

【0049】

データ転送部 202 は、パケット送受信要求格納部 210 にアドレスが格納されている場合、パケット送受信要求格納部 210 へ最初に格納されたアドレスを用いて、主記憶部 205 からパケット送受信要求を読み出す（ステップ S111）。そして、読み出したパ

ケット送受信要求に従って、MAC / PHY 206 のバッファから転送サイズで指定されたサイズ分のデータを読み出し、主記憶部 205 上の転送先 / 転送元アドレスで指定されたアドレスへ書き込む（ステップ S 112 , S 114 ）。主記憶部 205 への書き込みが終了した後、データ転送部 202 は、タイマ 203 のうちタイマ情報のコネクション識別子とタイマ種別で特定される時間計測を停止する（ステップ S 116 ）。

【0050】

以上のように、第2実施形態によれば、プロトコル処理を複数プロセッサで実行している場合においても、プロトコル処理と密接な関係を持つタイマの起動・停止の操作をプロセッサからオフロードするので、プロセッサの負荷を軽減することが可能となる。そして、プロセッサリソースを他プロトコル処理に割り当てることで、プロトコル処理の高速化が可能になる。なお、上記第2実施形態では、送信処理用プロセッサと受信処理用プロセッサを分けたが、送信処理が可能な複数のプロセッサを備えた構成であってもよいし、受信処理が可能な複数のプロセッサを備えた構成であってもよいことは明らかである。また、送受信が可能な複数のプロセッサを備えた構成であってもよいことも明らかである。

【0051】

また、第2実施形態のプロトコル処理装置は、プロセッサからのパケット送受信処理要求を一時的に格納するパケット送受信要求格納部 210 を有する。これにより、複数プロセッサの同時動作やプロトコル処理高速化によりパケット送受信要求が頻発した場合であっても、プロセッサの待ち時間を減少することが可能になる。

【0052】

以上説明したように、第1、第2実施形態によれば、プロトコル処理と関係を持つタイマの起動・停止をプロセッサからオフロードすることにより、プロセッサの負荷を軽減することが可能となる。そして、プロセッサリソースを他プロトコル処理に割り当てることで、プロトコル処理の高速化も可能になる。

【0053】

また、プロセッサからのパケット送受信処理要求を一時的に格納する手段を持つことにより、複数プロセッサの同時動作やプロトコル処理高速化によりパケット送受信要求が頻発した場合であっても、プロセッサの待ち時間を減少することが可能になる。

【0054】

[第3実施形態]

図6は、第3実施形態によるネットワークタイマ装置を含むプロトコル処理装置の構成例を示すブロック図である。

【0055】

図6において301はソフトウェアを実行するプロセッサである。302は送受信時のデータ移動などを行うデータ転送部である。303はネットワーク処理に必要な時間を計測するタイマである。304はデータ転送終了やタイマ満了通知などをプロセッサに通知する割り込みコントローラである。305はプロセッサの制御情報、データ、送受信データなどを保管する主記憶部である。306はネットワークと接続するためのデータリンク層および物理層のデバイス（以下、MAC / PHY）である。

【0056】

次に、第3実施形態によるプロトコル処理装置を説明する。第1、第2実施形態では、タイマの起動・停止の操作がプロセッサからオフロードされたが、第3実施形態ではタイマのタイムアップに伴う割り込み処理の一部がプロセッサからオフロードされる。

【0057】

第3実施形態のプロトコル処理装置の動作について図7のフローチャートを参照して説明する。プロセッサ301上で動作するユーザアプリケーションプログラムからTCPによるストリーム送信が要求されると、プロセッサ301は主記憶部305上にあるストリームデータからパケットを生成する（ステップS201）。パケット生成後、プロセッサ301はデータ転送部302に対するパケット送受信要求を主記憶部305へ書き込む（ステップS302）。ここで、パケット送受信要求は、例えば、

- ・データ転送部 302 にパケット送信処理を行わせるか、パケット受信処理を行わせるかを示す要求処理と、
- ・パケット送信処理要求時には送信するパケットの主記憶部 305 上の先頭アドレス、パケット受信処理要求時には受信するパケットを格納する主記憶部 305 上の先頭アドレスを示す転送先 / 転送元アドレスと、
- ・転送するデータ量を示す転送サイズとを含む転送情報を有する。第 3 実施形態のパケット送受信要求は、第 1、第 2 実施形態とは異なり、図 3 に示したような「タイマ情報」は持っていない。

【0058】

パケット送受信要求を主記憶部 305 へ書き込んだ後、プロセッサ 301 はパケット送受信要求の主記憶部 305 上の先頭アドレスをデータ転送部 302 へ通知する（ステップ S203）。

【0059】

データ転送部 302 は、通知されたアドレスを用いて、主記憶部 305 からパケット送受信要求を読み出す（ステップ S221）。そして、データ転送部 302 は、読み出したパケット送受信要求から、当該要求が送信要求か受信要求かを判定する（ステップ S222）。送信要求の場合、データ転送部 302 は、読み出したパケット送受信要求に従って転送先 / 転送元アドレスで指定されたアドレスから転送サイズで指定されたデータ量のデータを読み出し、MAC / PHY 306 へ転送する（ステップ S223）。受信要求の場合、データ転送部 302 は、MAC / PHY 306 のバッファから転送サイズで指定されたサイズ分のデータを読み出し、主記憶部 305 上の転送先 / 転送元アドレスで指定されたアドレスへ書き込む（ステップ S222、S224）。このようにデータ転送部 302 は、プロセッサ 301 からの送受信要求に応じて、メモリとデータリンク層との間のデータの転送処理を実行する第 1 の転送部として機能する。

【0060】

プロセッサ 301 は、タイマ 303 のタイマレジスタを設定し、当該プロトコル処理に必要なタイマを起動する。

【0061】

第 3 実施形態のタイマ 303 は、ネットワーク処理に特化したネットワークタイマ装置である。タイマ 303 は、内部に複数のタイマレジスタを持ち、それぞれのタイマレジスタには、カウントするタイマ値とともに、通信先を特定するためのコネクション識別子とタイマ種別を含むネットワークタイマ情報が登録される。プロセッサ 301 が、タイマ値とともにネットワークタイマ情報をレジスタに登録すると、タイマ 303 はタイマ動作を開始する。なお、本実施形態の場合、タイマの起動（Setup）、中断（Abandon）などのタイマ操作情報を有している。したがって、ステップ S204 において、送受信処理に関連したタイマの起動や中断を示すタイマ操作情報により、プロセッサ 301 は、タイマの起動や中断を制御できる。

【0062】

図 8 は、第 3 実施形態のタイマ 303 のタイマレジスタに登録される情報を示す図である。複数のタイマが起動される場合は、複数のタイマのそれぞれに対してタイマレジスタが設けられる。図 8 において、有効フラグフィールド 501 には有効フラグが格納される。この有効フラグの有効状態は計時中であることを示し、有効フラグの無効状態は計時停止状態或いは計時終了状態を示す。タイマ値フィールド 502 には計時するティック値が書き込まれる。本実施形態のタイマ 303 は、あるイベントを受けてからイベントを起すまでの時間、即ち経過時間を取り扱っている。

【0063】

コネクション識別子フィールド 503 にはコネクション識別子が格納される。タイマ 303 はネットワーク処理に特化しているタイマであるため、タイマに連動してコネクション識別子の情報を他モジュールの制御のために使用する。なお、コネクションとは通信相手先を特定するための情報であり、TCP / IP では自 IP アドレス、相手 IP アドレス

10

20

30

40

50

、自ポート番号、相手ポート番号の4つの情報を示し、コネクション識別子からコネクションを特定することが出来る。タイマ種別フィールド504は、それぞれのコネクションで使用されるタイマの種別を特定する情報が格納される。従来例でも説明したようにTCP/IPではコネクション毎に複数のタイマが用いられている。例えば、コネクション確立タイマ、再送タイマ、遅延ACKタイマ、持続タイマ、TIME_WAITタイマ、FIN_WAITタイマ、FIN_WAIT2タイマなど様々なタイプのタイマが使用される。タイマ種別フィールド504には、これらのいずれのタイマであるかを示すタイマ種別が格納されている。いずれかのタイマが満了になると、タイマ303はコネクション識別子とタイマ種別をデータ転送部302へ出力し、タイマレジスタのフラグを無効にする。

【0064】

10

中断 (Abandon) タイマ操作を受け付けた場合は計時の即時中断を意味し、コネクション識別子とタイマ種別によって特定されたタイマレジスタの有効フラグを無効状態にすることで計時を中断する。

【0065】

タイマ303は、上記のようにしてプロセッサからの指示により設定されたタイマの計数を開始する (ステップS211)。そして、タイマの計数が満了になった場合 (ステップS212)、対応するタイマレジスタに登録されているネットワークタイマ情報 (コネクション識別子とタイマ種別) がデータ転送部302に送出される (ステップS213)。

【0066】

20

コネクション識別子とタイマ種別を受け取ったデータ転送部302は、受け取った情報を基に、データ転送処理を実行したり、プロセッサ301への割り込み要求を行ったりする (ステップS225、S226)。このように、データ転送部302は、起動されたタイマを管理する。そして、データ転送部302は、タイマ満了となったタイマに対応するネットワークタイマ情報に基づいて送信すべきパケットを決定し、これを送信するべくメモリとデータリンク層との間のデータの転送処理を実行する第2の転送部として機能する。

【0067】

例えば、ステップS226においてデータ転送部302がデータ転送を行う場合は、コネクション識別子とタイマ種別からディスクリプタアドレスを特定する。ディスクリプタとはデータ転送を行うために必要な情報であり、事前に主記憶部305などに記憶されている。ディスクリプタには、例えば、データの送信の手順 (送信すべきアドレス範囲や送信先など) が記述されている。データ転送に先だってデータ転送部302は、主記憶部305から特定されたディスクリプタを読み込み、読み込んだディスクリプタに従ってデータ転送を行う。すなわち、主記憶部305上などにタイマ種別あるいはネットワークタイマ情報に対応したディスクリプタを予め用意しておく。そして、データ転送部302は、タイマの満了時にタイマが出力したコネクション識別子とタイマ種別に対応するディスクリプタを主記憶部305から読み出し、読み出したディスクリプタによりデータ転送処理を実行する。このようにすることで、データ転送部302は、タイマ満了に伴って自動的にデータ転送を開始することが出来る。また、タイマ種別によっては、後述するように、従来どおりにプロセッサに割り込み要求をかけることが望ましい場合もある。その場合は、データ転送部302が割り込み出力と共にタイマ303からの出力であることをプロセッサ301に通知するようにする (ステップS226)。すなわち、データ転送部302は、タイマ種別に基づいてプロセッサ301に割り込み要求をするか否かを判定する機能を有する。

30

40

【0068】

次にコンピュータ間の通信として広く使われているTCP/IPで用いられるネットワークタイマを上記構成のプロトコル処理装置で実現した場合の動作について説明する。

【0069】

図9は一般的なTCP/IP通信を行っている通信機器間のパケットの送受信状態をシ

50

ーケンス図で表したものである。図中、*s r c*とあるのはコネクションを能動的に確立しようとするアクティブ側の通信装置（以下、*s r c*側）を示す。また、*d s t*とあるのは*s r c*側からのコネクション確立要求に応じて通信を行うパッシブ側の通信装置（以下、*d s t*側）であることを示す。TCP/IPでは3wayHandshakeと呼ばれている仕組みによってコネクションを確立している。まず、*s r c*側がTCPパケット中の*s y n*フラグを立てたパケットを*d s t*側に送信すると共にコネクション確立タイマを起動する。送信したパケットの*s r c*側のIPアドレスとポート番号、*d s t*側のIPアドレスとポート番号の4つの変数はコネクション識別子に対応している。このコネクション識別子はユニークな番号であり、同じ識別子が別の4変数の組に割り当てられることはないので、コネクション識別子から*s r c*側と*d s t*側のIPアドレス、ポート番号を求めることも容易である。また、コネクション識別子にコネクションの通信状態を示す変数を対応させている。

10

【0070】

このコネクション確立タイマはコネクションを確立する時に1度使用するタイマでその後には使用することはない。また、このコネクション確立タイマが満了時に行われる処理も、コネクション確立のために用意した変数領域を解放する程度の処理である。このため、コネクション確立タイマの自動化に伴うメリットは少ない。よって、データ転送部302は、タイマ303より送信されたタイマレジスタからタイマ種別がコネクション確立タイマであることを検知すると、割り込みコントローラ304へプロセッサ301に対する割り込み要求であることを通知する。

20

【0071】

一方、コネクション確立タイマが満了にならずに*d s t*側からSYN+ACKの2つのフラグがついたTCPパケットを*s r c*側が受け取ると、TCPのステートはSYN_SENTと呼ばれる状態に移行する。そして、*s r c*側がSYN+ACKに対するACKパケットを送出したところで*s r c*側はEstablished、即ちコネクションが確立したことを表す。*d s t*側は*s r c*側のACKパケットを受け取ったところでコネクションが確立したことを知る。コネクションが確立されるとどちら側からパケットを送信しても問題ないが、図9では*s r c*側がパケットを送信した場合を示している。

【0072】

TCP/IPでは常にデータの送信を確認しながら通信を行うため、プロセッサ301は、パケット送信と同時に再送タイマを起動する。プロセッサ301は、再送タイマを起動するためにタイマ303に対してタイマレジスタにタイマ値とコネクション識別子、タイマ種別（＝再送タイマ）を登録し、タイマレジスタの有効フラグを立てる。送信したパケットに対するACKパケットが*d s t*側から返ってきた場合、受信確認できたものと見なされ再送タイマは停止される。即ち、当該再送タイマのタイマレジスタのフラグが無効となる。これは、上述した中断（Abandon）の操作情報を用いることで実現できる。パケットの送信が再開されると共に再送タイマは再び起動される。このときのタイマ値は、プロセッサ301により再設定されることになる。実際のTCP/IPでは通信状況などから再送タイマの値を動的に変化させる場合があるが、本実施形態においてはその動きは省略する。このように再送タイマはパケットを送信する毎に起動され、ACKパケットを受け取る毎に停止される。

30

40

【0073】

*s r c*側でACKパケットを受信しないと再送タイマが満了となる。タイマ303は、満了になった再送タイマのコネクション識別子とタイマ種別（この場合は再送タイマ）をデータ転送部302に通知する。この場合、すでに送信した送信データについてACKを受けていないことから、データ転送部302は、当該送信データを保持したままである。したがって、データ転送部302は、当該データ転送に使用したディスクリプタを再利用してACKポイントからのパケット再送信を開始する（ステップS226）。

【0074】

一方、*d s t*側ではパケットを受信する度にTCP/IP通信装置では遅延ACKタイ

50

マが起動される。遅延 A C K タイマは受信パケット毎に A C K パケットを送信すると通信トラヒックが増大してしまうために、ある程度受信パケットをまとめた A C K パケットを送信するように A C K パケットの送信を抑圧する為のタイマである。遅延 A C K タイマの満了にともなって、d s t 側から s r c 側に通信確認パケットである A C K パケットが送信される。この A C K パケットは現時点での通信状況を相手側に通知するためのパケットであるため、データ送信要求が別にある場合は除き、ペイロードを持たないことが多い。そこで主記憶部 3 0 5 上に予め A C K パケットを作成しておき、データ転送部 3 0 2 は、ステップ S 2 2 6 において、コネクション識別子の情報 (I P アドレス、ポート番号、シーケンス番号など) と合成して A C K パケットを生成する。そして、データ転送部 3 0 2 は、生成した A C K パケットを M A C / P H Y 3 0 6 に転送する。このようにすることで遅延 A C K タイマの満了に伴って、プロセッサ 3 0 1 を介さずに、自動的に A C K パケットが送信されることになる。こうして、上記第 2 の転送部としてのデータ転送部 3 0 2 は、ネットワークタイマ情報に含まれるタイマの種別が T C P / I P 通信における A C K タイマを示す場合、自端末が受信したシーケンス番号を相手端末に通知する A C K パケットを送出する。

10

【 0 0 7 5 】

次に持続タイマについて説明する。送信側の通信能力と比較して受信側の通信能力が低く受信側の通信の為のメモリを使い切った場合や、アプリケーションのデータ処理が受信バッファが蓄積するスピードに追いつかない場合に、受信 Window が ' 0 ' であることを送信側に通知することがある。これは ' 0 ' window 公告と呼ばれるもので送信先の受信バッファが無くなったため公告された側は送信を停止する。

20

【 0 0 7 6 】

しかし、永久に止めておくわけにもいけないので、ある持続タイマが満了となる間隔をもってプロービングパケットと呼ばれる探査用のパケットを d s t 側に送信する。プロービングパケットを送りつけることによって d s t 側は A C K パケットを送出し、受信 Window の最新値を送信側に通知する。プロービングパケットは事前に用意することが出来るパケットで受信側の性能が送信側と比較して劣っている場合に良く生成されるパケットであるため自動化することで他のコネクションに対して影響を与えずにプロービングパケットを送信することが出来る。すなわち、タイマ満了したタイマのタイマ種別が持続タイマであった場合は、データ転送部 3 0 2 はプロービングパケットを自動的に d s t 側に送信する。

30

【 0 0 7 7 】

以上、第 3 実施形態のプロトコル処理装置における再送タイマ、遅延 A C K タイマ、持続タイマの動作に対してタイマ満了時の処理について説明した。他の T C P タイマ、TIME_WAIT、FIN_WAIT、FIN_WAIT2 などのタイマはコネクション確立タイマと同様にコネクションのクローズ時に一度起動されるだけの場合が多い。そのため、頻繁にコネクションをオープン、クローズするシステムでない限りは割り込みによるプロセッサの処理で十分な処理速度を得ることが出来る。

【 0 0 7 8 】

以上説明したように、第 3 実施形態によれば、コネクション型通信において使用されるタイマに、コネクション識別子とタイマ種別からなるネットワークタイマ情報が付加される。このため、タイマ満了の通知を受けたデータ転送部がネットワークタイマ情報からディスクリプタを読み込みに行き、データ転送を実行することが出来る。そのため大幅にプロセッサ負荷を削減することができ、削減したプロセッサリソースをプロトコル処理に回すことで処理能力の向上を見込むことができる。

40

【 0 0 7 9 】

なお、第 3 実施形態に示したタイマ満了時の処理手順を第 1 実施形態または第 2 実施形態に適用可能であることはいうまでもない。

【 0 0 8 0 】

以上において、本発明が特定の実施形態に関して図示され、かつ、説明されたが、さら

50

に他の修正および改善が可能であることは言うまでもない。

本発明は前述の実施形態に示された特定の形式に限定されるものではなく、かつ添付の特許請求の範囲において本発明の精神および範囲から離れることのない全ての修正をカバーすることを考えていることが理解されるべきである。

【図面の簡単な説明】

【0081】

【図1】第1実施形態に係るシステム構成を示す図である。

【図2】第1実施形態によるプロトコル処理を説明するフローチャートである。

【図3】第1実施形態に係るパケット送受信要求を示す図である。

【図4】第2実施形態に係るシステム構成を示す図である。

【図5】第2実施形態に係るパケット送受信要求を示す図である。

【図6】第3実施形態に係るシステム構成を示す図である。

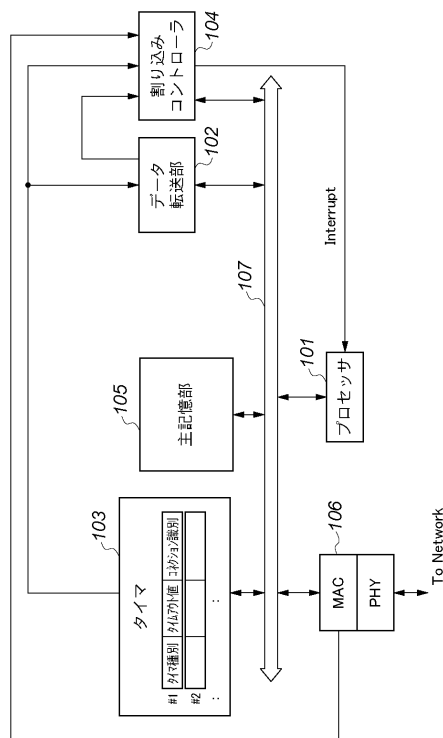
【図7】第3実施形態によるプロトコル処理を説明するフローチャートである。

【図8】第3実施形態に係るタイマレジスタの構成例を示す図である。

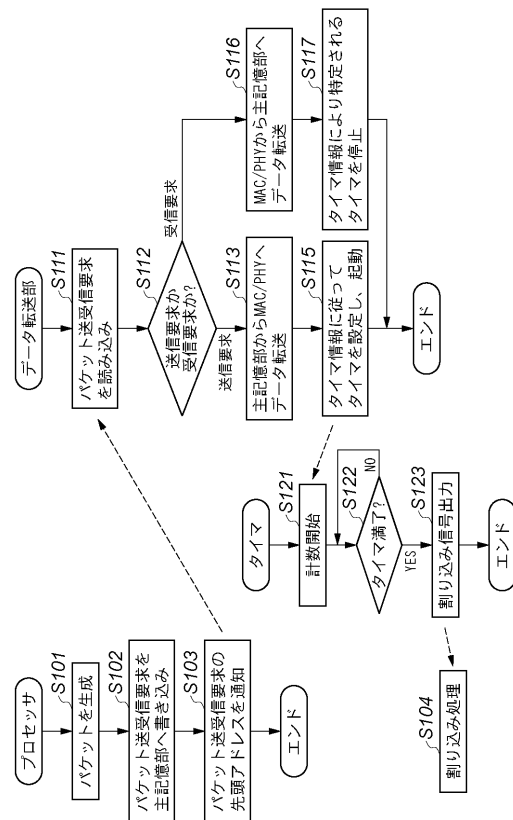
【図9】第3実施形態に係るシーケンス図の例である。

10

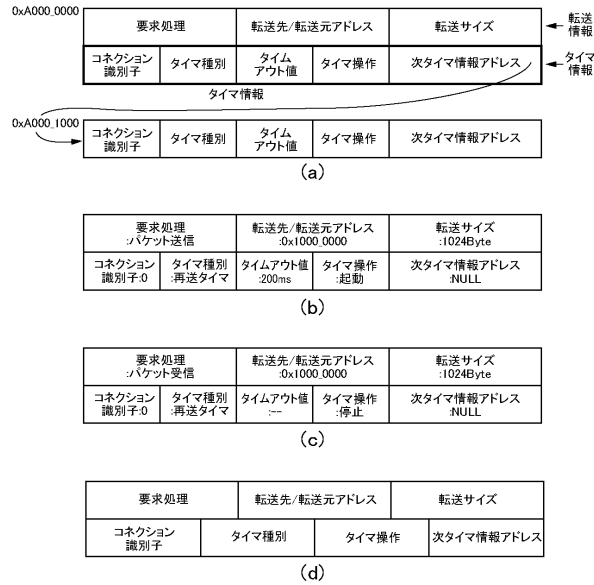
【図1】



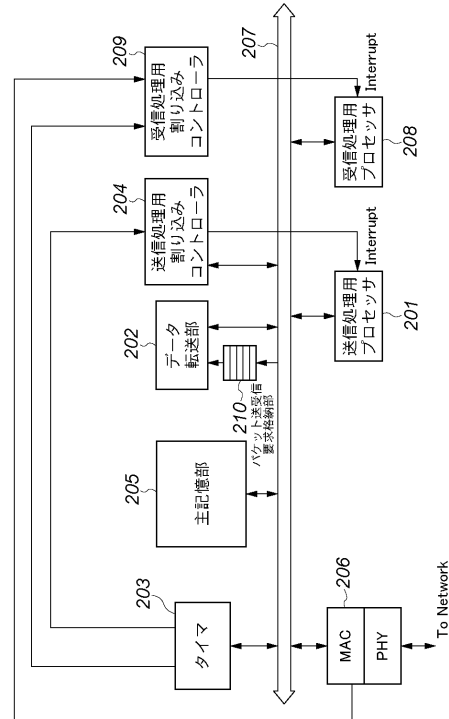
【図2】



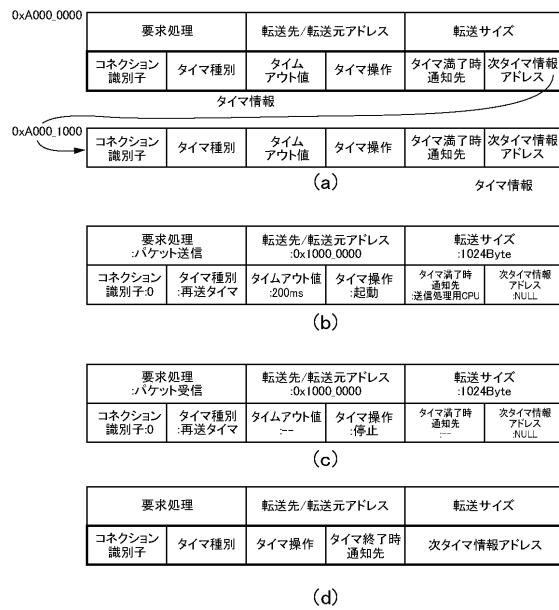
【図 3】



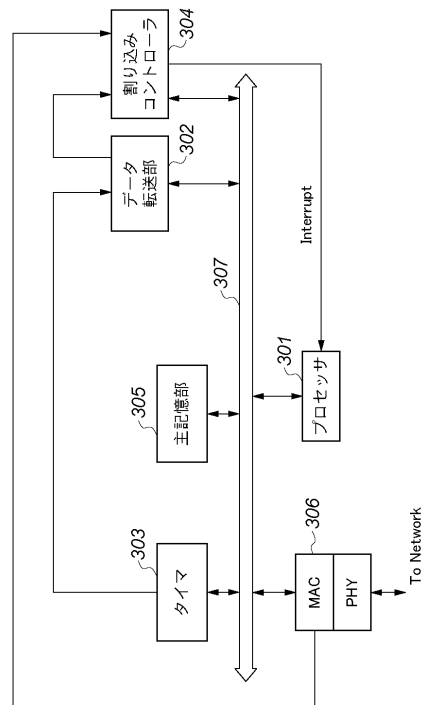
【図 4】



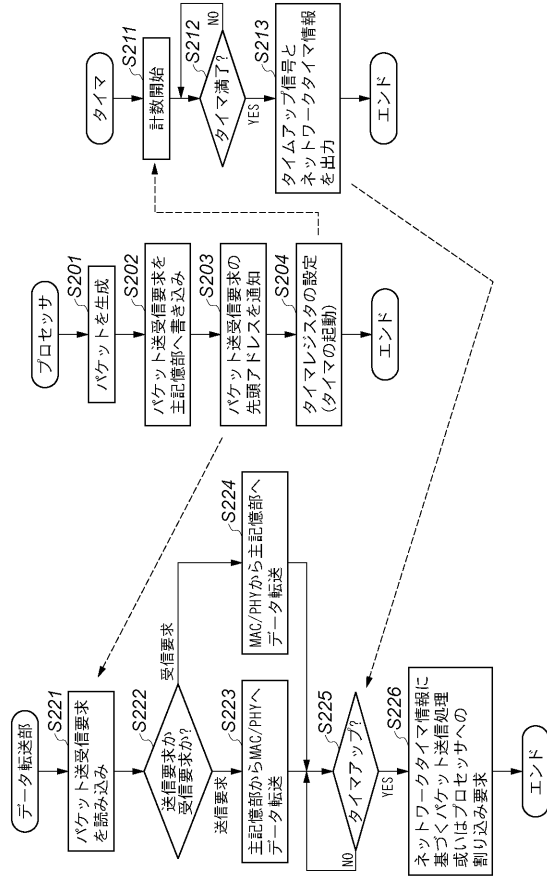
【図 5】



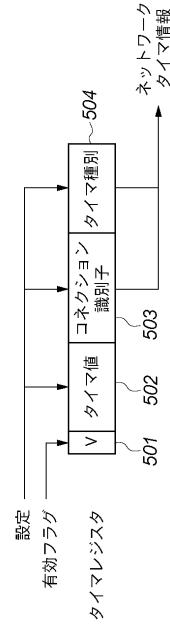
【図 6】



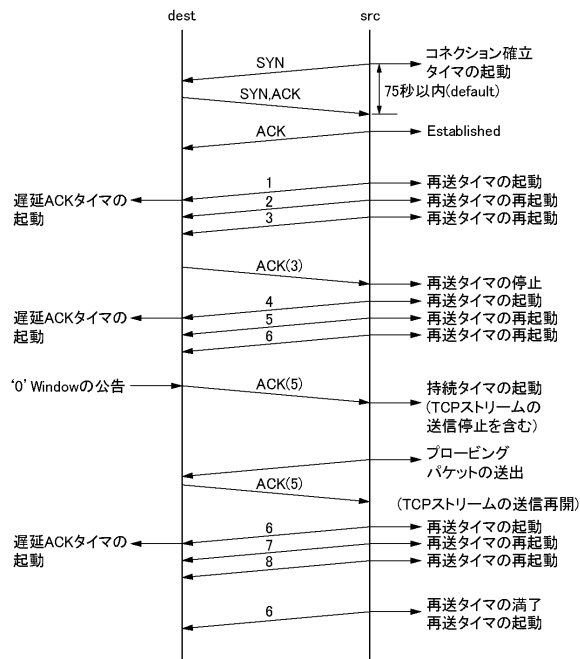
【図 7】



【図 8】



【図 9】



フロントページの続き

- (72)発明者 白石 大介
東京都大田区下丸子3丁目30番2号 キヤノン株式会社内
(72)発明者 森村 和彦
東京都大田区下丸子3丁目30番2号 キヤノン株式会社内

審査官 寺谷 大亮

- (56)参考文献 特開平03-273350(JP,A)
特開昭63-074351(JP,A)
特開平07-123128(JP,A)
特開昭63-042253(JP,A)

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