(19)日本国特許庁(JP)　(12)特　許　公　報(B2)　(11)特許番号

特許第5179218号
(P5179218)

(45)発行日　平成25年4月10日(2013.4.10)　(24)登録日　平成25年1月18日(2013.1.18)

(51)Int.Cl.			F I		
H04L	12/851	(2013.01)	H04L	12/56	200E
H04L	12/927	(2013.01)	G06F	13/38	340A
G06F	13/38	(2006.01)	H04L	12/56	100A
H04L	12/707	(2013.01)			

請求項の数 7　(全 15 頁)

(21)出願番号　特願2008-37692(P2008-37692)
(22)出願日　平成20年2月19日(2008.2.19)
(65)公開番号　特開2009-200611(P2009-200611A)
(43)公開日　平成21年9月3日(2009.9.3)
審査請求日　平成23年1月26日(2011.1.26)

(73)特許権者　000004226
日本電信電話株式会社
東京都千代田区大手町二丁目３番１号
(73)特許権者　504176911
国立大学法人大阪大学
大阪府吹田市山田丘１番１号
(74)代理人　100086232
弁理士　小林　博通
(74)代理人　100104938
弁理士　鵜澤　英久
(74)代理人　100140361
弁理士　山口　幸二
(74)代理人　100096459
弁理士　橋本　剛

最終頁に続く

(54)【発明の名称】ｉＳＣＳＩセッションのＴＣＰコネクション数制御方法、ｉＳＣＳＩホスト装置、およびｉＳＣＳＩイニシエータの構成プログラム

(57)【特許請求の範囲】
【請求項１】
ｉＳＣＳＩホスト装置とｉＳＣＳＩデバイス装置とがｉＳＣＳＩセッションによって接続されるときに
ｉＳＣＳＩホスト装置に構築されたイニシエータ機能をもってＴＣＰコネクション数を制御する方法であって、
ｉＳＣＳＩセッションをとおして送受信する一定のデータ量と当該データ量を送受信するのに要した時間との比を参照して、次に送受信する一定のデータ量を送受信するために使用するＴＣＰコネクション数を決定する第１ステップと、
前記決定に基づき次に送受信するために使用するＴＣＰコネクション数を増減制御する第２ステップと、
を有することを特徴とするｉＳＣＳＩセッションのＴＣＰコネクション数の制御方法。
【請求項２】
前記第１ステップは、ＴＣＰコネクション毎に一定のデータ量に振り分けて前記ｉＳＣＳＩデバイス装置と送受信を行うステップと、
前記ｉＳＣＳＩデバイス装置側から送信されたＴＣＰコネクション毎の送受信完了通知に基づき前記送受信に要した時間を求めるステップと、
を有することを特徴とする請求項１記載のｉＳＣＳＩセッションのＴＣＰコネクション数の制御方法。
【請求項３】

前記第１ステップは、ＴＣＰコネクション数を増加させながらデータ送受信のスループット値を計測し、該スループット値が前回値よりも低下したときにＴＣＰコネクション数の数値巾を抽出するステップと、
前記数値巾に黄金探索法を適用してＴＣＰコネクション数を絞り込んで、スループット値が最大となるＴＣＰコネクション数の最適値を決定するステップと、
を有することを特徴とする請求項１または２いずれか記載のｉＳＣＳＩセッションのＴＣＰコネクション数の制御方法。
【請求項４】
ｉＳＣＳＩデバイス装置とｉＳＣＳＩセッションによって接続されるイニシエータ機能が構築されたｉＳＣＳＩホスト装置であって、
ｉＳＣＳＩセッションをとおして送受信する一定のデータ量と当該データ量を送受信するのに要した時間との比を参照して、次に送受信する一定のデータ量を送受信するために使用するＴＣＰコネクション数を決定する手段と、
前記決定に基づき次に送受信するために使用するＴＣＰコネクション数を増減制御する手段と、
して機能することを特徴とするｉＳＣＳＩホスト装置。
【請求項５】
ＴＣＰコネクション毎に一定のデータ量に振り分けて前記ｉＳＣＳＩデバイス装置と送受信を行う手段と、
前記ｉＳＣＳＩデバイス装置側から送信されたＴＣＰコネクション毎の送受信完了通知に基づき前記送受信に要した時間を求める手段と、
して機能することを特徴とする請求項４記載のｉＳＣＳＩホスト装置。
【請求項６】
ＴＣＰコネクション数を増加させながらデータ送受信のスループット値を計測し、該スループット値が前回値よりも低下したときにＴＣＰコネクション数の数値巾を抽出する手段と、
前記数値巾に黄金探索法を適用してＴＣＰコネクション数を絞り込んで、スループット値が最大となるＴＣＰコネクション数の最適値を決定する手段と、
して機能することを特徴とする請求項４または５いずれか記載のｉＳＣＳＩホスト装置。
【請求項７】
請求項４～６記載のｉＳＣＳＩホスト装置として、コンピュータを機能させることを特徴とするｉＳＣＳＩイニシエータの構成プログラム。
【発明の詳細な説明】
【技術分野】
【０００１】
本発明は、ファイルサーバなどのｉＳＣＳＩホスト装置とストレージなどのｉＳＣＳＩデバイス装置とをＩＰネットワークを介して接続するｉＳＣＳＩの技術に関し、特にスループットを向上させる技術に関する。
【背景技術】
【０００２】
（１）ｉＳＣＳＩ（Ｉｎｔｅｒｎｅｔ　ＳＣＳＩ）
単体、もしくは複数のハードディスクドライブ装置の集合体、あるいは専用の制御部で複数のハードディスクドライブを制御するディスクアレイストレージ装置等から構成されるストレージ装置が実用化されている。
【０００３】
ストレージ装置（デバイス装置）とホスト装置（サーバ，ＰＣ）間を接続するインタフェース技術としてはファイバチャネル（ＦＣ），インフィニバンド（ＩｎｆｉｎｉＢａｎｄ），ＳＣＳＩ（Ｓｍａｌｌ　Ｃｏｍｐｕｔｅｒ　Ｓｙｓｔｅｍｓ　Ｉｎｔｅｒｆａｃｅ）などがある。

【０００４】
ＳＣＳＩインタフェースは近距離間を経済的に接続する用途として優れ、既に広く普及している。当該インタフェースでは、クライアントに相当する機能をイニシエータ，サーバに相当する機能をターゲットと呼び、イニシエータとターゲット間で，ＳＣＳＩコマンド（ＣＤＢ：ＣＯＭＭＡＮＤ Ｄｅｓｃｒｉｐｔｏｒ Ｂｌｏｃｋ）を交換する。
【０００５】
ｉＳＣＳＩ（ｉｎｔｅｒｎｅｔ ＳＣＳＩ）技術は、非特許文献１に示すように、２００４年にネットワークプロトコルであるＴＣＰ／ＩＰの上位プロトコルとして、ＩＥＴＦ（Ｔｈｅ Ｉｎｔｅｒｎｅｔ Ｅｎｇｉｎｅｅｒｉｎｇ Ｔａｓｋ Ｆｏｒｃｅ）で標準化された（ＲＦＣ３７２０）。このｉＳＣＳＩ接続は、ファイバチャネルなどの他のネットワークストレージプロトコルと違ってコンピュータにイーサネット（登録商標）インタフェースがあれば使用することができる。
【０００６】
現在、ｉＳＣＳＩ技術は、ギガビット以上の回線品質、特にデータセンター内、データセンター間や大規模ＬＡＮシステム等の基幹回線で利用されている。非特許文献１によれば、ｉＳＣＳＩ技術はＴＣＰコネクションを複数利用する方法を規定しているが、オプション機能のため必ずしも実装されているわけではない．
（２）ｉＳＣＳＩ高速化技術
一方で，広域・広帯域のＩＰネットワーク上で，ネットワーク遅延によるｉＳＣＳＩスループットの低下を防ぐ方法として，ＴＣＰコネクションの複数利用の有効性は知られている（非特許文献２．３）。また、ＴＣＰコネクションの複数利用の実現手段としてマルチリンクを用いる手法の非特許文献４やトランスポート層のプロトコルを変更する非特許文献５なども提案されている。
【０００７】
例えば、マルチリンクを用いる手法に関する非特許文献４では、ＶＰＮのマルチホーミング機能を用いて，並列ＴＣＰコネクションを複数の経路で確立させ，スループット向上を実現する手法が提案されている。
【０００８】
また、トランスポート層のプロトコルを変更する手法に関する非特許文献５では，複数のＬＡＮポートを利用して物理的にリンクを多重化し，ＴＣＰコネクションを異なる経路で確立することによりスループットの向上を図っている。
【０００９】
しかしながら，非特許文献４では、マルチリンクを利用できる環境に制限があり、非特許文献５ではマルチベンダー機器への接続が困難であるという問題を有している。
【非特許文献１】ＲＦＣ３７０ “Ｉｎｔｅｒｎｅｔ Ｓｍａｌｌ Ｃｏｍｐｕｔｅｒ Ｓｙｓｔｅｍｓ Ｉｎｔｅｒ－ｆａｃｅ （ｉＳＣＳＩ）”，Ａｐｒ ２００４
【非特許文献２】Ｑ．Ｋ．Ｙａｎｇ，“Ｏｎ ｐｅｒｆｏｒｍａｃｅ ｏｆ ｐａｒａｌｌｅｌ ｉＳＣＳＩ ｐｒｏｔｏｃｏｌ ｆｏｒｎｅｔｗｏｒｋｅｄ ｓｔｏｒａｇｅ ｓｙｓｔｅｍｓ，” ｉｎ Ｐｒｏｃｅｅｄｉｎｇｓ ｏｆ ｔｈｅ ２０ｔｈ Ｉｎｔｅｒｎａｔｉｏｎａｌ Ｃｏｎｆｅｒｅｎｃｅ ｏｎ Ａｄｖａｎｃｅｄ Ｉｎｆｏｒｍａｔｉｏｎ Ｎｅｔｗｏｒｋｉｎｇ ａｎｄ Ａｐｐｌｉｃａｔｉｏｎｓ （ＡＩＮＡ ２００６），ｖｏｌ．１，ｐｐ．６２９－６３６，Ａｐｒ．２００６．
【非特許文献３】Ｇ．Ｍｏｔｗａｎｉ ａｎｄ Ｋ．Ｇｏｐｉｎａｔｈ，“Ｅｖａｌｕａｔｉｏｎ ｏｆ ａｄｖａｎｃｅｄ ＴＣＰ ｓｔａｃｋｓ ｉｎ ｔｈｅ ｉＳＣＳＩ ｅｎｖｉｒｏｎｍｅｎｔ ｕｓｉｎｇ ｓｉｍｕｌａｔｉｏｎ ｍｏｄｅｌ，” Ｐｒｏｃｅｅｄｉｎｇｓ ｏｆ ｔｈｅ ２２ｎｄ ＩＥＥＥ／１３ｔｈ ＮＡＳＡ Ｇｏｄｄａｒｄ Ｃｏｎｆｅｒｅｎｃｅ ｏｎＭａｓｓ Ｓｔｏｒａｇｅ Ｓｙｓｔｅｍｓ ａｎｄ Ｔｅｃｈｎｏｌｏｇｉｅｓ （ＭＳＳＴ≡０５）}，ｐｐ．２１０－２１７，２００５．
【非特許文献４】Ｂ．Ｋ．Ｋａｎｃｈｅｒｉａ，Ｇ．Ｍ．Ｎａｒａｙａｎ，ａｎｄ

K. Gopinath, “Performance evaluation of Multiple TCP connection in iSCSI,” in Proceedings of the 24th IEEE Conference on Mass Storage Systems and Technologies, pp. 239－244, IEEE Computer Society, Sept. 2007
【非特許文献５】千島 望, 山口 実靖, 小口 正人, “ｉＳＣＳＩストレージにおけるＶＰＮ複数経路アクセス時の性能とＴＣＰパラメータ解析,” 電子情報通信学会第１８回データ工学ワークショップ ＤＥＷＳ２００７, Ｆｅｂ. ２００７.
【非特許文献６】L. Qiu, Y. Zhang, and S. Keshav, “On individual and aggregate TCP Performance,” in Proceedings of Internet Conference on Network Protocols, pp. 203－212, Oct. 1999.
【非特許文献７】T. Ito, H. Ohsaki, and M. Imase, “On parameter tuning of datatransfer protocol GridFTP in wide－area Grid computing,” in Proceedings of Second International Workshop on Networks for Grid Applications（GridNets ２００５）}, pp. 415－421, Oct. ２００５.
【非特許文献８】T. Ito, H. Ohsaki, and M. Imase, “GridFTP－APT: Automatic parallelism tuning mechanism for data transfer protocol GridFTP,” in Proceedings of 6th IEEE International Symposium on Cluster Computing and the Grid （CCGrid２００６）}, pp. 454－461, May. ２００６.

【発明の開示】
【発明が解決しようとする課題】
【００１０】
しかしながら、広域・広帯域のＩＰネットワークにおいてｉＳＣＳＩのスループットが低下する問題は広く知られているものの、この問題を従来は効果的に解決できなかった。
【００１１】
すなわち、広域・広帯域のＩＰネットワークを前提としたｉＳＣＳＩ接続では，スループット向上のため、利用する回線の帯域や遅延情報に基づき予めｉＳＣＳＩ／ＴＣＰなどのパラメータをチューニングし、あるいはサーバとストレージ間にプロトコル変換機を設置してスループットを改善するなど、運用設定が煩雑で、かつ経済的な運用が困難なおそれがあった。
【００１２】
また、このような改善は、固定的な情報に基づく設定であるため，帯域の変動に追随できず、効率的なデータ転送が難しく、またベストエフォート型の経済性の高い回線を利用できないおそれもあった。
【００１３】
そこで本発明は、広域・広帯域のＩＰネットワークにおいてｉＳＣＳＩのスループットが低下する問題を、（１）既存のストレージ装置を利用しつつ、（２）ネットワークの遅延，利用帯域などの利用環境を限定することなく、（３）経済的に解決することを課題としている。
【課題を解決するための手段】
【００１４】

ｉＳＣＳＩ技術は、一つのｉＳＣＳＩセッション内に複数のＴＣＰコネクションを確立して、複数のコマンドを並列に発行し、各ＴＣＰコネクションで同時にデータの送受信を行う機能を有しているものの、前記従来例にはＴＣＰコネクション数をどのように決定するかは記述されていない。

【００１５】

本発明は、並列ＴＣＰコネクションのスループットが多重度に関して上に凸の関数になる性質に着目して創作された技術的思想であり、ＴＣＰコネクションの多重度をスループットに対して最適化させて前記課題を解決している。

【００１６】

具体的には、請求項１記載の発明は、ｉＳＣＳＩホスト装置とｉＳＣＳＩデバイス装置とがｉＳＣＳＩセッションによって接続されるときにｉＳＣＳＩホスト装置に構築されたイニシエータ機能をもってＴＣＰコネクション数を制御する方法であって、ｉＳＣＳＩセッションをとおして送受信する一定のデータ量と当該データ量を送受信するのに要した時間との比を参照して、次に送受信する一定のデータ量を送受信するために使用するＴＣＰコネクション数を決定する第１ステップと、前記決定に基づき次に送受信するために使用するＴＣＰコネクション数を増減制御する第２ステップと、を有することを特徴としている。

【００１７】

請求項２記載の発明は、前記第１ステップは、ＴＣＰコネクション毎に一定のデータ量に振り分けて前記ｉＳＣＳＩデバイス装置と送受信を行うステップと、前記ｉＳＣＳＩデバイス装置側から送信されたＴＣＰコネクション毎の送受信完了通知に基づき前記送受信に要した時間を求めるステップと、を有することを特徴としている。

【００１８】

請求項３記載の発明は、前記第１ステップは、ＴＣＰコネクション数を増加させながらデータ送受信のスループット値を計測し、該スループット値が前回値よりも低下したときにＴＣＰコネクション数の数値巾を抽出するステップと、前記数値巾に黄金探索法を適用してＴＣＰコネクション数を絞り込んで、スループット値が最大となるＴＣＰコネクション数の最適値を決定するステップと、を有することを特徴としている。

【００１９】

請求項４記載の発明は、ｉＳＣＳＩデバイス装置とｉＳＣＳＩセッションによって接続されるイニシエータ機能が構築されたｉＳＣＳＩホスト装置であって、ｉＳＣＳＩセッションをとおして送受信する一定のデータ量と当該データ量を送受信するのに要した時間との比を参照して、次に送受信する一定のデータ量を送受信するために使用するＴＣＰコネクション数を決定する手段と、前記決定に基づき次に送受信するために使用するＴＣＰコネクション数を増減制御する手段と、して機能することを特徴としている。

【００２０】

請求項５記載の発明は、ＴＣＰコネクション毎に一定のデータ量に振り分けて前記ｉＳＣＳＩデバイス装置と送受信を行う手段と、前記ｉＳＣＳＩデバイス装置側から送信されたＴＣＰコネクション毎の送受信完了通知に基づき前記送受信に要した時間を求める手段と、して機能することを特徴としている。

【００２１】

請求項６記載の発明は、ＴＣＰコネクション数を増加させながらデータ送受信のスループット値を計測し、該スループット値が前回値よりも低下したときにＴＣＰコネクション数の数値巾を抽出する手段と、前記数値巾に黄金探索法を適用してＴＣＰコネクション数を絞り込んで、スループット値が最大となるＴＣＰコネクション数の最適値を決定する手段と、して機能することを特徴としている。

【００２２】

請求項７記載の発明は、請求項４～６記載のｉＳＣＳＩホスト装置として、コンピュータを機能させることを特徴としている。

【発明の効果】

【００２３】
請求項１～７記載の発明によれば、ＴＣＰコネクション数の最適値が自動的に算出され、当該値でＴＣＰコネクション数が運用される。したがって、運用設置が簡略化され，帯域変動にも追随でき、これにより与えられた回線で常に効率的なデータ転送が実現可能となり、経済性も向上する。
【００２４】
特に、請求項７記載の発明によれば、既存のｉＳＣＳＩホスト装置にプログラムをインストールすれば足り、システム構築・更新の経済性が一層向上する。
【発明を実施するための最良の形態】
【００２５】
以下、図１～図５に基づき本発明の実施形態を説明する。図１は、ｉＳＣＳＩイニシエータ１とｉＳＣＳＩターゲット２とがＩＰ（ｉｎｔｅｒｎｅｔ　ｐｒｏｔｏｃｏｌ）ネットワークを介して、ｉＳＣＳＩ接続される基本的な形態を示している。
【００２６】
前記ｉＳＣＳＩイニシエータ１は、ｉＳＣＳＩホスト装置を構成するファイルサーバに実装される一方、前記ｉＳＣＳＩターゲット２は、ｉＳＣＳＩデバイス装置を構成するディスクストレージ装置に実装されている。ここではｉＳＣＳＩ接続（ｉＳＣＳＩ　ｓｅｓｓｉｏｎ）の中にＴＣＰコネクションを複数形成することができ、ＴＣＰコネクション数を「Ｎ」と表記する。
【００２７】
前記ｉＳＣＳＩイニシエータ１の機能は、ファイルサーバの処理部（ＣＰＵ「Ｃｅｎｔｒａｌ　Ｐｒｏｃｅｓｓｉｎｇ　Ｕｎｉｔ」など）が、イニシエータ構成プログラムのプログラムコードを読み込んで、通信デバイス（例えばイーサネットインタフェースなど）を通じて実現構築されている。
【００２８】
なお、前記ファイルサーバは、コンピュータにより構成され、処理データなどを一時記憶可能なメモリ（ＲＡＭ），キーボードやマウスなどの入力デバイス，モニタなどの表示部などを有している。
【００２９】
図２は、前記ｉＳＣＳＩイニシエータ１の機能にＴＣＰコネクション数Ｎの制御機能（図２中のｉＳＣＳＩ－ＡＰＴ「Ａｕｔｏｍａｔｉｃ　Ｐａｒａｌｌｅｌｉｓｍ　Ｔｕｎｉｎｇ」３）が追加された構成を示している。この制御機能は、前記ｉＳＣＳＩイニシエータ１の機能と同様に、ファイルサーバの前記処理部などが前記イニシエータ構成プログラムのプログラムコードを読み込んで実現構築されている。
【００３０】
ここではｉＳＣＳＩイニシエータ１を実装したファイルサーバが、ｉＳＣＳＩターゲット２を実装したストレージ装置に書き込む（転送する）データを、ＴＣＰコネクション数Ｎが変動してもほぼ一定のサイズのデータ（チャンク）に分割しながら転送する。
【００３１】
このときチャンクは，「ｉＳＣＳＩ　ＰＤＵ」サイズの整数倍となるため，ＴＣコネクション数「Ｎ」に対して，各ＴＣＰコネクションにほぼ均等に「ｎ」個を振り分けるとすると以下の式１が成立する。
【００３２】
【数１】

$$\text{チャンクサイズ} = iSCSI\ PDU\ \text{数} \times n \times N \qquad \text{式(1)}$$

【００３３】
以下、図２に基づき動作例Ｓ１～Ｓ７を説明する。

（１）Ｓ１～Ｓ７の動作例

Ｓ１の処理：まず、ｉＳＣＳＩイニシエータ１の配置されるファイルサーバ内の上位層、例えばファイルシステムをとおして、ｉＳＣＳＩ層に書き込み要求のあった転送データを、ＴＣＰコネクション数Ｎに依存しない均一サイズのデータ（チャンク）に分割する。チャンクの大きさは、小さすぎるとＴＣＰ／ＩＰプロトコルの特性上、スループットを正確に反映できないため、スループットをほぼ正確に算出できる充分な大きさとする。

【００３４】

Ｓ２の処理：本発明の処理機能であるｉＳＣＳＩ－ＡＰＴ３が、ｉＳＣＳＩターゲット２（ストレージ）に対してｉＳＣＳＩ接続するために、初期のＴＣＰコネクションを確立し，「ｉＳＣＳＩ　ｌｏｇｉｎ」認証を行う（通常ここで接続されるＴＣＰコネクションは１本である。）。

【００３５】

Ｓ３の処理：Ｓ１の処理で生成したチャンクを、ｉＳＣＳＩのデータ転送単位である「ｉＳＣＳＩ　ＰＤＵ」サイズに分割し、分割した「ｉＳＣＳＩ　ＰＤＵ」をＴＣＰコネクションＮ本に対して、ほぼ均等「ｎ個」に振り分ける。

【００３６】

【表１】

N	n:iSCSI PDU数／TCPコネクション数N
1	840
2	420
3	280
4	210
5	168
6	140
7	120
8	105

Table 1:TCPコネクション数Nとチャンク分割の例

【００３７】

表１は、「Ｔａｂｅｌｅ　１」のＴＣＰコネクション数Ｎとチャンク分割の例を示している。ここでは「チャンク（ｎ）：ｉＳＣＳＩ　ＰＤＵ数／ＴＣＰセッション」を例に、ＴＣＰコネクション「Ｎ」の増加に伴い、１本のＴＣＰコネクションあたり「１／Ｎ」に減少することが示されている。

【００３８】

Ｓ４の処理：振り分けられた各ＴＣＰコネクション毎に転送する「ｉＳＣＳＩ　ＰＤＵ」（約ｎ個）を各々のＴＣＰコネクションをとおして転送する。このときｉＳＣＳＩ－ＡＰＴ３は、図３及び図４に示すように、すべてのＴＣＰコネクションの中で一番最初に転送を始めた時刻を記録する。

【００３９】

Ｓ５の処理：ＳＣＳＩターゲット２が、図３に示すように、各ＴＣＰコネクション毎に「ｉＳＣＳＩ　ＰＤＵ」受信完了通知、即ちＲ２Ｔ（Ｒｅａｄｙ　ｔｏ　ｔｒａｎｓｆｅｒ）を逐次送信する。

【００４０】

Ｓ６の処理：ｉＳＣＳＩ－ＡＰＴ３は、すべてのＴＣＰコネクションの中で、最後にＲ２Ｔを受信した時刻が記録し、図４に示すように、当該チャンクの転送時間から当該チャンク転送でのスループットＧ（Ｎ）を算出する。直後に，次のチャンク転送に設定するＴＣＰコネクション数Ｎを後述の内蔵アルゴリズム処理に基づいて決定する。

【００４１】

Ｓ７の処理：次のチャンク転送に設定するＴＣＰコネクション数Ｎに変更するため，ＴＣＰコネクションの増減制御を行う。

（２）内蔵アルゴリズム処理

ここではＳ６の処理で行うＴＣＰコネクション数Ｎの決定例を、図５に基づき説明する。ここでは並列ＴＣＰコネクションのスループット（チャンク／転送時間Ｔ）が、コネクション数Ｎの多重度に関して上に凸の関数となる性質が示されている。

【００４２】

【表２】

TCPコネクション数Nの初期値N_0	1
制御パラメータk	2

Table 2:図5に適用するパラメータ

【００４３】

表２は、図５の処理例に使用される「Ｔａｂｌｅ　２」を示している。この「Ｔａｂｌｅ　２」は、ＴＣＰコネクション数Ｎの初期値Ｎｏが「１」に設定され、制御パラメータｋが「２」に設定されている。

【００４４】

ＴＣＰコネクション数Ｎを最適化する処理は、以下のステージ１．２で構成されている。このステージ１には図５に示す「Ａ～Ｅ」の推移が該当し、ステージ２には図５に示す「Ｅ～Ｈ」までの推移が該当する。

（１）ステージ１

ステージ１は、ＴＣＰコネクション数Ｎを初期値$Ｎ_0$から、制御パラメータｋを乗じて増加させ、最適値Ｎｏｐｔの含まれるコネクション数（図５の横軸）のおおまかな巾（以降，ブラケットと呼ぶ）を探索する。この状態を式（２）に示す。

【００４５】

【数２】

初期値　N_0

$$N \leftarrow N_0 \qquad \text{式(2)}$$

【００４６】

そのうえで、Ｓ３～Ｓ６の処理で記載した方法に基づき、チャンク転送のスループットＧ（Ｎ）を算出する。算出したスループットＧ（Ｎ）が前回の算出値との間で以下の式（３）の関係となるまで、

【００４７】

【数３】

$$G(N) < G(N_{-1}) \qquad N_{-1}：\text{前回の}\ N\ \text{値} \qquad \text{式(3)}$$

【００４８】

以下の式（４）に基づき、

【００４９】

【数４】

$$N \leftarrow k \times N \quad (k > 1) \qquad \text{式(4)}$$

【００５０】
ＴＣＰコネクション数Ｎを増加させながら、チャンクのスループットＧ（Ｎ）を計測し，式（３）の評価（前回値との比較）を行う。式（３）が満たされた（ＴＲＵＥ）の場合（前回値よりスループットが低下した場合）、最初のブラケットには，式（５）に示すように、過去３回分の履歴にある３つのＮ値を組として定め、ステージ２へ移行する。

【００５１】
【数５】

$$(\text{前々回の } N \text{ 値, 前回の } N \text{ 値, 今回の } N \text{ 値}) = (N_{-2},\ N_{-1},\ N) = (l, m, r) \qquad \text{式(5)}$$

【００５２】
図５中、「Ａ」は初期値であり、「Ｔａｂｌｅ　２」に記載のとおり、「Ｎ＝N_0＝１」とする。次いで「Ｔａｂｌｅ　２」に記載のパラメータｋを用いて、ＴＣＰコネクションを２倍（図５中の「Ｂ」）にする。さらに次のチャンク転送では、Ｎ＝４（図５中の「Ｃ」），Ｎ＝８（図５中の「Ｄ」）と増やした段階で式（３）の評価が、式（６）となり、最初（１ｓｔ）のブラケット（２，４，８）を得る（図５中の「Ｅ」）。

【００５３】
【数６】

$$\text{TRUE} \leftarrow \{G(8) < G(4)\} \qquad \text{式(6)}$$

【００５４】
（２）ステージ２
ステージ２では、ステージ１で算出した最適値Ｎｏｐｔ（チャンクのスループットの唯一の極大点を示すＮの値）の含まれる最初のブラケットの巾（図５の横軸）に黄金探索法を適用して絞り込み、最適値Ｎｏｐｔを抽出する。これは主に非特許文献８の手法が用いられる。この非特許文献８は、ｉＳＣＳＩではなく、ＧｒｉｄＦＴＰという別の技術に黄金探索法が用いられている。以下、ステージ２の絞込みを具体的に説明する。

【００５５】
まず、最初に式（７）と式（８）の黄金比（Ｇｏｌｄｅｎ　Ｒａｓｉｏ）とを用いて、ステージ１で抽出した最初のブラケットから新たなＴＣＰコネクション数となる「Ｎ」を算出する。

【００５６】
【数７】

$$\text{TRUE or FALSE} \leftarrow \{G(N) > G(m)\} \qquad \text{式(7)}$$

【００５７】

【数８】

$$\text{Golden Rasio}: \ \nu = \frac{3-\sqrt{5}}{2} \approx 0.382 \qquad \text{式(8)}$$

【００５８】
ブラケット（ｌ，ｍ，ｒ）｝に対して、新たなＴＣＰコネクション数　Ｎを算出するため、式（９）を適用する（小数点以下は切り上げ）。
【００５９】
【数９】

$$N \leftarrow \begin{cases} l+(m-l)\nu & if \ \ m-l > r-m \\ m+(r-m)\nu & \text{上記以外} \end{cases} \qquad \text{式(9)}$$

【００６０】
次に，Ｓ３～Ｓ６の処理で記載した方法に基づき、チャンク転送のスループットＧ（Ｎ）を算出する。式（７）の関係が成り立つ（ＴＲＵＥ）場合、式（１０）により、新たなブラケットを導入する。
【００６１】
【数１０】

$$(l,m,r) \leftarrow \begin{cases} (m,N,r) & if \ m < N; \ \ \text{TRUE} \\ (l,N,m) & \text{上記以外}; \ \ \text{TRUE} \end{cases} \qquad \text{式(10)}$$

【００６２】
式（７）の関係が成り立たない場合（ＦＡＬＳＥ）は、式（１１）により、新たなブラケットを導出し、新たに導出したブラケットを用いて，式（９）から新たなＴＣＰコネクション数Ｎを算出する。
【００６３】
【数１１】

$$(l,m,r) \leftarrow \begin{cases} (l,m,N) & if \ m < N; \ \ \text{FALSE} \\ (N,m,r) & \text{上記以外}; \ \ \text{FALSE} \end{cases} \qquad \text{式(11)}$$

【００６４】
以上の過程を繰り返し，（ｌ，ｍ，ｒ）が連続する整数となった場合、式（１２）とする。
【００６５】
【数１２】

$$N_{opt} \leftarrow m \qquad \text{式(12)}$$

【００６６】
実際に，ステージ１で抽出した１ｓｔブラケット（２，４，８）を式（９）に当てはめると式（１３）となる。
【００６７】

【数１３】

$$N = 6 \leftarrow 5.528 \leftarrow (2, 4, 8) \qquad \text{式(13)}$$

【００６８】
Ｎ＝６で式（７）を図５の結果（縦軸）で評価すると、式（１４）が成立する。
【００６９】
【数１４】

$$\text{TRUE} \leftarrow \{G(6) > G(4)\} \qquad \text{式(14)}$$

【００７０】
よって，式（１１）で式（１５）の２ｎｄブラケットを構成する（図５の「Ｅ」）。
【数１５】

$$\text{2nd ブラケット}: (4, 6, 8) \leftarrow (2, 4, 8) \qquad \text{式(15)}$$

【００７１】
再び式（９）で「Ｎ」を算出すると、式（１６）となる。
【００７２】
【数１６】

$$N = 7 \leftarrow 6.764 \leftarrow (4, 6, 8) \qquad \text{式(16)}$$

【００７３】
Ｎ＝７で式（７）を評価（図５参照）し、式（１７）の結果から式（１０）で式（１８）の３ｒｄブラケットを構成する（図５の「Ｆ」～「Ｇ」）。
【００７４】
【数１７】

$$\text{FALSE} \leftarrow \{G(7) < G(6)\} \qquad \text{式(17)}$$

【００７５】
【数１８】

$$\text{3rd ブラケット}: (4, 6, 7) \leftarrow (4, 6, 8) \qquad \text{式(18)}$$

【００７６】
再び式（９）でＮを算出すると式（１９）となる。
【００７７】

【数１９】

$$N = 5 \leftarrow 4.382 \leftarrow (4, 6, 7) \qquad \text{式(19)}$$

【００７８】
Ｎ＝５で式（７）を評価（図５参照）し、式（２０）の結果から式（１１）で式（２１）の４ｒｄブラケットを構成する。
【００７９】
【数２０】

$$\text{FALSE} \leftarrow \{G(5) < G(6)\} \qquad \text{式(20)}$$

【００８０】
【数２１】

$$\text{4th ブラケット}: (5, 6, 7) \leftarrow (4, 6, 7) \qquad \text{式(21)}$$

【００８１】
ここでブラケット内の整数がすべて連続となったため、ステージ２を終了する。スループットを最大にするＴＣＰコネクション数の最適値Ｎｏｐｔとして、式（２２）に示す最後のブラケットの中央の数字である「６」が抽出できる。
【００８２】
【数２２】

$$Nopt = 6 \leftarrow (5, 6, 7) \qquad \text{式(22)}$$

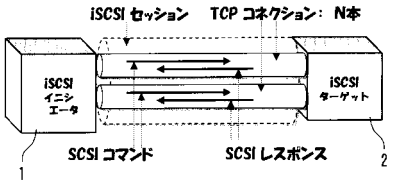
【００８３】
整理すると、図５の例では、ＴＣＰコネクション数の初期値N_0から最適値Ｎｏｐｔに移行するまで、式（２３）に示す遷移を行う。
【００８４】
【数２３】

$$N_0 = 1 \rightarrow 2 \rightarrow 4 \rightarrow 8 \rightarrow 6 \rightarrow 7 \rightarrow 5 \rightarrow N_{opt} = 6 \qquad \text{式(23)}$$

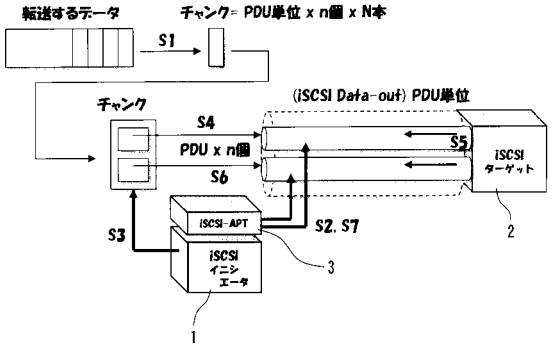
【００８５】
このように並列ＴＣＰコネクションのスループット（チャンク／転送時間Ｔ）が、コネクションＮの多重度に関して上に凸の関数になるという性質（図５の「Ａ」～「Ｈ」）を利用して、ＴＣＰコネクションの多重度をスループットに対して最適化させている（スループットが最大となるＴＣＰコネクション数を探索し、当該数のＴＣＰコネクション数にまで増減制御している。）。
【００８６】
すなわち、ｉＳＣＳＩイニシエータ１に組み込まれたＴＣＰコネクション数Ｎを自動調整するｉＳＣＳＩ－ＡＰＴ３の制御機能によって、予め使用するネットワークの遅延特性や帯域等の情報が無い場合でも、ｉＳＣＳＩ－ＡＰＴ３の処理機能を定期的に動作させることで、動作時のネットワークに最適化したＴＣＰコネクション数Ｎｏｐｔに自動追従さ

せている。
【００８７】
これによりｉＳＣＳＩ接続されるファイルサーバとストレージとの間のスループットの高速化を実現させている。この結果、ファイルサーバなどの運用設置を大幅に簡略でき、帯域変動にも追従可能となることから、与えられた回線で常に効率的なデータ転送が実現可能となる。
【００８８】
また、ファイルサーバの実装的にはｉＳＣＳＩイニシエータの修正で済むので、ＴＣＰコネクションの多重化オプションをサポートする市販のターゲット装置の利用でき、この意味で経済的なシステム構築・更新が可能である。
【００８９】
なお、本発明は、上記実施形態に限定されるものではなく、各請求項に記載した範囲において各種の変形を行うことが可能である。例えばデータ転送（送信）だけではなく、データの読み込み（受信）に関しても同様の効果が期待できる。
【００９０】
また、前記ｉＳＣＳＩ－ＡＴＰ３の制御機能を内蔵したｉＳＣＳＩイニシエータ１として、コンピュータを機能させるプログラムとして構築することもできる。
【００９１】
この場合には、コンピュータの処理部（例えばＣＰＵなど）がプログラムコードを読み出し、通信デバイスを通じてステージ１．２の処理が実行される。
【００９２】
したがって、既存のｉＳＣＳＩシステムにおいてファイルサーバに前記プログラムをインストールするだけで、本実施形態に係るｉＳＣＳＩイニシエータ機能を有するファイルサーバが構築でき、この点でシステム構築・更新の経済性が一層向上する。
【００９３】
このプログラムコードは、例えばＣＤ－ＲＯＭ，ＤＶＤ－ＲＯＭ，ＣＤ－Ｒ，ＣＤ－ＲＷ，ＤＶＤ－Ｒ，ＤＶＤ－ＲＷ，ＭＯ，ＨＤＤなどの記録媒体に格納される。また、前記プログラムを、インターネットサイトからダウンロードしてコンピュータに提供してもよい。
【符号の説明】
【００９４】
１…ｉＳＣＳＩイニシエータ
２…ｉＳＣＳＩターゲット（ｉＳＣＳＩデバイス）
３…ｉＳＣＳＩ－ＳＰＴ
Ｎ…ＴＣＰコネクション数
Ｇ（Ｎ）…スループット値
【図面の簡単な説明】
【００９５】
【図１】ＴＣＰコネクション数を複数利用するｉＳＣＳＩ接続の概略図。
【図２】本発明の実施形態に係るＴＣＰコネクション数制御機能（ｉＳＣＳＩ－ＡＰＴ）をｉＳＣＳＩイニシエータのｉＳＣＳＩ接続の概略図。
【図３】１つのＴＣＰコネクションでの「ｉＳＣＳＩ　ＰＤＵ」転送のシーケンス例を示し、チャンク転送時間算出のための時刻定義の概略図。
【図４】複数のＴＣＰコネクションを用いたチャンク転送での転送時間の定義を示す概略図。
【図５】チャンク転送のスループットの観測から最適なＴＣＰコネクション数を導出する処理過程例を示す概略図。

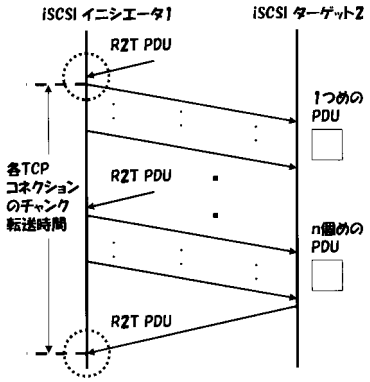
【図１】


iSCSI セッション
TCP コネクション：N本
iSCSI イニシエータ
iSCSI ターゲット
SCSI コマンド
SCSI レスポンス
1
2

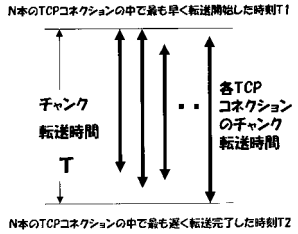
【図２】


転送するデータ
チャンク= PDU単位 x n個 x N本
S1
チャンク
(iSCSI Data-out) PDU単位
S4
PDU x n個
S5
S6
iSCSI ターゲット
iSCSI-APT
S2, S7
S3
iSCSI イニシエータ
1
2
3

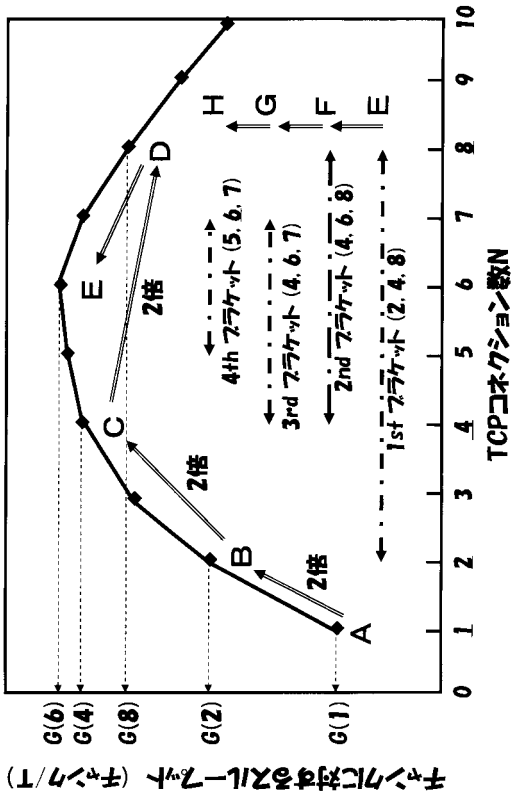
【図３】


iSCSI イニシエータ1
iSCSI ターゲット2
R2T PDU
1つめの PDU
各TCP コネクション のチャンク 転送時間
R2T PDU
n個め の PDU
R2T PDU

【図４】


N本のTCPコネクションの中で最も早く転送開始した時刻T1
チャンク 転送時間 T
各TCP コネクション のチャンク 転送時間
N本のTCPコネクションの中で最も遅く転送完了した時刻T2

【図５】


チャンクに対するスループット（チャンク/T）
TCPコネクション数N
0
1
2
3
4
5
6
7
8
9
10
G(6)
G(4)
G(8)
G(2)
G(1)
A
B
C
D
E
F
G
H
2倍
2倍
2倍
1st フラケット (2, 4, 8)
2nd フラケット (4, 6, 8)
3rd フラケット (4, 6, 7)
4th フラケット (5, 6, 7)

フロントページの続き

(72)発明者　野本　義弘
東京都千代田区大手町二丁目３番１号　日本電信電話株式会社内
(72)発明者　大崎　博之
大阪府吹田市山田丘１番１号　国立大学法人大阪大学内
(72)発明者　今瀬　真
大阪府吹田市山田丘１番１号　国立大学法人大阪大学内
(72)発明者　井上　史斗
大阪府吹田市山田丘１番１号　国立大学法人大阪大学内

審査官　矢頭　尚之

(56)参考文献　特開２００６－２７０３０３（ＪＰ，Ａ）
武田　裕子　ほか，ｉＳＣＳＩ　ｏｖｅｒ　ＶＰＮ環境における複数経路アクセス適応制御手法の提案と評価，電子情報通信学会　第１８回データ工学ワークショップ論文集 DEWS2007 E7-3，日本，電子情報通信学会データ工学研究専門委員会，２００７年　６月　１日，[online] DEWS2007 HIROSHIMA
伊藤　建志　ほか，ＧｒｉｄＦＴＰ－ＡＰＴ：データ転送プロトコルＧｒｉｄＦＴＰの並列ＴＣＰコネクション数調整機構，電子情報通信学会技術研究報告　IN2005-113，日本，社団法人電子情報通信学会，２００５年１２月　８日，第105巻　第472号，p.19-24
武田　裕子　ほか，ＶＰＮ上のｉＳＣＳＩ環境における複数経路アクセス適応制御手法の提案と評価，日本データベース学会Ｌｅｔｔｅｒｓ，日本，日本データベース学会，２００７年　６月，第6巻　No.1，p.129-132
藤原　啓成　ほか，広域IP網を介したiSCSI通信におけるプロトコルチューニングの一検討，第６８回（平成１８年）全国大会　講演論文集（１）　アーキテクチャ ソフトウェア科学・工学，日本，情報処理学会，２００６年　３月　７日，5A-5，p.1-55

(58)調査した分野(Int.Cl.，ＤＢ名)
Ｈ０４Ｌ　１２／５６
Ｇ０６Ｆ　１３／００
Ｈ０４Ｌ　１３／００