

1D-4

階層構造を持つ通信プロトコルの
並列処理方式に関する一検討

加藤 聡彦 藤長 昌彦 鈴木 健二

KDD研究所

1. はじめに

近年、100Mbpsを越える超高速ネットワークの導入に伴い、それらに接続されるコンピュータにおいて、伝送速度に対応する高性能な通信処理の実現が不可欠となっている。このためには、プロトコル処理用VLSIなどの専用ハードウェアの開発や、マルチプロセッサ構成の高性能ワークステーションを用いたプロトコルソフトウェアの実装などが有効である。これらの方法において高いスループットを達成するためには、通信プロトコルの処理全体を、独立な実行可能な部分に分割し、これらを複数のハードウェア・モジュールまたはプロセッサを用いて並列に処理することが必要となる。プロトコルの並列処理については、各種方式が検討されているが[1-3]、本稿では、階層構造を持つ通信プロトコルを対象として、複数の方式を組み合わせ、高い並列性を実現するとともに、分割されたユーザデータの受信時の処理遅延を防ぐ並列処理方式について述べる。

2. 通信プロトコルの並列処理の基本方針

階層構造を持つ通信プロトコルに対して、並列性の高い処理を実現するために、次のような基本方針を採用した。

(1) 階層構造を持つプロトコルに対しては、次のような並列処理方式が検討されている。

- 各層のプロトコル処理の並列化する方式
- 送受信するPDU(プロトコルデータ単位)毎に、複数層の一連のプロトコル処理を並列化する方法[1]
- 1つの層の中で、複数のPDUの作成処理や、1つのPDUの独立なパラメータ値の設定/解析処理などを並列化する方法[2]

高い並列性を実現するために、対象となるプロトコルに対して、これらの複数の方式を組み合わせ、並列実行可能なモジュールを抽出する。

(2) 層毎のプロトコル処理の並列化には、1つのPDUに対して、各層の処理をパイプライン的に実行する方法と、1つのPDUに対する複数層の処理を同時に行なう方法[3]とが考えられる。しかし、1つのPDUに対する各層のプロトコル処理は、送信時は上位から、受信時は下位から順に実行される必要がある場合が多い。例えば、PDUの受信時には下位のPDUのヘッダを解析して、それが正しい順序で到着したことを確認した後初めて上位層での処理が開始できる。そこでここでは、各層の処理をパイプライン的に実行する方法を用いることとする。

(3) 分割されたユーザデータを持つPDUの受信処理において、その層でユーザデータの組立て処理を完了した後に上位にPDUを渡す方法[3]を用いると、上位層での処理の開始に遅延が生ずる。これを防ぐために、分割されたユーザデータを受信した時点で、上位に通知する方法を採用する。

ただし、本稿では、OSIで使用される用語を用いるが、本方式はOSIプロトコルに制限されるものではない。

3. データ送信時における並列処理

前述の基本方針に基づき、データの送信時において、並列に実行可能なモジュールを抽出すると、次のようになる(図1参照)。ここでは、複数の層でのデータコピーを避けるために、各層からアクセス可能な共有メモリ上にPDUおよびサービスプリミティブを作成する[4]。

(N)層における1つの(N)-DATA reqプリミティブの処理が、並列に実行可能となる。この処理においては、コネクションを識別し、その状態に従って、(N)-PDUを作成し、フロー制御に基づいて(N)-PDUを送信し、そのコネクションの状態を更新するという手順をシークエンシャルに行なう。さらに、OSIの第6,7層のような複雑な構造の(N)-PDUに対しては、(N)-PDU作成の処理が並列に実

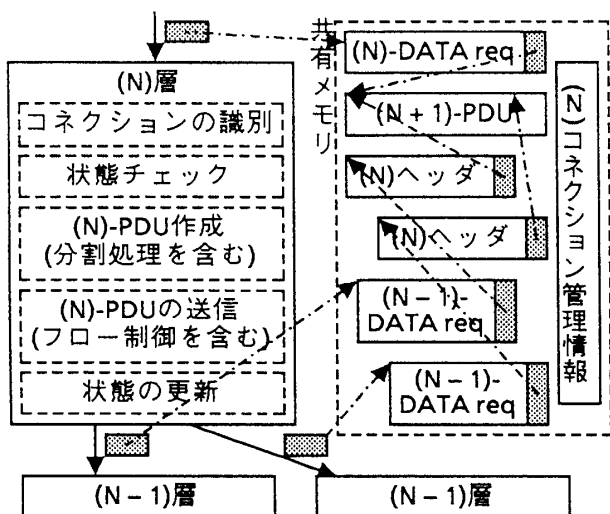


図1 データ送信時の並列処理手順

行されることも考えられる。次に、(N-1)層による(N-1)-DATA reqプリミティブの処理が、要求された個数だけ、並列に実行可能となる。ただし、この場合の複数の(N-1)-DATA reqの処理は、同一のコネクション上での処理であるため、各層のコネクション管理情報を用いて送信されるPDUの順序性を保つ必要がある。

4. データ受信時における並列処理

4.1 分割されたPDUの解析方法

1つの層において、ユーザデータが分割された場合は、分割された最初のPDUの運ぶユーザデータに、上位層のヘッダが存在する。このため、分割されたユーザデータを運ぶPDUを受信した場合、ユーザデータ自身と、それが上位PDUの先頭か/途中か/最終かの指定とを、上位層に通知することにより、図2に示すように、ユーザデータを分割したままで、上位に処理させることが可能となる。

- ① 分割の先頭の(N-1)-PDUであるため、(N)層は先頭に(N)ヘッダがあると解釈する。そのヘッダが分割の先頭の(N)-PDUを示すため、(N+1)層も先頭に(N+1)ヘッダがあると解釈する。
- ② 分割の最後の(N-1)-PDUであるため、(N)層は(N)-PDUの最終部分が受信されたと解釈する。
- ③ (N)層は先頭の(N)ヘッダを処理し、分割の最後の(N)-PDUであることを識別し、それを(N+1)層へ通知する。

4.2 データ受信時の並列処理の実現

受信側においても、送信側と同様に、各層に

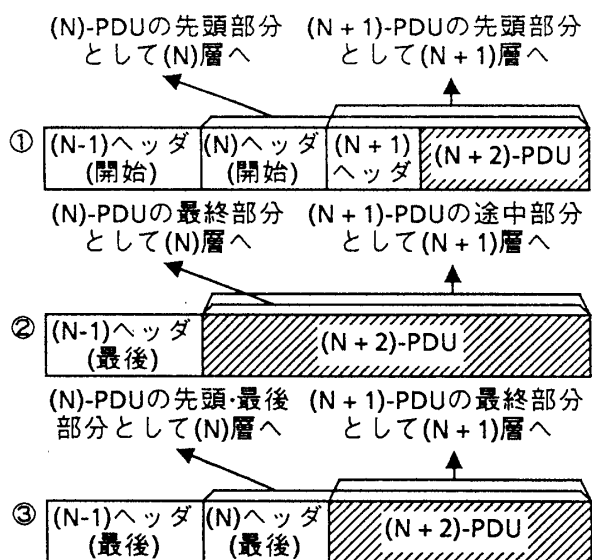


図2 分割されたPDUの解析手順

おけるデータ指示プリミティブの処理が並列に実行可能なモジュールとなる。この処理は、下位コネクションの識別、PDUの解析とPDUのコネクションへの関連付け、状態チェック、上位へのデータ指示プリミティブの通知(ユーザデータが分割された場合も含む)、そのコネクションの状態の更新などから構成される。この内、複雑な構造のPDU解析の処理は、必要に応じて並列化されることがある。

5. おわりに

本稿では、階層構造を持つプロトコルを対象とする並列処理方式について述べた。本方式は、層毎のプロトコル処理、PDUの送受信、層内の独立な処理に対する並列化を組み合わせ、高い並列性を実現するとともに、分割されたユーザデータをそのまま上位に通知することにより、上位層の処理の遅延を防ぐことを特徴とする。最後に日頃御指導いただくKDD研究所浦野所長、眞家次長に感謝する。

参考文献

- [1]: G. Neufeld, et.al., "Parallel Host Interface for an ATM Network," IEEE Network, July 1993.
- [2]: M. Bilgic, W. Wu and B. Sarikaya, "VASN1 - A High Speed Protocol Encoder / Decoder," Bilkent University Technical Report: CIS-9124, Sept. 1991.
- [3]: 石坂, 土田, 井出口, "OSIプロトコルにおける受信データの並列処理方式," 信学論B-I, Vol. J74-B-I, No.2, Feb. 1991.
- [4]: 加藤, 井戸上, 鈴木, "OSIプロトコル実装のためのユーザデータをコピーしないバッファ制御方式," 情処マルチメディア通信と分散処理研究会, 62-13, Sept. 1993.