

VIA 上の通信ライブラリの最適化

村山 和宏、落合 真一、山口 義一
三菱電機(株) 情報技術総合研究所

1. はじめに

近年のネットワークハードウェアの進歩および新しい通信方式の研究成果により、低遅延高速通信の実現が可能となった。中でも、新しい通信方式である VIA (Virtual Interface Architecture)は、OS のオーバーヘッドを削減した通信が可能であるため、ギガビットの NIC を用いて通信を行うと、通信速度 100MB/sec、遅延 7 μ 秒を実現することができる。そこで、大容量高速データ通信を必要とする信号処理システムの通信方式としてVIAを適用するため、VIA の通信ライブラリ関数の改良および測定、評価を行った。

2. システムの要求課題と解決策

2.1 システムの要求課題

我々がターゲットとする信号処理システムでは、1 計算機でデータ受信プロセスとデータ処理プロセスを並行して実行させることにより、応答時間を短縮させる。そこで、データ受信とデータ処理を並行して行うためには以下の 3 点が要求される。

- (1)通信処理に CPU を専有させない
- (2)高い通信速度
- (3)パケット損失のない通信

2.2 解決策

VIA には、ポーリングによってパケット受信を知るための関数と割込みによって受信通知を受けるための関数が存在する。2つの関数の受信方式の特徴を、パケット到着時間の間隔が短い／長い場合に分けてまとめる。

・パケット到着間隔が短い場合

表1:ポーリングと割込みの特徴(その1)

受信方式	通信速度	パケット損失	CPU 使用率
ポーリング	高	無	高
割込み	低	有	高

・パケット到着間隔が長い場合

表2:ポーリングと割込みの特徴(その2)

受信方式	通信速度	パケット損失	CPU 使用率
ポーリング	高	無	高
割込み	高	無	低

表 1、2 より、パケット受信間隔が短い場合にはポーリング、長い場合には割込みを使用すればよい。

3. 実装内容

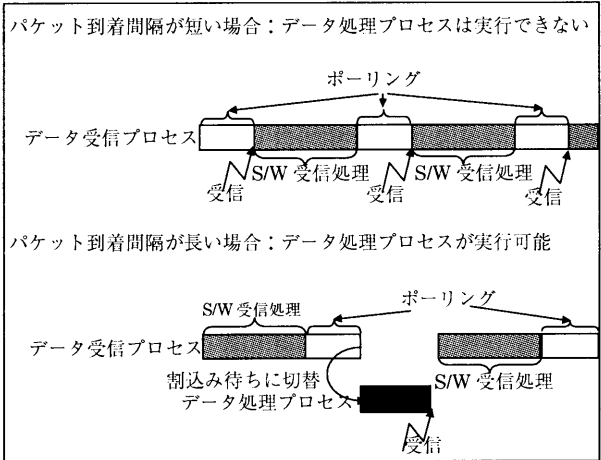


図1:動作手順

2 章で示した解決方針をもとに、パケット到着間隔に応じてポーリング・割込みを切り替えて受信を行う関数 I/F を追加した。本関数の動作手順は以下の通り：

- 一定時間ポーリングする。
- ポーリングしている最中にパケットが届いた場合には、データ受信処理を行い、1.へ戻る。
- 1.でパケットが受信できなかった場合には、割込みによる受信待ちに切り替える。切替後は他のプロセスが処理可能となる。

4. 測定と考察

4.1 測定の目的

Linux のような割込み応答性が保証されていない OS では 100 ミリ秒のオーダーでスケジューリングに遅れが生じることがしばしばあり、その際にパケットを損失する。このような場合、受信バッファサイズを大きくしてパケット損失を防止する方法と、ポーリング時間を長くしてパケット損失を防止する方法がある。

しかし、ポーリング時間をあまりに長くすると CPU 使用率が上昇して演算と通信の並行処理ができなくなり、確保できる受信バッファのサイズにも上限がある。そこで、測定により適切なポーリング時間と受信バッファサイズを決定する。

4.2 測定環境

測定環境は以下の通り。

- ・計算機…CPU:Pentium III 600MHz、メモリ 128MB
- ・NIC…Giganet 社 cLAN 1000
- ・OS…Redhat Linux 6.1 (カーネル 2.2.12)

4.3 測定結果

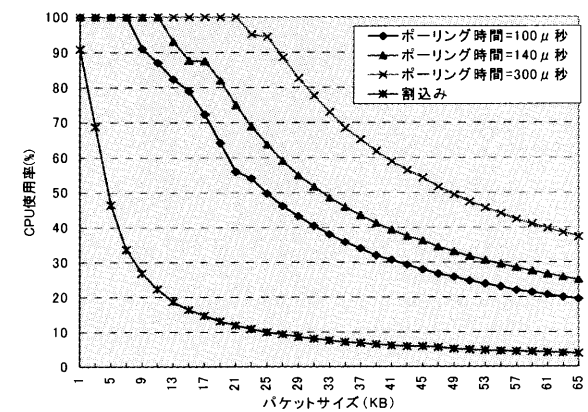


図 2:CPU使用率

上のグラフは、ポーリング時間を 100 μ 秒、140 μ 秒、300 μ 秒とし、1KB～65KB のパケットをバースト転送し、3 章で示した方式を用いて受信を行った場合の CPU 使用率の変化について示している。我々がターゲットとするシステムでは、50KB を中心にパケットを送信することが多く、グラフを見ると、ポーリング時間を 100 μ 秒、140 μ 秒とした場合では、50KB パケット転送時には CPU 使用率はそれぞれ 25%、31%であり、通信処理と演算処理の並行実行は可能である。しかし、ポーリング時間を 300 μ 秒とした場合では、50KB のパケット転送時の CPU 使用率は 50%と高い。

また、次の表は、ポーリング時間 100 μ 秒、140 μ 秒、300 μ 秒において、受信バッファが 7MB、10MB、15MB の場合のパケット損失の有無を示したものである。

表 3:受信バッファサイズとパケット損失の有無

	7MB	10MB	15MB
100 μ 秒	有(10～15KB にて 0.2%)	有(20～25KB にて 0.0001%)	無
140 μ 秒	有(20～25KB にて 0.00025%)	無	無
300 μ 秒	無	無	無

表 3 より、ポーリング時間 100 μ 秒、140 μ 秒、300 μ 秒でそれぞれ 15MB、10MB、7MB の受信バッファがあれば、パケットの損失なく通信を行うことができる。なお、

ポーリング時間が 100 μ 秒未満の場合ではバッファサイズをシステムの上限まで確保してもパケット損失が発生した。

以上より、我々がターゲットとするシステムでは、ポーリング時間を 100 μ 秒～300 μ 秒の間、バッファサイズを 7MB～15MB の間で変化させればシステム要求を満たすことができる。

4.4 考察

本測定結果および文献[1]のデータを元に、4.2 節と実行環境が異なる場合における、本測定結果との関係について述べる。

・CPU 性能が異なる場合:

VIA は、CPU を用いずに通信を行うため、CPU 性能が異なっても本環境の測定結果をそのまま流用することができる。

・NIC のバンド幅が異なる場合:

NIC のバンド幅が 2 倍になると、一定時間内に届くパケット数は 2 倍になるので、受信バッファを本環境で求めた値の 2 倍にすればよい。同様に、NIC のバンド幅が 1/10 であれば、受信バッファは本環境における値の 1/10 で済む。

・OS が異なる場合:

リアルタイム OS では割込み応答性が保証されているため、バッファサイズは本環境よりも小さくて済み、CPU 使用率も本環境よりも小さくて済む。

5. まとめ

本項では、我々がターゲットとする信号処理システムに対し、効率よく処理を行う方式について検討を行った。その結果、本方式の有効性を確認した。

本稿で提案した方式は、ポーリング時間を 100 μ 秒～300 μ 秒、バッファサイズを 7MB～15MB で変化させればシステム要求を満たすことができる。

参考文献

- [1]落合, 村山, 山口: “VI Architecture による分散リアルタイム環境の構築”, 情報処理学会 DPS 研究会, 2000 年 3 月.
- [2]村山, 落合, 山口: “高速インターコネクトにおける OS オーバーヘッド策現方式の検討”, 情報処理学会第 60 回全国大会, 2000 年 3 月.