

並列計算機 Ships1 のノード間結合装置の構築

†加藤 渉, †三浦 康之, †高野 誠一

湘南工科大学 工学部情報工学科

■ 1. はじめに

近年、計算機によって処理すべき分野が拡大している。より大規模な計算を必要とするようになり処理速度の飛躍的な向上が必要となる。また PC の高性能・低価格化に伴い、複数の PC を統合化し高い処理速度・信頼性を得ることを目的とした PC クラスタが実用化されている [1]-[3]。

PC クラスタを構築するための一般的な方法である Gigabit LAN は高いスループットを持っている。しかし、オーバーヘッドが大きい他、バッファのコピーによる遅延などにより細かい Data の転送を必要とする PC クラスタには向いていないという欠点がある。そこで、専用の通信機構として LVDS を用いて小規模並列計算機向けにこの問題を解決する。

湘南工科大学では、Gigabit LAN と LVDS の両方を使用し、それぞれの特性を生かした小規模かつ高性能な並列計算機の構築を目的とした Ships1 プロジェクトを進めている。

本稿では Gigabit LAN に代わりノード間の通信を行うための装置として、LVDS コネクタ付 FPGA ボードの開発状況を報告する。

■ 2. クラスタ型並列計算機

2.1 関連研究

専用の通信機構として Myricom 社の Myrinet[2] や筑波大学システム情報工学の Maestro2・3[3] など通信エラー率が低く、低レイテンシ及び高スループットに優れたネットワークインターフェイスの開発が進められている。これらの装置は大規模かつ低遅延並列計算機の構成が可能だが、専用のスイッチが必要になるなどコストが高くなる問題がある。

2.2 並列計算機 Ships1 の構成

Ships1[4][5] は、16 台の PC とネットワークインターフェイス (Gigabit LAN)、Gigabit スwitch で構成される。本研究では市販されている LVDS 付き FPGA ボードを使用する。LVDS を直接

網でリング型ネットワークに構成することにより、専用のスイッチは使用せずに小規模限定であるものの、低コストかつ高速な通信網の構築が可能になる。

■ 3. FPGA 回路の設計

3.1 FPGA ボードの仕様

図 1 に本研究で使用する FPGA ボードの内部構成を示す。本研究で使用する FPGA ボードは、ボードコネクタが 2 つあり、他の回路と接続することができる。LVDS は 1 対 1 の通信のため、複数ノードを結ぶ通信を行う場合にボードを複数用意する必要があり、複数のボード間通信はボードコネクタを経由して実現することが可能である。

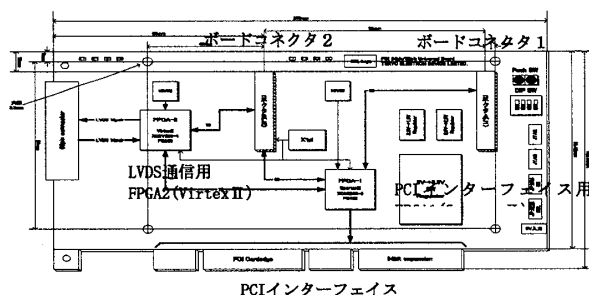


図 1 FPGA ボードの内部構成

3.2 FPGA 回路の開発状況

FPGA 回路の開発は下記の 4 つの順序で進める。

- 1) LVDS 通信を可能にする。
 - 2) 多対多のトラフィックに対応できるように LVDS を高速化する。
 - 3) 多対多の通信のためのルーティングの機能を載せる。
 - 4) 本体に LVDS の状態を知らせることにより、Gigabit LAN と LVDS の使用を使い分ける。
- 上記の機能を実現させるため、各 FPGA は下記の回路を実装する。

FPGA (Spartan2)

- 1) ・メモリと I/O デバイス間は、DMA 方式を用いる。また、Dual Port RAM を使用し、データ処理を高速化する。
- 2) ・クロックマネジメント回路 (DLL) を使用し、

FPGA 内で高速化する。

- 3) ・ボードコネクタを制御する回路を作成すると同時に、アドレス制御を行い、ルーティング機能を作成する。
- 4) ・LVDS の状態を PCI バスを通して CPU に送る回路を作成する。

FPGA (Virtex2)

- 1) ・Dual Port RAM を使用し、データ処理を高速化する。
・LVDS 制御回路を作成する。
- 2) ・クロックマネジメント回路 (DCM) を使用し、FPGA 内で高速化する。

FPGA (Spartan2) は Xilinx 社の PCI コアおよび東京エレクトロニクス社社の基本回路を用いて、DMA 制御方式を用いた PCI インターフェイスから Dual Port RAM への Data 転送を可能にした。LVDS 間通信の確認を行うため、一時的に Dual Port RAM に送られる Data を FPGA (Virtex2) へ転送する。図 2 に、FPGA (Spartan2) の構成を示す。

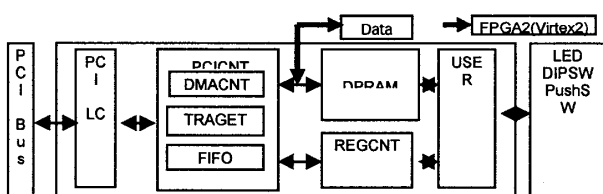


図 2 FPGA (Spartan2) の回路

FPGA (Virtex2) は送信側 PC と受信側 PC での LVDS 間通信の確認を行うため、送信用プログラムと受信プログラムを別々に載せている。送信用プログラムでは、Spartan2 から転送されてきた Data を Virtex2 のメモリ (FF) に格納し LVDS へ Data を送信する。LVDS へ正しい Data が転送できたか LED で確認する。図 3 に送信用回路を示す。

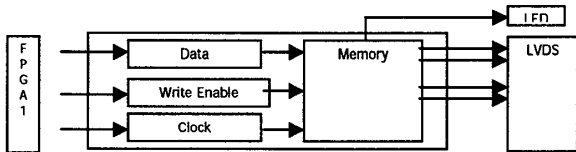


図 3 FPGA (Virtex2) の送信用回路

受信プログラムでは、LVDS から Data の受信を行いメモリ (FF) に格納する。LVDS ケーブルを通した Data を LED で確認する。図 4 に受信回路を示す。

3.3 動作状況

動作を確認は東京エレクトロニクス社からいただいたソフトを使用し、Spartan2 のメモリへ情報を書き込み確認する。Spartan2 から

Virtex2 への情報の転送を LED の点灯により正しく送られていることを確認し、Virtex2 同士 (LVDS 間) の通信も LED により確認することができた。

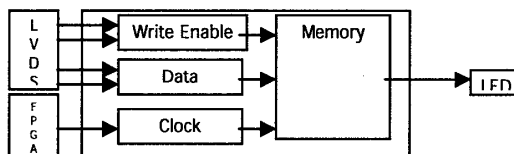


図 4 FPGA (Virtex2) の受信回路

4. 今後の予定

LVDS 間通信が正しく行えたのを確認した後、FPGA (Virtex2) の回路を送受信プログラムへと統合、DCM 回路を用いて高速化、FPGA (Spartan2) へ 64bit の情報を送受信可能に変更する。図 5 に変更後の FPGA (Virtex2) の構成を示す。

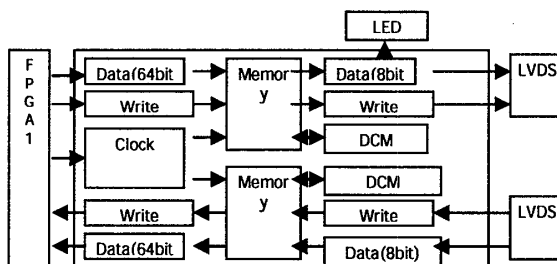


図 5 予定 FPGA (Virtex2) の回路

また、高速化のための変更もあわせて行う。具体的には、転送されてきた 8bit の Data を DMC 回路で 4 倍速にし、さらに DDR (Double Data Rate) 制御により 8bit の Data を 8 倍速で転送する。

■ 謝辞

本研究の一部は、東京エレクトロニクス株式会社インレピアム事業部の協力により行われた。

■ 参考文献

- [1] TOP500 Supercomputer Site
<http://www.top500.org/>
- [2] N. J. Boden, D. Cohen, R. E. Felderman, A. E. Kulawik, C. L. Seitz, J. N. Seizovic and WenKing Su. "Myrinet { A Gigabit-per-Second Local-Area Network". IEEE MICRO, Vol. 15, No. 1, pp. 29{36, February 1995
- [3] 青木圭一, 山際伸一, 和田耕一, 小野雅晃, Maestro2 クラスタネットワーク向けメッセージパッシングライブラリの開発と評価
- [4] 松尾成志, 岡本恵介, 大谷 真, 中小規模並列コンピュータ Ships1 の開発
- [5] 松原裕人, 和田 卓, 大谷 真, Ships1 におけるノード間接続装置の研究