

COREswitch におけるフレームフォワーディング方式

4 G - 1

川野哲生 高橋直久 丸山充 八木哲 小倉毅

NTT 光ネットワークシステム研究所

1 はじめに

我々は超高速データ通信方式の新方式の1つとして、MAPOS³⁾と呼ぶ新プロトコルの検討を行ないIETFにて既に提案済みである¹⁾。本稿では、MAPOSプロトコルに準拠して実現した、並列分散型高速通信スイッチCOREswitch⁴⁾における外部の通信回線対応のプロセッサ(CIF)のデータ転送/制御方式と遅延等の評価を示す。

2 COREswitch の概要

COREswitch は図1に示すように、外部の通信回線対応のプロセッサ(CIF)、システム全体の監視/制御用のプロセッサ(IFP)⁷⁾、クロスバススイッチ(XSW)⁵⁾、アービトレーションモジュール(ABT)⁶⁾からなる。ルーティング情報等の制御情報は制御バス(C-bus)を介してIFPから各CIFへと送られる。フレームのフォワーディング処理は、CIFで受信されたフレームデータが、XSWにより転送され、再びCIFから回線へと出力することにより行われる。フレームのフォワーディングの手順は以下の通りである。

1. CIFにてフレームデータを受信。
2. CIFにてフレームのアドレス部情報により宛先のCIF(NSP,SSPの場合はIFP)を決定し、XSW接続要求をABTに送り、ABTがXSWの接続を行う。
3. XSWを通してフレームデータを転送する。転送先CIFではフレームデータの先頭が到着した時点より回線への出力を開始する。
4. XSWの転送終了後、CIFはXSW開放要求をABTに送り、ABTがXSW接続の開放を行う。

3 CIF の構成

CIFの構成を図2に示す。CIFは、SONET/SDHのオーバーヘッド処理部(SONET/SDH O/H Proc.)、HDLCフレーム(HDLC)、送受信FIFO(TxFIFO,RxFIFO)、受信フレームのアドレス部とフレーム長を格納するReqFIFO、ルーティング制御を行うRSE、および、ルーティングテーブル(Routing Table)からなる。

現在、156M(OC-3c)版、および、622M版(OC-12c)版の2種のCIFが稼働中である。CIFでは、MAPOS ver. 1 および MAPOS-16²⁾をサポートし、CIF内部

Frame forwarding mechanism of COREswitch
Tetsuo KAWANO, Naohisa TAKAHASHI, Mitsuru MARUYAMA,
Satoru YAGI and Tsuyoshi OGURA
NTT Optical Network Systems Laboratories
9-11 Midori-cho 3-Chome Musashino-sh, Tokyo 180 Japan.

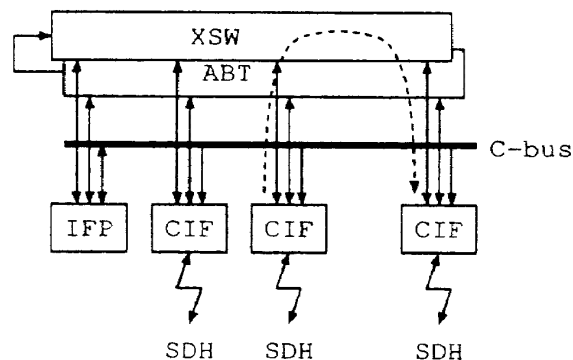


図1: COREswitch 概略図

に設けられた Routing table を用いて、フレームのフォワーディング先を決定する。また、受信バッファとして256KBのRxFIFO、送信バッファとして512KBのTx-FIFOを装備する。活線挿抜にも対応している。

CIFの特長を以下に挙げる。

- 受信部に store-and-forward 方式を採用：回線速度156M,622M,2.4Gの3タイプのCIFを想定し、XSWはポートあたり2.56Gbps X 2(送受信)の容量を持つ。フレームデータを受信側CIFにて一度バッファに格納し、フレーム全体のデータが揃った時点でXSWを介した転送を行う。そのため、XSW上は回線速度に関係なく2.56Gbpsの速度で転送が行われ、低速度のCIFからの転送により、XSWが占有されることがない。また、エラーフレームを受信した場合は、受信側CIFにて廃棄し、XSWを使用しない。
- XSW転送の単純化：XSW転送においては、フレームデータと共にフレームデータ受け取り側CIFのTx-FIFOへの書込み信号をXSWを経由して行う。そのため、XSW転送におけるデータの流れる方向は一方のみとなり、XSW転送経路上に任意段のラッチを挿入することができる。特に多ポートのXSWの場合、配線長が長く、また、ポートによる配線長の差が大きいが、COREswitchでは、それぞれのポートに任意段のラッチを挿入することができるため、多ポートXSWの設計が容易である。
- データと制御の分離：HDLCフレームの受信部にて、フレーム内のアドレス部の取り出し、および、フレーム長の計算を行い、フレームデータとは別にReqFIFOに格納する。RSEでは、ReqFIFOから取り出した、アドレス情報によりルーティン

テーブルを引き、XSW 設定リクエストを ABT へ送る。RSE は ABT から XSW 設定完了通知を受けると、フレーム長情報を元に、Rx FIFO の読み出し、および、転送先 CIF の Tx FIFO への書き込み信号を送出する。この様な方式により、ルーティングの際に高速なクロックで動作する Rx FIFO の出力側のデータを観測する必要がなく、Rx FIFO の出力側の制御も単純化される。

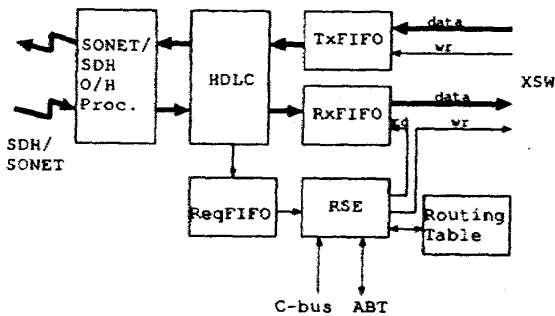


図 2: CIF 構成図

156M, 622M 版の CIF は共通の PCB 基板上に、5 個の FPGA および、SONET/SDH 用汎用部品等を実装している。基板サイズは 233 × 160mm である。

4 フォワード遅延評価

17ポート版 XSW (IFP 用 1ポート, CIF 用 16ポート), および、622M 版 CIF を用いた場合のフレームフォワード遅延の測定結果を表 1 に示す。ここでは、XSW 転送における競合が無い場合について、1kbytes のフレームデータの先頭が 622M の回線から入力を開始した時点から、フレームの先頭が再び 622M の回線から出力されるまでの時間を測定した。

COREswitch では受信側 CIF にて store-and-forward 処理を行っているため、フレームフォワード遅延はフレー

表 1: フレームフォワード遅延の測定値

項目	遅延時間
SONET/SDH O/H 処理 (送受信合計)	4.9 μs
HDLC 受信部遅延	0.3 μs
受信部フレームバッファリング (1kbyte 分)	13.4 μs
ReqFIFO 遅延	0.3 μs
ルーティング, および, スイッチ設定	0.6 μs
XSW 転送遅延	0.0 μs
TxFIFO 遅延	0.7 μs
HDLC 送信部遅延	0.4 μs
合計	20.6 μs

ム長の 1 次関数となる。フレーム長を $L(\text{byte})$ とすると、フレームフォワード遅延 $T_d(s)$ は、

$$T_d = \frac{L \times 8}{622 \times 10^6 \times \frac{1040}{1080}} + 7.2 \times 10^{-6}(s)$$

となる。最短フレーム長 (4 バイト) の場合のレイテンシは約 7 μs であった。622Mbps の回線では、7.2 μs は約 540 bytes のデータ、光速で約 2km の距離に相当し、SONET/SDH での運用を考慮すると十分に短い時間であると言える。

5 おわりに

本稿では、MAPOS プロトコルに準拠して実現した並列分散型高速通信スイッチ COREswitch における外部の通信回線対応のプロセッサ (CIF) について、その制御方式とフレームフォワード遅延の測定結果を示した。CIF はハードウェア構成の単純化により、622Mbps や 156Mbps という非常に高速な回線インターフェースにも関わらず、市販 LSI と FPGA を組み合わせることにより実現でき、基板サイズも 233 × 160mm と比較的コンパクトなものとなった。また、フレームフォワード遅延時間は、実用上十分に短いことが観測された。

今後、2.4G 版 CIF の設計製作、および、実使用環境下での評価を進めていく予定である。

謝辞

共に開発を進めていただいた、吉田敏明氏、小林正之氏、宮野入暁氏に感謝いたします。

参考文献

- 1) K. Murakami and M. Maruyama, "MAPOS — Multiple Access Protocol over SONET/SDH Version 1," RFC2171, June 1997.
- 2) K. Murakami and M. Maruyama, "MAPOS 16 — Multiple Access Protocol over SONET/SDH with 16 Bit Addressing," RFC2175, June 1997.
- 3) 村上, 高橋, 丸山, 八木, 小倉, 川野, "超高速データ通信プロトコル MAPOS の概要", 情処第 56 回全国大会論文集, 1998 年 3 月.
- 4) 高橋, 村上, 丸山, 八木, 小倉, 川野, "並列分散型高速通信スイッチ COREswitch", 情処第 56 回全国大会論文集, 1998 年 3 月.
- 5) 丸山, 高橋, 八木, 小倉, 川野, "COREswitch のハードウェアアーキテクチャ", 情処第 56 回全国大会論文集, 1998 年 3 月.
- 6) 小倉, 高橋, 丸山, 八木, 川野, "COREswitch におけるマルチキャスト方式", 情処第 56 回全国大会論文集, 1998 年 3 月.
- 7) 八木, 高橋, 丸山, 小倉, 川野, "COREswitch のソフトウェアアーキテクチャ", 情処第 56 回全国大会論文集, 1998 年 3 月.