

## COREswitch のハードウェアアーキテクチャ

3 G - 7

丸山 充 高橋直久 八木 哲 小倉 毅 川野哲生

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、MAPOS(Multiple Access Protocol over SONET/SDH)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<sup>1)</sup>。本発表では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のハードウェアアーキテクチャについて、全2重クロスバススイッチを用いたバックプレーンのデータ転送方式を中心に述べ、併せて得られた評価結果を示す。

## 2 COREswitch の構成と基本動作

COREswitchは図1に示すように、複数の通信回線対応プロセッサ(CIF)とシステム全体の監視/制御用のプロセッサ(IFP)がバックプレーン(BP)で接続されている。BP上ではそれぞれのプロセッサ間をデータ転送用のクロスバススイッチ(XSW)と制御用の64bit制御バス(C-bus)で接続している。またシステムにはXSWのアービトレーションを行なうためのアービトレーションモジュール(ABT)が1枚搭載され、ABTと各プロセッサとの間は個別に接続される。それぞれの主要諸元は表1に示すとおりである。

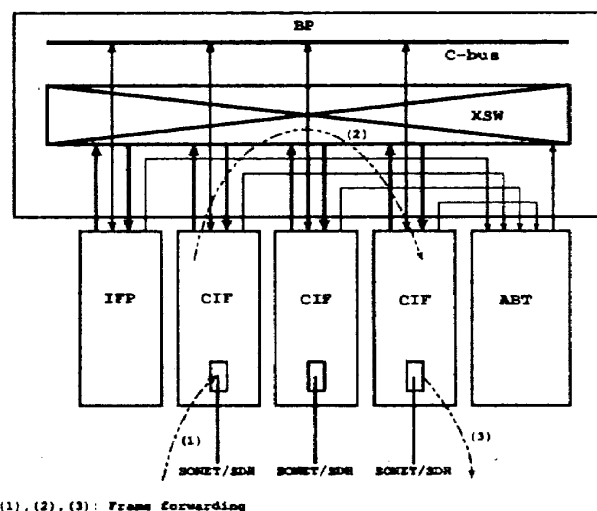
COREswitchの基本フォワード動作は図1に示すように、あるCIFが回線からMAPOSフレームを受信すると、CIF内でHDLCフレーム制御を行ない、アドレスフィールドを抽出して、検索転送制御部に渡す。検索転送制御部ではルートテーブル検索により、出力先のCIF番号を得てからABTに対して、データ転送要求を出す。ABTでは、BP上のXSWのリソースが空いていれば、XSWの接続制御を行なってから該CIFに転送可の信号を送る。CIFでは転送データにリンクレイヤのエラーが発生していない事を確認してからBPを経由して、出力先のCIFにフレームデータを転送する。出力先CIF

Hardware Design of COREswitch System.

Mitsuru MARUYAMA, Naohisa TAKAHASHI, Satoru YAGI,  
Tsuyoshi OGURA, and Tetsuo KAWANO

NTT Optical Network Systems Laboratories

9-11 Midori-cho 3-Chome Musashino-sh, Tokyo 180 Japan.



(1), (2), (3): Frame Forwarding

図1: COREswitch の構成

では転送されたフレームをHDLCレベルのコーディングおよび、FCS (Frame check sequence)を付加して、SONET/SDHのフレームに載せ、回線に送出する。

## 3 バックプレーンデータ通信方式

COREswitchシステム内のデータ通信は、CIFプロセッサ間のデータ転送を高速に行なうために以下の特徴を有する。

- ・高速なXSWを送受信を分離して使用
- ・タグ制御情報をデータ転送と同時伝送
- ・データ系はXSW、制御系はC-busに分離
- ・FIFO状態の先行監視

XSWは36bit幅17×171段のものを採用しておりプロセッサ間の転送は、XSWの接続方法でユニキャスト、マルチキャスト、ブロードキャスト、ループバック転送が可能である。各プロセッサとXSW間は、送受信を分離して32bit幅データを同期伝送している。また4bitの制御情報をタグデータとして、データ転送と同一方向に転送することで、制御情報のハンドシェイクのオーバヘッドを削減している。

表 1: COREswitch の主要諸元

項目	機能/諸元
システムサイズ	430mmx386mmx500mm
スロット数	17 スロット
ボードサイズ	233mx160mm
BP	XSW / C-bus
・ XSW	17 × 17 1段 36bit 幅
・ C-bus	最大 40Mbyte/s
CIF	SONET/SDH, リンク制御, アドレス検索, 転送制御
・ 接続回線	OC-12c, OC-3c シングル/マルチモードファイバ, 非シールドより対線
・ 入力 FIFO	256KByte
・ 出力 FIFO	512KByte
・ 回線制御	汎用 SONET/SDH LSI
・ HDLC 制御	FPGA
・ 検索転送制御	FPGA
・ スロット実装	活線挿抜対応
IFP	NSP,SSP プロトコル制御, システム管理制御
・ プロセッサ	Intel 960HD 33/66MHz
・ メモリ	RAM 128MByte, ROM 16MByte
・ バッファ	32MByte
・ 転送制御	FPGA
・ インタフェース	シリアル / Ethernet
ABT	クロスバススイッチ制御
・ 制御	FPGA
・ 接続要求	17 スロット
・ 要求信号	各スロット 4bit 幅

IFP が各 CIF のイニシャライズ/検索テーブルの書き換えなどの制御を行なう場合には、全て XSW とは分離した 64bit 幅の C-bus を使って行なう構成にしている。これにより、XSW をデータ系専用にして高効率で使う事ができる。また同時に、CIF 内部のデータ転送経路もデータ系/制御系に完全に分離して高速データ転送可能な構成にしている。

効率良いフォワーディング転送を行なうためには、システム内でデータブロックを避けなければならない。今回のシステムで、データ輻輳が起きる可能性のある箇所は以下の部分である。

- (1) 回線入力側 CIF の入力 FIFO 溢れ
- (2) 出力先 CIF に対するアービトレーションタイムアウト
- (3) 出力先 CIF の出力 FIFO 溢れ

ブロックが起きると HOL(head-of-the-line) ブロック問題によりシステムのトータルのフォワーディング性能が下がるため、早期にブロック状態を監視して対処する必要がある。今回我々は、回線入力側の CIF で、入力側 FIFO, 出力先 FIFO, 出力先へのアービトレーション状況を全て監視し、ブロックが起きた場合には、積極的に入力 FIFO の先頭フレームを廃棄する回路を付加する事で、システムとしての性能低下を避けている。

#### 4 バックプレーンの転送性能

BP 内 XSW と各プロセッサ間および BP 内のデータ転送は、36bit 幅の同期転送で実現されており、BP 上のクロックに同期する。本システムでは、クロック周波数最大 80MHz まで全てのスロット間でのデータ転送が問題なく行なえた。この時の転送速度は各スロットあたり、2.56Gbps の送信と受信が同時に行なえる事になる。これから 1 プロセッサあたり双方向 5.12Gbps の伝送が達成でき、XSW のトータルの転送速度は、43.52Gbps (2.56Gbps × 17) を実現できた。

次にデータ転送時の BP 上のオーバーヘッドを示す。CIF から転送要求が出て、XSW が設定されるまでのアービトレーション時間は 525ns である。また BP の転送遅延は、プロセッサから BP 上への入力ラッチ、XSW LSI の入力ラッチ、XSW LSI の出力ラッチ、BP からプロセッサへ出力するラッチの計 4 段で同期転送を行うために 80MHz 時は 50ns になる。この値は MAPOS フレームの伝送遅延と比較して十分に小さい値である。

#### 5 おわりに

我々が提案した MAPOS プロトコル準拠の並列分散型高速通信スイッチ COREswitch のハードウェア構成の概略を BP 上のデータ転送を中心に示した。その結果 BP は 80MHz の動作が可能であり、トータルスループット 43.52Gbps の転送性能が得られる事が分かった。これより本システムは OC-48c の回線速度までに対応できる性能を持つ事が明らかになった。今後、システムとしての評価を進めると共に、OC-48c 対応 CIF の設計を行なう予定である。

#### 謝辞

共に開発を進めていただく吉田敏明氏、小林正之氏、佐島隆博氏に感謝いたします。

#### 参考文献

- 1) Ken Murakami and Mitsutu Maruyama: "MAPOS-Multiple Access Protocol over SONET/SDH version 1" 他, IETF RFC2171-2176, June 1997.