

# Frame Switch エミュレータの実現

6N-3

- CORE-Switch -

八木 哲 川野哲生 丸山 充 村上健一郎

NTT ソフトウェア研究所

## 1 はじめに

本稿では、次世代計算機ネットワークのための高速データ通信方式、Frame switching<sup>1)</sup>のための検証環境の一つであるスイッチ・エミュレータについて述べる。エミュレータの主目的は、Frame switching で使用するプロトコルの検証であり、検証作業に伴う機能追加/変更の容易さが重要となる。現状では動作速度はさほど重要視していない。このため、基本的なフォワーディング処理と検証対象となるプロトコルの処理をアプリケーションレイヤで行なり構成を取った。

本稿では、先ずエミュレータのファームウェアの構成を示し、次にファームウェアが提供する、検証対象となるプロトコル処理プログラム用のI/Fと内部状態観測プログラム用のI/Fを示す。

## 2 ファームウェアの構成

### 2.1 ハードウェア部分

構築の容易さと不安定要因排除のため、SUN/SS5と互換性のある汎用のVMEのCPUボードを用いたWSクラスタ型の構成を取った(図1)。32K/64KByteという大きなMTU<sup>1)</sup>の扱いを容易にするため、CPUボード間のデータ転送はVMEbus<sup>4)5)</sup>を用いた。

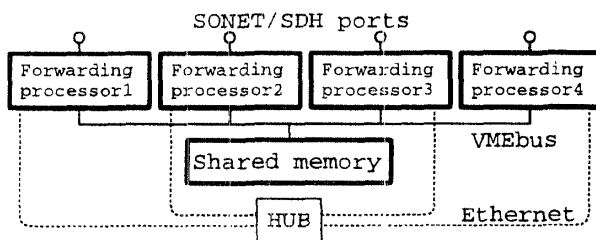


図1: ハードウェアの構成

- フォワーディングプロセッサ  
CPUボードのS-BUSスロット(最大2スロット)にSONET/SDHインタフェースカード<sup>2)3)</sup>を実装している。このCPUボードを対向接続した場合の転送レートの測定値は40Mbps程度である。この値

は、FastEtherカードを実装して対向接続した場合と同等である。

- 共有メモリ  
CPUボード間のデータ転送のバッファとして使用する。容量は32MByte。
- VMEbus  
バス要求の調停は、4つある要求信号線をラウンドロビンで監視できるため、最大4ボードまで公平なアクセスができる。VMEbusの転送レートは、MTUと同じ32/64MByteをDMA転送した場合に40MByte/sec前後であり、エミュレータ用のバックボーンとして十分な転送レートを備えている。
- Ethernet  
各フォワーディングプロセッサからの情報収集に用いる。

CPUボードのS-BUSスロット数とVMEbusの調停機能の制約から、エミュレータの最大ポート数は8である。

### 2.2 ソフトウェア部分

通常のSunOSにSONET/SDHドライバとVMEドライバをインストールし、フォワーディングプログラムをアプリケーションレイヤに実装した以下の構成を取っている(図2)。

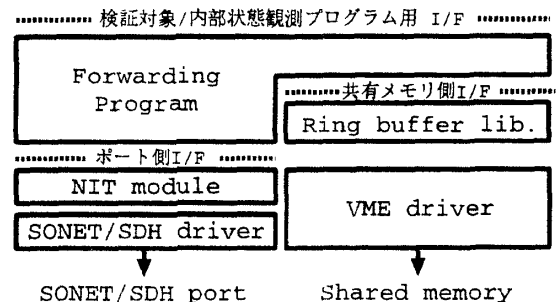


図2: ソフトウェアの構成

- ポート側I/F (SONET/SDH driver + NIT)  
エミュレータで使用するドライバは、全ての動作を制御可能とするためNITモジュールのみにデータを渡すように改造している。またブロードキャスト時の受信/送信バッファ間の転送も抑制している。

- 共有メモリ側 I/F (VME driver+Ring buffer lib.)  
共有メモリはリングバッファとして使用する。バッファは各 CPU ボード間ごとに独立しており、クリティカル・セクションを局所化している。各バッファの管理領域は 1DMA サイクルでアクセス可能なサイズであり、バッファへの書き込み/読み出しの排他制御を実現している。
- フォワーディングプログラム  
各フォワーディングプロセッサ上で同じソフトウェアが動作している。このプログラムは、1)NIT モジュールからバケットを読みだして VMEbus 上の共有メモリへ書き込む操作と、2)VMEbus 上の共有メモリからバケットを読みだして NIT モジュールへ書き込む操作を交互に行なっている (図 3)。

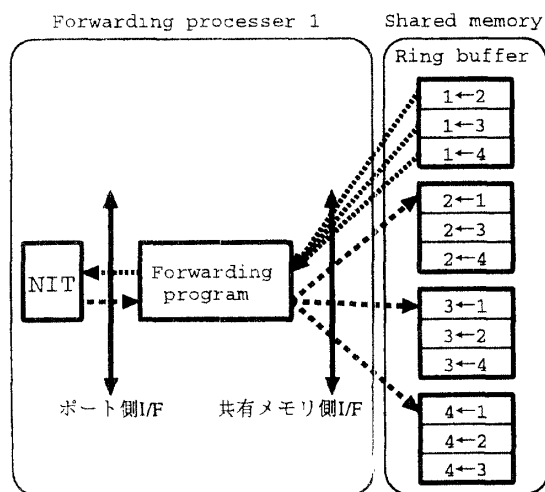


図 3: フォワーディング動作

リンクレイヤのアドレスとポートの対応を示すスイッチング用のテーブルは、VMEbus の競合を減らすため、各フォワーディングプロセッサに分散管理している。このための情報は、ARP<sup>6)</sup>の要求/応答バケット転送時にバケットからアドレス情報をスヌープすることで得ている。

### 3 検証支援環境

プロトコル検証のために作成した、a) 検証対象となるプロトコル処理プログラム用の I/F と、b) 内部状態観測プログラム用の I/F について示す。

#### a) プロトコル処理プログラム用の I/F

検証対象となるプロトコル処理プログラムは、ユーザ関数として定義する。このユーザ関数は、スイッチ宛のバケットを受信した場合に、受信したのバケットへのポインタ、バケット長、プロトコル番号を引数としてフォワーディングプログラムから呼ばれる。このユーザ関数からバケットを送信する場合は、あらかじめ用意されたフォワーディングプログラムのスタック領域にバケットを積む。このバケットは、フォワーディング処理時に、通常のバケットと同様に処理される。

#### b) 内部観測プログラム用 I/F

内部観測プログラムは、ユーザ関数として定義する。このユーザ関数は、バケットのフォワーディング処理を行なう時に、バケットへのポインタ、バケット長、プロトコル番号を引数としてフォワーディングプログラムから呼ばれる。観測値を指定のポインタが示す領域に記述すれば、以下のタイミングで、VMEbus 上の共有メモリ、もしくは各ボード内の共有メモリの特定領域に反映される。

#### ● VMEbus 上の共有メモリ

リングバッファの管理情報更新時。値の観測はバスサイクルによる排他制御が可能のため、エミュレータ全体のスナップショットが取得できる。

#### ● ボード内の共有メモリ

フォワーディング実施時。値の観測は Ethernet 経由で、任意の時に収集するか、反映時にフォワーディングプロセッサから送信する。ポート毎に送信/受信を契機にして状態が取得できる。

## 4 おわりに

本稿では、プロトコル検証のために作成したスイッチ・エミュレータについて、ファームウェアの構成と、検証のための支援環境を示した。本エミュレータでは、アプリケーションレイヤでフォワーディング処理や検証対象となるプロトコル処理を行なうため、検証作業に伴う機能の追加/変更と内部状態の観測が容易な構成が取れる。

今後、実際の検証作業をすすめる過程で、検証環境としての便利性/汎用性を高めると共に、動作速度についても留意していく予定である。また速度を追求したカットスルー型のスイッチの検討を行なう。

### 謝辞

日頃御指導をいただき、高橋直久リーダ、グループの皆様、ならびに多大な互助力をいただき吉田敏明氏、小林正之氏、佐島隆博氏に深謝します。

### 参考文献

- 1) 村上, 丸山, 八木, 川野, “超高速データ通信方式 Frame Switching の概要 - Frame over SONET/SDH -”, 情処学会第 54 回全国大会論文集 (3), Mar., 1997
- 2) 丸山, 川野, 八木, 村上, “Frame Switching 方式による通信インターフェースの実現と評価 - SONET-LAN -”, 情処学会第 54 回全国大会論文集 (3), Mar., 1997
- 3) 川野, 八木, 丸山, 村上, “Frame Switching 方式による HDLC フレーム LSI の実現 - CORE framer -”, 情処学会第 54 回全国大会論文集 (3), Mar., 1997
- 4) 日本モトローラ, “VMEbus アーキテクチャマニュアル Rev. C.1”, Oct., 1986
- 5) VME 64 Specification (ANSI/VITA 1-1994)
- 6) Douglas Comer 著, 村井, 楠本 訳, “第 3 版 TCP/IP によるネットワーク構築 Vol.1”, 共立出版, 1997