計算機アーキテクチャ 140－4
（2000．11．30）

LASN用 10Gbps/port 8x8 ネットワークスイッチ: RHiNET-3/SW

西 宏章[†2] 上野 龍一郎[†1] 多昌 廣治[†5]
稲沢 悟[†3] 西村 信治[†4]
工藤 知宏[†2] 天野 英晴[†1]

本稿では，新しいネットワーククラス LASN（Local Area System Network）のための LSI スイッチチップ RHiNET-3/SW の構成について述べる．

LASN はフロアやビルに配置されている複数の PC や WS を結合して，高い並列処理能力を獲得することを可能とする．LASN は SAN（System Area Network）と同等の低遅延と信頼性を持ちつつ，LAN（Local Area Network）と同等の配線長とトポロジの自由度を提供する．

RHiNET-3/SW は，8 個のポートを持つワンチップ CMOS スイッチであり，それぞれのポートは 送受各 10Gbps のバンド幅を持つ．再送機能により信頼性を保証するとともに，低遅延を実現している．また，64 の仮想チャネルを備えることにより大規模なネットワークに対応する．

RHiNET-3/SW: 10Gbps/port 8x8 network switch for LASN

Hiroaki Nishi ,[†2] Ryuichiro Ueno ,[†1] Koji Tasho ,[†5]
Satoru Inazawa ,[†3] Shinji Nishimura ,[†4] Tomohiro Kudoh [†2]
and Hideharu Amano[†1]

In this report, the architecture and the implementation of an LSI switch chip for an LASN (Local Area System Network), called RHiNET-3/SW, is presented. The LASN is a new class of network which enables high performance parallel processing by connecting PCs and WSs distributed on one or more floors of a building. It provides low latency reliable communication comparable to a SAN (System Area Network) as well as relatively free topology design and longer length of links comparable to a LAN (Local Area Network).

RHiNET-3/SW is a one-chip eight ports CMOS embedded array switch. Each port has a 10Gbps bandwidth in each direction, and the chip has 80Gbps aggregate throughput. To realize reliable communication by the physical layer, switch-to-switch retransmition mechanizm is implemented. 64 virtual channels with credit based flow control mechansims are provided to build a large system.

By using a large amount of on-chip memory, RHiNET-3/SW chips form a large sized, low latency, free topology network with reliable communication and a large bi-section bandwidth.

1. はじめに

クラスタによる並列分散処理は，安価な PC/WS クラスタ数十台から数百台でネットワークを構成し，大型計算機に匹敵する非常に高い性能を提供するシステムとして脚光を浴びている．

従来の高性能クラスタシステムは，計算機間の接続に Myrinet[2] などの System Area Network（SAN）を用いたものが多い．SAN は基本的にパケットを廃棄することのない低遅延大容量の通信ネットワークで，多くの並列アプリケーションで要求される通信性能を満たしている．しかし従来の SAN はリンク長やネットワークトポロジに制限があるため，計算機室内などに設置されたクラスタ専用のネットワークとして用いられてきた．

SAN に匹敵する性能を持つネットワークで，フロア内やビル内に分散した計算機群を接続することができれば，日常の業務に用いている計算機の余剰性能を用いて高性能並列処理環境を実現したり，別々の計算

†1 慶應義塾大学大学院理工学研究科計算機科学専攻
Department of Computer Science, Graduate School of Science and Technology, Keio University
†2 技術研究組合 新情報処理開発機構
Real World Computing Partnership
†3 日立通信システム (株)
Hitachi Communication Systems Inc.
†4 技術研究組合 新情報処理開発機構 光インターコネクション日立研究室
RWCP Optical Interconnection Hitachi Lab.
†5 シナジェテック (株)
Synergetech Inc.

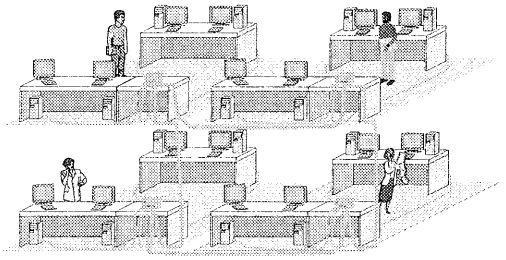

図 1 LASN の例
Fig. 1 Example of LASN

機室に設置された様々な計算機群を接続して統合した処理環境を実現することができると考えられる．

フロア内やビル内の計算機群を接続するためには従来 LAN が用いられてきたが，LAN には，遅延が大きいこと，大きな bi-section bandwidth を提供することが難しいこと，混雑時にはパケットを廃棄する可能性があること，パケットの到着順序が保証されないことなどの問題がある．

そこで我々は，LAN と SAN の両方の利点を追及した，LASN（Local AreaSystem Network）を提唱している [8]．LASN は図 1 に示すような形態をとり，パケットを廃棄しない低遅延大容量ネットワークという SAN の性質を保ちつつ，フロア内やビル内の計算機群を接続するのに十分なリンク長とトポロジの自由度を提供する．

近年の大容量ネットワークでは，光伝送が用いられるようになりつつある．光伝送を用いれば，10Gbps クラスの伝送が比較的容易に実現できる．

我々は，これまでに RHiNET-1/SW[9]，RHiNET-2/SW[8] と呼ぶ LASN 用スイッチを開発してきた．これらのスイッチは，Myrinet で用いられている slack buffer[2] に準拠した方式を用い，100m までの距離についてスキュー調整の必要がなくビット誤り率がほぼ無視できるほど優れている (10^{-20}) 光インタコネクションモジュールを用いた．

よりトポロジ自由度の高い大規模な LASN を構築する場合，このフロー制御方式では，リンクの往復の伝送遅延を見込んだメモリ量が必要となるためリンク長が ASIC 内部のメモリ搭載量で制限されること，光インタコネクションモジュールの構造が複雑で比較的高価であることなどの問題がある．そこで現在，credit based flow control を用い再送機構を持つスイッチ RHiNET-3/SW の開発を行なっている．credit based flow control を用いればリンク長に理論的な限界がなくなる．また，安価な光インタコネクションモジュールはビット誤り率が比較的劣るため再送機構を新たに設けた．

本稿では RHiNET-3/SW の構成について概略を述べる．

2. RHiNET の概要

RHiNET は，PCI バスに装着するネットワークインタフェース RHiNET/NI，ネットワークのスイッチとなる RHiNET/SW およびこれらを接続する光インタコネクトにより構成される [10]．

RHiNET におけるプロセス間通信の記述には，message passing モデル（MPI），共有メモリモデル（Open/MP）の両方を利用できる．ホストプロセッサには，専用の Linux の通信ライブラリおよびデバイスドライバを準備し，実際のプロトコル処理は RHiNET/NI によるリモート DMA で行われる．RHiNET/NI は，様々な並列処理のための通信のサポート機能を持ち，マルチタスク環境での zero-copy 通信を実現する．

RHiNET/NI が提供する並列処理のための通信サポートには，multiple writer protocol のためのハードウェア TWIN メモリ機構，isend/ireceive 待ち合わせ機構などがある [10]．ハードウェア TWIN メモリ機構は，リモートノードからデータをローカルホストのメインメモリへ転送する際に，TWIN メモリと呼ばれるネットワークインタフェース上のメモリにそのデータのコピーを保持しておき，ホスト上でデータに変更を加えて書き戻す際に TWIN メモリの内容と比較し，変更されている部分だけを書き戻すもので，Treadmarks[1] などで用いられる multiple write 機能をハードウェアにより提供する．isend/ireceive 待ち合わせ機構は，MPI の isend/ireceive などで必要となる送信側と受信側の待ち合わせ機構を NI が提供するものである．

RHiNET のスイッチは，LASN 特有の要求事項に対応するため，次のような方針に基づき設計を行う．

(1) レイテンシが大きい store and forward routing や マルチキャスト時のデッドロック回避が困難でチャネルの占有数が多い wormhole routing を利用せず，asynchronous wormhole routing[11] を用いる．

(2) パケットの廃棄や望まない順序の入れかえを行わず，ハードウエアでのエラーレートを低くすることで，上位プロトコル層における通信品質補償に必要な通信コストを極力小さくする．また，パケットの破棄を許さず，通信コストを抑えるため，デッドロックフリーを保障する．

(3) ビル内やフロア内に分散して配置された計算機を接続するために，ループを含むある程度自由なトポロジを許し，また，十分な延長距離を有する．

(4) 並列処理で要求される十分な bi-section bandwidth を確保する．

asynchronous wormhole routing 下では転送パス間にループ異存関係があった場合，デッドロックが生じる可能性がある．そこでパケットを廃棄することなく，デッドロックフリーとトポロジフリーを両立し，かつ柔軟性を持たせるため，縮約構造化チャネル法を

採用している．

構造化チャネル法[7]はネットワークの直径に等しい数の仮想チャネルを用意し，スイッチを経由するごとに異なる番号の仮想チャネル（VC）を用いることで，どのようなルーティングを行ってもデッドロックフリーとなる．構造化チャネル法では，パケットがスイッチを通過するごとに1だけ番号が大きい VC を使用する．このため構造化チャネル法は，ネットワークの規模（最大直径）が VC 数で制限される．

そこで，分岐のないスイッチ（他のスイッチと接続されているポートが2以下のスイッチ）を経由しても，異なる番号の VC を用いる必要がないことに着目して，縮約構造化チャネル法を提案した[9]．この方法では，全てのパケットは，他のスイッチへのリンクを3以上持つスイッチを通過した時にのみ，使用する VC の番号を増やす．これによって必要な VC 数を減らすことができる．

RHiNET にその他のデッドロック回避手法を採用することも可能である．例えば，トポロジフリーネットワークに spanning tree を構成し，up* もしくは down* routing を行うことでデッドロックを回避する手法[3]を RHiNET 上で用いることができる．spannning tree 上で up* もしくは down* routing を行う場合，デッドロックフリーを実現するために必要となる VC 数は1であるが，以下のような問題がある．

- 自由にルーティングを行うことができず，最短距離で通信できないノードの組が存在する．また，ルーティングの制限に伴うホットスポットが発生する可能性がある．
- spannning tree 下での up* もしくは down* routing の設定に必要な，トポロジ内部からプライマリなノードを選定し，そこから spanning tree を構成するという手順が煩雑である．
- ノードの追加削除，もしくはスイッチ障害の度に，spanning tree を張り直すため，張り直す前にネットワーク中に存在したパケットの救済が困難で，大規模システムには適していない．

構造化チャネル法によるルーティングは VC を多く必要とするが，この様な問題は発生しない．

また，RHiNET では，FIFO 性を保証する必要があるため適応型ルーティングを行わないが，RHiNET スイッチ自体はルーティング処理を付属のメインテナンスプロセッサに委ねることで適応型ルーティングにも対応することができる．

2.1 RHiNET-1

RHiNET-1 は LASN の最初のプロトタイプであり，PCI バスに接続するインタフェース RHiNET-1/NI と，ネットワークスイッチ RHiNET-1/SW を 133Mbps × 10bit 幅の転送バンド幅を持つ光インタコネクトで接続した構成を持つ．RHiNET-1/NI は，光インタフェースと，CPLD を用いたプロトコル制御用のハードウェアおよびアドレス変換テーブル，TWIN メモリなどのメモリを搭載している．

RHiNET-1/SW の中心は 0.35μm CMOS エンベッデッドアレイによる1チップスイッチ（ASIC）で，8 × 8 のクロスバを内蔵している．RHiNET-1/SW は自由なトポロジと，長距離のパケット転送に対応するために，外部に大容量の SRAM を持たせ，必要なチャネルのみをチップ内のバッファに割り当てる VCC（Virtual Channel Cache，仮想チャネルキャッシュ）[9]を採用した．この方法により3つの VC を格納するチップ内バッファと，8つの VC を格納するチップ外部バッファを用い，合計で11個の VC をもたせた．

2.2 RHiNET-2

RHiNET-2/SW の中心は，0.18μm CMOS エンベッデッドアレイを用いて構成した1チップスイッチで，VCC による速度低下をなくすため，すべてオンチップで16個の VC をもつ．また，VC を有効に用いるため，仮想ネットワーク機能を搭載し，リンクバンド幅を 8Gbps まで広げた．RHiNET-2/SW ASIC は，RHiNET-1/SW ASIC 同様 8 × 8 のクロスバを内蔵している．

一般に VC は単一ネットワーク上に互いに影響されない複数の仮想ネットワークを実現することにより，パケットのルーティングに優先順位を設けたり QoS を保証するなどの用途に用いられることが多い[4]．RHiNET では，VC をデッドロック回避に用いているが，ネットワークの直径が大きくない場合には，これらの VC を仮想ネットワークの実現に用いることができる．

構造化チャネル法では，本来スイッチを通過するごとにチャネル ID を1ずつ増やすが，RHiNET-2 ではこの増分を任意の値に設定可能とした．RHiNET-2/SW では各入力ポートに16の VC があるので，この増分をたとえば4にすれば，最大4回までスイッチを通過できる（実際には縮約構造化チャネル法により，これより多くのスイッチを通過可能）4つの仮想ネットワークを実現できる．この増分は任意の値に設定できる．

RHiNET のプリミティブは複数の request と acknowledge パケットにより構成されるが，これらの request や acknowledge パケットそれぞれに異なる仮想ネットワークを割り当てれば，request パケットに全く阻害されることなく acknowledge パケットを受け取ることができるようになった．従来ではプリミティブ間のデッドロック回避の為に，ネットワークインタフェースにイベントキューを設け，さらにプリミティブの発行数を制限する必要があったが，仮想ネットワークの利用によりこの制限が取り除かれる．

また，RHiNET-2/SW は，信頼性をあげるため，エラー訂正符号を付加した．

3. RHiNET-2/SW における問題点

RHiNET-2/SW を用いてよりノード数および設置場所の広がりがより大きなシステムを構築しようとすると，次のような問題が発生する．

- 仮想ネットワークの利用時の実効 VC 数:

プリミティブ間のデッドロック回避の為には仮想ネットワークが必要である．RHiNET-2/SW は16個の VC を持つが，RHiNET システムで要求される 4 つの仮想ネットワークに対応すると，1つの仮想ネットワーク当たりの実効 VC 数が 4 に制限され大規模システムの構築が困難である．

- 光インタコネクト:
  RHiNET-2/SW で用いた光インタコネクトは，スキュー調整の必要がなく非常にエラーレートが低いが，100m までの距離にしか対応できず，デバイスが比較的高価であった．RHiNET-2/SW の I/O は，この様な光インタコネクトに特化した設計となっている．
- ハンドシェイク:
  RHiNET-2/SW で採用されているハンドシェイク手法（slack buffer[2] に準拠）では，リンクの伝送遅延が大きくなるほど大きな内部メモリが必要となり 100m 以上の伝送距離に対応することが困難である．

RHiNET-2/SW では，これらの問題を解決するのに必要なハードウエアコストが膨大で実装が困難であった．RHiNET-2/SW では大規模ネットワークの構築可能性よりもレイテンシに重きをおいた設計とした．

4. RHiNET-3/SW

RHiNET-3/SW に搭載される ASIC チップは，リンクバンド幅を 10Gbps に広げた 1 チップスイッチで，ポート数は RHiNET-1, 2 と同じく 8 である．

RHiNET-3/SW は日立製作所デバイス開発センタ製の 0.14μm CMOS エンベッデッドアレイ ASIC を用いる．RHiNET-2/SW で用いた 0.18μm の ASIC と比較して約 3 倍程度の論理を実装できる．

4.1 RHiNET-2/SW における問題点の解決手法および拡張機能

RHiNET-3/SW では，RHiNET-2/SW における問題点を次のように解決し，機能拡張を行なう．

4.1.1 64 Virtual Channel

計算機間を接続して並列処理を行なう際には，4 程度の仮想ネットワークがあることが望ましい．この場合，デッドロック回避のための構造化チャネル法に必要な VC 数の 4 倍の VC が必要になる．そこで，RHiNET-3/SW では 64 （16 × 4）の VC を各入力に設ける．

4.1.2 より安価な光インタコネクションモジュールの利用

RHiNET-2/SW で用いた光インタコネクションモジュールでは，材料を考慮し精密な構造を持つことにより，チャネル間のスキューを 100m までの長さと 800Mbps/channel の伝送速度では問題にならない程度に押えている．また，DC レベルを伝送することができるため，特にコーディングを行なわずに任意のデータを伝送することができる．さらに，10^{-20} と BER(Bit Error Rate) が非常によいため，数千のノードを接続したネットワークでも実質的にエラーフリーであると考えることができる．しかし，材料，構造が精密であるために高価であるという欠点がある．

最近，スキュー値を保証せず，伝送帯域も制限されたタイプの光インタコネクションモジュールが多く使われるようになってきており，構造が単純であることもあり，安価に手に入るようになりつつある．この種のモジュールは BER が 10^{-12}～10^{-15} 程度と劣る．

そこで RHiNET-3/SW では，このような安価な光インタコネクションモジュールを用いるために，外部にスキュー調整とコーディングを行なう ASIC を接続することで対応する．エラーに対しては再送機構を設ける．

4.1.3 credit based flow control の採用

光ファイバは 5ns/m 程度の伝送遅延があるため，ハンドシェイクには，リンクの往復分のレイテンシと回路の動作時間を合わせただけの時間を要する．よって，受信側はパケット送信停止の要求を送信側に送った後でも，この間に受信するデータを受けとらなくてはならない．このため，受信側はリンク中にあるパケットを受信するのに十分な容量がパケットバッファに残っているうちに，送信側にパケット送信をこれ以上行わないように要求する必要がある．これは window によるフロー制御の一種であり，Myrinet で用いられている手法で slack buffer[2] と呼ばれる．RHiNET は構造化チャネル法を用いているため，パケットバッファは仮想チャネルごとに必要であり，送受信のハンドシェイク操作を仮想チャネルごとに行うように拡張する．この方式を拡張 slack buffer[9] および，RHiNET-1,2/SW で採用されている．

この方法ではリンク中にあるパケットを受信するのに十分な容量をスイッチ内部のメモリに持たせる必要があり，長距離伝送と仮想チャネル数の増大を考慮にいれると現在の ASIC では実装が困難となる．そこで，credit based flow control を採用する．RHiNET-3/SW では 8byte の payload を含む line と呼ばれる単位が最小ハンドシェイク単位である．credit based flow control は，line の通信数を管理し，送り先スイッチからいくつの line を処理したという credit を受け取るまで 次の line を出力しないことでハンドシェイクを行う．

図 2 に示すように，2 つのスイッチ A, B で通信した場合を想定する．まずはじめに，スイッチ A はスイッチ B のバッファ容量である 8 を credit として記憶している．line を出力する場合は，出力する line 同数の credit を消費し，credit を使いきると line を出力できないため，スイッチ B のバッファを溢れさせることがない．次に，スイッチ B がバッファから line を取り除くと，取り除いた line 数と同数の credit をスイッチ A に返される．スイッチ A はこの credit を受け取ることで，再び転送を再開する．

4.1.4 再送機構

大規模システムに対応するため，より信頼性を獲得する必要がある．そこで，CRC によるエラー検出と

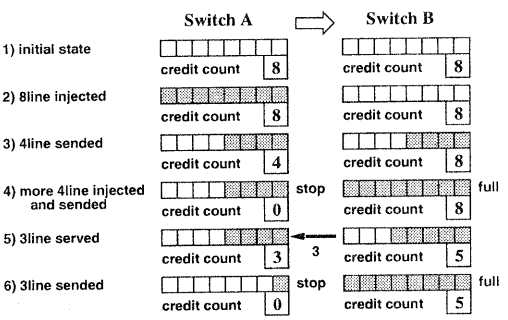

図 2 credit based flow control の手順
Fig. 2 Process of credit based flow control

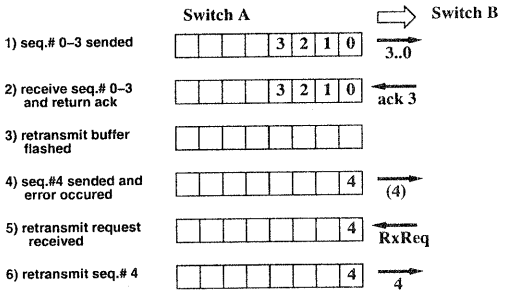

図 3 再送の手順
Fig. 3 Process of retransmitting

シーケンス番号による再送機構を取り入れ，エラー発生時はスイッチ間で迅速にリカバリを行う．

RHiNET-3/SW における再送の基本単位は MicroFrame（以下 MF，2line で構成される）である．有効な情報をもつすべての MF は一定期間内で固有のシーケンス番号（送信シーケンス番号）と CRC がつけられている．リカバリの対象となるのは，bit エラーが混入した MF，およびネットワーク中で消失した MF である．

図 3 に示すように，スイッチから出力される MF は，同時に再送バッファに蓄えられる．エラーが発生していない場合は，相手側スイッチから受信したシーケンス番号（受信シーケンス番号）を受け取ることで，再送バッファから該当する番号までの MF が捨てられる．ネットワーク中で MF に bit エラーが発生した場合は CRC により，また，MF が消失した場合にはシーケンス番号を調べることでエラーを検出し，再送要求を出す．再送要求を受け取ったスイッチは再送バッファの内容をもう一度ネットワークに流すことでエラーのリカバリを行う．

4.1.5 source routing の採用

RHiNET-2/SW は table routing のみサポートしており，source routing についてはノード数を限ればエミュレーションが可能である．一般に SAN では Myrinet 等に代表されるように source routing が採用されており，ユーザの便を考えると大規模システムにおける完全な source routing のサポートを考慮する必要がある．RHiNET/SW で source routing を行えば，どのように source routing エントリをパケットヘッダに記載してもデッドロックフリーを保証できるというメリットがある．

RHiNET-3/SW は 8hop までの source routing を行うことができ，source routing パケットと，従来通りの table routing パケットの混在を許す．

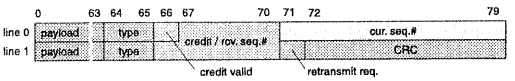

図 4 Micro Frame のフォーマット
Fig. 4 Format of Micro Frame

4.2 Micro Frame

RHiNET-3/SW における I/O 部の bit 幅は 80bit であり，payload として 64bit を確保すると残りは 14bit となる．このフィールドに効率よく情報を載せるため，credit の伝達，および再送に必要な情報は共通の bit フィールドを用いた．図 4 に示すように，1 MF は 2line で構成され，合計で 24bit の領域に line の種別，credit 情報もしくは受信シーケンス番号，その他に，送信シーケンス番号，CRC の情報が載せられる．

4.3 小規模化および高速化手法

RHiNET-3/SW は 64 個ものチャネルを備えるため，すべての資源を 64 個備えると膨大なハードウエア量が必要となる．そこで，出力のチェックを行うロジックなどの，通信と平行して処理可能で遅延にあまり影響しない部位は，必要数の 1/4 のみ備え共用することでゲート数の削減を図っている．共用した資源は時分割で利用されるが，要求のある分割単位には優先的に割り当てるなど高速化に対する配慮がなされている．

また，クロスバのアービトレーションと通信など，並列化して処理できるところは極力並列化している．クロスバのアービトレーションにおいても，最近まで割り当てられたチャネルは今後も利用される可能性が高いため，アービタはそのようなチャネルへは即座に応答できるよう，アービタがロックしたままにするなどの工夫がなされている．

4.4 内部構造

図 5 に RHiNET-3/SW の内部構造を概略示す．スイッチに入力された光インタコネクト受信モジュールからの 1.25GHz 8bit のパケットは，1:10 Demultiplexer により，125MHz 80bit に変換される．その後 elastic buffer により信号を内部クロックに同期させ，受信再送モジュールによりエラーチェックが行われる．

ルーティングモジュールでは，ルーティングテーブルメモリを参照してパケットの出力先，および VC メモリのアドレスを求め，VC メモリに書き込まれる．

VC controller ルーティング情報や VC 番号によりクロスバへのアービトレーションの要求，応答処理，

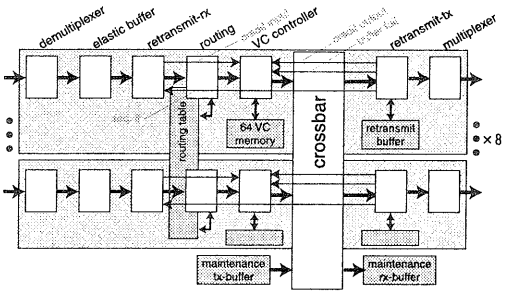

図 5 RHiNET-3/SW の構成図
Fig. 5 Structure of RHiNET-3/SW

エラー処理を行う.

アービトレーションに勝ったパケットはクロスバを抜けて, 送信再送モジュールにおいて, 内部の再送バッファに蓄えられると同時に, 適当なシーケンス番号とCRC が付加されてスイッチから出力される.

4.5 諸　　元

表 1 に RHiNET-3/SW の諸元を示す.

バッファ容量	80Kbyte/link
VC 数	64channel/link
ASIC	日立 DDC 製 0.14μ
バンド幅	10Gbps/port
ポート数	8
レイテンシ	250ns(予定)
論理部総ゲート数	1800KGates(予定)
I/O 動作周波数	1.25GHz
内部論理動作周波数	125MHz

表 1 RHiNET-3/SW の特徴
Table 1 Specifications of RHiNET-3/SW

5. 関連研究

High Performance Networking Forum (HNF) による, Gigabyte System NetworkTM (GSNTM - an HNF trademark, HIPPI6400)[6] に対応したスイッチとして, GENROCO による TSX-8864[5] がある. これは, 光および銅配線の GSN ポートを 8 まで持つことができ, トータルで 8Giga bytes/sec のスループットを持つ. 遅延は 3μ sec である.

6. ま と め

RHiNET-3/SW は credit based flow control を行う 64 個の VC を備え, 構造化チャネル法を利用することで, トポロジフリーネットワーク下でルーティングに依存せずデッドロックフリーを保証する. また, RHiNET-2/SW で用いた光モジュールよりも安価な光モジュールを利用でき, 再送機構の搭載するなど高機能なスイッチであるため, RHiNET-2/SW に比較してより信頼性の高い大規模システムを構築できる. また, 1 チップ ASIC スイッチで構成され, 80Gbps (10Gbps × 8) の通信バンド幅をもつ. 現在 RHiNET-3/SW は論理記述段階にあり, 今年度中に Engneering Sample を受け取る予定である.

謝辞 RHiNET/SW の実装においてご協力頂いた (株) 日立 DDC の大杉 浩三氏, 佐藤和善氏, 日立 CS (株) の原澤 克嘉氏, 圷 重人氏, 福田 周司氏, 日立 IT (株) の大杉 浩三氏に感謝致します.

参 考 文 献

1) Christiana Amza, Alan L. Cox, Sandhya Dwarkadas, Pete Keleher, Honghui Lu, Ramakrishnan Rajamony, Weimin Yu, and Willy Zwaenepoel. TreadMarks: Shared Memory Computing on Networks of Workstations. *IEEE Computer*, Vol.29, No.2, pp. 18–28, 1996.
2) N. J. Boden, et al. Myrinet – A gigabit-per-second local-area network. *IEEE Micro*, Vol.15, No. 1, pp. 29–36, 1996.
3) M.D.Schroeder, et al. Autonet: A high-speed, self-configuring local area network using point-to-point links. Technical Report SRC 59, DEC, 1990.
4) Lionel M. Ni and Philip K. McKinley. A Surbey of Wormhole Routing Techniques in Direct Networks. *IEEE COMPUTER*, Vol. 26, No. 2, pp. 62–76, 1993.
5) http://www.genroco.com.
6) http://www.hnf.org.
7) 堀江健志, 石畑俊幸, 池坂宏明. 並列計算機 AP1000 における相互結合網のルーチング方式. 電子情報通信学会論文誌, Vol. J75-D-I, No. 8, pp. 600–606, 1992.
8) 西宏章, 多昌廣治, 西村信治, 山本淳二, 工藤知宏, 天野英晴. Lasn 用 8gbps/port 8x8 one-chip スイッチ: Rhinet-2/sw. 2000 年記念並列処理シンポジウム (JSPP2000), pp. 173–180, 2000.
9) 西宏章, 多昌廣治, 工藤知宏, 天野英晴. 仮想チャネルキャッシュを持つネットワークルータの構成と性能. 並列処理シンポジウム JSPP'99, 第 99-6 巻, pp. 71–78, 1999.
10) 工藤知宏, 山本淳二, 建部修見, 佐藤三久, 西宏章, 天野英晴, 石川裕. PC 間ネットワークによる共有アドレス空間を持つ並列処理システム. 情報処理学会研究報告 ARC, 第 21- 21 巻, pp. 121–126, 1999.
11) 天野英晴. 情報系教科書シリーズ第 18 巻 並列コンピュータ. 株式会社昭晃堂, ISBN4-7856-2045-5, 1996.