

## 第 5 部

# IEEE1394 によるネットワーク相互接続 実験



# 第 1 章

## はじめに

近年、情報/通信/放送/家電/AV 等の各業界は「デジタル化」、「マルチメディア化」、「ネットワーク化」の大きな動きの中にある。今後これらの各分野は従来の垣根を越えて、「デジタルによる融合」の方向に向かっていくと見られる。現在この「デジタル化」の動きは、パソコンとネットワークの導入を核として、オフィス環境を中心に見られ、今後、更なる高度化、大容量化の方向に発展していくものと考えられる。

これらの動きは、放送や AV、家電を取り込んで膨大な家庭市場に進出してくるのは間違いないと考えられる。例えば、放送のデジタル化、AV 機器のデジタル化、家電製品のネットワーク化への動きがそれらの一例である。

このように、家電や放送、通信などの業界を巻き込んで、デジタル機器が家庭内に入ってくると、それらの機器を相互に接続する要求が出てくるのは自然である。例えばセットトップボックスとデジタル VTR を接続したり、デジタルビデオカメラで映した映像をパソコンで編集する場合等が典型的な例である。この時、従来のアナログケーブルに変わる「デジタルケーブルによる接続」が必要になる。これが将来の家庭内ネットワークの雛形になる可能性がある。

現在、「IEEE1394」が、この「デジタル AV ケーブル」、あるいは「ホーム AV ネットワーク」の有力候補となっている。

IEEE1394 は、数百 Mbps(ビット/秒) の速度を持つ高速バス(ケーブル) である。同期転送の機能をもっているため、遅延や揺らぎがなく、連続的にデータのやり取りを行う必要のある映像や音声の伝送に適している。その他、低コストでの供給が期待できること、プラグアンドプレイの機能があり家庭に受け入れられやすい等の理由から、AV 業界や CATV 業界等が IEEE1394 の採用を決定している。これらと並行して、デジタル映像信号の標準フォーマットである MPEG2 や、DV 等の転送方式も確定され、将来の AV 機器のデジタル接続は IEEE1394 を中心に展開されているといっても良い。

また、IEEE1394 採用の動きは放送/家電業界にとどまらない。パソコンのマルチメディア化を目指すマイクロソフトやインテルも、IEEE1394 の機能をサポートすることを発表している。また、将来のパソコンの拡張スロットの仕様である「デバイスベイ」も、これは IEEE1394 インタフェースを高速マルチメディアインタフェースとすることが確定しており、将来のパソコンのマルチメディア周辺機器の多くは IEEE1394 インタフェースを通

してパソコンに接続されると期待される。

このように、IEEE1394 は「家庭内 AV 網」、「パソコンのマルチメディアインタフェース」、「低コストな高速インタフェース」といった色々な側面を持っている。

家庭のインターネット接続を考える場合、インターネットと AV/放送の融合を考える場合、あるいはインターネット機器を低コストで提供することを考える場合、今後 IEEE1394 は非常に重要な位置を占めると考えられる。

このような背景から、WIDE では 97 年 9 月に、「IP over IEEE1394」WG を発足。「IPv4 over IEEE1394」、「IPv6 over IEEE1394」の検討と相互接続実験を目的とし、活動を行っている。

本報告では、次節にて IEEE1394 と、IETF で検討されている「IP over IEEE1394」スペックの概要の説明を行う。次に、WG 活動の報告として、各社(大学)の実装の説明、相互接続試験(WIDE 内、及び IETF)の結果報告、及び今後の課題と活動予定について説明する。

## 第 2 章

# IEEE1394 と”IP over IEEE1394” スペックの概要

### 2.1 IEEE1394 の技術概要

現在、IEEE1394 の仕様は IEEE1394-1995 として確定しており、現在は IEEE にて各種の拡張仕様が検討されている。

IEEE1394 は、現在 100/200/400Mbps の仕様が定められている。

IEEE1394 の最大の特徴は、同期転送 (QOS 転送) をするモードと、非同期転送 (ベストエフォート転送) をするモードの両方を持つ点である。このように、IEEE1394 は、同期データと非同期データの混在が可能なネットワークとして捉える事ができる。

IEEE1394 は 125  $\mu$ s 毎にタイムスロットが存在する。このスロットへの、一定帯域のパケットの送信権をあらかじめ予約しておくことにより、一定時間内に一定量のデータ送信が保証されることとなり、AV データ等のリアルタイムデータの送信が可能な仕様となっている (同期転送モード)。同期転送モードにおいては、「チャンネル番号」と呼ばれる仮想識別子を付けたパケットを、1394 バス上に伝送することで、通信が成立する (図 2.1)。

これに対して、同期転送で使われていない時間には、ベストエフォートでのパケット伝送ができる (非同期転送モード)。非同期モードでは、read/write/lock のトランザクション (要求、または応答) の形で、パケットの転送が行われる。更に、各々のトランザクションに対して、その受信ノードは、そのパケットを受信したことを送信ノードに知らせるための ack を返す形で通信が成立する (図 2.1)。

ここで、パケットの宛先/送信元アドレスには、イーサネットの MAC アドレス (ROM アドレス) とは異なり、論理的に各ノード毎に与えられた ID (ノード ID) が入る。このノード ID は、「バスリセット」と呼ばれる、IEEE1394 バスの立ち上がり時に一意に定められる値であり、バスリセット毎に値が変わる可能性がある。(なお、IEEE1394 では、ネットワーク上のノードの加入/脱退がある度に、バスの立ち上がりが繰り返される) また、パケットを送信する各ノード間で、公平に送出データ量を制御するフェアアービトレーション機構もある。

IEEE1394 はネットワーク接続される機器同士を専用ケーブルで互いに接続していく形

## 同期通信時のパケットフォーマット

...	チャンネル 番号 (6 bit)	...	データ
-----	------------------------	-----	-----

- 通信に先立ってチャンネル番号と帯域を確保
- パケットに送信元/宛先アドレスは付与しない
- マルチキャストが可能

## 非同期通信時のパケットフォーマット

宛先 ノードID (16 bits)	...	送信元 ノードID (16 bits)	オフセット (48 bits)	...	データ
--------------------------	-----	---------------------------	--------------------	-----	-----

$$\text{ノードID} = \begin{matrix} \text{バスID} & \text{物理ID} \\ \text{10 bits} & \text{6 bits} \end{matrix}$$

- ROMアドレスではなく、論理的なノードIDが入る
- ノードID = all"1" はブロードキャスト
- マルチキャスト機能は無い
- オフセットアドレスは宛先毎に一意に決定(レジスタ空間)

図 2.1: IEEE1394 ヘッダフォーマット

でネットワークを構成する。その際は、ディーゼーチェーントポロジとツリートポロジが許容されている。IEEE1394には、自動構成認識機能も含まれており、終端抵抗の設置等の処置も不要である。このため、ユーザは単純に機器同士を専用ケーブルで次々につなぐ形でネットワークを構成でき、いわゆる「プラグアンドプレイ」が可能な仕様となっている。

また、IEEE1394ではいくつかの拡張仕様を審議中である。このうちのいくつかを以下に紹介する。

- 1394.a

IEEE1394.aは、現在のIEEE1394-1995仕様を改善、強化するものであり、データ転送効率の向上、電源管理、AV機器用小型コネクタの仕様等がこれに含まれる。

”IP over IEEE1394”スペックに含まれる「非同期ストリーム」も、この拡張仕様に含まれる機能である。非同期ストリームとは、同期転送モードのパケットを、予め帯域を確保することなく、非同期転送のフェーズに送出するいわばIEEE1394の3つ目のモードとでも呼ぶべきモードである。IEEE1394は、非同期モードにおいて「ブロードキャスト」と「ユニキャスト」の2つの宛先指定しかできず、しかも「ブロードキャスト」は極力使用しないとの前提で作られている。そのため、IEEE1394上でマルチキャストを実現したい場合や、バス上の特定のノード(例えばIP通信を行うノード)のみにパケットを送出したい場合に、「非同期ブロードキャスト」が許されないため、同期モードで用意されている「チャンネル/マルチキャスト」の概念を非同期モードに適用しようというものと解釈すればよい。

- 1394.b

IEEE1394.b は、現在の IEEE1394-1995 仕様の欠点であるケーブル長 (4.5m) を長距離化するための仕様が含まれている。現在、プラスチック光ファイバ (POF) を始め、IEEE1394 の長距離化 (~ 100m) が検討されている。また、超高速 1394 も、この拡張仕様のターゲットであり、現在 800Mbps、1.6Gbps、3.2Gbps の各仕様が検討中である。

## 2.2 IP over 1394 の仕様

ここでは、IP over 1394 のドラフト仕様の概要を説明する。詳細はインターネットドラフト draft-ietf-ip1394-ipv4-08.txt を参照されたい。

ここで規定される IP over 1394 の主な特徴は次の通りである。

1. 通常の IP パケットは asynchronous のトランザクションとして伝送される
2. IP レベルとは別に、IEEE 1394 レイヤレベルのフラグメンテーション/ リアセンブリを定義 (ATM における AAL のようなものだが、IP over 1394 に固有)
3. ノード識別子として Node ID と EUI 64 の 2 つが存在 (ARP への影響)
4. IP ブロードキャストと ARP には、非同期ストリームを使用
5. そのため、非同期ストリーム用チャンネルを確保するための NPM(ネットワークプロトコルマネージャ) が必要。

IP over 1394 は、基本的に 非同期 write トランザクションを用いて通信を行う。ただし、IEEE1394 においては非同期の (リンクレイヤ) ブロードキャストを最小限にするべしとの方針から、IP のブロードキャスト、マルチキャスト、ARP においても、非同期リンクレイヤブロードキャストは使用せず、前節にて説明した「非同期ストリーム」を使って上記パケットは転送することが決まっている。

非同期ストリームでは、通信に先立ち、接続された IEEE1394 バス上の同期リソースマネージャから専用のチャンネル番号を獲得する必要がある。このため、

- 専用のチャンネルの獲得
- 獲得したチャンネル番号の IEEE1394 バス上の IP ノードへの分配

を目的として、「ネットワークプロトコルマネージャ(NPM)」が必要となる。逆に言うと、このネットワークプロトコルマネージャが確定しない間は、その IEEE1394 上では IP 通信ができないことになる。よって、この NPM の確定 (選定) は、IEEE1394 のバスリセッ

速度	非同期転送時	同期転送時
S100	512	1024
S200	1024	2048
S400	2048	4096

表 2.1: 伝送速度毎のリンクレイヤフレーム長

ト後、すぐに行われる。なお IEEE1394 バスに接続された IP ノードのうち、ノード ID が最も大きいノードが NPM となる。

さて、IEEE1394 における IP パケットの MTU サイズは 1500 バイトと定められている。これに対し、IEEE1394 では、100/200/400Mbps、あるいは将来のより高速な伝送速度と、様々な伝送速度をサポートしている。これらの伝送速度毎 (また、同期モード/非同期モード毎) に、転送できるリンクレイヤフレーム長が異なっている (表 2.1)。

このように、転送できるパケット長と MTU サイズの関係から、転送する IP パケットのサイズによっては、転送時にパケットのフラグメントを行う必要がある。そのための、フラグメントフォーマットが図 2.2 のように定まっている。



図 2.2: フラグメントヘッダ

各フィールドの意味は以下の通りである。

- lf; 表 2.2 の通り



lf	位置
0	Unfragmented
1	First
2	Last
3	Interior

表 2.2: lf フィールドの意味

- ether-type; IPv4, ARP, IPv6 等をこのフィールドで区別する。
- buffer-size; パケット受信する際に必要な受信側のバッファサイズを指定する
- fragment-offset; パケットに対するフラグメントのオフセット
- dgl; 送信するデータグラム毎にインクリメント
- signature; 基本的にノード ID が入る

このうち、dgl と signature の両領域については、IEEE1394 においてフレームの到着順の out-of-order が発生しうるため、これらの両領域をチェックすることで、同一データグラムのフラグメントを判別できるようになっている（同一データグラムについては、dgl と signature の両値も同じ値となる）。

また、ARP パケットの転送フォーマットは以下の通りである。なお、前述のように、ARP も IP ブロードキャストと同様に、非同期ストリーム内を転送されることになっている。

ここで注意すべきは、リンクレイヤアドレスとして、ネットワークカードの ROM アドレスに相当する uniqueID (EUI64 アドレス) と、前述の論理的なノード ID の両方が入る点である。その外、max-rec とはそのネットワークカードが受信できるフレーム長の最大値、spd とはそのネットワークカードがサポートできる転送速度の最大値であり、IEEE1394 に特有の値である。

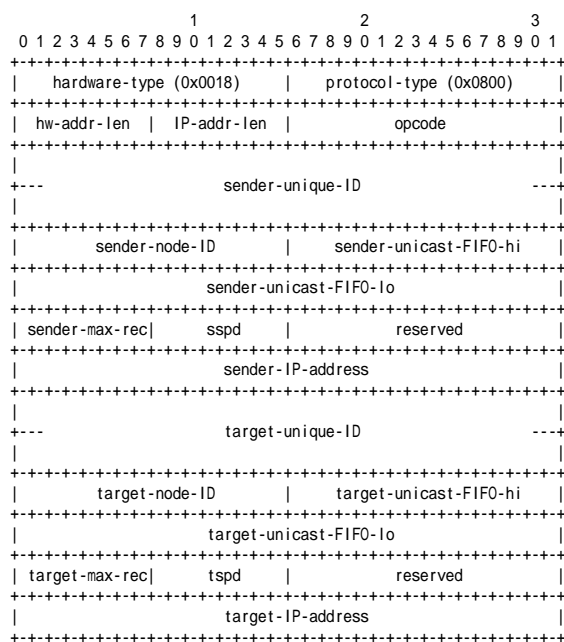


図 2.3: ARP パケットフォーマット

## 第 3 章

### WG 活動の報告

本年度の WG 活動としては、IPv4 over IEEE1394 スペックの実装、及びその相互接続試験 (WIDE 内、及び IETF での試験への参加)、合宿や研究会での BOF の開催、及びメイリングリストでの議論等を行った。本節では、これらの実装の紹介と、相互接続試験の結果報告等の活動報告を行う。

#### 3.1 実装

本節では、WG に参加しているメンバの IP1394 に関する実装について紹介する。

##### 3.1.1 東芝の実装

東芝は、家庭ネットワークの研究プラットフォームとして IP over IEEE1394 の研究開発を行っている [39][40]。東芝の IP over IEEE1394 環境の概要は以下の通りである

- ハードウェアプラットフォーム; PC-AT 互換機
- ハードウェア; 富士マイクロデバイス製 PCI ボード (SA2212)
- OS; BSD/OS バージョン 2.1/3.0
- 実装ソフトウェア; IPv4 over IEEE1394 ドラフトバージョン 7 準拠ソフト

ただし、上記ドラフトでは IP ブロードキャストと ARP について、非同期ストリームを介して行う様に規定されているが、東芝実装ではハードウェアの制約から、これらについては非同期 write ブロードキャストを用いた実装で、これを行っている。今回、東芝が行った実装の範囲は、バージョン 7 ドラフトの全範囲、即ちリンクレイヤフラグメント、ARP、ネットワークプロトコルマネージャの実装である。これまでの実装では、IP マルチキャストや、同期チャンネルへの IP パケットの転送については実装していない。開発した IP1394 のためのドライバのコード量は約 2500 行である (IEEE1394 向けファームウェアのコード量は除く)。

現在までの IP1394 実装の際の問題点を以下にあげる。

### 転送速度

netperf を用いた転送速度の測定結果は約 20Mbps である。測定条件は以下の通りである。

- ホスト 東芝 Equium(MMXPentium200MHz)
- 1394 リンクレイヤフラグメントサイズ 512 バイト
- MTU サイズ 1500 バイト
- TCP ソケットバッファサイズ;8192 バイト
- メッセージサイズ;8192 バイト

現在は、ボード上の DMA 機能を使用しておらず、データを 1 ワード単位で転送しているため、転送速度の向上のためには DMA 転送を考慮に入れる必要がある。

### FIFO とタイムアウト

使用した LSI に内蔵されている FIFO は比較的小さく、Asynchronous/Isochronous の送受信を合わせて 2KB である。この内蔵 FIFO の小ささの問題は、ほとんどの 1394LSI に共通する。

非同期的なコンピュータ通信のためにはバッファ容量が足りないと感じている。2KB のバッファではパケットを 512 byte にフラグメント化してさえ、たかだか 3 個しか入らない (4word のヘッダがある)。しかもこれでは Isochronous にはほとんどバッファを割り当てることができない。

Isochronous を使うことを考えたとき、NTSC の MPEG2 ストリーム 6Mbps を FIFO を 1KB 使ってハンドリングしたとすると 1 秒間の割り込み回数は約 700 回 14ms 弱に一回の割合となり、余裕を取ると割り込みの回数は約 10ms に一回となり、結構な負荷となる。それでも 1394 の帯域 200 Mbps に比べるとここで扱っている帯域はごく小さい。

FIFO の容量はシステムのパフォーマンスを左右するが、現在まで PCI ブリッジ採用の製品も含めて FIFO は 2KB 程度しか実装していないものがほとんどである。コンピュータ通信に使うにはバッファの量を増やすことが望まれる。

### ARP

動的に Node ID が変わるのは問題である。バスリセットによって一斉に ARP 要求が発生してしまう。大規模なネットワークでは、ARP による broadcast storm が上記のバッファオーバーフロー問題を起こす可能性もある。

### 3.1.2 ソニーの実装

ソニーの IP over IEEE1394 環境の概要は以下の通りである

- ハードウェアプラットフォーム; PC-AT 互換機
- ハードウェア; ソニー製 PCI ボード (外販せず)
- OS; FreeBSD 2.2.5
- 実装ソフトウェア; IPv4 over IEEE1394 ドラフトバージョン 7 準拠ソフト

この実装では、非同期ストリームを含め、上記ドラフトの全機能を実装している。これまでの実装では、IP マルチキャストや、同期チャンネルへの IP パケットの転送については実装していない。

### 3.1.3 電気通信大学の実装

電通大で実装中の IP over IEEE1394 環境の概要は以下の通りである

- ハードウェアプラットフォーム; PC-AT 互換機
- ハードウェア; 富士通研究所製 PCI ボード (外販せず)
- OS; FreeBSD 及び BSD/OS
- 実装ソフトウェア; IPv4 over IEEE1394 ドラフトバージョン 7 準拠ソフト

現在は FreeBSD、及び BSD/OS 上にドライバを実装している段階である。チップセットとして、Texas Instrument 社の PCILynx を使用している。

### 3.1.4 富士通研の実装

株式会社富士通研究所は、97年6月に IEEE1394/PMC (PCI Mezzanine Card) ネットワークアダプタボード (200Mbps) を試作し、このボードを用いて asynchronous モードによる IP over 1394 を実装し、ping, telnet, ftp などの IP アプリケーションを動作させた。この結果は、97年7月の IEEE 1394 BOF および 97年9月の WIDE 合宿で発表され [41]、97年11月に福岡で開かれた研究会ではデモが行われた。本実装は IETF の IP1394WG 発足前に行われたため独自仕様である。本章は富士通研の IP over 1394 方式の概略を説明し、97年9月の WIDE 研究会時点での実装評価結果を述べる。

#### IEEE1394/PMC ボード

97 年前半には実験に利用できる IEEE1394 機材が一般に流通していなかったため独自の IEEE1394 ボードを開発した。IEEE1394/PMC ボード (図 3.1) は PMC 規格に準拠したドータボードである。PMC は PCI と電氣的に互換であり、市販の PMC-to-PCI 変換アダプタを利用して本ボードを PC の PCI スロットで使用できる。

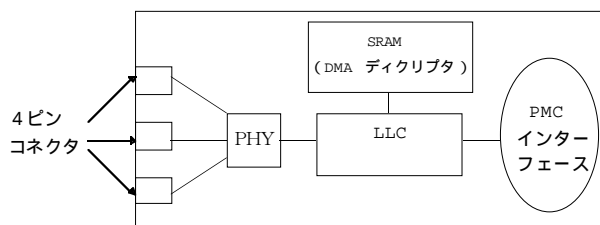


図 3.1: IEEE1394 PMC ボード

チップセットは TI の TSB12LV21 ( LLC ) と TSB21LV03 ( PHY ) を使用し、リンク転送性能は 200Mbps である。この LLC チップは PCI メモリ空間 ↔ LLC 内 FIFO 間の scatter-gather DMA 機能を持ち、例えば IP パケットがヘッダとボディの 2 つの mbuf に分かれていても 1 つの 1394 パケットにフレーミングできる。なお、1394 バスコネクタは電源ペアを抜いた 4 ピンタイプを使用している。

なお、本実装は富士通研が試作した PCI 準拠の Intelligent I/O ( I<sub>2</sub>O ) ボードである Comet に装着して行った。Comet は I/O ボード内部にプロセッサとメモリをもち、ホストプロセッサと独立にパケット処理を行うことができる。

### IP over 1394 実装

IP over 1394 の仕様は以下のとおりである。オペレーティングシステムは BSD/OS であるが、IP over 1394 処理は Comet ボード上で実現されている。

- Ethernet エミュレーション方式

IP packet をイーサフレームに入れてから 1394 パケットにカプセルリングする (図 3.2)。MAC アドレスは EUI64 から生成したダミーアドレスを使用する。1394 asynchronous broadcast write パケットにカプセルリングし、1394 レベルではフィルタリングせずすべての Ethernet フレームを受けようにする。こうすれば、ホストは IEEE1394 バスの存在を意識する必要が無く、バスリセットによるノード ID 変更を気にしなくてよい (ARP も Ethernet ARP を使用)。さらに、100Base-T でも 1394 でも BSD/OS 側のドライバは同じにできる。

- リンクフラグメントなし

1 つの 1394 パケットに 1 つの IP パケット全体を必ず入れ、リンクレイヤフラグメントはしない。必要があれば IP フラグメントを使う。従って、MTU は IEEE1394 規格の最大データサイズである 1024byte (200Mbps 時) までである。

1394 ヘッダ (16byte)	Ether ヘッダ (14byte)	IP ヘッダ (20byte)	IP データ
----------------------	-----------------------	--------------------	--------

図 3.2: IP over 1394 パケットフォーマット

一例として、本実装における IP パケット送信手順は図 3.3 のようになる。

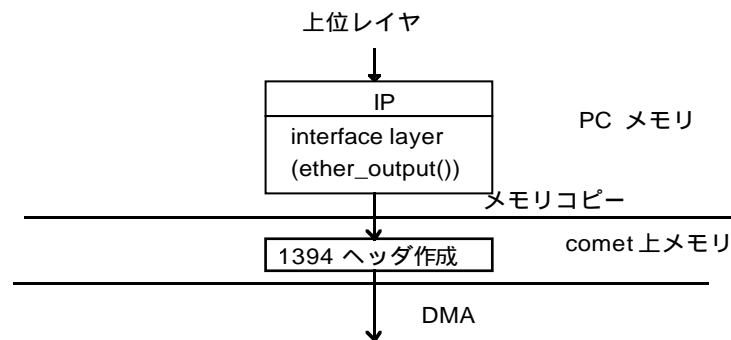


図 3.3: IP over 1394 送信

1. PC 側で IP packet を作ってイーサフレームに入れる。BSD/OS の ether\_output() を利用。
2. パケットを PC のメモリから Comet 上のメモリにプログラム I/O でコピー
3. Comet が 1394 asynchronous broadcast write パケットヘッダ付加。
4. LLC チップが Comet 上のメモリから 1394LLC FIFO へ DMA 転送
5. PHY チップから 1394 バスへデータ送信

Comet メモリ上には、1394 パケットをバッファリングするキューがあり、パケットが到着したかを Comet 上の RISC プロセッサがポーリングして 1 個 1 個処理するというオーソドックスなネットワーク処理になっている。

### ping による実装評価結果

97 年 9 月時点での富士通研実装の実装評価結果を述べる。IEEE1394 ネットワークを 1 hop (point-to-point) で構成し、ping の性能評価を行なった。97 年 9 月時点での初期実装は、現在の方式とは僅かに異なり、asynchronous broadcast write ではなく、asynchronous (non-broadcast) write パケットにイーサフレームをカプセル化する。その際、1394 asynchronous ヘッダのノード ID フィールドには static に固定値を設定していた。

まずパケットサイズを変更して ping した時の RTT ( 300 回の平均 ) を示す ( 表 3.1 )。比較対象として 100Base-T ( DEC 社 DE520 ) を用いた場合の結果も測定した。97 年 9 月の時点では TI のチップにバグがあり、長いパケットを飛ばせなかったため MTU を 256byte としている ( 今はバグが fix されている )。1394 の場合と 100Base-T の場合で Comet 内の処理はほぼ同等で、BSD/OS のドライバは全く同じである。IEEE1394(200Mbps) の方が RTT が短く、IEEE1394 の高速性が発揮された。また、IP ペイロードが 236byte の時と 256byte の時では IP フラグメントの結果 RTT が大きく異なる。この差から IP パケット一つの処理に要する IP 層、ドライバ、ネットワークの基本オーバーヘッドが 200  $\mu$  sec 程度と考えられる。

表 3.1: ping の ROUND TRIP TIME

IP ペイロード (byte)	1394 での RTT(ms)	100Base-T での RTT(ms)
64	0.514ms	0.552ms
128	0.579ms	0.622ms
192	0.610ms	0.667ms
236	0.624ms	0.671ms
256	0.828ms	0.915ms
512	1.063ms	1.126ms
1024	1.484ms	1.554ms

次に、1 つの ping パケットについて、ping をかける側のノードの LLC-PHY 間の信号をロジックアナライザで測定し、リンク性能を調査した。ping -s 248(つまり、IP データサイズは 256byte) コマンド実行における ping パケット送信処理手順およびその所要時間を表 3.2 に示す。

gap 待ちが 23  $\mu$  sec もかかる。実際のデータ転送時間が 15  $\mu$  sec なので、ハードウェアレベルで転送性能の 1/3 強しか出せていない。他のアプリケーションでも大容量データを一方に送る場合、これと同じ問題が起きる。本実験では MTU が小さいため、gap に対するデータ転送時間が短いのだと思われるかもしれないが、現在の IETF 案での MTU でもネットワーク転送性能が 1.6Gbps になれば gap の方が大きくなってしまふ。

#### 考察 1 : リンク層制御方式

ping RTT の結果から、IP フラグメント送信、1394 転送、IP フラグメント受信、IP フラグメント送信、1394 転送、IP フラグメント受信、という一連の処理が本実装では 200  $\mu$  sec 程度であることがわかる。この結果から片方向の IP 転送遅延は 100  $\mu$  sec 程度と考えられる。このように非常に簡単な実装を高速の RISC で on-cache で処理させても制御オーバーヘッドがある程度重荷になることがわかる。IETF の IP1394WG ではリンク層でフラグメントするという議論がなされているが、安易にフラグメントを採用すると性能が出ない



表 3.2: ping パケット送信ノードの処理手順

処理順序	処理内容	内容詳細	所要時間 (μ sec)
1	IP フラグメント 1 個目 (IP ペイロード 236byte)	送信側 LLC チップ arbitration	0.5
2		データ送信	12
3		ack パケット受信	1
4	gap 待ち	subaction gap 検出待ち	11
5		arbitration reset gap 検出待ち	さらに 12
6	IP フラグメント 2 個目 (IP ペイロード 20byte)	送信側 LLC チップ arbitration	0.5
7		データ送信	3
8		ack パケット受信	1

という結果になる可能性がある。この点に関して富士通研は伝統的なパケット通信方式ではなく、メモリベースの通信方式を提案している。1394 は本来バスプロトコルを採用しており、他のノードのメモリ空間に対するリード、ライト、ロックアクセスをサポートしている。この枠組みを積極的に用い、例えば送信側が受信側のバッファをまず必要なだけ獲得し、あとはパケットデータを複数のバストランザクションとして一方的に送り付けるようなことが可能である。このようにすれば送信処理、受信処理ともにほとんど LLC チップがハード的に処理してしまい、パケット転送完了通知のみがソフトに伝えられるので大きな IP データを効率良く転送できることが期待される。今後 1394 はますます高速化するが、一度のフレーム送信に 100 μ sec もの時間がかかっているのはリンク高速化の意味がない。400Mbps でさえ 100 μ sec では 5000bytes のデータが転送される。この性能に対応できるようなリンク層制御方式が望まれる。

#### 考察 2 : gap による性能低下

上述したように、asynchronous 転送の gap をデフォルトのままにしているのは折角の高速伝送性能を活かせない。これに対する解決策としては、gap の長さをデフォルトより小さくする方法がある。例えば、subaction gap は以下の式で与えられるが、

$$\text{subaction gap} = (27 + \text{gap\_count} \times 16) / 100 \mu \text{ sec}$$

デフォルトでは gap\_count が 63 である。この値は PHY チップを設定して小さくできるが、無条件には小さく出来ない。適切な gap\_count はルートからの最大ホップ数に依存して、IEEE1394 仕様書に定められている。従って、バス上のルートノードの位置が重要になる。ルートになるノードがバスリセットで動的に変わらないようにするには、ルートにしたい PHY チップの set\_force\_root ビット をセットせねばならない。実際の運用ではシステム管

理者がどのノードをルートにすべきかを意識しなければならないだろう。

#### まとめと今後の予定

富士通研の IPv4 over 1394 は独自方式で、機能的にも完全とはいえないが、IETF での検討が始まる以前の 97 年 7 月の時点で存在した数少ない動作する実装であり、性能的な観点から評価を行っているところに特徴がある。

現在 IETF で議論されている方式は IPv6 を対象としていないため近い将来には仕様変更を余儀なくされると考えられる。また 800Mbps 以上の高速性能が可能となる将来にむけて、IP 通信性能に重点をおいた検討は十分でない。そこで今後の研究としては IPv6 over 1394 の実装ならびに性能向上のための方式検討に重点をおく予定である。

## 3.2 相互接続試験

98 年 3 月に IETF ロサンゼルス会合に先立ち、ロサンゼルス郊外のシスコ社内にて、相互接続実験が開催された。これには、インテル、ソニー、東芝、サムソンの 4 社が参加した。しかし、サムソン社は実機を持ち込まず、相互接続試験はインテル、ソニー、東芝の 3 社間で行われた（インテルは、IP1394 を Win98 上に実装している）。その結果、3 社間での相互接続が確認され、IETF 会期中に相互接続のデモも実施された（東芝実装は非同期ストリーム未実装のため、デモは行わなかった）。

これに先立ち、ソニー、東芝の両社間での相互接続実験を 98 年 3 月に東芝研究開発センター内にて行った。主な結果は以下の通りである。

- IP ブロードキャストと ARP については、非同期ストリームではなく、非同期 write ブロードキャストにて相互接続した。
- ただし、ネットワークプロトコルマネージャ選択のプロセスについては接続実験を実施した。
- 接続に使用したマシンソニー 1 台、東芝 2 台の計 3 台。
- 相互接続を確認し、ping, ftp, telnet 等のアプリケーションの動作を確認した。
- 両者間のデータの転送速度の計測などは行えなかった。
- ドラフト v7 内の記述に間違い（フラグメントヘッダの値）があり、不整合があった。訂正して対処した。この記述の間違いは、ドラフト v8 においては訂正されている。

### 3.3 合宿や研究会での BOF の開催

#### 3.3.1 97 年 7 月研究会での BOF の開催

97 年 7 月に国立がんセンターで開かれた研究会において、IEEE1394 BOF を初めて開催した。参加者は 19 名。

ここでは、IEEE1394 の概要の紹介、富士通研が実装中の IP over 1394 の説明、参加者の IP1394 関係の活動の様子の情報交換が行われた。ここで、97 年 9 月の合宿にて再度 BOF を開催することとした。

#### 3.3.2 97 年 9 月合宿での BOF の開催

97 年 9 月に浜松で開かれた合宿において、再度 BOF を開催した。参加者は 26 名。

ここでは、まず参加メンバへの IEEE1394 に期待することへの意見交換を行った。それによると、ホームネットワークへの期待、及び低コストな高速ネットワークとしての期待に 2 分された。また、先に開かれた IETF での結果の紹介 (IP over IEEE1394 のスペックの紹介) を行った。

その後、今後の活動予定を討議し、以下のコンセンサスを得た。

- ワーキンググループとしての活動を申請する
- 活動対象として、IP over IEEE1394 の IETF 標準化活動への参画、実装、ならびに相互接続試験を行っていくこととする
- IEEE1394 の非同期転送、同期転送 (QoS)、ならびに IPv4/IPv6 のそれぞれをターゲットとする
- 98 年 5 月に IETF 準拠 IPv4 相互接続試験、99 年 3 月に IPv6 相互接続試験を目標スケジュールとする (目標時期については、98 年 3 月の合宿で変更された)

その後、ボードの了承を得て、97 年 9 月に WG としての活動を開始した。

#### 3.3.3 97 年 11 月研究会での BOF の開催

97 年 11 月に九州大学で開かれた研究会の Router BOF にて、富士通研版 IP over 1394 のデモを行った。同時に 1394 を用いたネットワーク実験について検討し、1394 over IP をやることを決めた。

### 3.3.4 97 年 12 月研究会での BOF の開催

97 年 12 月に慶應大学三田校舎で開かれた研究会において、BOF を開催した。参加者は 14 名。

ここでは、先だって行われた IETF ワシントン会合の報告と、マルチキャストについての議論が行われた。特に、マルチキャストについては、チャンネル番号の割当てと、「フローとチャンネル番号の対応関係の通知」についての議論、及び同期通信との関係についての議論が行われた。

### 3.3.5 98 年 3 月合宿での BOF の開催

98 年 3 月に浜松で開かれた合宿において、BOF を開催した。参加者は 25 名。ここでは、

- 最新スペック (07) の紹介
- 実装状況の報告 (東芝、富士通研、ソニー、電通大、ITRC)
- 相互接続試験 (東芝、ソニー) の結果報告
- 今年度の総括と来年度の予定

について、報告、及び議論を行った。

来年度の活動として、ワーキンググループを存続させるか、あるいは家庭ネットワークを活動内容に含めるか等も含めた議論された。その結果、本 WG での活動は、IPv4/IPv6 over IEEE1394 に閉じるものとし、その実装と相互接続を中心に行っていくことを確認した。また、ターゲットスケジュールも、「次回の合宿 (98/9) で v4/v6 の相互接続試験を行う」ことを目標に再設定した。

### 3.3.6 98 年 3 月合宿での 1394 over IP 実験

97 年 11 月研究会 (Router BOF) の結果をうけて、98 年 3 月合宿の合宿ネットワークに 1394 over IP 試作を組み込み、Digital Video(DV) 画像 2 チャンネルの中継を行った。

## 3.4 その他の活動

### 3.4.1 家庭ネットワーク

97 年 7 月の研究会において、研究論文「レジデンシャル環境におけるネットワーク相互接続方式」を発表した [39]。ここでは、家庭内ネットワークの要求条件をまとめ、複数

網環境における映像や音声転送を念頭に置いたシグナリング方式を提案した。ここでは、IEEE1394 の同期転送機能に着目し、ラベルスイッチ技術の制御に用いられる FANP[42] を通信資源確保ができるように拡張することで、複数網をまたがる通信資源の予約を行う。IEEE1394 の場合には、FANP にて確保するデータリンク資源(仮想チャネル資源)として、同期チャネルを用いる。シグナリング(いわゆる C プレーンデータ)には IP を用いるが、転送されるデータ(U プレーンデータ)は IP パケットとは特に限定しない点が特徴である。

### 3.4.2 1394 over IP の実験

IEEE1394 の高速な伝送性能(現在 400Mbps まで規格化、将来的には 3.2Gbps まで拡張予定)および同期転送モードによる 125  $\mu$  秒単位の高精度な QoS 機能は従来のネットワークにはみられないものであり、この高速性、機能をローカルエリアネットワークに応用することが期待されている。

ただしここで問題になるのは IEEE1394 のスケーラビリティである。本来 IEEE1394 は PC とディスクやビデオ装置などの周辺機器を接続する高速シリアルバスとして開発されたため、ケーブル長が最大 4.5m に限定される、一つの物理ネットワークに高々 64 個しかノードが接続できないという制限がある。

この制限を回避する為に Ethernet など既存のネットワークとのゲートウェイ(IEEE1394GW)を用いて遠隔の IEEE1394 ネットワークを結合する方法が考えられる。IEEE1394GW では 1394 パケット(同期、非同期)と IP や Ethernet フレームとのプロトコル変換を行うが特に同期パケットは変換遅延を一定にする(ジッタを低減する)必要がある。ジッタが大きいと、このジッタを吸収するだけのバッファリングをしなければならず、通信のリアルタイム性を損なうことになるからである。ところが従来の計算機ネットワークやルータなどでプロトコル変換を行う場合、アダプタからホストに割り込みを上げ、プロトコル変換を行い、アダプタにデータを送り返すという一連の処理が必要で、これを 125  $\mu$  秒という高精度で行うのは難しいのが現状である。

株式会社富士通研究所はプロトコル変換などの複雑なネットワーク処理を高速に実行できるネットワークアダプタ(Comet)を開発してこの問題を解決する実験を行った。Comet は、同時に二種類のネットワークインターフェース(ATM、100Base-T、Gigabit Ethernet、IEEE1394 から任意に選択できる)を持つプロトコル処理専用エンジンを搭載した PCI カードで、PC やワークステーションに搭載して使用する。

Comet ゲートウェイを用いて IEEE1394 同期転送モードのデジタルビデオ(DV)信号を IP 中継してビデオ再生を行なう実験が 98 年 3 月の WIDE 合宿(浜松)で行なわれた。実験の概略は以下の通りである。Comet に IEEE1394 と 100Base-T を装備し、DV 装置を IEEE1394 で接続する。DV カメラは 125  $\mu$  秒毎に 500byte のデータを同期転送モードで流すので、これを受けて IP カプセル化し、100Base-T 側からインターネットに送出する(IEEE1394 over IP - 図 3.4)。受信側では逆変換を行い、同期パケットを IEEE1394 バス

に転送、これをリアルタイムで再生する。Comet 内の処理時間合計は 40  $\mu$ 秒 / パケットであった。

合宿では二つの BOF 会場にそれぞれ DV カメラを設置し、この画像を合宿ネットワーク ( Gigabit Ethernet ) を介して配送、別の二箇所を受信表示したが、合宿期間中を通じて安定して動作した。また合宿後に同様の構成で ATM ヘゲートウェイする実験も行われ、広域 ( 接続距離 40Km ) でも良好な結果が得られた。

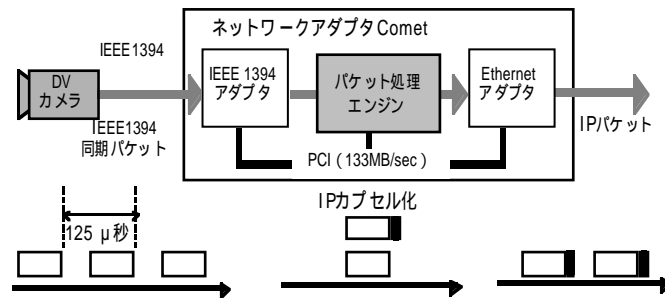


図 3.4: IEEE1394 同期パケットの IP カプセル化

## 第 4 章

### 今後の予定

前述のように 98 年 3 月の合宿において、今後の WG の予定について議論した。ここで、現在のチャータ (IPv4/IPv6 over IEEE1394、及び同期チャネルの利用) を再確認し、次回合宿での IPv4/IPv6 の相互接続実験を主な WG のターゲットとすることを確認した。そのための仕様検討の場としても、WG/メイリングリストを活用していく予定である。

また、ホームネットワークの議論については、その対象が特に IEEE1394 に限定されないことから、この WG での活動対象とは特にしないことを確認した。

