

500MHz高速光伝送の計算機入出力インタフェースへの適用

5M-6

伊藤 隆弘 長坂 庄市 斎藤 成一
三菱電機(株)

1 はじめに

複数計算機による二重系/多重系システムでは、各中央処理装置間での処理の引き継ぎや監視などを目的としてデータ共有装置、例えば磁気ディスクによる大容量記憶装置や半導体メモリによる高速アクセス記憶装置を使用する。

このような多重系システムにおいて、高信頼性を確保するには災害耐性向上も重要な課題である。即ち火災や事故による局所的な装置の故障や停電が即座にシステムダウンに到らない為の対策が必要とされる。

そのために各中央処理装置やデータ共有装置の設置場所を別棟に分散配置したり供給電力システムを分離するなどシステム構成の自由度を高める方策が求められる。電気的絶縁対策も重要である。これらに対する最も有効な実現手段は光伝送方式である。

光伝送の特長として、次の点があげられる。

- (1) 長距離伝送
光ファイバの低損失性により無中継長距離伝送が可能。
- (2) 高速データ伝送
光ファイバの広帯域特性により高速データ伝送が可能。
- (3) 電気的分離
光ファイバは電気絶縁体であるため設置間電位差対策と電気的ノイズの影響拡大防止が図られる。

これらの特長を利用して、中央処理装置とデータ共有装置間の入出力インタフェース部に伝送路周波数500MHzの光伝送システムを適用した。

本稿ではこの光伝送システムについてその技術的課題と具体的な実現方式について報告する。

2 技術的課題

(1) 高速データ転送の実現

中央処理装置と共有データ装置間のデータ転送では伝送路ネックにならないだけのデータ転送性能が要求される。光伝送部自体は十分な高速動作が可能であるが、その高速性を保ちつつ計算機入出力インタフェースとマッチングをとるための伝送プロトコルの検討が必要である。

(2) エラー検出方式

多重系システムの信頼性を向上するには部品単体の故障率の低減だけでなく、故障を確実に検出してリカバリ処理を行ない、システムとして継続運転する必要がある。そのためにはケーブルの断線やノイズによる瞬時エラーの検出及びその処理方式を実現する必要がある。

(3) 実装技術

計算機内部はマイクロプロセッサやLSIを含む高速デジタル回路で構成されており同一基板上に高速アナログ回路を搭載する場合は両者の相互干渉を防止する手段が必要である。回路設計技術とともに機構設計を含めた実装技術を検討する必要がある。

以上の課題に対する実現方式についてディスク装置への適用例を紹介する。

3 実現方式

3.1 システム構成

図1に多重系システム構成例を示す。各中央処理装置(CPU1---, CPU_n)はデータ共有装置であるディスク装置に対し、互いに非同期でアクセスできる。中央処理装置とディスク装置間インタフェースに光伝送を使用している。

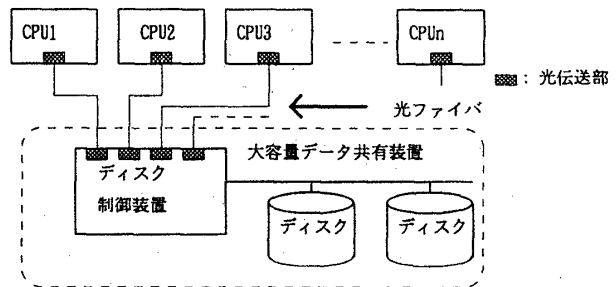


図1 システム構成図

図1に示す光伝送部は送信部と受信部から構成され、全二重シリアル伝送を行なう。光伝送部ブロック図を図2に示す。送信、受信ともデータの一時蓄積が可能なようにFIFO (First In First Out memory) インタフェースをとっている。

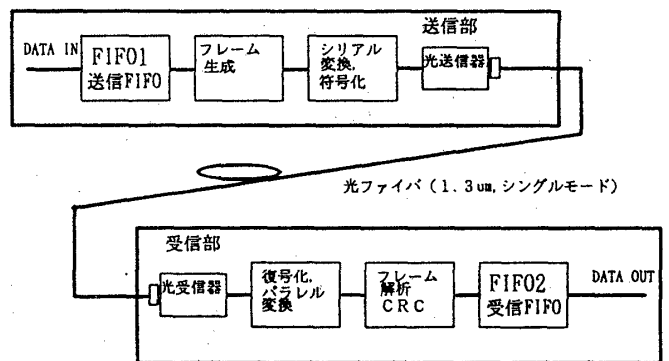


図2 光伝送部ブロック図

3.2 高速データ転送の実現方式

本システムでは1回の最大データ転送量が2KB、データ転送スループットとして20MB/s以上の性能目標とした。この高速データ転送を実現するために独自のフレーム処理方式を採用した。

ディスク装置へのデータ書き込みを例として動作概要を説明する。中央処理装置からの書き込みデータがFIFO1に格納されると、同時にフレーム生成部に対してデータ転送要求が出される。フレーム生成部はFIFO1からデータを読みだし、図3に示すような制御コード付のフレームを生成する。データ転送前のアイドル状態時は制御コード"00"を付加したアイドルデータが伝送路上を流れ、これにより相手側との同期が常に保たれている。1回に転送するデータの最後にはCRCデータが転送され、これがフレームの区切りとなる。次にこのデータをシリアル変換する。この際CMI (Coded Mark Inversion)方式による伝送路符号化を行なう。パラレル変換時の同期情報として各データの先頭ビットに対しCRV (Coding rule violation)変換をかける。

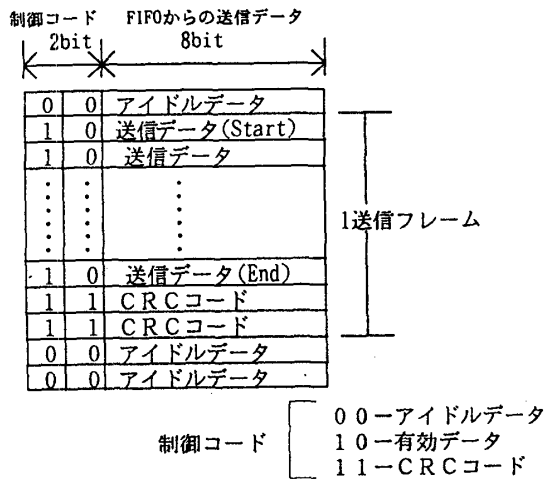


図3 フレーム構造

符号化されたシリアルデータは光送信器で光変換され光ファイバへ送り出される。受信側では送られてきたシリアルデータを光受信器によって電気信号に変換する。その後、CMI符号化されて制御コード付10ビットの平行データとしてフレーム解析部へ出力される。制御コードを解析し、有効データのみを受信側FIFO2に格納すると同時にこの有効データに対するCRCコードを生成する。

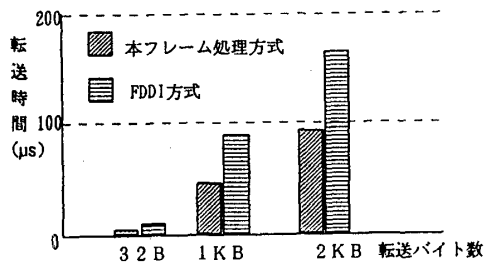
送信データ中のアイドルデータは受け捨てとなる。CRCコードが検出されたことで当該フレームの終了を認識し、CRCによるエラーチェックを行なう。

このフレーム処理方式の特長は以下のとおりである。

- (1) 光伝送部FIFOに対するデータ書き込み/読み出しと光伝送路上のデータ転送が非同期かつ並列動作できるため、転送オーバーヘッドが小さくなる。
- (2) データ転送中に送信側FIFOが空になった場合にはアイドル情報を流すことによって受信側をウェイト状態とすることが可能なので送信データが全て用意されていなくてもデータ転送が開始できる。

本データ転送性能は次のとおりである。

光伝送路周波数は500MHz。CMI符号化を行なっているため実効転送レートは250Mbpsとなる。フレーム構造が示すように10ビットデータ中2ビットが制御コードなので有効データ転送レートとしては25MB/sとなりディスク装置要求を満足する性能を得ることができた。尚、25MB/sはバックボード上のバス転送速度に匹敵する。比較例として計算機間光伝送として現在実用化されているFDDIと本フレーム処理方式との転送速度比較結果を図4に示す。



3.3 エラー検出方式

本システムで適用した光伝送路媒体の誤り率は 10^{-11} と非常に低い。が、部品故障など、万一エラーが発生した場合のエラー検出能力を高めるために次のエラー検出方式を併用した。

- (1) ビットエラー検出
CMI符号化における符号化規則チェックを行なうことで伝送路上での1ビット誤りを検出する。
- (2) 同期外れエラー検出
平行変換時の同期はずれ検出のためにCMI符号化方式で使用されるCRV方式を用いる。
- (3) CRC
フレーム単位にデータの正当性をチェックする為にCRCを行なう。
- (4) 受信断検出
光受信器の光パワーをチェックして光ファイバの断線、劣化や部品故障などを検出する。

3.4 実装技術

光伝送部はアナログ及びECL信号による500MHzの高速度動作回路で構成される。これを数10MHzのデジタルで動作する高密度な計算機制御カードに搭載する必要がある。両者の干渉や基板特性による影響を確認するためにSPICEシミュレーションと高密度実装に対する事前評価を行ない論理設計と基板設計へフィードバックした。500MHzクロックを例として図5に対策前の反射による歪み波形(a)と対策後の波形(b)を示す。

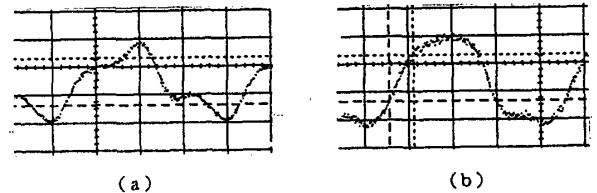


図5 500MHzクロック波形

具体的な実施項目を以下に示す。

- (1) 論理設計への反映
 - ・ファンアウトは1とし信号の分岐は行なわない。
 - ・クロックは差動で供給。
 - (2) 基板設計への反映
 - ・50オームのマイクロストリップライン層構成。
 - ・ECL系とTTL系のグラウンドのスリットによる分離。
 - ・スルーホールを介した別層への渡りを禁止。
- その他、最大パターン長制限、パターン曲り角度、平行線長に対する基準値もシミュレーション結果にもとづき設定された。以上の実装方式に従って製作されたボード例を写真1に示す。

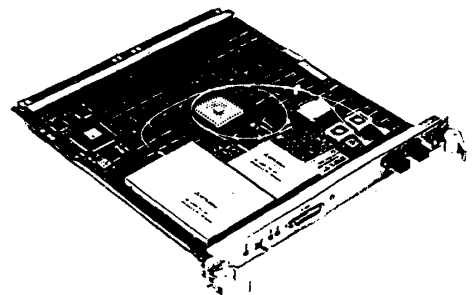


写真1 光伝送部搭載ボード

4 おわりに

計算機とデータ共有装置間の入出力インタフェースに伝送路周波数500MHzの光伝送システムを適用した。この分野では現在1GHzクラスのインタフェースも標準化が進められており、今後さらに高速なデータ転送方式を検討していく予定である。