

# リフレクティブ・メモリによる リアルタイム・ネットワーク

文書バージョン : 2010年1月  
文書番号 : GFT-630A



## リフレクティブ・メモリとは何か？

リフレクティブ・メモリ・ネットワークは、複数の独立したコンピュータがデータの共通セットを共有することを可能にするよう設計された特殊な共有メモリ・システムです。リフレクティブ・メモリ・ネットワークは、各接続システムに共有メモリ・セット全体の独立したコピーを置きます。各接続システムは、ローカル・メモリへのフル書き込み速度でこのローカル・データのセットにアクセスし、変更する完全な無制限の権限を有します。

データがリフレクティブ・メモリのローカル・コピーに書き込まれる時、高速ロジックは、図に図示されたように、リング・ネットワーク上の次のノードに同時に送ります。続く各ノードは、同時にこの新しいデータをそのローカル・コピーに書き込み、リング上の次のノードに送ります。メッセージがオリジナルのノードに戻った時、ネットワークから除去され、特定のハードウェアとノード数に応じて、ネットワーク上の各コンピュータは、数  $\mu\text{sec}$  以内に同じアドレスで同じデータを有します。ローカル・プロセッサは、ネットワーク・アクセスなしに、このデータをいつでも読み込むことができます。このスキームにおいて、各コンピュータは、常に共有メモリ・セットの最新のローカル・コピーを有します。示された4ノードの例では、全コンピュータがリフレクティブ・メモリに書き込まれたデータを受信するのに、 $2.1 \mu\text{s}$  かかります。\*

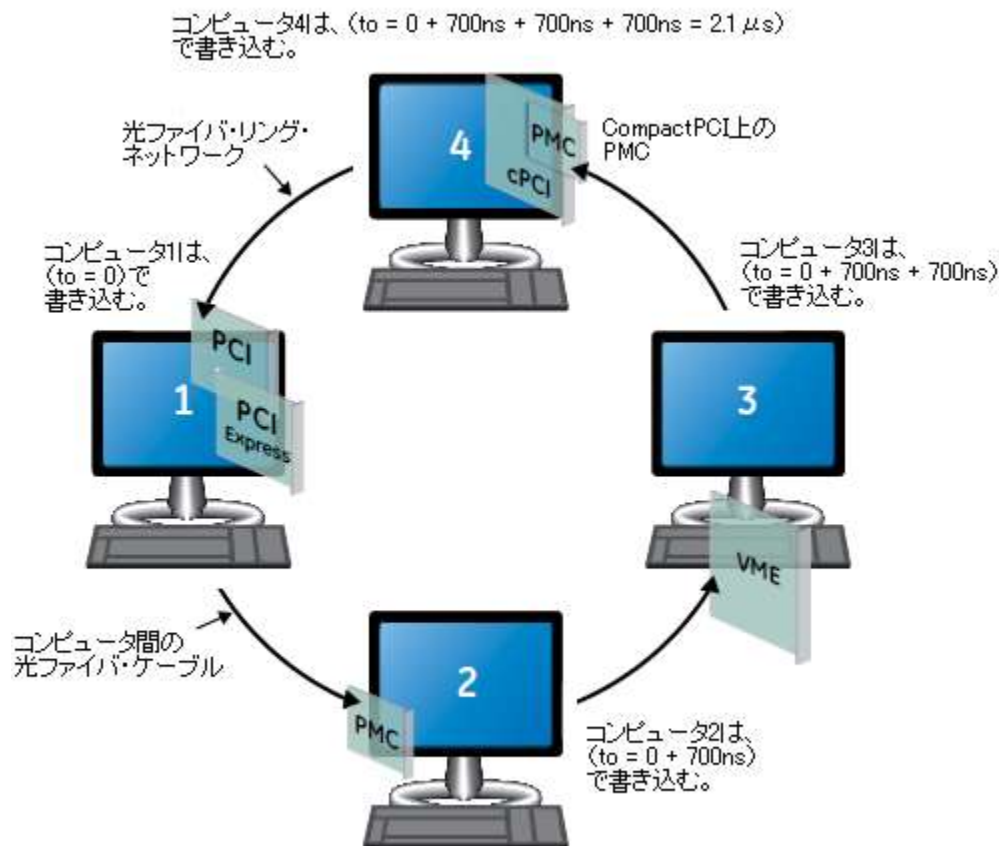


図1 リフレクティブ・メモリは、ネットワーク間の遅延が非常に低い。

\*この遅延は、ネットワーク・トラフィックがなく、ケーブル長が短く、最大の packetsize が可能なものと仮定して計算されています。ケーブル長とネットワーク・トラフィックは、遅延を増大させ得ますが、ネットワークの帯域を超過しない限り、遅延は、余り増加しません。

リフレクティブ・メモリ・ボード(ノード)は、ローカル・メモリ、組み込みインタフェース、及びホスト・コンピュータとリフレクティブ・メモリにアクセスを提供する調停ロジックから成ります。リフレクティブ・メモリ・ボードは、VME、及びPCI/PCI-X、CompactPCI、PCI Expressを含む各種コンピュータ・バス、又はPMCサイトをホストできる標準若しくは専用システムに物理的に設置又は接続できます。これにより、その相互運用性に関係なく、ほとんどのワークステーションとシングル・ボード・コンピュータをリフレクティブ・メモリ経由で接続することができます。

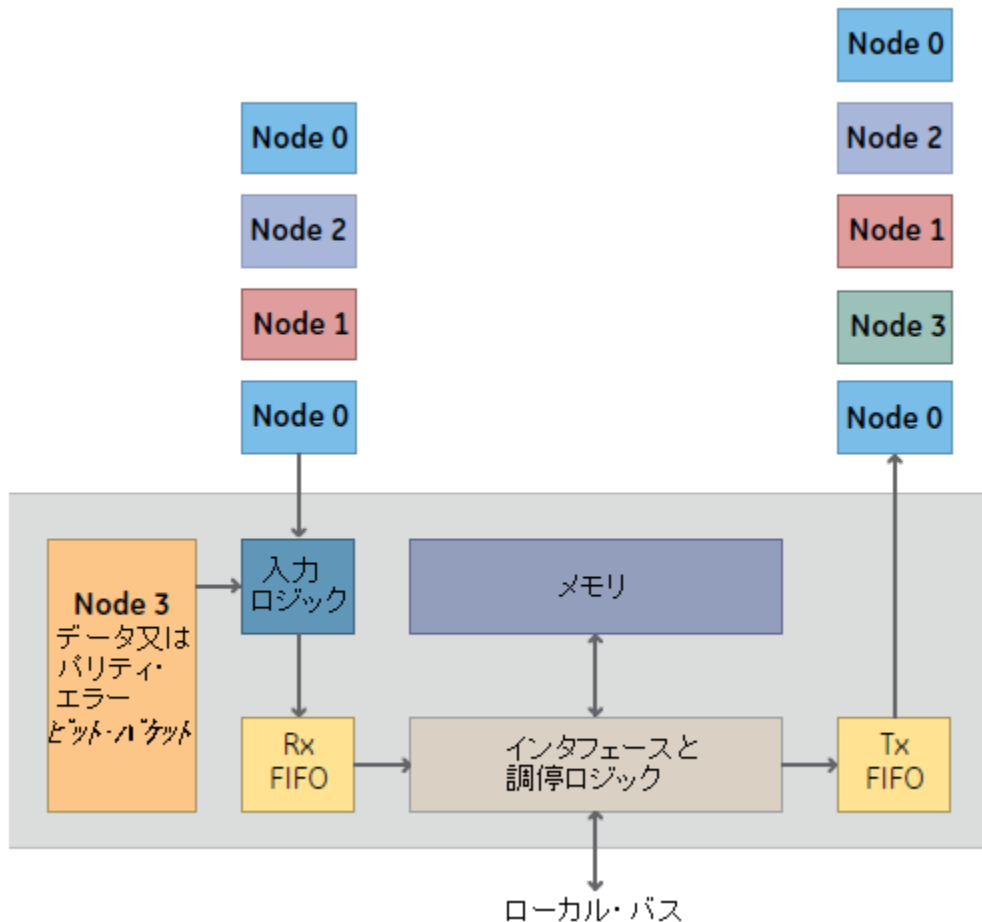


図2 リフレクティブ・メモリのデータ挿入。ネットワークからのデータは、組み込みロジックによりローカル・メモリに自動的に書き込まれ、次のネットワーク・ノードに送信される。

### どこでリフレクティブ・メモリを使うのか？

リフレクティブ・メモリは、コンピュータ又はProgrammable Logic Controllers (PLC) を一緒に接続するためにイーサネット、ファイバチャネル、又はその他のシリアル・ネットワークを使用するアプリケーションで使用できますが、全てのアプリケーションに理想的という訳ではありません。リフレクティブ・メモリは、リアルタイムでの相互作用が主な関心事であるシステムに最も適切です。

決定論、低遅延、及び高速通信が必要なシステムにおいて、典型的により低性能のハードウェアよりも高価なリフレクティブ・メモリ・ボードは、使用容易さの利益に加えて、性能に莫大なリターンをもたらします。

### どのようにリフレクティブ・メモリを使うのか？

リフレクティブ・メモリ・ネットワークの使用は、シンプルな数ステップのみを要求します。

- リフレクティブ・メモリを利用できるバックプレーン・スロット(VME、PCI Express等)に接続するか、PMCサイトを有するシングル・ボード・コンピュータ又はキャリア(VME、CompactPCI等)に接続し、ケーブルを接続。
- メモリに書き込む(リフレクティブ・メモリのグローバル・メモリは、標準RAMとしてコンピュータに現れます)。
- メモリを読み込む(ネットワーク上のリフレクティブ・メモリ・ボード)。

リフレクティブ・メモリのような低ソフトウェア、高速、ハードウェア・ドリブン・ネットワークの利益は、ネットワーク全体と個々のノード間の極めて低いデータ遅延です。この低遅延性能は、シミュレータのようなリアルタイム・システム構築時、最も重要です。

### 誰がリフレクティブ・メモリを使うのか？

リフレクティブ・メモリは、以下のものを含む数百のアプリケーションで使用されています。

- 航空機シミュレータ
- 自動試験システム
- 艦船及び潜水艦シミュレータ
- アルミニウム圧延制御/監視
- 発電所シミュレータ
- エンジン試験台
- 工業プロセス制御
- 高速データ取得
- 超地平線レーダー
- PLCユーザ

### 何故、リフレクティブ・メモリを選ぶのか？

リフレクティブ・メモリLAN又はリアルタイム・ネットワークは、リフレクティブ・メモリ・ボードの以下の1つ以上の特性により解決される必要性又は問題を設計者が抱えている時に通常構築されます。

- [決定論的データ転送](#)
- [高速性能](#)
- [使用の容易さ](#)
- [OS及びプロセッサ独立性](#)
- [経済性と利用できるtime-to-buildシステム](#)

- [標準LAN技術に対する長所](#)

#### 決定論的データ転送

リフレクティブ・メモリは、ハードウェア・ベース・ネットワークです。ノードに転送される全データは、ローカル・メモリに記憶され、他の全てのノードのメモリに自動的に配列されます。データ転送と関連したソフトウェアの遅延はなく、ハードウェアの遅延は最小です。遅延は、ハードウェア・レベルで科され、best-to-worstケース遅延の非常に小さな枠内で事前決定することができます。2つ以上のノード間の通信が完了する時間が保証されたリフレクティブ・メモリの決定論は、システム設計者が堅実な時間の枠内でデータ配送を保証できる効果的なリアルタイムLANを構築することを可能にします。このことは、シーケンシャル・アクションの保証されたスケジューリングを可能にし、データが失われないことを保障します。

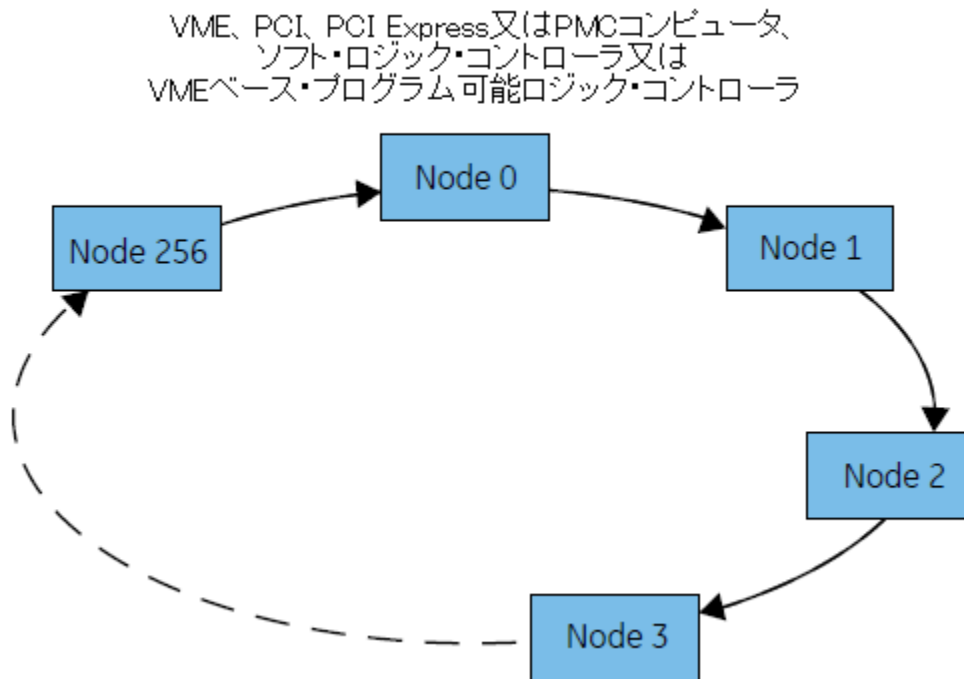


図3 リフレクティブ・メモリ・リング・アーキテクチャは、リアルタイムで最大256の独立ネットワーク・ノードを接続

#### 高速性能

リアルタイム性能の需要は、リフレクティブ・メモリ・ネットワークが、非常に高速で動作しなければならないことを要求します。いくつかの場合、この速度とその生み出すスループットは、リフレクティブ・メモリを使わずにいられない理由です。例として、GE Intelligent Platformsのリフレクティブ・メモリ製品は、動的パケット・サイズ(4~64 byte)、最小CPUオーバーヘッドのためのDMA能力、最大256 Mbyte DRAM、2.12 Gbaud光ファイバ接続、及び持続170 MB/sの転送速度を提供します。このことは、ほとんどのデータ・ネットワークに対して、速度とスループットの両面で有利です。

### 使用の容易さ

リフレクティブ・メモリほど設置及び運用が容易な高性能LANは、他にありません。理想的なネットワークは、各コンピュータが各コンピュータのメモリに同時アクセスすることを可能にします。リフレクティブ・メモリは、 $\mu$  sec単位で最大256台の他の全てのコンピュータのメモリの実コピーを各コンピュータ与えることにより、この理想に近付けています。このメモリはグローバルである以上、アクセスは、複数のコンピュータで同時に行えます。この共通メモリ空間への全CPUの書き込みアクセスは、ネットワーク内の全ノードに複製されます。リフレクティブ・メモリは、アプリケーションがソフトウェア・オーバーヘッドのペナルティなしでデータを共有できるように、このデータを透過的に監視し、複製します。

### OS及びプロセッサ独立性

リフレクティブ・メモリ・ハードウェアは、現在、VME、PCI/PCI-X、PMC、PCI Expressその他の各種フォーム・ファクタで利用できます。これにより、異なるバスを単一のリフレクティブ・メモリ・ネットワークで接続することができます。PMCスロットを有する組み込みシングル・ボード・コンピュータ(VME又はCompactPCI)は、ホスト・バックプレーン上のトラフィックをオフロードするために、リフレクティブ・メモリPMCボードを使用できます。ユーザは、どのデスクトップ・ワークステーション、シングル・ボード・コンピュータ、又はサーバも、CompactPCI、PCI/PCI-X、PCI Express、VMEバス・スロット、又はPMCサイトを利用できるどのコンピュータとも直接情報を共有できる高速ネットワークを構成することができます。

異なるbyteフォーマット(big-及びlittle-endianタイプ)を使用する異なるタイプのコンピュータを接続する時ですら、byteスワップは、リフレクティブ・メモリ・システムで問題ありません。PCIベース・リフレクティブ・メモリ・ボードは、スワップの言語を翻訳するために明確に設計されたハードウェアを含めることができます。このハードウェアは、迅速、効率的かつ反復可能な双方向変換を提供します。また、big-とlittle-endianタイプ間の変換に対して、プロトコル・オーバーヘッド又は時間的ペナルティはありません。リフレクティブ・メモリは、使用するOS又はマシーンに関係なく、同じ使用容易さを有します。

### 経済性と利用できるTime-to-Buildシステム

システム設計者は、より短時間でより能力があり、複雑なシステムを構築するよう、ますます求められています。そのような場合、ハードウェアのコストは、ハード/ソフトウェアのインテグレーションに費やす時間も、ソフトウェアと比較した時、最小です。これは、特に単一又は小規模なシステムで真です。市場への時間が緊要な側面である場合、リフレクティブ・メモリ使用時のシステムの構築と運用に要求されるソフトウェア人時への低い投資は、極めて有益かも知れません。これらの需要のあるシステムにおいて、リフレクティブ・メモリのシンプルな通信の読み書き方法は、市場への時間を十分改善し、並びにネットワーク上のデータ・スループットを改善します。

### 標準LAN技術に対する長所

リフレクティブ・メモリは、グローバル・メモリ、高速データ転送のような標準ネットワークに対する様々な機能、及びソフトウェア透過性を提供し、リフレクティブ・メモリをマルチコンピュータ通信に魅力的なソリューションにしています。従来の通信アプローチに対する追加の開発時間、試験、整備、文書及び特別なCPU要求のコストと比較して、リフレクティブ・メモリは、費用対効果のある

選択肢を提供します。

## アプリケーションの例

### フライト・シミュレーション

低遅延性能の重要性を示す1つの例は、複数の参加者が複合的な人工環境下で独立の変数として参加するインタラクティブな戦闘フライト・シミュレータです。このタイプのシミュレーションにおいて、独立の各コンピュータ・システムは、表示の生成、シミュレーションへの個々の参加者の入力の管理、地形又はその他の環境様相の生成、兵器システムの管理、又は他の各種機能を担当するかも知れません。複数の独立した参加者が極めて高速で高度に流動的な動きを遂行しているため、システムが現実の正確かつ真に迫ったシミュレーションを再現するのに十分な頻度で更新することが肝要となります。

GE Intelligent Platformsのリフレクティブ・メモリ・ネットワークにおいて、リフレクティブ・メモリ・ネットワーク上に置かれたメモリ書き込みは、次のコンピュータが700 nsec以下でデータとして受信することができます。これは、ローカルRAMに書き込まれ、次のリフレクティブ・メモリ・ボードに送信された時間から測定されます。

2機のシミュレートされたF-22が巡航速度1,070 mphで飛行していると仮定してみましょう。シミュレーション内のあるパイロットが過激な機動を行い、他のパイロットに回避行動を強いる場合、シミュレーションに迅速に再現されなければ、シミュレーションのリアリズムは失われるでしょう。

更新遅延を最小化することにより、リフレクティブ・メモリは、シミュレーションの完全性を保障します。ネットワークの速度は、シミュレーションが人間の参加者の知覚認識の精査に耐えることを可能にし、各機の兵器、制御及び航法システムのような電氣的「参加者」のサービス対するより高度な要求に合致します。

### アルミニウム圧延

リフレクティブ・メモリ・リアルタイム・ネットワークは、PLC制御アルミニウム圧延のオペレーションを向上させるために使用されました。分散PLCは、3,500 ft/分の圧延を監視及び制御するために使用されました。PLC制御ループは、アクチュエータが応答する前に、2~3 ftのアルミニウムを通過させてしまう分解能の応答を有していました。これらのアクチュエータは、厚さを可変させるために、アルミニウムに対する圧力を加算/開放するために応答していました。

リフレクティブ・メモリを使用して、独立のVMEベース・システム間の通信のシステムが、この応答時間を短縮するために開発されました。2~3 ftの分解能は、4 inchesにまで縮小され、莫大な浪費の削減と最終製品の品質改善をもたらしました。圧延に関連するデータは、PLCにインポートされ、リフレクティブ・メモリのメモリに直ちに書き込むことにより、独立のVMEコンピュータ・システムに送り、複雑な制御アルゴリズムを転送します。システムは、リフレクティブ・メモリに単にコマンドを書き込むことにより、PLCに出力制御データを送り返します。このアプリケーションにおいて、データは、PLC上のローカルのリフレクティブ・メモリにほぼ即時的に現われます。データ転送は早く、計算は、PLCがその制御ループを遅延なく走らせるくらい早くなります。このように、PLCは、工場の制御デバイスとリフレクティブ・メモリのメモリを読み書きするだけなので、その最大スキャ

ン速度で動作します。

### ロケット・エンジン試験台

複数のマイクロプロセッサを堅固に結合する分散コンピュータは、分散システムが構築と使用に経済的であるため、人気が成長しています。同時に、分散アプローチは、しばしば、単に最も実践的な選択です。ほとんどの分散アプリケーションは、各々特定の独立タスクを処理する専用コンピュータの部分に分解されます。このことは、非常に複雑なタスクでも、分割とコード化を容易にします。設計者は、コンピュータを戦略的に置くか、既存の空間の制約に合わせて置くこともできます。

例えば、ロケット・エンジン試験台は、各種パラメータを測定するために、数百台の変換器を使用します。オペレータは、試験への遅滞ない接続を必要としますが、安全上の理由のため、機器/視覚センターは、3,000 m離れて位置するかも知れません。適用を分散することにより、設計者は、データをデジタル化し、事前処理する試験台にコンピュータを設置することができます。そして、3,000 mに及ぶ数百のディスクリート線の代わりに、1つの高速リフレクティブ・メモリ・ネットワーク・リンクが、制御室内のメイン・コンピュータにデータを送り返すために要求されるだけです。この遠隔コンピュータは、その後、試験オペレータによる閲覧用にデータを分析、アーカイブ化、フォーマットし、モニターに表示します。

高速リフレクティブ・メモリ・リンクを使用することにより、オペレータは、接続により科される最小の遅延で、起こった変化を観測し、反応することができます。不安定な試験から安全な距離に制御スタッフとコア処理コンピュータを置くことにより、オペレータは、試験パフォーマンスを低下させることなく、人員と機材へのリスクを最小化することができます。

### 原子力発電所シミュレータ

原子力発電所シミュレータは、時折極めて遠距離に広がる著しい量の分散計算力を要求し、リアルタイムで不安定な大量のデータを共有する必要がある成長株の複合アプリケーションの代表です。選択肢のソリューションより遥かに低い遅延と高度の決定論を提供し、システムをより高速かつ堅実にする高度にスケラブルで、シンプルなソリューションであるため、これらのアプリケーションの多くは、リフレクティブ・メモリに切り替わっています。

リフレクティブ・メモリは、13のシミュレータによる起こり得る全ての条件下で、発電所の理解を発展させ、操作を学ばなければならないドイツの約2,000人の学生の訓練用に選ばれた技術です。シミュレータ訓練は、法的に要求され、オペレータが必要とする専門家のスキルを獲得し、恒久的に維持することにおいて、重要なコンポーネントであり、ドイツの原子力発電所の安全な運用に大きく寄与しています。

シミュレータは、リアルタイムで管理される最大27,000 I/Oを見ます。3つのVME Intelligent I/O Controller (IIOC) システムは、VME光ファイバ・リフレクティブ・メモリ・インタフェース経由で一緒に結合され、VME Data Communications Controllerにリンクされます。リフレクティブ・メモリ・インタフェースは、170 Mbyte/sの速度でData Communications ControllerとIIOCにデータを動かします。各IIOCは、I/Oスキャン及び更新、起動、リアルタイム、及びオフライン診断をサポートするために、



プロセッサとファームウェアを含みます。Data Communications Controllerは、リフレクティブ・メモリ・インタフェースを経由してホスト・コンピュータと通信します。システムは、ホストから多くのオフロードI/O関連タスクを可能にするIIOCで、ホスト・コンピュータとのデータの転送と同時にI/Oデータの並行処理をサポートします。

### シンクロトロン放射施設

先進の実験科学は、エレクトロニクスにおける物質の構造の調査のような「サブナノメートル」レベルで日常的に働いています。この小型化レベルで鍵となる技術は、電磁放射の形態である放射光です。放射光は、たんぱく質の結晶学、断層写真術、写真平板法、X線及び残留応力分析のような多くの種類の研究及び産業に理想的なツールであり、生命科学、医療、材料科学、分子環境科学、石油科学及び化学産業において応用されています。

リフレクティブ・メモリは、シンクロトロンのストレージ・リングにおける粒子のジッタを制御するため、フィードバックを迅速に提供するシンクロトン・アプリケーションで使用されています。放射光は、軌道内の磁力の微小変化のため動いている間、ジッタが生じるでしょう。リフレクティブ・メモリは、粒子が常に周回していることを保障するため、ジッタを制御するシステムで使用されています。粒子は、非常に高速で動くため、システムは、リアルタイムで、迅速に反応しなければなりません。

### 要約

リフレクティブ・メモリは、データ取得及び処理制御から先進シミュレーションにまで及ぶタイムクリティカルなアプリケーションにおいて、データを共有する最適な方法です。リフレクティブ・メモリ・ネットワークは、低遅延と決定論的性能に対してほとんどの通信技術を上回るリアルタイム・ネットワーク能力を提供します。リフレクティブ・メモリ・ネットワークは、複数の離れて位置するノードがリアルタイムで単一のデータ・セットを共有することを可能にするため、最小の更新遅延とアクセス制限なしでシステムを接続します。

GE Intelligent Platformsは、リアルタイム・ネットワークを専門とし、ほとんどのノードをリアルタイムで一緒に接続する完全なソリューションを開発しています。GE Intelligent Platformsのリフレクティブ・メモリ製品ラインは、使用するOS又はプロセッサから独立した決定論的性能を提供します。

*メモリを読み、メモリに書く。いたって、シンプルです。*

## 別紙1 リフレクティブ・メモリの機能と実践の詳細

### リング・アーキテクチャ・ネットワーク - データの衝突なし

GE Intelligent Platformsリフレクティブ・メモリ製品ラインは、光ファイバ・リング上で2.12 Gbaudの速度で動作するデータ挿入、リング・アーキテクチャ・ネットワークを提供します。リフレクティブ・メモリは、ほとんどのイーサネット・システムがそうであるような衝突ベース・バス調停システムではないため、データ・パケットのキューイングとチェックに要求される複雑さを回避します。

リフレクティブ・メモリ・リングは、適切な接続性を保障し、追加の負荷制限又は終端要求を科しません。ノード間の距離は、最大10 kmに及びます。これらの特性は、Gigabitイーサネットの約100 MB/s(プロトコル・オーバーヘッドを除く)と比較して、リフレクティブ・メモリに対して170 MB/sのデータ転送速度を可能にします。

### 割り込み能力により、ノード間のデータ同期が可能

リフレクティブ・メモリは、ネットワーク・ノードが他のノード又は全ノードに割り込むことを可能にします。最大4つの異なる割り込みを割り当てることができ、機能、優先度、及びベクトルに関して、割り込み毎にユーザ定義可能です。これらの割り込みは、メールボックス・割り込み又はネットワークに渡るデータ同期のような機能に対して使用できません。有効時、割り込みは、典型的に、データが転送された後、ネットワーク上のあるノード又は全てに割り込むために使用されます。

割り込みの使用は、要求されません。リフレクティブ・メモリは、その割り込みを無効にした状態で起動します。このことは、割り込みが必要とされない限り、リフレクティブ・メモリをソフトウェア透過にし、ドライバ・ソフトウェアを要求しません。

### バス・サポート

GE Intelligent Platformsの170 MB/sリフレクティブ・メモリ製品ラインは、VMEバス、PMC、CompactPCI、PCI Express及びほとんどのPC、ワークステーション及びシステムで使用されているPCIアーキテクチャ、並びにVMEベースPLCシステムでも利用できます。

### データ転送

リフレクティブ・メモリ光ファイバ・リング・ネットワーク上において、転送されるべきデータは、他のシステム情報と共にパケット内に置かれ、ノードからノードにパスされます。各ノードは、前のノードから受信したデータを記憶した後、パケット内のデータをネットワーク内の次のノードに再送信します。ノードによるデータ挿入は、いつでも起こり得ます。

データ・フローのパケットは、各ノードを通過するストリームで、リングを回ります。各ノード内において、パケットは分解され、エラーがチェックされ、データは、リフレクティブ・メモリに記憶されます。パケットは、その後、再編され、次のノードにパスされます。ローカルCPU又はDMAデバイスにより、データがリフレクティブ・メモリに書き込まれる場合、新しいパケットが形成され、ノードを流れるパケットのストリームに挿入されます。

データ・パケットがリングを完走し、オリジナルのノードに戻った時、認識され、ネットワークから除去されます。結果は、ネットワーク内のコンピュータ・メモリの全てへのデータ送信を完了するため、最小のハードウェア遅延で高度に効率的なリアルタイムLANとなります。

#### エラー管理技術は、優れたデータの完全性を提供

GE Intelligent Platformsのリフレクティブ・メモリ・ボードは、広範なエラー検知と通知設備を有します。ネットワークのエラー率は、システムの光部分で生み出されたエラーの率に関係します。この光エラー率は、光ファイバ・ケーブルの長さタイプに依存します。ノードがエラーを検知した時、誤った転送は、システムから除去され、実装されている場合、割り込みが生成されます。

この些細なエラーの潜在性ですら受け入れ難いシステムに対して、リフレクティブ・メモリは、各転送が2度送信される冗長転送モードで運用することができます。この動作モードにおいて、2つ目の転送が使用されている場合、エラーが検知されない限り、1つ目の転送が使用されます。両転送でエラーが検知された場合、ノードは、システムから転送を除去します。

光ファイバ・リフレクティブ・メモリ・ボードは、データがリングを適切に走っていることを確認するために使用できるネットワーク監視bitも用います。このbitは、ネットワーク遅延を測定するために監視することもできます（データがネットワーク上の全コンピュータのメモリに転送され、記憶される時間）。

#### 自動光ファイバ・バイパス・スイッチは、ネットワークの信頼性を向上

GE Intelligent Platformsは、ノードが故障した場合、リフレクティブ・メモリ・ネットワークの運用を継続できる光ファイバ・ハブを生産しています。これらの製品は、ノードの不具合がネットワーク全体をクラッシュさせないことを保障するため、動作が停止した場合、ネットワーク・ノードを自動的に迂回します。ハブは、カスケードすることができ、最大256ノードを有する管理ハブ・アレイを可能にします。各ポートは、シリアル光信号を再生成し、挿入損失とケーブル減衰の問題を除去します。信号の再生成は、ジッタも減少させます。

#### OS

各種OSのドライバ・サポートに関して、GE Intelligent Platformsのインストール・プログラムにより、ユーザは、各種プロセッサ/OS/ハードウェアの組み合わせを選択することができます。

## 別紙2 従来の分散コンピュータ用のネットワークの弱点

システム間でメッセージ又は大データのブロックを転送する多くの方法が存在し、各方法は、それ独自のユニークな能力と制限を有します。最もシンプルなデータ転送技術は、あるコンピュータから別のコンピュータのバックプレーンにCPU読み書き信号を転送するため、バス・リピーターを使用します。2つ目の技術であるDirect Memory Access (DMA)は、2つ以上のコンピュータのグローバル・メモリ間でデータを動かします。DMAは、ローカル・プロセッサからのバックプレーン制御を要求します。他の方法は、単一の共有グローバルRAM、並びにイーサネット及びGigabitイーサネットのような標準LAN経由でパスされるメッセージを含みます。

### バス・リピーター

バス・リピーターは、示されたように、あるコンピュータのCPUバックプレーンを別のコンピュータのCPUバックプレーンに接続します。この接続により、CPU間でメッセージをパスすることが可能になり、各CPUが他のコンピュータのリソースにアクセスすることも可能にします。バス転送がいつでも、コンピュータ1と2の間のどの方向でも起こり得る以上、バス調停サイクルは、2つのシステム間の読み書きサイクル毎に要求されます。

このアプローチの問題は、CPUがリモート・バックプレーンのリソースにアクセスしたい時毎に、先ずリモート・バックプレーンへのアクセスを要求した後、アクセスが認められるまで待機しなければならないことです。リモート・バックプレーンで行われる他のバス活動の優先レベルとタイプに応じて、これは数 $\mu$  secから数msecかかるかも知れません。このオーバーヘッド遅延は、データ転送を妨げるだけではなく、バックプレーンの要求を停止させ、リモート・アクセスが完了するまでこのバックプレーンの他の活動をブロックします。システムが互いに待つ時間が多くなればなるほど、積み重なる遅延は、リアルタイム・アプリケーションにとって法外なものとなります。

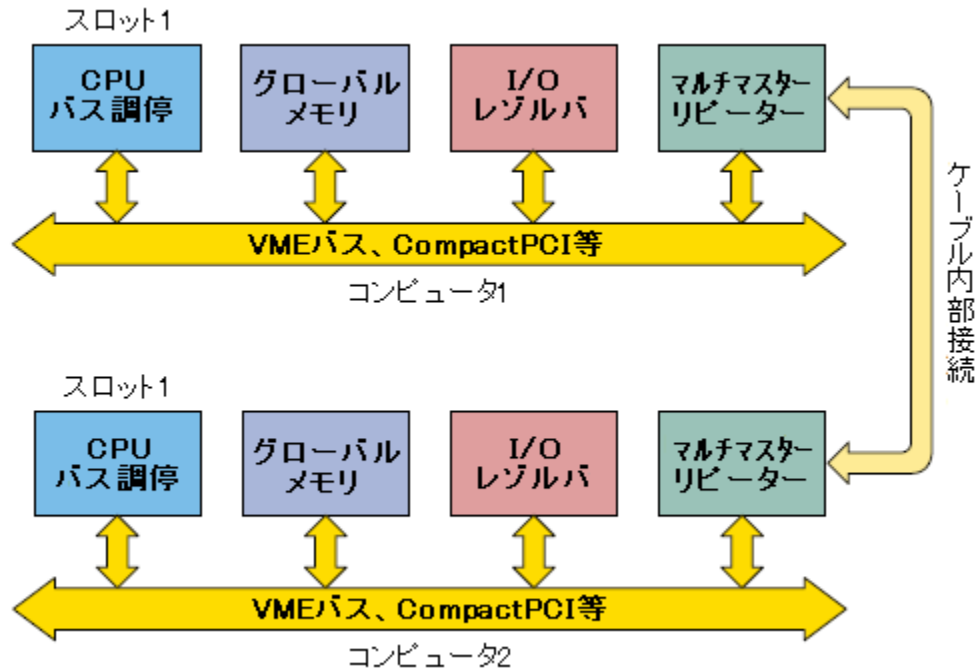


図4 バス・リピーター接続

#### Direct Memory Access (DMA)

バス・リピーターは、少量のデータ(数byte又はワード)をバックプレーンからバックプレーンに動かすには非常に効果的です。しかしながら、多くの分散マルチプロセッシング・システムでは、より大量のデータがブロック・パラメータの形態で各種CPU間で交換されます。これらのデータのブロックは、示されたように、DMAコントローラ・ボードを使用して、より効率的に動かすことができます。

これらの接続において、各システムのCPUは、それ自身のDMAコントローラ・ボード上のアドレス・レジスタとサイズ・レジスタを初期化します。このプロセスにおいて、オリジナルのDMAコントローラ・ボード上のアドレス・レジスタは、DMAコントローラがグローバル・メモリからパラメータ・ブロックを読み始めるところを示します。宛先DMAコントローラ・ボード上のアドレス・レジスタは、DMAコントローラがパラメータ・ブロックを記憶し始めるところを示します。一旦2つのCPUがその各々のDMAレジスタを初期化すれば、転送は自動的となり、CPUは、他の活動に注意を向けることができます。

DMA転送は、非常に高速で起こり得ます。つまり、一旦上記のオーバーヘッド・プログラミングとセットアップの全てが行われれば、送信元と宛先の両コンピュータは、データ転送にアクティブな関係を有しなければなりません。最も重要なことは、ブロック転送が完了する時毎に、次の転送を準備するため、DMAコントローラを設定できるように、両プロセッサが割り込まなければならないことです。これらのDMA転送が起こっている間、各ローカル・プロセッサは、そのDMAボードで利用できるバス帯域を共有しなければなりません。このセットアップは、一定の環境下では有

効ですが、頻繁な更新は、遅延を科す頻繁な割り込みを要求します。DMAとメイン・アプリケーション間のバス大域の分割は、ホスト・アプリケーション、並びにDMAプロセスに対するデータ・ボトルネックを作り出し得ます。

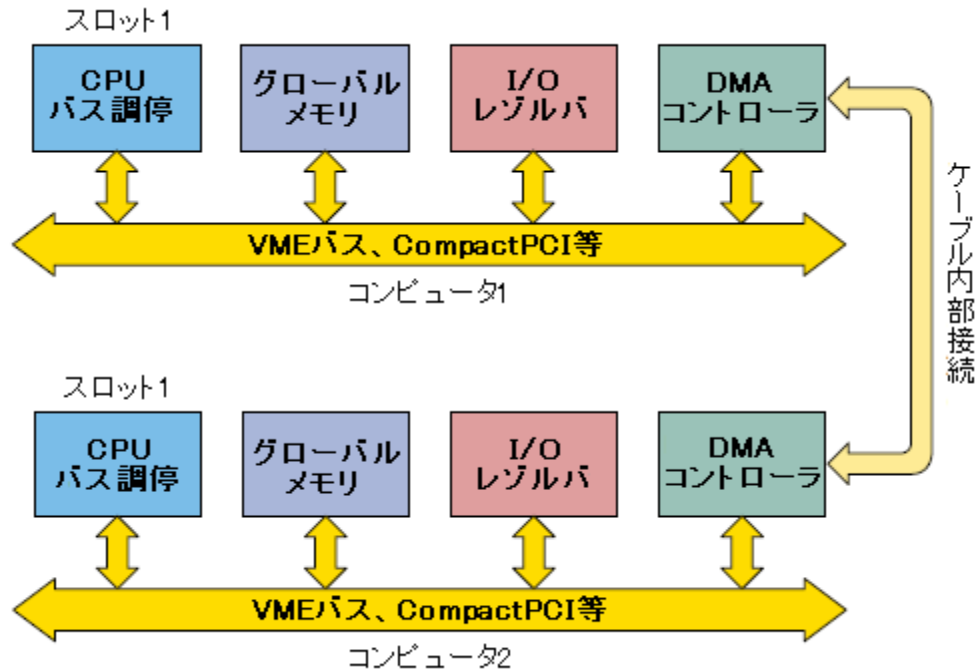


図5 DMAコントローラ接続

#### 共有(グローバル)メモリ経由のメッセージ・パス

3つ目の構成は、図示されたように、2台以上のコンピュータにグローバル・メモリの単一セットを共有させることです。典型的な共有グローバル・メモリ・シナリオは、同じバックプレーン・ベース・シャーシ(通常、VMEバス又はCompactPCI)に存在する2台以上のコンピュータです。これらのコンピュータの各々は、それ自身のメモリを有し、アクセスは、プロセッサのフル速度で起こります。コンピュータは、事前確立されたメッセージ・プロトコル・スキームを利用することにより、同じバックプレーンに存在するグローバル・メモリ・セットを経由して、互いに通信し、データを共有することができます。

このタイプのシステムにおいて、グローバル・メモリは、基本的に、数台のコンピュータ中で共有されるシングルポート・メモリであり、同じシャーシ内に存在する全てのコンピュータがアクセス可能ですが、このリソースへのアクセスは、調停されなければなりません。また、プロセッサ間通信は、典型的にローカル・メモリへのアクセスよりも遅い、バス・メモリ・カードの組み合わせの速度で行われます。個々のコンピュータは、情報の共有を促進する1つの貴重なリソースを争い、プロセッサが共有メモリへの自由なアクセスを有する時ですら、速度は低下します。

通信は、外部的に分散したコンピュータがリピーター、DMA、又はLAN経由でシングルポート・

グローバル・メモリに接続される時、より厄介なものとなります。トータルデータの遅延は、各プロセッサがメモリへのアクセスの順番を待たなければならないので、積み重なることとなります（新しいデータの書き込みでも、グローバル・メモリ経由の他のコンピュータからのメッセージの受信でも）。このシナリオにおいて、データ遅延（全てのコンピュータが新しいデータへのアクセスを得られる前にかかる時間として広義に定義できる）は、すぐに制御不能となり得ます。

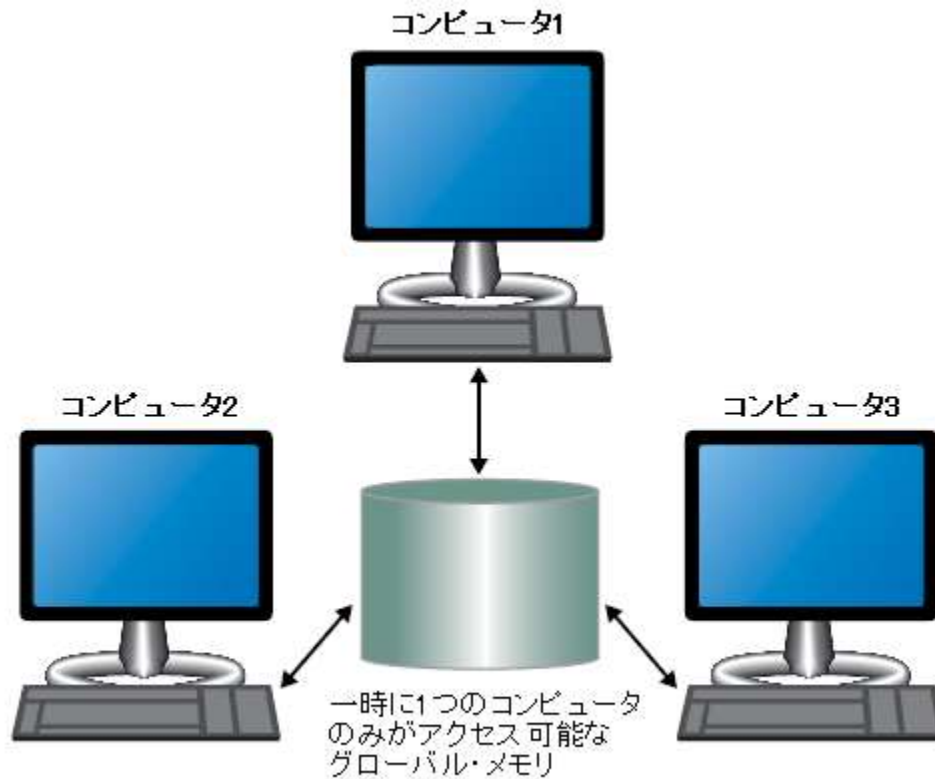


図6 グローバル共有メモリ・アーキテクチャ。  
一時に1つのノードのみが、共有メモリにアクセス可能

#### 従来のLocal Area Network (LAN)

コンピュータ間でデータを共有する最もなじみの方法は、10/100又はGigabitイーサネットのような通常のネットワーク標準に頼ることです。このアプローチでは、コンピュータは、物理的に分離され、ネットワーク経由で接続され、共通データベースは、該当標準ネットワーク上に送られたデータから保守されます。このことは、より広い接続性とより標準化された通信アプローチを可能にしますが、データ送信に対して少なからぬオーバーヘッドを追加します。また、イーサネットの調停スキームのため、決定論（通信がネットワーク上の特定の場所で利用できるようになる時間枠を定義する能力）が失われます。

イーサネットのようなLANプロトコルの通信オーバーヘッドは、使用可能なデータ・ペイロードを減らしつつ、一層の複雑さを加えます。一旦システムが数ノードを超えて成長

すれば、このオーバーヘッドは、共有メモリ・スキームが提供する利点を上回り得ます。他の例のように、これは、まだシングルポート・メモリ・アプローチで、1つのノードしか一時にデータベースを更新できません。LAN技術は、開発者にそのシステムを分散させることを可能にしていますが、シングルポート・メモリにアクセスするボトルネックを扱っておらず、今なお本質的に調停リソースです。

#### *Gigabitイーサネットの例*

以下の例は、2台のコンピュータ間でデータを共有するように要求されたプロセスを示します。これらのステップは、正規のイーサネット、並びにGigabitイーサネットLANに当てはまります。この例において、コンピュータAは、10の異なるタイプのセンサーから生データのサンプルを収集します。各タイプ20センサーでは、計200センサーが存在します。このデータは、コンピュータA自身のメモリに記憶された後、GUIによる処理と表示のため、コンピュータBに転送されます。

1. コンピュータAは、異なる間隔で各センサーのタイプのデータを収集します。それ故、これが余りに非効率である以上、データ全ての固定フォーマット・データ・ストリームを送りません。その代わりに、コンピュータAは、センサー・タイプにより、コンピュータBにデータを送ります。これを達成するために、コンピュータAは、コンピュータBが流入データを処理する方法を知るために、センサーのデータと共にセンサーのタイプと番号(1~20)を含めなければなりません。
2. これら2台のコンピュータ間には、これらのセンサーのタイプ/番号/データ・メッセージをエンコード及びデコードするアプリケーションが存在しなければなりません。コンピュータAは、各タイプのセンサーのデータの1つずつ、つまり、10の異なるタイプのメッセージをエンコードする方法を知らなければならず、コンピュータBは、10の異なるメッセージをデコードする方法を知らなければならぬのは明らかです。コンピュータBは、その後、該当メッセージの内容で活動します。
3. コンピュータAは、センサーのタイプ/番号/データ・メッセージを構築した後、該当メッセージをコンピュータBに送信しなければなりません。ネットワーク・ハードウェアとハードウェアのドライバ・ソフトウェアに頼ることにより、これを行います。コンピュータAは、これらの構築されたメッセージをネットワーク・アダプタにパスします。ネットワーク・アダプタは、その後、ネットワーク上に送信するため、このメッセージをデータ・パケットに再フォーマットします。アダプタ・ハードウェアは、コンピュータB上の受信ハードウェア・インタフェースが情報を得て、その正当性をチェックできるように、ルーティング・アドレス、エラー・チェック情報、及びその他のネットワーク・プロトコルのような情報を追加しなければなりません。
4. 情報の受信後、コンピュータBのネットワーク・アダプタ・ハードウェアは、パケットがそのまま、エラーなしで到着したことを確認するため、データ・パケットを読み込み、割り込みます。ハードウェア・アダプタは、その後、送信データをメモリに置く準備ができたことをコンピュータに通知します。コンピュータBは、その後、コンピュータAが構築したこのタイプ/番号/データ・メッセージをデコードします。コンピュータBは、実センサーのデータからセンサーのタイプとセンサー番号を分離するため、このメッセージをデコードしなければなりません。
5. コンピュータBは、どの特定のセンサー・タイプがこのメッセージを含んでいるのかを決定



するため、各種ケース又はケースのような宣言を通してセンサーのタイプと番号を決定し、20センサーの内どれからデータが来たのかも決定しなければなりません。この情報が引き出された後初めて、コンピュータAからの実データをコンピュータBのメモリに書き込ませることができ、このデータの実処理が始められます。

リフレクティブ・メモリで適用した同じ事例：

1. コンピュータAは、そのリフレクティブ・メモリ・ボード上のメモリに各センサーからの生データを置きます。各センサーは、メモリ内にそれ自身のユニークなアドレスを有します。
2. リフレクティブ・メモリは、コンピュータBのリフレクティブ・メモリ・ボードにこのデータを自動的に複製します。
3. コンピュータBは、今、ローカル・メモリで利用できるデータを有し、このデータの処理を始めることができます。

要するに、標準的なLANは、リアルタイム通信が要求される時、いくつかの欠点を有します。

- 転送速度が低い。
- データ遅延は、予測が難しく、典型的に、リアルタイム分散マルチプロセッシング・システムにとって余りに大きい。
- 多層化プロトコル・ソフトウェアは、遥かに価値あるプロセッサの時間を浪費する。

表1 オフ・ザ・シェルフ・ネットワーク技術の比較

特性	リフレクティブ・メモリ・ネットワーク (5565/5565PIORC/5565RC)	10/100イーサネット	Gigabitイーサネット
送信速度	2.1 Gbaud/s	10/100 Mbit/s	1000 Mbit/s
データ転送速度	170 MB/s	1/10 MB/s	100 MB/s
Endianデータ変換	Yes	No	No
ソフトウェア透過性	Yes	No	No
媒体	光ファイバ	同軸、UTP	光ファイバ
トポロジー	リング	リング、ハブ	リング
ネットワーク・データ送受信は、決定論的か？	Yes	No	No
ネットワーク転送スキーム	データ挿入	キャリア感知多重アクセス/衝突検知	トークン・パス
共有データへのメモリ・マップド・アクセス？	Yes	No: メッセージ・アプリケーションが構築されなければならない。	No: メッセージ・アプリケーションが構築されなければならない。
アプリケーションは、データを共有するように設計されなければならないか？	No	Yes: メッセージ・アプリケーション	Yes: メッセージ・アプリケーション
アプリケーションは、メッセージをエンコード/デコードしなければならないか？	No	Yes	Yes
アプリケーションは、エラー・チェック/処理再送信等を遂行しなければならないか？	No	Yes	Yes
CPUのオーバーヘッドは、共有データ機能をサポートすべきか？	No	Yes	Yes
CPUのオーバーヘッドは、送信ハードウェア・インタフェースで要求されるか？	No	Yes	Yes