

PPRAM-Link インタフェース・コアの開発

橋本浩二† 山崎雅也†
沖野晃一† 村上和彰††

本稿は、メモリ/ロジック混載型 ASSP「PPRAM」向けの標準通信規格「PPRAM-Link」九大業に基づいて開発中の「PPRAM-Link インタフェース・コア」について述べている。PPRAM-Link は PPRAM ノードならびに (PPRAM ノードを 1 個以上含む) PPRAM チップを 1 対 1 単方向パラレル/シリアル・リンクで接続し、1G バイト/秒 (パラレル・リンク当り) 以上の高速データ転送を可能とする。PPRAM-Link は論理階層および物理階層から構成され、論理階層はトランザクション、トランスミッション (フロー制御、キュー割当、リアルタイム転送、等を含む)、エラー検出、初期化の各プロトコルを有する。PPRAM-Link インタフェース・コアはこの論理/物理階層仕様に基づき、アドレス・デコーダ、バイパス FIFO、パケット入出力キュー、フロー・コントローラ、等から構成される。試作した PPRAM-Link インタフェース・コアのテストチップについても報告している。

Development of PPRAM-Link Interface-Core Circuit

KOJI HASHIMOTO ,† MASAYA YAMASAKI ,† KOHICHI OKINO †
and KAZUAKI MURAKAMI††

This paper outlines an interface-core circuit based on a draft standard for PPRAM-Link. The PPRAM-Link provides a high-bandwidth interface by using a collection of fast point-to-point unidirectional links, and it would be defined at 1 Gbyte/s (16-bit parallel). The PPRAM-Link consists of logical layer and physical layer; the logical layer defines several protocols such as transaction, transmission, error-recovery, initialization, and so on. The PPRAM-Link interface-core circuit, which are now developed at Kyushu University, consists of address-decoder, bypass-FIFO, packet-queues, flow-controller, and so on. A test chip of the PPRAM-Link interface-core circuit is also reported.

1. はじめに

PPRAM (Parallel Processing Random Access Memory) とは、メモリ/ロジック混載型システム LSI を基本構成要素として、それらを 1 個以上並列に相互接続することでコンピュータ・システムおよび電子機器システムを構築しようと言う「アーキテクチャ上の概念」である¹⁾。

PPRAM-Link とは、その PPRAM 型 LSI チップを相互接続するためのチップ間通信インタフェース* である²⁾。現在、PPRAM コンソーシアムにおいて、異なるベンダーの PPRAM チップ同士の相互接続性、

相互運用性、および、ソフトウェア可搬性を保証するために、この PPRAM-Link の標準仕様を策定している。標準化の対象となっているのは次の 3 つである (図 1 参照)。

- 物理階層: チップ内外に関わらず、PPRAM ノード (定義は後述) 間を相互結合する通信媒体に関する電氣的/機械的仕様を定める。
- 論理階層: PPRAM ノード間のトランザクション・プロトコル、フロー制御プロトコル、エラー検出プロトコル、初期化プロトコル、等を定める。
- API (Application Program Interface): PPRAM-Link インタフェース (定義は後述) を I/O デバイスと見做し、デバイス・ドライバ等の低レベル・ソフトウェアを開発する際に必要となるハードウェア独立な API を定める。

我々は、上記 PPRAM-Link の仕様策定と並行して、以下を目的として PPRAM-Link インタフェース・コアの開発を進めている。

- PPRAM-Link の実現可能性を示し、その性能およびコストを評価する。

†九州大学 大学院システム情報科学研究科 情報工学専攻
Department of Computer Science, Kyushu University
ppram@c.scce.kyushu-u.ac.jp
http://kasuga.scce.kyushu-u.ac.jp/~ppram

†† PPRAM コンソーシアム
PPRAM Consortium

* 正確には、後述するようにチップ内のノード間通信インタフェースも包含する。

- PPRAM-Link インターフェース・コアの実 LSI への搭載を容易にし、IP コアとしての流通を促進する。

本稿では、PPRAM-Link 論理階層仕様(九大案 0.1 版)に基づいた PPRAM-Link インターフェース・コアの概要、および、試作したテストチップについて報告する。

まず、2章で PPRAM-Link の概要および到達目標を明らかにする。3章にて論理階層仕様(九大案 0.1 版)の概要を述べる。4章で PPRAM-Link インターフェース・コアの内部構造に関して述べる。5章にて試作したテストチップについて概説する。最後に 6章で今後の課題等について述べ本稿のまとめとする。

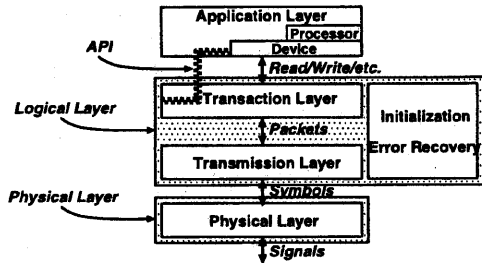


図1 プロトコル・スタック

2. PPRAM-Link とは

2.1 PPRAM ベース・システムの定義

PPRAM ノードとは、システムの基本構成単位であり、図 2 に示すように以下のものから成る。

- 0 バイト以上のメモリ
- 0 個以上のプロセッサ/ロジック
- 1 個の PPRAM-Link インターフェース

PPRAM チップとは、1 個以上の PPRAM ノードを含む LSI チップである。2 個以上の PPRAM ノードが存在する場合、それらは PPRAM-Link により相互接続される。

2 個以上の PPRAM チップから構成される PPRAM ベース・システムの場合、それらは図 3 に示すように PPRAM-Link により相互接続される。

PPRAM-Link とは上述の通り、チップ内外を問わず PPRAM ノードを相互接続し、各ノードに対して他ノードとの間の通信を可能とならしめる通信インターフェースのことである。

PPRAM-Link インターフェースとは、PPRAM ノードにあって PPRAM-Link との間のインターフェースをつかさどるロジックである。図 2 の下半分の部分がこれに相当する。

2.2 PPRAM-Link の目標

PPRAM-Link の性能および機能上の到達目標を次のように定めた。

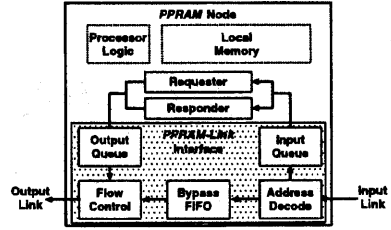


図2 PPRAM ノード

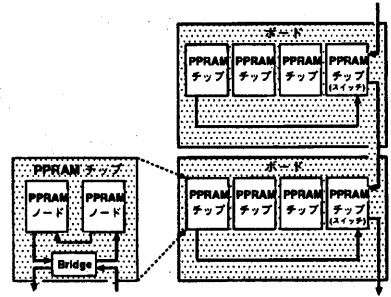


図3 PPRAM ベース・システム例

- 性能：システム性能として従来型の (DRAM と MPU とが別チップになった) システムを凌駕できるだけの、低レイテンシかつ高スループットの通信性能を提供する。スループット (バンド巾) に関する具体値としては、2000 年の時点でパラレル・リンク当り 1G バイト/秒以上、あるいは、シリアル・リンク当り 1G ビット/秒以上を目標とする。
- 機能：今日の I/O バスやシステム・バスに匹敵する機能を提供し、かつ複数の PPRAM チップが協調しながら効率的に並列/分散処理できるよう、今日の並列/分散システム向け相互結合網に相当する機能も提供する。主要なものは以下の通り。
 - IEEE Std 1212-1991 CSR^{*1} Architecture⁸⁾、および、IEEE Std 1596-1992 SCI^{*2,7)} に一部準拠。
 - NCC-NUMA^{*3} アーキテクチャの提供。すなわち、
 - * PPRAM ベース・システム全体でグローバルな 1 個の 64 ビット物理アドレス空間を提供。
 - * 上記グローバル物理アドレス空間上での通常のメモリに対する Read/Write という操作で通信を実現。
 - リングレット^{*4} を基本トポロジとするが、シ

*1 Control and Status Register

*2 Scalable Coherent Interface

*3 Non Cache Coherent, NonUniform Memory Access

*4 SCI の定義で、あるノードの出力リンクから同一ノードの入力リンクへと至る閉路。

システム全体としては特定のトポロジには依存しない。

- IEEE1394⁹⁾等の周辺バスで提供されているアイソクロナス (isochronous) 転送機能に対応したリアルタイム転送機能の提供。
- 活線挿抜機能の提供。

3. 論理階層

以下、論理階層仕様 九大案 (0.1 版⁵⁾, 0.11 版⁶⁾) を概説する。

論理階層は図 1 に示すように、次の 2 層から成る。

- (1) **トランザクション層**: 上位層のアプリケーション層から指示されたトランザクション (定義後述) をバケット (定義後述) の交換により実現する。
- (2) **トランスミッション層**: 上位層のトランザクション層におけるバケット送受信をシンボル (定義後述) の伝達により実現する。

さらに、両層にまたがるプロトコルとして、エラー検出プロトコルおよび初期化プロトコルが存在する。

3.1 トランザクション層

2 個のノード間のデータ交換をトランザクションと呼ぶ。1 個のトランザクションは、図 4 に示すように、一般に次の 2 個のサブアクションから成る (スプリット・トランザクション)。

- (1) **要求サブアクション**: 要求側ノードから応答側ノードへ要求を送る。
- (2) **応答サブアクション**: 応答側ノードから要求側ノードへ応答を返す。

さらに、1 個のサブアクションは、一般に次の 2 個のバケットから構成される。

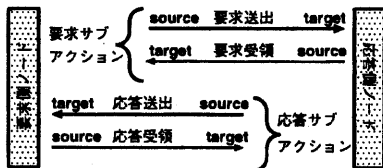


図 4 トランザクション

- (1) **送出バケット (send packet)**:
 - **要求送出バケット**: 要求側ノードから応答側ノードへ、要求を送出。
 - **応答送出バケット**: 応答側ノードから要求側ノードへ、応答を送出。
- (2) **受領 (echo) バケット**:
 - **要求受領バケット**: 応答側ノードから要求側ノードへ、要求を受領したことを通知。
 - **応答受領バケット**: 要求側ノードから応答側ノードへ、応答を受領したことを通知。

PPRAM-Link は、読出し (read)、書込み (write)、選択ワード書込み (writesw)、移動 (move)、ロック

(lock)、イベント (event)、アクティブ・メッセージ (actmes)、未使用 (unused) の 8 種類のトランザクション・セットを定める。

3.2 トランスミッション層

バケットを構成する不可分な単位は、16 ビット長のシンボルである。バケット長は最小 8 シンボル (16 バイト) で、常に 8 シンボル (16 バイト) の倍数となる。バケット・フォーマットについては、⁶⁾ 参照。トランスミッション層では、バケットをシンボル単位にソース・ノードからターゲット・ノードへと転送する。

通常、バケットとバケットの間には 1 つ以上の idle シンボルをはさむ。idle シンボルは、フロー制御に必要な情報を運ぶ手段としても用いられる。

フロー制御プロトコルとして、次の 2 つを定める。

- **バンド巾割当てプロトコル**: 優先度の高いトランザクションにバンド巾を優先的に割り当てる。
- **キュー割当てプロトコル**: 受信側ノード (つまり送出バケットのターゲット・ノード) のバケット受領キューの割り当て制御を行う。

また現在、リアルタイム転送プロトコルの仕様を策定中である。リアルタイム転送では一般に、次の 2 つの性質のいずれか、あるいは、双方を要求する。

- **低レイテンシ性**: 1 つのバケット転送に要するレイテンシがある要求値以下であること
- **定インターバル性**: 連続したバケットの送出がある一定間隔で行えること

4. PPRAM-Link インタフェース・コアの開発

PPRAM-Link インタフェース・コアは論理階層部と物理階層部から成る (図 5 参照)。

図 5 は、物理階層からトランスミッション層までを包括したインタフェース・コアである。この場合トランザクション層は、ソフトウェアまたは別回路で提供する。インプット・リンクから入力されたバケットは、自ノード宛ならばレスポンス/リクエストに渡し、それ以外はバイパス FIFO を介してアウトプット・リンクから隣接ノードに送出する。

インプット・リンクおよびアウトプット・リンクの入出力はそれぞれ次の通り。

- シンボル (16 ビット)
- **フラグ・ビット**: バケットの種類/長さを判別する目的で出力する 1 ビットの信号線。
- **クロック・ビット** (1 ビット)

図 5 の各ブロックに関して、以下に述べる。

4.1 アドレス・デコーダ

アドレス・デコーダの機能は次の通り。

- 入力シンボル列 (例えば図 6) からバケットのノ

* 後述するエラストティック・バッファでの周波数のズレの修正用に用いる、同期バケットとその 1 つ前のバケットの間には挟まない。

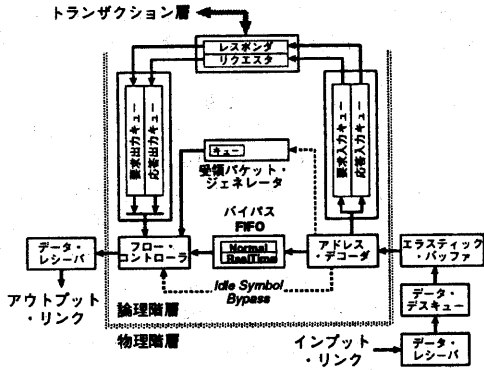


図5 PPRAM-Link インタフェース・コアの基本構造

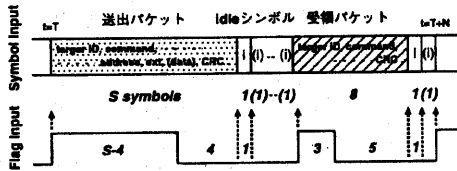


図6 シンボル列の入力例

ド ID およびコマンドを検出し、自ノード宛のバケットだけを取り出す。

- 自ノード宛送出バケットの入力ならば、受領 (echo) バケット・ジェネレータに受領バケット生成を指示すると同時に、バケットを入力キューに渡す。
- 自ノード宛受領バケットの入力ならば、出力キューに保存されている送出バケットを破棄。
- キュー割り当てプロトコルに基づく入力キューの管理。
- 自ノード宛バケットを対象にした CRC (CRC シンボルはバケットの最後尾に付加される) などのエラー検出。
- 入力シンボル列中の idle シンボルをフロー・コントローラへバイパス。

すなわちアドレス・デコーダは、入力シンボル列のバケット認識とその送り先制御を行う。

4.2 バイパス FIFO

バイパス FIFO には、自ノードが新たなバケットを送出している間にアドレス・デコーダから入力されたバケットを保存する。

バイパス FIFO 自身は一般的なクロック同期式 FIFO で、その容量はデッドロックを防ぐため最低でも「最大バケット長+1」必要である。またリアルタイム転送をサポートする場合は、バイパス FIFO の二重化が必要となる。

バケット入力の制御はアドレス・デコーダが行い、出力制御はフロー・コントローラが行う。

4.3 フロー・コントローラ

以下の優先順位で出力バケットを選択し、データ・

ドライバを介して送出する。

- (1) 受領バケット・ジェネレータにて生成された受領バケット
- (2) 出力キューに格納されたバケット
- (3) バイパス FIFO に格納されたバケット
- (4) フロー・コントローラ内の idle ジェネレータが生成する idle シンボル

出力バケットの選択は、フロー制御プロトコルに基づいて決定する。idle ジェネレータはアドレス・デコーダから idle シンボルを受け取ることでリンクの現在の状況を把握し、新たに idle シンボルを生成する。

フロー・コントローラでは、シンボルの送出と並行してフラグ・ビットおよびクロック・ビットを生成、送出する。

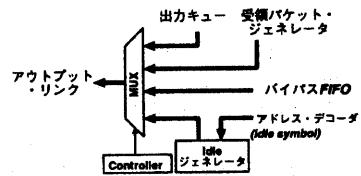


図7 フロー・コントローラ

4.4 入力キュー

入力キューには、自ノード宛送出バケットが格納される。入力キューの回路構成の選択肢として

- 要求バケット用と応答バケット用でキューの二重化
- キューの容量

が挙げられ、これらは回路面積とトランザクション層の動作速度とのトレード・オフで決定する。キューを二重化しかつ容量をバケット 4 個分とした場合を、図 8 に示す。

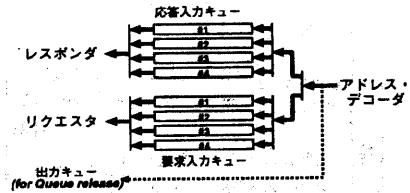


図8 入力キュー

4.5 出力キュー

出力キューは基本的に入力キューと同一構造である。出力キューはリクエスト/レスポンスからバケットを受け取り、フロー・コントローラに出力する。キューの削除は、アドレス・デコーダが受領バケットを受け取った時に行う。

4.6 エラスティック・バッファ

エラスティック・バッファとは、チップ外部からのシンボル/フラグ・ビットを取り込むタイミングを調整するための回路である。

インプット・リンクからのシンボルの入力、前ノードのクロックに同期している。一方、受信側ノードは自分自身のクロックでシンボルを受け取ろうとする。

すると時間の経過に伴い、シンボルが送られてくるタイミングとそのシンボルを受信する側のタイミングのずれが無視できなくなり、正常なパケット転送が継続できなくなる場合がある。これをエラスティック・バッファにより防ぐ。

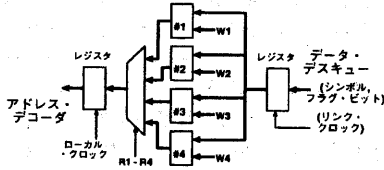


図9 エラスティック・バッファ

図9におけるリンク・クロックは前ノードからのクロック入力、W1 - W4 はリンク・クロックから生成された書き込みタイミング信号である。

これにより入力シンボルは、#1→4の順に書き込まれる。同様に#1→4の順にシンボルを読み出す。ここで書き込みに対して読出しが2ローカル・クロック・サイクル分遅い状態を維持すると、1クロック・サイクル程度の書き込み/読み込みのタイミングのズレを調整できる。

修復不可能なズレに対しては、入力途中のパケットの破棄をアドレス・デコーダに指示する。

4.7 データ・デスクュー

PPRAM-Link がパラレル・リンクである場合に、シンボルの各信号線間の到達時間の歪み (Skew) を極力低減させるための回路がデータ・デスクューである。

データ・デスクューはパケット・シンボルの各ビット線をサンプリングし、ビット線間の差に対応した遅延を挟み込む。

PPRAM-Link が想定する G バイト単位の高速度パラレル・リンクでは、データ・ドライバ/レシーバによる遅延、チップ内外の配線による遅延が問題となる。データ・デスクューによる遅延差の解消は必須である。

5. テストチップ

1997年6月に設計した PPRAM-Link インターフェイス・コアの機能縮小版テストチップについて概説する。図10がその回路構成図で、図11はそのチップ・レイアウト図である。

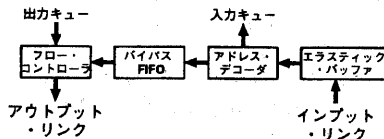


図10 試作した PPRAM-Link インタフェース・コア

今回のチップ試作では、時間的制約と使用可能なチップ面積の制限および、SRAM セルを記憶素子として採用しなかったため、準拠すべき PPRAM-Link 論理仕様を大幅に縮小した。

その変更後のパケット・フォーマットは以下の通り。

- シンボル長=8ビット, 最大パケット長=20シンボル (ただしフラグ・ビットの動作は論理仕様通り)。
- ノード ID, コマンド共に8ビット (=1シンボル)。
- 最大データ転送サイズ=16バイト。

フロー制御プロトコルの内、バンド巾割り当てプロトコルの一部を実現し、最低限のパケット転送を保証した。エラー検出および初期化プロトコルは実装しなかった。バンド巾割り当てプロトコルを最低限度での実装に留めると、本来バンド巾情報を内部に含む idle シンボルは、その含むべき内容がないためパケットとパケットの間に挟む以外の役割は持たなくなる。

結果としてアドレス・デコーダとフロー・コントローラは大幅に簡略化された。

物理階層についてはエラスティック・バッファの実装のみとし、インプット・リンクおよびアウトプット・リンク (リンク入出力ポートの I/O パッド) は通常の TTL 入出力とした。

試作チップの諸元は以下の通り。

- 0.5μm CMOS テクノロジ, メタル2層, ポリ1層
- スタンダード・セル方式 (合計セル数: 約300)
- チップ・サイズ (パッド含む): 4.8mm 角
- I/O ピン数: 76本 (電源関連含めて108本)
- I/O パッド: NEL* 提供の TTL 入力

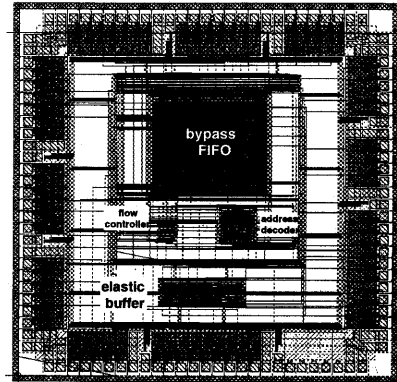


図11 チップのレイアウト図

以下に各ブロックの概要を述べる。

5.1 アドレス・デコーダ

図12にその回路構成を示す。シンボル(8ビット)とフラグ・ビットのバイライン・レジスタが回路の基本で、各レジスタの値を読み取って、パケットの送出先を決定する。その認識機構は以下の通り。

- パケットの先頭検出: #f3=0 and #f2=#f1=1。
- パケット・ノード ID: パケット先頭検出時の #s2。
- パケット・コマンド (パケット種類等の情報を含む): パケット先頭検出時の #s1。
- パケットの後尾決定: #f3=1 and #f2=#f1=0 の

* NEL: NTT electronics technology corp.

検出から数えて、送出バケットならば4シンボル後、受領バケットならば4シンボル後。

- idle シンボルの上書き：#f3=0 かつ #f2=0 ならば #3 を #2 で上書きし、どこにも送出しない。バケットとして認識されないシンボルは全て、idle シンボルとする。

バケットの直前の idle シンボルは上記の上書きの対象とはならず、バケットと共に送出される。

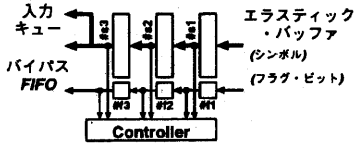


図12 試作チップにおけるアドレス・デコーダ

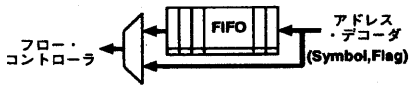


図13 試作チップにおけるバイパス FIFO

5.2 バイパス FIFO

D-フリップ・フロップで構成された、21(=最大バケット・シンボル数+1)の容量の FIFO である。フロー・コントローラの回路簡略化のため、フラグ・ビットも格納する。図13にその回路構成を示す。

5.3 フロー・コントローラ

バケットの入力は優先順に

- (1) バイパス FIFO
- (2) 入力キュー (チップ外)
- (3) idle ジェネレータ (idle シンボル中に情報は含まれない)

フラグ・ビットの生成機構を持たないため、フラグ・ビットの入力を必要とする。

5.4 エラスティック・バッファ

シンボル長が8ビットである事以外は、4.6節で示した構造の通りである。

6. おわりに

PPRAM-Link の概要およびその論理仕様について述べた。次に PPRAM-Link インターフェース・コアの基本構造を概説し、最後に試作したテストチップの概要を述べた。今後の予定および課題は以下の通り。

- チップ試作で得られた、現状の PPRAM-Link 論理仕様の問題点を整理し、現実のチップ設計を配慮した仕様とする。
- 物理階層部分のデータ・デスクューおよびレシーバ/ドライバ回路の開発。
- フロー制御およびエラー訂正プロトコルの完全な実装、およびアドレス・デコーダにおける遅延の削減。
- SRAM セルの開発。

- 試作チップに対する性能評価。

謝辞

常日頃貴重な御意見を頂く PPRAM コンソーシアムのメンバ諸氏に感謝致します。

また、日頃から御討論頂く、九州大学 大学院システム情報科学研究科 安浦寛人 教授、岩井原瑞穂 助教授、PPRAM プロジェクトのメンバ諸氏、ならびに、安浦・村上・岩井原研究室の諸氏に感謝致します。

本研究は一部、文部省科学研究費補助金基盤研究(A)(2)展開研究「メモリ/ロジック混載技術に基づく大規模集積回路システム・アーキテクチャの研究開発」(課題番号: 09358005)、富士通研究所奨学寄付金「プロセッサアーキテクチャ」による。また、今回のテストチップ試作は、東京大学大規模集積回路設計教育研究センター (VDEC) のご協力による。記して感謝致します。

参考文献

- 1) 村上和彰, 岩下茂信, 宮嶋浩志, 白川 暁, 吉井 卓, “メモリ-マルチプロセッサ一体型 ASSP (Application-Specific Standard Product) アーキテクチャ: PPRAM,” 信学技報, ICD96-13, CPSY96-13, FTS96-13, 1996年4月。
- 2) 村上和彰, 岩下茂信, 宮嶋浩志, “メモリ-マルチプロセッサ一体型 ASSP PPRAM 用標準通信インターフェース【PPRAM-Link Standard】 Draft 0.0 の概要,” 情処研報, ARC-119-27, 1996年8月。
- 3) 村上和彰, 吉井 卓, 岩下茂信, 宮嶋浩志, “PPRAM ベース・システム向け分散共有メモリ・システムの提案,” 情処研報, OS-73-2, 1996年8月。
- 4) 山崎雅也, 橋本浩二, 沖野晃一, 村上和彰, “PPRAM-Link 論理階層仕様 (九大案 0.1 版) の概要,” 信学技報, ICD97-24, 1997年5月。
- 5) 九州大学 PPRAM プロジェクト・チーム, PPRAM-Link 論理階層仕様書 (九大案 Ver 0.1), PPRAM コンソーシアム第1回合同分科会, 1997年4月10日。
- 6) 九州大学 PPRAM プロジェクト・チーム, PPRAM-Link 論理階層仕様書 (九大案 Ver 0.11), PPRAM コンソーシアム第2回合同分科会, 1997年5月21日。
- 7) IEEE Std 1596-1992, IEEE Standard for Scalable Coherent Interface (SCI), IEEE Computer Society, 1992.
- 8) IEEE Std 1212-1991, IEEE Standard Control and Status Register Architecture for Microcomputer Buses, IEEE, 1991.
- 9) IEEE Std 1394-1995, IEEE Standard for a High Performance Serial Bus, IEEE, Aug. 1996.