

# 超並列計算機 JUMP-1 における ハイビジョン画像表示システム

小畑 正貴\*<sup>1</sup>, 中條 拓伯\*<sup>2</sup>

\*<sup>1</sup>岡山理科大学 工学部 情報工学科, \*<sup>2</sup>神戸大学 工学部 情報知能工学科

e-mail: kohata@ice.ous.ac.jp

JUMP-1 は, プロセッサ間での通信/同期のための機能を備えた複数のクラスタを, RDT と呼ばれるネットワークで接続した分散共有メモリ型のアーキテクチャを持つ。クラスタと画像表示システムは高速シリアルリンクにより接続され, 分割画面データは並列転送される。本稿では画像表示システムの構成と表示方式, アクセス方式について述べる。

## An HDTV display system for the Massively Parallel Computer JUMP-1

Masaki Kohata\*<sup>1</sup>, Hironori Nakajo\*<sup>2</sup>

\*<sup>1</sup> Faculty of Engineering, Okayama University of Science, \*<sup>2</sup> Faculty of Engineering, Kobe University,

A massively parallel computer JUMP-1 consists of multiple clusters providing inter-processor communication and synchronization mechanism via a broad bandwidth inter-connection network called RDT.

We introduce an HDTV display subsystem configuration which consists of distributed frame memory connected via fast serial links each called Serial Transparent Asynchronous First-in First-out Link (STAFF-Link). In this paper, we describe the features and hardware configurations of the display subsystem and its access method from JUMP-1.

# 1 はじめに

現在、文部省科学研究費補助・重点領域研究において、分散共有メモリ型の超並列計算機のプロトタイプマシン JUMP-1[1][2] の開発が進められている。超並列計算機の重要な応用分野である科学技術計算においては計算結果の可視化のための画像表示装置が必要であるが、この場合、高速計算に見合う表示の高速性が要求される。

並列計算機における画像表示システムの形態を図1に示す。多くの並列計算機における画像表示システムは(a)の形態であり分割処理した部分画像を1ノードに集めてそこで表示する。この方式では、高画質画像や動画像のようにデータ量が多くなると、表示データの転送が計算のための通信を圧迫したり、表示ノードに通信が集中したりする問題点が生じる。

(b)は画面を分割してフレームメモリを分散させる方式であり、特にグラフィックス应用到重点を置いたマシンで使われている。プロセッサ側からは複数が同時にアクセスできるため、競合が軽減される。しかし分散配置されたフレームメモリからビデオ出力回路(D/A 変換器など)へのデータ転送に高速性(数百 MB/s)が要求されることから、複数ラックに及ぶような大規模システムへの拡張は実装上困難である。

(c)は分割したフレームメモリを遠隔に置く方式であり、大規模マシンの画像表示システムとして有効である。現在のところ、この形の画像表示システムは AP1000 の画像表示システムに見られる程度である。

(c)では特に、プロセッサとフレームメモリ間の通信リンクに対して高速性や実装容易性が要求されるが、JUMP-1では Serial Transparent Asynchronous First-in First-out Link (STAFF-Link)[3] と呼ばれる高速なシリアルリンクを用いる。

本稿では、JUMP-1 の入出力サブシステムの概略を述べた後、画像表示装置のハードウェア構成、表示方式、他のシステムへの応用について述べる。

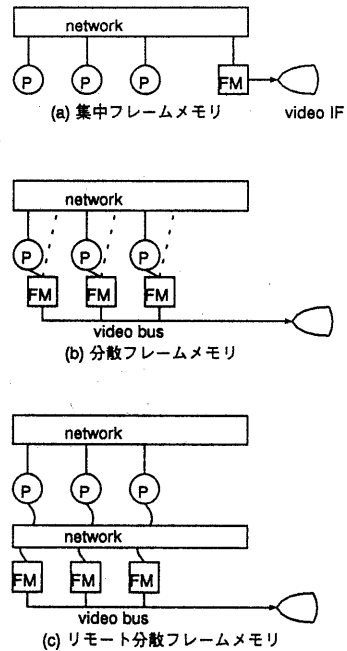


図 1: 並列計算機の画像表示システム

## 2 JUMP-1 の入出力アーキテクチャ

### 2.1 JUMP-1 の概要

JUMP-1 は分散共有メモリ型の並列計算機である。汎用プロセッサに加えて、Memory Based Processor (MBP)[4] と呼ばれる非局所処理に特化したプロセッサを導入し、MBP に通信や同期等の処理を実行させる。複数の PE と複数の MBP を組み合わせてクラスタを構成し、複数のクラスタを Recursive Diagonal Torus (RDT)[5] と呼ばれる相互結合網を介して結合する。JUMP-1 のクラスタ構成を図 2 に示す。RDT は 2 次元トーラス構造を基本とする相互結合網で、高い転送バンド幅と、メッシュやハイパーキューブのエミュレーションが可能である等の特徴を持つ。

### 2.2 入出力サブシステムの構成

入出力データは各クラスタに分散しているため、入出力装置との接続は複数のクラスタに分散した形態が望ましい。図 3 に JUMP-1 の入出力サブシ

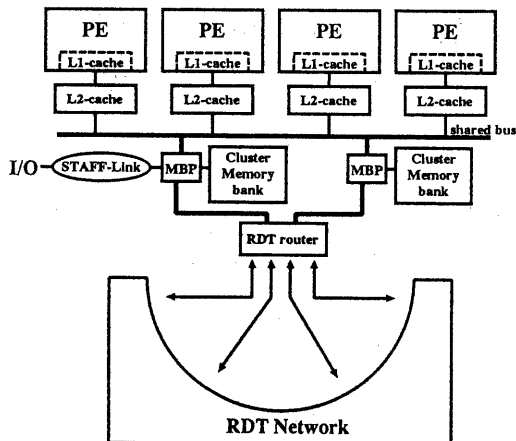


図 2: JUMP-1 のクラスタ構成

システムの構成を示す。STAFF-Link は高速シリアル伝送による通信リンクであり、システム全体が物理的に大きくても、柔軟かつ実装が容易である。

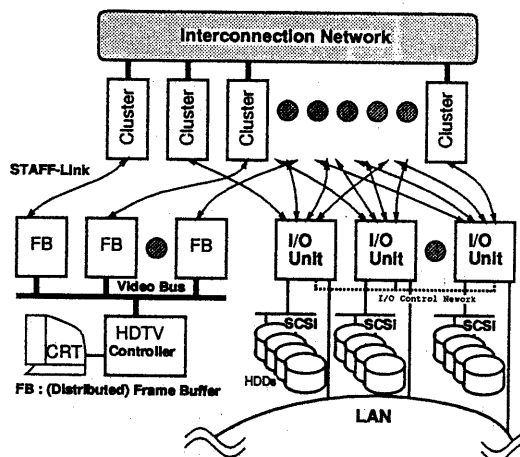


図 3: JUMP-1 の入出力サブシステムの構成

JUMP-1 における入出力サブシステムの特徴を以下に挙げる [6].

### 1. STAFF-Link による設置場所の柔軟性

入出力機器は本体に近い位置に設置するのではなく、比較的離れた場所に設置し、柔軟で高速な STAFF-Link を介して結合することによ

方式	ハイビジョン
画面構成	2048 × 1024 (規格は 1920 × 1035) 24 bit/pixel (1 画面あたり 6MB)
ドット周波数	74.25 MHz
水平周波数	33.75 KHz
垂直周波数	60 Hz (インターレースで 30 フレーム/秒)
メモリ容量	6MB × 2 バンク (ダブルバッファ)
分割数	16
ホストとの通信	17.5MB/s シリアルリンク (STAFF link) × 16

表 1: 画像表示システムの仕様

て、分散した共有入出力装置を構築する。

### 2. 共有入出力バッファ

入出力用のアドレス空間を本体の共有メモリマップ上にマッピングする。これによって、入出力機器を共有メモリとしてすべてのクラスタ間で共有することが可能となり、種々の入出力機器の特性を吸収することができる。この入出力のための共有メモリを共有入出力バッファと呼ぶ。

## 2.3 画像表示システムの構成

### 2.3.1 全体構成

表 1 に画像表示システムの仕様を、図 4 に全体構成を示す。ハイビジョン画質で動画像の表示が可能となるように仕様を決めている。

ハイビジョン 1 画面分のフレームメモリは、JUMP-1 のクラスタからは共有入出力バッファとして見え、クラスタからのアクセスは共有空間に対するリード/ライトアクセスとして実現される。その共有フレームメモリは 16 分割され、表示データが 16 クラスタから並列に転送される。

クラスタと共有フレームメモリ間は TAXI チップを利用した STAFF-Link により接続される。ハイビジョン規格の画面はドット周波数 74.25MHz で、1920×1035 ドットの構成である。今回は回路簡略化のためフレームメモリの構成を 2048×1024(2M ドット)としている。各ドット 24 ビットなので、1 画面の情報量は 6MB となり、毎秒 30 フレームで 180MB/s の情報量となる。175MHz の TAXI チップで最大 17.5MB/s の転送能力があり、これを 16 本使うと 280MB/s となるので、表示速度に見合うだけのデータ転送幅が確保できる。

16 枚のフレームメモリボードからディスプレイへのデータ転送はビデオバスによる。ビデオバスには

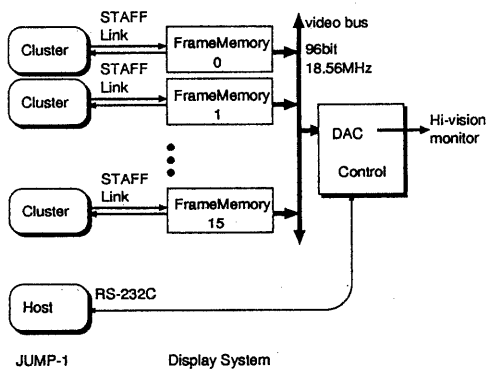


図 4: 画像表示システムの構成

汎用の VME バスを利用する。データ幅を 96 ビットにし 4 ドット分を同時に転送することによって、バス周波数は

$$74.25/4 = 18.56\text{MHz}$$

となる。

### 2.3.2 共有フレームメモリ

表示とデータ転送を並列に行うため、共有フレームメモリはダブルバッファ構成とし、一方が表示中に、他方に次の表示データをクラスタから転送する。フレームメモリボードの構成を図 5 に示す。1 ボードあたり

$$32\text{K ドット} \times 4 \text{ ページ} = 128\text{K ドット}$$

で、16 ボードで 2M ドットとなる。

制御回路は 2 個の FPGA (Field Programmable Gate Array) により実現される。1 個は JUMP-1 からのデータ転送を制御し、他方はビデオバスへのデータ転送を制御する。

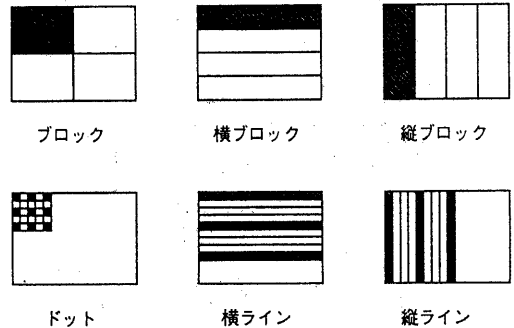


図 6: 分割モード

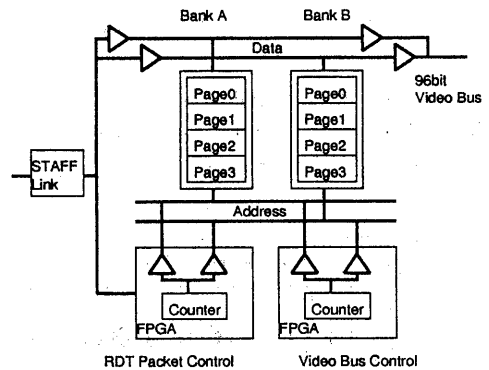


図 5: フレームメモリボード

画面分割は図 6 の 6 種類をサポートする。ここで、ラインモードは幅を持ったラインの繰り返しを含み、ドットモードは大きさを持った小ブロックの繰り返しを含むものとする。分割方法およびクラスタへの割り当ては静的で、プログラム実行前に分割モードを切替えるものとする。画面分割を変更すると、各フレームメモリからのデータの読み出し順序の変更が必要になるが、これはビデオデータ読み出し側の FPGA の再コンフィギュレーションによって行う。

全フレームメモリは JUMP-1 の共有メモリ空間にあるので、PE 上での画面分割とフレームメモリ上での分割が一致しなくても表示は可能であり、フレームメモリの分割モードは 1 種類でもかまわない。しかし RDT 上での表示データ転送を少なくするためには、両者はできるだけ一致している方がよい。

ビデオバス上では横方向に連続する 4 ドットを同

時に転送する構成になっているため、各フレームメモリには4ドット単位で分割して持たせるのがよい。ドットと縦ライン以外のモードでは連続した4ドットは同一のボードから同時に読み出されるが、純粋なドットモードと縦ラインモードでは連続した4ドットは4枚のボードから読み出さなければならない。

純粋なドットモードと縦ラインモードに対応することは可能であるが、ビデオバスの転送制御は複雑になるので、横4ドットだけは連続させることにする。従ってドットモードは1×4のブロックの繰り返し、縦ラインモードは4本単位での繰り返しとする。JUMP-1のプロセッサ数はフレームメモリの分割数より多く、JUMP-1上でドット分割した場合でも近くの画素は同じフレームメモリに集まることになるので問題はないと考える。

クラスタからのデータ転送は、1画面分のデータを適当なドット数単位で分割して、メモリへのライトアクセスとして行う。パケットは、RDTパケットに準拠し、画面に対応する共有空間にRGBデータを書き込むという形をとる。フレームメモリが接続されてないクラスタからも、共有入出力空間上の目的位置にライトすることで自動的にRDTを通してデータ転送が行なわれる。

誤り検出時の処理としては、表示切替えの周期が短い場合は非再送、長い場合は再送が適していると考えられるので両方用意する予定である。ただし、STAFF-Linkの誤り率が低い場合にはすべて非再送で処理する。

### 2.3.3 コントロールボード

コントロールボードの構成を図7に示す。ビデオバスからのデータをマルチプレクサを通してDA変換器(DAC)に入れ、モニタにつなぐ。同期信号等のタイミングはFPGAで作成する。FPGAのコンフィギュレーションによって、ハイビジョンのほか、SVGAなどの通常のグラフィックディスプレイにも表示が可能である。

FPGAのコンフィギュレーションや全体の制御を行うためマイクロプロセッサを使用する。このマイクロプロセッサをRS-232Cで宿主(コンソール)と接続し、コンフィギュレーションデータのダウンロードや表示モードの切替えなどを行うようにする。

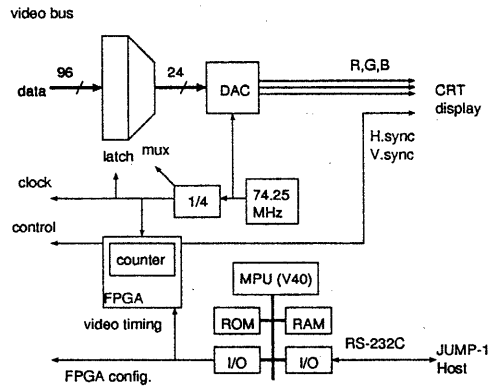


図7: 画像表示コントロールボード

実装は、VMEダブルハイトのボードを使い、16枚のフレームメモリボードとコントロールボードとを標準VMEラックに収める。STAFF-Linkの部分はドータボードとして実装する。

## 2.4 アクセス方式

### 2.4.1 ディスプレイキャッシュ

共有フレームメモリへのアクセスを考えると、表示の面では共有フレームメモリはビデオ回路の近くにあるのがよく、計算の面ではプロセッサの近くにあるのがよい(物理的にも論理的にも)。STAFF-Linkはメモリアクセスに比べると遅いので、画素データの読み書きのたびに直接フレームメモリをアクセスするのは非効率である。

そこで、図8のようにJUMP-1のクラスタメモリ内にディスプレイの表示イメージを持たせ、これを定期的に共有フレームメモリにブロック転送することで表示を一致させる方式をとる。この方式はディスクに対するディスクキャッシュと類似しており、ディスプレイキャッシュと呼ぶことにする。

### 2.4.2 データ転送と表示のタイミング

JUMP-1からのデータ転送は図9のように行なわれる。任意のクラスタメモリの表示データはRDTを通して任意の分割フレームメモリに転送できる。このため、プロセッサへの画面割付けと分散フレームメモリへの画面割付けは一致しなくてもよいが、

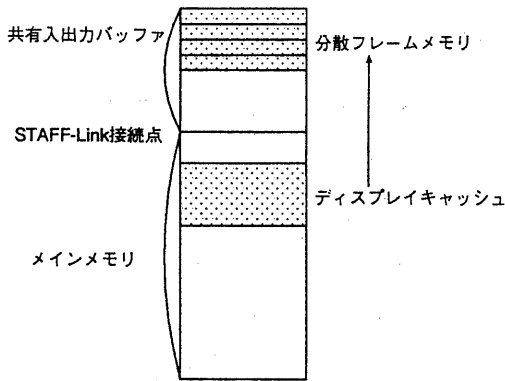


図 8: メモリ空間

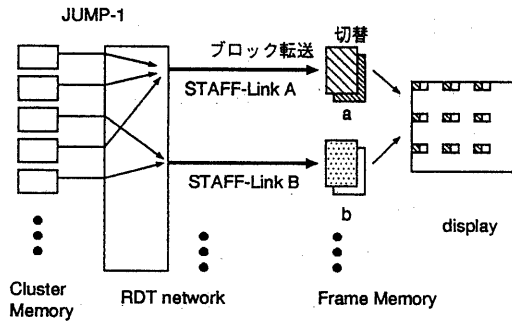


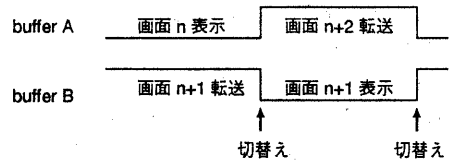
図 9: 表示動作

RDT 上の通信局所性を出すにはフレームメモリの分割モードを適当に選択するのがよい。

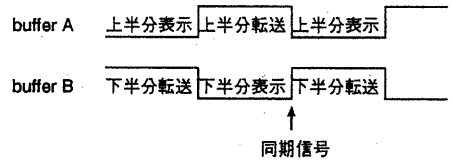
フレームメモリはダブルバッファになっているが、この切替え方式として以下の2つが考えられる。タイミングを図 10 に示す。

### 1. 画面切替え

片方を表示バッファ、他方を転送バッファとし、1画面更新毎に切替える。この方式では前の画面の一部だけを書き換えるということできないので、全画面データを転送しなければならない。次の画面の作成中にも表示画面は変化せず、切替え信号によって全画面は同時に新しい画面と切り替わる。データ転送後終了コマンドを特定のアドレスに送ることにより、フレームメモリのバッファ切替えを行う。データ転送時間は



(a) 画面切替え



(b) 画面分割

図 10: バッファ切替のタイミング

各ボードによって違ってくるため、切替は全ボードが終了コマンドを受けてからとなる。

### 2. 画面分割

1画面を2分割し、一方の表示期間中に他方に転送する。分割方法としては、上下分割、奇数ライン/偶数ライン分割、インターレースにおけるフィールド分割などが考えられる。バッファ切替えはビデオ信号に同期して定期的に行なわれ、画素値の変更はすぐ表示に現れる。転送は変更画素に対してのみ行なえば良い。ただし、実際は転送はブロック単位であるので、その画素を含むブロック転送となる。

画面の更新はディスプレイキャッシュのライトバックに相当する。このタイミングとしてユーザ指定とタイマ割り込みが考えられる。

1. ユーザ指定: プログラム内に共有フレームメモリへのブロック転送関数を記述する。
2. タイマ割り込み: JUMP-1 の OS が内部タイマ割り込みまたはフレームメモリ側からの割り込みにより定期的にブロック転送を実行する。

表示データ転送に関するパケットは、RDT パケットフォーマットに準拠する。1パケットで転送されるデータは32バイトである。

## 謝辞

本研究の一部は文部省科学研究費(重点領域研究(1)課題番号 04235130「超並列ハードウェア・アーキテクチャの研究」)による。

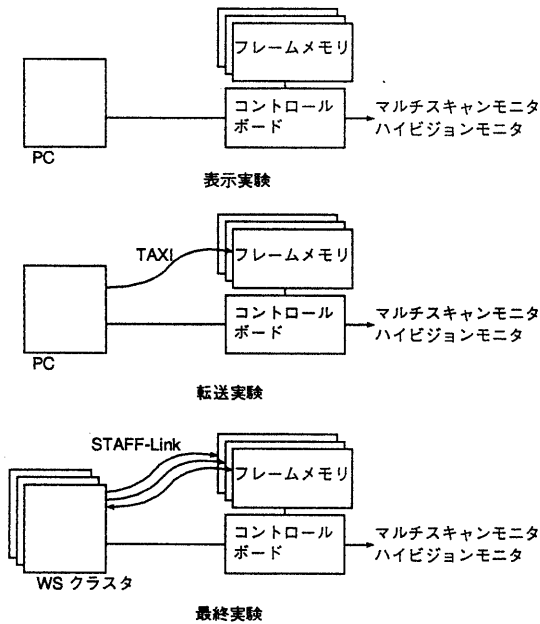


図 11: 開発手順

## 3 開発環境

現在のところ JUMP-1 本体も STAFF-Link も開発途中にある。この状況において、画像表示システムの開発を図 11 に示す手順で進めている。

まずコントロールボード上のマイクロプロセッサからテストパターンをフレームメモリに書き込み、表示部を完成させる。次にパーソナルコンピュータに実験用通信リンクボードを作成し、フレームメモリとの間で通信実験を行なう。実験用通信リンクは STAFF-Link で使用するのと同じシリアル通信チップ(TAXI)を使用するが、片方向とし、FIFO も省略する。パーソナルコンピュータからは RDT パケットに準拠した形式でブロック転送を行ない通信部を完成させる。

STAFF-Link ができた時点で実験用リンクをこれと交換する。JUMP-1 の代わりとして複数のワークステーションをネットワーク接続したワークステーションクラスタを用い、リモート分散フレームメモリとしての総合的な実験を行なう。ワークステーションクラスタとの組合せは、開発用というだけでなく実用にもなると考えている。

## 参考文献

- [1] 松本 尚, 平木 敬“超並列計算機上の共有メモリアーキテクチャ”, 信技報, *CPSY 92-26*, pp.47-55, Aug 1992.
- [2] 平木 敬, 天野 美晴, 久我 守弘, 末吉 敏則, 工藤 知宏, 中島 浩, 中條 拓伯, 松田 秀雄, 松本 尚, 森 眞一郎, “超並列プロトタイプ計算機 JUMP-1 の構想”, 情報処理学会計算機アーキテクチャ研究会報告 *ARC102-10*, pp.73-84, Oct 1993.
- [3] 中條 拓伯, 松田 秀雄, 金田 悠紀夫, “超並列計算機における ワークステーションクラスタ・ファイルシステム”, 情報処理学会計算機アーキテクチャ研究会報告 *ARC107-24*, Jul 1994.
- [4] 松本 尚, “局所処理と非局所処理を分離並列処理するアーキテクチャ” 第 49 回情報処理学会全国大会講演論文集 (6), pp.115-116 Oct 1991.
- [5] 文部省重点領域研究「超並列原理に基づく情報処理基本体系」第 3 回シンポジウム予稿集, pp.257-279, Sep 1993.
- [6] 中條 拓伯, 松本 尚, 小畑 正貴, 松田 秀雄, 平木 敬, 金田 悠紀夫, “分散共有メモリ型超並列計算機 JUMP-1 の入出力サブシステムのアクセス方式”, 情報処理学会計算機アーキテクチャ研究会報告 *ARC104-15*, pp.113-120, Jan 1994.
- [7] 岡田 勉, 中條 拓伯, 松本 尚, 小畑 正貴, 松田 秀雄, 平木 敬, 金田 悠紀夫, “超並列計算機 JUMP-1 における入出力サブシステムのアクセス方式”, 情報処理学会計算機アーキテクチャ研究会報告 *ARC107-23*, pp.177-184, July 1994.