

## 2 CPU 共有資源

### o はじめに

システムの両中央処理装置(CPU)は、メインフレームの中央メモリ、CPU間通信部、入出力部を共有する。これらのCPUで共通な領域を、以下のページで記述する。

### o 中央メモリ

中央メモリはいくつかの半導体RAMのバンクから成っている、そして、CPUとI/O部で共有されている。標準の中央メモリの大きさは、16バンクで2Mワード、か、32バンクで4Mワードである。

バンクは各々独立している; 連続にアドレス付けされたワードは、連続したバンクに置かれる各ワードは64bitのデータと8bitのチェック・ビットを持った72bitである。

中央メモリのサイクル・タイムは4クロック周期(CP)で38ns(ナノ秒)である。アクセス・タイム、中央メモリから操作レジスタへオペランドを読み出すのに必要な時間、はA(アドレス)とS(スカラ)レジスタへは、14CP(133ns)である。V(ベクトル)レジスタのアクセス・タイムは「17CP+ベクトル長」であり、そしてB(中間アドレス)かT(中間スカラ)レジスタへのブロック転送は「16CP+ブロック長」である。

B,T,VレジスタのためのCPUあたりの最大転送レートは、CPあたり3ワードであり; A,SレジスタのCPUあたりのレートは、2CPに1ワードである。

命令バッファへの命令の転送はCPあたり32パーセル(8ワード)のレートである。I/O部の転送レートはCPあたり2ワードである。

中央メモリは以下に要約した特徴があり、これからの段落で詳述する。

- ・両CPUから共有アクセスされる
- ・2Mか4MワードのICメモリ
- ・ワードあたり8エラー訂正ビットと64bitデータ・ビットを持つ
- ・16か32のインターリーブされたバンク
- ・バンクのサイクル・タイムは4CP
- ・1bitエラー訂正、2bitエラー検出(SECDED: Single Error Correction/Double Error Detection)
- ・CPUあたりB,T,Vレジスタについて、CPあたり3ワードの転送レート
- ・CPUあたりA,Sレジスタについて、2CPあたり1ワードの転送レート
- ・命令バッファについて、CPあたり8ワードの転送レート
- ・命令フェッチと交換を除く、すべてのメモリ・アクティビティとI/Oの並行転送レートは、CPあたり2ワード

### - メモリ構成

メモリはすべてのCPUで、高速性と効果的なアクセスを提供するように、構成されている。

メモリから/メモリへのデータ転送は、1bitエラーは訂正され、2bitエラーは検出される。(SECDED) 中央メモリ

は、各セクションに4か8バンクがある、4つのセクションから構成されている。2Mワード・システム(モデル 22)では16バンクにフェーズ化され、4Mワード・システム(モデル 24)では32バンクにフェーズ化される。

図2-1に示すように、各CPUは、4つのセクションへ入っていく独立したアクセス・パスに接続されている。このコンフィギュレーションは1クロック周期で8つまでのメモリ参照が可能である。

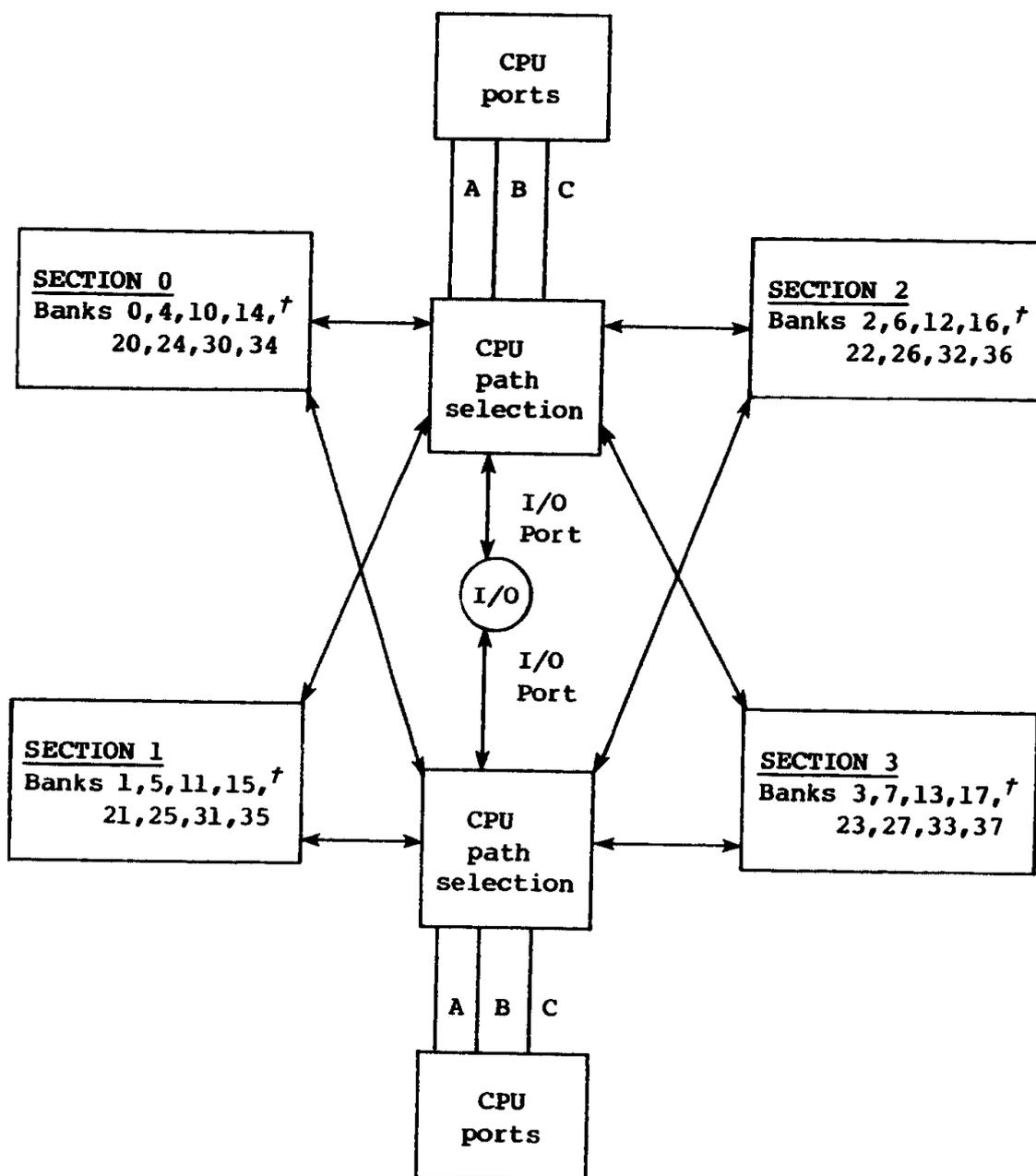


図2-1. デュアル・プロセッサ・市住むの中央メモリ構成 (pp.2-2)

--

+各セクション内の小さな番号の付いた4つのバンクは、16バンク・システムである。

- メモリ・アドレッシング

メモリ・アドレッシングはシステム・メモリ・アーキテクチャ(チップのサイズとバンク数)とメモリ量に依存している。6カラムと12カラムのデュアル・プロセッサ・システムのメモリ・アドレッシングを続く段落で述べる。

o 6カラム・メインフレームのメモリ・アドレッシング

図 2-2 に示すように、32 バンク・メモリ内のワードは、最大 22bit でアドレスされる。下位の 5bit は 32 バンクの 1 つを指定する。次の 14bit のフィールドは、チップ内を指定する。上位の 3bit はモジュール上のチップの 1 つを指定する。

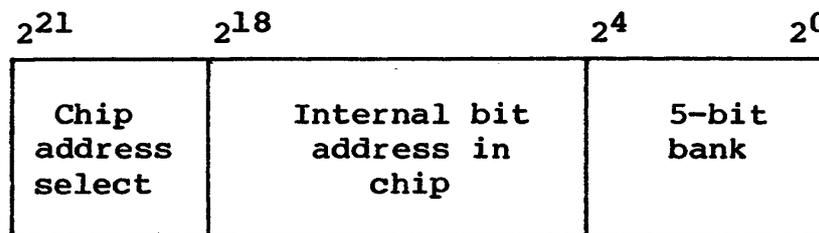


図 2-2. 6カラム・

メモリ・アドレス(32 バンク)

(pp.2-3)

--

図 2-3 に示すように、16 バンク・メモリ内のワードは、最大 21bit でアドレスされる。この場合、下位の 4bit は 16 バンクの 1 つを指定する。次の 14bit のフィールドは、チップ内を指定する。上位の 3bit はモジュール上のチップの 1 つを指定する。(+)

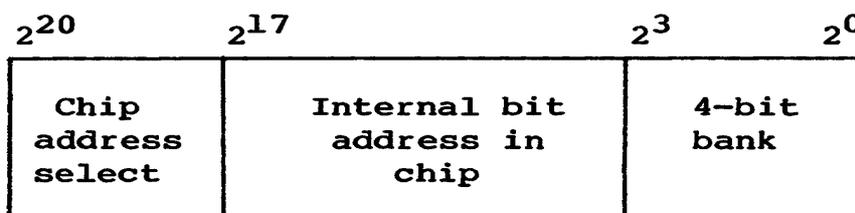


図 2-3. 6カラム・メモリ・アドレス(16 バンク)

(pp.2-3)

--

+ ハードウェアは 4bit のバンク・フィールドを使ってアドレスを組み立てる。ソフトウェアはメモリ・エラー訂正のためのアドレスを作るとき、交換パッケージから重みのついた 5bit を受け取る。上位 bit(右から左へ 0 から 4 数えた bit)は、エラー訂正のためのアドレスを組み立てる時に、ソフトウェアによって捨てられる。

--

o 12 カラム・メインフレームのメモリ・アドレッシング

図 2-4 に示すように、

32 バンク・メモリ内のワードは、最大 22bit でアドレスさ

れる。下位の 5bit は 32 バンクの 1 つを指定する。次の 12bit のフィールドは、チップ内を指定する。上位の 5bit はモジュール上のチップの 1 つを指定する。

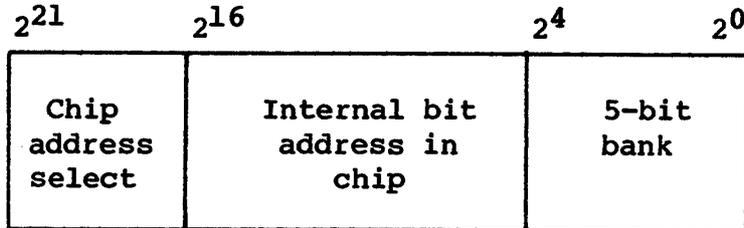


図 2-4. 12 カラム・メモリ・アドレス(32 バンク)

(pp.2-4)

--

図 2-5 に示すように、16 バンク・メモリ内のワードは、最大 21bit でアドレスさ

れる。この場合、

下位の 4bit は 16 バンクの 1 つを指定する。次の 12bit のフィールドは、チップ内を指定する。上位の 5bit はモジュール上のチップの 1 つを指定する。(+)

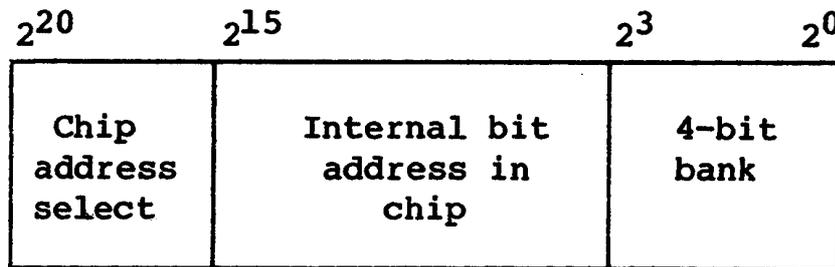


図 2-5. 12 カラム・メモリ・アドレス(16 バンク)

(pp.2-4)

--

+ ハードウェアは 4bit のバンク・フィールドを使ってアドレスを組み立てる。

ソフトウェアはメモリ・エラー訂正のためのアドレスを作るとき、交換パッケージから重みのついた 5bit を受け取る。上位 bit(右から左へ 0 から 4 数えた bit)は、エラー訂正のためのアドレスを組み立てる時に、ソフトウェアによって捨てられる。

--

## - メモリ・アクセス

両 CPU とも、ポート A, ポート B, ポート C, I/O として参照される 4 つのメモリ・アクセス・ポートを持つ。各ポートは、CP ごとに 1 つの参照を行う能力がある。ポート A, B, C は CPU レジスタ転送に使用される。

B, T とベクトル・メモリ命令は特定のメモリ・ポートへ発行する:

- ・ベクトル・リード(ブロック・リードのみ)、B リード命令(176,034)はポート A を使用する。
- ・ベクトル・リード(ブロック・リードのみ)、T リード命令(176,036)はポート B を使用する。
- ・ベクトル・ストア, B か T ストア命令(177,035,037)とスカラ命令(100-137)はポート C を使用する。

一度命令がポートへ発行されたら、ポートはその命令のための参照がすべて行われるまで予約される。

ブロック転送(V, B, T)の各要素のための参照は、ポートを通して順番に行われ、完了する。しかし、各参照は、個々に、起こりえる競合を検査しているため、転送のためのデータ・フローが連続でないこともありえる。もし、命令がビジーなポートを要求したら、発行はブロックされる。

転送の総合的な実行時間は、転送中に合った競合の型と回数に依存する。

\*\*\*\*\*

### 警告

並行したブロック・リードとライトは、ライトの前のリードや、リードの前のライトを検査しない(メモリが被っていると危険な状態(memory overlaphazard conditions)) ので、ソフトウェアがこのような状態の発生の検出と、逐次の実行を保証しなければならない。

\*\*\*\*\*

双方向メモリ・モード許可(0025)、双方向メモリ・モード禁止(0026)、メモリ参照完了(0027) 命令は、これらのケースを解決し逐次操作を保証する。もし、双方向メモリ・モードがクリアされていれば、ブロック・リードとライトは当該 CPU 内で同時に操作することは許されない。命令 0027 は、発行した CPU 内で、競合解決ステージを通過した、ブロック転送の先立つものすべての最後の参照と、転送されたデータが指定されたメモリかレジスタの場所に送られるのを、待つことができる。命令 0027 はプログラム中に必要のない、CPU 内か CPU 間での、逐次メモリ操作を保証するための機構をソフトウェアに提供する。

スカラ・メモリ参照の発行はポート A, B, C が利用可能であることを要求し、CPU 内でのブロック転送とスカラ参照の間で逐次操作を保証する。

スカラ参照競合は、実行の CP3 で検出される。もし、競合が発生したら一つ以上のスカラ参照の発行が許されている。もし先立つスカラ参照のための競合状態が存在しているならば、第 3 のスカラ参照は発行がホールドされる。

スカラ参照は、常に CPU 内で発行された順に実行する。命令 0027 はすべてのスカラ参照がそれを発行した CPU 内で、競合解決ステージを通過した時を検出する。

CPU I/O チャンネルの 1.5 は、各 CPU の I/O ポートを通してメモリを参照する。I/O ポートはポート A,B,C のアクティビティに無頓着にアクティブになることができる。

命令フェッチ要求が発生した時、8 つのメモリ・ポートからのすべての参照が禁止される。メモリが(0 から 3CP) 静かなとき、フェッチは進行し、次の 4CP(もし 16 バンクなら 6CP)で 32 バンクを参照する。そして、8 ポートの参照は許可される。

-----  
注意

スカラ・ストアが後続するフェッチ・シーケンスはある状態のもとでは、ストアの前に完了できる。これが起きるためには、しかしながら、out-of-buffer(バッファの外) 状態が、スカラ・ストアが実行の CP が 2 になる以前に発生していなければならない。out-of-buffer コンディションがスカラ・ストアが実行の CP 2 の以前に発生することは、もし、バッファの境界が分岐とクロスしなければありえる。これは、フェッチとストアがメモリの同じ領域に起こる時だけ、問題を提示する。

よって、動的なコーディングを利用するソフトウェアは、メモリの領域が命令バッファへとフェッチされる前に、生成されたコードが本当にメモリ中にあることを保証しなければならない。

(pp.2-6)

-----

交換は、その交換の要求が成される前に、CPU 内のすべてのアクティビティが完了することを要求する。

交換の要求が成された時、他の CPU の 4 つのメモリ・ポートからのすべての参照は抑制される。メモリが静か(0 から 3CP で) な時、交換は進行し、次の 21CP で 16 のバンクを参照する。各バンクがこの一回の間に 2 度、一回はリード、一回はライトのために、参照される。交換の直後に続くフェッチ要求が完了し、それから、他の CPU の 4 つのメモリ・ポートからの参照が許可される。

## - 競合解決

毎クロック周期の間、システム内のメモリ・ポートの参照は、メモリ・アクセスの競合が検査されている。

もし、参照の競合が発生していたら、参照はホールドされ、競合が解決されるまで、そのポートからのより進んだ参照は一切許されない。

メモリ・アクセスの競合は3つの型が起こりえる:

バンク・ビジー、同時バンク(**Simultaneous Bank**)、セクション・アクセス。

## - バンク・ビジー競合

-- バンク・ビジー競合は、CPU 内か CPU 間のポートの、一つの参照サイクル中の、並行した1つのバンクへの要求が引き起こす。この競合は、バンク・サイクルが完了した時に解決される。バンク・ビジー競合によって、CPU の全ポートは1か2か3CP ホールドされる。

## - 同時バンク(**Simultaneous Bank**)競合

-- 同時バンク(**Simultaneous Bank**)競合は、異なるCPUの2つ以上のポートが同じバンクを要求することで発生する。この競合の解決は優先度(後述の小章「メモリ・アクセス優先度」を参照)に基づいている。1つのCPUの全ポートは同時バンク競合で1CP ホールドされる。バンク・ビジー競合は常に同時バンク競合の後に続く。

## - セクション・アクセス競合

-- セクション・アクセス競合は、同じCPUの2つ以上のポートの、同じセクションのバンクへの要求が引き起こす。最も高い優先度のポートが進行することを許され、この競合に関する他のすべてのポートはホールドされる(後述の小章「メモリ・アクセス優先度」を参照)。ポートはセクション・アクセス競合で1CP ホールドされる。

## - メモリ・アクセス優先度

次の優先度がメモリ・アクセス競合を解決するために、使用される。

### •CPU 内優先度:

ポートA,B,C間の優先度は次の条件で決定される:

- 奇数増加のポートは、偶数増加のポートより、発行されたシーケンスに関わらず、常に高い優先度を持つ。
- 同じ型の増加(奇数でも偶数でも)を持つすべてのポート間の優先度は、発行の相対時間が優先度を決定する、1番に発行されたものが、最も高い優先度をもつ。

•CPU 間の優先度: 各4CPごとに、CPU間の優先度は変わる。

•I/O 優先度: I/OポートはCPU内では常に最も低い優先度である。

## o 16バンク・フェーズ化

命令フェッチの16バンク・フェーズ化の効果は、命令バッファの充填のための予測可能な2CPの増加である。

そうでなく、32バンクの代わりに16バンクであるための性能低下の合計は、メモリ競合の増加から大きな結果が生じるため、簡単には予測できない

メンテナンス用途のために、32バンク・システムは、上位半分だけか下位半分だけのメモリを使用して、16バンクだけで動作するように変更可能である。メンテナンスはバンク選択スイッチを上位か下位のバンクに設定することで行われる。

## o メモリ・エラー訂正

単一エラー訂正/2エラー検出(SECDED)ネットワークはCPUとメモリの間で使用される。SECDEDは、メモリへ書き込まれたデータがCPUへ一貫した正確さで戻されることを保証する(図2-6)。

もし、データ・ワードの1つのbitが変化しているとき、単一エラー変化が、データ・ワードが計算機に渡る前に、自動的に訂正される。もし、同じデータ・ワードの2nitが変化していたら、エラーは検出されるが訂正はされない。どちらのケースも、CPUは割り込まれることができる、エラー処理に割り込みを許可するオプションの選択に依存するが、3bit以上のエラーでは、結果は不確かである。

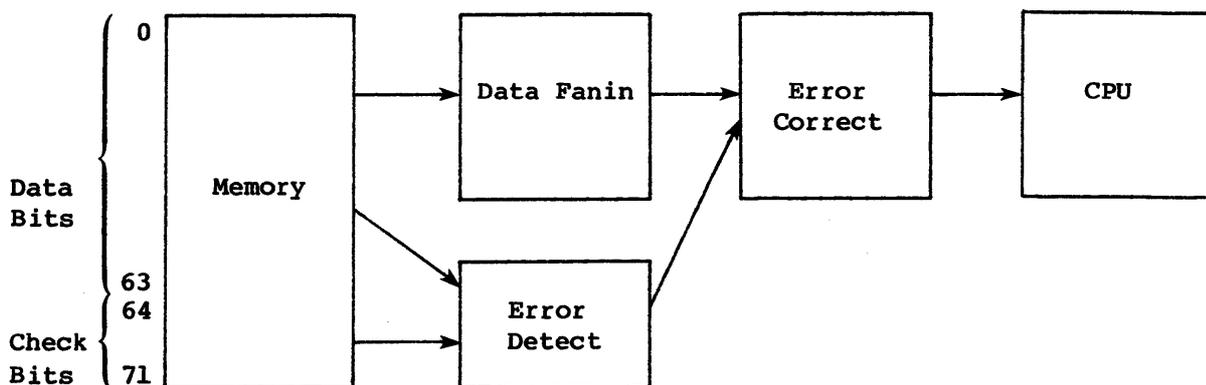


図 2-6. SECDED つきのメモリ・データ・パス

(pp.2-8)

--

SECDED エラー処理の概要は、R.W.Hamming(+ )によって発明されたエラー検出と訂正の符号に基づいている。

( + Hammin,R.W., "Error Detection and Correcting Codes," Bell System Technical Journal, 29, No. 2, pp.147-160(April, 1950). )

データがメモリに書かれる前に、8bit のチェック・バイトが 64bit のデータ・ワードに付加される。

8 チェック・ビットは、データ・ビットの指定されたグループのためにパリティ・ビットのように生成される。図 2-7 は、各チェック・ビットの状態を判定するために使用されるデータ・ワードのビットを示す。

水平の列中の X は、チェック・ビットの生成に寄与するデータ・ビットであることを示す。

このように、チェック・ビット 0 は

bit  $2^1, 2^3, 2^5, 2^7, 2^9, 2^{11}, 2^{13}, 2^{15}, 2^{17}, 2^{19}, 2^{21}, 2^{23}, 2^{25}, 2^{27}, 2^{29}, 2^{31}$  から  $2^{55}$  のグループのグループ・パリティを作るビットである。

8 チェック・ビットとデータ・ワードはメモリの同じ位置に格納される。メモリからの読み出し時、図 2-7 と同じ 64bit 行列が新しいチェック・ビットの集合を作るのに使用され、それらは古いチェック・ビットと比較される。比較結果の 8bit はシンδροーム・ビット(+)(S bit)と呼ばれる。これらの Sbit の状態は発生したすべてのエラーの兆候である (1=コンペアできていない)。もし、全シンδροーム・ビットが 0 なら、メモリ・エラーの無いことが保証される。

(+)Syndrome(シンδροーム):

ある型、状態などを同定するために、注目される特徴の集合。Webster New World Dictionary より。

		CHECK BYTE																							
		2 <sup>71</sup>	2 <sup>70</sup>	2 <sup>69</sup>	2 <sup>68</sup>	2 <sup>67</sup>	2 <sup>66</sup>	2 <sup>65</sup>	2 <sup>64</sup>	2 <sup>63</sup>	2 <sup>62</sup>	2 <sup>61</sup>	2 <sup>60</sup>	2 <sup>59</sup>	2 <sup>58</sup>	2 <sup>57</sup>	2 <sup>56</sup>	2 <sup>55</sup>	2 <sup>54</sup>	2 <sup>53</sup>	2 <sup>52</sup>	2 <sup>51</sup>	2 <sup>50</sup>	2 <sup>49</sup>	2 <sup>48</sup>
check bit 0									x									x	x	x	x	x	x	x	x
check bit 1								x		x	x	x	x	x	x	x	x								
check bit 2							x			x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x
check bit 3						x				x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x
check bit 4					x					x		x		x		x		x		x		x		x	
check bit 5				x						x	x			x	x			x	x			x	x		
check bit 6		x								x	x	x	x					x	x	x	x				
check bit 7	x									x			x		x	x		x			x		x	x	
		2 <sup>47</sup>	2 <sup>46</sup>	2 <sup>45</sup>	2 <sup>44</sup>	2 <sup>43</sup>	2 <sup>42</sup>	2 <sup>41</sup>	2 <sup>40</sup>	2 <sup>39</sup>	2 <sup>38</sup>	2 <sup>37</sup>	2 <sup>36</sup>	2 <sup>35</sup>	2 <sup>34</sup>	2 <sup>33</sup>	2 <sup>32</sup>	2 <sup>31</sup>	2 <sup>30</sup>	2 <sup>29</sup>	2 <sup>28</sup>	2 <sup>27</sup>	2 <sup>26</sup>	2 <sup>25</sup>	2 <sup>24</sup>
		x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x		x		x		x	
		x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x			x	x		
		x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x				
		x		x		x		x		x		x		x		x		x		x		x		x	
		x	x			x	x			x		x		x	x			x	x	x	x	x	x	x	x
		x	x	x	x					x	x	x	x					x	x	x	x	x	x	x	x
		x				x	x			x			x		x	x		x		x		x		x	
		x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x
		x	x	x	x					x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x
		x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x

図 2-7. エラー訂正行列 (pp.2-9)

メモリ中の単一ビットの状態の変化は、シンドローム・ビットの 1 になるべきものの数を奇数にさせる。二つのエラー(2bit がエラー)は、シンドローム・ビットの 1 になるものが偶数として現れる。

行列は次のように設計される:

- 全シンドローム・ビットが 0 なら、エラー無しが保証される。
- ただ 1 つのシンドローム・ビットが 1 ならば、対応するチェック・ビットはエラーである。
- もし 1 つ以上のシンドローム・ビットが 1 で、シンドローム・ビットの S0 から S7 のパリティが偶数ならば、データ・ビットかチェック・ビット中に、二つのエラー(または、ビットのエラーが偶数個)発生している。
- もし 1 つ以上のシンドローム・ビットが 1 で、全シンドローム・ビットのパリティが奇数ならば、単一で訂正可能なエラーが発生したことが保証される。シンドローム・ビットは、エラーの起こっているビットを特定するためにデコードできる。
- もし 3bit 以上のメモリがエラーであると、全シンドローム・ビットのパリティは奇数で、結果はあいまいである。

SECCDED のために使用される 8bit のチェック・バイトの生成、解釈に関するモジュールは、チェック・ビット格納の検証、チェック・ビット生成、エラー検出/訂正のために使用される論理を含む。SECCDED 保守機能についての情報については付録 D を参照のこと。

#### o CPU 間通信セクション

メインフレームの CPU 間通信セクションは 2 つの CPU の間の通信、制御、実時間クロックのために、特別なハードウェアを含む。

実時間クロック(RTC:real-time clock)、共有アドレス(SB)、共有スカラ(ST)、セマフォ(SM)レジスタは CPU によって共有される。これらのレジスタ、これらのソースとデストネーションと共に、は図 2-8 に示され、これからの段落で説明する。

#### - 実時間クロック(RTC)

メインフレームは、2 つの CPU に共有される、1 つの実時間クロック(RTC)レジスタを含む。プログラムは、クロック周期(CP)カウンタを使用することにより、正確に時間を計測することができる。このカウンタは 64bit 幅で、95 ナノ秒のクロック周期毎に 1 カウントが進む。クロックは、プログラム実行と同期して進むので、プログラムの実際の CP 数を計測することに使用できる。しかし、次のようなアプリケーション、終了時刻を読む前に割り込みが発生するような場合、計数は、別なタスクからのカウントを含む。

RTC レジスタと使用される命令は:

0014j0 RT Sj RTC レジスタに(Sj)を入れる  
072i00 Si RT (RTC)から Si へ送信

プログラムは、命令 072 を使用して CP カウンタを読み、それを命令 0014j0 でリセットする。  
 CP カウンタのロードか読み込みは全 CPU から同時に起きてもよい。もし、一つ以上の CPU がモニタモードなら、一つの CPU だけがこのレジスタに値を入れることを、ソフトウェアが保証しなければならない。

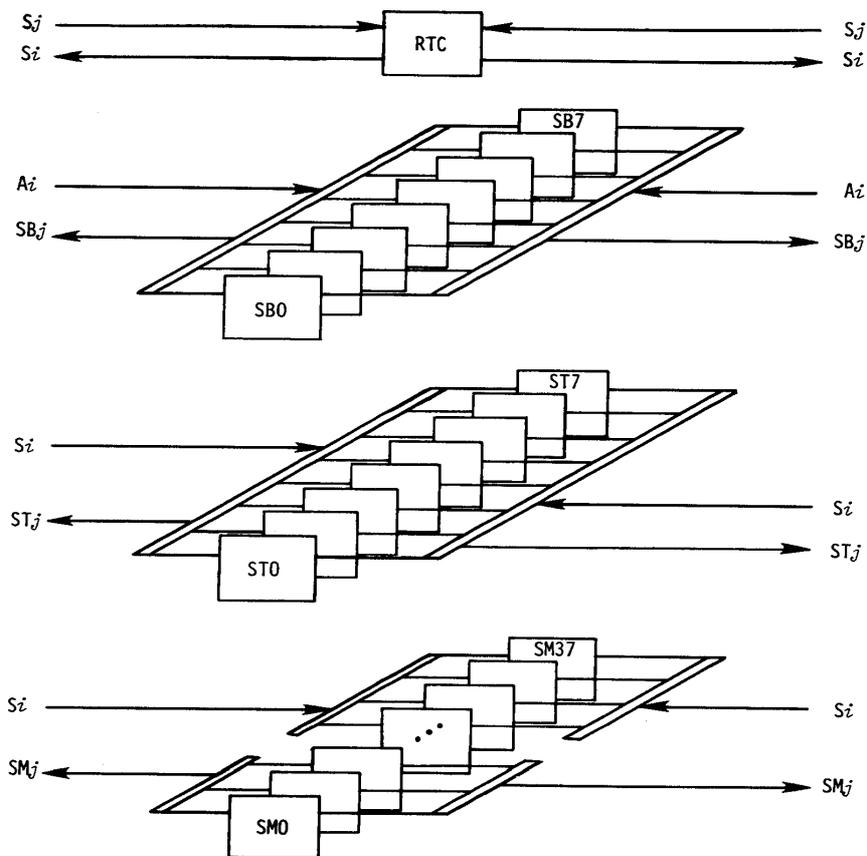


図 2-8. 共有レジスタと実時間クロック  
 (pp.2-22)

## o CPU 間通信と制御

3つの同様な共有レジスタの組が CPU 間の通信と制御に使用される。各組は、8つの 24bit 共有アドレス (SB)レジスタ、8つの 64bit 共有スカラ(ST)レジスタ、32個の 1bit セマフォ・レジスタを含んでいる。

各 CPU のクラスタ番号(CLN)レジスタは、(クラスタ指定中の)CPU によってアクセスされる共有レジスタの組を決定する。CLN レジスタは交換パッケージからか、または、CPU がモニタ・モードならば、命令 0014j3 を通してロードされる。

CLN レジスタは 4つの異なる値のうちの一つを持つことができる。値 1,2,3 は、共有レジスタの 3つの組の 1つに CPU がアクセスすることを許す。値 0 は、共有レジスタの CPU によるいかなるアクセスもできなくする。もし値が 0 なら、共有レジスタに対する命令は no-op になる、例外として、値を Ai か Bi に返す命令は戻り値が 0 となる。もし、両 CPU の中の CLN レジスタが同じ値(1,2,3)のセットされたら、2つの CPU は共通 SB,ST,SM の組を共有する。

## o 共有アドレスと共有スカラ・レジスタ

共有アドレス(SB)と共有スカラ(ST) レジスタは、CPU から他にアドレスやスカラの情報を渡すために使用される。これらのレジスタにはハードウェア予約が成されることはない。これらのレジスタへの制限されたアクセスのために必要な予約は、セマフォ(SM)レジスタの使用を通したり、共有メモリ設計による、ソフトウェアによって取り扱われなければならない。SB、ST レジスタへのアクセスの一つのハードウェア制限は、1つの CP では 1つの読み出しか 1つの書き込み動作しかできないことである。

SB,ST レジスタに使用される命令は:

026ij7 Ai SBj (SBj)を Ai へ送信  
027ij7 SBj Ai (Ai)を SBj へ送信  
072ij3 Si STj (STj)を Si へ送信  
073ij3 STj Si (Si)を STj へ送信

共有アドレス(SB)と共有スカラ(ST)レジスタの競合したアクセスは、表 2-1 に示される、無頓着なクラスタリングの状態のもとで発生する。例えば、もし CPU0 のリード命令と CPU1 のリード命令が同時の CIP で入ったら、競合が発生し、CPU1 は 1CP の間、発行がホールドされる。

## - セマフォ・レジスタ

セマフォ(sm)レジスタは、CPU 間の制御に使われる。これらのレジスタには、ハードウェアの予約は一切、成されない。SM レジスタのロード、リード、あるいは、ある SM レジスタへのセット、クリアはいつでも片方かまたは両方の CPU から起こすことが可能である。

テスト・アンド・セット命令(0034jk)は SM レジスタへだけの命令であり、ハードウェア・インターロックを含む。このインターロックは、両 CPU からの同じ SM レジスタへの、同時のテスト・アンド・セット動作を防ぐ。

テスト・アンド・セット命令は、まず、選ばれた SM レジスタの値をテストする。もし、値が 0 なら命令は発行しその SM レジスタを 1 にセットする。もし、値が 1 なら、命令は値が 0 になるまで発行をホールドされる。

クラスタ中の全 CPU がテスト・アンド・セット命令でホールドされた時、デッドロック割り込みが発生できる。

もし、両 CPU 内の CLN レジスタ が同じで、0 でなければ、両 CPU は同じクラスタに属し、両 CPU はテスト・アンド・セット命令の発行がホールドされるべきで、デッドロック割り込みを発生させる。これが起きたとき、クラスタ中の両 CPU とも、デッドロック割り込みを受け取る。

もし、両 CPU 内の CLN レジスタが異なるとき、2 つの CPU は違うクラスタにある。もし 1 つの CPU がテスト・アンド・セット命令の発行がホールドされると、その CPU はデッドロック割り込みを受け取る。クラスタ 0 (CLN=0) 内なら、デッドロック割り込みは発生しない。

表 2-1. デュアル・プロセッサでの共有レジスタのアクセス競合  
(pp.2-13)

SB or ST register operation		Hold issue 1 CP
CPU 0	CPU 1	
READ (first CP in CIP)	READ (first CP in CIP)	CPU 1
READ (not first CP in CIP)	READ (first CP in CIP)	CPU 1
READ (first CP in CIP)	READ (not first CP in CIP)	CPU 0
READ (not first CP in CIP)	READ (not first CP in CIP)	CPU 0
WRITE (first CP in CIP)	WRITE (first CP in CIP)	CPU 1
WRITE (not first CP in CIP)	WRITE (first CP in CIP)	CPU 1
WRITE (first CP in CIP)	WRITE (not first CP in CIP)	CPU 0
WRITE (not first CP in CIP)	WRITE (not first CP in CIP)	CPU 0
READ (Write issued 3 CPs before)		CPU 0
READ	(Write issued 3 CPs before)	CPU 0
	READ (Write issued 3 CPs before)	CPU 1
(Write issued 3 CPs before)	READ	CPU 1

割り込みが発生した時、通常、すでに NIP と CIP レジスタに入っている命令は、交換シーケンスが始まる前に発行することができる。もし、CIP レジスタ中でテスト・アンド・セット命令がホールド中で割り込みが発生したら、特別な交換起動シーケンスが開始される。このケースでは、NIP レジスタの命令と CIP レジスタ中のテスト・アンド・セット命令は捨てられ、プログラム・カウンタ(P) レジスタは捨てられたテスト・アンド・セット命令を指すように修正される。交換パッケージ内の、セマフォで待っていた(WS:Waiting on Semaphore)フラグがセットされ、割り込みが発生した時に、CIP レジスタ中でテスト・アンド・セット命令がホールドされていたことを示す。そして、交換シーケンスは開始される。

SM レジスタに使用される命令は:

0023jk SMjk 1,TS SMjk にテスト・アンド・セット  
0036jk SMjk 0 SMjk をクリア  
0037jk SMjk 1 SMjk をセット  
072i02 Si SM (SM)を Si に送信  
073i02 SM Si (Si)を SM に送信

#### o CPU I/O セクション

メインフレームの I/O セクションは 2 つの CPU から共有されている。メインフレームは、最大転送レート 1250MByte/秒、100MByte/秒、6MByte/秒で識別される、3 つの型のチャンネルをサポートする。

1 つの 1250MByte/秒チャンネルのペアは中央メモリと半導体ストレージ・デバイス(SSD)の間のデータ転送に使用される。これらのチャンネルは、128bit 幅で各方向に 16bit のチェック・ビットが使われる。10Gbit/秒を超える最大転送レートが、1250MByte/秒チャンネルで可能である。チャンネルは、2 つの並列 64bit チャンネルで、それぞれ SECDED 付きである; そして、ある状況下では、全幅のチャンネルは 2 つのエラーを訂正できる。

2 つの 100MByte/秒チャンネル・ペアは、中央メモリと I/O サブシステムの間でデータを転送する。100MByte/秒チャンネルは 64bit 幅で各方向に 8bit のチェック・ビットを使用する。データ・ワードは、DataReady(データ・レディ)と DataTransimit(データ・送信)の制御信号の制御のもと、16 のブロックで転送される。各 100MByte/秒チャンネルの、最大転送レートは約 850Mbit/秒である。

I/O サブシステムの CPU との通信は、コントロール・チャンネルの 4 つのペアを使い、各々、6MByte/秒の最大転送レートを持つ。

各 CPU からは、1 つの I/O ポートがある。チャンネルは、2 つの 6MByte/秒チャンネル・ペア、1 つの 100MByte/秒チャンネル・ペア、ポートあたり 1.5 個の 1250MByte/秒チャンネル、をもって、ポートへと、ハードウェア布線されている。各ポートは CP ごとに 1 ワードのレートで、データを転送することができる。100MByte/秒チャンネルと 1250MByte/秒チャンネルでは、バッファが参照を行うたびに、その参照は、完了するまで(通常 16 ワード分) ポートを保持する。

すべての I/O(100MByte/秒と 1250MByte/秒チャンネルを含む) は、メモリへの I/O ポートを使用する。これらのポートへのアクセスはスキヤナによって制御されている。すべての CPU のメモリ・ポート(ポート A,B,C)は I/O ポートよりも高い優先度を持っている。

I/O セクションのチャンネルの特徴は、以下に要約し、この章の残りで解説している。

- ・チャンネルあたりの最大転送レート **1250MByte/秒**のチャンネル・ペアが **1** つ
  - 各方向あたり、**128** データ・ビットで **16** のチェック・ビット付き
- ・チャンネルあたりの最大転送レート **100MByte/秒**チャンネル・ペアが **2** つ
  - 各方向あたり、**64** データ・ビット,**3** 制御ビット,**8** チェック・ビット付き
- ・チャンネルあたりの最大転送レート **6MByte/秒**チャンネル・ペアが **4** つ
  - CPU からの共有制御
  - 各方向あたり、**16** データ・ビット,**3** 制御ビット,**4** パリティ・ビット付き
  - データの消失検出
- ・チャンネルは **4** つのグループに分けられる、各グループは入力か出力かのチャンネルを含んでいる。
- ・チャンネル・グループはメモリによって均等に扱われる。(各グループは **4CP** 毎にスキャンされる)
- ・チャンネル優先度はチャンネル・グループによって解決される

#### - 半導体ストレージ・デバイスのためのデータ転送

データは半導体ストレージ・デバイス(SSD)とメインフレームの間を、**1250MByte/秒**チャンネルを使用して、直接に転送される。**1250MByte/秒**チャンネルは **128bit** 幅でソフトウェアを通してプログラムされる。SSD のポート **3** が、**CRAY X-MP** システムに接続される。

SSD のためのプログラミングの詳細は「半導体ストレージ・デバイス(SSD)リファレンス・マニュアル」,CRI 出版 HR-0031 に説明されている。

#### - I/O サブシステムのためのデータ転送

**100MByte/秒**チャンネル・ペアは、中央メモリと I/O サブシステムのバッファ I/O プロセッサ(BIOP)との間で、データを転送する。**2** 番目の **100MByte/秒**チャンネル・ペアは、中央メモリと、ディスク I/O プロセッサ(DIOP)か補助 IO プロセッサ(XIOP)との間でデータを転送できる(+). 各チャンネルは **64bit** 幅で約 **100MByte/秒**でデータを取り扱う。各チャンネルは単一エラー訂正/2 エラー検出(SECDED)のために、中央メモリが使うように、**8** チェック・ビットを付加して使用する。

**100MByte/秒**チャンネル・ペアの CPU 側は、**16** ワードのバッファの **1** つのペアを中央メモリから出るデータを流す(stream)ために、別のペアを中央メモリへデータを流すために、使用する。出力では、**1** つのバッファ・ブロックが I/O プロセッサ(IOP)に送られているとき、別なバッファは中央メモリから満たされている。同様に、入力では、**1** つのバッファ・ブロックが中央メモリに送られている間に、別な方が IOP から満たされている。

**100MByte/秒**チャンネル・ペアの IOP 側では、ローカル・メモリ(IO プロセッサのメモリ)へのデータの渡しはダブル・バッファ化されていて、**16bit** パーセルに分解されている。ローカル・メモリから CPU へ送信するための、チャンネル側のデータ渡しは、単純に **16bit** パーセルを **64bit** ワードに組み立てる。

I/O プロセッサは、**100MByte/秒**チャンネル・ペアと中央メモリとの接続を制御する。IOP は、チャンネルにおけ

るすべてのデータ転送の開始と、チャンネルに必要なエラー処理のすべてを実行する。100MByte/秒チャンネル・ペアには CPU 命令は必要ない。100MByte/秒チャンネル・ペアのプログラミングの詳細については「I/O サブシステム・リファレンス・マニュアル」,CRI 出版 HR-0030 に含まれている。

(+) 現在、ソフトウェアは、100MByte/秒チャンネル・ペアを使用した、XIOP へのデータ転送をサポートしていない。

### o 6 MByte/秒チャンネル

システムのための標準制御チャンネルが 6MByte/秒チャンネルである。各 6MByte/秒チャンネルは、フロントエンド・インターフェースのために使用される 16bit の非同期制御論理を持つ。

6MByte/秒チャンネルに使用される命令は次である:

0010jk CA,Aj Ak (Aj)で指定されるチャンネルの、現在アドレス(Current Address)(CA)レジスタを(Ak)にセットし、チャンネルをアクティベートする。

0011jk CL,Aj Ak (Aj)で指定されるチャンネルの、リミット・アドレス(Limit Address)(CL)レジスタを(Ak)にセットする。

0012jk CI,Aj (Aj)で指定されるチャンネルの、割り込みフラグとエラー・フラグをクリアする:

出力チャンネル k=0;MC をクリア, k=1;MC をセット。

入力チャンネル k=0;NOP, k=1;ホールドされている待ちをクリア。

033i00 Ai CI 送信チャンネル番号を Ai にする

033ij0 Ai CA,Aj チャンネル(aj)の送信アドレスを Ai にする

033ij1 Ai CE,Aj チャンネル(aj)のエラー・フラグを Ai にする

## - マルチ CPU プログラミング

6MByte/秒 I/O チャンネルはどちらの CPU からも操作でき、どちらの CPU もチャンネルへの命令を発行できる。CPU 間のハードウェア・インターロックは存在しない、よって、モニタ・モードの間は、ソフトウェアが、一時にただ一つの CPU だけが I/O をサービスするように保証しなければならない。命令 033 は、まったく独立で、インターロックなしで発行できる。

次の状態は、I/O 割り込みを発生する:

- どちらの CPU も交換の待ちに無い
- どちらの CPU もモニタ・モードに無い
- 割り込みが存在する

通常、6MByte/秒チャンネルからの割り込みは、最後に、そのチャンネルに対して割り込みクリア命令(0012)を発行した CPU へ導かれる。しかし、一時には 1 つの CPU の中だけで I/O 割り込みが発生するので、後述の状況(優先順序どおり)が、どの CPU へ割り込みを導くかを決定する。

1. 全 I/O 割り込みは、外部割り込み選択モード(Select External Interrupt Mode)をセットした CPU に、導かれる。
2. もしどの CPU も外部割り込み選択モードを選んでいなければ、割り込みは、テスト・アンド・セット命令の発行でホールドされている CPU に導かれる。
3. 状況 1 も 2 も存在しないか、または、もし両 CPU ともにそれらが存在していれば、割り込みは、最後に、チャンネルに向けて割り込みクリア命令を発行した CPU に導かれる。

o 6Mbyte/秒チャンネル操作

各入力や各出力チャンネルは中央メモリに直接にアクセスする。入力チャンネルは外部データをメモリに格納し、出力チャンネルはメモリからデータを読む。チャンネルの一番の仕事は、中央メモリの **64bit** ワードを **16bit** パーセルに、また、**16bit** パーセルを中央メモリの **64bit** ワードに変換することである。4 パーセルは1つの中央メモリ・ワードを作り上げ、パーセルのビットは表 2-2 に示す、メモリのビットの位置に割り当てられる。入力、出力の両操作では、常にパーセル **0** が最初に転送される。

入力か出力の各チャンネルは、データ・チャンネル(4 パリティ・ビット、16 データ・ビット、3 制御線)、64bit 組み立てや分解レジスタ、チャンネル現在アドレス(CA)レジスタ、チャンネル・リミット・アドレス(CL)レジスタ、を持っている。

3つの制御信号(レディ(Ready), 再開(Resume)、切断(Disconnect)) は、チャンネルを使ったパーセルの転送を同調させる。3つの制御信号に加えて、1つのペアの出力チャンネルは、マスタ・クリア線を持つ。付録 B は 6MByte/秒チャンネルの信号シーケンスについて記述している。

表 2-2. チャンネル・ワードの組み立て/分解  
(pp.2-18)

Characteristic	Bit position	Number of bits	Comment
Channel data bits	2 <sup>15</sup> -2 <sup>0</sup>	16	Four 4-bit groups
Channel parity bits		4	One per 4-bit group
CRAY X-MP word	2 <sup>63</sup> -2 <sup>0</sup>	64	
Parcel 0	2 <sup>63</sup> -2 <sup>48</sup>	16	First in or out
Parcel 1	2 <sup>47</sup> -2 <sup>32</sup>	16	Second in or out
Parcel 2	2 <sup>31</sup> -2 <sup>16</sup>	16	Third in or out
Parcel 3	2 <sup>15</sup> -2 <sup>0</sup>	16	Fourth in or out

--

I/O 割り込みは次で発生させられる:

- ・全出力チャンネルにおいて、もし(CA)が(CL)と同じになったとき、最後のパーセルの転送のための再開が割り込みをセットする。
- ・外部デバイス切断は、どの入力チャンネルでも受け取られ、チャンネルをアクティブにする。
- ・チャンネル・エラー状況が発生した(この章の後ろで説明)

割り込みを発生させたチャンネルの番号は、命令 **033** を使用することで判断できる。命令 **033** は、割り込み要求中で、最も高い優先度のチャンネルの番号を **Ai** に読み込む。最も低い番号の付いたのチャンネルが最も優先度が高い。次の、最上位優先度のチャンネルからの割り込みが、もしあって、それが感知された時、割り込み要求はモニタ・プログラムによってクリアされるまで続く。すべての割り込みは命令 **033** を通してどちらかの CPU で利用可能である。6MByte/秒チャンネルのチャンネル番号は 10(八進)から 17(八進)(10/11, 12/13, 14/15, 16/17 - 偶数は入力、奇数は出力)である。

o 入力チャンネル・プログラミング

入力操作を開始するには、CPU はプログラムする(図 2-9 参照):

- 1.チャンネル・リミット・アドレスを最後のワード・アドレス+1(LWA+1)にセットする。
2. チャンネル・カレント・アドレスを最初のワードのアドレス(FWA)にセットする。

カレント・アドレスの設定は、チャンネル・アクティブ・フラグのセットを起こす。そして、チャンネルは、データを受け取る準備ができる。4 パーセル・ワードが組み立てられたら、ワードは CA レジスタの内容のアドレス位置のメモリに格納される。ワードがメモリによって受け付けられた時、カレント・アドレスは 1 進められる。

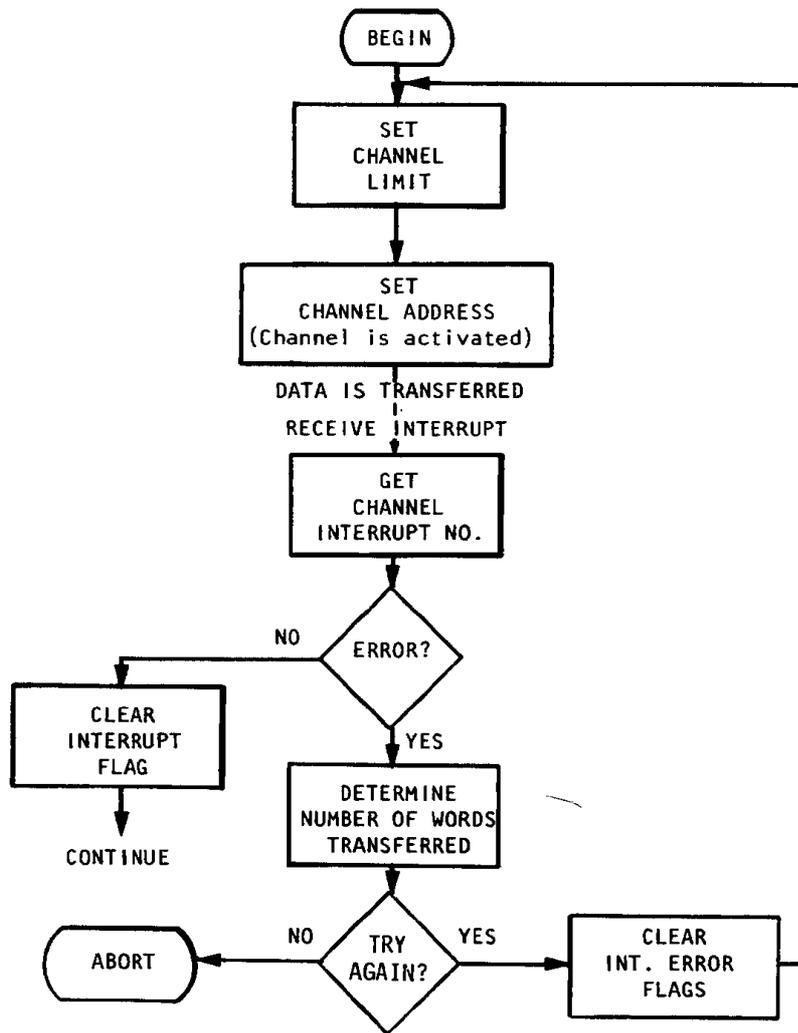


図 2-9. 基本 I/O プログラムのフローチャート  
(pp.2-19)

外部送信デバイスは、転送の終わりを示すために、切断信号を送る。切断信号が受け取られた時、チャンネル割り込みフラグ(**Channel Interrupt Flag**)がセットされ、部分的に組み立てられたワードのチェックするためのテストが実行される。もし、部分的なワードが見つかったら、ワードの正しい部分がメモリに格納され、未受信の下位のパーセルはゼロとして格納される。

割り込みフラグは、切断信号が受信された時か、チャンネル・エラー・フラグがセットされた時に、セットする。

#### o 入力チャンネル・エラー状態

入力チャンネル・エラー状態はパーセル・レベル(パリティ・エラー)か、チャンネル・レベル(予期しない **Ready**(レディ)信号)で発生する。パーセルでエラーが発生したとき、パリティ・フォールト・フラグが即座にセットされる。パリティ・フォールト・フラグは割り込みを生成せず、それは、保存され、切断が起きた時にエラー・フラグをセットする。よって、割り込みが受け付けられた時、プログラムは、エラー・フラグの状態をチェックしなければならない。エラーの後に格納されたパーセルは、すべて **0** になっている。

チャンネルがアクティブでない時(予期しない **Ready** 信号)に、**Ready**(レディ)信号が受信されたら、レディ状態はチャンネルがアクティベートされるまで、ホールドされる。このとき、**Resume**(再開)信号は送られる。エラー・フラグはセットされず、割り込み要求は生成されない。チャンネルが非アクティブな時レディ状態はホールドされるので、チャンネルをセットアップする前に、この **Ready**(レディ)信号をクリアすることが可能であり、時々、有利である。それは、特に、エラー後のチャンネルの再同期や **deadstart**(※訳注:謎)の時に。**Ready** 信号は、命令 **0012j1** を入力チャンネル(**Ai**)に使用することで、クリアできる。(※訳注 **Ai** は **Aj** の誤りか?) 命令 **0012j1** の発行前にホールドされているいかなる **Ready** 信号もクリアする。

#### o 出力チャンネル・プログラミング

入力操作を開始するには、**CPU** がプログラムする:

- 1.チャンネル・リミット・アドレスを最後のワード・アドレス+1(**LWA+1**)にセットする。
- 2.チャンネル・カレント・アドレスを最初のワードのアドレス(**FWA**)にセットする。

カレント・アドレスの設定は、チャンネル・アクティブ・フラグのセットを起こす。そして、チャンネルは、ワードは **CA** レジスタの内容でアドレス指定されたメモリから最初のワードを読む。ワードがメモリから受け取られた時、カレント・アドレスは **1** 進められ、データ転送を開始する。

各ワードがメモリから読まれ、カレント・アドレスが進められた後、**CA** レジスタと **CL** レジスタの内容を比較することによって、リミットのテストが行われる。もし、それらが同じなら、最後のパーセルの転送が終わった瞬間に、操作は完了である。

もしエラーが検出されたら、割り込みフラグもセットされる。出力チャンネルが検出するただ一つのエラーは、チャンネルが非アクティブの時に、**Resume**(再開)信号が受信することである。外部への応答は生成されない。

#### o プログラムされた外部機器へのマスタ・クリア

システムは、出力チャンネルを通して、マスタ・クリア信号(Master Clear signal)を外部機器に送ることができる。外部マスタ・クリア・シーケンスは次の通り:

##### 1. 0012jk

チャンネル・ペアの外部アクティビティが止まったことを保証するために、入力チャンネルをクリアする。

##### 2. 0012j1

チャンネル・ペアでの CPU のアクティビティが止まったことを保証するために、出力チャンネルをクリアする。マスタ・クリアをセット。

##### 3. Delay 1 (待ち 1)

デバイス依存;マスタ・クリア信号の持続時間を決定する。

##### 4. 0012j0

出力チャンネルをクリアする。マスタ・クリア信号をオフにする。

##### 5. Delay 2 (待ち 2)

デバイス依存;取り付けられた装置の初期化アクティビティが完了するための時間。

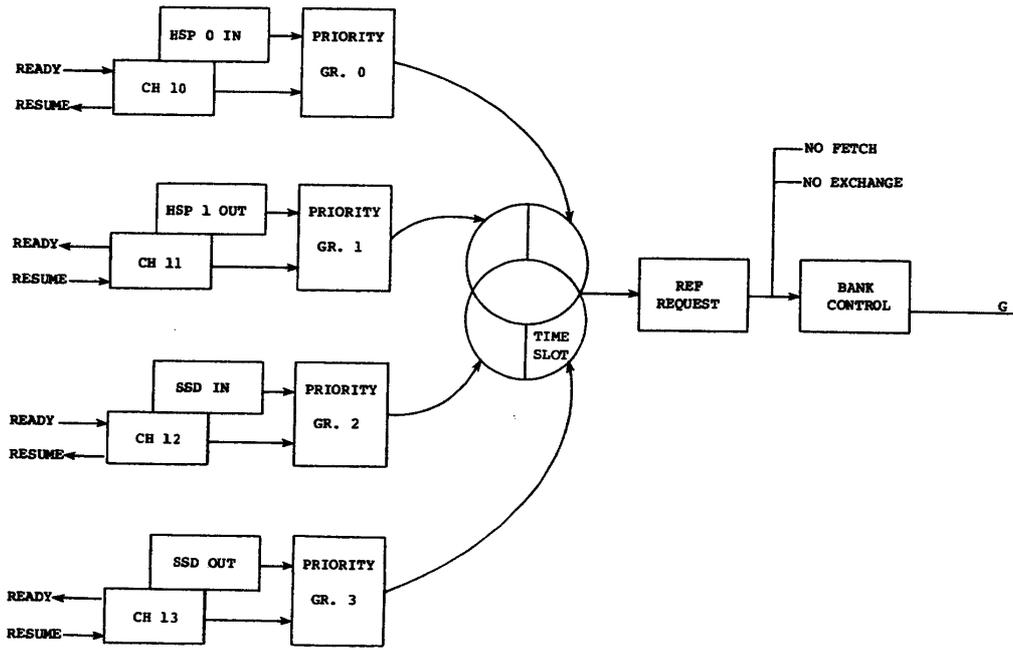
Cray リサーチでは、フロントエンド・インターフェースは、Delay1 と Delay2 は、各々、最小 80CP 必要である。

#### o メモリ・アクセス

後に示すように、4 つのチャンネル・グループの各々はタイムスロットが割り当てられており、メモリ要求は、4CP 毎に 1 度、スキャンされている。グループ内のもっとも低い番号のチャンネルが、もっとも高い優先度を持つ。次の 3CP の間、スキャナは、他の 3 つのチャンネル・グループからの要求を許す。よって、I/O メモリ要求は、毎 CP に発生することができる。

スキャナは、I/O 参照によって引き起こされた全メモリの競合によって停止する。また、スキャナは、バッファが参照している間(最大 16 ワード)、のブロック(100MByte/ 秒チャンネル)参照のためにも、止まる。

REFERENCE CONTROL



ADDRESS

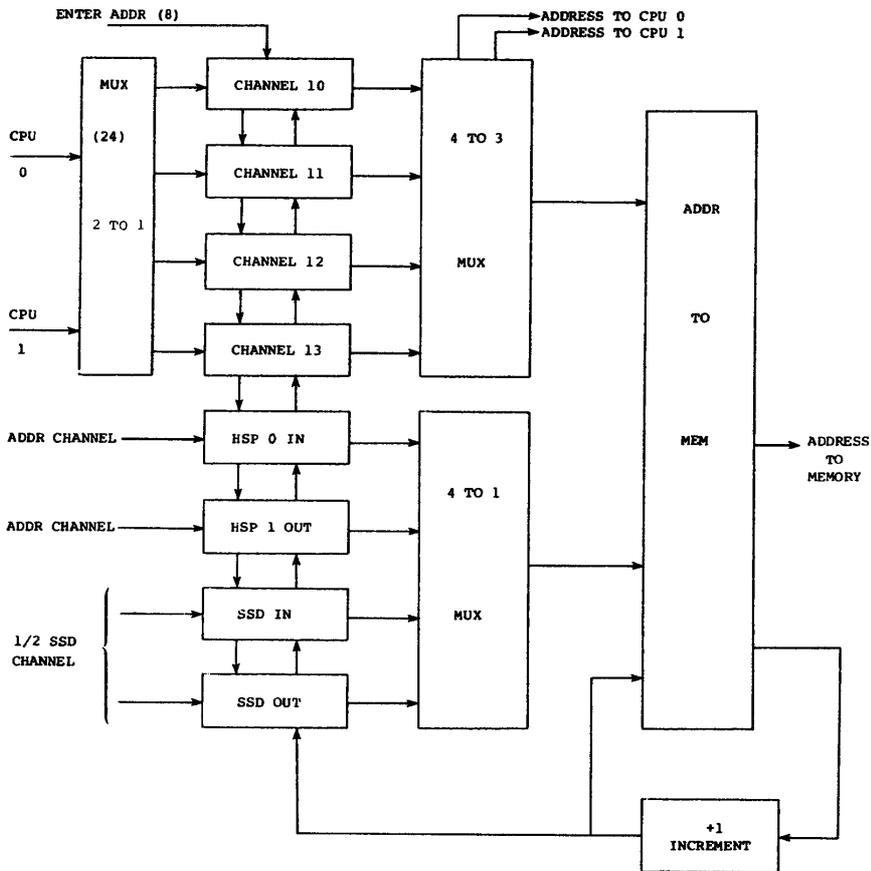
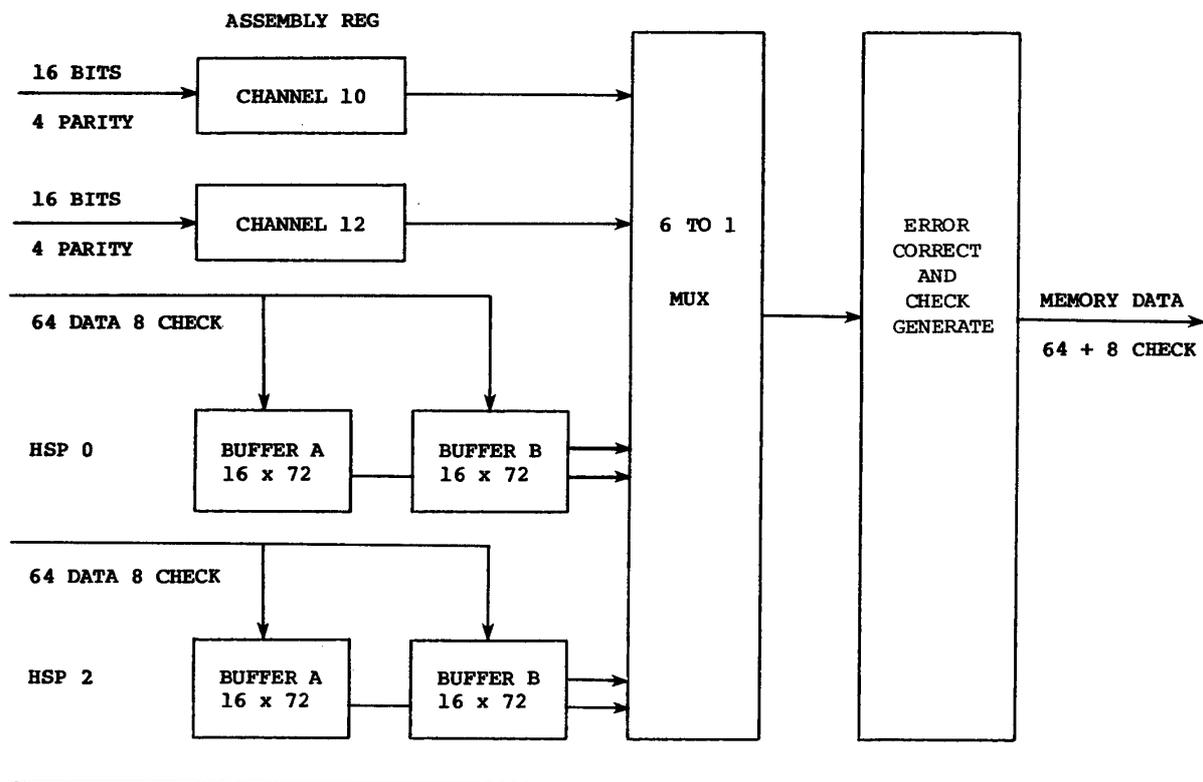


図 2-10. チャンネル I/O 制御(1 プロセッサ分のみ示す)  
(pp.2-22)

INPUT DATA PATH



OUTPUT DATA PATH

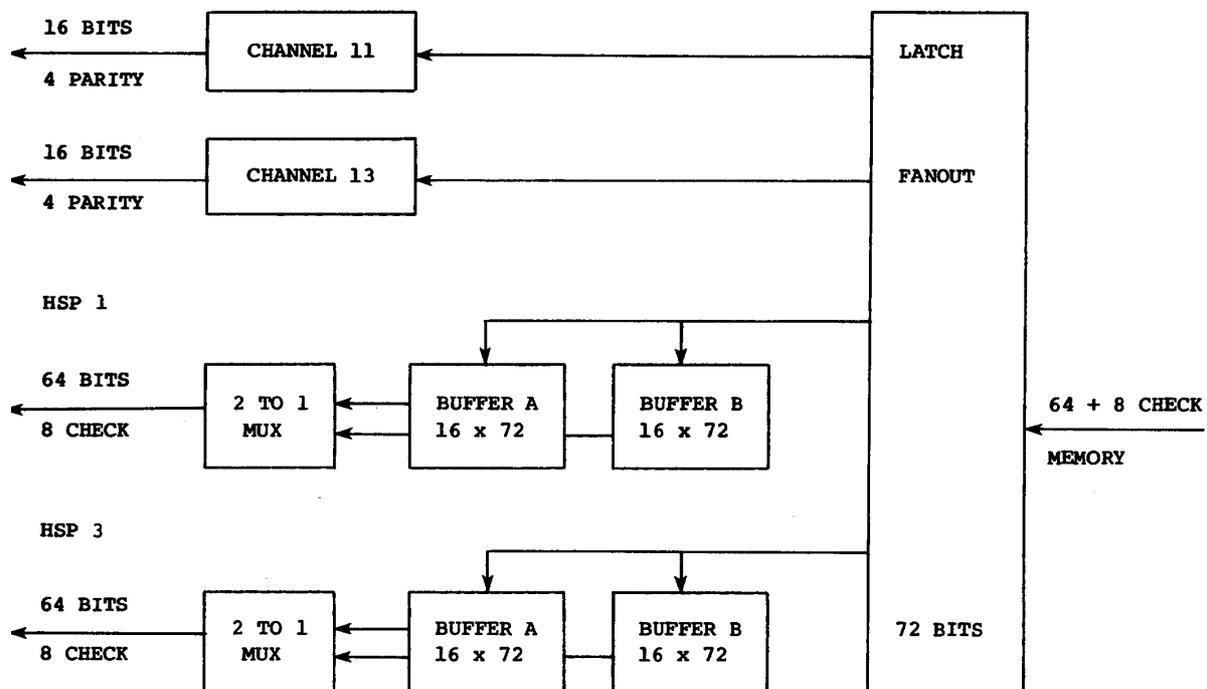


図 2-11. 入出力データパス  
(pp.2-23)

6Mbyte/秒チャンネルは10(八進)から17(八進)まで番号が付いている。100MByte/秒チャンネルは、各CPUの中で、0から3まで番号が付いている(SSDチャンネルは両CPUのチャンネル2と3を使用する)。チャンネルは次のようにグループ化されている:

	CPU0	CPU1
グループ0 入力チャンネル	0,10	0,14
グループ1 出力チャンネル	1,11	1,15
グループ2 入力チャンネル	2,12	2,16
グループ3 出力チャンネル	3,13	3,17

#### o I/O ロックアウト

I/O メモリ要求は交換シーケンスか命令フェッチ・シーケンスによって、ロックアウト(閉め出し)されることがある。

#### o メモリ・バンク競合

メモリ・バンク競合は、CPUのスカラ、ベクトル、I/Oのメモリの参照のために検査される。交換シーケンスか命令フェッチ・シーケンスが進行中の時、他のすべてのメモリ参照は、ロックアウトされる。

各メモリ・バンクは、4CP毎に、新しい要求を受け付けることができる。メモリ・バンク競合を検査するために、メモリ・アドレスの下位5bit(+)を、バンク・ビジー競合と他のメモリ参照に対してチェックする。

(+: 16バンク・フェーズ化の時は、4bit; 中央メモリの節を参照)

#### o I/O メモリ競合

メモリ・バンク競合の検査の前に、交換シーケンスか命令フェッチ・シーケンスが進行中でないことを、保証するためのチェックが行われる。もし、これらのどちらかの状況が存在していれば、I/O要求はホールドされる。I/O参照のアドレスの下位5bit(+)が、バンク・ビジー競合と他のメモリ参照に対してテストされる。もし、バンクが参照によってビジーなら、参照はホールドされ、スキャナは停止する。

(+: 16バンク・フェーズ化の時は、4bit; 中央メモリの節を参照)

## o I/O メモリ要求状態

I/O メモリ要求が処理されるためには、これから述べる状態でなければならない:

### ・I/O 要求

- ・バンクはビジーでない
- ・他のメモリ・ポートに同時の競合が無い
- ・フェッチ要求が無い
- ・交換シーケンスが無い

## o I/O メモリ・アドレッシング

すべての I/O メモリ参照は絶対指定である。CA と CL レジスタは 22bit であり、I/O アクセスはすべてのメモリへのアクセスができる。CA と CL レジスタへのセットは、モニタ・モードに限られている。I/O メモリ参照アドレスは、範囲エラーのチェックがなされない。

--EOF