

1 システム記述

o はじめに

CRAY X-MP/22と**CRAY X-MP/24**は、2つの中央処理ユニット(CPU)を含む強力で汎用の計算機システムである。

両方のCPUのランダムアクセス、半導体メモリ(RAM)と共有レジスタが、スカラーとベクトル処理の能力を効果的に使用することで、システムは高い並列処理率を実現している。

ベクトル処理は、順序のあるデータの集合への繰り返し操作の実行である。2つ以上のベクトル操作は一緒にチェインされる。2つ以上の操作は伝統的なスカラー処理をはるかに超える計算率で、9.5ナノ秒クロックごとに実行する。

スカラー操作は、ベクトル技術がまだすぐに適用できない問題に解決を提供することによってベクトル機能を補う。

周辺オプションは、特別な使用(表 1-1 参照)のためにシステムがコンフィギュアされることを可能にする。

2 プロセッサ・システムの中央メモリは64bitワードの2M(モデル 22)か4M(モデル 24)のいずれかが可能である。

システムはCray I/Oサブシステムの既存モデルと互換である。それは、大容量記憶ユニット、他の周辺デバイス、様々なホスト計算機などとの高い転送レートをもつメインフレームの処理率とマッチしている。

加えて、メインフレームとI/Oサブシステム、CRAYリサーチ製半導体記憶デバイスはシステムとコンフィギュレーション可能である。SSDは大きなデータファイルを繰り返しアクセスするプログラムの大きく改善されたスループットを提供する。

図 1-1 はCray I/OサブシステムとSSDを持ったメインフレーム・コンフィギュレーションを図解する。

この節はシステム・コンポーネントとコンフィギュレーションを説明する。表 1-1 はシステムの特徴のすべてを提供する。

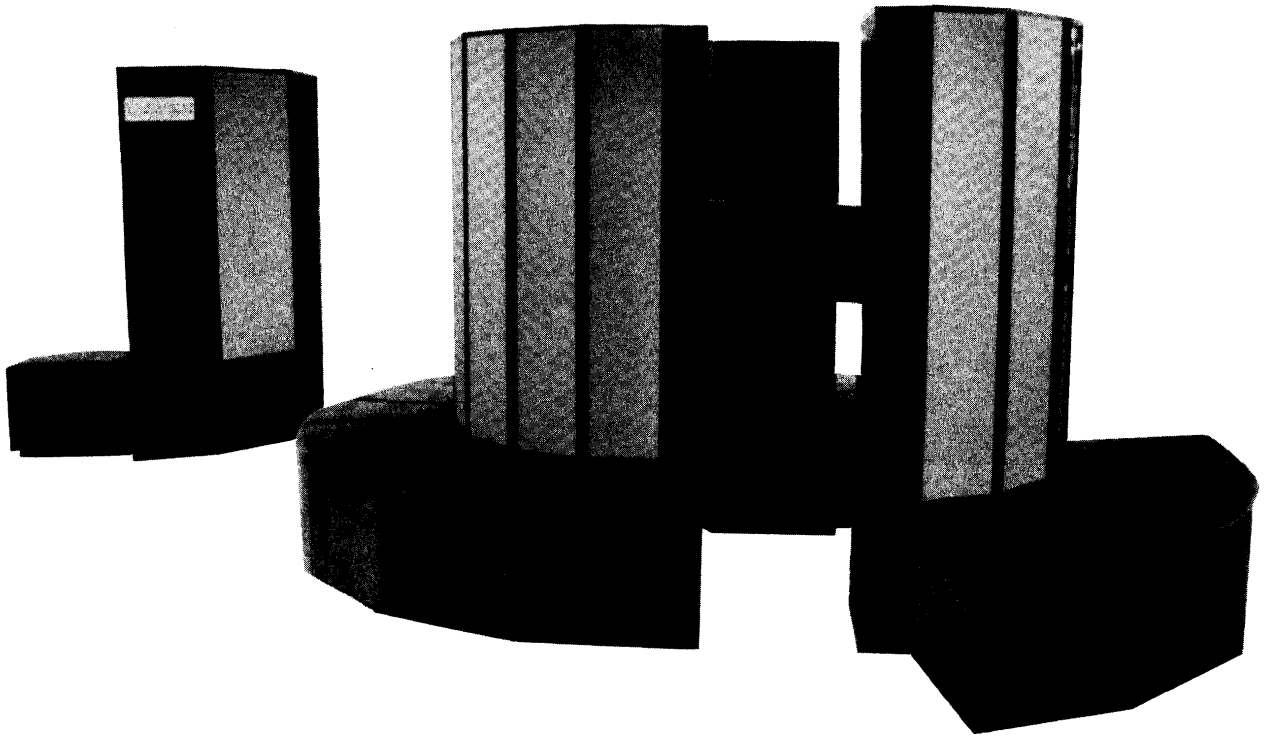


図 1-1. CRAY X-MP モデル 22 か 24 CRAY I/O サブシステムと SSD を持った 12 カラムのメインフレーム
(pp.1-2)

--

表 1-1. CRAY X-MP デュアル・プロセッサ・システムの特徴

<p>Configuration</p>	<ul style="list-style-type: none"> - Mainframe with 2 Central Processing Units (CPUs) - I/O Subsystem with 2, 3, or 4 I/O Processors - Optional Solid-state Storage Device (SSD)
<p>CPU speed</p>	<ul style="list-style-type: none"> - 9.5 ns CPU clock period - 105 million floating-point additions per second per CPU - 105 million floating-point multiplications per second per CPU - 105 million half-precision floating-point divisions per second per CPU - 33 million full-precision floating-point divisions per second per CPU - Simultaneous floating-point addition, multiplication, and reciprocal approximation within each CPU
<p>Memories</p>	<ul style="list-style-type: none"> - Mainframe has 2 million (model 22) or 4 million (model 24) 64-bit words in Central Memory
<p>Input/Output</p>	<ul style="list-style-type: none"> - One 1250 Mbyte per second Solid-state Storage Device (SSD) channel pair - Two 100 Mbyte per second channel pairs for interface to I/O Subsystem - Four 6 Mbyte per second channel pairs
<p>Physical</p>	<ul style="list-style-type: none"> - 64 sq ft floor space for 12-column mainframe; 32 sq ft floor space for 6-column mainframe. - 15 sq ft floor space for I/O Subsystem - 15 sq ft floor space for SSD - 5.25 tons, 12-column mainframe weight; 2.95 tons, 6-column mainframe weight. - 1.5 tons, I/O Subsystem weight - 1.5 tons, SSD weight - Liquid refrigeration of each chassis - 400 Hz power from motor-generators

o 慣例

以下の慣例が、このマニュアルでは使用される。

・イタリック文字(斜体字)

斜体の小文字, **jk** のような、は、変数(?変化する)情報を示す。

・レジスタ慣例

カッコの付いたレジスタ名は、このマニュアルではしばしば、「レジスタ ---の内容」という表現の省略した形式として使用される。例えば、「**(P)**への分岐」は、「レジスタ **P** の内容によって指されるアドレスへ分岐する」を意味する。

A,B,S,T,V レジスタの呼称は、広範に使用される。例えば、「**(Tjk)**から **Si** への転送」は、「指定子 **jk** で指定される **T** レジスタの内容を **i** 指定子で指定される **S** レジスタへ転送する」と意味である。

レジスタのビットには右から左へ、**2** のべきで番号が付いている。

S,V,T レジスタのビット 2^{63} は最上位ビットを表現している。**A** か **B** レジスタのビット 2^{23} は最上位ビットを表現している。**(A,B)** レジスタは **24** ビットである) 交換パッケージとベクトル・マスク・レジスタの番号付け慣例は例外である。交換パッケージは左から右へ番号付けられ、**2** のべきではない、しかし、ビット **0** が最上位ビットでビット **63** が最下位ビットであるというビット **0** から **63** である。ベクトル・マスク・レジスタは **64** ビットを持ち、各ビットは、ベクトル・レジスタのワード要素に対応している。ビット 2^{63} は要素 **0** に対応し、 bit^0 は要素 **63** に対応している。

・数の慣例

明示されているときを除き、このマニュアルでは数は十進数である。八進数は **8** サブスクリプト(添え字)で示す。例外として、レジスタ番号、チャンネル番号、命令バッファ中の命令パーセル、命令フォームは、サブスクリプト無し八進数で与えられる。

・クロック周期

CPU 計算時間の基本単位は **9.5ns**(ナノ秒)でクロック周期(**CP**(Clock period))として参照される。命令発行、メモリ参照、他のタイミングの考察は、通常 **CP** で測る。

o システム構成物

システムは、メインフレーム、I/O サブシステムから構成されている。大容量記憶デバイス、フロントエンド・インターフェース、オプションのテープ・デバイスはシステムの統合部品である。オプションで、Cray ソリッド・ステート・記憶デバイス(SSD)をシステムの部分とできる。

この装置をサポートするのは、冷却のためのコンデンシング・ユニット(凝縮器:熱媒体凝縮用の熱交換器)、システムに電源を供給するためのモータ発電機、メインフレームやI/O サブシステム、SSD のための配電ユニットである。システム構成物は以下に記述される。

- 中央処理装置(Central Processing Unit)

各 CPU は独立した制御部と計算部を持つ。

両 CPU で中央メモリと CPU 間通信と I/O 部を共有する。(CPU 部は後で説明する。)

図 1-2 は計算機の基本構成を図解する。図 1-3 は、システム中の一つの CPU の構成物と、制御とデータのパスを示す。図 1-4 は、メインフレームのシャーシを示す。

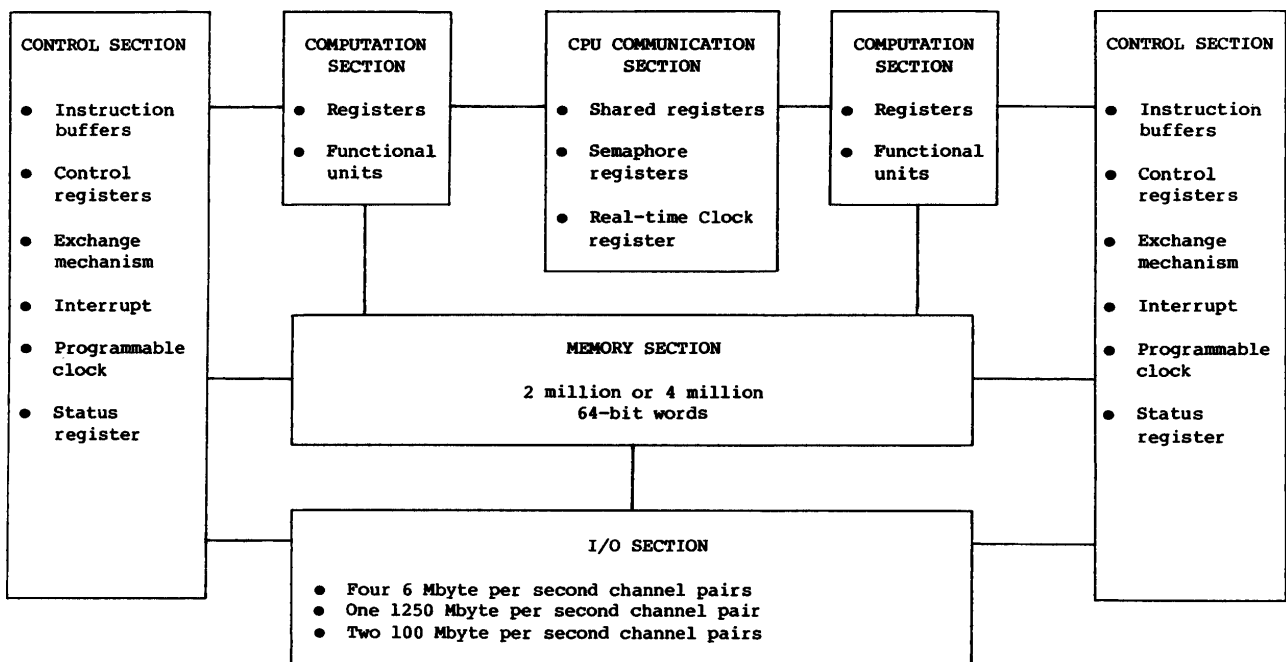
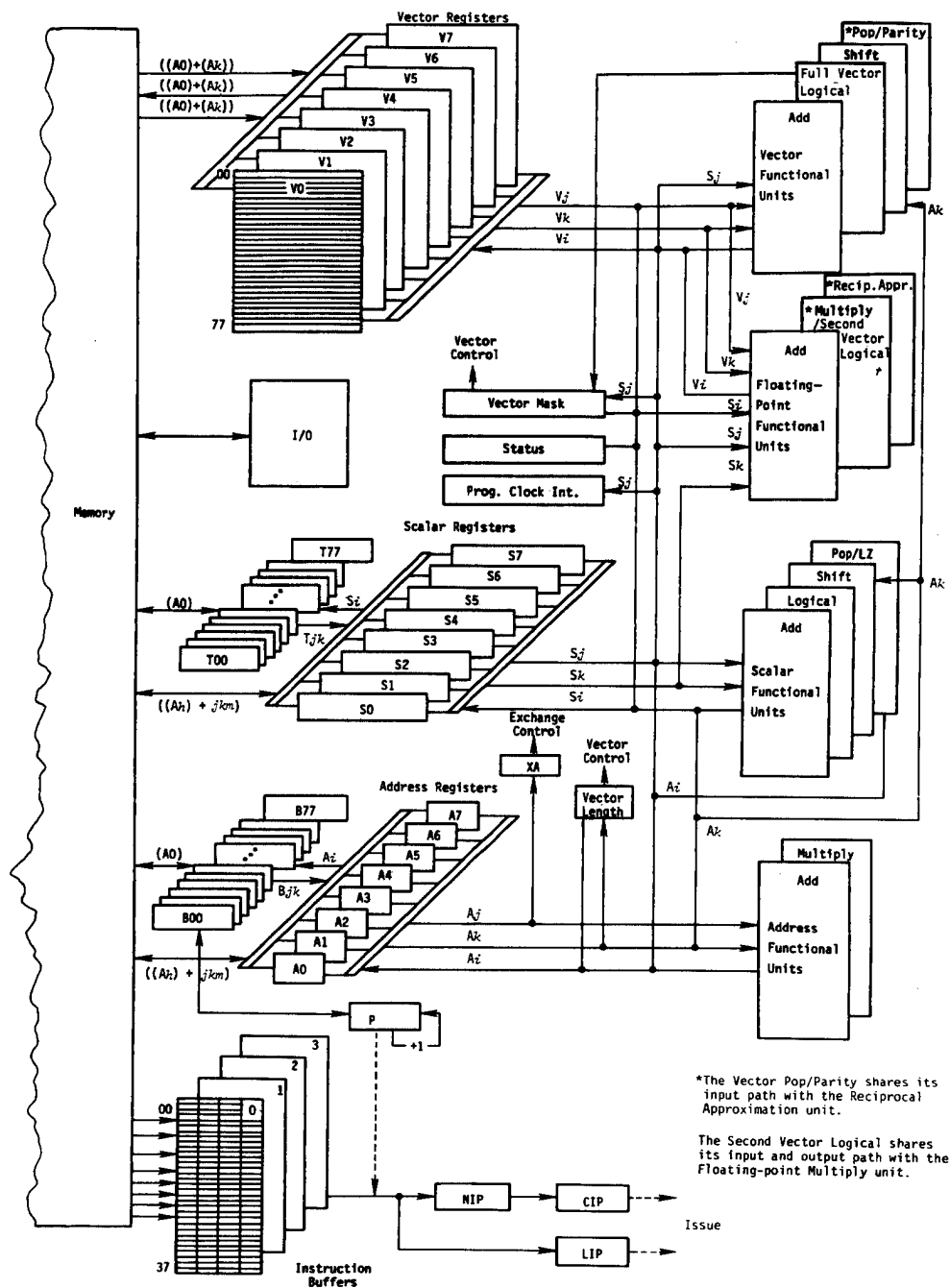


図 1-2.デュアル・プロセッサ・システムの基本構成

(pp.1-5)



*The Vector Pop/Parity shares its input path with the Reciprocal Approximation unit.
 The Second Vector Logical shares its input and output path with the Floating-point Multiply unit.

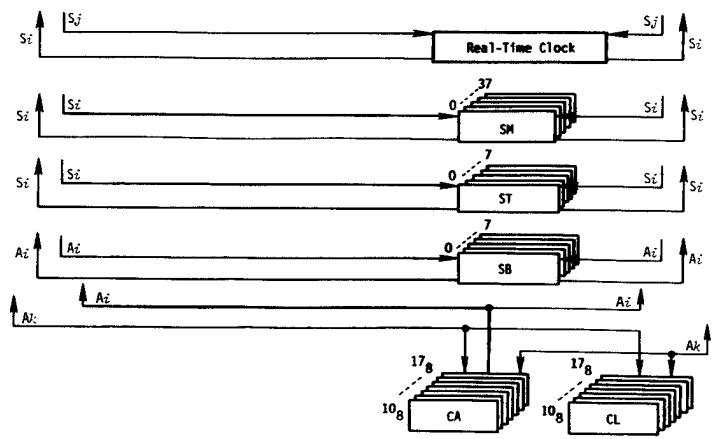


図 1-3.単一 CPU の制御とデータのパス
 ※ 第 2 ベクトル論理ユニットは、どの機械にも存在しない (pp.1-6)



図 1-4. Cray X-MP モデル 22, 24 の 6 カラム・メインフレームのシャーシ
(pp.1-7)

o インターフェース

Cray メインフレームは計算機ネットワーク中のフロントエンド計算機とともに使用するよう設計されている。フロントエンド計算機システムは自己完結で、それ自身のオペレーティング・システムで実行を行う。

標準インターフェースは **Cray** メインフレームの **I/O** チャンネルとフロントエンド計算機のチャンネルを結ぶ。フロントエンド計算機は、**Cray** に入力データを供給し、周辺装置に出力を配るためにそれを受け取る。

インターフェースはチャンネルの幅、機械のワード長、論理の電氣的レベル、制御信号の違いを補正する。

I/O サブシステムのマスタ **I/O** プロセッサは、フロントエンド計算機システムと、通信する。それは **6MBytes/秒** のチャンネル・ペアを通して **Cray** メインフレーム内のチャンネル・アダプタ・モジュールへ。通信は、フロントエンド・インターフェースを通して、通常は、フロントエンド計算機 **I/O** チャンネルを通して、フロントエンド計算機へ続く。

フロントエンド・インターフェースは独立したキャビネット(図 1-5)に収納され、ホスト計算機の近くに置かれている。その操作はフロントエンド計算機のユーザにも **Cray** のユーザにも見えない。

インターフェースの第 1 のゴールは **Cray** システムに接続されたフロントエンド・チャンネルの使用を最大限にすることである。

インターフェースに接続された **MIOP** チャンネルは、インターフェースに接続されたどのフロントエンド・チャンネルよりも速いので、インターフェースのバースト転送レートは、フロントエンド・チャンネルの最大レートによって制限される。

フロントエンド計算機へのインターフェースは、フロントエンド計算機が **Cray** メインフレームをサービスするために、次のことができるようにする:

- マスター操作ステーション
- ローカル操作ステーション
- ローカル・バッチ入力ステーション
- 一つの **Cray** チャンネルへいくつかのステーションをマルチプレックス(複合化)するためのデータ・コンセントレータ(データ集約機)
- リモート・バッチ入力ステーション
- インタラクティブ通信ステーション

フロントエンド計算機に取り付けられる周辺機器は **Cray** システムの使用に依存し、多様である。



写真
図 1-5. 標準的なインターフェース・キャビネット
(pp.1-8)

--

oI/O サブシステム

I/O サブシステム、図 1-6、は Cray X-MP 計算機システムのすべてのモデルで標準であり、それは 2 か 3 か 4 の I/O プロセッサ(IOP)、バッファ・メモリ、必要なインターフェースを持つ。フロントエンド計算機、周辺デバイス、記憶デバイスと I/O サブシステムのバッファ・メモリ間、または、バッファ・メモリと Cray メインフレームの中央メモリ間の、高速なデータ転送のために I/O サブシステムは設計された。

I/O プロセッサの 4 つの型が I/O サブシステムにコンフィギュアされるだろう:

マスタ IOP(MIOP)、バッファ IOP(BIOP)、ディスク IOP(DIOP)、補助(Auxiliary)IOP(XIOP)。すべての I/O サブシステムは、最低限、1 つの MIOP と 1 つの BIOP を持たねばならない。DIOP と XIOP の数はサイトによる。

I/O サブシステムの各 IOP はメモリ部、制御部、計算部、I/O 部を持つ。I/O 部は独立しており、サブシステムへの I/O 要求のいくつかの部分を取り扱う。各 IOP は、そのローカル・メモリへの DMA(direct memory access) ポートを 6 つ持っている。

マスタ I/O プロセッサ(MIOP) はフロントエンド・インターフェースと、ステーション(※)周辺機器の標準グループを制御する。周辺拡張器(Peripheral Expander)はステーション周辺機器を MIOP の 1 つの DMA ポートへインターフェースする。MIOP は、バッファ・メモリへ接続し、6MBytes/秒のチャンネル・ペアを超えてメインフレームへ接続する。MIOP は、I/O サブシステム全体のアクティビティを調整するために、Cray オペレーティング・システム(COS) と通信する。

(※「ステーション」という言葉は、ハードウェアとソフトウェアの両方を意味する。ステーションはフロントエンドへのリンクか、または(MIOP に対して)制限されたフロントエンドとして振舞う。)

バッファ I/O プロセッサ(BIOP)は、メインフレームの中央メモリと大容量記憶デバイスとの間のメイン・リンクである。大容量記憶からのデータは BIOP のローカル・メモリを経由し、メインフレームの中央メモリに 100MByte/秒のチャンネル・ペアを通して転送される。

ディスク I/O プロセッサ(DIOP)は付加ディスク記憶ユニットのために使用される。このプロセッサは、最大 16 までのディスク記憶ユニットが付けられるディスク・コントローラ・ユニットを 4 つまで取り扱うことができる。

DIOP は各コントローラのために 1 つの DMA ポートを使用する。1 つの DMA ポートはバッファ・メモリに接続し、他方の DMA ポートはメインフレームの中央メモリへの 100MByte/秒のチャンネルペアへ接続される。

補助 I/O プロセッサ(XIOP)は、ブロック・マルチプレクサ(多重化)・チャンネルのために使用され、最大 4 つの BMC-4 ブロック・マルチプレクサ・コントローラをインタフェースする。

各コントローラは 4 つまでのブロック・マルチプレクサ(多重化)チャンネルを取り扱える。

XIOP は、各コントローラのために 1 つの DMA ポートを、バッファ・メモリへの接続のために他の 1 つの DMA ポートを使用する。

I/O サブシステムのハードウェアは、BIOP と DIOP 間、I/O サブシステムの XIOP とメインフレームの中央メモリ間の同時のデータ転送ができる。(※)

(※:XIOP への 100MByte/秒のチャンネル・ペアをサポートするソフトウェアは、現在、無い)

Cray デュアル・プロセッサ・システムの CPU 入出力部はこのマニュアルの 2 章で説明する。
I/O サブシステムの完全な説明については、「I/O サブシステム・リファレンス・マニュアル」、CRI 出版
HR-0030、を参照のこと。

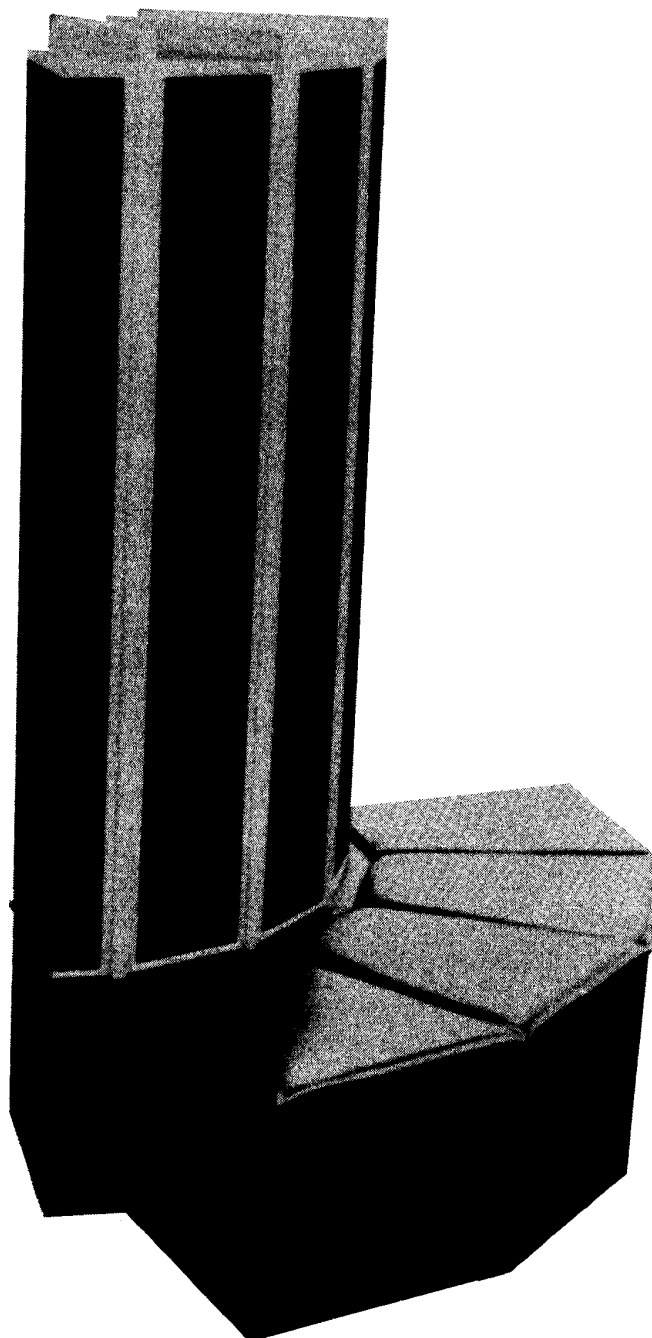


図 1-6. I/O サブシステムのシャーシ
(pp.1-10)

o ディスク記憶ユニット

大容量の記憶のために、システムは **Cray** リサーチ製のディスク記憶ユニット(DSU)を使用している。

ディスク・コントローラ・ユニット(DCU) は、ディスク記憶ユニットを、1 つの DMA ポートを通して、I/O サブシステムの I/O プロセッサとインターフェースする。4 つまでのディスク記憶ユニットが一つの DCU に接続できる。

I/O プロセッサとディスク・コントローラ・ユニットは、すべての DSU がデータを失わず、回転を飛ばすことなくフルスピードで動作しながら、DMA ポートと 4 つの DSU の間でデータを転送することができる。最少 2 つ、最大 48 の DSU が 1 つの I/O サブシステムにコンフィギュアすることができる。図 1-7 に **Cray DD-29** ディスク記憶ユニットを示す。ディスク・コントローラ・ユニットは I/O サブシステムのシャーシに収まる。

各 DSU は、コントローラに接続するために 2 つのアクセスを持つ。2 番目の独立データ・パスは他の **Cray** リサーチ製コントローラを通して、各 DSU へのために存在する。予約論理は各 DSU への制御されたアクセスを提供する。デバイスの動的な共有は、**Cray** オペレーティングシステム(COS)ソフトウェアではサポートされていない。大容量記憶サブシステムのより詳しい情報は、I/O サブシステム・リファレンス・マニュアル、CRI 出版 HR-0030、と大容量サブシステム・ハードウェア・リファレンス・マニュアル、CRI 出版 HR-0630 に含まれている。



図 1-7. DD-29 ディスク記憶ユニット
(pp.1-11)

o 半導体記憶デバイス

図 1-8 に示す、半導体記憶デバイス(SSD)はオプションであり、一時データ記憶のために使用される高性能デバイスである。それは、特殊な Cray インターフェース・ケーブルの組を通して、最大速度 1250MByte/秒で、データをメインフレームの中央メモリと SSD の間を転送する。

SSD リファレンス・マニュアル、CRI 出版 HR-0031、に書かれているとおり、実際のこれらの転送速度は、SSD メモリの大きさと、システムのコンフィギュレーションに依存する。

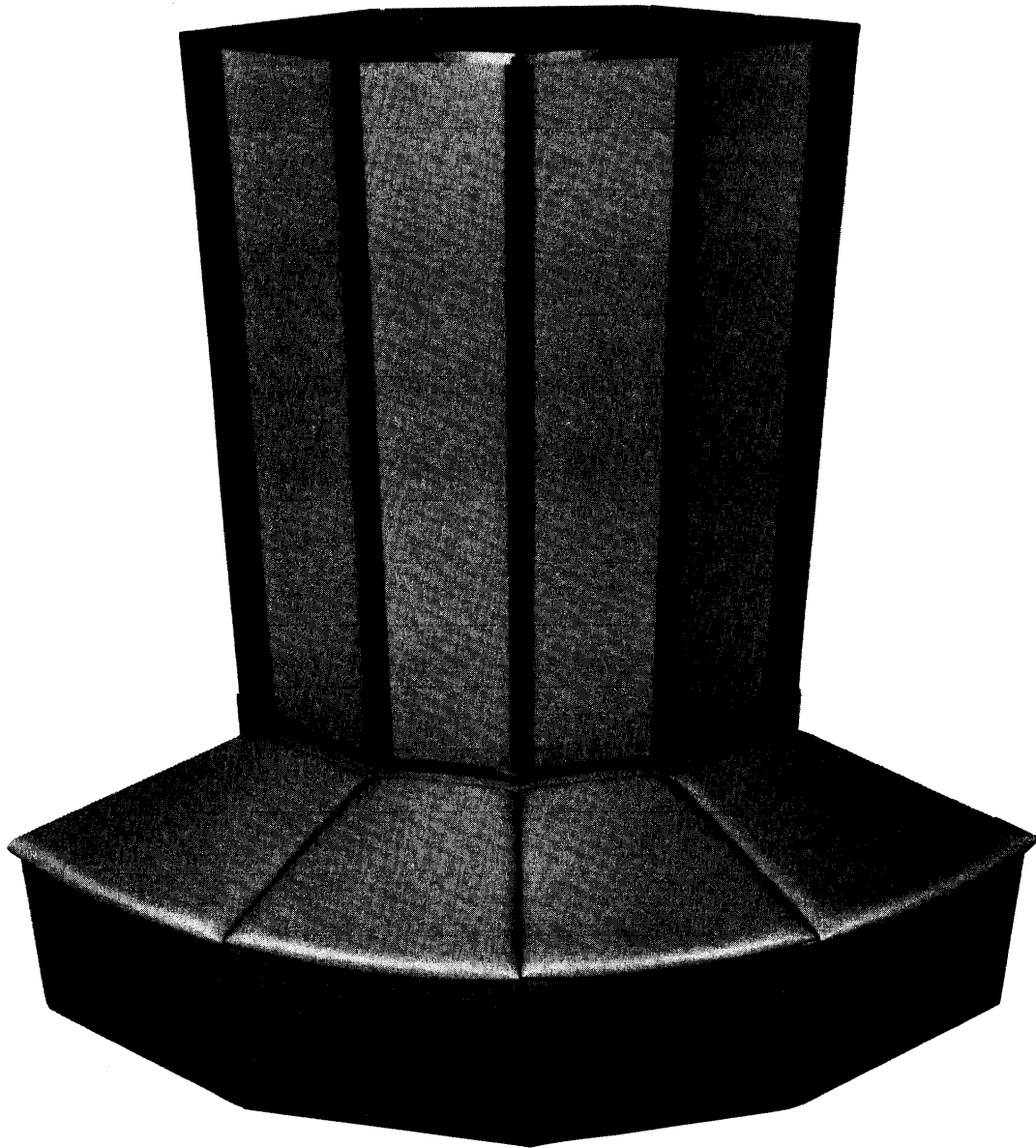


図 1-8. 半導体記憶デバイス・シャーシ
(pp.1-12)

0 コンデンシング・ユニット

コンデンシング・ユニット(図 1-9)(凝縮器:熱媒体凝縮用の熱交換器)は、計算機のシャーシを冷却するために使用され、25トンの2つの凝縮器からなる、冷却システムの主要な部分である。
熱は、計算機システムの部分ではない、第2レベル冷却システムによってコンデンシング・ユニットから取り除かれる。計算機を冷やすフロンが、熱を拾い上げ、コンデンシング・ユニット中の水まで運ぶ。

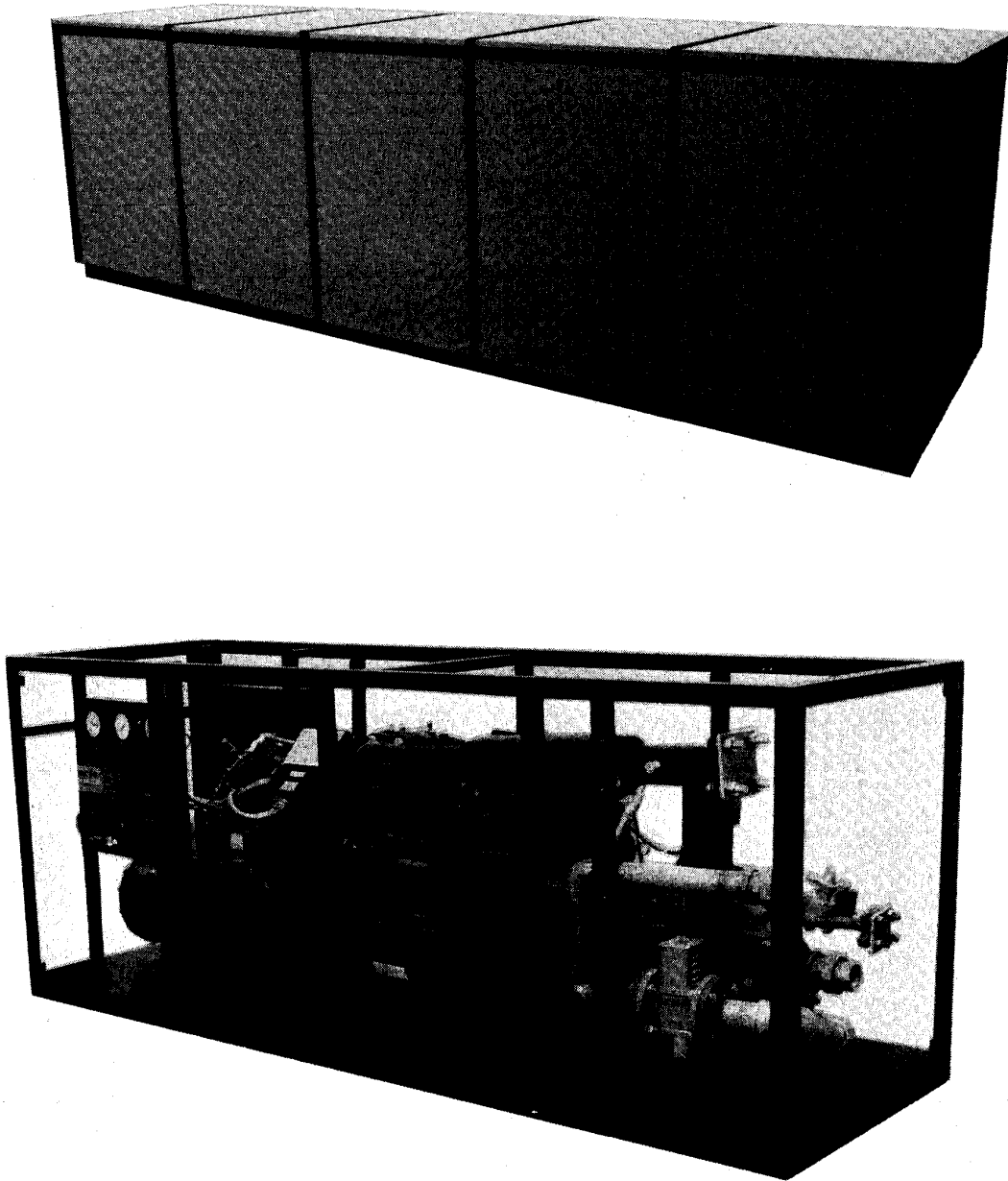


図 1-9. コンデンシング・ユニット
(pp.1-13)

o 配電ユニット

Cray メインフレーム、I/O サブシステム、SSD はすべて 400Hz の三相交流で動作している。Cray メインフレーム、I/O サブシステム、SSD は独立した配電ユニットを持っている。メインフレームのための配電ユニットは、メインフレームの各棚に電圧を調整するための設定可能なトランスを持っている。配電ユニットは、温度と電圧をモニタする装置を持ち、メインフレーム・シャーシの戦略的位置の温度をチェックする。自動警告とシャットダウン回路が、オーバーヒートや過冷却からメインフレームを保護する。モータ発電機とコンデンシング・ユニットのための制御スイッチはメインフレームの配電ユニットに取り付けられている。

小さな配電ユニットは I/O サブシステム・シャーシや SSD シャーシと同様の機能を果たす。

図 1-10、左はメインフレームのための、右は I/O サブシステムか SSD のための、配電ユニットである。



図 1-10. 配電ユニット
(pp.1-14)

o モータ発電ユニット

モータ発電機は主電源を、商用電力線から、システムで使用される400Hz電力に変換する。これらのユニットは商用電力線の瞬時や長期の電圧変動からシステムを隔離する。装置は2つか3つのモータ発電ユニットと制御キャビネットから成っている。図1-11が典型的なモータ発電機とその制御キャビネットである。

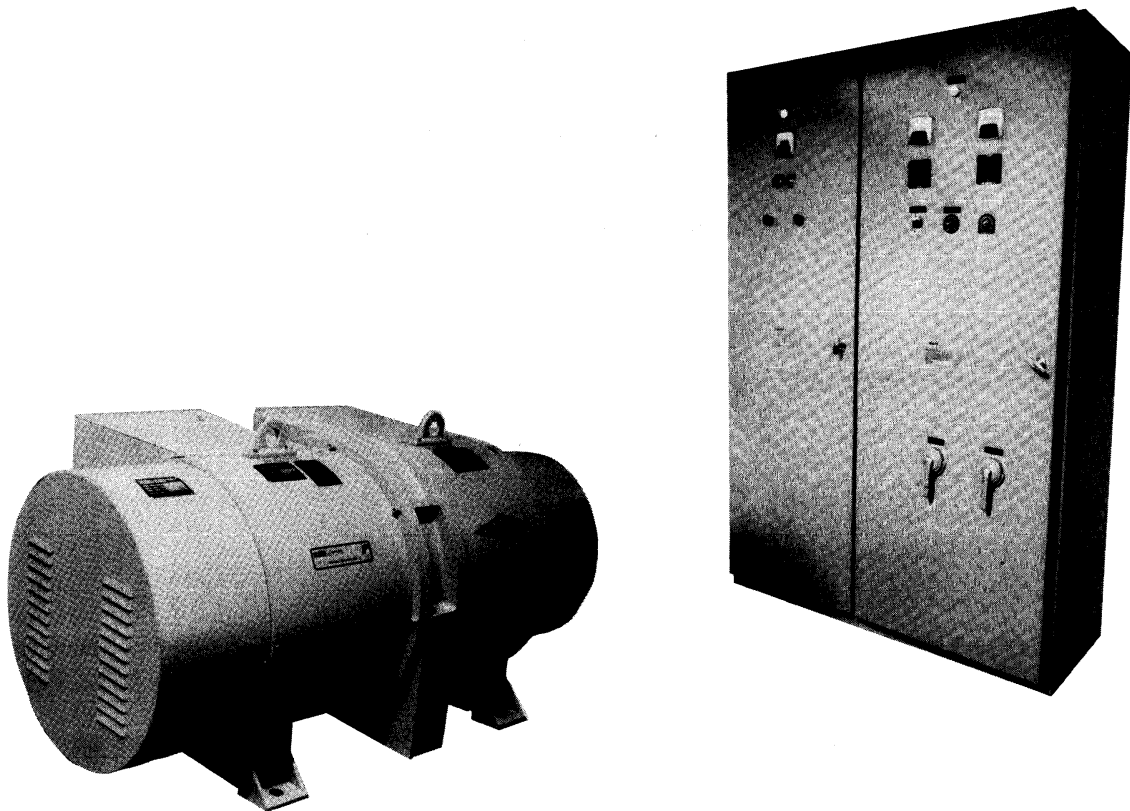


図1-11. モータ発電装置
(pp.1-15)

0 システム・コンフィギュレーション

図 1-12 と 1-13 は Cray X-MP 計算機システムのモデル 22 や 24 のための 2 つのコンフィギュレーションを説明している。

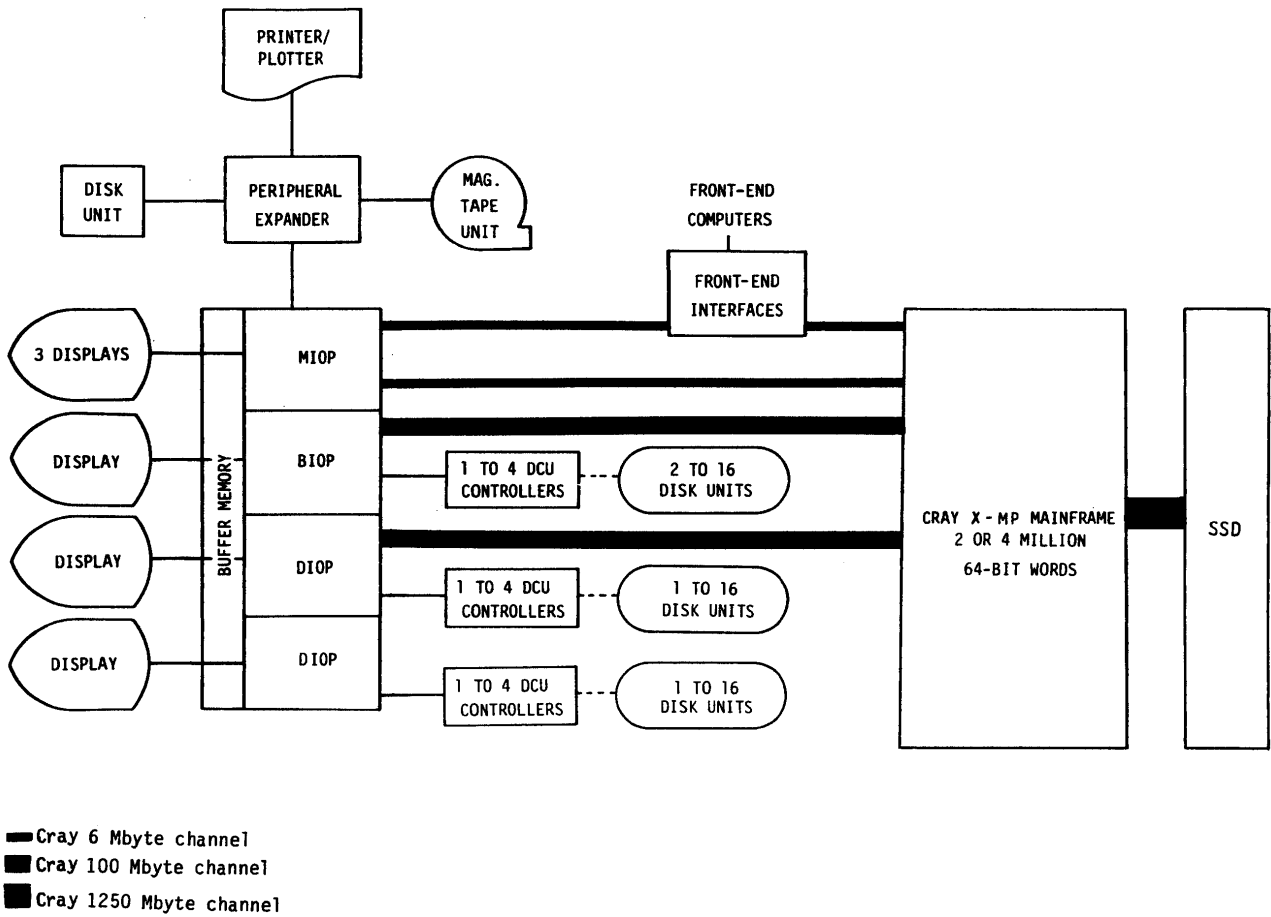
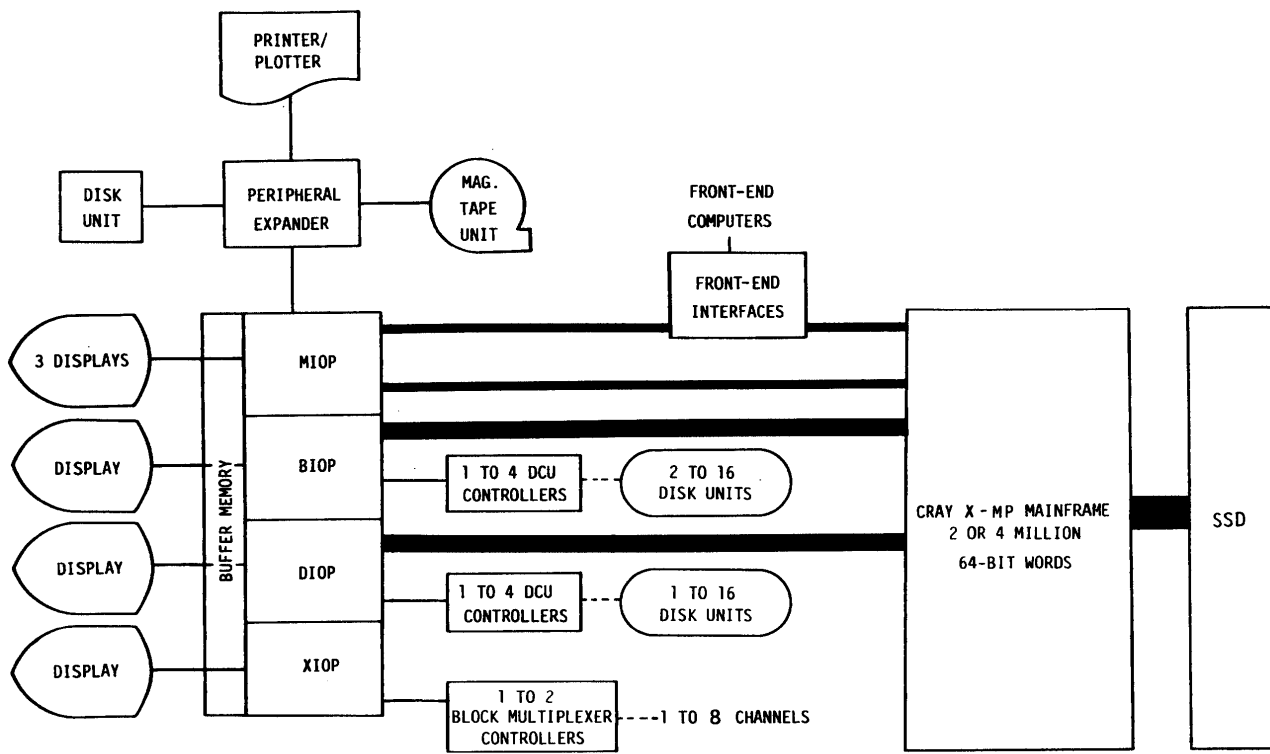


図 1-12. Cray X-MP デュアル・プロセッサ・システム、フル・ディスク容量のブロック図 (pp.1-16)

--



- Cray 6 Mbyte channel
- ▬ Cray 100 Mbyte channel
- Cray 1250 Mbyte channel

図 1-13. Cray X-MP デュアル・プロセッサ・システム、ブロック多重化チャンネル付きのブロック図 (pp.1-17)

--