

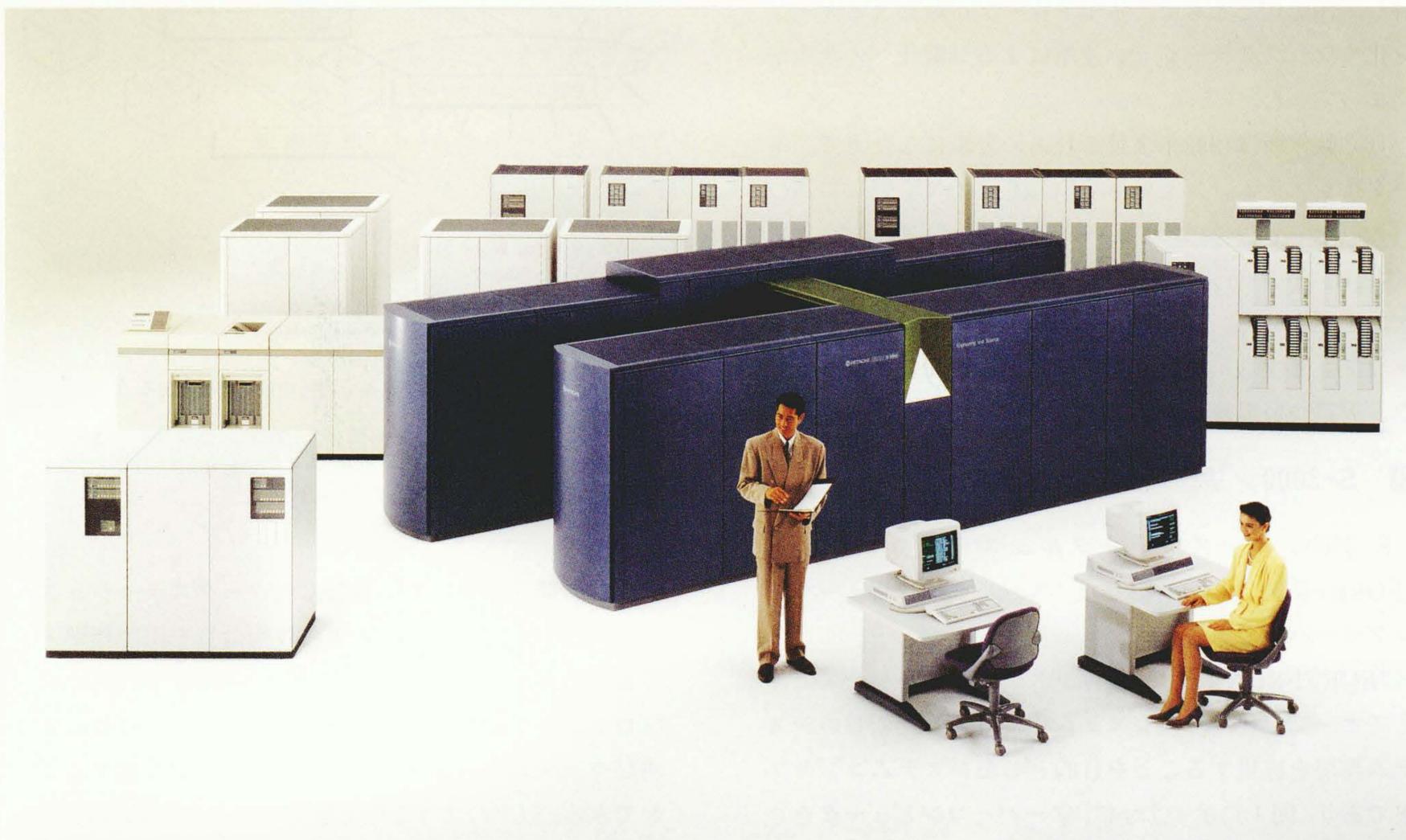
オープン化・分散化に対応した 超高速スーパーコンピュータ

Hitachi S-3000 Series Supercomputer, Its Operating Systems and Hardware

村山 浩* *Hiroshi Murayama*

福多謙治** *Kenji Fukuta*

山田光子** *Mitsuko Yamada*



スーパーコンピュータHITAC S-3000シリーズ S-3800モデル280
ing Point Operations Per Second)に達し、科学技術計算に威力を発揮する。

2台のマルチプロセッサで最大ベクトル性能が16 G FLOPS (Float-

気象予測や素粒子の解明など自然現象の探究から、大規模構造解析や流体解析による新技術・新製品の開発に至るまで、利用範囲がますます広がるスーパーコンピュータに対し、高性能化に加えてオープン化、分散コンピューティング環境への対応が強く求められている。

日立製作所はこれらの要求にこたえて、スーパーコンピュータHITAC S-3000シリーズを開発した。最大4台のマルチプロセッサ構成により、世界最高

速レベルの最大ベクトル性能32 G FLOPS (Floating Point Operations Per Second)を達成した。また、国際標準・規格に準拠したUNIXシステム^{*1)}HI-OSF/1^{*2)}-MJ (Hitachi OSF/1 Operating System)を採用し、業界標準通信プロトコルTCP/IP^{*3)}や通信媒体、UltraNet^{*4)}、HYPERchannel^{*5)}などをサポートし、マルチベンダ環境での柔軟なシステム構成を可能にした。

* 日立製作所 汎用事業部 ** 日立製作所 ソフトウェア開発本部

1 はじめに

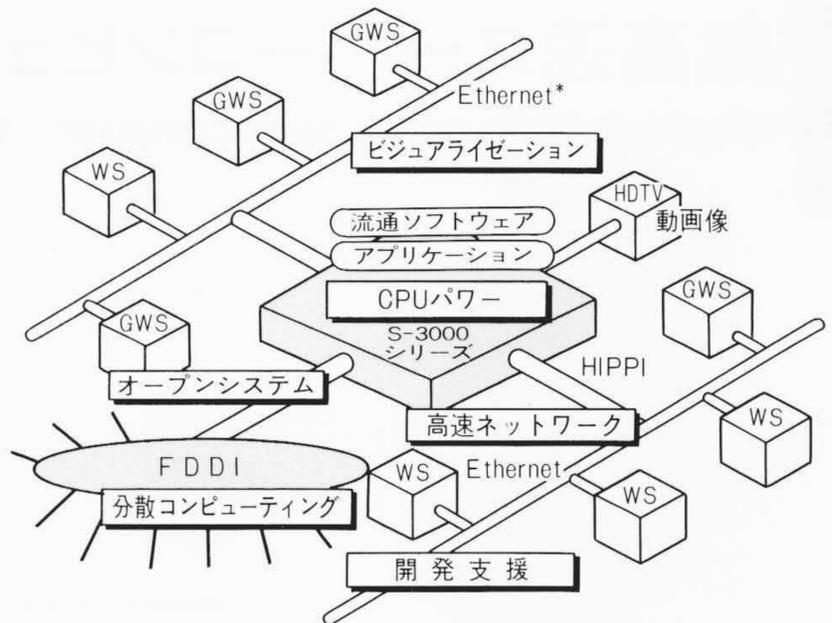
1983年10月に日立製作所が出荷した国産初のスーパーコンピュータや、第一世代とも言うべきその他の国産スーパーコンピュータは、1台のプロセッサ構成で、オペレーティングシステムが各社固有であった。しかし、近年スーパーコンピュータを利用した情報処理システムに対し高性能化とともに、ネットワーク接続によるオープン化やワークステーション連携による分散化への要求が高まってきた。

日立製作所は1992年3月これらの要求にこたえて、トータルシステムコンセプト“FOREFRONT^{※6)}”に基づいた「スーパーコンピューティングサーバ」であるHITAC S-3000シリーズ(以下、S-3000シリーズと略す。)を発表した。ここではS-3000シリーズのシステムコンセプトとオペレーティングシステム機能およびハードウェアについて述べる。

2 S-3000シリーズシステムコンセプト

S-3000シリーズはトータルシステムコンセプト“FOREFRONT”に対応した「スーパーコンピューティングサーバ」である。

“FOREFRONT”は、現行システムのエンハンスに加えてオープン化、分散システム化に対応した新しいシステム環境を提供することを目的としたシステムコンセプトであり、図1に示すように、スーパーコンピュータをコンピュータネットワーク中の一つのノードとして組み込み、高速ネットワークを介した超高速の計算サーバとして位置づけるものである。このようなネットワーク環境で、応答性が必要なプログラムのデバッグやチューニン



注：略語説明

GWS (Graphic Work Station)

WS (Work Station)

HDTV (High Definition Television)

HIPPI (High Performance Parallel Interface)

FDDI (Fiber Distributed Data Interface)

*(Ethernetは富士ゼロックス株式会社の登録商標である。)

図1 スーパーコンピューティングコンセプト S-3000シリーズはネットワーク中の一つのノードであり、スーパーコンピューティングサーバとして位置づけられる。

グなどの開発作業は慣れ親しんだワークステーションで行い、CPUパワーを必要とする技術計算を国際標準・業界標準のAPI(Application Programming Interface)やプロトコルで超高速計算サーバを利用する。計算結果は再びワークステーションで手軽にビジュアル化することができる。以上のような環境を提供することが、スーパーコンピューティングサーバとしてのコンセプトである。

S-3000シリーズはマルチプロセッサ構成などのハードウェアアーキテクチャの拡張、最新鋭の半導体・実装技術、ANSI(American National Standards)規格HIPPI(High Performance Parallel Interface)に準拠した高速入出力(HIPPI)機構などによって大規模計算の高速処理を実現した。また、ハイビジョン方式の動画像出力などによる可視化機能やUNIXシステムおよびオープンネットワークなどのオープン対応、プログラム開発支援などのワークステーションへの分散化によって上記コンセプトを実現し、スーパーコンピュータの適用範囲と使用形態を拡大した。

3 オペレーティングシステム

S-3000シリーズをサポートするオペレーティングシステムは、国際標準・国際規格に準拠したUNIXオペレーティングシステムHI-OSF/1-MJ(Hitachi OSF/1

※1) UNIXシステム：UNIXオペレーティングシステムは、UNIXシステムラボラトリーズ社が開発し、ライセンスしている。

※2) OSF/1：Open Software Foundation, INC.の商標である。

※3) TCP/IP：米国国防総省が標準化したプロトコルである。

※4) UltraNet：米国ウルトラネットワークテクノロジー社の製品名称である。

※5) HYPERchannel：住商エレクトロニクス株式会社の製品名称である。

※6) FOREFRONT：日立製作所が1992年3月、情報処理環境の変化に対応し、エンドユーザーの視点に立った新しい事業姿勢、製品計画、サービス提供、システム構築の考え方を提唱したトータルシステムコンセプトである。詳細は平成4年5月号を参照願う。

Operating System)と、汎(はん)用プロセッサとの高い互換性・親和性を持つVOS3/AS(Virtual Storage Operating System 3/Advanced System Product)の2種類があり、システム形態に応じた選択ができる。

また、仮想計算機システムVMS/AS(Virtual Machine System/Advanced System)により、両オペレーティングシステムを共存させて運用できる。これにより、一つのスーパーコンピュータを二つのオペレーティングシステムで効率よく使用することができる。

3.1 HI-OSF/1-MJシステム

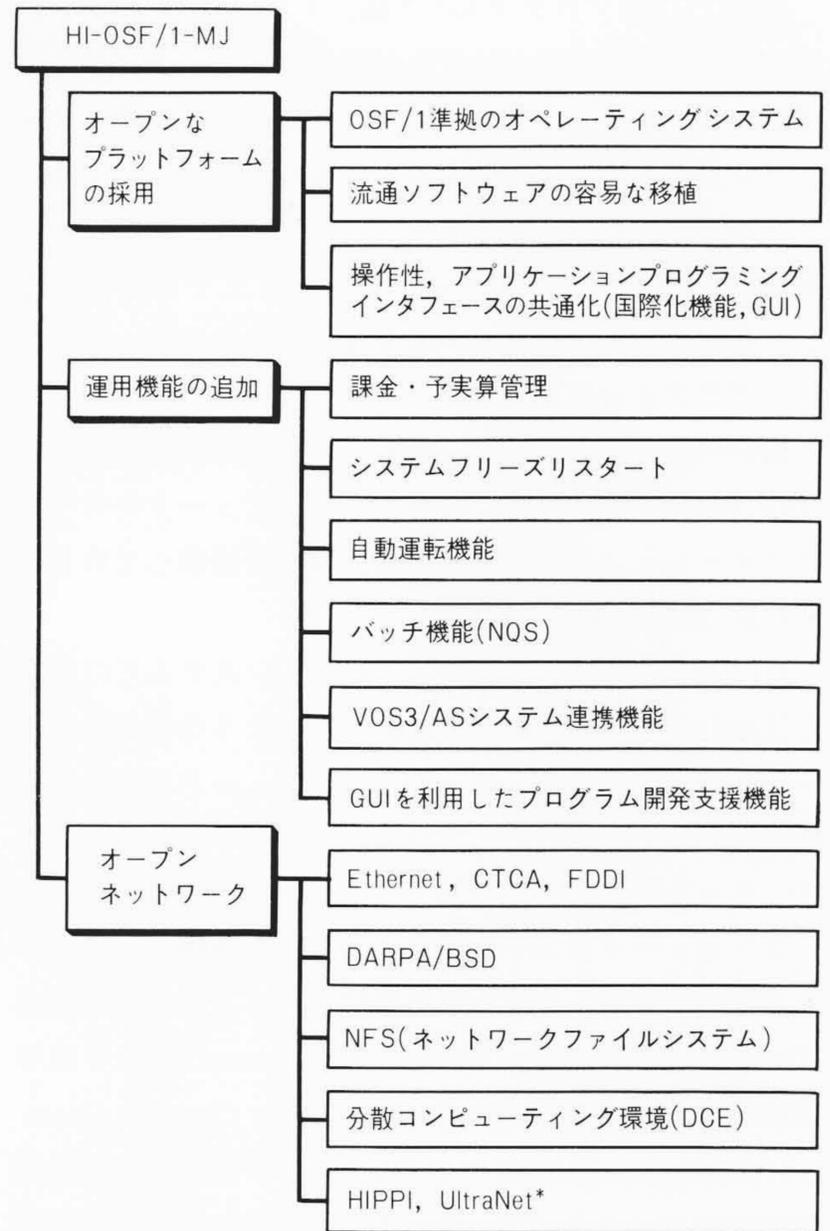
HI-OSF/1-MJの特長を図2に示す。HI-OSF/1-MJシステムは、各種国際標準・規格に準拠した最新UNIXシステムであるOSF/1をベースとし、インタオペラビリティ、ポータビリティ、スケーラビリティに富むオープンなプラットフォームを採用している。また、カーネル部には分散オペレーティングシステム対応のMach^{*7)} 2.5を使用し、ネットワークを核にしたマルチベンダシステムの環境にスーパーコンピューティングサーバとして、S-3000シリーズを取り込んでいる。

3.1.1 HI-OSF/1-MJの機能

HI-OSF/1-MJシステムは、マルチスレッド機能、機密保護機能、分散コンピューティング機能などのOSF/1の共通機能をサポートするとともに、スーパーコンピュータ向けの機能として、次のような機能をサポートしている。

- (1) マルチプロセッサ上で、同一または複数のベクトル処理プログラムを同時実行し、ベクトル処理のターンアラウンド時間の短縮やスループットの向上を図るためのスケジューリング機能。
- (2) 通常の対話処理での主記憶装置割り当て機能と、大規模技術計算プログラムが使用する大容量主記憶装置割り当て機能を効率よく共存させるメモリ制御。
- (3) 高速・大容量な拡張記憶装置をベクトルプロセスのスワッピング用の補助記憶装置や大規模計算ジョブでのデータファイル装置として使用する機能。
- (4) ベクトル化FORTRAN，並列化FORTRANおよびベクトル化Cなどのコンパイラ。
- (5) 膨大な数値シミュレーション結果の時系列変化を、動画像(アニメーション)としてビデオレート(30画面/s)でカラーモニタあるいはVTRに出力させる動画像出力

^{*7)} Mach：カーネギメロン大学で開発されたオペレーティングシステムである。



注：略語説明ほか HI-OSF/1-MJ (Hitachi OSF/1 Operating System)
 API (Application Programming Interface)
 GUI (Graphical User Interface)
 NQS (Network Queuing System)
 CTCA (Channel to Channel Adapter)
 FDDI (Fiber Distributed Data Interface)
 NFS (Network File System)
 DCE (Distributed Computing Environment)
 * (米国ウルトラネットワークテクノロジーズ社の製品名称である。)

図2 HI-OSF/1-MJの特長 S-3000シリーズ用オペレーティングシステムとしてHI-OSF/1-MJを選択することにより、スーパーコンピュータを分散環境のサーバとしてパーソナルコンピュータやワークステーションと同一の操作で利用できる。

制御機能。

- (6) 100 Mバイト/sを超えるデータ転送能力を持つ超高速ネットワークへの接続を可能にするHIPPI機構のサポート。

3.1.2 ワークステーションとの連携機能

プログラムの開発や実行を簡単に効率よく行うために、HI-OSF/1-MJシステムでは、優れたユーザーインタフェースを提供している。

- (1) プログラム開発支援

FORTRANプログラムの作成、コンパイル、リンク、デバッグおよび実行と強力な対話型ベクトルチューニングが、スーパーコンピュータ上だけでなくワークステーション上でも可能にした。これにより、図3に示すようにプログラムデバッグをワークステーション上で行い、実際のジョブをスーパーコンピュータ上で実行するなど、効率のよいプログラム開発環境が得られる。

(2) グラフィックス支援

操作の基本となるGUI(Graphical User Interface)にOSF/Motif^{※8)}を採用し、スーパーコンピュータからワークステーションまで統一した操作環境を提供している。

3.1.3 豊富な運用支援機能

HI-OSF/1-MJシステムは、UNIXシステムでの弱点と言われる運用性を高めるため、次のような機能をサポートすることにより、スーパーコンピュータを含めた大規模なシステムに対応している。

(1) ネットワーク管理機能

コンピュータシステム間での有機的な連携・分散ができるオープンネットワークを実現させるため、国際標準プロトコルOSI(Open System Interconnection)や業界標準プロトコルTCP/IP(Transmission Control Protocol/Internet Protocol)への対応、および国際標準通信媒体ISDN(Integrated Services Digital Network)、FDDI(Fiber Distributed Data Interface)や業界標準通信媒体HYPERchannel、UltraNetへの対応を行ってオープンインタフェースを実現した。

これにより、マルチベンダ環境での柔軟なシステム構築を可能にしている。

また、SNMP(Simple Network Management Protocol)により、ネットワークでの障害監視や運用での統計情報などを収集し、専用のサーバへ報告する機能もサポートしている。

(2) 分散コンピューティング機能

分散環境でのアプリケーションの利用、開発、保守のために統合化されたサービスを提供する分散コンピューティング環境(DCE: Distributed Computing Environment)をサポートしている。

ネットワークを経由した大規模な分散コンピューティング環境を構築でき、世界規模の環境からスーパーコン

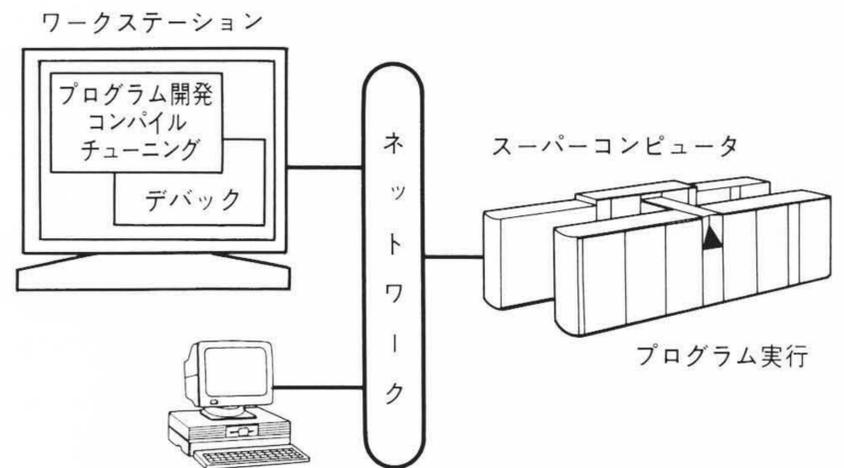


図3 ワークステーションとの連携 スーパーコンピュータで実行させるアプリケーションプログラムのコンパイル、デバックおよびチューニングをネットワーク上のワークステーションで効率よく行える。

ピュータを容易に利用できる。

(3) システム運用機能

システムフリーズリスタート、自動運転などのシステム運用機能により、大型システムでのセンターオペレーション、およびシステム運転の省力化と安全運転を実現している。

(4) バッチ運用機能

UNIXシステムは対話式で処理を行うのが一般的であるが、スーパーコンピュータを使用するプログラムがバッチジョブとしてCPUなどのリソースを効率よく利用して実行するように、NQS(Network Queuing System)と呼ぶバッチジョブ管理機能を備えている。

(5) NFS^{※9)}(Network File System)

ネットワーク内のクライアント側のユーザーが、ネットワークを意識せずクライアント側と同じ操作でサーバのファイルシステムをアクセスできる。

HI-OSF/1-MJシステムは、オープンシステムでありながらも、日立製作所独自の機能を追加し、ユーザーにさらに優れた開発環境を提供している。

3.2 VOS3/AS

VOS3/ASは、豊富な実績のある汎用大型プロセッサ用のオペレーティングシステムであるが、これにS-3000シリーズ用の制御プログラムであるプログラムプロダクトVOS3/HAP/ASを付加し、マルチプロセッサ、並列化FORTRAN、HIPPI機構などをサポートした。

(1) 汎用オペレーティングシステムの豊富な機能

※8) OSF/Motif: Open Software Foundation, INC.の商標である。

※9) NFS: 米国Sun Microsystems, Inc.の商標である。

システムの運用に関しては、汎用大型システムHITAC Mシリーズの持つ豊富な機能をそのまま利用することができる。したがって、汎用システムからの高い移行性、親和性および互換性を持っており、汎用機との併設や汎用機からの移行に適している。

(2) HI-OSF/1-MJとの連携

VOS3/ASとHI-OSF/1-MJは、図4に示すようにEthernet^{※10)}などのLANを介し、ネットワーク間プロトコル(TELNET, FTP)によって通信や転送ができる。また、CTCA(Channel to Channel Adapter)を経由したRPC(Remote Procedure Call)による連携も可能である。さらに、NFSによってHI-OSF/1-MJからVOS3/ASのファイルにアクセスできる。

4 ハードウェア

4.1 ラインアップ

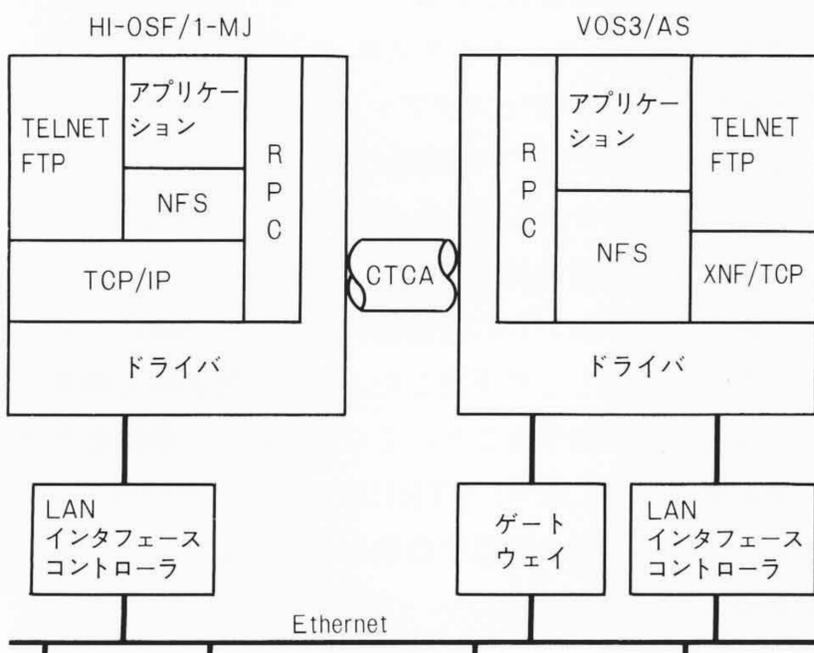
S-3000シリーズは、図5に示すように250 M FLOPSから32 G FLOPSまで幅広いベクトル性能範囲をカバーするS-3800プロセッサグループおよびS-3600プロセッ

サグループ全15モデルで構成する。上位機種種のS-3800は、1台から4台のプロセッサから成りフィールドアップグレードが可能である。モデル480は4台のマルチプロセッサで、世界最高速レベルの32 G FLOPSを達成している。価格性能比の良さを追求するS-3600は、空冷方式のユニプロセッサで、より緩やかな設備条件で使用でき、エントリーモデルから、業務の拡大に応じて容易にフィールドアップグレードが可能である。また、S-3000シリーズは従来のスーパーコンピュータHITAC S-810/S-820シリーズとは上方互換性があり、従来のソフトウェア資産を生かすことができる。

4.2 ハードウェア構成

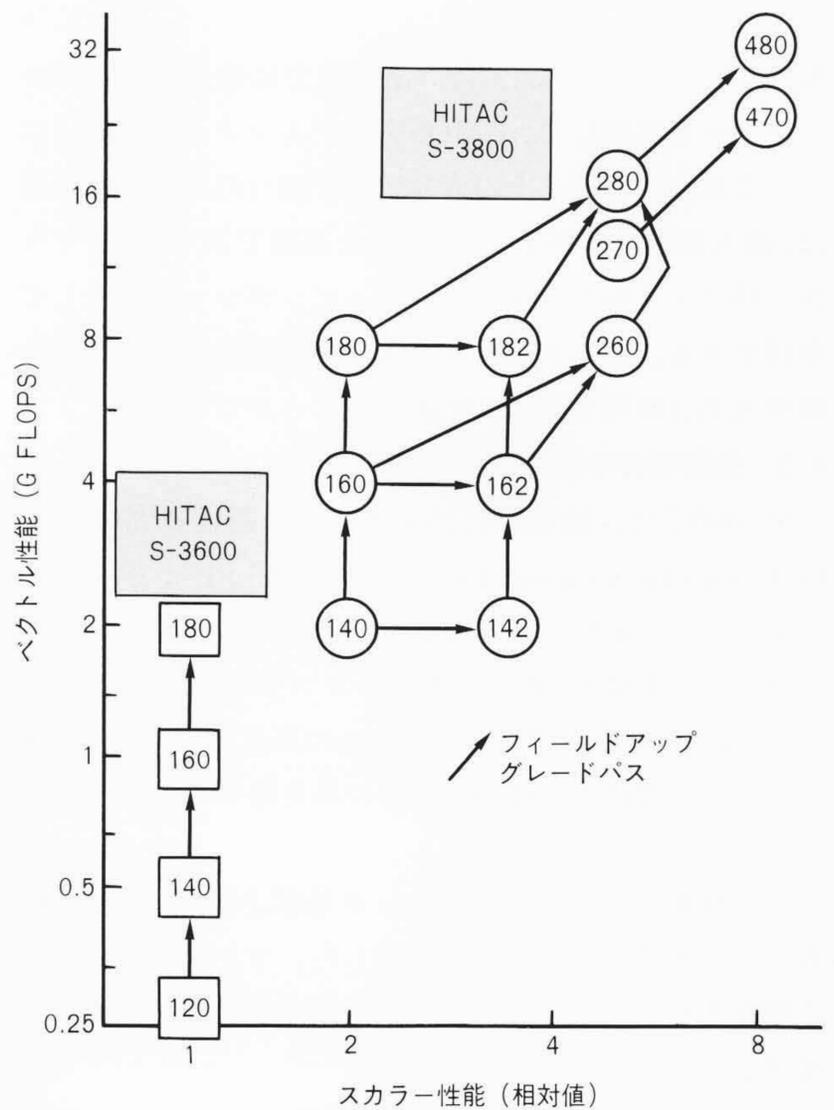
S-3800のハードウェア構成を図6に示す。

スカラー命令を実行するスカラープロセッサ、ベクトル命令を実行するベクトルプロセッサから成る命令プロセッサによって高速演算を実現した。高速・高集積なBI-CMOS(Bipolar-Complementary Metal Oxide Semiconductor)スタティックメモリ素子を使った最大メモリ



注：略語説明
 CTCA (Channel to Channel Adapter)
 RPC (Remote Procedure Call)
 NFS (Network File System)
 TCP/IP (Transmission Control Protocol/Internet Protocol)
 TELNET (リモートログイン機能)
 FTP (File Transfer Protocol)

図4 VOS3とHI-OSF/1-MJの連携 LANやCTCAを利用してVOS3/ASとHI-OSF/1-MJを接続することで、HI-OSF/1-MJ側からVOS3/ASのファイルや運用機能を利用できる。



注：図中囲み内の数字はモデル名を示す。

図5 HITAC S-3000シリーズのモデル S-3000シリーズは、水冷で最大ベクトル性能32 G FLOPSのS-3800と、空冷のS-3600の二つのラインから成る。

※10) Ethernet：富士ゼロックス株式会社の登録商標である。

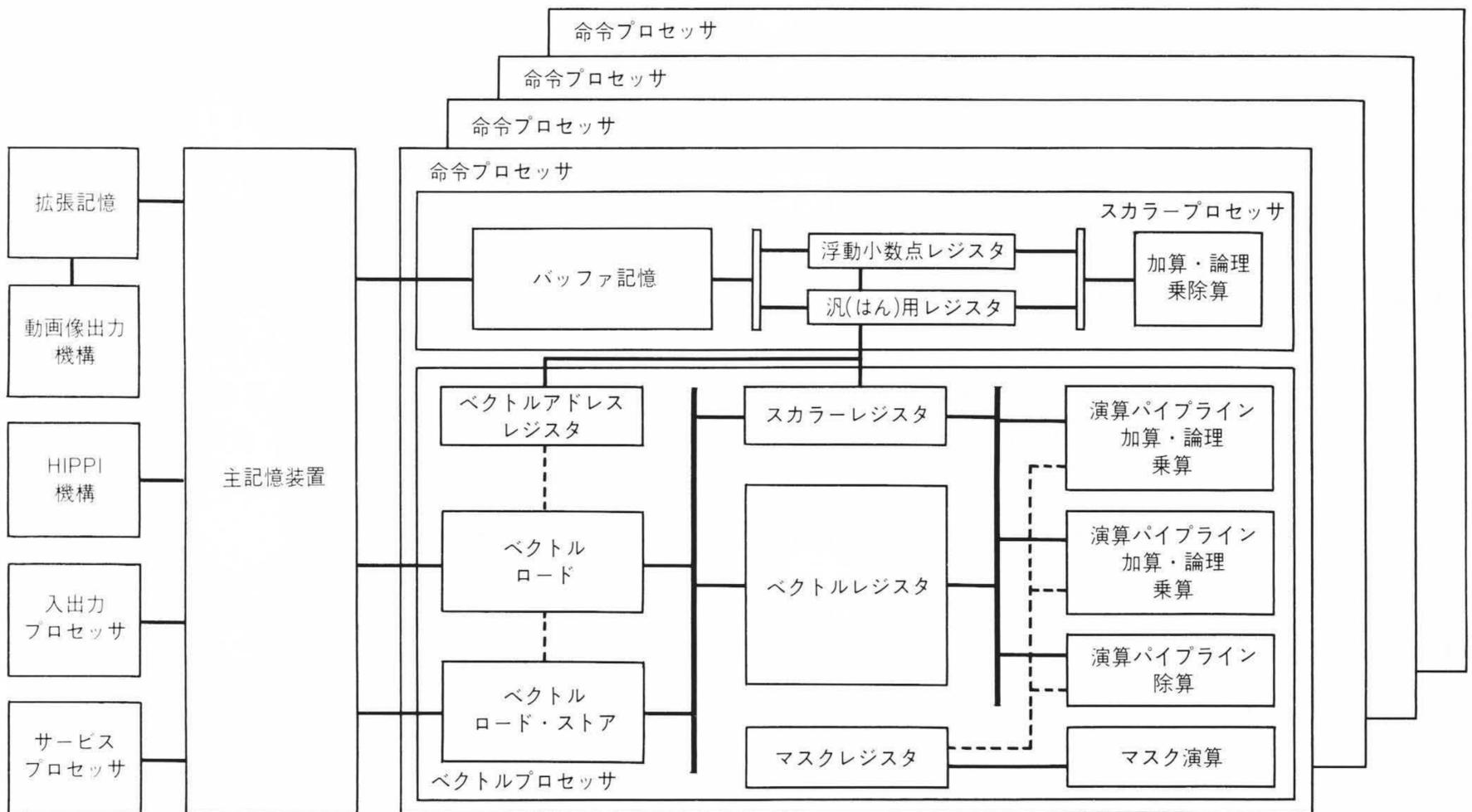


図6 S-3800のプロセッサ構成 最大4台の命令プロセッサが主記憶を共有する。

容量2Gバイトの主記憶は、演算能力に適合する高スループットを実現した。CMOSダイナミックメモリ素子によって最大64Gバイトの大容量化を図った拡張記憶装置は、最大転送速度が4Gバイト/sと高速であり、ユーザー用一時ファイルやジョブのスワッピングファイルとして使用できる。これらにより、大規模数値計算での処理時間や入出力時間を大幅に短縮することができる。

4.3 最先端の半導体と実装技術

S-3800では、論理部に25kゲート・遅延時間60psのECL(Emitter Coupled Logic)高速論理LSIを、ベクトルレジスタに9kゲート+18kビット・アクセス1nsの超高速メモリLSIを、また高速アドレス変換部に11kゲート+256kビット・アクセス1.6nsの高速メモリLSIを採用した。いずれのLSIも0.5μmの最先端半導体技術を使用している。

実装技術としては、セラミック基板上に高速LSIを搭載する高密度モジュールを採用した。また冷却方式は、高密度モジュールを直接水で冷却する高効率冷却方式を採用した。

4.4 高速入出力機構と動画像システム

他のスーパーコンピュータやワークステーションなどと高速にデータ転送を行うためのANSI規格HIPPIに準拠した高速入出力(HIPPI)機構を付加機構として用意

した。この機構は最大100Mバイト/sの高速な転送能力を持ち、ウルトラネットテクノロジーズ社の高速LANシステムUltraNetやディスクアレー装置と接続する。これにより、入出力のデータ転送能力を飛躍的に向上させ、マルチベンダ環境での多様なシステムニーズに対応できる。

計算結果の可視化はスーパーコンピュータでは欠かすことができないが、この可視化に有効な動画像出力システムを付加機構として用意した。拡張記憶装置を画像ファイルとして使用することにより、連続した動画をカラーモニタに出力したり、VTRに収めることができ、シミュレーション結果の確認を迅速かつ効率よく行うことができるようになった。

5 おわりに

以上、S-3000シリーズのオペレーティングシステムとハードウェアについて述べた。

スーパーコンピュータは、ネットワークを核にしたマルチベンダ環境での超高速計算処理サーバとして、高速性能や高信頼性とともオープン化、分散システム化に対応することが重要であり、今回の製品開発ではこの点に力を置いた。今後も、トータルシステムコンセプト“FOREFRONT”に基づいて、より良いスーパーコンピュータの開発を進めていく考えである。