

# ベクトルパラレルスーパーコンピュータ VPP5000シリーズのハードウェア

## Hardware of VPP5000 Series Vector-parallel Supercomputer Systems

### あらまし

本稿では高速科学技術計算処理を行うベクトルパラレルスーパーコンピュータシステム「VPP5000」シリーズのハードウェアを紹介する。本シリーズは、現行のベクトルパラレルアーキテクチャと上位互換をとりながら、実効性能向上のためのアーキテクチャ拡張を行い、ベクトルパラレルアーキテクチャを継続発展させている。また、64ビットアドレッシングを採用し、大規模処理への対応を図っている。本シリーズは9.6 GFLOPSのPE(Processing Element)を1台から最大512台まで結合することによって、9.6 GFLOPSから4.9 TFLOPSまでの幅広い性能レンジをカバーしている。さらに、I/O系では、高速な標準インタフェースであるPCIバスを採用し、高速デバイス、高速ネットワークへの対応を図っている。

### Abstract

This paper describes the hardware of the VPP5000 series of vector-parallel supercomputer systems for high-speed scientific calculations. The VPP5000 series are upward compatible with the current vector-parallel architecture and expand on it to improve performance. The series uses 64-bit addressing to enable large-volume data processing.

This series can comprise any number of 9.6 GFLOP processing elements (PEs) from 1 to 512, so it can cover the very wide performance range from 9.6 GFLOPS to 4.9 TFLOPS. Its I/O system employs a PCI bus (high-speed standard interface), so it can be connected to high-speed devices and used to configure high-speed networks.



中田達己(なかだ たつみ)

1987年名古屋工業大学電気情報工学専攻了。同年富士通入社。以来スーパーコンピュータの開発に従事。第三コンピュータ事業部第三開発部

## ま え が き

近年、科学技術計算分野の急速な進展に伴い、スーパーコンピュータは大学や研究機関のみならず、民間企業の技術開発や製品開発において欠くことのできない存在になっており、その高速演算能力、大容量メモリ、高速I/O処理に対する要求がますます高まっている。

このような状況のなかで、VPP5000シリーズは従来のVPPシリーズで培ったベクトルパラレルアーキテクチャを踏襲し、最先端テクノロジーの採用とともにアーキテクチャの改善を行うことで、最高性能のみならず実効性能を大幅に向上させ、優れた価格性能比のスーパーコンピュータを提供することをねらいとして開発した。

## 開発の方針

VPPシリーズは<sup>(1)</sup> 1 PE(Processing Element)専用機であるVPP5000Uと、512 PEまで拡張可能なVPP5000の二つのモデルから構成される。VPP5000Uは、価格性能比に優れ、設置面積2.7 m<sup>2</sup>の高性能計算サーバである(図-1(a))。またVPP5000は、スケーラブルなPE構成によって、最大4.9 TFLOPSの大規模科学技術計算を可能とするハイエンドマシンである(図-1(b))。

VPPシリーズは1プロセッサあたりの性能が優れたベクトルプロセッサを一次元の完全クロスバネットワークで数百台結合し、並列処理することを特徴とするベクトルパラレルアーキテクチャを採用している。

以下にVPP5000シリーズの開発の方針を示す。

実効性能向上のためのアーキテクチャ拡張

VPP5000シリーズは、ベクトルパラレルアーキテク

チャを踏襲し、アーキテクチャの拡張による大幅な実効性能向上を目標としている。そのために、以下の三つの並列処理技術の高いレベルでの融合を実現する。

## (1) ベクトル処理による並列技術

PEは、多数の演算を並列実行するベクトル処理を行う。

## (2) VLIWによる命令並列技術

PEのスカラユニットには、VLIW(Very Long Instruction Word)型RISCアーキテクチャを採用し、命令レベル並列処理を行う。

## (3) 複数PEによる並列技術

ベクトル型のPEを一次元の完全クロスバネットワークで接続し、分散メモリ型並列処理を行う。

大規模処理および超高速処理への対応

スーパーコンピュータに要求される大規模処理、および超高速処理に対応するためには、メモリの大容量化と高スループット化が必須である。VPP5000シリーズでは、DRAMと同等の集積度を持ち、かつクロック同期型でアクセス時間が高速なSDRAM(Synchronous DRAM)と、クロック同期型でアクセス時間がさらに高速なSSRAM(Synchronous SRAM)を採用する。

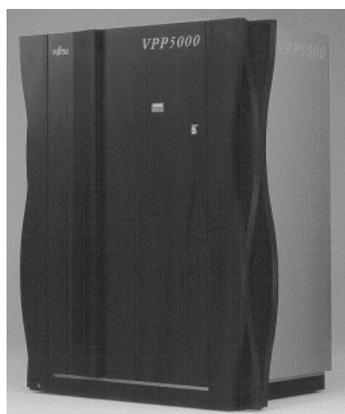
I/O処理の高速化

I/O処理については、以下の三つの高速化技術を実現する。

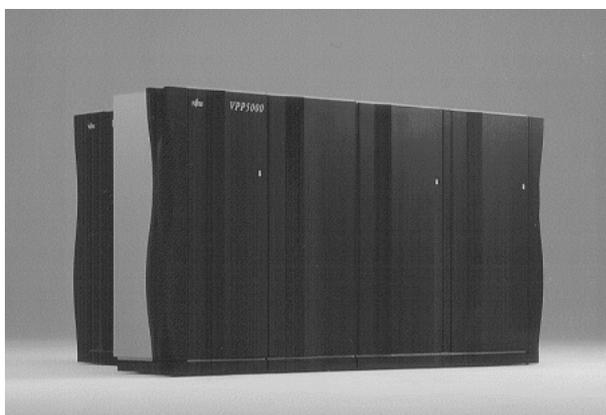
## (1) 単体PEの処理能力向上、およびメモリ容量拡大に見合うI/O処理能力の向上。

## (2) I/O処理を行うPE(IOPE)を複数台接続し、並列動作する機構(複数IOPE)によるシステムI/O処理能力の向上。

## (3) 高速デバイス、および高速ネットワークの新規サ



(a) VPP5000U



(b) VPP5000

図-1 VPP5000シリーズ  
Fig.1-VPP5000 series.

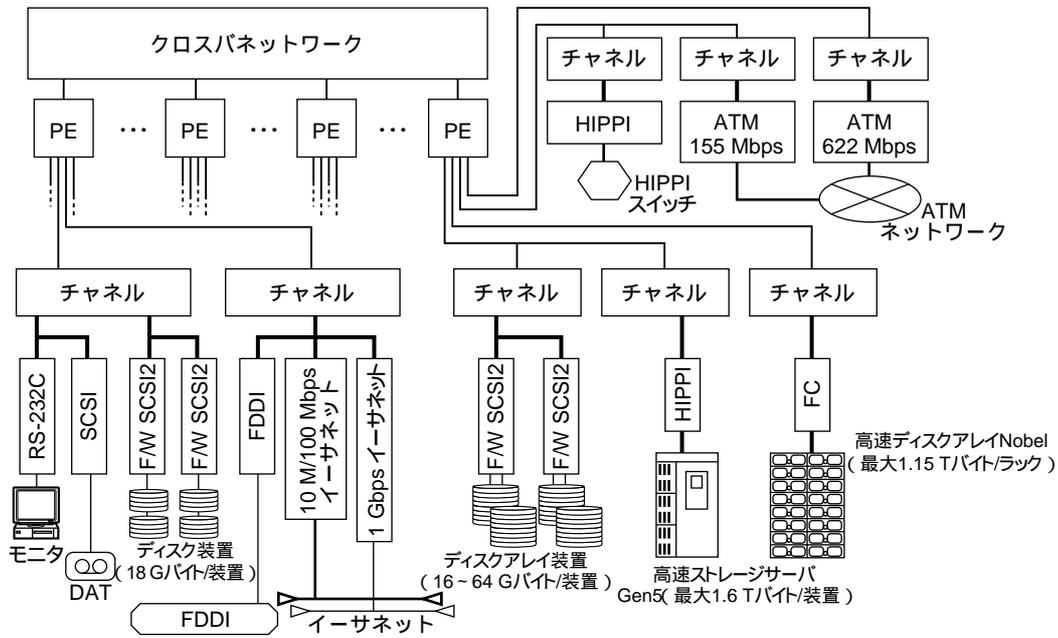


図-2 システム構成例

Fig.2-Example of system configuration.

ポート。

システム概要

システム構成

VPP5000シリーズのシステム構成の例を図-2に示す。システムはクロスバネットワークによって高速に接続された複数のPEと、PEに接続される複数のI/O装置から構成される。

システム実装

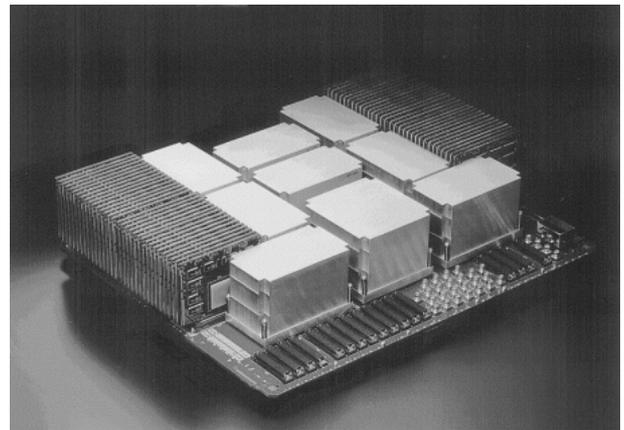
VPP5000Uモデルは一つの筐体の中に、PE1台とI/Oアダプタ、内蔵ディスク、サービスプロセッサなどを搭載する。

VPP5000モデルでは、最大4PEまでを実装可能なPE筐体、最大四つのI/Oバックパネルが実装可能なI/O筐体、最大32PEまでの結合を可能とするクロスバ筐体、システム全体の制御を行うSCC(System Control Interface)筐体から構成され、ユーザの要求する性能に応じてPE筐体、I/O筐体、クロスバ筐体を任意に増設することでスケラブルにシステムを構築することができる。

システムテクノロジー

VPP5000シリーズではゲート長0.22μm、平均負荷ゲート遅延時間24 ps、最大集積度3,300万トランジスタのCMOS LSIテクノロジーを採用している。

PEボード(図-3)には、このCMOS LSIを合計34個搭載したMCM(Multi Chip Module)と、64枚のRAMモジュールが高密度に実装されている。



CMOS LSI	34個
RAMモジュール (メモリ容量)	64枚 (2~16 Gバイト)

図-3 PEボード  
Fig.3-PE board.

構成要素

VPP5000シリーズの本体は以下に述べる各装置から構成される。

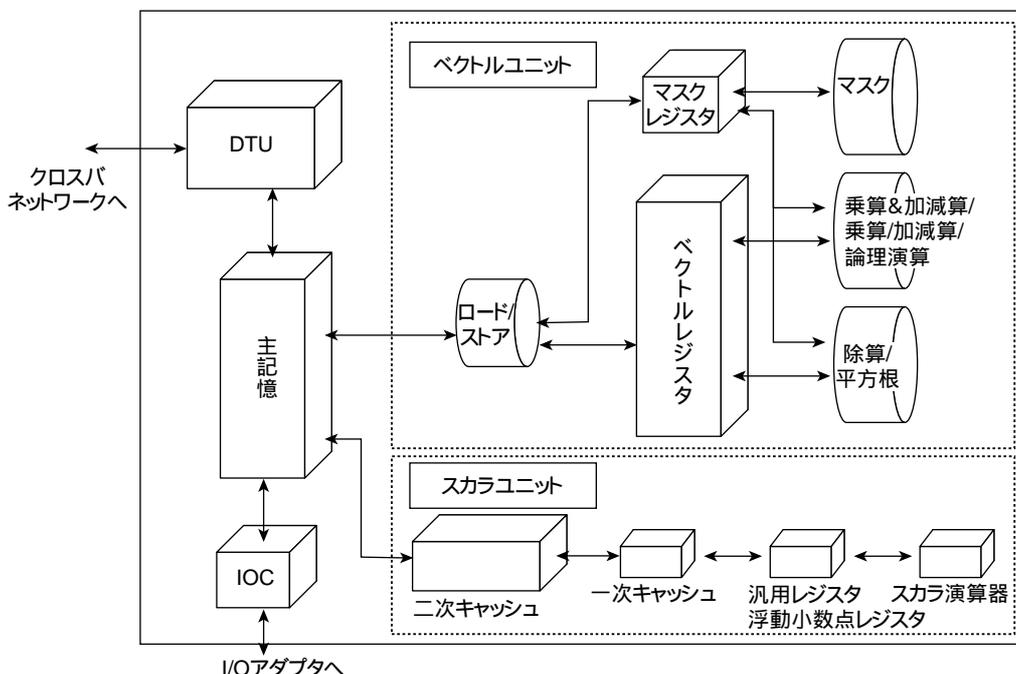


図-4 VPP5000 PE構成図  
Fig.4-VPP5000 PE configuration.

#### (1) PE

PEは以下のユニットから構成される(図-4)。

##### (i) スカラーユニット(以下, SU: Scalar Unit)

SUはスカラー演算器, 汎用レジスタ, 浮動小数点レジスタ, キャッシュなどから構成され, スカラー命令の実行, および割り込み処理などを行う。

##### (ii) ベクトルユニット(以下, VU: Vector Unit)

VUは複数のベクトルパイプライン演算器, ベクトルレジスタ, マスクレジスタなどから構成され, ベクトル命令を実行する。

##### (iii) 主記憶装置(以下, MEM: Memory)

プログラムやデータを格納する装置であり, VUが要求する大量のメモリアクセスを高速に処理する。

##### (iv) データ転送ユニット(以下, DTU: Data Transfer Unit)

クロスバネットワークを介して, PE間のデータ通信および同期処理を行う。

##### (v) I/O制御ユニット(以下, IOC: IO Controller)

各種I/Oアダプタを接続し, 入出力制御を行う。

#### (2) クロスバユニット(以下, XB: crossbar unit)

DTUを用いてPE間転送処理を行うクロスバネットワーク。

#### (3) サービスプロセッサ(以下, SVP: Service Processor)

本体装置とは独立した計算機システムであり, システム全体の制御, 診断, 保守などを行う。

## ハードウェア概要

### SU

命令レベルの並列処理と大規模処理を実現するために, VLIW方式の64ビットRISC型プロセッサを開発した。

VLIWは Intel社のIA-64アーキテクチャでも採用されている方式であり, 一命令語の中に複数の演算操作を指示することが可能となっている。並列実行を指示するフィールドへの演算割当はコンパイラが行う。そのため, スーバスカラー方式のようにハードウェアによる演算操作のスケジュールは必要なく, ハードウェアの軽量化および高速化が実現できる。

VPP5000シリーズで採用したVLIW方式の動作概略を図-5に示す。命令語長は128ビットであり, 一命令語の中に1~4個のスカラー操作(命令1)または最大2個のベクトル操作を割当て可能とした。複数サイクルを要する非同期操作(例えば主記憶アクセス操作や, 浮動小数点演算操作)を実行する場合でも最大性能を得るために, 先行する非同期操作(命令2)の実行完了を待たず後続操作(命令3)を発行する非同期実行機構を持つ。これにより, データの依存関係が崩れない範囲において, 非同期操作の実行順序を変更し, 並列効率を向上させている。

この他のSUの特徴を以下に示す。

(1) 分岐予測機構を用い, 分岐先命令の先行読み出しの高速化を図っている。また予測が困難な条件分岐に対

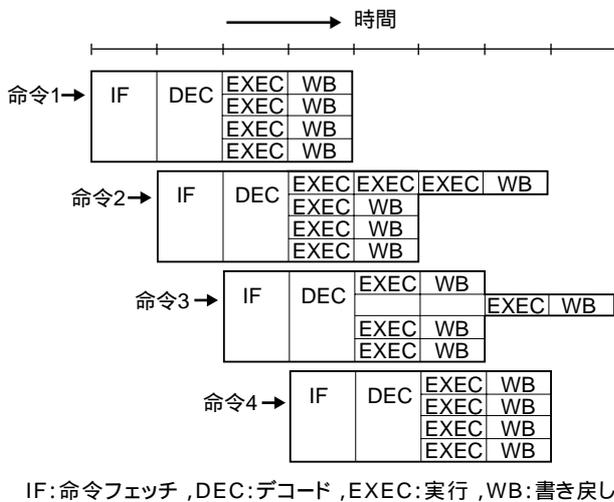


図-5 VLIW方式の動作概略  
Fig.5-Outline of VLIW operation.

しては、「条件付き無効化操作」による演算の選択的実行により、分岐操作そのものを減らし高速化を実現している。

- (2) 分岐操作を越えて命令の移動を行い、性能向上を図るコンパイラ技術「トレーススケジュール」をサポートするための命令を実装している。
- (3) 浮動小数点乗加算器、および高速な浮動小数点除算・平方根演算器を各2個装備している。また、その演算器に十分な演算データを供給するために、同時に最大4個(32バイト)の浮動小数演算データが供給できるキャッシュ性能を実現している。
- (4) データや命令をキャッシュにプリフェッチする命令を実装し、ソフトウェアによるキャッシュ制御を可能としている。
- (5) 一次キャッシュの容量を倍増するとともに、新規に2Mバイトの二次キャッシュを装備し、大規模データ処理の高性能化を実現している。

#### VU

VUは、1命令で複数の演算を実行するSIMD(Single Instruction Multiple Data)方式の代表であるベクトル処理方式を実現するユニットである。

ベクトルパイプライン演算器は、毎サイクル16個の演算をパイプライン処理する。本演算器には、乗加算、乗算、加算もしくは論理演算が可能な乗加算パイプラインと、除算もしくは平方根演算が可能な除算パイプラインと、主記憶装置との間のロードもしくはストアが可能なロード・ストアパイプラインと、条件つきで演算の実行を可能とするマスクパイプラインがあり、同時に複数の

パイプライン演算器を動作させることによって、複数のベクトルデータを並列に処理する。ベクトルレジスタの容量は128 Kバイト、マスクレジスタの容量は2 Kバイトである。

VPP5000シリーズのVUでは以下に示す各種改善により高い実効性能を実現している。

- (1) 乗加算演算器の採用や、平方根演算のハードウェア化による演算器の高速化。
- (2) ベクトルパイプラインの幅を太くし(16エレメント)、かつ本数を4本に絞ることで、最大のパイプライン性能を引き出す構成を採用。
- (3) ショートベクトル性能の向上をねらいとした、命令発信から演算完了までのトータルな演算時間(立上り時間)を短縮。
- (4) ランダムなメモリアクセスとなるインダイレクトアクセス系の高速化を実現するために、SDRAMのページモード制御を採用。

#### MEM

主記憶装置の記憶素子は、アクセスタイム45 nsの128 Mビット/チップ、および64 Mビット/チップのSDRAMと、アクセスタイム9 nsの16 Mビット/チップのSSRAMを採用した。RAMチップをRAMモジュール基板上に両面実装し、1 PEあたり64枚のRAMモジュールを使用している。またVUのメモリアクセス要求を満たすスループットを得るために、主記憶全体を独立にアクセス可能な512の単位(WAY)に分割し、76.8 Gバイト/秒のメモリスループットを実現している。

VPP5000シリーズでは以下のRAS機能によって主記憶の信頼性の向上を図っている。

#### (1) ECC機構

MEMにはECC機構を備えており、1ビットエラーの完全訂正、および2ビットエラーの完全検出を行っている。

#### (2) パトロール機構

間欠に発生する1ビットエラーの復旧のために、定期的に主記憶データを読み出し、エラーの有無を検査するパトロールを行っている。もしデータ読出し時に1ビットエラーが検出された場合には、データは訂正され主記憶に再度書き込まれる。

#### (3) ページ切離し

固定的なエラーを回避するために、2ビットエラーが発生した主記憶領域をページ単位で切り離す。

#### DTU

DTUは各PE内に搭載され、演算とは独立にXBを介してPE間通信を実行する。DTUはデータ転送処理部とPE間

同期処理部から成る。データ転送処理部は以下のような特徴を持つ。

- (1) データの送信，受信はそれぞれ独立に並行して行うことが可能である。
- (2) データ転送時のメモリアクセスとして，連続，等間隔，部分配列，間接の四つのパターンがあり，PE間転送の効率を上げている。
- (3) 転送データのメモリアドレス，および転送先PEアドレスを変換するアドレス変換機構を持ち，PE番号とメモリアドレスの仮想化を実現している。

PE同期化処理部は以下のような特徴を持つ。

- (1) 各PEのプログラム進行情報をすべてのPEに放送する機構と，この情報を受信して同期完了を検出する機構を持つ。
- (2) 各PE内に同期をとるべきPEグループを示すマスクレジスタを持ち，プログラムを任意のPEから成るグループで走行することを可能としている。これによって複数の並列プログラムを効率的に実行可能としている。

## XB

XBはすべてのPEのDTUと接続され，PE間のデータ通信路を提供しており，任意のPE間で最大3.2 Gバイト/秒×2(送信と受信)の転送スループットを持つ。最大512 PE接続時にはネットワーク全体のバンド幅は，1.6 Tバイト/秒になる。XBには以下の特徴がある。

- (1) 相手先PEが通信中でない限り通信可能な一次元完全クロスバ方式を採用しているために，ネットワークの競合が極めて少ない。
- (2) 各PE間の距離がすべて等しく，PEの使用状況に合わせて任意のプロセッサを選択し，グループ化してもネットワークの性質が変化しない。これにより，複数の並列プログラムの効率的運用が可能である。

## IOC

IOCはPEに接続される各種I/O装置のインタフェースを制御する。VPP5000シリーズのI/O処理能力は1 IOPEあたり最大3.0 Gバイト/秒であり，高速なI/O装置を多数接続可能である。

また，高速アダプタの標準インタフェースであるPCIバスを新規にサポートし，100 Mバイト/秒のファイバチャネルをインタフェースに持つ高速ディスクファイル装置(Noble)や，1 Gbps-イーサネット，ATM622などの高速ネットワークをサポートしている。

## SVP

SVPは，本体システムハードウェア，および通信回線などとの各種インタフェースを持ち，各種運用機能をサ

ポートする。以下にSVPの特徴を述べる。

### (1) 電源制御

システム全体の電源投入切断を行う。VPP5000においては各筐体単位での電源切断を行い，筐体単位の活性保守を可能としている。

### (2) 構成制御

OSの指示により，障害PEの切離しを行い，システムの運用を中断することなく，構成変更を可能とした。またシステムパーティション機能によってシステムを論理的に分割し，複数のシステムとして運用を可能としており，例えば運用システムとテストシステムを同時に使用可能としている。

### (3) 本体システムの監視

本体システムの動作状態を監視し，異常を検出した場合，状態をログ情報として収集し，通信回線を介して保守センタに通知することにより，故障の修復を迅速に行うことを可能としている。

### (4) 自動運転機能

あらかじめ設定された指示に従い，OSの初期プログラムロード(IPL)を行ったり，OSからの指示により，システムの電源を切断したりする，自動運転機能をサポートしている。

## 性 能

スーパーコンピュータの性能測定に広く使用されるLINPACKベンチマーク<sup>(2)</sup>での行列演算処理性能を示す。

ショートベクトル性能の改善を行ったVPP5000シリーズは，100円で1,125 MFLOPSと他社をはるかに凌駕する世界最高性能を達成し，1,000円の性能も8,759 MFLOPSと単1 PEでは世界最高性能を達成している(表-1)。

複数PEを使用する自由元の結果は，強力なPE間データ転送能力により，PE数に比例してシステム性能が向上することを示している(図-6)。並列化効率も極めて高く，32 PE構成時に296 GFLOPSと，ピーク性能の96.4%の実効性能を達成している。

## む す び

VPP5000シリーズの開発のねらい，システム概要，ハードウェア機能概要，性能について述べた。各種の強

表-1 単一プロセッサ処理性能 (MFLOPS)

	VPP5000 (SSRAM)	他社最高値
LINPACK 100円	1,125	856
LINPACK 1,000円	8,759	7,280

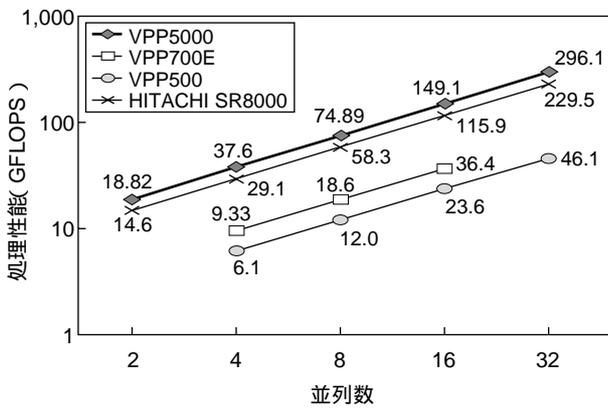


図-6 並列処理性能(LINPACK自由元)  
Fig.6-Performance of parallel processing.

化によって、現行のベクトルパラレルアーキテクチャと上位互換を保ちながら、世界一の実効性能を達成し、さらにスーパーコンピュータに要求される大規模処理への

対応を実現した。

スーパーコンピュータを利用した科学技術計算分野においては、その計算規模の増大に伴い、高速処理に対する要求はとどまるところがない。また、スーパーコンピュータの利用範囲も今後いっそう拡大するであろう。今後もこれに応えるために、さらなる演算性能・I/O性能の向上、主記憶容量の拡大、価格性能比の改善を図っていく所存である。

#### 参考文献

- (1) 内田：世界最高速スーパーコンピュータの開発．電気評論，平成十一年臨時増刊号第八四巻，pp.73-77(1999)
- (2) Jack J. Dongarra：Performance of Various Computers Using Standard Linear Equations Software.  
<http://www.netlib.org/benchmark/performance.ps>

