

解説

高速プロセッシングデータバス技術

1. Sun Ultra 2 の UPA 相互接続テクノロジー概説

An Outline of Sun Ultra 2 UPA Interconnect Technology by Hiroki INOUE (Product Marketing Dept., Marketing Div., 日本 Sun Microsystems K.K.).

井上 洋紀¹

¹ 日本サン・マイクロシステムズ(株)マーケティング本部

1. はじめに

Sun は 64Bit プロセッサである UltraSPARC を搭載したワークステーション Sun Ultra 1/ Ultra 2 を 1995 年秋に、ハイエンド・サーバとして Ultra Enterprise x000 を 1996 年春に発表した。

Sun Ultra 1/ Ultra 2 では新たに採用された 64Bit プロセッサ UltraSPARC により卓越した演算性能を実現すると同時に、Creator/Creator 3D グラフィックスシステムが優れたグラフィックス性能を発揮している。

Sun では Ultra Enterprise シリーズのハイエンド製品において、いままで汎用機で実現されていたデータ容量と性能そして信頼性について UNIX をベースで実現することに成功したと考えている。

PC の高性能化が進み、「PC とワークステーションの違いが不明確になった」という評価がある。確かに、以前の PC と比べると GUI を処理する能力を完全に身につけ、利用される範囲も非常に広がった。しかし、一部の「PC ワークステーション」と呼ばれる機種には、PC に高機能なグラフィックス・アダプタを装備するだけであり、いわゆるワークステーションとしての能力を発揮できるのか疑問を感じる。たとえば、Sun ではハイエンドワークステーションとしては 3 次元データを扱うような高レベルの CAD や、より高速なシミュレーションを行うことを前提としており、最大 14CPU 構成のマルチプロセッサマシンをデスクトップマシンとして実現している。PC のデスクトップマシンでは、4CPU より多いマルチプロセッサ構成の機種は極端に少ない。サーバではこの差がさらに拡がり、Sun では 64CPU 構成のモ

デルをラインナップしているのに対し、PC サーバではデスクトップと同一の 4CPU モデルをその最大構成としている場合が多く見受けられる。

本稿は、UltraSPARC を採用したワークステーションのなかでも Ultra 2 を中心にワークステーションに採用されている UPA 相互接続テクノロジーを解説していくことを目的とする。

UPA 相互接続テクノロジーは、デスクトップ PC などのアーキテクチャと比較するとまさにバスの役割をつかさどる部分であり、この UPA 相互接続をバス・アーキテクチャを比較することにより、高速プロセッシングをどのように実現しているか、4CPU 以上での MP 性能をどのように実現しているのか、そのテクノロジーと考え方について理解していただけるものと考えている。

2. Sun のラインナップと設計意図

現在 Sun では、UltraSPARC を採用したワークステーションとサーバを 1 から 64CPU までスケラブルにラインナップしている。すべて同じ Solaris 2.5 が動作し、同じバイナリで動作する。そして、その性能が CPU の数に従って推移する。

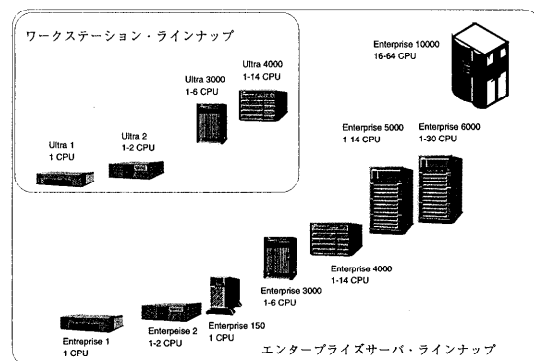


図-1 ワークステーションとサーバのラインナップ

表-1 ワークステーションの適用分野と求められる機能

分野, 用途	求められる機能, 性能
ソフトウェア開発	使いやすい OS 環境 デバック, コンパイルに十分な演算性能
機械系設計 (MCAD/MCAE)	高度なグラフィックス処理
電子設計 (EDA)	高度なグラフィックス処理 高い演算能力
出版/イメージ処理	グラフィックスとイメージ処理 高い演算能力
研究開発分野	使いやすい OS 環境

表-3 SPARCstation 20 と Ultra 2 の比較

アーキテクチャ コンポーネント	SPARCstation 20	Ultra 2 Model 2300	性能比
CPU	32 ビット 75MHz SuperSPARC-II プロセッサ×4	64 ビット 300MHz UltraSPARC プロセッサ×2	2 倍から 5 倍
CPU と外部 キャッシュの 接続	64 ビットデータ キャッシュミス時 に命令実行が滞る	128 ビットデータ キャッシュミス時 に命令実行が滞ら ない	2 倍以上
CPU/システム インタフェース	MBus 回線交換式 動作周波数 50MHz 最大 400MB/s	UPA パケット交換式 動作周波数 100MHz 最大 1600MB/s	3 倍から 5 倍
メモリ	144 ビットバス 600ns のレイテン シ 20MB/s で Bcopy BW を実行	576 ビットバス 170ns のレイテン シ 180MB/s で Bcopy BW を実行	2 倍から 4 倍
グラフィック スとの接続	動作周波数 25MHz 32 ビット SBus 接 続	動作周波数 100MHz 64 ビット UPA 接続	2 倍から 6 倍
ディスクイン タフェース	10MB/s の Fast SCSI	20MB/s の Fast/Wide SCSI	2 倍
ネットワーク インタフェース	10Mbps の イーサネット	100Mbps の Fast イーサネット	10 倍

ことが可能である。つまり、パフォーマンスが必要なユーザはより CPU の多いマシンを選びさえすればよい。

ここでワークステーションとサーバに求められる性能とは、それぞれどのようなものであるか、Sun がどのような用途での利用を想定して設計を行ったのかを考えてみたい。

まずは、ワークステーションとサーバがそれぞれの分野でどのような機能と性能を求められるのかをまとめてみる。

図-1 をみると、ワークステーションとサーバがかなり違って製品構成をしていると感じるかもしれない。しかし、サーバとワークステーションでは、その設計を何に重点を置くかが異なっていることが、2つの表からうかがえる(表-1, 表-2 参照)。もちろん、演算性能は両者ともに共通した重要な要素であるが、主にワークステー

表-2 サーバの適用分野と求められる機能

分野, 用途	求められる機能, 性能
企業基幹系システム	大規模データベース性能
業務アプリケーション	トランザクション性能
意思決定支援	バッチ処理性能
ファイルサーバ	ファイル入出力性能
電子メールサーバ	比較的小規模な I/O 性能
WWW サーバ	

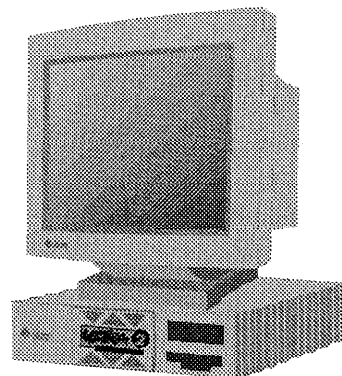


図-2 Ultra 2 外観図

ションではグラフィックス性能、サーバでは I/O 性能が重要となる。

図-1 の製品は、ワークステーションの用途として最適のパフォーマンスを発揮するように考えられた Ultra 1, Ultra 2 と、サーバとして最適なパフォーマンスを発揮する Ultra Enterprise 3000, 4000, 5000, 6000, 10000 ラインとに大分される。今回はデスクトップ・ワークステーションとして、Sun の主力製品である Ultra 2 の内部構造について説明する。

3. Sun Ultra 2

Sun Ultra 2 は CPU に UltraSPARC を 2 つ搭載することが可能なマルチプロセッサ対応のハイエンド・ワークステーションとして設計された。UltraSPARC は以前の Sun のハイエンドモデルである SPARCstation 20 に搭載されていた SuperSPARC 75MHz と比べると 2 倍から 5 倍のパフォーマンスを発揮する。

表-3 では CPU 以外のコンポーネントについて Ultra 2 と以前の製品である SPARCstation 20 とを比較する。

表のように Ultra 2 の設計時には単に高速の CPU を採用するだけではなくシステムとしての

全体的なパフォーマンスを向上させるためにさまざまなコンポーネントを強化している。図-3 はこれらのコンポーネントの接続を模式化したものである。それぞれ強化されたコンポーネントはシステムバスである UPA によって相互に接続されている。

周辺装置のパフォーマンス向上のためには、システムバスのパフォーマンスとそのアーキテクチャが重要である。ここでは、UltraSPARC, UPA について詳しく説明する。

3.1 UltraSPARC プロセッサ

Sun Ultra 2 に採用した、UltraSPARC の基本アーキテクチャは SPARC アーキテクチャ Version 9 である。

この Version 9 のアーキテクチャは 1994 年に「The SPARC Architecture Manual, Version 9」として発表された。Version 9 の特長としては「64 ビット仮想アドレスおよび 64 ビット整数データ」「フォルトトレラント」「高速のトラップハンドリング、コンテキストスイッチング」「ビッグエンディアンとリトルエンディアンのサポート」がある。もちろん過去の SPARC との互換性もアーキテクチャのなかで定義されている。

UltraSPARC は上記の SPARC Version 9 としての特長のほかに、以下のような特長をもち高いパフォーマンスを発揮する。

- 9 段のパイプライン
- 4 命令同時発行のスーパスケラ
- 動的分岐予測
- 16K バイト + 16K バイトのチップ内キャッシュ
- 512K バイトから 4M バイトの外部キャッシュ
- Visual Instruction Set

現在のマイクロプロセッサではごく一般的であるパイプライン、スーパスケラを実装することにより、同一時間内で処理する命令数を増やし、高性能を実現している。

CPU 本来の性能を出すためにはこれらの内部機能をいかに休みなく動かし続けるかが、重要となる。命令やデータを効率的に CPU に供給するのは最終的にキャッシュが行うことになる。したがって、キャッシュをどれだけ効率よく使えるようにシステムを構築するかが重要でありパフォー

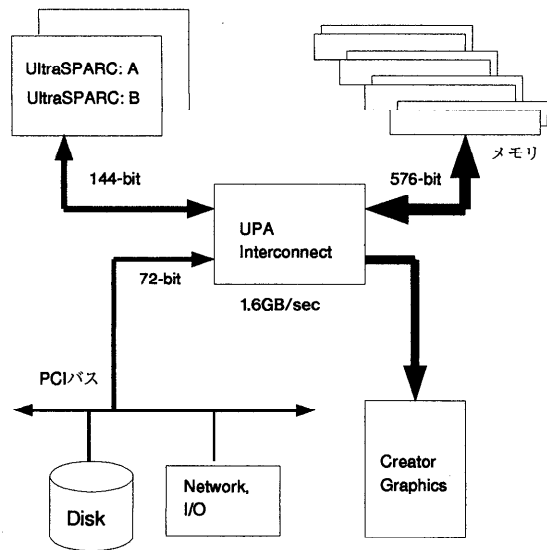


図-3 UPA 概略図

マンスを大きく左右することになる。

現在 PC とワークステーションのほとんどは、上記のキャッシュを一次キャッシュと二次キャッシュとして装備している。容量としては、一次キャッシュは 8K から 32K 程度であり、二次キャッシュのサイズは最低で 256KB から 1MB 程度が一般的である。もちろん一次キャッシュが十分に大きければ二次キャッシュを配する必要はない。しかし CPU 内部にさらに大きいキャッシュを装備すると CPU のシリコンのサイズを大きくしなければならず、チップ製造の効率が著しく落ち製造が困難になる。結果として CPU の価格上昇を招き、システム全体の価格をあげることになってしまう。このため、CPU 内部には必要最低限のキャッシュメモリを搭載し、CPU のシリコンとは別にキャッシュを配することが一般的である。

ここで、一次キャッシュと二次キャッシュの位置に注目し、システム構成図-4 と図-5 をみてみたい。

図-4 の構成の場合、同じバス上に二次キャッシュが接続されているため、CPU: A と CPU: B が同時に二次キャッシュの中を参照する必要が出たときには、どちらかの CPU が処理を中断しバスが解放されるのを待つ必要が出てくる。したがって、この構成では二次キャッシュでの衝突が、頻繁に発生する可能性が高く、マルチプロセッ

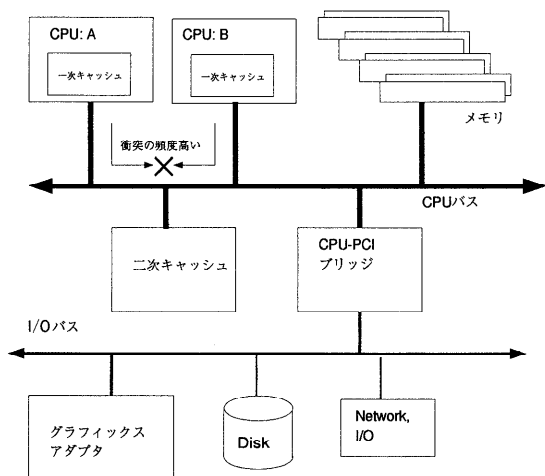


図-4 システム構成図

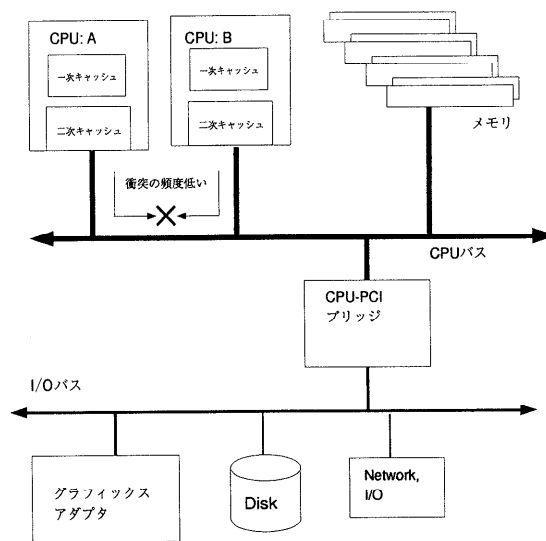


図-5 システム構成図

サ・システムでのパフォーマンスを発揮することが困難である。

筆者も Pentium ベースのマルチプロセッサ対応 PC を使用しているときに、1CPU モデルに CPU を追加することにより、ベンチマーク性能の劣化が起こることを体験している。このときも図-4 の要因が大きく影響していた。もちろん、すべてのプログラムで劣化がおこるわけではなくパフォーマンスの向上がみとめられた、ベンチマークプログラムも存在した。おそらく、CPU の一次キャッシュにそのプログラムがすべて収容できるサイズであることと、データの依存性がないことなどがこの条件となる。このようなプログラムは、ごく簡単なベンチマークプログラムでは見受けられるが、ハイパフォーマンスが期待される実際のアプリケーションにおいては、このようなアプリケーションはほとんど存在しない。

これに比べ、図-5 の構成のときには、それぞれの CPU が行っている演算に依存性がある場合にはバスを経由して、それぞれの CPU 側にある二次キャッシュの状態を変更する必要があるが、自身の二次キャッシュのアクセスでの衝突は起こらない。したがって、衝突頻度は激減する。現在主流となりつつ PC のマルチプロセッサ構成も図-5 の構成が主流になりつつあるようである。しかしながら、PC においてもそのスケーラビリティは 4CPU 程度から劣化が、報告されているようである。UltraSPARC の場合は、さらにデ

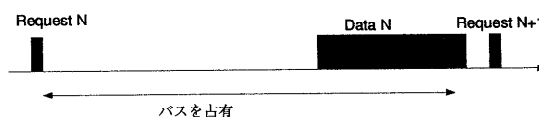


図-6 回線式交換型バス

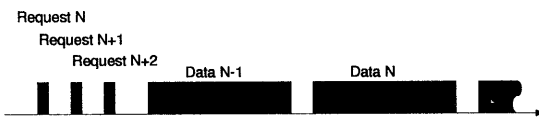


図-7 パケット交換式交換型バス

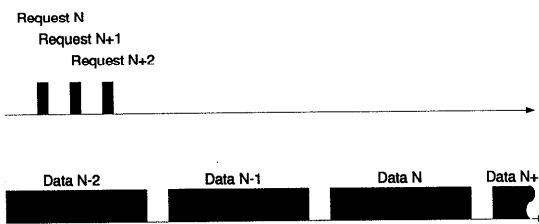


図-8 アドレスバスとデータバスの分離

ータ転送を高速化するために UPA (Ultra Port Architecture) 相互接続とよばれるクロスバーテクノロジーを採用している。複数の CPU とメモリのアクセスという観点から、次の節でこの UPA 相互接続について解説したい。

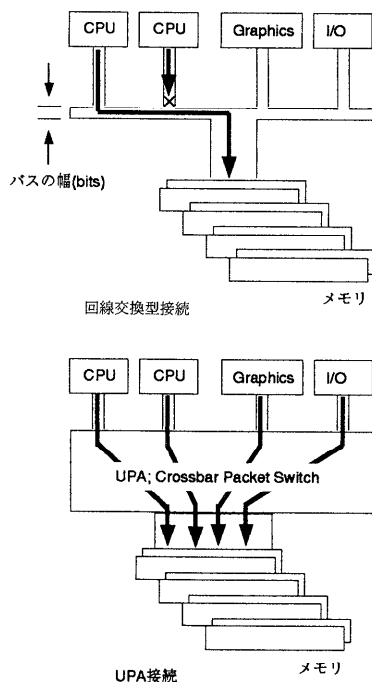


図-9 回線交換式接続とUPAによるパケット交換式接続

3.2 Ultra Port Architecture (UPA) 相互接続

図-3がUltra 2におけるシステム構成図となる。UPAは100MHzで動作する。これはPCのバス66MHzとくらべても高速である。さらに、UPAは回線交換式のバス結合ではなくパケット交換でクロスバー接続を行う。ここでは、このUPAについて解説する。

CPUはメモリや周辺機器に対して、データをReadあるいはWriteする。回線交換式のバスでは、図-6のようにRequestを出した後、実際のデータをバスにReadあるいはWriteを行う。

メモリをはじめとした、CPU以外のデバイスはCPUからみると非常に低速で、リクエストを出してから実際のデータのやり取りに遅延が生じる。通常回線交換式のバスでは、1つのリクエストを発行してから、データのやり取りが終了するまで、バスを占有することになる。

図-7のようなパケット交換式ではそれぞれのリクエストとデータのやり取りを文字どおりパケットとして取り扱う。これによりリクエストからデータ転送の間に任意のリクエストあるいはデータ転送を行うことを可能にしている。

この現象を模式的に描くと図-9となる。1つのCPUがメモリに対してアクセスを行っているときはほかのデバイスがデータのやり取りを行うことができない。回線交換型の接続では、決まった転送能力をさらに、複数のCPUで奪い合う格好になってしまい効率が悪くなる。これに対し、UPAによる接続は同時に複数のアクセスが可能である。

さらにUltra 2においては、リクエストを送るためのアドレスバスとデータを送るデータバスを別々に用意している。図-8に示すように、リクエストにおけるアドレス指定と、データの転送を互いに干渉されずに行うことができる。Ultra 2のUPAでは、100MHzで動作し、最大1.6GBのデータ転送を可能にしている。また、パケット交換を採用したことにより、実際のアプリケーション動作時にもその転送速度が劣化しないという特性がある。

また、PCアーキテクチャで転送速度を向上させるために各I/OにたいしてDMAデバイスを配することが一般的である。しかし、UPAとくらべると構成が複雑になるという欠点がある。

現在のPCのバス転送能力は500MB/sec.程度であるが、マルチプロセッサ構成の場合の実効転送能力は、半分程度になってしまう。Ultra 2ではUPAによって実現している。UPAではクロスバーパケットスイッチを使っていてCPUからの接続を並行して行うことができる。UPA接続を行うことにより、実際のアプリケーション動作時にも、最大の転送能力に近い能力を発揮し、それによりとくにマルチプロセッサ構成のときのパフォーマンスの劣化を防ぐ。

グラフィックスデバイスについてもPCなどではPCIデバイスとして、接続するのが一般的である。PCIバスはCPUバスの転送速度と比べると遅いバスになる。Ultra 2では図-3で示すとおりUPAに直接接続することにより、高速なグラフィックス環境を実現している。

4. おわりに

高性能コンピュータにおいてマルチプロセッシング構成はいまやごく一般的な技術となった。CPU単体の性能に限られるとなれば、あたり前の進化である。今回解説を試みたように、マルチ

プロセッサ構成のシステムは、バスのパフォーマンスによりシステム全体の性能が大きく左右される。

SPARC に代表される Sun のアーキテクチャはオープンである。Sun はオープン・システム自身を次のように定義している。

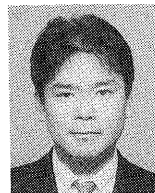
- 1) インタフェースが明確に定義されていること
- 2) その実装が複数存在すること
- 3) 技術のライセンスが無料もしくは適価で手にはいること、

言い換えると、互換機を作りやすい環境を作るために、インタフェースを定義しそれを安価で公開している。現実 Sun の互換機を製造販売するコンピュータメーカーがたくさん存在する。結果として SPARC のシステムはインタフェース・レベルでの統一性が保たれるだけでなく、どのレベルで互換性を保てばよいか明確であるため、設計の自由度がある。また、その定義の際に将来の拡張性を考慮していたことが、SPARC: Scalable Processor ARChitecture という名前からもうかがえる。

一方 PC のアーキテクチャは、製品の使用がたまたま公開されていて、それをすべて模倣することが前提となっている。PC の互換機においても、過去のソフトウェア資産を動作させることが第一の命題としてあり、CPU やバスのアーキテクチャをはじめ、数多くの制約事項がある中で、どのように工夫をして早くしていくかを追求する技術である。

今回は Ultra 2 について解説を行ったが、同様のアプローチをほかのマシンでも行い、ラインナップの構成を実現している。今後 CPU の単体性能に限界がくるとすれば Sun のような SMP 構成による高性能化のアプローチが注目されるであろう。また同様の理由でクラスタリングのテクノロジーが、注目され始めてきた。クラスタリングをしたシステムでも、システムの性能を左右するのはノード間の接続方式とそのスピードになると考えられる。当然これらは、アプリケーションプログラムの構造とも、深く関わり進化していくことになろう。またこの傾向は、コンピュータの単体性能やアーキテクチャのみでの比較がますます困難になり、アプリケーションやあるいは構築されたシステム全体としての評価をする必要が大きくなることを意味すると考えられる。日本ではコンピュータのハードウェアやアプリケーションでは、欧米に並ぶものがなかなか出てこないといわれることがしばしばあるが、ぜひ優れたシステムが構築され国際的に認知紹介されることを期待したい。

(平成 9 年 4 月 14 日受付)



井上 洋紀

1991 年日本サン・マイクロシステムズ(株)入社。ソフトウェアエンジニアリング部にて、ソフトウェア開発のサポート業務に従事。1997 年より、Workgroup Server と周辺機器の製品を担当。