



White Paper

Intel® Itanium® 2 プロセッサを搭載する SGI® Altix®サーバによる
エンジニアリングにおける高性能計算への要求への対応



目次

エグゼクティブ・サマリ.....	1
MCAE における解析レベルのブレイクスルー.....	2
業界をリードする MCAE 解析ソリューション.....	4
ALTIX®での MSC.NASTRAN™ベンチマーク(シングル CPU).....	4
スケーラビリティのベンチマークテスト.....	5
SGI Altix 3000 サーバ・プラットフォーム.....	7
SGI ALTIX システムのキー・テクノロジー.....	7
SGI LINUX®ツールによるパフォーマンスの向上.....	9
MSC.Nastran.....	10
バイナリ互換の実現.....	10
FFIO による I/O スループットの向上.....	10
結論.....	12

Intel® Itanium® 2 プロセッサを搭載する SGI® Altix® 3000 サーバと MSC.Nastran™ ソフトウェアの組み合わせは、極めて大きな自由度の有限要素解析 (Finite Element Analysis: FEA) を可能とするスケーラブルでハイパフォーマンスなソリューションです。

エグゼクティブ・サマリ

技術計算における有限要素解析 (FEA) では、その解析に対する要求は、厳しくなりつつあります。数十万から数百万自由度で設計解析を行うことが一般的となり、ハイパフォーマンス、高いスケーラビリティ、コスト効果の高いコンピューティング・ソリューションが求められています。専用開発されたプロセッサとプロプライエタリなオペレーティングシステムをベースとした従来のコンピューティング・ソリューションでは、この要求に十分に答えることが困難になってきています。SGI® Altix® 3000 サーバ・プラットフォームは、FEA やその他のエンジニアリング・アプリケーションの実行プラットフォームとして、ハイパフォーマンスでコスト効果の高いソリューションを提供しています。

Intel® Itanium® 2 プロセッサ・ベースの SGI Altix システムは、オペレーティングシステム (OS) として、64 ビット Linux® ディストリビューションを採用しています。この Linux 上で、代表的な FEA ソフトウェア MSC.Nastran を稼働させることで、解析モデルの自由度の増大に対応可能な非常にスケーラブルな独自のソリューションの提供を可能とします。Itanium 2 プロセッサの性能、SGI® NUMAflex™ テクノロジー、SGI の Linux OS 用機能拡張ツールは、Itanium 2 マイクロプロセッサ用にコンパイルされた MSC.Nastran ソフトウェアと連動して、大規模な解析問題の処理時間を短縮することができます。本ホワイトペーパーでは、FEA 問題解析における SGI Altix システムと MSC.Nastran のパフォーマンスを調査し、次に、産業用コンピューティングで高生産性を実現するこのソリューションのパフォーマンスを支えるテクノロジーについて解説します。

MCAE における解析レベルのブレイクスルー

製造業の設計担当者は、社内からと顧客から要望される製品を政府が規定する基準や規制にに応じて性能・機能性・安全性に優れたデザインを、コストを抑えて設計することが求められており、製品設計ステージにおける要求はますます厳しくなりつつあります。これは、自動車・航空宇宙産業における複雑なメカニカル設計での解析に特に顕著な傾向となっています。設計者は、長年の間、有限要素モデリング (Finite Element Modeling: FEM)、有限要素解析 (Finite Element Analysis: FEA)、衝突解析 (crash synthesis) に、コンピュータを使用したメカニカル・コンピュータ支援エンジニアリング (MCAE) を用いてきました。MCAE におけるメカニカル設計解析でさらに高いレベルが求められるにしたがって、この分野でのテクニカル・コンピューティングに対する要求は年を追うごとに厳しくなっています。

自由度が増大し続ける解析問題に対応するため、コンピューティング・ソリューションには格段に高度な解析を実行完了できる性能が求められるようになってきています。今日の自動車設計においては、数十万から数百万自由度のモデルが一般的となっており、航空宇宙産業における設計では数億自由度が要求されています。さらに近い将来には、数十億自由度の問題が実行されるようになるはずですが、今後、より大規模なデータセットが生成され、ますます集中する浮動小数点演算を処理する能力が要求されることとなります。こうなると、メモリ空間のアドレスに制限のある 32 ビット・コンピューティングの処理能力の限界を超えます。

このような大規模な問題の解析には、適正な時間内に正確な解析を完了するための先進のアルゴリズムを持つ非常にパワフルなコンピューティング・プラットフォームとソフトウェアが必要です。MCAE のためのコンピューティング・プラットフォーム・ソリューションを開発するにあたっては、正確性と解析完了までの時間が重要な考慮点となります。MCAE 解析で使用される現在のコンピューティング・ソリューションの大半は並列処理システムであり、プラットフォーム・ハードウェアには、特に大規模な問題の厳しい要求に対応できる次のような特徴を備えています。

- 高速プロセッサ、高スループット、強力な浮動小数点演算エンジン、大容量キャッシュ
- 拡張の容易さ
- 64 ビット・アドレス空間を持つ大容量メモリ
- 低レイテンシで高いシステムバンド幅
- ノード間的高速転送を可能にするインターコネクト

並列コンピューティング用に設計された解析ソフトウェアでは、解析時間を最短にするには、転送時間に対する計算時間の割合をできる限り高め、オペレーティングシステム (OS) はハードウェアの特長を最大限に活用できなければなりません。

従来、MCAE のためのコンピューティング・ソリューションは特別に開発された 64 ビット・プロセッサ・テクノロジーとプロプライエタリなオペレーティングシステムをベースに構築されており、このために高価なシステムとなっていました。現在では、Intel Itanium 2 プロセッサとオープンソースの Linux オペレーティングシステムをベースにする MCAE のためのパワフルな並列コンピューティング・ソリューションが開発されており、より低価格でメカニカル設計解析の要求に対応しています。これらのシステムは、複雑な問題をすばやく処理する 64 ビットの Itanium 2 プロセッサ用に最適化されたハードウェアとソフトウェアを稼動し、MCAE のための高速なソリューションを提供します。業界をリードするこのようなコンピューティング・ソリューションの 1 つが、MSC.Nastran 解析ソフトウェアを稼動する Itanium 2 プロセッサ・ベースの SGI Altix 3000 サーバ・プラットフォームです。

業界をリードする MCAE 解析ソリューション

SGIとMSCの両社は、FEAなどのバーチャル・プロダクト・開発(VPD: 仮想製品開発)のためのソリューションを長年にわたって提供してきました。これまでも MCAE 向けの数々の優れた製品が発売されてきましたが、中でも SGI Altix 3000 プラットフォームと Linux 対応の MSC.Nastran は、最近登場した画期的な製品です。SGI と MSC の組み合わせによるソリューションは、高度にスケーラブルな並列処理プラットフォームとして他製品よりも優れた価格性能比を実現しています。

ALTIX®での MSC.NASTRAN™ベンチマーク(シングル CPU)

MSC Software は、表 1にあるような様々な自由度の問題と異なった解析種類で、MSC.Nastran のベンチマークを行っています。これらのテストは、テスト対象システムと OS に対応した標準の MSC.Nastran を使用して、様々なハードウェア・プラットフォーム上で実行されています。SGI Altix システムでのテストでは、Linux 対応の標準の MSC.Nastran ディストリビューションが使用されました。テストされたシステムの全構成については、MSC Software の Web サイトに掲載されています。

自由度とアプリケーションの使用メモリ量、スワッチ・ディスク領域、総 I/O 量は、非常に複雑な関係にあります。実際には、各解析で必要とされる MSC.Nastran のメモリ量は、32 ビット・コンピューティングでの最大値となる 4GB(ギガバイト)を大きく下回っているにもかかわらず、一ケースを除いて、スワッチ領域として、確保される領域は、32 ビット・システムのアドレス可能領域を大きく超えています。これらのスワッチ領域をメモリ上に確保出来ない場合には、これらのスワッチ・データへのアクセスは、ディスクへのアクセスとなるため、性能が低下します。

表 1: MSC.Nastran™ v2004 シリアルテスト問題

Name	N Degrees of Freedom	Description	MSC.Nastran™ Solution	Memory Used	Scratch Disk Space	Total I/O (GB)
LGQDF	31,125	Cube with Interior	108	100 MB	600 MB	1,000
XLEMF	654,560	Car Body	111	400 MB	11 GB	400
XLOOP	486,573	Car Body	200	1 GB	28 GB	1,700
XXAFST	2,490,516	Propeller Housing	101	400 MB	11 GB	77
XLTFD	529,027	Car Body	108	450 MB	5 GB	209
XXCMD	1,584,622	Car Body	103	800 MB	43 GB	2,400

図 1 は、標準の Linux ディストリビューションで MSC.Nastran コードを実行するシングルプロセッサの SGI Altix システムによる解析完了時間(単位:秒)を表しています。Linux の標準ディストリビューションには、Linux 上でのパフォーマンス向上を実現するために SGI と MSC Software が共同開発した MSC.Nastran の最適化機能は含まれていません。これらのテスト結果は、SGI Altix システムにおける MSC.Nastran の基準値を示しています。

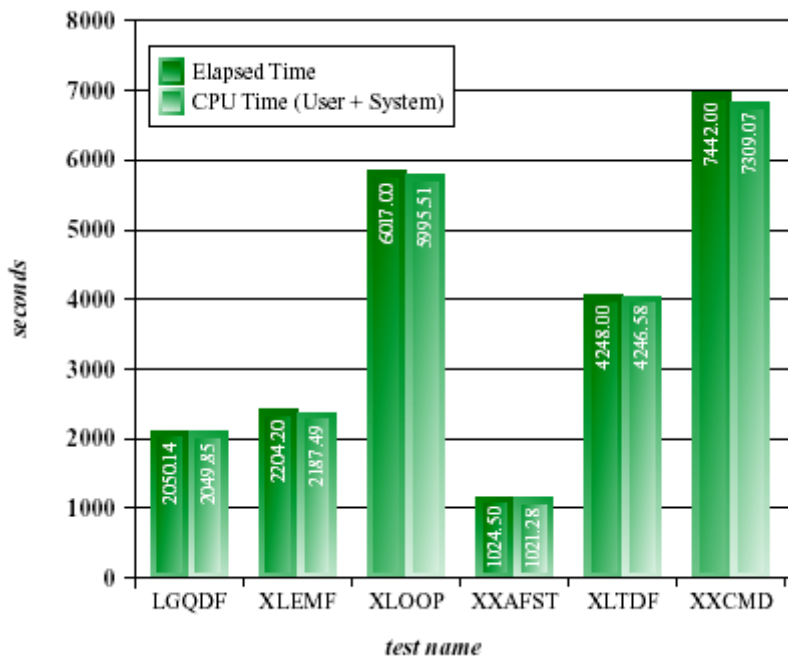


図 1: 標準の Linux ディストリビューション搭載 SGI® Altix® (シングル CPU) での MSC.Nastran™ の解析完了時間。
(出典: MSC Software, 2004 年)

スケーラビリティのベンチマークテスト

シングルプロセッサ・システムは小規模問題ではまずまずの結果を出すことが可能ですが、大規模問題になると完了までの時間は妥当な範囲を超えます。このため、現在のアプリケーション・コードは、拡張可能で複数プロセッサ・ノード上の並列処理をサポートし、より多くの計算リソースを持つコンピューティング・プラットフォームを活用できる構造になっています。ソフトウェアとハードウェアの両者がスケーラブルなシステムにおいては、図 2 に示されるように解析完了までの時間を大幅に短縮することが可能です。

シングル CPU でのテストと同じ MSC.Nastran の解析を、Altix システムに最適化された MSC.Nastran を利用して実行しました。ここでは、最大 8 プロセッサまでの並列処理を評価しています。ここで示すように、使用するプロセッサ数に応じて、スケーラブルな実行性能が得られており、並列処理によって、大幅な処理時間の短縮が可能となっています。

最適化された MSC.Nastran コードを実行する SGI Altix 3000 プラットフォームは、高度にスケーラブルなソリューションとなります。SGI テクノロジと、Itanium 2 プロセッサ用に最適化された MSC のコード、Linux OS 用の SGI の拡張機能の組み合わせによって、メカニカル・エンジニアリングでの解析において高いパフォーマンスが実現されています。

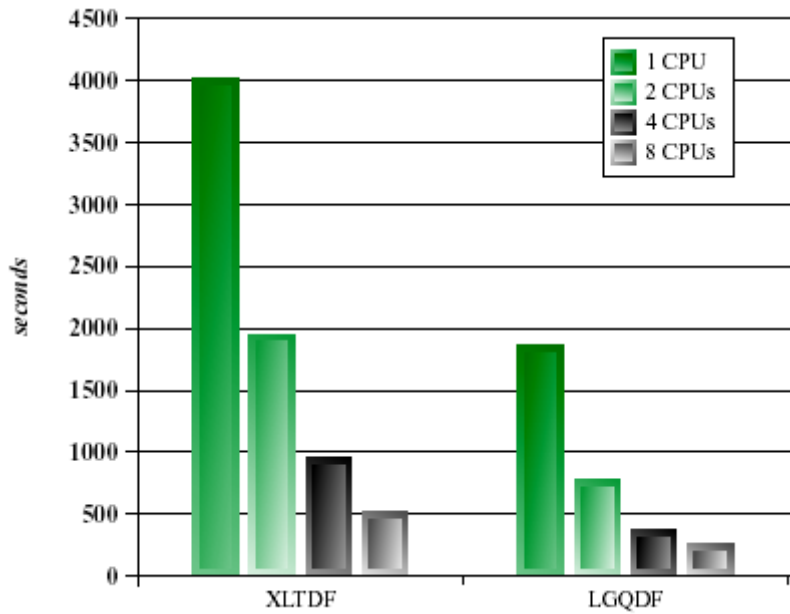


図 2: SGI® Altix® 3000 シリーズ・サーバと Intel® Itanium® 2 プロセッサ用に最適化された MSC.Nastran で MSC.Nastran™ SOL108 を実行したスケーラビリティ・テストの結果 (出典: SGI 社内でのテスト資料、2004)

SGI Altix 3000 サーバ・プラットフォーム

SGI は MCAE 分野で着実な経験を積み重ねてきました。SGI の従来の MCAE ソリューションのためのプラットフォームは、自社で開発・設計を行ってきた 64 ビット MIPS アーキテクチャのプロセッサを搭載し、SGI® IRIX® オペレーティングシステムで稼動する SGI® Origin® サーバ・シリーズでした。SGI Altix 3000 サーバ・プラットフォームは、64 ビットの Itanium 2 プロセッサを搭載し、64 ビットの Linux オペレーティングシステムで稼動する次世代の並列処理システムです。

SGI ALTIX システムのキー・テクノロジー

SGI Altix システムでは、FEA のためのハイパフォーマンスでスケーラブルなコンピューティング・プラットフォームであり、3 つの主要テクノロジーを装備しています。それは、SGI NUMAflex テクノロジー、9 MB L3 キャッシュ搭載 Itanium 2 プロセッサ、64 ビット対応 Linux オペレーティングシステムです。

SGI NUMAflex テクノロジーは SGI Origin システム向けに開発されたテクノロジーで、最大 512 プロセッサまでの共有メモリシステムを可能とし、プロセッサ・アーキテクチャやオペレーティングシステムが物理的にアドレス可能な共有メモリ・シングルアドレス空間への全プロセッサからのアクセスを可能とします。これは、非常に大規模な共有メモリ空間をユーザに提供することになります。この NUMAflex テクノロジーでは、オペレーティングシステムは、プロセッサ数に関係なく、システム全体の管理を行い、非常に大規模なシングルノードのサーバとして、ユーザがシステムを利用することを可能とします。

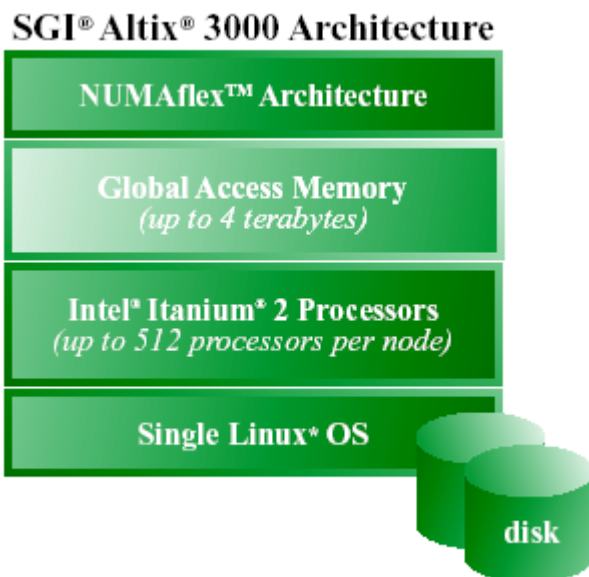


図 3: NUMAflex™ テクノロジーを採用した SGI® Altix® 3000 アーキテクチャ

	InfiniBand Architecture	Myrinet*	Gigabit Ethernet	Altix*
End-to-End Latency	7.6 ms	8 ms	60 ms	1.8 ms
Bandwidth	822 MB/s	250 MB/s	125 MB/s	1165 MB/s

表 2:レイテンシ/バンド幅の比較

SGI では、マルチプロセッサで構成されるノードを複数接続して、大規模なシステムの構築を可能とするために、SGI® NUMalink™ テクノロジーを開発しました。これは、現在のインターコネクト・テクノロジー(表 2)のパフォーマンスを超える大規模なスーパークラスター・システムのための超高速インターコネクト・テクノロジーです。NUMaflex と NUMalink のテクノロジーを使用することにより、数百プロセッサで利用可能な共有メモリを持つ、ハイパフォーマンスで高度にスケーラブルなシステムを SGI は構築することができます。SGI Origin システムで最初に採用されたこの NUMaflex テクノロジーを SGI Altix 3000 プラットフォーム・アーキテクチャにも適用し、新しいコンピューティングプラットフォーム・ファミリでも同じ共有メモリアーキテクチャを実現しました。

64 ビット・コンピューティングは、特に自由度の大きな解析モデルでは、必須となります。これらの問題では非常に大規模なメモリを必要とする大規模データセットが生成され、32 ビット・コンピューティングの限界とされる 4 GB(ギガバイト)を超えるメモリが必要となるケースも多数あります。Itanium 2 プロセッサの 64 ビット・アドレス空間は、アプリケーション・メモリのための大規模なメモリ容量を提供するばかりでなく、OS に対しても、ストレージを使用する代わりにメモリ内にスクラッチ・ディスク領域を確保することを可能にします。(前ページの表 1 に示される、MSC.Nastran ベンチマークが必要とするスクラッチ領域を参照してください。)メモリ上では、読みとり・書き込み操作がディスクよりもはるかに高速で行なわれるため、メモリ内にスクラッチ領域を持つと解析時間は短縮されます。さらに、Itanium 2 プロセッサの 9MB L3 キャッシュは大容量の L3 キャッシュに格納されたデータへの高速アクセスを可能とするため、パフォーマンスをはるかに向上させます。また、解析問題は浮動小数点演算であり、かつ、非常に多くの I/O が行なわれます(前ページの表 1 に示される総 I/O スループットを参照してください)。Itanium 2 プロセッサの大規模なコンピューティング・リソースには、2つの浮動小数点演算エンジン(それぞれクロックサイクルあたり2オペレーションを実行)、複数の 128 ビット・レジスタを持ち、6.4 GB/秒のスループットを実現し、MSC.Nastran ソフトウェアが必要とする高い浮動小数点演算性能を提供します。

産業および科学分野での Linux の普及、進化を続ける Linux OS、そして 64 ビット Linux ディストリビューションの登場によって、Linux は産業用コンピューティング・エコシステムの一翼を担い、今や、64 ビットのコンピューティング・ソリューションにおける主要コンポーネントとなっています。Supercomputing 2003 において、Silicon Graphics のシミュレーション・アプリケーション部門ディレクターである Christian Tanasescu は、自動車産業で導入されたハイパフォーマンス・コンピューティン

グシステムのトップ 20 の調査結果を挙げて、この業界での Linux の急速な普及を明らかにしました。標準ベースのハイパフォーマンスな 64 ビット Itanium 2 プロセッサと信頼性の高い 64 ビットのオープンソース OS の組み合わせは、SGI Altix 3000 プラットフォームなどの並列コンピューティングシステムで優れた価格性能を実現し、自動車産業やその他の MCAE 市場に貢献しています。

SGI LINUX®ツールによるパフォーマンスの向上

SGI NUMAflex テクノロジは、システム内のすべてのプロセッサからすべてのメモリへのアクセスを可能にします。コードを実行するプロセッサとデータを格納するメモリが同じノード内がない場合、メモリアクセスのレイテンシは高くなります。しかし、CPU リソースとメモリ・リソースが同一ノードにあるというプロセッサ・アフィニティが保たれた状態では、最大限のシステム・パフォーマンスが得られます。

標準の 64 ビット Linux ディストリビューションでは、SGI NUMAflex テクノロジの特性を認識しません。しかし、マルチプロセッサ・システムにおける OS のタスクは、プロセッサ・アフィニティを維持できるようにアプリケーション・コード(MSC.Nastran コード)をサポートします。そこで、SGI では Linux OS をツールによって機能強化し、エンドユーザがユーザプログラムと CPU リソースおよびメモリ・リソース間でアフィニティを確立できるようにしています。これらのツールには次のものがあります。

- CPUSET—CPU リソースとメモリ・リソースのパーティション化ソフト
- RUNON—特定 CPU(複数可)でのジョブ実行
- DPLACE—特定 CPU(複数可)へのジョブのバインド

CPUSET は、大規模システムの CPU とメモリをパーティション化します。これにより、同じシステム上で実行中の他ジョブへ特定のプロセスが影響を与えることを制限し、同時に、同一システムで実行中の他プロセスからこれらのプロセスに対する影響も制限します。

RUNON と DPLACE は、ユーザが指定した CPU セットでジョブを実行します。I/O バッファ・キャッシュを除いて、これらのコマンドにおけるデフォルトでのメモリ割り当てポリシーは、まずローカルメモリを使用し、ローカルメモリが使用し尽くされた場合は直近の物理メモリを使用するように設定されています。これらの機能によってプロセッサ・アフィニティを維持します。また、DPLACE コマンドは、指定された CPU へのプロセッサのバインドも行います。これは、I/O インテンシブなプロセスにおいて、OS が I/O 中断後にプロセスを別の CPU に移動しないようにする機能です。OS はプロセスの実行に最適な CPU をおさえない傾向があるためこの機能は重要となります。

MSC.Nastran のパフォーマンスを向上させるために開発されたこれらの機能強化は、SGI Altix 3000 システムと共に販売される Linux 対応 MSC.Nastran ソフトウェア・ディストリビューションでは、すでに組み込まれています。

MSC.Nastran

MSC.Nastran は業界をリードする FEA アプリケーションとして知られており、すでに、SGI Altix 3000 プラットフォーム用に Itanium 2 プロセッサへの移植が完了しています。MSC.Nastran は、SGI Altix 3000 プラットフォームと NUMAflex テクノロジーの様々な機能を余すことなく活かすことができるスケラブルなアプリケーションです。

Intel は MSC.Nastran コードの移植とチューニングにおいて MSC Software をサポートしました。大規模な L3 キャッシュなどの Itanium 2 プロセッサ・アーキテクチャの特長を活かせるようにコードを変更し、Itanium 2 プロセッサのために最適化された Intel の Math Kernel Library を広く活用し、さらに、ループ・アンローリング、ソフトウェア・パイプライニング、データ・プリフェッチングなどの Intel コンパイラの最適化機能を利用して、MSC.Nastran のパフォーマンスを最適化しています。

バイナリ互換の実現

産業分野の設計環境の多くはヘテロジニアスなネットワーク構成をとっています。このような環境では、FEM モデルのプリ/ポストプロセッシング、FEA ソリューション、コンピューティング、およびデータベースならびにファイルサービングが、すべて異なるコンピューティング・プラットフォームで実行され、ファイル形式間でバイナリの互換性が維持できないという問題が生じています。これは、UNIX®ではビッグエンディアン形式が採用されているのに対して、Linux と SGI Altix 3000 プラットフォームではリトルエンディアン形式が採用されていることによるものです。MSC.Nastran はバイナリの互換性を有効にするいくつかのソリューションを提供しており、ユーザはヘテロジニアスなネットワーク環境で容易にファイルを移動できます。

MSC.Nastran が提供する TRANS や RCOUT2 など数種類のツールを用いて MSC.Nastran のファイル形式を変換することも可能です。たとえば、RCOUT2を使用すれば MSC.Nastran OUTPUT2 ファイルをニュートラル形式(ON2)から UNIX ワークステーションのネイティブなバイナリ形式(OP2)に変換できます。しかし、変換オペレーションはワークフローを遅滞させる可能性があるため、バイナリファイルを直接あつかうほうがより望ましい方法です。Intel Itanium 2 コンパイラは「F_UFMTENDIAN 環境変数」をサポートしています。これは、通常使用される MSC.Nastran の Fortran の未フォーマットのバイナリファイル変換用ユーザプログラムとともに使用することができます。また、MSC.Nastran の最新バージョンでは、MSC.Nastran の ASSIGN コマンドによって、入出力でのエンディアン形式の選択が可能となっています。

FFIO による I/O スループットの向上

ソフトウェアとオペレーティングシステムは、NUMAflex テクノロジーを用いてメモリをキャッシュや I/O バッファとして使用して並列 I/O を非同期に実行できるため、I/O パフォーマンスが向上します。さ

らに、MSC.Nastran には FFIO と呼ばれる SGI のハイパフォーマンス I/O インタフェースが組み込まれており、I/O スループットを向上させています。FFIO は、図 4 に示されているようにジョブあたりの I/O パフォーマンスを格段に向上させます。

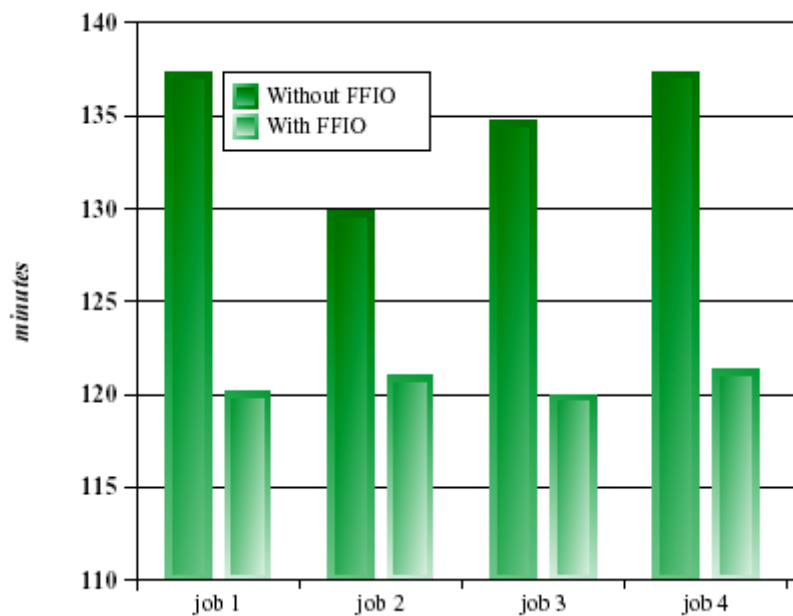



図 4: MSC.Nastran™ の I/O ジョブのスループット(出典: SGI, 2004)

結論

FEA アプリケーションで高生産性を実現するには、より自由度の大きな解析モデルの処理が要求される今日の複雑なFEA問題を迅速に処理することができる、ハイパフォーマンスでスケーラブルなコンピューティング・ソリューションが必要です。Itanium 2 プロセッサ・ベースの SGI Altix 3000 並列処理プラットフォームと MSC.Nastran FEA ソフトウェアの組み合わせは、解析やその他の MCAE アプリケーションのためのハイパフォーマンスでコスト効果の高いコンピューティング・ソリューションであり、高度に複雑な解析問題を迅速に解決します。

日本SGI株式会社

〒150-6031 東京都渋谷区恵比寿4-20-3 恵比寿ガーデンプレイスタワー31階

 TEL : 0120-161-086 FAX : 0120-161-087 <http://www.sgi.co.jp>



インテル株式会社

〒300-2635 茨城県つくば市東光台 5-6

<http://www.intel.co.jp/>

©2005 SGI Japan, Ltd. SGI、Altix、SGIのロゴマーク、およびSGIのキューブは日本SGIの登録商標です。NUMAflexは米Silicon Graphics, Inc.の商標です。Intel、Intel ロゴ、Itanium は、アメリカ合衆国およびその他の国におけるIntel Corporation またはその子会社の商標または登録商標です。本誌に掲載されているその他の商標については商標の所有者に所有権が属していません。(01/2005)