

HPC デベロッパー・フォーラム

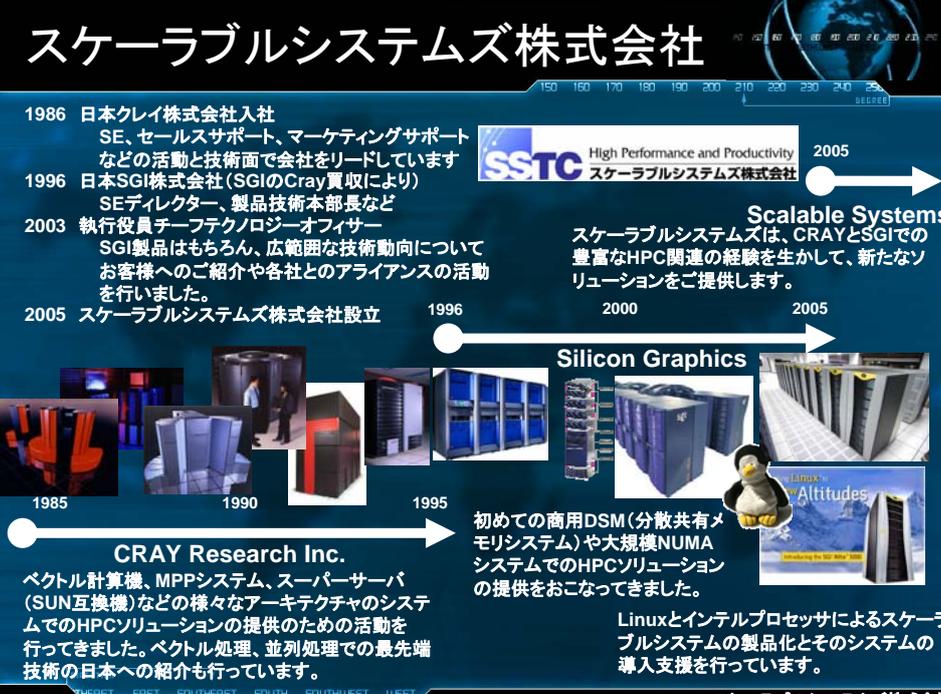


High Performance Computing 過去と現在、そして未来へ

スケーラブルシステムズ株式会社
代表取締役 戸室 隆彦

DIRECTION
NORTH EAST SOUTHEAST SOUTH SOUTHWEST WEST

スケーラブルシステムズ株式会社



1985 **CRAY Research Inc.**
ベクトル計算機、MPPシステム、スーパーサーバ (SUN互換機) などの様々なアーキテクチャのシステムでのHPCソリューションの提供のための活動を行ってきました。ベクトル処理、並列処理での最先端技術の日本への紹介も行っています。

1986 日本クレイ株式会社入社
SE、セールスサポート、マーケティングサポートなどの活動と技術面で会社をリードしています

1996 日本SGI株式会社 (SGIのCray買収により)
SEディレクター、製品技術本部長など

2003 執行役員チーフテクノロジーオフィサー
SGI製品はもちろん、広範囲な技術動向についてお客様へのご紹介や各社とのアライアンスの活動を行いました。

2005 スケーラブルシステムズ株式会社設立

1995 初めての商用DSM (分散共有メモリシステム) や大規模NUMAシステムでのHPCソリューションの提供をおこなってきました。

2000 **Silicon Graphics**

2005 **Scalable Systems**
スケーラブルシステムズは、CRAYとSGIでの豊富なHPC関連の経験を生かして、新たなソリューションをご提供します。

2005 **LinuxとIntelプロセッサによるスケーラブルシステムの製品化とそのシステムの導入支援を行っています。**

SSITC High Performance and Productivity
スケーラブルシステムズ株式会社

スケーラブルシステムズ株式会社

HPC デベロッパー・フォーラム

1. はじめに
 2. HPCシステムの歴史
 3. HPCシステムの課題
 - ソフトウェア
 - ハードウェア
- マイクロプロセッサ
- まとめとして



DIRECTION
NORTH EAST SOUTHEAST SOUTH SOUTHWEST WEST

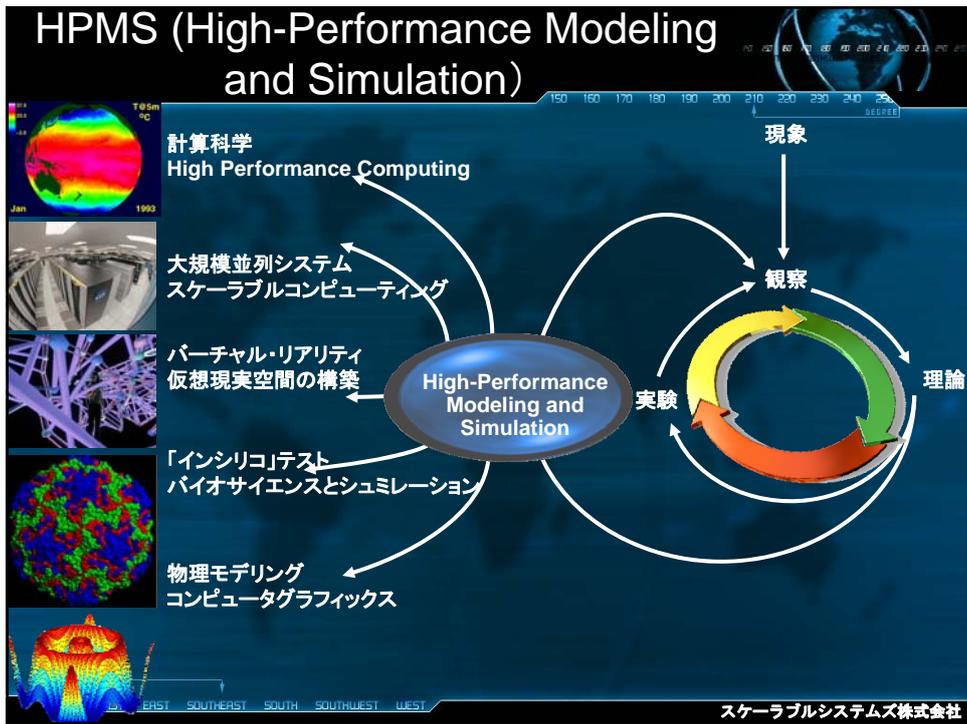
スケーラブルシステムズ株式会社

HPCシステム

- HPCからHPMS (High-Performance Modeling and Simulation)
 - 計算システム+ストレージ+可視化の統合システム
 - High Performance と High Productivity
- Capability .vs. Capacity
 - Capability → 単一ジョブの高速処理
 - Capacity → 複数ジョブの多重処理
- ハイエンドコンピューティングに関する課題
 - プログラミングモデル(Programming Productivity - Safety, Portability, Performance, Integrationなど)
 - 仮想化、IO、OS、APIなど様々な課題
- マイクロプロセッサの動向の変化

DIRECTION
NORTH EAST SOUTHEAST SOUTH SOUTHWEST WEST

スケーラブルシステムズ株式会社



このスライドは誰が?

Technical Computing Reduced Time To Insight

Computational Modeling

Real-World Data

Persistent Distributed Storage

Workflow Data Mining & Algorithms

Interpretation Insight

スケールシステムズ株式会社

HPCの歴史



150 160 170 180 190 200 210 220 230 240 250
DEGREE

Episode I The Phantom Menace



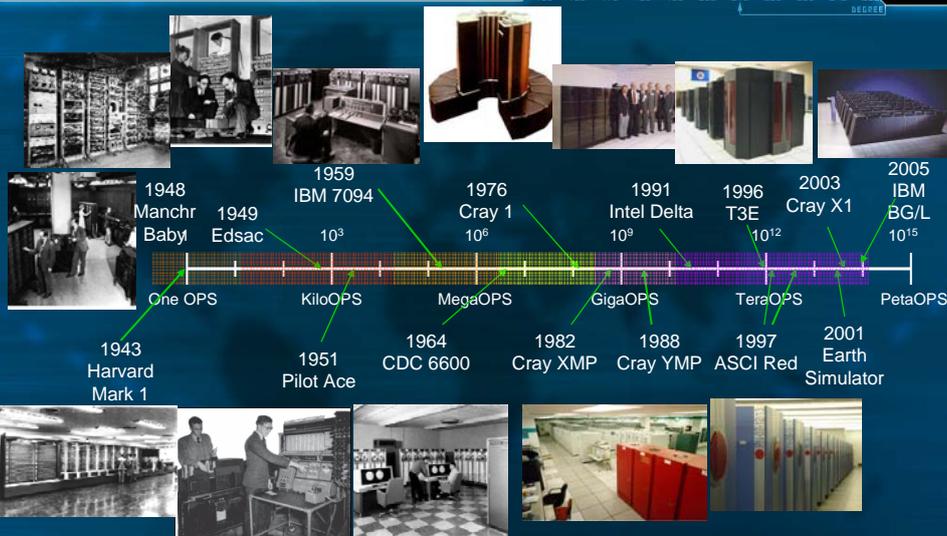
DIRECTION
NORTH EAST SOUTHEAST SOUTH SOUTHWEST WEST

スケーラブルシステムズ株式会社

過去60年間の進化



150 160 170 180 190 200 210 220 230 240 250
DEGREE



Scalar to super scalar to vector to SMP to DMP to massively parallel to hybrid designs

DIRECTION
NORTH EAST SOUTHEAST SOUTH SOUTHWEST WEST

スケーラブルシステムズ株式会社

様々なアーキテクチャのシステム

- Parallel Vector Processors (PVP)
 - NEC Earth Simulator, SX-6
 - Cray- 1, 2, XMP, YMP, C90, T90, X1
 - Fujitsu 5000 series
- Massively Parallel Processors (MPP)
 - Intel Touchstone Delta & Paragon
 - TMC CM-5
 - IBM SP-2 & 3, Blue Gene/Light
 - Cray T3D, T3E, Red Storm/Strider
- Distributed Shared Memory (DSM)
 - SGI Origin
 - HP Superdome
- Single Instruction stream Single Data stream (SIMD)
 - Goodyear MPP, MasPar 1 & 2, TMC CM-2
- Commodity Clusters
 - Beowulf-class PC/Linux clusters
 - Constellations
 - HP Compaq SC, Linux NetworX MCR



DIRECTION
NORTH EAST SOUTHEAST SOUTH SOUTHWEST WEST

スケールラブルシステムズ株式会社

並列計算機ベンダー

- | | | |
|---|---|--|
| <ul style="list-style-type: none"> • ACRI • Alliant • American Supercomputer • Ametek • Applied Dynamics • Astronautics • BBN • CDC • Cogent • Convex > HP • Cray Computer • Cray Research > SGI > Cray • Culler-Harris • Culler Scientific • Cydrome • Dana/Ardent/Stellar/Stardent • Denelcor • Encore • Eleksi | <ul style="list-style-type: none"> • ETA Systems • Evans and Sutherland Computer • Exa • Flexible • Floating Point Systems • Galaxy YH-1 • Goodyear Aerospace MPP • Gould NPL • Guiltech • Intel Scientific Computers • International Parallel Machines • Kendall Square Research • Key Computer Laboratories searching again • MasPar • Meiko • Multiflow • Myrias • Numerix | <ul style="list-style-type: none"> • Pixar • Parsytec • nCube • Prisma • Pyramid • Ridge • Saxpy • Scientific Computer Systems (SCS) • Soviet Supercomputers • Supertek • Supercomputer Systems • Suprenum • Tera > Cray Company • Thinking Machines • Vitesse Electronics • Wavetracer |
|---|---|--|



Requiescat In Pace

DIRECTION
NORTH EAST SOUTHEAST SOUTH SOUTHWEST WEST

スケールラブルシステムズ株式会社

Cray システム

- Cray-1 (1977)
 - 250 MFLOPS
 - 80 MHz
 - 1 MWord (64-bit)
- PC 8088 (1979)
 - 5 MHz
 - 1 MB RAM
- Modern PC (Pentium 4)
 - 3.2 GHz (Dual Core)
 - 12.8 GFLOPS
 - 4 GB RAM



<http://ed-thelen.org/comp-hist/CRAY-1-HardRefMan/CRAY-1-HRM.html>

スケラブルシステムズ株式会社



Cray 1 #6 from LLNL.
@ The Computer
Museum History
Center, Moffett Field



Cray X-MP/4



Cray-2

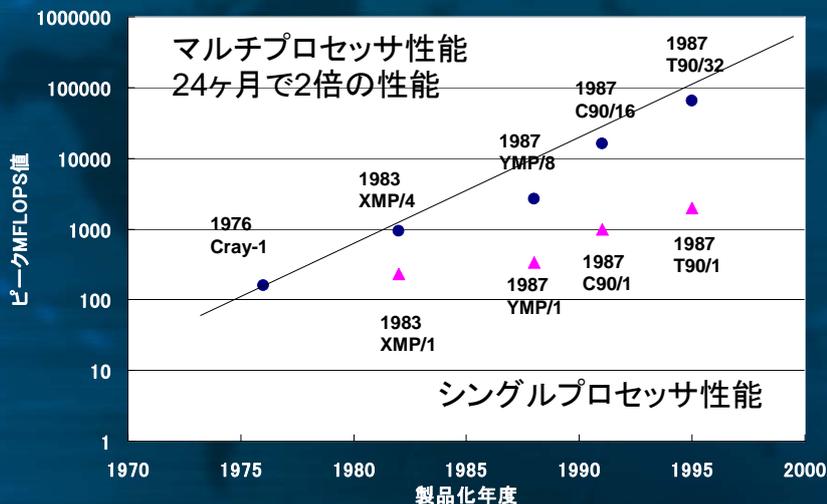


Cray-3 1995 processor
500 MHz
32 modules 1K GaAs IC
8 プロセッサ

DIRECTION
NORTH EAST SOUTHEAST SOUTH SOUTHWEST WEST

スケラブルシステムズ株式会社

Crayシステム:ピーク性能



スケラブルシステムズ株式会社

シングルプロセッサ性能: Linpack

UNIT = 10**6 TIME/(1/3 100**3 + 100**2)

$\frac{2}{3} N^3$ $2N^2$ ops time

Facility	TIME N=100	UNIT micro-secs.	Computer	Type	Compiler
NCAR	14.0	0.14	CRAY-1	S	CFT, Assembly BLAS
LASL	4.64	0.43	CDC 7600	S	FTN, Assembly BLAS
NCAR	3.54	0.56	CRAY-1	S	CFT
LASL	3.27	0.61	CDC 7600	S	FTN
Argonne	2.31	0.86	IBM 370/195	D	H
NCAR	1.91	1.05	CDC 7600	S	Local
Argonne	1.77	1.33	IBM 3033	D	H
NASA Langley	1.40	1.42	CDC Cyber 175	S	FTN
U. Ill. Urbana	1.34	1.47	CDC Cyber 175	S	Ext. 4.6
LLL	1.04	1.61	CDC 7600	S	CHAT, No optimize
SLAC	1.19	1.69	IBM 370/168	D	H Ext., Fast mult.
Michigan	1.09	1.84	Amdahl 470/V6	D	H
Toronto	1.77	2.59	IBM 370/165	D	H Ext., Fast mult.
Northwestern	1.77	4.20	CDC 6600	S	FTN
Texas	3.54	5.63	CDC 6600	S	RUN
China Lake	3.54	5.69	Univac 1110	S	V
Yale	7.53	7.53	DEC KL-20	S	F20
Bell Labs	10.1	10.1	Honeywell 6080	S	Y
Wisconsin	10.1	10.1	Univac 1110	S	V
Iowa State	10.2	10.2	Intel AS/5 mod3	D	H
U. Ill. Chicago	11.9	11.9	IBM 370/158	D	G1
Purdue	16.6	16.6	CDC 6500	S	FUN
U. C. San Diego	38.2	38.2	Burroughs 6700	S	H
Yale	49.9	49.9	DEC RA-10	S	F40

* TIME(100) = (100/75)**3 SGEFA(75) + (100/75)**2 SGESL(75)

スケラブルシステムズ株式会社

ベクトル計算機の性能

Q: なぜ、ベクトル計算機の性能が、マイクロプロセッサの性能のように向上しなかったのでしょうか？

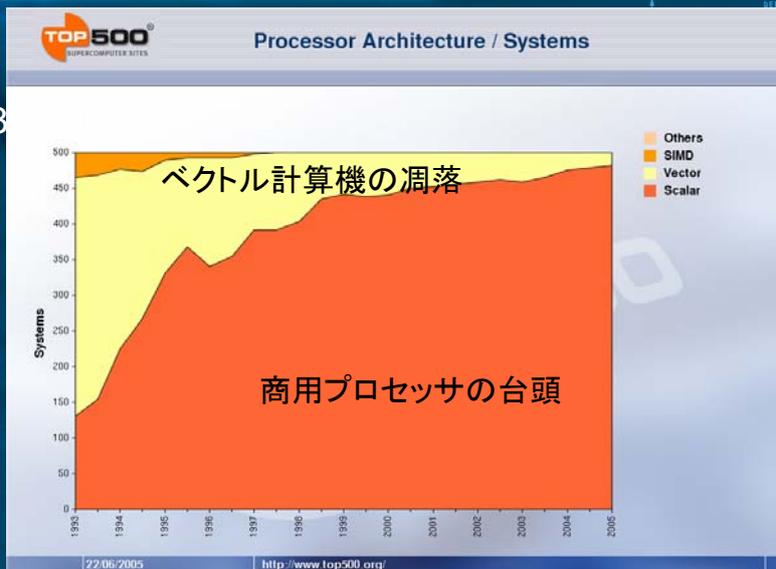
A: ベクトル計算機は、グローバル共有メモリに対する高い接続性能にその性能が依存していたために、このメモリ間接続の性能向上がボトルネックとなりました。

例: DRAMメモリの性能と仕様

1979: 標準DRAM	1999: 200 MHz SDRAM	1979→1999
16K bit	256 Mbit	X 16000
1-bit wide interface	16-bit wide interface	X 640
5 Mb/s uniform access BW	3200 Mb/s uniform access BW	X 500
2 Mb/s random access BW	1000 Mb/s random access BW	X 25

The Pahntom Menace

1993



HPCの歴史



150 160 170 180 190 200 210 220 230 240 250
DEGREE

Episode II Attack of the Clones

DIRECTION
NORTHEAST EAST SOUTHEAST SOUTH SOUTHWEST WEST

スケラブルシステムズ株式会社

イノベーションのジレンマ



150 160 170 180 190 200 210 220 230 240 250
DEGREE



増補
改訂
版
イノベーションのジレンマ

- 顧客の意見に熱心に耳を傾け、新技術への投資を積極的に行い、常に高品質の製品やサービスを提供している業界トップの優良企業が、その優れた経営のために失敗を招き、トップの地位を失ってしまうという、この逆説的なコンセプトは、ハーバード・ビジネス・スクールのクレイトン・クリステンセン教授が、同名の著書によって明らかにしたことで、有名になったものです。マーケティングにおいて最も基本とされる、顧客の意見に耳を傾け、顧客の求める価値提供を行なうことが、破壊的なイノベーションの前では逆にマイナス要素にさえなるというこの逆説は、変化の時代といえる現代においては新たなマーケティング課題であるといえます。

DIRECTION
NORTHEAST EAST SOUTHEAST SOUTH SOUTHWEST WEST

スケラブルシステムズ株式会社

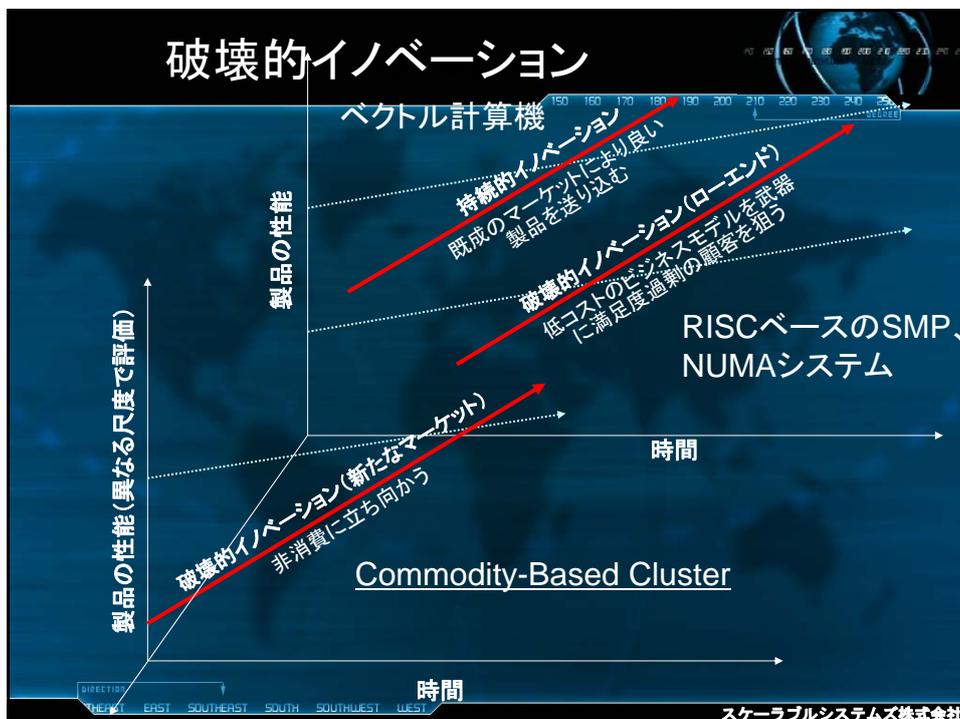
イノベーションのジレンマ



- クレイトン・クリステンセンの「イノベーションのジレンマ」
- 持続的イノベーションと破壊的イノベーションによる市場の動向を分析
- 持続的イノベーション
 - 技術革新が顧客の求める性能向上軸に沿っている
- 破壊的イノベーション
 - 既存顧客が求める性能とは異なる軸の性能(特性)

スケラブルシステムズ株式会社

破壊的イノベーション



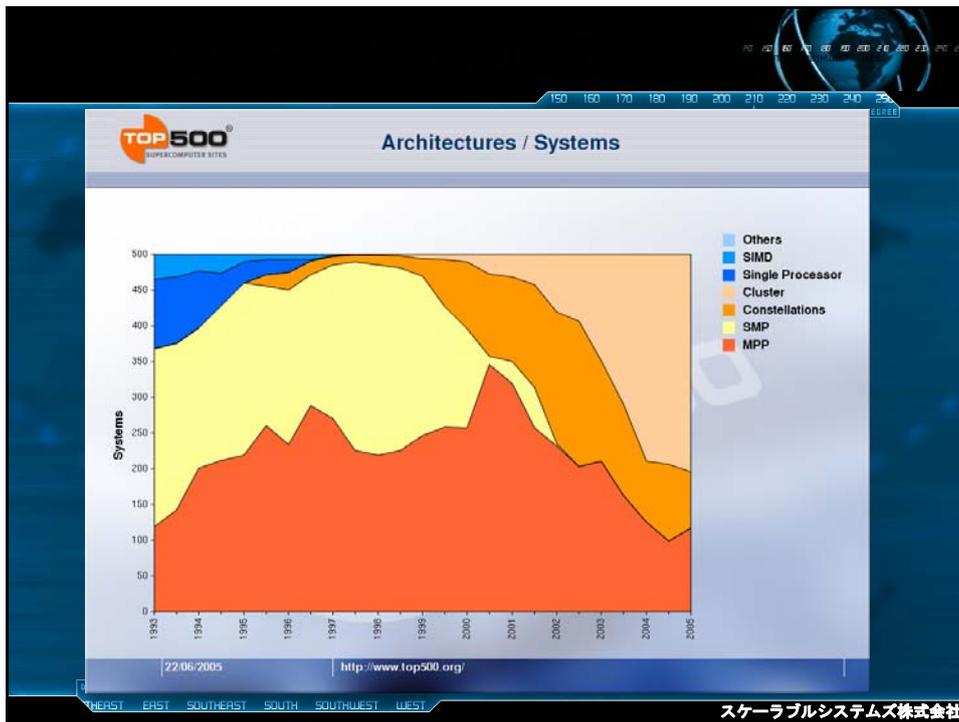
スケラブルシステムズ株式会社

Beowulf プロジェクト

Beowulf プロジェクト

<ul style="list-style-type: none"> ◆ Wiglaf - 1994 ◆ 16 Intel 80486 100 MHz ◆ VESA Local bus ◆ 256 Mbytes memory ◆ 6.4 Gbytes of disk ◆ Dual 10 base-T Ethernet ◆ 72 Mflops sustained ◆ \$40K 	<ul style="list-style-type: none"> ◆ Hrothgar - 1995 ◆ 16 Intel Pentium100 MHz ◆ PCI ◆ 1 Gbyte memory ◆ 6.4 Gbytes of disk ◆ 100 base-T Fast Ethernet (hub) ◆ 240 Mflops sustained ◆ \$46K 	<ul style="list-style-type: none"> ◆ Hyglac-1996 (Caltech) ◆ 16 Pentium Pro 200 MHz ◆ PCI ◆ 2 Gbytes memory ◆ 49.6 Gbytes of disk ◆ 100 base-T Fast Ethernet (switch) ◆ 1.25 Gflops sustained ◆ \$50K
---	--	---

スケラブルシステムズ株式会社



ベクトル計算機の逆襲

Episode V
The Empire Strikes Back

Sputnik: October 4, 1957

スケーラブルシステムズ株式会社

ベクトル計算機の逆襲

- 2002
 - 地球シミュレータ
 - コンピュータにおけるスプートニクショック
- 5,120 (640 8-way nodes) 500 MHz NEC
 - 8 GFLOPS per CPU (41 TFLOPS total)
 - 2 GB Memory per CPU (10 TB total)
 - 20 kVA power consumption per node



DIRECTION
NORTH EAST SOUTHEAST SOUTH SOUTHWEST WEST

スケーラブルシステムズ株式会社

課題

Episode III
Revenge of the sith

DIRECTION
NORTH EAST SOUTHEAST SOUTH SOUTHWEST WEST

スケーラブルシステムズ株式会社

HPCの現状



スケラブルシステムズ株式会社

HPCマーケット

- HPCマーケットでのHPCシステム構築及び製品は、次の3つのセグメントに分かれている
 - 一般商用システム (Commodity-based systems)
 - 一般のクラスタシステム (Dell HPCCなど)
 - 付加価値システム (Value-based systems)
 - 多くのSMPやNUMAシステム (SGI Altixなど)
 - 特定目的システム (Purpose-built systems)
 - アプリケーションと解析対象に合わせたシステム設計 (IBM BlueGene/Lなど)
- IDCなどのレポートでも、一般商用システムのHPCマーケットでの導入がもっともその成長が大きい
 - 付加価値システムの課題 (一般商用システムとの競争に対する対応、もしくは、新たな分野の開拓→ペタスケールコンピューティング)
 - HPCSプログラムは、この付加価値システムのベンダーにとっても、生き残りを賭けた戦い? (2006、July)

スケラブルシステムズ株式会社

HPCの二極分化

Going UP

'Peta-Scale'
コンピューティング

- 複雑なシステム構成
- 新しいプログラミングAPIの提案
- アプリケーション開発

Going DOWN

'Commodity'
コンピューティング

- 商用HW/SW
- オープンソース
- パーソナルクラスタ
- 商用アプリケーション
- マルチスレッド

DIRECTION
NORTH EAST SOUTHEAST SOUTH SOUTHWEST WEST

スケラブルシステムズ株式会社

HPCシステムの動向

国家プロジェクトと商用製品のギャップの拡大

Going UP

'Peta-Scale'
コンピューティング

- 複雑なシステム構成
- 新しいプログラミングAPIの提案
- アプリケーション開発

Peta-Scaleコンピューティングに求められる基本技術と現在のHPCの主要マーケットでの要求はあまりにも差が大きい
→HPCSシステムは、各社との'Commodity'のマイクロプロセッサではなく、独自のプロセッサを開発中

ハードウェアは、'Commodity'なものを利用して、SWの改善、サポート、利用技術のサポートが今後の主要マーケットでの成功の鍵となる

Going DOWN

'Commodity'
コンピューティング

- 商用HW/SW
- オープンソース
- パーソナルクラスタ
- 商用アプリケーション
- マルチスレッド

DIRECTION
NORTH EAST SOUTHEAST SOUTH SOUTHWEST WEST

スケラブルシステムズ株式会社

HPCシステムの現状分析

Good News !

“HPCシステムにおける問題は、たった2つ
だけである”

ソフトウェアとハードウェア

ソフトウェア : The Law of More.....

- システム規模とその複雑さの急速な増加・拡大
- ソフトウェアの準備が出来た時点でハードウェアは既に陳腐化し、次のシステムの導入の検討が進む...

ハードウェア : Moore's Law (ムーアの法則)

- 消費電力の問題のため、プロセッサの動作クロックを今までのペースで上げることは困難
- プロセッサとメモリの性能差の拡大によるCPUサイクルとのギャップ
- ピーク性能と実効性能のギャップの拡大

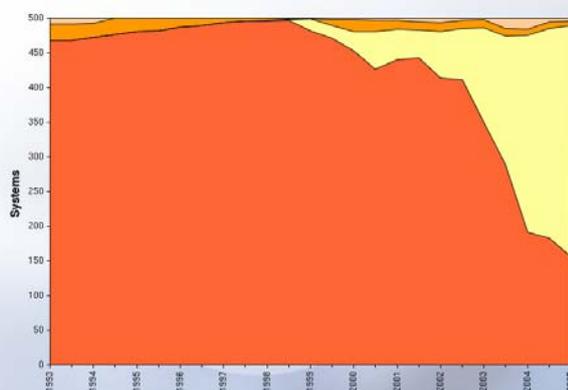
DIRECTION
HERST EAST SOUTHEAST SOUTH SOUTHWEST WEST

スケールシステムズ株式会社

HPCにおけるLinuxの活用

TOP500[®]
SUPERCOMPUTER SITES

Operating System / Systems



22.06.2005 | <http://www.top500.org/>

DIRECTION
HERST EAST SOUTHEAST SOUTH SOUTHWEST WEST

スケールシステムズ株式会社

ソフトウェア: The Law of More...

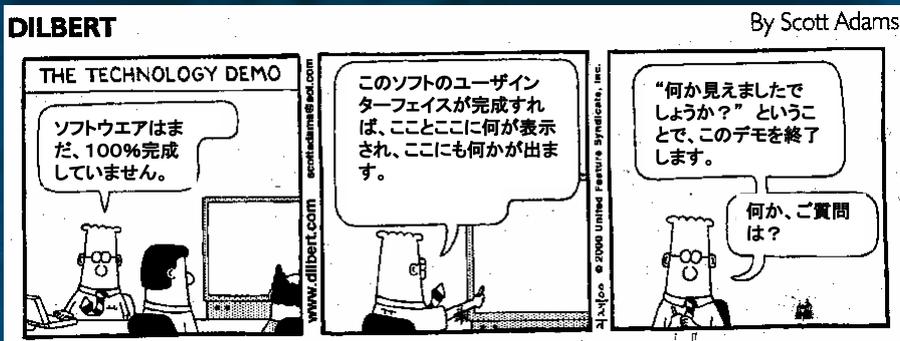
- 研究者は、より多くの時間 (More Time) をソフトウェアの開発のために必要としている
- 問題はより複雑 (More Complex) になり、そして、より多くのプロセッサ (More Processors) を利用して処理を行うには、より多くの困難 (More Difficult) が伴います

ソフトウェアに関する問題については、この資料では、詳しくは解析をしていません。

DIRECTION
NORTH EAST SOUTHEAST SOUTH SOUTHWEST WEST

スケーラブルシステムズ株式会社

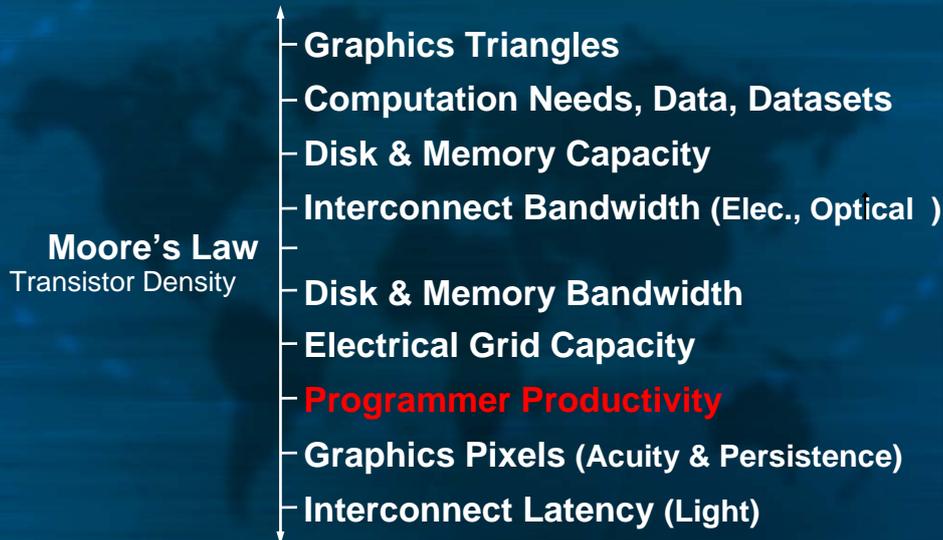
テクノロジーデモ



DIRECTION
NORTH EAST SOUTHEAST SOUTH SOUTHWEST WEST

スケーラブルシステムズ株式会社

技術の進歩の度合いの評価



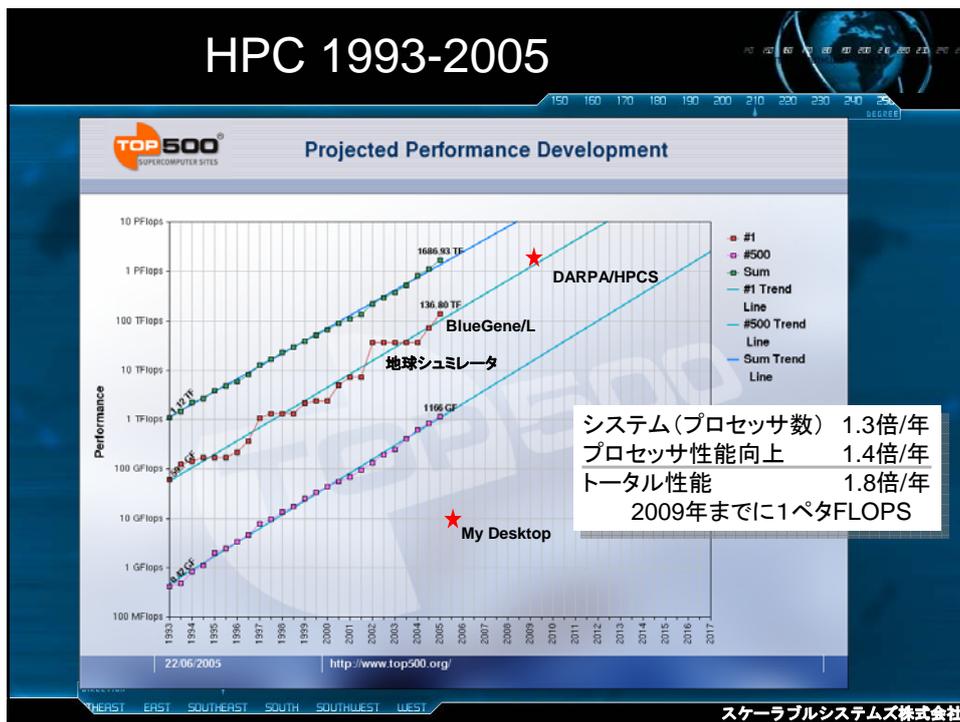
スケラブルシステムズ株式会社

ソフトウェア: The Law of More...

- 一般の商用製品を活用したクラスタソリューションでは、「Capacity」の実現は容易であるが、「Capability」の実現については依然として課題が多い
 - コストパフォーマンスの高いシステムの構築は可能だとしても、コストプロダクティビティの高いシステムの構築も課題
- 数百～数千プロセッサ構成のシステムの利用技術と解析対象の検討
 - 小規模、中規模問題の高速処理への対応
 - ソフトウェア開発の生産性
- 数プロセッサ～数十プロセッサをより簡便に、容易に利用できる技術
 - シングルプロセッサ、シングルスレッドを利用するのと同じように.....

スケラブルシステムズ株式会社

HPC 1993-2005



ペタスケールシステムの構築

- 今まで以上のアプリケーションのスケラビリティ
 - ~100,000プロセッサでのスケラビリティ(ピーク)
 - ~1,000プロセッサ(通常運用での利用?)
- 可用性
 - プロセッサ障害でのリカバリ(耐障害性やチェックポイント)
- 省電力
 - 消費電力あたりの性能を最大にし、高性能で低消費電力のシステム構築

ペタスケールシステムの構築

現在のテラ
FLOPS級の問題



‘複雑さ’の壁

ペタスケールシステム
の構築のための兆戦

Source: ORNL

- ソフトウェア(アプリケーション、OS、プログラミングAPIなど)の課題の克服が課題
- システムの複雑さと生産性

例:

Linpack Benchmark

- オリジナルベンチマークプログラム ~100ライン
- HPL ベンチマークプログラム ~10,000ライン (x100より複雑?)

NORTH EAST SOUTHEAST SOUTH SOUTHWEST WEST

スケラブルシステムズ株式会社

HPC Challenge Benchmarks

- <http://icl.cs.utk.edu/hpcc/>
- Benchmarks:
 - HPL
 - DGEMM
 - STREAM
 - PTRANS
 - RandomAccess
 - FFTC
 - Comm. bandwidth & latency

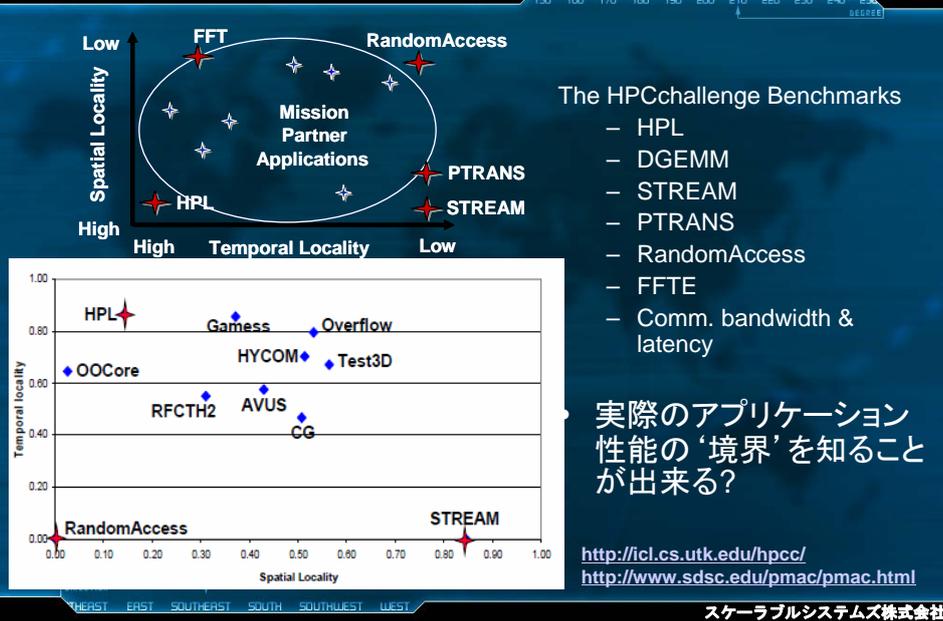


Linpackベンチマークが実際のアプリケーションの性能評価に全く役に立たないことは、常に言われていることです。IBMのBlueGene/Lや地球シミュレータ、NASAのColumbiaは、常にアプリケーションの性能で高い性能を示して、高い評価を受けています。

NORTH EAST SOUTHEAST SOUTH SOUTHWEST WEST

スケラブルシステムズ株式会社

システムの性能評価の問題



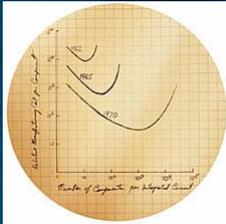
計算機の性能向上

- 動作周波数(クロック)の向上
 - 過去12年間で、Pentiumプロセッサの動作周波数は、60 MHz から 3,800 MHz にまでアップ
 - 現在までの高性能化の約80% はクロック周波数の向上によるもの

ハードウェアの問題 Moore's Law: ムーアの法則



Dr. Gordon Moore
(co-founder of Intel)



- インテルの共同設立者の1人である Gordon Moore 博士が、1965年4月19日号の「**Electronics**」誌に投稿した、「一定面積に集積されるトランジスタの数は12か月で倍増し、それに伴いトランジスタの動作速度が向上する」という予測（その後、1975年に Moore 博士はチップの複雑化を考慮してトランジスタ数の倍増ペースを24か月に修正）
- また、一般にはあまり知られていないがテクノロジーの進歩とともに製造コストが劇的に下落することも予測（左図）

指数関数的成長は永遠には続かない。しかしその永遠を先延ばしにすることはできる [英語: PDF 形式 2MB]

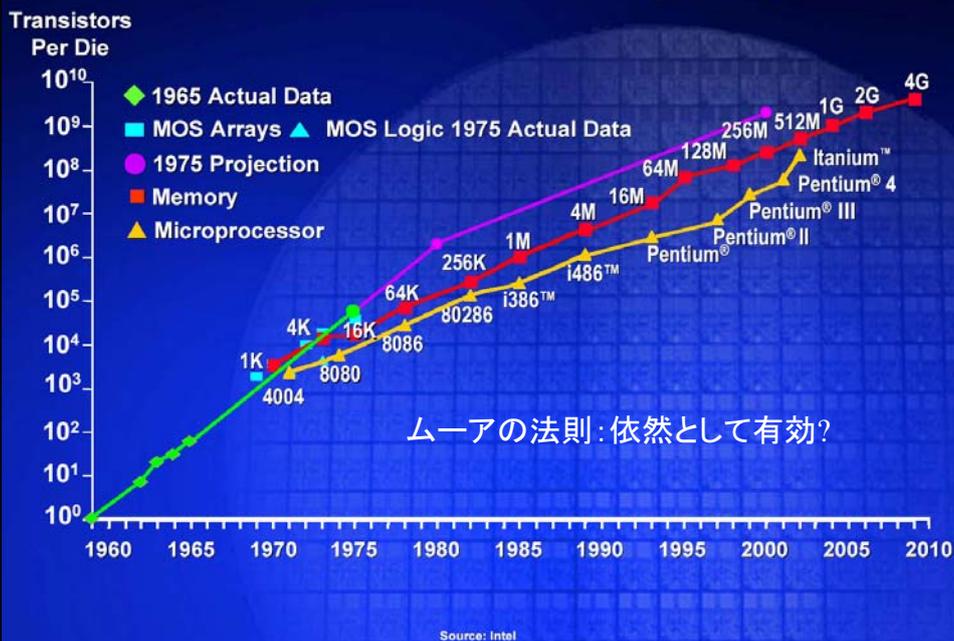
Gordon E. Moore, 2003年2月10日、ISSCC (International Solid State Circuits Conference) でのプレゼンテーション

<http://www.intel.co.jp/developer/technology/silicon/mooreslaw/index.htm>

DIRECTOR EAST SOUTHEAST SOUTH SOUTHWEST WEST

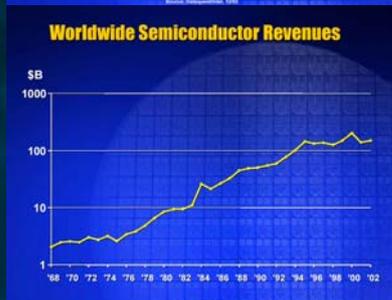
スケラブルシステムズ株式会社

Integrated Circuit Complexity



ムーアの法則=ビジネスモデル?

- Intelは「ムーアの法則」を元に会社全体の戦略が動いている。これまでも、そしてこれからもだ。ご存じの方も多いたろうが、ムーアの法則とは「18カ月ごとに半導体の集積度は2倍になる(あるいはトランジスタ数は2倍になる)」という、Intel名誉会長のGordon Moore氏が半導体黎明期に示したビジョンである。Intelはこのビジョンを現実の製品へと結びつける技術の研究開発を続けることで、今日のプロセッサ業界における地位を確立してきた。
- 翻って言うならば、ムーアの法則の破綻は、Intelの現在のコアコンピタンスを揺るがす大事件へと発展してしまう。Intelが今後もさらに発展するためには、ムーアの法則を生かし続けなければならない。いや、ムーアの法則を維持するだけでなく、さらにその法則を他の製品にも適用していかなければ、大きな成長が望めなくなってくる。
- インテル社 Pat Gelsinger氏



スケラブルシステムズ株式会社

性能向上の源泉は?

ハードウェアデバイス技術の進歩

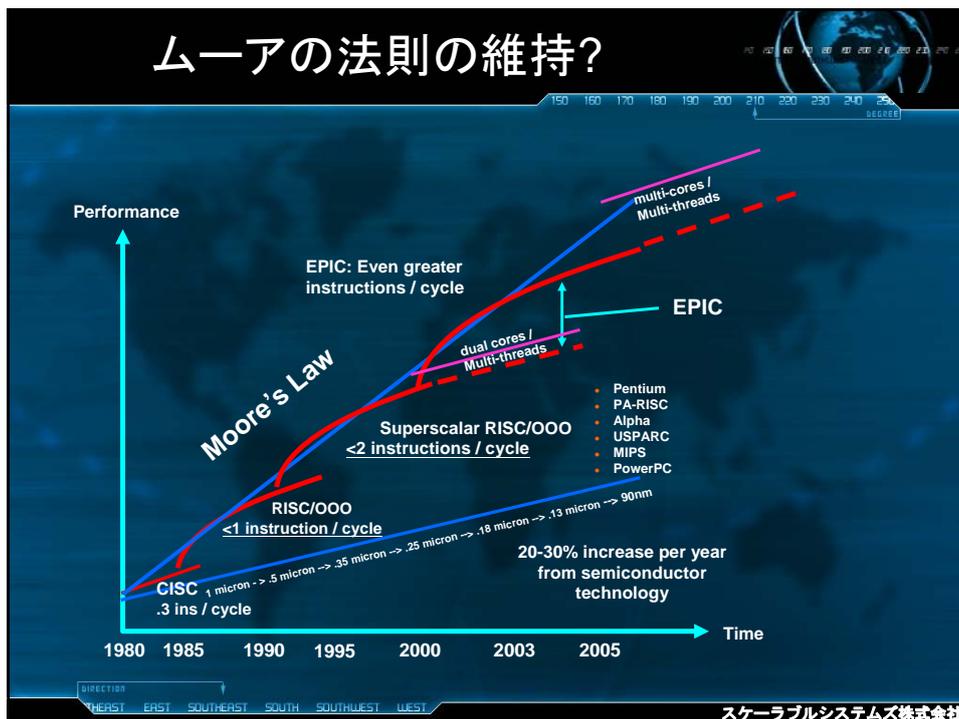
- ロジック回路のスイッチング速度の向上とデバイス密度
- メモリサイズの拡大とアクセス速度の向上
- 通信性能(バンド幅とレイテンシの向上)

コンピュータ・アーキテクチャ

- 命令発行・実行速度の向上
 - パイプライン化
 - 分岐予測
 - キャッシュ
 - Out-of-order など
- 並列性
 - 1サイクルでの命令実行数
 - 命令レベルでの並列性 (ILP)
 - ベクトル処理
 - プロセッサあたりコア数
 - ノードあたりのプロセッサ数
 - システムあたりのノード数

スケラブルシステムズ株式会社

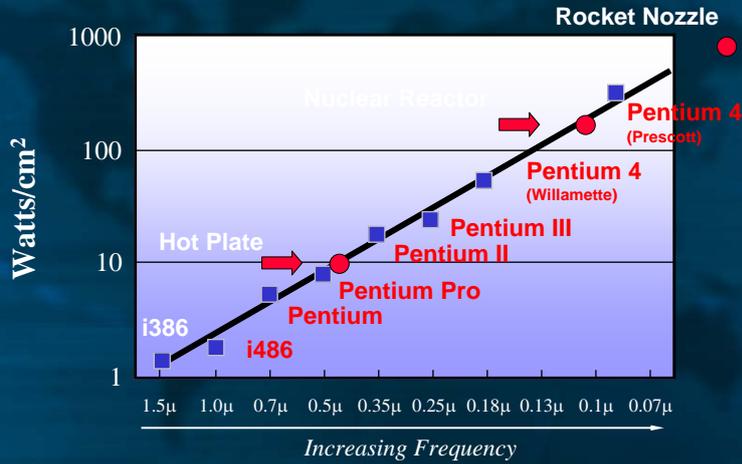
ムーアの法則の維持?



GHz競争

- 2000年に開催されたIEEE国際電子デバイス会議2000(2000 IEEE International Electron Devices Meeting: IEDM)において、インテル社は4億個以上のトランジスタを集積した、10GHz駆動のプロセッサが2005年までに実現可能だと発表しました。
 - 実際には、インテル社の最速プロセッサは、6ヶ月前に発表された3.8GHz (Intel Pentium 4)となっています。
- Prescottプロセッサの6xxシリーズ発表に際して、インテル社は、“adding value beyond GHz” のコメントを出しています。それ以降、インテル社の多くのドキュメントやプレスリリースは、この“adding value beyond GHz” についての内容を含んでいます。

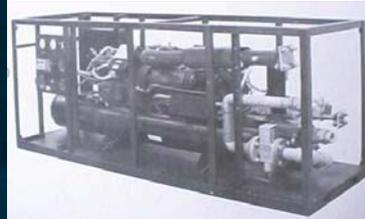
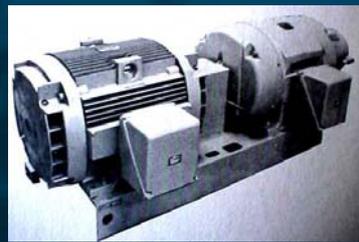
発熱の問題が深刻化



Bob Colwell氏の資料より抜粋

スケラブルシステムズ株式会社

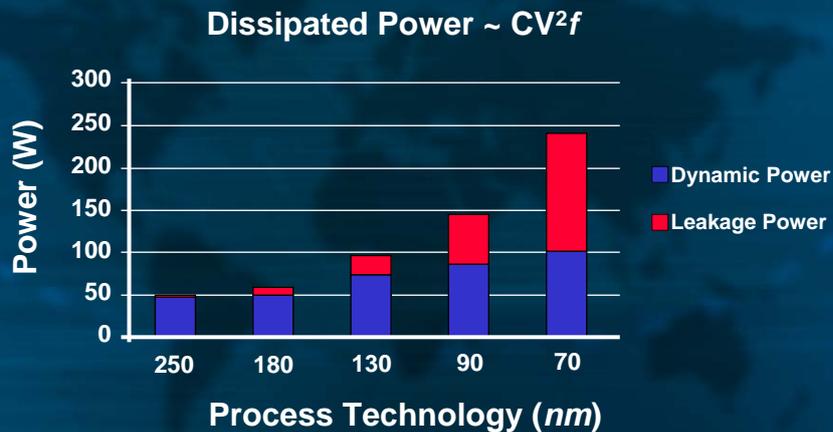
Cray-1:150 Kw. MG 及び熱交換器



DIRECTION
NORTH EAST SOUTHEAST SOUTH SOUTHWEST WEST

スケラブルシステムズ株式会社

消費電力におけるリーク電流の影響

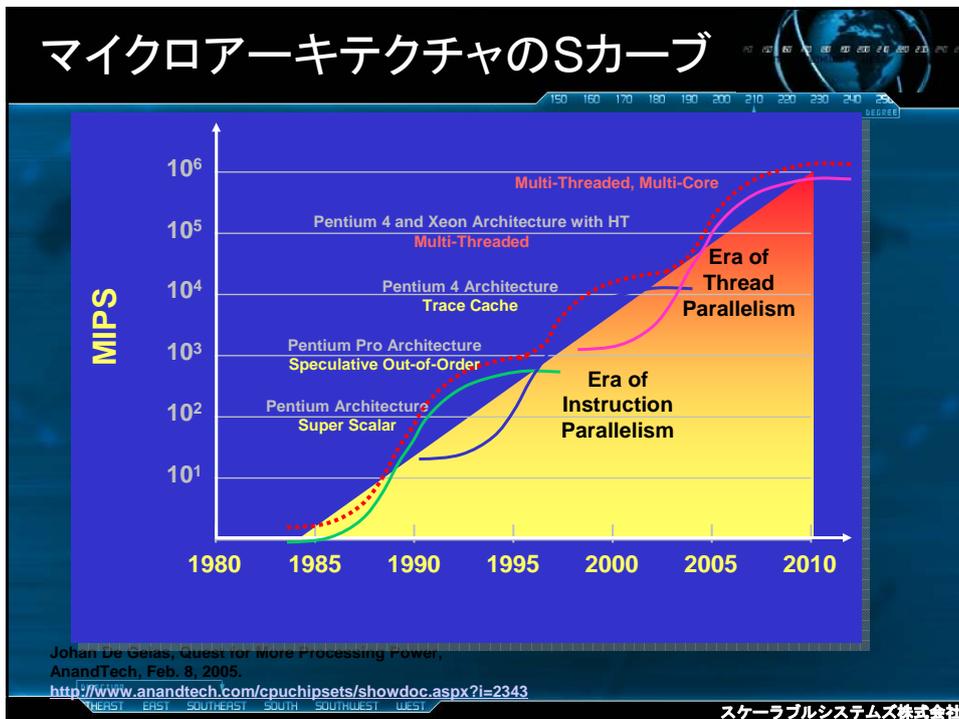
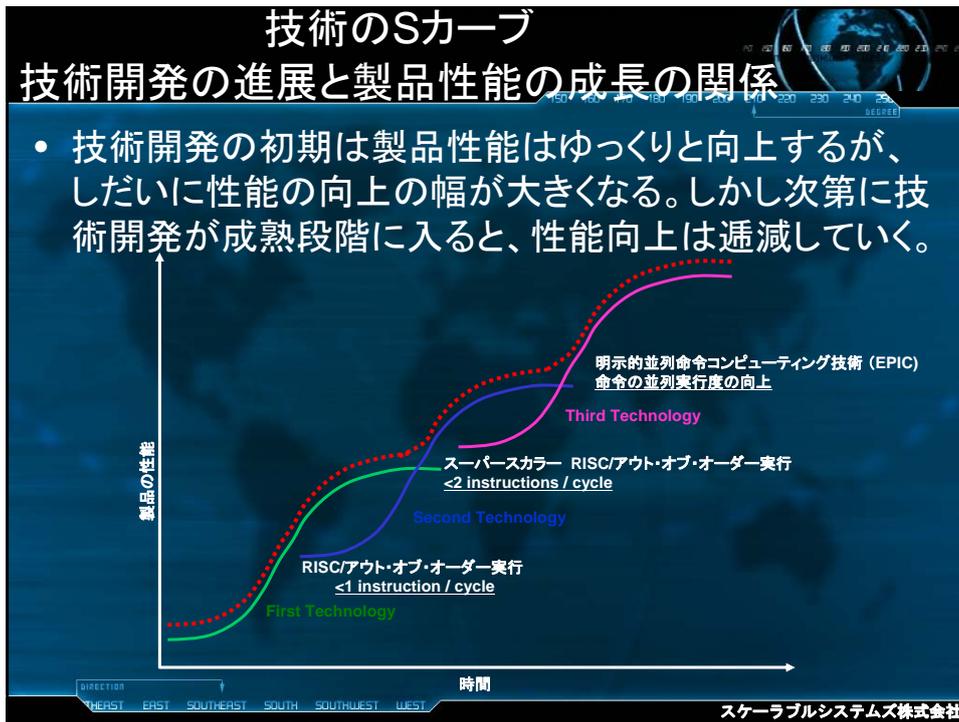


スケラブルシステムズ株式会社

計算機の性能向上

- 動作周波数(クロック)の向上
 - 過去12年間で、Pentiumプロセッサの動作周波数は、60 MHz から 3,800 MHz にまでアップ
 - 現在までの高性能化の約80% はクロック周波数の向上によるもの
- 命令実行の強化と最適化
 - より強力なインストラクションセット
 - 命令実行の最適化(パイプライン化、分岐予測、複数命令の同時実行、命令実行順序の変更など)

スケラブルシステムズ株式会社

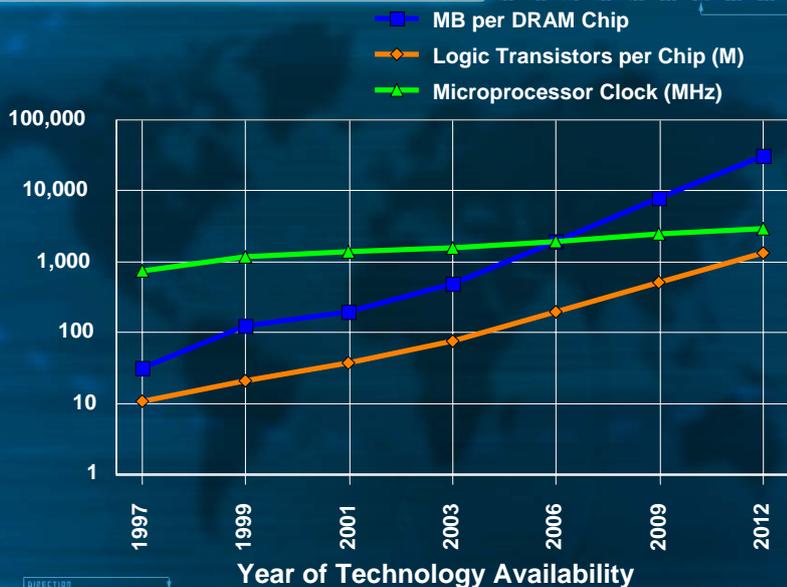


計算機の性能向上

- 動作周波数(クロック)の向上
 - 過去12年間で、Pentiumプロセッサの動作周波数は、60 MHz から 3,800 MHz にまでアップ
 - 現在までの高性能化の約80% はクロック周波数の向上によるもの
- 命令実行の強化と最適化
 - より強力なインストラクションセット
 - 命令実行の最適化(パイプライン化、分岐予測、複数命令の同時実行、命令実行順序の変更など)
- 大容量キャッシュ
 - プロセッサの速度とメモリレイテンシ(待ち時間)とバンド幅のギャップの拡大に対する対策・対応としての容量の拡張

スケラブルシステムズ株式会社

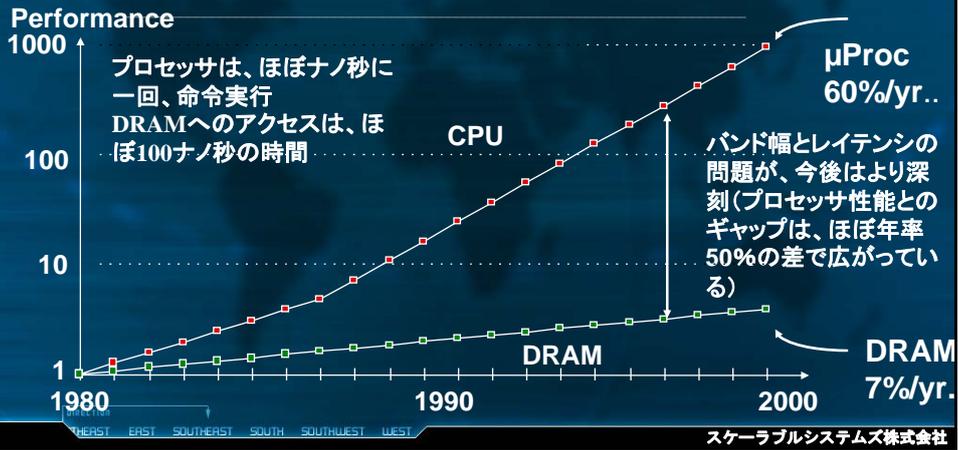
半導体の技術動向予測



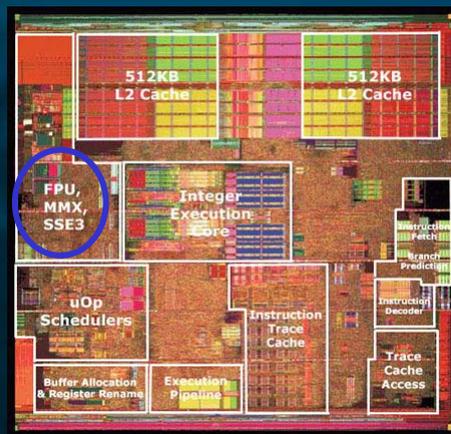
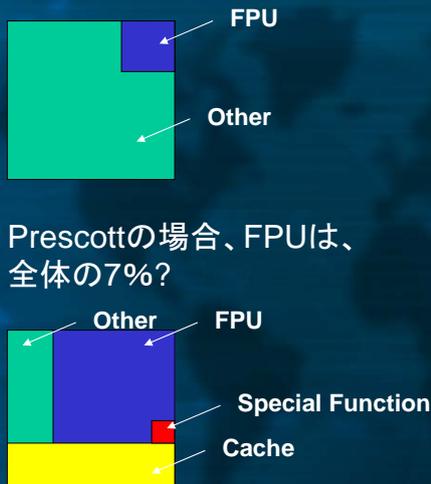
スケラブルシステムズ株式会社

性能ギャップの問題

- プロセッサ速度とメモリアクセスの速度差によって、プロセッサがより高速になったとしても、プロセッサはその演算能力を完全に使い切ることが出来ない



FPUのプロセッサに占める割合？



Pentium Prescott
90nm CMOS
3.4GHz

Processor for HPC?

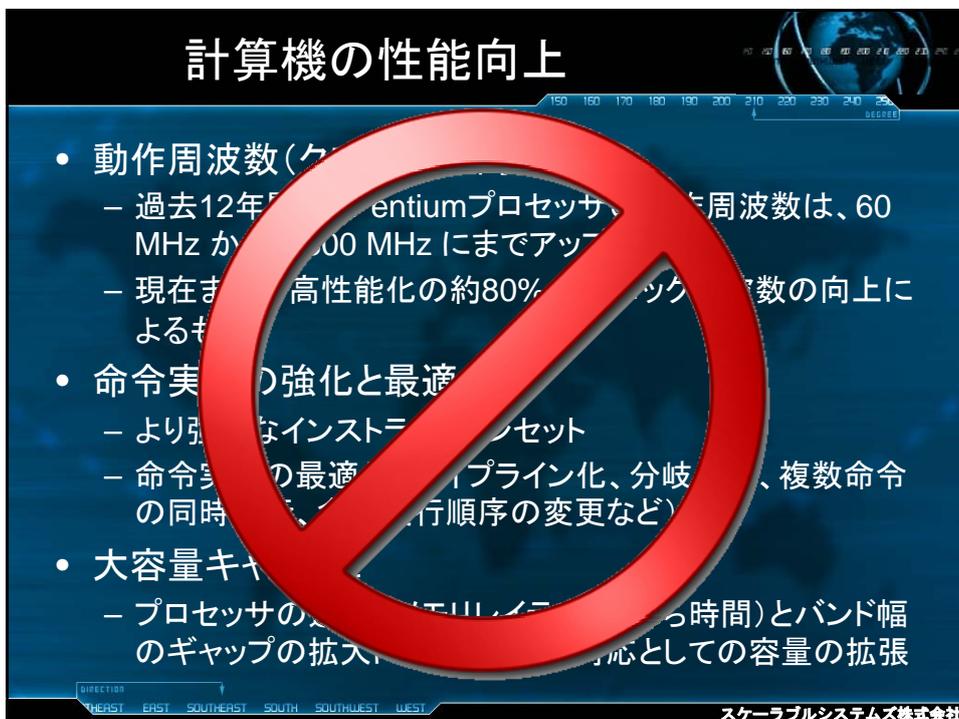
今後の発展と課題



Episode IV
A New Hope

スケラブルシステムズ株式会社

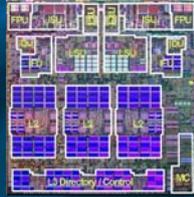
計算機の性能向上



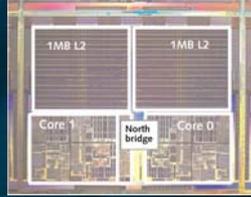
- 動作周波数(クロック)
 - 過去12年(2000年) Pentiumプロセッサの動作周波数は、60 MHz から 3 GHz にまでアップ
 - 現在まで、性能向上の約80%は動作周波数の向上によるもの
- 命令実行の強化と最適化
 - より強力なインストラクションセット
 - 命令実行の最適化(パイプライン化、分岐予測、複数命令の同時実行、実行順序の変更など)
- 大容量キャッシュ
 - プロセッサの動作時間(実行時間)とバンド幅のギャップの拡大(キャッシュとして容量の拡張)

スケラブルシステムズ株式会社

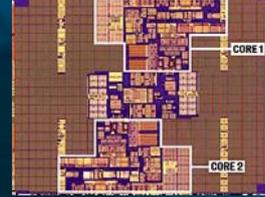
デュアルコアプロセッサ



IBM Power5
with 1.9MB L2



AMD Opteron
with 2MB L2



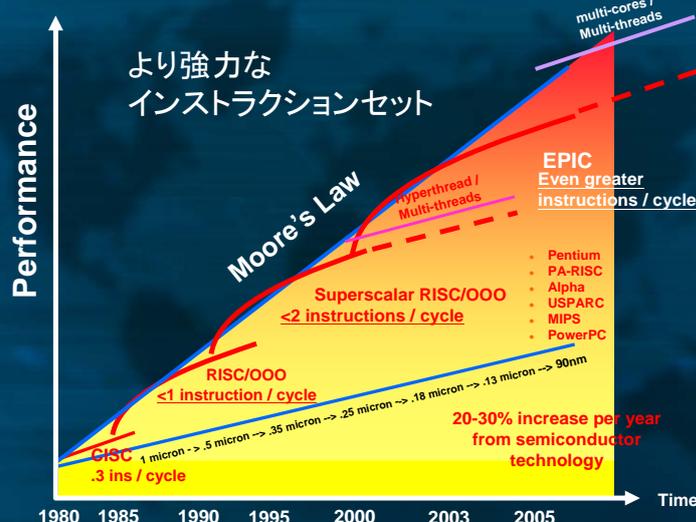
Intel Montecito
With 24MB L3

- チップ上のトランジスタのより有効活用が可能
- スレッドレベルでの並列処理を活用
- よりシンプルなプロセッサの設計が可能
- 将来のマイクロプロセッサはより多くのコアを実装可能
- 将来のマイクロプロセッサはより大容量のキャッシュの実装が可能

DIRECTION: NORTH, EAST, SOUTHEAST, SOUTH, SOUTHWEST, WEST

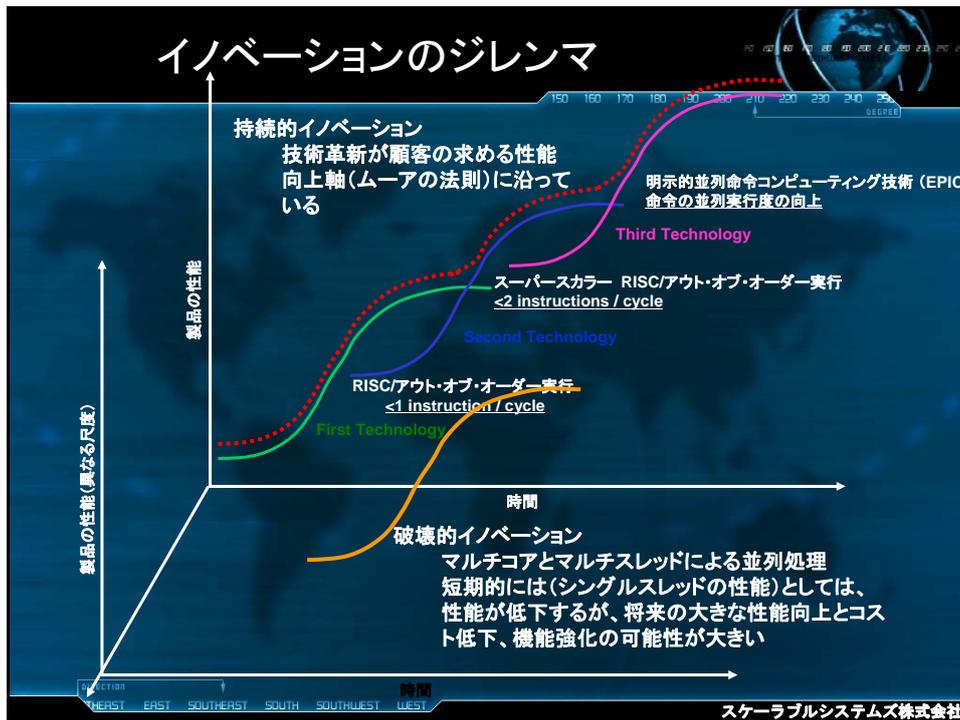
スケラブルシステムズ株式会社

命令実行の強化と最適化



DIRECTION: NORTH, EAST, SOUTHEAST, SOUTH, SOUTHWEST, WEST

スケラブルシステムズ株式会社



マルチコアの利点?

ワークロードの処理効率の向上

- マルチスレッドアプリケーション
 - 現在、多くのアプリケーション(データベース、WEB、科学技術計算)はマルチスレッド化
 - マルチコアプロセッサでは、これらのアプリケーションのマルチスレッドでの実行が容易に可能
- 複数ジョブの処理
 - システムでは、複数のワークロード同時に処理することが必要
 - マルチコアでは、これらのワークロードへの処理が可能

スケーラブルシステムズ株式会社

マルチコアの利点？

消費電力あたりの性能を最大にし、高性能で低消費電力のシステム構築が可能

- OS自身のマルチスレッド対応
 - OSのサービスもマルチスレッドで処理することで、より効率よく処理することが可能
- 仮想化
 - サーバのセキュリティや管理の強化
 - 管理するノード数を減らし、運用コストの削減を図る
- 最新のソフトウェア・テクノロジーの活用

DIRECTION
NORTH EAST SOUTHEAST SOUTH SOUTHWEST WEST

スケラブルシステムズ株式会社

大きな変革・・・しかし、容易ではない

マルチコアプロセッシング(または、汎用もしくは専用プロセッサをソケットに複数搭載可能なこと)は、Ethernetの誕生以来、ITインフラに対しての大きなインパクトをもたらします。

– *Multicore Processing: Disruption or Distraction for the IT Infrastructure?*, Vernon Turner, IDC, November 18, 2004.

デュアルプロセッサは、386プロセッサの発表以来、性能に関して最大の向上を実現します。しかし、このような性能向上には、ソフトウェアの最適化がプロセッサの性能をフルに発揮するためには必要です。

– *Readying Applications for New Server Technologies*, Martin Reynolds, Gartner Research, April 12, 2005.

DIRECTION
NORTH EAST SOUTHEAST SOUTH SOUTHWEST WEST

スケラブルシステムズ株式会社

並列性 (Parallelism) の利用



マルチタスク処理

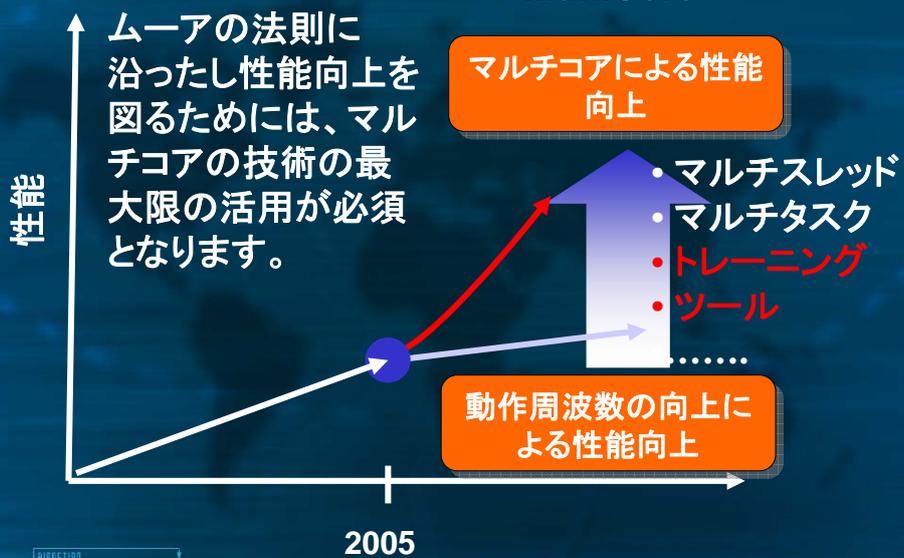
MPIなどによる複数のノード間での並列処理

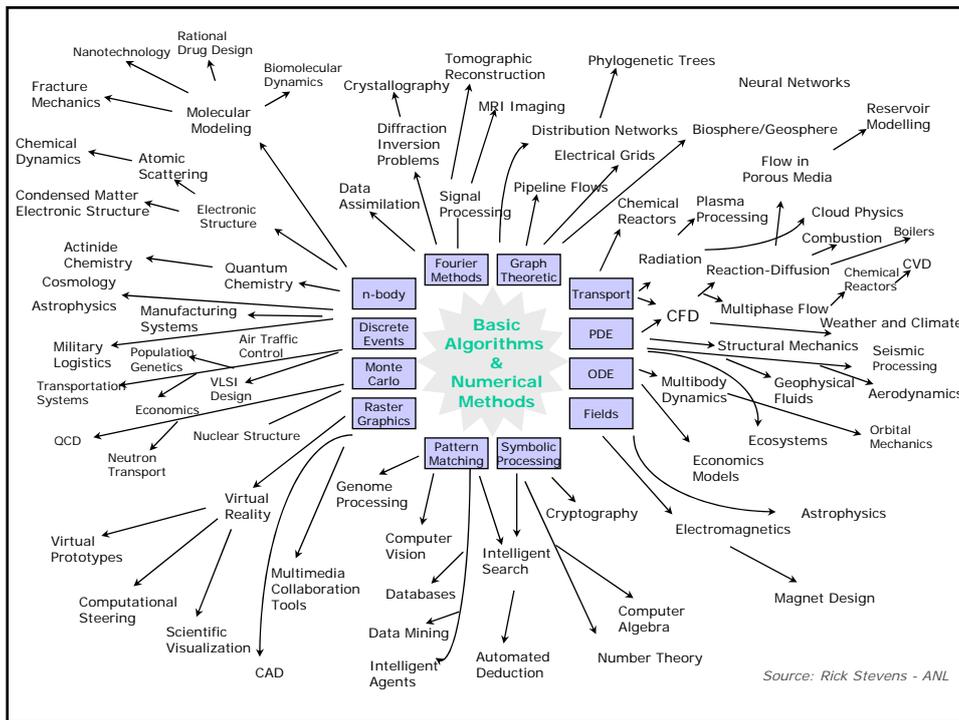
ユーザ及びコンパイラによるスレッドレベルでの並列処理 (TLP)

コンパイラによる命列実行レベルでの並列処理 (ILP)

ムーアの法則 (GHz から MC へ)

MultiCore





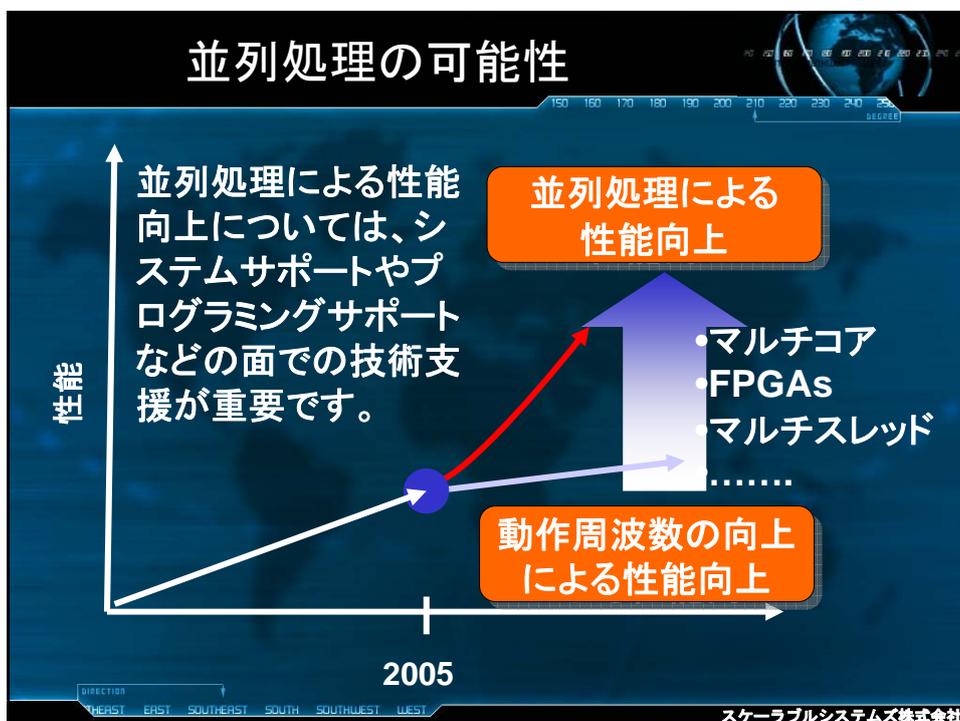
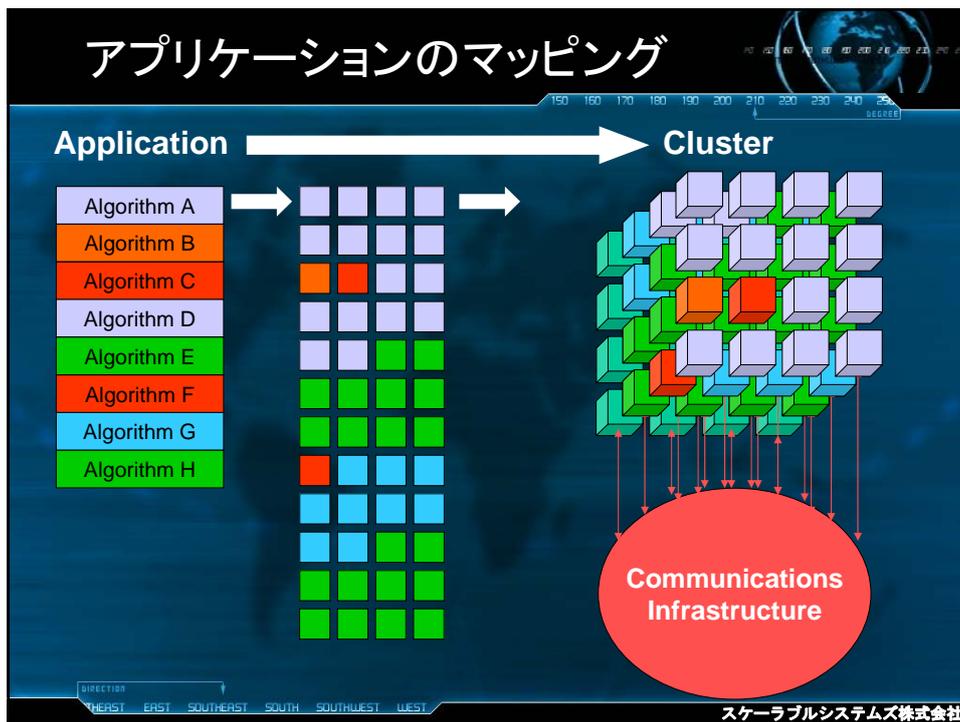
アプリケーションの実装

Algorithm A
Algorithm B
Algorithm C
Algorithm D
Algorithm E
Algorithm F
Algorithm G
Algorithm H

- 一般にアプリケーション毎に解析アルゴリズムは異なる
- これらの複数のアルゴリズムのアプリケーションをリアルタイムでハードウェアにマッピングする必要がある
- アルゴリズムごとに要求するコンピュータリソースはかなり異なる

DIRECTION
 NORTHEAST EAST SOUTHEAST SOUTH SOUTHWEST WEST

スケーラブルシステムズ株式会社



ビル・ゲイツ氏の基調講演 HPC goes mainstream

Computation Transforming The Sciences

Technical Computing

Earth Sciences

Life Sciences

Social Sciences

New Materials, Technologies & Processes

Math and Physical Science $E=MC^2$

Computer & Information Sciences

Multidisciplinary Research

NEARST EAST SOUTHEAST SOUTH SOUTHWEST WEST

スケーラブルシステムズ株式会社

Supercomputing Goes Personal From Microsoft HPC Presentation

	1991	1998	2005
System	Cray Y-MP C916 	Sun HPC10000 	Shuttle @ NewEgg.com 
Architecture	16 x Vector 4GB, Bus	24 x 333MHz Ultra-SPARCII, 24GB, SBus	4 x 2.2GHz x64 4GB, GigE
OS	UNICOS	Solaris 2.5.1	Windows Server 2003 SP1
Performance	~10 GFlops	~10 GFlops	~10 GFlops
Top500 #	1	500	N/A
Price	\$40,000,000	\$1,000,000 (40x drop)	< \$4,000 (250x drop)
Customers	Government Labs	Large Enterprises	Every Engineer & Scientist
Applications	Classified, Climate, Physics Research	Manufacturing, Energy, Finance, Telecom	Bioinformatics, Materials Sciences, Digital Media

Kyril Faenov (kyrilf@microsoft.com)
Director of High Performance Computing
Microsoft Corporation

NEARST EAST SOUTHEAST SOUTH SOUTHWEST WEST

スケーラブルシステムズ株式会社

マイクロソフトのHPCビジョン

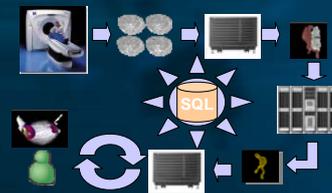
- スーパーコンピュータ上でのバッチ処理



- パーソナルクラスタ上でのインタラクティブ(対話)処理



- アプリケーション間での複雑なワークフローへの対応



DIRECTION
NORTHEAST EAST SOUTHEAST SOUTH SOUTHWEST WEST

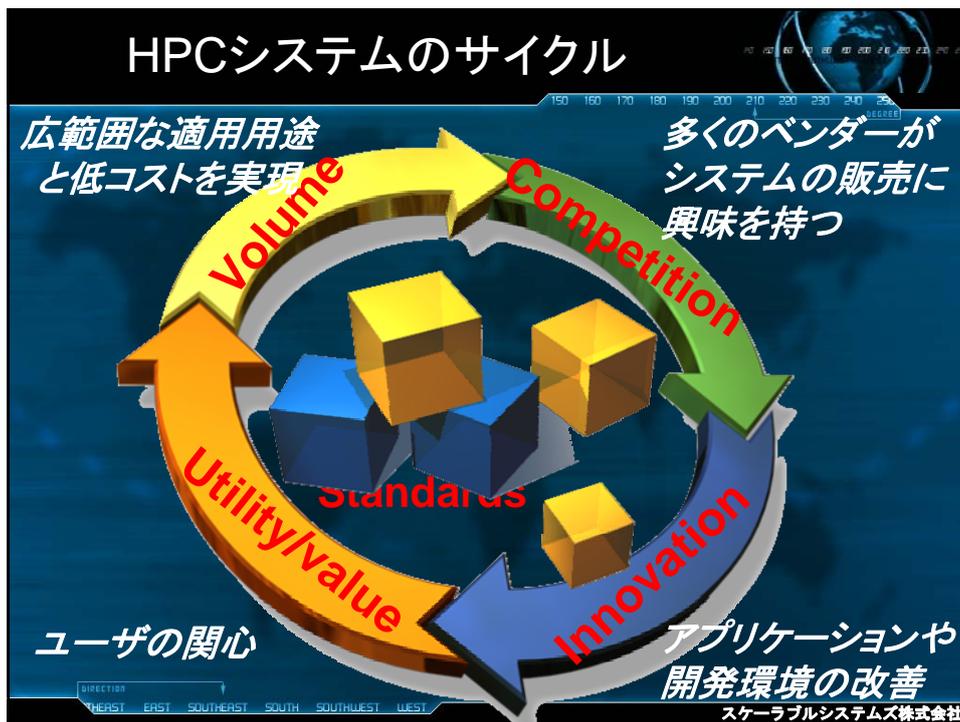
スケーラブルシステムズ株式会社

標準コンポーネントの利点

- 特定のベンダーからのシステムを組み合わせるのではなく、他社のシステムも含めて最適なシステムを選択が可能
 - スケーラブルSMP、ベクトル計算機、クラスタの幅広い選択肢
 - 64ビット、マルチコアマイクロプロセッサの性能向上を最大限に活用
- 標準コンポーネントの技術革新の活用
 - PCI-Expressや、FB-DIMMの利用技術

DIRECTION
NORTHEAST EAST SOUTHEAST SOUTH SOUTHWEST WEST

スケーラブルシステムズ株式会社



Breaking the 1-2K nodes Barrier !

<http://www.wilk4.com/misc/soundbreak.htm>

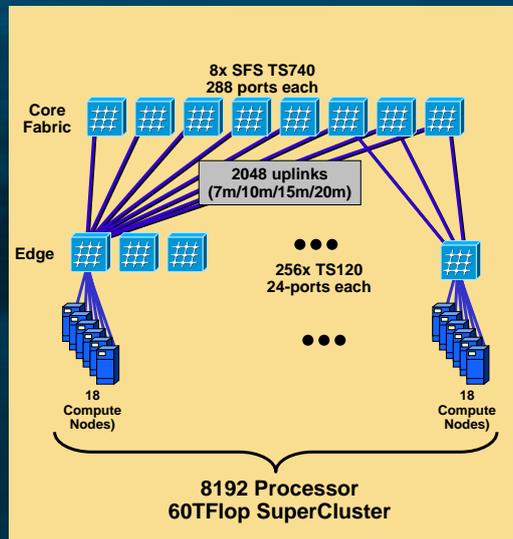
- 音の障壁, サウンド・バリアー (sound barrier)
飛行機が音速近くになると、衝撃波の発生によって、抵抗の増大、境界層の剥離など、設計・運用上のさまざまな障害(壁)に出合って、超音速飛行は不可能かと思われた時代があった(1947年ごろまで)ので、音の障壁といわれていた。

クラスタのノード数が、ある規模に近くなると、その構築や運用において、負担の増大、システムの安定稼働、スケーラビリティなど、設計・運用上のさまざまな障害(壁)に出合って、クラスタ構築は不可能と思われた時代があった(?)

スケラブルシステムズ株式会社

米国エネルギー省 サンディア国立研究所

- システム:
 - 4096 Dell Servers
 - 50% Blocking Ratio
 - 8 TS-740s
 - 256 TS-120s
- TOP500 (Nov.5th)
 - No.5
- 用途:
 - ‘Capability’ クラスタ
 - 標準コンポーネントでのシステム構築

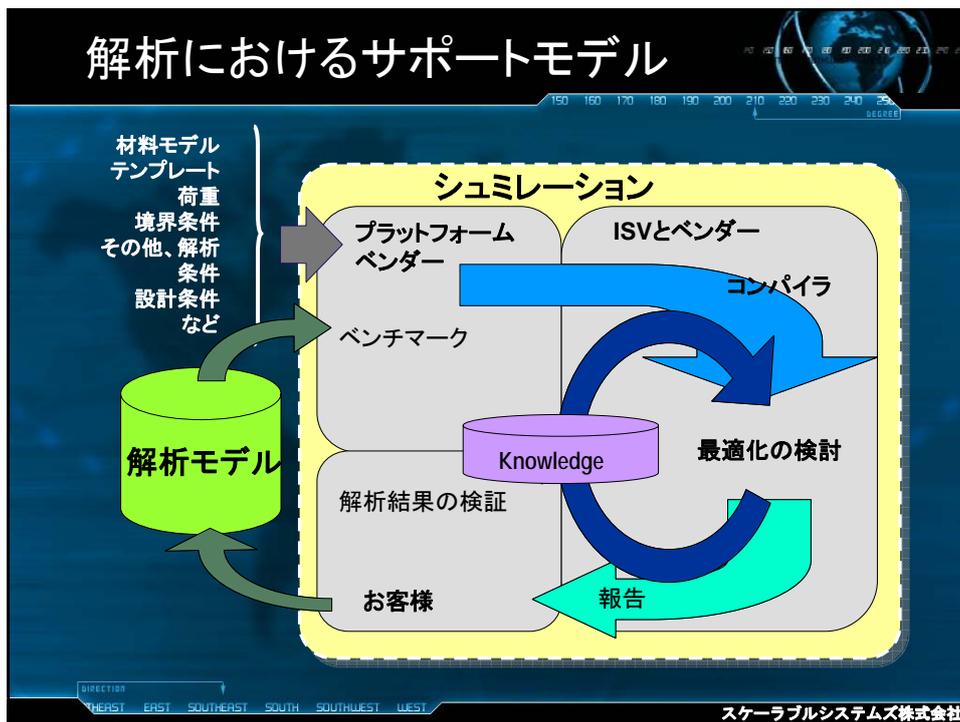


スケラブルシステムズ株式会社

まとめとして

- HPCも高付加価値のサービスが問われる時代
- 部分的な最速の追求はあまり意味がない
 - 無秩序なシステム
 - 十分な競争力を発揮出来ない
- ベンダー間の競争
 - コンポーネントの提案ではなく、複数の技術を組み合わせ、システムを提案し、それをサポートする能力での評価

スケラブルシステムズ株式会社



「テクノロジー」について

- 「テクノロジー」をどのようにとらえるか？
 - 企業経営基盤のコア要素
 - ユーザの本質的な課題を解決する戦略的な武器
- 今後のマーケットは、「テクノロジー」が牽引する
 - ただし、ユーザに何らかのメリットをもたらさない「テクノロジー」は、意味を成さない
 - テクノロジーを最適に組み合わせることで、問題解決のためのソリューションを提供

テクノロジーとITインフラを明確に説明することが今、求められています。

スケラブルシステムズ株式会社

さらに詳しい情報は.....



150 160 170 180 190 200 210 220 230 240 250 DEGREE

Episode VI Return of the Jedi

テクノロジーとITインフラを明確に説明することが今、求められています。

DIRECTION
NORTH EAST SOUTHEAST SOUTH SOUTHWEST WEST

スケーラブルシステムズ株式会社

さらに詳しい情報は.....



150 160 170 180 190 200 210 220 230 240 250 DEGREE

SSTC
Scalable Systems Co., Ltd.
スケーラブルシステムズ株式会社

<http://www.sstc.co.jp>
biz@sstc.co.jp

Enter >> Back to SSTC Home Page >>

Copyright 2005 Scalable Systems Co., Ltd. All rights reserved.

Technology Consultation
Business Overview

- 弊社のコンサルテーションに関するご提案資料もダウンロード可能です。(非公開WEBページ)別途、弊社に内容等については、お尋ねください。

お問い合わせ先:

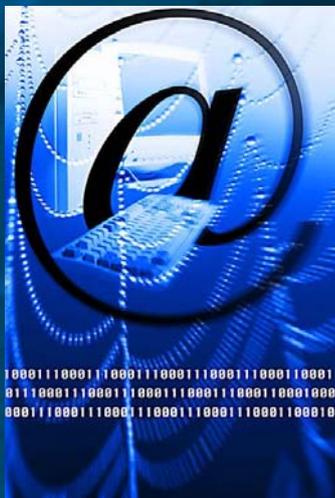
〒102-0083
東京都千代田区麹町3-5-2
BUREX麹町 8F
電話: 03-5875-4718
FAX: 03-3237-7612
E-mail: biz@sstc.co.jp
<http://www.sstc.co.jp>

www.sstc.co.jp/biz

DIRECTION
NORTH EAST SOUTHEAST SOUTH SOUTHWEST WEST

スケーラブルシステムズ株式会社

スケーラブルシステムズ株式会社



ハイエンドコンピューティングに関するコンサルテーションとして、幅広いサービスをご提供致します。

このサービスを最大限に活用していただくことで、コラボレーションによる「顧客志向」のコンサルテーションサービスをご提供できればと思っております。

DIRECTION
NORTH EAST SOUTHEAST SOUTH SOUTHWEST WEST

スケーラブルシステムズ株式会社

社名、製品名などは、一般に各社の商標または登録商標です。無断での引用、転載を禁じます。

In general, the name of the company and the product name, etc. are the trademarks or, registered trademarks of each company.

Copyright Scalable Systems Co., Ltd., 2005. Unauthorized use is strictly forbidden.

2005年11月

DIRECTION
NORTH EAST SOUTHEAST SOUTH SOUTHWEST WEST

スケーラブルシステムズ株式会社