

インテル株式会社
プラットフォーム&ソリューション マーケティング本部

intel.

Architectures

Cluster: MPP
Star: TIE

Constellation

Legend: Single, SMP, MPP, TIE

資料 www.top500.org

クラスタ・ソリューションの拡大！

intel.

2

HPC 500の傾向

- 1TFLOPSを超えるのは47位以上 (半年前 23位)
- 500位のシステムは195.8GFLOPS
- PC クラスタがすべての性能クラスに登場して来た
- TOP500の中にインテル社製のPCクラスタは55個
- Itanium® プロセッサアーキテクチャのクラスタは4個 (2 + 2)
- インストールされたFLOPS性能ベースで、IBM(31.8%)、HP(22.1%)、NEC (4.6%) の順

intel.

3

HPCのマーケット

- 100 ノード以上の計算機 (MPP)が大規模科学計算を支配
- 価格性能比ではPC (Pentium®4, Xeon™ プロセッサ)ベースのシステム
- ノードあたりの性能の低い RISCプロセッサが中間層を支配Itanium® 2 プロセッサへの置き換えを期待

**Xeon™、Itanium®2 プロセッサ
が次世代スーパーコンピュータの
構成 ノードとして期待されている**

intel.

4

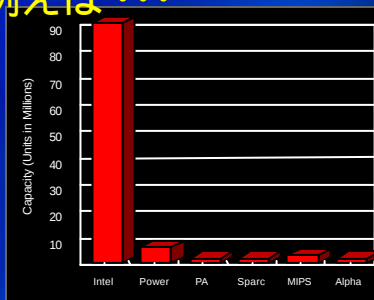
どうしてか？ 汎用技術の向上

• 性能価格比

— 様々なメーカー、様々なプラットフォームから選択

- 汎用技術の高速化
- 安定供給
- アップグレード
- 周辺機器
- 汎用ツール
- ソフトウェア資産

例えば・・・

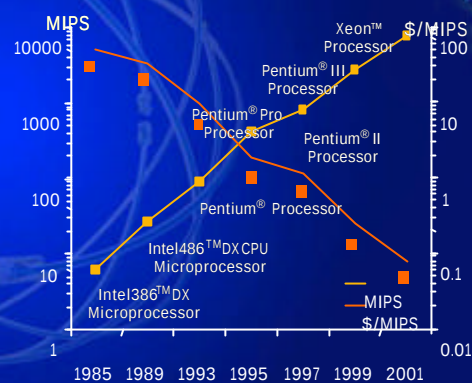


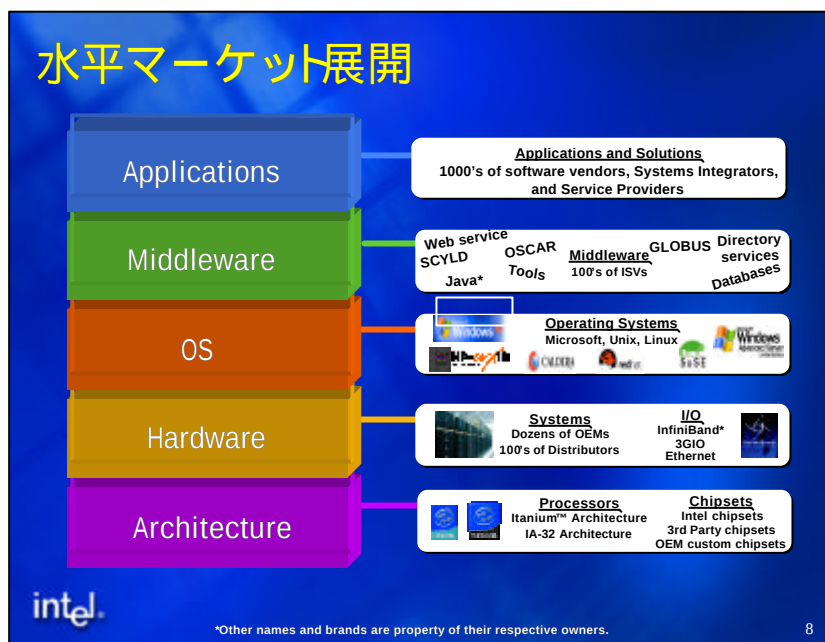
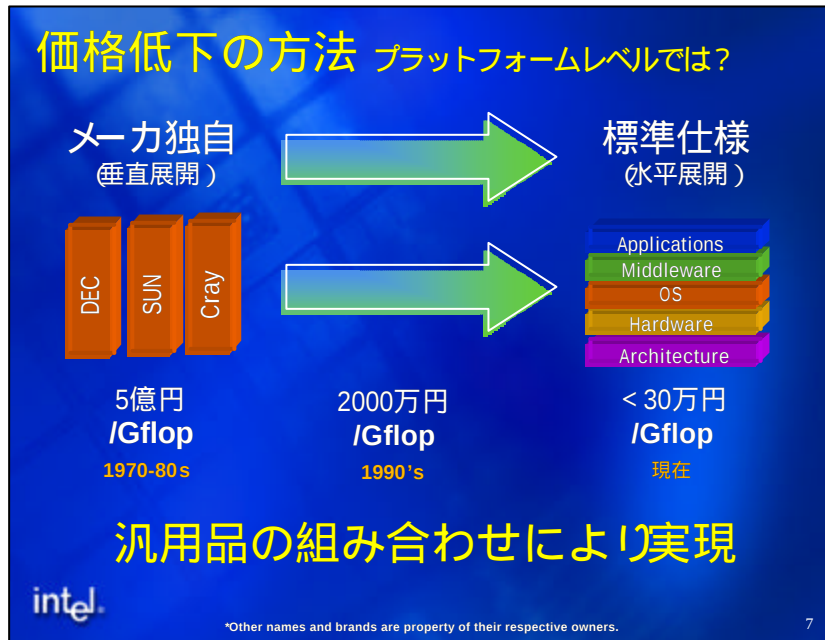
intel.

5

ムーアの法則

トランジスタは18ヶ月毎に2倍..





水平展開のメリット

⁽¹⁾
\$19.0B

(日本円に直すと、約2兆3千億円の投資)

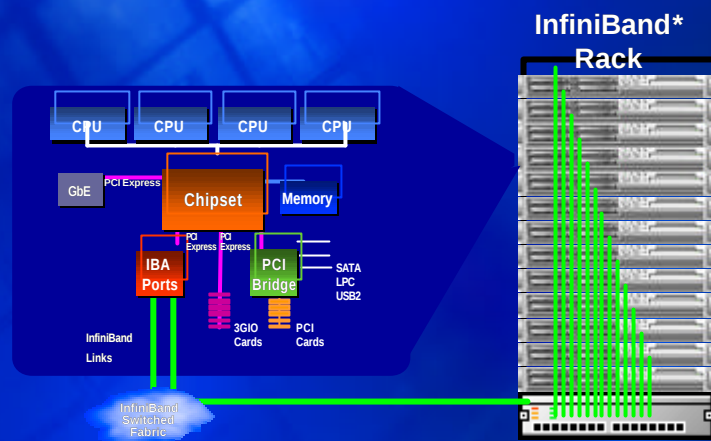
インテル・アーキテクチャに関わる研究開発費はトータル\$19.0B
一方、サン・マイクロシステムに関わる研究開発費はトータル\$1.6B(1)

intel.

1: 2000 R&D spending for Intel, Microsoft, Compaq, Dell, HP, SGI & IBM; Source: 2000 Annual Reports

9

クラスタ・プラットフォーム概要



intel.

10

汎用技術トレンド ハードウェア周辺

Processors

2010年には、
30GHz・2000億トランジスタ

Memory

DDR2 667/400,
DDR-II, DDR-III

Internal I/O

PCI, PCIX, PCIX DDR,
PCIExpress™

Storage Conn.

U320, U640,
S-ATA, SAS

Management

PMI1.5

Interconnect

10/100/G/10GbE net,
InfiniBand
(競合技術?)



InfiniBand is a trademark or registered trademark of Intel Corporation or its subsidiaries in the United States and other countries.

11

メモリ技術の向上

DDR-II & DDR-III Info

2001	2002	2003	2004	2005	>2006
DDR-I 333, 400 2.7GB/s-3.2GB/s 0.18/0.16/0.13 micron 184-pin 2.5V SSTL-2 4 Banks 2-bit prefetch 66 pin TSOP-III/FBGA 256, 512, 1GB容量		DDR-II 400, 533, 600, 667 3.2GB/s-4.8GB/s 0.09 micron 240-pin 1.8V SSTL-18 4 Banks/ 8 Banks? 4-bit prefetch Differential Data Strobe 60 balls FBGA Post CAS, OCD On-Die Termination 256, 512, 1GB, 2GB, 4GB容量			DDR-III 800, 1200 > 6.4GB/s 0.065 micron 1.5-1.2V 8-bit prefetch Differential Data Strobe 1GB, 2GB, 4GB, >4GB容量

SS Reports

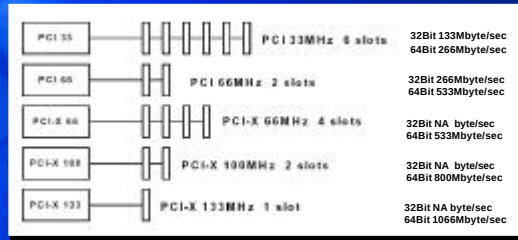
Source: <http://www.vr-zone.com/Home/news132/news132.htm#2129>



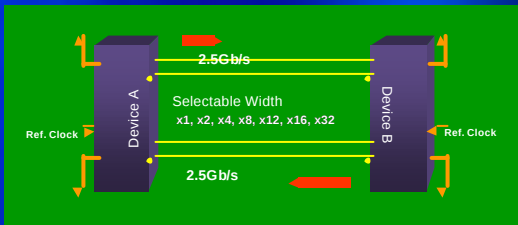
12

D 技術の向上 パラレルから高速シリアルへ

PCI-X



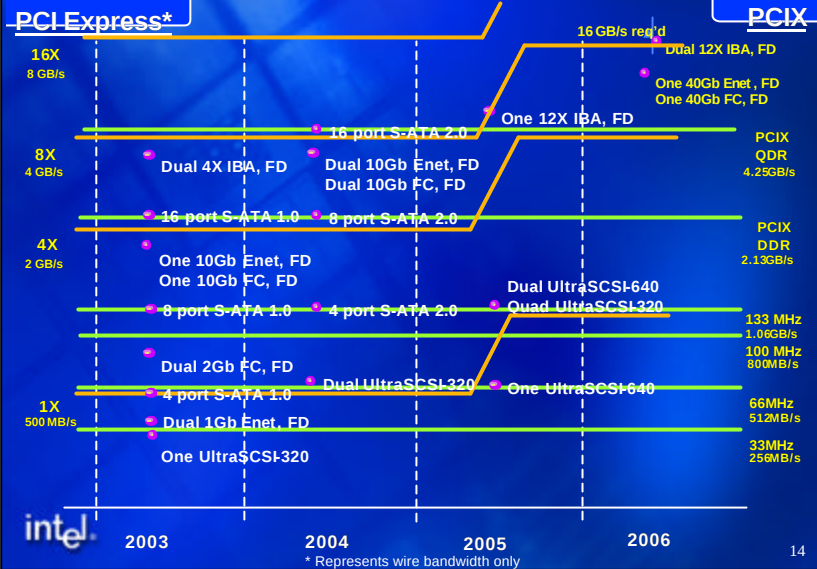
PCI-Express™



intel.

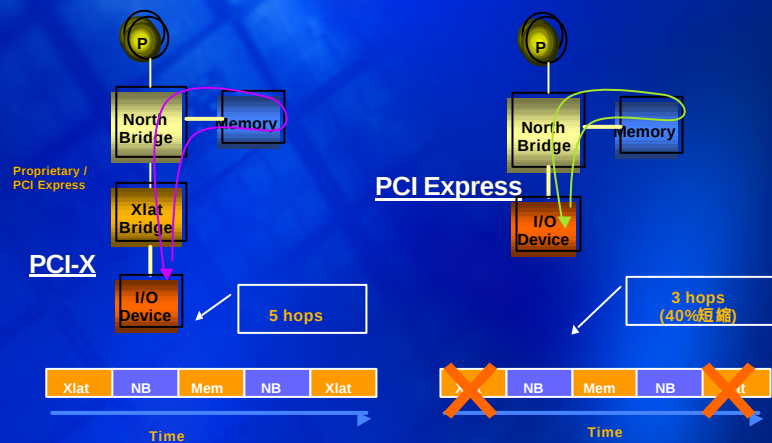
13

システム内 バンド幅の要求



14

アクセス遅延 アクセス遅延の軽減



intel.

15

信頼性の向上 エラー発見 機能比較

障害タイプ	PCI-X • 32-bit parity	PCI-X 2.0 • 32-bit parity • 64-bit ECC	PCI Express* • 32-bit CRC • LLR • Error Forwarding
Framing Error	不可	不可	収集可
Address/ Header	致命的障害として発見	致命的障害として発見	収集可
Data	発見可能、収集不可	収集可、複数ビット障害 収集不可	収集可
Forwarded Data Error	可能 localとforward 区別不能	可能 localとforward 区別不能	可能

intel.

16

インターコネクトの選択 インフィニバンド*

遅延の少ない 管理メッセージ

- ハードウェアによる RDMA で zero copy 転送
- ハードウェア管理の send/receive
- 複数 VLs 環境下での 高い優先順位転送

高い Data 転送能力

- 2.5Gb/s bandwidth 1x links
- 10Gb/s bandwidth 4x links
- 30Gb/s bandwidth 12x links

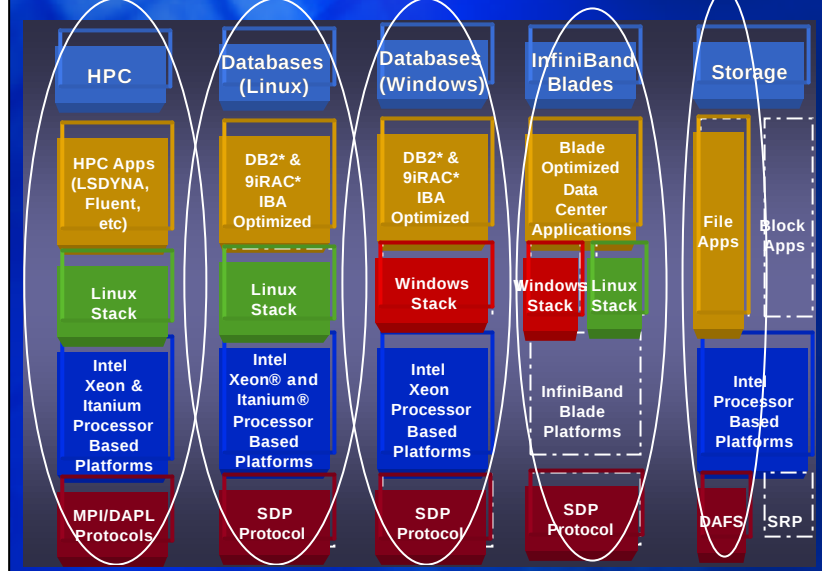
低い CPU 負荷

- アプリケーションから 制御できる API
- ハードウェアによる CRC とパケット再送
- ハードウェアによる パケットコントロール

intel.

17

インフィニバンド* 標準のインターコネクト



比較 InfiniBand* Architecture vs. Ethernet

	InfiniBand* Arch	Ethernet + TOE
端末間遅延	少ない遅延: μ Sec	数十 μ Sec かそれ以上
QoS	VLグループ間の優先順位設定可能	単一グループ設定
同時通信	100% 同時通信可能	802.3ad
Security	メモリ保護領域設定など HWでサポート	VLAN, Require IP-SEC
信頼性	HWによる障害時の自動パス移行	SWにより手動設定
RDMA	Native	Optional, 仕様策定中
HWの状況	4X+4X (2x10Gb) Q4 2002 12X (30Gb) 2003	1+1 GbE TCP/IP Q4 2002 10 GbE TCP/IP 2004
価格	100-200\$ for 2*10Gb ports を目標	2Gb TOE + memory $\approx$ 200\$ 10Gb 1000\$以上を予測

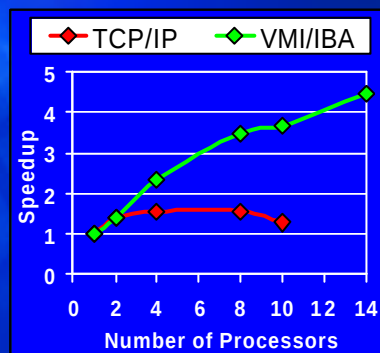
intel.

19

性能の一例 VMIをInfiniBandで

LS-DYNA*

乗用車の衝突解析
[7th International
LS-DYNA Users
Conference 2002,
Chaltas and Magro,
p. 18-7 (LSTC, 2002)]



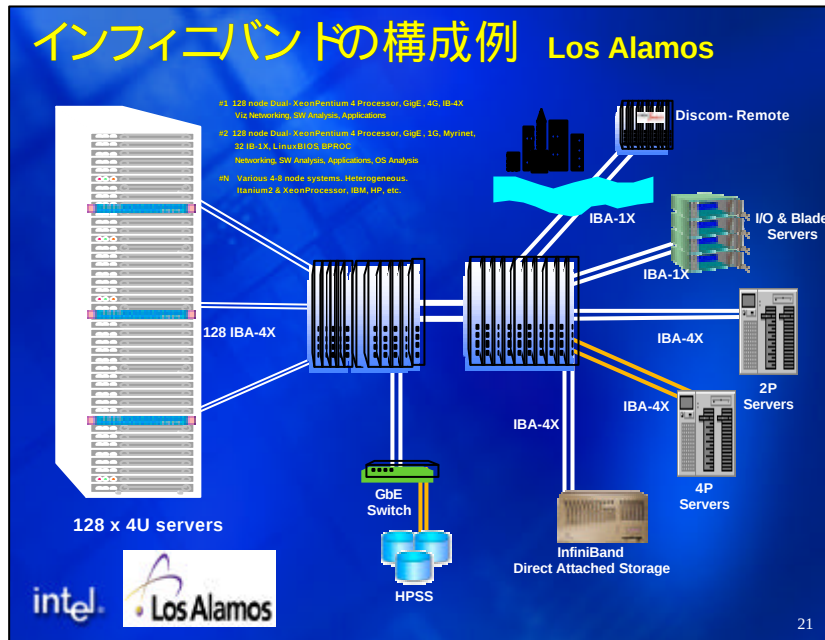
$$\text{Speedup} = \frac{\text{Single-processor run time}}{\text{Multi-processor run time}}$$

インフィニバンドの効果は絶大!

NC3A

intel.

20



管理機能 購入必須項目

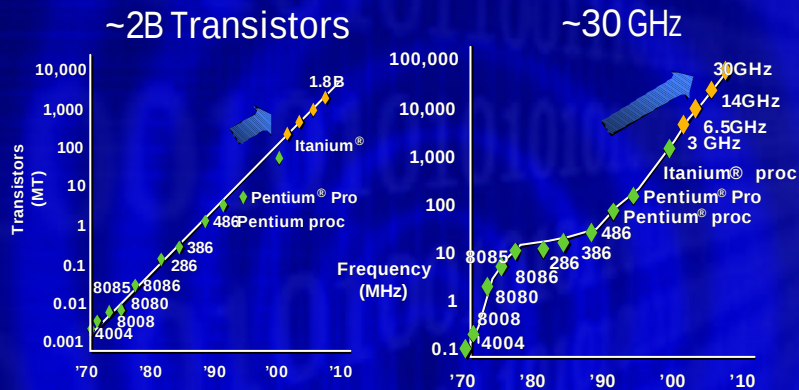
ノードを 管理できますか？

デモ

intel.

22

プロセッサ 10年の展望...



2010年...30GHz素子, 10 ナノメートル以下の製造技術、
1テラ命令/s 性能 by 2010」(1)

1) Pat Gelsinger, Intel CTO, Spring 2002 IDF

23

プロセッサ HPC分野での位置づけ



performance 2P,
4P, 8P, 8P+
high-end servers

Intel® Itanium™ プロセッサは 現在と将来の64-bit HPC 分野での計算要求を意識して作られた

例として 将来の革新的なHPC分野で使用される広大なメモリ空間を必要とする、大規模なSMP systems、複雑な high-end 浮動小数点アプリケーション、64-bit 整数アプリケーション



4P, 8P, 8P+
rack optimized,
pedestal servers

2P
rack optimized,
pedestal servers

Intel® Xeon™ プロセッサは 現在の計算要求のほとんどを占める32-bit HPC 分野を意識して作られた

例として 今日のHPCに関わる多くの 32-bit アプリケーション、分散アプリケーションサーバ、peer-to-peer のサーバ環境、高速32-bit 整数アプリケーション



UP, DP
Rack, blades, low
power blades

Intel® Pentium® 4 プロセッサは 分散環境の高性能端末 peer-to-peer 等の分野を意識して作られた

例として HPC に関わる端末アプリケーション、peer-to-peer 端末としての最高性能 (Java, XML...)

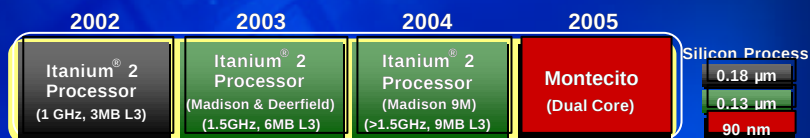
intel.

24

ロードマップ Itanium® プロセッサ ファミリ

• Itanium® プロセッサ ファミリ ロードマップの強化

- 競合環境に対応
- デュアル・コア製品のプル・イン
- 製品サイクル



• Montecito プロセッサでデュアル・コアを実現

- 同一パッケージ、同一バス・プロトコル
- 90nm製造技術
- 2005年にリリース予定

intel.

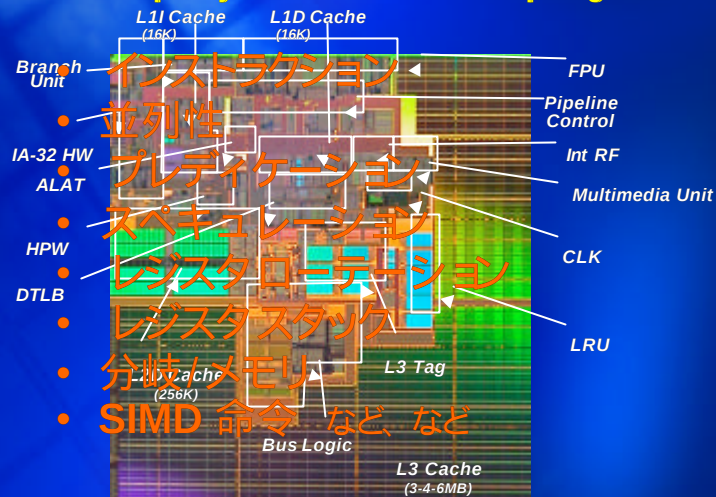
強力なラインアップの継続

25

EPIC

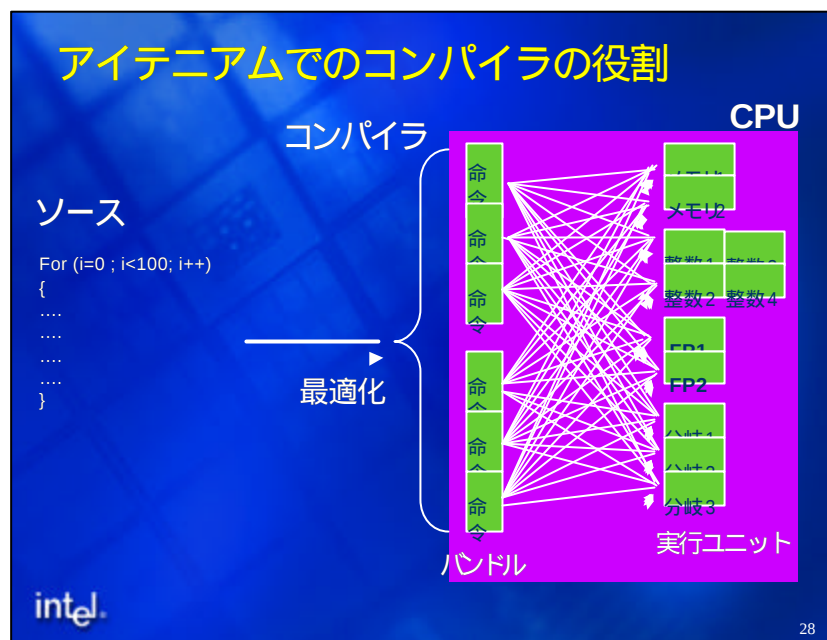
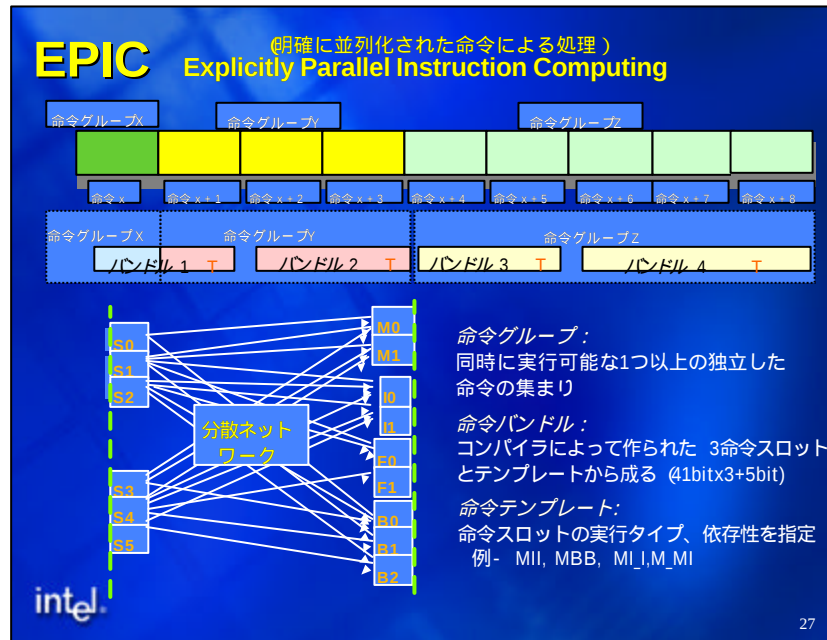
(明確に並列化された命令による処理)

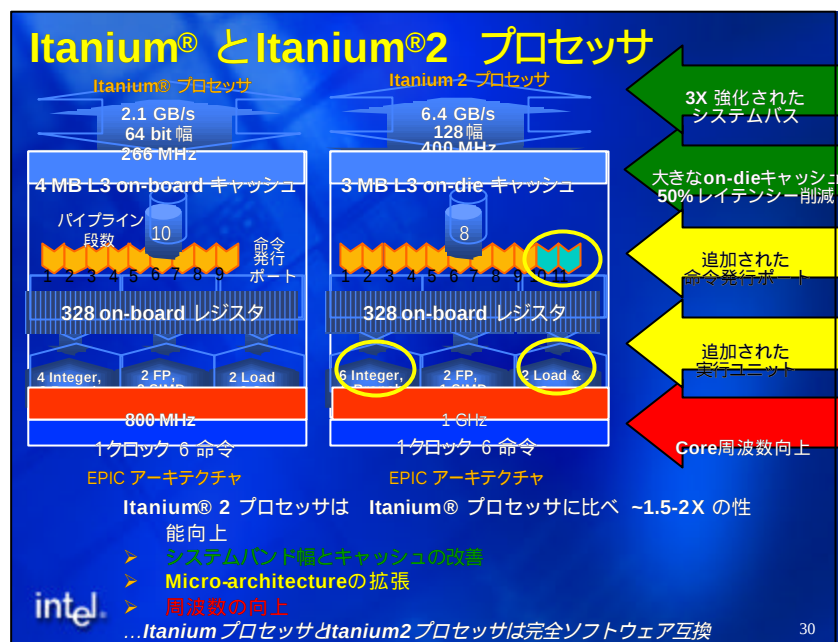
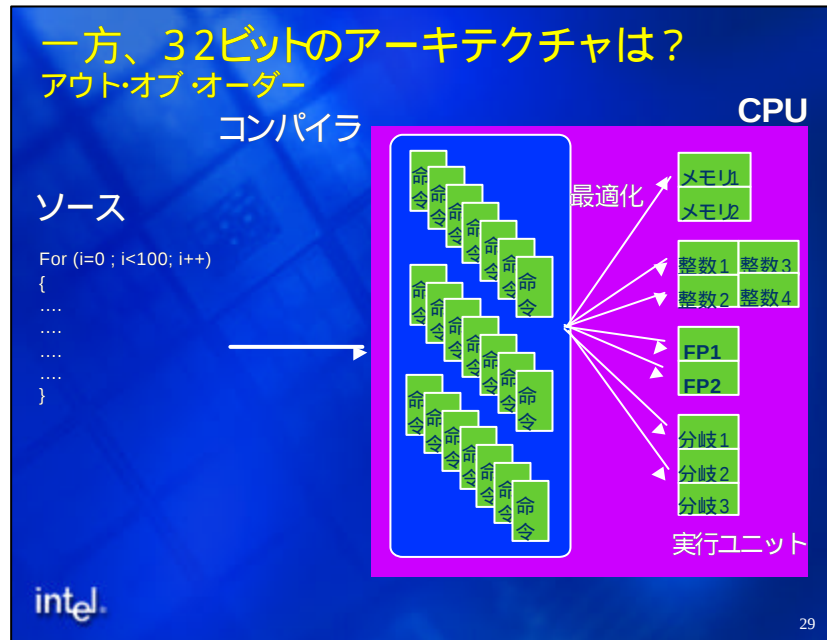
Explicitly Parallel Instruction Computing



intel.

26

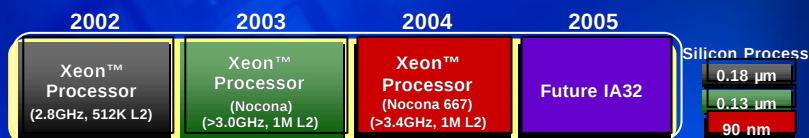




ロードマップ Xeon™ (DP) プロセッサ ファミリ

• Xeon™ プロセッサ ファミリ ロードマップの強化

- ハイパー・スレッディング技術
- コア周波数・システム・バス周波数の向上
- キャッシュ・サイズの向上



• 2003年では・・・

- 次世代 DPプロセッサ (Nocona)の投入
- コア周波数の強化： 3GHz強
- L2キャッシュの強化： 1MB



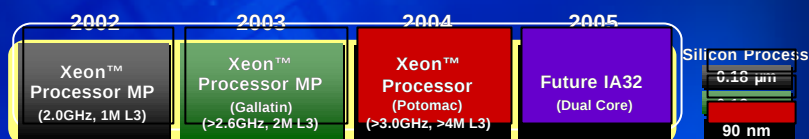
強力なラインアップの継続

31

ロードマップ Xeon™ MP プロセッサ ファミリ

• Xeon™ プロセッサ MPファミリ ロードマップの強化

- ハイパー・スレッディング技術
- コア周波数・システム・バス周波数の向上
- キャッシュ・サイズの拡大



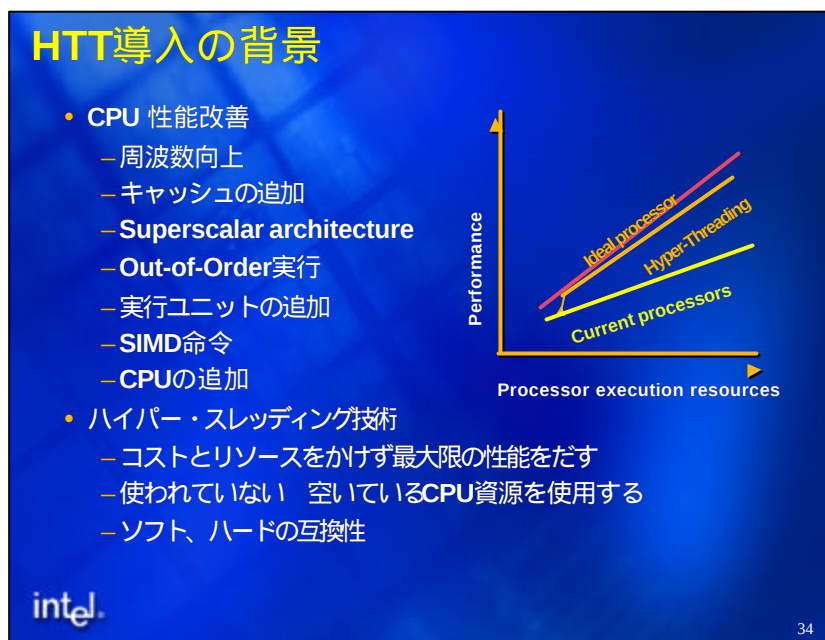
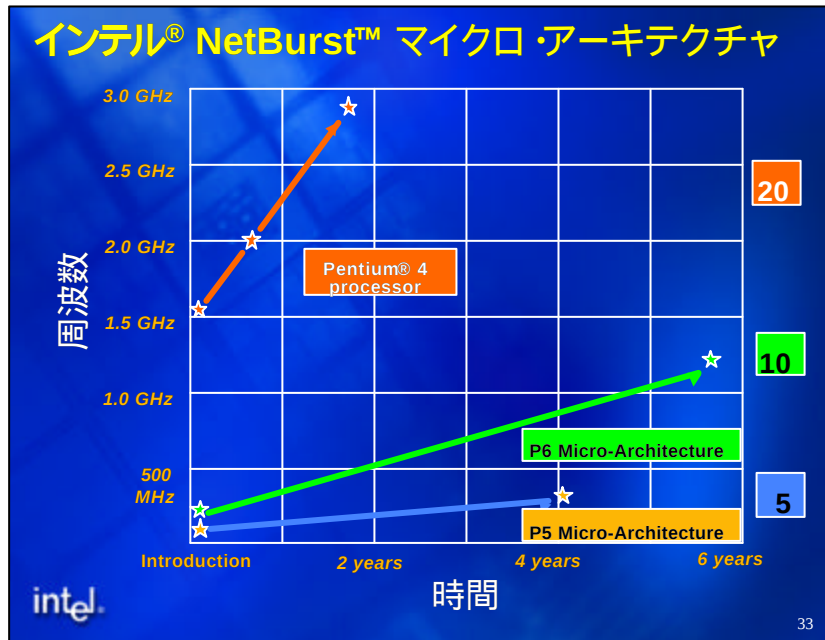
• 2003年では・・・

- コア周波数の強化： 2.6GHz強
- L3キャッシュの強化： 2MB



強力なラインアップの継続

32



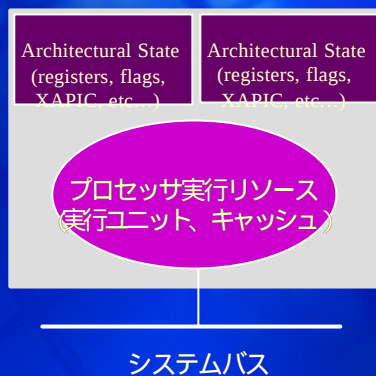
ハイパー・スレッディング技術 1個が2個に...

- 1個の物理プロセッサを2個の論理プロセッサにみせる
- 複数の命令の流れ (スレッド) を1個のプロセッサで同時に実行するSMTのインテル版
- 時間あたりに実行できる仕事を増やすことにより、プロセッサの効率を最大限に生かす
- 性能の改善や応答時間の短縮等をもたらす
- Intel® Pentium® 4 と Xeon™ プロセッサで利用可能

intel.

35

ハイパー・スレッディング技術 実装方法



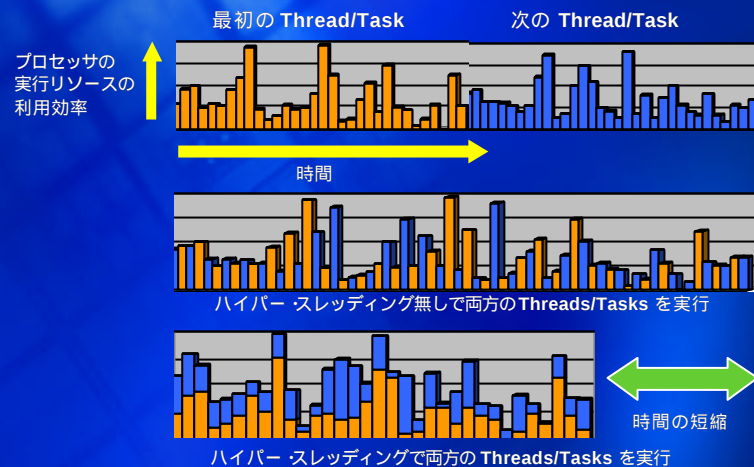
- 2個の論理CPU
- Architectural stateを持つ
 - レジスタ, フラグ, XAPIC, タイムスタンプ, etc.
- ほとんどのリソースを2個の論理CPUで共有
 - キャッシュ, 実行ユニット, etc.
- ハイパー・スレッディング技術は空いているリソースの使用率を高める技術

ハイパー・スレッディング技術で、
2-wayマルチプロセッサのアプリケーションを実行できる

intel.

36

ハイパー・スレッディング技術 動作コンセプト



プロセッサの実行リソースの利用効率の向上は
性能向上を意味する

37

インテル・ソフトウェア・ツール HPC性能最適化

• コンパイラ

- Intel アーキテクチャを採用した Windows* Linux* のプラットフォームで動作
- 最新のマイクロアーキテクチャに対応した最適化を行えるのが特徴
 - インテル® Xeon™ プロセッサ ファミリー
 - インテル® Itanium® プロセッサ ファミリー
- 自動ベクタライズと OpenMP のサポート



• ライブラリ

- Math Kernel Library
 - BLAS, LAPACK, FFT's
- Integrated Performance Primitives
 - ベクタライズされた超越関数群
 - cblas interface



• 解析ツール

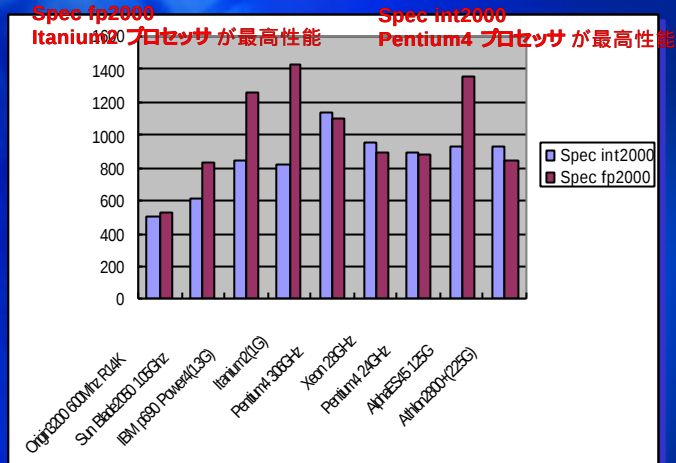
- Vtune™ を使った性能分析、最適化



<http://www.intel.com/software/products/>

38

基礎性能比較 SpecCPU2000

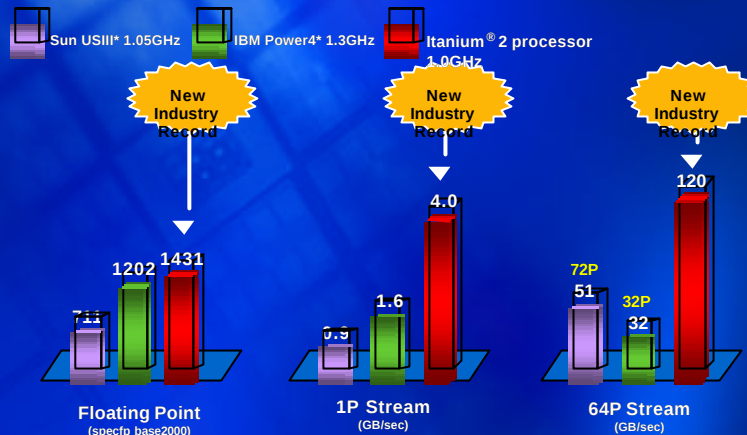


Results published at www.spec.org 12/25/02

intel.

39

科学技術計算性能比較 世界最速の性能



Source: <http://www.spec.org> for Sun results as of July 2, 2002. Itanium® 2 processor results measured on HP server using Itanium® 2 processor 1.0GHz with integrated 3MB L3 cache, Linux operating system, measured in the SPECint2000 and SPECfp2000.

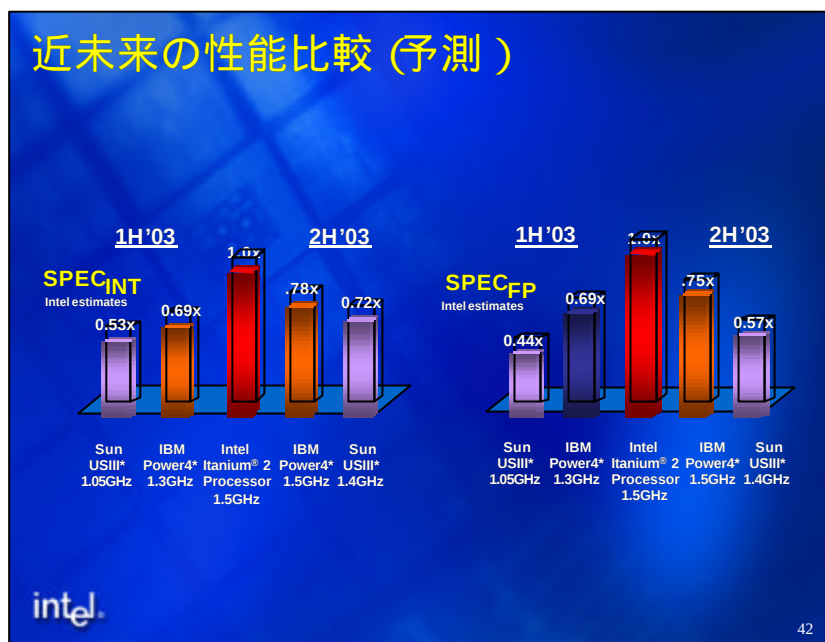
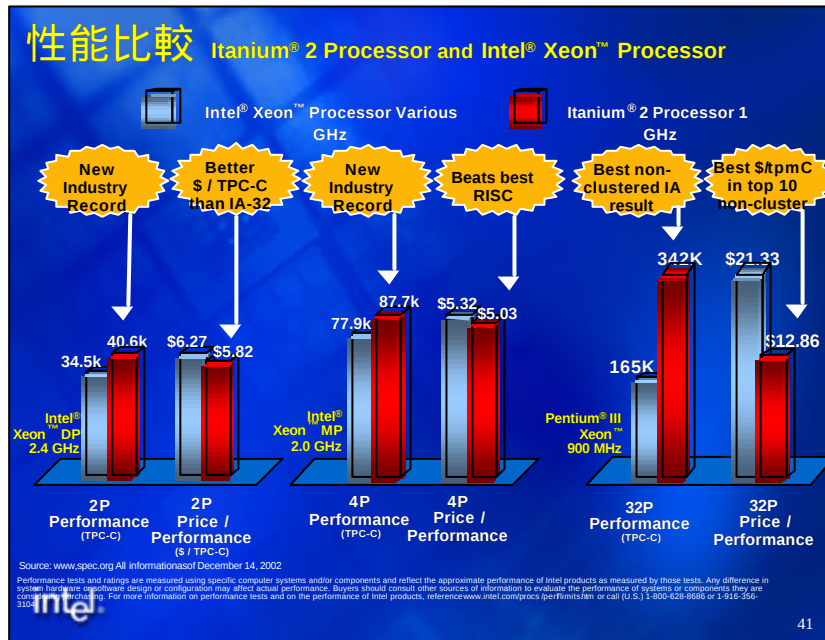
Source for Single processor Test: <http://www.spec.org> for Itanium® 2 processor and IBM Power4 results. Results obtained using the SPECint2000 and SPECfp2000.

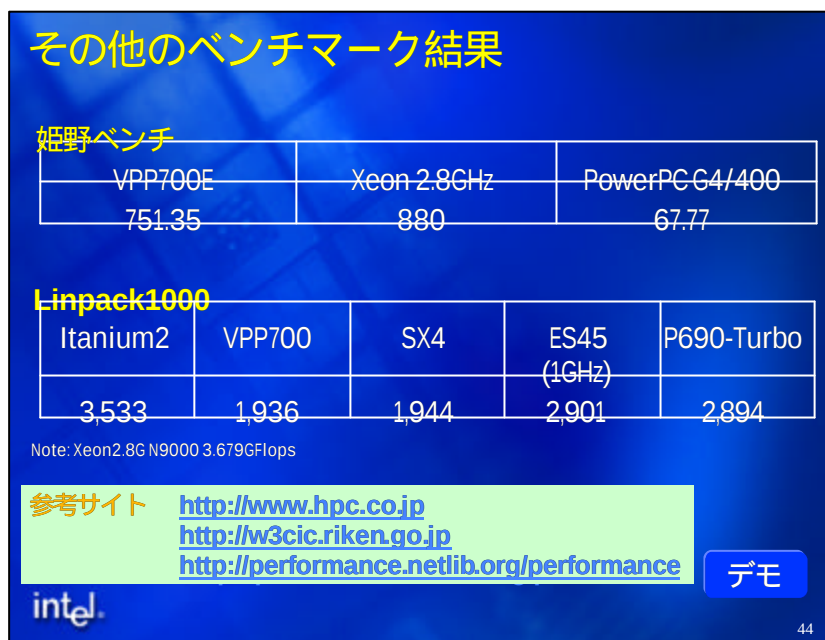
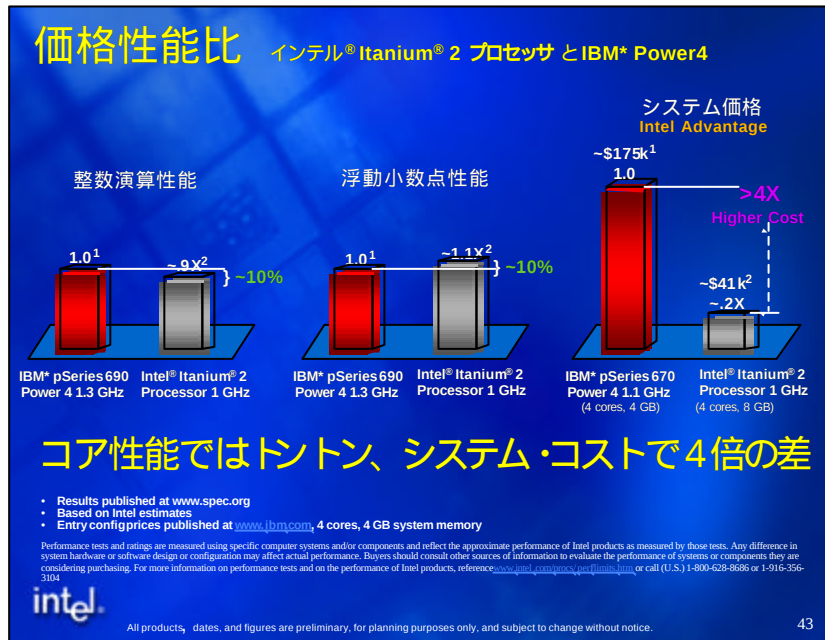
Source: Itanium® 2 processor measurements done on a 300 Sun Blade 1000 server using 64 Itanium® 2 processors, 12GB memory, Linux OS. Sun result of 120P from the SPECint2000 and SPECfp2000.

Performance tests and ratings are measured using specific computer systems and/or components and reflect the approximate performance of Intel products as measured by those tests. Any difference in system hardware or software design or configuration may affect actual performance. Buyers should consult other sources of information to evaluate the performance of systems or components they are considering purchasing. For more information on performance tests and on the performance of Intel products, reference www.intel.com/performance or call (U.S.) 1-800-426-0606 or 1-916-356-3104.

intel.

40





姫野ベンチのTop 20

Rank	MFLOPS	Maker	Machine	CPU	Clock	Compiler	Options
1	4705	NEC	SX-5	10GFLOPS	312.5MH	-	"-Chopt"
2	4401	FUJITSU	VPP5000	9.6GFLOPS	300MHz	VPP FORTRAN	"-O4 -Wv,-Me"
3	4282	NEC	SX-5	10GFLOPS	312.5MH	-	"-"
4	4282	NEC	SX-5	10GFLOPS	312.5MH	-	"-"
5	4270	HITACHI	SR8000/E1	9.6GFLOPS	-	-	"-"
6	4137	FUJITSU	VPP5000	9.6GFLOPS	300MHz	VPP FORTRAN	"-"
7	4137	FUJITSU	VPP5000	9.6GFLOPS	300MHz	VPP FORTRAN	"-"
8	3785	NEC	SX-5	8GFLOPS	250MHz	-	"-Chopt"
9	3701	FUJITSU	VPP800	8GFLOPS	250MHz	VPP FORTRAN	"-O4 -Mv,-Me"
10	3701	FUJITSU	VPP800	8GFLOPS	250MHz	VPP FORTRAN	"-O4 -Mv,-Me"
11	3566	HITACHI	SR8000	8GFLOPS	-	-	"-"
12	3395	NEC	SX-5	8GFLOPS	250MHz	-	"-"
13	3395	NEC	SX-5	8GFLOPS	250MHz	-	"-"
14	3262	FUJITSU	VPP800	8GFLOPS	250MHz	VPP FORTRAN	"-O4"
15	1722	NEC	SX5/S1	4GFLOPS	250MHz	f90	"-"
16	1722	NEC	SX5/S1	4GFLOPS	250MHz	f90	"-"
17	751	FUJITSU	VPP700E	2.4GFLOPS	-	FORTTRAN + MP	"-"
18	750	-	Self-made	Pentium4	2GHz (1.8GHz +)	Intel C/C++	"-O3 -tpp7"
19	685	-	Self-made	Pentium4	2GHz	Intel C/C++	"-O3 -tpp7"
20	652	-	Self-made	Pentium4	1.8GHz	Intel C/C++	"-O3 -tpp7"

intel.

*Other names and brands may be claimed as the property of others

45

2002年理研シンポジウム特別賞

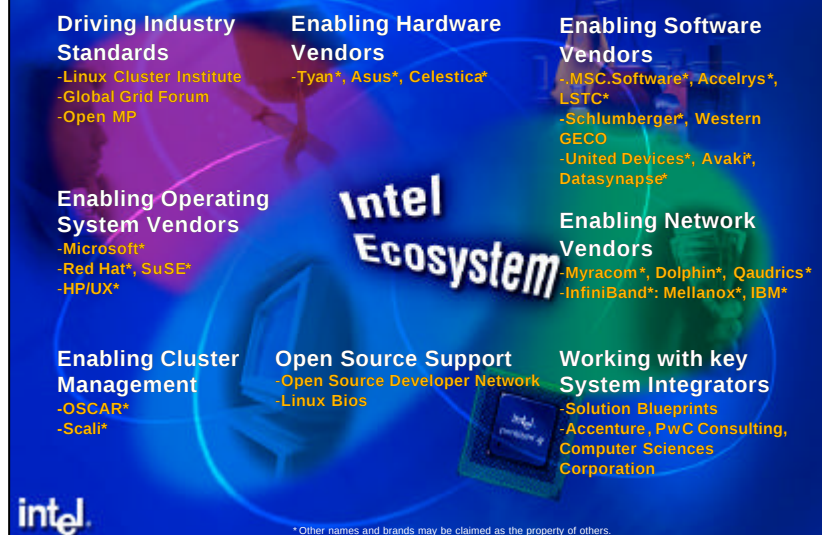
- スカラ・プロセッサで初めて1GFLOPS達成
- 1688.314MFLOPS



intel.

46

インテルのHPC業界への貢献



まとめ

- これからもCPU性能を向上
 - Itaniumプロセッサファミリの強化
 - Xeonプロセッサファミリの強化
- 汎用技術を利用したクラスタ技術の推進
 - CPUの性能を生かす高性能チップセット
 - PCI-Expressによるプラットフォーム内高速接続
 - InfiniBand*