

将来(1ペタフロップス超・2010年代前半を想定)の
超高速計算機に必要な要素技術の研究開発について

平成16年8月
(株)日立製作所

Contents

1 将来の超高速計算機システムについて

2 アプリケーションソフトウェアについて

3 必要な要素技術について

Contents

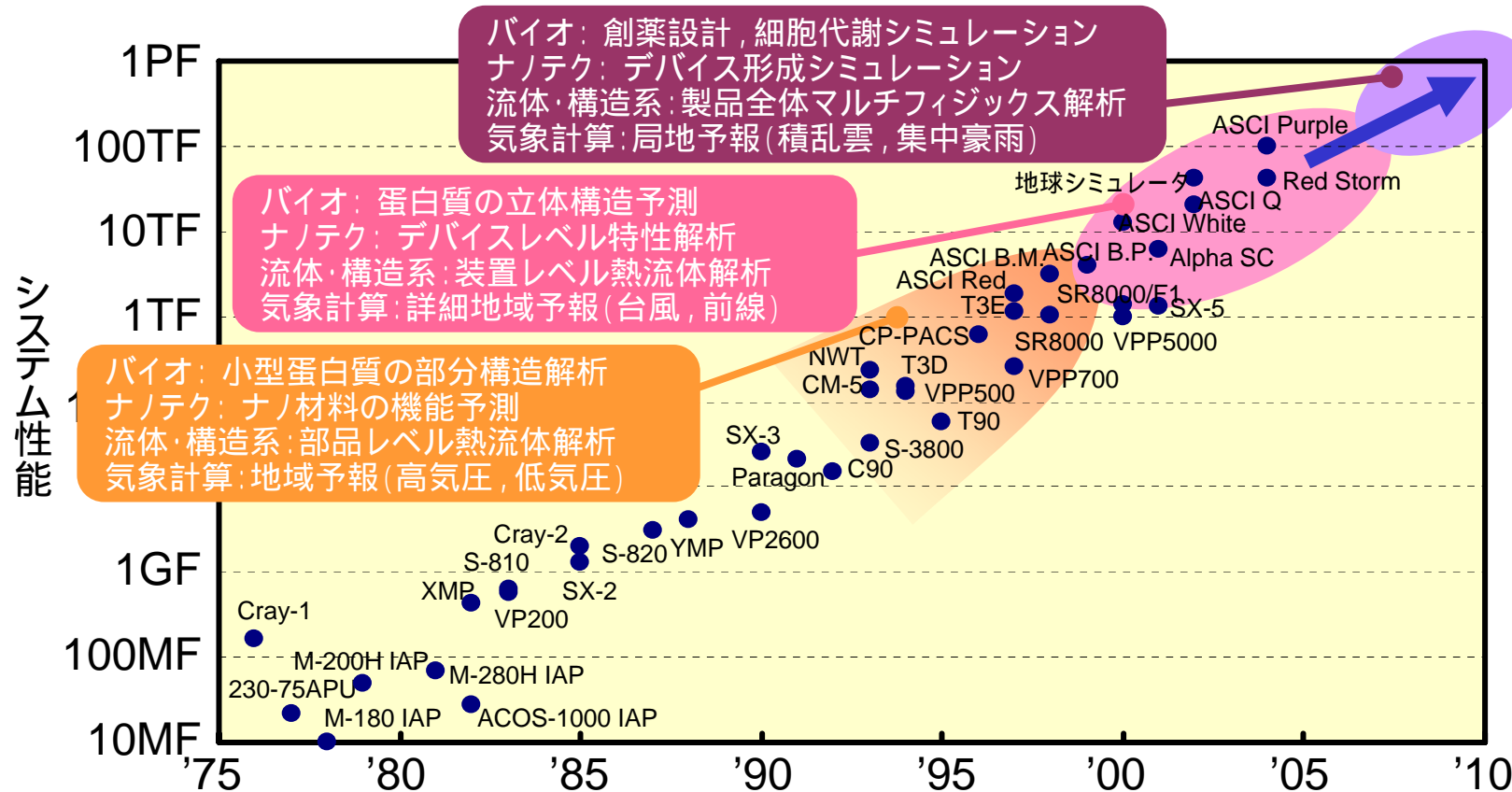
1 将来の超高速計算機システムについて

2 アプリケーションソフトウェアについて

3 必要な要素技術について

1 -1. 背景

- 2010年代の超高速計算機はペタフロップスクラスまで拡大
- 広範囲のアプリケーションに適用できる超高速計算機の実現は日本の科学技術, 産業分野における国際競争力向上につながると予想



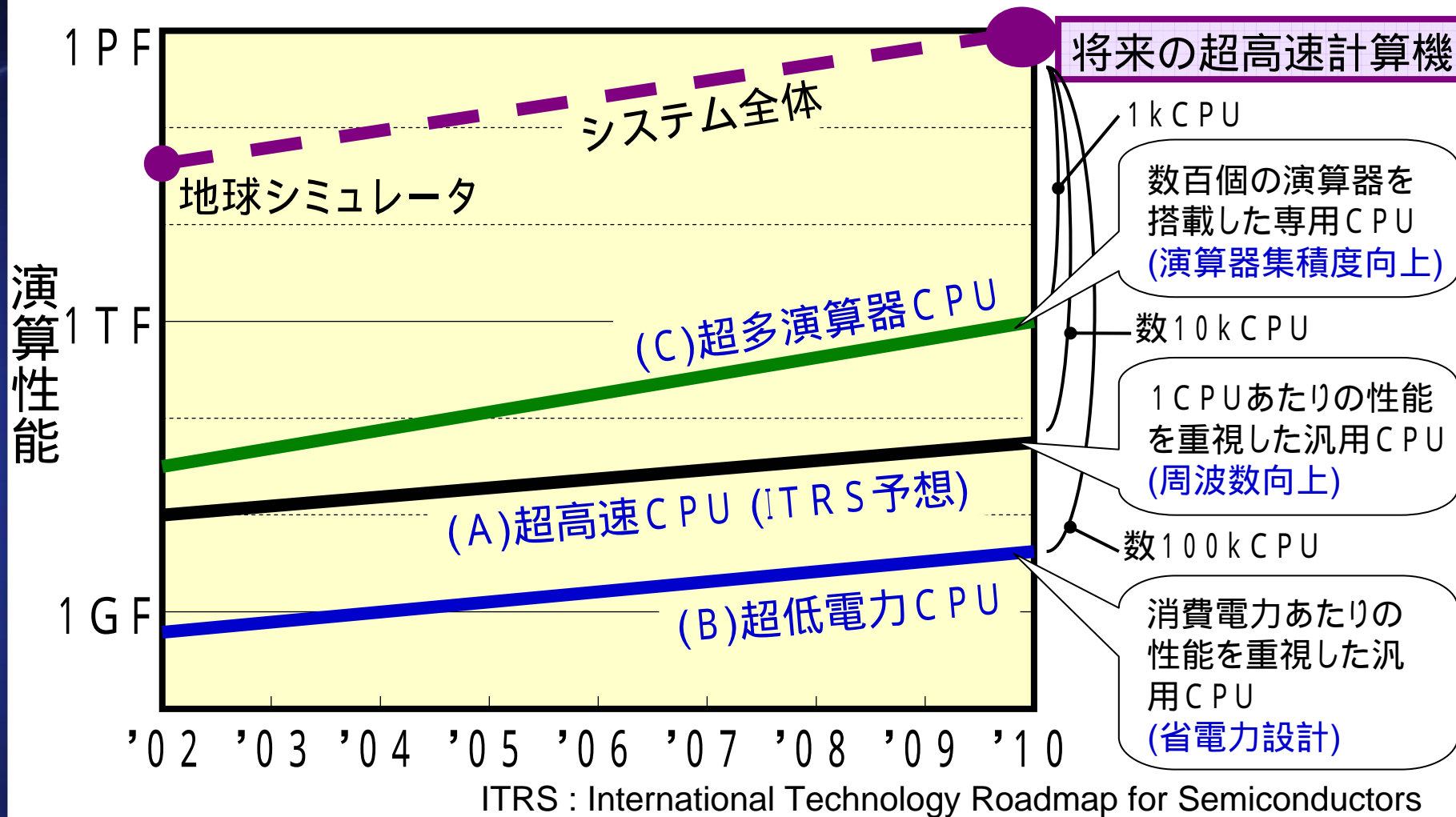
1 -2. 検討方針

- 2010年代の超高速計算機に必要となるブレークスルー技術をテクノロジトレンドより検討
- 以下に示すシステム特長の実現を前提とした
(今後, WG方針に従いターゲットシステム仕様は修正していく)

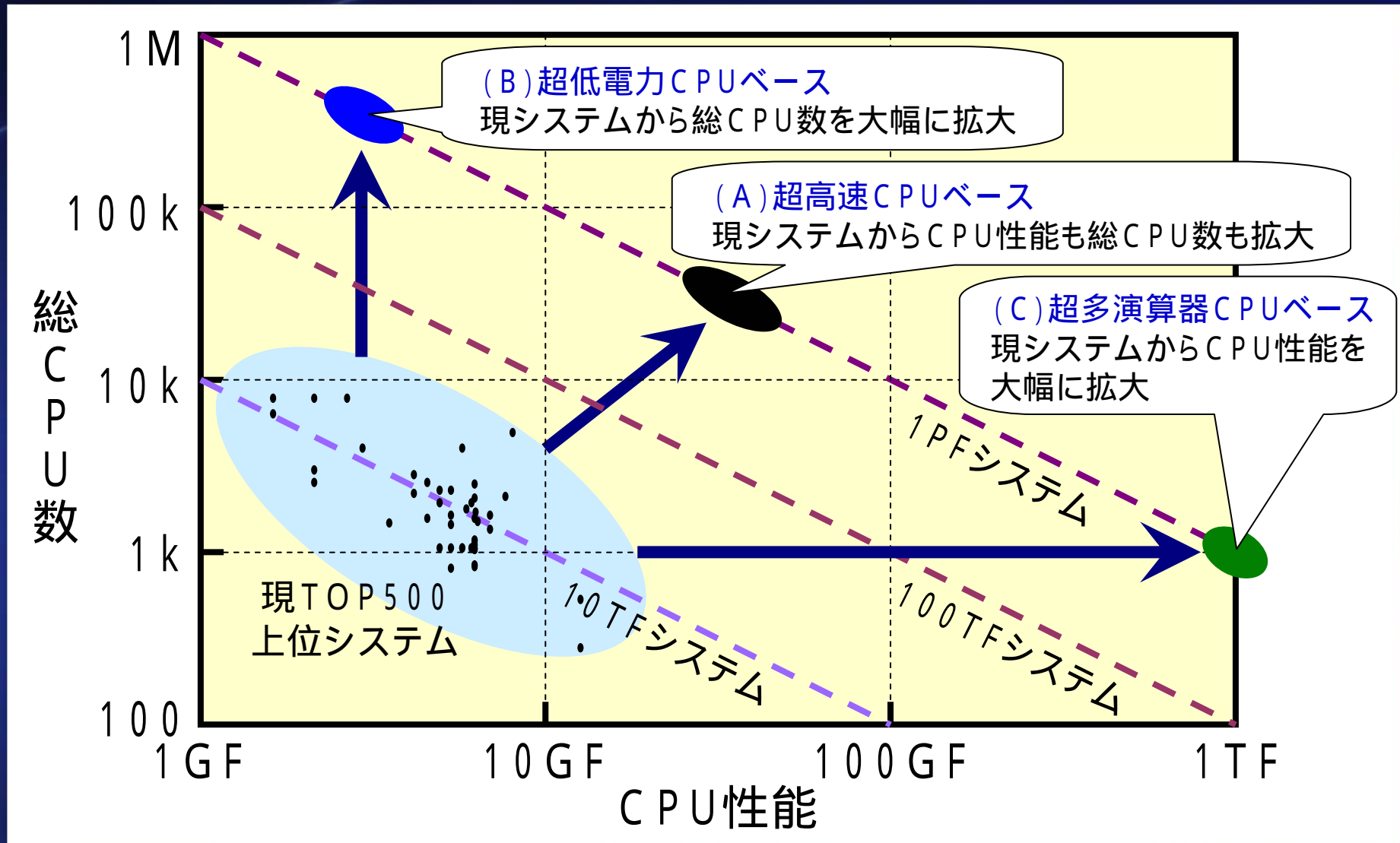
	システム特長	システム特長を支える主なハードウェア仕様
1	多種多様な分野のプログラムを実行できること	・0.5PB以上のシステム搭載メモリ ・ハイエンドサーバ並みの共有メモリ容量 ・Grid対応(数100GB/s以上の外部ネットワーク)
2	高い実効性能を発揮すること(効率30~50%目標)	・GFLOPSあたり2GB/s以上の実効メモリバンド幅 ・GFLOPSあたり0.1GB/s以上のシステム内ネットワークバンド幅
3	システム諸元が地球シミュレータと同程度に収まること	・消費電力:10MWクラス ・物理構成:500ラック程度

1-3. 超高速計算機のシステム・CPU性能トレンド

- 超高速計算機に適用するCPUとして典型的な3タイプを想定

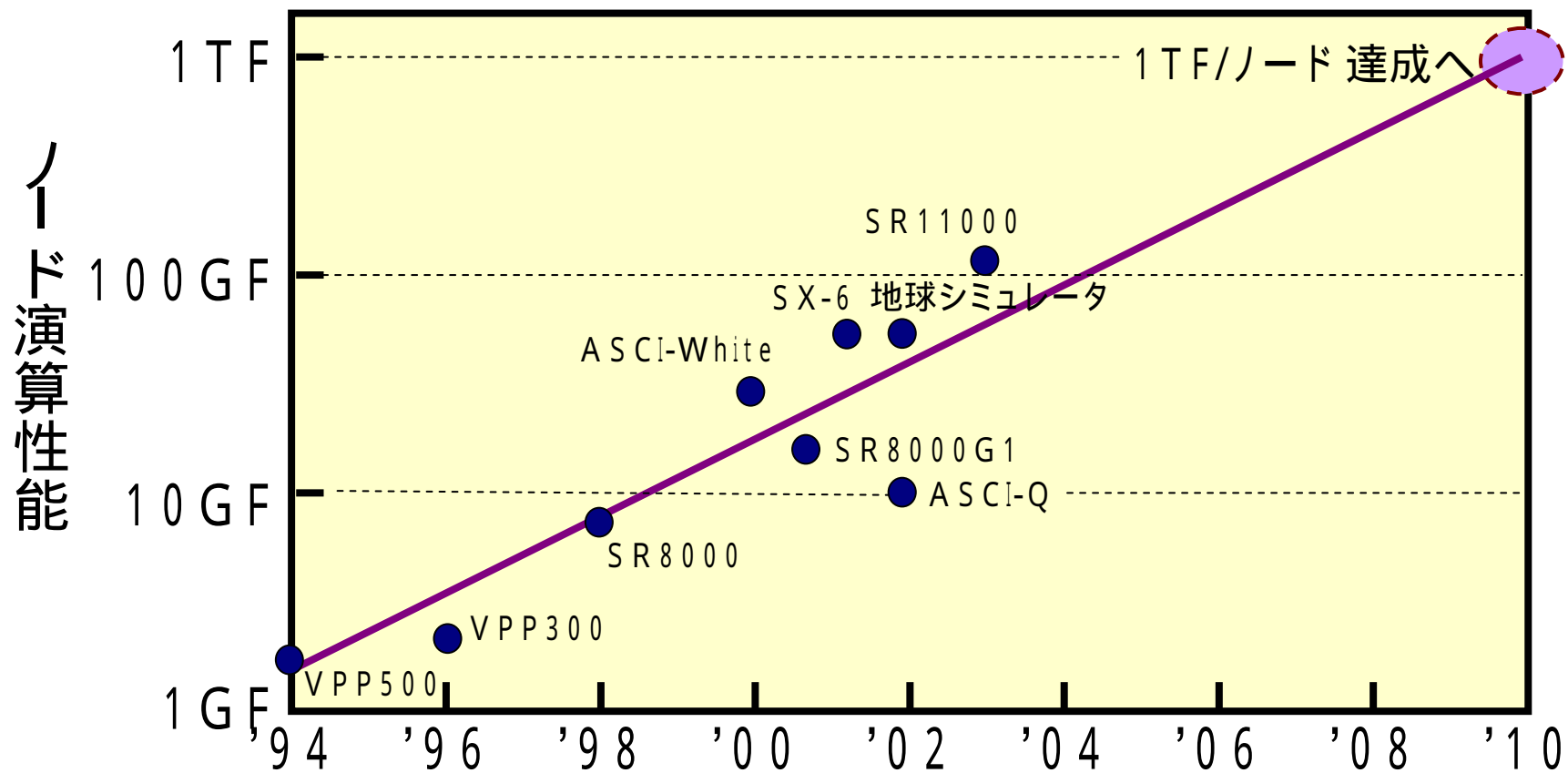


1 -4. 将来の超高速計算機と現システムとの関係

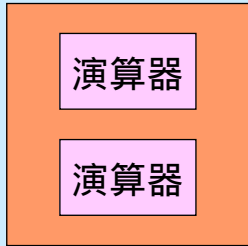
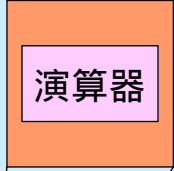
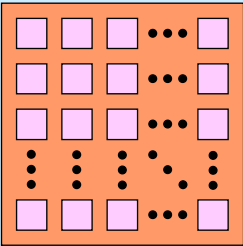
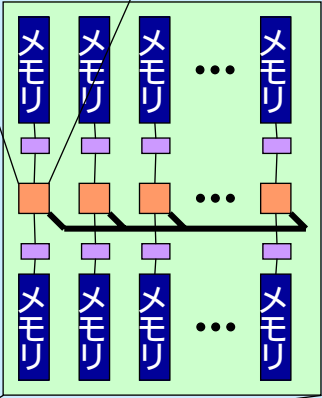
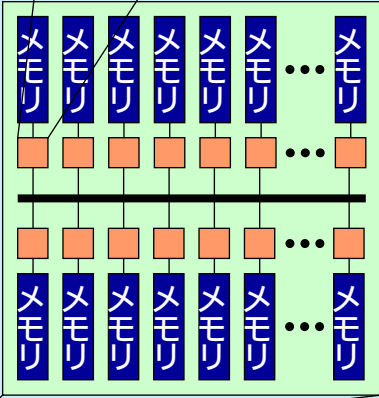
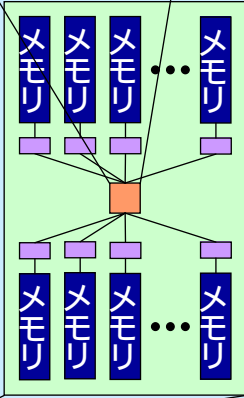
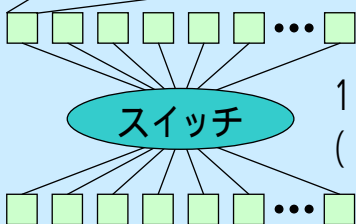
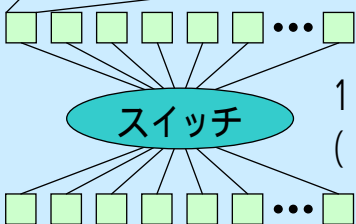
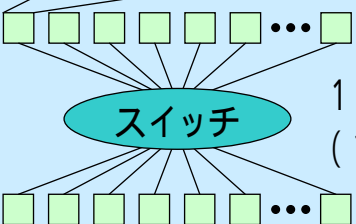


1 -5. ノード性能のトレンド

- ノード：メモリを共有するCPU群(システムの構成単位)
- 2010年のノード性能は1TFクラスとなると予測



1 -6. CPUタイプとシステム構成概要(ハードウェア)

項目	(A)超高速CPUベース	(B)超低電力CPUベース	(C)超多演算器CPUベース
プロセッサ	 <p>高周波数による高性能を重視 (数10GF)</p>	 <p>消費電力の低減を重視 (数GF)</p>	 <p>数百個の演算器を集積 (1TF)</p>
ノード	<p>数10CPU(1TF)</p> 	<p>数100CPU(1TF)</p> 	<p>1CPU(1TF)</p> 
ノード間ネットワーク	 <p>1kノード (1PF)</p>	 <p>1kノード (1PF)</p>	 <p>1kノード (1PF)</p>

1 -7. ハードウェアの課題(1)

項目	(A)超高速CPUベース	(B)超低電力CPUベース	(C)超多演算器CPUベース
プロセッサ	<ul style="list-style-type: none"> ・10 ~ 16 GHz ・40 ~ 64 GF ・200Wクラス 	<ul style="list-style-type: none"> ・1.5 ~ 2 GHz ・3 ~ 4 GF ・5Wクラス 	<ul style="list-style-type: none"> ・1.5 ~ 2 GHz ・1TF ・300Wクラス
<p>1 低消費電力と高性能との両立(デバイス, 論理, アーキテクチャ, ソフト連携)</p>			
	<ul style="list-style-type: none"> ・ブリッジ経由でメモリ接続 	<ul style="list-style-type: none"> ・プロセッサ - メモリ直結 	<ul style="list-style-type: none"> ・ブリッジ経由でメモリ接続
<p>3 超高速伝送(電気 or 光)</p>			
	(超高速電気 or 光伝送 20 Gbpsクラス)	(超高速電気 or 光伝送 20 Gbpsクラス)	(超高速電気 or 光伝送 100 Gbpsクラス)
ノード	<ul style="list-style-type: none"> ・1TF (20 CPU程度) ・メモリ512GB 	<ul style="list-style-type: none"> ・1TF (300 CPU程度) ・メモリ512GB 	<ul style="list-style-type: none"> ・1TF (1 CPU) ・メモリ512GB
<p>4 多CPU結合構成</p>			
	<ul style="list-style-type: none"> ・実効バンド幅2 B/F程度 ・レイテンシ1000 pcy ~ 	<ul style="list-style-type: none"> ・実効バンド幅2 B/F程度 ・レイテンシ200 pcy ~ 	<ul style="list-style-type: none"> ・実効バンド幅2 B/F程度 ・レイテンシ250 pcy ~
<p>5 高バンド幅・低レイテンシメモリ構成</p>			
	<ul style="list-style-type: none"> ・6 kWクラス 	<ul style="list-style-type: none"> ・4 kWクラス 	<ul style="list-style-type: none"> ・3 kWクラス

2 超多演算器アーキテクチャ

数字 ブレークスルー必須 主要項目

1 -8. ハードウェアの課題(2)

項目	(A)超高速CPUベース	(B)超低電力CPUベース	(C)超多演算器CPUベース
ノード間ネットワーク	・100GB/s×2(双方向) レイテンシ3μs程度	・100GB/s×2(双方向) レイテンシ3μs程度	・100GB/s×2(双方向) レイテンシ3μs程度
	6 超高バンド幅低レイテンシネットワーク		
	・多段 or 完全クロスバ	・多段 or 完全クロスバ	・多段 or 完全クロスバ
システム構成	・1PF ・20k CPU程度 ・メモリ 0.5PB ・1000ノード ・400~500ラック	・1PF ・300k CPU程度 ・メモリ 0.5PB ・1000ノード ・400~500ラック	・1PF ・1k CPU程度 ・メモリ 0.5PB ・1000ノード ・400~500ラック
	・トータル10MW超クラス	・トータル10MWクラス	・トータル10MWクラス
	8 10MWクラス実装・冷却技術		
I/O	・1TB/s程度(ストレージ) ・数百GB/s(外部NW)	・1TB/s程度(ストレージ) ・数百GB/s(外部NW)	・1TB/s程度(ストレージ) ・数百GB/s(外部NW)
9 I/Oインタフェース高速化技術			
高信頼化技術	・Non-Stop技術	・Non-Stop技術	・Non-Stop技術
10 超高信頼化技術(ハード+OS)			

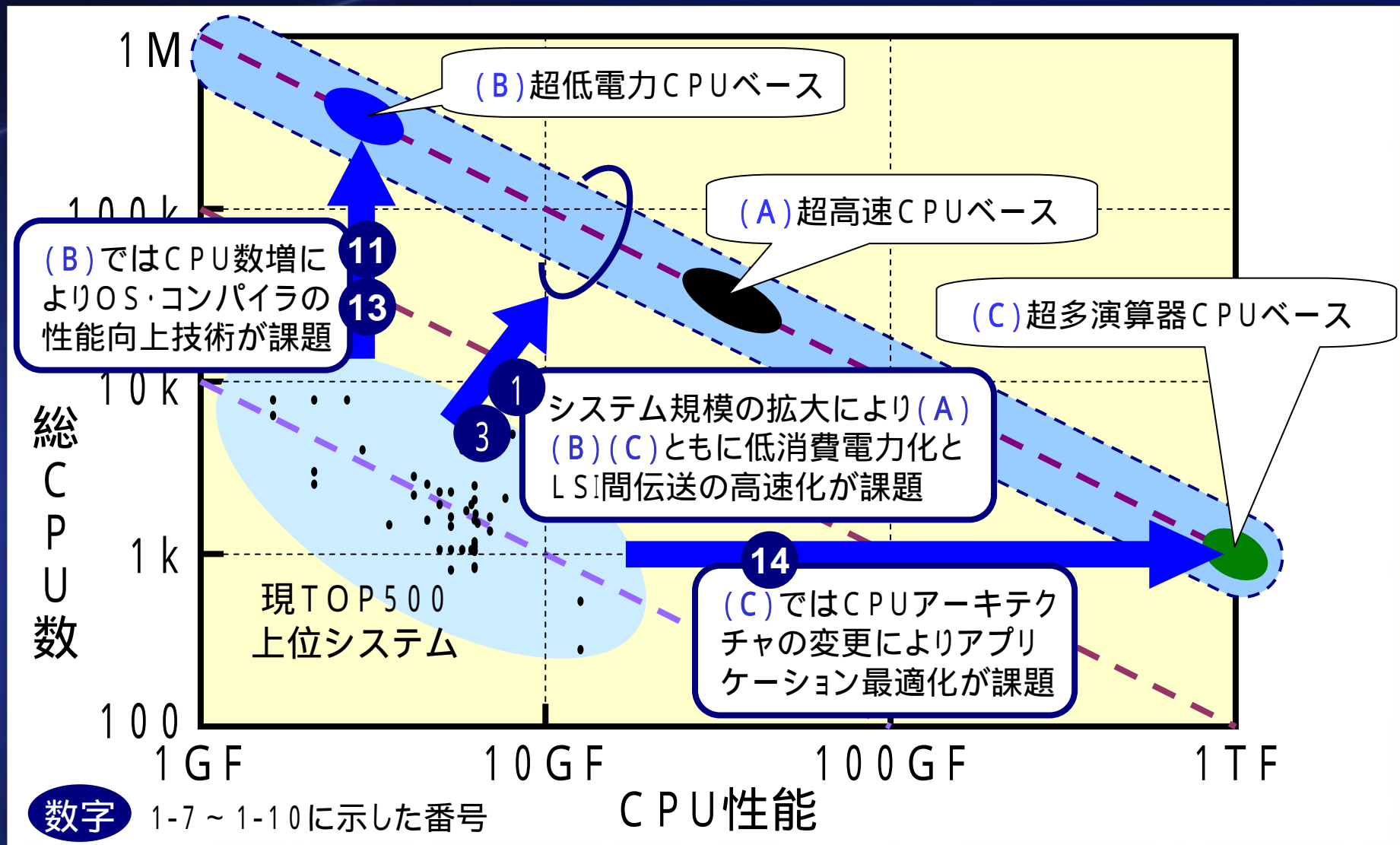
1 -9. ソフトウェアの課題

項目	(A)超高速CPUベース ・ノード 20CPU程度 ・1000ノード	(B)超低電力CPUベース ・ノード 300CPU程度 ・1000ノード	(C)超多演算器CPUベース ・ノード 1CPU ・1000ノード
OS (ノード内)	・汎用OSベース	・汎用OSベース (高多重SMP対応)	・汎用OSベース (専用CPU向けカスタム化)
11 ノード内スケラビリティ向上技術(OS)			
OS 及び クラスタ運 用管理ミド ルウェア (ノード間)	[1システム化実現時] ・超並列1システムOS [nシステム+ミドルウェア] ・スケラブル1システム ミドルウェア	[1システム化実現時] ・超並列1システムOS [nシステム+ミドルウェア] ・スケラブル1システム ミドルウェア	[1システム化実現時] ・超並列1システムOS [nシステム+ミドルウェア] ・スケラブル1システム ミドルウェア
12 ノード間スケラビリティ向上技術(OS+ミドルウェア)		14 専用CPU活用技術	
プログラミ ング言語	・レイテンシ隠蔽指向ノード 内自動並列化コンパイラ	・階層型ノード内自動 並列化コンパイラ	・専用CPU向けノード内自動 並列化コンパイラ
13 ノード内スケラビリティ向上技術(コンパイラ)			
開発ツール	・並列アプリ統合開発環境	・並列アプリ統合開発環境	・専用CPU向け統合環境
アプリケー ション	・高効率アルゴリズム	・高効率アルゴリズム	・専用CPU向けアルゴリズム ・高効率アルゴリズム

1 -10. ブレークスルー必要技術一覧

課題	カテゴリ
① 低消費電力と高性能との両立	デバイス,アーキテクチャ,ソフト
② 超多演算器アーキテクチャ	アーキテクチャ
③ 超高速伝送(光 or 電気)	伝送,実装
④ 多CPU結合構成	実装,アーキテクチャ
⑤ 高バンド幅・低レイテンシメモリ構成	実装,アーキテクチャ
⑥ 超高バンド幅低レイテンシネットワーク	伝送
⑦ 多ノードシステム向けスケーラブルネットワークポロジ	アーキテクチャ
⑧ 10MWクラス実装・冷却技術	実装
⑨ I/Oインターフェース高速化技術	伝送
⑩ 超高信頼化技術(ハード+OS)	アーキテクチャ,ソフト
⑪ ノード内スケーラビリティ向上技術(OS)	ソフト
⑫ ノード間スケーラビリティ向上技術(OS+ミドルウェア)	ソフト
⑬ ノード内スケーラビリティ向上技術(コンパイラ)	ソフト
⑭ 専用CPU活用技術	ソフト,アプリケーション

1-11. 主なブレークスルー必要技術の位置づけ



Contents

1 将来の超高速計算機システムについて

2 アプリケーションソフトウェアについて

3 必要な要素技術について

2 -1. 超高速計算機向けアプリケーションの動向

- 各種解析の個別実行から連携解析による製品開発力強化へ
- ミクロ材料物性からマクロ現象までの現象解明による創造提案へ
- 扱う現象の大規模化と複雑化に伴い計算量爆発の方向

各種従来手法

電磁場解析

音響解析

流体解析

構造解析

分子解析

バイオ解析

(1) 電気・機械系の競争力強化

いかに早く製品を創生するか

全体系マルチフィジックス連携解析

(2) ナノ・バイオ系の新分野開拓

何を創生するか, どう解析するか

ナノ-マクロ マルチスケール解析

2 -2. PFLOPSシステム向けアプリケーション(1)

対象分野	アプリケーション	概要	目的	期待されるブレイクスルー	経済的波及効果	
マルチスケール バイオ	ライフサイエンス	創薬設計シミュレーション (細胞代謝シミュレーション)	蛋白質と化合物の結合エネルギーを高精度に計算	均一な薬効設計と副作用低減	テーラード医療の実現	医薬品業界の国際競争力の向上
	有機分子応用製品	有機分子自己組織化シミュレーション	量子論自己組織化過程を計算	バイオチップなどの高機能設計	生体分子を応用した新概念製品実現	機能性デバイス市場の開拓・拡大
マルチスケール ナノテク	燃料電池	触媒化学反応シミュレーション	触媒存在下での化学反応ダイナミクスを計算	低コスト・高効率触媒材料設計	燃料電池コストの実用レベルへの到達	エネルギー消費構造の転換
	半導体プロセス, ディ스플레이	成膜プロセスシミュレーション	ナノスケール加工, 高分子液膜形成などの各種プロセスを計算	新デバイス・新ディスプレイの最適プロセス構築	プロセス開発の圧倒的時短・コスト短縮	半導体業界, ディ스플레이業界の国際競争力の向上

2 -3. PFLOPSシステム向けアプリケーション(2)

対象分野		アプリケーション	概要	目的	期待されるブレイクスルー	経済的波及効果
ナノスケール (マルチスケール)	超高密度磁気記録材料・デバイス	極薄潤滑・薄膜物性シミュレーション	ヘッド・ディスク間界面現象, 新記録材料特性を分子動力学計算	潤滑状態や膜物性, 記録方式の高精度予測	超高密度記録方式の実現	ストレージ装置のユビキタス化, 国際競争力の向上
	機械自動設計	イメージベース最適化シミュレーション	ボクセルをフレームワークとした熱流体・構造振動解析	自動設計による開発期間の短縮	製品丸ごと解析の実現	製造業全体の国際競争力の向上
流体構造系 (マルチフィジックス)	燃焼・ターボ機械	非定常燃焼・圧縮性流体シミュレーション	燃焼反応を伴う流体現象を高精度に計算	各種高効率流体機械(ガスタービン等)の開発	試作レス設計の実現による開発期間の大幅短縮	エネルギー・環境社会インフラの高度化, 国際競争力向上
気象計算 (マルチフィジックス)		高精度気象シミュレーション (メソモデル)	非静力学的地域気象予測	局地的異常気象の予測	集中豪雨予報による地域災害の軽減	都市の安全性向上

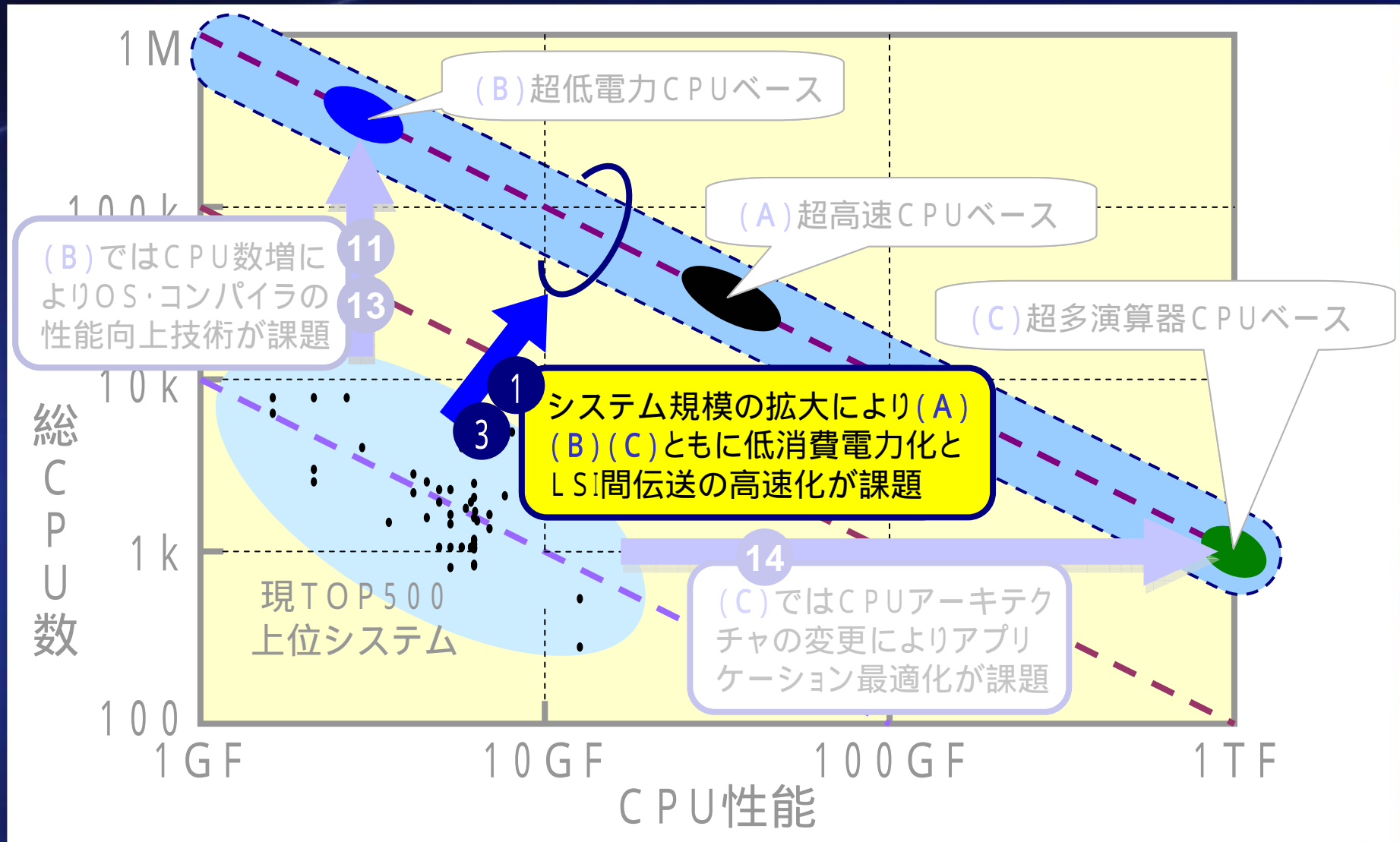
Contents

1 将来の超高速計算機システムについて

2 アプリケーションソフトウェアについて

3 必要な要素技術について

3-1. 主なブレークスルー必要技術(ハードウェア)



3 -2. 高性能・低電力システムの現在の技術の限界

1. ITRS (ロードマップ) にみる現在と将来のLSI

- 2004年: 90nmプロセス, ゲート長37nm, 4GHz
- 2010年: 45nmプロセス, ゲート長18nm, 15GHz

2. 問題点

- 微細化に成功しても, 性能, 消費電力とも目標未達
- リーク電流による消費電力増大が深刻化

従来のデバイス・回路の個別対策から, デバイス・回路・アーキテクチャ・ソフトの広範囲に連携した対策が必須

3. 期待されるブレークスルー

- 消費電力を増加させないデバイス高速化手法
- 消費電力を増加させないデバイス・回路・アーキテクチャ・ソフト連携による高速化手法

3 -3. 高性能・低電力システムの研究の現状

デバイスレベルでの低電力化・高性能化

- (1) 移動度向上(歪みSiなど)
- (2) 短チャネル効果の抑制手法(SOI, マルチゲート構造など)
- (3) パラメータ制御手法(しきい電圧制御など)
- (4) 新素材(Low-k, High-k)

デバイス・回路・アーキテクチャ・ソフト連携による低電力化

- (1) 動作回路の最小化(Clock Gatingなど)

実装技術による高性能化

- (1) 低消費電力高速伝送技術
- (2) 低損失素材(ケーブル, 基板, LTCC)
- (3) 低温冷却CMOS

SOI: Silicon On Insulator

LTCC: Low Temperature Co-fired Ceramic 低温焼成セラミックス

3 -4. 高性能・低電力システムに必要な研究開発項目

	技術項目	概要	主な課題
1	しきい電圧可変デバイス	MOSトランジスタのしきい電圧を制御してリーク電流を低減する	デバイス
2	動作速度最適化	高性能が必要な部分のみ高周波数・高速動作させる(#1との連携)	デバイス
3	オンチップスイッチ	チップ上のスイッチ・トランジスタにより動作不要部分の電源電圧の供給を止める	回路
4	低電力高速I/F	I/F回路の最適化により消費電力を増やさずにLSI間の超高速伝送を実現する	回路
5	並列処理	複数の機能ブロックの並列処理により、動作周波数を高めずに処理性能を向上する	アーキテクチャ
6	ソフトウェア制御	OSやアプリケーションの負荷に応じたシステム制御を動的におこなう	ソフトウェア

上記デバイス・回路・アーキテクチャ・ソフトを統合する技術の開発が必要

3 -5. 超高速伝送技術(電気)

1. ITRS(ロードマップ)にみる現在と将来の超高速伝送

- 2004年: 3.2 Gbps (差動信号線あたり)
- 2008 ~ 2013年: 10 Gbps (差動信号線あたり)

2. 課題

- LSI間伝送の性能向上には, 高速化と高密度化の両立が必要
- 数Gbps以上の高速化実現には多値伝送技術が有効
- 数100pinクラスの高密度化実現には低電力化が必須
- ケーブル, 基板, コネクタなど超高速伝送向け実装技術も必要

3. 期待されるブレイクスルー

- 超高速低電力多値伝送技術(差動信号線あたり数10Gbps級)
- 超高速実装技術(ケーブル, 基板, コネクタ)
- 超高速伝送を活かすデータ伝送方式(プロトコル等)

4. 技術の適用分野

- ITシステム(サーバ, ストレージ, ネットワーク), 医療機器など画像処理装置

3 -6. ノードアーキテクチャ (CPU結合構成、メモリ構成)

1. 課題

- システム全体の実効性能に対して、ノード性能の影響大
- 広範囲のアプリケーションで高いノード性能の達成が必須
- 汎用性の高いハードウェアアーキテクチャと、ハードウェアの特長を生かすコンパイラとを同時に検討することで、様々な特性のアプリケーションで高性能達成を狙う

2. 期待されるブレークスルー

- コンパイラ技術と連携した高効率ノードアーキテクチャ
 - (A)超高速CPUベース：メモリレイテンシ隠蔽方式
 - (B)超低電力CPUベース：多CPU結合方式
 - (C)超多演算器CPUベース：メモリバンド幅向上方式

3. 技術の適用分野

- 高性能サーバなど

3-7. 要素技術の研究開発スケジュール

