

秘

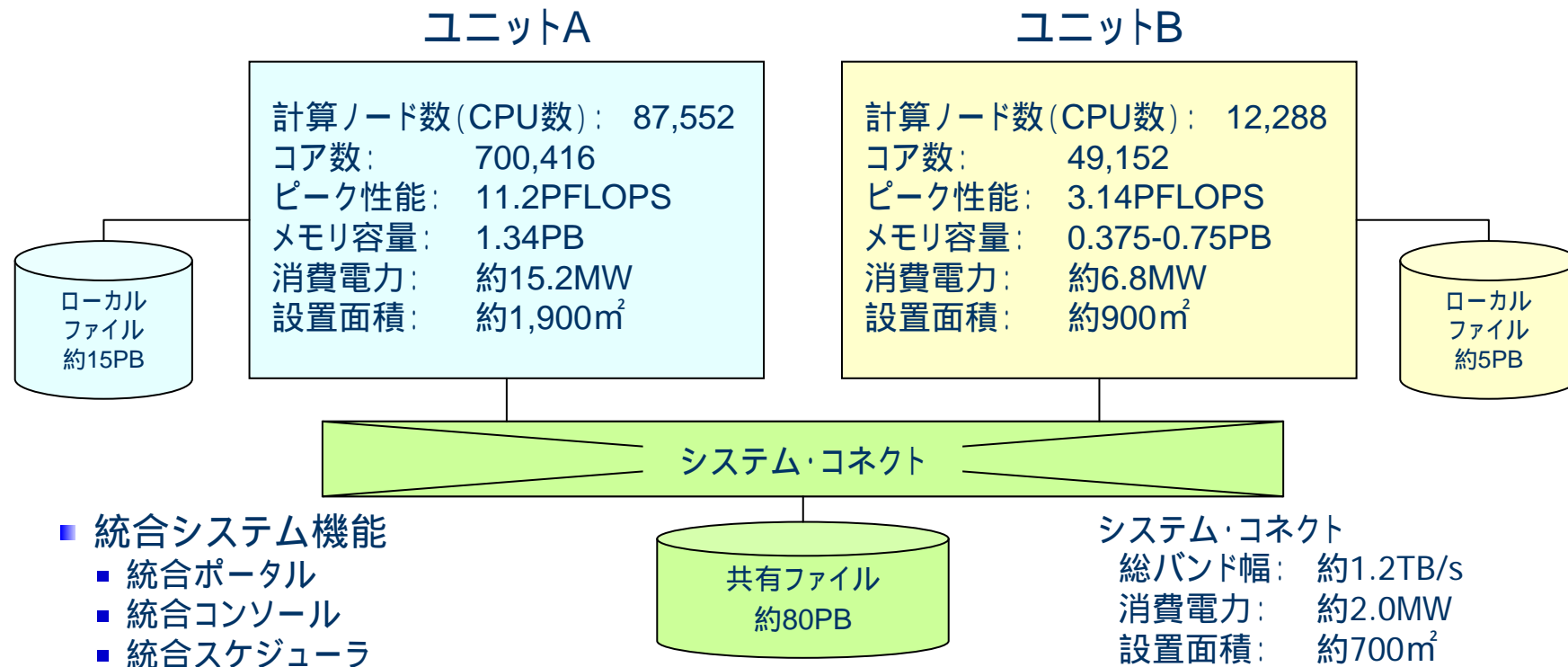
回収資料

(資料6-3別紙)

システム構成と要素技術について

統合汎用スーパーコンピュータシステムの構成

- ノード数 (CPU数): 99,840
- コア数: 749,568
- ピーク性能: 14.3PFLOPS
- メモリ容量: 1.7-2.1PB
- 磁気ディスク容量: 約100PB
- 消費電力: 約24MW (空調を除く)
- 設置面積: 約3,800m² (空調を除く)
- 電力性能比: 約1.68MW/PFLOPS
- 面積性能比: 約266m²/PFLOPS

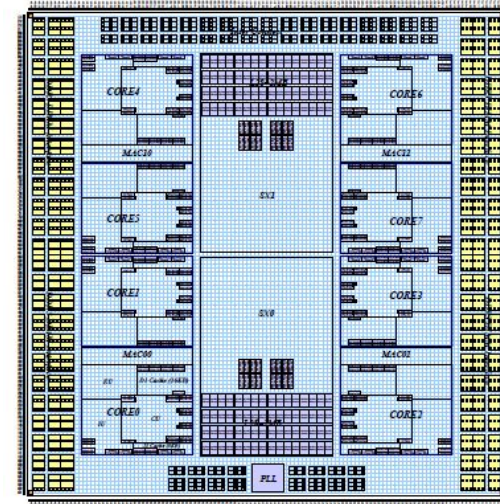


両ユニットの特長

	ユニットA	ユニットB
プロセッサ	<ul style="list-style-type: none"> ■ 1CPUチップ当り128GFLOPSのSIMD演算器 ■ 1CPU当り8コア構成, 動作周波数2GHzで駆動 ■ コア当りSIMD拡張演算器(4FMA, 4逆数近似等), FPレジスタ128本 (SPARC-V9規格の4倍) ■ 6MBのL2キャッシュを8コアで共有, ハードバリア機構 ■ 命令リトライによる高信頼性 ■ メモリバンド幅: 0.5B/FLOPS 	<ul style="list-style-type: none"> ■ 1CPUチップ当り256GFLOPSの高性能ベクトル演算器 ■ 1CPU当り4コア構成, 動作周波数2GHzで駆動 ■ コア当り8FMAx2セットの演算器と128KBの大容量ベクトルレジスタ ■ 8MBのL2キャッシュを4コアで共有しソフトウェアでも制御可能としたRDB (Reusable Data Buffering) 機能 ■ メモリバンド幅: ~ 1B/FLOP ■ 最大1TBまでの共有メモリ (NUMA) が利用可能
ネットワーク	<ul style="list-style-type: none"> ■ 18CPUを1セットとしたシャシ内を完全結合, シャシ間を3次元トラス結合した独自新規方式. ■ 隣接通信を重視した設計思想 ■ 次元毎に2シャシ単位で直方体分割することによるパーティション運用 	<ul style="list-style-type: none"> ■ 2段Fat tree (バイセクションバンド幅98TB/s) で384 Nノードを接続 ■ 光インターコネクション・ネットワークの採用 ■ 非同期転送, 同報機能, 高速バリア同期機能付きのデータ転送機能 ■ 入出力ポートの構成制御によるパーティショニング
ピーク性能決定の考え方	<ul style="list-style-type: none"> ■ 目標性能であるLINPACKの実効性能10PFLOPSを達成することを目標に設定. ■ ナノ・デバイスの高精細度シミュレーション等を実行するためには, 10PFLOPS超の理論性能が必要 (実効性能 約3-5PFLOPS). ■ 多くのユーザが利用可能なPCクラスタや並列サーバによる開発・実行環境からの連続性, プログラム移植性を重視. 	<ul style="list-style-type: none"> ■ 雲解像大気大循環モデルによる気候変動予測等には, 理論性能3PFLOPS超が必要 (実効性能 約1PFLOPS). ■ 地球シミュレータなどベクトルプロセッサ向けのプログラム資産の有効利用を図る.

【ユニットA】論理LSIの開発要素技術 (1/2)

論理LSIフロアプラン



- **高性能技術**
 - 富士通45nm半導体プロセス技術
 - マルチコアLSI技術(8コア)
 - HPC向け機能拡張コア
 - コア当たり4つの浮動小数点積和演算器
 - SIMD機構
 - 共有2次キャッシュ(6MB)
 - 高速バリア機構
- 高いピーク性能(128GFLOPS)に加え, 実環境でも高い性能を発揮

- **低消費電力技術**
 - アクティブ電力, リーク電力削減技術
 - 低温動作
- 消費電力/性能を当社従来比10分の1
- 小型 ~ 大型/ブレードサーバへ適用可能

低消費電力技術	主な狙い (: 該当)	
	アクティブ電力削減	リーク電力削減
クロック制御		-
RAMの電力削減		-
省電力ラッチの採用		
Vth最適化	-	
トランジスタサイズ最適化		
キャパシタセルのゲートリーク削減	-	
チップごとのVdd, Vbs制御		

【ユニットA】論理LSIの開発要素技術 (2/2)

高信頼性回路技術

- 高信頼技術
 - メインフレーム用高信頼性回路技術
 - ハードによるエラー検出 & 自己回復
 - 低温動作
 - 故障率の低減
- ➔ ペタスケール (およそ10万個のプロセッサ相当)のシステム運用を実現

■ 論理LSI開発スケジュール



		種別	エラー検出方法
内蔵RAM	L1命令\$	データ部	パリティビット
		タグ部	パリティビット
	L1データ\$	データ部	SEC・DED ECC
		タグ部	パリティビット
	L2\$	データ部	SEC・DED ECC
		タグ部	SEC・DED ECC
命令・データ mTLB			パリティビット
	ブランチヒストリ	タグ部, データ部	パリティビット
実行部	レジスタ	GPR, FPR, GUB, FUB, PC, PSTATE, 演算入出力ラッチ等	パリティビット
	演算器	ALU, SHIFT, 除算器, グラフィック演算器, 比較器	パリティ予測回路
		乗算器, FMA	剰余チェック回路 パリティ予測回路

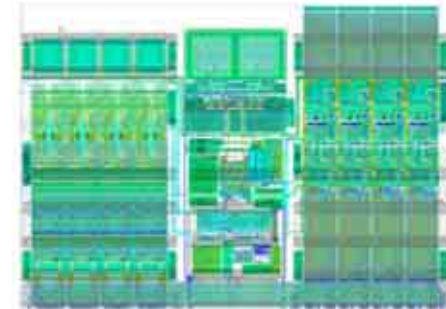
【ユニットA】インターコネクトの開発要素技術 (1/2)

■ シリアルI/Oマクロ技術

- 6.25Gbps高速シリアル伝送
- プリエンファシス, イコライゼーション技術の確立
- 小面積, 低消費電力を実現

→ 主流となりつつある, 高速シリアル伝送技術を発展
汎用シリアルバス方式の高速化に寄与

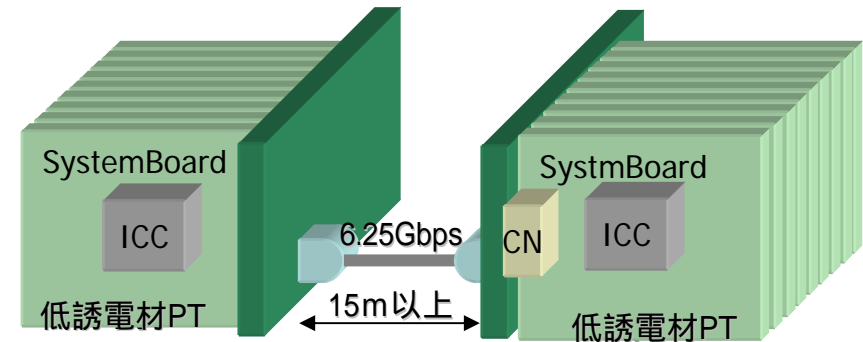
3.125Gbpsシリアルマクロ



■ 高速信号伝送技術

- 6.25Gbpsデータ伝送技術の開発
- 低損失コネクタの開発
- 低誘電材料を使用したPT板開発
- ラック間15m以上のデータ伝送を実現

→ 高速伝送が必要とされるIDCでの汎用サーバ, ブレードサーバへの応用



【ユニットA】インターコネクットの開発要素技術 (2/2)

■ ToFuインターコネクット

- 高帯域, 低遅延伝送プロトコル
- 耐故障性にすぐれたシステム運用機構
- 集合通信処理を高速化するMPI処理オフロード機構
- 大規模メモリシステムを実現する分散メモリ機構
- 論理ピーク性能100PetaFlopsを超えるスケーラビリティ

- ペタスケールコンピューティングに最適なインターコネクットの実現
 - HPCだけでなく汎用サーバへの転用も可能

■ インターコネクット開発スケジュール



【ユニットA】部品単位水冷技術 (1/2)

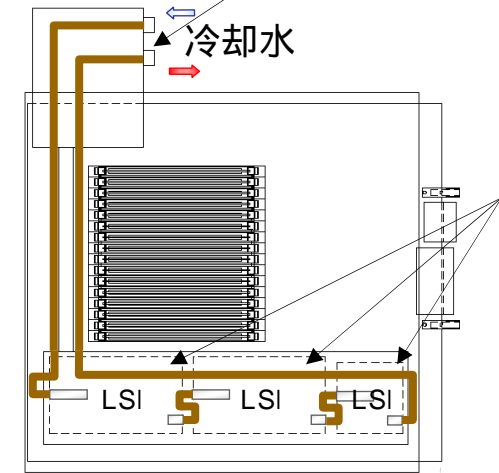
■ 高密度水冷機構

- 小型コールドプレート, 小型冷却水カプラを開発し, 従来のはない高密度実装を実現
- 空冷用ダクト空間削減による実装高密度化

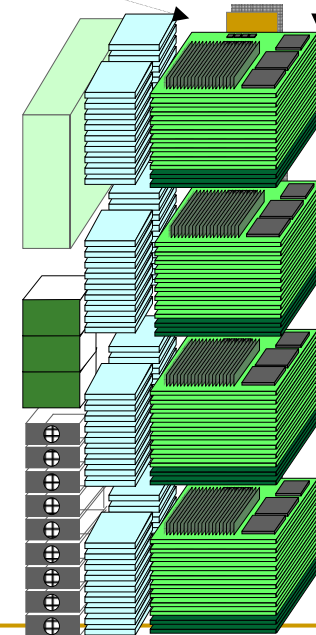
■ ボード活性交換との両立

- 故障したシステムボード(SB)の交換を, 他SBの稼働を継続したまま行える冷水制御, 漏水防止機構の開発

【システムボード(SB)】 冷却水カプラ



【計算機筐体】



コールドプレート

【ユニットA】部品単位水冷技術 (2/2)

- 半導体の稼働温度を下げて故障率を大幅に低減
 - 空冷装置と比較してCPUの固定故障率を約0.006倍へ低減(理論計算値)
- 半導体の稼働温度を下げてリーク電流を低減
- 水冷による高効率な冷却の実現
 - 効率良く熱を部屋外へ移動できるため、データセンターで生じている廃熱問題を解決
 - 冷却の高効率化は、計算機本体だけでなく空調設備の省電力化、静音化、省スペース化も実現
- 小型～大型/ブレードサーバへの展開
 - 省電力化、静音化、高密度実装へ寄与
- 部品単位水冷開発スケジュール

2007年度	2008年度	2009年度	2010年度
方式/部品設計	試作/評価	製造	実機検証

【ユニットA】 SIMD化コンパイラの開発要素技術 (1/2)

- SIMD機構の活用:
コンパイラの命令スケジューリング機能を応用
 - 並列化オーバーヘッドのない細粒度の並列実行
 - Basic, Extendの両ユニットで2並列実行
 - 2演算/1命令
 - Basic, Extendを独立使用し, 条件実行時も両ユニットを並行動作
 - SIMD機能の2演算を独立に使用し, 条件付演算もSIMD機構を活用して並列実行

```
DO I=1,N
  IF (条件(I)) then
    A(I)=B(I)+C(I)
  ELSE
    X(I)=Y(I)*Z(I)
  ENDIF
ENDDO
```



```
DO I=1,N,2
  IF (条件(I)) then
    IF (条件(I+1)) then
      A(I)=B(I)+C(I)  A(I+1)=B(I+1)+C(I+1)
      SIMD演算
    ELSE
      A(I)=B(I)+C(I)  X(I+1)=Y(I+1)*Z(I+1)
      BASIC側で計算  EXTEND側で計算
    ENDIF
  ELSE
    IF (条件(I+1)) then
      X(I)=Y(I)*Z(I)  A(I+1)=B(I+1)+C(I+1)
      BASIC側で計算  EXTEND側で計算
    ELSE
      X(I)=Y(I)*Z(I)  X(I+1)=Y(I+1)*Z(I+1)
      SIMD演算
    ENDIF
  ENDIF
ENDDO
```

- プリフェッチによるメモリアクセス高速化
 - L2, L1の両キャッシュ向けに 目的に応じてプリフェッチ命令を配置

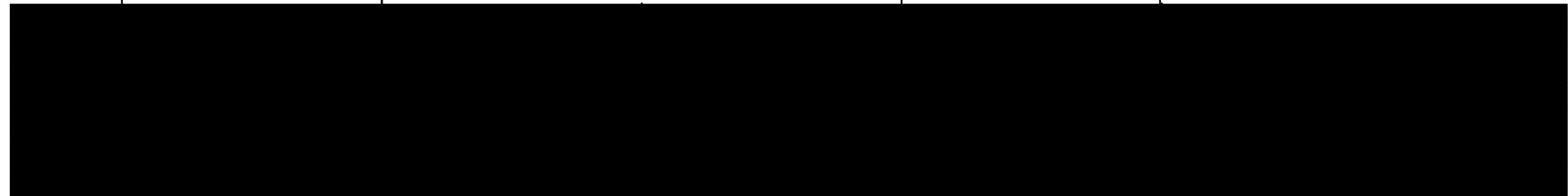
【ユニットA】SIMD化コンパイラの開発要素技術 (2/2)

- 自動並列化機能:
Venusの8コアを1つの高性能コアのように活用
 - 最内ループの並列化
 - コア間共有キャッシュ, コア間高速同期機構の活用
粒度の小さい最内ループも自動並列化
 - ベクトルより広範囲の適用可能性を追求
 - マルチスレッドの特徴
 - 制御オーバーヘッドを縮小し, 回転数の少ないループでも高速化が可能
回転数(ベクトル長)の確保が必須なベクトルに対する優位性

■ SIMD化コンパイラ開発スケジュール



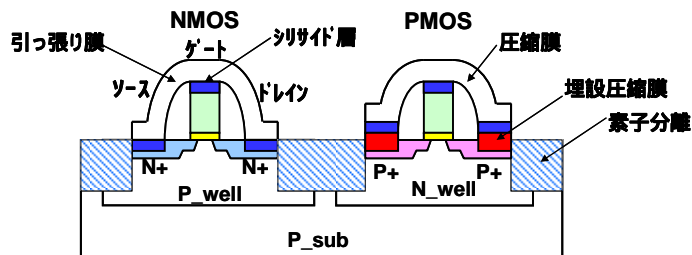
2007年度	2008年度	2009年度	2010年度
詳細設計	プログラミング/機能検証		実機検証 /チューニング



【ユニットB】論理LSI技術(1/2)

(1) 要素技術の内容

- ・最先端半導体プロセス
トランジスタ構造

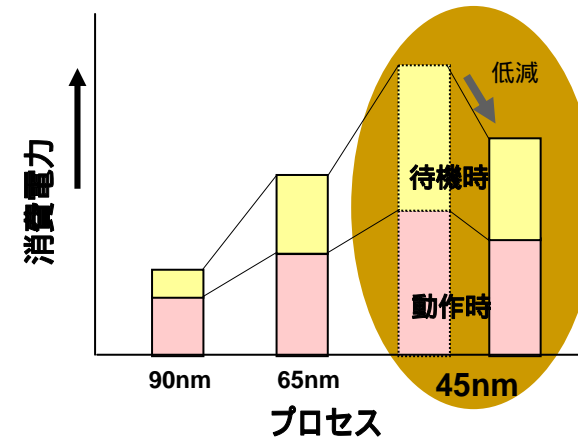


論理LSI概要

プロセス	45nm CMOS, 13層Cu配線
消費電力	最大140W
総トランジスタ数	最大7億
実装方式	フリップチップ, ヘアチップ実装
クロックサイクル	2GHz

- ・LSI回路技術
 - ✓高速・高集積LSIを実現する回路技術
 - ✓動作時/待機時両方の消費電力低減技術
- ・LSI設計技術
 - ✓高い設計品質の確保を目指した検証技術
 - ✓高速・高集積を実現するレイアウト設計技術

消費電力の低減

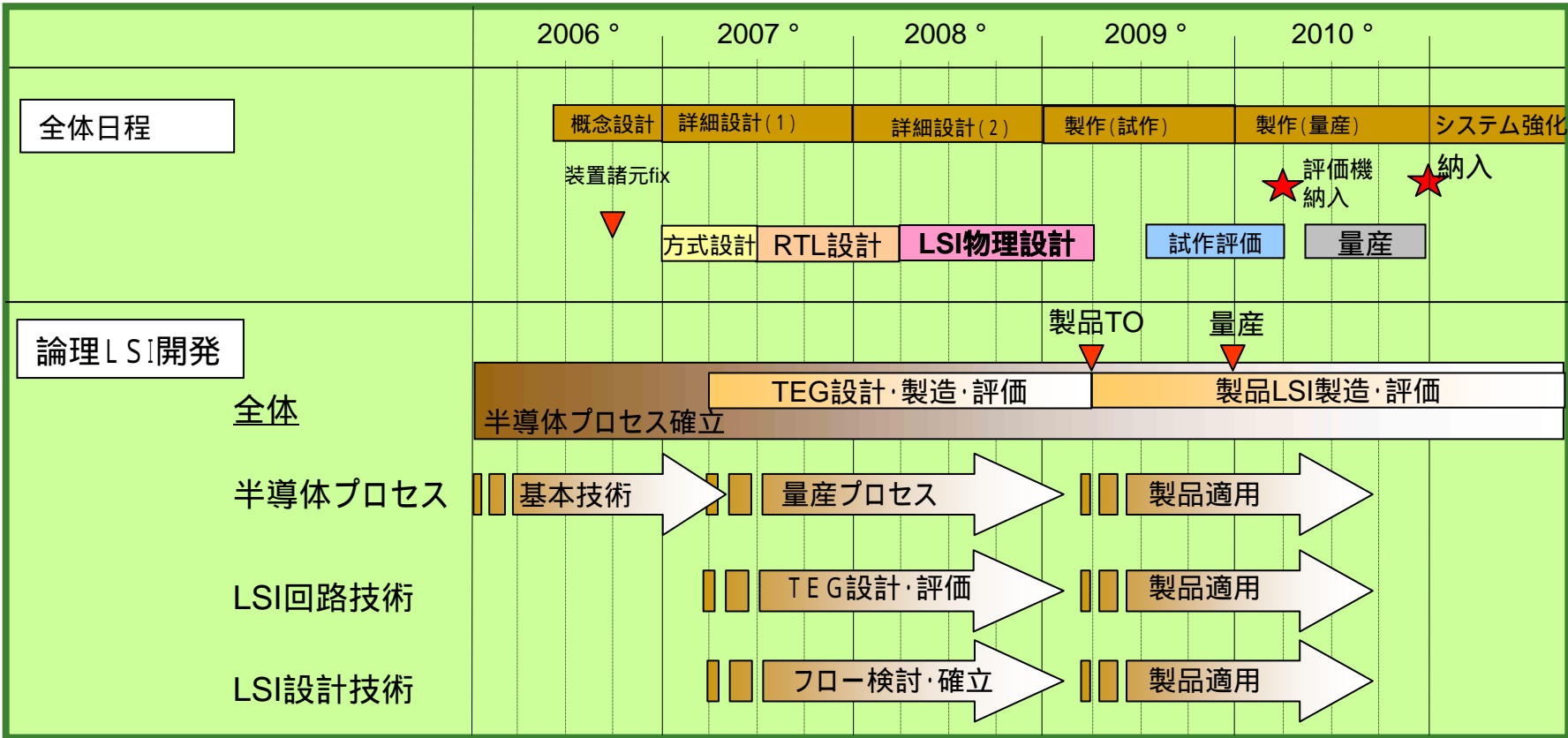


(2) 革新性, 発展性, 技術力強化への寄与

- ・最先端・高性能45nmCMOSプロセスの採用 (歪みシリコン, Low etcの先進技術適用)
- ・高速・高集積回路技術確立 (高速内部セル/SRAM, 高速インタフェース, 高速クロック分配etc)
- ・高速性と両立する低消費電力回路技術の確立 (マルチVth, クロックゲーティングetc)
- ・上記確立した技術を, 他の製品領域 (高性能サーバー, デジタル家電etc) に展開

【ユニットB】論理LSI技術(2/2)

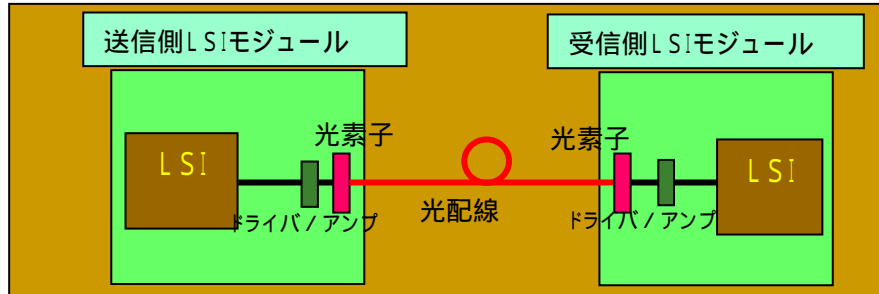
論理LSI開発日程	
TEG設計開始	; 2007年度/1Q末
製品LSI テープアウト	; 2009年度/1Q末
LSI製造量産開始	; 2009年度/4Q末



【ユニットB】光インターコネクト技術(1/2)

(1) 技術開発内容

LSI間の信号伝送を従来の電気伝送に変わって**光で伝送**, 部分水冷技術



具体的な研究項目

- ・高速光素子に関する開発
発光素子, 受光素子
- ・高密度実装に関する開発
高速光電気実装技術, 高密度光配線技術, 冷却モジュール技術, 他

(2) 革新性, 発展性, 技術力強化への寄与

開発目標

- ・高速化: 20Gbps超SerDes技術 **ITRS予測の2倍**
電気伝送の限界5~10Gbpsを打破
- ・小型化: 1000信号/LSI **従来の1/200**
- ・消費電力: **従来比1/100**
- ・高信頼度, 低コスト化



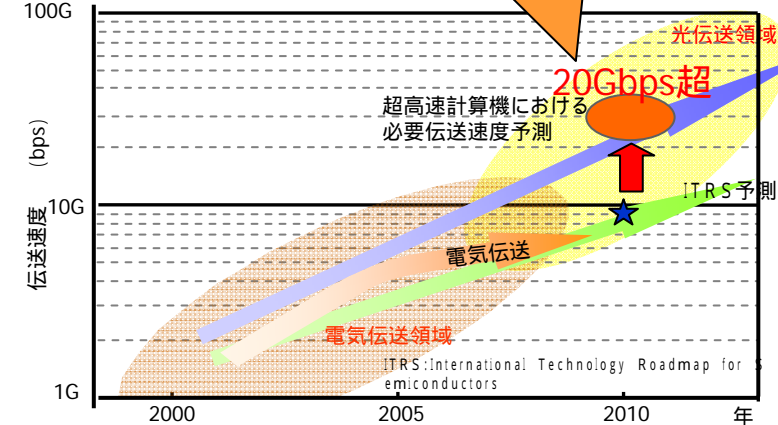
次世代スパコン実現に寄与

さらに

- ・ハイエンドサーバ, ファイルサーバネットワーク機器
 - ・PC, 携帯端末, 情報家電
 - ・医療機器, 車 ……
- への活用が期待される

光技術よりITRS予測の
2倍の高速化を目標

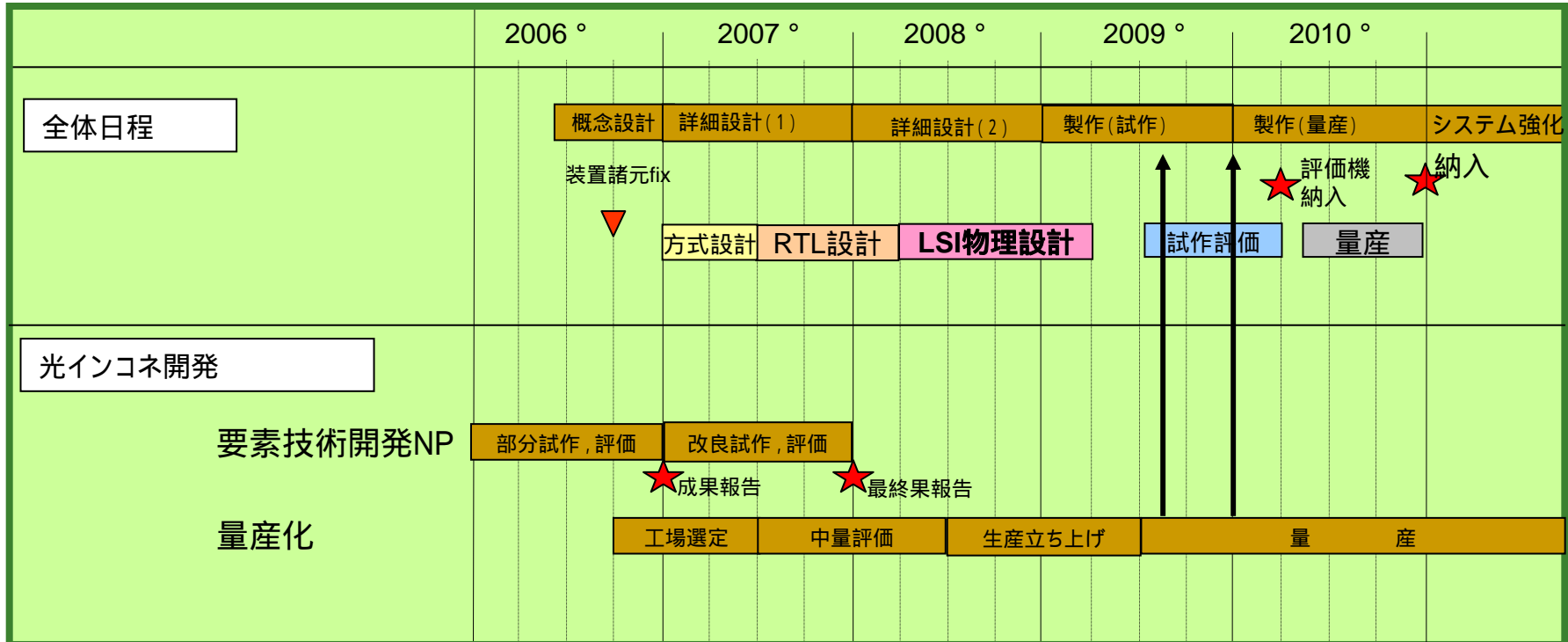
インターコネクト技術のトレンド



【ユニットB】光インターコネクタ(2/2)

光インターコネクタ開発日程

要素技術開発完了	: 2007年度/4Q末
インターフェイス仕様FIX	: 2007年度/4Q末
量産工場選定	: 2007年度/2Q末
中量評価完了	: 2008年度/2Q末
生産立ち上げ	: 2009年度/2Q末
量産開始	: 2009年度/3Q~



【ユニットB】キャッシュ対応ベクトル化コンパイラ(1/2)

(1) 技術開発内容

データをキャッシュに登録するかしないかをソフトウェア(コンパイラ)から制御する技術で、再利用性の高いデータを選んでキャッシュに登録することにより、キャッシュ容量の有効利用を図る。

具体的な技術項目

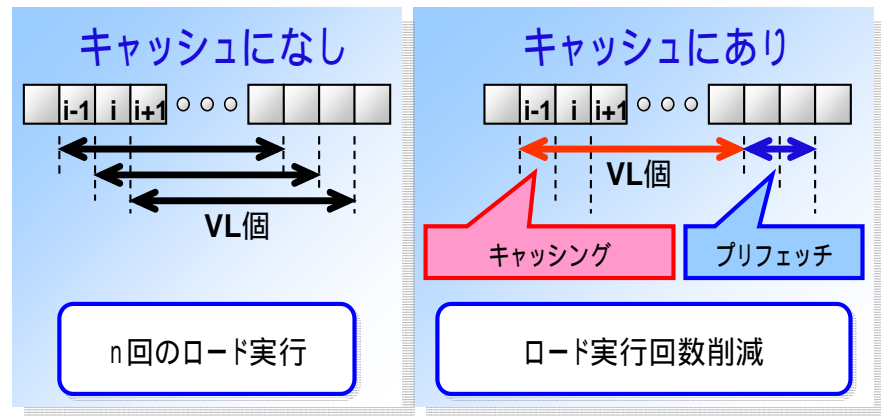
- ・自動的に適用配列を選択する技術
- ・自動的に非適用配列を選択する技術
- ・指示行で適用配列を指定する技術
- ・指示行で非適用配列を指定する技術

```
Program
DO i = 1, n
...
= ...+B(i-1)+...
= ...+B(i)+...
= ...+B(i+1)...
...
END DO
```

(2) 革新性, 発展性

キャッシュ対応技術により、メモリバンド幅維持と短レイテンシ化を実現。

- ・登録データの選択によるキャッシュの有効利用
 - メモリレイテンシ隠蔽
 - キャッシュ格納済み配列のロードバンド幅高
 - バンク競合率低減
 - 短ベクトル長時の実行性能向上
- ・SWプリフェッチ機能によるメモリレイテンシ隠蔽



【ユニットB】キャッシュ対応ベクトル化コンパイラ(2/2)

コンパイラ開発日程

基本設計完了 : 2007年度/4Q末
 機能・詳細設計完了 : 2008年度/4Q末
 製造・単体評価完了 : 2009年度/4Q末
 総合評価完了 : 2010年度/4Q末

