

富士通の最新プロセッサ SPARC64™ IXfxと今後の取り組み

2012年2月16日

富士通株式会社

次世代テクニカルコンピューティング開発本部

システム開発統括部 新庄

SPARC64™ IXfx

PRIMEHPC FX10



■ FX1および「京」からの技術を継承

- マルチコア向け超並列アプリの高い実行性能
- 最大23.2PFlops
- 高いスケーラビリティ

ハイブリッド並列
(VISIMPACT)
高機能SW

Tofuインターコネクト
命令セット拡張
(HPC-ACE)

PRIMEHPC FX10



FX1
SPARC64™ VII
40GFLOPS
4-core CPU
CY2008 ~

「京」(*)



SPARC64™ VIIIfx
128GFLOPS
8-core CPU
CY2010 ~



SPARC64™ IXfx
236.5GFLOPS
16-core CPU
CY2012 ~

*「京」は、理化学研究所が2010年7月に決定した「次世代スーパーコンピュータ」の愛称です。

PRIMEHPC FX10

- 2011年11月7日発表
- 高速・超大規模な計算環境を提供
 - 最大23.2 PFLOPS (98,304ノード, 1,024ラック, 6ペタバイトメモリ)
- ノード
 - プロセッサ: SPARC64™ IXfx
 - 236.5 GFLOPS, 16 cores
 - 高速・大規模メモリ
 - 85GB/s
 - 32 or 64 GB
- Tofuインターコネク
- 水冷オプション
 - 通常の空調用チラーからの冷却水でCDUが動作可能
 - 排気冷却ユニット(EXCU)により空冷部の排気熱を100% 除去可能



■ 高性能・省電力マルチコアCPU

■ 高い消費電力性能比

- マルチコア化の追求により, 高性能化と省電力化を両立
- 水冷方式の採用により, CPU温度を下げ, リーク電流を削減し, 省電力化を実現

■ マルチコアを一つのCPUとして扱う VISIMPACT に対応

■ アプリケーション性能の更なる高速化を追求し, HPC向けに強化

- HPC向け拡張機構 HPC-ACE
- 高いメモリバンド幅(最大 85 GB/s)

■ 高信頼

- 広範囲なエラー検出・自己修復機能, 命令リトライ機能

■ 主な仕様

	仕様
アーキテクチャ	SPARC64™ V9 + HPC-ACE
L1キャッシュ(コアあたり)	32 KB(D) / 32 KB (I), 2 way
L2キャッシュ(コア間共用)	12 MB, 24 way
浮動小数点演算器構成(コアあたり)	FMA x 4 (2 SIMD)
浮動小数点演算同時実行可能数(コアあたり)	8 (= 4 Multiply and Add)
ピーク性能	236.5 GFLOPS / 211 GFLOPS
メモリバンド幅	最大 85 GB/s (BF比 : 0.4) *1

*1) メモリバンド幅およびBF比は, メモリ構成等により異なります.

HPC向け拡張機構：HPC-ACE

(High Performance Computing - Arithmetic Computational Extensions)

■ SPARC64™ IXfx の命令セットアーキテクチャ

■ 準拠仕様

- SPARC-V9 仕様
- JPS (Joint Programmer's Specification): SPARC-V9拡張仕様

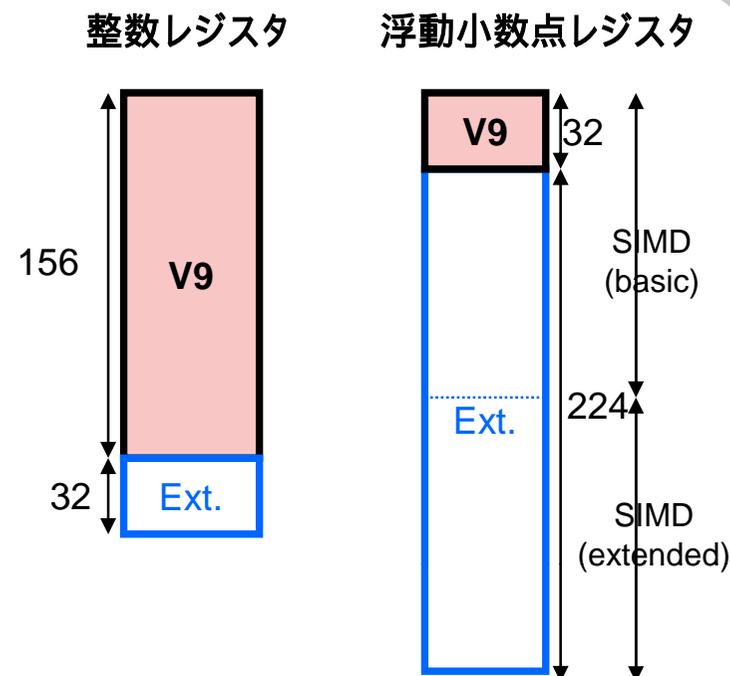
■ HPC-ACE：富士通独自のHPC向けアーキテクチャ拡張

- レジスタ拡張
 - 整数レジスタ, 浮動小数点レジスタ数の拡張による命令レベルの並列度向上
- SIMD (Single Instruction Multiple Data) 命令
 - 1つの命令での複数の演算を並列処理
- セクタキャッシュ (ソフトウェア制御キャッシュ)
 - 再利用性のあるデータを選択的にキャッシュに保持し, キャッシュヒット率向上

HPC-ACE : レジスタ拡張

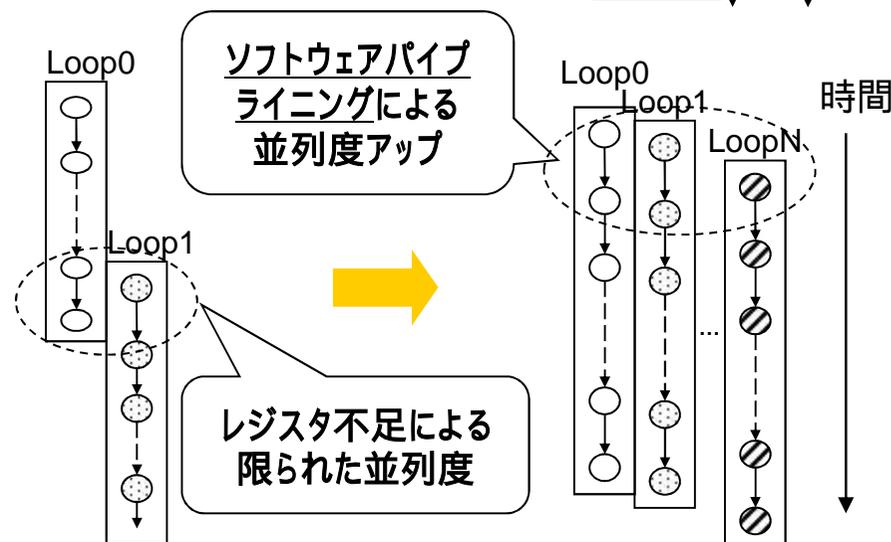
■ SPARC-V9からレジスタを拡張

- 整数レジスタ 156→188
 - 倍精度浮動小数点レジスタ 32→256
 - 下位の32個のレジスタは SPARC-V9 と共通
 - 浮動小数点レジスタはすべて同等
- 拡張浮動小数点レジスタは非SIMD
命令からもアクセス可能



■ 拡張の効果

- レジスタ不足で制限されている命令レベルの並列度の向上
- レジスタ不足によるレジスタ キャッシュ間のレジスタ内容の退避・復元回数の削減



HPC-ACE : SIMD

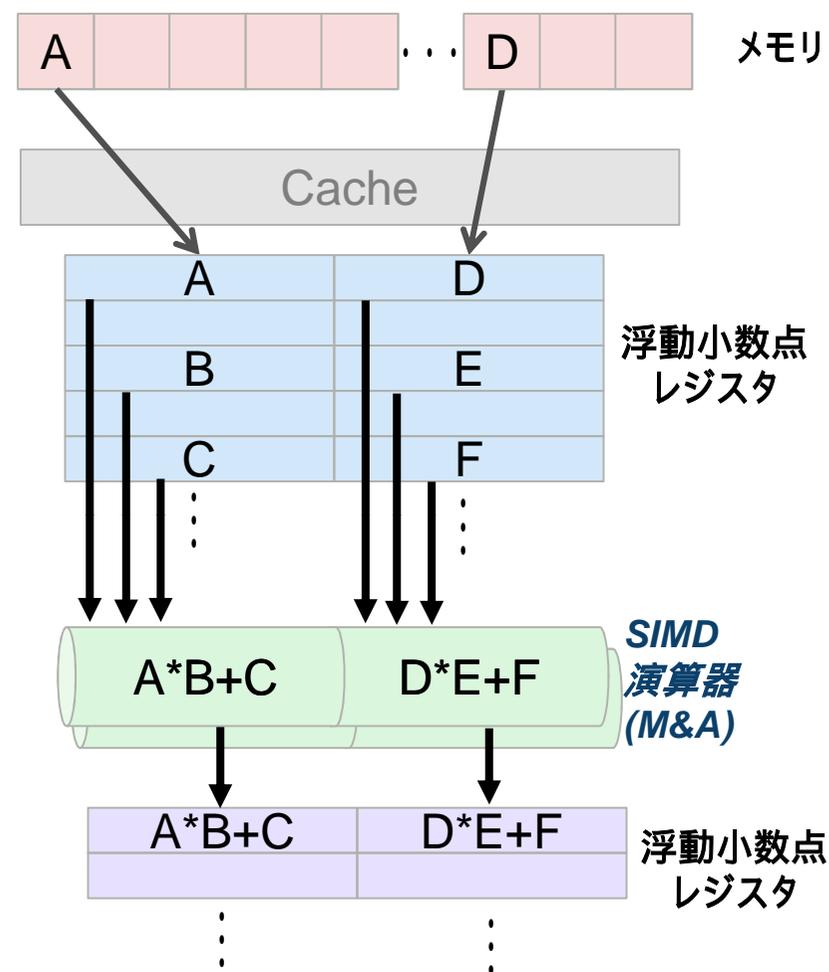
■ SIMD (Single Instruction Multiple Data)

- 1つの命令で複数の演算を並列処理すること

■ SPARC64™ IXfx のSIMDの特長

- 1つのSIMD命令で2つの浮動小数点演算 (単精度 or 倍精度)を並列処理
- 積和演算(M & A)をサポート
- 1つのコアで2つのSIMD命令を同時に実行

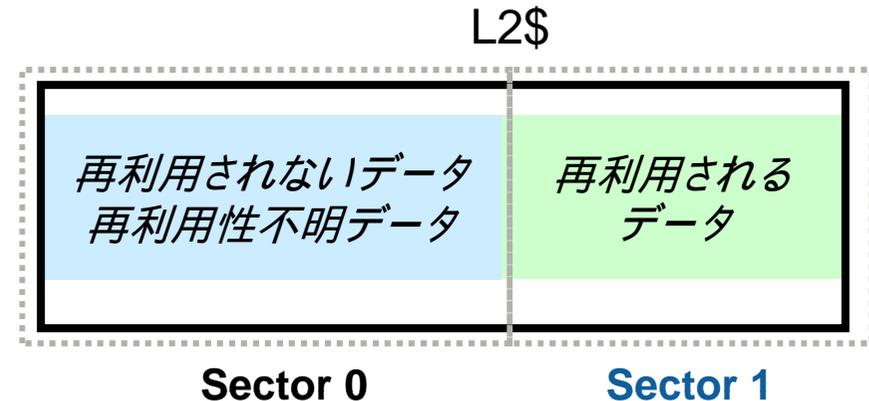
→ 1つのコアで8つの浮動小数点演算を同時に実行可能



HPC-ACE : セクタキャッシュ

■ 再利用性のあるデータを選択的にキャッシュに保持し, キャッシュヒット率の向上を図る

- キャッシュを二つのセクタ(Sector0,1)に分割
- 再利用されるデータはSector 1を利用
- その他のデータはSector 0を利用
- ➔ Sector 0 を使用するデータのアクセスで, キャッシュに残しておきたいSector 1のデータがキャッシュから追い出されることがなくなる



- Sector 1上に保持されるデータは, ユーザがコンパイラ指示行により指定可能

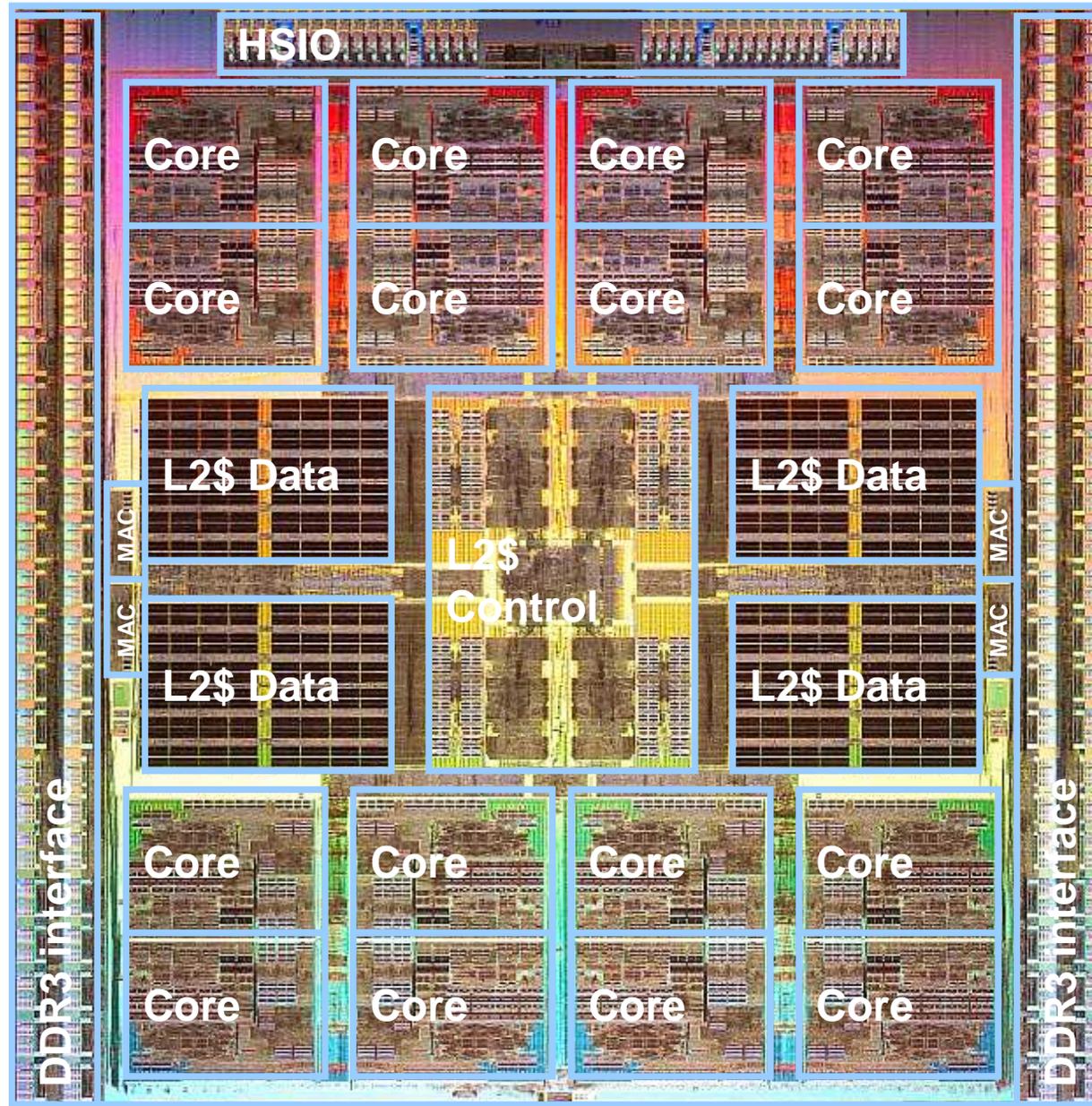
```
!ocl CACHE_SECTOR_SIZE(N1,N2)
!ocl CACHE_SUBSECTOR_ASSIGN(a)
do j=1,m
  do i=1,n
    a(i) = a(i) + b(i,j) * c(i,j)
  enddo
enddo
!ocl END_CACHE_SUBSECTOR
!ocl END_CACHE_SECTOR_SIZE
```

L2キャッシュの N wayを以下のように分割
N1 : Sector 0
N2 : Sector 1

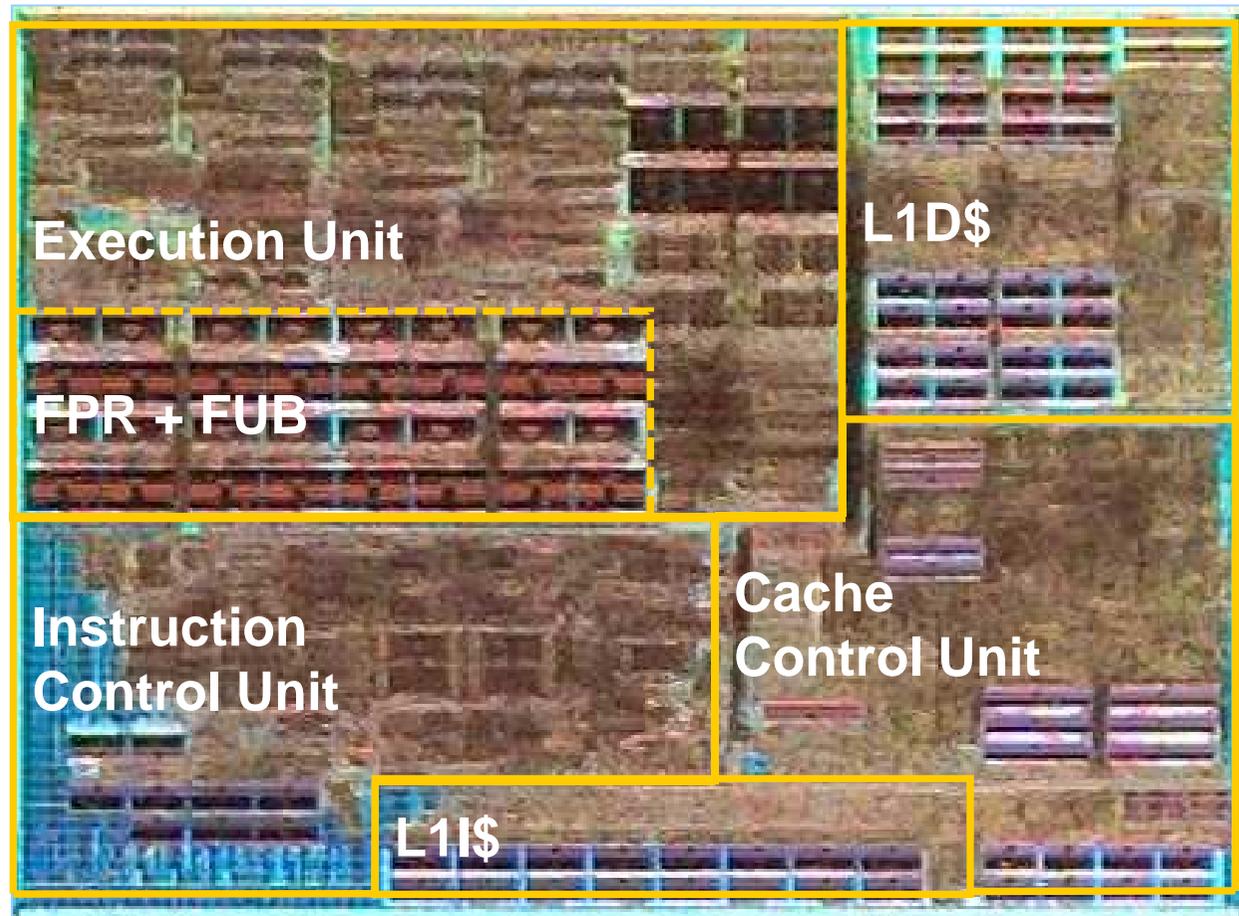
配列 a が b, c の参照によって
キャッシュから追い出されない

- 配列 a を Sector 1 に保持
- その他は Sector 0

SPARC64™ IXfx レイアウト



SPARC64 IXfx Coreレイアウト



SPARC64 IXfx specification

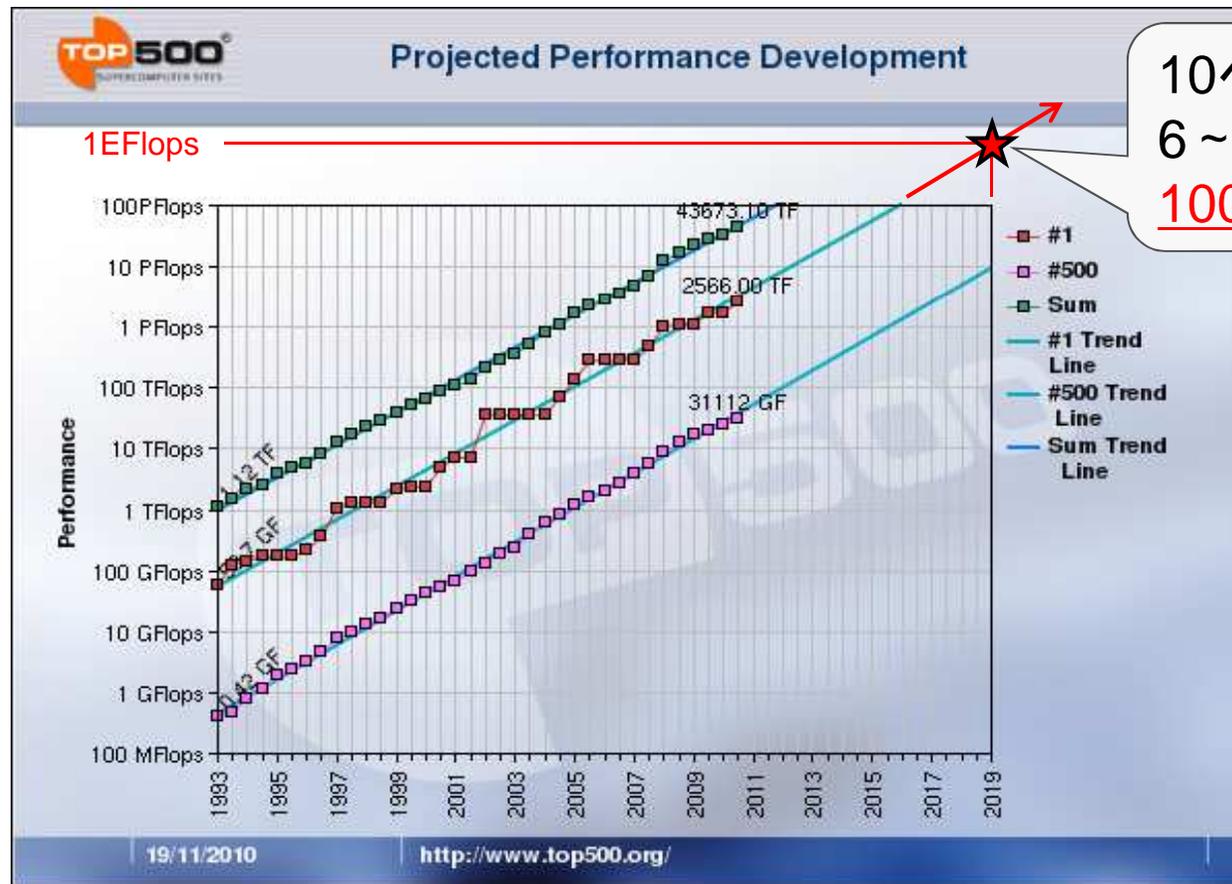


		SPARC64™ IXfx	SPARC64™ VIIIfx
Number of cores		16 cores	8 cores
Clock frequency		1.848 GHz / 1.650 GHz	2 GHz
Cache	L1	I: 32KB/core, D: 32KB/core	I: 32KB/core, D: 32KB/core
	L2	12MB (Shared cache)	6MB (Shared cache)
Peak performance		236.5 GF / 211 GF	128 GF
Memory throughput		85 GB/s	64 GB/s
Power consumption		110W	58W
Process		40 nm	45 nm
Die size		21.9 mm × 22.9 mm	22.7 mm × 22.6 mm
Number of transistors		Approx. 1.87 billion	Approx. 0.76 billion

今後の取り組み

エクサスケールの実現時期

- 2018～2020年頃に、1ExaFlopsの性能が達成される見通し
 - Top500リストより、世界「1位」スパコンの性能を外挿
 - 施設電力のキャパは現在と大差無し **電力当たり性能100倍が必要**

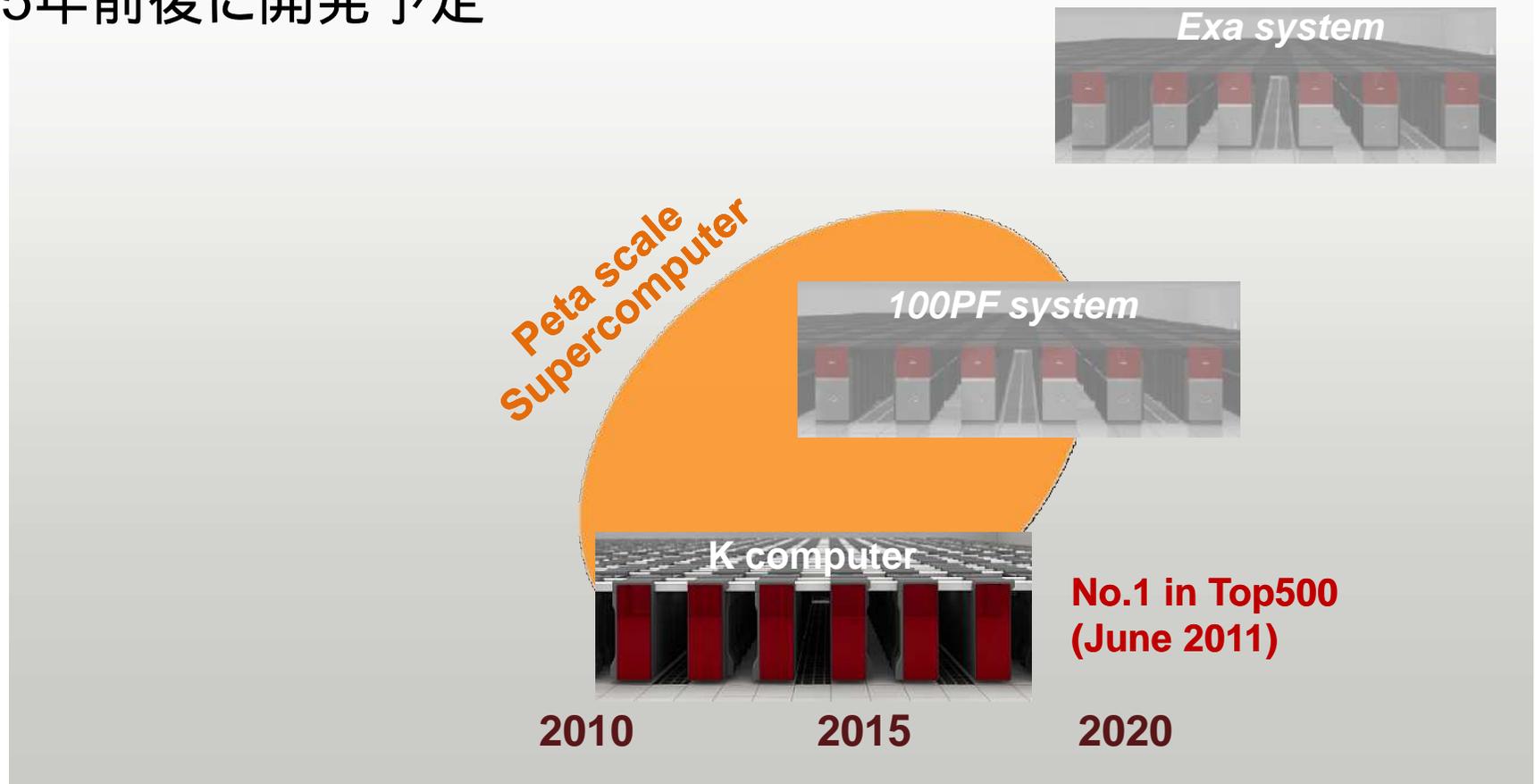


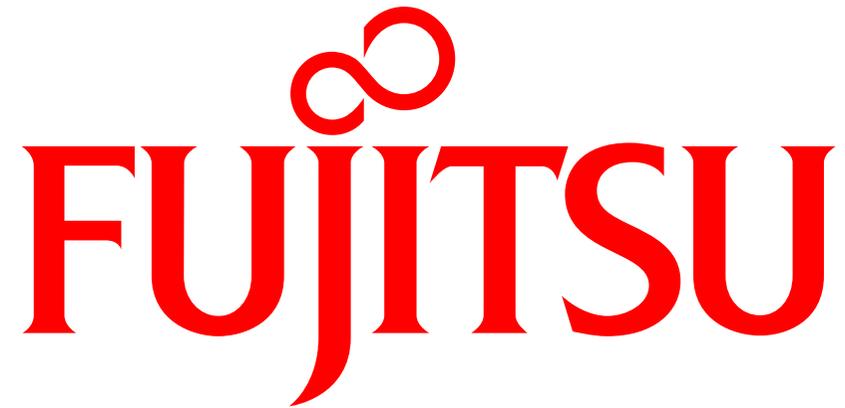
10ペタの
6～8年後に
100倍の性能

http://www.top500.org/lists/2010/11/performance_development

富士通の取り組み

- 富士通は、2020年頃のエクサスケールに必要となる100倍の電力性能を実現するため、今後も高性能・高密度・低消費電力技術の開発に注力して行きます
- 次のステップとして最大100PFLOPS級を目指したテクノロジーを2015年前後に開発予定





shaping tomorrow with you