3. 実例 2

Cell Broadband Engine のアーキテクチャ

増渕 美生
（株）東芝 セミコンダクター社 ブロードバンドシステムLSI開発センター
yoshio.masubuchi@toshiba.co.jp

鈴置 雅一
（株）ソニー・コンピュータエンタテインメント 半導体開発本部
suzu@rd.scei.sony.co.jp

Jim Kahle
IBM Corp.
jakahle@us.ibm.com

Sony グループ，東芝，IBM が共同開発を進めている次世代プロセッサの狙いとアーキテクチャについて述べる．Cell Broadband Engine は，ディジタルホームから分散コンピューティングまでの幅広い分野をターゲットとし，メディア演算・浮動小数点演算を中心とするデータ処理を高速に実行するため，異なるアーキテクチャのコアを複数搭載する非対称マルチコアプロセッサのアプローチをとっている．プロトタイプ実装では，4GHz を超える動作が確認されており，256GFlops を超える性能が実現される．2005 年 8 月には，アーキテクチャ仕様が公開された．

Cell プロジェクト

2001 年 3 月，Sony Computer Entertainment Inc./Sony，東芝，IBM は，次世代のプロセッサ開発を目指して，米国テキサス州オースチンにデザインセンタを開設した．以来，ピーク時で 300 名を超える技術者が参加して 4 年以上に渡り開発を行ってきた．2005 年 8 月には，Cell Broadband Engine と名づけたプロセッサのアーキテクチャを公開した．複数の非対称コアプロセッサを有することを特長とし，高速かつ多様な実装を可能とするフレキシブルなアーキテクチャ定義である．

本稿では Cell Broadband Engine の狙い，主な特長，および第 1 世代の実装例などについて述べる．

Cell Broadband Engine の狙い

プロジェクトの開始時点において，5 年後，10 年後のアプリケーションで必要とされる機能および性能を想定し，既存アーキテクチャの枠組みにとらわれることなく，これに最適なアーキテクチャを開発することを目指した．応用分野としては，ディジタルホームから分散コンピューティングまでの幅広い分野を視野に入れた．

ディジタルホームは，AV 機器を中心にディジタル化，ハイビジョン化が進む発展著しい分野であり，ホームエンタテインメント機器も含めて，主要ターゲットとした．ここでは，大量の AV データをリアルタイムで処理する能力が要求される．加えて，一般の人が使うための自然なヒューマンインタフェースが重要な要素であり，出力系としてのグラフィクス処理，入力系としての音声・画像などの信号処理，さらにはリアルワールドのシミュレーション処理など，さまざまなデータ処理能力が必要となる．

さらに，これらの機器は，ブロードバンドネットワークを通じてインターネットにつながり，有機的に大きなシステムの一部を構成していくことが想定された．ここでは，ホーム内の機器に対してサービスを提供するサーバが存在し，端末機器―サーバ間，サーバ相互間，さらには端末機器相互間などのさまざまな形態での分散コンピューティング機能が必要になってくる．

Cell Broadband Engine は，これらの幅広い応用分野で高い性能を発揮できる汎用性とフレキシビリティを持たせることを目指して，開発が行われた．

masubuti.indd 1244 05.10.24 6:48:47 PM

新しいアーキテクチャの要件

前章で述べたような狙いを実現するために，アーキテクチャ上求められる要件をまとめてみる．

第 1 に，高速データ処理能力が要求される．一般に，プロセッサの性能は，クロック周波数と並列度の積で決まる．今回も，これらを向上させることが，高性能化の鍵となることはいうまでもない．ただし，この際，処理のターゲットを明確にすることが重要である．面積などの物理的な制約の中で，万能なものはあり得ない．Cell Broadband Engine の場合は，前述のような応用ターゲットから，メディア演算・浮動小数点演算を中心に据え，これらの性能を向上させることを主眼に置いた．

この高速演算処理を実現させるためには，大量のデータを高速に通信する能力がなくてはならない．高速なプロセッサに対しては，それに見合ったバンド幅を持ったメモリが必要である．他のチップとの間でブロードバンドデータを送受信するための I/O，さらにはチップ内部のバスについても，高速性が要求される．

これらの性能向上を目指す上で大きな制約条件となるのが消費電力である．一般家庭用機器に使われるプロセッサは，これらの機器の電力プロファイルの範囲内に収めることが必須であり，この条件の中で最大の性能を発揮できるような高い電力効率が求められる．

また，AV データやヒューマンインタフェースの処理を実行する際には，リアルタイム性が重要になる．このため，演算実行時間が予測可能であること，および，複数タスク間で共有されるメモリバンド幅や I/O バンド幅などのリソースの割付け管理が可能であることが必要である．

さらには，幅広い分野でのさまざまな実装が可能なように，汎用性とスケーラビリティの高いアーキテクチャ定義が求められる．

新しいアーキテクチャの基本構成

前述のように，高いデータ演算性能を電力効率よく実現することが最大の要件である．このための基本的なアーキテクチャとして，シンプルなデータ処理プロセッサを複数個搭載するという構成をとった．このプロセッサは，Synergistic Processor Element（SPE）と呼ばれ，既存のアーキテクチャではなく，メディア演算・浮動小数点演算を高速に実行できる SIMD 演算を基本とする新しい RISC 型の命令セットアーキテクチャを持つ．これに加え，OS，I/O 処理，ユーザインタフェース，全体の制御などを受け持つ汎用プロセッサを搭載する．両者を同一アーキテクチャのプロセッサで処理するよりも，それぞれの特長を活かしたかたちで分担をする方が，効率がよいと考えたからである．後者は，既存のシステムソフトウェアやソフトウェア開発環境を利用できるという観点から，新規アーキテクチャではなく，既存の PowerPC アーキテクチャをベースとした．これを Power Processor Element（PPE）と呼ぶ．

このように，全体として非対称型のマルチコアプロセッサ構成をとることが，大きな特長となっている．

Cell Broadband Engine の構成

Cell Broadband Engine のアーキテクチャとしては，PPE・SPE の数，メモリ・I/O・内部バスの構成などは，各種の実装が可能な定義になっている．

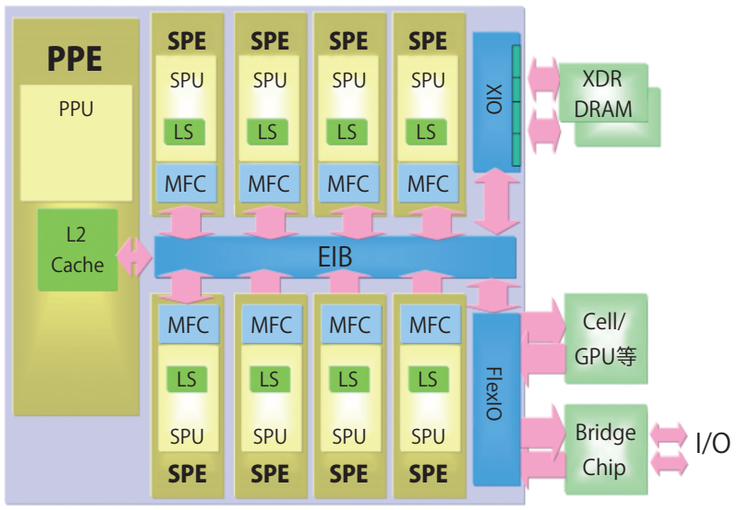
図-1 は，最初の実装である第 1 世代の Cell Broadband Engine の構成を示したものである [1]．この実装では，汎用プロセッサである PPE を 1 個，データ演算プロセッサである SPE を 8 個搭載している．メモリインタフェースとしては，XDR DRAM を 2 チャネル直接接続できるメモリコントローラを持つ．また，2 チャネルの I/O インタフェースを備えており，2 つの外部チップを接続できる．2 つのチャネルは，それぞれに割り当てるデータ幅をシステム構成によって変えることができる．これらの各モジュールは，Element Interconnect Bus（EIB）と呼ぶリングバスで接続され，すべてのデータ転送はこの EIB を介して行われる．

以下，各部の詳細について述べる．

Power Processor Element（PPE）

PPE は，64 ビット PowerPC アーキテクチャに準拠する汎用プロセッサコアであり，主に，OS，I/O 処理，ユーザインタフェース，全体制御などの汎用処理を実行する．**図-2** に，PPE のブロック図を示す．PPE は，In-Order の 2 ウェイスーパースカラ方式を採用しており，2 命令を同時に実行する．さらに 2 ウェイのハードウェアマルチスレッド機構を持ち，基本的には 2 つのスレッドを交互にデコードし発行する．キャッシュの構成は，L1 キャッシュが 32KB の命令キャッシュと 32KB のデータキャッシュ，L2 キャッシュが 512KB である．これらのキャッシュには，エントリをロックする機能があり，データを常駐させてアクセス速度の予測性を向上させることができる．これは，高速応答性，リアルタイム性を

masubuti.indd 1245 05.10.24 6:48:48 PM



PPE：Power Processor Element　SPE：Synergisitic Processor Element　EIB：Element Interconnect Bus

図-1　Cell Broadband Engineの構成

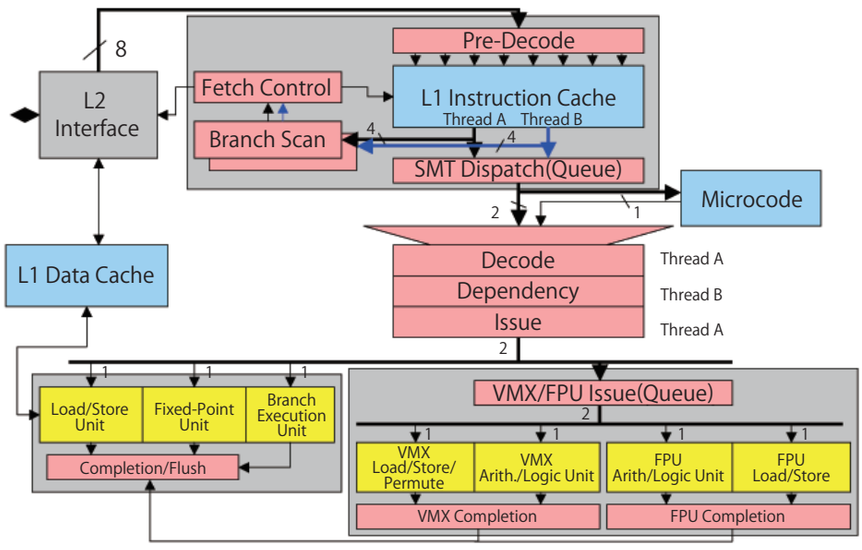


図-2　Power Processor Elementのブロック図

要求する場合に用いられる機能である．

なお，命令セットとしては，通常の PowerPC 命令群に加えて，メディア演算およびベクトル演算に適した Vector/SIMD Multimedia Extension 命令群を併せ持っている．

Synergistic Processor Element（SPE）

SPE は，データ演算処理向けの新しいアーキテクチャに基づくコアプロセッサである[2]．命令体系としては，RISC 型のロードストアアーキテクチャであり，SIMD

masubuti.indd 1246　05.10.24 6:48:54 PM

演算命令が基本となっている．また，今回の実装において特長的な点は，各 SPE ごとに 256KB のローカルストアを持っていることである．SPE の命令セットから直接アクセスできるメモリ空間をローカルストアのみとすることで，シンプルな構造で高速アクセスを実現するとともに，キャッシュと異なってアクセス時間の予測性を高くすることができ，リアルタイム性を要求する用途に適した構成となっている[3]．

命令フォーマットは，32 ビット固定長であり，最大 3 ソースレジスタと 1 デスティネーションレジスタを指定することができる．汎用レジスタとしては，128 本の 128 ビット幅レジスタファイルを有しており，SIMD 演算命令により 128 ビットデータに対して演算を施すことができる．たとえば，単精度浮動小数点演算は，32 ビット× 4 の構成で，4 並列の演算が実現できる．

なお，パイプライン実装依存の命令は持っておらず，将来に渡る互換性の維持にも考慮してある．

前述のように，SPE の命令セットから直接アクセスできるメモリ空間はローカルストアのみであり，主メモリは直接アクセスできない．命令およびデータは，DMA（Direct Memory Access）機能を用いて，主メモリからローカルストアに転送してから処理され，処理結果は同様に DMA により主メモリに戻される．**図-3** に，その様子を示す．各 SPE のローカルメモリは主メモリと同様にシステムメモリ空間上にマッピングすることができ，この空間上で DMA 転送を実行することによって，主メモリ—ローカルストア間のデータ転送が実現される．同様にして，異なる SPE のローカルストア相互間のデータ転送も実行可能である．なお，システムメモリは，PowerPC アーキテクチャをベースとしており，PPE・SPE 双方に，メモリ管理ユニット（MMU）を持っている．

DMA 転送は，プログラムの実行と並行して処理されるため，次のようにして転送時間を隠蔽することが可能である．たとえば，1 つのスレッドの中で，ダブルバッファリングをする．すなわち，ローカルストアのバッファ上のデータを利用して演算処理をしている最中に，別のバッファに対する DMA 転送を実行し，先の演算処理を完了した時点で使用するバッファを切り替えて，新たなデータに対する演算処理を実行するとともに，最初のバッファから結果をメモリに書き戻し，さらに新たなデータをロードする．また，別の方法としては，スレッド間のインタリーブを行うことも可能である．すなわち，演算処理を完了したスレッドをスリープさせ，別のスレッドをアクティブにするとともに，先の演算結果をメモリに書き戻し，新たなデータをロードする DMA 転送を実行する．

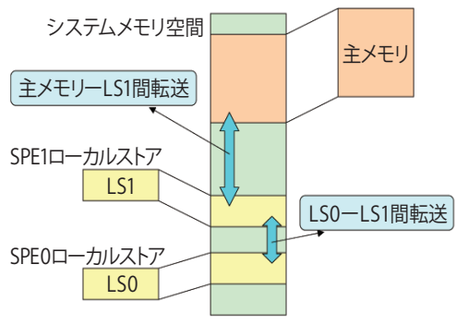


図-3　DMA機能を用いたデータ転送

図-4 に SPE のブロック図を，**図-5** に SPE のパイプラインを，**表-1** に主な命令の演算レイテンシをそれぞれ示す．SPE も，PPE 同様に In-Order の 2 ウェイスーパースカラ方式をとるが，2 命令を同時に発行するためには次のような条件がある．図-4 に示す通り，演算パイプラインは，2 つのグループに分かれており，これらを Even パイプ，Odd パイプと呼ぶ．Even パイプに属する命令が×××0 番地に置かれ，Odd パイプに属する命令が×××4 番地に置かれている場合に，これら 2 命令が同時発行される．こうすることにより，最小限のハードウェア機構によってデータ演算とローカルストアへのアクセスが並列実行でき，効率のよい処理が実現できる．

一般に，プロセッサの動作速度を上げようとすると，パイプラインが深くなり，分岐ペナルティが大きくなる問題がある．これを削減する手法としては，Branch Target Buffer や分岐予測機構などが一般的に知られている．Cell Broadband Engine では，次に示すようなアーキテクチャ上のサポート機能により，少ないハードウェアで実現できる手法をとっている．

1 つは，Hint-for-Branch 命令である．これは，特定の分岐命令の分岐先をあらかじめハードウェアに教えるために命令列中に挿入する命令である．次の例では，B 番地にある分岐命令の分岐先が L 番地であることを Hint-for-Branch 命令で指定する．

```
L:      ....
        Hint-for-Branch B, L;
        .... (compute cond)
B:      if cond goto L
        ....
```

今回の実装では，**図-6** に示す命令バッファを用いて，分岐先命令列の先読みを実行している．6 エントリの命令バッファのうち 2 エントリが，分岐先命令列用になっており，ここに Hint-for-Branch 命令で指定された分岐

masubuti.indd 1247　05.10.24 6:48:56 PM

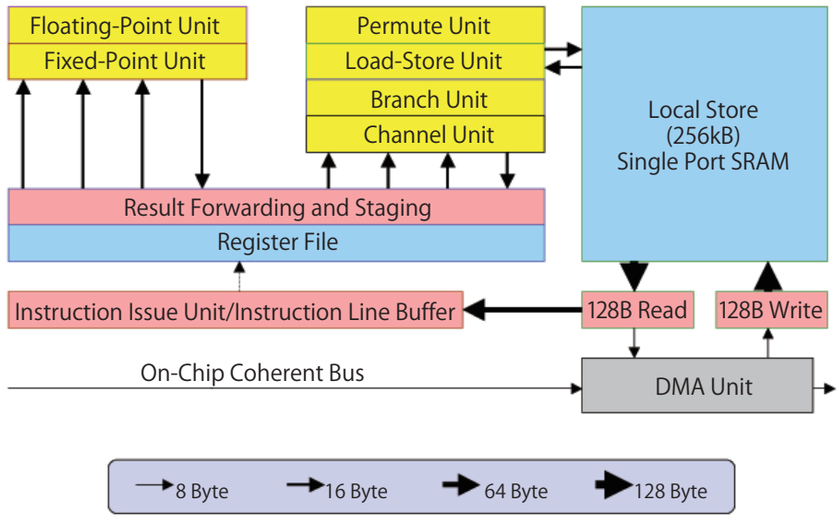


図-4　Synergistic Processor Elementのブロック図

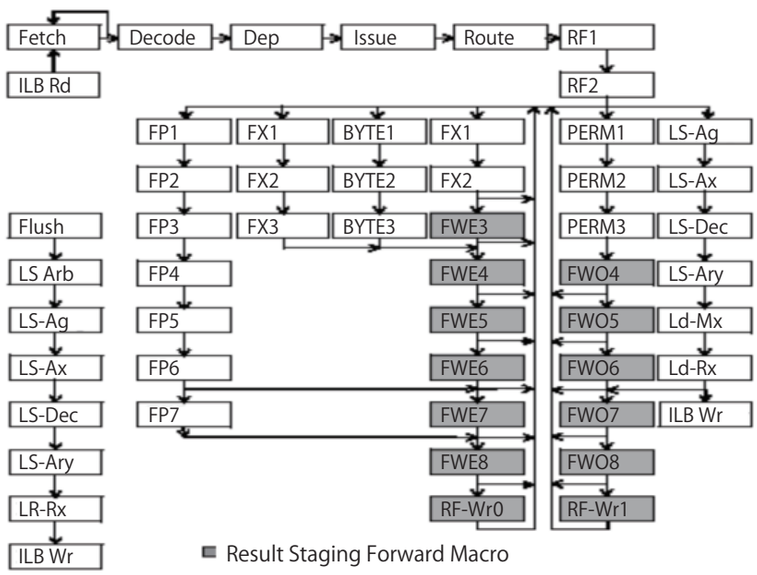


図-5　Synergistic Processor Elementのパイプライン

先命令列が格納される．これは，ソフトウェアで管理される Branch Target Buffer と考えることができる．

もう 1 つは，条件の then 側と else 側の両方を実行して，条件に応じていずれかの結果を残すことにより分岐命令をなくすもので，一般に条件実行命令として知られている手法である．Cell Broadband Engine には，Select 命令があり，条件の値によって，2 つの値のうちのいずれかを選択的に代入することができる．これによって，条件実行が実現できる．

メモリインタフェース

主メモリとして，2 チャネルの XDR DRAM を直結するためのオンチップ・メモリコントローラを持つ[4]．データ幅は，1 チャネル当たり 32 ビット + ECC，ビット当たりのデータレートは最大 3.2Gbps，メモリバンド幅は最大 25.6GB/s である．

masubuti.indd 1248　05.10.24 6:49:01 PM

Simple Fixed (FX)	2
Shift (FX3)	4
Single Precision (FP1-6)	6
Floating Integer (FP7)	7
Byte	4
Permute	4
Local Store (LS)	6
Channel	6
Branch	4

表-1　主な命令の演算レイテンシ

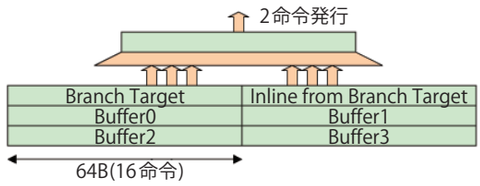


図-6　SPEの命令バッファ

I／Oインタフェース

I/O インタフェースとしては，2 つのチップを接続できるよう，2 チャネルのインタフェースを備えている．信号伝送技術として Rambus 社の FlexIO 技術を用いた単方向インタフェースとなっており，出力側 7 バイト，入力側 5 バイトの合計 12 バイトのデータ幅構成である．2 チャネル合計のバンド幅は最大 76.8GB/s（6.4Gbps 動作時）になる．

2 つのチャネルは，多様なシステム構成を可能とするために，次のような機能を持っている．

1 つは，2 つのチャネルのそれぞれに割り当てるデータ幅をバイト単位で変える機能である．これによって，システムや接続チップによって要求されるバンド幅に応じたインタフェース構成をとることが可能になる．

もう 1 つは，1 つのチャネルをコヒーレントインタフェースとして設定する機能である．この設定で，たとえばもう 1 つの Cell Broadband Engine を接続すると，共有メモリ型マルチプロセッサ（SMP）システムを構成することができる．すなわち，チップ間にまたがってメモリのコヒーレンシが保たれることになる．さらに，ブリッジチップを介して，より大規模の SMP システムを構成することも可能である．なお，通常の I/O インタフェースとしての設定では，コヒーレンシ機能はディセーブルされる．もう一方のチャネルは，I/O インタフェース専用となっており，通常，I/O 機器へのブリッジチップを接続することが想定されている．

内部インターコネクトバス

PPE，8 個の SPE，メモリインタフェースコントローラ，および I/O インタフェースコントローラは，Elemet Interconnect Bus（EIB）と呼ぶリングバスで相互に接続される．EIB は，物理的に時計回り 2 本と反時計回り 2 本，合計 4 本のデータリングバスで構成される．1 本当たりのデータ幅は 16B，動作周波数はコアプロセッサの 1/2 であり，ピーク時には 1 プロセッササイクル当たり 96B のデータバンド幅を持つ．すなわち，データ転送の送信元と受信元の組合せによっては，1 つのリング上で同時に複数のデータ転送が可能である．

その他の主な機能

消費電力の管理とシステム制御のために，温度監視機構とパワーマネージメント機能を持っている．温度監視用には，1 個のリニアセンサと 10 個のディジタルセンサを搭載しており，前者はサーマルダイオードの出力が外部に接続されるので，温度情報をシステム制御に使用することができる．後者は，設定温度で割り込みを発生

masubuti.indd 1249　05.10.24 6:49:04 PM

する機構であり，温度上昇時の早期ワーニングとして利用することができ，たとえばシステムソフトで負荷を調整するなどの対応が可能となる．

また，パワーマネージメント機能として，アーキテクチャ的に Active，Slow，Pause，State Retained Isolated（SRI），State Lost & Isolated（SLI）という 5 つの状態が定義されている．Active はフルスピードでの実行状態，Slow は動作周波数を落とした実行状態，Pause は一時的にコアプロセッサの実行動作を止めた状態，SRI は全体の実行動作を止めた状態（ただしステートは保持），SLI はステートも失って完全に止まった状態である．これらを用い，システムの負荷状況などに応じてシステムソフトにより状態遷移をすることによって，消費電力の調整が可能になる．

リアルタイム処理では，各タスクを一定時間内に完了させることが求められる．この場合，マルチタスク間で共有しているメモリや I/O などのハードウェアリソースに対するアクセス競合に起因する性能劣化が大きな問題となる．これに対応するために，リソース割り当て機能を備えている．すなわち，各タスクが要求するリソースを一括管理し，それぞれに割り当てるバンド幅を調整する機能である．これによって，タスク相互間の干渉によるリアルタイム性の問題に対応することが可能となる．

さらに，ハードウェアの仮想化機能を備えており，複数の OS を並行して実行することが可能である．前述のリソース管理機能は OS 間でも有効であり，これによって，たとえば Linux とリアルタイム OS を並行実行するシステム構成も可能となる．

各 SPE は，Isolated Execution Mode と呼ぶモードで動作させることが可能である．このモードでは，システムソフトを含めて外部からのアクセスが禁止される．また，このモードに遷移する際に，ロードするプログラムを認証する機構を備えており，これらによって，認証されたプログラムによる安全な処理が実現される．

Cell Broadband Engine の実装

第 1 世代の Cell Broadband Engine は，90nm SOI で実装されている．配線層構成は 8 層銅配線＋ローカルインターコネクト，Low-K 技術を用いている．チップサイズは，235mm^2，トランジスタ数は 235 百万個である．プロトタイプチップは，実験室レベルで 4GHz を超えるクロック周波数で動作することが確認されている．

まとめ

Cell Broadband Engine は，汎用処理を効率よくサポートするとともに，メディア演算・浮動小数点演算を中心とするデータ処理を高い性能電力効率で高速に実行するために，非対称マルチコアプロセッサのアーキテクチャをとっている．応用としては，ディジタルホームから分散コンピューティングまでの幅広い分野を想定し，多様な実装を可能とするフレキシブルな定義となっている．第 1 世代のプロトタイプ実装では，4GHz を超える動作が確認されており，このときの性能は 256GFlops を超える．アーキテクチャ定義は公開されており，今後，これらの特長を活かしてさまざまな分野に適用されること，さらには多様な実装がなされることが期待される．

参考文献

1）Pham, D. et al.: The Design and Implementation of a First-Generation CELL Processor, Proceedings of IEEE International Solid-State Circuits Conference, pp.184-185 (2005).
2）Flachs, B. et al.: The Microarchitecture of the Streaming Processor for a CELL Processor, Proceedings of IEEE International Solid-State Circuits Conference, pp.134-135 (2005).
3）Sakai, R. et al.: Programming and Performance Evaluation of the CELL Processor, Hot Chips 17 Conference Proceedings (2005).
4）Clark, S. et al.: Cell Broadband Engine Interconnect and Memory Interface, Hot Chips 17 Conference Proceedings (2005).

（平成 17 年 9 月 27 日受付）

注）：XDR，XIO，FlexIO は，Rambus 社の商標です．

masubuti.indd 1250 05.10.24 6:49:05 PM