

## 9つのCPUと2つのマトリクスプロセッサを搭載した ヘテロジニアス・マルチコア SoC の開発と評価

中島 雅美 石見 幸一 奥村 直人

梶井 規雄 山本 治 近藤 弘郁

株式会社ルネサステクノロジ システムコア技術統括部

〒664-0005 兵庫県伊丹市瑞原 4-1

E-mail: nakajima.masami2@renesas.com

あらまし 高性能・低消費電力が要求される様々なアプリケーション(認識, 推論, 計測, 制御, セキュリティなど)に対応するためのマルチコア SoC を開発した。本 SoC は, 3 種類のシンセサイザブルプロセッサ(8つの CPU:M32R, 2つのマルチバンク・マトリクスプロセッサ:MBMX, 1つコントローラ:M32C)を集積しており, それぞれ, 1GHz, 500MHz, 500MHz で動作する。これらのプロセッサは, 高帯域のマルチレイヤ・システムバスで接続されている。8つの CPU は, キャッシュ・コヒーレンス・メカニズムを用いたパイプラインバスで接続されており, また, バストラフィックを削減するために, 512kB の L2 キャッシュメモリを共有している。マルチバンク・マトリクスプロセッサは, 2リード/1ライトの演算とバックグラウンドでの I/O 動作が可能である。1GHz 動作の CPU は, 様々なアプリケーション, プロセステクノロジに適用可能なディレイモニタで構成されるディレイ・マネジメント・ネットワークを用いることで実現された。9つの CPU と 2つのマトリクスプロセッサを搭載した本コンフィギャラブル・ヘテロジニアス・アーキテクチャにより, 消費電力は 45%削減可能である。

キーワード ヘテロジニアス・マルチコアプロセッサ, コンフィギャラブルプロセッサ, マトリクスプロセッサ, ディレイモニタ, クロックディレイ・アジャスタ

## Design and Evaluation of a Heterogeneous Multicore SoC with 9 CPUs and 2 Matrix Processors

Masami NAKAJIMA Koichi ISHIMI Naoto OKUMURA

Norio MASUI Osamu YAMAMOTO and Hiroyuki KONDO

System Core Technology Div., Renesas Technology Corp.

4-1 Mizuhara, Itami-shi, Hyogo, 664-0005, Japan

E-mail: nakajima.masami2@renesas.com

**Abstract** A multicore SoC has been developed for various applications (recognition, inference, measurement, control and security) that require high-performance processing and low power consumption. This SoC integrates three types of synthesizable processors: eight CPUs (M32R), two multi-bank matrix processors (MBMX), and a controller (M32C). These processors operate at 1GHz, 500MHz and 500MHz, respectively. These three types of processors are interconnected on this chip with a high-bandwidth multi-layer system bus. The eight CPUs are connected to a common pipelined bus using a cache coherence mechanism. Additionally, a 512-KB L2 cache memory is shared by the eight CPUs to reduce internal bus traffic. A multi-bank matrix processor with 2-read/1-write calculation and background I/O operation has been adopted. The 1-GHz CPU is realized using a delay management network which consists of delay monitors that can be applied for any kind of application or process technology. Our configurable heterogeneous architecture with 9 CPUs and 2 matrix processors reduces power consumption by 45%.

**Keyword** Heterogeneous multicore processor, Configurable processor, Matrix processor, Delay monitor, Clock delay adjuster

## 1. はじめに

近年の多機能携帯電話や自動車に搭載される高性能 SoC においては、高度な認識、検索、ネットワーク機能が要求されている。このような多くのアプリケーションを実現するためには、大規模 SoC の開発が必要であり、そのために SoC の開発期間が長くなるという問題があった。そのような問題の解決のためには、ハードウェア・ソフトウェアの共通プラットフォームに基づいたアプリケーション・システム・ソリューションが必須である。上記要求に対応するため、これまでに、多くのマルチコア・アーキテクチャが開発されてきた [1] [2]。

我々は、従来技術よりさらに低消費電力動作が可能な SoC を開発した。本 SoC はヘテロジニアス・マルチコア・アーキテクチャを採用しており、複数の汎用プロセッサと並列処理能力が強化された複数のマトリクスプロセッサを搭載している。信号処理能力の強化のために、マルチバンク・マトリクスプロセッサ(The multi-bank Matrix Processor : MBMX)を採用している。MBMX は性能向上のために、演算と I/O 動作を同時に実行することが可能である。また本 SoC では、発熱やノイズ(シグナル・インテグリティや電圧降下)に対応するため、ディレイ・マネジメント・ネットワークを搭載している。本機能は、シンプルな回路で実現され、また、応答速度も速い。

## 2. アーキテクチャ

### 2.1. チップ概要

近年の SoC 開発においては、高いエネルギー効率が必要とされている。我々は、複雑なプログラムに対応するための 8 つの汎用 CPU(M32R)、大量のマルチメディアデータ処理に対応するための 2 つのマトリクスプロセッサ、チップ全体管理のためのコントローラ(M32C)、の 3 種類のプロセッサを集積した SoC を開発した。チップのシステムブロック図、仕様、写真を、それぞれ、図 1、表 1、図 2 に示す。周波数 1GHz で動作可能なシンセサイザブル 32bit RISC プロセッサは、それぞれ、浮動小数点演算装置、8K バイトの命令キャッシュメモリ、データキャッシュメモリ、ローカルメモリを搭載している。CPU は、共通バスで接続されており、この共通バスを介してキャッシュのコヒーレンシは保たれている [3]。さらに、複数の CPU がデータを共有したり、コヒーレンシを保ちながら演算を行うために、512K バイトの L2 キャッシュメモリを搭載している。8 つの CPU は異なるグループに分割可能である。CPU は同じグループの他の CPU からのミスヌープされる。CPU グループモードにおけるスヌープの動作を図 3 に示す。図 3(a)は、シングルグループモードであり、図

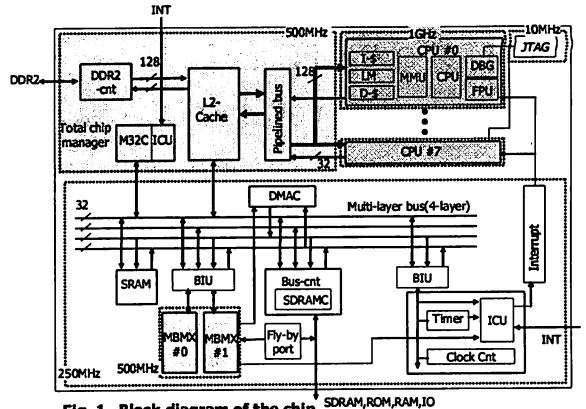


Fig. 1. Block diagram of the chip.

Processor	M32R(1GHz) x 8 I\$:8kB, D\$:8kB, LM:8kB, MMU, FPU M32C(500MHz) x 1 MBMX(38GOPS@500MHz) x 2 2b-PEx768, I-SRAM:32kB, D-SRAM:96kB
Memory	L2-cache:512kB Internal SRAM:32kB
DRAM I/F	DDR2 I/F x 1 SDRAM I/F x 1
Bus	Multi-layer bus(4-layer:4GB/s) Pipelined bus(8GB/s for read, 2GB/s for write) Fly-by bus(1GB/s)
Process	90nm Generic CMOS
Chip size	6.35mm x 6.35mm
Package	729FCBGA
Power	5.0W

Table 1 : Specifications of the chip.

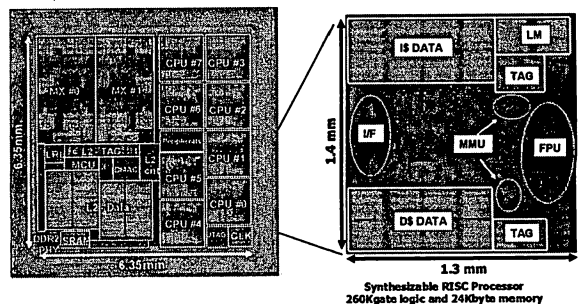


Fig. 2. Micrograph of the chip and the CPU.

3(b)はマルチグループモードである。

### 2.2. コンフィギュラブル・マルチコア・システム

図 4 に、コンフィギュラブル・ヘテロジニアス・マルチコアシステムを示す。本システムにおいては、マルチタスクで構成されるマルチアプリケーションが同時に実行される。



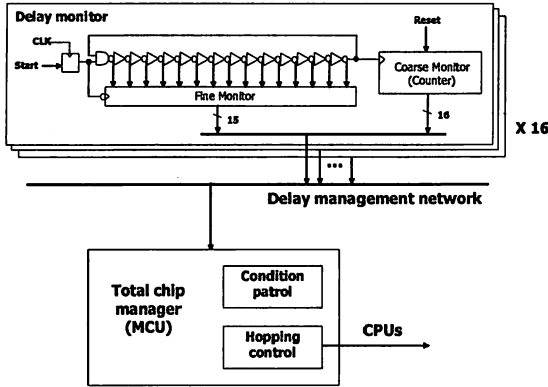


Fig. 7. Block diagram of the delay management network.

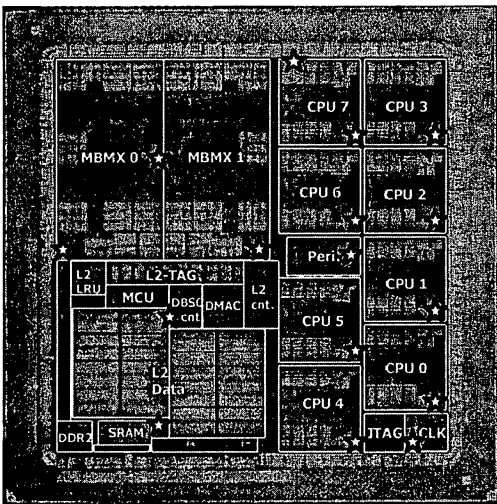


Fig. 8. Location of delay monitors. ☆ Delay monitor

この問題を解決するため、小型のデジタル・ディレイモニタを開発し、本 SoC の各プロセッサに集積した [6]。デジタル・ディレイモニタは、チップの温度ではなく、遅延時間を計測ため、温度、プロセステクノロジー、電源電圧などの調整が不要である。

デジタル・ディレイモニタは、図 7 に示すようにファインモニタとコースモニタから構成されており、1 クロックサイクルでの信号の伝播遅延を計測する。ファインモニタは、15 個のフリップフロップと 15 段のインバータ・チェーンから構成されている。コースモニタは、16bit のカウンタで構成されており、ファインモニタのインバータ・チェーンを伝播した回数を計測する。ディレイモニタは、ファインモニタの 15bit とコースモニタの 16bit の合計 31bit データを出力する。トータル・チップ・マネージャは、ディレイモニタの出力を読み出して、遅延時間に変換する。本 SoC のデジタル・ディレイモニタの計測制度は、20-50psec である。

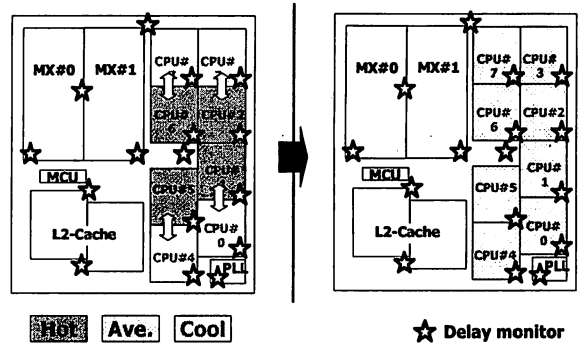


Fig. 9. Task hopping function.

Technique	FV control	Task hopping
Detect time	Several nsec	Several nsec
Control time	Several usec	1usec (Cache copy)
Area	0.1mm <sup>2</sup>	0.025mm <sup>2</sup> (for 1 monitor)
Design methodology	Analog hardmacro (Regulator)	Synthesizable

Table 2 : Comparison of power management techniques.

### 3.2. ディレイ・マネジメント・ネットワークとタスク・ホッピング

SoC に多くのプロセッサを集積すると、処理量に応じていくつかの場所が高温となる。発熱を平均化するために、高温の場所をトータル・チップ・マネージャにフィードバックする必要がある。図 7 に、16 個のディレイモニタとトータル・チップ・マネージャで構成されるディレイ・マネジメント・ネットワークを示す。ディレイモニタを集積した場所を図 8 に示す。

タスク・ホッピングの動作を図 9 に示す。トータル・チップ・マネージャが、ディレイ・マネジメント・ネットワークで計測された遅延時間から、各場所のフラグ(ホット、アベレージ、クール)を生成する。ハードウェアとソフトウェアの連携によるタスク・ホッピングでは、負荷の高い CPU の処理を他の CPU へ分散させ、高温の CPU の熱を下げることにより、チップ全体の性能を向上させる。これにより、性能を低下させることなく電源電圧を下げるが可能となり、消費電力の削減が可能である。

2 つの電源制御の方法の比較を表 2 に示す。1 つは、チップの処理量により周波数と電源電圧を制御する FV 制御である。温度はアナログ温度計によって計測されるため、回路規模は大きくなる。もう 1 つは、タスク・ホッピングである。CPU に割り当てられたタスクを他の CPU に移動させる場合、キャッシュメモリのデ

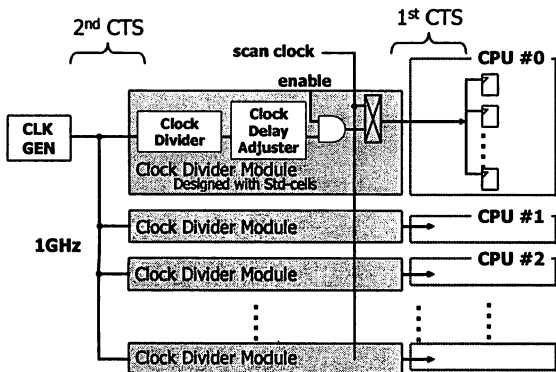


Fig. 10. Clock-tree design.

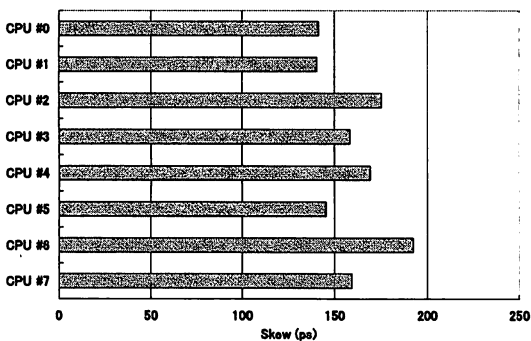


Fig. 11. Local clock skew.

ータのコピーバックが必要となる。8kB のデータのコピーに約 1usec を要する。電源電圧とバックゲートバイアスに対する複雑な制御を必要とする FV 制御と比較して、本技術では、小さい面積と高速なレスポンスを達成している。本技術を用いることで、電源電圧を 2%低下させることが可能となる。

### 3.3. クロックツリーの設計

高速 SoC のクロックツリーの設計においては、クロックスキューを適切に調整することが重要である。本 SoC では、クロックディレイ・アジャスタを含むクロックドライバは、各プロセッサの中央に配置されている。クロックツリーの構造を図 10 に示す。

クロックツリーの合成は、2つのステップから成る。第1ステップでは、クロックドライバから各プロセッサの全てのリーフ FF までのクロックツリーが合成される。第2ステップでは、クロックジェネレータから各クロックドライバまでのクロックツリーが合成される。本手法により、ローカル・クロックスキューの小さいクロックツリーが設計可能である。本 SoC のローカル・クロックスキューを図 11 に示す。各 CPU を 1GHz で動作させるために、ローカル・クロックスキューは 150ps 付近に調整されている。

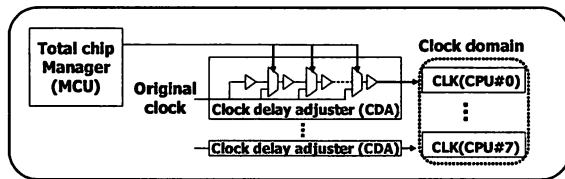
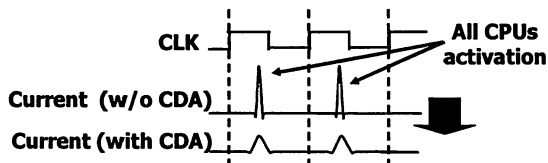


Fig. 12. Clock skew control and reduction of peak current.

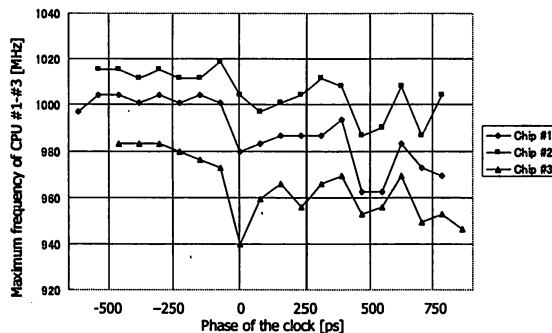


Fig. 13. Evaluation of the clock delay adjuster.

### 3.4. クロックディレイ・アジャスタ

高速 SoC を実現するためにクロックスキューを削減することは、一般的には良いことであるが[7]、小さいクロックスキューは、ピークノイズを発生させる。ダイナミックなピーク電流を削減するために、本 SoC では、各 CPU のクロックパス遅延をクロックディレイ・アジャスタを用いて調整している。クロックディレイ・アジャスタ(CDA)のブロック図を図 12 に示す。CDA は遅延エレメントのセットとセレクタから構成され、その分解能は 78ps である。CDA を用いることにより、各プロセッサのクロック位相は微調整可能となる。本回路の使用により、ダイナミックなピーク電流やノイズは効果的に分散される。

3つのチップで評価した CDA の結果を図 13 に示す。本評価では、7つの CPU の位相は合っており、1つの CPU の CDA の値だけを変化させている。図 13 の縦軸は変化させた CDA の値を示し、横軸は対象 CPU の最高周波数を示している。対象 CPU のピーク電流が他の 7つの CPU と重なった場合には、電圧降下が大きくなり、最高周波数は低下する。つまり、ピーク電流が重ならなければ、電圧降下が小さくなり、最高周波数は向上する。本 SoC では、クロック位相の調整により、約 4%(40MHz)の周波数向上が達成された。

図 13 では、周波数が低下する 2点が存在する。1つは 0ps 周辺であり、もう 1つは 500ps 周辺である。

前者では、全てのクロック位相が重なり、電圧降下が大きくなっている。後者では、SoCに搭載したSRAMがクロックの立下り近辺で多くの電流を消費するために、電圧降下が大きくなっている。

#### 4. まとめ

本チップは、90nmのCMOSプロセスで試作されている。コンフィギャラブル・ヘテロジニアス・マルチコアSoCは、従来の対称型マルチCPUシステムと比較し、45%の消費電力の削減が可能である。また、1GHz動作のCPUと76GOPSのMBMX(500Mhz動作)は、タスク・ホッピングにより実現された。これらの技術は、将来の様々なアプリケーションやプロセステクノロジーにおいても幅広く適用可能である。

#### 5. 謝辞

東北大学青木教授には、アプリケーション側からの技術サポートに深謝致します。

本研究は、新エネルギー・産業技術総合開発機構(NEDO)の支援を受けています。

#### 文 献

- [1] D. Pham et al., "The Design and Implementation of a First-Generation CELL Processor", ISSCC Dig. Tech. Papers, Paper 10.2, pp.184-185, Feb., 2005.
- [2] S. Rusu et al., "A Dual-Core Multi-Threaded Xeon Processor with 16MB L3 Cache", ISSCC Dig. Tech. Papers, Paper 5.3, pp.102-103, Feb., 2006.
- [3] S.Kaneko et al., "A 600MHz Single-Chip Multiprocessor with 4.8GB/s Internal Shared Pipelined Bus and 512kB Internal Memory", ISSCC Dig. Tech. Papers, Paper 14.5, pp.254-255, Feb., 2003.
- [4] M.Nakajima et al., "A 40GOPS 250mW Massively Parallel Processor Based on Matrix Architecture", ISSCC Dig. Tech. Papers, Paper 22.5, pp.410-411, Feb., 2006.
- [5] N.Uchida et al., "3D face recognition using passive stereo vision", Proc. of IEEE Int. Conf. on Image Processing, pp. II950-II953, Sep., 2005.
- [6] M. Nakajima et al., "Design of a Multi-Core SoC with Configurable Heterogeneous 9 CPUs and 2 Matrix Processors," VLSI Symposium on Circuits, Dig. Tech. Papers, Paper 2.2, pp. 14-15, Jun. 2007.
- [7] P. Vuillod, L. Benini, A. Bogliolo, and G. De Micheli, "Clock-skew optimization for peak current reduction," Proceedings of the 1996 International Symposium on Low Power Electronics and Design, pp.265-270, Aug., 1996.