

3次元積層プロセッサ・アーキテクチャ の研究動向

井上弘士[†]

あらまし：マイクロプロセッサの継続的な発展を可能にする新しいアプローチとして3次元積層デバイスの活用が国内外で注目を集めている。複数ダイを製造後に積層することにより、微細化に頼らない集積度の向上が可能となる。また、これまでの2次元実装LSIにおいては、回路の大規模化に伴いブロック間接続のための配線が長くなり、引いては動作周波数の低下や消費電力の増大を招くといった問題があった。これに対し、3次元積層LSIでは、3次元方向へ回路を集積することで短い配線長を維持しつつ、回路を大規模化できるといった利点がある。これに加え、DRAMとロジックのように異なるプロセスを製造した複数のダイを積層し、これらの間を多数の貫通ビア（TSV：Through Silicon Via）で接続することにより、従来の複数チップ構成では実現し得なかった極めて高いメモリバンド幅の活用が可能となる。本講演では、特にマイクロプロセッサとメモリに焦点を当て、3次元積層アーキテクチャの研究動向を紹介する。

Research Trends: 3D Integrated Microprocessor/Memory Architectures

Koji Inoue^{†1}

Abstract: 3D integration is one of the hottest topics in wide area such as VLSI device, computer architecture, design methodology, and so on. By means of stacking multiple dies, we can integrate more functions into a single chip. There are many advantages in this kind of stacking approaches. First, we can reduce the negative effects of long wires in terms of latency and power consumption. Second, it is easy to stack multiple dies which are fabricated in different process technologies, such as DRAM and high-speed logics. Third, we can achieve high on-chip communication bandwidth by exploiting TSVs (Through Silicon Vias). In this talk, I will introduce the trends in research targeting on 3D integrated microprocessor and memory architectures.

なぜ、3次元積層なのか？^{a*}

1970年代初頭に開発されて以来、マイクロプロセッサは現在までに目覚ましい発展を遂げてきた。並列処理や投機実行、オンチップ・キャッシュの大容量化、マルチコア化など、豊富なトランジスタ資源を有効利用することで劇的な性能向上を達成した。また、1990年代中頃からPDAや携帯電話といったモバイル機器が広く普及するようになり、バッテリー寿命の延長に対する要求が高まった。その結果、プロセッサの低消費電力化/低消費エネルギー化に関する研究・開発が加速し、現在では電源電圧の動的最適化など様々な技術が実用化されている。しかしながら、近年、マイクロプロセッサの発展を阻害する以下の問題がより顕在化してきている。

- **メモリウォール問題**：現在、1個のLSIチップに複数個のプロセッサ・コアを搭載した「マルチコア・プロセッサ」が主流となっている。また、順調に成長を続ける半導体微細化技術の進歩を背景に、搭載されるコア数は増加傾向にある。しかしながら、単純にコア数を増加するだけでは、それに見合った性能向上を達成できない場合がある。その主な理由として、「メモリウォール問題の顕在化」が挙げられる。コア数の増加に対してメモリバンド幅はスケールしないため、メモリ性能がボトルネックとなり、マルチコアの高い演算能力を最大限に活用することができなくなる。
- **ムーアの法則の終焉**：これまでのマイクロプロセッサの発展は半導体微細化技術の進歩に支えられてきた。プロセス世代が進むごとに周波数の向上や消費電力の削減、また最近では1個のチップに搭載するコア数の増加など大きなブレークスルーを実現してきた。しかしながら、トランジスタの微細化において物理限界を超えることは不可能であり、ムーアの法則が終焉を向かえた際には更なる高性能化/低消費電力化の実現が困難になる。

マイクロプロセッサはコンピュータ・システムの頭脳とも言うべき重要な構成要素であり、マイクロプロセッサ発展の限界はコンピュータ・システムそのものの限界を意味する。今後の更なる高度情報化社会を見据えた場合、膨大な量のデータを対象としたインテリジェントな情報処理が必要となる。このような要求を満足するためには、マイクロプロセッサの継続的な発展は必要不可欠である。

前述した諸問題を解決しマイクロプロセッサの継続的な発展を可能にする新しいアプローチとして3次元積層デバイスの活用が国内外で注目を集めている。複数ダイ

^{†1}九州大学
Kyushu University

を製造後に積層することにより、微細化に頼らない集積度の向上が可能となる。また、これまでの2次元実装LSIにおいては、回路の大規模化に伴いブロック間接続のための配線が長くなり、引いては動作周波数の低下や消費電力の増大を招くといった問題があった。これに対し、3次元実装LSIでは、3次元方向へ回路を集積することで短い配線長を維持しつつ、回路を大規模化できるといった利点がある。これに加え、DRAMとロジックのように異なる製造プロセスを経て作成した複数のダイを積層しこれらの間を多数の貫通ビア(TSV: Through Silicon Via)で接続することにより、従来の複数チップ構成では実現し得なかった極めて高いメモリバンド幅が活用可能となる。その他にも、配線複雑度の緩和によるマスクコスト削減、古いプロセステクノロジーを用いたダイのスタックによるコスト削減、といった利点も議論されている。

研究開発動向

大きな期待を集める3次元積層マルチコア・プロセッサであるが、その実用化に向けて解決すべき課題も多く存在する。現在、学术界や産業界において様々な研究開発が進められている。以下、プロセッサと大容量メモリの積層を前提とした研究開発動向を示す。

- **3次元積層プロセッサのアーキテクチャ技術**: 3次元積層技術の活用はメモリ・ウォール問題の直接的な解決策として大きな注目を集めている。マルチコア・プロセッサと大容量メモリを同一パッケージに集積することで、I/Oピン駆動といった高負荷容量の充放電を回避することができる。これにより、メモリアクセス・レイテンシが大幅に削減される。また、半導体パッケージ構造に起因する使用可能I/Oピン数の制約を受けることなく、プロセッサとメモリを接続するデータバスを容易に拡張できる。その結果、従来の個別チップ構成では実現が難しかった極めて高いメモリ・バンド幅が実装可能となる。1990年代後半に、プロセッサと大容量メモリ(DRAMなど)を1つのLSIに混載するアプローチが大きな注目を集めた[Patterson97][Murakami97][Inoue99]。しかしながら、同一ダイにロジックとDRAMを混載するアプローチでは製造コストが高くなる傾向にあり、その実用化には至らなかった。これに対し、TSVを用いた3次元積層技術ではこれらを個別プロセスで製造し、その後に積層することでプロセッサと大容量メモリの混載を可能にするため、比較的低コストで実現できる。現在はプロセッサと大容量メモリを3次元積層した商用チップは存在しないが、プロセッサ主要メーカーは3次元積層技術の導入を検討しており、今後数年以内の実用化が期待される。具体的な研究開発例としては、米国インテル社が発表した80コア搭載プロセッサ

と大容量メモリの積層[Azimi07]、慶応大学/ルネサステクノロジ/日立製作所による無線TSVを用いた組込みマルチコア・プロセッサ[Niitsu09]、NECによるネットワーク・オン・チップ(NOC: Network On Chip)を搭載した組込み向けメモリ[Saito09]などが挙げられる。また、3次元積層技術を用いることによるアーキテクチャへの最も大きな影響として、従来の階層メモリ構造(もしくはメモリ・アーキテクチャ)とは異なる新しいアプローチが実現可能になる点を挙げることができる。これに注目し、3次元積層を前提としたメモリ・アーキテクチャ技術に関する様々な研究が進められている。例えば、DRAMを用いた大容量キャッシュ・メモリの活用法[Black06][橋口09][Loh09]やメモリアクセス管理の最適化手法[Loh08]などが挙げられる。また、最近ではPRAM(Phase-change RAM)やMRAM(Magnetoresistive RAM)といった不揮発性メモリを積層したプロセッサ・システムに関する検討も始まっている[Zhou09][Sun09]。一方、メモリシステムのみならず、3次元積層技術を活用したアクセラレーション技術に関する研究[Woo08]や、ネットワーク・オン・チップに関する研究[Park08][松谷10]も今後の重要な要素技術として注目されている。

- **3次元積層プロセッサの性能/消費電力/温度解析**: 3次元積層マイクロプロセッサに関しては、特に「3次元化によりどの程度の高性能化を実現できるのか?」「低消費電力化は達成可能か?」「チップ温度上昇はどの程度問題になるのか?」が大きな関心事となる。このような疑問に対し、従来の2次元実装プロセッサとの性能/消費電力/温度比較の結果が報告されている[Liu95][Xie06][Loi06]。
- **3次元積層プロセッサの設計技術**: 3次元積層LSIにおいてはその設計空間がより広がるため、設計技術の確立は急務の課題である。また、2次元実装と比較して温度制約がより厳しくなるといった欠点も考慮する必要がある。このような背景の中、3次元積層LSI向け設計技術に関する研究開発が活発に行われている。例えば、3次元積層を考慮したキャッシュ・メモリの性能/消費電力見積りツール[Tsai05]や温度解析ツール、コスト見積り技術[Dong09]などがある。また、3次元積層LSIにおいてはLSIテストがより複雑となるため、この問題を解決すべく幾つかの研究開発が実施されるようになった[Lee09]。

上記以外にも多くの研究開発が行われており、3次元積層技術の活用はプロセッサ・システムに新しいパラダイムシフトをもたらす可能性を秘めている。

謝辞

本研究は一部、独立行政法人新エネルギー・産業技術総合開発機構（NEDO）若手グラント、ならびに、科学研究費補助金（課題番号 21680005）による。

参考文献

- [Patterson97] David Patterson, Thomas Anderson, Neal Cardwell, Richard Fromm, Kimberley Keeton, Christoforos Kozyrakis, Randi Thomas, and Kathy Yelick, “Intelligent RAM (IRAM): Chips that remember and compute,” Proc. of the 1997 International Solid-State Circuits Conference, p. 224–225, Feb. 1997.
- [Murakami97] K. Murakami, S. Shirakawa, and H. Miyajima, “Parallel processing ram chip with 256mb dram and quad processors,” Proc. of the 1997 International Solid-State Circuits Conference, pp. 228–229, Feb. 1997.
- [Inoue99] Koji Inoue, Koji Kai, and Kazuaki Murakami, “Dynamically Variable Line-Size Cache Exploiting High On-Chip Memory Bandwidth of Merged DRAM/Logic LSIs,” Fifth International Symposium on High-Performance Computer Architecture, pp.218-222, Jan. 1999.
- [Azimi07] M. Azimi, N. Cherukuri, D. N. Jayasmha, A. Kumar, P. Kundu, S. Park, I. Schoinas, and A. S. Vaidya, “Integration Challenges and Tradeoffs for Tera-scale Architectures,” Intel Technology Journal, Vol. 11, Issue 03, Aug. 2007.
- [Niitsu09] K. Niitsu, Y. Shimazaki, Y. Sugimori, Y. Kohama, K. Kasuga, I. Nonomura, M. Saen, S. Komatsu, K. Osada, N. Irie, T. Hattori, A. Hasegawa, and T. Kuroda, “An Inductive-Coupling Link for 3D Integration of a 90nm CMOS Processor and a 65nm CMOS SRAM,” Proc. of the 2009 International Solid-State Circuits Conference, pp. 480–482, Feb. 2009.
- [Saito09] H. Saito, M. Nakajima, T. Okamoto, Y. Yamada, A. Ohuchi, N. Iguchi, T. Sakamoto, K. Yamaguchi, and M. Mizuno, “A Chip-Stacked Memory for On-Chip SRAM-Rich SoCs and Processors,” Proc. of the 2009 International Solid-State Circuits Conference, pp. 60-62, Feb. 2009.
- [Liu95] C. C. Liu, I. Ganusov, M. Burtscher, and S. Tiwari, “Bridging the Processor-Memory Performance Gap with 3D IC Technology,” IEEE Design & Test of Computers, pp.556-564, Nov.-Dec. 2005.
- [Xie06] Y. Xie, G. H. Loh, B. Black, and K. Bernstein, “Design Space Exploration for 3D Architectures,” ACM Journal on Emerging Technologies in Computing Systems, Vol.2, No.2, pp.65-1-3, Apr. 2006.
- [Loi06] G. L. Loi, B. Agrawal, N. Sriyastaya, S. Lin, T. Sherwood, and K. Banerjee, “A Thermally-Aware Performance Analysis of Vertically Integrated (3-D) Processor-Memory hierarchy,” Design Automation Conference, pp.991-996, 2006.
- [Black06] B. Black, M. Annavaram, N. Brekelbaum, J. DeVale, L. Jiang, G. H. Loh, D. McCaule, P. Morrow, D. W. Nelson, D. Pantuso, P. Reed, J. Rupley, S. Shankar, J. Shen, and C. Webb, “Die Stacking (3D) Microarchitecture,” The 39th International Symposium on Microarchitecture, pp. 469-479, 2006.
- [Loh09] G. H. Loh, “Extending the effectiveness of 3D-stacked DRAM caches with an adaptive multi-queue policy,” The 42th International Symposium on Microarchitecture, pp.201-212, 2009.
- [Loh08] G. H. Loh, “3D-Stacked Memory Architectures for Multi-core Processors,” The 35th International Symposium on Computer Architecture, pp. 453-464, June 2008.
- [Zhou09] P. Zhou, B. Zhao, J. Yang, and Y. Zhang, “A durable and energy efficient main memory using phase change memory technology,” The 36th International Symposium on Computer Architecture, pp. 14-23, June 2008.
- [Sun09] G. Sun, X. Dong, Y. Xie, J. Li, and Y. Chen, “A Novel Architecture of the 3D Stacked MRAM L2 Cache for CMP,” The 15th International Symposium on High-Performance Computer Architecture, Feb. 2009.
- [Park08] D. Park, S. Eachempati, R. Das, A. Mishra, Yuan Xie, V. Narayanan, and C. Das, “MIRA: A Multi-layer On Chip Interconnect Router Architecture,” International Symposium on Computer Architecture, pp. 251-261, June 2008.
- [Woo08] D. H. Woo, J. B. Fryman, A. D. Knies, M. Eng, and H-H. S. Lee, “POD: A

3D-Integrated Broad-Purpose Acceleration Layer, " In IEEE MICRO special issue on Accelerator Architectures, Vol. 28, No. 4, pp.28-40, July/August, 2008.

[Tsai05] Y-F. Tsai, Y. Xie, N. Vijaykrishnan, and M. J. Irwin, "Three-Dimensional Cache Design Exploration Using 3DCacti,"IEEE International Conference on Computer Design, pp. 519-524, Oct. 2005.

[Dong09] X. Dong and Y. Xie, "System-level Cost Analysis and Design Exploration for 3D ICs," Asia and South Pacific Design Automation Conference, 2009.

[Lee09] H-H. S. Lee and K. Chakrabarty, "Test Challenges for 3D Integrated Circuits," In IEEE Design & Test of Computers, Special Issue on 3D IC Design and Test, Vol.26, Issue 5, pg. 26-35, Sept/Oct, 2009.

[橋口 09] 橋口慎哉, 小野貴継, 井上弘士, 村上和彰, "3次元DRAM - プロセッサ積層実装を対象としたオンチップ・メモリ・アーキテクチャの提案と評価,"情報処理学会研究報告, Vol. 2009-ARC-183, No.16, 2009年4月.

[松谷 10] 松谷 宏紀, 鯉淵 道紘, 黒田 忠広, 天野 英晴, "ワイヤレス 3-D NoC のための通信プロトコルの検討", 情報処理学会研究報告 2010-ARC-190 (SWoPP'10), No.17, Aug 2010.