ASME Advances in Electronic Packaging 97. vol 1. June, 1997. pp. 863-868.

To appear in the proceedings of the ASME Interpack 97, June 1997

IMPACT OF MCM'S ON HIGH PERFORMANCE PROCESSORS

Brian T. Davis, Claude Gauthier, Phiroze Parakh, Todd Basso, Charles Lefurgy,

Richard Brown and Trevor Mudge

Electrical Engineering and Computer Science Dept.
University of Michigan
1301 Beal Avenue
Ann Arbor, MI 48109-2122, USA
http://www.eecs.umich.edu:80/UMichMP/

ABSTRACT

Researchers at the University of Michigan, in collaboration with their partners from Motorola and Cascade Design Automation, are developing design methodologies and automated tools for use in implementing high clock rate digital "systems-on-an-MCM." The PUMA processor, a demonstration vehicle that executes a subset of the PowerPC instruction set, will be designed to operate with a 1 GHz clock. The PUMA will be implemented as a system-on-an-MCM from a set of complementary GaAs (CGaAs) chips that employ area interconnect for high bandwidth inter-chip connections on the MCM. In particular, the paper discusses design tools for systems-on-an-MCM, the optimization of inter-chip drivers, and techniques for optimizing the system performance given the delay of inter-chip crossing.

1. INTRODUCTION

There are a variety of reasons that MCMs are likely to be important for the next generation of electronic devices including small size, light weight, low power, device density and greater connectivity. Because MCMs can provide small size, with light weight while requiring less power, MCMs are attractive for use in portable computing and communications applications where miniaturization is desirable. In fact, this market, owing to its size, is likely to drive the research and demand for MCMs. However, MCMs will also have an important impact in high performance processors because of the high bandwidth, high-frequency connectivity they offer. This paper describes a project to design and build a 1 GHz processor that relies on this property.

To reach our goal of 1 GHz, the core chips of the PUMA system will be designed and fabricated in Motorola's complementary GaAs (CGaAs) technology. CGaAs has many of the properties of conventional silicon CMOS but it can run faster at lower power (it operates comfortably with a Vdd of 1-1.5V). Although CGaAs' higher speed should allow us to reach the clock rate we are aiming for, its low integration levels, typical of an experimental technology, are a potential obstacle to overall system performance.

Low integration levels require that the major functional blocks of the processor be partitioned into several separate chips. The challenge is to keep the inter-chip signals from running significantly slower than the chips themselves and, as a result, degrading the system performance. This can be attacked at two levels: 1) the architecture should be designed to minimize the number of inter-chip crossings that are on critical timing paths; and 2) use area interconnect and advanced MCM technology to limit the delay experienced on inter-chip crossings. For the PUMA system this means that what is typically referred to as the processor core is contained in a chipset of four die plus the necessary cache memory devices (see Figure 1).

The successful design and verification of a processor requires a range of sophisticated design tools. Many of these exist for silicon CMOS and are widely deployed in the computer

- Two level Cache: Split I and D through L2
- L1: direct mapped virtual index/virtual tags, Line size 16 Byte
- L1: 1 KB on-chip I-side / 16 KB off-chip custom CGaAs pipelined SRAM D-side
- L2 direct mapped virtual index/ virtual and physical tags, Line size 16 Byte
- L2: 1 MB of high-speed commercial pipelined SSRAM on each side

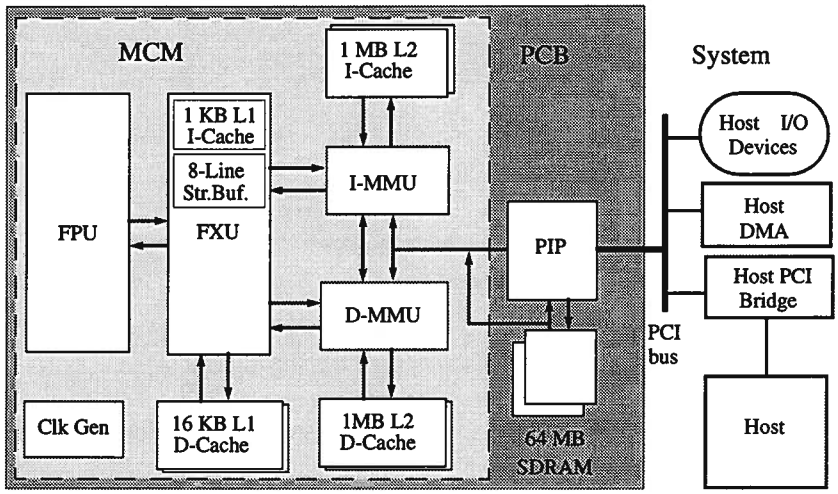

FIGURE 1. PUMA System Architecture

industry. Due to its similarity with CMOS, these tools can be readily adapted to the design of CGaAs chips. Unfortunately, tools for the design of MCM-based systems with area interconnect are much less developed. Indeed, an important enabling component to the widespread use of MCMs are tools that support the design of systems-on-an-MCM, where multiple die integrated onto an MCM to form a single system. An important part of our work has been to develop, in collaboration with Cascade Design Automation (CDA), design tools for this purpose.

The paper is organized as follows. The next section gives an overview of the system. Following that, the MCM packaging is discussed. In Section 4 our work on I/O circuit design for area interconnect is discussed. Section 5 examines the design tools need ed to design MCM-based systems. Section 6 discusses how performance can be optimized if the architecture level and the MCM level are considered together. Section 7 concludes the paper.

2. System Overview

Figure 1 is a block diagram of the PUMA system. The heart of the system is the single chip FXU, a 32-bit fixed point processor that implements a subset of the PowerPC instruction set. The first level (L1) instruction cache is a 1 KB structure located on the FXU, assisted by an 8 line stream buffer. Floating point instructions are executed in a separate 64-bit IEEE-compliant floating point co-processor chip (FPU). The L1 data cache is composed of four 4 KB cache die for a total of 16 KB of L1 cache. The management for memory beyond the first level of cache is split across two die, the IMMU and the DMMU, to allow for concurrent accesses. Each of these memory management units is master over ten commercial cache die which implement 1 MB of level-2 (L2) cache on each side. The memory management units interface via a shared bus to a main memory controller (PIP) that also allows access to the main memory via the PCI bus. The main memory is comprised of 64 MB of synchronous DRAM.

The FXU, FPU, IMMU, DMMU and the first level cache die are currently under design and, as noted earlier, will be fabricated in Motorola's CGaAs technology. Several subsystems from these chips are already taped out for fabrication, so that we can obtain early feedback on the viability of the CGaAs technology. The PIP memory manager and PCI interface has been fabricated in Hewlett Packard's 0.5 micron CMOS through MOSIS. The L2 pipelined SSRAM (synchronous SRAM) and the synchronous DRAM are both commercial commodity parts.

The packaging techniques used in this design are critical to achieving system performance; a focus of this project is area-

interconnect flip-chip MCM packaging. This approach allows for large numbers of I/O per die which are used to provide high bandwidth connections between die on the MCM. Other anticipated benefits of this packaging approach are smaller integrated circuits for a given functionality, faster logic paths due to less total interconnect, lower power dissipation because of reduced parasitic capacitance, and the ability to restrict interconnect to the MCM substrate for better signal propagation characteristics. The I/O circuits and CAD tool developments are crucial to the ability to design for this packaging strategy.

3. Packaging Strategy Overview

Primary objectives of the fine-pitch, flip-chip, area pad array packaging approach are to make possible high bandwidth connections to the memory subsystem by supporting high I/O counts, and to minimize interconnect length throughout the processor. ICs having peripheral pads can be flip-chip mounted, but like wire-bonded or TAB-mounted parts, they require wide on-chip power rails to assure stable voltages in the core. Additionally, peripheral pads can dictate long interconnect routes from the core to bonding pads. Conventional "retro-fit" flip-chip array packaging methods redistribute signals from the periphery to an array of bumps covering the chip surface. While there are some advantages to these approaches, such as eliminating the need to reroute the IC, the power distribution situation is not helped, interconnect is made even longer, and the potential for reduction in the area of the die is not realized.

Figure 2 is a cross-sectional view of the PUMA MCM using the pad array packaging approach, showing CGaAs circuits flip-chip mounted on the top side of the MCM, with a heat sink attached to the back.

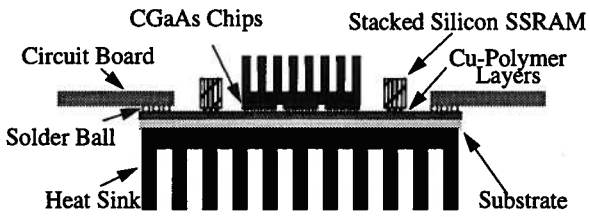

FIGURE 2. MCM Cross-section

Signal I/O pads will be distributed across the chip in an array of fine-pitch gold bumps. Each of these pads connects to receivers, drivers, and electro-static discharge (ESD) protection modules sized for MCM interconnect loads. These circuits will be distributed across the surface of the die in close proximity to the I/O bumps to minimize interconnect length between signal source or destination and the pad. Power pads will provide local Vdd and ground for each internal module, drastically reducing the size of power rails. Additional unused array sites can serve as thermal vias where needed.

One distinguishing feature of our packaging solution is the use of stacked commercial SRAM to minimize the area required on the MCM substrate as well as minimizing interconnect length to the second-level cache. The MCM will be mounted on a printed circuit board with solder bumps, to maintain a high I/O count.

4. MCM Signal I/O Circuits

In order to realize large numbers of high speed interconnect several issues must be addressed. Driver and receiver circuits must enable high speed point to point connections, while the number of I/O implies significant power consumption. Electromagnetic effects and termination schemes must also be considered, given the fast rise times (~100pS) inherent in CGaAs circuits. These electromagnetic effects include signal reflections, crosstalk, and self induced noise.

Self induced noise is caused by fast signal transition times. The use of bypass capacitors reduces this type of noise but not to an acceptable level. Reduced swing solutions reduce both self induced noise and dynamic power dissipation. Low swing driver and receiver circuits for use on an MCM, have already been designed and are currently being tested.

Crosstalk is the injection of one signal into another through parasitic capacitance or mutual inductance. Crosstalk effectively reduces signal rise and fall times. Cross talk can be minimized through careful routing on the MCM.

Reflections (ringing) are caused by an impedance mismatch at any point on the signal path. Reflection noise can distort data waveforms, and is particularly harmful when the reflection coincides with the data transitions at the receiver input. Conventional methods to reduce reflections include termination and reducing signal rise and fall times. Reducing rise and fall times is an undesirable solution as it slows the arrival of the output signal at its destination, degrading the data transmission rate. Passive termination is inappropriate for the proposed system due to the high static power dissipation. An alternative solution to this problem is to implement active termination. Several implementations exist for CMOS [1][2][3], however

CGaAs devices exhibit forward gate conduction at voltages exceeding 1V. A solution tailored to CGaAs that takes this and other CGaAs related issues into consideration is being studied. Alternatively, the length of the MCM signal traces can be chosen such that the reflections do not arrive at the receiver when the signal edge is in transition, but rather in the middle of the cycle when noise is less important.

Current mode, another possible signalling method under study, employs a low voltage swing driver and a low impedance receiver. Current mode signalling is inherently faster than voltage mode signalling because of the low input resistance of the receiver [4]. This type of signalling is very energy efficient since low voltage and current swings are employed. Implementations of driver and receiver circuits for CGaAs using current mode signalling are being studied.

5. CAD Tools

We have noted that allowing signal pads to be placed across the area of the die reduces routing length from the circuit to the bond pads, which in turn reduces area and power dissipation while enhancing performance. Further, array power pads eliminate the need for on-chip power distribution networks by providing local power taps for each module. This can lead to additional reduction in area (56% on a sample 500Mhz CGaAs core) and, thereby, fabrication costs [5]. Optimizing the layout of a circuit having distributed pads has been challenging because current CAD tools do not consider power rail sizing, pad allocation, signal routing and power distribution as interdependent [6][7]. Therefore, most optimizations are neglected until the final stage of the design cycle and are done manually.

What is desired for MCM design is a tightly integrated suite of CAD tools (Figure 3) which assist the designer in the power analysis, floorplanning, and the routing of array I/O designs.

In the system-on-an-MCM tool set that we are developing together with CDA, the power analysis tool aids Vdd and ground pad placement by graphically providing insight into the power dissipation and temperature rise of the design. The signal pad floorplanner analyzes the position of the area pads and blocks to recommend positions for the I/O buffer cells, thus improving system routing. The new placement is annealed and passed to a router for detailed routing. Net weights may be assigned to prioritize the allocation of area pads and thus optimize the placement of timing-critical I/O buffers. Finally, an

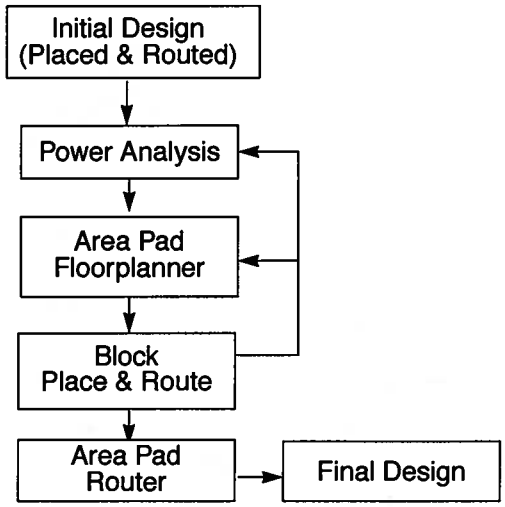

FIGURE 3. Design Flow for Area Interconnected Chip

area pad router is used to place and route bond pads across the face of the design. Signal pads are routed to ESD and buffer circuitry, while power pads are connected to local power rails. The area router is capable of using the top three metal layers to achieve this. Together, these tools address the pad placement and routing needs in the design of area interconnected flip-chips.

Also under development are CAD tools for partitioning a large design into multiple array interconnected flip-chips and their subsequent floorplan. This will permit the designer to minimize inter-chip routing through the MCM, thus reducing the power consumed by the drivers. Furthermore, performance increases will be realized by the reduced inter-chip latency.

With the growing complexity of MCM based circuits, we believe enhanced CAD tools, such as these developed by CDA and the University of Michigan, will become necessary for efficient MCM design.

6. Architectural Performance

Design decisions in the specification of the PUMA system were all supported by architectural studies performed using trace driven simulation. Before any significant specification decisions were made, simulations of the principal alternatives were performed.

Arch 1: I- 8k, off-chip, 3 cycle / D- 8k, off-chip, 3 cycle

Arch 2: I- 8k, off-chip, 3 cycle, next four lines prefetch / D- 8k, off-chip, 3 cycle

Arch 3: I- 1k, on-chip, 1 cycle / D- 16k, off-chip, 3 cycle

Arch 4: I- 1k, on-chip, 1 cycle, next four lines prefetch / D- 16k, off-chip, 3 cycle

Arch 5: I- 2k, on-chip, 1 cycle / D- 16k, off-chip, 3 cycle

Arch 6: I- 2k, on-chip, 1 cycle, next four lines prefetch / D- 16k, off-chip, 3 cycle

FIGURE 4. Average SPEC95 CPI for different PUMA configurations

The system we have chosen is Architecture 4 in Figure 4. It is the best performer given the size of instruction cache the FXU can support. Architecture 4 has a 1K byte on-chip instruction (I) cache and a 16 K bytes of data (D) cache off-chip. Since the decreased penalties due to cache misses are more significant, the overall effect of moving the first level data cache onto the MCM yields an increase in system performance.

The prefetching method in the studies of Figure 4 fetches the missed line and the next four lines. Recent studies suggest that a stream buffer of 8 cache lines, as mentioned with Figure 1 is better because it does not pollute the cache.

Architectural studies that include models for the delays due to chip crossings on the MCM are essential to assess the overall impact on system performance. Such studies have been advocated by us in earlier work [8][9][10], and they often indicate that the choice of a system-on-an-MCM yields superior performance to a system-on-a-chip. In our case, utilizing the high-speed interconnect available on an MCM allows increased performance from the memory system. This likely to be the case even with conventional silicon system although the die are likely to be much larger.

7. Conclusions

The widespread use of MCMs will require a number of technological advancements. Primary among these is the design of a suite of CAD tools which fully support the design of systems-on-an-MCM. The tools currently available to the digital hardware designer are inadequate for the task of designing an MCM. The tools generated for the PUMA system are a step in the right direction. While drivers and receivers have been widely studied, bringing signal traces onto the MCM expands the arena allowing higher speeds than were previously possible and is an area the deserves further investigation. The designers of MCM drivers and receivers can take advantage of the better signal propagation to reduce the power necessary for communicating over wide busses.

By utilizing an MCM, the interconnect to the upper levels of the memory system can be contained entirely within the MCM, and a large number of devices can be brought into close physical proximity. This allows the busses to be wider, and the interconnect to be run at a higher frequency. Using multiple die in a cache hierarchy enables significantly larger memory systems than are otherwise feasible. The PUMA processor being designed at the University of Michigan will rely on this and other attributes of MCMs to achieve high performance.

Acknowledgment: This work was supported by DARPA contract DAA H04-94-G-0327.

References

[1] Dolle, "A Dynamic Line-Termination Circuit for Multireceiver Nets," IEEE Journal on Solid State Circuits, vol 28, no. 12 Dec. 1993, pp. 1370-1373.

[2] R. Igarashi et al., "Effective low-power CMOS terminating circuit for coupled transmission lines," in *Proc 42nd Electronic Component and Technology Conf.*, San Diego, CA, pp. 814-822, May 1992.

[3] Kawahara et al., "Low-Power Chip Interconnection by Dynamic Termination," IEEE Journal on Solid State Circuits, vol. 30, no. 9 Sep. 1995, pp. 1030-1034.

[4] Seevinck et al., "Current-Mode Techniques for High-Speed VLSI circuits with Application to Current Sense Amplifier for CMOS SRAM's," IEEE Journal on Solid State Circuits, VOL 26, NO. 4 April 1991, pp. 525-535.

[5] R. Farbarik, X. Liu, M. Rossman, P. Parakh, T. Basso, R. Brown, "CAD Tools for Area-Distributed I/O Pad Packaging," *IEEE MCMC-97.*

[6] Banerjia, Glaser, Harvatis, Lipa, Pomerleau, Schaffer, Stansaski, Tekmen, Bilbro, & Franzon, "Issues in Partitioning Integrated Circuits for MCM-D/Flip-Chip Technology." *1996 MCM Conference*, pp. 154-159.

[7] Dehkordi, Powell, & Bouldin, "Development of a DSP/MCM Subsystem Assessing Low-volume, Low-cost MCM Prototyping for Universities." *1996 MCM Conference*, pp. 89-93.

[8] A.I. Kayssi, K.A. Sakallah, R.B. Brown, R.J. Lomax, T.N. Mudge, and T.R. Huff, "Impact of MCM's on system performance," *1991 Multichip Module W'shop*, University of California, Santa Cruz, March 1991, pp. 58–65.

[9] O.A. Olukotun, R.B. Brown, R.J. Lomax, T.N. Mudge, and K.A. Sakallah, "Multilevel optimization in the design of a high-performance GaAs microcomputer," *IEEE Jour. Solid-State Circuits*, vol. 16, no. 5, May 1991, pp. 763–767.

[10] O. Olukotun, T. Mudge, R. Brown, "Performance optimization of pipelined caches," *IEEE Trans. Computers*, to appear.