5.4 An Integrated Quad-Core Opteron™ Processor

J. Dorsey[1], S. Searles[1], M. Ciraula[1], S. Johnson[1], N. Bujanos[1], D. Wu[1], M. Braganza[1], S. Meyers[1], E. Fang[2], R. Kumar[2]

[1]AMD, Austin, TX
[2]AMD, Sunnyvale, CA

The next generation Opteron™ processor integrates 4 enhanced-performance x86 cores, each with 512kB L2 cache and an enhanced 128-bit FPU. The cores are integrated with a shared 2MB L3 cache and an enhanced on-chip memory controller that supports up to 4 16-bit HyperTransport™ links and a dual-channel 128-bit DDR2/DDR3 interface. The design contains nearly 450M transistors and is fabricated in a 65nm SOI CMOS process with dual stress liners and embedded SiGe for PMOS source/drains. The design uses 11 layers of copper interconnect (Fig. 5.4.1) that include advanced low-k dielectrics. In a $95W_{max}$ power envelope, the target frequency is 2.2 to 2.8GHz at 1.15V. The SoC chip is designed to facilitate maximum reuse of functional components and to provide the flexibility to create targeted variations.

The processor is based on a flexible clocking architecture designed to easily scale across die configurations. Each core contains its own PLL, clock distribution system, and power grid, which can be independently power/performance managed by varying the frequency and voltage. The core voltage and the individual core frequencies are independent of the on-chip Northbridge, allowing them to enter power-efficient states while the processor interface runs at full speed to service DDR memory and HyperTransport™ traffic. The clock distribution system is designed for a worst case systematic skew of 12ps in each core.

To provide clocking flexibility and reduce communication latency between the processor and the northbridge, an asynchronous FIFO buffer absorbs the global clock skews and clock rate variations. The FIFO enables a modular design style when building die with a large number of cores and is fundamental to minimizing the core-to-Northbridge latency. A synchronous mode is provided in the FIFO buffer for tester determinism. To provide reverse compatibility with previous AMD Opteron™ processors, this chip incorporates only one differential clock receiver to receive the 200MHz clock reference. The reference clock is distributed across the die to each core PLL, Northbridge PLL, HyperTransport™ link, and DDR memory interface. The reference clock network contains special power-supply-filtered buffers to reduce the clock jitter created by a large clock-tree network. Figure 5.4.2 shows the PLL locations and processor voltage domains.

For thermal control, each of the 4 cores contains 8 remote temperature sensors scattered across the core and connected to a thermal evaluation (TCEN) circuit. The Northbridge contains 6 additional remote sensors connected to a fifth TCEN circuit. The 5 TCEN circuits are connected to a global thermal control (TCON) circuit that instructs the remote controllers to collect thermal measurements and report the results. The remote sensor contains a diode array and additional control circuits. The temperature is determined by measuring voltage drops while forcing a range of currents. The TCEN block contains differential switched-capacitor integrators that perform the arithmetic operations and convert the analog temperature into a 9-bit digital value. A first-order $\Delta\Sigma$ converter is used for the A/D conversion. Additionally, amplifiers use auto-nulling circuits to cancel inherent offsets. The TCON tracks reported temperatures for comparison against predetermined threshold temperatures. Based on this analysis, the TCON can select from various power saving modes to reduce the die temperature.

The separate L1 instruction and data caches are each 64kB, 2-way set associative, and exclusive of the L2. ECC capable of 2-bit error detection and single-bit correction is maintained for each 64-bit word in the data cache. To support exclusivity, as well as updating the ECC on partial writes, the data cache supports a read followed by a write to the same memory address in one cycle.

The cache is implemented with a standard 6T memory cell by performing the read on the first half of the cycle followed by a self-timed write and precharge in the second half. The duration of the self-timed write is set based on extensive Monte Carlo analysis, encompassing local and lot-to-lot process variations. In addition, the duration of the write pulse can be modified after fabrication by programming electrical fuses. The process variation impact on write circuits and the fast access time requirements result in a memory cell size of $1.06\mu m^2$.

All of the L1 cache system components share a common memory cell and building blocks. The basic building block is shown in Fig. 5.4.3 and consists of two groups of 16 cells connected to a local sense circuit. As in [1], a domino read structure is chosen to reduce the access time sensitivity due to local process variation and to improve cell stability. In particular, short bit lines improve cell stability, producing less stress on the memory cell during a read. Unique aspects of the design include the NMOS push-pull writing technique and cross-coupled PMOS transistors to replace the traditional domino half-latch. Figure 5.4.4 illustrates the push-pull write, which improves cell stability by preventing the bit lines from being driven to V_{DD} while the wordline is asserted.

The 16-way set associative L2 and 32-way set associative L3 caches share a common bit-cell design. The read sensing structure is similar to that in the L1 cache. The cell size is $0.81\mu m^2$ and the gate lengths are extended to reduce static leakage. The L3 cache uses a modular design with variable latency regions. The latency regions and adjustable timing provide scalability across a wide range of L3 sizes. For both the L2 and L3 caches, data collection and transport is achieved via a serial-OR/AND mechanism.

The memory interface is handled through a DDR2/DDR3 combination PHY. The interface is capable of running DDR2 at 400 to 800Mb/s from a 1.7 to 1.9V supply or DDR3 at 800 to 1600Mb/s from a 1.4 to 1.6V supply. Support for SO-DIMM, U-DIMM, and R-DIMM is also included in the DDR PHY. The wide voltage range of 1.4 to 1.9V and the wide frequency range of 400 to 1600Mb/s, coupled with the unique architecture requirements of DDR3, required 88 unique DLL placements. Most of these DLLs have the capability of changing their delays for each write/read burst to provide optimal data versus strobe placement to each DIMM accessed.

The processor's I/O interface supports HyperTransport™ rev3 (HT3) specification [2], along with legacy HT1 and HT2 modes. The HT3 I/O is still a source-synchronous interface but supports dynamic skew compensation, transmit de-emphasis equalization and receive-side DFE ISI compensation. At data rate of 2Gb/s or below, it operates in legacy HT mode. It switches to HT3 mode operation at 2.4Gb/s and supports a peak data-rate of 5.2Gb/s. Figure 5.4.5 shows a block diagram of the I/O transceiver. The transmitter uses a voltage-mode driver with current-mode equalization to conserve power. The HT3 receive front-end employs on-chip AC coupling to address the wide input DC common range allowed by the specification. Baseline wander is compensated through a decision feedback mechanism.

Acknowledgements:
The authors gratefully acknowledge the global AMD team that made this design a reality.

References:
[1] J. Warnock, D. Wendel, T. Aipperspach, et al., "Circuit Design Techniques for a First-Generation Cell Broadband Engine Processor," *IEEE J. Solid State Circuits*, vol. 41, pp. 1692-1706, Aug., 2006.
[2] HyperTransport I/O Link Specification Revision 3.00, HyperTransport Consortium, http://www.hypertransport.org/tech/tech_specs.cfm.

 1-4244-0852-0/07/$25.00 ©2007 IEEE.

Layer	Pitch
Contacted Poly	250nm
M1	200nm
M2	200nm
M3	270nm
M4	270nm
M5	270nm
M6	400nm
M7	400nm
M8	400nm
M9	400nm
M10	800nm
M11	800nm

Figure 5.4.1: Interconnect pitches.

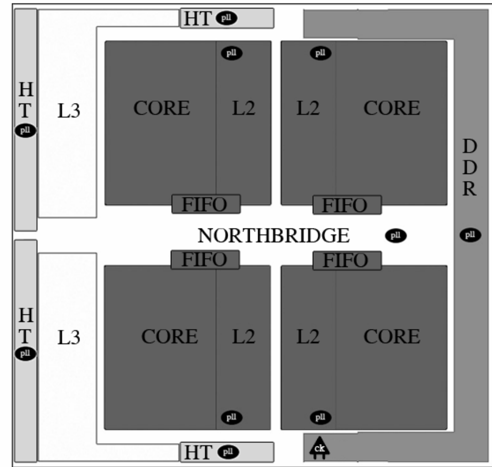

Figure 5.4.2: PLL locations and processor voltage domains.

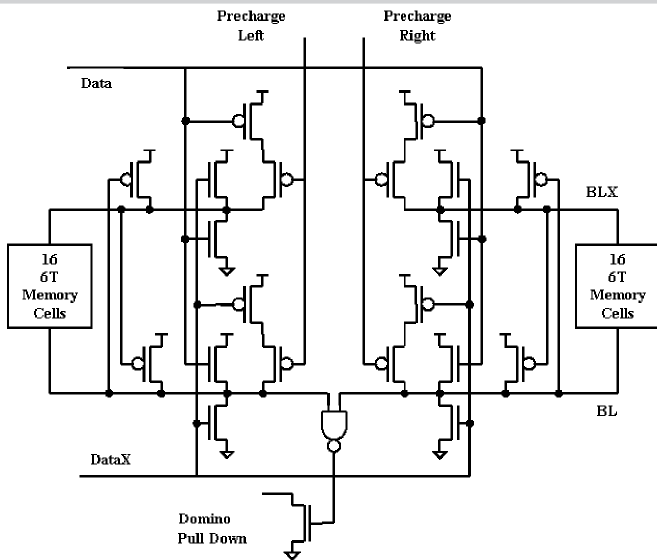

Figure 5.4.3: Cache building block with single-ended sense.

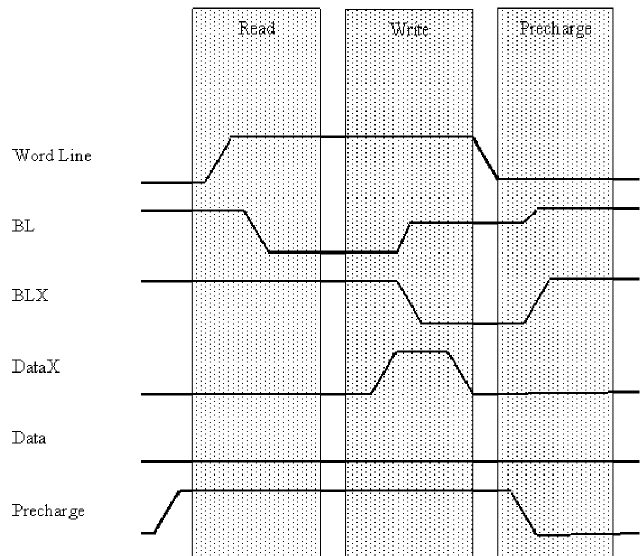

Figure 5.4.4: Read followed by wire timing.

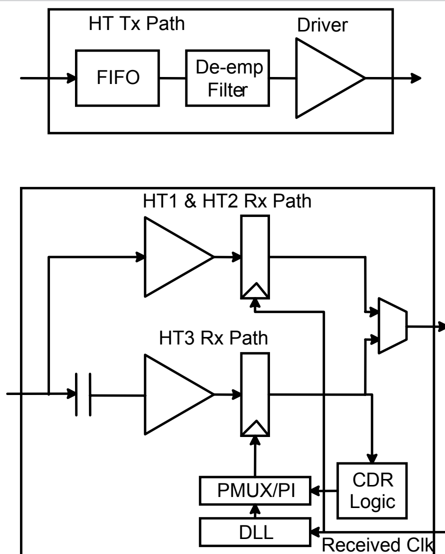

Figure 5.4.5: HT I/O transceiver block diagram.

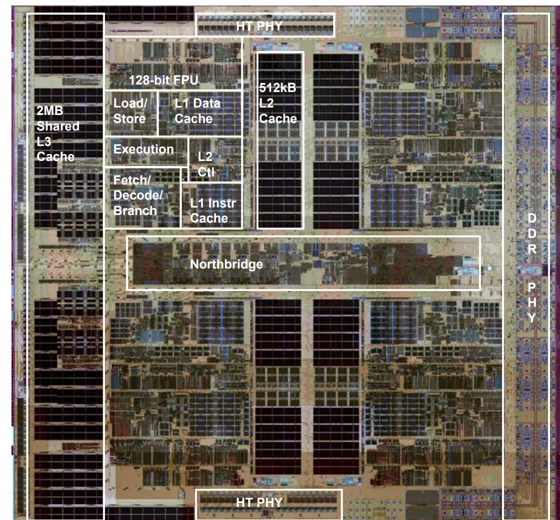

Figure 5.4.6: Die micrograph.