4.5 The Xeon® Processor E5-2600 v3: A 22nm 18-Core Product Family

Bill Bowhill[1], Blaine Stackhouse[2], Nevine Nassif[1], Zibing Yang[1], Arvind Raghavan[1], Charles Morganti[2], Chris Houghton[1], Dan Krueger[2], Olivier Franza[1], Jayen Desai[2], Jason Crop[2], Dave Bradley[2], Chris Bostak[2], Sal Bhimji[1], Matt Becker[1]

[1]Intel, Hudson, MA, [2]Intel, Fort Collins, CO

The next-generation enterprise Xeon server processor maximum configuration supports 18 dual-threaded 64b Haswell cores [1], 45MB L3 cache, 4 DDR4-2133MHz memory channels, 40 8GT/s PCIe lanes, and 60 9.6GT/s QPI lanes. The processor has 5.56B transistors on a 31.9mm×20.8mm die in Intel's high-κ metal-gate tri-gate 22nm CMOS technology [2] with 11 metal layers and achieves a 33% performance boost, on average, over previous generations [3]. Two additional metal layers enable area optimization and performance improvement. The design supports a wide range of configurations, including thermal design power ranging from 55 to 165W and frequencies ranging from 1.6 to 3.8GHz. Fig. 4.5.1 shows the processor block diagram of the 18 core die. The floorplan allows for core and cache communication via a ring interconnect, as well as the ability to deliver derivative designs with lower core counts. The 4th column in the 18-core die is removed to implement the 12 core chop, and the 3rd column is removed for an 8 core chop. Key architectural innovations include the addition of AVX2 technology, DDR4, and fully integrated voltage regulators (FIVR) [4] that enable per-core p-states and uncore frequency scaling.

Motherboard power delivery for high-performance microprocessors has become a significant design challenge that threatens the ability to cost effectively deliver high dynamic power at low voltage. To address this challenge, a fully integrated voltage regulator solution (FIVR) has been implemented, which enables several power-performance and platform optimizations. With FIVR, power is delivered to the die from a single mother-board voltage regulator (MBVR) at an elevated voltage of 1.8V, which results in a square-law reduction in power losses in the platform. FIVRs have a wide input voltage range of ±~200mV which simplifies the design of the platform delivery network and enables higher MBVR efficiencies. FIVR delivers cost reductions by reducing the number of MBVRs and the total number of VR phases required. The socket and package also require fewer power pins. The result is that total platform power is lower in the FIVR design, when compared to a classic MBVR solution at iso performance and cost.

Many FIVR domains are used on-die to regulate the voltage down to per-block optimized levels. Each FIVR domain is implemented as a 150MHz synchronous multiphase buck converter with up to 16 phases based on load. The FIVR LC output filter is implemented as package trace inductors (Fig. 4.5.2) and high-density on-die metal-plate capacitors. Up to 183 FIVR air core inductors, built into the package layers, increase package complexity. Inductors located in the pin field cavity use a circular inductor formed parallel to the package layers, while inductors outside this region use an inductor oriented perpendicular to the package planes. The multiple on-die voltage domains improve V/F optimization and enable cores to run at different V/F points concurrently. This capability improves performance in heterogeneous workloads, as individual cores can be placed in lower power states resulting in higher turbo frequencies for the remaining cores. In addition, process variation is mitigated as cores use different V/F curves to set operating points, which saves about 4W on average compared to a single V/F curve. Finally, voltage changes are faster and enable more frequent power-management optimizations.

Figure 4.5.3 shows the floorplan of the 31 FIVR domains. IO domains have separate FIVR digital and analog supplies, enabling independent power and supply-noise optimization for critical circuits. The cache logic and ring interconnect (CLR) voltage and frequency are independent of the cores. The CLR domain uses 3 regions, each with its own FIVR supply (VccR [0,1,2]). The 3 align with the boundaries of the derivative chops. The domain is power/performance optimized for a 10% slower frequency target than the cores. This optimization is motivated by the minimum operating voltage of the CLR being 50mV higher than the cores and its peak turbo frequency being lower than the cores. The lower design target is used to reduce power in the CLR logic. VccU is an always-on FIVR supply, which supplies the power control unit and global communication networks required to control the FIVRs. VccF is a FIVR infrastructure supply produced from a linear voltage regulator that operates before FIVR is active and enables the configuration of FIVR circuits.

The clock system architecture is modular in order to support the die chops. It includes up to 32 PLLs, as shown in Fig. 5.4.4(a). Two different types of PLLs are used: a self-biased PLL for all digital logic and DDR blocks and an inductor-based PLL for QPI and PCIe logic. Each core has its own clock domain, while the CLR domain is a single clock domain across three FIVR domains. All high-speed clock domain crossings connect to the CLR domain, as shown in Fig. 5.4.4(b). Clock synchronization uses a FIFO rate-matched buffer. The ring performance requires low-latency crossings of the 3 FIVR domains. A FIFO-like scheme provides >200ps of clock skew tolerance across these crossings, with no increase in latency. The global clock distribution is propagated under the common VccU supply and then level-shifted into each respective ring domain. Each CLR domain has an adjustable delay line in the clock distribution, enabling static clock compensation of process variation. Clock domains also include digital duty-cycle correctors.

This design is the first Intel CPU supporting DDR4 (1.2V), with data rates from 1333 to 2133MT/s and up to 3 DIMMs per channel. The same 4-channel memory IO also supports DDR3 (1.5V, 1.35V) from 800 to 1867MT/s and another single-ended signaling interface up to 3200MT/s. All analog circuits are multimodal to accommodate different signaling specifications and a wide range of data rates and voltages. The DLL supports 4× of frequency range (800-3200MHz) with optimal power and performance through bias trimming and capacitance tuning in delay cells. The transmitter (Tx) includes features such as two-tap voltage-mode constant-impedance de-emphasis-style equalization, current-mode (I-mode) swing boost, and a carefully implemented final driver and ESD, which significantly lower pad capacitance. Fig. 4.5.5(a) shows the diagram of the I-mode design, where additional pull-down current can be enabled to increase the voltage swing at the pad for WRITE operations. 5mA of current can increase far-end eye size by up to 20mV. The receiver (Rx), as shown in Fig. 4.5.5(b), includes a high-voltage tolerant amplifier, a 1st-order continuous-time linear equalizer (CTLE), a per-bit reference voltage, a per-bit de-skew, and separate sampling clocks for the even and odd data.

The digital and analog circuits were partitioned into two independent power supplies that can be individually adjusted to achieve maximum power efficiency. In normal operation, the analog supply powering critical analog circuits is set based on link margin at the particular data rate, while the digital supply is tuned based on the process corner and digital clock speed, saving 1.2W of active power. During various power states, they can be independently set to retention voltages or turned off for minimum idle (saving 2.1-2.8W) and leakage power. Finally, an innovative way to co-optimize performance and power was introduced by combining jitter and power into a single design metric. Overall, the memory interface consumes 40% less area and 38% less power than the previous product.

The high-speed I/O circuits are comprised of 40 lanes of PCIe (2.5/5.0/8.0GT/s), 4 lanes of direct-media interface (DMI) (2.5/5.0GT/s), and 60 lanes of QPI (6.4/7.2/8.0/9.6GT/s). In order to support the new 9.6GT/s operation, the CTLE on the Rx portion of the circuits used a new architecture to support higher bandwidth, increased peaking in the frequency response and greater amplifier linearity compared to previous circuits that operated at 8GT/s. See Fig. 4.5.6 for a block diagram. Furthermore, the clock circuitry implements a peaked frequency response small swing distribution across all the lanes to decrease power supply sensitivity and commensurately, jitter. Power consumption is decreased compared to the predecessor to 10.5pJ/b by moving significant digital logic from the analog supply to a lower voltage digital supply. Lane area is 10% smaller than the previous product.

Acknowledgements:
The authors thank the Intel Haswell server teams for their creativity, passion, and dedication in bringing this product to market.

References:

[1] N. Kurd, *et al.*, "Haswell: A Family of IA 22nm Processors", *ISSCC Dig. Tech. Papers*, pp. 112-113, Feb. 2014.

[2] C. Auth, *et al.*, "A 22nm High Performance and Low-Power CMOS Technology Featuring Fully Depleted Tri-Gate Transistors, Self-Aligned Contacts and High Density MIM Capacitors", *IEEE Symp. VLSI Tech.*, pp. 131-132, 2012.

[3] S. Rusu, *et al.*, "Ivytown: A 22nm 15-Core Enterprise Xeon Processor Family", *ISSCC Dig. Tech. Papers*, pp. 102-104, Feb. 2014.

[4] E.A. Burton, *et al.*, "FIVR – Fully Integrated Voltage Regulators on 4th Generation Intel Core SoCs", *IEEE Applied Power Electronics Conf.*, pp. 432-439, 2014.

 978-1-4799-6224-2/15/$31.00

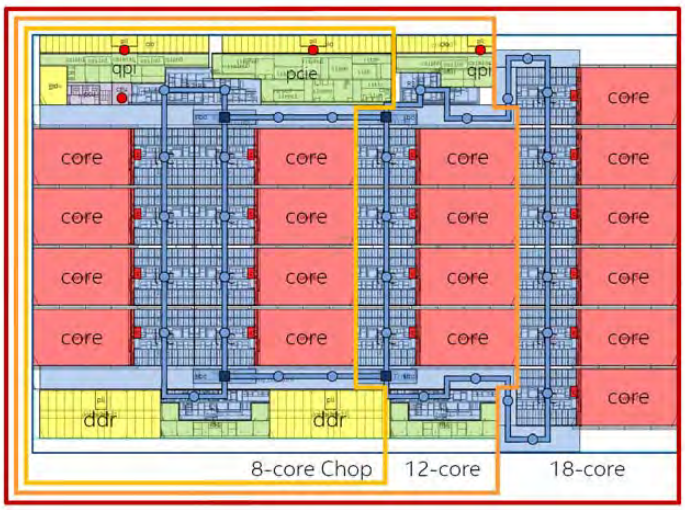

Figure 4.5.1: Processor block diagram and chop options.

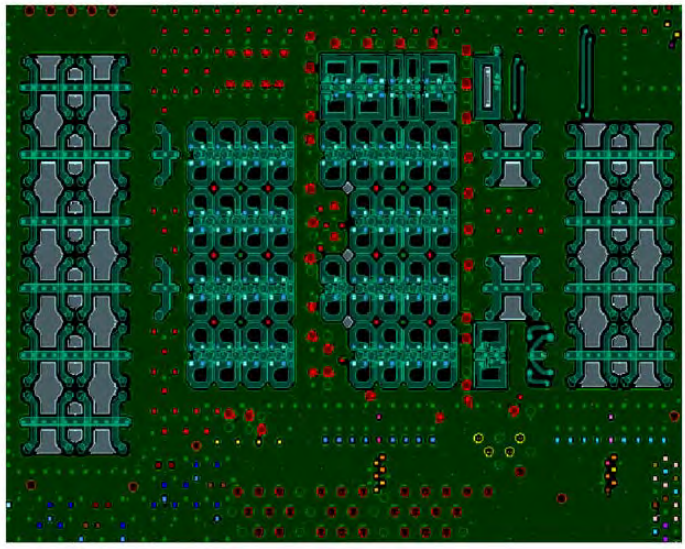

Figure 4.5.2: Package air core inductor (ACI) topologies over die shadow.

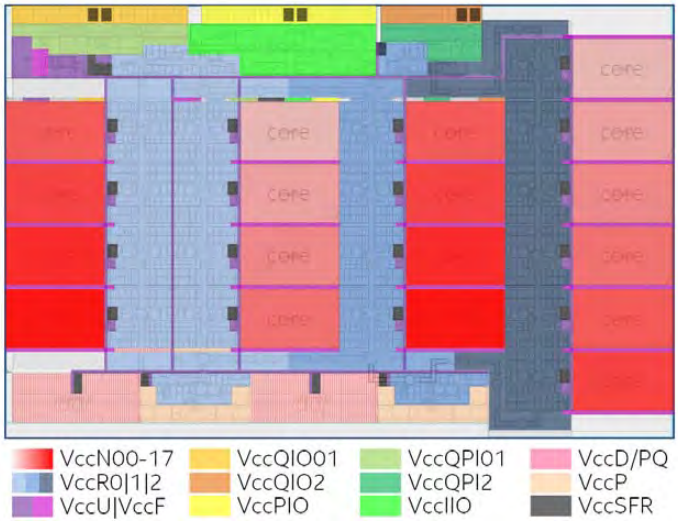

Figure 4.5.3: Fully integrated voltage regulator (FIVR) power domains.

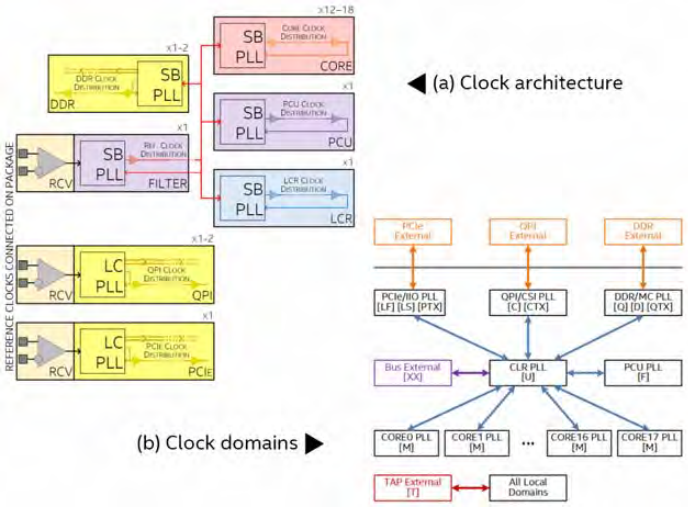

Figure 4.5.4: Clock architecture and clock domains.

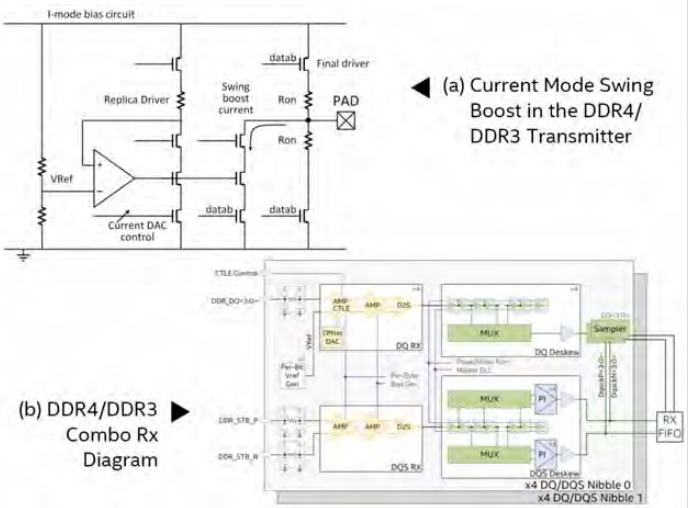

Figure 4.5.5: DDR4/DDR3 IO circuit architecture.

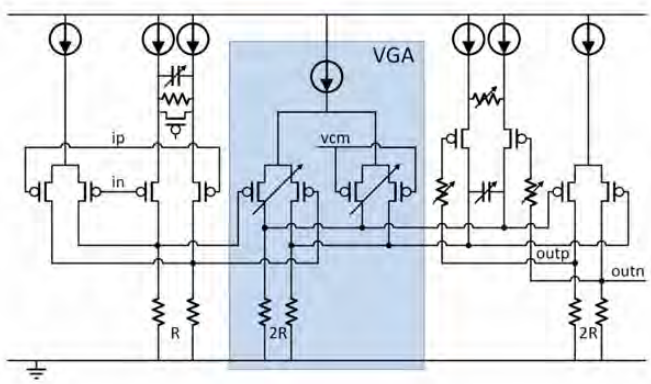

Figure 4.5.6: Continuous time linear equalizer (CTLE).

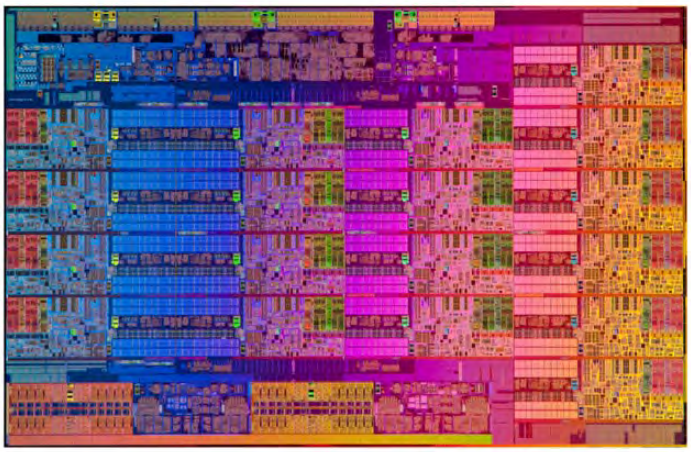

Figure 4.5.7: Intel® Xeon® processor E5-2600 v3 product family server die photo.

978-1-4799-6224-2/15/$31.00