

Zen: An Energy-Efficient High-Performance $\times 86$ Core

Teja Singh, *Member, IEEE*, Alex Schaefer, Sundar Rangarajan, Deepesh John, Carson Henrion, *Member, IEEE*, Russell Schreiber, *Member, IEEE*, Miguel Rodriguez, *Member, IEEE*, Stephen Kosonocky, *Senior Member, IEEE*, Samuel Naffziger, and Amy Novak, *Senior Member, IEEE*

Abstract—AMD’s next-generation, high-performance, energy-efficient $\times 86$ core, Zen, targets server, desktop, and mobile client applications with a 52% instructions per clock cycle (IPC) uplift over the previous generation. The increase in IPC complements a >15% process neutral reduction in CAC (switching capacitance). Performance and energy efficiency are further improved with various circuit techniques including write wordline boost, contention-free dynamic logic, supply droop detection with mitigation, a per-core frequency synthesizer, and a per-core integrated linear voltage regulator. Utilizing a 14 nm FinFET process, the Zen core complex unit consists of a shared 8 MB L3 cache and four cores.

Index Terms—14 nm, adaptive voltage and frequency scaling (AVFS), energy efficiency, finFET, high-frequency design, microprocessors, MIMcap, power management.

I. INTRODUCTION

ZEN is AMD’s next-generation high-performance, energy-efficient $\times 86$ Core designed for multiple market segments: laptop, desktop, and server [1, Fig. 1]. The architecture is built from-scratch and replaces AMD’s previous two core strategy of using Excavator [2] for high performance and Jaguar [3] for low power. Using Global Foundry’s energy-efficient 14LPP finFET process, making wide scale changes in architecture and updating the methodology allowed the team to create a single, scalable core that covers low-power and high-performance markets. The core design achieved a 52% increase in instructions per clock, while reducing process neutral CAC by 15% over the previous generation. Circuit techniques such as write wordline (WWL) boost, contention-free dynamic logic, supply droop detection with mitigation, and a per-core integrated linear voltage regulator further enhance both performance and efficiency.

II. ARCHITECTURE

Zen’s architecture is wider, has lower cache latency and supports higher bandwidths than previous generations. Major microarchitecture improvements include simultaneous multi-threading, enhanced hardware prefetching, and improved

Manuscript received May 6, 2017; revised July 25, 2017; accepted September 3, 2017. Date of publication November 28, 2017; date of current version December 26, 2017. This paper was approved by Guest Editor Muhammad M. Khellah. (*Corresponding author: Teja Singh*.)

T. Singh, A. Schaefer, S. Rangarajan, D. John, R. Schreiber, and A. Novak are with AMD, Austin, TX 78735 USA (e-mail: teja.singh@amd.com).

C. Henrion, M. Rodriguez, S. Kosonocky, and S. Naffziger are with AMD, Fort Collins, CO 80528 USA.

Color versions of one or more of the figures in this paper are available online at <http://ieeexplore.ieee.org>.

Digital Object Identifier 10.1109/JSSC.2017.2752839

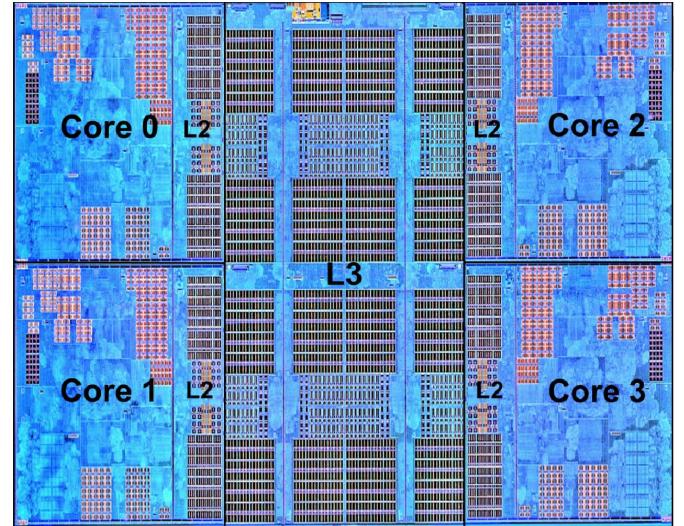


Fig. 1. Zen core complex die photograph.

branch prediction with two branches per branch target buffer entry. Items creating the instructions per clock cycle (IPC) uplift are shown in Table I [4]. Power was a main focus and the team tracked IPC and CAC very closely. Some architectural features were implemented that increased IPC while reducing CAC. For example, the Op Cache was utilized to store decoded instructions which increase the Ops/cycle, but also save power by reducing the pipeline length. Generally large IPC increases result in proportional increases in CAC [5]. However, for Zen, the team was able to actually decrease CAC in a process neutral fashion. This paper will walk through some of the approaches that enabled this achievement.

III. PHYSICAL DESIGN

Zen physical design contained major improvements in technology, design flexibility, and power. Zen uses a 12 layer telescoping metal stack containing 3 dual pattern layers and 4 1.25 \times layers. The design utilizes a 56 \times top metal layer and is AMD’s first implementation using metal-insulator-metal capacitors (MIMCap). The technology provides three V_T options with additional longer channel variants. The finFET process and modifications to internal timing closure methodologies allowed a wide range of voltage support and lower V_{MIN} than previous generations [2], [3]. Each Zen core has frequency and voltage control. The core has advanced speed,

TABLE I
MAJOR CONTRIBUTORS TO ZEN IPC UPLIFT [2]

Improved Core Design	Improved Cache System
Two Thread per Core	Lower latency L2 cache
Branch mispredict improved	Lower latency L3 cache
Large micro-op Cache	Faster load to FPU: 7 vs. 9 cycles
Wider micro-op dispatch 6 vs. 4	Better L1 and L2 data prefetcher
Larger Instruction Schedulers	Nearly 2X the L1 and L2 bandwidth
Integer: 84 vs 48 FP: 96 vs 60	
Larger retire 8 ops vs. 4 ops	5X Total L3 bandwidth increase
Quad Issue FPU	
Larger Queues: Retire 192 vs. 128 Load 72 vs. 44 Store 44 vs. 32	

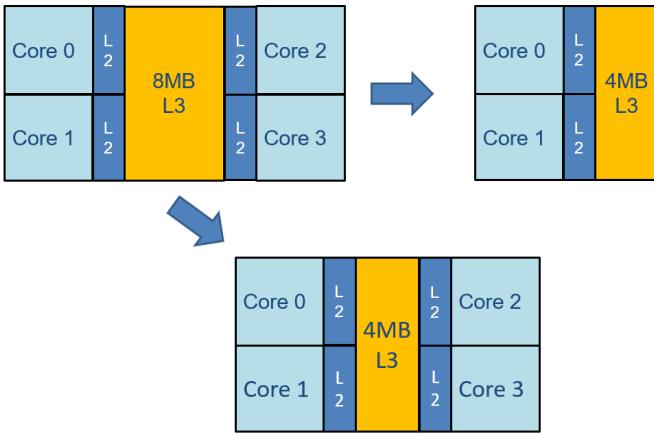


Fig. 2. Potential CCX variants.

voltage and temperature sensors that are used to dynamically change operating frequency and voltage. The core has third-generation adaptive voltage and frequency scaling (AVFS) [6] which is used for fine grained voltage and frequency control. Various sections of the design utilize different V_T devices based on their power profile.

A Zen Core complex, abbreviated CCX, is \sim 1.4 billion transistors and consists of 4 cores, 4 L2s and a shared 8MB L3 cache shown in Fig. 2. Each core has a private 512KB L2. The CCX is floor-planned such that the L3 acts as a crossbar which is a low latency solution for a four-core cluster. Multiple CCXs can be instantiated at the system-on-a-chip (SOC) level depending on the market segments. For client parts that do not require such a large L3 cache, a quad core 4MB L3 solution can be built. For lower power and cost sensitive markets, a dual core CCX with a smaller L3 cache can be built.

The CCX has various voltage domains to enable dense design and improved operating voltage range, as shown in Fig. 3. Real VDD (RVDD) is the core supply as distributed in the package and from this, local low drop-out regulators or LDOs, are used to create a VDD per-core. RVDD is the core LDO input and L3 peripheral logic supply, controlled by a high-efficiency platform voltage regulator with a granularity of 6.125 mV increments. RVDD is distributed to the cores through the package, while the core LDO provides a local per-core VDD supply at a finer granularity of 2–3 mV per step. VDD is routed in package metals across the Core.

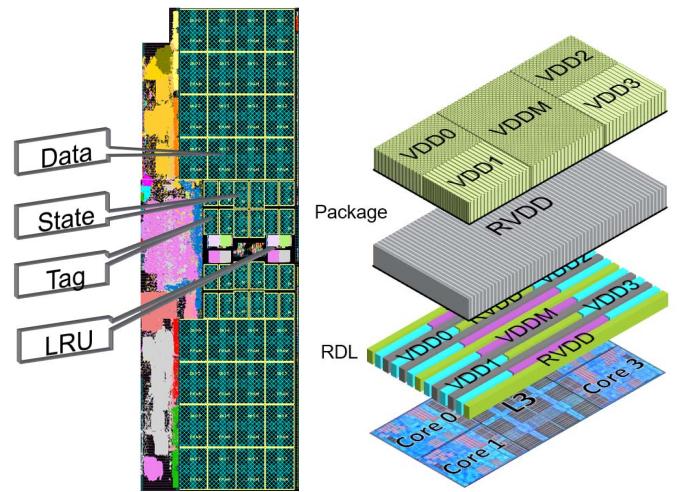


Fig. 3. L2 block diagram and CCX package stack.

The L2 and L3 static random access memory (SRAM) bitcells are powered by a separate voltage domain called VDD memory (VDDM) which enables dense SRAM and improved core VMIN. VDDM is supplied from an on-die low drop out regulator located at the SOC level. The VDDM package plane over the L3 is shared with the four VDD core domains; therefore, an on-silicon redistribution layer is used to route VDDM from the L3-Core boundary to the L2 SRAM macros. The non-SRAM logic over the L2 use VDD and the non-SRAM logic over L3 use RVDD.

IV. L2 AND L3 CACHES

The L2 is 8-way associative and strictly inclusive of the core instruction and data cache. The L2 supports a bandwidth of 32B to the IC/DC/L3 in each direction. The L2 contains three custom SRAM macros: the tag, the data and the state. The tag and data macros use high-current 6T bitcells powered by VDDM for density. The state macro uses an 8T bitcell for additional read and write bandwidth. The L2 state macro requires VDDM for stable operation due to write column muxing. Reliability, availability and serviceability features are included in the L2 for market scalability. Parity is checked for all ways during state and tag matching and full SECDED is performed on the selected hit or victim way. All L2 read

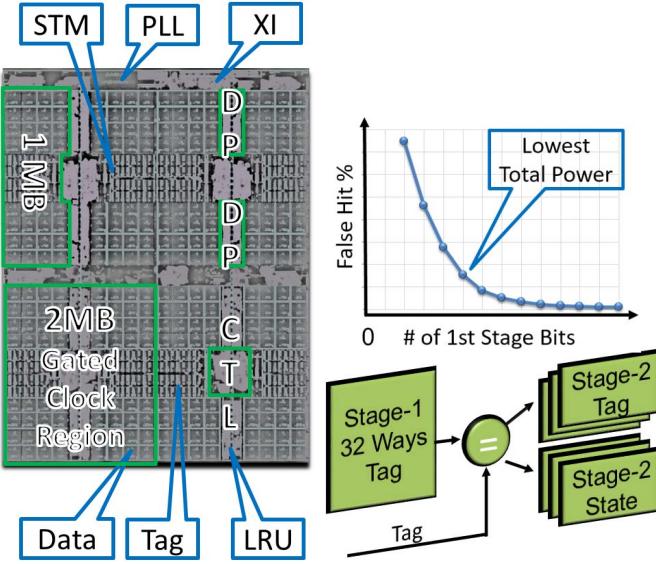


Fig. 4. L3 block diagram and shadow tags.

data is protected by DECTED. There is no performance penalty for implementing DECTED over SECDED because the calculation does not limit the read latency. DECTED C_{AC} is approximately $1.06 \times$ that of SECDED, so there is minimal power impact.

The $\sim 16 \text{ mm}^2$ L3 cache is organized into four slices which are interleaved by the low order address bits as shown in Fig. 4. While each L3 slice abuts to a given core, that core's data could live in any L3 slice. The shared phase-locked loop (PLL) resides in the L3 and creates clocks for all four cores and the L3. This methodology keeps the L3-Core interface fully synchronous. A level shifting FIFO is used on the Core-L3 interface to handle potentially different voltage levels and clock frequencies. Even if the cores are in C6 (i.e., power gated), the L3 can operate. In multi-CCX configurations, the L3 can flush itself. Power was a primary focus for the L3. To save leakage power, the L3 used mid V_T and high V_T devices and restricted low V_T exclusively for the clocks. The usage of high V_T was challenging because the L3 runs at core frequencies and voltages. The Zen clocking design introduced mesh gating. Four large clock meshes over each L3 slice are gated based on L3 activity. L3 clock mesh gating reduced L3 active power by 35% for average workloads and by 60% during idle. Additionally, the regions over L3 data macros have no clock mesh which contributed to a 40% reduction in clock load per area relative to the L2.

Special “shadow tag” macros reside in each L3 slice. Each slice duplicates the entries found in the L2 tag and state macros for the indexes contained within that slice. On an L2 miss or probe from another CCX, the shadow tags are checked in parallel to the L3 for valid data. This significantly reduces the bandwidth requirement on the L2 state and tags. The shadow tag consists of 3 custom 6T SRAM macros powered by VDDM. The stage1 shadow tag contains the LSBs of the L2 tag and all 32 ways (8 ways per core). If there is a hit in stage1, only those ways are enabled in stage2. If stage1 misses, the second stage is not accessed. In the

LEVEL SHIFTERS	SYMBOLS	USES
		• Clock level shift, isolating low
		• Clock level shift, isolating high
		• WL decode • WRCS decode • RDCSX decode • Various other VDDM decodes
		• Timing interlocks

Fig. 5. L2/L3 level-shifter variants.

graph in Fig. 4, as the bits used in stage1 increase, the power increases linearly, but the probability of a false hit drops exponentially. The graphic shows where the team optimized the number of stage1 bits for lowest total power. Accesses to the shadow tags use 76% less power compared to accessing all the L2 tags directly.

Zen supports a wide voltage range of operation, min–max range of $2.3 \times$ and up to a 500 mV gap between VDD and VDDM. To support high frequencies at V_{MAX} , most of the internal logic circuits remain in the VDD domain, requiring a large number of level shifters to convert decoded signals deep in the logic cone. A collection of level shifters shown in Fig. 5 is used in the L2 and L3 design. The top two level shifters in Fig. 5 are common level shifters, with isolation incorporated to handle different supply sequencings; the bottom two level shifters are pseudodynamic (PD) circuits with partially gated keepers that allow Z to be pulled low quickly, even when voltage source supply domain is significantly lower than voltage destination supply domain. These small, fast PD-level shifters are widely used in the designs, limiting the use of the common level shifters to a small handful per macro. As a result, the overall area overhead for incorporating the VDDM supply, including layout spacing penalties, averages under 1 percent across all the macros.

The placement and usage of these level shifters are further demonstrated in Fig. 6. The latch clock (*latchclk*) that captures macro index inputs is shaped with both VDD and VDDM clocks. The rising edge of *latchclk*, which determines the hold requirements of *index_low* and *index_high* inputs, is controlled solely by VDD, ensuring a hold time that will track with the $2.3 \times$ VDD range. The decoded indexes are driven to the PD WL drivers where they must hold until the *WLCLK* has fallen. The falling *latchclk* edge that opens the latches is shaped with a VDDM clock, ensuring this hold time is met without adding delay elements in the path.

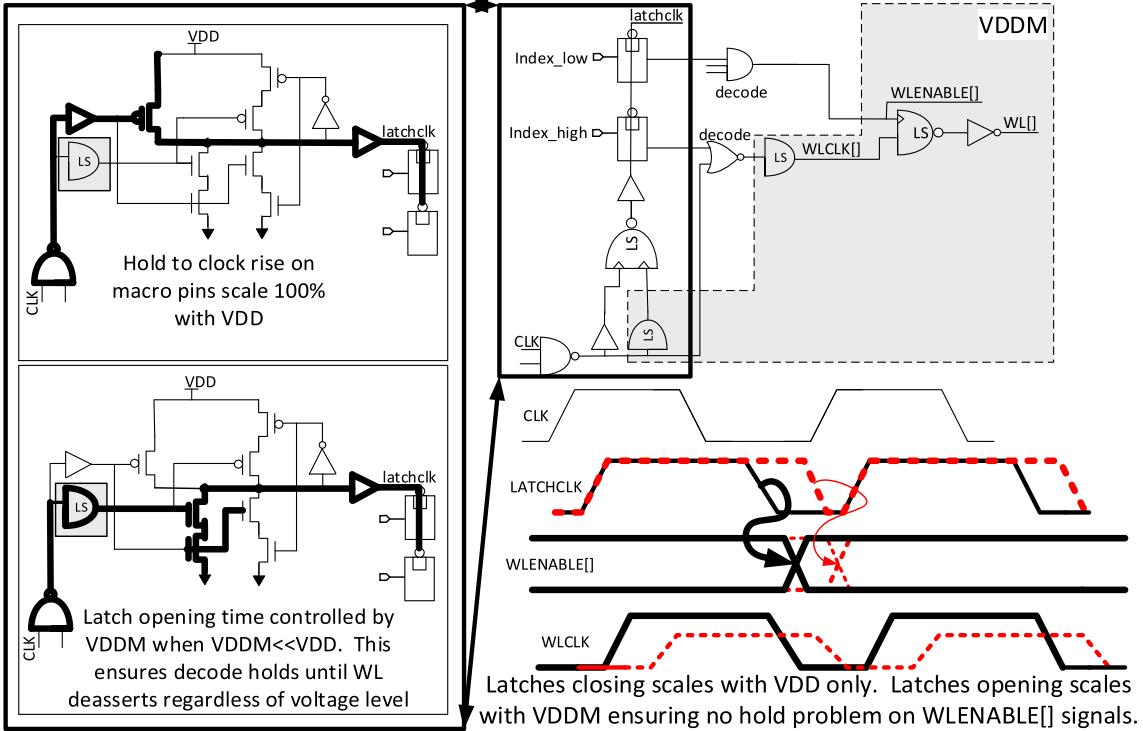


Fig. 6. Latch/clock interlock.

V. L1 ARRAYS

There are five size variants of the 8T bitcell L1 array macro that are used in ten different data structures including the instruction, operations, and data cache. Instruction and data cache lines can be invalidated by L2 probes and can be replaced based on a least recently used scheme. Zen uses a write back cache scheme which is more energy efficient than the write through scheme used in previous generations [2]. Additional bits are added in the data cache to achieve SECDED without requiring column muxed writes. The I/O bus is interleaved external to the macro in the semi-automated place and route (SAPR) regions achieving DECTED. Since the L1 does not require column muxed writes, nominal VDD can be used with a WWL boost assist scheme shown in Fig. 7. The L1 arrays are NOT placed on the dedicated memory supply, VDDM. Not having VDDM in the L1 macros reduces area, timing complexity, VDDM supply droop, and maximizes routing resources.

The WWL boost scheme is similar to the capacitive coupling techniques described in prior art [7] except that here full PMOS capacitors are used for boosting only one virtual supply corresponding to a group of WWL drivers. Compared to previous generations [2], a configurable JTAG signal allows for three different boost levels to the node labeled Virtual VDD based on total PMOS capacitor size. This provides flexibility for silicon tuning. The WWL boost scheme adds 3% area overhead to the macros but ensures that low voltage operation is limited by standard cell logic rather than the memory array. The increase in write power is 5% when WWL boost is enabled at low voltages.

Each macro contains a synchronous boost enable flop that enables/disables WWL boost. The boost enable flop in each macro receives its data input from the system management unit (SMU) which tracks all voltages. When a write occurs, the WWL is boosted above the core supply voltage if the boost enable flop contains a logic “1.” When the flop contains a logic “0,” the WWL is not boosted during macro write operations. The SMU ensures that the boost will not exceed V_{MAX} to avoid reliability concerns. On previous generations [2] the macro had to be disabled when switching to and from a boost state, but Zen has implemented a new two stage voltage transition that allows the macro to keep running as shown in Fig. 7. For example, at low voltages the SMU ensures that each macro’s boost enable flop contains a logic “1” when a write occurs. To transition the core to a V_{MAX}/F_{MAX} state, it first raises the core voltage to an intermediate level where WWL boost is not required, but if it is enabled, the boosted voltages will not exceed V_{MAX} . In the intermediate state, the boost enable is de-asserted. Then the core voltage can rise to higher levels without having any boosted WWLs exceed V_{MAX} . A similar two stage transition can be used when returning to lower core voltages where WWL boost is required. During these two stage voltage transitions, the macros can be accessed continuously.

Due to the relatively late arrival time of the boosted WWL, read after write hazard (RAW) timing to the same address is not guaranteed and is bypassed external to the macros to avoid an IPC penalty due to stalling. The bypass circuitry is essentially a multibit mux that normally selects the macro’s output, but will select the previous cycle’s macro input write

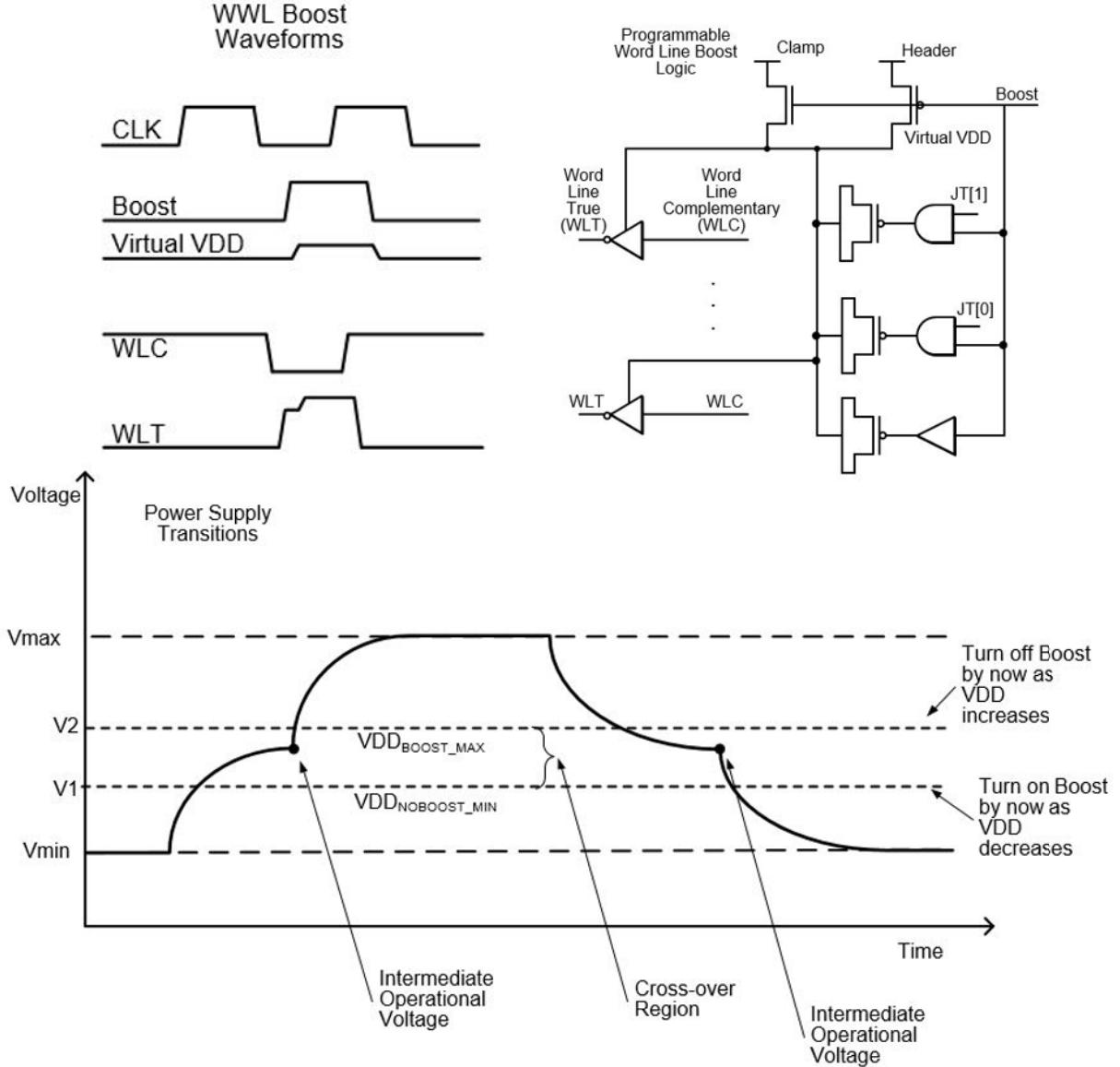


Fig. 7. L1 word line boost.

data whenever an RAW occurs to the same address. The muxes and associated control logic are instantiated external to the full custom macro in the SAPR regions. The registers that store the previous cycle's macro input write data are only clock gated when a macro write occurs.

Removing drive fights to enable a lower V_{min} is a priority on Zen. Previous generations used a traditional keeper on the L1 global bitlines in a domino read [8]. Zen implements a contention-free circuit shown in Fig. 8. It relies on a bus of *GBL* keeper enable signals, labeled as $KpEnGBL[m-1:0]$ in the schematic. Each of these signals are derived from the decoder as the logical NOR of their corresponding pair of *PchLBL* signals. The *PchLBL* signals are all driven to VSS during precharge and only one transitions high during reads causing only one $KpEnGBL$ signal to transition low corresponding to where an RWL is activated. If the global bitline node, labeled *GBL* evaluates to a 1, the PMOS stack

drives the GBL to maintain the precharge state. When the GBL pulls down, the bottom PFET in the stack is cutoff and the drive fight is removed ensuring excellent scaling to low voltages.

VI. LOW DROP-OUT REGULATORS

Previous generation power gating rings evolved into upper and lower banks on Zen with added support for digital low-drop-out voltage regulation. Fig. 9 highlights the header banks, the VDD package plane and the distribution of RVDD, VDD, and VSS bumps. A 15- μm -thick copper package layer is used to augment the redistribution of the header output current over the core. A high-efficiency platform voltage regulator module (VRM) is used to dynamically supply the highest voltage request from all cores running in the SOC. This in turn minimizes the drop out to the lowest voltage core, maximizing overall average system power delivery efficiency to greater

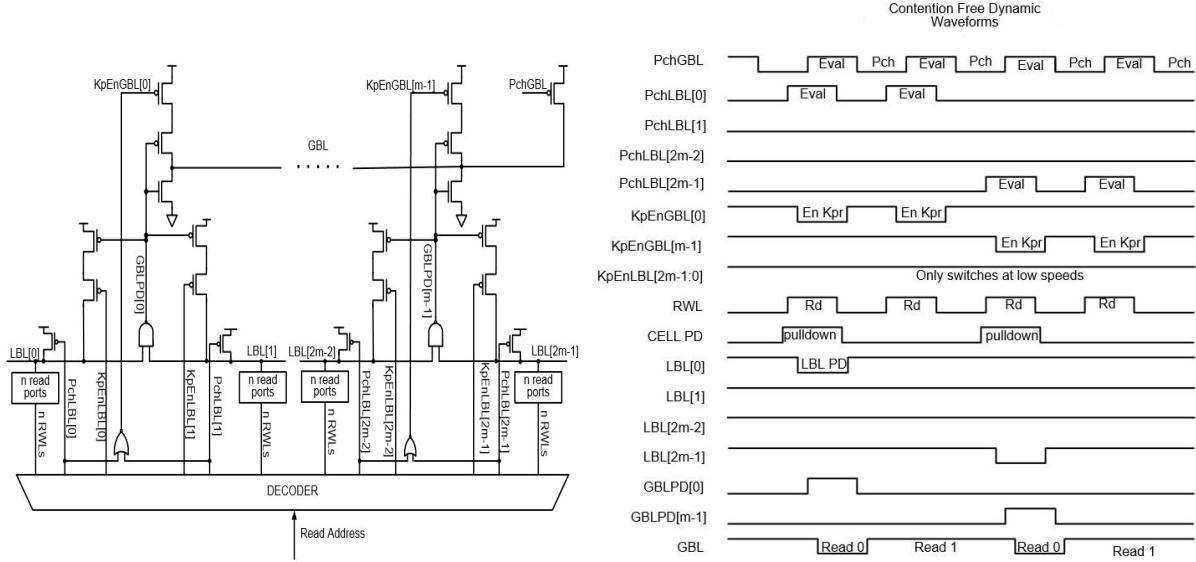


Fig. 8. L1 contention-free circuit.

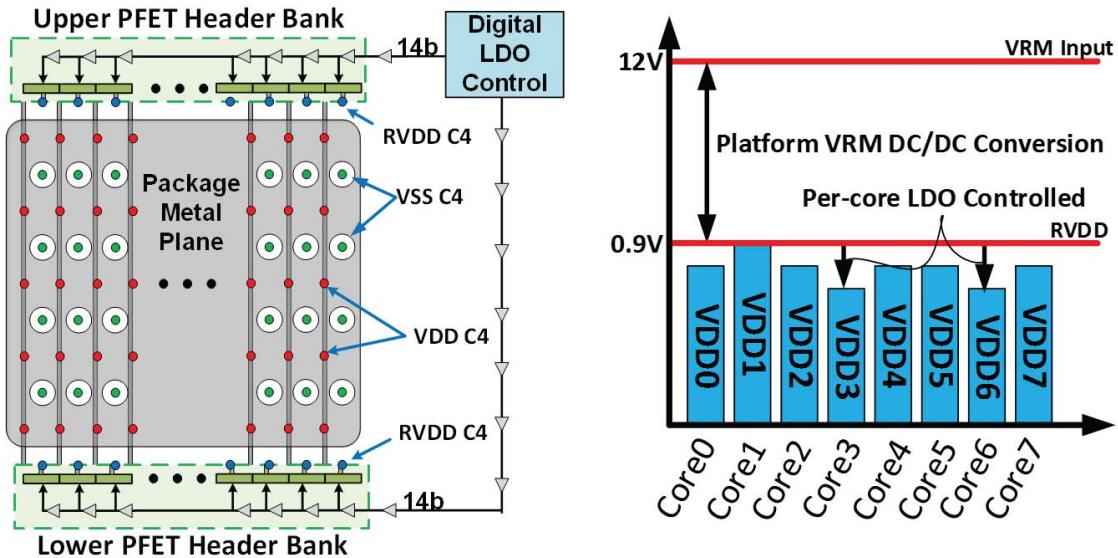


Fig. 9. LDO header banks and per-core voltage regulation.

than 90% relative to the base VRM efficiency. The bar chart in Fig. 9 shows 8 cores running at different voltages depending on the SOC power management.

Fig. 10 describes the dual-loop digital LDO implementation. A low-bandwidth, high-accuracy LDO loop uses a thermal design current-based power supply monitors (PSM) to measure VDD and RVDD voltages. It incorporates a fully digital configurable compensator clocked at 470 MHz, which calculates the required PFET header strength (*cntrlX[13:0]* in Fig. 10) required to match the VDD PSM measurement to a target PSM value supplied from the SMU. To cope with rapid di/dt events, a fast loop operates in conjunction with the low-bandwidth loop: a high speed fast droop detector monitors VDD and generates a signal (*DroopDetected* in Fig. 10) when it goes below a predefined threshold; *DroopDetected* is sent to top and bottom PFET header banks with <200 ps delay,

causing a predetermined number of headers to be turned on ($chargeInj[13:0]$ in Fig. 10). This mechanism rapidly shunts charge from the RVDD to the VDD rail to counteract the droop. An internal, calibrated Lookup-Table addressed by PSM RVDD and VDD voltage measurements automatically adjusts the $chargeInj$ strength depending on the operating conditions.

A simplified diagram of the fast droop detector is shown in Fig. 11. A Sigma-Delta modulator (SDM) translates a 10 bit voltage reference into a single bitstream that is then fed to a hard-macro low-pass filter and a single-ended voltage comparator where it is used to generate the droop threshold. A level-shifter supplied from a fixed voltage (VDD_AUX) and a low-pass filter are used to convert the bitstream from the SDM into a dc voltage at the input of the comparator circuit, which is made up of a chain of fast inverters supplied from

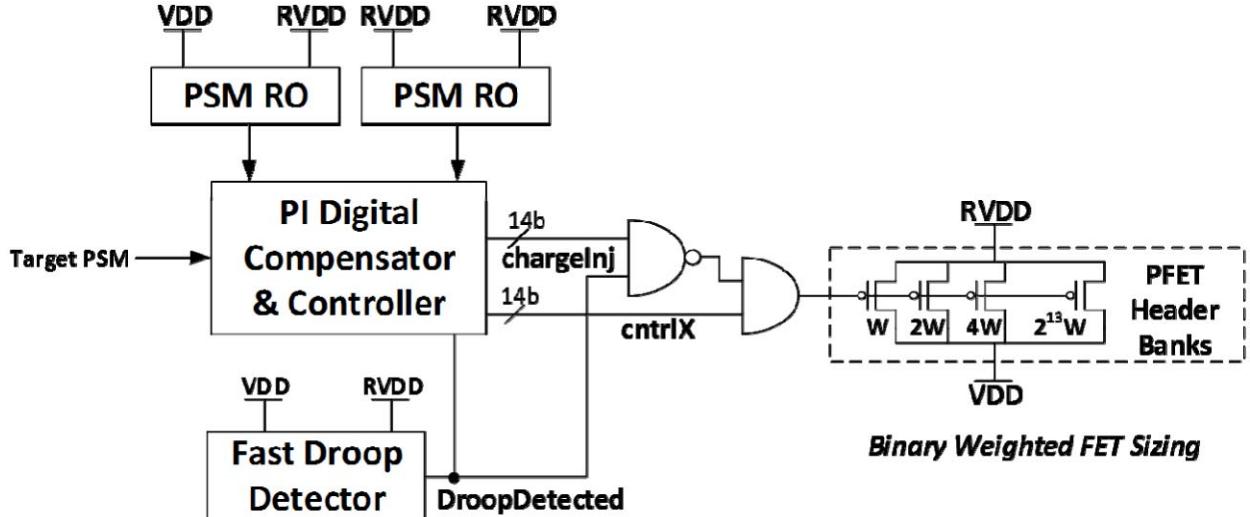


Fig. 10. Dual-loop digital LDO.

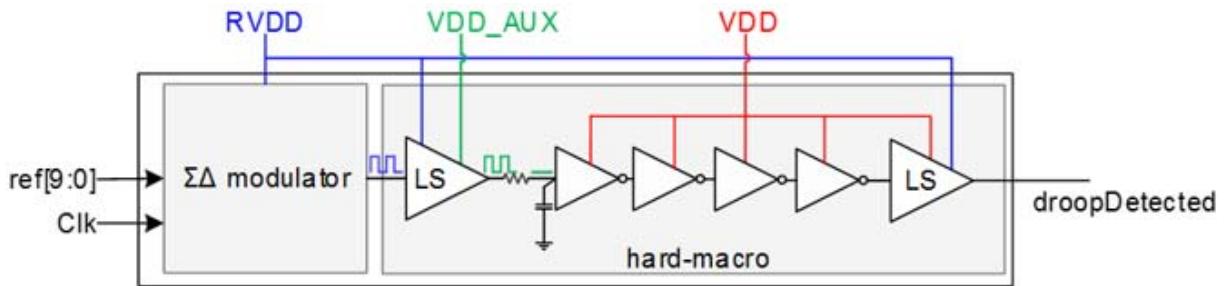


Fig. 11. LDO fast droop detector.

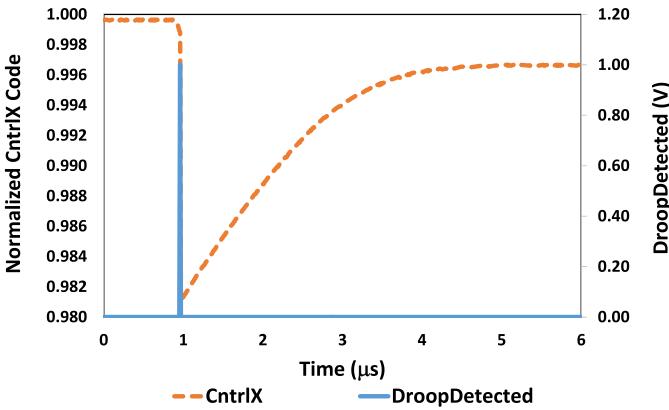


Fig. 12. Measured silicon response of CntrlX node.

VDD. The overall detection delay is <100 ps. Compared to previously reported approaches [9], this circuit uses almost no custom analog cells, except for the low-pass RC filter. Furthermore, a single SDM + fast droop detector can be used per core, adding minimal overhead. Functionally, the fast droop detector threshold is set below the target regulation voltage of the low-bandwidth loop, therefore triggering only when droop events take place. Note that in [9] each distributed regulator has a comparator operating

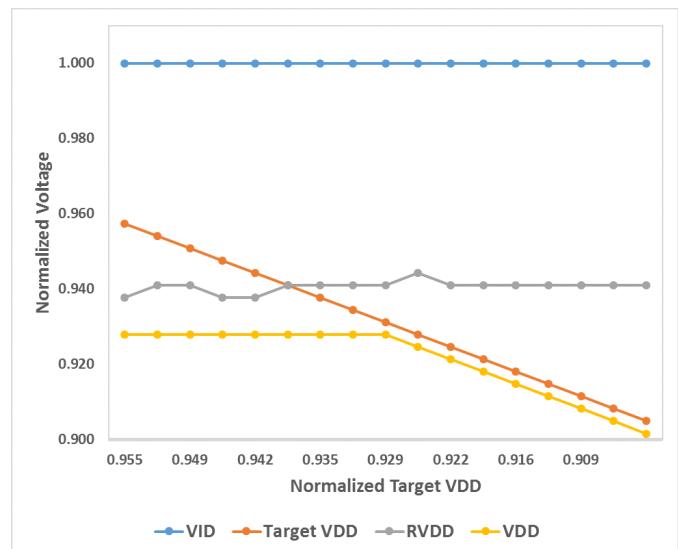


Fig. 13. Measured silicon results regulating a worst case stress pattern.

continuously at relatively high frequencies (1–2 GHz), decreasing LDO efficiency and increasing overhead.

Fig. 12 shows a silicon measured response of the (normalized) CntrlX code using an on-die debug state machine system

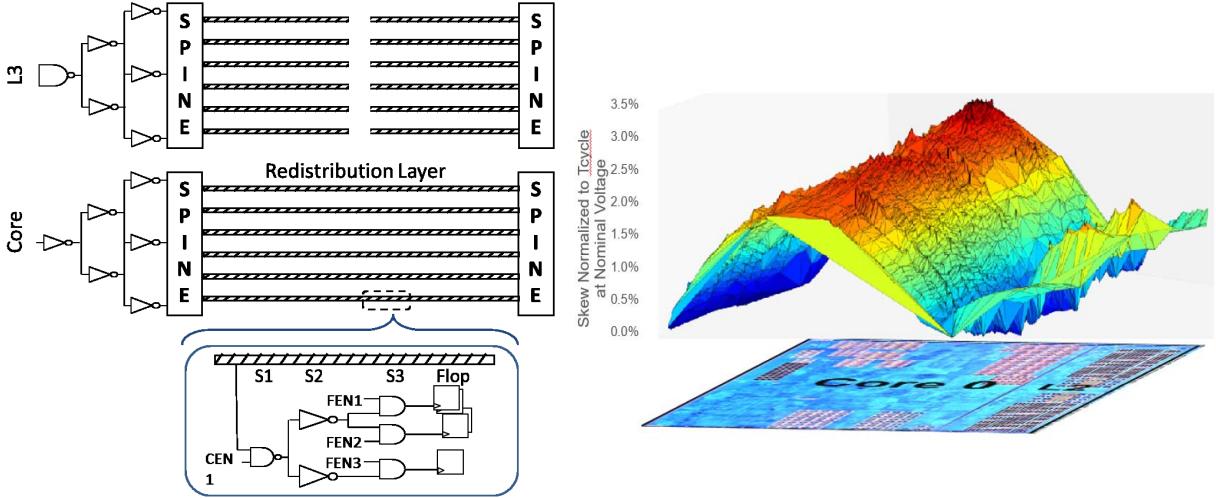


Fig. 14. Clock construction.

to acquire the real-time PSM signals. The solid line shows a droop event (*DroopDetected* signal) and the subsequent response of the fast loop, which quickly turns on more PFET headers to rapidly decrease resistance to counteract the droop. *DroopDetected* is also fed back to the low-bandwidth controller, causing it to quickly update *CntrlX* according to the *chargeInj* resistance level. The dotted line in Fig. 12 shows the response of the low-bandwidth loop bringing the system back to the average current level of the workload.

Fig. 13 shows PSM derived normalized voltages of the LDO system while regulating a worst-case stress pattern generating the maximum load current. In this measurement the input VRM is set with voltage identification (VID) code to produce a nominal drop-out voltage at idle. The worst stress pattern causes a load-line drop of the input RVDD voltage of the system. The “Target PSM” value is then swept from a voltage higher than RVDD (LDO input voltage) to a minimum operating voltage below RVDD. Fig. 13 demonstrates the robustness of the LDO system. When the target voltage is above the LDO input voltage the LDO saturates into a “self-bypass” mode, forcing a minimum header resistance. As the target PSM value is stepped lower, the LDO begins to regulate when the dropout is sufficient to sustain the target output voltage. The VDD signal indicates the LDO output voltage level, and shows correct regulation down to the V_{MIN} level.

VII. CLOCKING

The clocking methodology is well balanced between high-frequency specifications and energy efficiency. As shown in Fig. 14, Zen uses higher level metal to distribute the recombinant clock mesh. A configurable clock driver library is used to tune the mesh for skew and power. The flexible design allows for quick turnaround and silicon tuning. The diagram also shows a highly tuned clock skew plot superimposed over the CCX die plot. Zen requires a very tightly controlled clock mesh that is optimized for a high-frequency design. As shown at the bottom of the diagram, Zen uses a 4 logic level clocktree synthesis sourced from the mesh. This is a change from the

two logic levels used in a previous generation [2]. Zen uses coarse gaters with an efficiency >50%. Usage of coarse gating and gater cloning optimization reduced overall clock power as a percentage of total power by 30% over Jaguar [3] and 25% over Excavator [2].

A shared PLL is used per CCX and each core and L3 has a fine grain digital frequency synthesizer or digital frequency synthesizer (DFS) shown in Fig. 15. The fine grain DFS has programmability built in to tune the insertion delay up to 15% of the cycle and adjust duty cycle by 5% of cycle. As has been the case in high performance designs for over a decade [10], the Zen CCX mitigates the frequency impact of voltage droop with reactive frequency reductions. The fine grain DFS also allows each core and L3 to independently stretch their clock (Fig. 15). Since the efficacy of this approach is inversely proportional to the response time, the L3 and Core each have a coarse grain DFS to achieve only a 1 cycle delay from a voltage droop, to a reduction in the clock frequency which is at least a cycle faster than AMD’s previous generation [11] and the same or better than other industry implementations [12], [13]. The clocking for the caches and cores is all balanced to minimize domain crossing latency.

VIII. TIMING

In a change from previous generation timing methodologies, Zen was designed with multiple optimization points in mind. Looking at the normalized 14 nm frequency versus power plot (Fig. 16), the design covered the various market segments with three different optimization regions. Each region had a focus area and associated work effort level. The majority of the effort focused on the middle region and helped improve our perf/watt. The leftmost region was focused on gate dominated paths and roll-off at low voltages. Standard cells with high variation and poor scaling were banned. The rightmost Fmax region focused on RC dominated paths. Significant effort was put into wire engineering certain sections and particular nets to ensure high-frequency targets were met. The stars on the curve roughly show the design optimization points.

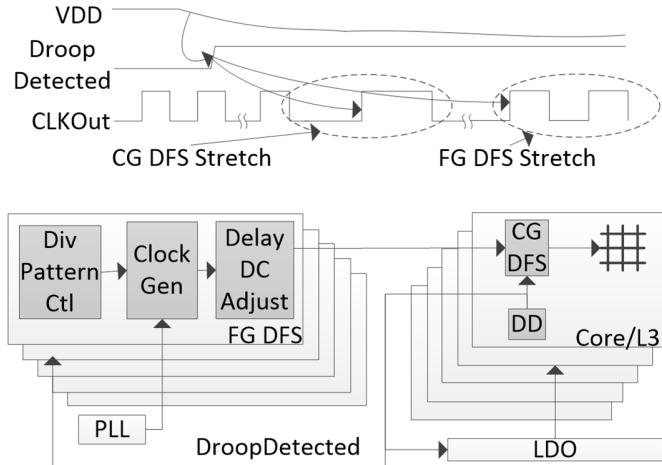


Fig. 15. Coarse grain (CG) and fine grain (FG) clock stretch.

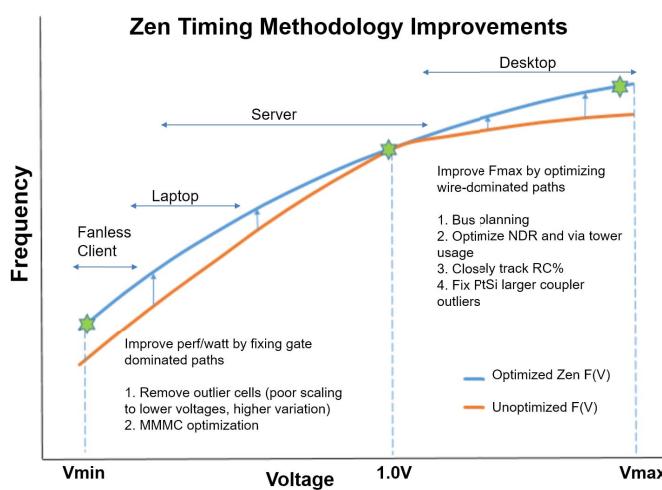


Fig. 16. Timing methodology.

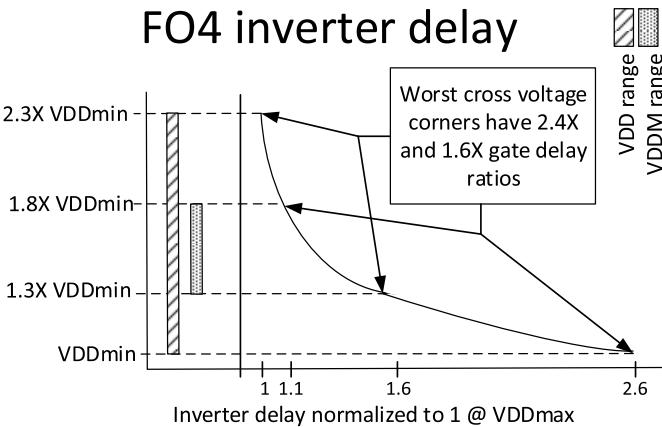


Fig. 17. Multi-voltage timing challenges.

The SRAM voltage, VDDM, adds timing complexity at voltage interfaces, as shown in Fig. 17. The worst-case FO4 inverter delay ratios on the two separate supplies are 1.6× when VDDM < VDD and 2.4× when VDDM > VDD. A path starting in VDD and ending in VDDM would have to close setup timing across low VDD to VDDM and hold timing across high VDD to VDDM. Also paths on VDDM

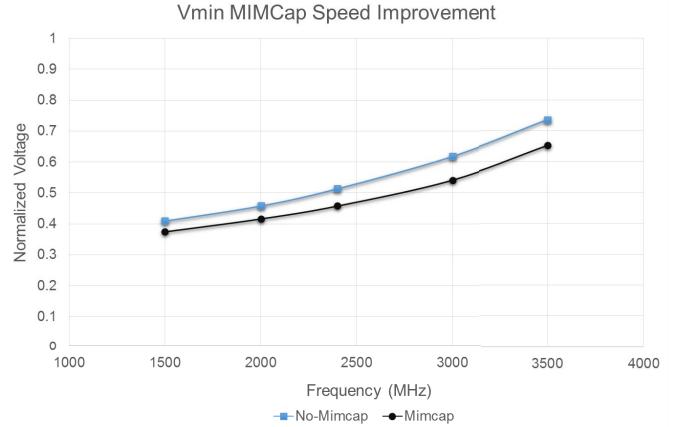


Fig. 18. Zen MIMCap improvement.

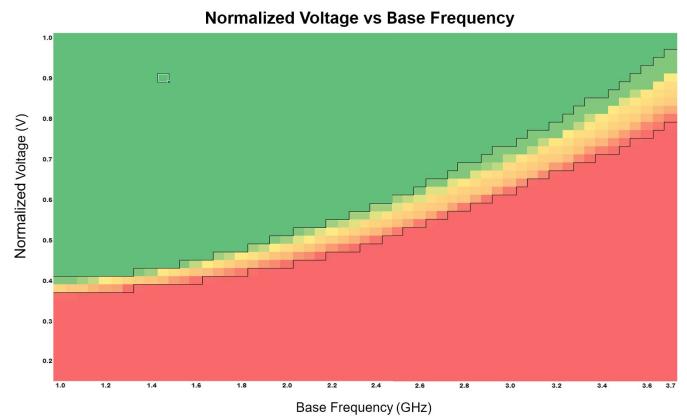


Fig. 19. 8c16t silicon shmoo.

have to be designed to run at VDD's Fmax even though VDDM is lower to ensure bitcell stability and long-term reliability.

Zen uses high-density MIM decoupling capacitors placed at the upper layers achieving ~45% area coverage in the Core. The team tried to optimize coverage, but there are several limitations to the coverage. There are technology required cutouts for power grid and bottom plate connectivity. Additionally, there are design required cutouts for sensitive nodes and timing critical regions. MIMCap is used primarily to mitigate droop events. The total decoupling capacitance provided by the MIMCap is ~6× the total decoupling capacitance provided by explicitly inserted decap cells (extrinsic FEOL decap). Decoupling can also be provided by inactive peer cells (intrinsic FEOL decap) which must be derated by activity so its total value is an upper limit. If MIMCap was not available, an additional 3–3.5 mm² per core of explicit decap cells would be needed to achieve comparable capacitance. As shown in Fig. 18, usage of MIMCap greatly improved Zen's frequency and overall energy efficiency.

Fig. 19 shows a silicon shmoo from an all-thread simultaneous multi-threading (SMT)-enabled Zen client SOC. The SOC has 2 CCX modules so it has 8 cores running with SMT-enabled totaling 16 threads. The shmoo shows the SOC hitting the all-cores base frequency of 3.6 GHz with headroom for overclocking. The 14 nm finFET technology and significant

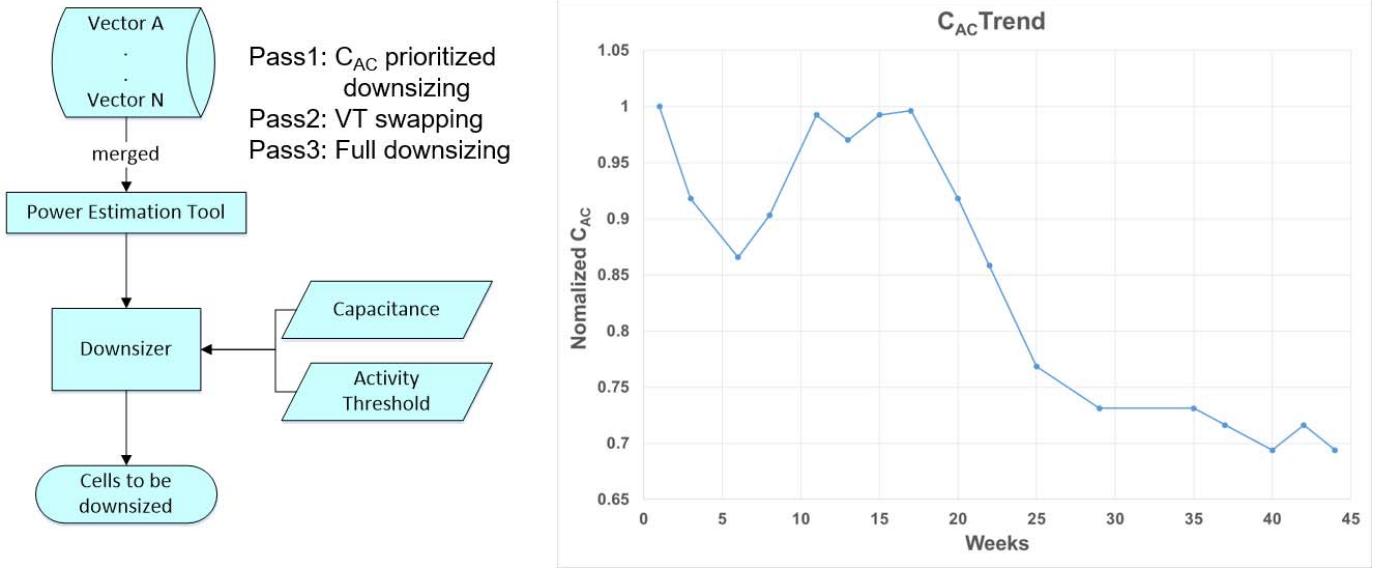
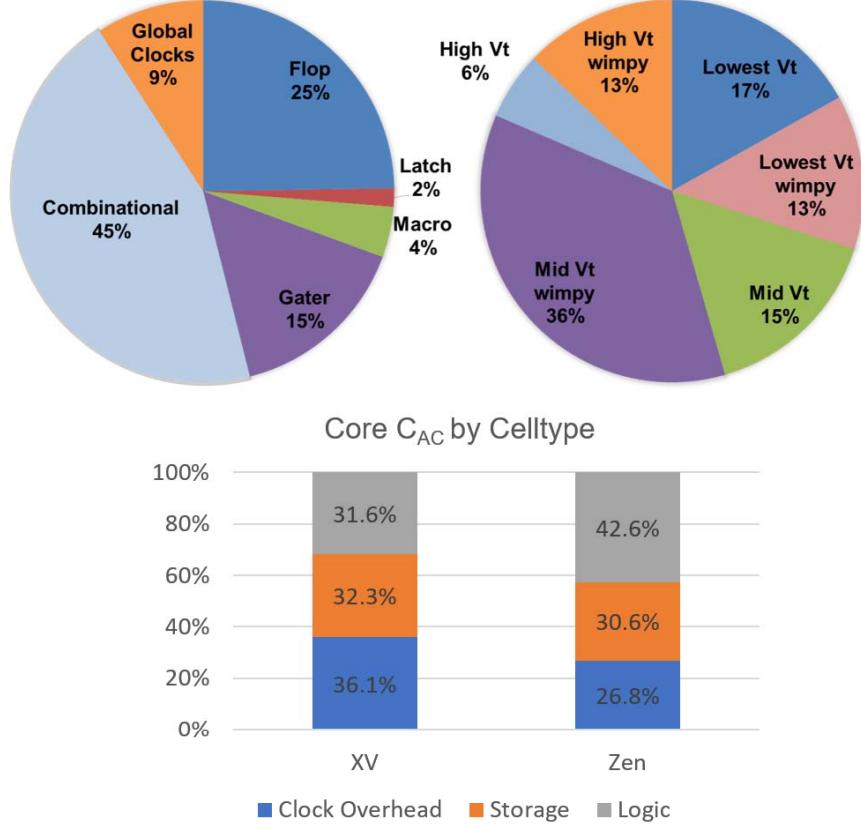
Fig. 20. Core optimization flow and C_{AC} trend chart.

Fig. 21. Core power breakdown.

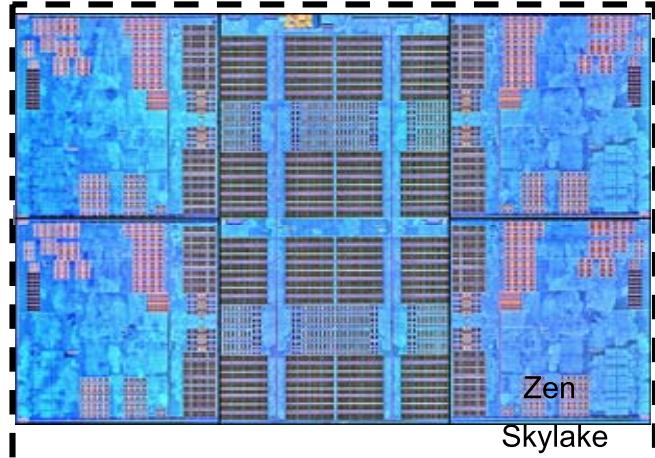
focus on design throughout the project allowed Zen to achieve this wide range of operation and lower V_{MIN} operation.

IX. POWER

The energy consumed per clock tick is proportional to $C \cdot V^2$ (with leakage power being left as overhead). Improving energy efficiency is therefore about doing more work per clock tick (IPC), while minimizing the amount of capacitance

switched to accomplish that work. The Zen optimizations essentially focused on minimizing any toggling of nodes that are not accomplishing useful work, which is done by aggressive fine grained clock gating, right-sizing gates and flops, and by minimizing overhead C_{AC} associated with clock distribution and sequentials.

A tight collaboration between the physical design team and the RTL team drove this elimination of useless toggling.



	Zen	Skylake [9]
Technology	14 nm	14 nm
Cores	4 Cores, 8 Threads	4 Cores, 8 Threads
Area	44 mm ²	49 mm ²
L2	512 kB 1.5 mm ² /core	256 kB, 0.9 mm ² /core
L3	8 MB, 16 mm ²	8 MB, 19.1 mm ²
CPP (nm)	78	70
Fin Pitch (nm)	48	42
1x Metal Pitch (nm)	64	52
Standard 6T SRAM (um ²)	0.0806	0.0588
Metal Layers	12 with MIM	13 with MIM

Fig. 22. Area comparison with state-of-the-art Skylake cores.

Initially, the team built micro benchmarks to locate high activity sections and to review gater efficiency. Later, to prevent re-introduction of excess toggling, the team wrote numerous power patterns and tracked them closely in physical design. The chart in Fig. 20 shows a C_{AC} burndown rate. The team has looked at timing burndown rates before, but this time the team treated power reduction and tracking on the same footing as timing. The chart here shows nearly a 30% power reduction in roughly a year once features were enabled.

The majority of the C_{AC} savings was achieved through architecture, physical design and clock gating. Clock gater enable logic cones were expanded to the extent that timing allowed to produce the most precise enablement or qualification. Redundant logic originally located in lower levels of repeated RTL hierarchical instances was removed from the lower level modules, and instantiated as little as possible. Logic was sometimes re-pipelined to reduce flop count. For example, initially data was read from two arrays of macros, stored in two banks of flops and then muxed in the next pipeline stage. Later, the macro C2Q timing improved. Then the mux was moved to the previous pipeline stage, immediately after the final mux stage between all macro instances, to remove a bank of flops.

Physical techniques such as structured design, buffer optimization, and wire engineering also minimized C_{AC} . The team used 4 bit flops and latches in structured arrays to share clock buffers. New V_T swapping and gate sizing flows were developed. A multi-pass C_{AC} aware approach employing

AMD's design rule check (DRC) aware V_T optimizer and power/timing downsizer aggressively optimized power. Pass1, which is shown in the diagram in Fig. 20, accounted for activity factors and did targeted downsizing on high activity nets. Pass2 did an aggressive DRC aware V_T swap algorithm and Pass 3 did the remaining timing aware downsizing.

The charts in Fig. 21 show the Core power breakdown for an average application. Flops and latches make up about 27% of the total, while gaters and clocks are around 24%. The remaining power is predominantly combinational logic. It is important to note, that as shown in the bar chart in Fig. 21, the aggressive reductions in clock and flop C_{AC} resulted in a much larger fraction of the total C_{AC} being used for computation (logic), rather than overhead. The 35% increase in C_{AC} fraction allocated to logic than on the prior core design is one of the reasons the core IPC could be increased so dramatically and still be lower power. While more than half of the devices used for the core are mid- V_T , the next highest usage, at about 30% of the total, are lowest V_T devices with about half of those having slightly longer channels (wimpy) to control leakage. The higher usage relative to previous generations is acceptable due to the improved characteristics of the 14LPP finFET devices and achieves performance and power goals.

Since area is indirectly related to power, it was also a focus during the project. Smaller area results in less leakage, shorter wire lengths for large data buses and typically has less total switching logic. Fig. 22 shows a core area comparison to the

Skylake core [14], [15]. It is clear that even with a less area efficient 14 nm process, the Zen core still occupies about 10% less area. Although the Zen core does not support AVX256, it does have twice the L2 cache size. In our floorplans, these two architectural differences would approximately cancel out, resulting in a fair comparison.

X. CONCLUSION

Zen is AMD's new, ground-up design core targeted for a wide range of applications: client, desktop, and server. A CCX configuration is 4 cores, 2MB L2, and 8MB L3. Versus AMD's previous generation, the design sees a 52% IPC increase while reducing $C_{AC} > 15\%$ process neutral. Various circuit techniques including WWL boost, contention-free dynamic logic, supply droop detection and mitigation, and a per-core frequency synthesizer and integrated linear voltage regulator further improve both performance and efficiency. Zen silicon shows 8c16t (8 cores with 16 threads running) base frequency of 3.6 GHz+.

REFERENCES

- [1] T. Singh *et al.*, "Zen: A next-generation high-performance ×86 core," in *IEEE Int. Solid-State Circuits Conf. (ISSCC) Dig. Tech. Papers*, Feb. 2017, pp. 52–53.
- [2] B. Munger *et al.*, "Carrizo: A high performance, energy efficient 28 nm APU," *IEEE J. Solid-State Circuits*, vol. 51, no. 1, pp. 105–116, Jan. 2016.
- [3] T. Singh, J. Bell, and S. Southard, "Jaguar: A next-generation low-power x86-64 core," in *IEEE Int. Solid-State Circuits Conf. (ISSCC) Dig. Tech. Papers*, Feb. 2013, pp. 52–53.
- [4] M. Clark, "A new X86 core architecture for the next generation of computing," *Hot Chips*, vol. 28, p. 6, Aug. 2016. [Online]. Available: http://www.hotchips.org/wp-content/uploads/hc_archives/hc28/HC28.23-Tuesday-Epub/HC28.23.90-High-Perform-Epub/HC28.23.930-X86-core-MikeClark-AMD-final_v2-28.pdf
- [5] M. Horowitz, E. Alon, D. Patil, S. Naffziger, R. Kumar, and K. Bernstein, "Scaling, power, and the future of CMOS," in *IEDM Tech. Dig.*, Dec. 2005, pp. 7–15, doi: 10.1109/IEDM.2005.1609253.
- [6] A. Grenat *et al.*, "Increasing the performance of a 28 nm x86-64 microprocessor through system power management," in *IEEE Int. Solid-State Circuits Conf. (ISSCC) Dig. Tech. Papers*, Feb. 2016, pp. 74–75.
- [7] R. V. Joshi, M. M. Ziegler, and H. Wetter, "A low voltage SRAM using resonant supply boosting," *IEEE J. Solid-State Circuits*, vol. 52, no. 3, pp. 634–643, Mar. 2017.
- [8] R. Jotwani *et al.*, "An x86-64 core implemented in 32 nm SOI CMOS," in *IEEE Int. Solid-State Circuits Conf. (ISSCC) Dig. Tech. Papers*, Feb. 2010, pp. 106–107.
- [9] Z. Toprak-Deniz, "Distributed system of digitally controlled microregulators enabling per-core DVFS for the POWER8 microprocessor," in *IEEE Int. Solid-State Circuits Conf. (ISSCC) Dig. Tech. Papers*, Feb. 2014, pp. 98–99.
- [10] T. Fischer, J. Desai, B. Doyle, S. Naffziger, and B. Patella, "A 90-nm variable frequency clock system for a power-managed itanium architecture processor," *IEEE J. Solid-State Circuits*, vol. 41, no. 1, pp. 218–228, Jan. 2006.
- [11] K. Wilcox *et al.*, "Steamroller module and adaptive clocking system in 28 nm CMOS," *IEEE J. Solid-State Circuits*, vol. 50, no. 1, pp. 24–34, Jan. 2015.
- [12] Y. YangGong *et al.*, "Asymmetric frequency locked loop (AFLL) for adaptive clock generation in a 28 nm SPARC M6 processor," in *Proc. IEEE Asian Solid-State Circuits Conf.*, Kaohsiung, Taiwan, Nov. 2014, pp. 373–376.
- [13] M. S. Floyd *et al.*, "Adaptive clocking in the POWER9 processor for voltage droop protection," in *IEEE Int. Solid-State Circuits Conf. (ISSCC) Dig. Tech. Papers*, Feb. 2017, pp. 444–445.
- [14] M. Bohr, "14 nm process technology: Opening new horizons," in *Proc. Intel Developer Forum*, 2014.
- [15] L. Gwennap, *Skylake Speedshifts to Next Gear*. Microprocessor Report, Sep. 2015.



Teja Singh (M'17) received the B.S.E.E. degree from The University of Texas at Austin, Austin, TX, USA, in 1999.

He was with Digital Equipment Corporation, Austin, ARM, Austin, Cadence, Austin, and Alchemy Semiconductor, Austin. He has over 20 years of high-performance, low-power microprocessor design experience across x86, MIPS, and ARM architectures. He was a Key Technical Lead in various programs, such as Griffin, Bobcat, Jaguar, and Zen, AMD, Austin. He is currently the Zen and Next-Generation Core Circuit Methodology Lead with AMD.



Alex Schaefer received the B.A.Sc. degree in electrical engineering from the University of Waterloo, Waterloo, ON, Canada, in 2000, and the M.A.S. degree in electrical engineering from the University of Illinois at Urbana-Champaign, Urbana, IL, USA, in 2001.

In 2002, he joined Advanced Micro Devices, Austin, TX, USA, as a Circuit Designer. He was with both Silicon Debug and CAD flow teams. He is currently the Analysis Lead for a next-generation product. His work primarily focuses on full custom arrays for CPU products from 65 nm down to 14 nm.



Zen family in the areas methodologies.



Deepesh John received the B.E. degree in electrical and electronics engineering from BITS, Pilani, Pilani, India, in 2004, and the M.S. degree in electrical and computer engineering from the University of Michigan, Ann Arbor, MI, USA, in 2006.

In 2006, he joined AMD, Austin, TX, USA, he was involved in high-speed custom memory array design, high-speed clock distribution, and custom circuit design methodologies on several generations of processor cores, where he is currently a Principal Member of Technical Staff. He is currently the Clock Design Lead for high performance x86 core, cache and graphics core, and the clocking architect for server SOCs.



Carson Henrion (M'97) received the B.S.E. degree in electrical specialty and the B.S. degree in engineering physics from the Colorado School of Mines, Golden, CO, USA, in 1997.

From 1997 to 2007, he was with Hewlett-Packard, where he was involved in SRAM, SerDes, silicon debug, and DDR custom circuits for PA-RISC microprocessors and chipsets. He was involved in the design and power optimization of L2 and L3 cache with AMD, since 2007, where he is currently a Senior Manager. He currently manages L2, L3, and SRAM macro design teams for AMD's next-generation microprocessor. He holds twelve U.S. patents and four other IEEE publications.



Stephen Kesonocky (M'90–SM'14) received the B.S., M.S., and Ph.D. degrees from Rutgers University, New Brunswick, NJ, USA.

He is currently a Senior Fellow Design Engineer with Advanced Micro Devices, Fort Collins, CO, USA, where he leads a low power advanced development team focusing on low power circuits for AMD's CPUs, GPUs, and APU products. He has authored or co-authored 70 publications and workshops, and an inventor on over 60 issued or pending U.S. Patents.

Dr. Kesonocky was the Program Chair/Co-Chair from 2005 to 2006, the General Chair/Co-Chair from 2007 to 2008, and a Technical Program Committee Member from 2001 to 2008 of the Symposium on VLSI Circuits, an Executive Committee Member from 2005 to 2015 of the VLSI Symposia, a Technical Program Committee Member of the International Solid State Circuit Conferences from 2002 to 2004 and 2010 to 2015, the International Solid State Conference Energy Efficient Digital Subcommittee Chair 2014 to 2015, and a Technical Program Committee Member of the International Symposium on Low Power Electronics and Design from 2001 to 2005. He was the IEEE Solid-State Circuit Society Membership Chair from 1998 to 2000, a member of the IEEE Electron Device Society Membership Committee from 1997 to 2005, and the Chair of the 1999 IEEE Technical Activities Board Focus Committee on retaining young members.



Russell Schreiber (M'14) received the B.S. degree in computer engineering from Penn State University, State College, PA, USA, in 2001, and the M.S. degree in electrical engineering from the University of Illinois, Urbana, IL, USA, in 2004.

In 2004, he joined AMD, Austin, TX, USA, he was involved in custom circuitry in L2 and L3 cache designs on several generations of processor cores, where he is currently a Principle Member of Technical Staff. He is currently the Lead Circuit Designer for the L2 and L3 cache hierarchy of

AMD's next-generation processor core. He is listed as an Inventor on 18 issued and pending U.S. patents.



Samuel Naffziger received the B.S.E.E. degree from CalTech, Pasadena, CA, USA, in 1988, and the M.S.E.E. degree from Stanford, Stanford, CA, USA, in 1993.

He is a Corporate Fellow with AMD responsible for power technology development, and has been the Lead Innovator behind many of AMD's low power features. He has been in the industry 29 years with a background in microprocessors and circuit design, starting at Hewlett Packard, moving to Intel and then at AMD since 2006. He has authored dozens of publications and presentations in the field. He holds 116 U.S. patents in processor circuits, architecture, and power management.



Miguel Rodriguez (S'06–M'11) was born in Gijon, Spain, in 1982. He received the M.S. and Ph.D. degrees in telecommunication engineering from the University of Oviedo, Oviedo, Spain, in 2006 and 2011, respectively.

From 2011 to 2013, he was a Post-Doctoral Research Associate with the Colorado Power Electronics Center, University of Colorado at Boulder, Boulder, CO, USA. In 2013, he joined Advanced Micro Devices, Inc., in Fort Collins, CO, USA, as a member of the Low Power Advanced Development Group. His current research interests include dc/dc conversion and digital control, analysis of power delivery networks, and linear and switched integrated voltage regulators for microprocessors and SoCs.



Amy Novak (SM'08) received the dual bachelor's degrees in biomedical and electrical engineering from Duke University, Durham, NC, USA, and the master's degree in computer engineering from the University of Texas, Austin, TX, USA, in 1996.

She is a Principal Member of Technical Staff with Advance Micro Devices, Austin, TX. She joined AMD in 1999 to focus on SRAM and register file circuit designs for x86-64 processors. Since 2012, her focus has been enablement and delivery of finFet technologies from all major foundries for high performance processor design. Prior to AMD she was with IBM, Austin, TX, for over 10 years, where she was involved in many disciplines including IO processor design and architecture, cache hierarchy, and enablement of the transition from planar to SOI technology. She holds ten issued patents in the areas of architecture, circuit design, and CAD enhancements to support processor design.

Ms. Novak was a member of the Technical Program committee for the International Conference on Computer Design for over 15 years.