Theme Article: Hot Chips

The AMD “Zen 2” Processor

David Suggs, Mahesh Subramony, and Dan Bouvier
Advanced Micro Devices Inc.

***Abstract*—The “Zen 2” processor is designed to meet the needs of diverse markets spanning server, desktop, mobile, and workstation. The core delivers significant performance and energy-efficiency improvements over “Zen” by microarchitectural changes including a new TAGE branch predictor, a double-size op cache, and a double-width floating-point unit. Building upon the core design, a modular chiplet approach provides flexibility and scalability up to 64 cores per socket with a total of 256 MB of L3 cache.**

The Zen 2 processor provides a single core design that is leveraged by multiple solutions, with focused goals for improving upon the predecessor Zen processor. The primary targets for the core were advancements in instructions per cycle (IPC), energy efficiency, and security. Building on the new core, the solutions for server and client aimed to promote design reuse across markets, increase core count, and improve IO capability. Achieving these goals required innovations in microarchitecture, process technology, chiplets, and on-package interconnect.[1]

CPU CORE

The Zen 2 CPU core, shown in Figure 1, has two primary areas of improvement over its predecessor, Zen. First, energy efficiency is doubled from a combination of technology and microarchitecture improvements. Second, IPC* is increased by approximately 15% from the microarchitectural changes in the in-order front-end, integer execute, floating-point/vector execute, load/store, and cache hierarchy.

ENERGY EFFICIENCY

The Zen 2 core microarchitecture originally targeted lower energy per clock cycle than the prior generation, independent of process technology improvements. This by itself was an aggressive goal, because increased IPC means more activity each cycle, resulting in added switching energy. Neutral energy per cycle would require additional design optimization, including improvements in branch predictor accuracy,

Digital Object Identifier 10.1109/MM.2020.2974217
Date of publication 17 February 2020; date of current version 18 March 2020.

*AMD Zen 2 CPU-based system scored an estimated 15% higher than previous generation AMD Zen based system using estimated SPECint_base2006 results. SPEC and SPECint are registered trademarks of the Standard Performance Evaluation Corporation. See www.spec.org. GD-141.

0272-1732 © 2020 IEEE

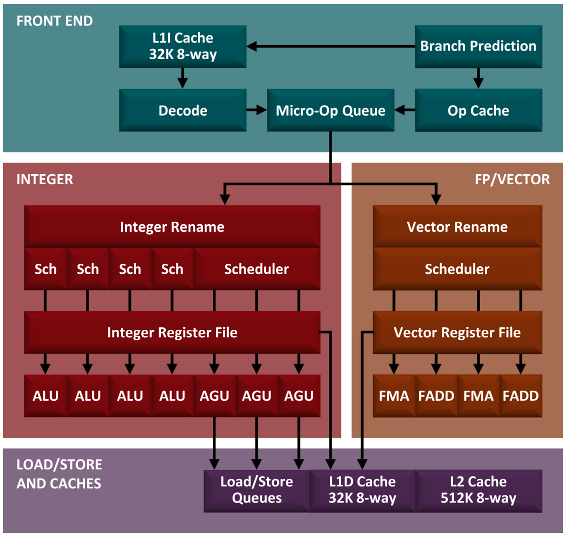

Figure 1. Core block diagram.

higher op cache hit rate, and a dedication to continuous clock and data gating improvements. Zen 2 achieved significantly better results than the original targets: an estimated 1.15× IPC increase while delivering a 1.17× improvement in energy per cycle. These two factors, together with an estimated 14–7-nm process technology improvement in energy per cycle of 1.47×, produce double the number of instructions per unit energy, as measured in silicon.[†]

SECURITY

Security is a top design consideration in modern processor design. Over the past few years, security researchers have identified an increasing number of security threats in processor microarchitecture. Some examples include Spectre v2 (indirect branch target injection) and Spectre v4 (speculative store bypass).[2] Once aware of these vulnerabilities, AMD provided mitigations for Zen, and AMD then extended the mitigations in Zen 2 with dedicated hardware implementations that improve performance.[3]

Beyond targeted mitigations for security issues, the Zen family includes conscious design choices that improve security. Specifically, the hardware checks permissions prior to consuming data, and these checks are performed prior to speculation with the data. This can prevent microarchitectural state from being updated unless permission checks pass, which allows the Zen family to be affected by fewer security vulnerabilities.[4]

PREDICTION, FETCH, AND DECODE

The in-order front-end of the Zen 2 core includes branch prediction, instruction fetch, and decode. The branch predictor in Zen 2 features a two-level conditional branch predictor. To increase prediction accuracy, the L2 predictor has been upgraded from a perceptron predictor in Zen to a tagged geometric history length (TAGE) predictor in Zen 2.[5] TAGE predictors provide high accuracy per bit of storage capacity. However, they do multiplex read data from multiple tables, requiring a timing tradeoff versus perceptron predictors. For this reason, TAGE was a good choice for the longer-latency L2 predictor while keeping perceptron as the L1 predictor for best timing at low latency.

The branch capacity in Zen 2 is nearly double that of Zen. The L0 BTB was increased from 8 to 16 entries. The L1 BTB was increased from 256 to 512 entries. The L2 BTB was increased from 4096 to 7168 entries. The indirect target array was increased from 512 to 1024 entries. The combination of improved conditional predictor and increased branch capacity allows Zen 2 to target a 30% lower mispredict rate than Zen.

The instruction cache and op cache configurations in Zen 2 are reoptimized for better performance. The op cache, containing previously decoded instructions, was doubled in capacity from 2048 to 4096 fused instructions. The L1 instruction cache was halved in size from 64 to 32 kB to make room for the larger op cache. This provides better overall performance and improved energy efficiency. The L1 instruction cache provides better utilization due to increasing associativity from four to eight ways. The op cache also covers more microarchitectural cases of instruction fusion, which increases effective throughput and op utilization throughout the core.

INTEGER EXECUTE

The Zen 2 core features a distributed execution engine, with separate schedulers, registers,

[†] Testing conducted by AMD Performance Labs as of 7/12/2019 with 2nd Generation Ryzen and 3rd Generation Ryzen engineering samples using estimated SPECint_base2006 results. PC manufacturers may vary configurations yielding different results. SPEC and SPECint are registered trademarks of the Standard Performance Evaluation Corporation. See www.spec.org

and execution units for integer and floating-point/vector operations. The integer engine operates on general-purpose registers and generates addresses for loads and stores. The floating-point/vector engine operates on vector registers. The Zen 2 integer engine focused on increasing issue width to provide more throughput and growing the out-of-order window size to expose more program parallelism.

The Zen core had the foundation for two loads and one store per cycle, but with just two address generation units (AGUs), Zen was not able to sustain this throughput in the steady state. The Zen 2 core adds a third AGU, unlocking this throughput potential and providing a more balanced processor.

A major component of window size is the scheduler queue size. Like its predecessor, Zen 2 has four fully distributed arithmetic-logic unit (ALU) scheduler queues, one per ALU. Zen 2 increases the size of each queue from 14 entries to 16 entries. The AGU scheduler remains the same size, but it is now upgraded from two separate, distributed 14-entry queues each feeding an AGU to a single 28-entry queue feeding all three AGUs. The unified AGU queue has more effective capacity due to removing the potential for queue imbalance. The unified queue is also better able to prioritize picking of the oldest ready ops, resulting in reduced mis-speculation from out-of-order loads.

Other window size components are increased, including growing the physical register file (PRF) from 168 to 180 entries and the re-order buffer (ROB) from 192 to 224 entries. These both work to allow more ops in the window, exposing more program parallelism.

Finally, the execution engine improves simultaneous multithreading fairness. It is possible for one thread with inherently low parallelism (for example, a pointer chase through main memory) to consume many of the ALU or AGU scheduler resources without benefit. The second thread may have high inherent parallelism but be unable to realize its performance potential due to insufficient scheduler resources. New fairness hardware detects this condition and slows the rate at which the low-parallelism thread can allocate into the scheduler. This gives the high-parallelism thread an opportunity to approach its potential without significant performance impact on the low-parallelism thread.

FLOATING-POINT/VECTOR EXECUTE

The Zen 2 floating-point/vector engine has doubled the data path width from 128 bits (Zen) to 256 bits. Both cores support AVX-256 instructions, but Zen double-pumps operations using its 128-bit data paths whereas Zen 2 supports native operation with its 256-bit data paths. The vector PRF width is also doubled to 256 bits. Registers can now be renamed on a 256-bit granularity instead of a 128-bit granularity. The effective capacity of the vector PRF is therefore doubled for AVX-256 code, even though the number of vector PRF entries remains the same at 160.

A significant consideration with physically doubling the data path is the potential for switching activity spikes that could cause electrical design current (EDC) specifications to be exceeded. A simplistic approach to mitigating this issue would be to immediately throttle frequency and reduce voltage when AVX-256 instructions are detected. However, this would unnecessarily penalize programs that make occasional use of AVX-256 instructions. To optimize performance, Zen 2 builds an intelligent EDC manager which monitors activity over multiple clock cycles and throttles execution only when necessary.

LOAD/STORE AND L1D/L2 CACHES

The load/store unit and level 1 data (L1D) cache provide more throughput and larger structures. An important component of overall window size, the store queue size was increased from 44 to 48 entries. The L2 data translation lookaside buffer was increased from 1536 to 2048 entries, now supporting 1-GB pages installed as splintered 2-MB pages.

The 32-kB 8-way L1D cache maximum throughput was increased in Zen 2 due to two factors. First, read and write bandwidth are doubled through an increase in width from 128 to 256 bits, matching the vector data path width. Second, the third AGU provides 50% more sustained load/store operations. Combined, these net Zen 2 three times the load+store bandwidth.

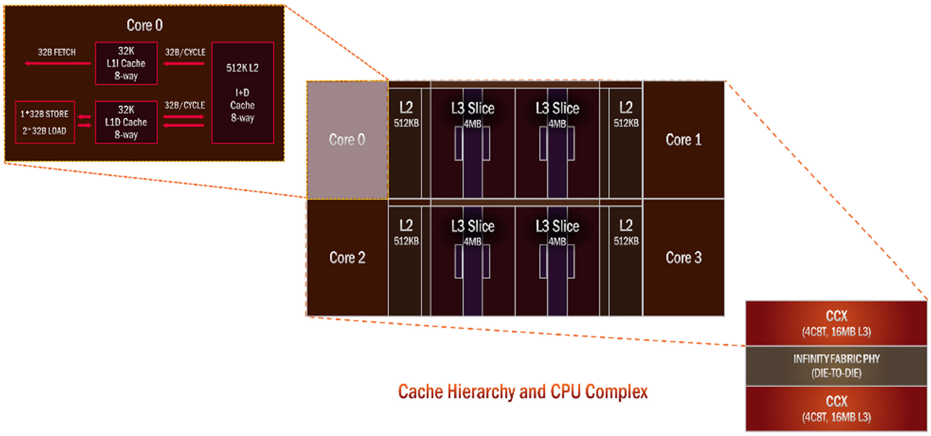

Figure 2. Cache Hierachy and CPU Complex.

The Zen 2 L2 remains 512 kB, and it is 8-way set-associative with 12-cycle load-to-use latency.

Zen 2 has new prefetch throttling capability that can reduce the aggressiveness of data prefetching when memory bandwidth utilization is high and prefetching is not being effective. This is a particularly important to performance for high core-count, constrained memory bandwidth processors such as those used in server or high-end desktop.

CORE COMPLEX AND L3 CACHE

A core complex (CCX) is composed of four Zen 2 cores and a shared level-3 (L3) cache. The L3 cache has four slices connected with a highly tuned fabric/network. Each L3 slice consists of an L3 controller, which reads and writes the L3 cache macro, and a cluster core interface that communicates with a core. The four slices of L3 are accessible by any core within the CCX. The distributed L3 cache control provides the design with improved control granularity. Each slice of L3 contains 4 MB of data for a total of 16 MB of L3 per CCX. The L3 cache is 16-way set-associative and is populated from L2 cache victims. The L3 is protected by DECTED ECC for reliability. A CPU Core Die (CCD) chiplet is composed of two CCXs, for a total of eight cores, 16 threads, and 32 MB of L3 cache as shown in Figure 2.

Chiplet Strategy: Challenges and Solutions

Delivering Zen 2 to market across multiple platforms in a short period of time was a key design challenge. While the leading edge 7-nm technology is a key element to “Zen2” performance and power efficiency, mitigating the cost was another challenge. The technology shrink factor did not apply equally across all circuits. Specifically scaling some analog circuitry, for example those used in system IOs, did not benefit enough compared to the technology cost increase. This led to the adoption of a chiplet strategy. This strategy defines SoCs using a hybrid process technology, allowing each chiplet to be manufactured in its optimal technology node. The SoCs built using the hybrid process technology married one or more of the second-generation area-optimized CCD chiplets in the advanced node and an IO-die chiplet in a mature node. This resulted in construction of cost-effective, high-performance SoCs as well as offering configurable solutions to broaden the product portfolio.

ON-PACKAGE DIE-TO-DIE INTERCONNECT

An important requirement to making the chiplet strategy viable was an optimized on-package die-to-die interconnect. The interconnect was required to support various product configurations while meeting power, bandwidth and latency metrics. A new on-package Infinity Fabric (IFOP) link was designed to meet these requirements and allow efficient communication between the core chiplet and the IO chiplet. The IFOP link was optimized for a short channel reach and responsible for carrying both data and control fabric communication. Each IO-die chiplet to CCD chiplet connection is made using an independent point-to-point instance of the IFOP link.

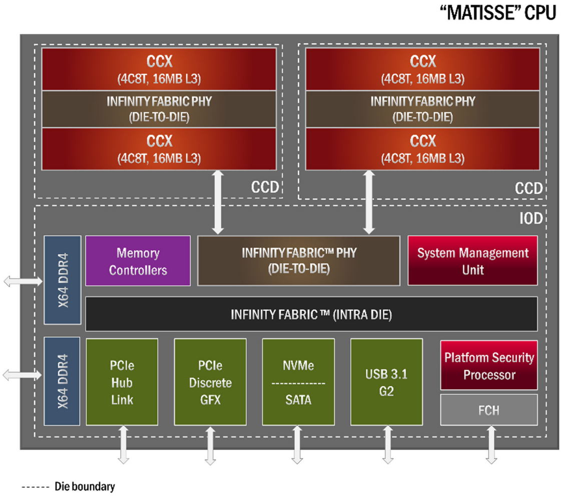

Figure 3. "Matisse" SoC.

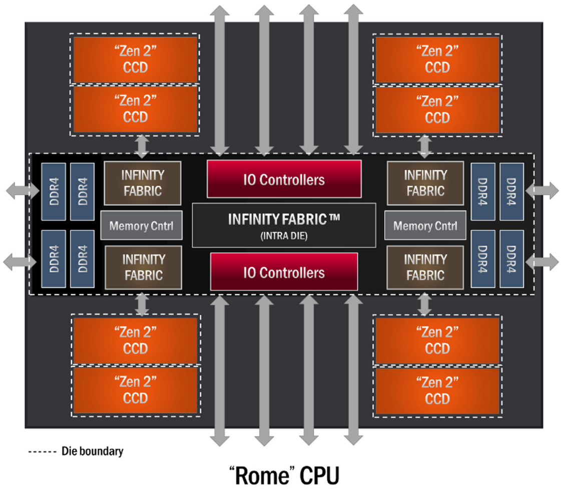

Figure 4. "Rome" SoC.

"MATISSE" SOC

The "Matisse" desktop SoC shown in Figure 3 utilizes up to two CCD chiplets with a single IO-die chiplet in an AM4 micro-PGA package. Copper pillars were used to attach the die to the package instead of traditional solder bumps to accommodate the signal connectivity requirements between the chiplets. This enabled managing mixed processes on the single package.

The IO-die features two 64-bit channels of DDR4-3200 memory for a peak bandwidth of 51.2 GB/s. The device includes 24 lanes of PCIe Gen4 lanes for a peak native IO bandwidth of 48 GB/s. The noncore clocking within the IO-die has more degrees of freedom when compared to its predecessor, allowing decoupling of the Infinity Fabric (FCLK) clock and DDR Memory Clock (MEMCLK). This enables more flexibility for power management. It also increased tuning flexibility for over-clocking the device when probing the limits of the system.

The "Matisse" SoC maintained backward compatibility to the AM4 socket/platform infrastructure. Voltage rail compatibility and IO connectivity compatibility was achieved by use of integrated low dropout regulators and firmware control.

"ROME" SOC

The "Rome" server SoC, shown in Figure 4, contains up to eight CCD chiplets paired with a single IO-die chiplet. "Rome" is packaged in an LGA package known as SP3. The SP3 package contains up to 1000 mm^2 of cumulative silicon area. The full assembly results in up to 64 cores with a total of 256 MB of L3 cache. A directory structure within the Server IO-die manages probe traffic efficiently.

All memory and IOs are hosted by a single IO-die resulting in significant generational latency improvement. Decoupled clocking within the IO-die enables efficient management IO-die thermal budget. This opens additional power budget for increased CPU frequency. With support for eight channels of DDR4-3200 memory and 64 lanes of PCIe Gen4, "Rome" delivers balanced memory and IO to support the high CPU core count.

"Rome" maintains backward socket-level compatibility with the first generation EPYC platform. "Rome" adds advanced platform security features like SEV-IO and SEV-ES. The chiplet approach allows "Rome" to deliver up to twice the socket-to-socket bandwidth, up to twice the IO-operations (IOPs) per socket, and up to twice the PCIe bandwidth. This results in near twice the application performance of the previous generation.

Reuse and modularity were important for time-to-market delivery of "Rome" and "Matisse." The CCD chiplet is used in both products. The IO-die chiplets, while unique for each SoC, use highly leveraged IP. The same IO-die used in "Matisse" was also packaged in a standalone BT1 BGA package as the X570 chipset. Paired with "Matisse" the X570 chipset opened first to market PCIe Gen4 connectivity in a premium desktop PC.

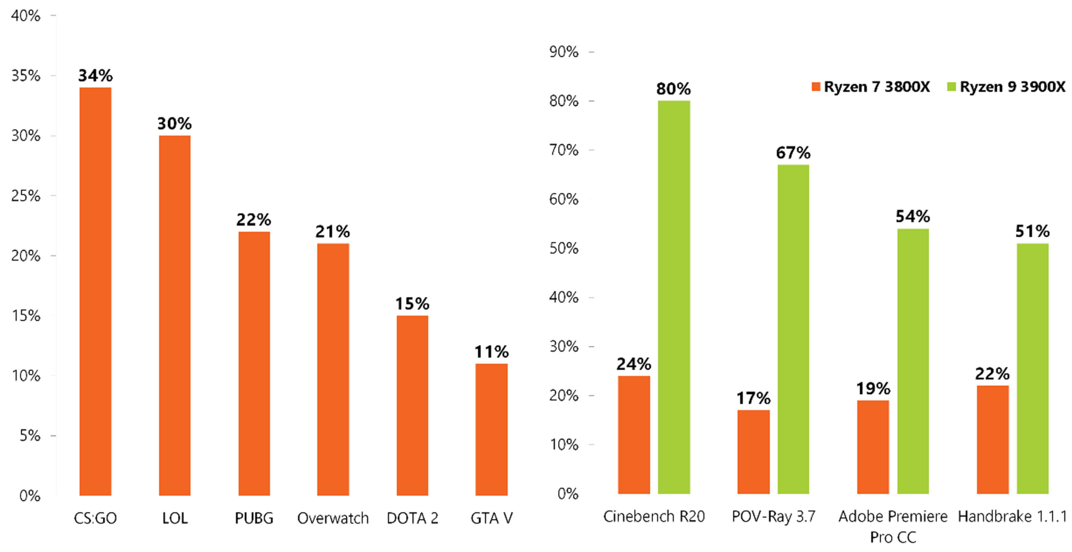

Figure 5. "Matisse" application and 1080p gaming performance compared to Ryzen 7 2700X.

PERFORMANCE

The resulting generational performance uplift of the "Matisse" SoC is very compelling, as shown in Figure 5. Refer to Table 1 for the configurations and system parameters for the performance measurements.

Cinebench R20 is a real-world cross-platform test suite to evaluate performance scalability reflecting advancements in CPU and rendering technologies. The Persistence of Vision Raytracer (POV-Ray) is a software tool utilized for rendering photorealistic 3-D images via raytracing. With higher scores being better, the third-generation Ryzen processor scores approximately 17% to 24% higher than its predecessor (Ryzen 7 2700X) using the same number of CPU cores. Further, "Matisse" scores approximately 67%–80% higher using the same thermal budget (see the last row of Table 1) as the predecessor, exploiting 50% more cores.

Table 1. Matisse performance evaluation hardware and software configuration.

CPUs	
Ryzen 9 3900X	12 cores 24 threads (Zen2)
Ryzen 7 3800X	8 cores 16 threads (Zen2)
Ryzen 7 2700X	8 cores 16 threads (Zen1)
Motherboard	AMD Reference Board
Memory	2×16 GB Dual-Rank DDR4-3200
Operating system	Windows 10 v1903
Security mitigations	Windows 10 v1903 Default
GPU	GeForce RTX 2080
Platform	AM4
Infrastructure Limits	TDC = 95A, EDC = 140A, TDP = 105W, PPT = 141.8W, Tjmax = 95C.

Adobe Premiere Pro content creation software for video editors uses a highly parallel CPU-accelerated software encoder. Handbrake is an open-source, GPL-licensed, multiplatform, multithreaded video encoder. With higher scores associated with lesser time to complete the task, the "Matisse" processor completes the task approximately 16%–18% faster when compared to its predecessor at the same core count and approximately 34%–35% faster using the same thermal budget (see the last row of Table 1).

The combination of high responsiveness, affordable prices and unit volume growth in emerging markets has made 1080p a steadfast option for gamers around the world. Using average frames per second (FPS) as a performance metric, the "Matisse" processor improves example game titles by approximately 11%–34% over its predecessor, shown in Figure 5. The higher clock-speeds, higher IPC, and larger L3 cache, combined with new synergy with the Windows scheduler provide the key ingredients for the performance uplift.

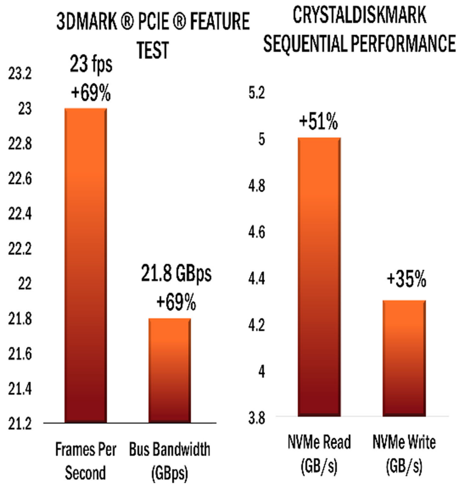

Figure 6. PCIe Gen4 vs Gen3 performance.

High-end post-production tools and certain categories of game VFX benefit highly from nonlinear edition (NLE) performance, which is limited by the IO bandwidth. The 3DMark PCIE Express feature test evaluates the performance of vertex animation across a field of wheat-like objects. CrystalDiskMark measures the sequential READ–WRITE performance of the storage system. With the introduction of native PCIe Gen4, the "Matisse" processor scores up to 69% higher in the feature test and approximately 35%–51% higher in the sequential performance tests over the same system running PCIe Gen3 (shown in Figure 6).

When compared to the first generation EPYC processors and using maximum core count, "Rome" delivers up to twice the socket-to-socket bandwidth, up to twice the IO-operations (IOPs) per socket, up to twice the PCIe bandwidth and nearly four times the peak theoretical FLOPs.

ESI Virtual Performance Solution (VPS) is used for crash simulations during the design of occupant safety systems primarily for the automotive industry. LS-DYNA is a general-purpose multiphysics, finite-element analysis program capable of simulating complex real-world problems. It is used by the automotive, aerospace, construction, military, manufacturing and bioengineering industries. Altair's PBS Professional is a fast, powerful workload manager designed to improve productivity, optimize utilization & efficiency, and simplify administration for HPC clusters, clouds and supercomputers. ANSYS Fluent is a computational fluid dynamics benchmark widely used in almost every industry sector.

Table 2. "Rome" application performance compared to 2P Intel Zeon Platinum 8280 power server.

Application	% Improvement
ESI VPS–NEON4M	Upto 58%
Altair Radioss 13.3.1	Upto 72%
LS-DYNA R9.3.0	Upto 79%
STAR-CCM+ 13.06.012	Upto 95%
ANSYS Fluent 19.1	Upto 95%

High performance computing applications benefit significantly from the scalability of "Rome" especially when used in a 2P configuration. "Rome" performed up to 58%–95% higher versus a 2P Intel Xeon Platinum 8280 power server,[‡] shown in Table 2.

CONCLUSION

Zen 2 powers the next generation of AMD desktop and server processor products. It features up to 2× instructions per unit energy compared to its predecessor. The products were designed as chiplet-based solutions enabling efficient targeting of technology while balancing power and cost. Significant leverage of the chiplet and underlying design facilitated faster deployment of a diverse product stack. As a result of the higher IPC, higher area utilization, improved security features, and higher power efficiency of Zen 2, AMD delivered high performance, efficient, and secure SoCs to market in the form of 3rd Generation Ryzen "Matisse" desktop processor and 2nd Generation EPYC "Rome" processor.

REFERENCES

1. D. Suggs, D. Bouvier, M. Subramony, and K. Lepak, "Zen 2," *Hot Chips*, vol. 31, 2019.
2. P. Kocher *et al.*, "Spectre attacks: Exploiting speculative execution," in *Proc. Symp. Secur. Privacy*, 2019, pp. 1–19.

[‡] Based on AMD internal testing of ANSYS FLUENT 19.1, lm6000_16m benchmark; LSTC LS-DYNA R9.3.0, neon benchmark; of Altair RADIOSS 2018, T10M benchmark; ESI VPS 2018.0, NEON4m benchmark; and Siemens PLM STAR-CCM+ 14.02.009, kcs_with_physics benchmark as of July 17, 2019 of a 2P EPYC 7742 powered reference server versus a 2P Intel Xeon Platinum 8280 powered server. Results may vary.

3. AMD, "Indirect branch control extension," 2019. [Online]. Available. https://developer.amd.com/wp-content/resources/Architecture_Guidelines_Update_Indirect_Branch_Control.pdf
4. AMD, "Speculation behavior in AMD microarchitectures," 2019. [Online]. Available. https://www.amd.com/system/files/documents/security-whitepaper.pdf
5. A. Seznec and P. Michaud, "A case for (partially)-tagged geometric history length predictors," *J. Instruction Level Parallelism*, vol. 8, 2006. [Online]. Available. https://www.jilp.org/howtoref.html

David Suggs is a Fellow with Advanced Micro Devices Inc. (AMD), Santa Clara, CA, USA, where he was the chief architect for the Zen 2 CPU core. Previously, he was the architect of the op cache and the instruction decoder for the Zen core. He has been working in architecture and design since 1993 on projects spanning CPU cores, north bridges, south bridges, DSPs, voice telephony, and PC sound cards. He received the M.S.E.E. degree from the University of Texas at Austin and the MBA degree from St. Edward's University. He is the corresponding author of this article. Contact him at david.suggs@amd.com.

Mahesh Subramony is a principal member of Technical Staff with Advanced Micro Devices Inc. (AMD), and was the SoC Architect for the third generation Ryzen "Matisse" desktop processors. Since 2003, he has been with AMD in architecture and design on SoCs spanning mobile, desktop, and servers. He received the M.S. degree in computer engineering from the University of Minnesota, Twin Cities, and the B.Tech. degree in electrical engineering from the College of Engineering, Thiruvananthapuram. Contact him at mahesh.subramony@amd.com.

Dan Bouvier is the Corporate VP and Client Products Chief Architect for Advanced Micro Devices Inc. (AMD), Ryzen products. He has defined the past five generations of AMD notebook and desktop processors. During his 32 year career, he has focused on high-performance processors, SoCs, and systems. Prior to joining AMD in 2009, he was a processor CTO for AMCC and before that the director of Advanced Processor Architecture for PowerPC processors at Freescale/Motorola. He received the B.S. degree in electrical engineering from Arizona State University. Contact him at dan.bouvier@amd.com.

Rejuvenating Binary Executables • Visual Privacy Protection • Communications Jamming
IEEE SECURITY & PRIVACY
Building Dependability, Reliability, and Trust
SOFTWARE EVERYWHERE
January/February 2016 Vol. 14, No. 1
Policing Privacy • Dynamic Cloud Certification • Security for High-Risk Users
IEEE Symposium on Security and Privacy
March/April 2016 Vol. 14, No. 2
Smart TVs • Code Obfuscation • The Future of Trust
ECONOMICS OF CYBERSECURITY PART 2
May/June 2016 Vol. 14, No. 3
S & P
IEEE computer society
CELEBRATING 70 YEARS
IEEE Reliability Society