AMD Ryzen Processor Family Analysis and Discussion

Muhammad Salman
Department of Electrical and Computer Engineering
Carleton University
Ottawa, ON Canada
muhammadbsalman@cmail.carleton.ca

Oussama Khallil
Department of Electrical and Computer Engineering
Carleton University
Ottawa, ON Canada
oussamakhallil@cmail.carleton.ca

ABSTRACT

In this research paper, we will analyze the AMD Ryzen Processor family, discuss the claims AMD makes for its products, research popular and in-demand AMD products, and provide our findings. We will then explain the theory science behind these products, such as changes in micro architecture, CPU performance times, and core/cache distribution etc. We will then discuss the properties of the infamous Zen architecture which AMD designs in-house for its processors. We will dive deeper into the CPU core and outline the phases instructions go through starting from front-end, to execution engine and Load/Store. AMD planned to improve on three areas when designing Zen 3: Single-threaded performance, Latency and Power efficiency. In the paper, we will also outline how these areas were improved in the CPU Core redesign and analyze how these improvements are projected on the performance of the Ryzen 5000 Series.

KEYWORDS

AMD, Ryzen, CPU, GPU, APU, Threads, Cores, Caches, Processing Times, IPC, Zen, Front-End, Load/Store, Branch Predictor, ALU, AGU, SSE, SIMD, TDP, PPT.

ACM Reference format:

Muhammad Salman, and Oussama Khallil. 2021. AMD Ryzen Processor Family Analysis and Discussion. Final report for SYSC5704 offered at Carleton University, Ottawa, Canada.

1 Introduction

AMD is one of the largest players in the processor and semiconductor space. As such, it has produced numerous high end processors in many spaces, including: Central Processing Units (CPUs), GPUs (Graphics Processing Units), and APUs (Accelerated Processing Units). These products have gained fame in the semiconductor and processor space, and are a go to product for many, including professionals, gamers, and amateurs alike. The focus of this paper will be the Ryzen family of processors, with the most recent release being the Ryzen 9 5900X/5950X CPUs.

There are many ways to assess the performance of a CPU, including: the operating frequency or clock, the IPC (Instructions Per Cycle), the cache/core performance and many more aspects.

AMD claims that the Ryzen 9 5900X is the ‘King of Hill,’ when it comes to the gaming industry, also adding that they are the ‘best in the world for PC gaming[1]. Though these claims are quite boastful, AMD explains the theory and design behind their processors, which will be analyzed in depth in a later section of this paper.

2 Model Overview

2.1 Cache Summary

Before diving into the current models, it is important to define the levels of cache on a CPU. From a very high level perspective, these caches can be seen interconnected in Figure 1 below.

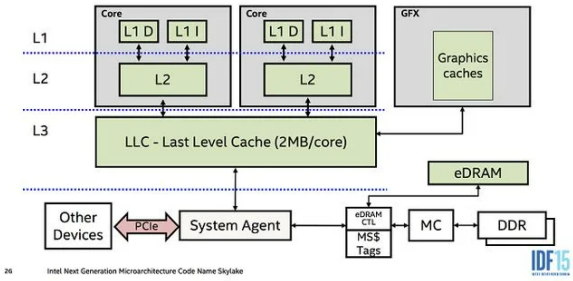

Figure 1: Cache Distributions on Standard CPU [1]

The L1 cache is also known as the fastest cache and can be found on the actual processor core itself. The data and instruction set are actually separated into two separate caches instead of one, hence also being the smallest cache in size. The L2 cache on the other hand is larger than L1, but as a result it is also slower. Both of these cache levels pertain to singular cores on the CPU chip, meaning they present a faster connection to the individual core itself, improving performance. Finally, the L3 cache serves as a general memory pool for the entire CPU package, and as such, is the slowest level of all the caches. Due to recent advancements in technology, the cache has moved from the previously being on the motherboard, to the actual CPU package itself [1].

2.2 Ryzen Applications

Although the Ryzen processor family is well-known for gaming, it also presents solutions for desktop, laptop, embedded applications. However, since the majority of gaming is done on desktop environments, let's take a deep dive into the AMD Ryzen 9 5900X Desktop Processor. A quick overview of some important specifications can be seen in the figure below.

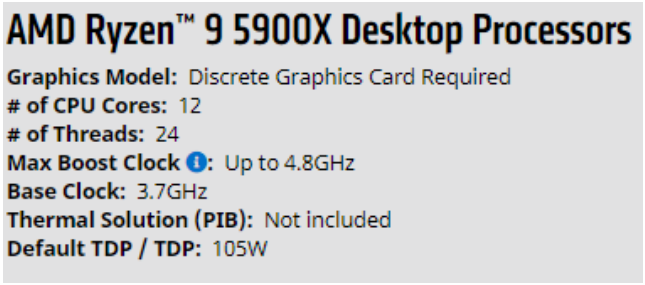

Figure 2: AMD Ryzen 9 5900X Specifications [2]

The number of cores is 12, while the number of virtual cores, or threads, is 24. A core is a physical component of hardware on the CPU package, just like the cache. A thread can be understood as an execution of a task in virtuality. For example, if two tasks, or threads, need to be executed, they can be done simultaneously on one core. This is also known as multi-threading. Multi-threading allows the cores to be used to their maximum potential efficiently, as well as speeding up various processes on many applications [3].

Another application present in these processors is the ability to enhance the operating frequency of the CPU. This technology is known as *Precision Boost 2* (PB2). When the CPU requires additional processing power or units, or has a larger amount of instructions to complete, the PB2 technology allows the Ryzen processor to temporarily raise its clock speeds, and subsequently decrease the time spent per cycle. Hence, a larger set of instructions can be executed in a shorter period of time. This is all done automatically through the CPU's intelligence, without requiring any type of input or command from the user [4].

3 Zen Microarchitecture

The Zen microarchitecture is developed by AMD in-house, and is the proprietary design used across their CPUs. The first generation of Zen was launched in 2017, while the most recent edition, Zen 3 was launched at the end of 2020 [5].

3.1 Zen 3 Overview

The most up to date processors released by AMD, such as AMD Ryzen 5000 series, use the Zen 3 architecture. The Zen 3 has a strong objective to improve gaming performance, hence, it seeks to improve not only CPU performance as a whole, but individual core performance as well [6].

Compared to Zen 2, a wide array of changes have been implemented, and as such, Zen 3 has been re-designed in a way. As a result, this has led to a higher IPC execution rate (upto 19% higher compared to Zen 2), as well as lower latency between inter-core communication, and core-cache communication [6].

3.2 Zen 3 Microarchitecture

The greatest difference between the Zen 2 and Zen 3 architectures as far as a hardware redesign is concerned, is the core-cache distribution. For example, Zen 2 utilized two 16MB L3 caches, each directly connected with four cores. Each individual cache only has access to 16MB of cache, and latency between the two sets of cores is greatly increased. This design is quite beneficial in terms of scalability, since adding another set of 16MB cache/4 cores in parallel would be a straightforward approach to increasing performance (such as AMD increasing to 64 cores in the past) [6]. Refer to the image below for a representation of Zen 2 microarchitecture.

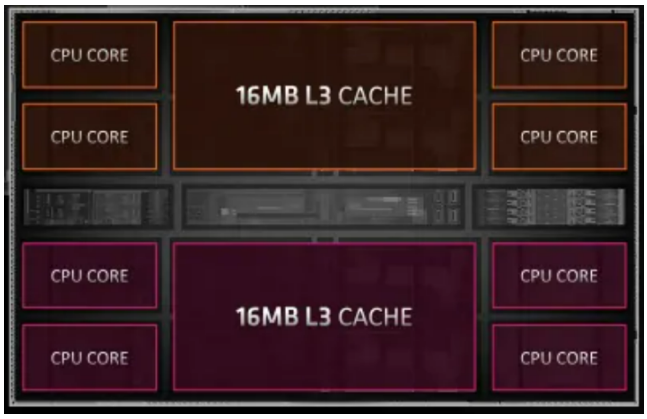

Figure 3: Zen 2 Microarchitecture High-Level [6]

However, as mentioned before, in order to increase gaming performance, performance per core also has to be improved. This served as the main inspiration behind the redesign of the Zen 2. As such, implementation of the Zen 3 does not change the *total* amount of cores or cache on the CPU package itself, rather, it simply rearranges it. The Zen 3 has one 32 MB L3 cache, with all eight cores having access to the same cache. By not splitting the cache-core distribution into two sections, the cache access present to each core subsequently doubled from 16MB to 32MB [6]. Refer to the image below for a representation of Zen 3 microarchitecture.

Figure 4: Zen 3 Microarchitecture High-Level [6]

Along with increasing gaming performance, the Zen 3 also decreases latency between core-core and core-cache communication, leading to a faster overall CPU package and individual core.

3.3 Zen 3 CPU Core: Front-End

In addition to the core-cache distribution, the large increase in performance of the Zen 3 microarchitecture is due to the redesign and increase of bandwidth in all three sections of the core: the front end, the execution engine, and the memory subsystem. AMD introduced various components (such as their new branch prediction flow and forwarding mechanisms) with the Zen architecture that eliminate the need for operations to go through the high power ALUs and decoders, increasing the overall power efficiency and throughput [7].

Each of the CPU cores consists of a front end that fetches instructions, decodes them, generates micro-operations (e.g. single arithmetics and memory operations), and fuses the micro-operations to dispatch them to the Execution Engine. Instructions

are either fetched from the L1I Cache or come from the micro-operations cache (on subsequent fetches) eliminating the decoding stage altogether [7]. The Zen 3 also includes an enhanced branch prediction unit that breaks instructions down into predicates, which improves the execution of instructions [7]. Branch prediction is a technique used to speed the execution of instructions on processors that use pipelining [8].

CPUs initially executed instructions one by one as they came in, but the introduction of pipelining meant that branching instructions could slow the processor down significantly as the processor waits for the conditional jump to be executed. The Branch Predictor breaks instructions down into predicates, then the CPU only executes statements if a predicate is true [8]. This helps the CPU to work much more efficiently reducing the backlog in the pipeline.

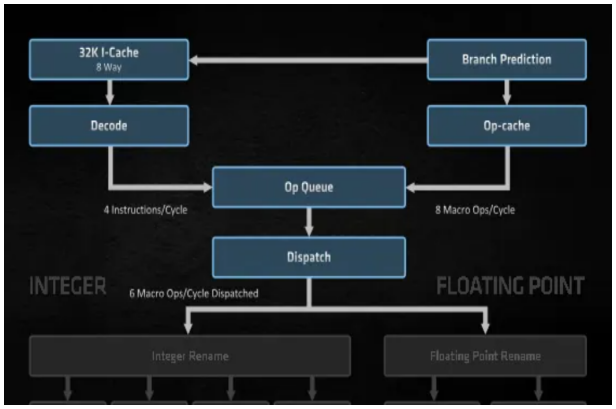

Figure 5: Zen 3 Front-End [6]

Between Zen 2 and Zen 3, the fundamental building blocks are the same. However, there is a list of enhancements in the Zen 3. The primary enhancements in the Zen 3's front end are listed below:

1. An improved branch predictor, with faster sequencing and switching of the Op-cache fetches and pipes, respectively [9].
2. Then, there's the no-bubble prediction. This simply means that the core has a much faster pipeline flush in case of a branch misprediction which usually stalls the pipeline, otherwise known as a bubble [9].
3. The L1 Branch Target Buffer (BTB) has also been doubled from 512 to 1,024 entries, further improving branch prediction [9].

In general, the Zen 3 front-end hasn't changed much, the branch predictor has seen some optimizations to improve prediction accuracy as listed above. However, the branch prediction unit is still a TAGE-based design as had been introduced in Zen 2. TAGE predictors provide high accuracy per bit of storage capacity[10].

3.4 Zen 3 CPU Core: Execution Engine and Load/Store

The execution engine in the Zen microarchitecture follows traditional core partitioning - every core is an independent core with its own floating-point/SIMD units and an L2 cache. The two sections are completely separate, each featuring separate registers, schedulers, queues, and execution units. The integer engine operates on general-purpose while the floating-point engine operates on vector registers. The Integer side splits up the micro-operations via a set of individual schedulers that feed the various ALU units. On the floating point side, there are different schedulers to handle the 128-bit and 256-bit floating-point operations.

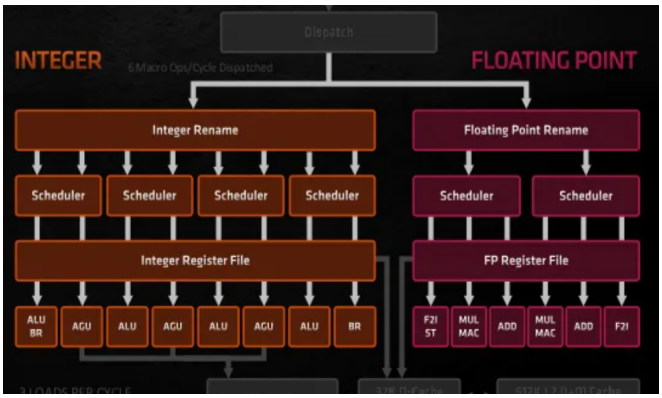

Figure 6: Zen 3 Execution Engine [6]

Changes on the Execution unit and the Load/Store of Zen 3 are bigger than the front-end enhancements, being the biggest factor in the +19% IPC gain. On the Integer side, even though the number of schedulers has reduced to 4, these new 24-entry schedulers are

more power-efficient. Also, the scheduler's entry capacity has gone up slightly from 92 to 96 [6]. The physical register can now hold 192 entries while the reorder buffer has also been increased from 224 to 256 [6].

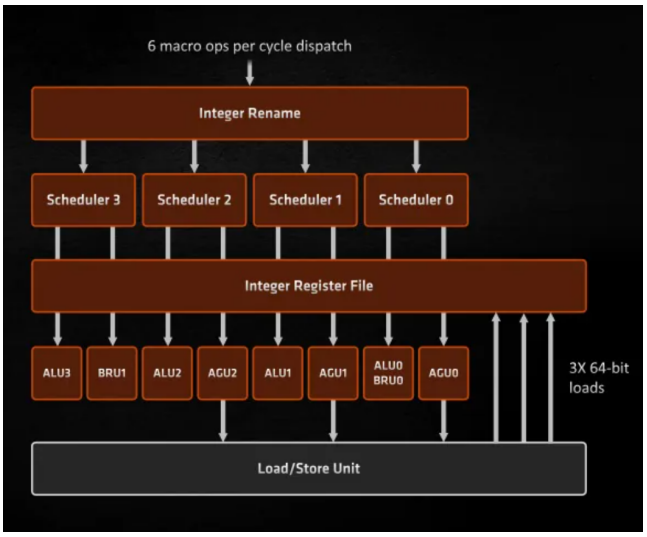

Figure 7: Zen 3 Integer Execution [6]

The total number of ALU execution units and AGU has stayed the same, however, there are 2 shared stores units, 1 shared branch unit, and 1 dedicated branch added to increase bandwidth. Another difference to the integer execution is the inclusion of wider, but fewer schedulers. According to AMD, this improves the ALU/AGU utilization by balancing the load across the various ports [9].

Similarly, the schedulers in the Floating Point execution have gotten wider, also the dispatch bandwidth is higher and the time to execute FMAC instructions has also been reduced from 5 cycles to 4, which is a major latency improvement.
Following the same trend in the previous front-end and the execution units, AMD increased the bandwidth in the memory subsystem. Data is fed into the L1D cache via the load and store queue (both of which were almost doubled in capacity) via the Address Generation Units (AGUs) at the rate of 3 loads and 2 stores per cycle, in comparison to 2 loads and 1 store per cycle in the case of Zen 2. However, the L1 to L2 cache transfer speeds are unchanged at 32 bytes/cycle x 2. The L1 fetch width is also the same as Zen 2 at 32 bytes [9].

3.6 Instruction Sets

The Zen microarchitecture implements the AMD64 (x86-64) instruction set, which is a 64-bit version of the x86 instruction set. The instruction set adds 64-bit addressing and expands register resources to support higher performance for recompiled 64-bit programs while being compatible with legacy 16-bit and 32-bit programs and operating systems.
The AMD64 application programming instructions are organized into four subsets [11]:

1. General-Purpose Instructions—These are the basic x86 integer instructions used in virtually all programs. Most of these instructions load, store, or operate on data located in the general-purpose registers (GPRs) or memory.
2. Streaming SIMD Extensions Instructions (SSE)—These instructions load, store, or operate on data located primarily in the YMM/XMM registers. SSE instructions perform integer and floating-point operations on vector (packed) and scalar data types.
3. Multimedia Extension Instructions—These include the MMX™ technology and AMD 3DNow!™ technology instructions.
4. x87 Floating-Point Instructions—These are the floating-point instructions used in legacy x87 applications. They load, store, or operate on data located in the 80-bit x87 registers.

In this section, we will outline the major new and improved instructions supported in the Zen 3 microarchitecture. The FMAC (Fused Multiply-Accumulate), which is a part of the extended SSE instruction set, is improved to reduce latency. In Zen 3, a single FMA takes 4 cycles with a throughput of 2/clock, compared to 5 cycles in Zen 2 [12]. This enhancement is very impactful in the use-cases of high-performance computation and machine learning. The Zen 3 microarchitecture adds AVX2 (256-bit execution) to support Vector-based AES and PCLMULQDQ operations [12]. In Zen 2, VAES and VPCLMULQDQ were limited to AVX \ 128-bit execution [12]. VAES has a latency of 4 cycles with a

throughput of 2/clock, while VPCLMULQDQ has a latency of 4 cycles, with a throughput of 0.5/clock [12]. The addition of AVX2 instruction support helps in increasing performance in the applications of cryptography and encryption/decryption [12].

4 Power

All the enhancements in the core mentioned above, such as wider buffers and increased bandwidth, result in the +19% gain in IPC but also come with the need for more power. When a core gets wider, supporting more IPC means more silicon has to be turned on all the time, and this influences leakage (static) power. Or in the case the processor gets used simultaneously, with more activity each cycle, then there is higher switching (dynamic) power.

In addition to the TSMC's 7nm process that AMD leverages, additional design optimization, including improvements in branch predictor accuracy, dedicated branch, and more power-efficient schedulers are important factors that have made Zen 3 remain power neutral. As claimed by AMD, the 7nm Zen 3 processors are 2.8 times more efficient than the competing Intel Core i9 10900K processor [6].

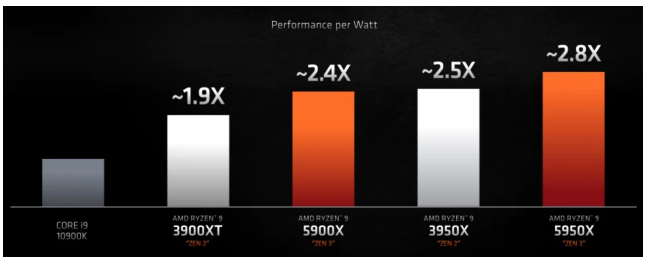

Figure 8: Zen 3 Power Efficiency Gain[6]

To analyze the power efficiency further, let's have a quick analysis of the Thermal Design Power (TDP) in the AMD Ryzen 5000 Series. TDP is measured in watts and refers to the power consumption under the maximum theoretical load [13]. Power consumption is less than TDP under lower loads. The TDP is the maximum power that one should be designing the system for in order to help with proper thermal solution selection [13]

As reported in an article by Anandtech, a computer hardware magazine, throughout their test most of the AMD Ryzen 5000 processors have a 105 W TDP, with a Package Power Tracking (PPT) setting of 142 W [14]. PPT is the power threshold that is allowed to be delivered to the socket. The only exception was the Ryzen 5 5600 since it is a 65 W processor, hence the PPT value is 88 W, and they reported only 76 W TDP, showing some of the efficiencies on the Ryzen 5 5600X [14].

As a result of the tests, the graph below (Figure 9) was generated by measuring the power consumption in the Ryzen 9 5950X. As can be seen in the graph two values are reported: total package power (for the whole chip), and the power solely used by the sum of cores, which includes the L3 cache [14]. The difference between the two includes the IO die as well as any chiplet-to-chiplet communications, PCIe, CPU-to-chipset, and DRAM controller consumption [14].

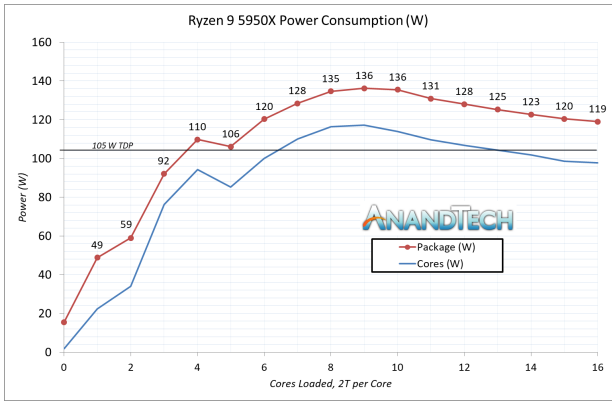

Figure 9: Ryzen 9 5950X Power Consumption [14]

From the graph, we observe a slow decrease in total package power consumption after 8-10 core loading. This is an indicator of how the processor has increased current density as it loads up the cores, and as a result, there's a balance between the frequency it can give, delivering the power, and applying the voltage in a consistent way [14]. We can also observe the difference between the two values of Package power and sum of Cores power increasing slightly, as more data is transferred over those off-chiplet

communications. This effect was also reported by Anandtech to be seen in the 5900X processor as well, perhaps indicating this is a feature of the dual chiplet design [14].

In comparison to other processors, the chart below (Figure 10) outlines the highest loaded peak power as reported by Anandtech's tests. In general, the AMD processors including the newer 5000 series and the older 3000 and 2000 series seem to be scoring lower peak power in comparison to the Intel chips. If we take the highest-end AMD Ryzen 5950X and compare it to the highest-end Intel Core i9 10900K, we see that there is an advantage of 1.77X on the AMD processor. Hence, it can be concluded that AMD leads the power efficiency by leveraging the TSMC 7nm process, the Zen 3 additional design optimization, and the chiplet design.

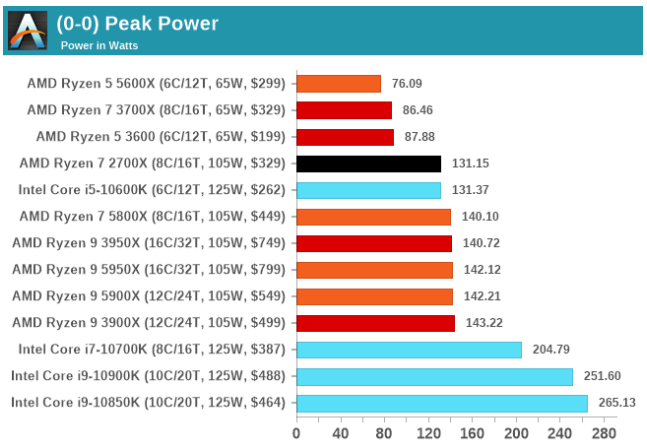

Figure 10: Peak Power Comparison [14]

5 Performance

The biggest performance milestone achieved in the Ryzen 5000 processors with the Zen 3 architecture is the +19% gain in Instructions per Cycle (IPC), as claimed by AMD in their launch event. This is a big uplift from Zen 2 to Zen 3, the acclaimed 19% increase is based on tests performed to compare between Ryzen 5 3800XT and Ryzen 5 5800X when both CPUs are at 4.0 GHz and using DDR4-3600 memory [15]. By comparison, Anandtech tested with industry-standard benchmarks at rated clock speeds and JEDEC supported memory, and were able to accurately achieve that +19% number [15].

With the new Zen 3 architecture and especially the core-cache distribution that doubles the amount of cache for each core, the latency has dropped significantly, boosting single-core performance higher than ever seen in the industry. For instance, in Techradar's (technology magazine) testing, the AMD Ryzen 9 5900X is around 20% faster than the Ryzen 9 3900X and the Intel Core i9 10900K in single-core benchmarks using the Cinebench test suite [16]. This obviously also results in faster gaming performance too, as the best PC games are still heavily dependent on clock speed and single-core performance [16].

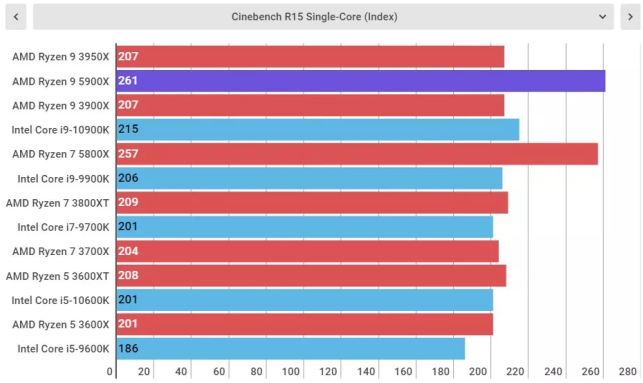

Figure 11: Cinebench Single-Core Comparison [16]

For gaming performance, Anandtech was able to run tests at 1920x1080 with all the quality settings on maximum, and the performance gain averaged around +10% [16]. The games tested were: World of Tanks, Strange Brigade, Red Dead Redemption, Civilization 6, and Far Cry 5 [16]. Also, when comparing the AMD Ryzen 9 5950X against the Intel Core i9-10900K, the average FPS advantage for AMD was +21% at CPU-limited scenarios, ranging from +2% to +52% [15].

6 Other Zen Architecture Products

The Zen architecture has become integrated in multiple products and markets that bring in a lot of value in performance and scalability. In April, 2021 AMD announced its new series of Ryzen 5000G (G

here stands for graphics) [17]. The APU (Accelerated Processing Unit) consists of the Ryzen CPU cores plus an integrated Radeon GPU. Available in August, the new APUs will feature up to eight cores and 16 threads with the Ryzen 7 5700G sticking to a 65W TDP (the same as the Ryzen 5600X) [17].

In contrast to their CPU counterparts, the APUs consist of single dies with integrated graphics and smaller caches. capable enough for mid range gaming.

In AMD's product line, the Ryzen Mobile 5000 series brings AMD's new Zen 3 architecture to the mobile gaming and ultraportable segments.

The Mobile 5000 series includes the H-series — which you'll find in laptops intended for gaming and content creation; and the U-series — meant for ultraportable notebooks [18].

On the H-series side, the big player is the Ryzen 9 5980HS, also with eight cores and 16 threads, but with 3.0Ghz clock speeds boosting up to 4.8GHz. AMD says these are "the fastest mobile processors you can get"[18].

Another big market that the Zen architecture intrudes is the console gaming market. The PlayStation 5 is powered by a custom system on a chip (SoC) designed in tandem by AMD and Sony, integrating a custom 7 nm AMD Zen 2 CPU with eight cores running at a variable frequency capped at 3.5 GHz [19].

The Xbox Series X is also powered by a custom 7 nm Zen 2 CPU with eight cores running at a nominal 3.8 GHz [19].

7 Conclusion

In conclusion, this paper summarizes the key components and specifications of the AMD Ryzen Processor, such as the instruction set, hardware cache and core design, and overall SoC analysis. As per the research presented in this paper, the latest Ryzen CPU based on the Zen 3 microarchitecture, Ryzen 9 5950X, can be seen as a remarkable CPU for gaming purposes overall.

ACKNOWLEDGMENTS

We would like to thank Professor Mustafa for organizing and teaching this course, as well as allowing us to participate in this research paper for the AMD Ryzen Processor.

REFERENCES

[1] MUO. 2021. *How Does CPU Cache Work? What Are L1, L2, and L3 Cache?*. [online] Available at: <https://www.makeuseof.com/tag/what-is-cpu-cache/> [Accessed 30 July 2021].

[2] AMD. 2021. *AMD Ryzen*. [online] Available at: <https://www.amd.com/en/processors/ryzen> [Accessed 30 July 2021].

[3] Guru99.com. 2021. *CPU Core, Multi-Core, Thread, Core vs Threads, Hyper-Threading*. [online] Available at: <https://www.guru99.com/cpu-core-multicore-thread.html> [Accessed 8 August 2021].

[4] AMD. 2021. *AMD*. [online] Available at: <https://www.amd.com/en/support/kb/faq/cpu-pb2> [Accessed 8 August 2021].

[5] Mujtaba, H., 2021. *AMD Next-Gen Zen 3 Based Ryzen 4000 Desktop And EPYC Milan Server CPUs Arriving Late 2020*. [online] Wccftech. Available at: <https://wccftech.com/amd-zen-3-ryzen-4000-desktop-and-epyc-milan-server-cpus-late-2020-launch/> [Accessed 8 August 2021].

[6] Zen 3 AMD. 2021. *Zen 3*. [online] Available at: <https://www.amd.com/en/technologies/zen-core-3> [Accessed 8 August 2021].

[7]Wikichip. 2021 *Zen - Microarchitectures - AMD.* [online] Available at: <https://en.wikichip.org/wiki/amd/microarchitectures/zen> [Accessed 8 August 2021].

[8] Techopedia. 2021 *Branch Prediction.* [online] Available at: <https://www.techopedia.com/definition/18062/branch-prediction> [Accessed 8 August 2021].

[9] Areej. 2020. *AMD Ryzen 5000 "Zen 3" Architectural Deep Dive.* [online] Available at: <https://www.hardwaretimes.com/amd-ryzen-5000-zen-3-architectural-deep-dive/> [Accessed 8 August 2021]

[10] David Suggs, Mahesh Subramony, and Dan Bouvier. 2019. *The AMD "Zen 2"*

Processor. [online] Available at: <https://ieeexplore-ieee-org.proxy.library.carleton.ca/stamp/stamp.jsp?tp=&arnumber=9000513> [Accessed 8 August 2021]

[11] AMD, AMD64 Architecture Programmer's Manual, Volume 1: Application Programming, p.4.

[12] Dr. Ian Cutress & Andrei Frumusanu. 2020. *New and Improved Instructions - AMD Zen 3 Ryzen Deep Dive Review: 5950X, 5900X, 5800X and 5600X Tested*. [online] Available at: <https://www.anandtech.com/show/16214/amd-zen-3-ryzen-deep-dive-review-5950x-5900x-5800x-and-5700x-tested/6> [Accessed 8 August 2021]

[13] Intel. 2019. *Thermal Design Power (TDP) in Intel Processors.* [online] Available at: <https://www.intel.ca/content/www/ca/en/support/articles/000055611/processors.html>

[14] Dr. Ian Cutress & Andrei Frumusanu. 2020. *New and Improved Instructions - AMD Zen 3 Ryzen Deep Dive Review: 5950X, 5900X, 5800X and 5600X Tested*. [online] Available at: <https://www.anandtech.com/show/16214/amd-zen-3-ryzen-deep-dive-review-5950x-5900x-5800x-and-5700x-tested/8> [Accessed 8 August 2021]

[15] Dr. Ian Cutress & Andrei Frumusanu. 2020. *New and Improved Instructions - AMD Zen 3 Ryzen Deep Dive Review: 5950X, 5900X, 5800X and 5600X Tested*. [online] Available at: <https://www.anandtech.com/show/16214/amd-zen-3-ryzen-deep-dive-review-5950x-5900x-5800x-and-5700x-tested/31> [Accessed 8 August 2021]

[16] Bill Thomas. 2020. *AMD Ryzen 5000 release date, price, benchmarks and specs.* [online] Available at: <https://www.techradar.com/news/amd-ryzen-5000> [Accessed 8 August 2021]

[17] Antony Leather. 2021. *AMD Announces Ryzen 5000 APUs: Zen 3, Up To Eight Cores With Radeon Graphics to Boost Budget Gaming.* [online] Available at: <https://www.forbes.com/sites/antonyleather/2021/06/01/amd-announces-ryzen-5000-apus-zen-3-up-to-eight-cores-with-radeon-graphics-to-boost-budget-gaming/?sh=77ceb6ad5b8a> [Accessed 8 August 2021]

[18] Joel Hruska. 2021. *AMD Launches New Ryzen 5000 Mobile Processors - ExtremeTech.* [online] Available at: <https://www.extremetech.com/computing/319525-amd-launches-new-ryzen-5000-mobile-processors> [Accessed 8 August 2021]

[19] Arjun Krishna. 2020. *Tech Analysis: How Are Xbox Series X And PS5 SoCs Manufactured*?. [online] Available at: <https://gamingbolt.com/tech-analysis-how-are-xbox-series-x-and-ps5-socs-manufactured> [Accessed 8 August 2021]