

Simultaneous Multithreading: Driving Performance and Efficiency on AMD EPYC CPUs

Mar 03, 2025



No Compromise: Driving Performance and Efficiency with AMD EPYC and SMT

One of the main pillars that vendors of Arm®-based processors often cite as a competitive advantage versus x86 processors is a keen focus on energy efficiency and predictability of performance. In the quest for higher efficiency and performance, Arm vendors have largely designed out the ability to operate on multiple threads concurrently—something that most enterprise-class CPUs have enabled for years under the technology description of “SMT”—which was also created in the name of enabling performance and efficiency benefits.

Arm ecosystem often claim that SMT brings security risks, creates performance unpredictability from shared resource contention and drives added cost and energy needed to implement SMT. Interestingly, Arm does support multi-threading in its Neoverse E1-class processor family for embedded uses such as automotive. Given these incongruities, this whitepaper intends to provide a bit more clarity to help customers assess what attributes of performance and efficiency really bring them value for their critical workloads.

What is SMT:

Simultaneous Multithreading (SMT) is a technology that allows a CPU core to execute multiple threads simultaneously. Since its inception, SMT has been implemented in many modern processors with varying numbers of threads. The most common approach is 2-way SMT, where two threads execute simultaneously per CPU core—versus each thread running to completion serially—as shown in Figure 1. This blog focuses on 2-way SMT, as implemented in AMD “Zen” processor cores.

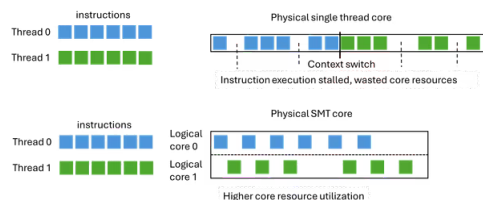


Figure 1: Single threaded processing flow compared to SMT process flow.

Benefits of using SMT:

SMT is a popular CPU feature because it offers several performance and efficiency benefits:

- **Improved core resource utilization:** SMT keeps cores busy by dynamically interleaving instructions from two threads across shared execution resources. Ideally, a CPU core would constantly execute instructions without breaks, but in reality, core stalls often occur, for example when waiting for data from memory after a cache miss or during a branch misprediction. SMT helps fill these gaps by allowing a second thread to use shared core resources while the other thread is stalled or otherwise awaiting data.
- **Increased throughput:** Simultaneous execution of two threads allows more instructions to pass through core pipelines in parallel, leading to increased Instructions Per Cycle (IPC) and better overall performance.
- **Energy efficiency:** SMT can improve performance without significantly increasing overall processor power consumption. For many workloads this translates to significant energy efficiency gains.
- **Ability to increase performance or capacity without incurring additional physical core-based license costs.**
- **Software support:** SMT has been around for more than 20 years, during this time the software ecosystem has embraced it and supported it from games to enterprise software, and the cloud. All modern Operating Systems are built to support SMT and do so effectively, by distributing threads optimally based on each CPU's core organization and NUMA domains. Software developers can choose to optimize their software and extract additional performance and energy efficiency from SMT, but there is no additional development effort necessary to support it. SMT works out of the box, and it is transparent to high level software.
- **Flexibility:** SMT can be enabled/disabled through the system BIOS for more permanent changes or during runtime in Linux®, allowing the administrator to choose the setting that best meets workload needs.

SMT design challenges:

While SMT can add significant performance to a core, it also presents hardware design challenges for silicon and system vendors to address:

- **Increased attack surface:** Virtually any feature of any component of any system must be considered an attack surface, and CPU and system vendors invest significant resources throughout the product lifecycle to understand potential vulnerabilities. Features that interoperate with highly privileged system resources get the highest levels of scrutiny and testing, and SMT is one such feature as it enables core resource sharing between the two threads, making it a tempting possible target for exploits such as side-channel attacks. Over the course of SMT's 20-year existence, CPU and systems vendors have identified and mitigated such threats through firmware updates and amendments to the core designs to address them in subsequent generations. In addition, AMD continuously works with the software community to identify and address any new potential security vulnerabilities across the processor feature set.

Share:



Article By

Apostolos Kotsiolis



- **Fair sharing of core resources for both threads:** Another challenge is providing good performance to both threads, while enabling a fair share of the core resources. CPU architects must decide which resources will be shared and how to efficiently schedule instructions from both threads while sharing the core's resources. The original "Zen" was designed from the ground up as an SMT-ready core, and subsequent generations build on the same principles:
 1. Running thread gets all resources when the other thread is sleeping.
 2. Each thread can fully utilize pipeline resources when the other thread is stalled.
 3. When SMT is enabled most of the core's resources are competitively shared between the two threads.

As shown in Figure 2, controls must be in place to meet these principles. To prevent starvation and ensure fairness, in-order queues are statically partitioned while parts of the out-of-order queues and branch prediction are watermarked for each thread, and the rest is competitively shared.

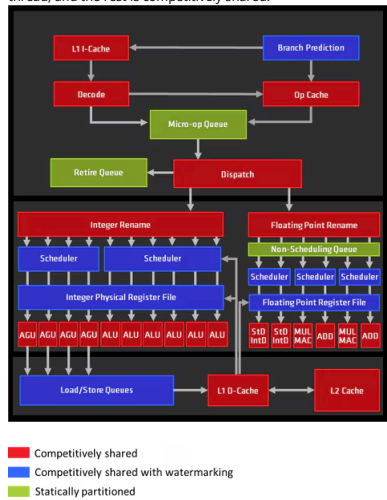


Figure 2: SMT-enabled resource sharing on the "Zen 5" core

How "expensive" is it to implement SMT?

From an end customer perspective, there is no material "cost" for utilizing SMT—it is a built-in function that most x86 customers can turn on or off. But in the very practical terms of semiconductor economics, anything that consumes transistor area on chip silicon or consumer energy when running represents a cost. And in terms of the cost to implement SMT, that cost is small, and typically offset by the gains it enables. For example, AMD engineers estimate that implementing Simultaneous Multithreading (SMT) takes less than 5% of the core area in the latest AMD "Zen 4" and "Zen 5" cores. This includes all the necessary logic to allow two threads to share the core's resources. In easy "manager math", SMT enables up to 384 threads while consuming less silicon area than 10 physical cores—that is strong ROI. Additionally, in cases where software is licensed based on the number of physical cores in the system, having the extra performance and capacity enabled by the availability of virtual cores/threads can enable *significant cost savings!* Now to dispel that pesky energy consumption myth.

SMT Enables Performance and Efficiency

AMD EPYC processors have established hundreds of performance and efficiency world records. These include workloads that benefit significantly from multithreading and SMT and a number that do not, such as a number of HPC and technical computing apps. Let's suppose we want a separate broad-based assessment of where SMT brings value and how AMD delivers the goods. Independent testing house Phoronix has done perhaps the most complete and consistent analysis of the value of SMT. The Phoronix test results³ for the "Zen 5" based AMD EPYC 9755 CPU showed big performance gains on a broad set of tested workloads, including databases, cryptography, and compression workloads—as shown in chart 1.

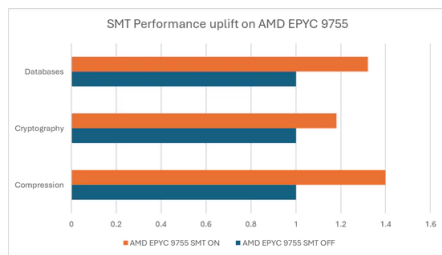


Chart 1: SMT Performance gains on AMD EPYC 9755 Systems.

These results are not surprising given that an earlier Phoronix analysis of SMT using prior generation AMD EPYC 9754 platforms identified similar performance and power efficiency gains.³ For those interested in workloads outside of the domains summarized in this chart, note that this site will provide a rather comprehensive detailed analysis of the 170 diverse tests. You'll find that while a few workloads in technical and high-performance computing do seem to prefer having exclusive use of all physical core resources, many workloads gain incremental performance with SMT enabled.

Importantly, when Phoronix tested 4th and 5th generation EPYC CPUs across a wide variety of workloads it also measured minimal to no difference in power consumption when SMT is enabled vs disabled.

"For workloads able to benefit from SMT, it's still a clear win with AMD EPYC 9005 processors. When looking at all of the CPU power consumption across 170+ benchmarks taking ~13 hours to complete, the data here shows no power consumption difference overall to having SMT enabled" - Phoronix

The significant SMT performance gains (often in the range of 30-50%) combined with virtually no or minimal change in power consumption means that energy efficiency is getting a boost—better performance per watt! SMT is a major contributor to energy efficiency on modern x86 superscalar CPUs such as 5th gen AMD EPYC, together with power management and dynamic frequency scaling. The following comments summarize the benefits:

"SMT enabled on the AMD EPYC 9575F on average led to just a 2 Watt increase to the CPU power consumption than when it was disabled." - Phoronix

Why do we see efficiency? When a core is in normal operation state (C0) executing instructions, a thread stall while waiting on data doesn't move it to a lower power state to save power, but having a second thread to fill in the gaps can make a big difference in performance. The increased instruction throughput may slightly increase power consumption, at the same time power efficiency often improves much more.

AMD EPYC and SMT: Still delivering great value after all these years



Simultaneous multithreading was developed in a time when core resources were quite precious—one, two or perhaps 4 cores per socket—and it was essential for customers to be able to squeeze as much processing out of them as possible. In an age where 5th Gen AMD EPYC processors offer up to 192 physical high-performance "Zen 5" cores per socket, it may seem natural to ask if these resources are still quite so precious and if SMT still carries value. If you ask any IT manager struggling to balance incredible growth in demand for compute resources and budgets, you'll likely hear a resounding "yes".

While physical cores are now quite plentiful, they are also still quite valuable for there is often a LOT of work to be accomplished and a significant other driver of IT solutions cost —software license costs—are often tied to the number of physical cores in the host server! The typical IT shop needs to get the most out of every resource—and having the flexibility to gain incremental compute capacity and performance with as few hardware resources as possible can deliver a powerful ROI. SMT is a compelling option: allowing relatively "free" performance boost where it can add value, but also easily disabled where it does not.

References:

1. [Exploring The Zen 5 SMT Performance With The AMD EPYC 9755 "Turin" CPU - Phoronix](#)
2. [SMT Remains Very Advantageous For 5th Gen AMD EPYC Performance Review - Phoronix](#)
3. [SMT Proves Worthwhile Option For 128-Core AMD EPYC "Bergamo" CPUs Review - Phoronix](#)

Related Blogs

[View All Blogs >](#)



Performance Profiling on AMD GPUs - Part 3: Advanced Usage — ROCm Blogs

Part 3 of our GPU profiling series guides beginners through practical steps to identify and optimize kernel bottlenecks using ROCm tools

October 22, 2025



AMD Driving the Next Frontiers of Science

AMD's full-spectrum AI infrastructure, combining HPC, AI accelerators, and open ecosystems, empowers researchers to tackle complex problems faster.

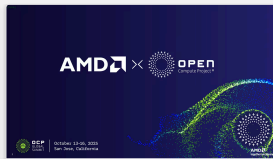
October 21, 2025



Open Standards for AI Scale: AMD and OCP Shape Infrastructure

AMD CTO Mark Papermaster shares how AMD and OCP are scaling AI with open standards, ROCm, and Helios. Read his keynote recap blog.

October 15, 2025



Open Ecosystems and Scalable Compute for the AI Era

At OCP Global Summit 2025, AMD showcases how open collaboration drives scalable, sustainable, and interoperable data center innovation.

October 14, 2025



Advancing Quantum Chemistry through HPC and VeloxChem

VeloxChem on LUMI's AMD GPUs, researchers at KTH, CSC, and AMD Silo AI achieve scalable simulations for drug discovery and materials science.

October 13, 2025



Subscribe to the latest news from AMD



Company

- [About AMD](#)
- [Management Team](#)
- [Corporate Responsibility](#)
- [Careers](#)
- [Contact Us](#)

News & Events

- [Newsroom](#)
- [Events](#)
- [Media Library](#)

Resources

- [Developer Central](#)
- [Blogs](#)
- [Case Studies](#)
- [Webinars](#)
- [Explore Resources](#)

Partners

- [AMD Partner Hub](#)
- [Authorized Distributors](#)
- [AMD University Program](#)

Investors

- [Investor Relations](#)
- [Financial Information](#)
- [Board of Directors](#)
- [Governance Documents](#)
- [SEC Filings](#)

[Terms and Conditions](#) | [Privacy](#) | [Trademarks](#) | [Supply Chain Transparency](#) | [Fair & Open Competition](#) | [UK Tax Strategy](#) | [Cookies Policy](#) | [Cookie Settings](#)

© 2025 Advanced Micro Devices, Inc.

