

4TH GEN AMD EPYC™ PROCESSOR ARCHITECTURE



together we advance_data center computing

*Fourth Edition
November 2023*

4TH GEN AMD EPYC PROCESSOR ARCHITECTURE

CONTENTS

INTRODUCTION	3
HYBRID MULTI-DIE ARCHITECTURE	4
Decoupled Innovation Paths	4
'Zen 4' Core	5
'Zen 4c' Core	5
Modularity Enables Innovation	5
Simplified Production	7
I/O Die Innovation	7
AMD Infinity Architecture	7
4TH GEN EPYC PROCESSOR CORES	8
'Zen 4' CPU Die	8
'Zen 4c' CPU Die	8
Double-Digit IPC Improvements	8
AVX-512 Instructions	8
Larger Addressable Memory	9
Security Enhancements	9
SYSTEM-ON-CHIP DESIGN	10
AMD Infinity Fabric™ Technology and the I/O Die SERDES	11
NUMA Considerations	12
MULTIPROCESSOR SERVER DESIGNS	13
Single-Socket Server Configurations	13
2-Socket Server Configurations	13
AMD INFINITY GUARD FEATURES	14
Cutting-Edge Security Features	14
AMD Secure Processor	14
CONCLUSION	16

INTRODUCTION



The information technology industry is changing rapidly, with many different workload facets demanding innovation that can satisfy their specific needs. High-performance computing and cloud applications need high-density CPUs with high core counts for highly parallelized workloads. Enterprise applications need a balance between CPU and I/O capability. Artificial intelligence, data analytics, high-performance computing, as well as structured and unstructured data applications are driven by the strength and speed of individual cores and accelerated mathematical functions. Certain cloud, manufacturing, healthcare, retail, telco, and other edge applications need highly efficient computing that can operate in challenging environments, from closets in retail stores to telephone switching and cellular transmission locations. And network infrastructure, security, and edge applications need cost- and power-optimized systems that can be deployed confidently in locations around the globe.

The design decisions we have made in the 4th generation of AMD EPYC™ processors have evolved a platform that can support all of these needs. The following goals have driven the design of the AMD EPYC 8004 and 9004 Series processors:

- **INSTRUCTIONS-PER-CLOCK (IPC)** improvements, ranging in double-digit increases across generations
- **EXCELLENT EFFICIENCY**, with leadership performance per watt

- **BALANCED ARCHITECTURE**, with high memory bandwidth and I/O capacity to match the CPU's voracious appetite for data
- **LOW LATENCY**, with a goal of reducing average latency with higher cache sizes and effectiveness
- **HIGH THROUGHPUT**, with a goal of reducing dynamic power to enable significantly higher core counts

This white paper describes the processor architecture that supports 4th Gen AMD EPYC processors and future enhancements that enable you to branch out and address a continuously widening universe of workload demands. Our hybrid, multi-chip architecture enables us to decouple innovation paths and deliver consistently innovative, high-performance products. The 'Zen 4' and 'Zen 4c' cores represent a significant advancement from the last generation, with new support for highly complex machine learning and inferencing applications. Our system-on-chip approach helps server vendors to accelerate their designs and get innovative products into customers' hands quickly. AMD EPYC processors are the only x86 server CPUs with an integrated, embedded security processor that is "hardened at the core" to help secure customer data whether in a central data center or distributed across locations at the network edge. Finally, this paper will review some of the design choices that enable no-compromise single-socket servers as well as some of the most powerful two-socket servers on the planet.

HYBRID MULTI-DIE ARCHITECTURE

The most important innovation in AMD EPYC processors is the hybrid multi-die architecture first introduced in 2nd Gen EPYC processors. We anticipated the fact that increasing core density in monolithic processor designs would become more difficult over time. One of the primary issues is the fact that the process technology that can create a CPU core is on a different innovation path than the technology that lays down the analog circuitry to drive external pathways to memory, I/O devices, and an optional second processor. These two technologies are linked together when creating monolithic processors and can impede the swift delivery of products to market.

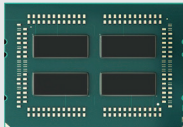
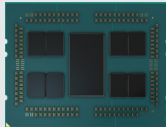
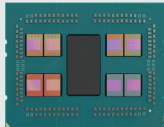
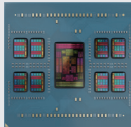
DECOUPLED INNOVATION PATHS

Second, third, and fourth-Gen AMD EPYC processors have decoupled the innovation paths for CPU cores and I/O functions into two different types of dies that can be developed on timelines

appropriate for what they need to do. For example, the 'Zen 4' cores are produced with 5nm technology and the I/O die is created with 6nm processes. This decoupling has enabled higher core densities within the same thermal envelope. Continuous improvement in instructions per cycle has yielded double-digit performance gains with every new generation (Table 1). The approach we have taken is more flexible and dynamic than trying to force all aspects of a processor into one fabrication technology. We believe that it is faster to deliver high-performance products and specialized CPUs to market by assembling modules into a processor than to create large, monolithic CPUs.

In today's 4th Gen AMD EPYC processors we use two different cores to address a range of workload needs by varying the type and number of cores and how we package them. The EPYC 9004 Series uses an SP5 form factor and processors within this series use either the 'Zen 4' or 'Zen 4c' core designs. The EPYC 8004 Series

Table 1: The multi-die architecture has enabled significant improvements for each processor generation since the beginning

	AMD EPYC 7001 'NAPLES'	AMD EPYC 7002 'ROME'	AMD EPYC 7003 'MILAN'	AMD EPYC 9004, 8004 'GENOA', 'SIENA'
				
Core Architecture	'Zen'	'Zen 2'	'Zen 3'	'Zen 4' and 'Zen 4c'
Cores	8 to 32	8 to 64	8 to 64	8 to 128
IPC Improvement Over Prior Generation	N/A	~24% ^{ROM-236}	~19% ^{MLN-003}	~14% ^{EPYC-038}
Max L3 Cache	Up to 64 MB	Up to 256 MB	Up to 256 MB	Up to 384 MB (EPYC 9004) Up to 128 MB (EPYC 8004)
Max L3 Cache with 3D V-Cache™ technology			768 MB	Up to 1152 MB
PCIe® Lanes	Up to 128 Gen 3	Up to 128 Gen 3	Up to 128 Gen 4	Up to 128 Gen 5 8 bonus lanes Gen 3
CPU Process Technology	14nm	7nm	7nm	5nm
I/O Die Process Technology	N/A	14nm	14nm	6nm
Power (Configurable TDP [cTDP])	120-200W	120-280W	155-280W	70-400W
Max Memory Capacity	2 TB DDR3-2400/2666	4 TB DDR4-3200	4 TB DDR4-3200	6 TB DDR5-4800

is designed for a smaller SP6 package, and is optimized for single-socket environments. This series uses the 'Zen 4c' core.

'ZEN 4' CORE

This core is optimized for high performance. Up to eight cores are combined to create a core complex (CCX) that includes a 32 MB shared L3 cache. This core complex is fabricated onto a die (CCD), up to 12 of which can be configured into an EPYC 9004 processor for up to 96 cores in the SP5 form factor. Compared to the prior generation, the 'Zen 4' core delivers an estimated 45% more integer and 75% more floating point performance for 64-core processors, which frees thermal envelopes to deliver more computing power.^{SP5-003B, SP5-004B}

'ZEN 4C' CORE

This core is optimized for density and efficiency. It has the exact same register-transfer logic as the 'Zen 4' core, but its physical layout takes less space and is designed to deliver more performance per watt. The 'Zen 4c' core complex includes up to eight cores and a shared 16 MB L3 cache. Two of these core complexes are combined onto a single die for 16 cores per die and a total of 32 MB of L3 cache per die (Figure 1). In the EPYC 9004 Series, up to eight of these dies can be combined with an I/O die to deliver CPUs with up to 128 cores in an SP5 form factor. In the EPYC 8004 Series, up to four 'Zen 4c' dies combine to deliver CPUs with up to 64 cores in an SP6 form factor designed for single-socket systems with small physical footprints and operating in environmentally challenging conditions.

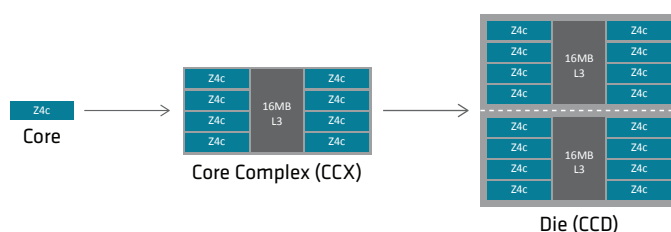


Figure 1: Relationship between core, core complex, and die

MODULARITY ENABLES INNOVATION

Our modular approach enables us to deploy the CPU die as a unit of innovation where we can create variants that are targeted to address specific workloads. It's a flexible unit that we can swap in and out to more closely balance computing power and efficiency requirements with workloads:

- **BALANCED WORKLOADS:** We use our 'Zen 4' core to address mainstream performance needs including application development, business applications, data management and

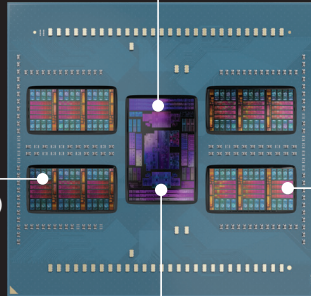
analytics, collaborative, and infrastructure applications. Each CPU die includes up to eight 'Zen 4' cores that are optimized for high performance per core. These share a 32 MB L3 cache. Up to 12 of these dies can be used to create a processor with up to 96 cores.

- **LICENSE-COST-CHALLENGED WORKLOADS:** When you pay per-core software license fees, you want to get the most performance from each core. For these, and for other workloads needing high per-core performance, we have created a range of high-frequency options with fewer cores and higher clock speeds. These CPUs, with 'F' at the end of the part number, use the same die but with only a small number of active cores per die. This spreads the thermal load across the processor package and enables us to increase the clock frequency. For example, the 16-core EPYC 9174F uses eight dies with only two active cores per die, enabling a base frequency of 4.10 GHz compared to our standard 16-core EPYC 9124 with a base frequency of 3.00 GHz. Our high-frequency processors are available with up to 48 cores.
- **MEMORY-INTENSIVE WORKLOADS:** Many technical workloads process models that require large amounts of memory, putting high demands on memory throughput and cache. These include RTL simulation, computational fluid dynamics, weather forecasting, and molecular dynamics. Some business applications fall into this category as well, including Java® enterprise middleware. In processors with AMD 3D V-Cache™ technology, we literally stack additional cache memory on top of each CPU die bringing the per-die total to 96 MB of L3 cache. This uses a direct copper-to-copper hybrid bonding process that enables more than 200 times the interconnect densities of current 2D technology and more than 15 times the interconnect technologies that use solder bumps for the connection.^{EPYC-026} This innovation delivers up to 1152 MB of L3 cache in our 4th Gen processors with 3D V-Cache.
- **COMPUTE-INTENSIVE WORKLOADS:** For some workloads, even 96 cores per processor may not be enough. These include cloud-native applications developed with containers, virtualized environments striving for the highest number of virtual machines or containers per server, and highly parallelized workloads including life sciences, chemistry, content rendering, and delivery. To address these needs, deploy up to eight density-optimized 'Zen 4c' dies with eight cores and a total of 1 MB of L2 and 32 MB of L3 cache per die. This brings the total core density to up to 128 cores per processor in the EPYC 9754 CPU. We offer one step down to 112 cores per processor in our EPYC 9634 part.
- **WORKLOADS RUNNING IN CHALLENGING ENVIRONMENTS:** Some workloads require more energy efficiency than cores, and they must be able to run in locations with challenging environmental conditions. For example, manufacturing, retail, healthcare applications often have to run in closets. Consider massive

4TH GEN AMD EPYC PROCESSOR ARCHITECTURES

I/O die
12 memory controllers
PCIe® Gen 5 controllers
Infinity Fabric™ controllers
SATA controllers
CXL™ controllers
AMD Secure Processor

CPU die
Up to 16 cores per die (8 shown)
Up to 12 dies per processor



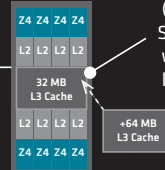
AMD EPYC 9004 SERIES PROCESSORS (16-96 CORES)

'Zen 4' CPU die
(up to 12 per processor)
8 'Zen 4' cores
1 MB L2 cache per core
Shared 32 MB L3 cache



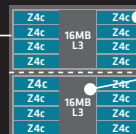
AMD EPYC 9004 SERIES PROCESSORS WITH 3D V-CACHE TECHNOLOGY

'Zen 4' CPU die with 3D V-Cache technology
(Up to 12 per processor)
Shared 32 MB L3 cache with 64 MB additional layered above (96 MB total)



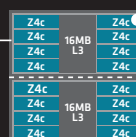
AMD EPYC 9004 SERIES PROCESSORS (122-128 CORES)

'Zen 4c' CPU die
(Up to 8 per processor)
16 'Zen 4c' cores
1 MB L2 cache per core
Total 32 MB L3 cache per core complex

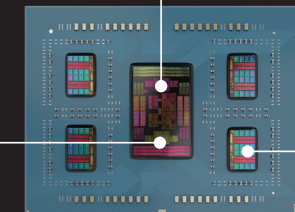


AMD EPYC 8004 SERIES PROCESSORS (8-64 CORES)

'Zen 4c' CPU die
(Up to 4 per processor)



I/O die
6 memory controllers and 96 PCIe lanes connected to SP6 package pins





amounts of parallel processing in telco and 5G buildings that require NEBS™ compliance. And some file servers and cloud environments need fewer, more efficient cores. The AMD EPYC 8004 Series deploys from one to four ‘Zen 4c’ cores in a smaller SP6 package that holds from 8 to 64 cores. The EPYC 8004 Series is designed for single-socket servers with high efficiency and low power enabling less cooling and better acoustics in constrained environments.

- **INNOVATIVE FUTURE:** The hybrid multi-die architecture puts us on a path of rapid innovation. Our 3D V-Cache technology adds cache memory to CPU dies today, but we don’t see ourselves limited to adding static capacity with our solderless connections. What if we could add specialized processing units to the CPU die?

SIMPLIFIED PRODUCTION

The multi-die architecture can help reduce waste in the fabrication process. When we place many (relatively) small CPU dies on a silicon wafer, the inevitable production flaws affect a small number of dies that fail testing and are not integrated into any processors. In comparison, if the wafer contains fewer, larger, monolithic processors, a single flaw can cause the entire processor to be rejected, reducing the overall yield in terms of average number of processors produced per wafer. This can contribute to higher costs.

I/O DIE INNOVATION

The I/O die is a place for parallel innovation. In the EPYC 9004 Series we have doubled the I/O bandwidth of the CPU from the past generation by incorporating PCIe® Gen 5 capabilities onto the I/O die. Not being satisfied with just doubling the I/O bandwidth, the I/O die supports 12 DDR5 memory controllers, AMD Infinity Fabric™ interconnects, SATA disk controllers, and Compute

Express Link (CXL™) 1.1+ memory controllers that can be flexibly assigned to specific functions at server design time. The I/O die is where the dedicated AMD Secure Processor resides, close to the memory controllers that manage the range of memory encryption mechanisms that are part of our AMD Infinity Guard^{GD-183} feature set.

The I/O die used in all 4th Gen AMD EPYC processors has 12 Infinity Fabric connections to CPU dies. Our CPU dies can support one or two connections to the I/O die. In processor models with four CPU dies, two connections can be used to optimize bandwidth to each CPU die. This is the case for some EPYC 9004 Series CPUs and all EPYC 8004 Series CPUs. In processor models with more than four CPU dies, such as in the EPYC 9004 Series, one Infinity Fabric connection ties each CPU die to the I/O die.

The exact same I/O die is used in all 4th Gen EPYC processors, however their capabilities vary with how the die is connected to the outside world. The larger SP5 package enables pins to carry 128 lanes of I/O bandwidth and 12 memory channels. The smaller SP6 package has fewer pins so the I/O die is connected to 96 lanes of I/O connectivity and six memory channels.

AMD INFINITY ARCHITECTURE

When creating a processor based on a hybrid, multi-chip architecture, the performance of the interconnect is of paramount importance. The heart of the AMD Infinity Architecture is a leadership interconnect that supports extraordinary levels of scale at every layer. Components communicate using AMD Infinity Fabric technology—a connection that is used between CPUs, between components in the multi-chip architecture, and to connect processor cores, memory, PCIe® Gen 5 I/O, and security mechanisms. As a result, the architecture delivers breakthrough performance and efficiency to deliver on the promise of next-generation computing.

4TH GEN EPYC PROCESSOR CORES

At AMD, our core design is an undertaking of continuous optimization. The ‘Zen 4’ core integrated into 4th Gen AMD EPYC processors comprises the first and only x86 server CPU built with 5nm fabrication technology. Because we build our server processors as part of a multi-chip architecture, the CPU die is a component that can be innovated and enhanced independently of the I/O die.

‘ZEN 4’ CPU DIE

The ‘Zen 4’ CPU die used in EPYC 9004 Series processors consists of up to eight cores, dedicated 1 MB L2 cache per core, and a 32 MB cache shared between the eight cores (Figure 2). This die can be augmented with 3D V-Cache technology to bring the L3 cache capacity to 96 MB.

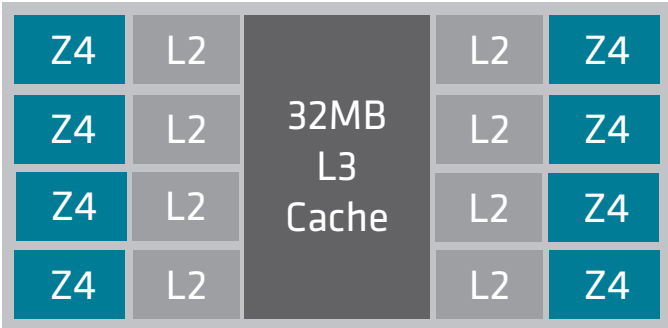


Figure 2: Layout of the ‘Zen 4’ CPU die with 8 cores per die

‘ZEN 4C’ CPU DIE

The ‘Zen 4c’ CPU core is designed for high density and energy efficiency. The same logic as the ‘Zen 4’ core is more closely packed in its core complex to build processors with up to 128 cores. The ‘Zen 4c’ CPU die holds two core complexes, each with eight cores each having 1 MB L2 cache and a shared 16 MB L3 cache (Figure 3). In EPYC 9004 Series, up to to eight of these dies can be attached to the I/O die for a total of up to 128 cores per processor for ultra-dense, high-performance systems. In the EPYC 8004 Series, up to four dies can be configured to create CPUs from 8 to 64 cores per processor. With its low per-core power requirement, the ‘Zen 4c’ CPU

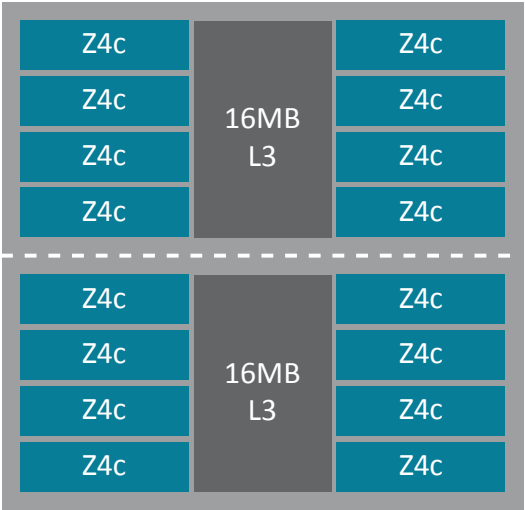


Figure 3: The ‘Zen 4c’ CPU die holds two core complexes for a total of 16 cores per die

die is used in the EPYC 8004 Series for traditional data centers as well as diverse edge servers, and innovative form factors targeted for manufacturing, healthcare, retail, and telco.

DOUBLE-DIGIT IPC IMPROVEMENTS

For each generation, we strive for double-digit percentage improvements in instructions per cycle, which we have been able to deliver with each new EPYC processor series (see Table 1). Improvements over the ‘Zen 3’ core include 1 MB L2 private cache per core, branch-prediction improvements, larger ,micro-operation cache, and deeper internal buffers.

AVX-512 INSTRUCTIONS

Many applications today strive to gain knowledge from data, and they repeat arithmetic calculations on large amounts of data. These workloads include:

- Machine learning and inferencing
- High-performance computing
- Computational fluid dynamics

- Finite element analysis
- Financial services
- Image and audio/video processing
- Cryptography and data compression

While most applications use a single instruction to operate on a single data element (SISD), these applications need parallel execution of multiple data elements directed by a single instruction (SIMD). Some codes, including HPC, financial services, and video processing use vectors of full-precision floating-point data. Machine learning and inferencing workloads are increasingly using half-precision arithmetic including 16-bit floating point and 8-bit integer operations to speed the flow of data and reduce the power needed to process large data sets.

AVX-512 is a set of instructions based on an SIMD model. As its name suggests, a single instruction operates on a 512-bit vector of 8-, 16-, 32-, or 64-bit data values. 4th GEN EPYC processors implement the full set of AVX-512 instructions used in 4rd Gen Intel Xeon CPUs such as BFLOAT16 and Vectorized Neural Network Instruction (VNNI). Our implementation of these data-heavy instructions enables applications that are hard coded for AVX-512 to work without modification.

Implementing the AVX-512 instruction required carefully balancing the 'Zen 4' and 'Zen 4c' core design goals. Both cores use 256-bit data paths internally, including 256-bit floating-point units, helping

to reduce core size and increase core density. 4th Gen EPYC processor cores use a 512-bit AVX register file, from which two 256-bit vectors are executed in sequential clock cycles (Figure 4). Adding this register increased core size slightly but overall it kept 'Zen 4' using 40% less area than the Intel "Sunny Cove" core.^{EPYC-041}

Another drawback that the 256-bit data path avoided is the need to throttle processor frequency to avoid thermal spikes when 512-bit floating-point operations are executed. Overall, our approach that favors core density even wins out on performance. The geometric mean of workloads tested by Phoronix shows the 96-core AMD EPYC 9654 outperforming the top-of-the-line Intel® Xeon® 8490H by ~19 percent.¹

LARGER ADDRESSABLE MEMORY

With the potential to expand memory through CXL controllers, virtual memory is now addressable through 57 bits, and a fifth level of nested page tables has been implemented to support this.

SECURITY ENHANCEMENTS

Each 'Zen' core generation builds upon the security features of the previous one, and they incorporate mitigations for vulnerabilities known at design time with no modifications necessary to application software. The original 'Zen' core has resisted certain side-channel attacks in part because of the tagging of memory to threads once read into the processor caches. This helps reduce the possibility of one thread being able to view another thread's data when in use in the processor. For the 'Zen 4' core we introduced the capability for guest operating systems in virtualized environments to run exclusively on one core—thus introducing further solutions that can help protect against side-channel attacks targeted at cached memory.

New support for virtualized environments includes secure multi-key encryption (SMKE) that enables hypervisors to selectively encrypt address space ranges on CXL-attached memory. Memory encrypted with SMKE can be accessed by the CPU across reboots, and the existing software encryption framework works seamlessly with CXL-attached memory as well independent of device implementation.

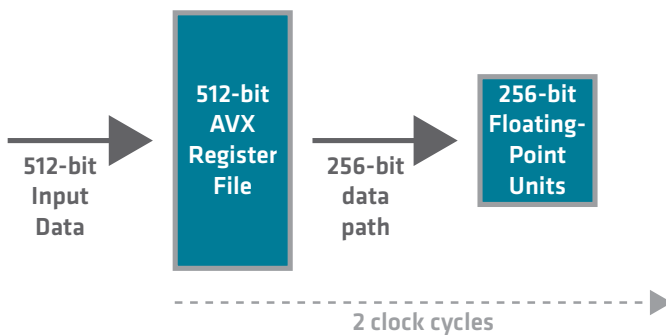


Figure 4: The 'Zen 4' core implements an energy-efficient AVX-512 instruction with 256-bit internal data paths

SYSTEM-ON-CHIP DESIGN

The I/O die (Figure 5) implements many of the functions that would normally be implemented with external chip sets, thus qualifying AMD EPYC processors as systems on chip (SOCs). This approach can help reduce server design complexity and power consumption because fewer chips outside of the CPU are needed. Our all-in philosophy means that every processor series has the built-in features that are described below. This takes the mystery out of CPU selection. Just choose the core count, frequency, and L3 cache size your workload requires, and the rest are included at no extra cost.

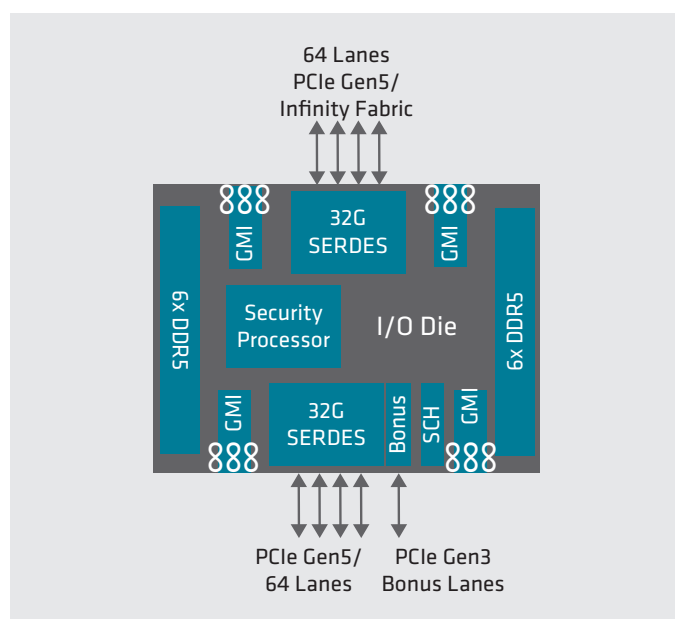


Figure 5: The I/O die implements many functions that would otherwise require external chip sets (EPYC 9004 Series shown)

- **DDR5 MEMORY CONTROLLERS**—12 in the EPYC 9004 Series, 50% more memory controllers than any other x86 processor.^{[EPYC-033](#)} Having more, and more high-performance x86 CPU cores creates a higher demand for memory, and more memory channels drive high memory bandwidth that keeps this equation in balance. Memory interleaving on 2, 4, 6, 8, 10, and 12 channels

helps optimize for both small- and large-memory configurations. The memory controllers include inline encryption engines for implementing AMD Infinity Guard features discussed below. In the EPYC 8004 Series, six memory channels are interleaved on 2, 4, and 6 memory channels.

- **UP TO 128 PCIe GEN 5 LANES IN A 1P CONFIGURATION; UP TO 160 LANES IN A 2P CONFIGURATION (EPYC 9004 SERIES); UP TO 96 IN THE EPYC 8004 SERIES.** The available PCIe Gen 5 lanes can be dedicated to support higher-level functions including up to 32 PCIe lanes (also 32 lanes in the 8004 Series) configurable as on-chip SATA controllers for massive disk capacity and up to 64 lanes (48 lanes in the 8004 Series) configurable as CXL 1.1+ memory controllers for cache-coherent memory expansion and support for persistent memory. In server designs, the bonus lanes are often used for access to performance-insensitive I/O such as to M.2 drives used for system boot.
- **UP TO 12 PCIe GEN 3 'BONUS' LANES** in a 2-socket EPYC 9004 Series configuration, or 8 lanes in a single-socket configuration (EPYC 8004 and 9004 Series).
- **2X FASTER AMD INFINITY FABRIC CONNECTIVITY IN THE EPYC 9004 SERIES** over the prior generation for CPU-to-CPU connectivity. (The EPYC 8004 Series is 1P only so this does not apply). Rather than invent new connectivity mechanisms that can delay time to market, we use the same physical interfaces for Infinity Fabric connections as for the PCIe Gen 5 I/O, with different protocols layered on the physical (PHY) layer. This affords server designers the freedom to trade off more PCIe I/O lanes in exchange for fewer interprocessor communication links. AMD supports use of 3 or 4 links each of which correspond to x16 PCIe physical connections. With Infinity Fabric protocols running on these interfaces, four links can support a maximum theoretical bandwidth of 512 GB/s between processors, which allows AMD to maintain a balanced architecture as we support updated DDR and PCIe interface standards.
- **UPDATED INFINITY FABRIC INTERFACE** offers up to 36 Gb/s for communication between each CPU die and the I/O die where a total of 12 Infinity Fabric connections can connect CPU dies to I/O

dies. The new ‘Zen4’ and ‘Zen 4c’ CPU dies can use one or two Infinity Fabric interfaces, allowing for double the CPU-core-to-I/O die bandwidth (up to 72 Gb/s) for processor models with four or fewer CPU dies. The 4th Gen EPYC I/O die offers great flexibility with twelve Infinity Fabric interfaces, also accommodating 4-, 8-, or 12-core dies depending on the performance and power requirements per customer use case. (This is known internally as the Global Memory Interface (GMI) and is labeled this way on many figures.)

- **INTEGRATED AMD SECURE PROCESSOR** that supports confidential computing with features including secure root of trust, secure memory encryption (SME), and secure encrypted virtualization (SEV).^{GD-183} This is discussed in a separate section below
- **A SERVER CONTROLLER HUB** can minimize the required chip set for basic server control functions. It includes direct USB connectivity, 1 Gb/s LAN-on-motherboard, and various UART and I2C and I3C bus connectivity.

AMD INFINITY FABRIC™ TECHNOLOGY AND THE I/O DIE SERDES

The use of the same physical layer to support I/O functions including AMD Infinity Fabric technology reflects our philosophy of using industry-standard, well-understood technologies that offer server designers flexibility to design innovative servers, and simplifies our CPU designs over inventing proprietary interconnects.

The PCIe Gen 5 I/O is supported in the I/O die by serializer-deserializer (SERDES) silicon with one independent set of traces to support each port of 16 PCIe lanes. The I/O die contains eight SERDES devices. In the EPYC 9004 Series, four are typically used to connect to a second processor and four to connect to I/O devices. Each of these devices can be customized so that the underlying PCIe Gen 5 PHY circuitry can be used for:

- Up to 4 links of Gen3 AMD Infinity Fabric connectivity (zero in the 1P-optimized EPYC 8004 Series)

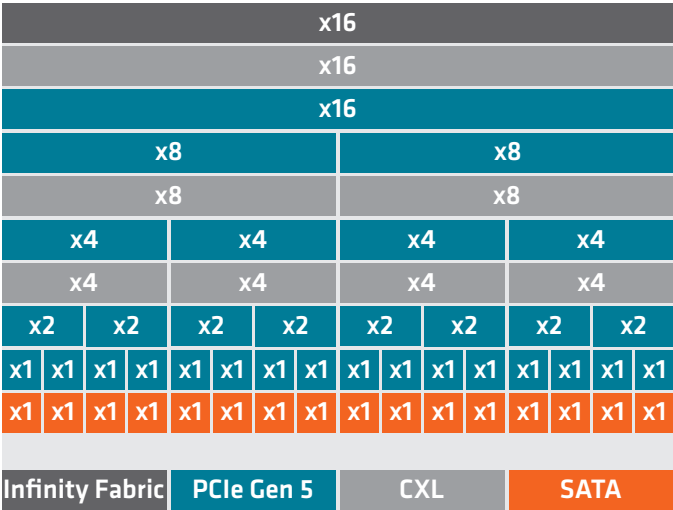


Figure 6: Idealized example of SERDES lane bifurcation options

- 128 lanes of PCIe Gen 5 connectivity to peripherals (96 lanes in the EPYC 8004 Series and up to 160 lanes in EPYC 9004 Series 2-socket designs)
- Up to 64 lanes that can be dedicated to CXL 1.1+ connectivity to extended memory (48 lanes in the EPYC 8004 Series)
- Up to 32 I/O lanes that can be configured as SATA disk controllers

The lanes in each SERDES can be bifurcated given constraints described in server design documentation. Each SERDES has specific constraints, for example some are restricted to PCIe and Infinity Fabric connectivity, while others enable the richer set of functions.

An idealized bifurcation diagram—no single SERDES provides all of them—is illustrated in Figure 6, indicating that the entire port can be dedicated to 16 lanes of Infinity Fabric, PCIe, or CXL connectivity. These can be broken down to various combinations of x8, x4, x2, and x1 bandwidth. For example, if SATA controllers share connectivity with PCIe on a SERDES, a maximum of eight x1 SATA controllers can be allocated. Or, as the diagram illustrates, CXL connections must use a minimum of four lanes.



NUMA CONSIDERATIONS

In a multi-chip architecture, there can be varying amounts of memory latency depending on the connectivity between memory controllers and CPU dies. This is known as non-uniform memory access, or NUMA. For applications needing to extract every last percent of latency out of memory accesses, they can take advantage of these varying latencies to create an affinity between specific address ranges and the CPU cores closest to that memory.

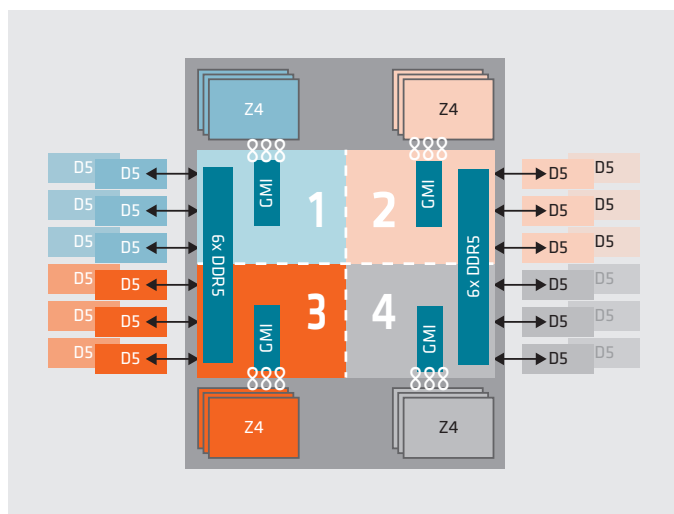


Figure 7: Dividing a 4th Gen AMD EPYC processor into four NUMA domains can give small performance improvements for some applications (EPYC 9004 Series processor shown)

In AMD EPYC 7001 Series processors, memory controllers were located on the same die with up to eight CPU cores, creating a tight affinity between the memory controlled by the die and the CPU cores on the die. When a memory controller had to request data destined for a different set of cores, the data had to pass from one die to another over an internal Infinity Fabric connection.

Beginning with AMD EPYC 7002 Series processors, non-uniform latency was reduced dramatically by locating memory controllers onto the I/O die. NUMA domains were flattened more by the move to 32 MB of L3 cache in the EPYC 7003 Series. In 4th Gen EPYC processors, optimizations to the Infinity Fabric interconnects reduced latency differences even further.

Still, for applications that need to squeeze the last one or two percent of latency out of memory references, creating an affinity between memory ranges and CPU dies ('Zen 4' or 'Zen 4c') can improve performance. Figure 7 illustrates how this works. If you divide the I/O die into four quadrants for an 'NPS=4' configuration, you will see that six DIMMs feed into three memory controllers, which are closely connected via Infinity Fabric (GMI) to a set of up to three 'Zen 4' CPU dies, or up to 24 CPU cores.

Most applications don't need to be concerned about using NUMA domains, and using the AMD EPYC processor as a single domain (NPS=1) gives excellent performance. The [AMD EPYC 9004 Architecture Overview](#) provides more details on NUMA configurations and tuning suggestions for specific applications.

MULTIPROCESSOR SERVER DESIGNS

This chapter applies only to the EPYC 9004 Series, as the EPYC 8004 Series is designed for single-socket servers only.

The flexibility of the SERDES enables the Infinity Fabric interconnects to share the same physical infrastructure of chip's PCIe I/O. In Figure 8, these are labeled as 'G' and 'P' links, each of which support 16 lanes of PCIe Gen 5 connectivity. In a single-socket configuration, all Infinity Fabric links are dedicated to PCIe I/O, affording 128 lanes of Gen 5 bandwidth on AMD EPYC 9004 Series processors.

SINGLE-SOCKET SERVER CONFIGURATIONS

AMD EPYC processors with no 'P' suffix can be used in single-socket and 2-socket configurations. Processor part numbers with a 'P' suffix

are optimized for single-socket servers by dedicating the 'P' links for PCIe I/O connections only. Figure 8 illustrates a single-socket configuration with two DIMMs per memory channel.

2-SOCKET SERVER CONFIGURATIONS

In these configurations, three or four 16-lane 'G' links are used to connect to the second processor. For I/O-intensive server designs, three links can be used as Infinity Fabric interconnects and one additional link from each CPU can be dedicated to PCIe Gen 5 I/O, bringing the server I/O capacity to 160 lanes (Figure 9).

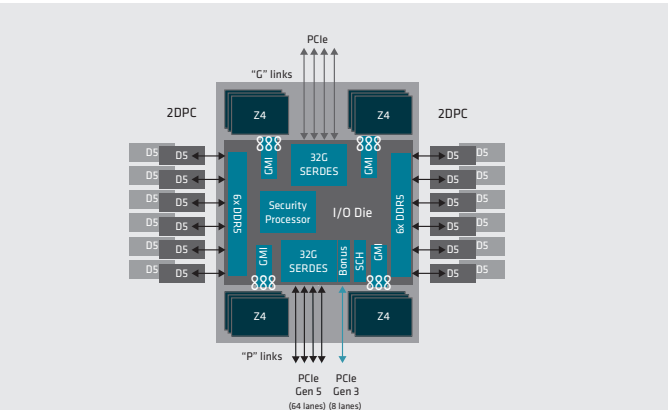


Figure 8: EPYC 9004 Series processor in a single-socket server configuration with all links dedicated to PCIe connectivity

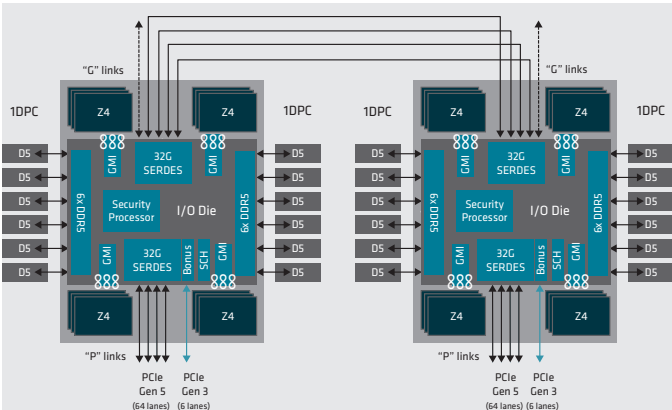


Figure 9: EPYC 9004 Series processors in a 2-socket configurations

AMD INFINITY GUARD FEATURES

Data is every organization's most precious asset, and AMD Infinity Guard security features are designed to help protect your data from malicious users, hypervisors, and even administrators. This approach can help mitigate the risks of attacks against physical DIMMs or attacks against guests in virtualized and hyperconverged environments.

CUTTING-EDGE SECURITY FEATURES

Cutting-edge security features are built into our processors, and, like our core designs, they are the outcome of continuous improvement. Figure 10 illustrates the generation-over-generation improvements we have made to help hypervisors increase the isolation of virtual machines. We are proud to report that select EPYC 9004 processors are on track for United States Federal Information Processing Standard (FIPS) 140-3 certification in 2023.

AMD SECURE PROCESSOR

Security features are managed by the AMD Secure Processor, a 32-bit microcontroller that runs a hardened operating system. The hardening process removes unnecessary components and applies previous security patches in the microcontroller to help reduce attack

surfaces. It provides cryptographic functionality for key generation and key management, and it supervises hardware-validated boot, where the foundation for platform security starts. AMD Infinity Guard security features must be enabled by server OEMs and/or cloud service providers to operate. Check with your OEM or provider to confirm support of these features. These include:

- **HARDWARE-VALIDATED BOOT** helps verify that the operating system or hypervisor software that you intended to load is what is actually loaded. The AMD Secure Processor loads the on-chip boot ROM that loads and authenticates the off-chip boot loader. The boot loader, in turn, authenticates the BIOS before any of the 'Zen' cores can execute the code. Once the BIOS is authenticated, the OS boot loader loads the operating system or hypervisor.
- **AMD SECURE MEMORY ENCRYPTION (SME)** can be used to encrypt all of main memory with no changes required to the operating system or application software. SME helps protect against attacks on the integrity of main memory (such as certain cold-boot attacks) because it encrypts the data. 256-bit AES-XTS encryption engines are built into 4th Gen EPYC processor memory controllers to help reduce performance impact during reading and writing of encrypted memory. These engines can be used to encrypt memory with either 128 or 256-bit keys. The new, 256-bit

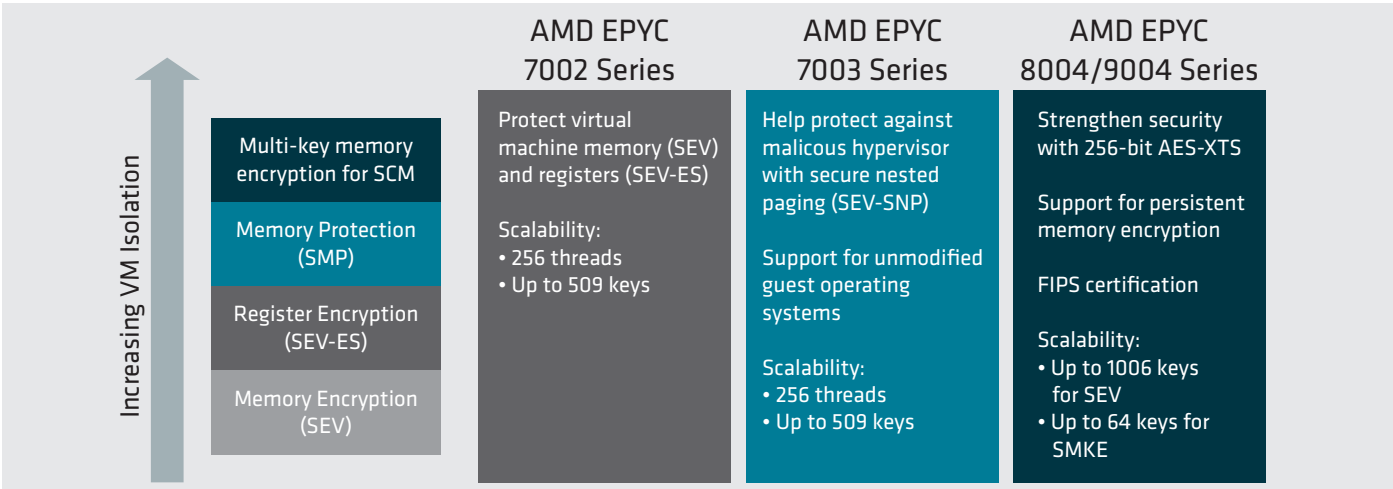


Figure 10: Each new AMD EPYC processor generation delivers more features to help isolate virtual machines

encryption option is integrated into the I/O die in order to support United States Federal Information Processing Standards (FIPS) 140-3 compliance. All of this is done without the encryption key being visible outside of the AMD Secure Processor.

- **AMD SECURE ENCRYPTED VIRTUALIZATION (SEV)** enables hypervisors and guest virtual machines to be cryptographically isolated from one another. Thus, if malicious software is successful in evading the isolation provided by the hypervisor, or if the hypervisor itself is compromised, reading memory from another virtual machine will expose only encrypted data for which the key is stored inside of the AMD Secure Processor and memory controllers. In 4th Gen AMD EPYC processors, up to 1006 keys can be used for virtual machine encryption.
- **AMD SECURE ENCRYPTED STATE (SEV-ES)**, introduced in 2nd Gen AMD EPYC processors, encrypts virtual machine state when interrupts cause it to be stored in the hypervisor. With this information encrypted with the virtual machine's encryption key, a compromised hypervisor is unable to view a virtual machine's registers.
- **AMD SECURE NESTED PAGING (SEV-SNP)** introduced in 3rd Gen AMD EPYC processors, builds on SEV and SEV-ES by adding strong encryption to virtual machine nested page tables to help prevent attacks such as data replay, memory remapping, and more—all with the goal to create confidential, isolated execution environments for virtual machines. With the 57-bit physical memory enabled by 4th Gen AMD EPYC processors, we have increased the page table depth that can be encrypted to five levels.
- **AMD SECURE MULTI-KEY ENCRYPTION (SMKE)**, introduced in 4th Gen AMD EPYC processors, enables fast encryption for storage-class memory, which helps data stored on CXL-attached memory to remain encrypted across a system reboot, helping protect even persistent memory from prying eyes.

This powerful set of security features, is enabled in turn by a multi-layered set of technologies accessible by all of the major hypervisor vendors. It is an innovative set of modern security features that help decrease potential attack surfaces as software is booted, executed, and processes your data. Built-in at the silicon level, AMD Infinity Guard features offer state-of-the-art capabilities to help defend against internal and external threats. Whether yours is a small- or medium-size business or an enterprise organization, implementing robust security features on premises or in the cloud is streamlined with AMD Infinity Guard.



CONCLUSION



AMD EPYC 8004 and 9004 Series processors demonstrate how our hybrid, multi-die architecture delivers real innovation and enables customer value with every new generation. Decoupling our core and I/O development processes enabled us to shrink the CPU die and optimize variants for performance or energy efficiency. Core density enabled by the 'Zen 4' core unleashes a voracious appetite for memory access and I/O capacity, and we set the table with a new I/O die that supports an industry-leading 12 DDR5 memory channels, 50% more than the prior generation. We also doubled our I/O and AMD Infinity Fabric™ throughput by basing them on PCIe Gen 5 interfaces and also added 'bonus' PCIe Gen 3 lanes for devices that are less performance sensitive. Support for domain-specific instructions such as AVX-512, and connectivity to next-gen GPU accelerators prepares AMD EPYC to excel in an increasingly important world of artificial intelligence and machine learning. Increased density also enables support for single-socket servers in environmentally challenged locations such as those in manufacturing, retail, healthcare, telco, and cloud. And if that weren't enough, support for CXL 1.1+ technology enables new Infinity Guard features to help protect even your persistent memory pools from prying eyes. We have raised the bar for data center computing once again with 4th generation of AMD EPYC processors.

END NOTES

For details on the footnotes used in this document, visit amd.com/en/claims/epyc, amd.com/en/claims/epyc3x and amd.com/en/claims/epyc4.

- 1 AVX-512 Performance Comparison: AMD Genoa vs. Intel Sapphire Rapids & Ice Lake” - January 2023, page 8, <https://www.phoronix.com/review/intel-sapphirerapids-avx512>
- EPYC-026 Based on calculated areal density and based on bump pitch between AMD hybrid bond AMD 3D V-Cache stacked technology compared to AMD 2D chiplet technology and Intel 3D stacked micro-bump technology.
- EPYC-033 AMD EPYC 9004 CPUs support 12 memory channels. Intel Scalable Ice Lake CPUs support 8 memory channels. $12 \div 8 = 1.5x$ the memory channels or 50% more memory channels per <https://ark.intel.com/>.
- EPYC-038 Based on AMD internal testing as of 09/19/2022, geometric performance improvement at the same fixed-frequency on a 4th Gen AMD EPYC™ 9554 CPU compared to a 3rd Gen AMD EPYC™ 7763 CPU using a select set of workloads (33) including est. SPECrate®2017_int_base, est. SPECrate®2017_fp_base, and representative server workloads.
- EPYC-041 ~40% less area measures Core + L2 Area: “Zen 4” = 3.84 mm² vs. “Sunny Cove” ~6.5mm². ~48% estimated SPECrate®2017_int_base CCC -D3 Jemalloc results based on internal AMD reference platform and Intel platform measurements of 11/10/2022. Comparison of estimated 1P AMD EPYC 9534 (537 est. SPECrate®2017_int_base, set to 270 Total TDP W, 64 Total Cores, AMD Est) is 1.48x the performance per watt of 1P Intel Xeon Platinum 8380 (363 est. SPECrate®2017_int_base, 270 Total TDP W, 40 Total Cores, AMD est.) for 1.48x the SoC performance/watt. SPEC®, SPEC CPU®, and SPECrate® are registered trademarks of the Standard Performance Evaluation Corporation. See www.spec.org for more information. OEM published scores will vary based on system configuration and determinism mode used (claim uses 270W TDP performance profile)
- GD-183 AMD Infinity Guard features vary by EPYC™ Processor generations. Infinity Guard security features must be enabled by server OEMs and/or Cloud Service Providers to operate. Check with your OEM or provider to confirm support of these features. Learn more about Infinity Guard at <https://www.amd.com/en/technologies/infinity-guard>.
- MLN-003 Based on AMD internal testing as of 02/1/2021, average performance improvement at ISO-frequency on an AMD EPYC™ 72F3 (8C/8T, 3.7GHz) compared to an AMD EPYC™ 7F32 (8C/8T, 3.7GHz), per-core, single thread, using a select set of workloads including SPECrate®2017_int_base, SPECrate®2017_fp_base, and representative server workloads.
- ROM-236 Based on AMD internal testing, average per thread performance improvement at ISO-frequency on a 32-core, 64-thread, 2nd generation AMD EPYC™ platform as compared to 32-core 64-thread 1st generation AMD EPYC™ platform measured on a selected set of workloads including sub-components of SPEC CPU® 2017_int and representative server workloads.
- SP5-003B SPECrate®2017_int_base comparison based on published scores from www.spec.org as of 03/31/2023. Comparison of published 2P AMD EPYC 9534 (1250 SPECrate®2017_int_base, 560 Total TDP W, 128 Total Cores, 2.232 Perf/W, <http://www.spec.org/cpu2017/results/res2023q1/cpu2017-20230116-33519.html>) is 1.45x the performance of published 2P AMD EPYC 7763 (861 SPECrate®2017_int_base, 560 Total TDP W, 128 Total Cores, 1.538 Perf/W, <http://www.spec.org/cpu2017/results/res2021q4/cpu2017-20211121-30148.html>) [at 1.45x the performance/W]. SPEC®, SPEC CPU®, and SPECrate® are registered trademarks of the Standard Performance Evaluation Corporation. See www.spec.org for more information.
- SP5-004B SPECrate®2017_fp_base comparison based on published scores from www.spec.org as of 03/31/2023. Comparison of published 2P AMD EPYC 9534 (1160 SPECrate®2017_fp_base, 560 Total TDP W, 128 Total Cores, \$17606 Total CPU \$, 2.071 Perf/W, <http://www.spec.org/cpu2017/results/res2023q1/cpu2017-20230116-33521.html>) is 1.75x the performance of published 2P AMD EPYC 7763 (663 SPECrate®2017_fp_base, 560 Total TDP W, 128 Total Cores, \$15780 Total CPU \$, 1.184 Perf/W, <http://www.spec.org/cpu2017/results/res2021q4/cpu2017-20211121-30146.html>) [at 1.75x the performance/W] [at 1.57x the performance/CPU\$]. AMD 1Ku pricing and Intel ARK.intel.com specifications and pricing as of 3/31/23. SPEC®, SPEC CPU®, and SPECrate® are registered trademarks of the Standard Performance Evaluation Corporation. See www.spec.org for more information.

© 2022–2023 Advanced Micro Devices, Inc. All rights reserved. All rights reserved. AMD, 3D V-Cache, the AMD Arrow logo, EPYC, Infinity Fabric, and combinations thereof are trademarks of Advanced Micro Devices, Inc. in the United States and/or other jurisdictions. CXL is a trademark of Compute Express Link Consortium, Inc. Intel and Xeon are trademarks of Intel Corporation or its subsidiaries. Java is a registered trademark of Oracle and/or its affiliates. NEBS is a trademark of Telefonaktiebolaget LM Ericsson. PCIe® is a registered trademark of PCI-SIG Corporation. SPEC, SPEC CPU, and SPECrate are trademarks of the Standard Performance Evaluation Corporation. See www.spec.org for more information. Other names are for informational purposes only and may be trademarks of their respective owners.
LE-85001-02 09/23

