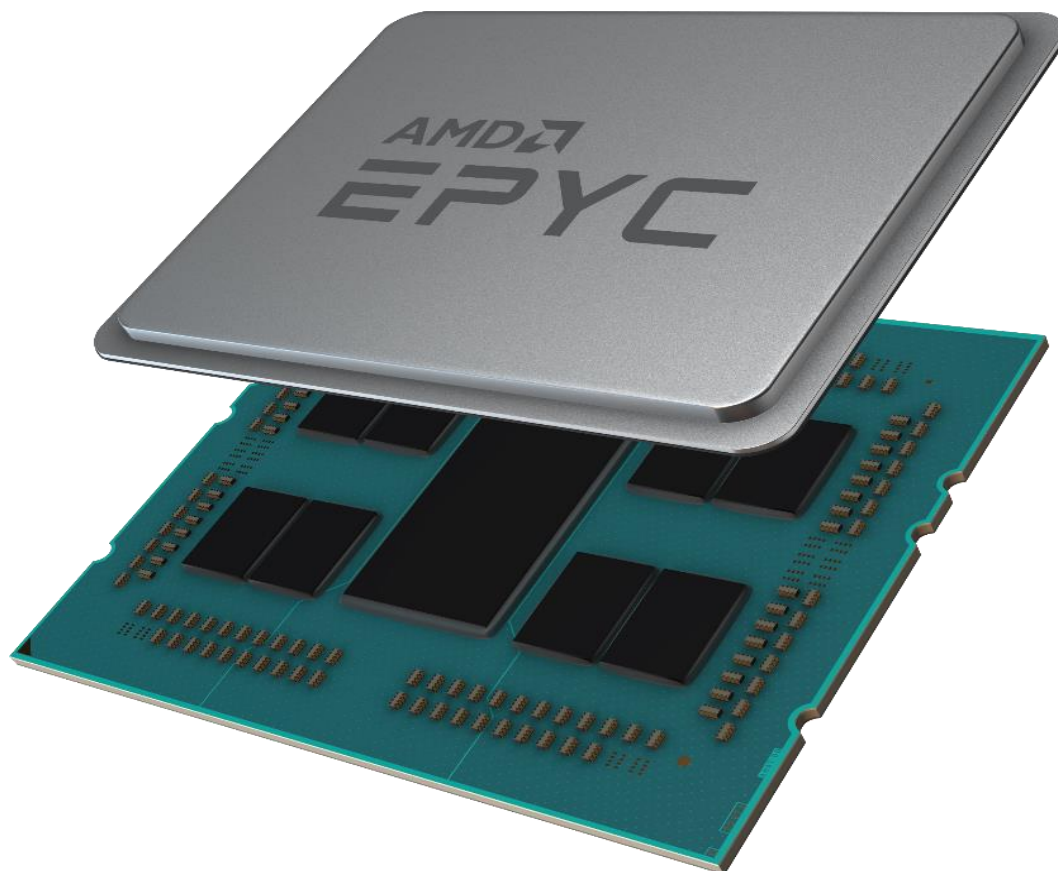


Introduction

The original AMD EPYC processor launched in 2017 as a revolutionary multi-die-on-package processor that allowed AMD to build a high performance and high throughput processor in a more efficient manner. Building off the success of the first generation, AMD has evolved the multi-die strategy, as the company now focused on a different form of multi-die packaging solutions in the second generation EPYC server processor. This new solution, which AMD refers to as hybrid multi-die, allows AMD to not only divide a potentially large die into small interconnected dies, but also specific functions can be fabricated in the most appropriate process node based on required cost and performance. This agile hybrid multi-die architecture decoupled the CPU and cache complex and I/O innovation paths, giving AMD the ability to deliver the best process technology for CPU cores and letting I/O circuitry develop at its own rate. With this new architecture, to date AMD has been able to set over 140 world records across workloads that include Analytics and Big Data; Enterprise; HPC; Virtualization and Cloud Workloads.



AMD EPYC 7002 Series server processor

This paper will look at the changes made from the first generation EPYC processor that improved cache performance and reduced overall memory access time variability.

Quick review of 7001 vs. 7002

The heart of the original EPYC (7001 Series) design was the AMD Infinity Fabric that ties the cores and the multiple die together on the multichip module (MCM) package. Each die had up to eight cores, 2 memory controllers, and PCIe controllers with up to four die per package, for up to 32 1st Gen “Zen” CPU cores per socket (Figure 1). The EPYC 7002 Series doubles the maximum number of CPU cores per socket, increases cache memory and doubling both the Infinity Fabric and PCIe bandwidth to support the additional processing capability. All of these features are efficiently interconnected by using an innovative hybrid multi-die packaging approach shown on the right side of Figure 1.

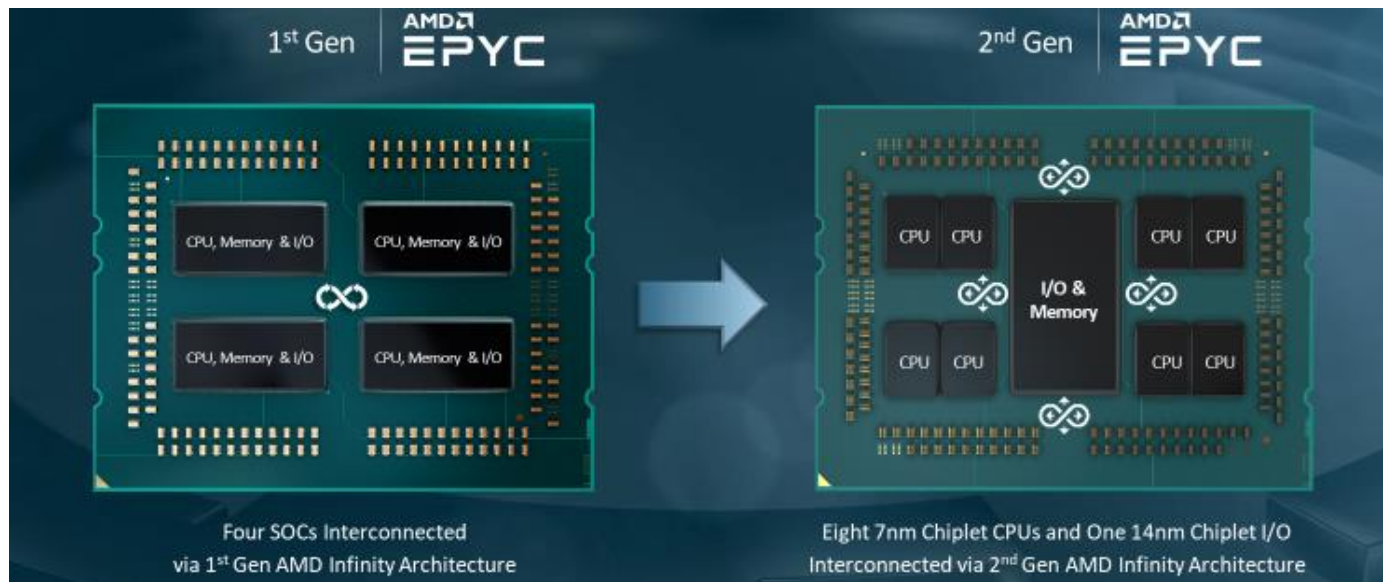


Figure 1. EPYC 2nd Gen. Hybrid Multi-Die package compared with 1st Gen. EPYC

AMD Infinity Fabric remains the bridge between these multi-die configurations and the system architecture is now called the Infinity Architecture. AMD boosted the performance on Infinity Fabric by increasing the aggregate Infinity Fabric bandwidth, which can lower latency for cross-CPU memory fetches (i.e. those fetches that pass through the I/O die). Internally, the hybrid multichip architecture connects the CPUs to memory through the I/O die. The AMD Infinity Architecture lowers average memory latency for single NUMA (non-uniform memory architecture) domains (per socket) and helps applications show better out-of-the-box performance. For those users with specialized workloads, further NUMA optimizations are available (see https://developer.amd.com/wp-content/resources/56745_0.80.pdf).

Feature	EPYC 7001 Series	EPYC 7002 Series
“Zen” Core	1 st Gen	2 nd Gen
Core Count	8 to 32	8 to 64
PCIe Lanes	128	128
PCIe Generation	Gen 3	Gen 4
Process Technology	14nm	7nm (CPU and Cache)
Power	120-180W	120-280W
Max. Memory Capacity	2 TB	4 TB

Table 1. EPYC 7001 Series (“Naples”) vs. EPYC 7002 Series (“Rome”) comparison

EPYC processors and all modern dual-socket processors have a non-uniform memory architecture (NUMA). The EPYC processor Infinity Architecture changed the NUMA profile between the first and the second generations, enabling a single NUMA domain per socket as the default setting reducing average memory access time and delivering excellent “out-of-the-box” performance for most customers.

EPYC Evolves to Hybrid Multi-Die

The hybrid multi-die architecture advanced in the AMD second generation EPYC processor (formerly code named “Rome”). The hybrid multi-die package decouples two main functions: that for the processor and cache cores, and that for the I/O die that supports security and communication outside the processor.

The AMD goal of balanced system performance allows it to choose the number of cores needed and have enough I/O, memory, and memory bandwidth resources to accomplish the system level task. The number and performance of the 2nd generation Zen (“Zen 2”) cores increased from the 1st generation EPYC and required faster Infinity Fabric, memory, and I/O to keep up with the addition computational workload.

Second-generation EPYC does also provide more overall cache per core, speeding the flow of data to the Zen 2 CPU, and higher-performance DDR-3200 memory for lower latency and additional bandwidth to keep all the cores “fed.”

With this architecture AMD uses cutting-edge 7nm technology for the Zen 2 CPU and cache dies (CCD) and improves performance independently of other SoC functions. The I/O functions, including the DRAM controller, PCIe 4.0 controller, and the die to die Infinity Fabric mesh network is located in a single 14nm I/O die. With this architecture, AMD can optimize manufacturing costs and deliver world-class performance at lower costs than a monolithic die. In addition, AMD offers the same memory support, I/O, and security features across the entire stack of EPYC parts. You don’t have to upgrade to get those features.

What is Non-Uniform Memory Access (NUMA)?

A multiprocessing (multi-die) architecture in which each processor is attached to its own local memory (called a NUMA domain) but can also access memory attached to another processor.

It is called “non-uniform” because a memory access to the local memory has lower latency (memory in its NUMA domain) than when it needs to access memory attached to another processor’s NUMA domain.

The advantage of this architecture is that it provides multiprocessor scalability, adds more memory bandwidth with the addition of more processors, and reduces the memory contention for CPUs if they are competing for access across a common bus (shared front-side bus).

NUMA Changes

The first generation EPYC (formerly code named “Naples”) NUMA domains worked for many workloads but were best when optimized for data locality. The tradeoff for using multiple die is that there is additional latency for memory access between the die across the package. The Infinity Fabric connections between die are distributed across the four first generation EPYC (“Zen”) die for balance. Each die had its own memory controller which meant that a DRAM

access (cache miss) might have to connect to a DRAM controller on a different die if the data wasn't specifically directed to in the DRAM attached to the CPU processing it (data locality). Modern operating systems support these NUMA associations and manage data locality and variable latencies, as explained in the sidebar. Dual socket systems also deal with data locality issue of the DRAM attached to the sockets.

The second generation EPYC processor has significant changes to the multi-die design that changes the memory latencies and NUMA locality. The new EPYC processor is designed with a single I/O die connected to up to eight 8-core "CCD" processor dies (see figure 1). With the memory controllers centralized in the I/O die, the processor cores have similar average latency to DRAM when in NPS=1 NUMA mode. While there still can be further optimizations of the DRAM controller location on the I/O die and the specific CCD die, the difference is significantly smaller than the original EPYC.

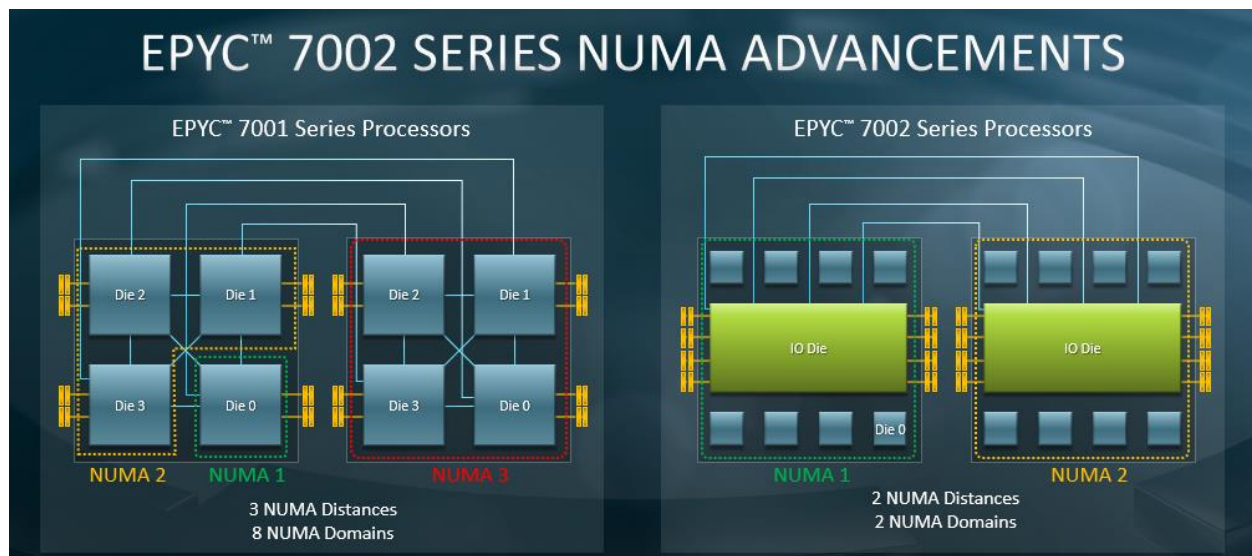


Figure 2. Reduced Number of NUMA Domains and Distances in 2nd Generation EPYC

The result of the new NUMA architecture is that average memory latency per socket out of the box is approximately 19% lower with the second generation EPYC processor (based on AMD internal testing in August 2019). Reducing average latencies make the second generation EPYC easier to deploy.

The 2nd generation EPYC processor offer more NUMA flexibility for specialized workloads. The new processor allows one, two or four NUMA domains per socket. The additional configuration options can be found in the Socket SP3 Platform NUMA Topology technical paper (see https://developer.amd.com/wp-content/resources/56338_1.00_pub.pdf)

Cache Design and a Larger L3

The 2nd generation EPYC processor system can have up to a total 576MB of cache in a two-socket system. Part of the changes from the first generation, AMD reoptimized the level one (L1) instruction cache with almost twice the bandwidth, and the cache is 8-way set associative

(increased from 4-way) for greater efficiency. The new EPYC processor also increased L3 cache size by two-fold in order to handle larger workloads. The L2 cache remains at 512KB. The size of the caches are a balance of die size and performance, but each new process node offers the chance to increase the amount of cache memory without a large die size penalty.

Four Zen 2 cores share an L3 slice in a logical “CCX” module. There are two CCX modules per CCD die as seen in Figure 3. AMD maintains coherency across all L3 cache (and memory) and uses a probe filter to track all the L3 cache slices. When a CPU core looks for data in its shared L3 cache and doesn’t find it (local cache miss), the probe filter will know if it is available in another L3. Depending on the state of data (clean, dirty, modified, etc.) it can move the data to the requesters local L3 or return the data or get the data from memory over the AMD Infinity Fabric connections.

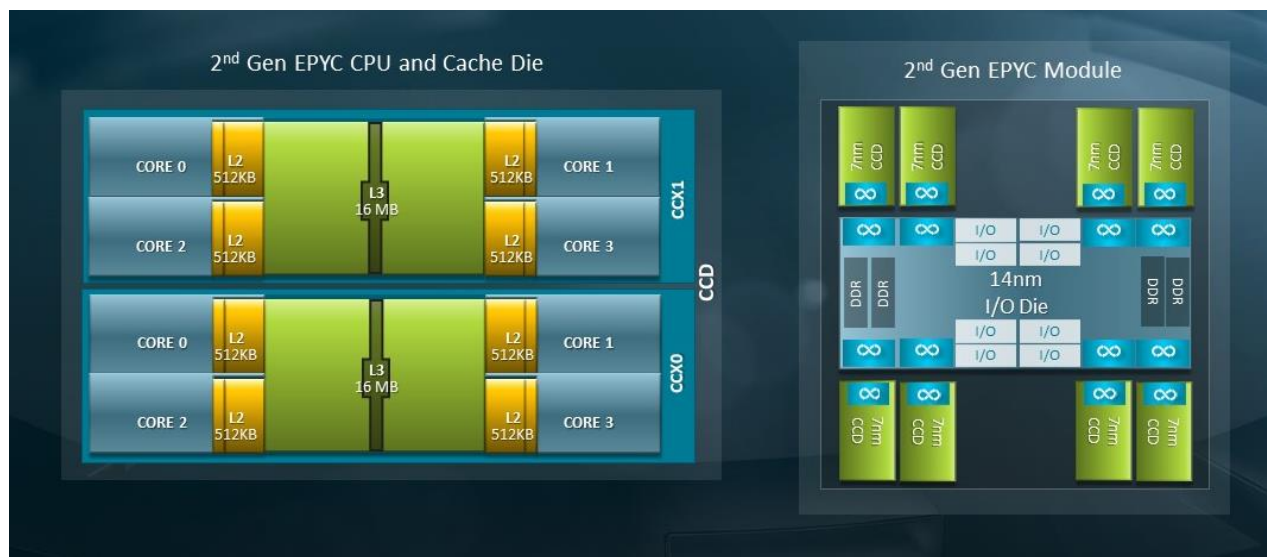


Figure 3 AMD "Rome" CCD and CCX Illustrated

Further optimizations can be achieved if the application is NUMA aware, or if the application can be “pinned” to a CCX by the operating system. These optimizations can be found in the Socket SP3 Platform NUMA Topology technical paper mentioned above

Conclusion

The EPYC processors are uniquely designed for today’s workloads, with larger memory arrays, more cores, and more I/O compared with other single or dual socket x86 platforms. The second generation of EPYC processors has set numerous records and now exhibits an excellent out-of-the-box experience across a variety of workloads. EPYC is good for HPC, Enterprise, SDI, cloud, private cloud, compute edge workloads – just about every workload. AMD has a list of world records to prove it (see <https://www.amd.com/en/processors/epyc-world-records>).

Copyright © 2020 TIRIAS Research. TIRIAS Research reserves all rights herein.

Reproduction in whole or in part is prohibited without prior written and express permission from TIRIAS Research.

The information contained in this report was believed to be reliable when written but is not guaranteed as to its accuracy or completeness.

Product and company names may be trademarks (™) or registered trademarks (®) of their respective holders.

The contents of this report represent the interpretation and analysis of statistics and information that is either generally available to the public or released by responsible agencies or individuals.

AMD, the AMD logo, EPYC, Infinity Fabric, and combinations thereof are registered trademarks of Advanced Micro Devices, Inc.