AMD EPYC 7003 Series Microarchitecture Overview

Publication	57075
Revision	3.0
Issue Date	Mar, 2022

Date	Version	Changes
Mar, 2021	2.0	Initial public release
Mar, 2022	3.0	Added AMD 3D V-Cache information

Audience

This guide provides a high-level technical overview of 3rd Gen AMD EPYC™ 7003 Series Processor internal IP.

Author

Chris Karamatas

Table of Contents

This page intentionally left blank.

Chapter 1 AMD EPYC™ 7003 Series Processors

AMD EPYC™ 7003 Series Processors are the third generation of AMD server-class processors. They feature AMD's new "Zen 3" core and are socket compatible with previous-generation AMD EPYC 7002 Series Processors. Some platforms may support both 2nd Gen and 3rd Gen AMD EPYC processors.

Note: Processor-compatible platforms may require a BIOS update.

1.1 AMD EPYC 7003 Series Processors

All 3rd Gen AMD EPYC processors are based on the same "Zen3" core & IP. This generation of AMD EPYC processors also includes:

- High frequency F-Series models that offer core performance (at a higher TDP) and an increased cache/core ratio.
- High cache X-Series models that feature 3D Vertical stacking (AMD 3D V-Cache) technology that triples the size of the L3 cache.

1.2 General Specifications

AMD EPYC 7003 Series Processors offer a variety of configurations with varying numbers of cores, Thermal Design Points (TDPs), frequencies, L3 cache sizes, etc. that complement AMD's existing server portfolio with further improvements to performance and value, as shown in Table 1-1.

3rd Gen AMD EPYC 7003 Series Processors	
Process technology	7nm
Max Processor speed*	4.1 GHz
Max number of cores	64
Max memory speed	3200 MT/s
Max memory capacity	4 TB per socket
Max L3 cache**	Up to 768MB
Peripheral Component Interconnect	128 lanes (max) PCIe Gen 4
* With AMD EPYC 7003F Series Processors	

Table 1-1: AMD EPYC 7003 Series Processors

1.3 Operating Systems

AMD recommends using the latest available OS version and updates. Please see AMD EPYC™ 7003 Series Processors Minimum Operating System (OS) Versions for detailed OS version information.

1.4 Processor Identification

As described in Figure 1-1:

- **Family:** 19h (identifies the "Zen 3" architecture)
- **Model:** 00-01h (identifies the Server Program)
- **Stepping:** May be used to further identify minor design changes

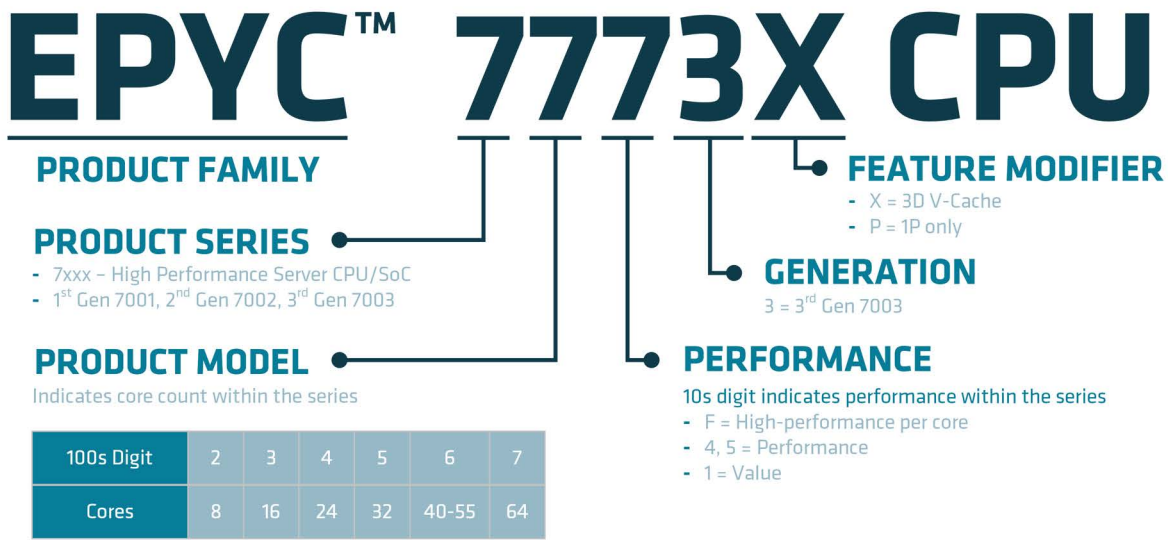

Figure 1-1: AMD EPYC SoC naming convention

Chapter 2

Microarchitectural Overview

AMD EPYC 7003 Series Processors incorporate compute cores, memory controllers, I/O controllers, RAS, and security features into an integrated System On a Chip (SoC). The AMD EPYC 7003 Series Processor retains the proven Multi-Chip Module (MCM) Chiplet architecture of prior successful AMD EPYC processors while making further improvements and upgrading the compute units to new "Zen 3" cores.

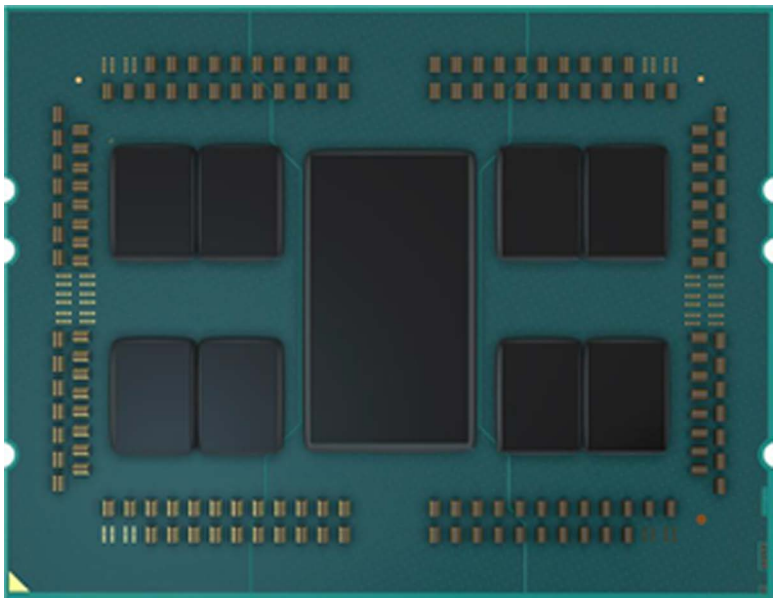

Figure 2-1: AMD EPYC 7003 configuration with 8 Core Complex Dies (CCD) surrounding central I/O Die (IOD)

2.1 "Zen 3" Core

The AMD EPYC 7003 Series Processor is based on new "Zen 3" compute cores. The "Zen 3" core is manufactured using a 7nm process and designed to provide an Instructions per Cycle (IPC) uplift over prior generation "Zen" cores. Each core includes an optimized 32 KB 8-way L1 I-cache and 32 KB 8-way L1 D-cache, as well as a private 512 KB unified (Instruction/Data) L2 cache. All caches use a 64B cache line size.

2.2 Core Complex (CCX) and Core Complex Die (CCD)

Figure 2-2 shows how up to eight "Zen 3" core compute units share a L3 or Last Level Cache (LLC). This grouping is called a Core Complex (CCX). Enabling Simultaneous Multithreading (SMT) on each core allows a single CCX to support up to 16 concurrent hardware threads, up to 4MB of L2 cache, and 32 MB of base L3 cache (or up to 96MB L3 on AMD EPYC 7003 OPNs with 3D V-Cache). AMD EPYC 7003 Series Processors, contain a single CCX within a single Core Complex Die (CCD).

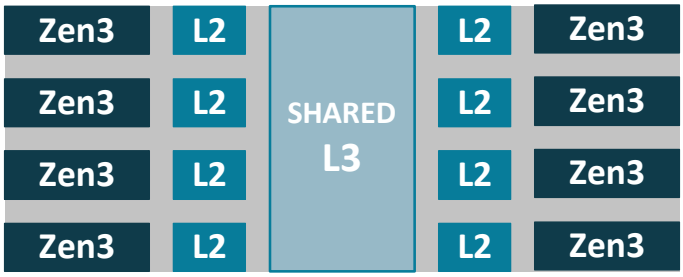

Figure 2-2: Eight Compute Cores sharing an L3 cache within a single Core Complex Die (CCD)

Note: AMD EPYC 7003 Series Processors with AMD 3D V-Cache technology have a 96MB shared central L3 cache.

2.3 AMD 3D V-Cache Technology

Some AMD EPYC 7003 Series Processor Orderable Part Numbers (OPNs) also include 3D V-Cache die stacking technology that enables denser, more efficient chiplet integration. AMD 3D Chiplet architecture stacks L3 cache tiles vertically to provide up to 96MB of L3 cache per CCD (and up to 768 MB L3 Cache per socket) while still providing socket compatibility with all AMD EPYC™ 7003 Series Processor models.

AMD EPYC 7003 Series Processors with AMD 3D V-Cache technology employ industry-leading logic stacking based on copper-to-copper hybrid bonding "bumpless" chip-on-wafer process to enable over 200X the interconnect densities of current 2D technologies (and over 15X the interconnect densities of other 3D technologies using solder bumps), which translates to lower latency, higher bandwidth, and greater power and thermal efficiencies.

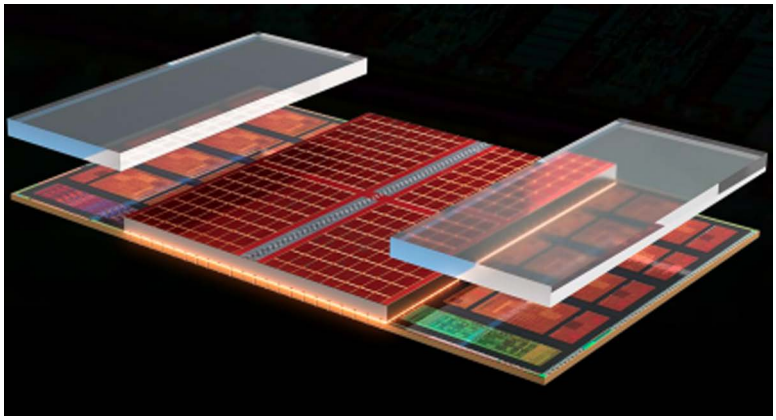

Figure 2-3: Central L3 tiles vertically stacked on the central L3 portion of the base CCX.

2.4 I/O Die (Infinity Fabric™)

The CCDs connect to memory, I/O, and each other through the I/O Die (IOD). Each CCD connects to the IOD via a dedicated high-speed Global Memory Interconnect (GMI) link. The IOD also contains memory channels, PCIe® Gen4 lanes, and Infinity Fabric links. All dies (or chiplets) interconnect with each other via AMD's Infinity Fabric Technology. The fabric clock (FCLK) can now run up to 1600Mhz and thus be coupled with DDR4-3200 Memory DIMMs, which also run at 1600MHz (MEMCLK), further improving memory latency.

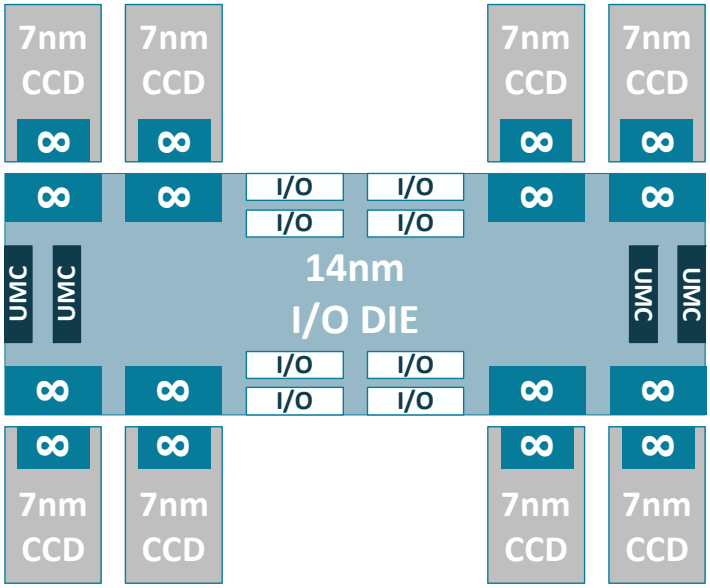

Figure 2-4: AMD EPYC 7003 Series Processor internal CCD, I/O, and Unified Memory Controller (UMC) topology to AMD Infinity Fabric

2.5 Memory and I/O

3rd Gen AMD EPYC processors bring additional performance capabilities and a new 6-way interleave mode to the memory subsystem. Each AMD EPYC 7003 Series Processor has 8 Universal Memory Controllers (UMC). Each UMC (or memory channel) can support up to 2 DIMMs per channel (DPC) for a maximum of sixteen DIMMs per socket. A 3rd Gen AMD EPYC processor can support 4TB of DDR4 memory.

The IOD has the flexibility to supports 4, 6, and 8 memory-channel configurations. 8 memory channels are most common and generally provide the best performance.

Each processor has eight x16-bit I/O links that provide the PCIe subsystem with up to 128 lanes of high-speed PCIe Gen4 I/O in single-socket platforms and up to 160 lanes in dual-socket platforms.

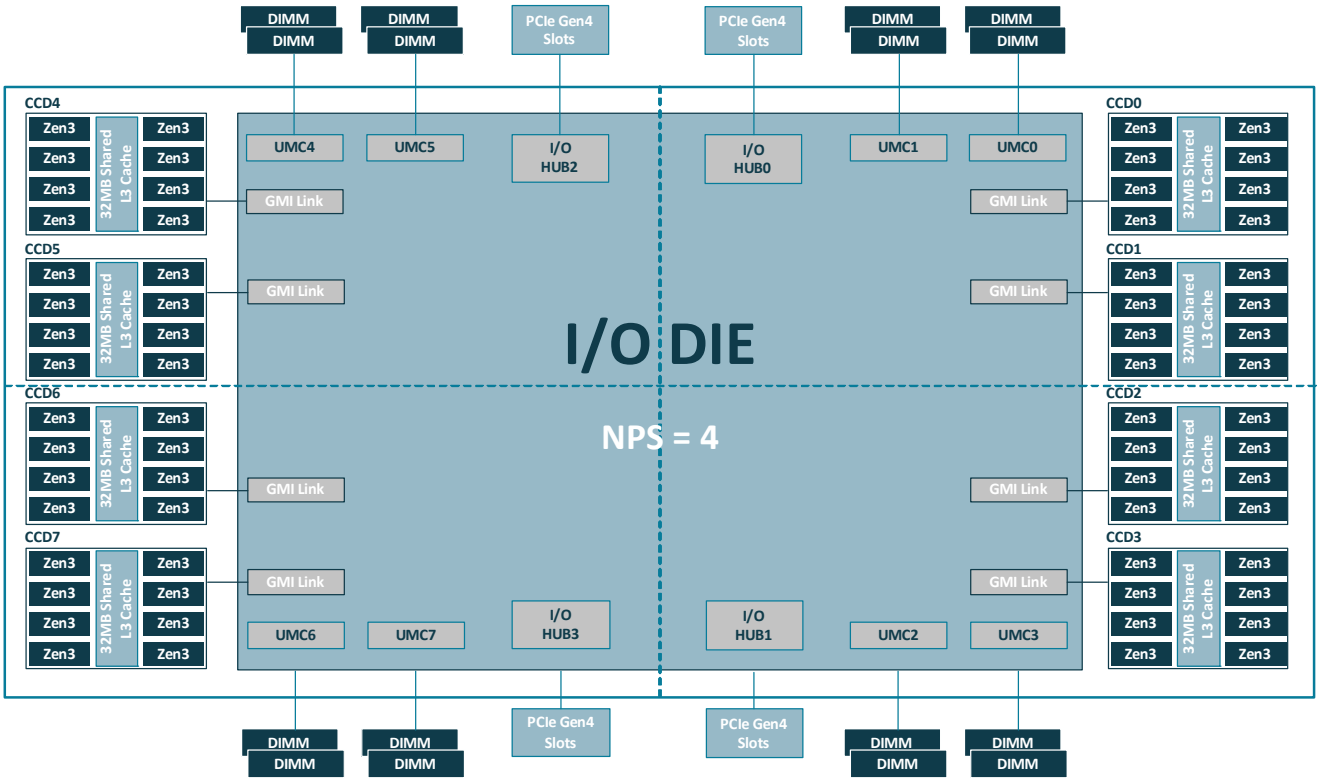

Figure 2-5: EPYC 7003 System on Chip (SoC) consists of 8 CCDs and central IOD

2.6 NUMA Topology

AMD EPYC 7003 Series Processors use a Non-Uniform Memory Access (NUMA) micro-architecture. Furthermore, using system BIOS settings a user can optimize this NUMA topology for their specific operating environment and workload.

Using **NUMA Nodes Per Socket** (NPS) BIOS setting, a system can be setup with different NUMA configurations.

If for example we set NPS=4, as shown in Figure 2-5 above, we can divide the processor into quadrants. Each quadrant would have 2 CCDs, 2 UMCs, and 1 I/O Hub as drawn in Figure 2-5. The closest processor memory-I/O distance is between the cores, memory controllers, and I/O within the same quadrant. The furthest distance is between a core and memory controller or IO hub in diagonal quadrants. Locality of cores, memory, and IO in a NUMA-based system is an important aspect when tuning for performance.

Furthermore, the NPS setting also controls the interleave pattern of the memory controllers. For each NUMA node, all channels within that NUMA node are interleaved. A setting of NPS=4 partitions the processor into four NUMA domains. Each logical quadrant of the processor is configured as its own NUMA domain. Memory is interleaved across the two memory channels in each quadrant. PCIe devices will be local to one of the four NUMA domains on the processor depending on the quadrant of the IO die that has the PCIe root for that device. 5

Meanwhile a setting of NPS=1, indicates a single NUMA node per socket. This setting configures all memory channels on the processor into a single NUMA domain, i.e., all the cores on the processor, all attached memory and all PCIe devices connected to the SoC are in one NUMA domain. Memory accesses are interleaved across all eight memory channels into a single address space.

As the NPS setting gets more granular, the number of channels interleaved decreases accordingly. A setting of NPS=2 configures 2 NUMA domains per socket, which interleaves corresponding four memory channels within the same 4 CCD NUMA domain. Half the cores and half the memory channels of each SoC are grouped together into one NUMA domain, and the remaining cores and memory channels are grouped into a second domain. Memory is interleaved across the four memory channels in each NUMA domain.

Additionally, certain environments performance may be further improved by associating workloads to compute cores which all share a single LLC. The LLC (Last Level Cache, or L3 cache) as NUMA BIOS Setting exposes this ability. Enabling this setting equates each CCD to a separate NUMA domain, as a unique L3 cache per CCD. A single 7003 processor with 8 CCDs would have 8 NUMA nodes.

In conclusion, a single EPYC 7003 Series Processor may support configurations ranging from a single NUMA node system, all the way up to up to 8 NUMA nodes per socket..

2.7 Dual-Socket Configurations

AMD EPYC 7003 Series Processors support single- or dual-socket system configurations. Processors with a ‘P’ suffix in their name are designed for one socket configurations. In two-socket configurations, both processors must be identical. Two different processor OPNs, or steppings, cannot be used in the same 2-socket system. See “Resources” on page 9 for additional information on the processor naming conventions.

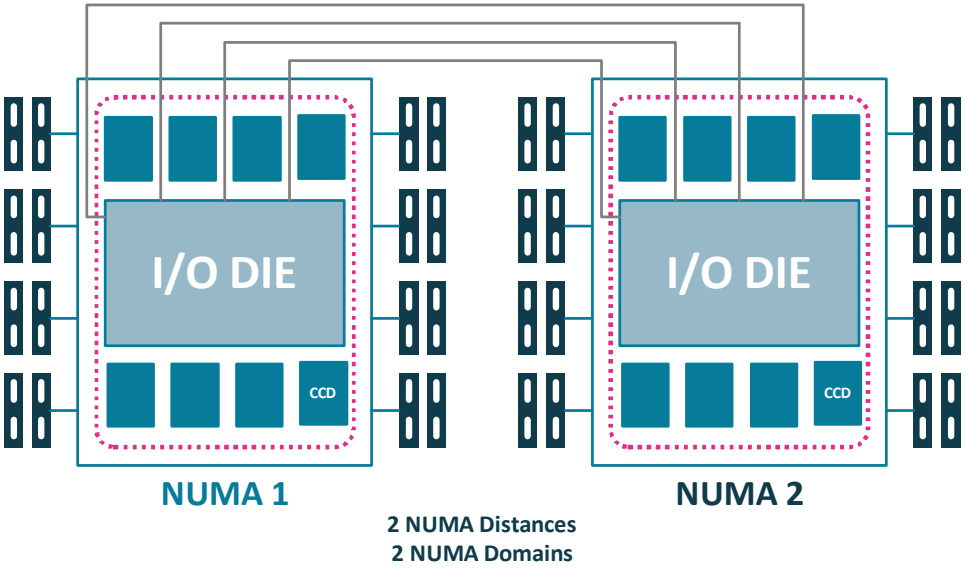

Figure 2-6: Two EPYC 7003 Processors connect through 4 xGMI links (NPS1)

In dual-socket systems, two identical EPYC 7003 series SoCs are connected via their corresponding Infinity Fabric, or External Global Memory Interconnect (xGMI), links. This creates a high bandwidth, low latency interconnect between the two processors. System manufacturers can use either 3 or 4 Infinity Fabric Links depending upon I/O and bandwidth system design objectives.

The Infinity Fabric links utilize the same physical connections as the PCIe lanes on the system. Each Link uses up to 16 PCIe lanes. In a typical dual socket system 64 PCIe lanes (4 links) from each socket will be reconfigured for Infinity Fabric connections therefore each socket still has 64 PCIe lanes remaining, the system still has a total of 128 PCIe lanes. In some cases, a system designer may want to expose more PCIe lanes for the system, by reducing the number of Infinity Fabric Links to from 4 to 3. In these cases, a system designer may allocate up to 160 lanes for PCIe (80 per socket), utilizing only 48 PCIe lanes per socket for Infinity Fabric links, instead of 64.

A dual-socket system has a total of 16 memory channels, or 8 per socket. Dual-socket systems can be configured in a variety of ways, including with 1, 2, 4, 8, or 16 NUMA domains.

Chapter 3 Appendix

3.1 New Software-Visible Features

The introduction of the "Zen3" core makes new ISA features available in AMD EPYC 7003 Series Processors, such as improved performance, security features, and RAS. Some of the new ISA features include:

- Shadow Stack CET (control-flow enforcement) technology
- AVX2: VAES & VPCLMULQDQ w/256-bit support
- Broadcast TLB invalidation (INVLPGB & TLBSYNC Instructions)
- Fast short rep movs
- Predictive Store Forward Disable
- Secure Nested Paging
- Process Control ID
- Memory Protection Keys for Users

Not all operating systems or hypervisors support all of these features. Please see the latest version of the *AMD64 Architecture Programmer's Manuals* or *Processor Programming Reference (PPR) for AMD Family 19h*.

3.2 Resources

- AMD EPYC™ Tech Docs and White Papers
- *Workload Tuning Guide for AMD EPYC™ 7003 Series Processors* (available from AMD EPYC Tuning Guides)
- Memory Population Guidelines for AMD EPYC 7003 Series Processors - Login required.
- Socket SP3 Platform NUMA Topology for AMD Family 19h Models 00h–0Fh - Login required.

This page intentionally left blank.