WRF® AND AMD 3D V-CACHE™ TECHNOLOGY
WEATHER FORECASTING

Powered by 4th Gen AMD EPYC™ Processors with AMD 3D V-Cache™ technology

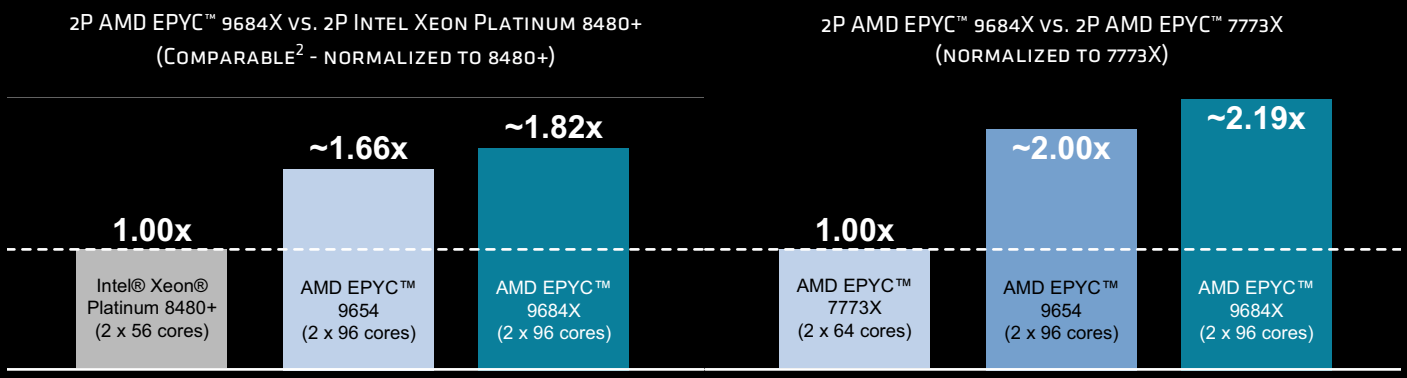
June 2023

AT A GLANCE

2P systems powered by 96-core 4th Gen AMD EPYC™ processors with AMD 3D V-Cache™ technology demonstrate outstanding high core count competitive and generational performance uplifts on WRF®.

PERFORMANCE HIGHLIGHTS

A single 2P 96-core 4th Gen AMD EPYC 9684X with AMD 3D V-Cache technology powered system demonstrates a ~1.82x uplift vs. a single 2P Intel® Xeon® Platinum 8480+ system on the conus-2.5km WRF workload. The 2P 96-core 9684X-powered system also shows a ~2.19x generational uplift vs. a 2P 64-core 3rd Gen AMD EPYC 7773X system.



KEY TAKEAWAYS

A 2P server powered by 4th Gen AMD EPYC 9684X (96-core) processors delivered a composite average WRF competitive performance uplift of ~1.82x compared to a 2P server powered by a Intel Xeon Platinum 8480+ (56-core) processors. The 2P 4th Gen AMD EPYC 9684X system also delivered a composite average WRF generational performance uplift of ~2.19x compared to a 2P server powered by 3rd Gen AMD EPYC 7773X (64-core) processors.

4th Gen AMD EPYC 9004 processors with AMD 3D V-Cache technology are available in 1P and 2P configurations and feature:

- Up to 1,152MB L3 cache vs. 384MB in standard 4th Gen AMD EPYC processors.
- Up to 4 links of Gen 3 Infinity Fabric™ at up to 32 Gbps.
- 12 memory channels that support up to 6TB of DDR5-4800 memory.
- Support for PCIe® Gen 5 at up to 32 Gbps.
- AVX-512 instruction support for enhanced HPC and ML performance.
- AMD Infinity Guard technology to defend your data.[3]

IN THIS BRIEF

SCALABILITY

The significant single-node performance advantage of AMD EPYC 9004 Series Processors with AMD 3D V-Cache technology becomes even more pronounced when these are deployed against realistic workloads in a multi-node technical computing context. Adding more computational nodes to a technical computing cluster reduces the portion of the dataset being processed by each node. Enough reduction allows each portion of the dataset to fit entirely within the L3 cache in each compute node, which causes a sudden performance boost called *superlinear scaling*. This behavior is not unusual for processors, but the industry-leading 1152MB size of the L3 cache in AMD EPYC 9004 Series Processors with AMD 3D V-Cache technology (3x the 384MB of the standard EPYC 9004 series L3 cache) show excellent scalability, including superlinear scaling.[6] AMD testing on an 8-node cluster shows the AMD EPYC 9684X delivering ~9 nodes worth of performance on WRF.

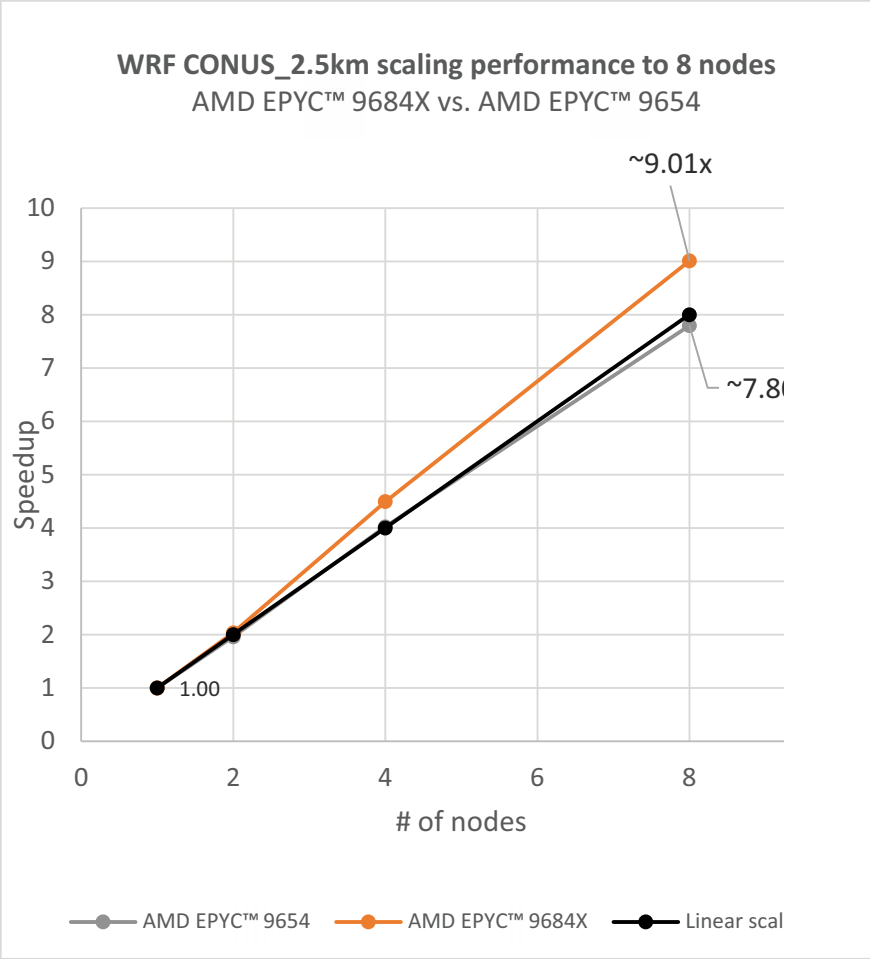


Figure 1: WRF 8-node scaling performance with AMD EPYC 9864X processors with AMD 3D V-Cache technology

AMD 3D V-CACHE™ TECHNOLOGY

Some AMD EPYC 9004 Series Processors include AMD 3D V-Cache™ die stacking technology that enables more efficient chiplet integration. AMD 3D Chiplet architecture stacks L3 cache tiles vertically to provide up to 96MB of L3 cache per die (and up to 1152 MB L3 Cache per socket) while still providing socket compatibility with standard AMD EPYC™ 9004 Series Processor models.

AMD EPYC 9004 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),[4,5] which translates to lower latency, higher bandwidth, and greater power and thermal efficiencies.

SYSTEM CONFIGURATION

AMD SYSTEM CONFIGURATION			
CPUs	2x AMD EPYC 7773X	2x AMD EPYC 9654	2 x AMD EPYC 9684X
Frequency: Base | Boost[7]	2.20 GHz | 3.50 GHz (up to)	2.40 GHz | 3.55 GHz (up to)	2.55 GHz | 3.70 GHz (up to)
Cores	64 cores/socket (128 threads)	96 cores/socket (192 threads)	96 cores/socket (192 threads)
L3 Cache	768 MB per CPU	384 MB per CPU	1152 MB per CPU
Memory	1.0 TB (16x) Dual-Rank DDR4 3200 64 GB DIMMs 1 DPC	1.5 TB (24x) Dual-Rank DDR5 4800 64 GB DIMMs 1 DPC	
NIC	25 Gb Ethernet CCX512-A ConnectX-5 (fw 16.35.2000)		
InfiniBand	200 Gb HDR ConnectX-6 VPI (fw 20.35.2000)		
Storage: OS | Data	Samsung MZQL21T9HCJR-00A07 1.92 TB		
BIOS Version	1009B	1007D	
BIOS Settings	SMT=OFF; NPS=4; Determinism=Power		
OS	RHEL 8.7 (kernel 4.18.0-425.3.1.el8.x86_64)		
OS Settings	amd_iommu=ON; iommu=pt; mitigations=off; clear caches; NUMA balancing=0; THP=on; CPU governor=Performance; C2 states=disabled		

Table 1: AMD system configurations

INTEL SYSTEM CONFIGURATION	
CPUs	2x Intel Xeon Platinum 8480+
Frequency: Base | Boost	2.00 GHz | 3.80 GHz (up to)
Cores	56 cores per socket (112 threads)
L3 Cache	105 MB per CPU
Memory	1.0 TB (16x) Dual-Rank DDR5 4800 64 GB DIMMs 2 DPC
NIC	25 Gb Ethernet CCX512-A ConnectX-5 (fw 16.35.2000)
InfiniBand	200 Gb HDR ConnectX-6 VPI (fw 20.35.2000)
Storage: OS | Data	Samsung MZQL21T9HCJR-00A07 1.92 TB
BIOS Version	ESE110Q-1.10
BIOS Settings	Hyperthreading=Off, Profile = Maximum Performance
OS	RHEL 8.7 (kernel 4.18.0-425.3.1.el8.x86_64)
OS Settings	processor.max_cstate=1; intel_idle.max_cstate=0; iommu=pt mitigations=off; clear caches; NUMA Balancing=0; randomize_va_space 0; THP=ON; CPU Governor=Performance

Table 2: Intel system configurations

TEST METHODOLOGY

WRF (Weather Research & Forecasting Model) provides a standard set of benchmarks that evaluate the performance of different platforms running WRF applications. These benchmark cases represent typical usage and cover a range of sizes. The uplift is calculated as the ratio of the systems under test (*sut*) to the reference systems (*ref*). In this Performance Brief, the Intel Xeon Platinum 8480+ processors were the *ref* system, and 4th Gen AMD EPYC processors were the *sut* for the competitive uplift. For the generational uplift, the 4th Gen AMD EPYC 9684X and 9654 systems were the *sut* and the 3rd Gen AMD EPYC 7773X system was the *ref*. The 4th Generation AMD EPYC processors tested include both standard AMD EPYC 9004 and AMD EPYC 9004X models with AMD 3D V-Cache technology. The total amount of variability between individual runs was <1%. The results presented in this Performance Brief are the average per-processor uplifts running the conus2.5km benchmarks. The systems tested were configured as shown in Tables 1 and 2, above.

FOR ADDITIONAL INFORMATION

Please see the following additional resources for more information about 4th Gen AMD EPYC features, architecture, and available models:

- AMD EPYC™ 9004 Series Processors
- AMD EPYC™ Tuning Guides
- AMD EPYC™ Products

REFERENCES

1. "Technical Computing" or "Technical Computing Workloads" as defined by AMD can include: electronic design automation, computational fluid dynamics, finite element analysis, seismic tomography, weather forecasting, quantum mechanics, climate research, molecular modeling, or similar workloads. GD-204
2. The Intel Xeon Platinum 8480+ is the highest-performing processor listed in the Performance General-Purpose category for 4th Gen Intel Xeon CPU models at https://download.intel.com/newsroom/2023/data-center-hpc/Intel-4th-Gen-Xeon_product_SKUs.jpg.
3. 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. GD-183
4. Based on calculated areal density and based on bump pitch between AMD hybrid bond AMD 3D V-Cache stacked technology ccompared to AMD 2D chiplet technology and Intel 3D stacked micro-bump technology. EPYC-026
5. Based on AMD internal simulations and published Intel data on "Foveros" technology specifications. EPYC-027
6. AMD defines "linear scaling" as an equal and proportionate application performance uplift relative to single node performance; that is, when scaling out to 2 nodes results in 2x the performance of a single node, scaling out to 4 nodes results in 4x the performance of a single node, and so forth. "Super-linear" scaling is when the performance uplift achieved by adding one or more node(s) is greater than linear. AMD allows a +/- of 2% margin of error when claiming linear or super linear scaling. GD-205
7. Maximum boost for AMD EPYC processors is the maximum frequency achievable by any single core on the processor under normal operating conditions for server systems. EPYC-18

AUTHORS

Alvaro Fernandez and Ashok Manikonda contributed to this Performance Brief.

RELATED LINKS

- WRF*
- AMD EPYC™ Processors
- AMD EPYC Technical Briefs

**Links to third party sites are provided for convenience and unless explicitly stated, AMD is not responsible for the contents of such linked sites and no endorsement is implied.*

AMD EPYC 9004 FOR WEATHER FORECASTING

4th Gen AMD EPYC CPUs deliver excellent per-core performance by taking advantage of fast CPU frequencies, low latency memory, and a unified cache structure. Meterologists can use AMD EPYC processors to perform complex weather forecasting tasks with ground-breaking high-performance computing and robust security features to deliver excellent results.

"ZEN 4" CORE & SECURITY FEATURES

Support for up to:

- 96 physical cores, 192 threads
- 1152 MB of L3 cache per CPU
- 96 MB of L3 cache per CCD
- 6 TB of DDR5-4800 memory
- Up to 128 1P, up to 160 2P PCIe® Gen 5 lanes

Infinity Guard security features[3]

- Secure Boot
- Encrypted memory with SME

WRF®

Developed and maintained by the National Center for Atmospheric Research (NCAR), the Weather Research & Forecasting (WRF) model has over 48,000 registered users in over 160 countries. WRF is a flexible and computationally efficient platform for operational forecasting across scales ranging from meters to thousands of kilometers.