LEADERSHIP XGBOOST PERFORMANCE
OUTPERFORMING 5TH GEN INTEL® XEON®

Powered by 4th Gen AMD EPYC™ Processors

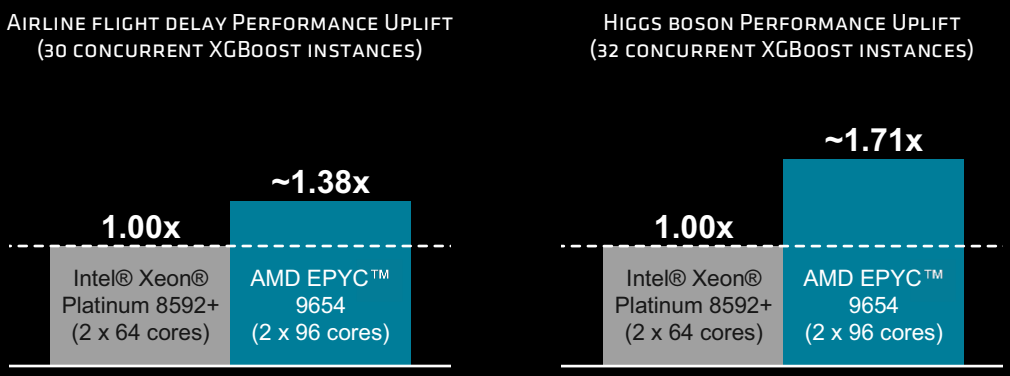
May 2024

AT A GLANCE

A dual-socket 4th Gen AMD EPYC™ 9654 system delivers ~1.38x airline flight delay prediction dataset and ~1.71x Higgs boson explosion dataset uplifts running XGBoost workloads versus a dual-socket 5th Gen Intel® Xeon® Platinum 8592+ system.

PERFORMANCE HIGHLIGHTS

A dual-socket 96-core AMD EPYC 9654 system shows ~1.38x and ~1.71x throughput performance uplifts in runs/hour on airline flight delay and Higgs boson datasets, respectively, versus a dual-socket Intel Xeon Platinum 8592+ system running multiple XGBoost workload instances utilizing all system cores.



KEY TAKEAWAYS

XGBoost (eXtreme Gradient Boosting) is a scalable high-performance learning method for machine learning training. XGBoost is an "ensemble" learning method that creates strong predictions from multiple weak predictions. It can handle large datasets and supports parallel processing that enables rapid training. It also gracefully handles missing values, which allows processing incomplete real-world data without needing significant pre-processing. This performance and versatility make XGBoost a good choice for a variety of applications. The examples in this performance brief predict the existence of Higgs bosons and whether a given airline flight will be delayed. The ~1.38x and ~1.71x uplifts delivered by the 2P 4th Gen AMD EPYC 9654 system showcase the ongoing performance leadership of AMD EPYC processors versus the latest competing processors.

General purpose 4th Gen AMD EPYC 9004 processors are available in 1P and 2P configurations and feature:

- Up to 96 cores (192 threads) per processor.
- Up to 384 MB L3 cache.
- Up to 4 links of Gen 3 Infinity Fabric™ at up to 32 Gbps.
- 12 memory channels per socket that support up to 6 TB of DDR5-4800 memory.
- Support for 128 (1P) and up to 160 (2P) PCIe® Gen 5 with up to 32 Gbps bandwidth.
- AVX-512 instruction support for enhanced HPC and ML performance.
- AMD Infinity Guard technology to defend your data.[1]

IN THIS BRIEF

AMD EPYC 9004 SERIES PROCESSORS

General purpose AMD EPYC 9004 Series Processors continue to redefine the standards for modern data centers. 4th Gen AMD EPYC processors are built on the innovative x86 architecture and "Zen 4" core. 4th Gen AMD EPYC processors deliver efficient, optimized performance by combining high frequencies, the largest-available L3 cache, 128 lanes of PCIe® 5 I/O (1P) and up to 160 lanes (2P), and synchronized fabric and memory clock speeds, plus support for up to 6 TB of DDR5-4800 memory. Built-in security features, such as Secure Memory Encryption (SME) and Secure Encrypted Virtualization (SEV-SNP), collectively known as AMD Infinity Guard, help protect data while it is in use.[1]

TEST METHODOLOGY

AMD tested the following two datasets:

- **Airline:** This training data contains airline logistics information, including whether a flight was delayed and, if so, by how much. The XGBoost model was trained on this data to predict whether a given flight will be delayed. You can obtain the data used for this test from https://s3.amazonaws.com/benchm-ml--main/train-10m.csv* (training; ~477.3 MB in size) and https://s3.amazonaws.com/benchm-ml--main/test.csv* (testing; ~4.8 MB in size).

- **Higgs boson:** This training data includes characteristics of subatomic particles and is used to train the XGBoost model to predict whether or not a given particle is a Higgs boson. You can obtain the data used for this test from the University of California at Irvine (https://archive.ics.uci.edu/ml/machine-learning-databases/00280/HIGGS.csv.gz*) and is ~1.2 GB in size (compressed).

The systems described in Tables 3 and 4 were configured to run XGBoost workloads on the airline and Higgs boson datasets described above. Tables 1 and 2 show the number of workload instances, instance pinning, instance sizing, and any idle system cores. The maximum number of parallel instances that could be run is determined by memory bandwidth capacity and optimal core utilization.

# OF CONCURRENT AIRLINE INSTANCES	INSTANCE PINNING	INSTANCE SIZE	# OF IDLE CORES PER SOCKET
16	• **Instance 1-8:** P0/mem0 • **Instance 9-16:** P1/mem1	• **AMD system:** 12 cores • **Intel system:** 8 cores	N/A

Table 1: Airline test configurations used for this performance brief

# OF CONCURRENT HIGGS INSTANCES	INSTANCE PINNING	INSTANCE SIZE	# OF IDLE CORES PER SOCKET
32	• **Instance 1-15:** P0/mem0 • **Instance 16-30:** P1/mem1	• **AMD system:** 6 cores • **Intel system:** 4 cores	N/A

Table 2: Higgs boson test configurations for this performance brief

SYSTEM CONFIGURATION

AMD NODE CONFIGURATION	
CPUs	2 x AMD EPYC 9654
Frequency: Base | Boost[2]	2.4 GHz | 3.7 GHz (up to)
Cores	96 cores/socket (192 threads) - 1 NUMA domain/socket
L3 Cache	384 MB per CPU
Memory	1.5TB (24 x Dual-Rank DDR5-4800 64GB DIMMs, 1DPC)
Storage	1.92 TB Samsung MZQL21T9HCJR-00A07 NVMe®
BIOS Version	1006C
BIOS Settings	SMT=OFF, NPS=1, Determinism=Power
OS	Ubuntu® 22.04.3 LTS

Table 3: AMD system configuration

INTEL NODE CONFIGURATION	
CPUs	2 x Intel Xeon Platinum 8592+
Frequency: Base | Boost[2]	1.9 GHz | 3.9 GHz
Cores	64 cores/socket (128 threads) - 1 NUMA domain/socket
L3 Cache	320 MB per CPU
Memory	1024 GB (16 x Dual-Rank DDR5-4800 64GB DIMMs, 1DPC)
Storage	3.2 TB SSDPF2KE032T10
BIOS Version	1.4.4
BIOS Settings	Hyperthreading=Disabled, Profile=Maximum Performance
OS	Ubuntu® 22.04.3 LTS

Table 4: Intel system configuration

CONCLUSION

Intel® released their 5th generation of Intel Xeon® Scalable processors (codenamed "Emerald Rapids") in December of 2023. Since 2017, AMD has released four generations of AMD EPYC™ processors. The uplifts shown in this performance brief demonstrate that the relentless focus by AMD on data center performance, power efficiency, and total cost of ownership continues to surpass the competition. The thriving AMD EPYC ecosystem includes over 250 distinct server designs and more than 800 unique cloud instances. AMD EPYC processors hold over 300 world records for performance across a wide spectrum of workloads, and Intel's latest release does not change this leadership position.

FOR ADDITIONAL INFORMATION

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

- AMD EPYC™ 9004 Series Processors
- AMD Documentation Hub

REFERENCES

1. 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
2. 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

AUTHOR

Rema Hariharan

RELATED LINKS

- AMD EPYC Processors
- AMD Documentation Hub

**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.*

SUPERB DATA ANALYTICS PERFORMANCE

Enterprises of all sizes rely on ever-growing datasets to query and analyze data to derive mission-critical business insights that support key decisions. Systems powered 4th Gen AMD EPYC™ 9xx4 processors deliver superb data analytics performance across even the most demanding workloads and dataset.

"ZEN 4" CORE & SECURITY FEATURES

General-purpose support for up to:

- 96 physical cores, 192 threads
- 384 MB of L3 cache per CPU
- 32 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[1]

- Secure Boot

Encrypted memory with SME