

4TH GEN AMD EPYC™ PROCESSORS OFFER LEADERSHIP MAXLINEAR COMPRESSION PERFORMANCE

AMD
together we advance_data center computing

*First Edition
August 2024*

Luke Chang



4TH GEN AMD EPYC™ PROCESSORS OFFER LEADERSHIP MAXLINEAR COMPRESSION PERFORMANCE

CONTENTS

INTRODUCTION - - - - -	3
About AMD EPYC 9004 Series Processors	3
About MaxLinear Panther III	3
SOLUTION OVERVIEW - - - - -	4
TEST METHODOLOGY AND RESULTS - - - - -	5
Performance	5
POWER CONSIDERATIONS - - - - -	7
SOFTWARE - - - - -	7
RELIABILITY, AVAILABILITY AND SERVICEABILITY (RAS) - - - - -	8
CONCLUSION - - - - -	8
REFERENCES - - - - -	8

INTRODUCTION

The ongoing migration to cloud computing coincides with exponential increases in the amount of data being generated, transferred, and processed. Modern host processors offer high core counts, fast memory bandwidth, and enhanced I/O connectivity. Even so, offloading some data security, compression, and data protection functions from the CPU can enhance platform performance while possibly reducing power consumption.

AMD EPYC™ processors deliver [world record](#) performance and power efficiency. [MaxLinear Panther III](#)* accelerators offer excellent data reduction and protection. This white paper showcases how combining 4th Gen AMD EPYC processors with MaxLinear Panther III accelerators that offload data compression/decompression along with encryption/decryption, duplicative enablement, cryptographic hash generation/authentication, and data protection using NVME PI generation/verification tasks from the processor(s) deliver exceptional compute, compression, and data defense performance to the datacenter, enterprise, and embedded markets.

ABOUT AMD EPYC 9004 SERIES PROCESSORS

AMD EPYC 9004 Series Processors continue to redefine the standards for modern datacenters. 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, up to 128 lanes (1P) or up to 160 lanes (2P) of PCIe® 5 I/O, 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

ABOUT MAXLINEAR PANTHER III

The MaxLinear Panther III accelerator offloads data reduction, deduplication, encryption/decryption, and data protection from the host processor(s). It delivers ultra-low latency, 200Gbps throughput, data reduction, data protection, and security acceleration for enterprise and hyperscale data centers in either a PCIe® Gen 4 x16 half-height or OCP 3.0 form factor.

MaxLinear states that the MaxLinear Panther III achieves breakthrough 12:1 data reduction performance and protection that

allows storage solutions to store 1/12 the data volume compared to the same amount of uncompressed data. The 200 Gbps throughput and ultra-low latency enable data access at rates that can enhance the speed and accuracy of big data workloads. These accelerators also support the exclusive MaxHash™ algorithm for independent hash block sizes and programmable offsets that enhance deduplication hit rates and optimize storage costs.

The feature-rich SDK contains APIs, drivers, and source code for easy incorporation with end application software and software-defined storage (SDS). All of this is available with real-time validation to give you “six 9s” reliability that safeguards data integrity and guards against data loss. Please read the [datasheet](#)* to learn more.

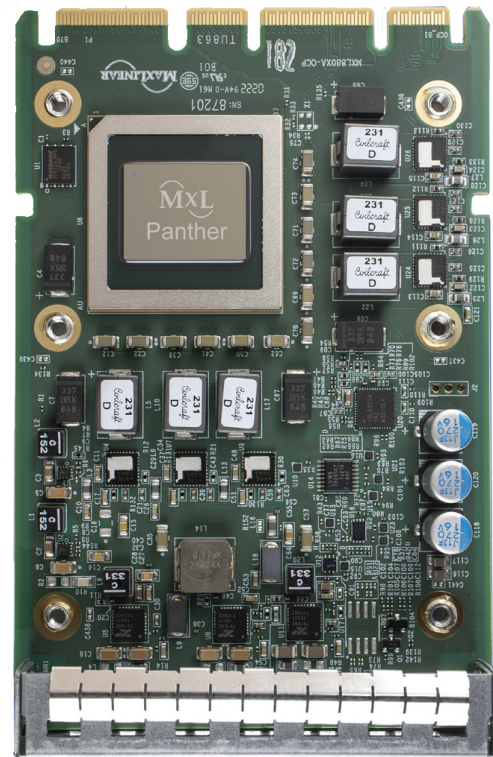


Figure 1: MaxLinear Panther III accelerator

SOLUTION OVERVIEW

The testing described in this white paper was conducted on systems configured as shown in Tables 1 and 2.

Item	Description
CPU	1 x AMD EPYC™ 9654
Frequency Base Boost ²	2.4 GHz 3.7 GHz
Cores Threads	96 192
L3 Cache	384MB per socket
Memory	384 GB (12 x 32 GiB DDR5 4800) 1 DPC
NIC	BRCM BCM5720 Gigabit NIC
Storage	2 x 500 MB SSD
Accelerator	MaxLinear Panther III
BIOS	RRR100AB
BIOS Options	default
OS	Ubuntu® 22.04
Kernel	6.5.0.-26-generic
OS Options	default
Software	MaxLinear SDK v1.0.1

Table 1: AMD system configuration

CPU	2 x Intel® Xeon® Platinum 8474C
Frequency Base Boost*	2.1 GHz 3.1 GHz
Cores Threads	48 96
L3 Cache	97 MB
Memory	512 GB (8 x 64 GB DDR5 4800) 1 DPC
NIC	Intel X710 10G
Storage	1 TB SATA
Accelerator	Intel QuickAssist Technology (Intel QAT)
BIOS	ESE114G2.10 3-16-2023
BIOS Options	default
OS	Ubuntu® 22.04 LTS
Kernel	5.15.0-101-generic
OS Options	intel_iommu=on, iommu=pt, transparent hugepage=never hugepagesz=1G, hugepages=32, hugepagesz=2M, hugepages=1024 default_hugepagesz=2M
Software	QAT 20.L.1.0.50-00003 driver DPDK version 23.11.0

Table 2: Intel system configuration

Testing used [Silesia Corpus](#)*, which is a set of files with differing characteristics that can be used to test compression algorithms.

TEST METHODOLOGY AND RESULTS

AMD tested the systems described above to determine the performance of the MaxLinear Panther III accelerator relative to Intel QAT. The following tests were performed:

- 108 QAT runs were performed on the Intel systems. These tests covered 12 Silesia Corpus data files across 9 deflate levels (Levels 1-9). The L9 (highest compression level) results were then used for study and comparison.
- 48 MaxLinear runs were performed on the AMD system. These tests covered 12 Silesia Corpus data files across 4 levels of deflate (Levels 1, 3, 6, and 9). The L9 (highest compression level) results were then used for study and comparison.

The test output indicated the compression ratio achieved for each of the 12 Silesia Corpus data files and the throughput in gigabits per second. AMD collected the results for each run on each component

and then calculated the average compression ratio and throughput results for each data file. The uplifts were calculated using geomeans normalized such that the Intel QAT system performance will always equal 1.00x for both the compression ratio tests and the throughput.

PERFORMANCE

Figure 2 shows the relative throughput of both systems performing Level 9 compression on each Silesia Corpus component.

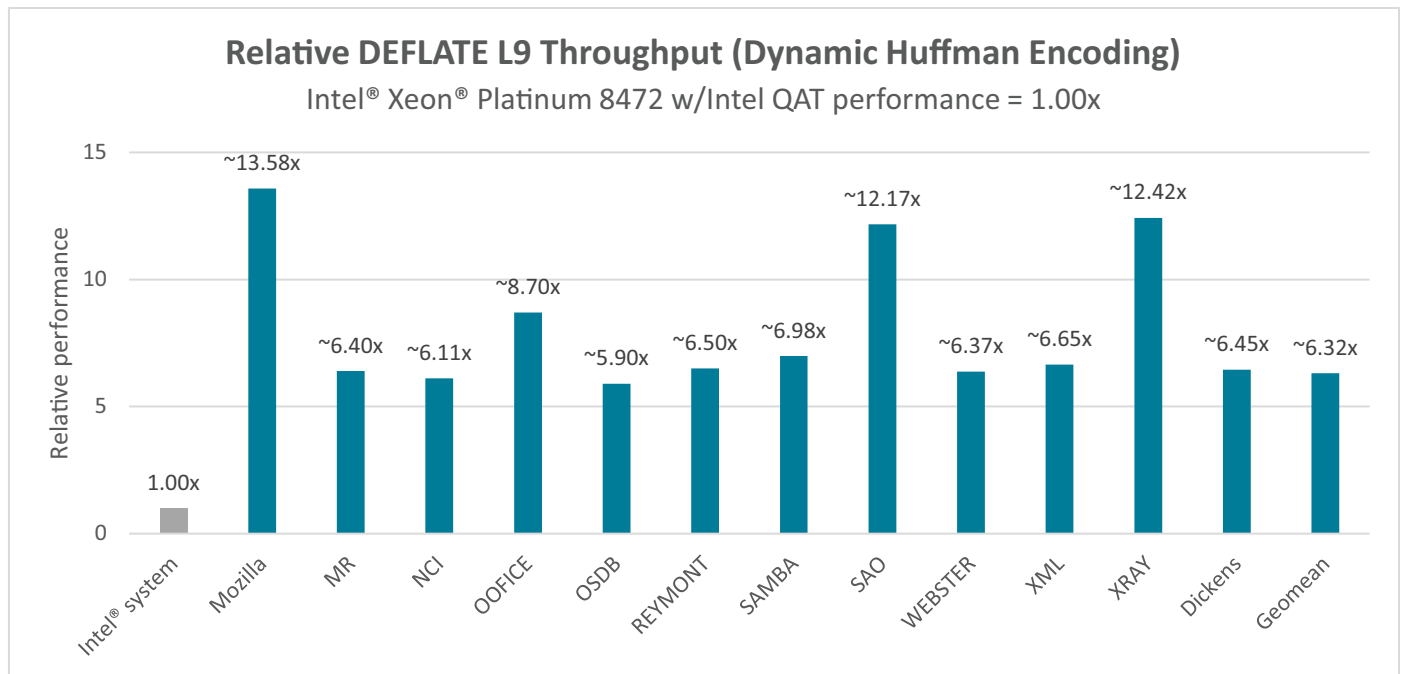


Figure 2: Relative L9 compression throughput performance vs. Intel Xeon Platinum 8472 system with Intel QAT

Compression ratio is another important metric because it measures how much the file can be compressed. Systems without offload capacity rely on the processor(s) to achieve high compression ratios. Adding the MaxLinear Panther III to the AMD EPYC processor-based system achieves compression ratios similar to those achieved using

AMD EPYC CPU cores only and up to a geomean compression ratio uplift of ~1.14x across all Silesia Corpus tests and a peak compression ratio uplift of ~2.04x on the Mozilla Silesia Corpus test versus the Intel system running Intel QAT. See Figure 3.

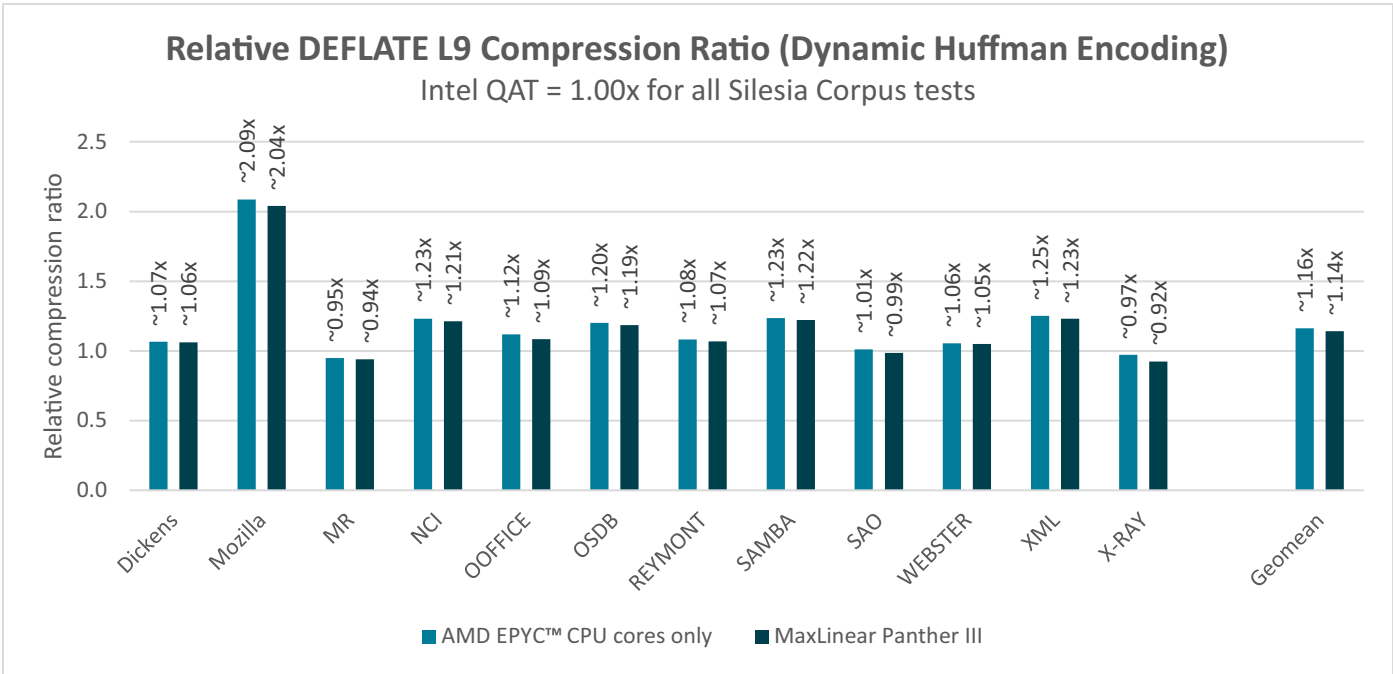


Figure 3: Deflate L9 compression ratio performance

The MaxLinear Panther III can also support data compression ratios beyond the Deflate Level 9 compression shown in Figure 3, which can further reduce data when deduplication is used in conjunction with compression to as much as the 12:1 ratio described above while lowering CPU utilization even further. This white paper does not compare performance beyond Deflate Level 9 because Intel QAT does not support this capability.

Note: The MaxLinear Panther III can compress and compute deduplication hash in a single pass.

Configuring the MaxLinear driver with the appropriate encryption algorithm allows the MaxLinear Panther III solution to compress in a single pass that preserves the rated 200 Gbps throughput. By contrast, the Intel QAT solution requires a compression pass followed by loading the OpenSSL security driver followed by a second pass to encrypt the compressed data. This multistage process requires the CPU to manage temporarily storing the compressed data, configuring the QAT driver with the appropriate encryption algorithm, encrypting the compressed data, and then protecting that data. The Intel QAT does not provide data protection or real-time validation.

POWER CONSIDERATIONS

The AMD EPYC processor and the MaxLinear Panther II card are discrete hardware components that can be independently power managed for optimal efficiency. The AMD EPYC processor performs efficient data processing and then hands that processed data to the MaxLinear Panther III for compression and encryption.

Powering down each component at the appropriate point(s) in this process preserves overall performance while optimizing power consumption because only one component requires high amounts of power at any one time.

SOFTWARE

The AMD and MaxLinear solution has broad software ecosystem support with standard software and a comprehensive Software Development Kit (SDK) stack that offers an easy path for utilizing the MaxLinear Panther III accelerator. The single MaxLinear SDK stack supports multiple compression, crypto, and data protection standards and works with standard data flow usage models. Upgrade and maintenance of the software stack is also straightforward and does not depend on other frameworks or OS kernel updates. The Maxlinear Panther SDK supports both kernel space and user space applications.

The QAT implementation requires multiple stacks with multiple drivers. For instance, the QAT compression driver is needed for compression workloads in the Data Plane Development Kit (DPDK) user space, and a separate QAT driver is needed for encryption and decryption with OpenSSL and DPDK.

RELIABILITY, AVAILABILITY AND SERVICEABILITY (RAS)

RAS is an important part of the server deployment and life cycle. Datacenters, enterprises, and embedded users share the goal of reducing platform downtime while enabling ease of access and service. The AMD/MaxLinear platform offers excellent reliability. Need to service a component? Power down that individual component and perform any needed service. The MaxLinear Panther III's plug-in form factor makes debugging or replacing hardware straightforward.

The MaxLinear Panther III provides real-time verification on all transform operations by applying inverse transform operations and comparing with input data. This coupled with support for data protection, ECC on the device, and eCRC on PCIe provides strong data protection on the device. Further, the Panther Software SDK supports software transform operations, which can act as a failsafe in the unlikely event of a hardware error.

Intel QAT is integrated with the CPU. This means that an accelerator issue will require taking the processor offline for debugging or replacement. The entire processor must also be taken offline for all firmware and/or software patches.

CONCLUSION

Installing MaxLinear Panther III accelerators in systems powered by 4th Gen AMD EPYC processors delivers outstanding performance for workloads that require data compression, deduplication, and encryption/decryption backed by optimal power usage, a mature and proven SW stack, and superior RAS compared to a comparable Intel QAT system.

REFERENCES

1. **GD-183A:** AMD Infinity Guard features vary by EPYC™ Processor generations and/or series. 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>.
2. **EPYC-18:** 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.

Luke Chang is a PMTS Systems Design Engineer at AMD.

DISCLAIMERS

*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.

The information contained herein is for informational purposes only and is subject to change without notice. While every precaution has been taken in the preparation of this document, it may contain technical inaccuracies, omissions and typographical errors, and AMD is under no obligation to update or otherwise correct this information. Advanced Micro Devices, Inc. makes no representations or warranties with respect to the accuracy or completeness of the contents of this document, and assumes no liability of any kind, including the implied warranties of noninfringement, merchantability or fitness for particular purposes, with respect to the operation or use of AMD hardware, software or other products described herein. No license, including implied or arising by estoppel, to any intellectual property rights is granted by this document. Terms and limitations applicable to the purchase or use of AMD products are as set forth in a signed agreement between the parties or in AMD's Standard Terms and Conditions of Sale. GD-18u

©2024 Advanced Micro Devices, Inc. All rights reserved. AMD, the AMD logo, EPYC, and combinations thereof are trademarks of Advanced Micro Devices. Ubuntu is a registered trademark of Canonical, Ltd. Other product names used in this publication are for identification purposes only and may be trademarks of their respective owners. Certain AMD technologies may require third-party enablement or activation. Supported features may vary by operating system. Please confirm with the system manufacturer for specific features. No technology or product can be completely secure.