[Public]

ZenDNN 6.1 Support Matrix

	zentf	**Zentorch**
Release Version	• 2.20.0.1 • 2.21.0.1	• 2.13.0.3 • 2.14.1.0
Release Date	Nov, 2026	Nov, 2026
manylinux Support (Based on Alma Linux GCC version 13.3)	manylinux_2_28	manylinux_2_28
glibc Version	2.28	2.28
Python Version	• 3.9 through 3.13 for 2.20.0.1 • 3.10 through 3.13 for 2.21.0.1	3.10 through 3.14
Native Framework versions	• TensorFlow 2.20.0 • TensorFlow 2.21.0	• PyTorch 2.13.0, 2.14.1 • vLLM 0.27.0 through 0.31.0
Compiler	GCC v14.2	GCC v14.2
Wheel File Release	Yes	Yes
Wheel file built with ZenDNN 6.1.0 and AOCL-DLP 5.2+	Yes	Yes
Wheel File Built with FBGEMM 1.2.0	Yes	Yes
C++ Interface Release	Yes	No
Java Interface Support	Yes, v1.2.0-SNAPSHOT	No
Compatible with OS	• Ubuntu 22.04, 24.04, 26.04 • RHEL 9.2, 9.5, 10.2	• Ubuntu 22.04, 24.04, 26.04 • RHEL 9.2, 9.5, 10.2*
Not Compatible with OS	N/A	N/A
Compatible H/W	4th, 5th , and 6th Gen AMD EPYC™ processors	4th, 5th , and 6th Gen AMD EPYC™ processors
Data Types	• FP32 • BF16	• FP16 • FP32 • BF16 • DA8W8 • DA8W4 • A16W4**

** Use libstdc++.so.6 library which supports GLIBCXX_3.4.30*
** DA8W8 - Dynamic INT8 activations with INT8 weights using symmetric quantization.*

DA8W4 - Dynamic INT8 activations with INT4 weights using symmetric quantization.

A16W4 - BF16 activations with INT4 weights supporting symmetric and asymmetric weight quantization.