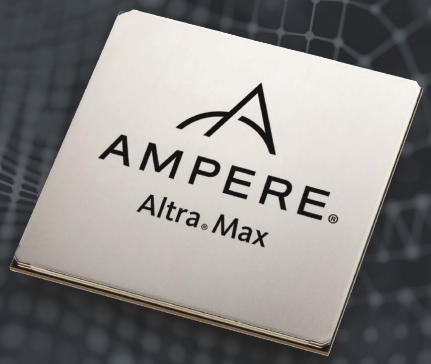
Workload Brief

x264 Video Transcoding on Ampere® Altra® Max



Ampere—Empowering What's Next

The Ampere® Altra® Max processor is a complete system-on-chip (SOC) solution that supports up to 128 high-performance cores with an innovative architecture that delivers predictable high performance, linear scaling and high energy efficiency. Online video continues to rapidly grow, driving usage of video transcoding to compress videos which greatly reduces both storage space and network bandwidth. We demonstrate Ampere Altra Max is ideal for running video transcoding using x264 by delivering both industry leading performance and power efficiency.

x264 on Ampere Altra Max

Ampere Altra Max is designed to deliver exceptional performance and power efficiency for applications like video transcoding. We use libx264 which implements the H.264/MPEG-4 AVC standard that is the most widely used today. Ampere Altra Max uses an innovative architectural design, operating at consistent frequencies with single-threaded cores that make applications more resistant to noisy neighbor issues. This allows workloads to run in a predictable manner with minimal variance. Additionally, the processors are designed to be highly power efficient. Together, this gives Ampere Altra Max outstanding performance and power efficiency running x264.

Benchmarking Configuration

We evaluate x264 using "vbench: a Benchmark for Video Transcoding in the Cloud, a benchmark for the emerging video-as-a-service workload", available at http://arcade.cs.columbia.edu/vbench. Vbench's 15 input videos were algorithmically selected to represent a large commercial corpus of millions of videos based on resolution, framerate, and complexity. We use the "Upload" and "Video on Demand" configurations to evaluate performance and power usage. Upload uses a single pass transcoding without degrading the input video quality which represents the initial upload encoding to a video service, requiring speed and quality. The Video on Demand (VoD) configuration uses a 2 pass transcoding that requires speed and improved compression without degrading video quality. The VoD first pass collects statistics used in the second pass to allocate more bits when encoding complex vs. simple frames.

Benefits of running x264 on Ampere Altra Max

- **Cloud Native**: Cloud Native: Designed from the ground up for 'born in the cloud' workloads like x264, Ampere Altra Max delivers up to 2.09x higher performance than the best x86 servers.
- **Energy Efficiency:** Energy Efficiency: With up to 128 energy-efficient Arm cores, Ampere Altra Max consumes up to 1.3x less power than leading x86 servers with better performance.
- **Lower Carbon Footprint:** Industry-leading performance and high energy efficiency result in Ampere Altra Max demonstrating up to 2.8x higher Performance/watt, leading to lower TCO and a smaller carbon footprint.
- **Scalable**: Ampere Altra Max processors delivering consistent performance at the socket level greater than the best x86 servers. This leads to much higher resistance to noisy neighbors in multitenant environments.

To maximize ffmpeg throughput, we run multiple ffmpeg instances equal to the number of CPU cores available on the socket, using one ffmpeg thread per instance. All ffmpeg instances are run on one socket with a dedicated CPU core using numactl to set affinity. We report the average time to transocde the 15 vbench input files for each ffmpeg process and the socket level power usage. To minimize OS overhead, the ffmpeg binary, and all input and output files are stored on a ramdisk. We compare Ampere Altra Max M128- 30 processor to Intel® Xeon® Platinum 8380 (Ice Lake) and AMD EPYC™ 7763 (Milan) running CentOS 8.4 with 4.18 kernel. We built the latest available versions of ffmpeg version and libx264 with gcc 11 on all platforms. See Additional Benchmarking Details description below for additional details including the ffmpeg commands run.

Transcoding Performance

Ampere Altra Max has the best transcode performance running ffmpeg using x264 compared to Intel Xeon 8380 and AMD EPYC 7763. In Figure 1, we plot the average transcoding time for each ffmpeg process showing Ampere Altra Max is 2.09x and 1.79x faster than Intel Xeon Platinum 8380 for the Upload and VoD configurations, respectively. Ampere Altra Max is 1.15x and 1.05x faster than AMD EPYC 7763 (Milan) for Upload and VoD.

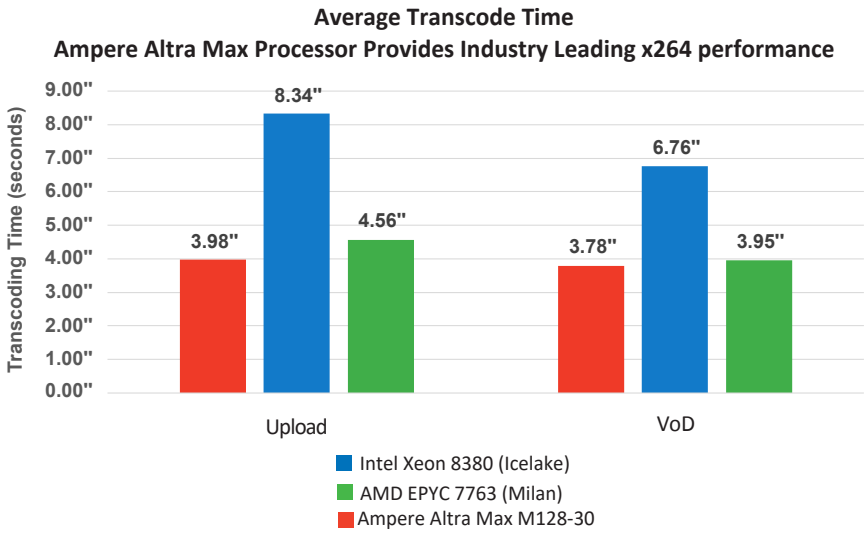


Figure 1. Average time to transcode

Transcoding Power Efficiency

In addition to the best transcoding performance, Ampere Altra Max is the most power efficient processor, reducing the carbon footprint of video transocding. In Figure 2, we plot the socket level power usage showing Ampere Altra Max is 1.17x more power efficient compared to Intel® Xeon® Platinum 8380 Processor (Ice Lake) and 1.24x vs. AMD EPYC™ 7763 (Milan) for the Upload configuration. For VoD, Ampere Altra Max is 1.22x more power efficient vs. Intel® Xeon® Platinum 8380 Processor (Ice Lake) and 1.29x vs. AMD EPYC™ 7763 (Milan).

Ampere Altra Max

- 128 Armv8.2+ 64-bit cores at 3.0GHz
- 64KB i-Cache, 64KB d-Cache per core
- 1MB L2 Cache
- 16MB-32MB System Level Cache
- Coherent mesh-based interconnect

Memory

- 8x72 bit DDR4-3200 channels
- ECC and DDR4 RAS
- Up to 16 DIMMs (2 DPC) and 4TB/socket

Connectivity

- 128 lanes of PCIe Gen4
- Coherent multi-socket support
- 4x16 CCIX lanes

System

- Armv8.2+, SBSA Level 4
- Advanced Power Management

Performance

- SPECrate®2017_int_base:350

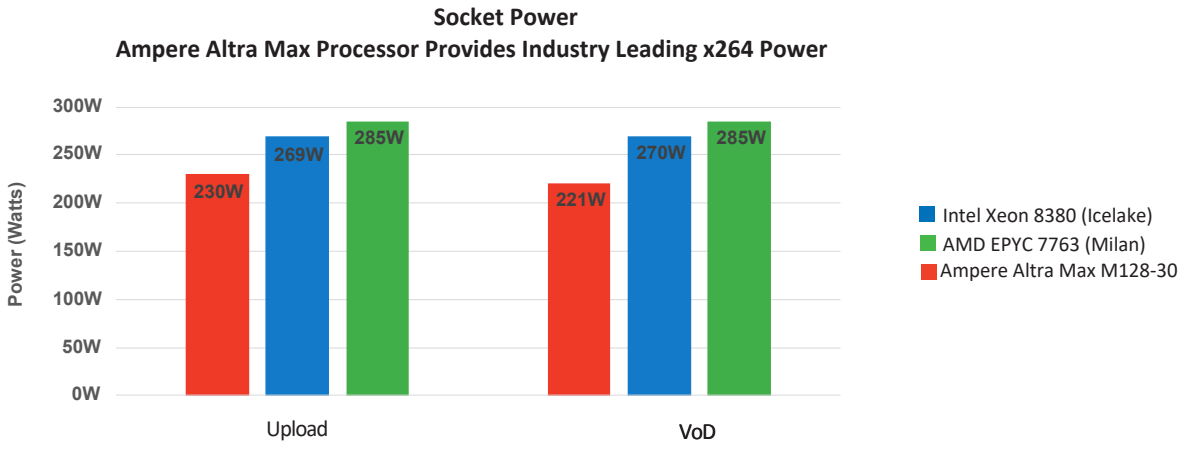


Figure 2. Socket level power

Benchmarking Results and Conclusions

Ampere Altra Max processors are a complete System On Chip (SOC) solution built for Cloud Native workloads, designed to deliver exceptional performance and energy efficiency for applications like video transcoding using x264. Ampere Altra Max delivers both industry leading performance and power efficiency running x264 with up to 2.09x faster performance compared to Intel Xeon 8380 and is up to 1.22x more power efficient. Compared to AMD EPYC 7763, Ampere Altra Max is up to 1.15x faster and up to 1.29x more power efficient. In additional to providing the fastest video transcoding, Ampere Altra Max innovative architecture that delivers predictable high performance with it's highly energy efficient design and reduces the carbon footprint of video transcoding.

For more information on this workload or other workloads our engineers have been working on, please visit: https://developer.amperecomputing.com/.

Additional Benchmarking Details
We used the latest ffmpeg version available, N-105446-g2e82c61055, git commit 2e82c610553efd69b4d9b6c359423a19c2868255 and the latest libx264, git commit obb85e8bbc85244d5c8fd300033ca32539b541b7. We used the vbench configurations specified in "vbench: a Benchmark for Video Transcoding in the Cloud, a benchmark for the emerging video-as-a-service workload", Andrea Lottarini, Alex Ramirez, Joel Coburn, Martha A. Kim Parthasarathy Ranganathan, Daniel Stodolsky, and Mark Wachsler (2018), available at http://arcade.cs.columbia.edu/vbench.
Upload configuration: numactl -m 0 -C $CORE ffmpeg -i $INPUT -c:v libx264 -threads 1 -y -loglevel quiet -crf 18 $OUTPUT
VoD configuration: The bitrate is set per input file as described in the vbench paper.
numactl -m 0 -C $CORE ffmpeg -i $INPUT -c:v libx264 -threads 1 -y -loglevel quiet -passlogfile ffmpeg2pass -pass 1 -f null -an -sn -b:v $BITRATE -preset medium /dev/null
numactl -m 0 -C $CORE ffmpeg -i $INPUT -c:v libx264 -threads 1 -y -loglevel quiet -passlogfile ffmpeg2pass -pass 2 -b:v $BITRATE -preset medium ${OUTPUT}

Ampere Computing® / 4655 Great America Parkway, Suite 601 / Santa Clara, CA 95054 / amperecomputing.com