Solution Brief

3rd Gen Intel® Xeon® Scalable Processor
Visual Cloud

Cost-Effective Growth of High-Definition and Ultra-High-Definition Video Services

ATEME achieves up to a 2x increase in high video transcoding density per node using the 3rd generation Intel® Xeon® Scalable processor, improving the key cost-per-channel metric for video content providers.[1]

Cable television, IPTV, and over-the-top (OTT) providers are all in the midst of building out their video headend capabilities to meet the growing subscriber demand for high-definition (HD) and ultra-high-definition (UHD) programming. Transcoding to provide these high-resolution services consumes far more computing resources per stream than standard-definition (SD) video, reducing the number of streams that can be handled per server and increasing the server count needed to support a given number of subscribers.

The number of streams that can be transcoded per server and delivered to customers in various formats therefore has direct and dramatic effects on the cost of providing services, and accordingly on profitability. Optimizing that density requires high performance in the video headend, which depends on both hardware and software. The business value of increasing the performance of both hardware and software ingredients is expected to continue growing in the near future.

Virtualizing the Video Headend

The ongoing industry trend has been for providers to virtualize the video headend as a means of increasing agility and reducing the cost per channel. And by supporting more streams per node, providers can either serve more customers or deliver higher-resolution content to the same customers with a given amount of headend infrastructure. Meeting the compute-intensive requirements of transcoding with CPU resources maximizes the flexibility of general-purpose servers and streamlines software development.

Virtualized headends offer other benefits as well, including very rapid channel installation handled by remote management capabilities instead of requiring an onsite engineer. This agility gives video delivery providers new flexibility to add or update channels to meet consumer demand for popular programming, offering dedicated channels for live events, upgrading channels to 4K, or adding new compression standards to support novel devices.

The functions of a video headend include acquisition, transcoding, metadata enrichment, scrambling, and multiplexing. To perform these functions, the video headend performs extensive processing on each video stream, including receiving and decoding the incoming video streams before encoding, transcoding, rate shaping, channel modulating, channel processing, and channel signal adapting the video for the subscriber's end device. Virtualizing the headend makes it all the more challenging to perform all these functions in real time with zero packet loss and low enough latency and jitter to ensure video quality and safeguard the viewer experience.

Ensuring reliable transmission over the network is also a challenge, with unforgiving quality requirements. Video streams are often transmitted in multicast to accommodate large numbers of subscribers, which requires flawless performance at the headend. After all, real-time video streaming makes no allowances for recovering dropped or malformed packets. Intel® Network Builders ecosystem partner ATEME has tested its TITAN© Live video headend on a server based on the 3rd generation Intel® Xeon® Scalable processors to meet these challenges.

ATEME TITAN Live

ATEME TITAN Live is a high video quality, high density software compression solution, designed for delivery of SD, HD, and UHD content. TITAN Live can ingest any channel to simultaneously produce multiple streams in real time and deliver live video compression for the converged headend.

As shown in Figure 1, TITAN Live offers comprehensive distribution features, which include encoding, transcoding, multiplexing, Statmux, and packager. TITAN Live is a true virtualized software-based solution that is optimized for Intel® architecture servers.

TITAN Live is based on ATEME's 5th generation STREAM compression engine, which delivers very high video quality at very low bitrates with accelerated parallel processing. With a powerful user interface, management system, and extensive set of APIs, TITAN Live can easily integrate with any ecosystem, reducing time to market and operational expenditures.

Next-Generation Server Platform

The ATEME TITAN Live virtual headend is built to scale transcoding performance and make efficient use of processing resources. The 3rd generation Intel Xeon Scalable processor provides advances in architecture and throughput that benefit ATEME TITAN Live as service providers work to maximize the number of live video streams delivered per server. The processor is available in a flexible range of SKUs, with a wide range of core counts, frequencies, features, and power levels.

The 3rd generation Intel Xeon Scalable processor optimizes transcoding throughput with ATEME TITAN Live through higher performance per core than predecessor platforms, enabled by L1 caches of up to 48 KB (50 percent larger than the previous generation). The processor also helps accelerate memory-bound transcoding workloads with increased memory bandwidth and support for up to 4 TB capacity per processor socket.

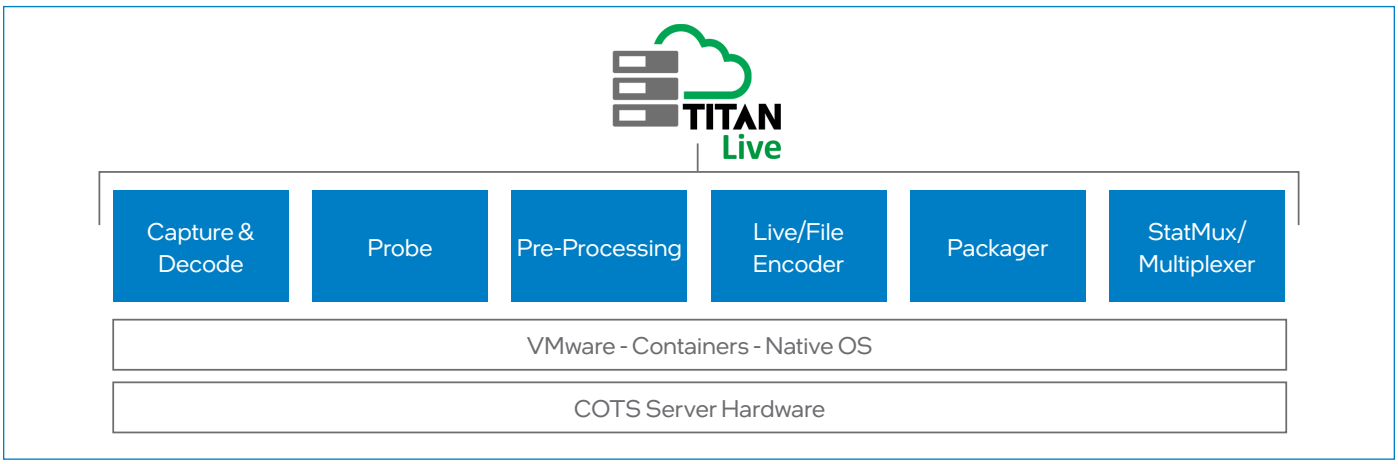
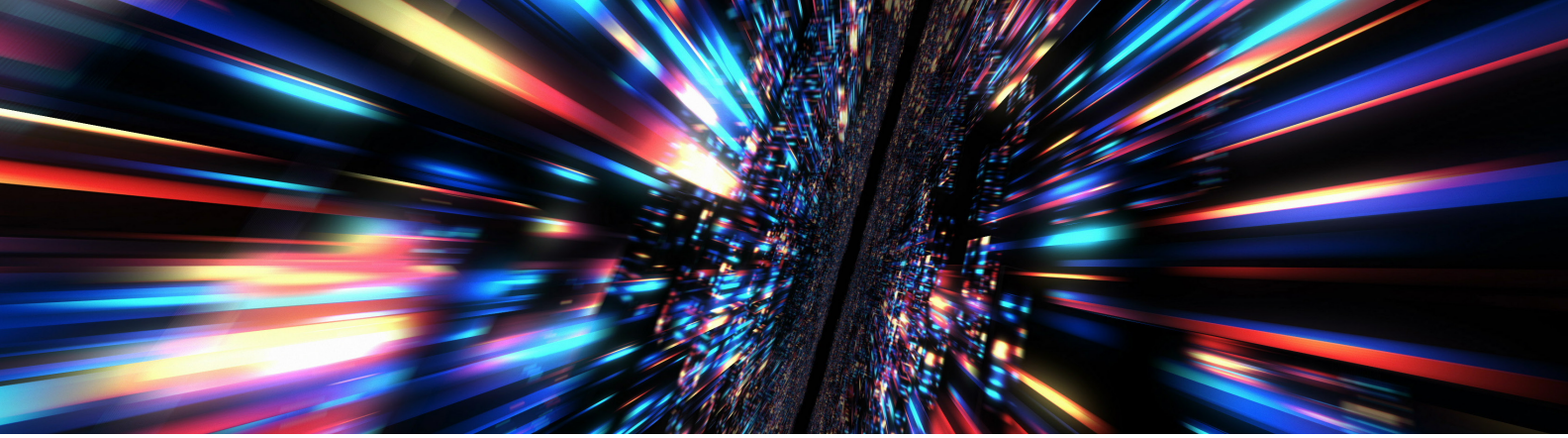


Figure 1. Block diagram of TITAN Live services.

3rd Generation Intel Xeon Scalable Processors

- **Flexibility from the edge to the cloud**, bringing AI everywhere with a balanced architecture, built-in acceleration, and hardware-based security.
- **Part of a complete set of network technology from Intel**, including accelerators, Ethernet adapters, Intel Optane persistent memory, FlexRAN, OpenNESS, Open Visual Cloud, and Intel® Smart Edge.
- **Engineered for modern network workloads**, targeting low latency, high throughput, deterministic performance, and high performance per watt.
- **Enhanced built-in crypto-acceleration** to reduce the performance impact of full data encryption and increase the performance of encryption-intensive workloads.
- **Hardware-based security** using Intel® Software Guard Extensions (Intel® SGX),[2] enhanced crypto processing acceleration,[2] and Intel® Total Memory Encryption.[2]



Test Setup

ATEME devised a test to show the impact of adopting 3rd generation Intel Xeon Scalable processors for HD and UHD transcoding. The devices under test (DUTs) were as follows:

- **DUT-1 (2nd generation Intel Xeon Scalable processor):** One node based on 2x Intel Xeon Gold 6230R processors (26 cores, 52 threads per socket)
- **DUT-2 (3rd generation Intel Xeon Scalable processor):** One node based on 2x Intel Xeon Gold 6330 processors (28 cores, 56 threads per socket)
- **DUT-3 (3rd generation Intel Xeon Scalable processor):** One node based on 2x Intel Xeon Gold 6338 processors (32 cores, 64 threads per socket)

The HD tests measured the transcoding performance of a 5 Mbits/s input stream into a 5 Mbits/s output stream. For UHD, the tests used a 20 Mbits/s input stream being transcoded to a 20 Mbits/s output stream. Processing the UHD files is compute-intensive because they run at 50 frames per second, so that the video stream features four times the resolution and two times the frame rate.

For simplicity of testing, input and output were streamed using the server's loopback interface. The performance of the encoding process relies heavily on two factors:

- Cache size and performance
- CPU frequency, instructions per cycle (IPC), and core count

Performance Outcomes: Higher Transcoding Performance per Node

Testing demonstrates the potential for significant increases in the number of streams that can be transcoded per node using the 3rd generation Intel Xeon Scalable processor, as shown in Figure 2.[1] These performance increases were observed across the following workloads:

- **HEVC Ultra High-Definition transcoding:** up to 100 percent more streams per node
- **Preprocessing:** up to 79.4 percent more streams per node
- **MPEG2 Standard Definition transcoding:** up to 42.2 percent more streams per node
- **H.264 High-Definition transcoding:** up to 50 percent more streams per node

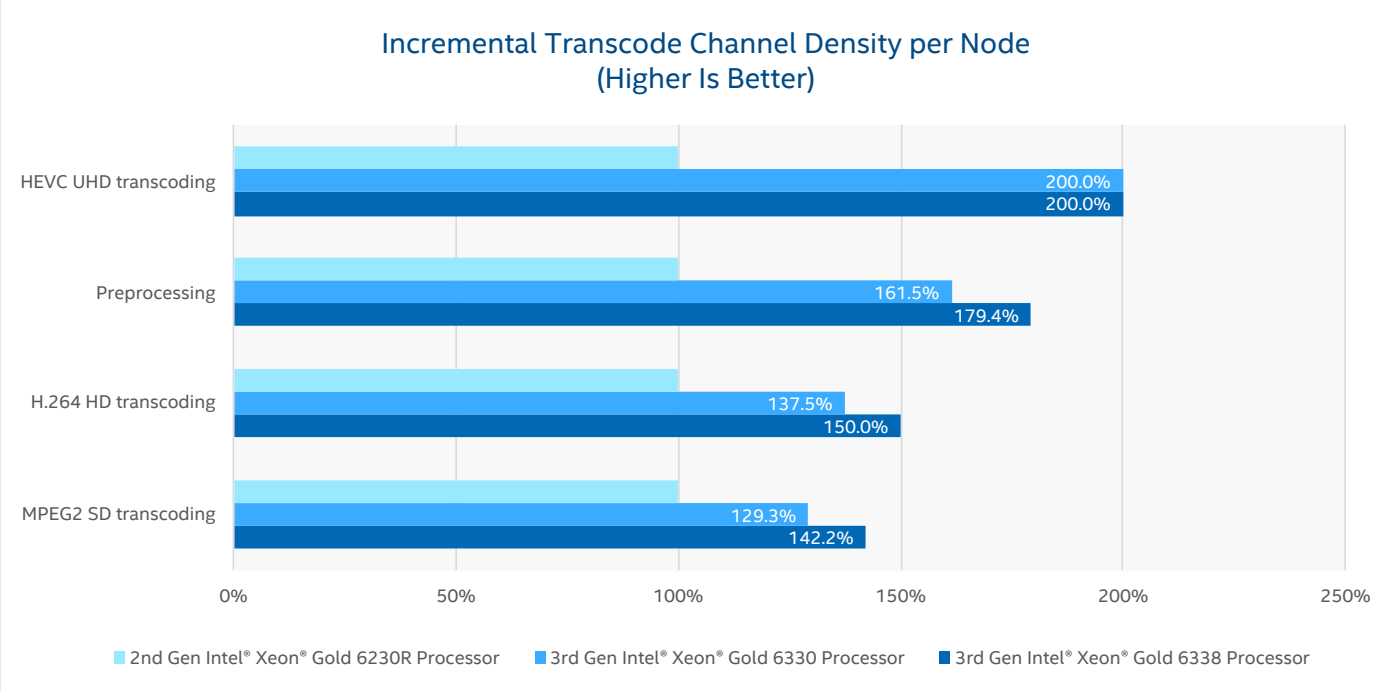


Figure 2. High channel density across video transcoding workloads.[1,3]

The advances built into the 3rd generation Intel Xeon Scalable processor help service providers optimize the amount of work they get out of every server, potentially reducing the number of servers they must procure, deploy, and maintain. That savings on both OpEx and CapEx helps expand operations further within budget constraints, positioning providers for an improved competitive position going forward.

Conclusion

Test results show the ATEME TITAN Live virtual headend offers up to 2x the density of transcode streams per node running on servers powered by 3rd generation Intel Xeon Scalable processors. This platform is based on a balanced architecture with higher performance per core than its predecessors, complemented by improvements in L1 cache size and memory bandwidth. Increased processing power enables more cost-effective deployment for video transcoding and other compute-intensive workloads for both today's needs and tomorrow's demands.

In these tests, this performance increased the number of channels supported by a single server to optimize the cost per server. The performance increases shown in these tests demonstrate that TITAN Live provides an outstanding basis for video delivery providers to build out infrastructure that will efficiently support ongoing growth in demand for HD and UHD content.

Learn More

- ATEME Titan Live
- Intel® Network Builders
- 3rd generation Intel® Xeon® Scalable processors
- Intel® Visual Cloud

Notices & Disclaimers

[1] Testing performed March 30, 2021, by ATEME.

DUT-1 (2ND GENERATION INTEL XEON SCALABLE PROCESSOR): One node based on 2x Intel Xeon Gold 6230R processors (26 cores, 52 threads per socket); 96 GB RAM total (12 modules @ 8 GB each, 2666 MHz); Intel Hyper-Threading Technology enabled; Intel Turbo Boost Technology enabled; Dell SSD T1WH8 (240 GB); Dell Intel Ethernet I350 Network Interface Card; microcode 0x05003003; BIOS version Dell 2.9.4; Debian 10 Buster, Kernel 4.19.104; TITAN Live.

DUT-2 (3RD GENERATION INTEL XEON SCALABLE PROCESSOR): One node based on 2x Intel Xeon Gold 6330 processors (28 cores, 56 threads per socket); 256 GB RAM total (16 modules @ 16 GB each, 2933 MHz); Intel Hyper-Threading Technology enabled; Intel Turbo Boost Technology enabled; Intel SSD D3-S4510 Series (960 GB); Intel Ethernet Converged Network Adapter X550T; microcode x270; BIOS version SE5C6200.86B.3021.D40.2103160200; CentOS 8.2, Kernel 4.18.0-193; TITAN Live.

DUT-3 (3RD GENERATION INTEL XEON SCALABLE PROCESSOR): One node based on 2x Intel Xeon Gold 6338 processors (32 cores, 64 threads per socket); 256 GB RAM total (16 modules @ 16 GB each, 3200 MHz); Intel Hyper-Threading Technology enabled; Intel Turbo Boost Technology enabled; Intel SSD D3-S4510 Series (960 GB); Intel Ethernet Converged Network Adapter X550T; microcode x270; BIOS version SE5C6200.86B.3021.D40.2103160200; CentOS 8.2, Kernel 4.18.0-193; TITAN Live.

[2] This technology is not supported when using Intel Optane persistent memory.

[3] Absolute performance numbers for transcode channel density per node.

	3rd Gen Intel® Xeon® Gold 6338 Processor	3rd Gen Intel® Xeon® Gold 6330 Processor	2nd Gen Intel® Xeon® Gold 6230R Processor
MPEG2 SD transcoding	165	150	116
H.264 HD transcoding	24	22	16
Preprocessing	70	63	39
HEVC UHD transcoding	4	4	2

Performance varies by use, configuration and other factors. Learn more at https://www.intel.com/PerformanceIndex.

Performance results are based on testing as of dates shown in configurations and may not reflect all publicly available updates. See backup for configuration details. No product or component can be absolutely secure.

Your costs and results may vary.

Intel does not control or audit third-party data. You should consult other sources to evaluate accuracy.

Intel technologies may require enabled hardware, software or service activation.

0521/DO/H09/PDF ♻ Please Recycle 341972-001US