ADVANTECH
Enabling an Intelligent Planet

socionext
for better quality of experience

White Paper

Hardware Accelerated Video Processing:

Building Agile, Scalable and Cost-Efficient UHD Solutions from Edge Contribution to Cloud Distribution

Table of Contents

The UHD Wave

The transition towards ultra-high definition (UHD) video is here and it brings an infrastructure revolution with it. As of 2017, there are around 50 Ultra HD channels airing globally and several over-the-top (OTT) services such as Netflix or Amazon already offer 4K plans. 4K TV prices are falling and that reflects on market penetration. 4K UHD TVs will make up over a third of all TV sales in 2017 and are expected to account for 60% of all shipments by 2020[1].

UHD greatly improves viewer experience. The hyper-realistic representation provided by 4K is achieved by packing 4x more pixels than currently mass deployed Full HD. And it is not just a matter of pixels, the WOW factor of UHD is built on other enriching technologies such as High Dynamic Range (HDR) and High Frame Rate (HFR) which have even further impact on the image realism. UHD doesn't end in 4K either, 8K is in the horizon. The Japanese public broadcaster NHK is testing its Super Hi-Vision channel which held the world's first 8K TV satellite broadcast at the 2016 Rio Olympics and plans to provide full 8K coverage of the 2020 Tokyo Olympics.

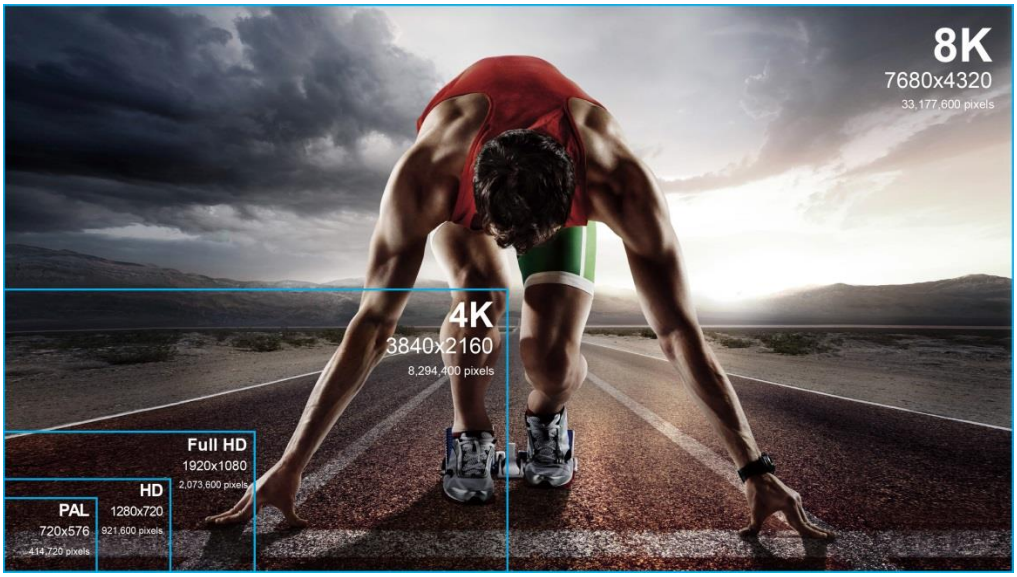


Figure 1. UHD Pixel Increase

On the infrastructure side, UHD deployment translates into a massive amount of video data that will need to be captured, processed and transmitted. In order to provide truly immersive UHD experiences that engage with the audience, media companies will need to optimize their video infrastructure to prepare for the upcoming UHD wave.

[1] Futuresource Consulting https://www.futuresource-consulting.com/Press-4K-UHD-TV-Uptake-Continues-0517.html & https://www.futuresource-consulting.com/Press-Global-TV-Market-1216.html

OTT Rules

The growing demand for over-the-top streamed content is disrupting broadcast, telecom and internet business models. Within the next five years, viewership hours of live-linear streaming OTT video will surpass those of traditional broadcast TV[2]. As shown in the image below, OTT delivered video content already has a significant impact on internet traffic and this will continue to grow with 80 percent of the world's internet traffic expected to be video by 2021[3].

The Online Video Era Drivers

		2016	2021
More Connected Devices		17.1 Billion	27.1 Billion
More Video Viewing		73% of Traffic	80% of Traffic
Increasing Video Definition		15% UHD Connected TV Sets	56% UHD Connected TV Sets
Faster Broadband Speeds		27.5 Mbps	53 Mbps

Mobile and social viewers will play a key role in increasing online video consumption with 5G being a key driver due to the broader bandwidth available. If 5G hold its promises, the mobile network could eventually support the whole content creation and consumption process, anywhere from contribution to mobile live streaming. This fundamental change in the video workflow brings great opportunities for media, telecom and internet companies who are searching for cost-effective ways to create and deliver high quality content across a broader range of devices.

[2] 2017 OTT Video Services Study http://www.streamingmedia.com/Research/7064-OTT-VIDEO-SERVICES-INNOVATION-OPPORTUNITY-MATURATION--TECHNOLOGY-TRENDS-IN-OTT-DELIVERY.htm
[3] Cisco VNI Global IP Traffic Forecast, 2016–2021 https://www.cisco.com/c/en/us/solutions/collateral/service-provider/visual-networking-index-vni/complete-white-paper-c11-481360.html

We are On Air

The WOW factor of UHD finds a perfect partner in live TV. The thrilling experience of live is not only relevant for events such as sports but also for TV shows which achieve higher viewership when their audience can interact online as the action happens. That is the case of the last episode of Game of Thrones Season 7 that reached a live audience of 12 million people. Producers highlighted the effervescence of social media live engagement which has greatly contributed to the show's success since its premier. Increasing audience engagement is also the reason why the leading online video service YouTube recently announced its new ultra-low latency mode that enables real-time interaction with the live audience and targets streaming of live events and shows. And last but not least, the community of online gamers also has a significant impact on the live streaming video industry. Live streaming services which come from a gaming background such as Twitch are gaining momentum. These demand the lowest latency and highest resolution to serve their subscriber community. On the technical side, low latency video processing adds to the challenges of UHD increased pixel count. The live factor requires real-time processing of high bitrate signals which brings the video processing infrastructure to the front line as the critical point to consider when looking to provide low latency UHD streaming services.

Immersive Experience is Next

Although 8K resolution is the natural evolution for broadcasting, there is another driving force for UHD which is pushing even harder. Virtual Reality (VR) and Augmented Reality (AR) have the potential to fully transform viewer experience and can now be achieved thanks to UHD technology. The life-like quality required for building hyper-immersive experiences meets its demands in 8K. From broadcasting and social media to gaming and enterprise, many industries can benefit from hyper-immersive VR. This is especially relevant for live events such as sports, concerts and trade shows. Virtual tickets are new revenue generators that can become an important part of the live entertainment industry income.

Predictions look good. Cisco forecasts that the number of VR headsets is expected to grow to nearly 100 million by 2021 and more than half of these will be connected to smartphones which will give them access to live streamed content worldwide. While media companies don't want to miss opportunities on what might be the next big thing, they are cautious as they are still on the learning curve of how to create and deliver content in a whole new way.

The Bandwidth Factor

The first consideration when looking at deploying UHD video services is the broader bandwidth required to deliver higher bitrate video signals to viewers. Looking at growing OTT adoption, it was foreseen from the beginning of UHD, even for 4K resolution, that current codecs wouldn't be

efficient enough to successfully stream UHD video over bandwidth constrained networks. In order to keep bit rates under control, the Joint Collaborative Team on Video Coding (JCT-VC) developed the successor of the widely adopted H.264/AVC. The new High Efficiency Video Codec (HEVC) or H.265 has demonstrated its ability to deliver the same subjective quality with a bit rate reduction of 50% over its predecessor[4] and it is gaining adoption with important players such as Netflix and Apple already supporting it.

This 50% bandwidth reduction is a key UHD enabler. HEVC can compress a broadcast-quality 4Kp60 4:2:2 10-bit feed at bitrates between 10 and 40 Mbps from its original uncompressed 12Gbps. However, HEVC's advantages come at the expense of increased computational complexity. 4K HEVC encoding computational load can be up to 40x that of Full HD AVC which outstrips the capacity of existing video processing solutions.

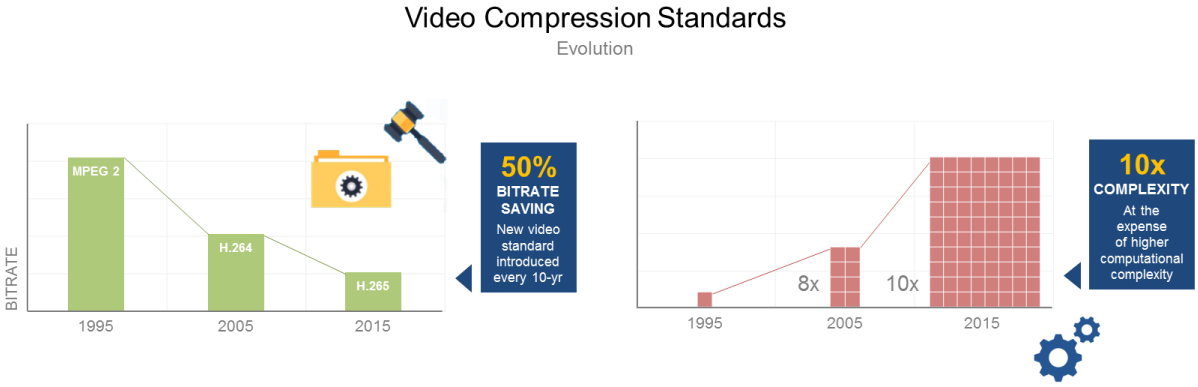


Figure 2. HEVC Bandwidth Gains vs Computational Costs

Alternatives to MPEG's standards-based codecs have been proposed by the Alliance for Open Media. VP9 is available and supported by Google on its YouTube and Android platforms. Its awaited successor AV1 is under development and is meant to compete with H.265 but we will still need to wait a while for it to be publicly available in order to have broader, independent results to check its efficiency against HEVC. In any case, the compute capacity required to code AV1 signals is expected to be at least the same as that required for HEVC if not higher which keeps the infrastructure computing capacity as an open challenge.

The IT Connection

In the same way that other industries are going through a major transformation, the media industry is looking forward to eliminate silos by adopting proven IT technologies such as virtualization and software programmability based on open interfaces that guarantee interoperability between vendors. Greater agility is enabled by software-defined everything. As software becomes an entity on its own, it can be decoupled from the underlying hardware bringing virtualization to play and all the associated benefits of commercial-off-the-shelf infrastructure. And as a final step, all these technologies come together to enable the transition to cloud based services which improve operational efficiency and bring greater economies of scale.

[4] BBC R&D http://www.bbc.co.uk/rd/blog/2016-01-h-dot-265-slash-hevc-vs-h-dot-264-slash-avc-50-percent-bit-rate-savings-verified

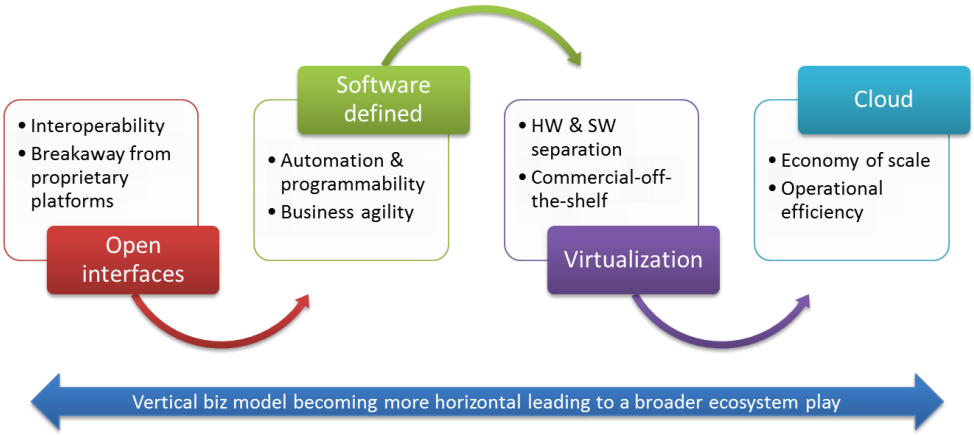


Figure 3. The Convergence of Broadcast, Telecom & IT

Cloud-based architectures enable rapid service provisioning which allows media companies to quickly adapt to fast-changing business requirements. In a cloud-based scenario, OTT video delivery is offered as a service, offloading the video processing and distribution hassle from content creators. Media companies can also choose a hybrid approach with live UHD services running on premises or on a private cloud and less critical or elastic on-demand services running on a public cloud. In any case, the adoption of virtual and software-defined architectures moves video processing applications to x86-based, commercial-off-the-shelf (COTS) servers where a way to integrate the processing power previously available in specialized video platforms needs to be found.

Video's nature is very different to IT application data. Even the latest server-class processing technology struggles to cope with the task of creating multiple OTT HEVC profiles for multiscreen live UHD video streaming in a software-defined, virtualized scenario. Different acceleration technologies are breaking into video developments as an effective way to offload compute-intensive encoding, decoding and transcoding tasks from server CPUs achieving better scalability and greener deployments. On the other hand, the introduction of acceleration cannot compromise the software flexibility provided by x86 based developments and the cost efficiencies provided by the use of mainstream servers.

The Power of Acceleration

The expected video tsunami is urging the development of advanced video processing hardware and algorithms that can efficiently cope with the massive compute required by low-latency UHD processing in real-life deployments. Silicon manufacturers are working on improved System-on-Chip (SoC) video processing technology that can provide the performance to handle broadcast-quality UHD processing at a fraction of the footprint and power required by software implementations.

This is the case of Socionext, a leader in hardware HEVC encoding established from former Fujitsu and Panasonic LSI businesses. Socionext has developed the MB86M31 SoC that supports professional-grade, real-time 4Kp60 H.265 encoding using the Main 10 profile and 4:2:2 subsampling in a single chip that consumes less than 7W. In addition to HEVC encoding, the new

MB86M30 adds support for AVC and MPEG-2 codecs and can also perform decoding and transcoding operations in the same size and power envelope.

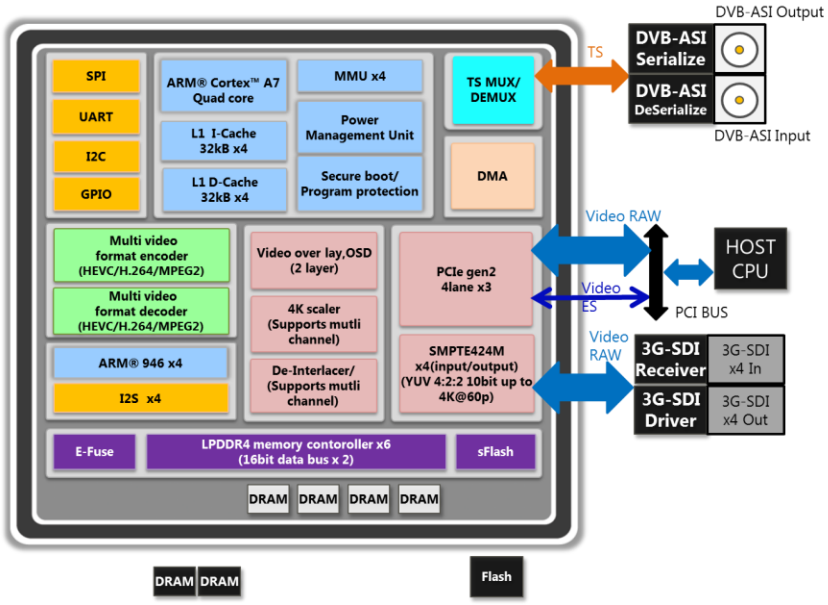


Figure 4. Socionext MB86M30 Block Diagram

This impressive piece of technology boosts performance of video processing applications by offloading encoding, decoding and transcoding tasks in hardware accelerated x86 based implementations. The performance, space and power gains are significant. Results for the popular x265 library show that optimized software-only HEVC implementations running on latest generation server-class processors require dual socket Intel® Xeon® Processor servers[5] for performing live encoding of 4K 10-bit video at 60 frames per second. The same operation can be performed by one single chip leveraging Socionext's encoding technology. When working together with a host CPU, accelerated HEVC encoders based on Socionext SoCs can achieve the same level of performance in compact appliances such as the Advantech VEGA 6300 Series consuming less than 75W.

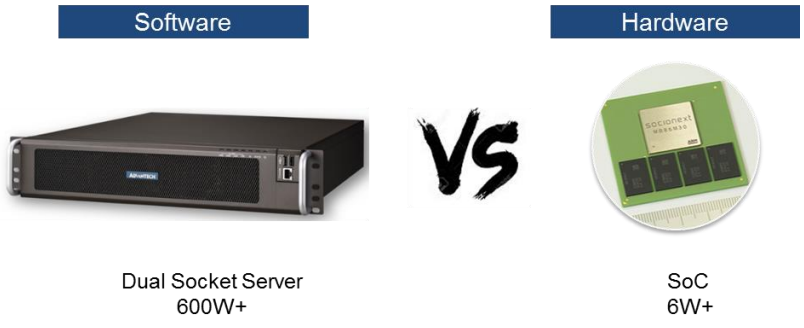


Figure 5. Live UHD HEVC Implementation Options

[5] x265 http://x265.org/haivision-demonstrate-breakthrough-performance-live-4k-hevch-265-software-encoding-2017-nab-show/ & https://builders.intel.com/docs/networkbuilders/accelerate-UHD-encoding-with-intel-xeon-scalable-processors.pdf

Scale to Infinity and Beyond

Hardware acceleration gains become critical in large scale deployments where hundreds or thousands of video streams need to be processed in a data center environment. When dealing with latest generation codecs in a scenario of increasing video resolution and consumption, software-only deployments do not scale nearly as well as hardware accelerated solutions. Accelerated solutions allow for a higher density, greener and more cost-efficient video infrastructure with savings that can reach several orders of magnitude in the case of Socionext based solutions. As an example, the new Advantech VEGA-3318 transcoding accelerator that integrates eight Socionext MB86M30 SoCs supports 8-channel low latency UHD HEVC processing in a single PCI Express adapter. Four VEGA-3318 accelerators can be integrated into a 1U data center server supporting 32 UHD or 128 FullHD HEVC ABR bundles per rack unit – the highest density available in the market.

Figure 6. High Density 1U Server for the Data Center supporting up to 32 UHD or 128 FullHD HEVC ABR Bundles

Sample FullHD ABR Bundle, 1 in -> 5 out	
Input	**Outputs**
1920x1080p60 HEVC/AVC ES stream, 4:2:0, 8 bits, 10Mbps	All HEVC ES streams, 4:2:0, 8 bits: 1280x720p60 / 3Mbps 1280x720p30 / 2Mbps 720x480p30 / 1.5Mbps 640x360p30 / 500Kbps 320x240p30 / 200Kbps
A 1RU Advantech/Socionext accelerated server supports 128 bundles	

A high level comparison with other implementation options for an HD ABR bundle is depicted below.

		Advantech/Socionext Accelerated Server	Hardware Accelerated CPU	High Performance Data Center CPU
AVC	720P ABR Density / RU	**256**	72	12
	Power / ABR Channel	**2.5W**	7.6W	62.5W
	Cost / ABR Channel	**1X**	2X	4X
HEVC	720P ABR Density / RU	**256**	36	2
	Power / ABR Channel	**2.5W**	15W	375W
	Cost / ABR Channel	**1X**	4X	24X

The Benefits of COTS

Over the past few years, the media industry has started a revolution to open the video infrastructure by leveraging IT technologies. By implementing open architectures that rely on virtual and software-defined environments, the industry can benefit from the flexibility and economy of scale of commercial-off-the-shelf servers based on General Purpose Processors (GPP). Even low latency UHD services that require massive compute capacity can benefit from flexible and agile GPP based solutions by integrating hardware acceleration. Providing COTS access to high-performance hardware encoding, decoding and transcoding technology is the objective of Advantech VEGA 3300 PCI Express video accelerators which combine the best of the hardware and software worlds. By leveraging VEGA 3300 acceleration, multi-channel, low-latency UHD HEVC encoding can be done in a single processor 1U server where the control plane is flexibly managed by the GPP and the heavy-lifting video processing tasks are offloaded to the acceleration card.

Figure 7. Advantech VEGA 3300 Series of UHD HEVC Accelerators supporting from 1-ch 4K HEVC encoding to 8-ch 4K HEVC transcoding

There are a few aspects to take into account for successfully bringing hardware acceleration into flexible and agile x86 based developments:

- *System integration*: both from a physical and functional perspective, hardware acceleration needs to be easily integrable into server architectures for reduced time to market and minimum development efforts. The use of the PCI Express bus guarantees physical compatibility and is the most commonly used approach to add plug-and-play acceleration components to server implementations offering high speed data transfers and avoiding the need for costly application-specific mechanical designs. These PCI Express acceleration cards can also integrate video or network interfaces such as 25GbE, 10GbE, 12G-SDI, 3G-SDI or HDMI ports in the same footprint which improves PCI Express bus utilization avoiding the need for extra interface cards. From a functional point of view, easily offloading encoding, decoding and

transcoding tasks to the acceleration component is achieved through a comprehensive application programming interface (API) that enables full access to all hardware functions as if they were another software building block.

- *Standards based*: virtualization and software programmability add complexity to the relationships between different hardware and software functional blocks that have now been dissociated. Video processing components need to be available via well-defined interfaces as a pool of resources independently from their physical location. Open interfaces avoid vendor lock-in and secure the application development investment and longevity. These need to be implemented at all levels, from the hardware, middleware and operating system to the application and cloud layers.

The Software Value

Advantech VEGA accelerators and integrated systems based on Socionext processing technology have been designed from the ground up as commercial-off-the-shelf, easy-to-integrate platforms with a threefold objective: boosting video performance, improving deployment efficiency and reducing development efforts. A dedicated team of video engineers has developed the Advantech VEGA software development kit (SDK) that supports industry-standard drivers, virtualization packages, hypervisors and cloud schemes including Linux, Windows, Intel® VT-d, KVM and OpenStack. This comprehensive software package comes with plug-ins for commonly used open coding libraries such as FFmpeg. This allows for minimum modifications in case of migrating the application software to or from a different hardware platform.

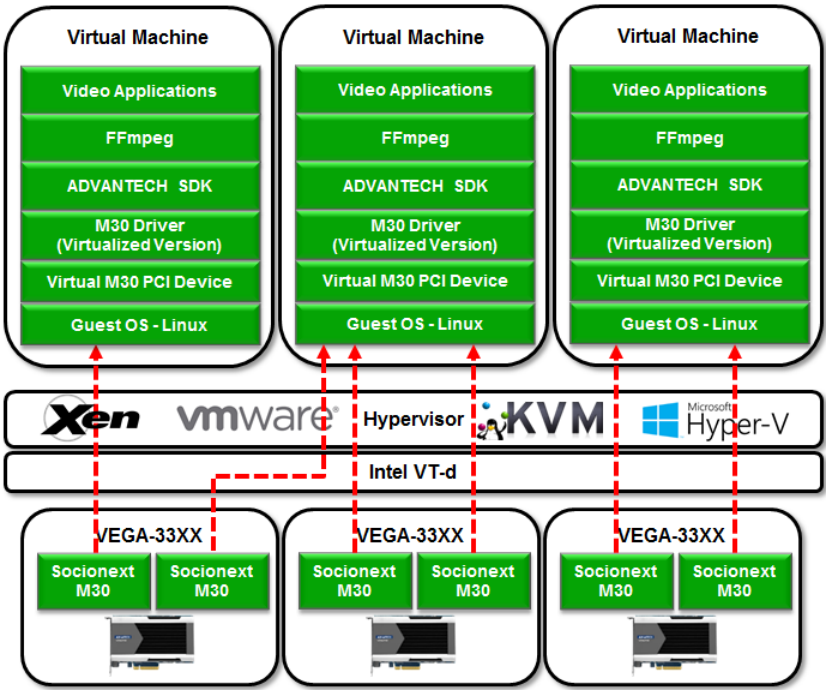


Figure 8. Block Diagram for Virtualized VEGA 3300 Series Accelerators

Advantech's extensive software package streamlines product development and reduces in-house integration efforts saving valuable software resources for application development. It allows for integration of VEGA virtualized platforms into video cloud deployments managed by OpenStack. Physical hardware resources can be assigned to virtual machines (VMs), controlled and monitored from the Advantech dashboard. VMs can discover and gain exclusive access to pre-given devices leveraging the same SDK used for non-virtualized environments and all these at virtually the same performance as bare metal implementations.

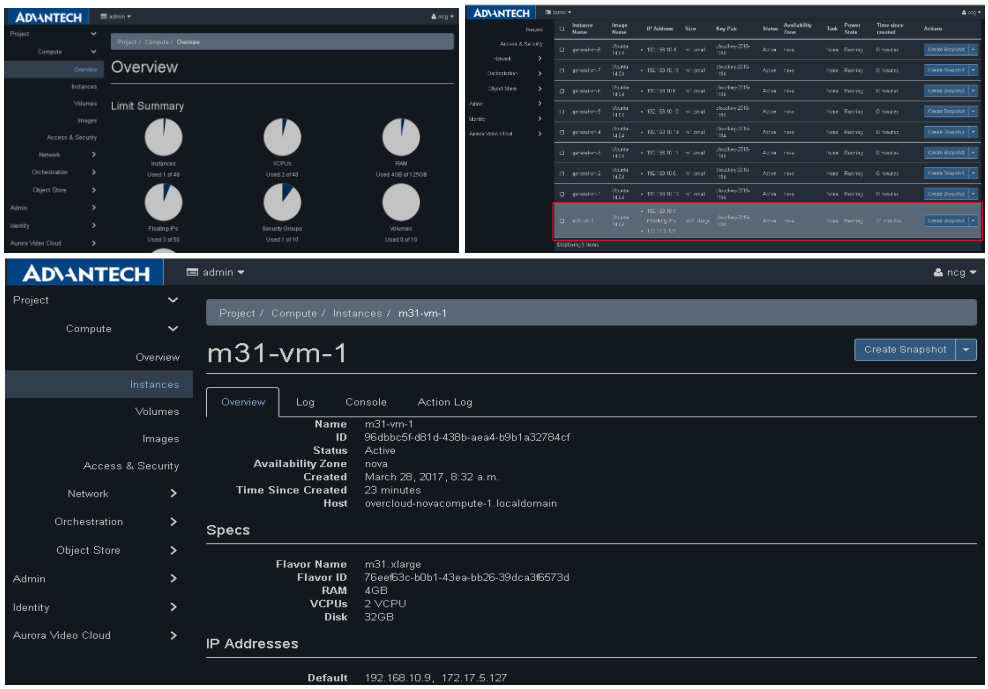


Figure 9. Advantech Control and Monitoring Dashboard for Virtual and Cloud Environments

The Customization Option

As mentioned before, the use of commercial-off-the-shelf products brings economy of scale benefits to open video solutions that can leverage powerful processing technology previously only available on specialized video platforms. However, there are cases where the one size fits all approach is not valid and different levels of customization to the standard product are required. These can go from minor middleware tweaks to mechanical changes or tailored branding. The technology provider capability to flexibly adapt to deployment or application specific requirements is an important factor to consider when choosing a long-term partner for developing complex video processing products.

Starting from commercial-off-the-shelf platforms, Advantech VEGA video accelerators and systems can be customized to meet particular physical, functional, environmental or logistic needs. These customization services include mechanical modifications, software changes, system integration, tailored branding, ruggedization and extended product lifecycles all the way up to full custom designs. Advantech owns the whole VEGA product life-cycle from in-house design and manufacturing to quality assurance and global logistics. Its strong team of hardware and software

video engineers accompanies customers in a partner relationship with the common objective of launching successful products to the market. As an example, the VEGA 7000 Series is a family of highly configurable servers that can integrate a wide range of video technologies to serve the media, telecom and internet industries. It has been optimized in density, power consumption and functionality to efficiently scale throughput of compute-intensive encoding, decoding and transcoding applications in live production, broadcast and OTT workflows.

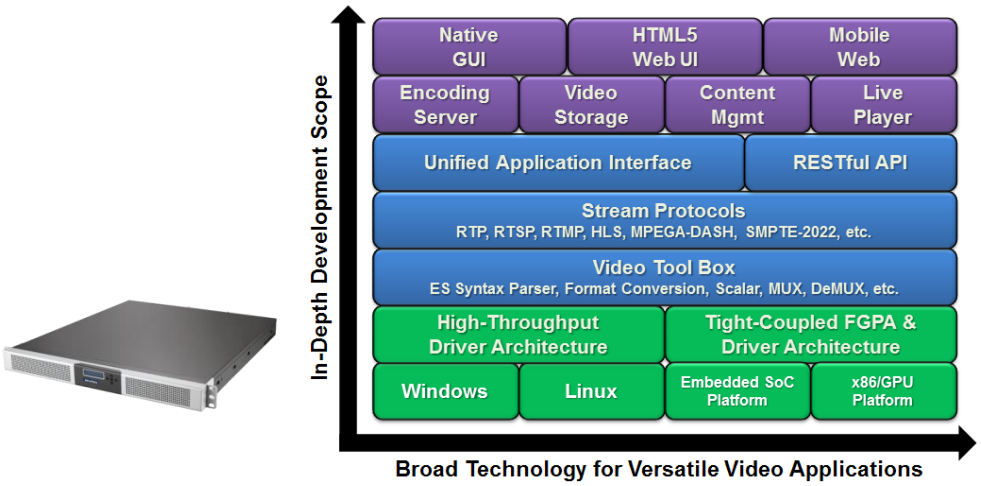


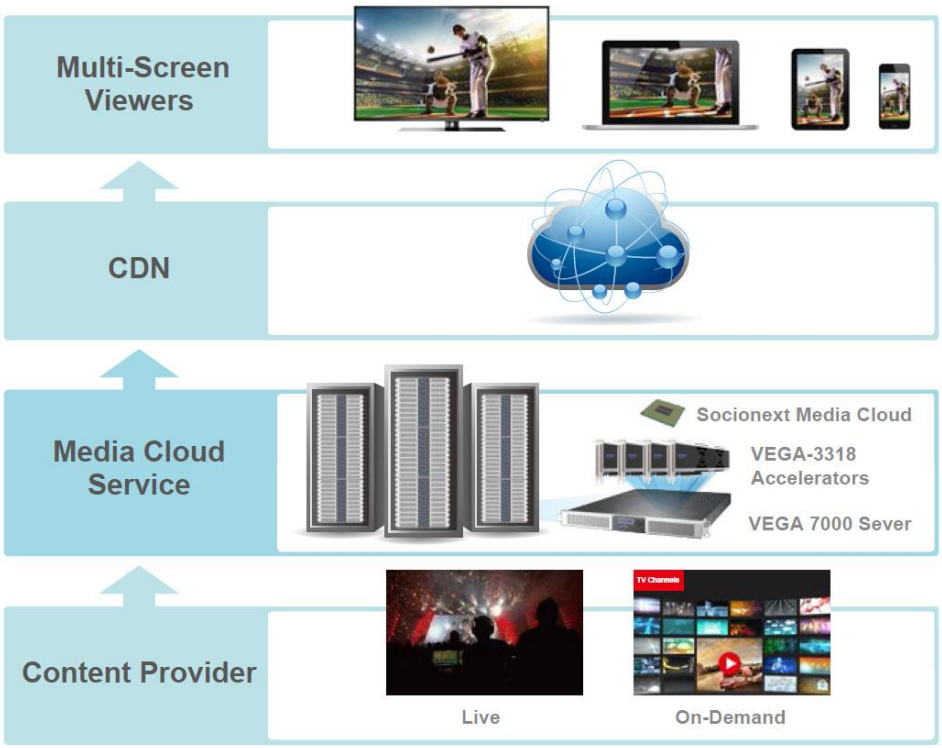
Figure 10. VEGA 7000 Series Development Framework

Use Case: High Density UHD Video Processing in the Data Center

Different aspects introduced throughout this white paper are to be considered when running video services in data center environments. These tend to be large scale deployments where power and space are major concerns. High density accelerators such as the VEGA-3318 target data center deployments providing the required performance boost to efficiently scale video processing applications running on x86-based servers, while keeping power consumption and footprint under control. The new VEGA-3318 is the first of its class able to perform professional-grade, low-latency transcoding of eight 4Kp60 video streams in an ultra-low-power and easy-to-integrate commercial-off-the-shelf PCI Express adapter that can be integrated into standard servers. It supports UHD, HD and SD formats and HEVC, AVC and MPEG-2 codecs including 10-bit profiles, 4:2:2 chroma subsampling and ABR streaming. Developers can easily integrate the VEGA-3318 into existing deployments or deploy pre-integrated VEGA 7000 servers for greener deployments with the following benefits:

- *Scalability and cost-efficiency*: VEGA-3318 accelerated solutions offer the highest density available in the market supporting low latency streaming of 32 UHD or 128 FullHD HEVC ABR bundles per rack unit with up to 20x size and power consumption reduction. This allows for energy and cost efficient cloud deployments that can meet the increasing resolution, mobility and speed demanded by online video viewers.

- *Software agility and better time to market:* service providers and equipment manufacturers can leverage Advantech's comprehensive software package that support Linux and Windows operating systems, FFmpeg, GStreamer, virtualization and OpenStack. Advantech VEGA video products are supported by a strong team of over 100 hardware and software engineers allowing VEGA-3318 customers to minimize their video processing development efforts and accelerate the roll-out of their cloud video solutions to catch the online video wave in a cost-effective and timely manner.



Use Case: Live UHD Video Processing at the Edge

The computational complexity of UHD is of great concern when having to deal with live encoding and decoding locally at the edge of the video cloud. The trend of moving video applications to the cloud opens great opportunities at the edge where services that run closer to subscribers can be optimized for improved latency, bandwidth and user experience. On the content creation part of the workflow, live production has an inherent proximity aspect to it with the requirement of being close to where the action happens. It is therefore subjected to the bandwidth, power and space constrains of field deployments.

Versatile platforms such as the new VEGA-7010 bring powerful, low-latency and efficient 4K HEVC acceleration to the edge of the video infrastructure providing high-performance and broadcast-quality media processing in a compact and low power format that can be easily deployed to run encoding, decoding or transcoding applications in regional production facilities, outside

broadcasting units, local hubs or mobile stations. The new VEGA-7010 is a highly configurable, 1U, short depth video server based on the Intel® Xeon® Processor E3-1200 v6 Product Family. Its four PCI Express slots provide high flexibility to fit a wide range of video accelerators such as FPGA or GPU cards supporting up to 300W PCIe power. When integrating Advantech's VEGA 3300 accelerators, the VEGA-7010 can perform multi-channel 4Kp60 HEVC 10-bit encoding, decoding or transcoding, including SDI and IP media capture. It supports dual redundant power supply units for high availability in a variety of edge media processing scenarios and integrates management features for remote control and monitoring.

Highest definition applications can also benefit from efficient edge computing platforms and the media industry's move to the cloud even for complex use cases such as 8K VR live streaming. The high bandwidth and massive compute required by 8K VR live production is not a concern when leveraging enabling technologies that allow for optimized bit rates and highest density video processing performance. The VEGA-7010 can be used in conjunction with the VEGA-3304 8K HEVC encoding accelerator to build hardware accelerated compact 8K live encoders for streamlined VR productions. Enjoying a reduced bandwidth, the encoded UHD streams can be sent to the cloud to be stitched and delivered as a service which minimizes onsite equipment and optimizes operations.

Figure 11. 1U 8K HEVC live encoder platform based on the VEGA-7010 server and award-wining VEGA-3304 accelerator

Conclusion

The arrival of UHD and the increasing demand for online video services are outstripping the processing capabilities of existing video infrastructure. At the same time, the industry is moving to open, cloud architectures that benefit from the programmability of software defined, virtual environments and the economy of scale of mainstream servers. Commercial-off-the-shelf access to unrivalled performance video processing acceleration such as the one provided by Advantech VEGA platforms based on Socionext technology combine the best of the video and IT worlds enabling agile, scalable and efficient deployments that can greatly contribute to the media, telecom and internet industries' objectives of improving operations and costs when building next-generation UHD video solutions.

Contact Information

Advantech

Hotline Europe: 00-800-248-080 | Hotline USA: 1-800-866-6008

Email: video.solutions@advantech.com

Regional phone numbers can be found at http://www.advantech.com/contact/

Socionext

Japan: +81-45-568-1015 | Hotline USA: 1-844-680-3453

Email: datacenter@us.socionext.com

Regional phone numbers can be found at http://www.socionext.com/en/contact/