

Virtual Video Transcoding in the Cloud

Jim Darroch, Senior Director, Software Product Marketing, Artesyn Embedded Technologies

Richard Dunphy, Software Architect, Intel

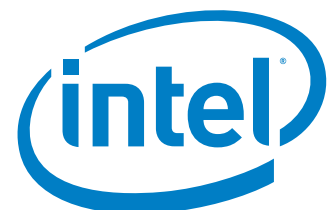
Franklin Flint, Global Telecommunications Strategist, Dell

The increasing density and high-quality processing demands from video applications is pushing broadcast and communications networks to the limit. Adding more equipment to handle these video streams is not economically viable. What's more, operators, service providers and content providers see the benefits of using standard servers in the cloud, and want to move away from special appliances or dedicated hardware. But standard servers currently are not optimized for video transcoding in the cloud.

The best solution is an accelerated video cloud: PCI Express video acceleration cards in standard servers.

The accelerated video cloud offers the ubiquity of standard server-based resources, but the added benefit of higher-performance and higher-density video processing needed to support today's users.

This paper presents the market drivers behind the need for an accelerated video cloud and discusses the implications of technologies such as OpenStack and other open source platforms in the context of a move towards SDN/NFV network architectures. Real-life example applications then give context to the performance and financial benefits outlined in the conclusion.



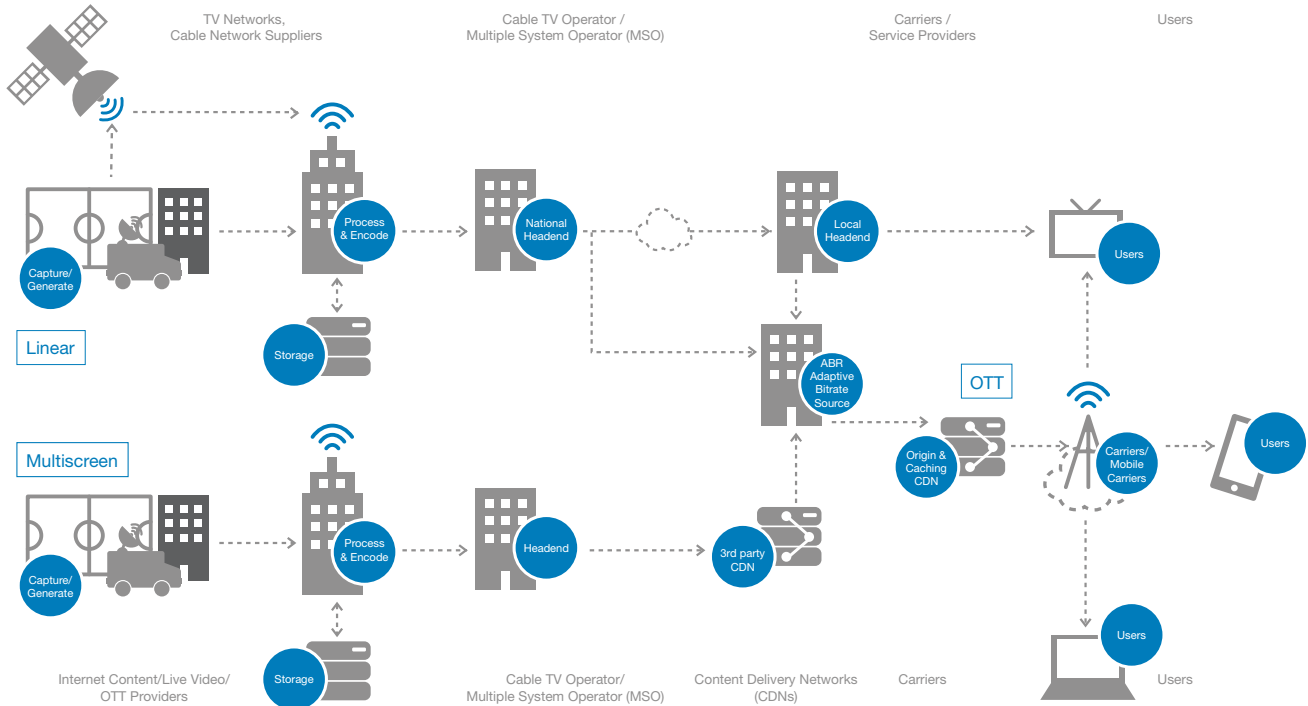
Virtual Video Transcoding in the Cloud

Contents

- Virtual Video Transcoding in the Cloud** 1
- Current Cloud Capability to Handle Video Traffic** 3
- The Solution – Accelerated Video Cloud** 4
 - The Desire to Use Standard Servers 4
 - Using Intel Processor Graphics 4
- OpenStack and What It Enables** 6
- SDN/NFV Standardization** 7
- SDN/NFV Control of Multiple Video Processing Resources** 7
- Example Applications: OTT Video Streaming** 8
 - Next Day TV OTT Content Delivery 8
- Conclusion: Benefits of This Approach** 11
 - Benefit #1: Reduced Capital Equipment Spending 11
 - Benefit #2: Power Savings and Reduced Overhead Cost 11
 - Benefit #3: Scalability 12
 - Benefit #4: Ease of Use through Ubiquity of x86 processing in the Cloud 12

Current Cloud Capability to Handle Video Traffic

Consumers are the major driving force behind the changing broadcast network landscape. User habits are changing from traditional broadcast to television video consumption (see top “Linear” segment of chart below) to dynamic, on-demand and mobile video viewing (see bottom “Multiscreen” segment of chart below).



This puts a strain on the networks, which are rapidly expanding their capability to handle both traditional cable/linear/broadcast video distribution and the more dynamic multiscreen video. Multiscreen video is driven by consumer preferences versus those of the operators. The bandwidth consumption of video cannot be ignored, especially judging by the size of its share of the network.

As an example Sandvine’s “Global Internet Phenomena Report”, June, 2014 reports real-time entertainment is responsible for over 63% of downstream bytes during peak period.

Peak Period Traffic Composition (North America, Fixed Access)

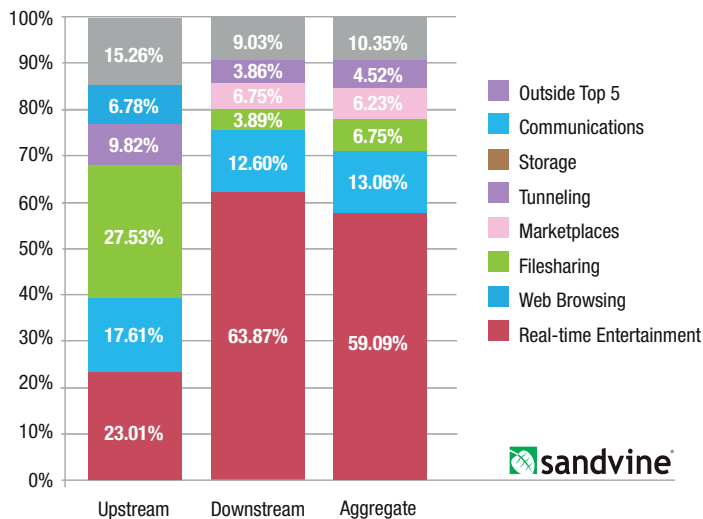


Figure 1 - Peak Period Aggregate Traffic Composition - North America, Fixed Access

The landscape of the carrier business is also changing for video, with a growing preference by multi-service operators (MSOs) and telecom service providers for the use of cloud technology to provision adequate processing, streaming and storage resources for the media market. However, the current performance level of video in the cloud is far from optimal.

Using the latest technologies from Intel®, Dell and Artesyn Embedded Technologies this paper will explore better ways to host video in the cloud, to ultimately achieve higher densities, with more flexibility in deployment, through virtual transcoding solutions.

The Solution – Accelerated Video Cloud

Typical cloud implementations are based on multiple identical servers that can be used as needed for varying tasks, with different scale and footprint requirements as specified by the application. The processors are NFVI (Network Function Virtualization Infrastructure) nodes in a sea of cloud computing resources.

In the case of video though, a server used for video streaming supports far less user capacity than a server used for web browsing. For this reason, adding video processing and transcoding resources to cloud servers is required.

With a growing preference by operators to use standard servers in the cloud, without the need for special appliances, and the increasing density and high-quality processing demands from video applications, the best solution is an accelerated video cloud. The accelerated video cloud offers the ubiquity of standard server-based resources, but the added benefit of higher-performance and higher-density video processing needed to support today's users.

The Desire to Use Standard Servers

Over the past 10 years, developments in cloud data center solutions have created an extremely cost effective business model both in terms of capital and operating expenditure. Using commercial-off-the-shelf (COTS) hardware and developing highly reliable and scalable software, which can be controlled with intelligent orchestrators, traditional data centers have become the norm for enterprise and over-the-top (OTT) Internet solution providers. The benefits of this model have yet to be realized in the telecommunications network due to long standing requirements for reliability, manageability, and availability, which carry a greater priority.

The effort to adopt well-established cloud technologies in the telecom network and address the unique requirements of that network was instigated by carriers and service providers under the auspices of the ETSI Network Functions Virtualization working group (more on that later) and now the solution providers are developing NFV solutions to bring cloud networking solutions to the carriers.

The benefits of this approach go beyond just offering lower costs for new paths to market. By taking advantage of the scale of COTS hardware and software, the network will require fewer resources to design, deploy, manage, and support. It will be easier to maintain using proven solutions. In addition, new services running in virtual platforms can be tested, piloted, and rolled out across the network more quickly, giving the carrier a significantly faster path to revenue and pleasing their customers who get access to new technologies sooner. The virtual network also allows for greater scalability. Improvements in power consumption, management, and support can be driven across data centers and, as an example, Dell platforms and services have been honed to provide the ideal balance of performance, density, cost of operation, and support that carriers desire.

The core operating component to these network data centers is the standard server. Standard servers are designed to provide a compelling financial situation, both in terms of capital and operating expenditure. They are often optimized for high reliability, performance and density and can be capable of operating in data center or telecom central office environments. For example, Dell's PowerEdge™ R720 2U COTS server, powered by two Intel® Xeon® E5 processors is offered with either AC or DC power supplies, standard configurations or NEBS Level-3 & ETSI compliant options to meet the various requirements of the telecommunications infrastructure deployment locations. Working closely with Artesyn, for the purpose of this white paper and solution, Dell recommended the carrier grade version of the PowerEdge™ R720 with NEBS L-3 compliance to meet the demanding requirements of the carrier network while providing the compatibility with the standard version which could be used in modern raised floor data centers. By combining the Dell PowerEdge™ R720 with the accelerated video cloud methodology proposed in this white paper, the benefits of both using standard servers and dramatically increasing network video processing performance can be achieved.

Using Intel Processor Graphics

In 2011, Intel® introduced Quick Sync Video (QSV) on their integrated processor graphics with the Intel 2nd Generation Core product line. QSV builds on the already available decoder hardware from previous generations of the Intel® HD processor graphics to include a flexible architecture for encoding of H.264 with video quality improvements and performance improvements in each future generation.

In 2013, with the Intel 4th Generation Core product line, Intel® introduced the Iris Pro processor graphics which added MPEG2 encoder capability, increased video quality with the addition of more motion estimation engines, and 128MB of embedded DRAM providing a high bandwidth memory capability for the GPU (70GB/s compared to 25GB/s for the dual channel DDR3 memory interface).

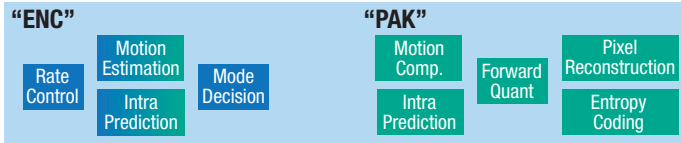
Quick Sync Video provides many advantages over other architectures when considering transcoding. Traditional hardware fixed functions provide solutions with low power, and high performance, but are unable to move quickly should there be an update required for the code. While software codecs provide complete flexibility at the expense of power and performance.

By utilizing hardware for the functions of a codec that don't change and software running on compute elements in the GPU for functions that can, Quick Sync Video enables a balanced solution to deliver high performance at low power while maintaining much of the flexibility required to improve the codec over time.

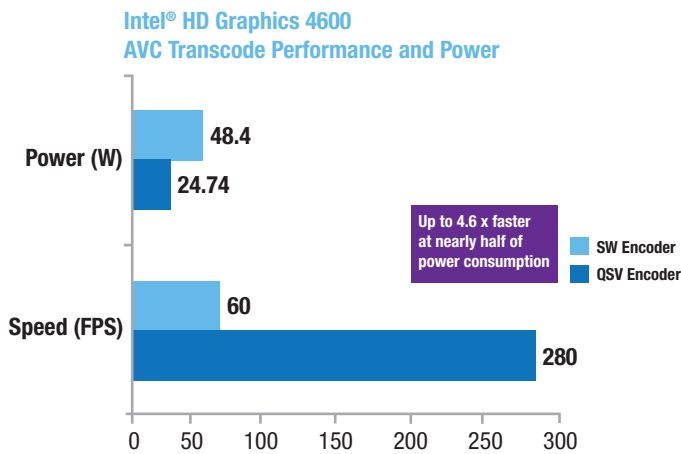
White	Performance	Power	Flexibility
Software	Low	High	High
GPGPU	Low	High	High
Traditional Fix Function hardware	High	Low	Low
Flexible Intel® Quick Sync Video	High	Low	Balanced

The Quick Sync Video solution consists of two main elements. The “ENC” which comprises of a hardware acceleration called the Media Sampler which provides for efficient motion search, and software running on the programmable Execution Unit array. And the “PAK” which reuses logic from the Multi-Format Codec Engine and provides a complete hardware unit to do the pixel reconstruction, quantization, entropy encoding, etc.

HYBRID 2-STAGE VIDEO ENCODER

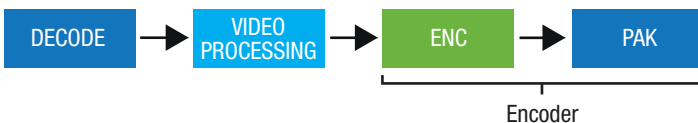


This combination of hardware and software allows Quick Sync Video to greatly reduce the power requirements for transcode on Intel Architecture while simultaneously increasing the performance capability.



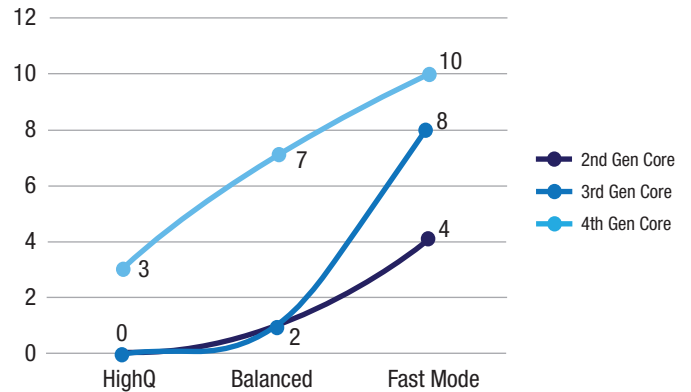
Internal measurements based on demo clip using Cyberlink* Media Espresso v6.7.3521, Fast Conversion Mode running on Intel® Core™ i7-4900MQ reference platform under Windows* 7 operating system.

The combination of these blocks with hardware acceleration for decode through the Multi-Format Codec Engine and video post processing such as Deinterlacing, DeNoise and some other filters, provides a four-stage transcode pipeline with all stages capable of running concurrently for high density of live/linear simultaneous transcodes as well as many times faster than real time or offline transcoding.

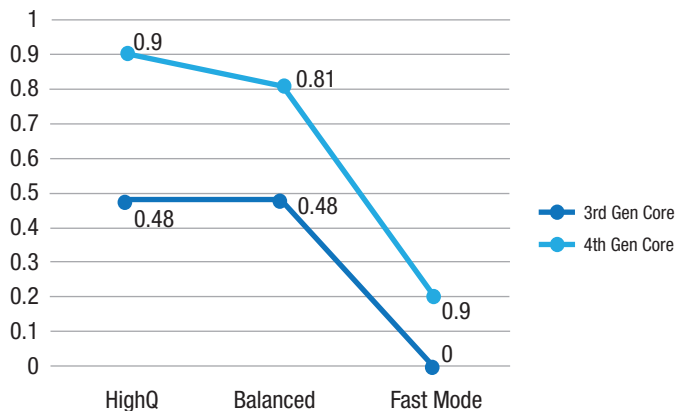


With a flexible encoder pipeline, Intel® is positioned to improve both performance and video quality from generation to generation. Between Intel® 2nd Generation Core and Intel® 4th Generation Core the performance in a “balanced” (between video quality and performance) mode has increased greatly while at the same time increasing the video quality at that mode. In general QSV in the Intel® 4th Generation Core has improved video quality at all levels over and above the previous generation.

Generational Performance improvements of Quick Sync Video for 1080p30 transcodes



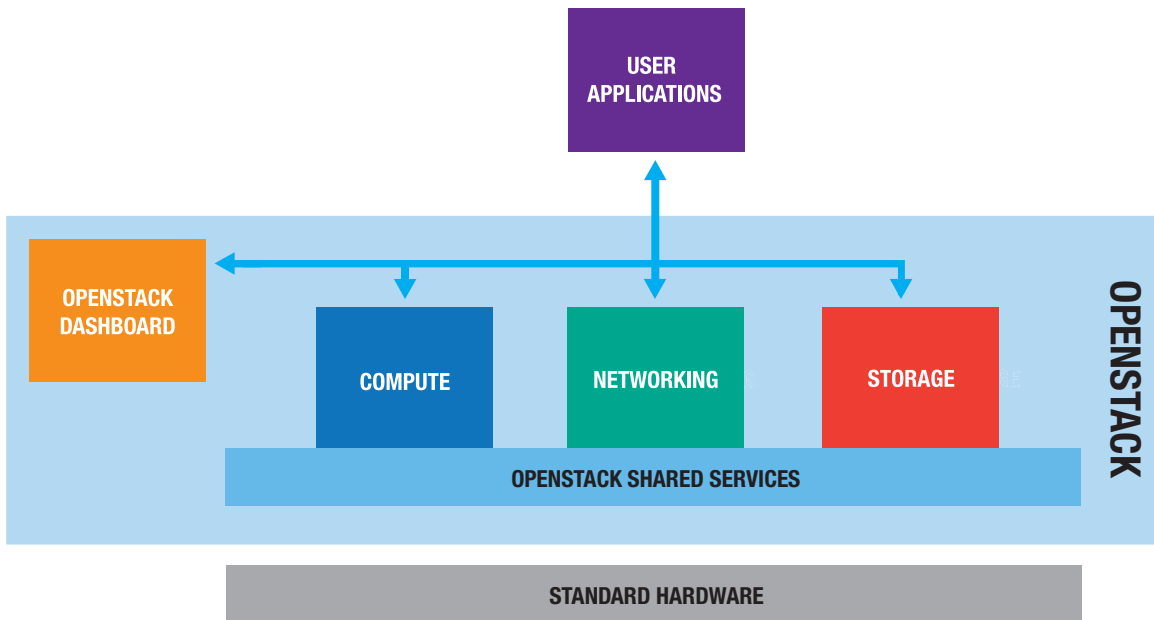
Video quality improvements for CQP (BD_PSNR)



CQP test: 163 progressive clips with different resolution. Average 50 frames in length. QP1=[22, 27, 32, 37], QPP=QP1+2, QPB=QP1+4. 3 B frames

OpenStack and What It Enables

OpenStack (openstack.org) is an open source cloud computing platform, which has seen unprecedented growth and support for implementing Infrastructure as a Service (IaaS). Since its introduction in mid- 2010, OpenStack has seen the involvement of more than 200 hardware, software and services providers.



OpenStack, from the top down, consists of:

- Application APIs which interfaces directly to users applications above, and to the resource layers below
- Resource layers, which provide a suite of standard resource elements (including compute, storage, and networking). These are the resources which are visible to the applications and management/orchestration applications which consume them, and are abstracted from the physical implementation in hardware
- Shared services – the glue layer between the hardware and the resource interfaces presented above. This is where virtualization exists – the hypervisor, which provides access to virtual machines (VMs) and guest operating systems, is a shared service amongst all managed compute elements
- The OpenStack dashboard - a management layer which allows the cloud service provider to assign resources to user applications (and provide auxiliary functions – such as billing)
- The shared services and (most importantly) resource elements reside on ‘standard hardware’

The objective of OpenStack is to provide a managed set of resources, completely independent of the underlying hardware. There are many advantages to this approach - services can be transitioned to alternative servers as network conditions and service demands change.

For many applications, this abstraction works very well. Some applications, however, have particular demands on compute or storage or network resources that tie them to more specific hardware instances. One example: video transcoding.

A note on ‘standard’ hardware - ‘standard’ does not mean one size fits all. Each physical resource must provide an agreed level of performance (be it CPU, network bandwidth, storage capacity - or ‘special’ hardware requirements).

These resources are presented in the management dashboard and can be assigned to user applications on demand. Applications must present a ‘profile’ of resources required to run; the dashboard allows service providers to assign matching resources to the application.

In the case of video transcoding, a pool of video-optimized resources will be made available via the dashboard. Policies are established with each user/application to govern access to the transcode functions, and this is enforced by the dashboard. Policies are derived from service levels, and are extremely flexible - providing elasticity of capacity, such as:

- Variable resource utilization, up to a cap
- Temporarily allow excess capacity
- Excess capacity provided at a premium cost

OpenStack has become of interest to Carrier Service Providers looking to implement Network Functions Virtualization (NFV) – a term applied to the decoupling of Network Applications from their underlying hardware. In other words, ‘Cloud Infrastructure for Telecom Applications’.

SDN/NFV Standardization

ETSI has established an Industry Standardization Group (ISG) to examine the need for NFV standards. Although OpenStack comes from the enterprise world, the concept of centralized orchestration of virtualized resources will be a key component of emerging NFV standards. It is probable that OpenStack, or a carrier grade version of it, will emerge as an essential technology.

The ETSI NFV ISG has produced Informative work on the structure of NFV – from infrastructure up to management and orchestration. As such, this is a description of recommended practice, rather than any standard on how NFV is implemented or how equipment and software from multiple vendors, utilized by many users and run by many service providers will actually interoperate. However, it is almost certain that the NFV ISG will continue its work for another two year term, with the goal of publishing normative standards for NFV.

The ISG itself is made up of a global who's-who of service providers, equipment manufacturers and independent software vendors. It's fair to say that NFV has very broad and committed industry support, and will be successful in recommending standard approaches to implementing a number of applications, including video.

SDN/NFV Control of Multiple Video Processing Resources

The orchestration of multiple, heterogeneous video resources available to many users is not an easy task. OpenStack has proved to be massively scalable in an enterprise cloud environment – the expectation is that any emerging NFV standard will do the same.

However, video as an application is different enough from enterprise cloud applications to merit some serious thought. Video delivery is the 'perfect storm' of resource consumption:

- Large volumes of storage are required to maintain a database of video content
- Transcoding video streams from source format into final delivery form (bit rates, video format, screen size, etc) is computationally expensive
- Traffic delivery towards end users is pretty-much real time; available bandwidth has to match the volume of traffic generated by the transcoders.

So, an orchestrator has to be aware of resources available to transcode video, and also the bandwidth required to get the resulting video packets through the network. This is a concern at the network edge (final delivery to a consumer device). It is also a concern for intermediate processing, where original, centralized content (generally from a producer or broadcaster) is transcoded and pushed towards multiple instances at the network edge (as close to the end consumers as possible).

One other consideration is the assumption that OpenStack controls ('orchestrates') virtual resources – in essence, virtual machines (VMs), made

available via a hypervisor layer which abstracts the application execution environment from the underlying hardware. A server can support many VMs and the resource is considered elastic in scale.

Again, video presents a problem. If transcoding is offloaded from a host CPU to an accelerator, then the OpenStack orchestrator needs to know that the accelerator is available (and is video-capable). Worse still, the accelerator architecture does not typically use VM technology – instead running directly on the CPUs primary operating system ('bare metal' has become a common term to describe the non-VM model).

How does OpenStack orchestrate 'execution' resources which map directly to hardware? Fortunately, there is a solution – OpenStack has a plugin (named 'Ironic') to orchestrate bare metal resources. Its northbound API is identical to the interface for managing VMs, but the southbound interface is aware that it manages a single hardware resource.

In time, the ETSI NFV group will standardize the interfaces and infrastructure required to do this. In the meantime, the complementary technologies of OpenStack and SDN plug the gap. While OpenStack allows the orchestration of resources, SDN makes use of the OpenFlow protocol to configure network switches to provide interconnect capacity consistent with the volume of video traffic to be transported. An SDN controller, such as OpenDaylight, helps with the orchestration of traffic flows.



The alternative is to simply provision the 'worst case' compute and network resources needed to process video to be 'always on'. This will always lead to excess capacity (and hence cost) in the network as resources are provided which may not be used the majority of the time.

In the Hulu model, video is processed and transported 'offline' in batches every day. The case of 'mass consumer events' is even more extreme – for example large sporting events, where live footage has to be cached and processed in real time.

By using a combination of NFV (OpenStack) and SDN, resources are only consumed – and paid for – when they are used. The elasticity of resource availability means that unexpected levels of demand can be catered for – without the need for up-front over-provisioning.

The existing technologies used here – OpenStack, OpenFlow, OpenDaylight are all open source projects, free to use for developers to implement these services.

Example Applications: OTT Video Streaming

Next Day TV OTT Content Delivery

A growing use case for video transcoding in the cloud is next day, or catch-up TV, where a content provider receives content for the day from various producers and makes it available to subscribers to view the following day. Such a provider can be receiving hundreds of hours of content every day that needs to be transcoded to multiple different formats for delivery to a multitude of devices. The result can be the need to transcode thousands of hours of video.

Take an example where the provider is receiving 200 hours of content from various producers. Depending on the devices supported, the provider could produce up to 100 different transcoded outputs of that content to address different codecs, resolutions, bitrates, etc. of any consumer devices it permits.

To make this example easier, let us assume for now the provider will perform 10 different outputs of 1080p30 H.264. Running on a standard 1RU dual-processor server, configured with dual Intel® Xeon® E5-2650V2 devices, the server should be capable of approximately 60 frames per second of transcoding per socket with X.264 (extrapolated from 33fps with fast mode, default CRF on 3.2GHz Intel® Core™ i7-4770R), or 120 frames per second per server when running without virtualization. However, in a cloud environment the transcoder will run in a virtual machine so we need to derate this number by about 10%, thus a total of approximately 108 frames per second, per server.

With 200 hours of content at 30 frames per second, the system needs to transcode 216 million frames across the 10 outputs. At a rate of 108 frames a second, a dual Intel® Xeon® E5-2650v2 server would require 556 hours to perform the task. Not really effective for next day TV. Using dual E5-2650V2 2RU servers, such as the Dell R720T, the workload would require 24 servers (>1 rack) running 100% flat out 24/7 to ensure the content could be delivered to consumers within 24 hours. Running flat out 24/7 is a sure way to cause failures in the data center, thus more servers are required to even out the load for reliability.

Comparison, Dell R720T with 2x Intel® Xeon® E5-2650V2 processors with/without 4x Artesyn SharpStreamer™ cards:

Next Day TV				
	BareMetal		Virtualized	
	E5-2650V2	i7-4650U	E5-2650V2	i7-4650U
FPS/Node	60.0	180.0	54.0	180.0
Number of sockets per board	2.0	4.0	2.0	4.0
FPS/Board	120.0	720.0	108.0	720.0
Number of boards per server	1.0	4.0	1.0	4.0
FPS/Server	120.0	2880.0	108.0	2880.0
Hours	200.0			
Streams	10.0			
Total FPS	216000000.0	216000000.0	216000000.0	216000000.0
Seconds Required	1800000.0	75000.0	2000000.0	75000.0
Hours Required	500	21	556	21
Num Servers	21	1	24	1
Improvement	21		24	
Power Per Server (W)	475	1056	475	1056
Total Power (W)	9979	1056	11405	1056
Power Improvement	9.45		10.8	

The alternative is to use the Artesyn SharpStreamer™ card in such a system. With 4 nodes of Intel® Core™ i7-4650U capable of each delivering 120-240 frames per second of 1080p transcode, a provider can squeeze much more efficiency out of each server. In this configuration, a single server with dual Intel® Xeon® E5-2650V2 and four Sharp Streamer cards can effectively provide almost 4000 frames per second when coupled with software on the CPU cores. For the purpose of comparing to the Intel® Xeon® E5-2650V2 software solution, we'll focus on the median of 180 frames per second per node at balanced quality mode (Intel(s) MediaSDK target usage=4), and thus 2880 frames per second processing across the quad PCIe cards. Such a solution would be able to process the 200 hours of content for the 10 individual outputs within 21 hours on a single server, a reduction of 24 times the servers, 11 times the power, and over 5 times the cost.

While 10x 1080p30 transcodes may not be truly representative of such deployments, it's conceivable that a provider would need to provide much more compute, e.g. a 1080p30 is roughly equivalent to a single 720p60. It should also be noted that 200 hours represents just a fraction of the total hours received by many of the content providers.

Real Time / Linear ABR Broadcast Transcoder Needs

Live TV viewing habits change throughout the day for consumers. IPTV providers today must cater for delivery to not only a known entity in their set top box, but also need to accommodate the multitude of devices on which consumers now watch their content, e.g. tablets, phones, 3rd party boxes like Roku™, Apple TV, and Amazon's Fire TV. Broadcast TV providers face similar challenges in providing their online TV portals. The outcome is that IPTV providers now need to be able to generate a lot of various transcoded formats in real time with minimal delay.

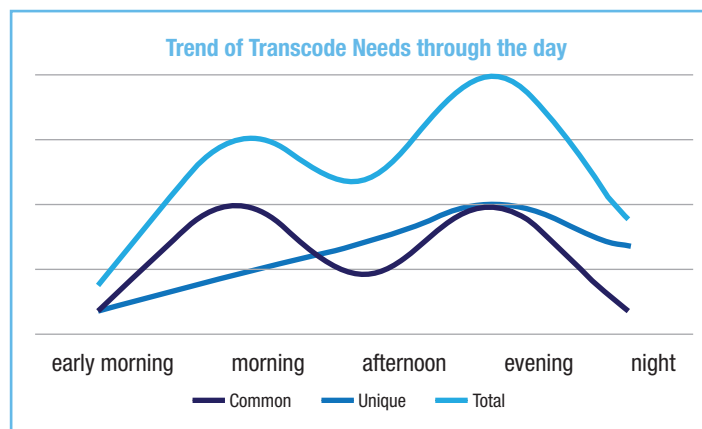
To accommodate congestion in the network most providers have switched to an adaptive bitrate technology, such as Apple's HLS, Silverlight, PlayReady, etc, which allows the consumer device to determine if it needs to switch to a different profile to ensure the content continues to play. In many cases consumers will tolerate a momentary drop in video quality, but re-buffering generally causes consumers to change channel, or change provider. Adaptive streaming attempts to help consumer devices to adapt to the changes in network speed and bandwidth by segmenting the video into multiple chunks of a certain time period, e.g. 2-4 seconds, and make these available to the client in a pseudo playlist called the manifest.

The manifest provides the client with data to show what profiles are available for certain time indexes, and what the necessary file to request is. The consumer device requests the file for the profile it desires and monitors the download time, if the time fails to meet the desired time to maintain play rate the device will request a file for a lower profile and monitor that,

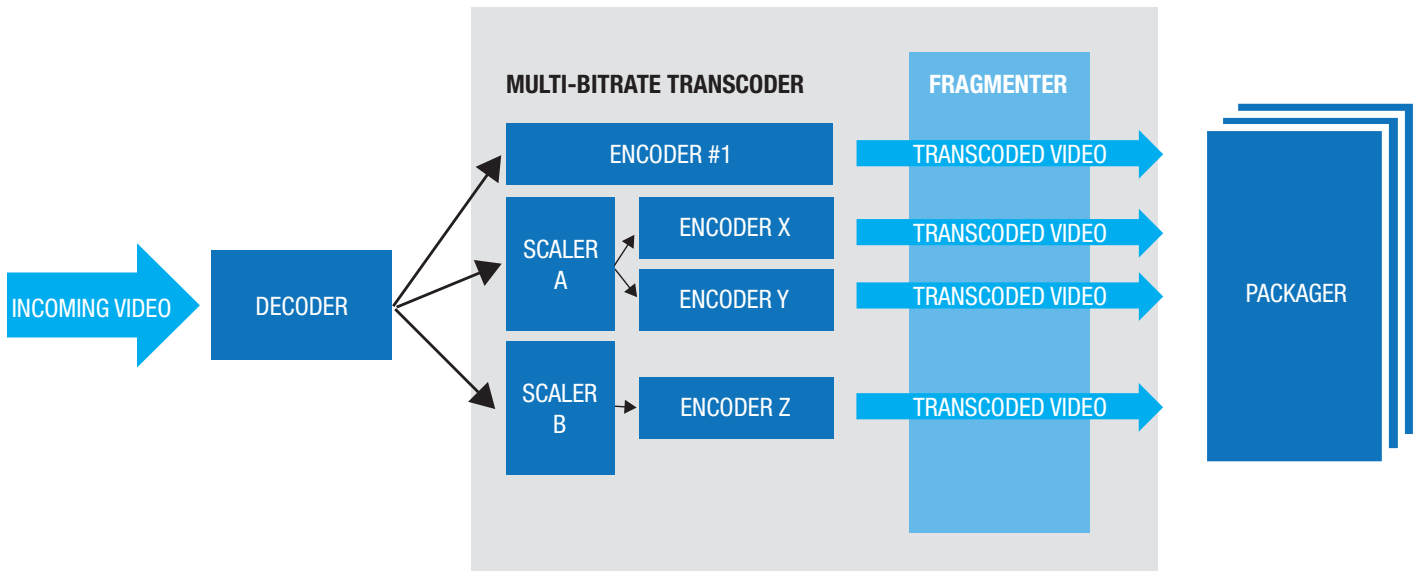
eventually a re-buffer may be required, but a well configured device will manage to get a profile downloaded in time for the player before re-buffering is required unless there are serious problems with the network.

The downside of adaptive streaming is the need to create the different profiles. In many cases the provider will need to handle multiple adaptive streaming technologies for the devices it wishes to target, but also needs to accommodate different resolutions, codecs profiles, bitrates and more that the various devices will support. This can lead to a lot of transcodes required for a single channel. A worst case scenario is that every device type permitted to access the content is online and accessing the content across every channel. The more channels there are the less likely it is to occur, but providers need to be aware of that peak number when planning their network.

Generally, throughout the day there will be a common set of transcodes that most if not all devices require and they can be wrapped in the various adaptive streaming protocols required as needed. There is another set of transcodes that are required to deliver the unique renditions required for specific devices, however this set is much more dynamic depending on the viewing habits. For example, many people use their TV with a set top box, or other when waking up, then move to more portable devices such as laptops, phones, etc.



When considering the trends across a number of channels there will be peaks on one set of channels, while there are troughs on others. Using virtualization on an Intel® Xeon™ based server, the system can bring online more transcoders as required, and configure them to produce the renditions required for the various devices of interest by implementing a multi-bitrate transcoder that decodes the incoming video, scales to required resolutions and encodes to the particular formats before sending to a fragmenter to partition the stream into segmented files and then finally forwarding to the packager(s) to wrap in the required packaging for the adaptive bitrate standard required by the consumer devices.



For an efficient multi-bitrate transcoder, the decode of the video should occur once and be used as a single reference for all encode outputs, and encoders can optimize to reduce the overheads of scaling across various output resolutions and motion search on those resolutions.

It is important that every output from the encoders is group of picture (GOP) and order (encode vs. display) aligned so that the resulting segments from the fragmenter are aligned correctly before handing to the packager.

The challenge for such a multi-bitrate transcoder server running software transcoders will be to ensure all the various renditions required are generated at the single server. If the renditions required exceed the capacity of the server the system will need to replicate the decoder for the incoming video to remove the need to pass raw baseband video between the systems (adding further latency) which can be require significant network bandwidth (approximately 500Mbps for 1080p30 8bit YUV content) per stream. Also the two systems will need to remain in synchronization to ensure the output renditions are GOP and order aligned, which is critical for successful fragmentation.

By using a system enabled with Artesyn SharpStreamer cards, the density provided allows for many more renditions, and many more channels to be accommodated on a single server. Where a Dell RT720p Dual Intel® Xeon® E5-2650V2 processor system can potentially output six separate 1080p30 streams, the same system outfitted with four SharpStreamer cards can accommodate up to 96 separate 1080p30 streams, a 16x increase in transcode capacity per server.

Also the power demands are seven times less on a SharpStreamer accelerated platform, requiring 1056W versus 7604W for the 16 servers it would require to process 96 streams.

A system enabled with SharpStreamer cards allows a provider to quickly adapt the network to the on-demand needs of the consumer devices.

Conclusion: Benefits of This Approach

Using the two scenarios described above there are many benefits that can be achieved through video acceleration in a virtual network.

Benefit #1: Reduced Capital Equipment Spending

The benefits of an accelerated approach mainly stem from the reduction in server footprint to the datacenter, and the reduced complexity to manage those resources. Network Function Virtualization enables providers to change the type and level of resources needed dynamically, and this applies to video transcoding as the VNF in the use cases above.

Number of Servers Required	Virtual Servers Required	SharpStreamer Card Accelerated Virtual Server Required
Next-day TV OTT Content Delivery (200 hours of content in 10 different formats)	24	1
Capital Equipment Impact of Acceleration: 83% less equipment cost		
Real-Time Broadcast ABR Transcoding (96 1080p streams)	16	1
Capital Equipment Impact of Acceleration: 74% less equipment cost		

The savings to the service providers in terms of CAPEX equate to spending 74-83% less on equipment alone.

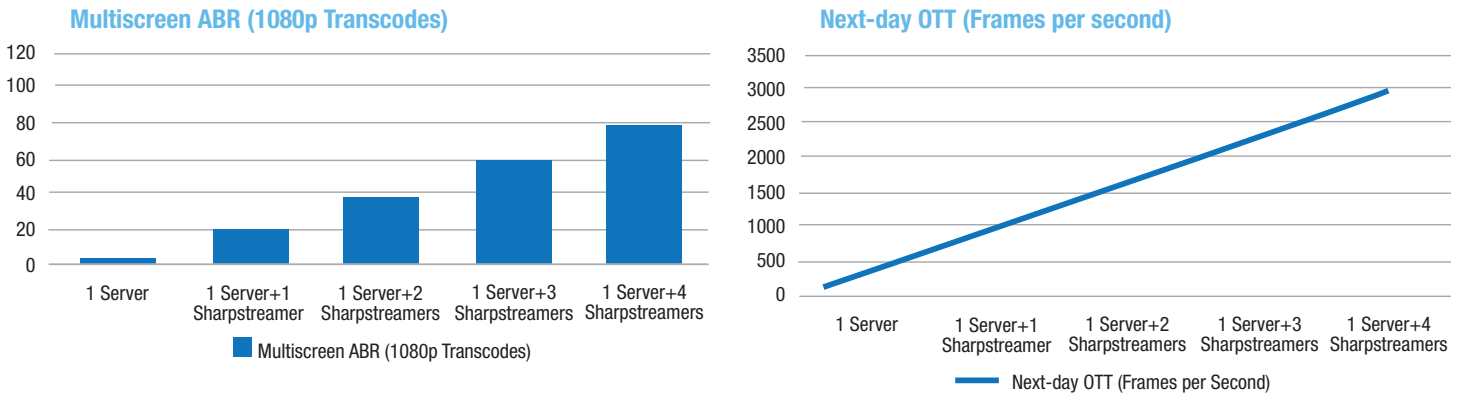
Benefit #2: Power Savings and Reduced Overhead Cost

Number of Servers Required	Virtual Servers Required Power	Power Costs per Year	SharpStreamer Accelerated Required Power	Power Costs per year
Next-day TV OTT Content Delivery (200 hours of content in 10 different formats)	11405W	\$9,991	1056W	\$925 (1.056kW*(hours/year)* \$.1 (energy unit cost/kWh)* 8760 (hours per year)
Power Cost Impact of Acceleration: 91% less power cost				
Real-Time Broadcast Multiscreen Transcoding (96 1080p streams)	7604W	\$6,661	1056W	\$925
Power Cost Impact of Acceleration: 86% less power cost				

The savings to service providers in terms of OPEX equate to spending \$925 per year versus \$6,661 or \$9,991 on an annual basis, a savings of 86-91% less power costs.

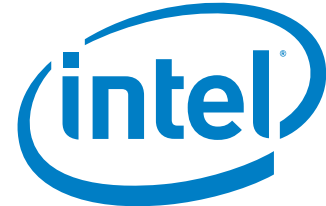
Benefit #3: Scalability

When network demands increase or decrease for video transcoding, this also allows scaling up and down of resources with lower cost, as the number of video transcodes can be addressed through add-on cards to a lower population of servers. Having a lower population of servers in the network contributes meaningful operating cost reductions as noted above. So, as service providers increase their services to provide premium OTT video services, add-on cards can gradually increase the density levels required without capital equipment expenditures as significant as traditional methods have offered to date.



Benefit #4: Ease of Use through Ubiquity of x86 processing in the Cloud

An x86-based way to solve the problem of video processing in the cloud has an important benefit for equipment vendors, in that Intel technology delivers a familiar and easy to use API to speed development and time to market. The Intel® Media SDK enables the transition from a pure software model to a media-offloaded acceleration model with the same capability to run Windows, Linux, QuickSync video and API libraries - yet in a higher-density capacity that delivers the maximum number of streams per rack unit for video applications.



© Copyright 2014 Artesyn Embedded Technologies, Inc., Dell Inc and Intel Corporation. All rights reserved.

Trademarks

Artesyn Embedded Technologies, Artesyn and the Artesyn Embedded Technologies logo are trademarks and service marks of Artesyn Embedded Technologies, Inc. Dell and the DELL logo are trademarks of Dell Inc.

Intel®, the Intel logo, and Intel® Xeon® are trademarks or registered trademark of Intel® Corporation or its subsidiaries in the United States and other countries.

All other product or service names are the property of their respective owners.

Reproduction of this material in any manner whatsoever without the express written permission of the authors is strictly forbidden.

Notice

While reasonable efforts have been made to assure the accuracy of this document, the authors assume no liability resulting from any omissions in this document, or from the use of the information obtained therein. The authors reserve the right to revise this document and to make changes from time to time in the content hereof without obligation of the authors to notify any person of such revision or changes. Electronic versions of this material may be read online, downloaded for personal use, or referenced in another document as a URL. The text itself may not be published commercially in print or electronic form, edited, translated, or otherwise altered without the permission of the authors. It is possible that this publication may contain reference to or information about products (machines and programs), programming, or services that are not available in your country. Such references or information must not be construed to mean that the companies intend to announce such products, programming, or services in your country.

Limited and Restricted Rights Legend

If the documentation contained herein is supplied, directly or indirectly, to the U.S. Government, the following notice shall apply unless otherwise agreed to in writing by Artesyn. Use, duplication, or disclosure by the Government is subject to restrictions as set forth in subparagraph (b)(3) of the Rights in Technical Data clause at DFARS 252.227-7013 (Nov. 1995) and of the Rights in Noncommercial Computer Software and Documentation clause at DFARS 252.227-7014 (Jun. 1995).

Virtual Video Transcoding in the Cloud white paper R1D1