RADVISION®
Delivering the Visual Experience®

SCOPIA®

White Paper

Delivering a Truly Scalable Visual Communications Solution

Scalable Video Coding Demystified –
The H.264/SVC Standard – Scalable Video Coding Benefits –
The RADVISION Solution

RADVISION®
Delivering the Visual Experience®

SCOPIA®

White Paper

The H.264/SVC standard is enabling the adoption of Scalable Video Coding in the visual communication market today. Scalable Video Coding (SVC) offers many benefits for existing video-based applications, as well as for future developments.

RADVISION introduces the first complete SVC-based solution, enabling a high quality visual experience for a wide-range of new and legacy end-point devices that can deliver high quality video under diverse network conditions.

The RADVISION solution is truly scalable, bringing the best of scalable video to the market, enabling mass deployment of visual communications for service providers, enterprises and system integrators.

Contents

Abstract

Video applications over the public Internet and corporate networks are gaining popularity. These, and other applications, including visual communications, are based on the transmission of video to various recipients located in different locations, using different equipment and connected to different networks.

Video coding standards have attempted to address the challenges presented in these use-cases by introducing scalability features that allow recipients, as well as other network elements, to adapt video streams to different capabilities. To date, however, all commercial solutions with scalability features have not been successful.

The introduction of H.264/SVC brought essential improvement, enabling for the first consumer products to be based on scalable video coding technologies.

However, deploying a scalable video coding solution is a lot more than video coding. A complete solution must deal with non-SVC infrastructure and legacy equipment.

RADVISION offers a real evolution to visual communication with its innovative solution based on SVC technology that allows connectivity for any user, using any equipment, on any network.

The RADVISION solution is truly scalable, bringing the best of scalable video to the market, and enabling mass deployment of visual communications.

Scalable Video Coding Demystified

Video applications, over the public Internet or corporate networks, have gained popularity in recent years. These and other applications, such as visual communications, surveillance and monitoring, telemedicine and eLearning, are based on the transmission of video to many recipients, located in different locations, using different equipment and connected to different networks.

Video coding standards have tried to address the inherent challenges of these use-cases by introducing scalability features, which allow recipients, as well as other network elements, to adapt the video stream according to their capabilities: CPU, display size, software support on the client side, available bandwidth, bandwidth variability on the network side and network conditions (packet loss).

The Scalability Types of SVC

Scalability can be implemented on several different aspects of video coding, such as resolution (frame size), frame rate (frames per second), and quality (bit rate). Different scalability types handle each of the above aspects.

A video stream is considered "scalable" when parts of the stream can be dropped, via a process known as "layer thinning", leaving a sub-stream which represents a different, yet equally valid, video stream.

A scalable video stream is comprised of a base layer, and a set of different enhancement layers, which – when combined with the base layer and other lower layers – yield a valid stream with different capabilities. These layers can be thinned, if needed. A stream that does not provide this property is referred to as a single-layer stream.

Spatial scalability allows sending a single video stream encapsulating several different resolutions. The stream may include various sub-streams, each representing a different resolution. For instance, a base layer may represent a qCIF (176x144) video, while enhancement layers may increase resolution to CIF (352x288) and 4CIF (704x576).

Temporal scalability allows sending a single video stream encapsulating several different frame rates. The stream may include various sub-streams, each representing a different frame rate. For instance, a base layer may implement a 7.5 frames per second (fps) video, while enhancement layers may increase frame rate to 15fps and 30fps.

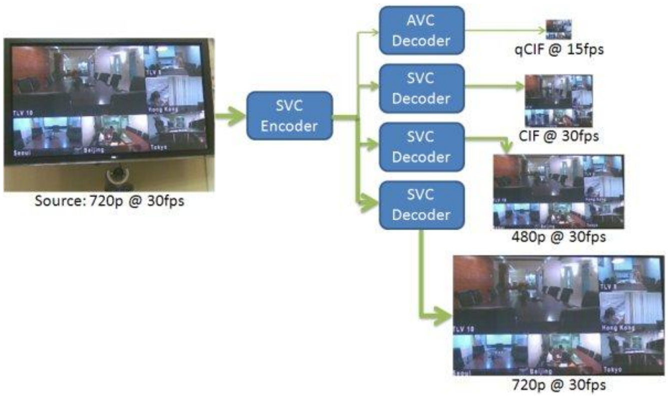

Fig. 1: on-the-fly adaptation of transmitted content

Quality scalability, also known as bit-rate or signal-to-noise (SNR) scalability, allows sending a single video stream encapsulating several different quality levels.

Quality is defined as the bit rate necessary for encoding a given stream, assuming a higher bit rate means better quality. The stream may include various sub-streams, each representing a different quality level. For instance, a base layer may represent a 128Kbps (Kilo bits per second) video, while enhancement layers may increase bit rate to 256Kbps and 512Kbps.

The different scalability types can also be combined, so many representations with different spatial-temporal-quality levels can be supported within a single scalable bit stream.

SVC Standardization

International video coding standards have played an important role in the success of digital video applications, past and present. As scalability is key for successful mass deployment of any application, scalable video existed in all video coding standards - including MPEG-2 Video, H.263, and MPEG-4 Visual - in the form of scalable profiles.

However, providing spatial and quality scalability using these standards significantly added complexity and significantly reduced coding efficiency – defined as the bit rate needed to maintain a given quality level or the quality in a given bit rate - as compared to the corresponding non-scalable profiles.

These drawbacks, together with the lack of signaling and transport standardization related to these standards, practically made the scalable profiles unusable. Therefore, It become the focus of work done in the recent SVC amendment of the H.264/AVC standard, known as H.264/SVC.

The H.264/SVC Standard

The H.264/MPEG-4 Advanced Video Coding (AVC) standard (known as H264/AVC) represents the state-of-the-art in video coding. Compared to prior video coding standards, it offers optimal coding efficiency, significantly reducing the necessary bit rate and offering higher quality at given bit rates. It also offers a very high level of implementation flexibility, which made it very popular in the consumer and corporate markets.

In October 2003 the Moving Picture Experts Group (MPEG) issued a call for proposals on Scalable Video Coding (SVC) Technology. On January 2005 MPEG and the Video Coding Experts Group (VCEG) agreed to standardize the SVC project as an amendment to the H.264/MPEG-4 AVC standard. On November 2007 the amendment was ratified, and is now known as the "SVC extension to H.264/AVC", H.264 Annex G or H.264/SVC.

H.264 SVC offers essential advantages that may help the standard succeed where its predecessors have failed before: in the area of mass commercial deployment.

These advantages include:

- Low impact on coding efficiency compared to single-layer coding for each sub-stream of the scalable bit stream
- Minimal increase in decoding complexity compared to single-layer decoding (due to the introduction of single loop decoding)
- Support for spatial, temporal and quality scalability
- Backward compatibility to H.264/AVC, as the H.264/SVC base layer can be decoded by any standard H.264/AVC decoder.
- Support for simple bit stream adaptations after the stream was encoded

H.264/AVC and H.264/SVC

The H.264/AVC standard covers the Video Coding Layer (VCL), as well as the Network Abstraction Layer (NAL). The VCL creates the coded representation of the source content, while the NAL formats the data and provides header information that enables simple and effective customization of the VCL data.

In addition, the H.264/AVC standard was followed by ITU-T RFC 3984, «RTP Payload Format for H.264 Video». This RFC covered the RTP payload format for H.264, which is used for video streaming over packet networks.

H.264/SVC is an amendment of the H.264/AVC standard. It changes the VCL and NAL accordingly, standardizing the coded representation of scalable content, as well as the necessary changes in the header information for scalable coding.

A draft for the RTP payload format for SVC Video is under work (draft-ietf-avt-rtp-svc-18, dated March 2009). An RFC, ITU-T RFC 5583, for SDP signaling ("Signaling media decoding dependency in Session Description Protocol (SDP)") was published on July 2009.

Scalable Video Coding Benefits

Scalable Video Coding allows applications to send a single video stream, encoded only once, but able to support a range of resolutions, bitrates and quality levels, via the process of discarding selected data ("layer thinning").

This means that clients with limited resources (display resolution, processing power, battery, etc.) can decode a subset of the original stream, which best suits their capabilities.

Note that layer thinning can be done not only by the receiving client, but also by any Media-Aware Network Element (MANE) in the system. Therefore, in multi-cast scenarios, terminals with different network conditions can be served by a single scalable bit stream.

As the transmitted stream is encapsulating different sub-streams representing different possible streams, the need for transcoding, as a means to adapting content to meet recipient capabilities is eliminated.

SVC and Error Resiliency

But SVC introduces another significant advantage over single-layer coding: error resiliency. Error resiliency is the ability of the video stream to survive extreme network condition, such as packet loss, with no decline in the perceived quality on the receiver side.

As temporal scalability introduces a layered structure of frame dependency, the ability of any network element, including a receiving client, to decode the bit stream even when some of the data is missing is greatly enhanced.

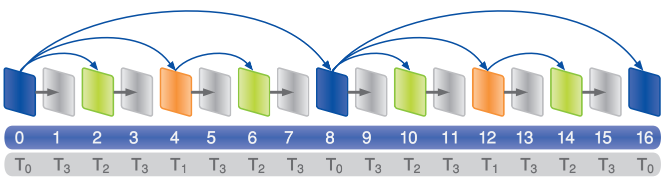

Fig 2: Temporal Scalability example

In the figure above, it is easy to see that if the odd-numbered frames (1, 3, 5, etc.) are lost, the rest of the stream is not adversely affected. If frames 2 or 6 are lost, only one other frame is affected. Therefore, this layered stream is much more error resilient than a similar single-layer stream.

Video from single-layers coders typically cannot be used over networks that have more than 1% packet loss. if sophisticated error correction schemes are implemented, perhaps 5%. But an H.264/SVC stream, due to the multi-layer structure explained above, can offer good visual quality, without artifacts, even with up to 40% packet loss.

Scalable Video Coding Applications

More and more applications are using the public Internet as their main means of transport. Among these are hosted services ("in the cloud"), consumer applications and global corporate networks, which rely on the public internet for inter-branch connectivity.

Error resiliency, therefore, becomes one of the biggest challenges for mass deployment of visual communications infrastructure, and a constant obstacle preventing a high quality experience.

Furthermore, in some cases the video transmission is not point-to-point, but point-to-multipoint (i.e. streaming) or multipoint (i.e. video conferencing). Therefore, the need for stream scalability becomes a key factor in application feasibility and success.

SVC offers great value to most video-based applications in exactly these two areas: error resiliency (through mainly temporal scalability) and high scalability (through spatial and quality scalability).

One may argue that scalability always comes with a cost and that in some cases scalability is not beneficial (and in fact, may reduce the quality of experience). However, used wisely and in the right configuration, scalability brings benefit to almost any use-case.

For example, in a point-to-point video call spatial scalability or quality scalability might not be effective, and the bit rate penalty may reduce the overall quality. However, temporal scalability can offer error resiliency even in a point-to-point scenario.

As any application that involves the transmission of a video stream to one or more recipients can achieve a significant gain from scalable video coding, application areas relevant for scalable video coding include: Point-to-point video calling, video conferencing, video streaming, and more.

Scalable Video Coding Solutions

As mentioned, Scalable Video Coding did not gain significant market attention with previous video coding standards. The Scalable Video Coding amendment to the H.264 Standard offers a real possibility to change that; however, on the way to deploying SVC systems, we still have a few important obstacles.

The primary obstacle for SVC deployment is legacy and non-SVC equipment. Although any H.264 SVC stream encapsulates a H.264 AVC stream (single-layer stream) in its base layer, this is a low quality stream, as it represents the lowest scalability level available.

This means that legacy and non-SVC systems, that do not support H.264/SVC, will be able to receive only the low resolution, low frame rate and low bit rate version of the stream. These include mobile handsets, most desktop conferencing and unified communications solutions, room systems and legacy infrastructure.

Another obstacle is the lack of standardization in using SVC for communication purposes. Important aspects, such as signaling (capabilities exchange), transport and error protection, are not yet fully standardized. This makes all current H.264 SVC systems proprietary.

Proprietary systems are "closed gardens". This means that they offer no interoperability, not with non-SVC systems and not even with other SVC-based systems, unless they all belong to the same vendor.

The "Gateway" Approach

To solve the problems with a SVC-based solution in non-homogenous deployments, which represent the majority of visual communications deployments, SVC-based systems use a gateway, in order to bridge between the SVC silo and the non-SVC world.

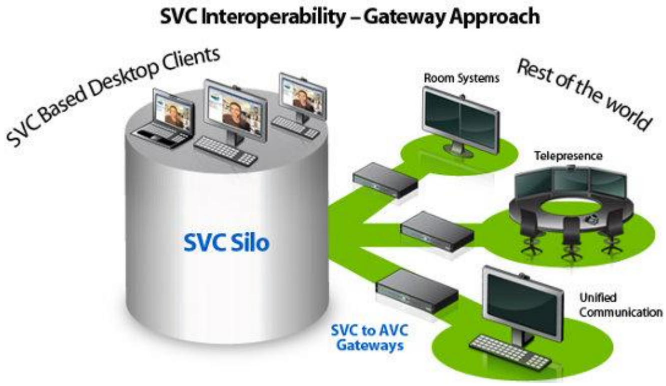

Fig. 3: the "Gateway" approach

Any client outside the SVC silo, be it a desktop client, room system, telepresence system or a mobile handset, requires a dedicated transcoding channel, to serve as a SVC-to-AVC bridge. This gateway is therefore complex in nature, not easily scaled and not cost-effective. On top of that, it increases latency and degrades quality of any call involving non-SVC clients - the majority of clients in today's market.

The RADVISION Solution

RADVISION's SVC solution is market-driven rather than technology driven - based on its experience as a leading vendor of visual communications infrastructures and expertise in innovation and standardization.

The video conferencing market has evolved significantly over the last several years, as video networks extended from meeting rooms onto desktops. RADVISION's MCU approach takes this evolution another step, allowing infrastructure to be scalable and fit any deployment.

RADVISION's MCU approach uses the existing MCU, software or hardware, to bring SVC-based and non-SVC infrastructure together. This approach brings the SVC benefits to the visual communications market, for both new and legacy systems.

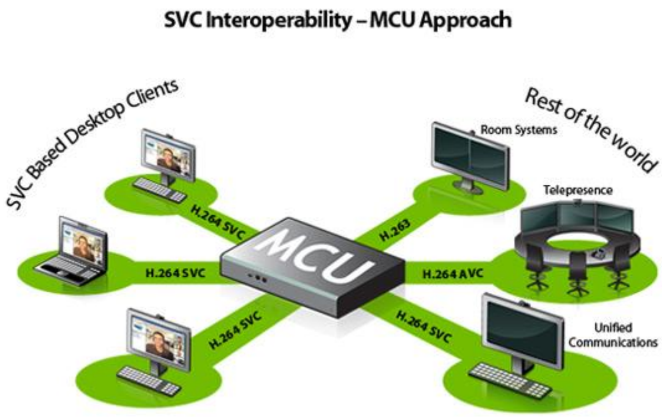

Fig 4: RADVISION's MCU Approach

Improving Error Resiliency

Current market needs require a solution that works with any device, without the need for additional components or transcoding, and without any increases in delay or decrease in quality for existing systems.

Through its new SVC-based technology, RADVISION offers a unique set of algorithms that allow full error resiliency over networks with up to 40% packet loss, while maintaining high quality and low delay.

Existing error resilience techniques can be generally categorized into two main groups: protective coding and correction codes. Both are discussed in length in "Maintaining A High Quality Experience In Unmanaged Networks", a RADVISION whitepaper.

In general, protective coding can yield a more resilient bitstream, but the compression is much less efficient (degraded quality); Using correction codes the stream can be protected without degrading the quality, but bandwidth and latency may be increased dramatically.

The RADVISION approach, however, uses Unequal Error Protection (UEP) methods, designed for data protection over erroneous channels, to provide full protection to the SVC layered structures. With UEP important data is protected, while less important data is less protected or not protected at all.

Exploiting the layered structure of SVC, RADVISION is able to protect the stream without adding a substantial amount of bandwidth (as in the protective coding approach), while maintaining a very high quality (as in the correction code approach).

RADVISION's UEP technology involves a range of error protection methods, including Forward Error Correction (FEC), packet retransmission and Video Fast Update (VFU). All methods are either standard-based or being standardized by RADVISION (e.g draft-galanos-fecframe-rtp-reedsolomon-mf-00 from July 2009).

By using the UEP scheme on top of H.264 SVC coding, RADVISION's solution enables customers to enjoy the highest quality of experience, even under error-prone network conditions.

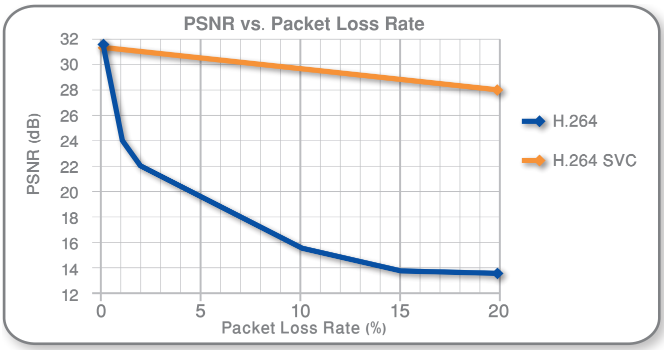

Fig 5: Quality in different Packet Loss rates

From figure 5 it is clear that while H.264 offer excellent quality (shown here in terms of Peak Signal-to-Noise Ratio, or PSNR), its quality declines significantly when packet loss rate increases. RADVISION's SVC-based solution offers similar excellent quality even at much higher packet loss rates.

The RADVISION SVC Offering

In 2009 RADVISION introduces the first line of visual communication infrastructure supporting SVC technology and offer a perfect match for many of today's use-cases:

- Point-to-point calling, such as from one desktop software client to another. The solution ensures high quality of experience, even if network conditions and/or client capabilities are limited.
- Remote calling scenarios, uch as when tele-workers are call from home into the company network. The RADVISION offering enables a high quality call, even if the home connection suffers from packet loss or limited bandwidth.
- Virtual MCU deployments, where MCUs are deployed in different geographic locations (branches) and are connected over dedicated lines or public Internet. The solution will enable a perfect connection between network elements, even if the paths are error prone. This allows standard non-SVC systems to enjoy the benefits of SVC through MCU cascading.

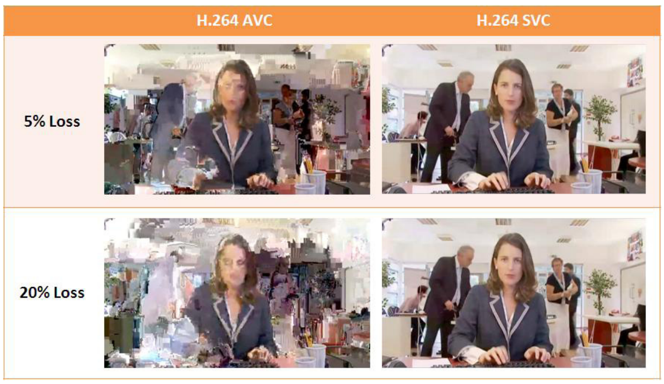

Fig 5: Quality in different Packet Loss rates

In addition to its infrastructure products powered by SVC, RADVISION will soon introduce the SVC SDK for vendors who wish to deploy a ready-made Scalable Video Coding engine within their products.

When connecting the RADVISION solution to legacy MCU equipment, as well as any new non-SVC equipment, the RADVISION approach allows for perfect interoperability, with no adverse effect on the user experience for the existing install base.

The RADVISION solution is truly scalable, bringing the best of scalable video and enabling real mass deployment of visual communications - from SMBs to very large enterprises deployments.

Conclusion

Scalable video coding offers a lot of benefits for video-based applications. The H.264/SVC standard is enabling the adoption of SVC in the visual communications industry. Having said that, SVC deployments are not without problems - mainly in terms of standardization and backward compatibility.

The RADVISION solution, focused on scalable video coding techniques combined with innovative UEP technologies, enables a consistently high quality experience, encouraging increased uptake of visual communications and IP telephony solutions in the enterprise arena.

The RADVISION solution delivers an optimized video experience, being transmitted to a wide range of end point devices, on a variety for network configurations. This integrated approach enables service providers, enterprises and system integrators to provide a high quality user experience.

RADVISION's SVC solutions bring substantial benefit to users, as error resiliency dramatically improves even under extreme network conditions, while maintaining the high level of interoperability RADVISION customers have learned to expect.

USA/Americas
T +1 201 689 6300
F +1 201 689 6301
infoUSA@radvision.com

EMEA
T +44 20 3178 8685
F +44 20 3178 5717
infoUK@radvision.com

APAC
T +852 3472 4388
F +852 2801 4071
infoAPAC@radvision.com

MKT-DI-210909-01