

Claims 1, 6, 11

Scale Video Coding LLC (“SVC”) provides evidence of infringement of claims 1, 6, and 11 of U.S. Patent No. 11,019,372 (hereinafter “the ’372 patent”) by Cisco Systems, Inc. (“Cisco”). In support thereof, SVC provides the following claim charts.

“Accused Instrumentalities” as used herein refers to at least the Cisco Meeting Server, Web App, Meeting Management, and associated products, along with associated hardware and/or software, including but not limited to other Cisco servers and related computer systems operated by Cisco that work in conjunction with the Cisco Meeting Server, Web App, Meeting Management, and associated products.

These claim charts demonstrate Cisco’s infringement by comparing each element of the asserted claims to corresponding components, aspects, and/or features of the Accused Instrumentalities. These claim charts are not intended to constitute an expert report on infringement. These claim charts include information provided by way of example, and not by way of limitation.

The analysis set forth below is based only upon information from publicly available resources regarding the Accused Instrumentalities, as Cisco has not yet provided any non-public information. An analysis of Cisco’s (or other third parties’) technical documentation and/or software source code may assist in fully identify all infringing features and functionality. Accordingly, SVC reserves the right to supplement this infringement analysis once such information is made available to SVC. Furthermore, SVC reserves the right to revise this infringement analysis, as appropriate, upon issuance of a court order construing any terms recited in the asserted claims.

SVC provides this evidence of infringement and related analysis without the benefit of claim construction or expert reports or discovery. SVC reserves the right to supplement, amend or otherwise modify this analysis and/or evidence based on any such claim construction or expert reports or discovery.

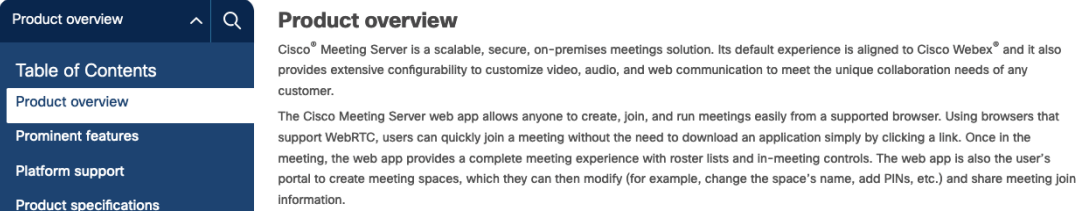
Unless otherwise noted, SVC contends that Cisco directly infringes the ’372 patent in violation of 35 U.S.C. § 271(a) by selling, offering to sell, making, and/or using, the Accused Instrumentalities. The following exemplary analysis demonstrates that infringement. Cisco makes, uses, sells, imports, or offers for sale in the United States, or has made, used, sold, or offered for sale in the past, without authority, without authority products, equipment, or services that infringe claims 1, 6, and 11 of the ’372 patent, including without limitation, the Accused Instrumentalities.

Unless otherwise noted, SVC believes and contends that each element of each claim asserted herein is literally met through Cisco’s provision of the Accused Instrumentalities. However, to the extent that Cisco attempts to allege that any asserted claim element is not literally met, SVC believes and contends that such elements are met under the doctrine of equivalents. More specifically, in its investigation and analysis of the Accused Instrumentalities, SVC did not identify any substantial differences between the elements of the patent claims and the corresponding features of the Accused Instrumentalities, as set forth herein. In each instance, the identified feature

of the Accused Instrumentalities performs at least substantially the same function in substantially the same way to achieve substantially the same result as the corresponding claim element.

To the extent the chart of an asserted claim relies on evidence about certain specifically-identified Accused Instrumentalities, SVC asserts that, on information and belief, any similarly-functioning instrumentalities also infringe the charted claim. SVC reserves the right to amend this infringement analysis based on other products made, used, sold, imported, or offered for sale by Cisco. SVC also reserves the right to amend this infringement analysis by citing other claims of the '372 patent, not listed in the claim chart, that are infringed by the Accused Instrumentalities. SVC further reserves the right to amend this infringement analysis by adding, subtracting, or otherwise modifying content in the "Accused Instrumentalities" column of each chart.

Claim Chart – Cisco

U.S. Pat. No. 11,019,372 Claims 1, 6, 11	Accused Instrumentalities
1. A video router, comprising:	The Accused Instrumentalities include a video router. Scalable video routers conforming to the H.264 standard perform a method of transmitting video signals by forwarding IP data packets containing scalable video information. https://www.cisco.com/c/en/us/products/collateral/conferencing/meeting-server/datasheet-c78-742168.html  https://www.cisco.com/c/en/us/products/collateral/conferencing/meeting-server/datasheet-c78-742168.html

	  sciencedirect.com/topics/engineering/scalable-coding to serve the various users' requirements. Usually all participants would have to agree on some minimal quality for the video conference, but this is not really satisfactory in many cases. With the advent of SVC, the MCU is not needed because transcoding is not needed. Using the SVC standard, so-called <u>video routers</u> can be substituted for the MCUs. These routers only have to <u>forward</u> and delete packets as appropriate to that user's connection and needs, so little extra delay is added. In effect, video packets can be routed by the video router based on their headers, just like other <u>network packets</u>.
a memory; and a processor, wherein the processor executes instructions stored in the memory to cause the video router to:	The Accused Instrumentalities include a video router, comprising a memory and a processor, wherein the processor executes instructions stored in the memory. All routers comprise a memory and a processor, wherein the processor executes instructions stored in the memory.

	https://www.learncisco.net/courses/icnd-2/network-management/router-internal-components-overview.html Router Internal Components Overview ← Ads by Google Report this ad Why this ad? ▾ I want you to take a moment and look at the components we have listed here. • CPU - Executes operating system instructions • RAM (Random access memory) - Contains the running copy of the configuration file. Stores the routing table. RAM contents lost when power is off. • ROM (Read-only memory) - Holds diagnostic software used when router is powered up. Stores the router's bootstrap program. • NVRAM (Nonvolatile RAM) - Stores startup configuration. This may include IP addresses (routing protocol, hostname of a router). • Flash memory - Contains the operating system (Cisco IOS) • Interfaces - Multiple physical interfaces that are used to connect the network. Fast Ethernet, Gigabit Ethernet; Serial; Management interfaces CPU, RAM, ROM, interfaces - those all sound very familiar to something else. A computer, a PC! All of these components we find inside the tower of our PC or inside of our laptop. So we have our router, we look inside, we have the same types of components. Let's go through these individually. Aux Console USB Flash NVRAM CPU System Control RAM USB Port Serial 0/0/0 Serial 0/0/1 FastEthernet 0/0 FastEthernet 0/1
receive a layered video data stream	The Accused Instrumentalities include a video router, comprising a memory and a processor, wherein the processor executes instructions stored in the memory to cause the video router to receive a layered video data stream. The accused instrumentalities conform to the H.264 standard, including Annex G (SVC). Scalable video routers conforming to the H.264 standard perform a method of receiving a multi-layered stream.

https://www.cisco.com/c/en/us/products/collateral/conferencing/meeting-server/datasheet-c78-742168.html

Product specifications

Table of Contents

Product overview

Prominent features

Platform support

Product specifications

Ordering information

Cisco Capital

Document history

Product specifications

Table 2 lists the specifications for Cisco Meeting Server.

Table 2. Specifications

Video standards

- H.261
- H.263 (+, ++)
- H.264 AVC (baseline and high profile)
- H.264 SVC
- WebM, VP8
- Microsoft RTV
- HTML5/WebRTC
- SIP, H.323, TIP
- BFCP
- RDP
- Far End Camera Control (FECC) passthrough
- RTMP and RTMPS

<https://www.cisco.com/c/en/us/products/collateral/conferencing/meeting-server/datasheet-c78-742168.html>

Feature Configuration Guide for Cisco Unified Communications Manager, Release 14 and SUs

Book Contents

Book Title Page

New and Changed Information

> Getting Started

> Remote Worker Features

> Remote Network Access

> Licensing

> Monitoring and Recording

> Call Center Features

> Voice Messaging Features

> Conferencing Features

> Placing Calls

> Receiving Calls

> Presence and Privacy Features

> Custom Features

> Device Management

> Advanced Call Processing

> QoS Management

> SIP Interoperability

Configure SIP Normalization and Transparency

Configure SDP Transparency Profiles

Configure Presentation Sharing using BFCP

Video Telephony

> Emergency Call Routing Regulations

Find Matches in This Book

Save

Trans

Chapter: Video Telephony

Updated: June 7, 2023

Was this Document Useful?

Yes No

Contact Cisco

[Open a Support Case](#)

(Requires a Cisco Service Contract)

> Chapter Contents

- Video Telephony Overview
- Video Telephony Support
- Video Network
- Video Telephony Configuration Task Flow
- H.323 Video
- Video Support
- Video Features
- QoS for Video Networks

Video Telephony Overview

Unified Communications Manager supports video telephony and thus unifies the world of voice and video calls. Video endpoints use Unified CM call-handling features and access a unified voice and video solution for dialing and connecting video calls.

The Unified Communications Manager video telephony solution offers these features:

- Supports video and video-related features, such as Far-End Camera Control (FECC)
- Supports multiple logical channels that are needed to allow the transmission of video streams
- Transmits midcall, media-related messages that are needed for video (that is, transmits commands or indications that are needed for video calls)
- Supports H.323, Skinny Client Control Protocol (SCCP), and Session Initiation Protocol (SIP)
- Enhances locations and regions to provide Bandwidth Management
- Provides serviceability information, such as Call Detail Records (CDRs), about video calls

Video Telephony Support

The following sections discuss the details of video telephony in the Unified Communications Manager environment.

- Video Calls
- Real-Time Transport Control Protocol Pass-Through in MTP Topologies
- Video Codecs

Video Calls

The typical video call includes two or three Real-Time Protocol (RTP) streams in each direction (that is, four or six streams). The call can include the following stream types:

- Video (H.261, H.263, H.263+, **H.264-SVC**, X-H.264UC, H.264-AVC, H.265, AV1 and VT Camera wideband video codecs)
- Far-End Camera Control (FECC) - Optional
- Binary Floor Control Protocols (BFCP)

https://www.cisco.com/c/en/us/td/docs/voice_ip_comm/cucm/admin/14SU2/cucm_b_feature-configuration-guide-for-cisco14su2/cucm_m_video-telephony.html

	← → ↻ sciencedirect.com/topics/engineering/scalable-video-coding Communicating Pictures: Delivery Across Networks David R. Bull, in Communicating Pictures, 2014 Scalable video encoding It is common in video transmission for the signal to be encoded and transmitted without explicit knowledge of the downstream network conditions. In cases where network congestion exists in a packet switched network such as the internet, routers will discard packets that they are unable to forward due to congestion. In some cases it is possible for packets within a stream to be prioritized or embedded such that the least important information is discarded before the more important data. This mechanism lends itself to scalable or layered encoding, where a video signal is composed of a hierarchical layering in terms of spatial resolution, temporal resolution or SNR. This enables devices to transmit and <u>receive multilayered video streams</u> where a base level of quality can be improved through the use of optional additional layers that enhance resolution, frame rate, and/or quality.
including a base layer and a set of enhancement layers ,	The Accused Instrumentalities include a video router, comprising a memory and a processor, wherein the processor executes instructions stored in the memory to cause the video router to receive a layered video data stream including a base layer and a set of enhancement layers.



AEROSPACE AND DEFENSE

Land, Naval and Airborne ▼

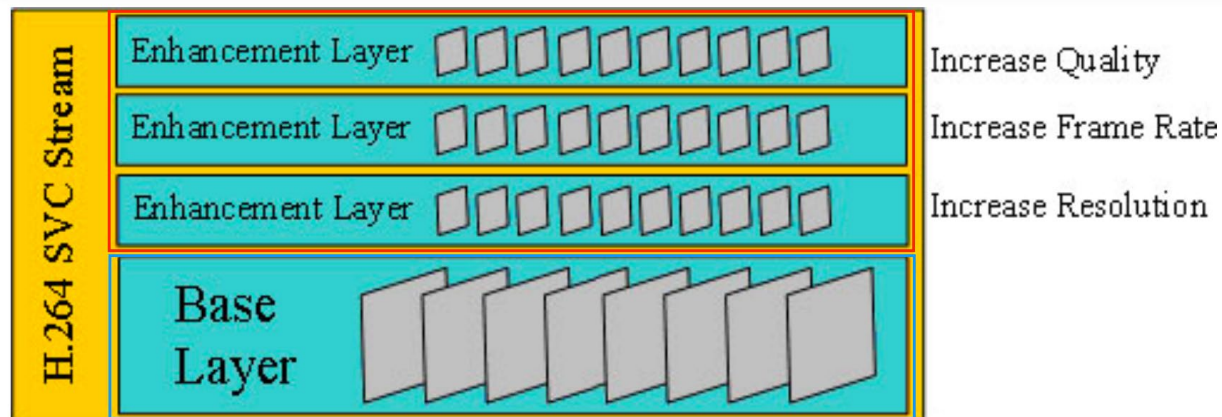


Figure 1: H.264 SVC Layering Technique

The “base layer” is the minimum decodable bitstream subset (or sub-bitstream), and every other layer is an enhancement layer.

G.3.4 base layer: A *bitstream subset* that contains all *NAL units* with the *nal_unit_type syntax element* equal to 1 and 5 of the *bitstream* and does not contain any *NAL unit* with the *nal_unit_type syntax element* equal to 14, 15, or 20 and conforms to one or more of the profiles specified in Annex A.

See ITU-T H.264 at 437.

← → ↺ [sciencedirect.com/topics/engineering/scalable-video-coding](https://www.sciencedirect.com/topics/engineering/scalable-video-coding)

Communicating Pictures: Delivery Across Networks

David R. Bull, in [Communicating Pictures](#), 2014

Scalable video encoding

It is common in video transmission for the signal to be encoded and transmitted without explicit knowledge of the downstream network conditions. In cases where network congestion exists in a [packet switched network](#) such as the internet, routers will discard packets that they are unable to forward due to congestion. In some cases it is possible for packets within a stream to be prioritized or embedded such that the least important information is discarded before the more important data. This mechanism lends itself to scalable or layered encoding, where a video signal is composed of a hierarchical layering in terms of [spatial resolution](#), [temporal resolution](#) or SNR. This enables devices to transmit and receive multilayered video streams where a base level of quality can be improved through the use of optional additional layers that enhance resolution, frame rate, and/or quality.

MPEG-4 Visual and H.264/AVC: Standards for Modern Digital Video

Berna Erol, ... Lowell Winger, in [The Essential Guide to Video Processing](#), 2009

10.3.6 Scalability

In addition to the video coding tools discussed so far, MPEG-4 Part 2 provides scalability tools that allow organization of the [bitstream](#) into base and [enhancement layers](#). The enhancement layers are transmitted and decoded depending on the bit rate, display resolution, network throughput, and decoder complexity constraints. Temporal, [spatial](#), [quality](#), complexity, and object-based Scalable video data streams coded according to the H.264 standard comprise a single bitstream that is further comprised of layers, wherein combinations of layers form “bitstream subsets” that may be decoded as video data streams.

G.3.54 scalable bitstream: A *bitstream* with the property that one or more *bitstream subsets* that are not identical to the scalable bitstream form another *bitstream* that conforms to this specification.

See Telecommunication Standardization Sector, International Telecommunication Union, H.264 Infrastructure of audiovisual services – Coding of moving video, 437 (Jun. 2019) (“ITU-T H.264”).

G.3.6 bitstream subset: A *bitstream* that is derived as a *subset* from a *bitstream* by discarding zero or more *NAL units*. A *bitstream subset* is also referred to as *sub-bitstream*.

See ITU-T H.264 at 437.

	Scalable video coding is specified in Annex G allowing the construction of bitstreams that contain sub-bitstreams that conform to this Specification. For temporal bitstream scalability, i.e., the presence of a sub-bitstream with a smaller temporal sampling rate than the bitstream, complete access units are removed from the bitstream when deriving the sub-bitstream. In this case, high-level syntax and inter prediction reference pictures in the bitstream are constructed accordingly. For spatial and quality bitstream scalability, i.e., the presence of a sub-bitstream with lower spatial resolution or quality than the bitstream, NAL units are removed from the bitstream when deriving the sub-bitstream. In this case, inter-layer prediction, i.e., the prediction of the higher spatial resolution or quality signal by data of the lower spatial resolution or quality signal, is typically used for efficient coding. Otherwise, the coding algorithm as described in the previous paragraph is used. See ITU-T H.264 at 4.
identify bandwidth-limited conditions of an internet protocol network between the video router and a plurality of video receivers,	The Accused Instrumentalities include a video router, comprising a memory and a processor, wherein the processor executes instructions stored in the memory to cause the video router to receive a layered video data stream including a base layer and a set of enhancement layers which identify bandwidth-limited conditions of an internet protocol network between the video router and a plurality of video receivers. Scalable video routers detect conditions in which bandwidth is limited. Such a bandwidth limitation could be caused by a congested state of a network over which packets must be sent, or a weak physical connection between two or more links in the network, a limitation in the capabilities of a given receiver, or some other network condition that would limit the bandwidth between one or more nodes.

Communicating Pictures: Delivery Across Networks

David R. Bull, in [Communicating Pictures](#), 2014

Scalable video encoding

It is common in video transmission for the signal to be encoded and transmitted without explicit knowledge of the downstream network conditions. In cases where network congestion exists in a [packet switched network](#) such as the internet, routers will discard packets that they are unable to forward due to congestion. In some cases it is possible for packets within a stream to be prioritized or embedded such that the least important information is discarded before the more important data. This mechanism lends itself to scalable or layered encoding, where a video signal is composed of a hierarchical layering in terms of [spatial resolution](#), [temporal resolution](#) or SNR. This enables devices to transmit and receive multilayered video streams where a base level of quality can be improved through the use of optional additional layers that enhance resolution, frame rate, and/or quality.

	  sciencedirect.com/topics/engineering/scalable-coding to serve the various users' requirements. Usually all participants would have to agree on some minimal quality for the video conference, but this is not really satisfactory in many cases. With the advent of SVC, the MCU is not needed because transcoding is not needed. Using the SVC standard, so-called <i>video routers</i> can be substituted for the MCUs. These routers only have to forward and delete packets <u>as appropriate to that user's connection and needs</u>, so little extra delay is added. In effect, video packets can be routed by the video router based on their headers, just like other <i>network packets</i>.
forward the base layer from the video router to at least two of the plurality of video receivers via the internet protocol network,	The Accused Instrumentalities include a video router, comprising a memory and a processor, wherein the processor executes instructions stored in the memory to cause the video router to receive a layered video data stream including a base layer and a set of enhancement layers which identify bandwidth-limited conditions of an internet protocol network between the video router and a plurality of video receivers which forward the base layer from the video router to at least two of the plurality of video receivers via the internet protocol network. Because the base layer is the minimum bitstream subset of a given scalable bitstream that is capable of being decoded, every receiver that receives a decodable bitstream subset is sent at least the base layer.

	G.8.8.2 Specification of the base layer bitstream Each scalable bitstream that conforms to this specification shall contain a base layer bitstream that conforms to one or more of the profiles specified in Annex A. This base layer bitstream is derived by invoking the sub-bitstream extraction process as specified in clause G.8.8.1 with dIdTarget being equal to 0 and qIdTarget being equal to 0 and the base layer bitstream being the output. NOTE – Although all scalable bitstreams that conform to one or more of the profiles specified in this annex contain a base layer bitstream that conforms to one or more of the profiles specified in Annex A, the complete scalable bitstream (prior to operation of the base layer extraction process specified in this clause) may not conform to any profile specified in Annex A. Rec. ITU-T H.264 (06/2019) 595
and selectively forward one or more of the set of enhancement layers, but fewer than all of the set of enhancement layers, to at least two of the plurality of video receivers through the internet protocol network based upon the identified bandwidth-	The Accused Instrumentalities include a video router, comprising a memory and a processor, wherein the processor executes instructions stored in the memory to cause the video router to receive a layered video data stream including a base layer and a set of enhancement layers which identify bandwidth-limited conditions of an internet protocol network between the video router and a plurality of video receivers which forward the base layer from the video router to at least two of the plurality of video receivers via the internet protocol network which selectively forward one or more of the set of enhancement layers, but fewer than all of the set of enhancement layers, to at least two of the plurality of video receivers through the internet protocol network based upon the identified bandwidth-limited conditions. If a given end destination can be reached at data speeds insufficient to transmit the complete scalable bitstream because of a bandwidth-limiting condition, the scalable video router will select and send all of the enhancement layers which can be sent at sufficient data speeds for each of the plurality of video receivers. For instance, as shown below, fewer than all of the set of enhancement layers are forwarded when priority_id is greater than pIdTarget, or temporal_id is greater than tIdTarget, or dependency_id is greater than dIdTarget (or dependency_id is equal to dIdTarget but quality_id is greater than qIdTarget) for any VCL NAL units or filler data NAL units in the bitstream.

limited conditions,	<u>The sub-bitstream is derived by applying the following operations in sequential order:</u> 1. Mark all VCL NAL units and filler data NAL units for which any of the following conditions are true as "to be removed from the bitstream": – priority_id is greater than pIdTarget, – temporal_id is greater than tIdTarget, – dependency_id is greater than dIdTarget, – dependency_id is equal to dIdTarget and quality_id is greater than qIdTarget. 2. Remove all access units for which all VCL NAL units are marked as "to be removed from the bitstream". 3. Remove all VCL NAL units and filler data NAL units that are marked as "to be removed from the bitstream". 4. When dIdTarget is equal to 0 and qIdTarget is equal to 0, remove the following NAL units: – all NAL units with nal_unit_type equal to 14 or 15, – all NAL units with nal_unit_type equal to 6 in which the first SEI message has payloadType in the range of 24 to 35, inclusive. 5. Remove all NAL units with nal_unit_type equal to 6 that only contain SEI messages that are part of a scalable nesting SEI message with any of the following properties: – sei_temporal_id is greater than tIdTarget, – the minimum value of (sei_dependency_id[i] << 4) + sei_quality_id[i] for all i in the range of 0 to num_layer_representations_minus1, inclusive, is greater than (dIdTarget << 4) + qIdTarget. 6. Remove all NAL units with nal_unit_type equal to 6 that contain SEI messages with payloadType equal to 24, 28, or 29. See ITU-T H.264 at 595. All sub-bitstreams that can be derived using the sub-bitstream extraction process as specified in clause G.8.8.1 with any combination of values for priority_id, temporal_id, dependency_id, or quality_id as the input shall result in a set of coded video sequences, with each coded video sequence conforming to one or more of the profiles specified in Annexes A and G. See ITU-T H.264 at 489.
----------------------------	---

	The representation of a particular scalable layer is the set of NAL units that represents the set union of the particular scalable layer and all scalable layers on which the particular scalable layer directly or indirectly depends. The representation of a scalable layer is also referred to as scalable layer representation. In the following specification of this clause, the terms representation of a scalable layer and scalable layer representation are also used for referring to the access unit set that can be constructed from the NAL units of the scalable layer representation. A scalable layer representation can be decoded independently of all NAL units that do not belong to the scalable layer representation. The decoding result of a scalable layer representation is the set of decoded pictures that are obtained by decoding the access unit set of the scalable layer representation. See ITU-T H.264 at 624.   sciencedirect.com/topics/engineering/scalable-coding to serve the various users' requirements. Usually all participants would have to agree on some minimal quality for the video conference, but this is not really satisfactory in many cases. With the advent of SVC, the MCU is not needed because transcoding is not needed. Using the SVC standard, so-called <i>video routers</i> can be substituted for the MCUs. These routers only have to forward and delete packets <u>as appropriate to that user's connection and needs</u>, so little extra delay is added. In effect, video packets can be routed by the video router based on their headers, just like other <i>network packets</i>.
and wherein the video router transmits the layered video data stream according to	The Accused Instrumentalities include a video router, comprising a memory and a processor, wherein the processor executes instructions stored in the memory to cause the video router to receive a layered video data stream including a base layer and a set of enhancement layers which identify bandwidth-limited conditions of an internet protocol network between the video router and a plurality of video receivers which forward the base layer from the video router to at least two of the plurality of video receivers via the internet protocol network which selectively forward one or more of the set of enhancement layers, but fewer than all of the set of enhancement layers, to at least two of the plurality of video receivers through the internet protocol network based upon the identified bandwidth-limited conditions, wherein each layer of the layered video data stream comprises data packets. Scalable video routers

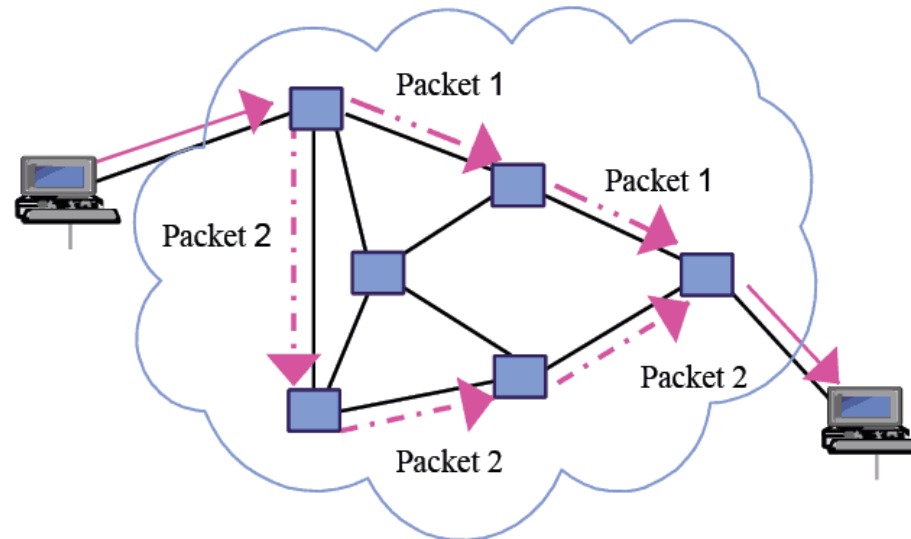
an internet protocol;
wherein each layer of the layered video data stream comprises data packets,

facilitate the distribution of scalable bitstreams of video information across a packet switched network, such as the internet. Data packets are the result of a process of dividing information into units so that said units may be transmitted across a packet switched network. At the destination, the data packets are reassembled to produce the information from which said data packets were derived.

→ ↻ networkencyclopedia.com/packet-switching/

What is Packet Switching?

Packet Switching is the process by which a networking or telecommunications device accepts a packet and switches it to a telecommunications device that will take it closer to its destination. Packet switching allows data to be sent over the telecommunications network in short bursts or “packets” that contain sequence numbers so that they can be reassembled at the destination.



Data packets that are transmitted across the internet are comprised of a header and payload/body, according to the internet protocol. The payload is comprised of the data intended for transmission, and the header is information facilitating said transmission. Moreover, the payload of a packet is often another packet, comprising another header

and another payload, as each packet layer is designed to be handled by a different part of the data distribution process. (NOTE: packet layers are not to be confused with scalable video layers).

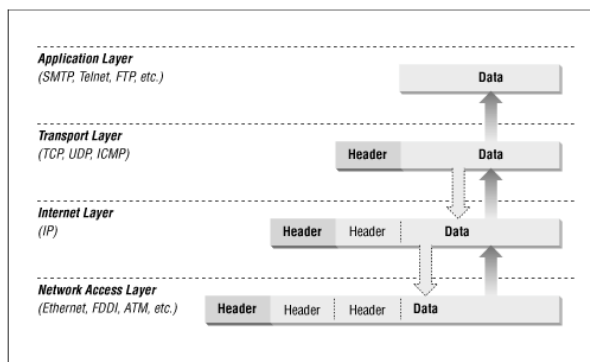
← → ↻ ⓘ Not Secure | web.deu.edu.tr/doc/oreilly/networking/firewall/ch06_03.htm ☆ W ⋮

Packets are constructed in such a way that layers for each protocol used for a particular connection are wrapped around the packets, like the layers of skin on an onion.

At each layer, a packet has two parts: the header and the body. The header contains protocol information relevant to that layer, while the body contains the data for that layer which often consists of a whole packet from the next layer in the stack. Each layer treats the information it gets from the layer above it as data, and applies its own header to this data. At each layer, the packet contains all of the information passed from the higher layer; nothing is lost. This process of preserving the data while attaching a new header is known as *encapsulation*.

At the application layer, the packet consists simply of the data to be transferred (for example, part of a file being transferred during an FTP session). As it moves to the transport layer, the Transmission Control Protocol (TCP) or the User Datagram Protocol (UDP) preserves the data from the previous layer and attaches a header to it. At the next layer, IP considers the entire packet (consisting now of the TCP or UDP header and the data) to be data, and now attaches its own IP header. Finally, at the network access layer, Ethernet or another network protocol considers the entire IP packet passed to it to be data, and attaches its own header. [Figure 6.2](#) shows how this works.

Figure 6.2: Data encapsulation



At the other side of the connection, this process is reversed. As the data is passed up from one layer to the next higher layer, each header (each skin of the onion) is stripped off by its respective layer. For example, the Internet layer removes the IP header before passing the encapsulated data up to the transport layer (TCP or UDP).

Additionally, the final payload of a scalable video data packet is a NAL Unit, which is also a unit of information comprising a header and a body. Scalable bitstreams that conform to the H.264 standard are transmitted as a sequence of NAL units.

	ip.hhi.de/imagecom_G1/savce/downloads/SVC-Overview.pdf <i>A. Network Abstraction Layer (NAL)</i> The coded video data are organized into NAL units, which are packets that each contains an integer number of bytes. A NAL unit starts with a one-byte header, which signals the type of the contained data. The remaining bytes represent payload data. NAL units are classified into VCL NAL units, which con-
each of which is encoded with a sequence number	The Accused Instrumentalities include a video router, comprising a memory and a processor, wherein the processor executes instructions stored in the memory to cause the video router to receive a layered video data stream including a base layer and a set of enhancement layers which identify bandwidth-limited conditions of an internet protocol network between the video router and a plurality of video receivers which forward the base layer from the video router to at least two of the plurality of video receivers via the internet protocol network which selectively forward one or more of the set of enhancement layers, but fewer than all of the set of enhancement layers, to at least two of the plurality of video receivers through the internet protocol network based upon the identified bandwidth-limited conditions, wherein each layer of the layered video data stream comprises data packets, each of which is encoded with a sequence number. Each data packet is identified by a sequence number which uniquely identified said packet among all other packets in the message. Every packet in the sequence contains an identification value that is one more than that of the previous packet in the sequence.

Packet Header Definition

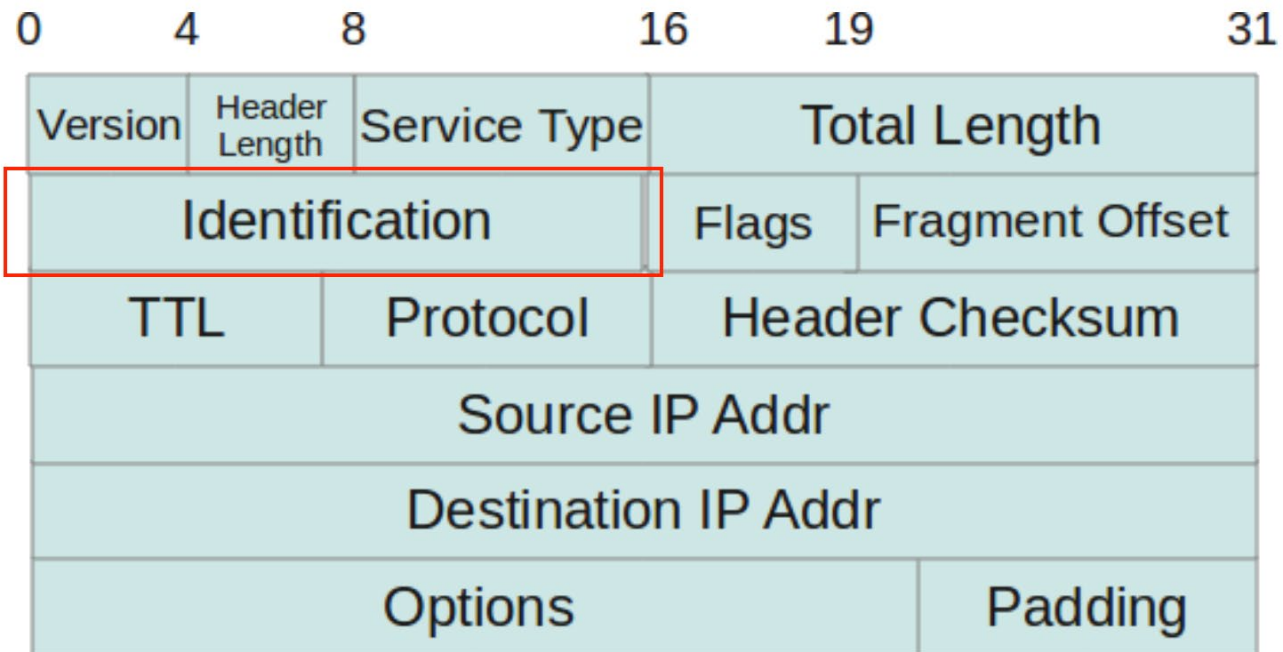
A *packet header* is the portion of an [IP](#) (Internet protocol) *packet* that precedes its body and contains addressing and other data that is required for it to reach its intended destination.

Packets are the fundamental unit of information transport in all modern computer [networks](#), and increasingly in other communications networks as well. They can be a fixed size or variable sizes, depending on the system. Regardless of their size, each packet consists of three main parts: a header, the body, also called the *payload*, and a *trailer*.

The header's format is specified in the Internet protocol. It normally contains 20 [bytes](#) of data, although an option exists within it that allows the addition of more bytes.

Among the contents of the header are the version of IP (which is always set to 4, because IPv4 is being used), the sender's [IP address](#), the intended receiver's IP address, the number of packets the message has been broken into, the identification number of the particular packet, the [protocol](#) (e.g. 1 for [ICMP](#), 2 for [IGMP](#), 6 for [TCP](#) and 17 for [UDP](#)) used, the packet length (on networks that have variable length packets), the *time to live* (i.e., the number of links or [hops](#) that the packet can be routed before being allowed to expire) and synchronization data (several [bits](#) that help the packet match up to the network).

IP Header



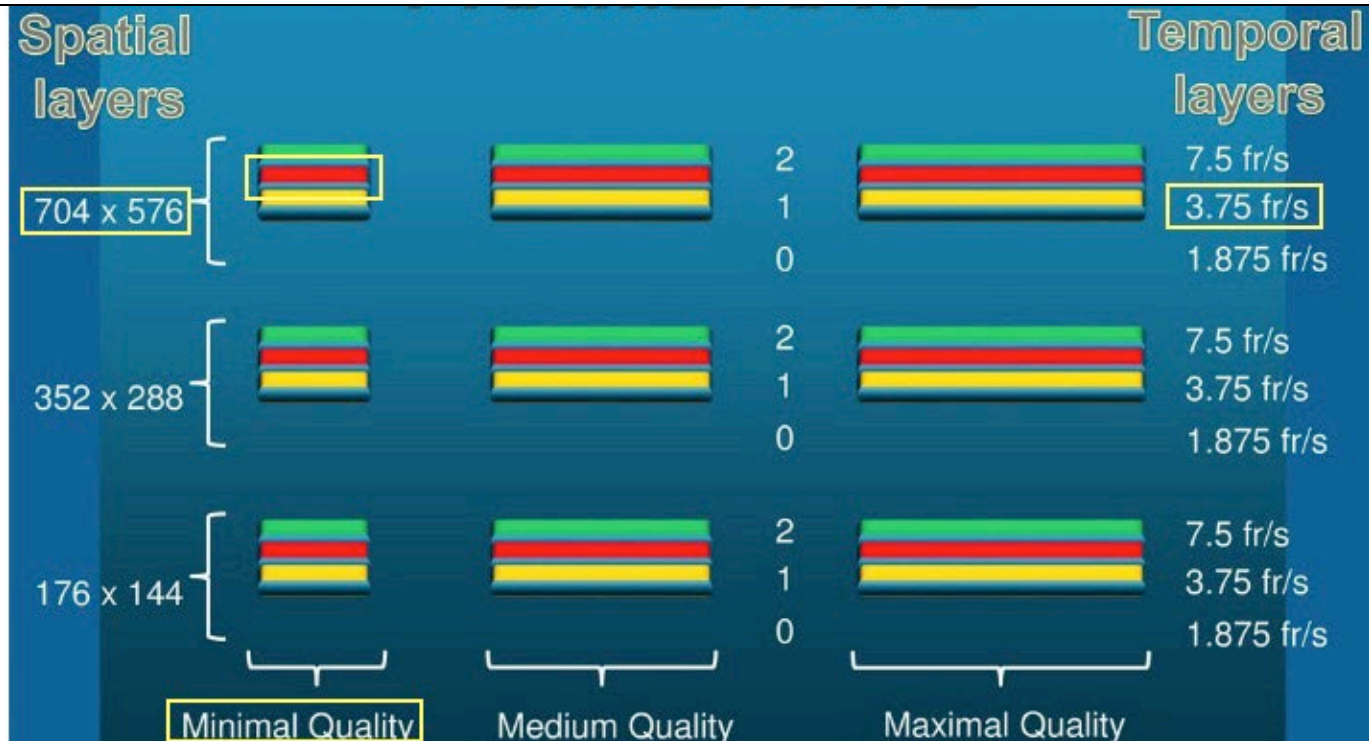
- **Identification(16 bits):** This field is used for uniquely identifying the IP datagrams. This value is incremented every-time an IP datagram is sent from source to the destination. This field comes in handy while reassembly of fragmented IP data grams.

and a layer identifier, and

The Accused Instrumentalities include a video router, comprising a memory and a processor, wherein the processor executes instructions stored in the memory to cause the video router to receive a layered video data stream including a

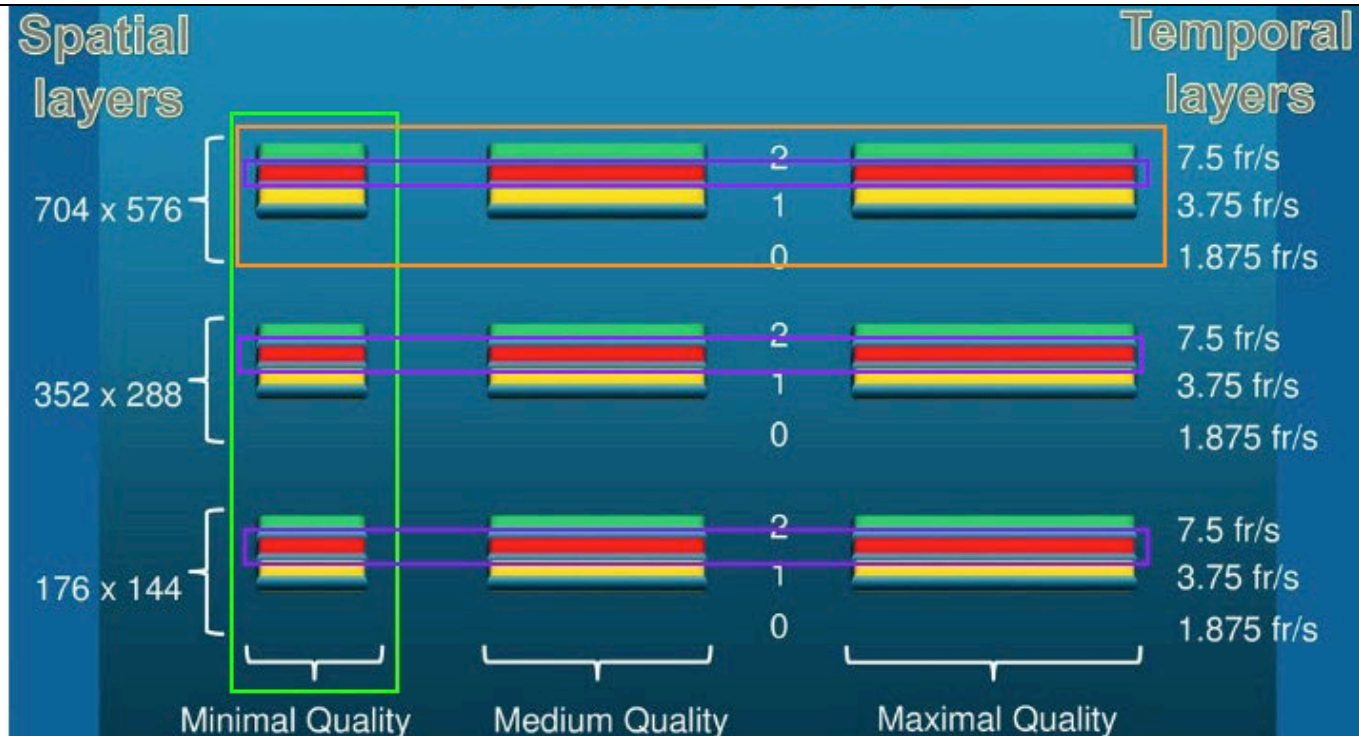
wherein the layer identifier for each data packet is based upon a layer to which the packet belongs.	base layer and a set of enhancement layers which identify bandwidth-limited conditions of an internet protocol network between the video router and a plurality of video receivers which forward the base layer from the video router to at least two of the plurality of video receivers via the internet protocol network which selectively forward one or more of the set of enhancement layers, but fewer than all of the set of enhancement layers, to at least two of the plurality of video receivers through the internet protocol network based upon the identified bandwidth-limited conditions, wherein each layer of the layered video data stream comprises data packets, each of which is encoded with a sequence number and a layer identifier, and wherein the layer identifier for each data packet is based upon a layer to which the packet belongs. <u>Each scalable layer is associated with a unique layer identifier as specified later in this clause. The representation of a particular scalable layer with a particular layer identifier layerId does not include any scalable layer with a layer identifier greater than layerId, but it may include scalable layers with layer identifiers less than layerId. The scalable layers on which a particular scalable layer depends may be indicated in the scalability information SEI message as specified later in this clause.</u> NOTE 3 – When all scalable layers for which scalability information is provided in the scalability information SEI message have sub_pic_layer_flag[i] equal to 0, the unique layer identifier values may be set equal to (128 * dependency_id + 8 * quality_id + temporal_id), with dependency_id, quality_id, and temporal_id being the corresponding syntax elements that are associated with the VCL NAL units of a scalable layer. See ITU-T H.264 at 624. NAL unit headers for scalable bitstreams that comply with the H.264 standard comprise a layer identifier, which is comprised of the set of values including priority_id dependency_id, quality_id, and temporal_id, according to the standard. dependency_id specifies a dependency identifier for the NAL unit. dependency_id shall be equal to 0 in prefix NAL units. The assignment of values to dependency_id is constrained by the sub-bitstream extraction process as specified in clause G.8.8.1. See ITU-T H.264 at 459. priority_id specifies a priority identifier for the NAL unit. The assignment of values to priority_id is constrained by the sub-bitstream extraction process as specified in clause G.8.8.1. See ITU-T H.264 at 458.
---	--

	quality_id specifies a quality identifier for the NAL unit. quality_id shall be equal to 0 in prefix NAL units. The assignment of values to quality_id is constrained by the sub-bitstream extraction process as specified in clause G.8.8.1. The variable DQId is derived by $\text{DQId} = (\text{dependency_id} \ll 4) + \text{quality_id} \quad (\text{G-63})$ When nal_unit_type is equal to 20, the bitstream shall not contain data that result in DQId equal to 0. See ITU-T H.264 at 459. temporal_id specifies a temporal identifier for the NAL unit. The assignment of values to temporal_id is constrained by the sub-bitstream extraction process as specified in clause G.8.8.1. The value of temporal_id shall be the same for all prefix NAL units and coded slice in scalable extension NAL units of an access unit. When an access unit contains any NAL unit with nal_unit_type equal to 5 or idr_flag equal to 1, temporal_id shall be equal to 0. See ITU-T H.264 at 459. For instance, the graphic below illustrates an example of a hierarchal virtual layer structure. The layers scale regarding multiple elements, such as resolution (spatial), frame rate (temporal), quality, etc. Every unique combination of the values that comprise the layer identifier indicates the relative location of a virtual layer within the hierarchy. For instance, the in the illustration below the convergence of the spatial value for “704 x 576”, the quality value for “Minimal Quality”, and the temporal value for “3.75 fr/s” comprises a single layer within the hierarchy.
--	--



Source, accessed May 18, 2020: <https://slideplayer.com/slide/13389021/>.

For instance, in the example below there are nine layers that are possible for each spatial value (e.g. 704 x 576), or for each quality value (e.g. minimal quality), or for each temporal value (e.g. 3.75 fr/s). Bitstream subsets can be derived from all of the NAL units of a given layer and all layers with a layer identifier less than the layer identifier of said layer, forming a hierarchy of virtual layers.



Source, accessed May 18, 2020: <https://slideplayer.com/slide/13389021/>.

2. The video router of claim 1, wherein the video router further selectively forwards the one or more of the set of enhancement layers based on video

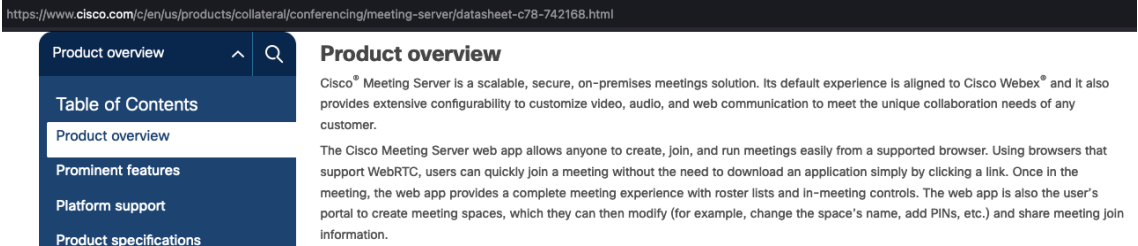
The Accused Instrumentalities include a video router that further selectively forwards the one or more of the set of enhancement layers based on video processing capabilities of the video receiver.

See claim 1, above.

STUNNING HIGH-DEFINITION (HD) VIDEO

Cisco WebEx® video supports all types of cameras, and with a supported HD webcam, you can achieve exceptionally crisp and clear video in all your meetings. Cisco® technology automatically adjusts the video throughout the meeting to the highest quality for each participant according to their available bandwidth and computer capabilities.



processing capabilities of the video receiver.	See https://www.meetingconnect.net/files/CiscoWebExVideoServices.pdf
6. A method for transmitting video signals, the method comprising:	The Accused Instrumentalities perform a method for transmitting video signals. For example, scalable video routers conforming to the H.264 standard perform a method of transmitting video signals by forwarding data packets containing scalable video information.  https://www.cisco.com/c/en/us/products/collateral/conferencing/meeting-server/datasheet-c78-742168.html

	  sciencedirect.com/topics/engineering/scalable-coding to serve the various users' requirements. Usually all participants would have to agree on some minimal quality for the video conference, but this is not really satisfactory in many cases. With the advent of SVC, the MCU is not needed because transcoding is not needed. Using the SVC standard, so-called <u>video routers</u> can be substituted for the MCUs. These routers only have to <u>forward</u> and delete packets as appropriate to that user's connection and needs, so little extra delay is added. In effect, video packets can be routed by the video router based on their headers, just like other <u>network packets</u>.
receiving a layered video data stream	The Accused Instrumentalities perform a method for transmitting video signals, the method comprising receiving a layered video data stream. For example, scalable video routers conforming to the H.264 standard perform a method of receiving a multi-layered stream:

https://www.cisco.com/c/en/us/products/collateral/conferencing/meeting-server/datasheet-c78-742168.html

Product specifications

Table of Contents

Product overview

Prominent features

Platform support

Product specifications

Ordering information

Cisco Capital

Document history

Product specifications

Table 2 lists the specifications for Cisco Meeting Server.

Table 2. Specifications

Video standards

- H.261
- H.263 (+, ++)
- H.264 AVC (baseline and high profile)
- H.264 SVC
- WebM, VP8
- Microsoft RTV
- HTML5/WebRTC
- SIP, H.323, TIP
- BFCP
- RDP
- Far End Camera Control (FECC) passthrough
- RTMP and RTMPS

<https://www.cisco.com/c/en/us/products/collateral/conferencing/meeting-server/datasheet-c78-742168.html>

29

SCALE VIDEO CODING, LLC
v. CISCO SYSTEMS, INC.

IPR2024-01371

Ex. 1007, p. 29 of 73

Feature Configuration Guide for Cisco Unified Communications Manager, Release 14 and SUs

Book Contents

Book Title Page

New and Changed Information

> Getting Started

> Remote Worker Features

> Remote Network Access

> Licensing

> Monitoring and Recording

> Call Center Features

> Voice Messaging Features

> Conferencing Features

> Placing Calls

> Receiving Calls

> Presence and Privacy Features

> Custom Features

> Device Management

> Advanced Call Processing

> QoS Management

> SIP Interoperability

Configure SIP Normalization and Transparency

Configure SDP Transparency Profiles

Configure Presentation Sharing using BFCP

Video Telephony

> Emergency Call Routing Regulations

Find Matches in This Book

Save

Trans

Chapter: Video Telephony

Updated: June 7, 2023

Was this Document helpful?

Contact Cisco
[Open a Support Case](#)
(Requires a Cisco Service Contract)

> Chapter Contents

- Video Telephony Overview
- Video Telephony Support
- Video Network
- Video Telephony Configuration Task Flow
- H.323 Video
- Video Support
- Video Features
- QoS for Video Networks

Video Telephony Overview

Unified Communications Manager supports video telephony and thus unifies the world of voice and video calls. Video endpoints use Unified CM call-handling features and access a unified voice and video solution for dialing and connecting video calls.
The Unified Communications Manager video telephony solution offers these features:

- Supports video and video-related features, such as Far-End Camera Control (FECC)
- Supports multiple logical channels that are needed to allow the transmission of video streams
- Transmits midcall, media-related messages that are needed for video (that is, transmits commands or indications that are needed for video calls)
- Supports H.323, Skinny Client Control Protocol (SCCP), and Session Initiation Protocol (SIP)
- Enhances locations and regions to provide Bandwidth Management
- Provides serviceability information, such as Call Detail Records (CDRs), about video calls

Video Telephony Support

The following sections discuss the details of video telephony in the Unified Communications Manager environment.

- Video Calls
- Real-Time Transport Control Protocol Pass-Through in MTP Topologies
- Video Codecs

Video Calls

The typical video call includes two or three Real-Time Protocol (RTP) streams in each direction (that is, four or six streams). The call can include the following stream types:

- Video (H.261, H.263, H.263+, **H.264-SVC**, X-H.264UC, H.264-AVC, H.265, AV1 and VT Camera wideband video codecs)
- Far-End Camera Control (FECC) - Optional
- Binary Floor Control Protocols (BFCP)

https://www.cisco.com/c/en/us/td/docs/voice_ip_comm/cucm/admin/14SU2/cucm_b_feature-configuration-guide-for-cisco14su2/cucm_m_video-telephony.html

30

SCALE VIDEO CODING, LLC
v. CISCO SYSTEMS, INC.

IPR2024-01371

Ex. 1007, p. 30 of 73

	← → ↻ sciencedirect.com/topics/engineering/scalable-video-coding Communicating Pictures: Delivery Across Networks David R. Bull, in Communicating Pictures, 2014 Scalable video encoding It is common in video transmission for the signal to be encoded and transmitted without explicit knowledge of the downstream network conditions. In cases where network congestion exists in a packet switched network such as the internet, routers will discard packets that they are unable to forward due to congestion. In some cases it is possible for packets within a stream to be prioritized or embedded such that the least important information is discarded before the more important data. This mechanism lends itself to scalable or layered encoding, where a video signal is composed of a hierarchical layering in terms of spatial resolution, temporal resolution or SNR. This enables devices to transmit and <u>receive multilayered video streams</u> where a base level of quality can be improved through the use of optional additional layers that enhance resolution, frame rate, and/or quality.
comprising a base layer and a set of enhancement layers;	The Accused Instrumentalities perform a method for transmitting video signals, the method comprising receiving a layered video data stream comprising a base layer and a set of enhancement layers. For example:



AEROSPACE AND DEFENSE
Land, Naval and Airborne ▼

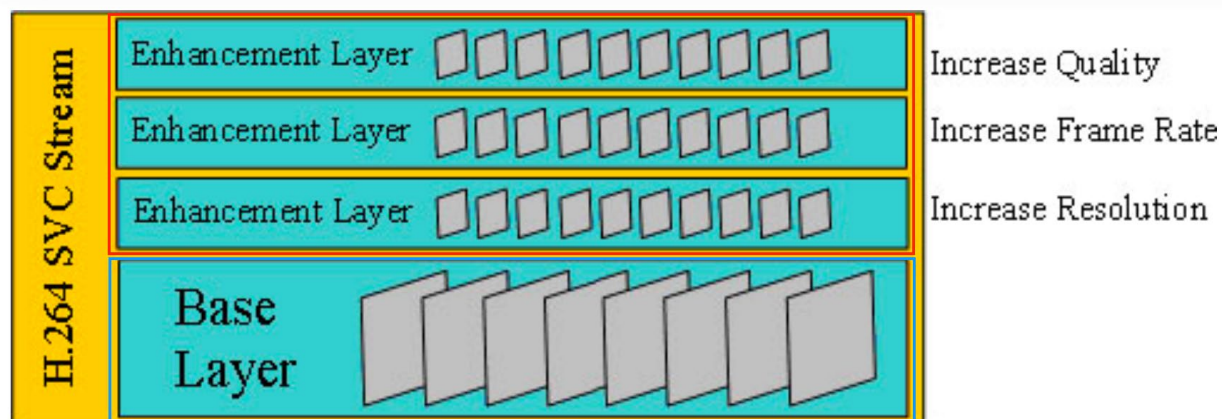


Figure 1: H.264 SVC Layering Technique

The “base layer” is the minimum decodable bitstream subset (or sub-bitstream), and every other layer is an enhancement layer.

G.3.4 base layer: A *bitstream subset* that contains all *NAL units* with the *nal_unit_type syntax element* equal to 1 and 5 of the *bitstream* and does not contain any *NAL unit* with the *nal_unit_type syntax element* equal to 14, 15, or 20 and conforms to one or more of the profiles specified in Annex A.

See ITU-T H.264 at 437.

← → ↺ [sciencedirect.com/topics/engineering/scalable-video-coding](https://www.sciencedirect.com/topics/engineering/scalable-video-coding)

Communicating Pictures: Delivery Across Networks

David R. Bull, in [Communicating Pictures](#), 2014

Scalable video encoding

It is common in video transmission for the signal to be encoded and transmitted without explicit knowledge of the downstream network conditions. In cases where network congestion exists in a [packet switched network](#) such as the internet, routers will discard packets that they are unable to forward due to congestion. In some cases it is possible for packets within a stream to be prioritized or embedded such that the least important information is discarded before the more important data. This mechanism lends itself to scalable or layered encoding, where a video signal is composed of a hierarchical layering in terms of [spatial resolution](#), [temporal resolution](#) or SNR. This enables devices to transmit and receive multilayered video streams where a base level of quality can be improved through the use of optional additional layers that enhance resolution, frame rate, and/or quality.

MPEG-4 Visual and H.264/AVC: Standards for Modern Digital Video

Berna Erol, ... Lowell Winger, in [The Essential Guide to Video Processing](#), 2009

10.3.6 Scalability

In addition to the video coding tools discussed so far, MPEG-4 Part 2 provides scalability tools that allow organization of the [bitstream](#) into base and [enhancement layers](#). The enhancement layers are transmitted and decoded depending on the bit rate, display resolution, network throughput, and decoder complexity constraints. Temporal, [spatial](#), [quality](#), complexity, and object-based Scalable video data streams coded according to the H.264 standard comprise a single bitstream that is further comprised of layers, wherein combinations of layers form “bitstream subsets” that may be decoded as video data streams.

G.3.54 scalable bitstream: A *bitstream* with the property that one or more *bitstream subsets* that are not identical to the scalable bitstream form another *bitstream* that conforms to this specification.

See ITU-T H.264 at 440.

G.3.6 bitstream subset: A *bitstream* that is derived as a *subset* from a *bitstream* by discarding zero or more *NAL units*. A *bitstream subset* is also referred to as *sub-bitstream*.

See ITU-T H.264 at 437.

	Scalable video coding is specified in Annex G allowing the construction of bitstreams that contain sub-bitstreams that conform to this Specification. For temporal bitstream scalability, i.e., the presence of a sub-bitstream with a smaller temporal sampling rate than the bitstream, complete access units are removed from the bitstream when deriving the sub-bitstream. In this case, high-level syntax and inter prediction reference pictures in the bitstream are constructed accordingly. For spatial and quality bitstream scalability, i.e., the presence of a sub-bitstream with lower spatial resolution or quality than the bitstream, NAL units are removed from the bitstream when deriving the sub-bitstream. In this case, inter-layer prediction, i.e., the prediction of the higher spatial resolution or quality signal by data of the lower spatial resolution or quality signal, is typically used for efficient coding. Otherwise, the coding algorithm as described in the previous paragraph is used. See ITU-T H.264 at 4.
identifying bandwidth-limited conditions of an internet protocol network between a video router and a plurality of video receivers;	The Accused Instrumentalities perform a method for transmitting video signals, the method comprising identifying bandwidth-limited conditions of an internet protocol network between a video router and a plurality of video receivers. For example, the Accused Instrumentalities identify bandwidth-limited conditions of an internet protocol network between a video router and a plurality of video receivers. Scalable video routers detect conditions in which bandwidth is limited. Such a bandwidth limitation could be caused by a congested state of a network over which packets must be sent, or a weak physical connection between two or more links in the network, or a limitation in the capabilities of a receiver, or some other network condition that would limit the bandwidth between one or more links.

Communicating Pictures: Delivery Across Networks

David R. Bull, in [Communicating Pictures](#), 2014

Scalable video encoding

It is common in video transmission for the signal to be encoded and transmitted without explicit knowledge of the downstream network conditions. In cases where network congestion exists in a [packet switched network](#) such as the internet, routers will discard packets that they are unable to forward due to congestion. In some cases it is possible for packets within a stream to be prioritized or embedded such that the least important information is discarded before the more important data. This mechanism lends itself to scalable or layered encoding, where a video signal is composed of a hierarchical layering in terms of [spatial resolution](#), [temporal resolution](#) or SNR. This enables devices to transmit and receive multilayered video streams where a base level of quality can be improved through the use of optional additional layers that enhance resolution, frame rate, and/or quality.

	  sciencedirect.com/topics/engineering/scalable-coding to serve the various users' requirements. Usually all participants would have to agree on some minimal quality for the video conference, but this is not really satisfactory in many cases. With the advent of SVC, the MCU is not needed because transcoding is not needed. Using the SVC standard, so-called <i>video routers</i> can be substituted for the MCUs. These routers only have to forward and delete packets <u>as appropriate to that user's connection and needs</u>, so little extra delay is added. In effect, video packets can be routed by the video router based on their headers, just like other <i>network packets</i>.
forwarding the base layer to at least two of the plurality of video receivers via the internet protocol network;	The Accused Instrumentalities perform a method for transmitting video signals, the method comprising forwarding the base layer to at least two of the plurality of video receivers via the internet protocol network. For example, because the base layer is the minimum bitstream subset of a given scalable bitstream that is capable of being decoded, every receiver that receives a decodable bitstream subset is sent at least the base layer:

	G.8.8.2 Specification of the base layer bitstream Each scalable bitstream that conforms to this specification shall contain a base layer bitstream that conforms to one or more of the profiles specified in Annex A. This base layer bitstream is derived by invoking the sub-bitstream extraction process as specified in clause G.8.8.1 with dIdTarget being equal to 0 and qIdTarget being equal to 0 and the base layer bitstream being the output. NOTE – Although all scalable bitstreams that conform to one or more of the profiles specified in this annex contain a base layer bitstream that conforms to one or more of the profiles specified in Annex A, the complete scalable bitstream (prior to operation of the base layer extraction process specified in this clause) may not conform to any profile specified in Annex A. Rec. ITU-T H.264 (06/2019) 595
and selectively forwarding one or more of the set of enhancement layers, but fewer than all of the set of enhancement layers, to at least two of the plurality of video receivers through the internet protocol network based upon the identified bandwidth-limited conditions;	The Accused Instrumentalities perform a method for transmitting video signals, the method comprising selectively forwarding one or more of the set of enhancement layers, but fewer than all of the set of enhancement layers, to at least two of the plurality of video receivers through the internet protocol network based upon the identified bandwidth-limited conditions. For example, if a given end destination can be reached at data speeds insufficient to transmit the complete scalable bitstream because of a bandwidth-limiting condition, the scalable video router will select and send all of the enhancement layers which can be sent at sufficient data speeds for each of the plurality of video receivers. For example, as shown below, fewer than all of the set of enhancement layers are forwarded when priority_id is greater than pIdTarget, or temporal_id is greater than tIdTarget, or dependency_id is greater than dIdTarget (or dependency_id is equal to dIdTarget but quality_id is greater than qIdTarget) for any VCL NAL units or filler data NAL units in the bitstream.

	<u>The sub-bitstream is derived by applying the following operations in sequential order:</u> 1. Mark all VCL NAL units and filler data NAL units for which any of the following conditions are true as "to be removed from the bitstream": – priority_id is greater than pIdTarget, – temporal_id is greater than tIdTarget, – dependency_id is greater than dIdTarget, – dependency_id is equal to dIdTarget and quality_id is greater than qIdTarget. 2. Remove all access units for which all VCL NAL units are marked as "to be removed from the bitstream". 3. Remove all VCL NAL units and filler data NAL units that are marked as "to be removed from the bitstream". 4. When dIdTarget is equal to 0 and qIdTarget is equal to 0, remove the following NAL units: – all NAL units with nal_unit_type equal to 14 or 15, – all NAL units with nal_unit_type equal to 6 in which the first SEI message has payloadType in the range of 24 to 35, inclusive. 5. Remove all NAL units with nal_unit_type equal to 6 that only contain SEI messages that are part of a scalable nesting SEI message with any of the following properties: – sei_temporal_id is greater than tIdTarget, – the minimum value of $(sei_dependency_id[i] \ll 4) + sei_quality_id[i]$ for all i in the range of 0 to num_layer_representations_minus1, inclusive, is greater than $(dIdTarget \ll 4) + qIdTarget$. 6. Remove all NAL units with nal_unit_type equal to 6 that contain SEI messages with payloadType equal to 24, 28, or 29. See ITU-T H.264 at 595. All sub-bitstreams that can be derived using the sub-bitstream extraction process as specified in clause G.8.8.1 with any combination of values for priority_id, temporal_id, dependency_id, or quality_id as the input shall result in a set of coded video sequences, with each coded video sequence conforming to one or more of the profiles specified in Annexes A and G. See ITU-T H.264 at 489.
--	--

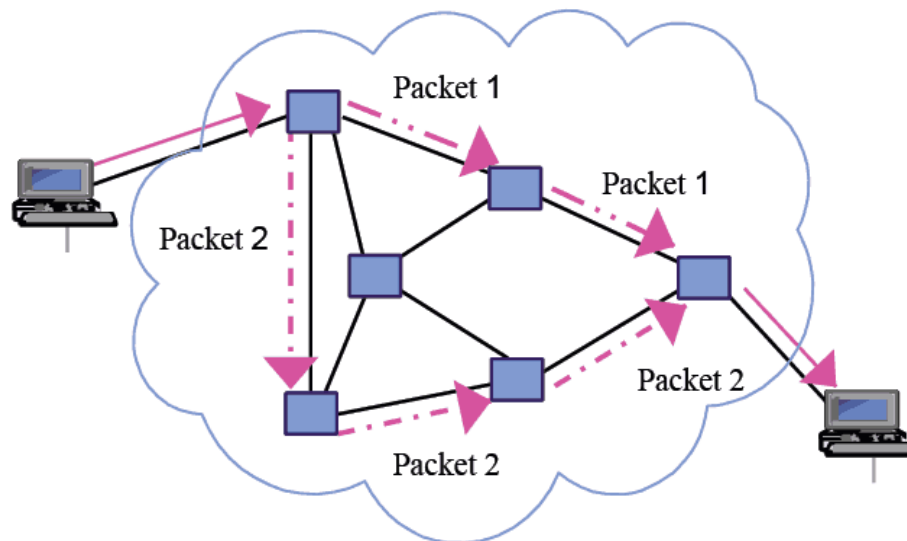
	The representation of a particular scalable layer is the set of NAL units that represents the set union of the particular scalable layer and all scalable layers on which the particular scalable layer directly or indirectly depends. The representation of a scalable layer is also referred to as scalable layer representation. In the following specification of this clause, the terms representation of a scalable layer and scalable layer representation are also used for referring to the access unit set that can be constructed from the NAL units of the scalable layer representation. A scalable layer representation can be decoded independently of all NAL units that do not belong to the scalable layer representation. The decoding result of a scalable layer representation is the set of decoded pictures that are obtained by decoding the access unit set of the scalable layer representation. See ITU-T H.264 at 624.   sciencedirect.com/topics/engineering/scalable-coding to serve the various users' requirements. Usually all participants would have to agree on some minimal quality for the video conference, but this is not really satisfactory in many cases. With the advent of SVC, the MCU is not needed because transcoding is not needed. Using the SVC standard, so-called <i>video routers</i> can be substituted for the MCUs. These routers only have to forward and delete packets <u>as appropriate to that user's connection and needs</u>, so little extra delay is added. In effect, video packets can be routed by the video router based on their headers, just like other <i>network packets</i>.
wherein the layered video data stream is transmitted according to an internet protocol;	The Accused Instrumentalities perform a method for transmitting video signals, wherein the layered video data stream is transmitted according to an internet protocol; and wherein each layer of the layered video data stream comprises data packets. For example, scalable video routers facilitate the distribution of scalable bitstreams of video information across a packet switched network, such as the internet, using internet protocol. Data packets are the result of a process of dividing information into units so that said units may be transmitted across a packet switched network. At the destination, the data packets are reassembled to produce the information from which said data packets were derived.

and wherein each layer of the layered video data stream comprises data packets,

→ ↻ networkencyclopedia.com/packet-switching/

What is Packet Switching?

Packet Switching is the process by which a networking or telecommunications device accepts a packet and switches it to a telecommunications device that will take it closer to its destination. Packet switching allows data to be sent over the telecommunications network in short bursts or “packets” that contain sequence numbers so that they can be reassembled at the destination.



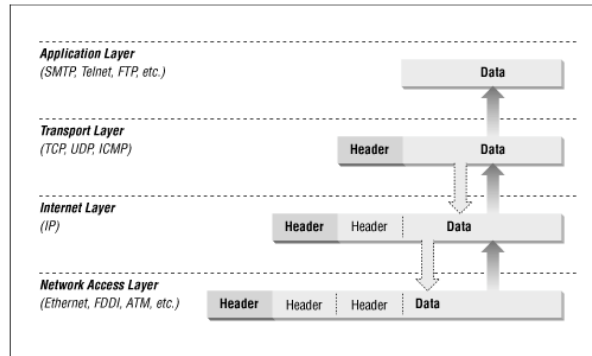
Data packets that are transmitted across the internet are comprised of a header and payload/body, according to the internet protocol. The payload is comprised of the data intended for transmission, and the header is information facilitating said transmission. Moreover, the payload of a packet is often another packet, comprising another header and another payload, as each packet layer is designed to be handled by a different part of the data distribution process. (NOTE: packet layers are not to be confused with scalable video layers).

Packets are constructed in such a way that layers for each protocol used for a particular connection are wrapped around the packets, like the layers of skin on an onion.

At each layer, a packet has two parts: the header and the body. The header contains protocol information relevant to that layer, while the body contains the data for that layer which often consists of a whole packet from the next layer in the stack. Each layer treats the information it gets from the layer above it as data, and applies its own header to this data. At each layer, the packet contains all of the information passed from the higher layer; nothing is lost. This process of preserving the data while attaching a new header is known as *encapsulation*.

At the application layer, the packet consists simply of the data to be transferred (for example, part of a file being transferred during an FTP session). As it moves to the transport layer, the Transmission Control Protocol (TCP) or the User Datagram Protocol (UDP) preserves the data from the previous layer and attaches a header to it. At the next layer, IP considers the entire packet (consisting now of the TCP or UDP header and the data) to be data, and now attaches its own IP header. Finally, at the network access layer, Ethernet or another network protocol considers the entire IP packet passed to it to be data, and attaches its own header. [Figure 6.2](#) shows how this works.

Figure 6.2: Data encapsulation



At the other side of the connection, this process is reversed. As the data is passed up from one layer to the next higher layer, each header (each skin of the onion) is stripped off by its respective layer. For example, the Internet layer removes the IP header before passing the encapsulated data up to the transport layer (TCP or UDP).

Additionally, the final payload of a scalable video data packet is a NAL Unit, which is also a unit of information comprising a header and a body. Scalable bitstreams that conform to the H.264 standard are transmitted as a sequence of NAL units.

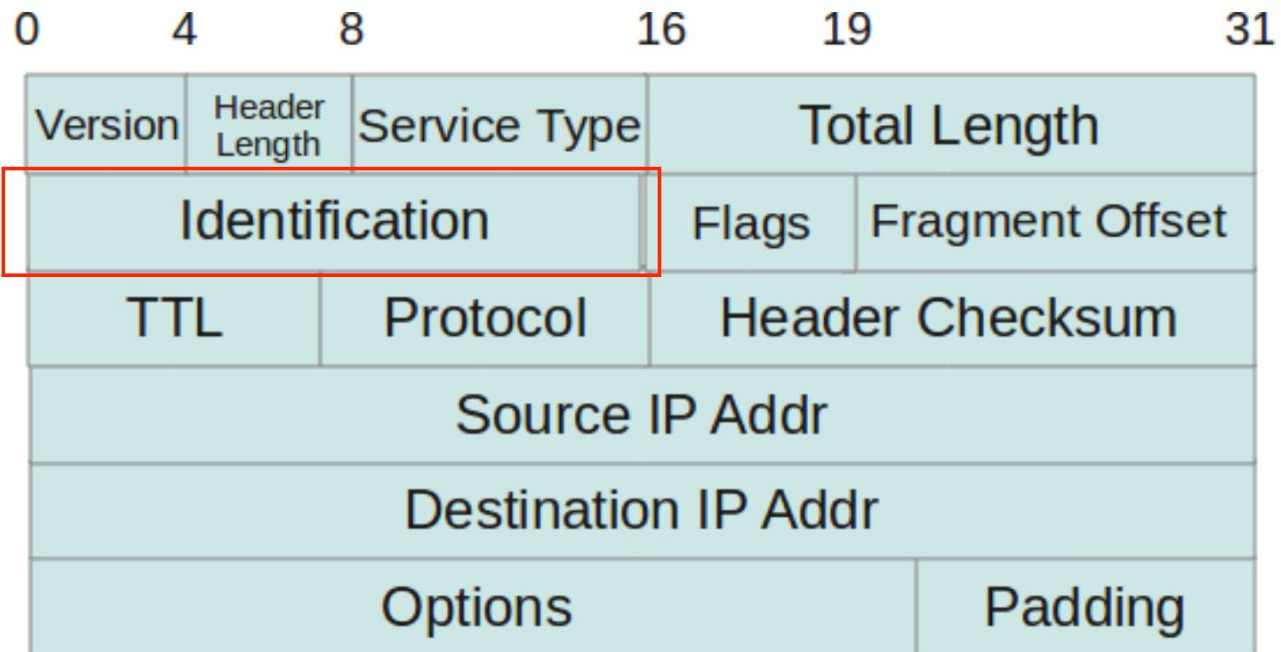
ip.hhi.de/imagecom_G1/savce/downloads/SVC-Overview.pdf

A. Network Abstraction Layer (NAL)

The coded video data are organized into NAL units, which are packets that each contains an integer number of bytes. A NAL unit starts with a one-byte header, which signals the type of the contained data. The remaining bytes represent payload data. NAL units are classified into VCL NAL units, which con-

each of which is encoded with a sequence number	The Accused Instrumentalities perform a method for transmitting video signals, wherein the layered video data stream is transmitted according to an internet protocol; and wherein each layer of the layered video data stream comprises data packets, each of which is encoded with a sequence number. For example, each data packet is identified by a sequence number which uniquely identified said packet among all other packets in the message. Every packet in the sequence contains an identification value that is one more than that of the previous packet in the sequence: ⓘ Not Secure linfo.org/packet_header.html Packet Header Definition <i>A packet header</i> is the portion of an IP (Internet protocol) <i>packet</i> that precedes its body and contains addressing and other data that is required for it to reach its intended destination. Packets are the fundamental unit of information transport in all modern computer networks, and increasingly in other communications networks as well. They can be a fixed size or variable sizes, depending on the system. Regardless of their size, each packet consists of three main parts: a header, the body, also called the <i>payload</i>, and a <i>trailer</i>. The header's format is specified in the Internet protocol. It normally contains 20 bytes of data, although an option exists within it that allows the addition of more bytes. Among the contents of the header are the version of IP (which is always set to 4, because IPv4 is being used), the sender's IP address, the intended receiver's IP address, the number of packets the message has been broken into, <u>the identification number of the particular packet</u>, the protocol (e.g. 1 for ICMP, 2 for IGMP, 6 for TCP and 17 for UDP) used, the packet length (on networks that have variable length packets), the <i>time to live</i> (i.e., the number of links or <i>hops</i> that the packet can be routed before being allowed to expire) and synchronization data (several bits that help the packet match up to the network).
--	---

IP Header



- **Identification(16 bits):** This field is used for uniquely identifying the IP datagrams. This value is incremented every-time an IP datagram is sent from source to the destination. This field comes in handy while reassembly of fragmented IP data grams.

and a layer identifier, and

The Accused Instrumentalities perform a method for transmitting video signals, wherein the layered video data stream is transmitted according to an internet protocol; and wherein each layer of the layered video data stream comprises

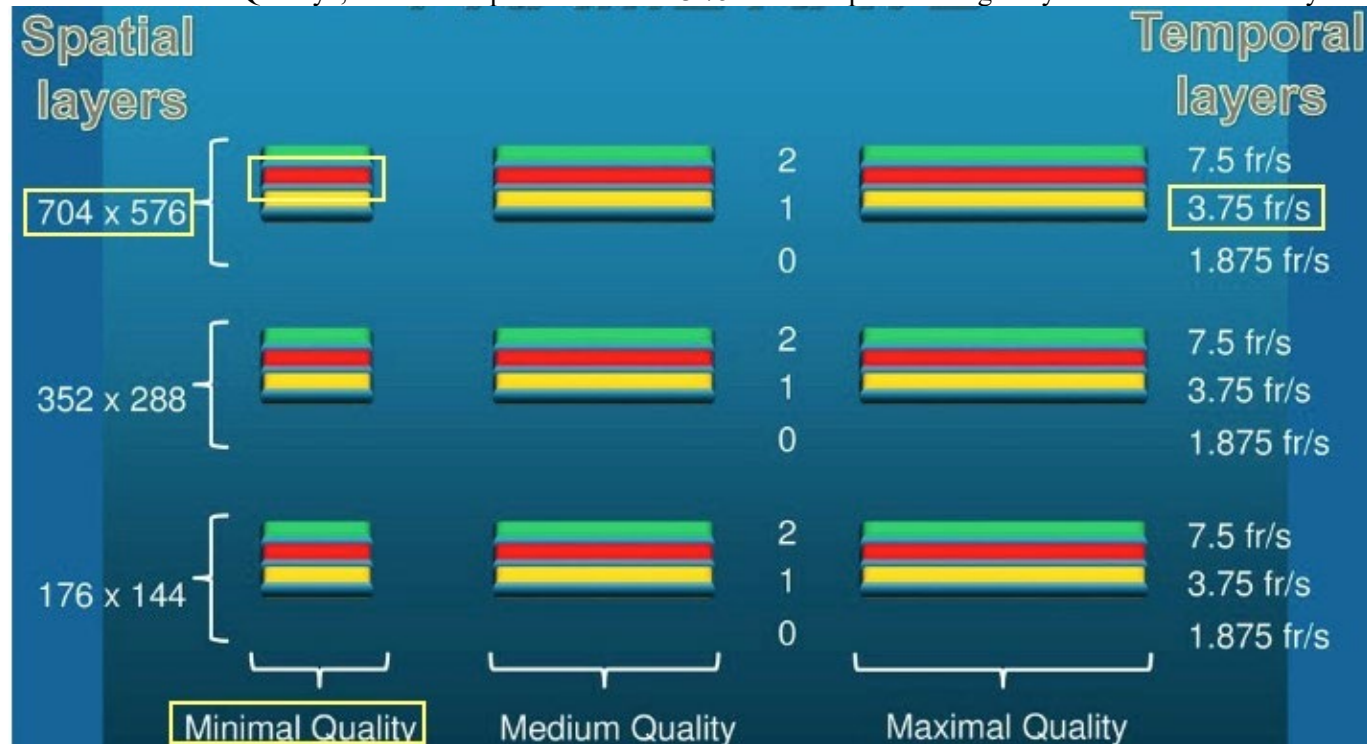
wherein the layer identifier for each data packet is based upon a layer to which the packet belongs.	data packets, each of which is encoded with a sequence number and a layer identifier, and wherein the layer identifier for each data packet is based upon a layer to which the packet belongs. For example: <u>Each scalable layer is associated with a unique layer identifier</u> as specified later in this clause. The representation of a particular scalable layer with a particular layer identifier <code>layerId</code> does not include any scalable layer with a layer identifier greater than <code>layerId</code>, but it may include scalable layers with layer identifiers less than <code>layerId</code>. The scalable layers on which a particular scalable layer depends may be indicated in the scalability information SEI message as specified later in this clause. NOTE 3 – When all scalable layers for which scalability information is provided in the scalability information SEI message have <code>sub_pic_layer_flag[i]</code> equal to 0, the unique layer identifier values may be set equal to $(128 * \text{dependency_id} + 8 * \text{quality_id} + \text{temporal_id})$, with <code>dependency_id</code>, <code>quality_id</code>, and <code>temporal_id</code> being the corresponding syntax elements that are associated with the VCL NAL units of a scalable layer. See ITU-T H.264 at 624. NAL unit headers for scalable bitstreams that comply with the H.264 standard comprise a layer identifier, which is comprised of the set of values including <code>priority_id</code>, <code>dependency_id</code>, <code>quality_id</code>, and <code>temporal_id</code>, according to the standard. dependency_id specifies a dependency identifier for the NAL unit. <code>dependency_id</code> shall be equal to 0 in prefix NAL units. The assignment of values to <code>dependency_id</code> is constrained by the sub-bitstream extraction process as specified in clause G.8.8.1. See ITU-T H.264 at 459. priority_id specifies a priority identifier for the NAL unit. The assignment of values to <code>priority_id</code> is constrained by the sub-bitstream extraction process as specified in clause G.8.8.1. See ITU-T H.264 at 458. quality_id specifies a quality identifier for the NAL unit. <code>quality_id</code> shall be equal to 0 in prefix NAL units. The assignment of values to <code>quality_id</code> is constrained by the sub-bitstream extraction process as specified in clause G.8.8.1. The variable <code>DQId</code> is derived by $\text{DQId} = (\text{dependency_id} \ll 4) + \text{quality_id} \quad (\text{G-63})$ When <code>nal_unit_type</code> is equal to 20, the bitstream shall not contain data that result in <code>DQId</code> equal to 0. See ITU-T H.264 at 459.
---	--

temporal_id specifies a temporal identifier for the NAL unit. The assignment of values to temporal_id is constrained by the sub-bitstream extraction process as specified in clause G.8.8.1.

The value of temporal_id shall be the same for all prefix NAL units and coded slice in scalable extension NAL units of an access unit. When an access unit contains any NAL unit with nal_unit_type equal to 5 or idr_flag equal to 1, temporal_id shall be equal to 0.

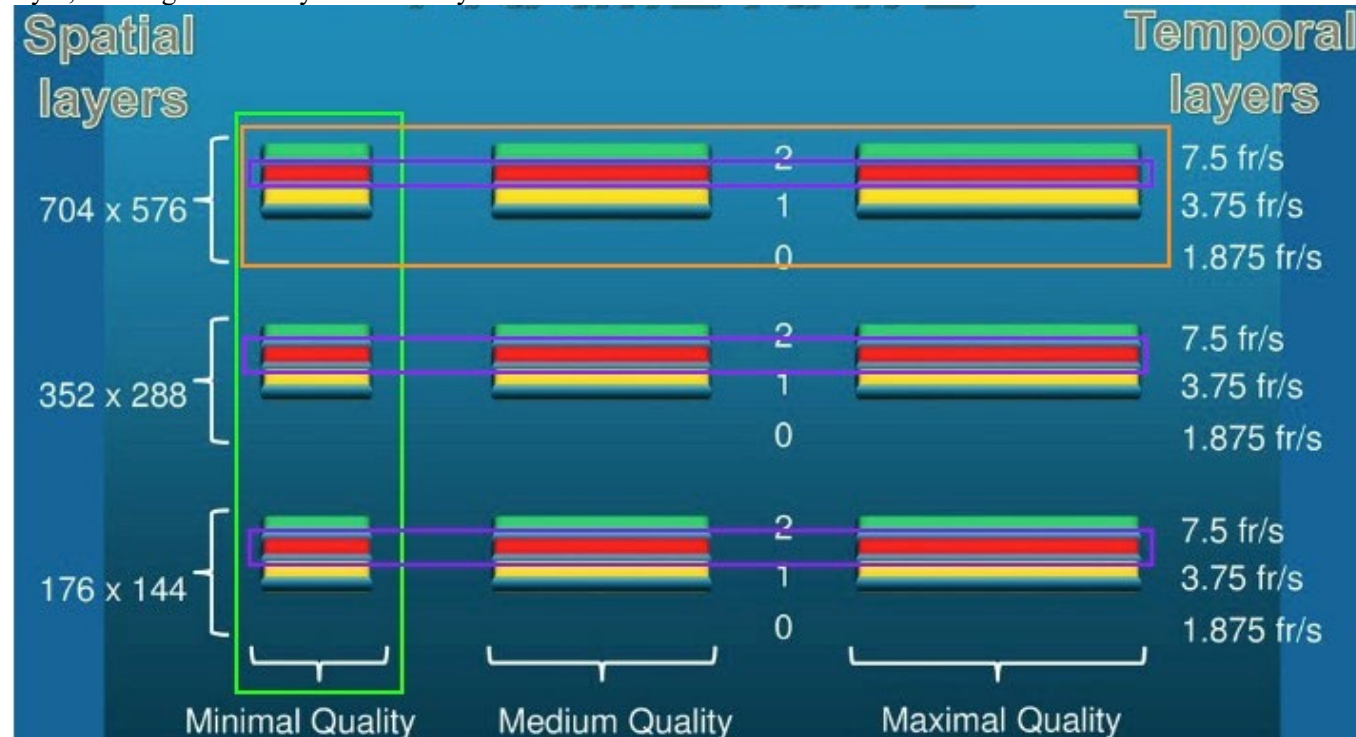
See ITU-T H.264 at 459.

For instance, the graphic below illustrates an example of a hierarchical virtual layer structure. The layers scale regarding multiple elements, such as resolution (spatial), frame rate (temporal), quality, etc. Every unique combination of the values that comprise the layer identifier indicates the relative location of a virtual layer within the hierarchy. For instance, the in the illustration below the convergence of the spatial value for “704 x 576”, the quality value for “Minimal Quality”, and the temporal value for “3.75 fr/s” comprises a single layer within the hierarchy.



Source: <https://slideplayer.com/slide/13389021/> (last accessed August 29, 2023)

For instance, in the example below there are nine layers that are possible for each spatial value (e.g. 704 x 576), or for each quality value (e.g. minimal quality), or for each temporal value (e.g. 3.75 fr/s). Bitstream subsets can be derived from all of the NAL units of a given layer and all layers with a layer identifier less than the layer identifier of said layer, forming a hierarchy of virtual layers.



Source: <https://slideplayer.com/slide/13389021/> (last accessed August 29, 2023)

7. The method of claim 6, wherein the selectively forwarding step further comprises:

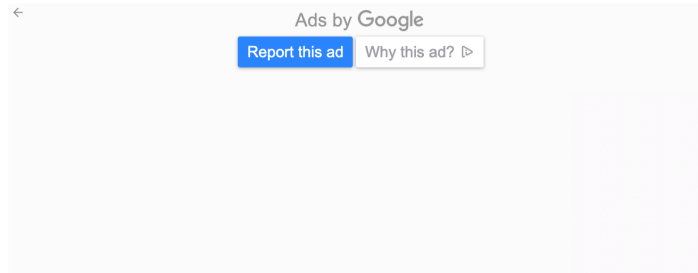
The Accused Instrumentality include a method for transmitting video signals wherein the selectively forwarding step further comprises selectively forwarding the one or more of the set of enhancement layers based on video processing capabilities of the video receiver.

See claim 6, above.

selectively forwarding the one or more of the set of enhancement layers based on video processing capabilities of the video receiver.	STUNNING HIGH-DEFINITION (HD) VIDEO Cisco WebEx[®] video supports all types of cameras, and with a supported HD webcam, you can achieve exceptionally crisp and clear video in all your meetings. Cisco[®] technology automatically adjusts the video throughout the meeting to the highest quality for each participant according to their available bandwidth and computer capabilities.  See https://www.meetingconnect.net/files/CiscoWebExVideoServices.pdf
11. A scalable video coding router, comprising:	The Accused Instrumentalities include a scalable video coding router. For example, scalable video routers conforming to the H.264 standard perform a method of transmitting video signals by forwarding data packets containing scalable video information.

	  sciencedirect.com/topics/engineering/scalable-coding to serve the various users' requirements. Usually all participants would have to agree on some minimal quality for the video conference, but this is not really satisfactory in many cases. With the advent of SVC, the MCU is not needed because transcoding is not needed. Using the SVC standard, so-called <u>video routers</u> can be substituted for the MCUs. These routers only have to <u>forward</u> and delete packets as appropriate to that user's connection and needs, so little extra delay is added. In effect, video packets can be routed by the video router based on their headers, just like other <u>network packets</u>.
a memory; and a processor, wherein the processor executes instructions stored in the memory to cause the <u>scalable</u> video <u>coding</u> router to:	The Accused Instrumentalities include a scalable video coding router comprising a memory and a processor, wherein the processor executes instructions stored in the memory. All routers comprise a memory and a processor, wherein the processor executes instructions stored in the memory.

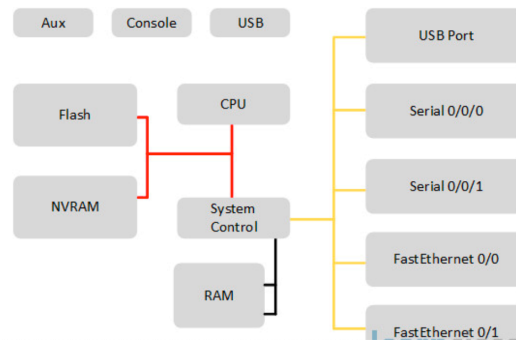
Router Internal Components Overview



I want you to take a moment and look at the components we have listed here.

- CPU - Executes operating system instructions
- RAM (Random access memory) - Contains the running copy of the configuration file. Stores the routing table. RAM contents lost when power is off.
- ROM (Read-only memory) - Holds diagnostic software used when router is powered up. Stores the router's bootstrap program.
- NVRAM (Nonvolatile RAM) - Stores startup configuration. This may include IP addresses (routing protocol, hostname of a router).
- Flash memory - Contains the operating system (Cisco IOS)
- Interfaces - Multiple physical interfaces that are used to connect the network. Fast Ethernet, Gigabit Ethernet; Serial; Management interfaces

CPU, RAM, ROM, interfaces - those all sound very familiar to something else. A computer, a PC! All of these components we find inside the tower of our PC or inside of our laptop. So we have our router, we look inside, we have the same types of components. Let's go through these individually.



receive a layered video data stream	The Accused Instrumentalities include a scalable video coding router comprising a memory and a processor, wherein the processor executes instructions stored in the memory to cause the scalable video coding router to receive a layered video data stream. The accused instrumentalities conform to the H.264 standard which perform a method of receiving a multi-layered stream. ← → ↻ sciencedirect.com/topics/engineering/scalable-video-coding Communicating Pictures: Delivery Across Networks David R. Bull, in Communicating Pictures, 2014 Scalable video encoding It is common in video transmission for the signal to be encoded and transmitted without explicit knowledge of the downstream network conditions. In cases where network congestion exists in a packet switched network such as the internet, routers will discard packets that they are unable to forward due to congestion. In some cases it is possible for packets within a stream to be prioritized or embedded such that the least important information is discarded before the more important data. This mechanism lends itself to scalable or layered encoding, where a video signal is composed of a hierarchical layering in terms of spatial resolution, temporal resolution or SNR. This enables devices to transmit and <u>receive multilayered video streams</u> where a base level of quality can be improved through the use of optional additional layers that enhance resolution, frame rate, and/or quality.
--	--

that comprises a base layer and a set of enhancement layers,

The Accused Instrumentalities include a scalable video coding router comprising a memory and a processor, wherein the processor executes instructions stored in the memory to cause the scalable video coding router to receive a layered video data stream that comprises a base layer and a set of enhancement layers.

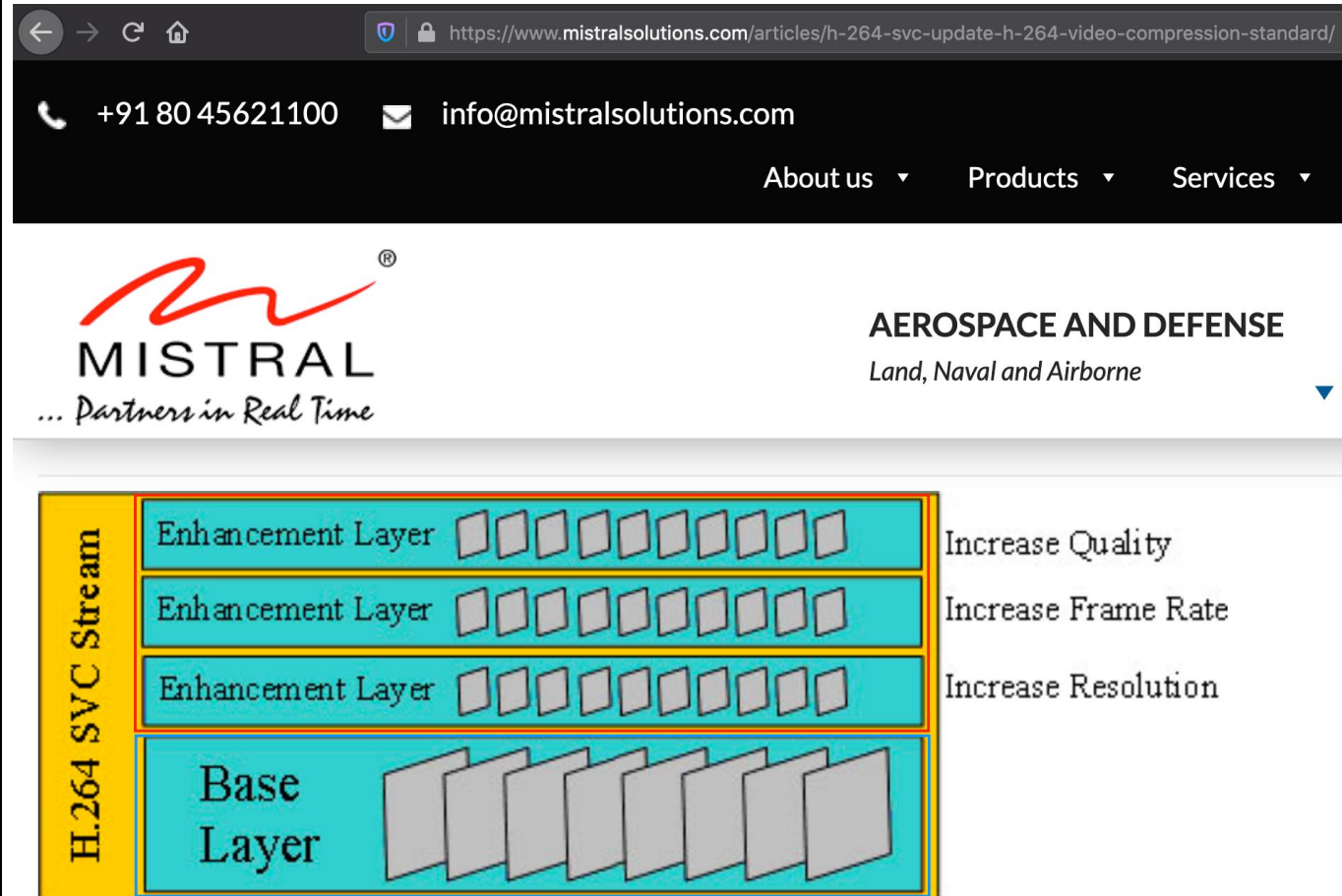


Figure 1: H.264 SVC Layering Technique

The “base layer” is the minimum decodable bitstream subset (or sub-bitstream), and every other layer is an enhancement layer.

G.3.4 base layer: *A bitstream subset that contains all NAL units with the nal_unit_type syntax element equal to 1 and 5 of the bitstream and does not contain any NAL unit with the nal_unit_type syntax element equal to 14, 15, or 20 and conforms to one or more of the profiles specified in Annex A.*

[CITE]

← → ↺ [sciencedirect.com/topics/engineering/scalable-video-coding](https://www.sciencedirect.com/topics/engineering/scalable-video-coding)

Communicating Pictures: Delivery Across Networks

David R. Bull, in [Communicating Pictures](#), 2014

Scalable video encoding

It is common in video transmission for the signal to be encoded and transmitted without explicit knowledge of the downstream network conditions. In cases where network congestion exists in a [packet switched network](#) such as the internet, routers will discard packets that they are unable to forward due to congestion. In some cases it is possible for packets within a stream to be prioritized or embedded such that the least important information is discarded before the more important data. This mechanism lends itself to scalable or layered encoding, where a video signal is composed of a hierarchical layering in terms of [spatial resolution](#), [temporal resolution](#) or SNR. This enables devices to transmit and receive multilayered video streams where a base level of quality can be improved through the use of optional additional layers that enhance resolution, frame rate, and/or quality.

MPEG-4 Visual and H.264/AVC: Standards for Modern Digital Video

Berna Erol, ... Lowell Winger, in [The Essential Guide to Video Processing](#), 2009

10.3.6 Scalability

In addition to the video coding tools discussed so far, MPEG-4 Part 2 provides scalability tools that allow organization of the [bitstream](#) into base and [enhancement layers](#). The enhancement layers are transmitted and decoded depending on the bit rate, display resolution, network throughput, and decoder complexity constraints. Temporal, [spatial](#), [quality](#), complexity, and object-based

Scalable video data streams coded according to the H.264 standard comprise a single bitstream that is further comprised of layers, wherein combinations of layers form “bitstream subsets” that may be decoded as video data streams.

G.3.54 scalable bitstream: A *bitstream* with the property that one or more *bitstream subsets* that are not identical to the scalable bitstream form another *bitstream* that conforms to this specification.

See ITU-T H.264 at 437.

G.3.6 bitstream subset: A *bitstream* that is derived as a *subset* from a *bitstream* by discarding zero or more *NAL units*. A *bitstream subset* is also referred to as *sub-bitstream*.

See ITU-T H.264 at 437.

	Scalable video coding is specified in Annex G allowing the construction of bitstreams that contain sub-bitstreams that conform to this Specification. For temporal bitstream scalability, i.e., the presence of a sub-bitstream with a smaller temporal sampling rate than the bitstream, complete access units are removed from the bitstream when deriving the sub-bitstream. In this case, high-level syntax and inter prediction reference pictures in the bitstream are constructed accordingly. For spatial and quality bitstream scalability, i.e., the presence of a sub-bitstream with lower spatial resolution or quality than the bitstream, NAL units are removed from the bitstream when deriving the sub-bitstream. In this case, inter-layer prediction, i.e., the prediction of the higher spatial resolution or quality signal by data of the lower spatial resolution or quality signal, is typically used for efficient coding. Otherwise, the coding algorithm as described in the previous paragraph is used. See ITU-T H.264 at 4.
identify bandwidth-limited conditions of an internet protocol network between the video router and a set of video receivers;	The Accused Instrumentalities include a scalable video coding router comprising a memory and a processor, wherein the processor executes instructions stored in the memory to cause the scalable video coding router to receive a layered video data stream that comprises a base layer and a set of enhancement layers which identifies bandwidth-limited conditions of an internet protocol network between the video router and a plurality of video receivers. Scalable video routers detect conditions in which bandwidth is limited. Such a bandwidth limitation could be caused by a congested state of a network over which packets must be sent, or a weak physical connection between two or more links in the network, or a limitation in the capabilities of a receiver, or some other network condition that would limit the bandwidth between one or more links.

Communicating Pictures: Delivery Across Networks

David R. Bull, in [Communicating Pictures](#), 2014

Scalable video encoding

It is common in video transmission for the signal to be encoded and transmitted without explicit knowledge of the downstream network conditions. In cases where network congestion exists in a [packet switched network](#) such as the internet, routers will discard packets that they are unable to forward due to congestion. In some cases it is possible for packets within a stream to be prioritized or embedded such that the least important information is discarded before the more important data. This mechanism lends itself to scalable or layered encoding, where a video signal is composed of a hierarchical layering in terms of [spatial resolution](#), [temporal resolution](#) or SNR. This enables devices to transmit and receive multilayered video streams where a base level of quality can be improved through the use of optional additional layers that enhance resolution, frame rate, and/or quality.

	  sciencedirect.com/topics/engineering/scalable-coding to serve the various users' requirements. Usually all participants would have to agree on some minimal quality for the video conference, but this is not really satisfactory in many cases. With the advent of SVC, the MCU is not needed because transcoding is not needed. Using the SVC standard, so-called <i>video routers</i> can be substituted for the MCUs. These routers only have to forward and delete packets <u>as appropriate to that user's connection and needs</u>, so little extra delay is added. In effect, video packets can be routed by the video router based on their headers, just like other <i>network packets</i>.
forward the base layer from the video router to at least two of the set of video receivers;	The Accused Instrumentalities include a scalable video coding router comprising a memory and a processor, wherein the processor executes instructions stored in the memory to cause the scalable video coding router to receive a layered video data stream that comprises a base layer and a set of enhancement layers which identifies bandwidth-limited conditions of an internet protocol network between the video router and a plurality of video receivers and forwards the base layer from the video router to at least two of the plurality of video receivers via the internet protocol network. Because the base layer is the minimum bitstream subset of a given scalable bitstream that is capable of being decoded, every receiver that receives a decodable bitstream subset is sent at least the base layer.

	G.8.8.2 Specification of the base layer bitstream Each scalable bitstream that conforms to this specification shall contain a base layer bitstream that conforms to one or more of the profiles specified in Annex A. This base layer bitstream is derived by invoking the sub-bitstream extraction process as specified in clause G.8.8.1 with dIdTarget being equal to 0 and qIdTarget being equal to 0 and the base layer bitstream being the output. NOTE – Although all scalable bitstreams that conform to one or more of the profiles specified in this annex contain a base layer bitstream that conforms to one or more of the profiles specified in Annex A, the complete scalable bitstream (prior to operation of the base layer extraction process specified in this clause) may not conform to any profile specified in Annex A. Rec. ITU-T H.264 (06/2019) 595
wherein the scalable video coding router forwards all of the set of enhancement layers to at least two of the video receivers in the set of video receivers with bandwidth-sufficient conditions,	The Accused Instrumentalities include a scalable video coding router comprising a memory and a processor, wherein the processor executes instructions stored in the memory to cause the scalable video coding router to receive a layered video data stream that comprises a base layer and a set of enhancement layers which identifies bandwidth-limited conditions of an internet protocol network between the video router and a plurality of video receivers and forwards the base layer from the video router to at least two of the plurality of video receivers via the internet protocol network wherein the scalable video coding router forwards all of the set of enhancement layers to at least two of the video receivers in the set of video receivers with bandwidth-sufficient conditions. To all of the end destinations (video receivers) that can be reached at data speeds sufficient to transmit the complete scalable bitstream (i.e. can be reached with sufficient bandwidth), the scalable video router will select and send all of the enhancement layers of the scalable video bitstream. For instance, as shown below, all of the set of enhancement layers are forwarded when priority_id is less than or equal to pIdTarget, temporal_id is less than or equal to tIdTarget, and dependency_id is less than dIdTarget (or dependency_id is equal to dIdTarget but quality_id is less than or equal to qIdTarget) for all VCL NAL units and filler data NAL units in the bitstream.

	<u>The sub-bitstream is derived by applying the following operations in sequential order:</u> 1. Mark all VCL NAL units and filler data NAL units for which any of the following conditions are true as "to be removed from the bitstream": – priority_id is greater than pIdTarget, – temporal_id is greater than tIdTarget, – dependency_id is greater than dIdTarget, – dependency_id is equal to dIdTarget and quality_id is greater than qIdTarget. 2. Remove all access units for which all VCL NAL units are marked as "to be removed from the bitstream". 3. Remove all VCL NAL units and filler data NAL units that are marked as "to be removed from the bitstream". 4. When dIdTarget is equal to 0 and qIdTarget is equal to 0, remove the following NAL units: – all NAL units with nal_unit_type equal to 14 or 15, – all NAL units with nal_unit_type equal to 6 in which the first SEI message has payloadType in the range of 24 to 35, inclusive. 5. Remove all NAL units with nal_unit_type equal to 6 that only contain SEI messages that are part of a scalable nesting SEI message with any of the following properties: – sei_temporal_id is greater than tIdTarget, – the minimum value of $(\text{sei_dependency_id}[i] \ll 4) + \text{sei_quality_id}[i]$ for all i in the range of 0 to num_layer_representations_minus1, inclusive, is greater than $(\text{dIdTarget} \ll 4) + \text{qIdTarget}$. 6. Remove all NAL units with nal_unit_type equal to 6 that contain SEI messages with payloadType equal to 24, 28, or 29.
wherein the scalable video coding router selectively forwards one or more of the set of enhancement layers, but fewer than all	The Accused Instrumentalities include a scalable video coding router comprising a memory and a processor, wherein the processor executes instructions stored in the memory to cause the scalable video coding router to receive a layered video data stream that comprises a base layer and a set of enhancement layers which identifies bandwidth-limited conditions of an internet protocol network between the video router and a plurality of video receivers and forwards the base layer from the video router to at least two of the plurality of video receivers via the internet protocol network wherein the scalable video coding router forwards all of the set of enhancement layers to at least two of the video receivers in the set of video receivers with bandwidth-sufficient conditions wherein the scalable video coding router selectively forwards one or more of the set of enhancement layers, but fewer than all of the set of enhancement layers,

of the set of enhancement layers, to at least two of the remaining video receivers in the set of video receivers based upon the identified bandwidth-limited conditions;	to at least two of the plurality of video receivers through the internet protocol network based upon the identified bandwidth-limited conditions. If a given end destination can be reached at data speeds insufficient to transmit the complete scalable bitstream because of a bandwidth limitation, the scalable video router will select and send all of the enhancement layers which can be sent at sufficient data speeds for each of the plurality of video receivers. For instance, as shown below, fewer than all of the set of enhancement layers are forwarded when priority_id is greater than pIdTarget, or temporal_id is greater than tIdTarget, or dependency_id is greater than dIdTarget (or dependency_id is equal to dIdTarget but quality_id is greater than qIdTarget) for any VCL NAL units or filler data NAL units in the bitstream.
---	--

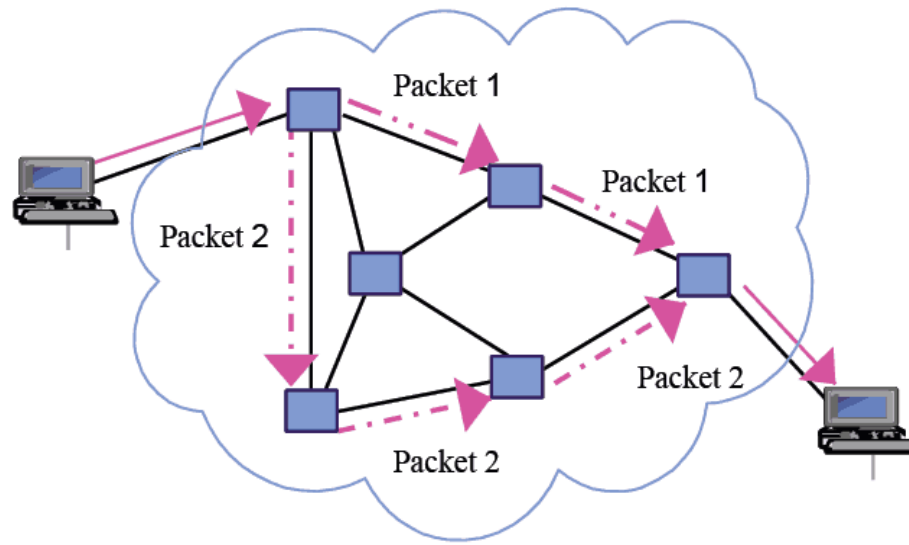
	<u>The sub-bitstream is derived by applying the following operations in sequential order:</u> 1. Mark all VCL NAL units and filler data NAL units for which any of the following conditions are true as "to be removed from the bitstream": – priority_id is greater than pIdTarget, – temporal_id is greater than tIdTarget, – dependency_id is greater than dIdTarget, – dependency_id is equal to dIdTarget and quality_id is greater than qIdTarget. 2. Remove all access units for which all VCL NAL units are marked as "to be removed from the bitstream". 3. Remove all VCL NAL units and filler data NAL units that are marked as "to be removed from the bitstream". 4. When dIdTarget is equal to 0 and qIdTarget is equal to 0, remove the following NAL units: – all NAL units with nal_unit_type equal to 14 or 15, – all NAL units with nal_unit_type equal to 6 in which the first SEI message has payloadType in the range of 24 to 35, inclusive. 5. Remove all NAL units with nal_unit_type equal to 6 that only contain SEI messages that are part of a scalable nesting SEI message with any of the following properties: – sei_temporal_id is greater than tIdTarget, – the minimum value of (sei_dependency_id[i] << 4) + sei_quality_id[i] for all i in the range of 0 to num_layer_representations_minus1, inclusive, is greater than (dIdTarget << 4) + qIdTarget. 6. Remove all NAL units with nal_unit_type equal to 6 that contain SEI messages with payloadType equal to 24, 28, or 29. See ITU-T H.264 at 595. All sub-bitstreams that can be derived using the sub-bitstream extraction process as specified in clause G.8.8.1 with any combination of values for priority_id, temporal_id, dependency_id, or quality_id as the input shall result in a set of coded video sequences, with each coded video sequence conforming to one or more of the profiles specified in Annexes A and G. See ITU-T H.264 at 595.
--	---

	The representation of a particular scalable layer is the set of NAL units that represents the set union of the particular scalable layer and all scalable layers on which the particular scalable layer directly or indirectly depends. The representation of a scalable layer is also referred to as scalable layer representation. In the following specification of this clause, the terms representation of a scalable layer and scalable layer representation are also used for referring to the access unit set that can be constructed from the NAL units of the scalable layer representation. A scalable layer representation can be decoded independently of all NAL units that do not belong to the scalable layer representation. The decoding result of a scalable layer representation is the set of decoded pictures that are obtained by decoding the access unit set of the scalable layer representation. See ITU-T H.264 at 595.   sciencedirect.com/topics/engineering/scalable-coding to serve the various users' requirements. Usually all participants would have to agree on some minimal quality for the video conference, but this is not really satisfactory in many cases. With the advent of SVC, the MCU is not needed because transcoding is not needed. Using the SVC standard, so-called <i>video routers</i> can be substituted for the MCUs. These routers only have to forward and delete packets <u>as appropriate to that user's connection and needs</u>, so little extra delay is added. In effect, video packets can be routed by the video router based on their headers, just like other <i>network packets</i>.
wherein the layered video data stream is transmitted according to an internet	The Accused Instrumentalities include a scalable video coding router comprising a memory and a processor, wherein the processor executes instructions stored in the memory to cause the scalable video coding router to receive a layered video data stream that comprises a base layer and a set of enhancement layers which identifies bandwidth-limited conditions of an internet protocol network between the video router and a plurality of video receivers and forwards the base layer from the video router to at least two of the plurality of video receivers via the internet protocol network wherein the scalable video coding router forwards all of the set of enhancement layers to at least two of the video

protocol, and wherein each layer of the layered video data stream comprises data packets,	receivers in the set of video receivers with bandwidth-sufficient conditions wherein the scalable video coding router selectively forwards one or more of the set of enhancement layers, but fewer than all of the set of enhancement layers, to at least two of the plurality of video receivers through the internet protocol network based upon the identified bandwidth-limited conditions wherein the layered video data stream is transmitted according to an internet protocol, and wherein each layer of the layered video data stream comprises data packets. Scalable video routers facilitate the distribution of scalable bitstreams of video information across a packet switched network, such as the internet. Data packets are the result of a process of dividing information into units so that said units may be transmitted across a packet switched network. At the destination, the data packets are reassembled to produce the information from which said data packets were derived.
---	--

What is Packet Switching?

Packet Switching is the process by which a networking or telecommunications device accepts a packet and switches it to a telecommunications device that will take it closer to its destination. Packet switching allows data to be sent over the telecommunications network in short bursts or “packets” that contain sequence numbers so that they can be reassembled at the destination.



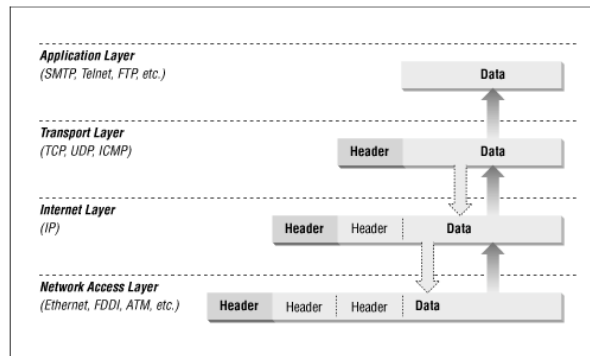
Data packets that are transmitted across the internet are comprised of a header and payload/body, according to the internet protocol. The payload is comprised of the data intended for transmission, and the header is information facilitating said transmission. Moreover, the payload of a packet is often another packet, comprising another header and another payload, as each packet layer is designed to be handled by a different part of the data distribution process. (NOTE: packet layers are not to be confused with scalable video layers).

Packets are constructed in such a way that layers for each protocol used for a particular connection are wrapped around the packets, like the layers of skin on an onion.

At each layer, a packet has two parts: the header and the body. The header contains protocol information relevant to that layer, while the body contains the data for that layer which often consists of a whole packet from the next layer in the stack. Each layer treats the information it gets from the layer above it as data, and applies its own header to this data. At each layer, the packet contains all of the information passed from the higher layer; nothing is lost. This process of preserving the data while attaching a new header is known as *encapsulation*.

At the application layer, the packet consists simply of the data to be transferred (for example, part of a file being transferred during an FTP session). As it moves to the transport layer, the Transmission Control Protocol (TCP) or the User Datagram Protocol (UDP) preserves the data from the previous layer and attaches a header to it. At the next layer, IP considers the entire packet (consisting now of the TCP or UDP header and the data) to be data, and now attaches its own IP header. Finally, at the network access layer, Ethernet or another network protocol considers the entire IP packet passed to it to be data, and attaches its own header. [Figure 6.2](#) shows how this works.

Figure 6.2: Data encapsulation



At the other side of the connection, this process is reversed. As the data is passed up from one layer to the next higher layer, each header (each skin of the onion) is stripped off by its respective layer. For example, the Internet layer removes the IP header before passing the encapsulated data up to the transport layer (TCP or UDP).

Additionally, the final payload of a scalable video data packet is a NAL Unit, which is also a unit of information comprising a header and a body. Scalable bitstreams that conform to the H.264 standard are transmitted as a sequence of NAL units.

	ip.hhi.de/imagecom_G1/savce/downloads/SVC-Overview.pdf A. <i>Network Abstraction Layer (NAL)</i> The coded video data are organized into NAL units, which are packets that each contains an integer number of bytes. A NAL unit starts with a one-byte header, which signals the type of the contained data. The remaining bytes represent payload data. NAL units are classified into VCL NAL units, which con-
each of which is encoded with a sequence number	The Accused Instrumentalities include a scalable video coding router comprising a memory and a processor, wherein the processor executes instructions stored in the memory to cause the scalable video coding router to receive a layered video data stream that comprises a base layer and a set of enhancement layers which identifies bandwidth-limited conditions of an internet protocol network between the video router and a plurality of video receivers and forwards the base layer from the video router to at least two of the plurality of video receivers via the internet protocol network wherein the scalable video coding router forwards all of the set of enhancement layers to at least two of the video receivers in the set of video receivers with bandwidth-sufficient conditions wherein the scalable video coding router selectively forwards one or more of the set of enhancement layers, but fewer than all of the set of enhancement layers, to at least two of the plurality of video receivers through the internet protocol network based upon the identified bandwidth-limited conditions wherein the layered video data stream is transmitted according to an internet protocol, and wherein each layer of the layered video data stream comprises data packets, each of which is encoded with a sequence number. Each data packet is identified by a sequence number which uniquely identified said packet among all other packets in the message. Every packet in the sequence contains an identification value that is one more than that of the previous packet in the sequence.

Packet Header Definition

A *packet header* is the portion of an [IP](#) (Internet protocol) *packet* that precedes its body and contains addressing and other data that is required for it to reach its intended destination.

Packets are the fundamental unit of information transport in all modern computer [networks](#), and increasingly in other communications networks as well. They can be a fixed size or variable sizes, depending on the system. Regardless of their size, each packet consists of three main parts: a header, the body, also called the *payload*, and a *trailer*.

The header's format is specified in the Internet protocol. It normally contains 20 [bytes](#) of data, although an option exists within it that allows the addition of more bytes.

Among the contents of the header are the version of IP (which is always set to 4, because IPv4 is being used), the sender's [IP address](#), the intended receiver's IP address, the number of packets the message has been broken into, the identification number of the particular packet, the [protocol](#) (e.g. 1 for [ICMP](#), 2 for IGMP, 6 for [TCP](#) and 17 for [UDP](#)) used, the packet length (on networks that have variable length packets), the *time to live* (i.e., the number of links or [hops](#) that the packet can be routed before being allowed to expire) and synchronization data (several [bits](#) that help the packet match up to the network).

IP Header

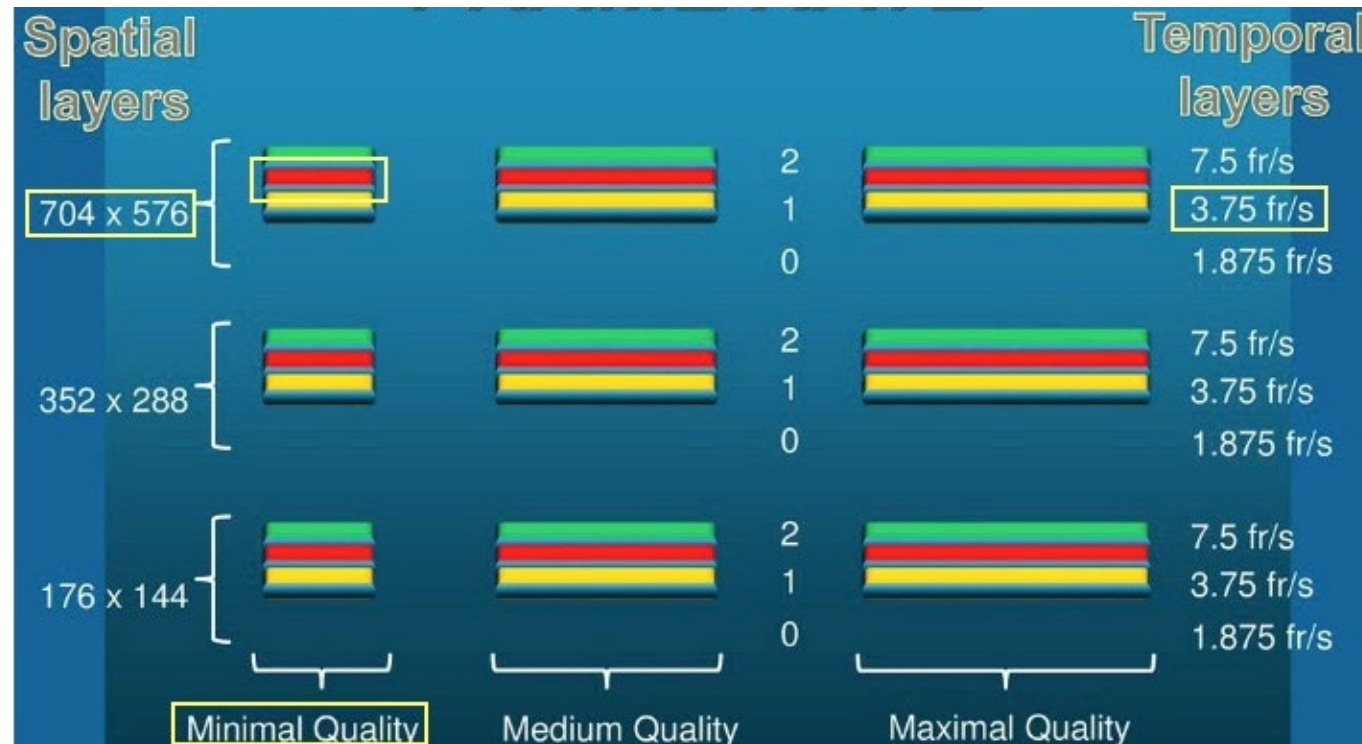
0	4	8	16	19	31
Version	Header Length	Service Type	Total Length		
Identification			Flags	Fragment Offset	
TTL		Protocol	Header Checksum		
Source IP Addr					
Destination IP Addr					
Options				Padding	

- **Identification(16 bits):** This field is used for uniquely identifying the IP datagrams. This value is incremented every-time an IP datagram is sent from source to the destination. This field comes in handy while reassembly of fragmented IP data grams.

and a layer identifier, and wherein the layer identifier for each data packet is based upon a layer to which the packet belongs.	The Accused Instrumentalities include a scalable video coding router comprising a memory and a processor, wherein the processor executes instructions stored in the memory to cause the scalable video coding router to receive a layered video data stream that comprises a base layer and a set of enhancement layers which identifies bandwidth-limited conditions of an internet protocol network between the video router and a plurality of video receivers and forwards the base layer from the video router to at least two of the plurality of video receivers via the internet protocol network wherein the scalable video coding router forwards all of the set of enhancement layers to at least two of the video receivers in the set of video receivers with bandwidth-sufficient conditions wherein the scalable video coding router selectively forwards one or more of the set of enhancement layers, but fewer than all of the set of enhancement layers, to at least two of the plurality of video receivers through the internet protocol network based upon the identified bandwidth-limited conditions wherein the layered video data stream is transmitted according to an internet protocol, and wherein each layer of the layered video data stream comprises data packets, each of which is encoded with a sequence number and a layer identifier, and wherein the layer identifier for each data packet is based upon a layer to which the packet belongs. <u>Each scalable layer is associated with a unique layer identifier as specified later in this clause. The representation of a particular scalable layer with a particular layer identifier layerId does not include any scalable layer with a layer identifier greater than layerId, but it may include scalable layers with layer identifiers less than layerId. The scalable layers on which a particular scalable layer depends may be indicated in the scalability information SEI message as specified later in this clause.</u> NOTE 3 – When all scalable layers for which scalability information is provided in the scalability information SEI message have sub_pic_layer_flag[i] equal to 0, the unique layer identifier values may be set equal to (128 * dependency_id + 8 * quality_id + temporal_id), with dependency_id, quality_id, and temporal_id being the corresponding syntax elements that are associated with the VCL NAL units of a scalable layer. See ITU-T H.264 at 624. NAL unit headers for scalable bitstreams that comply with the H.264 standard comprise a layer identifier, which is comprised of the set of values including priority_id dependency_id, quality_id, and temporal_id, according to the standard.
---	---

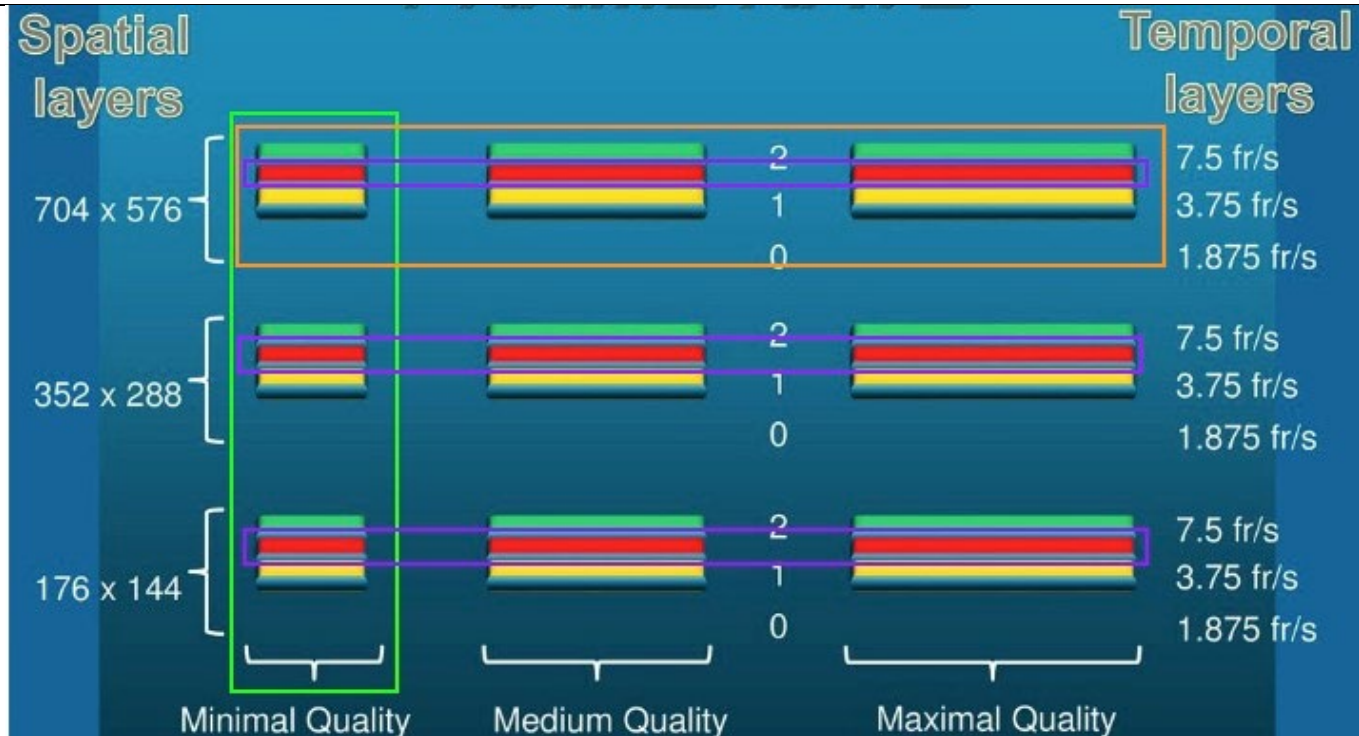
	dependency_id specifies a dependency identifier for the NAL unit. <code>dependency_id</code> shall be equal to 0 in prefix NAL units. The assignment of values to <code>dependency_id</code> is constrained by the sub-bitstream extraction process as specified in clause G.8.8.1. See ITU-T H.264 at 459. priority_id specifies a priority identifier for the NAL unit. The assignment of values to <code>priority_id</code> is constrained by the sub-bitstream extraction process as specified in clause G.8.8.1. See ITU-T H.264 at 458. quality_id specifies a quality identifier for the NAL unit. <code>quality_id</code> shall be equal to 0 in prefix NAL units. The assignment of values to <code>quality_id</code> is constrained by the sub-bitstream extraction process as specified in clause G.8.8.1. The variable DQId is derived by $\text{DQId} = (\text{dependency_id} \ll 4) + \text{quality_id} \quad (\text{G-63})$ When <code>nal_unit_type</code> is equal to 20, the bitstream shall not contain data that result in DQId equal to 0. See ITU-T H.264 at 459. temporal_id specifies a temporal identifier for the NAL unit. The assignment of values to <code>temporal_id</code> is constrained by the sub-bitstream extraction process as specified in clause G.8.8.1. The value of <code>temporal_id</code> shall be the same for all prefix NAL units and coded slice in scalable extension NAL units of an access unit. When an access unit contains any NAL unit with <code>nal_unit_type</code> equal to 5 or <code>idr_flag</code> equal to 1, <code>temporal_id</code> shall be equal to 0. See ITU-T H.264 at 459. For instance, the graphic below illustrates an example of a hierarchal virtual layer structure. The layers scale regarding multiple elements, such as resolution (spatial), frame rate (temporal), quality, etc. Every unique combination of the values that comprise the layer identifier indicates the relative location of a virtual layer within the
--	--

hierarchy. For instance, the in the illustration below the convergence of the spatial value for “704 x 576”, the quality value for “Minimal Quality”, and the temporal value for “3.75 fr/s” comprises a single layer within the hierarchy.



Source, accessed May 18, 2020: <https://slideplayer.com/slide/13389021/>.

For instance, in the example below there are nine layers that are possible for each spatial value (e.g. 704 x 576), or for each quality value (e.g. minimal quality), or for each temporal value (e.g. 3.75 fr/s). Bitstream subsets can be derived from all of the NAL units of a given layer and all layers with a layer identifier less than the layer identifier of said layer, forming a hierarchy of virtual layers.



Source, accessed May 18, 2020: <https://slideplayer.com/slide/13389021/>.

12. The video router of claim 11, wherein the video router further selectively forwards the one or more of the set of enhancement layers based

The Accused Instrumentalities include a video router wherein the video router further selectively forwards the one or more of the set of enhancement layers based on video processing capabilities of each of the set of video receivers.

See claim 11, above.

on video processing capabilities of each of the set of video receivers.	STUNNING HIGH-DEFINITION (HD) VIDEO Cisco WebEx® video supports all types of cameras, and with a supported HD webcam, you can achieve exceptionally crisp and clear video in all your meetings. Cisco® technology automatically adjusts the video throughout the meeting to the highest quality for each participant according to their available bandwidth and computer capabilities.  See https://www.meetingconnect.net/files/CiscoWebExVideoServices.pdf
--	---