



## **Optibase White Paper A Guide to MPEG-4**

What is MPEG-4? 1

---

---

So, what's the difference between MPEG-4 and MPEG 1 and 2? 1

---

---

Main Principles of MPEG-4 2

---

---

Alternative Streaming Solutions 5

---

---

MPEG-4 Applications 5

---

---

Market Potential of MPEG-4 6

---

---

Glossary 6

### **Copyright Information**

The contents of this publication may not be reproduced in any form by any means, in part or in whole, without the prior written permission of the publisher. The authors and publisher make no warranty of any kind with regard to this material, including, but not limited to, the implied warranties of merchantability and fitness for any particular purpose. Neither shall the authors or publisher be liable for any errors contained herein or for incidental or consequential damages in connection with the furnishing or use of this material. The information herein is subject to change without notice.

## What is MPEG-4?

---

MPEG-4 (ISO14496) is an ISO/IEC standard developed by MPEG (Moving Picture Experts Group). The first version of the MPEG-4 standard was finalized in October 1998 and became an international standard at the beginning of 1999. Although defined as one standard, MPEG-4 is actually a set of compression/decompression formats and streaming technologies that address the need for distributing rich interactive media over narrow and broadband networks.

The communication revolution triggered by the Internet, the advent of wireless devices and the promise of the Next Generation Internet (broadband Internet) underscores the importance of an international standard that defines a universal way of transmitting rich media. To this end, MPEG-4 aims to pave the way toward a uniform, high-quality streaming standard, that would replace the many proprietary streaming technologies in use today.

MPEG-4 has been designed to address the following issues:

- ⇒ **Interoperability.** The standard is not specific to any one platform but is designed for all platforms.
- ⇒ **Transport Independence.** MPEG-4 leaves the choice of transport mechanism up to the service provider. This allows MPEG-4 to be used in a wide range of networking environments.
- ⇒ **Compression and Transmission of Rich Media.** MPEG-4 has been designed for the low and mid bit-rate compression and transmission of rich media streams.
- ⇒ **Interactivity.** MPEG-4 allows content authors and viewers to influence how they interact with a stream.
- ⇒ **Scalability.** MPEG-4 allows for flexibility in the way multimedia streams are decoded. Decoding bit rate and resolution of content is adapted to the networking environment and display device. This quality is necessary when transmitting rich media over heterogeneous networks, as well as for applications where the receiver is not capable of displaying the full resolution or full quality images.
- ⇒ **Profiles.** MPEG-4 offers different technology profiles for different applications. In this way, service providers need not use the entire set of technologies, but only the sub-set that suits their applications needs.

## So, what's the difference between MPEG-4 and MPEG 1 and 2?

---

MPEG-1 and MPEG-2 are standards that focus on the compression and decompression of audio and video streams. Both standards address the needs of audio and video transport and synchronization. MPEG-1 was designed to provide a compression standard for media such as Video CD and CD-ROM, which have a typical playback rate of 1.2 Mbit/s. MPEG-2 was designed to provide higher quality for transmission applications, focusing mainly on Digital TV applications.

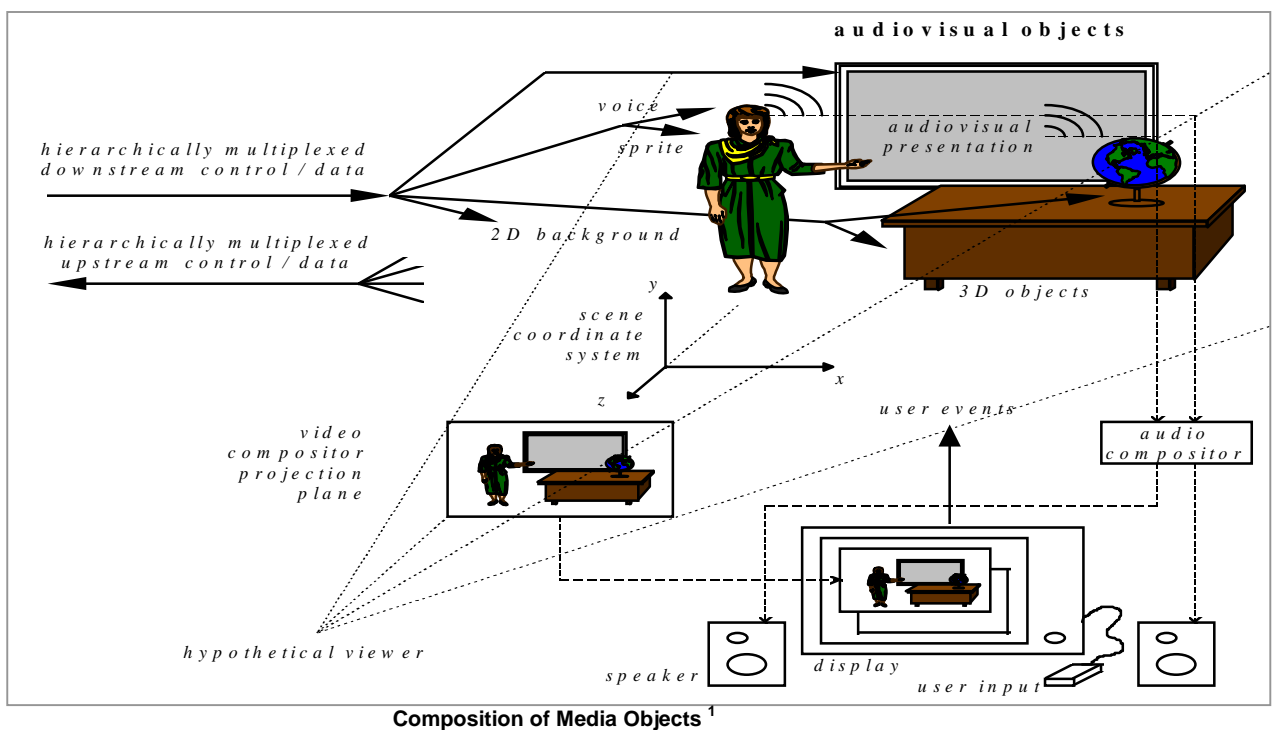
The major difference between MPEG-4 and MPEG-1 and 2 is the way MPEG-4 relates to the application level. MPEG-4 defines content that needs to be delivered over a network as a framework of media objects and scene descriptions. While MPEG-1 and MPEG-2 relate only to audio-video streams, MPEG-4 allows for the inclusion of other types of content such as animation, computer generated objects as well as video and audio. In MPEG-4, each component that comprises a multimedia scene is considered a media object. Each media object has spatial and temporal attributes that govern its behavior and location in the multimedia scene.

In addition to the concept of media objects, the MPEG-4 standard specifies that the transport mechanism of the multimedia stream need not be defined by the standard, but by the service provider or application developer. In contrast to MPEG-1 and 2, MPEG-4 defines streaming, synchronization and content rendering so as to accommodate bursty content delivery, scalable content delivery and to enable interactivity. Such requirements are intended to address the streaming of rich media over heterogeneous networks at bit-rates as low as 24 Kbit/s.

Although MPEG-4 covers more or less the same encoding range as MPEG-1 and MPEG-2, its target applications are different. MPEG-4 defines interactivity, scalability and streaming of rich media. Thus content compressed according to the MPEG-4 standard can be streamed over the broad or narrowband Internet, used in Interactive TV applications or streamed to wireless appliances such as cellular phones and PDAs (Personal Digital Assistants).

## Main Principles of MPEG-4

MPEG-4 aims to achieve its objectives by applying certain principles to the way data is represented. MPEG-4 relates to the components that comprise a multimedia scene as media objects. For example, a sound track, animation, video or image are all individual media objects. Media objects can be grouped together to form compound objects. These are the building blocks of multimedia scenes. But these media objects are only one part of an MPEG-4 stream. Additional information that governs how the objects are rendered on the screen and how they are transmitted over networks is also needed. For these purposes, MPEG-4 streams include Stream Description information and Coding information. The Screen Description information describes the relation between the media objects and how they are presented. The Coding information describes how the media objects are linked to the resources that are transmitting the media objects.<sup>1</sup>



## MPEG-4 Architecture

The MPEG-4 Standards comprises several core parts:

- ⇒ **MPEG-4 Systems.** This part of the standard describes, scene description, multiplexing, synchronization, buffer management and protection of intellectual property.
- ⇒ **Delivery Multimedia Integration Framework (DMIF).** This part of the standard defines rich media streaming.
- ⇒ **MPEG-4 Visual.** This part of the standard specifies the representation of natural and synthetic visual objects.
- ⇒ **MPEG-4 Audio.** This part of the standard specifies the representation of natural and synthetic audio objects.

<sup>1</sup> Slide by Philip A. Chou, in [An Overview of the MPEG-4 Standard](http://www.cseit.stet.it/ufv/leonardo/paper/isce96.htm#MPEG_4_or_multimedia_communications) by [Leonardo Chiariglione](#) at [http://www.cseit.stet.it/ufv/leonardo/paper/isce96.htm#MPEG\\_4\\_or\\_multimedia\\_communications](http://www.cseit.stet.it/ufv/leonardo/paper/isce96.htm#MPEG_4_or_multimedia_communications)



## **MPEG-4 Systems**

MPEG-4 Systems specifies the overall architecture of the standard and defines how MPEG-4 Visual and MPEG-4 Audio are integrated. In addition to dealing with multiplexing, synchronization and buffer management, MPEG-4 Systems introduces the concept of BIFS (Binary Format for Scenes). BIFS defines the interactive aspects of MPEG-4 content. Another foundation of MPEG-4 Systems is the framework of object descriptors. Object descriptors describe the elements that make up an MPEG-4 stream.

MPEG-4 relates to the components that comprise a multimedia scene as media objects. For example, a sound track, animation, video or image are all

grouped together to form compound objects. These are the building blocks of multimedia scenes.

All information relating to media objects, scene description or control information is contained in elementary streams. Elementary streams are information carriers. Elementary streams contain tags or pointers, called Object Descriptors, which determine how an MPEG-4 stream is decoded at the receiving station. Object Descriptors enable receiving stations to recognize the type of media being streamed and present it correctly. Object Descriptors identify the streams associated with one media object. This allows content authors to determine the hierarchy of media objects and apply meta information to the multimedia stream. All elementary streams are stored in the Sync layer. The Synchronization layer ensures that Elementary streams use a common system for conveying timing and framing information.

## **DMIF**

MPEG-4 has been designed for a wide range of applications and bit rates. The MPEG-4 standard is delivery unaware, and leaves decisions regarding the transport network up to service developers. For this reason, the standard deals with delivery and compression in two separate architectures. The architecture that governs delivery is called DMIF - (Delivery Multimedia Integration Framework). DMIF specifies how the MPEG-4 stream interfaces with different networking technologies and protocols. DMIF provides the overall delivery structure of MPEG-4 streams. DMIF covers areas such as billing, Quality of Service (QOS), broadcast requirements and interactivity. The bridge between DMIF and MPEG-4 Systems is called DAI (DMIF Application Interface).

The receiving station accesses the multimedia stream through the DAI. A DAI filter handles the request and determines the type of DMIF that is being requested based on the URL supplied by the application. An application can request more than one DMIF service, in accordance with the type of transport technologies needed. For example, one DMIF can specify IP multicasting while another can specify satellite broadcasts. In this regard, DMIF is designed to support simultaneous transmission of multiples streams over multiple transport technologies and protocols.

## **MPEG-4 Visual**

The MPEG-4 Visual standard allows encoding of natural (pixel based) images and video together with synthetic (computer generated) scenes. It also supports the compression of synthetic 2-D and 3-D graphic geometry parameters (i.e. compression of wire grid parameters, synthetic text). MPEG-4 Visual supports encoding bit rates between 5 Kbps and 10 Mbps, with resolutions from QSIF to Full D-1.

## **MPEG-4 Audio**

MPEG-4 Audio. This part of the standard specifies the representation of natural and synthetic audio objects. The standard defines audio coding tools that can encode at bit rates as low as 2 Kbps.

## Alternative Streaming Solutions

---

In addition to MPEG-4, other digital video and audio formats are currently used in streaming rich media today.

### **RealVideo and RealAudio from Real Networks**

RealVideo and RealAudio are video and audio compression technologies developed by Real Networks for media streaming over low bandwidths, mainly the Internet. Real Networks also offer client/server tools that allow RealVideo to be streamed over the Internet.

### **Microsoft Netshow/ Windows Media Technologies (WMT)**

WMT is a video and audio compression technology developed by Microsoft for streaming media over the Internet. This technology is incorporated into WMT server and client architecture. WMT uses certain elements of the MPEG-4 standard.

### **QuickTime**

QuickTime is a file format and streaming architecture that stores, streams and plays media files. Stored and streamed media is compressed with a multitude of compression schemes, both standard and proprietary (for example, H.263 or Sorensen). Streaming is implemented via the RTP protocol.

## MPEG-4 Applications

---

### ***Rich media streaming over the Internet***

MPEG-4 has several characteristics that make it the ideal standard for streaming rich media over the Internet.

- ⇒ For the narrowband Internet, applications can use content compressed at rates as low as 24 Kbit/s. For the broadband Internet, applications can use the same content encoded at higher bit rates.
- ⇒ The interactive nature of MPEG-4 means that MPEG-4 content can be used in advanced multimedia applications.
- ⇒ Because MPEG-4 allows for scalability, the same content can be streamed to different devices over heterogeneous networks.

### ***Video Streaming to Wireless Devices***

The MPEG-4 standard allows for streaming of very low bit rate content over all types of networks. In addition, MPEG-4 makes provisions for streaming in error-prone environments. These qualities are crucial when streaming rich content to wireless devices.

### ***Interactive TV***

Broadband broadcast applications can take advantage of the MPEG-4 standard to offer high-quality interactive content delivered over traditional TV networks or cable TV networks.

### ***Interactive Home Shopping***

MPEG-4's interactive character allows shoppers to evaluate goods online and place orders in real-time.

## ***Distance Learning and Training***

One of the keys to distance learning and training is the ability to transmit over different networking infrastructures and interactivity. In a corporate training scenario, MPEG-4 content can be broadcast via satellite to company branches in remote locations and over the LAN to employees at headquarters.

## **Market Potential of MPEG-4**

---

MPEG-4's strength lies in its interoperability and scalability, but many people have questioned MPEG-4's relevance for Internet applications in view of the widespread prevalence of Microsoft's Windows Media Technology and Real Networks streaming technologies. Despite the widespread prevalence of Real and WMT streaming technology, there is broad potential for MPEG-4, especially in the PDA sector. These appliances are based on chip sets and not downloadable software and they could form the basis for the broad introduction of MPEG-4 content.

Real and WMT technology focus to a large extent on the low bit-rate spectrum whereas MPEG-4 can greatly benefit broadband applications, which want to offer rich media streaming and interactivity.

## **Glossary**

---

### **Bandwidth**

The amount of data that can be transmitted in a fixed amount of time. For digital devices, the bandwidth is usually expressed in bits per second or bytes per second. For analog devices, the bandwidth is expressed in cycles per second, or Hertz (Hz).

### **BIFS**

Binary Format for Scenes. Architecture that describes scene composition information, both spatial and temporal.

### **Buffer**

Space allocated on a system's Random Access Memory (RAM) where data is stored temporarily until it is transferred to another part of the system. In streaming applications, buffers store video or audio data until there is enough information for the stream to be composed.

### **Compression Layer**

This layer takes care of media encoding and decoding.

### **Delivery Layer**

This layer makes sure that playback devices can access content regardless of the delivery technology used.

### **DMIF – Delivery Multimedia Integration Framework**

The part of the MPEG-4 standard that defines how multimedia streaming is managed.

### **Elementary Streams**

Streams that convey individual MPEG-4 media.

### **IP Multicast**

To send information over the Internet to a group of computers that share the same IP address. Multicasting is an efficient way to transmit information since the same message is sent once to an entire group.

### **Latency**

In networking, latency is the amount of time it takes a packet to travel from source to destination. Together, latency and bandwidth define the speed and capacity of a network. In digital video networking applications, latency is measured by the time it takes one frame of video to reach its destination.

### **Media Layer**

The component that carries out the decoding of the MPEG-4 content.

### **MP3.**

MP3 is the MPEG audio layer 3 standard. Layer 3 is one of three coding schemes (layer 1, layer 2 and layer 3) for the compression of audio signals defined by the MPEG committee. Layer 3 uses perceptual audio coding and psychoacoustic compression to remove the redundant parts of a sound signal. It also adds a MDCT (Modified Discrete Cosine Transform) that implements a filter bank, increasing the frequency resolution 18 times higher than that of MPEG audio layer 2.

### **MPEG-1**

ISO/IEC standard designed for low bandwidth of compressed digital video and audio.

### **MPEG-2**

ISO/IEC standard designed for transmission of high bandwidth compressed digital audio and video such as that used by broadcast television.

### **MPEG-4 Audio**

The part of the MPEG-4 standard that defines how natural and synthetic audio objects are coded.

### **MPEG-4 Systems**

The part of the MPEG-4 standard defines the overall architecture. MPEG-4 specifies scene description, multiplexing , synchronization, buffer management and management of intellectual property.

### **MPEG-4 Visual**

The part of the MPEG-4 standard that defines how natural and synthetic visual objects are represented and coded.

### **Multicast**

When data is transmitted to a defined group of recipients. For example, a video stream that is transmitted to a group of clients that share the same IP address. Standards such as TCP/IP allow users to join multicast groups.

### **Object Descriptors (OD)**

Unique identifiers that contain pointers to elementary streams.

### **Quality of Service (QoS)**

Quality of Service refers to the way data is transmitted between two hosts on a network. Networking protocols that offer QoS make sure that when information needs to be communicated, the sender requests a designated path with the network for a connection to the destination. The sender specifies the type, speed and other attributes of the call, which determine and guarantee the end-to-end quality of service.

### **Streaming**

A technique for transferring data so that it is received as a continuous real-time stream. Streaming refers mainly to audio and video data, which is time-dependent. Video files, especially, are very large and cannot be downloaded easily by home Internet users. Streamed data is transmitted by a server application and received and displayed in real-time by client applications. These applications can start displaying video or playing back audio as soon as enough data has been received and stored in the application's buffer.

### **Sync Layer**

The layer that specifies how elementary streams are packetized.

---

### **Bibliography**

Multimedia Systems, Standards, and Networks, Marcel Dekker, Inc, 2000

MPEG-4 Systems: Elementary Stream Management and Delivery, *Herpel C, Eleftheriadis A, Franceschini G*

MPEG-4: Why, What, How and When, *Pereira F*

MPEG-4 Systems: Overview, *Avaro O, Eleftheriadis A, Herpel C, Ganesh R, Liam W*

Binary Format for Scene (BIFS): Combining MPEG-4 media to build rich multimedia services, *Signes J*

Delivery and Control of MPEG-4 Content over IP Networks, *Basso A, Reha Civanlar M*

MPEG-4: Multimedia for our time, *Koenen R*

*The MPEG Home Page - <http://drogo.cselt.stet.it/mpeg/>*

An Overview of the MPEG-4 Standard, *Leonardo Chiariglione*