ITUPublications
Recommendations

International Telecommunication Union
Standardization Sector

Recommendation

ITU-T H.644.7 (09/2023)

SERIES H: Audiovisual and multimedia systems

Broadband, triple-play and advanced multimedia services – Content delivery and ubiquitous sensor network applications

Functional architecture for media processing services

ITU-T H-SERIES RECOMMENDATIONS

Audiovisual and multimedia systems

CHARACTERISTICS OF VISUAL TELEPHONE SYSTEMS	H.100-H.199
INFRASTRUCTURE OF AUDIOVISUAL SERVICES	H.200-H.499
MOBILITY AND COLLABORATION PROCEDURES	H.500-H.549
VEHICULAR GATEWAYS AND INTELLIGENT TRANSPORTATION SYSTEMS (ITS)	H.550-H.599
BROADBAND, TRIPLE-PLAY AND ADVANCED MULTIMEDIA SERVICES	H.600-H.699
Broadband multimedia services over VDSL	H.610-H.619
Advanced multimedia services and applications	H.620-H.629
Content delivery and ubiquitous sensor network applications	**H.640-H.649**
IPTV MULTIMEDIA SERVICES AND APPLICATIONS FOR IPTV	H.700-H.799
E-HEALTH MULTIMEDIA SYSTEMS, SERVICES AND APPLICATIONS	H.800-H.899

For further details, please refer to the list of ITU-T Recommendations.

Recommendation ITU-T H.644.7

Functional architecture for media processing services

Summary

Recommendation ITU-T H.644.7 specifies the functional architecture for media processing services. In particular, the scope of this Recommendation includes the domains and functional roles relationship, functional architecture and reference points for media processing services. Media processing services utilize a set of techniques including cloud computing, computing resource virtualization, and job queue processing to dynamically control and manage all kinds of computing resources, which improves scalability, flexibility, and availability.

History*

Edition	Recommendation	Approval	Study Group	Unique ID
1.0	ITU-T H.644.7	2023-09-13	16	11.1002/1000/15632

Keywords

Artificial intelligence (AI), audio processing, content delivery network (CDN), media processing service (MPS), video processing.

* To access the Recommendation, type the URL https://handle.itu.int/ in the address field of your web browser, followed by the Recommendation's unique ID.

FOREWORD

The International Telecommunication Union (ITU) is the United Nations specialized agency in the field of telecommunications, information and communication technologies (ICTs). The ITU Telecommunication Standardization Sector (ITU-T) is a permanent organ of ITU. ITU-T is responsible for studying technical, operating and tariff questions and issuing Recommendations on them with a view to standardizing telecommunications on a worldwide basis.

The World Telecommunication Standardization Assembly (WTSA), which meets every four years, establishes the topics for study by the ITU-T study groups which, in turn, produce Recommendations on these topics.

The approval of ITU-T Recommendations is covered by the procedure laid down in WTSA Resolution 1.

In some areas of information technology which fall within ITU-T's purview, the necessary standards are prepared on a collaborative basis with ISO and IEC.

NOTE

In this Recommendation, the expression "Administration" is used for conciseness to indicate both a telecommunication administration and a recognized operating agency.

Compliance with this Recommendation is voluntary. However, the Recommendation may contain certain mandatory provisions (to ensure, e.g., interoperability or applicability) and compliance with the Recommendation is achieved when all of these mandatory provisions are met. The words "shall" or some other obligatory language such as "must" and the negative equivalents are used to express requirements. The use of such words does not suggest that compliance with the Recommendation is required of any party.

INTELLECTUAL PROPERTY RIGHTS

ITU draws attention to the possibility that the practice or implementation of this Recommendation may involve the use of a claimed Intellectual Property Right. ITU takes no position concerning the evidence, validity or applicability of claimed Intellectual Property Rights, whether asserted by ITU members or others outside of the Recommendation development process.

As of the date of approval of this Recommendation, ITU had not received notice of intellectual property, protected by patents/software copyrights, which may be required to implement this Recommendation. However, implementers are cautioned that this may not represent the latest information and are therefore strongly urged to consult the appropriate ITU-T databases available via the ITU-T website at http://www.itu.int/ITU-T/ipr/.

Table of Contents

Recommendation ITU-T H.644.7

Functional architecture for media processing services

1 Scope

This Recommendation specifies the functional architecture and reference points for the media processing services.

2 References

The following ITU-T Recommendations and other references contain provisions which, through reference in this text, constitute provisions of this Recommendation. At the time of publication, the editions indicated were valid. All Recommendations and other references are subject to revision; users of this Recommendation are therefore encouraged to investigate the possibility of applying the most recent edition of the Recommendations and other references listed below. A list of the currently valid ITU-T Recommendations is regularly published. The reference to a document within this Recommendation does not give it, as a stand-alone document, the status of a Recommendation.

[ITU-T F.746.17] Recommendation ITU-T F.746.17 (2022), *Requirements for media processing services*.

3 Definitions

3.1 Terms defined elsewhere

This Recommendation uses the following terms defined elsewhere:

3.1.1 artificial intelligence (AI) [b-ITU-T F.749.13]: An interdisciplinary field, usually regarded as a branch of computer science, dealing with models and systems for the performance of functions generally associated with human intelligence, such as reasoning and learning.

3.1.2 cloud computing [b-ITU-T Y.3500]: Paradigm for enabling network access to a scalable and elastic pool of shareable physical or virtual resources with self-service provisioning and administration on-demand.

3.1.3 content delivery network (CDN) [b-ITU-T F.750]: A network optimized for delivering digital content.

3.1.4 infrastructure as a service (IaaS) [b-ITU-T Y.3500]: **Cloud service category** (see 3.2.10 in [b-ITU-T Y.3500]) in which the **cloud capabilities type** (see 3.2.4 in [b-ITU-T Y.3500]) provided to the cloud service customer (see 3.2.11 in [b-ITU-T Y.3500]) is an **infrastructure capabilities type** (see 3.2.25 in [b-ITU-T Y.3500]).

NOTE – The cloud service customer (see 3.2.11 in [b-ITU-T Y.3500)] does not manage or control the underlying physical and virtual resources, but does have control over operating systems, storage, and deployed applications that use the physical and virtual resources. The cloud service customer (see 3.2.11 in [b-ITU-T Y.3500]) may also have limited ability to control certain networking components (e.g., host firewalls).

3.2 Terms defined in this Recommendation

None.

4 Abbreviations and acronyms

This Recommendation uses the following abbreviations and acronyms:

AI Artificial Intelligence
ASIC Application Specific Integrated Circuit
CDN Content Delivery Network
CPU Central Processing Unit
DRM Digital Rights Management
GPU Graphics Processing Unit
HLS HTTP Live Streaming
HTTP Hyper Text Transfer Protocol
IaaS Infrastructure as a Service
IO Input/Output
MPS Media Processing Service
OSS Object Storage Service
PIP Picture in Picture
RTMP Real Time Messaging Protocol
URL Uniform Resource Locator

5 Conventions

None.

6 Overview

In accordance with the requirements for media processing services defined in [ITU-T F.746.17], the reference model for the media processing service (MPS) is shown in Figure 6-1. In a typical media processing service flow from end to end, video contents are first captured by the content provider. After the video is uploaded to the cloud server, media processing services begin to work. In order to adapt to different network environments and terminals, the video will usually be transcoded in the cloud. Some personalized value-added services are provided including the data statistics service, content review service and real-time watermarking service, and so on. After media processing, the video contents are ingested to the content delivery network (CDN), contents are then distributed and speeded up to the terminal devices. Video contents are finally decoded and played on the terminal devices.

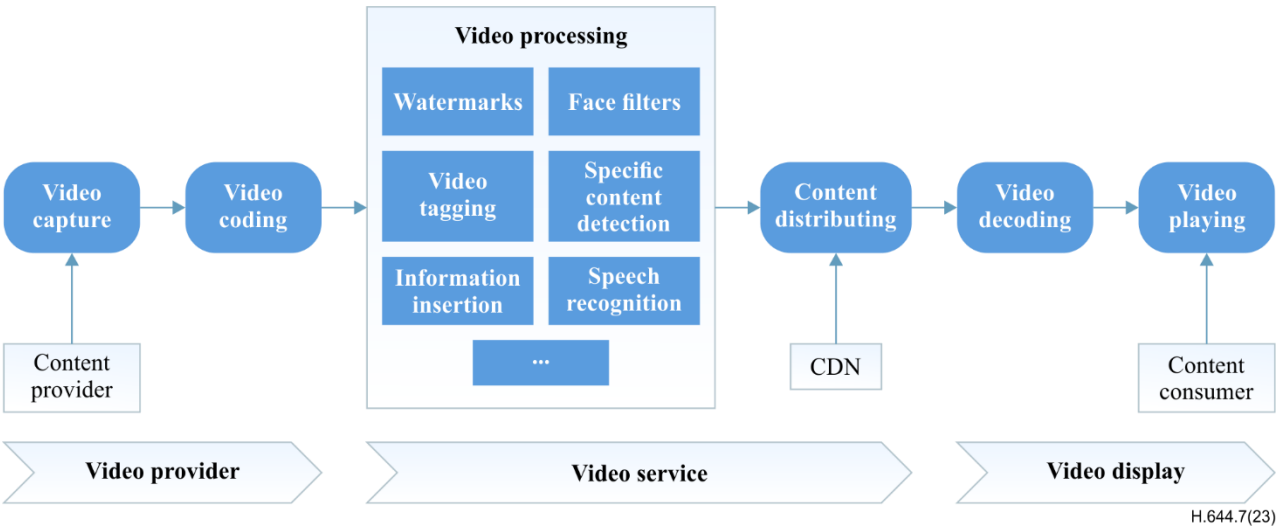

Figure 6-1 – Reference model for the media processing services

7 Domains and functional roles relationship

There are four main domains involved in the provision of MPS. Content is delivered from the content domain to the presentation domain, it is processed within the processing domain and passes to the delivery domain. This Recommendation is focused on the processing domain. In Figure 7-1 the greyed blocks are outside the scope of this Recommendation.

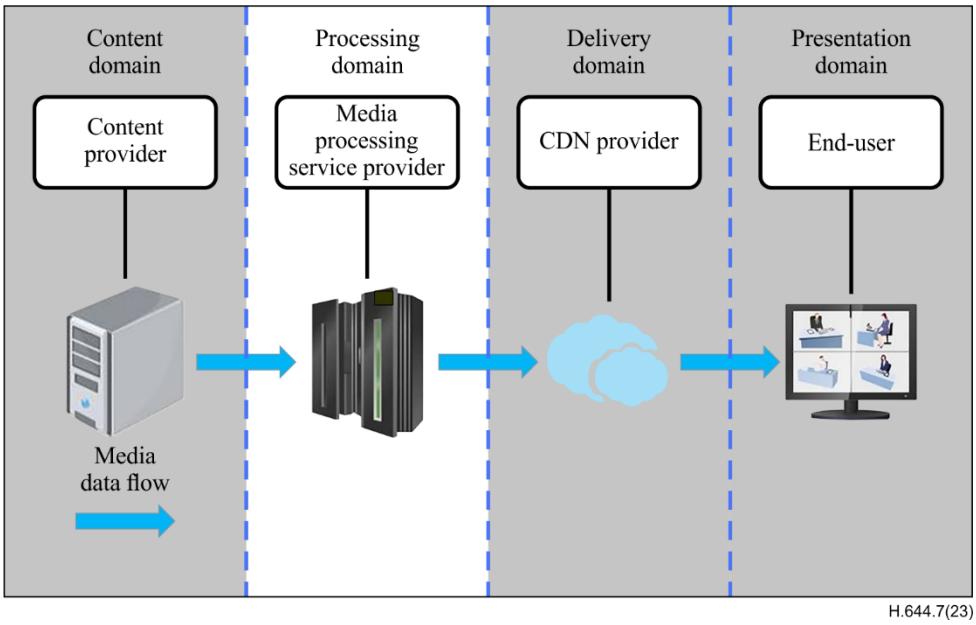

Figure 7-1 – Domains and functional roles relationship

Figure 7-1 shows the domains and the functional roles relationship described below:

Content provider: Provides video sources for MPS, and submits video sources in the form of jobs, which can be offline video (from storage) or streaming media (from CDN or a home network). It can include content management and content producing (such as streaming or files).

MPS provider: MPS is a cloud-end service, which can accept multiple jobs submitted by multiple content providers and perform corresponding job management and processing. It can also provide various media processing services for individual jobs, such as transcoding, snapshot, watermarking and related artificial intelligence (AI) processing, etc.

CDN provider: Provides a delivery network for the results of MPS, which can be used to support live service or on-demand service.

End-user: The actual end-user, who can play or download the corresponding processed video results. The end-user includes home network functions, decoding, and rendering, etc.

8 Functional architecture

This clause addresses the functional architecture for MPS. Figure 8-1 shows the functional architecture for the media processing services. Greyed blocks are outside the scope of this Recommendation.

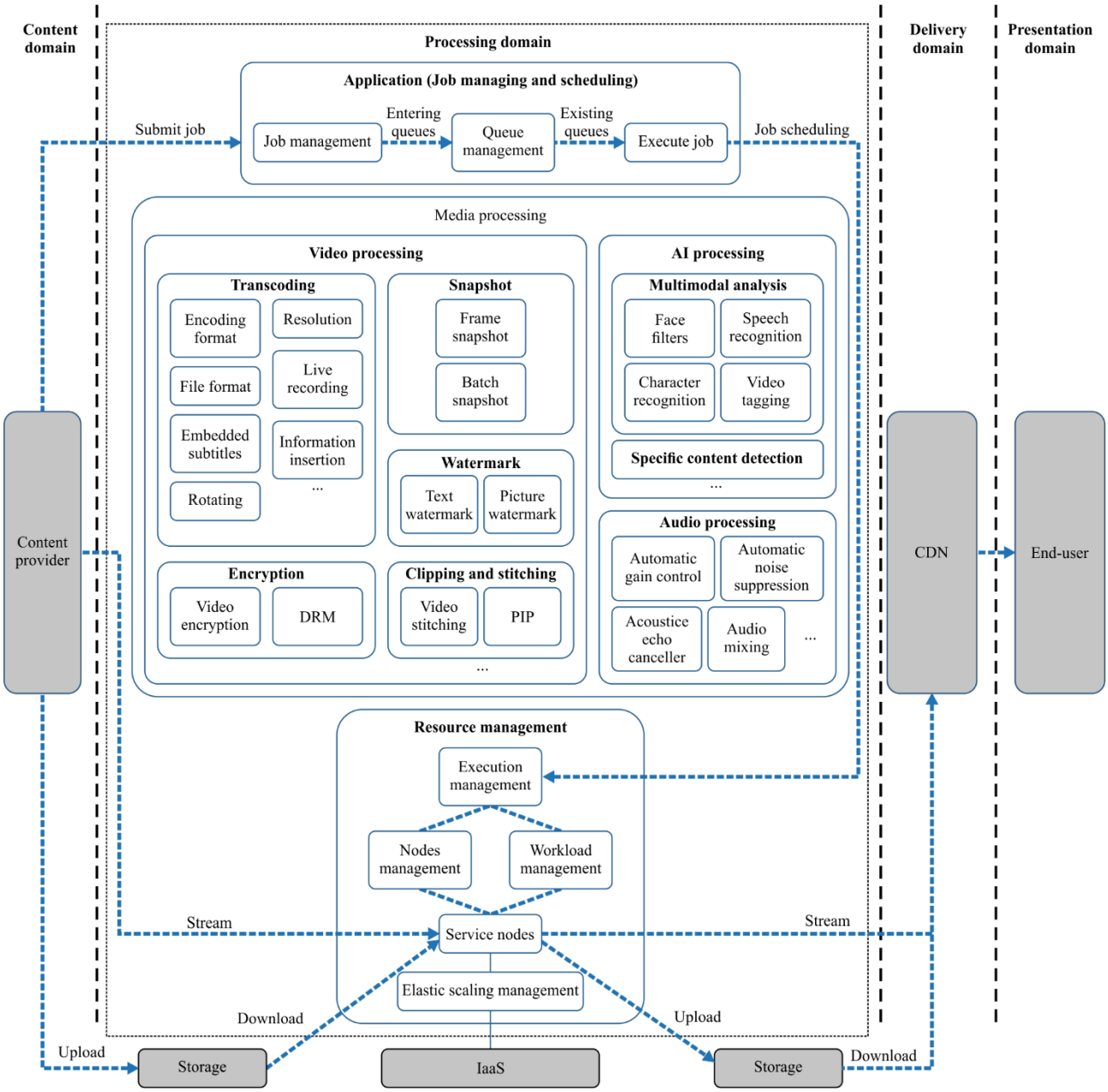

Figure 8-1 – Schematic of functional architecture for the MPS

The media processing domain consists of three main functional blocks: the application (job managing and scheduling) functional block, the resource management functional block, and the media processing functional block. Storage is used to store video files. Infrastructure as a service (IaaS) provides infrastructure such as containers, machines, etc.

The application (job managing and scheduling) functional block consists of three functional entities: job management, queue management, and job scheduling. Job management is responsible for the life cycle of the jobs. Queue management is responsible for the life cycle of queues and the functions of entering queues and exiting queues. Job scheduling is responsible for making decisions on when a job will be executed according to the information provided by the execution management function in the resource management functional block.

The media processing functional block consists of three main sub-functional blocks: video processing, audio processing and AI processing. Within these sub-functional blocks, video processing mainly includes transcoding, snapshot, watermark, encryption, clipping and stitching functional entities. Audio processing mainly includes automatic gain control, automatic noise suppression, acoustic echo canceller, and audio mixing functional entities. AI processing mainly includes multimodal analysis and specific content detection functional entities. The function of the media processing functional block is capable of expanding and upgrading with the development of technology.

The resource management functional block consists of five functional entities: execution management, nodes management, workload management, service nodes and elastic scaling management. Execution management evaluates the overall load of the resource pool through the information provided by the nodes management and workload management functional entities. After receiving the scheduled job, execution management selects the appropriate node for execution according to comprehensive information such as job type. Elastic scaling management is responsible for interacting with the infrastructure as a service (IaaS) layer and adjusting the level of the resource pool according to the highs and lows of service demand and so meet the peak demand of the service and improve the overall utilization rate. Service nodes is a set of service nodes, and it is a logical node deployed in a physical machine to execute the media processing service function.

9 Reference points

Figure 9-1 shows a schematic of the reference points, the greyed blocks are outside the scope of this Recommendation.

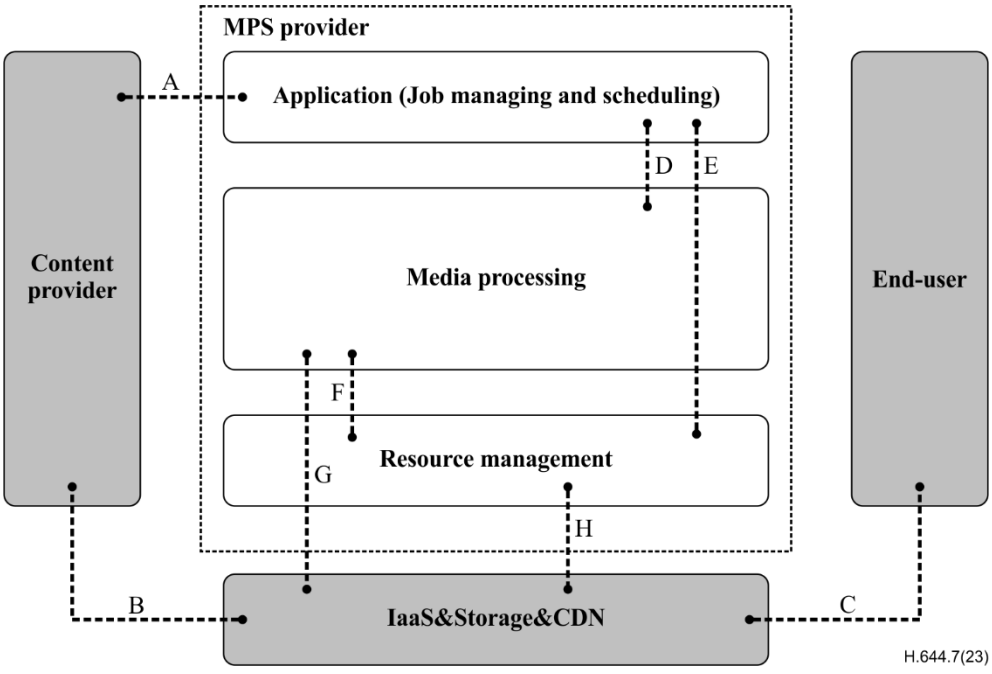

Figure 9-1 – Schematic of reference points

9.1 Reference point A

Reference point A is used to exchange information of the signalling link that indicates to the content provider to call the MPS for the cloud services, mainly concerning the submission of information related to the job in the batch, such as the video uniform resource locator (URL), the serial process of the video processing functional blocks and the input and output parameters of corresponding functional blocks, etc. Reference point A is represented as the signalling transmission link, which mainly includes the following:

– Submit job.

– Delete job.

– Query status of job queues.

– Modify priority of job.

9.2 Reference point B

Reference point B is used to expose the target URL of the source video content storage from the content provider. In the case of offline video, it is stored in the object storage space. In the case of streaming media content, it is ingested into the CDN edge node. Reference point B is represented as the data transmission link, which mainly includes the following:

– Video upload.

– Push/ingest streaming, following real time messaging protocol (RTMP), to CDN nodes or hypertext transfer protocol (HTTP) live streaming (HLS) back to source pulling, etc.

9.3 Reference point C

Reference point C is used for the end-user to get the processed video results through the object storage service (OSS) or to play the processed live streams through the CDN node. Reference point C is represented as the data transmission link, which mainly includes the following:

– Video download from the object storage service.

– Pulling streams from CDN nodes (following RTMP, HLS, etc.)

9.4 Reference point D

Reference point D is used to invoke the corresponding functional blocks for MPS according to the serial process of video processing in the job information. The result of each processing is used as the input of the next functional block for serial scheduling. The job management block manages the entire status of job processing and the status of the job queue. Reference point D is represented as the signalling control and service scheduling link, which mainly includes the following:

– Analysing serial process parameters of MPS.

– Scheduling requests of the MPS-related functional block.

– Requests for stopping some MPS-related functional processing.

– Requests for querying the status of the MPS-related functional block.

9.5 Reference point E

Reference point E is used to invoke the machine resource. In the job management, after analysing the MPS-related serial process, the corresponding virtual machine resource needs to be called and the corresponding MPS needs to be assigned to that virtual machine resource so that the MPS can be well executed. Because different MPS functional blocks require different virtual machine resources, such as graphics processing unit (GPU) cards, or a central processing unit (CPU) with multiple cores (e.g., 128 cores), or special application specific integrated circuit (ASIC) cards, etc., these need to start scheduling and allocation after the job information is parsed. Reference point E is

represented as the signalling control and resource scheduling link, which mainly includes the following:

– Allocation request for cluster resource.

– Release request for cluster resource.

– Status query request for cluster resource pool.

9.6 Reference point F

Reference point F is used to exchange information which indicates the dependency of a specific single type of resource for MPS, and for monitoring the operation status of this type of resource in real time, such as CPU consumption rate, disk utilization, etc. This information can be used for subsequent monitoring and correcting quota of demands concerning resource consumption. Reference point F is represented as the signalling monitoring link, which mainly includes the following:

– Timed requests for basic information about operation and maintenance for a single virtual machine, such as CPU utilization, memory, disk, network input/output (IO), and other real-time status information.

9.7 Reference point G

Reference point G is used to read/write the processed video in the video source from the object storage service (OSS) or CDN node, or read/write the information of the file which is in a middle status under some processing steps. For example, some videos are attached with additional metadata. When these videos are in service processing, these metadata are needed for use as input. Some specific functional blocks need the metadata (which is sourced from results of the previous functional block) as input for new processing, i.e., read and write metadata of the video (such as mask information about character outline, etc.). Reference point G is represented as the data transmission link, which mainly includes the following:

– Reading of video source or metadata.

– Writing of video source or metadata.

9.8 Reference point H

Reference point H is used for the resource management block, which depends on the underlying IaaS, to require the IaaS to support the ability of elastic expansion and contraction capacity of the virtual machine, and also provides related configuration information and service mirror deployment capabilities, etc. Reference point H is represented as the resource signalling control link, which mainly includes the following:

– Request of elastic expansion and contraction capacity of the virtual machine.

– Request of information of virtual machine.

– Request of service mirror deployment.

Bibliography

[b-ITU-T F.749.13] Recommendation ITU-T F.749.13 (2021), *Framework and requirements for civilian unmanned aerial vehicle flight control using artificial intelligence*.

[b-ITU-T F.750] Recommendation ITU-T F.750 (2005), *Metadata framework*.

[b-ITU-T Y.3500] Recommendation ITU-T Y.3500 (2014), *Information technology – Cloud computing – Overview and vocabulary*.

SERIES OF ITU-T RECOMMENDATIONS

Series A	Organization of the work of ITU-T
Series D	Tariff and accounting principles and international telecommunication/ICT economic and policy issues
Series E	Overall network operation, telephone service, service operation and human factors
Series F	Non-telephone telecommunication services
Series G	Transmission systems and media, digital systems and networks
Series H	**Audiovisual and multimedia systems**
Series I	Integrated services digital network
Series J	Cable networks and transmission of television, sound programme and other multimedia signals
Series K	Protection against interference
Series L	Environment and ICTs, climate change, e-waste, energy efficiency; construction, installation and protection of cables and other elements of outside plant
Series M	Telecommunication management, including TMN and network maintenance
Series N	Maintenance: international sound programme and television transmission circuits
Series O	Specifications of measuring equipment
Series P	Telephone transmission quality, telephone installations, local line networks
Series Q	Switching and signalling, and associated measurements and tests
Series R	Telegraph transmission
Series S	Telegraph services terminal equipment
Series T	Terminals for telematic services
Series U	Telegraph switching
Series V	Data communication over the telephone network
Series X	Data networks, open system communications and security
Series Y	Global information infrastructure, Internet protocol aspects, next-generation networks, Internet of Things and smart cities
Series Z	Languages and general software aspects for telecommunication systems

Published in Switzerland
Geneva, 2023