

Recommendation

ITU-T H.644.8 (06/2024)

SERIES H: Audiovisual and multimedia systems

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

Requirements and architecture for audio and video processing 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.644.8

Requirements and architecture for audio and video processing for media processing services

Summary

Given the requirements of recently developed video services, media contents are required to be processed before transmission through a content delivery network (CDN). For example, media contents are required to be transcoded adaptively according to the quality of the network. Some user generated contents (UGCs) are required to not include specific contents, while other image and text contents are required to be detected, located and identified in order to provide personalized services for consumers.

The media processing service (MPS) is based on the concept of cloud computing technology and provides a cost-effective, easy-to-use, elastic, on-demand and highly scalable method to process audio and video in various video-related services. Recommendation ITU-T H.644.8 identifies requirements and architecture for audio and video processing via a management platform for the media processing service.

History*

Edition	Recommendation	Approval	Study Group	Unique ID
1.0	ITU-T H.644.8	2024-06-13	16	11.1002/1000/15932

Keywords

Audio processing, management platform, media processing service, 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/>.

© ITU 2024

All rights reserved. No part of this publication may be reproduced, by any means whatsoever, without the prior written permission of ITU.

Table of Contents

	Page
1 Scope	1
2 References.....	1
3 Definitions	1
3.1 Terms defined elsewhere	1
3.2 Terms defined in this Recommendation.....	2
4 Abbreviations and acronyms	2
5 Conventions	3
6 Introduction	3
7 Management platform.....	4
7.1 Architecture	4
7.2 Core functions of platform	6
7.3 Platform operations and maintenance (O&M) functions	11
7.4 Platform generic capabilities	12
8 Audio processing subfunctional block	15
8.1 Basic requirements for service	15
8.2 Requirements for infrastructure-oriented adaptation capabilities	15
8.3 Key capabilities for audio processing.....	16
8.4 Key performance metrics	17
8.5 Requirements for supporting service operation capability	17
8.6 Requirements for supporting O&M capabilities	18
8.7 Requirements for supporting security assurance capability	18
8.8 Requirements for supporting completeness capability	19
9 Video processing subfunctional block.....	20
9.1 Basic requirements for service	20
9.2 Requirements for infrastructure-oriented adaptation capabilities	21
9.3 Key capabilities for video processing.....	22
9.4 Key performance metrics	24
9.5 Requirements for supporting service operation capability	24
9.6 Requirements for supporting O&M capabilities	25
9.7 Requirements for supporting security assurance capability	25
9.8 Requirements for supporting completeness capability	27
Bibliography.....	28

Recommendation ITU-T H.644.8

Requirements and architecture for audio and video processing for media processing services

1 Scope

This Recommendation specifies the requirements and architecture for audio and video processing via a management platform for the media processing service (MPS) and includes:

- Requirements and architecture for audio processing.
- Requirements and architecture for video processing.
- Requirements and architecture for the management platform.
- Other requirements and architecture for 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*.

[ITU-T H.644.7] Recommendation ITU-T H.644.7 (2023), *Functional architecture for media processing services*.

3 Definitions

3.1 Terms defined elsewhere

This Recommendation uses the following terms defined elsewhere:

3.1.1 media processing service (MPS) [ITU-T F.746.17]: A cloud service for transcoding multimedia data, which can provide a cost-effective, easy-to-use, elastic, and highly scalable method to convert audio and video into formats suitable for playing on PCs, TVs and mobile interfaces, and can also conduct deep learning based technology on massive data to analyse the content including text, voice and other scenarios of the media in multiple modes, and to achieve intelligent review, content understanding, intelligent editing, etc.

3.1.2 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.

3.1.3 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.

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

3.1.5 platform as a service (PaaS) [b-ITU-T Y.3500]: Cloud service category (see clause 3.2.10 in [b-ITU-T Y.3500]) in which the cloud capabilities type (see clause 3.2.4 in [b-ITU-T Y.3500]) provided to the cloud service customer (see clause 3.2.11 in [b-ITU-T Y.3500]) is a platform capabilities type (see clause 3.2.31 in [b-ITU-T Y.3500]).

3.1.6 infrastructure as a service (IaaS) [b-ITU-T Y.3500]: Cloud service category (see clause 3.2.10 in [b-ITU-T Y.3500]) in which the cloud capabilities type (see clause 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 clause 3.2.25 in [b-ITU-T Y.3500]).

3.2 Terms defined in this Recommendation

None.

4 Abbreviations and acronyms

This Recommendation uses the following abbreviations and acronyms:

2B	To Business
2C	To Consumer
ABR	Available Bitrate
AI	Artificial Intelligence
ASIC	Application Specific Integrated Circuit
CBR	Constant Bitrate
CDN	Content Delivery Network
CPU	Central Processing Unit
CU	Compute Unit
CU1	Unit of Reference CU
CUDA	Compute Unified Device Architecture
DDoS	Distributed Denial of Service
DRM	Digital Rights Management
DTS	Decoding Time Stamp
FPGA	Field Programmable Gate Array
FRC	Frame Rate Conversion
GPU	Graphics Processing Unit
HDR	High Dynamic Range
HTTP	Hyper Text Transfer Protocol
HTTPS	Hypertext Transfer Protocol Secure
IaaS	Infrastructure as a Service
ID	Identifier
IT	Information Technology
I/O	Input/Output
MPS	Media Processing Service
Mutex	Mutual exclusion

NAS	Network Attached Storage
O&M	Operation and Maintenance
OSS	Object Storage Service
PaaS	Platform as a Service
PC	Personal Computer
PIP	Picture in Picture
POSIX	Portable Operating System Interface
PTS	Presentation Time Stamp
QoS	Quality of Service
RAM	Random Access Memory
RTMP	Real Time Messaging Protocol
SaaS	Software as a Service
SDR	Standard Dynamic Range
SLA	Service-Level Agreement
SMS	Short Message Service
TV	Television
UGC	User Generated Content
VBR	Variable Bitrate
VM	Virtual Machine

5 Conventions

In this Recommendation:

- The keywords "functional block" indicate the top granularity used in function division in schematic of functional architecture for specific platform/system. It generally includes a further division.
- The keywords "subfunctional block" indicate the medium granularity used in function division in schematic of functional architecture for specific platform/system. It is included in a functional block. It can include more-subdivided subfunctional block(s) or functional entities.
- The keywords "functional entity" indicate the least granularity used in function division in schematic of functional architecture for specific platform/system. It is included in a functional block and maybe included in subfunctional block(s).

6 Introduction

Media processing service (MPS) supports regular media processing functions as well as the ability to convert video into structured data using the latest artificial intelligence (AI) technologies. It can meet the media processing needs of media files as well as video streams as input. The media processing capabilities are scalable and can be expanded and upgraded to follow technological developments.

7 Management platform

7.1 Architecture

Figure 7-1 (based on Figure 8-1 in [ITU-T H.644.7]) shows the functional architecture for the MPS. The blocks with grey background (i.e., blocks named as content provider, CDN, end-user, storage, and IaaS) are out of the research scope of this Recommendation.

NOTE – Based on Figure 8-1 in [ITU-T H.644.7], the following Figure 7-1 comprises a correction to the original figure with the positions of "Execute job" and "Job scheduling" switched as per a corrigendum.

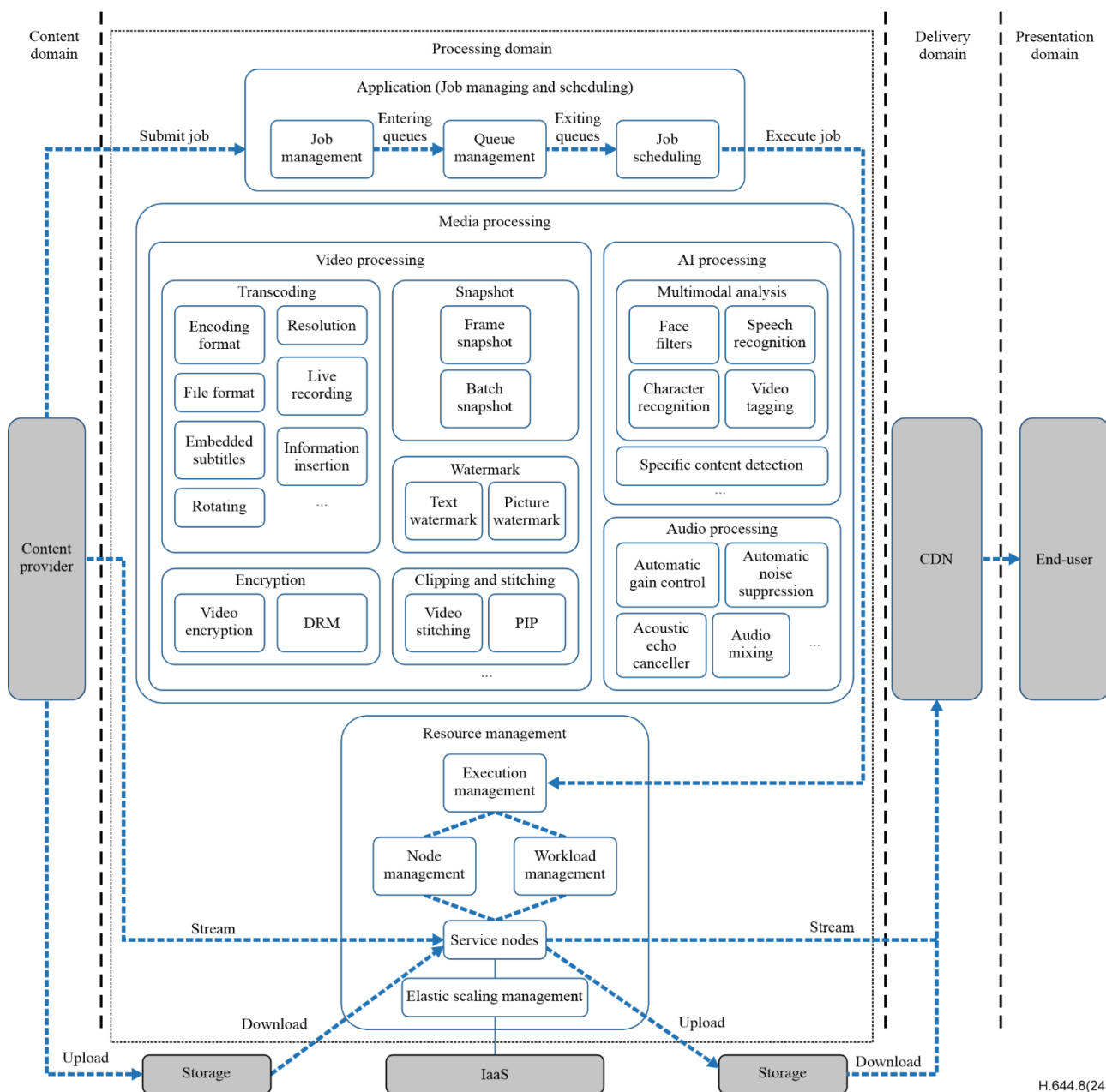


Figure 7-1 – Schematic of functional architecture for the MPS

The media processing domain is the core link of the MPS and belongs to the cloud-centric deployment. It 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 media 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 subfunctional blocks: video processing, audio processing and AI processing. 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 nodes management and workload management. 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 water level of the resource pool according to the peak and valley of the service, to meet the peak demand of the service and improve the overall utilization rate. Service nodes is a set of service nodes and is a logical node deployed in a physical machine to execute MPS.

The management platform is related to large-scale computing for MPS (belonging to the platform as a service (PaaS) layer), which should be able to provide large-scale computing capabilities by utilizing the IaaS layer as well as shielding the IaaS layer from the variability of different implementations. For the upper software as a service (SaaS) layer it is important to provide higher levels of domain abstraction, scalable processing and integrated capabilities, and the ability to support multi-tenant, elasticity, etc.

The management platform refers to the orchestration and configuration management of the functional processes of user' jobs, as well as the control of the status, progress, and resources consumed during the processing of services within functional modules. It is mainly oriented to business (2B) users. It is used to help business users effectively configure, monitor, control, and manage the service processes, job tasks, system resources, and related IT infrastructure for each functional module of the MPS.

The management platform, supported by the infrastructure, mainly includes core function modules such as job management, scheduling management (belonging to the application functional block), resource management, and cluster management.

Cluster management sits between the IaaS and resource management functional block in the overall architecture of Figure 7-1 (i.e., Figure 8-1 in [ITU-T H.644.7]), which is used to manage multiple IaaS instances (multi-vendor or multi-region). The system architecture is schematically shown in Figure 7-2.