DASH and MMT and Their Applications in ATSC 3.0

Yiling Xu[1], Shaowei Xie[1], Hao Chen[1], Le Yang[2], and Jun Sun[1]
(1. Shanghai Jiaotong University, Shanghai 200000, China;
2. Jiangnan University, Wuxi 214000, China)

Abstract

Despite the success of MPEG‑2 Transport Stream (TS) being used to deliver services in broadcast channels, the increase of on‑demand viewing of multimedia content over IP with browser‑centric media endpoints introduces a new requirement for more individualized and flexible access to content. This has resulted in alternatives to MPEG‑2 TS. While the needs of interactive broadcast services (such as personalized advertisement or selection of audio stream with a language suitable for a specific user) grow there is an active standardization work under going for the next generation broadcasting systems. To best enable a complete system of hybrid broadcast and broadband services, Advanced Television Systems Committee (ATSC) 3.0 has developed an enhanced broadcast transport method named Real‑Time Object Delivery over Unidirectional Transport (ROUTE)/DASH for delivery of DASH‑formatted content and non‑real time (NRT) data. Additionally, for broadcasting, ATSC 3.0 has also adopted MPEG Media Transport (MMT) standard, which inherits major advantageous features of MPEG‑2 TS and is very useful in real‑time streaming delivery via a unidirectional delivery network.This paper mainly describes features and design considerations of ATSC 3.0, and discusses the applications of the transport protocols used for broadcasting, i.e., ROUTE/DASH and MMT, whose comparative introductions are also presented in details.

Keywords

ATSC 3.0; ROUTE/DASH; MMT; Hybrid delivery; next‑generation broadcasting system

1 Introduction

The rapid increase in the volume of multimedia content, together with the needs of flexible and interactive multimedia broadcast services such as personalized advertisement, has changed significantly the multimedia service environments. To address the associated emerging challenges [1] for multimedia delivery, standardization has been taken for developing next generation broadcasting systems that are independent content delivery systems while being able to exploit broadband networks more frequently. The resultant Advanced Television Systems Committee (ATSC) 3.0 standard incorporates a number of innovative features not typical in a conventional broadcasting system, and it is the very first full IP‑based broadcasting standard. ATSC 3.0 assumes hybrid systems [2] with the media delivery channel consisting of broadcast and broadband networks.

Prior to ATSC 3.0, the most successful standards for multimedia content delivery are MPEG‑2 [3], proposed by MPEG and the associated ISO Base Media File Format (ISO BMFF) [4]. Specifically, the MPEG‑2 Transport Stream (TS) [5] provides very efficient methods of multiplexing multiple audiovi‑ sual data streams into one delivery stream according to consumption order, (**Fig. 1**). This makes MPEG‑2 TS an ideal solution for broadcasting multimedia contents in cases where a large number of users request the same content. However, MPEG‑2 TS may be inadequate to fulfill the emerging needs for more individualized and flexible access to the content brought by e.g., personalized on‑demand viewing of multimedia contents. This is also the case for the ISO BMFF, which stores metadata containing information for synchronized playback separately from the compressed media data. In particular, it is difficult to access a certain portion of the ISO BMFF content, e.g., to locate and retrieve an audio stream with a specific
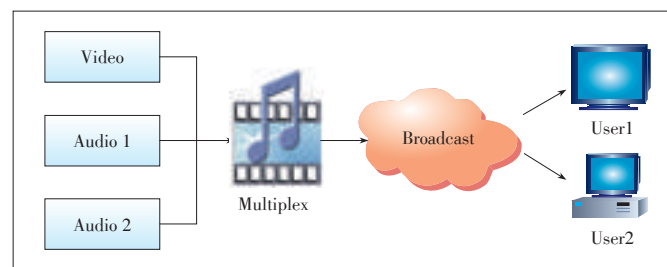
▲Figure 1. Multiplex of multiple audiovisual streams in MPEG-2 TS.

language during playback.

As a solution to immediate market demands, HTTP adaptive streaming [6], [7] is currently being used to deliver all broadband streaming IP content. Meanwhile, recognizing the drawbacks of existing standards and systems, MPEG developed the Dynamic Adaptive Streaming over HTTP (DASH) [8] standard by converging different HTTP adaptive streaming technologies with a focus on the adaptive streaming of media content over the legacy HTTP delivery environment [9]. Although HTTP streaming over TCP is suitable for broadband/unicast delivery, it is not an appropriate end‑to‑end delivery mechanism for broadcasting. To better support hybrid broadcast and broadband services, an enhanced broadcast transport method, referred to as Real‑Time Object Delivery over Unidirectional Transport (ROUTE) [10]/DASH, has been proposed for delivering DASH‑format content and non‑real time (NRT) data over broadcast channels.

MPEG-DASH has been successful in commercial online video markets [11], [12], and MPEG turned its attention to address the new challenges of multimedia content delivery in Internet delivery environments beyond HTTP. The MPEG Media Transport (MMT) standard [13] has been recently proposed as part of the ISO/IEC 23008 High Efficiency Coding and Media Delivery in Heterogeneous Environments (MPEG‑H) standard suite [14], [15]. It assumes IP networks with in-network intelligent caches close to the receiving entities. The caches actively pre‑fetch the content and also adaptively packetize and push cached content to receiving entities. MMT also assumes a network environment where content can be accessed at a finer grain with unique identifiers regardless of the specific service providers and their locations [16].

Recognizing the benefits and potential for flexible and interactive multimedia services in hybrid networks, ATSC 3.0 uses both ROUTE/DASH and MMT standards. In the remainder of this paper, we give a comparative review of DASH and MMT systems with introduction of their applications in ATSC 3.0. The discussion starts in Section 2 with the design considerations of ATSC 3.0 for the next generation broadcast and the employed signaling structure. The ROUTE/DASH and MMTP/MPU are presented in the Section 3. In Section 4, content delivery in broadcast is described in detail, including the media encapsulation, the delivery of streaming services and the NRT content, as well as the system models for content delivery. Section 5 introduces the hybrid delivery modes with ROUTE/DASH and MMTP/MPU used in broadcast. Section 6 concludes the paper.

2 ATSC 3.0 Overview

2.1 Receiver Protocol Stack

To address the need for a unified receiver protocol stack that can handle the diverse potential service types and delivery

methods, ATSC 3.0 adopts the receiver protocol stack depicted in **Fig. 2**. This structure has the advantages of allowing clean interface among the various layers and enabling the required functionalities via various delivery methods.

In Fig. 2, a typical ATSC 3.0 system contains three functional layers, namely, the physical layer, delivery layer and Application layer. More specifically, the physical layer comprises the broadcast and broadband physical layers. The delivery layer realizes data streaming and object flow transport functionality. The application layer enables various type of services, such as the digital television or HTML5 [17] applications.

ATSC 3.0 incorporates MMT and ROUTE in its Delivery Layer, both of which are carried out via a UDP/IP multicast over the broadcast physical layer. In particular, MMT utilizes the MMT Protocol (MMTP) [13] to deliver media processing units (MPUs), while ROUTE is based on MPEG DASH and Layered Coding Transport (LCT) [18] to deliver DASH segments. The delivery layer of ATSC 3.0 is also equipped with the HTTP protocol [19] with a TCP/IP unicast over the Broadband Physical layer. Non-timed content including the NRT media, electronic program guide (EPG) data, and other files can be delivered with ROUTE or directly over UDP. Signaling may be delivered over MMTP and/or ROUTE, while bootstrap signaling information is the means of the service list table (SLT). To support hybrid service delivery, in which one or more program elements are delivered via the broadband path, MPEG DASH over HTTP/TCP/IP is used on the broadband physical layer. Media files in ISO BMFF are used as the delivery, media encapsulation and synchronization format for both broadcast and broadband deliveries.

2.2 Functionality of ATSC 3.0 System

With ATSC 3.0 being a unified IP centric delivery system, content providers have the flexibility of exploring broadcast delivery, broadband delivery, or both to enhance efficiency and potential revenue. For instance, for broadcast services, the us-
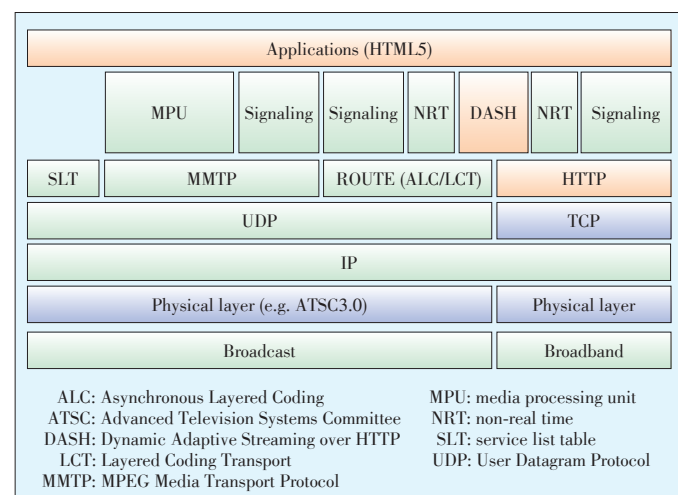
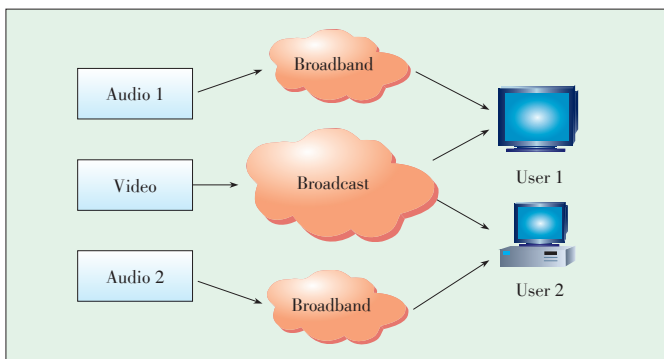▲Figure 2. ATSC 3.0 receiver protocol stack.

er experience can be improved via a browser-based user interface that enables interactive applications to run in the same playback environment as the streaming media. There are many opportunities for increased revenue from the placement of personalized advertisements. For broadband services, key benefits may be the opportunity to provide narrowly focused content more efficiently than broadcasting. This in turn enables more precise user targeting. There are also opportunities to improve channel capacity via e.g, always utilizing the more efficient delivery method. Therefore, this makes ATSC 3.0 an ideal option for hybrid services [20], which focus largely on either high-penetration contents that are best served by broadcast delivery or contents with narrower interest to be best carried over broadband delivery. ATSC 3.0 can aggregate the content of hybrid services from a variety of sources and deliver it through dynamically exploiting broadcast and broadband distribution channels (**Fig. 3**).

In addition, with MMT and DASH/ROUTE, services within ATSC 3.0 can be distributed as scalable streams consisting of a base layer and a number of enhancement layers [21]. The broadcaster has different means to transmit these multilayer content. For examples, the base layer content may be delivered via broadcasting, while the enhancement information is transmitted over the broadband network. Alternatively, delivering all layers content only via the broadcast network is also possible.

3 Signaling for ROUTE/DASH and MMTP/ MPU Service Delivery

3.1 Service List Table and Service Layer Signaling

Signaling information is carried in the payload of IP packets [22] with a well-known address/port and it is commonly referred to as Low Level Signaling (LLS). In ATSC 3.0, service list Ttble (SLT) is specified as the LLS and its functionality is similar to that of the program association table (PAT) of MPEG-2. To be more specific, SLT supports a rapid channel scan that enables a receiver first encountering the broadcast emission to build a list of all the received ATSC 3.0 services. SLT also pro-

vides bootstrap information that allows a receiver to locate the Service Layer Signaling (SLS). The SLS of each ATSC 3.0 service describes the service characteristics including its content components, where to acquire them as well as the device capabilities needed to produce a meaningful presentation of the service. The bootstrap information provided by the SLT contains the destination IP address and the destination port of the Layered Coding Transport (LCT) session or the MMTP session that carries the SLS for services delivered via ROUTE/DASH or MMTP/MPU.

Fig. 4 summarizes the relationship among the SLT, SLS and the ATSC 3.0 services. It can be seen that for broadcast delivery of ROUTE/DASH services (streaming and NRT), the SLS is carried by ROUTE/UDP/IP in one of the LCT transport sessions, while for the MMT/MPU streaming services, the SLS is carried by MMTP Signaling Messages. SLS is delivered at a suitable carousel rate to support fast channel join and switching. Under broadband delivery, the SLS is transmitted over HTTP(S)/TCP/IP.

3.2 Service Identification in ROUTE and MMTP Sessions

Each ROUTE session has one or more LCT sessions which carry, as a whole or in part, the content components that make up the ATSC 3.0 service. In streaming service delivery, an LCT session may carry an individual component of a user service, such as an audio, video or closed caption stream. Streaming media is partitioned per MPEG DASH into DASH segments. Each MMTP session contains one or more MMTP packet flows which carry MMT signaling messages for the content components. An MMTP packet flow may carry MMT signaling messages or components formatted per MMT as MPUs.

A ROUTE session is identified via the source IP address, destination IP address and destination port number. An LCT session associated with the service component(s) it carries is identified via the transport session identifier (TSI), which is unique within the scope of the parent ROUTE session. Properties common to the LCT sessions and properties unique to individual LCT sessions are given in a ROUTE signaling structure

▲Figure 3. Hybrid services with different delivery channels.
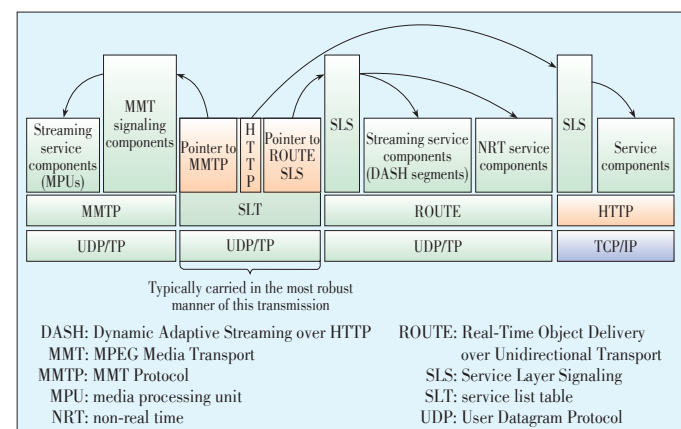
▲Figure 4. SLT references to services via SLS.

DASH and MMT and Their Applications in ATSC 3.0
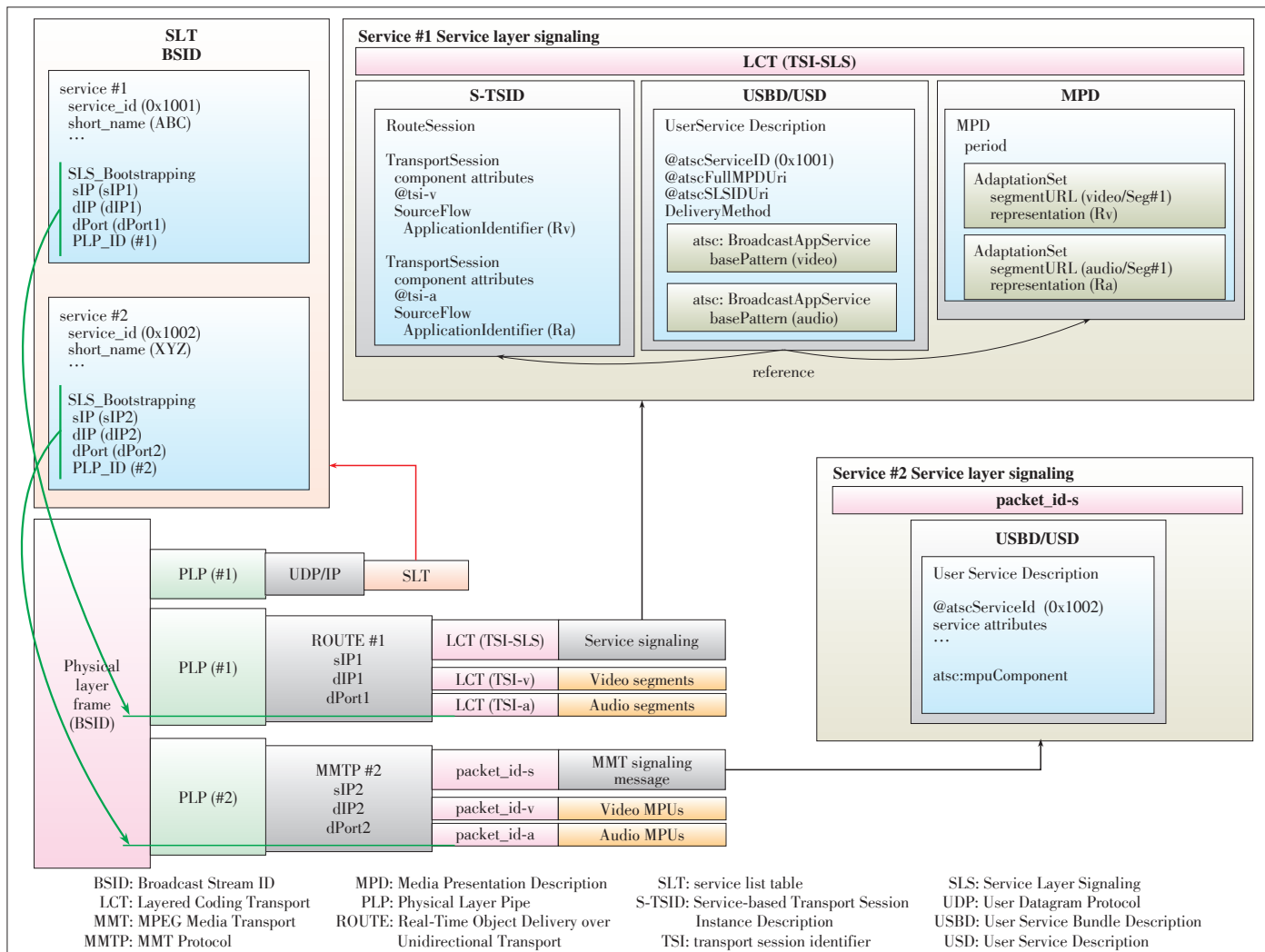Yiling Xu, Shaowei Xie, Hao Chen, Le Yang, and Jun Sun

called Service-based Transport Session Instance Description (S-TSID), which is part of the SLS. Each LCT session is carried over a single physical layer pipe. Different LCT sessions of a ROUTE session may or may not be contained in different physical layer pipes. The properties described in the S-TSID include TSI value and PLP_ID for each LCT session, descriptors for the delivery objects/files, and Application Layer Forward Error Correction (AL-FEC) parameters [23].

The MMTP session and packet flow are identified in a similar manner as the ROUTE session. In particular, the destination IP address/port number and the packet_id unique within the scope of the parent MMTP session are used. Properties common to each MMTP packet flow and properties of individual MMTP packet flows are included in the SLT. Properties for each MMTP session are given by MMT signaling messages, which may be carried within the MMTP session. These properties include packet_id value and PLP_ID for each MMTP packet flow.

Fig. 5 gives an example on the use of SLT to bootstrap the

SLS acquisition and also the use of the SLS to acquire service components delivered on either ROUTE sessions or MMTP sessions. In this example, ATSC 3.0 receiver starts acquiring the SLT and each service identified by service_id provides the following SLS bootstrapping information: PLP_ID, source IP address (sIP), destination IP address (dIP) and destination port number (dPort).

For service delivery in ROUTE, the SLS consists of the following metadata fragments: User Service Bundle Description (USBD), S-TSID and DASH Media Presentation Description (MPD). These are all pertinent to a certain service. The USBD/USD fragment contains service identification, device capability information, Uniform Resource Identifier (URI) [24] references to the S-TSID fragment and the MPD fragment, and metadata to enable the receiver to determine whether the transport mode is broadcast and/or broadband for service components. The S-TSID fragment provides component acquisition information associated with a service and mapping between DASH representations found in the MPD and the TSI corresponding to the

▲Figure 5. Exemplary use of service signaling for bootstrapping and service discovery.

BSID: Broadcast Stream ID
LCT: Layered Coding Transport
MMT: MPEG Media Transport
MMTP: MMT Protocol
MPD: Media Presentation Description
PLP: Physical Layer Pipe
ROUTE: Real-Time Object Delivery over Unidirectional Transport
SLT: service list table
S-TSID: Service-based Transport Session Instance Description
TSI: transport session identifier
SLS: Service Layer Signaling
UDP: User Datagram Protocol
USBD: User Service Bundle Description
USD: User Service Description

component of the service.

For service delivery using MMTP, the SLS contains only USBD fragments. These mainly reference the MMT signaling's MMT package table (MPT) message that provides identification of package ID and location information for assets belonging to the service. MMT signaling messages are delivered by MMTP according to the signaling message mode specified in [13].

To attain compatibility with MPEG DASH components, MMTP SLS needs to include USBD fragments used in ROUTE/DASH and even MPD fragments. This would lead to signaling redundancy. It is still unclear whether MPD should be carried as an SLS metadata fragment in ROUTE/DASH or in the payload of media presentation information (MPI) message, which is an instance type of MMT signaling messages. On the other hand, due to the USBD fragment being reused, USBD for MMT can give complete service bundle description for both media component formats, i.e., the DASH segment and MPU format. This means that MMTP/MPU can support the delivery of both media formats. As a result, MMTP/MPU is more inclusive and robust to diverse media component formats. In addition, MMTP/MPU defines more detailed USBD information about the service, including attributes such as @atsc:providerId, @atsc:ServiceCategory and @atsc:serviceStatus, and elements such as atsc:channel and atsc:componentInfo. At the cost of increased redundancy, MMTP/MPU signaling can achieve richer and more comprehensive media service.

4 Content Delivery

4.1 Overview

MPEG DASH is designed for communication between HTTP servers and DASH client(s). It specifies formats and methods for delivering streaming service(s), and describing a collection of media segments and auxiliary metadata through an MPD. The ROUTE protocol provides general broadcast delivery capabilities similar to the broadband delivery capabilities provided by HTTP. In addition to supporting the file-based streaming techniques used in DASH, ROUTE provides media-aware content delivery, which enables faster channel switch. On the whole, ROUTE/DASH is a broadcast-oriented, media-aware byte range delivery protocol based on MPEG DASH and LCT over UDP, with the use of AL-FEC. ROUTE/DASH can also operate with any media codec, including scalable media codecs provided that the appropriate codec specific file format is specified.

MMTP is a protocol designed to deliver ISOBMFF files and it is very useful in real-time streaming delivery via an unidirectional delivery network. MMTP has several distinct features such as:
- media-aware packetization of ISOBMFF files
- multiplexing of various media components into a single MMTP session
- network jitter removal at the receiver under the constraints set by the sender
- receiver buffer management by the server to avoid any buffer underflow or overflow and the fragmentation/aggregation of payload data
- detection of missing packets during delivery.

It is noteworthy that in ATSC 3.0, the use of AL-FEC in ROUTE is crucial for e.g. the large NRT file delivery, DVR applications and enhanced robustness for collecting small objects, while in MMTP, it is optional.

4.2 Media Encapsulation

In broadcast service, since ROUTE/LCT doesn't differentiate DASH Segments by type, it may introduce more service delay considering the dependency among the segments. On the contrary, MPU is self-contained, i.e., the initialization information and metadata required to fully decode the media data in each MPU is included in the MPU. In MMTP sessions, the media fragment type of the payload is known, leading to easy construction of the media data. These two properties of MMTP/MPU improve system performance in terms of resistance and robustness. For instance, with the knowledge on the fragment type in MMTP session, adaptive AL-FEC protection could be used on the basis of the significance of media fragments to achieve better user experience. In terms of random access and channel switch, MMTP/MPU acquires the service in MPU level faster due to self-contained attributes. By comparison, ROUTE/DASH may force users to wait until the next metadata segment arrives.

In addition, with the concept of Media Delivery Event (MDE), ROUTE/DASH users need more time to restore a service when abrupt data loss or error occurs in MDE starting with a Random Access Point (RAP), until the next MDE starts with a RAP. Another important issue is that live advertisement insertion and removal are easier for MMTP/MPU because there is no structural difference between an ad MPU and a program MPU.

Finally, for MMTP/MPU, each MPU contains a globally unique ID for media components and a sequence number to enable unique identification of each MPU regardless of the delivery mechanism. As for DASH, media segments and auxiliary metadata are referenced by Uniform Resource Locator (URL) through MPD, while in a ROUTE session, S-TSID provides the mapping between DASH representations found in the MPD and the TSI corresponding to the component of the service. The identification scheme used in MPU is more accurate and scalable, which benefits media allocation and recognition in media service.

4.3 Streaming Services

Flexible packaging and diverse delivery modes supported both in ROUTE/DASH and MMTP/MPU enable multiple types

of service components to be packaged and delivered in a way best for them. This feature enhances agility and effectiveness of delivery method and it also dramatically improves the QoS for the system and QoE of users.

4.3.1 Delivery in ROUTE/DASH

In the streaming service of ROUTE/DASH, no matter it is live content or pre-recorded content, the attribute 'type' of the MPD (MPD@type) should be set to "dynamic". As for the attribute 'minimumUpdatePeriod' (MPD@ minimumUpdatePeriod), when it is present, the receiver should get MPD updates carried in the LCT session. The objects delivered by LCT session of the ROUTE protocol shall be formatted according to the announcement in the MPD. The MPD and the described Media Presentation should conform to the ISO Base media file format as specified in [8].

In streaming services delivery using ROUTE/DASH, three different kinds of delivery mode are proposed, namely, Entity Mode, File Mode and Packaging Mode. The Entity Mode is used when it is not possible to determine the Extended File Delivery Table (EFDT) parameters in ROUTE prior to the object delivery. In this case, the EFDT parameters (as entity-headers) are sent in-band with the delivered object (as the entity-body) in the form of a compound object. The file/object metadata is carried by one or more entity-header fields associated with the entity-body. Meanwhile, Entity Mode enables partial or chunked delivery in the same way as HTTP, and it provides the means to reduce the sender delay and possibly the end-to-end delay as well. In the File Mode, the file/object metadata as represented by the EFDT would either be embedded within or be referenced as a separate delivery object. The file may also be sent in a progressive manner using the regular ROUTE sending operation in order to reduce sender delay. The Packaging Mode should be used as the delivery object format if the repair flow is used in conjunction with the source flow for streaming content delivery. It enables more robust AL-FEC recovery by applying FEC protection across a collection of delivered objects for enhanced time diversity and constant QoS.

4.3.2 Delivery in MMTP/MPU

Each content component is considered an MMT asset under MMTP/MPU. Each MMT asset is a collection of one or more MPUs with the same unique Asset ID. An MMT package is a collection of one or more assets, and an ATSC 3.0 Service can have one or more MMT packages. Both MMT packages and MPUs do not overlap in their presentation time.

Multiple assets can be delivered over a single MMTP session. Each asset is associated with a packet_id which is unique within the scope of the MMTP session. This enables efficient filtering of MMTP packets carrying a specific asset. Additionally, MMT signaling messages delivered to the receiver are used to designate the mapping information between MMT packages and MMT sessions.

Fig. 6 shows examplary mapping between a MMT package and an MMTP session. The MMT package has three assets: asset A, asset B and asset C and is delivered over two MMTP sessions. Although all the MMT signaling messages required to consume and deliver the MMT package should be delivered to the receiver, only a single MPT message is shown for simplicity. Alternatively, all the MMT assets of the MMT package and its associated MPT message can be multiplexed into a single MMTP session, as in the configuration of MPEG-2 TS, which together with the inclusion of the packet_id field valid range and Clock Relation Information (CRI) message makes MMTP/MPU backward-compatible with MPEG-2 TS.

In addition to a single MMT Package being able to be delivered over one or more MMTP sessions, multiple Packages may be delivered by a single MMTP session and multiple packages in the same service can be delivered simultaneously. MMTP provides the media-aware packetization of ISOBMFF files for real-time streaming delivery via an unidirectional delivery network.

4.4 NRT Content

In ATSC 3.0, the media content includes streaming service and NRT content. Streaming service can be delivered by either ROUTE/DASH or MMTP/MMT whereas NRT content can only be delivered over ROUTE/DASH because Generic File Delivery (GFD) mode specified in [6] cannot be used in ATSC 3.0 system.

File content comprising discrete media are considered to be NRT content. When delivering this kind of file content, the File Mode of ROUTE/DASH should be used. Before the delivery of the file, the file metadata (EFDT) should be informed to the receiver. In the delivery of the subsequent source flow, the delivered file references the LCT session, and the file Uniform Resource Identifier (URI). The file URI is used to identify the delivered EFDT. In this case, the receiver can use the EFDT to process the content file.

Besides common NRT content, service metadata belonging to an NRT content item of an ATSC 3.0 service can also be considered as NRT content (e.g., the SLS or Electronic Service Guide (ESG) fragments) from the application transport and IP delivery perspective. Their delivery also conforms to the princi-
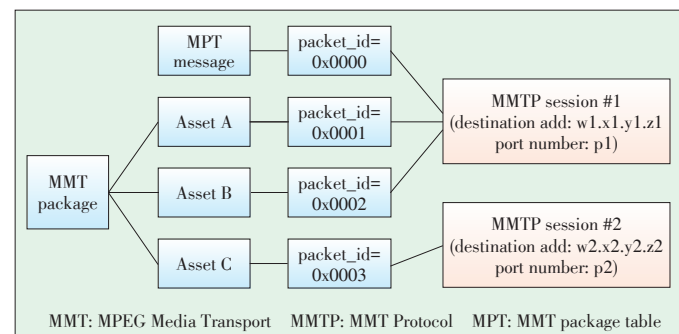
▲Figure 6. A package delivered over two MMTP sessions.

ples described above.

4.5 Synchronization

Regarding synchronization, DASH takes advantage of UTC [25] to fulfill the accurate requirement of wall clock. Since UTC can be established over the physical layer, a receiving device can obtain wall clock when connecting to ATSC 3.0 broadcast. In addition, servers of the same service can synchronize with respect to a common wall clock (UTC) source via broadcast or broadband networks. Notably, the broadcast - established wall clock is mainly used by servers for serving media components at the receiver.

In MMT, the synchronization of MPUs is also based on timestamps referencing UTC. The MPU_timestamp_descriptor as defined in [13] is used to represent the presentation time of the first media sample in terms of the presentation order in each MPU.

4.6 System Model for Media Delivery

4.6.1 ROUTE/DASH System

The generic ROUTE/DASH system model is shown in **Fig. 7**. ROUTE/DASH relies on the concept of discrete - time events, e.g., m bytes input to or n bytes output from the buffers at a specific instant. This results in no leakage rates specified. There is a transport buffer (TBn) is for a specific ROUTE session that may deliver multiple objects and related AL-FEC as encapsulated by ROUTE/UDP/IP. Objects for media services to be delivered are briefly stored in the ROUTE output buffer before they are consumed by the DASH client. The minimum size of this buffer is slightly smaller than that of the associated TBn, as the data in TBn would be wrapped in ROUTE/UDP/IP and may include AL-FEC packets. The output objects/files are decapsulated and decoded.

The EBn buffer is defined in MPEG systems as an elementary stream buffer. This buffer, when applied to ROUTE/DASH,

is integrated with the ISO BMFF file handler that holds data until it is parsed to the decoders. Given that there may exist multiple object/file streams in an LCT session, there may be several EBn(s) associated with a given LCT session. There may also be multiple media types within a single ISOBMFF file stream. As such, there are multiple decoders Dn connected to each EBn.

The task of scheduling media to the codec(s) at the receiver is handled solely by the ISOBMFF file handler (within the DASH client). Scheduling objects/files to the ISOBMFF handler is part of the DASH function. It consists of a series of steps that lead to the ISOBMFF handler receiving media delivery events (MDEs) that include byte ranges or object(s) that make contextual and temporal sense, and allowe the handler to deliver samples (media frames) to the codecs to fulfill the media presentation timeline.

The operation of the ROUTE/DASH system is defined such that none of the constituent buffers, namely, TBn, Ebn and the ROUTE Output Buffer, are allowed to overflow, and codec(s) cannnot stall due to lack of input media data. Each buffer has no data in the initialization stage and may likely become empty briefly during system operation. A notable aspect of the ROUTE/DASH system model is the lack of physical layer buffer. This is crucial for enabling ROUTE/DASH to perform at or close to the theoretical limit of the channel variation rate. ROUTE Transport Buffer Model is a loop-locked subsystem in which each module can proactively send feedback to others if needed without external assistance. This also means no extra signaling messages are necessary to support the operation of ROUTE transport buffer subsystem.

4.6.2 MMTP/MPU System

Fig. 8 shows the procedure of the content delivery, acquisition and playback of service in MMTP/MPU, which can be describe in the following steps:

1) The client acquires a service list table (SLT) to scan the channel information and service signaling messages to attain the service level information.

2) The user selects a service and identifies the corresponding MMT_package_id.

3) The channel information acquired in step 1) is used by an RF channel, and the PLP selection to precede the service selection. According to the service selection, an MMTP session carrying the corresponding MPT message is acquired.

4) In the referenced MMTP session, various content components and signaling messages are obtained through a few MMTP packets.
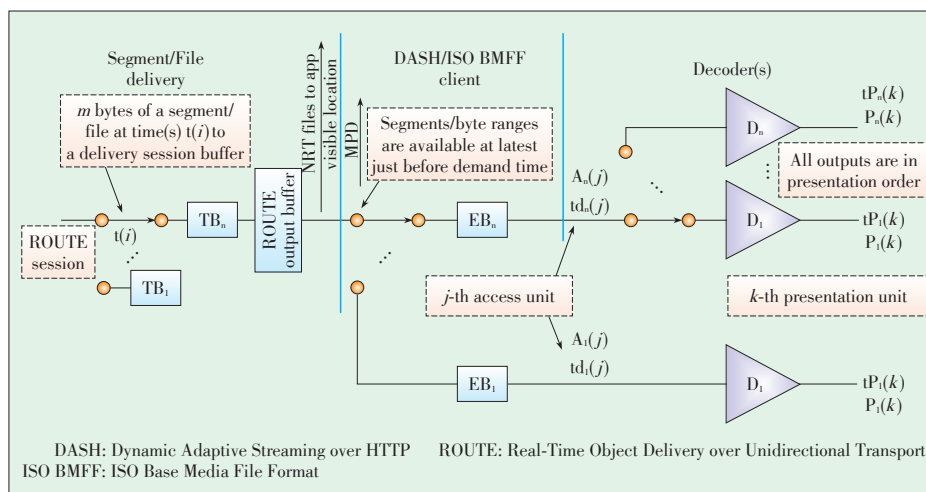
▲Figure 7. System model for ROUTE/DASH delivery.

DASH and MMT and Their Applications in ATSC 3.0

Yiling Xu, Shaowei Xie, Hao Chen, Le Yang, and Jun Sun
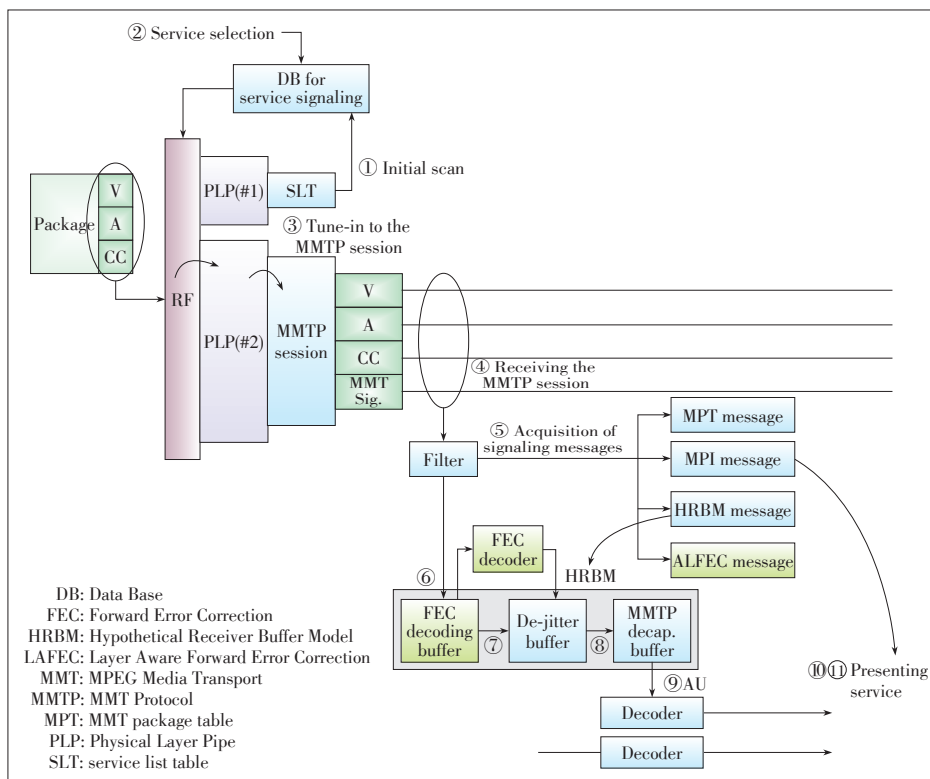
▲Figure 8. System model for MMTP/MPU delivery.

5) The type field of MMTP packet headers can inform the receiver to obtain corresponding signaling messages. The MP table extracted from the MPT message is processed to get the list of MMT assets comprising the selected service with the designated MMT_package_id. Other MMT signaling messages are processed if necessary, including the MPI, Hypothetical Receiver Buffer Model (HRBM) and AL-FEC.

6) From the MMTP packets received, different MMTP packets corresponding with each content component are filtered and stored in their corresponding FEC decoding buffers. MMTP packets with repair symbols corresponding to each content component are also received and stored separately.

7) MMTP packets received at the FEC Decoding Buffer are immediately copied into a corresponding de-jitter buffer. Missing packets can be detected using the packet_sequence_number of each MMTP packet. If a packet is not received before a predefined time specified by an AL-FEC message, it is considered missing and should be recovered by applying the AL-FEC code. The recovered packet needs to be copied to the de-jitter buffer immediately.

8) HRBM field specifies the amount time MMTP packets should spend in the de-jitter buffer.

9) MMTP packets of MPUs are processed to extract Access Units.

10) The first AU in an MPU is decoded by the appropriate decoder and presented at the time designated by the MPU_timestamp.

11) The next AU in the MPU is decoded and presented after the presentation of the first AU. This step is repeated until the last AU of the MPU has been decoded and presented.

For streaming service using MMTP/MPU, each MMTP packet with media data has an explicit indication of the boundaries of the media samples or sub-samples. As a result, MMTP packets with media data only carry minimum information about media data (such as a movie fragment sequence number and a sample number) needed to recover the association between the media data and the metadata. By contrast, ROUTE/DASH considers media segments just as payload, which indicates that more supplementary design in signaling and buffer model are needed.

In an MMTP/MPU system, HRBM is introduced so that a broadcast server can emulate the behavior of the receiver buffer, and any processing the receiver performs on the packet streams is within the reception constraints of the receiver. The HRBM message is signaled from the server to a client to guide the operation of receiver buffer subsystem. HRBM may shift more workload to the server-side, and it is possible that the subsystem does not work well in practice (e.g., buffer overflow or underflow may occur since the information provided in HRBM message may not match diverse environment situations). Moreover, the goal of HRMB is to achieve constant delay for the delivery system of the MMT receiving entity to synchronize the presentation of the media data, but this guideline may not apply for the situation that there exist vast delay gap among several transmission paths or networks. The de-jitter buffer of HRMB can mitigate the jitter introduced by multipath, multisource or multi-network though.

4.7 Rules for Session Presence

In ATSC 3.0, the rules regarding the presence of ROUTE/LCT sessions and/or MMTP sessions for carrying the content components of an ATSC 3.0 service are as follows:

1) For broadcast delivery of a linear service without application-based enhancement, the service's content components are carried by either
   - One or more ROUTE/LCT sessions, or
   - One or more MMTP sessions.

2) For broadcast delivery of a linear service with application-based enhancement, the service's content components are carried by:
   - One or more ROUTE/LCT sessions, and

DASH and MMT and Their Applications in ATSC 3.0
Yiling Xu, Shaowei Xie, Hao Chen, Le Yang, and Jun Sun

● Zero or more MMTP sessions.

The use of both MMTP and ROUTE for streaming media components in the same service is not allowed, and the selected protocol is specified in SLT.

3) For broadcast delivery of an application-based service, the content components are carried by:

● One or more ROUTE/LCT sessions.

The combination of sessions with the NRT content can only be carried by one or more ROUTE/LCT sessions (as described above). ROUTE/LCT can deliver almost all types of service components without the assistance of MMTP, but the inverse is not possible. Using ROUTE/LCT for delivery of service components appears to be more systematic than MMTP, while the latter is an alternative delivery solution for linear service in ATSC 3.0.

5 Hybrid Delivery Mode

Hybrid broadcast broadband TV (HbbTV) [26] is a globally initiative technology mainly developed by industry leaders. HbbTV aims at improving user experience when it comes to hybrid content, by harmonizing the broadcast and broadband delivery through connected TVs, set-top boxes and multiscreen devices. Combining elements of existing standards, including OIPF, CEA, DVB, MPEG-DASH and W3C, HbbTV specification defines a hybrid (broadcast and broadband) platform for the signaling, transport and presentation of enhanced or interactive applications on hybrid terminals, which include both a DVB compliant broadcast connection and a broadband connection to the Internet. In HbbTV, legacy DVB broadcast standards and systems are furthest maintained, while parts of already available standards based on IP network are largely referenced and adapted where necessary.

In order to integrate all the above technologies organically, HbbTV specification may need to afford intricate schemes due to the compatibility of data format, processing procedure and system architecture. Relatively, the design and development in ATSC 3.0 system are all based on IP network, and the transport protocols in application and physical layers are redesigned considering new requirements and scenarios. Therefore, ATSC 3.0 system could provide more systematic and fundamental solutions to support hybrid delivery of enhanced or interactive services.

This section describes hybrid delivery modes. One possible hybrid delivery operation employs ROUTE/DASH in the broadcast path and DASH over HTTP(s) in the broadband path. The other mode uses MMTP/MPU in the broadcast path and DASH over HTTP(s) in the broadband path.

5.1 Mode with ROUTE/DASH in Broadcast

The hybrid service delivery mode with ROUTE/DASH combines the broadcast delivery via ROUTE, and unicast delivery via HTTP. In some scenarios, e.g., due to device movement,

the broadcast signal may become unavailable, which would result in handover to broadband and may involve a subsequent return to the broadcast service. **Fig. 9** shows the hybrid mode operation within the DASH client.

The presence of an unmanaged broadband network is likely to introduce a variable amount of extra delay (latency) in the delivery, as the network congestion can impact the availability of the broadband-delivered content. ROUTE/DASH allows the broadcaster to employ techniques specified in the DASH standard to improve user experience in the hybrid delivery case. Buffering is very important and is handled in a slightly different way for different delivery modes. DASH segments for unicast broadband components are requested by receivers ahead of when they are needed, as indicated by the timeline set by the DASH MPD, and they are buffered until their presentation time. The broadcaster can ensure that broadband-delivered service components are made available to the broadcaster server well ahead of the broadcast-delivered components of the service to accommodate slower connections.

The receivers control delivery timing and buffering. Segments for broadcast components are delivered from the broadcast source according to the buffer model, which ensures that receivers get the segments soon enough to avoid decoder stall but not so soon as to reduce buffer overflow. The broadcast source uses the DASH MPD timeline to determine the appropriate delivery timing.

The possible scenario of switching from broadcast to broadband service access or vice versa involves the mobile ATSC 3.0 receiver, which due to user mobility, may move temporarily outside the broadcast coverage and only fallback service reception via broadband is possible. Support for such handover between different access modes may be indicated by the MPD fragment in SLS. In this case, as defined by the SLS protocols, the userServiceDescription.deliveryMethod element in the USBD fragment contains the child elements atsc:broadcastAppService and atsc:unicastAppService, which represent the broadcast and broadband delivered components of the named service. These broadcast and unicast delivered components may be substituted for one another in the case of handover from broadcast to broadband service access and vice versa.

Thanks to the uniformity of data encapsulation format in ROUTE/DASH, the system design and realization adapted for streaming service with app-based enhancement is simple and systematic, which means that there is few conversion in the me-

DASH: Dynamic Adaptive Streaming over HTTP   MPD: Media Presentation Description
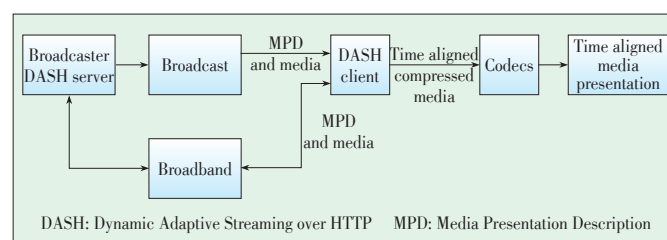
▲Figure 9. ROUTE/DASH hybrid delivery mode.

dia and signaling formats.

5.2 Mode with MMTP/MPU in Broadcast

The hybrid streaming in this mode over broadcast and broadband delivery is shown in **Fig. 10**. All the components of the system are locked to UTC for synchronization.

For the broadcast network, media data are encapsulated into MPUs, which are packetized into MMTP packets. For the broadband network, media data are encapsulated into DASH segments. When no components are delivered by broadcast, DASH MPDs are delivered over the broadcast network by a signaling message, and DASH segments are delivered via broadband by an HTTP session through the network interface of a regular HTTP server. This solution increases the system complexity and operational cost.

For the client, it is assumed that the MMTP packets delivered through the broadcast network are de-packetized and the media data are decoded by the appropriate media decoders, and that the DASH segments are delivered through the broadband network. To synchronize the presentation of a DASH segment delivered via the broadband network with an MPU delivered via the broadcast network, the presentation time of the DASH segment is represented by a timestamp referencing UTC.

Like the mode with ROUTE/DASH used in broadcast, when it comes to the switch frombroadcast to broadband, the client uses, before the transition, the latest MPD contained in an MPI message to make an HTTP request for a DASH segment. The received DASH segment is buffered and made available to the DASH client for decoding. For seamless handoff, the broadcast delivery delay should be adjusted so that it is equal to the broadband delivery delay; otherwise, service may freeze when switching from broadcast to broadband for the first time. Meanwhile, the DASH segment and MPU are inconsistent with each other in format, resulting in more processing time and delay in data conversion. Also, for handoff services, MPU boundary may overlap with DASH segment boundary, which can create barriers to seamless handoff and further degrade the service quality.
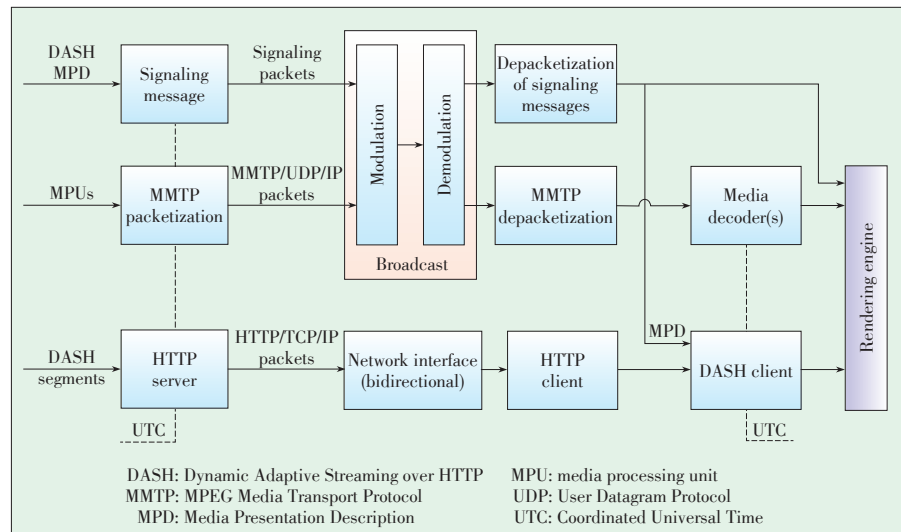
Becasue a transition can only happen on an MPU boundary, at least one DASH segment whose boundary matches with broadcast MPUs needs to be buffered in advance. In **Fig. 11**, the client has detected a loss in broadcast signaling for MPU n, and makes an HTTP request for the DASH segment cor-

responding to the lost nth MPU. There is no transition delay because MPU n-1 is presented during the time the client is receiving the designated DASH segment. The HRBM is used by the broadcaster to add a specific amount of delay required for the seamless transition. When the broadcast signal becomes available again, there is a seamless transition to the broadcast in this example.
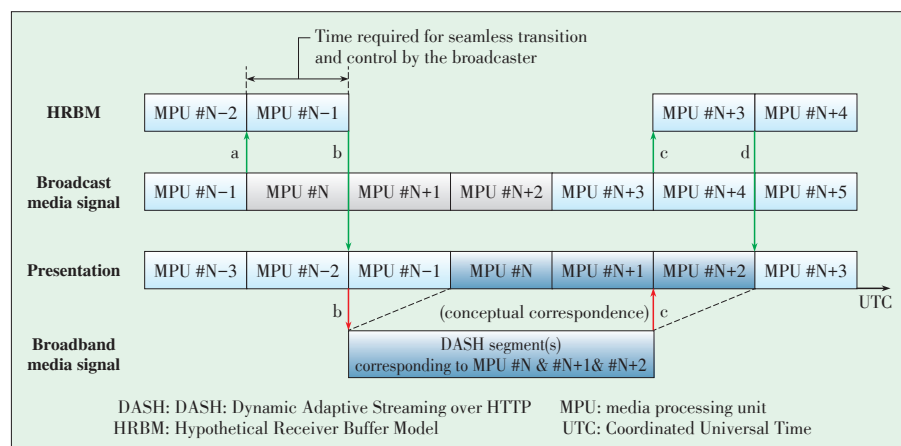
6 Conclusions

This paper described the features and design considerations of ATSC 3.0, a next-generation broadcasting system designed to address the emerging multimedia service delivery requirements with using broadcasting channels in combination with broadband networks. It also provided comparative comparative introductions and applications in details about ROUTE/DASH and MMTP/MPU adopted in the transport protocols used in ATSC 3.0 for broadcasting.

In ATSC 3.0, the ROUTE/DASH is derived from FLUTE

DASH: Dynamic Adaptive Streaming over HTTP
MMTP: MPEG Media Transport Protocol
MPD: Media Presentation Description
MPU: media processing unit
UDP: User Datagram Protocol
UTC: Coordinated Universal Time

▲Figure 10. MMTP/MPU hybrid delivery mode.

DASH: DASH: Dynamic Adaptive Streaming over HTTP
HRBM: Hypothetical Receiver Buffer Model
MPU: media processing unit
UTC: Coordinated Universal Time

▲Figure 11. Seamless transition: broadcast-broadband-broadcast.

[27] and DASH, FLUTE is designed for NRT data push services over unidirectional transport and DASH was developed to realize dynamic adaptive streaming over HTTP. It is the first time that ROUTE/DASH has been applied to a broadcasting system. MMTP/MPU was accepted as the MMT standard in 2012. In Japan, super hi‑vision test services are scheduled to begin in 2016, and commercial services are scheduled to begin in 2020 using MMT as a transport protocol for next‑generation broadcasting systems.

As developments of ROUTE/DASH and MMT standards, as well as next generation broadcasting system is still in progress, further study and verification of technologies adopted in ATSC 3.0 are still needed to be done.

References

[1] *2nd MMT Workshop in Kyoto: Presentations*, ISO/IEC JTC 1/SC 29/WG 11 N11200, MPEG, 2010.
[2] P. Podhradsky, "Evolution trends in Hybrid Broadcast Broadband TV", *55th International Symposium ELMAR-2013*, Zadar, Yugoslavia, Sept. 2013.
[3] *Information Technology — Generic Coding of Moving Pictures and Associated Audio Information: Part 1 Systems*, ISO/IEC 13818-1, 2013.
[4] *Information Technology — Coding of Audio-Visual Objects — Part 12 ISO Base Media File Format*, ISO/IEC 14496-12, 2012.
[5] *Digital Video Broadcasting (DVB), Generic Stream Encapsulation (GSE) Protocol*, ETSI TS 102 606 V1.1.1, Oct. 2007.
[6] R. Pantos and E. W. May. (2011, Oct. 2). HTTP Live Streaming [Online]. Available: http://tools.ietf.org/html/draft-pantos-http-live-streaming-06
[7] Microsoft. (2009, Sep. 8). IIS Smooth Streaming Transport Protocol [Online]. Available: http://www.iis.net/learn/media/smooth-streaming/smooth--streaming-transport-protocol
[8] *Information technology — Dynamic adaptive streaming over HTTP (DASH) — Part 1: Media presentation description and segment formats*. ISO/IEC 23009‑1, May 2014.
[9] I. Sodagar, "The MPEG-DASH standard for multimedia streaming over the internet," *IEEE Multimedia*, vol. 18, no. 4, pp. 62−67, Apr. 2011. doi: 10.1109/MMUL.2011.71.
[10] *ATSC 3.0 Management and Protocols − Signaling, Delivery, Synchronization, and Error Protection*. ATSC Working Draft S33-174r0, Dec. 2015.
[11] T. Stoekhammer, "Dynamic adaptive streaming over HTTP‑design principles and standards," in *Proc. Second Annual ACM Conference on Multimedia Systems*, New York, USA, 2011, pp. 2−4.
[12] C. Müller, S. Lederer, and C. Timmerer, "An evaluation of dynamic adaptive streaming over HTTP in vehicular environments," in *Proc. 4th Workshop on Mobile Video*, New York, USA, 2012, pp. 37−42. doi: 10.1145/2151677.2151686.
[13] *Information technology — High efficiency coding and media delivery in heterogeneous environments — Part 1: MPEG media transport (MMT)*, ISO/IEC 23008-1, 2nd Edition, Jun. 2015.
[14] S. Aoki, K. Otsuki, and H. Hamada, "New media transport technologies in super hi‑vision broadcasting systems," in *Proc. International Broadcasting Convention*, Amsterdam, Netherlands, Sept. 2013, pp. 7− 9. doi: 10.1049/ibc.2013.0029.
[15] Y. Lim, K. Park, J. Y. Lee, S. Aoki, and G. Fernando, "MMT: an emerging MPEG standard for multimedia delivery over the internet," *IEEE Multimedia*, vol. 20, no. 1, pp. 80−85, Mar. 2013. doi: 10.1109/MMUL.2013.7.
[16] S. Aoki, K. Otsuki, and H. Hamada, "Effective usage of MMT in broadcasting systems," in *IEEE International Symposium on Broadband Multimedia Systems and Broadcasting*, London, UK, Jun. 2013, pp. 1− 6. doi: 10.1109/BMSB.2013.6621756.
[17] W3C. HTML5: A vocabulary and associated APIs for HTML and XHTML [Online]. Available: http://www.w3.org/TR/2014/REChtml5-20141028/
[18] *Layered Coding Transport (LCT) Building Block*, IETF: RFC 5651, Oct. 2009.
[19] *Hypertext Transfer Protocol -- HTTP/1.1*, IETF: RFC 2616, Jun. 1999.
[20] Y. Lim, S. Aoki, I. Bouazizi, and J. Song, "New MPEG transport standard for next generation hybrid broadcasting system with IP," *IEEE Transactions on Broadcasting*, vol. 60, no. 2, pp. 160− 169, Jun. 2014. doi: 10.1109/TBC.2014.2315472.
[21] C. Müller, D. Renzi, S. Lederer, *et al.*, "Using scalable video coding for dynamic adaptive streaming over HTTP in mobile environments," in *Proc 20th European Signal Processing Conference (EUSIPCO)*, Bucharest, Romania, Aug. 2012, pp. 2208−2212.
[22] S. Aoki and K. Aoki, "Efficient multiplexing scheme for IP packets over the advanced satellite broadcasting system," *IEEE Transactions on Consumer Electronics*, vol. 55, no. 1, pp. 49−55, Feb. 2009. doi: 10.1109/TCE.2009.4814413.
[23] *Forward Error Correction (FEC) Framework*, IETF: RFC 6363, Oct. 2011.
[24] Joint W3C/IETF URI Planning Interest Group. (2001, Sept. 21). URIs, URLs, and URNs: Clarifications and Recommendations 1.0 [Online]. Available: https://www.w3.org/TR/uri-clarification/
[25] W3C. (1997, Sept. 15). Date and Time Formats [Online]. Available: http://www.w3.org/TR/1998/NOTE-datetime-19980827
[26] *Hybrid Broadcast Broadband TV*, ETSI TS 102 796 V1.3.1, Oct. 2015.
[27] *FLUTE - File Delivery over Unidirectional Transport*, IETF: RFC 6726, Nov. 2012.

Biographies

Yiling Xu (yl.xu@sjtu.edu.cn) received her PhD in electrical engineering from the University of Electronic Science and Technology of China. From 2004 to 2013, she was with Multimedia Communication Research Institute of Samsung Electronics Inc, where she focused on digital broadcasting, IPTV, convergence networks and next‑generation multimedia applications. Currently, she is with the Shanghai Jiaotong University, as a research associate. Dr. Xu has more than 70 patents and 10 academic papers. She is active in international standard organizations including DVB, MPEG, 3GPP, OIPF, OMA, and FOBTV.

Shaowei Xie (sw.xie@sjtu.edu.cn) received his BE degree in electronics and information engineering from Northwestern Polytechnical University, China in 2014. He is pursuing his PhD degree at the Institute of Image Communication and Signal Processing, Shanghai Jiao Tong University, China.His research interest is multimedia communication.

Hao Chen (chenhao1210@sjtu.edu.cn) received his BE degree in electronics and information engineering from Northwestern Polytechnical University, China in 2013. He is pursuing his PhD degree at the Institute of Image Communication and Signal Processing, Shanghai Jiao Tong University, China. His research interest is multimedia communication.

Le Yang (le.yang.le@gmail.com) received his BEng. and MSc degrees in electrical engineering from the University of Electronic Science and Technology of China, China in 2000 and 2003, respectively. He received his PhD degree in electrical and computer engineering from the University of Missouri, USA in 2010. Since 2011, he has been an associate professor with Jiangnan University, China. His research interests include sensor networks, passive localization, tracking and signal detection.

Jun Sun (junsun@sjtu.edu.cn) received his BS and MS degrees from the University of Electronic Science and Technology of China in 1989 and 1992, respectively, and the PhD degree from Shanghai Jiao Tong University, China in 1995. He is a professor with the Institute of Image Communication and Information Processing of Shanghai Jiao Tong University, China. His research interests include image communication, HDTV, and mobile communication.