ATSC Recommended Practice: Guide to the ATSC Mobile DTV Standard

Doc. A/154:2013
30 January 2013

Advanced Television Systems Committee
1776 K Street, N.W., Suite 200
Washington, D.C. 20006
202-872-9160

The Advanced Television Systems Committee, Inc., is an international, non-profit organization developing voluntary standards for digital television. The ATSC member organizations represent the broadcast, broadcast equipment, motion picture, consumer electronics, computer, cable, satellite, and semiconductor industries.

Specifically, ATSC is working to coordinate television standards among different communications media focusing on digital television, interactive systems, and broadband multimedia communications. ATSC is also developing digital television implementation strategies and presenting educational seminars on the ATSC standards.

ATSC was formed in 1982 by the member organizations of the Joint Committee on InterSociety Coordination (JCIC): the Electronic Industries Association (EIA), the Institute of Electrical and Electronic Engineers (IEEE), the National Association of Broadcasters (NAB), the National Cable Telecommunications Association (NCTA), and the Society of Motion Picture and Television Engineers (SMPTE). Currently, there are approximately 160 members representing the broadcast, broadcast equipment, motion picture, consumer electronics, computer, cable, satellite, and semiconductor industries.

ATSC Digital TV Standards include digital high definition television (HDTV), standard definition television (SDTV), data broadcasting, multichannel surround-sound audio, and satellite direct-to-home broadcasting.

Note: The user's attention is called to the possibility that compliance with this Recommended Practice may require use of an invention covered by patent rights. By publication of this document, no position is taken with respect to the validity of this claim or of any patent rights in connection therewith. One or more patent holders have, however, filed a statement regarding the terms on which such patent holder(s) may be willing to grant a license under these rights to individuals or entities desiring to obtain such a license. Details may be obtained from the ATSC Secretary and the patent holder.

Revision History

Version	Date
Initial version of A/154 approved (A/154:2011)	15 April 2011
Revision 1 approved (A/154:2013)	30 January 2013

Table of Contents

Index of Figures and Tables

ATSC Recommended Practice: Guide to the ATSC Mobile DTV Standard (Doc. A/154:2013)

1. SCOPE

The ATSC Mobile DTV service shares the same RF channel as the standard ATSC broadcast service described in ATSC A/53 ("ATSC Digital Television Standard, Parts 1 – 6"). The purpose of this Recommended Practice is to describe how the technology is documented in the Standard, explain what is enabled by in the Standard (technically and functionally), and to provide recommendations for the emission systems.

The initial release of this Recommended Practice (A/154:2011) provided an overview of the system and guidance for Parts 3, 4 and 6 of the ATSC Mobile DTV Standard, as documented in references [2], [3], and [4]. The second release of this RP (A/154:2013) provided guidance for Part 8, as documented in reference [6]. Guidelines covering additional Parts of A/153 are expected to be added via revision of this document at a later date.

1.1 Introduction and Background

The ATSC Mobile DTV system (also known as M/H, "mobile/handheld") provides for mobile/pedestrian/handheld broadcasting services using a portion of the ~19.39 Mbps ATSC 8-VSB payload, while the remainder is still available for HD and/or multiple SD television services. ATSC Mobile DTV is a dual-stream system—the ATSC service multiplex for existing digital television services and the M/H service multiplex for one or more mobile, pedestrian and handheld services.

ATSC Mobile DTV is built around a highly robust transmission system based on vestigial sideband (VSB) modulation coupled with a flexible and extensible Internet Protocol (IP) based transport system, efficient MPEG AVC (ISO/IEC 14496-10 or ITU H.264) video, and HE AAC v2 audio (ISO/IEC 14496-3) coding. The Standard describes the methodology for new services to be carried in digital broadcast channels along with current DTV services without any adverse impact on legacy receiving equipment.

A/153 is modular in concept, with the specifications for each of the modules contained in separate Parts. The major Parts are as follows:

- **Part 1** – "Mobile/Handheld Digital Television System"
- **Part 2** – "RF/Transmission System Characteristics"
- **Part 3** – "Service Multiplex and Transport Subsystem Characteristics"
- **Part 4** – "Announcement"
- **Part 5** – "Application Framework"
- **Part 6** – "Service Protection"
- **Part 7** – "Video System Characteristics"
- **Part 8** – "Audio System Characteristics"

Part 1 describes the overall ATSC Mobile DTV system and explains the organization of the standard.

Because of the complexity of the ATSC Mobile DTV system, it was recognized that guidelines for implementers would be beneficial. This Recommended Practice was developed to address this need.

1.2 Organization

This document is organized as follows:

- **Section 1** – Outlines the scope of this document and provides a general introduction
- **Section 2** – Lists documents referenced in this RP
- **Section 3** – Provides a definition of terms, acronyms, and abbreviations for this document
- **Section 4** – ATSC Mobile DTV system overview
- **Section 5** – RF transmission system characteristics
- **Section 6** – Data transport
- **Section 7** – Signaling data delivery and usage
- **Section 8** – Announcement data delivery and usage
- **Section 9** – Streaming data delivery
- **Section 10** – File delivery
- **Section 11** – Application framework
- **Section 12** – Service protection data delivery and usage
- **Section 12** – Video system characteristics
- **Section 13** – Audio system characteristics
- **Annex A** – Mapping of Service Map Table to Service Guide

2. REFERENCES

At the time of publication, the editions indicated were valid. All referenced documents are subject to revision, and users of this Recommended Practice are encouraged to investigate the possibility of applying the most recent edition of the referenced document.

2.1 Informative References

The following documents contain information that may be helpful in applying this Recommended Practice.

[1] IEEE/ASTM: "Use of the International Systems of Units (SI): The Modern Metric System," Doc. SI 10-2002, Institute of Electrical and Electronics Engineers, New York, NY.

[2] ATSC: "ATSC Mobile DTV Standard, Part 3 – Service Multiplex and Transport Subsystem Characteristics," Doc. A/153 Part 3:2009, Advanced Television Systems Committee, Washington, DC, 15 October 2009.

[3] ATSC: "ATSC Mobile DTV Standard, Part 4 – Announcement," Doc. A/153 Part 4:2009, Advanced Television Systems Committee, Washington, DC, 15 October 2009.

[4] ATSC: "ATSC Mobile DTV Standard, Part 6 – Service Protection," Doc. A/153 Part 6:2009, Advanced Television Systems Committee, Washington, DC, 15 October 2009.

[5] ATSC: "ATSC Mobile DTV Standard, Part 7 – AVC and SVC Video System Characteristics," Doc. A/153 Part 7:2009, Advanced Television Systems Committee, Washington, DC, 15 October 2009.

[6] ATSC: "ATSC Mobile DTV Standard, Part 8 – HE AAC Audio System Characteristics," Doc. A/153 Part 8:2009, Advanced Television Systems Committee, Washington, DC, 15 October 2009.

[7] IETF: “RTP Payload Format for H.264 Video,” Doc. RFC 3984, Internet Engineering Task Force, Reston, VA, February 2005.

[8] IETF: “RTP Payload Format for SVC Video,” Internet Draft draft-ietf-avt-rtp-svc-20.txt, Internet Engineering Task Force, Reston, VA, October 2009.

[9] IETF: “Signaling Media Decoding Dependency in the Session Description Protocol (SDP),” Doc. RFC 5583, Internet Engineering Task Force, Reston, VA, July 2009.

[10] ISO/IEC: ISO/IEC 14496-10 (ITU-T H.264), International Standard (2007), “Advanced video coding for generic audiovisual services” with Corrigendum 1 (01/2009), International Standards Organization, Geneva, Switzerland.

[11] IETF: “SDP Descriptors for FLUTE,” Internet Draft draft-mehta-rmt-flute-sdp-05.txt, Internet Engineering Task Force, Reston, VA, January 2006.

[12] OMA: “OMA DRM v2.0 Extensions for Broadcast Support,” Open Mobile Alliance, OMATS-DRM_XBS-V1_0, available from http://www.openmobilealliance.org

[13] OMA: “Service and Content Protection for Mobile Broadcast Services,” Doc. OMA-TS-BCAST_SvcCntProtection-V1_0, Open Mobile Alliance.

[14] ATSC: “ATSC Mobile DTV Standard, Part 2 – RF/Transmission System Characteristics,” Doc. A/153 Part 2:2009, Advanced Television Systems Committee, Washington, DC, 15 October 2009.

[15] ATSC: “Program and System Information Protocol For Terrestrial Broadcast and Cable,” Doc. A/65:2009, Advanced Television Systems Committee, Washington, DC, 14 April 2009.

[16] ISO/IEC: ISO/IEC 639.1, “Codes for the Representation of Names of Languages – Part 1: Alpha-2 Code, as maintained by the ISO 639/Joint Advisory Committee (ISO 639/JAC),” International Standards Organization, Geneva, Switzerland.

[17] ISO/IEC: ISO/IEC 639.2/B, “Code for the Representation of Names of Languages – Part 2: alpha-3 code, as maintained by the ISO 639/Joint Advisory Committee (ISO 639/JAC),” International Standards Organization, Geneva, Switzerland.

[18] IETF: “Tags for the Identification of Languages”, Doc. RFC 3066, Internet Engineering Task Force, Reston, VA, January 2001.

[19] IETF: “Tags for Identifying Languages”, Doc. RFC 4646, Internet Engineering Task Force, Reston, VA, September 2006.

[20] IETF: “Tags for Identifying Languages”, Doc. RFC 5646, Internet Engineering Task Force, Reston, VA, September 2009.

[21] ISO/IEC: ISO/IEC 8859-1:1998, “Information technology – 8-bit single-byte coded graphic character sets -- Part 1: Latin alphabet No. 1,” International Standards Organization, Geneva, Switzerland.

[22] IETF: “Network Time Protocol Version 4: Protocol and Algorithms Specification,” Doc. RFC 5905, Internet Engineering Task Force, Reston, VA, June 2010.

[23] OMA: “Service Guide for Mobile Broadcast Services,” Version 1.0, Doc. OMA-TS-BCAST_Service_Guide-V1_0, Open Mobile Alliance.

[24] IETF: “FLUTE – File Delivery over Unidirectional Transport,” Doc. RFC 3926, Internet Engineering Task Force, Reston, VA, October 2004.

[25] IETF: “Asynchronous Layered Coding (ALC) Protocol Instantiation,” Doc. RFC 3450, Internet Engineering Task Force, Reston, VA.

[26] IETF: “Layered Coding Transport (LCT) Building Block,” Doc. RFC 3451, Internet Engineering Task Force, Reston, VA.
[27] ATSC: “ATSC Digital Television Standard,” Doc. A/53 Parts 1–6:2007/2009, Advanced Television Systems Committee, Washington, D.C. 3 January 2007 and 7 August 2009.
[28] ATSC: “ATSC Mobile DTV Standard, Part 5 – Application Framework,” Doc. A/153 Part 5:2009, Advanced Television Systems Committee, Washington, D, 15 October 2009.
[29] ISO: “Information technology – Coding of audio-visual objects – Part 3: Audio,” Doc. ISO/IEC 14496-3:2009, International Standards Organization, Geneva, Switzerland.
[30] ATSC: “Recommended Practice – Techniques for Establishing and Maintaining Audio Loudness for Digital Television,” Doc. A/85, Advanced Television Systems Committee, Washington, D.C., 25 July 2011.
[31] ITU: “Algorithms to measure audio programme loudness and true-peak audio level,” Recommendation ITU-R BS.1770-3, International Telecommunications Union, Geneva, 2012.

3. DEFINITION OF TERMS

With respect to definition of terms, abbreviations, and units, this document follows the practice of the Institute of Electrical and Electronics Engineers (IEEE) as outlined in the Institute’s published standards [1]. Where an abbreviation is not covered by IEEE practice or industry practice differs from IEEE practice, the abbreviation in question is described in Section 3.3 of this document.

3.1 Treatment of Syntactic Elements

This document contains symbolic references to syntactic elements used in the audio, video, and transport coding subsystems. These references are typographically distinguished by the use of a different font (e.g., restricted), may contain the underscore character (e.g., sequence_end_code) and may consist of character strings that are not English words (e.g., dynrng).

3.2 Acronyms and Abbreviation

The following acronyms and abbreviations are used within this document.

ALC – Asynchronous Layered Coding

ATSC – Advanced Television Systems Committee

ATSC-M/H – ATSC Mobile and Handheld

AVC – Advanced Video Coding (ITU-T H.264)

BCRO – Broadcast Rights Object

bslbf – Bit string, left bit first

BSM – BCAST Subscription Management

CIT-MH – Cell Information Table for ATSC-M/H

CRC – Cyclic Redundancy Check

DRC – Dynamic Range Control

DIMS – Dynamic Interactive Multimedia Service

DNS – Domain Name System

ESG – Electronic Service Guide

FDT – File Delivery Table

FEC – Forward Error Correction
FIC – Fast Information Channel
FLUTE – File Delivery over Unidirectional Transport (RFC 3926)
GAT-MH – Guide Access Table for ATSC-M/H
HTTP – Hypertext Transfer Protocol
IETF – Internet Engineering Task Force
IP – Internet Protocol
IPsec – IP Security
LCT – Layered Coding Transport
LKFS – Loudness, K-weighted, relative to Full Scale, measured with equipment that implements the algorithm specified by ITU-R BS.1770 [31]. A unit of LKFS is equivalent to a decibel.
LTKM – Long Term Key Message
M/H – Mobile and Handheld
MDTV – Mobile Digital Television
NoG – Number of M/H Groups per M/H Subframe for a designated Parade
NTP – Network Time Protocol
OMA – Open Mobile Alliance
OMA-BCAST – Open Mobile Alliance Broadcast
PEK – Program Encryption Key
PS – Parametric Stereo
RI – Rights Issuer
RME – Rich Media Environment
RO – Right Object
ROT – Root Of Trust
RTP – Real-time Transport Protocol
SBR – Spectral Band Replication
SDP – Session Description Protocol
SEK –Service Encryption Key
SG – (Electronic) Service Guide
signed int – signed integer
SLT-MH – Service Labeling Table for ATSC-M/H
SMT-MH – Service Map Table for ATSC-M/H
SRTP – Secure Real Time Protocol
STKM – Short Term Key Message
SVC – Scalable Video Coding
TCP – Transmission Control Protocol
TEK – Traffic Encryption Key
TSI – Transport Session Identifier
UDP – User Datagram Protocol
uimsbf – Unsigned integer, most significant bit first

3.3 Terms

The following terms are used within this document.

Broadcast System – The collection of equipment necessary to transmit signals of a specified nature.

IP multicast stream – An IP stream in which the destination IP address is in the IP multicast address range.

IP stream – A sequence of IP datagrams with the same source IP address and the same destination IP address.

Local M/H Service – A Service which appears in one and only one MH Broadcast. Typically this is a Service created by a local broadcaster which will not be transmitted by another broadcast facility other than a repeater.

M/H Broadcast – The entire M/H portion of a physical transmission channel.

M/H Ensemble (or simply "Ensemble") – A collection of consecutive RS Frames with the same FEC codes, where each RS Frame encapsulates a collection of IP streams.

M/H Group – At the packet level, a collection of 118 consecutive MHE-encapsulated MPEG-2 transport packets delivering M/H Service data; also, the corresponding data segments after interleaving and trellis coding.

M/H Multiplex – A collection of M/H Ensembles in which the same IP protocol version is used for all the IP datagrams in the collection, and the IP addresses of the IP streams in the M/H Services in the Ensembles have been coordinated to avoid any IP address collisions. A single M/H Multiplex may include one or more M/H Ensembles.

M/H Parade (or simply "Parade) – A collection of M/H Groups that have the same M/H FEC parameters. Each M/H Parade carries one or two M/H Ensembles.

M/H Service – A package of packetized streams transmitted via an M/H Broadcast, which package is composed of a sequence of events which can be broadcast as part of a schedule.

M/H Service Signaling Channel – A single stream incorporated within each M/H Ensemble. The current version of the M/H-SSC uses an IP multicast stream to deliver M/H Service Signaling tables that include IP-level M/H Service access information.

M/H Slot – A portion of an M/H Sub-Frame consisting of 156 consecutive MPEG-2 transport packets. An M/H-Slot may consist of all main packets or may consist of 118 M/H packets and 38 main packets. There are 16 M/H Slots per M/H Sub-Frame.

M/H Subframe – One fifth of an M/H Frame; each M/H Subframe is equal in size to 4 VSB data frames (8 VSB data fields).

M/H TP – The term "M/H Transport Packet (M/H TP)" is used to designate a row of an RS Frame payload with two bytes header included. Thus, each RS Frame payload is composed of 187 M/H TPs.

MTU – The maximum sized datagram that can be transmitted through the next network.

Program – A collection of associated media streams that have a common timeline for a defined period. A program corresponds to the common industry usage of "television program."

Reference Receiver – A physical embodiment of hardware, operating system, and native applications of the manufacturer's choice, which collectively constitute a receiver for which specified transmissions are intended.

Rights Issuer – An entity that issues Rights Objects to OMA DRM Conformant Devices (as defined in OMA DRM [12]).

RI Stream –A stream of UDP packets with the common source and destination IP addresses and UDP port, containing RI Objects.

Rights Object – A collection of Permissions and other attributes which are linked to Protected Content (as defined in OMA DRM [12]).

RS Frame – Two-dimensional data frame through which an M/H Ensemble is RS CRC encoded. RS Frame is the output of M/H physical layer subsystem. One RS Frame contains 187 rows of N bytes each in its payload, where the value of N is determined by the transmission mode of M/H physical layer subsystem, and carries data for one M/H Ensemble.

UDP stream – A sequence of UDP/IP datagrams with the same destination IP address and the same destination UDP port number.

4. ATSC MOBILE DTV SYSTEM OVERVIEW

In very simple terms, the system achieves the robustness needed for mobile reception by adding extra training sequences and forward error correction. The total bandwidth needed for the ATSC Mobile DTV service depends on several factors, including the number and type of program services, the quality level, and level of robustness desired, typically ranging from less than one megabit per second to many megabits per second. The ATSC Mobile DTV system converts the current 8-VSB emission into a dual-stream system without altering the emitted spectral characteristics. It does this by selecting some of the MPEG-2 segments (corresponding to MPEG-2 Transport packets in the current system) and allocating the payloads in those segments to carry the M/H data in a manner that existing legacy receivers will ignore. Figure 4.1 is a high level block diagram showing the current 8-VSB and the new Mobile subsystems at the studio/transmitter side.

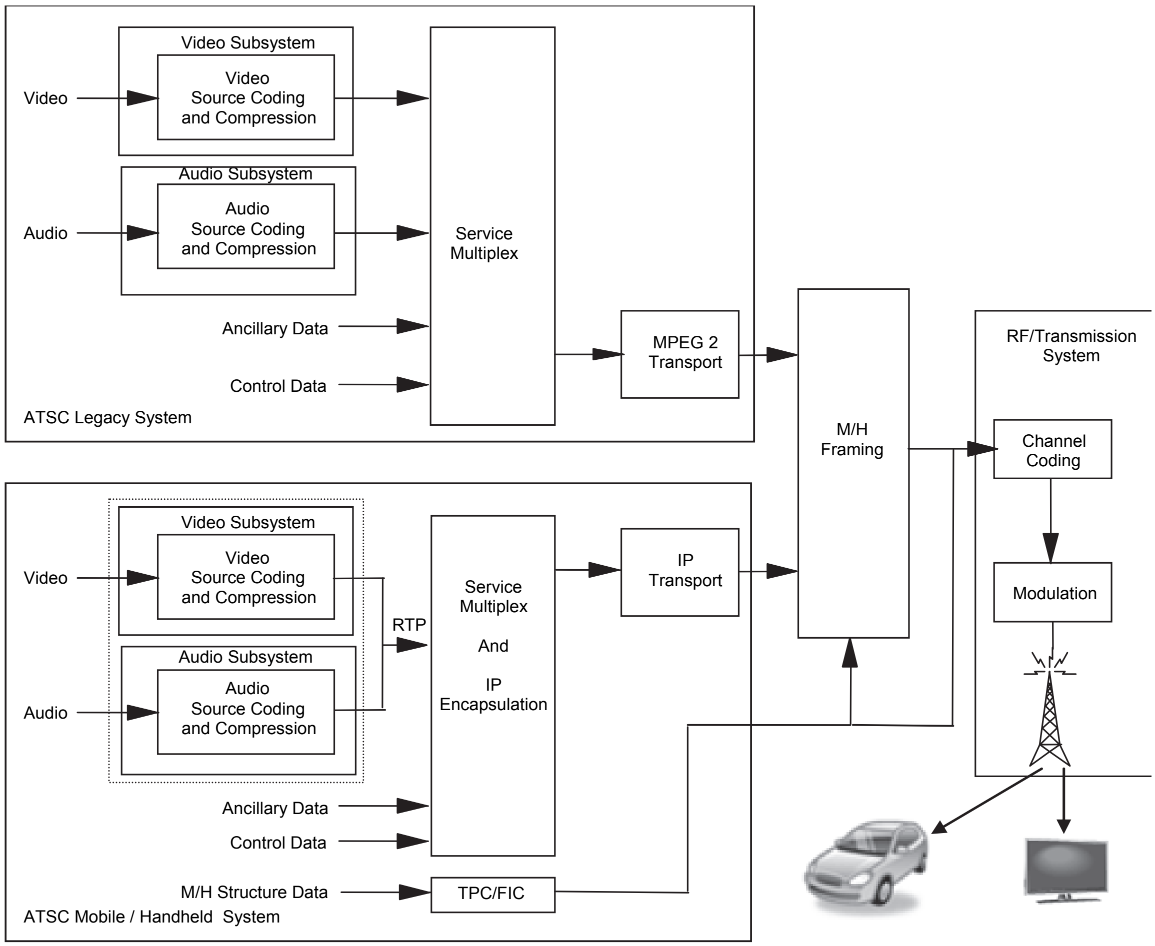

Figure 4.1 The ATSC Mobile DTV system.

A different view showing some of the functions and how they interrelate as part of the broadcast chain is shown in Figure 4.2.

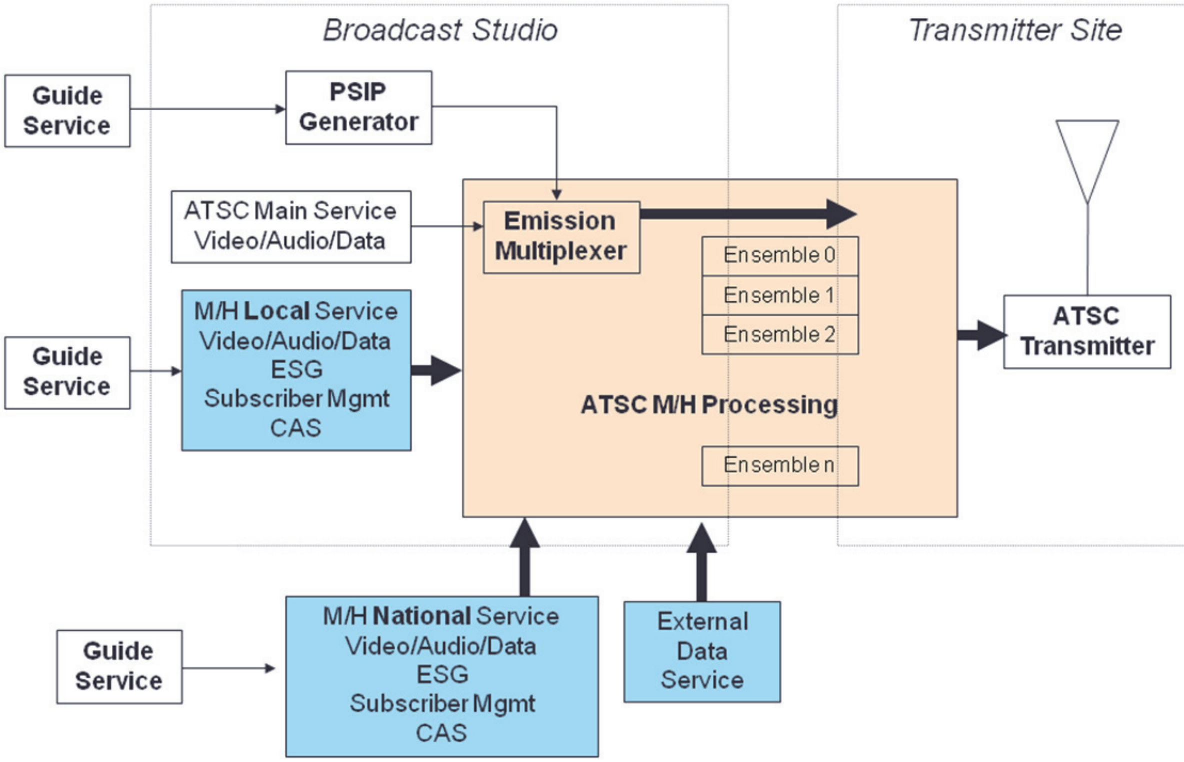

Figure 4.2 ATSC Mobile DTV broadcast system block diagram.

Another way to show the various parts and how they relate is a protocol stack. The overall ATSC Mobile DTV system broadcast protocol stack is illustrated in Figure 4.3. Consideration was given to the many system details that make such a signal compatible with legacy ATSC receivers, particularly audio decoder buffer constraints; but also such constraints as MPEG transport packet header standards, requirements for legacy PSIP carriage, etc. Those will be addressed in more detail in subsequent editions of this RP.

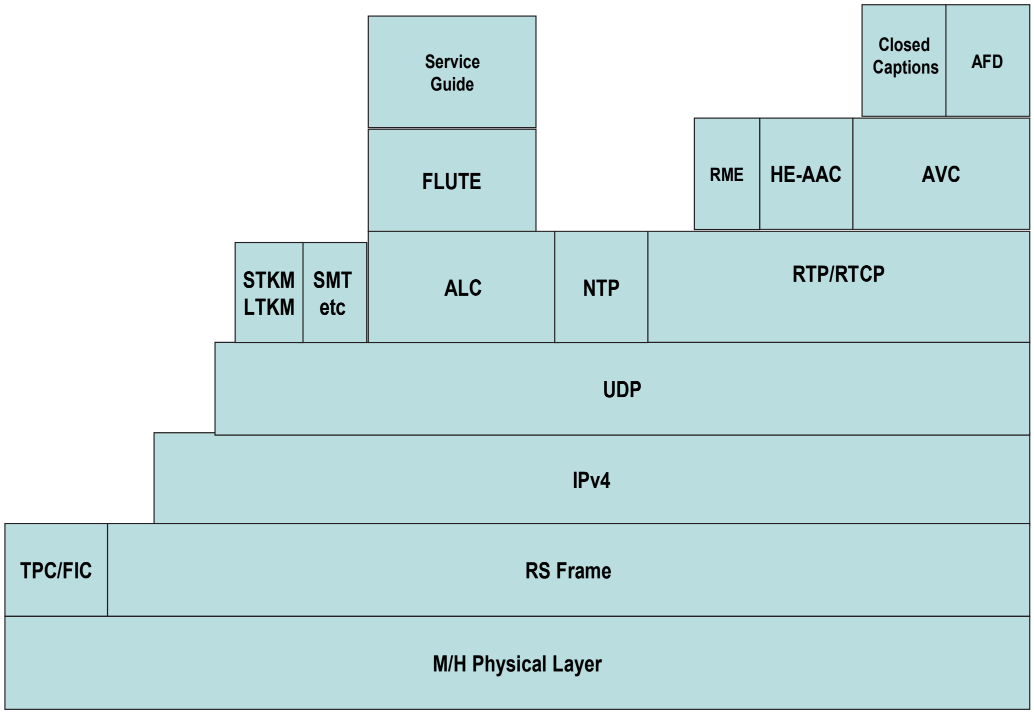

Figure 4.3 ATSC-M/H broadcast protocol stack**.**

4.1 Structure of the Standard

A/153 is modular in concept, with the specifications for each of the modules contained in separate *Parts*. (See Figure 4.4.) The individual Parts of A/153 are as follows:

Part 1 – "Mobile/Handheld Digital Television System"

Part 1 describes the overall ATSC Mobile DTV system and explains the organization of the standard. It also describes the explicit signaling requirements that are implemented by data structures throughout the other Parts.

Part 2 – "RF/Transmission System Characteristics"

Part 2 describes how the data is processed and placed into the VSB frame. Major elements include the Reed-Solomon (RS) Frame, a Transmission Parameter Channel (TPC), and a Fast Information Channel (FIC).

Part 3 – "Service Multiplex and Transport Subsystem Characteristics"

Part 3 covers the service multiplex and transport subsystem, which comprises several layers in the stack. Major elements include Internet Protocol (v4), UniDirectional Protocol (UDP), Signaling Channel Service, FLUTE over Asynchronous Layered Coding (ALC) / Layered Coding Transport (LCT), Network Time Protocol (NTP) time service, and Real Time Protocol (RTP) / Real Time Transport Control Protocol (RTCP).

Part 4 – "Announcement"

Part 4 covers Announcement, where services can optionally be announced using a Service Guide. The guide used in the standard is based on an Open Mobile Alliance (OMA) broadcast (BCAST) Service Guide, with constraints and extensions.

Part 5 – "Application Framework"

Part 5 defines the Application Framework, which enables the broadcaster of the audio-visual service to author and insert supplemental content to define and control various additional elements of the Rich Media Environment (RME).

Part 6 – "Service Protection"

Part 6 covers Service Protection, which refers to the protection of content, either files or streams, during delivery to a receiver. Major elements include the Right Issue Object and Short-Term Key Message (STKM).

Part 7 – "Video System Characteristics"

Part 7 defines the AVC and SVC Video System in ATSC Mobile DTV. Additional elements covered in this Part included closed captioning (CEA 708) and Active Format Description (AFD).

Part 8 – "Audio System Characteristics"

Part 8 defines the HE-AAC v2 Audio System in ATSC Mobile DTV.

While the basic capability for sending files is defined in Part 3, it is very general and broad. A constrained set of interoperability points are being defined in a separate ATSC activity (in TSG S/13) where non-real-time (NRT) file delivery requirements are being documented for use in both the ATSC Mobile DTV system and the current DTV (8-VSB) system.

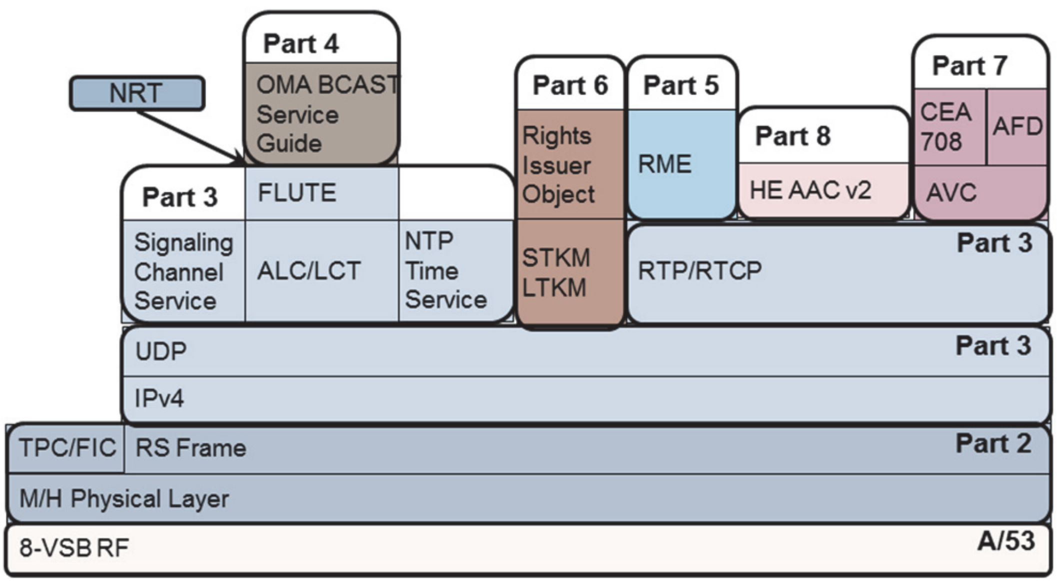

Figure 4.4 The overall A/153 documentation structure.

TSG/S4 divided the ATSC Mobile DTV task into four main elements, with most of the detailed work taking place in those sub-groups:

- **S4-1**, Physical Layer Group focusing on the RF, forward-error-correction, and legacy transport elements.
- **S4-2,** Management Layer Group focusing on transport, signaling, announcement, streaming and file delivery, service protection, and content protection.
- **S4-3**, Presentation Layer focusing on audio coding, video coding, and image formats.

- **S4-4**, Systems focusing on interface and project management issues (across all layers).

The layer stack organized by responsible sub-group is illustrated in Figure 4.5.

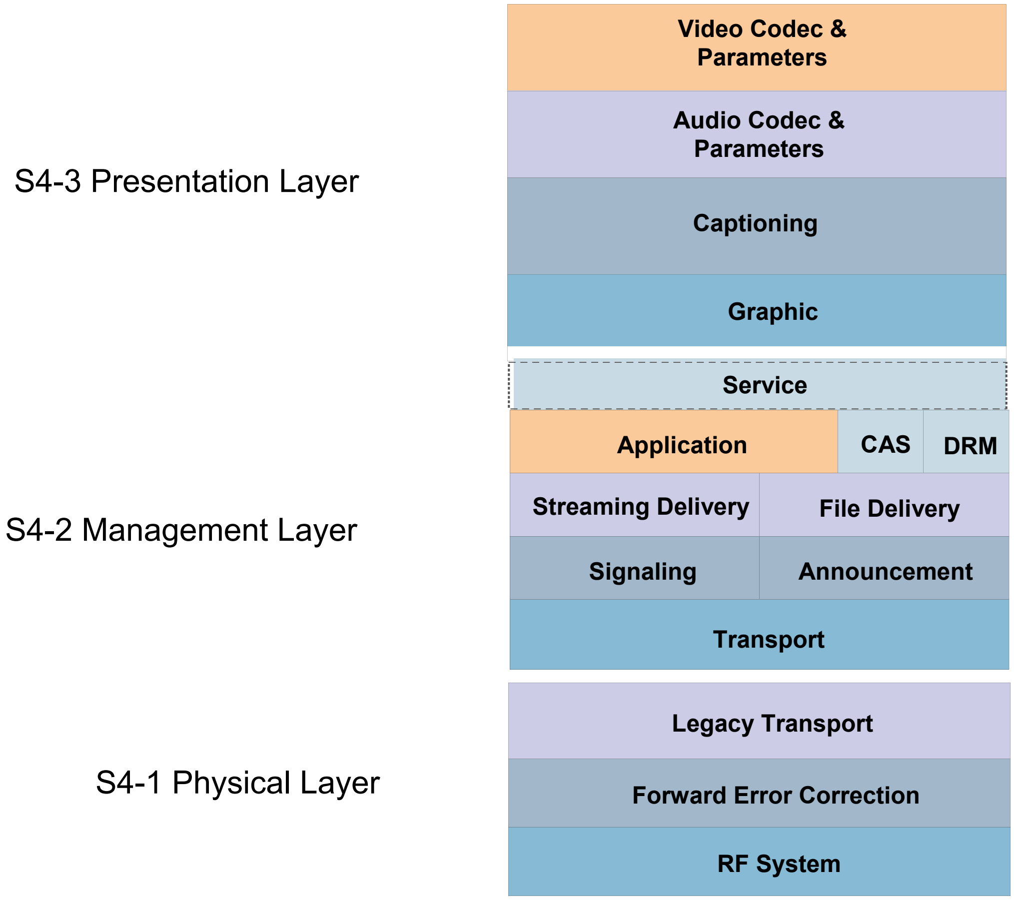

Figure 4.5 Functional overview of the ATSC Mobile DTV system organized by sub-group.

4.2 Physical Layer Overview

Central to the M/H system are additions to the physical layer of the ATSC transmission system that are easily decodable under high Doppler rate conditions. Additional training sequences and additional forward error correction (FEC) assist reception of the enhanced stream(s). These changes do not alter the emitted spectral characteristics.

ATSC Mobile DTV data is partitioned into *Ensembles*, each of which contains one or more *Services*. Each Ensemble uses an independent RS Frame (an FEC structure), and furthermore, each Ensemble may be coded to a different level of error protection depending on the application. Encoding includes FEC at both the packet and trellis levels, plus the insertion of long and regularly spaced training sequences into the data stream. Robust and reliable control data is also inserted for use by receivers. The system provides bursted transmission of the data, which allows the receiver to cycle power in the tuner and demodulator for energy saving. A simplified block diagram of the ATSC Mobile DTV transmission system at the transmitter side is illustrated in Figure 4.6.

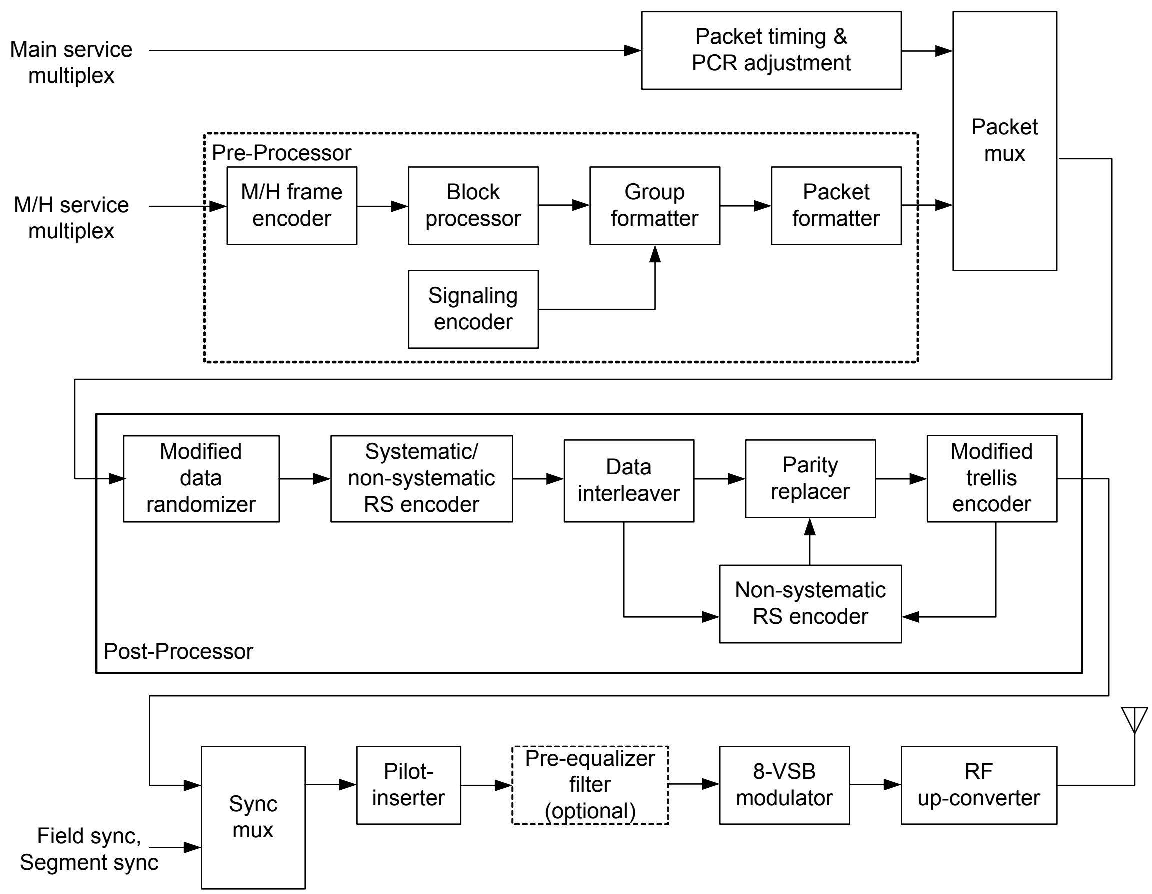

Figure 4.6 Block diagram of an ATSC Mobile DTV transmission system.

In the ATSC Mobile DTV physical layer system, the data is transferred by a time-slicing mechanism to improve the receiver's power management capacity. Each Mobile/Handheld (M/H) Frame time interval (~ 968 ms) is divided into five sub-intervals of equal length, called *Subframes*. Each Subframe is in turn divided into four sub-divisions of length 48.4 ms, the time it takes to transmit one VSB Data Frame. These VSB frame time intervals are in turn divided into four *Slots* each (for a total of 16 Slots in each Subframe).

The data to be transmitted is packaged into a set of consecutive RS Frames, where this set of RS Frames logically forms an Ensemble. The data from each RS Frame, which is transmitted during a single M/H Frame, is split up into chunks called *Groups*, and the Groups are organized into *Parades*, where a Parade consists of the Groups from up to two RS Frames but not less than one. The number of Groups belonging to a Parade is always a multiple of five, and the Groups in the Parade go into Slots that are equally divided among the Subframes of the Frame.

The RS Frame is the basic data delivery unit, into which the IP datagrams are encapsulated. While a Parade always carries a Primary RS Frame, it may carry an additional Secondary RS Frame as output of the baseband process. The number of RS Frames and the size of each RS Frame are determined by the transmission mode of the physical layer subsystem. Typically, the size of the Primary RS Frame is bigger than the size of Secondary RS Frame, when they are carried in one Parade.

The FIC is a separate data channel from the data channel delivered through RS Frames. The main purpose of the FIC is to efficiently deliver essential information for rapid Service acquisition. This information primarily includes binding information between Services and the Ensembles carrying them, plus version information for the Service Signaling Channel of each Ensemble.

4.3 Management Layer Overview

An "ATSC Mobile DTV Service" is similar in general concept to a virtual channel as defined in ATSC A/65 ("Program and System Information Protocol"). A Service is a package of IP streams transmitted through a multiplex, which forms a sequence of programs under the control of a broadcaster, which can be broadcast as part of a schedule. Typical examples of ATSC Mobile DTV Services include TV services and audio services. Collections of Services are structured into Ensembles, each of which consists of a set of consecutive RS Frames.

In general, there are two types of files that might be delivered using the methods described in the ATSC Mobile DTV system. The first of these is content files, such as music or video files. The second type of file that may be transmitted is a portion of the service guide. This includes long- and short-term keys for service protection, logos, and Session Description Protocol (SDP) files. In either case, the delivery mechanisms are the same and it is up to the terminal to resolve the purpose of the files.

A block diagram of the transport subsystem is illustrated in Figure 4.7.

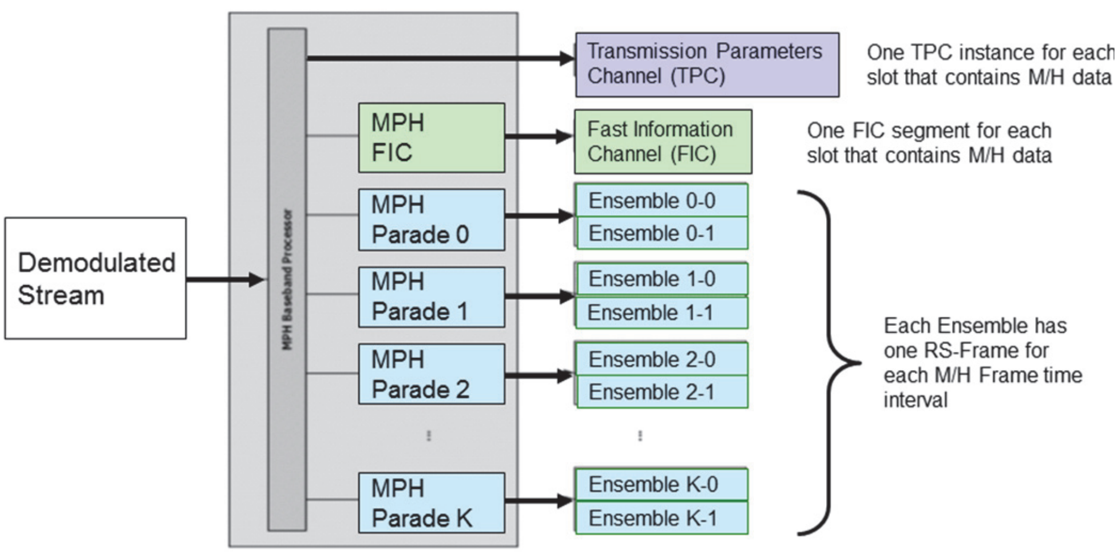

Figure 4.7 ATSC Mobile DTV physical transport system.

In the ATSC Mobile DTV system, the Services available on that system (or another system) are announced via the Announcement subsystem. Services are announced using a Service Guide. A Service Guide is a special Service that is declared in the Service Signaling subsystem. A receiver determines available Service Guides by reading the Guide Access Table. This table lists the Service Guides present in the broadcast, gives information about the service provider for each guide, and gives access information for each guide.

The ATSC Mobile DTV Service Guide is an OMA BCAST Service Guide, with certain constraints and extensions. A Service Guide is delivered using one or more IP streams. The main stream delivers the Announcement Channel, and zero or more streams are used to deliver the guide data. If separate streams are not provided, guide data is carried in the Announcement Channel stream. The Service Guide is designed so that it may also be delivered over a separate connection if a device has two-way connectivity.

Service Protection refers to the protection of content, be that files or streams, during its delivery to a receiver. Service Protection assumes no responsibility for content after it has been delivered to the receiver. It is intended for subscription management. It is an access control mechanism, only.

The ATSC Mobile DTV Service Protection system is based on the OMA BCAST DRM Profile. It consists of the following components:

- Key provisioning
- Layer 1 registration
- Long-Term Key Message (LTKM), including the use of Broadcast Rights Objects (BCROs) to deliver LTKMs
- Short-Term Key Messages (STKM)
- Traffic encryption

The system relies on the following encryption standards:

- Advanced Encryption Standard (AES)
- Secure Internet Protocol (IPsec)
- Traffic Encryption Key (TEK)

In the OMA BCAST DRM Profile there are two modes for Service Protection—interactive and broadcast-only mode. In interactive mode, the receiver supports an interaction channel to communicate with a service provider, to receive Service and/or Content Protection rights. In broadcast-only mode, the receiver does not use an interaction channel to communicate with a service provider. Requests are made by the user through some out-of-band mechanism to the service provider, such as calling a service provider phone number or accessing the service provider website.

The OMA BCAST 4-layer key hierarchy is illustrated in Figure 4.8.

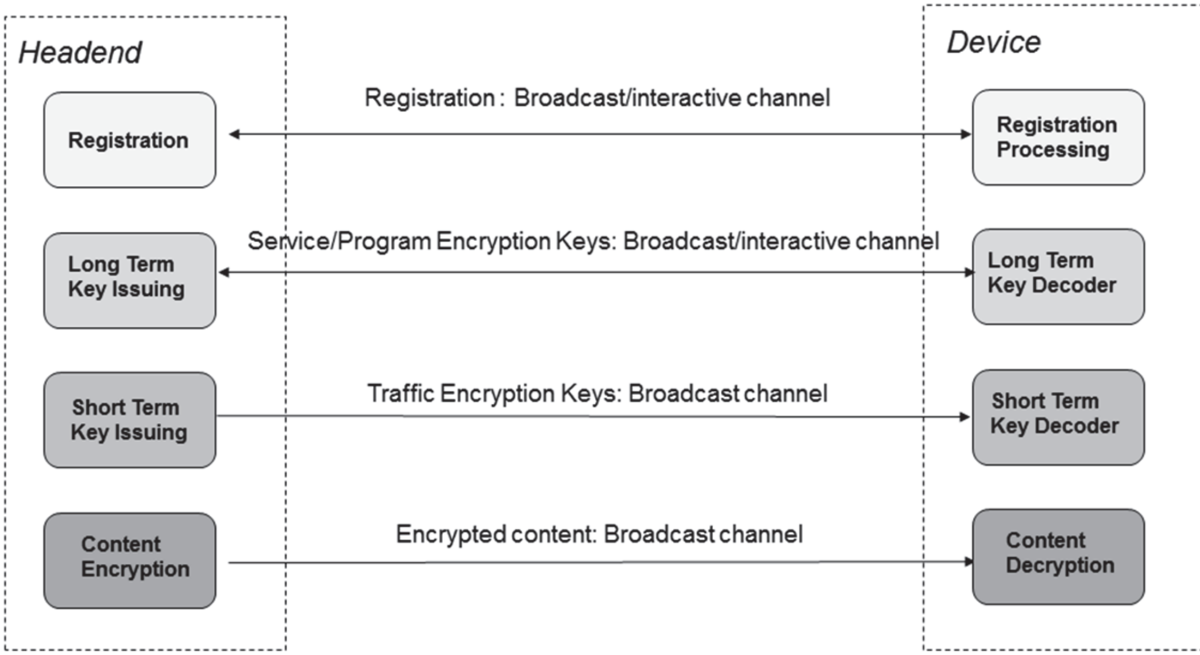

Figure 4.8 OMA BCAST 4-layer key hierarchy.

4.4 Presentation Layer Overview

The Application Framework is a toolkit that allows for the creation of graphical and syntactical elements to be added in conjunction with the delivery of audio and video. It differs from a middleware specification in that the Application Framework allows for an overlay on top of the audio/video plane rather than a true software control layer controlling all upper layers of the client. The Application Framework data is transmitted in-band alongside the audio and video.

This subsystem enables the broadcaster of the audio-visual service to author and insert supplemental content to define and control various additional elements to be used in conjunction with the audio-visual service. It enables the definition of auxiliary (graphical) components, layout for the service, transitions between layouts and composition of audio-visual components with auxiliary data components. Furthermore, it enables the broadcaster to send remote events to modify the presentation and to control the presentation timeline. The Application Framework further enables coherent rendering of the service and its layout over a variety of device classes and platforms, rendering of action buttons and input fields, and event handling and scripting associated with such buttons and fields.

The Application Framework is important because it allows the ATSC Mobile DTV system to expand beyond the mobile playback of video and audio, and provide a new set of tools for Internet-style personalization and interaction. This entertainment experience can capitalize on whatever data return channel(s) may be available.

The Application Framework is built around the OMA Rich Media Environment (OMA-RME). The OMA-RME, designed around a similar requirement set, is an umbrella standard encompassing elements in application creation, delivery, and control.

Video coding for the ATSC Mobile DTV system utilizes a layered approach that targets two types of devices:

Base Layer – AVC Baseline Profile at Level 1.3 with constraints from SVC.

Enhancement Layer – SVC Scalable Baseline Profile at Level 3.1 (optional). SVC provides an AVC-compatible base layer for increased coding efficiency. SVC may enable valuable features such as graceful service quality degradation in lossy environments and reduced channel change times.

Handheld devices support one resolution: 416 x 240, 16:9 aspect ratio (approximately). In-car devices support two enhanced resolutions: 624 x 360 and 832 x 480.

AVC is the international video coding standard ISO/IEC 14496-10 (MPEG-4 Part 10) "Advanced Video Coding," finalized in 2003. It supports a wide range of applications, from low-bit-rate mobile video streaming to high-bit-rate HDTV broadcast and DVD storage. AVC has been widely deployed.

The AVC coding toolkit of Profiles is illustrated in Figure 4.9.

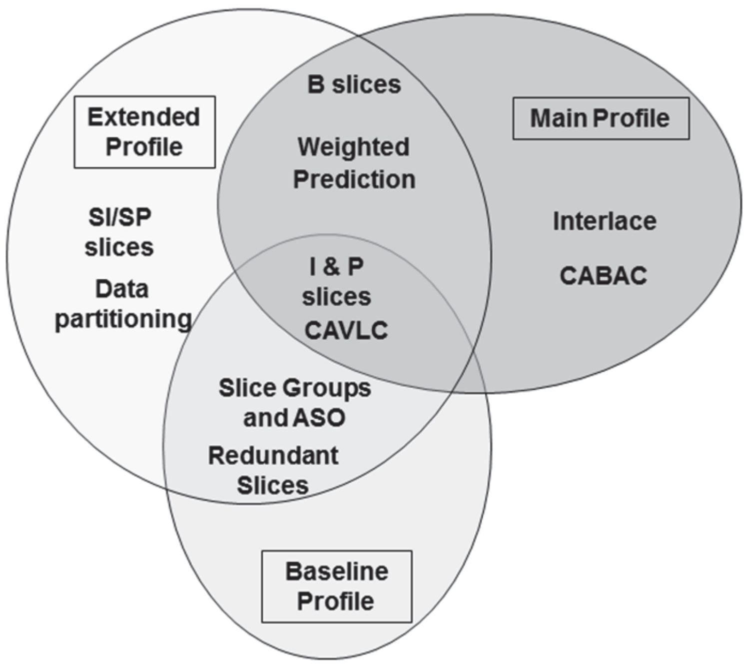

Figure 4.9 Profiles in the AVC coding standard.

Audio coding for the ATSC Mobile DTV system utilizes HE AAC v2, specified in ISO/IEC 14496-3 – MPEG-4 Audio. Constraints include the following:

- HE AAC v2 Profile, Level 2, maximum of two audio channels up to 48 kHz sampling rate.
- Bit rate and buffer requirements comply with the MPEG-4 HE AAC audio buffer model.
- Supported sampling frequencies are 32 kHz, 44.1 kHz, and 48 kHz.

HE AAC v2 is the combination of three audio coding tools, MPEG-4 AAC, Spectral Band Replication (SBR), and Parametric Stereo (PS). The HE AAC toolkit is shown in Figure 4.10.

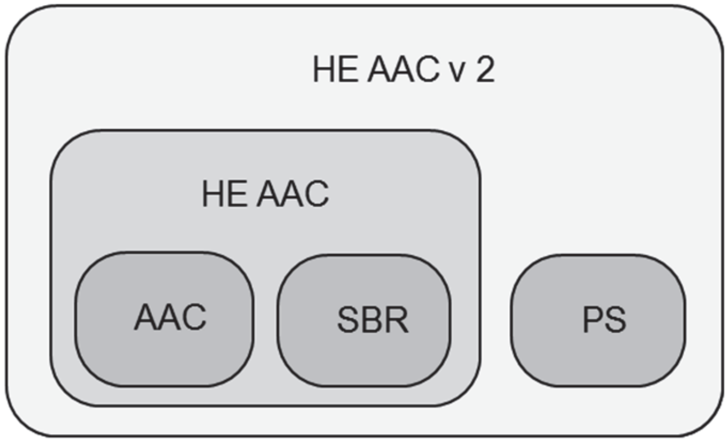

Figure 4.10 Toolkit of the HE AAC v2 system.

4.5 Systems Level Overview – Evolving the ATSC Mobile DTV System

Recognizing that technology will continue to move forward, A/153 contains mechanisms throughout its layers/subsystems to enable graceful evolution of the ATSC Mobile DTV system over time. This capability has been built into the "System Configuration Signaling" architecture. The goal of this system is to support a configuration change of the protocol stack used by the broadcaster. Key considerations included:

- The system must provide information about each piece of content and how it is transmitted
- A receiving device must be able to determine if it can support such a content before the content is exposed to the user

As designed, the signaling is multilayer and will support two types of changes:

Major Version Change: a non-backward-compatible level of change.

Minor Version Change: a backward-compatible level of change, provided the major version level remains the same. Decoders/receivers can assume a minor change does not prevent them from rendering content.

The following signaling requirements were established for the ATSC Mobile DTV system:

- Capable of signaling the addition of a new elementary subsystem. For example, a Digital Rights Management capability may be added.
- Capable of signaling the removal of an elementary subsystem. For example, service protection is removed and replaced with functionality that resides outside of the ATSC system; i.e., an out-of-band method.
- Capable of signaling the replacement of an elementary subsystem. For example, one encryption is replaced with another encryption—the black box operation is equivalent.
- Capable of signaling service compatibility in an expedient manner, where the receiver is able to determine if it can support a service within one complete frame time.
- Capable of signaling all functionality needed to support a service correctly (i.e., transport, file management, SVC sync, and so on).
- Capability to support the Service Guide not displaying an event that cannot be decoded by the receiver.

- Signaling of the elementary subsystem functions must be complete enough for the receiver to definitely determine if it can process the content.
- Capable of signaling multiple generations of service carried concurrently, in the same ATSC Mobile DTV emission.
- Capable of signaling services that are intended for the equivalent of a multicast group (target is subset of receivers grouped by activity).
- Capable of signaling legacy services with optional extensions, such that the legacy receiver ignores the optional functionality signaling, and supports the legacy portion of the service.
- Capability to change the System Configuration Signaling system protocol without adversely affecting products built to the original signaling protocol.
- Capability to support a receiver determining a channel is out of service.
- Capable of signaling that the protocol version of a single elementary subsystem has changed.
- Capable of communicating a code for the version of each elementary subsystem required to decode and correctly display the services offered.

The signaling approach is hierarchal, with the physical of RF layer being considered the bottom of the stack. Much of the signaling is defined as integral parts of the data structures. At the bottom-most layer a simple (one-bit) signaling means was established. A major change of the entire physical layer or the presence of a compatible enhancement can be signaled by use of one such bit of a number that are available, possibly concurrent with the use of other such bits signaling the presence of other mutually-compatible enhancements to the system. Other signaling for the RF layer is implemented with a simple version field in key data structures, each of which enable signaling of changes in the data structure above each.

At higher layers more signaling capability is established, reflecting the increasing likelihood of change in those layers as time progresses.

5. RF/TRANSMISSION SYSTEM CHARACTERISTICS

The RF/transmission system characteristics of the ATSC Mobile DTV Standard are defined in A/153 Part 2 [14]. M/H data is partitioned into Ensembles, each of which contains one or more services. Each Ensemble uses an independent RS Frame (an FEC structure) and, furthermore, each Ensemble may be coded to a different level of error protection depending on the application. M/H encoding includes FEC at both the packet and trellis levels, plus the insertion of long and regularly spaced training sequences into the M/H data. Robust and reliable control data is also inserted for use by M/H receivers. The M/H system provides bursted transmission of the M/H data, which allows the M/H receiver to cycle power in the tuner and demodulator for energy saving.

6. DATA TRANSPORT

6.1 Parades, Ensembles, and RS Frames

As specified in A/153 Part 2 [14], M/H data is delivered in Parades, each of which consists of one or two Ensembles, where each Ensemble consists of a sequence of RS frames with the same FEC codes. The payload of each RS Frame consists of a matrix of bytes, with 187 rows and N columns, where N depends on the transmission configuration of the ensemble to which the RS

Frame belongs. This matrix is augmented by additional rows of RS parity bytes at the bottom and then an additional 2 columns at the right containing 16-bit cyclic redundancy check sums (CRCs) for each row, as shown in Figure 6.1.

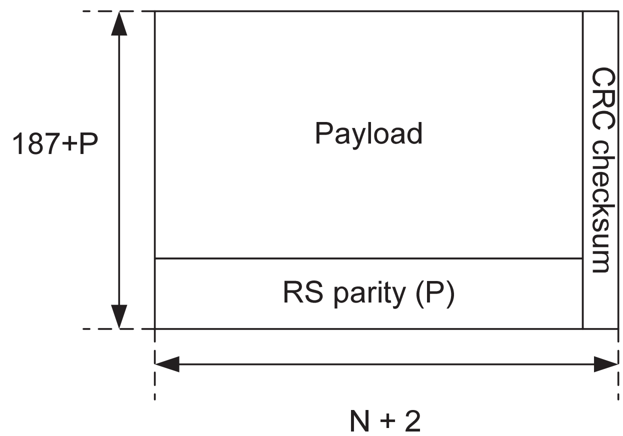

Figure 6.1. RS Frame payload byte array with augmentation.

As specified in Section 6.2 of A/153 Part 3 [2], each row of the payload of an RS Frame contains a 2-byte header with a 3-bit network_protocol field, a 1-bit error_indicator field, a 1-bit stuffing_indicator field, and an 11-bit pointer field. The network_protocol field gives the type of the network protocol packets that first appear in that row.

Version 1.0 of the ATSC M/H standard defines two types of network protocol packets in the RS Frames: IPv4 packets and "framed packets." IPv4 packets are inserted directly into the RS Frame rows. Framed packets are defined primarily for future proofing. They are not expected to be used initially. Each framed packet has a 32-bit header that gives the length and protocol of its payload, thereby accommodating many different network protocols. The codes for the protocols of the payload of framed packets are the same as those used for Ethernet packets.

The obvious advantage of allowing protocol packets that are expected to be widely used, such as IPv4, to be inserted directly into the RS Frames instead of encapsulating them in framed packets is that it saves the overhead of the 32-bit headers in the framed packets.

Later versions of the A/153 Standard might define additional types of network protocol packets that can be inserted directly into the RS Frames, such as IPv6.

While mixing the network protocol packet types in any given row is not prohibited, it is recommended that only one type be placed in one row and stuffing be used if the 'framed packets' are sent as well.

The Error Indicator field indicates whether it is known that there are error bytes in the row.

The Stuffing Indicator field indicates whether the row contains any stuffing bytes (as opposed to bytes that belong to network protocol packets).

The Pointer field gives the offset to the first byte of the first network protocol packet that starts in the row, if there is one. If the Pointer field has value 0x7FF, this indicates that there is no network protocol packet starting in the row.

The network protocol packets of each row are packed into the row end-to-end, with wrap-around from one row to the next row of the same Network Protocol type if necessary, where the next row of the same Network Protocol type may be in the same RS Frame or may be in the next RS Frame of the same ensemble.

Certain constraints are imposed on IPv4 packets, as listed in Section 6.3 of A/153 Part 3 [2].

6.2 Insertion of Data into the RS Frames

If multiple types of network protocol packets are being inserted into the RS Frames, then the module at the head end that is packing network protocol packets into RS Frames needs to be able to distinguish the different protocol types of the packets coming to it, so that it can insert them into distinct rows of the RS Frames. There is no restriction on the order in which the rows corresponding to different protocols are interleaved in an RS Frame, but it is a good idea to preserve the relative timing of the arrival of the packets at the head end, to the extent that is possible (rather than, for example, putting all the packets of one protocol at the beginning of the RS Frame and all the packets of another protocol at the end).

If it is absolutely essential for some reason to strictly preserve the arrival order of the packets between protocols, this can be done by starting a new row each time there is a change from one protocol type to another. However, this will result in decreased payload bandwidth, since partially full rows will be going out.

If packets are not arriving fast enough to fill up an RS Frame before it has to be emitted, then stuffing bytes can be used to fill partially or totally empty rows. Such stuffing bytes are inserted at the beginning of the rows, not the end. When a row contains any stuffing bytes, the Stuffing Indicator bit in the row header is required to be set to ‘1’. When this bit is set, then the value 0xFF in the first byte of the row indicates that it is the only stuffing byte in the row, the value 0xFEFF in the first two bytes of the row indicates that these are the only two stuffing bytes in the row, and any value less than 0xFEFF in the first two bytes of the row gives the number of stuffing bytes in the row.

6.3 Extraction of Data from RS Frames

The RS Frame payloads resulting from RS decoding and subsequent error checking may have some rows that contain errors, which should be identified with the Error Indicator bit in the row header. Any such rows should be discarded.

If the previous RS Frame had a network protocol packet that wrapped around at the end of the last row of that Frame (meaning the last row of a given protocol type), and if the first row of the current Frame has no error, then the receiver can simply pick up where it left off in extracting the packet.

However, if the current RS Frame is the first one to be received after selecting the current ensemble, or after the previous RS Frame was lost, or after the last row of any protocol type in the previous RS Frame was lost, then the receiver has to resynchronize with the network protocol packet boundaries. Similarly, if any row of the current RS Frame has errors, then the receiver has to resynchronize when starting to extract packets of any protocol type after the lost row.

In order to resynchronize, the receiver should scan forward until it locates a row where the Pointer field in the header has a value other than 0x7FF. It should then skip forward in that row by the number of bytes indicated by the Pointer field and start extracting a new protocol packet there. As long as no rows are lost in the current and subsequent Frames, the receiver can continue to extract protocol packets, using the length field in each packet to determine where the next packet starts (with wrap-around from row to row and RS Frame to RS Frame as needed). As soon as it encounters a missing row or RS frame, however, it has to resynchronize.

While extracting packets, the receiver should also check the Stuffing Indicator field in each row header, and skip stuffing bytes as indicated.

6.4 Services

The term M/H Service is used to designate "a package of packetized streams transmitted via an M/H Broadcast, which package is composed of a sequence of events that can be broadcast as part of a schedule." This is the same general concept as the term Virtual Channel used in the ATSC PSIP standard [15]. The primary difference is that the packetized streams are transmitted in UDP/IP datagrams, rather than in MPEG-2 transport stream packets.

Each service has a unique MH_service_id associated with it. This service ID is similar in general concept to the virtual channel number used in the ATSC PSIP Standard. The primary difference is that the range of values for an MH_service_id is split into two sub-ranges, one sub-range for local services (those broadcast only within a single broadcast area) and one sub-range for regional services (those broadcast in multiple broadcast areas). See Section 7.1.5.5 (Service Level Fields) and Annex B of A/153 Part 3 [2] for a full description of these sub-ranges.

All the components of a service are typically contained within a single Ensemble. However, there may be situations when it is desirable to have the components of a service split across multiple ensembles.

An example of a situation when a Service might be split across two Ensembles is when it is desired to deliver the basic service in an Ensemble with a high level of FEC, and an enhancement layer in a different Ensemble with a lower level of FEC. This will allow receivers in poor reception areas to receive the basic service, while receivers in good reception areas receive the enhanced service, and less bandwidth is used than if the basic and enhancement layer were both sent with a high level of FEC.

It is important to be aware of the requirement in the M/H standard that a service can only be split across multiple Ensembles if the set of Ensembles satisfies the conditions to be an "M/H Multiplex," which means that the same IP protocol version is used for all the IP packets in the set of Ensembles, and there are no IP address collisions among the Ensembles in the set.

Another "multi-ensemble" services situation can arise when it is desired to replicate a Service Guide service across all the ensembles in an M/H Broadcast. This situation is discussed in Section 8.4.

7. SIGNALING DATA DELIVERY AND USAGE

A/153 Part 2 [14] describes how to locate the signaling data structures. This section recaps the purpose and addresses the next levels of signaling.

7.1 Signaling Data Structures

7.1.1 Overview of Hierarchical Signaling Structure

The signaling structures of the M/H Standard are organized in a hierarchical fashion, with three levels:

- Transmission Parameter Channel (TPC) signaling data, at the bottom level
- Fast Information Channel (FIC) signaling data, at the middle level
- Service Signaling Channel (SSC) signaling data, at the top level

Table 7.1 summarizes the purpose and the usage of each of these three levels of signaling data.

Table 7.1 M/H Hierarchical Signaling Structures

TPC	**Purpose**	(1) Provide Parade configuration and decoding parameters. (2) Signal FIC-Chunk data version changes.
	Usage	Support physical layer demodulation and data extraction functions, and efficient upper level monitoring of changes to the FIC-Chunk data.
FIC	**Purpose**	Provide binding information between M/H Services and M/H Ensembles. Provide certain basic information about Services and Ensembles. (3) Signal SSC data version changes for each Ensemble.
	Usage	Support quick channel scans, rapid service acquisition and efficient monitoring of changes in the Service Signaling Channel.
SSC	**Purpose**	Provide access and decoding information for M/H Services. Provide certain other information about M/H Services. (3) Provide parameters for certain ancillary functions, such as Service Guide discovery, intermediate channel scans, and inter-transmitter hand-off.
	Usage	Support access and decoding of Services, display of minimal information about Services and current programming, and certain ancillary functions.

7.1.2 TPC Structure

Each slot that carries M/H data also carries an instance of the TPC data structure, which contains several categories of data:

1) Identification of the subframe and slot[1] containing it.
2) Certain high level parameters of the Parade to which the slot belongs:
   - Parade ID, which forms part of the Ensemble ID of the Ensemble(s) in the Parade.
   - Parade Repetition Cycle, which allows the receiver to tell when the next M/H Frame with data for this Parade will appear.
   - Parade Continuity Counter, which allows the receiver to tell whether an M/H Frame of this Parade has been missed or not.
3) RS modes and SCCC modes of the Parade to which the slot belongs.
4) Starting group number and number of groups in the Parade, and total number of groups used by all Parades together.
5) FIC-Chunk data version.

To allow receivers to decode the current M/H Frame efficiently, the data in the TPC instances contained in the first two subframes of each M/H Frame refer to the current M/H Frame.

To give receivers advance notice of changes, the data in categories 3 and 4 in the TPC instances contained in the last three subframes of each M/H Frame refer to the next M/H Frame. These instances also contain a “total number of groups” field referring to the current M/H Frame, since this information is needed to decode the FIC, and FIC decoding near the end of an M/H Frame can be essential to rapid service acquisition when tuning to a new RF channel.

7.1.3 FIC Structure

Each slot that contains M/H data carries an FIC Segment of length 37 bytes. The FIC data consists of FIC “Chunks” encapsulated in these segments.

---

[1] See A/153 Part 2 [14] for discussion of slot and other fundamental terms that describe the way the bearer layer redefines the meaning of certain 8-level symbols in the RF domain (as compared to A/53).

7.1.3.1 FIC-Chunk Structure

Each FIC-Chunk contains a list of all the Ensembles in the M/H Broadcast, and for each Ensemble it gives:

- Ensemble ID.
- Indicators of whether or not the Ensemble contains a GAT and/or SLT.
- List of all the Services in the Ensemble.

For each Service it gives:

- Service ID.
- Indicator of whether all the components of the Service are in this Ensemble, or whether some components are contained in other Ensembles, and if some are contained in other Ensembles, whether the components in this Ensemble are sufficient to provide a meaningful service by themselves or not. (For a multi-Ensemble service, the services containing the components in the different Ensembles all have the same value of the Service ID.)
- Status of the Service (active or inactive, "hidden" or not).
- Indicator of whether one or more critical components of the Service is protected (encrypted).

A Service is "inactive" when it is temporarily off the air. A Service is "hidden" when it is a test or proprietary service that is not intended for ordinary receivers, or is not intended to be selected for rendering, such as a Service Guide service. If an ordinary receiver attempts to render a "hidden" Service, the results are unpredictable.

The FIC-Chunk also has a "current/next" flag, indicating whether the FIC-Chunk applies to the current M/H Frame or the next M/H Frame. The following section explains how this is used.

7.1.3.2 Encapsulation of FIC-Chunks in FIC Segments

The FIC Segments each have a 2-byte header, leaving 35 bytes of payload for containing FIC-Chunk data. Therefore, for the purpose of encapsulating FIC-Chunks in the FIC Segments, each FIC-Chunk is divided into segments of size 35 bytes (except that the last segment of the FIC-Chunk might contain less than 35 bytes). There can be a maximum of 16 segments, limiting the size of a FIC-Chunk to 560 bytes.

Note that this limit on the size of a FIC-Chunk puts an implicit limit on the combined number of Ensembles and Services in an M/H Broadcast. The FIC-Chunk header takes up 5 bytes, each Ensemble header takes up 4 bytes, and each Service takes up 3 bytes. If, for example, the M/H Broadcast is configured with 12 Ensembles, then this implies that there can be no more than a total of 169 Services. As the number of Ensembles goes up, the number of Services that can be accommodated in the FEC-Chunk goes down, at a ratio of 4 fewer Services for each 3 additional Ensembles.

If the number of slots carrying M/H data in each subframe is 3 or less, then the number of Services that will fit in the FIC-Chunk is smaller, since the total number of FIC Segments in the entire M/H Frame will be less than 16. However, in this case the total bandwidth of the M/H data is also much lower, so the lower limit on the number of Services is not likely to be a problem.

The FIC Segment header indicates which FIC-Chunk segment it contains (segment number, starting from 0), and whether the segment belongs to the current or next version of the FIC-Chunk, and the total number of segments in the FIC-Chunk (last segment number). The FIC Segments carrying a FIC-Chunk should normally appear in order, but because of this labeling of the contents of each FIC Segment, the FIC-Chunk can be recovered even if the FIC Segments are received out of order.

The FIC Segment header also contains the FIC_chunk_major_protocol_version, which will be explained in Section 7.3.4.

Section 6.6.2 of A/153 Part 3 [2] contains a number of constraints and recommendations on the encapsulation of FIC-Chunks in the FIC Segments. Some of the key constraints are:

- If any segments of a FIC-Chunk are contained in an M/H Frame, then at least one complete set of the segments of that FIC-Chunk is required to be contained in the Frame; i.e., the delivery of a FIC-Chunk cannot cross M/H Frame boundaries. This does not prevent partial copies of FIC-Chunks from appearing in an M/H Frame, as long as at least one complete copy of each such FIC-Chunk appears in the Frame.
- At least one complete FIC-Chunk that signals the configuration of the next M/H Frame with valid M/H data is required to be delivered at the conclusion of each M/H /Frame.
- If it will fit, at least one complete FIC-Chunk that signals the configuration of the current M/H Frame is required to be delivered at the beginning of each M/H Frame.

See Section 7.3.4 for an explanation of how these rules apply in the presence of more than one major protocol version of the FIC-Chunk.

7.1.4 Service Signaling Channel (SSC)

The SSC is an IP multicast stream, which is used for carrying signaling tables within each ensemble. It uses the destination IP address 224.0.23.60 and the destination UDP port 4937, both of which were assigned to ATSC for this purpose by IANA (the Internet Assigned Numbers Authority, which is responsible for managing IP code points).

The signaling tables carried within the SSC have a syntax that is modeled closely on the syntax of MPEG-2 private sections, with the “generic section syntax” (long form of the section header). Each table can include up to 32 sections, with each section identified by the section_number field in the seventh byte of the section. The different tables appearing in the SSC can be distinguished from each other by the table_id field in the first byte of each section and the table_id_extension field appearing in the fourth and fifth bytes of each section. Each section of each table is carried as the payload of a single UDP/IP packet.

Each section has a current/next bit in the header, which tells whether it is part of a table that is currently in effect, or is part of a new table that will be the next version of the table to take effect. There is no requirement that any “next” versions of tables appear, but they can be useful to give receivers advance notice of upcoming changes, especially for time critical tables such as the SMT. If a “next” version of a table does appear, there is no limitation on how many times the “current” version can continue to appear before a change is made, but the next time the “current” table changes, the new “current” version is required to match the previous “next” version.

It is important to be aware that the signaling of descriptor loops in the tables of the SSC is slightly different from the usual practice in MPEG-2 tables. The length of a descriptor loop in an MPEG-2 table is usually indicated by a “descriptors length” field that gives the total number of bytes taken up by all the descriptors in the loop. A receiver knows that it has parsed all the descriptors in the loop when it has gone through the indicated number of bytes. In the ATSC M/H standard, the length of a descriptor loop in indicated by a “number of descriptors” field that gives the total number of descriptors in the loop. The receiver knows that it has parsed all the descriptors in the loop when it has gone through the indicated number of descriptors. The primary reason for this change was to reduce the number of bits required to signal the length of descriptor loops.

7.1.5 Service Map Table (SMT)

7.1.5.1 General Properties

Each ensemble contains an SMT to describe all the services appearing in that ensemble. (Strictly speaking, there is an SMT "instance" in each ensemble, with the different instances in different ensembles having different values for the table_id_extension field in the section headers, since the ensemble_id comprises 8 bits of the 16-bit table_id_extension.) The SMT may consist of one or more sections, each of which is in one IP/UDP packet. Each section has a header, followed by a loop with an entry for each service, followed by a descriptor loop for ensemble level descriptors. The entry for each service contains some fields with service level information, followed by a loop with an entry for each component of the service, followed by a descriptor loop for service level descriptors. The entry for each component of the service contains some fields with service level information, followed by a descriptor loop for component level descriptors. This hierarchical structure is illustrated in Figure 7.1.

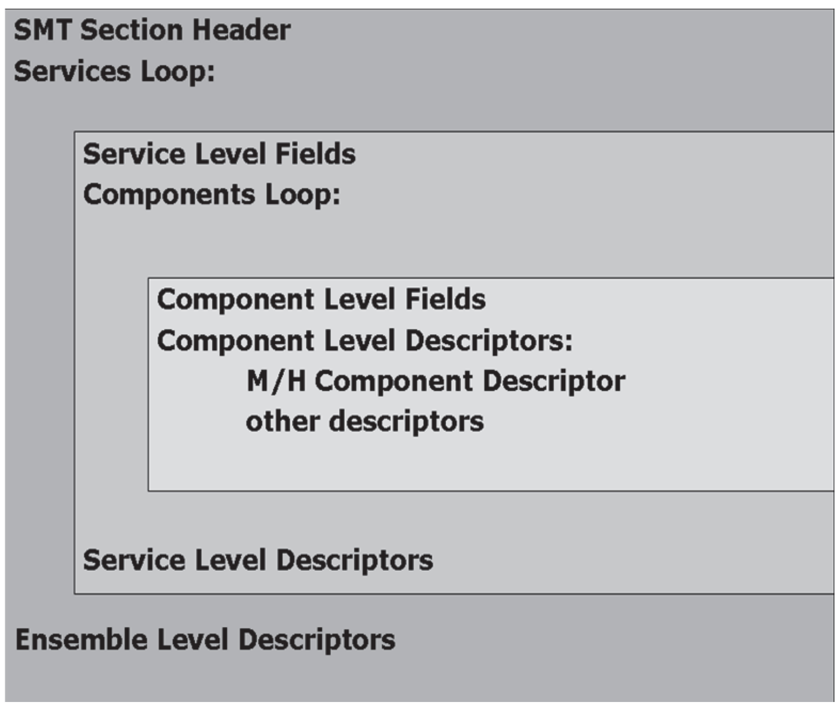

Figure 7.1 SMT hierarchical structure.

There is no restriction on how the services in an ensemble are divided up among sections of the SMT. All services can appear in one section (if they will fit, given that the size of a section is limited by the constraint of 1500 byte length MTU of IP packets in M/H), or each service can appear in a separate section, or anything in between. If different services of an ensemble come from different sources, it can be convenient for each source to generate one or more SMT sections for the services it is responsible for. Then the equipment that merges the services into the ensemble can merge these sections into the SSC intact, only ensuring that the section numbers are modified as needed to make them consistent.

The A/153 Standard states that whenever there is any change to the lineup or properties of the services in an ensemble, a new SMT reflecting the changes is required “to be inserted into the ensemble data stream of the ensemble at the point at which the changes occur.” In practice it may not be possible for the equipment generating the broadcast to ensure that the SMT appears at exactly the correct place, especially in early implementations, so receiver implementers should be alert to the possibility that the new version of the SMT may show up shortly before or shortly after the change. (Note that this is a good reason for the “next” version of the SMT to be transmitted ahead of time, so receivers know what to be watching for even if the new “current” version is a little late to appear.) However, implementers of broadcast equipment should make a strenuous effort to get the timing as close to correct as possible, since this will reduce the likelihood that receivers will experience a glitch when changes occur.

7.1.5.2 IP Address Signaling

Some of the more important fields that appear in the SMT are the destination addresses and ports where the components of the services can be found, and in some cases the optional source address of the packets as well. To provide both efficiency and flexibility, the source and destination addresses can be specified at the service level in situations where all or almost all the components share the same addresses, and they can be specified at the component level in cases where they are not specified at the service level, or where the service level default value needs to be overridden for certain individual components. The inheriting/overriding semantics as per the C language rules apply. Also, if a set of media streams with the same parameters are being delivered via the same destination address and consecutive destination ports, a single component with a port number and port count can be specified for the set, rather than having to treat them as multiple components. An example of a situation when this can be useful is a FLUTE file delivery session having multiple channels with the same destination address and consecutive port numbers.

The component level fields in the SMT include a port_num_count field and a component_destination_UDP_port_num field. The value of port_num_count gives the total number of UDP ports used by the component (whether for RTP, RTCP, etc.). As an example, a video component would require an RTP/RTCP pair (RTP Session), and therefore port_num_count would take on the value of 2. An audio stream would typically be configured the same way, RTP/RTCP and a value of 2. See Annex A for additional information including how this maps to a similar—but not identical—field in the BCAST SG. The UDP port numbers used by the component are numbered sequentially, starting from the value in the component_destination_UDP_port_num field.

When the optional source IP address field applies to a service or individual component, this indicates that source filtering is to be used for the stream(s) that make up the service or individual component. That is, the intent is that the receiver is not to process an IP packet in such a stream unless it has a source IP address that matches the specified unicast address, even if the multicast destination address and port matches those specified in the SMT.

Note that the source address is in general not reachable (in fact, it may be a private range IP address), nor is it used in the context of a Single Source Multicast network. It is provided purely for filtering purposes.

In general, it is recommended that source filtering not be used for most services, as M/H services are intended to have disjoint (destination IP address, port) combinations, and source filtering represents a small overhead to the receiver in this case.

However, the broadcaster needs to provide a source IP address for some components. In particular, a source IP address needs to be known to the receiver for any FLUTE component, including those within a service guide, as this address is part of the session ID within the FLUTE protocol.

Normally, such a service will have at most a single source IP address, specified at the service-level in the SMT. However, a mixture of A/V and FLUTE components within a service, or a combination of FLUTE components from different sources, is not ruled out by the A/153 Standard. Such cases would require component-level fields to be specified for some or all of the components; in general, a service-level source address field should not be provided if A/V components are present.

Although only IPv4 addresses are specified in this initial version of A/153, the SMT accommodates IPv6 addresses as well, in case some later version of A/153 includes them.

7.1.5.3 Supporting Future Extensibility

Two important aspects of the SMT in terms of facilitating future extensibility of the M/H standard are the service_category field semantics and the use of component descriptors to define the properties of components of a service, including a "wild card" (dynamic range) option for this descriptor.

7.1.5.4 Ensemble Level Fields and Descriptors

Currently the only substantive ensemble level field in the SMT is the ensemble_id that appears in the section headers. The primary purpose of this field is to allow a receiver to merge the SSCs of multiple ensembles together, if so desired, and still be able to identify which ensemble each SMT section refers to.

The only two ensemble level descriptors currently defined are the M/H String Mapping Descriptor and the M/H Original Service Identification Descriptor.

The purpose of the String Mapping Descriptor is to save bandwidth in situations when multiple fields in one or more of the SSC tables contain the same very long string, such as a long URL. A String Mapping Descriptor can be used to represent such strings with 1-byte string_id values, and such fields can be coded so that they can either contain the string itself or they can contain a reference to the string_id of the string.

The purpose of the M/H Original Service Identification Descriptor at the ensemble level is to identify the original service_id of a Service when it is rebroadcast with a new service_id (for example, to rebrand it with a local station brand). When it appears at the Ensemble level of an Ensemble containing a single service, it refers to that service. When it appears at the Ensemble level of an Ensemble containing more than one service, the semantics are currently undefined.

7.1.5.5 Service Level Fields

This section describes only those service level fields that have not been explained earlier, either in the explanation of the FIC-Chunk or the explanation of IP address signaling in the SMT.

For local services, the high-order 8 bits and low-order 8 bits of the MH_service_id can be treated much like major and minor channel numbers. For regional services, which are allocated on a national basis, or even an international basis, there is not necessarily any relationship between services that happen to have the same value for the high order 8 bits of the MH_service_id, so the major/minor channel number model is not useful. (See Annex B of A/153 Part 3 [2] for more details.)

The Short_MH_service_name is intended solely for display to viewers when the receiver does not have Service Guide information for the service or if there is inconsistency between the SMT and SG.

The semantics of the MH_service_category field in the SMT are slightly different from those of the corresponding service_type field in the Virtual Channel Table (VCT) of A/65 [15]. The M/H service_category field is solely intended to identify the general nature of the service, which can be used by receivers to decide which services to allow viewers to access directly (for example screening out Rights Issuer and Service Guide services, which are not intended for direct viewer consumption), or can be used by viewers to ignore services of types they are not interested in at the moment (for example ignoring audio-only programs when the viewer wants to watch a show). In particular, the service_category field does not imply any particular set of codecs or coding parameters for the service. To obtain that information, the receiver needs to parse the SMT.

The point of this is that many mobile/handheld receivers will have some mechanism for software updates, and their design should provide maximum flexibility for the introduction of new codecs in future versions of the A/153 Standard, by not having built-in assumptions about the codecs and parameters of particular categories of services. This one of many ways in which the M/H Standard has attempted to provide a high degree of extensibility.

7.1.5.6 Service Level Descriptors

The Rights Issuer Service descriptor is required for services with MH_service_category 0x03 (Rights Issuer services). It provides certain information necessary to support efficient rights issuer service carriage. More details can be found in A/153 Part 6 [4].

The optional Current Program descriptor gives the title, start time and duration of the service's current program. It is intended solely for display to the viewer when the receiver does not have up to date Service Guide information for the service or if there is inconsistency between the SMT and SG.

The optional Original Service ID descriptor gives the original service_id of a Service in situations when a service is rebroadcast with a new service_id (for example, to rebrand a regional service with a local station brand).

The Protection descriptor is required for each service that has service protection applied. It indicates the filtering option used for applying encryption and it applies to all the encrypted components of the service. (See A/153 Part 6 [4] for more details.)

The optional ATSC Content Labeling descriptor provides a unique identifier for the current content playing out on the service.

The ATSC Caption Service descriptor gives information about caption services that helps receivers present the captions properly to viewers. The A/153 Standard does not require it to be present, but it is strongly recommended that it be present whenever one or more caption services are present in the service, since without it receivers may not be able to present the captions properly to viewers.

The optional ATSC Content Advisory Descriptor provides the parental guidance ratings of current content playing out on the service.

The ATSC Genre descriptor gives the Genre classification(s) of the current content playing out on the service. It does not give Genre classification(s) for the service itself.

The content and usage of the ATSC Private Information descriptor is not specified in the M/H Standard, however it is specified in A/53 Part 3, Section 6.8.4 [27].

Several of the optional service-level descriptors provide information about the currently transmitted content for which the equivalent can be provided in a Service Guide Content fragment. This includes the Current Program, Content Labeling, Content Advisory and Genre descriptors. If both SMT and SG versions of the information are acquired by a receiver, the SMT will take precedence in the case of a mismatch.

Because the SMT is a frequently repeated table, a broadcaster that desires to conserve bandwidth may choose not to include some of the optional descriptors in the SMT, if the corresponding information is known to be present (and correct) in the SG. In doing so, the broadcaster should consider the existence of receivers that are able to make use of current content metadata but are not capable of acquiring the SG. The broadcaster should also bear in mind that acquiring the SMT is likely to be significantly faster than acquiring the broadcast SG when a receiver first tunes (or scans) to an ensemble.

7.1.5.7 Component Level Fields and Descriptors

The essential_component_indicator is the only component-level field in the SMT that is not part of the IP address signaling. This indicator simply signals whether or not the component in question is essential to rendering a meaningful service or not. If that indicator is set to ‘1’ for a component that the receiver does not know how to decode, then the receiver should not attempt to offer that service to the viewer.

Most of the information about service components is carried in the M/H Component descriptors, which identify the type and configuration parameters of each component of the service. Note that in general there may be more than one descriptor with the same component_type code, with some aspect different, such as when there is both an English and Spanish audio component of the service. They are intended to convey much the same media level information as would be conveyed in an SDP file for an Internet media application, but in a much more compact format, in order to save broadcast bandwidth.

A component descriptor contains a component_type field, a component_encryption_flag that indicates whether the component is encrypted (and if it is encrypted a list of the STKM streams that deliver the decryption keys), a text field to deliver transport parameters (to handle cases where a need for additional transport parameters is identified later), and an MH_component_data structure, the format of which depends on the encryption type.

The component_type field is based on the <fmt> field in a media (“m-“) line of an SDP file. The history of this SDP field is that originally it was used primarily for media streams delivered via RTP where the value of this field matched the value of the 7-bit payload type (PT) field in the RTP header, and IANA maintained a list of <fmt> values and their corresponding <fmt> values. However, it became clear some years ago that the number of component types in the Internet was growing so fast that it would quickly overwhelm the 7-bit field. Thus, IANA froze the list of <fmt> codes, and new media types since then have been represented via so-called “dynamic payload types” in the range from 96 to 127. With dynamic payload types the <fmt> is set to an arbitrary value in the range of 96 to 127, and an “a=rtpmap:” attribute is used in the SDP to map the <fmt> value to an actual media type.

For both static and dynamic payload types, an “a=fmtp:” attribute is used in an SDP to provide any necessary encoding parameters as a text string. The specific format of the encoding parameters depends on the media type.

In the interests of saving bandwidth, ATSC decided to assign static type codes to the coding formats of primary interest for ATSC M/H broadcasts, rather than using an “a=rtpmap:” text string, and to represent encoding parameters in a compact binary format, rather than using an "a=fmtp:" text string. Thus, the ATSC M/H Standard has a 7-bit component_type field and a set of component data fields, one for each component_type. The values to be used for the component_type field are taken from ranges that were left unassigned when IANA froze the list of <fmt> codes.

A/153 Part 3 [2] defines eight different specific component types and associated component data structures. It is expected that additional component types and their component data structures might be defined in future versions of the M/H Standard.

The M/H Standard also defines a "dynamic range" component data structure that can be used to provide support for media formats that are experimental, or directed to a niche market, or emerging in the market place but not yet incorporated into a new revision of the M/H standard, at the cost of using more signaling bandwidth than for the media formats specified in the standard. The component_type for a dynamic range component can be any value in the range 96 to 127, and the dynamic range component data structure then allows the media type and encoding parameters to be entered as text strings.

This primary motivation behind this overall approach to describing the components of a service is to provide as much future extensibility as possible.

Several of the M/H descriptors, such as the audio component descriptors and the Caption Service descriptor, each contain a field for ISO 639.2 [17] language codes.,Each of these fields carry one of the ISO 639.2-defined three-character codes for language encoded using the ISO/IEC 8859-1 [21] character set; e.g., "eng" (0x656E67). [2]

7.1.6 Service Labeling Table

The Service Labeling Table (SLT) is optional. It can be sent by a broadcaster to allow receivers to acquire limited information during a service scan about the services available in an M/H broadcast more rapidly than would be possible by acquiring the SMT of each ensemble of the M/H broadcast. The SLT provides service names and types, but not the decoding details found in the SMT, so its main purpose is to support a display to the user only. As discussed in A/153 Part 3 [2], the primary use case for which this table was designed is a medium speed channel scan that allows the service names for all channels to be presented, rather than just the major/minor numbers that can be extracted from the FIC.

It is recommended that a broadcaster only include the services intended to be made visible by the receiver's user interface in the SLT, thereby enabling selection by this label. For example, it is recommended that a broadcaster not include hidden services, service guide services, or rights issuer services.

The SLT also supports the signaling of service category, to allow this information to be indicated to the user after a medium speed channel scan. It is recommended that a broadcaster appropriately categorize each service to allow receivers to present the services as informatively as possible. In particular, a regular TV service should have the MH_service_category coded as Basic

[2] Language coding in descriptors differs from language coding in the Service Guide (SG). Section 8.1 and its subsections provide more information about language codes to be used for the descriptors and in the Service Guide. The Standard mandates the use of RCF 3066 for language coding. Additional guidance about matching of the language coding between the two systems for the SG and issues resulting from the different language coding methods that are used in the system is provided in (Section 8.1.2 and Annex A).

TV, and an audio-only service should have the MH_service_category coded as Basic Radio, rather than Unspecified.

The short_MH_service_name should be a label that brands the service. Repeating the standard-defined labels for service category in this field (e.g., Basic TV and Basic Radio) provides no additional information to a receiver.

A broadcaster can include additional service-level descriptors that allow a receiver to filter or further distinguish a service visually to the user. For example, a service with a particular theme, such as a cartoon channel or a news channel, might have a Genre descriptor in the SLT, and a service targeted to speakers of a particular language might have an ISO-639 language descriptor in the SLT.

The SLT for an M/H broadcast can be sent in more than one ensemble, but there is no advantage to doing so. A receiver can find an ensemble with the SLT during a medium speed service scan (if one exists) by inspecting the FIC as it is obtained during the scan.

It is recommended that receivers display all active, non-hidden services of category 0x01 (Basic TV), 0x02 (Basic Radio) and 0x00 (Unspecified) as these are all potentially renderable services and should be selectable by the user. If possible, the receiver should provide an indication to the user of the service category. As service category is non-normative, however, and as the SLT does not provide decoding parameter detail, the receiver cannot determine whether rendering is possible on the device until the SMT is acquired in the appropriate ensemble.

7.1.7 Guide Access Table

The Guide Access Table (GAT) lists the Service Guide (SG) data sources that contain the fragments describing the broadcaster’s M/H service offerings. As it is recommended that a minimal amount of SG data always be sent, the GAT should always be present. Furthermore as longer term service guides are recommended to be accessible via the interactive channel, it is recommended that the GAT list those SG data sources. For each SG data source the GAT gives the name of the SG data provider, plus the access parameters for the SG data source.

The access parameters are provided through MH SG bootstrap descriptors. Each such descriptor gives the type of network over which the SG data source can be accessed (the M/H Broadcast in which the GAT appears, a different M/H Broadcast, a non-M/H IP broadcast channel, or an IP-based interaction channel), and the specific access parameters, where the structure of the access parameters depends on the network type. A/153 Part 4 [3] requires that both broadcast and interactive SG data sources conform to the OMA BCAST SG specifications [23].

This means that a broadcast SG data source consists of FLUTE file delivery sessions that deliver SGDUs (containing metadata describing the programming) and SGDDs (indexes to the SGDUs). The OMA BCAST SG specification [23] requires that all the SGDDs are contained in a single FLUTE session, called the “Announcement Channel”. It is recommended that this session consist of only a single ALC/LCT channel (see Section 6.1.1 of [23]). The SGDUs can be in this same session or in other sessions.

Broadcast SG sources that appear in an M/H Broadcast are required to be in the form of M/H services (consisting of FLUTE file delivery components). Therefore, the access information for a SG data source in the M/H Broadcast containing the GAT consists of the Service ID of the service and the FLUTE Transport Session Identifier (TSI) of the Announcement Channel of the service. The access information for a SG data source in a different M/H Broadcast consists of the

Transport Stream ID (TSID) of that M/H Broadcast, together with the M/H Service ID of the service and the TSI of the Announcement Channel. The SG_bootstrap_data for these two cases provides this necessary information.

The situation is more complicated with regard to the access information for non-M/H broadcast SG sources. The problem is simplified somewhat by the fact that the SGDDs in the Announcement session not only list the SGDUs, but also give the source addresses, TSI values, and destination addresses/ports needed to access the FLUTE sessions containing them. It is also assumed that the receiver can access an IP multicast stream in the non-M/H broadcast network by simply knowing its IP address and port.

It is strongly recommended that the announcement session consists of a single FLUTE channel. In this case, then only a single IP destination address and port are needed, so the SG_bootstrap_data for the non-M/H broadcast case provides all the necessary information.

In any case, it is strongly recommended that the source IP address be provided, even though it is an optional field in the SG_bootstrap_data, since it is in general necessary to identify a FLUTE session. If the source IP address is not provided, it is essential for the SG service provider to ensure that the broadcast channel contains no other UDP datagrams that have the same destination address and port as the Announcement session, but different source IP addresses.

For a SG data source on a non-M/H interactive network, the M/H SG bootstrap descriptor provides a URL, and that is all a receiver needs to know in order to access it. The URL provided is used by receivers in the way specified for interactive Service Guide delivery in Section 5.4.3 of [23].

7.1.8 Cell Information Table

The Cell Information Table (CIT) is intended to allow a mobile device to make a relatively smooth transition when it crosses out of range of one transmitter into range of another, in the situation when the new transmitter contains essentially the same service as that to which the device is tuned. ("Essentially the same service" might mean, for example, that both stations are affiliates of the same network and are showing the same programming, although the time coordination of the two broadcasts may not be perfect, and some of the interstitial material may be different.)

To achieve the intended handoff between stations, an optional CIT instance in each Ensemble can list the services in the Ensemble and for each service give location, transmitter power, broadcast band, Ensemble ID, and service ID information for essentially the same service being broadcast by surrounding transmitters. The CIT also gives the location of the transmitter or transmitters (in the case of a single frequency network) to which the device currently is tuned. This information allows the device, especially if it has GPS capability so that it knows exactly where it is relative to its current transmitter and the surrounding transmitters, to decide which transmitter and service to switch to when the signal to which it is currently tuned becomes too weak.

Each CIT transmitter description loop contains a field named transmitter_AERP. This field is intended to carry a value that represents the maximum Effective Radiated Power (ERP) of the described transmitter, as adjusted for the height of its transmitting antenna. AERP is expressed in dBW (decibels relative to 1 watt) and can be found from one of the following formulas, depending upon the channel on which the described transmitter operates (as indicated by the PTC_num field):[3]

$$\text{for } 2 \leq \text{PTC\_num} \leq 13$$

$$\text{AERP} = \text{dBW} + 20 * \log (\text{H} / 305)$$

$$\text{for } 14 \leq \text{PTC\_num} \leq 51$$

$$\text{AERP} = \text{dBW} + 23.25 * \log (\text{H} / 305)$$

Where:
dBW = Maximum Effective Radiated Power (ERP) in decibels relative to 1 Watt
H = Height Above Average Terrain (HAAT) of transmitting antenna in meters

Receivers receiving an AERP value can determine expected signal levels from the described transmitter by treating it as having the specified AERP value as its maximum ERP and having an antenna height of 305 meters (1000 feet) above average terrain. It should be noted that, in order to determine the expected signal level in any given direction from a described transmitter, the associated AERP value must be used in combination with the transmitter_relative_pattern_depth and transmitter_null_positions fields that follow it.

7.1.9 Rating Region Table

The RRT used for ATSC M/H broadcasts is the one specified in A/65 [15] for regular ATSC broadcast. It is strongly recommended that a broadcaster not transmit an RRT if the RRT corresponding to that rating_region is fully defined in standards for that region. In particular, this applies to the RRT for rating_region 0x01. A/153 Part 3 [2] requires that the RRT instance (if present) be sent in the ensemble that carries the GAT (indicated with the GAT_ensemble_indicator in the FIC-Chunk). If no SG location is being signaled (which means that no GAT is necessary), but it is necessary to transmit an RRT, then it is recommended that an empty GAT be transmitted to allow the terminal to locate the RRT.

7.2 Timing of Signaling Data

The SMT is required by the A/153 Standard to be delivered at least once in each M/H Frame. Unless it is necessary to signal very rapid changes to service configurations, there is no reason to deliver it any more frequently than that.

[3] It should be noted that the formulas given for AERP were derived from the propagation data contained in Figures 9, 10, and 10b of Secion 73.699 of the FCC rules (47CFR73.699), using propagation statistics of F(50,90). They were optimized for a range of approximately 35 km from the transmitter, which is the region where propagation transitions from line-of-sight (LOS) to tropospheric in nature and where testing has shown that Mobile DTV signals tend to become less reliable. The height factor additions (and consequent AERP values) can be adjusted where testing of specific transmitters shows that the field strengths found at particular distances of interest would be predicted at different power levels emitted at the nominal 305 m antenna elevation.

If the SLT is provided at all, it is required by the A/153 Standard to be delivered at least once in each M/H Frame. There is no value in delivering it any more often than that.

The GAT is required by the A/153 Standard to be delivered at least once per minute.

Transmission of a RRT asserting Rating Region 1 is not required. If a broadcaster provides advisories associated with any other Rating Region value it is recommended that RRT be transmitted in the broadcast stream at least once every sixty minutes.

The A/153 Standard has no requirement on the repetition rate of the CIT. Since the CIT almost never changes, and since a device only needs this table when the user is already viewing a service and is approaching the edge of the broadcast area, there is little or no value in delivering it more often than once a minute.

7.3 Updates to Signaling Data

7.3.1 Update Propagation Mechanisms

Since each type of signaling data (TPC signaling data, FIC signaling data and M/H-SSC signaling data) has its own purposes and usages, as described in Section 7.1, each type of signaling data has its own factors that trigger updates. On the other hand, since these three types of signaling data are arranged in hierarchical manner, so that the updates in the signaling data of upper layer can be signaled by the signaling data of lower layer, updates in the upper layer signaling data propagates downwards and thus causes the updates in the lower layer signaling data. Figure 7.2 shows this update propagation while listing the factors that trigger the updates in each type of signaling data.

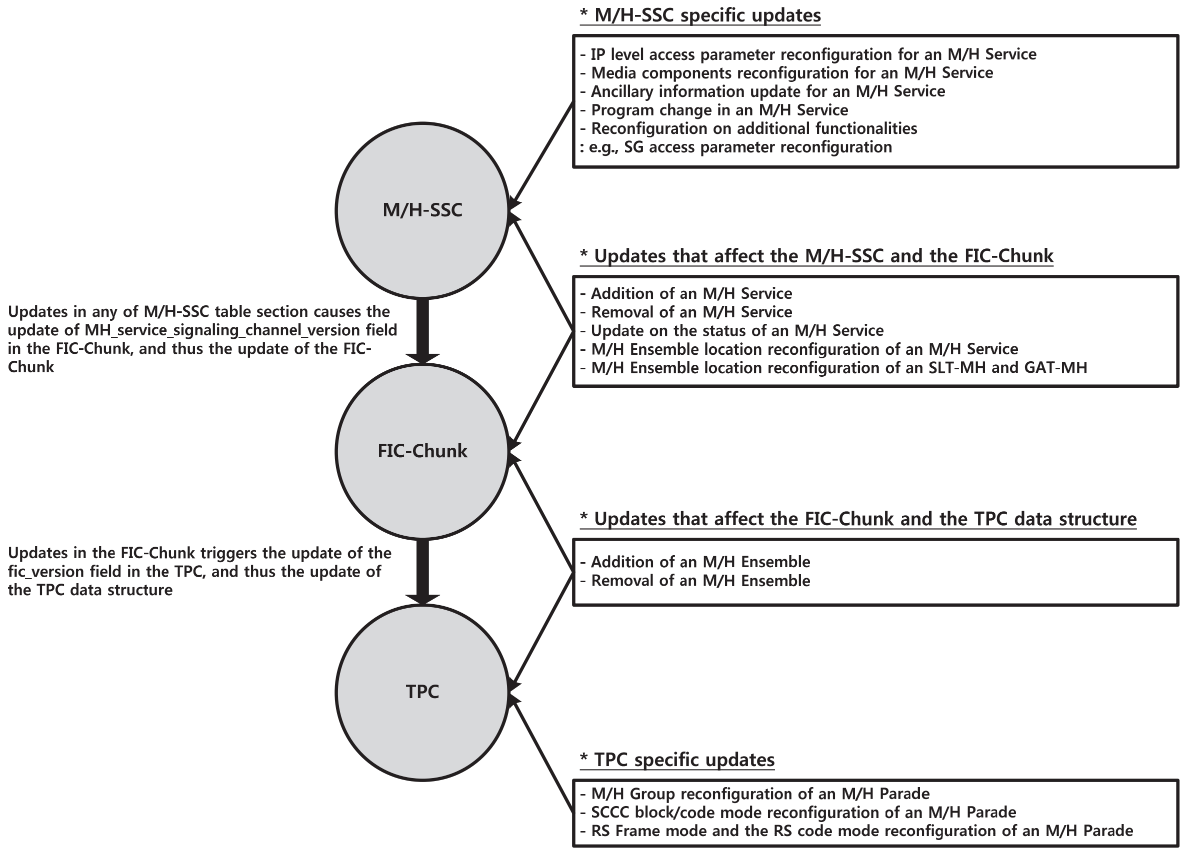

Figure 7.2 Signaling hierarchy and the factors triggering updates in the signaling data.

As shown in Figure 7.2, the system updates which are restricted to the physical layer operation only trigger updates of the TPC data structure. However, since the data version of the FIC-Chunk is signaled through the TPC data structure, any system update that causes the FIC-Chunk to update also triggers an update of the TPC data structure. Likewise, any system update that causes the M/H-SSC to update also triggers an update in the FIC-Chunk and in the TPC data structure.

Since the receiving device always monitors the TPC data structure during normal operation (in the time-slicing mode), the receiving device can always detect the update of the TPC data structure as soon as the updated TPC data structure is delivered. Similarly, since the receiving device always has access to the M/H-SSC of the M/H Ensemble containing the service that the device has selected, the receiving device can always detect updates of the M/H-SSC table sections in the scope of that M/H Ensemble. However, the tricky part is an update of the FIC-Chunk, since the receiving device does not have access to the FIC during normal operation, due to the data delivery mechanism of the FIC. Thus, to allow the receiver to pick up the updated signaling data in timely manner, it is quite important to signal the update of the FIC-Chunk before the actual update, so that the receiver device can be ready to access the FIC at the time of the updated FIC-Chunk delivery. For this purpose, it is strongly recommended to pay careful attention to the timing of the system configuration update and its signaling and to utilize the advanced signaling mechanism defined in the ATSC-M/H system.

Within an M/H Frame boundary, the TPC data structure delivered through the first and second M/H Subframe contains the signaling parameters that are applicable to the current M/H Frame, while the TPC data structure delivered through the third, fourth and fifth M/H Subframe contains the signaling parameters that are applicable to the next succeeding M/H Subframe. See A/153 Part 2 [14] Section 5.2.3 (Signaling in Advance).

The FIC-Chunk data structure need not be aligned with M/H Subframe boundaries, and the delivery units of the FIC-Chunk, the FIC-Segments, are interleaved across each M/H Subframe. The only restriction on the FIC-Chunk in terms of the M/H Frame and M/H Subframe is that a single FIC-Chunk has to be completed within a single M/H Frame. Of course, in most cases, a FIC-Chunk will fit into a single M/H Subframe, so there will be enough space to deliver both the FIC-Chunk signaling the current M/H Frame and the FIC-Chunk signaling the next M/H Frame. However, the mandatory signaling data that the FIC-Chunk is required to contain is the signaling data about the next succeeding M/H Frame. The FIC-Chunk containing the signaling data about the current M/H Frame can be delivered if and only if the FIC-Segments after packing the FIC-Chunk signaling the next M/H Frame have enough space left for the FIC-Chunk signaling the current M/H Frame.

Thus, both the TPC data structure and the FIC-Chunk data structure have the capability to signal the configuration of the next M/H Frame, as well as the current M/H Frame. With this “signaling in advance” characteristic, the example operation for the signaling data update is described with the following use case example.

7.3.2 Use Case A: Small Expansion of Ensemble and Services

1. The original configuration.
    - There is only one M/H Ensemble in the M/H Broadcast, and the NoG for the M/H Parade carrying the M/H Ensemble is one. Thus, the value of TNoG is equal to 5.
    - The number of M/H Services carried in the M/H Ensemble is 8. Thus, the size of the FIC-Chunk is small enough to fit into a single FIC-Segment, and there are 5 FIC-Chunks delivered within each M/H Frame, where the first two of them are signaling the current M/H Frame and the last three of them are signaling the next M/H Frame.
2. The broadcaster wants to add a two new M/H Services into this M/H Ensemble, and expand the size of the M/H Ensemble to accommodate them, making NoG equal to two.
    - Since the value of NoG for the M/H Parade carrying the M/H Ensemble is increased to two, the value of TNoG will be increased to 10.
    - However, since the total number of the M/H Services is 10, which results in the FIC-Chunk being too big to fit into a single FIC-Segment, the FIC-Chunk signaling the updated configuration will need two FIC-Segments for its delivery.

7.3.2.1 Recommended Operation Procedure

The configurations of three M/H Frames will be examined here, M/H Frame (x), M/H Frame (x + 1), and M/H Frame (x + 2), where M/H Frame(x) contains the old system configuration, and M/H Frame (x + 2) contains the new system configuration.

1. *The signaling configuration in M/H Frame* (*x*)

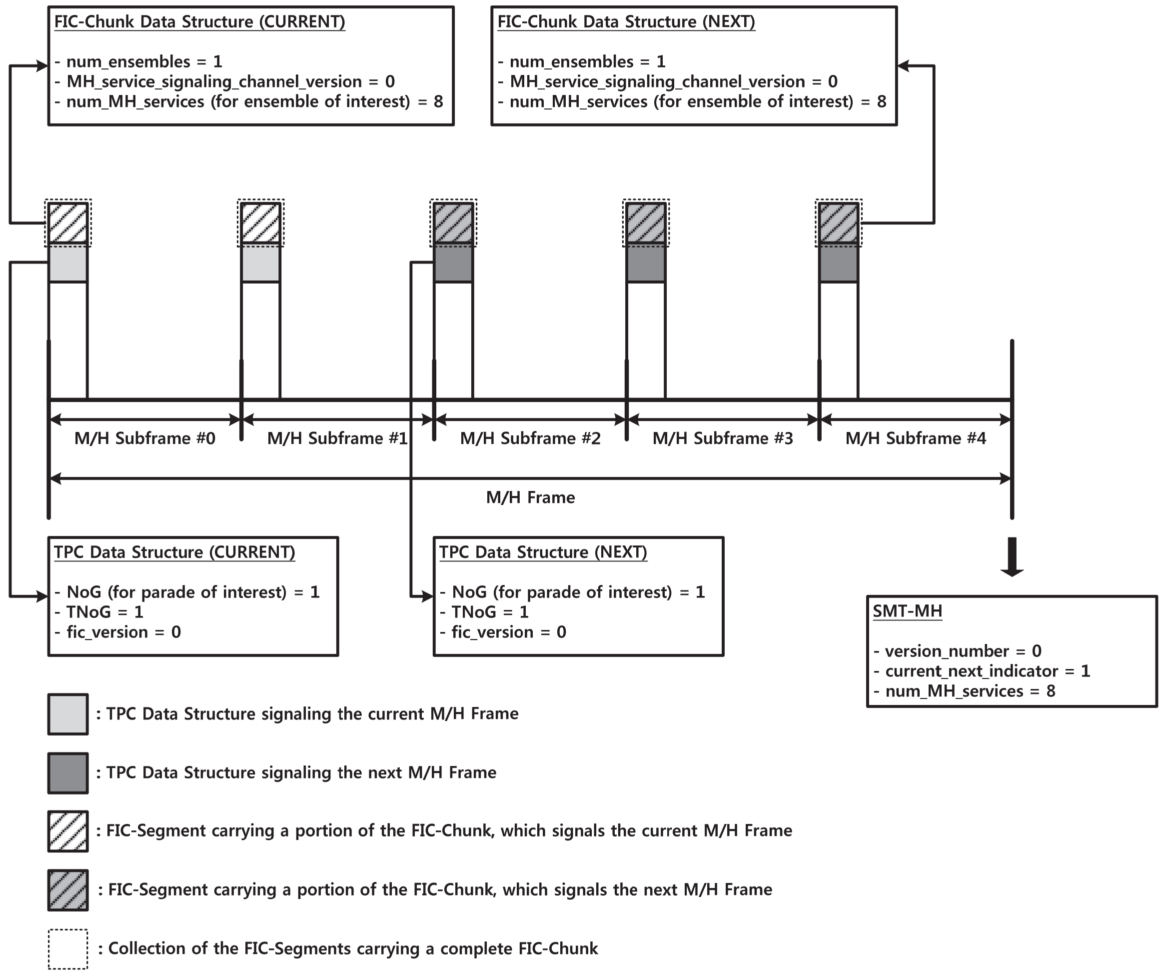

Figure 7.3 Signaling configuration in M/H Frame (*x*).

Figure 7.3 illustrates the configuration of the M/H Frame (*x*), which is two M/H Frame time before the actual system reconfiguration. As shown in the figure, the version_number of the SMT-MH in the M/H-SSC is 0, and the corresponding MH_service_signaling_channel has value of 0 also. Also, the fic_version field in the TPC Data Structure is set to 0. And note that in this M/H Frame(*x*), the FIC-Chunk is small enough to fit into a single FIC-Segment, and thus there are five FIC-Chunks delivered (two FIC-Chunks signaling the current M/H Frame(*x*) and three FIC-Chunks signaling the next M/H Frame(*x*+1)). Basically, nothing is happening yet in terms of the use case scenario.

2. *The signaling configuration in M/H Frame* (x + 1)

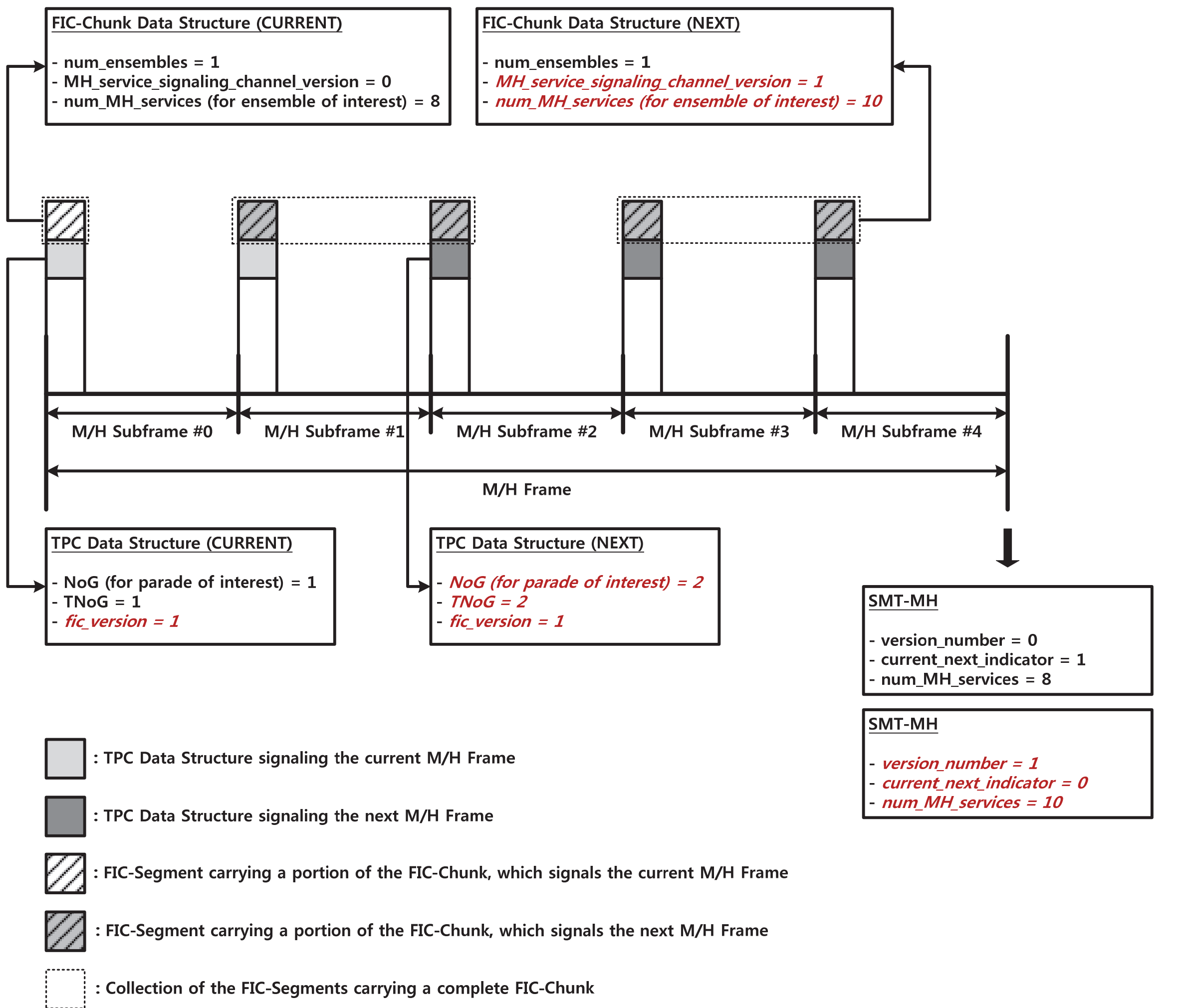

Figure 7.4 Signaling configuration in M/H Frame (x + 1).

Figure 7.4 illustrates the configuration of the M/H Frame (x + 1), which is one M/H Frame time before the actual system reconfiguration. As shown in the figure, even though the actual system reconfiguration has not yet been applied, the signaling data is starting to be updated.

- **SMT-MH**: Even though there is no actual update in the SMT-MH instance which has the current_next_indicator = '1', meaning the specific SMT-MH instance is currently applicable, another SMT-MH instance is delivered with current_next_indicator = '0', meaning the specific SMT-MH instance is not applicable yet, but will be applicable starting from the next succeeding M/H Frame. The purpose of having this upcoming SMT-MH is to enable the receiver device to get ready for the upcoming IP level system reconfiguration.
- **FIC-Chunk**: Since there is no actual update yet, the FIC-Chunk in this M/H Frame (x + 1) which contains the signaling information for the current M/H Frame (x + 1) remains the same as the FIC-Chunk of the previous M/H Frame (x), which was signaling the next M/H Frame (x + 1). However, the FIC-Chunk in this M/H Frame (x + 1) which contains the signaling information for the next M/H Frame (x + 2), is updated, since the M/H

Service reconfiguration will happen in the next M/H Frame ($x + 2$), and the data version of the M/H-SSC (MH_service_signaling_channel_version) will be updated in the next M/H Frame ($x + 2$). Note that the upcoming M/H Service reconfiguration in the M/H Frame ($x + 2$) will result in the addition to the number of M/H Services, and thus the FIC-Chunk signaling the next M/H Frame ($x + 2$), requires two FIC-Segments for its delivery as shown in the figure. As a result, the current FIC-Chunk appears only in the FIC-Segment in the M/H Subframe #0, while the other four FIC-Segments contain portions of the next FIC-Chunk, which are different FIC-Segment placements compared to the M/H Frame (x).

- **TPC**: Since the M/H Parade configuration will be updated in the next M/H Frame ($x + 2$), the TPC data structure contains the parameters of the next M/H Frame ($x + 2$) is updated, while the parameters applicable to the M/H Frame ($x + 1$) remain the same. However, the fic_version field of the TPC data structure is updated in both "CURRENT" and "NEXT" TPC data structure. The reason of having the uniform value of the fic_version across the entire M/H Frame is that the purpose of the fic_version is to notify the receiver device that there is an update in the FIC-Chunk within the M/H Frame boundary, so that the receiver device can escape from the time slicing mode and access the FIC.

3. *The signaling configuration in M/H Frame* (x + 2)

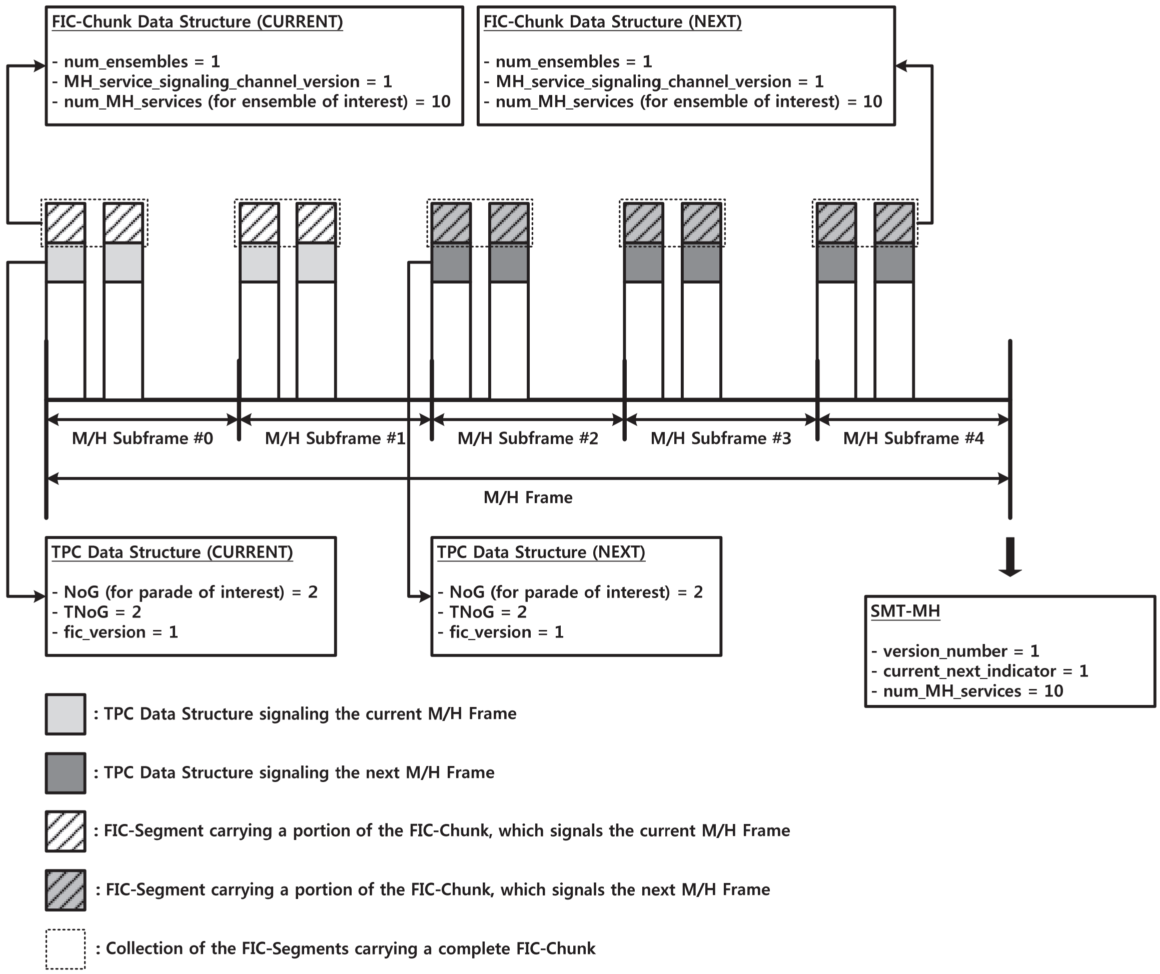

Figure 7.5 Signaling configuration in M/H Frame (x + 2).

Figure 7.5 illustrates the configuration of the M/H Frame (x + 2), where the actual system reconfiguration has happened. As shown in the figure, the signaling configuration is stabilized with updated signaling parameters, both for the current and the next M/H Frame.

The following recommendations apply to the signaling data update process:

- It is recommended to prepare and execute the signaling data ahead of the actual system reconfiguration, so that the receiver device can get ready for the reconfiguration and react in a timely way.
- It is recommended to transmit the SMT-MH instance with the current_next_indicator marked ‘0’ in the M/H Frame preceding the M/H Frame where the actual system reconfiguration happens.
- It is important to transmit the updated FIC-Chunk in the preceding M/H Frame as specified in A/153 Part 3 [2], so that the receiver device can get ready for the M/H Ensemble and M/H Service reconfiguration ahead of time.

- It is important to keep the signaling in advance rule for the TPC data structure as specified in A/153 Part 2 [14], so that the receiver device can get ready for the M/H Parade reconfiguration ahead of time.
- It is important to set the fic_version field of the TPC data structure across the entire M/H Frame, since the purpose of the fic_version is to signal the update in any FIC-Chunk within the M/H Frame boundary.

7.3.3 Use Case B: Large Expansion of Ensembles and Services

However, there might be a situation for the recommended operation described in the use case example above, when the new number of Ensembles and new number of Services are so big that the available FIC-Segments in the previous M/H Frame cannot contain the FIC-Chunk for the next M/H Frame at all. For example, in the use case example above, instead of adding two new services, if the broadcaster wants to add three new Ensembles with 55 M/H Services, then the five FIC-Segments in the M/H Frame ($x + 1$) cannot contain the FIC-Chunk for the next M/H Frame where the actual system reconfiguration happens. To handle such a case, the following is recommended.

- If the new numbers of Ensembles and Services are so big that the FIC-Chunk in M/H Frame ($x + 1$) for the next M/H Frame ($x + 2$) would not fit into the FIC-Segments present in M/H Frame ($x + 1$), then use two steps for the expansion. In the first step, in M/H Frame ($x + 2$), add the new ensembles, but do not add the new services. Put in each new ensemble only an SMT-MH with no services in it. In the second step, in M/H Frame ($x + 3$) add the new services. The ensembles that were added to M/H Frame ($x + 2$) in the first step will provide ample FIC Segments to hold the FIC chunks announcing the services in M/H Frame ($x + 3$).

7.3.4 Use Case C: Deleting Parades (and their Ensembles)

Normally the “next” fields in the TPC data for each group tell what the encoding parameters will be in the next M/H Frame for the Parade to which that group belongs (where “next M/H Frame” means the next M/H Frame which contains that Parade, which may not be the immediately succeeding M/H Frame in the case when the parade repetition rate for the Parade is greater than 1). However, the M/H Standard does not say what to put in the “next” fields of a group that belongs to a Parade which is being deleted, so that it will not be present in the next M/H Frame.

When a Parade is being deleted, the “next” FIC-Chunk in the current M/H Frame will no longer contain the data about the Ensembles belonging to that Parade. This change in the “next” FIC-Chunk will cause the fic_version field of the TPC to change, so receiving devices will know to acquire and check the FIC-Chunks in the current M/H Frame. Thus, receiving devices will learn that the Parade is being deleted, and therefore the values of the “next” fields in the TPC data for such a Parade are not important. Receiving devices are supposed to ignore all of them. It is therefore recommended to leave them at the same values they would have had if the Parade was not being deleted.

7.4 Multiple Protocol Versions of Signaling Data

A key objective of the M/H signaling subsystem is to facilitate evolution of the A/153 Standard, allowing it to accommodate future developments in technologies and business models.

This objective is met in part by the well-known MPEG-2 technique of providing length indicators for data structures such as tables, descriptors and designated portions of the FIC-Chunk, with the understanding that new fields may be added at the end of these data structures in future releases of the Standard. Thus, if an older receiver parses through all the fields it knows about, but finds it has not come to the end of the data structure as signaled by the length indicator, it can just skip over the rest. Such a receiver cannot take advantage of any new functionality enabled by the new fields, but it can continue parsing without losing its place. This enables multiple generations of receivers to use the same data structure without breaking.

However, there are certain possible changes for which such techniques do not suffice. One simple example is assigning a meaning to a reserved bit in a data structure. Such a change is backward compatible, in that legacy receivers are supposed to ignore reserved bits, regardless of their state. However, it is not forward compatible. If a newer receiver sees an instance of such a data structure that was generated by an older generation server, without any warning that it is an older version of the data structure, it is likely to interpret the reserved bit as having a meaning that could be misleading.

Another example is a situation where the basic structure of a data structure is changed so fundamentally that a legacy receiver cannot make any sense of it at all.

To provide flexibility for handling such changes, "protocol version" fields are defined for the TPC, the FIC-Chunk, the tables in the Service Signaling Channel, and the fundamental structure of an Ensemble, including its RS-Frame structure and its Service Signaling Channel location and structure. Each protocol version field has two subfields, a "major protocol version" subfield and a "minor protocol version" subfield. If a data structure changes in a backward compatible way, then the minor protocol version subfield is incremented, but the major protocol subfield remains the same. If it changes in a non-backward compatible way, then the major protocol subfield is incremented, and the minor protocol subfield is set back to 0. (The protocol version fields of the TPC, FIC-Chunk and SSC tables are contained in the respective data structures themselves. The protocol version field of the fundamental structure of an Ensemble is contained in the Ensemble entry of the FIC-Chunk.) Examples of backward compatible changes are adding a new field at the end of a data structure or assigning a meaning to one or more reserved bits. Examples of non-backward compatible changes are changing the semantics of an existing field, or fundamentally rearranging the fields of a data structure.

In order to deal successfully with this signaling approach, a receiver needs to be familiar with all the major/minor protocol versions that were released up to the time the receiver was developed (or most recently updated), since the receiver could encounter older versions of the data structures in some broadcasts. If it encounters a data structure with a major protocol version it knows about, it can proceed to parse it and use it. If the minor protocol version of the data structure is higher than the highest one the receiver knows about, then it could find it has to skip over some fields, using the length indicators. If the minor protocol version of the data structure is lower than the highest version the receiver knows about, then the receiver may have to ignore some fields that it knows about, but which were not defined at that lower protocol version.

If a receiver encounters a data structure with a major protocol version that it does not know about, then it will need to ignore it.

It is important to recognize that if a receiver encounters two instances of a data structure with two major protocol versions it knows about, it is best to parse both of them. If the data structure is an SMT, for example, the two different versions could be providing parameters for two different sets of services. If the receiver parses only one of the table instances, it could miss knowing about services the user would like to watch.

7.5 Service Scans by Receivers

For the purpose of this section, the following example broadcast configuration is used (see Table 7.2).

Table 7.2 Example ATSC-M/H Broadcast Configuration

RF Channel	TSID	M/H Service ID	Ensemble ID	Short Service Name	Service Category	IP/UDP
20	0x0019	0x1702 (23.2)	0x00	NBZ-M1	0x01	xxx.xxx.xxx.xxx/yyy
		0x1703 (23.3)	0x00	NBZ-M2	0x01	xxx.xxx.xxx.xxx/yyy
		0x1704 (23.4)	0x01	NBZ-M3	0x02	xxx.xxx.xxx.xxx/yyy
21	0x0021	0x1862 (24.2)	0x00	KU-M1	0x01	xxx.xxx.xxx.xxx/yyy
		0x1863 (24.3)	0x01	KU-SEC	0x03	xxx.xxx.xxx.xxx/yyy
		0x1864 (24.4)	0x02	GuideOne	0x08	xxx.xxx.xxx.xxx/yyy
22	0x0071	0x5102	0x00	KBE-M1	0x01	xxx.xxx.xxx.xxx/yyy
		0x5103	0x01	KBE-M2	0x01	xxx.xxx.xxx.xxx/yyy

In order to effectively navigate M/H Services, the receiving device is expected to maintain a list of 'services' that may be presented. This list may be used in service selection processes such as the Service Guide and service surfing. In the ATSC-M/H system, three different types of service scan process are possible for the receiving device:

- Fast service scan only through the FIC
- Intermediate service scan through FIC and SLT-MH
- Complete service scan through FIC and SMT-MH instance in all ensembles

7.5.1 Fast Service Scan only through FIC

The receiving device may perform the service scan process examining and processing only the signaling data contained in the FIC-Chunk of each M/H Broadcast. Since the expected processing time for the FIC-Chunk for a single M/H Broadcast is in the range of 200 milliseconds to ~ 2 seconds. (one M/H Subframe minimum, two M/H Frames maximum depending on the size of the FIC-Chunk that the receiver device needs to process), the service scan only through the FIC is the fastest way for scanning services. However, this FIC-only service scan can only provide the transport_stream_id, MH_service_id, and ensemble_id, which are not very friendly for presentation to the user. The example service map that the receiving device can build through the FIC-only scanning will be as shown in Table 7.3.

Table 7.3 Resulting Service Map Through FIC only Service Scan

RF Channel	TSID	M/H Service ID	Ensemble ID
20	0x0019	0x1702 (23.2)	0x00
		0x1703 (23.3)	0x00
		0x1704 (23.4)	0x01
21	0x0021	0x1862 (24.2)	0x00
		0x1863 (24.3)	0x01
		0x1864 (24.4)	0x02
22	0x0071	0x5102	0x00
		0x5103	0x01

Note that the M/H Services on RF channel 20 and 21 have their MH_service_id values in the range of 512d – 17919d (with a hexadecimal notation), which is the value range reserved for Local M/H Services, and thus the receiving device may present them as two-part channel decimal numbers (major.minor).

7.5.2 Intermediate Service Scan through FIC and SLT-MH

To complement the lack of information that is provided in the FIC-only service scan, while keeping the necessary processing time relatively short, the receiving device may perform the service scan examining the FIC-Chunk and the SLT-MH delivered through the MH-SSC. Since the ensemble location of the MH-SSC which contains the SLT-MH is signaled through the FIC-Chunk, the receiving device only needs to examine a single ensemble to access the SLT-MH in addition to examining the FIC-Chunk, and thus the service scan process time for a single M/H Broadcast would be in the range of 1.2 seconds to ~ 3 seconds. Table 7.4 shows the example service map that the receiving device can build through the service scan with FIC and SLT-MH on each M/H Broadcast.

Table 7.4 Resulting Service Map through FIC+SLT-MH only Service Scan

RF Channel	TSID	M/H Service ID	Ensemble ID	Short Service Name	Service Category
20	0x0019	0x1702 (23.2)	0x00	NBZ-M1	**0x01**
		0x1703 (23.3)	0x00	NBZ-M2	**0x01**
		0x1704 (23.4)	0x01	NBZ-M3	**0x02**
21	0x0021	0x1862 (24.2)	0x00	KU-M1	**0x01**
		0x1863 (24.3)	0x01	KU-SEC	0x03
		0x1864 (24.4)	0x02	GuideOne	0x08
22	0x0071	0x5102	0x00	KBE-M1	**0x01**
		0x5103	0x01	KBE-M2	**0x01**

Note that the some values of the service category information are informative only. Thus, the receiver device is recommended to defer categorizing those services until the device sees the SMT-MH entries and their component descriptors, other than to indicate that they are "radio" (audio only) or TV (video) channels. Note that the non-bold values of service_category indicate these services are used for delivery of rights issuer and service guide data, so these services would not normally be displayed to the user.

7.5.2.1 Complete Service Scan through FIC and SMT-MHs

The receiver may perform the complete service scan process for building the service map with more detailed information examining every SMT-MH on each M/H Ensemble. However, since the number of M/H Ensembles per M/H Broadcast could in extreme cases be as large as 32, the service scan process could take up to 33 seconds per M/H Broadcast. The example service map that the receiving device can build through the service scan with FIC and SMT-MHs is shown in Table 7.5.

Table 7.5 Resulting Service Map through FIC+SMT-MH only Service Scan

RF Channel	TSID	M/H Service ID	Ensemble ID	Short Service Name	Service Category	IP/UDP
20	0x0019	0x1702 (23.2)	0x00	NBZ-M1	0x01	xxx.xxx.xxx.xxx/yyy
		0x1703 (23.3)	0x00	NBZ-M2	0x01	xxx.xxx.xxx.xxx/yyy
		0x1704 (23.4)	0x01	NBZ-M3	0x02	xxx.xxx.xxx.xxx/yyy

21	0x0021	0x1862 (24.2)	0x00	KU-M1	0x01	xxx.xxx.xxx.xxx/yyy
		0x1863 (24.3)	0x01	KU-SEC	0x03	xxx.xxx.xxx.xxx/yyy
		0x1864 (24.4)	0x02	GuideOne	0x08	xxx.xxx.xxx.xxx/yyy
22	0x0071	0x5102	0x00	KBE-M1	0x01	xxx.xxx.xxx.xxx/yyy
		0x5103	0x01	KBE-M2	0x01	xxx.xxx.xxx.xxx/yyy

Note that the UDP/IP information could either be stored in the service map or be stored in the SMT-MH processing buffer and be fetched during the service acquisition process. Also note that the values of the service category which are informative only can be confirmed by examining the component descriptors of the SMT-MHs.

If the examination of component descriptors indicates that a service has essential components which the device cannot render, then display of such a service to the user should not occur and the resulting display appearance and sound, if any, would be undefined and in direct contradiction of the expressly signaled condition.

7.6 Service Acquisition by Receivers

Once a new service is selected, a service acquisition of the ATSC-M/H system is done through three steps:

1) RF tuning
2) Ensemble selection
3) M/H Service selection

The service acquisition process can be initiated by various user actions, such as following:

- Scrolling across services by pressing up/down button of the device.
- Service selection through the Service Guide.
- Punching the (two part) channel number manually.

The result of any of these actions is that the receiving device will have the service_id of the service to be acquired. The device can then access the service map obtained from a service scan and utilize the appropriate parameters for each step of the service acquisition process.

The RF tuning step is processed by triggering the RF front end utilizing the RF channel number (stored for that TSID).

The ensemble selection step includes a number of detailed functions. First, the receiving device is expected to collect FIC-Segments and parse the corresponding FIC-Chunk, whenever the device is tuning into the RF channel for service acquisition. Utilizing the information retrieved from the FIC-Chunk, the receiving device can confirm whether the ensemble configuration information stored in the service map database is valid or not and update the service map database if it is invalid. After confirming the ensemble configuration information, the receiving device can collect only the M/H Groups that belong to the M/H Ensemble of interest.

Upon the completion of the ensemble selection process, the receiving device will acquire the first RS Frame that belongs to the M/H Ensemble of interest, at the end of the M/H Frame when the ensemble selection process is completed or at the end of the next M/H Frame. When the receiving device gets the first RS Frame, the device should join the IP multicast stream of M/H-SSC and then collect the SMT-MH sections. Through the collected SMT-MH sections, the device gets the IP access information of the M/H Service of interest and the decoding parameters for the media components comprising the M/H Service of interest.

7.7 Receiver Handling of Updates to Signaling Data

As described in Section 7.3, the three types of signaling data (TPC, FIC, and M/H-SSC) are closely related and are hierarchically arranged in terms of the update mechanism and the update signaling mechanism. This section describes the expected behavior of the receiving device, when it encounters an update of each type of signaling data.

Any update of any table sections carried through the MH-SSC is signaled through two ways: 1) through the version_number field of each MH-SSC table section, 2) through the MH_service_signaling_channel_version field of the FIC-Chunk carried through the Fast Information Channel.

The purpose of having the update signaling of the MH-SSC in the FIC is to let the receiving device process the MH-SSC table sections only when there is an update in the table sections, so that the device can ignore unnecessary table section processing, and to let the receiving device notice updates of MH-SSC table sections belonging to M/H Ensembles different from the one containing the service the device has selected, so that the device can determine whether MH-SSC processing is needed or not when the device moves to a service in a different M/H Ensemble.

Whenever the ensemble configuration of the M/H Broadcast is updated, or the signaling data carried through the M/H-SSC of any M/H Ensemble is updated, the contents of the "current" FIC-Chunk are updated in the M/H Frame where the change occurs, and the contents of the "next" FIC-Chunk are updated in the M/H Frame immediately prior to the M/H Frame where the change occurs. When the receiving device is in normal time-slicing operation, only receiving the M/H Groups of the Parade containing the Ensemble containing the service being presented, the device does not have access to the FIC-Chunk, due to the fact that the FIC-Segments carrying the FIC-Chunk are spread across the M/H Groups regardless of the M/H Parade and M/H Ensemble, and interleaved over the entire M/H Subframe. That is why the TPC data structure, which is delivered in every M/H Group and therefore is always accessible by the receiver device, signals the data version of the FIC-Chunk through the fic_version field. When a receiving device sees the fic_version field of the TPC data structure incremented (modulo 32), the device should immediately exit time-slicing mode and collect the FIC-Segments for the M/H Frame, while rendering the M/H Service in parallel. The signaling data acquired from the updated "next" FIC-Chunk will give the desired information for the next M/H Frame.

The TPC data structure is unlike the other two types of signaling data (FIC and M/H-SSC) in that the receiving device is expected to keep monitoring it in normal operation (while rendering a service for presentation), so that the receiving device can be informed about Parade configuration updates or FEC mode updates for the Parade (Ensemble) containing the service that the receiver device has selected, and can be notified of updates in the FIC signaling data. Note that even the updates in the MH-SSC table section(s) are eventually signaled through the fic_version field of the TPC data structure, since the updates in any MH-SSC table section will increment the MH_service_signaling_channel_version in the FIC, which will cause the update of the fic_version field in the TPC.

If a change in Ensemble configuration includes the deletion of one or more Parades, the TPC "next" parameters for these Parades in the M/H Frame immediately preceding the deletion will be meaningless. They are to be ignored by receivers.

7.8 Time of Day

The A/153 Standard requires each ensemble to include time of day information at least every 10 minutes. The standard defines that it is to be carried using NTPv4 [22] in multicast server mode, using the IANA assigned IP address 224.0.1.1, and port number 123. Since it is being employed in a unidirectional system, not all of the fields in the NTPv4 RFC have relevance to the A/153 application. Following are some suggested settings for the various fields in the NTP packet.

LI – This field should be set according to NTPv4 Section 7.3. It may be used by the decoders to prepare for a leap second correction.

VN – This field should be set to 4 indicating NTPv4.

Mode – This field should be set to 5 to indicate broadcast mode.

Stratum – A/153 specifies signaling as though from a primary NTP server. A/153 requires that this field be initially set to 1 by the time source and not be changed by other equipment before transmission.

Maxstrat – Should be set to 16.

Poll – The standard requires an NTP packet to be sent every 10 minutes. The closest value that meets the requirements at the given resolution of this field would be 9, resulting in an 8.5 minute value.

Precision – This value should reflect the accuracy of the time of day clock.

Root Delay – Should be set to 0 and can be ignored by the decoder.

Root Dispersion – Should be set to 0 and can be ignored by the decoder.

Reference ID – Per A/153 is set to a value corresponding to a Stratum 1 NTP server and taken from Figure 12 of [22]. Use of any other value may cause receivers to ignore the NTP packet.

Reference Timestamp – Should be reset whenever a discontinuity occurs in the encoder clock. Decoders should note this and make appropriate adjustments to their internal clock.

Origin Timestamp – Should be set to 0 and can be ignored by the decoders.

Receive Timestamp – This field can be set to 0 by the encoder. It can be used by decoders at the lower packet processing level to assist upper levels in compensating for decoder processing delays.

Transmit Timestamp – Should be set to the encoder clock value.

Extension Fields – A Key Identifier and MAC are not defined in ATSC and may be omitted by encoders. Decoders may ignore them if present.

8. ANNOUNCEMENT DATA DELIVERY AND USAGE

In an ATSC-M/H system, the Service Signaling Subsystem data described in A/153 Part 3 [2] provides enough information for a receiver to display the available services to the user by name, and to decode the audio/video or data streams if the device supports the necessary capabilities. The signaling also may provide information suitable for display to the device user about the title of the current program in the service currently being watched, together with its genre and content advisory (parental guidance) rating.

However, it is recommended that a broadcaster enable a richer guide to the available content to be displayed on receiving devices, as many devices will be capable of more complex user interfaces. To do so, it is recommended that the broadcaster make use of the Announcement functionality described in A/153 Part 4 [3], which allows distribution of:

- Descriptive information about each M/H service and current program
- Descriptions and schedule for upcoming programs

This enhanced service metadata is provided using a hierarchy of Service Guide (SG) fragments (XML data structures) that together can cover current and future offerings, as defined in [23]. This collection of XML fragments is delivered in accordance with [23], with constraints and extensions as specified in A/153 Part 4 [3] and further recommended herein.

This document does not attempt to recap the parts of [23] relevant to the elements identified in A/153 Part 4 [3], as the reader is assumed to be familiar with them, but focuses on the aspects specific to M/H.

A Service Guide delivered within an M/H broadcast is transported using one or more multicast FLUTE/UDP/IP streams, as discussed in more detail below. Such a Service Guide is a special M/H Service that is declared in the Service Signaling subsystem, but will not typically be displayed as a service to the user of the receiving device, as the application on the receiver responsible for rendering the SG will not usually consider such M/H services as being “tunable”.

8.1 Language Treatment in M/H Announcements and Signaling

This section introduces how languages are handled in the A/153 Standard. There are two ways to code languages used by A/153, and both need to be accommodated and resolved. An important piece of this is the accurate selection of audio tracks signaled in the SMT from audio track options announced in the SG. Caption selections from the SG, in a similar manner, are expected to result in selection among caption services, which are found in the video stream (not in a text stream as announced).

In general, receiving devices should consider the ISO 639.1 [16] (alpha-2, or 2-character) language code equivalent to the corresponding ISO 639.2 [17] (alpha-3, or 3-character) code, and vice-versa. Compliant SG data will use an alpha-2 code, and compliant signaling fields (in various places in the SMT) will use the equivalent alpha-3 code, as described above. See http://www.loc.gov/standards/iso639-2/php/code_list.php for the mapping between alpha-3 and alpha-2 codes. Section 8.2.4.1 provides more complete information and processing recommendations and Annex A provides detailed examples.

8.1.1 Language Treatment in A/153 Part 4 (SG data)

A/153 Part 4 [3] inherits the OMA structure for language. A/153 Part 4 puts no additional constraints on the OMA-defined strings for either the elements themselves or the OMA LanguageSDPTag.

The OMA BCAST SG specification has AudioLanguage and TextLanguage elements (in both the Service and Content fragments), to describe the language of audio and for ATSC caption tracks embedded in the video. Both of these elements have a string value, which is freeform text, and a required LanguageSDPTag attribute, which is constrained to have the syntax of a language tag as specified in IETF RFC 3066 [18].

> Note that OMA BCAST SG uses RFC 3066, although it was replaced by RFC 4646 [19] and then by RFC 5646 [20]. The specifications described below are still the same in the newer RFCs.

Multiple instances of such elements may appear, with the values of the elements in different languages, labeled by the built-in xml:lang attribute. (For example, one instance of an AudioLanguage element for a French film could have the value “French,” while another could have the value “Français,” while another could have the value “Französisch”). However, in such a situation the value of the LanguageSDPTag is required to be the same for all elements.

RFC 3066 [18] states that a language tag is a Primary subtag followed optionally by additional information, and it places the following restrictions on the Primary subtag: A Primary subtag is a string of from 1 to 8 alpha characters. If a Primary subtag is a 3-letter subtag, it is interpreted according to the 3-letter assignments found in ISO 639 Part 2 [17], or assignments subsequently made by the ISO 639 Part 2 maintenance agency or governing standardization bodies. If it is a 2-letter subtag, it is interpreted according to the 2-letter assignments found In ISO 639 Part 1 [16], or assignments subsequently made by the ISO 639 Part 1 maintenance agency or governing standardization bodies. An ISO language tag is to be used for any language that has one. If it has both a 2-letter tag and a 3-letter tag, then the 2-letter tag is required to be used in the SG transmission, and is expected by the SG processing software in the receiver. The associated freeform text string (AudioLanguage for audio) is expected to be shown to the user, as the codes are not optimal for display.

8.1.2 Language Treatment in A/153 Part 3 (SMT Component Descriptors)

A/153 Part 3 [2] has language specified as a field in the HE AAC v2 Component Data structure, the Dynamic Range Component Data structure, and the Caption Service Descriptor. It requires that these language designation fields contain a 3-character code, in conformance with ISO 639.2, or assignments subsequently made by the ISO 639 Part 2 maintenance agency. This is the same format as in the ATSC PSIP Standard [15].The current list of RFC 4646 language tags can be found at http://www.iana.org/assignments/language-subtag-registry. If a language has both a “Bibliographic” (B) and “Terminology” (T) ISO-639-2 code, the “B” code is required.

In the case of no language present (music only), filling each of the three byte values with 0x00 is necessary since in the semantic definition of this field, A/153 Part 3 [2] states “In case of no language specified for this audio stream component, each byte shall have the value 0x00.” The ISO 639.2 standard’s code for signaling “no linguistic content” was not the option selected by A/153.

Content that was originally authored for ‘web’ usage may contain language codes per RFC 3066 [18] or its successors (RFC 4646 [19] and RFC 5646 [20]). As described in Section 8.1.1, these define a more complex language coding method consisting of a primary tag with the capability to add additional qualifiers. For example, both ‘en’ and “en-US” are valid per RFC 4646 and both of these (and all other ‘en’ variants) should be converted to ‘eng’ for insertion into fields that carry ISO 639.2 codes SMT

Stations should ensure that their systems support assigning three character codes to the descriptors that are the closest to the codes in the input objects for all languages of interest to the broadcaster and identify the presence of any language code without a defined mapping for operator resolution. ISO 639.2 [17] also contains the code “und”, meaning undetermined. The broadcast station equipment should be programmed to use “und” when the language is unknown. The ISO 639.2 code “mul” for "multiple languages" should be used in appropriate cases.

8.1.3 Mapping Language Tracks Between SMT and SG

Because OMA BCAST uses 2-letter and 3-letter codes per RFC 3066 , and A/153 Part 3-defined SMT exclusively uses 3-letter codes from ISO 639 Part 2, the broadcast system can be expected to be transmitting both kinds of codes for the same M/H Service. For example, any English-language program will have “en” in the SG and “eng” in the SMT in the compliant broadcast (see Annex A for more examples). The trivial case when they are not different is when a language has no alpha-2 code, as compliant SG data is then required to contain the appropriate alpha-3 code and will match the data field in the compliant SMT data exactly. This trivial case is not expected often in the countries where the M/H standard is likely to be used and needs no further explanation.

The objective of this part of the system design is to enable the receivers to be able to present the most up-to-date information about the languages available with this service. It was recognized that, practically speaking, updates of the SG might involve service providers to whom broadcasters send updates, with varying update latencies to incorporate those updates and return the SG fragments to be broadcast. Further, some receivers might not have SG data before tuning to an RF channel, and therefore some data is needed to aid a tuning decision. The SMT data can be acquired as soon as any data is acquired, enabling rapid presentation of a limited amount of information about available services, and to enable immediate selection of audio/caption services based on stored language preferences receivers may enable. The A/153 Standard reflects the judgment that signaling of language in the SMT can be most easily be directly maintained by the broadcaster and therefore will be the most accurate data; hence it establishes that in the event of conflict the SMT data takes precedence. It is expected that receivers will perform this check when a program is being viewed, so that if a viewer calls up the Service Guide, the language information for the current program in the Service Guide always matches what is actually present in the program.

There are two mapping functions that need to occur. The first is locating the audio/caption service that is signaled in the SMT based on a used making a selection while accessing the SG presentation. The other is updating the SG presentation of the current program descriptions using data from the SMT signaling.

A change to the audio/caption tracks in a service may not have propagated to the SG at the time a consumer accesses the SG. Therefore, for the most accurate real-time presentation of language choices to the user, it is expected that the receiver will compare the SMT codes in all programs in the currently available ensembles and modify (or create) the text strings for each service with a language string if the codes are found to be different. Then the values in the SMT are expected to result in logical replacement of the corresponding values for a given program in the SG data. However, as the SG's language coding usually will be different, a simple code value replacement cannot be expected to work in most cases. Annex A contains this and other mapping relationships. Caption language, for example, should be treated similarly.

8.2 ATSC-M/H-Specific Scoping of OMA BCAST SG

A/153 Part 4 [3] specifically lists a number of fragments of the OMA BCAST SG [23], and a number of their attributes and elements, that are to be supported in ATSC M/H broadcasts. Although broadcasters are not prohibited from transmitting an M/H SG service that includes attributes and/or elements that are not explicitly mentioned in the standard, this is not recommended. Broadcasters should not expect that all or any receivers will pay any attention to the additional information, even if it is valid per the OMA BCAST SG specification

8.3 SG Delivery

This section discusses the OMA BCAST facilities for SG delivery, and how certain aspects of the M/H environment interact with these facilities.

8.3.1 Delivery Protocol Overview

The ATSC Mobile DTV Standard requires that any SG service comply with the delivery specifications in OMA BCAST SG [23]. Familiarity with this specification is assumed here, but a few concepts are worth reviewing.

OMA BCAST broadcast SG delivery involves transmission of two types of files, using the FLUTE protocol [24] over multicast IP:

- XML files containing Service Guide Delivery Descriptor (SGDD) entries, a kind of table of contents to the guide data.
- Container files known as Service Guide Delivery Units (SGDUs) that transport XML fragments that provide the guide data itself.

The SG may also announce the delivery of auxiliary data, such as media files for user interfaces or interactivity. The only such data explicitly identified in A/153 Part 4 [3] is preview data to be used in specific ways (service icons, channel change interstitials) by the receiver.

There can be one or more FLUTE sessions involved in this process. The SGDDs are all delivered in a single FLUTE session, called the SG Announcement Channel, which consists of a single UDP/IP stream. The SGDUs can be delivered in the same FLUTE session or in one or more separate FLUTE sessions. The parameters of the FLUTE sessions delivering the SGDUs are described in the SGDDs. Auxiliary data files may be delivered in additional FLUTE sessions, which are described in Access fragments within the SGDUs.

The XML fragments acquired by any individual receiver may be only a subset of those transmitted, since certain kinds of filtering by the receiver are possible, as described below.

The details of SGDD processing follow the procedures described in OMA BCAST SG [23], subject to the constraints of A/153 Part 4 [3]. A receiver uses the SGDD entries to determine which SGDUs contain the fragments it wishes to acquire. It then uses the transport parameters provided by the SGDDs to acquire the SGDUs from the broadcast.

Each SGDU is a binary container wrapped around a set of XML fragments that provide the SG data itself. The details of SGDU processing follow the procedures described in OMA BCAST SG [23], subject to the constraints of A/153 Part 4 [3]. In particular, the constraints of Annex A of A/153 Part 4 [3] are required to be applied as appropriate, to allow the appropriate mappings to be made between signaling and SG elements.

The SGDUs can be transmitted in compressed form using gzip. When this is done, the Content-Encoding attribute of the SGDUs in the FLUTE FDT is set to “gzip”.

OMA BCAST also provides specifications for interactive SG delivery over a 2-way IP link. Such interactive SG delivery also uses SGDDs and SGDUs, typically requested via an interactive query language based on the grouping criteria.

8.3.2 SG Data Grouping and Filtering

The SGDD entries can include optional GroupingDescriptor elements that allow the receiver to decide whether to acquire and process the guide data within a particular SGDU. The fragments delivered in the SGDUs can be grouped by time window, genre, service, and/or service provider. Since SGDUs can be delivered in different FLUTE sessions, different SGDUs can be given different delivery speeds. For example, SGDUs containing fragments describing current and near-term programming can be cycled at a more rapid rate than SGDUs containing fragments describing programming farther in the future.

The subset of OMA BCAST SG elements identified in A/153 Part 4 [3] includes the ability to associate some SGDUs with a specific Broadcast Subscription Management identifier (BSMSelector element in an SGDD), identifying a specific service provider for which that portion of the guide is relevant. This is intended for situations when the broadcaster knows that a receiver device population exists for which that service provider code is provisioned, and wishes to support a “private” guide, since a generic receiver is expected to ignore such fragments.

There are several reasons why a broadcaster might take advantage of the fact that SGDUs can be transmitted on separate FLUTE sessions. First, SGDD entries describe the fragments in the SGDUs, which allow receivers to make some filtering decisions about the SGDUs to be acquired. Separate sessions allow a receiver to easily avoid processing of unneeded SGDU container files. Second, separate sessions make bandwidth allocation and control of acquisition time simpler on the transmission side.

Finally, some of the auxiliary data in an OMA BCAST SG is described in terms of separate delivery sessions, in large part because this data can be substantial. Image, animation or video files referred to by PreviewData fragments are delivered in FLUTE sessions that are described via Access fragments. Although not strictly required, it is recommended that these sessions are distinct from the SGDD and SGDU delivery session(s), to allow independent acquisition of files from these sessions and rapid acquisition of the core SG.

8.3.3 M/H-Specific Aspects of SG Delivery

There are a number of unique characteristics of the ATSC M/H environment, which affect various aspects of SG delivery:

- In any given viewing area the receiving devices have access to a number of M/H broadcast streams, produced by independent broadcasters. Each of these broadcast streams contains one or more services.
- Nearly all M/H broadcast streams carry SG data for the services in that broadcast stream, but most broadcast streams do not carry SG data for services in other broadcast streams.
- Each M/H broadcast stream is divided into a set of “ensembles”. Some receiving devices will only be able to receive data from one ensemble within one broadcast stream at a time. Thus, if a user of such a device is watching a TV service in one ensemble of one broadcast stream, the device will not be able to acquire SG data from any other broadcast stream, or from any other ensemble in that broadcast stream, at the same time.
- Some receiving devices can acquire SG data from the broadcast streams during times when they are not being used for TV watching. Others cannot. Even those that can acquire SG data during such “off” times might only be able to do so when they are connected to a charger, to avoid excessive battery drain.
- Many receiving devices have a two-way IP connection that provides access to an interactive SG source, but some will not have access to such an interactive channel.
- Many receiving devices have access to a non-M/H broadcast channel, which can be used to acquire SG data in broadcast mode, whether or not the device is being used to watch M/H services at the same time, but some will not have access to such a broadcast channel.
- An M/H broadcast has a special mechanism for signaling how SG(s) are to be acquired, using the Guide Access Table for ATSC-M/H (GAT-MH). This table lists SG data sources relevant to that M/H Broadcast, and gives access information for each source. This includes SG sources in the broadcast stream itself, SG sources in other M/H broadcast streams, SG sources in non-M/H broadcast streams, and interactive SG sources. The GAT is required to list all SG sources in the M/H broadcast stream itself. Listing other SG sources is optional. See A/153 Part 3 Section 7.4 [3] and Section 7.1.7 of this document for more details.

- Each SG service in an M/H broadcast stream is listed in the SMT, along with the other types of services. This provides an alternative way to discover those SG services that are contained in the broadcast stream itself. The SMT entry for an SG service is only required to provide the access parameters for the FLUTE Announcement Session. Providing access parameters for any other FLUTE sessions in the SG service is optional (since the SGDDs in the Announcement session can provide the access parameters of the other sessions).
- Some of the information in an OMA SG overlaps with information broadcast in the M/H signaling tables. The ATSC M/H standard dictates that the data provided via signaling take priority over the SG data, as it is likely to be more up to date. A receiver is thus expected to deal with any conflict that might arise between the two.
- When the SG is being delivered over M/H transport, the announcement of an interstitial image (an image to be displayed during the delay when switching from one service to another) is slightly different from the usual OMA BCAST SG practice. To provide this functionality in an OMA BCAST SG using only those elements identified in A/153 Part 4, the Service fragment of the “switched to” service needs to include a PreviewDataReference element, with its “idRef” attribute set to the id of a PreviewData fragment representing the image and its “usage” attribute set to “1” (representing “Service-by-Service Switching”). The “validTo” attribute of the PreviewData fragment should not be sent (to save a little bit rate); but if sent it should be set to a value well into the future, which informs the receiver that the intent is for the image to be cached until needed for use when switching to that service. The OMA BCAST Standard [23] establishes the meaning of ‘validTo’ as: “The last moment when this fragment is valid. If not given, the validity is assumed to end in undefined time in the future.” When the SG is being delivered over an interactive channel, the same method of announcing an interstitial image should be used.

8.3.4 Announcement of Services in Other Broadcasts

Annex A of A/153 Part 4 [3] requires that the globalServiceID attribute of a Service fragment encode the TSID of the M/H broadcast and the M/H Service ID of the service. If a broadcast SG is announcing services only in its own M/H broadcast, the TSID will be its own.

However, an SG can include metadata for services in M/H broadcasts other than its own. In this case, the globalServiceID attribute of each Service fragment is required to match the TSID of the broadcast in which the service is available.

> Note that a regional service is indicated in the SG with no TSID (the corresponding part of the globalServiceID is set to “0”), which means that it is not possible for a receiver to know *a priori* which M/H broadcast the metadata refers to. Broadcasters announcing a regional service have the responsibility to insure that such a service is identical in every broadcast and the receiver is presented with the same metadata for any instance of a given regional service it discovers in the signaling.

8.4 SG Delivery Recommendations

8.4.1 In-Band SG Data Delivery

It is valuable for a broadcaster to ensure that while a receiver is viewing any of the broadcaster’s services, it has access to service guide information on at least the current program and the next program for all of the broadcaster's services. Because of the limitations on some receivers described in Section 8.3.3, this leads to the following:

- **Strong recommendation**: Provide in every ensemble of the M/H broadcast current/next SG information for all the services in the broadcast.

This can be done by delivering such SG data as a set of identical SG services, one per ensemble, each of which has the same short name, source IP and TSI in the SMT, and which has the same provider name and Announcement session TSI in the GAT.

For broadcasters who want to provide more SG data in-band, the following two approaches are recommended, each with advantages and disadvantages;

1) Provide at least 24 hours worth of SG data in a single ensemble of the M/H broadcast, transmitted at a rate that meets the broadcaster's bandwidth target for such data.
2) Provide at least 24 hours worth of SG data in each ensemble of the M/H broadcast, transmitted at a rate that in the aggregate meets the broadcaster's bandwidth target for such data.

With approach 1, any receiver that can spend several hours a day scanning all the area broadcasts for SG data (e.g., during times when it is connected to a charger and is not being used to watch TV) will be able to cache enough SG data from such a daily scan that it can always display information about the current programming on the services of all broadcasters who follow this delivery strategy.

With approach 2, any receiver that watches any of the broadcaster's programming for a significant amount of time each day will be able to cache enough SG data for that broadcaster that it can always display information about the current programming on that broadcaster's services.

8.4.1.1 Multiple Broadcast Guides

The A/153 Standard allows the announcement of multiple SGs in an M/H broadcast. Although each SG has a provider name, the standard provides no specific semantics for selection among the multiple guides, and so a receiver is expected to acquire all the signaled SGs. At the very least, a receiver must acquire the SGDDs (possibly subsequently filtering out acquisition of some or all of the SGDUs). It is therefore recommended that multiple SGs not be delivered within the same broadcast unless the guides are being generated separately for some other reason.

8.4.1.2 Interactive SG Data Delivery

In order to optimize the tradeoff between access to information, bit rate allocations and rapid tuning from a SG display, use of the interaction channel is strongly recommended (in conjunction with the duplication of the SG fragments in each ensemble covered in Section 8.4.1).

If a broadcaster has arranged for SG data to be available from an interactive channel, it is critical to include the information about the interaction channel in the broadcast stream so that receivers can find it. This is a departure from SG operations as expected in other environments which expect acquisition of guide first from some out-of-band source, then tune. In ATSC, it is tune first (using the signaling acquired during a frequency scan), then acquire guides so next time it is a better experience.

The A/153 Standard allows the availability of guides (or portions of guides) over the interaction channel to be announced in two ways:

- The GAT includes an entry with a bootstrap descriptor of interaction channel type

- The broadcast SGDDs include an AlternativeAccessURL element that provides the bootstrap information for additional guide information, or for an alternative method of acquisition of the same data as in the broadcast SG, as specified in Section 5.4.1.5.4 of OMA BCAST SG [23].

8.4.1.3 SG Data Compression

It is recommended that the SGDUs be transmitted in compressed form using gzip. This requires indicating a Content-Encoding attribute of "gzip" in the FLUTE FDT.

8.4.1.4 Signaling

The SMT entry for an SG service only needs to include a component for the Announcements Channel, since the parameters for any other FLUTE sessions in the SG service are contained in the SGDDs in the Announcements Channel. However, it is recommended that the SMT entry for the SG include components corresponding to all the additional FLUTE sessions, as this will help ensure that none of the IP streams carrying the other FLUTE sessions get dropped at some point in the broadcast facility.

Whatever arrangements are made for providing SG services, it is recommended that each M/H broadcast include a GAT containing a list of all available SG services covering that M/H broadcast.

8.4.1.5 SG Data Partitioning

It is strongly recommended that partitioning on time window be used in the SG data to allow receivers to select SGDUs to process based on the time window they expect to find useful, in the event a broadcaster wishes to send more than the current and next programs being sent in all ensembles.

If the broadcaster makes use of time or service provider grouping, it is recommended that the SGDUs for different time blocks or different service providers be delivered in separate FLUTE sessions. Moreover, it is recommended that the Service fragments and service-associated fragments such as Access and PreviewData be partitioned into a separate FLUTE session from the Content fragments (or from the Content fragments corresponding to long-term program listings), so that the former can be configured for more rapid acquisition.

8.5 Receiver SG Acquisition Model

In order to understand the broadcaster options and tradeoffs for SG transmission and the resulting user experience, it is important to understand how receivers are expected to acquire the SGs they need within a broadcast market. This aspect is complicated by the fact that different receivers may have differing capabilities and strategies for SG acquisition. In particular, they may differ in:

- Number of tuners and R-S decoders (i.e., number of broadcasts and ensembles for which A/V and SG data can be simultaneously acquired).
- Ability to acquire SG data offline (i.e., while a user is not using the Mobile DTV application to watch TV).
- Ability to acquire an SG over the interaction channel.
- Desired level of intrusion on user media viewing experience (e.g., to initiate a re-scan for SG data).

A significant number of receivers capable of SG presentation can be expected to have the ability to process data from only one ensemble, and to do so only while a mobile-TV-specific application is being used by the device owner.

For user's TV viewing experience not to be interrupted to perform guide acquisition on such a device, it must be possible to acquire the guide and any updates without tuning away from the current ensemble.

8.5.1 Bootstrapping and Guide Set

A receiver is expected to determine the potential sources for SG data for each of the M/H broadcasts it detects, but is not expected to tune to each to immediately acquire the SG fragments[4]. These sources are determined by obtaining the GAT in each M/H broadcast, if present. (A receiver may also have information about other SGs, the delivery of which is by mechanisms outside the scope of the ATSC Mobile DTV Standard.)

The high maximum interval (60s) for GAT transmission permitted by A/153 Part 3 [2] means that a receiver is not expected to make GAT acquisition part of any routine signaling scan, whether FIC-only, SLT, or SMT. (The presence or absence of the GAT can be noted during these scans, however.) Therefore, it is likely that the GAT will be acquired in one of two ways and then cached:

- Opportunistically, if the GAT and FLUTE session for a broadcast is available on the ensemble to which the user is tuned on a single tuner, single-R-S-decoder device (the mode recommended herein).
- Global scan: As part of an SG scan and acquire operation (a very slow scan across all broadcasts, during which the user will be unable to view A/V media on a single-tuner device). This is not expected unless the unit is being charged or otherwise not being actively used.

8.5.2 SG Acquisition

Since the standard provides no normative way to distinguish one guide from another, at a minimum the receiver will attempt to acquire the SGDDs for each guide it is capable of acquiring. A receiver device may not be able to acquire all of the SGs signaled in the GAT; for example, the device may not support an interaction channel or a non-M/H broadcast receiver. Such devices will have a minimal capability (only the current and next programs from the current broadcaster can be expected to be displayable).

A device may choose to filter the guide data acquired based on time partitioning announced in the guide, if it can only make use of short-term data in its UI.

8.5.3 Data Merging and Conflict Resolution

8.5.3.1 Multiple Guides

At a certain point in this process, therefore, the SG application on the device is expected to have collected a set of OMA SG fragments from one or more SGs, referring to a range of services. These fragment sets are required to be merged together (in an OMA compliant device).

In the event of a conflict (multiple fragments for the same Service, for example), the receiver is expected to give priority to SG data from the same source as the service being described.

[4] Some receivers may, especially when encountering a broadcast market for the first time, instruct the user to leave the device on for a 'long' time to gather such data.

8.5.3.2 Announcement and Signaling Data

A broadcaster should always attempt to keep the signaling and SG data consistent. A receiver design may be able to deal with circumstances where the information differs, and is required to give priority to the signaling data per A/153 Part 3 [2] if a difference exists.

However, this is more than an either/or choice; as the receiver needs to merge the information from signaling about the services and current programs together, in whatever form necessary to display information to the user, and/or to provide the necessary codec information to the media player. Mismatches between the two sources increase the probability that a receiver will fail to make the proper associations.

Other types of mismatch are possible, though treatment of these is not as well-defined. For example, if a service is described in the SG that has a globalServiceID that matches the M/H URI schema, but the corresponding M/H Service ID is not present in the FIC of the broadcast with the corresponding TSID, or no such broadcast is currently available, the receiver behavior is undefined.

Note that it is possible that one or more of the SGs describes services available by means other than an M/H broadcast, and it may be reasonable to present these to the user. Receiver treatment of such services is out of scope of this document.

8.5.4 Caching Strategies

Once acquired, the GATs and the SGs are expected to be cached by receivers.

Updates to the GAT itself should seldom be needed, but the SG data will be periodically updated, raising the same issues as for initial acquisition in the case of devices not capable of background data acquisition (on separate RF signals).

To enable display of content between channel changes (before the new service is acquired), the PreviewDataReference of the service fragment can be sent with a multi-day expiration period. This enables a receiver to cache this element and display it when the associated channel is selected. Receivers are expected to display the most recently received PreviewData element (pointed to by the PreviewDataReference) marked as a channel change interstitial (usage=1) if no interstitial is available in the cache for the destination service.

8.6 Updates to Announcement Data

An OMA BCAST SG is updated as specified in [23]. This process involves the SG provider making available a new version of the SG, in which some or all of the fragments have been updated (per-fragment versioning is supported by the specification), removed or replaced by new fragments. It is the responsibility of the broadcasters to properly signal these updates, including taking into account the validity interval of the fragments themselves.

There are two reasons to update an SG:

- The data in the currently provided SG time window has been updated, for example to reflect a recent modification to the duration of an event.
- Data for a new time window is being provided, so as to guarantee a certain minimum amount of future programming information; this applies to the time windows of a partition as well as to a complete guide.

A broadcaster should only send the current program and the next program (for all services), but should send this in every ensemble (so a receiver need not tune away to get such data). The guide should then be updated every 15 minutes unless an unplanned schedule change happens. In the case of multiple partitions, no incremental update may be needed for the slower rate partitions unless the broadcaster wishes to hold the total time covered constant.

8.6.1 Broadcast SG

8.6.1.1 Transmission Behavior

A broadcast SG is updated by adding new SGDD and SGDU files to the FLUTE sessions of the Announcement Channel and any auxiliary FLUTE sessions in use. The usual mechanisms within FLUTE to indicate new or updated files, in addition to the versioning within the SG itself, as described in OMA BCAST SG [23], should be handled by the implementations and not require routine intervention by the broadcaster.

There is no signaling change required by A/153 Part 3 [2] or A/153 Part 4 [3] to indicate that the SG has been updated.

8.6.1.2 Receiver Behavior

Detection and acquisition of updated broadcast SG data, whatever the signaling and transmission mechanism, requires that a receiver be tuned to the broadcast and ensemble in which the notification and the data are transmitted.

In the case of a broadcast-only (no return channel) receiver with a single tuner and/or single R-S decoder, acquisition of SG updates from a broadcast SG would require interruption of the user A/V experience, unless the broadcaster transmits at least a short-term guide for all services in every ensemble, as recommended herein.

8.6.2 SG Delivered Over Interaction Channel

For delivery over the interaction channel, a client can detect updates to an SG by retrieving the SGDDs using an HTTP query formatted as defined in OMA BCST SG [23] and comparing the version information to the currently acquired SG.

For many single tuner/decoder devices, this check (and update of the SG data, if changed) should be possible without interrupting the A/V user experience.

9. STREAMING DATA DELIVERY

9.1 Streaming Data Delivery Mechanisms

In order to enable faster and more deterministic service acquisition and to enable accurate inter-component synchronization (i.e. lip sync), the A/153 Standard has adopted timing and buffer models similar to the MPEG-2 systems model.

9.1.1 Timing

In order to preserve the timing relationship between the encoder and decoders, the A/153 Standard has adopted a timing model similar to the MPEG systems model. In the MPEG systems model, the video, audio and any other timing related data in the encoder are synchronized to the encoder's local clock (STC). The encoder maintains the relationships between the individual streams by the use of Presentation Time Stamps (PTS) which are transmitted in the bit stream to signal the decoder when the appropriate time is to present the data. By assuming a constant delay through the transmission system, the encoder is then able to guarantee that the decoder's buffers will not overflow or underflow as long as these presentation time stamps are adhered to. It is necessary, however to ensure that the encoder and decoder clocks are synchronized, and for that, the MPEG transport stream carries a Program Clock reference (PCR), which is a timing indicator that the decoder can use to synchronize its STC. The PCR is a relative clock, which represents the STC of the encoder, and with it, the encoder can indicate its relationship to the PTS of the individual streams to establish how long the streams should remain in the decoder's buffers before presentation. This fixed relationship between the PCR and the various PTS's is what guarantees buffer conformance and data synchronization.

In A/153 the data delivery is via IP, and audio, video and data services are delivered via the Real-time Transport Protocol (RTP). Audio and video synchronization is achieved via the presentation timestamps present in every RTP packet. Each video and audio component is required to provide a continuous stream of RTP RTCP Sender Reports that provides the relationship between the component's RTP presentation timestamp and the system time clock (STC) of the encoder. In an ideal jitter free delivery system, the RTP/RTCP relationship of the various component streams would be sufficient to insure lip sync if synchronization is guaranteed between the various encoders. However, in a typical IP delivery system, delays encountered by the individual streams introduce jitter to each of the decoders. Because the timing recovery of each of the decoders is separate, this jitter is cumulative across the various streams, and lip sync will be adversely affected. To overcome this, A/153 requires the various encoders to be synchronized to a common STC, and a periodic NTP sender report placed in the bit stream to indicate this STC value to the receiver, allowing a common clock recovery mechanism to be shared among the various decoders. In A/153 each service is responsible for sending a single component stream of NTP packets that contains the current encoder system time clock in the NTP Transmit Timestamp field. A/153 has adopted NTPv4 [22]. Since it is being employed in a unidirectional system, not all of the fields have relevance to the A/153 application. Following are some suggested settings for the various fields in the NTP packet.

LI – This field should be set according to NTPv4 Section 7.3 [22]. It may be used by the decoders to prepare for a leap second correction.

VN – This field should be set to 4 indicating NTPv4.

Mode – This field should be set to 5 to indicate broadcast mode.

Stratum – A/153 specifies signaling as though from a primary NTP server. A/153 requires that this field be initially set to 1 by the time source and not be changed by other equipment before transmission.

Maxstrat – Should be set to 16.

Poll – The A/153 Standard requires an NTP packet to be sent every RS frame time, so a reasonable value for this field within the fields accuracy would be 1 second.

Precision – This value should reflect the accuracy of the encoder's 90 kHz clock.

Root Delay – Should be set to 0 and can be ignored by the decoder.

Root Dispersion – Should be set to 0 and can be ignored by the decoder.

Reference ID – Should be set to "GPS" or "PPS".

Reference Timestamp – Should be reset whenever a discontinuity occurs in the encoder clock. Decoders should note this and make appropriate adjustments to their internal clock.

Origin Timestamp – Should be set to 0 and can be ignored by the decoders.

Receive Timestamp – This field can be set to 0 by the encoder. It can be used by decoders at the lower packet processing level to assist upper levels in compensating for decoder processing delays.

Transmit Timestamp – Should be set to the encoder clock value.

Extension fields, a Key Identifier and MAC are not defined in ATSC and may be omitted by encoders. Decoders may ignore them if present.

In order to maintain the timing relationship of the various components of the service, the time relationship between the NTP and RTCP timestamps must be maintained, and the NTP arrival must be as jitter free as possible. Jitter can be introduced by instability in the encoder STC and by repositioning the NTP packets as they make their way through the various multiplexers, de-multiplexers and broadcast equipment in the system. The STC in A/153 is defined to be 90 kHz, which introduces an inherent granularity of approximately 11 microseconds. Since the encoder is required to have less than 500 ns of jitter, its contribution will not be noticeable. This leaves the multiplexer responsible for ordering the IP packets for transmission, the modulator, the decoder and the various equipment involved in routing and transporting the IP streams. In a typical 100Base-T system, 2 packets arriving at the same time can cause a delay of up to approximately 100 microseconds for one of the packets. If multiple packets arrive simultaneously, the delay is compounded, so good system design dictates that any IP transport devices in the system should be lightly loaded to avoid collisions.

In order to insure accurate re-creation of the encoder system time clock via the NTP stream any infrastructure in the transmission chain that delays the transmission of a NTP packet should adjust the NTP transmit timestamp value accordingly. This process will allow the NTP timestamp to be re-aligned to within plus or minus 1 90 kHz tick, but it will change the time relationship with the RTP/RTCP timing data in the bit stream, and since these timing errors are accumulative, a poorly designed system could create buffer underflow issues in the decoder. As a more extreme example, if the ATSC-M/H multiplexer buffers a NTP packet by 250 milliseconds to insert legacy data the multiplexer must add 1073741824 to the 32-bit fraction field of the NTP transmit timestamp. If this causes the fraction field to roll over then the 32-bit seconds field must be incremented by 1. This will allow the cadence of the NTP time to be maintained, but it will have the effect of reducing the presentation time by 250ms.

It is the IP multiplexer's responsibility to sort the incoming IP streams according to their services, place the data in the appropriate M/H slots and generate the Service Signaling Channel for the various ensembles. This will most likely require buffering the data. As long as the delay is constant across the various streams, the encoder's buffer assumption is not violated so no changes in timing would occur. However, A/153 requires that the NTP and RTCP packets cannot be split across RS frame boundaries, so adjustment of the location of these packets will most likely be required, and the NTP time must be adjusted appropriately to compensate for the shift. In addition, the encoder should insure that at least 1 NTP and 1 RTCP packet is present in every RS frame with the understanding that NTP and RTCP packets will not be split across RS frames.

It should also be noted that the addition of the Service Signaling Channel into the ensemble will consume data from the service. If the encoder does not allow for this additional bandwidth in its data generation, it could result in loss of data.

The decoder is to use the NTP transmit timestamp as the system time clock. Thus immediately upon receipt of the first NTP transmit timestamp this value should be used as the NTP system time clock starting point. Upon subsequent receipt of NTP packets the NTP transmit timestamp should be compared with the current system time clock. If the decoder is running significantly faster or slower an adjustment should be made to the decoder system time clock to keep alignment with the system time clock being transmitted in the NTP transmit timestamp.

Also note that the decoder decodes the ensemble in complete RS frame increments. Consequently, if an RTP packet is split across the RS frame boundary, it can take up to a full RS frame or approximately one second to finish decoding that data frame. This is something that the encoder should take into account when it is establishing its timing.

9.1.2 Buffer Model

The A/153 buffer model is similar to the buffer model described in Section 2.4 of ISO-IEC 13818-1 (MPEG 2 Systems) with the exception of redefining buffer sizes to accommodate IP packets, and the addition of a smoothing buffer in the front end to buffer the cadence of the MH Frames.

For A/153, the H.264 video base layer and AAC+ v2 audio layer are always transmitted in decoding order. (H.264 enhancement layers, which do support out of order transmission, will not be addressed here.) Given this, the buffering model can be very simple having only the concept of a transport buffer that holds compressed data upon reception, and stream component decoders. The component decoding processes must be able to decode a media presentation unit (frame) in exactly real-time or slightly faster. Thus each component media decoder must pull the presentation unit out of the transport buffer at least one (1) presentation unit time ahead of the RTP presentation timestamp for that presentation unit. This will insure that the presentation unit will be fully decoded at or slightly before the time it needs to be displayed.

In this model transport buffer fullness is managed by the encoder's transmission of the system time clock (STC) via the NTP stream. Using the timing model above the system time clock is managed to insure that the transport buffer on the client never overflows or underflows.

10. FILE DELIVERY

10.1 File Delivery Mechanisms

Files are delivered in ATSC M/H Broadcasts by means of FLUTE, a protocol specified in IETF RFC 3926 [24], which is designed for massively scalable delivery of files over a one-way IP multicast network.

FLUTE is based on the Asynchronous Layered Coding (ALC) protocol, defined in IETF RFC 3450 [25], which in turn is based on the Layered Coding Transport protocol, defined in IETF RFC 3451 [26].

This section describes the FLUTE protocol briefly, clarifying the concepts of FLUTE sessions and channels, and explaining their relationship to M/H service components. It also provides some guidelines to senders on organizing files into FLUTE sessions for delivery, and provides some guidelines to receivers on extracting files from FLUTE sessions.

10.2 FLUTE File Delivery Mechanism

When using the FLUTE protocol [24], files are delivered in FLUTE "sessions."

FLUTE inherits the definition of a "session" from ALC, which in turn inherits it from LCT. An LCT session consists of a collection of "channels" coming from the same sender (i.e., having the same source IP address), and having the same value of the TSI (Transport Session Identifier) in the LCT header. Each "channel" consists of LCT packets with the same destination IP address and port. There is no restriction in the number of channels in a session, nor on the IP addresses and ports used for the channels (except that for ATSC the addresses must be IP multicast addresses). Different sessions may use distinct IP address/port combinations for their channels, or they may use overlapping address/port combinations for their channels, relying on the source address and TSI value to distinguish the LCT packets belonging to the different sessions. (Usually different sessions use distinct IP address/port combinations.)

The advantage of having multiple channels in a FLUTE session comes in situations where different groups of receivers want to extract different sets of files from a FLUTE session. If the files are suitably divided among multiple channels, each receiver can listen to only the channel(s) containing files it is interested in, and let the filtering of files be done in the IP protocol stack, rather than in the application.

The LCT protocol supports delivery of packets containing segments of binary objects. Each packet has an LCT header containing a "Transport Session Identifier" (TSI), a binary number identifying what session that packet belongs to, and a "Transport Object Identifier" (TOI), a binary number identifying what object of the session that packet belongs to. The ALC protocol added mechanisms for Forward Error Correction (FEC), and the ability to identify the order in which the segments of an object need to be reassembled to recover the object. The FLUTE protocol added a file directory, somewhat similar to a computer disk directory. The FLUTE file directory, called a File Delivery Table (FDT), maps logical file names to TOI values, and allows a good deal of other information to be associated with each file, such as file size, content type, transport encoding type (e.g., gzip), MD5 signature, and FEC parameters used for the file. The FDT is delivered in-band, along with the data, distinguished by having TOI = 0 in the headers of the LCT packets delivering it.

A "component" in the SMT has associated with it an optional IP source address (which is mandatory for FLUTE components), an IP destination address, and up to 31 consecutively numbered UDP port numbers. (where "consecutive" can mean by increments of more than one for certain types of components, but means increments of one for FLUTE components). Each FLUTE component has a TSI value associated with it.

Thus, with the current FLUTE and SMT specifications, there is no correspondence between FLUTE sessions or channels and SMT components, or even services. A single FLUTE session could be shared among multiple SMT services, and it could contain multiple SMT components within a single service. A single SMT component could contain multiple FLUTE channels, as long as they all share the same IP destination address and have consecutive UDP ports. An SMT service could contain multiple FLUTE sessions.

The FLUTE RFC [24] recommends that for a FLUTE session delivered via multicast over the Internet, multiple channels can be used to provide congestion control. The different channels can all carry the same content at different data rates, and a receiver can join the channel that has the highest data rate that the device itself and its link to the server can handle. In this situation the multiple channels would typically have different destination addresses, since that would provide filtering by the routers, given the way IGMP works. However, in a DTV broadcast environment this type of congestion control is not very useful, and so there is no advantage to using different IP addresses for different channels.

Therefore, it is possible in an M/H Broadcast to make an M/H FLUTE component correspond to an IETF FLUTE session by making all the channels in a FLUTE session have the same destination IP address, consecutive UDP port numbers, and no more than 31 channels. This would not be a good idea for FLUTE sessions delivered over the Internet, but works fine for FLUTE sessions delivered via DTV. For simplicity of signaling FLUTE sessions, it is recommended that FLUTE sessions be organized this way.

10.3 Organizing Files into FLUTE Sessions

The first decision to be made when delivering files via FLUTE is how many sessions to use. The typical rule of thumb is that files for different applications go into different sessions. That is, if there is no overlap at all among the receivers to which two different sets of files are targeted, put the files in different FLUTE sessions.

The next decision is how many channels to use for each FLUTE session. In many cases the simplest solution is the best solution. Just use a single channel. As indicated above, if some of the receivers have are interested in only a subset of the files being delivered, or if most receivers will need a subset of the files very frequently, and the rest infrequently, it can be useful to put such subsets in different channels, to reduce the amount of application level filtering of packets that needs to occur. For example, putting the FDT in one channel and the data files in another can make good sense.

Having made these decisions, it is then necessary to decide what size segments to break the files into for transmission. This decision is often based on the payload size of the IP packets in the network. Smaller packets mean more protocol overhead. Larger packets can mean the packets need to be fragmented, which is another type of overhead.

10.4 Extracting Files from FLUTE Sessions

The typical steps for a receiver interested in extracting certain files from a given service are first to get the FLUTE access parameters from the SMT entry for the given service, then to join the IP multicast streams carrying the FLUTE sessions and extract the FDT. By looking up the desired files in the FDT, the receiver can determine which session(s) contain them, and their TOI values, and start collecting the packets containing them. The final step is to extract the file segments from these packets and reconstruct the files.

11. APPLICATION FRAMEWORK

The ATSC Mobile DTV application framework is described in A/153 Part 5 [28]. The Application Framework enables the broadcaster of an audio-visual service to author and insert supplemental content to define and control various additional elements to be used in conjunction with the M/H audio-visual service. It enables definition of auxiliary (graphical) components, layout for the service, transitions between layouts and composition of audio-visual components with auxiliary data components. Furthermore, it enables the broadcaster to send remote events to modify the presentation and to control presentation timeline. The Application Framework further enables coherent rendering of the service and its layout on a variety of device classes and platforms, rendering of action buttons and input fields, and event handling and scripting associated with such buttons and fields.

12. SERVICE PROTECTION DATA DELIVERY AND USAGE

Service protection in ATSC M/H is based on the OMA BCAST DRM Profile as described in [13]. A/153 Part 6 [4] is a profile of the OMA BCAST specification. ATSC selected particular options from the OMA BCAST specification. An example is the encryption of content, where OMA BCAST has allows three possibilities (IPSec, SRTP, and ISMACryp) and ATSC selected only one possibility, IPSec. The purpose of OMA BCAST DRM profile is to manage access to encrypted broadcast streams. It is used to control who gets access to those streams and for how long, operating as a Conditional Access System (CAS).

ATSC M/H is a broadcast system. Therefore, the default mode of operation for the ATSC M/H service protection solution is broadcast-only mode. In broadcast-only mode, messages are broadcast and requests can only be made to the service provider using an out of band process. The ATSC M/H service protection solution can also operate in interactive mode if an interactive channel is available and if the service provider supports it. In interactive mode, the interactive channel is used for registering receivers and for sending rights messages to the receiver. This saves bandwidth on the broadcast channel.

12.1 Service Protection Mode of Operation and Data Structures

12.1.1 Broadcast Streams Protection

The ATSC M/H service protection solution goal is to implement the business models a service provider wishes to use to sell access to encrypted content. To decrypt an encrypted stream, the receiver requires a set of keys which can be provided in real time with the content itself or in advance.

The various types of keys required to give access to the content are described in Section 12.2.2. These keys do not play the same role in the protection system but all are needed. Different types of keys have different lifetimes and the management of these different lifetimes and key synchronization is an important element in any implementation.

12.1.2 Four Layer Key Hierarchy

The 4 layer key hierarchy is a set of keys which, when used in combination, is used to grant access to encrypted channels:

- The Traffic Encryption Keys (TEKs) are used to directly encrypt broadcast audio or video streams using IPSec. IPSec takes the IP packets forming the stream and encrypts them. The lifetime of a TEK is short; they change every few seconds. This lifetime is called a crypto-period.
- TEKs are encrypted by a Service or Program Encryption Key (SEK and PEK respectively) and broadcast in Short Term Key messages (STKMs). An STKM is equivalent to an ECM in CAS in subscription TV systems. A receiver cannot decrypt the confidential part of the STKM, and therefore the associated encrypted channel, unless it has the appropriate SEK or PEK. Having a valid subscription gives a user the right to receive the appropriate SEK or PEK. SEKs are used to implement subscription business models. The lifetime of a SEK is often coordinated with the subscription renewal period; e.g., one month. PEKs are used for pay per view events occurring within the normal subscription service. Only customers who have received a PEK get access to the event. The value of SEK and PEK is high, as this allows accessing an encrypted event or an encrypted service for an extended time. Delivery and storage of these keys must therefore be done securely.
- The Rights Encryption Key (REK) is used to protect the delivery of SEKs and PEKs. SEK and PEK are delivered in Long Term Key Messages (LTKMs). An LTKM is equivalent to an EMM in CAS in subscription TV systems. The REK is delivered to a receiver via a registration protocol and is specific to that particular receiver. Any other receiver receiving the LTKM cannot decrypt the enclosed SEK as it does not own the right REK. The only exception is LTKMs used in broadcast mode (BCROs) which are addressed to broadcast groups. Broadcast-only receivers therefore require an additional set of keys to decrypt LTKMs addressed to groups.
- Both the receiver and the service provider need a public-private key pair. This key pair is unique for each receiver. If an interactive channel is available, during the registration, after a mutual authentication, the service provider can set the REK and broadcast group

keys to be used by the receiver to process LTKMs. Receivers that operate in broadcast-only mode receive the registration key set via the broadcast channel.

The relationships between messages and keys are summarized in Figure 12.1.

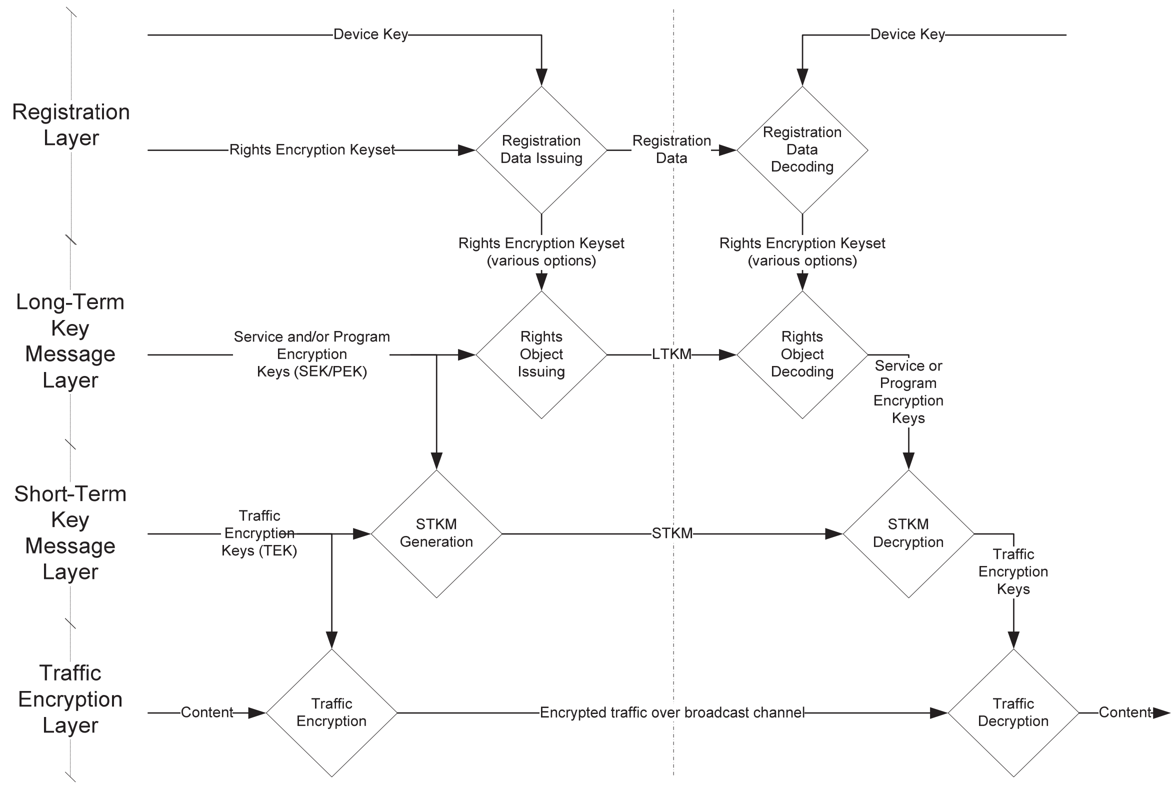

Figure 12.1 The relationships between messages and keys.

12.1.3 Trust Authority

The ATSC M/H service protection solution is built on a public key infrastructure (PKI). Proper operation requires a Certification Authority (CA) and an OSCP responder. In addition, out of bound processes are required to support broadcast-only receivers, as these receivers cannot send their certificates directly to the service provider. OMA defines the functions of the trust authority but does not supply one. ATSC does not supply one either.

CMLA is one entity that provides a trust model for OMA BCAST DRM profile as well as a legal framework for liability and compliance and robustness rules for CMLA licensed products. However, any entity that can provide the same requisite technical functions can be used as the trust authority for ATSC M/H services.

A service provider can be recognized by several trust authorities but this is not mandatory. A receiver, in most cases, is recognized by one trust authority only. Hence, if several trust authorities are used, this can lead to interoperability problems, as a receiver will not be allowed to access content if the service provider and the receiver do not share a common trust authority.

12.1.4 Broadcast groups

ATSC M/H service protection supports fixed broadcast groups of sizes 256 and 512. Broadcast groups are used only for broadcast-only receivers. The goal of groups is to reduce the bandwidth needed to broadcast LTKMs enabling receivers to access content. In the ideal case, one LTKM should enable a full group of 256 or 512 receivers accessing content. In order to optimize bandwidth usage, some care should be put into the assignment of receivers to broadcast groups. If most receivers in the group have the same subscription, they will require the same SEKs and this reduces the number of LTKMs required to reach all receivers with this subscription in the population. Mixed populations are possible in the same network; some groups may be of 256 receivers while other groups are of 512 receivers. Once a receiver is assigned to a group, it is very likely that, except if the group is compromised, it will remain in the group for its lifetime.

The size of the broadcast group registration response sent by the service provider to the receiver is between 330 and 1200 bytes for each receiver in a group. Factors influencing the size of the message are the presence of a signature or not, the presence of the signer certificate chain or not and the size of the group (256 or 512). Each receiver has a node position in the group. When a new receiver is added to a group, a registration message needs to be sent to that receiver with all group key information, other receivers do not need to be updated.

12.1.5 Interactive vs. Broadcast-Only Mode

In interactive mode, LTKMs and registration messages are delivered to receivers over the interactive channel. In broadcast-only mode, all key messages need to be delivered over the broadcast channel. In all cases, STKMs are delivered over the broadcast channel.

The advantage of broadcast-only mode is that a service provider does not rely on any other network for operating its services. The drawback is the broadcast bandwidth allocated to management of the subscriptions.

The advantages of interactive mode are as follows:

- LTKMs and registration messages do not need to be broadcast, thereby reducing broadcast bandwidth usage. There are no RI streams to schedule and signal on the broadcast channel.
- Broadcast groups are not required.
- Reception of messages by the receiver via the interactive channel is acknowledged therefore, the delivery is more reliable than over the broadcast channel.

Because of these advantages, it is advisable to use interactive mode when available. However, the ATSC M/H protection solution supports both modes and they can be used in parallel. Which modes to support is at the discretion of the service provider.

12.1.6 Broadcast Mode and Bandwidth Usage

Messages broadcast are divided in two categories: those that can only sent over the broadcast channel and those that could be sent over the interactive channel. The first category includes only STKM while the second category includes LTKM and registration messages. This section provides guidelines on the parameters which determine how much bandwidth is used in different scenarios.

The size of an STKM is between 100 and 130 bytes depending on the different parameters that are part of an STKM. In addition to the size of the STKM itself, there are two parameters influencing the required bandwidth for STKM:

- The number of STKM streams. If both the audio and video streams of a service are to be encrypted, then it is possible to share a single STKM stream between all encrypted streams which belong to the service. However, it is not advised to share STKM streams between multiple services (even though this may result in a small saving in bandwidth) because the baseCID carried in STKMs, which is used to identify the ROs associated with a service, should be unique per service.

- The repetition rate of STKM. The higher is the repetition rate, the shorter is the time needed for switching between channels. While very long STKM intervals are allowed, a receiver may not have been tuned to a service for that interval (up to ~18.2 hours) and sending an STKM in every RS frame carrying scrambled content would minimize access delay in such cases.

The size of a LTKM is between 100 bytes and 300 bytes depending on the whether a signature is used and the complexity of the access rights that is expressed in the LTKM. In addition to the size of the LTKM itself, in broadcast-only mode, there any many parameters influencing the bandwidth needed for LTKMs:

- The size of the broadcast group (256 or 512). More LTKMs are needed if groups of 256 receivers are used. If groups of 512 receivers are used, the receiver needs more space for the secure storage for the keys.
- The policy for allocating a receiver to a broadcast group. As explained in Section 12.2.4, receivers requiring the same LTKM should be in the same group. When an SEK has to be changed, all receivers needing it receive a LTKM. Hence one and only one LTKM has to be sent to each of the broadcast groups containing such a receiver.
- The broadcasting period and the repetition rate of LTKMs. In order to maximize the probability that a receiver effectively receive the LTKMs it needs, they need to be sent over a given period of time on a regular basis. The longer the period of broadcasting and the higher the rate are, the higher the probability a receiver gets LTKMs. The update period and retransmission rate is at the discretion of the broadcaster.

12.1.7 Parental Control

Parental control in ATSC M/H is controlled from the SG data. Therefore, if rating information is carried in STKMs (as defined in OMA BCAST DRM profile) it should only be displayed for information and not used to take any action by a receiver. (Some receivers may be able to accommodate editing of display of the announcement of programs in compliance with user settings, but this is not required by the A/153 Standard.)

12.1.8 ROAP Trigger Delivery

In interactive mode, LTKM and registration message delivery is triggered using an XML file called a ROAP trigger. ATSC M/H service protection [4] does not mandate any particular method of delivering these triggers to receivers although WAP Push is common in deployments involving mobile operators. ATSC M/H service providers will have to determine the most suitable delivery methods to their subscribers; e.g., downloading triggers from a portal managing subscriptions.

12.1.9 Signaling

OMA BCAST DRM Profile assumes that the OMA BCAST Service Guide (SG) is used to carry protection signaling information like whether a particular channel is encrypted. However, since ATSC M/H does not mandate the usage of an SG this information was also added to ATSC M/H Service Signaling; e.g., M/H component data for STKM.

While direct communication between the signaling and protection systems is generally not needed, protection information like details of STKM streams do need to be coordinated between the two.

12.2 Timing on Access Key Delivery

Timing issues with ATSC M/H service protection are concerned with key lifetimes and schedules for key message delivery over the broadcast channel.

12.2.1 Key Lifetimes

Setting the key lifetimes of TEKs and SEKs involves trade-offs that should be taken into consideration by service providers.

The lifetime of a TEK is recommended by security experts to be measured in seconds. As an example, in subscription pay TV, a crypto-period is around 10 seconds. The length of a crypto-period is defined in the IPSec scrambler and is at the discretion of the broadcaster (who directs the service provider). In the error case where no STKM can be generated (e.g., a server is down), the IPSec scrambler keeps on using the latest available TEK in order to avoid black screens on receivers. The lifetimes of this particular TEK and STKM are therefore extended. Once new STKMs can be generated, the normal operations are resumed and new TEKs are used by the IPSec scrambler. Fresh STKMs are then also added to the broadcast. In such case of error, the behavior of a receiver cannot be predicted after a maximum of 65,536 seconds (maximum lifetime of a TEK as set by the traffic_key_lifetime parameter in the STKM). Some receiver may stop descrambling content while others may keep on descrambling. The interruption of service should then be as short as possible (for example the STKM server can be redundant).

The lifetime of a SEK is measured in weeks or months. The SEK lifetime is often coordinated with the subscription billing period; e.g., one month if subscribers pay for one month's access at a time. This allows easier subscriber management since no additional message needs to be broadcast to the receiver to deny access to the content when a subscription is not renewed.

The lifetime of a PEK is usually measured in minutes and it often is the same as the event to which it gives access.

12.2.2 BCRO and Broadcast Message Delivery Scheduling

Broadcast-mode messages (LTKMs and registration messages) need to be scheduled and regularly re-transmitted to ensure that the messages are received by the devices that do not have an interactive channel. The re-transmission rate has a drastic impact on the bandwidth usage. To maximize efficiency, the re-transmission rate should be similar to the average daily viewing duration of a user.

The RI stream should include all LTKMs carrying a SEK or a PEK used to access currently broadcast content.

The RI stream should include all LTKMs carrying a SEK or a PEK that will be used in a period equal to half of the key (SEK or PEK) lifetime.

The RI stream should have a complete cycle with a duration equal to the average daily viewing duration . An average viewer can expect to get one update a day.

Broadcasters using an interactive channel for LTKM delivery need not transmit the RI stream. Other broadcasters will need to adjust the rate based on available bandwidth and accept that the available bandwidth will significantly constrain the number of non-interactive-capable receivers that can be kept authorized.

12.3 Extraction of Access Keys by Receivers

In some cases, the same receiver may process key messages from multiple service providers. This should not cause any problems since each service provider will have its own certificate that uniquely identifies it to the receivers (see Section 12.1.3). Each service provider can deliver its LTKMs separately to the receiver over the interactive or broadcast channel depending on the mode being used.

Another scenario that may arise is that the same encrypted channels are shared by different service providers who sell them in various channel packages. The protocol called SP4 allows implementing this case in various ways:

- The service providers do not share the associated STKM streams. Only TEK have to be shared, each service provider can use its own SEK and its own LTKMs. In this mode, each service provider has the full control on its ATSC M/H service protection system.
- The service providers also share the associated STKM streams. Each STKM stream is encrypted with a set of SEKs that also need to be shared with all the service providers who are issuing LTKMs for those channels. In this mode, one service provider has to host the full ATSC M/H service protection system and other service providers only keep the LTKM generation.

The OMA BCAST DRM profile specification does not define how the specification should be implemented on receivers but rather the behavior expected from receivers. It defines a secure function that handles all sensitive operations, but does not specify how it is implemented. Only TEKs can be exposed outside of this secure function on the receiver (because TEKs are needed for real-time decryption).

It is the responsibility of the Trust Authority to publish a set of compliance and robustness rules. As an example, CMLA compliance and robustness rules define which keys and other information should be integrity or confidentiality protected on the receiver.

13. VIDEO SYSTEM CHARACTERISTICS

The ATSC Mobile DTV video system characteristics are described in A/153 Part 7 [5]. The system uses MPEG-4 Part 10 AVC and SVC video coding as described in ITU-T Rec. H.264 | ISO/IEC 14496-10 [10], with certain constraints.

14. AUDIO SYSTEM CHARACTERISTICS

The ATSC Mobile DTV (MDTV) audio system characteristics are described in A/153 Part 8 [6]. The system uses HE AAC v2 audio coding as documented in ISO/IEC 14496-3 [29], with certain constraints. HE AAC v2 is used to code mono or stereo audio and is a combination of three specific audio coding tools, AAC Low Complexity (LC), Spectral Band Replication (SBR), and Parametric Stereo (PS).

14.1 Loudness and Dynamic Range Control

Section 5 of A/153 Part 8 specifies that the audio content at the input to the HE AAC v2 encoder has a target measured loudness value of –14 LKFS. There is a footnote to this requirement indicating that methods to measure loudness are explained in the ATSC Recommended Practice A/85, [30], which should be referred to for specific information. In addition, A/153 Part 8 Section 5.1 recommends that Dynamic Range Control (DRC), as defined in ISO/IEC 14496-3 [29], not be transmitted in the bitstream.

14.1.1 Loudness Control

All audio content for MDTV should be measured and (if necessary) adjusted by suitable means[5] to a target loudness of –14 LKFS, before it is sent to the HE AAC v2 encoder. Because this target value is fixed, there is no need to use audio metadata to signal its value.

14.1.2 Dynamic Range Control

When MDTV audio is transmitted without Dynamic Range Control metadata (as recommended for MDTV), the MDTV receiver typically will reproduce the audio with the dynamic range of the content as presented to the HE AAC v2 encoder. This is unlike AC-3 audio for the main service, where the transmitted bitstream generally includes audio metadata for dynamic range control that provides AC-3 receivers with the ability to change the audio dynamic range to suit the capabilities of the consumer's audio system or the user's preference.

In preparing audio content for encoding with HE AAC v2, the broadcaster should optimize the dynamic range of the content before encoding to be both appropriate for the type of content and also suitable for the intended MDTV receiver[5]. It should be noted that the specified –14 LKFS target loudness level reduces the available headroom for program peaks when compared to the main service using AC-3, so a simple linear gain increase of content prepared initially for the main service may not be suitable for repurposing or preparing content for MDTV. In addition, reduced dynamic range may be desirable for services to mobile devices.

14.1.3 Applicability of Loudness and Dynamic Range Control Techniques Described in A/85

Some of the techniques for managing audio loudness and dynamic range control for MDTV are similar to those for the main ATSC service using the AC-3 codec, as described in A/85 [30]. However, the constraints on Mobile DTV audio defined in A/153 Part 8 [6] necessitate that some techniques needed for managing loudness and dynamic range audio for Mobile DTV are different from those described in A/85 [30].

The following points should be noted when considering techniques for MDTV:

- References throughout A/85 [30] to the AC-3 codec's use of metadata, including dialnorm and the AC-3 Dynamic Range Control (DRC) system, do not apply to audio transmitted in the MDTV service and should be disregarded. Audio for MDTV uses the HE AAC v2 codec and does not rely on metadata for loudness and dynamic range control.
- A/85 [30] Section 4 (The AC-3 Multichannel Audio System) does not apply.
- A/85 [30] Section 5 (Loudness Measurement) applies in its entirety, with the exception of the reference in the first paragraph of 5.2 to A/53 and the dialnorm value in the AC-3 stream which should be disregarded.
- A/85 [30] Section 6 (Target Loudness and True Peak Levels for Content Delivery or Exchange) does apply to MDTV in respect to delivery or exchange of content through the network and studio portions of the distribution chain (assuming such distribution serves both a main service and an MDTV service). Section 6 does not apply to distribution to the consumer (either for MDTV or the main service).
- A/85 [30] Section 7 (Metadata Management Considerations Impacting Audio Loudness) does not apply to MDTV.

---

[5] Specific processing methods for use by the broadcaster for loudness control and dynamic range control prior to audio encoding are not part of this recommended practice and manufacturers of devices designed for this purpose should be consulted for more information.

- The portions of A/85 [30] Section 8 (Methods to Effectively Control Program-to-Interstitial Loudness) that refer specifically to AC-3 with the use of dialnorm do not apply to MDTV. However, the solutions described in Section 8.1.1, a), b), and c) are applicable, with the deletion of the words "and that long term loudness matches the dialnorm value". The recommendations in A/85 Section 8.3 item 3) are also applicable to MDTV.
- A/85 [30] Section 9.1 (AC-3 Dynamic Range Control System "Reversible") in general does not apply to MDTV. However, Section 9.1.5, AC-3 DRC: Choosing "None", includes a statement in bullet #3 that is applicable to MDTV, in which case:

  Dynamic range should be controlled in another fashion (rather than using DRC metadata) by the operator or by the program originator.
- A/85 [30] Section 9.2 (Dynamic Range Processing With Metadata Interface "Reversible") does not apply to MDTV.
- A/85 [30] Section 9.3 (Dynamic Range Processing Without Metadata Interface "Irreversible") does apply to MDTV, except that in the last line, the reference to AC-3 should be changed to HE AAC v2.
- A/85 [30] Section 9.4 (Consumer Experience) does not apply to MDTV. It is possible that MDTV receivers may incorporate methods that do not rely on metadata for optimizing the consumer experience and adjusting the dynamic range of content to suit the device. This is outside the scope of this recommended practice.
- A/85 [30] Section 10 (Audio Monitoring Setup) does apply in general to audio for MDTV, except that references to AC-3, dialog normalization and dynamic range control in the first paragraph of 10.1 are not relevant to MDTV.
- A/85 [30] Annexes A, B, C, D, are all non-codec specific and do apply in general to audio for MDTV.
- A/85 [30] Annex F (AC-3 Dynamic Range Control Details) does not apply to MDTV.
- A/85 [30] Annex G (AC-3 Metadata Parameters) does not apply to MDTV.
- A/85 [30] Annex H (Quick Reference Guide for Station and MVPD Engineers) is written largely around the requirements for setting AC-3 dialnorm correctly and, as such, does not apply to MDTV, although the parts relating to measurement are codec and metadata independent and have general applicability.
- A/85 [30] Annex I (Quick Reference for Audio Mixers and Editors Creating Content) has general applicability to audio content that may be intended for both a main service and an MDTV service.
- A/85 [30] Annex J (Requirements for Establishing and Maintaining Audio Loudness of Commercial Advertising in Digital Television when Using AC-3 Audio Codecs) does not apply to MDTV.
- A/85 [30] Annex K (Requirements for Establishing and Maintaining Audio Loudness of Commercial Advertising in Digital Television when Using Non-AC-3 Audio Codecs) applies to MDTV, with the proviso that the "Operator-selected loudness target value" mentioned in K.5 should be –14 LKFS.

Annex A: Mapping of Service Map Table to Service Guide

A.1 INTRODUCTION

Section 5.1 of A/153 Part 4 [3] states: "Where the Service Guide and the M/H Service Signaling Channel [16] contain conflicting metadata, and there exists a defined semantic mapping between the two metadata items, the Service Signaling Channel metadata shall take precedence."

The purpose of this Annex is to define a semantic mapping from metadata items in the Service Map Table (SMT) to metadata items in the Service Guide (SG). The mapping includes not only fields In the SMT proper, but also fields in descriptors that can appear in the SMT. The mapping is in the form of a table that lists fields of the SMT in the left column and gives the corresponding elements and attributes of the SG in the right column, or in some cases gives fields of a Session Description Protocol (SDP) message from an Access fragment in the right column. Explanatory notes appear at the end of the table as needed to clarify the mapping, for example in cases when the mapping involves a data transformation. References to these notes appear in the table.

The table includes only those fields that have substantive content, and omits those fields that contain only coding artifacts or versioning information (since there is no attempt to coordinate version numbers between the SMT and the SG).

The designation "N/A" in the right hand column indicates that the ATSC M/H signaling descriptor, element or attribute does not have a corresponding value in the OMA BCAST SG schema.

The entries in the right hand column show the full path to the element, attribute, or SDP field in the SG, starting with the top level element in the appropriate SG fragment. The usual dot notation is used to show child elements and attributes. In the case of a field in an SDP message, the entire line containing the field appears, enclosed in parentheses to identify the field clearly. In all cases, the target element, attribute or field in the right hand column is in bold font, to make the intended target clear.

ATSC M/H Signaling Field	OMA BCAST Element/Attribute
Service Map Table	
ensemble_id	N/A (not signaled in SG)
num_MH_services	N/A (not an explicit parameter in SG)
MH Services Loop	
MH_service_id	Service.**globalServiceID** (Note 1)
multi_ensemble_service	N/A (not signaled in SG)
MH_service_status	N/A (not signaled in SG)
SP_indicator	N/A (not an explicit parameter in SG)
short_MH_service_name	Service.**Name** (Note 2)
MH_service_category	Service.**ServiceType**
num_components	N/A (not explicitly signaled in SG)
IP_version_flag	Access.AccessType.BroadcastServiceDelivery. SessonDescription.SDP(or SDPRef). (o=<username> <sess-id> <sess-version> <nettype> **<addrtype>** <unicast-address>) Access.AccessType.BroadcastServiceDelivery.

	SessonDescription.SDP(or SDPRef). (c=<nettype> **<addrtype>** <connection-address>) (Note 3)
service_source_IP_address	Access.AccessType.BroadcastServiceDelivery. SessonDescription.SDP(or SDPRef). (o=<username> <sess-id> <sess-version> <nettype> <addrtype> **<unicast-address>**)
service_destination_IP_address	Access.AccessType.BroadcastServiceDelivery. SessonDescription.SDP(or SDPRef). (c=<nettype> <addrtype> **<connection-address>**) where the “c=” line is at the session level
Components Loop	
essential_component_indicator	N/A (not signaled in SG)
port_num_count	Access.AccessType.BroadcastServiceDelivery. SessonDescription.SDP(or SDPRef). (m=<media> <port>/**<number of ports>** <proto> <fmt>) (Note 4)
component_destination_UDP_port_num	Access.AccessType.BroadcastServiceDelivery. SessonDescription.SDP(or SDPRef). m=<media> <**port**>/< count> <proto> <fmt>)
component_source_IP_address	N/A (Note 5)
component_destination_IP_address	Access.AccessType.BroadcastServiceDelivery. SessonDescription.SDP(or SDPRef). (c=<nettype> <addrtype> **<connection-address>**) where the “c=” line is at that media level
num_component_level_descriptors	N/A
component_level_descriptor_loop	
num_MH_service_level_descriptors	N/A
MH_service_level_descriptor_loop	
num_ensemble_level_descriptors	N/A
ensemble_level_descriptor_loop	
SMT Ensemble Level Descriptors	
MH_original_service_id_descriptor	
MH_original_service_id	N/A (not signaled in SG)
MH_string_mapping_descriptor	N/A
SMT MH Service Level Descriptors	
MH_rights_issuer_descriptor	
MH_service_protection_version	N/A
kms_type	Access.KeyManagementSystem.**kmsType**
rights_issuer_URI	Access.KeyManagementSystem. **PermissionIssuerURI**
MH_current_program_descriptor	
current_program_start_time	Content.**StartTime**
current_program_duration	Content.**EndTime** – Content.**StartTime**
title	Content.**Name**

MH_original_service_id_descriptor	
MH_original_service_id	N/A (not signaled in SG)
Content_labeling_descriptor	
metadata_application_format	N/A
metadata_application_format_identifier	N/A
content_reference_id_record	Content.**globalContentId** (Note 6)
Caption_service_descriptor	
number_of_services	N/A (not signaled explicitly by SG)
caption services loop	
language	Content.**TextLanguage.languageSDPTag** (Note 7)
caption_service_number	N/A
easy_reader	N/A
wide_aspect_ratio	N/A
Content_advisory_descriptor	
rating_region_count	N/A
rating region loop	
rating_region	Content.**ParentalRating** (Note 8)
rated_dimensions	N/A
rating dimension loop	
rating_dimension_j	Content.**ParentalRating** (Note 8)
rating_value	Content.**ParentalRating** (Note 8)
rating_description_text	Content.ParentalRating/**ratingValueName**
ATSC_genre_descriptor	
attribute_count	N/A (not signaled explicitly in SG)
attribute loop	
attribute	Content.Genre.href.<ATSC genre-cs URI>:**TermId**
	(Note 9)
SMT Component Level Descriptors	
MH_component_descriptor	
component_type	Access.AccessType.BroadcastServiceDelivery. SessonDescription.SDP(or SDPRef). (m=<media> <port>/<number of ports> <proto> **<fmt>**
component_encryption_flag	N/A
num_STKM_streams	N/A
STKM stream loop	
STKM_stream_id	N/A
MH_component_data	See component data structures below

H.264/AVC Component Data (component_type **35**)	Access.AccessType.BroadcastServiceDelivery. SessonDescription.SDP(or SDPRef). (m=**video** <port>/<number of ports> **RTP/AVP 35** a=rtpmap:**35 H264**/**90000**) (Note 10)
profile_idc	Access.AccessType.BroadcastServiceDelivery. SessonDescription.SDP(or SDPRef). (a=fmtp:**35** sprop-parameter-sets=**<SPS>**
constraint_set0_flag	Access.AccessType.BroadcastServiceDelivery. SessonDescription.SDP(or SDPRef). (a=fmtp:**35** sprop-parameter-sets=**<SPS>**)
constraint_set1_flag	Access.AccessType.BroadcastServiceDelivery. SessonDescription.SDP(or SDPRef). (a=fmtp:**35** sprop-parameter-sets=**<SPS>**)
constraint_set2_flag	Access.AccessType.BroadcastServiceDelivery. SessonDescription.SDP(or SDPRef). (a=fmtp:**35** sprop-parameter-sets=**<SPS>**)
AVC_compatible_flags	Access.AccessType.BroadcastServiceDelivery. SessonDescription.SDP(or SDPRef). (a=fmtp:**35** sprop-parameter-sets=**<SPS>**)
level_idc	Access.AccessType.BroadcastServiceDelivery. SessonDescription.SDP(or SDPRef). (a=fmtp:**35** sprop-parameter-sets=**<SPS>**)
AVC_still_present	N/A
AVC_24_hour_picture_flag	N/A
SVC Enhancement Layer Data (component_type **36**)	Access.AccessType.BroadcastServiceDelivery. SessonDescription.SDP(or SDPRef). (m=**video** <port>/<number of ports> **RTP/AVP 36** a=rtpmap:**36 H264-SVC**/**90000**) (Note 11)
profile_idc	Access.AccessType.BroadcastServiceDelivery. SessonDescription.SDP(or SDPRef). (a=fmtp:<fmt> sprop-parameter-sets=**<SPS>**)
constraint_set0_flag	Access.AccessType.BroadcastServiceDelivery. SessonDescription.SDP(or SDPRef). (a=fmtp:**36** sprop-parameter-sets=**<SPS>**)
constraint_set1_flag	Access.AccessType.BroadcastServiceDelivery. SessonDescription.SDP(or SDPRef). (a=fmtp: **36** sprop-parameter-sets=**<SPS>**)
constraint_set2_flag	Access.AccessType.BroadcastServiceDelivery. SessonDescription.SDP(or SDPRef). (a=fmtp: **36** sprop-parameter-sets=**<SPS>**)
SVC_compatible_flags	Access.AccessType.BroadcastServiceDelivery. SessonDescription.SDP(or SDPRef). (a=fmtp: **36** sprop-parameter-sets=**<SPS>**)
level_idc	Access.AccessType.BroadcastServiceDelivery. SessonDescription.SDP(or SDPRef). (a=fmtp: **36** sprop-parameter-sets=**<SPS>**)
layer_id	Access.AccessType.BroadcastServiceDelivery. SessonDescription.SDP(or SDPRef). (a= mid:**<layer_id>**) (Note 12)
max_temporal_id	N/A (TID from NAL unit headers in bitstream)
max_dependency_id	N/A (DID from NAL unit headers in bitstream)

max_quality_id	N/A (QID from NAL unit headers in bitstream)
num_directly_dependent_layers	**n** in mapping of directly_dependent_layer_id[i] values below
directly dependent layer loop	
directly_dependent_layer_id[i]	Access.AccessType.BroadcastServiceDelivery. SessonDescription.SDP(or SDPRef). (a=depend:<fmt> lay **directly_dependent_layer_id[1]**:<fmt> … **directly_dependent_layer_id[n]**:<fmt>
HE AAC v2 Component Data (component_type **37**)	Access.AccessType.BroadcastServiceDelivery. SessonDescription.SDP(or SDPRef). (m=**audio** <port>/<number of ports> **RTP/AVP 37** a=rtpmap:**37** **mpeg4-generic**/<clock rate/<# channels) (Note 13)
ISO_639_language_code	Content.**AudioLanguage.languageSDPTag** (Note 7)
rap_flag_present	Access.AccessType.BroadcastServiceDelivery. SessonDescription.SDP(or SDPRef). (a=fmtp:<fmt> … randomAccessIndicator=**rap_flag_present)**
RTP_clock_rate	Access.AccessType.BroadcastServiceDelivery. SessonDescription.SDP(or SDPRef). (a=rtpmap:<payload type> <encoding name>/**<clock rate>**/ num_audio_channels)
constant_duration	Access.AccessType.BroadcastServiceDelivery. SessonDescription.SDP(or SDPRef). (a=fmtp:<fmt> … constantDuration=**constant_duration** …)
sampling_rate	Access.AccessType.BroadcastServiceDelivery. SessonDescription.SDP(or SDPRef). (a=fmtp:<fmt> … config=**config …)**
audio_service_type	N/A
audio_channel_association	N/A
num_configs	N/A
Configs loop	
profile_level_id	Access.AccessType.BroadcastServiceDelivery. SessonDescription.SDP(or SDPRef). (a=fmtp:<fmt> … profile-level-id=**profile_level_id** …)
num_audio_channels	Access.AccessType.BroadcastServiceDelivery. SessonDescription.SDP(or SDPRef). (a=rtpmap:<payload type> <encoding name>/<clock rate>/ **num_audio_channels**)
config()	Access.AccessType.BroadcastServiceDelivery. SessonDescription.SDP(or SDPRef). (a=fmtp:<fmt> … config=**config** …)
FLUTE File Delivery Data (component_type 38)	Access.AccessType.BroadcastServiceDelivery. SessonDescription.SDP(or SDPRef). (m=**application** <port>/<number of ports> **FLUTE/UDP)** (Note 14)
TSI	Access.AccessType.BroadcastServiceDelivery.

	SessonDescription.SDP(or SDPRef). (a=flute:tsi:**<tsi>)**
session_start_time	Access.AccessType.BroadcastServiceDelivery. SessonDescription.SDP(or SDPRef). (t=**<start-time>** <stop-time>)
session_end_time	Access.AccessType.BroadcastServiceDelivery. SessonDescription.SDP(or SDPRef). (t=<start-time> **<stop-time>**)
tias_bandwidth	Access.AccessType.BroadcastServiceDelivery. SessonDescription.SDP(or SDPRef). (b=tias**:<bandwidth>**)
as_bandwidth	Access.AccessType.BroadcastServiceDelivery. SessonDescription.SDP(or SDPRef). (b=as**:<bandwidth>**)
FEC_encoding_id	Access.AccessType.BroadcastServiceDelivery. SessonDescription.SDP(or SDPRef). (a=FEC-declaration: <fec-ref> encoding-id**=<enc-id>** instance-id=<inst-id>)
FEC_instance_id	Access.AccessType.BroadcastServiceDelivery. SessonDescription.SDP(or SDPRef). (a=FEC-declaration: <fec ref> encoding-id=<enc-id> instance-id=**<inst-id>**)
STKM Data (component_type 39)	Access.AccessType.BroadcastServiceDelivery. SessonDescription.SDP(or SDPRef). (m=**application** <port>/<number of ports> **UDP vnd.oma.bcast.stkm**) (Note 15)
MH_service_protection_version	N/A
encryption_type	Access/**EncryptionType**
kms_type	Access/KeyManagementSystem/**kmsType** and Access.AccessType.BroadcastServiceDelivery. SessonDescription.SDP(or SDPRef). (a=fmtp:vnd.oma.bcast.stkm … kmstype=**kms_type**)
STKM_stream_id	Access.AccessType.BroadcastServiceDelivery. SessonDescription.SDP(or SDPRef). (a=fmtp:vnd.oma.bcast.stkm … streamid=**STKM_stream_id>**)
base_CID	Service/**baseCID** and optionally Access.AccessType.BroadcastServiceDelivery. SessonDescription.SDP(or SDPRef). (a=fmtp:vnd.oma.bcast.stkm … baseCID=**base_CID**)
srv_CID_extension	Access.AccessType.BroadcastServiceDelivery. SessonDescription.SDP(or SDPRef). (a=fmtp:vnd.oma.bcast.stkm … srvCIDExt=**srv_CID_extension**)
prg_CID_extension	Access.AccessType.BroadcastServiceDelivery. SessonDescription.SDP(or SDPRef). (a=fmtp:vnd.oma.bcast.stkm … prgCIDExt=**prg_CID_extension**)
num_rights_issuers	N/A
Rights issuers loop	
rights_issuer_URI	Access/KeyManagementSystem/ **PermissionIssuerURI**

LTKM Data (component_type 40)	Access.AccessType.BroadcastServiceDelivery. SessonDescription.SDP(or SDPRef). (m=**application** <port>/<number of ports> **UDP vnd.oma.bcast.ltkm** (Note 15)
MH_service_protection_version	N/A
kms_type	Access/KeyManagementSystem**/kmsType** and Access.AccessType.BroadcastServiceDelivery. SessonDescription.SDP(or SDPRef). (a=fmtp:vnd.oma.bcast.ltkm … kmstype=**kms_type**)
LTKM_stream_type	N/A
Rights issuer loop	
rights_issuer_URI	Access/KeyManagementSystem/ **PermissionIssuerURI**
OMA-RME DIMS Data (component_type 41)	Access.AccessType.BroadcastServiceDelivery. SessonDescription.SDP(or SDPRef). (m=**video** <port>/<number of ports> **RTP/AVP 41** a=rtpmap:**41 richmedia+xml**/<clock rate>) (Note 16)
version_profile	Access.AccessType.BroadcastServiceDelivery. SessonDescription.SDP(or SDPRef). (a=fmtp: **41** … Version-profile=**version_profile**)
level	Access.AccessType.BroadcastServiceDelivery. SessonDescription.SDP(or SDPRef). (a=fmtp: **41** … Level=**level**;)
NTP Timebase Stream Data (component_type 42)	(Note 17)
version	N/A
Dynamic Range Type Data (component_type 96-127)	
general_media_type	Access.AccessType.BroadcastServiceDelivery. SessonDescription.SDP(or SDPRef). (m=**<media>** <port>/<number of ports> <proto> <fmt>**)**
ISO_639_language_code	Content/**AudioLanguage.languageSDPTag** (Note 7)
media_type	Access.AccessType.BroadcastServiceDelivery. SessonDescription.SDP(or SDPRef). (a=rtpmap:<payload type> **<encoding name>**/<clock rate> [/<enc params>])
decoding_parameters	Access.AccessType.BroadcastServiceDelivery. SessonDescription.SDP(or SDPRef). (a=fmtp:<format> **<format specific parameters>**)

Note 1: The mapping from an ATSC M/H service_id to an OMA BCAST Service Guide globalContentId element is defined in Annex A Section 1 of A/153 Part 4 [3]. It is important to recognize that the requirement for SMT values to take precedence over SG values is not necessarily meaningful for this field, since this field provides the linkage between

corresponding services in the SMT and SG. If the value for a service in the SG is wrong, the linkage is broken. There is no straightforward way to know what the value in the SG should be, since it is not clear which service it is supposed to represent.

Note 2: The short_MH_service_name field in the SMT is at most 16 bytes in length and is UTF-8 encoded. The Service.Name element of the SG is unbounded in length, can have multiple instances in different languages, and no encoding is specified in either the OMA BCAST SG standard or the ATSC M/H standard. It is recommended that UTF-8 encoding always be used for the XML messages carrying the SG fragments, and that one of the Service.Name elements for each service have the same length and value as the short_MH_service_name in the SMT for the same service.

Note 3: The IETF SDP RFC allows different components of an IP multicast session to have different IP address types. However, different IP address types in a single service are not allowed for ATSC M/H services.

Note 4: The value of port_num_count maps to "<number of ports>" as shown above, but the mapping requires a multiplier for some media types. The reason is that the port_num_count in A/153 gives the actual number of ports used, whereas in some cases the value of <number of ports> in the SDP data gives the number of sets of ports used. For example, port_num_count is twice that of "<number of ports>" for RTP/RTCP streams, such as audio and video streams, since in this case the <number of ports> in the SDP data gives the number of RTP sessions, which is the number of RTP/RTCP pairs. On the other hand, the value of port_num_count is equal to the <number of ports> for many other media types, such as FLUTE sessions or STKM streams. The default value for <number of ports> in the SDP data is 1 when no value is specified. For an RTP session (i.e. RTP/RTCP pair) this corresponds to a value of 2 for the port_num_count in the SMT; however, port_num_count must be specified and therefore has no default value.

Note 5: The IETF SDP RFC does not allow different components of an IP multicast session to have different IP source addresses. However, different source IP addresses in different components of a single service are allowed for ATSC M/H services.

Note 6: The mapping from a content label in an ISO 13818-1 Content Labeling descriptor to an OMA BCAST Service Guide globalContentId element is defined in Annex A Section 2 of A/153 Part 4 [3].

Note 7: The value in an ATSC M/H language field maps to the Primary subtag of an OMA BCAST SG languageSDPTag attribute. As noted above, the Primary subtag of an OMA BCAST SG LanguageSDPTag string is always a 2-character or 3-character ISO 639 code for the language. To conform to RFC 3066 (and therefore to OMA BCAST SG, and therefore to A/153), the 2-character code is used if it exists. Since the languages most likely to be used in the countries with M/H broadcasts all have a 2-letter ISO 639.1 code, the transmitted LanguageSDPTag normally consists of a 2-letter Primary-subtag. In the general case this might be followed by additional subtags, conforming to IETF 3066.

The receiver mapping to be used for matching tracks is therefore defined as between the ISO 639.2 values in the SMT to the Primary subtag of the LanguageSDPTag in the corresponding SG AudioLanguage or TextLanguage elements.

It is recommended that receivers have a list of ready-to-be-displayed Language names associated with their 2-letter and 3-letter codes. Then the stored value is simply displayed for a service component when needed (based on SMT if different from SG) and the table is referenced to select services. Receivers are recommended to store a set of such relationships, but if a text name is not available, updating the SG with the 3-letter string is recommended as it is somewhat easier to understand than a two-letter string.

For mapping from the 2-letter ISO-639.1 Primary subtags appearing in the SG, to the selected language track, the 2-letter tags should be converted to the corresponding 3-letter ISO-639.2 tags before the track matching process (using the Bibliographic code).

For language track selection when more than one audio or caption track is available, and none of the available tracks match the viewer's preset preferences, the receiver should present the viewer with the language names to make a choice. The only time when it is expected that the receiver would present a 3-letter ISO 639 code from the SMT to the viewer for selection is when no SG data is available and that 3-letter code in the SMT does not appear in the receiver's list of mappings from language codes to language names.

The following examples (assuming English) illustrate how the matching and precedence rules are supposed to enable the correct label to be displayed to the user:

Example 1:

- SMT: ISO_639_language_code = "eng"
- SG: languageSDPTag = "en"; AudioLanguage string value = "English"

The 3-character ISO 639.2 code "eng" (English) corresponds to the 2-character ISO 639.1 code "en", so no change is needed.

Example 2:

- SMT: ISO_639_language_code = "eng"
- SG: languageSDPTag = "es"; AudioLanguage string value = "Spanish"

The 3-character ISO 639.2 code that corresponds to the 2-character ISO 639.1 code "es" (Spanish) is "spa", so there is a conflict. Should change the AudioLanguage string value to "English" for display of the current program in the SG.

Example 3:

- SMT: ISO_639_language_code = "und"
- SG: languageSDPTag = "en"; AudioLanguage string value = "English"

The SG value "en" (English) does not conflict with the SMT value "und" (Undetermined). The difference in codes could simply mean that the source for the SG had more information than the source for the SMT. Should not change the values in the SG.

Example 4:

- SMT: ISO 639 language code = 0x000000
- SG: languageSDPTag == "en"; AudioLanguage string value == "English"

The SG value "en" (English) conflicts with the SMT value 0x000000 (Not Applicable), since the latter asserts that the audio has no linguistic content. Should change the SG value "en" to "zxx", and change the AudioLanguage string value to "No Language" for display.

If the selected user interface language for the SG is some other language, then the AudioLanguage string value for display would typically be in that other language.

Note 8: The mapping from an ATSC Content Advisory descriptor to an OMA BCAST Service Guide Parental Guidance element is defined in A/153 Part 4 Section 6.10.3 [3]. Note that in the US (rating region 0x01), the descriptor in the SMT will not include a rating_description_text field (as there is not one), per CEA-766. The text that would normally correspond in the OMA BCAST SG (ratingValueName attribute) is also modified as described in Part 4, based on the text strings present in CEA-766.

Note 9: The value of the “attribute” field in an ATSC_genre_descriptor in the SMT is the code value of an entry in Table 6.20 of the ATSC PSIP Standard [15]. The value of the “href” attribute of an OMA BCAST Service Guide Genre element is a reference to the TermId of a term in the ATSC genre-cs Classification Scheme defined in A/153 Part 4 [3]. The terms in the genre-cs scheme correspond to the entries in Table 6.20, and the TermId of a term in the genre-cs scheme is identical to the code value of the corresponding entry in Table 6.20.

Note 10: The representation of parameters for AVC streams in an SDP message is specified in IETF RFC 3984 [7], which in turn references ISO/IEC 14496-10 (ITU-T H.264) [10] for details of the encoding of the individual parameters into the <SPS>.

Note 11: The representation of parameters for SVC streams in an SDP message is specified in IETF Internet Draft draft-ietf-avt-rtp-svc-20.txt [8], which in turn references ISO/IEC 14496-10 (ITU-T H.264) [10] for details of the encoding of the individual parameters into the <SPS>.

Note 12: The use of layer_id for signaling decoding dependencies in an SDP message is specified in IETF RFC 5583 [9].

Note 13: The representation of parameters for HE AAC v2 streams in an SDP message is described in A/153 Part 8 Annex A [6].

Note 14: The representation of parameters for FLUTE sessions in an SDP message is specified in IETF Internet Draft draft-mehta-rmt-flute-sdp-05.txt [11].

Note 15: The representation of parameters for STKM and LTKM streams in an SDP message is defined in OMA BCAST “Service and Content Protection for Mobile Broadcast Services” [13], Section 10.1.

Note 16: The representation of parameters for OMA RME streams in an SDP message is specified in A/153 Part 5 Section 6.9 [28].

Note 17: There is no completely standard way to describe the NTP timebase stream in SDP, and A/153 Part 4 does not specify one. Receivers are therefore expected to process this information based on the component level destination IP address and port communicated in the SMT.