Second Report of the Video Programming Accessibility Advisory Committee on the Twenty-First Century Communications and Video Accessibility Act of 2010

Video Description

April 9, 2012

Table of Contents

Table of Figures

I. INTRODUCTION/OVERVIEW

This is the Video Description report which is one of three separate reports that make up the Second Report of the Federal Communications Commission's Video Programming Accessibility Advisory Committee on the Twenty-First Century Communications and Video Accessibility Act of 2010. The First Report addressed issues relating to closed captioning of video programming delivered using Internet protocol. This report addresses issues relating to video description among other items.

On October 8, 2010, President Obama signed the Twenty-First Century Communications and Video Accessibility Act of 2010 ("CVAA" or "the Act"), which amended certain sections of the Communications Act of 1934.[1] Among other provisions, this Act required the Federal Communications Commission ("FCC" or "Commission") to reinstate video description rules that had been deleted as the result of a judicial proceeding.[2] In addition, the Act required the Commission to establish an advisory committee -- the Video Programming and Emergency Access Advisory Committee ("VPAAC").[3] Under the Act, the VPAAC, among other things, was charged with developing and submitting to the Commission a report on video description within 18 months of the date of enactment.[4]

With respect to video description, the Act requires the report to include the following:

- A recommended schedule of deadlines for the provision of video description.[5]
- An identification of the performance objectives for protocols, technical capabilities, and technical procedures needed to permit content providers, content distributors, Internet service providers, software developers, and device manufacturers to reliably encode, transport, receive and render video descriptions of video programming, except for consumer generated media, delivered using Internet protocol or digital broadcast television.
- An identification of additional protocols, technical capabilities, and technical procedures beyond those available as of the date of enactment of the Twenty-First Century Communications and Video Accessibility Act of 2010 for the delivery of video descriptions of video programming, except for consumer generated media, delivered using Internet protocol or digital broadcast television that are necessary to meet the identified performance objectives.
- A recommendation for technical standards to address the performance objectives identified.

---

[1] Twenty-First Century Communications and Video Accessibility Act of 2010, Pub. L. No. 111-260, 124 Stat. 2751 (2010) (amended 2010).

[2] *Id*. at § 202(f)(1); In a *Report and Order* issued on August 25, 2011 the FCC reinstituted the rules. In re Implementation of the Twenty-First Century Communications and Video Accessibility Act of 2010, Report and order, 26 FCC Rcd 11847 (Aug. 25, 2011), corrected, 2011 FCC LEXIS 4285 (Sept. 1, 2011) and 2011 FCC LEXIS 4102 (Oct. 5, 2011).

[3] *Id.* at § 201(a).

[4] *Id.* at §201(e)(2).

[5] . The FCC rules issued on August 25, 2011, established the calendar quarter beginning July 1, 2012 as the compliance date for the new rules.

- A recommendation for any regulations that may be necessary to ensure compatibility between video programming, except for consumer generated media, delivered using internet protocol or digital broadcast television and devices capable of receiving and displaying such programming, except for consumer generated media, in order to facilitate access to video descriptions.

The VPAAC created a working group, Working Group 2, to assist in completing this task. Working Group 2 was co-chaired by Paul Schroeder [6] (American Foundation for the Blind) and Charlie Kennamer (Comcast). This Working Group had been meeting by telephone every two weeks, and more frequently as needed.

II. BACKGROUND AND HISTORY OF VIDEO DESCRIPTION

In 1984, the Secondary Audio Program ("SAP") standard was established in the United States by the National Television Systems Committee ("NTSC"). SAP was established as part of Multichannel Television Sound ("MTS") standard as an auxiliary audio channel for analog television that could be broadcast or transmitted both over the air and by cable television providers.

Initially, the primary network and cable broadcasting application of SAP was for the voluntary transmission of a secondary language program dialogue audio track, such as the Spanish translation of an English language program.

With the realization that SAP could also be used for delivery of other program related audio services, video description for broadcast and cable television was born.

Video description consists of a narrator talking through the presentation, describing what is happening on the screen during the natural pauses in the audio, and sometimes during dialogue if deemed necessary. Key visual elements described in the narration include characters in the scene, location of the scene, what the characters are wearing, colors, gestures and scene changes which, when described, engage a viewer who is blind or visually impaired with the story.

The following are significant events in the history of video description:

- 1974 - While working on his broadcasting master's thesis in "television for the blind," Gregory Frazier develops the concepts underlying video description.
- 1982 - The Metropolitan Washington Ear works with the producers of the PBS "American Playhouse" television broadcast to simulcast video description on radio reading services.
- 1985 - TV station WGBH, the Boston, Massachusetts, member of the Public Broadcasting Service (PBS), begins investigating uses for the new technology of stereophonic television broadcasting (Multichannel Television Sound), which allowed for a monophonic audio channel, called the Secondary Audio Program ("SAP").
- 1987-1988 - The Metropolitan Washington Ear works with the WGBH Educational Foundation, Public Television Playhouse, Inc., and PBS in a year-

[6] In January 2012, Paul Schroeder of AFB replaced Brad Hodges of AFB as Working Group 2 co-chair.

long nationally broadcast test of what would become Descriptive Video Services. For the first time, synchronized, pre-recorded video description was broadcast via satellite (on the SAP channel) for the season's 26 "American Playhouse" productions.

- 1990 - WGBH Educational Foundation launches Descriptive Video Services (DVS®), a subsidiary to provide video description for television viewers.
- 1997 - The American Foundation for the Blind (AFB), publishes the seminal work *Who's Watching?: A Profile of the Blind and Visually Impaired Audience for Television and Video*, by Jaclyn Packer, and Corinne Kirchner. Based on a survey of blind and visually impaired people, this publication provided detailed demographic information about the experience with and interest in video description as well as viewing habits and preferences among this population. The survey found that blind and visually impaired individuals watch television at comparable rates to the general population. The report also addressed the real life consequences of lack of full access to television programs.
- 1998 - Congress amends the Rehabilitation Act to require Federal agencies to make their electronic and information technology accessible to people with disabilities. Beginning in June 2001, most film, video, multimedia, and information technology produced or procured by Federal agencies must include video description.[7]
- 1999 – The FCC announces its Notice of Proposed Rulemaking for phased-in video description for television.[8]
- 2000 - The FCC adopts rules requiring major broadcast networks and cable companies in the top 25 television markets to provide 50 hours of described programming per quarter effective April 2002.[9]
- 2002 – The U.S. Court of Appeals for the District of Columbia reverses the FCC ruling requiring video description for television, finding that the FCC had acted beyond the scope of its authority in adopting those rules.[10] CBS and PBS continue to provide approximately the same 50 or more hours of described programming per quarter. Other broadcast and cable networks also continue to provide varying amount of described programming.
- 2007 - Audio Description Coalition publishes the first version of its Standards for Audio Description and Code of Professional Conduct for Describers.
- November 2008 - The Described and Captioned Media Program (DCMP) and the American Foundation for the Blind publish the Description Key. To meet the need for guidelines for description of educational media, DCMP, under a grant from the U.S. Department of Education, commissioned AFB to undertake research and conduct a consensus-based process among experts to create the

[7] 36 CFR § 1194.24: *Electronic and Information Technology Accessibility Standards*, Architectural and Transportation Barriers Compliance Board, 65 FR 80500-01, 2000 WL 1860548 (F.R.).
[8] In the Matter of Implementation of Video Description of Video Programming, NPRM, 14 FCC Rcd 19845 (1999).
[9] In the Matter of Implementation of Video Description of Video Programming, Report and Order, 15 FCC Rcd 15230 (2000).
[10] *Motion Picture Ass'n of Am. v. FCC*, 309 F.3d 796, 806–07 (D.C.Cir.2002).

"DCMP Description Key."

- 2009 - American Council of the Blind launches its Video Description Project to offer training, establish standards, encourage growth, disseminate information, and encourage studies of video description.
- 2010 - President Obama signs into law the "Twenty-First Century Communications and Video Accessibility Act of 2010." The Act, among other things, requires the FCC to reinstate with certain revisions the video description rules that were overturned in court in 2002.
- 2010 – The FCC announces the establishment and appointment of members of the Video Programming Access Advisory Committee, an advisory committee required by the CVAA.
- 2011 – The FCC reinstates the video description rules, pursuant to the CVAA.

A. Technique

The rights to create a video description track must be obtained from the program content owner or licensee, before the production process can begin. The content owner or licensee is not always the same entity as the distributor.

Once those rights are obtained, video description writers screen a program and then write a script describing visual elements which are important to understand what is occurring on screen and the plot as a whole. They carefully time the placement and length of the description to fit within natural pauses in the dialogue. For example, in the opening credit sequence of the children's series Arthur on PBS, the description has been performed as follows:

> *"Arthur is an 8-year-old aardvark. He wears round glasses with thick frames over his big eyes. He has two round ears on top of his oval-shaped head. He wears red sneakers and blue jeans, with a yellow sweater over a white shirt."*

The length of descriptions and their placement in the program are largely dictated by what can fit in natural pauses in dialogue, or between other critical sound elements. Once recorded, placed and mixed with a copy of the original soundtrack, the description track is then "laid back" to the master on a separate audio track or to its own master.

Video descriptions for digital television are carried as a secondary audio service during a broadcast. They are typically accessed from a menu in the viewer's home television receiver or set-top box. In some cases, a user can set the audio stream carrying descriptions as the default choice for audio, if that service is available for a particular program. Further details regarding how video descriptions are carried by the various industry segments are described in Section 3, below.

B. Role of the Federal Communications Commission

The CVAA required the FCC to reinstate rules regarding the provision of video description by television broadcasters and multichannel video program distributors ("MVPDs"), with certain modifications.[11] On August 24, 2011, the Commission released rules implementing this provision.[12]

These rules require MVPDs (on those systems with 50,000 or more subscribers) to provide 50 hours of video-described programming each calendar quarter on the top five non-broadcast networks[13]. Affiliates of the top four broadcast networks in the top 25 television markets must also provide 50 hours of video-described programming each calendar quarter.[14] To count toward satisfying the 50-hour per quarter video description requirement, video-described programming on the top four broadcast network stations and the top five non-broadcast networks must appear on prime time or children's programming. In addition, all MVPDs and network-affiliated television stations must pass through video descriptions contained in the programming they carry if "technically capable" of doing so.[15] Full compliance with the rules is required beginning on July 1, 2012.

The FCC's Order reinstating the video description rules mentions the VPAAC in two respects.

Paragraph 50 of the Order states that:

> 50. *Quality Standards*. The *NPRM* sought comment on whether we should adopt quality standards for video description. The majority of commenters that address this question are strongly opposed to the imposition of quality standards of any kind. Other commenters do support the imposition of quality standards, with some pointing to the possible adoption of such standards in the closed captioning context as a demonstration of the need for rules. Nonetheless, we decline to adopt any such standards at this time. We acknowledge that our capacity to adequately judge description quality could benefit from practical experience as entities begin implementing these rules. Nonetheless, given the quality issues that have arisen in the closed captioning context, we will invite comments on the quality of video description when we conduct the inquiry that will inform our first report to Congress under the CVAA. We also recommend that the VPAAC consider this issue, and will include any analysis they provide in the same report. If necessary, we will revisit this issue at a later date.

Paragraph 51 of the Order states that:

> 51. *Program Selection*. In the *NPRM*, the Commission sought comment, for informational purposes, on how programs are likely to be chosen for description. The majority of commenters that address this question are strongly opposed to the Commission seeking information about

[11] CVAA, Sec. 202(a).

[12] In re Implementation of the Twenty-First Century Communications and Video Accessibility Act of 2010, Report and order, 26 FCC Rcd 11847 (Aug. 25, 2011), corrected, 2011 FCC LEXIS 4285 (Sept. 1, 2011) and 2011 FCC LEXIS 4102 (Oct. 5, 2011).

[13] For these purposes, the FCC found that the top five non-broadcast networks currently are USA, Disney Channel, TNT, Nickelodeon/Nick at Nite, and TBS.

[14] Television stations affiliated with the top 4 networks in markets 26 -60 are also subject to the 50-hour-per-quarter obligation beginning July 1, 2015.

[15] The rules include an exemption to the pass-through requirement if the second audio stream is being used for another program-related purpose (such as Spanish language programming.)

program selection even for informational purposes. Given the fact that only a small subset of programming will be required to be video described, the Commission also asked whether we should require that the availability of video description on certain programs be publicized in a certain way. All commenters agree that this information should be widely and clearly available, and most agree that this will occur without the need for regulation. We decline, at this time, to require that the availability of video description on certain programs be publicized in a certain manner. Nonetheless, we expect that programmers, stations, and systems will provide this information to viewers in an accessible manner, including on their websites and to companies that publish television listings information. We recommend that the VPAAC consider this issue and analyze industry best practices. In particular, we recommend that the VPAAC consider how broadcasters provide notice to MVPDs as to which programming is video described, and how effective that notice is. Both NAB and NCTA indicated that use of the ISO-639 language descriptor might be appropriate, but that the issue can be resolved through industry coordination. We recommend the VPAAC examine whether this coordination has been successful.

III. TECHNICAL CAPABILITIES, PROTOCOLS AND PROCEDURES

A. Introduction

While the FCC's requirement to provide video description will not be effective until the third quarter of 2012, video description is voluntarily provided today by some program providers and can be viewed by over the air audiences and by MVPD customers. This section describes the chain through which video description currently reaches the consumer. It begins with the production of the described programming by the video programming provider, which is then received by and distributed through over-the-air digital television, cable television, Internet Protocol television, or Direct Broadcast Satellite to the consumer's home. The rendering and reception of video description is the final step in the process.

MVPDs and over-the-air broadcasters send video description in a second audio stream that is shared with alternate languages like Spanish. The VPAAC has identified no technical impediments to the reliable transport of video description throughout this chain. Figure 1 below provides a high level summary of how content, and associated video description, is provided today.

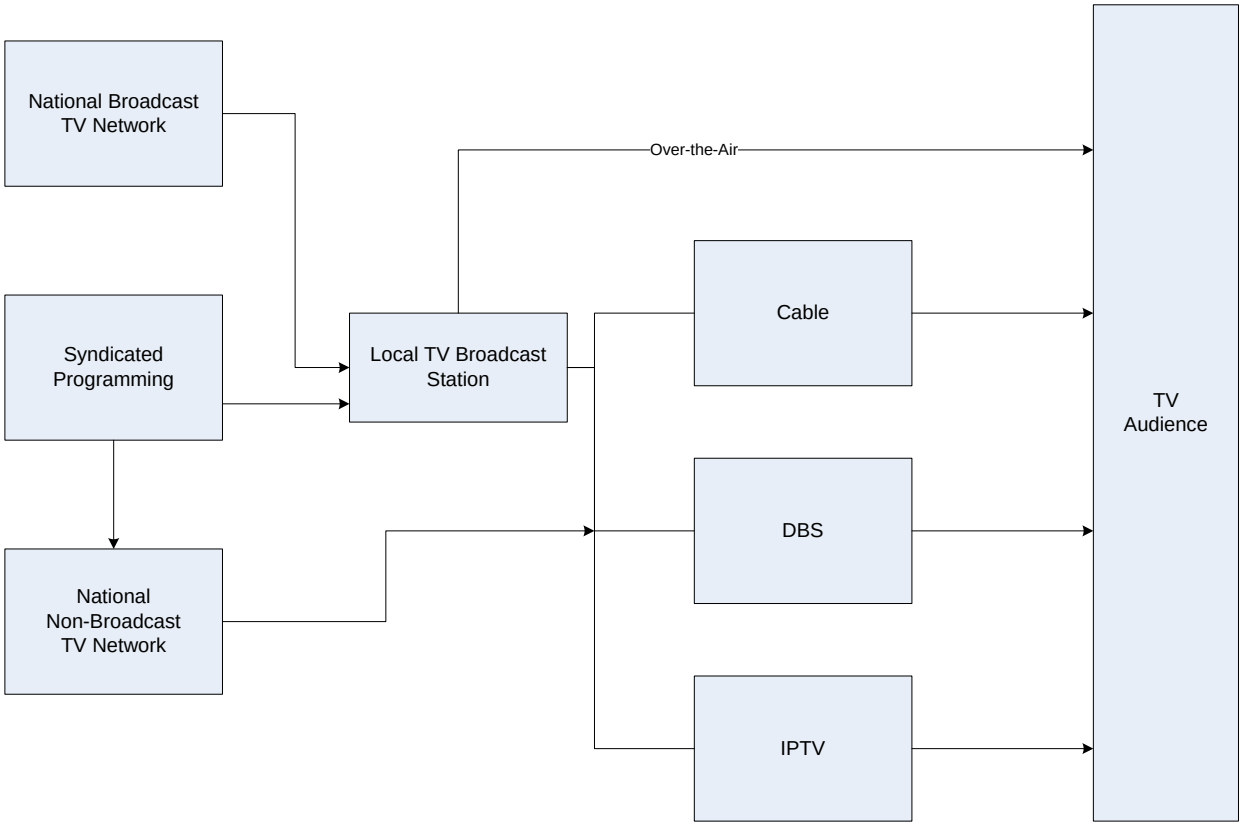

Figure 1 High Level Distribution Summary

Link to accessible textual description of Figure (1)

1. Production

Video description can be created for a wide variety of program genres, including scripted drama, children's programming, documentaries, movies, mini-series and scripted comedy. While the specific steps and timing of the video description process can vary widely depending on a number of factors, we provide here a general overview of that process.

The video description production process begins with the securing of rights from the video programming owner (VPO) to video describe the program. The VPO, or a video programming distributor (VPD) or video programming provider (VPP) as its licensee, typically secures the production services required for the creation and delivery of video description from an outside accessible media services company. In the current marketplace there are a number of these services (both not-for-profit and for-profit), which range from full service accessible media agencies providing video description, closed-captioning and other related accessible media production services, to standalone video description agencies providing a single service.

The initial delivery of a rough cut of the program to the video description agency occurs about a week prior to the first scheduled program air date, although individual network and producer delivery patterns can vary widely. Upon receipt of this rough cut the video description

writers start working on a script. [16] The final version of the program may not be available to the video description agency until close to air date, sometimes the day before, or day of, air. At that time, the video description agency must conform and/or revise their video description script to accommodate changes or additions to the program and complete final voicing and audio mixing before delivering the completed video description content to the client. The video description agency's final product for delivery to the client must be a pristine broadcast quality full mix, a balanced combination of the main program audio and the video description.

The video description agency typically delivers a completed full mix of the video described program audio track back to the client via File Transfer Protocol (FTP), usually in the Waveform Audio File (.wav) format. Upon delivery, the client then retrieves the video description file from the FTP location to utilize it in either a linear (videotape) or non-linear (tapeless) video editing environment. The video description production element is then integrated into the air tape or digital program file. Once program formatting is completed by the client, the air-ready program content with video description is delivered to a cable or broadcast network's distribution facility for play-out to the viewing audience with video description included in the broadcast at the scheduled time(s) of air.

2. Distribution via Digital Broadcast Television to Over-the-Air Audiences

ATSC standards[17], and associated FCC rules, specify the transport and signaling of audio services within the U.S. broadcast system. ATSC A/52, *Digital Audio Compression Standard,* defines the AC-3 audio codec and the descriptor used to signal the characteristics of audio services in the broadcast emission. ATSC A/53, *ATSC Digital Television Standard,* Part 5 specifies certain constraints on the AC-3 audio codec for use in the ATSC digital television system. These standards together describe the structure of the audio tracks and the syntax and semantics of the associated information tables.

The ATSC digital television system is built on MPEG-2 standards defined by ISO/IEC[18]. For each service in the multiplex, the Program Map Table (PMT) (defined in MPEG-2 *Systems,* ISO/IEC 13818-1) contains an AC-3 audio descriptor associated with each audio track. Video

---

[16] Video description agency writer recruitment and training starts with skilled professional writers who are taught in the unique style of writing for a blind and visually impaired audience. The main goal of a video description writer is to provide the blind and visually impaired audience the information needed to better and more fully understand the final product. The writers are trained to describe during pauses where there is no dialogue, or sometimes where there is music (only if the music is not pertinent to the storyline). The writers typically describe from the program picture, not from the script or audio. By describing only the content that they see, they are able to provide the blind or visually impaired consumer with a more complete experience. It is the describer's goal to match the timing of their descriptions to a program's visual content in close proximity to the visual's appearance, using the first natural pause or break in the primary program audio (e.g., dialogue, sound effects, music, etc.) to insert the description. As a result and to necessarily varying degrees, video descriptions are often performed as asynchronous with the picture, preceding, coinciding or in close placement after a particular program scene. The desired end result is intended to provide a blind or visually impaired consumer the ability to follow the flow of the program.

[17] ATSC Standards can be found here: http://www.atsc.org/cms/index.php/standards/published-standards

[18] MPEG standards can be purchased directly from ISO by email (sales@iso.org), from their website (http://www.iso.org/iso/home.html) or from a National Body. Some MPEG standards are publicly available (http://standards.iso.org/ittf/PubliclyAvailableStandards/index.html).

description (or audio description) is known as "Visually Impaired" (VI) audio in the ATSC standards. As defined in A/53 Part 5:2010, Section 6.4, "The VI associated service is a complete program mix containing music, effects, dialogue, and additionally a narrative description of the picture content. The VI service may be coded using any number of channels (up to 5.1)." VI tracks are identified by a value of '2' in the bitstream mode parameter (formally known as "bsmod").

The AC-3 audio descriptor also identifies the spoken language of the VI audio track.

In addition to ATSC A/53 Part 5 and A/52, ATSC A/65, *Program System Information Protocol (PSIP) Standard*, defines Event Information Tables (EIT) that are used to announce information about scheduled programming, including the presence of types of audio streams in the electronic program guide (EPG) of broadcast television. Thus it may be possible for viewers to determine in advance whether or not a given program will have a video description service.

Today, most MVPDs depend on the presence of an ISO 639 descriptor identifying the broadcaster's audio channel carrying Descriptions as "SPA" (Spanish), in addition to the AC-3 descriptor[19]. This additional descriptor is used by MVPD set-top boxes to identify and select the secondary audio service.

[19] SCTE-54: *Digital Video Service Multiplex and Transport System Standard for Cable Television*, can be found on SCTE's website: http://www.scte.org/standards/Standards_Available.aspx

3. Distribution via Digital Broadcast Television to Cable Television Customers

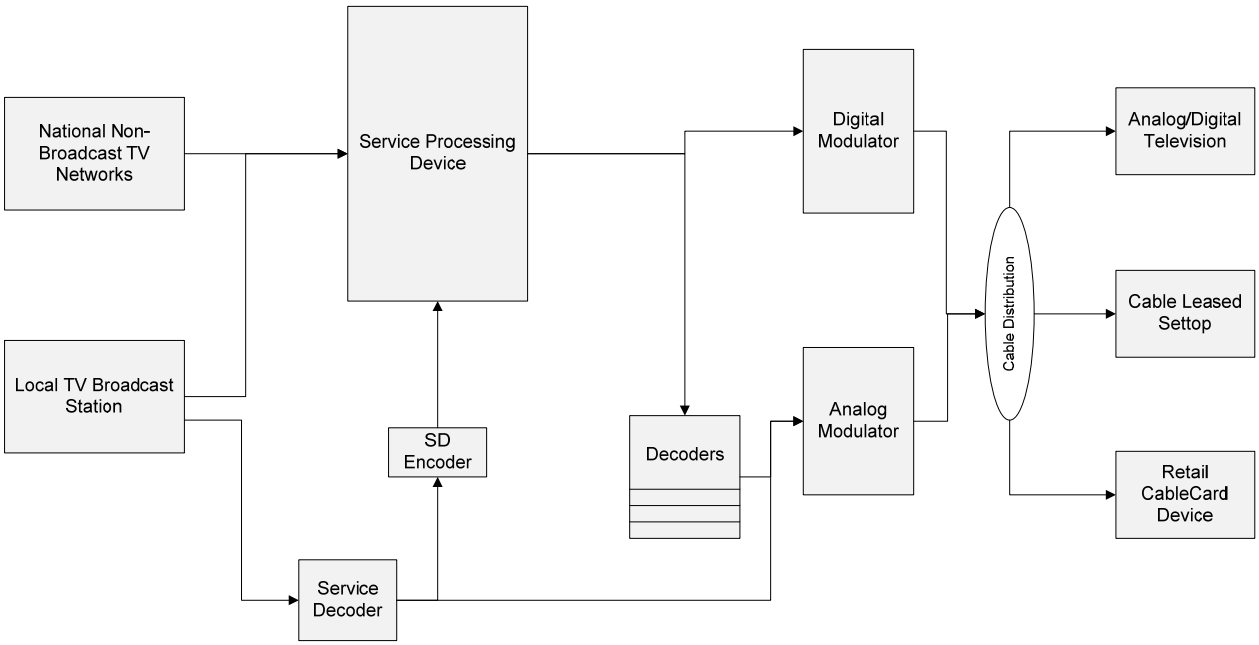

Figure 2 Detailed Distribution Path – Cable

Link to accessible textual description of Figure (2)

Cable systems use a mixture of fiber and coaxial cable to provide bi-directional signal paths to the customer. This "hybrid" fiber-coaxial (HFC) architecture effectively segments the cable system into a number of parallel distribution networks. The HFC architecture is beneficial to the cable operator because it improves signal performance and reliability, increases available bandwidth, and generally is easier to maintain than older architectures, which relied solely on coaxial cable. Although network designs vary, the HFC architecture in any particular community is typically based on a three-level topology, which includes a headend, one or more distribution hub(s), and multiple fiber nodes.

At the top level of hierarchy is the "headend." The headend serves as the origination and processing point for most of the signals essential to the programs and services carried in the cable system. Signals are received at the headend from local terrestrial television broadcast station antennas and/or direct fiber feeds, parabolic satellite antennas, terrestrial microwave links, and fiber links to and from the Internet, local exchange carriers, and other networks. Sometimes content is generated locally. Each of these signals is locally processed and converted for distribution to distribution hubs or directly to the customer.

Distribution hubs typically serve as intermediate signal processing points in the HFC architecture. Signal transport between the headend and the distribution hubs occurs over optical baseband and/or broadband links. Distribution hubs are fed by a redundant fiber optic "ring" architecture for service availability and reliability. Depending upon the size and design of the system, one or several distribution hubs may be placed on the ring. A distribution hub will then feed a number of nodes, depending on the customer density.

Individual fiber nodes provide the interface among the optical signal, the coaxial cable trunk, and distribution cables. The fiber node converts the optical signal into a radio frequency (RF) signal for the "last mile" distribution to the customer.

Two-way operation of the cable system is essential to support services such as high-speed Internet access, telephony and on-demand video services. The HFC architecture makes this aspect of the network practical and reliable for the cable operator because the cable system is segmented into small groups of homes passed. In the coaxial portion of the network, the signal is amplified in both the forward and reverse directions. This is achieved over a single cable by partitioning based on frequency spectrum.

Cable operators provide many channels of high-definition digital, standard-definition digital and analog programs to their customers. Standard- and high-definition digital programs are typically comprised of an MPEG-2 compressed video transport stream and AC-3 compressed audio transport stream. Analog programs are compliant with the NTSC ("National Television Systems Committee") analog color television standard used throughout North America.

6-MHz in-band channels are used to distribute the digital programs using QAM (Quadrature Amplitude Modulation) and the analog programs using VSB-AM (Vestigial Sideband Amplitude Modulation) for video signals and frequency modulation for audio signals. The 6-MHz in-band channels follow a consistent channel spectrum plan defined in ANSI-CEA-542-C, *Cable Television Channel Identification Plan*, and services may be carried either in the clear or protected using conditional access technology. Out-of-band forward data channel(s) and out-of-band reverse data channel(s) are also present in the system to carry service information to/from the set-top box or cable-ready television receiver.

Cable operators ingest programs containing video descriptions at the headend, or in some cases at a distribution hub. These programs typically come from two sources: local television broadcast stations and national non-broadcast program networks. For some content that is delivered in analog format, cable operators will encode it to a digital format, and conversely, for content that is only available in digital, some of that content will be decoded into analog. The purpose of these conversion steps is to assure that both analog and digital versions of the select content are available in the customer's home to display on various devices. The carriage of both analog and digital versions of select content is typically referred to as 'digital simulcast' within the cable community. Programs from these sources are locally processed and converted, or modulated, to a channel, then combined into a complete spectrum through a process known as frequency division multiplexing before distribution to the customer (See Figure ***2***).

Cable operators and many program networks and local broadcast affiliates currently support up to two audio streams for each program they distribute. The first audio stream contains the main or primary audio track for the program, typically in English; the second audio stream contains the video description or a second language (typically French or Spanish). The program network and local broadcast stations are relied upon to manage which service - video description

or second language -appear in a given second audio stream.[20]

In cases where the cable provider distributes a standard- or high-definition digital program containing video descriptions on one or more of its analog channels (*i.e.* digital simulcast), the descriptions or second language is typically transmitted in the analog SAP (Secondary Audio Program) channel. SAP is a monophonic, single, auxiliary audio channel for analog television that is part of the MTS (Multichannel Television Sound) standard developed by NTSC. In cases where the cable provider distributes a standard- or high-definition digital program containing video descriptions on one or more of its QAM (*i.e.* digital) channels, the video description is transmitted in a secondary audio stream.

The availability of the two audio streams in digital programming is announced by using a "descriptor" as defined in the relevant ATSC and SCTE standards. Descriptors are (generally) optional, variable-length data elements that can add standards-defined data elements to MPEG-2 private table sections. In a typical cable system, each of the two audio streams has been labeled by an *ISO_639_language_descriptor* in order to announce the availability of each audio stream as well as identify its language. For example, English language = "eng" and Spanish language = "spa". The cable operator's set-top box will identify the language value contained within the ISO-639 descriptor and select the appropriate audio stream based on the user's selection.

In a typical configuration, the audio stream containing the program's main or primary audio will be identified by an ISO-639 descriptor using a value of *eng*. The second audio stream containing video description or a second language will be identified by an ISO-639 descriptor using a value of *spa* (Note: the ISO-639 descriptor does not have an "audio service type" field or equivalent that can signal the presence of video descriptions). To access this second audio stream, a cable customer must select "Spanish" in the set-top box user interface. Providing video description in this way continues the longstanding practice in the analog environment of utilizing a single secondary audio channel in addition to the primary audio channel. Given the continued presence of analog versions of digital broadcast stations on cable systems, it also would result in a common approach to video description across analog and digital content sources. As noted above, local broadcast stations and program networks determine which audio service – alternative languages or video description – will appear in the second audio stream.

[20] In the case of a local television broadcast station, emergency information may also be carried in the second audio stream.

4. Distribution via Digital Broadcast Television to DBS Customers

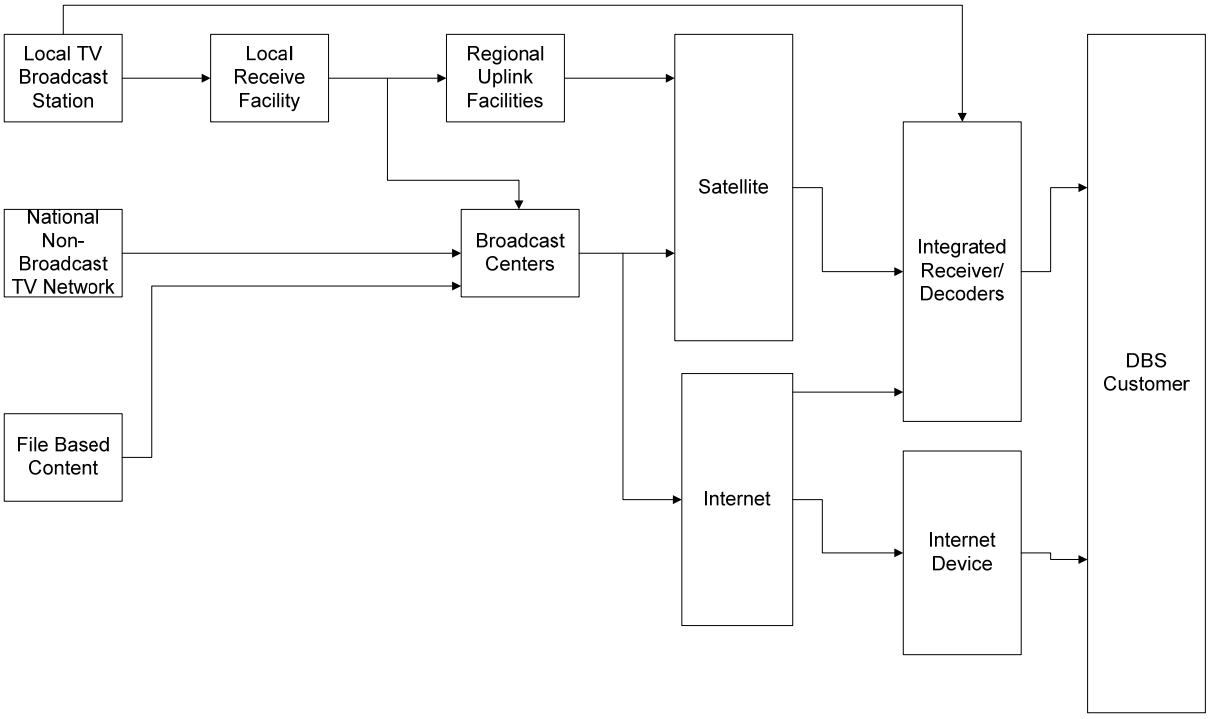

Figure 3 Detailed Distribution Path – DBS

Link to accessible textual description of Figure (3)

Direct Broadcast Satellite (DBS) service providers use geostationary communications satellites to deliver multichannel video programming services to customers. Like digital over-the-air (OTA) transmission, DBS satellite delivery is broadcast only. The two most popular US DBS service providers are DIRECTV and DISH. DBS digital television distribution architecture is shown in Figure 3.

Both DBS service providers provide a combination of SD (standard definition) and HD (high definition) channels to their customers. Standard definition channels are typically carried using MPEG-2 video compression and stereo MPEG-1 audio. High definition channels are typically carried using MPEG-4 Part 10 video compression and AC-3 audio. DBS systems combine video signals together into a single satellite transponder using VBR (variable bit rate) encoding and statistical multiplexing of multiple video sources. Compared to the alternative, constant bit rate encoding, VBR enables overall better video quality in a given amount of bandwidth. Corresponding audio signals are multiplexed using pre-allocated constant bit rates with the video signals into the broadcast signal. DBS service providers also transmit system and program data to customers including, for example, electronic program guide (EPG) data. The EPG allows the customer to see what content is available, and includes information that allows consumer equipment to receive and process the content.

National channels broadcast to all subscriber locations from DBS service providers' Broadcast Centers. National channels are typically received at the Broadcast Centers over satellite or direct fiber. When the DBS service provider has the applicable rights, advertisements may be inserted into the channel by specialized equipment in the Broadcast Center; otherwise, advertisements and program material in the feed received at the Broadcast Center are passed through. The channels are encoded into their respective signal formats (*e.g.*, HD MPEG-4 Part 10 video and AC-3 audio), multiplexed together, and then uplinked to a DBS satellite. National channels are broadcast across the whole country using DBS satellites' national (or "CONUS", for contiguous US) beams.

At least one, and typically both, of the DBS service providers carries local channels in each of the 210 Designated Market Areas (DMAs) as defined by Nielsen. A DBS operator is not required to provide local channels in any DMA, but if it decides to do so, it is required to carry any qualifying local station that elects mandatory carriage. Alternatively, a station can choose to negotiate a carriage agreement with a DBS operator. In order to collect local signals in a DMA, a DBS operator will have in place a Local Receive Facility (LRF) for SD programming and another LRF, possibly in a different location, for HD programming. In a number of DMAs, the LRFs are shared between DIRECTV and DISH. As there are no employees stationed at LRFs, the equipment therein is designed to operate with high reliability under consistent reception conditions. For its part, a local station must deliver its broadcast signal to the appropriate LRF with sufficient quality. After signals are received in the LRF, the video is re-encoded as HD MPEG-4 Part 10 and/or SD MPEG-2 while the audio is passed through as AC-3 for HD and/or re-encoded in stereo MPEG-1 audio for SD. The video and audio of the resulting channels are multiplexed together, delivered to one of a half-dozen or so regional uplink facilities, and then broadcast via satellite to customers in that DMA. Local channels are typically broadcast using spot beams, which focus the signal onto the desired region of the country. Unlike national channels, advertisements are not inserted by the DBS providers into local channels.

Additional content, such as Pay-Per-View movies, Video-On-Demand movies and TV shows, is typically delivered as data files to DBS service providers' Broadcast Centers. This file-based content may arrive on physical media or via a digital network (such as direct fiber). The content is processed into the relevant DBS signal formats and can then be delivered both via satellite broadcast and via Internet transmission to customers.

DBS service providers currently carry up to two audio streams for each of the SD and HD programming channels they distribute. When a second audio stream for a channel is carried by the DBS service provider, it remains configured in the DBS network full-time. The program network and local broadcast stations that deliver two audio streams manage what service appears in each stream. Typically, the first audio stream will contain the main or primary audio for the program in English and the second audio stream will contain Spanish language audio. The second audio stream might contain video description, a second language (e.g. French or Spanish), or potentially, emergency information; however, if none of these are available, the

program network and local broadcast station typically delivers its primary English audio in the second audio stream. This practice has been found to help prevent DBS customers from incorrectly concluding that there is a problem with the program or their equipment.

DBS Integrated Receiver/Decoders (IRDs) receive the satellite transmissions, decode them and deliver them to television displays via common TV interfaces (*e.g.*, HDMI). Many DBS IRDs (often referred to as "satellite set-top boxes") are also capable of receiving local over-the-air-broadcast signals and integrating information about those channels with other channels in the EPG. More recently, newer DBS IRDs can be connected via customers' broadband services to the Internet to receive file based content that is also integrated in the EPG. When a channel includes more than one audio service, DBS customers typically chose the alternate audio service through the IRD's menus. Many DBS IRD models include digital video recording (DVR) capability as well. DBS DVRs store content received from any source (whether from the DBS satellite, OTA, or Internet) in internal memory. Customers can then pause, rewind, and fast forward content as well as record content for later playback. DBS providers have also begun to provide some of their services to their customers' internet capable devices via Internet streaming and downloads.

5. Distribution via Digital Broadcast Television to IPTV Customers

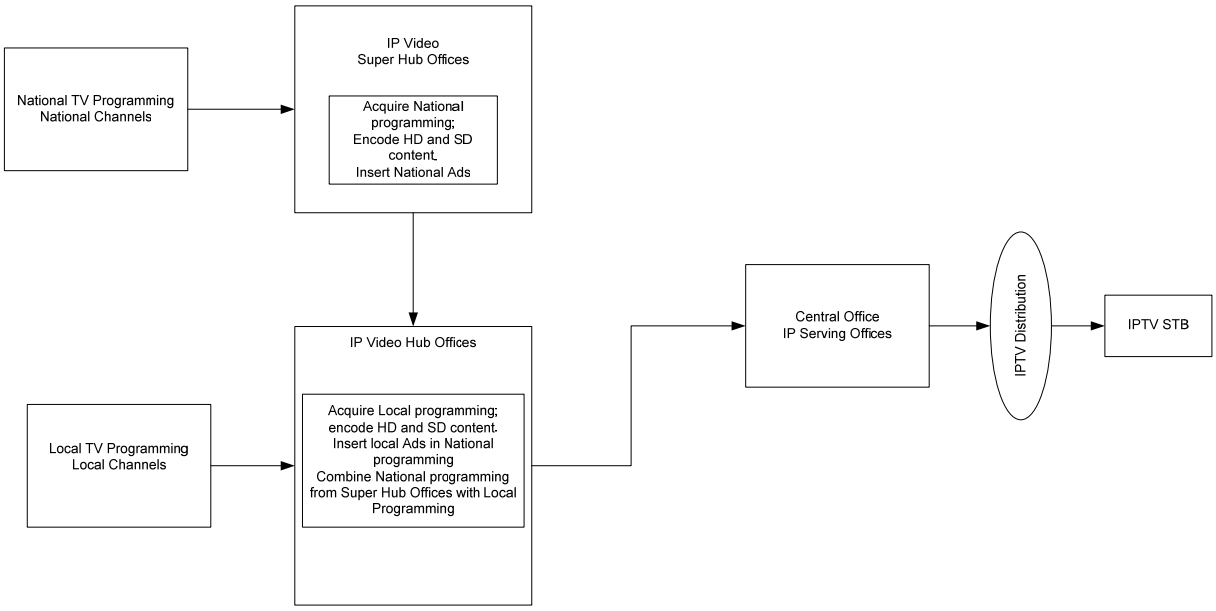

Figure 4 Detailed Distribution Path – IPTV

Link to accessible textual description of Figure (4)

Internet Protocol Television, (IPTV) is a technology used by some Multichannel Video Programming Distributors (MVPDs) to deliver TV services. For IPTV systems both the core (backbone) network and access networks rely on Internet Protocol. Internet Protocol (IP) networks enable the delivery of high-quality video, advanced functionality, and interactive applications. Video content is transported over a managed, two-way IP network and can be delivered to the customer's home using a combination of fiber and/or copper in the last mile of the access network. Use of Digital Subscriber Line (xDSL) technology enables delivery of video over copper phone line. Internet Protocol video networks are able to be efficient in utilization of network capacity when the access network is copper by only delivering to a customer's premises and Set Top Box (STB) only that content requested by the customer. When the access network is fiber the large network capacity makes it easier to multicast hundreds of channels to all STBs simultaneously.

The architecture of the typical IPTV system deployed in the United States features a hierarchical structure comprising of one or two national head- ends serving many regional head-ends across the country, which in turn distribute content via several community based offices to the homes served. The national head- end (*i.e.*, Super Head End (SHE) or Super Hub Office

(SHO)) is used for acquisition of national content via satellite receivers, processing the content as necessary (for example, encryption, format conversions between HD and SD encoded content, inserting national advertisement) and for the packing and formatting of the content for distribution to the regional head- ends (*i.e.*, Video Hub Offices, VHO) over a core IP backbone network (*i.e.*, Long Haul Network, LHN). Each VHO essentially serves a local market by extracting from the SHE/SHO distribution the specific national channels applicable to the region, aggregating it with local content (*i.e.*, local broadcast, PEG (public, educational, governmental) channels), inserting local advertisement, and serving On Demand content to subscribers in the region. The VHO is also the entry point for the reception of Emergency Alert System (EAS) messages. – EAS alerts are processed in the VHO and notification is distributed to subscribers' receivers located in the EAS zones served by the VHO.

Beyond the VHO, the architecture for the IPTV systems in the US differs. In the AT&T U-verse IPTV system, the aggregated output from the VHO generally passes through an Intermediate Office (IO) to a serving Central Office (CO) and finally to a neighborhood node, called a Video-Ready Access Device (VRAD) serving a set of local subscribers over a copper phone access line. In the Verizon FiOS IPTV system, the video leaving the VHO is transported over fiber to the CO and that traffic is merged with Internet IP traffic and plain old telephone service (POTS) traffic and carried over a fiber to the premises (FTTP) access network to the subscriber's home. This FTTP access network enables operation in a hybrid mode, i.e., a combination of the traditional Cable TV signal distribution concurrently with a purely IPTV signal distribution.

The functionalities provided by the IPTV systems require complex client-server software applications on the STB (client) as well as on the servers located in the VHO. Furthermore, because of the two- way nature of the IP network, every message between the client and server is encrypted to ensure protection of customer information and video content.

IPTV systems generally have finite limits on the services available, although other services that can be provided over an IP system, such as home broadband Internet access and VoIP services. For the video service itself, the limits of the MVPD's network architecture can impact the delivery of HD programming, On Demand services, and can lead to limits on the number of available channels.

IPTV networks generally use state of the art video and audio compression technologies to efficiently deliver simultaneous channels of IPTV programming per household. As for The AT&T's U-verse IPTV network, it supports delivery of a Surround Sound Main Audio Program (MAP) on HD channels and a 2- channel Secondary Audio Program (SAP) for both HD and SD programming. Often, the SAP is used to carry an alternate dialogue track for Spanish language audiences, weather information, or audio from a live event.

IPTV providers currently support up to two audio streams for each program they distribute. The first audio stream contains the main or the Surround Sound MAP. The second audio stream contains a 2-channel SAP, which is part of the MTS for both HD and SD programming. IPTV providers incorporate language descriptors and audio type in the PMT to describe the audio stream. SD and HD encoders transcode the program material into MPEG4 (H.264) Video and AC-3, AAC, or Dolby Digital Plus audio, creating the elementary streams for the IPTV system. Based on the capabilities of the encoders, the appropriate language codes, are set using ISO 639-2 (3 letters) or ISO 639-1 (2 letter) standards.

IPTV providers such as AT&T U-verse need to be able to determine what audio selection to present at the program level. Therefore, additional information is collected for each audio stream to determine its display name and position, such as the audio type, Packet ID (PID), audio stream format (AC3, PCM, MPEG1, etc.), the order that the stream appeared in the PMT, and other parameters. The audio type field can be provided by the encoder, and this field can represent special audio commentary for people who are hearing or visually impaired. When the audio type is provided, a string is appended to the language display name. For example, Language Descriptor 'eng' and Audio Type = 3 ("Visually Impaired commentary," can be displayed as "English Visually Impaired Commentary").

IPTV customers may select the Alternate Language or "Spanish" in the set-top boxes user interface to navigate to the audio stream containing either the Spanish or video descriptions audio track. This implementation uses a single secondary audio channel in addition to the primary audio channel, thus applying a common approach to accessing descriptions across the content source. The program network and local broadcast stations determine which audio service, alternative languages or video descriptions, will appear in the second audio stream.

IV. POST-CVAA DEVELOPMENTS FOR DELIVERY OF VIDEO DESCRIPTION

The preceding sections of this report outline the mechanisms for delivering video description service across a multitude of today's platforms and architectures. While video description services are already being provided to consumers today, the delivery methods may be further optimized to offer more robust usability and improved functionality.

The VPAAC arrived at the following findings:

- The majority of the content creation, distribution, and consumption chain remain limited to two discrete audio services. Of these two services, English language typically occupies the primary audio service, while the secondary audio service delivers alternate languages and/or video description service.
- Discussions within the VPAAC emphasized the importance of identifying which programs are scheduled to contain video description. How program providers currently provide information specific to their programs to intermediaries including, but not limited to, Tribune and Rovi, which in turn aggregate that information for content distributors, was also discussed. [21] The broadcaster or MVPD that utilizes the services of these companies provide this data to populate a program guide in the majority of receiving devices.
- Another step towards enhancing the usability of video description is the ease of navigation to the secondary audio service where video description service is typically carried. The capability of a receiving device to navigate to the

[21] Rovi Corporation and Tribune Media Services are two separate companies that collect, store and distribute television and movie listings information as a service to other companies including cable, satellite and IPTV video providers.

secondary audio service of a video program depends upon how that content is signaled in the encoded transport stream.

- Given that broadcast, cable, DBS, and IPTV providers use varied signaling mechanisms, a more unified signaling mechanism would likely yield more consistent results. Furthermore, at present, no method is currently in use for unambiguously signaling video description; and for delivering video description disambiguated from secondary language. Due to the complexity of the challenge, this situation is unlikely to change significantly in the near future.
- Industry members of the VPAAC stressed that any abrupt change to this signaling mechanism would likely result in significant legacy equipment issues, and subsequent user confusion. Any consideration of future signaling changes should take these legacy concerns into account.
- The Advanced Television Systems Committee standardized the AC-3 Descriptor which allows unambiguously identifying the audio service type (Main, description for the visually impaired, clean dialog for the hearing impaired, etc.) present in each audio stream. At this point in time, the features within this descriptor to unambiguously identify video descriptions are not in use to any significant extent. This may be an appropriate long-term solution for delivery systems that support AC-3, but needs to be considered in other delivery environment Standard Development Organizations. Building upon the ATSC's work, consumer receiving devices can be built in accordance with the recommendations in CEA-CEB21, *Recommended Practice for Selection and Presentation of DTV Audio*, a bulletin published by the Consumer Electronics Association in June, 2011. Such receivers offer the user an operating mode in which video description services are decoded whenever they are available. When a program has no video description, the regular main audio is decoded instead. Elements of the industry are discussing possible transition scenarios towards support for this solution in broadcast, but a pan-industry solution, which may not be the same as this, needs to be agreed. Appropriate treatment of video description will require not only establishing signaling method(s), but also infrastructure development. The latter is a long-term development that implicates not only devices, but also distribution infrastructure; and, potentially different solutions may need to be developed for service delivery in non-AC-3 environments.

- The VPAAC also discussed longer term implementation scenarios that could enhance video description service delivery capabilities. While potentially advantageous, several aspects must be considered when evolving content delivery, including: bandwidth efficiency, ease of creation, and user experience; all of which need wide industry adoption and acceptance. A few examples of these implementation scenarios were discussed and are as follows:

- Adding an additional audio channel. While considered by many to be a straight-forward approach, this option yields the lowest bandwidth efficiency, has limited user experience enhancements, and would require a large-scale re-tooling of the content delivery chain. Moreover, this approach is of limited utility until such time as consistent pan-industry signaling method(s) and infrastructure development are established.

- Utilizing the receiver to perform the mixing of primary audio with a video description narrative audio "sub-stream." This approach does offer potential for some bandwidth savings, but has not been adopted in the United States by either content creators or receiver manufacturers. Thus, a significant legacy issue exists without a workable migration strategy. No bandwidth savings or other benefits could be achieved until a full migration to this technology could be completed, which would require the very lengthy process of a complete change-out of all deployed receiving devices.
- "Late-binding," in which the content is sent only with the user-selected audio. This remains in very preliminary stages of development, is useful in non-broadcast (unicast) environments only, would require significant overhaul of many distribution platforms, and would require significant bandwidth increases to allow for more unicast delivery.
- Voice synthesis, which consists of the narrative track being generated by voicing software, is also in its infancy stages. While this is an exciting technology that could potentially yield many benefits for consumers and industry alike, there remain significant technical and performance concerns which would result in long lead times before this could be a viable solution.

Of course, the above list represents only a few examples of alternatives that are possible in future implementations, and are dependent upon coordination of every sector of industry, as well as being subject to adoption by the user community.

As the content distribution model continues to evolve, newer and perhaps even better options may begin to surface, and it remains critical to foster unrestrained innovation in the area of creation and delivery of video description services.

V. FINDINGS AND RECOMMENDATIONS

The VPAAC has come to the following findings and recommendations for certain aspects of creating and delivering video description. The VPAAC endeavored to provide the best guidance

to the FCC with the following section, while attempting to achieve a broad consensus on all critical issues. However, while some items did achieve consensus, there are a few items that did not. The following represents the primary points of focus, and the resulting consensus position, or where consensus was not achieved, the differing positions are noted.

A. Availability of information

The VPAAC does agree on the general principal to provide information about programming that is video described, however there were differing positions that highlight variations in both how expansive the information availability should be, as well as timing of implementation. Below are both the consumer and industry positions which highlight those similarities as well as the differences.

1. Consumer position

The consumer members of the VPAAC recognize the importance of making widely available information about what programs are video described. Consumers represented on the VPAAC explained that this is likely to be a critical factor in the successful roll-out of video description, and expressed a strong recommendation that the entities required to provide described programming pursuant to the FCC's rules, must also provide information about described programs on their websites, provide this information to programming information distributors such as Rovi and Tribune, and consider alternative ways of ensuring that blind and visually impaired consumers have access to such information. They agree with the need to ensure that blind and visually impaired consumers can easily find information about which programs are video described and recognized that both short- and long-term solutions will be needed.

Accordingly, the consumer members of the VPAAC make the following recommendations:

Beginning no later than July 1, 2012, the top four commercial television broadcast networks (ABC, CBS, Fox, and NBC), and top five national non-broadcast networks, (USA, the Disney Channel, TNT, Nickelodeon/Nick at Nite, and TBS) should provide information on their web sites indicating which programs they are airing with video description. To ensure that the information can be accessed, it must be provided in a manner that is accessible to and usable by individuals who are blind or have a visual impairment.

Beginning no later than July 1, 2012, commercial television broadcast stations affiliated with one of the top four commercial television broadcast networks (including those that must pass through video description provided by the network), and multichannel video programming distributor (MVPD) systems that serve 50,000 or more subscribers and carry one of the top five national non-broadcast networks, (along with MVPD systems of any size that must pass through video description on each broadcast station and non-broadcast network they carry,) should provide information on their web sites indicating which programs they are airing with video description. To ensure that the information can be accessed, it must be provided in a manner that is accessible to and usable by individuals who are blind or have a visual impairment.

The consumer members of the VPAAC are informed that the principal television programming information distributors, Tribune Media Services, and Rovi are capable of including description in their program data distribution. Therefore, they recommend that entities required to provide described programming should ensure that the principal television programming information distributors are informed in a timely manner of programs that include description. In addition, covered entities that make program guide information available should include information about video described programs.

To ensure the broadest possible distribution of program information, the consumer members of the VPAAC encourage covered entities to work with the disability community to establish means by which information regarding programs with video description can be made accessible, usable, and searchable online and through other means such as by telephone using an automated Integrated Voice Response (IVR) system.

2. Industry position

The industry members of the VPAAC recognize the importance of making information about what programs are video described widely available. Consumers represented on the VPAAC explained that this is likely to be a critical factor in the successful roll-out of video description.

The industry members of the VPAAC agree with the need to ensure that blind and visually impaired consumers can easily find information about which programs are video described and recognize that both short- and long-term solutions are needed. Accordingly, the industry members of the VPAAC make the following recommendations:[22]

- The top four commercial television broadcast networks (ABC, CBS, Fox, and NBC) and top five national non-broadcast networks (USA, Disney Channel, TNT,

[22] The intent of these recommendations is for industry best practices, not for additional FCC rules.

Nickelodeon, and TBS) should provide information on their web sites indicating which programs they are airing with video description.

- The industry members of the VPAAC understand that it is possible for the principal television programming information distributors, Tribune Media Services and Rovi, to include description information in their program data distribution. Therefore, they recommend that, to the extent they provide information to one or both principal television programming information distributors about their television listings, broadcast and non-broadcast networks that are required to provide described programming identify in a timely manner which of their listed programs include description. A longer term solution is expected when covered entities make program guide information accessible. (See the VPAAC report on User Interface.)

- To ensure the broadest possible distribution of program information, the industry members of the VPAAC encourages industry to work with members of the blind and visually impaired community to establish additional means by which information regarding programs with video description can be made accessible, usable, and searchable online and through other means.

B. Content signaling

The VPAAC agrees on the need for a more user-friendly mechanism to allow the carriage of multiple audio services, but did not come to agreement on the concept of a clearly defined timeframe to achieve such mechanism.

1. Industry position

The industry members of the VPAAC encourage industry to identify a pan-industry end-to-end solution to unambiguously identify critical elements of video description signaling, at least for AC-3 support environments, and a transition plan towards such, allowing achievable time-frames for signaling and for infrastructure enhancements.

They do not recommend any single format, protocol, or standard for multiple audio services., Moreover, many of the issues identified with respect to improving access to video-described programming may be addressed through other provisions of the CVAA, which call for industry to identify methods to enable audible access to user interfaces, program guides, and device menus that will enhance navigation to secondary audio services. The industry members of the

VPAAC conclude that as long as the video description service is readily and predictably available, no single technical method for carriage of video description is required.

2. Consumer position

The consumer members of the VPAAC recommends that the FCC consider how best to facilitate a transition, including timelines, for an orderly transition to support a strategy by broadcasters, MVPDs and consumer electronics manufacturers to deliver multiple simultaneous ancillary audio services, so that both Spanish (or other alternate languages) and video description could be provided for the same program.

C. Point of contact

The VPAAC discussed consumers' interest in having MVPDs and local broadcast stations designate a point of contact to address issues that may arise relating to the provision of video described programming. Consumers pointed to the FCC's rules regarding a contact person for closed captioning issues as the model. While this issue is not part of the VPAAC's mandate, consumers' desire for assurance that distributors will be able to respond to inquiries about video description is recognized. Therefore, the VPAAC recommends that MVPDs work to ensure that appropriate customer care representatives are able to assist blind and visually impaired customers in resolving issues relating to access to video description, e.g., when the description audio isn't available. The VPAAC further recommends that MVPDs and local broadcast stations establish processes to address both immediate and longer term customer concerns regarding their video description service. For immediate customer concerns, the VPAAC recommends that MVPDs and local broadcast stations provide either the customer service number, or alternatively, contact information and procedures similar to those articulated in the Commission's closed captioning rules, and that any calls or inquiries received by MVPDs and local broadcast stations using this contact information where a distributor is not immediately available should be returned or otherwise addressed within 24 hours. *See* 47 CFR Sec. 79.1(i).

D. Persistence of secondary audio service

The VPAAC recommends that best efforts be undertaken to ensure that the secondary audio channel should not be programmed with silence. It is further suggested that main program audio be inserted in the secondary channel when neither descriptions, nor an alternate language, nor

emergency information (as proposed by the VPAAC report on Emergency Information) are present.

E. Descriptions on internet content

The VPAAC was not able to fully explore the topic of delivery of video description over the internet. The brief summaries here illustrate areas of agreement among many members that further effort would be required to develop internet technologies to accommodate consistent delivery of programming with description. Areas of disagreement include the extent of technology development, as well as what programming could be provided over the internet with description. Finally, we include an alternative view that internet delivery is beyond the scope of the VPAAC.

1. Consumer position

Internet streaming technologies used for content distribution will have to add features to move to a state of being "Description Ready." Many streaming platforms, especially those that use adaptive bit-rate technologies, will have to be modified and add functionality to support the incremental audio component(s) needed to become capable of storage, serving, transport, and user selected playback of video description.

Currently, many streaming systems have the limitation of packaging one video with one audio stream as the final asset for serving to customers.

Today, the *one audio* approach is the primary audio for delivery over the sometimes-bandwidth-constrained Internet pipe. Enabling user selection of video description would require new system designs, some of which are in development and exploit the new HTML5 web architecture, so that the end-user could select video description on an asset-by-asset basis with the server sending the proper video and selected audio to each user. Once new system designs are developed and deployed, content would need to be packaged with added video description tracks that may have been derived from and transmitted via broadcast or other means. We recommend such deployment, once developed and fully tested, focus on new (post-rule) content initially, applied to the programming subject to FCC video description requirements for broadcast and cable networks.

2. Industry position

Internet streaming technologies used for content distribution will have to add features to move to a state of being "Description Ready." Many streaming platforms, especially those that use adaptive bit-rate technologies, will have to be modified and add functionality to support the incremental audio component(s) needed to become capable of storage, serving, transport, and user selected playback of Video Description.

Currently, many streaming systems have the limitation of "packaging" one video with one audio stream as the final asset for serving to customers. Today, the *one audio* approach is the primary audio for delivery over the sometimes-bandwidth-constrained Internet pipe. Enabling a user's selection of Video Description would require new system designs – so that the end-user could select Video Description on an asset-by-asset basis with the server sending the proper video and selected audio to each user. Once new system designs are deployed, their already-packaged content archives will have to be re-packaged before being capable of adding the Video Description 'tracks' that may have been transmitted via broadcast or other means. Even presuming the Video Description tracks are 'as-aired' in a previous exhibition, the repackaging of the content archives is a significant work effort.

3. Alternate Industry position

At least one Industry member believes that it is premature and unnecessary to make recommendations or suggest technical approaches to video description on the Internet. The CVAA does not grant authority to the FCC in this area. As a result, this report should focus on the complex issues at hand in order to meet the required delivery of video description through broadcast and MVPD systems.

F. Description quality

The VPAAC has spent considerable time and effort on the topic of description quality. While there was agreement on the fact that some best-practices should be included in this report as a guide, the VPAAC was unable to come to a consensus on recommendations regarding quality of description.

1. Consumer position

Consumer members with vision loss believe that description quality is very important in the successful implementation of video description. Consumer members submitted information about description quality gathered from best-practice guidelines. The guidance information and further resources are listed below. In addition, given the importance of ensuring high-quality video description, the FCC is urged to consider undertaking a Notice of Inquiry on issues related to description quality.

The essential elements of video description have been developed and refined by various providers for nearly 30 years. These basic elements of effective video description include:

- Describe what is most essential for the viewer to know in order to understand and appreciate the scene being described.
- Prioritize description so that it can be provided within the pauses between program dialog or critical sound elements including sound effects, music, and program ambiance.
- Insert descriptions of key visual elements (actions, costumes, gestures, facial expressions, scene changes, and onscreen text) during pauses in program dialog.
- To ensure consistency, use the language in the program as a guide to vocabulary, sentence structure and pacing.
- Main program audio and added video descriptions should be properly and actively mixed so that the descriptions can be clearly heard when voiced and so that main program audio quality is maintained. Every effort should be made to produce a clean finished copy, i.e., removing extra ambient sounds and ensuring descriptive tracks are appropriate to the action with which they are aligned.

Other guidance that has been found to be effective for description includes:

- Clarity: Use "everyday" terms, and limit the use of slang or jargon unless appropriate to the content/image being described.
- General to the Specific: Start generally, creating a context, then move to details to enhance understanding and appreciation.
- Vocal delivery should be appropriate to the nature of the material being described. Narrators' voices must be distinguishable from other voices in a production, but they must not be unnecessarily distracting, as with recognizable celebrity voices or inappropriate accents.
- Identify color when it is vital to the comprehension of content; use basic colors and shades.
- Describe objectively without personal interpretation, censorship, or comment.

It is recommended that the entities covered by the mandate for video description, ABC, CBS, Fox, NBC, USA, the Disney Channel, TNT, Nickelodeon/Nick at Nite, and TBS, provide description that adheres to best-practice guidelines.

2. Industry position

Industry members of the VPAAC recommend that entities affected by the requirement provide good quality video description. The following resources represent current published and ongoing discussions regarding best practices in video description:

3. Resources

The following resources represent a consensus list of references among the Consumer and Industry positions.

- Described and Captioned Media Program - Description Key http://www.dcmp.org/descriptionkey/

  The Described and Captioned Media Program, with funding from the U.S. Department of Education, established the "DCMP Description Key." The Description Key is a set of guidelines based on a review of best practices to facilitate video description of educational media.

- ITC/OFCOM Guidance for Standards on Audio Description http://www.ofcom.org.uk/static/archive/itc/itc_publications/codes_guidance/audiodescription/index.asp.html

  This document, produced in the United Kingdom in 1996, represents the first guidance provided for the production of audio description.

- American Council of the Blind "Audio Description Project" Guidelines/Best Practices http://www.acb.org/adp/ad.html

  A work in progress, this document is the first effort in the United States to produce user-focused guide to quality audio description. The document is a collation of commentary from a range of description guidelines produced around the world and is coordinated by the ACB's Audio Description Project committee and a group of experienced description consumers and practitioners. Further, the document is available on a "wiki" website for commentary by description enthusiasts around the world.

- Accessible Content Best Practices Guide for Digital Environments: Descriptive Video http://www.mediac.ca/pdf/DVBPGDE_Version_1.docx

  This document, produced in Canada, is the most recent effort to establish best practices and will be updated and reviewed on an ongoing basis.

- Effective Practices for Description of Science Content within Digital Talking Books http://ncam.wgbh.org/experience_learn/educational_media/stemdx

  WGBH's National Center for Accessible Media (NCAM) received a National Science Foundation (NSF) grant to research and document effective practices for providing meaningful descriptions of non-text science content for post-secondary students or scientists who have vision loss, in collaboration with the American Foundation for the Blind. The research resulted in publication of "Effective Practices for Description of Science Content within Digital Talking Books"

APPENDIX A - CALM ACT

Previous sections of this report detail the method(s) for how video description is created by the content producer, and how each broadcaster and MVPD carry video descriptions and associated metadata to the consumer. It is important to note that those parties adding descriptions to an audio track may need to consider the requirements set forth in the CALM Act ("Commercial Advertising Loudness Mitigation Act") [23], and the associated FCC Rules. [24] The CALM Act required the FCC to incorporate into its rules on December 2012 a reference to the Advanced Television Systems Committee's ("ATSC") recommended practice A/85 *Techniques for Establishing and Maintaining Audio Loudness for Digital Television*.[25] All broadcast advertising-supported programming, including that containing video descriptions, will need to conform to the techniques defined in ATSC A/85 to control audio loudness in commercial advertisements. Further, it is understood that other programming is required to comply with ATSC A/85 by specific agreements and arrangements between individual MVPDs and content providers.

The method for adding video descriptions to programming is achieved by creating a unique audio channel composed of the original dialogue and sound effects with the additional dialogue of the descriptions. This inclusion of additional dialogue to the original soundtrack will yield a unique audio channel with characteristics that differ from the original soundtrack. It is expected that these new characteristics will need to independently conform to ATSC A/85. What is unknown at this time is if, after the inclusion of the video descriptions, the narrative audio will generally be considered to be the "anchor element"[26] for purposes of compliance with ATSC A/85. If so, the audio level of the narrative channel may be different from that of the primary audio channel, and adjustments in program loudness in the narrative channel may need to be made to ensure it conforms to ATSC A/85. It is expected that programming containing video descriptions can easily meet the CALM requirements so long as the method of loudness measurement accounts for the additional narrative audio and unique audio mix of the video description service.

---

[23] The Commercial Advertisement Loudness Mitigation ("CALM") Act, Pub. L. No. 111-311, 124 Stat. 3294 (2010) (codified at 47 U.S.C. § 621).

[24] 47 C.F.R. §§73.682(e) and 73.8000 (broadcast rules) and 47 C.F.R. §§76.602 and 76.607 (MVPD rules).

[17] *In re Implementation of the Commercial Advertisement Loudness Mitigation (CALM) Act*, Notice of Proposed Rulemaking, MB Docket No. 11-93 (May 27, 2011) ("Notice").

[26] A/85, Section 3.4 (Definitions) defines "Anchor Element" as "Anchor Element – The perceptual loudness reference point or element around which other elements are balanced in producing the final mix of the content, or that a reasonable viewer would focus on when setting the volume control."

APPENDIX B - ACCESSIBLE TEXT FOR DIAGRAMS

This section provides an accessible textual description of the figures within the body of the document.

A. Figure 1. High Level Distribution Summary

Described programming begins with three possible entities:

- National Broadcast TV Network
- Syndicated Programming
- National Non-Broadcast TV Network

Both the National Broadcast TV Network and Syndicated Programming flow to the Local TV Broadcast Station; Syndicated Programming may also be sent to a National Non-Broadcast TV Network. National Non-Broadcast TV Network programming and Local TV Broadcast Stations feed to three types of distributors:

- Cable Distribution
- DBS (satellite) Distribution
- IPTV (Internet) Distribution

In addition, a Local TV Broadcast Station may also go out "Over-the-Air"; these local feeds, along with National Broadcast TV Network and National Non-Broadcast TV Network signals, may flow directly to WEB Based Video Sites (represented by dotted lines on the flow chart).
Finally, all signals are received by the TV Audience via:

- CECB (Coupon-Eligible Converter Box) STB (Set-Top Box) for Over-the-Air broadcasts (these programs may also be delivered directly to the TV Audience)
- Cable STB for Cable Distribution
- DBS STB for DBS Distribution
- IPTV STB for IPTV Distribution
- PC/Internet Device for WEB Based Video Sites

Link back to document

B. Figure 2. Detailed Distribution Path – Cable

The diagram depicts video described programming sources into a typical cable headend, processing flows within the headend, and final distribution out of a cable headend to the cable customer. Video described programming sources fit into two categories:

- Local Broadcast TV
- National Non-Broadcast TV Network

The signals from both categories flow into the headend service processing device which multiplexes digital services together into transport streams and inserts advertising spots on select channels.

The Local Broadcast TV station sources flow to additional headend equipment which decode and convert the broadcaster's digital signal into an analog version of the signal. The analog signal then flows to an analog modulator and into a standard definition digital video encoder, which in turn flows to the central service processing device.
All of the digital signals from the service processing device flow into digital modulators. Many of the standard definition non-broadcast TV networks will also flow from the service processing device to service decoders which decode these digital signals to analog signals. These analog signals in turn flow to the analog cable modulators.
The outputs of both the analog and digital modulators are combined and distributed to the cable customer over the service provider's hybrid fiber/coaxial cable distribution system until final reception of the signal at the customer's premises.
The devices that receive this programming fall into three categories:
- Analog or Digital televisions
- Cable Leased set top boxes
- Retail set-top box devices

Link back to document

C. Figure 3. Detailed Distribution Path – DBS

Described programming begins with three sources:

- Local TV broadcast stations, whose signals are received either over-the-air (OTA) or by direct fiber
- National Non-Broadcast TV Networks, whose signals are received either by satellite or direct fiber
- File Based Content, which is delivered to a Broadcast Center either by physical media or via a digital network such as direct fiber

The diagram shows, for these three sources, the subsequent distribution paths to the end customer:

- Local TV broadcast station signals are collected at the individual DMAs, processed in Local Receive Facilities and then delivered via direct fiber to a Broadcast Centers or Regional Uplink Facilities to be uplinked to specific satellites
- National Non-Broadcast TV Networks and File Based Content are both processed at a Broadcast Center and uplinked to specific satellites

The satellites receive the various transmissions and retransmit the signals back down to customers' IRDs.
The diagram shows that IRDs (in some cases) are capable of also receiving digital OTA signals.
The diagram also shows that programming services may also be delivered from a Broadcast Center via the Internet to either a customer's IRD or Internet Device.

Link back to document

D. Figure 4. Detailed Distribution Path – IPTV

The diagram depicts video described programming sources into a typical IPTV Network Distribution system, processing flows from the Super Hub office to the Video Hub Regional offices, and the final distribution of the IP Central Offices to the IPTV customers. Video described programming sources fit into two categories:

- National TV Programming
- Local TV Programming

The video and audio signals from the National Network Programming category flows into the Super Hub Offices (SHOs) for Processing which includes the Video acquisition, processing and encoding of HD and SD content into the appropriate format and the insertion of national ads. The Super Hub Office (SHO) aggregates all the national content and processes it for distribution to the Video Hub Offices (VHOs).

The Local TV programming sources flow into the VHOs where the HD and SD content is received and encoded to the proper format. Each VHO serves a local market area and performs many of the same functions for locally generated programming that the SHO performs for national content, e.g., acquisition and processing of local channels and the insertion of local ads. The VHOs also serve as the distribution points for video on demand (VoD) content.

For AT&T, the aggregated output from the VHO generally passes through an Intermediate Office (IO) to a serving Central Office (CO) and finally to a neighborhood node, called a VRAD in AT&T's architecture, serving a set of local subscribers. The devices that receive the programming are the IPTV set-top boxes.

Link back to document