CONSIDERATIONS WHEN DELIVERING CABLE TV TO IP CONNECTED CONSUMER ELECTRONICS

Mark Francisco
Comcast Cable Corporation

Abstract

The number of connected consumer devices is expected to grow to more than 2.5 billion worldwide by 2014. The ability to directly deliver traditional Cable TV services to the most relevant classes of devices can provide greater opportunity, value and consumer satisfaction. The variation across and within multiple device classes present challenges when adapting content and services to the network link and capabilities of the video player.

This paper will introduce a variety of relevant device classes, their user interaction presentation and playback ecosystems. It will catalog cable TV features (i.e. closed-captioning, iTV, SAP and others) and discuss mechanisms to adapt to the device classes. An explicit objective is to maintain a good quality television viewing experience equivalent to that of a STB connected TV.

OBJECTIVES

Television as an Application

Cable television and similar Multi-channel Video Programming Distributors (MVPD) are accustomed to the concept of an Additional Outlet (AO) delivery model. AO's are typically extensions of the primary screen television service to additional televisions in the home and carry the same programming and features. The AO may offer a subset of the primary outlet services if the Cable Receiver or Television Receiver has limited capabilities. The growing popularity of online video viewing on connected consumer devices suggests direct support of an AO model to these devices is of consumer interest and benefit.

Connected consumer electronics include one or more IP-connection technologies, and the ability to integrally display video or support an external display through HDMI or similar digital interface. The devices include a user interface controlled by a remote control, keyboard/mouse, integrated buttons or touchscreen. The devices navigate and play video through an applications interface. The application environment may be standard or proprietary, with open or controlled access by service providers. Some of the most notable and popular video-supporting application environments today include (but not limited to):

1) Adobe Flash running within a PC browser
2) Apple iOS
3) Android
4) HTML 5
5) Microsoft Silverlight running within a PC browser

Adapting cable services to these environments requires a change of the traditional definition of television as a device or service to television as an application.

Principles of Television

With the objectives of extending MVPD services to connected consumer electronics, the characteristics that define television should be discussed. The following characteristics should be adapted in keeping with the traditional delivery model:

Television is available continuously and without interruption – Consumers expect television to arrive with the powering on of the device and remain until it is powered off. This characteristic will stress battery-operated devices and IP networks without broadcast capability. Heuristics should be employed to provide an instant, "always on" experience while conserving power and network capacity when the user is not actively viewing.

Television is Live – Live events are the cornerstone of television and maintain its highest viewership. Live television has explicit scheduling, minimal propagation delay and may experience high concurrent viewership. Balancing robust delivery methods, which may employ buffering against the desire for minimal delay from live, is a challenging task, particularly when delivering over unmanaged (capacity) network segments. The number of live television events delivered over the Internet using unicast delivery has increased, aided by emergent adaptive delivery methods.

Television is Multi-channel - The predominant behavior of television viewers is to watch programming at the first time of airing on the host network. Despite the lack of relevance of channel numbers identifying the means of tuning in the programming, established norms have maintained identity of networks via their historical frequency assignment. In other words, channel surfing, while increasingly supplanted by more effective discovery methods such as electronic program guides, search and recommendations, remains core to television viewing behavior. The requirement to zap, or rapidly change channels may be addressed through application or delivery techniques or more likely obsoleted through more effective discovery methods.

Television is Immersive – In the 20th century, television emerged as a focal point for living room gatherings of family and friends to watch live radio shows adapted to video. It subsequently offered prime-time entertainment, news and live sports in appointment-based viewing events that individuals would consider when planning their day. Only recently has television viewing become somewhat personal and associated with multi-screen, multi-tasking consumption. To satisfy traditional viewing habits, television should occupy the user's primary attention and deliver an immersive soundstage.

Television has Features and Control Requirements – Television programs are delivered with synchronous data and multi-program audio. The synchronous data may include teletext, alternate audio, program metadata, enhanced applications, alternate programming insertion triggers and control instructions. Programming may be copy protected. These features may include closed-captioning, emergency alerts, advertising insertion, descriptive audio, alternative language, parental advisories, rights management and interactive television. Several elements may be required through programmer contractual agreements and applicable transmission regulation.

While the television experience may be tailored to the specifics of a consumer device, usage and environment, adaptation of the core principals is possible and worthy of technical definition.

TELEVISION APPLICATION PLATFORM

Television has historically been delivered to NTSC or ATSC-based receivers with minimal variation. Viewer operation has moved from mechanical elements to remote controls and only recently been augmented with web-based discovery and control. Consumer devices offer challenges to traditional interaction and consumption and provide the opportunity to accelerate creation of new experience models.

Environmental Considerations

A number of factors differ between conventional HFC-based cable distribution of QAM video and IP delivery to consumer electronics. Most notable is the difference between managed and predictive capacity of MPEG-2 MPTS delivery over QAM versus the varying and contention-based method of delivery of IP networks. The second key difference is the variation in consumer electronics regarding video display capability, including resolution, frame rate, aspect ratio, and audio/video CODEC. Translation of video and audio formats is a requirement for most device categories. This may involve both spacial and temporal changes for video, and dynamic and encoding changes for audio.

User behaviors with connected devices may differ significantly from traditional television. Personal portable equipment such as smartphones and tablets are designed for mobile activity in short, but frequent intervals. Brief start-up time is a key requirement, as well as the ability to adapt to frequent changes in network and occasional loss of connectivity.

The third and perhaps most challenging area of adaptation is the concept of delivery to unmanaged devices. Consumer devices have a variety of operating systems, software stacks, application provisioning methods and native video pipelines. Ensuring application integrity and content security may involve detection of the characteristics of the device as configured and creation of secure enclaves within the device for provisioning of security related elements such as content keys.

TECHNICAL REQUIREMENTS

The reference architecture of Figure 1 proposes a system to realize the principles and environmental conditions. Elements are either under service provider control or customer supplied; therefore the two categories of requirements will be discussed separately. The requirements discussed below are derived in support of linear video applications, which may include broadcast content or content originating from video on demand servers. File-based delivery of content to consumer devices is a mature application and not described. While the technical solutions are proposed and discussed, the implementation timeframes and costs are not.

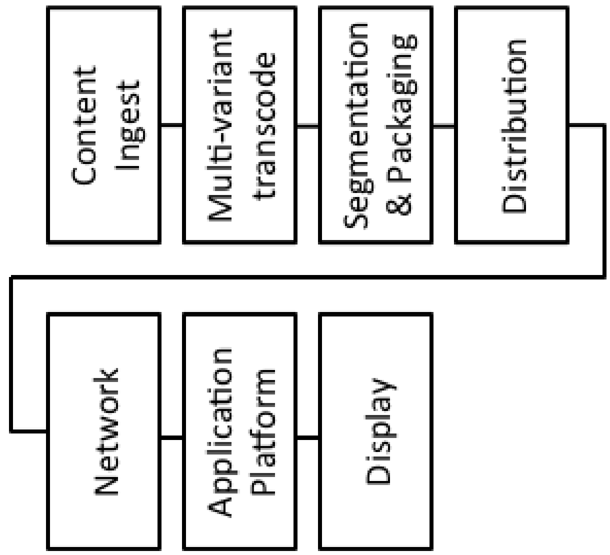

Figure 1. Reference Architecture

Service Provider Elements

Content arrives from many sources with most formatted for broadcast or Cable VOD distribution. This primarily MPEG-2 content requires transcoding into multi-variant H.264 profiles to reach the widest variety of devices and network conditions. Live content delivery across unmanaged networks requires adaptive delivery. The number and range of stream variants will be discussed in conjunction with the Technology Options section.

To maintain all associated metadata oriented features of the incoming video streams, the stream-associated information must be mapped from incoming transport to delivery technology to the client player protocol. Multiple translations will be presented within the Technology Options section. Content encoding is the element of the network with the lowest scalability and greatest use of custom hardware, lending to its separation from other elements in the delivery path.

The delivery path for video to consumer devices is expected to involve a home network, often wireless. TCP delivery of video content will be used, as home routers do not adapt well to UDP or multicast transport. HTTP delivery is widely adopted for video delivery to Internet connected devices as it is stateless, error resilient and traverses NAT.

Content security is an integral part of segmentation & packaging. Most content security clients rely on three main elements; content encryption, key management and authentication. Elements that need to be considered when defining a content security solution include integrity, ability to cache or store the content and device compatibility. Many popular content security systems rely on one of two AES-128 block cipher methods, cipher-block chaining (CBC) or counter (CTR). Both modes need to be considered due to fragmentation of support in the most popular connected devices.

Choice of key management and authentication systems is often dictated by the capability of the client. While a single key management / DRM system may be considered, an alternative is to unify the content encryption and adapt the key delivery to the client's native capability.

Customer Supplied Elements

The most widely deployed Internet connected video devices are PCs. This is followed by and soon to be eclipsed by smartphones. The third most popular connected video device is the game console. Other quickly growing categories include tablets, connected TVs and connected Blu Ray Players. Digital Media Receivers such as the Apple TV and Roku round out the most popular and relevant device list. With little exception, all include either an integrated display or support an HDMI-connected television.

Categorically, small-screen devices such as smart phones support low resolutions of 480p30 or less and require content encoded in H.264 baseline profile. Mid-size devices such as PCs, Tablets and digital media receivers typically support up to 720p30 and support H.264 baseline or main profile. Some high end PCs, digital media receivers and most connected TVs and BluRay players support up to 1080p60 resolution, which requires H.264 High Profile.

Content security in consumer devices is implemented as a link protection such as DTCP-IP or TLS (SSL), through application-based DRM or natively in the platform. Very few connected devices have native DRM capability that is available to service operators.

The application environment is an important aspect of adapting services to the consumer device. Most devices tie their application-provisioning environment to their native OS. The most popular application environments include Apple iOS, Android OS, Windows and MacOS. The application store process defined by Apple iOS and game consoles are the most constraining due to the requirement for certification prior to launch, while others are very flexible. In addition to application framework, many devices support presentation frameworks such as Silverlight or Flash.

Consumer-supplied elements provide the target platforms for delivery of Cable TV. The following sections "thread the needle" with content transformation, delivery method and application to complete the adaptation.

TECHNOLOGY OPTIONS

The core element required to deliver cable TV to connected consumer devices is the adaptive delivery technology. With the derived requirements stating the delivery will utilize HTTP, a number of ecosystems are candidates and will be introduced here.

HTTP Live Streaming

HTTP Live Streaming (HLS) was designed by Apple and submitted for standardization to the IETF. It is the most commonly used smartphone and tablet media framework in the United States and has been supported on all Apple iOS devices since the introduction of the iPhone 3G in July, 2008. HLS was initially used for streaming of file-based assets from sites such as YouTube and ABC but has since been applied to live video delivery. All video delivered to iOS devices must be presented using HLS in order to gain approval for 3G network use. Other mechanisms used for WiFi delivery include progressive MP4 download and PIFF delivery. Apple additionally limits content security to AES-128 CBC stream encryption with TLS key delivery.

HLS provides a robust delivery mechanism that can traverse most networks, but suffers from considerable latency and lacks features such as seamless content splicing and trick modes. It has not seen widespread adoption beyond iOS and Mac OS. HLS should be considered where necessary, as iOS devices are currently the most relevant for live video delivery, based on deployment and user interest.

A key benefit of HLS is the use of MPEG-2 transport stream for its container. This provides easier conversion from cable TV services but more importantly provides compatibility with all methods of carrying metadata such as captioning, ad triggers, and content advisories.

DLNA

The Digital Living Network Alliance (DLNA) specifies a set of device and media profiles that allow sharing of media between content sources and playback devices within the home. The Alliance assembles externally defined standards into interoperability guidelines and provides a certification program for manufacturers to receive approval to use the DLNA logo.

The DLNA is working on additional guidelines that will allow MVPDs to adapt subscriber content for delivery to DLNA players. These guidelines, which are being developed by the Alliance with service provider input, are planned for imminent release.

More than 8,500 devices have been DLNA certified but currently few support DLNA content protection. Adoptions of future guidelines are necessary to ensure DLNA certified devices are able to receive subscriber content. Cablelabs' OpenCable Home Networking Tru2Way (OCHN) extensions are DLNA device compatible, offering a good transition from in-home gateway delivered content to network IP sources of video.

The DLNA does not currently specify any adaptive delivery methods. It is anticipated that DLNA would include adaptive formats in updates to its guidelines.

Flash Streaming

Flash adaptive streaming is currently the most widely deployed Internet video technology as it is supported on 99% of PCs. In most cases, the stateful RTMP streaming mechanism is not usable through firewalls. The system defaults to a progressive download model in these cases. Flash adaptive streaming has found limited adoption in smartphone and tablet products. Progressive download methods are not well suited for live content delivery.

Adobe has recently released HTTP Dynamic Streaming (HDS), a version of Flash Streaming that supports dynamic streaming over HTTP connections. HDS requires Flash

10.1 or AIR v2 or later and is incompatible with RTMP origin servers. A content security technology called Flash Access is available for HDS streaming.

Flash support must be considered due to its widespread availability on PCs and integration into the Flash graphics presentation environment, although Silverlight is seeing growing adoption as an alternative. Advertising is almost exclusively distributed using Flash technologies.

WebM

Google has created an open source media framework entitled WebM. WebM is supported in Chrome Browsers, Android OS, and the Gstreamer open source media player. It includes support for an open source CODEC created by On2 Technologies called VP8. WebM has a plug-in structure with a relatively small group of components available for file-based on demand streaming. Live tools and features are not available at this time.

MPEG-DASH

MPEG-DASH is a multi-media delivery platform based on HTTP. DASH will deliver both MPEG-4 file and MPEG-2 TS based content. It is likely that HLS, DECE, 3GPP, and PIFF / SmoothHD will be supported by the DASH standard when completed. The MPEG-DASH specification is currently an ISO Draft International Standard (ISO/IEC DIS 23001-6) with an anticipated completion in July 2011 and release by end of year.

MPEG-DASH addresses HTTP delivery of streaming video/audio with adaptive features and supports both live and file-based streaming. It standardizes the container description information to ensure interoperability between servers and clients. DRM is not explicitly defined although support of DRM metadata is included in the description, making this a complement to the DECE content security framework.

MPEG-DASH shows promise as a useful and flexible method of adaptive delivery to the widest variety of consumer devices. DASH is proposing an HLS compatible profile in addition to a PIFF compatible profile.

IIS Smooth Streaming

IIS Smooth Streaming is a component of Microsoft's Protected Interoperable File Format (PIFF), a common file structure and adaptive delivery method for Silverlight and other HTTP clients. While PIFF is an open standard as per Microsoft's Community Promise license, Smooth Streaming is currently an element bound to IIS Origin servers. PIFF is based on the ISO MPEG-4 file format specification. Metadata may be carried as timed tracks. This will require mapping elements such as captions, ad triggers and advisory data from the traditional MPEG-2 transport stream mechanism to XML-based timed tracks.

PlayReady is the default DRM for IIS Smooth Streaming and supported in the Silverlight environment. Currently, Silverlight is implemented in Windows and MacOS, Xbox and PS3, many connected TVs and Windows Mobile Smartphones, but currently lacking in iOS and Android devices.

Technology Option Summary

A number of viable technology ecosystems have been introduced here, all with comprehensive feature sets but incomplete market availability and in some cases incomplete definition. The market is undergoing rapid, evolution and introduction of candidate technologies, further complicating selection. PIFF and MPEG-DASH show the most promise from a broad-scale adaptability but their ultimate adoption is unknown. Due to difference in core encryption methods between HLS, which specified AES-128 CBC and PIFF, which specified AES-128 CTR, common encryption will be a challenge. It may be possible to use a common key delivery method to decrypt content delivered in the device's native cipher mode.

PERFORMANCE CHARACTERIZATIONS

To achieve the goals of fast video acquisition time and uninterrupted playback, tradeoffs between elements related to video quality and elements related to robust delivery are required. Elements related to video quality include resolution, compression profile, frame rate, and bitrate. The elements related to robust delivery include GOP size, segment size, and the step size between variants. Ultimately, the client will determine the limits on the values for these items and will contain internal heuristics, which may dictate the most favorable combinations.

A set of experiments was conducted using a transcoder, web server and iPad to explore the impact these parameters have on an actual device. The results included in Figure 2 and 3 show a direct relationship between elements that increase the size of segments and the stream acquisition time. While results tend towards improved performance more frequent segmentation and smaller file size, other issues may result when serving a larger number of small files, given overhead requirements on a per file basis. In order to balance the requirements of video quality and scalability, it may be possible to direct the initial content acquisition to use a fast-access profile and allow it to adapt to higher quality profiles a short time after the start of streaming.

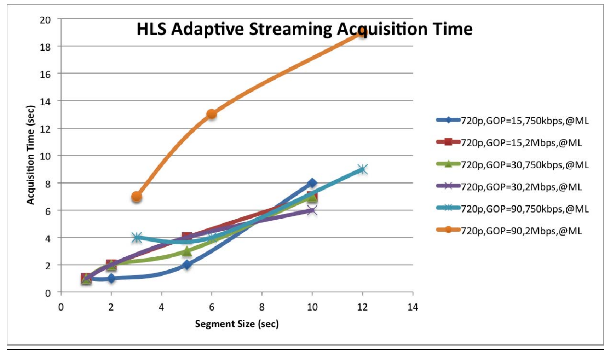

Figure 2. HTTP Streaming Acquisition Time – 720p Resolution

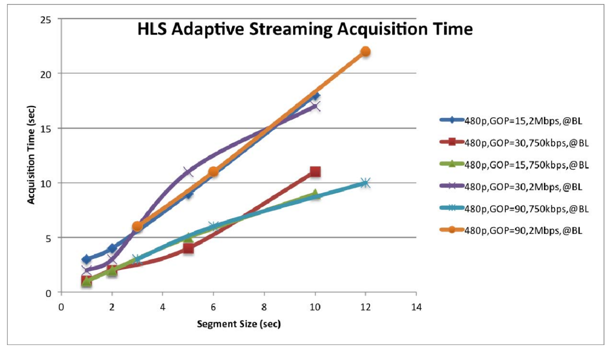

Figure 3. HTTP Streaming Acquisition Time – 480p Resolution

Other performance topics to discuss include captions and secondary audio programming. The method to convert captioning differs between transport stream-based (i.e. PIFF) and file-based (i.e. HLS) adaptive technologies. When using a transport stream-based adaptive technology, ATSC-A53 or SCTE-21 carriage of broadcast EIA-708 captions is possible. File-based adaptive technologies require conversion of captions to W3C timed-text. In either case it will be the responsibility of the client video player to display the captions. The client video player may require specific treatment in the client application to enable captioning. In a similar fashion, SCTE-18 ad splicing or SCTE-35 emergency alert triggers may be maintained in the transport stream, or mapped to a specific track in the MPEG-4 file structure. The application data delivery is not currently standardized in any file-based adaptive technology. Again, the client application is responsible to act upon the application data and take necessary actions, such as to display an ad or change to an emergency broadcast channel.

Alternate audio is standardized in transport-stream delivery although not currently handled in Apple's iOS player. Audio is limited to a single alternate in MPEG-2. PIFF offers a mechanism called late binding, where one of any number of audio alternatives may be joined to the video stream within the player environment. This feature is of interest particularly to support descriptive audio for hearing impaired viewers.

SUMMARY

A method of adapting Cable TV delivery to IP connected consumer devices is presented. Due to fragmentation in the video delivery technologies supported by popular devices, a multi-ended architecture is recommended. With the scalability and simplicity of container adaptation, a set of encoding variants can readily be packaged on a per stream basis. A replication of Cable TV on connected consumer devices can be achieved with appropriate mapping of channel associated metadata and application handling in the client.

REFERENCES

[1] HTTP Live Streaming, draft-pantos-http-live-streaming-06, March 31, 2011, http://tools.ietf.org

[2] Portable encoding of audio-video objects The Protected Interoperable File Format (PIFF) http://go.microsoft.com/?linkid=9682897

[3] HTTP Dynamic Streaming, Adobe Website, http://tinyurl.com/25mtnwj

[4] DLNA Interoperability Guidelines (member only), http://www.dlna.org

[5] DLNA Enables Premium Commercial Content Across Home Networks, http://tinyurl.com/3jownmf

[6] Webm, an open web media project, http://www.webmproject.org/

[7] Information technology — MPEG systems technologies — Part 6: Dynamic adaptive streaming over HTTP (DASH), http://www.iso.org

[8] US Connected Devices Sales (2009-2015), Parks Associates