Delinearizing Television – An Architectural Look at Bridging MSO Experiences with OTT Experiences

Bhavan Gandhi, Varma Chanderraju, & Jonathan Ruff
Motorola Mobility, Inc.

Abstract

Advent of over the top (OTT) content services by providers such as Netflix, Hulu and Vudu has dramatically altered the media consumption experience and with it the expectations of consumers. OTT services and in some cases cable provider services such as Xfinity TV supplant traditional linear and on-demand offerings. However, despite the availability of all these choices and services, the end-user's media consumption experience is disjoint and detracts from traditional lean-back TV watching. There is an opportunity to build solutions that provide a more cohesive, unified and intuitive user experience for the end-user. This paper describes architectural and system details of a system capable of delivering such an experience.

BACKGROUND

Consumer Experiences Trends

Traditional or linear TV watching, contrary to anecdotal views, is not dead. The ability to access over-the-top (OTT) content has not led to a mass exodus away from television[1]. Looking at absolute numbers, a recent Nielsen study found that people watch an average of 32 hours and 47 minutes per week of traditional TV compared to 27 minutes a week of watching video online[2]. Regardless of absolute numbers, the trend is that people are watching more video than ever before; this includes online, on portable devices, and traditional TV sets as well[3].

It is fair to say that the TV watching experience is evolving; it has been changing from a single-source, single-device experience to one of a multi-source, multi-device experience. In this new world, the incumbent service operator (i.e., MSO) is still the richest single source for linear, broadcast entertainment. Even in this ecosystem, the content delivery and experience infrastructure has been changing to accommodate and support varied devices and content formats.

The increasing number of content sources brings about the more radical experience changes. In some cases, the source is the content provider; other sources include the growing number of internet-based providers of entertainment content such as Netflix and Hulu, who we refer to as OTT operators. The addition of these content sources is fragmenting the user experience and taking it from a lean-back experience to forcing the consumer take an active role in discovering and consuming content from their various sources or subscriptions, while keeping mindful of the devices on which the content is playable. Experience fragmentation is caused by the user having to be cognizant of their subscriptions and the applications from which the content can be discovered and then played. These applications tend to be provider specific, so discovery of content within the provider's offerings requires being in their application. In the confines of the living room, this experience starts diverging from being lean-back. Even outside of the living room, there is a meaningful need to have a central hub for content discovery and an easy way of consuming the content.

Towards Convergence

The objective of making entertainment content lean back in this changing marketplace is attained primarily by unifying the linear, on-demand, and OTT content discovery process. This should be without

regard to the client device that is being used for consuming the content. And, once discovered, content playback should be easy too. Playback capabilities of the specific device should be transparent to the user.

There are a number of client-based applications on various devices that are attempting to achieve unification. Applications such as Fanhattan[4] are trying to unify the content discovery process across OTT content stores such as Netflix, Hulu, Amazon, iTunes, etc., as well as what is generally available linearly. One shortcoming is that linear TV is regional and subscriber dependent; therefore it does not reflect a subscriber's view into available content from their subscription. When it comes to fulfillment, OTT content can be consumed readily if the appropriate clients are supported on the device, however, fulfillment and consumption of broadcast content (through an MSO) is not supported even if the subscriber has a subscription.

Google TV[5] is a client device / application play in the living room to unify the content consumption experience. It is an application platform running Android OS that has the capability of interfacing to the Internet and to the set-top-box. Generally, OTT provider applications (e.g., Netflix) can run on the device to access the repository. There is also an attempt to converge the experience around television, movies, and shows through the Google TV application. This application strives to unify the content discovery from live content and the web; the application allows you to consume the content without regard to the content source (broadcast or internet). If the content is live, the Google TV box tunes the STB to the appropriate channel, and if it is available over the Internet it can be streamed to a client player that supports the appropriate security protocols. The disadvantage of this approach is that the convergence is done at the Google TV client; as such, this experience cannot be replicated across the increasing varying number of devices that are also being used to consume entertainment content.

Server-side (or cloud-based) unification of content discovery and federated content delivery and playback has the potential of bringing to bear the best of the Internet and marrying it with the best of broadcast television. Not only does it unify the content discovery, it also has the flexibility to support applications and experience across a variety of fixed and mobile devices. This is an evolution over unifying content discovery and centralizing content delivery that is espoused by Tranter[6]. We espouse centralizing discovery but federating the fulfillment.

ARCHITECTURAL CONVERGENCE

Architecturally modularizing and separating the content discovery (metadata), control, and data delivery provides flexibility in creating new services. This enables service providers to aggregate content information from multiple sources and to create tools and services that let end-users browse, search, discover, access and control content consumption across their ensemble of devices. This is key to creating a more unified user experience. The end-user is provided a unified content discovery mechanism through an application or guide interface. We have developed a cloud-hosted metadata service that aggregates, normalizes, and correlates content metadata from disparate sources and provides RESTful interfaces to applications. With the prescribed architecture, we allow for secure registration and sign-up to the user's set of subscribed services, whether they are linear broadcast or OTT, and secure access to the user's desired content. One of our primary goals is to achieve a unified discovery experience for the user while simultaneously ensuring that the content distributors or service operators can exercise and enforce their content rights over their media assets.

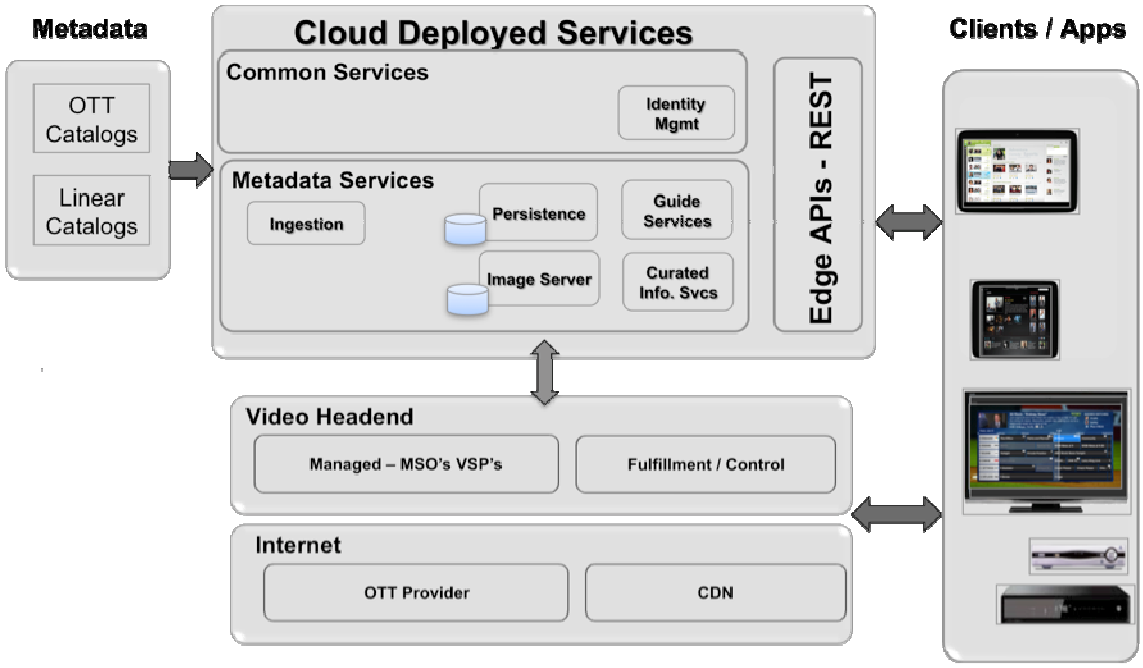

Figure 1. Modular Architecture supporting Converged Experiences

Figure 1 shows a high-level architectural system diagram that disaggregates content discovery from content fulfillment and control. The Cloud Deployed Services is responsible for hosting metadata across the varied sources and providing a unified view for discovery to client-facing applications.
For Internet based OTT content, the URL to play the content may be hosted as part of the Metadata Services or accessed using OTT provided APIs in the client application. OTT content is typically hosted and fulfilled through a content delivery network (CDN).

Taking a similar model with MSO hosted content channels and on-demand content, the client applications need a mechanism to access the content originating from the Video Headend. In our system, we use the Metadata Services as a proxy for passing minimal yet essential control information. Our approach is to publish the URL of the Fulfillment/Control services that the client application can access to discover the video channel (URL). This is modeled after the Internet OTT. By this approach, the data plane and control layers provide flexibility with how the content is delivered and what client devices are supported. The data plane layer is highly flexible and can exercise complete control over video playback on client devices. By minimizing the interdependencies between content discovery, delivery, and ultimately playback, we reduce the likelihood of linear and on-demand information sourced from a provider from getting stale.

In part, the function of an intelligent Fulfillment / Control module is to advertise its location and to enable appropriate content playback. Essentially it acts as a media broker between a user application(s) and supported clients. Once the client application / device accesses the location (URL) of this brokering service and stations that are supported, the client player can then tune to the selected content channels for the supported device type. The inherent assumption is that users are subscribers and the devices are registered with the service operator (and ultimately the Video Headend).

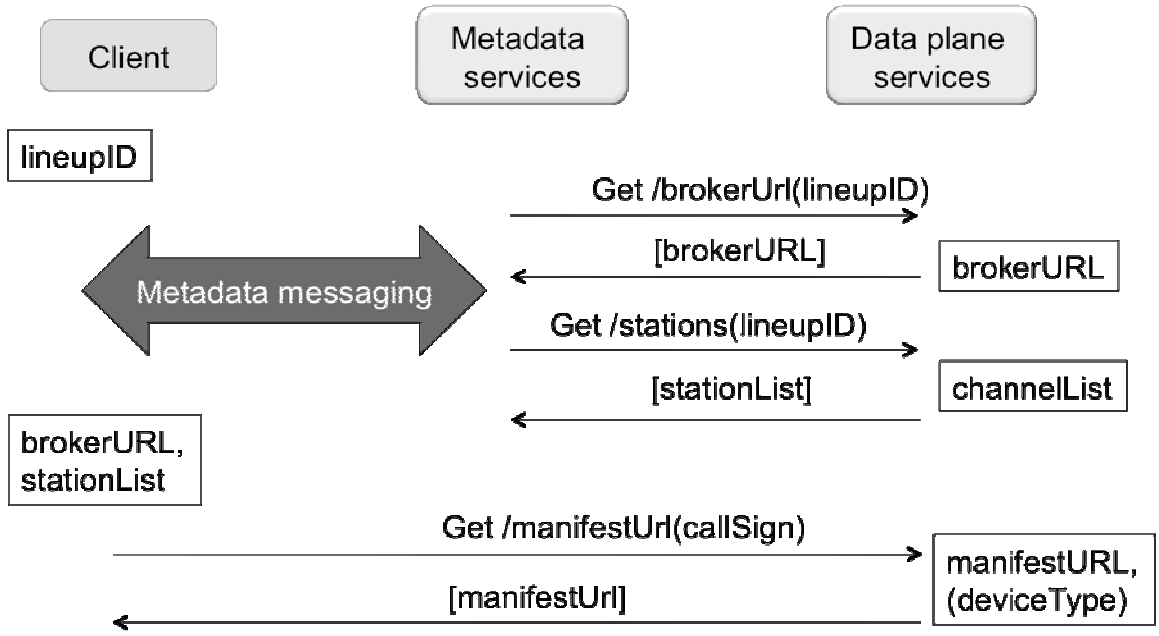

Figure 2. Interaction Diagram between Client, Metadata, and Data Services

Different models can be envisioned depending on whether all the services are hosted and operated by a single operator, or whether there are different operators for the different services.

An example set of interactions is shown in **Figure 2**. The client application first interacts with the Metadata Services to find the address of the appropriate Data plane that is servicing the delivery of the specific lineup (through lineupID). The Metadata services, through interaction with Data Plane services, can find out where the client application should point for getting access to the appropriate linear channels. Given this information (brokerURL, stationList), the application can interact directly with the data plane (Video Headend) to access the linear content (manifestURL) supported on the accessing device. This modularized system provides the advantages of unified content discovery as well as the ability to support an ensemble of devices. Also, since the service providers are ultimately responsible for fulfilling the content to the user / subscriber, they can control access. This allows flexibility in deploying and operating the system.

SOLUTION & DEPLOYMENT ECOSYSTEM

The entire system is built on a proven and robust technology stack that makes use of the latest web services technologies. The system has also been designed to support flexible deployment scenarios.

At a high level the de-linearizing (or converged) television ecosystem is comprised of several subsystems and logical modules that each encapsulate specific functionality. These include:

- Unified Metadata Services subsystem
- Portal & User Interface subsystem
- User Management subsystem
- Device Management subsystem
- Network DVR subsystem that includes scheduling, recording, and archiving
- Fulfillment subsystem
- Dataplane subsystem (Video Headend) that encapsulates transcoders, recorders & streamers

A key element of our solution is the ability to provide a set of unified metadata services to our clients. This allows the clients (internal and external) to rely on a single entity for all content metadata needs irrespective of where the source data is derived from and irrespective of the type of the source metadata (linear, non-linear, broadcast, VOD, OTT etc).

The unified metadata services subsystem is capable of assimilating highly unstructured, inconsistent and incomplete data sourced from dozens of individual metadata sources and turning the data into a consistent, complete and usable set of metadata services that can be consumed by a variety of clients. The unified metadata service exposes feeds and APIs for clients to access linear TV data including scheduling data, lineups, stations and non-linear data such as series, episodes, movies, news programs etc. The unified metadata service subsystem uses a host of techniques including customized data-collection agents (or ingestors) that are continuously tuned to be in sync with ever-changing data publishing formats used by the data-providers that we ingest data from. Examples of data sources include metadata providers such as TMS & Rovi, OTT providers such as Hulu & Netflix. Data-clustering and data-classification techniques are used to normalize and classify related data sourced from multiple metadata providers. The metadata capture and classification has evolved over time and has been enhanced using heuristic learning. In addition a powerful set of editorial tools provide the means and capability to further refine the data through manual oversight. Manual oversight is only required for a small fraction of the data. At the storage layer metadata services use a combination of SQL and NoSQL storage to ingest, classify, store and archive metadata.

Metadata services are deployed in the cloud and expose clean, simple and flexible REST APIs to consumers of the metadata services. JSON is the preferred format for the output of these REST APIs as most clients are browser-based rich-web applications.

Another key element of the solution is the usage of a powerful and flexible framework that is capable of driving the presentation elements on the client (device) side as well as providing an SDK (both on the client side and on the server side) that service operators (or 3rd parties) can use to extend, adapt and customize the look, feel and functionality of the solution. The framework helps abstract different form factors and input paradigms from user-experience developers and assists with ultimately creating uniform experiences across various device types (TVs, Tablets, Smartphones). The choice of client-side technologies (HTML5, JS, CSS) makes it possible to address devices running mobile operating systems such as Android and iOS to STBs running custom Linux images.

The server side components of the solution are deployed in Linux OS, virtualizable, developed on a Java EE platform, conform to the MVC architectural model, make heavy use of the Spring framework and use Hibernate to abstract the storage layer. The server side components are highly modular and communicate with each other mostly using REST APIs. This lends itself to flexible deployment options that can take into account different business and technical requirements that drive the deployment choices of service operators.

CONCLUSION

Being aware that video consumption is evolving and a continued user need for being easily entertained, we have architected and developed a system that allows content from multiple sources to be discovered and consumed on any device. Such a system

requires the disaggregation and modularization of the discovery and fulfillment processes. We have used the latest in web technologies to create a flexible eco-system for deploying and operating the system. This allows content to be discovered and consumed from both traditional TV and Internet OTT sources in a unified way.

ACKNOWLEDGEMENTS

Our colleagues Anthony Braskich and Stephen Emeott contributed the specification of the interaction between the Metadata and Data Plane services, including **Figure 2**.

REFERENCES

1. Robertson, A. (Feb. 9, 2012). The Verge. In "*Nielsen: most Americans still pay for traditional TV, but a growing minority go broadband-only*." Retrieved March 22, 2012, from http://www.theverge.com/2012/2/9/2787037/tv-internet-streaming-video-viewing-survey-2012-nielsen

2. Schonfeld, E. (Jan. 8, 2012). Tech Crunch. In"*How People Watch TV Online And Off."* Retrieved March 22, 2012, from http://techcrunch.com/2012/01/08/how-people-watch-tv-online/

3. Indvik, L. (Oct. 20, 2011). Mashable Entertainment. In "*Americans Are Watching More Video Online – and Everywhere Else*." Retrieved March 22, 2012, from http://mashable.com/2011/10/20/nielsen-video-tv-study/

4. (n.d.). Fanhattan. Retrieved March 19, 2012, from http://www.fanhattan.com/

5. (n.d.). Google. Retrieved March 19, 2012, from http://www.google.com/tv

6. Tranter, Steve, 2011. *"Putting the Best of the Web into the Guide,"* SCTE Cable-Tec Expo Proceedings, Atlanta, GA.