OPTIMIZATION AND EFFICIENT UTILIZATION OF ADAPTIVE BITRATE (ABR) VIDEO USING BIG DATA AND PREDICTIVE ANALYTICS

SRIDHAR KUNISETTY
JEFFREY TYRE
ROBERT MYERS

ARRIS™

TABLE OF CONTENTS

INTRODUCTION

Adaptive Bitrate (ABR) protocols have become the de-facto technology for delivering Over the Top (OTT) video to multiscreen devices such as smart phones, tablets, etc. ABR streaming is also used by operators to deliver linear content over IP to the home's second-screen devices, and it's making its way to the primary TV screen in the form of next generation IP set-top boxes (STBs) and Customer Owned and Maintained (COAM) devices. While the intent is to enable all TV subscriber devices to be ABR-based, with that comes the need for new real-time data processing and data repository platforms that can store and process network analytics and viewership data from millions of these HTTP-based devices.

Actionable knowledge derived from analyzing the vast amount of data of home devices provides valuable information that not only allows us to optimize the network by addressing the operational issues, but also enables saving valuable network bandwidth by predicting the best usage of resources. Harnessing Big Data technologies in real-time enables a feedback loop, using things such as multicast channel maps, that can be changed dynamically at a granular level – at a specific time of the day in a specific geographical area – to better utilize the network bandwidth and reduce operational complexity. In addition, historical data analysis used in conjunction in EPG metadata can be used to predict several important factors such as user surfing patterns, channel popularity, and channel change probability at granular level. This knowledge can be applied in real-time to optimize ABR delivery and directly impact operator network bandwidth savings while maintaining subscriber Quality of Experience (QoE).

This paper provides an overview into how harnessing Big Data enables insight into user behavior and corresponding video consumption, and how Predictive Analytics, among other things, enables identification of the best resources for ABR video delivery to optimize operator network bandwidth usage.

Recent research efforts into optimizing IP Video delivery using ABR led to promising solutions such as M-ABR (Multicast assisted ABR) and C-ABR(Cloud assisted ABR). The ideas discussed in this paper are directly applicable in the optimization and effective utilization of M-ABR and C-ABR.

IP VIDEO AND ABR ENABLING SOLUTIONS

ABR in IP Video Networks

Since the early deployment of IP-based video networks, various technologies have emerged to help cope with the variability associated with delivering video over non-

deterministic, best effort networking. HTTP progressive download has historically been a popular means of video delivery over the Internet. Yet in managed environments, Pay TV service operators have traditionally used MPEG-2 TS as the transport mechanism for video over IP networks.

Today, however, a newer technology called adaptive streaming has emerged. Adaptive streaming promises to enable videos to be delivered over unmanaged networks with a very high quality of experience, and is thus applicable to both Internet video environments and managed video networks that are seeking to extend the delivery of premium content to devices other than the television set. HTTP adaptive streaming has proven to be the technology of choice for many types of video delivery—both over the public Internet and what has traditionally been managed video network environments.

Unlike previous streaming technologies such as progressive download, adaptive streaming introduces the ability to dynamically react to changes in network conditions by switching to a video encoded at a different bitrate. This ability to adapt in real time more accurately reflects the dynamic conditions of today's networks, content, and devices. With users streaming more premium long form content, it is natural to expect that there will be fluctuations in the amount of bandwidth available during a two-hour movie, for example. Adaptive streaming is recognition of this fact and enables viewers to watch this premium content with a superior QoE.

Adaptive streaming works by leveraging the same content encoded in various bitrates—in a range that reflects the expected quality of the content itself, the network performance, and the screen resolution desired. For example, a video could be encoded in bitrates ranging from 300 Kbps (YouTube-quality online video) up to 6 Mbps or higher (high quality streaming content to the TV). A typical video is encoded in as many as eight different profiles, depending on the range of devices and their display sizes desired to be supported. Each of these files are then further segmented or "chunked" into short segments (typically two seconds long) that are each precisely time-stamped.

As the video is delivered, the HTTP client maintains a communication channel with the adaptive bit rate server. The client downloads these chunks as individual files, which are buffered by the client, decoded, and played out as a continuous presentation of video and audio to the viewer. During the viewing session, the client player monitors the rate at which the buffer is filling and can thereby infer the performance of the network.

If there is degradation in network performance, the client can request that chunks be delivered from one of the lower bitrate files. This is all seamless to the viewer since each source file is chunked and time-stamped in the same, very precise intervals – so there is no visible interruption or hesitation when switching to a different bit rate. Likewise, if the player detects an improvement in performance, it can request HTTP file segments from one of the higher bitrates.

Adaptive Bitrate (ABR) protocols have become the mainstay of multiscreen devices like tablets, smart phones, gaming devices, and smart TVs for accessing OTT video content. Because of their explosive popularity, it is highly desirable for an operator to provide existing services to these IP Video, multiscreen devices. The ABR protocols have been optimized to deliver a wide range of bitrates for varying screen sizes from thumbnails to 4K ultra-HD, operating over IP Network connections with fluctuating bandwidth and service levels.

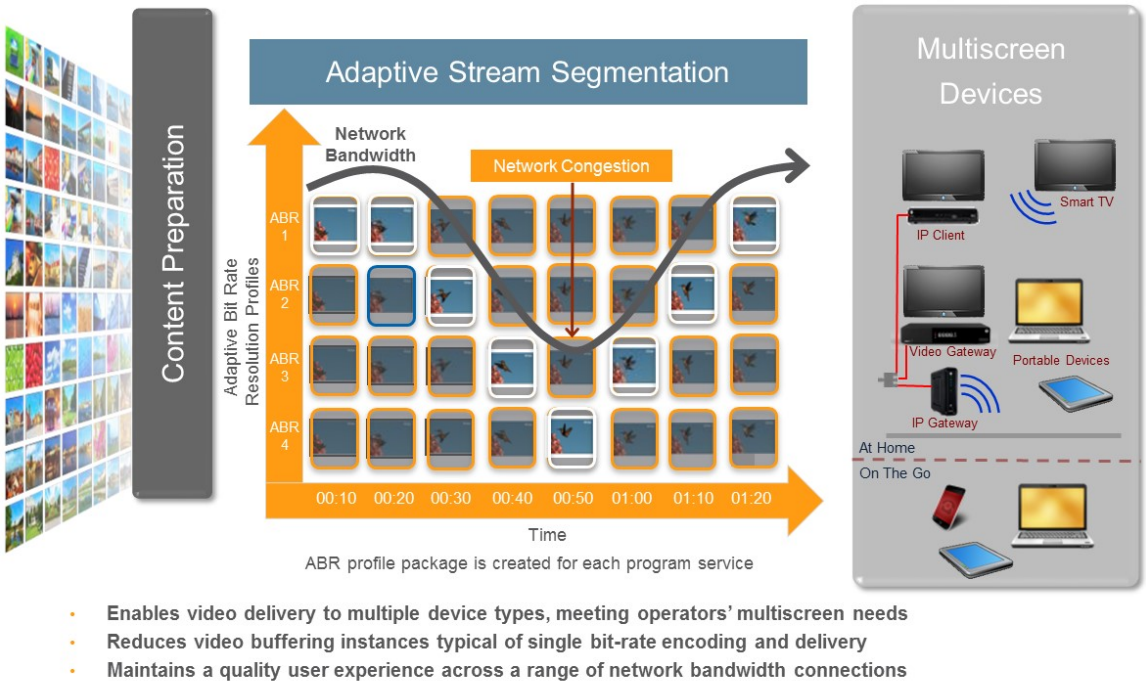

Figure 1 – ABR Ecosystem in IP Video Networks

Adaptive Bitrate streaming is already commonly used today by pay TV operators to deliver video content over IP to the home to second-screen subscriber devices such as PCs, laptops, tablets, and mobile phones. ABR streaming is now making its way to the primary TV screen in the home in the form of next generation IP STBs. We expect that pay TV operators will continue and accelerate their migration strategies to HTTP, ABR, and other IP Video technologies in terms of multiscreen strategies that can unify the full range of subscriber experience and behavior.

Unified Multiscreen Analytics and Operational Dashboards

Operational, predictive, and monetization aspects are all supported using a Big Data Analytics platform.

Understanding the operational issues becomes essential, especially for new technologies, so that we can address those issues to improve efficiency, reduce cost,

and provide better user experience. Actionable insight derived from collecting, processing, and analyzing the huge amount of data from the millions of home gateways provides several levels of valuable information that allows us to better manage the network by understanding the operational issues. The processed data is exposed as dashboards and charts to provide a visual view into the different operational and usage aspects of the system.

For example, we can capture client data statistics across all screens in the home, such as video start time, video start failure, exits before video starts, number of successful playbacks, and client buffering or stalling occurrences helps. This can be used to extract subscriber viewership patterns and tolerance for Quality of Service (QoS) levels and corrective actions can be taken by operators to improve overall subscriber QoE.

By unifying data collection and processing, instead of doing it in silos, from all components – from the headend to the end clients - we can not only provide a unified operational dashboard that allows correlations with different aspects of the system across the board.

C-ABR (Cloud Assisted ABR)

A key problem in service operator's transition to leveraging IP video networking is that ABR clients by design act independently, trying to use the maximum network bandwidth available. In conventional ABR video delivery, the ABR client determines the bitrate decisions based on its own interpretation of network conditions. Each client manages its own requirements for maintaining video / audio playout with no insight into system-wide requirements, leading to a demonstration of many inappropriate behaviors at the individual client level in terms of inconsistent content presentation to the viewer. The ABR client's "greedy" behavior leads to significant unfairness, instability, and inefficiencies relative to the larger population of ABR clients [2]. These ABR client traits are often not in line with an operator looking to offer a true "managed" ABR video service with the associated QoE.

By adding some cloud based decision making into the solution (as opposed to isolated, client-based control in the typical IP video deployment using ABR streaming), the operator can regain control to provide a first rate video service with better QoE while retaining the key underlying benefits of adaptive protocols. We refer such a solution as C-ABR (Cloud-assisted ABR). In a C-ABR solution, the key ABR decision making as to which bitrate to send to a client is controlled from the server side in the cloud. The system level intelligence in the server understands the state for every client; the available bandwidth for each client; and a "reasonable" visual quality of the video for a given size of display and attributes of video etc. Based on that intelligence the server controls what bitrate each client gets. This increases network utilization and provides significantly better fairness between clients.

For C-ABR to be effective and make an informed decision, the key is to understand the state of each client, the network, content that is being consumed, etc. This is where analytics come into picture, by helping C-ABR to make an informed decision. Data analytics can be used to optimally manage media delivery across a large and diverse population of multiscreen, such as ABR clients deployed for many different types of TV services: Linear, VOD, Network DVR, etc.

C-ABR can not only assist in providing efficient ABR video stream delivery over IP Video Networks, it can be a fundamental cornerstone in the operator's tool set in supporting a strategic migration to Web-based analytics platforms. ABR client telemetry data can provide operational monitoring and be readily integrated into operator data logging systems. Integration with customer back office systems presents opportunity for long term trending, periodic data analytics and monetization. The technologies and ideas described further below can assist in the overall planning for leveraging Cloud / NFV (Network Functions Virtualization) systems as part of a unified, IP Video service deployment strategy.

M-ABR (Multicast Assisted ABR)

M-ABR is a solution that helps in conserving the bandwidth while facilitating ABR streaming at the same time.

Figure 2 illustrates the major components of a Multicast-assisted ABR solution based on CableLabs specifications [1].

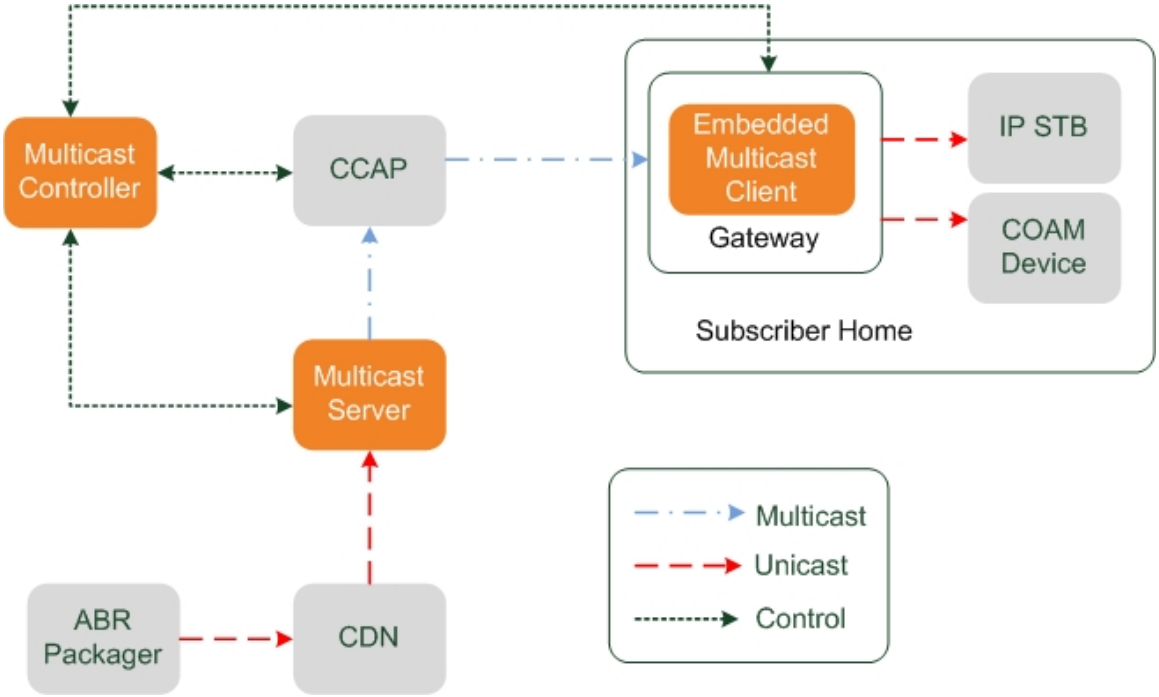

Figure 2 – Major Components in a M-ABR System

The major components in the baseline M-ABR architecture are:

1. Multicast Controller
2. Multicast Server
3. EMC (Embedded Multicast Client aka Multicast Client)

The Multicast Server pulls content from the CDN as a normal ABR client would do, and sends the content as a multicast transport over an IP routed network.

The Multicast Controller is responsible for defining the channel maps, mapping of specific ABR delivered streams, for the Multicast Servers to multicast delivery of the ABR content and provide these channel maps to edge devices so they know what services are available and how to connect to them. The Multicast Controller determines what channels are multicast by instructing M-ABR Servers to setup and subsequently tear down multicast streams as ABR stream viewership changes.

The EMC is an application running in the gateway (M-ABR Edge Device) that manages all aspects of the multicast inputs received from the multicast server and their conversion to HTTP unicast ABR streams delivered over a local network (e.g. subscriber home network) to ABR clients for playout.

The following describes the process that occurs when an ABR player client makes a request for content through the gateway. When EMC receives an ABR video segment requests from ABR player clients which have been proxied to it:

1. EMC checks if the ABR client URL request matches a stream that is currently being multicast and in the gateway's local cache. If not, EMC requests the video segment via HTTP unicast and returns it to ABR client. In this capacity, EMC acts as a traditional CDN transparent web cache.
2. If the video segment requested by the ABR client is present in the local cache, EMC returns it to the ABR client immediately. A video segment could already be present in the local cache due to a previous ABR client request or based on a scheduled operator request to pre-cache content based on projected program viewership such as for live sports events.
3. If the multicast stream has not been joined, the EMC joins the multicast stream to receive video segments via multicast, then converts these from NACK-Oriented Reliable Multicast (NORM) to an ABR format (e.g. HLS, DASH ISO), and stores these as ABR segments in the gateway's local cache.
4. All subsequent ABR segment requests are delivered from the gateway cache based on step #1 above.
5. EMC periodically checks for updates to the configuration and multicast stream channel map.

Due to the bandwidth limitations and other considerations, it is not practical to deliver multicast traffic for all the channels [3] [4]. M-ABR system achieves efficient use of network bandwidth by multicasting only the *popular channels*. This is where analytics come into picture, by identifying the popular channels and other information.

BIG DATA & PREDICTIVE ANALYTICS

Overview

The advent of the Internet and the resulting data that needs to be managed, stored, and analyzed, over cheap commodity hardware that can easily scale horizontally, led to the emergence of modern day Big Data technologies. Terms typically used to characterize Big Data are volume, velocity, and variety. Volume refers to the sheer amount of data that is being created continuously. Velocity is the speed with which this data can be processed and analyzed for a meaningful use. Variety is the different types and formats of data that is collected.

The core tenets of any Big Data system are its ability to collect a variety of large amounts of data, use parallel processing to analyze and process that information to derive meaningful insight, and to provide access to the derived information through dashboards, reports, and APIs [5]. All this can be achieved at near linear scalability - as the data size increases, one can throw more hardware at it and see processing complete in the same amount of time – and at the same time being tolerant to failures on any of the computer nodes processing that data.

Big Data related technologies have evolved over the years and most of the popular ones, such as Hadoop, MapReduce, YARN, SpringXD, and Spark are available as open source making it easier for adoption. While Hadoop is good at handling Big Data, it primarily deals with batch-oriented data processing and storage. Traditional Hadoop technologies are not well suited for real-time data processing needs. In recent years, the open source Spark framework has become the defacto choice for processing and analyzing big data in real-time. Spark also makes the life of a data scientist easier by providing APIs for statistics, machine-learning, graph processing, etc. – all in a single framework while providing linear scalability. Spark provides strong integration with Hadoop ecosystem and in many cases both of them are used together – Spark for real-time data processing & Hadoop for batch processing and to store the data. A Spark cluster, running on multiple nodes, provides the Analysis & Processing Layer the ability to horizontally scale as the data traffic increases.

A typical Analytics System is broadly classified into three different layers:

- Data Ingest Layer: This layer facilitates collection of multi-formatted information from a variety of data sources in real-time, doing parsing and pre-processing as

needed and passing it to the Analysis & Processing Layer. Optionally, a copy of the raw data is stored in Hadoop for long-term storage needs

- Analysis & Processing Layer: This is where all the real-time processing and analytics happen. It consists of different processes all running asynchronously in parallel. One set of processes operate on the raw data, processes them, and stores the results. Another set of processes perform the next level of processing by doing summarizations, correlations, deeper analysis, etc.
- Data Output Layer: The resultant data is exposed through multiple methods – dashboards and charts that provide a visual view into the different operational and usage aspects of the system and REST APIs that can be programmatically consumed by various services to act on the analyzed data

While Big Data provides an infrastructure to collect the data and provide a scalable platform to process and store the data, predictive analytics provides a way to leverage all of that information and gain tangible new insights by recognizing patterns in data to project probability. Predictive analytics encompasses a variety of techniques such as predictive modeling, machine learning, and data mining that analyze current and historical data to make predictions about future or otherwise unknown events.

In business, predictive models exploit patterns found in historical and transactional data to identify risks and opportunities. Organizations use predictive analytics in a variety of different ways by applying machine learning (ML) and artificial intelligence (AI) algorithms to optimize business processes and uncover new statistical patterns. Predictive models often perform calculations during live transactions, for example, credit card fraud detection.

Predictive models are models of the relation between the specific performance of a unit in a sample (referred to as the “training data”) and one or more attributes of the unit. The objective of the model is to assess the likelihood that a similar unit in a different sample will exhibit the specific performance. Models capture relationships among many factors which help in guiding decision making.

Big Data technologies such as Spark are making predictive analytics systems more accessible and easier to implement.

Applicability to ABR Enabling Solutions

A well-instrumented ABR ecosystem mirrors the web world in terms of volume, velocity, and variety of data that needs the processing power of a Big Data system in terms of collecting the data, processing, and storing. A typical operator has a few million subscriber homes, each of them having a home gateway, along with one or more STBs and second-screen devices connected to the home gateway. Each of these millions of home-based subscriber devices

are constantly producing events. For example, we have a system that is generating millions of data events every minute. In addition to the devices in the subscriber's home, the components in the video delivery side also produce huge amount of useful data.

Analyzing this data provides insights into the ABR ecosystem generating a feedback loop as shown in Figure 3 to help in the automation for fixing the operational issues and making key decisions in terms of network bandwidth savings, QoS, and more.

As mentioned in the C-ABR and M-ABR sections, analytics play a role in their decision making. Subscribers' home gateways are instrumented to collect user interaction and content consumption information to derive insight into user behavior with respect to video consumption. Analyzing historical data in conjunction in EPG metadata can be used to predict several important factors such as user surfing patterns, channel popularity, and channel change probability at granular level (at a specific time of the day in a specific geographical area). This knowledge can be applied to optimize ABR delivery and directly impact operator network bandwidth savings while maintaining subscriber QoE.

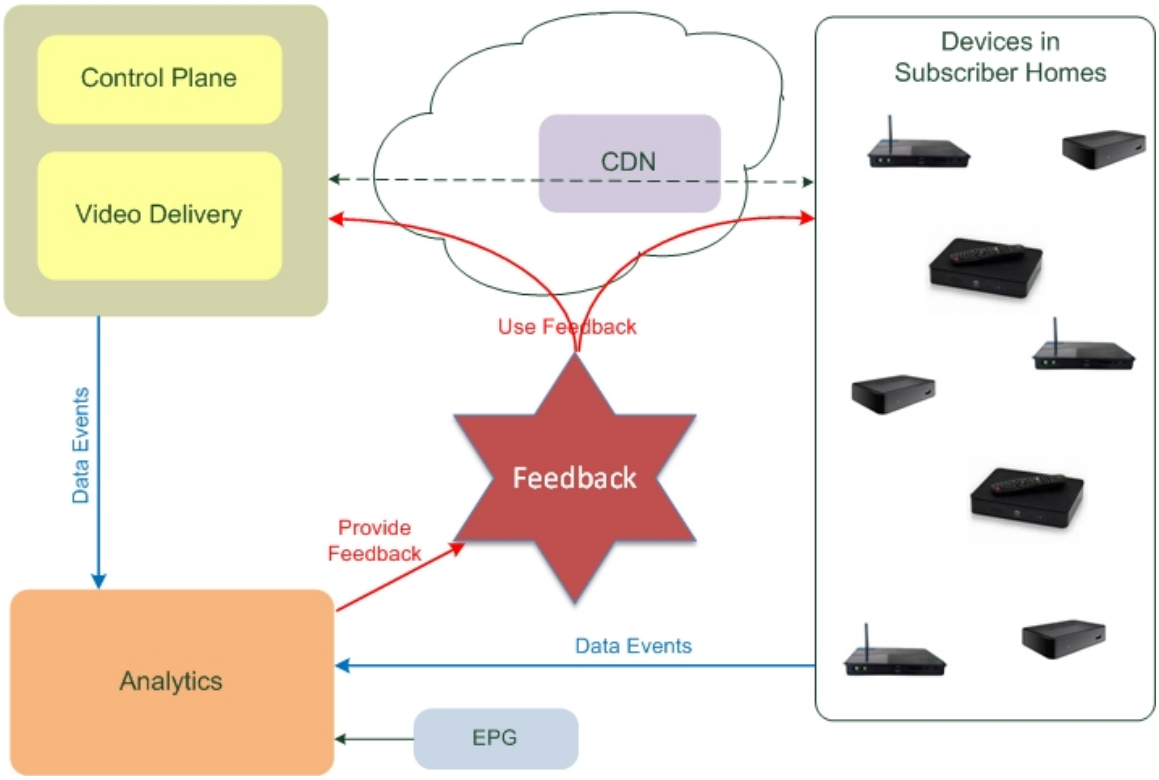

Figure 3 – Analytics Providing a Feedback Loop

Identifying and Addressing Operational Issues

Data is collected constantly from millions of devices (home gateways, COAM devices) in subscriber's home and also from the components in the content delivery side, such as Multicast Server. The data that is collected and processed provides actionable value to the operator.

Analyzing this data provides insights into the ABR solutions for understanding operational issues and generating a feedback loop to help in the automation for fixing the operational issues that were detected. In order for this to be effective, the collected data needs to be analyzed in real-time. For example, if we detect that the network latency is beyond a particular threshold, then this information can be used for corrective actions immediately.

Some of the generic high-level goals include:

- Understanding usage of different components, services, and network
- Understanding and classifying network and system limitations by analyzing trends of system usage with respect to scalability, capacity planning, and bandwidth management
- Creating dashboards to provide business intelligence, and APIs for downstream usage
- Examining the performance and workload of the CMTS/CCAP

While the generic goals cover all generic aspects of various solutions, each particular solution would have its own specific goals. For example, the goals for the operational aspects of the M-ABR [3] solution include:

- Answering key questions about the performance of multicast over DOCSIS, including the telemetry of any packet loss detected by the EMC
- Evaluating the workload within the EMC to perform the functions of the "transparent proxy cache"
- Examining the performance of the CMTS/CCAP for multicast delivery and implications related to its support for IGMP

Note that while the above mentioned goals specifically mentions cable network related technologies such as CCAP/CMTS and DOCSIS, similar goals are valid for Passive Optical Network (PON) network technologies such as OLT (Optical Line Terminal).

In order to take specific actions to optimize the network and provide a better QoS, we would need to answer specific questions. For example, in the case of M-ABR [3], the operational data would help in analyzing and answering questions such as:

- How often does a file segment experience any packet loss?
- Is the packet loss characterized by factors such as time of day, downstream capacity, CMTS type/configuration, aspects of the HFC access network, IGMP traffic management, etc.?
- How does channel surfing affect the IGMP behavior of joins and leaves and when does it become more of a burden and less of a benefit?
- How quickly can initial tuning to a new channel result in moving from unicast video segments to multicast (cached) video segments?
- How long should the EMC remain on a multicast stream before it determines there is no benefit to caching stream segments?
- What is the timing of a cache segment to client request?
- What is the behavior of client change in quality using alternate, non-multicast variant playlists?

The operational data is also exposed in a visual form through dashboards charts. This visual representation allows the operation to get a quick overview of the operations aspects of the system. For example, the charts in Figures 4 and 5 provide insights into an M-ABR system [3]. Figure 4 shows the number of client sessions that exist on a per multicast channel (stream) basis over a specified time period, termed as the Popularity of a Channel. It can help to determine which channels are more popular and when. This data can be used to optimizing the network and QoS for the popular channels and to determine what channels are more appropriate for multicast. Figure 5 shows the peak and the average time that it took for the multicast client to receive a segment from the multicast stream. This value is dependent on other factors such as duration of the segment, the available bandwidth and the multicast rate percentage. If this value becomes greater than the segment duration it could indicate an issue somewhere in the network or in the multicast server.

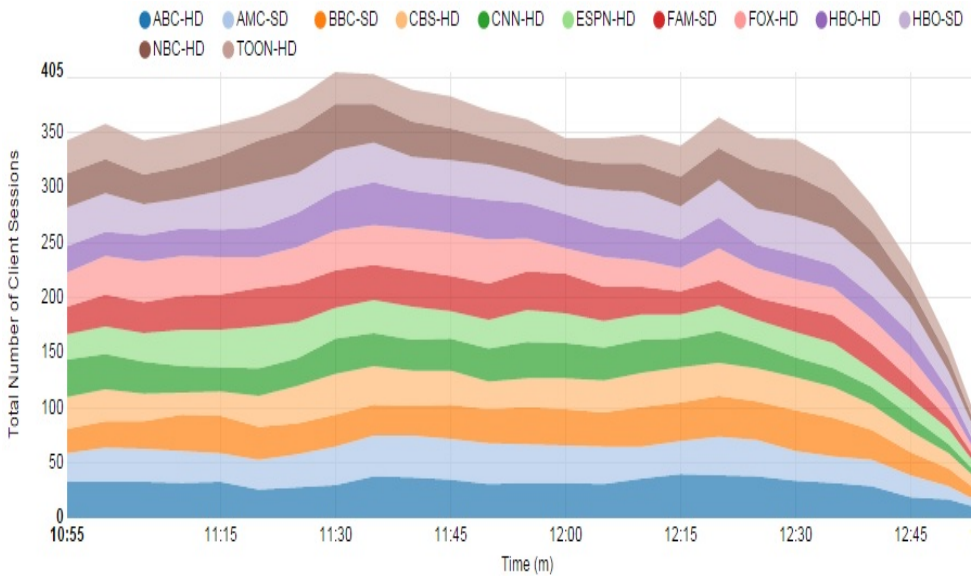

Figure 4 – Channel Utilization

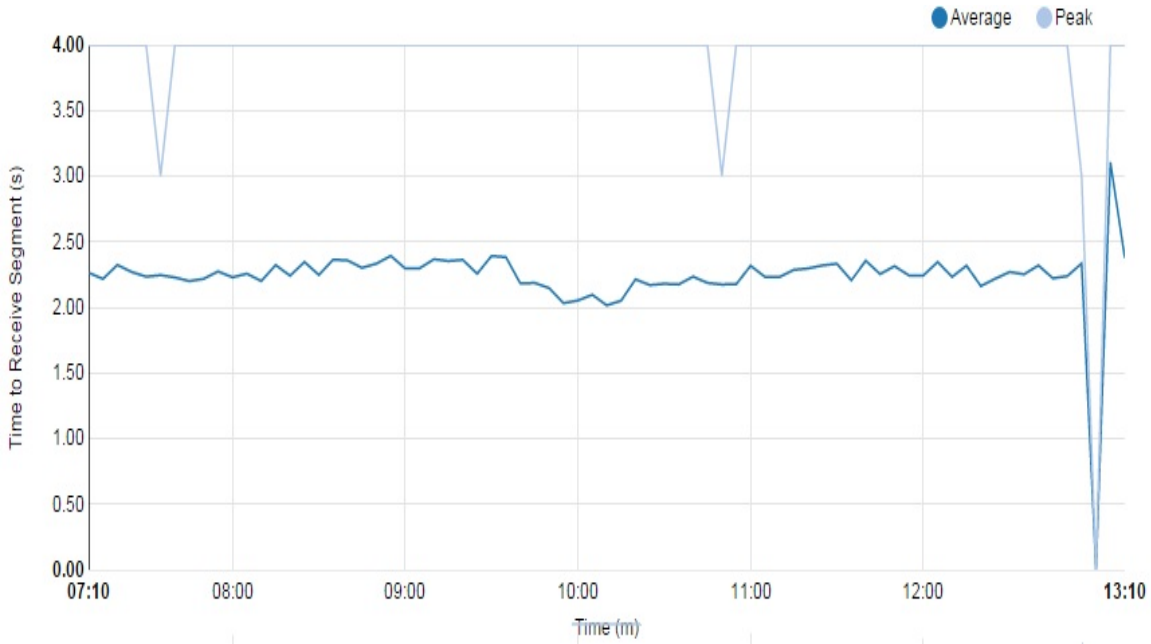

Figure 5 – Network Latency

Using Predictive Analytics in M-ABR

M-ABR system improves efficiency and saves bandwidth & network costs in IP Video delivery to multiscreen devices, by streaming some channels via multicast instead of unicast.

Due to capacity limits and constraints [4] (e.g., DOCSIS network constraints at the provider), the number of channels that can be multicast are limited (e.g. the top 20

popular channels can be multicast out of the available 500+ channels). So, there is a need to identify the channels that are the best candidates to multicast. Popular channels (i.e., more number of people watching) are the best candidates as they result in more network bandwidth saving. Channel popularity can change dynamically with time and geographic region (aka zone).

One of the straight-forward mechanisms to identify popular channels is through real-time monitoring of channel usage. This approach has drawbacks:

1. It is costly and resource intensive (requires frequent heartbeat or equivalent techniques) to monitor channel usage in real-time.
2. It becomes counterproductive in the case of flapping – channels are swapped in-and-out very frequently in short durations, thus increasing the problem rather than solving it.

Regarding the first drawback, collecting and processing heart beats (channel usage statistics) very frequently (say, every 10 seconds) from millions of subscriber gateways is costly and resource intensive. Predictive analytics can be used to predict the best channels to multicast that change in time and geography (zone) i.e., for each time period of a day, predict the multicast channel map at a neighborhood/zone level granularity. This solution can complement (or replace) the traditional real-time channel monitoring mechanism to save on costs and resources. For example, if we know that CNN is going to a popular channel at a particular time, then we can add CNN automatically to the multicast group at the appropriate time and the real-time channel monitoring mechanism can skip CNN and monitor other channels thereby reducing the number of channels it monitors

The second drawback - channel flapping – can also be addressed using predictive analytics. In the group of channels that are selected to be multicast, the channels are ordered by popularity (the first channel in the most popular and the last one the least popular in that group). Say, the group size is 10 for a top 10 list. The channels at the top portion (say top 80%) are relatively stable. As time progresses, the position of a channel in the top may change, but it will still remain in the top 10 list. However, the channels at the bottom portion (say bottom 20%) will vary frequently due to viewers changing/flipping channels. This changing/flipping channels can result in a channel (in the bottom portion of popular channels list) to drop out of the top 10 list and be back into the top 10 list within a few seconds. Typically, when a channel is flapped-out, the M-ABR system will stop multicasting that channel and it will then be available only via unicast. When a channel is flapped-in, the M-ABR system will start multicasting it. If this in-and-out happens on a channel within a few seconds, then it creates a problem where multicast is not efficiently used and switching between the protocols creates additional overhead and potential unicast bursts.

Figure 6 provides an overview on analytics collecting data, generating channel maps and feeding this information to the multicast controller, which uses this information to decide on the list of channels that would be delivered using multicast.

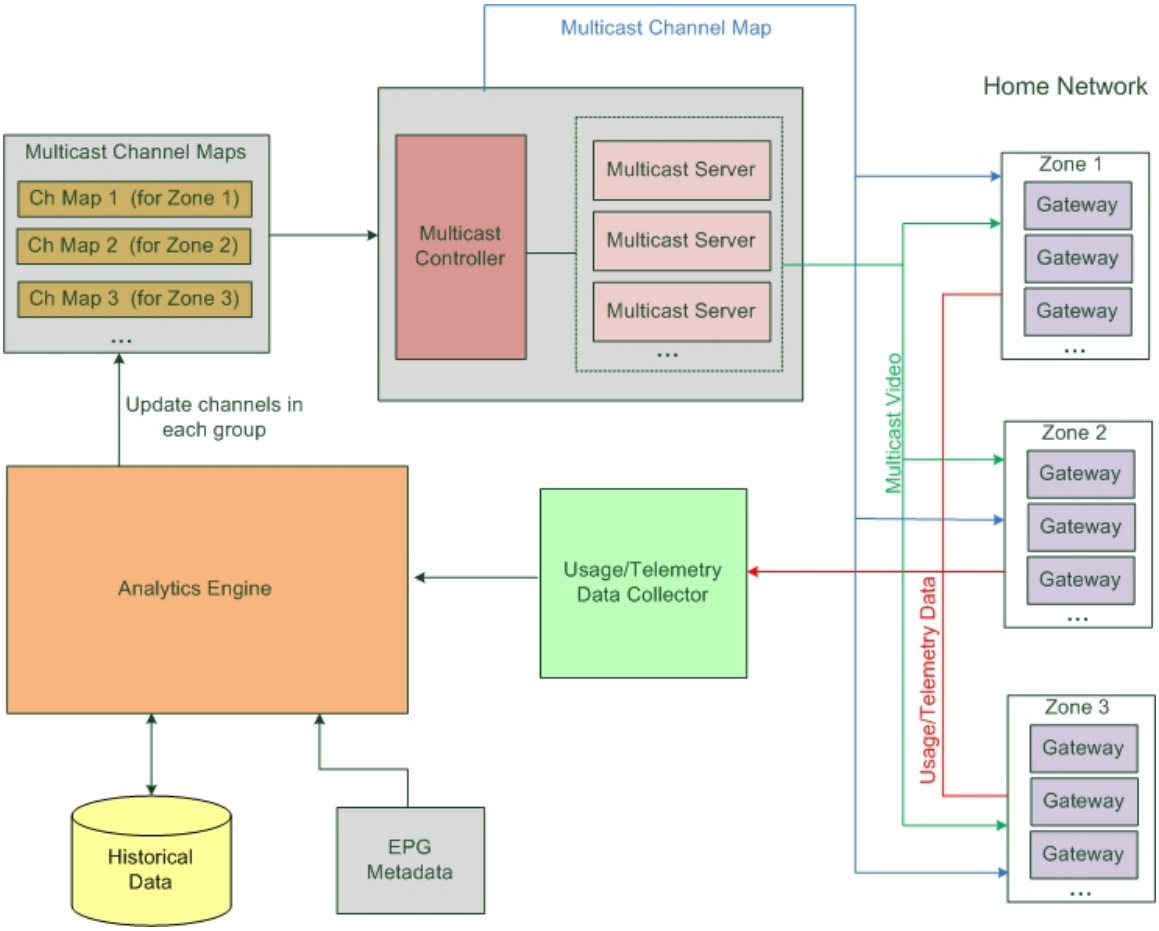

Figure 6 – Predictive Analytics in M-ABR

The analytics system keeps collecting usage and telemetry data from all the subscriber's gateways. This data is stored and over time we get a considerable amount of historical data which is analyzed to find user behavior, usage patterns and correlations. Note that for privacy reasons, Personally Identifiable Information (PII) are anonymized.

A variety of techniques are used to make different types of predictions taking different. Each of those techniques takes different data points into consideration to make predictions.

The use of external data, such as EPG metadata augments the prediction capabilities by allowing us to peek into more granular level of usage patterns. The EPG metadata provides details such as genre, rating, cast, title, channel, timestamp, etc. This data helps to build a program popularity model and allows us to derive insights into popularity of a program (and the corresponding TV channel) and the subscriber's

likes/dislikes. For example, if we see that a subscriber (end user) regularly is tuning into CBS channel at 7 PM on Mondays, from the EPG metadata we can find that the program during that time is Big Bang Theory and derive that the viewer likes that program and can predict that he/she is likely to watch that program when it telecasts at a future date/time and predict a given program's usage in future.

Understanding usage patterns provides information to predict channel usage. The historical data is analyzed to find patterns such as:

- Binge watching vs. random vs. following a schedule (time of a day)
- Favorite programs & channels
- Frequent channel surfing vs. steady watching vs. surfing during commercials/ads
- Favorite types - sports vs. news vs. reality shows
- Automated/scheduled recordings.

Channel change prediction finds the probability when a channel will be changed (users leaving & users joining) using historical data with input from program popularity and usage patterns. It could predict events such as:

- Channel join (the probability that a viewer tunes to a channel at a given time of day)
- Channel leave (the probability that a viewer leaves to a channel at a given time of day)
- Channel flipping during an ad (SCTE-35 markers is one way to know when an ad occurs)
- Whether channel change coincides with the change in the program (usually around the half hour or hour time boundary)
- Viewership of a channel in the next few minutes
- Program change based on EPG metadata

The viewership and channel surfing patterns of individual users are aggregated to predict an informed decision about a channel change at any geographical zone and time frame. This allows us to address flapping i.e., we can predict if a channel needs to be kept in the popular list (e.g., top 10 list) even though the viewership drops temporarily or vice-versa.

Based on all the above mentioned prediction models, without using any heart beat kind of mechanisms the analytics system can build channel maps – for a given zone and for a given time of the day – as shown in Figure 7 and provide that information to the multicast controller which uses this data to decide on the channels to be multicast.

Figure 7 – Example Showing Multicast Channel Maps: Time & Zone View

Augmenting Decision Making in C-ABR

In a C-ABR solution, the key ABR decision of deciding which bitrate to send to a client is controlled from the network side in the cloud, as opposed to a typical ABR solution where each client decides how much bandwidth is available independent of all other clients. In parallel, subscriber QoE can be maintained by maximizing network delivery capacity based on the dynamic measurement of video quality on a per video stream basis. In order to do make a network-wide optimized bandwidth capacity and subscriber QoE decisions, collection of data in real-time from all the end-clients (e.g., subscriber's COAMdevices) is required.

The Big Data infrastructure is used to capture client data statistics across all screens in the home, such as video start time, video start failure, exits before video starts, number of successful playbacks, and client buffering or stalling occurrences. This data is used to extract subscriber viewership patterns and tolerance for QoS levels. Bandwidth measurements can provide analytics such as the Wi-Fi bandwidth achieved in a home by client type (Mobile, STB, Gateway, Smart TV, etc.), which can be used for tailoring media delivery to such client as well as being leveraged for longer term service assessment.

Using this data, the analytics system would augment the decision making for the C-ABR, by understanding the state for every client; the available bandwidth for each client; and a "reasonable" visual quality of the video for a given size of display and attributes of video etc. Based on that intelligence, C-ABR can enforce a network-wide bandwidth capacity policy that allocates bandwidth by ABR client or class, with precision potentially down to the specific ABR stream bitrate that each client gets over a given duration of time.

The net effect increases overall network bandwidth and resource utilization while providing significantly better fairness between clients. As a cloud-based solution, C-ABR

can dynamically scale processing requirements according to network capacity being consumed by ABR clients yet within the guidelines of network capacity provisioned by the service operator.

Unmanaged ABR client deployments have been extensively studied and reported in earlier research papers[2] .

CONCLUSION

As operators migrate to IP Video delivery, they should look into adopting promising solutions such as M-ABR (Multicast Assisted ABR) and C-ABR(Cloud Assisted ABR) in order to optimize the network and improve QoS and QoE, while still enabling ABR streaming to TV subscriber home devices.

This paper provides an overview into how harnessing Big Data enables insight into user behavior and corresponding video consumption and how Predictive Analytics can analyze historical data to predict several important factors such as user surfing patterns, channel popularity, and channel change probability at granular level (at a specific time of the day in a specific geographical area). This knowledge can be applied to optimize ABR delivery and directly impact operator network bandwidth savings while maintaining subscriber QoE.

Also, understanding the operational issues becomes essential for these new technologies so that we can address those issues to improve efficiency and reduce cost. Big Data provides tools and technologies to collect, process and store data that are constantly being produced from the millions of home-based subscriber devices. Analyzing this data and providing a feedback loop into the ecosystem helps in the automation of addressing the operational issues.

While this paper has shown a few possibilities in terms of helping in decision making and addressing operations issues by using Big Data and Analytics, operators can think of many other possibilities that are specific to their solutions. For example, Analytics can help predict traffic volume which allows us to plan and manage the capacity of resources in the cloud (CDN capacity, Multicast Servers, VOD Servers, and Network Resources / CMTS Service Groups).

ABBREVIATIONS

ABR	Adaptive Bitrate
C-ABR	Cloud assisted ABR
CCAP	Converged Cable Access Platform
CMTS	Cable Modem Termination Systems
COAM	Customer Owned and Maintained
DOCSIS	Data Over Cable Service Interface Specification
EPG	Electronic Programming Guide
IP	Internet Protocol
M-ABR	Multicast assisted ABR
OTT	Over the top
SABR	Smart ABR
STB	Set Top Box
QoE	Quality of Experience
QoS	Quality of Service

RELATED READINGS

- **ABR Delivery Architectures and Virtualization** – this paper discusses two emerging trends in video processing delivery, namely, migration of various video processing functions to the network cloud to leverage advances in virtualization and dynamic packaging techniques for adaptive bitrate (ABR) delivery of video.
- **Effective Utilization of Multicast ABR Using Big Data and Real-time Analytics** – This paper provides an overview into how Big Data and real-time analytics can enable insight into video consumption and network operational aspects to make M-ABR more effective. By leveraging actionable insight, service providers can better manage their networks, while increasing bandwidth utilization and reducing operational complexity.
- **Smart ABR: The Future of Managed IP Video Services** – This paper compares in detail the SABR system to traditional unmanaged ABR delivery as well as a system with enhanced CMTS QoS. With SABR, operators can significantly increase their IP Video capacity while gracefully handling congestion and providing an improved user experience.

MEET ONE OF OUR EXPERTS: Sridhar Kunisetty

Sridhar Kunisetty, Distinguished Engineer, has 20 years of experience in various technical and management roles at ARRIS, Motorola, Oracle, Commerce One, and iPass & Connectbeam. His expertise is in Analytics, Web Services, Cloud, and Databases technologies. Sridhar is a lead inventor on several patents. When he gets a break, Sridhar enjoys biking & running and has participated in several marathons. Sridhar holds an MS in Computer Science & Engineering from University of Florida, Gainesville and a Bachelor's Degree in Computer Science & Engineering from National Institute of Technology (NIT), Warangal, India.

REFERENCES

(1) IP Multicast Adaptive Bitrate Architecture Technical Report; CableLabs document OC-TR-IP-MULTI-ARCH-V01-141112, Nov 2014

(2) SMART ABR: The future of managed IP Video services; John Ulm, Ajay Luthra, Praveen Moorthy, Mark Schmidt, and Haifeng Xu; 2013 Cable Show Technical Session

(3) Effective utilization of M-ABR (Multicast assisted ABR) using Big Data and Real-time Analytics; Sridhar Kunisetty, Jeffrey Tyre, and Robert Myers; INTX 2016 Spring Technical Forum

(4) Challenges delivering Multiscreen Linear TV Services: Multicast assisted ABR to the Rescue; John Ulm, SCTE Cable-Tec Expo, Fall 2014

(5) TV INSIGHTS - Application of Big Data to Television"; Bhavan Gandhi, Alfonso Martinez-Smith, & Doug Kuhlman, IBC 2015

