Adding the community to channel surfing: A new Approach to IPTV channel change

The MIT Faculty has made this article openly available. Please share how this access benefits you. Your story matters.

<table>
<thead>
<tr>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>As Published</td>
<td><a href="http://dx.doi.org/10.1109/CCNC.2009.4784953">http://dx.doi.org/10.1109/CCNC.2009.4784953</a></td>
</tr>
<tr>
<td>Publisher</td>
<td>Institute of Electrical and Electronics Engineers</td>
</tr>
<tr>
<td>Version</td>
<td>Final published version</td>
</tr>
<tr>
<td>Accessed</td>
<td>Sun Apr 07 10:19:31 EDT 2019</td>
</tr>
<tr>
<td>Citable Link</td>
<td><a href="http://hdl.handle.net/1721.1/60008">http://hdl.handle.net/1721.1/60008</a></td>
</tr>
<tr>
<td>Terms of Use</td>
<td>Article is made available in accordance with the publisher's policy and may be subject to US copyright law. Please refer to the publisher's site for terms of use.</td>
</tr>
<tr>
<td>Detailed Terms</td>
<td></td>
</tr>
</tbody>
</table>
Adding the Community to Channel Surfing
A new Approach to IPTV Channel Change

Marie-José Montpetit
Motorola Media Lab Resident
Cambridge, MA USA
mmontpetit@motorola.com
Henry Holtzman,
MIT Media Lab
Cambridge, MA USA
Holtzman@media.mit.edu
Herb Calhoun
Motorola H&NM TO
Dallas, TX USA
Herb.Calhoun@motorola.com
Dan Grossman
Motorola ARTC
Malboro, MA USA
Dan.Grossman@motorola.com

Abstract—While IPTV architecture and user interfaces have received a lot of interest in recent years much less focus has been given to the issues of channel surfing in the IPTV era. IPTV is rapidly transitioning from a Telco solution into a generic service over any broadband access network. Traffic issues that relate to channel change are moving to the mainstream and are the talk of a number of standards organizations. This paper proposes a new approach to channel change. It uses peering and viewing statistics to reduce channel change delays and improve the user quality of experience.

I. INTRODUCTION

Channel change delays are well known to impair the viewing experience of users. The response time given by analog tuners is hard to replicate in the digital domain. Especially for IPTV, where the channel change is an Internet Protocol (IP) message, network delays can impair channel change especially at high network loads. When a keypress is detected the signaling has to travel all the way back to the headend video server to be processed. The server then sends the right channel downstream. Network delays caused by Transport Control Protocol windowing and slow start issues may be exacerbated by heavy traffic resulting in poor performance for the channel change.

Many solutions to Fast Channel Change exist and have been deployed and submitted for standardization [1, 2]. They rely on upper layer traffic and content management like modifying rates and compression ratios [1] or marking packets as high priority [2]. But the messaging still moves across most of the network. Could there be another way?

II. IPTV BASICS AND VIEWING STATISTICS

IPTV is TV delivered over an IP network but the IPTV experience differs in many ways from the web video experience. On the web, user are used to search for content and click on a URL. People watching TV expect to be able to channel surf even with the best of on screen user guides. They expect this surfing to be fast. They are used to the RF delays of analog boxes and similar delays from some digital box.

In IPTV, live (linear) TV uses multicasting. One does not in fact “change a channel” a set top “joins” a stream from a content source in the video headend. In principle, a channel change message has to travel all the way across the network to a remote IGMP server (router) and back for the change to be “executed” and the user be joined to the requested stream.

The message is essentially another IP message and can and will experience congestion in the network especially in "mass channel change events" like the start or end of a popular show or sporting event.
However, it is well known that there are only a small number of channels that are actually been watched at any time (see Figure 1) inside a local community. Hence there is a high probability that someone else is watching the channel one wants to switch to. This is the basis for multicasting for large audience but it remains true at smaller granularity (100s of users). Figure 1 was created with 350 users and the correlation is high: only a small subset of channels has all the viewers and most channels are not watched [3]. This phenomenon of highly correlated viewership will be used to improve channel change as will be seen in the next sections.

Figure 1. Viewership distribution

III. CHANNEL ACQUISITION DELAYS

Channel change delays are created by how fast a STB can acquire the TV stream. The ability of a decoder to acquire and render video stream encoded using H.264 (MPEG-4) is dependent upon 3 main contributors. These are the session signaling, network signaling (may include tuning to a physical channel in some networks) and basic video transport protocol. The network signaling is the topic of this paper and includes IGMP delays. The session signaling includes all of the information needed to access the video content including any subscriber authorization, Conditional Access (CA) and Digital Rights Management information as well as the device processing (server loading etc). The transport includes the basic H.264 protocols which include picture, frame and slice structures (types and ordering). This will set the minimum delay that the channel change can achieve.

The transport for H.264 encoded video defines an Instantaneous Decoder Refresh (IDR) where all of the information for a picture is refreshed as intra-coded. Reference [4] defines this as:

"An encoder sends an IDR (Instantaneous Decoder Refresh) coded picture (made up of I- or SI- slices) to clear the contents of the reference picture buffer. On receiving an IDR coded picture, the decoder marks all pictures in the reference buffer as unused for reference. All subsequent transmitted slices can be decoded without reference to any frame decoded prior to the IDR picture. The first picture in a coded video sequence is always an IDR picture."

Intra coded slices (or frames if the whole frame is a slice) do not rely on any previous frames information.

In digital transmission, I-frames provide a “baseline” for a slice which is then modified by subsequent P-frames until the next IDR picture event.

Additionally at an IDR any P slices do not rely on any previous slices information. So as an example using NTSC frame rates, if the I picture is sent once per GOP, where the GOP repetition rate is 30 frames (1 sec), and an IDR repetition rate is 5 seconds, the worst case time for new channel acquisition would be 60 frames or 2 seconds. The best case, if the change happened exactly at the IDR would be 3 frames.

The standards organizations have provided further guidance. Once a transport is acquired, the acquisition time (also termed time at which clean displayable pictures start) depends on PAT (Program Association Table) and PMT (Program Map Table) frequency followed by the location of H.264 video acquisition points. Standards require (Program Association Table (PAT) to be repeated 100 ms and Program Map Table (PMT) every 400 ms. For H.264 based networks, SCTE and DVB standards specify what are called Random Access Points (RAP) which are the H.264 acquisition points. SCTE requires a RAP every 1 second and this is an I or a IDR picture. In addition, 500 ms from the acquisition of RAP all presented pictures are expected to be fully displayable.

In addition H.264 adds a layer of complexity of 3 pictures after the acquisition. SCTE and DVB standards added this conditioning using the RAP for H.264 while this was not needed for MPEG-2. So the worst case acquisition time using this standard would be 100 ms for PAT plus 400 ms for PMT plus 1 second for RAP repetition plus 500 ms or 2 seconds total. Well
designed encoders can reduce the 500 ms to 99 ms (3 frames) for a worst case acquisition time of approximately 1.6 seconds. This created the lower boundary for the channel change delay. The goal of the channel change is to get as close as possible to this boundary by reducing the other elements of the delay namely the network and authorization delays. As mentioned earlier, RF transmitted H.264 has the same channel “capture” delay, the difference between IPTV and RF is the network signaling added to the tuning time.

IV. USING PEER GROUPS FOR CHANNEL CHANGING

The proposed solution wants to replicate the RF tuner performance by moving the channel change from edge of the core network router to the edge of the access network.

This approach takes advantage of the formation of neighborhood peer to peer groups based on shared common physical transmission to facilitate fast channel changes in Internet based television.

When the neighborhood group is formed the peering client collect information on the users via, for example, expanded Dynamic Hash Tables (DHT) [5] well known in the software and peer to peer communities. The tables contain information to associate the group to a common group identity and encryption key to secure content and prevent snooping by non authorized users. This table is also updated when new channels are joined by members of the community and maintained/managed as close to the community as possible (super peer, gateway, access point etc.) and retransmitted to the group.

With this setting, there is no need to modify channel change messaging but to use standard IGMP protocols [6] and to leverage the statistical information about TV viewship in a neighborhood. This approach keeps the channel switching close to where the message originated and executes the change in a “super peer”: community server, GPON, base station, access point etc. in order to minimize channel change traffic and meet network delays of the order of 50ms. This is achieved by using capabilities already offered by these devices (proxying or snooping for example and eventually routing) to enhance how the IGMP signaling is processed.

With this solution, any channel attachment message for a channel already in the community is not relayed upstream instead is used at the edge to connect the requestor to an already existing transmission. This can be achieved in a network access device (NAD), namely a gateway, access point or even eventually community server. So, those channels already watched locally are processed locally, new channels requests are sent to the servers but with a large reduction of cross network traffic.

The group management also ensures that NADs with full access to the feed from the video server minimizes the loading on lower capacity peers. The system can also be extended to support more advanced television application including user to user communications and near-VOD applications.

Hence the Monitor, STB and NAD will be the only participants in a channel change and the channel information table information can be sent asynchronously to keep all peers updated. Hence the budget for channel change can be summarized as:

1. STB: request processing < .1 seconds
2. NAD: request processing, response < .1 seconds
3. NAD: addition of channel not in downstream < .3 seconds (3 layers of processors)
4. STB – acquisition time (from the last section) < 2 seconds
5. Monitor – video scaling (16:9, 4:3) < .5 seconds

A total budget for IPTV channel change of less than 3 seconds in the worst case.

V. GPON IMPLEMENTATION

In order to prove the concept, a first implementation of the P2P FCC intends to leverage advances in optical networking technology to create a combination group access point/local switch with the following elements:

- GPON/OLT technology for fast transmission
- Fast state of the art optical switching methods
- Peer to peer group formation and management
- IGMP snooping and spoofing at the GPON
- Group conditional access and private/public identities embedded in the DHT

A passive optical network (PON) is a point to multi-point network architecture that delivers a common signal set by means of optical splitters at each drop off of a single fiber. Encryption is used to preclude eavesdropping except for shared signals (such as Linear TV). Upstream signals time share a separate common fiber, equal capacity being allocated to upstream signals by the Optical Line Termination (OLT) “ranging function”.

A Gigabit PON (GPON) provides a downstream capacity of 2.2 Gbps which is sufficient for 300 8 Mbps High Definition (HD) channels, approximately 2 unique HD channels to each premises. Within the proposed system the downstream signals can be shared between premises to conserve downstream bandwidth. Above the OLT all channels are delivered to the OLT with capacity left for VOD, voice and data services.

For channel switching it is proposed that device uses current capabilities for IGMP proxying and switches directing one user on a frequency directly onto a time and frequency port already watched in the neighborhood and allows P2P transmission of the new stream. In order to do this the GPON needs to keep a table of ports, multicast addresses and authentication information as well as other ancillary information as needed. This is just a slight extension of existing capabilities.

When the neighborhood group is formed the user management server at the GPON creates the tables from the DHT information. This information is also associated with a group identity and encryption key to secure content and prevent snooping by non authorized users. Essentially the GPON becomes a “super peer”.

The GPON monitors channel attachment messages in the community and it appends new channel information to the table with associated metadata and GPON port number for the time/frequency allocation. Any attachment message for a channel already in the community is not relayed upstream instead the optical switching allows the watched channel to be shared. The table is updated and the information shared in the P2P community asynchronously when new group members join.

The operations can be summarized as:
1. A viewer is watching Linear TV channel
2. The viewer chooses to move to a different Linear TV channel
3. The viewer selects the destination channel, the STB receives selection
4. The STB sends the multicast IP “join” request upstream (to GPON/OLT)
5. The GPON OLT receives the “Join”, replies with the multi-cast address of the request channel if it is in the table of “watched channel”; if not it sends the message to the IGMP server in the cloud.
6. For the case that the channel was not in downstream, the OLT inserts it into the stream and removes un-used channel. The OLT knows which channels are un-used because its received changes to new channels (or shutdowns) from all viewers. It is believed there will always be enough room due to the fiber capacity and limited residences per fiber.

VI. CONCLUSION

This paper presented a new approach to channel change that leverages viewing statistics and local information. While the solution was developed with optical networking as the first implementation, the table based solution proposed can be used over a large number of access networks where the community of viewers is significant and when the number of channels offered is not exceeding the network instantaneous capacity. This approach also exemplifies a trend in the design of the next generation Internet: use community based and social networking to improve performance and security while keeping standard and simple procedures and protocols.

ACKNOWLEDGEMENTS

The authors would like to thank Mandayam Narasimhan, Erik Elstermann, John Shumate and Steve Anderson of Motorola for their valuable inputs on decoding delays and Bill Weeks for his continuing support to this work. Thanks to the MIT CIPS and Fred Leonberger for the
opportunity offered to share this work with the optical networking research community.

REFERENCES