Abstract—This article describes a novel architecture for delivering personalized IPTV experiences to the end users. The framework leverages the IP Multimedia Subsystem (IMS) to quickly enable new innovative, multi-media services. We describe the session flows for a single IPTV session switch between multiple set top boxes (STBs) and between a STB and a PC/mobile device SIP user agent. We also demonstrate how supplementary services like caller ID and advanced telephony features can also be migrated across multiple devices.

Index Terms—IPTV, IMS, handoff

I. INTRODUCTION

IP Multimedia Subsystem (IMS) is a service subsystem within the next generation network (NGN) architecture currently planned for mobile and fixed multimedia services, standardized by the 3rd Generation Partnership Project (3GPP) [1]. IMS promises a scalable integrated platform that enables new services and provides for the combination of telecommunications and Internet services.

This paper addresses the problem of transferring IPTV sessions from one device to another. Session mobility is the capability that allows a user to transfer an ongoing communication session from one device to another. There are several use cases for session mobility namely a) maintaining continuity for mobile users (e.g. a user leaving work has the need to switch his IPTV desktop session to the mobile terminal), b) achieving better quality of service (e.g. improve voice quality by switching from mobile terminal to fixed terminal or improving video screen resolution), c) avoiding a loss of session (e.g. avoiding battery depletion of a mobile by switching to a fixed terminal) and d) lowering the communication cost (e.g. by switching access from cellular to WiFi or fixed line).

Included in the scope of this paper are:
1. The ability for a participant to transfer and retrieve an active media session to one device.
2. The ability for a session to be split across multiple devices (i.e. Complete transfer or partial media specific transfer)

3. The ability to also transfer supplementary services (e.g. CallerID) to the destination device along with the IPTV media.

In order for session mobility to occur, destination devices in the vicinity of the current device need to be located. The techniques for device discovery are handled by mechanisms not detailed in this paper. Existing protocols like Bluetooth or SIP service location protocol (SLP) perform this function adequately.

There are several performance metrics that need to be assessed.

1. Transfer delay – To ensure a smooth transition the transfer delay should be kept at a minimum. The delay is due to the time for media buffering or application setup.

2. Media disruption – Frame losses and packet drops during switching can have an undesired impact on the quality of the multimedia session. In this paper we do not quantify the video/audio quality as a result of these losses.

3. Destination device capabilities – Display size, video resolution and codec mismatch should be reconciled before transfer. Media adaptation or transcoding is a key issue while handing off.

The attributes of a NGN IMS subsystem include:

1. Access Agnostic Infrastructure – services are independent of underlying access network.

2. Full mobility - transparent connectivity across heterogeneous networks, protocols and access mechanisms.

3. Always on, always available capabilities via sessions that cross networks and devices, automatically and transparently.

4. User-centric context, both device and context-sensitive.

5. Personalized context-aware applications catered to the needs of an individual or a group of individuals.

6. Flexible user interface enabling users to achieve their goals efficiently.

7. Privacy, safety, and security of information to safeguard business and consumer integrity and protect the digital rights of content creators.

The opportunity is to provide a highly integrated solution for seamless, networked-based media over three screens (TV, mobile devices, and personal computers). The objective of this
paper is to enable portability of video services through IMS. In particular, several handoff scenarios between wireline set top boxes and wireless handsets will be shown. There are several reasons for using an IMS core. Some of the reasons are as follows:

1. Core service network which is independent of access technology
2. Same application and service is available from any access method or device.
3. Ability to migrate and deploy across fixed and mobile users
4. Standards allow scalable deployment of new services
5. Evolution to combined services for enhanced user experience (presence, messaging, address book)
6. Security in IMS is built-in - identity management, authentication, authorization and service access
7. Centralized user profiles shared between applications
8. Architecture designed for scalability and redundancy
9. Common solution to achieve Quality of Service
10. Flexible Charging for multimedia and combined services
11. Common Provisioning

Section II provides a quick overview of the various types of session mobility that are used in the context of media switching. Section III discusses the specifics of the IPTV infrastructure using an IMS core. Section IV investigates service details, including typical session flow scenarios that we believe are novel in the context of IPTV. Section V describes a prototypical implementation of the above mentioned scenarios in a realistic environment.

II. RELATED WORK

Bodzinga et al. discuss how IMS and IPTV service platforms can be interworked and integrated to reduce network complexity and provide a flexible network for novel differentiated services [1][2]. An important aspect that IMS addresses is redundancy and scalability; these issues within IMS are addressed in [3]. This is particularly attractive for operators that deploy services which scale to a large subscriber base. The basis for SIP session mobility is well covered in the IETF draft [4]. The mechanisms developed in this paper follow the general handshaking principles of this draft but repurposed within the context of IMS.

1) SIP Session Mobility- IETF RFC [4]
Session Mobility involves five types of components: A correspondent Node (CN), a Mobile Node (MN), one or more local devices used as targets for session transfer, Directory Agent (DA), and, optionally, a transcoder. The Correspondent Node (CN) is a basic multimedia endpoint being used by a remote participant and may be located anywhere. It may be a SIP User Agent (UA), or a Public System Telephone Network (PSTN) phone reachable through a gateway. The Mobile Node (MN) is a mobile device, containing a SIP UA for standard SIP call setup, as well as specialized SIP-handling capabilities for session mobility and a User Agent (UA) for discovering local devices.

2) 3GPP Voice Call Continuity - Standardization of Voice Call Continuity (VCC) between WLAN and GSM access networks is underway in 3GPP and may be considered as part of wireless/wireline convergence. As next generation networks (NGNs) evolve and are deployed, there will be increasingly a need to provide continuity of service between different types of access networks and provide a consistent user experience. The User Equipment will typically be a Dual-Mode (Mobile CS and SIP PS over WiFi) device that could operate in either network domain, depending upon availability and preferences.

For voice and real time multimedia the goal is to provide call continuity across different access networks which utilize different technologies and support the migration from legacy to IP based communications. In addition to the IPTV session switch there is also a need for supplementary service treatment like incoming call routing features, caller ID display etc. For data, session continuity translates to maintaining VPN tunnels such that the user is always connected as the visited network access changes.

In this paper we make a distinction between nomadic usage where users are intermittently connected to the access network (normal data interaction mode (HTTP) etc.) and seamless mobility where consumers are watching IPTV in one viewing station and simultaneously switching to another device. In a previous paper, Kalmanek et al. studied an IMS based architecture to support seamless voice mobility [5]. What we describe next is an IMS based architecture that supports IPTV. We are specifically interested in IPTV session mobility within the context of the NGN architecture.

III. SYSTEM ARCHITECTURE

Figure 2 and Figure 3 show a layered model and a high-level IPTV network architecture being supported by an IMS infrastructure. Three functional layers are defined, namely, the Service layer, the Control layer and the Media layer, as shown in Figure 1. The layered architecture facilitates interoperability amongst different vendor solutions and maintains ease of service creation. The IPTV service layer provides multimedia services to the end user by means of the IPTV Application Platform (IAP), as shown in Figure 2. The IAP implements the portal with which a user interacts and includes functionalities like the electronic service guide (ESG), VoD etc. The IAP interacts with the IPTV Terminal Function (ITF) that handles display and interactivity functions for users. It also performs functions such as content encoding/decoding and buffering for both unicast and multicast streams. The system is divided into a number of logically separated parts, namely, the home network, access network, aggregation network, and the service provider domain.
A. IPTV user profiles

In the IMS IPTV architecture, personalization is an important feature. To achieve personalization at the application level (i.e., personalized EPG’s, advertisements, or even personalized blended communication services), every user has an IPTV profile. The relation between the IPTV profile and the IMS profile depends on the availability of a home IMS gateway (HIGA) [4]. The HIGA is a functional block with an attached ISIM card reader, which can be deployed in the residential gateway or any other networked consumer equipment. The HIGA translates home signaling, whether SIP, UPnP or perhaps pure HTTP to IMS signaling. It also takes care of NAT traversal and secure connectivity with the P-CSCF in the IMS domain, as well as identity, device subscription, and management inside the home domain and towards the IMS core.

In a home network domain without an IMS gateway, every IPTV account needs to have IMS public/private ID pairs for users of the system. These are used to log into the IMS domain. However, since the TV is a social device, one which many users are quite often watching TV together, the IMS IPTV STB contains a default user which represents the household itself (for example, sip:family@op.com). In this way, the STB can be configured to use the default household user to log in to the IMS domain, personalizing itself with the default values for the whole family. It can also be configured to make the initial login of a personal user (i.e., sip:dad.family@op.com). Personal user profiles may be protected with a PIN that needs to be typed in via a remote control to select a particular profile.

If the household contains a HIGA, the family members can choose whether they want to have full IMS identity, one which enables them full communication capabilities supported by IMS, or they opt to just have an IPTV profile, that will use the default IMS household identity for authentication purposes. The IPTV profile information that needs to be shared between different IMS services is stored in the IPTV XDMS database [5]. This database is accessed using XCAP [6], which works over HTTP. These profiles can be shared by different users and other stakeholders within the IPTV system.

B. The Multicast Data Channel (MDC)

The multicast data channel is a special IP multicast pipe that allows the IPTV application server (or any other authorized node) to transmit information to all STBs registered with the IPTV service. Each STB joins this special multicast group upon startup and keeps listening to it for as long as it is powered up. The MDC can be separated on different multicast groups for special purposes.

The MDC carries information wrapped in an XML schema, which provides the ability to differentiate the various pieces of information, like:

1. Electronic program guide information, a link to download the EPG from a server, the EPG itself in XML form, or even EPG updates.
2. Interactivity triggers, in the form of scheduled actions to be displayed/executed in the STB; an HTML page, showing a pop-up with information or triggering a special interactivity mode in the STB.
3. Firmware upgrades. The MDC can carry an order for all STB’s to download an upgrade, immediately or schedule to some appropriate time.
4. Alert or emergency messages, which should be shown as immediate pop-ups in the STB; these may not disappear until the user acknowledges them.

For each piece of information in the MDC, there is a tag with a timestamp, which marks the validity of the information; there
is also a tag that marks whether the information is included in the XML content (i.e. the EPG goes inside the XML-wrapped content), or whether it should be obtained via some other means (i.e. an HTTP GET to a particular server, or a file transfer of some type). The XML wrapper contains a set of tags to identify the desired receivers of the content. In this way, the MDC can contain information that it has tagged to be received by a subset of the STB population. Typical tags that can be used include the following:

1. Channel being watched: only STB’s currently displaying that channel will react upon the information.
2. Age: only STB’s whose active user falls into the age range will react.
3. Region: only STB’s located in the specified region will react.
4. Gender: only STB’s whose active user has the desired gender will react (if field is populated)

The filtering of the received information according to the XML tags happens in the STB. In this way, users can decide how much personal information they want to configure in their personal profiles in the STB.

C. IPTV control plane

The control plane of the IPTV architecture can be divided into a set of functions, like session setup, media flow setup, media flow control, and non-media related functions. The choice of protocols for each function tries to re-use existing standards or IETF standardized protocols when possible. The key components of the IMS control layer are the x-CSCFs and the HSS. The S-CSCF evaluates all originating and terminating messages and may, based on information of the service, link-in during session setup any number of IMS Application Servers (ASs) to perform desired IPTV services. For all IPTV related SIP messages originating from the STB, the IPTV AS will be linked-in. For IPTV, the HSS keeps triggers and filter information for the IPTV Public Service ID (PSI) or the service identifier. The information is stored and conveyed on a per IMS Application Server basis. This means that IMS ASs are allocated dynamically and that SIP messages will be routed all the time to the same IMS Application Server. The S-CSCF downloads rules and triggers, upon user registration, from the HSS. Figure 3 shows an IPTV implementation architecture using the appropriate IMS entities.

IV. SESSION FLOW SEQUENCE DIAGRAMS

Two session handoff modes are available namely a) Mobile Node Control mode (MNC) and Session Handoff mode (SH).

A. Session transfer to a single device - MNC mode

In Mobile Node Control mode, the Mobile Node uses third-party call control. It establishes a SIP session with each device used in the transfer and updates its session with the CN, using the SDP parameters to establish media sessions between the CN and each device, which take the place of the current media session with the CN. The shortcoming of this approach is that it requires the MN to remain active to maintain the sessions.

B. Session transfer to a single device – SH mode

In session handoff mode the session is completely transferred, relinquishing control of both media and signaling to another device.

Figure 4 - Flow for transfer to a single device (MNC mode)

Figure 5 - Session handoff mode transfer to a single device

Session handoff mode uses the SIP REFER method. This message is request sent by a "referer" to a "referee," which "refers" it to another URI, the "refer target," which may be a SIP URI to be contacted with an INVITE or other request, or a non-SIP URI, such as a web page. This URI is specified in the "Refer-To" header. The "Referred-By" header is used to give the referer's identity which is sent to the refer target for authorization. Essential headers from this message may also be encrypted and sent in the message body as S/MIME to authenticate the REFER request. Figure 5 shows the flow for...
transferring a session.

C. Session transfer in IMS IPTV with associated applications

Transferring an IMS IPTV session with all its associated embedded applications requires the IPTV AS to be in control of the state information to be transferred to the new device. For example, one possible application could be callerID on screen. The callerID on screen embedded application is personalized to the user logged in to the IMS IPTV service, so that only those users that have requested the application will see the callerID popup in the device (viewing station). The callerID works by sending a SIP SUBSCRIBE for a callerID event to the IPTV AS, or a dedicated callerID AS. The SUBSCRIBE contains the information of the tel: URI that the user wants to receive callerID information for. The AS checks the rights of the user to receive callerID information for the requested number, and accepts or rejects the subscription. Figure 6 shows the control events for a callerID display.

When a VoIP call is received for the family VoIP phone, the S-CSCF forks the SIP INVITE towards the AS (the IPTV AS or the dedicated CallerID AS), which retrieves the To and From URI’s and answers with a 488 service unavailable. At the same time, the original INVITE is forwarded to the recipient IP phone, which can proceed with the normal call handling. The AS then checks the active subscriptions to the called URI, and generates a SIP NOTIFY that contains the caller URI to each one of the active subscriptions. In the example depicted in the figure, both a CallerID on screen in the STB and a pager device will receive the NOTIFY and show the desired information.

Once the CallerID application is running on a device as part of an IPTV session a session transfer needs to take into consideration transferring the state of the called application. Figure 7 shows the steps of a session transfer from a STB to a PC. This scenario could correspond to a family in which a member is watching a personalized TV session in the living room TV but when new members arrive, he decides to move his TV session to his PC in his room.

The session transfer depicted in the Figure 8 corresponds to a session handoff from one device to another. When the user logged in the STB wants to send his personalized TV session to his mobile, he uses a menu in the STB that allows to send the session to a predefined PC client, a predefined mobile or a different sip: URI.

The user configures the sip URI’s of his mobile when building his configuration profile for the IMS IPTV service. When the user selects the ‘Transfer to PC’ menu item, a SIP REFER is triggered towards the IPTV AS with a Refer-To header pointing to the sip URI of the PC client. The IPTV AS receives the REFER and checks for the required rights to transfer the session to the new device. If the rights are correct then the AS answers with a 202 Accepted message and retrieves the state information of the user’s IMS IPTV session. The state information contains the media characteristics as defined by the sdp negotiation in the associated INVITE to set up the media pipe, the channel being watched and position in the media (if watching VoD), which is kept by the AS and updated by the client when the channel zapping stabilizes, and the associated state for the session-embedded applications. In this case the application’s state is the callerID tel URI that the user has subscribed to.

The IPTV AS then sends a SIP INVITE towards the new device (the receiving PC client), with an SDP that describes the media to be transferred as well as the other elements of the session state information, and in parallel commands the
required elements in the media plane to prepare the media for the characteristics of the new device if needed. The interaction with the media plane occurs via the Mr interface towards the Multimedia Resource Function controller (MRFC). The MRFC is a specialized media control node which is able to locate the required asset in a media distribution or caching overlay, and control (using the Mp interface) the Media Resource Function Processor (MRFP) needed for transcoding the asset to the required format for the new device. The receiving client then completes the INVITE dialog and sends a SUBSCRIBE to the event associated with tel URI received as part of the INVITE from the IPTV AS. When the INVITE dialog completes then the IPTV AS instructs the required media nodes to transmit the media to the new device and send a SIP NOTIFY to the referrer node announcing that the transfer was completed successfully.

Upon receiving the NOTIFY with the successful completion of the REFER, the referring client can choose to terminate his own IMS IPTV session or keep it active. This behavior can be chosen by the referrer when sending the SIP REFER, so that the session transfer instead is a session replication.

Once the session has been transferred to the new device, any NOTIFY to the callerID watcher triggered by an external VoIP call will be automatically sent to the new device and to the old the device (if the referrer decides to replicate instead of transfer) or only to the new device if the referring session has been terminated. Figure 9 shows the callerID message popping up in the screen.

V. EXPERIMENTS

In order to evaluate the performance and feasibility of our IMS IPTV session mobility ideas, we have implemented a proof-of-concept prototype using off-the-shelf components. Figure 10 shows the components of our prototype. The home domain contains two IMS IPTV clients running in two emulated STB in Linux PC machines. The first client represents the STB in the family living room and is connected to a TV. The second client runs only on a PC screen and corresponds to the personal PC of the user doing the session transfer. A VoIP phone is as well connected to the home domain and registered in the IMS operator as tel:+1234@operator.com. The home gateway is connected to a router to which the different network nodes are attached. The IMS core corresponds to a P-CSCF and a S-CSCF both running in a linux PC, with an emulated HSS. We have chosen to use the OpenSER SIP router to provide the two CSCF’s. The IPTV AS is running in another linux machine, while our media plane is provided by a set of VLC (Video LAN media server) instances running in a separate PC. Live TV (i.e. multicast MPEG2-TS flows) is streamed from the media server; VLC commands are sent via a telnet interface from the AS to control the VoD assets.

With this test bed we have performed a set of experiments to evaluate the transfer delay of our architecture, with the following results:

- **Processing delay** for the session transfer in the AS - It represents the time it takes from the moment the user presses the menu ‘Transfer to my PC’ until a message pops up in the PC screen asking to accept the session transfer. The time is in order of 450 ms, which corresponds to the REFER, state check in the AS and INVITE being sent.

- **Media startup delay** in the client - It corresponds to the time it takes from the moment the receiving client accepts the session transfer until the media appears on the screen. This time contains as well the average waiting time for an I-frame in the client media player and the buffering the client. For live TV (i.e. MPEG2-TS multicast) the waiting time is around two seconds, which corresponds to a GOP size of approximately five seconds of MPEG-2 encoded media. For VoD, the time corresponds to the AS instructing the media server to start the streaming and locate the point where the VoD was when the REFER event was received and its value is under three seconds for the VoD assets that we are using (H.264 encoded RTP streams with a movie length of five minutes)

VI. CONCLUSIONS

This paper evaluated IPTV session switching between different viewing stations with the aid of IMS. We addressed a key technical issue namely, the migration of a complete service context (e.g. the currently watched video session and its caller ID service on IPTV) of a user between his devices. Call flows for the transfer of the media session and supplementary
services were discussed. A prototype developed to demonstrate this concept showed media startup delays during switching for live TV and VoD to be less than three seconds while processing delays were less than a second. The described architecture highlights the capabilities of the IMS subsystem to deal with complex service interaction scenarios, where service blending over multiple devices and access networks greatly enhances the overall user experience.

VII. REFERENCES


