Analysis and Development of an MHP Application for Live Event Broadcasting and Videoconferencing

Kristof Demeyere, Tom Deryckere, Luc Martens
Wireless and Cable group, Department of Information Technology, IBBT-Ghent University
Kristof.Demeyere@UGent.be, Tom.Deryckere@intec.UGent.be, Luc.Martens@intec.UGent.be

Abstract

The introduction of Interactive Digital Television (iDTV) is about to completely change the user experience of television in the living room. In our opinion, the iDTV infrastructure lends itself perfectly to support live event broadcasting and videoconferencing, both enhanced with interactive applications. These services have a vast application domain which includes plain videoconferencing but also video surveillance, T-learning and T-health. For decades people have been familiar with television thus it only seems natural to use this medium for live audiovisual communication. This article discusses different architectures and shows an implementation for the Multimedia Home Platform.

Keywords
Multimedia Home Platform, MHP, Interactive Digital Television, Live event, Videoconference, VodCast

1. Introduction

1.1 Interactive Digital Television (iDTV)
Interactive Digital Television is gradually replacing his analogue predecessor whom everyone has been so familiar with for a long time. iDTV offers superior image and sound quality. On top of that, it enables the user to interact with the broadcast programs. The magic box that unlocks the wealth of new services is the set-top box (STB). An STB is capable of decoding the broadcast digital video signal and provides local as well as regional interactivity.

1.2 Multimedia Home Platform (MHP)
MHP is the first common open middleware platform being deployed at large scale in set-top boxes. It provides the application developer with a number of API's suitable to develop a variety of applications irrespective of the underlying hardware of the STB. Applications developed for this Multimedia Home Platform will be interoperable with different MHP implementations resulting in a horizontal market. The MHP standard considers following profiles: Enhanced Broadcast profile, Interactive Broadcast profile and Internet Access profile. The first two profiles are incorporated in MHP specification 1.0 while the last one is present in the more recent MHP specification 1.1. Each profile refers to an application area thus describing the required set-top box capabilities. Enhanced Broadcast profile combines digital broadcast of audio and video with applications providing local interactivity. This profile does not require the STB to have a return channel. Interactive Broadcast profile extends the Enhanced Broadcast profile by enabling services with regional interactivity. To enable this regional interactivity requires the presence of a return channel. The last profile, Internet Access, requires a more sophisticated STB with more processing power and memory. This profile extends Interactive Broadcast by providing internet services. Most of today's set-top boxes labelled MHP compliant, currently implement MHP specification 1.0.2 or 1.1.2 (ETSI, 2002, 2005).

1.3 Live Event Broadcasting & Videoconferencing
Live event broadcasting and videoconferencing both require regional interactivity so at least the Interactive Broadcast profile has to be supported. With both services custom audiovisual content, as requested by the user, has to be delivered to the STB. This audiovisual content can range from the video from a participant in a conference to a specific camera angle video stream resulting in a wealth of application scenarios. In general this audiovisual stream will originate from a packet-switched network. Looking at a typical STB there are two ways to receive audiovisual content: the return channel and the broadcast channel. Given the flexibility by which each user should be able to start a video conference or select a live event to watch, it seems preferable to use the return channel. In that case the return channel should provide a sufficient bandwidth to ensure an acceptable video quality. Videoconferencing also requires that each participant is able to capture his own video and audio stream in order to transmit this media to other participants.

2. MHP Specification Limitations & STB capabilities

The MHP specifications provide a rich set of features for the application developer to use but lacks support of specific functionality required to implement live event broadcasting and videoconferencing. Although the application developer is bounded by the features provided by the API, it is worth taking into consideration the capabilities of current set-top boxes.

2.1 Return Channel
The return channel is the path through which the set-top box can communicate with the outside world. Basically, this return channel can take any form, ranging from a low-speed internet connection through a PSTN modem or GPRS, to a high-speed internet connection through a cable modem. The MHP specification defines a set of

interaction protocols accessible to MHP applications depending on the supported profile. Figure 2.1 illustrates this protocol stack (ETSI, 2002).

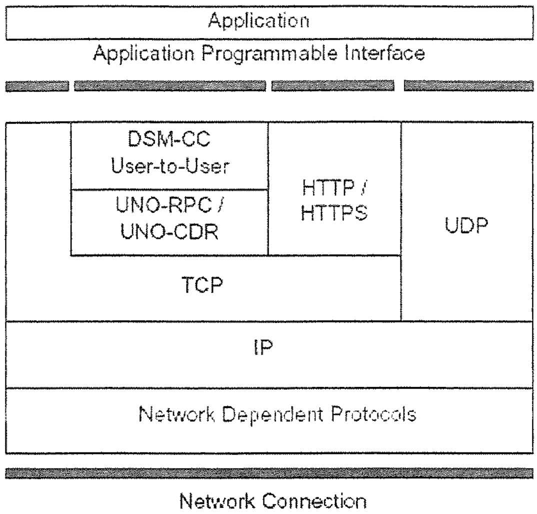

Figure 2.1 Return Protocol Stack

It should be noted that most set-top boxes do not implement al of these protocols. To stream live video to the STB it seems sufficient to use UDP over IP. However, when it comes down to using this return channel to receive content such as video or audio, there are limitations. Delivering streaming media through the return channel is currently not supported by MHP (ETSI, 2002, p. 52). It is however interesting to note that some STB manufacturers have or are about to equip their boxes with this functionality. The MHP standard can be improved by enabling playback of audiovisual content delivered through the return channel. This improvement would take the MHP standard closer to IPTV and MHP for IPTV has been proposed (Piesing, 2006).

There are some ways to try to work around the restriction. One way could be to render successive frames of a video sequence in the background layer or the graphics layer. This requires encoding the full video sequence as MPEG-2 I-frames but this is not an appropriate way to deliver video due to bandwidth limitations and latency in the network. Displaying the frames in the graphics layer would impose additional problems. Both methods are likely to fail in delivering an acceptable video quality. Moreover, this approach forces the separation of audio and video which might result in synchronisation problems at playback. Extending the JMF 1.0 API (Sun Microsystems, 2005) included in the MHP specifications could be another approach. The JMF API provides functionality to control how media is to be decoded and displayed. DataSources, as described in the JMF API, fetch media data and deliver it to a Player which decodes and presents the media. Providing a custom made DataSource could make it possible to receive media through the return channel. Unfortunately, such extensions are not supported.

As the return channel is not available for live media delivery, we propose an alternative by transferring the streaming media coming from a packet-switched network into the broadcast. This requires however that this content is correctly signalled in the broadcast and the STB is informed where to look for that specific content. This approach is the one used for developing this MHP application.

2.2 Peripherals

In the current MHP specifications there is no mention of other devices for the user to interact with but remote control and keyboard. Set-top boxes now available on the market have only a limited amount of support for peripherals like keyboards. STB with video camera support are rarely available. For videoconferencing this is of course a required feature. Although there is currently no support for such capture devices, it can be expected that future MHP specifications and set-top boxes will support numerous peripherals amongst which for example a USB webcam. It should be noted that in case an STB provides a connector for a capture device, the STB will have to deal with the encoding of the captured media in a format understandable for the intended receiver. It is unlikely that STB manufacturers will equip their set-top boxes with advanced hardware encoders, because this would substantially increase the cost of the STB unit. There are however several video coding standards such as H.261 and H.263, providing acceptable quality combined with affordable hardware.

The absence of a capture device can be solved by using intelligent peripherals incorporated in a home network together with the STB. The incorporation into the home network could range from basic IP connectivity to full integration into an existing home network using OSGI or UPnP. Some scenarios based on MHP and OSGI describing the interaction between digital TV receivers and home networks have been illustrated (Tchakenko & Kornet, 2004). The Open Cable Application Platform (OCAP) already provides an extension for home networking (CableLabs, 2005). Inside the DVB MHP group, the MHP Home Networking (MHP-HN) group is addressing this topic amongst others.

The solution we have chosen is based on incorporating the STB in the home network simply by connecting it to the residential gateway. The solution for the capture of live video and audio can be implemented in different ways. The first implementation uses custom capture software running on a host equipped with a webcam. The STB could discover and configure such capture software by broadcasting specific predefined messages on the home network. This capture software and the interaction with it should be implemented in such a way that the MHP application has the impression that the device is directly connected to the STB. Such a software module will be used in the architecture for videoconferencing. Another way is to use an ethernetcam to capture the video and audio. The ethernetcam is then located in the home network along the STB just like the capture software mentioned above. The ethernetcam is supposed to have the ability to capture and stream in MPEG-2 format. This way the stream might be incorporated in the broadcast. Such devices are available on the market but they are expensive.

3. Architecture Exploration

The following paragraph briefly discusses several possible architectures. Some of the architectures rely on special set-top boxes or use features that are not included in the MHP specifications. Nevertheless it is interesting to mention those solutions as they are likely to appear on the market soon.

3.1 Embedded Solution

STB manufacturers are incorporating delivery of media over IP in the STB design (so called IP STB) or extending current set-top boxes with special ASIC modules. In this case the architecture for any video delivery service such as videoconferencing or live event broadcasting is quite straight forward. The STB can be directly supplied with the media content it requested by streaming to the STB through the return channel. In this case a peer to peer architecture seems self-evident.

3.2 IP Triple Play in HFC

Some STB vendors have announced set-top boxes provided with the functionality to receive two signals through two separate tuners. One tuner will be used to receive the broadcast MPEG-2 Multi Program Transport Stream (MPTS), while the other one will be used to receive the (Euro-)DOCSIS downstream (Bugajski, 2004, p. 5). These STB, labelled dual feed set-top boxes, will be able to play media content delivered through (Euro-)DOCSIS by using the same MPEG-2 decoder as used to process the broadcast. Delivering media content to the STB in this case again requires only streaming the content to the STB using its IP address. In the backend the CMTS then inserts the media in the broadcast where the STB can receive it by using a tuner to process the (Euro-)DOCSIS downstream. A peer to peer architecture for this approach seems the right choice.

Both cases illustrate the migration to IP Television and enable many new applications including videoconferencing, video on demand and live event broadcasting. Embedded solutions may have greater flexibility because their range of supported video formats may exceed that of dual feed set-top boxes who could be forced to use the MPEG-2 decoder to process content.

3.3 MPEG-2 Transport Stream Pipe

Although the solutions mentioned above do have a large potential, most of the deployed set-top boxes do not have the ability to receive streaming media over IP. The MHP application under development is intended to run on currently deployed set-top boxes. For that reason the media content requested by the STB will be inserted in the broadcast and accessed by the STB in about the same way as it always does to view a service. Assuming that the media is located somewhere on a networked host, the architecture of figure 3.1 can be used in the backend to insert that stream into the broadcast.

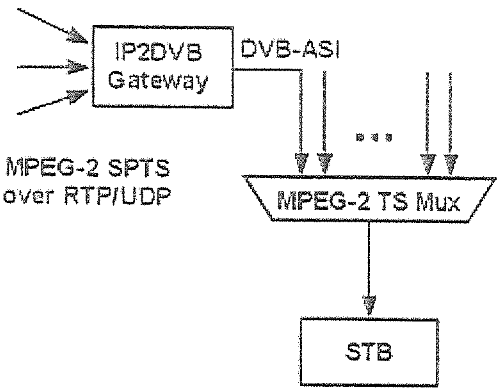

Figure 3.1 MPEG-2 TS Pipe

To establish the bridge between the IP network over which the media is delivered and the DVB network an IP2DVB gateway software module will be used. The software module runs on a host equipped with a DVB-ASI PCI card. Media content is being streamed to this host over RTP/UDP and the gateway module subsequently feeds this content to the DVB-ASI card. The ASI output will eventually be multiplexed with other DVB streams by a hardware MPEG-2 Transport Stream multiplexer (TS Mux) in order to form a valid MPEG-2 Multi Program Transport Stream (MPTS). The content at the input of the IP2DVB gateway could be delivered in any format of choice providing great flexibility at the expense of increased complexity due to transcoding. However, in the architectures for videoconferencing and live event broadcasting we assume that the media content being delivered is already coded as a MPEG-2 audiovisual stream and encapsulated in a MPEG-2 SPTS (Single Program Transport Stream), relieving the gateway module of the transcoding task.

In order for the STB to be able to find that specific inserted content it will need a DVB locator pointing to the service containing that video content. A DVB locator consists of the original network ID, the transport ID and the service ID of a service. These ID's inform the MPEG-2 decoder where to look for the service containing the video stream in the broadcast. When multiple MPEG-2 SPTS are combined to form a MPTS, an operation referred to as multiplexing, there can be a remapping of those ID values. Either it should be guaranteed that there is no remapping of the PID's in any of the streamed MPEG-2 TS streams containing video, or there should be some form of signalling to inform about that remapping. The first case facilitates informing the STB about the locator as the originator of the content can perform that task. However, this also means that there has to be some entity in charge of assigning PID values for the various content deliverers to use. This is where the centralized IP2DVB scheduler comes into play. The second case is more difficult to achieve; if the content deliverers are free to choose their

PID values, the IP2DVB gateway should also be capable of performing PID remapping.

The proposed architecture is a centralized one and relies on two central entities: the IP2DVB gateway and scheduler. Both will be used in the architectures for live event broadcasting and videoconferencing. The proposed architecture has several advantages. First of all it is compatible with current MHP compliant set-top boxes and as such imposes no additional costs for users. Secondly, for broadcasters it requires a minimum investment cost as it is an almost complete software solution. Furthermore the system is quite flexible: services containing requested media can be added and removed dynamically, making it ideal for video conferencing. An important issue for this architecture is resource dimensioning in terms of necessary transport streams or services for media like live video streams. Taking into account the number of users connected to a headend and presuming a certain blocking probability, cable operators can estimate the number of required channels by applying the B-formula or C-formula of Erlang (Parkinson, 2002; Hills, 2002) in the same way a telephone operator does. The latency of the proposed architecture will depend on the performance of the IP2DVB gateway. It should be remarked that there are hardware replacements for the IP2DVB gateway available which can reduce the latency at expense of less flexibility.

4. Architecture for Live Event Broadcasting

As mentioned earlier, live event broadcasting comprises broadcasting music concerts, sport events or even live lectures. Before any user can choose to view such a live event, he or she has to know about the event. To publish such information the usage of VodCasts is suited. Podcasting is a popular way to advertise available audio content. Podcasting consists of creating audio content, placing it on a server host and publishing the presence of the content by using PodCasts (which are an extension of RSS feeds). A so called aggregator collects information about available content, informs the user and fetches the content on request of the user. Using the same principle, one can advertise video content, also known as VodCasting. This approach has been chosen to publish live events. The architecture is depicted in figure 4.1.

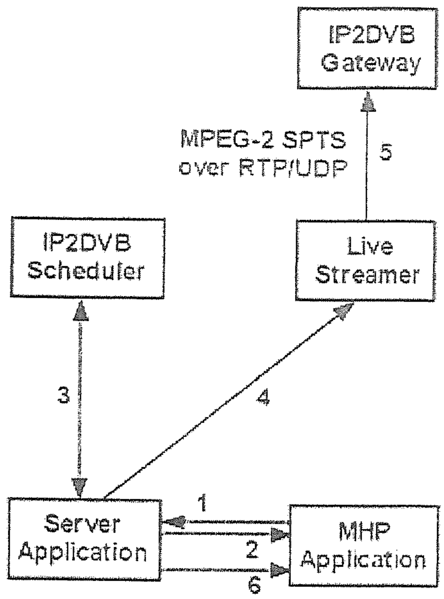

Figure 4.1 Live Event Broadcasting Architecture

The server application comprises everything needed to make the live event available to interested users. This application consists of a regular web server and a communication module. The web server can be queried by the MHP application. It answers to HTTP queries with the HTTP page of the content provider containing anchors to VodCast feeds describing the live events. In case a live event is requested the communication module negotiates with the IP2DVB scheduler to obtain the necessary information to start the transcoding and streaming. This communication involves receiving the IP address of the IP2DVB gateway and negotiating PID values for audio, video and PMT of the service. It then provides this information to the live streamer. This live streamer controls a capture device and is registered with the server application.

The live streamer performs the capturing, transcoding and streaming. The system is quite plug and play as new Live Streamers can be added to the system and registered with the server by simply supplying a VodCast feeds. The file gives a description about the live event and information about how the server application can locate and contact the live streamer.

The MHP application itself consists of an aggregator module, graphical user interface and XML parser. The aggregator module will query a server on behalf of the user and presents all the VodCast feeds present in the HTTP page in a catalogue on the television screen. It will use the XML parser to process the VodCast feeds. When the user requests to view a live event listed on the HTTP page, the aggregator will query the server for using the standard HTTP protocol.

A typical scenario shown in the illustration above comprises the following actions:

The user inputs a URL for the MHP application to query. The aggregator performs the query and processes the HTTP response message resulting in a catalogue appearing on screen.

The user selects a live event to view which results in the aggregator querying the server for this content using HTTP.

The server application negotiates through the communication module with the IP2DVB scheduler before starting streaming the content to the IP2DVB gateway.

The server application contacts the live streamer and provides the information it received from the IP2DVB scheduler.

The streamer application starts streaming live media to the IP2DVB gateway.

The server application sends a HTTP response message to the MHP application containing the DVB locator which causes the STB to tune to that specific service.

Although for this architecture only the Interactive Broadcast profile is required, the Internet Access profile would truly unlock the potential of VodCasting for iDTV. The architecture described above can not only be used for live event broadcasting but it can also be applied to provide Video On Demand (VOD). There is no restriction to what content that can be delivered on demand. The only requirement is that it is transcoded if necessary and inserted in a valid MPEG-2 SPTS before being streamed to the IP2DVB gateway.
The transcoding can be done by using open-source software like VideoLAN VLC. This media player provides the necessary transcoding functionality in combination with the flexibility to set various parameters like PID values for audio, video and PMT.

5. Architecture for Videoconferencing

The architecture for videoconferencing is based on that of VoIP. The description of the session in terms of parameters like connection or media information relies on the Session Description Protocol (SDP) (Handley & Jacobson, 1998). To initialize a multimedia session between participants, VoIP uses the Session Initiation Protocol (SIP) (Rosenberg et al, 2002). SIP will also be used in the third generation wireless networks. By using an MHP compliant set-top box as SIP endpoint a video call between a handheld mobile phone and a television set seems plausible. The architecture described below uses SDP and SIP to describe and initialize a videoconference session. The usage of SDP however will be limited to passing a DVB locator and other information as explained below. Figure 5.1 shows the architecture.

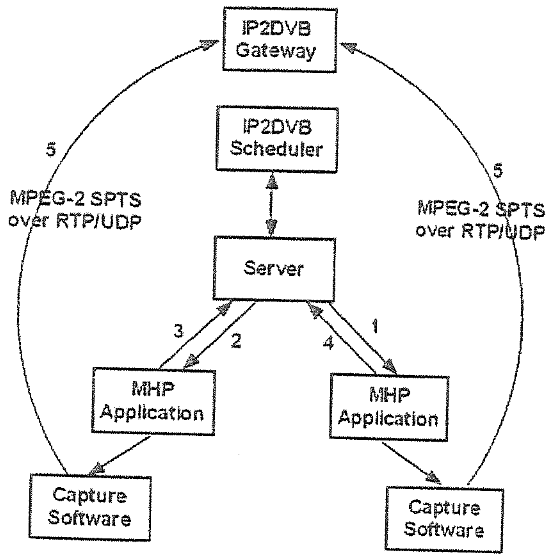

Figure 5.1 Videoconferencing Architecture

To handle SIP register requests from users the architecture has a SIP registrar which is located in the central server. Note that this server also includes a location service database containing the location of the registered users. This server will perform the lookup in case of an invitation requested by a user. Before the media session can be initialised, this server also has to negotiate with the IP2DVB scheduler in order to obtain the DVB locator that can be used.

The MHP application consists of a SIP module, a network module and graphical user interface. The SIP module will perform the registration with the registrar, service invitations ordered by the user and will listen for incoming invitations from other users. The network module handles the detection and possible configuration of the capture devices through the capture software located on a host in the home network. The capture software module will act as a peripheral connected to the STB. It will perform the capturing, transcoding and streaming of audio and video. This component could be replaced by for instance an Ethernet webcam.

After registration and configuring the capture module, the scenario depicted in the illustration above consists of the following steps:
A user wishes to invite another user for a videoconference session. The SIP module issues this command at the central server.
The server looks up the whereabouts of the invited user and forwards the invitation.
The SIP module of the invited user informs the user about the invitation. On acceptation, it responses to the invitation with an acknowledgement message.
The server forwards this message to the inviting user.
Both MHP applications order the capture software to stream to the IP2DVB gateway and tune to the service containing the live stream sent by the other participant.
Note that not all the SIP signalling messages are shown.
Furthermore there are two remarks. First of all, this architecture requires that the invitation message passed

by the server is extended with SDP information before being forwarded. This SDP information is limited to a DVB locator and configuration parameters. The locator presents the service containing the live stream from the other participant. The configuration parameters are meant to provide the capture software with necessary information for streaming. These include IP address of the IP2DVB gateway and PID values to form the MPEG-2 SPTS. The server extends the acknowledgment message with similar SDP information. Secondly, although the server adds SDP information before the addressed user accepts the information, this should not be a problem. On refusal of the invitation, the scheduling can easily be made undone when passing the refusal back to the initiating user.

Conclusions

Extending the Multimedia Home Platform with new video services like live event broadcasting and video conferencing cannot only improve the user experience of iDTV but has also the possibility of introducing new business models that can help the iDTV development. The here presented architecture, enables these new services by bringing the flexibility of an IP network and the high bandwidth of a DVB broadcast network together.

References

European Telecommunications Standards Institute (ETSI): Digital Video Broadcasting (DVB) - *Multimedia Home Platform (MHP) Specification 1.0.2*, Doc.nr.: TS 101812 v1.2.1, June 2002.

European Telecommunications Standards Institute (ETSI): Digital Video Broadcasting (DVB) - *Multimedia Home Platform (MHP) Specification 1.1*, Doc.nr.: TS 101812 v1.2.1, May 2005.

Sun Microsystems, Inc. (1994-2005). *Java Media Framework (JMF) 1.0 API Specifications.* From http://java.sun.com/products/java-media/jmf/1.0/

Piesing, J. (2006, March). *Roadmap for MHP and Related Technologies* Presented at DVB World 2006 – The World of Challenge.

CableLabs®: Open Cable Application Platform (OCAP) Specification OCAP Home Networking Extension, Doc.: OCAP-HNEXT1.0, Mai 2005

Tkachenko, D., & Kornet, N. (2004). *Convergence of iDTV and Home Network Platforms.* Russia, St.Petersburg: St.Petersburg State Polytechnic University.

Bugajski M. (2004). *Convergence of Video in the Last Mile.* Technical Paper for Jornadas 2004 Conference, Chile, Arica.

Hills, M. (2002, September). *Traffic Engineering for the 21st Century.* Retrieved March 11, 2006, from http://www.htlt.com/articles

Parkinson, R. (2002). *Traffic Engineering Techniques in Telecommunications.* Retrieved March 11, 2006, from http://www.infotel-systems.com

Handley, M., & Jacobson, V. (1998, April). *SDP: Session Description Protocol.* Retrieved November 20, 2005 from http://www.ietf.org/rfc/rfc2327.txt

Rosenberg, J., Schulzrinne, H., Camarillo, G., Johnston, A., Peterson, J., Sparks, R., et al. (2002, June). *SIP: Session Initiation Protocol.* Retrieved November 20, 2005, from http://www.ietf.org/rfc/rfc3261.txt

4th EUROPEAN CONFERENCE ON INTERACTIVE TELEVISION

EuroITV 2006

itv Beyond Usability, Broadcast, and TV

PROCEEDINGS

Edited by

Prof. Georgios Doukidis
Dr. Kostas Chorianopoulos
Dr. George Lekakos

25-26 May, 2006
Athens, Greece

ELTRUN: the E-Business Center
Department of Management Science and Technology
Athens University of Economics and Business (AUEB)