WHITE PAPER

IPTV-Based Video Telephony

Deploying IP Set-Top-Boxes for Live Interactive Video Services

NOTICE

USA/Americas
17-17 State Highway 208,
Suite #300
Fair Lawn, NJ,
07410-2819 USA
Tel: +1 201 689 6300
Fax +1 201 689 6301
info@RADVISION.com

EMEA
Abbey House
450 Bath Road
Longford, Heathrow
UB7 0EB
Tel: + 44 208 757 8817
Fax: +44 208 757 8818
infoUK@radvision.com

Asia/Pacific
Suite 2901 29/F China
Resources Building,
26 Harbour Road
Wanchai, Hong Kong
Tel +852.2801.4070
Fax +852.2801.4071
apacinfo@RADVISION.com

About This Document

This document describes the market need for IPTV-based video telephony and live interactive video services, and offers an overview of the technological challenges and solutions to develop the IP Set-Top-Boxes of the future for IP video telephony and other live, person-to-person, interactive video services from the home.

Table of Contents

The Case for IPTV-based Video Telephony

With the advent of support for HD (High Definition) resolutions on decoders, and System-on-Chip (SoC) integrated decoders rather than more cumbersome software solutions, the market had begin to adapt IPTV-based interactive video services. Today these services include VoD (Video on Demand) and EPG (Electronic Program Guide). But with emerging next-generation networks and technologies, the ground is ripe for introduction of full interactive video services to residential and business customers based on IPTV and IP set-top-box models.

Telco TV & IPTV – Current Status

As consumers continue to search for complete service packages that include voice, video and data, they will naturally begin to consider telco TV. In addition, telecom operators with significant broadband deployments seeking new sources to generate revenue will find IPTV-based telco TV a natural extension of current network infrastructure. IPTV is a system where digital television service is delivered using IP network infrastructure. In the residential market, IPTV is often provided in conjunction with Video-on-Demand (VoD) and may be bundled with other Internet services, such as Web access and VoIP. The commercial bundling of IPTV, VoIP and Internet access is referred to as "Triple Play." IPTV is delivered to residential subscribers using an IP set-top-box (IP STB).

The development and delivery of value-added services is key to the long-term viability of the telco TV (IPTV) business model. These services will be built upon IP STBs that deliver full interactivity.

Today's IP STBs allow only a limited range of interactive services. These include VoD (Video-on-Demand), mentioned before, PVR (Personal Video Recorders) and EPG (Electronic Program Guide). These services are either already available, or in late stages of development. Therefore, IP STB manufacturers have very little to work with to differentiate their IP STB offerings from the competition.

This is where the technology that enables manufacturers to turn conventional IP STBs into full-fledged, live interactive video service platforms comes in. RADVISION's technology plays an important role in this trend and offers all the technology and expertise needed to develop live face-to-face interactive video services on STB platforms.

IP STB of the Future

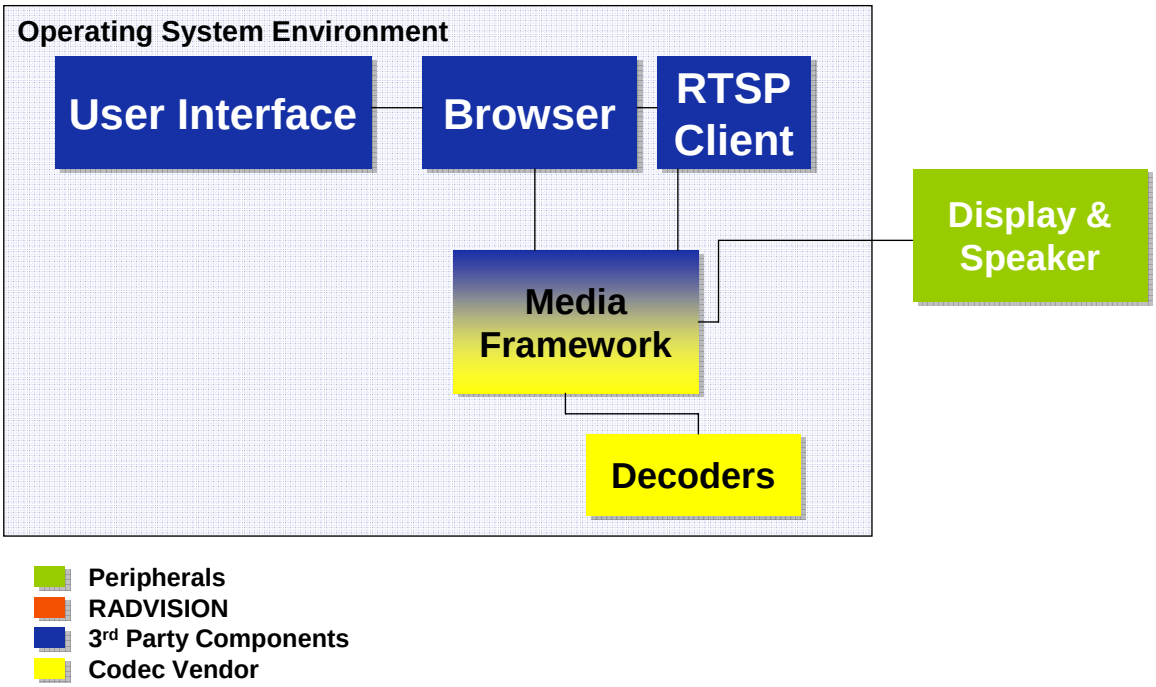
RADVISION's vision is to transform the IP STB into a true, interactive and rich communication center – the household's communication epicenter. This can be achieved by adding video conferencing technologies to the IP STB itself, which enables a variety of new ways to use the IP STB.
According to recent market research published by Frost & Sullivan (October 2006): "**The adoption of common standards for IPTV interoperability is crucial** if there is to be mass adoption of interactive visual communications in the business and consumer markets. **The implementation of open standards developed by video conferencing** offers the means of connecting these disparate islands of users in a manner similar to that of the IM federation, which is already successfully uniting the IM population."

Not only should interactivity not be limited to content; live interaction between people is a springboard for an entirely new paradigm of communication.

The Technology to Make it Happen

A New IP STB Architecture

If we take a peek inside today's IP STB architecture, we see that it is designed to decode incoming media streams and display them.

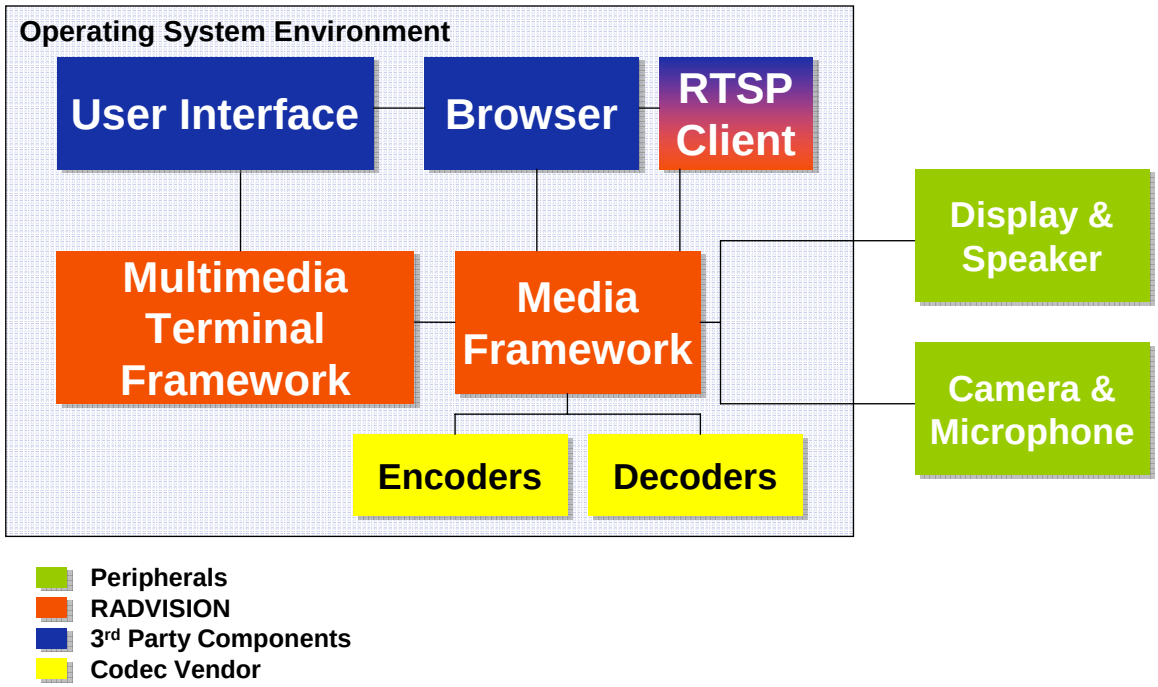


This is achieved via a dedicated Media Framework, capable of decoding multiple media streams and synchronizing them with one another. These components interact with the STB's browser and user interface.

To enable live interactive services on the same platform, such as video conferencing, several changes must be made:

- Encoders must be integrated into the system, with a sophisticated Media Framework that will synchronize incoming streams as well as outgoing streams.
- Camera and microphone peripherals must be connected to encoders.
- A call control solution is needed, such as the RADVISION Multimedia Terminal Framework, that can interact with the user interface and the media framework.

This architecture is illustrated in the following diagram:



RADVISION Technology – Powering the Paradigm Shift

With over a decade of experience developing and deploying video conferencing technology, RADVISION is powering a new level of interactivity designed especially for the burgeoning telco TV and IPTV industry. These technology solutions include:

Video clients for chipset manufacturers, middleware vendors and STB OEMs

Protocols and technology infrastructure for STB platform developers

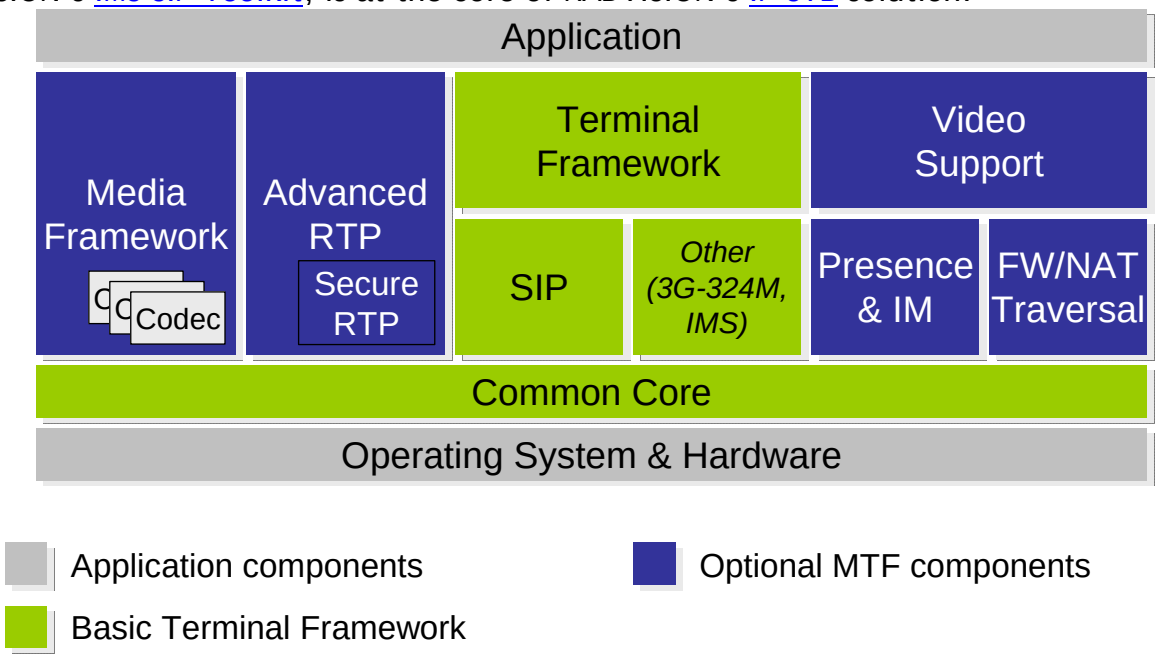
Network infrastructure and equipment for multipoint video conferencing, interactive video services and connectivity to mobile handsets

These solutions encompass the entire spectrum of technology required to develop IP STBs that offer all the live interactive video functionality required to differentiate from competing products. This technology can be found in the following RADVISION products:

- RADVISION Multimedia Terminal Framework
  Call control framework for audio and video conferences
- RADVISION Advanced RTP Toolkit
  RTP solution that includes SRTP (Secure-RTP) for privacy support
- RADVISION SIMPLE Client
  for developing Presence and IM applications
- RADVISION IMS SIP Toolkit
  Support for IMS in all audio and video calls, enabling interoperability for wireline and wireless operators
- RADVISION NAT Traversal Toolkit
  Built-in STUN client enabling NAT traversal for home deployments
- RADVISION 3G-324M Toolkit
  Circuit Switched video telephony support enabling direct connectivity to 3G mobile handsets
- ProLab™ Testing Suite
  Traffic emulation and analysis for signaling and media over IP to assist in development and pre-production stages

RADVISION's Multimedia Terminal Framework Solution

The Multimedia Terminal Framework, a call control layer developed on top of RADVISION's IMS SIP Toolkit, is at the core of RADVISION's IP STB solution.



The Multimedia Terminal Framework uses an operating system abstraction layer called the "Common Core", which makes it suitable for virtually any operating system. It supports Embedded Linux and Windows CE as well as a wide range of other operating systems.

The Multimedia Terminal Framework provides the fundamental building block to add video conferencing and other interactive services to IP STBs. The Video Support module adds video conferencing, while Presence & IM and NAT Traversal components add these features respectively. The media flow is handled by the Advanced RTP and the Media Framework, which together with the Terminal Framework provide a complete, interoperable solution for both the signaling and the media.

Integrating Video Telephony in IP STBs

Enabling video conferencing in the set-top-box requires tight integration with the STB chipset.

The first step requires connecting new peripherals to the STB, such as the camera and microphone. Next the codecs must be updated to those that support video conferencing. For example, today's STBs usually include MPEG2, MPEG4/AVC or VC1 decoders. Video conferencing uses other codecs sets and requires integration of algorithms. These include:

Video:

- H.263 – the most common codec used today for video conferencing. This widely used codec delivers relatively good video quality and has moderate computation requirements.
- H.264 (MPEG4/AVC) – one of the most advanced codecs standardized to date. This codec delivers superior video quality, but uses more computational CPU resources than other video codecs.

Audio:

Video conferencing requires a set of codecs for speech that are not usually found in STB platforms. These include G.711, G.723.1 and G.729, although a wide range of other speech codecs may be required as well.

Algorithms

Today's STBs do not need to handle outgoing media channels. The only algorithms required to perform their limited tasks today are jitter buffering, lip synchronization and packet loss concealment. To support video conferencing, acoustic echo cancellation and noise suppression algorithms are also necessary.

The Technological Challenge

The most difficult part of adding video conferencing technology to an IP STB is the video encoder element. As a rule of thumb, encoding usually requires 200% more processing power than decoding. This means that architecture designed solely for decoding will not be able to handle the encoding task.

There are three viable options to integrate encoders into STB chipsets. The choice depends on the chipset selected and architecture employed.

Technological Solutions

Option 1: Embed hardware encoders on the chipset

This option embeds the encoding element into the chipset itself, for a system-on-chip (SoC) solution. This requires a dedicated DSP, or hardware encoder, to be integrated as part of the chipset.

An external camera can be connected via USB, or via another connection to the chipset.

Pros
- Easy integration. Once the system-on-chip is available, developers need only access it.
- Optimal long-term solution for STB platforms that support video conferencing.

Cons
- Higher initial platform costs due to the required additional encoders.
- Most STB chipsets are not inherently designed for encoding.

Option 2: Embed software encoders

Similar to Option 1. This option embeds a software encoder into the middleware itself.

An external camera can be connected via USB, or via another connection to the chipset.

Pros
- Easy integration. The provided encoders are optimized for the specific processor

Cons
- Higher initial platform costs due to CPU requirements for encoding
- Integration is required for every single platform because encoders must be optimized for each separate platform

Option 3: Camera with built-in encoder

This solution does not require the STB to have an encoder, nor be capable of running one.

An external USB camera is used, which is capable of providing encoded data instead of only raw data.

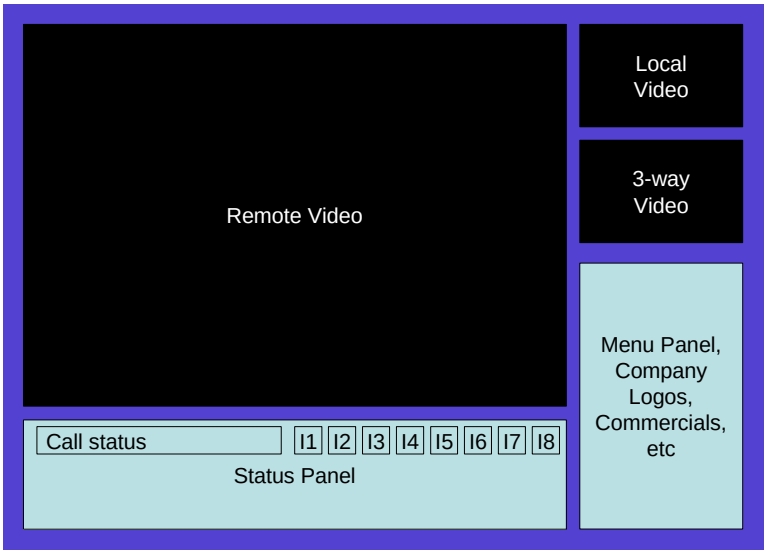
Pros
- Fits any STB that has a USB connection for a camera
- An optional solution – STBs can be deployed for regular use; encoder-equipped cameras can be added as required
- The sense of perceived privacy is greater; the camera can be unplugged from the STB at any time

Cons
- External camera is not an integral part of the STB itself
- Specialized drivers are required to this type of camera

The Next Step: The Optimal User Interface

Video conferencing on an IP STB requires integration with a Web browser to deliver the optimal user experience. This can be accomplished in a variety of ways, such as seamless integration into the Web browser itself. This enables instant video dialing while surfing the Internet, as well as the ability to receive incoming calls while viewing a TV show.

The example below illustrates a user interface capable of handling a dedicated video conference call and TV sharing.

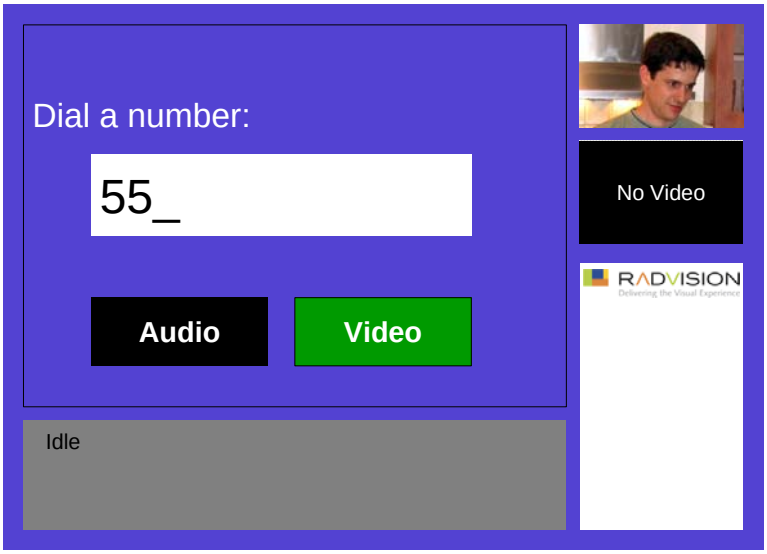


The example includes a main viewing screen which is open while in a call. The video screens on the side can display local or remote views, as well as streamed TV broadcasts, shown in parallel to the call.

The status panel is used to display the status of calls and other activities, in addition to other indicators that can occur during a call – displayed as needed. (I1 – I8). These include the various states of the call, such as hold, mute, call waiting, etc. Call status is a textual indication of what is occurring at any point in time.

The Menu itself is a list of strings; each can cause an action on its own or open a dialogue box, instead of the remote video.

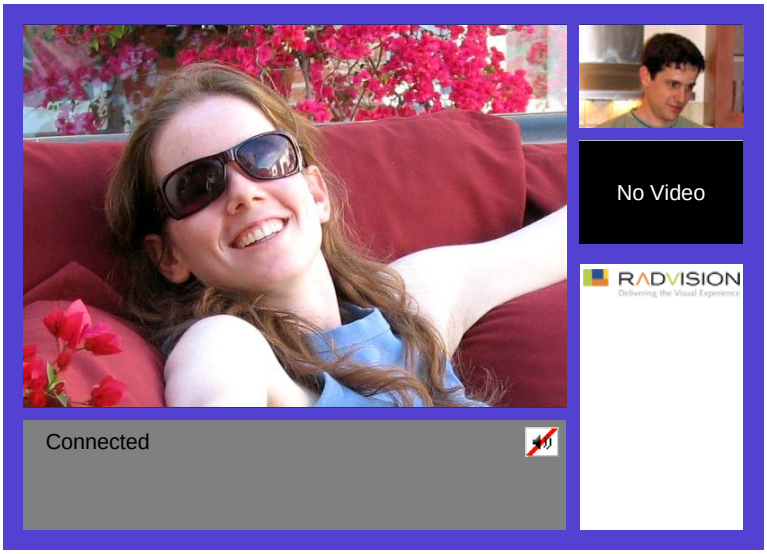
If a user wishes to dial out from the STB, the view of the main screen will change accordingly.



In the screen above, the "Remote Video" window, which doubles as an action screen, is currently being used to dial out a call. In this case, the local video can be displayed (as is the case above).

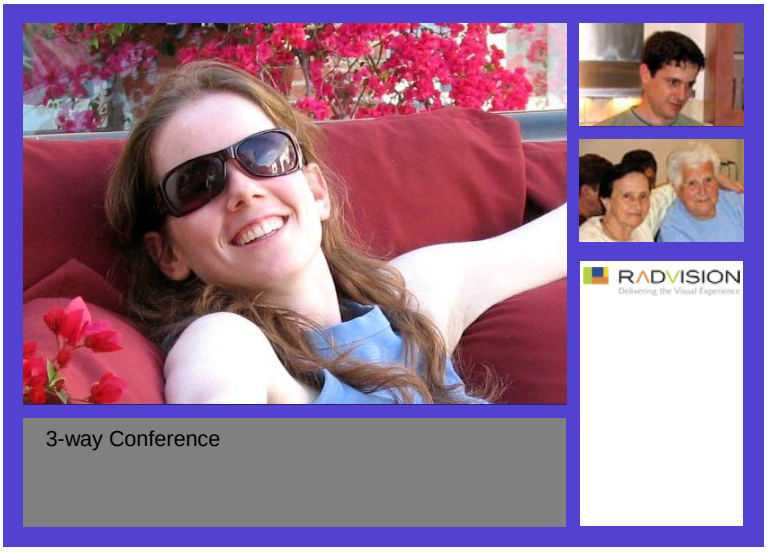
The "Audio" and "Video" buttons can be selected (only one) by pressing either the left or right on the remote control's navigation panel.

Once a call is connected, the screen is split into the remote and local views (below).

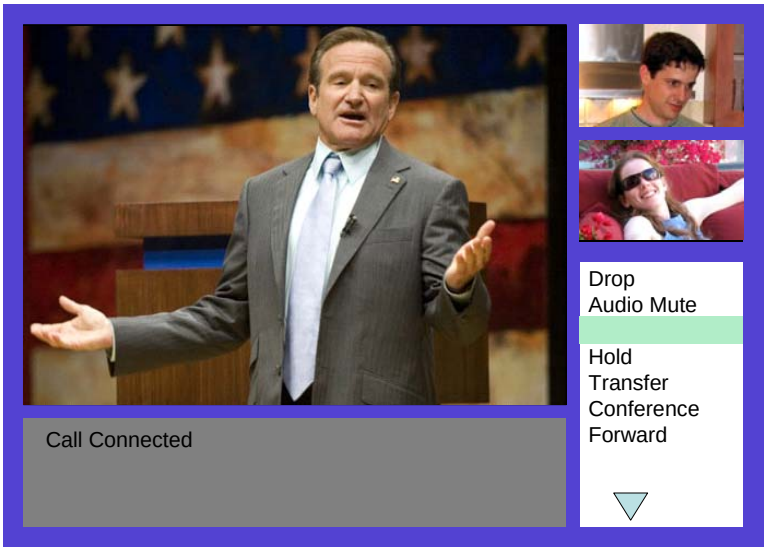


This example shows how a two-way video call would be implemented, with no TV or streaming content . There is no menu available at the moment and you can see that the audio is muted by the indicator in the gray area.

The same layout can also be used for a three-way video call.



or for a 2-way video conference along with a TV show.



In this example, the menu is enabled and is visible. The video conference is situated on the right side of the screen, in smaller frames, allowing users to continue watching the movie or TV show on the main screen.

A Plethora of Live Interactive Possibilities

While a very compelling service, video conferencing is not the only application that can be delivered using an IP STB. In fact, video conferencing is a "service enabler." It opens opportunities for a wide range or live interactive possibilities and a new set of usage models for the STB.

A number of these new service possibilities are outlined below:

Interactive Video Calling

This is the basic service enabled by video conferencing. It allows people to conveniently dial from the comfort of their home to friends and family through their television set, receive voice and video calls in an easy-to-use manner. Video calling can be implemented between STBs or to any other SIP client, such as a desktop of wireless device. This service can be enhanced with call recording and playback to enable a local, intelligent voice mail system. And above all – it can be upgraded easily to support IMS – the mandated Next Generation Network (NGN) architecture for both fixed and mobile networks.

Three-Way Conferencing

Three-Way Conferencing can be supported from the STB itself and does not require any external Multipoint Conferencing Unit (MCU) – if the STB's platform can handle the workload. This compelling value added service brings together several family members or friends, and can be further enhanced using RADVISION's MCU to enable seamless "n-way" multipoint conferencing.

Click-to-Dial via the Web

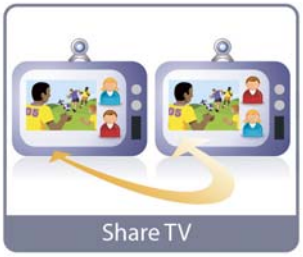
Using the browser to surf through Internet sites on their TV, users can click on dynamic links to immediately receive personalized video calls to sales agents or other representatives. This enables easy access to contact centers with no need to dial a number. This provides an excellent way to upgrade a shopping channel's user experience.

Presence and IM

STBs can leverage upon the wide deployment of IM and presence systems to deliver this familiar method of communication between family and friends. Dialing through presence information instead of memorizing numbers and the possibility of sending instant messages, instead of calling them, enhances the user experience.

Quad-Play Connection to Mobile

STBs also support quad-play, by adding mobile connectivity solutions and enabling the use of an STB as the communication appliance of choice while at home or when communicating with mobile phones.



Share TV

As STB platforms become more robust, the possibility of managing several media streams simultaneously on a split screen enables TV viewing while conducting video calls at the same time. This adds an entirely new dimension to sports broadcasts and gaming.

TV Shopping

Shopping channels can be enhanced to enable viewers to dial a sales representative and make a purchase directly through the TV set, and in a face-to-face conversation. The phone is no longer required.

Conclusion

Expanding on the success of IPTV-based interactive video services such as VoD (Video on Demand) and EPG (Electronic Program Guide), the ground is ripe for introduction of full interactive video services to residential and business customers based on IPTV and IP set-top-box models. Open standard video conferencing offers the means of connecting people – in real-time – to foster mass adoption of interactive visual communications in the business and consumer markets. There are many viable, revenue-generating value added services that can be deployed based on the IP STB foundation, which will be a springboard for an entirely new paradigm of communication.