THE BEST OF 2008

Technical Review

BBC iPlayer

DAB Open Source Handheld

SVC Video Compression

EBU HD Codec Tests

EBU P2P Media Portal

EBU Audio over IP

EBU·UER

Welcome

The best of 2008

Welcome to the "Best of 2008" selection of articles which have previously been published electronically in EBU Technical Review at **http://tech.ebu.ch**

Since 1998, the Technical Review has been published on-line, four times per year. This venture has been very successful because it has expanded the readership. It has been very gratifying to meet people who have "discovered" the on-line version. Some of these new readers are not directly connected with the EBU but they work in related areas, such as suppliers of hardware or software to the broadcasting industry.

Our electronic publishing service includes an archive of the Technical Review, dating back to 1992. You can access these archive articles by clicking on the "Archive / Thematic Index" link, in the navigator frame to the left of the screen. This archive contains a wealth of excellent articles on many different topics. Recently we added a "Hot Topics" section to the on-line version which brings together relevant articles from the archive on topics such as "HDTV in Europe" and "Broadcasting to Handhelds".

The on-line version also includes a list of abbreviations used in the Technical Review over the past 16 years or so. New terms are continually being added to the list which you can download as a PDF file by clicking on the "Abbreviations" link in the navigator frame.

At the beginning of the year 2000, the EBU abandoned the printed version of EBU Technical Review. Nevertheless, it was recognised that electronic publishing could not entirely replace "hard copies". This state of "Nirvana" will not arrive until we have computers that can match this paper publication in terms of portability, weight, readability, quality, price and (last, but not least) battery life!

As there is still considerable value in paper publication, the EBU decided to re-publish some of the articles in an annual printed edition of EBU Technical Review. Whereas the on-line version is available only in English, the annual edition is also available in French.

If you enjoy reading the articles in this publication, remember to consult the on-line version of EBU Technical Review at **http://tech.ebu.ch/**

Finally, do not forget to give this URL to your friends and colleagues!

Lieven Vermaele Director
EBU Technical Department

Evolution of the BBC iPlayer

Anthony Rose
Controller, Vision and Online Media Group, BBC

For more than ten years, EBU Members have been developing and refining their websites in order to enhance and augment their core radio and television broadcasting activities. The web is no longer merely an information medium (providing textual and pictorial information) but has become an audiovisual content-distribution medium for the internet-connected PC user – for both linear (scheduled) programmes ("channels") as well as for non-linear ("on-demand") programmes.

The BBC's development of the iPlayer is undoubtedly one of the best examples of how broadcasters can exploit the internet as a new media delivery mechanism. It can thus serve as a blueprint for other broadcasters to develop their broadcast services on the internet.

This article is based on a series of phone-calls in August 2008 between Franc Kozamernik (EBU Technical) and Anthony Rose, BBC Controller Vision & Online Media Group, which includes the iPlayer.

For the uninitiated, some background information on the iPlayer is provided in the box on *Page 5*.

***Franc Kozamernik (FK)**: There is a lot of interest among EBU Members in the BBC's iPlayer developments. The EBU Delivery Management Committee (DMC) set up a Project Group D/WMT (Web Media Technologies) chaired by Paola Sunna (RAI), in order to develop and evaluate a similar development termed the "EBU Media Player", which will be capable of delivering all kinds of content including the streams received from satellite, terrestrial, cable and IPTV channels as well as VoD and catch-up TV. What advice could you give to the group?*

Anthony Rose (AR): The biggest problem in developing services such as the iPlayer is typically not so much the website and media playout, or even the transcoding system, but rather the metadata and the content ingestion.

In the case of the proposed EBU project, a key design question is whether it will be an automated system that will capture content from satellite or another source, or whether there will be a team who manually process and ingest content.

***FK**: Our provisional idea is that our system will be fully automated. The system will allow users to find content via a variety of categories and other criteria. The metadata used will be broadcast via DVB-SI and TV Anytime, as appropriate.*

AR: There are a number of important questions which I think need to be addressed before one starts a project such as this. For example, is it the EBU's intention that each broadcaster creates their own website to where users can access this captured content, or will the EBU provide a so-called white label solution, which means that the

EBU develops a fully working website, which each broadcaster can then "skin" or brand to make it look like their own site? Will broadcasters need to arrange rights clearances for each territory, or can the EBU arrange this on behalf of all? Where will you source detailed metadata from (e.g. actor names, full programme descriptions, etc.)? Would you pull it from the DVB feed or will editors log in separately to apply enhanced metadata?

Perhaps the EBU project is more similar to the Redux project developed by BBC Research rather than the iPlayer? Redux is a technical trial using a fully-automated media ingest and capture system, is largely built on open-source technologies, and does not use DRM. Redux is being used within the BBC as a means for transcoding and providing content to BBC platforms. It is a very convenient and flexible input system.

In contrast, the BBC iPlayer is a well-staffed 24/7 operation with significant viewer traffic. We make sure that a comprehensive metadata scheme is exactly right. An essential asset of iPlayer is the right level of content protection for files and streams, as well as geo-protection, to address licence fee and content owner issues.

It is more likely that Redux can be made available to the EBU for testing rather than the iPlayer, given the sheer amount of resources that have been spent on making the iPlayer a viable commercial product. Many European broadcasters approach us with an interest in licensing the iPlayer. The question is whether they want a complete end-to-end system or whether they want individual pieces of the iPlayer production system, playout system or website. We spent several million pounds of taxpayers' money and could not give away that technology readily. However, in the case of Redux, the investment is substantially lower and the technology could perhaps be more readily available to 3rd-party broadcasters.

***FK**: We would like you to focus on the iPlayer now. What was the BBC's motivation to develop the iPlayer?*

AR: Our motivation for designing the iPlayer has been to develop a consumer proposition to satisfy the end user, i.e. the BBC listener and viewer, in an age where people are acquiring their entertainment from the internet, not just from their TV set. What does the user really want? They do not care about codecs and metadata taxonomy, they want to find content that interests them. We did not want the iPlayer to become a regular video-sharing site like YouTube or a music store like iTunes, where people would need to sort through thousands of programmes to find one of interest. This is a different use case. The reason why people like to come to the iPlayer site is because it allows them to find a particular programme that they missed on TV or radio. They want to catch up with what they know exists but were unable to enjoy at the time of broadcast. It is possible that over the coming months and years, the iPlayer will become a general browsing proposition, with demand driven by you or your friends rather than by the linear broadcast schedule. However, today it is focused on catching up with regularly-scheduled BBC radio and TV programmes.

When we launched iPlayer streaming at Christmas 2007, the home page had only six featured programmes and that was all. The BBC marketing team chose these six featured programmes. If you liked one of these programmes, you were in luck as this was exactly what you could easily find, right on the home page. The problem was if you did not want one of those programmes, you had to do a bit of work and browse by category, by day or search

The BBC iPlayer in a nutshell

The iPlayer is a web application – available at www.bbc.co.uk/iplayer – that allows internet users in the UK to download and stream BBC television and radio programmes for up to 7 days after the broadcast.

Users are able to download and stream programmes as soon as they have been broadcast on BBC TV and Radio. Users can keep downloads and watch them as many times as they like during the following 30 days.

For selected series, all episodes of the series are available for up to 13 weeks, known as Series Stacking. The iPlayer will in due course allow users to subscribe to a programme series and automatically download each programme after it is broadcast.

Recently, simulcast streaming was added, allowing users to watch TV live in addition to the on-demand catchup services.

The iPlayer services can be accessed on broadband internet-connected devices such as PCs, Apple Macs and Linux computers as well as Apple iPhone, Nintendo Wii and Sony PS3 gaming consoles, Nokia N96 mobile phones, Windows Media compatible portable media players, and Virgin Media set-top boxes.

To follow new developments of the iPlayer, go to www.bbc.co.uk/blogs/bbcinternet/iplayer/.

The iPlayer now has over 1 million users per day, and up to 1.7 million stream and download requests each day. The iPlayer should reach the 300 million play-request milestone early in 2009.

by name and so on. That might have been a complex (or even unsuccessful) operation, so we tried to make it easier to find a programme.

The first home page design was essentially *"the BBC chooses what you watch"*. Then we added a "most popular" zone on the home page – this was about what other viewers (rather than the BBC) recommended that you should watch. And then we also added a *"just in"* feature for those items that have just arrived and *"the last chance"* feature for items that would disappear soon. Finally, we also added a *"more like this"* option as a sort of recommendation system (similar to that used by Amazon). These content-selection mechanisms proved to be extremely useful and popular among iPlayer users.

***FK**: The iPlayer does not use any ratings, as opposed to ZDF's Mediathek in Germany. Why?*

AR: Indeed, we have considered adding a rating mechanism, but we feel it's only useful where applying a rating is a means of recommending that programme to your friends, rather than rating the programme in the way it's done on YouTube. If you have a video website with a million videos, possibly uploaded by the users themselves and often of mediocre quality, then you need a rating system so that users can say which are worth watching and which are not. In contrast, when you only have 600 programmes of professional quality, it adds little value to invite viewers to rate them. For example, how do you rate a Parliamentary channel? Rating BBC programmes would not add much value for the iPlayer user. In one sense, the programmes are all pretty good and marketed for different demographics.

However, we need to develop more personal recommendations – which programmes are good for "you". When we changed the site by adding the above selection mechanisms such as most popular, we made it much easier to find programmes. Before launching these features, we asked ourselves whether:

- people would watch more programmes because they can find more programmes;
- they would make fewer page views (because navigation is better);
- they would make more page views (because they may browse more, as browsing is easier);
- people would watch more programmes but would watch for less time (they may see recommendations for other programmes and would just click on something else before finishing the current programme).

Before we introduced these recommendation changes, there were about ten web-page views for every programme played. After these changes were introduced, the number of pages viewed dropped by 30% while the number of programmes played went up by 30%. These numbers showed that our changes actually helped people to find their programmes more easily. Finally, the number of page views per programme watched settled to about five and stayed there.

It is interesting that the average viewing time per programme did not change. We found that people watch a programme they chose for an average of 22 minutes. We also found that, on average, people watched two programmes per day, giving an average viewing time of about 40 minutes per person per day. About 35% of programmes are viewed all the way to the end. This is an excellent outcome, because our programmes are usually 30 or 60 minutes long.

***FK**: What is the editorial relationship between the BBC website and iPlayer? How are they differentiated?*

AR: The iPlayer is a destination within the BBC website. In many cases a given programme is available both within iPlayer and elsewhere on the BBC site, allowing users to discover and view the programme in the context in which they were browsing the BBC site. For example, most people used the BBC sports site rather than iPlayer for the Beijing Olympics. We're promoting the iPlayer as the home for long-format content. The sports site, the news site and other BBC sites are typically focused on shorter formats, like news clips and programme trailers. They also cover live events such as the Opening Ceremony at Beijing: live streaming was watched by over 100,000 simultaneous users on the www.bbc.co.uk website. A total stream capacity of 45 Gbit/s was provided by the Akamai content distribution network (CDN). For video coding, the On2 VP6 Flash format was used.

The consumption of Olympic programmes on the iPlayer was also very good. Many people who could not watch the Olympic events while broadcast on terrestrial, cable or satellite networks were able to use the iPlayer and watch those programmes delayed. For example, the Opening Ceremony was the most-viewed programme on iPlayer. It added more than 20 percent to the iPlayer traffic after the event [1].

***FK**: How would you describe the structure of the iPlayer system? Which are the principal layers?*

Abbreviations

CDN	Content Delivery Network
CPU	Central Processing Unit
DRM	Digital Rights Management
DVB-SI	DVB - Service Information
HTTP	HyperText Transfer Protocol
IP	Internet Protocol
ISP	Internet Service Provider
LLU	Local Loop Unbundling
PoP	Point of Presence
RTMP	(Adobe) Real-Time Messaging Procol
RTMPE	RTMP – Encrypted
RTSP	Real-Time Streaming Protocol
VoD	Video-on-Demand
WMV	(Microsoft) Windows Media Video

---

[1] *This interview was held during the Olympic Games. During the second week, as the Games moved into the final stage, the iPlayer consumption even increased by about 40%.*

AR: The iPlayer basically contains four layers, as follows:

- iPlayer destination portal site – this is what everybody sees;
- embedded media player – a Flash player which is used for media playout both in iPlayer and across the BBC site;
- media production – to create the content that can be used by the Flash player and is invisible to most people;
- a media distribution system.

***FK**: Could we start perhaps with the latter one first, please?*

AR: For On2 VP6 streaming, we currently use the Akamai CDN, whereas for H.264 streaming we currently use the Level 3 CDN, which is one of the biggest CDNs in the USA (in August 2008, Akamai did not provide for H.264 streaming).

***FK**: Why does the BBC iPlayer not use a Peer-to-Peer solution?*

AR: The BBC has explored a range of distribution solutions, but P2P does not currently provide the optimal proposition for streaming. First, viewers do not want to install any specific plug-ins. Currently to use P2P you need to install extra software. Second, P2P uses a computer's CPU and bandwidth, and most users generally do not like it.

If you are going to download some content via BitTorrent, you may agree to use P2P, and many people are happy to trade their bandwidth for free content. But in the case of the BBC, where people have to pay a licence fee of £130 a year, some are less than happy if we require that they use their bandwidth and install special software. This is especially true for people with low bandwidth and those who pay additional charges if they exceed a certain download limit. There were definite and substantial benefits from using P2P two years ago, but in that time the price of bandwidth has declined dramatically, such that today the use of P2P no longer provides substantial benefits. Of course nothing stands still in the technology world and, in a year or two, P2P may again be the preferred choice.

Of course, we know about Octoshape, Rawflow and a few others, and we have investigated using them for iPlayer distribution. But we are very happy with our current CDN-based streaming system; you click on play and the stream starts to render in about 300 ms. The only reason not to use a direct streaming facility could be cost and potential savings.

For downloading, we currently use the Kontiki P2P system which currently gives us a bandwidth saving of about 60 percent, so it halves our bandwidth bill for downloads. But we have to run a very complex server farm to make up the cost associated with it.

Actually the BBC is running a massive server farm itself, with over 200 computers, and we have 92 percent free peering. In fact, our bandwidth really does not cost us very much, at least not for downloads. If you look at all these various pieces, you wonder about the benefits of P2P. We believe that P2P works really well in some cases, particularly if you have a few programmes or a few files which are downloaded by many people, because then there is a good peering efficiency. It does not work well if you have an enormous catalogue, because the downloaded file only resides with a few peers.

For the Kangaroo project[2], P2P might work well for the 50 most popular programmes, but it will not be optimal to use P2P for a catalogue with lots of items.

We believe that the right approach is not P2P but caching at the edge of the network. We only have 500 hours a week of video content, which means that one TB of storage is enough to store our entire catalogue. This can be more efficiently done by simply putting a caching service in our network.

It may be a solution that, for the primary proposition, the user need not install any plug-ins. But you can have a secondary proposition which could offer better quality (say, high-definition TV). In this case, the use of a P2P plug-in may be

2) *Wikipedia article:* http://en.wikipedia.org/wiki/Kangaroo_(video_on_demand)

justified, because distribution costs for HD streaming are very high and could be significantly lower by using P2P.

***FK**: How about a combination of P2P and CDN, which is now increasingly used by both CDN and P2P providers?*

AR: With the iPlayer we have a bandwidth bill which is not insignificant, but it is something we can afford. We do something like 100 TB per day of streaming traffic. This is a fairly significant amount of traffic. The cost of bandwidth is falling very rapidly and there is a lot of competition between the CDNs.

At the moment the cost is not too excessive. But imagine in a year or two when we have a TV set-top box with an integrated iPlayer and millions of people using it, and each of them consuming 1.6 Mbit/s for a TV stream. The bandwidth required would be 10 times what it is now. Obviously, if this happens we will have a problem. The question is, what is the best solution for this problem. Is it P2P, should we make this new box with P2P or shall we build an edge-caching solution in conjunction with other broadcasters and ISPs? We do not know the answer but we need to build an agile architecture that allows different transport layers to be plugged in. We should separate the delivery layer from the content delivery formats, the DRM and the download manager, so that we can flexibly glue in different propositions at short notice as needed.

We are of course monitoring the developments of Tribler and other future-generation P2P approaches, as well as hybrid systems where P2P is backed to a caching box, for example.

***FK**: The BBC is renowned for its trials on IP multicasting. Could that be an option for iPlayer distribution too?*

AR: BBC Research has been trialling IP multicasting for a while. Different parts of the BBC may have slightly different objectives. In our case, we just want the iPlayer to work for everybody: go to the iPlayer site, find a programme on the home page, click it and play it. Other parts of the organization, such as BBC Research, look further into the future, and would like ISPs to build IP Multicast in their networks. Of course, we would like this as well, but the reality today is, as the UK statistics indicate, that only 5 percent of users are multicast enabled. It is probably not worthwhile to put much effort into making a multicast system for such a small number of IP multicasting-enabled users. It is really a chicken and egg situation.

Nevertheless, we are considering in the forthcoming months to use JavaScript or other means to detect if users are multicast-enabled, and if so, we may be able to give these users a higher quality stream. If they are not multicast-enabled, they would only get a lower quality stream. In this way, both ISPs and the users would have an incentive to introduce multicasting. The users are likely to choose those ISPs that have been able to upgrade their routers and can offer higher quality streams.

***FK**: There has been recently a lot of noise in the UK about the increase in network load caused by the iPlayer traffic. It seems that some ISPs have filed complaints with the telecom regulator?*

AR: The press largely misrepresented the situation by saying that due to the iPlayer, the internet will collapse and everything will come to an end. Of course, this is not true. We spent a lot of time talking to ISPs and we continue to meet with them regularly. The reality is that about 7% of peak UK internet usage is due to the iPlayer. So, the iPlayer service is only a small fraction of the overall traffic and will certainly not cause internet failure.

In the UK, there are three classes of ISP delivery networks: cable (example: Virgin Media), LLU (Local Loop Unbundling) and IP stream.

The cost of reaching the end user with cable is very low. In the case of LLU, the ISPs invested a lot of money in putting some equipment in the local exchange, resulting in a very low cost-per-bit. The third class, so-called IP stream, is a rented bandwidth from BT Wholesale.

If you are looking for some figures, there are in total about 5000 points of presence (POPs) around the UK. About

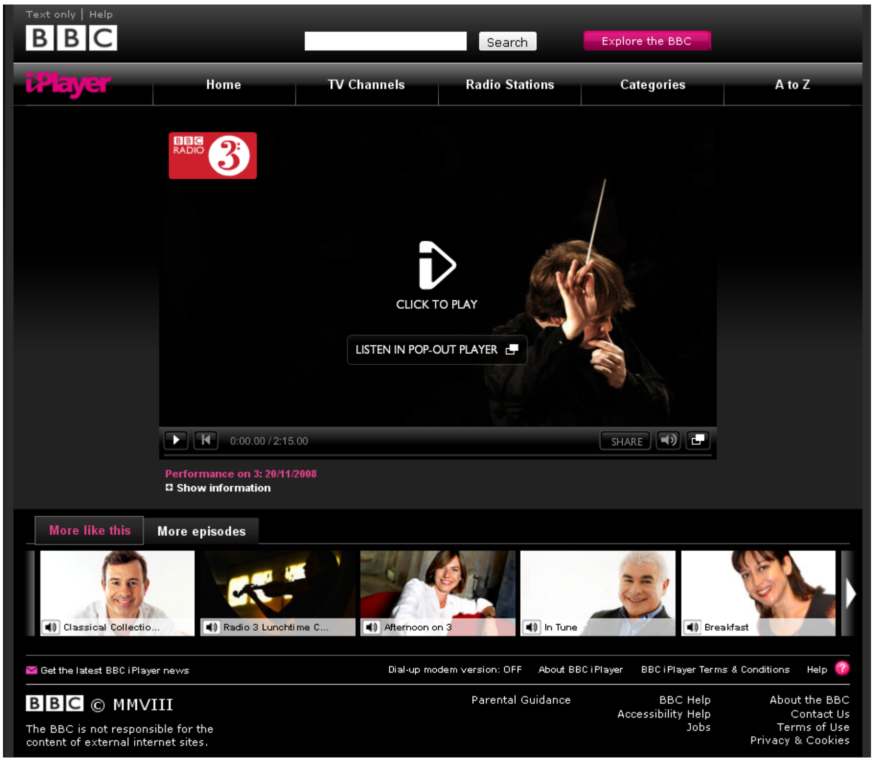
1500 of them are LLU enabled. About 30% of users are on cable. For cable and LLU the cost is relatively low, while for IP stream the cost of bandwidth is very high. This hurts those ISPs. There is no problem with the amount of bandwidth as the iPlayer is no way near reaching the bandwidth limit. However, our audience statistics show that iPlayer usage peaks in the hours between 6 and 11 p.m., which is also peak traffic for ISPs. The ISPs license the bandwidth for IP stream, based on peak usage. For this reason, iPlayer traffic is costing those ISPs. It is not just iPlayer, all traffic from YouTube, Facebook and other services is costing them. Our statistics indicate that this traffic is even larger than the iPlayer's traffic.

The situation is quite complicated as some ISPs like Virgin Media (cable) are offering 50 Mbit/s packages. This encourages people to use more bandwidth. Virgin Media is happy with the iPlayer and higher bandwidth consumption. Other ISPs that offer an IP Stream service are less happy because the iPlayer traffic is costing them more.

***FK**: So the situation is very complex, isn't it? How do you plan to resolve it?*

AR: The future lies in tiered services. What we need to do is to create the iPlayer services at different quality levels and then let ISPs offer different bandwidth propositions to users. For example, the user who enjoys higher bandwidth connections would pay more, and those who are satisfied with lower bandwidth connections would pay less. Of course, nobody should get a worse experience than today. We were offering streaming initially at 500 kbit/s. Today we are also offering 800 kbit/s and in three months time we might be offering 1.5 Mbit/s.

Some people will stay with 500 kbit/s, so they will not be able to experience our high-quality streams. If you sign up with Virgin, you will be on a 20 Mbit/s plan and you can download a film in 6 minutes, rather than in one hour if you only have a 2 Mbit/s line. So we could introduce a new scalable business model. For example, the user can get a good quality



iPlayer service for, say, £10 a month but for £20, a much better iPlayer quality would be available.

If we can create iPlayer in tiers, then ISPs will be able to work out how to sell that. Every content provider should create such quality tiers and then ISPs will be able to build business models around these propositions. This can lead to win-win situations and ISPs will see video services as a profit centre rather than a cost burden.

***FK**: Which bitrates are actually being used for streaming and downloading?*

AR: Back at Christmas 2007, we started with 500 kbit/s for live streaming and 1.2 Mbit/s for downloads coded in WMV (Windows Media Video). Now, we have introduced 800 kbit/s as well. In the future there should be no difference between downloads and streams but we are going to make a range of different bitrates, for example, 500, 800 and 1500 kbit/s.

The other thing we are going to do is pre-booking. The user will be able to download automatically a programme during the night. If you leave your computer on and if, for example, you watched Dr Who last week and the week before, it is likely that you will want to watch Dr Who next week. For ISPs, peak bandwidth is very expensive, but it is cheap during the night. We know that our top 20 programmes account for about 70 percent of all our bandwidth. In this way, most of our programmes could be delivered during the off-peak hours, downloaded and stored on the user's local hard drive. Thus, peak bandwidth usage could be significantly reduced. This is really a mixed economy where the difference between streaming and downloading is getting blurred.

In this scenario, our programmes will all be DRM'd and you will be able to either stream them or download them. A person with a good network connection will be able to stream, whereas the user with a poorer connection speed will download it and watch after the download completes or even during downloading.

The prime user experience is and will always be the iPlayer website. Imagine you go to the iPlayer website and you want to play something. Of course, you should not look at your hard drive to find out what is on it, your web page should now be smart enough to find out whether the programme is already stored on your local hard disk, and if it is, play it from there, rather than from the BBC server. This complete seamless integration of on-line and local playout is what we would like to implement in 2009. Another advantage is that users can simply unplug their computer and watch the downloaded programme offline, for example, while on an airplane.

FK: Recently the BBC introduced the H.264 codec for the iPlayer and some users complained about poor accessibility. Why?

AR: H.264 requires more processing power and better graphics cards. We have spent quite some time looking at this problem. There are a few H.264 compression settings that produce brilliant results but which require a high-end computer and graphics card. If you have a dual-core processor with a high-end graphics card, it looks fantastic, you can do HD at 4 Mbit/s. However, if you have a low-end portable computer, the quality is terrible, with the video running at 10 frames per second or less. So you need to carefully select the profile you use to ensure the video plays back seamlessly on a wide variety of target computers.

H.264 allows for three profiles – Base, Main and High – and for each profile you can turn on different features. We have gone for Main profile and we also turned on hardware scaling for full-screen playback, as the default. In fact, we have now found that H.264 does not use more CPU power for the configuration we have chosen, compared to the On2 VP6 codec. Rather, the contrary is true in full screen mode and, because we use hardware acceleration, it uses less CPU power. The answer is that, if you are not careful, H.264 is unplayable on low-end machines, but if you choose carefully, H.264 could be a pretty good user proposition. It is bit more complicated than that because the older Mac computers have problems with H.264 and can play On2 VP6 more successfully. With some older computers there is a problem. But with newer computers, again if you choose wisely, you can actually get a better experience.

MPEG-2 is old and no longer in the running, as bitrate requirements are far too high. Two other candidates for encoding are Microsoft VC-1 and On2 VP6 or indeed On2 VP7. Many people have evaluated these, and other codecs, and the outcome is that H.264 is generally thought to be the winner. But it is not always that clear cut. For lower-end computers, On2 VP6 is the best choice. On the other hand, if you are targeting Windows computers and full-screen playback, I think Microsoft has done a really good job with the Windows renderer, so that VC-1 plays back beautifully, even on lower-end Windows machines, but it does not work well for the Mac.

Microsoft Silverlight is a cross-platform application but it does not yet have the hardware-rendering capability that Windows Media Player has, which is unfortunate.

FK: Is this issue the reason why the BBC also considers Adobe AIR?

AR: Adobe AIR works fine with H.264 and is a clear candidate for the download solution with its DRM system, partially because we have a requirement to be fully cross-platform, and AIR runs on PC, Mac and Linux.

FK: Broadcasters often face the problem of codec licensing. What is your experience?

AR: iPlayer is now using H.264 and the question of licensing does not arise. If you use Flash, Adobe's agreements cover the playback licence fees. The BBC believes that there is no H.264 per-stream fee involved.

FK: BBC Research is developing an open source, licence-free codec called "Dirac". The EBU plans to evaluate its technical merits, as many EBU Members are potentially interested in using it for internet delivery. Does the iPlayer have any plans to migrate to Dirac?

AR: At the moment we believe Dirac is probably better focused on high-quality video encoding rather than on internet transmission. If you look at what is needed for successful internet transmission and for putting in the production workflow (using TeleStream, AnyStream or some workflow software), you need a codec that you can put in the workflow software. Then, you need a streaming server with a CDN that understands that particular codec format. You also need to have a rights protection model (DRM) and the user's computer needs a plug-in with a good renderer that can do frame-rate adjustments and so on. So, there are actually quite a lot of pieces that need to come together.

Currently Dirac is a stand-alone encoder and has not yet been worked into the different workflows. The Dirac player is not quite apt for real time on lower-end machines. There is no integration with CDNs and no plug-in has been developed, as of yet. Therefore it is premature for Dirac to be a consumer proposition at the moment but that will come with time.

FK: How important is Digital Rights Management (DRM) for the iPlayer?

AR: It is too narrow to look only at streaming and downloading. For general analogue or digital broadcast we do not have any DRM or any obfuscation, so people can do what they want, whenever, with the content received. Live broadcasting is readily recordable and there is no attempt to prevent people from recording it.

As far as streaming on the internet is concerned, we do not use DRM (in the conventional sense of the word) but we use some stream obfuscation technologies. Essentially, a stream must remain a stream, it must not become a download. So if a stream remains a stream, we believe we do not need to DRM it. In order to prevent a stream from turning into a download, we use technologies such as RTMP or other technologies that make sure a stream remains a stream.

FK: What experience with using RTMP do you have?

AR: If you link to a media file served from an HTTP server, your media player will pop up and begin playing it and that is called a progressive download. The played file would probably end up in your browser's cache and it would be very easy to copy this link and place it in another application which lets you save it. The problem with this approach is that it becomes easy to save a file that is meant to be streamed only. So we do not do that. Instead, a lot of companies offer streaming solutions which do not let you easily save the file. It will let your media player throw away the segments of the file after they've been played, rather than allowing them to be saved to your hard drive.

Microsoft has a solution and the product is called MMS. Then there is RTSP (Real Time Streaming Standard) which is an open standard, and Adobe has a proprietary standard called RTMP (Real Time Messaging Protocol) and another one, RTMPE, which is an encrypted version. The latter one offers better protection but requires more CPU power on the user's machine. Currently we do not see the need for it, as there is no widespread evasion or hacking. We monitor regularly whether content hacking occurs and, at the moment, this is not the case. Also, as the same programme was broadcast in the clear the evening before, the cost benefit is not there and we do not really see the need to DRM our streaming content.

Now, for downloading our position is different. For downloading, we have to DRM our files for two reasons. First, the rights holders expect that the content will be available in the UK only. Second, content must only be available for a limited amount of time, so it can be commercially exploited, as is the case with BBC Worldwide's licensing of the Top Gear programme. Broadcasters in the USA who pay BBC Worldwide millions of pounds for broadcast rights would probably pay less if there was no DRM, as the content would be available elsewhere. This is the main reason why the rights holders demand DRM. In addition, it is a requirement of the BBC Trust (the BBC governing body) that files are only available for 30 days after download and seven days after being broadcast. So these are the reasons why we have to apply DRM to downloads.

Not all content owners however demand DRM. For example, we do not need DRM for our parliamentary channel. However, with time and usage restrictions still in force, we do need to apply it. We have, of course, the open source community saying that we should not use DRM at all.

FK: You clarified why DRM should or should not be used for the iPlayer content, but then the question is which DRM do you use to control iPlayer usage?

AR: The open source community criticises us for using Microsoft DRM and tells us we should use an open-source DRM solution. We have done a lot of due diligence and we have investigated all the viable DRM solutions. We have met with companies that develop them and we looked at the technologies themselves and evaluated them. The reality is that, until quite recently, Microsoft was the only viable one. It is free, secure and approved by Hollywood labels and approved by rights holders. It is easy to put on servers and clients. The problem is, however, that it is Windows only.

Other companies with DRM, for instance Apple, do not give access to the DRM system. The only way to allow content to be available using Apple DRM is to put content on the iTunes store and that really means disaggregating our content. Therefore, we do not have BBC iPlayer

content available in the iTunes store. Apple would like us to give them our content and put it in a bucket with a million other programmes. For us that is equivalent to the BBC taking the content of BBC 1 programmes and giving it to competitors to put on their sites. This is clearly not acceptable. We have asked Apple for access to the DRM but so far they have not given us access.

The good news however is that other companies like Adobe are developing cross-platform DRM products. Adobe AIR now has DRM available for the PC, Mac and Linux. We hope to have a cross-platform solution by the end of this year based on Adobe AIR and Adobe DRM.

***FK**: iPlayer services are not available outside the UK. At my home in Switzerland I received a message "Not available in your area". Why do you constrain iPlayer to the UK territory?*

AR: Two reasons: one is the rights reason. Licence holders sell their content in each territory. Traditional broadcasts are geographically targeted by the transmitter radiation and TV is generally very short range. But on the internet, streams can go anywhere. Licensing models change dramatically, they are still limited by, or are working within, a TV broadcast framework. The BBC is licensed to broadcast in the UK and these are the licence rights we typically acquire.

The other reason is less obvious: public services are funded by licence-fee payers in the UK. As there is always a distribution cost on the internet, it is not fair for a licence payer in the UK to pay for distribution to someone in the USA watching the content. Even in cases where we have rights to broadcast outside the UK or make content available outside the UK, we would not do it in such a way that UK licence payers fund the distribution. BBC Worldwide might fund it or may cover the distribution costs or may have ads to support the model. For these two reasons, we need geo-locking.

***FK**: Which geolocation system do you use and how effective is it?*

AR: The answer is pretty simple. We use look-up tables of UK IP addresses, stored in a Quova database. These lists are regularly updated. We check the user's IP address and if it is located in the UK it is good and, if not, we say "sorry you can't have it".

Why do we not use the Akamai Geolocation database? First, it would lock us into exclusively using Akamai and we do not want to use Akamai for all services. In fact, H.264 content is now being distributed via Level 3 Communications Inc. It is strategically better that we have our own central control system. Second, we need to maintain the whitelists and blacklists, so for example sometimes we want to set up a proxy to try and access iPlayer outside the UK, so we need the means to control this ourselves and not to rely on Akamai.

Another reason for not relying on a CDN company's geo-location service is that we really want to alert the user that the video won't be available to them as soon as they view the iPlayer web page, rather than waiting for them to click the Play button and receiving a streaming error.

We really need to know the geo-location at the time we render the web page, so that we can give the user a nice message saying that the content is not applicable to the user: "Sorry you are not in the UK, you cannot play TV but you can play radio". If we just relied on the CDN company's streaming service to enforce the geo-location, then the user would receive a stream error message and no explanation why they cannot see the content.

Things are getting more complex now with 3G access. For example, you may have a roaming arrangement with Vodafone UK. If you are in France, our system may think that you are still in the UK, even though you are actually in France. This is a new challenging area. It is not a widespread problem yet because roaming access is so expensive that it would probably cost you a fortune to receive BBC programmes abroad via a mobile phone and hence few people try. However, we will need to tweak the IP lists and work with 3G vendors to make sure that you are in the UK, even if Vodafone UK has a roaming agreement with France.

***FK**: What kind of arrangements do you have with the ISPs to provide you with the users' IP numbers?*

AR: Quova makes those arrangements and it regularly updates the look-up tables. There is a way for ISPs to also update this information. We are quite happy with these arrangements; there is 99.9% effectiveness.

***FK**: You have ported the content to mobile devices such as the Nokia N96 mobile phone. Can you please outline the process for doing this?*

AR: We have addressed the content creation not only for PCs but also for some portable devices. Previously, if you used Windows Media Video (WMV) files and downloaded them onto your portable media player, the WMV files may either have been refused by the device or they were played with significant frame dropping. As of September, we are now creating content specifically for Windows Media compatible mobile devices. We are creating a special low-resolution version which is small enough to download and play nicely on these devices.

The standard resolution on these devices is 320 by 240 pixels. At the moment the resolution of our main PC profile is 720 by 544 non-square pixels, which gives best quality on a PC but it is not suitable for small devices, so we plan to make a number of special encoded formats for these portable devices. As far as downloading these formats is concerned, we will offer a number of different options. These formats may still be primarily available from the iPlayer site intended for a PC, but we will develop several custom websites intended for downloading content to different mobile and portable devices, such as the Nokia N96 etc.

For certain devices which we think offer a great user experience, we plan to design a special version of the site. Such a site

Anthony Rose is Controller of the Vision & Online Media Group at the BBC, where he heads a team of over 200 people who are responsible for the BBC iPlayer, embedded media player, social media, syndication, programme websites and other projects within the BBC's Future Media & Technology division.

Mr Rose joined the BBC in Sep 2007, prior to which he was at Kazaa/Altnet. During his six years with them, he worked on a host of projects and patents covering P2P networks, DRM-based content publishing and social networking services.

Prior to joining Kazaa/Altnet, Anthony Rose was Vice President for Technology at Sega Australia New Developments, developing real-time 3D animation and 3D graphics engines.

will tailor the content automatically to the characteristics of the mobile device (screen size, resolution, etc). The first of those devices was the iPhone. So if you go to the iPlayer site on an iPhone, you get a nicely tailored web version. The BBC will produce a custom version of the site for a selected number of other mobile/portable devices, so that the media will play automatically in the right format for that device.

***FK**: Do you plan to bring the iPlayer to STBs and consumer devices such as TV sets?*

AR: The answer is yes. The challenge is that these devices often try to aggregate different content into one portal. To the extent that the box is just a playout device like Windows Media Extender devices[3], the answer is broadly that we would like the iPlayer content to be there. To the extent that device manufacturers are able to offer the iPlayer site experience, we would like to work with them. But, to the extent that they would like to take the BBC programming and put it in their own interface, broadly speaking, that does not work for us. It is not acceptable for the BBC to just give away its content to other websites that can then build a consumer business proposition around it.

If you Google "BBC IPTV", you will see announced plans to work on IPTV set-top boxes that are already open and available to either everyone or selected parties. This is a very good second-generation IPTV proposition. One of the problems is that often there are not many of these boxes on the market and it is really very hard to get onto these STBs. In other words, it creates a huge amount of work for us but few consumers would use it. The cost benefit really does not work out, which is why we are currently not working on this project. Today there are quite a few different providers and the market is still relatively small, say, several hundred thousand subscribers. But this may change in the future.

***FK**: Thank you for clarifying the most burning issues relating to the iPlayer. I am sure that the EBU Members will find this article very interesting and useful. Should they have any further questions, could they approach you directly?*

AR: Sure. You can give them my email address if appropriate.

---

3) ***Windows Media Center Extender** is a set-top box which is configured to connect via a network link to a computer running Microsoft Windows XP Media Center Edition or Windows Vista to stream the computer's media center functions to the Extender device. This allows the Media Center and its features to be used on a conventional television or other display device. The household's Media Center can be physically set up in a location more appropriate for its role, instead of being in the living room. Additionally, with an Extender, the Media Center can be accessed at the same time by several users. The Xbox 360 gaming console is a very popular example of a Media Center Extender.*

Note from FK: Since the interview (which took place in August 2008), the iPlayer has been under-going constant software development, with new features and functionality added almost every week.

At the Microsoft Professional Developers Conference in October, Anthony Rose successfully demonstrated the syncing of iPlayer content across computers and mobile devices, using a Microsoft desktop application called Live Mesh cloud, which is based on cross-platform Silverlight technology. The application automatically synchronizes downloaded shows across all the iPlayer-compatible devices on the person's Mesh network. That includes Mac computers, which also have a download client for iPlayer.

The new prototype iPlayer also featured several social-networking features, such as lists of the most popular shows watched by friends on the MSN Messenger list and updates on which shows each of the contacts had watched and downloaded. iPlayer users will also be able to rate scenes from the show as they go along – using the "Lovemeter" – which shows the parts of shows that people like the most.

Erik Huggers, the BBC's Director of future media and technology, recently stated[4] that the success of the iPlayer is proof that the corporation is right to bet its future on the internet. He stated that the online TV catch-up service has served 248 m items of content since it launched officially on Christmas Day 2007. The iPlayer service that is available through Virgin Media's cable service alone has served 49 m videos since June 2008. The soap drama, EastEnders, which pulls in an average of 18.9 million TV viewers each month on BBC1 and BBC3, attracts around 457,000 viewers on the iPlayer. The CBBC digital channel programme, MI High, has a far higher proportion of viewership on the iPlayer: it has a TV audience of 145,000, while 30,000 watch it on the iPlayer. Huggers insisted that the online audience did not cannibalise the TV audience. The iPlayer is popular during office hours through the day but, as viewership peaks in the evening around 9 pm, heavy usage typically continues for an hour longer than TV viewing.

The BBC's user data shows that the iPlayer is used by a range of ages. 15- to 34-year-olds account for 37% of viewers and 35- to 54-year-olds account for 43%. A further 21% of users are aged 55 or over and Huggers credited the iPlayer's popularity to it being easy to use.

The priority is to make the iPlayer available on as many digital platforms as economically possible. PC users still account for the vast majority of iPlayer viewers with 85% of the audience, with Nintendo Wii and Linux both accounting for 1%. The popularity of the iPhone and iPod Touch had taken the BBC future media and technology team by surprise. Apple Mac users now account for 10% of iPlayer viewers, while iPhone and iPod Touch owners account for a further 3%.

Here are three conclusive quotes from Erik Huggers:

> *" The situations we're seeing are interesting – mum and dad are watching linear TV in the living room but the kids are watching in a different way ... on the iPhone, iPod Touch or a laptop. "*

> *" Having seen all this and understanding more about the success of the service, the sort of users, when they watch it and what they watch ... I think the BBC is absolutely betting on internet protocol in a way where it's not just for the distribution side of what the internet enables. "*

> *" We are completely re-engineering the way in which we make fantastic programming. "*

First published: 2008-Q4

---

4) *Guardian story:* www.guardian.co.uk/media/2008/nov/07/bbc-erikhuggers

Open source Handhelds

– a broadcaster-led innovation for BTH services

François Lefebvre (Project Leader), Jean-Michel Bouffard and Pascal Charest
Communications Research Centre, Canada

Emerging *Broadcasting to Handhelds* (BTH) technologies could be used to convey much more than the usual audio or video programming. For a long time now, broadcasters have imagined and standardized many new multimedia and data applications which, deplorably, did not succeed in the market.

In the first part of this article, we suggest that the *open source handhelds* which have become prominent as a consequence of recent technological trends, could also bring the emergence of broadcaster-led applications on mobile devices. In the second part, we will introduce the Openmokast project and describe how the CRC was able to produce, with very limited resources, the first open mobile-phone prototype, capable of receiving and presenting live broadcasting services.

There is a growing enthusiasm today for BTH services as presented in EBU Technical Review by Weck & Wilson [1]. These services can either use broadcast standards such as DAB/DMB and ATSC-M/H or the standards proposed by the mobile telecommunications industry such as DVB-H or MediaFLO. These technologies provide efficient delivery mechanisms for up-to-date information, popular media and valuable data services to mobile users. In the present interest for rich media convergence, BTH and mobile technologies have complementary features. BTH infrastructures promise to transmit large volumes of valuable content in one-to-many communications while wireless telecommunications networks could provide the channels for one-to-one exchanges.

This interest has led to the development of many new standards for BTH services in the context of DAB: MOT transport, Broadcast Web Site (BWS), SlideShow, DLS, TopNews, EPG and TPEG. But only a few of these standards have actually been implemented on commercial receivers. For example, BWS, which could be used to provide very attractive information services, is generally not supported on current receivers. In the remainder of this article, we will refer to those unsuccessful applications as the *missing apps*.

The stagnation of BTH technological advances could be explained by the innovative and competitive wireless communications ecosystem that is thriving today. Several kinds of new wireless communications technologies are emerging in the quest to reach mobile users wherever they are with a maximum throughput.

It appears as if BTH is standing at a juncture between broadcasters and mobile network operators (MNOs). Their business models are in conflict. Broadcasters are naturally inclined to pursue and extend their current free-to-air (FTA) services which are monetized by public funding, licence fees and advertising. MNOs, on the other hand, plan to deploy BTH services to generate new revenue streams through cable-like subscriptions and pay-per-view models.

The Internet is also challenging the traditional broadcasting model. It has lead to rapid innovation cycles that produce direct and disruptive benefits to end-users. For example, webcasting, peer-to-peer streaming and podcasting all represent new and attractive alternatives to broadcasting.

We suggest here that one of the key limiting factors for broadcaster-led innovation lies in the broadcasters' limited control over the implementation of their standards into receivers. The market for broadcast receivers is horizontal. Implementing broadcaster-led standards into mobile phones is an even bigger challenge because these devices are part of a vertical market. That is, MNOs have a firm control over which feature sets get implemented into "their" devices. Understandably, they will likely promote their own feature sets before those of broadcasters. As a consequence, we can expect that BTH innovations that are broadcaster-led are less likely to be implemented into mobile phones. Von Hippel's theory [2] would suggest that as empowered "users" of the mobile phone technology, MNOs stand in a much better position to innovate and compete.

The following section will introduce emerging trends that could create new opportunities for broadcaster-led innovations in the BTH space. Please see *Appendix A: "Anatomy of a Handheld"* for an overview of the components and terminology that will be referred to throughout the remainder of this article.

Note: Throughout this paper, the term handhelds is used to refer to generic pocket-sized computing devices which may provide connectivity to any kind of network. On the other hand, mobile phones are a specialized category of handhelds that connect to MNOs' infrastructures.

Trends towards broadcaster-led innovation

Specialized hardware goes generic

The manufacturing of handhelds is expensive. Mass-production is imperative to reach profitability. This results in steep barriers to entry for newcomers in the field. Fortunately, the relentless push of Moore's law has permitted the implementation of generic, compact and powerful integrated circuits. These components can be produced affordably and run on little power. This benefits smart phones too.

Inside today's mobile devices, specialized hardware components are increasingly being replaced by generic ones. Flexible application processors (APs) can be re-used in the design of many types of devices, thereby lowering the overall production costs associated with specific applications.

Functionality goes software

In early-generation mobile phones, user applications were performed by low-level software loaded directly on the device's permanent storage. Only basic tasks were supported: dialler, phonebook, device settings or ringtone selection. There was no room for new applications. Nowadays, generic and powerful APs coupled with large storage capacities have led to better capabilities. This has also significantly raised the importance of software in handhelds.

As a consequence, functionality features – based on software – present much lower barriers to entry than for hardware because of the fundamental properties of software:

- it can be duplicated at no cost;
- it can be distributed instantly and at no cost;
- it can be developed with low-cost or free tools;
- it can be modified, fixed and enhanced with no impact on the manufacturing chain.

Software goes open

The software industry is experiencing changes as open source software (OSS) is gradually entering the domain. With OSS, the global level of collaboration adds significant value to the software development chain. Individual programmers and organizations work at innovating and solving issues with a "let's-not-reinvent-the-wheel" approach. Instead, they build together a solid common core which sustains their respective business models.

In our opinion, the qualifying term "open source" is not in itself very significant. The important factor is to apply the proper rights to the code, once it is written. This is where software licensing comes into play. Projects are said to be free/libre Open Source Software (FLOSS) when their licensing terms offer flexibility in the usage rights while remaining accessible to the largest possible community. The Free Software Foundation classifies software licences and promotes those that are, in accordance with their criteria, genuinely open. The Open Source Initiative (OSI) is also a recognized reference body that leads the reviewing and approving of licences that conform to the Open Source Definition [3].

There is an obvious trend in motion now. Industry has adopted many open source software solutions. Several OSS products are widely deployed: GNU/Linux, Apache, OpenOffice, Firefox, Asterisk, Eclipse and MySQL. Companies using these tools see several benefits in the form of flexibility, independence, ability to fix, to enhance and to tweak the software that they need. For all of those involved, OSS is a key to success.

The popularity of OSS has even reached one of the technology-sector bastions. We can now purchase consumer electronics (CE) products that do "run" on OSS. Here are some important examples of such devices:

- The Linksys WRT54G Wi-Fi wireless router [4] is a product for which the GNU/Linux firmware was released. Since this release, users have enhanced its functionality to enterprise-grade routers.

- Neuros [5], an Internet set-top box manufacturer, will shortly release its next product called OSD2. Neuros relies partly on users and external developers to create or integrate new applications for their platforms which also use OSS.

- The DASH Express [6] is a new type of GPS car navigation system with a two-way telecommunication data channel. It can receive real-time traffic data information and can also offer Internet search for points of interest. The DASH device came on the market a year ago and seems to be very successful in the USA. Interestingly, DASH is based on the FIC Inc. Openmoko first hardware release. The DASH is a good example of a vertically-integrated device. It is a great case study of how OSS can be used with the right licensing scheme on closed business models.

- Some developers have also enabled OSS frameworks on top of closed CE devices. The Rockbox project [7] creates open source replacement firmware for several brands of portable digital audio players like Apple, Archos and iRiver. Audio codecs are amongst the numerous features added: FLAC, WavPack, AC3 (A/52), AAC/MP4 and WMA. These were not available on the iPod models sold by Apple.

Handhelds go open source

Today, there is an important push towards OSS on handhelds promoted by many industry players. Their interest to participate in the mobility value chain is motivated by the current trends described above. Some key projects with the potential to impact BTH are described in this section.

Openmoko

Openmoko [8] is an OSS project that was initiated by First International Computer, Inc. (FIC Inc.), an important manufacturer of motherboards for personal computers. Early on, the company made the bet that their best option to become competitive with their new smartphone products was to open up a complete software stack that would enable and leverage user innovation. In July 2007, FIC Inc. released its first "developer preview" prototype called the Neo 1973 with the Openmoko software stack.

To many developers, the Neo represents the first truly open mobile phone platform. The Neo incorporates several interesting connectivity options: GSM, GPRS, GPS and Bluetooth. Interestingly, the Neo 1973 initially shipped with a primitive software stack that did not even allow the completion of phone calls. We had to wait a few more months before the "phoning feature" became available through software updates.

The Neo FreeRunner (GTA02), was the next version of the device. It was released in June 2008 with enhanced usability, hardware improvements and Wi-Fi connectivity.

Openmoko was conceived to enable various business models. Openmoko Inc., the company, does sell devices for profit. Independent developers can sell proprietary software applications thanks to the LGPL licence which covers the Openmoko software stack. With the DASH Express, FIC Inc. has also hinted that vertically integrated devices could be constructed on Openmoko. With such a framework, both closed and open applications are on a level playing field for competition. In this environment, users are just "one click away" from one or the other type of applications. The end-user will decide which best fits his/her needs.

Another interesting fact is that shortly after the release of the FreeRunner, another developer community was able to successfully port the Qtopia OSS distribution onto the device. Qtopia had been developed some years before by Trolltech, a company acquired by Nokia in 2008. At the current time, Koolu, a Canadian company, announced that it will port Android (another OSS distribution introduced in the next section) to the FreeRunner by the end of November 2009. This shows how organic and efficient an OSS ecosystem can be. Just a few months after its introduction, the FreeRunner device can already host several new software platforms.

Android

Android [9] was originally conceived to be the fundamental building block of new mobile devices. It is a software distribution that includes an operating system, a middleware layer and some key applications. It is being developed by the Open Handset Alliance (OHA), a consortium of 34 members including Google, HTC, T-Mobile and other important players in the field.

The Android Java-based software development kit (SDK) was released in November 2007 to allow application development long before any Android device was even produced. Since then, an application development contest sponsored by Google was initiated to stimulate the development of new Android applications. A total of US$5 million was awarded to 50 submissions featuring the most innovative applications [10]. The first Android-based mobile phone (T-Mobile G1) was released in October 2008 in selected markets.

Initially, the degree of openness of Android was limited despite the fact that the SDK was available for free. The OHA recently decided to release the Android open source project [11] which will encompass most components of the Android platform. With the Android OS and virtual machine becoming open, new exciting development projects are now possible. This could even lead to the creation of new hardware components. Details to come about the licensing and the governance of the project will ultimately reveal to what extent Android will be open. But in its current configuration, Android still presents a compelling option for the design of open source handhelds.

Other open mobile platforms

Other open platforms which do not include mobile phone network interfaces are available to create new handhelds as well. These devices represent potential candidates for broadcast reception.

The Nokia Internet tablet computers based on Maemo [12] are such devices. The Maemo platform provides lots of functionalities and shows great potential for many useful usage scenarios. Maemo is a software platform that is based on OSS projects such as Debian GNU/Linux and GNOME.

The Ubuntu Mobile Edition is another example of a GNU/Linux based software stack that could fulfil the requirements for new handhelds. This effort was launched by Canonical Inc. to support the development of an OSS distribution for mobile internet devices.

Another potential major advance for OSS on mobile phones could come from Nokia. The company announced the creation of the Symbian Foundation and the release of its Symbian OS as an open source software [13]. This OS was designed for mobile devices and comes with libraries, user interface frameworks and reference implementations. Symbian, with a market share currently surpassing 50%, is still the platform deployed on most smartphones in the world.

The Openmokast project

None of the open software frameworks and devices encountered in our study did support digital broadcasting hardware. Since BTH is a main field of research for our group at the CRC, we were motivated to explore the possibility of integrating broadcast functionality into such an open device. When FIC Inc. announced in February 2007 the launch of their open mobile phone prototype using the Openmoko framework, we decided to initiate the **Openmokast** ("OPEN MObile broadKAST-ing") project in our lab.

The aim of the project was to integrate a DAB receiver in a fully-functional mobile phone. We would design, build and test, with a live DAB signal, a prototype capable of decoding typical DAB audio services as well as some of the missing apps. Based on our previous experience in the lab, we chose to work with GNU/Linux and other OSS packages. We have learned that using OSS accelerates the integration of a prototype by reusing common SBBs found in those packages. In order to build a final product, we had to find a DAB reception platform and integrate it into the prototype. Other major software components such as the receiver control unit, the bitstream demultiplexer and the decoder had to be developed from scratch. We even had to manufacture a physical extension to the original handset to be able to embed the small USB DAB receiver and its required antenna into our prototype.

The Openmokast software platform

CRC-DABRMS is a stable software platform developed previously at the CRC to control commercial computer-based DAB receivers. This original effort provided access to raw DAB bitstreams on typical personal computers. CRC-DABRMS can decode signalling information contained in the fast information channel and dispatch desired sub-channels to various types of outputs. This in turn permitted the demonstration and testing of new applications geared toward DAB but not yet standardized. This system had been implemented for Windows and GNU/Linux platforms.

The GNU/Linux version of CRC-DABRMS was ported to the Openmoko platform

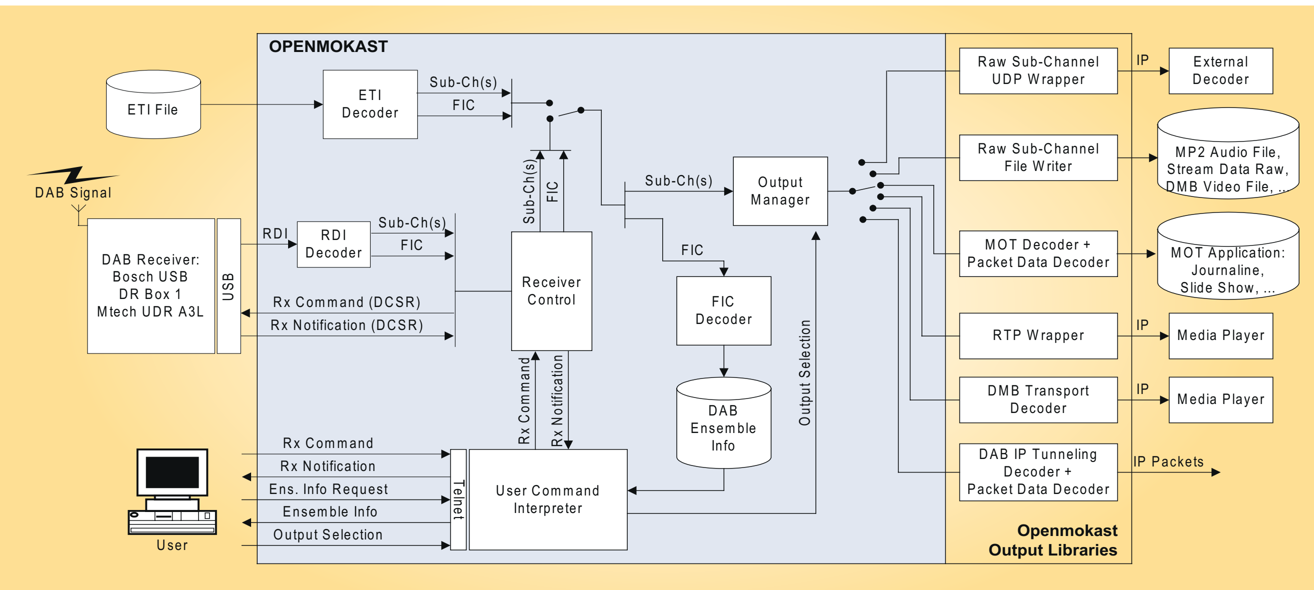
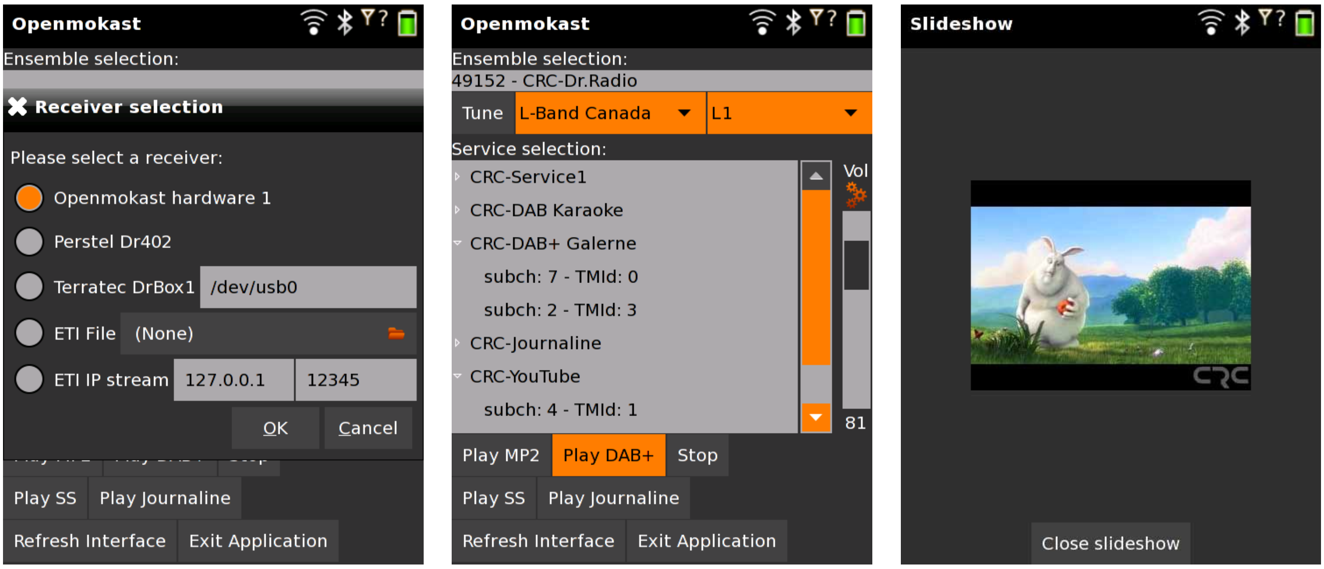


Figure 1
Architecture of the Openmokast middleware



(a) (b) (c)

Figure 2
Screenshots of a running Openmokast device

and renamed Openmokast. Porting it involved recompiling the application for the new target AP. It also meant adapting the code while verifying that all required libraries would be available on the new platform at runtime. The architecture of the Openmokast middleware is shown in *Fig. 1*.

The original interface of Openmokast was the command line. Later we developed a GUI using Openmoko's GTK libraries. Screenshots of a running Openmokast are shown in *Fig. 2*. At start-up, Openmokast presents a menu where the input device must be selected *(Fig. 2a)*. The system can also accept, as input, a locally-stored DAB multiplex file. This feature enables off-line testing and development of new applications without the need for a physical receiver and a live signal.

Different applications were either developed or were direct integrations of existing OSS projects. The standard DAB radio application was done with an HTTP wrapper which forwards MP2 audio to the Mplayer media player. A package for Mplayer was readily available. The DAB+ application was constructed in the same manner except that the transport protocol had to be removed prior to forwarding the AAC+ stream to Mplayer.

Two data applications were integrated by re-using source code made available under the OSS project called Dream [14]. Dream is an SDR receiver for DRM which includes an MOT transport protocol decoder, as well as two of the *missing apps*: Journaline and Slideshow *(Fig. 2c)*. The packet mode decoder had to be developed in-house. Since we were able to re-use code from other OSS projects, these applications could be developed quickly and efficiently. This experience reinforced our views that OSS carries significant benefits for developers.

The DAB receiver

A fundamental component required in the design of the prototype was a DAB receiver. Our early analysis suggested that a USB receiver could be a good fit for the unit. First, the FreeRunner has one USB port available. We also knew that the FreeRunner's AP was not powerful enough to perform DAB signal demodulation in real-time. We therefore searched for a USB component where this task could be performed efficiently on silicon.

We first tested the USB Terratec DrBox. The device drivers (DABUSB) for this device were already included in the Openmoko kernel. This should have made the integration task easier but unfortunately, the firmware for the unit could not be uploaded correctly. We had similar experiences with other USB receivers that did not succeed our initial test runs.

An alternative in our quest for the appropriate receiver was to obtain development kits offered by chipset makers. We contacted many companies but none could provide the components required. In general, their offerings did not meet the needs of a research development project like ours. Those companies were unwilling to support our initiative which would likely not lead to significant device sales. Furthermore, the development kits offered are expensive and the USB device drivers provided are usually built for Windows instead of GNU/Linux, and the usage of those kits requires the signature of non-disclosure agreements. We did not want to follow such a path of development and were stalled for a while.

We continued our research and found an off-the-shelf product which could fulfil our requirements. The MTECH USB key model UDR-A3L had all the features we needed. The libusb library was chosen to communicate with it. Libusb is a Unix suite of user-mode routines that control data transfer between USB devices and the host systems. We were able to establish basic communications between the MTECH key and the FreeRunner USB port. A reference to this key was added as a new input source in Openmokast (hardware 1 in *Fig. 2a*).

Prototype construction

The openness practices promoted by the Openmoko project went beyond our initial expectations. We found that even the CAD drawings for the FreeRunner mechanical casing were released to the public and available freely for download. Using those drawings for reference, we then modified the original mechanical design and produced a clip-on plastic extension which provides the extra space inside the device to embed the stripped down USB key. *Fig. 3* shows this extension.

To manufacture the physical extension part, we contracted the services of a 3D printing shop. We sent them the modified Pro/Engineer formatted 3D file and received the finished ABS part within 48 hours for the price of US$90. We were happy to find that just like with OSS, even mechanical hardware components of a prototype benefit from the "democratization" of production means. The CRC will release the schematics of this extension under a non-restrictive *creative commons* licence.

Fig. 4 shows the thickness of the final Openmokast prototype.

The MTECH receiver was connected directly onto the internal USB test points on the printed circuit board. In this configuration, it could draw its power from the device's main battery. This setup also had the advantage of freeing the external USB connector on the handset. This port remains the most convenient way to recharge the device. Once disassembled, the FreeRunner and extension exposed enough internal free space to install the dedicated L-Band antenna.

Figure 3
ABS extension for FreeRunner

Figure 4
Comparison of thickness between the original device and the Openmokast prototype

Figure 5
Inside the Openmokast prototype

Figure 6
The final Openmokast prototype

Fig. 5 shows the prototype's internals while *Fig. 6* shows the final product. Notice the colourful "skin" with our organization's brand as well as the Openmokast logo.

Some results

We are satisfied with the overall test results of the Openmokast prototype so far. This device was put together by a small team in a short period of time and it has performed well since its release. The form factor of the device is appreciated by current testers. Some DAB missing apps which are usually not available on commercial receivers could be demonstrated.

The Openmokast prototype was introduced at the IBC 2008 exhibition and was hailed as the first open mobile broadcasting handset.

The overall performance of the receiver component is good. The MTECH provided good signal reception in the L-Band and Band III frequency ranges, under various conditions. The device could receive signals coming from either the CRC-mmbTools LiveCD transmitter [15] or from standard commercial equipment.

The total CPU computational load measured for real-time DAB and DAB+ audio decoding on the FreeRunner was low. The GNU/Linux tops utility was used to estimate the processing cycles for two audio decoding scenarios (*Table 1*).

Table 2 shows the power autonomy estimates based on measurements made during three different usage scenarios.

The Openmokast software was installed and tested on typical GNU/Linux PCs.

	Codec type	Bitrate (kbit/s)	Mplayer	Openmokast	DAB+	Total
Scenario 1	Musicam	192	12.70%	1.00%		13.70%
Scenario 2	HE-AACv2	64	14.30%	0.60%	0.60%	15.50%

Table 1
CPU usage for two audio decoding scenarios (%)

	Source	Measured consumption	Estimated autonomy
Scenario 1	Receiver in Band III	650-670 mA	1h49
Scenario 2	ETI file	220-230 mA	5h20
Scenario 3	None	190 mA	6h19

Table 2
Openmokast power consumption and autonomy with 1200 mAh battery

Abbreviations

ABS Acrylonitrile Butadiene Styrene (thermoplastic)
AP Application Processor
ATSC Advanced Television Systems Committee
www.atsc.org
BTH Broadcasting To Handhelds
BWS (DAB) Broadcast Web Site
CAD Computer-Aided Design
CPU Central Processing Unit
DAB Digital Audio Broadcasting (Eureka-147)
www.worlddab.org
DLS (DAB) Dynamic Label Segment
DMB Digital Multimedia Broadcasting
www.t-dmb.org
DRM Digital Radio Mondiale
www.drm.org
DVB Digital Video Broadcasting
www.dvb.org
EPG Electronic Programme Guide
FLOSS Free/Libre Open Source Software
FTA Free-To-Air
GPRS General Packet Radio Service
GPS Global Positioning System
GSM Global System for Mobile communications
GTK Graphical user interface Tool Kit
GUI Graphical User Interface
HTTP HyperText Transfer Protocol
IP Intellectual Property
LGPL (GNU) Lesser General Public Licence
www.gnu.org/licenses/lgpl.html
MNO Mobile Network Operator
MOT (DAB) Multimedia Object Transfer
OHA Open Handset Alliance
www.openhandsetalliance.com
OS Operating System
OSI Open Source Initiative
www.opensource.org
OSS Open Source Software
SDK Software Development Kit
SDR Software Defined Radio
TPEG Transport Protocol Experts Group
www.tisa.org

The code could execute successfully on an emulator of the Neo 1973 available through the Openmoko project. In this setup, both Openmoko as well as the Openmokast GUI could be tested. We could not repeat a similar test for the FreeRunner configuration since no emulator for this device exists to date.

Openmokast's software could also be tested directly on two different GNU/Linux distributions, without the need for an Openmoko device emulator. We can conclude from these experiments that we could deploy Openmokast on other platforms and operate DAB devices in those environments.

Opening Openmokast

During this project, we had to face the issue of Intellectual Property (IP) in relation to OSS implementations. In fact, it is important to realize that implementing most of today's international open broadcast standards implies using proprietary IP. As a consequence, an implementation like Openmokast has to include proprietary algorithms or techniques to fully implement such a standard. Most of the time, licence fees must be paid to the rightful holder for each implementation of technology "sold".

Consequently, standards including proprietary IP appear to be a barrier to OSS implementations for two reasons. OSS projects are given away for free and do not generate revenues. It is also impossible to control the distribution of such software. Therefore, it would be very hard to collect any licence fees. Some countries have recognized the need for free standards and are promoting the adoption of new standards that are freely available and that can be implemented at no charge. The issue of open source and standardization was discussed in a report commissioned by ETSI in 2005 [16].

One possible solution for implementations of non-IP-free standards in OSS is to design frameworks with licensing schemes that allow the integration of IP-free as well as non-IP-free modules. With this approach, the non-IP-free components must be extracted from the overall distribution. A carved-up framework can be widely distributed along with the IP-free modules. The non-IP-Free components have to be distributed separately with provision for the correct attribution of the licensing fees. We plan to use this approach for the Openmokast project. The framework could be distributed openly without the module that requires special licensing (MP2 and HE-AACv2 decoding, etc..).

Conclusions

In this article, we have identified trends shaping the current environment that can impact the development of new BTH services. For a long time now, Moore's law has been the driver for advances in hardware and functionality that previously was provided by specialized circuits – but is now done on generic chips. The shift to software in the implementation of powerful handhelds is also a well-established trend. With software and its flexibility, developers have been able to create in a few months, thousands of new innovative applications for an Android or an iPhone.

We believe that a transformational force may be lying beyond the possibilities created by Moore's law and flexible software. The most important trend that we have identified during our study is that Open Source Software, once integrated on handhelds, carries an enormous potential for innovation. This is a revolution in the making that could reach its true potential once the right mix of collaboration and openness is found. Broadcasters, if they choose to follow this trend, could find new opportunities to promote their technologies. This could catapult broadcaster-led applications in the foreground and push further the deployment of new BTH networks.

In this article, we presented several open handset projects, similar consumer electronics devices and the Openmokast prototype developed at the CRC. We believe, following our study, that broadcasters have an opportunity now to sponsor the development of broadcaster-led handhelds. If they did, chipset manufacturers could be encouraged to participate in the process. In fact, the Openmokast framework could easily be adapted to support other technologies such as DVB-T or ATSC-M/H by exploiting current building blocks redundancy.

In an open device with both BTH and mobile telecommunications network interfaces, we could hope that some synergies would happen. It is a simple matter of providing software on the hand-held to bridge those two networks. With Openmokast, we can claim at last that we have reached convergence.

In an attempt to support our vision of convergence and the new opportunities that arise from it, the CRC plans to release, as open source software, several tools which were developed for the Openmokast project (www.openmokast.org).

Acknowledgements

The authors would like to thank their colleague Martin Quenneville for his work on the Openmokast prototype. Particular thanks should go to Karl Boutin for his thorough review and to André Carr, René Voyer, Bernard Caron, Silvia Barkany and Bruce Livesey for their comments and support.

References

[1] C. Weck and E. Wilson: **Broadcasting to Handhelds – an overview of systems and services**
EBU Technical Review No. 305, January 2006

[2] E. von Hippel: **Democratizing Innovation**
MIT Press, 2006

[3] www.opensource.org/docs/osd

[4] http://en.wikipedia.org/wiki/Wrt54g

[5] http://wiki.neurostechnology.com/index.php/OSD2.0_Development

[6] www.dash.net

[7] www.rockbox.org

[8] www.openmoko.com or

www.openmoko.org

[9] http://code.google.com/android

[10] http://code.google.com/android/adc_gallery

[11] http://source.android.com

[12] http://maemo.org

[13] www.symbianfoundation.org

[14] http://drm.sourceforge.net

[15] P. Charest, F. Lefebvre: **Software DAB Transmitter on Live CD** Published in the IASTED WOC 2008 conference proceedings, Québec, QC, May 2007

[16] Rannou & Soufron: **Open Source Impacts on ICT Standardization** Report – Analysis Part – VA6, ETSI

Appendix A: Anatomy of a handheld

Fig. 7 depicts typical hardware and software components on today's handhelds. The hardware building blocks (HBBs) operate around an application processor (AP). Given the portable nature of handhelds, special features support mobility: wireless connectivity, GPS and accelerometers are such examples.

Although APs can run most processing tasks, some of the "heavy lifting" has to be performed by specialized processors such as Digital Signal Processors (DSPs). The AP just does not have the processing power needed. Besides, even if it could manage the computational tasks, the related energy consumption requirement would be prohibitive. Instead, DSPs are used and these tasks can be efficiently accomplished while consuming less energy.

One such set of DSPs are the wireless Hardware Building Blocks (HBBs) found on the receiver presented in *Fig. 7*. As an example, the broadcast receiver front-end filters the signal then digitizes and translates it. The signal then emerges at

François Lefebvre joined the Broadcast Technologies research branch at the Communications Research Centre, Canada, in 1999 to lead its Mobile Multimedia Broadcasting team. Since then, he has contributed to numerous national and international standardization efforts and R&D projects. His recent work has focused on creating and developing open software building blocks for next-generation mobile broadcasting networks, devices and applications.

Mr Lefebvre graduated from Laval University in Electrical Engineering where he also completed his M.A.Sc. in 1989. He then moved to Europe where he worked for ten years as an engineer in R&D laboratories and as a freelancer on several multimedia and Internet projects in Germany.

Jean-Michel Bouffard graduated in Computer Engineering (B.A.Sc.) in 2003 from Sherbrooke University, Canada. He then joined the Broadcast Technologies research branch at the Communications Research Centre, in Ottawa, where he is involved in projects related to mobile multimedia broadcasting systems.

Mr. Bouffard's expertise in multimedia communications and his interest in convergence led to studies about the applications of the participatory Web concepts to broadcasting and mobility. In the same vein, his current occupation focuses on enabling and promoting broadcasting on open mobile devices. He is also completing his M.A.Sc in Systems and Computer Engineering at Carleton University.

Pascal Charest graduated in Computer Engineering (B.A.Sc.) in December 2000 from Laval University, Quebec City, Canada. He then joined the Broadcast Technologies research branch at the Communications Research Centre, in Ottawa, as a Research Engineer. He has been involved in several research projects in the area of mobile multimedia broadcasting.

Mr Charest's recent work has focused on system building blocks and open source software, with emphasis on encoding, decoding, modulation, device control and system integration. In addition, he is a member of the Club Linux Gatineau where he has served both as the Vice-President and President in the past.

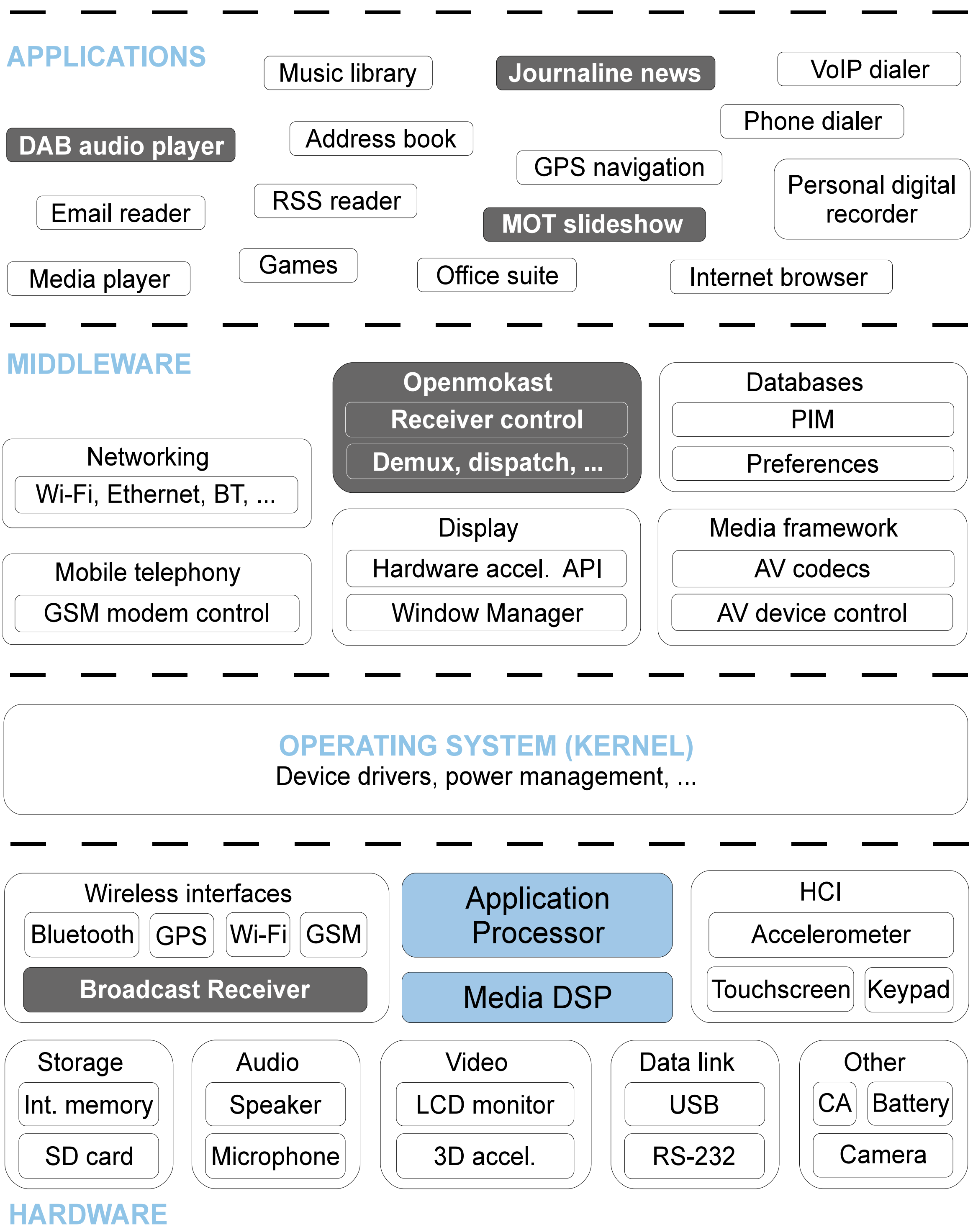


Figure 7
Typical hardware and software components of a contemporary handheld

an intermediate frequency or directly at baseband. At this point, a DSP performs the demodulation to produce a bitstream that can be processed by the AP.

Media processing normally requires some degree of powerful yet efficient processing and is performed using DSPs. Most current systems include specialized media processing units to decode media streams such as H.264 video and AAC audio.

Another typical hardware component that needs consideration in the design of a handheld is the conditional access HBB. This function usually relies on hardware to perform access management for pay services. Fortunately, since this is not a requirement for FTA services, the design of broadcaster-led handhelds is simplified.

The rapid evolution of APs makes us believe that in the foreseeable future, software defined radios (SDRs) will perform the demodulation of broadcast and other signals in real-time, with versatile wideband front-ends and A/D converters on the AP itself.

Three main levels of software are also depicted in *Fig. 7*: the operating system (OS), the middleware and the application layer. There is no clear separation between the layers and some software building blocks (SSBs) could actually overlap two or three of those layers. A complete im-plementation, including all SBBs required to operate a device, is often referred to as a software stack or a distribution. It provides the device with all of the basic software needed to operate.

In the software stack described above, the lower layer components provide their "Application Programming Interfaces" (APIs) to the upper layer components. Device drivers present the APIs of HBBs to upper software components.

First published: 2008-Q4

SVC – a highly-scalable version of H.264/AVC

Adi Kouadio
EBU Technical

Maryline Clare and Ludovic Noblet
Orange Labs, France Telecom R&D

Vincent Bottreau
Thomson Corporate Research

Scalable Video Coding (SVC) is a recent amendment to the ISO/ITU Advanced Video Coding (H.264/AVC) standard, which provides optional but efficient scalability functionalities on top of the high coding efficiency of H.264/AVC. In addition to bringing a cost-efficient solution to the delivery of different formats of the same content to multiple users, it can be used to provide a better viewing experience (enhanced content portability, device power / content-quality adaptation, fast zapping times and fluid forward / rewind functions, efficient error retransmission, etc.).

This article describes the potential of SVC, in terms of applications and performance. A brief overview of SVC functionalities, as well as practical use cases, are given in the following sections. Different performance evaluations, based on test results, are also described.

1. Introduction

Providing content everywhere is a major goal for video service providers. In addition to legacy broadcast TV, consumer video applications today span:

- IPTV (over managed networks and the open Internet);
- catch-up TV;
- Video-on-Demand services (VoD);
- Mobile TV;
- Web 2.0 content (including user-generated) and media platforms, etc.

All these new video applications are becoming a reality, thanks to developments in transmission, storage and compression technologies. Another enabler for this diversity of services is the strong penetration of end-user devices such as HDTV flat-panel displays, portable multimedia players (PMPs) and 3G mobile devices – and the availability of broadband Internet access ... providing high bandwidth connectivity into the home (xDSL, FTTH), within the home, between devices (Ethernet, Wi-Fi, Power-Line Communication) and outside of the home (3G, 4G, WiMax).

Nevertheless, providing content everywhere in such an environment, while achieving cost efficiency, is still a challenge for video service providers. Enabling such services implies the implementation of content-repurposing bricks in the service architecture, for transcoding between:

- *multiple image formats* (QCIF, CIF, QVGA, VGA, SD, HD);
- *multiple bitrates* (variable or constant

according to the access networks);
- *multiple frame rates* (50 Hz, 25 Hz, 12.5 Hz) and;
- *different delivery platforms (with different coding schemes)*.

Repurposing (transcoding) the content at any point in the chain generates extra cost, either for the service provider, the consumer or the network operator. It also alters the user experience by being an additional hurdle to content portability, which does not necessarily preserve the included DRM (Digital Rights Management). Other alternatives, such as content simulcasting, would result in higher bitrate requirements.

Fortunately, while many different video codecs were used in the past (depending on the targeted device for a given service), there is today a general trend towards H264/AVC [1]. This codec has not only been widely implemented in set-top boxes, it is also soon to be generalized in mobile devices as well as in Portable Media Players (PMPs); it has even been introduced recently in the Adobe Flash 9 and Apple QuickTime media players.

With H264/AVC generalization, transcoding will soon be limited to format and bitrate adaptation, and there will be less need for several output codecs. This is a key point in considering the SVC scalable extension of H264/AVC [2].

Considering the continuously increasing number of possible combinations between formats and bitrates, smart content adaptation today becomes a key issue for achieving the "content everywhere" target.

Because of all these considerations, scalability and flexibility are key points for the near future of video services, whether these are new services or the evolution of existing services. Such scalability is needed not only at the architecture and infrastructure levels, but also at the content level.

Scalable Video Coding provides the appropriate tools to efficiently implement content scalability and portability. It is the latest scalable video-coding solution, and has been standardized recently as an amendment to the now well-known and widespread H.264/AVC standard [1] by the *Joint Video Team*[1] (JVT). Other video scalability techniques have been proposed in the past, (even standardized as optional modes for MPEG-2 [3] and MPEG-4 - Part 2 [4]) but they were less efficient and more complex; moreover, because of the (then) lack of market need for scalability (video services limited to standard-definition broadcasts), they were never used.

In the following sections, we will first give a brief technical overview of SVC functionalities. The second part will outline different practical use-cases of the standard while the third part will describe preliminary performance evaluation results. The fourth and last part will describe ongoing work and the action taken by the EBU and other standardization bodies (MPEG, JVT) to extend the standard and provide more clarity on SVC performances.

2. SVC overview

Scalable coding consists of compressing a digital video into a single bitstream in such a way that other meaningful and consistent streams can be generated by discarding parts of the original compressed stream. Those sub-streams can be directly interpreted at different bitrates, different resolutions or different time scales.

SVC organizes the compressed file into a **base layer** (BL) that is H264/AVC encoded, and **enhancement layers** (EL) that bring additional information about quality, resolution or frame rate *(see Fig. 1)*. This implies that SVC base-layer streams can be decoded by H.264/AVC products (set-top boxes, PMPs), thus ensuring backward compatibility for consumers not having the SVC upgrade. More information about H264/AVC can be found in [1].

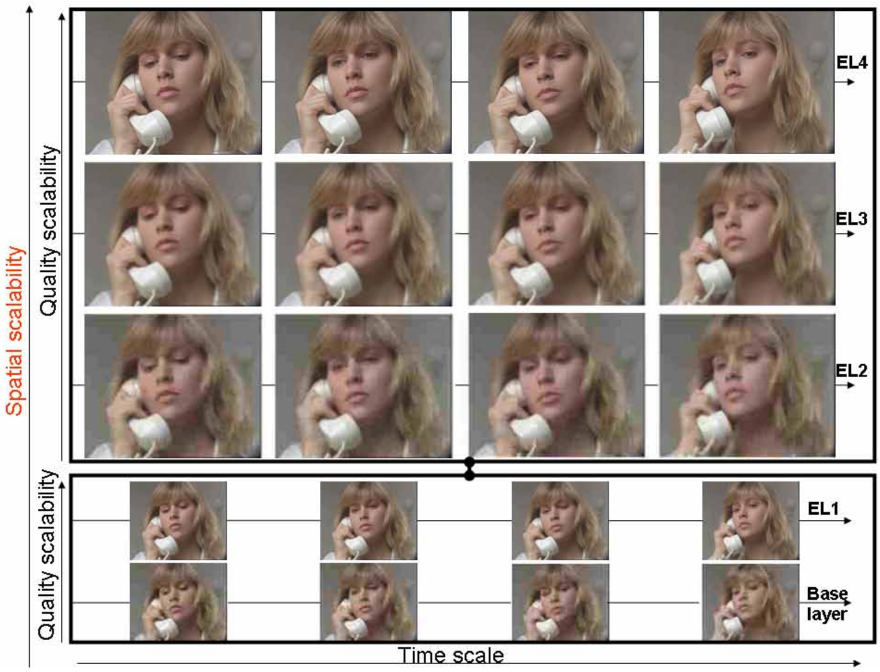


Figure 1
Overview of SVC layer structure (EL = enhancement layer)

[1] *The Joint Video Team is a joint working group from the ITU VCEG group and the ISO/IEC MPEG group at the origin of the H.264/AVC standard.*

SVC provides spatial, quality and temporal scalability types *(see Section 2.1)* that can be combined at each level. The enhancement layers can be fully hierarchical, or not:

- a layer can be generated ("predicted") from another layer and also be a prediction for yet another one;
- or a layer can be a prediction for two layers that are not hierarchically inter-dependent.

Such encoding parameters depend on the targeted application.

A compressed video bitstream is made up of Network Abstraction Layer (NAL) units, and each enhancement layer corresponds to a set of identified NAL units. A NAL unit is a packet with a header of a few bytes (containing information about the payload) and a payload corresponding to the compressed information. A set of successive NAL units, sharing the same properties, forms a NAL access unit.

Depending on the context, the enhancement layers may (or may not) be transmitted by the network, and may (or may not) be decoded by the end user device. In the first case, the network integrates some adaptation modules – deciding what to transmit and what to filter (for instance, depending on the network bandwidth characteristics). In the latter case, the terminal extracts the layers it can exploit. Such an adaptation mechanism is based on packet selection / dropping.

Setting up a service based on SVC technology implies two important considerations, *decision* and *adaptation*, which are further discussed in *Section 2.3*.

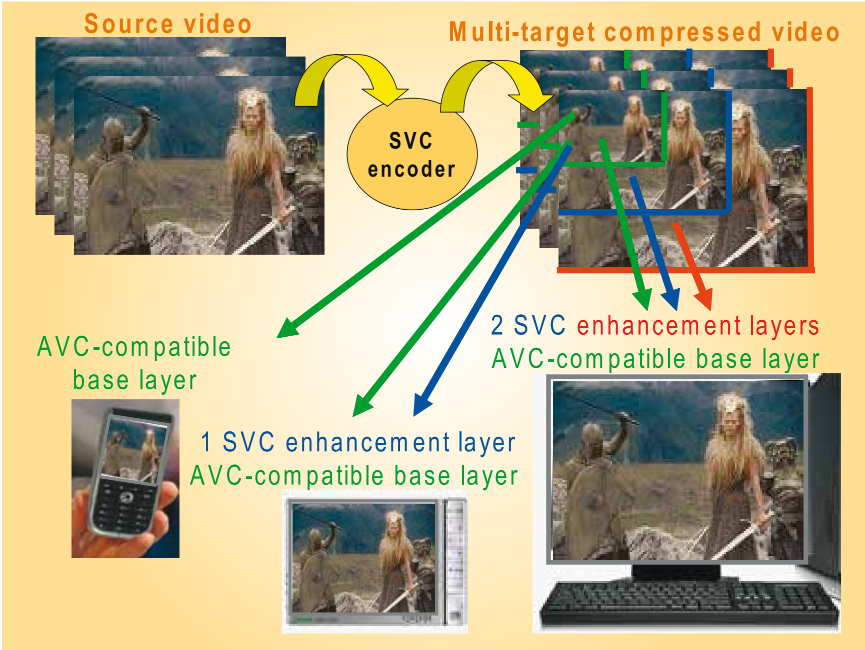


Figure 2
Spacial scalability

2.1. Scalability

2.1.1. Spatial scalability

Spatial resolution gives the video horizontal and vertical dimensions in pixels, resulting in several well-known "video formats" such as QCIF (176 x 144 pixels), CIF (352 x 288), SD (720 x 576) and HD (1280 x 720 up to 1920 x 1080).

The SVC standard's ability to embed 4:3 and 16:9 picture aspect ratios is, for example, a very important spatial scalability feature, typically when considering SD/HD broadcast. It should be noted that, depending on the standard profile in use, the ratio between layers can be fixed to a restricted set of values *(see Section 2.2)*.

Spatial scalability is provided by filtering / upsampling mechanisms and inter-layer predictions (Motion data prediction, intra-texture prediction and residual signal prediction). Each spatial enhancement layer (EL) is referred to as a *dependency representation*. A predicted EL always indicates the reference layer representation where it was originally predicted. EL macroblocks are predicted from reference layer macroblocks. They inherit motion vector values and other prediction data (texture and residual) from the appropriate reference layer macroblocks, after normative scaling and merging processes.

Spatial scalability can typically be used for transmission of the same video bitstream to PCs and portable devices *(see Fig. 2)*, or to SD and HD television sets.

2.1.2. Temporal scalability

Temporal scalability defines the difference in the number of images per second (expressed in Hz). Typical frequencies in Europe are 50 Hz, 25 Hz or 12.5 Hz.

SVC extends the tools already provided by H264/AVC (hierarchical P or B slices coding), structuring the bitstream into a hierarchy of images, thus allowing the easy removal of the lower level(s) in the hierarchical description.

Temporal scalability can typically be used in video transmissions over mobile networks where bandwidth capacity can change very often, or if the target terminal has very low CPU capacities: in such cases, it is interesting to drop the enhancement layers and send only the base layer (which could, for example, contain only half the number of images per second).

2.1.3. Quality scalability

Quality scalability is often referred to as SNR (Signal-to-Noise Ratio) scalability and is intended to give different levels of detail and fidelity to the original video, while having the same spatial and temporal definitions. In an SVC-compressed bitstream, each spatio-temporal layer can have different levels of quality – each of them bringing additional detail and accuracy.

It is up to the encoding process to decide whether more detail will be added to random parts of the video images, or to some specific parts of given images. Differences in quality levels can be medium (MGS, Medium Grain Scalability) or large (CGS, Coarse Grain Scalability). CGS provides a quality difference of about 25% between two layers while MGS offers 10%. MGS uses a modified high-level signalling, which allows a switching between different MGS layers in any access unit, and the so-called key picture concept, which allows the adjustment of a suitable trade-off between drift and enhancement layer coding efficiency for hierarchical prediction structures.

With the MGS concept, any enhancement layer NAL unit can be discarded from a quality-scalable bitstream, and thus packet-based quality-scalable coding is provided. The possibility of very fine granularity (FGS, Fine Grain Scalability), resulting in bitstreams that can be truncated anywhere, has been considered during the joint MPEG/ITU standardization process, but was not finally selected – as the finer the quality scalability is, the more complex the encoding/decoding process. Alternatively, the MGS concept allows the EL transform coefficients to be distributed among several slices. Thus, the information for a quality refinement picture that corresponds to a certain quantization step size can be distributed over several NAL units corresponding to different quality refinement layers.

Quality scalability can typically be used for:

- HD transmission to customers that are eligible for HD (full quality) and people not eligible for HD quality, but still equipped with HD screens (top enhancement layer is dropped);
- or for extra refinements when the bandwidth increases in mobile environments *(see Fig. 3)*.

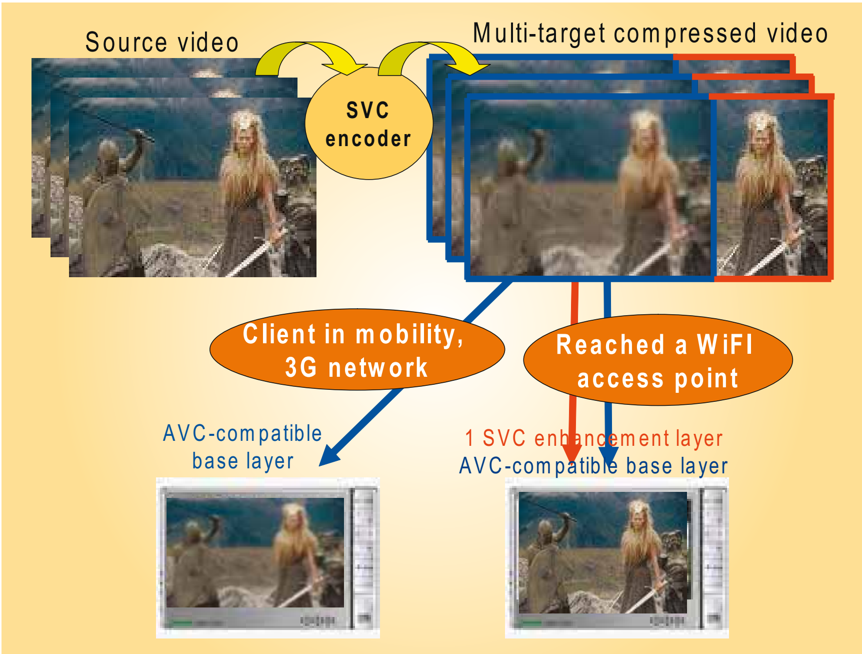


Figure 3
Quality scalability

2.1.4. Interlaced and progressive scalability

SVC inherits interlace tools from the H.264/AVC system: **Paff** (Picture adaptive frame/field) and **Mbaff** (Macroblock adaptive frame/field). SVC specifies four types of scanning-mode scalability – with some subjected to coding restrictions, depending on the interlaced coding mode of the base and enhancement layers:

- progressive to interlaced;
- interlaced to interlaced;
- interlaced to progressive;
- progressive to progressive.

More information on scanning-mode scalability can be found in [5].

2.2. Profiles and Levels

Profiles define the set of coding tools (for example, arithmetic or run length entropy coding, etc.) that can be used to build up the stream, while ***levels*** specify the constraints on key coding parameters, such as the number of macroblocks, the bitrate.

The SVC extension specifies three new scalable profiles, which are closely related to H.264/AVC profiles:

- Scalable Baseline Profile: aimed at low complexity applications;
- Scalable High Profile: aimed at broadcasting and video storage applications;
- Scalable High Intra Profile: aimed at professional applications.

The levels are the same as the H.264/AVC levels. However, the characteristic number of macro-blocks per second in an SVC stream is calculated according to the number of layers in the stream *(see formula below)*.

$$N = \begin{cases} \frac{1}{2} * \sum_{i=1}^{L-2} NL_i + NL_L, & L > 2 \\ NL_2 \quad , & L = 2 \end{cases} \quad \text{where } L > 0$$

with,
N = Total number of Macroblocks.
L = Total number of layers in the stream.
NL_i = Number of Macroblocks on layer i

Tools\Profiles	Scalable Baseline	Scalable High	Scalable High Intra
Base Layer profile	AVC baseline	AVC High	AVC High intra
CABAC and 8x8 transform	Only for certain levels	Yes	Yes
Layer spatial ratio	1:1 , 1.5:1 or 2:1	Unrestricted	Unrestricted
I, P and B slices	Limited B slices	All	Only I slices
Interlaced Tools (Mbaff & Paff)	No	Yes	Yes

Table 1
SVC profiles – main differences

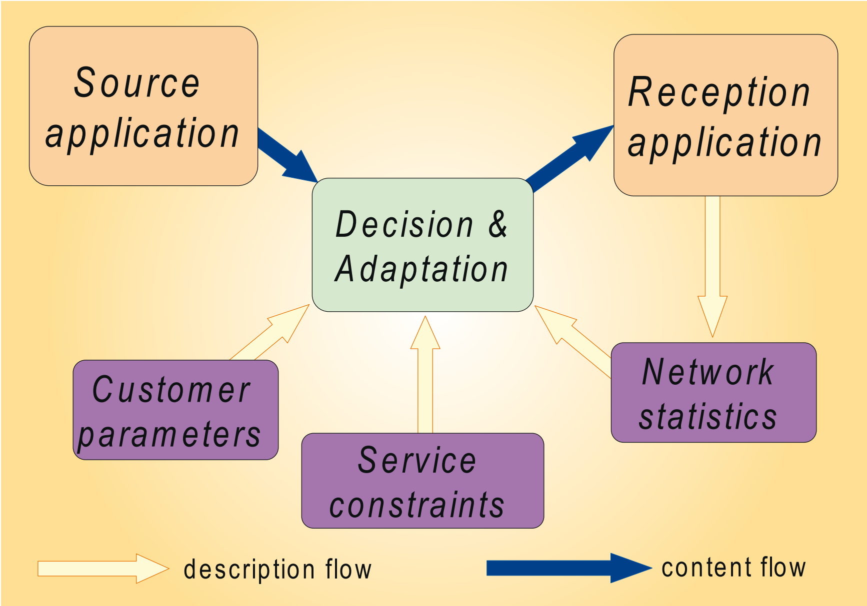


Figure 4
Bitstream adaptation

2.3. Overview of a service architecture with SVC

Once an SVC bitstream is generated, its scalability properties enable it to best match the transmission conditions over a network path to a given location and end-user device. One or more adaptation mechanisms can be implemented somewhere in the content delivery network – between the video streaming source and the end-user device. Such an adaptation process needs to obey a decision mechanism, based on the complete context at the time the service is offered (terminal properties, subscription characteristics, available bandwidth, error rate, DRM, etc.). Typically, in an IMS/TISPAN environment, it shall be noticed that previously-mentioned contextual information can be processed in very close relationship with not only access-control-related functions of IMS/TISPAN, but also with user-descriptive data.

Depending on the application, the decision and adaptation processes might not be implemented in the same equipment. Data flows in decision and adaptation mechanisms are illustrated in *Fig. 4*. Such a decision can be static (made once only, at the start of the transmission) or dynamic, and can be implemented in different parts of the service architecture, for example:

- At encoding – by splitting the video units into different bitrate levels, all assigned a different output stream, and sent to user groups having heterogeneous network / terminal capacities;
- At a VoD server – by adapting the video transmission bitrate in a unicast session, or by using scalability properties for providing a better trick mode user experience;
- At a network node – by reorganizing the video units to allow video streaming over a sub-network with different characteristics;
- At the edge – where a DSLAM can dynamically select video units (i.e. packets) for QoS, channel change, or just eligibility management;
- At the home gateway – for adaptively taking into account home networking conditions.

3. Envisaged use cases

As mentioned earlier, the current audiovisual landscape and its permanent evolution generate a strong need for scalability. We have already identified a few use cases where SVC would be of some benefit; now we present a few that are consensually considered as potentially providing much benefit.

3.1. Quality of Service and enhanced user experience

3.1.1. Fast zapping, fluid forward and rewind

The goal here is to improve the customer experience when using functions such as channel change or video fast forward / rewind. If the current video is displayed at a given bitrate and is decomposed on, at least, a base layer and an enhancement layer, then switching to the lowest layer allows us to either quickly visualize the next channel (TV channel switch) or have a more fluid fast forward mechanism (VoD function).

This is explained by the fact that less information is needed to describe lower-layer images, so the available bandwidth is used to transmit more, but smaller, images *(see Fig. 5)*. Of course, the new decoded stream (a new channel in the case of a switch, or further sequences of the video in the case of a fast forward) offers a lower quality until both layers can finally be sent ... lower resolution images are, in a first phase, received and enlarged artificially to fit to the screen dimensions. It is, however, demonstrated that the human visual system needs a couple of seconds before it is really sensitive to quality, so having a black image (channel switch) or a non-fluid forward is more annoying than quickly seeing information – even if this leads to a poor quality image for up to two seconds.

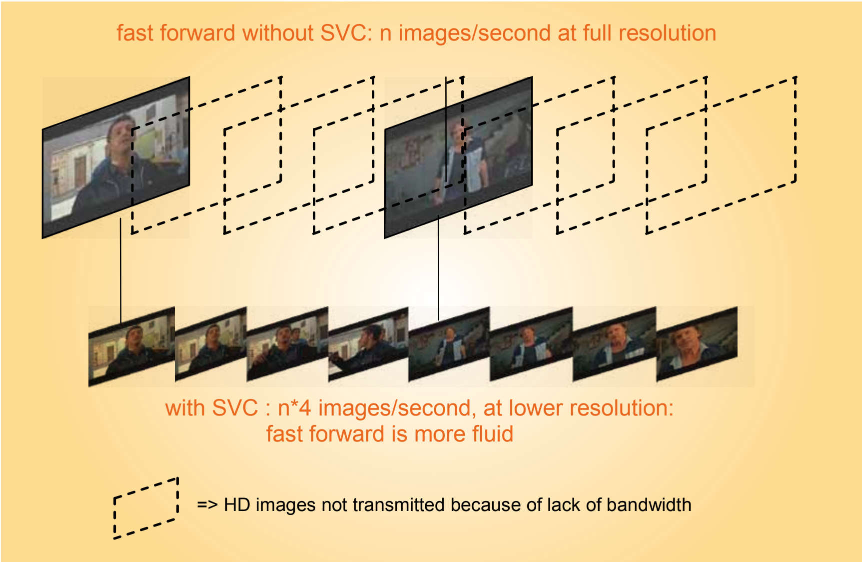


Figure 5
SVC for faster forward at lower resolution

Abbreviations

3GPP 3rd Generation Partnership Project
ATSC Advanced Television Systems Committee (USA)
AVC (MPEG-4) Advanced Video Coding, part 10 (aka H.264)
BL Base Layer
CABAC Context-Adaptive Binary Arithmetic Coding
CAS Conditional Access System
CDN Content Delivery Network
CGS Coarse Grain Scalability
CIF Common Interchange Format (352x288 pixels)
DLNA Digital Living Network Alliance www.dlna.org
DMB Digital Multimedia Broadcasting www.t-dmb.org
DRM Digital Rights Management
DSLAM Digital Subscriber Line Access Multiplexer
DSP Digital Signal Processor / Processing
DVB Digital Video Broadcasting www.dvb.org
EL Enhancement Layer
FGS Fine Grain Scalability
FTTH Fibre To The Home
GoP Group of Pictures
HDTV High-Definition Television
HHI Heinrich Hertz Institut (German R&D lab)
IC Integrated Circuit
IMS IP Multimedia Subsystem
IPTV Internet Protocol Televison
JSVM Joint Software Verification Model
JVT (MPEG/VCEG) Joint Video Team
LCD Liquid Crystal Display
Mbaff Macroblock adaptive frame/field
MGS Medium Grain Scalability
MOS Mean Opinion Score
NAB National Association of Broadcasters (USA) www.nab.org
NAL Network Abstraction Layer
Paff Picture adaptive frame/field
PLC Power-Line Communication, also written PLT, BPL ...
PLT Power-Line Transmission/Telecommunication, also written PLC, BPL ...
PMP Portable Multimedia Player
PoP Point Of Presence
PSNR Peak Signal-to-Noise Ratio
QCIF Quarter Common Intermediate Format (176x144 pixels)
RWTH Rheinisch-Westfälische Technische Hochschule
SAMVIQ Subjective Assessment Methodology for Video Quality
SDTV Standard-Definition Television
SNR Signal-to-Noise Ratio
SSIM Structural Similarity Metric
STB Set-Top Box
SVC (MPEG-4) Scalable Video Coding
SWOT Strengths, Weaknesses, Opportunities, Threats
TISPAN Telecoms & Internet converged Services and Protocols for Advanced Networks
WG-IPTV Walled Garden IPTV

3.1.2. Integration of retransmissions at constant bitrate

In the case of frequent transmission errors, different mechanisms are set up to correct them and improve the quality of service. One of these mechanisms consists of retransmitting those packets that never reached the terminal side. However, such mechanisms either require extra bandwidth or they slow down the transmission rate because of the information overhead.

An interesting feature of SVC is the ability to use the enhancement layer as a retransmission layer: in the case of errors, retransmission packets will be inserted in such layers, thus providing error correction at constant bitrate *(see Fig. 6)*. The only price to pay is to accept a loss of quality since less enhancement data is then received – some enhancement data being replaced by resent base-layer data. Such a mechanism guarantees reception of a minimum level of information quality, by providing maximal protection to this information.

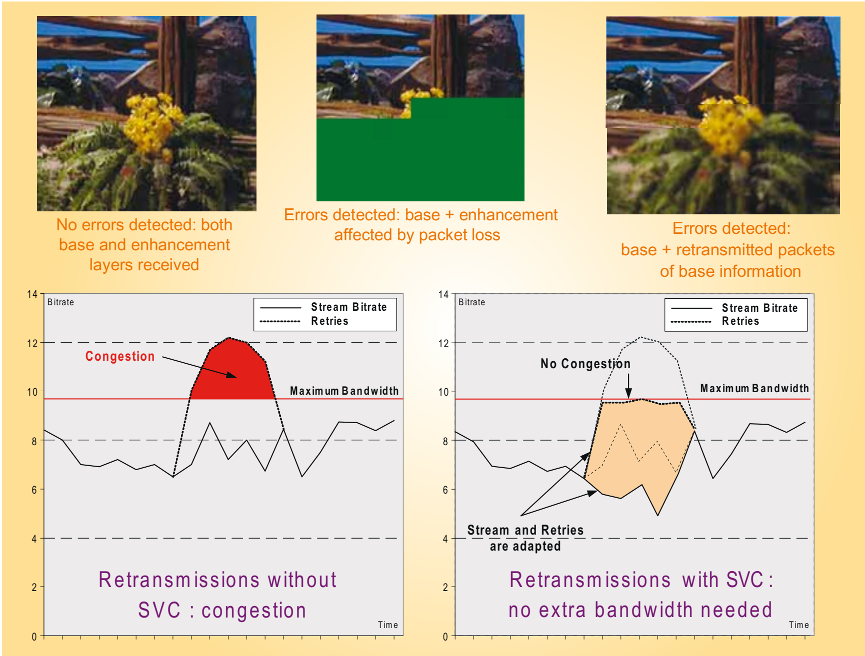


Figure 6
SVC for integrating retransmissions at constant bit-rate

Figure 7
SVC for maintaining service in mobility

3.2. Continuity of service in mobile environment

Mobile reception cannot rely on a stable bandwidth (it can drop when the network cell is overloaded or when the customer experiences a hand-over), but customers are highly annoyed by service ruptures. The advantage of having a video decomposed into layers is that they can simply be dropped and later reinserted, depending on the available bandwidth *(see Fig. 7)*.

Such capacity is not only used in the case of difficult transmission over a single network, it can also be used for the same video transmitted through different types of networks and towards different types of terminals: the same video is sent with a base layer only to mobile phones, but PCs connected via 3G cards to the same network can receive the full video quality (base + enhancement).

3.3. Portability for Video on Demand: Open Internet illustration

SVC offers an extremely simple way of using the "same video" source on different terminals. No operation such as transcoding is necessary, so there is no extra computing time and no loss of quality. This becomes extremely useful when you don't know in advance where and how you will finally use the video content you are interested in. SVC allows you to buy the content, start watching it on a given terminal and finish watching it on another. A very important consequence for content owners is the fact that DRM is preserved, which is an essential advantage when compared to transcoding.

An immediate illustration of this characteristic is the naturally-convergent video platform called the Internet ! You can access its portals from PCs, but also now (and soon even more) from mobile phones. As shown in *Fig. 8*, this results in an obvious need to watch given content on a platform that is not decided in advance (i.e. at downloading time).

3.4. Cost optimization for on-demand long tail management

As on-demand catalogues become larger and larger in terms of available titles / references, the costs for preparing and repurposing such content become higher and higher, even when using

Figure 8
SVC for easy content portability management

automated workflows. Moreover, this cost inflation is also related to the target devices and bitrates (access networks).

While there are known solutions to the cost model for the most popular content, the question of efficiently amortizing the preparation costs of the "long tail" (less popular content) still remains, even though the associated owners' rights are less important than for block-busters.

The above-mentioned costs correspond to the following tasks, needed for the general process of providing on-demand content:

- content capture (from tape, turn-around, file, etc.) and editing may require indexation means as well as metadata extraction / generation;
- content preparation (encoding and transcoding) and metadata processing;
- content integrity checking;
- content packaging (including protection);
- content provisioning;
- content ingest.

Content preparation and checking are important costs within this process. Multiplying the instances to be generated, in order to address multiple devices through multiple access networks, increases the overall on-demand processing costs, especially for the long-tail content – even with automated workflows.

We believe that using SVC instead of multiple H.264/AVC instances can reduce the capital and operational expenditures for the aforementioned content processing.

Moreover, we believe that SVC also allows more efficient schemes for content management within a content delivery network (CDN). Actually, it allows us to use access-network characteristics to manage the content distribution down to the edge. In that context, the combination of spatial (multiple devices) and SNR scalabilities (access-network capabilities) are probably the most anticipated features.

Finally, storage costs are lowered because the SVC global version is leaner than the cumulated AVC files of embedded versions.

3.5. Fine adaptation to xDSL residential eligibility

3.5.1. Optimal quality for a given eligibility level: illustration with HD screen without HD eligibility

SVC helps here to define intermediate ADSL eligibility levels, to which specific quality can be offered. This can be used, for instance, to provide good SD programme quality to customers who cannot reach HD bitrates, but who can still have a much better picture quality than the one provided by the SD eligibility threshold.

The programme can be SVC-encoded, with an H264/AVC SD base layer, plus a first quality and / or resolution enhancement layer – at a bitrate that is halfway between SD and HD bitrates – and a second enhancement layer for HD customers (see *Fig. 9*). This would help improve the bad quality noticed by customers viewing programmes on their HD screens if the quality of the video they receive is not really HD.

Such a feature can also apply to FTTH deployments. There will indeed be a period of time when FTTH is not available everywhere. SVC can provide a natural way to deliver the same content at different bitrates, qualities and resolutions, so when FTTH is present, an ultimate enhancement layer can be dedicated to the FTTH bandwidth capacities, thus allowing premium HD content to homes that can receive it.

3.5.2. Mutual dynamic access to bandwidth inside a home

SVC SNR scalability provides an efficient means for achieving a better simultaneous Walled Garden-IPTV (WG-IPTV) and Internet experience for the end user by adapting the Walled Garden IPTV video bitrate, depending on the bandwidth necessary for achieving the correct user experience on the Internet side within reasonable limits. Moreover, if the

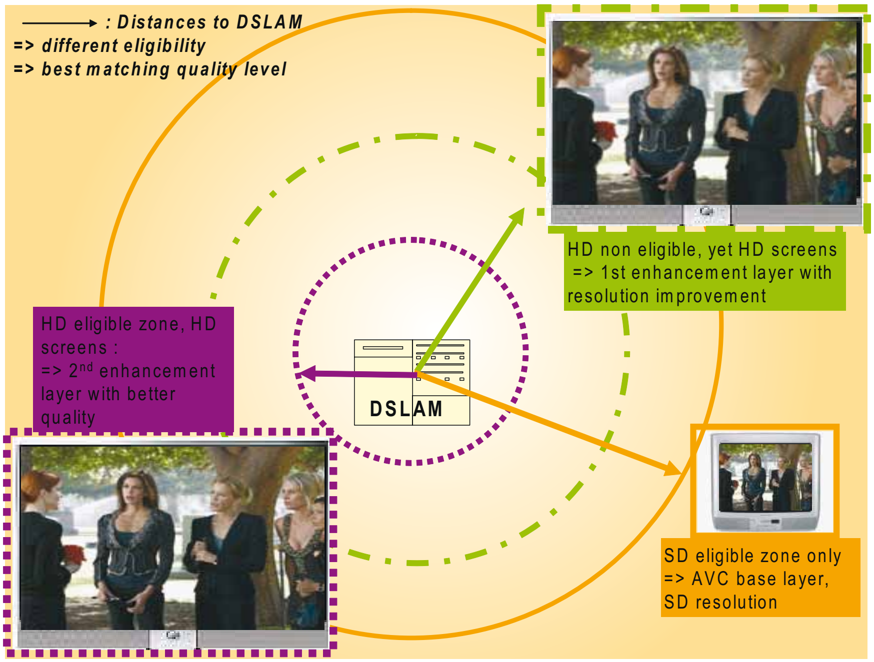


Figure 9
SVC for fine eligibility level management

internet video is encoded with SVC SNR scalability, it can also be adapted so that the impact over the WG-IPTV stream is not perceptible (including with fade-in/ fade-out mechanisms at the transitions). This adaptation can be either dynamically managed on the network side, or under full user control. With single-layer video technologies, the implementation of such dynamic mechanisms would require a transrating operation to be performed somewhere – and probably at the edge – which obviously does not seem to be interesting in terms of network architecture and density for massive processing and cost-efficiency.

An extension of this use case is bandwidth dynamic adaptation, for example if a second TV set (SDTV) is turned on when another (HDTV) is already being used. In this case, bandwidth can be shared between both programmes, providing optimal trade-off in quality to each, depending on their properties and on the target terminal characteristics.

3.5.3. Vector for efficient Premium HDTV Services enabling

The consumer adoption of H.264/AVC set-top boxes is still in progress. Bringing a completely new scalable system to market would imply a new set of products and interoperability issues with existing systems. Simulcasting several streams would require more bandwidth than using SVC on its own and would also require the user to tune in to the channel that provides the required quality/resolution.

SVC base layer compatibility with H.264/AVC would enable consumers who do not have SVC functionality to still see their usual programmes (SDTV or HDTV [1080i/25 or 720p/50]) if the latter is provided as the base layer of the SVC stream. SVC-compatible user will be able to access, for example, higher quality (1080p/50) signals stored in the enhancement layer as a premium service.

It is worth mentioning that the support of interlaced/progressive is very important for the broadcasting industry in order to enable a smooth (for consumers) and efficient (for broadcasters) transition from an SD to an HD (hopefully completely progressive) audiovisual landscape.

3.6. Other use cases

Even though they appeared to us as being of lower priority at the moment, other identified use cases should also be further investigated.

- Premium content incentive (“teasing”): show free content that is missing essential information (e.g. players in a soccer game) but introduce this information as soon as content is paid for.
- P2P streaming: introduce SVC as a tool to take advantage of information multiple-source distribution.
- User generated content: typical content accessed by heterogeneous terminals and through different networks, which could make the most of SVC.
- Provisioning and video preparation: how to analyse, index, pre-process, check and assign DRMs to a single video within its different layers.
- Video mail server optimization: allow access to only the lower version of the video when transmitted through

Definitions

720p/50	High-definition progressively-scanned TV format of 1280 x 720 pixels at 50 frames per second
1080i/25	High-definition interlaced TV format of 1920 x 1080 pixels at 25 frames per second, or 50 fields (half frames) every second
1080p/50	High-definition progressively-scanned TV format of 1920 x 1080 pixels at 50 frames per second

a mobile network, and full resolution when viewed via a residential access, or appropriate intermediate versions according to network and terminal conditions (SVC saves storage of intermediate versions).

- Handheld terminal battery optimization: switch to lower resolution if autonomy (battery life) is lower than a predefined threshold (and inform the customer that he or she still has a given number of minutes left for visualization at the current resolution, and a bigger number of minutes at a lower resolution).
- Heterogeneous terminals and access networks for videoconferences, e-learning and video surveillance: SVC allows us not to impose the lowest network and terminal capacities on the rest of the participants.
- Efficient signal monitoring on contribution links by decoding only a low resolution instance of the signal instead of decoding the full video stream.

4. Technology evaluation

The MPEG committee defined a "requirement" document for SVC prior to the actual standardization process, in which the most important requirements were:

- To provide a standard compatible with the state-of-the-art, i.e. H264/AVC. This point has been fully addressed since each SVC file base layer is H264/AVC encoded;
- To compress, in a single stream, different versions (e.g. different resolutions or different qualities) of the same video:
  - more efficiently than if the different versions are separately H264/AVC-encoded, then used together (Simulcast) – this point is addressed;
  - with 10% maximum of additional information compared with an H264/AVC encoded stream of maximum resolution/quality. This point depends on the use case scenarios,. Basically, the more levels included, the better gain compared to separate compressions, but the more diverse the spatial layers, and extra information is required.

Assessing a new technology means defining tests to evaluate its performance against alternative existing solutions in the same context. Defining SVC tests is not that easy because it is not a new compression standard that you can compare to another older one, but a new way to compress multiple representations of the same information. Next, once the target application is chosen, we need to define the embedded representation of the information, i.e. the different ways the video can be exploited: spatial enhancements (at which resolutions?), quality enhancements (up to which quality?), temporal enhancements (which frequencies?) ... or a mix of all these enhancements?

All of this implies that the comparison depends on the targeted application:

- For multicast service environments, we can compare an SVC stream transmission to the sum of the simulcasted information encoded with H264/AVC as illustrated in *Fig. 10*. Simultaneous transmission of the yellow-coloured streams (on the left of the diagram) has to be compared with transmission of the unique blue-coloured stream on the right – keeping in mind that it requires stream adaptation/extraction to be performed on it, somewhere between the stream production area and the end device.
- For VoD services, we can compare SVC streams with the sum of the encoded stored files (storage of the yellow-coloured streams compared to storage of the light blue-coloured stream, noting that the indexing is easier in the "multiple-versions-in-a-single-file" solution).
- For content portability, we can compare the SVC capability with the transcoding applied to a first encoded version of the video.

SVC performance evaluations on particular use cases have been conducted by different research labs, industries, and the JVT to quantify SVC performances both objectively (metric-based) and subjectively (visual quality). A summary of those test results is described in the following paragraphs.

4.1. Performance evaluations

4.1.1. Visual quality assessment test by Orange Labs

Orange decided to state the criteria that were essential for current audiovisual services and to evaluate difficult situations in order to have a low anchor and recognize that only better results can be expected in the near future.

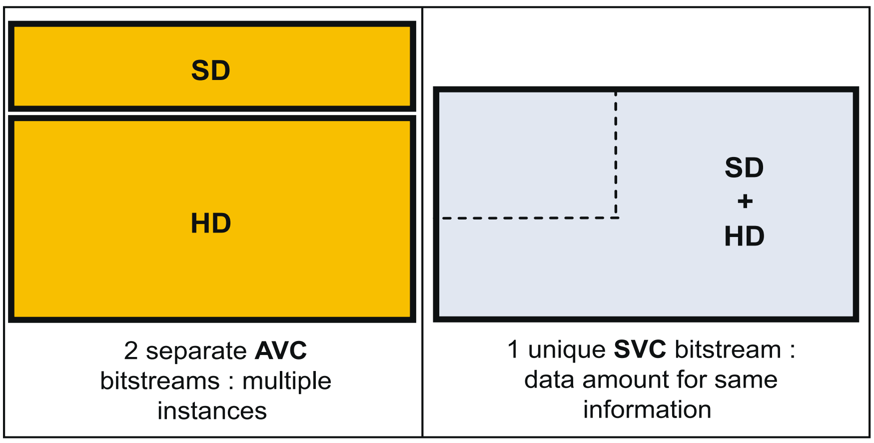


Figure 10
H264/AVC single layer vs. SVC single stream

Orange chose the "ADSL constant eligibility level" criteria as a mandatory point for introducing new technologies; in other words, Orange decided to evaluate what it would mean to provide SVC-encoded TV and video at the current service bitrates. Orange then had to compare a single layer H264/AVC encoded-decoded HD file *(bottom left yellow file of Fig. 10)* with the completely encoded-decoded HD SVC file containing an embedded SD base layer *(blue file of Fig. 10)*.

It is important to notice that, unlike evaluating the "classical" compression method where we compare the bitrate difference for a given fixed quality, here the quality loss at constant bitrate is being evaluated. Beware, it doesn't mean that SVC decreases the quality when compared to previous methods (otherwise, what is the point of considering it to replace previous methods!). It means that this is an evaluation of the side effects caused by scalability: the extra cost required for the "maximum" version (no attention being given to the fact that "minimum" versions are then intrinsically transmitted too since they are embedded in the file) compared to the way that such a version is encoded with other methods.

Tests have been performed at different bitrates: 12 Mbit/s, 10 Mbit/s, 8 Mbit/s, 6 Mbit/s and 4 Mbit/s.

For testing bitrate *n*, an H264/AVC file was encoded completely at n Mbit/s, and the SVC file had a base layer for SD always encoded at 2 Mbit/s, plus an enhancement layer encoded at *n*-2 Mbit/s.

More precisely, the tested scenarios were:

a) H264/AVC (720p/50) versus SVC (576i/25, 720p/50)
b) H264/AVC (1440x1080i/25 [2]) versus SVC (576i/25, 1440x1080i/25)

These figures are summarized in *Table 2*.

[2] *1440 samples per line are subsampled from a 1920 samples/line source signal.*

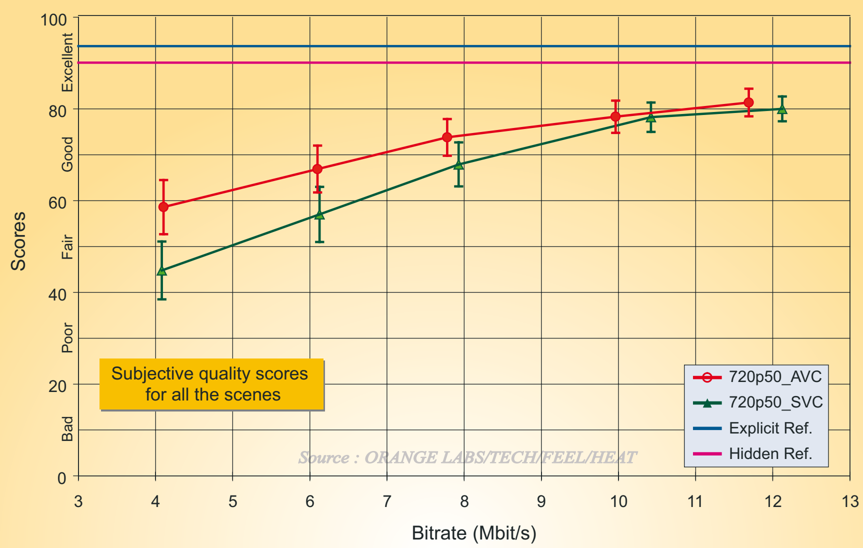
They used MPEG-ITU JSVM (Joint Software Verification Model) to generate both the H264/AVC and SVC files.

The visual quality assessment was performed on a 46-inch LCD display, following the SAMVIQ (Subjective Assessment Methodology for Video Quality) method [6][7].

Fig. 11 and *Fig. 12* show both sets of test results.

Several conclusions can be drawn from these tests:

- The SVC stream (with embedded SD and HD) provides similar to better visual quality than the single-layer HD AVC stream for bitrates greater than 7 Mbit/s. Transmitting both SD and HD as single-layer AVC would require a higher bitrate (the sum of the respective streams' bitrates) than the SVC stream.
For an HD AVC single-layer bitrate of less than 7 Mbit/s, SVC needs, for providing a similar visual quality, an extra bitrate corresponding to the SD AVC single-layer one.
- In critical cases, the loss between H264/AVC single representation and SVC maximum representation can reach up to 10 MOS (Mean Opinion Score) points, which is a noticeable difference.



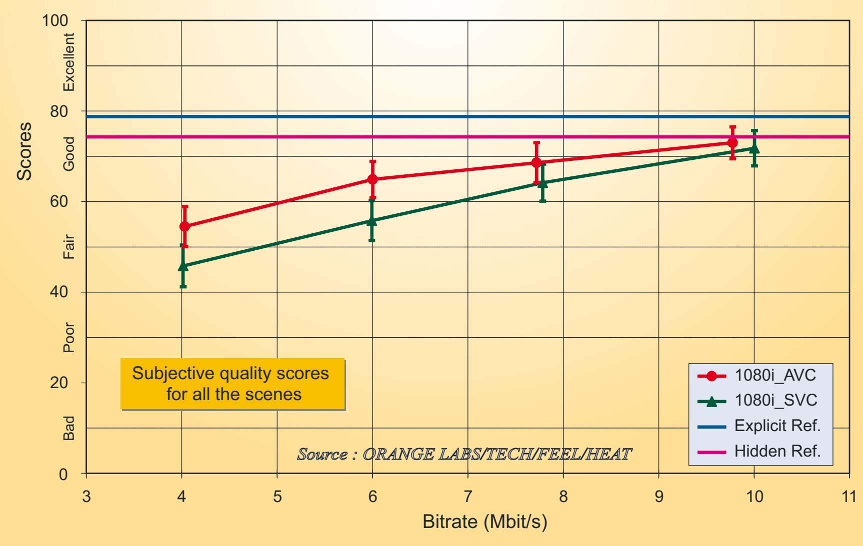


Figure 11 *(upper)*
Results for tests A

Figure 12 *(lower)*
Results for tests B

	H264/AVC bitrate (Mbit/s)	SVC bitrate (Mbit/s)		
		Base	Enhancement	Global
1280x720p/50	4	2	2	4
	6	2	4	6
	8	2	6	8
	10	2	8	10
	12	2	10	12
1440x1080i/25	4	2	2	4
	6	2	4	6
	8	2	6	8
	10	2	8	10

Table 2
SVC/AVC bitrate figures for visual comparisons

- In non-critical cases, the difference is always noticeable, but not penalizing, since the range of appreciation is generally maintained (e.g. "excellent", "good").
- SVC performs better if the base layer is of good quality since all enhancement predictions rely on it.
- If a sequence is originally interlaced, then SVC performs better with interlaced-to-interlaced scalability than when mixing interlaced with progressive layers.
- If a sequence is originally progressive, then SCV performs better with progressive-to-progressive scalability than when mixing interlaced with progressive layers.

Please note that all these results have been obtained with only the publicly-available versions of SVC software (MPEG/ITU JSVM). Needless to say, these versions are not as optimized as future in-dustrial implementations will surely be when available.

The tests were designed with the very precise goal of identifying the impact on quality when replacing H264/AVC with SVC at constant eligibility (i.e. constant bitrate, CBR) for current services. This constant bitrate affects the highest level, here HD resolution, since we impose that the SD base layer within the SVC file is encoded at today's SD eligibility level, i.e. 2 Mbit/s. This is indeed the way to ensure compatibility with existing services because the SD layer can simply be decoded by those having an H264/AVC currently-deployed decoder and no SVC decoder. So we simulate here a base that fulfils current TV service requirements.

4.1.2. Performance evaluation by Thomson's Corporate Research

Since January 2005, H.264/AVC-based HD broadcasting has been rolling out in the market with millions of receivers being shipped (DirecTV, Echostar, BSkyB,...). The HD formats are either 720p/50-60 or 1080i/25-30.

For premium services (e.g. Sports), broadcasters want to broadcast 1080p/50-60 while maintaining the existing customer base. Broadcasting of 1080p/50-60 SVC bitstreams enhances video resolution from 720p/50-60 or 1080i/25-30. Enhanced set-top boxes could be provided to premium customers, with additional 1080p/50-60 SVC (and H.264/AVC) capability – without the need to exchange recently shipped set-top boxes.

a) 720p/1080p versus 1080i/1080p comparison

Today, broadcasters are mainly using 1080i/25-30 for HDTV. Tomorrow, source capture will be more and more 1080p/50-60. SVC enables a potential migration towards 1080p/50-60 using a backward-compatible solution with H264/AVC (base layer). The main remaining question for the H.264/AVC base layer is either to encode it as 720p/50-60 or 1080i/25-30? Thomson Corporate Research has performed a first objective assessment using the SVC reference software and looking at rate distortion curves.

Test conditions

We have used the JSVM software for a 2-layer case:

- either 720p50/1080p50: progressive-to-progressive inter-layer prediction;
- or 1080i25/1080p50: interlace-to-progressive inter-layer prediction using coded field pictures.

The encoder settings were as follows:

- Scalable High profile;
- Hierarchical GoP size 4 (IbBbP coding structure);
- Intra period = 32;
- 100 first frames;
- H.264/AVC base layers encoded at fixed (constant) bitrates (CBR)
  - R0 = 4 Mbit/s;
  - R1 = 6 Mbit/s;
  - R2 = 8 Mbit/s;
  - R3 = 10 Mbit/s;
- For each rate, SVC enhancement layers were encoded using quantization parameter QpEL = QpBL + {0, 4, 8, 12}.

Results

Some results of these tests are given in *Figs 13* to *18* on the next two pages.

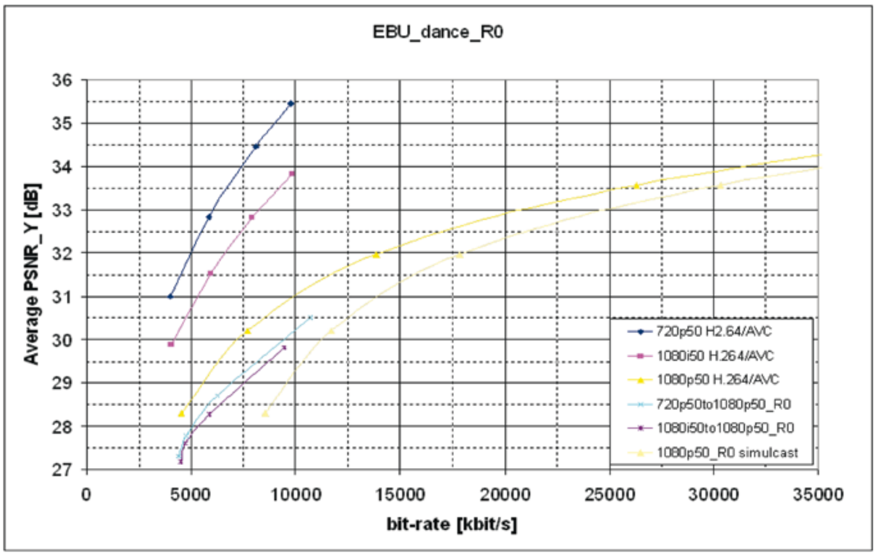


Figure 13
Results for sequence EBU_dance at rate R0

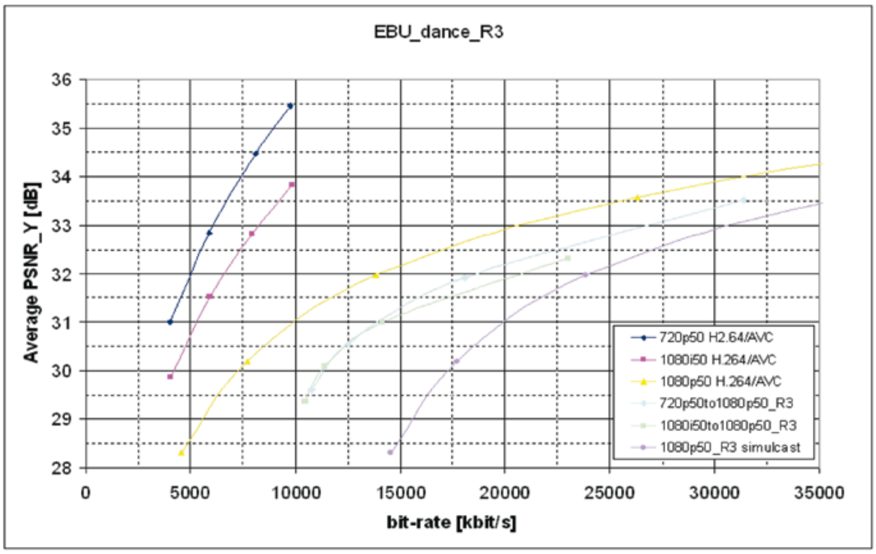


Figure 14
Results for sequence EBU_dance at rate R3

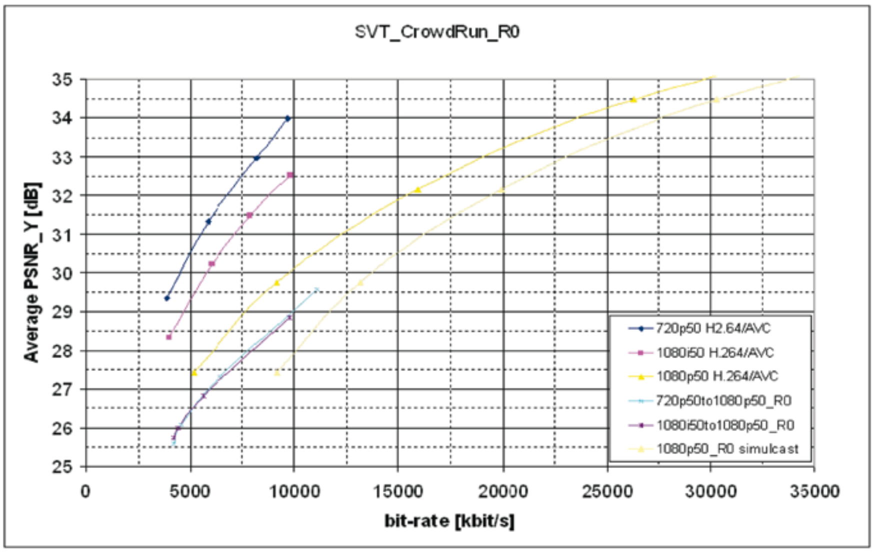


Figure 15
Results for sequence SVT_CrowdRun at rate R0

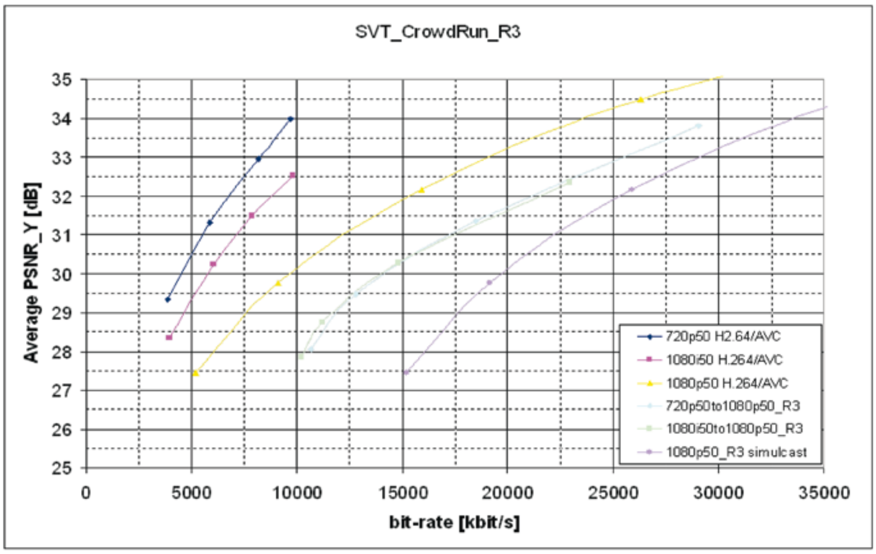


Figure 16
Results for sequence SVT_CrowdRun at rate R3

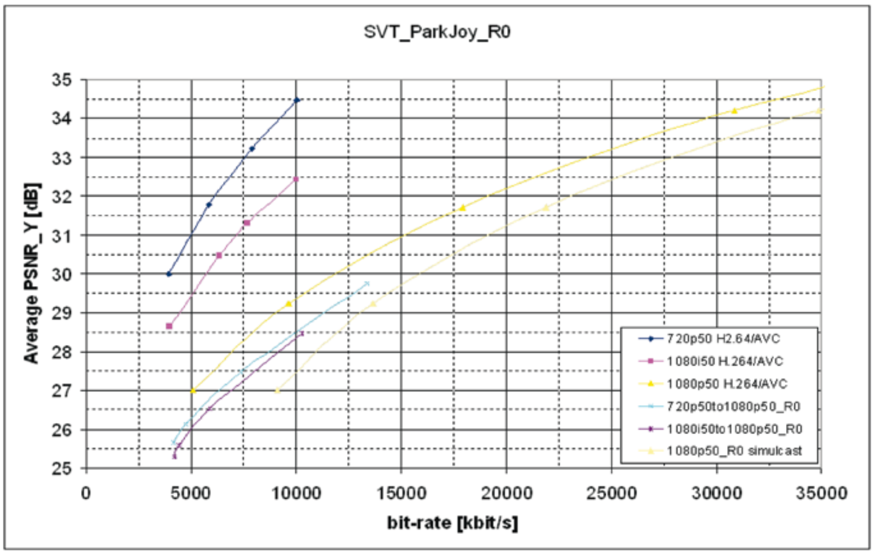


Figure 17
Results for sequence SVT_ParkJoy at rate R0

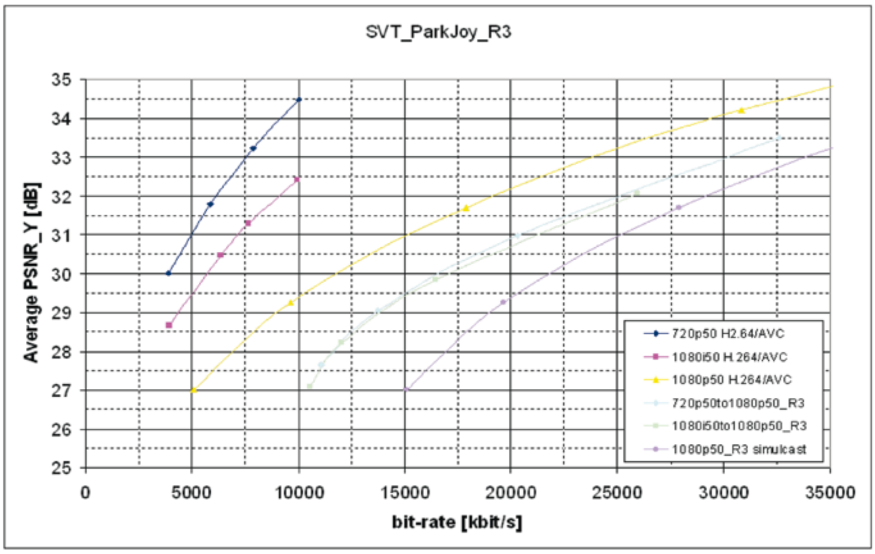


Figure 18
Results for sequence SVT ParkJoy at rate R3

Conclusions

Using an H.264/AVC 720p/50 or 1080i/25 base layer with a 1080p/50 enhancement layer provides approximately the same PSNR (Peak Signal to Noise Ratio) results, in terms of compression efficiency (a difference of less than 0.5 dB). Furthermore, it can be noted that the gain in bitrate compared to simulcast is shown and confirmed (~30%).

Knowing that PSNR doesn't always well correlate with visual perception, further subjective visual quality assessments should be conducted on a wider set of test sequences to assess the PSNR results.

4.1.3. Performance evaluation by JVT

Early this year, the JVT group released a report on SVC performance evaluation over a set of typical test cases [8]. The tests included:

- objective quality evaluation using the PSNR and SSIM (Structutral SIMilarity metric) methods;
- subjective quality evaluation using two assessment methods based on ITU-R Recommendation BT.500. [9]

It has to be noted that the different test scenarios only considered progressive-to-progressive image-format scalability, while discarding interlaced-to-progressive (SDI to 720p/50 or 1080p/50), or progressive-to-interlace (less probable), or even interlaced-to-interlaced (SDi to 1080i/25) scalability, which might be valuable for the broadcasting industry.

The overall conclusion of the test was that SVC will provide a 17 to 34% bitrate gain compared to simulcast, depending on the application. On the visual quality side, for critical sequences it needs up to 10% more bitrate to achieve similar or even better quality than the single-layer stream. For further information, please refer to the following document [8].

4.2. Implementation and complexity

At the moment, only the reference software implementation, JSVM, of the standard (now in version 9.12) is freely available. Research labs such as HHI or the industry are constantly developing optimized versions of the encoders.

Encoder-complexity evaluations still need to be done, even if it doesn't seem be an issue with the software version.

Additional decoder complexity in comparison with H.264/AVC is believed to be limited. Indeed, the SVC design specifies a single motion compensation loop (by imposing constrained intra prediction within reference layers). Thus, the overhead in decoder complexity for SVC compared to single-layer coding is smaller than that for prior video-coding standards, which all require multiple motion compensation loops at the decoder side. Additionally, each quality or spatial enhancement layer NAL unit can be parsed independently of the lower layer NAL units, which may further help in reducing the decoder complexity.

In order to keep track of the changes in software development and to always provide an up-to-date version of the JSVM software, a CVS server for the JSVM software has been set up at the *Rheinisch-Westfälische Technische Hochschule* (RWTH) in Aachen, Germany. The CVS server can be accessed using WinCVS or any other CVS client. The server is configured to allow read access only, using the parameters specified in *Table 3*. Write access to the JSVM software server is restricted to the JSVM software coordinators group.

authentication:	pserver
host address:	garcon.ient.rwth-aachen.de
path:	/cvs/jvt
user name:	jvtuser
password:	jvt.Amd.2
module name:	jsvm or jsvm_red

Table 3
CVS access parameters

5. Standardization information

As the audiovisual ecosystem has evolved in multiple dimensions (services, terminals, network access) and the services may also be rendered through retail devices, the need today for interoperability has never been so critical. Standardization bodies become interested in a technology when many service operators and industrial companies begin to take an interest in it. Thus, many standardization bodies and open forums have announced working groups on SVC:

- **DVB**: DVB is of considerable influence in the world of audio/video codecs, and the DVB H264/AVC groups have included SVC in their work programmes for mobile broadcast (DVB-H), 1080p and IPTV. The IPTV groups are now considering SVC in their content-delivery task forces.
- **3GPP**: SVC will be presented for its impact on architecture.
- **Open IPTV Forum**: since the first phase relies on currently available codecs and is near to being completed, a second phase will be started soon and we have offered to include SVC in working topics for both device codecs and IPTV infrastructure.
- **ITU Focus Group IPTV**: liaisons are active with DVB on the codec topics.
- **EBU**: the Delivery Management Committee (DMC) started a project group D/SVC in June 2008 to investigate the objective and subjective performance of SVC in broadcasting. The group is open to both EBU Members and the industry.

There is still a need to see if:

- **DLNA** (Digital Living Network Alliance, home network standardization activities) would be interested in including SVC as part of a work item.

Adi Kouadio obtained a B.Sc. and M.Sc. in communication systems at the Swiss federal institute of Technology (EPFL) in Lausanne, Switzerland. After working for Fastcom Technology SA on several projects relating to video-based object recognition, he joined the EBU Technical Department in 2007. Here, he is heavily involved in studies and evaluations of compression systems for various broadcasting applications (HDTV production, contribution and distribution).

On behalf of the EBU, Mr Kouadio liaises with the ISO/IEC JTC1/SG26/WG1 (JPEG) and ISO/IEC JTC1/SG26/WG11 (MPEG) working groups.

Maryline Clare graduated from INSA (*Institut National des Sciences Appliquées*, an engineering school in France) in 1989. She gained 10 years of still-image coding experience at Canon Research France and was heavily involved in the JPEG2000 standardization effort, during which time she was both "Head of the French Delegation" and the "Transform - Quantization - Entropy coding" group co-chair.

Ms Clare has continued working in the advanced video compression domain, first in Canon Research France then in Orange Labs, France Telecom R&D in Rennes. Starting with the AVC standard (Advanced Video Coding, MPEG-4 - Part 10, ITU H.264), she then closely followed its scalable extension SVC (Scalable Video Coding), on which she is currently leading a project in Orange Labs.

Ludovic Noblet received an Electronics and Computing Systems Dipl.-Ing. from the *École Polytechnique de l'Université de Nantes*, France in 1992. He started his career at Alcatel where he was in charge of introducing internet technologies within the Alcatel private network (Alcanet). He then moved to Thomson Corporate Research in October 1994 where he was involved in designing and contributing to the development and success of three successive generations of MPEG-2 encoders and one generation of MPEG-2 decoder. In 2002, he started working on the very first H.264/MPEG-4 AVC encoding implementations for Thomson's first generation of SDTV and HDTV AVC encoders.

In 2004, Mr Noblet joined France Telecom as an IPTV architect and senior technical advisor for the introduction of H.264/MPEG-4 AVC and high-definition within the Orange TV service. In September 2006, he was appointed head of the "Advanced Video Compression" team at Orange Labs.

Since December 2004, Mr Noblet has been the Orange representative at DVB, in both commercial and technical ad hoc groups for defining the use of A/V codecs in DVB applications. He also represents Orange at the Open IPTV Forum on the same topics.

Vincent Bottreau received the french state degree of Physics Engineer from the *École Nationale Supérieure de Physique Strasbourg* (ENSPS), France, in 1999. He also received the DEA (diploma validating the first year of a Ph.D. programme) in Photonic and Image Processing from the *Université Louis Pasteur*, Strasbourg, France, in 1999.

From 1999 to 2003, Mr Bottreau was with Philips Research in Suresnes, France where he worked as a Research Scientist in the Video and Communication group. In 2003 he has joined IRISA as a Research Expert in the TEMICS team. Since 2005 he has been an R&D engineer at Thomson R&D, France. His research activities include video coding, with a special focus on motion estimation, scalability and transcoding. He is in particularly involved in the ITU and MPEG standardization activities and, more specifically, on SVC.

We believe so, since today there is a considerable lack of standardization for retail devices. DLNA might be a good candidate since it deals with home networking matters.

- **TISPAN** would need to investigate if and how SVC may be of impact over the next generation network architecture it is defining, mostly in terms of flows between functional blocks, not only in the content traffic plane, but also in the control one.
- **ATSC** (US) and **DMB** (Korea) have announced doing some work on SVC, or have included requirements for scalability functionalities.

6. Conclusions

SVC seems to be of interest and very promising for current and future audiovisual services, especially in the area of:

- Video-on-Demand, to reduce the costs otherwise associated with the generation of multiple formats of the same video, especially for the long-tail management.
- Vector for efficient dissemination of premium HDTV services (1080p/50-60).
- Mobile video transmissions, to ensure a better continuity of service.
- Finer adaptation to home parameters such as device characteristics (HD screens) and available network access bandwidth.
- Quality of experience such as channel-switching time, fast forwards and rewinds.
- Dynamic bandwidth access management.
- Open Internet IPTV applications.

SVC is a good candidate technology for achieving the "content anytime, anywhere" target. However, further studies need to be conducted to further assess its efficiency. Among them we include:

- Other visual assessments to provide comparisons with alternative solutions to SVC (i.e. simulcasting, multiple description coding, transcoding).
- Reference test sequences and full source-quality video should be provided in all common media formats (CIF, QCIF, SDTV, HDTV, etc.).
- Visual assessments should be redone when optimized encoders are available because there is obviously room for improvement.
- Encoder complexity evaluation.

7. References

[1] Thomas Wiegand, Gary J. Sullivan, Gisle Bjontegaard and Ajay Luthra: **Overview of the H.264/AVC Video Coding Standard**
IEEE Transactions on Circuits and Systems for Video Technology, Vol. 13, No. 7, pp. 560-576, July 2003.

[2] Heiko Schwarz, Detlev Marpe and Thomas Wiegand: **Overview of the Scalable Video Coding Extension of the H.264/AVC Standard**
IEEE Transactions on Circuits and Systems for Video Technology, Special Issue on Scalable Video Coding, Vol. 17, No. 9, pp. 1103-1120, September 2007, Invited Paper.

[3] ITU-T Recommendation H.262 and ISO/IEC 13 818-2 (MPEG-2): **Generic Coding of Moving Pictures and Associated Audio Information – Part 2: Video**
ITU-T and ISO/IEC JTC 1, 1994.

[4] ISO/IEC 14 496-2 (MPEG-4 Visual Version 1): **Coding of audio-visual objects - Part 2: Visual**
ISO/IEC, Apr. 1999.

[5] Edouard Francois, Jérome Viéron and Vincent Bottreau: **Interlaced coding in SVC**
IEEE Transactions on Circuits and Systems for Video Technology, Vol. 17, No. 9, pp 1136-1148 September 2007.

[6] EBU BPN 056: **SAMVIQ – Subjective Assessment Methodology for Video Quality**
Report by EBU Project Group B/VIM (Video In Multimedia), May 2003.

[7] F. Kozamernik, V. Steinman, P. Sunna and E. Wyckens: **SAMVIQ – A New EBU Methodology for Video Quality Evaluations in Multimedia**
IBC 2004, Amsterdam, pp. 191 – 202.

[8] N9577: **SVC Verification Test Report**
ISO/IEC JTC 1/SC 29/WG 11, Antalya, TR – January 2007.

[9] ITU-R Recommendation BT.500-11: **Methodology for the Subjective Assessment of the Quality of Television Pictures**
Question ITU-R 211/11, Geneva, 2004.

[10] TD 531 R1 (PLEN/16), Draft new ITU-T Rec. H.264 (11/2007) | ISO/IEC 14496-10 (2008) Corrigendum 1: **Advanced video coding for generic audiovisual services: Corrections and clarifications**

First published: 2008-Q2

HDTV production codec tests

Massimo Visca (RAI) [1] **and Hans Hoffmann (EBU)**
EBU Project Group P/HDTP

To address the need for more efficient HDTV studio compression systems, vendors have recently introduced new HDTV studio codecs. In 2007, an EBU project group investigated these codecs and this article describes the methodology used for the tests and summarizes the results obtained.

The introduction of HDTV in Europe requires that broadcasters renew their production equipment. Whilst HDTV equipment in the past has targeted tape-based solutions, the user requirements for modern HDTV production workflows are file-based, non-linear and non-real-time – with shared access via networks and servers – and, last but not least, they have to be cost-effective. These requirements put challenges on the video compression format applied in mainstream HDTV production equipment – particularly on the trade-off between the data rate and video-quality headroom.

To meet these requirements, the industry has offered several codec solutions for mainstream HDTV production and the EBU decided to test these codecs in its P/HDTP (High Definition Television Production) project group. New codecs were provided for these tests by:

- AVID (DNxHD);
- GVG/Thomson (Infinity J2K);
- Panasonic (AVC-I);
- Sony (XDCAM HD422).

Existing legacy systems were also included in the evaluations in order to gather an understanding of the improvements achieved by new technology over legacy systems.

All tests on the new codecs were performed with each of the vendor's products individually (i.e. non-comparative tests) and the test plan and the results obtained were discussed between the EBU project group and the individual vendors.

The tests were carried out between spring and August 2007 and were conducted by several key EBU Members of the P/HDTP project. The AVID tests were performed by the IRT (Germany), the GVG/Thomson and Sony tests by the RAI (Italy), and the Panasonic tests by the EBU (Switzerland) with the assistance of RTVE (Spain). A representative of the vendor concerned was present at each test and at the subsequent expert viewing sessions.

1. Scope of the tests

The EBU codec tests focused on the image quality which the individual compression algorithm would provide after multi-generation processing. This codec performance is certainly only a part of the overall equipment performance of a recorder/server or camcorder device, camera, etc. But, in particular for HDTV with its inherent demand for high quality images, it is a very important parameter. The multi-generation codec assessment – by means of introducing pixel shifts after each generation – simulates how the production chain affects the images as a result of multi-compression and decompression stages.

For SDTV, the agreed method of multi-generation testing was to visually compare the 1st, 4th and 7th generations (including pixel shifts after each generation) with the original image under defined conditions (reference video monitor, particular viewing environment settings, and expert viewers). The same method was adapted to the current HDTV codec tests. In addition, an objective measurement – the PSNR – was calculated to give some

[1] *Massimo Visca acted as project coordinator and team leader for the codec tests.*

general trend indication of the multi-generation performance of the individual codecs.

2. Algorithms under test

The framework of the tests was aimed at investigating the performance of practically all the HDTV compression algorithms available on the market or under development in the manufacturers' laboratories. The test plan included both the so-called "Legacy" algorithms, applied in the most widely used systems since the start of HDTV production, and the "New" algorithms that were planned for market launch in the shorter term at the time of testing: some of these, at the time of writing this article, are now commercially available.

2.1. *"Legacy" algorithms*

The main features of the video compression algorithms employed in the "Legacy" equipment are summarized in *Table 1*.

2.2. *"New" algorithms*

The new HDTV systems – AVID (DNxHD), GVG/Thomson (Infinity J2K), Panasonic (AVC-I) and Sony (XDCAM HD422) in alphabetical order – employ a wide range of compression algorithms, differing both in terms of bitrate and in the mathematical tools used to perform the compression itself.

It should be noted that this article provides only the information about these algorithms that is necessary for a general understanding of their functioning. The following Tables provide some basic information (bitrate, bit depth, etc.) but, for a complete understanding, the reader should refer to the bibliography and to any official information provided by the manufacturers. Moreover, it is worth underlining that whilst these parameters provide some objective information about the system resources (e.g. the bitrate vs. storage capacity and network bandwidth), their correlation with the available picture quality is much more difficult, if not impossible, to determine from them. This is the reason for the large effort expended by the P/HDTP group in testing real implementations of the algorithms.

	Video Bitrate (Mbit/s)	Bit depth	Subsampling	Compression	Format	SMPTE standard
HDCAM	116.64	8	1440 Y 480 C_b/C_r	DCT based (Intra)	1080i/25 1080p/25	SMPTE 367M-368M
DVCPRO	100	8	1440 Y 480 C_b/C_r	DV based (Intra)	1080i/25 1080p/25	SMPTE 370M-371M
	100	8	960 Y 480 C_b/C_r	DV based	720p/50	
HDCAM-SR	≈440	10	NO	MPEG-4 SP (Intra)	1080i/25 1080p/25 720p/50	SMPTE 409-2005
HDCAM@35	35	8	1440 Y 720 C_b/C_r 4:2:0	MPEG-2 (GoP)	1080i/25 1080p/25	–

Table 1 – Video compression algorithms employed in the "Legacy" equipment

2.2.1. AVID (DNxHD)

Name	Video Bitrate (Mbit/s)	Bit depth	Subsampling	Compression	Format	SMPTE standard
DNxHD	120	8	NO	DCT based (Intra)	1080i/25 1080p/25	SMPTE VC-3
	115	8	NO	DCT based (Intra)	720p/50	SMPTE VC-3
DNxHD	185	8	NO	DCT based (Intra)	1080i/25 1080p/25	SMPTE VC-3
	175	8	NO	DCT based (Intra)	720p/50	SMPTE VC-3
DNxHD	185	10	NO	DCT based (Intra)	1080i/25 1080p/25	SMPTE VC-3
	175	10	NO	DCT based (Intra)	720p/50	SMPTE VC-3

Table 2 – AVID DNxHD video compression codec parameters

2.2.2. GVG/Thomson (Infinity J2K)

Name	Video Bitrate (Mbit/s)	Bit depth	Subsampling	Compression	Format	Compression standard
Infinity	50	10	NO	Wavelet based (Intra)	1080i/25 1080p/25 720p/50	JPEG2000
Infinity	75	10	NO	Wavelet based (Intra)	1080i/25 1080p/25 720p/50	JPEG2000
Infinity	100	10	NO	Wavelet based (Intra)	1080i/25 1080p/25 720p/50	JPEG2000

Table 3 – GVG/Thomson video compression codec parameters

2.2.3. Panasonic (AVC-I)

Name	Video Bitrate (Mbit/s)	Bit depth	Subsampling	Compression	Format	Compression standard
AVC-I	54.272	10	1440 Y 720 C_b/C_r 4:2:0	AVC (Intra)	1080i/25 1080p/25	High 10 Intra Profile
	54.067	10	960 Y 480 C_b/C_r 4:2:0	AVC (Intra)	720p/50	High 10 Intra Profile
AVC-I	111.820	10	NO	AVC (Intra)	1080i/25 1080p/25	High 4:2:2 Intra Profile
	111.616	10	NO	AVC (Intra)	720p/50	High 4:2:2 Intra Profile

Table 4 – Panasonic AVC-I video compression codec parameters

Name	Video Bitrate (Mbit/s)	Bit depth	Subsampling	Compression	Format	SMPTE standard
XDCAM HD50	50	8	NO	MPEG-2 GoP L=12 M=3	1080i/25 1080p/25 720p/50	–

Table 5 – Sony MPEG-2 Video Compression codec parameters

2.2.4. Sony (XDCAM HD422)

This algorithm employs a Long GoP (Group of Pictures) with L=12 and M=3, i.e. the GoP structure is IBBPBBPBBPBBI. This feature has some important implications on the testing of multi-generation performance, as described in detail in paragraph 3.2.3.

3. Methodology

In order to evaluate the performance of the different HDTV algorithms in a production environment, the very classical approach of the multi-generation (cascading) test was used. This method has been extensively used in all EBU tests since the introduction of compression algorithms in the SDTV production environment, starting from Digital Betacam up to the more recent IMX and DVCPRO systems. It is considered by the broadcast community to be a reliable methodology which is able to provide repeatable and stable results for easy interpretation.

The method is simply based on the performance of different compression-decompression steps using the algorithm under test, in order to simulate the cumulative effect of the artefacts introduced by the compression algorithm on the picture quality.

This method was originally devised to stress traditional tape-based equipment, where each copy implied a decompression and compression step. It could perhaps be considered as not fully reflecting the workflow of a future IT production-based infrastructure, where the necessity to perform a cascading of compression could be reduced. Nevertheless, considering that in the present production scenario, cascading is still a common process, the method allows the evaluation of the so-called "quality headroom" in the system. There is a good knowledge base in the evaluation of results using this method of assessment and it was agreed to continue to use it for analysis of the performance of compression algorithms implemented in the HDTV equipment under test.

3.1. Selection and shooting of test sequences

The selection of the test sequences to be used in any test is always a very critical issue; even ITU-R Rec. BT.500 – the most important reference document, containing all the procedures for picture quality evaluation based on subjective assessment – provides only a simple guideline, stating that the sequences have to be "critical, but not unduly so". Obviously, a "biased" selection of test sequences can be used to "drive" the test. The only way to solve this problem is to select a very large amount of material, in order to include sequences that cover all the possible features in terms of:

- high-frequency content (details), motion portrayal, colorimetric, contrast, etc.;
- indoor and outdoor shooting;
- different kinds of content, i.e. natural and artificial objects, texture, skin tones, etc.

Moreover, it is better if at least a subset of the test material is brand new, in order to avoid any kind of "optimization" of the new compression algorithm using familiar sequences.

As such a library was not available at the time of testing, it was necessary to shoot brand new sequences. The shooting was performed using state-of-the-art technology (Canon FJS Series prime lens, Sony HDC 1500 camera and uncompressed storage on a DVS server via HD-SDI). The result was a large portfolio of test sequences, each of 10s duration, satisfying the above-mentioned criteria.

All the sequences were shot in different formats – i.e. 1080p/50, 1080i/25, 1080p/25 (with and without shutter) and 720p/50 – making the best effort in order to guarantee the same conditions of lighting and exposure.

Note: The 1080p/50 sequences were down converted to obtain 1080i/25 or 720p/50 versions; they were therefore not directly used in these tests but they will provide a library that is available for future needs, such as the comparative test of the present 1.5 Gbit/s scenario with the future 3 Gbit/s.

Some sequences were singled out from material originally shot in one single HDTV format only. Other sequences were converted from film, or from rendered graphics or even taken from the archive in PAL format. All these sequences were converted into the three HDTV formats included in the tests to obtain an even larger portfolio of 10s-long test sequences.

Note: All the technical details (equipment, software, etc.) relevant to the conversion process are available from the authors, upon request.

For the 1080i/25 and 720p/50 formats, fourteen 10s-long test sequences were concatenated one after the other (no gray or black frames included) to form a single clip.

For the 1080p/25 (shutter on) formats, only eight 10s long test sequences were then concatenated to form a single clip. The total amount of test material employed, its different origins and the criteria of selection, guaranteed the completeness and formal correctness of the tests. Some frames extracted from the sequences are shown in *Fig. 1*.

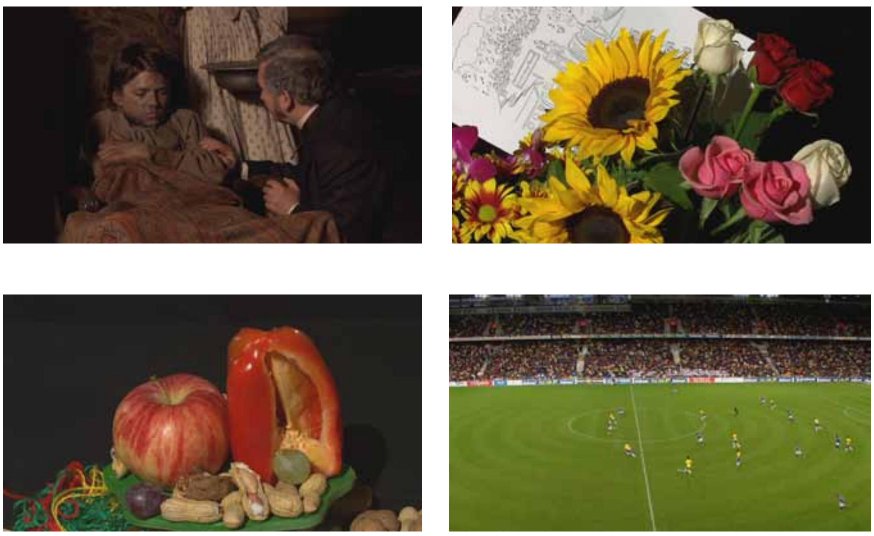


Figure 1
Still shots captured from four of the test sequences used

3.2. Standalone chain

The "standalone chain" comprised a cascading of several compression and decompression processes of the same algorithm; each pair of compression and decompression processes is usually referred to as a single "generation". In order to simulate different production scenarios and to investigate different features of the algorithm under test, the chain may or may not include processing between each generation, as explained below. As already mentioned, each algorithm under test was subjected to a multi-generation process up to the seventh generation.

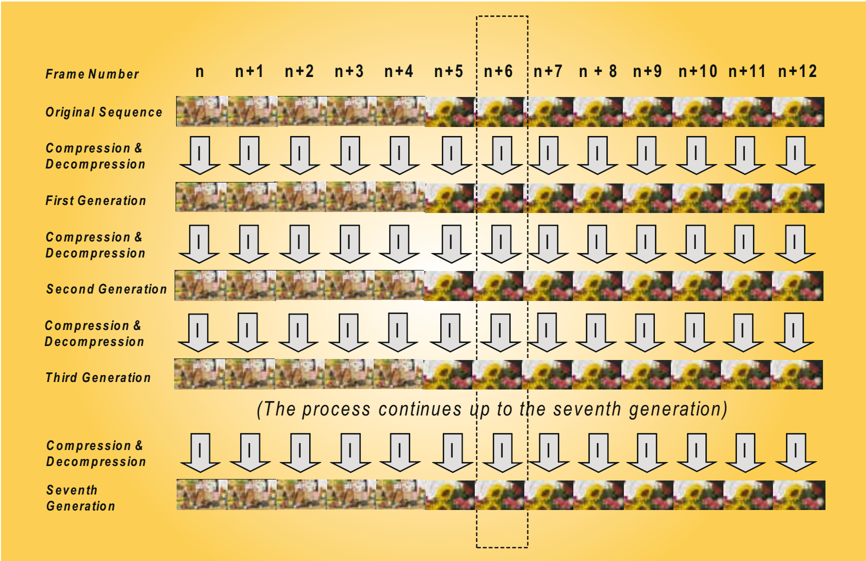


Figure 2
Standalone chain (without spatial shift) for INTRA codec

3.2.1. Standalone chain without processing

The standalone chain without processing simply consisted of several cascaded generations of the codec under test, without any other modifications to the picture content apart from those applied by the codec under test, as summarized in *Fig. 2*.

This process accurately simulates the effect of a simple dubbing of the sequence and is usually not very challenging for the compression algorithm. In fact, the most important impact on the picture quality should be incurred at the first generation, when the encoder has to eliminate some information, but the effect of the subsequent generations should be minimal as the encoder should basically eliminate the same information already deleted in the first compression step. Nevertheless, this simple chain can provide useful information about the performance of the sub-sampling filtering that is applied, or about the precision of the mathematical implementation of the code.

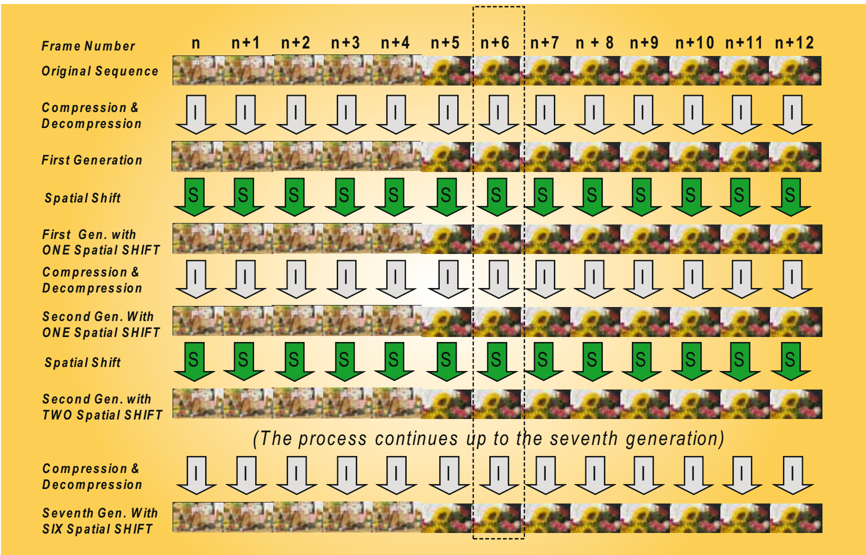


Figure 3
Standalone chain (with spatial shift) for INTRA codec

3.2.2. Standalone chain with processing

In a real production chain, several manipulations are applied to the picture to produce the master, such as editing, zoom, NLE and colour correction. Therefore, a realistic simulation has to take into account this issue. As all these processes are currently feasible only in the uncompressed domain, the effect of the processing is simulated by spatially shifting the image horizontally (pixel) or vertically (lines) in between each compression step, as summarized in *Fig. 3*.

Obviously, this shift makes the task of the coder more challenging, especially for those algorithms based on a division of the picture into blocks (e.g. NxN DCT block), as in any later generation the content of each block is different to that in the previous generation.

The shifts were applied variously using software or hardware, but the method used was exactly the same for all the algorithms under test. The shifts are summarized in *Table 6* and the process is summarized in *Fig. 4*.

For the horizontal shift (H), a "positive" shift means a shift towards the right, "negative" towards the left. Only even shifts are performed to take into account the chroma subsampling of the 4:2:2 format.

Shift between	Spatial Shift
First and second generation	+4H and +4V
Second and third generation	+2V
Third and fourth generation	–2H
Fourth and fifth generation	–2H
Fifth and sixth generation	–2V
Sixth and seventh generation	–4V

Table 6 – Spatial (vertical and horizontal) applied between each generation

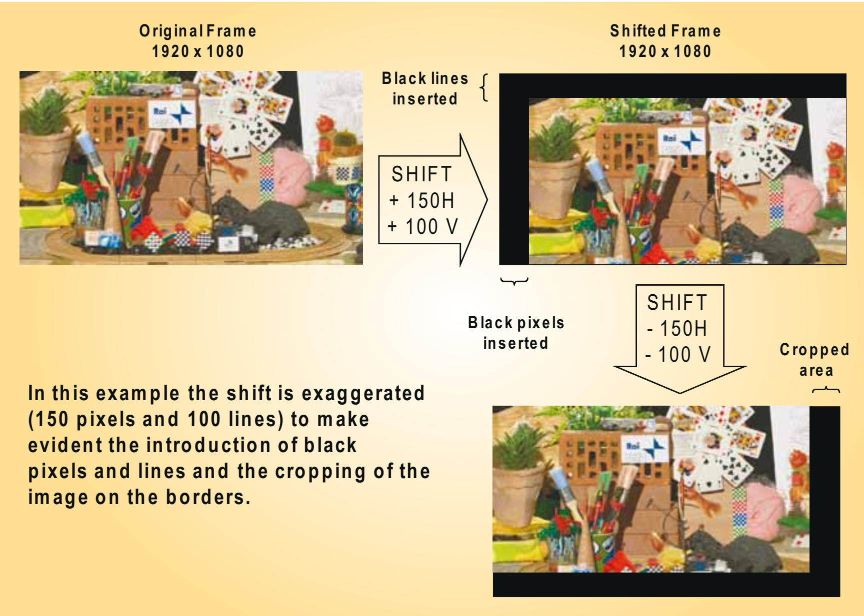


Figure 4
The shift process in the standalone chain

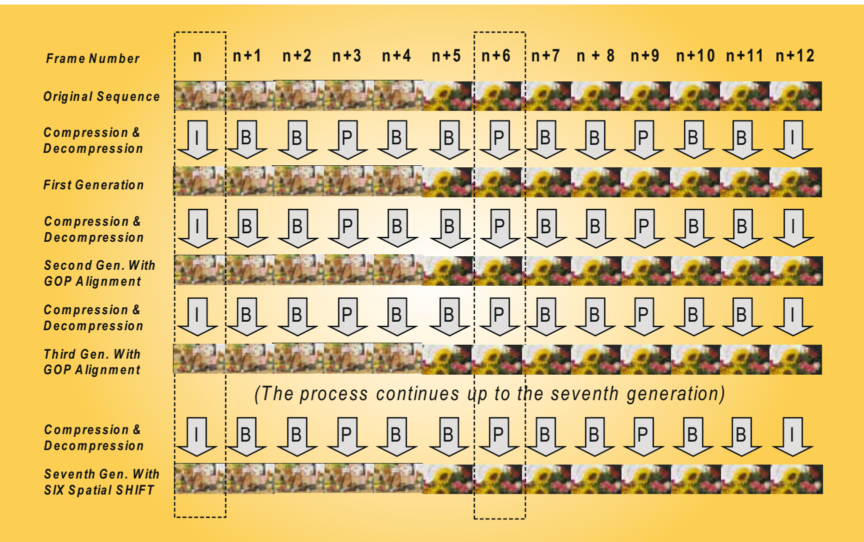


Figure 5
Standalone chain with GoP alignment (without spatial shift) for INTER codec

For the vertical shift (V), a “positive” shift means a down shift, “negative” an up shift. The shift is applied on a frame basis and is always an even value. For progressive formats, the whole frame is shifted by a number of lines corresponding to the vertical shift applied, while for interlaced formats each field is shifted by a number of lines corresponding to half the shift applied. For example, a shift equal to +2V means two lines down for progressive formats and 1 line down for each field of an interlaced format.

The shift process introduces black pixels on the edges of the frame if/when necessary.

3.2.3. Standalone chain for GoP-based algorithms

As shown in *Table 5* (on page ??), the XDCAM HD50 system exploits the MPEG-2 motion compensation tools and, in particular, Long GoP coding with L=12 and M=3 (i.e. IBBPBBPBBPBBI); even if each MPEG-2 encoder applies a rather sophisticated strategy to allocate its bitrate resources on the different kinds of pictures, all the MPEG-2 algorithms usually guarantee the best quality for the Intra picture (I), a reduced quality for the Predicted picture (P) and, in the same manner, an even lower quality for the Bidirectional picture (B).

Therefore, the GoP structure has some important implications on the way the standalone chain has to be realized, and introduces a further variable in the way the multi-generation can be performed – depending on whether the GoP alignment is guaranteed between each generation (GoP aligned) or not (GoP mis-aligned).

As explained in *Fig. 5*, the GoP is considered to be *aligned* if one frame of the original picture that is encoded at the first generation using one of the three possible kinds of frame – Intra, Predicted or Bidirectional – is again encoded using that same kind of frame in all the following generations: for example, if frame *n* of the original sequence is always encoded as Intra and frame *n*+6 as Predicted. It is therefore possible to have only one multi-generation chain with "GoP alignment".

On the contrary, if this condition is not guaranteed, several conditions of GoP mis-alignment are possible; considering the GoP length L=12, for the second generation 11 different GoP mis-alignments are possible, then for the third generation 11 by 11 and so on, making the testing of all the possible conditions unrealistic. It was therefore agreed to apply one "temporal shift" equal to one frame between each generation, as explained in *Fig. 6*, so that the frame that is encoded in Intra mode in the first generation is encoded in Bidirectional mode in the second generation and, in general, in a different mode for each following generation. It is interesting to underline that the alignment of the GoP in the different generations was under control (not random) and that this was considered the likely worst case as far as the mis-alignment effect is concerned, and was referred to in the documents as "chain without GoP alignment".

Taking into account also the necessity to simulate the effect of manipulation by means of the spatial shift, it was agreed for the GoP-based system (XDCAM HD50) to consider and to realize four different possible standalone chains up to the seventh generation, as follows:

- Multigeneration chain with GoP alignment (without spatial shift) *(see Fig. 5)*
- Multigeneration chain without GoP alignment (without spatial shift) *(see Fig. 6)*
- Multigeneration chain with GoP alignment AND spatial shift (*see Fig. 7*)
- Multigeneration chain without GoP alignment AND spatial shift (*see Fig. 8*)

All the abovementioned chains were carried out on the "Legacy" systems and on the "New" sys-tems. The resultant sequences were all stored in YUV 10-bit format.

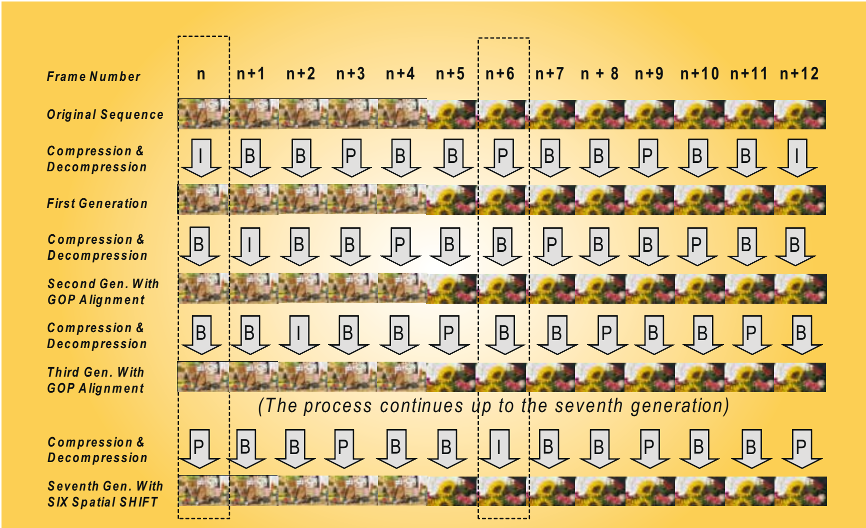


Figure 6
Standalone chain without GOP alignment (without spatial shift) for INTER codec

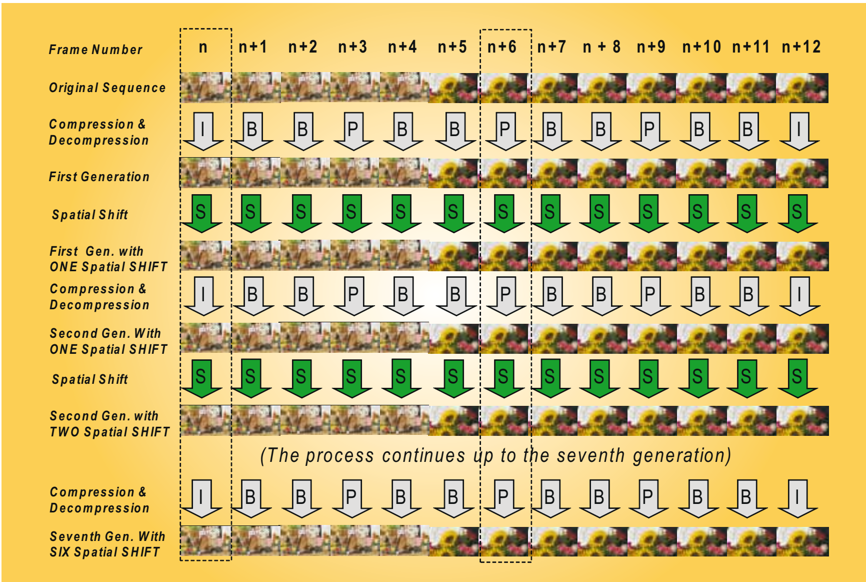


Figure 7
Standalone chain with GoP alignment AND with spatial shift for INTER codec

Note: For the XDCAM @35 chain it was not possible to obtain the GoP alignment control and therefore this multi-generation has to be considered "random" in terms of GoP alignment)

The tests on the AVID DNxHD codec were performed by the IRT, the tests on the GVG/Thomson codec were performed by the RAI using an Infinity prototype camera, the tests on the Panasonic AVC-I codec were performed by the EBU using a prototype encoder, and the Sony XDCAM HD50 was tested by the RAI using a prototype encoder. Engineers from the respective manufacturers attended the tests for at least the first day to ensure the correct working of the equipment.

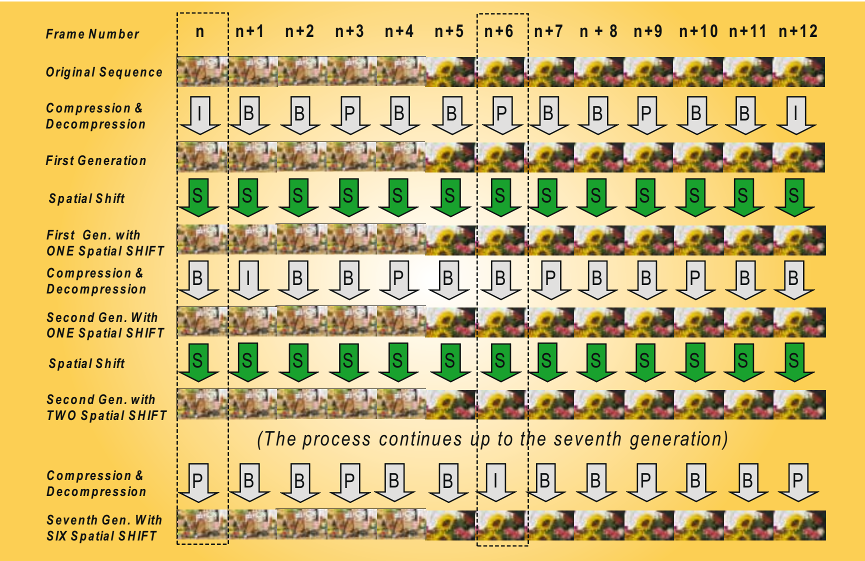


Figure 8
Standalone chain without GoP alignment AND with spatial shift for INTER codec

All the sequences (original and those subjected to compression) were stored in uncom-pressed form (YUV10 format) on two video servers that were able to be run in parallel. The sequences were displayed simultaneously in a vertical split-screen condition (original on one side, those subjected to compression on the other side). During the expert viewing, the T bar of the mixer that provided the split-screen effect (a Panasonic type AV-HS300) was moved to compare the same areas of different versions of the same sequences, as was found necessary. The position of the split was made more evident by applying a just-noticeable vertical white bar at the transition between the two images; a real example is shown in *Fig. 9*.

The viewing distance was marked on the floor and was set to 3H, where H defines the vertical dimension of the display, and the observers were asked to respect this viewing distance. Sometimes a closer viewing distance, e.g. 1H, was used to closely observe small details and artefacts and, when used, this condition was clearly noted in the report.

It should be made clear that this method allows the evaluation of very small impairments, near to the visibility threshold, and it must be considered a very severe analysis of the picture quality.

As mentioned, the tests focused on the performance of the algorithms at the first, fourth and seventh generations, comparing the picture quality with the original (headroom evaluation) or with legacy system (improvement provided by the new system). A test list, which summarized in the form of tables all the different comparisons planned, was prepared and discussed in advance.

Each expert was provided with a paper copy of the test list, so she/he was always completely aware of what sequence was displayed in each part of the screen.

For example, in the case of a comparison between the "Original" sequence and the

4. Analysis of the performance of the algorithms

The analyses of the performance of the algorithms was performed both using objective measurements (PSNR) and visual scrutiny of the picture (i.e. expert viewing), as described in the following sections. These two methods provide different kinds of information and they are considered to be complementary.

4.1. Objective measurements

The PSNR has been computed via software and obviously applied a procedure to reestablish the spatial alignment between the original and the de-compressed version of the test sequence. Moreover, it skipped 16 pixels on the edges of the picture to avoid taking measurements on the black pixels introduced during the shift.

The formula used to evaluate the PSNR via software was:

$$PSNR = 10Log_{10}\left(\frac{Vpeak^2}{\frac{\sum_{i=1}^{Ncol}\sum_{j=1}^{Nlin}(ori(i,j)-cod(i,j))^2}{Ncol * Nlin}}\right)$$

where: *ori (i,j)* = original frame,
cod(i,j) = manipulated frame,
Ncol = horizontal resolution in pixels,
Nlin = vertical resolution in pixel and
Vpeak = $2^{10} - 1$ = 1023.

The results are expressed in dB.

It is well known that PSNR does not correlate accurately with the picture quality and thus it would be misleading to directly compare PSNR from very different algorithms. On the other hand, this parameter can provide information about the behaviour of the compression algorithm through the multi-generation process.

4.2. Expert viewing

Analysis of the algorithm performance (from the point of view of picture quality) was carried out using so-called "expert viewing". Even if this method is not formally included in any ITU Recommendation, it is very often used as it provides fast and consistent results. Under the generic name of "expert viewing" are included several kinds of analysis of the picture quality evaluation.

For the purposes of these P/HDTP tests, it was agreed in advance with manufacturers to use the following interpretation of what "expert viewing" would entail.

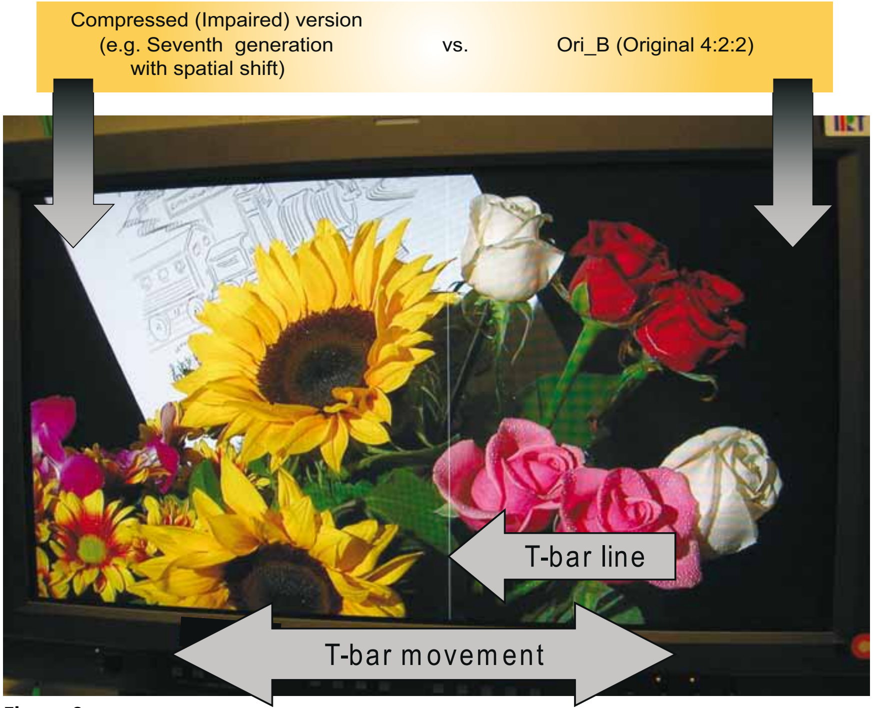


Figure 9
Standalone chain with GoP alignment AND with spatial shift for INTER codec

version subjected to seven generations of algorithm "A", including a spatial shift between each generation, the expert was provided with the following table:

Compressed (Impaired) version (e.g. Seventh generation with spatial shift)	vs.	Ori_A (Original 4:2:2)
Ratio: loss due to seven generations with post-processing		

The expert was aware that "A" meant a specific vendor and was given full details of each test condition (bitrate, GoP aligned or not, etc.). The table also included a short sentence describing the "rationale" of the test and, in order to avoid any misunderstanding, the position (right or left) of the sequences was the same on the paper table and on the display; c.f., the Original sequence on the right side and the processed sequence on the left side in the above table.

All the experts were formally requested to refrain from expressing their opinions during the sessions, in order to avoid biasing other people. Only after the complete analysis of the test sequence (for its full length) was a discussion started to summarize in a few sentences the opinions of the different experts. Sometimes it was not possible to get an agreement on the visual analysis of a sequence and in this case the sequence was repeated. If even in this case an agreement was not possible, the situation was noted in the report.

Best efforts were made to guarantee that the panel of experts was comprised of the same people during the different expert viewing sessions; this condition was readily met during each single day and almost perfectly so during different days. One day of expert viewing was dedicated to each manufacturer, and representatives of the individual vendor took part in the expert viewing as well.
The results of the expert viewings were collected in a series of EBU BPN documents; BPN 076 (results for Avid DNxHD), BPN 077 (results for GVG/Thomson JPEG2000), BPN 078 (results for Panasonic AVC-I) and BPN 079 (results for Sony XDCAM HD50). These documents are pub-lished by the EBU exclusively for its Members.

The reader should be aware that, due to the complexity and framework of the tests, only a deep analysis of these documents can provide a complete appreciation of the results.

4.2.1. Display

The following displays were used during the tests:

- CRT 32" Sony: Type BVM-A32E1WM
- CRT 20" Sony: Type BVM-A20F1M
- Plasma Full HD 50": Type TH50PK9EK Panasonic
- LCD 47": Type Focus

The displays were connected through HD-SDI interfaces. The displays were aligned according to the conditions described in ITU-R BT.500-11 and the room conditions were set accordingly.

The final assessment was always done while considering the quality perceived on the CRT displays.

Nevertheless, there was a general agreement that the flat-panel displays, both LCD and plasma, magnified the impairments.

5. Results

It was agreed between the EBU project group and the vendors to make the reports about the test details available to EBU Members only. In late 2007, the results of the test were published as BPN076 to BPN079, as noted above.

Note: Due to the importance of the subject, vendors and the EBU agreed to provide some preliminary results in a PowerPoint presentation given at the IBC-2007 conference, before the actual BPN reports became available to EBU Members. This PowerPoint presentation contained a short summary of the test results in tabular form. The published reports BPN076 to BPN079 contain a much larger framework of test conditions than shown in the IBC-2007 PowerPoint. Neither the test reports nor the PowerPoint tables are intended, or suited, for comparative studies. The tabular form of the PowerPoint presentation did not include information about whether the

Abbreviations

720p/50	High-definition progressively-scanned TV format of 1280 x 720 pixels at 50 frames per second
1080i/25	High-definition interlaced TV format of 1920 x 1080 pixels at 25 frames per second, i.e. 50 fields (half frames) every second
1080p/25	High-definition progressively-scanned TV format of 1920 x 1080 pixels at 25 frames per second
1080p/50	High-definition progressively-scanned TV format of 1920 x 1080 pixels at 50 frames per second
AVC	(MPEG-4) Advanced Video Coding, part 10 (aka H.264)
DCT	Discrete Cosine Transform
GoP	Group of Pictures
ITU	International Telecommunication Union www.itu.int
ITU-R	ITU - Radiocommunication Sector www.itu.intpublications/sector.aspx?lang=en&sector=1
JPEG	Joint Photographic Experts Group www.jpeg.org
MPEG	Moving Picture Experts Group www.chiariglione.org/mpeg/
NLE	Non-Linear Editing
PSNR	Peak Signal-to-Noise Ratio
SMPTE	Society of Motion Picture and Television Engineers (USA) www.smpte.org
YUV	The luminance (Y) and colour difference (U and V) signals of the PAL colour television system

tests were conducted with or without pixel shift.

The EBU Production Management Committee then subsequently concluded in a recommendation (EBU R124-2008)[2] that:

> ***For acquisition applications an HDTV format with 4:2:2 sampling, no further horizontal or vertical sub-sampling should be applied. The 8-bit bit-depth is sufficient for mainstream programmes, but 10-bit bit-depth is preferred for high-end acquisition. For production applications of mainstream HD, the tests of the EBU has found no reason to relax the requirement placed on SDTV studio codecs that "Quasi-transparent quality" must be maintained after 7 cycles of encoding and re-coding with horizontal and vertical pixel-shifts applied. All tested codecs have shown quasi-transparent quality up to at least 4 to 5 multi-generations, but have also shown few impairments such as noise or loss of resolution with critical images at the 7th generation. Thus EBU Members are required to carefully design the production workflow and to avoid 7 multi-generation steps.***

The EBU recommends in document R124-2008 that:

- If the production/archiving format is to be based on I-frames only, the bitrate should not be less than 100 Mbit/s.
- If the production/archiving format is to be based on long-GoP MPEG-2, the bitrate should not be less than 50 Mbit/s.

Furthermore, the expert viewing tests have revealed that:

- A 10-bit bit-depth in production is only significant for post-production with graphics and after transmission encoding and decoding at the consumer end, if the content (e.g. graphics or anima-tion) has been generated using advanced colour grading, etc.
- For normal moving pictures, an 8-bit bit-depth in production will not significantly degrade the HD picture quality at the consumer's premises.

[2] *R124-2008:* http://tech.ebu.ch/docs/r/r124.pdf

6. Bibliography

[1] ITU-R BT.500-11: **Methodology for the subjective assessment of the quality of television pictures**
ITU-R Geneva, 2003.

[2] ITU-R BT.709-5: **Parameter values for the HDTV standards for production and international programme exchange**
ITU-R Geneva, 2002.

[3] SMPTE 296M: **1280 x 720 Scanning, Analogue and Digital Representation and Analogue Interface**
SMPTE New York, 2001.

[4] SMPTE 292M: **Bit-Serial Digital Interface for High-Definition Television Systems**
SMPTE New York, 1998.

[5] SMPTE 274M: **1920 x 1080 Scanning and Analogue and Parallel Digital Interfaces for Multiple Picture Rates**
SMPTE New York, 1998.

[6] Sony: www.sonybiz.net and search for XDCAM-HD 422

[7] Panasonic: www.panasonic.com/business/provideo/p2-hd/white-papers.asp

[8] Avid: www.avid.com/dnxhd

[9] Thomson Grass Valley: http://thomsongrassvalley.com/products/infinity/

Acknowledgements

The authors wish to acknowledge the considerable work carried out by Giancarlo De Biase (RAI), Dagmar Driesnack (IRT), Adi Kouadio (EBU) and Damian Ruiz Croll (RTVE) during the codec testing.

First published: 2008-Q3

Technical trial of the EBU P2P media portal

Franc Kozamernik
Senior Engineer, EBU

The EBU P2P Media Portal (EBUP2P) was developed as a demonstration tool for EBU Member organizations to show their television and radio channels internationally.

A 6-month trial of the portal was set up by EBU Project Group D/P2P in the first half of 2008 to evaluate P2P (Peer-to-Peer) technology – via the example provided by the company Octoshape. This article reports on the outcome of the trial.

EBU Project Group D/P2P (Peer-to-Peer), chaired by Frank Hoffsümmer (SVT), was set up in 2006 in order to:

- evaluate newly-emerging peer-to-peer technologies;
- formulate potential EBU requirements for such systems;
- to carry out technical experiments and trials on a working P2P system.

The Group was encouraged by the large attendance at a seminar organized by the EBU Technical Department in February 2006 in Geneva, which identified significant interest among EBU Members in studying this new technology.

The initiative for an EBU Media Portal came from the Group D/P2P at its meeting in June 2007. The Group proposed that EBU Members should come together and set up a common Internet portal through which Members' TV and radio channels could be made available to the general public worldwide. The portal – if ever established permanently as a non-commercial collaborative EBU project – could mean much more than just a technically advanced and innovative project: it could become a one-stop shop

What is Peer-to-Peer?

In this article, P2P is considered as an Internet *media distribution* system which relies on end-users' computers to propagate content through existing computer networks. Such a P2P system has nothing to do with napsterization or illegal content-sharing. Quite the contrary, P2P offers an attractive possibility for broadcasters to distribute their content efficiently across the Internet.

As P2P does not require any special emission infrastructure to be installed, the investments and maintenance costs are significantly lower than those of more traditional Content Distribution Networks (CDNs) which may use several thousands of special streaming servers. In addition, P2P does not have a single point-of-failure, so its service reliability is very high.

On the downside, P2P is a relatively new technology which requires a lot of further studies and hands-on experience in order to turn an interesting technical innovation into a viable business proposition.

window of Members' programming achievements.

Prior to the advent of P2P, EBU Members were extensively using different server-client approaches such as CDNs (Content Distribution Networks), IP Multicasting and a great variety of proprietary streaming solutions from Microsoft, Real Networks and QuickTime. A common characteristic of all these systems is that they are relatively expensive ... because broadcasters must pay an ISP (Internet Service Provider) for each connected stream (i.e. user). Thus, the more users there are, the higher the ISP's costs will be. Effectively, broadcasters pay more because of their own success.

As an example, in 2005, the Eurovision Song Contest was distributed via the Internet using a CDN system serviced by Akamai. Being a highly popular event, it generated a lot of interest worldwide (several tens of thousands of Internet users). The calculated cost of video streaming the event was about one CHF per user per hour, which amounted to around CHF 100'000 (€ 65'000) for the 3-hour show.

P2P may change this paradigm radically. Video can be streamed via P2P at a cost that is below € 0.05 per GB[1]. As a result of the competition from P2P, the cost of CDN services has also dropped significantly but it is still much more expensive than P2P by an order of magnitude.

In addition to the cost factor, P2P technologies have many advantages compared to other Internet distribution systems: no central server streaming "farm" is required and there is no central point of failure (assuming a decentralised, distributed tracker). However, P2P is still in its infancy and many challenges are still to be resolved. Setting up EBUP2P is only the first step in the direction of gathering experience and solving potential issues relating to P2P.

Requirements for an Internet media-distribution system

Broadcasters have the following requirements for any Internet distribution system they might wish to use:

- low distribution cost (ideally independent of location, time, quality and number of users);
- reliable delivery (no glitches or interruptions, reasonable end-to-end latency, fast zapping);
- high quality levels – SD, even HD (including multichannel audio if required);
- large channel capacity (in principle, there are no frequency spectrum constraints as in conventional broadcasting);
- the largest number of concurrent users possible (several hundreds of thousands of concurrent P2P users have been successfully demonstrated).

The P2P trial

In order to make the whole operation manageable, we had to limit the number of Member participants to about ten (see *Table 1*). The P2P system chosen for the trial was provided by Octoshape, already described in an earlier edition of EBU Technical Review[2].

		Channel	Organization	URL	Comment
TV	1	HR	Hessischer Rundfunk (HR)	www.hr-online.de	
TV	2	DW - TV	Deutsche Welle (DW)	www.dw-world.de	
TV	3	TV SLO 1	RTV Slovenia (RTVSLO)	www.rtvslo.si	
TV	4	TV SLO 2	RTV Slovenia (RTVSLO)	www.rtvslo.si	
TV	5	24H tve	RTV Spain (RTVE)	www.rtve.es	
TV	6	DOCU tve	RTV Spain (RTVE)	www.rtve.es	Discontinued
TV	7	TV Ciencia on-line	TV Portugal	www.tvciencia.pt	Since May 08
TV	8	iTVP	Polish TV (TVP)	www.itvp.pl	Since May 08
TV	9	Taiwan TV			Since June 08
Radio	10	radio-suisse jazz	SRG-SSR	www.srg-ssr.ch	MP3, 192 kbit/s
Radio	11	radio-suisse pop	SRG-SSR	www.srg-ssr.ch	MP3, 192
Radio	12	radio 3	RNE	www.rne.es	WM, 96
Radio	13	radio classica	RNE	www.rne.es	WM, 128
Radio	14	youfm	HR	www.hr-online.de	MP3, 160 – up to the end of April 08
Radio	15	RSI	RTVSLO	www.rtvslo.si	WM, 192
Radio	16	Val 202	RTVSLO	www.rtvslo.si	WM, 192

Table 1
Member participants in the EBU P2P Portal trial

1) *Octoshape states a price of 2 Eurocents per GB on its website: http://www.octoshape.com*

2) *Visit: http://tech.ebu.ch/docs/techreview/trev_303-octoshape.pdf*

Basic portal description

During the trial, there was a highly-visible vignette "P2P Media Portal Trial" on the EBU's home page, which took the user to the portal's own home page[3]. Access to the portal was open to all users worldwide and *Fig 1* shows an earlier design of the home page.

During the trial, the design of the web portal underwent continuous improvements, both graphically and in terms of accessibility and user friendliness. *Fig. 2* shows the current graphical design of the portal, as produced by Nathalie Cullen from the EBU's Communication Service.

Each Member participant is represented by an icon which emulates their logo. On the top right side, there are eight icons representing TV channels, while below them are six radio channels. In this new design, users during the trial had the flexibility of using either an embedded player (which starts playing automatically when you first open the page) or Windows Media Player (which opens in a separate adjustable-size window).

Underneath each logo we put a link to the Member's URL, allowing the users to consult the schedule of programmes and access some additional information about the channel concerned.

On the bottom of the page we included a temporary link "Your Comments", allowing the users to send in reports and comments about their viewing experiences. During the trial we received more than a hundred comments which were all consistently positive about the user experience.

The page also included information about what the user needed in order to access the portal content.

These are as follows:

- a suitable broadband connection;
- a PC running Windows, Mac or Linux;
- Internet Explorer (IE) or Firefox browser;
- Active X (in the case of IE);
- Windows Media Player (video and audio);
- MP3 player;
- Octoshape plug-in.

The trial plan

The first issue in the process of establishing an operational EBU Media Portal was to identify a suitable Internet distribution technology and to agree some operational parameters.

A technical trial Group consisting of ten EBU Members held a kick-off meeting on 10 July 2007 in Geneva, in order to define the technical and operational parameters based on peer-to-peer (P2P) technology[4]. Once these parameters had been set, the trial could start informally in autumn 2007 but officially it started in January 2008. It then continued for six months until the end of June 2008.

The Trial Group was coordinated by the Author and operated under the auspices of the D/P2P Group. It held three meetings in order to supervise the development and monitor the technical quality of the portal.

Figure 1
EBU P2P Media Portal – Windows Media Player only *(Courtesy: Nathalie Cullen, EBU)*

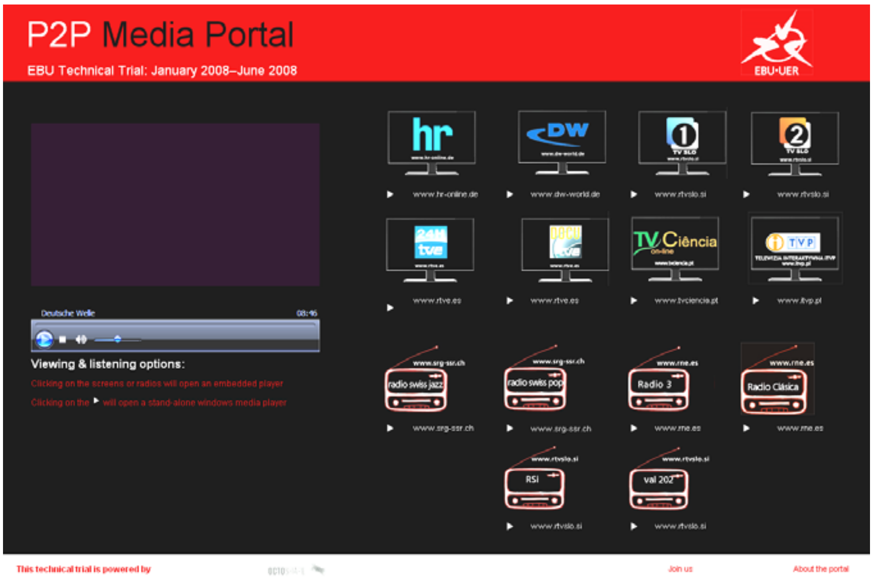


Figure 2
Final design of the EBU P2P Media Portal *(Courtesy: Nathalie Cullen, EBU)* **– which uses either Windows Media Player or an embedded player in the browser**

[3] *The deep link to the Media Portal can still be found here: http://www.ebu.ch/members/EBU_Media_portal_Trial_1.php*

[4] *Octoshape P2P technology was selected on the basis of some comparative tests conducted by Project Group D/P2P during IBC 2006 and also as a result of Eurovision Song Contest events from 2005 onwards, for which Octoshape was very successfully used.*

	Test	Description	Participants
1	Accessibility of contents from the EBU web site	Working links, smooth zapping of channels, only one stream at a time	all
2	Overall quality performance (continuity and reliability) under different broad-band network conditions	Different upstream and downstream capacity, different network load (resulting in packet loss and jitter)	all
3	Scalability: 200 and 700 kbit/s	Number of concurrent users all	all
4	Different computer platforms/ operating systems	PC, Mac, Linux	IRT
5	Different browsers	IE, Firefox, etc	IRT
6	Geolocation (dynamic)	Displaying the messages for non-availability of streams	?
7	Rights Management	Different DRM system including MS DRM, DVB CPCM	?
8	Audience statistics	Immediate return of data for each broadcaster	All, EBU
9	Pre-roll advertising	Octoshape inserts a up to 3 s pre-roll ad. Content of add should be mutually agreed by channel owner and Octoshape	HR, TVP
10	Audio Watermarking [a]	Embed about 100 bit/s	?
11	Flash codec [a]	In addition to WM, we should test Flash codec	?
12	On-demand delivery [a]	Video on-demand play out of files	?

Table 2
Technical test planned

[a)] *To be tested in the second phase (subject to agreement with Octoshape)*

The principal objective of this trial was to assess whether or not the Octoshape P2P technology is an efficient and reliable platform for live streaming of Members' television and radio services. We also seized the opportunity to try some peripheral services such as the graphical design and accessibility of the Portal (e.g. how the user accesses the site and changes channels – zapping), along with video encoding, geolocation, pre-roll advertising, etc

Based on the above system and the operational requirements, a test plan was developed, as shown in *Table 2*.

Technical specification of the trial

The agreed technical parameters for the audio and video streams is given in *Table 3*.

Players: Windows Media and embedded Octoshape player.

Distribution via Internet: P2P and HTTP.

Television	Video		
	Codec	Bitrate	Aspect ratio & resolution
	MS Windows Media	about 700 kbit/s	4:3 480 x 360 px; 16:9 520 x 360 pixels
	Audio		
	Codec	Bitrate	Stereo/Mono
	MS Windows Media 9	64 kbit/s	48 kHz, stereo (A/V) 1-pass CBR
Radio	Audio		
	Codec	Bitrate	Stereo/Mono
	MS Windows Media 9	96/128/192 kbit/s	44.1/48 kHz, stereo (A/V) 1-pass CBR
	Mpeg Layer 3	192 kbit/s	

Table 3
Technical parameters used in the EBU P2P Media Portal

Abbreviations

CBR	Constant Bit-Rate	**EPG**	Electronic Programme Guide	**P2P**	Peer-to-Peer
CDN	Content Delivery Network	**FS**	Full Scale	**SDI**	Serial Digital Interface
CE	Consumer Electronics	**HTTP**	HyperText Transfer Protocol	**WM**	(Microsoft) Windows Media
DRM	Digital Rights Management	**ISP**	Internet Service Provider	**XML**	eXtensible Markup Language

Distribution of work

The following is the list of tasks that each participant had to accomplish in order to make the trial a successful operation.

Octoshape

Octoshape provided the following services to the Portal:

- Information to the EBU to enable EBU Members to encode their streams in WM;
- Octoshape-specific Source Signal Solution (SSS) software to all Members for encoding their material;
- If required by a Member, performed encoding (or asked third party to do it);
- Information to EBU to inform EBU Members how to send streams to Octoshape;
- User's plug-in (Octoshape-specific) with regular updates;
- Powered the Portal by providing P2P services for live streaming;
- Provided audience statistics on request to all participants;
- Provided geolocation services (if required);
- Provided pre-roll advertising (if required).

EBU staff

The EBU had the following responsibilities:

- Ensured that the technical, legal and programming interests of EBU Members were fully respected and taken into account;
- Coordinated the evaluation process;
- Developed a dedicated website in coordination with Octoshape, and according to Octoshape requirements, and provided a link to the EBU home page;
- Ensured a proper design of the web page, with constant improvements to the look and feel;
- Ensured a balanced (non-discriminatory) visibility to all the channels involved;
- Provided all the information required for the end user to access all the channels and other information required;
- Enabled the user to download the latest version of the required media players and the Octoshape plug-in;
- Reported regularly to the various EBU bodies on the progress of the technical trial;
- Ensured some publicity for the portal in order to maximize the use of the site;
- Promoted the portal at relevant events, EBU seminars, conferences and trade shows;
- Discussed common strategies for pre-operational and regular services with Octoshape, i.e. the future steps and business model options to be considered.

EBU Members – Trialists

The EBU Members that participated in the trial coordinated their activities through the EBU Technical Department. Participants conducted the following activities:

- Provided the content, i.e. selected the TV/radio channels or any other streaming content they wished to publish on EBU website[5];
- Granted rights to the EBU for their content to be published on the EBU website;
- Encoded their streams (by using appropriate technologies);
- Forwarded their streams to Octoshape;
- Defined the coverage constraints (if and when required) and instructed Octoshape how to apply geolocation filtering;
- Provided a link to the EBU website on their own website, so that users in their country could easily access other EBU Members' content;
- In the case of pre-roll adverts, Members performed editorial control of the ad content.

Legal matters

Throughout the course of the trial, legal matters – particularly copyright issues – played an important role in our discussions. Initially, EBUP2P thought we could be considered a kind of re-broadcaster of EBU Members' content and be treated as a cable network. To this end, the EBU (as owner of the portal) needed explicit permission from Members to re-broadcast their channels. If required, the EBU also needed to clear the rights issues with the collecting societies. All participants were required to sign a rights clearance form that there were no legal obstacles for the EBU to make the TV channel(s) available to the general public, free of charge and in unchanged form and simultaneously with the terrestrial broadcasts of these channel(s).

A later discussion showed that, in practive, the end user who clicks on the icon on the EBU website to access a certain TV channel is merely redirected to the original stream of the actual content pro-vider. Therefore, the EBU is not re-transmitting the stream. It was agreed to publish a disclaimer which reads:

> ***It should be noted that the users are redirected to the original stream of the actual content provider and the EBU is not re-transmitting the stream.***

Another solution for a future portal would be for the EBU to simply provide the links to the Members' web pages containing embedded players.

Evaluation results

Audience

The table below shows how many unique users were able to join the trial across all channels offered during the January - June

---

5) *Generally all webcasts should be available for 24 hours a day but it is up to the individual Members to decide on their webcast times. In the latter case, information should be given about the broadcast times.*

2008 period. It also shows the aggregate time (in hours) of user media consumption for each month.

Month	Unique users	Total hours
Jan	9'072	32'375
Feb	8'157	39'777
March	8'652	44'898
April	7'031	41'519
May	3'808	26'247
June	2'936	24'184

Fig. 3 shows an example of the audience variations on two Hessicher Rundfunk channels, one TV and one radio, during the political campaign before the elections at Hessen, Germany. Members of political party SPD were able to watch the TV stream via our P2P stream, as they did not have TV sets in their offices.

Please note that the Octoshape system was limited to serve no more than 600 concurrent users, as the broadcaster did not notify Octoshape in advance (see the section on "Scalability" below).

Accessibility

All web links were working satisfactorily, zapping of channels was smooth and only one stream at a time was available (as required). Some occasional glitches (e.g. missing sound, lip-sync problems) occurred due to poor encoding. The Spanish Documentary channel is not available outside of Spain, due to copyright, and the geolocation filtering was applied to implement this. The end user was informed about this via a message that popped up "This stream cannot be viewed in your country".

Some users had problems with downloading the Octoshape plug-in. Some people, particularly those located in large corporations (including some large broadcasting organizations) could not download the plug-in at all, and consequently were not able to access the Portal services. This element of the Octoshape system needs to be considered, as it represents an obstacle to audience acceptance.

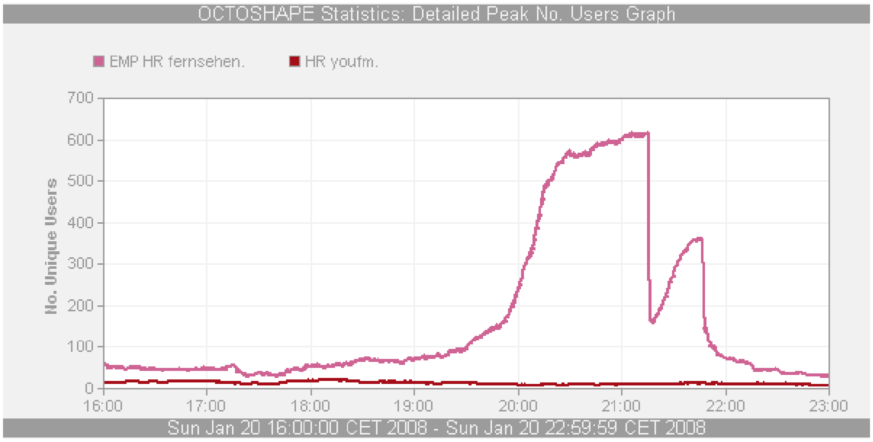


Figure 3
Audience variations on HR's radio and TV channels

Overall quality performance

There has been no evidence that network traffic load, asymmetricity or last-mile issues affected the overall quality performance of the Octoshape system in any significant manner. It should be stressed however that only large-scale laboratory tests, allowing for controlled repeatability of results, could yield scientifically-valid results. According to the reports received, our experience about the service quality was positive – we can deduce that Octoshape performed correctly on all networks.

The service quality therefore mainly depends on the encoding quality. We detected some errors performed by Members in encoding video material, in particular regarding the correct aspect ratio when the source material was produced in HDTV (16:9).

Should there ever be a regular service, the question of correct encoding requires extremely careful consideration. Differences between different sources should be avoided, so that zapping from one channel to another does not result in level and other differences. Broadcasters should adopt a common set of coding parameters. Square pixels should be consistently used.

Resolutions:
4:3 – 512 x 384 pixels
16:9 – 512 x 288 pixels

De-interlacing problems should be handled in the hardware (encoder card). Complexity coding should be enabled. If possible, performance could be enhanced by a 2-step encoding. Use of SDI sources is highly recommended while composite sources (to reduce cross-colour and cross-luminance interference) is to be avoided

Audio level: 0 dB FS level (full scale).

Scalability

On a major Internet event in May, HR experienced a service breakdown. Octoshape explained that the 600[6] user limit was configured on the Octoshape P2P network by default (which means that the service was effectively cut down when more than 600 peers joined the network). This was explained as a normal precautionary measure, as Octoshape does not know the network (ISP) limits which may vary from one network to another. Octoshape is however able to scale bandwidth to the ISP limits.

Such a service breakdown could have been avoided if Octoshape had been informed in advance of an event where it was likely that a larger number of peers may join.

[6] *Octoshape explained that the number 600 is really arbitrary and can well be set to a much higher value if required (especially for live events).*

Franc Kozamernik graduated from the Faculty of Electrotechnical Engineering, University of Ljubljana, Slovenia, in 1972.

He started his professional career as an R&D engineer at Radio-Television Slovenia. Since 1985, he has been with the EBU Technical Department and has been involved in a variety of engineering activities covering satellite broadcasting, frequency spectrum planning, digital audio broadcasting, audio source coding and the RF aspects of various audio and video broadcasting system developments, such as Digital Video Broadcasting (DVB) and Digital Audio Broadcasting (DAB).

During his years at the EBU, Mr Kozamernik has coordinated the Internet-related technical studies carried out by B/BMW (Broadcast of Multimedia on the Web) and contributed technical studies to the I/OLS (On-Line Services) Group. Currently, he is the coordinator of several EBU R&D project groups including B/AIM (Audio in Multimedia), B/VIM (Video in Multimedia) and B/SYN (Synergies of Broadcast and Telecom Systems and Services). He also coordinates EBU Focus Groups on Broadband Television (B/BTV) and MultiChannel Audio Transmission (B/MCAT). Franc Kozamernik has represented the EBU in several collaborative projects and international bodies, and has contributed a large number of articles to the technical press and presented several papers at international conferences.

Octoshape can scale down the quality from 700 kbit/s to 200 kbit/s[7] either automatically or manually.

In order to implement the scalability and ensure continuous services (even when the number of users increases), it is necessary to encode the streams in two (or three) bitrates. In this way, Octoshape can perform an automatic switch to a lower bitrate, as soon as the number of peers reaches a certain limit.

It should be pointed out that dual bitrate encoding can be implemented by a single PC.

Octoshape has already demonstrated on a number of occasions that it is a very scalable system, e.g. for the Eurovision Song Contest during which it was able to support around 45'000 simultaneous streams without any problems.

Computer platforms and browsers

There was no evidence of any problems resulting from the use of different computer platforms, operating systems and browsers.

Geolocation

Throughout the trials, the Group gave very serious consideration to issues relating to geolocation filtering, as this tool is essential to limit coverage to specific areas, mainly for copyright reasons. The geolocation system must obey very strict requirements in terms of security, accuracy and reliability, in order to prevent any leakage of content outside the granted zone.

The Octoshape geolocation system is an advanced commercial product called "IP2location". This system enables identification of the geographic location and Internet domain name by means of an IP address. The IP2location database is used to match an incoming IP address to the country, region, city, latitude, longitude, zip code, Internet Service Provider (ISP), time zone, network speed and domain name of the Internet user. Octoshape merely provides an interface to this database. Octoshape believes that the geolocation system they are using provides excellent accuracy and security. In the many years that they have been using this system, no difficulties have been experienced whatsoever. If necessary, the Octoshape system would be able to interface to any other geolocation system including Akamai or Quova if required. Members agreed that EBUP2P should meet the highest level of geolocation performance and security as required.

In the case when geofiltering is applied, copyrighted content is only available within the copyrighted zone, whereas outside this zone a message board should be displayed on the screen. The text on the message board may say the following:

> ***"Due to copyright restrictions, the currently broadcast programme is only available within the authorized zone. Your location lies outside this zone, therefore at the moment you have no access to the content. We apologize for any inconvenience."***

Preferably the above message is shown in the national language and in English; other languages may be added of course. Geofiltering can only be applied to video, while leaving the audio available.

The scheduling of geofiltering could be two-fold: [a] pre-scheduled (pre-determined start and end times) or [b] flexible (time stamps or manual activation of geofiltering).

7) *The lower limit of 200 kbit/s has been increased to 350 kbit/s to improve the lower quality limit.*

Rights management

Digital Rights Management (DRM) was not tested: it is independent of the P2P system in use.

Pre-roll advertising

A pre-roll system is an example of a business model in which the end user receives all video and audio streams for free. For the testing of EBUP2P, Hessischer Rundfunk used a 5s pre-roll service. Octoshape confirmed that they were able to provide a pre-roll advertising technology which is very similar to the Zattoo one. However, if pre-roll is to be used for an operational service, several related open questions apply:

- Who provides the pre-roll ad content?
- Should the ad content be adapted to the destination market?
- Is the ad content both channel- and zone-specific (leading to different pre-rolls for different markets)?

Preliminary conclusions

The principal conclusions of this trial can be summarized as follows:

- EBUP2P represents a state-of-the-art technical solution and fulfils all the tested technical and operational requirements – in terms of the service quality, scalability, video and audio quality, accessibility, security and user-friendliness.
- EBUP2P has no technical limitations regarding the number of radio and TV channels to be accommodated in a Portal. Members can flexibly join in and opt out at any time.
- EBUP2P can fulfil our requirements concerning copyright, by applying territorial filtering (ge-olocation) and watermarking.
- EBUP2P enables a number of business models.
- EBUP2P is future proof and will be extended towards CE (consumer electronics) devices.

In spite of the very limited human and financial resources available for conducting the EBU P2P Media Portal trial, we brought the trial to a successful end, according to the schedule planned. The participants in the trial performed a large number of technical and operational tests.

The number of participating EBU organizations was restricted to about ten. We could not accept more participants, as our logistic resources are so limited. Also the number of end users were quite modest, as we did not carry out any significant promotion of the portal.

These EBU tests cannot be considered rigorous scientific tests. They were more akin to “proof of concept” and “experience-gathering” evaluations. Octoshape is not the only commercial P2P system available in the market but we selected it because our previous experience with this system was positive.

The main conclusion of the trial is that Octoshape is an excellent Internet distribution system for carrying audio and video streams to PC users. The system is scalable, reliable, easy to manage and interoperable with a number of codecs, operating systems, browsers and geolocation systems. In the course of the project a large number of issues were successfully resolved, although several issues were left open for future activities (see the next section).

The P2P Media Trial has shown that Octoshape can be used by our Members as a viable and technically appropriate system for the distribution of audio and video streams across the Internet. It can be used either as a standalone distribution system or in combination with some CDN or IP Multicasting technologies.

Required future work

Running a Media Portal is a complex issue and technologies evolve very rapidly. It is not enough to show that the P2P distribution system functions correctly and according to our expectations. For a possible future commercial portal, the following additional issues may also be considered:

- Watermarking (and Fingerprinting) – optionally embed audio watermarking signals to detect the originator of the content (if required);
- Allow for both customized media players (which open in a separate page) and embedded players;
- Optionally embed DRM in the stream (if required to control consumption of the media received) [8];
- Develop an XML-based template for EPGs and optionally provide an EPG (daily, weekly) for each channel;
- Provide additional coverage of special events (if required);
- Extension of portal services to embrace content downloading and VoD (on-demand) services (documentaries, archives, recorded sports events, etc);
- Hybrid TV receivers with broadband (Ethernet or Wi-Fi) connection: embedded P2P client in commercial TV sets and set-top boxes (such as in DVB);

Acknowledgements

The EBU would like to thank all participants for their very kind cooperation in this field trial. Particular thanks should go to the Octoshape staff – Stephen Alstrup, Theis Rauhe and Lasse Riis – for their fast and efficient resolution of all problems that may have arisen during the tests. Many thanks also go to Nathalie Cullen from the EBU Communication Service for the attractive and functional web design of the portal. And last (but not least), thanks should also go to the EBU management for providing continuous support and encouragement for conducting this project.

8) *Participants will be able to apply different DRM (digital rights management) tools in order to protect their content.*

First published: 2008-Q3

Streaming audio contributions over IP – a new EBU standard

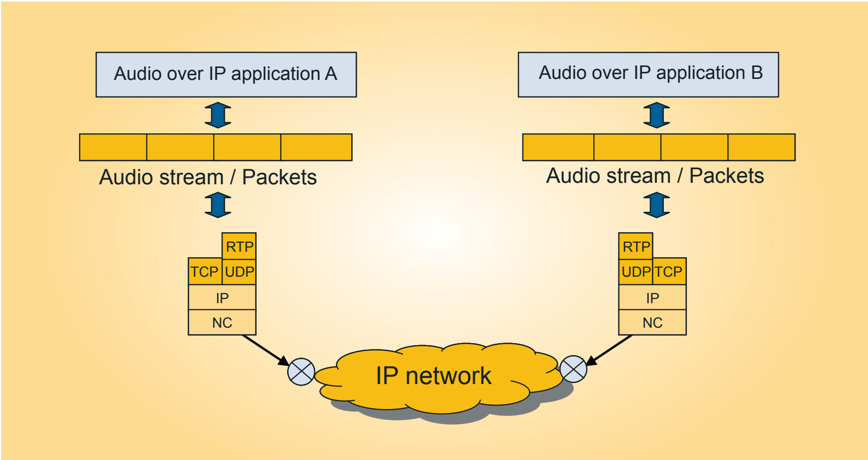
Lars Jonsson
Swedish Radio

Mathias Coinchon
EBU Technical Department

Audio-over-IP end units are increasingly being used in radio operations for the streaming of radio programmes over IP networks, from remote sites or local offices into main studio centres. The IP networks used can be well-managed private networks with controlled Quality of Service. However, the open Internet is increasingly being used also for various types of radio contribution, especially over longer distances. Radio correspondents will have the choice in their equipment to use either ISDN, the Internet via ADSL or other available IP networks to deliver their reports. ISDN services used in broadcasting will be closed down in some countries.

The EBU has created a standard for interoperability in a project group, N/ACIP (Audio Contribution over IP). This standard, which has been jointly developed by members of the EBU group and manufacturers, is published as EBU Tech 3326-2007. The standard has quickly been implemented by the manufacturers. A "plug test" between nine manufacturers, held in February 2008, proved that earlier incompatible units can now connect according to the new standard.

Internet Protocol (IP) is used worldwide on the Internet and also on the IP-based corporate or private networks used by broadcasters. It is independent of the underlying data transmission technology and many IP adaptations exist for the physical layers, e.g. Ethernet, ATM and SDH over copper, fibre or radio links.



Applications can communicate in a standardized way over interconnected IP networks. Host computers can be reached almost instantly wherever they are located. Fixed or temporary connections for audio contributions can share similar types of applications. The connection is established by dialling a number or an e-mail-like name. The audio stream is then sent using standardized protocols (SIP, RTP, UDP). Many types of audio coding formats can be used at various bitrates. Higher bitrates will allow stereo audio or multi-channel with linear PCM coding. The permitted maximum audio bitrate will depend on the bandwidth and if the network has a good QoS.

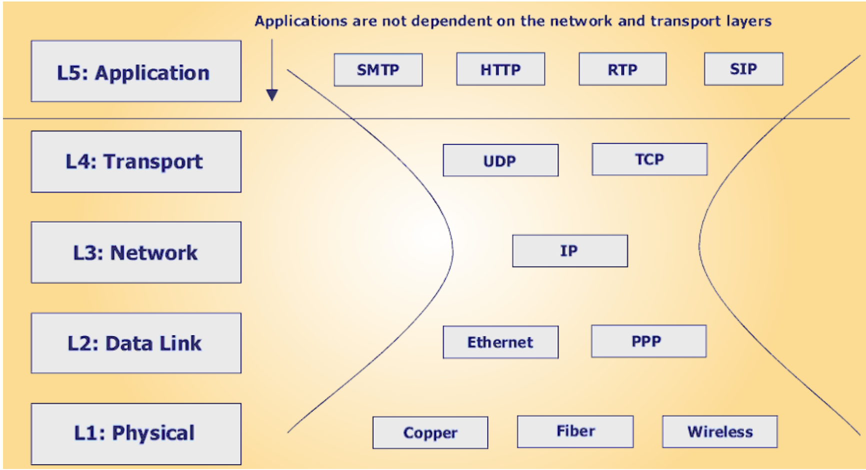
The open Internet, as well as closed IP networks, are constantly being improved and are moving towards higher bandwidths, which will allow for the transfer of high-definition TV pictures and high quality audio to all consumers. It will be possible to use both corporate IP networks and the Internet for audio-only contribution and distribution. The cost of using networks will decrease and the Quality of Service will be improved. NGN (Next Generation Networks) is the common name for improved public networks that will offer considerably higher bandwidth.

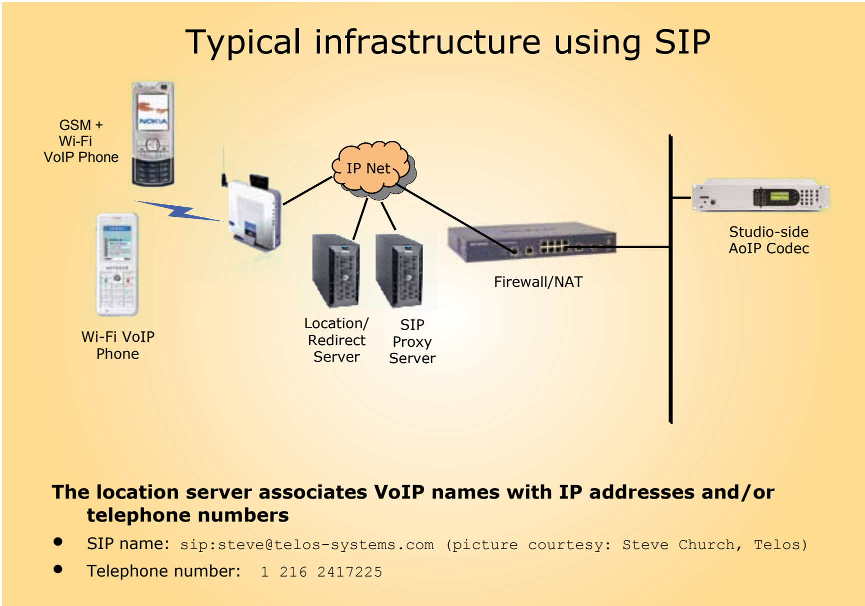
Mobile IP networks will be available with good population coverage in many countries. 3G/UMTS and LTE systems (4G) will soon offer higher bitrates upstream. WiMAX and WLAN hotspots are other possible evolutions which may offer solutions for radio contribution. However, the robustness required for live contribution may not always be guaranteed in these systems because of interference and shared access. The added value of these new networks will be near-instant access for news reporters in all urban areas. The improvements gained by using the presence functionality in SIP (Session Initiation Protocol), will be that the reporter can easily be reached almost anywhere in the world with one identity, irrespective of the platform being used, such as a mobile phone or Audio over IP codec.

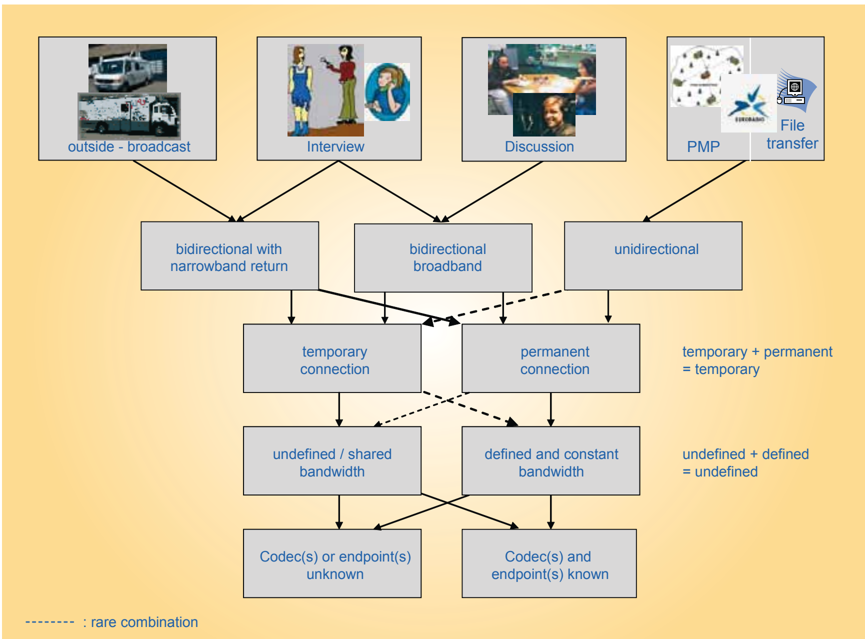
Types of connections

Two types of connections can be used:

- Permanent connections – which are generally based on managed private







networks with constant and well-known bandwidth and Quality of Service. For permanent connections, the audio codec types are usually known in advance.

- Temporary connections – which may be based on previously-unknown networks with shared and unknown bandwidth over the Internet or over temporary leased private networks. The codecs and endpoint may be unknown. The audio codec type can be found through negotiation using the SIP and SDP protocols.

EBU standard

In the past, Audio over IP end units from different manufacturers have not been compatible. Based on an initiative coming from German vendors and broadcasters, the EBU started a project group called N/ACIP (Audio Contribution over IP). One of its tasks is to suggest a method for interoperability. After a joint meeting between broadcasters and manufacturers in September 2007, an interoperability recommendation was published, which manufacturers have already implemented [1].

In a plug-test held at the IRT in Munich in February 2008, nine manufacturers demonstrated interoperability. The EBU is working on a reference implementation software, which can be used to verify compatibility.

The requirements for interoperability are based on the use of RTP over UDP for the audio session and SIP for signalling. The packet payload audio structure is defined by IETF RFC documents for commonly-used audio formats in radio contribution, such as G.722, MPEG Layer II and linear PCM. By using SIP, a negotiation can be made to automatically find a common audio coding system on unknown units at each end.

The EBU standard suggests using RTP over UDP rather than TCP. A one-way RTP stream with a small header (low overhead) is more suitable for audio transfers. Moreover, RTP over UDP sometimes has higher priority than TCP in routers.

Further information on the EBU N/ACIP project can be obtained at : http://www.ebu-acip.org

Types of equipment

Different types of equipment can be identified:

- General contribution equipment – meant for all types of contribution (fixed or remote);

Example of an ISDN and Audio over IP fixed end unit

- Portable contribution equipment – used mainly for monophonic speech contribution at low bi-trates.

Example of a portable ISDN and Audio over IP unit

Networks

Telecom operators are offering an increasing number of IP-based services. Traditional services based on ATM, PSTN, ISDN, SDH and PDH will gradually be phased out or will become expensive niche products. They will tend to be replaced by all-over-IP services carried over copper, fibre or wireless links.

Some EBU members have experience of real-world testing of Audio over IP codecs over various types of IP networks. Determining the most suitable type of network from the solutions offered by providers requires careful evaluation.

In particular, measurements and real-world tests with audio must be made by the broadcaster in order to analyse the service level agreement and to verify the performance offered by the network provider. The network must be tested with the applications that are to be used. Long-term testing (months) of uninterrupted audio throughput is recommended.

The distinction between well-managed IP networks and the open Internet is important. On the open Internet, no mechanisms yet exist to achieve a good QoS. The Internet is a “best effort” network with no guaranteed Quality of Service at all. Over a ten-year period the packet loss, delay and jitter over the Internet have slowly improved, but the network performance still poses a major problem to the developers of Audio over IP units.

Last mile access

There are many access solutions to connect the end user to the Internet or private IP networks. Here’s an overview:

Fibre optics

- This is the highest quality access solution, offering low error rates and low delays. It is ideal for contribution purposes but is still expensive and not widespread.
- Some cities have started to deploy FTTH (Fibre to the Home) access.

Abbreviations

3G	3rd Generation mobile communications
ADSL	Asymmetric Digital Subscriber Line
AoIP	Audio over IP: broadband audio
ATM	Asynchronous Transfer Mode
HSDPA	High-Speed Downlink Packet Access
HSUPA	High-Speed Uplink Packet Access
IETF	Internet Engineering Task Force www.ietf.org
IP	Internet Protocol
ISDN	Integrated Services Digital Network
LTE	Long Term Evolution (4th generation mobile networks)
PCM	Pulse Code Modulation
PDH	Plesiochronous Digital Hierarchy
PPP	Point-to-Point Protocol
PSTN	Public Switched Telephone Network
QoS	Quality of Service
RFC	Request For Comments (IETF standard)
RSTP	Real-Time Streaming Protocol
RTCP	Real-Time Control Protocol
RTP	Real-time Transport Protocol
SAP	Session Announcement Protocol
SDH	Synchronous Digital Hierarchy
SDP	Session Description Protocol
SDSL	Symmetric Digital Subscriber Line
SIP	Session Initiation Protocol
SMTP	Simple Mail Transfer Protocol
TCP	Transmission Control Protocol
UDP	User Datagram Protocol
UMTS	Universal Mobile Telecommunication System
VoIP	Voice-over-IP: narrowband audio
WiMAX	Worldwide interoperability for Mobile Access

Copper with xDSL (ADSL, ADSL2+, VDSL, SDSL)

- This type of access is now in widespread use to make Internet connections. Business providers also have solutions for connecting to private IP networks using xDSL with a guaranteed Quality of Service. This type of solution is pre-ferred to straightforward Internet access for contribution purposes.
- ADSL: Asymmetrical uplink/downlink bitrate.
- SDSL: Symmetrical uplink/downlink bitrate. This type of access is preferred for contribution because of the higher uplink bitrate for sending.

Mobile communication: 3G/ UMTS, HSDPA, WiMAX

- Mobile high bitrate accesses to the Internet are starting to emerge. However the problem is the lack of guaranteed Quality of Service. In many cases the access is shared among users of the same radio cell. So, at the moment, this is not ideal for contribution despite the great advantage of mobility. Many operators also filter the traffic and block access for Voice/Audio over IP.
- Some operators plans to offer access to private IP networks in the future. We may expect that future solutions will have better Quality of Service.

Satellite

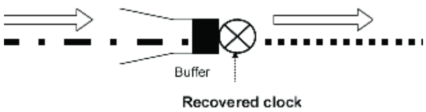
- It is possible to get Internet access through a satellite by using a transmitter system for the return channel. It is used in remote location where all other access technologies are not available. DVB-RCS (Return Channel over Satellite) is used in most cases. The access is generally shared by users and so it is difficult to have a guaranteed Quality of Service. Inmarsat BGAN is also another option.
- It may be possible in the future to have access to private satellite networks with enhanced Quality of Service.
- The delay for the transmission over satellite is generally long (about 500ms roundtrip delay). It is also necessary to have a direct view to the satellite.

Wi-Fi

- Wi-Fi is not really a last mile access but more a home network solution. It is however available as a last mile access in some cities (sometimes for free).
- The frequency band is shared without coordination by many users and also other systems (microwave ovens, DECT telephones). So it is impossible to have guaranteed Quality of Service on such accesses but they may give good results for a local access in not-too-crowded places.

Synchronization

Variations in the delivery time of packets occur, mainly due to varying delays in routers and the sharing of available capacity with other data traffic. Buffering is required to compensate for this variation *(see the diagram)*. There is generally no clock transported so it must be reconstructed at the receiving end. Many different clock recovery algorithms exist. The difficulty is to estimate the slow clock drift correctly and separate it from the short-time network jitter. Streaming audio with high quality is dependant on a guaranteed and stable clock rate at both the sending and receiving ends. Another possibility is to use an external clock source, such as GPS or from a common non-IP network clock.



Delay

Buffers at the receiving end can introduce a considerable amount of delay. The delay buffer size is a trade-off between an acceptable delay and a reliable transmission. In addition, the IP network itself has a delay, from a few tenths of milliseconds in well-managed networks up to 500 ms or more on very long distances over the open Internet. The audio encoding itself may give delays of just a few milliseconds for PCM up to more than hundreds of milliseconds for some bitrate-reduced coding formats.

The time to fill in packets must also be considered: longer packets will mean increased delay, especially at low bitrates. In the case of a two-way conversation, a total round-trip delay which is less than 50 ms is generally preferred, otherwise a conversation becomes difficult, especially when persons from the general public are interviewed. Experienced reporters may be less sensitive to higher values of delay. When using Audio over IP in combination with video contribution, lip sync will be an issue.

Conclusions

The continuous development of IP networks, combined with more sophisticated Audio over IP end units will lead to more use of this technology in the future. The EBU group N/ACIP has made a proposal for interoperability. Connections over the Internet with different types of telephony and professional units for broadcasting will improve telephone audio quality and worldwide access for reporters. Small handheld units and also software codecs in laptops or mobile phones will provide very efficient tools for reporters. SIP will provide a very powerful way of finding the other end, and negotiate a suitable audio coding format. Fixed Audio over IP units will begin to replace older synchronous point-to-point equipment for contribution of stereo or multichannel audio.

First published: 2008-Q1

Lars Jonsson was born in 1949 and received an M.Sc. in Electronic Engineering at the Royal Institute of Technology in Stockholm in 1972. He joined the Research Team of Swedish Radio and, in the early days, worked on the development of video and RF systems. Later he worked in Operation and Development at the Local Radio Company. He rejoined the SR development team in 1992.

During the last decade, Mr Jonsson has worked on digital audio quality issues, archiving and the audio computer infrastructure within Swedish Radio. He is a member of several working parties within the Audio Engineering Society and the EBU.

Lars Jonsson is currently the chairman of the EBU Audio Contribution over IP working group, N/ACIP.

Mathias Coinchon was born in 1975 and graduated in 2000 in communication systems engineering from the Swiss Institute of Technology in Lausanne (EPFL), Switzerland, and the Eurecom Institute in Sophia-Antipolis (France). He developed his diploma thesis at BBC R&D in Kingswood Warren, studying and developing propagation analysis solutions for Digital Radio Mondiale (DRM) field trials. He then joined as technical project manager, a startup company called Wavecall, which is active in the development of a physical propagation prediction tool for the mobile telecommunication industry.

After Wavecall, Mr Coinchon spent four years at RSR public swiss radio, first with responsibility for contribution and then in charge of a group dealing with distribution, contribution and IT networks. During this period, he was also actively involved in the technical study group of the Swiss broadcasting corporation (SRG-SSR idée suisse) for the re-launch of DAB digital radio in Switzerland.

Since 2006 Mathias Coinchon has been a senior engineer in the EBU Technical Department. He is currently secretary of the N/ACIP and N/VCIP groups dealing with audio and video contribution over IP. He is also involved in digital radio matters and is vice-chairman of the WorldDMB technical committee. His other areas of work include audio, distribution, open-source software, IP-based TV studios and traffic information. In his spare time, he is involved in helping a community radio station.

Visit EBU Technical's new website at :
http://tech.ebu.ch

Find EBU Technical - Publications, including EBU Technical Review, at
http://tech.ebu.ch/publications