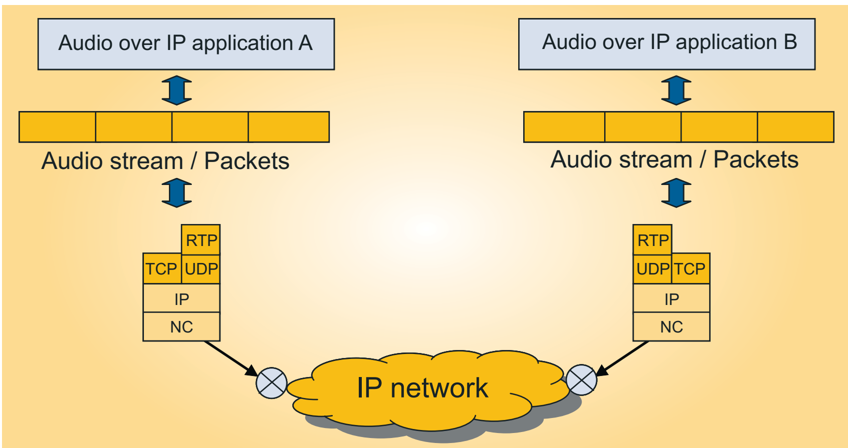
Streaming audio contributions over IP — a new EBU standard

Lars Jonsson
Swedish Radio

Mathias Coinchon
EBU Technical Department

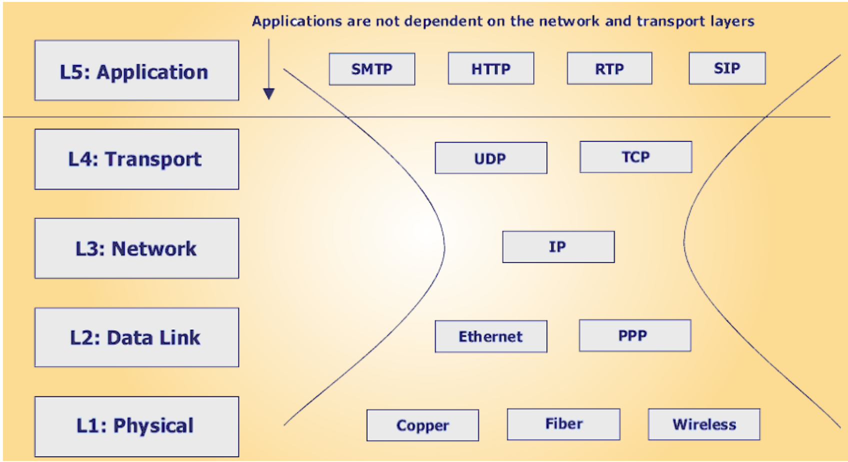
Audio-over-IP end units are increasingly being used in radio operations for the streaming of radio programmes over IP networks, from remote sites or local offices into main studio centres. The IP networks used can be well-managed private networks with controlled Quality of Service. However, the open 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.

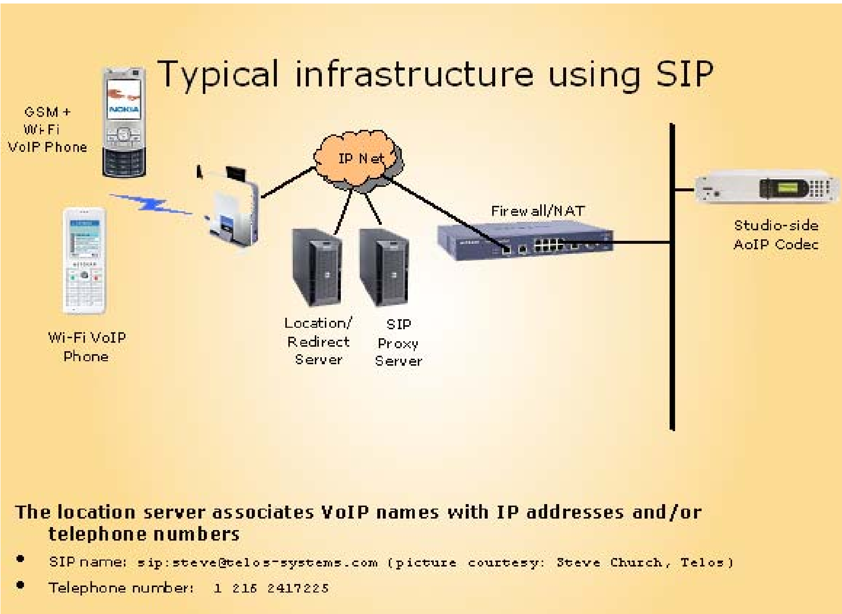


Internet Protocol (IP) is used worldwide on the Internet and also on the IP-based corporate or private networks used by broadcasters. It is independent of the underlying data transmission technology and many IP adaptations exist for the physical layers, e.g. Ethernet, ATM and SDH over copper, fibre or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



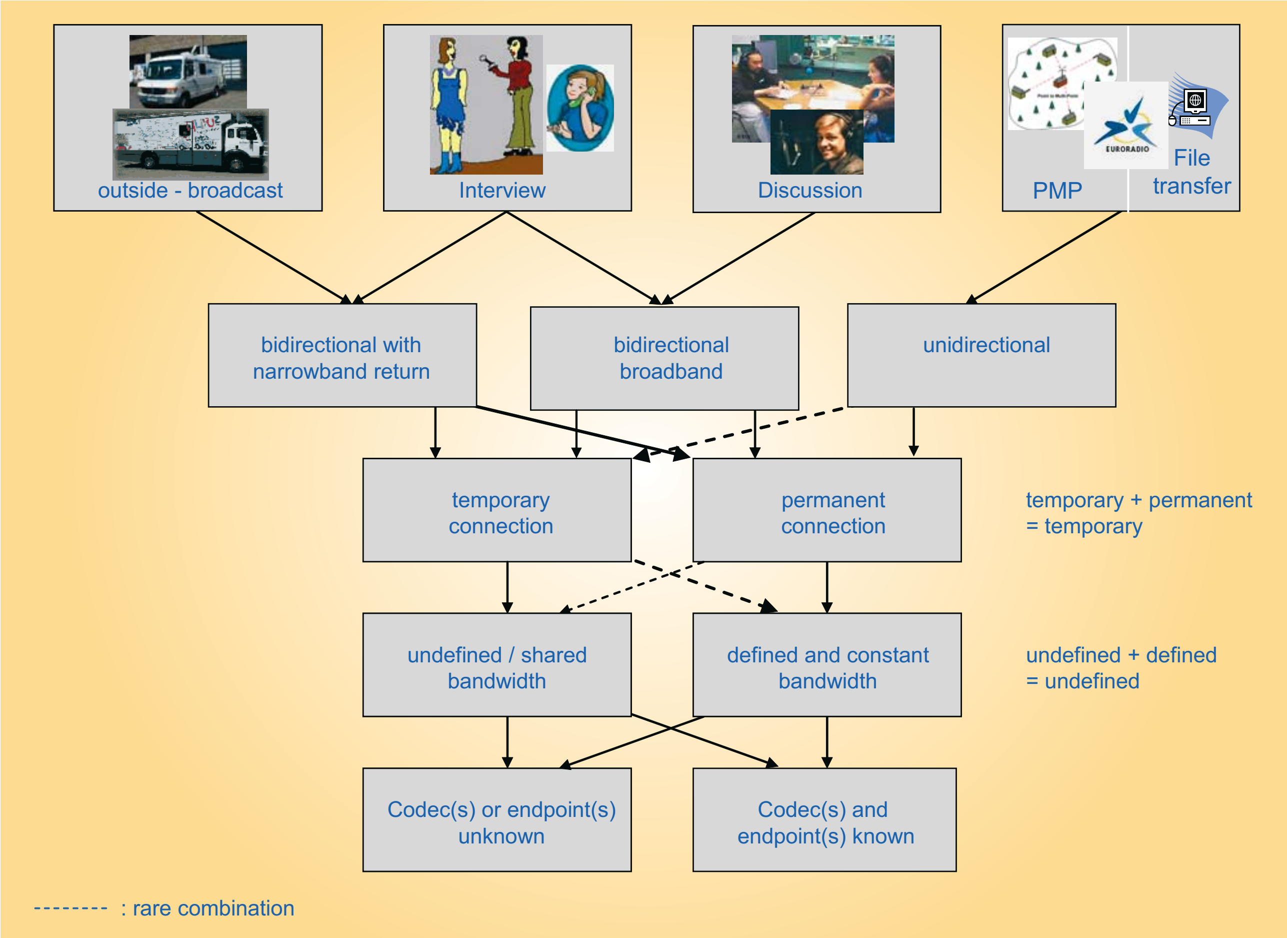
Mobile IP networks will be available with good population coverage in many countries. 3G/ UMTS and LTE systems (4G) will soon offer higher bitrates upstream. WiMAX and WLAN hotspots are other possible evolutions which may offer solutions for radio contribution. However, the robustness required for live 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.



Some of the different types of operations for Audio over IP

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 bitrates

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.

Abbreviations

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

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.

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 preferred 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

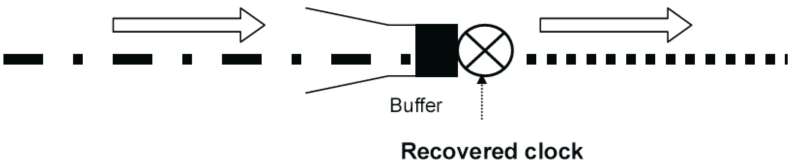
- It is possible to get Internet access through a satellite by using a transmitter system for the return channel. It is used in remote location where all other access technologies are not available. DVB-RCS (Return Channel over Satellite) is used in most cases. The access is generally shared by users 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.

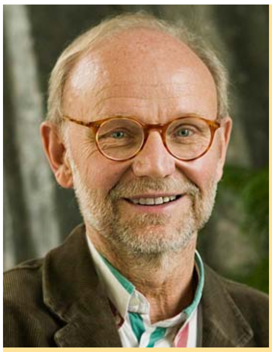
Synchronization

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



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.

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.