

It has been mentioned that sound and vision signals do not take the same amount of time to pass through the CODEC due to the dissimilar delays through the video and audio compression electronics. As a consequence, the sound signals are delayed to synchronise them with the vision. The effect of this delayed sound is to produce a most objectionable echo when conferencing with a remote site. If this echo is not reduced, effective conferencing becomes impossible. The reason that this occurs is explained below.

As shown in Figure C. 1, sound (voices) from the local site A are initially delayed by the CODE part of local CODEC A, then delayed further by the DECODE part of CODEC B at the remote site before being radiated by the loudspeaker to the remote audience. A proportion of these delayed voices from the local site are picked up by the remote site's microphone and fed back via CODE B and DECODE A now even further delayed to the local site's loudspeaker. This sound (now significantly delayed) is then picked up again by microphone A together with the live (un-delayed) voices from A, which generates the characteristic echo. These echoes continue throughout the conference unless precautions are taken to minimise



Network processing equipment other than CODECs can also introduce delays. This is particularly true for IP networks where the signals have to pass through several items of equipment in the transmission chain. The latest audio coding termed MPEG4 AAC-LD has however been specially designed to significantly reduce this transit delay.

## Links

[1] <http://community.ia.net/system/files/images/vc-intro-c1.jpg>