

Voice calls over an ATM-based DSL access network can provide the same quality as calls over today's PSTN.

## Quality of calls transported over a DSL access network

### Introduction

Digital Subscriber Line (DSL) solutions exploit the fact that the local loop (copper pair) has more capacity than is required to transport a single legacy telephone call. Up to now, standard xDSL solutions have focused mainly on data services, such as high speed Internet access. However, the Voice over DSL (VoDSL) solution proposes to take advantage of the high data bandwidth that is available to carry additional voice band lines together with the regular data stream. Telephones connected to these additional voice lines are referred to as VoDSL phones (see *Figure 1*).

One of the main questions with respect to the real-time transport of voice over packet-based (access) networks is whether the same quality can be attained as on today's circuit-switched networks. We examine the subjective quality for calls between a VoDSL phone and a (traditional) phone, as well as for calls between two VoDSL phones. Additionally, for calls between two VoDSL phones, today's scenario in which the two DSL access parts are connected via the Public Switched Telephone Network (PSTN), is compared with a Next Generation

Network (NGN) scenario in which the two access networks are connected via an Asynchronous Transfer Mode (ATM) network. In the latter scenario, the voice payloads are transported over ATM from end to end.

### Reference Network

Today's reference network is shown in *Figure 2*. At the customer premises side, an Integrated Access Device (IAD) multiplexes the data and voice streams over the DSL link using ATM as the transport layer. The data flow is carried as far as a Broadband Remote Access Server (BRAS) which terminates the ATM Virtual Circuit (VC) using ATM Adaptation Layer 5 (AAL5), thus providing access to the Internet. All voice band flows (typically between 2 and 32) are multiplexed on a single dedicated AAL2 ATM virtual connection up to a Voice GateWay (VGW), which terminates the ATM VC and provides access to the PSTN using standard V5 (European Telecommunications Standards Institute; ETSI) or GR303 (American National Standards Institute; ANSI) interfaces. The Loop Emulation Service (LES) defines protocols to

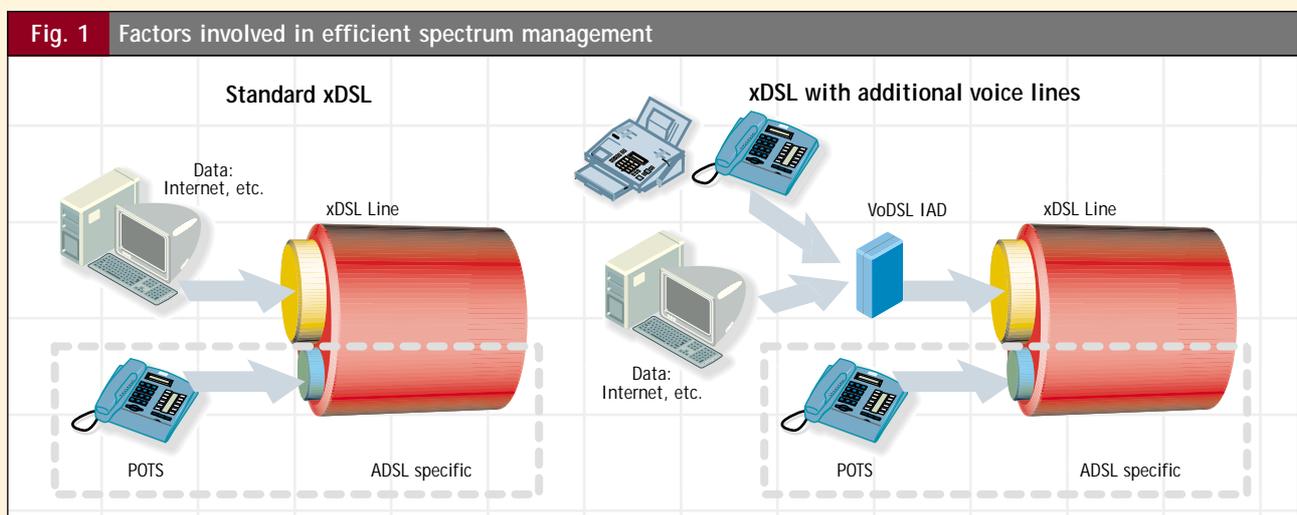
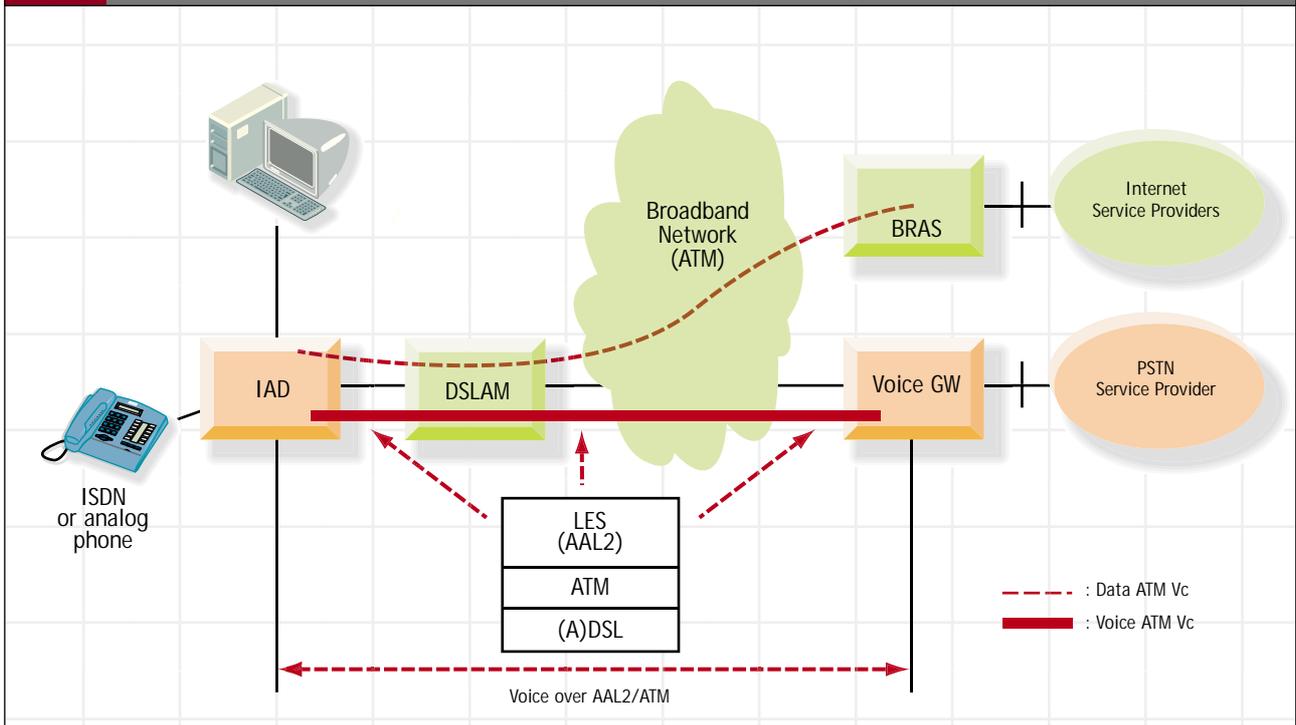


Fig. 2 End-to-end reference network as deployed today



control the AAL2 channels between the IAD and VGW. For calls between two VoDSL phones, we also consider an NGN scenario in which the AAL2 ATM VC is not terminated in a VGW but runs through the ATM network from one IAD to the other. This solution assumes that switched AAL 2 connections are supported on the ATM network and that a signaling gateway in the DSL Access Multiplexer (DSLAM) intercepts and interprets the signaling associated with a voice call. One could, of course, think of a number of other relevant NGN scenarios in which, for example, one party uses a mobile phone for which other codecs have to be considered, or in which part of the transport network is an Internet Protocol (IP) network. Although interesting, these scenarios are outside the scope of this paper.

From a quality point of view, the IAD acts as a VGW. If analog telephone sets are connected to the IAD, then it also performs the conversion from four to two wires in a (4-to-2-wire) hybrid and digitizes the incoming voice signal (i.e. transforms it into the G.711 format).

In one direction, the VGW encodes the G.711 voice signal (if a low bit rate codec is used) and packs the code words into ATM cells using AAL2. In addition, voice activity detection can be used to avoid sending ATM cells when there is no voice signal. In the other direction, the VGW dejitters the received ATM cell stream and decodes the received payloads. The VGWs also perform echo control.

The IAD and the DSLAM (de) multiplex the ATM connections carrying data between the data device (e.g. a PC) and BRAS, and the ATM connection, carrying voice between the VGWs. The latter should have the guarantee of low delay and low jitter. For simplicity, we assume that a Con-

stant Bit Rate (CBR) category service transports the ATM cells carrying voice. However, real-time Variable Bit Rate (rt-VBR) could also be used for voice as long as it meets the low delay, low jitter requirements. The IAD, DSLAM and any other traversed ATM switches give priority to ATM cells containing voice over cells containing data.

### Using the E-Model as an Evaluation Tool

The difference in quality between a VoDSL call and a PSTN call is caused by the Mouth-to-Ear (M2E) delay and distortion introduced during the call.

The E-model [1] is a standardized tool for predicting and quantifying the subjective quality of a telephone call based on its characterizing transmission parameters, such as the codec used, M2E delay and packet loss. The E-model combines the impairments caused by these transmission parameters into a rating  $R$ , which can be used to predict subjective user reactions, such as the Mean Opinion Score (MOS) and maps to certain quality classes defined in [2], as shown in *Table 1*. For a call to be of "traditional quality", its rating  $R$  should be above 70 [3].

Tab. 1 Speech transmission quality categories defined in ITU-T Recommendation G.109

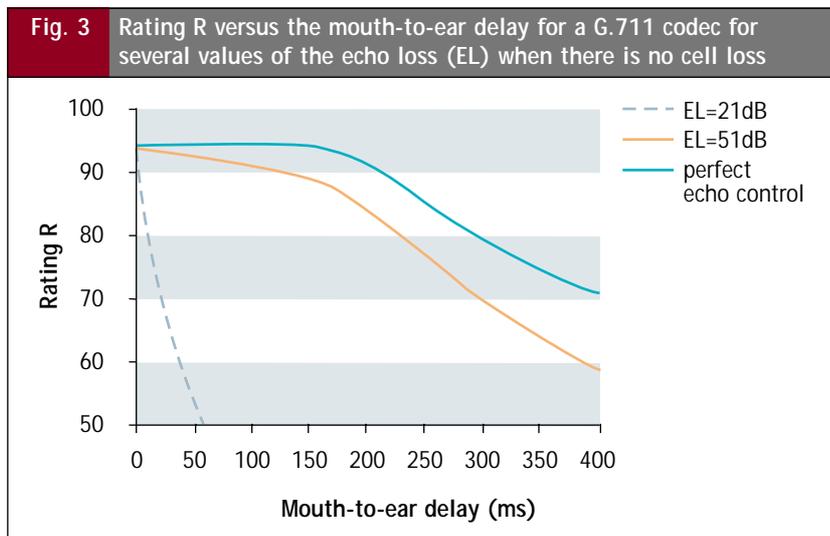
R-value range	100 - 90	90 - 80	80 - 70	70 - 60	60 - 0
Speech transmission quality category	best	high	medium	low	(very) poor

We use the E-model to compare the quality of a PSTN call, a VoDSL call as it works today, and a call between two VoDSL phones over an ATM network. The third scenario is a valid candidate for VoDSL in an NGN. The only E-model parameters that matter for this comparison are the M2E delay and the distortion since they are the only ones that differ between the three scenarios.

### Influence of Mouth-to-Ear Delay

If there is a delay in the voice signal the called party receives from the caller, the rating  $R$  decreases by a certain amount. This impairment is the sum of two contributing impairments: echo and loss of interactivity. Echo may occur in the 4-to-2-wire hybrids in the local PSTN, in analog telephone sets or in the IAD at the customer premises. The amount of echo is expressed by the echo loss, which is the ratio of the level of the original signal to the level of the echoed signal expressed in dB. A typical value for the echo loss in the hybrids of a PSTN is 21 dB [4]. Using an echo controller can increase the echo loss to 51 dB in compliance with the standards [5], or can even remove the echo completely (referred to as perfect echo control).

The best obtainable quality when using the G.711 format corresponds to about  $R = 94$ . Figure 3 shows the influence of the M2E delay on the  $R$  rating for different values of echo loss.



### Influence of Distortion

If the voice signal is distorted, the rating  $R$  decreases (i.e. the  $R$  curves are shifted downwards in Figure 3) by an amount equal to the distortion impairment. Distortion has two origins: encoding of the voice signal and payload loss during the transport of voice payloads. When the G.711 codec is used end to end, there is no codec impairment. Encoding the voice signal into G.726 at 32 kbit/s decreases the quality by 7 points on the  $R$  scale. In today's scenario with VoDSL phones at both ends, however, this encoding occurs twice since the signal has to be transported over the PSTN between the two access networks in G.711 format. In this case, symmetric use of the G.726 codec decreases the quality by

14 points on the  $R$  scale. The NGN scenario does not suffer this burden since, with only an ATM network in between, voice can be transported from end to end using the low bit rate format.

It is assumed that the ATM backbone, the DSL link and the dejittering buffer are configured in such a way that there is little or no cell loss.

## Quality Evaluation

### Mouth-to-Ear Delay Components

The M2E delay, together with echo loss and distortion, determine the subjective quality, that is, the rating  $R$  of the call.

Packetization delay  $T_{pack}$  is defined as the time needed in the ingress VGW to collect the voice information transported in one AAL2 packet. Here we consider AAL2-lite, such that one AAL2 packet fits precisely in the payload of one ATM cell. This gives a  $T_{pack}$  of 5.5 ms for the G.711 codec and a  $T_{pack}$  of 11 ms for the 32 kbit/s version of the G.726 codec.

The encoding delay  $T_{enc}$  is the time needed by the Digital Signal Processor (DSP) in the VGW to digitize the voice signal and encode it from the G.711 format into the low bit rate codec format. Likewise, the decoding delay  $T_{dec}$  is the time needed by the DSP in the VGW to decode

the voice from the low bit rate codec format into G.711 format and possibly to convert it back into an analog signal.

The encoding and decoding delays depend on the DSPs used in the VGWs and the number and priority of the other processes (e.g. echo control) running on the DSPs. For the sum of both,  $T_{DSP}$ , a value of 13 ms was measured when a G.711 codec was used. This rather high value is a result of other processes, such as tone detection, running on the DSP. These are fundamental delays that do not change when the processing speed of the DSP increases because the signal is delayed for a certain time, referred to as the look ahead, to enable algorithms running on the DSP to distinguish between in-band signaling or fax tones and voice.

In the case of a high capacity (155 Mbit/s or more) ATM access network with permanent virtual paths / virtual circuits giving priority to voice over data (with data constituting the bulk of the traffic), it is reasonable to assume that the only bottlenecks that contribute to the queuing delay in the access network  $T_{que,ac}$  are the IAD (upstream) and the DSLAM (downstream). We assume that the VGW in the IAD uses the same  $T_{pack}$  for all voice sources connected to it and that no voice activity detection is used. In this way we can treat all DSL scenarios at the same time: Asymmetric Digital Subscriber Line (ADSL), Single pair High speed Digital Subscriber Line (SHDSL) and differently configured link capacities. The relation

$$T_{que,ac} \leq T_{pack} \quad (1)$$

holds in every stable system. (The serialization delay is included in  $T_{que,ac}$ ; this is a worst case assumption.) In the NGN scenario, the ATM network (with switched VCs) introduces an additional queuing delay  $T_{que,ATM}$ . An upper bound for this delay is agreed upon when the connection is set up (a 2 ms delay has been assumed); the call is blocked when the network cannot meet the required delay and jitter conditions. As is common practice in the PSTN today, the nodes in the NGN should be properly dimensioned in order to keep call blocking probabilities sufficiently small.

Transporting a bit over the DSL link involves a bit pipe delay  $T_{DSL}$ , which is determined mainly by the chosen interleaving depth. It equals about 1 or 2 ms for SHDSL, where there is no interleaving, and for the ADSL fast mode, and 20 ms for the ADSL interleaved mode.

The dejittering delay  $T_{jit}$  is needed in the egress gateway to compensate for the difference in queuing delay of voice payloads. If the first payload of a particular VoDSL call is delayed for more than the maximum queuing delay and the dejittering buffer is read at a constant rate (one payload every  $T_{pack}$  ms), it is guaranteed that every payload will arrive in time in the dejittering buffer. Therefore, if in today's access network:

$$T_{jit} = T_{pack} \quad (2)$$

then no payloads are lost in the dejittering buffer.

The delays discussed above (apart from  $T_{que,ATM}$ ) are introduced in the access network and are counted twice in today's VoDSL phone to VoDSL phone scenario. In the NGN scenario,  $T_{pack}$  and  $T_{DSP}$  occur only once, but  $T_{DSL}$  has to be counted twice. The payloads risk queuing at two multiplexing points in the access: in the IAD (upstream) and in the DSLAM (downstream). Since dejittering is only performed in the NGN at the receiver after the payloads have traversed two access parts plus a (switched) ATM network, Equation 2 becomes:

$$T_{jit} = 2T_{pack} + T_{que,ATM} \quad (3)$$

The final delay component is the propagation delay  $T_{prop}$ . Take 5  $\mu$ s per km for propagation in optical fibers at the speed of light and multiply the distance between the two calling parties by a value between 1.5 and 2, the so-called "engineering factor" which takes account of the fact that the fiber does not follow a straight line. This gives the worst case M2E as:

$$T_{DSP} = 3T_{pack} + T_{DSL} + T_{prop} \quad (4)$$

for VoDSL phone to phone;

$$2(T_{DSP} + 3T_{pack} + T_{DSL}) + T_{prop} \quad (5)$$

between two VoDSL phones in today's architecture; and

$$(3T_{pack} + T_{que,ATM} + T_{DSL}) + (T_{DSP} + 2T_{pack} + T_{DSL}) + (T_{que,ATM} + T_{prop}) \quad (6)$$

between two VoDSL phones when the intermediate PSTN is replaced by an ATM network. This delay is worst case in that:

- The bottleneck link is fully loaded with CBR voice traffic.
- The first payload to arrive is the slowest possible.
- The dejittering delay is chosen equal to its maximum value so that no payloads are lost in the dejittering buffer.

### Quality of Different VoDSL Scenarios

All calls traversing an IAD use the same codec, packetization delay and dejittering delay; the VoDSL phone to VoDSL phone scenarios are symmetric and all scenarios assume standard echo control (an echo loss of at least 51 dB). It is also assumed that the DSL links are configured so that impulsive noise does not cause a noticeable cell loss on the DSL link. This can be achieved by using a conservative setting for the tolerable Bit Error Ratio (BER) or by using the interleaved mode for ADSL if larger noise pulses are expected.

Take  $T_{prop}$  as 5 ms for a local call and 50 ms for a long distance call (about 5000 km; for example, a call from the US to Europe).

First, consider the G.711 codec. For  $T_{pack} = 5.5$  ms, Table 2a gives the M2E delays and R ratings for the dif-

Tab. 2 M2E delay and the rating R for different VoDSL phone to phone and VoDSL phone to VoDSL phone scenarios						
	VoDSL-to-PSTN		VoDSL-PSTN-VoDSL		VoDSL-ATM-VoDSL	
	M2E delay (ms)	rating R	M2E delay (ms)	rating R	M2E delay (ms)	rating R
Local	36.5	93	68	92	53.5	92
Long distance	81.5	91	113	90	98.5	91
(a) G.711 at 64 kbit/s, fast mode						
	VoDSL-to-PSTN		VoDSL-PSTN-VoDSL		VoDSL-ATM-VoDSL	
	M2E delay (ms)	rating R	M2E delay (ms)	rating R	M2E delay (ms)	rating R
Local	54.5	92	104	90	89.5	91
Long distance	99.5	91	149	89	134.5	89
(b) G.711 at 64 kbit/s, interleaved mode						
	VoDSL-to-PSTN		VoDSL-PSTN-VoDSL		VoDSL-ATM-VoDSL	
	M2E delay (ms)	rating R	M2E delay (ms)	rating R	M2E delay (ms)	rating R
Local	53	85	101	77	81	84
Long distance	98	84	146	75	126	83
(c) G.726 at 32 kbit/s, fast mode						
	VoDSL-to-PSTN		VoDSL-PSTN-VoDSL		VoDSL-ATM-VoDSL	
	M2E delay (ms)	rating R	M2E delay (ms)	rating R	M2E delay (ms)	rating R
Local	71	85	137	75	117	83
Long distance	116	83	182	72	162	81
(d) G.726 at 32 kbit/s, interleaved mode						

ferent ADSL fast mode (or SHDSL) scenarios. In every case the 25 ms delay, which is the limit above which echo control is needed, is exceeded. Thus echo control was assumed from the beginning. For all calls, the M2E delay remains below the interactivity bound of 150 ms. “Best” quality is achieved in all cases (see *Table 1*).

With interleaved ADSL (*Table 2b*), the delays are 18 ms and 36 ms longer for the VoDSL phone to phone scenario (denoted VoDSL to PSTN) and for the VoDSL phone to VoDSL phone scenario, respectively. In view of the non-perfect echo control, this results in a slightly lower quality. Next, consider the G.726 codec. *Tables 2c* and *2d* show the resulting ratings and M2E delays. All VoDSL phone to phone calls achieve a “high” quality ( $83.1 < R < 85.2$ ). Even in the worst scenario encountered – VoDSL phone to VoDSL phone with interleaving and a PSTN in between – traditional quality ( $R > 70$ ) can be guaranteed.

Results show that the NGN scenario, in which voice payloads are transported end to end in ATM format, can guarantee a substantially higher voice quality for calls between two VoDSL phones with symmetric use of the G.726 codec in the access parts. This is because it does not involve the transcoding (from G.726 to G.711 and back) that is necessary in today’s intermediate PSTN network. Alternative NGN scenarios that avoid transcoding (e.g. an IP network connecting the two access networks) are, of course, also possible, but are not considered here.

### Conclusion

We have studied the quality of voice calls switched over a DSL/ATM access network to either the PSTN or an ATM network (both referred to as VoDSL calls). All scenarios assume standard echo control (an echo loss of at least 51dB [5]) in all traversed VGWs.

With respect to quality, DSL/ATM access introduces more delay than traditional PSTN access. It is assumed that there is negligible cell loss in the ATM network, the DSL link and the dejittering buffer, so distortion only depends on the low bit rate codec used.

The quality in various VoDSL scenarios was compared with standard reference quality classes, as well as with today’s PSTN quality. The voice quality evaluation found the following:

- It is necessary to deploy echo controllers complying with ITU-T Recommendation G.168 in the VGWs in both the IAD and the central office.
- As long as standard echo control is implemented, the mouth-to-ear delay stays well below a tolerable limit.
- If the same low bit rate codec is used in both access parts for a call between two VoDSL phones, then the NGN scenario in which voice is transported end to end in ATM format can guarantee a better voice quality.

In all scenarios, results show that it is possible to guarantee the traditional PSTN quality. ■

### References

1. “The E-model, a computational model for use in transmission planning”, *ITU-T Recommendation G.107*, December 1998.
2. “Definition of categories of speech transmission quality”, *ITU-T Recommendation G.109*, September 1999.
3. D. De Vleeschauwer, J. Janssen, G. H. Petit, F. Poppe, “Quality Bounds for Packetized Voice Transport”, *Alcatel Telecommunications Review*, 1st Quarter 2000, pp 19–23, January 2000.
4. “Speech processing, Transmission and Quality aspects (STQ): Overall transmission plan aspects for telephony in a private network”, *ETSI EG 201 050*, February 1999.
5. “Digital network echo cancelers”, *ITU-T Recommendation G.168*, April 1997.



**Maarten J.C. Büchli** is a research engineer participating in the QoS, Traffic and Routing Strategies project within the Network Architecture team of the Alcatel Network Strategy Group in Antwerp, Belgium



**Danny De Vleeschauwer** is a research engineer participating in the QoS, Traffic and Routing Strategies project within the Network Architecture team of the Alcatel Network Strategy Group in Antwerp, Belgium.



**Jan Janssen** is a research engineer participating in the QoS, Traffic and Routing Strategies project within the Network Architecture team of the Alcatel Network Strategy Group in Antwerp, Belgium.



**Guido H. Petit** is Director of the Network Architecture team of the Alcatel Network Strategy Group in Antwerp, Belgium.



**Annelies Van Moffaert** is a research engineer participating in the QoS, Traffic and Routing Strategies project within the Network Architecture team of the Alcatel Network Strategy Group in Antwerp, Belgium.



Security