Mastering Voice over DSL: Network Architecture

A TECHNOLOGY WHITE PAPER

CopperCom

Executive Summary

This paper is the second in a series of technical publications on the topic of Voice over DSL. It is intended for network planners and technologists who are responsible for the planning and design of local broadband and voice access networks. The paper follows up the market requirements for voice over DSL that were described in the previous publication, and offers a detailed discussion of solutions for the delivery of multiple lines of local telephone service over Digital Subscriber Line connections, focusing on network architectures and the transport of voice and signaling.

About the Author

Martin Taylor is chief technology officer at CopperCom, a privately held venture-funded startup based in Santa Clara, CA. Prior to joining CopperCom in March 1998, he was vice president of network architecture for Madge Networks, and responsible for strategic planning of products and technology. While at Madge, he developed the MadgeOne architecture, a solution for integrating voice, data and video communications over a common LAN infrastructure which employed a call control server and distributed switching in place of the traditional PBX. Before joining Madge Networks, he designed passive optical network systems and LAN/PBX integration solutions at GEC-Plessey Telecommunications.

Contents

1 Introduction

The 1996 Telecommunications Act has had a profound impact on the evolution of telecom technology, and nowhere more so than in the local access network. For some years now, consumers and businesses alike have benefited hugely from competition in long distance and wireless networks. But for local access, residential users and small businesses are still largely reliant on monopoly service providers and traditional voice service delivery technologies. With the advent of new voice over DSL technology, all that is about to change.

DSL uses existing copper local loop cabling to bring cost-effective broadband communications to residential and small business customers. This bandwidth provides access to local packet networks that efficiently handle both voice and data. With appropriate equipment on each end of this packet network, multiple lines of local voice access can be delivered inexpensively alongside high speed data on a single copper pair. That's voice over DSL or VoDSL for short.

This paper focuses on technical architectures and solutions for deploying VoDSL to bring bundled or integrated voice and data services to small business and residential customers.

2 Market Requirements for Voice over DSL

The market requirements that shape the technical characteristics of VoDSL solutions have been covered in detail in a previous publication from CopperCom – "Complete DSL: Requirements for Public Multi-line Telephone Service Delivery over the DSL Access Network."

These requirements can be summarized as follows:

True Telephony – the local phone service delivered by VoDSL must be indistinguishable from a regular POTS service delivered over a conventional copper loop in all respects. The human voice must be clearly reproduced without discernible echo or delay, and the service must support the connection of all conventional phone instruments including tone and pulse dial handsets. The service must also support the connection of local switching equipment such as key systems and PBXs, and offer full performance for dial-up data communications devices including fax machines and modems. All CLASS, Custom Calling and Centrex features supported by the public network must be supported transparently over VoDSL.

Seamless Integration – the VoDSL network solution must leverage existing investments in DSL Access Multiplexers (DSLAMs), the packet or cell switching infrastructure that supports these DSLAMs, and Class 5 switches that provide voice service. The elements of the VoDSL network solution must interoperate with existing interfaces on these devices.

Flexible and Efficient – VoDSL technology must make efficient use of the limited bandwidth available on DSL connections to customers, and employ appropriate techniques including voice compression, silence removal and dynamic bandwidth allocation to maximize the value of the available bandwidth.

Easy to Provision and Manage – the solution for managing the VoDSL network must integrate cleanly into existing Operations, Administration and Maintenance systems without disrupting the well-established models for the provisioning of voice and data services.

99.999% Reliable – though based on the use of data networking technology to transport voice, the VoDSL solution must take into account the very high levels of availability that are expected and demanded of voice services.

3 Network Topologies for VoDSL

A generic network topology for VoDSL is illustrated in figure 1. DSL service is delivered over conventional copper loops from DSL Access Multiplexers (DSLAMs) in the Central Office. For those customers who receive only data services over DSL, these loops are terminated at the customer premises with a DSL modem or router. For combined voice and data services, the DSL loop is terminated typically by a device that provides integrated voice and data access. Such devices typically offer an Ethernet port for data and multiple analog POTS ports for voice.

The DSL connection to the customer makes use of a packet protocol such as ATM or frame relay to support voice and data. The DSLAM serves as a packet concentrator, delivering traffic from multiple customers over a high-speed uplink to a metropolitan or regional packet network. The principle data service that is offered to DSL customers is Internet access, so the packet network is connected to the Internet, typically through a device known as a Subscriber Management System. Connections to enterprise data networks may also be present, to support telecommuters and home-based workers.

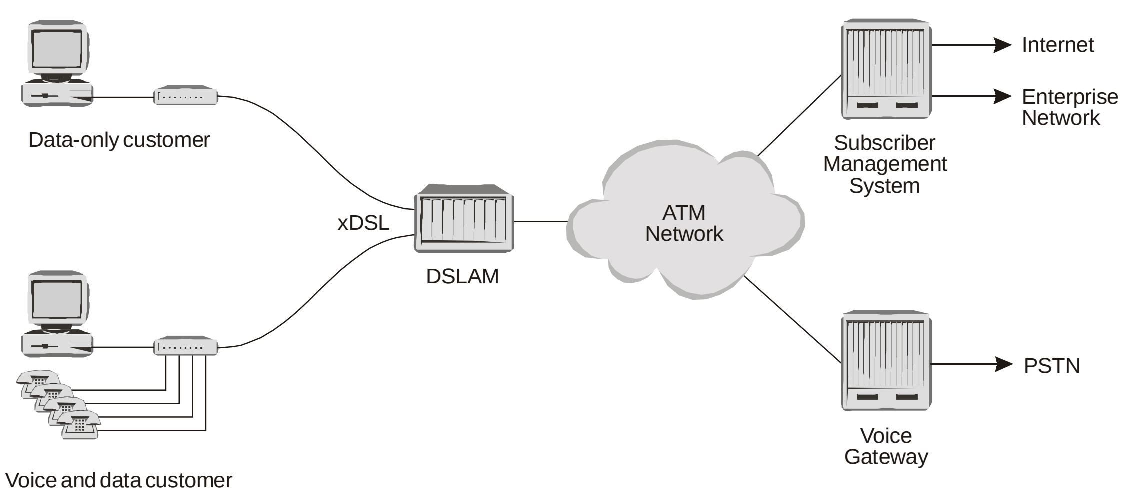

Figure 1: Generic Architecture for Voice over DSL

Voice services are delivered to DSL customers by means of a voice gateway which connects the public switched telephone network (PSTN) to the packet network. Digital voice streams are converted into packet format for transport over the packet network between the voice gateway and the integrated access device on the customer premises. The voice gateway connects to the PSTN via a Class 5 switch. Since the voice gateway represents a digital access network from the point of view of the Class 5 switch, the connection between the gateway and the Class 5 switch typically makes use of a standard interface for digital loop carrier systems, such as GR-303, TR-008 or V5.

The generic network topology just described can be implemented in a variety of forms, to suit the specific requirements of service delivery. The two main architectural variants, the "centralized" and the "distributed" architectures, meet the needs of different kinds of service providers by locating the voice gateway optimally in relation to the other network elements.

3.1 Centralized Architecture

The centralized architecture (see figure 2) meets the needs of service providers who wish to deliver voice services from a centrally located Class 5 switch via DSLAMs in multiple different Central Offices. This is normally the case with competitive local exchange carriers (CLECs) who provide DSL service from DSLAMs which are installed in collocation cages in an incumbent's Central Offices.

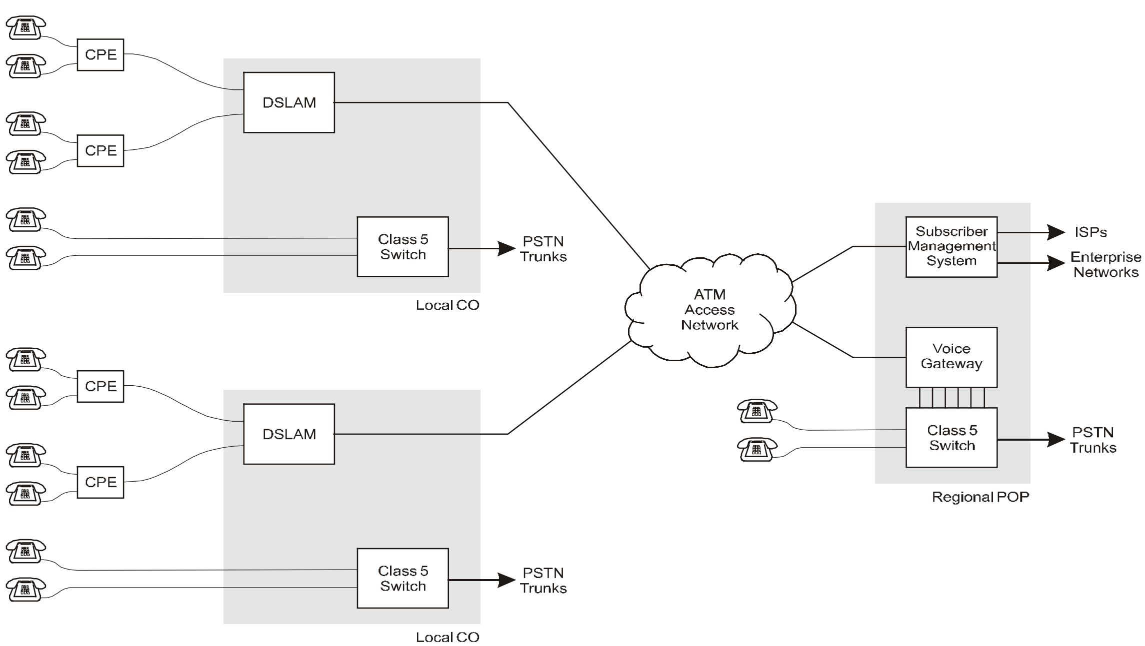

Figure 2: Centralized architecture for Voice over DSL

In this case, the packet network which aggregates traffic from multiple DSLAMs is used to back-haul both voice and data traffic to one or more locations where connections are made to the Internet and to the PSTN.

The benefit of this approach is that a single Class 5 switch can serve the voice needs of customers who are spread over a large metropolitan area. The packet network concentrates voice traffic from multiple different DSLAMs into a single gateway connected to the PSTN, which enables the service provider to offer voice services over an extensive geographic area with a modest initial investment, even in the early phases of deployment when market penetration is low.

3.2 Distributed Architecture

The distributed architecture meets the needs of service providers who own both DSLAMs and Class 5 switches in the same Central Offices, and who wish to drop off voice traffic from the DSL network locally in each Central Office. This architecture is appropriate for incumbent local exchange carriers (ILECs) who wish to serve their VoDSL customers from the same Class 5 switch as they would use to provide POTS services to these customers.

In the distributed architecture, there is a requirement to separate voice traffic from data traffic between the DSLAM and the regional packet network. Data traffic will be passed over the packet network to a central location where connections into the Internet are made, while voice will be handed off locally to the Class 5 switch.

The separation of voice and data packets requires a packet switching function; however the installation of a separate packet switching device in each CO to accomplish this is generally regarded as highly undesirable.

There are two different solutions to this architectural problem. The first solution is to use an enhanced voice gateway that provides “data pass through” functionality, as illustrated in figure 3(a). This type of gateway connects to both the DSLAM uplink and the regional packet network. Packets coming from the DSLAM are examined to see if they contain voice or data. Voice packets are dropped off inside the gateway and converted to circuit traffic for connection to the Class 5 switch, while data packets are passed on to the regional data network. In the reverse direction, data packets arriving from the regional packet network are merged with voice packets generated from PSTN circuits, and the combined packet stream is sent to the DSLAM.

The second solution is to use a “switching DSLAM” that offers two or more high-speed uplinks, as illustrated in figure 3(b). The packet connections in the DSLAM are configured to direct data packets out of one uplink to the regional packet network, while voice packets are directed out of the other uplink into a voice gateway for handoff to the Class 5 switch.

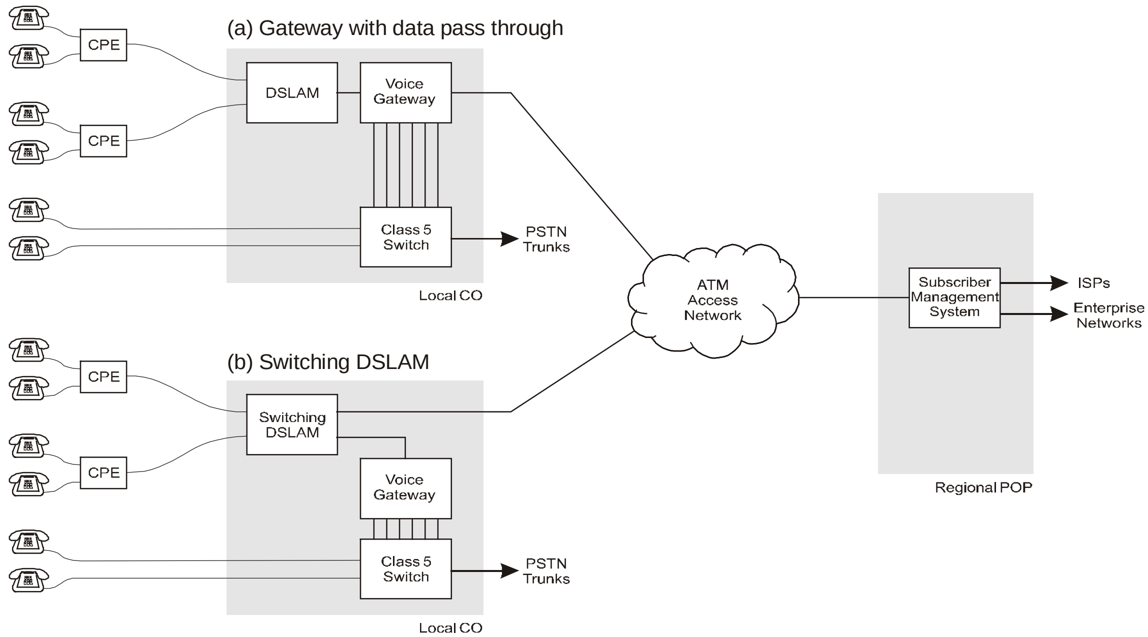

Figure 3: Distributed architecture for Voice over DSL

4 Packet Protocols for DSL Access Networks

In the discussion so far, we have referred to "packet networks" without defining what kind of packet we are talking about. In the context of DSL access networks, the term "packet" could refer to an ATM cell, a frame relay frame, or an Internet Protocol (IP) packet. IP packets are typically carried as payload within ATM cells or frame relay frames. To further confuse the picture, packet switching within the regional packet network may be carried out at the ATM or frame relay layer by means of cell switches or frame switches, or it may be carried out at the IP layer by means of routers. And to make matters even more complicated, frame relay networks may be interconnected with ATM networks via frame-to-cell interworking functions.

The delivery of combined voice and data services over DSL is implicit in the concept of VoDSL. The data services that are delivered over DSL are almost exclusively IP-based, so IP is universally supported by DSL access networks. But for VoDSL, the question that needs to be answered is: what kind of packet should be used for voice transport?

Before answering this, we need to look in more detail at how existing DSL access networks are constructed. And following the principles of one of the key market requirements identified above, which states that VoDSL solutions should overlay and not displace existing access network architectures, we need to identify a solution for packet voice that is compatible with current practice.

4.1 Current Practice in DSL Access Networks

4.1.1 ILEC Architectures for DSL Access

The vast majority of ILEC deployments of DSL use Asymmetric DSL (ADSL) based on the T1.413 standard, supporting ATM cell transport. The preferred protocol stack for data traffic is IP over PPP over ATM, although other methods such as RFC1483 bridging and routing may also be supported. ADSL DSLAMs perform ATM-based traffic concentration and support ATM-based uplinks. These uplinks may be connected into a regional ATM network which provides further traffic aggregation, or they may terminate directly in a Subscriber Management System which supports IP-based connections to Internet Service Providers.

Many ILECs plan to deploy G.lite in the future, alongside ADSL. G.lite is a variant of ADSL that promises to offer a lower cost solution for the mass market, and like ADSL, it will be deployed to support an ATM cell transport. G.lite can be accommodated within the existing architectural framework that supports ADSL, and no changes are necessary.

ILEC DSL access networks could support voice over ATM (VoATM), or voice over IP (VoIP). The requirement to separate voice and data traffic between the DSLAM and the regional packet network to support the centralized architecture suggests that, whichever method is used, the voice and data should travel on different ATM virtual circuit connections (VCCs). Voice and data could travel on the same VCC if VoIP were to be used, but this would imply that either the DSLAM or the gateway would have to reassemble all ATM cells into IP packets to determine which packets contain voice and which contain data. This process would add substantially to the cost and complexity of the solution.

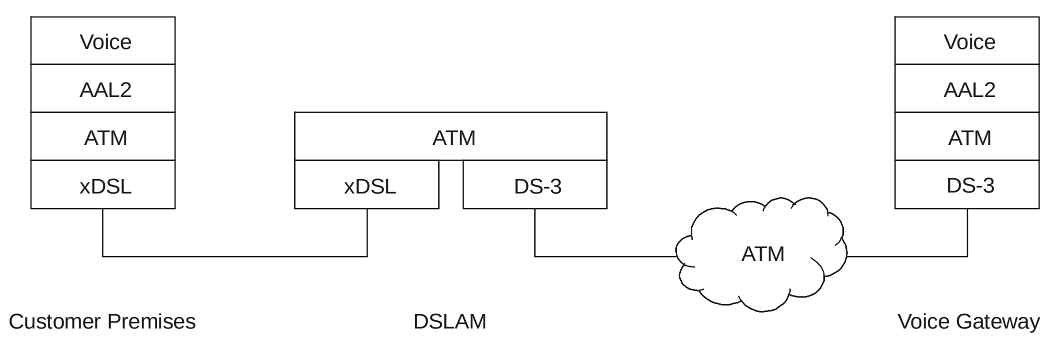

Figure 4: Protocol stacks for Voice over ATM AAL2

4.1.2 CLEC Architectures for DSL Access

CLEC networks tend to show a greater variety of different architectures than those of the ILECs.

Currently, the majority of CLEC DSL access networks use Symmetric DSL (SDSL) DSLAMs that support frame relay transport over DSL. The great majority of these DSLAMs are connected to regional ATM networks via frame-to-cell interworking functions, which may exist either in the DSLAM itself, or in the ATM switch that is providing edge access to the DSLAM.

Since IP packets pass transparently through the frame-to-cell interworking function, VoIP appears to offer a straightforward solution for voice transport in these kinds of networks. Alternative techniques are, superficially anyway, more complex, because the customer premises device must support frame relay while the centralized voice gateway is connected via ATM. This implies that voice over frame relay (VoFR) is needed on the DSL connection, and that the frame-to-cell interworking function must map frames containing voice into cells containing voice so as to be compatible with the voice gateway. However, while VoIP may appear to be simpler, it has some serious drawbacks for voice over DSL, as we'll see later.

By no means all CLECs use frame-based DSLAMs. A small but growing number of CLECs are deploying ATM-based DSLAMs, which are of course interconnected by regional ATM networks. ATM-based DSLAMs are sometimes a little more costly than frame-based equivalents, but the difference may be more than offset by the advantages that ATM-based DSLAMS can offer when voice traffic is to be supported.

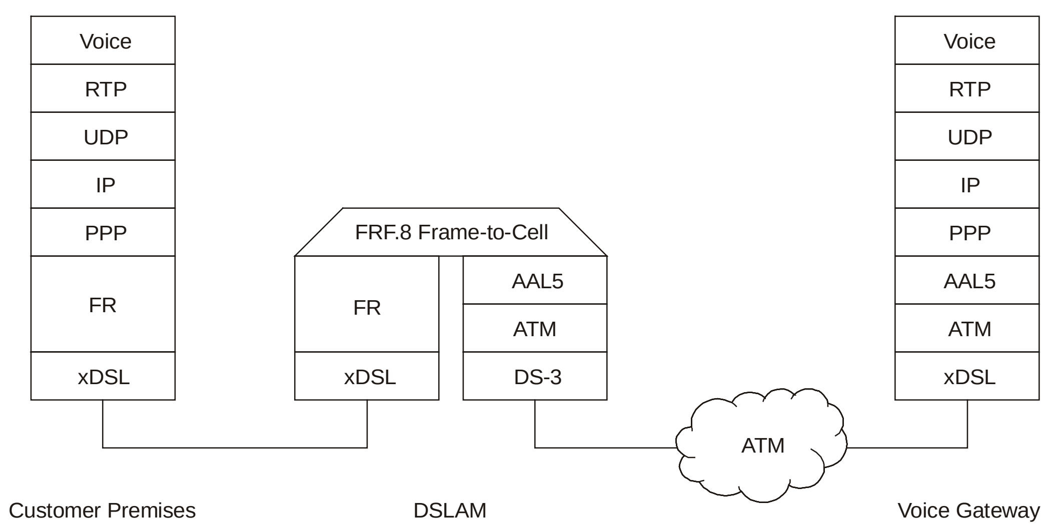

Figure 5: Protocol stacks for Voice over IP with frame-based DSLAM

4.2 Voice Over ATM Solutions

4.2.1 Circuit Emulation

Traditional solutions for voice over ATM have been based on a technique known as "Circuit Emulation," where a byte-interleaved frame structure such as T1 or E1 is converted into a stream of ATM cells that are transmitted at a constant rate. Circuit emulation is based on the use of the ATM Adaptation Layer 1 (AAL1) protocol, and is widely used today for the bulk transport of circuits such as T1 and T3 over ATM networks.

Circuit emulation does meet one of the basic requirements of VoDSL, which is that multiple voice channels share the same ATM VCC, but it fails to meet some other key requirements, in particular:

- Support for a mix of compressed voice traffic and uncompressed, clear channel traffic simultaneously on the same VCC. This is necessary to support the efficient transport of voice in compressed form over the bandwidth-constrained DSL connection, while at the same time supporting voiceband data traffic from fax machines and modems at full speed.
- Support for silence removal, a technique that offers bandwidth savings of up to 50% in the core of the regional packet network.

Another requirement of VoDSL, namely support for dynamic bandwidth allocation between data and voice, can be supported using circuit emulation techniques, but the technology required to implement it is very complex.

4.2.2 Loop Emulation

A better solution for VoDSL may be found in the form of ATM Adaptation Layer 2 (AAL2), a newer technique that was designed to support the multiplexing of real-time packet streams over an ATM VCC. AAL2 defines a packet header for voice packets (or other kinds of real-time media such as video) which contains a length indication, a channel identifier and a means of identifying what kind of voice encoding the packet contains. The facilities provided by AAL2 offer an excellent basis for a VoDSL solution.

The details of AAL2 are set out in the ITU-T specification I.363.2, which was published in September 1997. This specification covers the AAL2 packet format, and the method by which a stream of AAL2 packets may be assembled into ATM cells. Also relevant to VoDSL solutions are the ITU-T specifications I.366.1 and I.366.2. The I.366.2 document details the packet formats needed for transport of voice and call control signaling using AAL2, while I.366.1 describes a method for carrying data packets (containing, for example, network management protocols) over the same AAL2 VCC as the voice packets.

In April 1999, the ATM Forum recognized the importance of VoDSL as an application of voice over ATM technology, and approved a new work item entitled "Loop Emulation Service Using AAL2" to specify how these ITU-T standards should be applied to support the delivery of voice over DSL. Approval of this specification is expected early in 2000.

4.3 Voice over Frame Solutions

As discussed earlier, the VoDSL solution for customers served by frame-based DSLAMs is complicated by the fact that most such DSLAMs are connected to ATM networks via frame-to-cell interworking functions.

At the customer end, voice must be inserted into frame relay frames for transmission on the DSL connection. One approach for achieving this might be the Frame Relay Forum specification for voice over frame relay, FRF.12.

Frame-to-cell interworking is defined by the Frame Relay Forum specification FRF.8, which describes how a frame relay payload is mapped to an ATM AAL5 payload. AAL5 is the normal ATM adaptation layer for packet-mode data.

If we were to make use of these established processes, then the voice traffic in the ATM network would be carried in the FRF.12 format as AAL5 payload, an entirely different format from that which was described for the pure voice over ATM solution above. There would therefore be major differences between voice gateways designed to handle the two different approaches.

An alternative approach to this problem would be to use the AAL2 packet format for the transport of voice, and to carry each AAL2 packet as a frame relay payload. At the FRF.8 frame-to-cell interworking function, the AAL2 packet would be mapped into one or more ATM cells, with the AAL5 protocol trailer attached to the end. If the AAL2 packet size is chosen appropriately, then the AAL2 packet with its AAL5 protocol trailer fits exactly into one ATM cell. From the point of view of a voice gateway designed for voice over ATM AAL2, this looks like a familiar packet format, the only difference being the presence of the AAL5 protocol trailer. Provided that the voice gateway can be configured to ignore this AAL5 trailer attached to each voice packet, the same gateway can be used for both voice over end-to-end ATM networks and those that use voice over frame on the DSL connection.

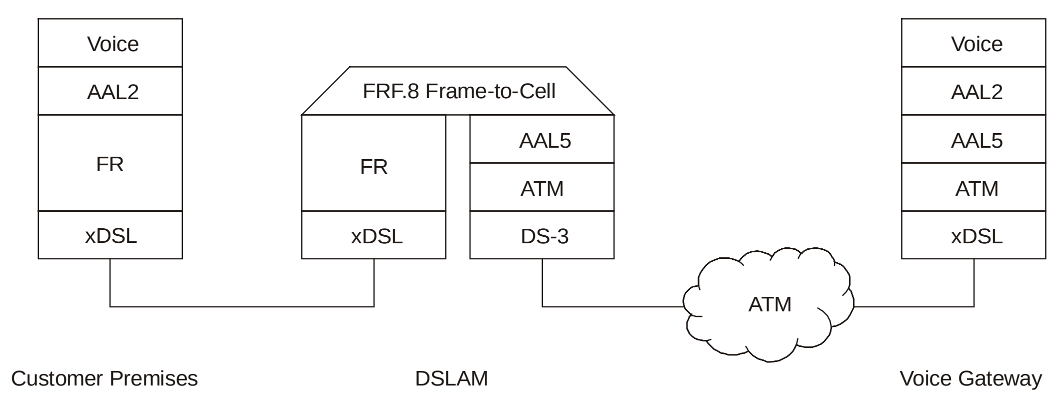

Figure 6: Protocol stacks for Voice over AAL2 with frame-based DSLAM

4.4 Voice Over IP Solutions

Since all real DSL access networks support IP traffic, regardless of whether they are based on ATM or frame relay over DSL, it would appear that VoIP can offer a single universal solution for all VoDSL needs. However, as we shall see, there are some issues with VoIP which have led to the emergence of two distinctly different approaches to IP voice transport that are relevant to the needs of VoDSL: the Internet telephony approach and the IP trunking approach.

4.4.1 The Internet Telephony Protocol Stack

The most widely accepted protocol stack for Internet telephony is based on the Real-Time Protocol (RTP) running over the User Datagram Protocol (UDP) over IP.

RTP provides identification of the voice encoding format within the voice packet, together with a sequence number and timestamp to enable a synchronous voice stream to be re-constructed from a stream of RTP packets. UDP provides a port number which enables multiple voice streams to be multiplexed between two IP end-points. IP, of course, provides the source and destination address that enables the IP network to switch packets from one end-point to another.

The Internet telephony protocol stack is widely understood and implemented, and its use as the basis of a VoDSL solution would have the advantage of direct compatibility of packet format between the access network and future IP-based long distance telephony networks. These advantages, combined with the ability of VoIP to operate over any kind of DSL access network, ought to make it the clear choice for VoDSL. However, there are some major difficulties to be overcome, one of which is bandwidth efficiency. We'll deal with that issue here, and cover some of the others in a later section.

Bandwidth efficiency is a major concern for VoDSL. One of the clear market requirements for VoDSL solutions is to make efficient use of the limited bandwidth available on DSL connections so as to deliver as many voice circuits as is practical, while meeting the objective of toll quality voice. Unfortunately, the Internet telephony protocol stack does not fare well when measured against this market requirement.

The issue here is that of protocol overhead. To stay within acceptable limits for access network transmission delay, packet voice solutions have to utilize small packets that contain no more than around 15-20 ms of encoded voice. If we assume that voice is compressed 2:1 using the ADPCM 32 kbps algorithm, the maximum packet size we could use would be 80 bytes, which is equivalent to 20 ms of encoded voice. The RTP, UDP and IP protocol headers add respectively 12, 8 and 20 bytes to this. For transmission over ATM, as would be needed for all ATM-based DSL access networks, a PPP header (2 bytes) and an LLC header (4 bytes) need to be added. The 8-byte AAL5 trailer is then added to the end of the packet, which is then padded so that the entire packet is an integral number of 48-byte cell payloads for insertion in ATM cells. In the case of the 80-byte voice packet we have described, the total length before padding is calculated as follows:

Content	Bytes	
-----------	-------	
LLC header	4	
PPP header	2	
IP header	20	
UDP header	8	
RTP header	12	
Voice payload	80	
AAL5 trailer	8	

Sub-total	134	
Add AAL5 padding	10	

Total ATM payload	144	(3 x 48)
Cell headers	15	(3 x 5)

Grand total	159	

Thus to transmit 80 bytes of voice payload it is necessary to transmit a total of 159 bytes including ATM cell overhead. This is equivalent to a bandwidth efficiency of almost exactly 50%. A single channel carrying 80-byte voice packets encoded with ADCPM 32 kbps would consume 64 kbps on the DSL connection. By contrast, voice over ATM AAL2 using a 44-byte voice payload size consumes less than 40 kbps per voice channel, and has a packetization delay of only 11 ms compared with 20 ms for the Internet telephony example illustrated. When a 44-byte voice packet is used with the Internet telephony protocol stack, efficiency falls to around 40%.

4.4.2 IP Trunking

Protocol overhead is widely acknowledged as an issue for the IP telephony protocol stack, and a number of solutions have been proposed to improve bandwidth efficiency for VoIP. In the VoDSL environment, one of these solutions in particular can be applied effectively, taking advantage of the fact that the path between the integrated access device at the customer premises and the voice gateway connecting to the PSTN is a point-to-point connection that is required to carry multiple voice channels. This solution is known variously as IP trunking or RTP multiplexing, and it is described in the Internet Draft draft-ietf-avt-multiplexing-rtp-00.txt.

With IP trunking, each IP packet contains both UDP and RTP as for Internet telephony. But in this case, the RTP payload can contain multiple voice packets, one from each active channel. Each voice packet is preceded by a 2-byte “mini-header” which identifies the length of the packet, and which channel it belongs to.

If only one voice channel is active, then the protocol overhead of IP trunking is nearly the same as if the Internet telephony stack were to be used. When multiple voice channels are active, bandwidth efficiency improves as the protocol overhead is shared between the active channels.

The efficiency that is achieved varies according to whether IP trunking is running over frame-based DSL or ATM-based DSL. For typical packet sizes used in VoDSL applications, the bandwidth efficiency of IP trunking over frame relay is approximately the same as that of voice over AAL2 when five channels are active. The efficiency of IP trunking over ATM-based DSL is compared with that of AAL2 in figure 7, where it can be seen that IP trunking approaches, but does not exceed the efficiency of AAL2 as the number of active channels increases towards 16.

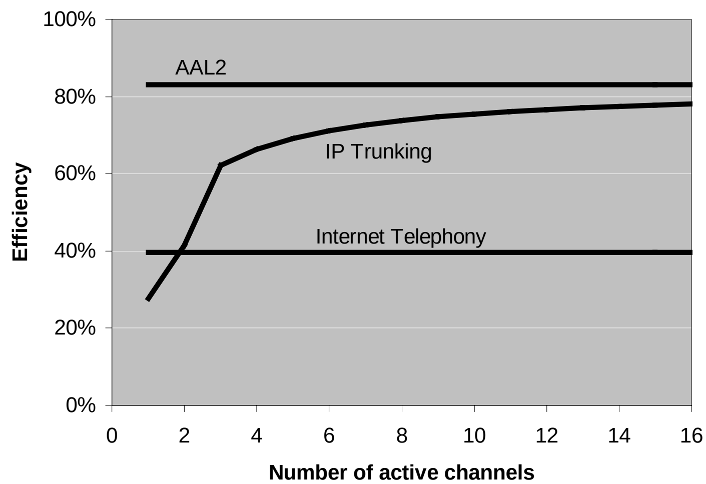

Figure 7: Bandwidth efficiency of Voice over IP and AAL2 on ATM-based DSL (44-byte voice packets)

4.5 Voice over IP vs. Voice over ATM or Frame

The discussion above has focused on the issue of bandwidth efficiency as it relates to VoIP, and has shown that the IP trunking technique offers major improvements over the Internet telephony protocol stack. But bandwidth efficiency is by no means the only issue for VoIP in the context of voice over DSL.

There are three other major areas of concern for VoIP: security, performance and administration.

Before getting into a detailed discussion of these issues, it is necessary to point out that VoIP can be carried over the cell- or frame-based DSL connection in two different ways. The simple way is to transport all IP packets, regardless of whether they are carrying voice or data, over the same ATM or frame VCC. The alternative way is to set up separate VCCs for voice and data packets.

Let’s deal with the simple way first.

All IP traffic, regardless of whether it contains voice or data, is carried between the integrated access device at the customer premises and the IP network over a single VCC. Voice packets are therefore directed to the voice gateway by means of their destination IP address. This creates an immediate security concern; it will be necessary for the voice gateway to implement some kind of session-level security procedure to prevent service theft and denial of service attacks. The implementation of session-level security will require the assignment of user identities and passwords, adding very substantially to the burden of administration for service delivery.

Carrying voice and data packets on the same VCC also creates a performance concern, because the router or routers that separate voice and data traffic and direct voice traffic towards the voice gateway may suffer congestion at times of heavy data load. It may be possible to address this concern by prioritizing voice packets with the aid of the Type of Service (TOS) field in the IP header, but the use of TOS and the queuing behavior to be applied to each TOS class is not standardized and, in any case, does not provide guaranteed performance.

There's an additional problem with voice and data sharing the same VCC, which has to do with variability of queuing delay. Even if the voice and data packets are placed in separate queues for transmission over the DSL connection, with voice packets getting priority, the transmission of large data packets will monopolize the connection for considerable periods of time. For instance, a 1500 byte IP packet will take more than 30 ms to transmit on a 384 kbps DSL connection. This means that a voice packet that arrives in the transmission queue could wait anywhere between zero and 30 ms before it can be sent. Voice packets therefore experience 30 ms of jitter in addition to that which is introduced by variable queuing delays in the switches and routers in the core of the network. The jitter buffer at the receiving end must allow for this, the result being that overall one-way transmission delay in this case is likely to be in excess of 50 ms, an unacceptable amount of delay for many service providers.

The alternative way to transport voice and data packets over a DSL connection is to carry them on separate VCCs. This approach has the advantage of reducing the scope for breaches of security, since the voice VCC is a point-to-point connection between the integrated access device and the voice gateway. It also improves the outlook for guaranteed voice performance, because the VCC that is provisioned to transport voice packets can be set up with the appropriate traffic class (Constant Bit Rate or Variable Bit Rate Real Time) to give guaranteed Quality of Service. Not only that, if the underlying transport is ATM, the voice packet transmissions can be interleaved with data packet transmissions, avoiding the extra jitter effect of large data packets. On the other hand, there is now the additional necessity for the voice gateway to provide an IP address for the separate VCC into each integrated access device.

The security and performance concerns around the use of a single VCC to transport both voice and data packets are sufficient to deter most voice service providers from adopting this approach. The alternative approach means that we have to provision a separate VCC from each integrated access device to the voice gateway. But having created all these point-to-point links, we should be asking the question: what value does the use of IP bring to justify the additional overhead that it imposes?

Using native ATM or frame-based transport for voice of course means that we have to provision two VCCs to each customer – one for voice and one for data. But once these links are provisioned, we have a bandwidth-efficient and inherently secure solution that requires no IP address administration and no administration of user identities or passwords. And if we define appropriate QoS parameters for the voice VCC, we have no performance concerns either.

To summarize: the universality of IP ought to make it the preferred transport for packet voice over DSL, but a detailed examination of the realities brings one to an entirely different conclusion. As a result, the great majority of VoDSL deployments for the foreseeable future are likely to be based on native ATM or frame-based voice transport. This is true of both distributed (ILEC) and centralized (CLEC) VoDSL network architectures.

5 Requirements on the DSL Access Network

The delivery of voice over DSL places certain requirements on the DSL access network infrastructure over and above those which apply if a data-only service is to be delivered. These requirements are generally not difficult to meet, but attention clearly needs to be paid to them.

5.1 ATM-based DSLAMs

ATM-based DSLAMs must provide Quality of Service on a per-virtual circuit connection basis. In practice this means the voice and data must be handled via different queues both at the DSL ports in the direction towards the user, and at the high-speed uplink port in the direction towards the network. The queues must be managed such that, provided the voice traffic does not exceed the provisioned cell rate, voice will always be given priority over data.

5.2 Frame-based DSLAMs

Frame relay does not support the same set of Quality of Service capabilities as ATM. Nevertheless, frame-based DSLAMs can support mixed voice and data traffic successfully provided that prioritization is applied on a per-virtual circuit connection basis. As is the case for ATM, this means that voice and data VCCs need to be handled through different queues on all DSLAM ports, with voice being given priority.

There is an additional requirement on frame-based DSLAMs which arises from the variable packet size property of frame relay. Where there is a mix of large data packets and small voice packets on a frame-based connection, voice packets will suffer from large variations in queuing delay even if the voice queue is given priority. This issue was identified above in relation to voice over IP, and it applies also to the case of voice over frame.

The delay variation occurs because large packets take appreciable amounts of time to be transmitted over DSL connections, and once a data packet transmission has commenced, it must proceed until the entire packet has been sent. This is in contrast to ATM, where large data packets are segmented into small, fixed-length cells which can be interleaved with voice packets.

In a typical VoDSL situation, where the DSL bandwidth is 384 kbps and the data service supports the transport of 1500-byte IP packets, the voice packets will experience a variable queuing delay or jitter of about 30 ms. This arises because a voice packet that is queued for transmission when the line is not busy will be sent immediately, whereas a voice packet that arrives in the queue just after a 1500-byte data packet has started transmission will have to wait until the entire packet has been sent. The transmission time for a data packet is calculated as (number of bits in the packet) / (line rate), or in this case 1500 * 8 / 384 = 31.25 ms.

If voice packets suffer variability of queuing delay in excess of 30 ms, then the jitter buffer at the receiver has to accommodate at least this amount of delay, which is additive to the packetization delay and other queuing delays in the network. The result is likely to be a total one-way transmission delay in the access network of 50 ms or more, which many voice service providers will regard as unacceptable.

The solution to this problem is frame fragmentation, a technique that involves breaking large data packets down into a number of smaller fragments, permitting a much finer level of granularity in the interleaving of voice and data packets on the DSL connection. A specification for frame fragmentation has been published by the Frame Relay Forum as FRF.12. The frame fragmentation procedure must be carried out at both ends of the DSL connection, in the integrated access device and in the DSLAM. A maximum fragment size of 256 bytes should provide acceptable jitter performance for voice.

5.3 Traffic Management

The bandwidth consumed by voice packets in DSL access network varies with time. When a phone is on-hook, the VoDSL channel that supports that connection is idle, and no voice packets are being sent on this channel in either direction. If the VoDSL system supports silence removal, then the packet stream on a given VoDSL channel is only active when speech activity is present on this channel.

The dynamic nature of bandwidth utilization in a VoDSL system supports statistical multiplexing gains in the DSL access network. Since it is very unlikely that all users of phones connected to VoDSL systems are off-hook at the same time, the overall network capacity can be dimensioned to support a load that corresponds with the likely peak load. For a VoDSL system that supports thousands of phone connections, the likely peak load may be as little as 15-20% of the theoretical maximum.

Taking advantage of this statistical multiplexing gain involves setting appropriate rules for the traffic parameters that are specified when the virtual circuit connections for voice are provisioned. There are two approaches that could be applied here.

5.3.1 Use of rt-VBR Service Category

The first approach is to provision a virtual circuit connection from each integrated access device to the voice gateway, and specify the connection as having the rt-VBR (real-time Variable Bit Rate) service category. With rt-VBR, you can specify a Peak Cell Rate (PCR) and a Sustained Cell Rate (SCR) which correspond with peak and mean load on the VCC. In addition, you need to specify a Maximum Burst Size (MBS) which effectively corresponds to the number of cells that can be sent over the VCC at Peak Cell Rate before the network determines that the traffic contract has been exceeded and starts dropping cells.

The rt-VBR service category is intended to enable ATM networks to realize statistical multiplexing gain with “bursty” traffic. In the case of VoDSL, there are actually two levels of burstiness. The first level corresponds with the variable number of phones that are off-hook at any time, while the second level corresponds with the speech bursts that are present on each active channel. If the VoDSL system does not support silence removal, then only the first level of burstiness applies.

Burstiness that relates to the number of phones that are off-hook at any time involves bursts that may last many minutes, and may involve the transmission of thousands or even tens of thousands of cells in excess of the average load. While it would theoretically be possible to specify a value of MBS that covers such long-duration bursts, in practice most ATM switches do not support such large values of MBS. This means that the value of SCR specified for a VoDSL connection must be almost the same as the PCR in order to support a situation where all the customer's phones are in use simultaneously.

If the VoDSL system does support silence removal, then both long-term (phone usage) and short-term (speech activity) burstiness is present. In these circumstances, the difference between the specified value of SCR for a rt-VBR connection and the value of PCR may be much greater. Since the total amount of bandwidth resource that is reserved in the network for rt-VBR connections is approximately equal to the sum of the SCR values for these connections, the difference between SCR and PCR delivers the statistical gain.

5.3.2 Voice Path Aggregation

The limitations on the use of rt-VBR on a per-VCC basis arise from the statistics of phone usage spread over a relatively small number of lines. For a population of, say, 16 lines to a single customer, there is a definite likelihood that all 16 lines will be in use for some small proportion of the day. But if we take a population of 5,000 lines spread across hundreds of customers, the likelihood that all these lines will be in use at any one time is vanishingly small.

We can take advantage of these statistical effects with VoDSL by aggregating large numbers of virtual circuits that support voice into virtual paths, and provisioning the virtual paths with traffic parameters that match the statistically likely peak load across all the circuits in the path. In a DSL access network, the paths that lead from individual customers' integrated access devices to the voice gateway take the form of a branching tree, with the voice gateway as the root. As traffic traverses this tree from the customer "twigs" towards the voice gateway "root", the branches have progressively more telephony channels in them, and the statistics become more and more favorable to over-booking of the bandwidth.

With voice path aggregation, the dimensioning of PCR for the virtual paths at each level of branching is based on classic telephony traffic engineering principles. These make use of standard statistical rules that are known as "Erlang tables."

Note that not all ATM switches support the provisioning of traffic parameters at the virtual path level. This may be a consideration in the selection of ATM switches to build VoDSL networks.

6 Voice Compression and Echo Cancellation

Voice over DSL provides a cost-effective means of delivering multiple lines of telephony to small business customers who can't quite justify the cost of a leased T1 access line. Customers who require more than 16 phone lines can make a business case to lease a T1 access line, so it follows that the greatest opportunity for VoDSL lies in addressing the market for 16 lines and below.

In the public network, voice is normally carried over 64 kbps circuits. If an uncompressed 64 kbps voice stream is packetized for transport over a DSL access network, the packet overhead will bring the bandwidth required for each voice channel to almost 80 kbps.

With ADSL, upstream bandwidth is limited to 640 kbps, equivalent to a maximum of 8 uncompressed voice channels. And if all these channels were active simultaneously, there would be no bandwidth left over for data traffic.

SDSL fares a little better, since it may be capable of delivering as much as 1.1 Mbps in each direction. Most real deployments of SDSL, though, top out at 768 kbps, which is equivalent to 9 uncompressed voice channels.

To meet the market requirement for up to 16 lines, then, it is quite clear that some kind of voice compression is going to be needed.

6.1 Voice Compression Algorithms

The choice of voice compression algorithms for VoDSL solutions is the result of a number of trade-offs between voice quality, transmission delay and cost of implementation. Table 1 identifies four standard voice encoding schemes that may be appropriate for VoDSL.

All of the algorithms listed in Table 1 meet the requirements of toll quality voice transmission, as measured by Mean Opinion Score (MOS). The tests that are used to arrive at a MOS value for an algorithm are subjective and so they vary from one test laboratory to another; the scores shown in the table are comparable as they are from the same laboratory.

G.726 ADPCM at 32 kbps is widely used today in telecom networks, particularly for international calls. It is a simple algorithm with relatively low delay, and it has been well characterized for "tandeming," where multiple stages of compression and decompression cause progressive deterioration of voice quality. With ADPCM at 32 kbps, it is possible to support up to 16 voice channels over 640 kbps upstream ADSL, and 18 or more channels over 768 kbps SDSL. Hence the use of ADPCM 32 kbps encoding enables VoDSL solutions to meet the market need for up to 16 lines of telephony over a single DSL connection.

G.728 is a low delay, high quality coder operating at 16 kbps. This combination of features should be ideal for VoDSL, but unfortunately cost is an issue with this coder. It requires a good deal of processing power, about 3 times that needed for ADPCM 32 kbps, and is subject to significant royalty payments as it is based on patented technology.

G.729A is widely used for Internet telephony. It combines relatively low complexity with excellent voice quality, but its major disadvantage is delay. An encoding delay of 25 ms, combined with a packetization delay of 44 ms, puts its total delay beyond acceptability for most access network applications. With a smaller packet size, the delay performance can be improved somewhat, but there are other issues with G.729A, in particular its lack of transparency to DTMF tones which are used for tone dialing. When using G.729A, it is necessary for each end to detect DTMF tones and to convert them to digital values that are sent in separate packets, while at the same time suppressing the transmission of the distorted tone. At the far end, the digital DTMF value has to be decoded and the appropriate DTMF tone regenerated. A similar demodulation and re-modulation process is needed to deal with FSK signals which are used to transmit Caller ID information to the customer. These processes add to the complexity, and hence the cost, of a G.729A solution.

Figure 8 shows how the number of lines of telephony that can be delivered over a DSL connection falls away with the length of the loop. It can be seen that 32 kbps encoding is necessary to address the 16-line requirement, and that a 16 kbps encoding scheme can extend the reach for the delivery of 16 lines from about 10 kilofeet to about 13 kilofeet.

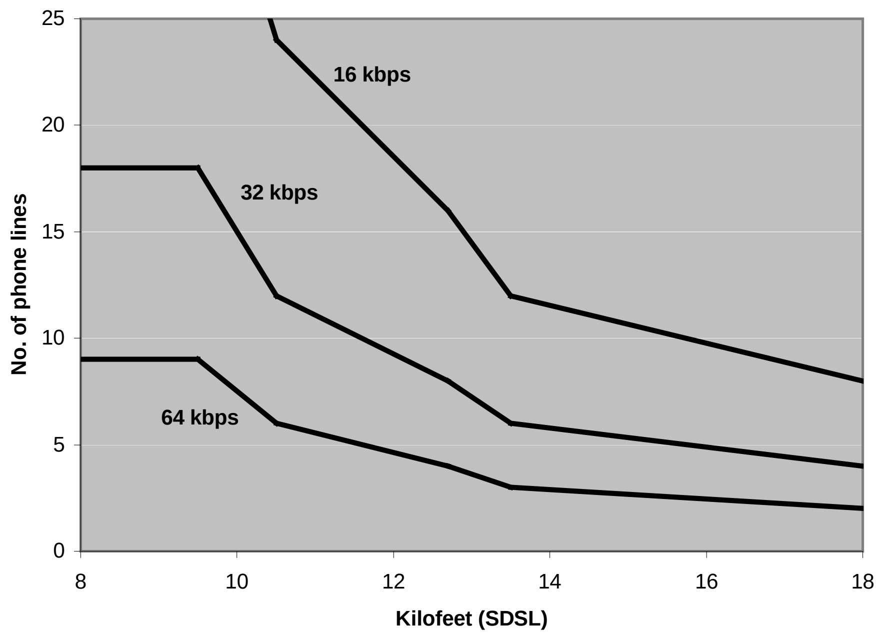

Figure 8: Lines and reach with different levels of voice compression

There is considerable value in using a 16 kbps voice coder for VoDSL, both to extend the reach of service delivery and to maximize the amount of data bandwidth that is available on the DSL connection when multiple voice channels are in use. However, the only standard 16 kbps coder that meets the quality criteria carries with it a cost burden that may be hard to justify in some VoDSL applications. An alternative solution is to use proprietary algorithms that are less processor-intensive than G.728 while offering similar performance.

6.2 Fax and Modem Transmissions

All voice compression schemes impact the performance of fax and modem transmissions. With ADPCM running at 32 kbps, the fastest speed at which modems and fax machines will establish a connection is 4800 bps, while for G.728 the limit is 1200 bps. G.729A won't support a fax or modem transmission at any speed.

There is clearly a requirement for VoDSL solutions to support fax and modem traffic at full speed. The general solution to this is to support uncompressed 64 kbps transport for such traffic. This means that VoDSL solutions need to support a mix of compressed and uncompressed channels, so as to take advantage of the bandwidth savings obtained by compressing voice, but allowing fax and modem traffic to pass unimpeded with full performance.

The selection of voice coding algorithm on each telephony channel over the DSL connection can be carried out statically, by provisioning the lines with or without compression, or it can be done dynamically, by detecting the presence of fax or modem traffic on a line and changing coding algorithm on the fly to uncompressed PCM.

6.3 Echo Cancellation

All telephone networks that deliver analog POTS service generate a small amount of echo. This is caused by imperfections in the device that combines the transmit and receive paths into a single pair of wires at the point where the analog service meets the digital phone network. A small proportion of the signal being transmitted towards the analog phone is reflected at this device (called a "line hybrid") and travels back towards the person speaking.

For local calls, the echo is not noticed by the speaker because there is negligible delay between the sound of the echo and the sound of the outgoing speech. For long distance calls, the echo would be audible because it is delayed with the respect to the outgoing speech by the round-trip transmission time over the network. If this round-trip time approaches 30 ms, then the echo can become noticeable. For this reason, the PSTN inserts echo cancellation devices in the network on long distance calls to remove the echo signal.

When voice is compressed and packetized for transmission over a DSL access network, transmission delays are incurred which are comparable with, or greater than, the threshold for noticeability of echo. It is therefore essential that the integrated access device and the voice gateway work to remove this echo in order to meet the requirements for acceptable speech quality.

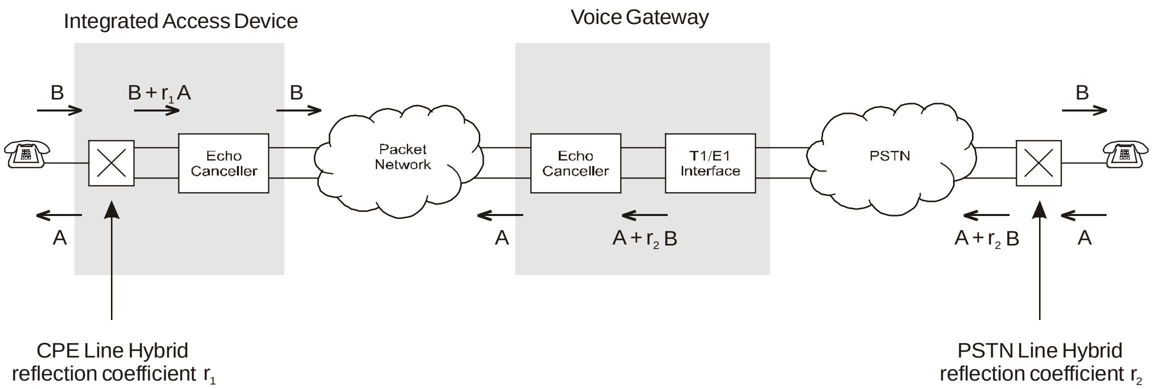

Figure 9: Echo cancellation in a VoDSL network

7 Signaling Transparency

The transparent support of all CLASS, Custom Calling and Centrex features is a key requirement for VoDSL. These features are not only expected and demanded by customers, but they also generate substantial revenues with very high margins for service providers.

In determining how a VoDSL solution can meet the requirement for transparent delivery of these features, it is instructive to examine how digital loop carriers operate in this regard. This is relevant because the VoDSL system is effectively emulating a digital loop carrier system from the point of view of the Class 5 switch.

Digital loop carriers present analog POTS interfaces to the customer, and are connected to Class 5 switches by a DS1-based digital interface such as GR-303 or TR-008. The signaling activity between the switch and the customer takes place in two forms: *supervisory* signaling and *voiceband* signaling.

7.1 Voiceband Signaling

Voiceband signaling comprises tones, such as DTMF tones generated by the keypad during dialing, or FSK tones that convey Caller ID information from the Class 5 switch. All the voice encoding schemes that were discussed above, with the exception of G.729A, are transparent to these tones. As already described, G.729A requires special steps to be taken to ensure that tones are re-generated at the far end.

Provided that the VoDSL system can transparently convey DTMF tones (for dialing) and FSK tones (for Caller ID) over the voice channels between voice gateway and integrated access device, then the calling features that depend on these tones will all operate just as they would on a regular digital loop carrier system.

7.2 Supervisory Signaling

Supervisory signaling is concerned with analog line states. On-hook and Off-hook are supervisory states which have meaning in the direction from the customer towards the switch, while in the opposite direction, the Ringing supervisory state causes the phone to ring, while the Idle supervisory state applies the rest of the time. These are the states that apply to regular Loop Start lines; Ground Start lines used for PBX connections have additional states.

At the analog ports of a digital loop carrier, the supervisory signaling states are represented as voltage and current conditions. For instance, on-hook places an open circuit on the line, while off-hook applies a specific impedance that causes a current to flow in the loop.

These analog supervisory states exist at the telephony ports of the integrated access device in the customer premises, and must be conveyed over the packet network to the voice gateway and hence to the Class 5 switch. The interface between the voice gateway and the Class 5 switch is digital, and the supervisory states must therefore be transmitted over this interface in some appropriate digital form.

7.2.1 Channel-Associated Supervisory Signaling

In networks based on North American standards, the supervisory states are conveyed between the voice gateway and the Class 5 switch by means of a 4-bit codeword that is repeatedly transmitted on each voice channel. This codeword is referred to as the "ABCD bits."

To ensure that the VoDSL system provides the same degree of feature transparency as a digital loop carrier, the best approach for North American networks is to transport the ABCD bits intact between the Class 5 switch and the integrated access device, and to perform the mapping between the ABCD bit values and the analog line states in the integrated access device. This approach guarantees that the VoDSL system is completely transparent to all CLASS, Custom Calling and Centrex features.

An alternative approach is to map the signaling states as conveyed by the ABCD bits into messages that can be sent over a common signaling channel from the voice gateway to the integrated access device. This common signaling channel could be based on standards such as Q.931, which defines the messages for ISDN subscriber signaling. However it is not possible to define a straightforward mapping between ABCD signaling states and Q.931 messages that supports full transparency of all features – Distinctive Ringing and Caller ID in particular do not map cleanly. This means that VoDSL systems which use common channel signaling between the voice gateway and the integrated access device have to employ some proprietary message formats, which has some worrying implications for third-party interoperability testing.

7.2.2 Common Channel Supervisory Signaling

In networks based on European standards, supervisory signaling is conveyed between the voice gateway and the Class 5 switch by means of messages sent over a common signaling channel. The logical approach for these networks is to extend the message-based supervisory signaling channel between the Class 5 switch and the voice gateway all the way to the integrated access device. Unfortunately this is not as simple as it sounds, because the supervisory signaling between the Class 5 switch and the voice gateway is all carried in a single channel, and this has to be de-multiplexed for distribution to all the different integrated access devices that are served by the gateway.

An alternative approach for European-based environments is to convert the messages to supervisory states in the gateway and then convey these states via ABCD bits to the integrated access devices. In the discussion of Channel-Associated Signaling above, it was pointed out that ABCD bit signaling from the switch does not map cleanly to standard common signaling channel messages towards the customer. In the converse situation, however, it is possible to map common signaling channel messages from the switch cleanly into ABCD bit values. Hence a VoDSL system that uses CAS between gateway and integrated access devices can meet the needs of both North American and European markets without any requirement for proprietary signaling features.

8 Network Management

There is not space here to cover the management of VoDSL networks in any detail. That subject will be covered in a later white paper in this series. However it is worth taking a moment to identify the main issues that the management solution will have to deal with.

For the purposes of management, we can view a VoDSL network as a kind of distributed digital loop carrier. This means that, to a large extent, we can follow existing management practice as it relates to digital loop carriers. However the packet network that links the VoDSL gateway to the integrated access devices in the customers' premises presents some additional challenges in arriving at a management solution.

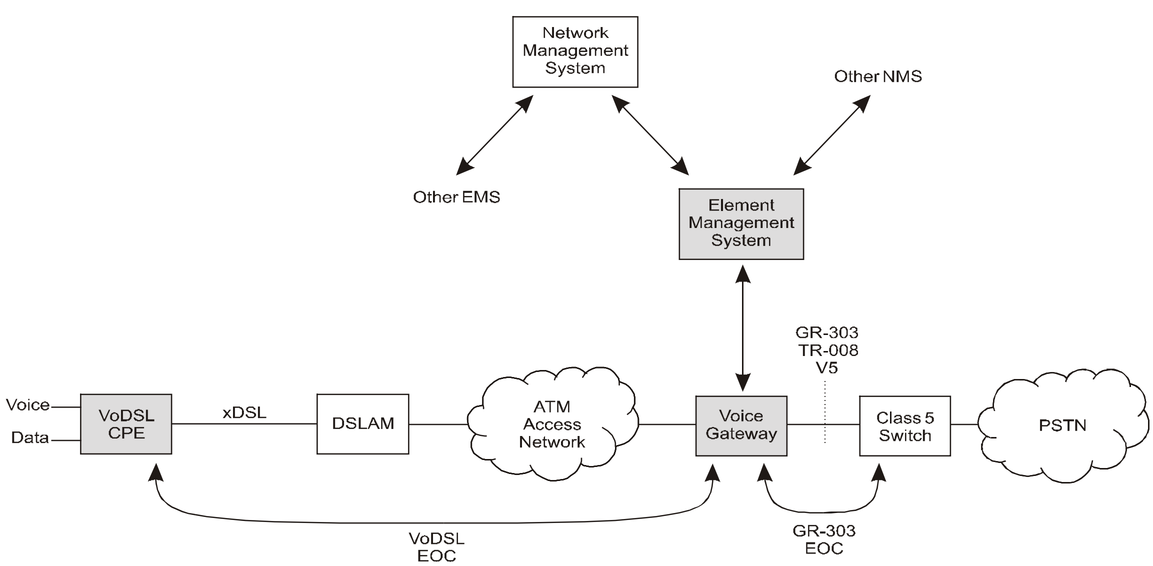

Figure 10: Network management architecture for Voice over DSL

Digital loop carriers are "network elements" and are typically managed by an "element management system" or EMS. In the case of VoDSL, an EMS communicates with multiple voice gateways, and provides application support for configuration, provisioning and alarm gathering. The EMS in turn communicates with a "network management system" or NMS, which communicates with EMS applications that support other aspects of the DSL access network, including the DSLAMs and the ATM switches.

Where a VoDSL voice gateway is connected to a Class 5 switch via a GR-303 interface, an "embedded operations channel" or EOC exists between the switch and the gateway. A limited set of management operations is supported over the GR-303 EOC, including status queries, loop test control and alarm reporting.

Some VoDSL systems also provide an embedded operations channel between each integrated access device and the voice gateway. It is possible for this "VoDSL EOC" to share the same virtual circuit connection as the voice traffic, avoiding the need to provision an additional VCC just for management. The VoDSL EOC allows the voice gateway to configure those aspects of the integrated access device that are critical to voice service delivery, and to obtain performance information, alarms and inventory information from the integrated access device.

9 Conclusion

Building a successful DSL access network to support the delivery of voice services is not simple, but nor is it formidably complex. In this paper we have discussed the principle architectural questions that impact on the design of a VoDSL network solution, and these can be summarized as follows.

- Is voice to be handed off to local Class 5 switch at each DSLAM location, or backhauled across the packet network to some central point of presence? For local voice hand-off, a Gateway with data pass-through feature or a switching DSLAM is required.
- Is the DSL service based on ATM or frame? Both technologies can support voice effectively, but frame-based DSLAMs require prioritization and frame fragmentation support to approach the voice performance of the best ATM-based DSLAMs. In either case, native voice over ATM or frame provides the greatest bandwidth efficiency and the only means to guarantee voice service quality.
- How many lines of derived voice are to be supported over a single DSL connection? If there is any requirement for more than 8 lines, the voice gateway and integrated access devices must support voice compression. Whether voice is compressed or not, echo cancellation is essential.
- What interfaces to Class 5 switches need to be supported? Robbed-bit style interfaces such as GR-303 and TR-008? Or message-based interfaces such as V5? Whichever signaling scheme is used between integrated access device and voice gateway, care should be taken to verify that complete signaling transparency is achieved such that all CLASS, custom calling and Centrex features are supported, preferably without the need for proprietary signaling messages.
- How is the network to be managed? Particular attention should be paid to the management of the point of service delivery, which is the integrated access device. An effective VoDSL Embedded Operations Channel between integrated access devices and the voice gateway is a pre-requisite here.

Dealing with these questions carefully is critical to a successful deployment of voice over DSL. But armed with the right answers, service providers can move forward with confidence to take advantage of the huge opportunities afforded by this exciting new telecommunications technology.