VOIP IN A BROADBAND ACCESS NETWORK

An analysis of the issues facing service providers deploying voice services in an IP access network

A MetaSwitch™ White Paper

Document VWP-002-0101

EXECUTIVE SUMMARY

OVERVIEW

The term "Voice over IP" (VoIP) describes the transport of voice over IP based networks – including the Internet. Voice can be transported reliably and securely using IP and higher-level protocols – RTP/RTCP (Real Time Protocol/Real Time Control Protocol) with UDP/IP to transport digitized voice, and MGCP, Megaco, SIP and H.323 for signaling and device control.

To date, VoIP has been mainly deployed in enterprise networks or as a trunking technology to reduce transport costs in voice backbone networks. VoIP can also now be deployed to provide next generation access network solutions for a complete end-to-end carrier VoIP network – though there are a number of issues that a service provider must consider in order to provide a PSTN-quality service in this way.

This White Paper outlines the architecture and benefits of VoIP in a carrier broadband access network as part of new service deployment or a PSTN evolution strategy. It then goes on to examine the main potential issues in such deployments and the key decisions a service provider must make in order to successfully use VoIP in this role.

CONCLUSIONS

VoIP, as provided over a broadband access network, has several advantages over traditional, time-division-multiplexed service provided over the PSTN (Public Switched Telephone Network). It is particularly effective as a platform for rolling out new services or as a means to converge voice a data networks onto a single cost-reduced network.

The key network design choices a carrier must make when utilizing VoIP in the broadband access network are as follows.

- What services need to be offered, for example full PSTN equivalence, or a more restricted "cheap second line" service.
- How much bandwidth is available in the last mile network, which will affect the choice of voice codec, packetization period, and where to use compression to best meet the service goals.
- Whether echo control is required to ensure that voice quality is not hampered by the inherent delay in IP-based access networks.
- Which signaling protocols support the service set required. PSTN-equivalence typically requires use of device control protocols such as MGCP or H.248, rather than service protocols such as H.323 or SIP.
- The types of end user terminals supported – POTS phones, PC clients, IP Phones or PBXs.
- Whether the Quality of Service requirement for voice requires that packetized voice be prioritized over data traffic, typically using Diffserv or MPLS.

- The security risks must be clearly identified and appropriate techniques employed to ensure that the call agents, in particular, are protected from attack. This is a particular problem if direct SIP traffic is supported, as opposed to POTS phones connected to trusted edge devices.
- Lawful interception requirements in many countries will prevent a public carrier from allowing direct connection between IP phones. All calls must be routed via an access gateway that hides any intercepts in place.
- IP addressing becomes more complex if the service is to be offered wholesale to other carriers (e.g. in conjunction with several data ISPs). Typically, appropriate proxies and gateways must be added to allow for voice and data traffic to flow over different IP address domains.

The best solution to these issues, which all interact to some degree, must be tailored to suit the particular service set a carrier wishes to offer – but the solutions do exist and VoIP can be successfully deployed in the carrier broadband access network with careful network design.

NOTICE

This white paper is provided for informational purposes only. MetaSwitch and Data Connection make no warranties, express or implied, as to the accuracy of the content. MetaSwitch and Data Connection are registered trademarks of Data Connection Limited and Data Connection Corporation. All other brand and product names are trademarks or registered trademarks of their respective owners.

CONTENTS

1. OVERVIEW OF VOICE OVER IP

At its simplest, Voice over IP is the transport of voice using the Internet Protocol (IP). However, voice over IP doesn't imply the use of the public Internet. Private, managed IP networks are required to guarantee security, quality and reliability.

IP, by itself, merely provides an unreliable datagram service. Higher layer protocols are required to deliver a voice service.

- Real Time Protocol / Real Time Control Protocol (RTP/RTCP) using UDP/IP as a transport protocol is used to transport the digitized voice.
- One of a number of signaling or device control protocols is used to set up and tear down voice calls, including MGCP (Media Gateway Control Protocol), Megaco (also known as H.248), SIP (Session Initiation Protocol) and H.323.

1.1 BROADBAND ACCESS NETWORK

VoIP can be deployed in many different network segments. To date, it has been mostly deployed in the backbone and enterprise networks. The broadband access network introduces additional constraints and issues discussed in section 3.

Figure 1 presents a generalized view of a broadband access network in a fully distributed functional form.

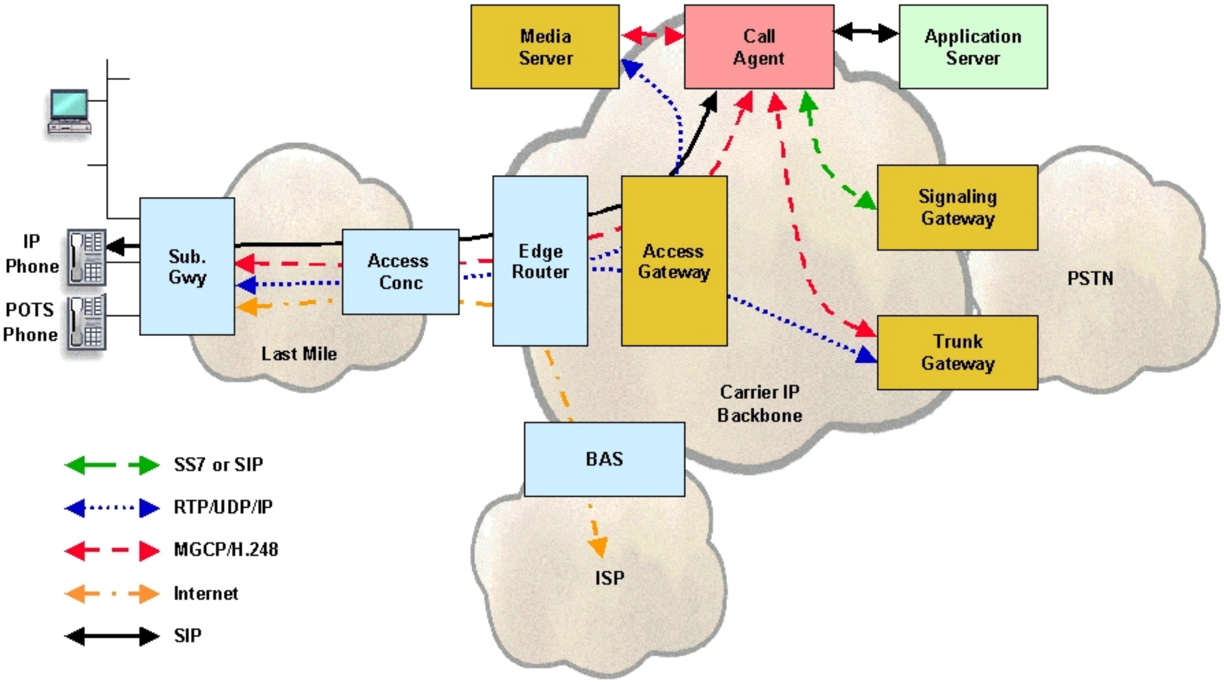


Figure 1: Fully Distributed Broadband Access Network

In general, there is no single, agreed understanding of the functional elements in a broadband access network, and in a real deployment, many of the functions are likely to be combined. The following sections outline the function of each of the elements in Figure 1.

1.1.1 Subscriber Gateway

This element, also known as the residential gateway, IAD or MTA, is one type of media gateway. It terminates the WAN (Wide Area Network) link (DSL, T1, fixed wireless, cable etc) at the customer premises and typically provides both voice ports and data connectivity. Usually, it uses a device control protocol, such as MGCP or Megaco, under the control of the call agent.

1.1.2 Access Concentrator

This element terminates the WAN links used over the "last mile", and concentrates voice and data traffic, typically at the service provider premises. For example, in a DSL network, this is a DSLAM; in a cable network, a CMTS.

The access concentrator may absorb subscriber gateway functions in some architectures. For example, some DSLAMs (or next generation DLCs) can support direct POTS connections and will packetize the analog voice under the control of the call agent.

1.1.3 Edge Router

The edge router routes IP traffic onto the carrier backbone network. It may also provide other functions, such as header (de)compression and multiplexing. This element may be combined with the access concentrator.

1.1.4 BAS (Broadband Access Server)

The key function of the BAS is to provide access from the carrier backbone network to an ISP's (Internet Service Provider's) network. Typically, it provides subscriber management and authentication functions. This element may be combined with the edge router.

1.1.5 Trunk Gateway

The trunk gateway sits between the carrier IP backbone and the TDM (Time Division Multiplexing)-based PSTN. It provides transcoding from the packet-based voice, VoIP onto a TDM network. Typically, it uses a device control protocol, such as MGCP or Megaco, under the control of the call agent.

The trunk gateway can be combined with the access gateway, signaling gateway, media server or call agent in a single, integrated device.

1.1.6 Access Gateway

The main function of the access gateway is to protect voice resources in the core network from misuse, though it also incorporates other voice-specific functions such as lawful intercept capability. Inclusion of access gateway function is one of the main differences between a carrier-class VoIP architecture and a less secure enterprise network. The access gateway can be combined with the trunk gateway, signaling gateway, media server or call agent in a single integrated device.

The access gateway uses a device control protocol, such as MGCP or Megaco, under the control of the call agent. Some of the functions this device can provide are

- connecting local ‘on-net’ calls
- lawful interception – intercepting call content to provide to Law Enforcement Agencies
- policing use of voice Differential Services Code Points (DSCPs) across the core network to trunk gateways or other access gateways.

1.1.7 Media Server

This element is also called an announcement server. For voice services, it uses a control protocol, such as MGCP or Megaco, under the control of the call agent or application server.

Some of the functions this device can provide are

- playing announcements
- mixing – providing support for 3-way calling etc
- codec transcoding
- tone detection and generation
- interactive voice response (IVR) processing
- fax processing
- Voice Activity Detection.

The media server’s function may overlap with that of the access gateway, and the elements can be combined. The media server can also be combined with a trunk gateway, signaling gateway or call agent.

1.1.8 Signaling Gateway

This element acts as a gateway between the call agent signaling and the SS7-based PSTN. It can also be used as a signaling gateway between different packet-based carrier domains. It may provide signaling translation (between SIP and SS7) or simply signaling transport, (i.e., carriage of SS7 over IP). It has a protocol interface to the call agent.

The signaling gateway can be combined with the access gateway, trunk gateway, media server or call agent in a single integrated device.

1.1.9 Call Agent

This element is also called a media gateway controller, or a softswitch. It provides the call logic and call control signaling for one or more media gateways, maintaining call state for every call on each media gateway. Many call agents include service logic for CLASS services (such as Caller ID and Call Transfer), and may interact with application servers to supply services that are not directly hosted on the call agent.

The call agent can be combined with the access gateway, trunk gateway, media server or signaling gateway in a single integrated device.

1.1.10 Application Server

This element provides the service logic and execution for one or more applications or services that are not directly hosted on the call agent. For example, it may provide voice mail or conference calling facilities. Typically the call agent will route calls to the appropriate application server when a service is invoked that the call agent cannot itself support.

1.1.11 Networks

The last mile network, carrier backbone, ISP and PSTN networks shown in Figure 1 may all be owned and operated by different companies. In many cases, the carrier backbone and ISP networks will be separate and will use separate IP addressing schemes. The implications of this for IP addressing and security trust models are highlighted in section 3.8.

2. BENEFITS OF VOICE OVER IP

Migrating to Voice Over IP in the access network has many benefits.

IP is ubiquitous. Whatever the access network type – whether Frame Relay or ATM-based DSL, T1, fiber or fixed wireless – IP is invariably available for the transport of data, and with it packetized voice.

For service providers examining the business case for VoIP, it is clear that the consolidation of voice and data in one network significantly reduces cost, for a number of reasons.

- IP leverages data network capacity.
- IP equipment is typically faster and cheaper than ATM or TDM equipment – a gap that is increasing rapidly every few months.
- Re-routing of IP networks (e.g. with MPLS) is much cheaper than, say, SDH protection switching.
- IP equipment has lower management costs.
- The open standards on which VoIP is built foster increased vendor competition.

More importantly, VoIP is the key technology that will enable next generation networks to deliver innovative converged voice and data services that are difficult or impossible to implement with the current PSTN infrastructure.

- IP-based internet applications, such as email and unified messaging, may be more easily integrated since the data and voice services are delivered on a single IP plane.
- The flexibility of next generation platforms allows for the development of new services.
  - Development cycles are typically shorter than for TDM-based equipment.
  - VoIP products, unlike legacy TDM switches, often support open service creation environments which allow third party developers to invent and deliver differentiated services.

Whatever the justifications, most service providers recognize that VoIP is the direction of the future – though it is not always clear how to migrate from today's networks to reap the promised rewards. The remaining sections of this paper examine the challenges facing service providers starting out down this path.

3. ISSUES IN A BROADBAND ACCESS NETWORK

In order to deploy a carrier-grade VoIP service in a broadband access network, various issues need to be addressed.

- Service set to be offered, and the types of end user terminal supported.
- Choice of signaling protocol.
- Quality of Service (QoS).
- Bandwidth utilization.
- Echo Control.
- Reliability.
- Security.
- IP address domains.
- IP and PC phones
- Fax support.
- Auto-configuration.
- Lawful interception.

The following sections give an outline of each of these issues, options for resolving them and recommended solutions

3.1 SERVICE SET

At the risk of stating the very obvious, the crucial first decision facing the designer of a VoIP network is the service set that needs to be supported. This could range from a minimal set of services for “teen line” offerings alongside data services, through to full PSTN equivalence and advanced services for carriers wishing to replace their current infrastructure with a new converged network for all subscribers.

An important part of the service design is choosing the types of end user terminal that are to be supported. Possible choices include

- POTS “black phones”
- PBXs and key systems
- PC soft-clients (including web-based applications)
- IP phones.

The choice of service and terminals is outside the scope of this paper, but does interact with many other network design decisions. Where this interaction is particularly important, we have highlighted it in the discussion of other issues below.

3.2 SIGNALING PROTOCOLS

Numerous different signaling protocols have been developed that are applicable to an access network. They fall into two classes.

- Device control protocols e.g. MGCP and H.248 (Megaco).
- Service protocols, e.g. SIP and H.323.

3.2.1 Device Control Protocols

Media Gateway Control Protocol (MGCP) and H.248 (Megaco) are access device control protocols. They deal with functions like on/off hook handling, detecting hook-flash and Multi-Frequency (MF) tones, and basic announcements. They are effectively identical in purpose and very similar in syntax. Both are used by call-aware control entities to control call-unaware devices.

MGCP has become the de facto industry standard and has been adopted by the International Softswitch Consortium (ISC) (see www.softswitch.org). Network Call Signaling (NCS), which is based on MGCP, has been adopted by CableLabs (see www.cablelabs.com) as the basis of their PacketCable VoIP standards.

H.248 has the blessing of formal industry standards bodies such as the ITU (International Telecommunications Union) and IETF (Internet Engineering Task Force) and is newer than MGCP.

MGCP has much wider industry acceptance as a device control protocol and is the current protocol of choice. H.248 has been predicted to take over in six months time for the last 18 months, but has not yet achieved the same level of acceptance. However H.248 is rapidly becoming the protocol of choice for control of trunk gateways owing to the ongoing work in the ITU, MSF and other bodies to ensure it scales sufficiently well to handle these larger boxes.

3.2.2 Service Protocols

Session Initiation Protocol (SIP) and H.323 are peer-to-peer service protocols for establishing sessions between communicating entities. The key difference from device control protocols is that the end-points, typically subscriber gateways or IP phones, are required to maintain call state.

SIP is regarded as a “light” protocol that primarily deals with session establishment and uses other protocols, such as Session Description Protocol (SDP), for capability negotiation. In contrast, H.323 is derived from ISDN, and is regarded as a “heavy” protocol. In practice, as the SIP protocol continues to evolve, and as it starts to address the same function as H.323, it is becoming “heavier”.

While H.323 has been historically deployed, and whatever the technical merits, SIP is now replacing H.323 as the end-to-end protocol of choice for intelligent end-points.

3.2.3 Conclusion

At present, the service protocols are generally unable to support a full PSTN-equivalent service set, though ongoing standards work is slowly closing this gap. Hence a carrier who wishes to deploy full PSTN service equivalence should currently use one of the device control protocols.

3.3 QUALITY OF SERVICE (QOS)

Quality of Service (QoS) is one of the biggest, if not *the* biggest issue facing the deployment of VoIP in a broadband access network. The PSTN provides a very high-quality service; the speech quality is high and no perceptible echoes, noticeable delays, or annoying noises are heard on the line. This is a tough standard to meet.

Voice quality is very sensitive to three key performance criteria in a packet network.

- Delay.
- Jitter.
- Packet loss.

Unfortunately, IP, by its nature, provides a best-effort service and does not provide guarantees about the key criteria. There are three solutions to the problem of providing guaranteed QoS.

- Allocate more bandwidth.
- Implement a QoS protocol, such as Diffserv, RSVP or MPLS to guarantee prioritization of voice media streams over best-effort data.
- Carry voice media streams on a logically separate path from data.

Allocating more bandwidth in the access network (for example, running fiber to every subscriber) is not economically viable. Therefore, this section will discuss the remaining options for providing guaranteed QoS in the access network.

3.3.1 Diffserv

Diffserv (Differentiated Services) is a relatively simple means of prioritizing different types of traffic – in particular, prioritizing voice traffic above data traffic. Basically, Diffserv makes use of the IPv4 Type of Service (TOS) field (and the equivalent IPv6 Traffic Class field) to specify the Differentiated Services Code Point (DCSP). Each DSCP is mapped to a particular forwarding type, known as Per-Hop Behavior (PHB), according to the policy set for the network.

Diffserv defines two types of PHB; expedited forwarding (EF) and assured forwarding (AF).

Expedited forwarding PHB can be used to build a low-loss, low-latency, low-jitter, assured-bandwidth, end-to-end service through Differential Service domains. Such a service appears to the endpoints like a point-to-point connection or a “virtual circuit”. EF is provided by ensuring that queuing delays at each transit node are removed. This is done by assigning to a given traffic stream a minimum departure rate from each transit node that is greater than the pre-agreed maximum arrival rate.

Assured forwarding PHB is a service in which packets from a given source are forwarded with a given probability, provided that the traffic from that source does not exceed some pre-agreed maximum. There are four AF classes. Each class is allocated a certain amount of forwarding resources (buffer space and bandwidth) in each transit node. Within each AF class, IP packets are marked with one of three possible drop precedence values. In case of congestion, the drop precedence of a packet determines the relative importance of the packet within the class. A congested transit node tries to protect packets with a lower drop precedence value from being lost by preferably discarding packets with a higher drop precedence value.

Diffserv is fairly straightforward to implement and has already been widely deployed. It requires the following.

- Media gateways must set the DSCPs to differentiate high-priority traffic, such as voice, from lower-priority traffic, such as data.
- All transit nodes – including the subscriber gateway, access concentrator and edge router – need to implement an appropriate queuing mechanism, such as weighted fair queuing.

3.3.2 RSVP

RSVP (Resource Reservation Protocol) is a more complicated protocol that enables resources to be reserved for a session before any attempt is made to send or receive traffic. Though more complicated than Diffserv, RSVP provides QoS guarantees equivalent to a circuit-based network.

RSVP currently offers two levels of service.

- Guaranteed – which is equivalent to circuit emulation.
- Controlled load – which is equivalent to a best-effort network under no-load conditions.

RSVP sets up via an explicit path through the network, using a signaling protocol, and reserves resources for the flow at each transit node. These reservations are ‘soft’ and need to be refreshed periodically.

RSVP is harder to implement than Diffserv. Subscriber gateways must implement the RSVP protocol and set up paths through the network for voice traffic. In addition, RSVP usually requires all routers and access concentrators to be upgraded.

3.3.3 MPLS

MPLS (Multi-Protocol Label Switching) is not primarily a QoS solution, although it can be used to support the QoS requirements. Instead, it is a new switching architecture, which is used instead of IP routing. Standard IP routing requires each router to examine the IP header and determine the next hop, normally based on the IP address. MPLS takes a different approach. It attaches a label to each packet at the ingress point of the network. The packet and its label are passed to the next node, which examines the label and determines next hop and label to use. The key difference from IP routing is that the label is determined at the point of ingress and can be chosen based on criteria such as destination and QoS requirements. It can then force a packet to take a specific route through the network, which can be important for ensuring QoS.

For MPLS to work, the labels used need to be distributed to all the nodes. Several protocols have been defined to do this, including an extended version of RSVP and LDP (Label Distribution Protocol). The path a particular packet takes through the network is referred to as a Label Switched Path (LSP).

MPLS is harder to implement than Diffserv. It can also be used in conjunction with Diffserv so that, for example, the subscriber gateways use only Diffserv in the last mile network, but the edge router maps the different DSCPs into traffic engineered LSPs across the backbone. Such a solution avoids the need for MPLS-awareness outside the core network equipment, and is typical of the QoS architecture envisaged for many carrier VoIP networks.

3.3.4 Packet Fragmentation

It is also necessary to minimize delay in the access network. One key issue is that packets are transported serially, and once a packet has started transmission, no other packets can overtake it. Data packets are much larger than voice packets and can take a significant time to transmit. For example, transmitting a 1500 byte packet over a 256 kbps upstream ADSL link takes approximately 45 milliseconds. This introduces unacceptable delay and jitter to the voice service sharing that link.

The solution to this problem is to fragment large data packets so that higher-priority voice traffic does not have to wait in a queue. There are two mechanisms available for fragmentation.

- Use a Layer 2 protocol that provides fragmentation, such as ATM or Frame Relay. For example, service can be provided using two ATM Permanent Virtual Circuits (PVCs), one for the voice path and one for the data path. As ATM cells are always 53 bytes, the transmission delay is minimal. Note that using only a single PVC does not resolve the issue, as it is not possible to interleave voice and data packets on a single PVC.
- Fragment the data packets using IP fragmentation. This requires the subscriber gateway and access concentrator to fragment large data packets into small IP packets before transmitting them across the access network. However, this approach does impact performance within the network, either requiring the Broadband Access Server (BAS) to reassemble data packets before forwarding them into the network, or imposing a higher packet forwarding load on routers throughout the network.

3.3.5 Conclusion

Diffserv is the most suitable QoS protocol for the access network. It is lightweight and straightforward to implement, yet provides adequate guarantees of QoS to provide high-quality voice services.

RSVP and MPLS are much more complicated and the additional guarantees provided are not required in an access network, which typically shares a point-to-point non-routed link between data, voice and possibly video traffic. However MPLS can be usefully deployed in the core network and interoperates with Diffserv in the access network.

In addition, one of the mechanisms for data fragmentation must be implemented. Using two PVCs is a good technical solution but negates some of the benefits of using VoIP. In particular, it increases the configuration and management requirements. However, IP fragmentation is undesirable, as it greatly increases the number of IP packets in the data network and could impact the service provided. The choice of solution is likely to be dependent upon the particular network architecture, and subscriber gateways need to be capable of implementing either.

3.4 ECHO CONTROL

This is a subject that could fill a complete white paper on its own, so we will restrict the discussion here to a short summary of the issues that will be familiar to many readers already.

The current PSTN echoes a user's own voice back into the earphone of their handset, caused by the 2-wire to 4-wire bridge in the far end central office. This echo is not noticeable provided that the total round-trip delay is kept below about 150ms. Much effort has been expended on the current TDM-based PSTN to ensure that this is achieved or, if it cannot be, that echo cancellation DSPs are provided to remove the echo and hence improve the voice quality. However many national networks assume very tight limits, sometimes less than 20ms, on the total delay that is permitted in the access network in order to optimize the number of locations where echo cancellers are deployed, for example only at international gateway switches.

Unfortunately VoIP access networks using long packetization periods and low-rate codecs to maximize bandwidth efficiency cannot meet these tight delay budget restrictions. Hence echo cancellation should be deployed on all calls that are passed out to the PSTN via a trunk gateway. If this is not done, the voice quality will be perceived as patchy. For example, local calls may be fine owing to the short delay in the PSTN, but inter-state calls may be affected by echo problems.

Most trunk gateways incorporate echo cancellation DSPs. However the carrier should check that sufficient DSP is provided for the mix of calls they envisage using or, ideally, for all calls so that there can never be any shortage of DSP to cause intermittent voice quality issues. If a carrier already implements echo cancellation on all calls, for example at the trunking layer in a mobile network, the VoIP trunk gateways may not need to include echo cancellation DSPs.

3.5 BANDWIDTH UTILIZATION

In a VoIP network digitized voice is transported using real-time protocol (RTP). A typical voice sample is less than 100 bytes, but the combined headers are at least 40 bytes.

For example, using G.726-32 with a 15-millisecond packetization period generates 60-byte voice samples. Assuming IP is carried over ATM AAL5, the voice packet also requires

- a layer 2 transport header – such as RFC 2684 (obsoletes RFC 1483) = 8 bytes
- the IP, UDP and RTP headers = 40 bytes
- a AAL5 trailer = 8 bytes.

The full voice packet occupies 116 bytes(shown in Figure 2 below), which requires three ATM cells (each carrying up to 48 bytes), giving ~62kbps of bandwidth with a transmission efficiency of 37%. This eliminates the theoretical bandwidth saving from G.726-32 compared with using 64kbps G.711.

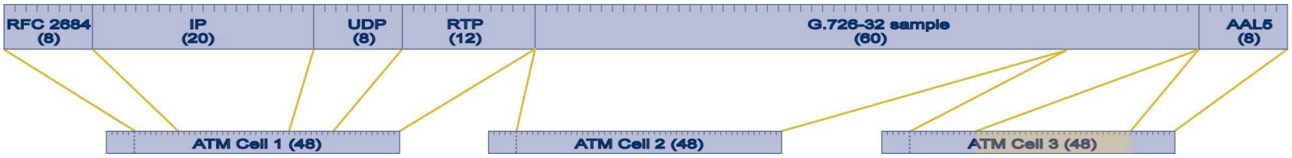


Figure 2: Uncompressed Voice Packet

Given that the access network is of limited bandwidth and one of the advantages of using VoIP is that it should be possible to use lower bit codecs to save bandwidth, a mechanism for reducing the overhead is required. The main approach is to implement IP header compression.

3.5.1 Header Compression

IP header compression is defined in IETF RFCs 2507, 2508 and 2509, and allows for compression of the RTP, UDP and IP headers. This reduces the RTP, UDP and IP header overhead from 40 bytes to an average of 4 bytes.

For the same G.726-32 example discussed above, the compressed header voice packet now occupies 80 bytes (shown in Figure 3 below), which requires two ATM cells, giving ~42kbps of bandwidth with a transmission efficiency of ~75%. The full header still flows on the first packets so bandwidth must be allocated to allow for this, but subsequent packets just send the differences between the RTP/UDP/IP header and the previous header.

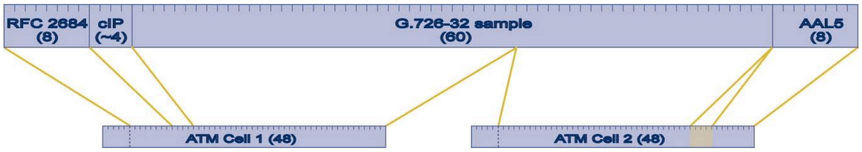


Figure 3: Compressed Voice Packet

The downside is that header compression requires state information at each endpoint, and consumes additional CPU and storage on each endpoint.

When using a Layer 2 protocol with a fixed cell size, such as ATM, header compression does not automatically reduce the bandwidth required: bandwidth saving only occurs if the number of cells necessary to send each voice sample is reduced, because the cells are zero-padded. In the example given above, the number of ATM cells required to transmit each voice sample is reduced from three to two, but the packetization interval does need to be selected carefully to maximize the bandwidth efficiency when traversing a cell-based network.

3.5.2 Conclusion

Header compression provides a useful bandwidth saving in a broadband access network at the expense of requiring some CPU overhead per endpoint. A key decision is where these endpoints should be located – with typical choices being the subscriber gateway at one end and either the access concentrator or edge router at the other. For best results, the packetization period for voice must be carefully chosen to match the cell boundaries of any cell-based network technology, such as the ATM layer underlying DSL.

In general, it is not worth trying to use header compression in the backbone network, where bandwidth is less scarce than in the access network. If a carrier is only supporting POTS lines from a combined access concentrator and subscriber gateway (for example, a voice-capable DSLAM or DSL-capable DLC), there may be no need to support compression anywhere in the network.

3.6 RELIABILITY AND SCALABILITY

The PSTN achieves five-nines reliability, equivalent to fewer than five minutes per year downtime, and it handles millions of simultaneous calls. A VoIP network needs to achieve similar levels of reliability and scalability.

The required reliability and scalability can be achieved in a VoIP network by using redundant and load-sharing equipment and networks. The call agent, access gateway, trunk gateway, signaling gateway and media server need to be fault tolerant. Fault tolerance requires

- redundant hardware
- redundant network connections
- hot-swap capability
- no single point of failure
- software and firmware upgradeable without loss of service.

3.7 SECURITY

The PSTN has been very resistant to security attacks and has not suffered from significant problems since the introduction of SS7 out-of-band signaling. A VoIP broadband access network must address three key security issues.

- Denial of service.
- Theft of service.
- Invasion of privacy.

3.7.1 Denial of Service

A denial of service attack prevents legitimate users of a network from accessing the features and services offered by that network. Denial of service attacks are extremely difficult in the PSTN but all too common in packet networks. There have been several successful attacks on web servers on the Internet, even including the White House.

In a complicated network, there are many possible denial of service attacks. Some examples include sending false signaling messages so that a call agent is fooled into believing that a party has gone on-hook and bombarding a device with pings or other packets so frequently that it has no spare processing power to process legitimate requests.

3.7.2 Theft of Service

Theft of service attacks are aimed at the service provider, where the attacker simply wants to use a service without paying for it. The most common form in the current PSTN is called subscriber fraud, where a subscriber sets up an account with a service provider using false billing information, for example a stolen credit card. Other forms of theft are more technical, often utilizing “black boxes” or similar to fool the network into providing free service.

In a VoIP access network, bandwidth is a key resource, and is important to billing for calls. Therefore, the network needs to be protected from subscribers stealing bandwidth.

3.7.3 Invasion of Privacy

Subscribers to the PSTN expect that their calls are private, and that no third party can eavesdrop (with the exception of lawful interception). The PSTN achieves this privacy mainly by physical security mechanisms – the wire from a subscriber's home is only connected to the local exchange or digital loop carrier and cannot easily be accessed.

This is not the case with cordless phones or first-generation cell phones. It is easy to eavesdrop on these calls, using a radio receiver – as some celebrities have found out to their cost.

3.7.4 Security Model

A VoIP broadband access network must be designed to address security threats. For each interface, the following must be considered.

- Authentication and Non-repudiation.
- Access control.
- Integrity and Confidentiality.

Authentication and Non-repudiation

Fundamentally, the problem is that elements in a VoIP network are identified by an IP address, but it is quite possible to 'spoof' a source IP address and steal someone else's identity.

The various elements in the network must authenticate each other's identity. In particular, this applies to the subscriber gateway and the call agent. The call agent must authenticate the subscriber gateways under its control to ensure that nobody is fraudulently using another subscriber's identity to obtain service. This also provides non-repudiation as the subscriber cannot legitimately claim that another party has made a call or sent a particular fax.

It is less important for subscriber gateways to authenticate the call agent, but it is possible to imagine a scenario in which a hacker has compromised the security of a call agent and set up a dummy call agent in order to eavesdrop.

Authentication in an IP network is addressed by IPSec, where the IP authentication header (AH), using transport mode, provides the required level of authentication. Both MGCP and Megaco support the use of IPSec for the transport of signaling messages.

Using IP authentication requires the distribution of encryption keys. Keys may be distributed via periodic updates or manually – for example, by shipping subscriber gateways with a key already installed.

Access Control

Access control is required to ensure that only authorized subscribers are permitted to use the telephony services. For example, subscribers who have not paid their bills may have service disconnected or only be permitted to call emergency services. In addition, access control is required to prevent subscribers from exceeding their permitted media bandwidth allocations. Access control requires control of both the signaling and the media flows.

The call agent is responsible for controlling the signaling flows. For example, it can take a subscriber gateway out of service or reject all call attempts other than to emergency services.

The control of the media flows requires an access gateway function, and is one of the key differences between a carrier-class VoIP architecture and less secure enterprise networks. The access gateway function is responsible for policing all media flows to and from subscriber gateways to ensure that only authorized flows are permitted. It silently drops any media flows that are not authorized, to prevent denial of service and theft of service attacks. With Diffserv, the policing involves checking the DSCP of each packet and routing or dropping the packet as required.

Integrity and Confidentiality

A subscriber expects calls and faxes to be private and unintercepted and not changed by a third party.

It is easier to meet this expectation in a VoIP network using a point-to-point architecture for the last mile – DSL, for example – than in a VoIP network using shared media such as cable or fixed wireless for the last mile.

In both cases, the carrier IP backbone must be protected from unauthorized access and denial of service attacks. This protection is most easily achieved using a private, managed IP network with strictly limited firewall access or no access to any public networks. If SIP access to the network is supported using public IP addresses (or from a different IP address domain), it is strongly recommended that the call agent is not given a public IP address. Instead, SIP proxies at the network edge can gateway public SIP access into the secure private network. These proxies can also act as "fuses" to limit damage to the service received by the carriers own customers during a concerted denial of service attack.

In a shared media network, it may also be necessary to encrypt the media to prevent eavesdropping on the last mile. This could be done with an IPSec encapsulation header, as in the packet cable standards.

3.7.5 Conclusion

It requires significant effort to make a VoIP broadband access network secure.

It is not possible to give a one-size-fits-all solution for the security model a carrier should employ in a VoIP broadband network. At a minimum a carrier must

- determine the security domain boundaries in the network, which includes deciding whether subscriber gateways are trusted devices

- secure the backbone network, in particular the call agent, against attack using firewalls and sacrificial SIP proxies at the network edge
- implement IPSec or similar authentication of subscriber gateways and other non-trusted entities
- police media flows into the core network using an access control function in the edge router or access gateway to prevent high priority voice bandwidth being usurped for other purposes that have not been authorized.

Encryption may also be required to prevent eavesdropping.

3.8 IP AND PC PHONES

One of the key benefits of moving to a VoIP network is the ability to support PC soft-clients and other IP- telephony-capable devices that can offer a much richer interface, for example using web-based GUIs. These “soft IP phone” applications are typically attached to a subscriber gateway via local Ethernet, 802.11b or HomePNA local area network (LAN).

The PC-based nature of these devices places some additional requirements on the VoIP network, though some of what follows also applies to dedicated IP phones connected via a LAN.

3.8.1 Best Effort Service vs. Guaranteed Service

It is perfectly possible to use an IP phone to connect to an ITSP (Internet Telephony Service Provider) using the Internet. Indeed, competition rules may require that the carrier allow their customers to do this even if they also subscribe to a voice service from that carrier. However, such calls receive only best effort service and will typically be low quality. This allows the carrier to compete with ITSPs by offering guaranteed high voice quality for calls placed using their own call agent and VoIP broadband access network versus the much more variable quality an ITSP can deliver over the public internet.

Note that calls placed via ITSPs do not involve any of the elements in the VoIP broadband access network. In Figure 1 (in section 1.1), the traffic follows the Internet connection to a separate data network.

3.8.2 Signaling

IP and PC phones typically use one of the various service protocols described in section 3.2.2, such as SIP.

3.8.3 Media Prioritization

To provide guaranteed quality of service, the subscriber gateway must route the media packets from the locally attached IP phones onto the VoIP network. To prevent denial of service and theft of service attacks, each call needs to be authorized by the call agent, and only packets for authorized calls should be routed on to the VoIP network with the voice DSCP. Put another way, the subscriber gateway should not trust the DSCP marking given to it by LAN-connected IP phones.

3.9 IP ADDRESS DOMAINS

It is quite likely that the voice and data services on a converged network may use disjoint IP addressing domains, especially where the service is offered wholesale to multiple ISPs.

As the subscriber gateway is typically attached to a both networks, it must be capable of routing packets between the IP address domains – both of which could be using locally administered IP address ranges rather than public IP addresses. This places some additional requirements on the functions supported in the subscriber gateway in order to manage the interface between the two domains.

3.9.1 Firewall, NAT and Proxy Function

The subscriber gateway must provide firewall, network address translation, and proxy function between the VoIP network and the locally attached LAN.

The firewall must prevent all packets except signaling messages and voice packets from locally attached IP phones from entering the VoIP network.

SIP has built-in support for proxy function. The SIP proxy accepts requests from the SIP client and forwards them, after some translation. This function is required in the subscriber gateway to provide address translation.

Similarly other protocols require address translation, including RTP, MGCP, Megaco and H.323. Typically these protocols will need to be added as Network Address Translation (NAT) applications, although it is notoriously difficult to apply NAT to H.323.

3.10 FAX AND MODEM SUPPORT

The PSTN supports fax and modem calls, and is very reliable. Calls connect on almost every attempt and rarely fail. A VoIP network must provide a similarly reliable fax and modem service. However, fax and modem traffic imposes some additional constraints beyond voice traffic.

Compared to voice traffic, fax and modem traffic is much more sensitive to packet loss but less sensitive to overall delay. In addition, lower-bit-rate codecs are optimized for voice traffic and cannot transport fax or modem traffic.

T.38 defines how fax can be sent in an IP network as pure data, independent of the voice traffic. However, it is a relatively recent standard and requires the use of either a T.38 capable Subscriber Gateway or fax machine.

Alternatively, fax and data can be supported successfully over an IP network by switching to a high bit-rate codec (such as G.711). The media gateways need to detect a fax or modem call, monitor for a 2100 Hz answer tone and switch to G.711 for fax or modem calls. Silence suppression and echo cancellation may also need to be turned off.

Note that the detection and switch to G.711 needs to be performed in a timely manner, to allow the fax / modem to train at the highest possible data rate.

3.11 AUTO-CONFIGURATION

One significant difference between a POTS (plain old telephone service) network and a VoIP network is that intelligent subscriber gateways now reside on the customer premises. These complex devices need to be configured, unlike a POTS phone, so auto-configuration of subscriber gateways becomes important as the network scales up.

The configuration requirements for a subscriber gateway vary depending upon the network architecture, but can include the following.

- IP address, subnet mask and default IP gateway.
- Name of primary and secondary call agent.
- Call progress tones, including frequency, cadence and power.
- Analog line configuration, e.g. loop start/ground start, voltages on line.
- SNTP server address and time offset.
- Trivial File Transfer Protocol/ File Transfer Protocol (TFTP/FTP) server address.
- Virtual LAN ID if using Ethernet.
- ATM PVC parameters if using ATM.
- Frame Relay PVC parameters if using Frame Relay.
- FTP user ID and password.

Some of these requirements can be addressed using DHCP, but others require some form of management interface such as SNMP, LDAP or UPnP (Universal Plug and Play).

Considerable work has been done in the DSL Forum to address auto-configuration of DSL equipment, but to date the issue of auto-configuration in VoIP networks has not been addressed.

3.12 COST OF SUBSCRIBER GATEWAY

While IP networking equipment usually is cheaper than equivalent TDM or ATM equipment, VoIP requires significant processing power to implement all the various protocols. Even compared to a VoATM network, the processor requirements for a subscriber gateway can be much higher, leading to a greater cost for VoIP-capable subscriber gateways. However, the cost of the chipsets required to implement VoIP is falling and is expected to continue to do so.

3.13 LOOP TESTING

The PSTN has extensive capabilities for remote line (loop) testing to minimize the necessity for technicians to attend street cabinets or customer premises. The testing capability includes both testing the line unit and the distribution wiring. A VoIP broadband access network must provide similar capabilities.

The line unit comprises the electronics that drive the loop. For a POTS line, the line unit includes the codec, the line hybrid that converts from four-wire to two-wire operation, the ringing generator and the loop current detector, which detects on-hook and off-hook conditions on the line.

The distribution wiring includes all wiring between the line unit and the customer's phone equipment. If the line unit is located in the customer premises, then the distribution wiring comprises only the wiring within the customer premises. If the line unit is located in a remote terminal, such as a street cabinet or vault, then the distribution wiring includes also the loop between the remote terminal and the customer premises.

PSTN remote testing capability addresses three distinct types of test.

- Distribution wiring test. This tests the wiring from the line unit (such as the distribution frame or DSLAM) to the customer equipment for safety, shorts, off-hook phones etc.
- Channel media test. This tests the path that the path between the line quality and the customer equipment performs to an acceptable quality level.
- Channel signaling test. This tests that the line unit can apply ring voltage and detect off-hook correctly.

The loop test procedures developed for the existing PSTN may not be directly applicable to a VoIP broadband access network, depending on the nature of the access devices and technologies chosen. Many recent DSL and line driver chipsets include at least a subset of the test capabilities outlined above, thus avoiding the separate test heads that characterized early DSL deployments. However standards work is still required to make the management of this capability easy in operational networks, for example via standardized SNMP MIBs. This work is currently underway in several forums, but is not yet complete.

In the meantime, a carrier building a VoIP Access network should check what line test capabilities are available via the element management system of the subscriber and access gateways they have chosen.

3.14 LAWFUL INTERCEPTION

Historically, lawful interception (wiretapping) of telephone conversations has been a relatively well-defined and straightforward process. Typically, a law enforcement agency applied to a court for an order to tap a particular phone number. Once the agency had the order, it served that order on the provider of the telephone service for the number to be tapped. The service provider then put a tap on the circuit, extracted all the necessary information and passed it to the law enforcement agency. The introduction of VoIP complicates this process considerably.

The law varies according to location (in the United States, the relevant legislation is the Communications Assistance for Law Enforcement Act - CALEA). The following requirements are typical.

- No wiretap is permitted without a court order (although this is not true in all countries).
- Wiretaps apply to phone numbers, not particular suspects.
- Wiretaps fall into two categories.
  - Call detail – a tap in which the details of the calls made and received by a subscriber

  are passed to the law enforcement agency. (Referred to as pen register and trap and trace in the U.S.).
  - Call content – a tap in which the actual contents of a call are passed to the law enforcement agency.
- The suspect must not detect the tap, so the tap must occur within the network and not at the subscriber gateway. Also, the tap may not be detectable by any change in timing, feature availability or operation.
- A suspect may be tapped by more than one agency. The taps are separate, and the various agencies are not aware of each other's taps. The taps do not have to be of the same category.
- It is the responsibility of the telecommunications carrier to originate or terminate calls to provide lawful interception.

As described in section 2.1, VoIP networks typically contain separate call agents and media gateways. The call agent is responsible for all call control and is the element that collects all the details of the calls required in a call detail tap. However the call agent does not see the call content, so call content must be collected elsewhere in the network.

The requirement to be able to tap the content of calls leads to the conclusion that all calls, whether they remain within the carrier's IP network or access another network (e.g. PSTN) must be routed via a device capable of duplicating the content and passing it to law enforcement. This function is not currently available in access concentrators and edge routers, which are normally unaware that the IP traffic is voice. Therefore, an access gateway is required as shown in Figure 1.

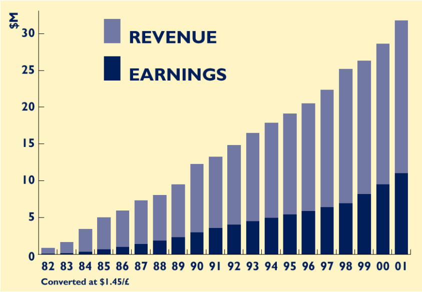
4. ABOUT METASWITCH

As a division of Data Connection, MetaSwitch leverages over 20 years' experience supplying communications technology and support to the leading service providers including Verizon, SBC and BT, and major equipment vendors.

Our VoIP expertise is derived from success developing world-leading products including the core protocols (MGCP, Megaco/H.248, MPLS, IP Routing, SIP, …), applications (unified messaging, conferencing, …) and next generation switching technology (MetaSwitch VP3500). The MetaSwitch VP3500 is easy to deploy and enables carriers to deliver reliable, toll-quality VoIP services in the access and backbone networks, with a full range of Class 5 subscriber services and PSTN interoperability.

Data Connection is a relentlessly profitable and stable private company, creating a basis for long-term investment and growth that ensures our ability to fund ongoing product investment and deliver first-class customer support.

MetaSwitch has offices in Alameda (California), Dallas (Texas), Reston (Virginia), and Enfield (North London), Chester and Edinburgh in the UK.



Converted at $1.45/£

For further information on how MetaSwitch can help service providers implement a successful VoIP migration strategy, visit **www.metaswitch.com**.

5. CONCLUSION

IP is ubiquitous and cost-effective. As shown in this paper, it can be successfully utilized in broadband access networks even to offer full PSTN-equivalent services – provided that the network is designed carefully to match the needs of the target service set. By analyzing their needs against the network design choices described above, a carrier can

- deploy new converged voice and data services
- remove the need to manager separate voice and data networks
- utilize cheaper IP-based backbone equipment to carry voice
- reap the benefits of a standards-based and highly flexible network architecture, giving a competitive market between equipment vendors and a wide range of equipment for different market niches.

As an example, a carrier wishing to migrate their current PSTN infrastructure away from TDM-based equipment that is rapidly becoming obsolete could deploy the following network design today:

- Services supported are POTS phones, connected to DSLAMS or DLCs that convert the analog traffic to VoIP flows in the core, and packetized voice for second-line services from residential DSL IADs.
- Header compression is implemented between the IAD and the network edge (for example, the DSLAM, router or CMTS), using Diffserv to prioritize the voice in the access network and MPLS to transport voice across the core.
- The IADs and DSLAMS (for POTS customers) are controlled using MGCP or H.248, with all calls routed via an edge access gateway to allow for intercept services.
- The voice network is a private network. The IADs are not trusted and hence are authenticated by the call agent using IPSec.

There is still a lot of work to be done, but it does feel as though the standards and implementations of VoIP are reaching maturity. The industry need is evident, and the equipment to meet that need is ready. We are already seeing VoIP deployments move from secondary to primary lines, and to carrier scale deployments. VoIP is set to take over the mantle of TDM in the long-term evolution of carrier voice networks.

GLOSSARY

AAL2	ATM Adaptation Layer 2, a media-bearing protocol for Voice over ATM
AAL5	ATM Adaptation Layer 5, a media-bearing protocol for Voice over IP over ATM
BLES	Broadband Loop Emulation Service, the standard for Voice over ATM/DSL signalling
CALEA	Communications Assistance for Law Enforcement Act
DLC	Digital Loop Carrier
DSCP	Differential Services Code Points defined by Diffserv and used in QoS
DSL	Digital Subscriber Line
DSP	Digital Signal Processor or Processing
FTP	File Transfer Protocol, an IETF defined IP file transfer protocol
GR-303	A digital signalling protocol used between DLC and Class 5 switches in North America
H.248	An ITU protocol for media gateway control, equivalent to Megaco
IAD	Integrated Access Device, customer premises equipment providing DSL data and voice connectivity
Megaco	An IETF protocol for media gateway control, equivalent to H.248
MF	Multi-Frequency signaling
MGCP	Media Gateway Control Protocol
MIB	SNMP management database, defined by the IETF
MTA	Multimedia Terminal Adapter: customer premises equipment in a cable network
NAT	Network Address Translation, an IP address translation technique used in firewalls
NCS	Network Call Signaling, adopted by CableLabs as the basis of the PacketCable VoIP standard
NEBS	Network Equipment Building Standards
NGN	Next Generation Network
POTS	Plain Old Telephone System
PSTN	Public Switched Telephone Network
PVC	ATM Permanent Virtual Circuit
RTP	Real Time Protocol, a media-bearing protocol for Voice over IP
SIP	Session Initiation Protocol
SS7	Signalling System 7
TDM	Time Division Multiplexing
TFTP	Trivial File Transfer Protocol, an IETF defined IP file transfer protocol
V5.2	A digital signalling protocol used between DLC and Class 5 switches outside of North America
VoB	Voice over Broadband