M. Ali Salman and Terry Hodgkinson

The Future of IP–PSTN Interworking

As the Internet continues to grow at a phenomenal rate, there is an emerging desire for it to be able to support voice communications, sometimes referred to as* voice over IP *(VoIP). Although the Internet is some way from being able to support carrier-class voice quality, interworking of PSTN and VoIP is currently receiving much attention. The approach being adopted at the moment is the use of a media gateway box, which enables carrier PSTN standards to be interworked with VoIP protocols. In the longer term, however, and depending upon the actual traffic mix and volumes, such an approach may lead to an expensive and unscalable global telephony service. The particular issue addressed by this paper is the need to understand the technical and economic pros and cons associated with interworking different telephony service standards rather than developing a new single global standard.

The aspects this paper concentrates on include: the current approach of VoIP provisioning based on gateway/gatekeeper interworking and the limitations presented to the network in terms of cross-network signalling. It will compare the cost and complexity to other evolving networks such as public and corporate PSTN and mobile networks. The paper will examine how the current growth trends, which are being fuelled by data traffic, may influence future network architectures and protocols and hence the VoIP implementation options.

Introduction

Traditionally, voice services, or *telephony services* as they are more commonly known, have been carried over circuit-oriented networks. These networks have been evolved and optimised on the assumption that telephony would essentially continue to remain the predominant service carried. In contrast, the Internet, which is based on connectionless networking principles and uses the TCP/IP (transmission control protocol/Internet protocol) protocol suite, has been evolved on the assumption that it would always carry a very diverse range of services. It was inevitable, therefore, that the Internet would eventually aspire to provide voice services, and it is no surprise that this has now happened.

Dr. M. Ali Salman:
British Telecommunications plc
Room 136, B29,
BT Adastral Park, Martlesham Heath,
Ipswich, IP5 3RE
UK
Tel: + 44 1473 642509
Fax: + 44 1473 643906
Email: ali.salman@bt.com
Dr Terry Hodgkinson:
British Telecommunications plc

This capability to support voice services is commonly known as *voice over IP* (VoIP). VoIP is currently in its infancy and its real value will only become apparent once the new wave of Internet voice-related services have matured. It is predicted that as VoIP standards become more established voice-related services will become increasingly prolific. The accuracy of this prediction will have to be judged by the passage of time. But assuming it to be true, it is likely that there will be a virtuous circle of improving the network to support new voice services which will inspire newer services that will stimulate further network improvements, and so on. The final outcome of this will eventually be a network that is optimised for supporting a diverse range of applications of which VoIP will be a sub-set. This is the complete opposite to how the traditional telephony network has evolved.

However, for the foreseeable future, the dominant issue will almost certainly continue to be interworking VoIP with traditional telephony services. It is the aim of this paper to indicate why this will be so and to highlight the interworking issues that will be encountered. It will then look at ways for improving/eliminating the need for interworking, and use these observations to identify how to seamlessly remove all need for interworking in the 'ideal' future network.

What is changing?

Traditional telephony networks are typically described by the following four key points:

- based on homogeneous technology end-to-end;
- homogeneous service; that is, telephony;
- other services spoofed to look like telephony; that is, use of modem for carrying data; and
- well understood traffic characteristics and usage patterns.

However, with the explosive growth in the Internet, user expectations and demands will change. For the future the typical network is more likely to be described by the following three key points:

- based on heterogeneous mix of technologies end-to-end;
- heterogeneous mix of applications/services; and
- unpredictable mix of different types of traffic statistics and usage patterns.

The increasing heterogeneity and unpredictability will demand that networks should be made as independent of services and applications as is practicable. The current plans for application programmable interfaces (APIs), capable of opening up the

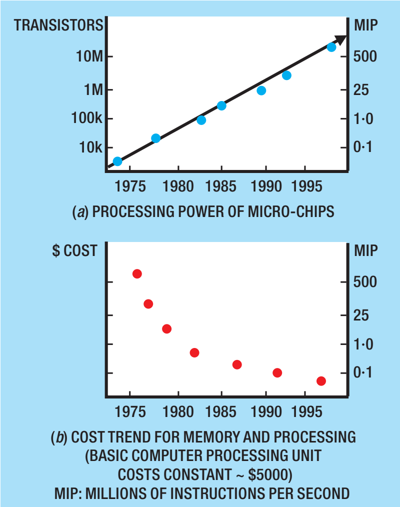


Figure 1—Moore's Law (source: http://www.intel.com/intel/museum/25anniv/hof/moore.htm)

network to various applications, may help, but the key feature will be to migrate all application/service-related features into the host machines. Traditionally, the cost and the extremely limited processing power of the host have prohibited this. However, this is unlikely to be the case in the future as memory and processing power costs continue to fall. This is especially true for processing power despite the fact that the processing power continues to increase according to Moore's Law, see Figure 1 (data from Intel web site). The network simplifications offered by migrating all application/service specific features into host machines will be considered in more detail later.

Typical Network Scenarios for VoIP

The incumbent public switched telephone network (PSTN) will have to be able to interwork with new VoIP data networks if it is to offer its users global connectivity. In other words, if someone using a telephone wants to speak to someone using a PC with voice-service capabilities, global connectivity cannot be claimed if the PSTN prohibits such communication. To offer global connectivity, special interface equipment is needed, namely *signalling gateways*, *media gateways* and *media controllers*: these are discussed in more detail later. These gateways effectively provide the necessary features needed to support PSTN and VoIP interworking. A further reason for interworking is the emergence of different mid-band and broadband network services, carrying mainly data, requiring access through both PSTN and the Internet.

Figure 2 shows the main two scenarios of interworking in VoIP:

- core network, and
- standard PSTN access interworking with the IP network.

Interworking the core networks has different requirements from interworking in the access networks. Depending on the size of the network, and how extensive the VoIP is going to be, the proliferation of VoIP interworking is likely to be in the access networks due to the fact that core networks only deal with aggregated traffic.

Interworking IP and standard PSTN equipment requires the installation of additional servers, gateways and control equipment. These gateways and the features they support are described below.

Gateway Features for Core Interworking

The interworking of IP and PSTN has made it necessary to increase the complexity of the signalling protocols in order to provide call and bandwidth control across the two networks. Figure 3 depicts the current signalling architecture for network elements interfacing PSTN and IP networks.

User access to the PSTN is carried by Signalling System No. 7's (SS7's) user-to-network signalling interfaces (UNI), through to the signalling point on the IP network side. ISUP information is carried between the two networks to provide features for network-to-network Interfacing (NNI).

As for deployment, the biggest market for VoIP in the short-term is expected to be in the intranet and virtual private network (VPN) domains where bandwidth provisioning is more easily managed and controlled. The issue of quality of service (QoS) for PSTN voice over the IP network will remain the major barrier to VoIP deployment in the public network.

Signalling gateway

This is the box that performs the mapping of SS7 signalling formats

Figure 2—High-level architecture for current VoIP

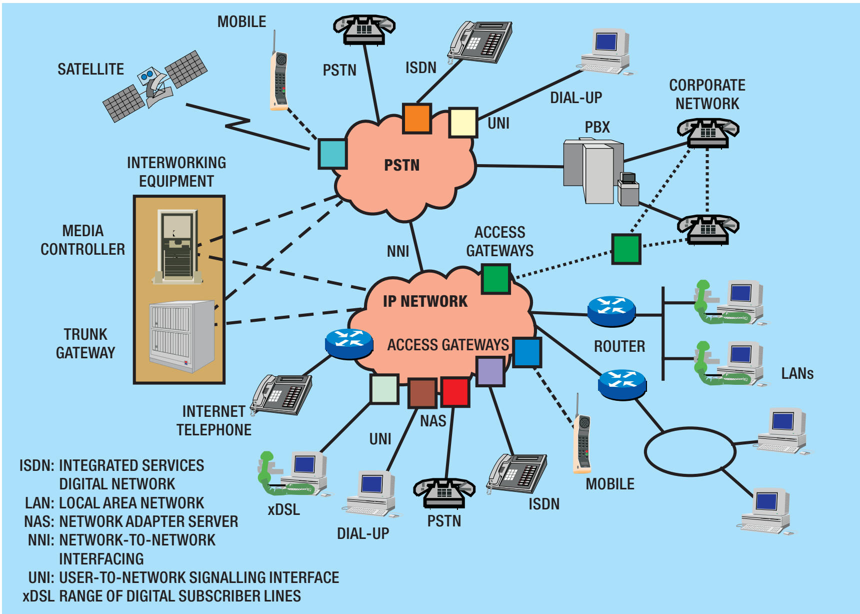
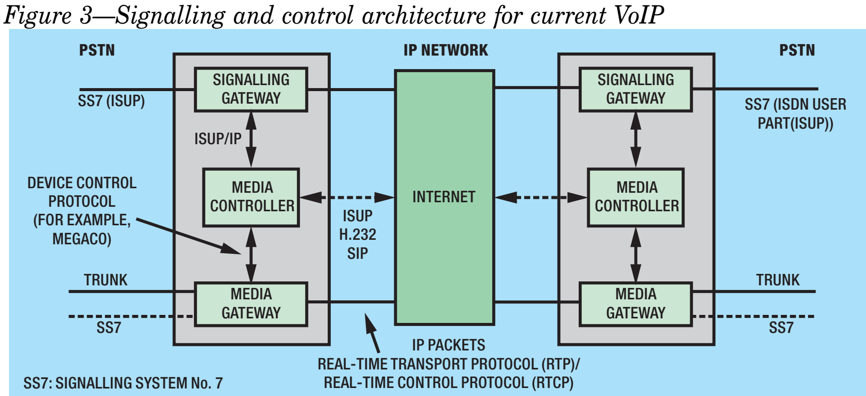


Figure 3—Signalling and control architecture for current VoIP



into IP compatible formats. SS7 is a global common-channel signalling standard defined by the International Telecommunications Union (ITU-T). It defines the protocols by which PSTN network elements exchange information over a digital signalling network for both fixed and cellular networks. The signalling network is used primarily for the following functions:

- call set-up and clear down;
- efficient use of voice circuits;
- intelligent number processing—local number portability, (0800) services etc.;
- intelligent services—call forwarding, call waiting, three-way calling etc.; and
- authentication and roaming services for mobile communications.

ISDN user part (ISUP) protocols are used to set-up, modify and release (among others) trunk circuits that carry voice and data between the calling parties. Signalling connection control part (SCCP) provides global title translation to associate an address (for example, dialled 0800 number) with a destination signalling point code and an application unique number. SS7 messages can be encapsulated (*tunnelled*) within IP and used by the signalling controller for routing within the network. The media gateway looks up the SS7 contents and establishes the required transcoding scheme to be employed.

The PSTN is still a cornerstone for the current VoIP architecture. The SS7 signalling used in the PSTN is also taken as the de facto signalling that new VoIP equipment will have to interwork with. Therefore, the claim that 'VoIP will cause the death of the PSTN' is still far from accurate. The PSTN, its core switches and many of its access technologies will still be used for the foreseeable future.

Media controllers (MCs)

Gateway controllers perform the majority of the call control and intelligence, to aid call progression within the network. For example, H.323 gatekeepers perform telephone-number-to-IP-address mapping and can act as platforms to run services for VPN applications. Media controllers implement one or more of a set of developing protocols: H.323, SIP, MGCP etc.

The main functions provided by these elements include *address translation, connection control* and *bandwidth management.* The use of these boxes has proven most useful in the corporate and leased lines and scenarios where bandwidth is more easily controlled.

Signalling between media controllers can be carried using ISUP with a special extension for IP, using the H.323 protocol with extension for PSTN or using session initiation protocol (SIP)[4] with extension for PSTN. This is an active area of discussion within the Internet Engineering Task Force (IETF), where the standards have not yet been fully agreed.

Protocols such as ITU's H.323 are commonly used in media controllers to provide mapping and various call set-up control functions between network servers. The IETF is proposing a new equivalent 'lighter weight' SIP to perform similar functions.

Media gateways

Gateways are end points in the network providing ports that connect to IP and circuit switched network users. Currently, each customer needs two ports, in order to be able to send and receive. Gateways can also carry out translation functions between video and audio codecs etc. Gateways are mainly used to establish links between PSTN terminals and IP or ISDN-based terminals (among others). Media gateways provide basic functions necessary for analogue access (UNI) and network-to-network (NNI) for core interworking. These functions include:

- *IP encapsulation/de-encapsulation* To traverse from PSTN to the IP network, ISUP is carried over IP between the signalling gateway and the media controller. This uses the ISUP information to issue control signals to the media gateway using protocols such as MGCP. Control functions may include connection creation, connection clear down, notification information etc. The media path that connects one gateway to another is used to transfer voice data from one end box to another using protocols such as RTP/RTCP protocols.
- *Encoding of analogue signals* This includes the digitisation of analogue signals and the packetisation of the resulting data.
- *Supervisory tones* The network user needs continued feedback on the progress in the call set-up stage. Dialling tones, busy tones, line-down tones etc. are generated according to the status of the call. Tone codes are provided by the switches which connect the various gateways.

In brief, the combination of the media gateway and controller provide the key features that have traditionally been placed in the network. In order to establish interworking between two different networks, encoding is essential due to inherent analogue access. Encapsulation is also needed due to interworking two different signalling standards (for example, SS7 and IP). Supervisory tones are generated in the gateway to compensate for the local exchange functionality. These features have been replicated in order to facilitate the communication between the various network systems. However, the ultimate price is paid in making the network service dependent. This will limit future service deployment and the flexibility in enriching current service features.

Network Futures

Basic network functions

Migrating certain network features into the host machine will enable the network to be service independent. This requires that there is minimal interworking and that supervisory tones are generated in the host machines. The key issues for PSTN and VoIP interworking are those of IP packetisation, common addressing, supervisory tones and signalling interworking. The impact of migrating these features into the host is designed to reduce the need for the gateway and its features, thus eliminating the need for interworking. This section will give an outline on how this could be achieved.

IP packetisation

Assuming that eventually all host machines will send a digital stream into the network rather than an analogue one. The need for IP encapsulation within the network will not be a network requirement. Moving these functions to the host side will enable their use by new applications, without much impact on the network functionality. This would also eliminate the encapsulation function from the media gateway.

Encoding

For the same reason as above, encoding can now be done by the host, although host-to-host communication may take place to determine what to use. This does not, however, involve the network. Signalling and messaging functions such as those carried out by H.323, SIP, TCAP etc. can now reside in the host, eliminating the need for them to be in the MC, or in other network nodes.

Service specific supervisory tone handling

These functions can be moved to the host side, enabling new applications to use a richer set of tones. These could now be audible, visual, text or any other form. Historically, tones were placed in the network nodes because that was the only way to use them. Modern server-based applications could generate the tones, as configured by the end user. Moving supervisory tones to the host eliminates another major function that media gateways provide. However, a 'unified' set of supervisory codes remain to be provided by the network, and may be carried by signalling messages.

Signalling interworking

Current developments in IP and other network technologies require a fresh look at providing a global 'lightweight' signalling system. Instead of 'interworking' technologies, there is now a clear need for an 'internetworking'[2] signalling solution that spans across all technologies. The new architecture will provide a network level intelligence that is built around 'general-purpose' servers that are independent from the network switching fabric. The future network can be based on heterogeneous technologies but still needs to be controlled by a 'global' homogeneous signalling system, sparing signalling mapping that occurs at every interface point in the network. Current network trends show that every new technology tried to be the universal solution for today's problems. This does not work, requiring us to think of ways to 'internetwork' different technologies rather than 'interwork' them.

It is, however, understood that it is more difficult to design a 'universal' signalling system, given existing network complexity. This remains essential if the network 'future proof', 'cost' and 'complexity' are to be kept under control.

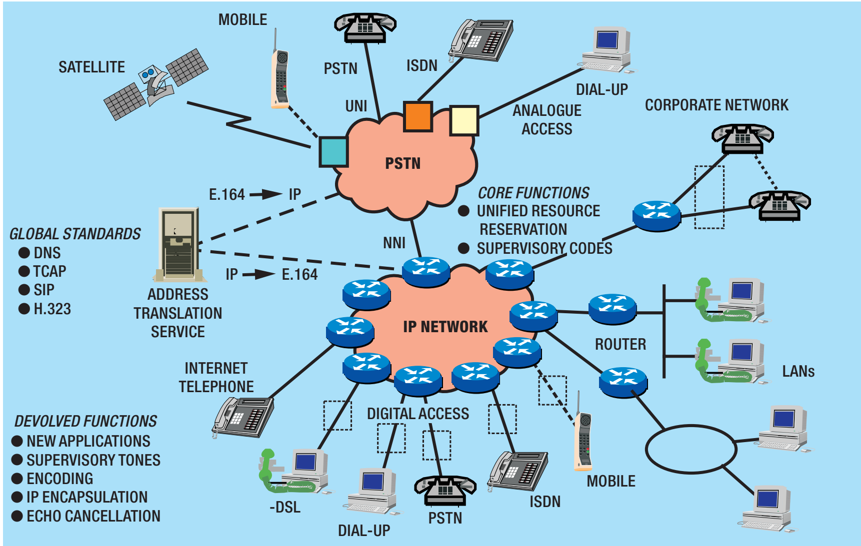


Figure 4—An example of universal 'multi-service' heterogeneous network

Basic signalling functions

Comparing IP protocol functionality with that of SS7 and that of asynchronous transfer mode (ATM) reveals that there are close similarities between the various stacks. The major differences are in the layer at which certain functions have been assigned. There is enough evidence that these signalling stacks may converge to provide one globally accepted standard. Salient features of such a standard include:

- *Messaging for call control and service association* A 'global' addressing scheme that is capable of locating network nodes and servers. Messaging systems such as SS7's SCCP, DNS and SIP all attempt to achieve similar functions. TCP-like functionality (error checking, re-transmission etc.) can be moved to this layer in the network. MTP2 in SS7 signalling will not need to do re-transmission as it is done at the TCP level. Applications will choose between the various messaging systems, resulting in the need to keep several systems at the user host machine. Different services can also create their own formats, but all will be encapsulated through one layer (the IP layer). IP does not make assumptions about the underlying technology.
- *Resource allocation* A network-wide reservation system will provide resource allocation functions. RSVP-like signalling may be used to provide reservation control for the generated traffic. Packets could also carry enough information to the network about their resource requirements on a hop-by-hop basis. Re-transmission of failing packets will have to be handled correctly, by higher layers.

Resulting network

It is quite plausible that by migrating functions such as re-transmission and error corrections to the TCP layer and functions such as address resolution protocols to the session layer we could provide service features that are not affected by the underlying network. IP can be used as the sole 'internetworking' protocol carrying all traffic (encapsulating or otherwise) to 'technology independent' lower layers. This 'functional separation' will allow any service feature to be carried by any underlying network technology using a 'single' universal signalling standard. Necessary information may be carried by packets to enrich the way this is finally implemented. Figure 4 depicts the way some of these changes may be achieved:

The future network (as depicted in Figure 4) has mainly devolved functionality between the network and the host. Functions such as IP encapsulation, encoding, tones etc. are now performed by the host eliminating the need for gateway devices such as gateways and MCs. Hosts can now connect directly to a router network. Duplication of 'messaging' functions at the user host may become inevitable, in the same way that computer applications now share a desktop PC. Interworking various network technologies will not be an issue any more.

Cost of Interworking PSTN and VoIP

Current arrangements for VoIP include the use of at least three networks (PSTN, IP and the SS7 signalling). This implies the use of at least three sets of standards, interworked to provide a voice service. The complexity in terms of network elements, numerous protocol stacks, signalling and stove-piping arrangements puts hurdles in the way of building a cost-effective network. When considering the broadband data network solutions, the number of interworking networks can mount to dozens. If not addressed, complexity may be the biggest factor in hindering the expected take-up of VoIP.

A simpler network signalling arrangement will result in savings in money and complexity. The internetworking proposals in this paper should help towards bringing in reasonable saving through eliminating intermediary devices and introducing a 'unified' signalling stack. This will encourage re-use of resources within the network, and simplify resource reservation across the network. The penalty that will be paid, is the duplication of software addressing and messaging schemes. We foresee that several of these applications (for example, SIP, H.323, DNS) will form a global standard and will co-exist in the host station.

The future network will serve all existing telecommunications applications and will be simple enough to provide seamlessly interconnected networks, majoring on IP and capitalising on existing network infrastructure. The move towards a multi-service network will yield added technical and economic benefits. The network operator should not worry about extra operation, management and investment every time a new network technology is introduced. This architecture introduces a single user box (for example, a router) to provide network access, including QoS, security, and reusability of resources.

Conclusion

Digitisation at the host into IP packets removes the need for gateways to encapsulate into IP packets. Use of a single global addressing scheme removes the need for layer 3 address translations. The telephone number can still be used but this would now have become a layer 4 address and effectively it is an alternative form of uniform resource locator (URL).

Having removed the need for IP encapsulation and address mappings, the only feature now provided by a gateway is signalling interworking. However, because layer 4 signalling messages can be evolved independently of the network—for example, DNS, SIP, TCAP—and other features of user parts not associated with resource reservation, they can be ignored from the point of view of network interworking issues. SS7 becomes recast in a use as 'fit-for-purpose' collection of protocols that are defined independently. For example, interworking DNS and SIP is not an issue for the network and they can co-exist completely independently. The basic reason for duplication is that this communication is purely server–server interaction, and does not require the network to have knowledge of the messages carried.

Because the fundamental principles of IP networking make no assumptions about underlying technologies, the technology-specific aspects of SS7 standards become redundant and can be discarded. All that remains is resource reservation and supervisory code information. By adopting a single, end-to-end globally accepted resource reservation protocol and supervisory messages, all need for signalling interworking is removed. Service specific tone generation would be located within host machines with the appropriate mapping between the supervisory message and the tone/visual display now being under the control of the host application. At this point the gateway becomes redundant and PSTN telephony and VoIP have converged to become the same.

References

1 Comer, Douglas E. Internetworking With TCP/IP. Vol 1: Principles, protocols, and architecture. Prentice Hall, Inc., Upper Saddle River, New Jersey 07458, USA.

2 Peterson, Larry L.; and Davie, Bruce S. Computer Networks, A Systems Approach. Morgan Kaufmann Publishers, Inc., San Francisco, California, USA.

3 Halabi, Bassam. Internet Routing Architectures. New Riders Publishing, 201 West 103rd Street, Indianapolis, IN 46290, USA.

4 Schulzrinne, Henning; and Rosenberg, Jonathan. Signalling for Internet Telephony. Department of Computer Science technical report CUCS-005-98.

Biographies

Dr. M. Ali Salman
British Telecommunications plc

Dr. M. Ali Salman gained his first degree in Computer and Communications from Essex University in 1984. He obtained a Ph.D. in adaptive systems from Loughborough University in 1988. He joined BT Laboratories in 1989 to work on traffic and SS7 management, utilising object-oriented knowledge-based systems. He joined the Intelligent Systems Unit in 1992 to develop network design solutions. In 1996, in the Network Technology Centre, he investigated issues of Internet protocols and contributed to the VoIP investigations. He recently moved to the Internet and Data Networks Centre focusing on IP solutions for future core networks.

Terry Hodgkinson
British Telecommunications plc

Terry Hodgkinson joined BT Laboratories in 1975, and since then has been involved in many aspects of optical communications and data networking research. From 1975–92 his research covered various aspects of optical transmission, particularly coherent optical detection methods, fibre polarisation effects and fibre amplifier performance modelling. Since 1992, his research interests have concentrated on data networking, and IP and ATM in particular, and how they might converge. Currently, he is leading a team researching the issues beyond IP/ATM networking as they are known today. The key aim is to identify new networking philosophies/technologies/architectures for achieving very efficient and flexible future networks. He has been awarded a D.Sc. Higher Doctorate and jointly awarded the ECOC '82 prize and the 1984 Rank Optoelectronics Prize. He has written 66 conference/journal papers and three book chapters and filed 10 patents in the fields of optical communications and data networking. He is a Member of the Institution of Electrical Engineers (IEE) and a Chartered Engineer.