Everything over IP — an overview of the strategic change in voice and data networks

S A Karim and P Hovell

This paper investigates the changes that are required in networking technology for ‘everything over IP’ to become a reality. Initially the changes that are taking place in the telecommunications industry are reviewed. This review ranges from a discussion about the companies installing massive global IP networks to the emergence of novel routeing technologies, e.g. multi-protocol label switching (MPLS) and terabit router technologies. The role of existing telecommunications operators is then discussed, along with the reasons why they are developing interworking and intelligence layers based on distributed computing principles to support all their networks — mobile, fixed, broadband and IP.

1. Introduction

In the 1990s, IP, the Internet protocol with its background in the US military, became the single most important driving force within the computing industry. IP has now engulfed the telecommunications industry and is emerging as a major driving force in all areas of future telecommunications development and deployment.

Imagine a network that can be built and dimensioned to carry all domestic voice traffic in a single industrially developed country, and a single operator being able to carry such traffic with both the ability to undercut the incumbent operator, and the modern technology to be able to combine all voice and data traffic on a single network. Such a dream is a modern day reality [1].

We are now in an environment where seamless multimedia and information transport to all business and home users is becoming a part of our everyday communications requirements.

In the past decade we have seen the technological advances in optical transmission. The introduction of wave division multiplexing (WDM) and dense wave division multiplexing (d-WDM) has given a bandwidth explosion to existing operators who have been supplied with the ability to replace ageing ‘out-of-capacity’ networks; it has also given new operators the chance to break into a once closed market. This is already happening around the globe. IP is creating that driving force.

In the telecommunications industry the hype surrounding IP-based services has brought the concept of voice and data to be combined on to a single network. This then brought the startling assertion that the growth in the technology was moving so rapidly that the Internet could swallow not only voice telephony but the whole of telecommunications, and then came the scare stories of how the connection-oriented circuit-switched network was doomed, and only had a three-year life span, as IP routers and perhaps terabit routers would do the job better and faster. The reason for this was that the future of telecommunications would be all-IP networks, connected by high-speed routers instead of 64 kbit/s narrowband voice switches.

The basic functionality of the Internet has been the WWW and a browser for effective electronic publishing of information and electronic distribution of media. The later sections of this paper will show what is driving the future networks that are being built today, the developments that are underpinning the growth, the router technology and network expansion of the new players, and what future networks could look like and what services they could offer.

2. Current networks

What is not beyond doubt is the need for the existing circuit-switched networks to change with the changing traffic they have to carry. On the traditional PSTN, call-holding times have increased beyond any network planner’s wildest imaginations. Dial-up IP access has created severe network congestion — this is very acute in North America with local calls being free. In Western Europe such issues are becoming problematic and the circuit-switched network has simply never been designed to carry bursty IP traffic with circuits permanently left open. The ability to have an all-IP network, removing the 64 kbit/s time-slot, stops the problem of network congestion. The ability to deliver to a large corporate organisation an intranet able to encompass all IP data traffic, and then carry, on that same network, all voice traffic, brings down the cost of ownership and reduces the cost of network management. In fact, the term 'large-scale IP' has been coined, where the network is all IP, i.e. today's Internet but with the ability to carry telephony as well.

Figure 1 shows the structure of the Internet today and the relationships between the customers, the Internet service providers (ISPs), and the parts with which we are mainly concerned, the transit networks. The customers have dial-up or leased circuits to the ISPs. The network operators who peer with other networks have invested large amounts in the network infrastructure to keep up with the demand and growth in traffic and the sheer volume of users. Clearly the Internet is very much USA-centric, and for European network operators large circuits are needed to connect Europe to the USA. Again, capacity planning in Europe has been difficult, with many operators having seen unprecedented demand for Internet connectivity.

It is this global IP network that is becoming the future communications network. As with many new technologies there are new players entering the arena that are building large transit networks; many of these companies have only been in existence for three or four years but within that time have established themselves as major forces within the Internet. For a traditional network carrier with voice traffic being its dominant business, IP and data is the brave new world. Big networks with large volumes of traffic will be run by the dominant Internet players of the future.

3. The brave new world

Future networks will have a plethora of new applications/ services:

- WWW — agents, push technology, homework, community computing,
- telephony, videotelephony,
- home services:

  — video on demand, multicast, staggercast, games, banking, home shopping, security, monitoring, audio distribution, video distribution, software, etc,

  — access to intranets,

  — global, home working, mobility,
- extranets — communities of interest, electronic commerce.

From the above it can be seen what types of service a new all-IP network would be expected to support/carry. What can also be seen is that many of the services/applications are new and not currently supported by today's conventional 64 kbit/s switched networks. The major consideration is that IP is driving the application development — simply look at the growth of the WWW, from a system in the early 1990s for academics to share information, to the modern way of buying books via a credit card. Security problems are no longer a worry to a lot of people. In less than five years the way we trade for simple items has changed. If trading for simple items is now so different, it is not too hard to see that the whole concept of communications is changing, and the way we use communications is clearly going to have a massive impact on the communications providers and operators. From this

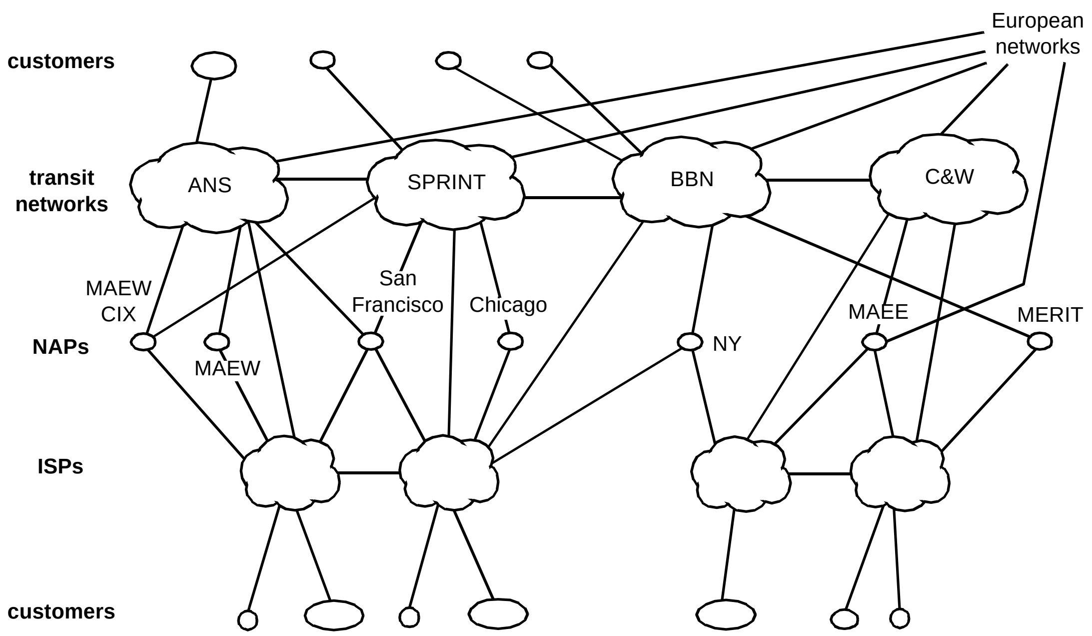

Fig 1 The Internet today.

changing landscape, the customers are demanding greater speed, more functionality and reliability, and, of course, they expect quality of service:

- bandwidth prices falling — customers expect double the bandwidth for the same price per annum,
- data dominating growth markets — growth in traffic will come from data rather than voice,
- Internet explosion — creating significant IP networks and innovative applications,
- internetworking costs falling — encouraging by-pass of traditional correspondent networks,
- capital distribution costs falling — data networks have lower capital, with costs that are equivalent to capacity voice networks.

Taking these factors into account, how does a network operator decide what technology to deploy, especially when the mind-set of most existing telecommunications operators is very much resolved towards connection-oriented, circuit-switched networks? One simple way is to see what is happening around the world, and compare the growth in data traffic to the growth in voice traffic.

Internet revenues [2] have been growing at more than 100% a year compared with core telecommunications businesses growth of 5% per annum. Data has been growing at approximately 20% per year compared with 8% for voice. Demand for Internet access is increasing at the staggering rate of 1000% a year — see Fig 2 for the growth in Internet hosts. By the year 2000, half of all the bandwidth in the world will be Internet traffic; the other half will be everything else, i.e. including voice. Extend that growth rate for another few years, and by 2003, the Internet will be more than 90% of the bandwidth and by 2004, more than 99%. You could argue we will not even know voice is in there except for its quality of service demands. In terms of bandwidth, voice will be insignificant. Of course extrapolating current growth can be dangerous and misleading, but even if the projections are slightly inaccurate, the same conclusion can be arrived at, but a few years later.

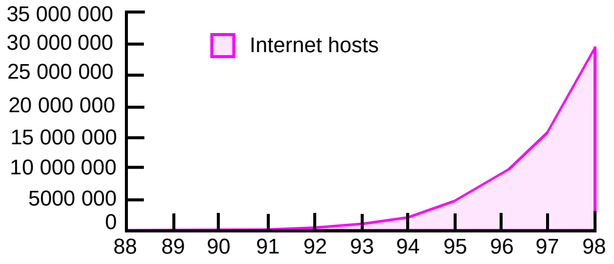

Fig 2 Growth in Internet hosts.

3.1 Current trends

There is a popular belief that the Internet works on a different time-scale to the rest of the telecommunications world — 1 Internet year = 7 telecommunications years. This is true in terms of technological change and commercialisation of the Internet.

New technologies will be able to meet such challenges, as mentioned earlier, with d-WDM systems and terabit routers.

However, it is important to see where the traditional business of real-time voice fits into the 'brave new world' of IP networks and the current maturity of voice over IP (VoIP) technology. As can be seen from Fig 3, the revenues from IP telephony are expected to rise dramatically over the next few years.

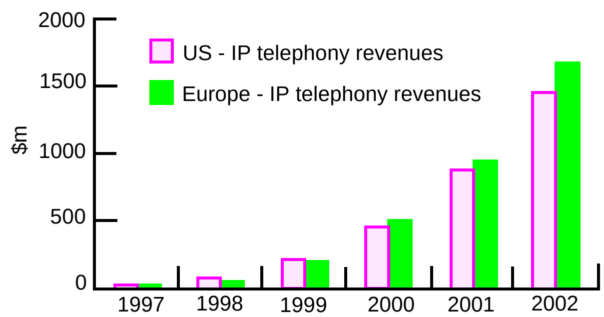

Fig 3 A view on IP telephony growth.

3.2 The IP place for voice communications

VoIP is not a concept of the 1990s; it was first put forward in the 1970s [3]. Over 20 years later all major software vendors, such as IBM and Microsoft, have bundled the voice-enabled IP software into their operating systems.

For real-time network communications, there is a requirement for synchronous delivery of the data, transmitted on a reliable circuit (a fact that we have taken for granted for the past 20 years). However, with IP technology, data is transported in packet format, i.e. it is inherently non-deterministic. Therefore, for multimedia traffic to be sent over an IP network, some additional protocols and some form of quality of service (QoS) within the network is required [4, 5]. The way QoS can be supported in an IP network is discussed later.

As voice and data begin to merge, voice needs to be maintained at the quality to which users all over the world have become accustomed. Long delays and poor quality is not acceptable to PSTN users. Some people consider voice to be the 'killer' application for the Internet. The introduction of Internet telephony gateways that connect the PSTN to the Internet, made by companies such as VocalTec

[6], enables the corporate LAN to be connected to the PBX and thus to the PSTN.

Also one has to consider that the telephone and PC have been rapidly converging for the past four years. Toll-quality voice requires a 100-MHz Pentium PC, but in reality most PCs sold today are at minimum 350 MHz and multimedia-capable. Thus nearly all PCs in existence are capable of good-quality IP voice communications:

- Ethernet moves from 10 Mbit/s to 100 Mbit/s, to 1 Gbit/s,
- PSTN/IP gateways @ $1000 per port and falling,
- PBX feature functionality built into the server.

As shown in Fig 4, one has to ask the question where does the 64 kbit/s circuit-switched phone sit in this arrangement? The corporate LANs/intranets have the PCs all networked, and the PCs are voice enabled. Traditional telephony as we know it is nowhere to be seen.

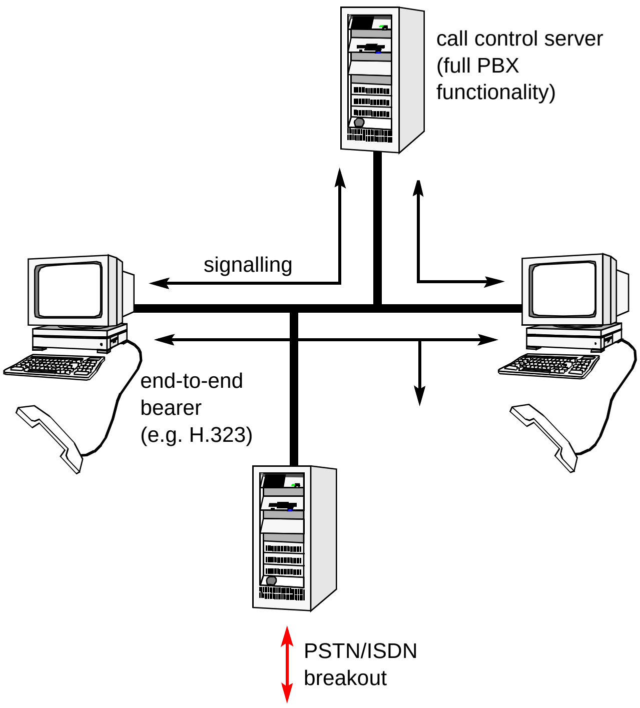

Fig 4 LAN/PBX convergence.

4. Technology

The technology used to build such networks comes from the data networking community with no influence from the voice transmission area; but with technology becoming more and more sophisticated, the machinery to move packets, the router, has been evolving with the protocols to carry voice traffic. It is by no coincidence that a lot of router companies have been bought by traditional switch vendors, and new router start-ups have received venture capital from switch vendors and network operators — perhaps to protect their core business of selling voice switches.

The router at the end of the day is a box (see Fig 5) that simply routes IP packets, similar to voice routeing, but it is totally connectionless.

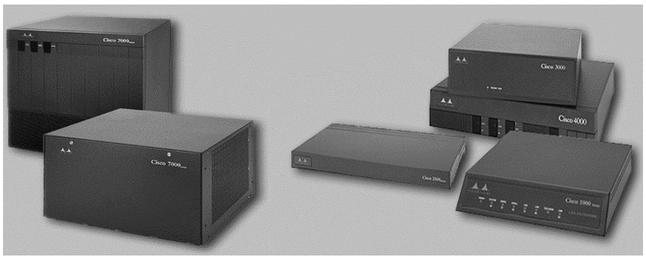

Fig 5 Typical router hardware.

4.1 *Quality of service*

For a number of years mechanisms to implement some form of quality of service in IP virtual private networks (VPNs) have been investigated. The idea of extending the RSVP mechanism [7], that was proposed for the LAN environment, into the WAN environment has more or less failed. The idea of requesting a certain quality of service from the network that would cause all the interconnection points (routers, etc) to co-operate and reserve the requested bandwidth is not really practicable when the connection traverses lots of different commercial networks. Currently the most promising mechanism for implementing QoS is for the customer to set the type-of-service (TOS) bits in the IP header to indicate the level of QoS the particular packet requires. The TOS bits can be set to indicate that the information is either real time (VoIP) or requires a set QoS, e.g. bronze, silver or gold. The TOS information is used by the access router to first ensure that the customer's data is within the agreed contract and then to place information with similar QoS requirements in similar output queues. The basic idea is that high priority/low latency information is given the highest precedence and hence gets priority for access to the transmission line. One technique that is currently being standardised is multi-protocol label switching (MPLS) (see Fig 6) or Cisco's [8] proprietary Tag implementation.

The basic idea with MPLS is that a number of label-switching routes are set up between the various routers in a network with each router having a label-forwarding table. On ingress to an MPLS network the destination address of the incoming packet is evaluated and a label added to the packet that indicates the next router and the priority of the packet. All packets to the same next router and with the same QoS requirements are assigned the same label. The output queue of the ingress router and the links between routers are engineered so that set amounts of bandwidth are allocated to certain labels and hence QoS-type packets. By this mechanism the packet travels from router to router with

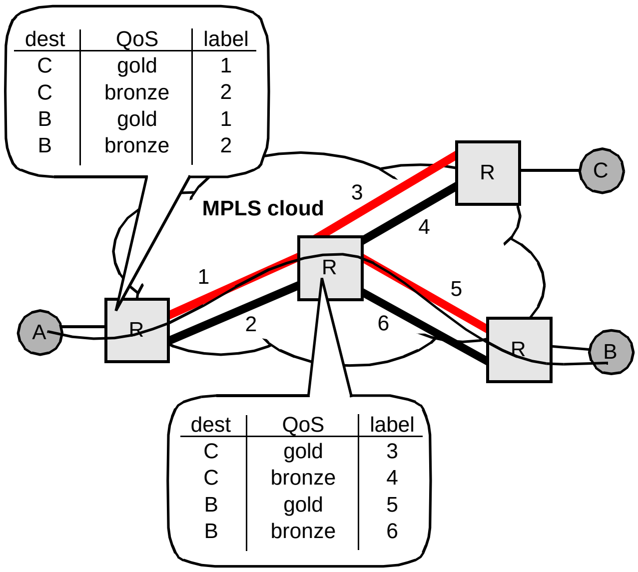

Fig 6 Example of MPLS network.

its label being swapped but always going via a traffic-engineered route for its particular QoS requirements. At the egress router the last label is stripped off and the packet is transmitted to its destination (as shown in Fig 6). This structure can also be used to separate traffic on different VPNs in that they go via different label routes that cannot interfere with each other.

4.2 *Technology onwards*

With QoS being so critical on bandwidth-short circuits, the need for resource reservation protocols is important. However, even with the arrival of ultra-fast circuits, RSVP may not seem so important when bandwidth is so freely available, but other issues, such as router speed, become critical.

Such technology is moving at the same rate as the Internet. However, the router market until very recently was not growing in line with the bandwidth market. The fear was that the routers would not be able to cope with the projected traffic growth. Cutting-edge router technology was just able to meet the 622 Mbit/s bandwidth challenge, but when you have fibre optics capable of carrying 3.2 Tbit/s, then 622 Mbit/s seems rather small. In fact, networks capable of carrying 10 Tbit/s are being designed around the world. Such network capacity five years ago would simply have been inconceivable. Yet today such plans are commercial reality.

At this point it seems appropriate to mention the terabit router. The ability to route vast quantities of IP packets is becoming a limiting factor, since, if the fibre is carrying terabits of data, congestion will build at the megabit router. Consider Moore's Law, which is from Dr Moore of Intel Corp (the desktop PC industry was driven by Moore's Law), which says that the computer's productivity/power relative to cost would double every month, and it does — this is why a computer is obsolete as soon as you buy it. Now, if we apply the same statement to the Internet (using as the analogy bandwidth demand or bandwidth growth), IP backbone bandwidth in general has been doubling, not every 18 months, but every three and a half months for the past three years, i.e. growth by a factor of 10, or 1000% per year. That scaling challenge has never been faced before. There has never been an industry or a technology (or anything) that has had that kind of growth curve. The increase in the speed of silicon was not sufficient to meet the demands of the Internet backbones, and the router vendor community has faced the dilemma for the past 18 months that eventually it would be impossible to route the traffic growth and the term terabit router became the buzzword of the technology venture-capital community.

Figure 7 shows the phenomenal growth in Internet traffic compared to PSTN. Development of conventional 64 kbit/s PSTN switches is relatively straightforward because of the slow growth of the traffic. Router development to handle Internet traffic is, however, far more complex because of the much faster growth in traffic. In fact, a technological leap is required to keep up with the traffic growth.

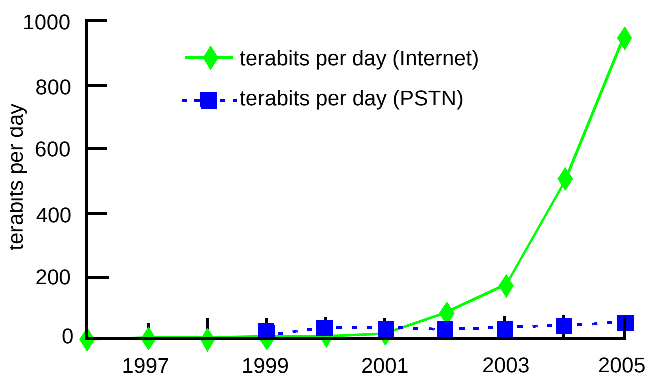

Fig 7 The reason why terabit routers are required for the Internet.

The potential need for such huge routers is due to the success in the Internet. Companies such as Pluris [9] and Juniper Networks [10] are technological research leaders in this new field. They received funding from switch vendors and other telcos bringing further credibility to the argument for the need for faster routers.

5. The future

Since the current challenges of scalable VPNs and providing QoS to support real-time services (VoIP) have been solved, the next big development within the IP arena will be to fully support mobility. Mobility comes in many flavours, from simple IP access via GSM (generalised packet radio service — GPRS) to the all-encompassing

Universal Mobile Telecommunications System (UMTS) environment.

With GPRS, if a user's terminal is switched on and within GSM coverage, they will have IP connectivity, the location management being performed within the GSM home/visitor location registers and the IP data being tunnelled to the active GSM base-station. The UMTS goal is far more all-encompassing in that it combines all access media and includes a 'virtual home environment' concept — you will receive your same service set whatever the access mechanism. In theory you can therefore use a fixed access, radio LAN, GSM, UMTS air interface, satellite, etc, and still receive your particular service set. Handover between these various access types will be possible, but the service may change in quality as you go from a high bandwidth access mechanism to one with less capability. The concept is that you could be in the office with several active sessions via a hyperLAN access, and then, as you go out of the building, you would be seamlessly handed over to, for example, a GPRS access mechanism. The lower bandwidth available from the GPRS access may cause degradation in the service quality of certain services, e.g. the size and frame rate of video may be reduced, or the coding standard of voice may change.

Terminals are another area where there is very active development. The recent announcement of Symbian [11], a combination of Psion [12] and mobile telephony manufacturers, is an example of the commercial drive to integrate portable computing devices with mobile phones. Other consortia, like Bluetooth [13], are actively working on cheap ($5), secure, short-range radio transmission equipment; volume silicon is expected by 2000. Another area of active development within the IETF is '*ad hoc* networking'. The concept is that all devices become routers and jointly form a network without any fixed infrastructure. Communication between devices is via whatever media is available — Bluetooth, infra-red, etc. Access to the fixed network would be via a device within the '*ad hoc* network' that has the capabilities and is within range of some fixed infrastructure.

Once all the above technologies are in place several new scenarios are possible. Small portable devices capable of all computing and voice communication needs will be available. These devices will use whatever communications structure is available. This could be an '*ad hoc* network' that they form themselves or they could access the fixed network via a variety of access mechanisms.

Bluetooth-enabled devices will allow 'home networks' to be developed where the TV, HiFi, video recorder, games terminal, computer, etc, are all linked together; as a further goal more mundane domestic appliances could be connected to the home network for control and maintenance purposes.

The view of the future is that IP will be the enabling technology for a new range of portable personal communications devices. These will place an ever-increasing demand for bandwidth and QoS on the IP networks of the future.

6. The requirement for a middleware layer

A number of the newer IP/Internet applications and services require an intelligence layer that is similar to the IN layer currently employed for existing 64 kbit/s switched networks. In some instances, VoIP relies even more heavily on an IN layer than existing conventional telephony. It was soon realised that a number of transport mechanisms (existing 64 kbit/s switched, IP, mobile, etc) all required a similar intelligence layer. The concept of a common generic service plane, or middleware layer, was conceived (see Fig 8) that had one set of IN services that were accessible by all transport layers. In practice, the middleware layer is based on a large distributed computing engine that has a number of components that can be brokered. Information from a particular transport layer is converted to a common format via a thin shim layer between the middleware and the particular transport layer. Once the data is in the middleware environment, whatever service the data requires can be found. Examples of services would be user authentication, number translation, personal number screening.

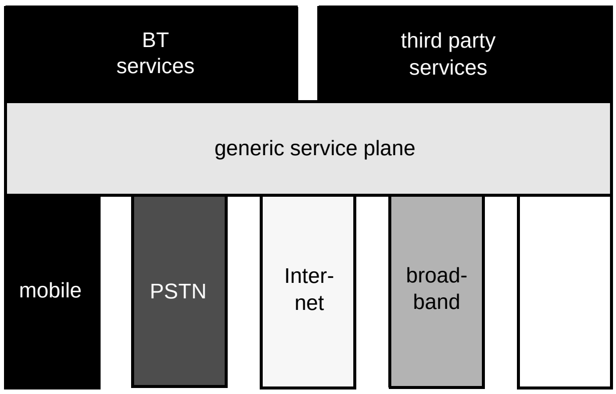

Fig 8 The concept of a common IN layer to support all transport networks.

7. Conclusions

This paper has reviewed the development of IP from its early life as a military communications protocol to the multimedia protocol of choice that it has become today. Today's IP networks have been discussed and developments to support QoS, increase speed and make VPNs scalable via MPLS have all been described. The next great challenge for IP networks — to fully support mobility and the use of a distributed computing platform to form a common IN layer for all transport networks — has also been explored.

References

1 Chetterjee S: 'Requirements for success in gigabit networking', Communications of the ACM, 40, No 7, pp 64—73 (July 1997).

2 http://www.upside.com/texis/mvm/story?id=3512c4520

3 Cohen D: 'Specifications for network voice protocol', IETF 742 22 (November 1977).

4 ITU-T Recommendation Q.41: 'Mean one way propagation time' Melbourne (1988).

5 Rudkin S, Grace A and Whybray M W: 'Real-time applications on the Internet', BT Technol J, 15, No 2, pp 209—225 (April 1997).

6 http://www.vocaltec.com

7 Braden R et al: 'Resource reservation protocol (RSVP), version 1, fundamental specification', IETF RSVP Working Group, work in progress (draft ietf rsvp spec-14) (November 1996).

8 http://www.cisco.com

9 http://www.pluris.com

10 http://www.juniper.net

11 http://www.symbian

12 http://www.psion.com

13 http://www.bluetooth.com

Asad Karim AMIEE and Grad Inst P, graduated from the University of Liverpool in 1994 with a BSc(Hons) in Maths and Physics. In 1995, after completing an MSc in Computer Science also from Liverpool with a thesis on the performance characteristics of relational database technology and OO databases for service control points (SCPs), he joined BT's Network Intelligence Engineering Centre. On joining BT Laboratories, he spent three years looking into fault-tolerant software systems, object relational databases for SCPs, high-performance databases for multimedia services and worked on a EURESCOM collaborative project, before then joining the large-scale IP networks project. He has recently joined the BT Technology Directorate where he is working on the UMTS Project on the IN and IP aspects of Third Generation Mobile.

Peter Hovell graduated from the Royal Military College of Science in 1978 with a BSc (Hons) degree in Electrical Engineering. After a number of years working for the Civil Aviation Authority on the development of radar processing equipment, he joined BT. Initially he worked on the system specification and design of novel local access mechanisms before becoming involved with two RACE projects dealing with broadband ATM demonstrators. During the next few years he was associated with a number of projects constructing broadband demonstrators at BT Laboratories. For the last two years he has managed a corporate project investigating the design of very large-scale IP networks. This area of work has investigated all the up and coming IP technologies of terabit routers, MPLS, IPv6, QoS, UMTS requirements, etc.