

NGN and the Packetizing of Telecommunications

In the early 1980s the introduction of optical fiber into the long haul plant triggered the sudden extinction of analog carriers. With the drop of a pin, Sprint committed exclusively to optical transmission, and shortly thereafter AT&T wrote off its entire analog plant. In the space of a few short years the nation's long distance telecommunications infrastructure was digitized because of optical technology.

Perhaps the theory of evolution by punctuated equilibria applies to telecommunications progress. Seen from a distance of time and perspective, the deployment of technology seems to take giant leaps forward, only to stabilize for another period of years. Now as we approach the millenium we appear to be poised on the brink of another punctuation with the introduction of the concepts embodied in the next generation network. Packet switching may lead to the extinction of the circuit-based network as we know it today, leading to an entirely new framework for the generation of new services that integrate voice, video, and data.

Why Packets?

The idea of packet switching has been with us for almost 40 years. In 1962 Paul Baran at Rand Corporation proposed a system of sending messages over a network of computers in the form of packets. The elegance and flexibility of Baran's proposal had immediate appeal to researchers, but while packets seemed well suited to data transmission, the application to continuous, delay-sensitive voice appeared awkward. Of course, voice was the dominant traffic on the telecommunications network of that day, and would be for many years to come. Moreover, it would be a long time before the day when progress in digital processing technology would make packet switching economically viable. But that day has come – the traffic is turning to data and the economics are turning to packets.

Packet communications has the advantages of flexibility, a common standard and protocol suite, an ability to integrate services, and the potential for lower expense relative to a circuit-switched network. The latter point deserves some further discussion. Even though packets require transmission overhead in their headers and in the buffering of real-time voice, they save transmission capacity relative to circuit switching because of their more efficient voice coding.

In today's circuit-switched network voice is coded using 64 kilobit per seconds, and employs a full-time, full-duplex channel. In contrast, voice can be transmitted on a packet network with equivalent quality at about 8 kilobits on a half-duplex channel, and with savings during the many periods of silence. Taking

this transmission advantage, and accounting for the need for overhead and buffering, simulations at Telcordia show that about ten voice channels can be carried on average over a single DS0 circuit. Thus there is a potential savings of a factor of ten relative to circuit switching on transmission capacity. Moreover, packet switches, i.e., routers, are less expensive per voice channel than are the usual circuit switches.

Neither the efficiency or cost argument in favor of packets is intrinsic to the concept itself, but rather to the timing and force of the packet "wave." We are comparing a "green field" design using packets to an historic and inefficient circuit network. If we were able to redesign the circuit network to use efficient voice coding, and able to mobilize the telecommunications industry to build a new generation of fast circuit switches using the latest technology, then the costs of the two networks might well be comparable. However, that isn't happening. The reality is that today packets are winning the popularity and economic battles.

The Popularity and Resulting Economics of Packets

A recent study at Telcordia of the economics of packet networks showed a cost advantage of 20-40% for the equipment costs in packet technology relative to circuit technology. However, it is important to realize that this advantage is fast increasing because of the exponentially-declining costs of packet routers. While circuit switching costs are also decreasing, they are doing so at a much slower pace. One estimate is that routers are doubling their cost effectiveness every 20 months, as compared with a very slow 80 months for circuit switches. The point is very simple, but profound -- the world is working on packet technology and not circuit technology. In technology today it is necessary to "ride the wave" of popularity, because that is where the economics will be most advantageous. Because of the growth of the Internet, that wave today is with packets.

There is universal agreement in the industry that data and voice traffic today are about equal, and that data will shortly overwhelm voice, making packets ever more the natural container for the content of the network. While voice traffic is growing at 3-5% annually, data appears to be growing at a rate in excess of 100%. Unfortunately, it is almost impossible to determine exactly how much of today's traffic is data and what the growth rate is. How do we define "data"? Does facsimile count? How do we measure the modem traffic? Although certain figures have become industry folklore -- that data traffic and voice traffic are now equal, and that the growth rate of data is about 300% annually -- there is no reliable, accurate assessment of these estimates.

On the other hand, it is possible to get reasonable statistical estimates of the size and growth rate of the Internet. For the last decade this growth rate, as is well known, has been approximately 100% annually. Telcordia maintains a capability that provides continuous estimates of the size of the Internet, which is available on a public website at www.netsizer.com. The Netsizer methodology involves

the sampling of about 100,000 randomly-generated IP addresses each day, and sending a Domain Name Server query for each to discover whether there is a registered owner. From this information a reliable indication of the number of active hosts can be determined. At the end of March 1999, that number was about 53 million hosts computers. Since users behind modem pools are not seen, the number might be multiplied by about 3 to determine actual users. Netsizer shows that the growth of the Internet during the past year has slowed somewhat to a current annual growth rate of approximately 60%.

Thus we see that the traffic on the network is inevitably turning to data, that packets offer efficiency in capacity and cost, and that the momentum of the Internet -- the great promulgator of this technology -- seems unstoppable.

Packet Networks versus Circuit Networks

The inevitable argument arises: packet technology may be cheaper, but the equipment costs in telecommunications are small relative to the operational expenses. Basically, telecommunications is a service business, dominated by people costs, not technology. While there is a great deal of truth in this argument, it does not obviate the advantages of migration to an IP network. (Hereafter, we shall often refer to a packet network as an IP -- Internet Protocol -- network, even though the IP packets may be carried within the network as ATM cells.) Even a small advantage in cost can be parlayed to market power. Moreover, the cost advantage is constantly increasing, and the flexibility of an IP infrastructure promises new value-added service opportunities.

The cost of equipment in a telecommunications company appears in the annual depreciation expense. How large this expense is relative to the operational expense depends greatly on whether the company is a local exchange carrier or an interexchange carrier. The IXCs have a greater percentage equipment cost than do the LECs, who have much greater expense associated with the maintenance and operation of the loop plant. At Telcordia we have developed an economic model for a fictional end-to-end carrier comprising aspects of both LEC and IXC expenses. For this model the equipment cost (annual depreciation) is about 22% of the total expense. Thus packet technology could give an advantage in total expense of 20-40% of that 22% depreciation cost. While not overwhelming, it is still significant in itself.

This assumes, of course, that the operational expense in an IP network will be the same as that in a circuit network. While there is no hard data that proves otherwise, there are indications that there might be substantial differences. It is possible that packet networks are cheaper than circuit networks to operate. Clearly, they are different in some profound ways.

First, and of primary importance, is the difference in nature between a circuit in today's voice network and a "circuit" in an IP network. Dial tone in the voice

network guarantees a known bandwidth and quality. There is no equivalent of "dial tone" in an IP network, and the definitions of bandwidth and quality are complex. A "circuit" does not guarantee bandwidth. Instead of dial tone there is call admission, and instead of the historic measure of voice quality, MOS (mean opinion score), there is QoS (Quality of Service) and a set of SLAs (Service Level Agreements) that is yet to be defined. It is clear that the packet world will be more complicated in some fundamental ways, but these same complications open the way for a wider variety of bandwidths and services, and for the benefits of more customized pricing structures, both for the benefit of the user and the provider.

The operational methodology for an IP network will need to be much more responsive than that traditionally used in voice telephony. In today's local area networks, computer users are automatically assigned addresses upon connection. The IP address is customer-provided. In contrast, voice telephone numbers today are "hard wired". Changing or adding a subscriber involves rewiring and engineering, whereas in the data world this is a matter of computerized registration. Moreover, all operations and services in a data network are traditionally configured from the desktop, and are expected to take place in milliseconds, rather than the minutes or hours involved in service changes in voice telephony. Again, there are complications and possibilities. There is a demand for increased automation and responsiveness, but also the concomitant possibility of decreased operational expense.

Voice on IP

To reiterate, the future traffic on the network will be overwhelmingly data. But even as data overwhelms voice, the main revenue that supports the network infrastructure comes from voice. Clearly, both from an economic and service standpoint, voice will be an essential component of the network traffic. If the network is to be packetized, then voice must be carried seamlessly and with accustomed quality and service features.

In the years since the proposal of packet communications, engineers have argued about the efficacy of carrying voice on packets. In order to carry voice on packets, the voice samples must be segmented and buffered in a way such as to convey the perception of a continuous stream to the listener. If individual packets are lost or delayed too much, there is a break in the speech. Worse, if the latency of the whole stream is more than about a third of a second, then it becomes difficult to carry out a two-way conversation -- the problem that became familiar to users of synchronous satellite telephony.

Meeting these restrictions was difficult in the early days of IP-telephony, which started as a hobby phenomenon with push-to-talk telephony from home PCs. In fact, the delays associated with the multimedia drivers in PC operating systems and in queuing for modem transmission exceed the latency bounds even before

the packets arrive in the transmission network. Thus, while the speech sounded poor, it was acceptable only as a demonstration and as a way to avoid the tariffs of long distance telephony.

The bad quality of these early PC-telephony systems is not intrinsic to IP-telephony. There is no technical reason why IP-telephony should not sound as good as conventional, circuit-switched telephony. In fact, it could be better. Even much better. In IP-telephony the voice coding is a matter of end-to-end agreement, rather than being hard-wired at 64 kilobits per second as in today's telephony. If users want high fidelity, stereo voice, they can have it.

However, there are several problems in throwing voice telephony on top of the Internet as an infrastructure for the nation's voice communications. First, we have the inherent latency in typical PC implementations. Then there is the uncontrolled quality of the Internet itself, where packet losses and delays might cause the voice quality to suffer. But perhaps these problems are more amenable to solution than the fact that for some time to come, a minority of telephone users will be connected via computers on the Internet. Most calls will take place between ordinary "black telephones", or possibly with an ordinary telephone at one end or the other. It will be absolutely necessary to have a system that seamlessly allows IP-telephony calls regardless of the premises equipment being used. Furthermore, it will be necessary to provide some of the service features, such as 800-number dialing, that are a familiar adjunct to today's telephony. All of these objections are overcome in the architecture being proposed for the Next Generation Network

The Next Generation Network

The Next Generation Network, or NGN, marries the three essential networks -- the PSTN, the Internet, and the SS7 network. The signaling network (SS7) contains the databases and intelligence for call service features, as well as the capability for call setup and takedown. The Internet can be used for packet transmission and for access to the computer side of the world. The PSTN is the gateway to the world of telephony. Together their strengths are complementary.

The key new ingredient is the call agent, as described in companion papers in this issue. The call agent mediates between the three networks, passing messages back and forth between networks, telephones, gateways, and appliances to enable IP-telephony sessions to take place. The call agent uses MGCP (Multimedia Gateway Control Protocol), now being standardized by the IETF (Internet Engineering Task Force), to communicate the control messages.

The NGN architecture is flexible and open, and supports the development and operation of other gateway functions, such as accounting gateways and trunking gateways. It uses each network for its own strengths, and facilitates the migration to a packet network at a time when the world is largely circuit-switched.

The problem of latency in IP-telephony can be overcome by having the voice coding within the network.

In the near future we envision a world in which the network may be viewed as a set of inter-operating, communicating, applications -- a world in which the user has a more intimate view of the network, and where configurations are dynamic. We envision a world in which the network databases are self-configuring, cached, cloned, and replicated. We envision a world in which services are composed from trusted software objects that migrate over the net. We envision mobility services that adapt and dispense capabilities according to individual environments. We envision peer-to-peer service activation, and the migration of intelligence in and out of the network. And, eventually the mass market having broadband digital access, whether it is by ADSL, cable modem, or wireless. It will happen, and the NGN is only the first step on an exciting path to a powerful, adaptable future.

R.W. Lucky