Internet Telephony

Henning Schulzrinne   Jonathan Rosenberg

February 4, 2000

Contents

Chapter 1

The Circuit-Switched Telephone System

Confound a telephone, anyway. It is the very demon for conveying similarities of sound that are miracles of divergence from similarity of sense.
Mark Twain, *A Connecticut Yankee in King Arthur's Court*
”Telephone companies are organisations with digital networks which deliver analogue service to users who buy modems and fax machines to carry data and documents.” (by Markus Schuetz ¡Markus.Schuetz@ascom.ch¿)
A while back, German scientists dug 50 meters down in the underground and discovered small pieces of copper. After studying these pieces for a long time, Germany announced that the ancient Germans 25,000 years ago had a nation-wide telephone network. Naturally, the Russian government was not that easily impressed. They ordered their own scientists to dig even deeper. 100 meters down they found small pieces of glass and they soon announced that the ancient Russians 35,000 years ago already had a nation-wide fiber net. Ethiopian scientists were outraged. They dug 200 meters down in the underground, but found absolutely nothing. They happily concluded that the ancient ethiopians 55,000 years ago had cellular telephones.
I have always wished that my computer would be as easy to use as my telephone. My wish has come true. I no longer know how to use my telephone.
Bjarne Stroustrup (author of the C++ programming language)
OLD TELEPHONES never die, they just stop ringing.
”This ’telephone’ has too many shortcomings to be seriously considered as a means of communication. The device is inherently of no value to us.” –Western Union internal memo, 1876.
TELESCOPE, n. A device having a relation to the eye similar to that of the telephone to the ear, enabling distant objects to plague us with a multitude of needless details. Luckily it is unprovided with a bell summoning us to the sacrifice.

1.1 Internet and the Telephone Network – a Comparison

As communications networks, both the Internet and the public switched telephone network share basic requirements, such as universal connectivity, yet they differ in a number of design assumptions.

Service: Even with the addition of some video services, the service model of the PSTN remains dominated by the provision of a narrowband analog channel. If both end systems are digital, it is possible to establish a bit-transparent 64 kb/s channel between them, but this service has to be declared explicitly at call setup time.

Even though it is labeled as a connection-oriented network, the PSTN is really a hybrid network, with a packet-based control network setting up voice circuits. Indeed, any network which offers switched rather than pre-wired (virtual) circuits must be such a hybrid.

End system: With few exceptions such as ISDN phones, phones are "stimulus-controlled" end system that are not generally aware of the state of the service. They simply generate signals, such as dial pulses or DTMF, and render others, such as audio from the line or ring voltage. A phone does not know whether the button pressed is for entering calling card information, invoking a service, dialing a subscriber number or answering prompts for an interactive voice response (IVR) system. This is generally referred to as the "dumb end system" model.

In Internet applications, the end system has a notion of session state, e.g., for a TCP connection or a DNS lookup. This "service awareness" makes Internet end systems "intelligent", although most AI researchers would probably not want to claim sending email as an outgrowth of their work. As always, intelligence is relative…

Note that it is quite possible to build an Internet-based phone system with "dumb" end systems similar in spirit to traditional telephones, as discussed in detail in Chapter **??**.

The location of session awareness governs where services can be implemented. For example, services such as distinctive ringing need to be implemented in the network if the end system simply has a means for generating alerting tones.

Control: Telephone and similar services are based on model of a session or call, with three phases of setup of a session context, information transfer and teardown. (An alternative is the transaction model, consisting of a set of requests and responses. A transaction at the application layer may well set up a transport session, such as a TCP connection.) A session always requires setup at the participating end systems and may require setup in the network. The PSTN combines the two in its signaling protocols, while the Internet, due to its connectionless model, allows to separate establishing sessions at the end system from those within the network. For example, an Internet phone call over a best-effort network may only establish an end-system session, or a number of application sessions may all share the same resource reservation session.

Separating the end-to-end from network session setup has the advantage that data flows do not have follow the path of end-to-end session establishment packets. For example, with call forwarding, there is a danger that voice data has to travel first to the original destination, and then to the final destination. This is often called tromboning or hair-pinning, based on the shape of the circuit.

Both Internet telephony and circuit-switched telephony use out-of-band signaling, but while Internet signaling is *logically* separate, i.e., carried as set of packets distinguished, say, by their port numbers, circuit-switched telephony typically uses separate *physical* circuits for signaling traffic. These signaling circuits may well exist within the same fiber or conduit as the voice circuits, but cannot share bandwidth with the voice circuits. Multiplexing control and voice traffic has the advantage that the transmission delays are those of the high-speed data pipe, not of a lower-speed signaling circuit.

Interfaces: Telephone systems are defined by a set of interfaces, derived from its history in electrical interfaces. These interfaces incorporate all layers of the protocol stack, from the electrical connection to the application. Examples of such interfaces include the S, T and U interfaces for ISDN service. Interfaces are classified as either user-to-network (UNI) or network-to-network interfaces (NNI), where the former expose only a small subset of network functionality. The difference between the two also causes a disparity

Only two questions need to be answered when connecting an end system or another network to the Internet: Is there a common electrical or optical interface for connecting the two? Do both sides "speak" IP? In principle, any IP end system is a functionally equal participant in the network, although other hosts will trust only certain Internet-connected machines to provide services such as DNS lookups or routing. Thus, the endless debate as to which services need to be provided by end system vs. "the network" is largely meaningless in the Internet context. This statement has to be qualified for two reasons: systems behind a firewall and those without a permanent IP address are second-class citizens and cannot, in general, provide services, but rather are limited to being clients. (It is interesting to note that the UNI-NNI division in the phone network is largely meant to protect the network from malicious users, while firewalls serve to protect end systems from attacks from within the network.)

The indications provided to the end system in analog circuit-switched telephony are designed to be directly interpreted by humans, making unambiguous interpretation by machines difficult. For example, machine end systems have to detect tones to track call progress, where such tones often differ depending on the location called. Some information is only available as spoken language, which makes auto-dialers using calling cards, for example, rather brittle.

Naming: With the advent of local number portability, both systems now distinguish between an identifier that labels the communications endpoint and an address which provides guidance where to find this end point. The phone system has a structured numbering space for each, but the identifier is effectively a random number with a country identifier, while the Internet equivalent, domain names, are organization-based. Section **??** covers numbering in detail.

Resource management: In the telephone network, a subscriber either obtains a dedicated switched connection, or nothing. Thus, resource commitment and connectivity are inextricably linked. In a best-effort Internet, connectivity can exist from a few bytes every hour on up. Also, in analog telephone systems, resources cannot be relinquished during the call, so that putting a call on hold does not conserve network resources and resources are consumed even when, say, waiting in silence for the next available agent. This implies that services like call waiting or voice mail forwarding have to be provided by local switch of the person using the call waiting service, while they can be provided by the end system in a system that separates the two.

Regulation: From its introduction, telephone service has been heavily regulated as a public utility, even in countries where service was not provided by a government-sanctioned monopoly. Regulation was primarily national, with different technical and operational standards for each. Even as recently as the introduction of common channel signaling (Section 2.9.2) or ISDN, many regional and country variations developed. Telephone customer premises equipment (CPE) such as telephones or PBXes has to be approved by certification bodies (e.g., according to Part 68 of the FCC rules) before a subscriber can attach them to the telephone network. The PSTN is part of the public service infrastructure (Section 1.8) and must support this role, e.g., by allowing no-charge emergency calls from pay phones. Provision of telephone service requires a government license even in those countries, such as North America and Europe, that have opened parts of the PSTN infrastructure to competition.

1.2 A Brief History of the Telephone

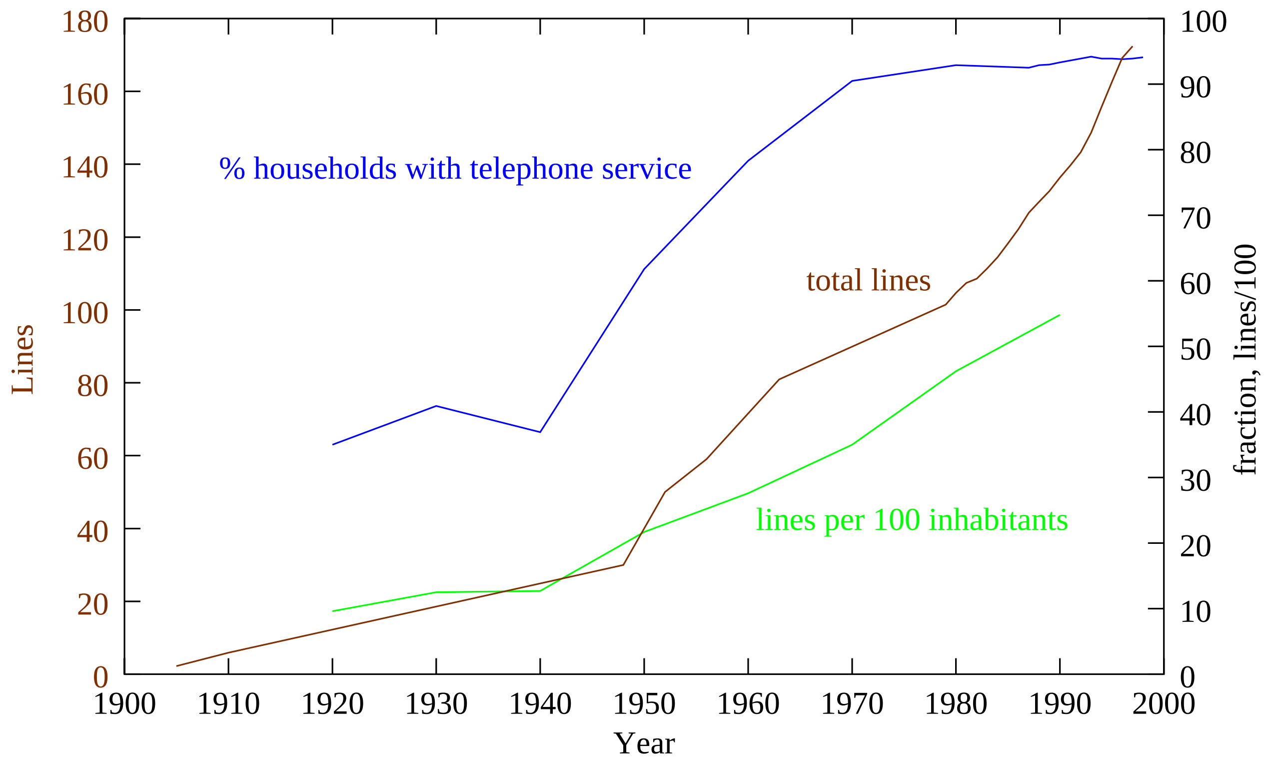

Figure 1.1: Growth of U.S. telephone system

1858	first Atlantic telegraph cable, fails after 26 days
1860s	voice transmission experiments by Philip Reis
1861	coast-to-coast telegraph in the United States
1866	second Atlantic telegraph cable
1876	invention of telephone (Alexander G. Bell)
1878	first central office (New Haven, Connecticut)
1881	long-distance line between Boston and Providence, Rhode Island
1989	first pay-phone (Hartford, Connecticut)
1891	Strowger switch patented
1915	first transcontinental telephone (New York – San Francisco)
1920's	first automatic switches
1927	public demonstration of television transmission over wires
1927	overseas radio telephone
1936	coaxial cable between New York and Philadelphia
1938	first crossbar central office (Brooklyn, New York)
1941	push-button phone
1951	long-distance dialing (Englewood, New Jersey)
1956	TAT-1 transatlantic cable (35 circuits)
1957	pagers
1958	modem developed
1959	direct dialing from Washington, D.C.
1962	digital transmission (T1)
1962	commercial introduction of pagers
1963	touch-tone (DTMF) service
1964	first fax machine
1965	AT&T 1ESS analog switch
1965	first telecommunications satellite ("Molniya", Soviet Union)
1968	U.S. 911 common emergency number introduced
1970	international DDD (London and New York)
1974	Westar Satellite for domestic service, United States
1977	AT&T 4ESS digital switch
1978	first cellular phone testing
1980s	Signaling System #7 (out-of-band signaling)
1982	INMARSAT satellite for maritime use
1984	Divestiture: break-up of AT&T into RBOCs
1987	ISDN trials in the United States
1989	in-flight telephone service (Skyphone)
1996	last electromechanical central office retired

The growth of the U.S. telephone system can be measured in at least three different ways, as shown in Fig. 1.1: by the total number of access lines, the number of access lines per 100 inhabitants, and the fraction of households with telephone service. With the increasing number of households with second lines, the total number of access lines grows more rapidly than the fraction of households with telephone service. This fraction has held fairly constant at around 93-94% since about 1980. A fourth measure, the amount of usage of each telephone line, has not changed significantly since the 1980s, rising from 46 to 57 minutes per day, with three quarters of that increase attributable to increased long-distance calls.

Space does not permit a more detailed treatment of this fascinating topic. The reader may find the book [1] edited by Ithiel de Sola Pool helpful. Also, Amy Friedlander [2] summarizes the history of the telegraph and telephone infrastructure from 1837 to 1940. The most thorough, month-by-month accounting appears in *Events in Telecommunications*, published by AT&T (1992). Some of these can be found at http://www.webbconsult.com/timeline.html.

1.3 Architecture of the Telephone System

The telephone system is effectively a network for connecting end systems via an analog channel with a bandwidth of approximately 3.5 kHz[1]. Digital end systems can reach other digital end systems at either 56 or 64 kb/s.

The telephone system consists of access lines, switches and transmission facilities. This section will highlight some of the important characteristics of these components.

1.3.1 The Classical Hierarchical Phone Network

Neither interconnecting all telephones directly with each other nor through a single switch scales to a network of more than a few thousand phones. Thus, early on, a hierarchy of switches developed. Telephone lines are connected to central offices (COs), which in turn are connected to each other via *trunks*. (Typically, T1/E1 or T3/E3 transmission facilities are used for trunks, as described in Section 1.3.5 below.) Roughly, since most access lines are used only a small fraction of the time (see Section **??**), one trunk circuit is provided for each ten access lines.

Switches that only serve trunks rather than access lines are often called *tandem switches*, although, strictly speaking, the term is reserved for trunk switches interconnecting end offices. The hierarchy is organized into five classes, from regional center, class 1, to sectional center, primary center, toll center down to the end office (class 5). The most common designations are "class 4" and "class 5" offices, for tandem and end office, respectively. (In the old AT&T system, 4ESS and 5ESS served in those two roles.)

The Bell System had 10 regional centers, 52 sectional and 148 primary centers, but 508 toll centers and 9803 end offices [3]. A calls was routed hierarchically until it reached a tree node that had both originating and terminating end office as its child. High-usage trunks could bypass the hierarchy for efficiency. Roughly speaking, the regional and sectional centers were transformed

[1] More precisely, from about 200 Hz to 3.4 kHz

into a single level for most long-distance carriers. AT&T, for example, has 134 switches in its network. However, circuits are set up so that never more than three 4ESS switches are involved within the long-distance network. An average access switch serves 8,600 access lines.

Other countries have similar hierarchies, often reflected in their numbering system. For example, a call in Germany to area code 02232 would go to the Cologne office, which handles all 02…area codes, is then passed on the Bonn office, which handles 022…codes and finally to the office in the small city of Brhl, identified by area code 2232. When mechanical switches were used, subsequent digits in the area code were handled by the next switch, rather than the originating switch.

Note that non-interconnected networks still exist. For example, one traveller reports from Azerbaijan in 1989: "…Later on, I would discover that there were even more than these four networks [the normal public network, the national security network, a network connecting state governmental offices throughout the entire Soviet Union and a system connecting the governmental offices within the Republic of Azerbaijan] in Azerbaijan as numerous private networks existed, all of which were operating more efficiently than the public system. For example, many top government officials have 8 or 9 phones on their desk. The President's assistant mans 19 separate telephones." (*Azerbaijan International*, Summer 1994)

1.3.2 U.S. Post-Divestiture Architecture

In January 1984, the old Bell System was dismantled, dividing the network into local and long-distance providers. AT&T, part of the old Bell System, became one of many competing interexchange carriers (IXC), while local service was split into Bell Operating Companies, so called Baby Bells. In addition, the independent local monopoly telephone companies like GTE, Contel, United Telecommunications and about 1500 small-town telephone carriers continued to provide local service. While the 1984 divestiture achieved competition in long-distance service, the 1996 Telecommunications Act had as one of its goals to introduce local competition between the incumbent local exchange carriers (ILECs) and competitive local exchange carriers (CLECs). In 1998, CLECs had a market share of about 5%, and mostly resold lines owned by the ILEC, either with switching or as "unbundled network elements" (UNEs). Some ILECs provide their own transmission facilities to large business customers.

A long-distance phone call is routed to the point-of-presence of the presubscribed long-distance carrier of the caller, or to the carrier identified by the carrier code dialed after the "10" or "1010" prefix.

1.3.3 End Systems and Access Lines

Fig. 1.4 shows a simplified schematic of a classic (Western Electric 500-style) telephone [3]. In electronic phones, the transformer and balance network have been replaced by electronic components, but the basic architecture remains the same. The ringer is permanently connected to the phone line, while the remainder is only active when the phone receiver is taken off the hook[2] The switch in the central office detects the decreased resistance of the line and then supplies dial tone or stops ringing for incoming calls. A circuit that uses loop-shorting for signaling is called a loop-start (LS) circuit. One problem with loop-start lines occurs when both sides seize the line at the same time, that is, for both an outgoing and incoming call during pauses in ringing. This situation is called "glare" and is avoided for PBX trunks by using "ground start", where the switch grounds one of the wires to the PBX.

Note that the phone only uses two wires for both sendind and receiving audio signals. A "hybrid" transformer, shown as windings A, B, and C, directs the signals to either microphone or receiver. The transformer is designed so that a small amount of voltage from the microphone "leaks" to the receiver. This *sidetone* makes the phone appear to be working. Since the loop resistance can vary significantly, the loop equalization network, consisting of a fixed and a current-dependent resistor (varistor), reduces the receive and transmit levels for short loops. In the central office, a similar hybrid splits the two signal directions; the remainder of the phone system always uses "four-wire" circuit, i.e., a separate channel for each direction. Since the hybrids are not perfect, there is some coupling of the voice signal from the talker through the hybrid at the four-wire-to-two-wire interface at the far end CO, so that the talker hears himself after one round-trip delay ("talker echo"). If an additional reflection occurs, "listener echo" is experienced. The impact of this echo depends on the amplitude and delay. Connections with more than 45 ms of roundtrip delay, corresponding to about 1800 miles distance, need additional echo suppression or cancellation [4] [3, p. 673].

Echo suppression [5] cuts off the speech path from B to A if A is talking, thus eliminating talker echo, operating the system in half-duplex mode. However, this makes interrupting a speaker very difficult. Thus, modern telephone systems use *echo cancellation*. Echo cancellation is "a voice-operated device placed in the four-wire portion of a circuit and used for reducing near-end echo present on the send path by subtracting an estimation of that echo from the near-end echo." [6] A phone system with echo cancellers on both sides is shown in Fig. 1.2. Since the characteristics of the hybrid depend on the two-wire circuit, the echo canceller is implemented as an adaptive filter [7] that subtracts a delayed and filtered version of the far-end waveform from the send path.

In addition, speaker phones and can cause acoustic echo, where the sound from the speaker is reflected on walls and returns, with additional delay and distortion, to the microphone. Sound may bounce around the room several times, so that even small rooms may have Thus, speaker phones typically have a built-in echo suppressor or canceller, as illustrated in Fig. 1.3. Typical conference phones can deal with delays of up to 400 ms[3].

Telephones dial either by rotary dial or "touch tones". Rotary dials simply interrupt the line at roughly 10 pulses per second, with the digit "0" generating 10 pulses. [8] Tone dialing uses a pair of tones, called dual-tone multifrequency signaling, for each digit, as shown in Fig. 1.5. Each digit is at least 40 ms long.

When one side has hung up, the telephone switch has no way of signaling this to the other side

[2]The terminology of "hooks" goes back to the first phones that indeed had hooks for the receiver rather than a cradle.

[3]A concert hall has a reverberation time of two seconds!

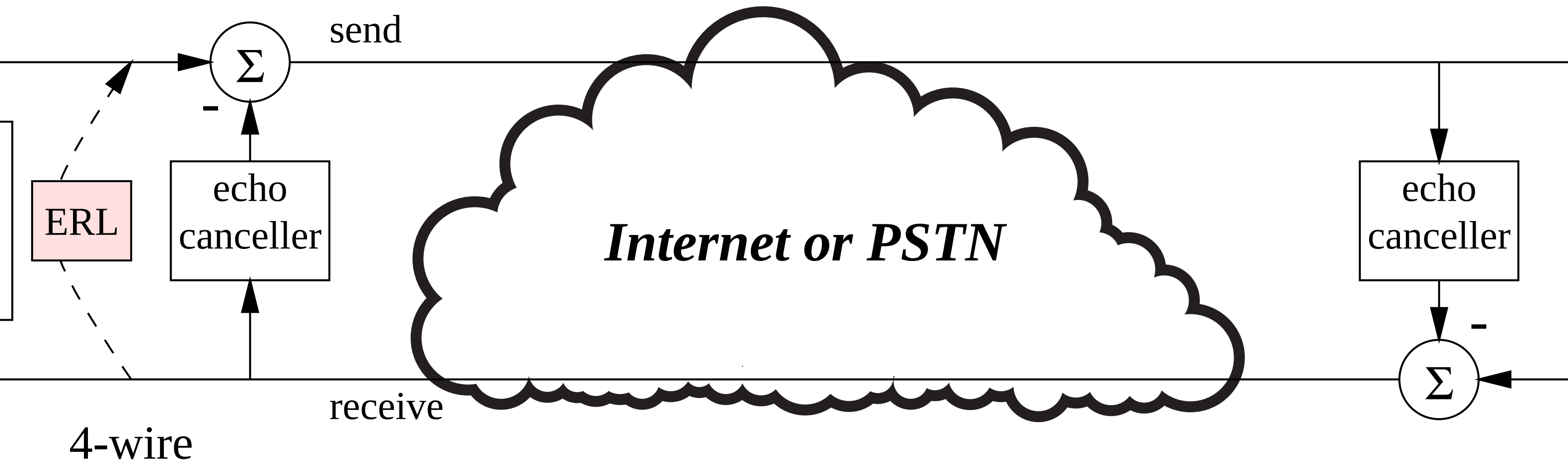

Figure 1.2: Cancelling hybrid-induced echo

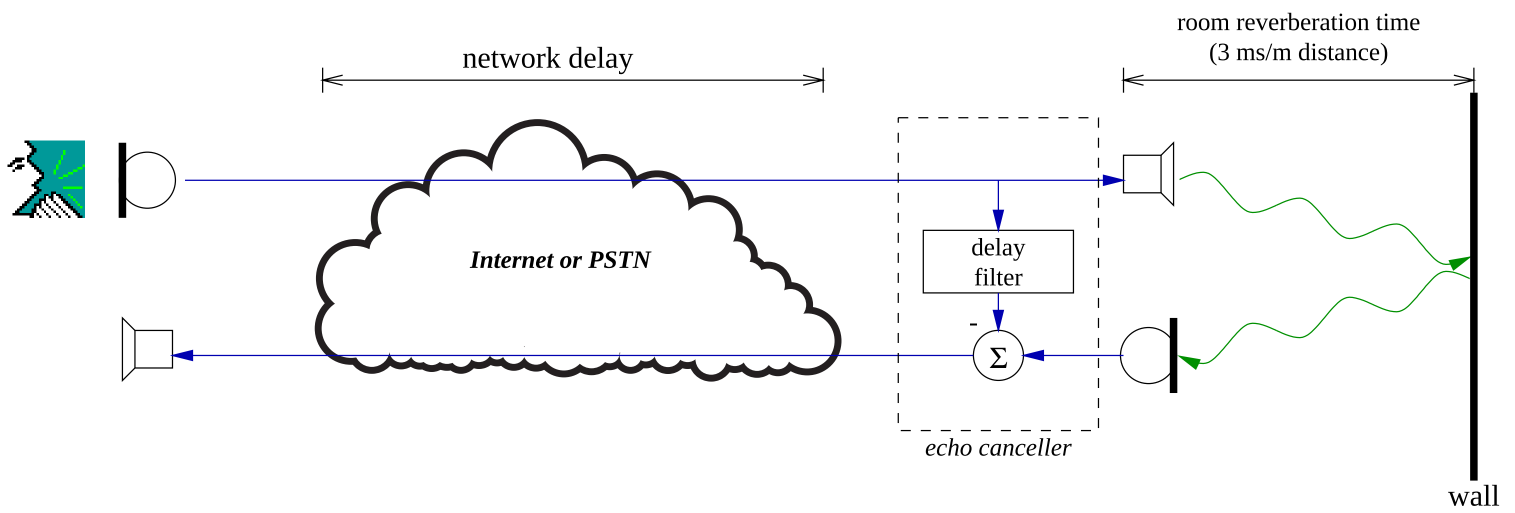

Figure 1.3: Acoustic echo

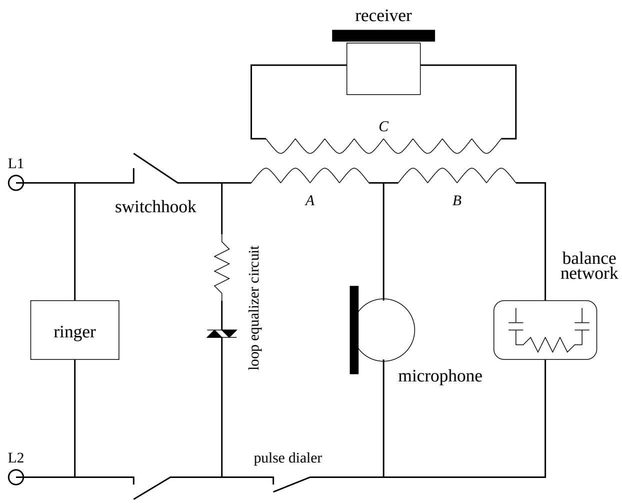

Figure 1.4: Schematic of an analog phone

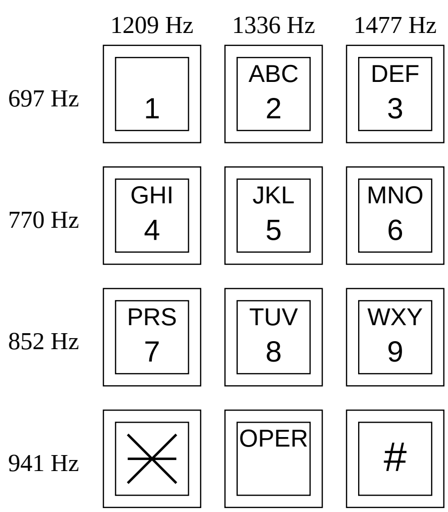

Figure 1.5: Tone-dialing frequencies

of loop-start lines. After about ten seconds, the switch simply supplies dial tone. Also, if the caller abandons the call during ringing, this can only be detected after a full ring pause of six seconds. With ground-start loops, the CO signals disconnects by removing battery from the trunk.

All but the smallest PBXs are generally connected to the central office by either ground-start analog lines, direct inward dialing (DID) trunks, or so-called E&M ("ear & mouth") trunks. The latter is a four-wire interface with two additional signal wires. The PBX grounds the M wire to indicate off-hook, while the CO switch uses the E wire. For digital trunks, the E and M signals are carried as signaling state bits for each circuit.

Telephones are connected by wire of gauge 19 (thickest) to 26 (thinnest) to the central office. Diameters of 26 and 24 ga. are most common. A 26 ga. wire has a diameter of 0.016 inches (0.4 mm) thick and has a loop resistance of 81.62 Ω/1000 ft, while 24 ga. (0.5 mm diameter) wiring has a resistance of 51.34 Ω/1000 ft. Between 6 and 2700 wire pairs are combined into one cable, with two pairs run to each house. The median length of a phone line is 1.7 miles, or, to use the unit commonly used, 9 kilofeet (kft), with 95 percent of all loops shorter than 5.2 miles (27.5 kft). Distances beyond 18 kft, which 15 percent of all loops, are not suitable for ISDN and are thus generally avoided for new construction.

In many cases, the local loop is no longer terminated at the central office but rather at a *subscriber loop multiplexer* (SLM). The SLM statistically multiples up to 80 subscribers into a single T1 line. The common SLC-40 and SLC-96 circuit multiplexers map analog circuits directly into voice channels. The SLC-96 directly feeds the digital signal to a digital CO switch via four T1 lines and is thus called a digital loop carrier (DLC) system. Since a T1 line can be carried in two pairs, instead of running 96 pairs to the CO, only eight are needed.

Access lines are connected to *line cards* in the local central office. Line cards provide BORSCHT:

Battery: The switch provides about 48 V direct current to the telephone, independent of the electric utility. When the receiver goes off-hook, the voltage drops to 5 to 10 Volts. The short-circuit current is only about 20 to 80 mA, so that only low-power electronics can be fed from loop current.

Overvoltage protection: Circuitry is needed that protects the switching equipment from lightning strikes or fallen power wiring and limits the amount of current that can flow if the subscriber loop is shorted.

Ringing: Ringing is 20 Hz AC current at around 90 Volts, with a cadence of 2 s on and 4 s off.

Supervision: Supervision detects whether the phone is on-hook or off-hook. If the current into the phone line exceeds 2 to 5 mA, the phone is assumed to be off-hook and the switch provides dial tone.

Test: The switch has to be able to test for continuity and should be able to inject test tones from the line interface into the switch.

For Internet telephony gateways, the BORSCHT circuitry, particularly the need to supply 48 V and 90 V, is a major component of the cost, but integrated circuits exist that encapsulate all but some of the overvoltage protection in a single package. Typical subcriber line interface circuitry

(SLIC) also performs two-wire-to-four-wire conversion and may integrate analog-to-digital and digital-to-analog conversion.

Calling Number Delivery

The calling number delivery (CND) or "caller id" service delivers the caller's directory number to the called party's two-wire loop. In addition, the current time and date is transmitted. CND transmits its data using a 1200-baud modem signal (V.23) between the first and second ring, using two frequencies for binary 0 and 1.

1.3.4 Switches

A local telephone switch consists of the actual voice switching part, the control logic and a signaling system. The signaling system recognizes when one of the lines goes "off-hook" and applies dial and other tones.

Telephone tones differ from country to country, and even within a country, depending on the technology. Typically, they consist of two pure tones mixed together, or a tone modulated by another one. Some common ITU (E.180 [9]) and U.S. tones are shown in Table 1.1, with additional details on their use found in ITU Recommendation E.182 [10]. RFC XXX [**?**] summarizes uses of tones and their characteristics. Country-specific tones are documented in ITU Recommendation E.180 Supplement 2 [11].

In most cases, Internet telephones try to emulate the same tones as their circuit-switched ancestors. Special information tones precede spoken announcements such as "The number you have reached is not in service". In the U.S., slight frequency and duration variations are used to encode five different conditions.

Tone	frequencies, Hz	volume	cadence
ITU dial tone	425		continuous
U.S. Dial tone	350+440	-13 dBm	continuous, -13 dBm
ITU ringing tone	425		0.67–1.5 s on, 3–5 s off
U.S. ringing tone	440+480	-19 dbm	2.0 s on, 4.0 s off
ITU busy tone	425		slow cadence
U.S. busy tone	480+620	-24 dbm	0.5 s on, 0.5 s off
Congestion tone	480+620		0.25 s on, 0.25 s off
Receiver off-hook	1400+2060+2450+2600		0.1 s on, 0.1 s off
Special information tone (SIT)	950, 1400, 1800		each tone 0.33 s

Table 1.1: Standard ITU and U.S. telephone tones

A large telephone switch such as the Lucent 5ESS-2000 can serve up to 92,000 trunks and 750,000 busy hour call completions (BHCA), with plans in place for capacity expansion to 250,000 ports and 2,500,000 BHCA. Typical configuration for central offices are shown in Table 1.2.

	urban	suburban	rural
Number of switching systems	2.3	1.3	1.0
Area served (square miles)	12	110	130
Busy hour CCS/telephone line	310	270	210
Intraoffice calling (%)	31	54	66
Working lines	41,000	11,000	700
Trunks	5,000	700	35

Table 1.2: Average wire center parameters [3]

1.3.5 Transmission

Each phone call is carried as 64 kb/s bit stream, with bits flowing regardless of whether the sender talks or not[4]. The speech signal is encoded at a sampling rate of 8 kHz, with eight bits per sample. The encoding is a simple *companding* that maps sample amplitudes from 0 to 8159 to an 7-bit table entry, with a roughly logarithmic scale [12, 4]. There are two such encoding schemes, A-law and μ-law, where the former is found in European countries, while the latter is used in North America and Japan. The encoding is often also referred to by its ITU Recommendation name, G.711 [13] or called PCM coding, even though this is not really technically accurate. μ-law encoding offers a signal-to-noise ratio of 39.3 dB for a full-range signal and a dynamic range of 48.4 dB. This is roughly equivalent to that of FM radio, except that the audio bandwidth is far lower.

The 8 kHz sampling period yields the basic clock period, 125 μs, that is found throughout the digital telephone system, even when no voice is being transmitted.

A number of these digital signals are then multiplexed into a single frame. For example, a T1 circuit consists of 24 voice channels, with one byte per channel. This packaging of channels into a single digital stream is called *time-division multiplexing* (TDM). A *frame* consists of these voice channels plus one or more synchronization bits. For example, a T1 trunk has 193 bits per frame, for a raw data rate of 1.544 Mb/s (Table 1.3). In some T1 trunks, the least significant bit of each channels is “robbed” every sixth frame to signal whether the circuit is on-hook or off-hook, so that the transmission is not completely bit-transparent.

While the lower transmission speeds are synchronous and aligned on frames, higher transmission speeds use the Synchronous Optical Network (SONET) technology [14, 4, 15] that offer more management capabilities. SONET frames contain section, line and path overhead for synchronization and maintenance, comprising roughly 5% of the bandwidth. SONET frames are still sent at intervals of 125 μs, in conformance with the basic time interval of telephony.

Table 1.3 shows these higher transmission rates as OC-x. More precisely, the OC-x designation refers to the optical signal, while STS-x and STM-$x/4$ refers to the corresponding electrical signal and bit pattern, but hardly anybody seems to use the latter terminology.

Transmission facilities are usually bundles of fibers, since the construction cost of about $30,000 to $80,000 per mile does not depend on the number of fibers – optical fiber costs only about $100 per mile.

[4] On international trunks, silence detection and compression are used.

Name	voice channels	bit rate (Mb/s)	medium	remark
DS1 (T1)	24	1.544	2 twisted pairs	U.S.
E1	30	2.048		Europe
DS2	96	6.312	2 twisted pairs	U.S. (rarely used)
E3	480	34.368		Europe
DS3 (T3)	672	44.736	fiber	U.S.
OC-3	2340	155.520	fiber, UTP	SONET
OC-12	9360	622.080	fiber	SONET
OC-48	37440	2488.320	fiber	SONET
OC-192	149760	9953.280	fiber	SONET

Table 1.3: Digital TDM Signals

1.3.6 Lawful Intercept

In the United States and many other countries, law enforcement authorities require carriers to give them access to either call details such as the numbers called by a suspect ("peg") or the actual voice stream. In the U.S., wireline and wireless carriers have to upgrade their digital switches to allow such intercept if authorized by court order, as mandated by the Communications Assistance for Law Enforcement Act (CALEA). Depending on the technology, the voice of the suspect and his conversation partner are either mixed, or carried on separate circuits. In some cases, only the voice of the suspect is recorded. Some of the technical issues that arise in balancing the needs of law enforcement and the privacy of third parties not subject to an intercept order include:

- What happens if the suspect forwards his calls to another number?
- If the suspect is put on hold, conversations of the other parties should not be recorded.
- Since foreign governments will not look kindly upon their own citizen on their own soil being intercepted by, say, the FBI, how do you determine when a suspect with a mobile phone is actually in the United States? (This issue arose recently with satellite phones and with a Canadian carrier wanting to provide service in the United States.)

1.3.7 Characteristics of the U.S. Network

The U.S. telephone system consists of 169.2 million access lines (1998) terminating in 19,134 exchanges (1996). These access lines comprise about $2.614 \cdot 10^9$ km of wire on 19 million poles, 280,000 km of trenches and 1,455 million km of ducts (1998). In 1997, 83.9% of these lines are served by digital switches and the rest by analog electronic switches. In addition, there were 60.8 million cellular telephone subscribers.

For international circuits, about half of the activated circuits are used for telephony, the remainder for private-line data service, primarily Internet service. For countries like Japan and South Korea, the fraction of telephone circuits drops to around 38% [16].

1.3.8 Telephone Numbers

Over the year, telephone numbers have acquired a range of meanings. They still primarily identify a particular instrument and its location within the telephone network by area code and (in the U.S.) exchange. Numbers also can refer to a user, through 700 personal numbers, or a whole group of telephones ("hunt groups"). With the advent of number portability, telephone numbers have become even more overloaded, as there are routing numbers, identifying the attachment point of the telephone subscriber, and directory numbers. This is further complicated by various private numbering plans.

In addition to routing calls, numbers also indicate whether a call incurs per-minute charges (by dialing a zero, in many European countries, or one, in the U.S., as the first digit) or whether the callee pays for the call, as in 800 ("freephone") calls or incurs additional charges ("900" calls).

Unfortunately, phone numbers do not convey the type of communication that is possible, i.e., whether there's a regular phone, a mobile phone, a fax machine, modem, pager or TDD (telecommunications device for the deaf, a teletype) at the other end. This is particularly annoying for automated callers, as anybody who has been repeatedly called by a fax machine on a regular phone line and couldn't whistle the 2,100 Hz answer tone back can attest to. Recently, phone numbers have Thus, phone numbers combine the functionality as MAC addresses and IP addresses in the Internet.

In some countries such as Italy, Finland, Germany, mobile phones have special area codes, as calls to those numbers cost more than to regular phones. In the U.S., some regions reserve area codes for mobile and pagers, primarily to avoid having to reassign existing home and business phone numbers, but this is not predictable.

The structure of telephone numbers is given by ITU Recommendation E.164 [17]. E.164 limits international telephone numbers to 15 digits. Recommendation E.123 [18] governs the textual representation of telephone numbers in the customary form of +49 2232 1234, with the "+" indicating the international access code (generally, 00 outside the United States), which varies from country to country. (It is coincidence that the U.S. long distance prefix and the U.S. country code are the same.)

The North American Numbering Plan (NANP) encompasses country code "1", i.e., the United States and its territories, Canada, Bermuda, and many Caribbean nations. It uses fixed-length ten-digit numbers, with the first three digits as area code, the next three as the "exchange" (central office) and the last four as the "line number" within the exchange[5]. In many cases, large cities now use "overlay" area codes, so that a single geographic area has two or more area codes. In the early days, the first two digits of exchange numbers were spelled out as letters, as in LE4 (524) for Leonia, New Jersey. To avoid confusion with long-distance numbers or operator services, exchanges cannot start with 0 or 1 and cannot end in 11, allowing a total of 792 exchanges within each area code.

Until 1995, area codes could only have 0 or 1 as their middle digit, to allow switches to distinguish them from exchanges. (Area codes with a middle zero were used for states originally assigned a single area code, such as 701 for North Dakota. Exchanges could not start with the dig-

[5]Line numbers of the form 9XXX are used by pay phones. Phone numbers in movies, for example, always use XXX-555-XXXX, with XXX-555-01XX as the new official "demo" number.

its 0 or 1, since local phone numbers originally had the form LOcust 4-5678, where the first two letters were derived from the name of the city area. The digits 0 and 1 have no corresponding letters.) [19] In addition, three-digit numbers of the form "N11" are reserved for directory assistance (411), repair (611), emergency services (911) and similar purposes.

The final form of numbers are "vertical service codes" starting with a "*", which are used to invoke features and services such as callback (*69), call trace (*57), anonymous call rejection (*77) or call forwarding (*72).

The telephone numbering system has a number of problems:

- As discussed earlier, telephone numbers do not identify the type of device.

- Due to their limited numbering space, telephone numbers are scarce, so that different subscribers typically share a single number. (Compare, for example, with the number of email addresses per person or family.)

- Dialing rules differ from location to location. For example, in some locations, the area code always has to be dialed, even for local calls, while in others it is not allowed to dial one's own area code. Sometimes a prefix "1" is used for all calls to another area code, while elsewhere only calls that incur toll charges are allowed to use a "1". (The *Industry Numbering Committee Uniform Dialing Plan* proposes uniform 10-digit dialing, but implementation is uncertain.)

- Access codes, e.g., for non-PBX "outside lines", long-distance or international calls, differ widely and change periodically.

- Area codes in the U.S. change frequently, with 44 new area codes in 1997 and 28 new areas codes in 1998. In other countries, digits are added for individual cities or the whole country.

- The length of a phone number cannot be determined without detailed knowledge of the country, city or local exchange, except for countries like the U.S. or France with fixed-length phone numbers (10 and 8 digits, respectively). This is a particular problem for international calls, where local switches in the calling country often cannot determine the end of the dial string, so that they have to wait a few seconds[6]

  Many countries have area codes that differ in length, with short codes for big cities. These area codes employing a prefix numbering scheme (see Section 1.3), so that area codes of a small city near a major city shares the first few digits. For example, area codes in Great Britain have between three and five digits, while German area codes have between two and five digits, not counting the initial zero that indicates a long distance call. Berlin, for example, has the area code 030 and no other city has an areacode starting with those digits. Smaller cities around Berlin have area codes with four digits (such as 0331 for Potsdam near Berlin) or five digits (e.g., 03301 for Oranienburg, also near Berlin).

[6] MCI has a two-second wait; in other systems, *overlap dialing* is used where each digit triggers a signaling message and the end system indicates when the digit string is complete) after the last dialed digit (or until # is pressed).

The number of digits can vary even within a single city. For example, the operator of the Technische Hochschule Darmstadt has the phone number "06151/16-0", while extensions within the University replace the zero by a four-digit number. Smaller countries like Hong Kong have a single eight-digit number,without area codes at all. The maximum number of digits ininternational phone numbers is limited to 15 as of 1997.

- With local number portability, discussed below, telephone numbers become essentially random ten to fifteen-digit strings of numbers, which are much harder to remember than seven-digit numbers plus a recognizable prefix.

1.3.9 Number Portability

Most telephone numbers are tied to geography, i.e., moving from one town to another means changing telephone numbers. (It is often possible to maintain the same number when moving within reach of the same wiring center. Area codes 8YY[7] for toll-free numbers and area code 900 for information services are geographically portable since they are mapped into a routable phone number; numbers in area code 500 were designed to provide portable numbers for "follow-me", but there are far to few for any widescale deployment.

Beyond geographical portability, the advent of competition has made portability across providers of great interest. toll-free numbers were the first to be made portable in 1993, with a database administered by Database Service Management, Inc. When a subscriber makes a toll-free call, the originating local exchange carrier queries this database to find out which carrier handles the number and forwards the call to that carrier. The carrier may then use time-of-day or origin to map the toll-free number to a routable number. Every toll-free call also delivers the caller's number and information about the type of caller through Automatic Number Identification (ANI), even if the caller has disabled caller id.

Toll-free numbers are now inexpensive enough that they are available to residential customers and small businesses at no monthly charge. Unlike Internet domain names, toll-free numbers are assigned on a first-come, first-served basis and cannot be legally bought and sold, avoiding the problem of "domain name squatting".

The more difficult numbering portability problem arises when subscribers want to change their local exchange carrier without changing their phone number. Without this *local number portability* (LNP), it is unlikely that anybody except those moving into a new home would switch telephone companies. Currently, a subscriber can only maintain his number when changing carriers, rather than moving to a different location, but the same technology could be applied for geographic mobility as well. LNP is also used to support service changes, for example from POTS to ISDN or mobile.

For LNP, a subscriber has now two numbers, the routing number (LRN), representing the physical location and carrier of the called party and the directory or called-party number (CdPN), which is the "public" and permanent identifier of the called party. Each carrier is assigned one or more blocks consisting of an area code and exchange, i.e., 10,000 numbers. In a rough way,

[7] 800, 877, 888, with 855 and 866 to follow

this scheme is similar to the division of labor between domain names and IP addresses. Domain names other than personal email addresses such as aol.com are independent of the Internet service provider, while most smaller organizations have to change their IP addresses if they change providers. (Larger organizations that have their own Autonomous System (AS) number can change IP providers without changing their IP addresses. It is apparently a whole lot easier to tell people that their area code has changed than to reconfigure thousands of computers...)

Lockheed Martin CIS (Communications Industries Services) maintains the LNP database for Canada and the U.S. and serves as the Number Portability Administration Center (NPAC). The NPAC has host-to-host CMIP (Common Management Information Protocol, the OSI version of SNMP) network management links to about 300 local exchange carriers (LECs). The LECs submit updates, which are then broadcast in real-time as they become effective. Both the carrier losing and the carrier gaining a customer send an update such as "Directory number 201 555 0100, which used to be served by Tin Can Telco, is now served by the Wet Strings Telephone Company, with a routable number of 201 123 4567" to the NPAC. Once the physical loop has been connected to the new carrier, the update is broadcast to all LECs, who then update their SCPs.

The database is structured into seven RBOC regions plus Canada. The interface specification is designed by an FCC advisory committee, the North American Numbering Council (NANC).

When a caller makes a local call, the originating carrier queries the database and gets back the LRN from the called party number. The LRN then routes the call to Wet Strings, in our example, which recognizes its own prefix. It restores the original dialed number, which was saved in a signaling parameter earlier and routes the call.

Long-distance calls could either be translated by the originating exchange or be routed to the apparent destination based on the dialed number, with the destination performing the translation. Clearly, the originating switch needs to make the translation if geographic portability is to work. However, each SCP will have to have access to a database of almost 200 million entries once LNP becomes widely deployed.

In the United States, the rule is that the $N-1$st switch, i.e., the switch just before the final destination, should perform the database "dip". This is the originating local carrier for local calls and the interexchange carrier for long-distance calls.

Other mechanisms for providing local number portability are described by Nilsson [20]. They include "route to pivot" (RTP), where each "call is routed to the exchange to which the dialled number was originally assigned. If the number has been ported from the switch, the call is released back to a previous switch (the 'pivot' switch) in the call path, for rerouting to the new location" [20].

SS7, described in Section **??**, also uses phone numbers as "global titles" to route TCAP messages. Prior to LNP, STPs could use the first six digits to route the message, but now they have to ask the LNP database for the correct LRN.

Internet telephony gateways have to know whether they are given a number that has already been translated or that still needs translation. Thus, Internet telephony protocols using telephone numbers should allow to distinguish the two meanings of these numbers if they are integrated into the telephone network rather than just end systems that look like PBXs or telephones.

1.3.10 How reliable is the phone system?

Telephone reliability can be measured in a number of ways. For example, one could look at the downtime of individual switches or access lines. The Federal Communications Commissions (FCC) tracks reliability by requiring telephone operating companies to report outage incidents that affect 90,000 or more customers for at least 30 minutes. According to Bellcore, 972 such outages were reported between 1992 and 1997.

As a global indicator of the severity of telephone network failures, ANSI committee T1A1 defined an logarithmically scaled *outage index*, similar to say, the Richter scale or hurricane strengths. The index is calculated from the outage duration, the number of subscribers potentially affected, the date and time the outage started and the number and type of services affected, such as intra-office calls, 911, or long-distance. For example, a local switch with 30,000 lines failure for 30 minutes on a weekday has an outage index of 1.92. 1998 had an aggregated outage index of 1608.

The Network Reliability Steering Committee, which is part of the Alliance for Telecommunications Industry Solutions (www.atis.org), also collects and disseminates information about network outages. According to their reports [21, 22], the mean time between such outages is 2.6 days, with each outage lasting on average 2.1 hours (median 2.9 hours). The leading causes of failures were facilities (45% of all outages, 48% of total outage minutes and 51% of the aggregated outage index for the last five years), local switches (18% of outages), common channel signaling (CCS) (13%) and central office power (7.3%), with causes like overload, digital cross connect (DCS) failures and tandem switch failure rare. The main causes of facility outages, in turn, were cable dig-ups (58%), also known as backhoe fade[8]. Another cause are faulty cable electronics (8%). Natural disasters, while rare, cause the longest outages, with a mean of 13.4 hours. Facility outages also last significantly longer than problems localized within the central office, with a median of 4.8 hours. On average, each line is unavailable three minutes per year [23], with 34 seconds of what as scheduled maintenance.

Outages are not evenly distributed over the day. For example, most scheduled maintenance is performed in the early morning hours. Snow [23, 24] computed the number of line-hours lost as a function of the time of day and found that most the outages with largest potential impact were between midnight and 1 am.

Also, the ARMIS (Automated Reporting Management Information System) records switch outages for local carriers, among many other parameters[9]. For example, Bell Atlantic reported that in 1998, their 180 switches in New York with 20,000 or more lines suffered a combined downtime of 628 minutes, corresponding to a failure probability of $6.6 \cdot 10^{-6}$ or a reliability of 99.9993%. (This reliability level is often referred to as "five nines".) The switch hardware itself can be an order of magnitude more reliable than that, with other outages due to operator error, scheduled maintenance or other problems.

Another measure that reflects the impact of outages is *defects per million*, which measures how many calls per million did not go through the first time because of a network procedural, hardware or software failures [25]. For example, during 1997, AT&T's network had 173 defects-per-million or a reliability of 99.98 percent.

[8] AT&T's term for backhoes is unauthorized cable locators.

[9] ARMIS table 43-05, Section III and IV

nines	fraction	unavailabity per year
1	90%	36.5 days/year
2	99%	3.65 days/year
3	99.9%	8.8 hours/year
4	99.99%	53 minutes/year
5	99.999%	5 minutes/year
6	99.9999%	32 seconds/year

Table 1.4: Reliability expressed as “nines”

Unfortunately, there is no equivalent measure of Internet reliability. It appears unlikely, for example, that the success rate of clicking on a hyperlink would be particularly instructive, since failures there are often due to server problems. Some indication of this overall service probability are available from Inversenet.com. According to their studies, a total of about 5% of web retrievals fail because either the server or network malfunctioned. (This does not count pages that have moved or disappeared.) As an indication of common reliability, some carriers are offering 99.5% availability as a premium service. On the other hand, an ISP (AppliedTheory) claims that its network reliability has been 99.97%. It is not clear whether this counts interconnection failures or access link failures.

America Online reports [26] scheduled and unscheduled outages of 1% or 88 *hours* a year for 1996, down from 3.5% or 307 hours the year before.

The lower Internet reliability is not necessarily due to inferior hardware. Typical local network switching equipment has actualized MTBFs of 170,000 to 210,000 hours (19 to 24 years); from anecdotal evidence, hubs and Ethernet switches rarely fail as a whole, rather, individual interfaces may.

Recent reported large-scale Internet outages were due to either misconfiguration, as in the case of the AOL BGP router collapse in 1996, or a local power failure without adequate uninterruptable power supplies, as when a major POP on the Stanford University campus was out of service for a better part of a day. Many of the Mbone routing failures are due to router misconfiguration, for example injecting all unicast routes into the multicast routing protocol. Also, it is clearly harder to maintain telephone-level uptime when traffic doubles every few months and host counts double yearly.

However, there are a number of obvious improvements that are necessary for Internet services to approach telephone-level reliability:

- Restoration from fiber failures has to take place at the physical (e.g., SONET) layer, since only that can guarantee restoration times on the order or 50 ms [27, 28, 29]. Intradomain linkstate routing protocols such as OSPF can recover from link failures in 200 to 300 ms according to some, while Cisco claims about 6 to 10 seconds [30]. Recovery from router or interface failures take much longer, since they rely on periodic keepalive packets. Restoring circuits at higher layers takes much longer. For example, BGP for Internet routing appears to converge in several minutes, while the AT&T FASTAR system can restore 90 to 95 percent of telephone circuits within two to three minutes.

- Software upgrades should be possible without taking down a router or switch.

- Router configuration must be made simpler, with checking against local rules that make catastrophic failure less likely.

- Closer integration of network functionality into the operating system should reduce end-system difficulties. Much of the complexity of configuring current PCs for Internet usage appears to stem from having to configure the modem and multiple protocol stacks. Widespread use of DHCP [31] and IPv6 autoconfiguration [32], as well as eliminating the modem, should make the network invisible.

- While backbone networks feature redundant links, POPs and access links are often single point of failures. Also, different providers often peer at only a small number of points. Particularly the latter problem must be remedied to increase the number of true end-to-end alternate paths.

In the long run, tools like traceroute and ping, as well as relatively simple management protocols (SNMP) and applications with built-in reporting mechanisms such as provided by RTCP (Chapter **??**) probably make Internet service more manageable than many traditional POTS installations.

1.3.11 What is the maximum tolerable delay for a phone call?

Internet telephony, due to packetization and delay jitter in the network (see Chapter **??**), will suffer from additional delays beyond the speed-of-light propagation delay. Thus, there has been renewed interest in determining the end-to-end delay budget for phone calls. In traditional telephone systems, delays are caused by speed-of-light delays. These delays are largest in submarine coax cables, about 6 μs/km and smaller for fiber, at $5\,\mu$s/km and terrestrial coax and radio transmission, about $4\,\mu$s/km. Thus, copper wiring is actually "faster" than fiber. Satellite systems in low-earth orbit (LEOs) add about 5 ms of delay, while a geosynchronous satellite such as Inmarsat parked at 36,000 km altitude adds 260 ms of delay. ITU Recommendation G.114 [33] summarizes delays incurred by other elements in the phone system, such as multiplexers and echo cancellers, which are all below 1 ms. Another major source of delay are mobile systems; their speech compression and framing algorithms can add delays of up to 80 ms.

There are two delay thresholds, one caused by echos, the other by the difficulty of sustaining a conversation when delays are long. We discuss the how delays cause echo problems in Section 1.3.

When delays increase beyond the echo threshold to above about 100 ms one-way delay, conversation begins to suffer subtly. The problem occurs in the "hand-over" between the two parties. If one party makes a small, unconscious pause to give the other side a chance to get a word in edgewise, the remote party will not hear the pause until the first speaker has already continued after the brief silence. Then, the second speaker may speak up, but just ends up interrupting the first speaker, who might back off, just to repeat the cycle. This problem is very similar to that in Ethernet, where the collision sensing multiple access (CSMA)/collision detection (CD) mechanism is used. For that reason, Ethernets are generally limited to span a few hundred meters. In Ethernets, collisions destroy both packets, while double-talking is mostly just annoying[10]. Beyond a certain threshold, the system becomes half-duplex, where conversation hand-off mechanisms such as saying "over" are needed.

G.114 considers delays up to 150 ms acceptable for most applications and declares delays above 400 ms as unacceptable for network planning purposes. Clearly, most users will prefer a two-satellite path between mobile end system, with its delay of upward of 600 ms, preferable to no communication at all, but that's no reason to design an Internet telephone system that way.

A number of studies have investigated the effect of delays in echo-free telephone conversations in more detail, dating back to 1963 [34, 35, 36, 37, 38, 39]

COMSAT [33] asked callers whether they experienced difficulties in conversation when one-way delays ranged from 45 to 300 and 500 ms, but with echo cancellers instead of echo suppressors. No significant differences arose until the delay reached 500 ms, where the fraction reporting difficulties doubled to 15.8%. Studies with echo suppressors had shown that more than 60% experienced difficulties at delays of 500 ms. However, a task-oriented study [33] measuring the fraction of users reporting that overall quality or interruptability was poor or worse shows significant increases above about 300 ms. At 500 ms, about 15% rated interruptibility poor, rising to 60% at 700 ms and 80% at 1250 ms. A 1991 Bellcore study reported in [33] saw mean opinion scores from 3.8 (good) for no delay, to 2.8 (fair) for 250 ms and 2.2 (poor) for 500 ms.

NTT measured detectability thresholds for conversational tasks, with trained staff being able to detect delays of 50 ms on a task requiring subjects to read out random numbers in turn as quickly as possible. On the other hand, delays in free conversation could be detected only at about 350 ms by trained subjects.

There do not appear to be any studies that evaluate whether customers would tolerate higher delay in return for lower per-minute toll charges or higher bandwidth, e.g., for better speech quality or video.

1.4 Telephony Signaling

Signaling is an umbrella term for the control part of a (telecommunications) network. It comprises primarily the establishment, modification and tear-down of calls or connections, but also the auxiliary functions necessary to support this, such as distributed applications for controlling service or network management related to these. In this section, we will describe how telephone signaling works, while Internet telephony signaling is covered in Chapter **??** and Chapter **??**.

The telephone signaling system has evolved from *in-band* to *out-of-band* signaling. In in-band signaling, also known as channel-associated signaling (CAS), the signaling information is carried as tones or voltages in the same circuit as voice. The signaling between the switch and analog phones or PBXes still uses CAS today. For example, early signaling systems used a fixed, 2600 Hz tone to indicate that the trunk was idle. If one of the switches wanted to seize the trunk to set up a

[10]although the tolerance seems to vary by culture. . .

circuit, it would drop the tone and then transmit dialing information using multi-frequency tones, similar to DTMF used today. Unless filters were inserted, simple tone generators could be used to trick the phone network into making "free" calls. CAS is slow, with call setup times of between ten and twenty seconds, and makes it difficult to do number translation or non-hierarchical routing of phone calls.

A simple out-of-band signaling mechanism associates a separate pair of wiress with each voice circuit, where these separate wires indicate whether one side of the circuit wants to seize the circuit. This is, essentially, the mechanism used for E&M ("ear and mouth") signaling. In its digital form, the state of these extra wires is encoded in bits embedded in the voice stream. E&M signaling seems to be primarily used between the local switch and PBXs, as well as between PBXs.

Most out-of-band signaling in use in the PSTN is also known as common-channel signaling (CCS), since signaling information for all voice channels is carried together in a separate digital channel. While one could presumably invent analog common channel signaling, CCS emerged when packet-based multiplexing made it feasible to aggregate and separate a large number of logically different information streams on a single wire.

In the telephone network, we have two different, but related CCS system, ISDN user-to-network or access signaling, known as Digital Subscriber Signaling System 1 (DSS1), and Signaling System #7 (SS7) for network-to-network signaling. Both use protocol stacks that are similar to the standard seven-layer model, but are not directly comparable and do not use either OSI or Internet protocols. While DSS1 is strictly for point-to-point networks and thus needs no network layer, SS7 does need to forward messages through a series of switches.

Development of SS7 and DSS1 started in the 1980s, with most specifications completed by 1992.

While it may be possible to build a general data network on top of the lower SS7 layers, SS7 is designed for highly reliable delivery of individual messages, and would not work well for the delivery of bulk data. It is, however, possible to carry the upper parts of these signaling protocol stacks across IP, using either TCP or specialized transport protocols. The IETF SIGTRAN working group is investigating different mechanisms for this purpose.

The ISDN signaling protocol stack is shown in Fig. 1.7, while the SS7 stack is shown in Fig. **??**. Both stack consist of a set of layers providing message encapsulation, routing and reliability, and an application or user layer that provides the actual signaling capabilities.

XXX ISDN XXX ISDN signaling, described in ITU Recommendation Q.931 [40], uses a separate channel on the access link. *Basic rate ISDN* (BRI) consists of two 64 kb/s "bearer" (B) channels and one 16 kb/s data (D) channel, usually abbreviated as 2B+D. Bearer channels carry the application data, either voice or data. *Primary rate ISDN* (PRI) has 23B+D in the North America and 30B+D in Europe, where the D channel in a PRI interface is simply a 64 kb/s voice channel.

At the application layer, SS7 consists of a call setup part, primarily ISUP, and a service part, TCAP. While ISUP sets up calls and generally travels along the same path as the voice path to be set up, TCAP is used by switches to resolve and translate addresses and make other queries. When interoperating with Internet-based telephone services, TCAP stays hidden inside the PSTN; thus, we focus our discussion on ISUP below.

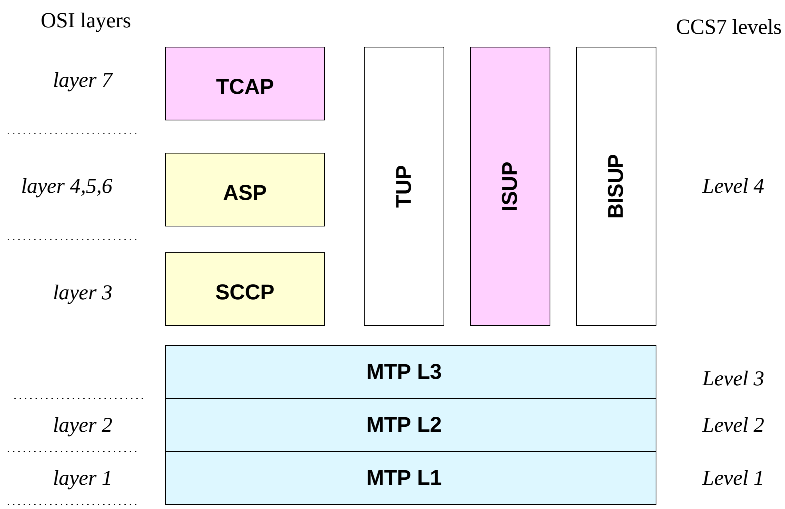

Figure 1.6: SS7 Protocol Stack

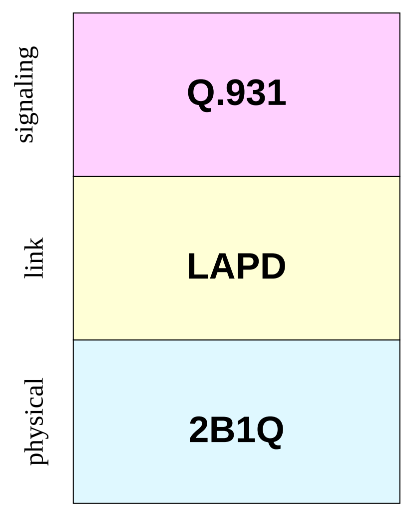

Figure 1.7: ISDN Signaling Protocol Stack

1.4.1 MTP

The lower layers of the SS7 protocol stack have similar functions as the link layer (say, PPP), IP and TCP in the Internet architecture, although the layering differs somewhat.

The Message Transfer Part (MTP) consists of three levels providing physical, link and network-layer functionality. The physical layer, MTP Level 1, supports 56 kb/s, 64 kb/s and 1.544 Mb/s links. Framing and error correction are done by MTP Level 2. Packets, or *signal units* in MTP parlance, are delineated by a “0111 1110” flag byte with bit stuffing, as in HDLC, where a zero-bit is inserted after every set of five consecutive one bits. The maximum transport unit (MTU) can be 280 bytes long, of which eight bytes are header and trailer bytes. The packet trailer consists of a 16-bit CRC checksum and the flag byte. The header has a Service Information Octet (SIO), roughly equivalent to the protocol number in IP, two 7-bit packet sequence numbers, similar to those in TCP, a length field and flag bits. The payload part is called signaling information field (SIF).

MTP Level 2 is a reliable, sequenced datagram protocol, with hop-by-hop reliability. It uses go-back-N with positive and negative acknowledgements. The window size is given by the sequence number lengths, at 127 packets. If the signaling point does not have any data to send, it transmits fill-in signal units containing no SIF. The checksum is then used to monitor the link for bit errors.

Flow control is hop-by-hop. The receiver withholds acknowledgements and sends a status message indicating congestion, throttling the sender.

MTP Level 3 routes packets between *signaling endpoints*, connected by links and *signal transfer points*, the SS7 equivalent of “routers”. End points are addressed by point codes, with MTP Level 3 containing an origination and destination point code (OPC and DPC). These addresses are similar to MAC or IP addresses. In ANSI SS7 networks, point codes have a structure similar to classful IPv4 addresses, with a division into network, cluster and member, similar to the network, subnet and host division for IP addresses. ITU-T point codes are drawn from a flat 14-bit address space. A signaling link selectin (SLS) field with four bits selects one of a set of alternate links for load balancing. In the Bellcore/ANSI standard, these addresses are 24 bits long, with a five-bits link selection field. The routing label makes up the MTP Level 3 header.

Signaling end point are typically connected by a set of redundant links in a quad configuration, as shown in Fig. 1.8. SS7 does not have a routing protocol in the sense of, say, OSPF or BGP, so that routes are pre-configured depending on the available links. This routing mode is called *quasi-associated*.

1.4.2 ISUP

The ISDN User Part (ISUP) of signaling system #7 provides the signaling for basic circuit-switched services, such as setting up and terminating calls. The packet format of ISUP message is shown in Fig. **??**. Note that, unlike more recent ITU protocols like H.323, neither Q.931 nor ISUP use ASN.1. The packet format is a combination of the fixed-sized structures found in protocols like IP or TCP and the type-length-value format found in protocols like RSVP. It also contains a somewhat peculiar pointer-based data structure for the mandatory-variable part, which contains those message parameters that are mandatory, but each have variable length.

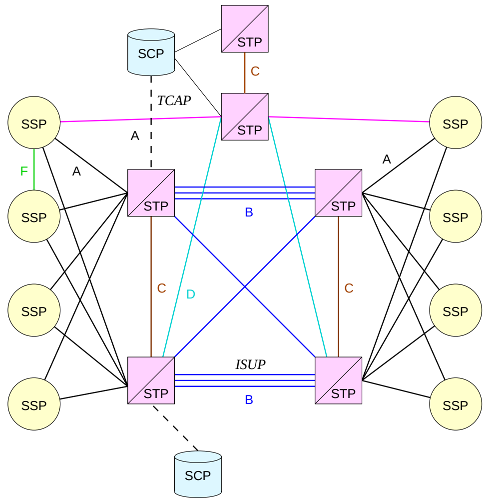

Figure 1.8: SS7 Signaling Network

The ISUP message identifies the circuit that it wants to set up with a circuit identification code (CIC). This CIC XXX

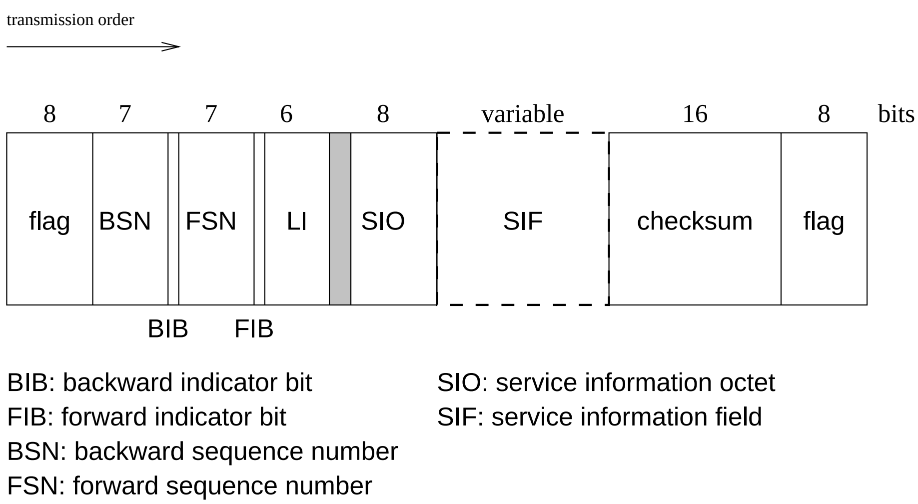

Figure 1.9: MTP level 2 packet

A typical call setup involving both ISDN and ISUP is shown in Fig. 1.11.

[ADDITIONAL DETAIL AND FIGURES OF IAM, REL, ETC. TO BE ADDED.]

ISUP comes in a large number of regional flavors, including versions standardized by ANSI/Bellcore for the North American network, ETSI for Europe, NTT for Japan and ITU Rec. Q.767 [41] as the “international ISUP”. Within these regional flavors, there are national variations. Recent ITU “White Book” versions have added a compatibility mechanism that indicate to nodes with earlier versions whether to simply pass on the extension, discard the extension or drop the call.

If two circuit-switched networks are to be connected via an internet, ISUP packets may need

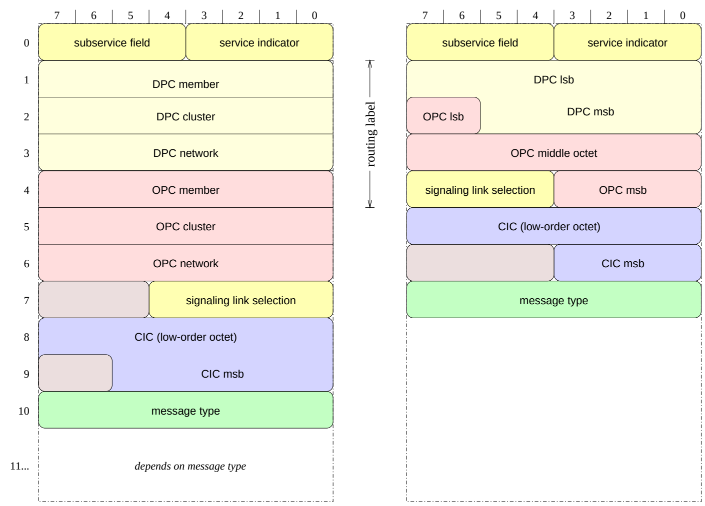

Figure 1.10: ISUP packet format

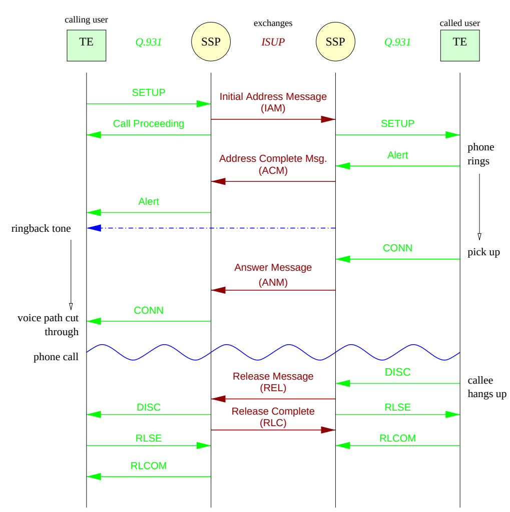

Figure 1.11: ISDN and ISUP call setup and teardown

to be tunneled across the IP cloud, without intermediate Internet signaling nodes having to know about the details of ISUP. Section **??** discusses the simple mechanism that allows SIP to carry ISUP messages.

1.5 Services

This section describes many of the features and services found in circuit-switched telephone networks. We see the enumeration of features as a convenient notational short-hand. It is not clear that an actual Internet telephony implementation should proceed feature by feature, as described in Section **??**. Also, many of these features do not necessarily map cleanly into the Internet, at least not without sacrificing the generality of Internet transport.

Study Group (SG) 11 of the International Telecommunications Union Telecommunications Standards Sector (ITU-T) published its accumulated descriptions of services and service features in Annex B of ITU-T recommendation *Q.1211: Introduction to Intelligent Network Capability Set 1* [42]. Since these service descriptions were compiled from a number of disparate sources, the document acknowledges that they may be self- and mutually-inconsistent. Thus, we have applied some editing to the descriptions below to reflect, as best as we could, reality.

Recently, ITU Study Group 11 has written a follow-up document, *Q.1221: Introduction to Intelligent Network Capability Set 2* [43]. This document has not yet been formally ratified or released by the ITU; I address it, in less exhaustive detail, in section 1.7.

For some features, we comment on how Internet telephony systems might provide this functionality. In other cases, this depends on the signaling approach being used and we have to defer the discussion to the description of H.323 (Ch. **??**), SIP (Ch. **??**) and the device control protocols (Ch. **??**). [NOTE: MAYBE BETTER TO DO A COMPARISON AFTER COVERING ALL THREE.]

TIA Technical Requirements document 41.3 also outlines features and how to support them using Internet telephony.

1.5.1 Capability Set 1: Service Features

Q.1211 divides the services it describes into two broad categories: “services,” which are what an Intelligent Network vendor would actually wish to provide to customers; and “service features,” which are lower-level building blocks used to construct the services.

This section describes all the service features listed in Q.1211, Annex B, section 2; Section 1.6 describes the services built out of these service features (from Annex B section 1), and any unique aspects of them in the Internet telephony environment.

For each service, Q.1211 lists a number of service features which are considered either core to it, those without which it would not be useful, or optional, which provide added value to the feature. For each service feature, this section lists the services which Q.1211 specifies use that feature. Similarly, in section 1.6, each service lists its component service features.

See table 1.5 for a summary of the characteristics of each service feature.

Service Feature			**Characteristics**	
Code	Name	Section	Location	Call Time
Authentication				
AUTC	Authentication	1.5.4	End / Proxy	Setup
AUTZ	Authorization code	1.5.5	Proxy	Setup
Billing				
PRMC	Premium charging	1.5.34	-	-
REVC	Reverse charging	1.5.36	-	-
SPLC	Split charging	1.5.37	-	-
Filtering				
OCS	Originating call screening	1.5.31	End / Proxy	Setup
TCS	Terminating call screening	1.5.38	End / Proxy	Setup
Forwarding				
CD	Call distribution	1.5.7	Proxy / Redirect	Setup
CF	Call forwarding	1.5.8	Proxy / Redirect	Setup
CFC	Call forwarding on busy/don't answer	1.5.9	Proxy / Redirect	Setup
ONE	One number	1.5.29	Proxy / Redirect	Setup
ODR	Origin dependent routing	1.5.30	Proxy / Redirect	Setup
PN	Personal numbering	1.5.33	Proxy / Redirect	Setup
TDR	Time dependent routing	1.5.39	Proxy / Redirect	Setup
Translation				
ABD	Abbreviated dialling	1.5.2	End / Proxy / Redirect	Setup
ATT	Attendant	1.5.3	End / Proxy / Redirect	Setup
PNP	Private numbering plan	1.5.35	End / Proxy / Redirect	Setup
User interface				
CW	Call waiting	1.5.16	End	Setup
CRG	Customized ringing	1.5.21	End	Setup
Other				
ACB	Automatic call back	1.5.6	End	Setup
GAP	Call gapping	1.5.10	Proxy	Setup
CHA	Call hold with announcement	1.5.11	End	In call
LIM	Call limiter	1.5.12	End	Setup
LOG	Call logging	1.5.13	All	All
QUE	Call queueing	1.5.14	End / Proxy	Setup
TRA	Call transfer	1.5.15	End	In call
CUG	Closed user group	1.5.17	End / Proxy	Setup
COC	Consultation calling	1.5.18	End	In call
CPM	Customer profile management	1.5.19	End / Proxy / Redirect	Indep. of call
CRA	Customer recorded announcement	1.5.20	End	In call
DUP	Destinating user prompter	1.5.22	End	In call
FMD	Follow-me diversion	1.5.23	End	Indep. of call
MAS	Mass calling	1.5.24	Proxy	Setup
MMC	Meet-me conference	1.5.25	Other	Setup
MWC	Multi-way calling	1.5.26	End	Setup
OFA	Off-net access	1.5.27	All	All
ONC	Off-net calling	1.5.28	Proxy	Setup
OUP	Originating user prompter	1.5.32	All	Setup

1.5.2 Abbreviated Dialing (ABD)

Used for: ABD (core); ACC (core); AAB (optional); CCC (optional); VPN (optional)

> Abbreviated dialing allows the definition of short (e.g., two digit) digit sequences to represent the actual dialing digit sequence for a public or private numbering scheme.

In Internet telephony, an end system would typically do this work; either by storing an internal table of locally-defined shortcut addresses for the actual addresses it would send, or (for setups more analogous to VPNs) by having end systems configured to consult a local database server (running, e.g., LDAP) for address-translation queries.

1.5.3 Attendant (ATT)

Used for: VPN (optional)

> This allows VPN users to access an attendant (operator) position within the VPN for providing VPN service information (e.g, VPN numbers) by dialing a special access code.

An Internet telephony end system needs only to be configured with an address of an appropriate local operator to translate the special access code to the actual local address of an attendant, or some address which will resolve to that address.

1.5.4 Authentication (AUTC)

Used for: FPH (optional); SEC (core); VPN (optional)

> This allows verification that a user is allowed to access certain options in the telephone network.

1.5.5 Authorization code (AUTZ)

Used for: ACC (core); AAB (core); CCC (core); UPT (core); VPN (optional)

> This allows a user (typically in a VPN) to override the restrictions placed on the system from which calls are made.

1.5.6 Automatic call back (ACB)

Used for: CCBS (core)

> This feature allows the called party to automatically call back the calling party of the last call directed to the called party.

This can be handled entirely by end systems, which need only store and make available to the user the previous invitations they have received. Note that this is not restricted to only the last call; the number of old invitations a Internet telephony end system remembers is bounded only by its local storage.

1.5.7 Call distribution (CD)

Used for: CD (core); DCR (core); FPH (optional); MAS (optional); PRM (optional); SPL (optional); VOT (optional); UAN (optional); VPN (optional)

> This service feature allows the served user to specify the percentage of calls to be distributed among two or more destinations. Other criteria may also apply to the distribution of calls to each destination.

1.5.8 Call forwarding (CF)

Used for: CF (core)

> This service feature allows the user to have his incoming calls addressed to another number, no matter what the called party line status may be.

1.5.9 Call forwarding on busy/don't answer (CFC)

Used for: CRF (optional); FPH (optional); PRM (optional); SCF (core); SPL (optional); UAN (optional)

> This service feature allows the called user to forward particular calls if the called user is busy or does not answer within a specified number of rings.

1.5.10 Call gapping (GAP)

Used for: FPH (optional); PRM (optional); SCF (core); SPL (optional); UAN (optional)

> This feature allows the service provider to restrict the number of calls to a served user to prevent congestion of the network.

The intended scenario of this service feature is that large numbers of people may simultaneously call the same destination address, for instance because it was announced on television, and the network needs to ensure that its servers and signalling network are not overloaded.

The simplest case of this is when the overloaded server does not have the necessary resources to completely fulfill the request, but it can still process it and send a basic response.

1.5.11 Call hold with announcement (CHA)

Used for: VPN (optional)

> The call hold with announcement service feature allows a subscriber to place a call on hold with options to play music or customized announcements to the held party.

This can be handled simply by switching the media which is being sent to the remote party. It can either originate from the same end system, or from a media server, perhaps triggered by a media server control protocol such as RTSP [44]. (RTSP servers are normally restricted from sending media to a third party, but if the RTSP server and SIP server trust each other this could be overridden.)

1.5.12 Call limiter (LIM)

Used for: CRD (optional); FPH (optional); MAS (optional); PRM (optional); SPL (optional); VOT (optional); UAN (optional)

> This service feature allows a served user to specify the maximum number of simultaneous calls to a served user's destination. If the destination is busy, the call may be routed to an alternative destination.

1.5.13 Call logging (LOG)

Used for: ABD (optional); ACC (optional); AAB (optional); CD (optional); CF (optional); CRD (optional); CCBS (optional); CON (optional); CCC (optional); DCR (optional); FMD (optional); FPH (optional); MCI (core); MAS (optional); OCS (optional); PRM (optional); SEC (optional); SCF (optional); SPL (optional); VOT (optional); TCS (optional); UAN (optional); UPT (optional); UDR (optional); VPN (optional)

> This service feature allows for a record to be prepared each time that a call is received to a specified telephone number.

Obviously, any element of an Internet telephony system may log anything it wishes, if it has someplace to store the log.

1.5.14 Call queueing (QUE)

Used for: CRD (optional); FPH (optional); MAS (optional); PRM (optional); SPL (optional); VOT (optional); UAN (optional); VPN (optional)

> This service feature allows calls which would otherwise be declared busy to be placed in a queue and connected as soon as the free condition is detected. Upon entering the queue, the caller hears an initial announcement informing the caller the call will be answered when a line is available.

A system could do queueing entirely at an end system, accepting the call from a signaling level and directing the call's media to a server which plays appropriate announcements about the queued call's state. When the call is de-queued to be picked up, the queueing server either transfers the call or proxies the signaling to the appropriate end system while directing the media transmission there. This solution is more similar to how call queueing is handled in the traditional telephone network. Other solutions depend on the Internet signaling protocol.

1.5.15 Call transfer (TRA)

Used for: VPN (optional)

> The call transfer service feature allows a subscriber to place a call on hold and transfer the call to another location.

1.5.16 Call waiting (CW)

Used for: CCBS (optional)

> This service feature allows a subscriber to receive a notification that another party is trying to reach his number while he is busy talking to another calling party.

Due to the separation of signaling and media in Internet telephony, this feature is entirely an end-system issue if peer-to-peer signaling such as H.323 or SIP is used. An end system which receives a call invitation while in a call may alert the user however it wishes, if it so chooses.

1.5.17 Closed user group (CUG)

Used for: VPN (optional)

> This service feature allows the user to be a member of a set of VPN users who are normally authorized to make and receive calls only within the group.

Both making calls and receiving calls within a group are restrictions which in Internet telephony require administrative control of end systems. The end systems can either directly enforce these calling restrictions, or they can insist that all their signalling go through a fixed local proxy, which enforces these rules.

Alternately, if the desired closed group corresponds to the end systems on some particular part of the underlying network topology, firewalls could keep calls restricted to that sub-network.

1.5.18 Consultation calling (COC)

Used for: CON (optional); VPN (optional)

> The consultation calling service feature allows a subscriber to place a call on hold, in order to initiate a new call for consultation.

Initiating new calls is possible at any time in Internet telephony. Placing a call on hold is a matter of either ignoring its media, if bandwidth is not an issue; or, more efficiently, sending it a re-invitation with media turned off. In either case, the local end system would likely either stop sending media or transmit a recorded message to the remote party.

1.5.19 Customer profile management (CPM)

Used for: ABD (optional); CD (optional); CF (optional); CRD (optional); CON (optional); DCR (optional); FMD (optional); FPH (optional); MAS (optional); OCS (optional); PRM (optional); SEC (optional); SCF (optional); SPL (optional); VOT (optional); TCS (optional); UAN (optional); UPT (optional); UDR (optional); VPN (optional)

> This service feature allows the subscriber to real-time manage his service profile, i.e. terminating destinations, announcements to be played, call distribution, and so on.

Features that reside in end systems can obviously be configured at these end systems transparently. Features that reside in the network, in proxies or redirect servers, can be configured through any number of means; one under current development is a Call Processing Language described in Section **??**.

1.5.20 Customer recorded announcement (CRA)

Used for: CDR (optional); FPH (optional); MAS (optional); PRM (optional); SPL (optional); VOT (optional); UAN (optional); UPT (optional); VPN (optional)

> This service allows a call to be completed to a (customized) terminating announcement instead of a subscriber line. The served user may define different announcements for unsuccessful call completions due to different reasons (e.g., caller outside business hours, all lines are busy).

A server receiving an incoming call can direct the media component of a call it accepts to any device which can send appropriate media, including, for instance, an RTSP server to play back media. (See the note about RTSP in section 1.5.11.)

1.5.21 Customized ringing (CRG)

Used for: FPH (optional); PRM (optional); SPL (optional); UAN (optional); VPN (optional)

> This service feature allows the subscriber to allocate a distinctive ringing to a list of calling parties.

Internet telephony end systems can easily do this based on the incoming request. Note that the distinctive ring does not need to be based only on the identity of the calling party; rings could be assigned based on call priority, the caller's organization, or any other aspect of the invitation. Sophisticated "rings" such as text-to-speech announcement of the calling party's name are also possible.

1.5.22 Destinating user prompter (DUP)

Used for: ABD (optional); FPH (optional); PRM (optional); SPL (optional); UPT (optional)

> This service feature enables to prompt the called party with a specific announcement. Such an announcement may ask the called party enter an extra numbering, e.g., through DTMF, or a voice instruction that can be used by the service logic to continue to process the call.

It is not clear that this is necessarily the best solution to this problem in an Internet environment; for example, if an HTML viewer is available, browsing via HTML/HTTP would probably make for a much better user experience than listening to a voicemail tree. However; for voice-only environments, an end system can implement the same functionality as in the traditional telephone network.

The DTMF codes could either be sent literally as audio data, or encoded using the specific DTMF encoding described in Section **??**.

1.5.23 Follow-me diversion (FMD)

Used for: FMD (core); UPT (core); VPN (optional)

> With this service feature, a user may register for incoming calls to any terminal access.

1.5.24 Mass calling (MAS)

Used for: FPH (optional); MAS (core); VOT (core)

> This service feature allows processing of huge numbers of incoming calls, generated by broadcasted advertisings or games.

This feature is similar to call gapping (section 1.5.10), except that rather than prevent the overloading calls from going through, they are processed in “early” in a distributed manner. Typically this is intended for services such as televoting.

This can be handled through similar techniques as were described for call gapping. In particular, since server addresses are resolved through DNS, load can be distributed over multiple servers with DNS load-balancing.

1.5.25 Meet-me conference (MMC)

Used for: CON (optional)

> This service feature allows the user to reserve a conference resource for making a multi-party call. At a specified date and time, each participant in the conference has to dial a designated number in order to have access to the conference.

1.5.26 Multi-way calling (MWC)

Used for: CON (core)

> This service feature allows the user to establish multiple, simultaneous telephone calls with other parties.

Internet telephony end systems may initiate as many simultaneous calls as it wishes, subject to its available network bandwidth.

1.5.27 Off-net access (OFA)

Used for: VPN (optional)

> This service feature allows a VPN user to access his or her VPN from any non-VPN station in the PSTN by using a personal identification number (PIN).

This feature consists of two issues: security authorization, and access to a VPN’s services such as the numbering plan. The former issue is best handled at a lower network layer, using appropriate data access tunneling protocols, to traverse the firewall; the latter can then be accomplished by simply treating the node like any other node on the intranet.

(We assume here that we substitute “Public Internet” for “PSTN” in the service feature description to achieve an equivalent Internet telephony feature. Connecting to an Internet VPN from the PSTN is a modem dialup running PPP or a similar protocol.)

1.5.28 Off-net calling (ONC)

Used for: VPN (optional)

> This service feature allows the user to call outside the VPN network.

This is exactly the firewall penetration problem; calls destined outside the intranet use the firewall's proxy server to reach the outside destination.

1.5.29 One number (ONE)

Used for: CD (core); CRD (core); FPH (core); PRM (core); SPL (core); UAN (core)

> This feature allows a subscriber with two or more terminating lines in any number of locations to have a single telephone number. This allows businesses to advertise just one telephone number throughout their market area and to maintain their operations in different locations to maximize efficiency. The subscriber can specify which calls are to be terminated on which terminating lines based on the area the calls originate.

The problem of assigning requests to "nearby" (in a network sense) servers is the wide-area service location problem, and is currently being researched in its general case [45].

1.5.30 Origin dependent routing (ODR)

Used for: CD (optional); DCR (optional); FPH (optional); MAS (optional); PRM (optional); SPL (optional); VOT (optional); UAN (optional); UDR (optional)

> This service feature enables the subscriber to accept or reject a call, and in case of acceptance, to route this call, according to the calling party geographical location. This service feature allows the served user to specify the destination installations according to the geographical area from which the call was originated.

Note that in this context, "routing" refers to the selection of an endpoint, not the network path a call takes to that endpoint. (Q.1211 is inconsistent in this usage.) In an Internet context, this is essentially the same service as One Number in section 1.5.29.

In the Internet environment, geographic location does not tend to be terribly relevant for network purposes, except for inter-continental connections. Thus, presumably some other criterion, such as shortest network path or cheapest acceptable QoS, will be used instead; again, this is the wide-area service location problem.

1.5.31 Originating call screening (OCS)

Used for: FPH (optional); MCI (core); MAS (optional); OCS (core); PRM (optional); SPL (optional); VOT (optional); UAN (optional)

> This service feature allows the served user to bar calls from certain areas based on the District code of the area from which the call is originated.

In an Internet context, the corresponding address attribute to the District code would be either the DNS domain or the IP network of the originating address. Either a SIP proxy, H.323 gatekeeper or an end system could filter requests on these attributes, or the entire address. Alternately, the filtering could be performed on the media's destination network address.

However, it is important to realize that an unauthenticated Internet telephony signaling request provides no guarantee that it actually came from the party associated with the address claimed in the source field. The only way this can be verified is with cryptographic signing of the request, along with an infrastructure for distribution of public keys. Media addresses are more likely to be meaningful in the absence of any authentication (at least if the remote party does indeed appear to be receiving the media), though media could in principle be passing through a forwarding gateway.

Note that Q.1211's choice to name this service feature "Originating call screening" contradicts every other usage of the term. Normally, it is called "Terminating call screening", as indeed it is in the service cited in section 1.6.24. The distinction is whether the name is chosen based on the fact that the screening is done at the termination point, as is normally done, or on the fact that the filters act upon the originating address, as is apparently the case here.

1.5.32 Originating user prompter (OUP)

Used for: ACC (core); AAB (core); CCC (core); FPH (optional); MAS (optional); PRM (optional); VOT (optional); UAN (optional); UPT (optional); VPN (optional)

> This service feature allows a served user to provide an announcement which will request the caller to enter a digit or series of digits via a DTMF phone or generator. The collected digits will provide additional information that can be used for direct routing or as a security check during call processing.

It is not clear that this service per se is useful in the Internet context.

1.5.33 Personal numbering (PN)

Used for: UPT (core)

> This service feature supports a UPT number that uniquely identifies each UPT user and is used by the caller to reach that UPT user.

Reaching a user through a personal address is accomplished simply by a proxy or redirection server which locates the user.

There are three possible levels of scoping for personalized SIP addresses. The one that will be the most common is specialized usernames. Analogous to the current e-mail forwarding servers, services can be set up which permanent SIP addresses to their subscribers, redirecting to actual end-system addresses. The domain is that of the service provider. An intermediate sort of permanent personal addressing would be vanity DNS domains; permanent SIP addresses in the user's chosen name could be set up much as personal web pages can be set up today.

Finally, a provider could choose to set up a permanent addressing scheme of host-independent naming. End systems could access this information when encountering addressing schemes which do not conform to the standard SIP addresses; this is analogous to Netscape's Netcenter name lookup, for instance.

1.5.34 Premium charging (PRMC)

Used for: PRM (core)

> This service feature allows for the pay back of part of the cost of a call to the called party.

This is (in the U.S.) 900-number service; see section **??** for discussion of billing in the Internet context. This is perhaps the simplest case of billing, as in this case the telephone company is simply acting as a settlement authority so the called party can bill the caller for their services. In an Internet environment this service could either be negotiated directly between the two parties, or some external settlement authority trusted by both parties could be used.

1.5.35 Private numbering plan (PNP)

Used for: VPN (core)

> This service feature allows the subscriber to maintain a numbering plan within his private network, which is separate from the public numbering plan.

Since in the Internet environment, addressing is always controlled by independently administered systems, this largely becomes trivial. If a numbering plan should be hidden or partially hidden from the public, a proxy can pretend to know nothing about the private addresses when they come from the outside; or proxies or end systems can re-write addresses to maintain a mapping from local to global names. Call routing for internal calls can be based either on the internal addresses or performed transparently by internal routers.

For instance, many systems re-write internal e-mail addresses today so that usernames specified without corresponding hosts are delivered as though they were addressed to the user at the local domain. A similar re-writing system could easily be set up for Internet telephony systems.

1.5.36 Reverse charging (REVC)

Used for: FPH (core)

> This service feature allows the service subscriber (e.g. freephone) to accept to receive calls at its expense and be charged for the entire cost of the call.

This is (U.S.) toll-free-number service; see Section **??** for a discussion of billing in the Internet context.

1.5.37 Split charging (SPLC)

Used for: SPL (core); UPT (core)

> This service feature allows for the separation of charges for a specific call, the calling and called party each being charged for one part of the call.

This is a generalization of the previous billing problems, with multiple parties responsible for payment. Again, see Section **??**.

1.5.38 Terminating call screening (TCS)

Used for: TCS (core)

> This service feature allows the user to screen calls based on the terminating telephone number dialed.

This can easily be done by a proxy which can force itself to be on the outgoing signalling path of an end system, or by administrative control over an originating end system (see section **??**).

As with OCS (section 1.5.31), the use of the term “Terminating call screening” for this service feature in Q.1211 contradicts every other usage of the term. It is much more commonly known as “Originating call screening,” since it occurs at or near the origination point of the call.

1.5.39 Time dependent routing (TDR)

Used for: CD (optional); DCR (optional); FPH (optional); MAS (optional); PRM (optional); VOT (optional); UAN(optional); UPT (optional); UDR (optional); VPN (optional)

> This services feature allows the served user to apply different call treatments based on the time of day, day of week, day of year, holiday, etc.

Any server or gatekeeper with a clock could be programmed to make decisions of this sort.

1.6 Capability Set 1: Services

This section describes all the services listed in Q.1211, Annex B, section 1. These services are built out of the service features listed in section 1.5.1. For each service, we list which service features are core (essential) to it, or optional, according to Q.1211.

See table 1.6 for a summary of the characteristics of each service.

1.6.1 Abbreviated dialling (ABD)

Uses: ABD (core); LOG (optional); CPM (optional); DUP (optional)

> This service is an originating line feature that allows business subscribers to dial others in their company using, e.g., only four digits even if the calling user's line and the called user's line are served by different switches.

This is entirely covered by the descriptions of the component service features; see particularly the discussion of the ABD service feature in section 1.5.2.

1.6.2 Account card calling (ACC)

Uses: ABD (core); AUTZ (core); LOG (optional); OUP (core)

> The account card calling service allows subscribers to place calls from any normal access interface to any destination number and have the cost of those calls charged to the account specified by the ACC number.

This is another generalization of the previous billing problems, namely the separation of the responsibility for payment with the recipient of a service. See Section **??**.

1.6.3 Automatic alternative billing (AAB)

Uses: ABD (optional); AUTZ (core); LOG (optional); OUP (core)

> This service allows a user to call another user and ask him to receive the call at his expenses.

This is another generalization of billing; again, see Section **??**.

Service		**Characteristics**		
Code	Name	Section	Location	Call Time
Authentication				
SEC	Security screening	1.6.17	All	All
Billing				
ACC	Account card calling	1.6.2	-	-
AAB	Automatic alternative billing	1.6.3	-	-
CCC	Credit card calling	1.6.9	-	-
FPH	Freephone	1.6.12	-	-
PRM	Premium rate	1.6.16	-	-
SPL	Split charging	1.6.22	-	-
UPT	Universal personal telecommunications	1.6.26	Proxy / Redirect	Setup
Filtering				
OCS	Originating call screening	1.6.15	End / Proxy	Setup
TCS	Terminating call screening	1.6.24	End / Proxy	Setup
Forwarding				
CD	Call distribution	1.6.4	Proxy / Redirect	Setup
CF	Call forwarding	1.6.5	End / Proxy / Redirect	Setup
CRD	Call rerouting distribution	1.6.6	End / Proxy	Setup
DCR	Destination call routing	1.6.10	Proxy / Redirect	Setup
FMD	Follow-me diversion	1.6.11	Proxy / Redirect	Indep. of call
SCF	Selective call forwarding on busy/don’t answer	1.6.18	End / Proxy / Redirect	Setup
UAN	Universal access number	1.6.25	Proxy / Redirect	Setup
Multi-party				
CON	Conference calling	1.6.8	End / Other	In call
Other services				
VOT	Televoting	1.6.23	End	In call
UDR	User-defined routing	1.6.27	-	Setup
VPN	Virtual private network	1.6.28	-	-
Translation				
ABD	Abbreviated dialling	1.6.1	End / Proxy / Redirect	Setup
User interface				
CF	Call forwarding	1.6.5	End / Proxy / Redirect	Setup
Other				
CCBS	Completion of calls to busy subscriber	1.6.7	End	Setup
MCI	Malicious call identification	1.6.13	End / Proxy	Setup
MAS	Mass calling	1.6.14	Proxy	Setup

Table 1.6: This table indicates the characteristics of each service, and sorts them into rough categories. For the “location” column, “End” means the service can be provided at an end system, “Proxy” means at a proxy server, and “Redirect” at a redirection server.

1.6.4 Call distribution (CD)

Uses: CD (core); LOG (optional); CPM (optional); ONE (core); ODR (optional); TDR (optional)

> This service allows a subscriber to have incoming calls routed to different destinations, according to an allocation law which may be real-time managed by the subscriber.
>
> Three types of law may exist:
>
> - circular distribution, where the calls are routed to the different locations with a uniform load;
> - percentage distribution, where the calls are routed to the different locations according to a percentage;
> - hierarchical distribution, where the first location to be chosen is the first met in the priority list.
>
> In addition, congestion at one location may cause overflow calls to be rerouted to an alternate location.

1.6.5 Call forwarding (CF)

Uses: CF (core); LOG (optional); CPM (optional)

> Call forwarding allows the called user to forward calls to another telephone number when this service is activated. With this service, all calls destined to the subscriber's number are redirected to the new telephone number.
>
> This service is under control of the subscriber and can be activated/deactivated by the subscriber.
>
> When this service is activated, the subscriber's line will receive an alerting ring, "reminder ring," to indicate that the service is activated.

Call forwarding itself is described in section 1.5.8 in the discussion of the service feature of the same name.

Reminder ring is somewhat ill-defined in Internet telephony, due to the separation of end systems and addresses; if a proxy is doing the forwarding, it is not clear what end system should be alerted of the forward, since call forwarding is just one specific case of the general user-location problem. A proxy could be configured with a specific end system to alert, however; though no signals are currently defined in SIP to request that an end system perform a reminder ring, it would be a relatively simple extension to the protocol, and several have been informally proposed.

If an end system is doing forwarding on its own, of course, it may perform a reminder ring if it so wishes.

1.6.6 Call rerouting distribution (CRD)

Uses: LIM (optional); LOG (optional); QUE (optional); CPM (optional); ONE (core)

> This service permits the subscriber to have his incoming calls encountering a triggering condition (busy, specified number of rings, queue overload or call limiter) rerouted according to a predefined choice: the calls may be rerouted to another destination number (including pager or vocal box), rerouted on a standard or customized announcement, or queued.

This is all easily possible in a proxy or end system; proxies can also handle the additional case when a device is not responding, whereas end systems keep enough call state to do call limiting more reliably. For control, see the description under CPM in section 1.5.19.

1.6.7 Completion of calls to busy subscriber (CCBS)

Uses: ACB (core); LOG (optional); CW (optional)

> This service allows a calling user encountering a busy destination to be informed when the busy destination becomes free, without having to make a new call attempt.

(This is noted in Q.1211 as not actually being possible to implement using only "Type A" (single-ended, single-point-of-control) service features, the type of features that Capability Set 1 specifies.)

The full semantics of this feature are somewhat elaborate: the calling user is alerted to the fact that the destination is willing to accept the call, and can re-initiate the call (automatically) at his or her leisure, before the called user is alerted to the call attempt. This cannot be implemented with the basic SIP standard, since the protocol does not provide any way to convey that information.

There are two proposed ways to create a feature which simulates this behavior to some extent. One is simply polling the destination periodically with new call attempts, and then alerting the calling user once it is available; another is specifying "Call-disposition: queue" when the call is placed (as specified in the SIP Call Control draft), to request that the call be queued rather than rejected if the destination is not available. Both these strategies have the disadvantage that the calling party might forget about the call, or get involved in another one, in which case a call would be placed but the calling party would not be present when the called party picked up.

It has also been proposed that the current standards work on Presence Information Protocols [46] could also be used to implement this feature, since the semantics of "the individual is available for communication" are similar. One of the proposals for this functionality involves using some simple extensions to SIP.

Most existing traditional telephone systems also cannot implement this service in all cases, especially between systems under separate administrative control, as their signalling protocols do not support it.

1.6.8 Conference calling (CON)

Uses: LOG (optional); COC (optional); CPM (optional); MMC (optional); MWC (core)

> Conference calling allows the connection of multiple parties in a single connection. The number of parties connected simultaneously will vary based on bridging requirements.
>
> **Conference calling add-on** This service allows the user to reserve a conference resource for making a multi-party call, indicating the date, time, and conference duration. Once the conference is active, the user controls the conference, and may add, drop, isolate, reattach or split parties.
>
> **Conference calling meet-me** This service allows the user to reserve a conference resource for making a multi-party call, indicating the date, time, and conference duration. In due time, each participant in the conference has to dial a special number whcih has been attached to the booked conference, in order to access the conference bridge.

There are a number of ways that conference calling can be handled in Internet telephony systems. The one requiring the least effort from end systems is an architecture similar to conference calling meet-me; the multiple parties all connect to a specified address, which controls a conference bridge. This conference bridge will perform mixing of the media from each station, and send it out to all the other stations involved in the call.

1.6.9 Credit card calling (CCC)

Uses: ABD (optional); AUTZ (core); LOG (optional); OUP (core)

> The credit card calling service allows subscribers to place calls from any normal access interface to any destination number and have the cost of those calls charged to the account specified by the CCC number.

This is yet another case of the billing problem. See Section **??**.

1.6.10 Destination call routing (DCR)

Uses: CD (core); LOG (optional); CPM (optional); ODR (optional); TDR (optional)

> This service allows customers to specify the routing of their calls to destinations according to
>
> - time of day, day of week, etc.;
> - area of call origination;
> - calling line identity of customer;

- services attributes held against the customer;
- priority (e.g. from input of a PIN);
- charge rates applicable for the destinations;
- proportional routing of traffic

See the service feature descriptions, particularly CD, ODR, and TDR in sections 1.5.7, 1.5.30, and 1.5.39; all of this can be done in a proxy or redirect server, possibly under the control of a call processing language.

1.6.11 Follow-me diversion (FMD)

Uses: LOG (optional); CPM (optional); FMD (core)

> This service allows the subscriber to remotely control his call forwarding capabilities, basically the number to which the calls are forwarded, from any point in the network.

1.6.12 Freephone (FPH)

Uses: AUTC (optional); CD (optional); CFC (optional); GAP (optional); LIM (optional); LOG (optional); QUE (optional); CPM (optional); CRA (optional); CRG (optional); DUP (optional); MAS (optional); ONE (core); ODR (optional); OCS (optional); OUP (optional); REVC (core); TDR (optional)

> Freephone allows the served user having one or several installations to be reached from all or part of the country, or internationally as appropriate, with a freephone number and to be charged for this kind of call.

See all the relevant service feature descriptions, particularly REVC (section 1.5.36), and the general discussion of billing in section **??**.

1.6.13 Malicious call identification (MCI)

Uses: LOG (core); OCS (core)

> Malicious call identification allows the service subscriber to control the logging (making a record) of calls that are received that are of a malicious nature.

Any entity through which Internet telephony signaling messages pass and which has some sort of storage can log information about these messages. If legal non-repudiability is needed, a customer can have a trusted third party perform proxying and logging for them.

See under OCS (section 1.5.31) about the complexities of guaranteeing the authenticity of SIP messages, however.

1.6.14 Mass calling (MAS)

Uses: CD (optional); LIM (optional); LOG (optional); QUE (optional); CPM (optional); CRA (optional); MAS (core); ODR (optional); OCS (optional); OUP (optional); TDR (optional)

> Mass calling involves instantaneous, high-volume traffic which is routed to one or multiple destinations. Calls can be routed to these destination numbers based on various conditions, such as the geographical location or the time of day.

See the discussion of the MAS service feature in section 1.5.24.

1.6.15 Originating call screening (OCS)

Uses: LOG (optional); CPM (optional); OCS (core)

> This services allows the subscriber to authorize outgoing calls, through the use of a screening list. This list may be managed by the subscriber. The user may override the restriction by giving a PIN.

Note that unlike the service feature with the same name, the Q.1211 summary for this service describes it correctly; thus, the proper core service feature for this service is described under TCS, in section 1.5.38. See that section for a full description of the implementation of this service.

1.6.16 Premium rate (PRM)

Uses: CD (optional); CFC (optional); GAP (optional); LIM (optional); LOG (optional); QUE (optional); CPM (optional); CRA (optional); CRG (optional); DUP (optional); ONE (core); ODR (optional); OCS (optional); OUP (optional); PRMC (core); TDR (optional)

> This service allows to pay back part of the call cost to the called party, considered as an added value service provider.

See the discussion of the PRMC service feature in section 1.5.34, and of billing in general in **??**.

1.6.17 Security screening (SEC)

Uses: AUTC (core); LOG (optional); CPM (optional)

> This capability allows security screening to be performed in the network before an end-user gains access to the subscriber's network, systems, or applications. Access code abuse detection is a capability which will generate a report on the invalid access attempts: how many, over what time period, by whom, and from where.

Facilities which implement authorization codes should have this kind of auditing in place; other Internet services which require authorization or other security approval should have a similar security infrastructure.

1.6.18 Selective call forwarding on busy/don't answer (SCF)

1.6.19 Selective call forwarding

1.6.20 Call forwarding on busy

1.6.21 Call forwarding on don't answer (no reply)

Use: CFC (core); GAP (core); LOG (optional); CPM (optional)

> This service allows the called user to forward particular pre-selected calls if the called user is busy or does not answer within Y seconds or X rings.

These services can easily be handled by either a proxy, which could also handle "selective call forwarding on device not responding", or an end system. They could be under control of a call processing language.

The "Selective call forwarding" case can also be handled by a redirect server.

1.6.22 Split charging (SPL)

Uses: CD (optional); CFC (optional); GAP (optional); LIM (optional); LOG (optional); QUE (optional); CPM (optional); CRA (optional); CRG (optional); DUP (optional); ONE (core); ODR (optional); OCS (optional); SPLC (core)

> This service allows a split charging, the calling and the called party being each charged for one part of the call.

See the discussion of SPLC in section 1.5.37, and of billing in general in **??**.

1.6.23 Televoting (VOT)

Uses: CD (optional); LIM (optional); LOG (optional); QUE (optional); CPM (optional); CRA (optional); MAS (core); ODR (optional); OCS (optional); OUP (optional); TDR (optional)

> Televoting enables subscribers to survey public opinion using the telephone network. Persons wishing to respond to an opinion poll can call advertised televoting numbers to register their votes. The charging is to the discretion of the service subscriber.

It is not clear in an Internet context how useful this actually is; a web page or email would be a much more natural way to present this service, and would be more efficient as it would not require setting up a voice path. However, an end system could implement this service without any intrinsic signalling support being necessary. See the OUP service feature in section 1.5.32.

1.6.24 Terminating call screening (TCS)

Uses: LOG (optional); CPM (optional); TCS (core)

> Terminating calls may be controlled by the terminating call screening capability. This allows the subscriber to specify that incoming calls be either restricted or allowed, according to a screening list and optionally, by time of day control.

Note that unlike the service feature with the same name, the Q.1211 summary for this service describes it correctly; thus, the proper core service feature for this service is described under OCS, in section 1.5.31. See that section for a full description of the implementation of this service.

1.6.25 Universal access number (UAN)

Uses: CD (optional); CFC (optional); GAP (optional); LIM (optional); LOG (optional); QUE (optional); CPM (optional); CRA (optional); CRG (optional); ONE (core); ODR (optional); OCS (optional); OUP (optional)

> This service allows a subscriber with several terminating lines in any number of locations or zones to be reached with a unique directory number. The subscriber may specify which incoming calls are to be routed to which terminating lines, based upon the area the call originated.

This is the same as the ONE service feature (section 1.5.29), with other features optionally "behind" it.

1.6.26 Universal personal telecommunications (UPT)

Uses: AUTZ (core); LOG (optional); CPM (optional); CRA (optional); DUP (optional); FMD (core); OUP (optional); PN (core); SPLC (core)

> This service provides personal mobility by enabling a user to initiate any type of service and receive any type of call on the basis of a unique and personal network-independent number, across multiple networks, at any user-network access (fixed, movable or mobile), irrespective of geographic location, limited only by terminal and network capabilities.

For Internet Telephony features, this is essentially the same as the FMD service feature (section 1.5.23) with authentication; once billing is resolved, the Split Charging issue needs also of course to be resolved.

1.6.27 User-defined routing (UDR)

Uses: LOG (optional); CPM (optional); ODR (optional); TDR (optional)

> This capability allows the subscriber to specify how outgoing calls, from the subscriber's location, shall be routed, either through private, public, or virtual facilities or a mix of facilities, according to the subscriber's routing preferences list. These lists will typically apply to individual lines or to several lines at the subscriber's location.

Note that despite the fact that both this and other services refer to the ODR and TDR features, they are referring to quite different things: what network is used to place a call, vs. the destination of that call. See section 1.5.30.

For most non-dysfunctional networks, the choice of network for connectivity is only of concern for media flows, not signalling flows. As such, this is not properly the domain of SIP, but rather that of the mechanisms which control the routing of media packets. Normally this would most likely be under the control of administrators who set up a network's routing table, but users could dynamically control it through loose source routing or a QoS routing protocol.

In somewhat unusual circumstances, however, a user may wish to be able to specify what networks signalling flows cross. For instance, the network's standard routing might cross networks that *are* dysfunctional; or security might be so important that encryption is not sufficient, for instance if the number of call attempts made is sensitive information. In these cases, the same techniques mentioned above for media routing could also be used for signal routing, since the underlying network does not, of course, distinguish between them.

1.6.28 Virtual private network (VPN)

Uses: ABD (optional); ATT (optional); AUTC (optional); AUTZ (optional); CD (optional); CHA (optional); LOG (optional); QUE (optional); TRA (optional); CUG (optional); COC (optional); CPM (optional); CRA (optional); CRG (optional); FMD (optional); OFA (optional); ONC (optional); OUP (optional); PNP (core); TDR (optional)

> This service permits the subscriber to build a private network by using the public network resources. The subscriber's lines, connected on different network switches, constitute a virtual PABX, including a number of PABX capabilities, such as private numbering plans (PNP), call transfer, call hold, and so on.

In an Internet telephony environment, this would consist of two parts: an IP-level Virtual Private Network, and an internal Internet telephony system running over it. The exact definition of an IP VPN is still a matter of some debate, but it would seem to include such features as quality of service support, network-level encryption, and on-demand dynamic bandwidth allocation, so as to provide the illusion of a private intranet between geographically separated locations. See Paul Ferguson's white paper *What is a VPN* [47] for a thorough discussion of this.

On this intranet, various private numbering plans can be maintained (see PNP, section 1.5.35), and within it any of the service features described in this document can be implemented much as they would be on the public Internet or a physically contiguous intranet.

1.7 Capability Set 2

The draft of recommendation Q.1221 introduces a large number of new services and service features beyond those listed in Q.1211. Many of these do not need to be addressed here in any great detail: a large number of them concern "Service Management Services," which are used to allow direct user control of the network's services. Since the architecture of Internet Telephony puts the network's services under the user's control intrinsically, addressing these services is not directly necessary.

Capability Set 2 also lists some services which in Capability Set 1 were merely service features, and it repeats some service features from CS1 with clearer or more cleanly-separated descriptions. The descriptions of those service features are not repeated here.

The other significant services and service features which are new in CS2 are addressed below.

1.7.1 Wireless services

A large number of the CS2 services are designed to support wireless (mobile) networks. Internet telephony systems can either use mobile IP or application-layer signaling (Section **??**).

1.7.2 Inter-network services

There are a number of services of Capability Set 2, such as Internetwork freephone (IFPH), Internetwork televoting (IVOT), and so forth, which are essentially services of Capability Set 1 extended so that they are guaranteed to work between multiple networks.

As its name implies, the Internet is intrinsically a collection of multiple networks. Internet telephony signalling is inherently multi-domain, and all the services described in sections 1.5.1 and 1.6 which operate between multiple machines will work regardless of whether those machines are on the same subnetwork or opposite sides of the planet.

1.7.3 Multimedia

> This service allows a subscriber to receive or send an integrated communication consisting of mixtures of voice, data, image, and video information. A key capability will be the ability to synchronize and control devicery of information from disparate sources (e.g. voice and data). This will include controlling delivery from multiple sources to a single recipient and from a single source to multiple recipients.
>
> The subscriber will also desire the ability to tailor a particular service to the type of terminating device or subscriber preference (i.e. turning off the video feed). Another key aspect of this service is that additional capabilities may be requested during the call (i.e. adding data capabilities to an existing voice connection).

The Internet's media delivery protocol, RTP [48], can deliver many simultaneous synchronized media formats; it can transparently perform either point-to-point or multicast delivery; it can disambiguate multiple media senders without their previous coordination; and many other features besides.

1.7.4 Call pick-up

> This service feature enables a user to associate a call request to an already alerting call. The alerting call awaits answer while the user originating call pick-up signals to the network a desire to connect to the alerting call. The network then connects the call parties.

1.7.5 Calling name delivery

> This service feature gives to the network operator the capability to display/announce the name of the calling party to the calling name delivery user (the called party) prior to answer, thus allowing this user to screen or distinctively answer the call.

1.8 Emergency Call Services

The telephone system has become the primary means of summoning emergency assistance, dating back to the early days of its development. (In cities, street-side call boxes, using telegraph technology, were the only other mechanism to call for help, but these are easy targets for pranks.)

Initially, each town and city had its own number for the fire and police departments. Starting in 1957, proposals were made to establish a universal emergency number, with 9-1-1 as the common number; the first 911 call was made in 1968 in Haleyville, Alabama. In the 1970s, enhanced services were added which allowed emergency operators to determine the address and phone number of the person calling, even if the person was too panicked, young or injured to provide directions. In 1999, nearly 93% of the population of the United States can call for help with 911, with 95% offering enhanced 911 services. However, 911 is still not universally available in rural areas, since only 50% of the geographic area is covered by 911 service. Canada also uses 911 as the universal emergency number, while other countries usually still have separate numbers for police and fire department.

Enhanced 911 service (E911) consists of three functions: identifying the number of the caller, his or her street address and routing the call to the right emergency operator. The emergency operator is located at a public safety answering point (PSAP). Automatic number identification (ANI) is similar to the system used to deliver the caller's number to the owner of a toll-free number and allows the emergency operator to call back if the call gets disconnected. Automatic location identification (ALI) retrieves the subscriber record based on the directory number, with a format given by the Master Street Address Guide (MSAG). The subscriber record contains information about the street address and other location information. (Indeed, in many rural communities, one of the harder parts of installing 911 service is to assign unique street addresses.)

Based on the ANI information, a selective router looks up the numbers of police, fire and emergency medical services responsible for the subscribers location. An emergency service number (ESN) summarizes this information and makes it easy to transfer a call to the appropriate emergency service.

- The emergency operator needs to be able to call back even if the directory number is being forwarded to another number.

- The emergency operator should be able to override any other calls or reach a caller that has been disconnected, but then left the phone off the hook. Services such as call-waiting or three-party calls are restricted when making a 911 call.

- While address information is sufficient for single-family homes and apartments, it is far too coarse for larger businesses [49]. Indeed, a business may well have locations spread around several jurisdictions, with a single phone number. Until recently, most PBXs would only deliver the main switchboard number via ANI, not the actual extension making the call. Even if it were to provide the station number, the ALI database needs to be updated with the correct room and building information. It is likely that the FCC will require multi-line telephone systems (MLTS) at least in buildings with more than 40,000 square feet.

By October 2001, wireless phone services will have to be equipped so that emergency crews can find the caller within one minute. Systems based either on installing global positioning system (GPS) in the handset or triangulation from different base stations have been proposed [50, 51, 52].

Internet telephony offers the opportunity not only to supply current emergency services, but also to add functionality. For example, multimedia communications can be used to provide video images or biometrics. The call setup can provide additional medical background, without having to store the information in a central database. Due to the faster call setup times and the ease of redirection, calls can be forwarded and transferred to emergency personnel en route to the caller or a nearby PSAP, in case the primary PSAP serving the caller's location is overloaded. Unlike current PSAPs, which require dedicated equipment, an Internet-based PSAP could be set up anywhere Internet access is available, providing easy relocation should the PSAP be unavailable during natural disasters. The ability to indicate language capabilities of the caller can help route the call to an operator, without the additional delay of having a general operator try to ascertain the language of the caller.

However, Internet-telephony 911 services face a number of difficulties. To begin with, there is no obviously no universal identifier for emergency services, unless they are simply gatewayed into the phone system.

There are two approaches to providing Internet emergency services: either using the Internet end-to-end or by gatewaying the call to a regular 911 service. Gatewaying will certainly be the more common approach for many years to come. However, the location of the gateway may be far removed from the location of the caller, so the gateway has to be able to insert the caller's phone number instead of its own.

However, the gateway may not have any information on the phone number that the caller used to reach the Internet, even assuming she used a modem at all, rather than, say, a hotel Ethernet or cable modem.

The routing mechanism that receives 911 calls from this gateway now may have to deal with phone numbers far outside its local area, since the call to the gateway may have originated at a hotel thousands of miles away from the subscriber's home location.

1.9 The Economics of the Telephone System

Telephone services are, by far, the most important sector of the whole communications industry and are a major component of the economy of developed nations. For example, total toll service revenues for U.S. long distance carries in 1998 was 88.6 billion dollars, with total toll revenue of 98.6 billion dollars, which was 85% of all U.S. telecommunications service revenues. About 1.02 million people work in the telephone communications industry, not counting manufacturing.

About 500 million households, approximately 34% of all households, in the world have telephone service. The other statistic used to characterize telephone penetration is the number of telephone lines per 100 inhabitants. This *teledensity* was at 12.80 in 1996 for the world at large, with most developed countries having densities of 40 to 60.[11]

1.9.1 Prices for Telephone Calls

Over the years, three trends can be observed in long-distance rates. First, per-minute rates have more or less continuously declined in both absolute and relative terms, particularly for international calls to countries where there is telecommunications competition (Fig. 1.12). (For example, the best generally available per-minute rate from the U.S. to the U.K. is roughly half of the rate to Germany and a third of the rate to Poland, even though the distances are comparable. Germany has only recently allowed competition in the provision of telephone service.)

Secondly, distance is playing an increasingly small factor in prices. For example, a one-mile call from New Jersey to Manhattan costs the same 10 cents as a call from San Francisco to London. Currently, due to lack of competition, prices can be highest when calling just outside the local calling area (Tab. 1.7). The removal of distance as a factor in telephone rates is called *postalization*, since it makes phone calls similar to how letters are priced.

Table 1.8 shows the capacities, measured as 64 kb/s circuits, and costs of the transatlantic cables. (The number of simultaneous phone calls is larger since silence suppression and voice compression are used to increase capacity.) From 1983 to 2000, the transatlantic capacity has grown by 64% per year, with voice path costs declining by 41% per year. The cost of a minute of transatlantic calls is derived from the investment cost assuming a typical utilization of 8 hours per day, with 50% of circuits are not activated. The cost per circuit is the annual investment cost for a 25-year lifetime.

Third, while there used to be large differences between peak and off-peak rates, the distinction has all but vanished in the United States and is decreasing elsewhere. In 1994, the German Telecom, for example, had six different rates throughout the day, with even a special "moonshine rate"[12] from 2 am to 5 am. Customers were sent a little multi-colored cardboard slide rule, with

[11] The ITU publishes the *Yearbook of Statistics* that provide this information.

[12] They did use the English term for a while. . .

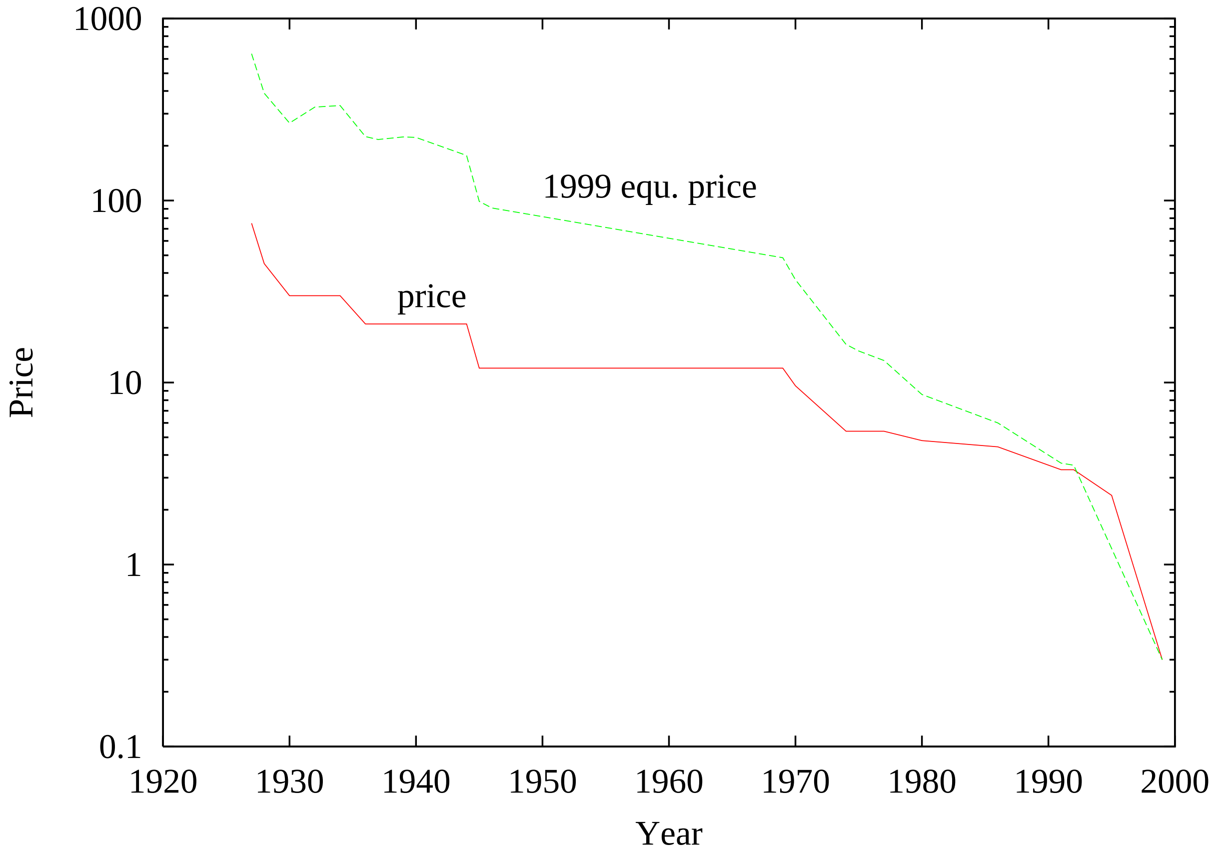

Figure 1.12: Standard daytime telephone rates from New York to London for a three-minute call (courtesy A. Odlyzko)

		Philadelphia		Chicago		San Francisco	
	mult.	first	add.	first	add.	first	add.
1984	1.59	0.57	0.34	0.62	0.43	0.74	0.49
1985	1.54	0.54	0.35	0.58	0.40	0.69	0.46
1986	1.51	0.51	0.33	0.55	0.38	0.65	0.42
1987	1.45	0.33	0.24	0.36	0.30	0.40	0.32
1988	1.40	0.29	0.21	0.31	0.27	0.36	0.28
1989	1.33	0.27	0.21	0.30	0.26	0.32	0.27
1990	1.26	0.22	0.215	0.24	0.239	0.25	0.25
1991	1.21	0.22	0.215	0.24	0.239	0.25	0.25
1992	1.18	0.21	0.21	0.23	0.23	0.25	0.25
1993	1.14	0.22	0.22	0.23	0.23	0.25	0.25
1994	1.11	0.23	0.23	0.24	0.24	0.25	0.25
1995	1.08	0.26	0.26	0.27	0.27	0.28	0.28
1996	1.05	0.27	0.27	0.28	0.28	0.30	0.30
1999	1.00	0.10	0.10	0.10	0.10	0.10	0.10

Table 1.7: AT&T basic long distance rates from New York, business day; first and additional minutes, with multiplier to arrive at equivalent 1999 prices (courtesy A. Odlyzko)

which they could figure out what the current price might be. This has now been reduced to three rates.

Finally, an increasing portion of the average consumer phone bill is fixed, including the national access fee ("presubscribed interexchange carrier charge", $1.07 in 1999) and the FCC subscriber line charge ($3.50 in 1999) which pays local carriers to originate and terminate long-distance calls [53] or charges for local-number portability. In addition, the local carriers charge the interexchange carriers approximately 3.71 cents per minute, of which 0.82 cents are for originating access and 0.16 cents for terminating access. These are averages, as charges differ for each local exchange carrier.

An average household spends $67 per month on telephone service, or about 2.3% of total expenditures. 70% of this amount is spent on long-distance service. Neither the fraction of household expenses nor the ratio of local to long-distance expenditures has changed significantly since 1980. Through the "lifeline" program, the cost of telephone service is subsidized for about 5.3 million low-income households, with a subsidy between $5.25 and $10.50 per month, for a total cost of 304 million.

1.9.2 Infrastructure Costs

The cost of providing telephone service can be roughly divided into infrastructure, such as buildings, switching and transmission, overhead, access (the wires from residences and businesses to the central office) and operation support systems (OSS). Fig. 1.13. The figure shows indicates that only a small fraction of the cost is caused by the component discussed most frequently, the tele-

System	Year	technology	cost	circuits	$/circuit	$/minute
TAT-1	1956	coax cable	$49.6M	40	213,996	2.443
TAT-2	1959	coax cable	42.7M	44	167,308	1.910
TAT-3	1963	coax cable	50.6M	79	111,027	1.267
TAT-4	1965	coax cable	50.4M	62	140,238	1.601
TAT-5	1970	coax cable	70.4M	648	18,773	0.214
TAT-6	1976	coax cable	197.0M	3,200	10,638	0.121
TAT-7	1983	coax cable	180.0M	3,821	8,139	0.093
TAT-8	1988	fiber optic	360.0M	6,048	10,285	0.117
TAT-9	1992	fiber optic	406.0M	10,584	6,628	0.076
TAT-10	1992	fiber optic	300.0M	18,144	2,857	0.033
TAT-11	1993	fiber optic	280.0M	18,144	2,667	0.030
TAT-12	1996	fiber optic	378.0M	60,480	1,080	0.012
TAT-13	1996	fiber optic	378.0M	60,480	1,080	0.012
Gemini	1998	fiber optic	520.0M	214,920	371	0.004
AC-1	1998	fiber optic	850.0M	483,840	304	0.003

Table 1.8: Transatlantic cable systems [53]

phone switch. However, it should be noted that Internet telephony may indirectly affect the cost of OSS and overhead, e.g., by allowing customers to self-provision services.

Unfortunately, there are few publically available studies that directly detail the cost of infrastructure, transmission and switching components. As an order-of-magnitude example, a trunk switch with 100 T1 lines costs about $2.5 to $3 million[13], while each port on a TDM line switches costs roughly $400 to $500. On average, a PBX port costs about $567 in 1996, including the telephone set, while other sources quote typical local phone switches at $100 per port.

Most of the cost of installing fiber is digging trenches. Fiber itself costs roughly $1/foot, while trenching costs can vary between less than $1 in rural areas to more than $150 per foot in urban locations. If strung between telephone poles, trenching costs are replaced by pole access costs, roughly $1 per foot. [54] Other sources cite costs of $30,000 to $80,000 for laying one mile of long-distance fiber [55].

We have attempted, in Table **??** to compare the cost of switching voice bits via traditional circuit-switched means and by Ethernet, switches and routers.

The table illustrates that port costs are roughly independent of the speed, i.e., a 64 kb/s port costs the same as a 100 Mb/s port.

This comparison obviously always has to remain apples-to-oranges. For example, telephone switches can switch more than 100,000 subscriber lines, while most packet networks have far fewer ports per switch[14] A typical modern design has workgroup Ethernet switches in a hallway telecommunication closet, gathering all "horizontal" wiring. The wiring closets are connected via "vertical" wiring to larger switches, typically with 100 Mb/s, probably upgraded to 1 Gb/s ports

[13] *America's Network*, March 1999

[14] A typical modern Ethernet switch has 24 10/100 Mb/s ports, with stackable switches providing about 100 ports and large chassis-based switches up to 500 10/100 ports.

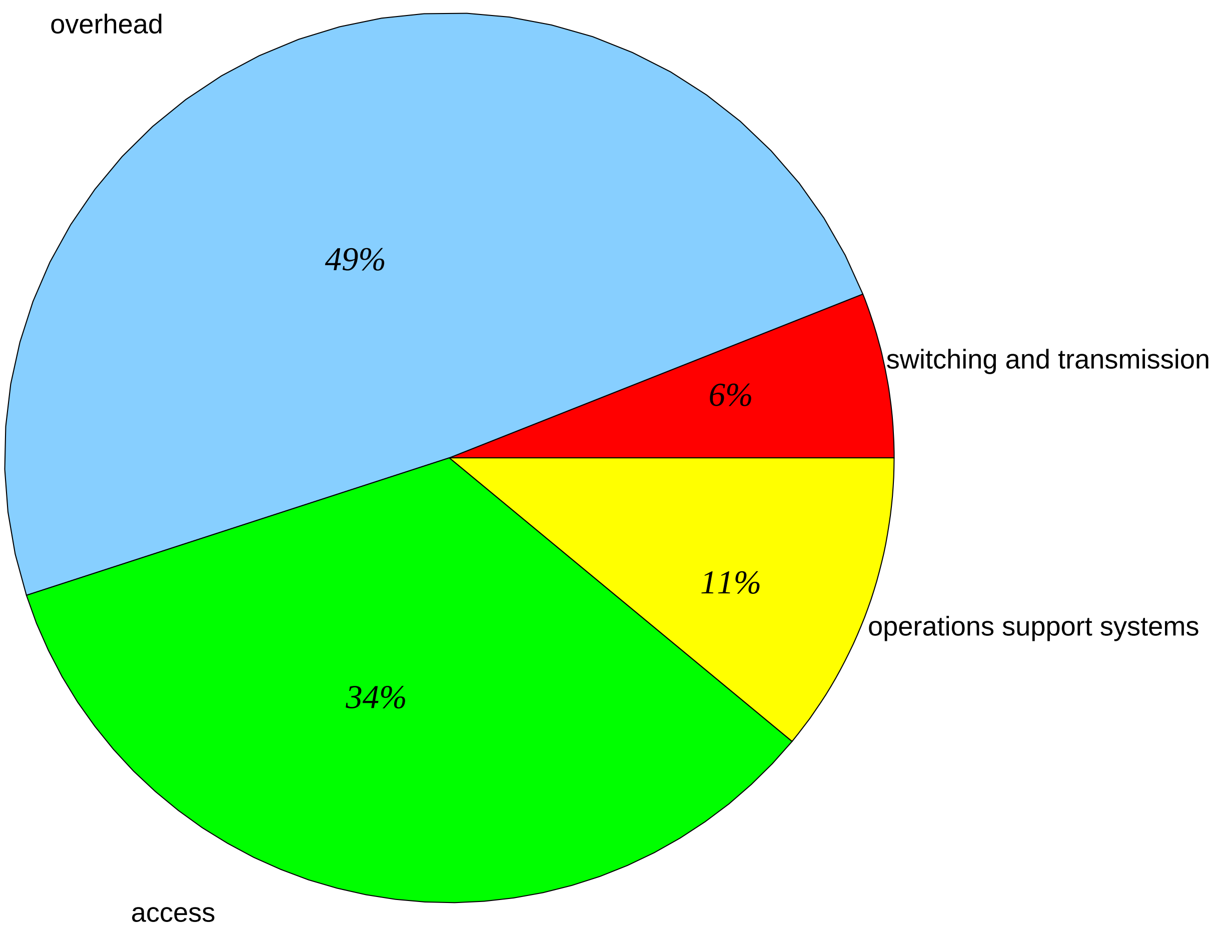

Figure 1.13: Estimated cost components for providing telephone service

Device	port speed	port cost	cost/64 kb/s
8-port Ethernet hub	10/100 Mb/s	8	0.008
24-port Ethernet switch	10 Mb/s	55	0.35
8-port Ethernet switch	100 Mb/s fiber	474	0.30
8-port Ethernet switch	1 Gb/s	1187	0.08
24-port Ethernet switch with Gigabit uplink	10/100 Mb/s	141	0.09
12-port Gigabit Ethernet switch	1 Gb/s	1350	0.08
100 T1 circuit switch	1.5 Mb/s	25,000	1041
5ESS local (no AIN), 5000 lines	64 kb/s	300	300
5ESS local (AIN), 20,000 lines	64 kb/s	175	175
Small PBX (few hundred lines)	64 kb/s	1,000	1,000
Large PBX ($>$ 5000 lines)	64 kb/s	500	500

Table 1.9: Cost comparison of circuit and packet switching

in the next five years. A recent campus topology, as deployed at Columbia University, is shown in Fig. 1.14. The core switches can handle up to 130 Gigabit Ethernet ports.

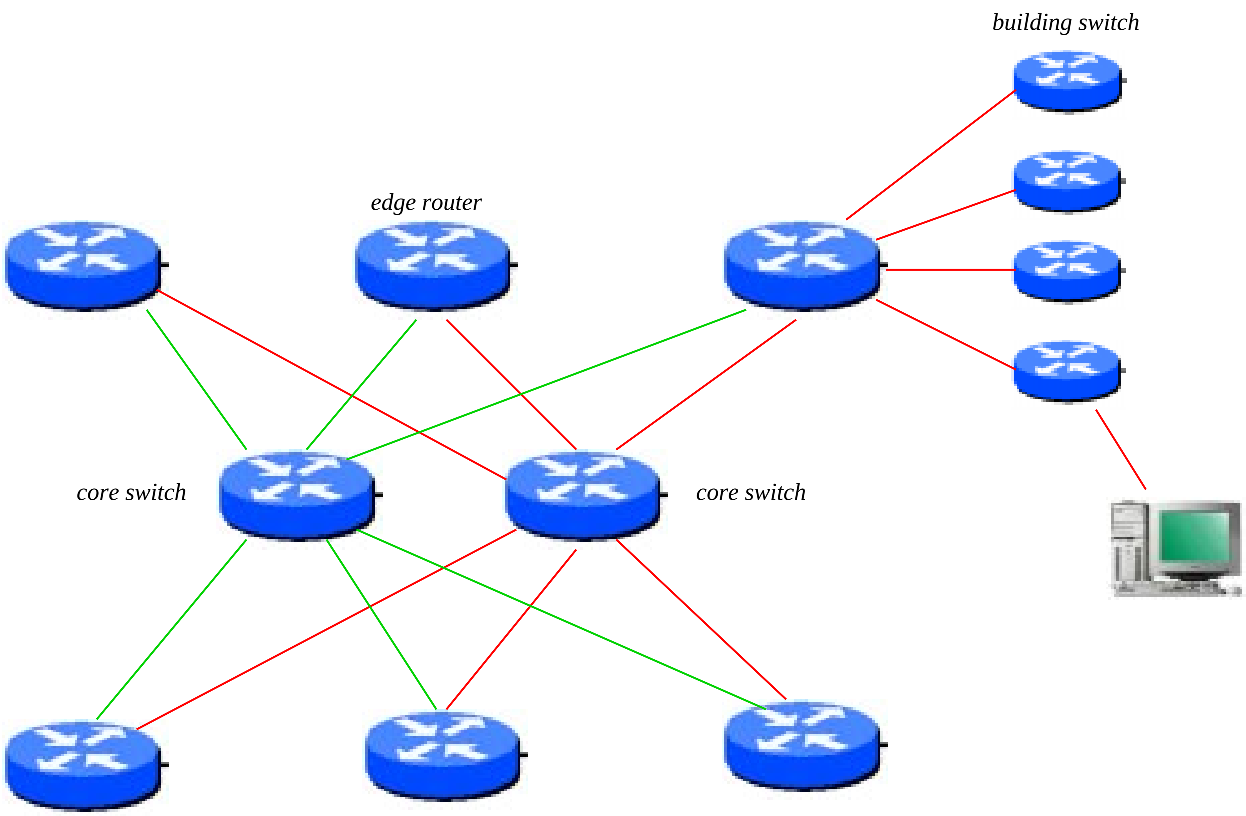

Figure 1.14: Ethernet-based redundant LAN architecture

Thus, in this design, packets would traverse four ports before leaving campus. A campus design that could replace a traditional PBX might look like Fig. 1.15. Here, Internet telephones are connected to Ethernet hubs. (In a hub, the total bandwidth for all attached systems is limited to the Ethernet bandwidth of 10 Mb/s.) Hubs are connected to Ethernet switches via standard unshielded twisted pair (UTP) cabling. Each Ethernet switch multiplexes up to about 100 ports into a single 100 Mb/s fiber uplink to another Ethernet switch, which in turn feeds into a Gigabit Ethernet switch. The sample design supports a total of 49,152 Internet telephones. The design is non-blocking, in that the total bandwidth from all sources is always less than the uplink bandwidth. Since it is not feasible to have a single broadcast-domain for thousands of end systems, the switches higher up in the hierarchy act as "layer 3 switches", i.e., simple IP routers, with each port as its own IP network. In mid-1999, such as design would cost approximately $22 per port.

Note that this design would have to be modified to serve as a next-generation access network in a suburban environment, since the transmission distances of twisted pairs and multimode fiber may not sufficient. This simple design also ignores the issue of power protection in a distributed environment. Telephone guidelines call for continuous operation through eight hours of power outage. (In the U.S. and Western Europe, there are about 15 power outages per year lasting a total of about 100 minutes, where 90% last less than 5 minutes and 99% less than one hour[56].) Figure 1.15 indicates estimates for per-port power consumption. A closet serving 768 lines would consume a maximum of 1200 W and thus would need 9600 Wh of backup capacity. This would take a stack of valve-regulated lead acid batteries weighing 350 kg. Clearly, lower-power designs for hubs and switches will be desirable.

An additional problem is that right now, the aluminum member of the coaxial cable can carry only about 4 W of power, while primary-line telephony and a cable modem would draw 8 to 12 W.

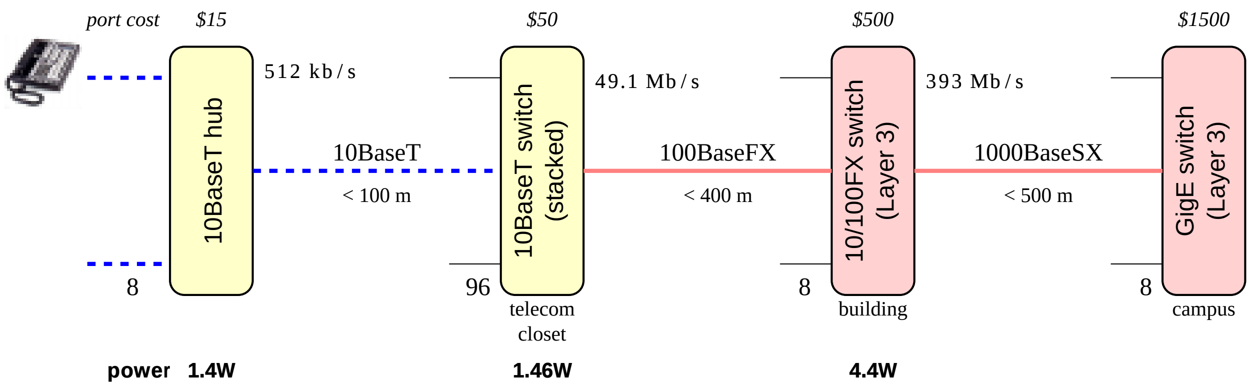

Figure 1.15: Hypothetical Ethernet-based distributed PBX

1.10 Internet and Telephony Standards

A number of organizations write standards and recommendations for telephone service, telecommunications and the Internet.

Standards organizations used to be divided into "official" and "industry" standards organizations, where the former were established by international treaty or law, while the latter were voluntary organizations founded by companies or individuals. Examples of such treaty-based organizations include the International Telecommunications Union (ITU, www.itu.int), that in 1993 replaced the former International Telephone and Telegraph Consultative Committee (CCITT). The CCITT's origins are over 100 years old. National organizations include the American National Standards Institute (www.ansi.org).

When telecommunications were largely a government monopoly, the ITU was roughly the "parliament of monopoly telecommunications carriers", with a rough one-country, one-vote rule. (Rumours have it that some small island nations could be convinced of the technical merits of a big-country proposal with indications of discounts on equipment or service...) Now, membership appears in the ITU to be open to just about any manufacturer or research organization willing to pay its dues. Thus, today there is no substantial practical difference between these different major standardization organizations. Standards are not laws or government regulations and obtain their force if customers require that vendors deliver products based on standards.

Volunteer organizations with broad technical scope include the Internet Engineering Task Force (www.ietf.org), the International Multimedia Teleconferencing Consortium (IMTC) (www.imtc.org), the IEEE (www.ieee.org) and the Telecommunications Industry Association (TIA) (www.tiaonline.org). We will discuss the IETF in detail in Section 1.10.1. The IMTC has a membership of 150 companies and aims to "promote, encourage, and facilitate the development and implementation of interoperable multimedia teleconferencing solutions based on open international standards". Its work is primarily based on ITU standards.

The IEEE is a professional organization for engineers and researchers in the area of electrical and computer engineering and related disciplines. Its standardization activities in communications focus primarily on Ethernet, in the 802.x series of standards.

The TIA is a United States national trade organization membership of 900 large and small companies that "provide communications and information technology products, materials, systems, distribution services and professional services in the United States and around the world." Its origin are in the telephone world.

More specialized industry consortia relevant to Internet telephony include the ATM Forum (www.atmforum.org) for mostly data-oriented ATM services, the Frame Relay Forum (www.frforum.com/, the Soft Switch Consortium (www.softswitch.org) or the World Wide Web Consortium (www.w3c.org), primarily for recommendations related to web-specific technologies such as HTML and XML.

For ITU standards, national bodies such as ANSI or carrier-organizations such as Bellcore often provide "cleanup" work, specifying a *profile* that make one choice among the many offered by the standard or adopt them to the installed base, government regulations or preferences in a particular country. For example, the International Multimedia Teleconferencing Consortium (IMTC) (www.imtc.org) wrote a profile of H.323 for Internet telephony that mandated the support of certain audio codecs. Profiles may also simply list a set of standards necessary to build a interoperable device. An example is the Telecommunications Industry Association (TIA, www.tiaonline.org) that is compiling a recommendation for a business Internet telephone, citing a host of related standards from Ethernet to call control.

The ITU also relies on other, specialized standards organizations for input. For example, the ATM Forum contributed many of the technical details of ATM standards, or RTP, an IETF standard, was integrated into H.323.

All of these organizations have in common that they tend to specify interfaces or protocols, but not the application programming interfaces (APIs) that software uses to access these protocols. (There are some exceptions, e.g., the IETF has specified socket extensions for IPv6 or a generic security API.) Microsoft performs this task for 95% of the operating systems, while vendor consortia such as Xopen () serve this role for Unix-based systems.

Standardization organizations differ in their procedures. While almost all of them try to operate largely by consensus, most traditional organizations, including the ITU, ANSI and IEEE, have formal voting procedures, usually requiring a super-majority such as two-thirds approval before a document can become a standard. The IEEE, for example, requires a three-fourths majority. As explained below, the IETF, on the other hand, has no formal voting mechanisms.

Most standardization organizations are partially financed by selling their standards. However, many ITU documents, for example, are available in the form just before standardization, with a bit of ftp digging. It has been argued that charging for standards slows adoption of standards, but it certainly is more of a burden for smaller or non-profit organizations.

Until the early 1990s, a largely separate world existed for telecommunications standards and for

Internet standards, with the former primarily focused on either physical and link-layer standards or telephone services. However, these two worlds are now intersecting, with some amount of culture clash and jockeying for position. However, this intersection is not altogether new – and previous overlaps were generally not without pain, for the participants, and confusion, for those building and buying networks. Examples include the dual standards for email, X.400 and Internet email (SMTP for email transport and related mail format standards), which persisted well into the late 1990's until the Internet standards came to dominate. For many years, interoperability standards had to be worked out and gateways deployed to allow X.400 users to send email to Internet-standards-based users and vice versa. A similar split occurred in the area of network management, between SNMP on the Internet side and CMIP on the telecommunications side. CMIP is still used in some telephone switches and cross-connects and plays a role in the implementation of local number portability (Section **??**). A third example occurred in the specification of the next-generation Internet protocol that can cure the Internet address exhaustion problem afflicting IPv4. In 1992, it was proposed [57] that CLNP, the ISO connectionless transport layer, should become the next generation Internet network-layer protocol, with UDP and TCP rather than ISO's TPx as the transport layer. CLNP allows variable-length network addresses much longer than the four byte IPv4 addresses. Only after a fair amount of discussion (a euphemism in this context) was a new Internet protocol designed from scratch and designated as IPv6. In the context of Internet telephony, the evolution of H.323 (Chapter **??**) and SIP (Chapter **??**) is a similar example, with the outcome still uncertain.

1.10.1 The Internet Engineering Task Force (IETF)

The IETF serves as a standardization body for Internet-related protocols. This generally comprises the network layer and above, although the IETF has also standardized a link layer, the Point-to-Point Protocol (PPP), and how to carry IP packets on a variety of link layers, from Ethernet [58, 59], to ATM [60, 61] to carrier pigeons [62, 63]. The IETF does not generally APIs or performance requirements, but does provide guidance on operational issues, such as security requirements.

The IETF differs from the standardization organizations discussed earlier in two important aspects: open membership and free publication. The other organizations mentioned earlier all require organizations to be members, generally at $10,000 and above per year. Anybody that has paid the IETF meeting fee can walk into an IETF meeting and is registered as an individual, not the representative of a company. (Not that it keeps people from walking in without the badge...) Meeting fees support the IETF secretariat that organizes meetings and processes Internet drafts (see below).

The IETF meets three times a year, with roughly every third meeting in a non-U.S. location. Attendance has grown from a few dozen to roughly 2,000 (Fig. 1.16), with most attendees still drawn from the United States, even if the meeting is at a European location (Fig. 1.17). Individual working groups may decide to hold interim design meetings between the full IETF meetings.

The IETF is organized into working groups, which in turn are grouped into areas: applications, general, Internet, operations and management, routing, security, transport, and user services. Reflecting the relative maturity of different parts of the Internet, the applications area currently has the most working groups, 26, while the user area that writes introductory documents has dwindled

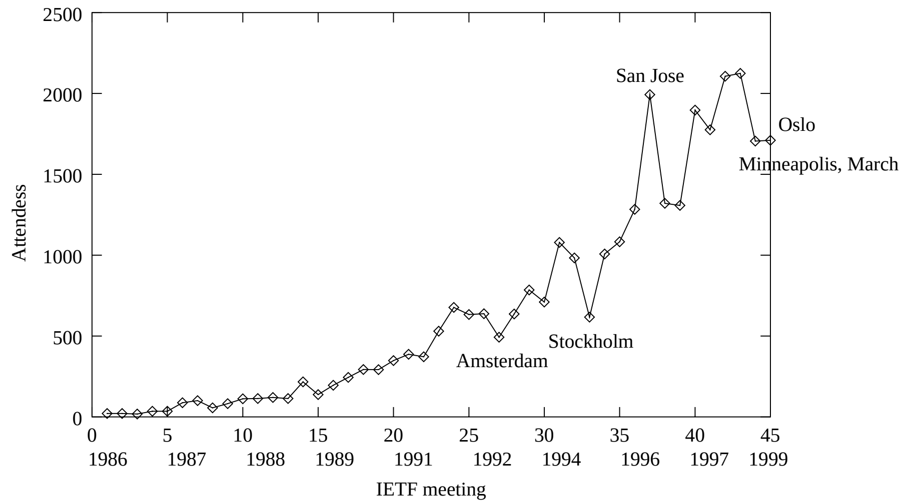

Figure 1.16: IETF meeting attendance

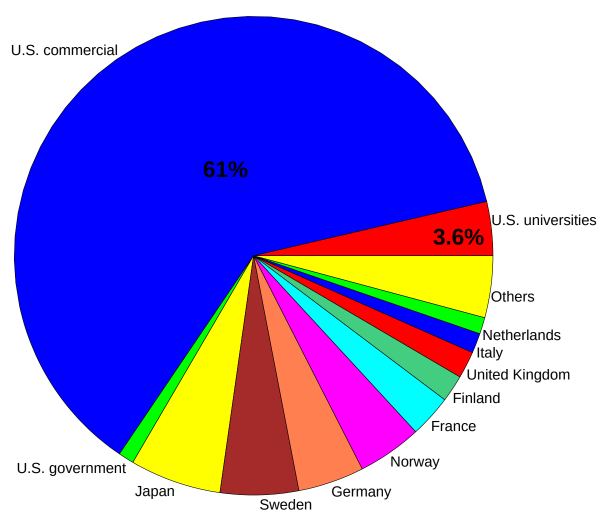

Figure 1.17: Attendees by country at the 45th IETF Meeting in Norway, July 1999

to 4. Currently, the working groups related to Internet telephony are all found in the transport area, even though protocols like SIP are at the application layer and the transport-layer status of RTP is subject to debate. The following working groups are active in the area of Internet telephony:

XXX JONATHAN XXX

Each area is headed by an area director or a team of area directors, which then form the IESG. The IESG is involved throughout the life cycle of a specification, from approving the working group to making the final determination as to when to declare the output of a working group an Internet standard.

The IETF publishes two series of documents, RFCs and Internet drafts. Both are available free of charge.

RFCs ("request for comments", although this derivation has become largely meaningless) are the archival document series of the IETF, dating back to April 1969. All except some early RFCs are available on-line, without charge, as plain text. RFCs can be either standards, in one of the three standardization stages described later, a best-current-practice document (BCP), experimental, informational documents or historical. Most, but not all, Internet standards and experimental RFCs are the output of working groups, while anybody is free to submit a technical document as a candidate for an informational RFC. Generally, unless the document has glaring technical faults or is competing with IETF standardization activity, the RFC editor will publish informational RFCs after some informal review. Informational RFCs allow common Internet protocols originating outside the IETF to be documented in an easily accessible forum.

RFCs are never modified once published. This differs, for example, from the practice of the ITU, where the protocol is identified by the standards label, e.g., X.400 as a document in the X-series, with "vintage" designations such as X.400-1992 to describe the version. As for wine, later vintages are not necessarily improvements.

Internet drafts are temporary working documents that expire and disappear from the on-line archives after six months, unless they are either in "last call" (see below) or about to be published as an RFC. Typically, an Internet draft goes through several versions before it gets published as an RFC, although many Internet Drafts are abandoned or merged into other documents without ever reaching that stage. The Internet draft name identifies whether the draft is an "official" work item of a working group, as in `draft-ietf-avt-rtp-new-04.txt`, or a submission by an individual author, as in `draft-antti-gsm-sms-url-04.txt`.

The IETF standards process is described in RFC 2026 [64]. The process distinguishes three stages of Internet standardization: proposed, draft and Internet standard. Formally, a specification can become a draft standards if "at least two independent and interoperable implementations from different code bases have been developed, and for which sufficient successful operational experience has been obtained." In practice, most proposed standards now already have two or more implementations, even if not complete. If done rigorously, features found in the Proposed Standard that do not have two independent implementations need to be removed when the specification is elevated to Draft Standard. Particularly for more complicated standards, this feature removal is not always performed rigorously. Many RFCs have languished as Proposed Standards for many years, even though they are widely implemented. For example, the Network News Protocol (nntp) has been a proposed standard since 1986.

The highest level of maturity is the designation "Internet standard". "A specification for which significant implementation and successful operational experience has been obtained may be elevated to the Internet Standard level.. . . A specification that reaches the status of Standard is assigned a number in the STD series while retaining its RFC number." [64] The same RFC can change status from proposed to draft or Internet Standard, but it is far more common that the protocol specification is revised as it gets promoted, usually to fix various mistakes and add clarifications based on implementation experience.

Roughly, the process from idea to standards-track RFC proceeds as follows:

1. The author(s) need to decide whether the new protocol or problem fits into an existing working group. Standardization does not require working group review and some drafts become standards without having gone through a working group, but this is rare for all but small additions to existing standards. If it fits into an existing working group, the next three steps are skipped.

2. If there is no working group, a birds-of-a-feather (BOF) session may be scheduled at the next IETF meeting. Often, a mailing list is set up to discuss the scope of the problem before hand.

3. If a sufficient number of people express interest at the BOF, the IESG may charter a new working group and appoint one or more chairs to lead it.

4. The working group has a charter, possibly revised periodically, and a timeline, usually covering about a year or two, although some working groups, such as the Audio-Video Transport (AVT) or IP security (ipsec) working groups, have been in existence since 1992[15]. All working groups have open mailing lists and mailing list archives.

5. Usually a single author or a small design team writes an Internet Draft, in ASCII and possibly PostScript.

6. The author should then ask chairs of relevant working group whether this is to be an official work item of the WG. If yes, it gets the name `draft-ietf-wg-something-00`, if not, the author can submit it as personal contribution and it will be named `draft-gore-internet-00`, assuming your name is Gore. Also, it is an individual submission to a working group, say, ipv6, it may be labeled `draft-gore-ipv6-internet-00`. If the former, it will appear on the WG web page, if the latter on the list of individual contributions. In either case, it will be announced on the IETF mailing list.

   Draft are submitted to the Internet drafts repository by sending it to internet-drafts@ietf.org, with copy to the chairs of the working group, if applicable. It will appear in the Internet drafts directory within a few days – and disappear automatically after six months unless refreshed with a new version.

[15]The oldest working group still active is rumored to be PPPEXT, cranking out PPP extensions since 1989, with no end in sight.

7. Authors may then ask (or be asked) to present the important items of the draft at the next IETF meeting. The presentation should concentrate on the open issues rather than being a tutorial on the protocol – attendees are expected to have read the draft. Draft need to be submitted well ahead of the IETF meeting, as there is a cut-off date several weeks ahead of the meeting week. During the weeks before the IETF meeting, several hundred Internet Drafts may be submitted.

8. Typically, a draft will go through a number of revisions, with names like draft-ietf-wg-something-01, based on comments from the working group or early implementation experience. A dozen or more iterations do occur for larger documents. For example, SIP went through twelve Internet drafts before becoming a proposed standard.

   Once the technical content is considered stable and the presentation in reasonable editorial shape, the authors ask the chair(s) of the WG for a "last call" period of about a month within the WG. The "working group last call" is meant to focus the attention of the working group. The reaction to the last call is also an indication of the roughness of the consensus that the specification should move forward. Working group members are expected to speak up or forever hold their peace.

9. Next, the working group chair submits it to the IESG for an IETF "last call" of approximately one month.

10. The IESG decides after the last call at one of their monthly meetings whether the last call and their internal reviews indicates that the document is ready for publication;

11. If it is, it is declared a Proposed Standard and announced to the IETF mailing list, with some comments about its technical history or the working group process.

12. The authors then work with the RFC editor to produce the RFC. This can take between one month and several months. The RFC editor maintains a status summary on a web page that provides rough guidance as to where the RFC is in its production cycle. Authors will not be given an RFC number until the process is complete.

13. Just before the RFC is published, the authors will be given 48 hours in which to correct any typographical errors or formatting problems. After that, the RFC will be announced to the IETF mailing list and will be available through the various on-line RFC archives.

The open membership of the IETF makes voting as practiced by other standardization organizations difficult, as ballot-stuffing would be hard to prevent. However, there is informal consensus process where working group chairs gauge and declare "rough" consensus. However, various approximations such as "straw polls", show-of-hands or "humming" are used in larger working groups to measure the strength of support for an idea or a document. Often, depending on the circumstances, a vocal minority can either block progress or a not-so-vocal minority can be steam-rolled, without establishing their true size. Voting also has a more philosophical problem in deciding technology preferences, since not all potential voters are equally qualified or equally affected by a vote [65].

Due to the looser organization, the IETF standardization process can vary significantly in duration. While every working group has a schedule laying out a timeline for publishing documents, it appears that these are very rarely met, except for the simplest documents. Progress can be hampered by document editors that loose interest or get drafted into other tasks, working groups distracted by too many documents proceeding in parallel or by discussions within the IESG and last-minute change requests from the IESG. From the time a document is "promoted", it may take another month or two until the RFC is published.

One of the more important decisions in determining the success of a standard is timing. Generally, standardization works only if the basic technology and the design choices are reasonably well understood. For example, there should be a number of research prototypes that demonstrate the feasibility of functionality and algorithms. However, if there are too many deployed products already, vendors have a natural interest in shaping the final outcome to be as close as possible to their existing implementation, while others may have an interest in *not* giving the first-to-market too much advantage. Also, as the working group size increases, too many perspectives and inputs may encourage feature creep or endless mailing list debates. Some IETF working groups have mailing lists with a thousand members and meetings with several hundred attendees.

In the IETF, standards are generally written by a small team of one to maybe six authors. Indeed, unlike in the IEEE or ITU, authors are given credit (or blame) by name on "masthead" of the RFC.

1.11 Additional Resources

Bellamy [4] provides a good overview of access, switching and transmission in the modern telephone system. A detailed description of the American telephone system as of 1984 can be found in the book "Engineering and Operations in the Bell System" [3]. (It appears to be out of print, unfortunately.) The most detailed introduction to the local network is contained in Telcordia's (formerly Bellcore's) "Notes on the Network" [8]. van Bosse [66] describes telephony signaling, while Kessler and Southwick [67] focus on ISDN signaling and transport and Faynberg *et al.* [68] offer a more high-level view of Intelligent Networks.

The Internatioal Telecommunications Union (ITU) (www.itu.int) publishes an annual survey titled *ITU Direction of Traffic: Trends in international telephone tariffs* that summarizes volume and growth of telephone traffic. The ITU web site contains basic indicators such as population, gross domestic product (GDP), main telephone lines and main lines per 100 people.

The Common Carrier Bureau in the U.S. Federal Communications Commission (www.fcc.gov) publishes an annual summary of telephone-related statistics called "Trends in Telephone Service", as well as data on network reliability. The North American Numbering Plan Administrator web site (www.nanpa.com) describes the assignment of area codes, while the Number Portability Administration Center (www.npac.org) has information on local number portability. Reports related to emergency call services are found at the National Emergency Number Association (www.nena9-1-1.org).

Cornell University has a searchable copy of the United States Code at www4.law.cornell.edu/uscode/, including Title 47 for radio and wireline communications and Title 18, Chapter 121 for lawful intercept of telephone communications. Links and articles related to computer and communications privacy can be found at http://www.hotwired.com/Lib/Privacy/.

The Network Reliability Steering Committee in the Alliance for Telecommunications Industry Solutions (www.atis.org) summarizes and analyses reliability data.

A number of journals and magazines cover telephony-related developments, including the *IEEE Communications Magazine* from the more technical angle, while trade magazines such as *America's Network* and *Telephony* focus on the business side.

1. Polling your international friends and colleagues, name some of the technical differences between your local phone service and that of other countries. Areas to consider: phone numbering, dial tone, busy tone, phone jacks, ...

2. Argue why local service should be priced at a flat monthly rate, independent of usage. Argue why it should be metered by minute. What might be a good argument for metering by call, regardless of duration (as in New York City)?

3. Gather call durations from mobile phone bills. Does it differ significantly from the statistics given in this chapter? (To get a larger sample, aggregate information across your class.)

4. List the telephone services provided by your local phone company. Which services work within the local exchange, within the country, everywhere? For example, does caller id work universally? Can you transfer a call to any number or just a local number?

5. Find out where the central office is closest to your home. This may be more difficult than it seems.

6. Does your phone company know when your home town first got telephone service?

7. Plot the usage of area codes in the United States. Predict when the last area code will be used up.

8. Electrical engineering question: compute the resistance of 18 kft of 24 gauge telephone wire.

9. Using a PC's audio recording functionality, measure the round-trip delay of a long distance call. Hint: it may be easier if both sides have two phone lines.

10. Can you model the cost function (Table **??**) by a polynomial? If so, when do you predict will the real cost of a minute of long distance drop below 1 cent?

Chapter 2

Internet Telephony

Utility is when you have one telephone, luxury is when you have two, and paradise is when you have none.
Doug Larson

2.1 Introduction

Imagine a phone call in the year 2010. The caller, Alice, picks up the receiver (a few things are hard to improve upon…) and speaks the name of the person she wants to call, Bob. She doesn't know where Bob is located, whether at home or using a mobile phone. Bob happens to be at home, with a small screen phone built into the refrigerator. (It also serves as the kitchen web browser and Internet radio.) Alice's picture appears on the screen and her name is announced to Bob. Bob decides to talk to her and picks up the handset. During the call, he wants to show her the new Internet toaster and picks up a small wireless video camera that joins the call.

Next, Alice wants to call her friend Carol. However, Carol is in a meeting and has instructed her phone server that personal calls are to be forwarded to her voicemail. Alice sees a web page that indicates Carol's schedule for the day, as Alice's web server recognizes her as a friend of Carol's. Alice follows the link on the web page that allows her to leave voice mail for Carol. When Carol gets home, she checks her private email and finds a voice message from Alice. She clicks on the link in the message to return Alice's call.

Next, the telemarketer Trudy tries to reach Bob. Bob subscribes to the Can-Spam service that screens all calls. Unfortunately (for Trudy), Trudy's company is already listed with Can-Spam, so that the call is forwarded to a web page indicating that Bob prefers not to receive solicitations for time-share condos.

This is one vision of *Internet telephony*. Internet telephony can be defined as the provision of telephony services using the Internet or Internet protocols, either end-to-end or in conjunction with circuit-switched telephony.

Internet telephony is also called *voice-over-IP* (VoIP) or ”IP telephony” (IPtel). *Fax over IP* (FOIP) describes the subset of Internet telephony concerned with carrying Fax transmissions over the Internet, either in real-time or a store-and-forward, email-based mode. (We will touch on fax-over-IP only occasionally.)

Since the beginning of digital communications in the 1960s, research in communications has had one over-arching theme: the integration of all forms of communications in a single infrastructure and service. Integration has been attempted several times in the history of communications, but the Internet offers the first realistic hope of unifying all electronic communications services, including radio, television and telephony, in a single network and single end system.

Services that we perceive as separate, such as email, radio and television, may blend into a continuum of services, characterized by four parameters: media, reach, synchronicity and symmetry. Media describes the type of data transmitted, for example audio or text. Reach describes the number of participants, from two to millions. The synchronicity parameters describes the tolerable delay between sender and receiver, with asynchronous communications such as email and real-time conversation forming the two extremes. Communications can be either symmetric, with all participants sending and receiving data, or asymmetric to various degrees.

As a replacement and enhancement for traditional telephone services, Internet telephony has to support audio and has a small reach, is useful primarily for synchronous communications (with adjunct services such as voicemail for asynchronous services) and mostly symmetric. However, as we will see later, the technologies used to support this basic service can also cover ranges of communications parameters beyond that.

2.2 Internet Telephony Architecture and Overview

Internet telephony builds on a dozen or more protocols and mechanisms, integrating the work of twenty years of Internet evolution. Before delving into the details of the foundational protocols, this chapter will trace how a basic Internet phone call takes place. Later chapters will then fill in the details.

2.2.1 Media Gateways

A (media) gateway connects the Internet with the telephone network. Gateways come in two flavors, depending on the nature of their telephone interfaces. A *line gateway* connects to one or more central-office or PBX lines, either analog or a digital trunk, such as a T1. The *phone gateway*, on the other hand, connects a number of traditional analog, key-system or digital PBX phones to an Internet interface. The latter system has to generate ring tone and provide line voltage to attached telephones.

Gateways can vary in size from a single interface to possibly thousands of phone lines. Gateways with a small number of lines are often implemented as plug-in ISA or PCI cards that use a regular PC for control and configuration, transferring voice data across the PC system bus to a network interface card (NIC). A *residential gateway* consists of a stand-alone box with connections for one or two analog phones, either with an Ethernet interface to connect to a cable or DSL modem or integrated into the modem.

Larger gateways can be integrated with remote access gateways, so that a single channel in a digital trunk or analog phone line can either be used to connect to a modem or to a voice-over-IP

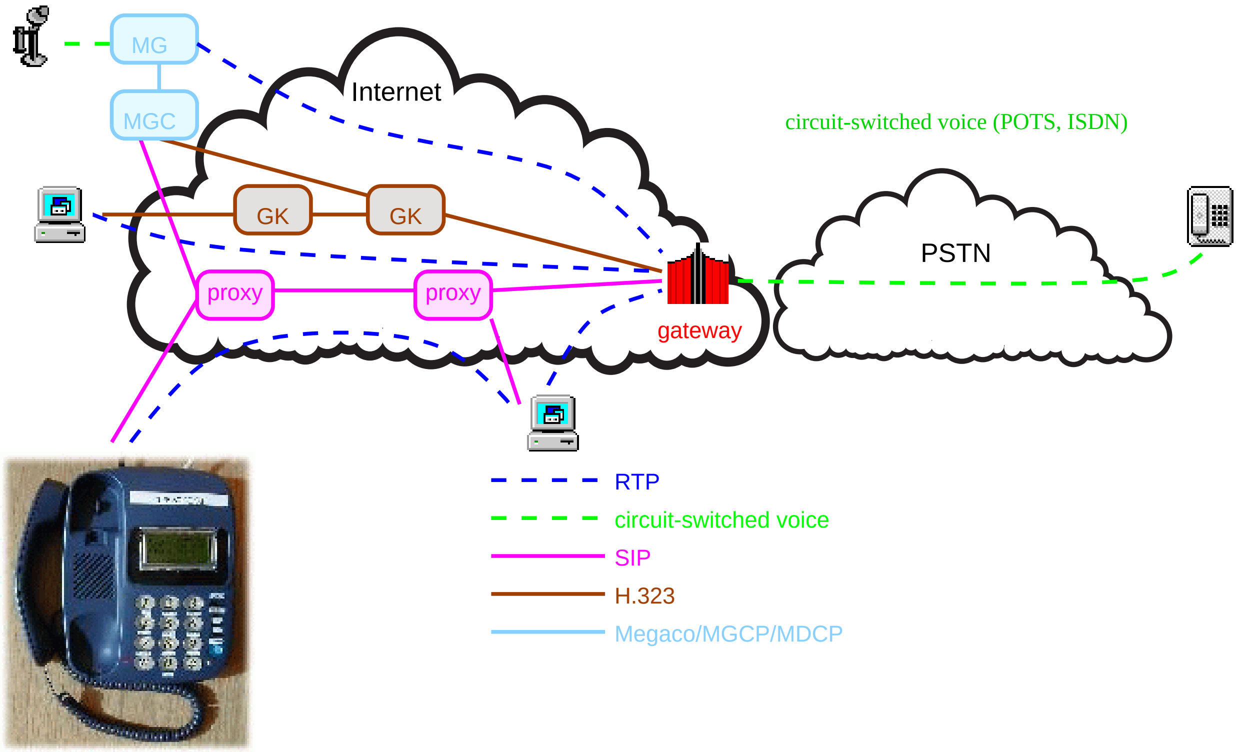

Figure 2.1: Internet telephony components

gateway port. Gateways have also been integrated into GSM base stations or as line cards in class-5 telephone switches or PBXs. (Indeed, it appears that every major modern PBX architecture now offers a voice-over-IP interface, treating it pretty much like another wide-area interface, similar to a T1 or T3 interface.)

In all of these architectures, a small number of channels or lines share a single digital signal processor (DSP), a CPU optimized for processing audio samples. The DSP encodes the audio samples received from the telephone channel into a compressed audio format and packetizes them. In the opposite direction, it decompresses the content of packets received. Often, the DSP adds the Internet protocol headers, including the transport layer. The packets are then sent across the system backplane bus to an Ethernet or other packet interface. If the gateway also serves as an access platform, the DSP can then also act as a software modem.

In such gateways, call setup ("signaling") and media are separated, with the call setup handled by a standard processor and operating system. Each DSP, linecard or a whole assembly of interfaces is labeled abstractly as a *media gateway* (MG). An Ethernet phone can also be treated as a media gateway.

A *media gateway controller* (MGC), typically a workstation, receives and generates call setup requests. This model is often called a master-slave model, with the MGC acting as the master. The MGC can be physically separated from the MGs that it controls. For example, one can imagine an architecture where a single MGC controls all the gateways for a carrier, including large multi-line gateways to the legacy phone network or residential gateways. Two master-slave protocols are currently under discussion: MGCP and Megacop. We will return to them in Chapter **??**.

2.2.2 "Black Phones"

Below, we will refer to a "traditional" or "legacy" circuit-switched phone as a *black phone*, in deference to the typical color of the Western Electric 500 set and its predecessors. We label both a single-purpose hardware device, such as an Ethernet phone, and software running on a general-purpose workstation or PC as an Internet phone. (It has become popular to call single-function devices connected to the Internet "Internet appliances", but a phone would probably resent being compared with toasters and washers.) Where the distinction is significant, we use the term *Ethernet phone*, since almost all of these stand-alone devices will connect to the Internet via an Ethernet local area network.

Among Internet phones, we refer to a *smart phone* as one which keeps state and communicates peer-to-peer, as described above, while we will label a phone dumb if it simply reacts to commands from a central controller, reflecting the master-slave model described earlier in this chapter. The details of the messages exchanged will be covered in the protocol chapters.

Function	required?	protocol
network layer	m	IPv4 or IPv6
IP control	m	ICMP
multicast subscription	o	IGMP
address translation	m	ARP
auto configuration	o	DHCP
transport	m	UDP, possibly TCP
name resolution	o	DNS
media transport	m	RTP
signaling	m	H.323, SIP, MGCP, ...
resource reservation	o	RSVP, diff-serv, ...
directory services	o	LDAP, whois
dynamic software updates	o	tftp
multicast address allocation	o	MADCAP
session announcements	o	SAP

Table 2.1: Protocols required for an Internet phone ("o": optional, "m": mandatory)

2.2.3 Internet Phone Calls Internet Phone

We discuss the simplest Internet phone call first, namely from one smart Internet phone to another. In order to call somebody, the caller needs the Internet-equivalent of a phone number. Depending on the protocols used to establish the call, the caller might be identified by his or her email address, an identifier that looks like an email address, but is used just for phone calls, a numeric IP address, a

URL similar to those used to identify web pages, a telephone number or some other identifier, such as a personal or company name. These identifiers can be divided into two classes, namely those that contain an Internet domain names and those without. For those without, for example, telephone numbers, a *directory service* needs to be contacted that translates this identifier to one which gives the Internet telephone a hint where to send requests. The Internet telephone can either perform this translation itself or delegate the task to a server. Clearly, the Internet phone needs to be configured with the address of the server. A number of generic directory services are in use in the Internet, including the ph (phonebook) service and whois, but the Light-Weight Directory Access Protocol (LDAP) is probably the most general and commonly available for looking up email addresses. There are also a few global email and telephone directories, for example, InfoSpace.

Let's assume that the Internet phone has obtained an address containing an Internet host name or address. It then needs to map this address to an Internet address, either an IPv4 or IPv6 address. This mapping will in almost all cases be performed by the Domain Name System (DNS), although some systems might have a local table that stores common mappings. (For example, it might be a good idea to store the mapping of the local computer support service so that one can call them up if the DNS server has failed or is misconfigured . . .)

In order to send the request, the phone has to translate the IP address to a link-layer address. If it is connected to an Ethernet or to a system that offers an Ethernet interface, such as a cable modem or digital subscriber line (DSL) modem, it will use the Address Resolution Protocol (ARP) to find the Ethernet address belonging to the IP address or, if the destination is outside the local network, that of the default router.

The phone then sends a call setup request using a signaling protocol, currently either SIP or H.323. The call setup request may either reach the other phone directly, or it may get forwarded by one or more signaling servers. (These are called gatekeepers in H.323 and proxies in SIP.) This allows the caller to reach somebody whose current IP address or host name they do not know. For example, consider trying to call Alice who is a subscriber of an Internet service provider (ISP) at home, but also has an Internet phone at her office. Just like Alice can read email from home, work and while traveling, as long as firewalls allow it, it would be nice if she could be reached via the same address regardless of which Internet device she is using. Thus, instead of addressing the call to a particular IP address, it is preferable that Alice gives out a "generic" address, say alice@acme.com. Calls are then first directed to the Internet telephony server at acme.com, which hopefully knows where Alice can be found at the moment. This server then sends on the call setup request to the current location, a location closer to the caller or indicates this location to the caller. For example, the call request to alice@acme.com might be sent on to alice@engineering.acme.com after the Acme server looks up Alice in the corporate database. The engineering server then might have more detailed knowledge which machine Alice happens to be logged in on that day and passes the request to that machine. (The details depend on the signaling protocol. When we discuss SIP in Chapter **??**, we describe how servers can ring several locations at once, if they are not exactly sure where Alice might be at the moment.)

The behavior of each of the intermediate servers can be either preprogrammed, e.g., a simple table lookup, or it can depend on a script or other program. In Chapter **??**, we describe several different approaches to making the behavior of servers depend on the identity of the caller, time of day or more complicated algorithms.

Once the call reaches the current location, the Internet phone alerts Alice to the incoming call. In some circumstances, however, it is preferable that the two devices about to establish a phone call first make sure that there are enough network resources on the path between them. It would be rather disruptive if Alice picks up the phone and then receives an error message saying "sorry, all packets are taken", as the modern equivalent of "sorry, all circuits are busy". Thus, some systems allow that the end systems attempt to reserve network resources, e.g., using the Resource Reservation Protocol (RSVP). The phone only rings if that reservation succeeds. Similarly, if caller or callee insists that the call be cryptographically secured against eavesdroppers, the negotiation about suitable encryption mechanisms should take place before the phone rings.

Alternatively, if Alice is unavailable, the call might be redirected to an Internet voicemail service. That voicemail service acts just like an Internet telephone. It is also possible to move the playback and recording functionality into a separate media server. This media server is not fundamentally different from the servers used to deliver audio-on-demand and Internet radio. It has been suggested to use the Real-Time Streaming Protocol (RTSP) (Chapter **??**) or a special MGCP or Megaco package for finer-grained control.

Once Alice has picked up the phone, literally or by the poor substitute of clicking on some pop-up menu button, actual audio data can flow between the two parties. The conversation may also encompass several media, such as a video stream, a shared application, e.g., to edit a document jointly, or even a video game. While the call setup request may have visited any number of server on its way to Alice's phone, the media streams should take the direct route from caller to callee, to avoid unnecessary delay and reduce network load.

Different components of the call may use different transport and higher-layer protocols. For example, audio and video are likely to use the Real-Time Transport Protocol (RTP, described in Chapter **??**) carried in UDP packets, while a shared application may distribute updates via TCP.

During the call, either side may add media to the call or drop media. The signaling messages can either traverse the same set of servers as the original call setup request or go directly, bypassing these servers.

In a simple setup, Internet phones are identified as today's phones, namely by the location of the device. The device might be labeled by its IP address or host name. However, Internet telephony also makes it possible that somebody's identity is only temporarily associated with a physical device. For example if Alice logs into a shared workstation or PC, say, venus.engineering.acme.com, that computer can then notify a server that Alice is currently reachable at that address. The computer simply *registers* Alice and associates her identifier with its network address.

Ethernet phones or public Internet phones could use a variety of identification mechanisms. For example, they might allow the device to be adopted by having a user swipe a magnetic stripe or smart card, touch it with an "i-Button"[1] or type in a personal code. A different system uses a PalmPilot PDA to identify the user. In the future, it might be sufficient to pick up the phone and speak one's name, with speaker recognition software establishing the speaker's identity. If most communications devices of the future are wireless, the mobile device may need to register as it enters a new network.

[1] An i-Button is a small (16 mm) computer chip encased in a steel button. The chip is personalized with information identifying its bearer and may even contain a small processor.

For a dumb Internet phone, some of the details change. Here, the phone has a signaling relationship only with its assigned media controller. The phone informs the media controller when the caller picks up the phone; in turn, the media controller instructs the phone to play dial tone and collect digits. After dialing, the digit string is passed up to the MGC, which then completes the call to the other side. During the call, it instructs the phone to either play tones or receive packet audio from a designated address. If the called party picks up, the MGC tells the MG which IP address to send audio packets to.

A telephone subscriber expects to be able to plug in a phone and make calls without having to spend hours configuring addresses and other parameters. Thus, there is a need for a phone to automatically learn its IP address, the address of nameservers and default routers. The Dynamic Host Configuration Protocol (DHCP) [31] can retrieve this information from a central server. In the absence of DHCP, broadcast ICMP router advertisements can be used to find the default router on the local network. IPv6 nodes can also generate their own local IP addresses and obtain a site prefix from a local router, further simplifying configuration.

At the signaling layer, the Internet phone needs to know the address of a local telephone server so that it can register with that server. Adding an option to DHCP is being discussed. However, it may not be necessary to configure the address of the local telephone server if that server is located within the same local area network and can be found by sending a multicast query to a well-known address.

For some calls, additional protocols are needed. For example, if a phone wants to initiate a multicast conference, it needs to allocate an available IP multicast address, using the MADCAP [**?**] protocol. To increase reliability, a network provider or a local network administrator may want to monitor and manage an Internet telephone using the SNMP network management protocols.

2.2.4 Calling a Black Phone from the Internet

For the next few years, most telephone calls will still terminate on a traditional phone. Clearly, Internet telephones need to be able to call these 600 million phones. As described above, the Internet connects to the public switched telephone network (PSTN) via Internet telephony gateways. From the perspective of the Internet phone, connecting to a gateway is no different than connecting directly to another Internet phone. Indeed, it may not even be aware that there is a black telephone at the other end, since this telephone may be addressed by an Internet telephony identifier, with the gateway acting as a stand-in for that address.

However, while the location of the Internet phone is given, there are many possible locations for a gateway. The only condition is that the gateway can dial the PSTN number. In many cases, one would want to locate the gateway close to the PSTN destination, particularly if Internet-based long-distance is cheaper than long-distance or international PSTN calls. On the other hand, if the Internet connection to a particular destination is congested, it may be necessary to find a gateway closer to the caller. In addition, capacity constraints and the need for a pre-existing business relationship may further constrain the selection of gateways. We discuss the gateway selection problem in detail in Chapter **??**. Gateway selection can be divided into two parts, namely the distribution of information about gateways and the ability of end systems to access a local snapshot of this information. It appears likely that telephony servers will participate in the distribution of gateway information. The servers will then route calls accordingly. However, the “routers” distributing gateway reachability information may also be separate from the telephony servers, with a protocol like the Service Location Protocol (SLP) [69], LDAP or OSP retrieving the best gateway for a given destination.

For example, in SIP, a call for the phone number (201) 555 0100 would be directed to the URL tel:+1-201-555-0100. The Internet telephone forwards this call to a designated server, which in turn queries the gateway selection protocol for a list of appropriate gateways. It then redirects or forwards the call to one of the gateways, by rewriting the URL as sip:12015550100@net2phone.net, for example.

2.2.5 A Black Phone Calls an Internet Phone

There are two options when calling an Internet phone: two-stage dialing or assigning telephone numbers to Internet phones. Calls from black phone to black phone, with the Internet as a wide-area transport mechanism, work in the same way, except that they add the gateway location part described in the previous section. *Two-stage dialing* works similar to how calling card calls are made. The caller dials a fixed number, probably a “freephone” (800) number or a local access number. An interactive voice response system prompts for the destination identity and possibly a calling card number. The destination identity could be either a regular phone number, for the Internet-in-the-middle case, or some number assigned to the Internet phone subscriber. This number could be the callee’s regular phone number or a number made up by the service provider. One can imagine dialing an IP address (e.g., 128059019141 for the host with the address 128.59.19.141) or using the letters on the dial to spell the user’s host name. The latter is likely to be very tedious, but auto-completion from a list of address book entries may make it feasible.

Two-stage dialing has the disadvantage that calls may reach the destination in a very roundabout way, unless there is some caller-based translation of the 800-number, for example. It is obviously also somewhat inconvenient.

As an alternative, we can assign each Internet phone a regular phone number or an 800-number, but one not used by a black phone. When looking up the number in the normal fashion, the telephone switch finds the destination switch, which happens to be an Internet telephony gateway. The gateway then uses an Internet-based lookup mechanism to find the URL, host name, or IP address currently assigned to the telephone number. A DNS-based mechanism is currently being developed, as described in Chapter **??**. It is also possible to store this mapping in a database and access it using, for example, LDAP. It has even been suggested [70] to carry IP addresses in telephone signaling messages, so that traditional telephone databases can be used.

2.3 A comparison of peer-to-peer and master-slave architectures

As shown in the call example above, there are fundamentally two ways to set up phone calls, namely either peer-to-peer or master-slave. As we will see, however, the two approaches are likely to coexist in many Internet telephony systems.

Peer-to-peer protocols set up calls between equal partners, with the signaling end point maintaining full call state. In other words, the signaling end point is aware of whether the call is ringing, being answered, transferred, or terminated. In the traditional phone system, ISDN signaling for end points or ISUP signaling for network-to-network signaling are examples of peer signaling, although ISDN signaling is different in that signaling messages are not exchanged directly between the calling and the called phone, but rather between each phone and the local switch.

Master-slave signaling is also related to *stimulus signaling*. Stimulus signaling is widely used by digital PBXs, where the desktop phone simply sends a message to the PBX whenever the user presses a key or lifts the handset. The PBX then performs the corresponding function, but the phone has no state for calls or knowledge of features. The rough equivalent in computing are protocols such as the X window system.

Traditional central office telephone switches are often split into a central controller and a number of remote offices, which are then controlled, operation by operation, by the central switch. If one were to draw an analogy to such traditional architectures[2], a client device implementing a master-slave protocol is the equivalent of a line card in a telephone switch which routes calls via a distributed switching fabric, namely an IP network. On the other hand, a client device implementing a peer signaling protocol is the equivalent of a PBX. A media gateway controller or call agent implementing a master-slave protocol is a “class-5” central office or PBX. A SIP proxy or H.323 gatekeeper occupies a somewhat similar niche as a “class-4” central office.

Up to a certain size, an Internet telephone network could use a single MGC, roughly equivalent to having a single large switch or PBX. However, for calls between providers, MGC have to communicate with each other, on a peer-to-peer basis. Thus, one possible architecture employs a master-slave protocol in each region, e.g., for each POP or cable head-end, and then has MGCs call each other using a peer-to-peer protocol.

The Packet Cable initiative is standardizing two architectures for voice-over-IP on cable, namely distributed call signaling (DCS), using SIP as its peer-to-peer protocol, and network-controlled signaling (NCS), using MGCP as its master-slave control protocol.

Feature implementation differs across the two architectures. Since the master-slave approach is very similar to a digital PBX, typical PBX features are readily implemented. Features such as call transfer or call parking are simplified, since the controller knows the state of all participants. On the other hand, this approach to feature implementation suffers from the same problem as PBXs, in that certain features are only available for calls within the same MGC or work differently across MGCs.

Another difference between the two approaches is the user interface. In the current version of MGCP, for example, the user can communicate with the system only through DTMF events and possibly a display. Thus, features have to be invoked in the same fashion as for current analog phone systems, by various digit combinations. For example, *06 activates call forwarding. The display is controlled by the same mechanism, called ADSI (Analog Display Services Interface), used by advanced analog phones, except that the characters are transmitted directly in the signaling messages rather than through a modem. In the peer-to-peer model, the user interface is up to the

[2]suggested by John Pickens

developer of the Ethernet phone or the software application. This makes integration with other end system services, such as address books or logging applications easier, since the phone does not have to do "screen scraping" to extract information.

If the MGC is provided by the carrier or ISP, only that provider can offer standard telephony services such as caller id, call forwarding or call blocking, just like in the existing telephone system. Given that Ameritech, one of the RBOCs, received more than a billion dollars from additional services such as caller-id in 1996, this is clearly of some interest, particularly as per-minute charges for phone calls may disappear.

A centrally-controlled arrangement makes it somewhat more difficult to be reachable through multiple VoIP service providers simultaneously, since each IP telephone or RGW cannot be controlled by more than one MGC. For peer-to-peer systems, a single phone may well be reachable as, say, both alice@att.com and alice@sprint.com.

Adding features to an Internet phone using peer-to-peer signaling requires updating software in the end system. For traditional PCs, it is hard to do this transparently and without user intervention, unless the phone software is a Java applet. For Ethernet phones, phones can automatically download new executables across the network, e.g., using the trivial file transfer protocol (tftp [71]), as long as the software server is reachable and not hidden by a firewall.

2.4 Hybrid PBX, IP Centrex and Other VoIP Architectures

So far, we have said little about where the Internet telephony components are located and who runs them. Even at this early stage of Internet telephony, a whole range of different architectures have been proposed, including IP PBXs, voice-over-cable, and IP Centrex. Each architecture can also be described by where the boundaries between Internet and circuit-switched communications are crossed and whether the end systems are black phones or IP phones.

An *IP PBX* replaces a more traditional circuit-switched PBX. In its purest incarnation, smart IP telephones connect to an Ethernet or other local area network. All calls are carried over IP, with calls to black telephones directed to gateways distributed across the Internet. The only central infrastructure is a signaling server that performs call routing and call filtering. This signaling server can be located either on the premises of the organization or be operated by an application service provider. An arrangement where the server is located outside the organization and operated by a third party is sometimes called *IP Centrex*, analogous to a traditional phone company offering PBX-like services from its central office. One difference between classical Centrex and IP Centrex is that the server can be located far away from the organization, since voice packets do not need to visit the server. A single server may well be responsible for a number of organizations, possibly with a local dial and naming plan for each organization. This arrangement is then somewhat similar to the virtual private networks (VPNs) that exist in some large organizations, where branch offices appear to all be part of the same PBX. In an Internet context, such a VPN may consist of a private name space, allowing shortened dialing, as well as traffic separation in routers and encryption for all IP communications within a VPN. Traffic separation ensures that each VPN receives a guaranteed share of the link bandwidth. In a related IP Centrex architecture, the server implements a media gateway controller, remotely controlling IP terminals.

In an IP Centrex architecture, the operator of the signaling server may also provide gateway services to the PSTN. For example, as a near-term architecture, one could imagine placing a gateway in each building, with VoIP only between either IP telephones or small peripheral gateways. In this near-term architecture, the service provider can avoid dealing with emergency calls. This architecture is shown in Fig. 2.3. In the figure, small enterprises have a single IP connection to their Internet service provider (ISP), which operates the gateway to the PSTN.

Since most companies already have analog or digital telephones, another architectures suggests itself, the hybrid PBX. In the hybrid PBX, users keep their analog phones and fax machines. These phones are connected to one of three devices:

Traditional PBX with external gateway: A standard PBX simply connects to a VoIP gateway with a trunk interface, e.g., a T1. The PBX remains unaffected, except that the IP "trunk" is selected based on standard least-cost routing mechanisms.

Traditional PBX with built-in IP interface: In addition to any traditional trunk interfaces, the PBX is outfitted with, typically, an Ethernet interface. The interface converts between packets and digital time-division multiplexing (TDM) calls. Most vendors of current PBXs have started to sell such interfaces.

Packet-based PBX: In such a PBX, the analog or digital line interfaces convert speech to packets, which is then switched internally by an Ethernet or other packet switch. Currently, this type of device appears to be being developed primarily for smaller PBXs with a few lines, where the packets are switched internally on a standard computer bus, such as a PCI bus. Fig. 2.2 shows such an architecture on the left hand side.

Gateways often have a fallback mechanism that switches the locally connected phones to an analog line should power fail or if the user dial an emergency number. This architecture is pictured in Fig. 2.2. Here, calls to the San Francisco corporate headquarter (415 area code) are routed through the four-line gateway. Similarly, the PBX in the figure has both a traditional T1 connection to the phone company as well as an Internet connection. Calls to the branch office in New Jersey (201 area code) are automatically routed via the intranet, while all other calls take the traditional route. This approach is an updated version of classical *toll bypass* mechanisms utilizing leased lines between branch offices.

Since some of these gateways replace traditional telephone switches, they are sometimes called *soft switches*, emphasizing their programmability. Unfortunately, there is uniform definition of this device at the moment.

For residential users, the signaling server or MGC will typically be operated by the cable company or the DSL service provider. However, one could also imagine an architecture where each household has one MGC that controls all phones.

2.5 Internet Telephony End Systems

One major disadvantage of Internet telephony is the cost of the end systems. It is hard to build packet voice "telephones" requiring no external power that operate over low-grade twisted pair

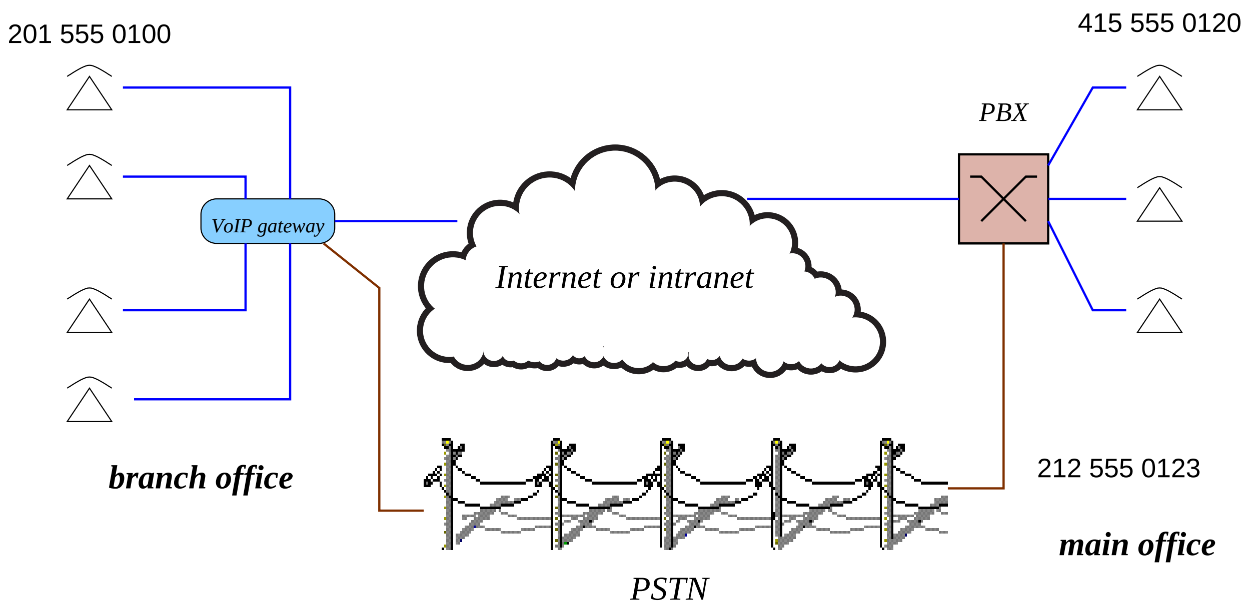

Figure 2.2: Hybrid PBX

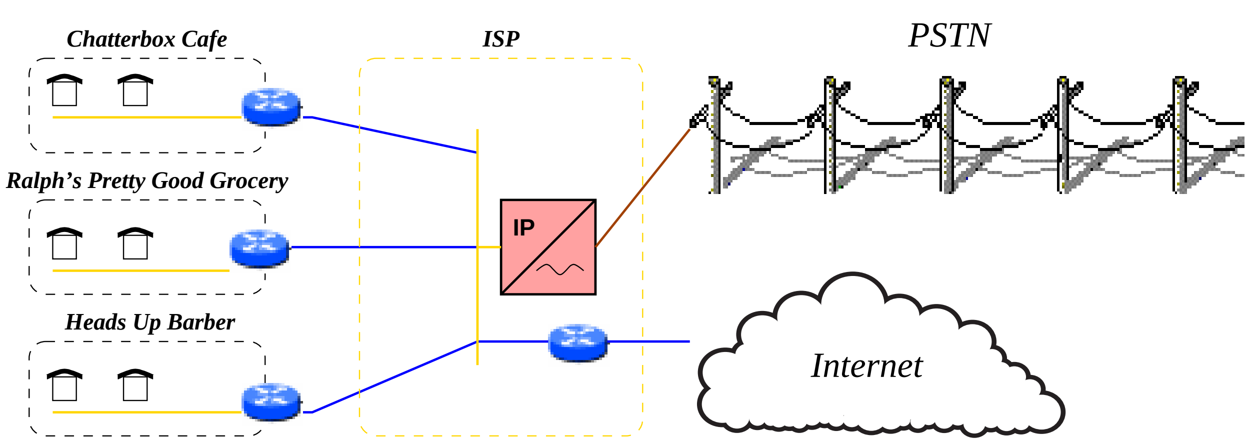

Figure 2.3: IP Centrex service

wires several miles long at the $20 to $100 price point of a basic analog phone. However, Internet phones are not restricted to being applications on personal computers. They are natural applications for network computers and "Internet appliances". Also, a voice-only "packet phones" can be built with a single DSP with on-board A/D and D/A conversion and serial or Ethernet interface (Fig. **??**). For business applications, Internet appliances with Ethernet interfaces are probably most appropriate, as existing business phone (Cat-3) wiring can carry at least 10BaseT (10 Mb/s) Ethernet. For home use, a simple RS-422 or RS-485 serial interface can operate over a distance of 4,000 feet at 100 kb/s, while the ISDN 2B1Q line coding can cover 18 kft (5.5 km) at 144 kb/s. Section **??** describes a possible residential access structure.

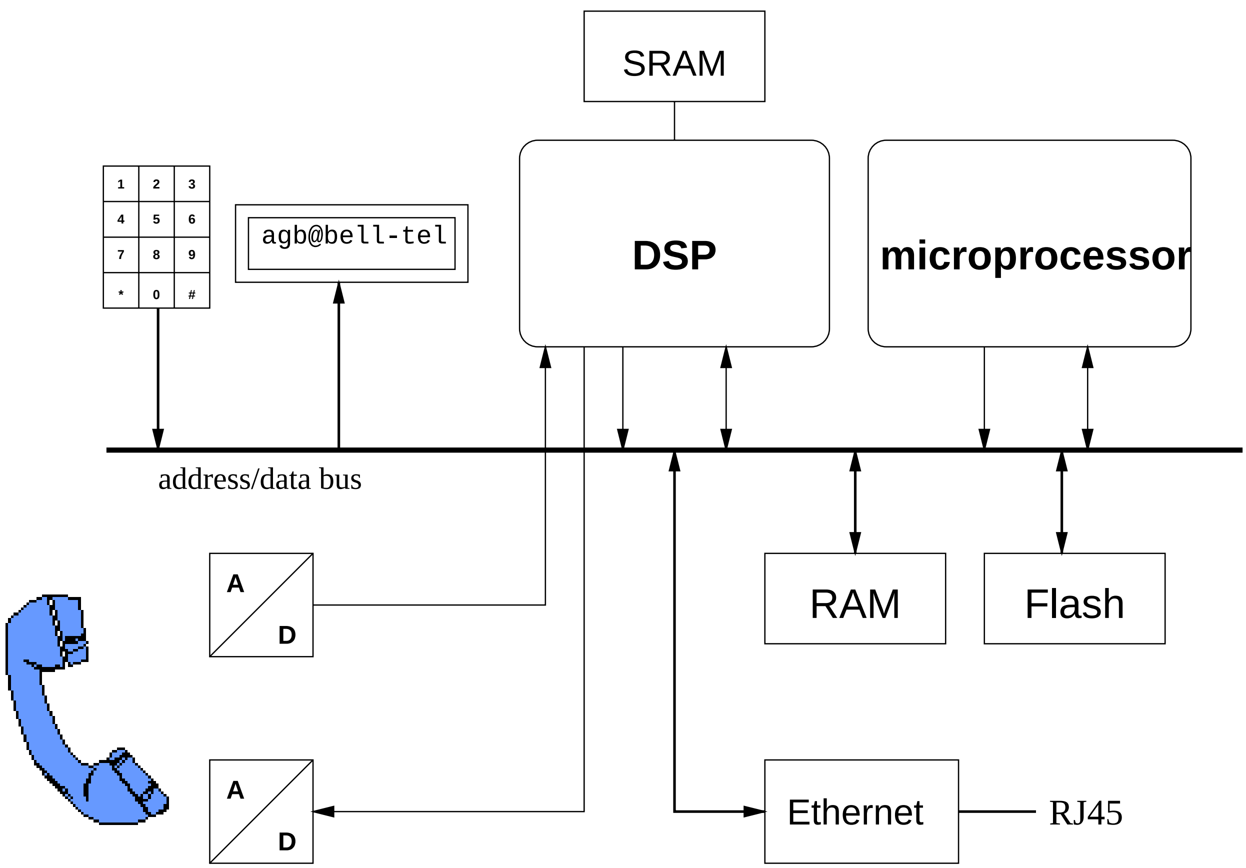

Figure 2.4: Block diagram of an Ethernet-connected Internet phone

With the description above, we can assemble a list of protocol parts that are needed to build a basic Internet phone. Table 2.1 lists protocols likely to be found in an Internet phone. A basic Internet phone can be implemented without TCP, since all required upper-layer protocols can work with UDP. DNS can be omitted if all name resolution is delegated to a local server, whose IP address is hard-wired into the phone.

If the phone is to participate in multicast conferences, it needs to be able to subscribe to multicast groups, using IGMP. To initiate a multicast conference, it needs to be able to allocate a multicast address, using MADCAP.

In addition, Internet phones may need cryptographic capabilities to protect the privacy of call setup messages and the actual communications. Also, cryptographic methods are likely the preferred way of ascertaining the identity of a caller, rather than trusting a phone company to identify the caller.

Ethernet phones may not be limited to making phone calls. For example, they could be outfitted with embedded web browsers, chat or email capability. Currently, the LCD is the major cost of such a device, however, limiting the display capabilities of low-cost devices. There is thus a continuum of devices, from a single-function phone that only makes audio phone calls, to a full-fledged Internet host supporting the whole range of Internet services.

The hardware components of a hypothetical smart Ethernet phone are shown in Fig. 2.4. In this example, the Ethernet phone has two processors, a digital signal processor (DSP) for compressing audio and echo cancellation and a general-purpose microprocessor for protocol handling and user interface. Both find their software in a flash ROM, which can be reprogrammed remotely. Typical DSPs do not have cache memory, so the example shows that it uses static RAM (SRAM), which is significantly faster than typical dynamic RAM (DRAM), but also much more expensive. The analog-to-digital (A/D) and digital-to-analog (D/A) converters translate the microphone output to digital samples and feed the receiver speaker. An Ethernet interface connects the phone to a local area network.

2.6 A Brief History of Internet Telephony

Research in transmitting voice across a packet network dates back to the early ARPAnet days. Cohen [72] refers to cross-continental packet voice experiments that took place in 1974 between USC/ISI and MIT's Lincoln Lab, compressing audio samples with a continuously-variable slope differential (CVSD) codec and carrying voice packets using the Network Voice Protocol (NVP) [73, 74, 75]. NVP was later formally specified in RFC 741 [76]. In December of 1974, additional tests were performed between Lincoln Labs and Chicago, this time using a low-bitrate linear predictive codec (LPC). In January 1976, the first Internet audio conference was demonstrated between ISI, Lincoln Labs, Chicago, SRI, coordinated via a new control protocol, the Network Voice Control Protocol (NVCP). In the same year, the first satellite-based packetized speech transmission took place between Lincoln Labs and both NTA in Norway and University College London. This agrees with J. W. Forgie's account [77], describing low-bit rate voice conferences being carried out in 1976. Some of the ideas were patented in April 1977 ("Packet Transmission of Speech", U.S. patent 4,100,377) by J. Flanagan of the British Telephone Laboratory. Other early work was done by Magill [78].

The early 1980s saw experiments of transmitting low-bitrate voice across mobile radio [79, 80, 81] and satellite [82] packet channels. A packet video standard followed in 1981 [83]. The idea of connecting a regular telephone to a packet audio system dates back to at least 1983 [84]. In 1983, Xerox PARC build the first Ethernet "phone" [85], consisting of two 8088 processors, analog-to-digital and digital-to-analog converters, an Ethernet interface and a serial port for keyboard and terminal. Control was and voice mail provided by a central server located on the same local area network [86, 87, 88, 89, 90]. Another Ethernet voice system was developed by DeTreville and

Sincoskie at Bellcore [91, 92]. The idea of carrying voice over Ethernet was first presented by Shoch, one of the co-authors of the Ethernet specification, in 1980 [93], with additional early work by a number of authors [94, 95, 96, 97, 98, 99, 100, 101, 102].

For wide-area service, the CCITT recommendation G.764 [103] for packet voice over HDLC, the CCITT link layer protocol, was published in 1990. It later also formed the basis for carrying voice in frame relay frames.

Packet audio/video should be set apart from the approach to voice/data integration that provides fixed-bandwidth circuits on multiple access networks [104, 105]. The last such attempt is probably the Iso-Ethernet (IEEE 802.9) specification developed around 1996. The 802.9 network combined a regular 10 Mb/s Ethernet and 96 ISDN channels operating at 64 kb/s each, but was never commercially successful.

System implementations of packet voice terminals are described in [77, 106, 107]. Packet radio experiments are featured in a paper by Spilling and Craighill [108]. Surveys on packet voice performance are presented in an article by Barberis et al. [106].

The early experiments were limited by the need to build specialized audio hardware and operating system extensions. Interest in packet audio increased in the early 1990s as more and more workstations came equipped with built-in toll-quality (Sun SPARCstations, DEC workstations) or CD-quality (NeXt) audio hardware support. Soon, there existed a fair number of simple programs that utilized the SPARCstation audio hardware to communicate between two workstation on a local net, for example vtalk (Miron Cuperman, OKI) or PhoneTalk (Patrik Nises and Joakim Wettby, Royal Institute of Technology, Stockholm). Programs designed for multiple-party connections across wide-area networks include VT [109] and vat (Van Jacobsen and Steve McCanne, LBL) [110, 111]. At the time, a number of commercial products already used medium-bitrate packet voice to more effectively utilize leased private lines, extending the concept of the traditional data-only multiplexer [112].

Numerous other voice/data integration schemes have been studied, usually combining a circuit-switched path for voice and a packet-switched path for data, possibly with bandwidth traded between the two. Examples include work by Miyahara and Hasegawa [113]. Economic studies comparing alternative network strategies were performed by Gitman and Frank [114]. A network protocol specifically designed for continuous media such as voice was proposed in three versions, ST [115], ST-II [116, 117] and ST-II+ [118], but it failed to gain acceptance beyond some experimental networks.

Larger-scale packet voice and video experiments were first performed within DARPA experimental IP network called DARTnet in February 1991, using first ISI's vt program and then later LBL' vat (August 1991) and UMass' nevot [119], with a primary emphasis on multicast conferences. In August 1992, the technology was made visible to a larger technical audience at the IETF meeting in San Diego [120]. The multicast overlay network set up for these tests became known as the Mbone [121, 122]. While it was meant as a temporary kludge until "native" multicast became available, it persists, with variations, until today, but with fairly limited use. (Multicast is a part of the triad of attempts to enhance network layer functionality, along with resource reservations and mobile IP, that has taken far longer than anticipated to see deployment.)

In addition to audio, in November 1992, Ron Frederick of Xerox PARC developed the first Mbone video application, nv [123]. In October 1993, a video application, vic, incorporating the

H.261 coding standard (see Chapter **??**) was released by Steve McCanne and Van Jacobson at LBL [124].

Since NVP was not really designed for modern Internet requirements and not suitable for multicast, the Audio Video Transport working group of the IETF started work in 1992 to standardize a common transport sub-layer for real-time continuous media, leading to the Real-Time Transport Protocol (RTP). It was made an IETF Proposed Standard in November 1995.

Most applications were written for Unix workstations, with the audio and video tool CUSeeMe [125, 126] written at Cornell University around 1993 for MacOS as the exception. (This tool has spawned a whole culture of its own, combining mostly black-and-white video with text chat.)

Initially, sessions were set up by emailing the multicast address and port number to all participants. Since this did not scale, a session directory, called sd, was developed at LBL. It periodically multicast brief descriptions of audio and video sessions at a well-known multicast address. The description format, somewhat modified, is used today as the Session Description Protocol described in Chapter **??**.

Largely independent of the Mbone-related efforts, a number of proprietary tools and applications for conferencing were developed as part of academic and industrial research projects, e.g., Lyceum (Sun), JVTOS and MMC (Technical University Berlin) [127, 128] as well as products, such as Intel ProShare, Sun ShowMe, or Insoft Communique. These were generally designed for small, centralized conferences on a single hardware platform.

An alternative mechanism, called Virtual Places (tm), for setting up sessions was developed at Ubique, in Rehovot, Israel, in 1994. Here, the Mosaic web browser was modified to trigger RTP sessions when two users with such a modified browser were looking at the same web page.

A second generation of "Mbone tools" was developed at University College London, for example rat for packet audio [129, 130, 131] and sdr for session announcements. However, there was little emphasis on point-to-point communications and no widely available tools for setting up on-demand phone calls. Also, all of these tools were primarily developed for Unix platforms and thus never reached much beyond the core Internet technical community. The tools also suffered from the lack of widely available high-quality audio codecs suitable for 14.4 kb/s modems, the most common modem speed at the time.

In February 1995, Vocaltec, a small start-up in Rehovot, Israel, released the first commercial Internet telephony application for Windows 3.1, initially designed to appeal to those wanting to make "free" phone calls across the Internet. However, since most residential users were still access the Internet by dial-up modem, this proved to have limited appeal. A caller would have to call up, via regular phone, or email the intended callee when they would like to converse, so that the callee could log in at the agreed-upon time. Alternatively, in a mode reminiscent of ham radio, people using the application would automatically appear in a centralized directory and thus be made visible to other users currently logged in. Microsoft's NetMeeting later adopted a similar technique, called the ILS (Internet Location Service). A number of other companies soon introduced similar PC-based Internet telephony tools, each restricted to calling those using the same tool.

These fledgling efforts apparently got the attention of the America's Carriers Telecommunications Association (ACTA), a trade association of smaller carriers and telephone service resellers. In March 1996, they filed a petition with the Federal Communications Commission (FCC) to ban the sale of Internet telephony software until it could be determined which traffic would be suitable for the Internet and how Internet telecommunications (i.e., voice) traffic could be assessed local access charges, at the time 2.5 cents per minute at the called and calling end. To the authors' knowledge, the petition has not been acted upon.

The first version of the ITU recommendation H.323, based upon the ISDN conferencing standards and originally intended as extending these to LANs, was published in early 1996. Shortly thereafter, in May 1996, Microsoft released its first conferencing application, NetMeeting, using T.120, an ITU data conferencing protocol, and H.323.

The first commercial Internet telephony service provider was Delta Three in 1996, followed by Net2phone, iBasis and Telematrix. It became clear early on that not every Internet telephony company would be able to place gateways in every country or local-dialing area. Thus, exchanges that allowed individual owners of gateways to receive compensation via a central clearinghouse, were conceived, starting around 1997. Around that time, the first trade-shows and industry conferences signaled the transition of Internet telephony from purely a lab curiosity to a commercially interesting application.

Beyond the development of RTP and work on multicast conferences, Internet telephony was not really on the radar screen of the IETF until relatively late. Initial work on protocols for creating smaller sessions was done by Schooler [132, 133, 134] and one of the authors [135], which led to two protocol proposals for session creation protocols, soon merged into the Session Initiation Protocol (SIP). SIP had been in development since 199?, and was finally "blessed" as a proposed standard in February of 1999, with interoperability events in February, August and December of that year.

2.7 Who's on First: Telephone Traffic or Internet?

In describing the transition from a world dominated by circuit-switched communications services such as phone and fax to a world where almost all communications is carried by the Internet, it is appropriate to take stock how far we have come.

The current Internet and intranets are making the transition from being a convenient additional means of communications that one can easily do without to an essential communication tool. Many in the technical and educational fields can probably function and continue to work reasonably well without an outside phone connection or PBX for a few hours, but are severely inconvenienced if the internal or external networks are unreachable.[3] Many engineers and researchers now receive far more email per day, often more than a hundred messages, than phone calls, faxes and postal mail combined. The importance is still largely limited to the technical community, however. A 1997 Future/Gallop Organization poll of 972 workers from large companies found that on average, respondents received 31.8 phone calls, 13.6 e-mails, 11.2 voice mails and 8.8 faxes per day [137].

As Table 2.2 shows, the volume of email is dwarfed by daily postal deliveries and long-distance phone calls. Also, only 32% of those who go online say they would miss services [138].

However, just as telephone and fax have started to displace international postal mail, the Inter-

[3]Compare this to the effect of a prolonged phone outage documented by Wurtzel and Turner [136].

Table 2.2: Communications volume per day (United States)

means of communications	year	millions/day
US Postal Service, first class mail	1998	630
Email (U.S.), including spam	1999	9400
Email, non-commercial	1999	2100
GSM short messages	1999	33
U.S. local phone calls	1998	1515
U.S. long distance calls	1998	274
U.S. outbound international calls	1997	11.6
ATM transactions	1999	29.3
UPS daily deliveries	1998	12.4
AOL email	1999	42
AOL instant messages	1999	309
Federal Express	1998	2

Table 2.3: Intensity of communications activities

	year	source	minutes/day
Internet	1995	[139]	57
America Online (AOL)	1997	[140]	55
long-distance calls, US	1980	[16]	8
long-distance calls, US	1997	[16]	14
local phone calls per loop, US	1980	[16]	38
local phone calls per loop, US	1997	[16]	42
Television, Sweden	1996	[141]	120
television, USA	1996	[141]	240

net could have similar displacement effects as it becomes ubiquitous. (The amount of international mail from the US has been *declining* since 1992 by about 5 to 7% annually, while international call volume rose by 17% p.a. between 1992 and 1997.)

Network	date measured	traffic	Gb/s
NSFNET	end of 1994	15 TB/month	0.046
Internet backbone	late 1994	20 TB/month	0.061
U.S. Internet backbone [142]	late 1998	5,000-8,000 TB/month	15.4–24.7
Other U.S. public data networks [142]	late 1998	1,000 TB/month	3.1
U.S. private line data [142]	late 1998	4,000-7,000 TB/month	12.3–21.6
AT&T frame relay network	late 1997	5.7 TB/day	0.53
MCI Internet backbone traffic	Nov. 1997	140 TB/week	1.8
U.S. local phone calls [16]	1997	2,683 GDEM	327
U.S. intrastate toll calls [16]	1997	404 GDEM	49
U.S. interstate toll [16]	1997	525 GDEM	64
U.S. international outbound [16]	1997	22.6 GDEM	2.8
U.S. international inbound [16]	1997	9.1 GDEM	1.1
World telephony	Nov. 1996		600

Table 2.4: Telephone and Internet traffic volumes; GDEM: giga dial equipment minutes (10^9 DEMs)

Table 2.7 compares traffic statistics for Internet and telephony services. Telephone traffic is commonly measured in dial equipment minutes (DEM), that is, minutes that the local switch is busy with a given call. A single phone call racks up two DEMs for each minute of "talk time". The world telephone traffic is based on an estimate that each of the approximately 640 million phone lines in the world are used about 20 minutes per day.[4]

The table shows that telephony traffic still dwarfs Internet traffic. However, on the transatlantic and transpacific links between the U.S. and Europe and Japan, Internet traffic already approaches voice traffic. For example, in 1997, the total installed capacity for Internet services between the U.S. and Japan was 650 Mb/s as compared to 400 Mb/s for telephony. As another example, in late 1997, the data volume between the U.S. and Sweden was about twice the volume of voice traffic. (Among other reasons, it appears to be easier to send an email to Japan than trying to communicate across the language and time zone barrier.)

Due to the lower utilization of data lines [142], capacity comparisons paint a slightly different picture. According to estimates, there is a total of 375 Gb/s of U.S. long-distance voice capacity, 150 Gb/s of Internet capacity, 80 Gb/s for other public data networks and 400 Gb/s for private data networks [143].

Also, while landline telephone usage grows by a few percent each year, Internet traffic is

[4] The 165 million United States phone lines are each used about an hour a day, a number that has only increased by about 18% since 1980. The amount of interstate long-distance calls has doubled from 4 to 8 minutes during the same time, however.

roughly doubling each year (with growth of 1,000% in 1995 and 1996), even though the number of Internet connected hosts has recently increased by "only" about 30% or so per year. In addition, a large fraction of the international voice traffic is caused by fax machines. Compared to voice, fax service can be most readily replaced by Internet services. Another estimate of the growth of total data and voice traffic is shown in Fig. 2.5 [144]. (For reasons unknown, the data traffic in the figure is significantly higher than in the figures above, underlining the difficulty of arriving at accurate estimates.)

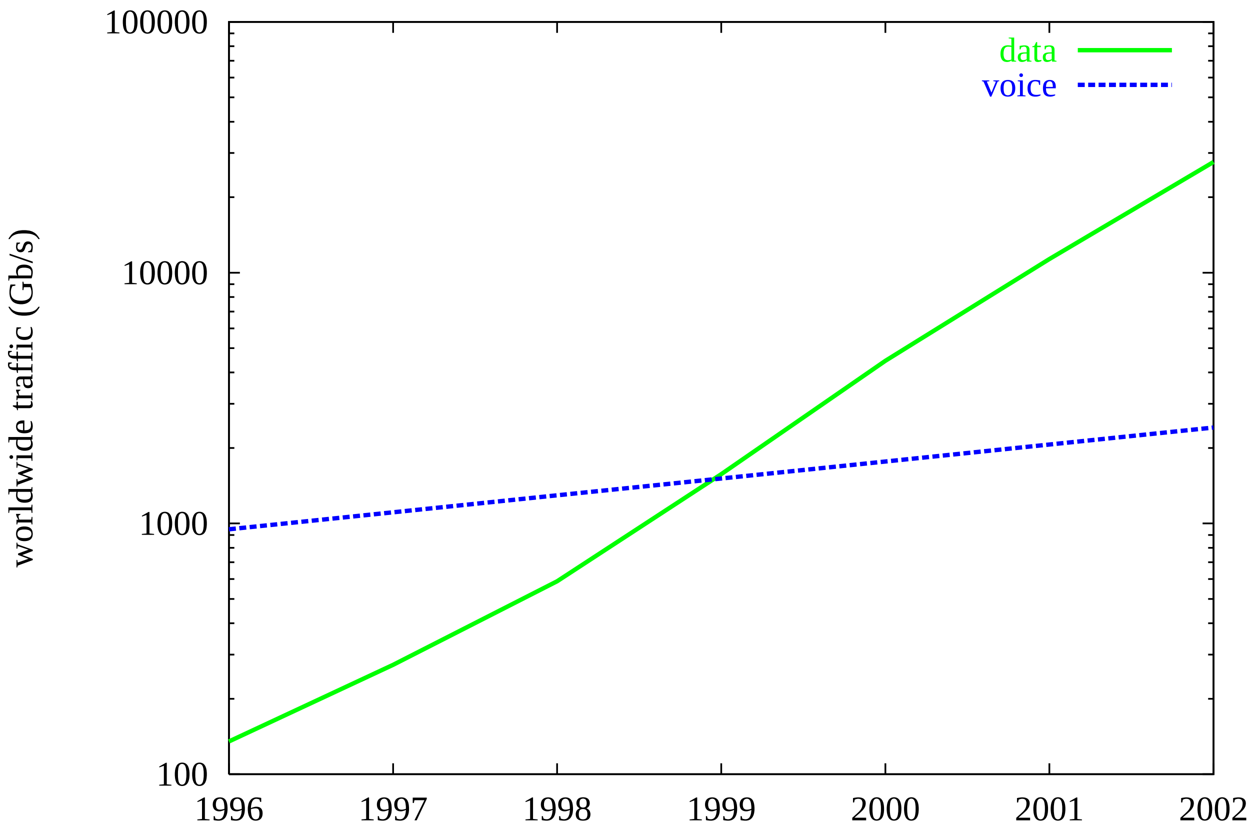

Figure 2.5: Voice vs. data traffic

2.8 Why Internet Telephony?

The use of the Internet to carry traditional "telecommunications" services such as telephony, fax, radio and television is just beginning. This marks a pronounced shift from the traditional use of the Internet, where it offered new means of communications, such as email or file transfer, rather than competing with dedicated networks.

Before discussing the protocols and algorithms that make this possible, it is instructive to take a closer look at the motivation for these changes. This section tries to motivate why Internet telephony may well replace the "innards" of the current telephone system. The emphasis of different motivations may well change over the years.

For brevity, we will use the term "Internet telephony" both for carrying voice and fax across the Internet. It should be understood, however, that once the Internet is used to carry voice, adding other media such as video or sharing applications is mostly a matter of adding bandwidth.

What makes Internet telephony different from computer-based conferencing? Internet telephony is dominated by two-party calls set up on demand rather than in advance, will probably remain dominated by carrying voice and has to interoperate with a huge installed base of "end systems" (telephones and switches) dating back to the 19th century. Rather than desktop computers running graphical user interfaces, most end systems even in a world dominated by Internet telephony will have to be simple, cheap single-function devices. Also, service expectations are far higher than for computer services. After all, people expect to use the telephone to call the fire department and to notify the electric utility when line power has been lost.

The Internet can be seen as the latest step towards a unified, integrated-services network. Traditionally, service integration has taken place at the trench level (several fibers sharing a single duct), at the physical layer (e.g., wave-length division multiplexing (WDM) or synchronous optical networks (SONET)) or at the link layer (ISDN and ATM). Service integration at the network layer offers the same packet-based service platform with different lower-layer technologies spanning about seven orders of magnitude in link speed, from a 300 baud modem to 10 Gb/s and beyond. IP is the only network technology working across seven orders of magnitude of bandwidth.

Compared to the current circuit-switched network controlled by a separate signaling network, using Internet technology to provide telephony services has a number of advantages:

Compression: Internet telephony allows the parties to use the encoding most appropriate for their quality needs. They may, for example, decide that for an international call, they would trade lower cost for full toll quality, while a reporter calling in her story to the radio station may go for full FM broadcast quality with little regard for price. Even without quality degradation, 5.3 kb/s (G.723.1) to 8 kb/s (G.729) are sufficient to support close to toll quality, while the current landline phone network requires 64 kb/s[5]. This flexibility also has the advantage that during severe network overload, e.g., after a natural catastrophe, telephone customers can still communicate at about 3 kb/s, increasing network capacity twenty-fold.

Compression primarily benefits the network provider if services are priced at a flat rate. However, it also allows better use of scarce network access resources, for example, a single copper twisted-pair operated as a "consumer-grade" digital subscriber line (640 kb/s downstream, 90 kb/s upstream) can now support about ten simultaneous voice calls.

Silence suppression: Sending audio as packets makes it easy to suppress silence periods in voice conversations, reducing bandwidth consumption by about half. The savings are even more pronounced for multi-party voice conferences or for voice announcement systems. In a telephone conference call, all participants use 64 kb/s in both directions, even though they, for most of the time, only make use of the incoming voice bits.

[5] Voice is compressed to lower bit rates on some international routes and in so-called statistical multiplexers.

It should be added that silence suppression and voice compression compensate for the lower bit efficiency of packetized communications and Internet telephony in particular. As we will see in more detail later, a typical voice packet may contain 50 bytes of voice samples and 40 bytes of packet headers. In low-speed access networks, header compression, discussed in Chapter **??**, can reduce this overhead to less than 10%.

Multimedia: Attempts to integrate multiple media into circuit-switched communications have been made, e.g., by ISDN-based video conferencing [145], but they have been largely limited to corporate environments. They also tend to be closed systems, without the ability to add new media types. For example, it would not be feasible to add a distributed video game to a conference. ISDN teleconferencing is limited by the multiplexing mechanism to fixed-rate channels, which is not very efficient for video and data. This is primarily due to the difficulty of multiplexing several channels into a single basic rate or primary rate channel.

Because IP is entirely packet-based and thus has multiplexing as its core functionality, media communication is not limited to a single fixed-rate communications channel as it is in circuit-switched network. Internet telephony can, as appropriate for the environment in which it is being used, use very-low-bitrate speech encodings, or high-bandwidth video. Multiple media sessions can also be used in a single call, and these media sessions will inherently multiplex the communications channel between the endpoints. Bandwidth usage can even vary dynamically within a call depending on network conditions, with end systems stepping down to a lower-bandwidth encoding as a network becomes more loaded, then restoring higher quality once resources are again available.

However, more effort is required to guarantee a continuous bandwidth, as discussed in Chapter **??**.

Multicast: IP supports network-level multicast protocols, without requiring application-level devices such as bridges. This enables a number of features both at the signaling and media levels. At the signaling level, it is possible to support a number of features such as “reach any member of a group,” without needing a server to distribute the request explicitly. More interestingly, media can also be multicast; this allows multi-party conferences to be established, in a bandwidth-efficient way, without the need for a conference bridge; and the transition between “multi-party telephone calls” and “large-scale conferences” can be made seamlessly, with no distinction necessary between the two.

UNI vs. NNI protocols: In both analog and digital telephony, there are different protocols and interfaces between the end system and the network compared to the protocols that connect switches and carriers. The former are often called user-network interfaces or UNI, the latter network-network interfaces or NNI. This distinction largely disappears for Internet telephony. (There are some exceptions in H.323 for RAS and inter-gatekeeper communications.) Indeed, Internet telephony does not make a strong distinction between user devices and network devices; a device sending a request typically is not aware (and does not need to be aware) of whether it is communicating with to a signaling server or an end system. Because of this unification, Internet telephony deployment can scale from a few individuals running their own end systems, to a giant organization providing elaborate services and user location features; and these two organizations can interoperate cleanly. What's more, this means that even a customer of a large provider can choose to bypass the provider if his current needs don't require its services; for simplicity, flexibility, reliability, or privacy reasons, users can choose to communicate with each other directly end-to-end rather than through intermediate servers, without any need to modify their end systems.

International standards: For most of the 20th century, telephone networks were operated as national monopolies, whether state-owned or state-regulated. Each national telephone network has its own technical standards, ranging from different phone jacks and tones to signaling protocols, converging only slowly. (Even for the most recent protocols, ISDN and ISUP, there are several national variants, e.g., the ANSI and ETSI versions.) So far, there is only one set of Internet standards, both for the lower layers and the telephony-specific protocols, greatly simplifying product development and deployment. In the Internet space, the greater danger are company-proprietary protocols, particularly those promulgated by operating system vendors with captive customer bases.

Parallel networks: By their nature, circuit-switched networks, if they are to enable communication among large numbers of people, require some sort of parallel signaling mechanism which enables circuits to be established. Because communication channels cannot be constantly maintained between every pair of stations that might wish to communicate, this parallel mechanism must be "self-routed" — an originating node specifies the destination of its signaling request, and the network sees to it that the request arrives at its destination; a circuit is established while this process takes place. The Internet, however, is inherently self-routing. Both signaling and media are sent off into the network through the same mechanism; thus there is no need for two parallel infrastructures to be maintained.

Out-of-band transmission: Additionally, because of the end-to-end nature of the Internet, the paths by which signaling and media traverse the network can be widely disparate. While in the PSTN signaling and media can indeed travel by separate routes, so-called "out-of-band" signaling, the architecture of that network still requires the two types of data to traverse the same administrative domains. In the Internet, by contrast, the routes which signaling and media traverse can be entirely disparate — only the end points of the two paths need to be the same. Media packets are normally sent end-to-end — thus traveling over the "natural" route the Internet's low-level routing protocols have established between the endpoints — whereas signaling can travel across many servers which can provide elaborate third-party services.

This separation greatly simplifies call forwarding and number portability, i.e., the ability to keep one's phone number when changing local telephone companies (see Section 1.3.9). Allowing signaling to travel without carrying the bearer channel alongside with it avoids the problem of having the call double back on itself when being forwarded, called *tromboning*.

Capability labeling: Internet telephony protocols allow for capability labeling of end systems. In traditional networks, one often encounters the problem of a voice caller accidentally reaching a fax machine or modem, or vice-versa. Internet telephony, by contrast, prevents this in two ways: first, since the media type specifications for voice and fax differ, a voice-only end system will immediately reject the call with an "unsupported media type" error. On a broader scale, an end system can identify itself by the type of communication it supports; when a caller is searching for a destination, it can specify the type of communication desired in the call, and thus network devices can automatically resolve and prevent incompatible calls. It should be noted that ISDN has a limited ability to categorize the "bearer capability" of end systems.

Traffic separation: Sending faxes across a circuit-switched network is rather inappropriate, as this is delay-insensitive, but loss-sensitive traffic. In addition, typical fax machines use only 9.6 kb/s of an voice channel that could support 56 kb/s or 64 kb/s[6]. Thus, fax traffic should be separated from voice traffic as close to the fax machine as possible and converted into either email messages or a TCP connection [146].

User identification: Standard telephone service offers, for a price, caller id indicating the number or, occasionally, name of most callers[7]. However, during a bridged multi-party conference, there is no indication of who is talking. The real-time transport protocol (RTP) (Chapter **??**) used for Internet telephony easily supports talker indication in both multicast and bridged configurations and can convey more detailed information if the caller desires.

Number supply: Phone numbers are in short supply or at least cost money. Often, a company has to resort to a two-level addressing scheme such as extensions; individual members of a household rarely have their own phone numbers, although distinctive ringing services have been introduced. Subscribers cannot easily have both a "private" and a published phone number, where the latter might only be answered during business hours. All these restrictions disappear with Internet telephony, as discussed in Section **??**. The ability to have "throw away" Internet telephony addresses offers new opportunities for privacy, but also raises the same potential "spam" problems as in today's email unless strong authentication is provided.

Proxies can easily map numerous identifiers to a single end system address, so that additional Internet telephony identifiers do not consume scarce IPv4 addresses.

User interface: Most telephones have a rather limited user interface, with at best a two-line liquid crystal display or, in the public network, cryptic commands like "*69" for call-back. Advanced features such as call-forwarding are rarely used or customized, since the sequence of steps is typically not intuitive, with little guidance except for sequences of tones. The graphical user interface offered by Internet telephone devices can be more readily customized and offer richer indications of process and progress. It is also easy even for dedicated Internet telephony "appliances" to be configured through a web page.

[6] Again, some transoceanic circuits convert fax modem signals to digital data and thus reduce the bandwidth.

[7] The phone company sends a low bit-rate modem signal containing the phone number to the phone between the first and second ring.

Security: While the Internet is generally considered insecure, it is actually far easier to tap a standard analog phone interface than a typical switched-Ethernet installation. The telephone network almost exclusively relies on physical security, i.e., keeping unauthorized persons from connecting to the telephone transmission or switching infrastructure. This works reasonably well only if all parts of the network and their operators can be trusted.

Beyond physical security similar to the existing phone network, Internet telephony makes it easy to routinely encrypt all signaling and media traffic on the network end-to-end, i.e., without having to trust intermedia entities. (A 90 MHz Pentium computer can encrypt data using the DES (Digital Encryption Standard) at a rate of 10.7 Mb/s.) Users only have to expose enough information to route the call to third parties, but no content.

Unless governments impose legal restrictions, traditional phone tapping, for both legal and illegal purposes, will become extremely difficult. If the government imposes restrictions on cryptography, it will become difficult for the law-abiding citizen, but adding an "illegal crypto" charge probably does not particularly concern those engaging in felonies. Detecting whether a particular voice channel is encrypted is rather difficult, particularly since sophisticated criminals can hide, using a process called steganography, an encrypted stream in an innocuous-sounding audio stream.

Privacy: The only real means of privacy in the phone system are unlisted numbers[8] or a secretary that screens calls. For Internet telephony, it is relatively easy to provide authentication of callers, by means of passwords or cryptographic certificates. Thus, denying a former significant other telephone privileges no longer requires changing one's phone number, but merely removing the person from the "not-forwarded-to-voicemail" list.

Anonimity: Non-800 phone numbers reveal a great deal about the location of a caller, even without access to phone company records or an inverse directory. Internet telephony addresses can be location independent; IPv4 addresses can be traced to the Internet service provider or modem bank, but it is relatively straightforward to set up anonymizer services that hide this information.

Computer-telephony integration: Computer-telephony integration (CTI) [147] allows computers to either control an analog or digital telephone or a circuit-switched PBX from a computer. For controlling end systems, TAPI [148], the Telephony Application Programming Interface, is commonly used in call centers, with a Java-based version called JTAPI developed more recently. Because of the complete separation of data and control paths and the separation of end systems, computer-telephony integration (CTI) [149] is very complex, with specifications for PBX-based control [150] running to 3,300 pages. All the call handling functionality can be much more easily accomplished once the data and control path pass through intelligent, network-connected end systems.

Signaling protocol richness: Internet telephony signaling protocols are significantly more expressive than those of the PSTN. This is particularly true compared to the limited signaling

[8] About half of California's numbers are unlisted.

of tones and hook signals available to two-wire analog telephones. Rich signaling in Internet telephony eliminates many previous limitations on feature development. For example, an end system no longer needs to indicate its desire to transfer a call through an elaborate sequence of switchhook and DTMF tones; it can explicitly indicate to its partner the party to which the call should be transfered.

Furthermore, Internet telephony signaling is extensible, and can be extended while maintaining compatibility. As new signaling properties or events are invented, they can be added to the existing protocol in ways which can interoperate cleanly with existing implementations, either by providing richer information about the signaling information or by allowing fine-grained control over what features are required to be understood in order to understand a signaling message successfully. Internet telephony devices can also query each other to determine what properties and parameters they support. As new signaling elements and capabilities are developed, the network will be able to evolve gracefully to support advanced features without needing to undergo painful universal upgrades of an entire system.

Shared facilities: Many corporations and universities already have high-speed local area networks. Given the low bit rate of packet voice, voice and low-bit-rate video can be readily supported on a well-designed (switched) LAN, even without explicit quality-of-service support.

Advanced services: From first experiences and protocols, it appears to be far simpler to develop and deploy advanced telephony services in a packet-switched environment than in the PSTN (public switched telephone network) [151, 152]. Internet protocols that support standard CLASS (Custom Local Area Signaling Services) [153] features take only a few tens of pages to specify. They can replace both the user-to-network signaling protocols such as Q.931 as well as the network signaling (ISUP, Signaling System 7) and, through cryptography, can be made at least as secure as the existing network. We will discuss such services in detail in Chapter **??**.

In the Internet, application-layer intelligence resides in end systems, which are typically replaced much more frequently and have an higher aggregate processing power than typical telephone switches. It is also easier to deploy services one-by-one, rather than waiting for the whole network to be upgraded. (On the other hand, implementing services in switch adjuncts makes them available immediately to all subscribers, regardless of the intelligence of the end system.) Due to implementation diversity, it may also be less likely that software faults in implementations of new features would bring down the whole network. (The Internet can also be used for service creation in the circuit-switched telephone system [151].)

Programmability: For all but the largest telephone customers, phone service offers very few opportunities for customization. Customers can choose only from a menu of services, with limited ability to change parameters such as phone numbers. With Internet telephony, the notion of just modifying a menu of services can change to programming services, with full generality. For example, users can easily make the behavior on incoming calls depend on time-of-day, the identity of the caller, number of ring cycles, any password entered and whether the person has called (or emailed) before. Since services reside on the Internet, it is easier to connect telephone services to other Internet-resident data, such as personal address books or appointment calendars. We will discuss two approaches to programming Internet telephony services in Chapter **??**, namely a call processing language designed for handling calls and a programming interface similar to that used by web servers to create dynamic content. We anticipate that this allows different interfaces, from very basic choices similar to today, but configured via the web rather than via interactive voice response, to full-fledged user-written programs and scripts with access to network databases.

Programming Internet telephone services at multiple locations, possibly implemented by different parties, raises new issues of *feature interaction*, where the end-to-end behavior may not be easy to predict.

Additionally, because user programmability is now possible, the new phenomenon of *features created by amateur feature designers* arises. Because new services can be created and deployed with much the same level of ease that, for example, web services can be created today — a simple service can be put together by a reasonably experienced programmer in a matter of hours — they may be created by programmers who may not consider feature interaction issues thoroughly, either through ignorance or expediency. Such distributed problems may be dismissed as the "just desserts" of customers of incompetent feature designers, but unfortunately other service providers will have to inter-operate with such services.

This is already evident, to a limited extent, with programmable mail filters, except that once telephony moves to the Internet, one cannot, as a last resort, call up the person one wishes to reach to tell them that their script is broken. Maybe somebody will offer a send-a-postcard service that delivers a "Sorry, I tried to reach you, but all I get is *the number you have reached is currently experiencing a core dump*."

We will describe additional complications in Section 2.12.

Cost: One of the primary drivers for the commercial introduction of Internet telephony is the ability to offer cheaper long-distance calls. Compared to traditional telephony, costs can be reduced by improved bandwidth efficiency, ease of building services and more cost-efficient switching (see Section 1.9). However, another, probably short-lived, attraction is that currently, Internet service providers do not have to pay the local access charge imposed on long distance carriers. This charge, at an average of 4.92c per minute (1998), constitutes close to half of the cost of an average long-distance call.[9] The pure bandwidth cost of a phone call (1998) is approximately one to two cents a minute[10].

Separation of bit transport and service provision: In Internet telephony, services can be provided by third parties — organizations dedicated only to providing services, with no intention of providing actual voice or multimedia transport — as easily as they can be by the original provider, and indeed providers may well specialize into service provision or data

[9] In addition, telephone lines are charged the monthly Federal Subscriber Line Charge, currently approximately $3.50/month (1998). Also, they are charged a Presubscribed Interexchange Carrier Charge of $0.53/month.

[10] Beck, private communications

transport, as these are rather separate tasks. This is already happening for data-oriented Internet services, where web hosting, email services and Internet access are available from separate parties. There appear to be two reasons to separate services, namely increased efficiency of sharing servers and access lines for web hosting, and the ability to keep the same address even when changing network providers. (The latter is particularly important since many network providers are only available regionally.)

The broadly distributed environment also introduces some new possibilities in terms of trust models for Internet telephony. It is relatively easy in Internet telephony for a customer to proxy all his calls through a service which, for example, automatically blocks calls from known telemarketers. A traditional telephone company does not have much interest in providing such a service — and few customers would likely trust a telephone company to provide it reliably, as telemarketing calls provide the company with revenue[11]. The introduction of the distributed network allows users to have trust relationships with organizations other than their service provider.

The separation of bit transport and service provision offers new mechanisms of paying for these services, as well as competition. Currently, these services cost between $2 and $10 a month, probably far in excess to the actual cost of providing the software. In an Internet environment, these services could either be purchased like software or be partially financed via advertising, similar to the web email services available today.

For example, services such as caller id, call waiting or forward-on-busy can currently only be offered by the telephone company. In an Internet context, these services can be implemented by the end system or any service provider.

Opportunity for CLECs: CLECs (Competitive local exchange carriers) or CAPs (competitive access providers) are telecommunication carriers that compete with the traditional phone company, also known as the ILEC (incumbent LEC). In the United States, the RBOCs (regional Bell operating companies) and GTE are ILECs, while companies like MFS and some long-distance carriers have attempted to become CLECs. Since it is prohibitive to dig up the street for a second set of phone wires or run wires on poles, most CLECs lease the local loop, the copper wire from the local phone switch to the residence or business. However, local competition has not happened nearly as fast as anticipated. It has been suggested that cable TV companies and utilities may use their networks, as discussed in Section **??** to offer Internet access and Internet telephony services.

2.9 Protocol Architecture

We now present an overall Internet protocol architecture that can support telephony and other continuous-media services such as “Internet radio” and “Internet TV” [122].

[11] USWest does provide a service that announces to the caller that the party called does not appreciate phone solicitations, for $6.95 a month…

2.9.1 Data Transport

For transporting real-time data across the Internet, the accepted end-to-end protocol is the Real-Time Transport Protocol (RTP) [48, 154]. It is also used by the ITU-T H.323 teleconferencing recommendation. RTP is a thin protocol providing support for applications with real-time properties, including timing reconstruction, loss detection, security and content identification. An associated control protocol called RTCP provides support for real-time conferencing for large groups within an internet, including source identification and support for gateways (like audio and video bridges) and multicast-to-unicast translators. It offers quality-of-service feedback from receivers to the multicast group as well as support for the synchronization of different media streams.

While UDP/IP is its initial target networking environment, efforts have been made to make RTP transport-independent so that it could be used, say, over CLNP, IPX or other protocols. RTP is currently also in experimental use directly over AAL5 using native ATM services. This protocol stack is shown in Fig. 2.6. RTP is described in detail in Chapter **??**.

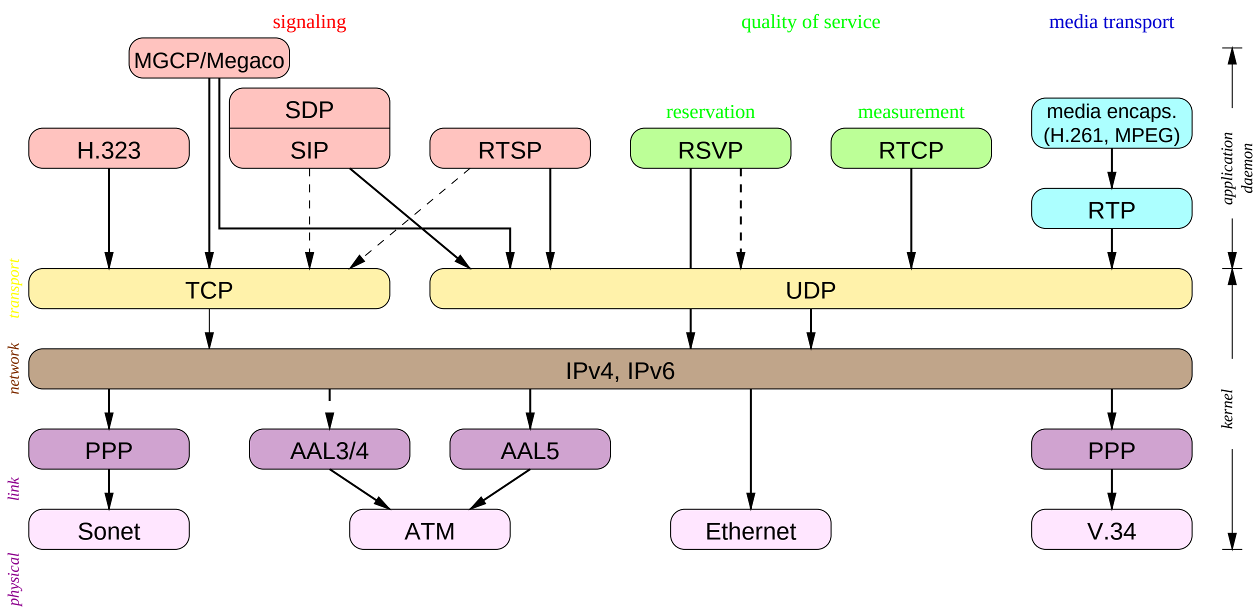

Figure 2.6: Internet-telephony protocol stack

For low-speed links and highly compressed audio, the combined stack consisting of IP, UDP and RTP add 40 bytes to every packet, while 20 ms of 8 kb/s audio only take up 20 bytes. Thus, just like for TCP, header compression is desirable [155].

As an end-to-end protocol, RTP cannot guarantee a certain quality of service. The resource reservation protocol, RSVP [156], might be used to allocate resources to either individual streams or a group of streams. Unlike in the telephone network, establishing connectivity and allocating resources are two distinct operations. Typically, the caller would first “ring” the callee. Only if the called party wishes to communicate and once a set of media has been agreed upon, would RSVP be used to reserve resources. This runs the risk of being denied the necessary network resources, but in any network dimensioned for telephone service, voice calls should be blocked extremely infrequently. Given that the resource needs are only known once the parties agree on the media and their quality, this order is also the only feasible one.

Since packet audio flows have a relatively low, constant bit rate, with deterministic arrival patterns, it may make more sense to treat all voice calls on a link as a single stream for scheduling rather than managing several thousand individual reservations, with the attendant state space and refresh overhead. As indicated earlier, it may be sufficient to simply give this class of service priority. In some cases it may also be advantageous for the end systems to reserve only the minimum necessary bandwidth, and obtain additional throughput and improved audio quality through best-effort flows.

2.9.2 Signaling

In the architecture described here, the traditional telephony signaling protocols, Q.931 for ISDN user-to-network signaling and SS7 for network-to-network signaling, are replaced by a single, much simpler signaling protocol. One such solution, the session initiation protocol (SIP) (Chapter **??**), can establish multimedia conversations with one or more parties. Instead of telephone numbers, it uses addresses of the form user@domain or user@host. In many cases, this address would be identical to a person's email address. Table **??** compares the properties of email addresses to standard telephone numbers.

SIP offers the standard PBX, ISDN [67] or CLASS functionality, including call forwarding, call waiting, caller ID, call transfer, camp-on[12], call park [13], and call pickup [14]. Many of these features actually require no signaling support at all, but can be implemented by end system software. SIP is designed as a variant of HTTP/1.1[158], which allows easy reuse of HTTP security and authentication, content labeling and payment negotiation features.

We used the SIP features to implement a calendar-based call handler. The call processing software accesses a user's personal appointment calendar and answers the phone accordingly. The user can define categories of callers and preset, based on the calendar entry, whether and where their calls are forwarded. The information released to the caller if calls are not forwarded may range, for example, from "is currently not available" to "John Smith is in a meeting until 3 pm in room 5621 with Jane Doe", depending on the caller's identity. In the near future, this will be integrated with the call processing language, a state-based scripting language that allows to construct voice-mail systems or automatic call handling systems in a few lines of code. It also manages the translation between ISDN calls and Internet telephony calls.

2.10 Naming and Addressing

As we discussed in Section **??**, telephone numbers have acquired a range of meanings and have a number of disadvantages, both from a technical as well as from a human-factors perspective.

[12]"Camp-on allows an attendant-originated or extended call to a busy single-line voice station to automatically wait at the called station until it becomes free while the attendant is free to handle other calls." [157]

[13]"Allows user to put a call on hold and then retrieve the call from another station within the system". [157]

[14]"Allows stations to answer calls to other extension numbers within the user specified call pickup group" [157]

Moving to Internet telephony allows us to revisit the issue of naming and addressing and to separate the functionalities of routing, device identification, naming, privacy and billing. A good naming scheme would yield short, easily remembered names that have some predictable correspondence to a person's civil name or role. These names should be as permanent as desired by the owner.

Identifiers should be easy to [159]

- Devise (what is my identifier for X?);
- Memorize (what was the identifier for X?);
- Guess (what could be the identifier for X?);
- Understand (what does X refer to?);
- Correct (spelling errors);
- Identify with (nice identifier, isn't it?);
- Manipulate (write, type, pronounce, ...).

One of the advantages of the Internet is that names, as opposed to IPv4 addresses or telephone numbers, are cheap and plentiful. Currently, there are six types of personal Internet addresses, either in the form of web URLs or email addresses:

Employment-based: Here, the domain part of the address reflects the employer of a person. Email and web pages are generally only forwarded for brief periods after an employee leaves. Employers obviously keep their domain name even if they change Internet service providers.

Affiliation-based: Many professional organizations, such as IEEE, the IEEE Computer Society and ACM, offer email aliases for dues-paying members.

ISP-based: It appears that most private email is delivered to addresses associated with Internet Service Providers. Indeed, it is probably one of the strongest reasons why people stick with their first Internet provider even though they may not be particularly happy with its service. Making ISP-based names the basis for visible Internet telephony names seems to be a step backwards compared to the local number portability efforts in the PSTN.

Hosting and forwarding services: The problems with ISP-based addresses lead to the emergence of web-based Internet mail services that offer only mail delivery. Since these services mostly rely on advertising rather than subscriber revenue, users are more likely to keep their account for longer periods of time[15]. This model can be readily transferred to Internet telephony, although it is less clear how advertising would be inserted. (Maybe the web-page delivery mechanism described in Section **??** will be of use here.) The primary problem with this mechanism is that all users are mapped into a small number of domain names, making naming less intuitive. Personally, the first author is less worried here.

[15] However, there is no guarantee that such a service might not be tempted to start charging for the service once it has captured enough users...

Own domain name: The only way to acquire a name permanently is to acquire a domain name. A domain name is the only truly portable phone number. Currently, names in the `.com`, `.net` or `.org` domain cost $35 per year. Other domains, such as `.nu`, are also readily available. In the near future, a number of registrars will be competing for name registrations, with a central registry for maintaining the database. This is likely to lower the price to the point that it is cheaper than a vanity license plate. Once the price barier is removed, scarcity of names becomes a more pressing issue, since only one Jones (family) can own `jones.com`. However, see the domains `smith.org` or `smith.net` for examples of trying to maintain a single domain for similarly-named people.

In the United States, a geographic naming hierarchy is available through the `.us` top-level domain. However, these names have been used primarily by state and city agencies, such as `state.nj.us` or `mta.nyc.ny.us` since registration by individuals is cumbersome. The addition of a domain for personal names, such as `.nom`, has been proposed as part of the transition of Internet name assignement to a competitive arrangement, but this only addresses the overlap between corporate and personal names, and does not solve the scarcity of "good" names.

Just like start-up companies now pick their name such that their `.com` is not taken yet, maybe in the future, married couples will select the domain name or use a naming consultant to find an available name... Shortly thereafter, the ownership of a domain name will turn up in divorce court.

Government-assigned name: It may be feasible to have a government entity assign names that are either permanent, similar to a social security number in the United States, or that is associated with the person's place of residence. At one point, for example, the United States Postal Service had suggested that it might manage such addresses. For example, addresses would take the form president@pennsylvania_avenue.washington.dc.us, but clearly this would be cumbersome.

Thus, unless a geographically-based naming mechanism can be built, we will likely end up with a combination of names based on affiliation and forwarding services. A likely outcome is that people will accumulate names throughout their lifetime, each of which will stay with them, with forwarding to their current address. A student may start as `jane.lee@lvjusd.k12.ca.us`, move to `j.lee@cs.umass.edu` and then `jlee@bell-labs.com` and `jane.lee@acm.org`. Having multiple addresses also simplifies call filtering, but one would hope that the current privacy mechanism of unlisted numbers can be replaced by more selective mechanisms of caller identification.

A truly permanent name that survives changes in jobs, name changes and moving from place to place would likely be about as mnemonic as a social security number and still would require a central registry, presumably run by a government organization. (It appears that there is an inherent conflict between identifier permanency and privacy. The more permanent and unique an identifier is, the more likely it is to be used to tie together records, thus making it easier for advertisers and government to accumulate dossiers on individuals.)

Generally, while it would be possible to use web-style URLs, such as `acme.com/~alice`, email-like addresses are preferable. First, it is likely that users already have email addresses that could then serve as a first guess at the user's Internet telephone address. Secondly, web URLs mix "little-endian" and "big-endian" in the same identifier, with the scope of the name first increasing from left to right in the host name and then decreasing in the part after the slash.

It has been suggested that the difficulties in naming could be circumvented by using directories such as PH [160] or LDAP [161, 162]. Unfortunately, while there are a number of existing directories for finding people and organizations, their use is rather tedious, with several steps of finding the directory, providing information, then filtering the listings provided, ascertaining that the "Alice Smith" found is indeed the correct one and transferring the information to the desired program. Most directories are either based on phone directories, self-registration or limited to entries from a particular organization. Most businesses would not want you to have to go to a directory to find them – you might find their competition instead. Thus, while global directories are useful, applications typically provide a more limited local "address book", with user-chosen nick names.

feature	phone number	email
mnemonic	no (except for some 1-800)	*name@organization*
multiple per person	no	easy
avg. characters	$\approx$ 10-12	22
location-independent	1-700	yes: j.doe@ieee.org
carrier $\neq$ naming	maybe	yes
directory	411, 1-555, switchboard.com	LDAP [161]

Table 2.5: Comparison of (U.S.) telephone numbers and email addresses
(Note: average email address size is estimated from the IEEE ComSoc TCCC mailing list, an international list with about 500 members.)

2.11 Internet Telephony Services

In the earlier discussion, we had argued that one of the prime motivators for deploying Internet telephony is the ability to create new services. Internet telephony services can be roughly grouped into four categories, namely Internet versions of existing services such as call forwarding, Internet integration services,

2.11.1 "Classical" Telephony Services

Classical telephony services, described in detail in Section **??**, can be implemented in Internet telephony. The implementation details differ between the master-slave and peer-architectures and among the signaling protocols, so we defer discussion of implementation to the chapters on SIP (Chapter **??**), H.323 (Chapter **??**) and MGCP (Chapter **??**).

For the peer model, many services that used to be implemented in telephone switches, can now migrate to the end system. This includes features such as distinctive ringing or call waiting. In Internet telephony, there is no limit to the number of calls that can be present at the end system; the only limit is the number of simultaneous media streams. Thus, call waiting simply needs to temporarily shut off media streams, while the call state itself remains. Similarly, camp-on does not consume resources in either the caller's or the callee's system; thus, one could camp on for three weeks until the caller arrives home from vacation.

Other services, such as call filtering and call forwarding, can be implemented either by IP telephony servers or end systems. However, it appears likely that instead of enumerating conditions such as "call forward busy" or "call forward unconditional", calls will simply be handled by a script in the server or end system, as described in Chapter **??**. This allows to make the behavior depend on the time of day, the identity of the caller, the urgency of the call, the number of other calls pending or in progress and any number of other variables. This avoids the current problem, where each variation on a feature generates a new feature.

Table 2.6 summarizes where services may be implemented. In the table, "*69" describes the call return service, while "call curfew" blocks incoming and outgoing calls during specified time intervals.

	end system	server
caller id	x	–
call forwarding, follow me	x	x
three-way calling	x	–
distinctive ringing	x	–
69	x	?
no solicitation	x	x
do not disturb	x	x
call curfew	?	x

Table 2.6: Implementation of services in a peer-to-peer architecture

2.11.2 Internet Integration Services

In the traditional phone system, integrating the telephone with other applications is generally cumbersome. Computer-telephony integration (CTI) can control single lines if suitable end systems or PBXs are available, but it is limited to placing and receiving calls.

Internet telephony can make integration with other Internet applications seamless if the signaling protocol supports it. The services we describe below are currently only feasible in their generality with SIP (Chapter **??**).

Candidates for integration include web and related technologies, email, chat services such as Internet Relay Chat (IRC), multi-user environments such as MUDs, streaming audio and video, network games, and presence information, in addition to remote-object services such as Corba and DCOM and remote-procedure call applications such as Sun RPC.

One can readily imagine new services building upon this integration. For example, an enhanced 911 service would allow the caller to give the emergency operator temporary access to medical records or detailed in-building directions.

Internet applications can be integrated “by reference” or “by value”, i.e., by including references (URL) to other Internet services or by carrying other Internet content in protocol messages. Most Internet services are addressable by URLs. An Internet telephone can then forward calls to any such service. For example, an incoming phone call could be redirected to a web page containing information about how to reach the person called or replace the traditional “press 1 for sales, press 2 for service” interactive voice response (IVR) interaction. This service is sometimes called web-IVR. Together with web forms or Java applets, the caller can provide additional details that might help route the call. Forwarding to email or chat may also be useful if the called party does not wish to be disturbed or prefers a silent interaction.

Mixing internet services can also enrich the pending phone call. For example, while waiting for an agent to pick up, web pages can update the caller on the expected waiting time or provide alternatives to talking to a human agent. One could even set up a conference among all those waiting so that they can chat, replicating the more social experience of waiting in a real line. Through web forms, the caller could also choose the music on hold rather than being subjected to the music choice of the company called (or to the traffic and weather reports of a far-away city…).

The call itself can contain far more information about the caller. For example, it is easy to add the vCard electronic business card [163] or a URL with the caller’s picture to a SIP call setup request.

The ability to route audio and video to different addresses allows services such as “dubbed” phone calls where the video gets transmitted directly, while the audio is mediated by a human translator.

Currently, the action of making a phone call is clearly visible as such, with devices dedicated to that function. In the future, Internet phone calls may be made without the caller necessarily thinking of this as a phone call. In the simplest form, clicking a button or link on a web page or email initiates a call, in a service known as “click-to-call”. In virtual worlds, two avatars may be able to hear and see each other if they move within shouting distance in their VRML coordinates.[16]

We tend to think of broadcast, radio and TV, and telephony as completely separate services, with different devices, user expectations and content. Current telephone services generally do not scale well to large groups, with even modest group sizes requiring the prior manual allocation of a conference bridge, distributing call-in numbers and passcodes. However, in the Internet we can have the same end systems and protocols serve group communications with sizes ranging from two to millions. Indeed, among the protocols listed in Table 2.1, almost all are applicable to distribution or broadcast services. (The only exception are protocols like TRIP (Chapter **??**) for interfacing with the PSTN.) Making a phone call to a friend differs from distributing an Internet TV show primarily in how participants find out about the intended communications, namely either by explicit invitation or through some announcement mechanism such as a web site, session announcement protocol (Section **??**) or mailing list. Even this distinction can get blurred. As we will see in

[16] A similar service, using proximity on a web page as a trigger for calls, was first implemented in a system built around 1994 by Ubique in Israel.

Chapter **??**, the Session Initiation Protocol (SIP) can be used to invite somebody to an Internet broadcast[17], without the "TV station" being aware of this.

Internet integration also allows the blending of applications. For example, while currently voice mail and email are mostly handled by separate applications, it is likely that in the future a voicemail message will simply appear to the callee as yet another email message. We will discuss the implementation of Internet voice mail in Chapter **??**. Treating voicemail as email offers easy forwarding, filing into folders and ready transition to videomail. The integration of fax, email, and voice mail into a single application has become known as *unified messaging*.

Other functionality currently used to manage email can also be drafted into enriching telephone functionality. For example, address books, email logs and email "spam" filters can probably serve Internet phone calls, in particular, if the naming mechanism of the two means of communications is similar (Section **??**). Thus, the now separate worlds of synchronous communications, phone calls, and asynchronous communications, email and fax, will be much more closely integrated.

2.12 Feature Interaction Problems

On a network level, the characteristics of the Internet introduce some new complications. First of all, the fact that *media packets travel end-to-end*, without being interceptable by intermediate servers, means that intermediate servers can no longer implement a number of features transparently. For instance, ordinary signaling servers cannot listen in on calls to collect digits ("press '#' for new call"); perhaps more significantly, they cannot perform "pipe-bending" services, where an intermediate system moves one endpoint of a call from one end system to another — for example, to transfer a call — without explicitly informing the end systems of the new locations to which they should send their media packets.

Another related complication is the fact that *end systems have control of call state*. While this introduces many new possibilities for general feature creation and deployment, it also complicates issues in situations when the network wants to be able to impose control contrary to the expressed desires of an end system. For example, in traditional telephone networks, 911 (emergency) calls are usually handled specially, so that end systems cannot hang them up; the emergency operator must hang up the call before the line is cleared. If the end system controls its own states, however, it is impossible for the network to enforce this without the end system's cooperation.

Several new features of Internet telephony protocols also have the potential for dramatic feature interaction consequences with existing protocols. Probably the most dramatic of these is what is known as the *forking proxy*. A signaling server, or proxy server, can take an existing call request and transmit it in parallel to several other devices. We discuss some examples of complex interactions that can occur with this feature in section **??**.

Another new feature is *request expiration*. A request, when it is placed, can specify how long it should be considered valid — a user might want a call to only ring for the equivalent of four rings, for example — but services on subsequent signaling servers may be programmed to do different things when the expiration time elapses.

[17] Sticklers for terminology will note that in the Internet, services resembling radio and TV are actually implemented using multicast, as we will discuss in Chapter **??**.

The Internet's *lack of address scarcity* can also complicate some common features. In traditional telephone networks, where telephone numbers are difficult to obtain, a telephone number can be used, reasonably effectively, as a representative of a party's identity for such purposes as incoming or outgoing call screening. In the Internet, however, "throw-away" addresses become easy to use; someone wishing to evade a block on their address can switch to another one with minimal effort.

Related to this problem is the Internet's *trust model*. In the PSTN, telephone users generally assume that they can trust their telephone company to provide accurate information, that their telephone company will not reveal private information to third parties when inappropriate, and that the wire leading out of their house indeed connects to the telephone company and no one else. Telephone carriers, meanwhile, can assume that the signals they get from a subscriber line are indeed coming from that subscriber; and signals they get from other telephone companies are reliable and secure. All these assumptions break down when end-to-end connectivity is introduced and anybody can become an Internet Service Provider. Forging communications becomes relatively straightforward when packets may be sent from any location on the network to any other, and intercepting them, while somewhat more difficult, is still significantly more tractable than on a telephone network, due to Internet characteristics such shared-bandwidth communications channels and dynamic routing protocols. While protocols for strong authentication and encryption have been developed, deployment of a key infrastructure which would enable large-scale trust is still a long way off[18]

Additionally, features like "caller I-D blocking" become much more difficult when users cannot trust the network not to reveal calling information to recipients — and indeed cannot reliably distinguish whether they are communicating with a "network" or a "recipient."

2.12.1 Applicability of existing feature interaction work

Existing work on feature interactions is applicable to the Internet environment in some circumstances. If we consider the framework of Cameron et al. [164], single-component interactions (those where all the interacting features are implemented on the same network component) are largely the same in the Internet environment as they are in traditional telephone networks, and we expect the techniques developed to resolve these interactions to work in the new environment.

An example of single-component interaction that can be dealt with in the Internet as it is in the PSTN is Cameron et al.'s Example 1, the interaction between *Call Waiting* and *Answer Call*. These two features have conflicting definitions of what should occur when a call attempts to reach a busy line: to signal the user with a tone, or to connect the calling party to an answering service, respectively. If, in an Internet telephony environment, both these services are deployed in the same device, or in multiple devices controlled by the same organization, techniques for resolving their interaction would carry over naturally from the PSTN.

Multiple-component interactions, however, are much more complicated for Internet telephony. The problem arises as features are designed and deployed by providers who do not cooperate,

[18]In addition, many of the existing user-level certification services simply assure that the presenter of the signed request can indeed be reached by the (email) address indicated, but do not associate a legal or civil identity with a key.

and have no interest in doing so; therefore, feature interaction resolution techniques such as those in [reference something or other] which depend on being able to describe features globally, and resolve their interactions at the time they are designed, are no longer practically applicable. (This is, of course, a growing problem in the PSTN as well, as increasing numbers of providers enter the market.)

2.12.2 Examples of new interactions in Internet telephony

Several varieties of new feature interactions appear in Internet telephony which either do not appear or are not as severe in traditional telephone networks. We categorize these into two types of interactions: *cooperative* interactions are those where all the parties who implement features would consider the others' actions reasonable, and would prefer to avoid an interaction if it were possible. *Adversarial* interactions, by contrast, are those where the parties involved in the call have conflicting desires, and one is trying to subvert the other's features. Roughly, cooperative interactions correspond with those that [164] describes as single-user multiple-component (SUMC) interactions; adversarial interactions are more commonly multiple-user multiple-component (MUMC) or customer-system (CUSY) interactions.

Cooperative interactions

"Cooperative" feature interactions are multiple-component feature interactions where all the components share a common goal — typically, allowing the caller to communicate with his or her intended called party — but have different and uncoordinated ways of achieving that goal. These conflicting implementations can interact in ways that can prevent the most desirable means of communication from occurring, even though it would be possible given the state of the parties involved; and can result in surprising or unpredictable consequences of deployed services.

Example 1 *Request Forking* and *Call Forward to Voicemail*

Request Forking allows an Internet telephony proxy server P to attempt to locate a user by forwarding a request to multiple destinations, A and B. The call will be connected to the first destination to pick up, and the call attempt to the others will be canceled. The interaction arises when the user to be reached is currently located at A, and another, B, has had its calls forwarded to a voicemail system. The call to B will be picked up first, as it is an automated system, and thus P will connect the call from B and cancel the call from A. The caller will never be able to reach the actual human.

Example 2 *Multiple Expiration Timers*

A SIP request may specify a length of time for which the request is valid. Difficulties arise, however, if several servers are programmed to have special behavior if the timeout elapses before the call has been definitively accepted or rejected. For example, one proxy server P_1 may be programmed to forward a call to a voicemail server when the expiration has elapsed, whereas another server P_2 may respond with a web page describing giving alternate ways of contacting the destination. If P_1 is earlier in the call path of P_2, the former server considers the latter server's response to be a definitive response to the call; and if P_2's response arrives at P_1 before its own timer expires, P_1 will forward that response back to the original caller rather than triggering its own expiration behavior. The two timers have the same nominal expiration period (the length of time specified in the request); which one executes first depends on factors such as processing time and the precision of the two servers' clocks. Therefore, there is a race condition of which of the two expiration-related services will be executed.

Example 3 *Camp-on* and *Call Forward on Busy*

Camp-on allows a caller who reaches a busy destination to continue to re-try that destination periodically until the line becomes free. However, if the destination has *Call Forward on Busy*, the call is forwarded to some alternate destination in this case, and the caller never receives the busy indication; thus there is no way to trigger the camp-on service. This is an interaction which can also arise in the PSTN, but it is more serious in Internet telephony for several reasons. First of all, because Internet telephony places so much additional power and call state knowledge into end systems, Call Forward on Busy is likely to be triggered by intelligent services implemented an end system, which may not be aware that the other party is attempting to camp on. PSTN switches which try to camp on will generally also be the location where Call Forward on Busy is implemented, and thus can resolve the interaction locally. Furthermore, camp-on services in the Internet will generally need to be globally usable; users will not accept camp-on services which work only within one provider's network, so state cannot be shared easily among servers in a private manner either.

Adversarial interactions

"Adversarial" feature interactions, by contrast, are those where several of the parties involved — the caller, the destination, and/or either endpoint's administrator — disagree about something having to do with the call, typically about whether it should be allowed to be completed. These can be more difficult to resolve reliably than cooperative interactions, because generally parties attempting to subvert others will find ways to lie to them, or bypass them. They are also more complicated because users will generally be quite upset if the network allows their expectations about security or privacy to be violated.

Example 4 *Outgoing Call Screening* and *Call Forwarding*

Outgoing Call Screening blocks calls at an originating party based on the address to which a call is placed. However, even if a Call Screening service blocks calls to an address X, another signaling server, downstream from the location where the blocked is imposed, may forward calls originally directed to a non-blocked address Y to the blocked address X. This interaction also appears in the PSTN, of course (and this description is largely taken from [164]), but the ability to easily change addresses and get easy call forwarding on the Internet makes this problem much more significant in the Internet environment.

Example 5 *Outgoing Call Screening* and *End-to-end connectivity*

Because the Internet provides end-to-end connectivity, enforcement of *Outgoing Call Screening* policy is difficult for another reason. A signaling server cannot force calls to be placed through it; because the Internet telephony UNI and NNI protocols are identical, and because any device can talk to any other, an end system can be programmed to communicate directly with the remote party, bypassing local administrative controls entirely.

Example 6 *Incoming Call Screening* and *Polymorphic Identity*

Incoming Call Screening allows a called party — either in a signaling server or an end system — to reject calls from certain callers automatically. Because Internet telephony addresses are cheap, however, and because the caller can switch the identity he presents in his call request, he can easily alter the address he presents as his own in order to evade the screening lists the destination has programmed her phone to reject.

Example 7 *Incoming Call Screening* and *Anonymity*

Even in the absence of a malicious caller, *Incoming Call Screening* can be complicated by a caller's legitimate desire for anonymity. Because the trust model of the Internet does not allow a user to be sure that a network provider will hide the information like caller ID, if a user wishes to be anonymous he must avoid sending all identifying information in the signaling information in the first place — and for assured anonymity will likely have to use an anonymizing server run by a trusted third party, which will hide all information, including the sender's IP address for transmission of media packets and signaling. In the PSTN, a destination switch can easily apply *Incoming Call Screening* and *Caller I-D Blocking* services simultaneously; and both the caller and destination can trust this switch to apply their service reliably. In Internet telephony, however, there are not generally such mutually-trusted third parties, so for anonymous calls the critical information is simply not to the network. There is no reliable way to screen anonymized calls other than simply rejecting all of them.

2.12.3 New Approaches for Managing Internet Interactions

Though Internet telephony brings about new feature interactions, it also presents new possibilities for managing or resolving these interactions. The flexibility of the signaling protocols, and the underlying infrastructure of the Internet, can be exploited to resolve or prevent interactions in a manner which maintains and extends the powerful new characteristics of the Internet telephony architecture.

Explicitness

Many of the interactions which we have categorized as "cooperative" can be prevented or made less likely by making explicit the actions being taken, and their desired effects. Because the Internet telephony protocols are extensible, it is possible to add parameters which tell downstream servers what actual actions are desired; such parameters are currently being standardized [165]. If a call is intended to only reach a human, for instance, it is possible to specify that the call should not be forwarded to a station which has registered with a "voicemail" attribute; intelligent services which would otherwise forward a call to voicemail should know to return a "not currently available" status code instead. Similarly, a call wishing to camp on to the actual user to be contacted could specify "do-not-forward" so as to get back a "busy" response rather than have the call be forwarded against their wishes. The difficulty with this solution is that it can complicate the creation of services significantly; service creators need not only to determine what it is they wish to do, but to determine whether those actions are compatible with the preferences the caller specified with the call. Also, this explicitness requires that the receiver know about the attributes the caller desires; a call may specify "want to reach only the family goldfish," but the recipient is unlikely to be able to do anything useful with this if "goldfish" is not a recognized category.

Universal Authentication

Many of the problems introduced by polymorphic identities and identity forging can be resolved by insisting on strong authentication of requests. Whereas a generic address can easily be used once and thrown away, and indeed a user can claim to be someone else, the barrier toward obtaining certificates giving actual signed identity information is much higher, and presumably widely-trusted certification authorities can be relied upon to be sufficiently consistent in their identification of users that call screening services can use this information to block callers. Unfortunately, all of this infrastructure fails if users accept non-authenticated calls; and authentication is far from being sufficiently widespread enough for it to be practical to accept only authenticated ones. However, we hope that the growth of Internet telephony will help be a driving force for widespread authentication to finally become widely deployed on the Internet.

Network-level Administrative Restriction

Administrative restrictions in the Internet cannot generally be reliably applied at the application level. If users have end-to-end connectivity available, it is not generally possible to prevent them from taking advantage of this connectivity by imposing restrictions solely at the application layer. Therefore, network-layer administrative restrictions such as firewalls must be used to limit end-to-end connectivity in order to impose administrative controls; these restrictions also have the advantage that they automatically apply to *all* Internet services, not just a limited subset of them. Network-level and application-level restrictions can also be used in concert; for instance, an Internet telephony signaling server, if it decided to allow a call, could instruct a firewall to open up the appropriate ports to allow the media associated with the call to flow.

Verification Testing

Finally, the most direct way of ensuring correct operation of features is to test them directly. It is for third parties to establish services which automatically, at your request, place calls to you with various parameters or conditions enabled, to allow you to confirm explicitly that your features work the way you desire. As such providers gain more experience into the sorts of conditions that are likely to cause problems with services, they can expand their suites of testing tools to cover more esoteric interaction conditions. Thus, it should be possible to verify features and resolve their interactions in the real environment in which they are deployed, rather than attempting to analyze and categorize all possible consequences of a feature beforehand.

2.13 Interoperation with Circuit-Switched Telephony

It is clear that the 700 million or so telephones in the world will not be converted any time soon to packet telephones. (There are only about 200 million computers of all kinds in the world.) Dumb, cheap end systems will continue to be appropriate in many circumstances. The corner gas station is not likely to install a coin PC next to the car vac and the telephone in the Mojave desert [166]. Thus, interoperation between "classical" telephony and packet telephony is an important consideration. This can take place in various ways, as shown in Fig. **??**. We can distinguish the following cases, which can be combined as needed:

end-to-end packet: End systems such as network computers, dedicated "Internet phones" or PCs packetize audio; the packets are delivered to one or more similar end systems for playback.

tail-end hop off: Packet networks are used for long-haul voice transmission, while standard circuit-switched voice circuits connect the CPE (telephones) to the packet telephony gateways. This can be used both for individual voice circuits as well as for PBX interconnect. Tail-end hop off allows to bypass long-distance phone companies as well as to connect POTS (plain old telephony service) devices to packet audio end systems. There exist "Internet telephony set-top boxes" that connect to a cable modem via an Ethernet interface. This device has a subscriber loop circuit that allows to plug in the regular house phone line. Dial tone and ring current are generated within the device and make it appear to old-style analog phones that they are connected to the central office.

local packet delivery: Voice is generated by packet audio end systems, but carried as circuit-switched voice over leased or public facilities in the wide area. An example is the "packet PBX" connecting to the PSTN.

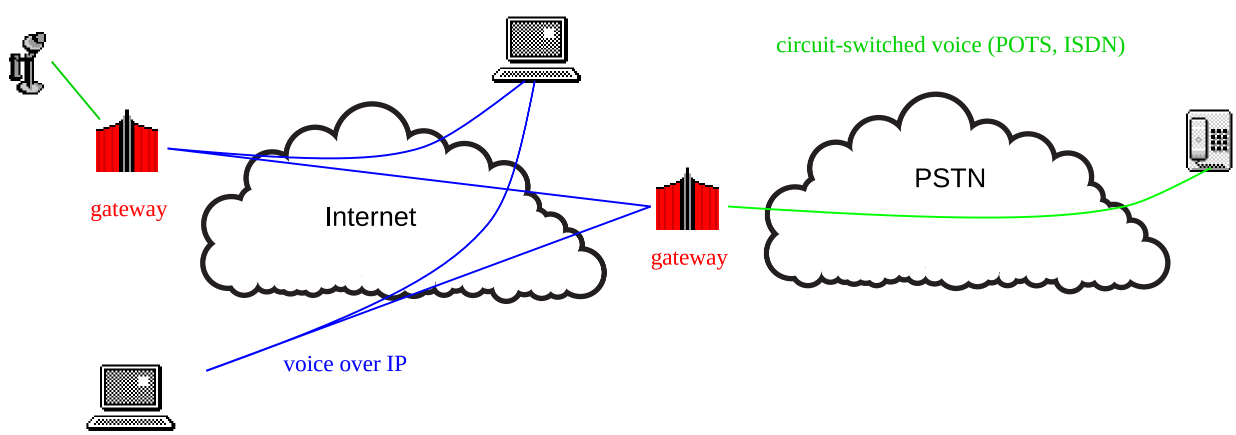

Figure 2.7: Possible interactions between Internet telephony and POTS

2.14 Fax over IP

There are two mechanisms for transmitting facsimile over the Internet: as an email attachment (store-and-forward) or in real-time. Facsimile is simply scanned text and images, usually scanned with a resolution of 100 dots per inch, compressed with run-length encoding and then transmitted with a special modem circuit at up to 9.6 kb/s [167]. (Higher-resolution fax, called G4, is available for ISDN, but seems to be little-used.)

Store-and-forward fax is simply an email attachment using MIME [168]. The fax image is transmitted in a special compressed image file format, the Tagged Image File Format (TIFF) [169].

Real-time fax creates the illusion as if the receiving fax machine is directly connected to the sending fax machine via a circuit. That is, fax pages are printed while the sending fax machine is still scanning pages. Fax calls can be set up just like Internet phone calls, using either SIP or H.323. Unfortunately, the ITU has designed a special-purpose protocol, T.38 [170], for carrying fax, even though it would have been fairly straighforward to create a fax payload type for the Real-Time Transport Protocol, the transport sublayer used for audio and video. The issues for real-time transmission for fax are fairly similar to those for voice, although end-to-end delay obviously matters less, as long as packet delay variations are compensated for. However, fax is more loss sensitive than packet voice, so that forward error correction is called for. The higher delay tolerance has led to Internet fax being a more widely used commercial service than Internet telephony. Many Internet telephony gateways also support real-time fax transmission.

A number of companies operate gateways that “print” email messages on regular fax machines. One of the first available email-to-fax services was set up as a cooperative venture [171, 172, 173], but these services are now run commercially, either charging or carrying sponsor advertisements. There are also now services that provide personal fax numbers to subscribers. Faxes received on these numbers are delivered via email.

2.15 Radio and Television

We mentioned earlier in this chapter that Internet telephony shares almost all protocols with multimedia distribution services like Internet radio and TV. The only major difference is that radio and TV programs are announced or listed on web pages, rather than starting communications by calling up somebody. However, it is not difficult to imagine a service where one calls up friends to invite them to watch a TV show together (Section 2.11). We only hope that your local TV station does not call up during sweep weeks....

Given the similarity, it is quite plausible that the same end systems serve as both telephones and radios. Thus, we briefly describe here what it would take to integrate radio services into an Internet environment. Compared to Internet telephony, Internet radio is already a popular service. According to a survey conducted in late 1999 by Edison Media Research for the Arbitron ratings service, only about four million listeners listen to the about 3,000 Internet radio stations each week [174]. Out of these, all but 240 simply simulcast their programming on the Internet. Currently, most of these are streamed either directly from the station’s web server or by specialized streaming media companies. Since this architecture delivers individual streams to each listener, it scales poorly to large Internet populations. In the long term, it appears likely that streams will be "staged" to servers based at ISP POPs or distributed via multicast.

The FCC reports that as of September 1999, 12,615 AM/FM radio stations (4,783 AM, 5,766 commercial FM, 2,066 educational FM) are broadcasting in the United States. In other countries, particularly of smaller area and with government-run or public stations, the number of radio stations is much smaller, for example, only about 100 major stations in Germany. If all such stations were to be made available via Internet multicast at FM quality (56 kb/s), this would take 680 Mb/s. (Servicable mono sound quality can be achieved with about 16 kb/s.) Since many of these stations broadcast identical programs, the likely number of channels transmitted nationwide is much lower. More realistically, offering the 31 channels of audio carried by the DirecTV satellite service would add only 1.7 Mb/s of Internet traffic, while the 45 FM channels in New York City would add 2.5 Mb/s.

Carrying radio services over the Internet allows improved directory services, easy addition of side information such as content labeling, bandwidth diversity for different kinds of programming as well as the reception of more diverse programming particularly in more sparsely populated areas. It also allows regional programming to be heard worldwide, similar to shortwave radio. However, unlike shortwave, its distribution does not depend on atmospheric conditions and its reception not hampered by steel-frame buildings.

Improving upon some of the recent European FM radio services such as Radio Data Service (RDS), Internet radio can offer much better auxiliary content labeling. Labeling makes it easy to have receivers assemble custom programs for listening at some later time ("time-shifting") [175, 176].

Similarly, given enough bandwidth, there are substantial advantages of distributing entertainment video, both broadcast and video-on-demand services, using Internet protocols, in particular MPEG over RTP [177]. This, in combination with an Internet-based stream control protocol such as RTSP [44, 178], appears to be both more flexible and simpler than using MPEG as a new network protocol and DSM-CC as a control protocol.

2.16 Problems and Experiments

- The United States has about 200 million people above the age of 12. Assume that each teenager and adult makes about an hour of long distance and an hour of local calls per day. Compute the total bandwidth needed to carry these conversations assuming that each phone call takes a bandwidth of 32 kb/s. Why is this computation not good enough to design a network?

- Experiment with an existing Internet telephony application. Can you measure how long it takes to set up a call? What is the delay? (Hint: You need two workstations in the same room for this experiment. You might find a stereo cassette recorder handy that can record audio from both the sending and receiving side. Alternatively, you can use a sound recorder tool on a third workstation to record both.)

- Find out (e.g., from manufacturer's web pages) how much power a typical PC, laptop and

Ethernet hub consume. A typical 12-Volt car battery has a capacity of 50 Ah or 600 Wh. In other words, this battery can sustain a load of 30 W for 20 hours or a load of 100 W for 6 hours. How long could you run the PC and other equipment on that battery? Find out how long typical office UPS (uninterruptible power supplies) last. What are the implications for Internet telephony?

- Can you suggest how E911 (enhanced emergency) service might work for Internet telephony? In this service, the dispatcher can see which address you are calling from. Depending on the address, a different 911 calling center will be contacted. What are the difficulties for dial-up users?

- Find out how many outside phone lines your workplace has and estimate how many minutes a day each employee makes or receives outside phone calls. In a typical PBX, about a third of calls are internal, with the remainder split evenly between incoming and outgoing external calls. A typical business line on a PBX is in use 10 minutes of the hour. This is referred to us a load of 6 CCS. CCS is a measure of telephony traffic representing hundred call seconds per hour. How much does your workplace differ?

- In many countries, you can choose a long-distance telephone provider or even use a different one for each phone call by dialing a prefix (10 and a five-digit number or 10-10 and a three-digit number in the United States). How could one implement a similar service for Internet telephony?

- Find out if your Internet service providers reveals or guarantee its reliability. If so, is reliability measured?

- Design an experiment that tests reachability from your workstation to a group of sites by periodically “ping”ing it or running `traceroute` to get more details. To reduce traffic, you should only run the test every few minutes. In your experiments, do you find that connectivity problems are local, in the middle of the network or at the destination?

Appendix A

Glossary

NEED TO FIND THE SOURCES FOR ALL OF THESE; SOME ARE SWEETWATER, I THINK.

16QCIF: High-resolution video format used by H.263 with 1408 by 1152 pixels.

4QCIF: Video format used by H.263 with 702 by 576 pixels.

4:1:1: A commonly used term for a component digital video format The numerals 4:1:1 denote the approximate ratio of the sampling frequencies of the single luminance channel to the two color (chrominance) channels. For every four luminance samples, there is one sample of each color channel.

4:2:2: A commonly used term for a component digital video format that is often used as an input format for image compression. The details of the format are specified in the CCIR (ITU-R) 601 standard document. The numerals 4:2:2 denote the approximate ratio of the sampling frequencies of the single luminance channel to the two color (chrominance) channels. For every four luminance samples, there are two samples of each color channel.

A/D, ADC: Analog-to-digital converter. Converts analog signals to their digital equivalent. $\longrightarrow$D/A.

ADSI: Analog Display Services Interface, a specification for advanced screen telephony.

ADSL: Asymmetric digital subscriber line. Variant of DSL where the downstream bandwidth (from central office to subscriber) is significantly larger than the upstream bandwidth.

AES: Audio Engineering Society.

AES/EBU: Informal name for a digital audio standard established jointly by the AES and EBU organizations. The sampling frequencies for this standard varies depending on the format being used; the sampling frequency for D1 and D2 audio tracks is 48 kHz.

AGC: Automatic gain control. Automatically adjust the amplifier gain according to the input signal amplitude.

Aliasing:

Analog: Any signal that varies continuously as opposed to a digital signal, which can only represent discrete levels.

Application sharing:

Audio bridge: In telecommunications, a device that mixes multiple audio inputs and feeds back composite (summed) audio to each stations, minus that station's input. Also known as a mix-minus audio system when used in radio call-in programs, where the caller receives the mix-minus signal.

Audio-follow-video (AFV): An operational mode in which audio and video sources are switched together so that when the operator selects the video source, the audio input simultaneously and automatically switches to the same source.

Balanced: A signal carried on two conductors, with symmetrical levels with respect to a common reference point, typically signal ground.

Bandwidth: The complete range of frequencies over which a circuit or electronic system can function with minimal signal loss, typically less than 3 dB. The information carrying capability of a particular television channel. In PAL systems the bandwidth limits the maximum visible frequency to 5.5 MHz, in NTSC, 4.2 MHz. The CCIR (TIU-R) 601 luminance channel sampling frequency of 13.5 MHz was chosen to permit faithful digital representation of the ⟶PAL and ⟶NTSC luminance bandwidths without ⟶aliasing.

BHCA: Busy Hour Call Attempts. Performance measure for the signaling and control parts of a switch. The busy hour is "the consecutive 60-minute interval with the highest levels of measurement or derived load used for traffic engineering." [8]

BNC: A coaxial connector primarily used for component video in studios and on some computer monitors as well as measurement instruments. A BNC connector is mechanically and electrically superior to a cinch connector.

BORSCHT: Battery, overvoltage protection, ringing, supervision, test.

Bus, buss: In audio and video, a central conductor for the primary signal path. A signal path to which a number of inputs may be connected for feed to one or more outputs. A mixer used for studio recordings often has 8 stereo buses.

CCS: Common Channel Signaling. Telephony signaling that uses a common channel, separate from those used for transmitting voice, for establishing and controlling telephone calls. The most common CCS protocol is ⟶SS7.

CCS: Hundred call seconds. Unit of telephone traffic per hour. 3600 CCS is the maximum utilization per telephone line.

CD: Compact disk. Disk of xx" diameter that holds about XXX Mb or 74 minutes of digital audio recorded at a sampling rate of 44.1 kHz and 16 bits resolution.

CIF: Common interchange format; one-half the height and width of CCIR601.

Chrominance: That portion of the video signal which contains the color information (hue and saturation). Video picture information contains two components: ⟶luminance (brightness and contrast) and chrominance (hue and saturation). Often abbreviated as C.

Clipping: The distortion of an audio and video signal once it reaches a certain threshold. Analog systems clip "soft", in that distortion increases beyond a certain signal volume, while digital systems clip "hard", in that they cannot represent signal values outside the quantization range.

Cinch connector: A connector used for consumer unbalanced line-level audio interfaces on tape decks and CD players and for composite video on VCRs. Also known as an RCA connector. It consists of a concentric ring for signal ground and a center pin for the signal.

Coaxial cable: A cable which has a metallic noise shield surrounding a signal-carrying conductor. In television, the cable impedance is 75 Ω. Coaxial cable for Ethernet has an impedance of 50 Ω. Coaxial cable can carry high-bandwidth signals such as video over distances of up to a kilometer. Often terminated with ⟶BNC connectors.

D1: A component digital videotape recording format that conforms to the specifications set in the CCIR (ITU-R) 601 standard. [179]

D2: An 8-bit composite digital videotape recording format in which the composite video signal is digitized by sampling it at the rate of four times the frequency of subcarrier (4fsc). The 4fsc frequency in NTSC is 14.3 MHz and 17.7 MHz in PAL. [179]

D3: An unofficial term for a composite digital videotape recording format invented by Panasonic. [179]

D connector: A type of connector that has a trapezoidal shell resembling a "D." Often used with 9 and 25 pins for serial data connections and with 15 pins for computer monitors.

D/A, DAC: Digital to analog converter.

DAT: Digital audio tape. Digital stereo tape recording at 44.1 and 48 kHz sampling rates.

dB: A relative measure of voltage or power, named for Alexander Graham Bell is a tenth of a bel. Given two voltages V_1 and V_2,. their ratio is expressed as $20 \log U_1/U_2$. Given two power lovels P_1 and P_2, their ratio can be expressed as $10 \log P_1/P_2$. A ratio of 3 dB corresponds to a doubling of power, while 1 dB is the smallest humanly detectable change in volume. Also used as a measure of loudness, relative to the softest audible sound. A loudness of 74 dB is commonly used as a reference for the level produced by a speaker at a distance of 1 ft, while a microphone at 1 in distance to the speaker's mouth receives a sound pressure level (SPL) of 94 dB.

dBA: A measure of power with a frequency weighting called the A filter.

dBm: A measure of power relative to 1 mW. 0 dBm equals 1 mW.

dbm0: Power relative to 1 mW, at the transmission level point (TLP).

dBrnC: A measure of relative noise power, with C-message weighting, expressed in decibels. It is measured relative to the reference noise of -90 dBm = $10^{-12}W$.

dBu: A mesure of voltage relative to 0.775 V. Subtract 2.2 dB to convert from dBu to dBV.

dBV: Voltage referenced to 1 Volt.

DDD: Direct distance dialing; long-distance calls dialed by a subscriber without operator assistance.

DCS: Digital cross connect. Connects telecommunications circuits, such as Sonet, T3 and T1 lines, with each other, so that the network can be reconfigured without changing physical connections. Changes are made electronically, but at timescales of minutes to days rather than on a call-by-call basis.

DID: Direct inward dialing; provides a regular phone number for users of a $\longrightarrow$PBX, with the dialed number conveyed to the PBX.

Digital milliwatt: Level of a digital signal that produces 0 dbm0 of analog voltage in a telephone system, generated by a digital sine wave with maximum amplitude of 5768. (The digital levels are set such that the maximum power of a digital signal is 3 dbm0 or 2 mW.)

Distribution amplifier (DA): A device that replicates an audio signal to several outputs, preventing the loading by the input impedances to effect signal amplitude or frequency characteristics. Often used to distribute audio or video across several amplifiers, tape decks and studio monitors.

DSL: Digital subscriber line. Uses existing local telephone loops to deliver high-speed data access. $\longrightarrow$ADSL

DTMF: Dual-Tone Multiple Frequency. Subscriber signaling scheme using two-tone signals for dialing and invoking services. (Section 1.3.3)

DVD: Digital video disk. Uses a plastic disk of the same size as a $\longrightarrow$CD, but records XXX.

flow: distinguishable set of packets, e.g., all packets with the same destination address. Typically identified by a subset of the five-tuple protocol (UDP, TCP, ...), source address, source port, destination address, destination port, possibly also by the IPv4 TOS byte.

Foley: Sound effects.

(Video) Frame: A complete video picture composed of two fields (two complete interlaced scans of the monitor screen). A frame consists of 525 interlaced horizontal lines of picture information in NTSC, 625 in PAL.

FXO: Foreign exchange office. Analog telephone interface that looks like a telephone to the connected device. For example, PBXs may use an FXO interface to connect to the central office.

FXS: Foreign exchange station. Analog telephone interface that provides ringing and loop current to an attached telephone.

G.711: Audio coding that encodes each audio sample as 8 bits and produces a 64 kb/s bit stream at a sampling rate of 8 kHz. G.711 maps logarithmically between a 16-bit input sample and the 8-bit code word, trading dynamic range for resolution at higher volumes. It is the most common audio coding in the digital $\longrightarrow$PSTN. Two different quantization curves are used: μ-law in North America and Japan, A-law elsewhere.

G.723.1:

G.726:

G.729:

Gen-lock: To phase-lock the frame timing of one piece of video equipment to another. Allows mixing of different video signals and smooth transitions from one source to the next.

Grey8: Uncompressed image format where each pixel is represented by one byte of data, with 256 shades of grey.

GSTN: Global Switched Telephone Network. The interconnected system of generally reachable telephones, both wired and wireless. Also known as $\longrightarrow$PSTN. (The term GSTN was supposed to replace the monopoly-era term PSTN, but does not seem to have caught on.)

H.323:

H.324:

HDTV: High-definition television.

Horizontal resolution: Chrominance and luminance resolution (detail) expressed horizontally across a picture tube. This is usually expressed as a number of black to white transitions or lines that can be differentiated. Limited by the bandwidth of the video signal or equipment. [179]

ILEC: Incumbent Local Exchange Carrier. One of the original monopoly local telephone companies ("Baby Bells").

Interlaced: . Opposite: progressive scan.

ISUP: ISDN User Part; part of the Signaling System #7 telephone signalling protocol suite.

IXC: Interexchange carrier, i.e., long-distance telephone service provider, as opposed to $\longrightarrow$LEC.

Key: Also called key source or key cut. A signal that can be used to electronically "cut a hole" in a video picture to allow for insertion of other elements such as text or another video image. The key signal is a switching or gating signal for controlling a video mixer which switches or mixes between the background video and the inserted element. [179] Chroma-keying defines the key region by a particular color, typically blue, in the primary video signal.

Landline: Part of the telephone system that uses fibers, coax and twisted pairs, as opposed to wireless transmission.

LEC: Local Exchange Carrier; telephone operating company that provides local telephone service; ⟶ILEC, CLEC, IXC.

Line level: A voltage level for connecting tape decks, CD players, microphone mixers and preamplifiers to power amplifiers. Typically around 1 V, with consumer equipment using -10 dBV (unbalanced, 0.316 V) on Cinch connectors and professional equipment +4dBm (balanced, 1.23 V) on ⟶XLR connectors or ⟶TRS connectors. Often symbolized by a circle with an arrow pointing in (input) or out (output).

LMDS: Local multipoint distribution system (LMDS) is a broadband fixed wireless technology that can be used to deliver analog video or digital data. It uses spectrum at 25 GHz or higher.

LNP: Local number portability; the ability to maintain one's telephone number when changing providers and possibly locations.

Luminance: The measurable, luminous intensity of a video signal. Differentiated from brightness in that the latter is nonrneasurable and sensory. The color video picture information contains two components: luminance (brightness and contrast) and chrominance (hue and saturation). The photometric quantity of light radiation. [179] Often abbreviated as Y.

MC (multipoint controller):

MCU (multipoint control unit):

M/E: Mix/effects. A subsystem of a video production switcher where a composite of two or more images can be created. Each M/E typically includes crosspoint buses, keyer(s), and mixer.

Mic(rophone) Level: The level (or voltage) of signal generated by a microphone. A typical dynamic microphone generates about 0.2 mV (-75 dBV) at a sound pressure level of 74 dB, while a condenser microphone generates about between 1 and 3.1 mV (-60 to -50 dBV).

MII Format: A second-generation component video format invented by Panasonic for use in videotape recorders. The signal set consists of separate Y, scaled R-Y, and scaled B-Y signals. The M refers to the way in which the tape is routed through the recording mechanism.

Mix-minus: ⟶audio bridge.

MMDS: Multichannel Multipoint Distribution Service. Fixed wireless access service operating in the 2.1 to 2.7 GHz frequency band.

Monochrome: Black and white video. A video signal that represents the brightness values (luminance) in the picture, but not the color values (chrominance).

MTP: Message Transfer Part. Physical, link and network layer of Signaling System #7. ⟶ISUP.

NTSC: Organization that formulated standards for the NTSC television system. Now describes the American system of color telecasting which is used mainly in North America, Japan, and parts of South America. NTSC television uses a 3.579545 MHz subcarrier whose phase varies with the instantaneous hue of the televised color and whose amplitude varies with the instantaneous saturation of the color. NTSC employs 525 lines per frame and 59.94 fields per second. [179]

Numbering Plan Area (NPA): "A specific geographical area identified by a unique NPA code. The boundaries of an NPA code are usually within a state, province or subdivision of another country within the North American Numbering Plan." [8]

outbound proxy: SIP proxy that handles calls originating from local users.

PAL: (phase alternate line) The name of the color television system in which the E'v component of burst is inverted in phase from one line to the next in order to minimize hue errors that may occur in color transmission. PAL-B (also called PAL-I) is a European color TV system featuring 625 lines per frame, 50 fields per second, and a 4.43361875 MHz subcarrier. It is used mainly in Europe, China, Malaysia, Australia, New Zealand, the Middle East, and parts of Africa. PAL-M is a Brazilian color TV system with phase alternation by line, but using 525 lines per frame, 60 fields per second, and a 3.57671149 MHz subcarrier. ⟶NTSC Most modern TV receivers can receive both formats, but in black-and-white only; VCRs can typically record and play only one or the other.

PANS: Pretty amazing new services. Usually used jokingly for fancy telephony services.

PBX: Private Branch Exchange. Privately-owned telephone switch that serves a business or other private organization, although it often uses the same technology as switches owned by the telephone company.

PCS: Personal communications services; name for second-generation digital cellular systems operating in the 900 MHz and 1.8 GHz bands.

Peak meter: Amplitude measurement device that displays the peak excursion of an audio signal. Used to prevent ⟶clipping in digital recording.

Phono connector: ⟶Cinch connector; not to be confused with a *phone* connector (⟶TRS connector)

Pixel: The smallest distinguishable and resolvable area in a video image. A single point on the screen. In digital video, a single sample of the picture. Derived from the words picture element. Also known as pel.

POTS: Plain old telephony service. Typically analog basic phone services; as opposed to ISDN or intelligent network services [179]. ⟶PANS

Progressive scan: All lines in a frame are scanned consecutively. Typically used for computer displays, as it gives better quality for (moving) text.

PSTN: Public Switched Telephone Network. The interconnected system of generally reachable telephones, both wired and wireless. Also known as ⟶GSTN.

Q.931: ISDN user-to-network signaling protocol.

QCIF (Quarter CIF) : 176x144 pixels[1]; common H.261 and H.263 video resolution. ⟶CIF.

RBOC: Regional Bell Operating Company. AT&T was split in 1984 into seven RBOCS ("baby bells"): NYNEX, Bell Atlantic, Ameritech, BellSouth, Southwestern Bell Telephone, US West and Pacific Bell. Some of these have since merged.

RCA connector: ⟶Cinch connector.

RGB: Component video that separates the red, green and blue color components. Sometimes found on high-end computer monitors and video projectors. Typically terminated with ⟶BNC connectors.

RGsB: Like RGB, but the green color channel carries the synchronization signals.

RGB15: Color mode where each image pixel is represented by two bytes (16 bits) of data, with 5 bits each for red, green and blue; the most significant bit is unused. This color mode is also referred to as "high color" or 5:5:5.

RGB24: Color mode where each image pixel is represented by three bytes, one each for red, green and blue, and otherwise has the same properties as RGB32.

RGB32: Color mode where each image pixel is represented by four bytes, one each for red, green, blue, plus one byte of padding. Each pixel can thus assume 256 shades of each color, for a total of 16.7 million colors. This color format is also referred to as "true color".

RS-170A: A document prepared by the Electronics Industries Association (EIA) describing recommended practices for NTSC color television signals in the United States.

RTCP (Real-Time Control Protocol): See Ch. **??**.

[1]The exact pixel dimensions may vary between applications.

RTP: Real-Time Transport Protocol; see Ch. **??**. Also: "route to pivot", a number portability mechanism.

RTSP: Real-Time Stream Protocol. See Ch. **??**.

Scan converter: Device that converts the signal of ⟶VGA monitors to a signal suitable for analog (NTSC or PAL) video. Often used to record the output of computer monitors on video tape. This conversion is necessary since video monitors use a different horizontal synchronization frequency (32XXX kHz as compared to 16XXX for NTSC and PAL) as well as frame rate. However, due to the poor quality, it is usually preferable to capture screen windows digitally or use *application sharing*.

SECAM: Sequential couleur avec memoire (sequential color with memory). A color television system with 625 lines per frame and 50 fields per second, developed by France and the USSR. Color difference information is transmitted sequentially on alternate lines as an FM signal.

S-VHS: Analog video tape format that offers a resolution of approximately 400 TV lines and stereo audio.

S-Video: Consumer component video format. Typically found on S-VHS VCRs and camcorders. An S-Video connector, typically a small round 4-pin DIN-type, carries chrominance and luminance signals.

SCN: Switched circuit network; used to describe the telephone network.

SIF (standard interchange format) : 320 x 240 pixels; one quarter the size of a full ⟶NTSC frame.

SIP: Session Initiation Protocol, an Internet-standard protocol for setting up conferences and phone calls in the Internet [180].

SS7: Signaling System #7, the most common example of a ⟶CCS protocol.

SMPTE time code: Time code that conforms to SMPTE standards. It consists of an eight-digit number specifying hours:minutes:seconds:frames. Each number identifies one frame on a videotape. SMPTE time code may be of either the drop-frame or non-drop frame type. In GVG editors, the SMPTE time code mode enables the editor to read either drop-frame or non-drop frame code from tape and perform calculations for either type (also called mixed time code). XXX [Explain different types.]

SNR: Signal-to-noise ratio. The SNR relates how much stronger a signal is than the background noise. Usually expressed in ⟶decibels (dB).

Subcarrier (SC) : In NTSC or PAL video, a continuous sine wave of extremely accurate frequency which constitutes a portion of the video signal. The subcarrier is phase modulated to carry picture hue information and amplitude modulated to carry color saturation information.

The NTSC subcarrier frequency is 3.579545 MHz, and the PAL-I frequency is 4.43361875 MHz. A sample of the subcarrier, called color burst, is included in the video signal during horizontal blanking. Color burst serves as a phase reference against which the modulated subcarrier is compared in order to decode the color information. [179]

sub-QCIF: Video format for H.263 (required) and H.261 (optional) with 128 by 96 pixels.

SVGA:

Sync: The portion of an encoded video signal that occurs during blanking and is used to synchronize the operation of cameras, monitors, and other equipment. Horizontal sync occurs within the blanking period in each horizontal scanning line, and vertical sync occurs within the vertical blanking period. [179]

Tail End Hop Off (TEHO): In a private network, a call which is carried over flat rate facilities (Intermachine Trunks or IMT) to the closest switch node to the destination of the call, and then connected into the public network as a local call. (itmedia)

TCP: Transmission control protocol. Transport protocol in the Internet protocol stack that offers a reliable, sequenced, flow- and congestion-controlled byte stream.

Time code: Timing code laid down on videotape to give each frame a unique number so as to ensure exact transitions during editing.

Tromboning: Sending traffic which comes from a fixed station and is destined for a mobile network in the same country via a second country to take advantage of beneficial accounting rates for termination of international traffic on mobile networks. (Oftel) Also: doubling back of calls along the same route.

TRS connector: Tip-ring-sleeve connector, also known as a phone connector, since it was used on telephone switchboards. Found in diameters of 1/4” (6.3 mm) for stereo headphones and professional balanced audio, 1/8” (3 mm) on computer sound cards (unbalanced) and personal stereo headphones. For balanced connectors, the tip carries the X, the sleeve X and the ring signal ground. For unbalanced connectors, the tip connects the left channel, the ring the right channel.

Twisted pair: A cable composed of two small insulated conductor twisted together. Since both wires have nearly equal exposure to any interference, the differential noise is slight. [179] Used in so-called category-5 “unshielded twisted pair” (UTP) cable to carry telecommunications signals of up to 150 Mb/s.

UTP: Unshielded twisted pair. Wiring used for telephone and data connections. “Category-3” is standard commercial telephone cable and supports bandwidths of XX MHz, sufficient for 10 Mb/s Ethernet (10BaseT), while “category-5” (CAT5) cabling has a bandwidth of 100 MHz and can support 100 Mb/s Ethernet (100BaseT).

XLR connector: 3-pin round connector used in professional audio equipment for connecting microphones to mixers and mixers to amplifiers. Two pins (X, X) carry the balanced audio signal, the center pin (pin X) is signal ground.

Vertical sync pulse: The synchronizing pulse at the end of each field which signals the start of vertical retrace. G

Video monitor: A high-quality television set (without RF circuits) that accepts video baseband inputs directly from a TV camera, videotape recorder, etc.

Video switcher (production switcher, video mixer): Device that accepts inputs from a variety of video sources and allows the operator to select a particular source to be sent to the switcher's output(s). May also include circuits for video mixing, wiping, keying, and other special effects.

VGA:

Vertical sync pulse: The synchronizing pulse at the end of each field which signals the start of vertical retrace.

VU meter: Volume-unit meter, a type of meter used to indicate average audio amplitude [181]. Generally better suited for analog recording; for digital recording a peak meter is preferred.

Watt: A measure of electrical power. The power expended when one ampere of current flows through a resistance of one ohm. The unit of electric power required to do work at the rate of 1 joule per second. Calculated by multiplying voltage (volts) times current (amperes).

XGA:

Y: $\longrightarrow$luminance. Y is derived from the red, green, and blue signals as 0.30 R + 0.59 G + 0.11 B.

Y, R-Y, B-Y: Color difference signal designation. Y corresponds to the luminance signal, R-Y corresponds to the red minus luminance signal, and B-Y corresponds to the blue minus luminance signal. These signals are derived as follows:

$$\begin{aligned} Y &= 0.3R + 0.59G + 0.1B \\ R - Y &= 0.7R + 0.59G - 0.1B \\ B - Y &= 0.89B - 0.59G - 0.3R \end{aligned}$$

Y, U, V: PAL luminance and color difference components. U and V are the names of the B-Y and R-Y color differences signals (respectively) when they are modulated onto subcarrier. Often also used for the digital components derived from both NTSC or PAL signals.

A.1 Related Resources

http://www.sweetwater.com/insync/wftd/wftd-list.html

Bibliography

[1] I. de Sola Pool, ed., *The social impact of the telephone.* Cambridge, Massachusetts: MIT Press, 1977.

[2] A. Friedlander, *Natural Monopoly and Universal Service: Telephones and Telegraphs in the U.S. Communications Infrastructure, 1837-1940*. Washington, DC: CNRI, 1995.

[3] R. F. Rey, ed., *Engineering and Operations in the Bell System*. Murray Hill, New Jersey: AT&T Bell Laboratories, 1984.

[4] J. Bellamy, *Digital Telephony*. New York: John Wiley & Sons, 1991.

[5] International Telecommunication Union, "Echo suppressors," Recommendation G.164, Telecommunication Standardization Sector of ITU, Geneva, Switzerland, Nov. 1988.

[6] International Telecommunication Union, "Echo cancellers," Recommendation G.165, Telecommunication Standardization Sector of ITU, Geneva, Switzerland, Mar. 1993.

[7] S. S. Haykin, *Introduction to Adaptive Filters*. New York: Macmillan Publishing Co., 1984.

[8] Bellcore, *Notes on the Networks*. Piscataway, New Jersey: Bellcore, 1997.

[9] International Telecommunication Union, "Technical characteristics of tones for telephone service," Recommendation Supplement 2 to Recommendation E.180, Telecommunication Standardization Sector of ITU, Geneva, Switzerland, Jan. 1994.

[10] International Telecommunication Union, "Application of tones and recorded announcements in telephone services," Recommendation E.182, Telecommunication Standardization Sector of ITU, Geneva, Switzerland, Mar. 1998.

[11] International Telecommunication Union, "Various tones used in national networks," Recommendation Supplement 2 to Recommendation E.180, Telecommunication Standardization Sector of ITU, Geneva, Switzerland, Jan. 1994.

[12] N. S. Jayant and P. Noll, *Digital Coding of Waveforms—Principles and Applications to Speech and Video*. Englewood Cliffs, New Jersey: Prentice-Hall, 1984.

[13] International Telecommunication Union, “Pulse code modulation (pcm) of voice frequencies,” Recommendation G.711, Telecommunication Standardization Sector of ITU, Geneva, Switzerland, Nov. 1998.

[14] C. Partridge, *Gigabit networking*. Reading, Massachusetts: Addison-Wesley, 1993.

[15] International Telecommunication Union, “Sub stm-0 network node interface for the synchronous digital hierarchy (sdh),” Recommendation G.708, Telecommunication Standardization Sector of ITU, Geneva, Switzerland, June 1999.

[16] Common Carrier Bureau, “Trends in telephone service,” tech. rep., Federal Communications Commission, Washington, D.C., Sept. 1999.

[17] International Telecommunication Union, “The international public telecommunication numbering plan,” Recommendation E.164, Telecommunication Standardization Sector of ITU, Geneva, Switzerland, May 1997.

[18] International Telecommunication Union, “Notation for national and international phone numbers,” Recommendation E.123, Telecommunication Standardization Sector of ITU, Geneva, Switzerland, 1998.

[19] W. H. Nunn, “Nationwide numbering plan,” *Bell System Technical Journal*, Vol. 31, pp. 851–859, Sept. 1952.

[20] G. Nilsson, “Number portability: A networking perspective,” *Telecommunications*, July 1997.

[21] R. F. Albers, “Annual report 1998,” report, Network Reliability Steering Committee, Alliance for Telecommunications Industry Solutions, Washington, DC, Jan. 1999.

[22] R. F. Albers, “Macro-analysis: 1st quarter 1999,” report, Alliance for Telecommunications Industry Solutions, Network Reliability Steering Committee, Washington, DC, June 1999.

[23] A. P. Snow, “Internet implications of telephone access,” *IEEE Computer*, Vol. 32, pp. 108–110, Sept. 1999.

[24] A. P. Snow, “A reliability analysis of local telecommunications switches,” in *Proc. of Sixth International Conference Telecommunications Modeling and Analysis*, pp. 288–295, 1998.

[25] AT&T, “AT&T sets the industry standard for network reliability,” Mar. 1998. http://www.att.com/network/standrd.html.

[26] D. Lazarus, “AOL’s email system crashes twice in one day,” *Wired News*, Jan. 1997.

[27] R. Kawamura, “Architectures for ATM network survivability,” *IEEE Communications Surveys*, Vol. 1, 4th Quarter 1998.

[28] A. Felekis and D. Milis, “Emerging technologies for fibre network survivability,” *Surveys and Presentations in Information Systems Engineering (SURPRISE)*, Vol. 4, 1996.

[29] T.-H. Wu, “Emerging technologies for fiber network survivability,” *IEEE Communications Magazine*, Vol. 33, pp. 60–75, Feb. 1995.

[30] Cisco Systems, “Packet-over-sonet/sdh,” application note, Cisco, 2000.

[31] R. Droms, “Dynamic host configuration protocol,” Request for Comments (Proposed Standard) 1541, Internet Engineering Task Force, Oct. 1993.

[32] S. Thomson and T. Narten, “IPv6 stateless address autoconfiguration,” Request for Comments (Proposed Standard) 1971, Internet Engineering Task Force, Aug. 1996.

[33] International Telecommunication Union (ITU), “Transmission systems and media, general recommendation on the transmission quality for an entire international telephone connection; one-way transmission time,” Recommendation G.114, Telecommunication Standardization Sector of ITU, Geneva, Switzerland, Mar. 1993.

[34] J. W. Emling and D. Mitchell, “Effects of time delay and echoes on telephone conversations,” *Bell System Technical Journal*, Vol. 42, pp. 2869–2891, Nov. 1963.

[35] G. K. Helder, “Customer evaluation of telephone circuits with delay,” *Bell System Technical Journal*, Vol. 45, pp. 1157–1191, Sept. 1966.

[36] A. G. Vartabedian, “The effects of transmission delay in four-wire teleconferencing,” *Bell System Technical Journal*, Vol. 45, pp. 1673–1688, Dec. 1966.

[37] P. T. Brady, “Effects of transmission delay on conversational behavior on echo-free telephone circuits,” *Bell System Technical Journal*, Vol. 50, pp. 115–134, Jan. 1971.

[38] N. Kitawaki, T. Kurita, and K. Itoh, “Effects of delay on speech quality,” *NTT Review*, Vol. 3, pp. 88–94, Sept. 1991.

[39] E. T. Klemmer, “Subjective evaluation of transmission delay in telephone conversation,” *Bell System Technical Journal*, Vol. 46, pp. 1141–1147, July–August 1967.

[40] International Telecommunication Union, “Digital subscriber signalling system no. 1 (dss 1) - isdn user-network interface layer 3 specification for basic call control,” Recommendation Q.931, Telecommunication Standardization Sector of ITU, Geneva, Switzerland, Mar. 1993.

[41] International Telecommunication Union, “Application of the ISDN user part of ccitt signallling system no. 7 for international ISDN interconnections,” Recommendation Q.767, Telecommunication Standardization Sector of ITU, Geneva, Switzerland, 1991.

[42] International Telecommunication Union, “General recommendations on telephone switching and signaling – intelligent network: Introduction to intelligent network capability set 1,” Recommendation Q.1211, Telecommunication Standardization Sector of ITU, Geneva, Switzerland, Mar. 1993.

[43] International Telecommunication Union, “Introduction to intelligent network capability set 2,” Recommendation Q.1221, Telecommunication Standardization Sector of ITU, Geneva, Switzerland, Sept. 1997.

[44] H. Schulzrinne, A. Rao, and R. Lanphier, “Real time streaming protocol (RTSP),” Request for Comments (Proposed Standard) 2326, Internet Engineering Task Force, Apr. 1998.

[45] J. Rosenberg, H. Schulzrinne, E. Guttman, and R. Moats, “WASRV architectural principles,” Internet Draft, Internet Engineering Task Force, Feb. 1998. Work in progress.

[46] L. Dusseault and M. Calsyn, “Presence information protocol requirements,” Internet Draft, Internet Engineering Task Force, Feb. 1998. Work in progress.

[47] P. Ferguson and G. Huston, “What is a VPN?,” white paper, Cisco Systems; Telstra Internet, Mar. 1998.

[48] H. Schulzrinne, S. Casner, R. Frederick, and V. Jacobson, “RTP: a transport protocol for real-time applications,” Request for Comments (Proposed Standard) 1889, Internet Engineering Task Force, Jan. 1996.

[49] J. Crollick and J. Combs, “9-1-1 and enterprise IP telephone systems,” TIA TR-41.1 contribution TR41.3.4/9908-11, SCC Communications, Temple Terrace, Florida, Aug. 1999.

[50] C. Drane, M. Macnaughtan, and C. Scott, “Positioning GSM telephones,” *IEEE Communications Magazine*, Vol. 36, pp. 46–59, Apr. 1998.

[51] S. Tekinay, E. Chao, and R. Richton, “Performance benchmarking for wireless location systems,” *IEEE Communications Magazine*, Vol. 36, pp. 72–76, Apr. 1998.

[52] J. H. Reed, K. J. Krizman, B. D. Woerner, and T. S. Rappaport, “An overview of the challenges and progress in meeting the E-911 requirement for location service,” *IEEE Communications Magazine*, Vol. 36, pp. 30–37, Apr. 1998.

[53] Common Carrier Bureau, “Trends in telephone service,” tech. rep., Federal Communications Commission, Washington, D.C., Feb. 1999.

[54] T. H. Sloane, M. Miller, and D. Sorenson, “Fiber finances,” *Telephony*, Vol. 238, pp. PTN16–PTN20, Jan. 2000.

[55] R. King, “Big pipes, bigger problems,” *Tele.com*, Nov. 1997.

[56] A. P. Conversion, “Effect of ups on system availability,” Technical Note T26, American Power Conversion, 1998.

[57] R. Callon, "TCP and UDP with bigger addresses (TUBA), a simple proposal for internet addressing and routing," Request for Comments (Informational) 1347, Internet Engineering Task Force, June 1992.

[58] C. Hornig, "Standard for the transmission of IP datagrams over ethernet networks," Request for Comments (Standard) 894, Internet Engineering Task Force, Apr. 1984.

[59] J. Postel, "Standard for the transmission of IP datagrams over experimental ethernet networks," Request for Comments (Standard) 895, Internet Engineering Task Force, Apr. 1984.

[60] J. Heinanen, "Multiprotocol encapsulation over ATM adaptation layer 5," Request for Comments (Proposed Standard) 1483, Internet Engineering Task Force, July 1993.

[61] M. Laubach and J. Halpern, "Classical IP and ARP over ATM," Request for Comments (Proposed Standard) 2225, Internet Engineering Task Force, Apr. 1998.

[62] D. Waitzman, "Standard for the transmission of IP datagrams on avian carriers," Request for Comments (Experimental) 1149, Internet Engineering Task Force, Apr. 1990.

[63] D. Waitzman, "IP over avian carriers with quality of service," Request for Comments (Informational) 2549, Internet Engineering Task Force, Apr. 1999.

[64] S. Bradner, "The internet standards process – revision 3," Request for Comments (Best Current Practice) 2026, Internet Engineering Task Force, Oct. 1996.

[65] P. Heywood, M. Jander, E. Roberts, and S. Saunders, "Standards: The inside story," *Data Communications*, Mar. 1997.

[66] J. G. van Bosse, *Signaling in Telecommunications Networks*. Telecommunications and Signal Processing, New York, New York: Wiley, 1998.

[67] G. C. Kessler and P. V. Southwick, *ISDN: concepts, facilities and services*. New York, New York: McGraw-Hill, 1997.

[68] I. Faynberg, L. R. Gabuzda, M. P. Kaplan, and N. J. Shah, *Intelligent Network Standards: their Application to Services*. New York: McGraw-Hill, 1997.

[69] J. Veizades, E. Guttman, C. Perkins, and S. Kaplan, "Service location protocol," Request for Comments (Proposed Standard) 2165, Internet Engineering Task Force, June 1997.

[70] C. A. Polyzois, K. H. Purdy, P.-F. Yang, D. Shrader, H. Sinnreich, F. Mnard, and H. Schulzrinne, "From POTS to PANS – a commentary on the evolution to internet telephony," *IEEE Network*, Vol. 13, pp. 58–64, May/June 1999.

[71] K. Sollins, "The TFTP protocol (revision 2)," Request for Comments (Standard) 1350, Internet Engineering Task Force, July 1992.

[72] D. Cohen, "On packet speech communication," in *Proceedings of the Fifth International Conference on Computer Communications*, (Atlanta, Georgia), pp. 271–274, IEEE, Oct. 1980.

[73] D. Cohen, "Specifications for the network voice protocol," Technical Report ISI/RR-75-39 (AD A023506), USC/Information Sciences Institute, Marina del Rey, California, Mar. 1976. Available from DTIC).

[74] D. Cohen, "Issues in transnet packetized voice communications," in *Proceedings of the Fifth Data Communications Symposium*, (Snowbird, Utah), pp. 6–10 – 6–13, ACM, IEEE, Sept. 1977.

[75] D. Cohen, "A protocol for packet-switching voice communication," in *Proc. of Computer Network Protocols Symposium*, (Liege, Belgium), Feb. 1978.

[76] D. Cohen, "Specifications for the network voice protocol (NVP)," Request for Comments 741, Internet Engineering Task Force, Nov. 1977.

[77] J. W. Forgie, "Voice conferencing in packet networks," in *Conference Record of the International Conference on Communications (ICC)*, (Seatle, Washington), pp. 21.3.1–21.3.4, IEEE, June 1980.

[78] D. T. Magill, "Adaptive speech compression for packet communication systems," in *Conference record of the IEEE National Telecommunications Conference*, pp. 29D–1 – 29D–5, 1973.

[79] S. A. Mahmoud, W.-Y. Chan, J. S. Riordon, and S. E. Aidarous, "An integrated Voice/Data system for VHF/UHF mobile radio," *IEEE Journal on Selected Areas in Communications*, Vol. SAC-1, pp. 1098–1111, Dec. 1983.

[80] N. Shacham, E. J. Craighill, and A. A. Poggio, "Speech transport in packet-radio networks with mobile nodes," *IEEE Journal on Selected Areas in Communications*, Vol. SAC-1, pp. 1084–1097, Dec. 1983.

[81] Anonymous, "Special issue on packet switched voice and data communication," *IEEE Journal on Selected Areas in Communications*, Vol. SAC-1, Dec. 1983.

[82] G. Falk, S. J. Groff, W. C. Milliken, M. Nodine, S. Blumenthal, and W. Edmond, "Integration of voice and data in the wideband packet satellite network," *IEEE Journal on Selected Areas in Communications*, Vol. SAC-1, pp. 1076–1083, Dec. 1983.

[83] R. Cole, "PVP - a packet video protocol," W-Note 28, Information Sciences Institute, University of Southern California, Los Angeles, California, Aug. 1981.

[84] I. H. Merritt, "Providing telephone line access to a packet voice network," Research Report ISI/RR-83-107, Information Sciences Institute (ISI), Marina del Ray, California, Feb. 1983.

[85] D. C. Swinehart, L. C. Stewart, and S. M. Ornstein, "Adding voice to an office computer network," in *Proceedings of the IEEE Conference on Global Communications (GLOBECOM)*, (San Diego, California), pp. 392–398 (11.4), IEEE, Nov. 1983. also Xerox Report CSL-83-8.

[86] D. C. Swinehart, "Telephone management in the Etherphone system," in *Proceedings of the IEEE Conference on Global Communications (GLOBECOM)*, (Tokyo, Japan), pp. 1176–1180, IEEE, Nov. 1987.

[87] P. T. Zellweger, D. B. Terry, and D. C. Swinehart, "An overview of the Etherphone system and its applications," in *Proc. of 2nd IEEE Conference on Computer Workstations*, (Santa Clara, California), pp. 160–168, Mar. 1988.

[88] D. B. Terry and D. C. Swinehart, "Managing stored voice in the Etherphone system," *ACM Transactions on Computer Systems*, Vol. 6, pp. 3–27, Feb. 1988.

[89] D. C. Swinehart, "System support requirements for multi-media workstations," in *Proc. of the SpeechTech '88 Conference*, (New York, New York), pp. 82–83, Apr. 1988.

[90] P. T. Zellweger, "Active paths through multimedia documents," in *Document Manipulation and Typography – Proceedings of the EP'88 Conference on Electronic Publishing, Document Manipulation and Typography* (J. C. van Vliet, ed.), (Nice, France), Cambridge University Press, Apr. 1988.

[91] J. DeTreville and D. W. Sincoskie, "A distributed experimental communications system," *IEEE Journal on Selected Areas in Communications*, Vol. SAC-1, pp. 1070–1075, Dec. 1983.

[92] E. J. Donofrio, "The experimental communication system for voice/data on the Ethernet," in *Proceedings of the IEEE Conference on Global Communications (GLOBECOM)*, Vol. 2, (New Orleans, Louisiana), pp. 848–851 (28.2), IEEE, Dec. 1985.

[93] J. F. Shoch, "Carrying voice traffic through an Ethernet local area network – a general overview," in *Proc. of the IFIP WG 6.4 International Workshop on Local-Area Computer Networks*, (Zurich, Switzerland), Aug. 1980.

[94] P. Ravasio, R. Marcogliese, and R. Novarese, "Voice transmission over an Ethernet backbone," in *Local Computer Networks: Proceedings of the IFIP TC 6 International In-Depth Symposium on Local Computer Networks* (P. Ravasio, G. Hopkins, and N. Naffah, eds.), (Florence, Italy), pp. 39–65, IFIP, North Holland, Apr. 1982.

[95] T. A. Gonsalves, "Packet-voice communication on an Ethernet local computer network: an experimental study," Technical Report CSL-82-5, Xerox Palo Alto Research Center (PARC), Palo Alto, California, Mar. 1982.

[96] F. Tobagi and N. Gonzalez-Cawley, "On CSMA-CD local networks and voice communication," in *Proceedings of the Conference on Computer Communications (IEEE Infocom)*, (Las Vegas, Nevada), pp. 122–127, IEEE, March/April 1982.

[97] G. J. Nutt and D. L. Bayer, "Performance of CSMA/CD networks under combined voice and data loads," *IEEE Transactions on Communications*, Vol. COM-30, pp. 6–11, Jan. 1982.

[98] I. Chlamtac and M. Eisenger, "Voice/data integration on Ethernet — backoff and priority considerations," *Computer Communications*, Vol. 6, pp. 236–244, Oct. 1983.

[99] T. A. Gonsalves, "Packet-voice communication on an ethernet local computer network: An experimental study," *ACM Computer Communication Review*, Vol. 13, pp. 178–185, Mar. 1983. SIGCOMM '83 Symposium.

[100] C. Ziegler, G. Weiss, and E. Friedman, "Implementation mechanisms for packet switched voice conferencing," *IEEE Journal on Selected Areas in Communications*, Vol. 7, pp. 698–706, June 1989.

[101] E. Friedman and C. Ziegler, "Packet voice communications over PC-based local area networks," *IEEE Journal on Selected Areas in Communications*, Vol. 7, pp. 211–218, Feb. 1989.

[102] J. Azevedo, J. A. Martins, and P. G. de Oliveira, "An experimental study on voice/data integration over an ethernet LAN," in *Proceedings of the International Conference on Information Networks And Data Communications (INDC) 1994*, p. 19, 1994.

[103] International Telecommunication Union, "Voice packetization - packetized voice protocols," Recommendation G.764, Telecommunication Standardization Sector of ITU, Geneva, Switzerland, 1990.

[104] J. R. Brandsma, A. A. M. L. Bruekers, and J. L. W. Kessels, "PHILAN: a fiber-optic ring for voice and data," *IEEE Communications Magazine*, Vol. 24, pp. 16–22, Dec. 1986.

[105] L. M. Casey, R. C. Dittburner, and N. D. Gamage, "FXNET: a backbone ring for voice and data," *IEEE Communications Magazine*, Vol. 24, pp. 23–28, Dec. 1986.

[106] G. Barberis, M. Calabrese, L. Lambarelli, and D. Roffinella, "Coded speech in packet-switched networks: Models and experiments," *IEEE Journal on Selected Areas in Communications*, Vol. SAC-1, pp. 1028–1038, Dec. 1983.

[107] A. A. Kapaun, W.-H. F. Leung, G. W. R. Luderer, M. J. Morgan, and A. K. Vaidya, "Wideband packet access for workstations: integrated voice/data/image services on the Unix PC," in *Proceedings of the IEEE Conference on Global Communications (GLOBECOM)*, Vol. 3, (Houston, Texas), pp. 1439–1441 (40.6), IEEE, Dec. 1986.

[108] P. Spilling and E. Craighill, "Digital voice communications in the packet radio network," in *Conference Record of the International Conference on Communications (ICC)*, (Seattle, Washington), pp. 21.4.1–21.4.7, IEEE, June 1980.

[109] E. M. Schooler and S. L. Casner, "A packet-switched multimedia conferencing system," *SIGOIS (ACM Special Interest Group on Office Information Systems) Bulletin*, Vol. 10, pp. 12–22, Jan. 1989.

[110] V. Jacobson, "Multimedia conferencing on the Internet," in *SIGCOMM Symposium on Communications Architectures and Protocols*, (London, England), Aug. 1994. Tutorial slides.

[111] V. Jacobson and S. McCanne, "vat - LBNL audio conferencing tool," July 1992. Available at http://www-nrg.ee.lbl.gov/vat/.

[112] L. T. Corley, "Bellsouth trial of wideband packet technology," in *Conference Record of the International Conference on Communications (ICC)*, Vol. 3, (Atlanta, Georgia), pp. 1000–1002 (324.2), IEEE, Apr. 1990.

[113] H. Miyahara and T. Hasegawa, "Integrated switching with variable frame and packet," in *Conference Record of the International Conference on Communications (ICC)*, Vol. 2, (Toronto, Canada), pp. 20.3.1–20.3.5, IEEE, June 1978.

[114] I. Gitman and H. Frank, "Economic analysis of integrated voice and data networks: A case study," *Proceedings of the IEEE*, Vol. 66, pp. 1549–1570, Nov. 1978.

[115] J. W. Forgie, "ST - a proposed internet stream protocol," Internet Engineering Note IEN 119, MIT Lincoln Laboratory, Cambridge, Massachusetts, Sept. 1979.

[116] C. Topolcic, "Experimental internet stream protocol: Version 2 (ST-II)," Request for Comments (Experimental) 1190, Internet Engineering Task Force, Oct. 1990.

[117] C. Topolcic, "ST II," in *Proc. International Workshop on Network and Operating System Support for Digital Audio and Video (NOSSDAV)*, ICSI Technical Reports, (Berkeley, California), Nov. 1990. TR-90-062.

[118] L. Delgrossi and L. Berger, "Internet stream protocol version 2 (ST2) protocol specification - version ST2+," Request for Comments (Experimental) 1819, Internet Engineering Task Force, Aug. 1995.

[119] H. Schulzrinne, "Voice communication across the Internet: A network voice terminal," Technical Report TR 92-50, Dept. of Computer Science, University of Massachusetts, Amherst, Massachusetts, July 1992.

[120] S. Casner and S. Deering, "First IETF Internet audiocast," *ACM Computer Communication Review*, Vol. 22, pp. 92–97, July 1992.

[121] H. Eriksson, "MBONE: The multicast backbone," *Communications ACM*, Vol. 37, pp. 54–60, Aug. 1994.

[122] H. Schulzrinne, "Internet services: from electronic mail to real-time multimedia," in *Proc. of KIVS (Kommunikation in Verteilten Systemen)* (K. Franke, U. Hübner, and W. Kalfa, eds.), Informatik aktuell, (Chemnitz, Germany), pp. 21–34, Gesellschaft für Informatik, Springer Verlag, Feb. 1995.

[123] R. Frederick, "Experiences with real-time software video compression," in *Sixth International Workshop on Packet Video*, (Portland, Oregon), Sept. 1994.

[124] S. McCanne and V. Jacobson, "vic: A flexible framework for packet video," in *Proc. of ACM Multimedia '95*, Nov. 1995.

[125] M. Sattler and G. Horman, *Internet TV With Cu-Seeme.* SAMS, 1995.

[126] J. Becker, "Cyberpornographia cuseemeata: A field guide and management strategies," tech. rep., Cornell University, Nov. 1996.

[127] M. Altenhofen, J. Dittrich, R. Hammerschmidt, T. Kppner, C. Kruschel, A. Kckes, and T. Steinig, "The BERKOM multimedia collaboration service," in *Proc. of ACM Multimedia*, (Anaheim, California), pp. 457–463, Aug. 1993.

[128] T. Gutekunst and B. Plattner, "Sharing multimedia applications among heterogeneous platforms," in *Second International Conference on Broadband Islands*, (Athens, Greece), 1993.

[129] I. Kouvelas, V. Hardman, and A. Watson, "Lip synchronisation for use over the internet: Analysis and implementation," in *Proceedings of the IEEE Conference on Global Communications (GLOBECOM)*, (London, England), Nov. 1996.

[130] A. Sasse, V. Hardman, I. Kouvelas, C. Perkins, O. Hodson, A. Watson, M. Handley, J. Crowcroft, D. Harris, A. Bouch, M. Iken, K. Hasler, S. Varakliotis, and D. Miras, "Rat (robust-audio tool)," 1995.

[131] V. Hardman, A. Sasse, M. Handley, and A. Watson, "Reliable audio for use over the internet," in *Proc. of INET'95*, (Honolulu, Hawaii), June 1995.

[132] E. M. Schooler, "The connection control protocol: Specification (version 1.1)," technical report, USC/Information Sciences Institute, Marina del Ray, California, Jan. 1992.

[133] E. M. Schooler, "The connection control protocol: Architecture overview (version 1.0)," technical report, USC/Information Sciences Institute, Marina del Ray, California, Jan. 1992.

[134] E. M. Schooler, S. L. Casner, and J. Postel, "Multimedia conferencing: Has it come of age?," in *Proceedings of the 24th Hawaii International Conference on System Science*, Vol. 3, (Hawaii), pp. 707–716, IEEE, Jan. 1991.

[135] H. Schulzrinne, "Simple conference invitation protocol," Internet Draft, Internet Engineering Task Force, Feb. 1996. Work in progress.

[136] A. H. Wurtzel and C. Turner, "Latent functions of the telephone: what missing the extension means," in *The Social Impact of the Telephone* (I. de Sola Pool, ed.), pp. 246–261, Cambridge, Massachusetts: MIT Press, 1977.

[137] G. Gillespie, "Warning: Too much info hazardous to your health," *IEEE Institute*, June 1997.

[138] Times Mirror Center for The People and The Press, "Survey about computer use," Oct. 1995.

[139] Emerging Technologies Research Group, "The american internet user survey," survey, Find/SVP, New York, Oct. 1996. Reported in *Tampa Tribune*, January 12, 1996.

[140] S. Hansell, "Now, AOL everywhere," *New York Times*, July 1999.

[141] D. Lavin, "Tailoring new products is becoming more complex," *Wall Street Journal*, Aug. 1996. European Technology Section.

[142] A. M. Odlyzko, "The current state and likely evolution of the internet," in *Proceedings of the IEEE Conference on Global Communications (GLOBECOM)*, (Rio de Janeiro, Brazil), pp. 1869–1875, Nov. 1999.

[143] K. G. Coffman and A. M. Odlyzko, "The size and growth rate of the internet," *First Monday*, Vol. 3, Oct. 1998.

[144] R. Rosenberg, "Hard to beat – ATM is the carrier's silver bullet," *America's Network*, May 1998.

[145] International Telecommunication Union, "Narrow-band visual telephone systems and terminal equipment," Recommendation H.320, Telecommunication Standardization Sector of ITU, Geneva, Switzerland, May 1999.

[146] C. Malamud and M. Rose, "Principles of operation for the TPC.INT subdomain: Remote printing – technical procedures," Request for Comments (Experimental) 1528, Internet Engineering Task Force, Oct. 1993.

[147] K. Asatani, "Standardization on multimedia communications: Computer-telephony-integration-related issues," *IEEE Communications Magazine*, Vol. 36, pp. –, July 1998.

[148] Microsoft Corporation, "Tapi: Windows telephony application programming interface," 1998.

[149] C. Strathmeyer, "Feature topic: Computer telephony," *IEEE Communications Magazine*, Vol. 34, Apr. 1996.

[150] Versit Consortium, “Computer telephony integration (CTI) encyclopedia,” Tech. Rep. Release 1.0, Versit Consortium, Oct. 1996.

[151] C. Low, “The internet telephony red herring,” in *Proceedings of Global Internet*, (London, England), pp. 72–80, Nov. 1996.

[152] C. Low, “Integrating communication services,” *IEEE Communications Magazine*, Vol. 35, pp. –, June 1997.

[153] P. Moulton and J. Moulton, “Telecommunications technical fundamentals,” technical handout, The Moulton Company, Columbia, Maryland, 1996. see http://www.moultonco.com/semnotes/telecomm/teladd.htm.

[154] H. Schulzrinne, “RTP profile for audio and video conferences with minimal control,” Request for Comments (Proposed Standard) 1890, Internet Engineering Task Force, Jan. 1996.

[155] S. Casner and V. Jacobson, “Compressing IP/UDP/RTP headers for low-speed serial links,” Request for Comments (Proposed Standard) 2508, Internet Engineering Task Force, Feb. 1999.

[156] L. Zhang, S. Deering, D. Estrin, S. Shenker, and D. Zappala, “RSVP: a new resource ReSerVation protocol,” *IEEE Network*, Vol. 7, pp. 8–18, Sept. 1993.

[157] G. W. Weber, “Centrex feature comparison for small business,” memorandum, AT&T, San Ramon, CA, July 1993. see also http://www.first-tel.com/ltlfeat.htm.

[158] R. Fielding, J. Gettys, J. Mogul, H. Frystyk, and T. Berners-Lee, “Hypertext transfer protocol – HTTP/1.1,” Request for Comments (Proposed Standard) 2068, Internet Engineering Task Force, Jan. 1997.

[159] M. J. Drst, “Internationalized internet identifiers,” in *Proc. of 11th International Unicode Conference*, (San Jose, California), Sept. 1997.

[160] R. Hedberg and P. Pomes, “The CCSO nameserver (ph) architecture,” Request for Comments (Informational) 2378, Internet Engineering Task Force, Aug. 1998.

[161] W. Yeong, T. Howes, and S. Kille, “Lightweight directory access protocol,” Request for Comments (Draft Standard) 1777, Internet Engineering Task Force, Mar. 1995.

[162] T. A. Howes and M. C. Smith, *LDAP: Programming Directory-Enabled Applications with Lightweight Directory Access Protocol*. Indianapolis, Indiana: Macmillan Technical Publishing, 1997.

[163] F. Dawson and T. Howes, “vcard MIME directory profile,” Request for Comments (Proposed Standard) 2426, Internet Engineering Task Force, Sept. 1998.

[164] E. J. Cameron, N. D. Griffeth, Y.-J. Lin, M. E. Nilson, W. K. Schure, and H. Velthuijsen, "A feature interaction benchmark for IN and beyond," *Feature Interactions in Telecommunications Systems, IOS Press*, pp. 1–23, 1994.

[165] H. Schulzrinne and J. Rosenberg, "SIP caller preferences and callee capabilities," Internet Draft, Internet Engineering Task Force, Mar. 1999. Work in progress.

[166] P. L. O'Connell, "If a payphone rings, who will answer?," *New York Times*, p. D10, May 1998.

[167] International Telecommunication Union, "Procedures for document facsimile transmission in the general switched telephone network," Recommendation T.30, Telecommunication Standardization Sector of ITU, Geneva, Switzerland, July 1996.

[168] H. Alvestrand, "A MIME body part for FAX," Request for Comments (Proposed Standard) 2159, Internet Engineering Task Force, Jan. 1998.

[169] L. McIntyre, S. Zilles, R. Buckley, D. Venable, G. Parsons, and J. Rafferty, "File format for internet fax," Request for Comments (Proposed Standard) 2301, Internet Engineering Task Force, Mar. 1998.

[170] International Telecommunication Union, "Procedures for real-time group 3 facsimile communication over IP networks," Recommendation T.38, Telecommunication Standardization Sector of ITU, Geneva, Switzerland, June 1998.

[171] M. Rose and C. Malamud, "An experiment in remote printing," Request for Comments (Experimental) 1486, Internet Engineering Task Force, Aug. 1993.

[172] C. Malamud and M. Rose, "Principles of operation for the TPC.INT subdomain: Remote printing – administrative policies," Request for Comments (Informational) 1529, Internet Engineering Task Force, Oct. 1993.

[173] C. Malamud and M. Rose, "Principles of operation for the TPC.INT subdomain: General principles and policy," Request for Comments (Informational) 1530, Internet Engineering Task Force, Oct. 1993.

[174] C. Simon, "The web catches and reshapes radio," *New York Times*, Jan. 2000.

[175] J. Brassil, S. Garg, and H. Schulzrinne, "Program insertion in real-time IP multicasts," *ACM Computer Communication Review*, Vol. 29, pp. 49–68, Apr. 1999.

[176] A. Dutta and H. Schulzrinne, "MarconiNet - an architecture for internet radio and TV networks," in *Proc. International Workshop on Network and Operating System Support for Digital Audio and Video (NOSSDAV)*, (Basking Ridge, New Jersey), pp. 241–245, June 1999.

[177] D. Hoffman, G. Fernando, and V. Goyal, “RTP payload format for MPEG1/MPEG2 video,” Request for Comments (Proposed Standard) 2038, Internet Engineering Task Force, Oct. 1996.

[178] H. Schulzrinne, “Real-time services in the internet?,” in *Panel discussion at the International Conference on Network Protocols (ICNP)*, (Columbus, Ohio), Oct. 1996.

[179] Telect, “Audio & video terminology.” http://www.telect.com/corporat/glossary.cfm.

[180] M. Handley, H. Schulzrinne, E. Schooler, and J. Rosenberg, “SIP: session initiation protocol,” Request for Comments (Proposed Standard) 2543, Internet Engineering Task Force, Mar. 1999.

[181] H. A. Chinn, D. K. Gannett, and R. M. Morris, “A new standard volume indicator and reference level,” *Bell System Technical Journal*, Vol. 19, pp. 94–137, Jan. 1940.