

# インターネット電話のしくみ

小田 稔周

(株) KDD 研究所

## インターネット電話とは

近年のパーソナルコンピュータ (PC) の発展とインターネットの爆発的な普及と高速化とともに、インターネット上での音声や画像を表現媒体としたリアルタイム通信を実現する技術が急速に進展しており、そのような通信の一形態として、音声通信を対象としたインターネット電話が衆目を集めている。

インターネット電話 (Internet Telephony) とは、一般に、IP プロトコルを用いたネットワーク (以下、IP ネットワークと呼ぶ) を情報伝達媒体として利用する電話のことを示す。すなわち、電話機相当の端末を用い、音声信号を IP パケット化し IP プロトコルを用いる LAN、イントラネット、インターネットを通して転送する音声通信の総称として用いられる。IP ネットワークを利用する点を明示する用語として、IP 電話 (IP Telephony) や VoIP (Voice over IP) がほぼ同義語として用いられている。本稿では、そのような IP ネットワークを利用する音声通信を実現する技術について述べる。

IP ネットワークを用いて電話サービスを実現する形態は、利用する電話端末とその組合せによって大別される。インターネット電話の最初概念は、PC に電話機相当の機能を提供するソフトウェアとハードウェアを組み込んで PC 間でリアルタイムに音声通信を行う形態として登場した。1994 年頃から各種の PC 電話製品が市場に現れている。この通信形態は、音声、データ、画像を組み合わせたマルチメディア会議通信の研究開発が基盤であり、リアルタイム音声通信、すなわち電話を提供する部分が切り出されたものと捉えられる。

これに続き、公衆電話網とインターネットを相互接

続する概念が登場し、両端の通信端末として従来の電話機を用いインターネットを中継して電話を実現する形態が、IETF (Internet Engineering Task Force) の RFC (Request For Comments) 1789 (1995 年 4 月) にまとめられた。

端末として PC あるいは電話機を用いる 4 つの基本接続シナリオを図-1 に示す。シナリオ 1 は電話機相当の機能を有する PC (以下、PC 電話端末と呼ぶ) 間を接続する形態であり、シナリオ 2 は一方が PC 電話端末で他方が電話機である形態、シナリオ 3 は両端が電話機で IP ネットワークを中継して接続する形態である。また、シナリオ 4 は、PC 間の音声コネクションを回線交換網により中継して接続する形態である。シナリオ 2, 3, 4 では、公衆電話網 / ISDN や企業内電話網などの回線交換網とパケット交換網である IP ネットワーク間で通話接続の橋渡しを行う装置が必要である。そのような装置を、一般に、インターネット・テレフォニー・ゲートウェイ (Internet Telephony Gateway), ある

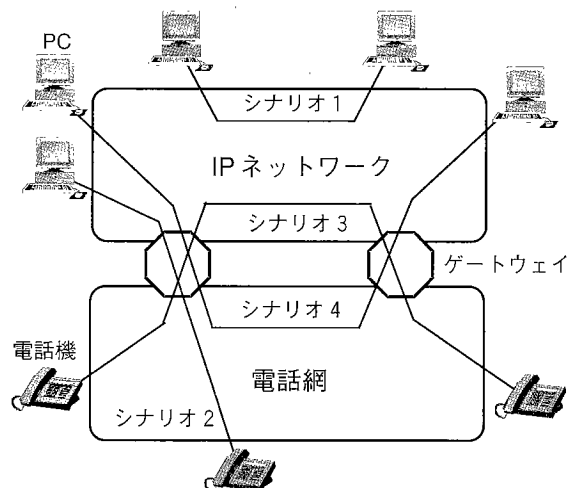


図-1 インターネット電話の基本接続シナリオ

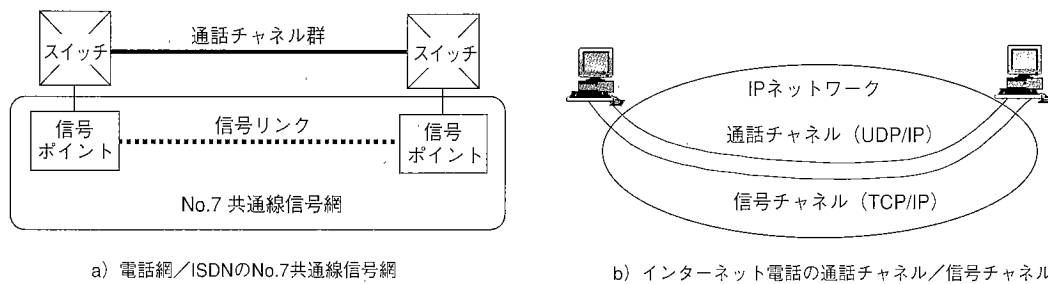


図-2 電話網/ISDNとインターネット電話の通話チャンネル/信号チャンネル構成の違い

いは単にゲートウェイと呼ぶ。ここに示した接続シナリオは、基本的な接続シナリオであって、それらを組み合わせたさらに多彩な接続形態が考えられる。2点間の通信に限らず、3点間以上を同時に接続する電話会議の形態も、基本形態の組合せとして各種考えられる。

## これまでの電話網との違い

### 帯域割当方式の違い

インターネット電話と従来の回線交換による電話網との大きな相違点として、まず、情報を伝達するネットワークでの帯域割当方式の違いが挙げられる。物理回線を割り当てる回線交換網では、接続される呼に対してエンド・エンドで固定的な帯域が割り当てられ、基本的には割当可能な帯域が枯渇するまで呼接続要求が受け付けられる。このため、到着トラフィックの増加により、呼接続率は低下するが、接続に成功した呼の割当帯域は影響されない。一方、IPネットワークは、基本的に呼受付制御機構を提供せず、帯域共有によってベストエフォート型のパケット転送機能を提供するのみで、エンド・エンドで確定的な帯域が割り当てられるわけではない。このため、トラフィック増加により、IPネットワークにおいて呼接続率が低下することはないものの<sup>\*</sup>、各通話において実効的に使用できる帯域は減少する。このように、使用するネットワークでその性能、すなわち、接続性能と情報転送性能が異なるため、結果としてユーザが感知する通信品質（接続品質と通話品質）が異なることとなる。

### 番号体系の違い

また、従来の電話網では、番号計画を規定するITU-T (International Telecommunication Union-Telecommunication Standardization Sector) 勧告E.164に則った電話番号体系が通信当事者を識別する手段として用いられ、さらに、サービスの識別、電話サービス事

業者の識別、課金先の識別などにも用いられる。すなわち、電話網では、電話機から番号をダイヤルするという形式のユーザインタフェースで統一されている。これに対し、インターネット電話では、IPネットワークを経由する部分でその両端点を識別する必要があり、そのベースになるのは関与する機器のIPアドレスである。通常、インターネットの各種アプリケーションでは、メールアドレスやURL (Uniform Resource Locator) などのalias (別名) を用いて最終的に通信の対象となる機器のIPアドレスを指定できるメカニズムが具備されており、ユーザが直接的にIPアドレスを毎回指定して通信を行うことは少ない。インターネット電話においても、同様のメカニズムが求められ、ネーミングサーバにより個人氏名とIPアドレスをマッピングするものがあるが、今後、さらにさまざまな工夫が施されるであろう。現在、ITU-Tでは、インターネット電話のためのグローバルサービス番号 (国番号に相当する番号) 割当の必要性について議論されている。そのような番号割当が採択されると、PC電話端末などインターネット電話端末として用いる端末に電話番号相当の番号を新たに付与できることとなり、政治的なインパクトが大きいですが、サービス面ではユーザの利便性を向上させる一手段となる。

### 通話チャンネル/信号チャンネル構成の違い

電話網とインターネット電話では、通話チャンネルと信号チャンネルの構成の点でも異なる。近年の電話網やISDNは、図-2に示すように、通話チャンネル網とNo.7共通線信号網が物理的に分離されている。これに対し、インターネット電話では、通話チャンネルと信号チャンネルは論理的に分離されているが、IPネットワーク上に統合的に実現される。このため、インターネット電話は、IPネットワーク上に実現される画像通信やWebアプリケーションなどの多様なインターネット・アプリケーションとの親和性が高く、マルチメディア通信としての発展性だけでなく、インターネット上に構築される各種のサービスとの融合が容易である。

その他、設備のコスト構造やインターネット電話サービス提供事業者間での相互接続の方法なども、電話

<sup>\*</sup> IPネットワーク上に構築されるインターネット電話アプリケーションの機能として、呼接続管理サーバ等により呼受付制御が行われる場合がある。また、ゲートウェイ等の処理能力限界により同時接続呼数が制限される。

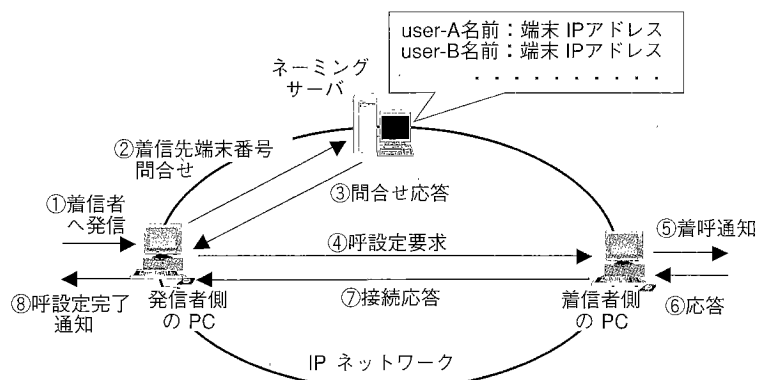


図-3 PC間での呼設定手順の概略

網とインターネット電話での大きな相違点である。また、電話網は24時間継続してサービスされることが基本となっているが、PC電話端末を用いるインターネット電話では、そのPCが動作状態にあることが必要である。

### インターネット電話における通話接続のしくみ

インターネット電話における通話は、一般の回線交換型通信と同様に、発着信端末間に通話チャンネルを設定する呼設定フェーズ、その通話チャンネルを用いて会話が行われる情報転送フェーズ、通話チャンネルの切断を行う呼解放フェーズから構成される。

#### PC電話端末間での呼接続

PC電話端末を用いた上記の接続シナリオ1における呼設定フェーズの原理を図-3に示す。発信側ユーザはPC上の電話インタフェースを用い、たとえば、電話着信先ユーザの名前を指定して、呼接続を要求する(①)。PC電話端末は、着信先PC電話端末のIPアドレスを入手するため、あらかじめユーザの名前と対応するPC電話端末のIPアドレスを保持するネーミングサーバに問い合わせる(②)。これに対する応答を受け(③)、発信PC電話端末は着信先IPアドレスとポート番号を指定して、呼設定信号を送出する(④)。着信側では呼設定信号を受けると着信情報が表示され(⑤)、着信側ユーザが応答する(⑥)。その結果、発着PC間で通話チャンネルが設定され(⑦)、通話状態であることが表示される(⑧)。当然、発信時に着信先PCのIPアドレスが分かっている場合には、それを直接入力することで、ネーミングサーバへのアクセスを省略できる。

#### インターネット中継電話における呼接続

次に、電話網とIPネットワークを接続するゲートウェイを導入して電話機—電話機間でインターネット電

話を行う形態(上記接続シナリオ3)について説明する。現在のゲートウェイの多くは、電話網から見ると、一種の電話端末として扱えるように構成されており、ここでもその場合を想定する。図-4に、呼設定の基本的な流れを示す。発信ユーザは電話機からゲートウェイの電話番号をダイヤルし(①)ゲートウェイと接続して(②)、ゲートウェイからの音声ガイダンス(③)に従って、着信先の電話番号をダイヤルする(④、⑤)。ゲートウェイは、指定された着信先電話番号に対応する着信側ゲートウェイのIPアドレスを得るため、電話番号とゲートウェイの対応情報を保持する番号変換サーバ(後述するITU-T勧告H.323におけるゲートキーパなど)に問い合わせる(⑥、⑦)。発信側ゲートウェイは、取得した着信側ゲートウェイのIPアドレスと呼制御プロトコル用のポート番号を指定して、着信先電話番号と発信側ゲートウェイの通話チャンネル用ポート番号指定を含む呼設定信号を送出する(⑧)。これを受信した着信側ゲートウェイは、電話網に対し、着信先電話番号のダイヤル・トーンを送出して着信先電話機へ発呼する(⑨)。着信ユーザが電話機をオフフックすることで応答し、通話チャンネル設定が完了して(⑩、⑪)、通話フェーズへ移行する。ここで、着信側ゲートウェイが、着信電話機からの応答を受信する前に、発信側ゲートウェイに対して接続応答(着信側ゲートウェイの通話チャンネル用ポート番号指定を含む)を返す構成や、さらに着信側から発信側への通話チャンネルを設定する構成も可能である。

上記の場合、電話網交換機とゲートウェイは、ユーザ・網間プロトコル、すなわち、加入者線プロトコルにより接続されている。これに対し、交換機とゲートウェイをNo.7信号プロトコルで接続する構成も考えられ、そのような機能を有するゲートウェイ製品も市場に登場してきている。このようなゲートウェイを適用すれば、電話網とIPネットワークとの接続点を電話網中継系内にとることができ、ゲートウェイを中継交換機と同様に扱え、より柔軟なルーティングが可能と

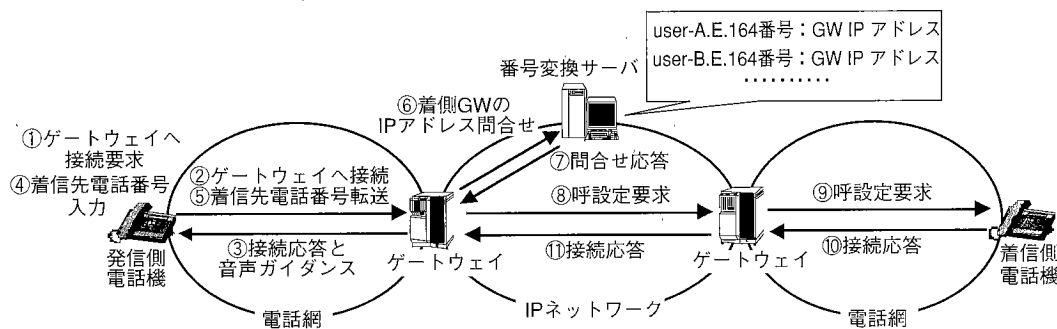


図-4 IPネットワークを中継する場合の呼接続手順の例

なる。

#### 異なるインターネット電話装置間での相互運用に向けて

IPネットワーク内での呼処理プロトコルとして、これまでは、ベンダごとに固有なプロトコルが用いられ、異なる製品間での相互運用性は必ずしも実現されていなかった。一方、ITU-Tにおいて、IPネットワーク等を用いるリアルタイム・マルチメディア通信用プロトコルとして策定されているH.323は、当然ながら、インターネット電話へ適用可能であり、標準プロトコルとして有力視されている。PC電話端末やゲートウェイ相互間で相互運用性を確保するため、現在、H.323準拠の機器開発が活発に進められている状況である。

H.323は、呼制御プロトコルH.225、マルチメディア通信システム制御プロトコルH.245およびコーデック規定からなるプロトコル・セットである。H.225は、ISDNのユーザ網間呼処理プロトコルQ.931、Q.932をベースとした呼処理プロトコル、端末とゲートキーパ間のRAS (Registration, Admission and Status) プロトコル、およびRTP/RTCP (Real Time Protocol/Real Time Control Protocol) の使用を規定している。H.245は、端末間での端末能力情報の交換や通話信号用論理チャネルのオープン/クローズなど、システム制御のためのプロトコルを規定している。H.323における呼接続の大まかなステップは、1) RASプロトコルによる呼の受付、2) Q.931呼設定信号シーケンス (Setup, Call Proceeding, Alerting, Connectなどの信号) による呼の設定、3) H.245による端末能力情報の交換と通話チャネルのオープン、4) 通話チャネルを用いた情報転送、5) H.245セッションの終了、6) Q.931信号 (Release Complete信号) による呼の完了から構成される。また、H.323をベースとした呼転送などの付加サービスについても研究が進められており、勧告H.450.1, 450.2, 450.3がすでに規定されている。

しかしながら、音声通信のみに適用するには、H.323

の全仕様は複雑であるため、より信号数や処理量の少ないプロトコルの標準化を求める声もある。特に、通話チャネル設定までのプロトコルの簡素化は、後述の接続品質向上の点で重要であり、H.323においても、平成10年3月に策定された第2版には高速接続手順 (Fast Connect Procedure) が盛り込まれ、Q.931呼設定信号シーケンスのみにより通話チャネルの設定を可能としている。今後、技術面だけでなく市場性も絡みながら標準化が進展していくこととなろう。

#### 呼接続にかかわる品質

公衆電話網/ISDNでは、呼処理面での通信品質、すなわち、接続品質として、呼処理にかかわる処理遅延時間と呼損率が重要であり、ITU-TのEシリーズ勧告にそれらの目標値が規定されている。たとえば、ITU-T勧告E.721では、ISDNにおける呼設定遅延時間 (Post-Selection Delay) を平常負荷時の国際接続で平均8sec以下と規定している。インターネット電話に対しても同様な規定を適用すべきか、あるいはどのような接続品質規定を定めるべきかについては今後の研究が待たれるところであるが、使用される呼処理プロトコルやネットワーク構成法、サービス提供形態等によって、接続品質は変化し、サービスクラスを選択肢拡大としてそのまま市場に受け入れられる可能性もあろう。

#### 音声転送のしくみ

情報転送フェーズでは、電話端末から入力される音声信号がIPネットワークを介して着信端末まで届けられ、再生される。PC電話端末を想定した場合に、音声信号がパケット化されて転送されるしくみを図-5に示す。図-5には、呼処理信号の送受信処理のしくみも同時に示した。

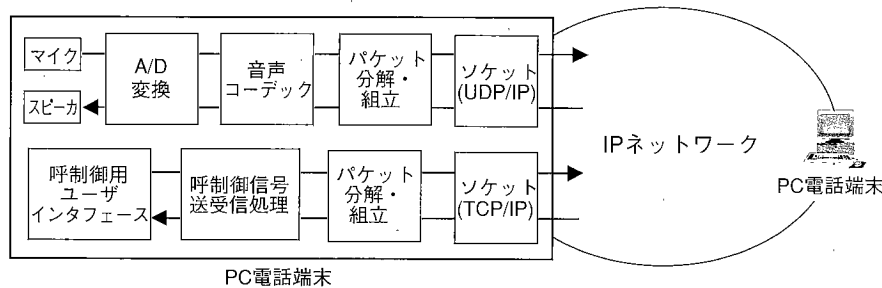


図-5 音声転送のための機能構成例 (PC電話端末の場合)

### 各種音声コーデックの利用

PCに接続されたマイクから音声が入力され、A/D変換装置(サウンドカードなど)によりPCM信号に変換される。インターネット電話では、PCM信号をそのままパケット化して転送する場合もあるが、低レート音声コーデックを適用して、必要帯域を低減させる場合が多い。ゲートウェイでは、32kbps ADPCM (Adaptive Differential PCM) や 13.2kbps GSM (Global System for Mobile communications) が用いられる例があるが、相互運用性の観点から、表-1に示すH.323準拠コーデックが主に使用されよう。音声符号化装置の出力信号はバッファに蓄積され、これを元に送信用PDU (Protocol Data Unit) が構成される。通常は、これにRTPヘッダ、UDPヘッダ、IPヘッダを付加して、音声信号の送信IPパケットが構成される。一方、受信側では、受信した音声IPパケットのペイロード部から符号化音声信号を取り出し、復号化する。なお、RTPヘッダ(固定部12バイト)、UDPヘッダ(8バイト)、IPヘッダ(IPv4 20バイト)で合計40バイトとなり、特に低レート音声コーデックではこのオーバーヘッドによる効率低下が大きい。そのため、これらヘッダの冗長性を除去するヘッダ圧縮手法の研究も進められている。

ゲートウェイを用いて電話網とIPネットワークを接続する場合は、ゲートウェイにおいて上記と同様なメカニズムにより、電話回線上の音声信号と音声IPパケット間の変換が行われる。電話加入者線にゲートウェイを接続する場合、大まかには、図-5に示したマイク・スピーカ部と呼制御用ユーザインタフェースを、電話加入者線の通話信号送受信機能と信号プロトコル処理機能でそれぞれ置き換えた機能構成となる。加入者線信号プロトコルには、PBXとの接続も考慮すると、アナログ/ISDN/T1/E1インタフェースなどがあり、接続対象に応じて選択することとなる。

### IPネットワーク性能の通話品質へのインパクト

インターネット電話では、IPネットワークを用いるため、パケット損失、パケット遅延、パケット遅延ゆ

らぎ、パケット順序逆転などのIPネットワーク固有の網性能によって、通話品質が変化する。パケット損失は、使用される音声コーデックに依存するが、多くの場合、クリック音として知覚される。パケット損失への対処として受信側で補償する方法などがある。パケット遅延は、会話における発声タイミングの乱れやエコーの原因となる。また、パケット遅延ゆらぎは、その程度が大きい場合、受信側の音声信号バッファのオーバーフロー、アンダーフローを発生させ、音声信号の欠落などの原因となる。パケット順序逆転は、受信側での対処に依存するが、音声信号誤りなどとして処理される。

エンド・エンドのパケット転送遅延は、PC電話端末またはゲートウェイ内部での処理遅延、ルータ内処理遅延、送信遅延、伝播遅延が要因となる。電話接続に関するエンド・エンドの信号伝送遅延時間要件はITU-T勧告G.114に規定されており、エコーキャンセラを具備した通話回線では、音声伝送遅延時間が150msec以下であればほとんどのユーザアプリケーションで許容され、ユーザアプリケーションへの影響を知った上で運用される場合400msecまで許容可能とされている。現状のインターネット電話システムでは、低レート音声コーデックにおける符号化時間間隔と処理遅延、バッファリング遅延によるPC電話端末またはゲートウェイ内部でのパケット化遅延が大きな遅延要素となっているが、DSPの利用やCPUの高速化などにより、さらに

表-1 H.323準拠の音声コーデック

ITU-T勧告	コーデック略称	レート
G.711	PCM	64kbit/s
G.722	7kHz ADPCM	48~64kbit/s
G.723.1	MP-MLQ/ACELPの複合	6.3kbit/s/5.3kbit/s
G.728	LD-CELP	16kbit/s
G.729	CS-ACELP	8kbit/s

MP-MLQ: Multi-Pulse Maximum Likelihood Quantization  
 ACELP: Algebraic Code-Excited Linear Prediction  
 LD-CELP: Low-Delay Code-Excited Linear Prediction  
 CS-ACELP: Conjugate-Structure Algebraic-Code-Excited Linear-Prediction

遅延短縮化が進むであろう。一方、今後は、従来の高品質な電話サービスから低品質で低廉なサービスまで、サービス品質に幅が生じることとなり、さまざまなサービスクラスが生まれてくるものと思われる。

大陸間では伝播遅延が大きいため、特にアナログ回線を含む電話網と接続するケースでは、エコー対策が必要である。電話機とゲートウェイ間は、多くの場合、近距離通話として接続されるため、電話網におけるその区間では通常エコー対策は施されていない。したがって、ゲートウェイにエコーサプレッサまたはエコーキャンセラを具備させる方策が採られる。

## サービスの高度化：IPによるサービス総合網への発展

インターネット電話は、IPネットワークを介して利用できる多種多様な通信アプリケーションや情報処理アプリケーションとの連携が容易であるという点が大きな特徴である。

呼のルーティング面では、PC電話端末あるいはIPネットワーク上で入手可能な情報（ユーザに関する情報など）や、既存電話網やISDNにおけるインテリジェントネットワークの資源を利用することで、PC電話端末あるいは電話機を自動的に選択して着信させるなど、状況に応じた柔軟で適応的なルーティングが可能となる。また、番号変換についても、各種のデータベースをIPネットワーク上に構築する、あるいはIPネットワーク経由でアクセスすることにより、柔軟性を向上させることが容易である。たとえば、氏名、IPアドレス、メールアドレス、URL、E.164番号などを相互に対応させることで、電話発着信操作における多様性が拡大する。課金の面では、異なるサービス・アプリケーション間での課金情報の統合的処理などにより、柔軟な方式が容易に実現可能である。

また、ルーツや定義はさておき、インターネット電話はCTI（Computer Telephony Integration）の一形態として語られる場合が多い。PBXをインターネット電話ゲートウェイで置換したり（un-PBX化）、あるいは両者を連携させることにより、VoIPベースのコールセンターシステムやユニファイド・メッセージング・システムを構築可能であり、これによって、より経済的に電話と情報処理業務の連携を図り、さらなるオフィス業務の効率化が可能になるものと期待されている。

インターネット電話は、ここ2、3年の間にゲートウェイ製品の性能向上と市場形成が進展し、同時にインターネットを中継網に利用する長距離電話サービスが出現してきたことで、現実のものとなった。現在は、市場はまだ小さく、需要を創造するための投資が行われている時期であり、従来の市場ライフ・サイクルのモデルでいえば、導入期に当たるものと考えられる。現

在のインターネット電話サービスは、電話機間を接続する電話サービスが中心であり、経済性の面が強調されている。しかし、“IPネットワーク・ベース”であることは、低価格の電話サービス提供のみにとどまらず、インターネットの高速化に伴って、IPネットワーク上で電話からリアルタイム・マルチメディア通信への発展、さらには情報処理も含めたサービス総合化が容易に進展しうるポテンシャルをインターネット電話が有することを意味しているのではなからうか。

## 引用したITU-T勧告の一覧

- E.164: The international public telecommunication numbering plan
- E.721: Network grade of service parameters and target values for circuit-switched services in the evolving ISDN
- G.114: One-way transmission time
- G.711: Pulse code modulation (PCM) of voice frequencies
- G.722: 7kHz audio-coding within 64 kbit/s
- G.723.1: Dual rate speech coder for multimedia communications transmitting at 5.3 and 6.3 kbit/s
- G.728: Coding of speech at 16 kbit/s using low-delay code-excited linear prediction
- G.729: Coding of speech at 8 kbit/s using conjugate-structure algebraic-code-excited linear-prediction
- H.225: Call signalling protocols and media stream packetization for packet based multimedia communication systems
- H.245: Control protocol for multimedia communication
- H.323: Packet based multimedia communications systems
- H.450.1 Generic functional protocol for the support of supplementary services in H.323
- H.450.2 Call transfer supplementary service for H.323
- H.450.3 Call diversion supplementary service for H.323
- Q.931: ISDN user-network interface layer 3 specification for basic call control
- Q.932: Generic procedures for the control of ISDN supplementary services

(平成10年11月16日受付)