SLAVKO ŠARIĆ, D. Sc.
Fakultet prometnih znanosti
Vukelićeva 4, 10000 Zagreb, Republika Hrvatska
E-mail: slavko.saric@fpz.hr
ANTO BILOBRK, B. Eng.
Cedevita d. o. o.
Planinska bb, 10000 Zagreb, Republika Hrvatska
E-mail: anto.bilobrk@cedevita.hr
DAVOR NAĐ, B. Eng.
Hrvatska televizija
Prisavlje 3, 10000 Zagreb, Republika Hrvatska
E-mail: dnad@email.htnet.hr

Technology and Management of Traffic
Review
U. D. C.: 656.07:621.395.658
Accepted: Jun. 16, 2004
Approved: Oct. 27, 2004

# VOICE TRANSMISSION OVER IP NETWORKS

*ABSTRACT*

*Voice transmission over IP networks (Voice over Internet Protocol) represents one of the ways in which voice and data networks are integrated. The future development is based on the explosion of the Internet as the means of communication, with the openness of standards and the readiness of the equipment manufacturers to accept such standard and to unify it. The service providers find interest in introducing new services that are not based only on voice transmission, but voice becomes only one of the applications that are realised over the IP networks. Voice transmission over the IP technology is at the moment not at the level of the existing quality of services, but the coming solutions in the near future will enable VoIP as the standard operative solution. The advantages are reflected in the increase of income keeping the current users and attracting new ones, investments into infrastructure will maximize the opportunities for packet service development, strengthen customers' loyalty and reduce operative expenditures, the services will be widespread – long-distance international services or national services of calling cards can be located at almost any European, African, or Central-Eastern country and directed to almost 100 percent of the population. With the advantages of low initial costs of the new POPs (Points of Presence), the service providers can suddenly expand their presence to many countries or destinations.*

*KEY WORDS*

*IP telephony, protocol architecture, Call center*

## 1. INTRODUCTION

The Internet telephony or IP telephony is the transmission of voice using the Internet Protocol (IP). In the mid-nineties, the professional circles started to speak about the transmission of speech using the IP network as a new method of voice communication that would substantially reduce the costs of intercity and international talks. The first applications (program packages for multimedia computers) allowed the Internet users to make telephone calls towards other active Internet users. At the same time the first service providers for the IP telephony started to appear.

During the first years of its existence, the IP telephony became extremely interesting to all the manufacturers of the telecommunication equipment as well as to the new potential investors who recognized a good opportunity for making profit. Gateways (VGW – Voice Gateway) have been introduced, and they allow transmission of voice from fixed and mobile public telephone networks towards data networks, as well as the transmission of voice from PCs towards telephone terminals.

Unlike PSTN which applies the channel switching, the VoIP system applies the networking of the packet switching, with the digitisation, compression and packetization of the voice signal. This method improves the level of utilization of the voice channel during conversation.

Wide application of the Internet telephony requires standardization, which allows communication between the programs and the devices of various manufacturers in the same network. The first specification that has been defined is the H. 323 regulated by the ITU (International Telecommunications Union), and new standards and architectures are appearing, such as the protocol for session initiation (SIP – Session Initiation Protocol) and MEGACO (Media Gateway Control) architectures[1].

In its early development phase the VoIP technology provides low level of service quality as compared to the analogue telephony, but it has been improving fast and today the quality of VoIP telephone call over virtual private network (VPN) equals the quality of an analogue call.

The Internet telephony implies the change in the method of designing the telecommunication structure by making the telephony just an additional service available on the Internet, thus simplifying significantly the whole infrastructure in a single important segment. The same principle, but to a much smaller volume is also valid for the private communication infrastructure which is typically present in various types of business and other types of organizations.

## 2. IP TELEPHONY

The Internet telephony is transmission of speech in the packets over the Internet Protocol (IP) and for the transmission of packets any packet network that uses this protocol can be used, such as the Internet, Intranet or local networks (LAN- Local Area Network)[2]. The relation between a network with channel switching and package networks used by the IP telephony is in the occupancy of the voice channel and the quality of voice. The network with the channel switching occupies during the whole conversation the fixed bandwidth channel of 64 kbit/s although the actual utilisation of the channel lies at as little as 50 percent, taking into consideration the period with no traffic when one does not speak because of listening to the other speaker or pausing between words. The application of VoIP technology makes it possible to use this unused bandwidth for other purposes by using the *voice activity detection* (VAD). VAD functions by detecting the intensity of speech in decibels (dB) and deciding when speech transmission should be interrupted. The usual method of operation is that VAD detects the fall in the voice amplitude and then waits for a certain period of time before stopping the setting of voice frames into packets. This defined period of time is known as *hangover* and usually takes 200 ms.

Although analogue communication is ideal for human communication, analogue emission is not efficient in recovering from line interferences since the noise gets stronger by amplifying the signal. It is much easier with digital samples that consist of bits 1 and 0, regarding the separation of noise on the line. Therefore, when the analogue signal is regenerated, sound is obtained as the digital sample, thus improving the quality of speech. However, every time when there is conversion from digital to analogue and vice versa, the form of speech becomes less accurate and the D/A conversion must be handled with great care. A simplified service process (telephone – telephone and telephone – IP terminal) are presented in Figure 1.

The gateway receives and sends the signals from the telephone network and towards it, interprets the signals, and by using the signalisation for VoIP (e. g. H. 323 or SIP) initiates the connection over the IP networks towards other IP terminals, such as IP telephones or personal computers with adequate IP telephone program, or towards gateways. H. 323 servers for call control (GK - Gatekeeper) or SIP servers will receive signalling information with the called number from the gateway, then, by means of address mapping tables, they will connect the end point and its current IP address with the telephone number and with this information allow the gateway to establish a direct me-

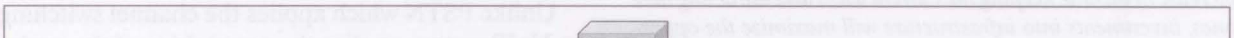


Figure 1 - VoIP network scheme

dia connection. If the end point is outside the IP network, the server determines which determines which gateway is responsible for the transmission of speech into the PSTN network to which the called number belongs. If the terminal with the called number is in the IP network, the voice content is exchanged directly between the gateway and the terminal[3].

After having transmitted the voice or fax signal over the PSTN network to the gateway, the signals are processed for the transmission over the IP network. The processing means digitization, if the incoming signal is in the analogue form, compression and packetization, cancelling echo and eliminating silence. The gateway compresses the speech signal for two reasons: in order to reduce the necessary transmission capacities and in order to reduce the delays in the IP network. One of the necessary requirements in the transmission of speech is the transmission in real time, since this is a service very sensitive to delays. The VoIP delay or latency is the period of time necessary for the speech produced by the speaker to reach the listener. Three types of delays are characteristic for the telephone networks: propagation delay (caused by the speed of light in the networks based on optical fibres or copper), processing delay and serialization delay. The processing delay defines several different causes of delay (packetization, compression, packet exchange) and the causing factors are the devices that send frames over the network. The serialization delay is the only component of the end-to-end delay. The G. 114 recommendation given by the International Telecommunication Union Telecommunication Standardization Sector (ITU-T) specifies that for high quality of voice the end-to-end delay should not exceed 150 milliseconds (Round Trip Delay, delay of the whole path). One of the ways in which end-to-end delay can be influenced is by means of *jitter*, the variation of time between sending and receiving the packet. Although many manufacturers decide to use the static jitter buffers, well-designed dynamic jitter buffers are the best mechanism to use in the packet-based voice networks because the dynamic jitter buffer adapts its size to match the delay variations between sending and receiving of the last several packets.

Furthermore, after the signal has arrived to the gateway, the gateway performs the same operation in the reverse order for the packets arriving from the IP network. Both operations (arrival to the telephone network/departure from the telephone network) are possible at a time, allowing thus two-way (full-duplex) conversation.

In case of IP terminal – IP terminal and IP terminal – telephone services, the call is initiated by the IP telephone program on these devices by using the H. 323 or SIP protocol. H. 323 Gatekeeper or SIP server will connect the called number with the IP address of the called terminal or with the gateway that transmits this call further to the PSTN network.

Encoding is necessary in transmitting the voice signal. The encoding algorithm has to operate in real time, has to provide satisfactory quality and has to be able to recover the lost packets. In voice transmission the lost packets are not resent, unlike data transfer, because in real-time voice transmission there should be no retransmission. Retransmission would cause additional delays that would be unacceptable in this type of communication. In order to avoid interruptions in speech, the receiving party must generate the speech signal at the time during which the lost packet should be reproduced.

Apart from *jitter* that would reduce the delay to an acceptable measure, it is necessary to use the routers and LAN switches with the installed mechanisms for service differentiation (Diffserv). By using this mechanism, the voice transmission receives higher priority compared to data traffic, and in this way the voice packets stay shorter in the element buffers waiting to be routed. Routers have to feature extremely high speed performances of routing and checking the routing tables, because voice over IP networks is transmitted in very small packets (smaller than the data packages) and such packets are very numerous in arrival. Small packets are used to reduce the delay on the receiving side caused by the conversion of signals into their original form and in order to increase the quality of service (QoS) in case of lost or damaged IP packets (large packets increase the time of reproduction – greater quantity of lost information).

## 3. PROTOCOL ARCHITECTURE

Protocol architecture in IP telephony consists of a series of protocols (Figure 2). The Internet Protocol (IP), Transmission Control Protocol (TCP), User Datagram Protocol (UDP) and the protocols of lower layer which serve the higher layer protocols necessary for the transmission of speech by means of packets over IP telephony.

- *IP (Internet Protocol)* – The Internet Protocol is a protocol of the network layer, connectionless, does not provide safe communication regarding accurate transmission of data. IP implements two basic functions: addressing and fragmenting. It receives the data from the upper layers, adds the header and forwards these to the lower layer. All the routing data are contained in the IP header. If the data from the transport layer exceed the maximal value a channel can accept, IP performs fragmentation and re-assembles the packet. The Internet Protocol treats every packet as an independent entity and these packets are called Internet protocol datagrams. They contain the information about

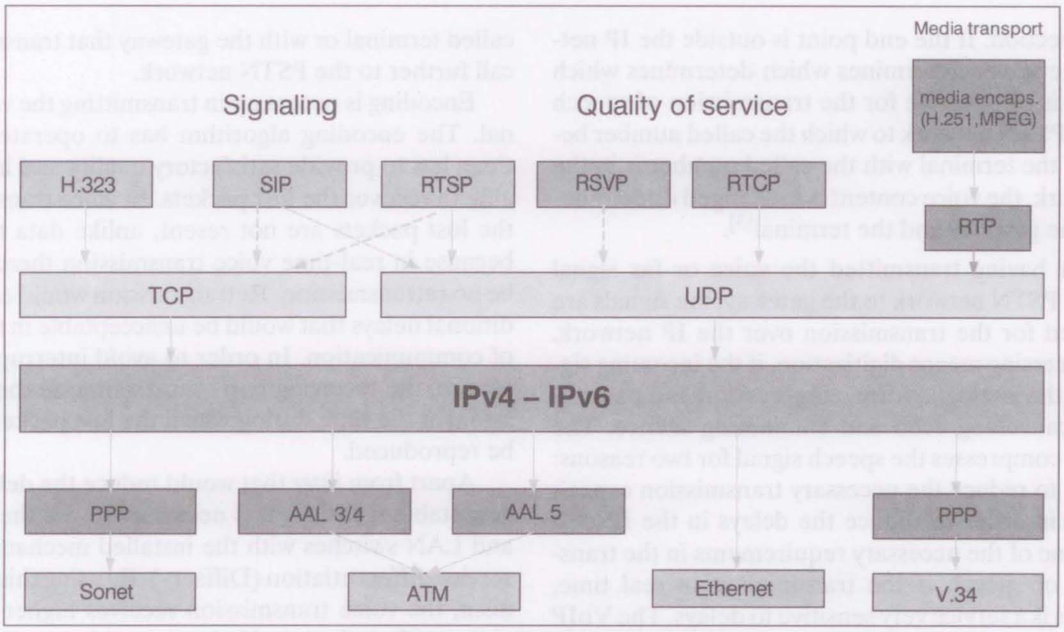


**Figure 2 - Protocol architecture of IP telephony**

the origin and destination, but they do not contain any other information that would determine their path through the network. Although network communications use data flows, the Internet traffic flows over datagrams, and the flows are emulated by means of higher layer protocols – TCP.

- The new version of the protocol, Ipv6, is slowly replacing the current version Ipv4. The constraints of Ipv4 result from its purpose for simple applications (file transfer, electronic mail and working on remote PCs), whereas today the address space may be even running out or the routing tables increase as result of the growth of the Internet. Apart from scalability Ipv6 provides the basis for the solution of other problems, such as the security, mobility and quality of service. The most important changes compared to Ipv4 include: larger address space (128 bit) simplification of header format, possibility of expanding the header by new types of headers, quality of service, simplified routing and security mechanisms on the network layer. The transition from one version is carried out gradually since the protocols are not compatible although they occupy the same place in the protocol structure.
- *TCP (Transmission Control Protocol)* – Protocol for the control of transmission provides the connecting service of the octet transport into a connectionless IP. TCP insures reliable transport, establishes end-to-end connection, and has installed time control and retransmissions. TCP contains the following functions: basic data transport, addressing and multiplexing, reliability, flow control, connection control, and priority and security. The logical connection between the processes is defined by the pair of 16-bit transport addresses, the so-called ports. TCP is suitable for reliable data exchange, but the real-time applications have to install their own time controls and flow control mechanisms because the accuracy of information is important for data transfer, and in *real-time* transmission the timely delivery of information is more important.
- *UDP (User Datagram Protocol)* – simple protocol which provides connectionless, unreliable service of UDP packet transport above IP. The UDP packets format is much simpler compared to TCP. The reliability mechanisms are built on the layers above the UDP protocol, but this protocol continues to be suitable for real-time information transfer. The reliability of this protocol is the responsibility of the application that uses it.
- *H. 323* – protocol for multimedia communication, defined specification by the ITU in version 1 (1996), version 2 (accepted in 1998, modified version 1 with the aim of accepting the packet-oriented multimedia communications, especially Voice over IP and Voice over ATM), version 3 (accepted in 1999) version 4 (2000) and version 5 (2003).
- *SIP (Session Initiation Protocol)* – protocol for the establishment, transfer and interruption of the session between two users or multimedia conferences, founded by IETF (Internet Engineering Task Force).
- *RTSP (Real Time Streaming Protocol)* – the suggested IETF protocol for standardisation of streaming media on the Internet. It operates on top of RTP insuring control mechanisms and taking care of higher level matters such as establishing of connection and permitting access. RTSP is try-

ing to provide the same services for the audio and video data flows as those provided by HTTP for text and graphics. It has been intentionally designed so as to have a similar syntax and functions as HTTP so that it can be expanded by some HTTP mechanisms. RTSP differs from HTTP in several points: HTTP is a *stateless* protocol, whereas RTSP keeps the state (identifier of the meeting) for every current display. HTTP is basically an asymmetric protocol where the customer sends a request, and the server answers, whereas in RTSP both the server and the customer can send requests. RTSP messages are usually sent by an independent channel, rather than by the one used by data. They may be sent by continuous transport connections, or a single connection can be established upon request, or one may use connectionless mode.

– *RTP (Real-time Transport Protocol)* – IP-based protocol which provides support for real-time data transmission (audio and video). RTP provides time reconstruction services, identifies lost packets, security and identification of contents. It was initially created for multi-destination (multicast) transmission of real-time data, but can be used also for individual (unicast) transmission, one-way transmission (*Video-on-Demand*), and for interactive services such as the Internet telephony. RTP is usually implemented in the very application as well as the solutions for the recovery of the lost packets and the control of congestion.

– *RTCP (Real-time Transport Control Protocol)* – real-time transmission control protocol, through control information packets together with RTP provides the following services: supervision of the quality of service and control of congestion (primary role, provides feedback to the application about the quality of data distribution), identification of sources (sources are identified by means of randomly selected 32-bit identifiers), synchronisation of different media (contains data about the real time of the source data and packet time codes) and the scaling of the control information.

– *RSVP (Resource Reservation Protocol)* – protocol for the reservation of network resources, allows the receiving side to request certain quality of end-to-end service for its flow of data. Real-time applications use RSVP for the reservation of the necessary resources in network routers along the transmission route so that the requested frequency bandwidth would be really available once the transmission starts. The reservation requests originate from the receiver. This protocol is part of the Internet Integrated Service (IIS) of the model which provides delivery without real-time guarantee and controlled connection sharing.

– *SDP (Session Description Protocol)* – protocol for the description of sessions, founded by IETF (Internet Engineering Task Force).

### 3.1. Signalling Protocols

Signalling protocols in IP telephony are H. 323 (set of protocols) and SIP (Session Initiation Protocol). Signalling structures have to provide the set-up, control and release of calls and connections, scalability, quality of connection according to the requests of the end devices, flexibility, and network control[4].

– *H. 323* – a set of protocols which control the call set-up, encoding and decoding of multimedia contents, call flow and their release. The H. 323 standard consists of the following components and protocols: H. 225 –Call Signalling, H. 245 – Media Control, G. 711, G. 722, G. 723, G. 728, G. 729 – audio codec (coder/decoder), H. 261, H. 263 – video codec, T. 120 – Data Sharing. The H. 323 system elements are: terminals, gateways, service and control servers (gatekeepers) and multipoint control units (MCU).

The terminals provide point-to-point and multipoint conferences for audio data (option: video and data). They have to have a system control unit, media transmission, audio codec and packet-based network interface.

Gateway connects to PSTN or ISDN network, has the function of a translator between the audio, video and data transmission format, as well as the communication systems and protocols. Gateway is not necessary unless interconnection of SCN (Switched Circuit Network) is required.

Service and control server (gatekeeper) provides services of access control and address translation for terminals or gateways. It is logically separated from other H. 323 network elements.

Multipoint control units are devices which allow two or more terminals or gateways to hold conferences with audio and/or video sessions.

Although many implementations of H. 323 use today TCP as the transport signalling mechanism, H. 323 v. 2 allows basic UDP transport. There are three main areas of control: *RAS Signalisation (Registration, Admission and Status Signalling)* – providing pre-call control, *CCS (Call Control Signalling)* – a channel for call signalisation, used for establishing connections, maintenance and cutting off the call and *Media Control and Transport* – provides reliable H. 245 channel which transmits media control messages.

– *SIP (Session Initiation Protocol)* – signalling protocol used for creating, modifying, and terminating multimedia communications in IP networks. The telephone call is a type of multimedia communication, and the IP telephones with SIP customer, PCs with SIP telephone application or SIP gateways between IP and PSTN network are used as end terminals for creating such calls. SIP is a textual protocol founded at HTTP, it can use TCP or UDP as the transport protocol. SIP defines two types of messages: requests and responses, and all the messages consist of a header and a body. It contains six basic types of requests – methods: *Invite* (method for contracting the communication parameters), *Ack* (method for confirmation of a new connection), *Options* (method for obtaining information about the possibilities of the server), *Register* (informs the server about the current location of the user), *Cancel* (cancels parallel searches) and *Bye* (client is about to leave the session). When the server receives a request, it sends the response to the customer. Every type of response is assigned a certain code number. There are 6 main types of code answers (1xx, 2xx, ... , 6xx). When SIP architecture uses servers, there are two ways of their operation: proxy servers or redirect servers. There is a possibility to operate without using the servers.

# 4. INTERNET TELEPHONY IN CROATIA

One of the first implementations of IP telephony in Croatia is the call-centre of REGOS. Apart from telephoning over IP network, a sophisticated and complex system of a distributed call-centre on IP network was implemented.

The call-centre is a software application that in interaction with the telephone switch controls the telephone calls and collects and processes various statistics. Another important factor of the call-centre are the agents. These are the operators that receive telephone calls. The agents apply for operation at the call-centre application which forwards the calls to the agents. The application knows at any time which agent is free, which received least calls that day or had the longest break. The agents can also be grouped. Thus e. g. the caller calling from Split can be automatically routed to the agents that are responsible for the sales in the Dalmatian region. If all of them are occupied at that moment, the call can be rerouted to the free agent in the group "heads of sales" or to the supervisor.

Additional possibilities include recording of the conversation, conference connections, various real--time statistics, various reports, etc.

The agents who receive the calls are distributed in all the seven centres in Croatia, and the operative headquarters is at Regos. The same channels are used to switch calls and other important information such as the agent status, data about the caller, various statistics, even correspondence between the agents and the supervisor. Besides, simple and fast access to the central SQL database is provided.

Cisco solutions have been selected for this project. Apart from the call-centre, REGOS has introduced IP telephony into their own telephone system, where some thirty Cisco IP telephones are controlled by the Call Manager program application. At every location there is a router with VoIP card that establishes connection with the telephone switch. In this way the IP telephones can communicate with the classical telephone network. All the calls arrive in the call-centre where they are converted into IP packets and forwarded by a server with the Call Manager to the IPCC server and further to a free agent in the call-centre. Agents receive calls on their IP telephones or PCs and as necessary forward them over IP network to the called agent in the institution. Call-centre applications are related to the telephone switch to which they are connected.

The call-centre uses 20 ISDN channels and four agents which is also the maximum number of calls that can be received at a time, while other calls have to queue and are received when one of the agents is free. The connections between the locations are made by Frame relay lines of a speed of 128 kbit/s at dislocations, and in the centre the speed is 1 Mbit/s. Voice is carried by a coder G. 729 which compresses one telephone conversation to a speed of 8 kbit/s, adding the RTP/UDP/IP header (40 octets) that can be compressed to two octets. With the header of the Frame relay protocol the transfer band occupies 10.8 kbit/s which means that a dozen telephone conversations can be realised over a single connection of 128 kbit/s (as compared to the classical telephony of 64 kbit/s where two conversations are possible).

Apart from the call-centre the network serves all the telephone services thus realising substantial financial saving. If the need arises in the future, there is the possibility of data transfer thus achieving full integration of the systems among institutions.

# 5. CONCLUSION

The basic advantage of the IP telephony is the reduction of operative costs and the telecommunication costs. The IP telephony allows significant savings in the international telephone traffic towards fixed and mobile networks, as well as in the conversation between branch offices, either by using the Internet infrastructure or by using the private network (Virtual Private Network). The costs incurred by IP telephony can be as much as up to 50 percent lower compared to the classical telephony, with the unchanged quality of the connection. For the networks of companies the unification of the voice and data networks means saving because the infrastructure requires fewer additions, moves and changes (the Internet is available anywhere – the companies and branch offices do not depend on the locations), and the IP networks are easier for construction and maintenance.

At the same time the IP telephony increases the income of the service providers, keeping the existing users and attracting the new ones. The reduction of costs is reflected also through the expenditures in the future since the same infrastructure is used for data, voice and, later, video services not requiring the construction of a separate network infrastructure to introduce additional services.

One of the advantages of the IP telephony is its simple implementation, scalability and user-friendliness. The service implementation does not change the existing network condition and the telephone calls are made in the same way as before, and the users do not have to change their telephoning habits.

With future additional services and judging by the present trends, the IP telephony service will be widely used and available to almost any citizen of Europe, Africa and Central-Eastern countries, which will certainly lead to the increase in the service providers' incomes and a reduction in the users' costs.

**SLAVKO ŠARIĆ**, D. Sc.
Fakultet prometnih znanosti
Vukelićeva 4, 10000 Zagreb, Republika Hrvatska
E-mail: slavko.saric@fpz.hr
**ANTO BILOBRK**, B. Eng.
Cedevita d. o. o.
Planinska bb, 10000 Zagreb, Republika Hrvatska
E-mail: anto.bilobrk@cedevita.hr
**DAVOR NAĐ**, B. Eng.
Hrvatska televizija
Prisavlje 3, 10000 Zagreb, Republika Hrvatska
E-mail: dnad@email.htnet.hr

## *SAŽETAK*

### *PRIJENOS GOVORA IP MREŽAMA*

*Prijenos govora IP mrežama (Voice over Iinternet Protocol) predstavlja jedan od načina kojim se integriraju govorne i podatkovne mreže. Budući razvoj zasnovan je na eksploziji interneta kao načina komuniciranja, uz otvorenost standarda i spremnost proizvođača opreme da prihvate takav standard i unificiraju ga. Davatelji usluga nalaze interes uvođenja novih usluga koje nisu bazirane samo na prijenosu govora, već je govor samo jedna od aplikacija koja se ostvaruje putem IP mreža. Prijenos govora putem IP tehnologije trenutno nije na razini postojeće kvalitete usluge međutim nadolazeća rješenja u skoroj budućnosti omogućit će VoIP kao standardno operatorsko rješenje. Prednosti se očituju kroz: povećanje prihoda uz zadržavanje postojećih i privlačenje novih korisnika, investicije u infrastrukturu će maksimizirati prilike za razvoj usluga u paketu, osnažiti lojalnost korisnika i smanjiti operativne izdatke, usluge će biti široko rasprostranjene - međunarodne usluge na velike daljine ili nacionalne usluge pozivnih kartica (calling card) mogu biti pozicionirane u gotovo svakoj europskoj, afričkoj ili srednje-istočnoj zemlji i usmjerene na gotovo 100% stanovništva. Uz prednosti niskih početnih troškova novih POP-a (Points of Presence), pružatelji usluga mogu naglo proširiti svoje prisustvo na mnoge zemlje ili odredišta.*

### *KLJUČNE RIJEČI*

*IP telefonija, protokolarna arhiktektura, Call centar*

## LITERATURE

[1] **Taylor, T.**, Megaco/H. 248: *A New Standard for Media Gateway Control*, IEEE Communications Magazine, 2000.

[2] **Sinnreich, H., Johnston, A. B.**: *Internet Communications Using IP*, Wiley & Sons, Inc., New York, 2001.

[3] **Grbia, G., Pavelia, B.**: *Prijenos govora IP mrežama*, ETK Revue 13(2001)1, Zagreb, 2001

[4] **Kumar, V., Korpi, M., Sengodan, S.**: *IP Telephony with H. 323*, Wiley & Sons, Inc., New York, 2001.

[5] **Sloane, A., Lawrence, D.**: *Multimedia Internet Broadcasting*, Springer, 2001.

[6] **Knight, R. R., Norreys, S. E., Harrison, J. R.**: *Bearer-independent call control*, BT Technology Journal Vol. 19, No. 2, 2001

[7] http://www.cisco.com

[8] http://www.nexcom.hr

[9] http://www.vm-mreže.hr

[10] http://www.ericsson.com

*Figures made by the program VISIO 2003. Used Internet pages under [7], [8], [9], [10].