

# Network dimensioning for voice over IP

Tuomo Hakala

tuomo.hakala@kolumbus.fi

## Abstract

This article concentrates in the issues of network dimensioning for voice over IP (VoIP). The network under dimensioning is an IP network between VoIP user devices. First, a short introduction to VoIP in general is given. Second, the issues in network dimensioning for VoIP are identified. Third, bandwidth requirements of VoIP are calculated. Fourth, basic approaches to Quality of Service are discussed and finally conclusions are drawn.

## 1. Introduction

VoIP represents the best opportunity so far for voice and data convergence and it is now one of the fastest-growing industries [12]. An IP network with mixed voice and data makes the network management easier than managing separate voice and data networks. A VoIP call uses less bandwidth than a circuit-switched call. VoIP makes new services possible.

IP networks, like the current Internet, offering only best-effort service, cannot satisfy the Quality of Service (QoS) requirements of VoIP. This is primarily because of the variable queuing delays and packet loss during network congestion [12].

The end-to-end Quality of Service of VoIP is composed of factors related to the network and factors related to the applications. Factors related to the network are [13] [1]:

- Network delay
- Network jitter
- Network packet loss and desequencing

Factors related to the applications are:

- Overall packet loss
- Jitter buffers
- Codec performance

- Overall delay

Currently there are several approaches to improve the audio quality of VoIP [7]:

- **Integrated Services (IntServ)** is a stateful approach where resources are reserved in the network before data starts to flow along the reserved path. [8]
- **Differentiated Services (DiffServ)** is a stateless approach where real-time traffic is marked to get preferred treatment in the network. [9]
- **Forward Error Correction (FEC)** algorithms reduce the impact of data loss by sending redundant data along with the audio data. The redundant data helps to reconstruct lost data. [10] [14]
- **Loss Concealment** algorithms try to reduce the impact of data loss by replacing the lost audio with an approximation. [11]

Forward Error Correction and Loss Concealment algorithms are methods used in the VoIP user devices. IntServ and DiffServ are methods used in the IP network.

## 2. The issues

During an average conversation, each party usually talks only about 35 percent of the time. Most of the techniques used to transform voice into data, the codecs, have the ability to detect silences. With this voice activity detection, data is transmitted only when needed. When several conversations are multiplexed on a single transmission line, statistical multiplexing can be used which leads to more efficient use of **bandwidth**.

When a VoIP packet is transferred through an IP network, it will experience **delay** that is caused by:

- Transmission delay between the nodes, depends on the frame size and the transmission speed
- Queuing delay in the nodes because of buffering

- Switching and processing delay in the nodes, the time to switch a frame from an input port to an output port
- Propagation delay, depends on the characteristics of the transmission media and the distance between the nodes

The use of statistical multiplexing means that the delay of sent packets within a conversation will vary. This varying delay is called **jitter**. The jitter must be minimized in the network and the remaining jitter needs to be corrected by the receiving side using jitter buffers to make the original speech intelligible. Jitter buffers increase the overall delay.

Several technologies enable the use of statistical multiplexing and mixing of voice and data on the same transmission lines. Such technologies are voice over frame relay, voice over ATM and VoIP. VoIP is the most flexible technology because it does not require virtual channels to be set up between the parties having a conversation. Also, VoIP scales in terms of connectivity much better than frame relay or ATM.

In IP networks, routers are the devices that execute the statistical multiplexing functionality. IP packets belonging to the same conversation may use different routes having different delays and therefore they may be received in different order than in which they were sent. This is called **desequencing**.

When an overflow of the buffers in the network nodes occurs because of network congestion, there will be some **packet loss**, which must be handled by the receiving side. It makes no sense to resend part of speech because of the overall delay. [1]

The bandwidth required by VoIP must be calculated considering the bandwidth requirements of a single conversation and the number of conversations on each link in the network. Acceptable packet loss and the buffering capacity of the nodes in the network must be considered as well. Delay and jitter must be minimized in the network.

The receiving side must take care of the remaining network jitter and the desequencing of packets. The Real-time Transport Protocol (RTP) was designed to allow the receiver to do the correction [4]. RTP includes:

- Information on the type of data transported
- Timestamps
- Sequence numbers

Real-Time Control Protocol (RTCP) allows the conveyance of feedback on the quality of the transmission and it can also carry information on the identity of the participants [4]. RTP and RTCP are mostly used on top of User Datagram Protocol (UDP) [3], which provides the use of a port number and a checksum. The use of UDP enables also the use of IP multicast i.e. sending packets to IP multicast addresses. This means that a RTP stream generated by a single source can be received by several destinations. [1]

### 3. Bandwidth requirements

Vocoders that support Voice Activity Detection (VAD) use less bandwidth in silence periods (m) than in active speech periods (M). M and m values for popular coders are shown in Table 1 [1].

Table 1: M and m values for popular codecs

Codec	M (kbit/s)	m (kbit/s)
G.723.1 (5.3 kbit/s)	8	3,73
G.723.1 (6.4 kbit/s)	9,07	3,73
Lucent SX7003P	20,27	13,87

The M values in Table 1 include transport overheads of RTP, UDP and IP headers that are shown in Table 2 [1].

Table 2: Transport overheads

Protocol	Overhead (octets)
IPv4 (Internet Protocol version 4) [2]	20
UDP (User Datagram Protocol) [3]	8
RTP (Real-time Transport Protocol) [4]	12

Activity rate a, in a one-way network bitrate during a voice conversation, is the proportion of active speech intervals of the whole time of the conversation. An average value is usually 0,35 but to be on the safe side a=0,5 should be used in the calculations.

Average bitrate of a single one-way voice channel during a conversation is

$$\text{Average\_bitrate} = Ma + m(1-a) \quad (1)$$

where

M = active bitrate (kbit/s)

m = silence bitrate (kbit/s)

a = activity rate (%)

For N simultaneous conversations using the same coder and with no buffering requirements for the nodes in the network, the average one-way bitrate is

$$N * \text{Average\_bitrate} = N(Ma + m(1-a)) \quad (2)$$

This formula gives the bandwidth needed when zero packet loss is required.

Buffering in the network increases jitter and therefore reduces interactivity. It is good practice to dimension VoIP links considering that there is no buffering in the network. This leads to some overprovision for slow links, but this overhead can be used by non real-time traffic in an IP network designed for both voice and data [1]. In an IP network with mixed voice and data the bandwidth requirements of VoIP are small compared to the bandwidth used for data in today's IP networks.

When bandwidth needs to be reserved for voice in an IP network designed for both voice and data, information needs to be gathered in order to know who phones where, how often and how long. When an existing circuit switched telephone network is planned to be realized by using VoIP, this information can be derived from existing phone bills during a reference period [1]. When VoIP network is planned for a new implementation and no statistics or phone bills are available, calculations of the voice traffic can be done using the Erlang model [1].

An optimal route on the network is chosen for each of the calls considering the cost of each link per unit of bandwidth. After this, it is possible to calculate the number of simultaneous calls on each link at any given time. The peak number of simultaneous busy hour calls is used to calculate the needed link bandwidth for zero packet loss using the formula (2).

## 4. Delay, jitter and packet loss

When the bandwidth required by VoIP is calculated for zero packet loss and no buffering is assumed in the

network, the network delay and jitter are minimized. In an IP network with mixed voice and data traffic, some mechanism must be used to ensure that the bandwidth calculated for VoIP is not used by other real-time traffic or non real-time traffic. When calculations for VoIP are done assuming zero packet loss in the network, somehow it must be taken care of that the buffers in the network nodes are not filled with packets of other traffic types which would cause VoIP packets to be dropped causing packet loss. Also, when the calculations for VoIP are done assuming that there is no buffering in the network nodes, because buffering would lead to increased delay and jitter, it must be somehow taken care of, that VoIP packets get sent first to the outgoing link even though there are packets of other traffic type in the buffers.

First of all, VoIP traffic must be somehow differentiated from other traffic types in the network so that it can be treated better. The nodes in the IP network, the routers, can differentiate traffic according to source and destination IP addresses, protocol type, port numbers and by the Differentiated Services (DS) field. The DS field means the type of service (TOS) byte in IPv4 and the traffic class byte in IPv6.

There are two basic approaches in an IP network with mixed voice and data traffic that can be used to improve the quality of VoIP [7]:

- Integrated Services (IntServ) is a stateful approach where resources are reserved in the network before data starts to flow along the reserved path. [8]
- Differentiated Services (DiffServ) is a stateless approach where real-time traffic is marked to get preferred treatment in the network. [9] [5]

### 4.1. Integrated Services (IntServ)

IntServ model proposes two service classes in addition to best-effort service: guaranteed service and controlled-load service. Guaranteed service is for applications requiring a fixed delay bound. Controlled-load service is for applications requiring reliable and enhanced best-effort service. [12]

IntServ requires that resources are explicitly managed for each real-time application. Routers must reserve resources (e.g. bandwidth and buffer space) in order to provide specific QoS for each packet flow. This requires flow-specific states in the routers. [12]

The four components of IntServ are:

- **Flow specification** - *Flowspec* describes the characteristics of the flow and it has two separate parts, *Tspec* (describes flow's traffic characteristics) and *Rspec* (specifies the service requested from the network)
- **Signaling protocol** - e.g. Resource ReSerVation Protocol (RSVP) [6]
- **Admission control routine** - determines whether a request for resources can be granted.
- **Packet classifier and scheduler** - packets entering a router are classified and put in the appropriate queue and then scheduled accordingly.

## 4.2. Differentiated Services (DiffServ)

In DiffServ model traffic entering an IP network is classified, marked, policed and shaped at the edge of the network.

The packets are then assigned to different behavior aggregates (BA). Each BA is identified by a single DiffServ CodePoint (DSCP). Users request a specific performance level per packet by marking the DiffServ field of each packet with a specific DSCP value which specifies the Per-Hop-Behavior (PHB) within the provider's network. Packets are forwarded within the core of the network according to the PHB.

The four components of DiffServ are [12]:

- **Services** - Characteristics of packet transmission in one direction over a path in a network are defined by a service. DiffServ can be provided by two approaches:
  - **Quantitative DiffServ** - QoS is specified in deterministically or statistically quantitative terms of throughput, delay, jitter and/or loss.
  - **Priority based DiffServ** - Services are specified in terms of a relative priority of access to network resources.
- **Conditioning Functions and PHB** - A user and a service provider must have a service level agreement (SLA) in place that specifies the supported service classes and the amount of traffic allowed in each class. Individual packets have DiffServ (DS) fields that indicate the desired service and these DS fields can be marked at hosts or at the access router or at the edge router in the service provider network. Packets are classified, policed and possibly shaped at the ingress of the service

provider network according to the rules derived from the SLA. Between domains, service provider networks, DS fields may be remarked, if so defined in the SLA between the two service providers. These traffic control functions at hosts, access routers or edge routers are generically called traffic conditioning. Per hop behavior (PHB) are defined to allocate buffer and bandwidth resources at each node among traffic streams. PHB is applied to a DiffServ behavior aggregate and a DiffServ-compliant node.

- **DS CodePoint** – DS field means the type of service (TOS) field in IPv4 and the traffic class byte in IPv6. Six bits of this DS field are used as a codepoint (DSCP) to select the PHB for a packet at each node.
- **A node mechanism for achieving PHB** – Buffer management and packet scheduling mechanisms are used in nodes to achieve a certain PHB. PHBs are defined as behavior characteristics relevant to service provisioning policies instead of particular implementation mechanisms. Various implementation mechanisms may be suitable for a particular PHB group.

## 5. Conclusions

The issues to be considered in network dimensioning for VoIP are bandwidth, delay, jitter, desequencing and packet loss. The bandwidth required by VoIP must be calculated considering the bandwidth requirements of a single conversation and the number of conversations on each link in the network. Acceptable packet loss and the buffering capacity of the nodes in the network must be considered as well. When the bandwidth required by VoIP is calculated for zero packet loss and no buffering is assumed in the network, the network delay and jitter are minimized. The receiving side must correct the remaining network jitter and the desequencing of packets. In an IP network with mixed voice and data traffic, some mechanism must be used to ensure that the bandwidth calculated for VoIP is not used by other real-time traffic or non real-time traffic. There are two basic approaches to achieve this: Integrated Services (IntServ) and Differentiated Services (DiffServ). IntServ is a stateful approach where resources are reserved in the network before data starts to flow along the reserved path. DiffServ is a stateless approach where real-time traffic is marked to get preferred treatment in the network.

## References

- [1] Hersent, Olivier; Gurle, David; Petit, Jean-Pierre: IP Telephony, Packet-based multimedia communications systems; Great Britain, 2000, [www.awl.com/cseng/](http://www.awl.com/cseng/), ISBN 0-201-61910-5
- [2] Postel, Jon: Internet Protocol, RFC 791, September 1981
- [3] Postel, Jon: User Datagram Protocol, RFC 768, 28 August 1980
- [4] Schulzrinne, Henning; Casner, Stephen L.; Frederick, Ron; Jacobson, Van: RTP: A Transport Protocol for Real-Time Applications, RFC 1889, January 1996
- [5] Blake, Steven; Black, David L.; Carlson, Mark A.; Davies, Elwyn; Wang, Zheng; Weiss, Walter: An Architecture for Differentiated Services, RFC 2475, December 1998
- [6] Mankin, A.; Baker, Fred; Braden, Bob; Bradner, Scott; O'Dell, Michael; Romanow, Allyn; Weinrib, Abel; Zhang, Lixia; Resource ReSerVation Protocol (RSVP), Version 1 Applicability Statement, Some Guidelines on Deployment, RFC 2208, September 1997
- [7] Trends in the Internet Telephony, <http://www.fokus.gmd.de/research/cc/glone/projects/ipt/> (11 March 2001)
- [8] IETF Integrated Services (IntServ) Working Group charter, <http://www.ietf.org/html.charters/intserv-charter.html> (11 March 2001)
- [9] IETF Differentiated Services (DiffServ) Working Group charter, <http://www.ietf.org/html.charters/diffserv-charter.html> (11 March 2001)
- [10] Speech Property-Based FEC (SPB-FEC), <http://www.fokus.gmd.de/research/cc/glone/products/voice/spb-fec/> (11 March 2001)
- [11] Adaptive Packetization / Concealment (AP/C), <http://www.fokus.gmd.de/research/cc/glone/products/voice/apc/> (11 March 2001)
- [12] Li, Bo; Hamdi, Mounir; Jiang, Dongyi; Cao, Xi-Ren: QoS-Enabled Voice Support in the Next-Generation Internet: Issues, Existing Approaches and Challenges; IEEE Communications Magazine, April 2000
- [13] TELECOMMUNICATIONS AND INTERNET PROTOCOL HARMONIZATION OVER NETWORKS ETSI PROJECT – TIPHON, [http://webapp.etsi.org/tbhomepage/TBDetails.asp?TB\\_ID=291&TB\\_NAME=TIPHON](http://webapp.etsi.org/tbhomepage/TBDetails.asp?TB_ID=291&TB_NAME=TIPHON) (12 March 2001)
- [14] Padhye, Chinmay; Christensen, Kenneth J.; Moreno, Wilfrido: A New Adaptive FEC Loss Control Algorithm for Voice Over IP Applications; IEEE 2000