VoIP over Satellite

An EMS Technologies Canada Technical Notes

Revision 1-2

Change Record

Revision	Date	Description
1-0	4 December 2003	First Release with the collaboration of R.Green and R. Desloges
1-1	17 February 2004	Updates following J. Landovskis's comments
1-2	28 May 2004	Updated following internal EMS reviews

1 INTRODUCTION

Interest in Voice over IP (VoIP) has increased steadily over the past few years. Many organisations view VoIP as a viable way to implement packet voice in terrestrial network, it is the case as well for a satellite network. They want to use/implement VoIP because their network can be consolidated, service can be converged and VoIP provides toll-bypass. The network is consolidated, e.g. voice, data and video are carried over a single network infrastructure, and therefore simplifies the network management. Service convergence allows enhanced functionality to be implemented through the coupling of multimedia services. Finally, the toll-bypass allows long-distance calls to be placed with no charge.

2 VOIP OVER SATELLITE ISSUES

Many aspects of the Voice over IP system over satellite have to be considered. The following list presents some issues of the VoIP over satellite:

- The codec used for VoIP should use the minimum bandwidth with acceptable quality (less then 16 Kbps including all overheads).
- The VoIP data should be prioritised from the other IP traffic data, because the voice over IP is real-time and jitter intolerant.
- The overhead of the VoIP packets should be reduced.
- The VoIP solution should be transparent to the use of private or public IP Addresses.
- It should be possible to use VoIP in a satellite system supporting 2 satellite hops.
- The VoIP should work in a congested network.

3 VOIP COMPONENTS

It is important to consider all the factors that will affect voice quality. The codecs, voice activity detection (VAD), source delay, delay budget and voice quality are the main factors to consider when designing a VoIP network. The VoIP Protocols will be defined to understand better the bandwidth used in a Voice stream.

3.1 VOIP CODECS

The codecs are used to digitalize and compress the analog voice. Coding techniques for telephony and voice packet are standardized by the ITU-T in its G-series recommendations. The following table lists some coding standards.

Table 1 Codecs Voice Bit Rate

Coding standard	Voice Bit Rate (Kbps)
G.711 PCM	64
G.726 ADPCM	32
G.728 LD-CELP	16
G.729 CS-ACELP	8
G.729 x 2 Encodings	8

Coding standard	**Voice Bit Rate (Kbps)**
G.729 x 3 Encodings	8
G.729a CS-ACELP	8
G.723.1 MP-MLQ	6.3
G.723.1 ACELP	5.3

The column Voice bit Rate represents only the voice bandwidth, which does not take into account the IP/UDP/RTP overheads included in all voice frames. On top of this overhead, for satellite communications, the overhead of DVB-RCS is added for the return link and the overhead of the MPEG over DVB-S is added on the forward link.

3.2 VOICE ACTIVITY DETECTION (VAD) AND COMFORT-NOISE GENERATION (CNG)

A feature called VAD, which is built into VoIP codec, causes the gateway to transmit when speech starts and cease transmitting when speech stops. During silences, it generates white noise so those callers do not mistake silence for a disconnected call. By suppressing packets of silence, VAD enables you to handle more calls by minimising bandwidth requirements. For VoIP bandwidth planning, VAD typically reduces bandwidth requirements by 35 percent. Enable VAD if you wish to allocate more bandwidth to other types of traffic. In this case, when VAD is enable, the Comfort-Noise Generation prevents uncomfortable dead silence on the receiving end.

3.3 VOICE DELAYS

There are two distinct types of delay: fixed and variable. Fixed delay components add directly to the overall delay on the connection. Variable delays arise from queuing delays in the buffers. These buffers create delays, called jitter, across the network. Variable delays are handled via the playout jitter buffer at the receiving VoIP Gateway. Fixed delays are the compression algorithmic delay, packetization delay, serialization delay (which is negligible), propagation delay and the Jitter delay at the VoIP Gateway. The variable delays are the queuing/router delay (which is negligible) and DVB-RCS delay.

3.4 VOICE QUALITY

Voice Quality is defined by the Mean Opinion Score (MOS), Perceptual Speech Quality Measure (PSQM) and Perceptual Evaluation of Speech Quality (PESQ). Described in ITU-T P.800, MOS is the most well-known measure of voice quality. It is a subjective method of quality assessment. MOS is the most relevant test, because it is humans who use the voice network and it is humans whose opinions count. They then rank the voice quality using the following scale: 5 – Excellent, 4 – Good, 3 – Fair, 2 – Poor, 1 – Bad. The following table shows the relationship between codecs and MOS scores.

Table 2 Codecs MOS

Coding standard	**MOS**[1]
G.711 PCM	4.1
G.726 ADPCM	3.85

[1] MOS values from a terrestrial network.

Coding standard	**MOS**[1]
G.728 LD-CELP	3.61
G.729 CS-ACELP	3.92
G.729 x 2 Encodings	3.27
G.729 x 3 Encodings	2.68
G.729a CS-ACELP	3.7
G.723.1 MP-MLQ	3.9
G.723.1 ACELP	3.65

Described in ITU-T P.861, the PSQM algorithm uses a psycho-acoustic model that aims to mimic the perception of sound in real life. The algorithm functions by comparing the original signal with the signal that was sent and received. PSQM provides an output in the range 0 to 6.5, where 0 indicates a good channel, and 6.5 indicates a very bad channel. If the input and output are identical, the algorithm is designed to produce a perfect score. Similarly, the objective is that if the input and output have inaudible difference, the score should not be degraded.

The PESQ algorithm uses a psycho-acoustic model that aims to mimic the perception of sound in real life. The algorithm functions by comparing the original signal with the signal that was sent and received. PESQ provides an output in the range –0.5 to 4.5, although in most cases the output range will be between 1.0 and 4.5. A score of 4.5 indicates a good channel, and 1.0 indicates a poor channel.

3.5 VOIP PROTOCOLS

Voice over IP uses the Real Time Protocol (RTP, RFC 3550) to transport the voice. This protocol uses an IP Header and UDP header to transport the RTP data. The Real Time Control Protocol (RTCP) is used to monitor the quality of the real-time session. The Session Initiation Protocol (SIP) is defined by IETF, and described in the RFC 3261. The SIP is an application layer control (signalling) protocol for creating, modifying and terminating sessions with one or more participants. The H.323 protocol is an International Telecommunication Union (ITU) standard. The H.323 comes from existing standards, i.e. H.320 and Q.931 (used for ISDN network). The H.323 was defined to facilitate the interoperability between ISDN and new technologies vendors. The H.323 protocol includes the H.225 Registration, Authentication and Status (RAS), H.225 Control protocol and the H.245 Control protocol. There is as well the Media Gateway Control Protocol (MGCP) which is another signalling protocol, it is the standard for cable network, and it is popular with Voice over ATM. EMS has tested VoIP with the H.323 protocol used with RTP and RTCP.

4 EMS VOIP ARCHITECTURE

The Figure 1 illustrates the EMS VoIP over Satellite architecture. This VoIP architecture uses a VoIP Gateway for H.323 protocol. This VoIP architecture allows phone calls from PSTN to satellite phone, from satellite phone to PSTN and from satellite phone to satellite phone.

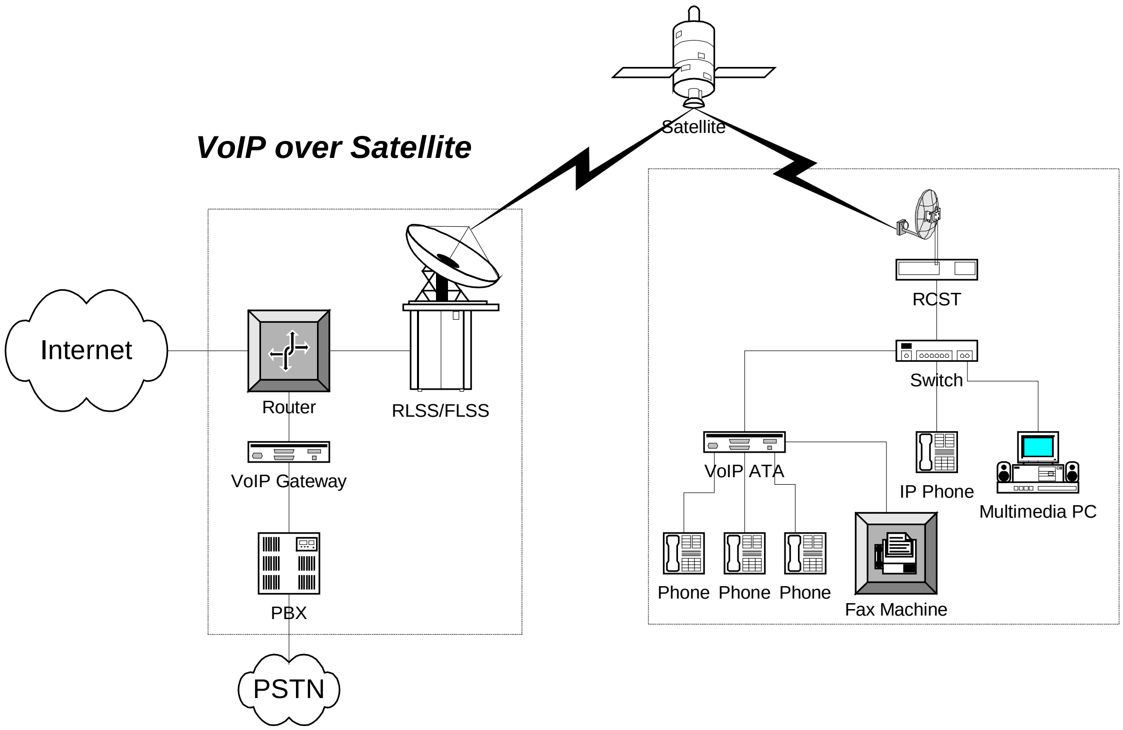

Figure 1 EMS VoIP Architecture

To support the VoIP over Satellite Architecture, many sub-systems are involved. The following table lists all these sub-systems:

Table 3 VoIP Sub-systems

Sub-systems	Features
Multimedia PC	PC which emulate a phone
IP Phone	VoIP Phone which support H.323 (in this example)
VoIP Analog Telephone Adapter (ATA)	Handset to Ethernet adapter that interfaces regular analog phones with IP based telephony networks. Converts voice into IP data packets. Includes Ethernet RJ-45 port and some (2-4 or more) RJ-11 Foreign Exchange Subscribers (FXS) ports. Supports many codecs (G.711, G.723, G.729, etc). Can be configured to use either H.323 or SIP call signalling protocol. Supports Type of Service (ToS) bit for Quality of Service (QoS), Supports Voice Activity Detection (VAD) with on/off control and supports Comfort Noise Generation (CNG)
Phones	Regular Analog phones
Fax	Regular Fax machine
Ethernet Switch	Regular Ethernet switch with 10/100 Fast Ethernet ports
RCST	Interface to the Customer LAN and to the satellite network.

Sub-systems	Features
Return Link Sub-System (RLSS)	Controls and manages the access of the Satellite Terminals on the return link. Receive upstream traffic from the Satellite Terminals. Support a scheduler for guaranteed and dynamic capacity assignment.
Forward Link Sub-System (FLSS)	Controls and manages the access of the Satellite Terminals on the forward link. Transmit downstream traffic to Satellite Terminals.
VoIP Gateway	Integrates Voice and data networking. Supports H.323 and SIP. Supports many codecs (G.711, G.723, G.726, G.728, G.729). May include a complete range of VoIP interfaces ((FXS), Foreign Exchange Office (FXO)) and Ethernet RJ-45 port.
PBX	Regular PBX used to connect to the Public Switched Telephone Network (PSTN)

The terminal has to support multiple queues to prioritise VoIP data over other IP data. The Router or the IP-DVB Encapsulator in the FLSS can be used to prioritise the VoIP data over other IP data in the forward link.

To meet the requirement of using less than 16 Kbps for the VoIP stream and to reduce the overhead, only the G.723 can be used. On the return link, the G.723 codec can be configured with one, two, three or four voice frames per IP packet. Please refer to Table 5 for the details. On the forward link, the G.723 codec has to be configured with three voice frames per IP packet. Please refer to **Table 6** for the details.

To develop a VoIP solution with the use of private and public IP Address, the NAT device has to support the different VoIP Protocols.

The selected VoIP ATA has to support the Voice Activity Detection (VAD) feature to allow a utilisation of capacity for other IP data when there is no voice stream to send. The VAD feature is important for scenarios with guaranteed capacity on the return link. The selected VoIP ATA has to support the capability to disable the Voice Activity Detection for scenarios with non-guaranteed capacity on the return link.

5 VOIP IN DVB-RCS

There are a number of capacity request mechanisms defined in DVB-RCS. EMS supports the following: Guaranteed Capacity assignment called Constant Rate Assignment (CRA), Dynamic Capacity assignment based on volume, called the Volume Based Dynamic Capacity (VBDC), Dynamic Capacity assignment based on rate, called the Rate Based Dynamic Capacity (RBDC) and Dynamic Capacity assignment based on the free capacity left, called Free Capacity Assignment (FCA). Each type has unique properties and is used for different types of traffic.

Table 4 DVB-RCS Assignment

Assignment Type	Traffic Properties
CRA	Real-time jitter intolerant such as VoIP
RBDC	Near real-time jitter tolerant such as streaming video
VBDC	Non real-time applications such as FTP
FCA	Non real-time applications and opportunistic request mechanism

The CRA is reserved and guaranteed at the logon time of the Terminal. The Terminal does not need to ask for the CRA. The RBDC is guaranteed and reserved at the logon time of the Terminal. The Terminal has to dynamically request RBDC slots, depending on the output rate of data to transmit over the satellite. The VBDC is not guaranteed and not reserved at the logon time of the Terminal. The Terminal has to dynamically request VBDC slots, depending on the number of packets to transmit. The Terminal does not request FCA. The RLSS gives FCA to terminals if there are unused slots available and terminals allowed FCA.

5.1 BANDWIDTH UTILISATION

The following table shows the VoIP bandwidth calculation (including overhead and DVB-RCS) for the return link.

Table 5 Return Link Bandwidth Calculation

Codec	MOS (Max 5)[2]	No. Voice Frames	Frame Duration (secs)	Frame Size (bytes)	Voice BW, bits/sec	Headers/Trailer (bytes)				Total (Frame size + Headers) (Bytes)	No. ATM cells	No. Cells Per sec.	Total BW on the Return Link (Kbps)	Return Link Inf. Rate (Kbps)
						RTP	UDP	IP	AAL Trailer					
G.723 ACELP	3.65	1	0.030	24	6400	12	8	20	8	72	2	66.67	28	32
G.723 ACELP	3.65	2	0.060	48	6400	12	8	20	8	96	2	33.33	14	16
G.723 ACELP	3.65	3	0.090	72	6400	12	8	20	8	120	3	33.33	14	16
G.723 ACELP	3.65	4	0.120	96	6400	12	8	20	8	144	3	25.00	10	16

[2] MOS value from a terrestrial network.

The following table shows the VoIP bandwidth calculation (including overhead and MPEG over DVB-S) for the forward link.

Table 6 Forward Link Bandwidth calculation

Codec	MOS (Max 5)[3]	No. Voice Frames	Frame Duration (secs)	Frame Size (bytes)	Voice BW, bits/ sec	Header/Trailer (Bytes)							Total (Frame Size + Headers) (Bytes)	Total BW on the Forward Link (Kbps)
						RTP	UDP	IP	LLC-SNAP	DSM-CC	CRC	MPEG		
G.723 ACELP	3.65	1	0.030	24	6400	12	8	20	8	13	4	4	93	50
G.723 ACELP	3.65	2	0.060	48	6400	12	8	20	8	13	4	4	117	25
G.723 ACELP	3.65	3	0.090	72	6400	12	8	20	8	13	4	4	141	16

6 VOIP TESTS

The VoIP Tests have been performed with guaranteed bandwidth and non-guaranteed bandwidth on the return link. The forward link was configured with a guaranteed bandwidth and a maximum shared bandwidth. In all tests, there was no congestion on the return and forward links. The EMS RLSS with the EMS SIT have been used for all the tests. The Transmission rate was 240 Kbps on the return link.

6.1 VOIP TESTS WITH GUARANTEED BANDWIDTH (CRA)

For this test with guaranteed and non-guaranteed capacity, the G.723 codec was used with four voice frames per IP packet, VAD was not enable to have a constant flow on the return link. The following table summarises the VoIP Tests with guaranteed bandwidth.

Table 7 VoIP Tests with guaranteed bandwidth (CRA)

TEST NO.	CRA (Kbps)	VBDC (Kbps)	Nb phone call (s) per RCST	Result (phone call)	Bandwidth on the Return link[4] (Kbps)	Return Link Inf. Rate, (Kbps)
1	16	224	1	Very Good	10	16
2	16	224	2	Very Good	21	32

The RCST was configured with an inactivity timeout set to 60 sec. After 60 sec of inactivity (no data to send over the satellite), the return link is released. As soon as there is traffic to send on the return link, the RCST re-acquires the return link and it takes around 3-4 sec before to send the first IP packet over the satellite.

[3] MOS value from a terrestrial network.
[4] Refer to the Table 5 for the detail.

6.2 VOIP TESTS WITH NON-GUARANTEED BANDWIDTH (VBDC)

For this test with only non-guaranteed capacity, the G.723 codec was used with two voice frames per IP packet, VAD was not enable to have a constant flow on the return link, which is required to avoid big jitter. The following table summarises the VoIP Tests with non- guaranteed bandwidth.

Table 8 VoIP Tests with non-guaranteed bandwidth (VBDC)

TEST NO.	CRA (Kpbs)	VBDC (Kbps)	Nb phone call (s) per RCST	Result (phone call)	Bandwidth on the Return link (Kbps)	Return Link Inf. Rate, (Kbps)
3	0	64	1	Good	14	16
4	0	64	2	Good	28	32

6.3 VOIP TESTS WITH NON-GUARANTEED BANDWIDTH (VBDC) AND DATA

EMS performed some VoIP tests with data sent and received at the same time. The Voice over IP has to be prioritised over the IP traffic data. In the return link, the RCST has to support prioritisation on queues, based on filters. On the forward link, the router has to support prioritisation queues (or the IP-DVB Gateway). The RCST was configured with two queues, one for the voice and one for the regular IP traffic.

The Codec used is G.723.1 with 2 voice frames per IP packet. Voice Activity Detection has to be disabled to generate constant VoIP stream. A constant VoIP stream is required to avoid big jitter.

Table 9 VoIP Tests with non-guaranteed bandwidth (VBDC) and data

NO	CRA / RBDC (Kbps)	VBDC (Kbps)	Nb phone call (s) per RCST	Result (phone call)	Bandwidth on the Return link (kbps)	Return Link Inf. Rate, (kbps)	Download 1 MB	Upload 1 MB
5	0	240	1	Good	14	16	720 Kbps	100 Kbps

7 CONCLUSION

This document presents a summary of the various VoIP Components, VoIP Protocols, Codecs, VAD, Source delays, Voice quality to understand the VoIP terminology. A VoIP architecture has been proposed and EMS used this VoIP architecture to execute some VoIP Tests over one satellite hop.

The VoIP tests results have been presented with preliminary VoIP tests results with CRA and VBDC, and VBDC only. The first VoIP test result with guaranteed bandwidth (CRA and non-guaranteed bandwidth VBDC) was good, but the problem is that we have to guarantee bandwidth (CRA) during all the session. The VoIP test result with non-guaranteed bandwidth only was good. In all tests, EMS recommends to use the G.723 codec with two voice frames.