Interoperability: voice and audio standards for Space missions

Dr. Osvaldo Peinado
German Space Operations Center (GSOC)
Oberpfaffenhofen
82234 Wessling
Germany

Voice communication and audio data exchange are essential services for supporting space mission operations. They are especially important in a cooperative international environment in human space flight programs like the International Space Station (ISS). Between humans, one of the most common information exchanges is voice communication. In human space flight, the primary function of voice communication is to support the mission executed by astronauts, cosmonauts, taikonauts, and other space travelers.

The international voice and audio industries have many standards and interfaces for acquiring, recording, and distributing live and recorded audio. That flexibility can lead to complexity when attempting to share or monitor voice or audio from acquisition to monitoring or recording locations. Voice communications was once the realm of analog technology; now voice is encoded and exchanged in digital form. The digital voice systems have advantages but also disadvantages. Voice and audio communications in space missions serve to the international supporting ground teams as well as the astronauts. The astronauts live and work in space facilities developed and launched by different countries. The ground supporting teams belong to different space agencies and private companies as well. For mission safety and success, interoperability of voice communications is crucial.

The Consultative Committee for Space Data Systems (CCSDS) decided that was necessary to have a document to define the standard for communications between control centers in different parts of the world, between spacecrafts, spacecraft and the control centers, between cosmonauts, taikonauts and astronauts wearing different space suites to the spacecraft and to the control centers.

This paper describes the setup of voice communication in a space mission context, points out special requirements and operational approaches, and defines the transmission, coding, interface, and quality parameters needed for space mission support. It provides system designers with a subset of the larger industry set of standards from which to choose, depending on the application and purpose of the voice system.

Nomenclature

BW	=	Bandwidth
CODEC	=	Code-Decode
Col-CC	=	Columbus Control Center
HOSC	=	Huntsville Operations Support Center
MCC-M	=	Mission Control Center Moscow

MCC-H = Mission Control Center Houston
PTT = Press to talk
QoE / QX = Quality of Experience, the subjective measure of a user experience with a service.
QoS = Quality of Service
UHF = Ultra High Frequency
VoIP = Voice over IP

I. Introduction

THE big difference between a phone call or a teleconference and a voice system is the usage of voice loops, where a user can have several loops simultaneously using roles to access those voice loops having to use PTT.
The first goal of an international standard is to promote interoperability and cross support among cooperating space agencies to reduce operation costs by sharing facilities and experiences.
The second goal is to reduce the cost to the various agencies of performing common functions by eliminating unjustified project-unique design and development.
A Mission Control Centers (MCC) may have different voice configurations.
- Several space agencies operating different voice system communicating to each other into one facility (MCC-M, COL-CC) and to externals.
- One space agency using several voice system supporting different aspects of the same mission (HOSC) and supporting externals communications.
- One space agency with one voice system communicating with other MCC having a different voice systems.
- Space agencies with one voice system using other MCC voice system to communicate with astronauts in Space.

All these different configurations should work together and it is here were a Standard fulfills its objective.
An example of that case is the common development of the voice system used now in all the NASA and European Space centers related to the ISS program.
By having interoperable and compatible voice systems, intensive testing and voice communications issues can be avoided, thus reducing costs and improving reliability.

II. General description

Voice is considered mission critical. That means its have the highest priority in a Space Mision. To create a new international standard for a voice system used for human space flight is very challenging. With the rising costs of space operations and the budget limitations of all the spaces agencies, the international cooperation is the only way to go. In the ISS or in an outpost in the Moon or Mars, it is mandatory that taikonauts, cosmonauts and astronauts can talk to each other's . Also the different control centers in Japan, China, Korea, Germany, America and Russia need to talk to each other. Now, the digital technologies are replacing the current and established communications.

Each country wants to support their own industry and new national standards. Besides of these national interest the politics still playing a big role. To put on the table Europe, China, Russia and the USA and to get a consensus of how to define a new voice standard is not an easy task.
Voice and audio communications in space missions serve to the supporting personnel and the astronauts as well. The astronauts live and work in space facilities developed by different countries. The supporting personnel belong to different Space Agencies and private companies as well. For mission safety and success, interoperability of voice communications is crucial.
Between humans, one of the most common information exchanges is voice communication. For purposes of this paper, the following descriptions apply:

- voice is defined as the transport of human voice using low BW resources for transmission (typically 300-3000Hz);
- audio may contain human voice, experiments noises, sounds, or music, all of which need more quality and more BW to be transmitted (typically 20-20kHz);
- audio files either may contain recorded human voice or any kind of audio. Because they are files, they are transmitted like any other files, using file transfer standards.

A voice system is a complex and dynamic system that can involve several hundred of users distributed around different buildings in a MCC or in different MCCs around the world.

A voice system can be divided in different segments according to the covered distance:

- Earth segment; including the communications inside of a MCC, between MCCs and to externals.
- MCC to Spacecraft in Short-haul segment (including Lunar operations);
- Spacecraft to Astronauts and to MCC in an EVA (extra vehicle activity)
- Spacecraft to spacecraft (rendezvous and docking operations)
- MCC to Spacecraft Long-haul segment (including Mars operations).
- Emergency communications

All these segment have different requirements and usually different technologies. The purpose to have a standard is to make all of them interoperable for all the users in Space and on the Earth.

For all these reasons to create a new standard for voice and audio communications is needed.

III. Voice System Concept

This article is heavily oriented to the voice communications used for human space flights. Unmanned space missions are less complex than crewed ones, and can be seen as a subset of the recommendations needed for human space flight. Human space missions are very dynamic and voice formats, voice loops and supporting personal are changing quite often, a good coordination related to the voice formats and loops between space agencies is mandatory.

In Satellite space missions, the voice system has a similar structure only for the LEOP. For operations the system needs only a few voice loops and no big changes are needed. Also voice formats are normally not necessary.

A Voice Format is a collection of voice loops grouped together to meet the requirements of a particular situation or mode of operation. Voice formats are typically negotiated between two or more MCCs to prioritize loop connectivity to best match communication needs with the available channels. Because in each MCC there are many more internal voice loops than available circuits to transport the voice loops between the MCCs, different operations modes must be supported by different voice formats. Each format contains a limited number of voice loops matched to the available channels for a specific operational aspect between MCCs. These formats are interchangeable and dynamic; they need to be clearly defined between MCCs. Examples of voice formats are: nominal operations; Joint Simulation and System Test Format. It is extremely important to reserve a number of channels in each format to maintain normal operations during a test or simulation. The exchange and activation of a predefined voice format should take maximal a few seconds.

A flight control room or a launch control room generally consists of a controlled access space in which flight control team (FCT) personnel work and communicate audibly through voice communications equipment, where low-latency and high capacity voice intercommunication is desired. In a MCC there are several flight control rooms, mission evaluation rooms, and multi-purpose support rooms, each housing members of the flight control team in numbers ranging from 10 to 60 persons. Keysets provide the user interface to a rich set of user functionality. User control of the individual voice loop volume complicates the user interface and the voice system. But it is necessary.

Besides that, access through remote operations is becoming more, where FCT members are not collocated a control room but rather in their individual offices, ad-hoc geographically locations (e.g. Emergency backup centers), or even working from home.

Given the real-time nature of their work and the extent of their voice communications, FCT personnel generally require low latency, non-blocking voice equipment with fast access times through defined and administered roles. Complexity of mission support, and the training required to prepare the FCT result in high capacity of always-on conference loops ready for immediate use, thereby easing schedule constraints of voice resources.

Voice summation, the creation of a coherent voice conference is possible with analog mixing equipment, and with digital waveform codecs such as G.711 [1]. But lossy voice codecs such as G.729 or G.722 cannot readily be ‘summed’ into a voice conference. They must be decoded, mixed, then re-encoded, which results in loss of voice quality. New VoIP systems used the same concept of voice summation, but using software and not a matrix.

Transcoding at interface boundaries is often required to accommodate local and regional difference in telecommunications infrastructures and end user instrumentation, e.g., E1 to T1 from Europe to the US.

Conventional voice communication technologies assume full-duplex or bidirectional communication paths, whereas for crew safety voice communications must operate independent of return path.

Voice communications with a one-way light-time delay of more than 5 seconds results in the participants engaging in message oriented conversation as opposed to being dialog oriented conversation.

From launch to orbit, as many as 400 keysets may be engaged at an MCC, accessing some 200 voice loops. While most FCT members listen to multiple voice loops concurrently, they are typically limited to talk on only one loop, whereas the CapCom and FD (Capsule Communicator and Flight Director respectively) often use a multi-talk mode – the ability to press PTT and have their voice included in up to 8 conference loops simultaneously.

Figure -1: The STS Flight Control Room (voice keysets in white circle); source: Wikimedia

Figure -2: The German ISS Flight Control Room (Voice keyset in white); source: DLR

Voice communications must be recorded and affixed with appropriate time and flight meta-information for future use and play back. Recording of both individual keysets and whole voice loops is generally required. Archival storage is required.

High reliability, maintainability and availability (RMA) are the hallmarks of the voice communications equipment intended for real-time mission operations support. Internally redundant architectures are necessary to reach the levels of RMA required.

To share voice loops among different centers, a voice loop in one center's voice switch is connected by commercial carrier to a voice loop in another center's voice switch, with each switch then serving the population of users in its respective center.

The figures 1 and 2 shown the MCC-H and Col-CC ISS control rooms. Inside of the white circles the keysets are indicated. Both systems are 100 % compatible and the development was done cooperating between both Space agencies. But as you can see the form of the keysets are completely different.

In spite of the differences of the interface to the user, the functionalities and the technology to communicate between the MCC is the same, because they use the same standard.

IV. Communications requirements

The exchange of voice data between ground-based personnel working in space missions can be achieved by commercial telecommunications services such as leased circuits like PDH (E1/T1), or Voice over Internet Protocol (VoIP). Both technologies can be combined inside and outside of an MCC. Carriers on the ground can use different technologies like Internet or Multi-Protocol Label Switching (MPLS).

When VoIP is used, the originating voice codec should be maintained from source to destination to prevent degradation resulting from transcoding. Examples would be to use IP/UDP/RTP/G.711 or IP/UDP/RTP/G.729.
If real-time voice communications either are impossible (like in Mars missions) or not needed, audio files should be exchanged using file transfer such as File Transfer Protocol (FTP), CCSDS File Delivery Protocol (CFDP [E 8]) over Delay-Tolerant Networking (DTN), or other approved standards.

A. Transport Technologies

Digital voice is transported using packets (VoIP) or frames (E1/T1). For ground communications there are currently two major possibilities: using TDM (Time Division Multiplexing) with E1/T1 interfaces or VoIP (over E1/T, MPLS or Internet). These technology options are described here:

- Time Division Multiplexing (TDM): Each voice sample time slot in the TDM frame is called a 'channel'. In European systems, TDM frames contain 30 digital voice channels (E1); in American/Japanese systems, the frames contain 24 channels (T1). Both standards (E1/T1) also contain extra bits (or bit time slots) for signaling and synchronization bits.
- TDM using PDH (Plesiochronous Digital Hierarchy) is still the more commonly used technology to transport voice with E1/T1 carriers (references [2], [3]).
- TDM over Internet Protocol (TDMoIP) is the emulation of TDM over a PSN (Packet Switched Network). TDMoIP is a type of pseudowire.
- VoIP is the current technology for new voice systems. VoIP can run over Internet or over dedicated networks.
- S/G Voice communication primarily uses Radio Frequency (RF) links, with different frequencies (S-band, Ku-band, and Ultra High Frequency [UHF]).

B. Protocols Used for Voice Transmission

Depending of which kind of voice system will be implemented in the control center and the system's functionality, one or a combination of the following protocols shown in table 1 will typically be used.

Protocol	Notes	Where Defined	Ref #
UDP (User Datagram Protocol)	None.	RFC 768 (STD 6)	[5]
RTP (Real-time Transport	RTP is used for MCC and space communications.	RFC 3550	[6]

Protocol)			(STD 64)	
RTCP (Real-time Control Protocol)		RTCP’s primary function is to provide feedback on the QoS.	RFC 3550 (STD 64)	[6]
SRTP (Secure Real-time Transport Protocol)		SRTP has a sister protocol called Secure RTCP (SRTCP)	ECP 3711	[7]
TCP/IP (Transmission Control Protocol/Internet Protocol)		TCP/IP is appropriate for exchange of audio files. TCP/IP is NOT recommended for voice communications.	RFC 675	[8]
E1/T1		E1/T1 protocols support 30 (E1) or 24 (T1) digital channels simultaneously using TDM.	ANSI T1 ITU-T G.704/G.732	[2], [3]
VoIP Signaling Protocols				
Session Control Protocols	Session Initiation Protocol (SIP)	SIP is designed to serve VoIP services over Internet Protocol (IP). SIP is an Application-Layer protocol that can operate over UDP or TCP. SIP is currently (2014) at version 2.	RFC 3261	[9]
	H.323	H.323 is an ITU-T recommendation for call signaling and control, multimedia transport and control, and BW control for point-to-point and multi-point conferences for any packet network.	ITU-T H.323	[10]
Media Control Protocols	Media Gateway Control Protocol (MGCP)	MGCP is one implementation for controlling media gateways on IP networks and the Public Switched Telephone Network (PSTN).	RFC 3435	[11]
	Megaco (H.248)	H.248 follows the guidelines of the Application Programming Interface (API) MGCP architecture and requirements described in RFC 2805.	RFC 2805	[12]

Table 1: Protocols Used for Voice Transmission

C. Voice and Audio Communication baseline

A voice communication system should provide the following:

- point-to-point voice communication;
- voice loop or voice conference (These are implemented by voice summation, which is the mixing of multiple voice sources into a single conference [voice] loop);
- voice switch (or voice matrix);
- voice recording and playback of audio files in a standard format such as Advanced Audio Coding (AAC) [14];
- voice/text conversion and viceversa;
- keysets;
- controlled access to control rooms and voice system;
- user roles;
- voice Formats;

A keyset provides capabilities for selection of loops (see figures 1 and 2), as well as for audio input and output. Interaction with the keyset for loop selection is usually a physical interaction by pressing buttons. The interaction components of the device may be dedicated hardware buttons for each loop, or a software user interface shown as a touchscreen. Physical interaction is preferred for higher user awareness. For actual audio input, the physical interaction is done by actively pressing a PTT button to activate the microphone of the keyset for transmission of the voice signals to the selected loop. The keyset may be either an integrated or multipiece separated unit, or it can be a software application (soft keyset). The complete device is called a keyset and provides all necessary functionality for end-user interaction with the voice conferencing core.

V. Voice System Services

The voice system is connected to external facilities where a different device is used for voice communications, generally a microphone or a phone. Synchronization with video for short distances (earth – moon) and voice to text conversion for long distances (e.g. Mars) are requirements that every voice system need to fulfill.

D. Lip Synchronization (Lip Sync) for Public Affairs

For public-relations events, an external facility communicates with the astronauts in space. The voice path normally uses S-band, and video uses Ku-band. The paths are different and have different latencies and need to be synchronized. Also the external facility uses a microphone or a telephone that is patched to a voice loop.

The final synchronization is done manually. The astronauts and the MCC test the whole path a few minutes before the event starts. The video personnel on the ground synchronize the video with the audio using the predefined values of the encoders. The personnel then either add or subtract steps of 10 milliseconds at a time to the encoder configuration until the lips are synchronized with the audio.

Automatic lip sync can be made possible if timestamps are used in the video and voice systems; however, this requires special equipment. Ideally, the synchronization of the video and audio should be done onboard and sent embedded from space. But the upload of the voice need to be synchronized anyways.

Some of the cameras are already using embedded audio (HD-SDI or SD-SDI) on the ISS.

E. Voice Recording/Playback

Voice recording and playback are an essential part of space missions. Classic uses of recording are to store voice data for post analysis during failure investigation, for reconstruction of issues, for public relations activities, and for training purposes. Another common use of voice recording is short-term playback (10 sec.) of unclear communication to avoid misunderstandings and frequent repetition over the channels.

Based on the use cases, three different types of recording are used: user-specific recording (or keyset recording), loop-specific recording (e.g., dialog between different positions), and interface-specific recording.

F. Voice-to-Text Conversion

Voice-to-text, also known as Automatic Speech Recognition (ASR) [4] or text-to-speech (TTS), has become a standard application for many smartphones, browsers, cars, and dictation devices.

When long transmission time becomes an issue, voice-to-text software should be considered; voice should be converted to text and sent to the MCC as a text file. The reply could be sent in the same way and could optionally be converted back to voice.

G. Voice and Audio Communication Security

Security concerns in voice and audio communication are voice privacy, audio data integrity, and source authentication:

- Voice privacy means that the voice loop should not be heard by any other than the intended recipient(s). It can be achieved by employing protected data transport at lower layers, restricting distribution, or using cryptographic techniques.
- Audio data integrity means that the content of the audio/voice should not be altered between source and destination.
- Source authentication is the requirement that the audio should be attributable to a known origin. It can be achieved together with the privacy implementation.

Each user has an assigned role, and each role have a predefined group of loops with Talk/listen or Monitor permissions. Each user has a user-id and password to access the assigned role into the voice system. The access control is generally done via a keyset and maintained in a system like LDAP (Lightweight Directory Access Protocol).

Due to the fact that the voice loops should be recorded and there are time stamps on them, they can be correlated with the LDAP access to the voice system and in this way is possible to identify the person that was talking, the voice loop and the content.

VI. VOICE SYSTEM: Parameter Considerations

The following parameters must be taken into account in the definition and implementation of a voice system:

- The network: internally in a MCC normally there is a dedicated network for the voice traffic or the operational network can be used for VoIP having separated VLANS (Virtual local Area Network). In this case the channel used for voice should have the highest priority in the network traffic.
- For the external network or WAN: packet-switched versus circuit-switched systems (packet-switched systems are steadily replacing circuit-switched systems);
- users and permissions associated with the login (e.g., the roles Talk/Listen/Monitor [T/L/M]); the Monitoring feature is used in many console positions. Especially in the externals ones (e.g. Payload operator) just to hear in a voice loop but without permission to talk. For the positions have talk/listen/monitoring permissions, the monitoring button is often used to not disturb the voice loop for talking in another loop simultaneously or to talk in a side conversation.
- voice formats and activation procedures, a common case when more voice loops are needed than available circuits to transmit them; it is extremely important that the voice formats are clearly defined and agreed between the MCCs before use. Exchange or activation of voice formats must be also coordinated and planned for each activity (e.g a Simulation).Changing voice loops within a format is a dynamic task – good coordination is mandatory.
- PTT must be implemented, manually using a button or pedal, or automatic using VoX (Voice Level Detection or Voice Activity Level or Voice Operated eXchange);
- codecs selections and settings (including lossless, lossy, codecs with less compression like G.711 [1] are preferred)
- minimization of codec conversions;
- synchronization (clocking);
- latency requirements (e.g., 10 to 200 msec for VoIP systems inside of one control room). The latency requirements for a system running inside a control room must be as low as possible. For systems communicating with external locations or different rooms inside of a facility, it is not a problem to have a higher latency; VoIP systems have a larger latency due to the processing delay (the encoding and decoding of voice);
- frequencies and channelization (it must be clearly agreed between the MCCs);
- signaling; CAS (Channel Associated Signaling) or CCS (Common Channel Signaling). It is important for the S/G loops using digital voice.
- voice encryption and voice private channel setup (restricted talk configuration for critical voice loops; e.g., only the CapCom can talk on the air-to-ground loop during launch).
- private conferences: private conferences occur with family and doctors. Authorized parties are connected to the private conference loop. Voice tech personnel also check and disable other interfaces as may be necessary to ensure a private conference. Once the parties are connected and voice checks complete, the voice techs themselves disengage from the private loop.
- quality monitoring (Depending on the case, QoS and/or QoE will be needed). If the Internet is used for the voice transmission, quality cannot be guaranteed. Quality of service is related to the network. Quality of Experience is related to the human perception, should be 5x5; Voice quality measure is possible through automated and subjective human evaluative methods. The long standing telecom standard MOS (Mean Opinion Score) is 4.0 to 4.2 for G.711 encoded voice. Wideband codecs (e.g., G.722.2) may result in higher comparative scores, while narrow band digital compression codecs (e.g., G.729) may result in lower comparative scores.
- etiquette and protocol (roles and positions to talk, a good training is mandatory);
- alternative methods (consider secondary paths such as IP telephony—e.g., Skype, VSee—in the ISS for voice conference which use different protocols and etiquette).

- individual voice loop controls including talk/monitor, monitor only, and volume level;
- each conference loop key is stateful, retaining whatever the current configuration settings are over its default configuration, as the user selects other voice loop pages on which that voice loop is found;
- VoIP hardware and software keysets deployed where low latency voice performance is not a driving requirement;
- high capacity includes 1,600 active conference loops, with 3,000 stored conference definitions;
- for critical flight control use, the central voice switch must have internally redundant voice paths such that no single failure can remove a voice loop from use.
- Packet Loss – Normally occurs in IP networks as a result of network congestion, where incoming packets are thrown away. Within RF networks, packet loss may be due to signal loss or fade, or a too tight link budget. The impact on voice of an occasional random packet loss is negligible, but the impact for excessive packet loss is poor voice quality or loss of voice communications entirely.
- Real-time or near-real-time voice communication is applicable for relatively short distances, spanning at most a planet and its moon(s). For longer distances, audio files exchange should be used.
- For audio files, an established standard should be used like ACC or Mpeg3.
- For the transmission of Audio files, especially for long haul missions, DTN (Delay/Disrupt Tolerant Network) protocol should be used.
- Emergency Communications: in the event that communications with central modules is somehow interrupted, emergency communications capabilities should exist, whether to orbiting satellites or at a low level directly back to earth.
- Latency in the Earth to Moon segment is about two seconds, and thus the voice processing delay begins to impact total latency. A lunar crew member talking with an earth based support team member should expect a five second interval, at the least, between the end of their utterance and the beginning of a response utterance from Earth.
- As the operations of a lunar or Mars outpost may increase in complexity, crew population, science and in-situ activities; over-subscription of the long-haul bandwidth may become an issue. Bandwidth utilization must be well planned for and traffic prioritization schemes considered and employed. Voice traffic must have the highest priority.

H. VOICE SYSTEM: example of an MCC

The following context diagram provides a view of a typical centralized voice switch and ancillary equipment supporting human space flight. The diagram describes the voice communication subsystem used at NASA's MCC-H. The block titled External Voice Distribution Equipment provides the equipment necessary to support transcoding among voice and transport technologies such as G.711 to G.729, T1 to E1.

The following bullets provide more detail to the figure 3:

- The central switch complex consists of a Voice Switch and its Conference Summation components. Voice summation remains an area of legacy equipment. (New digital systems are using the same concept based on software, but with greater delays). High performance and low latent voice intercommunication systems utilize Time Division Multiplexing (TDM) for the 'voice bus' and G.711 or codecs for voice encoding and summation. A voice loop 'talker' consumes one slice of the TDM voice bus to carry their voice. Another slice of TDM carries the summed voice loop audio. The central switch either decodes the two audio voice slices and mixes them in an analog fashion, or provides a code book look-up for "adding" the voice quanta together. The resulting audio is the 'summed' voice of the talker and the voice loop, and this audio is placed back into the voice loop slice of the TDM voice bus for all to hear.

- Interfaces with the central voice switch include:
- Keysets. Hardware keysets may be T1 or VoIP connected, software keysets are only VoIP connected.
- External interface; T1 with typical telecommunications signaling options. Additional peripheral equipment may translate and transcode as required for external systems. For instance, a single DS0 channel of a T1 may be defined to carry a specific voice conference for interface with an external secondary voice switch such as a PABX (Private Branch Exchange) or a third party VoIP switch or even the central voice switch of another agency.

- Voice recording equipment provides for both loop and individual keyset recording. Voice recording includes time stamp for storage and retrieval capability.
- Air to Ground Voice Equipment (AGVE) must accommodate one-way voice communications for crew safety. In the event that voice communication is not actively being received from a spacecraft, AGVE must continue to forward voice communications, for the crew may be receiving but cannot transmit.
- Adhoc VoIP centers are centers supporting operations or external facilities that participate in Public relations events with the Astronauts in the Spacecraft.

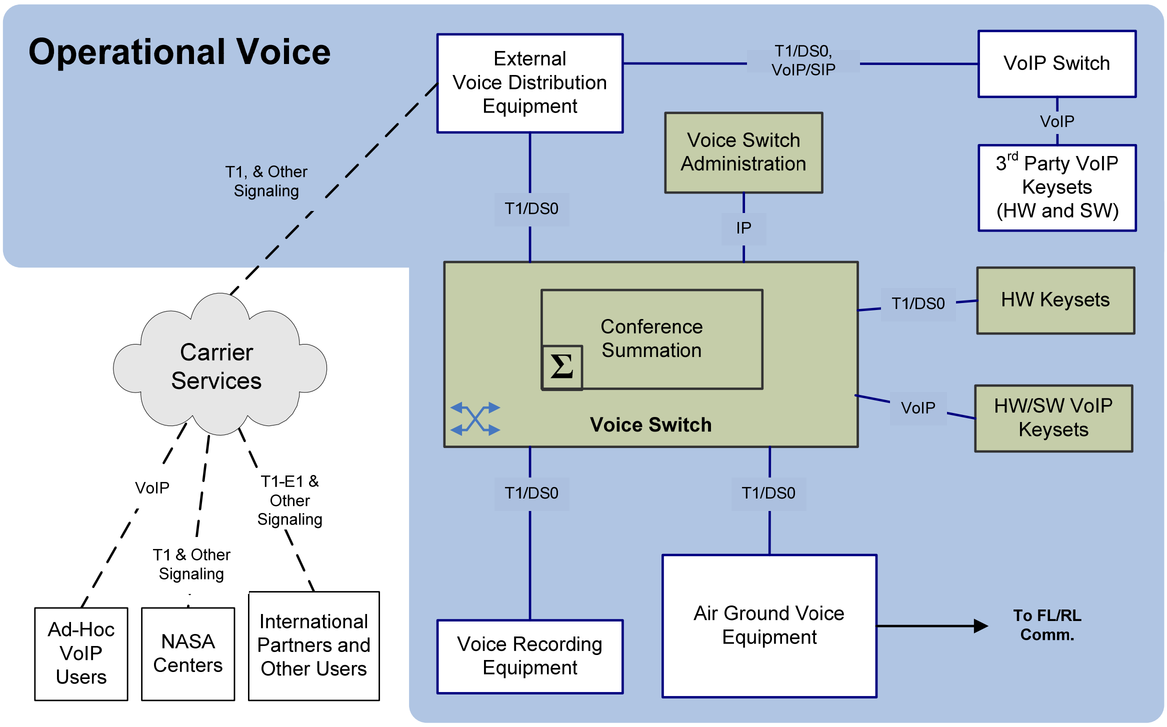

Figure -3: Mission Operations Voice Context Diagram

The US Space Shuttle employed a Dissimilar Voice (DV) communications during launch and ascent, and during landing. This voice channel through analog UHF radio with Mission Control was delivered synchronously with the Operational Air to Ground channel (A/G-1). The threshold for synchronicity was 7 ms, where if the voice signals of DV and A/G-1 under waveform analysis wander further apart than 7 ms, automated delay equipment adjusted accordingly the DV channel to bring them back into synchronicity. New Rockets still use UHF radio communications.

The International Space Station (ISS) utilizes MRELP (Modified Residual Excited Linear Predictive) a codec with data rate at 9.6 Kbps) for space to ground voice communications (S/G 1 and 2). The MRELP voice system was intended only for the early stages of ISS construction after which it was to become the backup voice system. However a primary voice system was never installed.

Two additional voice channels for ISS use VoIP G.729 within RTP/UDP/IP, used for backup communications or for PR events (S/G 3 and 4). ISS also supports ham radio communications primarily as a public affairs capability. UHF voice communications are used during proximity and docking operations with approaching crewed spacecraft.

VII. Conclusion

In this paper a short description of the problems to generate and agree to an international voice standard are presented. The high complexity of a voice system and the many internfaces used for human space operations is discussed. After the agreement to have a voice standard, the general requirements needed for a voice system are presented and also the different segments that a voice system must to cover. Also an example of a system in a mission control center and its interfaces is explained.

A voice system is mission critical, and must be works when everything else fails; it is a central system of a control center, especially for human space flight. The new digital technologies and the strong relationship to the network speed and reliability open new possibilities but also big challenges related to quality of voice transmission and security.

For PR event voice in getting more and more embedded in the video, but for communications is still needed as a separated system.

This discussion is still ongoing but at least an Agreement between all the mayor space Agencies was reached and will be soon the new international standard for voice and audio.

A standard must be continually reviewed and updated according to the new technologies.

References

[1] *Pulse Code Modulation (PCM) of Voice Frequencies.* G.711. International Telecommunication Union-Telecommunication Standardization Sector (ITU-T), November 1988.

[2] *Standard for Telecommunications—Network-to-Customer Installation*. ANSI T1.403-1989. ANSI, 1989.

[3] *Synchronous Frame Structures Used at 1544, 6312, 2048, 8448 and 44 736 kbit/s Hierarchical Levels*. G.704. ITU-T, October 1998 and *Characteristics of Primary PCM Multiplex Equipment Operating at 2048 kbit/s*. G.732. ITU-T, November 1988.

[4] *Coding of Speech at 8 kbit/s Using Conjugate-Structure Algebraic-Code-Excited-Linear Prediction (CS-ACELP).* G.729. ITU-T, June 2012.

[5] J. Postel. *User Datagram Protocol (UDP)*. RFC 768. ISI, August 1980.

[6] H. Schulzrinne, et al. *RTP: A Transport Protocol for Real-Time Applications*. RFC 3550. ISOC, July 2003.

[7] M. Baugher, et al. *The Secure Real-time Transport Protocol (SRTP)*. RFC 3711. ISOC, March 2004.

[8] V. Cerf, Y. Dalal, and C. Sunshine. *Specification of Internet Transmission Control Program*. RFC 675. December 1974.

[9] J. Rosenberg, et al. *SIP: Session Initiation Protocol*. RFC 3261. ISOC, June 2002.

[10] *Packet-based Multimedia Communications Systems.* H.323. ITU-T, December 2009.

[11] F. Andreasen and B. Foster. *Media Gateway Control Protocol (MGCP)*. RFC 3435. ISOC, January 2003.

[12] N. Greene, et al. *Media Gateway Control Protocol Architecture and Requirements*. RFC 2805. ISOC, April 2000.

[13] *Coding of Speech at 16 kbit/s Using Low-delay Code Excited Linear Prediction.* G.728. ITU-T, June 2012.

[14] *Information Technology—Generic Coding of Moving Pictures and Associated Audio Information—Part 7: Advanced Audio Coding (AAC).* International Standard, ISO/IEC 13818-7:2006. 4th ed. Geneva: ISO, 2006.