The Consultative Committee for Space Data Systems

Report Concerning Space Data System Standards

VOICE COMMUNICATIONS

INFORMATIONAL REPORT

CCSDS 706.2-G-2

GREEN BOOK

December 2018

The Consultative Committee for Space Data Systems

Report Concerning Space Data System Standards

VOICE COMMUNICATIONS

INFORMATIONAL REPORT

CCSDS 706.2-G-2

GREEN BOOK

December 2018

AUTHORITY

Issue:	Informational Report, Issue 2
Date:	December 2018
Location:	Washington, DC, USA

This document has been approved for publication by the Management Council of the Consultative Committee for Space Data Systems (CCSDS) and reflects the consensus of technical panel experts from CCSDS Member Agencies. The procedure for review and authorization of CCSDS Reports is detailed in *Organization and Processes for the Consultative Committee for Space Data Systems* (CCSDS A02.1-Y-4).

This document is published and maintained by:

CCSDS Secretariat
National Aeronautics and Space Administration
Washington, DC, USA
E-mail: secretariat@mailman.ccsds.org

FOREWORD

Through the process of normal evolution, it is expected that expansion, deletion, or modification of this document may occur. This Report is therefore subject to CCSDS document management and change control procedures, which are defined in *Organization and Processes for the Consultative Committee for Space Data Systems* (CCSDS A02.1-Y-4). Current versions of CCSDS documents are maintained at the CCSDS Web site:

http://www.ccsds.org/

Questions relating to the contents or status of this document should be sent to the CCSDS Secretariat at the e-mail address indicated on page i.

At time of publication, the active Member and Observer Agencies of the CCSDS were:

Member Agencies

- Agenzia Spaziale Italiana (ASI)/Italy.
- Canadian Space Agency (CSA)/Canada.
- Centre National d'Etudes Spatiales (CNES)/France.
- China National Space Administration (CNSA)/People's Republic of China.
- Deutsches Zentrum für Luft- und Raumfahrt (DLR)/Germany.
- European Space Agency (ESA)/Europe.
- Federal Space Agency (FSA)/Russian Federation.
- Instituto Nacional de Pesquisas Espaciais (INPE)/Brazil.
- Japan Aerospace Exploration Agency (JAXA)/Japan.
- National Aeronautics and Space Administration (NASA)/USA.
- UK Space Agency/United Kingdom.

Observer Agencies

- Austrian Space Agency (ASA)/Austria.
- Belgian Federal Science Policy Office (BFSPO)/Belgium.
- Central Research Institute of Machine Building (TsNIIMash)/Russian Federation.
- China Satellite Launch and Tracking Control General, Beijing Institute of Tracking and Telecommunications Technology (CLTC/BITTT)/China.
- Chinese Academy of Sciences (CAS)/China.
- Chinese Academy of Space Technology (CAST)/China.
- Commonwealth Scientific and Industrial Research Organization (CSIRO)/Australia.
- Danish National Space Center (DNSC)/Denmark.
- Departamento de Ciência e Tecnologia Aeroespacial (DCTA)/Brazil.
- Electronics and Telecommunications Research Institute (ETRI)/Korea.
- European Organization for the Exploitation of Meteorological Satellites (EUMETSAT)/Europe.
- European Telecommunications Satellite Organization (EUTELSAT)/Europe.
- Geo-Informatics and Space Technology Development Agency (GISTDA)/Thailand.
- Hellenic National Space Committee (HNSC)/Greece.
- Hellenic Space Agency (HSA)/Greece.
- Indian Space Research Organization (ISRO)/India.
- Institute of Space Research (IKI)/Russian Federation.
- Korea Aerospace Research Institute (KARI)/Korea.
- Ministry of Communications (MOC)/Israel.
- Mohammed Bin Rashid Space Centre (MBRSC)/United Arab Emirates.
- National Institute of Information and Communications Technology (NICT)/Japan.
- National Oceanic and Atmospheric Administration (NOAA)/USA.
- National Space Agency of the Republic of Kazakhstan (NSARK)/Kazakhstan.
- National Space Organization (NSPO)/Chinese Taipei.
- Naval Center for Space Technology (NCST)/USA.
- Research Institute for Particle & Nuclear Physics (KFKI)/Hungary.
- Scientific and Technological Research Council of Turkey (TUBITAK)/Turkey.
- South African National Space Agency (SANSA)/Republic of South Africa.
- Space and Upper Atmosphere Research Commission (SUPARCO)/Pakistan.
- Swedish Space Corporation (SSC)/Sweden.
- Swiss Space Office (SSO)/Switzerland.
- United States Geological Survey (USGS)/USA.

DOCUMENT CONTROL

Document	Title	Date	Status
CCSDS 706.2-G-1	Voice Communications, Informational Report, Issue 1	September 2010	Original issue, superseded
CCSDS 706.2-G-2	Voice Communications, Informational Report, Issue 2	December 2018	Current issue

CONTENTS

CONTENTS (continued)

1 INTRODUCTION

1.1 PURPOSE AND SCOPE OF THIS DOCUMENT

This CCSDS Informational Report presents an overview of voice communications in crewed space operations. The voice communications system for satellite operations is a strongly reduced version of that used for crewed space flight. Only Launch and Early Orbit Phase (LEOP) operations can be seen as having a similar complexity, but there is still a reduced version of a voice system dedicated to talking to astronauts. This document has been prepared by the Voice Working Group of the CCSDS Space and Internetworking Services (SIS) area.

1.2 DOCUMENT STRUCTURE

This document has the following major sections:

– Section 1 contains administrative information, definitions, and references.

– Section 2 describes the current state of voice communications supporting crewed and non-crewed spaceflight.

– Section 3 describes the technical drivers for the future of voice communications.

– Annex A contains acronyms and abbreviations.

– Annex B contains short descriptions of codecs.

– Annex C summarizes protocols used for voice transmission.

– Annex D discusses communications quality and important definitions.

1.3 DEFINITIONS

Within the context of this document the following definitions apply:

audio: Sonic information capable of being recorded or transmitted (typically 20 Hz–20 kHz). In addition to human voice, audio may contain experiments' noises, sounds, or music, all of which need better quality and higher data rates than voice to be transmitted.

audio files: Digital files that may contain recorded human voice or any kind of audio. Because they are files, they are transmitted, as with any other files, using file transfer standards.

bandwidth, BW: The rate of data transfer or throughput, measured in bits per second (bit/s) or multiple of bit/s (kb/s, Mb/s, Gb/s).

codec: Coder-decoder, in the context of voice communications.

encoding: Analog-to-digital conversion or compression of digitally represented audio data.

five-by-five: Of an audio signal, having sufficient strength and clarity to be subjectively considered to be of good quality (see annex D).

interoperability: The technical capability of two or more systems or components to exchange information via a common set of procedures or protocols, and to understand the same voice and data formats while using the same or compatible protocols.

keyset: A device or software interface that provides capabilities for selection of loops, as well as for audio input and output. (See annex D.)

mission control center, MCC: A facility that manages crewed space flights.

NOTE – This book uses MCC because it is the term used for crewed space flight. Mission Operations Center (MOC) is the term used primarily for satellite missions and is thus used in many CCSDS books.

press to talk, PTT (*also* push to talk, press to transmit): The capability to switch from voice reception mode to transmit mode.

sample rate (codecs): MHz or kb/s rate at which the data is sampled. For example, the G.711 codec algorithm samples the data at an 8-kHz rate (8 bits per sample) or 64 kb/s, which is the standard Pulse Code Modulation (PCM) data stream rate.

summation: The mixing of multiple digital or analog voice sources into a single digitally encoded conference loop.

voice: Audio produced by human vocal organs (typically 300–3000 Hz), capable of being transmitted using low-data-rate resources.

voice format: 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 MCCs to prioritize loop connectivity to best match communication needs with the available channels. (See annex D.)

voice loop or voice conference: The result of summation, a single conference containing the voices of all participating ‘talkers’.

user roles and permissions: Parameters that prescribe a user’s talk, listen, and monitor interactions with the audio system. (See annex D.)

voice operated exchange, VOX (*also* voice operated switch, voice level detection, voice activity level): A switch that operates when sound above a certain threshold is detected. (See annex D.)

1.4 REFERENCES

The following publications are referenced in this document. At the time of publication, the editions indicated were valid. All publications are subject to revision, and users of this document are encouraged to investigate the possibility of applying the most recent editions of the publications indicated below. The CCSDS Secretariat maintains a register of currently valid CCSDS publications.

[1] *Pulse Code Modulation (PCM) of Voice Frequencies*. ITU-T Recommendation G.711. Geneva: ITU, 1988; and

Pulse Code Modulation (PCM) of Voice Frequencies—Appendix II: A Comfort Noise Payload Definition for ITU-T G.711 Use in Packet-Based Multimedia Communication Systems. ITU-T Recommendation G.711—Appendix II. Geneva: ITU, February 2000.

[2] *Lossless Compression of G.711 Pulse Code Modulation*. ITU-T Recommendation G.711.0. Geneva: ITU, 2009.

[3] *CCSDS File Delivery Protocol (CFDP)*. Issue 4. Recommendation for Space Data System Standards (Blue Book), CCSDS 727.0-B-4. Washington, D.C.: CCSDS, January 2007.

[4] *CCSDS Bundle Protocol Specification*. Issue 1. Recommendation for Space Data System Standards (Blue Book), CCSDS 734.2-B-1. Washington, D.C.: CCSDS, September 2015.

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

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

[7] *Packet-Based Multimedia Communications Systems*. ITU-T Recommendation H.323. Geneva: ITU, 2009.

[8] *Licklider Transmission Protocol (LTP) for CCSDS*. Issue 1. Recommendation for Space Data System Standards (Blue Book), CCSDS 734.1-B-1. Washington, D.C.: CCSDS, May 2015.

[9] *Coding of Speech at 16 kbit/s Using Low-Delay Code Excited Linear Prediction*. ITU-T Recommendation G.728. Geneva: ITU, 2012.

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

[11] *Rationale, Scenarios, and Requirements for DTN in Space*. Issue 1. Report Concerning Space Data System Standards (Green Book), CCSDS 734.0-G-1. Washington, D.C.: CCSDS, August 2010.

[12] *Methods for Subjective Determination of Transmission Quality*. ITU-T Recommendation P.800. Geneva: ITU, 1996.

[13] F. Andreasen and B. Foster. *Media Gateway Control Protocol (MGCP) Version 1.0*. RFC 3435. Reston, Virginia: ISOC, January 2003.

[14] N. Greene, M. Ramalho, and B. Rosen. *Media Gateway Control Protocol Architecture and Requirements*. RFC 2805. Reston, Virginia: ISOC, April 2000.

[15] *AOS Space Data Link Protocol*. Issue 3. Recommendation for Space Data System Standards (Blue Book), CCSDS 732.0-B-3. Washington, D.C.: CCSDS, September 2015.

[16] *Proximity-1 Space Link Protocol—Physical Layer*. Issue 4. Recommendation for Space Data System Standards (Blue Book), CCSDS 211.1-B-4. Washington, D.C.: CCSDS, December 2013.

[17] *Encapsulation Service*. Issue 2. Recommendation for Space Data System Standards (Blue Book), CCSDS 133.1-B-2. Washington, D.C.: CCSDS, October 2009.

[18] *IP over CCSDS Space Links*. Issue 1. Recommendation for Space Data System Standards (Blue Book), CCSDS 702.1-B-1. Washington, D.C.: CCSDS, September 2012.

2 VOICE COMMUNICATIONS IN FLIGHT OPERATIONS

2.1 INTRODUCTION

This document provides an overview of voice communications in support of crewed spaceflight. The voice and audio communications used for non-crewed space flight can be considered as a subset of the ones used for crewed space flight. In this case the number of loops, except for the LEOP, is far less complex and not as dynamic. Also, the use of voice formats is usually not needed. In satellite missions, the voice loops used are almost static; they change from mission to mission but not to the extent that those used in crewed space flight do. All the other concepts related to users, roles, permissions, keyset configuration, etc., are applicable here as well. Voice communications addresses many combinations of users, grouped largely by where these users reside. Different locations dictate different requirements and constraints upon the voice communication services that must be provided. This document establishes four *segments* for voice communication, each with unique characteristics:

– Earth segment (between terrestrial facilities, including emergency communications and SAR);

– lunar or Mars segment (between lunar or Mars facilities, astronauts, and Earth, including emergency communications and SAR);

– short-haul segment (between terrestrial and space-borne facilities, including emergency communications); and

– long-haul segment (between terrestrial and space-borne facilities, including emergency communications).

The purpose of this document is to establish a shared understanding of the technical drivers that affect each of these segments, in order to inform subsequent work in defining or adopting standards for voice communication to serve one or more segments.

The use and nature of voice communications in human spaceflight operations evolved largely from the operational concepts of military voice communications. The early astronaut was a military jet pilot who endured short, lonely flights into low Earth orbit. Today's astronauts come from diverse disciplines, and their flights are typically neither short nor lonely. On the International Space Station (ISS), with its full crew compliment of 6, one finds pilots, to be sure, but also, and perhaps more importantly, engineers, scientists, medical doctors, and more. At one time, ISS hosted a crew of 7 from the US Space Shuttle, bringing its population to 13 personnel for a short time, and resulting in a significantly increased demand for the limited ISS voice communication resources. Clearly there has been significant evolution in both the requirements for and the constraints upon voice communication: from a single user (analog), short-duration, to multi-user, multi-spacecraft, and very long duration operation with substantial variation in load over time.

As ground-based flight control teams grew in size, high-capacity, high-performance voice switch and conferencing equipment was introduced. This equipment was often custom developed, or modified Commercial-Off-The-Shelf (COTS) products.

As the cost and complexity of space missions increased, joint international missions among multiple space agencies emerged. Digital voice conference systems as well as commercial telephony carriers and their capabilities were employed to provide connectivity among multiple flight centers for real-time mission operations support by the participating agencies.

Globally, the following describe some of the fundamental challenges for the future of voice communications in human spaceflight:

a) Flight operations personnel currently work in close proximity in collocated environments such as flight control rooms, where low-latency and high-capacity voice intercommunication is desired. However, greater access through remote operations is becoming a driving force, where flight operations team members are not collocated but rather may be found in their individual offices, ad-hoc geographic locations, or even working from home.

b) Voice summation, the creation of a coherent voice conference, is possible with analog mixing equipment, and with digital waveform codecs such as G.711 (reference [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, and then re-encoded, resulting in loss of voice quality. The new voice systems technologies are replacing the analog mixing equipment with digital ones and with pure software solutions as well.

c) Transcoding at interface boundaries is often required to accommodate local and regional differences in telecommunications infrastructures and end-user instrumentation, for example, E1 to T1 from Europe to the US or Japan, or G.722 to/from G.729 between cellular carriers. This may impact voice quality as transcoding may occur between lossy codecs.

d) Conventional voice-communication technologies assume full-duplex or bidirectional communication paths, whereas for crew safety, voice communications must operate independently of return path. And further, network transport technologies also require bidirectional communication, which again should not be employed for communications with in-flight spacecraft.

e) Voice communications with a one-way light-time delay of about five seconds results in the participants’ engaging in message-oriented conversation, as opposed to dialog-oriented conversation.

2.2 VOICE COMMUNICATIONS

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 through voice-intercommunications equipment. Keysets provide the user interface with a rich set of user functionality. FCT members may participate concurrently in multiple conferences, or voice loops, while talking on one selected voice loop. Certain authorized personnel may talk on more than one voice loop. User control of the individual voice loop volume control complicates the user interface and the intercommunications equipment.

Figure 2-1: ISS Flight Control Room at NASA Mission Control Center[1]

[1] Source: NASA.

Figure 2-2: The Russian ISS Flight Control Room[2]

Given the real-time nature of their work and the extent of their voice intercommunications, flight control 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 intercommunications must be recorded and affixed with appropriate time and flight meta-information for future use and playback. 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 intercommunications equipment intended for real-time mission-operations support. Internally redundant architectures are typically necessary to reach the levels of RMA required.

At a mission control center, 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. Centralized voice equipment provides the high-performance, high-capacity voice intercommunications necessary. From launch to orbit, as many as 400 keysets may be engaged, accessing some 200 voice loops. While most FCT

[2] Source: Wikimedia.

members listen to multiple voice loops concurrently, they are typically limited to talk on only one loop, whereas the Capsule Communicator (CapCom) and Flight Director (FD) often use a multi-talk mode, the ability to activate Press To Talk (PTT) and have their voices included in up to eight conference loops simultaneously.

To share voice loops among different centers, a voice loop in one center's voice switch is connected by a 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.

Given the operational concepts of a flight control team, the following summarizes suitable voice communication requirements:

– hardware keysets for dedicated appliance-like performance: collocated personnel require low latency voice among themselves;

– up to 10 pages each containing up to 24 voice loops; user configured with user- or role-defined defaults; the number of pages and voice loops per page can vary in different voice systems;

– individual voice loop controls, including talk/monitor, monitor only, and volume level;

– stateful conference loop keys, retaining whatever the current configuration settings are over their default configuration, as the user selects other voice loop pages on which that voice loop is found;

– Voice over Internet Protocol (VoIP) hardware and software keysets deployed whenever low-latency voice performance is not a driving requirement;

– voice latency requirements inside control room:

  • PTT to talk should be less than 15 ms,

  • Talk and voice summation should be less than 150 ms;

– restricted talk configuration for critical voice loops; for example, only the CapCom can talk on the Space-to-Ground (S/G) loop during launch;

– high capacity, including 1,600 active conference loops, with 3,000 stored conference definitions;

– expandable capacity, to 5,000 connections, whether keysets or external signals;

– internally redundant voice paths on the central voice switch for critical flight control use, so that no single failure can remove a voice loop from use.

The following context diagram provides a view of a typical centralized voice switch and ancillary equipment supporting human spaceflight. The diagram well describes the voice communication subsystem used at NASA's JSC. 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.

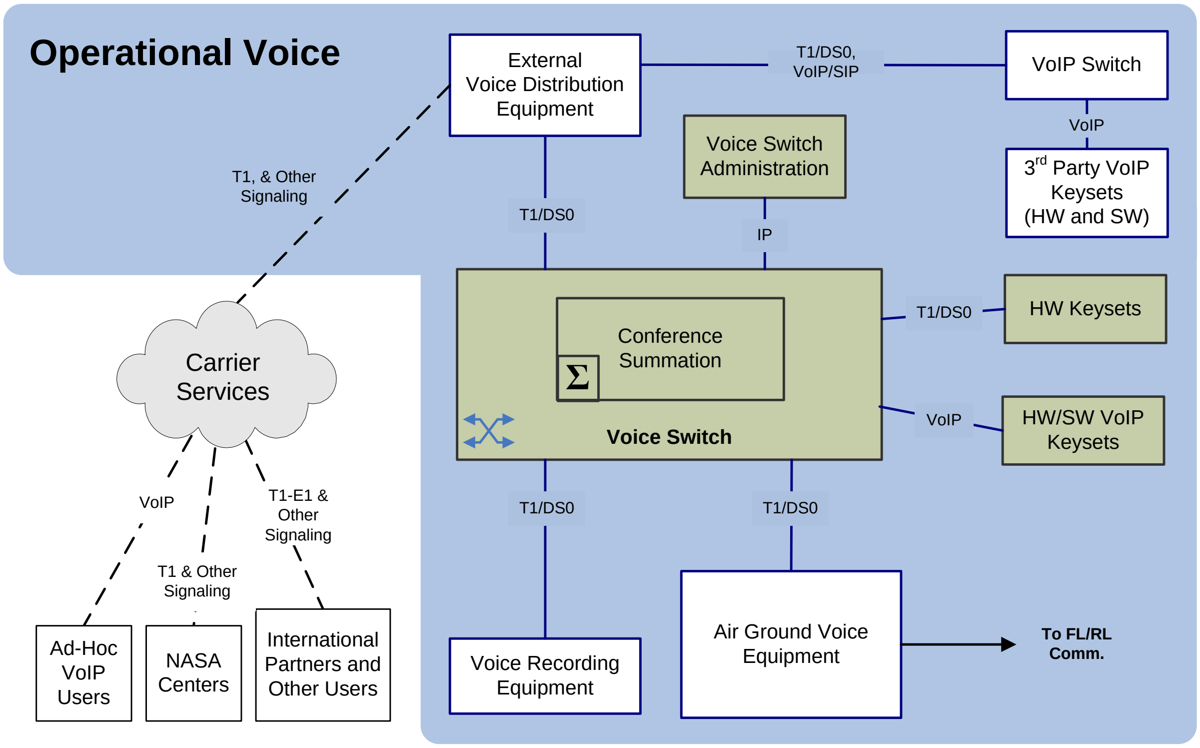


Figure 2-3: Mission Operations Voice Context Diagram

The following items provide more detail to the above diagram:

– The central switch complex consists of a Voice Switch and its Conference Summation components. This is Time Division Multiplexing (TDM) -based equipment that provides very low-latency, high-performance, and high-capacity voice intercommunications.

– Interfaces with the central voice switch include

  - Hardware keysets may be T1 or VoIP connected, software keysets are only VoIP connected.

  - External interface is 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 Private Automatic Branch Exchange (PABX), 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 Inter-Range Instrumentation Group (IRIG) timecode time stamp for storage and retrieval capability. Voice recording is typically stored digitally on a storage area network.

- Local site administrator control PCs communicate with the voice switch via Internet Protocol (IP).

Air-to-Ground Voice Equipment (AGVE), provides interconnect between the central voice switch and the air-to-ground communications network that communicates with spacecraft. Included in this is any necessary translation and transcoding.

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.

2.3 AUDIO COMMUNICATIONS

Non-real-time communications are expected to be handled by the exchange of audio files.

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 such as 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. For purposes of this document, the following descriptions apply:

– *voice* is defined as the transport of human voice using low BW resources for transmission (typically 300-3000 Hz);

– *audio* may contain human voice and experiments' noises, sounds, or music, all of which need more quality and more BW to be transmitted (typically 20 Hz–20 kHz);

– *audio files* may contain either recorded human voice or any kind of audio. Because they are files, they are transmitted like any other files, using file transfer standards.

For audio file transmission any method for file transfer can be used; CCSDS would use CFDP (reference [3]) and for long distances DTN (references [4] and [8]).

CFDP currently supports four classes of file transfer. Going forward, SIS recommends using one of the following two classes:

– Class 1—Unreliable CFDP Transfer over BP or directly over LTP-Red;

– Class 2—Reliable CFDP Transfer over BP or directly over LTP-Red or LTP-Green.

Each class can be supported over the DTN Bundle Protocol (BP—reference [4]) and DTN Licklider Transport Protocol (LTP—reference [8]). For CFDP, revisions have been proposed that could lead to deprecation of Classes 3 and 4 in favor of using CFDP over BP or LTP.

If Class-1 (Unreliable) CFDP is used over Bundle Protocol, then the expectation is that CFDP will invoke the Bundle Protocol reliable delivery option.

More information related to DTN can be found in *Rationale, Scenarios, and Requirements for DTN in Space* (reference [11]).

2.4 GROUND VOICE SYSTEMS

2.4.1 GENERAL

Most of the current voice systems used in the space agencies are based on an analog hardware voice matrix. New voice systems are going into the IP communications and have matrix based on IP (hardware or software).

Keyset devices typically have a headset with a microphone that converts the voice waveform into an analog signal. A voice codec is then used to convert this analog voice signal into a digitally encoded version. It accomplishes this by using algorithms to convert the analog voice waveform into a digital format or encoding. Choosing which algorithm to use involves evaluating voice quality verses bandwidth consumption, among other factors. Codecs may simply provide quantized waveform representation, sampled at appropriate rates. Other complex codecs are more CPU intensive, performing psycho-acoustic analysis and prediction of the waveform rendering parameter sets that are a very compressed description of sampled voice, for example, G.729.

The benchmark coding scheme is PCM (reference [1]). PCM was a product of the development of digital telephony in the 1970s, which led to the T-carrier systems that are widely deployed today. PCM samples the voice signal at a rate of 8,000 samples per second and then assigns each sample one of the 256 discrete levels using eight-bit code. This yields a data rate of 64 kb/s (8,000 samples/second × 8 bits/sample = 64000 bits/second). The discrete level is assigned in a logarithmic manner as opposed to linear which yields greater resolution with low signal levels, but the digital output is a constant 64 kb/s. In 1988 PCM encoding was standardized by the International Telecommunications Union (ITU) as Recommendation G.711 (reference [1]).

The disadvantage of PCM/G.711 is that the bandwidth is a constant 64 kb/s. Speech signals contain some forms of redundancy that could be removed to help compress the output data rate which would in turn reduce the required bandwidth. This is where other bandwidth-reducing ITU codec standards have been developed, some of which are noted here with their corresponding bandwidth: G.722.1 (24/32 kb/s), G.723.1 (5.3/6.3 kb/s), G.726 (16/24/32/40 kb/s) G.728 (16 kb/s) and G.729 (8 kb/s).

Bandwidth saving codecs do come with some caveats. Guaranteed packet delivery becomes more important with the higher-compression voice codecs. This is because more voice data is being squeezed into each packet, and the loss of a packet between the source and destination keyset will result in the loss of that part of the voice stream. This could be problematic in mission support when a series of numbers is being communicated.

In addition, codecs with significantly reduced bandwidth usage have the disadvantage of increased latency due to the sampling of a period of voice and the processing of that sample period.

Voice Summation

Voice summation remains an area of legacy equipment but is slowly being replaced by the new digital systems (VoIP). High-performance and low-latency voice intercommunication systems utilize TDM for the 'voice bus', and G.711 or other waveform quanta codecs for voice encoding and summation. A voice loop 'talker' consumes one slice of the TDM voice bus to carry voice. Another slice of TDM carries the summed voice loop audio. The central switch either decodes the two audio voice slices and either 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.

Voice summation with lossy codecs such as G.729 or G.722 is problematic, for the voice data must be decoded to a waveform codec like G.711, or even to analog voice, in order to create a conference, or voice loop. It is simply not possible to 'add' lossy codec voice data together. With each encoding/decoding of a lossy codec, voice quality is degraded.

Wideband Voice Codecs

Wideband codecs, such as G.722.2, are available. Telecom standard toll quality voice (G.711) is considered narrow-band with a frequency range of 300 to 3,400 Hz, whereas wideband codecs are 50 to 7,000 Hz. This provides better audio acuity for plosives (p, t, k, d, and b) and fricatives (s, f, v, and z).

While wideband codecs are used in some VoIP internet telephony applications and in some cellular systems, conventional telecommunications infrastructure supports only narrowband and requires transcoding at boundaries, resulting in the loss of the extra frequency data of the wideband codecs. However, the use of wideband codecs is on the increase, and this will prove a challenge to mission voice communications in the future.

2.4.2 VOICE OVER IP SYSTEMS

2.4.2.1 General

The Marshal Space Flight Center uses the most advanced and reliable VoIP voice system for operations in the space business, called Internet Voice Distribution System (IVoDS). The IVoDS is used to receive, sum, and distribute mission voice conferencing data (i.e., 'loops'). The system uses a standard T1 facility to connect and exchange voice data between other NASA centers, international partners, and ancillary equipment (voice recorders, telephone system, and other sources of voice within MSFC).

The system is made up of several main subsystems:

2.4.2.2 Conferencing Server

The conferencing server has the system data base, the VoIP System manager application (DICES/COTS), all required system monitoring, inventory, and logging facilities. The HOSC Server is an High Availability (HA) and Fault Tolerant (FT) class. The T1 gateways and all user Keysets (Clients) connect back to the VoIP Manager. The System uses an Administration Terminal (SAT) for control, status and DB entry. The terminal software runs on MS Windows PC or a Linux workstation/PC. The conferencing server communicates with other network services for user authentication (LDAP, RSA or other), DNS, Time, Auditing, Logging, Backup, and other available network services.

2.4.2.3 Functional Overview/Processing

The IVoDS is a system that utilizes IP communications between the end user keysets (Clients), the central processor (Conference Server) and the T1 Gateways (T1 channels to IP Packets). The IVoDS is a Quintron COTS product called DICES VoIP and is a proprietary product.

The IP protocol used is TCP/IP. The VoIP Server utilizes three main processes to receive, sum and distribute audio. The Input Manager receives audio and control inputs from the T1 Gateways and Keysets. The Input Manager receives audio and forwards the audio samples to the Summation Manager. The Summation process adds all the current ‘talkers’ of a given loop and hands that output to the Distribution Manager. The Distribution manager forwards the summed voice to all users (T1s and Keysets) monitoring or subscribed to the audio. The Conference Server, T1 Gateways and Keyset Clients utilize dynamic buffers to smooth the audio quality, which can be degraded by network traffic delivery issues.

2.4.2.4 T1 VoIP Gateway

The gateway has 2 basic functions: the TDM side and the IP side. The TDM interfaces with external equipment solely over a standard T1/E1 and can be configured to support all framing, coding, and signaling protocols. The T1 card supports 4 ports. The network/IP side performs voice exchange between the configured T1 voice channels and the conference server. The VoIP application converts TDM to IP and manages the associated signaling for all channels.

2.4.2.5 User Software Client

The software client can be run on a PC or dedicated hardware solution.

2.4.2.6 Hardware Keysets

Hardware keysets have full color touch screens. The client can be configured to support as few as 4 keys or as many as 80 keys. The client supports multiple pages (i.e., page 1 through 10) and provides many visual indicators for user awareness and operator selection. The HOSC Keysets utilize a 12-inch touch screen and display 45 keys. Every loop can be monitored simultaneously; talking is limited to one loop at a time.

2.4.2.7 Software Keysets (Running on User PC or Workstation)

Users download the software from the HOSC Portal. A client is available for PC, MAC, or Linux operating systems. The software client has all the same functionality as the hardware keyset. Users can access the system from any IP network.

Most of the ISS Payload user community using the IVoDS gain access using the commodity internet service provider networks. These networks have very low service level guarantees and often can lead to lost packets, delays and disconnects. The system provides features to compensate for these conditions, such as timers and buffering to offset network disruptions.

Figure 2-4: Example of the Keyset Layout

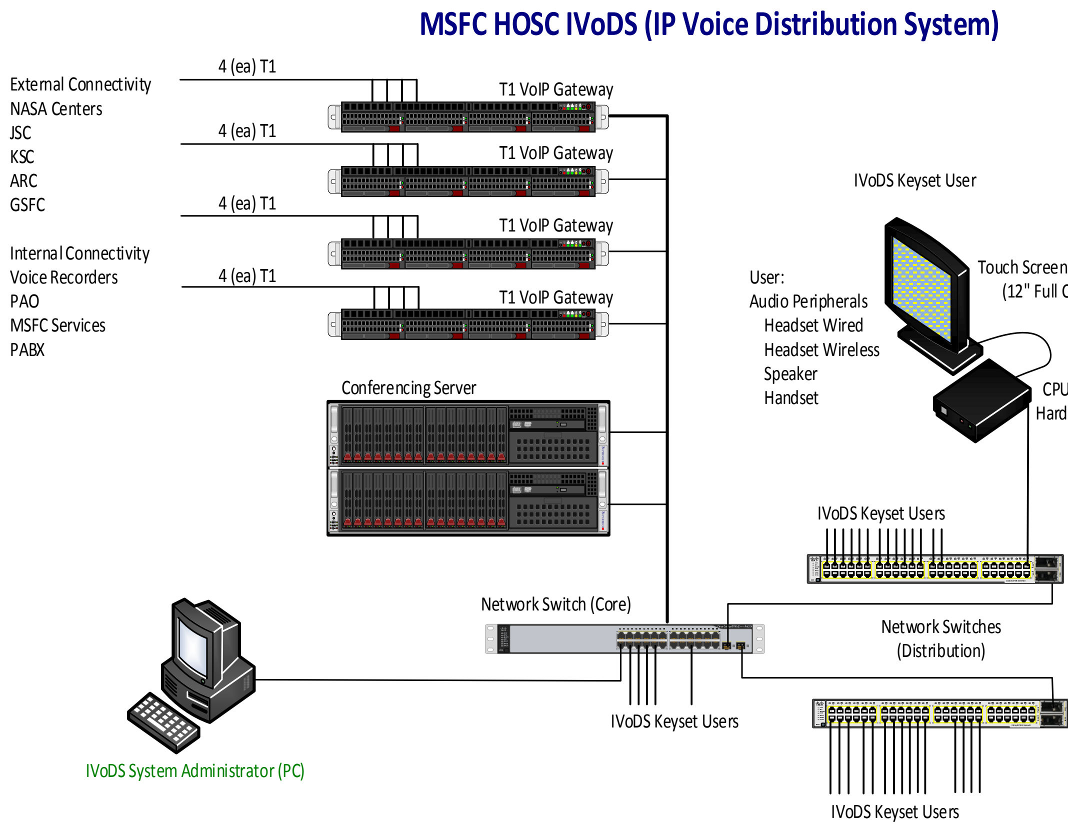


Figure 2-5: Example of the Keyset Layout

2.4.2.8 System Considerations

When implementing an IVoDS or equivalent system, the overall cost will be impacted by the operational concept and user base the system is going to support. A real-time support system should be implemented with a minimal amount of potential Single Points of Failure (SPoFs). A SPoF can exist in the underlying network as well as the choice of server class. Additionally, the user audio experience depends on the quality of audio peripheral equipment.

The considerations below need to be taken into account:

– Conferencing server options are

  • standard class server—limited redundancy and not fault tolerant, inexpensive, availability 95.9 percent;

  • high end class—fully integrated software and hardware, is highly available and fault tolerant, availability achieves 99.999 percent, relatively higher cost.

– The server choice depends on the site operations but should be procured with the VoIP application loaded, installed, verified, and licensed.

– T1 gateway:

  • standard server with some redundancy;

  • supports four T1 ports with availability of 99.5 percent;

  • can be purchased with the VoIP application.

NOTE – MSFC system is configured for 40 T1 ports.

– user keysets:

  • dedicated CPU and touch screen;

  • can be procured and loaded with the VoIP Client software and license.

NOTE – MSFC system is Licensed for 165 Keysets.

– software client:

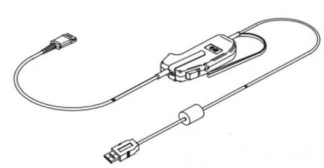
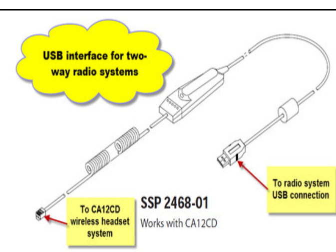
  • can be run on any PC, MAC, or Linux system/work station;

  • clients are licensed separately.

– headset information.

NOTE – Other options are available, the items below are specific to MSFC.

Table 2-1: Headset Devices Examples

Monaural with noise cancellation headset top	
USB with PTT headset bottom (this is the sound card)	
USB with PTT adapter (this is the sound card and it is used to connect to the wireless units, handsets, or speakers)	

Wireless units (base and wireless transmitter) Concerning about the picture to right: the coil cord with the PJ Plug is not used (the SSP 2468 is connected in its place)	

2.4.2.9 Latency, Jitter, and Packet Loss

Voice quality is affected by latency, jitter, and packet loss.

Latency is the delay of voice communications due to the following:

a) transport delay:

   1) light time delay,

   2) ground network transport delay;

b) processing delay: the encoding and decoding of voice, along with the associated buffering of analog or digital voice signal;

c) packet size buffer delay: where applicable, latency may be incurred when multiple samples of voice data are collected prior to transport;

d) boundary delay: the buffering, processing, and possible transcoding at the boundaries between transport domains.

Jitter is variations in the interval and arrival time of individual voice packets. Queuing and multiplexing of disparate data types and sizes within a communication channel results in some manner of jitter. Transport Layer framing may have an impact on jitter.

Packet loss, in which incoming packets are discarded, normally occurs in IP networks as a result of network congestion. 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 of excessive packet loss is poor voice quality or loss of voice communications entirely.

2.4.2.10 Voice Quality

Voice quality assessment is possible through automated and subjective human evaluative methods. The long-standing telecom standard Mean Opinion Score (MOS), defined in ITU P.800 (reference [12]), 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 narrowband digital compression codecs (e.g., G.729) may result in lower comparative scores.

The Diagnostic Rhyme Test/Diagnostic Acceptability Measure (DRT/DAM) testing provides for measuring intelligibility of voice communications.

Whereas the quality standard has been the MOS of G.711 encoded voice, the increasing use of wideband codecs will likely result in a new expectation for quality voice communications. As terrestrial systems evolve to handle wideband codecs, the use of such codecs in human spaceflight can be expected.

2.4.2.11 Quality of Experience (QoE)

The console users carry out voice tests regularly before important mission phases. The user normally uses the expression ‘5 × 5’, or ‘5 × 3’, or ‘3 × 3’, indicating the quality of the voice communication. This range, from 1 indicating intelligible communication to 5 indicating good voice quality, is specified in ITU-T P.800.

Quality of service is related to the network. Quality of Experience is related to human perception, and should be ‘5 × 5’ (see annex D).

2.4.2.12 Meta Information for Voice Data

Traditionally, the archiving of voice data related to human spaceflight has been an administrative function. Voice is recorded with accurate date and time information, and that information is correlated with mission events by way of a method outside of the voice archive. Internet protocols provide for additional voice meta information, such as the source identifiers in Real-time Transport Protocol (RTP), and voice extensions in Extensible Markup Language (XML) can provide richer meta information to a voice archive.

2.4.2.13 Recording and Storage of Voice Data

Codecs that are most effective store voice data at high compression rates without significant quality loss to optimize the efficient use of storage facilities. However, the low cost of storage and archival media may allow the storage of voice data in its captured format.

2.4.2.14 Private Conferences

Private conferences occur with family and friends, a flight surgeon, or other participants in situations in which private voice communication is desired by the participants or required by law. Executing a private conference is largely an administrative function of the voice technicians for Shuttle (in the past) and ISS. Function keys provide the voice technician the ability to switch a forward link and return link voice channel from a common air-to-ground or space-to-ground voice loop to a private conference loop. Authorized parties are then connected to the private conference loop. Voice technician personnel also check and disable other interfaces as may be necessary to ensure a private conference.

Once the parties are connected and voice checks are complete, the voice technicians disengage from the private conference and monitor voice performance only through audio meter readout of the circuits involved. The private conference continues as scheduled or until a designated party calls to notify the voice technicians that the private conference has ended.

2.5 SPACECRAFT AUDIO COMMUNICATIONS

2.5.1 SOYUZ

The Russian Soyuz capsule is the only one used at the moment to transport people to the ISS. Soyuz has two VHF channels:

– ISS VHF1 is used for S/G communication over Russian ground stations (RS ISS S/G 1 MH21) and also used as an alternate link for S/G communications when ISS is over NASA VHF ground stations. Uplink: 139.208 MHz; Downlink: 143.625 MHz.

– Soyuz VHF2 is used for S/G over Russian ground stations (RS ISS S/G 2 MH24) and space-to-space communication. This is the only VHF frequency pair carried on Soyuz. Soyuz VHF2 communication through NASA VHF ground stations is authorized for emergency use only per Flight Rule (D3-1). Uplink: 130.167 MHz; Downlink: 121.750 MHz.

Soyuz, when approaching proximity operations are done manually, severs the connection to the ground and uses only the VHF/S-band connection from spacecraft to spacecraft.

2.5.2 COMMERCIAL CREW (CYGNUS AND SPACEX DRAGON)

2.5.2.1 General

ISS standard communications are used when the spacecraft closes to within tens of kilometers to the ISS and ship-to-ship voice communications are established.

Communication and telemetry monitoring will be shared between the Commercial Vehicle Control Center (CVCC) and the ISS mission control facilities, Mission Control Center - Houston (MCC-H). MCC-H Mission Authority will be established to ensure ISS, spacecraft, and crew safety. When in close proximity to the ISS, after receiving approval from both the spacecraft and MCC-H, the spacecraft begins a final approach to a NASA-specified docking port on the ISS.

2.5.2.2 Voice Communication in Breathing Apparatus

The CTS (Crew Transportation System) shall provide voice communication between each of the crew members and the Voice Communication in Breathing Apparatus

2.5.2.3 Suit Communications

The suit shall provide two-way voice communications between each of the crew members and the CVCC. Rationale: When the crew is suited, they will not be able to hear each other. Additionally, the vehicle audio speakers will not be heard/intelligible, particularly in a vacuum. In these cases, the suit must allow for communication between crew members and between the crew and the CVCC.

2.5.2.4 Accommodate NASA Personnel at CVCC

The CTS shall accommodate 4 mission-essential NASA personnel at the CVCC with capability to monitor the pre-launch and mission-operations activities and communicate (via voice loops and other means) with NASA personnel in ISS MCC-H.

2.5.2.5 Communications with CVCC after Landing

The spacecraft shall provide 2 hours of cumulative post-landing two-way voice communications over a 24-hour period between the crew and the CVCC, until recovery/rescue forces have removed the crew from the vehicle. The crew should have periodic two-way voice communications with the operations team sufficient to aid in spacecraft safety and coordinate the recovery or search and rescue and other crew-survival efforts during this timeframe. The communication will need to be maintained until the search and rescue, or recovery, teams have recovered the crew.

For nominal and off-nominal landing scenarios, 5 minutes of continuous two-way voice every hour for 24 hours will satisfy search-and-rescue needs. This requirement does not preclude the use of portable devices as part of a strategy to provide two-way voice communications between the crew and the CVCC until recovery forces have removed the crew from the vehicle.

2.5.2.6 Spacecraft Voice Communication with Recovery/Rescue Forces

The spacecraft shall provide two-way voice communication between the crew and the recovery/rescue forces from the time the recovery/rescue forces are within direct line of sight of the spacecraft until the crew has been removed from the spacecraft.

The spacecraft shall be capable of communicating on International Air Distress (IAD) frequency of 121.5 MHz and Military Air Distress (MAD) frequency of 243.0 MHz.

The crew should have two-way voice communications with the recovery/rescue forces that will be conducting the search and rescue and other crew survival efforts during this timeframe. The communication will need to be maintained until the search and rescue or recovery teams have recovered the crew for nominal and off-nominal landing scenarios. It is assumed that it will take 4 hours for the rescue and recovery forces to remove the crew from the spacecraft. The CTS should develop and coordinate other special frequencies for use beyond the IAD and MAD frequencies for all operations that are not considered search and rescue. This requirement should not require the crew to open the egress hatches to maintain communication with recovery/rescue forces because the spacecraft may be in a contingency landing situation, such as a water landing, which would preclude opening the hatches

2.5.2.7 Voice Communication with Crew

The CTS shall provide single failure tolerant two-way voice communication between the CVCC (s) and the spacecraft crew from pre-launch through landing and during aborts.

Two-way voice communication between the CVCC (s) and the crew is required to execute the mission and resolve anomalies, should they occur. The intent of this requirement is to provide single failure tolerance in both the ground and onboard communications systems in order to ensure communications availability during all flight phases. This requirement is not meant to imply 100-percent continuous communication for all phases of flight. On orbit communications, coverage requirements will be dictated by the specific design requirements of the spacecraft.

2.5.2.8 Communications Coverage

The CTS shall provide communications coverage (two-way voice and telemetry) between the integrated space vehicle and CVCC(s) during at least 90 percent of the powered ascent flight phase and 65 percent during the entry flight phase to supported landing sites.

Historically, the ascent and entry phases of human space flight have been the timeframe of greatest risk for LOC (loss of communication). For powered ascent, there are a multitude of abort options and timely systems responses that ground personnel can assist with, leading to the requirement for near-continuous communications that can be accommodated by ground- and/or space-based communication assets. For entry, although the risk is high, there are fewer options available for the crew, and thus the requirement for continuous communication is less than that for ascent. Entry communications coverage to supported landing sites is intended to limit the number of trajectories for nominal entry assessment. The entry communication analysis should be focused on critical events (separations, chute deployment key navigation events) and the final phase of landing where the risk is the highest.

2.5.2.9 Private Audio

The CTS provides for a private audio communications capability between the NASA crew and the NASA flight surgeon during non-docked, on orbit operations.

This private, or secured, communication provides for privacy of medical information between the NASA crew and the NASA flight surgeons. Private voice communications are not required during ascent and entry due to the short nature of these flight phases. Voice communications are considered private with respect to those on the ground. No additional impacts to vehicle design should result from establishing private communication, as the information is not required to be private between crewmembers.

2.5.2.10 Integrated Voice Communications during ISS Proximity and Docked Operations

The CTS provides simultaneous two-way voice communications between the ISS, CTS spacecraft, CVCC, and ISS Mission Control Center during:

a) free flight when the CTS spacecraft is within 10 kilometers of the ISS;

b) docked operations when the crew is in the CTS spacecraft.

The intent of this requirement is that the CTS enables a multi-party voice loop, conferencing the ISS, CTS spacecraft, CVCC, and ISS Mission Control Center when the CTS spacecraft is within 10 kilometers of the ISS, to support nominal and contingency operations. This requirement provides voice communication when the crew is isolated in the CTS spacecraft while attached to ISS (i.e., Safe Haven) and enables conferencing of all the operations entities (ISS crew, CTS spacecraft crew, CVCC flight control team, and ISS flight control team) to support proximity and docking operations. This does not require direct point-to-point communication between all parties.

2.5.2.11 Communication System Design

Audio Volume Control

The spacecraft needs to provide a volume control from 5 to 100 percent of maximum for each audio channel carrying voice communications.

The crew needs to have the ability to adjust volume for each audio channel in order to hear and communicate under expected background and radio noise conditions. For example, volume control with 10 steps of 6 dB increments provides acceptable range. Adjustability down to 5 percent is significant, given the maximum audio level needed for the launch noise environment, as it ensures that crew can adjust volume to a comfortable level in the lower SPL on orbit environment and have adequate speech-to-ambient-noise margin. Conversely, the system must ensure that the crew does not miss radio calls due to low audio volume or elevated background noise (e.g., alarms present, multiple active communications channels). Visual indicators of audio signal can be an effective means of alerting crew to radio calls.

For commercial crew vehicles, there are no redundancy requirements for the voice system and S/G voice loops.

2.5.3 ORION

The US Orion program specified a VoIP voice communications capability with the Orion capsule, G.729 in RTP/UDP/IP for air-to-ground voice communications through CCSDS Encapsulation Service (reference [17]) and multiplexed with command and telemetry. Initial Extra Vehicular Activity (EVA) voice communications capability was specified as tethered Linear Pulse Code Modulation (LPCM) to the Orion capsule. Additional voice communications capabilities were to include Emergency Communications (EC), DV, and Search and Rescue (SAR) voice communications

Emergency Communications. EC was an outgrowth of the Apollo 13 experience, in which a near catastrophic failure led to a severe power shortage that required many on-board systems to be powered down. The concept for Orion was that some systems may be powered down, including the primary radio equipment. Low power EC equipment could then be engaged to transmit voice as well as a very limited amount of command/telemetry at a rate of (perhaps) 12.4 kb/s.

EC voice was encoded as G.729 data and transmitted over CCSDS AOS (reference [15]) Virtual Channel Access (VCA) service with no IP/UDP/RTP protocol wrapping.

Dissimilar Voice. DV provides assured voice communications through a parallel voice communications link with the crew during pre-launch, launch, and ascent. Both ground operations and mission operations utilize DV. For Orion, DV was specified as G.729 encoded voice exchanged through S-Band RF from ground stations at the launch site and other selected sites as required to accommodate ascent trajectories. CCSDS AOS VCA service was the expected transport packaging. Given the latency and jitter expected with the communications systems, the Orion Program determined that onboard synchronization of DV with the primary voice communications was required. This was to be accomplished through the processing of the respective ‘presentation times’ associated with each frame of voice data, given their respective communication channels.

2.5.4 INTERNATIONAL SPACE STATION

The International Space Station (ISS) utilizes Modified Residual Excited Linear Predictive (MRELP), a codec with data rate at 9.6 kb/s for space-to-ground voice communications. As originally designed, 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. Nowadays this system is still used for the S/G 1 and 2 loops.

Additional voice channels for ISS are implemented using VoIP G.729 within RTP/UDP/IP. These channels are the current S/G 3 and 4 voice loops.

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.

2.6 VOICE CONNECTIVITY AMONG AGENCIES

2.6.1 GENERAL

Voice communication is one of the crucial mission applications running over the International Ground System (IGS) Wide Area Network (WAN). The mission control centers are:

- Columbus CC (COL-CC);
- MCC-Houston (MCC-H);
- Huntsville Operations Support Center (HOSC);
- European Astronaut Center (EAC);
- MCC-Moscow (MCC-M).

Universities and other payload users use the Internet with services like Premium IP (PIP).

Each center has one or more voice conferencing systems. These conferencing systems interconnect using synchronous TDM-based (E1/T1) or VoIP based interfaces. MCC-M uses 12 voice loops exchanged using Cisco-based VoIP with Col-CC.

The following table details the voice loops that may be configured between the various control centers of different space agencies. Multiple configurations provide for the mapping of a large number of voice loops over a smaller number of physical circuits bridging between the agencies to accommodate different scenarios and mission phases according to the corresponding voice formats.

Table 2-2: Voice Channels

CSA – MCC-H	Six major configurations provide concurrent access to between 64 and 66 voice channels
JAXA – MCC-H	Ten major configurations provide concurrent access to 40 voice channels
COL-CC – MCC-H	Three major configurations provide concurrent access to up to 48 voice channels
MCC-M – MCCH	Three major configurations provide concurrent access to up to 48 voice channels
COL-CC – MCC-M	One configuration provides concurrent access to 12 voice loops
HOSC – COL-CC	Two major configurations provide concurrent access to 24 voice channels
HOSC-MCC-H	Two major configurations provide concurrent access to 48 voice channels
HOSC	IVODS, connections using the Internet and T1 to centers in Europe, to Col-CC, and payload users

Presently the E1/T1 synchronous voice communication is transported over MPLS. For the HOSC, Internet is also used. That can be changed in the future.

All these voice channels can be mapped to different voice loops, as is defined in the corresponding voice format.

2.6.2 VOICE-TO-TEXT CONVERSION

Voice-to-text, also known as Automatic Speech Recognition (ASR) (references [9] and [5]), along with the reverse process known as 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.

There are several voice-to-text and TTS software products available. Product selection is beyond the scope of this book.

2.6.3 VOICE AND AUDIO COMMUNICATION SERVICES

2.6.3.1 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, have different latencies, and need to be synchronized.

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.

An 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.

Most of the new cameras are already using embedded audio (HD-SDI or SD-SDI) on the ISS; the downlinked audio is embedded in the video, but the audio going to the ISS is not always synchronized with the video. In those cases, in order to avoid echoes or artifacts, the lip sync is recommended.

For some public-affairs events with educational institutions where VSee or Skype are used, the Lip Sync is not necessary because the audio is embedded in the video. The quality is lower than that of the audio used for TV or Governmental Events, but acceptable.

2.6.3.2 Voice Recording/Playback

2.6.3.2.1 General

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. A common use is short-term playback 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, loop-specific recording, and interface-specific recording.

2.6.3.2.2 User-Specific Recording (Keyset Recording)

User-specific/keyset recording is the sum of all audio traffic for a specific user and is dependent on the user's loop selection. User-specific recording represents the presence of all audio signals at a user position.

Keyset recording and playback of a user position is used for a short time span only; a user is able to replay the last minutes of all audio arriving at the keyset. It is mainly used for replays of unclear communication.

2.6.3.2.3 Loop-Specific Recording (Loop Recording)

Loop recording is defined as the sum of all audio traffic for a specific loop. It is user independent. Loop recording represents the presence of an audio signal within a loop.

Loop recording is the most common recording for space mission operations. All audio of a loop is stored separately within the recording system. This way, loop-specific investigations are possible (e.g., dialog between different positions).

Loop recording is very demanding in terms of storage capacity and correct time-tagging to allow replays of specific events based on a given time span.

2.6.3.2.4 Interface-Specific Recording (Interface Recording)

Interface recording represents the presence of an audio signal at an interface and is able to identify the direction of the signal (outgoing/incoming).

Interface recording is a technical support recording to verify audio sources within loops. It is aligned to loop assignments of channels.

Interface recording is mainly used to identify failure sources such as noise bursts within a loop, open microphone conditions, or other audible misalignments.

2.6.3.2.5 Etiquette and Protocol

To have access to the voice system in space operations, etiquette and protocol are mandatory (roles and positions to talk).

The correct functioning of the end-to-end voice system requires some support from the environment, particularly to support availability and reliability of the networks, but also to ensure that the people operating the system do so according to established rules. That means console operators must to be certified in the use of the voice system, as well as trained in the etiquette and protocol.

Each space agency has its own training programs for using its voice system, but the etiquette and protocol are common for all the space agencies. The definitions are in the NASA SATERN (System for Administration, Training, and Educational Resources for NASA) educational system. This system is accessible for all the users of a voice loop in the ISS program; this training is mandatory.

For satellite missions, similar training sessions are provided for each space agency and are very similar in content; this training is also mandatory for each participant of a space mission.

There are rules for how to talk, how to state a position, how to ask which position to talk to, which reduced and specific vocabulary should be used, etc.

3 TECHNICAL DRIVERS

3.1 INTRODUCTION—OVERALL APPROACH

The table and illustrations below describe the overall scenario for lunar missions and can be transposed for Mars missions.

In short, voice communications can be summarized in four defined segments:

a) Earth segment (between terrestrial facilities, including emergency communications and SAR);

b) lunar or Mars segment (between lunar or Mars facilities, astronauts, and Earth, including emergency communications and SAR);

c) short-haul segment (between terrestrial and space-borne facilities, including emergency communications); and

d) long-haul segment (between terrestrial and space-borne facilities, including emergency communications).

Table 3-1: The Voice Segments

Voice Segments	Description	Technical Constraints
Earth ground	Control center system and interfaces to other control centers and remote users, emergency communications, SAR	– conference loop capability – point to point – latency for co-located personnel – criticality, availability – bidirectional network – continual comm.
lunar or Mars ground	Voice communications for EVA, Habitat, Rover; spacecraft, emergency communications, SAR	– conference loop capability – point to point – latency for co-located personnel – criticality, availability – bidirectional network – continual comm.
short-haul: near Earth, near Moon, near Mars	Voice communications through orbital nodes to and among ground nodes, whether around the Earth, Moon, or Mars; emergency communications, SAR	– point to point – criticality and availability – channelization and bandwidth – intermittent to continual comm. – bidirectional to non-bidirectional

Voice Segments	Description	Technical Constraints
long-haul: Moon to Earth, Mars to Earth	Mars to Earth voice communications, emergency communications, SAR	– point to point – criticality and availability – channelization and bandwidth – intermittent to continual comm. (DTN should be used) – bidirectional to non-bidirectional

Figure 3-1 depicts the following:

– lunar ground segment;

– lunar short-haul segment;

– Moon-to-Earth long-haul segment.

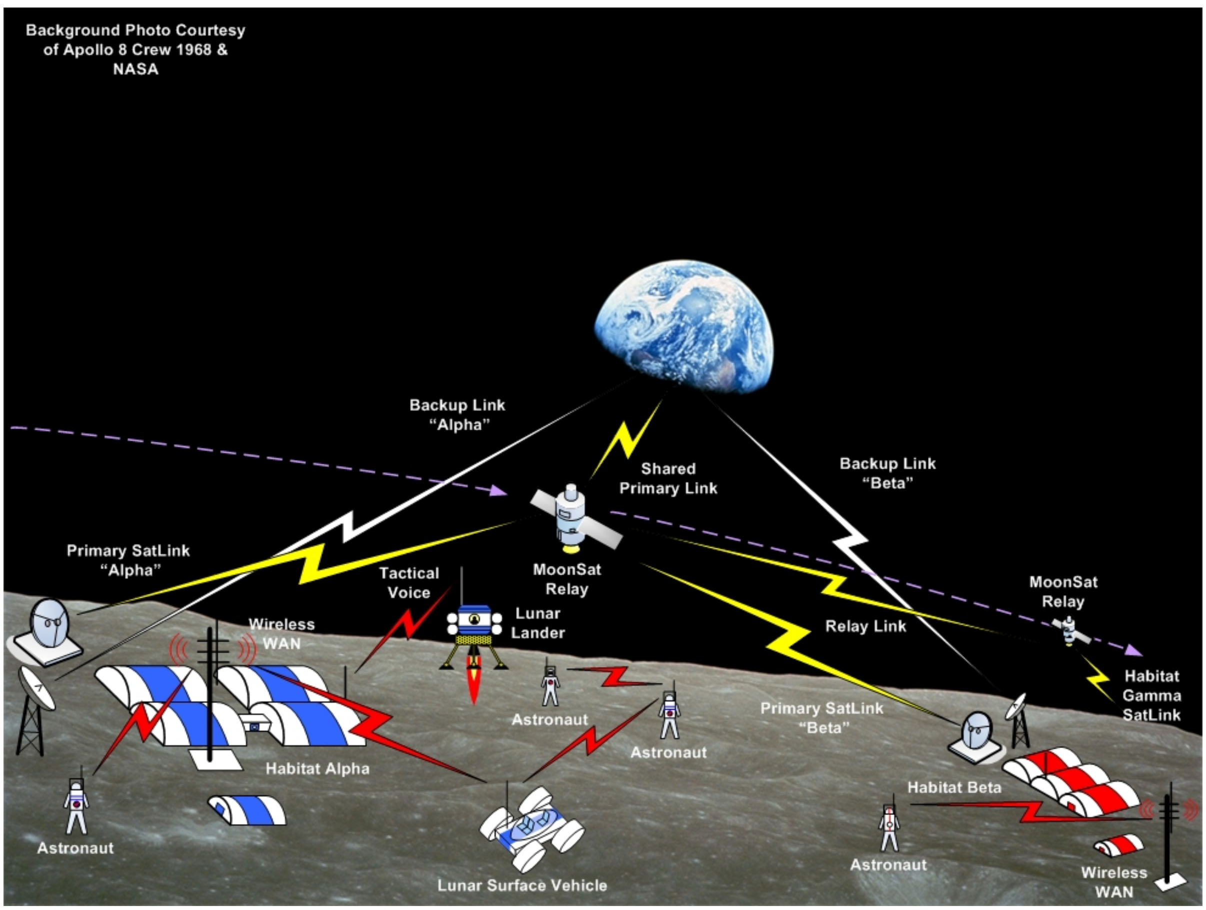


Figure 3-1: Lunar Surface Operations

3.2 EARTH SEGMENT

3.2.1 GENERAL

The Earth segment is characterized by a collection of control centers, many having their own voice conferencing systems connected with each other via a limited number of parallel voice channels. Each center may have a number of remote voice users connected via VoIP/Internet, or MPLS, or other means. Commercially available equipment tends to dominate the Earth segment in the telecommunications, Internet, and networking domains.

3.2.2 KEY DRIVERS

Latency. For co-located users, for example, an FCT in a flight control room, voice latency must be near to the normal sound travel time among the users. For point-to-point connections, latency becomes much less an issue.

Channelization. Channelization of voice communications needs to be provided between centers for configuration management and security reasons.

Criticality/Availability. Voice systems and their interfaces are critical capabilities. Experience shows that when they do not work, back-ups are immediately necessary. Hence not only should voice systems be on failure-tolerant equipment, but their external interfaces should be supported by redundant communication channels and have easy if not automatic fail-over.

Table 3-2: Earth Segment Communications Options

Layer	Mechanism	Comments
Physical	E1, T1, 802.3	
Data Link	MPLS, Ethernet	
Network	IP (SDP, RTCP, RTP, SIPv2 [reference [6]]), DTN	For interoperability point of view
Transport	TCP/UDP (IP), TDM, RTP, other	
Session		PTT
Presentation layer	G.711, G.729, G.722, G.728, etc.	
Application	Voice summation Voice Recording/playback	

Codec. A number of compression schemes exist and are employed today. It is in the interest of interoperable agencies to adopt a common scheme for compression to avoid multiple protocol conversions at interface boundaries, which results in a degradation of signal and the possible introduction of audio artifacts.

Connectivity. VoIP is used in most of the new voice systems. Other technologies have a place in voice connectivity or distribution:

– MPLS: does not allow circuit based emulation, but has acceptable QoS/performance;

– IP (Internet): characterized by poor quality and packet loss but is ubiquitous and relatively inexpensive.

3.2.3 FUNCTIONAL ASPECTS

3.2.3.1 Voice Summation and Conferencing

It is likely that the requirement for large voice-conferencing systems in the various control centers will remain for the foreseeable future. In these centers the following important factors remain:

– the number of keysets (i.e., users);

– the number of voice loops (voice conferences);

– the number of keysets (users) connected to any individual voice loop;

– the number of voice loops to which any individual keysets may be connected;

– the latency experienced by co-located users on common voice loops;

– the number of channels allocated to various external interfaces.

The high performance, high capacity central voice matrix remains TDM-based voice processing and switch equipment. VoIP is used as a distribution technology, and it is replacing the matrixes.

3.2.3.2 Channelization

It is important to define a limit for the number of channels available between centers, as an unlimited number of channels would quickly allow the definition of too many loops between centers. Managing the number of loops between control centers is a key configuration issue.

This is where the definition and use of voice formats becomes extremely important.

As previously explained, 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.

Different operations modes must be defined 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 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.

3.2.3.3 Security

On the Earth segment, security for voice links is important for:

– ensuring access control to the voice systems;

– ensuring the privacy of medical conferences;

– ensuring the privacy of 'private' calls (crew/family conferences);

– ensuring the confidentiality of voice communications related to critical operations or events;

– controlling access to voice records.

Voice system access is traditionally based on username/password, and sometimes on digital certificates (e.g., via an LDAP server). Voice recognition could be a future method for authentication.

For ensuring privacy or confidentiality, different techniques have traditionally been used, such as voice scrambling and, more recently, encrypting of digital voice communications. It should be noted that the need to store encrypted voice links in an encrypted format (and being able to decrypt the recorded voice) implies specific challenges in key management and implies metadata is carried by the voice signal, which includes the key ID.

3.3 LOCAL SEGMENT (LUNAR OR MARS ON GROUND)

3.3.1 GENERAL

Local segments, whether lunar or Mars ground segments, are likely to be dominated by the use of existing commercial telecommunications and networking capabilities adapted from the Earth ground segment, like communications over copper wire, optical fiber, WiFi, radio, or Bluetooth. To overlay these technologies upon hardware and the Physical Layers of communications in harsh environments will be a significant challenge. A local-segment voice does not take into consideration communications with other celestial bodies.

3.3.2 KEY DRIVERS

This segment, in the global scenario, is defined by a local infrastructure composed of vehicles such as rovers (crewed or non-crewed), crew and teams in suits (EVAs), habitat modules, science modules, and (perhaps) communications infrastructure modules. The main driver is the criticality of the voice channels given the dangerous activities performed by the crew during EVAs:

– voice communications for EVA/rover need to be redundant;

– the voice intercommunications interface may be voice activated to select voice channels by EVA/rover;

– emergency voice channels and equipment should be available for the astronauts during EVAs and for the rescue teams.

Table 3-3: Local Segment Communications Options

Layer	Mechanism	Comments
Physical	802.3, 802.16, Wireless LAN, RF links, Proximity 1 (reference [16])	UHF radio
Data Link		
Network	IP (SDP, RTCP, RTP, SIPv2), DTN	DTN only considered from interoperability point of view (multiagency)
Transport	TCP/UDP (IP)	
Session		
Presentation layer	G.711, G.729, G.722, G.728, etc.	
Application	Voice summation Voice Recording	

3.3.3 FUNCTIONAL ASPECTS

3.3.3.1 Voice summation

It is required for voice loops to be shared among the members of a working team on an EVA. Voice latency of the voice loop should be very low for those co-located. For example, two suited crew members working as a team may desire low latency voice between themselves, while the communications link to distant support personnel may experience a far longer latency.

In addition, it is clear that as ground operations grow in complexity, multiple voice-loops capability is necessary to support multiple and concurrent activities.

3.3.3.2 Secure Communications

Secure and confidential communications should be built into the capability.

3.3.3.3 Emergency Communications

In the event that communication with central modules is somehow interrupted, emergency communication capabilities should exist, whether to orbiting satellites or at a low level directly back to Earth.

3.3.3.4 Mesh Capability to Provide Redundancy

Interoperable nodes or mesh-capable nodes would be ideal for a lunar or Mars local segment. In this concept, any node can relay voice communication traffic to another node. This would allow EVA1 to communicate with the habitat via EVA2, in case EVA1 experienced loss of signal with the habitat.

3.4 SHORT-HAUL SEGMENT

3.4.1 KEY DRIVERS

Bandwidth and channelization begin to be an issue with short-haul segments, though more so with long-haul segments. For short-haul segments, bandwidth must be shared with other traffic, and with prioritized traffic.

Table 3-4: Short Haul Communications Options

Layer	Mechanism	Comments
Physical	RF transmission, optical communications	
Data Link	AOS, CCSDS Encapsulation Service, Proximity 1	(See references [15], [16], and [17].)
Network	IP (SDP), BP	(See reference [18].)
Transport	UDP (RTCP, RTP, SIPv2IP), VOIP (RTP, UDP, IP, CCSDS Encapsulation, AOS	(See references [15], [17], and [18].)
Session		
Presentation layer	G.711, G.722, G.728, G.729, AAC	
Application	Voice summation, Recorder	

There are a number of options that may be explored for voice services in a local lunar or Mars segment. Some will parallel the technologies employed for voice communications on Earth. The following is a representative but by no means an exhaustive list:

– Option 1: Voice over DTN (references [4] and [8]);

– Option 2: Voice over AOS (multiplexed voice traffic) (references [15]);

– Option 3: VoIP over CCSDS Encapsulation packet (references [17], and [18]);

– Option 4: VoIP over typical network topologies.

3.5 LONG-HAUL SEGMENT

3.5.1 KEY DRIVERS

Several key drivers exist in regard to long-haul segments:

– latency;

– channelization, bandwidth utilization, and traffic prioritization;

– secure communications;

– intermittent, bidirectional, and unidirectional connectivity.

Table 3-5: Long Haul Communications Options

Layer	Mechanism	Comments
Physical	RF transmission	
Data Link	AOS, CCSDS Encapsulation Service	(See references [15] and [17].)
Network	DTN, CCSDS BP	(See references [4] and [8].)
Transport	AMS/CCSDS, SMTP/CCSDS, FTP/CCSDS	
Session		file transfers
Presentation layer	MP3, AAC	
Application	E-mail, message service, CFDP	

3.5.2 FUNCTIONAL ASPECTS

3.5.2.1 Latency

Latency in the Earth-to-Mars long-haul segment is a factor of the light time necessary to traverse the distance. The one-way light time to Mars ranges from a minimum of about 5 minutes to a maximum of about 20 minutes. This latency will have an impact on the nature of voice communications. In this case, the voice communications with a one-way light-time delay of about five seconds results in the participants engaging in message-oriented conversation, as opposed to dialog-oriented conversation. Only file transfers using DTN can be used for establishing a communication.

Latency in the Earth-to-Moon long-haul segment is about two seconds, and thus the voice processing delay begins to impact total latency. Lunar crew members talking with an Earth-based support team member should expect at least a five-second interval between the end of their utterance and the beginning of a response utterance from Earth.

3.5.2.2 Channelization, Bandwidth Utilization, and Traffic Prioritization

As the operations of a lunar or Mars outpost may increase in complexity, crew population, science and in-situ activities, and over-subscription of the long-haul bandwidth may become an issue. Bandwidth utilization must be long planned for and traffic prioritization schemes considered and employed.

3.5.2.3 Secure Communications

Secure communications are necessary to support the safety of the crew and to support the private nature of personnel health and family conversations. In addition, as science and in-situ work increases, the associated data may be proprietary in nature and thus should be communicated through secure channels.

3.5.2.4 Intermittent, Bidirectional, and Unidirectional Communications

Long-haul communications links to distant relay satellites and outposts may be intermittent in nature. Earth-based assets used in the communications link may not have 24×7 line of sight with the distant node. In addition, the connection may be of such a latent nature that it should perhaps be considered a unidirectional link.

For this application, bundle protocol should be considered (DTN).

ANNEX A

ACRONYMS AND ABBREVIATIONS

Term	Meaning
AGVE	air-ground voice equipment
AOS	Advances Orbiting Systems
ATM	asynchronous transfer mode
ATV-CC	automated transfer vehicle control center
BRI	basic rate interface
CCSDS	Consultative Committee for Space Data Systems
CELP	code excited linear predictive
CFDP	CCSDS File Delivery Protocol
CODEC	coder-decoder
CVCC	commercial vehicle control center
DSP	digital signal processing
DSS	Digital Speech Standard
DTN	Delay/Disruption Tolerant Networking
DV	dissimilar voice
DVIS	digital voice intercommunications subsystem
ETSI	European Telecommunications Standards Institute
FCT	flight control team
IEEE	Institute of Electrical and Electronic Engineers
IMS	IP multimedia subsystem
IP	Internet Protocol
IP-MUX	IP multiplexer-demultiplexer
IPX	Internetwork Packet Exchange
IRIG	Inter-Range Instrumentation Group
ITU-T	International Telecommunication Union – Telecommunications
JSC	Johnson Space Center
kb/s	kilobits per second
LDAP	Lightweight Directory Access Protocol
LPCM	linear pulse code modulation

Term	Meaning
MOVE	mission operations voice enhancement
MPLS	multi-protocol label switching
MRELP	modified residual excited linear predictive
NEO	near-Earth object
PABX	private automatic branch exchange
PTT	press to talk
PVC	permanent virtual connection/circuit
QSIG	Q signaling
RAD	RAD Data Communications
RMA	reliability, maintainability, availability
RTCP	Real-time Transport Control Protocol
RTP	Real-time Transport Protocol
SDP	Session Description Protocol
SIPv2	Session Initiation Protocol version 2
SNMP	Simple Network Management Protocol
STS	Space Transport System
SVC	switched virtual circuit
TCX	transform coded excitation
TDAC	time domain aliasing cancellation
TDM	time division multiplexing
TDMoIP	TDM over IP
TDRS	Tracking and Data Relay Satellite
UHF	ultra high frequency
VoIP	voice over IP
WAN	wide area network

ANNEX B

CODEC SHORT DESCRIPTIONS

B1 OVERVIEW

The following is a selection of short descriptions of existing voice codecs. Other codecs exist and are not included here.

B2 G.711

G.711 is international standard PCM stream encoding telephone audio on 64 kb/s channel:

- ITU–T G.711 Recommendation compliant;
- sample rate of 8 kHz, 8 bits per sample;
- T1 platform compatible and available on any Digital Signal Processor (DSP) and other platforms;
- normally 64 kb/s, can expand output rate to 104 and 112 kb/s;
- A-law or µ-law compressor output;
- selectable frame/buffer memory size according to system;
- usually simple API interface;
- compliant with non-eXpressDSP DSP standard;
- products need to be code re-entrant supporting multi-threading;
- dynamic memory allocation, can also allow static memory allocation;
- should be portable to any platform;
- usually thought of as good for resource- (bandwidth-) constrained operations.

B3 G.722

G.722 is a wideband voice codec in that the audio samples are twice that of G.711 and other traditional telecom codecs:

- ITU-T B.722 Recommendations compliant;
- sample rate of 48 kb/s, 56 kb/s, or 64 kb/s, 14 bits per sample;
- based on Sub-Band Adaptive Differential Pulse Code Modulation (SB-ADPCM);

- effective audio and speech compression for storing voice, digital circuit multiplication and telephony operations;
- provides direct interface with the PCM 8 kHz sampled data;
- can process sample to sample, or from blocks of different length;
- compliant with TI's eXpressDSP;
- code is normally re-entrant, supporting multi-threading and dynamic memory allocation; can also enable static memory allocation;
- variants include G.722.1 and G.722.2;
- considered effective for situations in which bandwidth is constrained and easy to integrate with applications.

B4 G.723.1

G.723.1 is a voice codec for Digital Circuit Multiplication Equipment (DCME) applications, audio/video conferencing, and other multi-media devices:

- ITU-T G.723.1 Recommendations compliant;
- sampling rate of 8 kHz, 16 bits per sample;
- high rate based on Multi-Pulse Maximum Likelihood Quantization (MP-MLQ), low rate on Algebraic Code Excited Linear Prediction (ACELP);
- annex versions support Voice Activity Detection, and Comfort Noise Generation;
- code is normally re-entrant and supports multithreading and dynamic memory allocation.

B5 G.726 (BASED ON ADPCM)

G.726 is used primarily overseas for trunks of the telephone network, as well as for voice data storage and other telephony applications:

- ITU-T G.726 Recommendations compliant;
- bit rates include 16 kb/s, 24 kb/s, 32 kb/s, and 40 kb/s;
- based on Adaptive Differential Pulse Code Modulation (ADPCM);
- sample-based or block-based input; A-law, 3-law, and 14-bit uniform 8 PCM input/ output;
- can process blocks of different lengths;
- MIPS and memory optimized versions are available;

– compliant with TI's eXpressDSP;
– code is normally re-entrant and supports multi-threading and dynamic memory allocation; also allows direct interface to enable static memory allocation;
– optimized for resource-constrained applications.

B6 G.728

G.728 is a low latency codec used for voice data storage and voice communications:

– ITU-T G.728 and G.729 Annex 1 compliant;
– bit rate at 16 kb/s stream rate;
– based on Low Delay Code Excited Linear Prediction (LDCELP);
– sample-based or block-based analog input;
– low latent at 0.625 ms frame size with 0.625 ms algorithmic delay;
– direct interface with 8 kHz PCM sampled data;
– compliant with TI's eXpressDSP;
– code is normally re-entrant and supports multi-threading with dynamic memory allocation;
– allows direct interface to enable static memory allocation;
– considered good for high quality speech, low MIPS requirements and ease of integration.

B7 G.729.1

G.729.1 is a primary codec for VoIP applications:

– ITU-T G.729.1 compliant;
– scalable bit rate 8-32 kb/s;
– has multiple annexes (A, B, C, D, E, F, G, H, I, J, C+);
– annex J provides a wide-band version, in which the frequency range is 50 Hz to 7 kHz, also known as G.729.1;
– outputs a frequency range of 50-4000 Hz at 8 kb/s and 12 kb/s rates;
– supports digital signal sample rate of 16 kHz and 8 kHz;
– uses Code Excited Linear Predictive (CELP) algorithm, and TDAC algorithm;

– considered very high-quality audio speech results, and a robust codec system.

B8 CVSD

Continuously Variable Slope Delta (CVSD) modulation is a voice encoding method used primarily in the military for digitally encrypted voice communications, an option for Bluetooth service:

– not an ITU recommendation;

– CVSD encodes voice at 1 bit per sample, to bit rates of 9.6 kb/s to 128 kb/s;

– immune to noise, robust to bit and synchronization losses, highly portable, and highly optimized.

B9 GSM

Global System for Mobile (GSM) is a popular voice encoding for mobile communications:

– ETSI 1987 compliant;

– GSM 06.10 Regular Pulse Excited-Long Term Prediction (RPE-LTP) Linear Predictive Coder;

– sample rate of 8 kHz results in 200 Hz -3.4 kHz audio at 13 kb/s;

– considered highly optimized code for situations where resources are constrained;

– extensions include: Enhanced Full Rate (EFR) GSM; Adaptive Multi-Rate Narrow Band (AMR-NB); Adaptive Multi-Rate Wideband (AMR-WB); a Hybrid ACELP/TCX Technique GSM AMR WB+.

B10 EVRC

Entrance Variable Rate Code (EVRC) provides improved error performance in variable rate operations:

– TIA-EIA-IS-127 compliant;

– Relaxed Code Excited Linear Predictive (RCELP) algorithm, modified for variable rated operations, and for robustness in the Code Division Multiple Access (CDMA) environments;

– rates include 9.6 kb/s, 4.8 kb/s, and 1.2 kb/s;

– considered highly optimized and ideal for resource constrained applications.

B11 iLBC (LOW BIT NARROWBAND)

Internet Low Bitrate Codec (iLBC) provides somewhat higher voice quality over G.729 while being robust to packet loss:

– conforms with and exceeds (in quality) G.729A, and G.723.1;

– operates at 13.3 kb/s, and 15.2 kb/s rates; frame size is 30 ms for 13.3 kb/s, and 20 ms for 15.2 kb/s;

– the code has been optimized for constrained resource applications**.**

ANNEX C

PROTOCOLS USED FOR VOICE TRANSMISSION

Depending on which kind of voice system will be implemented in the control center, and on the system's functionality, one or a combination of the protocols shown in table C-1 will typically be used. These protocols are referenced from the normative sections of this document.

Table C-1: Protocols Used for Voice Transmission

<table>
<tr><th colspan="2">Protocol</th><th>Notes</th><th>Where Defined</th><th>Ref #</th></tr>
<tr><td colspan="2">User Datagram Protocol (UDP)</td><td>None.</td><td>RFC 768 (STD 6)</td><td></td></tr>
<tr><td colspan="2">Real-time Transport Protocol (RTP)</td><td>RTP is used for MCC and space communications.</td><td>RFC 3550 (STD 64)</td><td></td></tr>
<tr><td colspan="2">Real-time Control Protocol (RTCP)</td><td>RTCP's primary function is to provide feedback on the QoS.</td><td>RFC 3550 (STD 64)</td><td></td></tr>
<tr><td colspan="2">Secure Real-time Transport Protocol (SRTP)</td><td>SRTP has a sister protocol called Secure RTCP (SRTCP)</td><td>ECP 3711</td><td></td></tr>
<tr><td colspan="2">Transmission Control Protocol/Internet Protocol (TCP/IP)</td><td>TCP/IP is appropriate for exchange of audio files.
TCP/IP is NOT recommended for voice communications.</td><td>RFC 675</td><td></td></tr>
<tr><td colspan="2">E1/T1</td><td>E1/T1 protocols support 30 (E1) or 24 (T1) digital channels simultaneously using TDM.</td><td>ANSI T1
ITU-T G.704/G.732</td><td></td></tr>
<tr><th colspan="2">VoIP Signaling Protocols</th><td></td><td></td><td></td></tr>
<tr><td rowspan="2">Session Control Protocols</td><td>Session Initiation Protocol (SIP)</td><td>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.</td><td>RFC 3261</td><td>[6]</td></tr>
<tr><td>H.323</td><td>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.</td><td>ITU-T H.323</td><td>[7]</td></tr>
</table>

Protocol		Notes	Where Defined	Ref #
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	[13]
	Megaco (H.248)	H.248 follows the guidelines of the Application Programming Interface (API) MGCP architecture and requirements described in RFC 2805. It was declared obsolete in 2008 by IETF.	RFC 2805	[14]

ANNEX D

EXPANDED DISCUSSION OF COMMUNICATIONS QUALITY AND IMPORTANT DEFINITIONS

D1 COMMUNICATIONS QUALITY

Five-by-five is the best of 25 possible subjective responses used to describe the quality of communications, specifically the signal-to-noise ratio. As receiving stations move away from an analog radio transmitting site, the signal strength decreases gradually, causing the relative noise level to increase. The signal becomes increasingly difficult to understand until it can no longer be heard as anything other than static.

Five-by-five is the best possible Readability and Signal Strength Report, previously used by some radio services.

The term has its origins in the Q code used for commercial radiotelegraph communication and was later adopted by other radio services, especially amateur radio. The first number is the answer to the question ‘How do you receive me?’ (Q code QRK), typically answered with ‘I am receiving [1–5]’, in which 1 is unreadable and 5 is perfect. The second number is the answer to the question ‘What is the strength of my signals?’ (Q code QSA), typically answered with ‘The strength of your signals is [1–5]’. ‘5 by 5’ was the answer to the QRK/QSA questions to indicate the best quality signal. ‘5 by 5’, meaning ‘the best quality’, was applied to situations other than radio communication by analogy.

The U.S. Army Radio Operator's Manual, FM 24-6 (1945 version) defines the report format as follows:

Report	Readability	Signal Strength
1	Unreadable	Scarcely perceptible
2	Readable now and then	Weak
3	Readable but with difficulty	Fairly good
4	Readable	Good
5	Perfectly readable	Very good

FM 24-6 further states, “Readability and signal strength reports indicate the desired method of transmission and should be used in that sense. Readability and signal strength reports are not exchanged unless ‘3’ or less, lack of any report being assumed to indicate satisfactory communications.”

The modern equivalent of that manual is ACP131(F), which includes these entries:

Signal	Question	Answer, Advice, or Order
QRK	What is the intelligibility of my signals (or those of...)?	The intelligibility of your signals (or those of...) is... 1) bad. 2) poor. 3) fair. 4) good. 5) excellent.
QRK	What is the intelligibility of my signals (or those of...)?	The intelligibility of your signals (or those of...) is... 1) bad. 2) poor. 3) fair. 4) good. 5) excellent.

In voice procedure (the techniques used to facilitate spoken communication over two-way radios), a transmitting station may request a report on the subjective quality of signal they are broadcasting. In the militaries of the North Atlantic Treaty Organization (NATO) countries, and in other organizations, the signal quality is reported on two scales: the first is for signal strength, and the second for signal clarity or ‘readability’. Both these scales range from one to five, in which one is the worst and five is the best. The listening station reports these numbers separated with the word ‘by’. ‘Five by five’ therefore means a signal that has excellent strength and perfect clarity, the most intelligible signal possible.

‘Five by five’ (occasionally written "'5 by 5", ‘five-by-five’, ‘5 × 5’, ‘5-by-5’ or even just ‘Fives’), by extension, has come to mean ‘I understand you perfectly’ in situations other than radio communication.

This ‘5 × 5’ formulation is used by all the space agencies to do a short voice test before starting an activity or when problems appear during real time operations.

D2 IMPORTANT DEFINITIONS

D2.1 KEYSET

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 Press-To-Talk (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.

D2.2 USER ROLES AND PERMISSIONS

Each user of a voice system has one or more assigned roles. Each role has a predefined number of loops that can be used, with the corresponding permissions (T/L/M).

A role can also have a predefined format or a look and feel of the touchscreen and the pages and loops available for this role.

Users can and should have the option to configure their assigned keysets with their particular voice levels for the microphone and speaker stored on it.

Each voice loop shows the permissions for its particular role (T/L/M). Monitoring is used in many console positions, especially in external positions (e.g., Payload operator) in which there is permission to listen to a voice loop but not to talk. For the positions that have talk/listen/monitor permissions, the monitor button is often used so as not to disturb the voice loop for talking in another loop simultaneously or to talk in a side conversation.

There is no difference between listening and monitoring, but the permission and the name on the button or key is different.

Normally a user or role has the option to listen to or monitor several loops but permission to talk in only one.

D2.3 VOICE FORMAT

Because in each MCC there are many more internal voice loops (up to 5000 loops) than available circuits to transport the voice loops between the MCCs (typically 24, 48, or 92 channels), different operations modes must be defined 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. An example of a voice format is 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.

For the ISS program, voice formats are defined in Operational Interface Procedures (OIPs) stored as internal NASA documents that are used for all the space agencies participating in the project.

The voice formats are exchanged between agencies in a predefined Excel table. The integration in each voice system can be different, but the matching of voice loops (which can have different names in different space agencies) must be accomplished before the start of operations.

To change a voice format or a voice loop, an Engineering Change Request (ECR) needs to be raised and the change discussed between the agencies that will be affected; for a given change, that could be two or more control centers.

D2.4 VOICE OPERATED EXCHANGE

Press to talk is mandatory for the use of keysets in human space flight. It can be implemented physically via a button or pedal, or via VOX, where the voice level is monitored and the PTT is activated automatically. VOX (voice operated exchange, voice level detection, or voice activity level) may be implemented on keysets. If VOX is implemented, the VOX threshold must be adjustable between −26 and −94 dB; the adjustment of the VOX threshold must be fixed for each user.

The VOX system should also have a hang system (VOX time out) from 1 to 3 seconds to remain engaged during brief speech pauses. It must be possible to save and restore the VOX threshold settings for each individual user in a console position.