TECHNICAL REPORT

ISO/TR 16056-2

First edition
2004-07-01

Health informatics — Interoperability of telehealth systems and networks —

Part 2:
Real-time systems

Informatique de santé — Interopérabilité des systèmes et des réseaux de télésanté —

Partie 2: Systèmes en temps réel

Reference number
ISO/TR 16056-2:2004(E)

ISO copyright office
Case postale 56 • CH-1211 Geneva 20
Tel. + 41 22 749 01 11
Fax + 41 22 749 09 47
E-mail copyright@iso.org
Web www.iso.org

Published in Switzerland

Contents

Page

List of Figures

List of Tables

Foreword

ISO (the International Organization for Standardization) is a worldwide federation of national standards bodies (ISO member bodies). The work of preparing International Standards is normally carried out through ISO technical committees. Each member body interested in a subject for which a technical committee has been established has the right to be represented on that committee. International organizations, governmental and non-governmental, in liaison with ISO, also take part in the work. ISO collaborates closely with the International Electrotechnical Commission (IEC) on all matters of electrotechnical standardization.

International Standards are drafted in accordance with the rules given in the ISO/IEC Directives, Part 2.

The main task of technical committees is to prepare International Standards. Draft International Standards adopted by the technical committees are circulated to the member bodies for voting. Publication as an International Standard requires approval by at least 75 % of the member bodies casting a vote.

In exceptional circumstances, when a technical committee has collected data of a different kind from that which is normally published as an International Standard ("state of the art", for example), it may decide by a simple majority vote of its participating members to publish a Technical Report. A Technical Report is entirely informative in nature and does not have to be reviewed until the data it provides are considered to be no longer valid or useful.

Attention is drawn to the possibility that some of the elements of this document may be the subject of patent rights. ISO shall not be held responsible for identifying any or all such patent rights.

ISO/TR 16056-2 was prepared by Technical Committee ISO/TC 215, *Health informatics*.

ISO/TR 16056 consists of the following parts, under the general title *Health informatics — Interoperability of telehealth systems and networks*:

— *Part 1: Introduction and definitions*

— *Part 2: Real-time systems*

Introduction

Delivery of health care services by means of telehealth is advancing rapidly. Telehealth enables providing these services with the use of information and telecommunications technologies. This includes a broad spectrum of capabilities including acquisition, storage, presentation, and management of patient information (represented in different digital forms such as video, audio, or data), and communication of this information between care facilities with the use of communications links.

Telehealth interactions may be carried out in three ways: **real-time**, **store-and-forward** or with the use **media streaming** methods. While real-time interactions imply that all parties directly participate in the telehealth session, store-and-forward interactions involve sending, reviewing, and returning an opinion over a period of time. Streaming is a method of delivery real-time or stored data such as audio, video, documents, still images, or other data type across networks with a reasonable amount of Quality of Services (QoS). With streaming, a receiving system can start displaying (or playing) the data before the entire content arrives.

Real-time telehealth sessions usually involve **synchronous** data transmission while store-and-forward can usually be regarded as **asynchronous**. Streaming uses time-synchronized streams of continuous media during transmission. However, data presentation uses buffering, if the receiving system receives data more quickly than required. If the data is not received quickly enough, the presentation of the data is interrupted.

Interoperability of telehealth systems and networks is critical in ensuring the telehealth technology serves well the care recipients and providers and meets their expectations. While this requirement is essential to the long-term sustainability of telehealth, interoperability is difficult to achieve. There are many reasons that make telehealth interoperability difficult, however, the following three need urgent addressing: (1) too broad definition of telehealth, (2) lack of standards specifically designed for telehealth, and (3) collaboration between the information technology and telecommunications industries.

There are multiple definitions of telehealth. The services provided by telehealth cover a broad spectrum of activities ranging from videoconferencing through exchange of health information to providing care services in emergency and complex clinical cases. From a technology perspective, the scope of these services is too broad and this makes it difficult to develop telehealth standards and products.

There is no ‘official’ telehealth standard. The telehealth industry uses high-level health care guidelines and technical standards developed for various technology sectors including multimedia conferencing, information technology, data communications, and security. These guidelines and standards focus on functional and operational requirements and do not address interoperability. To further complicate the problem, all of these standards as well as the telehealth needs and practices are rapidly changing.

Telehealth, more than any other recent development, bridges the boundaries between telecommunications and information technologies. The business goals and attitudes of these two industries are different. Telecommunications industry has a history of regulation, standardization, and control of the customer premises equipment. Interoperability and reliability have been the key factors to growth. The information technology industry (the desktop computing industry in particular) has achieved success through encouraging innovation, diversity, and tremendous cost-efficiency not always paying attention to interoperability aspects of the technology. The marriage of these two cultures and the integration of their respective technologies proved to be challenging.

To address the needs for interoperable telehealth systems and networks, telehealth services must be clearly defined in terms of their scope and interrelationships with other health-related services, a set of telehealth-specific standards must be developed, and subsequently implemented by the respective industries.

This two-part ISO Technical Report addresses interoperability issues in telehealth systems and networks. This document has been structured as follows:

Part 1: Introduction and definitions. Covers an introduction to telehealth and includes the definitions of telehealth, interoperability, and related terms.

Part 2: Real-time systems. Defines the scope of the technical standards related to real-time applications, (including video, audio, and data conferencing), identifies gaps and overlaps in the standards, defines requirements for interoperable telehealth systems and networks, and identifies building blocks for interoperable telehealth solutions.

This Technical Report is to be complemented by two other documents that will cover interoperability of store-and-forward and media streaming telehealth applications.

The target users of these documents are care providers and health care organizations, telehealth equipment vendors and implementers of telehealth solutions, professional organizations, and governments.

Health informatics — Interoperability of telehealth systems and networks — Part 2: Real-time systems

1 Scope

This Technical Report entitled *Interoperability of telehealth systems and networks – Part 2: Real-time systems* builds on the introduction to telehealth described in *ISO/IEC TR 16056-1 Health informatics - Interoperability of telehealth systems and networks - Part 1: Introduction and definitions* and focuses on the technical standards related to real-time applications, (including video, audio, and data conferencing) and interoperability aspects of telehealth systems and networks.

Specifically, this document addresses four main areas:

i) **Standards for real-time telehealth systems**. This Technical Report describes the technical standards related to real-time telehealth applications, including audio, video, and data conferencing capabilities. It also identifies gaps, overlaps and inconsistencies in the standards, and provides some guidance about how they need to evolve.

ii) **Interoperability issues in telehealth applications.** This Technical Report examines interoperability aspects of real-time multimedia conferencing standards and telehealth products, and identifies areas of concerns from the interoperability perspective that need to be resolved.

iii) **Requirements for interoperable telehealth systems and networks.** This Technical Report defines interoperability requirements at different levels of interaction between telehealth systems and provides some guidelines on how interoperability can be achieved.

iv) **Framework for interoperable architectures.** This Technical Report identifies interoperable building blocks for telehealth solutions and interactions between these building blocks, and explores the possibility of standardization of these building blocks.

The scope of the Technical Report does not include conformity and interoperability tests or functional specifications for telehealth systems and networks.

2 Normative References

This Technical Report incorporates by dated or undated reference, provisions from other publications. These normative references are cited in the appropriate places in the text, and the publications are listed hereafter.

For dated references, subsequent amendments and revisions of any of these publications apply to this ISO Technical Report only when incorporated in it by amendment and revision. For undated references, the latest edition of the referenced document (including any amendments) applies.

CEN/TC 251/N99-097 (1999)	*Health Informatics - Interoperability of Healthcare Multimedia Report Systems. Final draft CEN Report*
ISO/IEC 17000:2004	*Conformity assessment – Vocabulary and general principles*
ITU-T Recommendation G.711 (1988)	*Pulse code modulation (PCM) of voice frequencies*.
ITU-T Recommendation G.722 (1993)	*7 KHz audio - coding within 64 kbit/s*.
ITU-T Recommendation G.728 (1992)	*Coding of speech at 16 kbit/s using low-delay code excited linear prediction.*
ITU-T Recommendation H.221 (1993)	*Frame structure for a 64 to 1920 kbit/s channel in audiovisual teleservices.*

ITU-T Recommendation H.230 (1997)	*Frame-synchronous control and indication signals for audiovisual systems.*
ITU-T Recommendation H.242 (1996)	*System for establishing communication between audiovisual terminals using digital channels up to 2 Mbit/s*.
ITU-T Recommendation H.243 (1997)	*Procedures for establishing communication between three or more audiovisual terminals using digital channels up to 1920 kbit/s*.
ITU-T Recommendation H.224 (1994)	*A real time control protocol for simplex applications using the H.221 LSD/HSD/HLP channels*.
ITU-T Recommendation H.281 (1994)	*A far end camera control protocol for videoconferences using H.224.*
ITU-T Recommendation H.233 (1996)	*Confidentiality System for Audiovisual Services.*
ITU-T Recommendation H.234 (1996)	*Encryption key management and authentication system for audiovisual services.*
ITU-T Recommendation H.320 (1996)	*Narrow-band visual telephone systems and terminal equipment*.
ITU-T Recommendation T.120 (1996)	*Data protocols for multimedia conferencing.*
ITU-T Recommendation T.121 (1996)	*Generic application template.*
ITU-T Recommendation T.122 (1993)	*Multipoint communication service for audiographics and audiovisual conferencing service definition.*
ITU-T Recommendation T.123 (1994)	*Protocol stacks for audiographic and audiovisual teleconference applications.*
ITU-T Recommendation T.124 (1995)	*Generic conference control.*
ITU-T Recommendation T.125 (1994)	*Multipoint communication service protocol specification.*
ITU-T Recommendation T.126 (1995)	*Multipoint still image and annotation protocol.*
ITU-T Recommendation T.127 (1995)	*Multipoint binary file transfer protocol.*

3 TERMS AND DEFINITIONS

For the purposes of this Technical Report, the following definitions apply.

3.1
accreditation
third party attestation related to a conformity assessment body conveying formal demonstration of its competence to carry out specific conformity assessment tasks

3.2
A-law
variant of the G.711 audio encoding used primarily in North America and Japan

NOTE Related terms include μ-law and G.711

3.3
asynchronous transmission
transmission of individual bytes without time-dependency between the bytes

3.4
audiographics terminal
terminal that has audio and graphics capabilities, but no video capability

3.5
audiovisual terminal
terminal that has audio, video, and graphics capabilities

3.6
basic rate interface
BRI
ISDN service comprising two B (bearer) channels operating at 64 Kbps each and one D (data) channel operating at 16 Kbps

3.7
call
point-to-point multimedia communication between two H.32x endpoints

3.8
call setup
process of establishing a group of communication users and includes the initialization of any shared application and other resources which the user may require to be available

3.9
call signalling channel
reliable channel used to convey call setup messages following Q.931

3.10
call teardown
process of ending a call and freeing any resources reserved for that call

3.11
centralized multipoint conference
conference call in which all participating terminals communicate in a point-to-point fashion with an MCU

3.12
certification
third-party attestation related to products, processes, systems or persons

NOTE 1 Certification of a management system is sometimes also called registration.

NOTE 2 Certification is applicable to all objects of conformity assessment except for conformity assessment bodies themselves, to which accreditation is applicable.

3.13
channel service unit
CSU
interface used to connect a terminal or computer to a digital medium in the same way that a modem is used for connection to an analogue medium

3.14
charge coupled device
CCD
device used in cameras as an optical scanning mechanism.

NOTE It consists of a shift register that stores samples of analog signals. An analog charge is sequentially passed along the device by the action of stepping voltages and stored in potential wells formed under electrodes. The charge is moved from one well to another by the stepping voltages.

3.15
COder/DECoder
COmpression/DECompression
CODEC
hardware and/or software used with interactive video systems that converts an analog signal to digital, then compresses it so that lower bandwidth telecommunications lines can be used

NOTE The signal is decompressed and converted back to analog output by a compatible CODEC at the receiving end. The compression method (algorithm) may be proprietary or standards -based.

3.16
common intermediate format
CIF
ITU-T standard video picture scanning format where information is stored in luminance (brightness) and two color difference (chrominance) components

NOTE CIF represents 352 pixels/line by 288 lines/image for luminance and 176 pixels/line by 144 lines/image for chrominance. See also QCIF.

3.17
composite video
type of video signal in which all information -- the red, blue, and green signals, and sometimes audio signals as well, are mixed together

NOTE Composite video is used by NTSC-compliant devices (see NTSC Standard).

3.18
conformity assessment
demonstration that specified requirements relating to a product, process, system, person or body are fulfilled

NOTE Conformity to a set of specifications is a prerequisite to interoperability. However, conformity to the specifications alone does not guarantee interoperability of systems.

3.19
data service unit
DSU
device used in digital transmission for connecting a CSU to data terminal equipment (a terminal or computer), in the same way that a modem is used for connection to an analogue medium

Note See also CSU.

3.20
decentralized multipoint conference
conference in which the participating terminals multicast to all other participating terminals without an MCU

3.21
endpoint
terminal, gateway, or MCU

3.22
G.711
ITU-T recommendation for the digital representation of speech up to 3.4 KHz of frequency producing a 64 Kbps data stream

NOTE Commonly used in telephone networks. It comes in two variants: A-law and μ-law.

3.23
G.722
ITU-T recommendation for the digital representation of audio up to 7 KHz of frequency producing a 64 Kbps data stream with a much higher quality than G.711

3.24
G.728
ITU-T recommendation for the digital representation of audio producing a 16 Kbps data stream producing near-telephone quality audio.

3.25
gatekeeper
H.323 entity that provides address translation, control access, and sometimes bandwidth management to the LAN for H.323 terminals, gateways, and MCUs

3.26
gateway
H.323 entity, which provides real-time, two-way communications between H.323 terminals on the LAN and other ITU terminals on a WAN, or to another H.323 gateway

3.27
generic conference call
GCC
set of conference services described in the ITU-T T.124 Recommendation

3.28
H.221
ITU-T recommendation defining how to multiplex video and audio into frames using 64-1920 Kbps channels for switched and leased network services, excluding packetized networks

3.29
H.225D
ITU-T recommendation that specifies messages for call control including signaling, registration and admissions, and packetization/synchronization of media systems

3.30
H.230
ITU-T recommendation that specifies the frame-synchronous control and indication signals for audiovisual systems

3.31
H.231
ITU-T recommendation that specifies the multipoint control unit

3.32
H.235
ITU-T recommendation that defines the security framework used to provide authentication, encryption, and integrity for H.323 systems

3.33
H.242
ITU-T recommendation that specifies how to establish the communication between audiovisual terminals using digital channels with speeds up to 2 Mbps

3.34
H.243
ITU-T recommendation that specifies the establishment of communication between three or more audiovisual terminals using digital channels with speeds up to 2 Mbps

3.35
H.245
ITU-T recommendation that specifies messages for opening and closing channels for media streams, and other commends, requests and indications between two H.323 endpoints

3.36
H.261
ITU-T recommendation that specifies the video encoding and compression algorithm for two video resolutions: 352 x 288 CIF and 176 x 144 QCIF

NOTE H.261 is used in both H.320 and T.120.

3.37
H.263
ITU-T recommendation that specifies a new video codec for video over packet-switched networks or POTS

NOTE H.263 optimizes H.261 for very low bit rate of video coding below 64 Kbps. H.263 provides better motion compensation, more accurate motion vectors, optimized quantization for very low bit rates, and arithmetic coding.

3.38
H.310
family of ITU-T standards ratified in 1995 that describes the technical specifications for adapting narrow-band ISDN visual telephone terminals, as defined in H.320, to broadband ISDN (BISDN) and ATM environments

NOTE H.310 adds the MPEG-2 video-compression algorithm that provides MPEG-2 video quality.

3.39
H.320
family of ITU-T standards, ratified in 1990, that specifies how voice and video conferencing systems communicate over ISDN or leased networks, using a bandwidth from 64 Kbps to 1920 Kbps

3.40
H.323
family of ITU-T standards, ratified in 1996, that extends H.320 to computer networks, including LANs and the Internet

NOTE H.323 supports both point-to-point and multipoint operations. In addition, H.323 shares many components of the H.32x specification, such as the H.261 video codec, the G.711 audio codec, the H.263 video codec, G.722, G.723 and G.728. As a new feature, H.323 specifies a gatekeeper component that allows LAN administrators to manage video traffic for QoS. The H.323 specification also defines a LAN/H.320 gateway that permits a H.323 node to interoperate with H.320/H.324 terminals.

3.41
H.324
family of ITU-T standards, ratified in 1996, that allows video conferencing over standard analog phone lines with features similar to those in H.320

NOTE The H.324 standard uses H.263, which contains a better codec for POTS than H.261. H.263 is an improved version of H.261 that adds a 128 x 96 sub-QCIF (SQCIF) format. By using a 28.8 or 36.6 Kbps modem, H.263 may produce frame rates approaching those achieved by H.320 systems over ISDN.

3.42
interoperability
the ability of two or more systems (computers, communication devices, networks, software, and other information technology components) to interact with one another and exchange information according to a prescribed method in order to achieve predictable results

3.43
interoperability testing
an assessment of the ability of two or more systems to interact with one another and exchange usable electronic data

NOTE As conformity to the specifications alone does not guarantee interoperability of systems, interoperability testing is required to assess the ability of two or more systems to interact with one another and exchange usable electronic data. Interoperability testing does not include assessment of performance, robustness or reliability nor does it measure the conformity of an implementation. Two systems can be interoperable but still not compliant to the standard or specification.

3.44
µ-law
variant of the G.711 audio encoding used primarily in North America and Japan

NOTE See also G.711 and A-law.

3.45
multipoint control unit
MCU
endpoint on the LAN which enables three or more terminals and gateways to participate in a multipoint conference

NOTE The MCU includes a mandatory MC and optional MPs.

3.46
multipoint controller
MC
An entity that provides for the control of three or more terminals in a multipoint conference

3.47
multipoint processor
MP
entity that provides for the processing of audio, video, and/or data streams in a multipoint conference

NOTE The MP provides for the mixing, switching, or other processing of media streams under the control of the MC.

3.48
multipoint conference
conference between three or more terminals, which may be on the LAN or on the circuit switched network

3.49
NTSC Standard
standard for television broadcasting established by the National Television Standards Committee (NTSC)

NOTE Used in North America, Japan and some other countries. NTSC format: Lines / frame: 525; Frames per second (fps): 30; Interlace ratio: 2:1; Aspect ratio: 4:3; Color matrix equation: Y = 0.3*R + 0.59*G + 0.11*B; I = 0.6*R – 0.28*G - 0.32*B; Q = 0.21*R – 0.52*G + 0.31 * B; where R = red, G = green, and B = blue.

3.50
point-to-point protocol
protocol defined in RFC 1661, the Internet standard for transmitting network layer datagrams (e.g. IP packets) over serial point-to-point links

3.51
primary rate interface
PRI
ISDN service comprising 23 B (bearer) channels operating at 64 Kbps each and one D (data) channel operating at 16 Kbps

3.52
pulse code modulation
technique of used for the digital sampling of sound

NOTE The input waveform with a bandwidth up to 4.0 KHz is sampled at the recommended rate of 8,000 samples per second. Each sample is converted to one of 212 digital values and then compressed on either the A-law or the µ-law. This sampling scheme is adequate for voice communication.

3.53
quality of service
QoS
set of network technologies that enable a network to handle data traffic with a minimum amount of negative effects in a network environment used by many other users

NOTE Subscribers of QoS specify requirements in service-level agreements (SLAs) regarding throughput, packet loss, latency, and jitter.

3.54
quarter common intermediate format
QCIF
represents 176 pixels/line by 144 lines/image for luminance and 88 pixels/line by 72 lines/image for chrominance

NOTE See also CIF.

3.55
real-time streaming protocol
RTSP
application-level protocol that establishes and controls one or more time-synchronized streams of continuous media

NOTE RTSP has been designed to serve up multimedia from a cluster of hosts and acts as a network remote control for multimedia servers.

3.56
real-time transport protocol
RTP
data communication protocol capable of delivering real-time data such as live or interactive audio and video over IP packet-switched networks

NOTE RTP runs over UDP and uses its multiplexing and error checking features.

3.57
specified requirement
need or expectation that is stated

NOTE Specified requirements can be stated in normative documents such as regulations, standards and technical specifications. Specific requirements are intended to define some feature of a real implementation and offer the possibility of testing.

3.58
synchronized multimedia integration language
SMIL
enables simple authoring of interactive audiovisual presentations

NOTE SMIL is typically used for rich media/multimedia presentations, which integrate streaming audio and video with images, text or any other media type.

3.59
synchronous transmission
data communications in which transmissions are sent at a fixed rate, with the sending and receiving devices synchronized

3.60
T.120
family of ITU-T standards, ratified in 1996, that defines collaborative document sharing and whiteboard activities

NOTE The T.120 standards provide the audiographic portion of the H.320, H.323 and H.324 families. They also work independently as an audiographic conference for a low-bandwidth channel. The whiteboard capability provides document-sharing functions for multiple users so that they can simultaneously view and annotate a document with pens, highlighters and drawing tools. This specification also allows data-only T.120 sessions when no video communications are required or provided. In addition, T. 120 supports multipoint meetings in which the participants use different transmission media.

3.61
T.121
ITU-T standard that provides a generic application template (GAT), which specifies a common set of guidelines for building application protocols and the management facility that controls the resources used by the application

NOTE T.121 also describes how an application protocol, such as T.127 for file transfer, performs the following functions:

- Registers itself with the conference.
- Applies its capabilities locally and remotely.
- Interoperates and negotiates capabilities with other applications.

To ensure application consistency, T.121 is a required standard for products developed under T.120. The ITU also recommends that non-standard applications incorporate T.121 to provide product interoperability.

3.62
T.122
ITU-T standard that defines the multipoint services, which allow one or more participants to send data as part of a conference

NOTE These multipoint services are implemented by T.125, which provides the mechanism for transporting the data. Together, the T.122 and T.125 standards make up the T.120 multipoint communication services (MCS).

3.63
T.123
ITU-T standard that defines the transport and sequencing of data, and for controlling the flow of data across networks, including connect, disconnect, send, and receive functions

NOTE For data transport, T.123 defines a series of network interface profiles. Also, T.123 provides an error-correcting mechanism that ensures accurate and reliable data delivery. T.123 Annex B, an addition to the T.123 data conferencing standard, also defines the protocol for secure data conferencing.

3.64
T.124
ITU-T standard that provides the generic conference control (GCC) for initiating and administering multipoint data conferences.

NOTE The GCC performs the following functions:

- Serves as the information centre, directing users and data in and out of conferences and monitoring progress so that the latest conference information is always available,
- Maintains lists of conference participants and their applications; the GCC identifies compatible applications and features so that products can interoperate, and

- Tracks MCS resources so that conflicts do not occur when conference participants use multiple application protocols, such as T.127 for file transfer and T.128 for application sharing.

3.65
T.125
ITU-T standard that specifies how data is transmitted within a conference

NOTE T.125 defines the private and broadcast channels that transport the data, and ensures accurate and efficient communication among multiple users. T.125 also implements the multipoint services defined by T.122.

3.66
T.126
ITU-T standard that specifies how an application sends and receives whiteboard information, in either compressed or uncompressed form, for viewing and updating among multiple conference participants

NOTE The role of T.126 is to manage the multi-user workspace provided by the whiteboard.

3.67
T.127
ITU-T standard that defines how files are transferred simultaneously among conference participants

NOTE T.127 (also known as T.MBTF for Multipoint binary file transfer) enables one or more files to be selected and transmitted in compressed or uncompressed form to all or selected participants during a conference.

3.68
T.128
ITU-T standard that specifies the program sharing protocol, defining how participants in a T.120 conference can share local programs. Specifically, T.128 enables multiple conference participants to view and collaborate on shared programs.

3.69
T.134
protocol that provides point-to-point and multipoint distribution of text messages within the T.120 conference

NOTE It provides real-time or near-real-time text communications for those applications where audio communication is not available.

3.70
T.135
protocol that allows a user to reserve and control multipoint conference resources

NOTE It defines conferencing reservation protocols in a T.120 environment, typically between a client application and a scheduling system which reserves resources for multipoint control units (MCUs) or bridges.

3.71
T.136
protocol that specifies how Remote Device Control and configuration may be performed using T.120 as the transport protocol

3.72
T.140
protocol for multimedia application text conversation

NOTE The protocol for text chat within T.120, goes with T.134.

3.73
T.AVC
protocol that describes the control of audio and video capabilities present in a desktop or videoconference

NOTE This standard extends the capabilities offered by H.320.

3.74
T.RDC
recommendation that provides control of remote audio and video devices during a conference

NOTE A relatively new recommendation, T.RDC is an extension of H.281 for far-end camera control.

3.75
telehealth
use of telecommunication techniques for the purpose of providing telemedicine, medical education, and health education over a distance

NOTE See GATES 1994.

3.76
telemedicine
use of advanced telecommunication technologies to exchange health information and provide health care services across geographic, time, social and cultural barriers

NOTE See Reid 1996.

3.77
terminal
endpoint system, which provides for real-time, two-way communications with another terminal, gateway, or MCU

NOTE A terminal must provide audio and may also provide video and/or data.

3.78
testing of conformity
determination of whether one or more characteristics of an object of conformity assessment fulfils specified requirements, according to a procedure

NOTE "Testing" typically applies to materials, products or processes. The primary output of conformity testing is a test report, which includes the specified requirements, the actual results of testing, and the conformity status (i.e., whether or not the given product passed the test).

3.79
transport control protocol/internet protocol
TCP/IP
de facto standard ethernet protocols incorporated into 4.2BSD Unix

NOTE TCP/IP was developed by DARPA for internetworking and encompasses both network layer and transport layer protocols. While TCP and IP specify two protocols at specific protocol layers, TCP/IP is often used to refer to the entire DoD protocol suite based upon these, including telnet, FTP, UDP and RDP.

3.80
user datagram protocol
unreliable networking layer that sits at the same level of the networking stack as TCP

3.81
videoconferencing
electronic form of communications that permits people in different locations to engage in face-to-face audio and visual communication. Also, a collection of technologies that integrate video with audio, data, or both to convey in real-time over distance for meeting between dispersed sites

3.82
video streaming
method of delivery of multimedia data (e.g. audio, video, images, text, alphanumeric data, time-series, waveform data) across the networks with a reasonable amount of QoS

NOTE The receiving system presents (displays or plays) the data while the data is being transmitted in the background. Typically no storage of data occurs during streaming. The following protocols have been created by the IETF and W3C to achieve data streaming:

- RTP (Real-time Transport Protocol),
- RTSP (Real-time Streaming Protocol), and
- SMIL (Synchronized Multimedia Integration Language).

3.83
zone
collection of all terminals, gateways, and MCUs managed by a single gatekeeper

NOTE A zone must include at least one terminal and may include LAN segments connected using routers.

4 Abbreviations

ACR	American College of Radiologists
ADSL	Asynchronous Digital Subscriber Line
ANSI	American National Standards Institute
ATM	Asynchronous Transfer Mode
B-ISDN	Broadband ISDN (See H.310)
BRI	Basic Rate Interface
CCD	Charge Coupled Device
CIF	Common Intermediate Format
CMS	Control, Management and Signalling
CODEC	COder/DECoder (also COmpression/DECompression)
CSU	Channel Service Unit
DARPA	Defense Advanced Research Projects Agency (USA)
DOD	Department of Defense (USA)
DSU	Data Service Unit
GCC	Generic Conference Call
IETF	Internet Engineering Task Force
IP	Internet Protocol
ISDN	Integrated Services Digital Networks
ITU-T	International Telecommunications Union – Telecommunications
LAN	Local Area Network
MC	Multipoint Controller
MCU	Multipoint Controller Unit
MP	Multipoint Processor

NTSC	National Television Standards Committee
POTS	Plain Old Telephone System
PRI	Primary Rate Interface
QCIF	Quarter Common Intermediate Format
QoS	Quality of Service
RTCP	Real-Time Control Protocol
RTP	Real-time Transport Protocol
RTSP	Real-time Streaming Protocol
SCN	Switched Circuit Network
SW56	Switched 56 Network
TCP	Transmission Control Protocol
UDP	User Datagram Protocol
WAN	Wide Area Network

5 Multimedia Conferencing Standards

5.1 General

Real-time multimedia communication and conferencing technology provides for audiovisual communication as well as document sharing, including text, tables, and images. The video, audio, and data are encoded and compressed prior to being transmitted, and decompressed and decoded at the receiving end. Hardware or software codecs are used for performing encoding/compression and decoding/decompression functions.

In addition to the point-to-point conferencing, multipoint conferencing can be established using an MCU. An MCU can also broadcast a conference from one site to other sites.

Standards that have been established, include:

a. H.320 – for videoconferencing over circuit-switched networks such as ISDN, Switched 56, and leased lines.

b. H.321 – for multimedia conferencing in B-ISDN environments. B-ISDN depends on ATM switching, which offers the considerable advantage of guaranteed QoS. H.321 also involves H.310, which is the recommendation for broadband audiovisual communications systems and terminals.

c. H.322 – for using H.320 equipment on guaranteed quality of service LANs (Iso-Ethernet).

d. H.323 – a family of specifications for audiovisual communication over packet-switched networks, LANs and the Internet.

e. H.324 – a family of specifications for videoconferencing over POTS lines.

f. T.120 specifications for data and audiographic conferencing.

The H.3xx and T.120 recommendations are considered of prime importance for real-time telehealth applications.

The standards and corresponding communication networks are illustrated in Table 1.

Table 1 - Summary of the multimedia conferencing standards

	H.320	**H.321**	**H.310**	**H.322**	**H.323**	**H.324**
Network	Circuit-Switched Networks (ISDN)	B-ISDN (ATM)	B-ISDN (ATM)	Packet-Switched Networks with Guaranteed Bandwidth LANs (e.g. Iso-Ethernet)	Packet-Switched Networks with no Guaranteed Bandwidth (e.g. Ethernet, Internet)	Analog PSTN
Network Interface	I.400	AAL I.363 AJM I.361 PHY I.400	AAL I.363 AJM I.361 PHY I.400	I.400 or TCP/IP	TCP/IP	V.34 V.90
Video Compression	H.261	H.261	H.261* H.262* H.263 MPEG-2*	H.261	H.261* H.263	H.261* H.263*
Audio Compression	G.711* G.722 G.728	G.711* G.722 G.728	G.711* G.722 G.728 MPEG-1* MPEG-2	G.711* G.722 G.728	G.711* G.722 G.723 G.728 G.729	G.723.1* G.729
Data	T.120	T.120	T.120	T.120	T.120	T.120 T.434 T.84
Multiplex	H.221	H.221	H.221 H.222	H.221	H.225	H.223
Signaling	H.230 H.242	H.230 H.242	H.230 H.245	H.242	H.230 H.245	H.245
Call Setup	Q.931	Q.931	Q.2931	Q.931	Q.931	V.25
MCU	H.243	H.243	H.231 H.243	H.231 H.243	H.323	N/A
Camera Control	H.224 H.281	H.224 H.281	H.224 H.281	H.224 H.281	H.224 H.281	H.224 H.281
Security	H.233 H.234	H.233 H.234	H.233 H.234	H.233 H.234	H.233 H.234 H.235	H.233 H.234 H.235

* Mandatory feature

5.2 H.320 RECOMMENDATIONS

H.320 is an “umbrella” standard composed of specifications that have been defined for videoconferencing over circuit-switched services, such as ISDN or Switched 56. H.320 specifies the audio and video communications by recommending requirements for processing audio and video streams, formats for compatible audio and video inputs and outputs as well as the communication protocols and synchronization of the audio and video signals.

Three types of information (including video, audio, and data) can be combined and then transmitted to the remote site(s). For transmission, each of these data types are encoded in a unique way, combined together with the necessary control information, compressed, and then sent over to the network.

Audio is delayed on the assumption that the audio stream is related to the video stream and, since video compression takes longer than audio compression, a delay must be introduced into the audio process to maintain audio-video synchronization.

At the receiving end, the combined data is decompressed, decoded, and a delay is applied to the audio data stream before video, audio and data are presented.

H.233 and H.234 define the privacy of transmission between H.320 terminals. H.233 describes encryption requirements and refers to a number of possible algorithms (DES, FEAL, and BCRYPT) and H.234 describes the authentication and key management methods.

Figure 1 illustrates the interrelationship of the components of the H.320 Recommendations and shows the connections of an H.320 terminal with external devices.

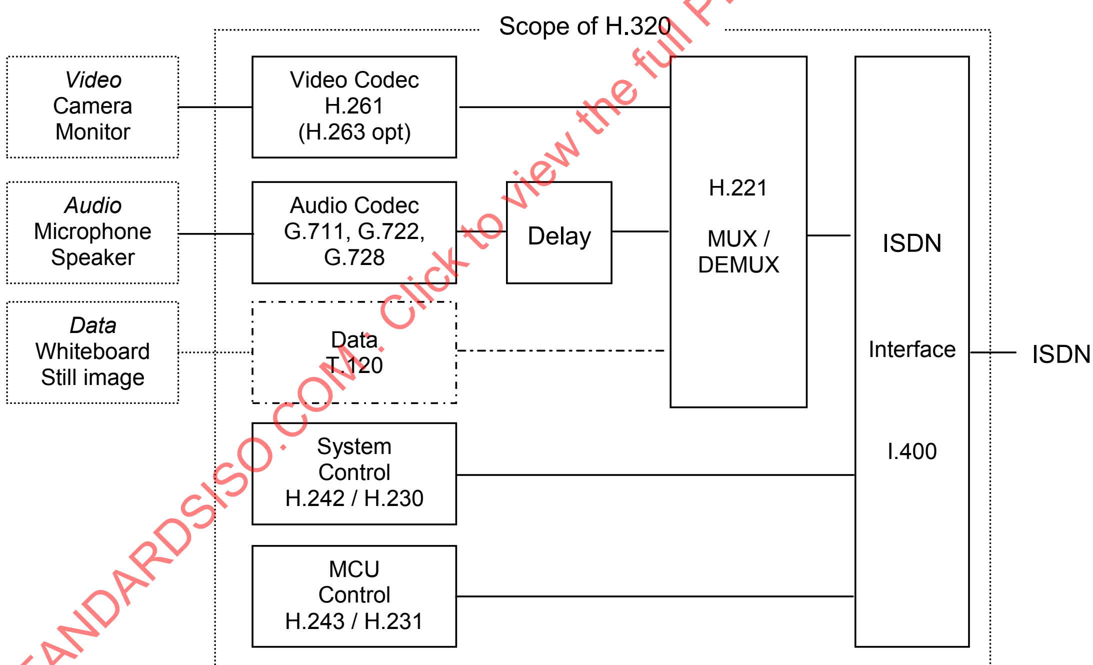

Figure 1 - The H.320 Terminal

The H.320 Recommendations alone do not define how data is to be encoded, thus proprietary data transmission can be implemented. This deficiency has been addressed in the T.120 standard, but T.120 is optional in H.320 and may or may not be included in telehealth equipment.

H.320 defines bandwidth usage over the range 64 Kbps to 1920 Kbps. Within this range, only certain specific bandwidths are allowed and these are restricted to particular multiples of the various ISDN channels (i.e., in multiples of 64 Kbps for 64 - 384 Kbps range, in multiples of 384 Kbps for 384 – 1,920 Kbps range).

The H.231 and H.243 Recommendations define multipoint connections between the H.320 terminals. While H.231 defines an MCU that serves as a bridge in multipoint (three or more) connections and provides audio mixing and video switching capabilities, H.243 defines the communication protocol between an H.320 terminal and an MCU. Two or more MCUs can be cascaded as illustrated in Figure 2.

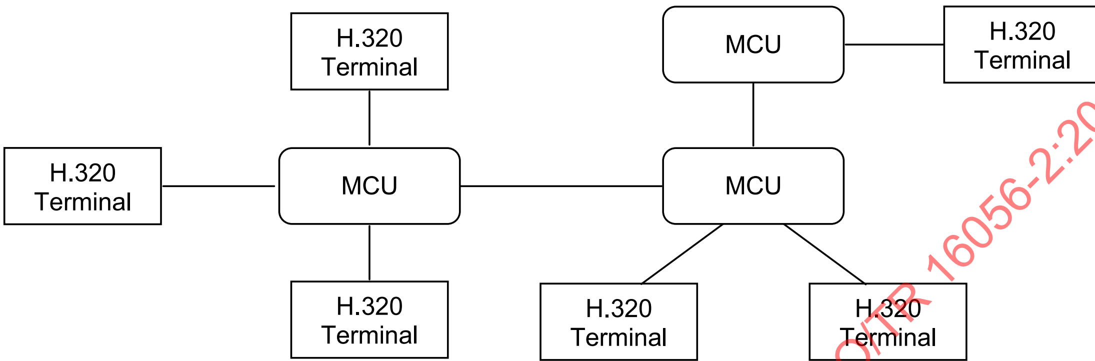

Figure 2 - A multipoint configuration

Table 2 presents a summary of the capabilities of the H.320 standard system.

Table 2 - Summary of the H.320 Recommendations

Feature	**Mandatory minimum**	**Optional maximum**
Resolution (pixels)	QCIF (176 x 144)	**CIF (352 x 288), 2CIF, 4CIF**
Frame rate (fps)	Up to 7.5	Up to 30
Transmission bit rate (Kbps)	56/64	2,048
Motion compensation - Encoder (pixels) - Decoder	 No Yes	 Yes (30 x 30 pixel search) Yes
Pre/post processing	No	Extensive
Audio coding	G.711	G.722, G.728
T.120 data conferencing	No	Extensive

The H.320 codecs operate at a constant bit rate (CBR) to accommodate the way the circuit-switched networks work. This is despite the inherently varying information content of a motion video sequence being depended on changing image complexity, degree of motion, and frequency of scene changes. The internally varying rate in these codecs is smoothed by buffering and dynamic control of codec parameters such as frame rate and quantizer step size to ensure that the buffer neither empties nor overflows. Such codecs operate at a fixed bit rate, but with variable quality.

5.3 H.321 AND H.310 RECOMMENDATIONS

The H.321 Recommendations describe the technical specifications for adapting narrowband H.320-compliant terminals to B-ISDN. While narrowband ISDN operates at rates equal or less than the primary rates (e.g. 1.544 Mbps), the B-ISDN operates at rates above the primary rates.

As described in I.121, the transfer mode is ATM in which the data is transmitted in a series of fixed-size blocks called cells.

ATM operates in a connection-oriented mode and provides QoS. At connection setup, the user of the ATM layer negotiates QoS traffic requirements. Examples of these requirements can be a bit rate or bounds on the delay variation or cell loss. The network allocates resources at the start of the call to meet the QoS requirements. The network may not accept the connection if the QoS of existing connections cannot be maintained. Throughout the connection duration, the network polices the traffic so that the QoS is delivered. Network resources are released at the end of the call.

Figure 3 depicts a generic H.321 terminal.

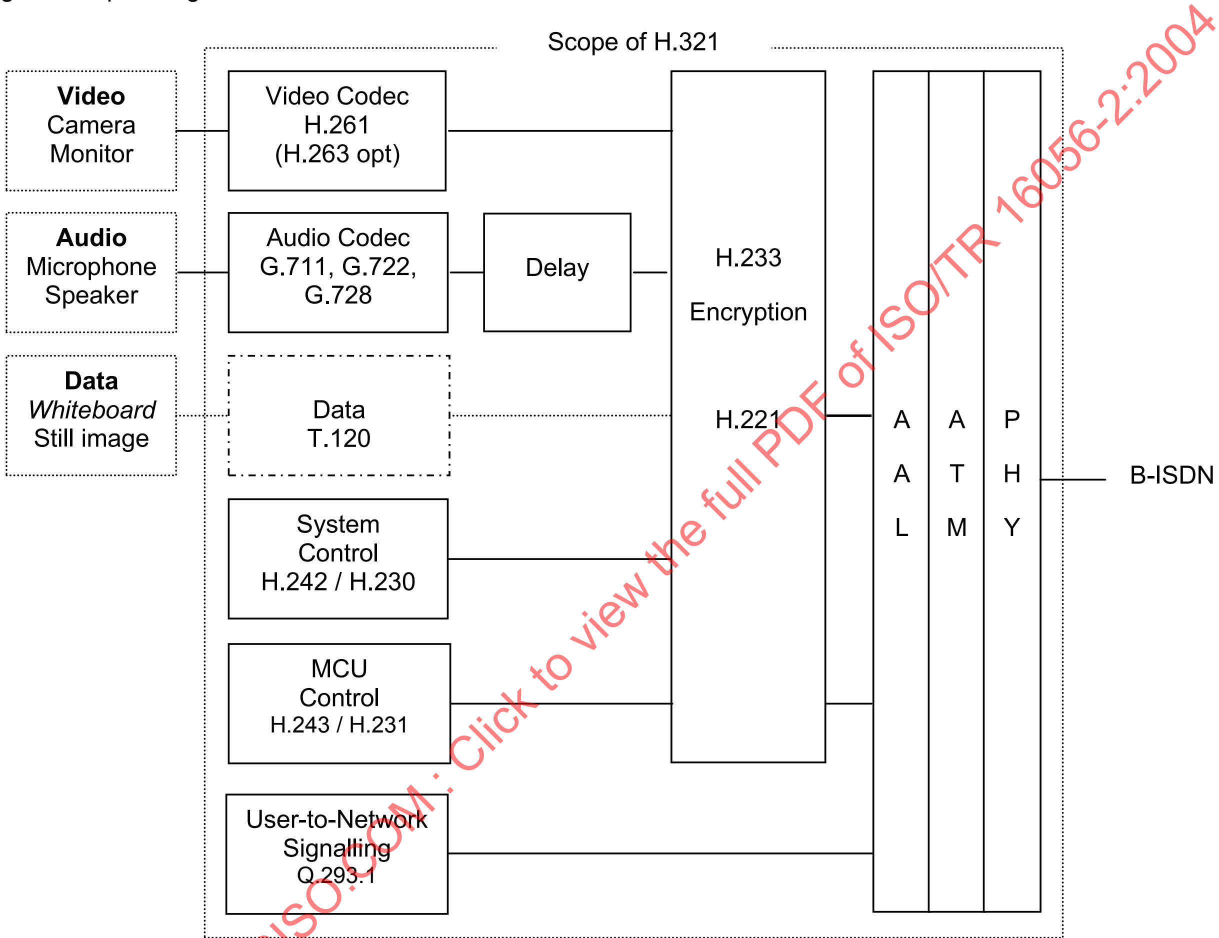

Figure 3 - The H.321 Terminal

The ATM adaptation layer (AAL) enhances the ATM layer to support the functions required to provide QoS for multimedia streams, including:

1. Control of transmission errors
2. Segmentation and reassembly of higher layer information into ATM cells
3. Handling of cell loss
4. Handling of cell delay jitter
5. Transfer of timing information

The H.321 terminal conforms to the requirements of H.320. The optional H.263 video codec, which was designed for the lower bit-rate terminals, is also a part of some H.321 terminals as it offers higher resolution video, such as 4CIF and 16 CIF, which are attractive for H.321 terminals running at the higher bandwidth.

To provide higher resolution terminals capable of handling H.262/MPEG-2 video and MPEG-1 audio, the H.310 Recommendations were developed. The interworking between H.310 and H.321/H.320 terminals is a mandatory requirement. It is achieved through a common set of H.320/H.321 functions defined in the H.310 Recommendations.

The ATM networks support variable bit rate (VBR) coded video allowing the transmitted bit rate to dynamically reflect the information content of the video signal. A VBR codec can therefore maintain a fixed-quality of transmission of multimedia data streams.

As indicated in Table 1, the H.310 Recommendations incorporate H.321 functionality. Therefore, there is no need for a gateway between H.310 and H.321 terminals on the B-ISDN.

5.4 H.323 Recommendations

In 1996, the ITU published the H.323 family of standards by incorporating multimedia communication and conferencing capabilities over packet-based networks, LANs and the Internet/Intranet. H.323 is a comprehensive standard that covers the selection of audio and video codecs, shared applications, call control, and system control, allowing for interoperability of H.323 and H.320 products.

The H.323 components include:

a. terminals to provide real-time communications and act as clients in a computer network.

b. gateways to provide services to H.323 clients so that they can communicate with non-H.323 systems. The services provided by a gateway include:

    i. translation between transmission formats, for example, between H.225 and H.221,

    ii. translation between communications procedures, for example, between H.245 and H.242, and

    iii. call setup and clearing for both the LAN and the telephone switched-circuit network.

c. gatekeepers to provide services to terminals in the network, in particular:

    i. bandwidth management to allocate, monitor, and control available bandwidth for registered endpoints,

    ii. name translation from symbolic names to IP addresses and vice versa, and

    iii. control services for registered H.323 endpoints.

d. MCUs to provide audio mixing and video switching capability between three or more terminals and support other control functions required in multipoint conferencing, such as H.245 negotiations between terminals to determine their capabilities for audio and video processing.

H.323 supports many types of multipoint conferences, including centralized, decentralized, and hybrid multipoint conferences. The use of the gateways, gatekeepers, and MCUs depends on the number and type of the terminal systems participating on the conference, the type of multipoint conference, and the type of the networks the terminals communicate with.

5.4.1 H.323 Terminals

The mandatory features of the H.323 Terminal are similar to those of the H.320 Terminal. In this manner, a minimum level of communications between H.323 and H.320 is provided without having to use transcoding gateways. Delay in the receive path is provided for synchronizing audio and video and jitter buffering.

Figure 4 illustrates the interrelationship of the components of the H.323 Recommendations and shows the connections of an H.323 Terminal with external devices.

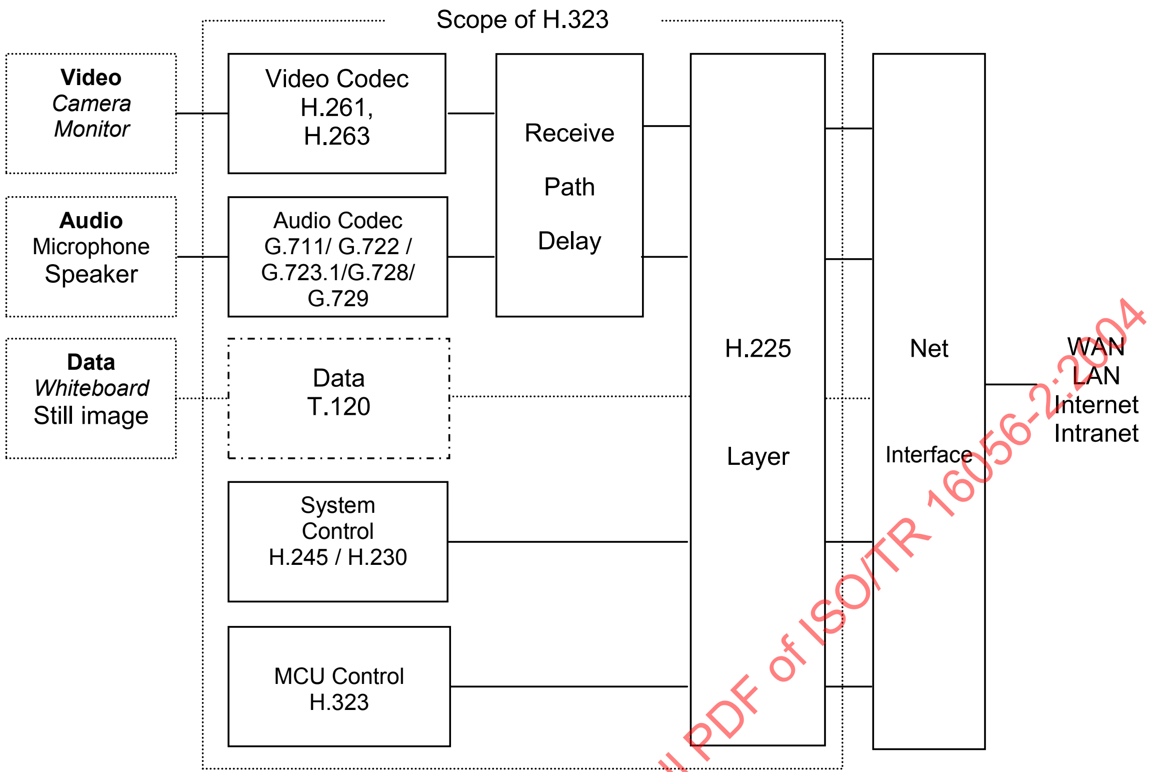

Figure 4 - The H.323 Terminal

Table 3 presents the H.323 and T.120 Recommendations in relation to the OSI Model.

Table 3 - The H.323 and T.120 protocol stack

OSI		Audio	Video	Terminal control and management			Data
Application	7	G.711, G.722, G.728, G.723, G.729	H.261, H.263	RTCP H. 225 RAS Channel	H.225 Call Sign	H.245 Control Channel	T.126/T.127 T.124/T.125 T.122
Presentation	6						
Session	5	RTP			X.224		T.123
Transport	4	UDP (Unreliable Transport)			TCP (Reliable Transport)		
Network	3	IP					
Data Link	2	Media Access Control (IEEE 802.3)					
Physical	1	Physical (IEEE 802.3)					

Since audio and video streams can tolerate some errors but cannot accept delays, they use the RTP running over the more efficient but unreliable transport of the UDP. Most terminal control and management functions and data conferencing use the reliable transport of the TCP with extended error checking and data delivery checking mechanisms.

Table 4 presents a summary of the capabilities of the H.323 standard system.

Table 4 - Summary of the H.323 Recommendations

Feature	Mandatory minimum	Optional maximum
Resolution (pixels)	QCIF (176 x 144)	CIF (352 x 288), 2xCIF, 4xCIF
Frame rate (fps)	Up to 7.5	Up to 30
Transmission bit rate (Kbps)	56/64	2,048
Motion compensation - Encoder (pixels) - Decoder	 No Yes	 Yes (30 x 30 pixel search) Yes
Pre/post processing	No	Extensive
Video coding	H.261, if provided	H.261, H.263
Audio coding	G.711 G.723.1 or G.729, if bandwidth is < 64 Kbps	G.722, G.728, G.723.1 G.729
T.120 data conferencing	No	Extensive

5.4.2 Gateways

The gateway, as per the H.246 Recommendations, provides all of the electrical and logical translation functions that are required to provide communications between an H.323 Terminal on the packet-switched networks and an H.320 Terminal on the circuit-switched network. Some gateways may provide interoperability with other terminal types, including H.324 and H.310.

The gateway operates like an H.323 Terminal or MCU on the packet-switched networks and an H.320 Terminal on the circuit-switched networks, with the protocol conversion functions done in the middle.

The implementations of the H.323 gateway vary significantly from a simple device capable of handling a single call between the packet-based and the circuit-based network to a very complex device capable of handing multiple calls, performing transcoding of the audio and video, interfacing to several different networks, and performing other functions such as the gatekeeper or the MCU. These differences contribute to the interoperability problems in the telehealth networks.

5.4.3 Gatekeeper

The gatekeeper provides network management functions for H.323 systems. It sets policies regarding the use of network resources and carries out these policies e.g. granting the H.323 Terminals permission for making a call using a specific amount network bandwidth.

Another function of the gatekeeper is to provide a directory service and address translation. It allows the users to enter alias addresses (e.g. facility names, telephone numbers), which are translated into network IP addresses for the H.323 Terminals to use. The gatekeeper also performs numerous housekeeping functions such as logging calls, keeping accounting information, reporting, and billing.

The gatekeeper is an optional component of the H.323 system. However, if it is included on the network, the H.323 Terminal must use it.

5.4.4 Multipoint control unit

The MCU is composed of two parts: the Multipoint Controller (MC) and the Multipoint Processor (MP). The MC is responsible for H.245 negotiations common operating modes between all terminals, and opening and closing of logical channels for the media stream distribution. The MP provides the media stream processing functions such as audio mixing and video switching, conversion between different codecs and bit rates, and multicasting.

H.323 supports centralized, decentralized, and hybrid multipoint conferences. Centralized conferences must use an MCU composed of an MC and audio, video, and data MPs. All terminals send audio, video, data and control streams to the MCU using point-to-point communication. The MC centrally manages the conference using H.245 control functions and the MCs perform the audio mixing, audio switching and data distribution, and send the processed data streams to the participants.

Decentralized multipoint conferences use multicasting instead of an MCU whereby the participating terminals multicast audio and video to other participating terminals without using the MCU. However, if data conferencing is involved in a multipoint conference, then the MC is used to provide the H.245 functionality and MP is used to centrally process the shared data. Hybrid multipoint conferences use a combination of centralized and decentralized features. Some terminals may operate in a centralized mode others in a decentralized mode, and the MCU is used to bridge the two modes.

The implementations of the MCU vary significantly from a simple system providing multipoint conferencing capabilities for homogeneous terminals to a sophisticated system capable of supporting different types of conferences with heterogeneous terminals and networks participating in these conferences and providing a comprehensive set of presentation functions (e.g. continues presence) and conference management capabilities (e.g. voice activation, chair control). These differences contribute significantly to the interoperability problems in the telehealth networks.

5.5 H.324 Recommendations

The H.324 Recommendation addresses and specifies a common method for sharing video, voice, and data simultaneously using a modem connection over a single analog (POTS) telephone line. It also specifies interoperability under these conditions, so that videophones, for example, based on H.324 will be able to connect and conduct a multimedia session. Support for each media type is optional, but if supported, the ability of use a specified common mode is required so that all terminals supporting that media type can interoperate.

Figure 5 illustrates the interrelationship of the components of the H.323 Recommendations and shows the connections of an H.323 terminal with external devices.

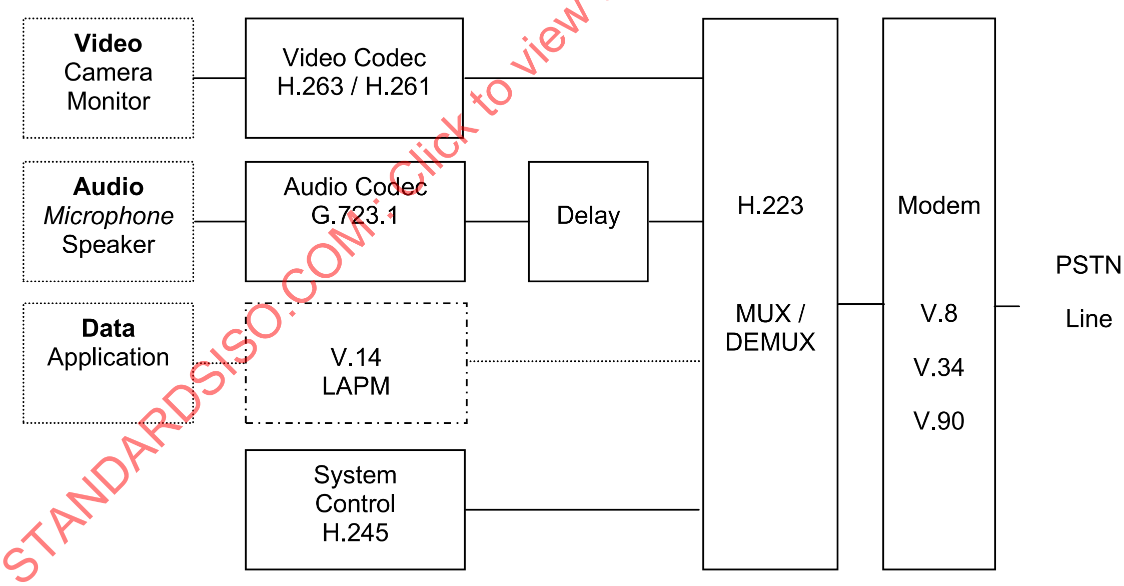

Figure 5 - The H.324 Terminal

H.324 terminals are typically implemented in stand-alone devices such as video telephones. They also can be integrated into personal computers.

5.6 T.120 RECOMMENDATIONS

5.6.1 T120 OVERVIEW

The T.120 Recommendations contain a series of communication and application protocols and services that provide support for real-time, multipoint data communications. These multipoint facilities are important building blocks for a whole range of collaborative applications, including desktop data conferencing, and multi-user applications.

A high-level overview of the T.120 standard is presented in Figure 6.

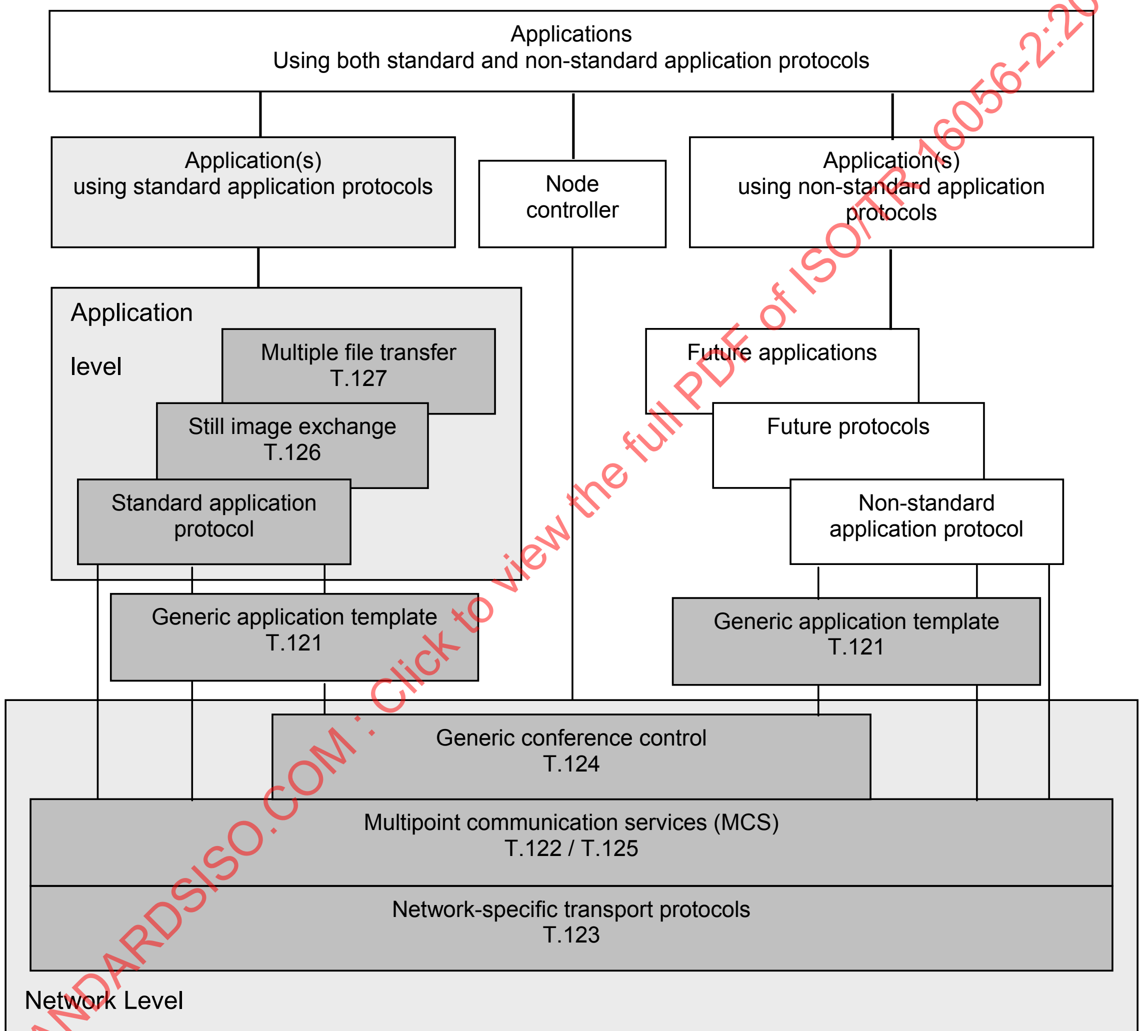

Figure 6 - The T.120 Recommendations

5.6.2 T.120 LAYER COMPOSITION

T.120 is composed of several layers:

a. **Transport Protocol** layer supports the network-specific transport stack for each supported network such as ISDN, POTS, LAN, and the Internet.

b. **Multipoint Communication Services** (MSC, T.122 / T.125) provides multipoint connection-oriented transport connections.

c. **Generic Conference Control** (GCC, T.124) provides services for establishing and terminating conferences, managing conferences, handling application registries and the overall conduct of the conference. GCC also provides coordination between data sharing and the audio/video in multipoint conferences.

d. **The Generic Application Template** provides a guide for the development of the T.120 application protocols for the top sub-layers. It is mandatory for standardized applications and highly recommended for non-standard application protocols.

e. The protocol sub-layer may include:

   i. **T.126** multipoint still image and annotation protocol for collaborative data sharing,

   ii. **T.127** multipoint binary file transfer protocol for transmitting files in a multipoint conference, or

   iii. **T.Share** application-sharing protocol for permitting participants in a T.120 conference to share applications.

The node controller depicted in Figure 6 is the command and control entity for T.120 and is responsible for administering network-level events, including the management of conference connections, participants, and conference data. This controller takes command of the other T.120 layers, particularly the transport layer, and uses the GCC, MCS, and other protocol services to manage the entire conference. The node controller acts as the translator, ensuring that events are interpreted and ordered correctly. The node controller itself is outside the scope of the T.120 Recommendations.

5.6.3 T.120 Interoperability

One of the most important features of the T.120 infrastructure is the interoperability of products and services that support the standard. T.120 systems can accommodate multiple independent applications concurrently using a multipoint network environment. The endpoint systems can be in circuit-switched networks, packet-switched networks or LANs using a variety of common topologies, including cascaded, star, and daisy-chain connections.

The interoperability of T.120 products is measured on two levels: networking and applications. The T.122, T.123, T.124, and T.125 standards make up the networking level of T.120. Products and services that meet these standards have the necessary infrastructure to do the following:

a) establish and maintain conferences without any platform dependence,

b) manage multiple participants and programs, and

c) send and receive data accurately and securely over a variety of supported networking connections.

The T.126 and T.127 standards make up the applications level of T.120. These standards ensure that the electronic whiteboard and file transfer applications developed under T.120 can interoperate across platforms and networks, and within multipoint conferences.

T.120 addresses lack of standardization for data services in H.320 / H.323. Using the T.120 Recommendations it is possible to design data conferencing services that will interoperate between different suppliers' equipment over heterogeneous networks (circuit-switched, packet/cell-switched, and LANs).

T.120-based implementations, however, still fail to deliver high-quality audio and video conferencing services in conjunction with data conferencing services. The separation of audio, video, and data into distinct services, with audio and video being treated in a very different way, is not justifiable from the telehealth applications viewpoint.

6 Telehealth Applications

6.1 General

Telehealth systems, as defined in this report, are capable of handling different data types including:

a. **Audio:** Usually represented as a one-dimensional array. Examples include speech, sound of heartbeat at different stress levels;

b. **Video:** Usually represented as a series of two-dimensional arrays. Examples include video captured with a video camera or video signals captured while performing ultrasound of heart, endoscopy of gastric ulcer, or coronary angiography;

c. **Numeric Values:** Integer or floating-point numbers obtained by direct measurement. Examples include systolic and diastolic pressure, blood volume, or calculations such as average weight;

d. **Vectors (waveform):** Represented as one-dimensional arrays of values obtained by direct measurement, such as electrocardiograms, or by calculation, such as a cardiac left ventricular volume curve, a histogram;

e. **Images:** Two-dimensional array of integers, sometimes floating-point numbers. Examples include digitized radiographs and MRI images;

f. **Time Series:** A series of multidimensional arrays synchronized in time. Examples include angiography cine and ultrasound cine;

g. **Graphics:** A set of co-ordinates of the vertices created as part of image processing ('region of interest') or a sketch illustrating the anatomical site of some relevant organ;

h. **Text:** Free text composed of strings, integer or floating-point values. Examples include patient demographic, labels, and annotations; and

i. **Documents:** A structured text suitable for computer processing. Examples include physicians' notes, and interpretation reports. A document containing a structured text may be tagged to convey information about its order, structure, pattern, and format. Examples include HTML or XML documents.

Telehealth systems may exchange any combination of the above data types as an asynchronous or synchronous traffic.

Telehealth systems are designed to provide various health services over a distance. Those telehealth systems are identified with the 'tele-' prefix followed by the health-related discipline, for example teleeducation, telepsychiatry, teleradiology, and telecardiology. Tele comes from an ancient Greek word, meaning 'distant'.

The telehealth systems and their capabilities of handling different data types, along with some information indicating the source of the data types, their approximate size, and the encoding standards are given below.

6.2 TELELEARNING

A telelearning system facilitates the provision of education and training services to health care professionals or patients. It is typically a room-based videoconferencing system with some additional attachments such as a scanner, VCR, a document camera or a computer.

Table 5 presents functions performed by a telelearning system, along with some information regarding data handling, size of the data type or bandwidth required and standards used to enclose the data.

Table 5 - Summary of functionality of a telelearning system

Function	Data type / source	Size (MB) or bandwidth (Kbps)	Encoding standards
Video acquisition, encoding and compression	Video / Video camera Video / Document camera	64-2000 Kbps [(a)]	Video standards H.26x
Pre-recorded video encoding and compression	Video / VHS	64-2000 Kbps [(a)]	Video standards H.26x

Real-time oral communication	Audio / Microphone	2.4-1410 Kbps [(a)]	Audio Standards G.7xx
Pre-recorded oral communication	Audio / VHS	2.4-1410 Kbps [(a)]	Audio standards G.7xx
Still image acquisition, encoding, and transmission Binary file transfer Application sharing Text chat	Binary file / Miscellaneous (e.g. video camera, document camera, digital camera, scanner)	64-2000 Kbps [(a)]	Multipoint data communications standards T.120
System Control	Control command / Codec	3.5-350 Kbps	Signalling standards H.230 H.242/H.245
Multipoint Conferencing	Control command / Codec		Session Standards
Security	Security algorithm-specific / Codec or a specialized security device		Encryption Standards H.233/H.234
[a] Depending on compression ration and signal quality			

6.3 TELECONSULTATION

A teleconsultation system facilitates the provision of specialty services (e.g. mental health, cardiology, or dermatology) by health care specialists to patients and/or primary care physicians. The system typically provides real-time media conferencing capabilities along with the ability to handle medical specialty-specific data streams (e.g. heart and lung sounds, ECGs, video and still images captured with the use of specialized scopes).

Table 6 presents functions performed by a teleconsultation system along with some information regarding data handling, size of the data type or bandwidth required and standards used to enclose the data.

Table 6 - Summary of functionality of a teleconsultation system

Function	**Data type / source**	**Size (MB) or bandwidth (Kbps)**	**Encoding standards**
Video encoding and compression	Video / Video camera Video / Document camera	64-2000 Kbps [(a)]	Video standards H.26x
Pre-recorded video encoding and compression	Video / VHS	64-2000 Kbps [(a)]	Video standards H.26x
Real-time oral communication	Audio / Microphone	2.4-1410 Kbps [(a)]	Audio standards G.7xx

Pre-recorded oral communication	Audio / VHS	2.4-1410 Kbps [(a)]	Audio standards G.7xx
Still image Binary file transfer Application sharing Text chat	Binary file / Miscellaneous (e.g. specialized scopes video camera, digital camera)	64-2000 Kbps [(a)]	Multipoint data communications standards T.120
Acquisition and transmission of heart and lung sounds	Analog sound signals / acoustic stethoscope or Digital signals / a digital stethoscope	Varies	Audio standards G.7xx or data communications standards T.120
Acquisition and transmission ECG	Waveform signals / ECG system	Varies	Proprietary
System Control	Control command / Codec	3.5-350 Kbps	Signalling standards H.230 H.242/H.245
Multipoint Conferencing	Control command / Codec		Session standards
Security	Security algorithm-specific / Codec or a specialized security device		Encryption standards H.233/H.234
[a] Depending on compression ration and signal quality			

7 INTEROPERABILITY ISSUES

7.1 GENERAL

This section examines the interoperability aspect of the real-time multimedia conferencing standards and telehealth equipment, describes the areas of concern from the interoperability perspective, and identifies the interoperability issues.

The relationship between parties that contribute to interoperability issues is depicted in Figure 7.

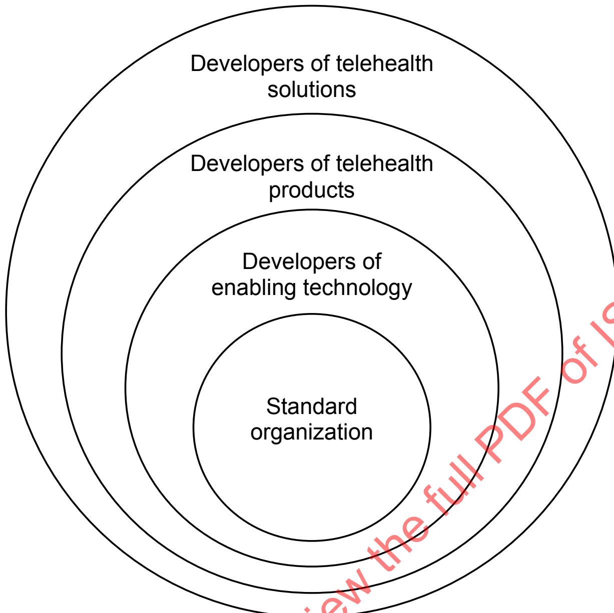

Figure 7 - Parties contributing to telehealth interoperability

The T.3xx and T.120 Recommendations developed by ITU-T are converted into a series of enabling technologies. Developers use these technologies to develop multimedia conferencing products. Since the standards only provide a development framework and do not define specific implementations, developers use the enabling technologies to develop H.3xx/T.120-compliant products with unique value-added features. Network developers, system integrators, and service providers select from an array of these products to implement a complete telehealth network capable of providing multimedia communication services.

Since each of these parties have different perspective on interoperability, they contribute to the list of interoperability issues in the telehealth systems and networks.

This document considers three sources of interoperability issues in telehealth networks:

1. The standards or recommendations developed by ITU-T and other standard setting organizations.
2. The enabling technologies and telehealth products.
3. The telehealth solutions that are the actual implementations of telehealth products interlinked via communication networks.

7.2 STANDARD-RELATED SOURCES OF INTEROPERABILITY ISSUES

The H.3xx and T.120 Recommendations, when completed and fully implemented will assist in achieving interoperability between different suppliers' telehealth equipment over the digital networks. However, there are still deficiencies in the standard specifications and/or their implementation by the manufacturers of the telehealth products. The following are some examples of such problems.

7.2.1 Loose adaptation of previous protocols

One of the requirements for new specifications is backward compatibility with the older and more established standards, for example, H.323 Recommendations must be backward compatible with the H.320 Recommendations. However, the new specifications must take into account unique requirements of the new domain, e.g. H.323 must include the requirements of the packet-based networks. Quite frequently the ‘building blocks’ of the older specifications are being modified to include the features of the new environment in a loosely manner.

Example - Q.931

This call signaling protocol was borrowed from ISDN and subsequently modified for use in H.323 products. When used in H.323, the layout of the original Q.931 protocol was kept but the meanings of many fields were redefined for H.323. In addition, many fields in the original Q.931 remain undefined for H.323. Nevertheless, there are attempts to use these fields in the exact way they are used in ISDN. Developers working on H.323 deliberately interpret and adjust Q.931 because packet-based networks are inherently connectionless and rely on different architectures using completely different components and procedures. When products use an original Q.931 implementation and run it as a part of H.323 implementation, they fail to comply with H.323 in the strict sense of the standard.

7.2.2 Loosely defined encoding and decoding methods

Part of the current interoperability problems can trace their origins to how the H.323 standard was defined. H.323 incorporates H.225, RTP and RTCP. Each of these protocols needs to be encoded according to a consistent formula.

Example - H.225 vs. RTP/RTCP

H.225 is based on ASN.1 for encoding and decoding. Yet, RTP and RTCP, which come from the IETF, are based on TLV (Type Length Value) formula. This leads to interoperability problems among multimedia conferencing products that use these different encoding and decoding formulas.

7.2.3 Inconsistent definition of mandatory/optional requirements

While the specifications for the optional components are complete, the definition of how and when to use of them is much less precise. This leads to the implementation of different subsets of the H.32x/T.120 Recommendations and consequently to problems with interoperability of the telehealth systems.

Example 1 - RTP/RTCP

These protocols were approved by the IETF prior to the development of H.323. RTP/RTCP is a mature protocol with many details and fields for mandatory and optional procedures (especially RTCP), which are partially overlapping, with other protocols defined within the H.323 Recommendations. Some developers conform to the RTCP guidelines, while others follow H.323 so they may choose to omit certain fields when interpreting the standard. Therefore, one implementation may differ and fail to correspond to other implementations leading to interoperability problems. For example, in RTCP, the "cname" field assigns a unique name to a media stream and it is mandatory. This field is redundant in the H.323 Recommendations. In H.323, stream identification and association is done by means of H.225 (using H.245). Some H.323-compliant products do not use the "cname" field. Others products that follow the original RTP/RTCP expect this field in the message. This prevents systems from interoperating.

Example 2 - Selection of audio codecs

H.323 dictates that if the available bit rate is more than 64 Kbps, the master (head) codec is G.711. On the other hand, if the available bit rate is less than 64 Kbps, the rules are different. Specifically, if it is a voice-only call, the audio codec must be G.729, and if the call is voice and video, the codec is G.723 (G.723 compresses voice more and uses less bandwidth than G.729 or G.711, leaving more room for the video portion of the call).

In order to select the right audio codec, the two endpoints need to be aware of the bandwidth available, yet the ability to detect the data rate is not explicitly a requirement of an H.323-compliant terminal.

7.2.4 Gaps in the recommendations

Completeness of the standards is an issue that is well recognized, yet, subsequent releases of the H.323/T.120 Recommendations have gaps in the definition of important features.

Example 1 - Security
H.323 v.1, rectified in 1996, did not fully address security. Early implementers explored use of security features defined at the IP layer. Subsequently, the new H.235 standard was developed as part of H.323 v.2 of the standard. The security aspects of the products that are compliant with H.323 v.1 are likely not interoperable with the security features of the products compliant with H.323 v.2.

Example 2 - Multiplexing
There is no definition currently of interoperable media stream multiplexing in the H.323 Recommendations or in the RTP/RTCP. Although some proposals have been discussed both in VoIP forum and in IETF, there is no clear direction to follow at this time. Until a standard technique is agreed upon, this is a source of interoperability issues in the future.

7.2.5 No specifications for data encoding in the H.320/H.323 Recommendations

The audio and video streams are continuous and synchronized. Data, on the other hand, are regarded as only occasionally required, therefore considered optional.

Example - H.32x and T.120
The H.320/H.323 Recommendations do not provide specifications for data encoding, thus proprietary data transmission is frequently implemented. Data encoding specifications are included in the T.120 Recommendation but T.120 is not part of H.320/H.323 and may or may not be implemented in telehealth systems.

7.2.6 Evolving T.120 Recommendations

The T.120 Recommendations addresses the deficiencies of H.320 and H.323, namely the lack of standardization of data conferencing services, and the lack of reliable and generic data transfer protocol independent of the underlying network infrastructure and data being transferred. The intention of the T.120 recommendation is neither to specify an Application Programming Interface for application development, nor to define specific application-level functions. The focus is on standardization of the data protocol level, ie the information transmitted and received across the network.

Using the T.120 recommendation carefully, it is possible to design audio, video and data conferencing services that will interoperate between different suppliers' equipment over circuit-switched networks (e.g. ISDN) or packet-switched networks (e.g. frame relay). However, there are still several areas where the interoperability aspect of T.120-compliant standards (and consequent T.120-compliant products) has to be examined carefully. This includes specifications for:

- T.Sound – transfer of audio objects,
- T.MPTV – transfer of MPEG encoded visual objects,
- T.PRIV – application communication encryption services, and
- T.RDC – Remote Device Control. Mechanisms are provided for controlling far end cameras and audio devices (e.g. audio volume). Specifications for control and management of other devices (e.g. VCRs and scanners), are still being developed. Also, the management and control of network devices such as MCUs, gateways, and switches are under consideration.

7.2.7 Loosely defined specifications for an MCU

Most MCUs provide support for video, audio, and T.120 data. Some of them, however, support only a subset of audio codecs (e.g. G.711 and G.722), providing transcoding as a substitute. This approach may introduce some latency or loss of quality. Some of them may mix audio data in half-duplex only.

7.2.8 Evolving/undefined specifications for Gateways and Gatekeepers

The H.323 specifications for gateways and gatekeepers are still evolving, therefore, their implementations differ significantly. To make the matter worse, some gateway/gatekeeper services specified in the H.323 standards are mandatory, some of them are optional, and some are mandatory only when a gateway or gatekeeper are present in an H.323 network. This leads to the development of proprietary products and significant differences in functionality of products manufactured by different vendors.

Example 1 - Varied capabilities
A gateway can be very simple and handle only a single call between the packet-based and the circuit-switched networks. It can also be very complex, handling many simultaneous calls, performing transcoding of the video and audio streams, and interfacing to several different circuit-switched networks. It may also include other functions such as the gatekeeper or the MCU.

Example 2 - Undefined gatekeeper-to-gatekeeper communications
Address resolution is defined in H.323 for specific topologies. In H.323 v.2 the concept of aliases has been expanded to support different addressing schemes. The H.323 v.1 list was expanded with new H.323 v.2 formats. However, there is no clearly defined way for gatekeepers to deal with both of these schemes, nor does the standard indicate if these are to be combined when communicating with another gatekeeper. This will lead to interoperability problems.

7.2.9 LOOSELY DEFINED BANDWIDTH MANAGEMENT

Both H.320 and H.323 standards support the bandwidth-sharing feature, ie available bandwidth is divided up amongst video, audio, and data. However, the allocation of bandwidth and bandwidth management differ between the H.320 and H.323 products. For the H.320 products, only certain specific bandwidths are allowed and these are restricted to particular multiples of the ISDN channel types. The H.323 products use a gatekeeper that has been defined to provide the bandwidth management functions. The gatekeeper can limit the number of users or the amount of bandwidth available to concurrent H.323 connections within a network to ensure that other traffic can be served. While the H.323 standards define how communication with a gatekeeper occurs, they do not specify how a gatekeeper should provide its service.

7.2.10 POOR STANDARDIZATION OF MEDIA CONFERENCING

The T.120 Recommendations for multipoint communication services (MCS), generic conference call (GCC), and generic application template (GAT) are poorly standardized, thus many implementation details are left for interpretation.

Example - Definition of MCS over a variety of communication stacks
It focuses on transmission of data only and ignores the transmission of audio and video. The separation of video, audio, and data into distinct services, with audio and video not being a part of the specifications, is not acceptable from an application point of view. Also, the rigid call management model of GCC may well reflect an important conferencing style, but is likely to be restrictive for telehealth applications.

7.3 PRODUCT-RELATED SOURCES OF INTEROPERABILITY ISSUES

Telehealth vendors strive for the development of products that are compliant with respective standards. However, there are numerous factors that contribute to interoperability problems. The following are some examples of such problems.

7.3.1 TIME DELAY BETWEEN THE STANDARDS (EVER-EVOLVING) AND PRODUCTS

Backward compatibility is required by the ITU-T. While each new version of recommendations represents a significant improvement in terms of features and functionality, it also seeks to "fix broken parts" of the preceding version. Some telehealth vendors who implemented the advanced features included in the older version of the standards have difficulties with a smooth migration to the new version of the standards.

Example - Association among the H.323 protocols
The issue of tying (or associating) the H.323 messages and protocols in the context of the same call is handled differently between H.323 v.1 and H.323 v.2. In H.323 v.1, implementers could choose one of two options when making the association between different messages in various protocols. They could use either:

a) ConferenceID and CRV original Q.931 fields, or

b) Source and destination call signalling Address, and AnswerCall fields that are defined in RAS and Q.931.

In H.323 v.2, the association of different protocol messages is defined much more precisely using CallIdentifier. The CallIdentifier field was added to all of the H.323 protocols, and as a result, every message that belongs to the same call is identifiable and evident. Sets of messages are easily "associated" with their partners. Another benefit of changing to CallIdentifier is that the new technique covers all the topologies, thus simplifying life for developers and implementers going forward.

While this approach smilingly simplifies the implementation of H.323-compliant products, in order to be backward compatible with H.323 v.1 protocol, all subsequent product implementations will need to support all 3 of the call association methods defined to date.

7.3.2 IMPLEMENTATION OF A DIFFERENT SUBSET OF THE STANDARDS

Most telehealth vendors implement only a subset of the standard recommendations for numerous reasons (e.g. due to cost, complexity, or conformity with the previous version of their products). While these products are compliant with the respective standards (or a subset of those standards), their interoperability in a multi-vendor environment is questionable.

Example - Implementation shortcuts
Some H.323 products may not support some of the bandwidth-scalable features of H.323, such as H.263 video codec and G.723.1 audio codec. In addition, within H.263 there are a few different “flavors” of H.263, namely: H.263 (draft), H.263 (final) and H.263+. Having implemented only the mandatory ISDN-based H.261 video codec and the G.711 audio codec or a different version of the H.263 codec leads to interoperability problems.

7.3.3 PROPRIETARY SOLUTIONS

Telehealth vendors claim conformity of their systems with H.3xx standards but additionally support proprietary codecs, particularly audio codecs, to achieve higher quality conferencing. These systems work well only when connected to systems that support the same proprietary solution. When working in a multi-vendor environment, they rely on a parameter negotiation task to establish a set of commonly used parameters/algorithms between the systems. If that task fails, the systems typically fail to establish a successful connection or deliver poor quality videoconferencing services.

7.4 IMPLEMENTATION-RELATED SOURCES OF INTEROPERABILITY ISSUES

Network developers, system integrators, and service providers involved in the implementation of telehealth solution also contribute to interoperability issues. The following are some examples of such problems.

7.4.1 UNDEFINED SPECIFICATIONS FOR GATEKEEPER COMMUNICATION

As mentioned above, the gatekeeper-to-gatekeeper communications is undefined in H.323 v.2.

Example - Address resolution
Address resolution is defined in H.323 for specific topologies. In H.323 v.2 the concept of aliases has been expanded to support different addressing schemes. The H.323 v.1 list was expanded with new H.323 v.2 formats. However, there is no clearly defined way for gatekeepers to deal with both of these schemes, nor does the standard indicate if these are to be combined when communicating with another gatekeeper. Network developers are looking for generic solutions to address this issue. However, when different address resolution schemes are implemented, this will lead to interoperability problems.

7.4.2 GATEWAY TO GATEKEEPER COMMUNICATION

Gatekeeper to gateway communication, which resides at the application level, is another source of interoperability problems. In the case when one segment is defined by the terminal and gatekeeper (see (b) below) and the other segment is from the gatekeeper to a number of possible gateways (see (a) below), the gatekeeper has to be able to pass information (e.g. address resolution and routing information) back and forth between the gateway and the terminal. In this case the following applies:

a. The messages between the gateway and the gatekeeper are defined using RAS. The gateway broadcasts what bandwidth it has, what kinds of interfaces to the circuit-switched network it can provide for calls/users, and what prefixes should be dialled to get access to a specific service;

b. When the calling terminal wants to go through a gateway, it explicitly asks the gatekeeper for the gateway access. This necessitates that the terminal knows what kind of service it wants to receive in terms of address resolution and bandwidth/quality of service. This is communicated from the user interface to the network entities using RAS and Q.931 call signalling.

Because of inconsistencies in the definition of RAS and Q.931, the network applications may interpret the messages is different ways leading to interoperability problems.

7.4.3 A DECENTRALIZED MULTIPOINT CONFERENCE

In this case the terminals are responsible for selecting one or more of the received video streams for display, and for mixing the audio streams. However, there are a number of undefined situations that present themselves during a multipoint session. Examples include:

a. There are no specifications who suppose to issue the commands in Q.931, H.245 or both;

b. There are no specifications for the terminals how to detect which multipoint modes are supported.

Telehealth solution developers have implemented different solutions to address these issues. These, however, lead to interoperability problems.

8 INTEROPERABILITY REQUIREMENTS

8.1 GENERAL

Interoperability has different meanings, depending on which level in the architectural model it is considered:

- Interconnectivity – a seamless connection of different networks in a global infrastructure;
- Interworking – a seamless access, mapping, conversion, and configuration of services totally transparent to the applications;
- Interchange and interpretation – a support for interchanging information and the business-context-dependent interpretation of that information between different applications.

From the network and product suppliers' point of view, interoperability covers interconnectivity and interworking respectively, while from the user's point of view interoperability primarily means the correct exchange and use of business data.

Interoperability can be achieved in a variety of ways, including:

- Implementation of system and network components that interoperate at the desired level;
- Implementation of gateways, translators, or bridges that allow interoperability of disparate systems;
- Implementation of interoperable 'building blocks' from which telehealth solutions can be composed.

This section describes the requirements for interoperable solutions that use the first two approaches.

Figure 8 depicts examples of telehealth services and associated media data types. The rows represent Telehealth systems dedicated to providing specific telehealth services. The columns represent various Media Data Types required for the provision of these services. The planes represent different Network Types across which telehealth services are provided.

Wireless networks

Telephone networks

Circuit-switched networks

Packet-switched networks

	Audio	Video	Images	Graphics	Alphanumeric	Docs	Waveform	Time-series
Teleradiology	√	√	√	√	√	√		√
Telepsychiatry	√	√	√	√	√	√		
Telecardiology	√	√	√	√	√	√	√	√
Teledermatology	√	√	√	√	√	√		
Telehomecare	√	√	√		√		√	
Telelearning	√	√	√	√	√	√	√	√
…								

Figure 8 - Interoperability dimensions for real-time telehealth services

Telehealth systems provide various health services such as radiology, cardiology, and health education. Depending on their purpose, these systems are equipped with devices that are capable of handling various media data types (e.g. audio, video, waveform data) and performing various functions (e.g. acquisition, storage and management, encoding, compression, and transmission at the sending site and receiving, decompression, decoding, and presentation at the receiving site).

Telehealth systems are interconnected via various communications links or networks (e.g. circuit-switched networks, packet-switched networks, telephone networks, or wireless networks).

The task is to define interoperability requirements for end-to-end-telehealth solutions composed of telehealth systems capable of providing various functions and handling different media data types interoperating across heterogeneous networks.

These requirements cover several functions performed by real-time telehealth systems, including:

a. Call setup;

b. Acquisition, processing, and transmission of multimedia content in a full duplex or a half-duplex mode;

c. Controlling near- and far-end devices; and

d. Call termination.

These functions, along with a brief description of the interoperability dimensions across which the implementation of these functions differ, are discussed below. These differences in implementation cause interoperability problems in telehealth systems and networks.

8.2 CALL SETUP

The implementations of this function differ across telehealth systems and network types interconnecting the participating telehealth systems.

Indeed, in the ISDN network environment, the call is established using the D-channel. Subsequently a single B-channel is opened for bidirectional communication to transmit the H.221 framing data structure. It enables the significance of the bits to be understood by relating each of its position in the frame. At this point the terminal is initialized to work in audio-only mode. In the next phase the terminals exchange their capabilities whereby each terminal specifies the modes that it is capable of supporting. If additional channels are to be used, the terminal that established the initial call sets up the supplementary channels. With the channels to be used now established, the terminals switch to the common operating mode negotiated earlier (as per H.242) and the channels are synchronized to compensate for differences in latency between the constituent channels.

Telehealth systems operating in the IP-based networks handle this function quite differently. They use IP addresses instead of phone numbers and communicate the conference control parameters via TCP/IP. They use H.225 transmission format and H.245 communications procedure for signalling, registration, admissions, packetizing, multiplexing/demultiplexing, and synchronization of media streams.

To provide interoperability across different telehealth systems and network types, telehealth systems must be able to:

a. Establish a call over a variety of supported networking connections;

b. Determine whether the telehealth products can open channels and pass data after the call is established;

c. Accept a call using default codecs and negotiate a suitable set of codecs; and

d. Ensure that all control sequencing runs correctly.

Interoperability achieved with the use of the Call Setup function is at the interconnectivity and interworking level.

8.3 ACQUISITION, PROCESSING, AND TRANSMISSION OF MULTIMEDIA DATA

As indicated in Table 1, support for audio, video, and data codecs varies significantly between telehealth systems compliant with different H.3xx Recommendations. The implementations of this function differ significantly across all three dimensions: the telehealth systems, media data types, and network types.

From a user perspective this is the most important function, since at this stage the systems are engaged in exchanging data streams. During this phase, it is permissible to change modes, for example to change audio codec or introduce data conferencing capability into the conference call.

To provide interoperability across different telehealth systems, media data types, and network types, telehealth products must be able to:

a. Support acquisition, encoding / conversion, compression, and transmission of various Media Data Types defined in standard recommendations (e.g. H.323) at the sending site;

b. Support receiving, decompression, decoding / conversion, and presentation of Media Data Types defined in standard recommendations (e.g. H.323) at the receiving site;

c. Adapt to different transmission rates in the case when the systems are interconnected via different circuits e.g. ISDN and Switched 56 circuits (speed matching);

d. Adapt to differences in a codec mode (e.g. audio mode) from G.711 μ-law to G.711 A-law;

e. Support selection of audio, video, and data sources;

f. Combine (multiplex) and separate (demultiplex) control information and different data types (e.g. video, audio, and data) at the sending and receiving sites respectively;

g. Support data conferencing using the ITU-T T.120 series standards in conjunction with a video conferencing mode, including: