Public Safety Statement of Requirements

for Communications & Interoperability

Office for Interoperability and Compatibility
Department of Homeland Security

Volume II, Version 1.2

August 2008

Quantitative

This page intentionally left blank.

Homeland Security
Science and Technology

Command, Control and Interoperability: Communication, Interoperability and Compatibility

Defining the Problem

Emergency responders—police officers, fire personnel, emergency medical services-need to share vital voice and data information across disciplines and jurisdictions to successfully respond to day-to-day incidents and large-scale emergencies. Unfortunately, for decades, inadequate and unreliable communications have compromised their ability to perform mission-critical duties. Responders often have difficulty communicating when adjacent agencies are assigned to different radio bands, use incompatible proprietary systems and infrastructure, and lack adequate standard operating procedures and effective multi-jurisdictional, multi-disciplinary governance structures.

OIC Background

The Department of Homeland Security (DHS) established the Office for Interoperability and Compatibility (OIC) in 2004 to strengthen and integrate interoperability and compatibility efforts to improve local, tribal, state, and Federal emergency response and preparedness. Managed by the Science and Technology Directorate, and housed within the Communication, Interoperability and Compatibility thrust area, OIC helps coordinate interoperability efforts across DHS. OIC programs and initiatives address critical interoperability and compatibility issues. Priority areas include communications, equipment, and training.

OIC Programs

OIC programs, which are the majority of Communication, Interoperability and Compatibility programs, address both voice and data interoperability. OIC is creating the capacity for increased levels of interoperability by developing tools, best practices, technologies, and methodologies that emergency response agencies can immediately put into effect. OIC is also improving incident response and recovery by developing tools, technologies, and messaging standards that help emergency responders manage incidents and exchange information in real time.

Practitioner-Driven Approach

OIC is committed to working in partnership with local, tribal, state, and Federal officials to serve critical emergency response needs. OIC's programs are unique in that they advocate a "bottom-up" approach. OIC's practitioner-driven governance structure gains from the valuable input of the emergency response community and from local, tribal, state, and Federal policy makers and leaders.

Long-Term Goals

- Strengthen and integrate homeland security activities related to research and development, testing and evaluation, standards, technical assistance, training, and grant funding.
- Provide a single resource for information about and assistance with voice and data interoperability and compatibility issues.
- Reduce unnecessary duplication in emergency response programs and unneeded spending on interoperability issues.
- Identify and promote interoperability and compatibility best practices in the emergency response arena.

This page intentionally left blank.

Publication Notice

Abstract

This document contains the assembled requirements for a system of interoperable public safety communications across all local, tribal, state, and Federal "first responder" communications systems.

Change Log

Version	Date	Changes
1.0 Draft	April 2006	Initial Document (speech only)
1.0 Draft	July 2006	Initial Document—one of three (speech, video, network) Volume II documents broken out for review. Introduction, References, and Acronyms are general to speech, video, and network pieces which will become one Volume II document at publish time.
1.0	October 2006	Review input and copy edit changes have been incorporated.
1.0	May 2007	Changed branding from SAFECOM to the Office for Interoperability and Compatibility (OIC). Version and date were not revised since no content changes were made.
1.1	April 2008	While the October 2006 release included the first video requirements from the *PS1 narrow field of view tactical* experiment, the April 2008 release incorporated video requirements from tests in the *PS2 wide field of view tactical* experiment and the *PS3 tactical and live surveillance* experiment. Removed procedure and experiment appendices, and published them as technical reports. See Appendix B.2 for titles.
1.2	August 2008	New speech transmission performance section listing relationships between example requirements concerning intelligibility, speaker identification, and detection of dramatized urgency (Section 1.4). New requirements for Interaction of Minimum Bit Rate and a Maximum Packet Loss Ratio for tactical and live surveillance video (Section 4.3), and for recorded surveillance video (Section 4.4).

Acknowledgements

OIC extends its sincere appreciation to the many public safety practitioners, individuals, and government organizations that directly contributed to the creation of the Public Safety Statement of Requirements (PS SoR) for Communications and Interoperability (C&I).

Contact Information

Please send comments or questions to: S&T-C2I@dhs.gov

This page intentionally left blank.

Introduction

> *"In times of emergencies, the public looks to government, particularly their Public Safety officials, to act swiftly and correctly, and do the things which must be done to save lives, help the injured, and restore order. Most disasters occur without warning, but people still expect a rapid and flawless response on the part of government. There is no room for error. Whether involving a vehicle accident, crime, plane crash, special event, or any other Public Safety activity, one of the major components of responding to and mitigating a disaster is wireless communications. These wireless communications systems are critical to Public Safety agencies' ability to protect lives and property, and the welfare of Public Safety officials."*

This statement comes from the highly regarded *Public Safety Wireless Advisory Committee (PSWAC) Final Report*, presented to the Chairman of the Federal Communications Commission (FCC) and the Administrator of the National Telecommunications and Information Administration (NTIA) in September 1996.[1] The PSWAC Final Report defined and documented critical public safety wireless communication needs in 1996, and projected anticipated needs through the year 2010. The report focused on the requirements for communications resources and the radio frequency spectrum to support those requirements. While the report mentioned the crucial need to promote interoperability, its emphasis was clearly on the necessity of taking immediate measures to alleviate spectrum shortfalls. Fortunately, for public safety and for the benefit of all Americans, the report spurred the allocation of precious spectrum for use by public safety practitioners.

Unfortunately, the communication challenges for those working on the front lines in public safety have not been eliminated. In fact, at a time when more attention is being paid to interoperability among different disciplines and jurisdictions within the community, there still exists fundamental communication deficiencies within disciplines and jurisdictions as practitioners strive to perform the most routine and basic elements of their job functions. Agencies must be "operable," meaning they must have sufficient wireless communications to meet their everyday internal requirements before they place value on being "interoperable," meaning being able to work with other agencies.

This document, the *Public Safety Statement of Requirements (PS SoR) for Public Safety Communications and Interoperability*, is the natural follow-on to the PSWAC Final Report, but differs in three ways, as follows:

- First, the PS SoR is not keyed to the issue of spectrum allocation, but focused on public safety requirements from a broader perspective. Operational and functional requirements delineated in the PS SoR are not based on a particular approach or technology.
- Second, the PS SoR was developed eight years after the PSWAC Final Report was published. While the Final Report did not explicitly identify specific technological approaches along with the stated requirements, it is important to realize that advances in technology have helped to fashion the way practitioners think about their jobs over the years. Because practitioners expect more from technology today, their needs and desires have been affected, sometimes subtly, by industry advances and solutions that exist in today's commercial and consumer world. Additionally, current technological advances promote technically advanced thinking about what the practitioner may be able to expect 15 years from now. For instance, the possibility that technology refresh cycles could

1. In 1994, the FCC and NTIA established PSWAC to evaluate the wireless communications needs of local, tribal, state, and Federal public safety agencies through the year 2010, as well as to identify problems and to recommend possible solutions.

be dramatically reduced for public safety based on these advances is extremely attractive. That said, the methodologies used and the general projections made in the PSWAC Final Report remain as valid today as when they were first published. Based on the rapid changes and potential of technology, the PS SoR addresses current requirements and future requirements for the next 5 to 20 years.

- Third, the PS SoR emphasizes the "information" aspects of communications; that is, the need for the wireless exchange of data, video, and other non-voice mediums. The need for voice communications was clearly made in the PSWAC Final Report, as well as the need for additional bandwidth for other data resources. The PS SoR defines the information requirements of public safety practitioners more explicitly to guide how practitioners will use information resources in the field in mission-critical situations.

Scope

The PS SoR is currently a two-volume set. This volume, Volume II, provides detailed quality of service methods of measurement for the applications and services identified in Volume I of the PS SoR [28]. (Appendix B lists footnoted references.) It also provides network parameters to specify the minimum acceptable performance of public safety communications systems carrying these services. This document provides performance requirements for specific public safety applications. In addition, this document provides a common understanding between manufacturers and practitioners for specifying and maintaining these applications and services.

The initial application of this document is for mission-critical speech and video services, in addition to specifying network performance parameters to meet these applications' quality of service needs. Future revisions of this document will provide detailed quality of service metrics for the balance of the applications and services.

DHS Research Reports Supporting Requirements

The following Department of Homeland Security (DHS) technical reports describe methods for measuring the quality of service for different types of public safety communication. These measurement methods were used to determine the quantitative performance values found in this volume.

- Measurement of Speech Transmission Suitability —This report describes a laboratory study to determine how suitable or unsuitable different speech transmission systems would be for mission-critical communications in public safety operations. Specifically, public safety first responders listened to and evaluated a large number of recordings of speech transmission systems [32].
- Video Acquisition Measurement Methods—This report describes important video acquisition (i.e., camera) performance parameters for public safety applications [33].
- Network Measurement Methods—Graphs in this report plot packet loss probability versus the number of public safety communications devices (PSCDs) sharing the network, and expected delay versus the number of PSCDs sharing the access network to describe how speech and video applications carried over the network could effect network performance [34].
- Tactical and Surveillance Video Quality Experiments—This report describes laboratory studies to investigate the level of quality required for the following public safety video applications: Narrow field of view, tactical, live surveillance, and recorded surveillance; Wide field of view, tactical, live surveillance, recorded surveillance; live tactical and surveillance [35].

- Speech Intelligibility and Detection of Voice Characteristics—This report characterizes the relationships between speech intelligibility, speaker identification, and the detection of dramatized urgency in a speaker's voice across a wide range of simulated speech processing conditions. Experiment results indicate that for the speech processing conditions considered here, detection of dramatized urgency is the most robust property, speaker identification is less robust, and speech intelligibility is the least robust [36].
- Task-based Live and Recorded Surveillance Video Quality Tests—This report describes laboratory studies to investigate the level of quality required to perform tasks that include recognizing people, lettering, or other objects in the video content, regardless of the perceptual quality [37].

Intended Audience

The PS SoR focuses on the functional needs of public safety first responders—Emergency Medical Services (EMS) personnel, firefighters, and law enforcement officers—to communicate and share information as authorized when it is needed, where it is needed, and in a mode or form that allows the practitioners to effectively use it. The communications mode may be voice, data, image, video, or multimedia, the latter including multiple forms of information.

Because functional requirements are the focus of the PS SoR, it does not specify technologies or business models (i.e., whether requirements should be addressed through owned products and systems or via commercial services). Similarly, the PS SoR does not specify infrastructure, except to note that consistent with first responder operations, it is assumed that terminal links to and from practitioners are wireless unless stated otherwise.

The PS SoR addresses a number of complementary objectives. Most importantly, it is rooted in the goal of improving the ability of public safety personnel to communicate among themselves, with the non-public safety agencies and organizations with whom they work, and with the public that they serve. The PS SoR can also assist the telecommunication interoperability and information-sharing efforts by and among local, tribal, state, and Federal government agencies, and regional entities, by delineating the critical operational functions and interfaces within public safety communications that would benefit from research and development investment and standardization.

The PS SoR can assist Federal programs that work with public safety practitioners to assist wireless interoperability at all government levels to develop a comprehensive vision for public safety communications that satisfies the defined needs. This vision can be reinforced by developing Federal grant programs that promote government research and development, as well as investment in communications equipment and systems, in a manner consistent with the PS SoR.

The PS SoR provides information that can assist the communications industry to prioritize its research and development investment and product and service development strategies so that they are aligned with public safety communications needs.

The PS SoR is intended to be fully consistent with the National Incident Management System (NIMS)[2] as defined by the Federal Emergency Management Agency (FEMA) in DHS. Any inconsistency between this document and NIMS is a discrepancy, and will be addressed in later version of the document.

2. Developed by the Secretary of Homeland Security at the request of the President, NIMS integrates effective practices in emergency preparedness and response into a comprehensive national framework for incident management. NIMS will enable responders at all levels to work together more effectively and efficiently to manage domestic incidents no matter what the cause, size or complexity, including catastrophic acts of terrorism and major natural disasters.

Finally, the PS SoR can be used to clearly identify public safety operational issues so that discussions regarding existing and proposed regulations and laws can be dealt with expeditiously by regulatory and legislative bodies.

Organization of this Volume

Section 1	**Measuring Speech Transmission Performance** describes the speech performance parameters.
Section 2	**Mission-Critical Speech Transmission Requirements** lists requirement values for speech performance parameters.
Section 3	**Measuring Video Performance** describes video performance parameters.
Section 4	**Video Performance Requirements** lists requirement values for video performance parameters.
Section 5	**Reference Model for Network Performance** describes a path-based reference model.
Section 6	**Measuring Network Performance** describes the methods used to measure network performance.
Section 7	**Network Requirements** lists requirement values for network performance parameters.
Appendix A	**Glossary and Acronyms** lists terminology and acronyms used in this document.
Appendix B	**References** identifies books, standards, and online references for this document.

Contents

Publication Notice . . . . . **v**
Abstract v
Change Log v
Acknowledgements v
Contact Information v
Introduction . . . . . **vii**
Scope viii
DHS Research Reports Supporting Requirements viii
Intended Audience ix
Organization of this Volume x
1 Measuring Speech Transmission Performance **1**
1.1 Mission-Critical Speech Transmission Services 1
1.2 Reference Model for Speech Performance Measurements 1
1.3 Speech Transmission Factors 3
1.3.1 Speech Coding 3
1.3.2 Packetized Transmission of Digitized Speech Data 4
1.3.3 Transducers 4
1.3.4 Voice Activity Detection 4
1.3.5 Echo Control 5
1.3.6 Encryption 5
1.3.7 Mouth-to-Ear Delay 5
1.3.8 Background Sound 6
1.4 Intelligibility and other Speech Considerations 6
2 Mission-Critical Speech Transmission Requirements **9**
2.1 Mouth-to-Ear Delay 9
2.2 Packet Loss 11
2.2.1 Requirements for Mission-Critical Speech: 70 Percent Suitability 11
2.2.2 Requirements for Mission-Critical Speech: 80 Percent Suitability 12
2.2.3 Requirements for Mission-Critical Speech: 90 Percent Suitability 13
3 Measuring Video Performance **15**
3.1 Mission-Critical Video Services 15
3.2 Reference Model for Video Performance Measurements 16
3.3 Video System Parameters 17
3.3.1 One-Way Video Delay 17
3.3.2 Control Lag 18
3.3.3 Luma Image Size and Interlaced Versus Progressive Scan Type 18
3.3.4 Chroma Sub-Sampling Factors 19
3.3.5 Aspect Ratio 19
3.3.6 Frame Rate 20
3.3.7 Acceptability Threshold 20
3.4 Video Acquisition Parameters 20
3.4.1 Resolution 21

3.4.2 Noise . . . 21
3.4.3 Dynamic Range . . . 21
3.4.4 Color Accuracy . . . 21
3.4.5 Capture Gamma . . . 22
3.4.6 Exposure Accuracy . . . 22
3.4.7 Vignetting . . . 22
3.4.8 Lens Distortion . . . 23
3.4.9 Reduced Light and Dim Light Measurements . . . 23
3.4.10 Flare Light Distortion (Under Study) . . . 23
3.5 Video Transmission Parameters . . . 23
3.5.1 Parameters for Measuring Calibration Errors . . . 23
3.5.2 Parameters for Measuring Coding/Decoding Impairments . . . 24
3.5.3 Parameters for Measuring Impact of Network Impairments . . . 26
3.6 Video Display Parameters . . . 27
4 Video Performance Requirements . . . 29
4.1 Target Size and Scene Complexity . . . 29
4.2 General Public Safety Video Requirements . . . 29
4.3 Tactical and Live Surveillance Video Requirements . . . 31
4.3.1 Example Tactical Video Scenarios . . . 31
4.3.2 Example Live Surveillance Video Scenarios . . . 31
4.3.3 Tactical and Live Surveillance Video Transmission Requirements . . . 31
4.4 Recorded Surveillance Video Requirements . . . 32
4.4.1 Example Recorded Surveillance Video Scenarios . . . 33
4.4.2 Recorded Surveillance Video Transmission Requirements . . . 33
4.4.3 Feature Recommendations for Forensic Video Analysis . . . 34
5 Reference Model for Network Performance . . . 37
5.1 Mission-Critical Network Services . . . 37
5.2 Path Model Definition . . . 38
5.3 Path Model Parameters . . . 42
5.3.1 Medium Access Control . . . 42
5.3.2 Propagation . . . 43
5.3.3 Channel Data Rate . . . 43
5.3.4 Public Safety Communications Device . . . 43
5.3.5 First Responder's Vehicle . . . 43
5.3.6 Jurisdiction Communication Tower . . . 44
5.3.7 Generic Nodes . . . 44
5.3.8 Node Delay . . . 44
5.3.9 Area Networks . . . 45
5.3.10 Number of Nodes . . . 46
5.3.11 Protocols . . . 46
5.3.12 User Applications . . . 47
6 Measuring Network Performance . . . 49
6.1 Factors Affecting Network Performance . . . 49
6.1.1 Noise . . . 49
6.1.2 Interference . . . 50
6.1.3 Packet Collisions . . . 51

6.1.4 Packetization . . . 51
6.1.5 Queuing . . . 51
6.1.6 Packet Loss and Retransmission . . . 52
6.2 Packet Loss Ratio Computations . . . 52
6.2.1 Dedicated Channel . . . 53
6.2.2 Slotted Aloha . . . 53
6.3 End-to-End Packet Transfer Delay Computations . . . 55
6.3.1 Link Delays . . . 56
6.3.2 Node Delays . . . 56
6.3.3 Medium Access Delays . . . 56
7 Network Requirements . . . 59
7.1 User-Perceived Quality of Service . . . 59
7.2 Speech Applications . . . 60
7.2.1 Packet Loss Requirements . . . 60
7.2.2 End-to-End Delay Requirements . . . 61
7.2.3 Path A . . . 63
7.2.4 Path B . . . 63
7.2.5 Path C . . . 64
7.2.6 Path D . . . 65
7.2.7 Path E . . . 67
7.2.8 Path F . . . 68
7.2.9 Path G . . . 69
7.3 Video Applications . . . 70
7.3.1 Packet loss Requirements . . . 70
7.3.2 End-to-End Delay Requirements . . . 71
7.3.3 Path A . . . 72
7.3.4 Path B . . . 73
7.3.5 Path C . . . 74
7.3.6 Path D . . . 75
7.3.7 Path E . . . 77
7.3.8 Path F . . . 78
7.3.9 Path G . . . 79
7.4 Summary for All Area Networks . . . 80
Appendix A Glossary and Acronyms . . . 83
Appendix B References . . . 87
B.1 Book and Standards References . . . 87
B.2 Online References . . . 89

This page intentionally left blank.

Figures

Figure 1: Digital Speech Transmission Reference Model . . . 2
Figure 2: Rating as Delay Varies for Two G.711-Based Speech Transmission System . . . 10
Figure 3: Video Performance Measurements Reference Diagram . . . 16
Figure 4: Two Example Video Transmission Systems, with Reference Points Identified . . . 17
Figure 5: Natural Network Hierarchy . . . 38
Figure 6: Link Diagram . . . 39
Figure 7: Hierarchical Reference Paths Based on "Natural Network Hierarchy" . . . 41
Figure 8: Peer Reference Paths (by Links) Based on "Network Diagram Link Descriptions" . 41
Figure 9: Protocol Stack for End User's PSCDs . . . 46
Figure 10: Nodes and Links Composing a Path, with Numerical Identifiers . . . 53
Figure 11: General Performance Requirements for User-Perceived Quality of Service . . . 59
Figure 12: Speech Maximum Packet Loss Ratio Requirements . . . 61
Figure 13: Speech Maximum End-to-End Delay Requirements . . . 62
Figure 14: Video Maximum Packet Loss Ratio Requirements . . . 71
Figure 15: Video Maximum End-to-End Delay Requirements . . . 72

This page intentionally left blank.

Tables

Table 1: Requirements Related to Speaker Identification and Detection of Emotions 7
Table 2: Speech Requirement Set 1 12
Table 3: Speech Requirement Set 2 13
Table 4: Speech Requirement Set 3 14
Table 5: General Public Safety Video Performance Requirements 30
Table 6: Tactical and Live Surveillance Video Performance Requirements 31
Table 7: Recorded Surveillance Video Performance Requirements 33
Table 8: Video Feature Recommendations for Forensic Video Analysis 34
Table 9: Symmetrical and Asymmetrical Network Path Types 40
Table 10: Packet Loss Requirements for Percentages of Satisfied Practitioners 60
Table 11: Speech Path A Network Performance Parameter Requirements 63
Table 12: Speech Path B Network Performance Parameter Requirements 64
Table 13: Speech Path C Network Performance Parameter Requirements 65
Table 14: Speech Path D Network Performance Parameter Requirements 66
Table 15: Speech Path E Network Performance Parameter Requirements 67
Table 16: Speech Path F Network Performance Parameter Requirements 68
Table 17: Speech Path G Network Performance Parameter Requirements 69
Table 18: Video Path A Network Performance Parameter Requirements 72
Table 19: Video Path B Network Performance Parameter Requirements 73
Table 20: Video Path C Network Performance Parameter Requirements 74
Table 21: Video Path D Network Performance Parameter Requirements 76
Table 22: Video Path E Network Performance Parameter Requirements 77
Table 23: Video Path F Network Performance Parameter Requirements 78
Table 24: Video Path G Network Performance Parameter Requirements 79
Table 25: Maximum Allowable Packet Loss and Delay for Each Type of Area Network 80

This page intentionally left blank.

1 Measuring Speech Transmission Performance

This section describes technical components that affect the ability to provide full-duplex public safety speech transmission services. Packetizing speech as digital data for wired and wireless transmission supports the following public safety communications goals:

- Highly flexible system architectures
- Different operating modes for Personal Area Networks (PANs), Incident Area Networks (IANs), Jurisdiction Area Networks (JANs), and Extended Area Networks (EANs)
- Speech, video, and data communications across networks

1.1 Mission-Critical Speech Transmission Services

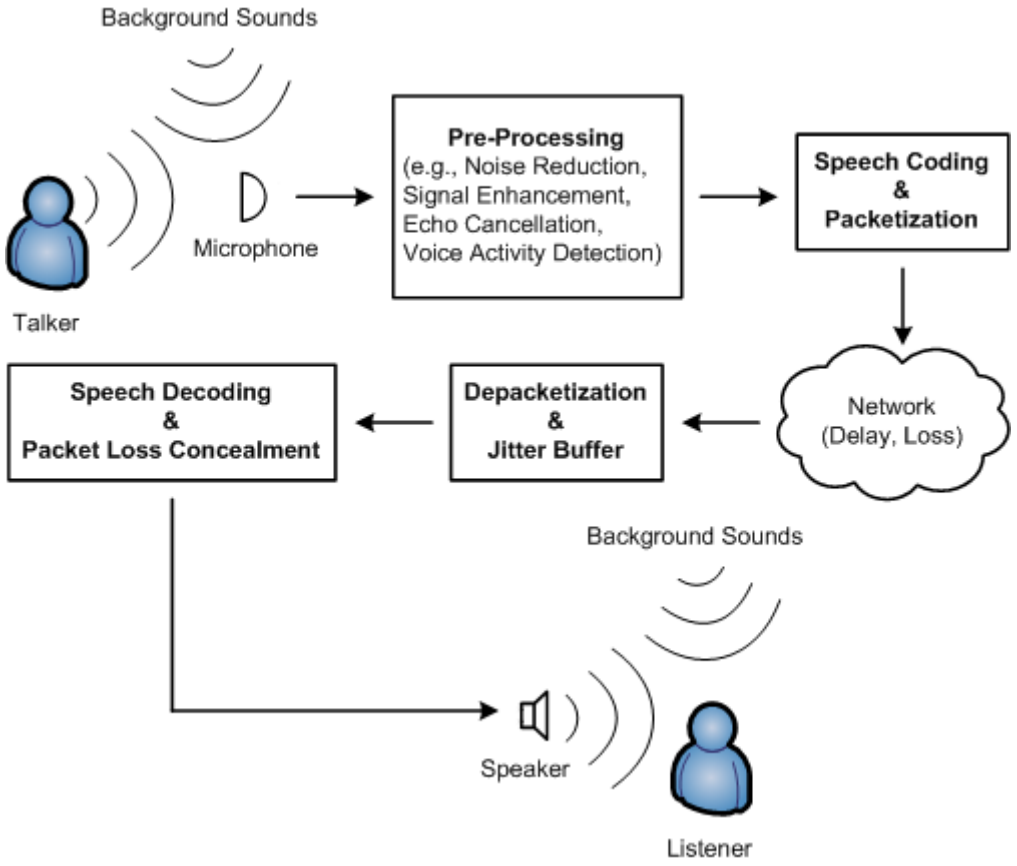
Services that provide speech transmission between two locations to support public safety operations have become firmly established over the decades. A classic example is a land-mobile radio (LMR) system that allows mobile practitioners to speak with each other, and with a dispatcher. These systems are typically half duplex and require each practitioner to "push to talk." Analog systems were once pervasive, and those systems have been migrating to digital operation in more recent years.

The public safety communications community is looking towards a future family of full-duplex public safety speech transmission services. This family of services would be securable, robust, and scalable for operation over PANs, IANs, JANs, and EANs. The services would transmit speech with a quality that is suitable for mission-critical communications.

1.2 Reference Model for Speech Performance Measurements

Figure 1 provides a high-level depiction for one direction of a generic, full-duplex, packetized, digital speech transmission system.

Figure 1: Digital Speech Transmission Reference Model



The quality of the speech transmission delivered by such a system will depend on equipment factors including, but not limited to the following elements:

- Microphone
- Pre-processing (noise reduction, signal enhancement, echo cancellation, speech-activity detection) algorithms
- Speech coding algorithm
- Packet size
- Jitter-buffer size and playout algorithm
- Packet loss concealment algorithm
- Earpiece or loudspeaker

In addition, the quality of speech transmission will depend on the following operational factors including, but not limited to:

- Network delay statistics and network loss statistics
- Background sound types and levels at each end
- Characteristics of the talkers, listeners, and messages

This section discusses these factors and provides some related choices. The DHS technical report, "Measurement of Speech Transmission Suitability," [32] provides a detailed description of the laboratory study conducted for this PS SoR to determine an initial set of requirements for network loss statistics, such that speech delivered is "suitable for mission-critical speech communications," as judged by public safety practitioners. The results necessarily reflect a spread of opinions among the practitioners. Section 2 provides the resulting set of initial requirements which are valid under the necessary conditions and constraints presented in Section 1 and the "Measurement of Speech Transmission Suitability" report.

1.3 Speech Transmission Factors

In the most general case of speech transmission using packetized transmission of digital data, there are many factors to consider. It was not possible to consider all possible combinations of these in a single laboratory study. Experts did, however, consider each factor independently and made informed choices to arrive at the most relevant and practical set of factors to include in the laboratory study. These choices reflect the need to balance the goal of creating a realistic and relevant environment against the goal of developing a well-controlled and performed laboratory study. The next subsections discuss each of the speech transmission factors and the related choices.

1.3.1 Speech Coding

Public safety operations are often conducted in environments with significant background sound. While noise-canceling microphones can often reduce the level of undesired background sounds entering the speech coder, they cannot always eliminate them. In some cases, the background sounds may contain important information, and accurate transmission of those sounds, along with speech, can be desirable.

Operations can be enhanced by knowledge of exactly who has transmitted speech and what emotional state is represented by his or her speech.

Thus for future services, speech coders that can very accurately transmit the fine nuances (e.g. talker-specific attributes and emotional-state specific attributes) of a human voice or voices along with arbitrary background sounds are necessary. (Note that transmitting both speech and background sounds allows for the optional use of signal processing-based noise reduction techniques at the receiving location, if and when the practitioner at the receiving location desires.) Many speech coders rely on mathematical models of a single speech signal to achieve bit rate reduction. Such coders can attain very low bit rates, but are not intended to accurately transmit fine speech and background sound nuances, and in general, cannot accurately transmit such nuances.

The highest fidelity is attained by directly coding the actual speech-plus-background sound waveform, making few, if any, assumptions about that waveform. The companded pulse-code modulation (PCM) speech coders specified in ITU-T Recommendation G.711 [19] give one efficient approach. These speech coders have very low encode and decode complexity and delay. They operate at 64 kbps. Appendix I of Rec. G.711 [18] includes a packet loss concealment (PLC) algorithm. This algorithm minimizes distortions caused by lost channel data, and is well-suited to networks where channel data is packetized and some packets may be lost. As detailed in the "Measurement of Speech Transmission Suitability" report, the laboratory study focuses on G.711 speech coding accompanied by this PLC algorithm.

Other methods of making G.711 speech coders more robust to packet loss include PLC algorithms and multiple description, or diversity, coding algorithms. Development work in this field continues. However, G.711 Appendix I is the only approach known to be fully disclosed and codified in a formal way, and thus

it is most suitable for the laboratory study. If other algorithms are eventually demonstrated to provide higher robustness, it may then be possible to relax the packet loss requirements given in Section 2.

1.3.2 Packetized Transmission of Digitized Speech Data

In packetized data transmission, delivery of all packets within a specified time window is not always guaranteed. Packets that fail to arrive within the required time may be considered lost. In addition, packets that are not lost may experience dissimilar transmission delays. This means that a stream of packets sent at uniform time intervals may be received at non-uniform time intervals. Since speech decoders generally require data at uniform time intervals, a jitter buffer is often used to provide such a data stream, at the cost of some additional fixed delay. For a given network configuration, increasing network traffic can increase network congestion, which in turn can lead to an increase in lost packets and a wider variation in transmission delays. The laboratory study requires some model of these processes.

Network models with varying levels of detail and complexity are available. More detailed models may better reproduce specific network behaviors, while less detailed models can show wider applicability. Given that no specific network details are presently available, it seems prudent to use a very basic and general model for the combined effects of the network and the jitter buffer. Thus this study treats the network and the jitter buffer together as a single black box that can be parameterized (for at least tens of seconds) by a pair of packet loss parameters. These are fundamental properties and thus they provide a basic yet relevant model.

The two packet loss parameters are packet loss ratio and packet loss correlation. When packet loss correlation is zero the packet loss process is random. As packet loss correlation is increased, the loss of packets becomes more bursty, and it becomes more likely that multiple packets will be lost in succession. In practice, packetized data networks can exhibit random or bursty packet loss patterns. The model used in the laboratory study to represent the network and jitter buffer is detailed in the “Measurement of Speech Transmission Suitability” [32] report. The model is applied to two different packet sizes; those that contain data representing 10 milliseconds (ms) of speech signal, and those that contain 40 ms.

1.3.3 Transducers

The determination of suitable acoustical and electrical properties for transducers (e.g., microphones, loudspeakers, and earpieces) used in public safety operating locations is a potential topic for future studies. Such studies will need to carefully account for the range of acoustic environments that may be present at these locations.

In the present laboratory study, we assume that these transducers are not limiting performance. Rather, G.711 speech coding and packet loss are the limiting factors in attaining “suitability for mission-critical voice communications.” We use studio-quality microphones and loudspeakers in the laboratory work to ensure that these transducers do not limit performance.

1.3.4 Voice Activity Detection

Voice activity detection (VAD) may be used in full-duplex speech transmission systems to prevent a speech coder from generating a full rate data stream when no one is talking. The challenges in VAD include accurate detection of voice activity, particularly the starts of utterances, and especially in the presence of significant background sounds. VAD is not appropriate for cases that require transmission of the most accurate representation of the acoustic environment.

The appropriateness of VAD and the consequences of imperfect VAD operation is a potential subject for future investigations. The present laboratory study does not use any VAD. Thus the results are not confounded by issues of VAD performance.

1.3.5 Echo Control

In full-duplex communications, there is the possibility a practitioner will be distracted by an echo of his or her own voice. These echoes can be of electrical or acoustical origin. Echoes can exacerbate the effects of delay since increased delay makes echoes more audible and annoying. In systems where delay and echo levels can cause annoyance, echo cancellers are typically deployed to minimize the levels, and thus the annoyance, of the echoes. The laboratory study supporting the requirements in Section 2 does not include any sources of echoes, or any echo control devices. Thus the results are not confounded by echo or echo control issues.

1.3.6 Encryption

Some applications require encryption to keep speech transmissions secure. The laboratory study described in the "Measurement of Speech Transmission Suitability" [32] report assumes that any encryption system used is transparent to the data stream, even when data packets are lost. If an encryption system requires additional data handling capabilities, or increases the data transmission delay, this must be considered separately.

1.3.7 Mouth-to-Ear Delay

Mouth-to-ear delay identifies the elapsed time between a sound leaving a talker's mouth and arriving at a listener's ear. (This is different from call-setup delay, which is the time associated with the initial establishment of communications between the parties involved.) In face-to-face communications and in many speech transmission systems, the mouth-to-ear delay is either imperceptible or negligible. In other systems, the mouth-to-ear delay can be significant, and can even be large enough to impair the communications attempted by the two parties. For the speech transmission systems considered here, the mouth-to-ear delay is likely to be dominated by the following types of delays:

- Packetization delays
- Network queuing delays
- Network transmission delays
- Jitter buffer delays

Other contributions to the mouth-to-ear delay that are likely to be negligible in most implementations include G.711 encoding and decoding delays and the mouth-to-microphone and loudspeaker-to-ear acoustic propagation times.

As a practical matter, it is necessary to treat mouth-to-ear delay separately from other speech transmission factors in the present study. In potential future studies, interactions between mouth-to-ear delay and other factors might be evaluated through laboratory studies of human subjects engaged in conversation tasks rather than the listening task employed in the "Measurement of Speech Transmission Suitability" laboratory study. Such studies require real-time implementations of each system that is to be evaluated, and each real-time implementation must support speech traffic in two directions. These studies are thus significantly more complex than studies that use listening tasks.

1.3.8 Background Sound

Background sound in public safety operations is often, if not always, present at some level. Background sound types and levels can vary greatly between locations (e.g., dispatch center, parked patrol car, fire truck responding to alarm). In general, the background sound environment at the talking location and the environment at the listening location have the potential to influence the perception of transmitted speech. A rigorous exploration of the variables of talking location background sounds, as well as listening location background sounds, could fill numerous potential future studies. In the "Measurement of Speech Transmission Suitability" laboratory study, a single background sound level and type is simulated for the talking location. A single background sound type and level type and two different sound levels are simulated at the listening location. See the "Measurement of Speech Transmission Suitability" [32] report for full details.

1.4 Intelligibility and other Speech Considerations

Mission-critical speech transmission systems must preserve message intelligibility. These systems also should aim to successfully transmit attributes of the speaker's voice. If transmitted successfully, this can allow a listener to identify the speaker or to confirm the purported identity of the speaker (speaker identification or SID).

SID can be particularly important for officials who must communicate rapidly to accomplish time-critical emergency operations. If speakers can be identified implicitly based on transmitted attributes of their voices, the additional overhead associated with explicit identification ("This is Officer Smith speaking.") can be avoided. In addition, if it is possible to detect that a speaker is not as claimed, this could be a very important discovery.

Officials sometimes monitor multiple transmissions with partial attention while also performing other important duties. If one of several other officers speaking displays a shift in emotional state via his or her voice, detecting that shift can be very important. When such a shift is detected it could be important to then commit full attention to that specific speaking officer in order to provide support.

The report [36] provides detailed descriptions of a set of experiments designed to characterize, relative to each other, the intelligibility, SID, and detection of emotion properties of speech transmission systems. In the context of the work described in [36], it is clear that speech transmission system impairments reduce word intelligibility much more quickly than they reduce the ability to perform SID. It is also clear that speech transmission system impairments reduce word intelligibility much more quickly than they reduce the ability to detect the emotional state called "Dramatized Urgency" (DU). In short, SID and detection of DU show a greater robustness to speech transmission impairments than word intelligibility does.

Not only is word intelligibility more robust than SID and detection of DU, but it is also arguably more fundamental and important to speech communications than SID and detection of DU. Together these observations mean that intelligibility requirements will dominate over SID and detection of DU requirements. In other words, if a speech transmission system provides sufficient intelligibility, then it will very likely also provide for sufficient SID and detection of DU.

Table 1 lists relationships between example requirements concerning intelligibility, SID, and detection of DU. Experiments in [36] studied these three tasks under six conditions, C1 through C6. In [36], the C1 normalized task performance (NTP) scores provide a best-case reference point (baseline) for each of the three tasks. In Table 1, C1 corresponds to C1 in [36], and Table 1 presents C1 as a benchmark condition. C2 through C6 in Table 1 correspond to C2 through C6 in [36], and are compared against the benchmark.

Speech processed with C2 results in a transmission quality that is lower than speech processed with C1, but higher than speech processed by C3 (and so on). See [36] for other details and assumptions.

Table 1: Requirements Related to Speaker Identification and Detection of Emotions

	Better ← Transmission Quality → Worse					
	Experimental Condition					
Parameter	**C1 (benchmark)**	**C2**	**C3**	**C4**	**C5**	**C6**
Word Intelligibility	Baseline (NTP score)	0 Percent Lower (than baseline)	15 Percent Lower	15 Percent Lower	30 Percent Lower	85 Percent Lower
Speaker Identification	Baseline	0 Percent Lower	5 Percent Lower	10 Percent Lower	15 Percent Lower	30 Percent Lower
Detection of Dramatized Urgency	Baseline	0 Percent Lower	5 Percent Lower	5 Percent Lower	10 Percent Lower	20 Percent Lower

Here is an example to aid in interpretation of Table 1. If there is a requirement that allows word intelligibility degradation of no more than 30 percent lower than the C1 NTP score for intelligibility, then C5 is relevant, and the table indicates that SID would meet a requirement that allows degradation of no more than 15 percent lower than the C1 NTP score for SID, and detection of DU would meet a requirement that allows degradation of no more than 10 percent lower than the C1 NTP score for DU.

Laboratory experiments like those described in [36] are important because they provide a level of control over speaking, listening and speech processing conditions that allows one to extract meaningful results. This would not be possible in a typical field environment. While laboratory experiments are essential to research progress, it should also be noted that laboratory experiments can be less realistic than the actual field environment.

2 Mission-Critical Speech Transmission Requirements

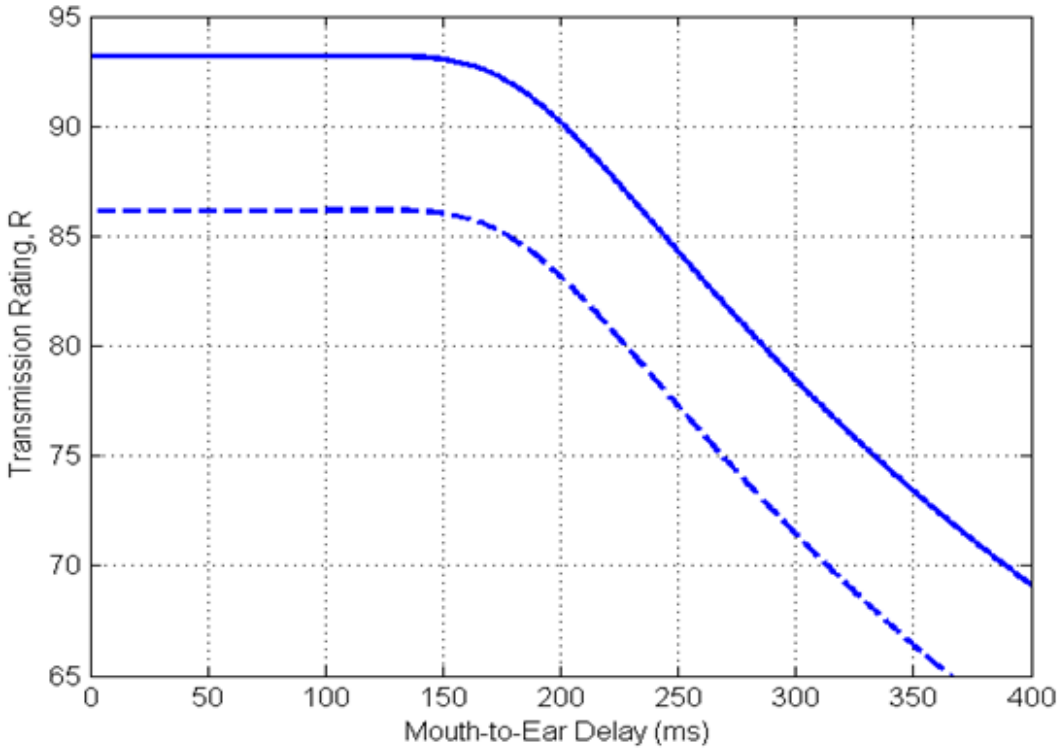
This section outlines a requirement for mouth-to-ear delay in speech transmission systems used for mission-critical communications. This requirement is based on existing results in the literature. The section also provides a set of packet loss requirements that when met, are expected to provide speech transmission that is suitable for mission-critical communications. Each requirement is related to an estimated percentage of practitioners that will find the results to be suitable for mission-critical communications. These requirements are based on the laboratory study described in the "Measurement of Speech Transmission Suitability" [32] report.

2.1 Mouth-to-Ear Delay

The issue of mouth-to-ear delay is addressed here primarily through information contained in ITU-T Recommendations G.107 (the E-Model) [24] and G.114 [22]. These recommendations reflect what is known about the relationship between mouth-to-ear delay, and practitioner satisfaction in the telecommunications context. They are based on extensive research conducted by various telephone operating companies and administrations around the world, over a period of years.

Figure 2 expands on Figure 1 of G.114. The solid line on this graph matches that of G.114 Figure 1, is generated by equations given in G.107, and relates mouth-to-ear delay to the Transmission Rating factor, R. This solid line treats the case where mouth-to-ear delay and G.711 encoding are the only significant factors impairing speech transmission. G.107 further indicates that $90 \le R$ will result in "very satisfied" users, $80 \le R < 90$ will result in "satisfied" users, and $70 \le R < 80$ will result in "some users dissatisfied."

Figure 2: Rating as Delay Varies for Two G.711-Based Speech Transmission System



The dashed line in Figure 2 is also generated by the equations given in G.107 [24], but for the case of packetized G.711 speech coding, with 10 ms packets, G.711 Appendix I PLC, and a random packet loss ratio of 2 percent. This result also requires the use of a packet loss robustness factor, found in Appendix I of G.113 [19]. This second curve is simply a shifted version of the first curve, and this is a visual manifestation of the key underlying principle used in G.107; "psychological factors on a psychological scale are additive." In the context of mission-critical speech transmission requirements, the principle indicates that perceived speech transmission degradation due to mouth-to-ear delay and due to G.711 coding, packet loss, and PLC combined, are additive.

Based on the extensive research behind G.107, this principle is expected to hold for all speech transmission systems included in this study. Thus, each system-specific curve showing R versus mouth-to-ear delay is simply a downward-shifted version of the original curve, with greater downward shifts associated with more highly impaired systems. The key feature of interest, common to all of these curves, is that they show nearly no negative effects of mouth-to-ear delay until that delay is about 150 ms.

Based on the information available at this time, the largest permissible mouth-to-ear delay for mission-critical communications is 150 ms, and this assumes that no audible echoes are present. The use of any greater value would first require laboratory evaluation (using conversation tasks rather than listening tasks) of the combined effect of that delay and the other impairments allowed by the packet loss requirements given in Section 2.2.

We can restate this conclusion and further highlight the logic behind it using the language of additive degradations introduced above. When meeting the packet loss requirements given in Section 2.2, it is possible to "use up the entire degradation budget." No portion of the degradation budget remains for

mouth-to-ear delay. This means that mouth-to-ear delay must remain in the range that does not add any degradation (i.e., 0 to 150 ms.). It is natural to ask if the degradation budget can be used differently; one might seek to use a less stringent delay requirement coupled with more stringent packet loss requirements. The definitive answer can only be found through laboratory studies of human subjects engaged in conversation tasks, where the relevant mouth-to-ear delay and the packet loss characteristics are simulated in real time for both directions of the conversation.

It is also natural to ask if the mathematical tools provided with G.107 [24] could provide an alternative to the required laboratory studies. At this time the answer is no. G.107 and the associated recommendations do not cover all of the higher packet loss ratio or non-random packet loss cases of interest, nor do they cover the case of 40 ms G.711 packets. An additional complication arises from the fact that at present, there is no well-established relationship between the R values produced by G.107 and the notion of "suitable for mission-critical communications."

Finally, recall from Section 1.3.7 that for the speech transmission systems considered here, the mouth-to-ear delay is likely to be dominated by packetization delay, network queuing delays, network transmission delays, and jitter buffer delay. In other words, the 150 ms delay budget must be allocated among these various sources of delay.

2.2 Packet Loss

The packet loss requirements given in Table 2, Table 3, and Table 4 are based on analysis of 12,320 votes collected in the laboratory study, as described in the "Measurement of Speech Transmission Suitability" [32] report. We view these votes to be samples of the pool of all possible votes that the entire body of public safety practitioners in this country, hearing all possible messages, could cast. If the samples (votes collected in this study) are representative (with respect to parameters that affect the votes) of the larger pool, then we can use the collected votes to find an estimate of the votes in that larger pool. Specifically, we can find estimates for the fraction of "yes" votes. The fraction of "yes" votes is the fraction of public safety practitioners that find a speech transmission to be "suitable for mission-critical communications."

The combinations of packet loss ratio, packet loss correlation, and packet size detailed in the "Measurement of Speech Transmission Suitability" report, define a total of 79, G.711-based speech transmission systems. In each table, systems with the attributes marked by "*" are expected to meet the requirements stated in the paragraph preceding each table. Table 2, Table 3, and Table 4 address attaining estimated "yes" votes from 70, 80, and 90 percent of the users, respectively. As expected, as requirements become more and more stringent (moving from 70, to 80, to 90 percent estimated "yes" votes), only lower and lower levels of packet loss and packet loss correlation will support the requirement. As described in Section 2.1, a single mouth-to-ear delay requirement of 150 ms appears in all three tables.

2.2.1 Requirements for Mission-Critical Speech: 70 Percent Suitability

It is expected that at least 70 percent of public safety practitioners will judge a speech transmission system "suitable for mission-critical communications" when the system:

- Has acceptable packet loss ratio and packet loss correlation combinations as identified by cells marked with "*" in Table 2.
- Conforms to the constraints and assumptions detailed in Section 1 and the "Measurement of Speech Transmission Suitability" report [32].

Table 2: Speech Requirement Set 1

	Packet Size = 10 ms				Packet Size = 40 ms			
	Packet Loss Ratio				Packet Loss Ratio			
Packet Loss Correlation	0%	2%	5%	10%	0%	2%	5%	10%
0.0	*	*	*	*	*	*	*	
0.1	*	*	*	*	*	*	*	
0.2	*	*	*	*	*	*	*	
0.3	*	*	*		*	*	*	
0.4	*	*	*		*	*		
0.5	*	*	*		*	*		
0.6	*	*	*		*	*		
0.7	*	*			*	*		
0.8	*	*			*	*		
0.9	*	*			*	*		

Mouth-to-Ear Delay Requirement: No greater than 150 ms.

2.2.2 Requirements for Mission-Critical Speech: 80 Percent Suitability

It is expected that at least 80 percent of public safety practitioners will judge a speech transmission system "suitable for mission-critical communications" when the system:

- Has acceptable packet loss ratio and packet loss correlation combinations as identified by cells marked with "*" in Table 3.
- Conforms to the constraints and assumptions detailed in Section 1 and the "Measurement of Speech Transmission Suitability" report [32].

Table 3: Speech Requirement Set 2

	Packet Size = 10 ms				Packet Size = 40 ms			
	Packet Loss Ratio				Packet Loss Ratio			
Packet Loss Correlation	0%	2%	5%	10%	0%	2%	5%	10%
0.0	*	*	*		*	*		
0.1	*	*	*		*	*		
0.2	*	*	*		*	*		
0.3	*	*	*		*	*		
0.4	*	*	*		*	*		
0.5	*	*			*	*		
0.6	*	*			*	*		
0.7	*	*			*	*		
0.8	*	*			*	*		
0.9	*	*			*	*		

Mouth-to-Ear Delay Requirement: No greater than 150 ms.

2.2.3 Requirements for Mission-Critical Speech: 90 Percent Suitability

It is expected that at least 90 percent of public safety practitioners will judge a speech transmission system "suitable for mission-critical communications" when the system:

- Has acceptable packet loss ratio and packet loss correlation combinations as identified by cells marked with "*" in Table 4.
- Conforms to the constraints and assumptions detailed in Section 1 and the "Measurement of Speech Transmission Suitability" report [32].

Table 4: Speech Requirement Set 3

	Packet Size = 10 ms				Packet Size = 40 ms			
	Packet Loss Ratio				Packet Loss Ratio			
Packet Loss Correlation	0%	2%	5%	10%	0%	2%	5%	10%
0.0	*	*			*	*		
0.1	*	*			*	*		
0.2	*	*			*	*		
0.3	*	*			*	*		
0.4	*	*			*			
0.5	*				*			
0.6	*				*			
0.7	*				*			
0.8	*				*			
0.9	*				*			

Mouth-to-Ear Delay Requirement: No greater than 150 ms.

Note that jitter buffer design can be used to trade loss and delay. That is, when the attained delay is smaller than the delay requirement but the packet loss requirements are not met, increasing the size of the jitter buffer will increase the delay and may (depending on network loss and delay variation statistics) reduce packet loss. On the other hand, delay can be reduced by making the jitter buffer smaller, but this will typically (depending on network loss and delay variation statistics) increase the packet loss. In potential future work, an additional goal might be to separate out jitter buffer operations to arrive at a set of pure network requirements.

Note also that these results address only a single communications link between two points. In potential future work, one might develop models of practitioner-generated traffic and apply them, building on the present results to develop more general network requirements for larger numbers of practitioners communicating between larger numbers of network nodes.

3 Measuring Video Performance

Video applications are important in the mission of public safety. They will only increase in importance over time. The initial video performance effort in this document focuses on mission-critical video services. Mission-critical video services include applications in tactical public safety situations where there is a potential risk to human life (i.e., either to the lives of the first responders or to the individuals the first responders are aiding).

3.1 Mission-Critical Video Services

Qualitative descriptions of example mission-critical video services are:

- Ground-based and aerial video taken at the scene of a fire or other emergency sites to provide immediate tactical firefighting response, to coordinate rescue efforts, and to help distant EMS staff estimate required medical support, treatment, etc.
- Specialized non-visual video (such as infrared [IR]) to warn of spreading fire, heat sources, etc.
- Robotics video at an emergency site to control robotics devices and to assist with tactical decision making by the incident commander.
- Video in support of telemedicine taken by EMS staff at the scene of a fire or other emergency sites to help distant medical personnel evaluate patient condition and treatment, etc. Telemedicine techniques may require high-resolution video or pictures to allow viewing a patient's burns, skin and bone details, etc.
- Video taken at the scene of a stakeout, a traffic stop, or an arrest to send assistance in the case of trouble. This video may also be recorded for later use as evidence, for further investigation purposes, and to document officer conduct.
- Video used for a mutual aid operation; where there is a requirement to rapidly assess damage caused by the disaster, sending on-site views to recovery coordination officials at remote command posts. Real-time video may also be used by robotics operators and search and rescue teams, where the situation is too risky for first responders.

Mission-critical video services may include special situations and constraints that should be considered when developing a complete system specification. In these cases, the performance parameters in this document may not adequately characterize the required system performance. It is thus advisable to consider including additional performance parameters and specifications. Special situations and constraints may include, but are not limited to, the following:

- Where the environment lighting can range from very bright lighting to no lighting, or where there is an extremely wide range of lighting within the video scene
- Where the video needs to be encrypted to preserve a patient's Health Insurance Portability and Accountability Act (HIPAA) rights, and to protect first responders' tactical operations from those without a need to know
- Where the video may have to be stored, either at the scene or at a remote location, with very high resolution and clarity so it may be used as evidentiary material
- A remote control (zoom, pan, focus, etc.) video collection system

3.2 Reference Model for Video Performance Measurements

Figure 3 provides a reference model for specifying video performance measurements. To fully quantify the user-perceived quality of service, the performance of three primary video subsystems must be specified.

1. The Video Acquisition Subsystem—That normally consists of a camera system and may also include a built-in video coder.
2. The Video Transmission Subsystem—May include a network and its associated interfaces, encryption, etc., and may also include the video coder and decoder and a video storage medium.
3. The Video Display Subsystem—Includes a monitor, playback computer, etc. May also include a built-in video decoder.

The exact demarcation for each of the three subsystems can vary from application to application due to integration of various functions within the end user's equipment. The approach adopted here is to specify performance parameters for the application as a whole (i.e., System Parameters—see Section 3.3), as well as to specify performance parameters that are unique to each video subsystem (i.e., Acquisition Parameters—see Section 3.4; Transmission Parameters—see Section 3.5; and Display Parameters—see Section 3.6). These generic sets of performance parameters are capable of characterizing the performance of different public safety applications. However, some public safety applications may only use a selected subset of these performance parameters. Figure 3 depicts a reference diagram for the performance measurements. The letters in the figure denote measurement access points that may or may not be available on all video systems.

Figure 3: Video Performance Measurements Reference Diagram

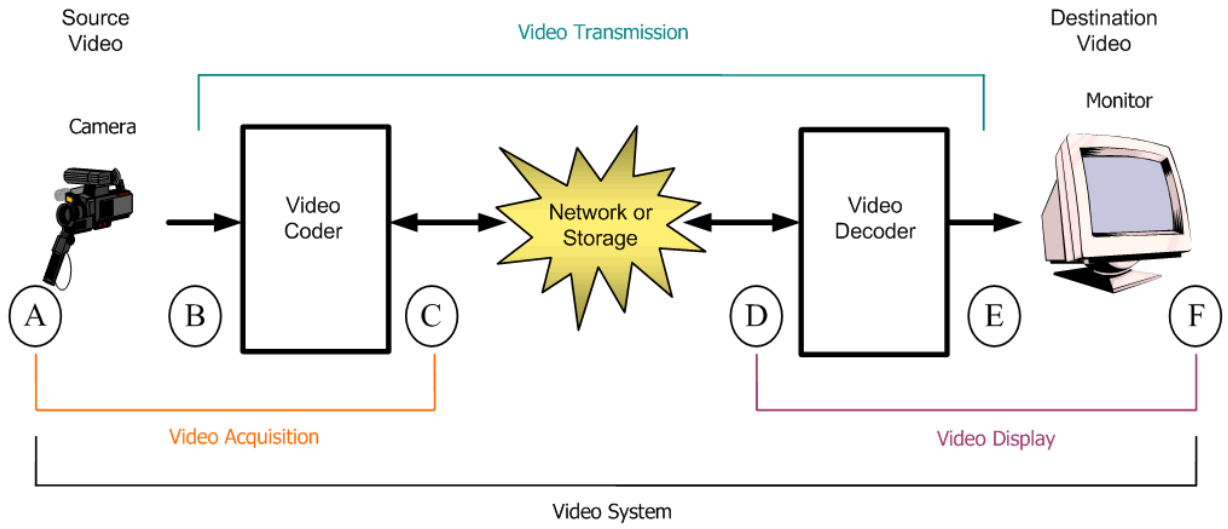
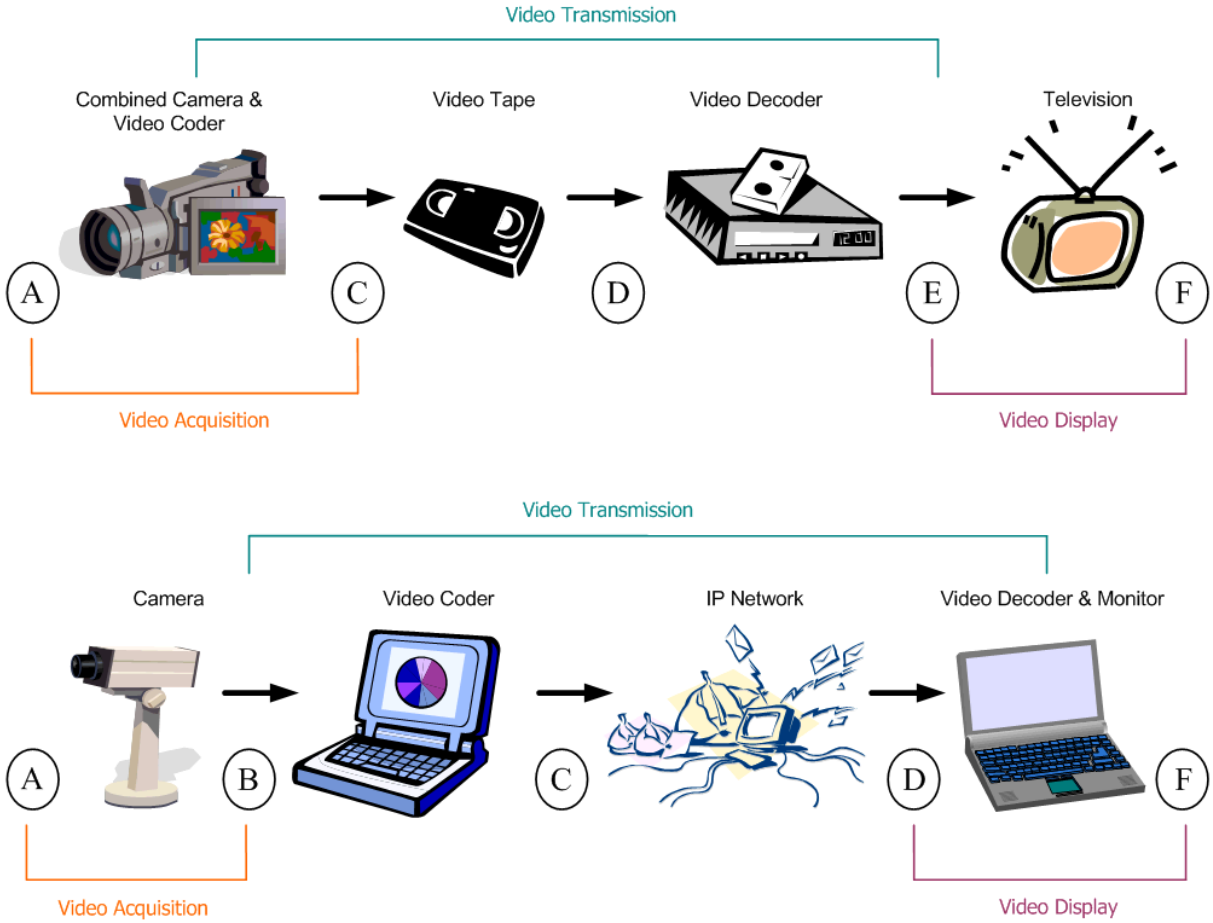


Figure 4 shows two example video systems with reference points identified. In the upper system, access point B is inside the camcorder and probably not available. In the lower system, access point E is inside the computer and may or may not be available. Some video may travel over multiple networks or storage medium.

Figure 4: Two Example Video Transmission Systems, with Reference Points Identified



3.3 Video System Parameters

Parameters in this section characterize the performance of the entire video system, from point A to point F in Figure 3. When a specification for a video system parameter is included, all the video subsystems (video acquisition, video transmission, and video display) must support this specification.

3.3.1 One-Way Video Delay

For interactive video services, an important performance parameter is the length of the time delay that is required to send video through the video system. Since coding and decoding (see Figure 3) can add substantial delay, some compression algorithms will not be suitable for public safety applications that have a low video delay requirement. ITU-T Recommendation P.931 [17] contains the recommended standard method of measurement for measuring the one-way video delay of the video transmission subsystem (i.e., from points B to E in Figure 3). To obtain the total one-way video delay (from points A to F in Figure 3), one must add to this the one-way delays of the video acquisition and video display subsystems. The video delays of the video acquisition and video display subsystems. The manufacturers of the equipment specify the video delays of the video acquisition and video display subsystems.

For some public safety applications, the two-way delay (i.e., round-trip delay) may be the more important specification. To obtain two-way delay, the one-way delay in each direction should be measured separately and then combined. This approach is taken because the delay may not be the same in both directions.

3.3.2 Control Lag

For some interactive applications there may be a control lag. Control lag is defined as the lag time between a user's request and the time that the request is actually implemented. For example, there may be a control lag between a controller requesting a camera zoom, and the implementation of that request at a remote surveillance site. In this case, control lag would measure the time delay required to actually implement the request on the remote end, but the controller would not actually see the change in the video scene until the new video scene was transmitted back to the controller (i.e., after undergoing a one-way video delay). Thus, the two-way delay (i.e., round-trip delay) for this application would be the sum of the control lag and the one-way video delay. Control lag is highly dependent on the type and nature of the control being employed. When required, the manufacturer or service provider should be able to provide a specification for control lag.

3.3.3 Luma Image Size and Interlaced Versus Progressive Scan Type

The luma video signal is the black-and-white portion of the video picture, denoted as Y. The image size used to represent the luma video signal is an important video system parameter since this will limit the perception of resolution. Luma image sizes will be specified as <pixels horizontally> by <frame lines vertically>.

Luma image size should be considered as an upper limit of usefulness for what *could* be achieved by an optimal video system. For example, just because a video system *could* produce, say, 352 by 288 useful pixels, does not mean that it actually *does*. After coding and decoding, the user might only see the equivalent of 176 by 144 pixels! As another example, High-Definition TV (HDTV) monitors commonly display an image size that is less than 1920 by 1080 (e.g., video might be displayed at only 1366 by 768). In light of the above discussion, additional methods for quantifying resolution and video fidelity need to be specified. (See Section 3.4, Section 3.5, and Section 3.6.)

Scan type specifies whether the image is interlaced (i) scan, or progressive (p) scan. With interlaced scan, the video frame consists of two interlaced fields; one field contains the even numbered lines and the other field contains the odd numbered lines. Fields are updated sequentially one field at a time. With progressive scan, the entire video frame is updated at the same time. Interlaced scan can produce smoother looking motion for a given frame rate (see Section 3.3.6), since the individual fields are updated at twice the frame rate. Progressive scan has advantages for reading text and displaying images on computer monitors.

Common luma image sizes and scan types are listed below:

- HDTV, 1080i—1920 by 1080, HDTV, interlaced video.
- HDTV, 720p—1280 by 720, HDTV, progressive video.
- NTSC, 525i—525-line interlaced video (horizontal image size depends upon the sampling rate). The National Television Systems Committee (NTSC), U.S. standard definition television (of the 525 video lines, only 486 contain picture information). Note that there are various flavors of analog NTSC video, including composite (lowest quality), s-video (higher quality), and component (highest quality).

- ITU-R Recommendation BT.601 [13] (i.e., Rec. 601)—720 by 486 (for 525i), interlaced video. The studio quality sampling format for standard definition television signals. The electrical signal used to transport the digitized video is commonly known as SDI (Serial Digital Interface), and is defined by the Society of Motion Picture and Television Engineers (SMPTE) 259M [30].
- VGA—640 by 480, Video Graphics Array (VGA), a progressive square-pixel video format used by computer monitors.
- CIF—352 by 288, Common Intermediate Format (CIF), progressive video. Used by video conferencing equipment.
- SIF—360 by 240, Source Input Format (SIF), progressive video. Used to encode Rec. 601 video at 1/4 resolution.
- QVGA—320 by 240, Quarter VGA (QVGA), progressive video. Used by personal digital assistants (PDAs).
- QCIF—176 by 144, Quarter CIF (QCIF), progressive video. Used by low-resolution video conferencing equipment.
- QSIF—180 by 120, Quarter SIF (QSIF), progressive video.

The luma image size and scan type of a video system are defined by the manufacturer and by service provider specifications.

3.3.4 Chroma Sub-Sampling Factors

For video signals that contain color information, it is common to sample the chroma signals (e.g., the blue and red chroma signals, denoted as C_B and C_R) at a lower rate than the luma signal (i.e., the luma signal, denoted as Y). The reason is the human visual system is not as sensitive to this color information. For instance, the chroma signals are normally sub-sampled by a factor of two in the horizontal direction or sub-sampled by a factor of two in both the horizontal and vertical directions—with only minimal impact on perceived quality. The chroma sub-sampling factors of a video system will be specified as <horizontal factor> by <vertical factor>. Hence, the chroma images (C_B and C_R) will be smaller than the luma image (Y) by the chroma sub-sampling factors. For example, Rec. 601 video has chroma sub-sampling factors of 2 by 1. This is because the C_B and C_R signals are sub-sampled by a factor of 2 horizontally and 1 vertically (i.e., no vertical sub-sampling) with respect to the Y signal.

The chroma sub-sampling factors of a video system are defined by the manufacturer and service provider specifications.

3.3.5 Aspect Ratio

Correct aspect ratio should be maintained when the video is passed through multiple pieces of equipment (i.e., acquisition, transmission, and display). Aspect ratio for a video picture is defined as the ratio of the displayed image width divided by the displayed image height, expressed as <horizontal width> : <vertical height>. Common aspect ratios are 4:3 (NTSC) and 16:9 (HDTV). Use of a simple video scene that contains a square can verify that aspect ratio is maintained throughout the video system (i.e., by measuring and dividing the horizontal and vertical sides of the square).

3.3.6 Frame Rate

Frame Rate (FR) is the rate at which a video system can produce unique consecutive images called frames. FR is measured in frames per second (fps). For example, an NTSC video system can display 59.94 interlaced fields per second, so the FR of this system is 59.94/2, which is approximately 30 fps. (Since NTSC is an interlaced scan system, half of the picture is updated every 1/59.94 seconds. Thus, 2/59.94 seconds are required to update the entire frame (see Section 3.3.3).

Unlike NTSC, where FR is a fixed characteristic of the video standard, the FR of many new video systems can be independently specified to achieve the desired motion rendition. Some video codecs (i.e., coder-decoder pairs as shown in Figure 3) have even adopted an adaptive approach to FR, particularly when the transmission bandwidth is fixed or constrained. Thus, when the scene contains still or nearly still video, a high FR can be used (e.g., 30 fps). However, as scene complexity increases (e.g., lots of motion and detail), the video codec drops back to a slower FR (e.g., 10 fps). This results in an improvement in the user's overall perception of quality.

In this document, FR will always refer to instantaneous FR. In other words, a specification for a minimum frame rate of 10 fps means that the video system must take no longer than 100 ms to present a new video frame. ITU-T Recommendation P.931 [17] contains the recommended standard method of measurement for measuring the FR of the video transmission subsystem (i.e., from points B to E in Figure 3). To obtain the effective FR for the whole video system (from points A to F in Figure 3), one must compare this FR to the FRs of the video acquisition and video display subsystems, and use the minimum FR over all three subsystems. The FRs of the video acquisition and video display subsystems are normally constant (i.e., not time varying) and specified by the manufacturers of the equipment.

3.3.7 Acceptability Threshold

The acceptability threshold for a video system is defined as the lower bound on the probability, with 95 percent confidence, of obtaining acceptable video clips for a given public safety video application. The acceptability threshold is measured by conducting controlled, subjective evaluations of video clips using viewer panels of public safety practitioners. (See the "Video Acquisition Measurement Methods," technical report [33] for an example procedure.) These subjective assessments should be conducted in accordance with ITU-R Recommendation BT.500 [12].

3.4 Video Acquisition Parameters

Video acquisition parameters measure the performance of the video acquisition subsystem (see Figure 3 and Figure 4), which reflects the creation of the video imagery itself. For some systems, the camera and coder are distinct, and thus video acquisition spans from point A to point B. For other systems, the camera performs coding, and thus video acquisition spans from point A to point C—that is, the camera and coder cannot be separated.

Wherever possible, existing video acquisition performance metrics that industry commonly uses have been specified. However, some public safety applications present unique video acquisition requirements that may require development of new performance metrics. Thus, this section is likely to evolve over time as additional public safety applications are examined.

The video acquisition parameters are identified as either *Primary* or *Secondary*. Primary parameters (listed first) are more important. They should be specified for most public safety applications that require a video

acquisition subsystem. Secondary parameters can be specified when they are applicable to meeting a specialized public safety requirement.

3.4.1 Resolution

(Primary Video Acquisition Parameter; method of measurement in Section 4.1 of the "Video Acquisition Measurement Methods" [33] report)

Resolution, as in sharpness, is the ability to resolve fine spatial detail in the video picture. Resolution will be quantified by MTF50P, the spatial frequency where the modulation transfer function, i.e., contrast, drops to 50 percent of its peak value. MTF50P is measured by analyzing near-vertical and near-horizontal edge responses of the video camera to the International Standards Organization (ISO) 12233 test chart [10] (e.g., Figure 1 in the "Video Acquisition Measurement Methods" report) under standard lighting conditions. MTF50P is then converted into line widths per picture height (LW per PH) to produce a measure of total image resolution.

3.4.2 Noise

(Primary Video Acquisition Parameter; method of measurement in Section 4.2 of the "Video Acquisition Measurement Methods" [33] report)

Noise is the unwanted random spatial and temporal variations (e.g., "snow") in the video picture. One method of measuring noise is to capture and analyze images of the Kodak Q-14 test chart (e.g., the top strip chart in Figure 2 in the "Video Acquisition Measurement Methods" report). The Q-14 test chart consists of 20 patches with densities from 0.05 to 1.95 in steps of 0.1. Noise and signal-to-noise ratio (SNR) can be measured for each patch. SNR tends to be worst in the darkest patches. Several lighting conditions with various intensities (e.g., standard, reduced, dim) and color temperatures (e.g., tungsten, daylight) may be required to adequately characterize noise. Noise can be measured using a similar approach with the GretagMacbeth ColorChecker (the bottom checkerboard chart in Figure 2 in the "Video Acquisition Measurement Methods" report). It is a standard color chart consisting of 24 patches: 18 color and 6 grayscale. Noise and SNR are used in the calculation of dynamic range, as described in Section 3.4.3.

3.4.3 Dynamic Range

(Primary Video Acquisition Parameter; method of measurement in Section 4.3 of the "Video Acquisition Measurement Methods" [33] report)

Dynamic range (DR) is the range of luminance levels (from lowest to highest) that can be captured with reasonable quality, without clipping, by the video acquisition system. Two methods will be presented to measure dynamic range: an indirect method that infers or extrapolates dynamic range using a Kodak Q-14 reflection test chart (e.g., Figure 2 in the "Video Acquisition Measurement Methods" report, top strip), and a direct method that uses transmission test charts (e.g., Figure 16 in the "Video Acquisition Measurement Methods" report).

3.4.4 Color Accuracy

(Primary Video Acquisition Parameter; method of measurement in Section 4.4 of the "Video Acquisition Measurement Methods" [33] report)

Color accuracy is the ability to reproduce colors with minimal chromatic distortion so that they are as close to real-life as possible, given the color-space limitations of the video standard being used (e.g., NTSC, Advanced Television Systems Committee (ATSC)). A GretagMacbeth ColorChecker test chart (the bottom checkerboard chart in Figure 2 in the "Video Acquisition Measurement Methods" report) is used to measure color accuracy. Several lighting conditions with various intensities (e.g., standard, reduced, dim) and color temperatures (e.g., tungsten, daylight) may be required to adequately characterize color accuracy.

3.4.5 Capture Gamma

(Secondary Video Acquisition Parameter; method of measurement in Section 4.5 of the "Video Acquisition Measurement Methods" [33] report)

Capture Gamma is a measure of camera contrast. It is the average slope of the equation that relates scene luminance to image pixel level, approximately, log(pixel level) = (Capture Gamma) × log(luminance). For image files intended for display on devices with display gamma = 2.2, Capture Gamma should be approximately 0.45.

3.4.6 Exposure Accuracy

(Secondary Video Acquisition Parameter; method of measurement in Section 4.6 of the "Video Acquisition Measurement Methods" [33] report)

Exposure accuracy is the ability of the video acquisition system to properly match the grayscale tonal levels of the scene being shot. All video cameras can be set for automatic exposure, which controls the shutter speed, lens aperture, and gain combination that is used to achieve proper exposure. Some video cameras may have a manual override that lets the user select the aperture. It is anticipated that most public safety applications will use video acquisition systems with automatic exposure. Several lighting conditions with various intensities (e.g., standard, reduced, dim) and color temperatures (e.g., tungsten, daylight) may be required to adequately characterize exposure accuracy for these systems. Exposure accuracy is measured by photographing the GretagMacbeth ColorChecker or Q-14 test charts (Figure 2 in the "Video Acquisition Measurement Methods" report) against a gray background (which affects the automatic exposure setting), and comparing pixel levels for a range of gray patches from light to dark gray with standard values. Exposure accuracy may be affected by illumination history: it may change following exposure to bright light.

3.4.7 Vignetting

(Secondary Video Acquisition Parameter; method of measurement in Section 4.7 of the "Video Acquisition Measurement Methods" [33] report)

Vignetting, which identifies light falloff and uniformity, is the reduction in image brightness at the edges of the image versus the center of the image. Illumination in many inexpensive optical systems is not uniform: it decreases with distance from the image center. Some camera modules compensate for this digitally. Vignetting can interfere with other performance parameter measurements, particularly those that use charts consisting of patches of known density or color that span a good portion of the image. For instance, it may be necessary to compensate for vignetting to obtain a valid dynamic range measurement.

3.4.8 Lens Distortion

(Secondary Video Acquisition Parameter; method of measurement in Section 4.8 of the "Video Acquisition Measurement Methods" [33] report)

Barrel and *pincushion* are terms that describe two types of lens distortion. Barrel distortion is a lens distortion that produces greater magnification at the center of the image versus the edges of the image, resulting in rectilinear grid lines (see Figure 3 in the "Video Acquisition Measurement Methods" report) that are bowed outward. Pincushion distortion is a lens distortion that produces less magnification at the center of the image versus the edges of the image, resulting in rectilinear grid lines (see Figure 3 in the "Video Acquisition Measurement Methods" report) that are bowed inward. Lens distortion can be measured from an image of a square or rectangular grid or from a simple rectangle near the image margins.

3.4.9 Reduced Light and Dim Light Measurements

(Secondary Video Acquisition Parameter; method of measurement in Section 4.9 of the "Video Acquisition Measurement Methods" [33] report)

Unlike still cameras, long exposures are not an option with video cameras. Thus, whereas a still camera can create a high quality photograph (e.g., noise free) by increasing the exposure time, the maximum exposure time for a video camera is the reciprocal of the video frame rate. Hence video performance at low light levels (e.g., noise) cannot be inferred from measurements at high light levels.

3.4.10 Flare Light Distortion (Under Study)

(Secondary Video Acquisition Parameter; method of measurement under study in Section 4.10 of the "Video Acquisition Measurement Methods" [33] report)

Flare light distortion is caused by light that bounces between lens elements and off the interior barrel of the lens. It reduces the usable dynamic range of the camera system under adverse lighting conditions, such as might occur in a night-time police traffic stop where spotlights are used.

3.5 Video Transmission Parameters

The video transmission subsystem includes everything that occurs after the video has been rendered by the camera until just before the video is displayed on the monitor (from point B to point E in Figure 3). Video transmission can be significantly more complex than what is depicted in Figure 3. Instead of a simple coder/network/decoder, the video transmission subsystem could route the video through an H.264 [25] video coder, over a congested IP network, through an H.264 decoder, through an MPEG-2 [11] coder to a DVD recorder, to be decoded with a PC's DVD player and rendered by a computer that pauses the playback occasionally.

In most cases, the video transmission subsystem is a major contributor to video impairments.

3.5.1 Parameters for Measuring Calibration Errors

In addition to introducing video delay, a video transmission system may introduce other fixed distortions due to improper calibration. These calibration errors include spatial scaling of the picture (both horizontal and vertical), spatial shifts of the picture (both horizontal and vertical), a reduction of the picture area (valid region), and changes in gain (contrast) and level offset (brightness) of the video signal. For some

applications, distortions to the video signal that result from calibration errors may not be important. For other applications (e.g., telemedicine), very accurate system calibration may be required.

This section describes a set of parameters that may be used to objectively measure the proper calibration of video transmission systems. One application of these measurements is to "tune" and correct potential calibration problems *before* the video transmission system is deployed. Ideally, network errors (see Section 3.5.3) should not be present when conducting the calibration measurements discussed in this section, as they may adversely affect the measurement accuracy.

3.5.1.1 Gain

Gain is a multiplicative scaling factor that has been applied to all pixels of an individual image plane (e.g., Y, C_B, and C_R) by the video transmission subsystem. Gain of the luma signal (Y) is commonly known as contrast. The ideal gain of the video transmission subsystem is 1.0 (i.e., no multiplicative scaling). The recommended method of measurement for gain is given in ANSI T1.801.03-2003.

3.5.1.2 Level Offset

Level offset is an additive factor that has been applied to all pixels of an individual image plane (e.g., Y, C_B, and C_R) by the video transmission subsystem. Level offset of the luma signal (Y) is commonly known as brightness. The ideal level offset of the video transmission subsystem is 0 (i.e., no additive shift). The recommended method of measurement for level offset is given in ANSI T1.801.03-2003.

3.5.1.3 Valid Region

Valid region is the rectangular portion of the image that is not blanked or corrupted by the video transmission subsystem. The valid region only includes those image pixels that contain usable picture information. Ideally, the valid region size is equal to the luma image size for the video standard being used (see Section 3.3.3) The recommended method of measurement for valid region is given in ANSI T1.801.03-2003.

3.5.1.4 Spatial Shift

Spatial shift is the shift of the image in the horizontal and/or vertical directions, measured in pixels. Spatial shift is considered positive when the video images exiting the video transmission subsystem are shifted to the right or down with respect to the video images entering the video transmission subsystem. The ideal spatial shift of the video transmission subsystem is zero (i.e., no spatial shift). The recommended method of measurement for spatial shift is given in ANSI T1.801.03-2003.

3.5.1.5 Spatial Scaling

Spatial scaling is an expansion or shrinkage of the image in the horizontal and/or vertical directions. The ideal spatial scaling of the video transmission subsystem is zero (i.e., no spatial scaling). If spatial scaling is present, the aspect ratio (Section 3.3.5) will be affected. The recommended method of measurement for spatial scaling is under study.

3.5.2 Parameters for Measuring Coding/Decoding Impairments

Coding and decoding are a reality of today's digital video systems. Coding entails compression, which enables a video service to be transmitted using a bandwidth that cannot accommodate the non-compressed video signal. The flip side of this coin is that some of the quality of the video signal may be lost. There is a

tendency to demand or require "uncompressed" video or completely lossless coding for particularly important applications. However, in light of recent advances in video coding technology, the requirement for lossless coding should be carefully examined. Consider that most high-definition (HD) video cameras output a digital signal that has been compressed *with some loss*. Likewise, all HD recording medium in common use perform some degree of *lossy* compression. The resulting video is nearly perfect, but it cannot be called "uncompressed". Very high-quality codecs can reduce the transmission and storage bandwidths dramatically while causing only a minute drop in the perceived video quality. On the other end of the spectrum, some types of coding losses might be unacceptable, or some types of coding loss might be acceptable only because it enables a new video service to be available that otherwise would not be.

Motion and spatial detail jointly determine the compressibility of video scenes, so this is a key detail that must be considered in any method of measurement for quantifying coding/decoding impairments. Thus, the performance measurements in this section use actual video scenes content. This scene content should be selected to span the full range of motion and spatial detail that is required for the given public safety application.

3.5.2.1 Lossless Impairment

Lossless video transmission means that the video stream entering the video transmission subsystem (i.e., point B in Figure 3) is bit-identical to the video stream leaving the video transmission subsystem (i.e., point E in Figure 3). Peak-Signal-to-Noise-Ratio (PSNR), as defined in the Alliance for Telecommunications Industry Solutions (ATIS) Technical Report T1.TR.74-2001 [2], is the recommended method of measurement for measuring lossless impairment. For the video transmission system to be truly lossless, the measured noise must be zero. This will produce an infinite PSNR, as you will be dividing by zero. Lossless impairment, as specified by PSNR, can also be effective for quantifying minor impairments to the uncompressed video stream (i.e., these impairments cannot include lossy video compression—see Section 3.5.2.2).

3.5.2.2 Lossy Impairment

Lossy video transmission means that the video stream leaving the video transmission subsystem (i.e., point E in Figure 3) has undergone lossy compression when compared to the video stream entering the video transmission subsystem (i.e., point B in Figure 3). The General Video Quality Model (henceforth abbreviated as VQM_G), as standardized by ANSI T1.801.03-2003[3] [1], will be used to specify the amount of perceptual lossy impairment. The recommended method of measurement is as follows:

1. Select at least eight video scenes with durations of 8 to 12 seconds each that span the range of scene content for the public safety application being deployed (e.g., tactical video—see Section 4.3). These scenes should be of the highest possible quality (uncompressed recording formats are recommended). Further, each scene should contain content that is substantially different from the other scenes being used. Scene characteristics to consider include a range of spatial detail, motion, color, and lighting levels. Development of a standard set of scenes to use for this purpose is under study.
2. Inject the scenes from step 1 into the video transmission subsystem (i.e., point B in Figure 3) and record the scenes from the output of the video transmission subsystem (i.e., point E in Figure 3). The recording of the output scenes should be of the highest possible quality—uncompressed recording formats are recommended.

3. VQM_G has also been internationally standardized by ITU-T as Recommendation J.144 [23] and by ITU-R as Recommendation BT.1683 [14].

3. Compute VQM_G for each pair of source and destination video streams from steps 1 and 2, respectively. The nominal range of VQM_G values is from 0 (i.e., no perceptual impairment) to 1 (i.e., maximum perceived impairment).[4]
4. The lossy impairment is computed as the average of the VQM_G values from all scene pairs.

3.5.3 Parameters for Measuring Impact of Network Impairments

Impairments present in the network (from point C to point D in Figure 3) can significantly affect the perceived quality of the video service (point F in Figure 3). Network impairments can impact video quality in many ways including brief appearances of false image blocks and/or strips, a frozen image that resumes with a loss of content (i.e., skip), a frozen image followed by fast forwarding through the missing content, a sudden drop in image resolution, the image replaced with a blank screen, etc.

This section provides a set of network performance parameters that are known to have a potential impact on video quality. Specification of acceptable levels for these network parameters can only be made after the video coder and decoder are selected. This is because the efficiency of the video coding algorithm, the use of error concealment by the decoder, and the use of forward error correction (FEC) and/or retransmission methods by the coder/decoder pair, all influence the impact that network impairments have on the final perceived quality. Thus, specification of these network parameters *must always* be associated with a specific coder-decoder (codec) pair configuration.

3.5.3.1 Coder Bit Rate

Coder bit rate is the amount of information (in bits per second) output by the video coder to the network (point C in Figure 3), excluding all transport and protocol overhead and retransmissions. For the purposes of this document, coder bit rate will always refer to the minimum instantaneous bit rate. Coder bit rate is specified by the manufacturers of the coder equipment.

3.5.3.2 Packet Loss Ratio

Packet loss ratio (PLR) is the fraction expressed in percentage (from 0 to 100 percent) of packets lost by the network (from point C to point D in Figure 3). The recommended method of measurement for PLR is given by the Internet Engineering Task Force (IETF) RFC2680 [8].

3.5.3.3 Packet Size

Packet size (PS) is the size of an IP packet in octets, including all overhead as well as payload information. The influence of PS on video transmission quality is currently under study. One consideration is to force the network errors to occur at the same time slice for two identical video streams that have been encapsulated with different packet sizes. This would allow direct comparisons of the effects of PS on video quality. Another consideration is the decreasing goodput (e.g., useful application information that is transmitted by the network) that might be available to the video coder when packet sizes are reduced. A third consideration is how the video decoder will be affected by losing information in different locations of the coded video stream, and how this might affect error concealment algorithms, if they are present.

4. For extremely impaired video sequences, VQM_G can produce values greater than 1.0 but this is not common.

3.6 Video Display Parameters

Video display parameters measure the performance of the video display subsystem (see Figure 3 and Figure 4), which reflects the presentation of the video imagery to the user. For some systems, the display and video decoder are distinct, and thus the video display spans from point E to point F. For other systems, these functions are inseparable, and thus the video display spans from point D to point F. Wherever possible, existing video display performance metrics that are commonly used by industry should be specified. However, some public safety applications present unique video display requirements that may require development of new performance metrics. Thus, this section is likely to evolve over time as additional public safety applications are examined.

Video display parameters are under study.

This page intentionally left blank.

4 Video Performance Requirements

This section provides requirements for public safety video applications as follows:

- General public safety video requirements (Section 4.2)
- Tactical and live surveillance video requirements (Section 4.3)
- Recorded surveillance video requirements (Section 4.4)

4.1 Target Size and Scene Complexity

"Target" refers to something the viewer hopes to discern or identify in a video scene. Depending on the situation or application, the target of interest to the viewer will be small or large. This distinction affects performance requirements for video quality. Testing illustrated a marked difference in the performance requirements for targets that occupy fewer than 0.3 percent of the total pixels in the video frame, and those that are larger. Therefore, the requirements for the video applications described here are specified for two cases, which appear in Table 6 and Table 7:

- Small target (less than 0.3 percent of total pixels)
- Large target (greater than 0.3 percent of total pixels)

The requirements in Table 6 and Table 7 for small and large targets are further broken down by scene complexity (low and high). "Complexity" applies to the effect of the homogeneity of the spatial and temporal video information on the parameters under study. The study [37] behind the requirements in Table 6 and Table 7 addresses compression, so in this case scene variance is the "complexity." A different example might be camera optics. In that case, the dynamic range of color and light levels may constitute the "complexity" of the scene.

Compression complexity is a function of motion and small details in the video content. The complexity score is a measure of object motion, diverse clutter, and spatial frequency information in a video scene or clip. The complexity score indicates how much processing the video CODEC will have to perform to encode and decode the video clip. A high complexity rating means that a scene or clip contains multiple independent motions, clutter, and higher spatial frequency information.

4.2 General Public Safety Video Requirements

Table 5 specifies performance requirement values applicable to tactical and live or recorded surveillance video applications, for the following video systems:

- Entire video system (Section 3.3)
- Video acquisition subsystem (Section 3.4)
- Video transmission parameters (Section 3.5)
- Video display subsystem (Section 3.6).

Requirement values for video performance parameters were obtained by conducting a survey and controlled subjective evaluations of video clips, using viewer panels of public safety practitioners (see the "Video Acquisition Measurement Methods" [33] report).

Table 5: General Public Safety Video Performance Requirements

Parameters	Performance Requirements
Video System (Section 3.3)	**Values**
One-Way Video Delay (Section 3.3.1)	Maximum of 1 Second
Control Lag (Section 3.3.2)	Not Specified
Luma Image Size and Scan Type (Section 3.3.3)	Minimum of 352 by 240, Progressive Scan[a]
Chroma Sub-Sampling Factors (Section 3.3.4)	Maximum of 2 by 2
Aspect Ratio (Section 3.3.5)	Not Specified
Frame Rate (Section 3.3.6)	Minimum of 10 fps
Acceptability Threshold (Section 3.3.7)	Minimum of 0.7
Video Acquisition (Section 3.4)	**Values**
Resolution (Section 3.4.1)	Under Study
Noise (Section 3.4.2)	Under Study
Dynamic Range (Section 3.4.3)	Under Study
Color Accuracy (Section 3.4.4)	Under Study
Capture Gamma (Section 3.4.5)	Not Specified
Exposure Accuracy (Section 3.4.6)	Not Specified
Vignetting (Section 3.4.7)	Not Specified
Lens Distortion (Section 3.4.8)	Not Specified
Reduced Light and Dim Light Measurements (Section 3.4.9)	Not Specified
Flare Light Distortion (Section 3.4.10)	Not Specified
Video Display (Section 3.6)	**Under Study**

a. Before being displayed, this image size must be up-sampled by a factor of 2 in both the horizontal and vertical directions using an up-sampling process that uses pixel interpolation.

4.3 Tactical and Live Surveillance Video Requirements

Tactical video is used in real time during an incident by public safety personnel to make decisions on how to respond to that incident. Live surveillance video is used in real time to make decisions about whether a situation requires a response.

4.3.1 Example Tactical Video Scenarios

Example tactical video scenarios include:

- Video used to provide the incident commander with situation information, such as 1) a camera carried by a public safety practitioner, looking for victims, into a burning building; 2) a body-worn camera during a SWAT raid; and 3) an aerial camera following a suspect on foot.
- Close-up videography from a camera on a robot being used to dismantle a bomb.
- Aerial videography used during wildfire suppression.
- Aerial video to pursue an automobile.

4.3.2 Example Live Surveillance Video Scenarios

Example live surveillance video scenarios include:

- A sweep of the whole incident scene to aid decision-making on how to deploy personnel.
- Monitoring an entry way for suspicious activity.
- Cameras covering a football stadium that can pan and zoom, which allow a remote police officer to look for disturbances.
- Cameras in a federal judge's residence that is under protective surveillance.
- Cameras in a bank lobby that can show where people are and their activities.
- Videography at a stakeout of a suspected drug house.
- An aerial camera used to perform a sweep of a large public event.
- Automated analysis that can usually discern the license plate number of a car driving at high speed.

4.3.3 Tactical and Live Surveillance Video Transmission Requirements

This section identifies a set of performance measurements for live tactical and surveillance video systems. The performance parameters and their recommended values are listed in Table 6 according to those that apply to the video transmission subsystem (Section 3.5).

Table 6: Tactical and Live Surveillance Video Performance Requirements

Parameters	Performance Requirements
Video Transmission (Section 3.5)	**Values**
Gain (Section 3.5.1.1)	$0.95 \leq \text{Gain} \leq 1.05$

Table 6: Tactical and Live Surveillance Video Performance Requirements (Continued)

Parameters	Performance Requirements	
Level Offset (Section 3.5.1.2)	$-10 \leq \text{Level Offset} \leq 10$ (for video systems with 255 quantization levels for each image plane)	
Valid Region (Section 3.5.1.3)	Minimum 95 Percent of the Luma Image Size (horizontally and vertically)	
Spatial Shift (Section 3.5.1.4)	Maximum of 2 Pixels (horizontally and vertically)	
Spatial Scaling (Section 3.5.1.5)	Under Study (No spatial scaling is preferred)	
Lossless Impairment (Section 3.5.2.1)	Not Required	
Lossy Impairment (Section 3.5.2.2)	Maximum of 0.41	
Interaction of Minimum Bit Rate (BR) (Section 3.5.3.1) and a Maximum Packet Loss Ratio (PLR) (Section 3.5.3.2) with Error Concealment	**Small Target**	
	Low Complexity ■ PLR of 15 Percent with a BR of 512 Kbps	**High Complexity** ■ PLR of 10 Percent with a BR of 512 Kbps
	Large Target	
	Low Complexity ■ PLR of 5 Percent with a BR of 256 Kbps	**High Complexity** ■ PLR of 5 Percent with a BR of 256 Kbps
Interaction of Minimum Bit Rate (BR) (Section 3.5.3.1) and a Maximum Packet Loss Ratio (PLR) (Section 3.5.3.2) without Error Concealment	**Small Target**	
	Low Complexity ■ PLR of 0.1 Percent with a BR of 256 Kbps	**High Complexity** ■ PLR of 0 Percent (None) with a BR of 512 Kbps
	Large Target	
	Low Complexity ■ PLR of 0.1 Percent with a BR of 256 Kbps ■ PLR of 0.5 Percent with a BR of 512 Kbps	**High Complexity** ■ PLR of 0.1 Percent with a BR of 256 Kbps ■ PLR of 0.5 Percent with a BR of 512 Kbps
Packet Size (PS) (Section 3.5.3.3)	Under Study	

4.4 Recorded Surveillance Video Requirements

Recorded surveillance video is stored for later use and analysis.

4.4.1 Example Recorded Surveillance Video Scenarios

Example recorded surveillance video scenarios include:

- Documenting routine incidents such as an uneventful prisoner transfer or an ordinary traffic stop
- Video from a residence that is under protective surveillance
- Video from an ATM camera
- Video from a convenience store camera recording all people who pass through a door
- Documenting crime scene investigation
- Documenting felonies
- Documenting officer investigations
- Recording an unknown suspect with the intention of identifying them through facial recognition and movement patterns
- Discerning the license plate number of a moving car

4.4.2 Recorded Surveillance Video Transmission Requirements

This section identifies a set of performance measurements for recorded surveillance video systems. The performance parameters and their recommended values are listed in Table 7, according to those that apply to the video transmission subsystem (Section 3.5). All of the measurements are currently under study.

Table 7: Recorded Surveillance Video Performance Requirements

Parameters	Performance Requirements
Video Transmission (Section 3.5)	**Values**
Gain (Section 3.5.1.1)	Under Study
Level Offset (Section 3.5.1.2)	Under Study
Valid Region (Section 3.5.1.3)	Under Study
Spatial Shift (Section 3.5.1.4)	Under Study
Spatial Scaling (Section 3.5.1.5)	Under Study
Lossless Impairment (Section 3.5.2.1)	Under Study
Lossy Impairment (Section 3.5.2.2)	Under Study
Interaction of Minimum Bit Rate (BR) (Section 3.5.3.1) and a Maximum Packet Loss Ratio (PLR) (Section 3.5.3.2) with Error Concealment	Under Study

Table 7: Recorded Surveillance Video Performance Requirements (Continued)

<table>
<tr><th>Parameters</th><th colspan="2">Performance Requirements</th></tr>
<tr><td rowspan="4">Interaction of Minimum Bit Rate (BR) (Section 3.5.3.1) and a Maximum Packet Loss Ratio (PLR) (Section 3.5.3.2) without Error Concealment</td><td colspan="2">Small Target</td></tr>
<tr><td>Low Complexity
PLR of 0.1 Percent with a BR of 256 Kbps
PLR of 0.2 Percent with a BR of 384 Kbps
PLR of 0.5 Percent with a BR of 512 Kbps</td><td>High Complexity
PLR of 0.1 Percent with a BR of 256 Kbps
PLR of 0.2 Percent with a BR of 384 Kbps
PLR of 0.5 Percent with a BR of 512 Kbps</td></tr>
<tr><td colspan="2">Large Target</td></tr>
<tr><td>Low Complexity
PLR of 0.1 Percent with a BR of 256 Kbps
PLR of 0.2 Percent with a BR of 384 Kbps
PLR of 0.5 Percent with a BR of 512 Kbps</td><td>High Complexity
PLR of 0.1 Percent with a BR of 256 Kbps
PLR of 0.2 Percent with a BR of 384 Kbps
PLR of 0.5 Percent with a BR of 512 Kbps</td></tr>
<tr><td>Packet Size (PS) (Section 3.5.3.3)</td><td colspan="2">Under Study</td></tr>
</table>

4.4.3 Feature Recommendations for Forensic Video Analysis

Forensic Video Analysis (FVA) is the scientific examination, comparison, or evaluation of video in legal matters. If there is a possibility that tactical video will be used for FVA, give consideration to including the features listed in Table 8.

Table 8: Video Feature Recommendations for Forensic Video Analysis

Feature	Description
FVA Support	Like recorded surveillance video systems, tactical video systems should record video to support FVA when necessary.
Export of Proprietary Compression to Standard Format	Video systems with proprietary compression algorithms should export uncompressed video to a standard format, such as Audio Video Interleave (AVI).
Export of Standard Compression	Video systems with standard compression algorithms, such as MPEG-2 or H.264, should export either compressed video in that standard format or uncompressed video to a standard format, such as AVI.
Export of Data to Standard Format	Video systems that associate data with the video, such as time, date, or Pin number used to open a secure door, should export that data to a standard format.
Atomic Clock Synchronization	Video systems that record time information should be automatically synchronized to standard time taken from a U.S. atomic clock.

Table 8: Video Feature Recommendations for Forensic Video Analysis (Continued)

Feature	Description
Automated Authentication	Video recording systems should have an automated authentication mechanism. Video authentication should be attached to the video sequence when the video is first recorded. Preferably, video recording equipment should use a digital video signature that the America Bar Association (ABA) has standardized approved.
Swappable Recording Medium	The video recording medium (e.g., hard drive) in the video system should be easily swappable without disabling the system. Thus, in the event of removal, an alternative recording medium should be available.

5 Reference Model for Network Performance

This section provides an understanding of the constraints imposed by the applications and the usage scenarios. Its purpose is to reveal insights into the resulting network performance trends.

5.1 Mission-Critical Network Services

A communications network permits the transmission of information from one location to another. The type of information transported depends on: the usage scenario such as the number and type of devices and their location, and the application considered, such as tactical video or speech. These choices impose many constraints on the communications network.

You can take at least two approaches to choosing a communications network. The first approach is to design a network customized to the type or types of information to be distributed (i.e., based on application requirements). The second approach is to use an existing network design and examine how to accommodate the requirements of a specific application. However, regardless of the network design approach you choose, the central question remains: how well does a network distribute the information generated by a particular application, or in other words, how well does a network meet the application's quality of service requirements? The requirements provided in this document offer useful planning information for network engineers building and maintaining communication networks for public safety.

This section describes a reference network model based upon a communication path within the System of Systems concept described in PS SoR Volume I [28]. This model is consistent with public safety communications networks that will be partitioned based on functional and jurisdictional boundaries and constraints (i.e., a system of systems), and whose primary objective is to transport information between public safety communications devices (PSCDs). The DHS technical report, "Network Measurement Methods," [34] lists the results obtained using the network model described here.

A network, or system, consists of a group of nodes corresponding to individual communications devices, and links that connect the nodes to each other. The arrangement, configuration, or topology of nodes and connecting links is based on many variables. Using any specific physical or geographical representation of a network topology would provide only one example topology out of an almost endless set of possible network configurations. To keep our performance analysis tractable, therefore, we use a path-based reference model. This considers the path that information will traverse through the network between a pair of PSCDs.

This network performance reference model considers all possible paths, using the hierarchy network model presented across the various area networks. Every path will not be applicable in all situations. Some paths might seem very unlikely to occur. However, from an engineering perspective, it is important to consider the performance requirements of a range of possible scenarios to support all necessary cases, even if some are unlikely.

This section provides a definition of the path model first, followed by a discussion of the parameters considered in the model. These include the function, capacity, number, and performance characteristics of the nodes as well as the connecting links that directly influence the path characteristics.

5.2 Path Model Definition

A path is the set of links and nodes that the information traverses from the originating PSCD to the destination PSCD. Links in the path represent physical cables that connect pairs of nodes (e.g., fiber optic lines linking high-speed routers), or they can correspond to point-to-point or one-to-many radio links in the case of wireless networks.

Generally, if we consider a path of interest, it has two types of nodes: end nodes and transit nodes. The end nodes are PSCDs and represent the origination or termination points for the connection. Transit nodes, also known as intermediate nodes, provide for the distribution, or routing, of the information stream carried by the connection.

In addition to the links and nodes along any given path, area network boundaries are defined by the administrative, jurisdictional, and coverage areas of various network segments that constitute the larger network. Figure 7 and Figure 8 illustrate this path-based model in the context of the PS SoR Volume I [28] hierarchical reference network illustration (Figure 5) and link-based description (Figure 6).

Figure 5: Natural Network Hierarchy

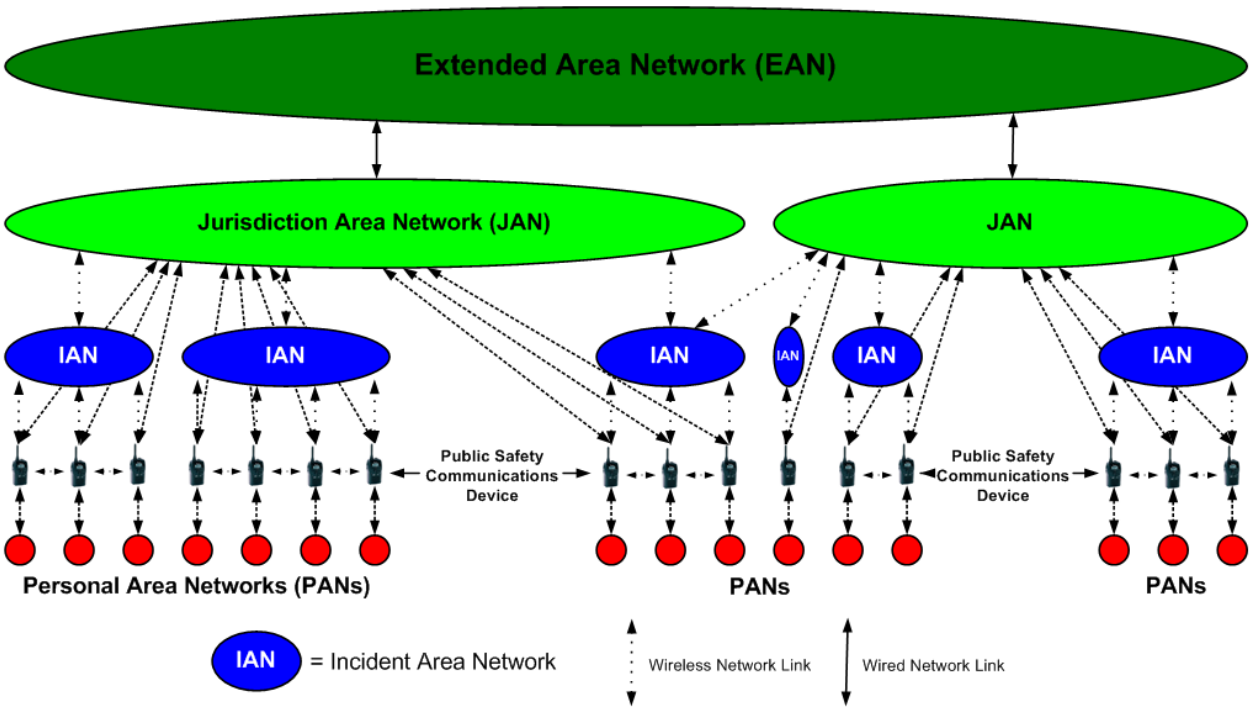


Figure 6: Link Diagram

For this study, we develop two sets of path models we refer to as Model A and Model B, using the basic network architecture from PS SoR Volume I [28]. Model A assumes the connection follows a strictly hierarchical path from the source PSCD through progressively larger area networks to an inflection point, after which the connection runs back down the hierarchy to the destination PSCD. The longest path allowed by Model A runs from the source PSCD on a personal area network (PAN), to the local Incident Area Network (IAN), then to the local Jurisdiction Area Network (JAN), next to the Extended Area Network (EAN), then to the destination JAN, destination IAN, and finally the destination PSCD and PAN.

Model B assumes that the hierarchical path includes peer-level communication links, i.e., from one IAN to another IAN.

In Model A, the communication path takes the least number of levels and the least number of links between the originating and terminating PSCDs. Two distinct sets of hierarchical paths exist. The first set contains all the symmetrical reference paths where the sequence of area network types traversed from the source PSCD to the inflection point is the reverse of the sequence of area network types traversed from the inflection point to the destination PSCD. An example is PAN-IAN-JAN-IAN-PAN. The second set contains all the asymmetrical reference paths, where the sequence of area network types traversed from the source PSCD to the inflection point is not the reverse of the sequence of area network types traversed from the inflection point to the destination PSCD. An example is PAN-IAN-JAN-PAN.

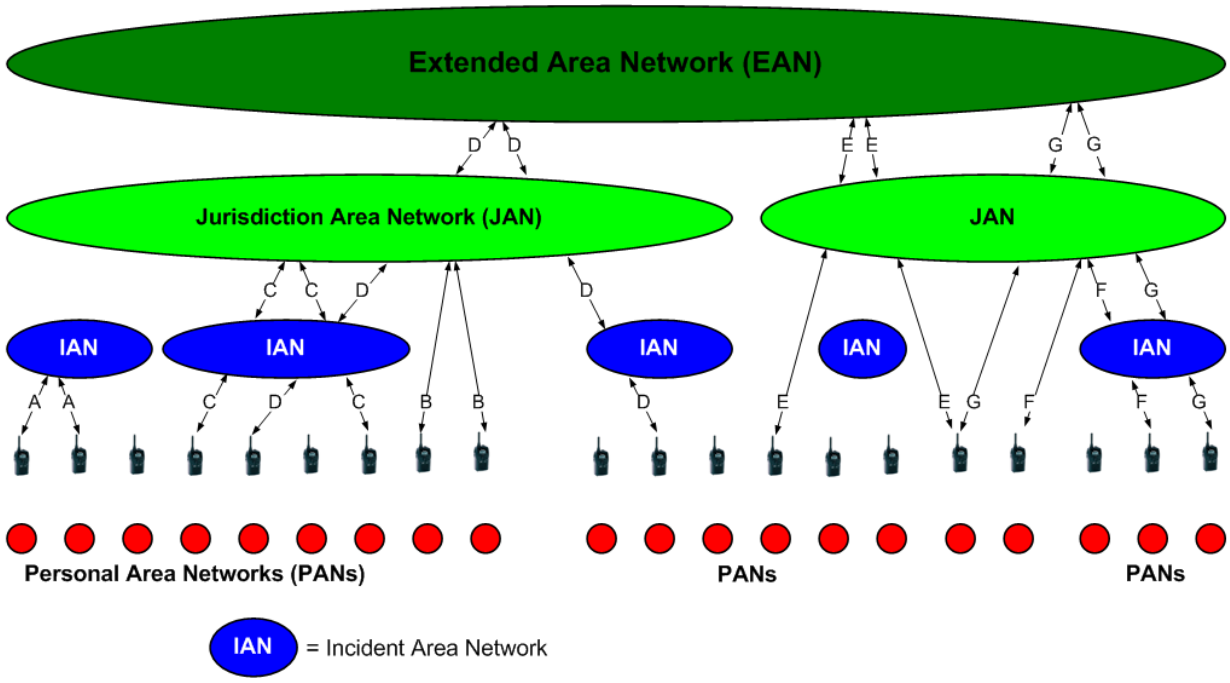
Figure 7 illustrates the following symmetrical and asymmetrical path types in Table 9.

Table 9: Symmetrical and Asymmetrical Network Path Types

Path	Description
A Symmetrical	Public safety communications device (PSCD) via first responder's vehicle (FRV) to PSCD (PAN-IAN-PAN)[a] (involves one IAN; two wireless links)
B Symmetrical	PSCD via jurisdiction communication tower to PSCD (PAN-JAN-PAN) (involves one JAN; two wireless links)
C Symmetrical	PSCD via FRV to jurisdiction communication tower to FRV to PSCD (PAN-IAN-JAN-IAN-PAN) (involves two IANs and one JAN; four wireless links)
D Symmetrical	PSCD via FRV to jurisdiction communication tower to EAN to another jurisdiction communication tower to FRV to PSCD (PAN-IAN-JAN-EAN-JAN-IAN-PAN) (involves two IANs, two JANs, and one EAN; six wireless links)
E Symmetrical	PSCD to jurisdiction communication tower to EAN to another jurisdiction communication tower to PSCD (PAN-JAN-EAN-JAN-PAN) (involves two JANs and one EAN; four wireless links)
F Asymmetrical	PSCD via FRV to jurisdiction communication tower to PSCD (PAN-IAN-JAN-PAN) (involves one IAN and one JAN; three wireless links)
G Asymmetrical	PSCD via FRV to jurisdiction communication tower to EAN to another jurisdiction communication tower to PSCD (PAN-IAN-JAN-EAN-JAN-PAN) (involves one IAN, two JANs and one EAN; five wireless links)

a. Public safety network types: extended area network (EAN), incident area network (IAN), jurisdictional area network (JAN), personal area network (PAN).

Figure 7: Hierarchical Reference Paths Based on "Natural Network Hierarchy"



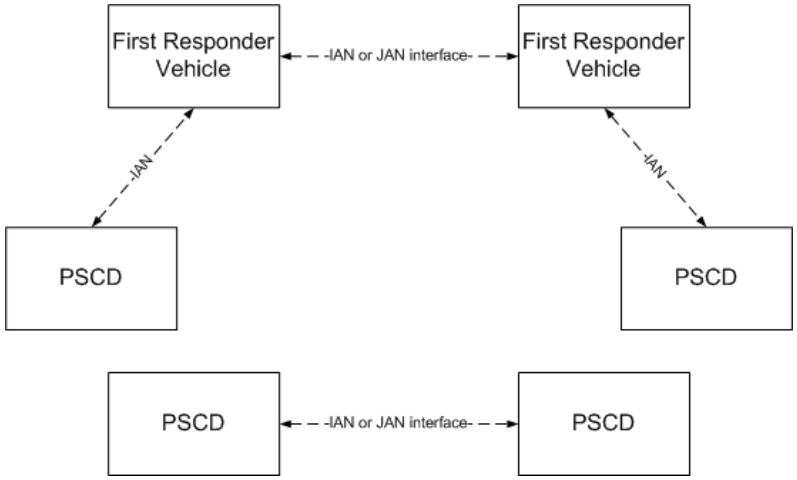
In Model B, the communication paths include peer-to-peer communication links between PSCDs and FRVs (IAN) communications equipment. Figure 8 shows the following peer-to-peer path types:

- PSCD to PSCD

  (where the PSCD to PSCD uses either the IAN or JAN interface with no infrastructure)
- PSCD to FRV to FRV to PSCD

  (where the PSCD to FRV uses the IAN, and the FRV to FRV could use either the IAN or JAN interface)

Figure 8: Peer Reference Paths (by Links) Based on "Network Diagram Link Descriptions"



5.3 Path Model Parameters

The parameters characterizing the path are the links, nodes, and area networks that make up the path and directly influence the path characteristics. The links are characterized by the raw data rate available, and the signal propagation time through the different media types. Nodes are characterized by the average time to queue and process a packet and the protocols used from the ingress point to the egress point, including those used to access the transmission on the link. Area networks describe network segments composed of nodes that the path traverses.

This section identifies some of the most important parameters affecting path performance.

5.3.1 Medium Access Control

A Medium Access Control (MAC) protocol is designed to permit access to a shared medium on a fair and equitable basis. The design of a MAC will consider characteristics of the medium or link being shared, the number and arrangement of nodes accessing the medium, and application traffic characteristics. Many MAC protocols exist. For example, some well-known MACs are those defined in IEEE Std. 802.3 [5] (commonly known as Ethernet) and IEEE Std. 802.11 (WiFi) [3].

Our discussion of network performance is technology-neutral, so the PS SoR considers only the generic functions and features of a MAC. A MAC's two important functions are regulating the transmission of packets on the shared medium, and dealing with loss in case packets do not make it through.

The two general categories of MACs are time division multiple access (TDMA) and carrier sensed multiple access (CSMA). A TDMA MAC divides the shared medium so that, once access is granted to a node, it is guaranteed access until the communication session is completed. Thus, the contending method is on a session-by-session basis to gain the initial granting of access for the communication. From the viewpoint of the node, the session is either a success (was granted access to the medium) or a failure (was denied access to the medium). This contending method happens in the CSMA MAC, but on a packet-by-packet basis.

Many flavors of CSMA allow multiple transmitting stations to share a communications medium in an uncoordinated fashion. Slotted Aloha is one the simplest examples. Unlike other more sophisticated versions of CSMA, Slotted Aloha does not require users to monitor the channel to see if it is busy before they begin transmitting. Users must schedule packet transmissions using a common clock that segments time into regular slots such that the length of a slot is equal to the amount of time required to transmit a packet. Any slot can potentially be used by any node. Thus it is possible for more than one node to transmit a packet in a given slot. If this happens, all packets transmitted in that slot collide and are assumed to be lost.

Another function performed by a MAC protocol is dealing with packet loss. Packets can be lost because of corruption by noise or interference, or because they have collided with packets that are being sent by other users on a shared channel. In some cases, a retransmission mechanism is implemented. This requires a procedure to transmit some kind of acknowledgment if the transmitter is incapable of sensing the channel or hidden terminals that are out of the transmitter's detection range. A collided packet is held by its transmitting station until it is successfully transmitted, or the amount of time allotted for it to be sent expires, or a set maximum limit on the number of attempts is reached. These retransmissions have the effect of increasing the traffic on the channel, thereby increasing the likelihood of collisions. While retransmissions are typically handled at the layer that is closest to the medium, in some cases this retransmission function may be performed in higher layers, such as the transport layer.

5.3.2 Propagation

The physical signal is transmitted over the link between two nodes as an electromagnetic wave. As such, there is a propagation delay that reflects the time the signal takes to travel through the link's medium, whether the medium is an optical fiber, or the air between two radio antennas. The propagation delay is the ratio of the distance between the link endpoints to the speed of light through the link medium.

The speed of light in a transmission medium, such as air in the case of a wireless link, is $c_{medium} = c_{vacuum}/n_{medium}$, where c_{vacuum} = 299,792,458 m/s is the speed of light in a vacuum, and n_{medium} is the index of refraction of the medium. The index of refraction is a function of the frequency of the electromagnetic wave that is propagating through the medium. In the case of air, the index of refraction is approximately 1.0003 over a wide range of frequencies, while the index of refraction of the core of an optical fiber at the frequencies used for optical communications is often around 1.48. Thus, for example, a radio link between a FRV and a JAN tower that is 6 km (3.73 miles) away has a propagation delay of approximately 0.02 ms.

5.3.3 Channel Data Rate

The channel data rate represents the maximum rate at which data can be transmitted onto the channel or link. It is usually dependent on the rate the transmitter in the node can send data. It is expressed in the number of bits that can be sent in a 1-second interval, e.g., kbps.

5.3.4 Public Safety Communications Device

The origination and destination nodes in the reference path are PSCDs. A PSCD is assumed to have at least three separate physical wireless interfaces as shown in Figure 6. Link 1 and Interface 1 between the PSCD and its PAN constitute a local, very short distance wireless communications link, which must support from very low to very high data rates. Interface 2 uses Link 2 and Link 3 to respectively connect the PSCD to a FRV or to another PSCD. This is a pair of medium-range wireless links. When peer PSCDs are using Link 3 to communicate, they are doing so on a separate channel from that used by Link 2. Link 3 does not contend with or share the same channel as the link with the FRV or jurisdiction communication tower. Interface 3 and Link 5 between the PSCD and the JAN communication tower must allow the PSCD to transmit over long distances, since they are used when the PCSD is out of range of the IAN associated with the FRV.

The issue of RF resource limitations is another reason for assuming multiple physical wireless interfaces. It is not efficient to have a single device dividing its time while using the same resource. It will already need an extensive amount of overhead to permit the automatic discovery of devices and access points and handovers. This overhead can exceed resource limits and will affect the performance of the applications' data, especially service class 0.

5.3.5 First Responder's Vehicle

The first responder's vehicle (FRV) represents another node type in the reference path model. Examples of FRVs include a patrol car, a truck, an ambulance, a van, and a fire truck. This node's wireless access point is designed to link, or interface, with the PSCD, peer FRVs, and the jurisdiction communication tower (i.e., using Links 2, 4, and 6 and Interfaces 2 and 3 in Figure 6). Since the purpose of the FRV is to provide a mobile access point for any PSCD associated with the FRV, and to provide a communication path back to the jurisdiction communication tower if the PSCD cannot reach the tower alone, the FRV must support at least two physical wireless technologies.

Even though Link 4 uses Interface 2 along with Link 2 and Link 3, when it is present it will be using a channel that is not the one generated by the FRV for the IAN. This is to reduce the probability of hidden nodes, to reduce contention, and to better coordinate resources. Link 4 can be considered an aggregate of the IAN for the FRV. If all communications in the IAN use Link 4 to reach their end destinations, then Link 4 is the aggregate of the entire IAN. If some of the communication sessions stay within the IAN, then Link 4 is only a partial aggregation.

5.3.6 Jurisdiction Communication Tower

The jurisdiction communication tower represents a node in the reference path and connects to two links using one interface, according to Figure 6. Thus, one wireless technology serves both the individual PSCDs and the FRVs. Link 7 joins jurisdiction communication towers and is assumed to be point-to-point. It is used for redundancy or when Link 8, which is wired, is not available.

5.3.7 Generic Nodes

Generic nodes include access gateway, interworking gateway, distribution, and core nodes. These nodes implement generic network functionality, such as routing packets and relaying signaling information, and do not have any additional functionality specific to public safety communications.

5.3.8 Node Delay

Node delay consists of several components, including processing, packetization, look-ahead, and transmission delays. In the reference model we use a constant delay for each node type that represents the sum of these node delays.

5.3.8.1 Processing Delay

Processing delay represents the time a device takes to do any of its tasks associated with forwarding a packet (e.g., read or write memory location, execute an instruction). Therefore, the faster the device can do its tasks or the smaller number of tasks it has to do, the less delay will be introduced.

The source node processing delay consists of the coder delay and the algorithmic delay. Coder delay is the delay associated with encoding raw sampled data, such as 64 kbps pulse code modulated (PCM), to produce a lower bit-rate data stream that retains most of the quality of the original signal. The amount of delay depends on the type of coder that is being used, but the worst-case coder delay is typically in the range of 10-20 ms for speech applications. Video applications have bigger delay. Unencoded speech, described in ITU-T recommendation G.711 [15], has no coding delay. Algorithmic delay, or look-ahead delay, is described in Section 5.3.8.3.

5.3.8.2 Packetization Delay

The source PSCD, which is the first node on the path, imposes additional application-specific processing delays to convert application data into packets. This is the case with packetization delay, when sampled speech signals or images are turned into packets of data before they are sent across the network to the destination PSCD.

5.3.8.3 Look-Ahead Delay

The source PSCD also introduces look-ahead delay, which results from the necessity of compression algorithms to look ahead from the block of data they are processing to the next block of data as part of the

algorithm. G.711's lack of encoding means this delay does not occur if raw 64 kbps speech is being transmitted. The amount of algorithmic delay depends on the encoder being used; for example, the G.723.1 coder introduces a look-ahead delay of 7.5 ms.

5.3.8.4 Transmission Delay

Each node also imparts a transmission delay, which is simply the amount of time required to send a packet on a link from one node to the next node. It is inversely proportional to the data rate of the transmitter on the link. For example, if we consider a 1500-byte frame transmitted on an 11 Mbps link, the transmission delay is 1500 x 8/11e6 = 1.091 ms.

5.3.9 Area Networks

This section describes various assumptions about area networks (i.e., PAN, IAN, JAN, and EAN), based on the information in PS SoR Volume I [28] and other relevant documents from various standards bodies. We use these assumptions to generate default parameter values for sizing the various network segments that a path traverses.

5.3.9.1 PAN

The PAN, as currently defined in PS SoR Volume I, does not use wireless links to communicate between its sensors and PSCD. Instead the PAN is defined as a single link between an intermediate device that aggregates all of the data from the sensors and the PSCD by means of a wired link. A maximum distance of 2 meters (approximate average height of a human) for the length of the wired communication link is assumed (with an average distance being 1 meter, where the aggregating device is located on the torso and the PSCD is handheld). This link represents the separation between the intermediate device and the PSCD.

5.3.9.2 IAN

We assume the coverage area for the IAN (one mobile FRV and one or more PSCDs) depends on the technology in use and environmental conditions. However, a minimum coverage radius must be at least 250 meters, which is the greater of the following two reference points:

- Minimum fire hose length is 244 meters (800 feet) in NFPA 1901 [26].
- Minimum distance between incident at high-rise and base location is 60 meters (200 feet) in NFPA 1561 [27].

5.3.9.3 JAN

We assume the coverage area for the JAN communications base station with a tower (PSCDs and mobile vehicles) depends on the technology in use and environmental conditions. Since this is most likely a fixed network infrastructure, coverage of the entire jurisdiction can be accomplished by systematic placement of base stations (access points), provided sufficient resources are available. Therefore, we assume that the range of the JAN is from 5 kilometers to 110 kilometers. Based on a sample of current paths of communication towers, we use an estimating factor of 2 to multiply the actual distance to represent the networking communications path's distance within the JAN.

5.3.9.4 EAN

Since the EAN represents the embedded network infrastructure, we use the ITU-T Y.1541 [22] model for a network section (NS) for calculations with the user-to-network interfaces (UNIs) corresponding to the

interface between a JAN and the EAN. We assume a worst-case distance of 12,000 kilometers (within the continental United States).

5.3.10 Number of Nodes

The number of nodes on a path consists of the number of PSCDs, FRVs, and jurisdiction communication towers on the path. We derive the following default values from the assumptions presented in PS SoR Volume I [28]:

- The number of PSCDs per PAN is 1, by definition of a PSCD.
- The number of PSCDs per JAN is 30, based on statistical data of less than 100 sworn officers for 95 percent of the law enforcement agencies and an 8-hour work shift.
- The number of IANs is 15, 4, 10, or 3 per JAN, depending on the FRV's grouping.
  - The number of IANs per JAN can be as high as one for every PAN PSCD, or one per vehicle. (We assume the following occupancy levels per vehicle: car: 2, van: 8, ambulance: 3, fire truck: 10.)
  - Given that there are 30 PSCDs per JAN, the number of IANs within a JAN is based on the type of FRV and its assumed PSCD occupancy. For example, 30 PSCDs per JAN and 2 PSCDs per patrol car, yielding 15 IANs per JAN.
- The number of JANs is a maximum of 50,000.

  (There are 18,000 law enforcement agencies; 32,000 fire, EMS, local, state, tribal, and other organization agencies; and 100 Federal agencies.)
- The number of EANs is 1, as currently defined in PS SoR Volume I.

5.3.11 Protocols

To provide information concerning overall network packet and application overhead, we assume use of the Internet Protocol version 6 (IPv6), User Datagram Protocol (UDP), and Real-Time Transport Protocol (RTP). To provide MAC information, we assume use of the TDMA and Slotted Aloha protocols. Figure 9 shows the protocol stack at the source and destination PSCDs.

Figure 9: Protocol Stack for End User's PSCDs

Application Layer (speech or video over RTP)
Transport Layer (UDP)
Network Layer (IPv6)
MAC Layer (TDMA or Slotted ALOHA)
Physical Layer (coax, fiber, wireless)

5.3.11.1 Internet Protocol Version 6

Internet Protocol Version 6 (IPv6), described in the Internet Engineering Task Force (IETF) Request For Comments 2460 [7], is currently being deployed as a major overhaul to the current Internet network layer

technology (IPv4). With its new features, especially support for quality of service differentiation, it is expected to satisfy the requirements of any application for many years to come. An IPv6 packet header contains 20 bytes more than the IPv4 packet header does; most of this extra overhead is devoted to supporting the 128-bit source and destination addresses that were developed to preclude address space exhaustion. We assume the average node processing delays recorded for IPv4 apply to IPv6.

5.3.11.2 User Datagram Protocol

User Datagram Protocol (UDP), defined in RFC 768 [6], supplies a method for distinguishing among multiple applications by providing source and destination part identifiers in an 8-byte header. UDP also includes a 16-bit checksum to check for errors in the UDP header and payload. UDP does not provide guarantees of successful packet delivery. This is in contrast to the Transmission Control Protocol (TCP), which provides a reliable transport mechanism above the network layer. We use UDP in this analysis since the two applications (speech and video) favor timely delivery at a cost of potentially lost packets, instead of reliable delivery at a cost of potentially long delays.

5.3.11.3 Real-Time Transport Protocol

Real-Time Transport Protocol (RTP), defined in RFC 3550 [9], is a protocol framework designed for real-time applications, such as speech and video. RTP exists above the transport layer and applies a header to the application data that is 12 bytes long at minimum, although it can be up to 72 bytes long if the full set of up to 15 CSRC (Contributing SouRCe) identifiers is used. RTP allows additional header extensions, but their use is discouraged. The RTP header length can result in a significant amount of overhead, so the RTP framework allows for header compression on point-to-point links. The compression algorithm removes information that appears in the header of every packet data stream, and can reduce a 40-byte header to 2 to 4 bytes.

RTP usually runs over UDP (although it runs over TCP if it carries RealAudio data), which has no acknowledgement mechanism like the one in TCP. For this reason, RFC 3550 also defines the Real-Time Control Protocol (RTCP), for creating and transmitting sender and receiver reports that contain descriptions of the data sent and received and that act as acknowledgements.

5.3.12 User Applications

From the view of the network, a user application is modeled as a traffic generator. In other words at regular time intervals the user application produces a certain-sized packet (in bytes). Using these packet size and generation interval parameters, we can calculate an average application data rate, which gives us the offered load. Packet size and generation interval parameters are sufficient if the user application is a constant bit rate service. Additional parameters may be necessary to model a user application other than a constant bit rate service.

The next two sections describe tactical speech and video as constant bit rate user applications.

5.3.12.1 Example Speech Application

ITU-T G.711 [15] defines a speech encoding scheme known as Pulse Code Modulation (PCM), which produces an 8-bit sample every 125 microseconds. This results in an application data rate of 64 kbps. Since we assume the network is a packet network, and not a circuit-switched phone system, a number of samples must be grouped together to form an application packet. The packet has RTP, UDP, and IPV6 headers applied, and is given to the network to deliver. We assume that neither the size of the packet containing the original speech sample(s), nor the number of speech samples in a packet, changes over any link or section

in the transmission path. That is, no fragmentation occurs along any link on the path. We use two speech sampling packet sizes: 80 samples per packet and 320 samples per packet.

5.3.12.2 Example Video Application

The H.264 [25] and MPEG-2 [11] standards define different video encoding and transmission schemes. The output application data rate is modified so that it produces an average constant bit rate. We assume a 600-byte packet for H.264 video, and a 1358-byte packet for MPEG-2. The video application packet is encapsulated using the protocol stack of RTP (IETF RFC 3550 [9]) over UDP over IPv6. We also assume that the size of the packet containing the original video application packet does not change over any link or section in the transmission path. That is, no fragmentation takes place along any link on the path.

6 Measuring Network Performance

PS SoR Volume I [28] lists the following end-to-end PAN, IAN, JAN, and EAN upper-bound performance metrics:

- Packet transfer delay
- Packet delay variation
- Packet loss ratio
- Packet error ratio

This section provides an overview of several major factors affecting network performance, and describes the methodology used to create upper-bound network performance measures for end-to-end packet transfer delay and packet loss ratio. A future PS SoR version will provide measures for packet delay variation and packet error ratio.

6.1 Factors Affecting Network Performance

This section lists the most common factors that can impact the network performance. These include noise and interference as well as delays due to queuing and packet collisions.

6.1.1 Noise

Every packet is ultimately a long string of individual bits (zeros and ones) that are used to generate a modulated waveform that is transmitted through the air or over wires from a sending station, or source, to a receiving station, or destination. Individual bits in a received packet can be corrupted by noise on the communications channel. Two general types of noise cause errors: internal noise and external noise. Appendix A of Ziemer and Tranter's *Principles of Communication: Systems, Modulation, and Noise* [31] provides a detail discussion of noise sources.

6.1.1.1 Internal Noise

Internal noise originates within communications devices themselves. It is caused by a variety of physical phenomena, which we briefly list here:

- *Thermal noise* (also known as Johnson noise or Nyquist noise) occurs because of random electron motion due to heat that occurs in any conducting or semiconducting material. Thermal noise power levels do not change with respect to frequency; the power density near baseband is the same as the power density at very high frequencies. For this reason, this type of noise is also known as white noise.
- *Shot noise* (also known as Schottky noise) occurs in electronic devices like diodes or transistors. It arises because each electron that gets swept across the junction where two different types of semiconducting material meet carries a fixed amount of charge. The electrons cross the junction at random times and the result is a randomly fluctuating current exiting the device that can be modeled as a constant average current plus a random component.
- *Generation/recombination noise* occurs in semiconductors because of random generation and absorption of free negative charges (electrons) or free positive charges (holes) by the bonds

between the silicon and dopant atoms in the material. These events happen randomly over time, so this kind of noise can be thought of as a kind of shot noise.

- *Temperature-fluctuation noise* occurs because elements of electronic systems such as resistors and transistors get hot, but their temperature is not constant. This is because the rate at which these devices give heat to the surrounding environment varies randomly over time. As the temperature of the device changes, so does the amount of thermal noise. Thus, temperature-fluctuation noise can be thought of as an additional noise component within the thermal white noise process itself.
- *Flicker noise* (also known as pink noise or 1/*f* noise) is noise whose power is greater at lower frequencies than at higher frequencies. This type of noise has been observed in many types of electronic devices, from vacuum tubes to field effect transistors. The exact cause of this type of noise is still a matter of debate in the research community.

6.1.1.2 External Noise

External noise originates in the environment surrounding the communications devices. External noise can be generated by natural phenomena in the earth's atmosphere, mechanical devices, or energy sources in deep space. Examples of atmospheric effects include lightning strikes and auroras: the latter are generated by charged particles from the sun striking the earth's magnetic field and then traveling along field lines into the atmosphere. Human-made devices that generate noise include rotating machinery, automobile ignitions, and discharge from overhead power lines. Cosmic noise sources include the sun, other stars, pulsars, and quasars. While the sun is obviously the dominant noise source because of its close proximity to the earth, the more distant objects are collectively a significant source of noise due to their sheer numbers. Cosmic objects are broadband noise sources, and typically generate signals in the MHz to GHz portion of the spectrum.

A number of these external noise sources, particularly lightning and ignitions, generate noise that is highly impulsive: that is, a plot of the noise signal strength vs. time will show many large isolated spikes. You can observe this phenomena if you listen to a radio station transmitting in the AM band when a lightning storm is in the vicinity. The audible cracks and pops superimposed on the station's signal are the impulses generated by individual lightning events. As you might expect, these impulse events can introduce localized but severe errors in received packets.

6.1.2 Interference

In addition to channel noise, communications systems must contend with interference produced by other communications devices that operate in the same frequency band, but are not necessarily attempting to communicate directly with either the transmitting or receiving device of interest. This type of interference is known as radio frequency interference (RFI). The following are several different sources of RFI:

- *Co-channel interference* (CCI) occurs when another carrier is present on the channel being used. This kind of interfering signal is not from another user attempting to communicate with a shared access point; it can be due to a user on another network who is trying to communicate with a nearby access point. This tends to occur more often in urban environments, where different providers' wireless systems are clustered close together.
- *Adjacent channel interference* (ACI) is caused by transmitters whose center frequency lies outside the passband of the receiver, but whose spectrum still overlaps the passband. If there is enough overlap, the additional energy entering the receiver can cause bit errors.
- *Inter-symbol interference* (ISI) is produced by the mechanism that turns the ones and zeros composing the bit stream into a modulated signal for transmission to the receiver. In an ideal

situation, the signal energy associated with a given bit would be confined to the time interval corresponding to that bit. However, because the shaping filters used to generate the transmitted waveform are not perfect, the transmitted energy associated with each bit bleeds into the bit intervals that occur after the bit in question. Communications systems designers use channel equalization to reduce the amount of ISI; this is preferable to more brute force methods such as simply increasing the transmitter's power level.

6.1.3 Packet Collisions

Packet collisions are another form of interference that occurs in a shared medium such as a radio channel. Signals from two or more stations overlap while the stations are attempting to transmit data simultaneously. Because the signal strengths of the contending stations are generally of the same order of magnitude, the probability that the receiver will be able to determine correct bit values and associate them with the correct transmitter is very low. In this analysis, we assume that if any overlap occurs between two or more packets, all the colliding packets are lost. For this reason, MAC layer contention resolution protocols, beginning with Slotted Aloha in the early 1970s, and continuing with various types of CSMA protocols today, were designed to increase throughput and reduce the probability of collisions in situations where multiple users contend for access to a shared transmission medium.

6.1.4 Packetization

Packetization delay occurs because the source node needs time to collect several blocks of application data into a payload and form a packet with the necessary protocol headers. For simple coders like G.711 or G.726, [16] the packetization delay is computed as the ratio of the payload size (not including headers) in bits to the sampling rate in bits per second. For example, a G.711 coder that generates 8,000 8-bit samples per second and produces a payload of 640 bits (80 samples) incurs a packetization delay of 10 ms.

6.1.5 Queuing

Network equipment uses queues, also known as buffers, to hold packets for transmission at a later time. In particular, they are employed when a node receives more data than it can send out; in such a situation the node may chose to drop the data, causing packet loss, or place it in a queue until it can be sent on an outgoing link.

The time that a packet spends in a queue depends on many factors. The primary determiners of queuing delay are the queue utilization and the queue polling discipline. The former is related to the ratio of the arrival rate of packets into the queue to the rate at which they are transmitted on the outgoing link. If packets arrive at the queue according to a Poisson process, so that the amount of time between arrivals is exponentially distributed with mean $1/\lambda$ seconds, and if time to service a packet is distributed according to some arbitrary distribution with mean $1/\mu$ seconds and standard deviation σ_S seconds, the average queuing delay, *W*, measured in seconds is

$$W = \frac{1}{\mu} + \frac{1}{\lambda}\frac{\rho^2 + \lambda^2\sigma_s^2}{2(1-\rho)}$$

where:

$\rho = \lambda/\mu$ is the queue utilization.

Note that as ρ gets close to 1, *W* goes to infinity, meaning that a queue will tend to back up when packets arrive nearly as fast as the queue can get rid of them.

Queuing delay can also be affected by queue polling to support differentiated quality of service for different applications. Weighted fair queuing (WFQ) is used to make sure that packets associated with delay-sensitive applications do not linger too long at a given node. WFQ assumes some mechanism exists to classify packets by application type (speech, data, etc.). The node uses this information to sort arriving packets into different queues; each queue has its own priority level. The node is then able to pull packets from high-priority queues more often on average than from low-priority queues, allowing delay-sensitive applications to enjoy better service than low-priority applications like routine file transfers.

6.1.6 Packet Loss and Retransmission

Packet loss and retransmissions have a significant impact on the overall network performance. While packets lost in the lower layers may not necessarily affect the packet loss ratio experienced by the application (due to retransmission mechanisms in lower layers), the time taken to retransmit lost packets is a direct contributor to the end-to-end delay. A packet can be lost due to link-related issues such as collisions with other packets at the MAC layer or data corruption caused by bit errors. Packets can also be lost because of node-related issues, especially buffer overflow. The penalty for retransmitting lost packets is increased delay.

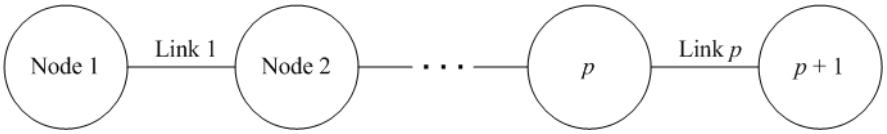
Packet retransmissions can be managed by the MAC layer (layer 2) or by the transport layer (layer 4, specifically if some variety of TCP is used). Many MAC protocols, from Slotted Aloha through CSMA, are designed to allow retransmission after a random delay when a packet is lost due to collision. For example, the IEEE 802.11 layer 2 protocol supports packet retransmission with an exponential backoff algorithm; after each collision, the expected waiting time before the next transmission attempt increases by a factor of 2. High packet loss rates will thus produce a large average delay. Some versions of Slotted Aloha or CSMA allow a maximum number of retransmissions before the packet is declared lost; at this point, higher-layer protocols like TCP may attempt to retransmit the packet from the source node.

In contrast to the MAC, which is managed on a link-by-link basis, TCP uses acknowledgement messages from the receiver node to the transmitting node to allow the latter to identify TCP segments that have been lost in transit so they can be resent. TCP also uses a slow start mechanism that increases the number of packets that can be sent at one time every time the transmitter receives a packet acknowledgement. If there are many losses, the number of packets that TCP will allow to be sent per second will remain small, thus increasing the average delay over the link.

6.2 Packet Loss Ratio Computations

This section describes the computations used to obtain an upper bound for the packet loss ratio as experienced by an application through the reference network model described above. Packet loss occurs when a packet is lost in transit, or when it is so corrupted by bit errors that it arrives at its destination but in an unusable state. The packet loss ratio is defined as the probability that a transmitted packet never reaches its destination.

Figure 10: Nodes and Links Composing a Path, with Numerical Identifiers



Given a path consisting of a set of p links $\{L_1, L_2, \ldots, L_p\}$ as shown in Figure 10, where the packet loss probability associated with link L_i is $PL_{link}(i)$, the success or failure of a transmission of a packet on a given link is independent of what happens on the other links on the path. The total packet loss probability for the path is 1 minus the probability that the packet is not lost on any of the *p* links. The probability that the packet is not lost on any of the links is the product of the complements of the loss probabilities for each of the links. Therefore, the path loss probability for a packet is:

$$PL_{path} = 1 - \prod_{i=1}^{p} (1 - PL_{link}(i))$$

where:

PL_{path} is the probability that a given packet will be lost somewhere on the path between the source and destination PSCDs.

We compute upper bounds for the packet loss probability, $PL_{link}(i)$, for two types of link layers. In the first case we examine a dedicated channel, in which individual users are able to reserve bandwidth and do not have to contend for access to the receiver. This type of channel sharing is common among long-range access networks such as IEEE 802.16 [4] networks. In the second case we consider Slotted Aloha, which is a CSMA MAC protocol and a simpler version of the protocol used by IEEE 802.11 networks.

6.2.1 Dedicated Channel

If we have a dedicated channel with time division multiplexing (TDM), and the offered data rate is less than or equal to the channel rate, the loss rate for packets is zero since there is no contention. If the offered load is greater than the channel data rate, the packet loss rate for a new user is 1 because it is blocked from accessing the channel. We assume there are no retransmissions at the transport level, and we define *G* to be the offered load from the user population, normalized by the channel data rate. The following expression gives the worst case packet loss probability from the perspective of the application:

$$PL_{link}(i) = \begin{cases} 0, G \le 1 \\ 1, G > 1 \end{cases}$$

This means that if any TDMA link on the path is oversubscribed, in the worst case, all packets will be lost and the attempted connection will be blocked.

6.2.2 Slotted Aloha

Slotted Aloha has been analyzed extensively in the technical literature, although most analyses compute throughput as a performance metric, rather than the collision probability. Our analysis follows the development in Rom and Sidi, *Multiple Access Protocols: Performance and Analysis*, [29] as a starting point.

First, we describe the assumptions we used in our model. We assume there is a single receiver and a finite number of stations, given by the parameter *M*, that are all located at the same distance from the receiving station. The user population does not change and none of the users move. There are no hidden terminals, that is, every station is in within transmitting range of every other station. We assume that packets arrive at each station according to a Poisson arrival process with a mean arrival rate of *g* packets per second. All packets have the same length, and require *T* seconds to transmit, given the channel data rate. For the case where the MAC protocol is Slotted Aloha, described in Section 5.3.1, we assume that a station holding packets for retransmission will send its packet in a given packet transmission interval with probability σ. Because of these assumptions, it follows that σ, which is the probability that a single station transmits during a given interval of time equal to the packet transmission time *T*, is related to the other parameters by the expression $gT = M\sigma$.

The length of a slot in seconds, *T*, is computed from the network parameters as:

$T = (L / C) / 1000$

where:

T is the length of the slot in seconds.

C is the channel data rate in kilobits per second.

L is the packet length in bits.

To obtain an upper bound for the packet loss probability $PL_{link}(i)$, we assume there are no retries for collided packets at layer 2, and that UDP is in use at the transport layer. Thus, PL_{link} on the link is equal to the packet collision probability, P_{coll}. To get P_{coll}, we first must compute σ, the probability that a station transmits in a slot of length *T*. σ is related to the offered load *G*.

G is the total offered bandwidth from the *M* stations using the channel normalized with respect to the channel data rate *C*, by:

$G = M\sigma$

where:

G is the total offered bandwidth from the *M* stations normalized with respect to the channel data rate *C*.

M is the number of stations.

σ is the probability that a station transmits in a slot of length *T*.

We get the collision probability by determining the fraction of slots in which the channel is busy that has multiple stations transmitting. To do this, we can look at a single slot and compute the probability that two or more stations are transmitting in the slot given that the slot is not idle (i.e., at least one station is transmitting in the slot). Letting N_T be the number of transmitting stations in the slot, this gives us:

$$PL_{link} = P_{coll} = \Pr\{N_T \geq 2 \mid N_T \geq 1\} = \frac{\Pr\{N_T \geq 2\}}{\Pr\{N_T \geq 1\}}$$

$$= \frac{1 - \Pr\{N_T = 1\} - \Pr\{N_T = 0\}}{1 - \Pr\{N_T = 0\}}$$

$$= \frac{1 - M\sigma(1-\sigma)^{M-1} - (1-\sigma)^M}{1 - (1-\sigma)^M} = \frac{1 - G\left(1 - \frac{G}{M}\right)^{M-1} - \left(1 - \frac{G}{M}\right)^M}{1 - \left(1 - \frac{G}{M}\right)^M}$$

where:

M, σ, and G are as defined above and we use the fact that $G = M\sigma$.

This expression is the probability of at least two stations transmitting in a slot divided by the probability of at least one station transmitting in the slot. It can also be thought of as the ratio of the number of slots with collisions to the number of slots that are active in a long time interval. In other words, if we look at a long time interval and count the number of slots that were active (i.e., had at least one station transmitting), and see what fraction of them had collisions (i.e., 2 or more stations transmitting at the same time), we would get a good estimate for the collision probability, P_{coll}. Increasing the length of the time interval will improve the estimate.

6.3 End-to-End Packet Transfer Delay Computations

In this section we discuss computations to generate upper bounds on end-to-end delay for a given application, using the network model described in the preceding section. The end-to-end delay is measured from the originating PSCD to the terminating PSCD.

Many phenomena affect the end-to-end packet transfer delay; some of these effects, like propagation through free space, produce fixed delays, while others, such as access and queuing, produce random delays. The causes of path delay include the length and type of each link, the number of links and the number of nodes on the path, and the processing speed of each node. Other causes of delay are the time a packet must spend in various nodes' data buffers, and medium access delays resulting from multiple users having to contend for access to a shared channel.

We derive an expression for the path delay that is a function of the delays associated with the individual links and nodes that lie on the path between the originating and terminating PSCDs. We define the component delays, and develop expressions for average link delays, based on the type of MAC in use on the link. As was the case in the discussion of the packet loss metric, we restrict the MAC types to TDMA and Slotted Aloha.

We have a path consisting of p links $\{L_1, L_2, \ldots, L_p\}$ and n nodes $\{N_1, N_2, \ldots, N_{p+1}\}$, as shown in Figure 10, where Node 1 is the originating PSCD and Node $p + 1$ is the terminating PSCD. The total expected one-way delay for the path, D_{path}, is the sum of the expected delays associated with each of the nodes and links. It is given by:

$$D_{path} = \sum_{i=1}^{p} D_{link}(i) + \sum_{j=1}^{p+1} D_{node}(j) + \sum_{j=1}^{p+1} D_{access}(j)$$

where:

$D_{link}(i)$ is the average delay associated with the *ith* link in seconds.

$D_{node}(j)$ is the average delay associated with the *jth* node in seconds.

$D_{access}(j)$ is the average delay associated with accessing the *jth* link in seconds.

6.3.1 Link Delays

The delay incurred on the link depends on the link distances and the signal propagation through the media type. The signal propagation is the time to send a signal through a particular medium per distance covered. For each link in a path, signal propagation delay is calculated by dividing the link length in meters by the signal propagation speed through the medium. For a terrestrial coaxial cable, the signal propagation delay is calculated by taking the link multiplied by a unit delay of 4 ms per km. Signals in optical fiber cables travel a little slower (5 ms per km). (See reference ITU-T G.114 [22]). For example, a 2000 km fiber link has a propagation delay of 10 ms.

An expression for $D_{link}(i)$ is as follows:

$$D_{link}(i) = l(i) / s(i)$$

where:

$D_{link}(i)$ is the average delay on the *ith* link in seconds.

$l(i)$ is the length of link *i* in meters.

$s(i)$ is the propagation speed on the *ith* link in meters per second.

6.3.2 Node Delays

The nodes on the path will also add processing delay to the total time required to send a packet. The originating PSCD and terminating PSCD are nodes, but so too are the FRVs, the jurisdiction tower, and any other device that lies on the path. The average total delays for four types of network nodes (access gateway, interworking gateway, distribution, and core) are given in ITU-T Y.1541 [20]. The average total delay includes queuing and processing delays. This gives us a figure for the node delay, $D_{node}(i)$. We assume that $D_{node}(i)$ is constant per node type.

Some of the processing delay is determined by the protocol stack. The model includes the amount of various protocol overhead in the delay computations. This overhead is added to the amount of data generated per packet by a given user. This is used to determine the total data rate of the traffic generated for the network links and nodes to service. From this we can compute delays that include the effect of overhead.

6.3.3 Medium Access Delays

The medium access delay is mainly affected by the MAC protocol, the channel data rate, and the load on a link. This section considers two generic MAC protocols: TDMA (dedicated) and Slotted Aloha.

6.3.3.1 Dedicated Channel (TDMA)

TDMA divides the given link capacity into channels by assigning users to different slots in a large frame that the MAC layer assembles and transmits. A given number of slots is available to accommodate a

number of communication sessions at a given quality of service. In this situation, the medium access delay is the average time a station must wait for an assigned slot for a communication that is permitted access. However, once the capacity of the link is reached, no new communication sessions will be accepted. Thus, the delay for an unsuccessful communication attempt is infinite, or at least as long as necessary for established communication sessions to release enough resources so that the new session can get a channel. Because we are interested in the worst case, we model the delay as infinite if the number of stations in an area network exceeds what the network can support, given its channel capacity.

An expression for $D_{access}(i)$ is as follows:

$$D_{access}(i) = \begin{cases} \frac{T}{2}, G \le 1 \\ \infty, G > 1 \end{cases}$$

where:

$D_{access}(i)$ is the average delay to access the *ith* link in seconds.

T is the slot length in seconds.

G is the offered load from the user population normalized to the channel data rate.

6.3.3.2 Slotted Aloha

To get an upper bound for the average delay associated with using Slotted Aloha, we assume the worst-case scenario, in which there is no limit to the number of times that collided packets can attempt retransmission. A simple analysis that results in an expression for the average delay can be found in Rom [29]; we cover the main points here.

If a packet is successfully transmitted on the first attempt, its delay is a single slot interval. If the first attempt fails, the station that is attempting to transmit the packet goes into a state known as *backlog*, in which it holds the packet and attempts to retransmit it after a random delay. As long as the backlogged station's attempts to transmit the packet that it is holding fail, it continues to schedule the packet for another attempt after a random delay. Because there is no limit to the number of times a backlogged station can attempt to transmit a packet, eventually the packet will be transmitted successfully (as long as the probability of a collision is not 1), and the backlogged station will move on to the next packet. Because of this policy, the average access delay experienced by a packet whose first transmission attempt is unsuccessful is one slot plus the average number of slot intervals that the packet waits for while the packet is held by the station (i.e., the packet is backlogged within the station).

By using Little's Theorem, we find the average number of slot intervals that a packet spends in backlog is the ratio of the average number of stations in backlog, $\overline{N}$, to the rate that stations become backlogged, b. So the average delay, in slots, is:

$$E\{D_{access}\} = Pr\{\text{success on 1st try}\} \cdot 1 + Pr\{\text{failure on 1st try}\}\left(\frac{\overline{N}}{b} + 1\right)$$

The throughput of a Slotted Aloha system, S, is the rate at which packets leave the station and are received by the destination node. Some of this throughput is from backlogged stations and the rest is from stations that succeed on the first attempt. The probability that a packet goes into backlog is b/S, the ratio of the backlog rate to the throughput. Therefore, we get:

$$E\{D_{access}\} = \frac{S-b}{S} \cdot 1 + \frac{b}{S}\left(\frac{\bar{N}}{b} + 1\right) = \frac{S + \bar{N}}{S}$$

To get the expected delay, we need to get the average number of backlogged stations. We note that the throughput *S* is equal to the offered load $G = M\sigma$, minus the load corresponding to unsuccessful transmissions from the backlogged stations, which is $\bar{N}\sigma$. So:

$$E\{D_{access}\} = 1 + \frac{G-S}{S\sigma} = 1 + \frac{M}{S} - \frac{1}{\sigma}$$

The throughput *S* as a function of *G* and *M* is given by:

$$S = G\left(1 - \frac{G}{M}\right)^{M-1} = M\sigma(1-\sigma)^{M-1}$$

This gives us the average delay in slot intervals:

$$E\{D_{access}\} = 1 + \frac{G-S}{S\sigma} = 1 + \frac{1}{\sigma(1-\sigma)^{M-1}} - \frac{1}{\sigma}$$

Multiplying this quantity by the length of a slot gives us the average delay in seconds:

$D_{access}(i) = T * E\{D_{access}\}$

where:

$D_{access}(i)$ is the average delay in seconds.

T is the slot length in seconds.

$E\{D_{access}\}$ is the average delay in slots.

Note that for the case where there is a single user ($M = 1$), the expected delay is one slot interval, which is what we would expect since no other stations are contending for access to the channel, and every packet transmission attempt is successful.

7 Network Requirements

This section provides network performance requirements to support the tactical speech and video application requirements, also known as user-to-user-perceived quality of service requirements. In addition, this section provides a summary for all area network performance requirements.

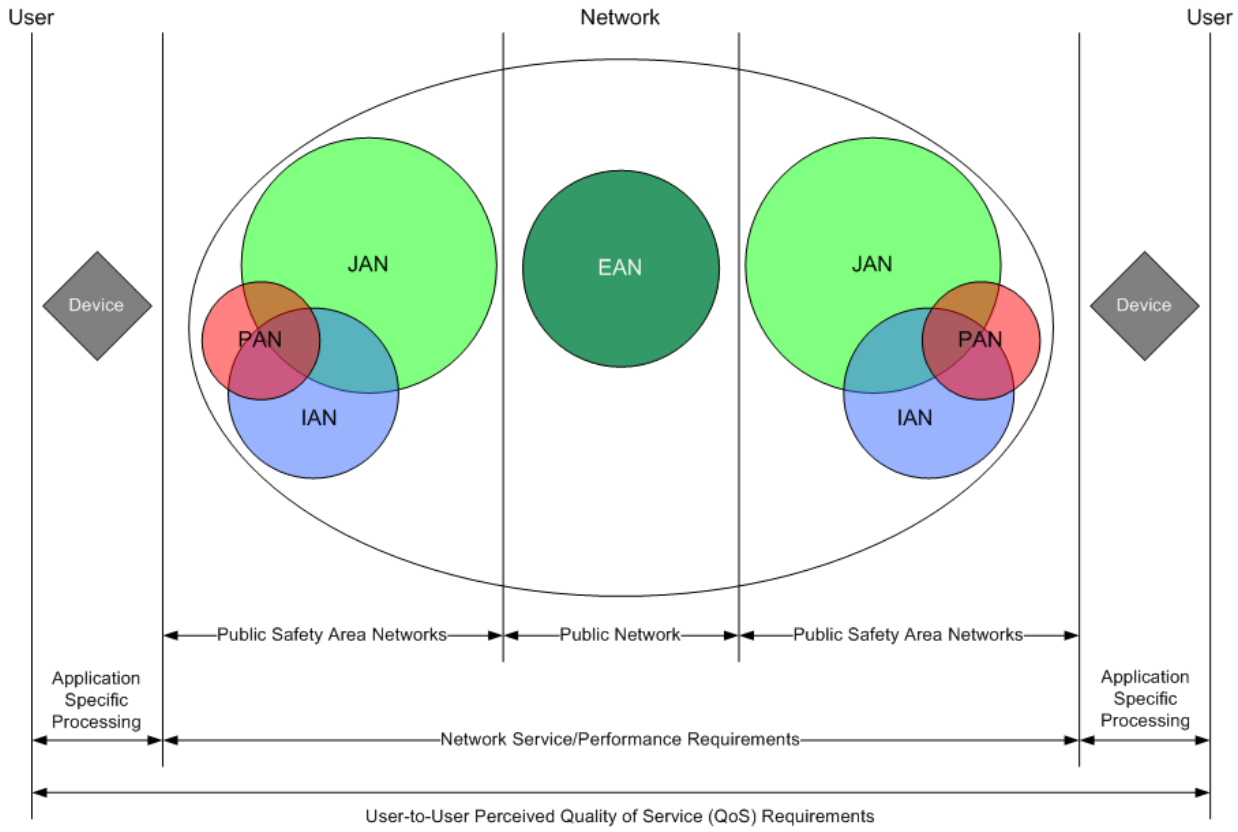
7.1 User-Perceived Quality of Service

Figure 11 illustrates the general performance requirements for user-to-user-perceived quality of service. These requirements can be broken into two components: the application-specific processing component, and the network service component. While performance requirements imposed by the application-specific processing component are important to consider, they depend on the type of application and thus are generally outside the scope of this study. We are mainly concerned with the network service component, which can be further divided into two major segments, as follows:

- The public network (i.e. EAN)
- The various public safety area networks detailed in Section 5.3.9

Depending on the path, a subset of the available area network types will compose the public safety component of the network segment.

Figure 11: General Performance Requirements for User-Perceived Quality of Service



Following is the step-by-step approach we use to provide upper bounds for the network performance requirements:

1. We use the path-based network reference model and the measurement methodology described in Section 5 and Section 6, respectively, to calculate upper bounds for the packet loss and the end-to-end delay for Paths A through G, inclusive.
2. We compare these path bounds to the user-to-user-perceived performance requirements established in PS SoR Volume I [28]. In this comparison, we account for the application-specific processing requirement for the given application type.
3. In case the path upper bounds obtained in step 1 exceed the user-to-user-perceived performance requirements, the path bounds are set to the user-to-user-perceived performance requirements, minus the application specific processing requirements.
4. We obtain the public safety area networks performance allocations from each path's upper bounds by grouping the links and nodes in the path into area networks.
5. We obtain upper bounds for area network performance requirements by comparing the set of area network performance requirements for all paths, and selecting the maximum values from the set.

7.2 Speech Applications

The network performance requirements necessary to support speech applications are based on the user-perceived quality of service study for speech applications described in the "Measurement of Speech Transmission Suitability" [32] report. In addition, the maximum acceptable mouth-to-ear packet loss ratio and end-to-end transit delay were obtained from ITU-T G.711 "Pulse Code Modulation of Voice Frequencies" [15]. These metrics are summarized in Section 7.2.1 and Section 7.2.2 for packet loss and delay, respectively. Section 7.2.3 through Section 7.2.9 provide the network performance budget for Paths A through G, respectively, in terms of the packet loss and end-to-end delay.

7.2.1 Packet Loss Requirements

The packet loss requirements described in Section 2 for speech depend upon several factors such as packet size, speech decoding, and post processing algorithms (e.g., packet loss concealment, or PLC). This section describes packet loss requirements for packet sizes of 80 and 320 bytes corresponding to packet interarrival times of 10 ms and 40 ms, respectively.

Table 10 summarizes the packet loss requirements considered in this study for different percentages of satisfied public safety practitioners, assuming a packet loss correlation of 0.0.

Table 10: Packet Loss Requirements for Percentages of Satisfied Practitioners

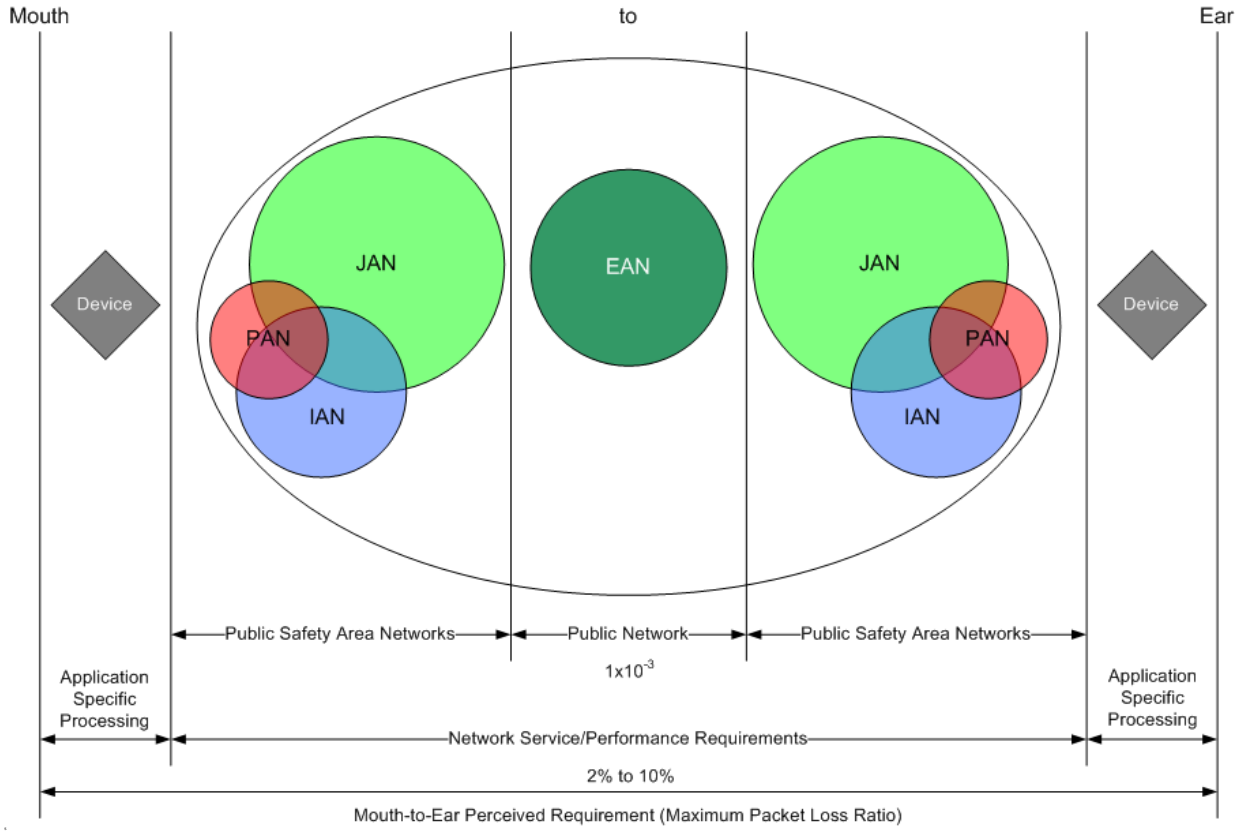
Percentage of Satisfied Practitioners	Packet Voice Sample Size	Packet Loss Percent Requirements
70 percent	80 bytes	10 percent
	320 bytes	5 percent

Table 10: Packet Loss Requirements for Percentages of Satisfied Practitioners (Continued)

Percentage of Satisfied Practitioners	Packet Voice Sample Size	Packet Loss Percent Requirements
80 percent	80 bytes	5 percent
	320 bytes	2 percent
90 percent	80 bytes	2 percent
	320 bytes	2 percent

Figure 12 illustrates the budget allocation for speech packet loss ratios. We use a network packet loss budget requirement of 1×10^{-3} specified in ITU-T Y.1541 [21] for the EAN.

Figure 12: Speech Maximum Packet Loss Ratio Requirements



7.2.2 End-to-End Delay Requirements

Given a maximum mouth-to-ear, end-to-end delay of 150 ms and three segments of the mouth-to-ear path (PSCD; PAN, IAN, and JAN; EAN), Figure 13 illustrates the recommended maximum end-to-end delay allocations, presented in terms of path and other input parameters. The value of "x" that represents the delay contributed by the application-specific processing device includes the packetization delay described in Section 6.1.4, and is bounded by 0.375 ms according to ITU-T G.114 [22] for PCM.

We use a network budget requirement for end-to-end transit delay of 100 ms given in ITU-T Y.1541 [20] for the EAN.

Figure 13: Speech Maximum End-to-End Delay Requirements

7.2.3 Path A

Path A consists of two PANs and one IAN. This path is one in which it is very easy to allocate the link budget for the various network performance parameters.

The PAN component is stable with a very small contribution to network performance.

The IAN component, assuming a fixed node (FRV) delay, has its major performance-affecting factors: the choice of MAC protocols, choice of packet size, number of PSCDs, and channel rate of the link. The packet loss probability, however, is affected significantly by the choice of MAC and the number of PSCDs.

Table 11: Speech Path A Network Performance Parameter Requirements

Speech Path A	Calculated				Requirement	
	80-Byte Packets		320-Byte Packets		80-Byte Packets	320-Byte Packets
Details	**Minimum**	**Maximum**	**Minimum**	**Maximum**	**Upper Bound**	**Upper Bound**
End-to-End Delay	11 ms	49 ms	11 ms	115 ms	45 ms**	17 ms**
Packet Loss Probability	0.00001	0.3944	0.00001	0.3944	0.1*	0.05*
PAN						
End-to-End Delay	<< 1 ms	<< 1 ms	<< 1 ms	<< 1 ms	1 ms	1 ms
Packet Loss Probability	N/A	N/A	N/A	N/A	0.001	0.001
IAN						
End-to-End Delay	11 ms	49 ms	11 ms	115 ms	45 ms**	17 ms**
Packet Loss Probability	0.00001	0.3944	0.00001	0.3944	0.1*	0.05*

Note*: Upper bound based on user-perceived mouth-to-ear delay study

Note**: Upper bound influenced by packet loss probability

N/A: Not included in the model

7.2.4 Path B

Path B consists of two PANs and one JAN. This path, unlike path A, may possibly contain multiple links within the JAN (i.e., not simply a communication path from PSCD to a single jurisdictional communication tower to another PSCD).

The PAN component is stable with a very small contribution to network performance.

The JAN component, assuming a fixed node (jurisdictional communication tower) delay, has as its major performance-affecting factors: the choice of MAC protocols, choice of packet size, number of PSCDs, and the channel data rate on the link.

Table 12: Speech Path B Network Performance Parameter Requirements

Speech Path B	Calculated				Requirement	
	80-Byte Packets		320-Byte Packets		80-Byte Packets	320-Byte Packets
Details	Minimum	Maximum	Minimum	Maximum	Upper Bound	Upper Bound
End-to-End Delay	7 ms	122 ms	11 ms	320 ms	40 ms**	60 ms**
Packet Loss Probability	0.00001	0.3944	0.00001	0.3944	0.1*	0.05*
PAN						
End-to-End Delay	<< 1 ms	<< 1 ms	<< 1 ms	<< 1 ms	1 ms	1 ms
Packet Loss Probability	N/A	N/A	N/A	N/A	0.001	0.001
JAN						
End-to-End Delay	7 ms	122 ms	11 ms	320 ms	40 ms**	60 ms**
Packet Loss Probability	0.00001	0.3944	0.00001	0.3944	0.1*	0.05*

Note*: Upper bound based on user-perceived mouth-to-ear delay study

Note**: Upper bound influenced by packet loss probability

N/A: Not included in the model

7.2.5 Path C

Path C consists of two PANs, two IANs and one JAN.

The PAN component is stable with a very small contribution to network performance.

The IAN component, assuming a fixed node (FRV) delay, has as its major performance-affecting factors: the choice of MAC protocols, choice of packet size, number of PSCDs, and the channel data rate on the link. The packet loss probability, however, is affected significantly by the choice of MAC and the number of PSCDs.

The JAN component, like the IAN component, is significantly affected by the choice of MAC, number of PSCDs and FRVs, packet size, and link channel rate.

Table 13: Speech Path C Network Performance Parameter Requirements

Speech Path C	Calculated				Requirement	
	80-Byte Packets		320-Byte Packets		80-Byte Packets	320-Byte Packets
Details	Minimum	Maximum	Minimum	Maximum	Upper Bound	Upper Bound
End-to-End Delay	27 ms	∞	39 ms	∞	32 ms**	39 ms**
Packet Loss Probability	0.0014	1	0.0014	1	0.1*	0.05*
PAN						
End-to-End Delay	<< 1ms	<< 1ms	<< 1ms	<< 1ms	1 ms	1 ms
Packet Loss Probability	N/A	N/A	N/A	N/A	0.001	0.001
IAN						
End-to-End Delay	11 ms	∞	12 ms	∞	11 ms	12 ms
Packet Loss Probability	0.0001	1	0.0001	1	0.0007	0.007
JAN						
End-to-End Delay	7 ms	∞	16 ms	∞	11 ms	17 ms
Packet Loss Probability	0.0162	1	0.0162	1	0.0163	0.0163

Note*: Upper bound based on user-perceived mouth-to-ear delay study

Note**: Upper bound influenced by packet loss probability

N/A: Not included in the model

7.2.6 Path D

Path D is similar to Path C, except it contains the EAN.

The PAN component is stable with a very small contribution to network performance.

The IAN component, assuming a fixed node (FRV) delay, has as its major performance-affecting factors: the choice of MAC protocols, choice of packet size, number of PSCDs, and channel data rate on the link. The packet loss probability, however, is affected significantly by the choice of MAC and the number of PSCDs.

The JAN component, like the IAN component, is significantly affected by the choice of MAC, number of PSCDs and FRVs, packet size, and link channel rate.

The EAN is assumed to satisfy the network performance objectives given in ITU-T Y.1541 [20]. We use a network budget requirement for the packet loss ratio of 10^{-3}, and an end-to-end transit delay of 100 ms given in ITU-T Y.1541.

Table 14: Speech Path D Network Performance Parameter Requirements

Speech Path D	Calculated				Requirement	
	80-Byte Packets		320-Byte Packets		80-Byte Packets	320-Byte Packets
Details	**Minimum**	**Maximum**	**Minimum**	**Maximum**	**Upper Bound**	**Upper Bound**
End-to-End Delay	54 ms	∞	74 ms	∞	61 ms**	74 ms**
Packet Loss Probability	0.0014	1	0.325	1	0.1*	0.05*
PAN						
End-to-End Delay	<< 1ms	<< 1ms	<< 1ms	<< 1ms	1 ms	1 ms
Packet Loss Probability	N/A	N/A	N/A	N/A	0.001	0.001
IAN						
End-to-End Delay	11 ms	50 ms	12 ms	115 ms	11 ms	12 ms
Packet Loss Probability	0.0001	0.2218	0.0001	0.2218	0.0001	0.0001
JAN						
End-to-End Delay	10 ms	∞	16 ms	∞	10 ms	16 ms
Packet Loss Probability	0.0163	1	0.0163	1	0.0163	0.0163
EAN						
End-to-End Delay	20 ms	N/C	20 ms	N/C	100 ms***	100 ms***
Packet Loss Probability	0	N/C	0	N/C	0.001***	0.001***

Note*: Upper bound based on user-perceived mouth-to-ear delay study

Note**: Upper bound influenced by packet loss probability

Note***: Upper bound from ITU-T Y.1541

N/A: Not included in the model

N/C: Kept constant

7.2.7 Path E

Path E consists of two PANs, two JANs, and the EAN.

The PAN component is stable with a very small contribution to network performance.

The JAN component, assuming a fixed node (jurisdictional communication tower) delay, has as its major performance-affecting factors: the choice of MAC protocols, choice of packet size, number of PSCDs, and the channel rate of the link.

The EAN is assumed to satisfy the network performance objectives given in ITU-T Y.1541 [20].

We use a network budget requirement for the packet loss ratio of 10^{-3}, and an end-to-end transit delay of 100 ms given in ITU-T Y.1541.

Table 15: Speech Path E Network Performance Parameter Requirements

Speech Path E	Calculated				Requirement	
	80-Byte Packets		320-Byte Packets		80-Byte Packets	320-Byte Packets
Details	Minimum	Maximum	Minimum	Maximum	Upper Bound	Upper Bound
End-to-End Delay	33 ms	264 ms	40 ms	660 ms	100 ms**	140 ms**
Packet Loss Probability	0.00001	0.3944	0.0003	0.3944	0.1*	0.05*
PAN						
End-to-End Delay	<< 1ms	<< 1ms	<< 1ms	<< 1ms	1 ms	1 ms
Packet Loss Probability	N/A	N/A	N/A	N/A	0.001	0.001
JAN						
End-to-End Delay	8 ms	122 ms	11 ms	320 ms	40 ms	65 ms
Packet Loss Probability	0.0003	0.3944	0.0003	0.3944	0.0508	0.0264
EAN						
End-to-End Delay	20 ms	N/C	20 ms	N/C	100 ms***	100 ms***
Packet Loss Probability	0	N/C	0	N/C	0.001***	0.001***

Note*: Upper bound based on user-perceived mouth-to-ear delay study

Note**: Upper bound influenced by packet loss probability

Note***: Upper bound from ITU-T Y.1541

N/A: Not included in the model

N/C: Kept constant

7.2.8 Path F

Path F consists of two PANs, one IAN, and one JAN.

The PAN component is stable with a very small contribution to network performance.

The IAN component, assuming a fixed node (FRV) delay, has as its major performance-affecting factors: the choice of MAC protocols, choice of packet size, number of PSCDs, and channel data rate on the link. The packet loss probability, however, is affected significantly by the choice of MAC and the number of PSCDs.

The JAN component, assuming a fixed node (jurisdictional communication tower) delay, has as its major performance-affecting factors: the choice of MAC protocols, choice of packet size, number of PSCDs and FRVs, and the channel rate of the link.

Table 16: Speech Path F Network Performance Parameter Requirements

Speech Path F	Calculated				Requirement	
	80-Byte Packets		320-Byte Packets		80-Byte Packets	320-Byte Packets
Details	Minimum	Maximum	Minimum	Maximum	Upper Bound	Upper Bound
End-to-End Delay	17 ms	∞	27 ms	∞	22 ms**	29 ms**
Packet Loss Probability	0.0007	1	0.0165	1	0.1*	0.05*
PAN						
End-to-End Delay	<< 1ms	<< 1ms	<< 1ms	<< 1ms	1 ms	1 ms
Packet Loss Probability	N/A	N/A	N/A	N/A	0.001	0.001
IAN						
End-to-End Delay	11 ms	49 ms	12 ms	115 ms	11 ms	12 ms
Packet Loss Probability	0.0001	0.2218	0.0001	0.2218	0.0007	0.0007
JAN						
End-to-End Delay	10 ms	∞	16 ms	∞	11 ms	17 ms
Packet Loss Probability	0.0163	1	0.0163	1	0.0164	0.0163

Note*: Upper bound based on user-perceived mouth-to-ear delay study

Note**: Upper bound influenced by packet loss probability

N/A: Not included in the model

7.2.9 Path G

Path G consists of two PANs, one IAN, two JANs, and the EAN.

The PAN component is stable with a very small contribution to network performance.

The IAN component, assuming a fixed node (FRV) delay, has as its major performance-affecting factors: the choice of MAC protocols, choice of packet size, number of PSCDs, and channel data rate on the link. The packet loss probability, however, is affected significantly by the choice of MAC and the number of PSCDs.

The JAN component, assuming a fixed node (jurisdictional communication tower) delay, has as its major performance-affecting factors: the choice of MAC protocols, choice of packet size, number of PSCD and FRVs, and the channel rate of the link.

The EAN is assumed to satisfy the network performance objectives given in ITU-T Y.1541 [20].

We use a network budget requirement for packet loss ratio of 10^{-3}, and an end-to-end transit delay of 100 ms given in ITU-T Y.1541.

Table 17: Speech Path G Network Performance Parameter Requirements

Speech Path G	Calculated				Requirement	
	80-Byte Packets		320-Byte Packets		80-Byte Packets	320-Byte Packets
Details	**Minimum**	**Maximum**	**Minimum**	**Maximum**	**Upper Bound**	**Upper Bound**
End-to-End Delay	44 ms	∞	63 ms	∞	52 ms**	65 ms**
Packet Loss Probability	0.0007	1	0.0165	1	0.1*	0.05*
PAN						
End-to-End Delay	<< 1ms	<< 1ms	<< 1ms	<< 1ms	1 ms	1 ms
Packet Loss Probability	N/A	N/A	N/A	N/A	0.001	0.001
IAN						
End-to-End Delay	11 ms	49 ms	12 ms	115 ms	11 ms	12 ms
Packet Loss Probability	0.0001	0.2218	0.0001	0.2218	0.0007	0.0001
JAN						

Table 17: Speech Path G Network Performance Parameter Requirements (Continued)

Speech Path G	Calculated				Requirement	
	80-Byte Packets		320-Byte Packets		80-Byte Packets	320-Byte Packets
Details	**Minimum**	**Maximum**	**Minimum**	**Maximum**	**Upper Bound**	**Upper Bound**
End-to-End Delay	10 ms	∞	16 ms	∞	10 ms	16 ms
Packet Loss Probability	0.0163	1	0.0163	1	0.0164	0.0163
EAN						
End-to-End Delay	20 ms	N/C	20 ms	N/C	100 ms***	100 ms***
Packet Loss Probability	0	N/C	0	N/C	0.001***	0.001***

Note*: Upper bound based on user-perceived mouth-to-ear delay study

Note**: Upper bound influenced by packet loss probability

Note***: Upper bound from ITU-T Y.1541

N/A: Not included in the model

N/C: Kept constant

7.3 Video Applications

The network performance requirements necessary to support video applications are based on the performance requirements for video applications described in Section 4. This section addresses MPEG-2 and H.264 encodings and the maximum acceptable eye-to-eye packet loss ratio and end-to-end transit delay given. These metrics are summarized in Section 7.3.1 and Section 7.3.2 for packet loss and delay, respectively. Section 7.2.3 through Section 7.2.9 provide the network performance budget for Paths A through G, respectively, in terms of the packet loss and end-to-end delay.

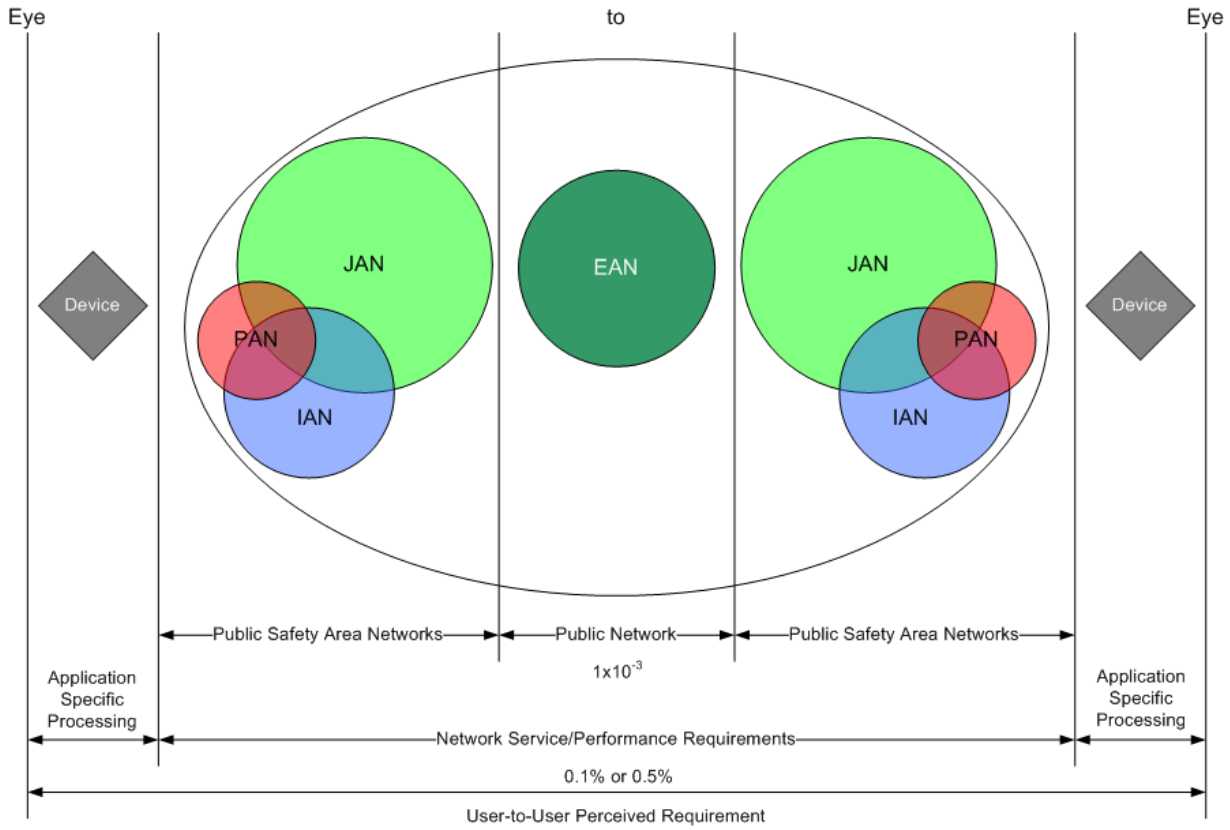
7.3.1 Packet loss Requirements

Section 4 lists video application transmission requirements that include end-to-end packet loss requirements. Figure 3 provides a video reference model illustrating the video transmission portion of a video system. The reference lines in Figure 14 where the device has a common interface with the PAN, are equivalent to the C and D reference points in Figure 3 and Figure 4. Section 3.5.3.2 describes packet loss for these C and D reference points.

Figure 14 shows the packet loss probability across its many components for the video application. From Table 6, we consider a maximum eye-to-eye packet loss ratio of 0.1 and 0.5 percent for H.264 and MPEG-2 encodings, respectively.

We use a network budget requirement of 10^{-3} given in ITU-T Y.1541 [20] for the EAN.

Figure 14: Video Maximum Packet Loss Ratio Requirements



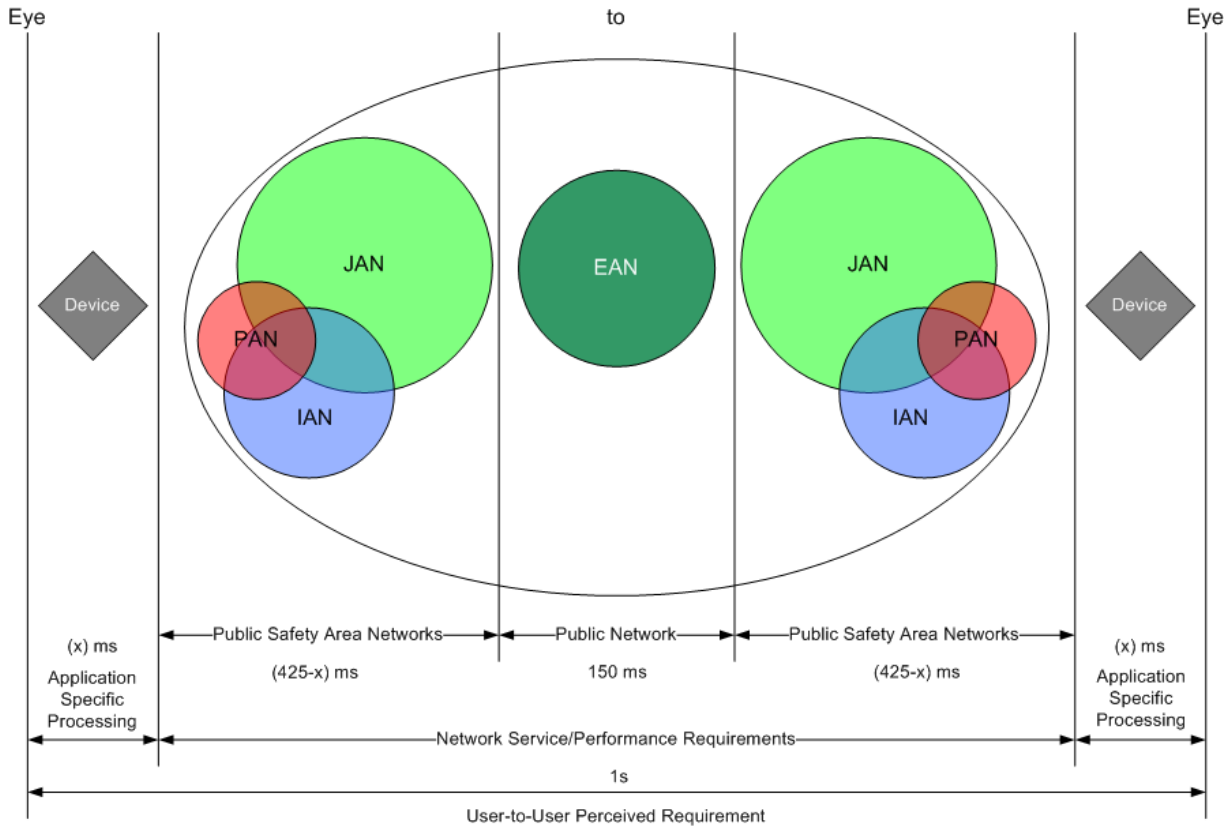
7.3.2 End-to-End Delay Requirements

Section 4 lists video application transmission requirements that include end-to-end delay requirements. Figure 3 provides a video reference model illustrating the video transmission portion of a video system. The reference lines in Figure 15 where the device has a common interface with the PAN, are equivalent to the B and E reference points in Figure 3. Section 3.3.1 describes one-way delay for these B and E reference points.

Given a maximum eye-to-eye, end-to-end delay of 1 second noted in Table 5, Figure 15 illustrates the recommended maximum end-to-end delay allocations presented in terms of path and other input parameters. The value of "x," which represents the delay contributed by the application-specific processing device, includes the packetization delay for video.

We use a network budget requirement for end-to-end transit delay of 150 ms given in ITU-T Y.1541 [20] for the EAN.

Figure 15: Video Maximum End-to-End Delay Requirements



7.3.3 Path A

Path A consists of two PANs and one IAN. This path is one in which it is very easy to allocate the link budget for the various network performance parameters.

The PAN component is stable with a very small contribution to network performance.

The IAN component, assuming a fixed node (FRV) delay, has as its major performance-affecting factors: the choice of MAC protocols, choice of packet size, number of PSCDs, and channel rate of the link. The packet loss probability, however, is affected significantly by the choice of MAC and the number of PSCDs.

Table 18: Video Path A Network Performance Parameter Requirements

Video Path A	Calculated				Requirement	
	H.264		MPEG-2		H.264	MPEG-2
Details	**Minimum**	**Maximum**	**Minimum**	**Maximum**	**Upper Bound**	**Upper Bound**
End-to-End Delay	11 ms	42 ms	11 ms	227 ms	11 ms**	11 ms**
Packet Loss Probability	0	0.6239	0	0.9518	0.001*	0.005*

Table 18: Video Path A Network Performance Parameter Requirements (Continued)

Video Path A	Calculated				Requirement	
	H.264		MPEG-2		H.264	MPEG-2
Details	**Minimum**	**Maximum**	**Minimum**	**Maximum**	**Upper Bound**	**Upper Bound**
PAN						
End-to-End Delay	<< 1 ms	<< 1 ms	<< 1 ms	<< 1 ms	1 ms	1 ms
Packet Loss Probability	N/A	N/A	N/A	N/A	0.001	0.001
IAN						
End-to-End Delay	11 ms	42 ms	11 ms	227 ms	11 ms**	11 ms**
Packet Loss Probability	0	0.6239	0	0.9518	0.001*	0.05*

Note*: Upper bound based on user-perceived mouth-to-ear delay study

Note**: Upper bound influenced by packet loss probability

N/A: Not included in the model

7.3.4 Path B

Path B consists of two PANs and one JAN. This path, unlike path A, can contain multiple links within the JAN (i.e., not simply a communication path from PSCD to a single jurisdictional communication tower to another PSCD).

The PAN component is stable with a very small contribution to network performance.

The JAN component, assuming a fixed node (jurisdictional communication tower) delay, has as its major performance-affecting factors: the choice of MAC protocols, choice of packet size, number of PSCDs, and the channel rate of the link.

Table 19: Video Path B Network Performance Parameter Requirements

Video Path B	Calculated				Requirement	
	H.264		MPEG-2		H.264	MPEG-2
Details	**Minimum**	**Maximum**	**Minimum**	**Maximum**	**Upper Bound**	**Upper Bound**
End-to-End Delay	7 ms	100 ms	7 ms	655 ms	8 ms**	10 ms**
Packet Loss Probability	0	0.6239	0	0.9518	0.001*	0.005*

Table 19: Video Path B Network Performance Parameter Requirements (Continued)

Video Path B	Calculated				Requirement	
	H.264		MPEG-2		H.264	MPEG-2
Details	**Minimum**	**Maximum**	**Minimum**	**Maximum**	**Upper Bound**	**Upper Bound**
PAN						
End-to-End Delay	<< 1 ms	<< 1 ms	<< 1 ms	<< 1 ms	1 ms	1 ms
Packet Loss Probability	N/A	N/A	N/A	N/A	0.001	0.001
JAN						
End-to-End Delay	7 ms	100 ms	7 ms	655 ms	8 ms**	10 ms**
Packet Loss Probability	0	0.6239	0	0.9518	0.001*	0.005*

Note*: Upper bound based on user-perceived mouth-to-ear delay study

Note**: Upper bound influenced by packet loss probability

N/A: Not included in the model

7.3.5 Path C

Path C consists of two PANs and two IANs and one JAN.

The PAN component is stable with a very small contribution to network performance.

The IAN component, assuming a fixed node (FRV) delay, has as its major performance-affecting factors: the choice of MAC protocols, choice of packet size, number of PSCDs, and channel rate of the link. The packet loss probability, however, is affected significantly by the choice of MAC and the number of PSCDs.

The JAN component, like the IAN component, is significantly affected by the choice of MAC, number of PSCDs and FRVs, packet size, and link channel rate.

Table 20: Video Path C Network Performance Parameter Requirements

Video Path C	Calculated				Requirement	
	H.264		MPEG-2		H.264	MPEG-2
Details	**Minimum**	**Maximum**	**Minimum**	**Maximum**	**Upper Bound**	**Upper Bound**
End-to-End Delay	30 ms	∞	34 ms	∞	36 ms**	37 ms**

Table 20: Video Path C Network Performance Parameter Requirements (Continued)

Video Path C	Calculated				Requirement	
	H.264		MPEG-2		H.264	MPEG-2
Details	**Minimum**	**Maximum**	**Minimum**	**Maximum**	**Upper Bound**	**Upper Bound**
Packet Loss Probability	0.008	1	0.0303	1	0.001*	0.005*
PAN						
End-to-End Delay	<< 1ms	<< 1ms	<< 1ms	<< 1ms	1 ms	1 ms
Packet Loss Probability	N/A	N/A	N/A	N/A	0.001	0.001
IAN						
End-to-End Delay	11 ms	42 ms	11 ms	227 ms	13 ms	14 ms
Packet Loss Probability	0	0.3867	0	0.7804	0	0
JAN						
End-to-End Delay	10 ms	∞	14 ms	∞	10 ms	10 ms
Packet Loss Probability	0.004	1	0.0153	1	0	0

Note*: Upper bound based on user-perceived mouth-to-ear delay study

Note**: Upper bound influenced by packet loss probability

N/A: Not included in the model

7.3.6 Path D

Path D is similar to Path C, except it contains the EAN.

The PAN component is stable with a very small contribution to network performance.

The IAN component, assuming a fixed node (FRV) delay, has as its major performance-affecting factors: the choice of MAC protocols, choice of packet size, number of PSCDs, and channel rate of the link. The packet loss probability, however, is affected significantly by the choice of MAC and the number of PSCDs.

The JAN component, like the IAN component, is significantly affected by the choice of MAC, number of PSCDs and FRVs, packet size, and link channel rate.

We assume the EAN satisfies the network performance objectives given in ITU-T Y.1541 [20]. We use a network budget requirement for packet loss ratio of 10^{-3}, and an end-to-end transit delay of 100 ms given in ITU-T Y.1541.

Table 21: Video Path D Network Performance Parameter Requirements

Video Path D	Calculated				Requirement	
	H.264		MPEG-2		H.264	MPEG-2
Details	**Minimum**	**Maximum**	**Minimum**	**Maximum**	**Upper Bound**	**Upper Bound**
End-to-End Delay	59 ms	∞	68 ms	∞	65 ms**	67 ms**
Packet Loss Probability	0.008	1	0.0303	1	0.001*	0.005*
PAN						
End-to-End Delay	<< 1ms	<< 1ms	<< 1ms	<< 1ms	1 ms	1 ms
Packet Loss Probability	N/A	N/A	N/A	N/A	0.001	0.001
IAN						
End-to-End Delay	11 ms	42 ms	11 ms	227 ms	13 ms	14 ms
Packet Loss Probability	0	0.3867	0	0.7804	0	0
JAN						
End-to-End Delay	10 ms	∞	14 ms	∞	10 ms	10 ms
Packet Loss Probability	0.004	1	0.0153	1	0	0
EAN						
End-to-End Delay	20 ms	N/C	20 ms	N/C	100 ms***	100 ms***
Packet Loss Probability	0	N/C	0	N/C	0.001***	0.001***

Note*: Upper bound based on user-perceived mouth-to-ear delay study

Note**: Upper bound influenced by packet loss probability

Note***: Upper bound from ITU-T Y.1541

N/A: Not included in the model

N/C: Kept constant

7.3.7 Path E

Path E consists of two PANs, two JANs, and the EAN.

The PAN component is stable with a very small contribution to network performance.

The JAN component, assuming a fixed node (jurisdictional communication tower) delay, has as its major performance-affecting factors: the choice of MAC protocols, choice of packet size, number of PSCDs, and the channel rate of the link.

We assume the EAN satisfies the network performance objectives given in ITU-T Y.1541 [20]. We use a network budget requirement for the packet loss ratio of 10^{-3}, and an end-to-end transit delay of 100 ms given in ITU-T Y.1541.

Table 22: Video Path E Network Performance Parameter Requirements

Video Path E	Calculated				Requirement	
	H.264		MPEG-2		H.264	MPEG-2
Details	**Minimum**	**Maximum**	**Minimum**	**Maximum**	**Upper Bound**	**Upper Bound**
End-to-End Delay	33 ms	219 ms	33 ms	1330 ms	35 ms**	38 ms**
Packet Loss Probability	0	0.6239	0	0.9518	0.001*	0.005*
PAN						
End-to-End Delay	<< 1ms	<< 1ms	<< 1ms	<< 1ms	1 ms	1 ms
Packet Loss Probability	N/A	N/A	N/A	N/A	0.001	0.001
JAN						
End-to-End Delay	7 ms	100 ms	7 ms	655 ms	8 ms	9 ms
Packet Loss Probability	0	0.3867	0	0.7804	0.0005	0.0018
EAN						
End-to-End Delay	20 ms	N/C	20 ms	N/C	100 ms***	100 ms***
Packet Loss Probability	0	N/C	0	N/C	0.001***	0.001***

Note*: Upper bound based on user-perceived mouth-to-ear delay study

Note**: Upper bound influenced by packet loss probability

Note***: Upper bound from ITU-T Y.1541

N/A: Not included in the model

N/C: Kept constant

7.3.8 Path F

Path F consists of two PANs, one IAN, and one JAN.

The PAN component is stable with a very small contribution to network performance.

The IAN component, assuming a fixed node (FRV) delay, has as its major performance-affecting factors: the choice of MAC protocols, choice of packet size, number of PSCDs, and channel rate of the link. The packet loss probability, however, is affected significantly by the choice of MAC and the number of PSCDs.

The JAN component, assuming a fixed node (jurisdictional communication tower) delay, has as its major performance-affecting factors: the choice of MAC protocols, choice of packet size, number of PSCDs and FRVs, and the channel rate of the link.

Table 23: Video Path F Network Performance Parameter Requirements

Video Path F	Calculated				Requirement	
	H.264		MPEG-2		H.264	MPEG-2
Details	**Minimum**	**Maximum**	**Minimum**	**Maximum**	**Upper Bound**	**Upper Bound**
End-to-End Delay	20 ms	∞	24 ms	∞	23 ms**	24 ms**
Packet Loss Probability	0.0048	1	0.0184	1	0.001*	0.005*
PAN						
End-to-End Delay	<< 1ms	<< 1ms	<< 1ms	<< 1ms	1 ms	1 ms
Packet Loss Probability	N/A	N/A	N/A	N/A	0.001	0.001
IAN						
End-to-End Delay	11 ms	42 ms	11 ms	227 ms	13 ms	14 ms
Packet Loss Probability	0	0.3867	0.0002	0.7804	0	0
JAN						
End-to-End Delay	10 ms	∞	16 ms	∞	11 ms	17 ms
Packet Loss Probability	0.0008	1	0.0032	1	0	0

Note*: Upper bound based on user-perceived mouth-to-ear delay study

Note**: Upper bound influenced by packet loss probability

N/A: Not included in the model

7.3.9 Path G

Path G consists of two PANs, one IAN, two JANs, and the EAN.

The PAN component is stable with a very small contribution to network performance.

The IAN component, assuming a fixed node (FRV) delay, has as its major performance-affecting factors: the choice of MAC protocols, choice of packet size, number of PSCDs, and channel rate of the link. The packet loss probability, however, is affected significantly by the choice of MAC and the number of PSCDs.

The JAN component, assuming a fixed node (jurisdictional communication tower) delay, has as its major performance-affecting factors: the choice of MAC protocols, choice of packet size, number of PSCDs and FRVs, and the channel rate of the link.

We assume the EAN satisfies the network performance objectives given in ITU-T Y.1541 [20]. We use a network budget requirement for the packet loss ratio of 10^{-3}, and for an end-to-end transit delay of 100 ms given in ITU-T Y.1541.

Table 24: Video Path G Network Performance Parameter Requirements

Video Path G	Calculated				Requirement	
	H.264		MPEG-2		H.264	MPEG-2
Details	**Minimum**	**Maximum**	**Minimum**	**Maximum**	**Upper Bound**	**Upper Bound**
End-to-End Delay	48 ms	∞	53 ms	∞	52 ms**	53 ms**
Packet Loss Probability	0.0048	1	0.0184	1	0.001*	0.005*
PAN						
End-to-End Delay	<< 1ms	<< 1ms	<< 1ms	<< 1ms	1 ms	1 ms
Packet Loss Probability	N/A	N/A	N/A	N/A	0.001	0.001
IAN						
End-to-End Delay	11 ms	42 ms	11ms	227 ms	13 ms	14 ms
Packet Loss Probability	0	0.3867	0	0.7804	0	0
JAN						
End-to-End Delay	10 ms	∞	14 ms	∞	10 ms	10 ms
Packet Loss Probability	0.004	1	0.0153	1	0	0

Table 24: Video Path G Network Performance Parameter Requirements (Continued)

Video Path G	Calculated				Requirement	
	H.264		MPEG-2		H.264	MPEG-2
Details	Minimum	Maximum	Minimum	Maximum	Upper Bound	Upper Bound
EAN						
End-to-End Delay	20 ms	N/C	20 ms	N/C	100 ms***	100 ms***
Packet Loss Probability	0	N/C	0	N/C	0.001***	0.001***

Note*: Upper bound based on user-perceived mouth-to-ear delay study

Note**: Upper bound influenced by packet loss probability

Note***: Upper bound from ITU-T Y.1541

N/A: No included in the model

N/C: Kept constant

7.4 Summary for All Area Networks

Based on the results for speech and video applications in Section 7.2 and Section 7.3, a set of requirements are created. These involve the maximum allowable packet loss and delay for each type of area network to meet the end-to-end quality of service requirements for both applications. Table 25 lists the upper bounds for them. In the event that area networks are combined together, use the path upper bounds instead.

Table 25: Maximum Allowable Packet Loss and Delay for Each Type of Area Network

Speech and Video Paths	Requirement
Details	Upper Bound
PAN	
End-to-End Delay	1 ms
Packet Loss Probability	0.0001
IAN	
End-to-End Delay	14 ms
Packet Loss Probability	0
JAN	
End-to-End Delay	65 ms
Packet Loss Probability	0

Table 25: Maximum Allowable Packet Loss and Delay for Each Type of Area Network (Continued)

Speech and Video Paths	Requirement
Details	**Upper Bound**
EAN	
End-to-End Delay	100 ms
Packet Loss Probability	0.001

This page intentionally left blank.

Appendix A Glossary and Acronyms

A

ABA
America Bar Association

ACI
adjacent channel interference

A/D
analog-to-digital

ANSI
American National Standards Institute

ATIS
Alliance for Telecommunications Industry Solutions

ATSC
Advanced Television Systems Committee

AVI
Audio Video Interleave (file format)

B

BR
bit rate

B&W
black and white

C

C&I
Communications and Interoperability

CBR
constant bit rate

CC
color correction (a type of lens filter)

CCD
charge-coupled device

CCI
Co-channel interference

CIF
Common Intermediate Format

codec
coder-decoder

CRI
color rendering index

CSMA
carrier sensed multiple access

CSRC
Contributing SouRCe

D

DAT
digital audio tape

dB
decibel

dBA
dB A-weighted sound pressure level

DHS
Department of Homeland Security

DOC
Department of Commerce

DR
dynamic range

DSC
digital still camera

DU
dramatized urgency

E

EAN
extended area network

EC
error concealment

EMS
Emergency Medical Services

F

FCC
Federal Communications Commission

FEC
forward error correction

FFT
fast Fourier transform

fps
frames per second

FR
frame rate

FRV
first responder vehicle

FVA
forensic video analysis

G

GOP
group of pictures

H

H.264
Also known as, MPEG-4 Part 10, or AVC (advanced video compression). A digital video codec standard that achieves very high data compression.

HD
high-definition

HDTV
high-definition television

HID
high intensity discharge (a type of photography lamp)

HMI
halogen metal iodide (a type of photography lamp)

HRC
hypothetical reference circuit

Hz
Hertz

I

i
interlaced video display scan (e.g., 1080i)

i3a
International Imaging Industry Association

IAN
incident area networks

IEC
International Electrotechnical Commission

IEEE
Institute of Electrical and Electronics Engineers

IETF
Internet Engineering Task Force

I-frame
Intra-coded frame

IP
Internet Protocol

IPv6
Internet Protocol version 6

IR
Infrared

ISI
inter-symbol interference

ISO
International Standards Organization

ITU
International Telecommunication Union

ITU-R
ITU Telecommunication Standardization Radiocommunications

ITU-T
ITU Telecommunication Standardization Sector

J

JAN
jurisdictional area network

K

K
Kelvin

kHz
kilo-Hertz

L

LMR
land-mobile radio

LUT
look up table

LW per PH
line widths per picture height

M

MAC
Medium Access Control (a network protocol)

MMCN
Multimedia Computing and Networking Conference

MOS
Mean Opinion Score

MPEG
Moving Pictures Expert's Group

ms
millisecond

MTF
modulation transfer function

MTF50P
MTF, where contrast drops to 50 percent of its peak value

N

NAL
network abstraction layer

ND
neutral density (a type of lens filter)

NFPA
National Fire Protection Association

NS
network section

NTIA
National Telecommunications and Information Administration

NTP
normalized task performance

NTSC
National Television Systems Committee

O

OECF
optoelectronic conversion function

OIC
Office for Interoperability and Compatibility

P

p
progressive video display scan (e.g., 720p)

PAN
personal area network

PCM
pulse-code modulation

PDA
personal digital assistant

PLC
packet loss concealment

PLR
packet loss ratio

PS
packet size

PSCD
public safety communications devices

PSNR
peak signal-to-noise ratio

PS SoR
Public Safety Statement of Requirements

PSWAC
Public Safety Wireless Advisory Committee

PSWC&I
Public Safety Wireless Communications and Interoperability

Q

QCIF
Quarter CIF. See CCI.

QSIF
Quarter SIF. See SIF.

QVGA
Quarter VGA. See VGA.

R

RFC
Request for Comments

RFI
radio frequency interference

RTCP
Real-Time Control Protocol

RTP
Real-time Transport Protocol

S

SD
standard definition

SED
Systems Engineering and Development

SFR
spatial frequency response

SID
Speaker identification

SIF
Source Input Format

SMIA
Standard Mobile Imaging Architecture

SMPTE
Society of Motion Picture and Television Engineers

SNR
signal-to-noise ratio

SPL
sound-pressure level

S&T
Science and Technology

T

TCP
Transmission Control Protocol

TDM
time division multiplexing

TDMA
time division multiple access

TIA
Telecommunications Industry Association

TV
television

U

UCL
University College London

UDP
User Datagram Protocol

UNI
user-to-network interface

V

VAD
voice activity detection

VCL
video coding layer

VCEG
Video Coding Experts Group

VGA
Video Graphics Array

VHS
Video Home System (recording media)

W

WB
white balance

WFQ
weighted fair queuing

Appendix B References

B.1 Book and Standards References

[1] (ANSI, 2003) American National Standards Institute, ANSI T1.801.03-2003, "American National Standard for Telecommunications—Digital Transport of One-Way Video Signals—Parameters for Objective Performance Assessment."

[2] (ATIS, 2001) Alliance for Telecommunications Industry Solutions, ATIS T1.TR.74-2001, "Technical Report on Objective Video Quality Measurement Using a Peak Signal-to-Noise Ratio (PSNR) Full Reference Technique."

[3] (IEEE, 1999) Institute of Electrical and Electronics Engineers, Standard for Local and Metropolitan Area Networks, IEEE Std. 802.11, 1999 (Reaffirmed June 12, 2003), "Part 11: Wireless LAN Medium Access Control (MAC) and Physical Layer (PHY) Specifications."

[4] (IEEE, 2004) Institute of Electrical and Electronics Engineers, Standard for Local and Metropolitan Area Networks, IEEE Std. 802.16, 2004 (Including IEEE Std. 802.16-2001, IEEE Std. 802.16c-2002, and IEEE Std. 802.16a-2003), "Part 16: Air Interface for Fixed Broadband Wireless Access Systems."

[5] (IEEE, 2005) Institute of Electrical and Electronics Engineers, Standard for Local and Metropolitan Area Networks, IEEE Std. 802.3, 2005 (Revision of IEEE Std. 802.3-2002, including all approved amendments), "Part 3: Carrier Sense Multiple Access with Collision Detection (CSMA/CD) Access Method and Physical Layer Specifications."

[6] (IETF, 1980) The Internet Engineering Task Force, IETF Standard 6/RFC 768, "User Datagram Protocol."

[7] (IETF, 1998) The Internet Engineering Task Force, IETF RFC 2460, "Internet Protocol, Version 6 (IPv6) Specification."

[8] (IETF, 1999) The Internet Engineering Task Force, IETF RFC2680, "A One-Way Packet Loss Metric for IPPM."

[9] (IETF, 2003) The Internet Engineering Task Force, IETF RFC 3550, "RTP: A Transport Protocol for Real-Time Applications."

[10] (ISO, 2000) International Organization for Standardization, ISO 12233, "Photography—Electronic Still-Picture Cameras—Resolution Measurements."

[11] (ISO/IEC, 2000) International Organization for Standardization, ISO/IEC 13818 (commonly known as MPEG-2), "Information Technology—Generic Coding of Moving Pictures and Associated Audio Information."

[12] (ITU-R, 2002) Recommendations of the International Telecommunication Union, Radiocommunication Sector, ITU-R Recommendation BT.500, "Methodology for the Subjective Assessment of the Quality of Television Pictures."

[13] (ITU-R, 1995) Recommendations of the International Telecommunication Union, Radiocommunication Sector, ITU-R Recommendation BT.601, "Studio Encoding Parameters of Digital Television for Standard 4:3 and Wide-Screen 16:9 Aspect Ratios."

[14] (ITU-R, 2004) Recommendations of the International Telecommunication Union, Radiocommunication Sector, ITU-R Recommendation BT.1683, "Objective Perceptual Video Quality Measurement Techniques for Standard Definition Digital Broadcast Television in the Presence of a Full Reference."

[15] (ITU-T, 1988) Recommendations of the International Telecommunication Union, Telecommunication Standardization Sector, ITU-T Recommendation G.711, "Pulse Code Modulation (PCM) of Voice Frequencies."

[16] (ITU-T, 1990) Recommendations of the International Telecommunication Union, Telecommunication Standardization Sector, ITU-T Recommendation G.726, 1990, "40, 32, 24, 16 kbps Adaptive Differential Pulse Code Modulation (ADPCM)."

[17] (ITU-T, 1998) Recommendations of the International Telecommunication Union, Telecommunication Standardization Sector, ITU-T Recommendation P.931, "Multimedia Communications Delay, Synchronization, and Frame Rate Measurement."

[18] (ITU-T, 1999) Recommendations of the International Telecommunication Union, Telecommunication Standardization Sector, ITU-T Recommendation G.711 Appendix I, "A High-Quality Low-Complexity Algorithm for Packet Loss Concealment with G.711."

[19] (ITU-T, 2002) Recommendations of the International Telecommunication Union, Telecommunication Standardization Sector, ITU-T Recommendation G.113 Appendix I, "Provisional Planning Values for the Equipment Impairment Factor Ie and Packet-Loss Robustness Factor Bpl."

[20] (ITU-T, 2002) Recommendations of the International Telecommunication Union, Telecommunication Standardization Sector, ITU-T Recommendation Y.1541, "Network Performance Objectives for IP-Based Services."

[21] (ITU-T, 2002) Recommendations of the International Telecommunication Union, Telecommunication Standardization Sector, ITU-T Recommendation Y.1541, Appendix X, "Speech Quality Calculations for Y.1541 Hypothetical Reference Paths."

[22] (ITU-T, 2003) Recommendations of the International Telecommunication Union, Telecommunication Standardization Sector, ITU-T Recommendation G.114, "One-Way Transmission Time."

[23] (ITU-T, 2004) Recommendations of the International Telecommunication Union, Radiocommunication Sector, ITU-T Recommendation J.144, "Objective Perceptual Video Quality Measurement Techniques for Digital Cable Television in the Presence of a Full Reference."

[24] (ITU-T, 2005) Recommendations of the International Telecommunication Union, Telecommunication Standardization Sector, ITU-T Recommendation G.107, "The E-Model, a Computational Model for Use in Transmission Planning."

[25] (ITU-T, 2005) Recommendations of the International Telecommunication Union, Radiocommunication Sector, ITU-T Recommendation H.264, "Advanced Video Coding for Generic Audiovisual Services."

[26] (NFPA, 2003) National Fire Protection Association, NFPA 1901, "Standard for Automotive Fire Apparatus."

[27] (NFPA, 2005) National Fire Protection Association, NFPA 1561, "Standard on Emergency Services Incident Management System."

[28] (PS SoR Volume I, version 1.2, 2006) *Public Safety Statement of Requirements for Communications & Interoperability*, Volume I: Qualitative, Version 1.2, October 2006.

[29] (Rom, 1990) R. Rom. and M. Sidi, *Multiple Access Protocols: Performance and Analysis*, Springer-Verlag, New York, 1990.

[30] (SMPTE, 1997) Society of Motion Picture and Television Engineers, SMPTE 259M, "Television—10-Bit 4:2:2 Component and 4fsc Composite Digital Signals—Serial Digital Interface."

[31] (Ziemer, 1985) R. E. Ziemer and W. H. Tranter, *Principles of Communications: Systems, Modulation, and Noise*, Houghton Mifflin Co., Boston.

B.2 Online References

[32] Department of Homeland Security, Public Safety Communications Technical Report, "Measurement of Speech Transmission Suitability," DHS-TR-PSC-07-01, October 2007. Available at: http://www.safecomprogram.gov/SAFECOM/.

[33] Department of Homeland Security, Public Safety Communications Technical Report, "Video Acquisition Measurement Methods," DHS-TR-PSC-07-02, October 2007. Available at: http://www.safecomprogram.gov/SAFECOM/.

[34] Department of Homeland Security, Public Safety Communications Technical Report, "Network Measurement Methods," DHS-TR-PSC-07-03, October 2007. Available at: http://www.safecomprogram.gov/SAFECOM/.

[35] Department of Homeland Security, Public Safety Communications Technical Report, "Tactical and Surveillance Video Quality Experiments," DHS-TR-PSC-07-04, October 2007. Available at: http://www.safecomprogram.gov/SAFECOM/.

[36] Department of Homeland Security, Public Safety Communications Technical Report, "Speech Intelligibility and Detection of Voice Characteristics," DHS-TR-PSC-08-05, August 2008. Available at: http://www.safecomprogram.gov/SAFECOM/.

[37] Department of Homeland Security, Public Safety Communications Technical Report, "Task-based Live and Recorded Surveillance Video Quality Tests," DHS-TR-PSC-08-06, September 2008. Available at: http://www.safecomprogram.gov/SAFECOM/.