QOS AND SECURITY CONSIDERATION FOR PCS SERIES VIDEOCONFERENCING SYSTEM

Table of Contents

1. QoS

On the current Internet, an estimated 5% of packets are lost. Packet loss causes video frame freeze, error propagation and audio glitches in video conferencing sessions. The QoS (Quality of Service) and packet loss robustness features are very important in video conferencing in lossy "best-effort" IP network environments.

The main objective of QoS features used in conventional video conferencing endpoints is to "conceal" the error. For example, certain technology is used to shuffle image data for transmission so that packet loss does not destroy wide areas of the video image. Other technology is used to increase "intra" macroblocks that do not require previous frame information, so that the packet loss error does not propagate over a long period but the increased amount of intra macroblocks requires more bits to carry video frames, resulting in a lower frame rate.

Sony has developed new QoS features using a different approach. Sony's approach is to "recover" the error, not to "conceal" it, thereby maintaining the quality of real-time communication. With the Sony PCS terminal and its QoS features, users need not be concerned about bandwidth and network quality; the PCS terminal automatically adjusts its bandwidth, buffering size and algorithm to maintain high-quality conferencing.

The following two major functions are implemented on the PCS terminals; PCS-1, PCS-11, PCS-TL30, PCS-TL50, PCS-G50 and PCS-G70 to provide the above-mentioned features:
— Real-time ARQ (Automatic Repeat reQuest)

— ARC (Adaptive Rate Control)

Furthermore the PCS-G50 and PCS-G70 support FEC (Forward Error Correction). FEC is a function which directly recovers the lost data by using parity packets attached to the data.

Firstly, the newly developed functions are described in the following sections. The mechanism to select those functions according to the condition of the network is then described in the next section. Network level QoS, IP Precedence, Type of Service and Differentiated Services, which are commonly used to deliver high-quality video and audio service over IP networks and are adopted in the PCS terminals, are then briefly introduced in the following section.

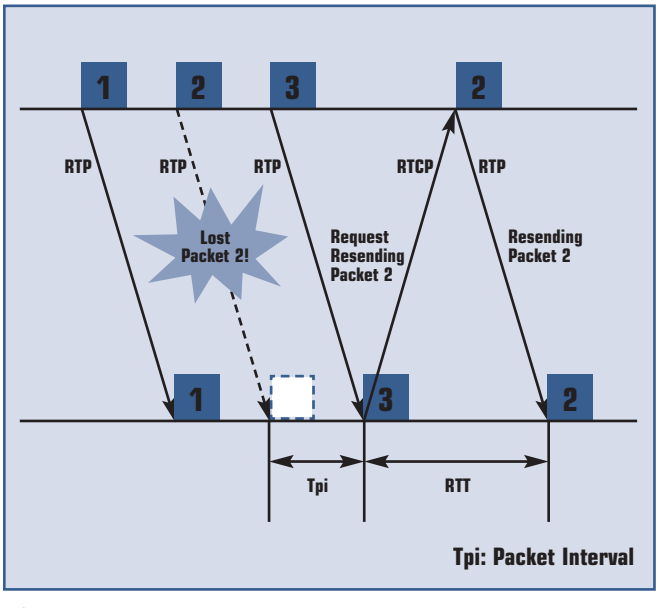
1.1. Real-time ARQ (Automatic Repeat reQuest)

The Real-time ARQ mechanism automatically resends lost packets and reorders received packets. Thanks to Real-time ARQ, the PCS terminal can recover almost completely from a packet loss rate of 10% while maintaining the minimum latency adaptively determined according to the network environment.

Figure 1.1 describes how Real-time ARQ works. The packet loss of an RTP packet of a video/audio bit stream is detected on the receiver side. A "resend request" is then sent to the transmission side using an RTCP (RTP Control Protocol) packet. On the transmission side, the transmitted packet is held, in preparation for resending according to the resend request.

By using an RTCP packet, the Round-Trip Time (RTT), or network latency, can be measured. If the RTT is large, it would be a waste of network traffic to send the "resend request" knowing that the resent packet would not arrive in time for decoding. Sony's Real-time ARQ mechanism is capable of determining whether or not the resending packet would arrive in time, and adaptively selects the optimum algorithm according to the RTT and packet loss rate.

Figure 1.1 Real-time ARQ resending diagram



A) Minimum latency with optimum packets buffering

In general, resending packets requires a receiver buffer for reordering of packets, and hence increases the system latency. Sony's PCS terminal has a variable-length reordering buffer for rearranging a packet in order of the sequence number of RTP headers. The size of the reordering buffer is optimized according to the measured network RTT and PLR (Packet Loss Rate). In an environment without packet loss, the reordering buffer is minimal and hence the communication latency is not affected.

B) Adaptive algorithm switching

In addition to the buffer size, the resending algorithm itself is adaptively switched according to the RTT and PLR. For example, if the network latency (RTT) is very large, packet resending is not used, but audio redundant transmission is used instead so that smooth conversation can be maintained.

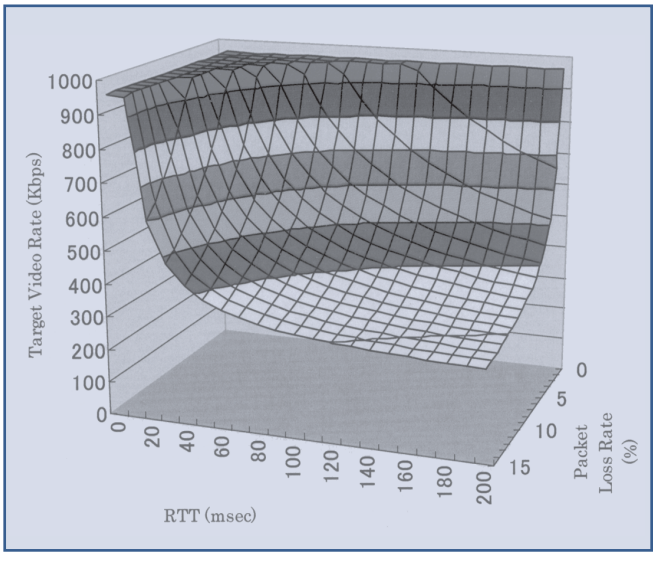
C) MCU and presentation data

Real-time ARQ works in combination with the internal MCU function. When the MCU function is used, each link between the terminals and the PCS terminal with MCU function uses ARQ to recover from packet losses. Additionally, when the PCSA-DSB1S and PCSA-DSM1 (Data Solution Box) are used, presentation data can also be recovered from packet losses.

1.2. Adaptive Rate Control (ARC)

Adaptive Rate Control is a mechanism used to slow down the video bit rate if network congestion occurs. Unlike conventional similar methods, the PCS terminal adopts TCP-friendly rate control in consideration of adjustment to other TCP traffic (e.g. FTP) which has a rate control function itself. TCP-Friendly Rate Control (TFRC) is designed for unicast flows operating in the Internet environment and competing with TCP traffic. The TCP throughput equation in IETF RFC3448 is a function of packet loss rate and RTT should be suitable for use in TFRC. Figure 1.2 is an example of the calculated target video bit rate.

Figure 1.2 Adaptive Rate Control (for example, max_video_rate is 960Kbps)



1.3. Forward Error Correction (FEC) for Real-time Communications

FEC is another mechanism for recovering packet loss on the Internet.

The idea of FEC across the packets is to transmit the parity packets for the receiver to reconstruct lost packets. The Reed-Solomon (RS) code is selected in our system, which is perfectly suited for recovering erasure errors such as packet loss. Across the packets, RS *(n, k)* encodes *k* information symbols (each symbol per packet) into *n* symbols so as to construct the *n-k* parity packets. Here, the symbol size of all RS codes is set to eight for the convenience of accessing information in bytes.

Figure 1.3 shows an example of the recovery procedure.

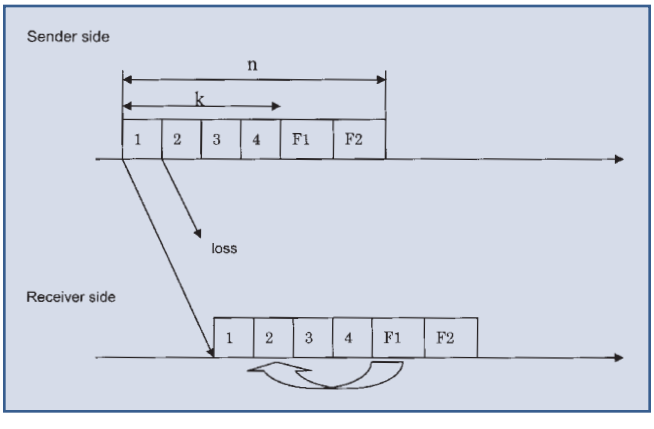
The first four packets, numbered 1 to 4, are the data packets, and the latter two packets, F1, F2, are the parity packets, forming one RS block. By using the FEC, a receiver can recover any lost pattern where the number of lost packets is no more than 2.

For instance, if packet 2 is lost, the receiver can recover it by using the F1 packet.

Compared with the Real-time ARQ mechanism, the performance of FEC is not affected by the amount of RTT; this is one of the advantages of FEC that allows it to be used in a long RTT environment. By constructing a short packet matrix of RS block and adapting an original, fast calculating RS algorithm, FEC decoding can be performed quickly so that PCS terminals can use FEC for real-time communication applications.

Since FEC always sends the parity packets, the overhead of the transmission rate is larger than that of the Real-time ARQ transmission.

Figure 1.3 FEC recovery diagram



1.4. Selection of QoS mode

As discussed in the previous sections, the three modes are selected according to the setup of QoS setting in PCS-G50 and PCS-G70. The following four modes are supported and can be selected by the user. FEC is supported only for video data.

Mode	Operation
ARQ	Only ARQ is activated
FEC	Only FEC is activated.
FEC & ARQ	Both ARQ and FEC are activated
Hybrid	ARQ/FEC/FEC & ARQ mode are selected according to the value of RTT

1.4.1. ARQ mode

ARQ mode for video data has two states: Standby and Active.

ARQ mode for audio data has three states: Standby, Active, and Audio double transmission.

In audio double transmission mode, audio data is doubly transmitted in the network with large delay to prevent any increase of the delay caused by retransmitted packets. When ARQ mode is set to on, ARC is automatically set to on.

1.4.2. FEC mode

The FEC mode is enabled when both the sender side and the receiver side are in the FEC mode. ARC is automatically turned on in the FEC mode. The ratio of the number of original packets to the number of parity packets is fixed.

1.4.3. FEC & ARQ mode

In mixed mode, the packet loss is recovered by using ARQ mode when the packet loss cannot be recovered by using FEC. This mode becomes active when both FEC and ARQ mode are set to on.

Because the number of original packets and that of the parity packets in a single FEC block are fixed in FEC, it is easy to identify whether the packet loss can be recovered or not by simply looking at the number of packets in the block.

Where the received original packet is S_ORG, the original packet is ORG, and the received parity packet is PAR, the packet loss cannot be recovered if the following equation becomes true:

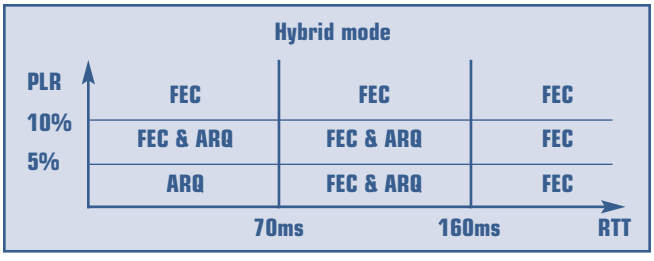
$$S\_PAR - (ORG - S\_ORG) < 0$$

Because the packet loss cannot be recovered, a resend request is issued.

1.4.4. Hybrid mode

When the terminal is set to Hybrid on and the remote party is also set to Hybrid on, Hybrid mode can be activated. When Hybrid mode is activated, the modes are selected automatically as shown in the following figure. When both the sender and receiver are set up in the Hybrid mode, the system works in the Hybrid mode.

If the receiver is set up in any mode other than the Hybrid mode, the system works in the FEC & ARQ mode.



The threshold values of the above figure might be changed in the future versions.

1.5. Network level QoS

The PCS terminal can enter the values of IP Precedence, Type of Service and Differentiated Services. The TOS (Type of Service) field in the IP header is used for either defining IP Precedence and Type of Service or DSCP (Differentiated Services Code point) bits of Differentiated Services. The usage of the field for either service is up to the service administrator of the network.

• IP Precedence and Type of Service

0	1	2	3	4	5	6	7
Precedence			Delay	Through-put	Reliability	Minimum Cost	CU

CU: Currently Unused

The value of IP Precedence and Type of Service can be defined in the Setup menu.

IP Precedence bits The IP precedence bits indicate the priority with which a packet is handled as follows:

1 1 1 : Network control

1 1 0 : Internetwork control

1 0 1 : Critical

1 0 0 : Flash override

0 1 1 : Immediate

0 1 0 : Priority

0 0 1 : Routine

Delay bit Used when the time taken for a datagram to travel from the source PCS terminal to the destination PCS terminal or host (latency) is important. The network provider is requested to select a route with the minimum delay when this bit is set to on.

Throughput bit Used when the volume of data transmitted in any period of time is important. The network provider is requested to select a route producing the maximum throughput when this bit is set to on.

Reliability bit Used when it is important that the data arrives at the destination without requiring retransmission. A network provider is requested to select a route with maximum reliability when this bit is set to on.

Minimum cost bit Used when minimizing the cost of data transmission is important. The network provider is requested to have datagrams routed via the lower-cost route when this bit is set to on.

Unless the value of this field is set by the administrator, the default value of this field is all 0.

• Differentiated Services

The Differentiated Services (Diffserve) architecture is based on a network model implemented over a complete Autonomous System (AS) or domain. As this domain is under administrative control, it is possible to take provisions to establish clear and consistent rules to manage traffic entering and flowing through the networks that conform to the domain. Diffserve defines a field in the IP header called the Differentiated Services Code point (DSCP) which is a six-bit field as shown below.

0	1	2	3	4	5	6	7
DS5	DS4	DS3	DS2	DS1	DS0	CU	CU

CU: Currently Unused

The type of service field of the IP header is used to define DSCP. Hosts in a network that supports DiffServe make each packet with a DSCP value. Routers within the DiffServe network use these values to classify the traffic into distinct service classes according to DSCP value. Thus, routers control packets not on a flow-by-flow bases but by traffic classes based on DSCP marking. Since the

routers are not required to maintain any elaborate state information to identify the flows, the routers can handle a large number of flows. PHB (Per Hop Behavior) is defined according to the traffic classes based on DSCP marking. For example, if the routers receive packets with DSCP = 101110, which means expedited forwarding (EF), the routers are requested to forward the packets for low latency and low loss service. Assured Forwarding (AF), which defines three levels of drop precedence is also defined and is used to help provide the necessary minimum bandwidth.

The default value of this field when the field is used as Differentiated Services is 000000, meaning simply that the best effort service is provided.

2. Security

- Secure communication is established on IP and ISDN connection
- Secure communication is authenticated by password
- Access to IP services is controlled by password
- Administrative operations are protected by password
- Incoming call can be rejected

2.1. Secure communication is established on IP and ISDN connection

Sony videoconferencing equipment achieves secure communication on H.323 and SIP connections by encrypting audio and video stream data. Sony equipment encrypts PC presentation data via DSB (Data Solution Box) as well as audio and video stream. The encryption algorithm is AES (Advanced Encryption Standard) the operational mode is CBC (Cipher Block Chaining), and the cipher key for the AES algorithm is 128 bits.

AES algorithm has been selected by NIST (National Institute of Standards and Technology) as its next-generation encryption algorithm and provides improved security compared to DES (Data Encryption Standard) algorithm. The main operations of AES algorithm are permutations and substitutions and the length of the encrypted data is the same as that of the input data. Therefore, AES operations do not affect the packet size of the media stream. The AES algorithm of Sony equipment has been tested with the test data NIST provides.

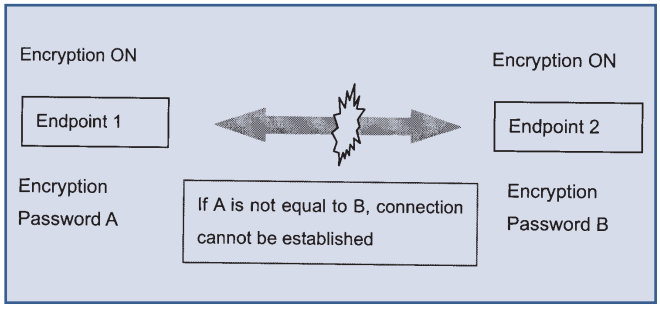
Even more, Sony products support ITU-T standardized encryption procedures H.235. Sony products are able to do secure communication with other manufactures' videoconferencing products. In H.235 mode the packet size will increase because the mode of AES changes.

Sony equipment also utilizes encryption function on H.320 connections. ISDN connection is a closed connection and the security level is much higher than an IP connection. And Sony products use proprietary operating system and construct proprietary firmware on this OS. Therefore, even though attackers could access Sony equipment via ISDN network, they cannot invade the IP network.

Sony products support encryption procedures on H.320 connections conforming to ITU-T standards H.233 and H.234 so that they can securely communicate with other vendors' videoconferencing products.

2.2. Secure communication is authenticated by password

Sony equipment requires the password when doing secure communication and the unit has to be on. Users must enter a password with 13 to 20 alphabetical characters, numerals, or symbols. If the passwords at both sites do not match, communication cannot be established. This means secure communication is authenticated by password.



2.3. Access to IP services is controlled by password

Sony endpoints provide several IP services such as FTP, Telnet, HTTP and SNMP. Access to these services is controlled by password. Users must enter the user name and the password to log on to these services. SNMP requires 'Community' name input instead of user name and password. The remote firmware upgrade function utilizes FTP service and users must enter the user name and the password to upgrade the firmware via IP network.

2.4. Administrative operations are protected by password

Sony endpoints support the following passwords. It is assumed that a system administrator will set these passwords up on the setup menu if necessary.

— Administrator Password

In case a system administrator sets the 'Administrator Password', users must enter the password in order to change any settings in the setup menu or phonebook.

— Superuser Password

In case a system administrator sets the 'Superuser Password', users must enter the password in order to make any changes to the phonebook.

— Remote Access Password

In case a system administrator sets the 'Remote Access Password', users must enter the user name and the password to access the endpoint via a web browser in order to control it. It is also possible to access it via a web browser as administrator or superuser.

PCS-G50 and PCS-G70 models provide 'Phone Book Modification Password' and 'Save Settings Password' instead of 'Superuser Password'.

2.5. Incoming call can be rejected

If users do not want to connect with an incoming call, they can select manual answer mode and reject those incoming calls. And users can reject all incoming calls during a multipoint conference by the setting they select.