



# Cisco ClearPath

## Whitepaper



D1498501

August 2012

# Contents

<b>1</b>	<b>Introduction</b>	<b>3</b>
<b>2</b>	<b>What is ClearPath?</b>	<b>4</b>
2.1	ClearPath design goals.....	4
<b>3</b>	<b>Packet loss descriptions</b>	<b>5</b>
<b>4</b>	<b>How ClearPath works</b>	<b>7</b>
4.1	Dynamic bit rate adjustment .....	7
4.2	Long Term Reference Frames .....	9
4.3	Forward Error Correction .....	10
4.4	Packet loss technology comparison.....	13
4.5	Packet loss scenarios, and example technologies used.....	14
<b>5</b>	<b>Additional technologies</b>	<b>15</b>
5.1	Packet pacing.....	15
5.2	Disposable frames.....	15
5.3	Packetization .....	15

## Figures

Figure 1: How ClearPath uses the RTCP Receive Report.....	8
Figure 2: Video stream with normal infra-frame repair .....	9
Figure 3: Video stream with ClearPath LTRF and Repair-P frames .....	10
Figure 4: Video stream with no FEC applied.....	11
Figure 5: Video stream with added video aware FEC Level 1 .....	11
Figure 6: Illustrates different types of FEC levels .....	12

# 1 Introduction

With the introduction of High Definition (HD) video and very high quality telepresence experiences, user expectations of video quality have dramatically increased when communicating visually. Users expect to communicate using telepresence with internal colleagues and with external customers, suppliers and partners, in addition to staying visually connected while on the road – all while maintaining the best experience possible. Today, with video becoming pervasive and more and more people using video as mobile workers, the type and quality of the connected network is often unknown. In addition Business to Business communication is increasingly using the public internet, which has no guarantee of network quality. The quality of the call will then depend on the enterprises' links to the Internet, as well as the Internet backbone transport quality.

As a high quality network is the most important requirement for a high quality video call, making telepresence calls in a non-optimal IP network environment can degrade the experienced audio and video quality. There are a number of mechanisms on the network layer that can address these challenges, and these should be implemented where possible. In cases where these mechanisms are not available, for instance over the Internet, Cisco's ClearPath technology comes into play. ClearPath defines a set of media resilience mechanisms that greatly increase the audio and video quality experienced by the user in the event of disruptive network conditions.

## 2 What is ClearPath?

Packet loss occurs for various reasons; therefore ClearPath is a dynamic technology that implements a number of media resilience mechanisms that reduce the negative effects of packet loss. Good meeting experiences in conditions with up to 10% packet loss can be achieved by the ClearPath implementation. The media resilience mechanisms in ClearPath can be divided into three main sets of tools:

- ▶ **Dynamic bit rate adjustment** related to the severity of packet loss. This entails adapting the call rate to the variable bandwidth available: downspeeding or upspeeding the call based on the packet loss experienced.
- ▶ **Long Term Reference Frames:** A method for the encoder and decoder to re-synchronize after packet loss without the use of an intra-frame. A repair P-frame can be used instead of a traditional intra-frame when packet loss occurs, resulting in approximately 90% less data being transmitted to rebuild the frame.
- ▶ **Video aware Forward Error Correction (FEC):** Protecting the most important data (typically the repair P-frames) using redundancy to make sure that the far end receives them.

ClearPath uses the technologies above where appropriate to provide the best possible user experience.

ClearPath is designed to be independent of the call setup protocol, and can be used by endpoints using H.323, SIP and XMPP. All the media resilience mechanisms within ClearPath result in an encoded bit stream that is H.264 compliant.

### 2.1 ClearPath design goals

ClearPath is designed to:

- ▶ React quickly to changes in available bandwidth.
- ▶ Make the quality (as experienced by a user) appear more stable over time when there is packet loss.
- ▶ Detect whenever a significant reduction in produced bit rate does not reduce the packet loss observed on the link, and not reduce bit rate in this case.

## 3 Packet loss descriptions

Loss is defined as packets that did not arrive at the decoder (that is, they were dropped somewhere along the network path). A telepresence system measures packet loss by comparing the sequence numbers of the RTP packets it receives against the sequence numbers it expected to receive. Packet loss can occur anywhere along the network path for a variety of reasons. Some common reasons are:

- ▶ Layer-1 errors on the physical interfaces and cables along the path, such as a malfunctioning optical interface.
- ▶ Mis-configured network interfaces along the path, such as Ethernet speed/duplex mismatches between two devices.
- ▶ Bursts of packets exceeding the buffer (queue) limit or policer configurations on network interfaces along the path, such as Ethernet switches with insufficient queue depth or oversubscribed backplane architectures, or WAN router interfaces which police traffic to conform to a Service Provider's contractual rates.
- ▶ A poor wireless network connection either because of distance to the access point or general network congestion.

A closely related metric is late packets, which are packets that arrive but exceeded the decoder jitter buffer (that is, they arrived too late to be decoded) and therefore were discarded (dropped) by the receiving telepresence system. Lost packets and late packets both result in the same outcome - reduction in video quality.

The traditional approach to solving these issues could be divided into three types of packet loss technologies:

- ▶ Decoder concealment - hiding artifacts when receiving video with packet loss.
- ▶ Down speeding - reducing the call rate to use less bandwidth.
- ▶ Use of intra-frames - the traditional way of handling video with packet loss.

Decoder concealment can lead to visual results such as a partially complete frame, discontinuities in the frame, motion jump-back or the build up of divergence artifacts over time.

Down speeding can eliminate packet loss in cases where the root cause of congestion is over-subscription and the bandwidth used by the call is making a significant impact. Even in cases of random packet loss, a lower rate call should not have as many indications of packet loss as a higher bandwidth call with the same packet loss rate.

An intra-frame is a frame sent from encoder to decoder that has all of the information required to build an image without reference to earlier frame. Sending an intra-frame either necessitates sending a lot more information and therefore spending more processing time and bandwidth than a normal frame, or limiting the bits and time spent on the frame by sending a frame of lower than usual quality.

Occasional intra-frames are not intrusive to the user; however when they are used to recover from packet loss they often are. Commonly, when sending an intra-frame, a lower quality is chosen that can result in pixilation. Using a larger frame of better quality would be paid for by frozen video

(probably showing a broken frame) while the larger frame is transmitted. Recovery frames may experience loss themselves, leading to a chain of frames sent while multiple attempts to resynchronize take place, resulting in a pulsing of video commonly associated with packet loss.

## 4 How ClearPath works

This section discusses the different media resilience mechanisms within ClearPath in more detail.

### 4.1 Dynamic bit rate adjustment

The need for dynamic bit rate adjustment can be understood with two common examples:

- ▶ Typical networks (public networks in particular) offer an available capacity for carrying UDP data that varies with time. Generating more traffic than the network can handle leads to packet loss, possibly having unfavorable effects on the audio/video channels rendered at the receiving end.
- ▶ Overly aggressive bandwidth settings at call start up may cause the produced (or received) traffic to overshoot the capacity of the current host network. Typically this happens if a system is moved from one network (for example, a corporate LAN) to a different network (such as a home DSL connection) without making changes to either provisioned or configured bitrates.

The traditional approach to solving these issues is to have systems observe packet loss on incoming streams and issue a request from the decoder for the far end system to perform "down speeding" (reducing the call rate to a new lower value) if packet losses are severe and persist over time. The requests to the far end system are either sent as a SIP Re-Invite message or an H.323 Flow Control message, depending on the system's native protocol. Some of the issues that are experienced with this approach include interworking control messages (SIP/H.323) and delay/responsiveness. Down speeding a call will ultimately result in a lower call resolution and therefore quality.

Most modern implementations now support RTCP; therefore it is possible to change this approach from having the receiver actively request call rate changes to making the sender side take action proactively. By inspecting RTCP Receiver Reports (hereafter called RRs), the sender is regularly notified of the reception statistics at the far end system. If RRs indicate packet losses, it is possible that the network is overloaded and that the traffic generated by the sender should be reduced.

With ClearPath, the sender side can automatically adjust its produced bit rate by using RRs (see the figure below). An advantage of this approach is that it does not involve any additional signaling (decisions are taken at the sender side without explicitly informing the receiver). Also, if RRs are scheduled fairly frequently (~5sec intervals are common), this potentially provides a system that is more responsive to changes in available capacity. Cisco also uses RRs to dynamically increase the bandwidth subsequent to down speeding when packet loss is no longer detected.

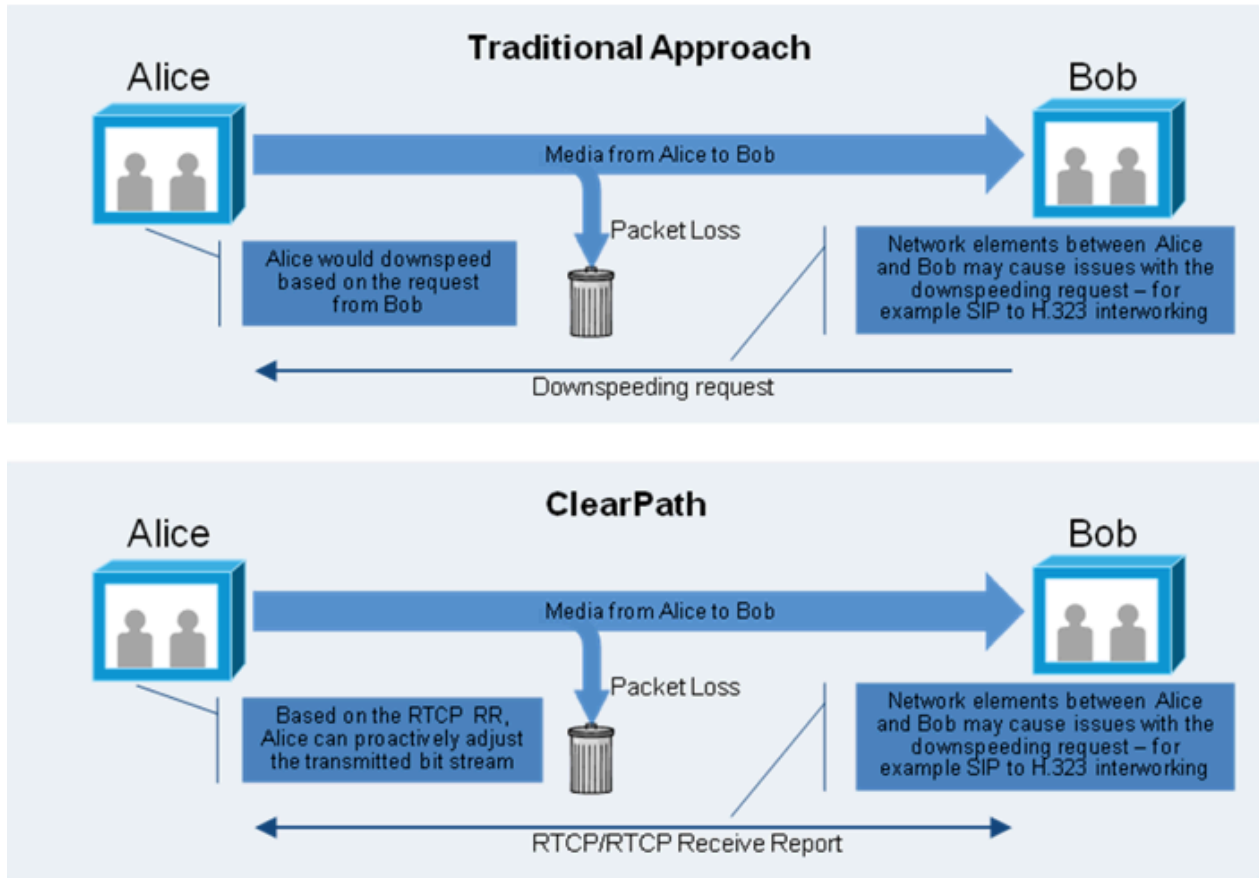


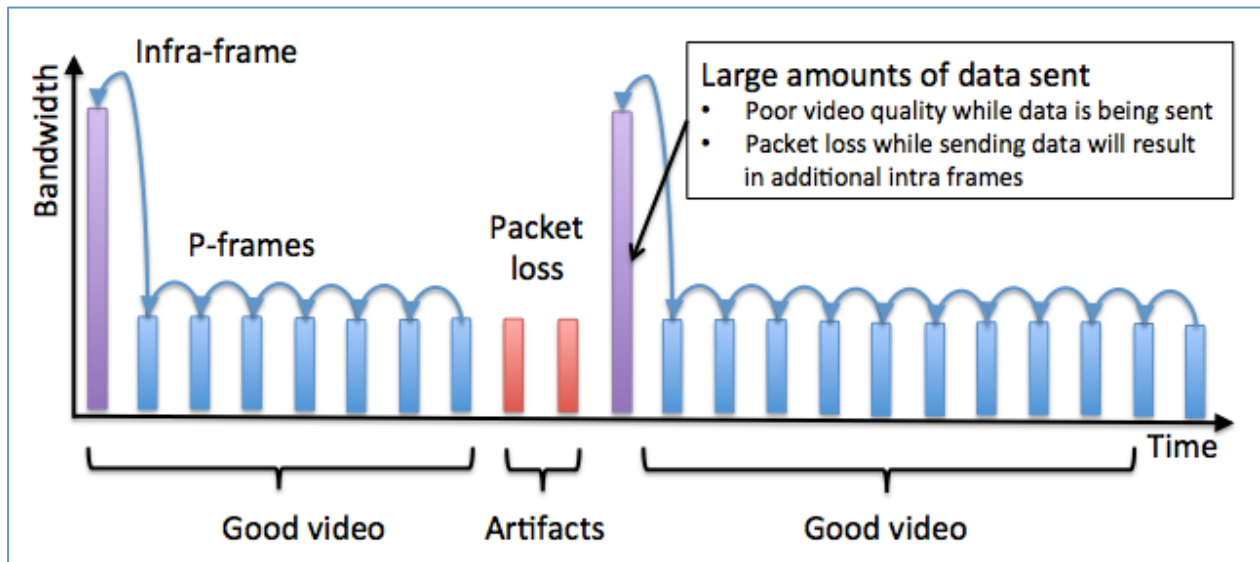
Figure 1: How ClearPath uses the RTCP Receive Report



## 4.2 Long Term Reference Frames

A Long Term Reference Frame (LTRF) is a reference frame that is stored in the encoder and decoder until they receive an explicit signal to do otherwise. (Up to 15 LTRFs are supported by H.264.)

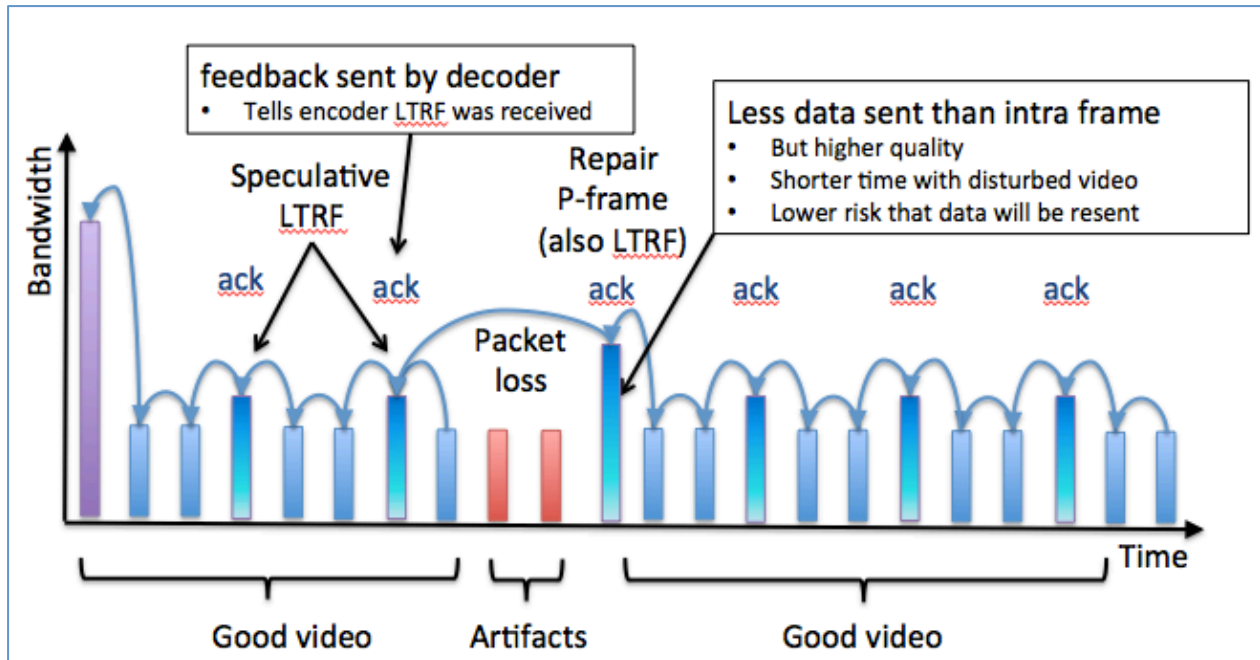
Typically (without LTRF) an intra-frame is used for encoder/decoder re-synchronization after packet loss (see the figure below).



**Figure 2: Video stream with normal infra-frame repair**

This is where LTRFs can provide benefits over normal infra-frames as an alternative method for encoder/decoder re-synchronization. Typically, the encoder inserts LTRFs periodically and at the same time instructs the decoder to store one or more of those LTRFs (see the figure below).

A repair P-frame uses a previous LTRF that has been decoded correctly as a reference. The repair P-frame is used in response to a missing frame or of its reference frame. Because the acknowledged LTRF is known to have been correctly received at the decoder, the decoder is known to be back in-sync if it can correctly decode a repair P-frame.



**Figure 3: Video stream with ClearPath LTRF and Repair-P frames**

Many video systems such as broadcast systems lack a backchannel and therefore cannot implement adaptive and feedback-based resilience mechanisms. However, because telepresence is a two way communication ClearPath is able to use a backchannel in order to allow the LTRF mechanism to work.

While an encoder can speculatively send LTRFs as a basis for future repair, it cannot reliably use them unless they are known to have been successfully received. This is achieved with a backchannel which acknowledges to the encoder the successful receipt and decode of LTRFs on which future repair can be based, and also signals when such frames have been lost. ClearPath includes mechanisms for repeat messaging and notification of a frame that the decoder did not receive, such that both ends are in agreement about which frames can be used for repair.

Recovery involves encoding new frames in the sequence based on a retained reference frame that does not immediately precede the new frames. LTRF recovery provides an exact and complete repair that is permanent going forward, unless there is further loss in subsequent frames.

---

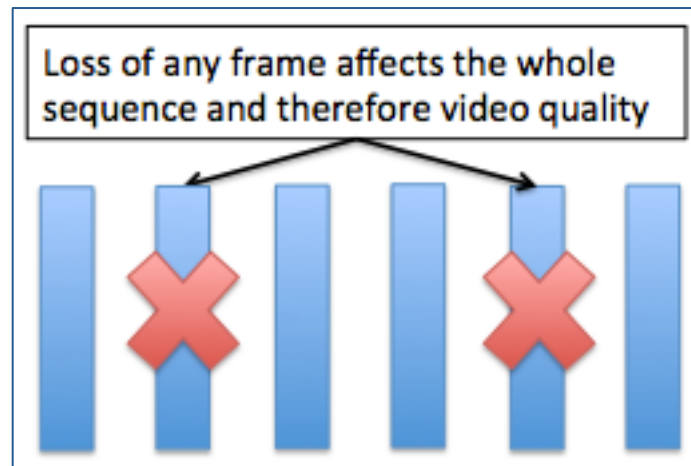
**Note:** This is not the same as decoder substitution of a received frame for the intended reference frame, which is a method that can be used for short-term concealment of loss. This provides a facsimile of the intended frame, but it is not identical to the frame at the encoder, and over time the two sequences will diverge, resulting in unpleasant visual artifacts.

---

### 4.3 Forward Error Correction

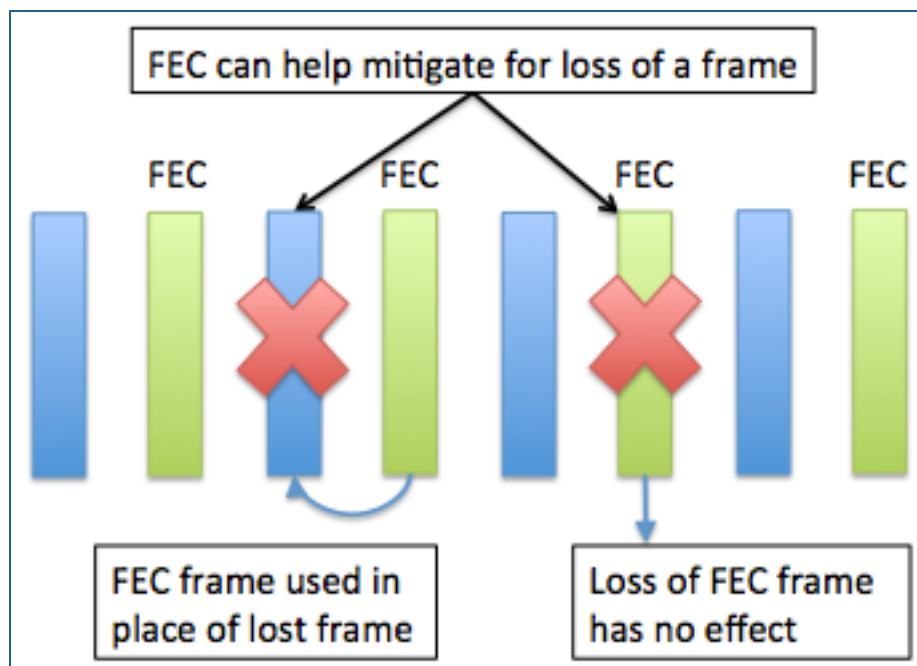
Forward Error Correction (FEC) is accomplished by adding redundancy to the transmitted information using a predetermined algorithm. This is achieved by sending redundant video data in band alongside normal video data (see the figures below). The redundancy allows the receiver to detect and correct a limited number of errors occurring anywhere in the message without the need to ask the sender for additional data. FEC gives the receiver an ability to correct errors without needing a

reverse channel to request retransmission of data, but this advantage is at the cost of a fixed higher forward channel bandwidth.



**Figure 4: Video stream with no FEC applied**

ClearPath uses a fixed percentage of the bandwidth for FEC/redundancy. FEC does not affect the overall latency of a call and may be disabled if it is not efficiently resolving packet loss problems.



**Figure 5: Video stream with added video aware FEC Level 1**

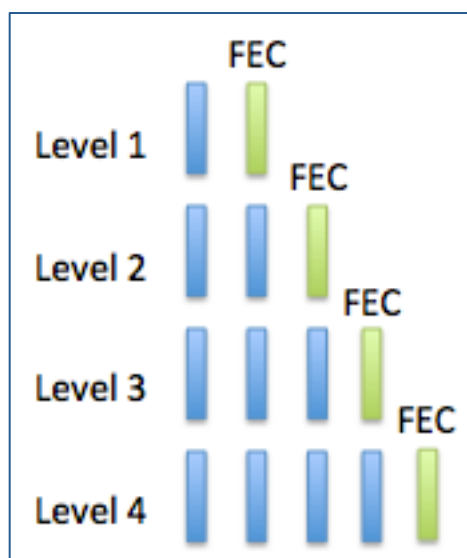
Within ClearPath, different FEC levels are used dynamically, based on the loss conditions; different levels of protection are applied to different classes of frame based on their importance in enabling sequence repair (see section 4.2).

A priority order of “FEC actions” exists in low to high priority; for example, “FEC protect Repair-P frames” might be the highest priority and “FEC protect long term reference frames” next, through to “FEC protect disposable frames” as low priority.

Example: The video encoder tags the video packets to be one of:

- ▶ High importance
- ▶ Medium importance
- ▶ Low importance

Packets of high importance may be protected with FEC level 1, packets of medium importance may be protected by FEC level 4 and packets of low importance may not be protected at all—applying an unequal error protection based on the level of importance of the video packets.



**Figure 6: Illustrates different types of FEC levels**

With each FEC level (n) you can lose 1 out of n-1 video packets without reducing the video quality.

## 4.4 Packet loss technology comparison

Packet loss technology	Strength	Weakness
Decoder concealment	Short-term receiver side concealment mechanism, can be quite effective for low-motion scenes, less so for high motion.	Ultimately, artifacts will emerge if concealment is not replaced by repair in a short time window
Down speeding	Standards-based, works in a mixed environment.	Does not up speed if packet loss was temporary, this can leave the call quality lower than the optimal. Will only have a substantial effect if the packet loss is being caused by a congested network.
Dynamic Bit Rate Adjustment	Similar to down speeding but with the added benefit of being able to up speed the call again if packet loss subsides. Can adapt to changing network conditions.	Will only have a substantial effect if the packet loss is being caused by a congested network.
LTRF	Reduces the amount of data sent by up to 90%. Helpful in all packet loss scenarios.	Requires support on both sender and receiver, and a backchannel for frame acknowledgement.
Video aware FEC	Very effective when the call rate is high in two ways: <ul style="list-style-type: none"> <li>- Low packet loss: removes quality reduction</li> <li>- High packet loss: protects repair P-frames for great overall performance.</li> </ul>	Not efficient in low call rate scenarios.

This table shows that a robust packet loss technology must use a variety of packet loss technologies, not just a single approach, and must do so in a dynamic fashion because packet loss is dynamic in nature.

## 4.5 Packet loss scenarios, and example technologies used

Mixed environment	Decoder concealment and down speeding
ClearPath capable products	Long Term Reference Frames should always be used. No negative side effects.
Bursty packet loss	Repair P-frames
Packet Loss in Low bandwidth calls	Repair P-frames
Packet Loss in High bandwidth calls	Repair P-frames and FEC
Constant packet loss	Dynamic bit rate adjustment if it helps; if not repair P-frames and FEC

## 5 Additional technologies

There are several other elements of resilience that are important; these techniques are not all fully defined in ClearPath today but are under consideration for future iterations.

### 5.1 Packet pacing

Video traffic is traditionally bursty in nature, and bunching packets is likely to exacerbate loss on a network link that is near capacity. Packet pacing ensures that packets are more evenly spaced in time based on the negotiated call bandwidth. Pacing does have an impact on latency so it is possible to adaptively pace to a higher bitrate than the actual call rate (in the absence of loss) reverting to a more cautious rate when loss is detected.

### 5.2 Disposable frames

In H.264 a frame can be marked as not used for reference so it does not enter the reference picture buffers. If frame  $n$  (dependent on frame  $n-1$ ) is disposable, then frame  $n+1$  must be predicted from frame  $n-1$  (or earlier).


Disposable frames can be added without a negative effect on video quality. These frames are automatically resilient, because their loss is transient and has no effect on subsequent frames; the sequence is self-repairing to loss within only those frames. Because they are not kept, they can also be simpler and in fact some steps need never be actually performed (quite literally, they need never exist in the decoder as a reconstructed video frame because they will not be used subsequently). Furthermore, networks can discard these frames without causing long-term problems; therefore they provide a means for networks to mitigate congestion (if they are signaled in a layer visible to the network).

### 5.3 Packetization

RTP provides packetization mode 1, which allows both aggregation and fragmentation. Aggregation places multiple complete H.264 Network Abstraction Layer Units (NALUs) in a single packet; fragmentation splits a large NALU across several packets. No packet can contain both aggregated and fragmented NALUs.

NALU fragmentation weakens resilience, because (unrecovered) loss of a packet renders further packets in the fragmented slice un-decoded. Slice headers contain important information for whether the current picture is an LTRF or not, and if the whole frame is coded in a single slice using fragmentation, then there is only one packet that contains the information that updates the state of the decoded picture buffer. Loss of the first packet of a frame is serious, with the decoder not knowing whether a frame needs to be NACK'ed or not (LTRFs only).

On the other hand, NALU aggregation aides resilience because it allows NALUs to be bound together such that slice data cannot be received without the important information that goes alongside it. This information can include things such as sequence/picture parameter sets, or the Supplemental Enhancement Information (SEI) message that reminds the decoder of decoded picture buffer state changes it may have been unaware of.



THE SPECIFICATIONS AND INFORMATION REGARDING THE PRODUCTS IN THIS MANUAL ARE SUBJECT TO CHANGE WITHOUT NOTICE. ALL STATEMENTS, INFORMATION, AND RECOMMENDATIONS IN THIS MANUAL ARE BELIEVED TO BE ACCURATE BUT ARE PRESENTED WITHOUT WARRANTY OF ANY KIND, EXPRESS OR IMPLIED. USERS MUST TAKE FULL RESPONSIBILITY FOR THEIR APPLICATION OF ANY PRODUCTS.

THE SOFTWARE LICENSE AND LIMITED WARRANTY FOR THE ACCOMPANYING PRODUCT ARE SET FORTH IN THE INFORMATION PACKET THAT SHIPPED WITH THE PRODUCT AND ARE INCORPORATED HEREIN BY THIS REFERENCE. IF YOU ARE UNABLE TO LOCATE THE SOFTWARE LICENSE OR LIMITED WARRANTY, CONTACT YOUR CISCO REPRESENTATIVE FOR A COPY.

NOTWITHSTANDING ANY OTHER WARRANTY HEREIN, ALL DOCUMENT FILES AND SOFTWARE OF THESE SUPPLIERS ARE PROVIDED "AS IS" WITH ALL FAULTS. CISCO AND THE ABOVE-NAMED SUPPLIERS DISCLAIM ALL WARRANTIES, EXPRESSED OR IMPLIED, INCLUDING, WITHOUT LIMITATION, THOSE OF MERCHANTABILITY, FITNESS FOR A PARTICULAR PURPOSE AND NONINFRINGEMENT OR ARISING FROM A COURSE OF DEALING, USAGE, OR TRADE PRACTICE.

IN NO EVENT SHALL CISCO OR ITS SUPPLIERS BE LIABLE FOR ANY INDIRECT, SPECIAL, CONSEQUENTIAL, OR INCIDENTAL DAMAGES, INCLUDING, WITHOUT LIMITATION, LOST PROFITS OR LOSS OR DAMAGE TO DATA ARISING OUT OF THE USE OR INABILITY TO USE THIS MANUAL, EVEN IF CISCO OR ITS SUPPLIERS HAVE BEEN ADVISED OF THE POSSIBILITY OF SUCH DAMAGES.

Cisco and the Cisco Logo are trademarks of Cisco Systems, Inc. and/or its affiliates in the U.S. and other countries. A listing of Cisco's trademarks can be found at [www.cisco.com/go/trademarks](http://www.cisco.com/go/trademarks). Third party trademarks mentioned are the property of their respective owners. The use of the word partner does not imply a partnership relationship between Cisco and any other company. (1005R)

Any Internet Protocol (IP) addresses and phone numbers used in this document are not intended to be actual addresses and phone numbers. Any examples, command display output, network topology diagrams, and other figures included in the document are shown for illustrative purposes only. Any use of actual IP addresses or phone numbers in illustrative content is unintentional and coincidental.

© 2012 Cisco Systems, Inc. All rights reserved.