Video Packets and The Need for a Specialized Video Content Delivery Network

Justin Trugman
justin@blinkcdn.com

Charles Bethin
charles@blinkcdn.com

Mukund Iyengar
mukund@blinkcdn.com

Sid Ahuja
sid@blinkcdn.com

blinkcdn

Introduction

In this white paper, the team of blinkcdn discusses types of frames and how it impacts the users' quality of experience (QoE). We believe that this knowledge about video is needed to guarantee the best user experience in current video applications and will be even more important in the future as the standard video resolution increases.

Frames & Group of Pictures

Video compression has come along way in recent years. Currently, there are three types of frames: I frames, P frames and B frames. Each serve their own purpose in a Group of Pictures (GOP)

Intra frames , more commonly referred to as I frames, carry the full video still image. They are essentially a still shot of that exact moment in the video. They are largest type of frame in terms of data and because of that, they have the largest impact on QoE if they are dropped. In addition to carrying the most data, they are also the reference frame for all other frames in the group of pictures. So if an I frame is lost or corrupted, not only do you lose or corrupt that frame, but all other frames in the GOP become corrupted as well.

Predictive-coded frames, or P frames, predict future frames based off of the I frame in it's group of pictures. Since they reference an I frame, it takes up less data than an I frame, typically being about twenty to seventy percent of the size of an I frame. This type of frame carries the second most amount of data in the group of pictures, and, since it is a reference frame for B frames, would also greatly impact the user's QoE if lost or corrupted.

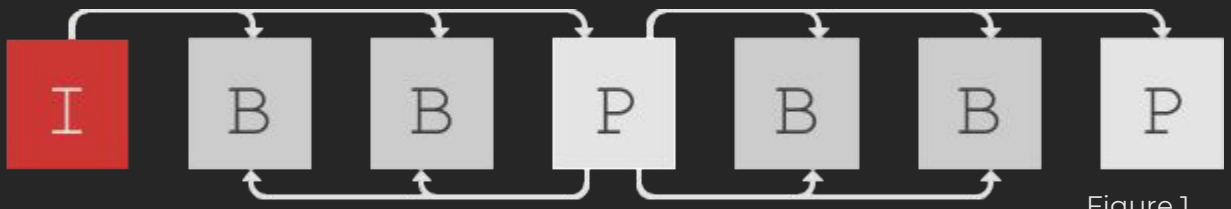

Figure 1

Bi-directionally predictive-coded frames, also known as B frames, carry the least amount of data in the GOP. They essentially carry the difference in motion between the last I or P frame. B frames are not referenced by any other frame in the GOP, meaning that if one is lost, only that specific frame would be effected and not any other frames. Therefore, the impact a B frame loss has on the user's quality of experience is significantly lower than an I frame or P frame loss.

All of these frame come together in what is called a Group of Pictures. A group of pictures contains an I frame and all the other frames before the next I frame. A group of pictures' size varies between twelve to fifteen frames contingent on numerous different factors. These factors, however, are perpendicular to the purpose of this paper.

As shown in Figure 1, B frames not only reference the preceding I frame or P frame, but also the future frame in the buffer. Figure 1 also demonstrates the gravity losing an I frame has. As shown in Figure 1 and previously explained, all the other frames in the group of pictures are affected by the loss of an I frame.

Frame Loss

So what happens when a frame is lost?

The short answer is a that a video artifact occurs. A video artifact is a visible distortion on a video frame in relation to the original video frame.

There are four types of video artifacts: a slice error, blocking or pixelization, ghosting and a freeze frame. A slice error occurs when at least one packet of any type of frame is lost. The result of this is that a horizontal “slice” of the frame the user sees is blurred. If it was an I frame or P frame packet than the rest of the group of pictures will be affected, if it was a B frame then the error will only affect that one frame.

Blocking or pixelation occurs when either a whole I or P frame is lost. This causes a pixelation effect in a large block of the screen and will affect not only the I or P frame, but the whole group of pictures. The only way for this type of error to be fixed is to either receive a new I frame or P frame depending on which frame is lost.

Ghosting is one of the worst types of errors, it occurs when a whole I frame is lost or corrupted. Since the I frame in the group of pictures is lost, the P and B frames reference the I frame from the previous group of pictures. This can cause a drastic effect on the video because it is adding new information to an old scene and can cause a massive amount of blur.

When a multiple frames are lost in a continuous stream a freeze frame artifact occurs. What happens is that the previously received frame is “frozen” on the screen until a frame is received.

Blinkcdn is using this knowledge to help develop our network and give the user the best QoE possible.

References

Greengrass, Jason, John Evans, and Ali C. Begen. "Not all packets are equal, part i: Streaming video coding and sla requirements." IEEE Internet Computing 13, no. 1 (2009): 70-75.

Greengrass, Jason, John Evans, and Ali C. Begen. "Not all packets are equal, part 2: The impact of network packet loss on video quality." IEEE Internet Computing 13, no. 2 (2009): 74-82.

Mukerjee, Matthew K., David Naylor, Junchen Jiang, Dongsu Han, Srinivasan Seshan, and Hui Zhang. "Practical, real-time centralized control for cdn-based live video delivery." ACM SIGCOMM Computer Communication Review 45, no. 4 (2015): 311-324.

Liu, Xi, Florin Dobrian, Henry Milner, Junchen Jiang, Vyas Sekar, Ion Stoica, and Hui Zhang. "A case for a coordinated internet video control plane." In Proceedings of the ACM SIGCOMM 2012 conference on Applications, technologies, architectures, and protocols for computer communication, pp. 359-370. ACM, 2012.