

Sharper Perspective Driving Superior Performance



Utilizing Unicast and Multicast Video Streams for Physical Security Systems over Wired and Wireless Networks

Effectively integrating real-time video streams over wired and wireless networks in physical security systems is a technological challenge. This report examines the issues systems integrators face and the approach Applus Technologies, Inc. takes to successfully deliver best-in-industry physical security systems. Two different methods are utilized in delivering streaming video over Ethernet communications: Unicast and Multicast. Each method has distinct advantages and disadvantages that must be weighed when customizing an enterprise security system to effectively meet the business objectives specific to each client.

1. Introduction

Advances in digital video encoding are enabling a trend away from analog surveillance video systems to IP-based digital surveillance video systems. IP based networks are more cost effective than analog and are also more flexible and enable more rapid deployment of security systems.

There are two methods for transmitting video across an IP network: Unicast and Multicast. Unicast is the more established method of delivering video over Ethernet networks and is a one-to one connection between the client and the server. Unicast uses IP delivery methods such as Transmission Control Protocol (TCP) and User Datagram Protocol (UDP), which are session-based protocols. When a client connects using unicast to a surveillance camera or encoder, that client has a direct relationship to the server. Each unicast client that connects to the server takes up additional bandwidth. For example, if you have 10 clients all playing 1024-kilobits per second (Kbps) streams, those clients as a group are taking up 10,240 Kbps. If you have only one client playing the 1024 Kbps stream, only 1024 Kbps is being used.

Multicast is the newer and more network efficient method of delivering video. Multicast is a true broadcast. The multicast source relies on multicast-enabled network devices utilizing Internet Group Management Protocol (IGMP) to forward the packets to all client subnets that have clients listening. There is no direct relationship between the clients and the surveillance video feeds. Each client that listens to the multicast adds no additional overhead on the network. In fact, the surveillance camera sends out only one stream per multicast station. The same load is experienced on the network whether one or 1,000 clients are listening.

Multicasting requires packet replication at various network locations; therefore, network devices must be able to perform this function without impairing the line-rate throughput of the device. This consideration becomes increasingly important as the network scales, since larger multicast trees require more replication.

2. Complexity of Video

Analog video or RAW video can use up to 165 Mbps or 165 million bits per second, which creates excellent video quality but that size video stream would be impractical and complicated to deploy on today's networks. To solve this problem, surveillance cameras today typically feature one or all of the following video compression formats: MJPEG, MPEG-4 and H.264.

Motion JPEG is a video compression technique known as still image compression which can be represented in a sequence of JPEG images. This method is an intraframe-only compression technique that typically utilizes a compression ratio of 1:20 or lower. It does, however, typically provide a higher quality image than MPEG-4 or H.264 but will come at a much higher bandwidth usage of close to 5.0 to 10.0 Mbps, depending on your settings for compression ratio and frames per second.

JPEG is inefficient, using more bits to deliver equal quality, compared to other, more recently developed formats. Since the development of the original JPEG standard in the early 1990s, technology improvements have made intraframe compression possible. Technology improvements can be found in the designs of MPEG-4 Part 2, which uses frequency-domain prediction of transform coefficient values, and in H.264/MPEG-4 AVC, which uses spatial prediction and adaptive transform block size techniques and more sophisticated coding than what was practical when the first JPEG design was developed. These new developments make MJPEG appear outdated and inefficient to deploy large scale video networks. In addition, MJPEG video can only be streamed utilizing unicast, which eliminates the flexibility of using multicast on your network to reduce bandwidth congestion as network conditions can change over time.

In MPEG-4 and H.264 (also referred to as MPEG-4 Part 10 or AVC) video compression is based upon the same technique that is used in JPEG. In addition, it also includes techniques for efficient coding of a video sequence.



Figure 1—a three picture sequence of MJPEG

Consider the video sequence displayed in Figure 1. The picture to the left is the first picture in the sequence followed by the picture in the middle and then the picture to the right. When displayed, the video sequence shows a man running from right to left with a house that stands still.

In Motion JPEG each picture in the sequence is coded as a separate unique picture resulting in the same sequence as the original one. In MPEG-4 and H.264 video only the new parts of the video sequence is included together with information of the moving parts. The video sequence of Figure 1 will then appear as in Figure 2. But this is only true during the transmission of the video sequence to limit the bandwidth consumption. When displayed it appears as the original video sequence again.



Figure 2—a three picture sequence of MPEG-4/H.264

This is further explained by the use of I, B and P frames. Figure 3 shows all three frame types in a group of pictures (GOP), or a sequence of frames that starts with a key frame and includes all frames until the next key frame. Briefly, an I frame is entirely self contained, and is compressed solely with intra-frame encoding techniques, typically a technology like JPEG, which is used for still images on the web and in many digital cameras.

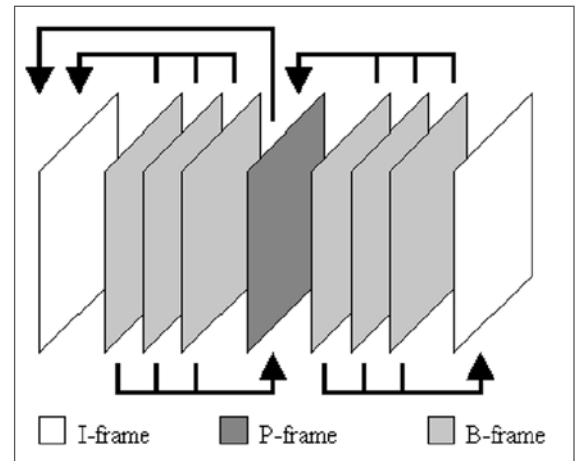


Figure 3—I, B and P-Frames

P and B frames are “delta” frames that “refer” to other frames for as much of their content as possible. Imagine a talking head video. A P frame will look back to a previous I or P frame for regions in the frame that haven’t changed, and only encodes what has changed between the frames. In a surveillance scenario of an interior doorway, very little changes happen so the P-frame tells the player, in effect, “just copy the walls and doors from that I-frame and then use these new pixels around the person walking into the doorway.”

This is why this type of scenario video compresses so efficiently; there’s so much inter-frame redundancy that the delta frames contain very little new information. In a busy outdoor city street intersection surveillance scenario, delta frames contain much more original content, which makes compressing down to the target data rate much tougher because you have cars moving, people walking on streets, and trees moving in the wind.

Back to our frame types. By definition, a P-frame looks backwards to a previous P or I frame for redundancies, while a B-frame can look backwards and forwards to previous or subsequent P or I frames. This doubles the chance the B-frame will find redundancies, making it the most efficient frame in the GOP.

As illustrated on the previous page, it is important to do not lose any I frames when utilizing compressed video standards. Even with the slightest packet loss in any I frame packets can diminish subsequent P frames and result in complete video loss. This is why it is important in certain network situations, depending on network reliability and media, to carefully weigh the advantages and disadvantages in utilizing compressed video formats as compared to MJPEG.

3. Complications of Video over Wireless

Extending video over wireless networks presents certain challenges compared to wired networks. These challenges arise because wireless introduces a new set of network performance characteristics including variable data rates, packet loss and multipath, causing over all latency and unreliability.

To understand the fundamental challenges of video over wireless, it is first important to understand the characteristics of each of these factors. A substantial difference between wireless and a wired is the relative unreliability of the underlying Layer 2 transport. To put it simply, wireless loses a lot more packets than wired. The first reason for packet loss is collisions—that is, two wireless devices attempting to transmit at the same time. Wireless uses a shared half-duplex medium, and while the “listen-before-talk” medium access method tries to avoid collisions, they cannot be totally prevented. This situation is worsened by non-802.11 devices, which may operate in the same band as 802.11 Devices. Most of these devices do not follow the “listen-before-talk” algorithm and therefore, collisions are common.

A second reason for packet loss is that wireless 802.11 transmissions are subject to short-term signal loss (referred to as fades). These fades can be caused by absorption from intervening objects in the environment (e.g. people) or reflections of waves in the environment, which accidentally cause signal cancellation. A third and smaller factor in packet loss is that wireless systems hunt for the best transmission data rate by trying different rates and therefore, some packets are lost during the search process.

Given the combination of collisions, fades and data rate selection, it is not at all uncommon for wireless to operate with an underlying packet error rate (PER) that can approach 5 percent. To compensate, wireless uses a retransmission mechanism whereby packets that are not successfully received and acknowledged are resent. This mechanism generally serves to reduce

the final packet loss rate (PLR) to less than 0.1 percent. However, these retransmissions result in jitter and eat into overall network throughput, both of which can impact QoS. And even after retransmissions, the final PLR is still much higher than is typically observed on wired connections.

A substantial difference between wireless and a wired is that the data rate of transmission over wireless varies over time, and depends on the distance of the client from the access point. This stands in contrast to traditional wired system, in which if a wired connection is operating at 100 Mbps today, it will operate at 100 Mbps tomorrow. As a result of the variations in data rates, the throughput of individual video flows and the capacity of the overall network changes with time.

As can be imagined, the variable throughput and capacity present a challenge to the traditional QoS approach of bandwidth reservation and admission control. For example, consider a client that is operating at 54 Mbps, and requesting a video stream of 10 Mbps. The system determines the necessary airtime for the new stream can be accommodated and admits the stream. But now the client moves away from the access point, and the data rate of the client drops to 6 Mbps. Now the video stream cannot be supported. In this sense, sending video over a wireless network has some similarity to sending video over the public Internet, where throughput and the user’s experience can vary widely over time.

Today, there is no efficient MAC Layer protocol in the IEEE 802.11 standard. The IEEE 802.11 protocol implements positive acknowledgment to provide reliable transmission of unicast packets over 802.11 but not for multicast packets.

The packet error rate plays a prominent role for wireless multicast traffic. For multicast transmissions (with multiple receivers), wireless does not provide a retransmission mechanism. As a result, the PLR for multicast traffic is equal to the PER. In other words, it would not be uncommon for wireless multicast traffic to experience a packet loss rate of at least 5 percent. This is a significant problem for video, where loss of even a single packet can result in an error that propagates for many video frames. In fact, this problem can be intensified by the more cameras that you add onto a wireless network which could potentially bring video traffic to a standstill. For this reason, it is quite normal for multicast video applications that work on a wired network to fail completely when they operate on a wireless network.

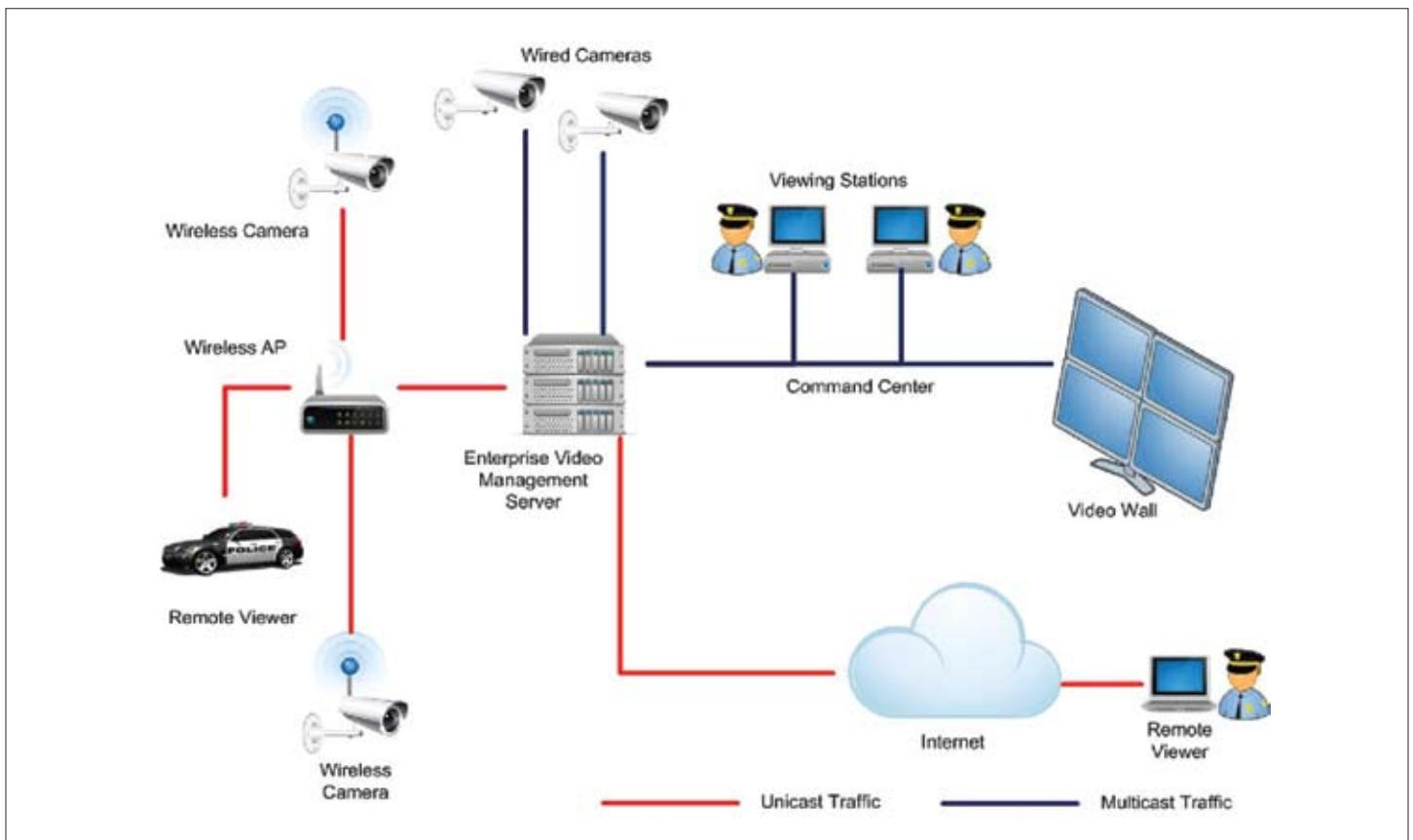


Figure 4—example of a Hybrid Unicast and Multicast Video Surveillance System

4. Choosing the Right Software Application

A critical step in planning a video surveillance system is in selecting the right video management system. Many video surveillance video management systems make claims of being enterprise ready and versatile in ways it can receive video from surveillance cameras and deliver video to its clients. It is best to test out several video surveillance software manufactures and run network performance tests with several cameras and network devices monitoring and then logging what network traffic is being broadcast on the networks. You will find that every software manufacturer generates its own unique traffic on the network and may not work in certain environments such as wireless. For example, some manufacturers even though their setting is for unicast video may generate the unicast via a multicast video steam, which can cause unplanned multicast traf- fic on the network.

It is preferable to select an application that is open architecture and one that is able to distribute video in a variety of ways such as multicast and unicast from the cameras to the server-side of the software application. The most important thing is to test before deploying any network video recording software before deploying it to a customer's site.

5. Using a Hybrid Network Approach to Build Your Enterprise Security System

The most efficient way of delivering video services is by utilizing a hybrid network design that will allow for multicast transmissions where bandwidth can be saved on the network and utilizing unicast transmissions where multicast transmissions cannot be delivered reliably such as the Internet and wireless networks. This factor makes it especially important to choose a video management system with the capability to receive mul- ticast and unicast video streams from the camera and then rebroadcast the streaming video using the proper method to the intended live viewer.

References

Axis Communications. White Paper—An Explanation of Video Compression Techniques.

Axis Communications Online: http://www.axis.com/files/whitepaper/wp_videocompression_33085_en_0809_lo.pdf

Cisco Systems. White Paper — Overview of IP Multicast.

Cisco Online: http://www.cisco.com/en/US/tech/tk828/technologies_white_paper09186a0080092942.shtml

International Journal of Video & Image Processing and Network Security IJVIPNS Vol: 9 No: 10. White Paper— Video Multicasting in Campus Networks

Online: <http://www.ijens.org/97910-2727%20IJVIPNS-IJENS.pdf>

IEEE Transactions on Circuits & Systems For Video Technology, Vol. 12, No. 6, June 2002 White Paper— Multicast and Unicast Real-Time Video Streaming Over Wireless LANs

Online : http://www.kozintsev.net/papers/journal_01.pdf

Author

David Engel

Senior Solutions Engineer

Applus Technologies, Inc.

dengel@aplustech.com

Applus 

Technologies

**Sharper Perspective
Driving Superior Performance**

444 North Michigan Avenue

Suite 1110

Chicago, IL 60611

phone: +1-847-616-6122

email: solutions@aplustech.com