TECH NOTE

IP Surveillance Camera Bandwidth

About The Tech Note

Introduction
This document includes an introduction to the fundamentals for Video Surveillance.

Conventions
This document contains notices, figures, screen captures, and certain text conventions.

Figures and Screen Captures
This document provides figures and screen captures as an example. These examples contain sample data. This data may vary from the actual data on an installed system.

Units of Measurement
Units of measurement in this publication conform to SI standards and practices.

Jan 01, 2022

Introduction

Bandwidth is the most crucial element of ethernet networking for video surveillance systems. Without careful planning ahead of time, the surveillance system might end up with a bandwidth bottleneck. This not only causes video packet loss, delay, or jitter but also degrades video quality, or even worse, inhibits recording of critical incidents. Bandwidth also determines the storage capacity requirements for a given retention period.

Understanding video bandwidth takes an in-depth knowledge of several fields. This technote is interned to provide fundamental knowledge of what affects video surveillance performance. We break down each of the following:

- **Measuring Bandwidth**
- Bits and Bytes
- Kilobits, Megabits and Gigabits
- Bit Rates
- **Video Compression and Bandwidth**
- Motion JPEG
- H.264
- H.265
- Constant and Variable Bit Rates (CBR and VBR)
- Constant Bitrate
- Variable Bitrate
- Camera Bandwidth Consumption
- Video Resolution
- Frame Rate
- Scene Complexity
- **Bandwidth and Recorder**
- LAN and WAN
- Dedicated Switch
- Switch Recommendation
- Sizing Networks for Video Surveillance

What is Bandwidth?

IP video is transmitted as a stream of data that contains the image, audio, and control data of the camera. The amount of data that has to be sent per second is called bandwidth. It is commonly measured in Mbit/s, which makes it easy to compare to the bitrate capacity of an Ethernet link. For example, 10 Mbit/s is called Ethernet, 100 Mbit/s is Fast Ethernet, and 1,000 Mbit/s is Gigabit Ethernet. Another measuring unit is MByte/s, equal to 1/8th of the bitrate because there are 8 bits in a byte.

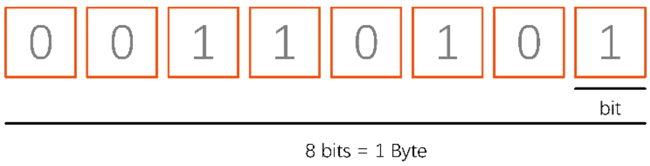
1 Mbit/s = 1,000 Kbit/s = 125 Kbyte/s

1 Gbit/s = 1,000 Mbit/s = 125 Mbyte/s

A 1920 x 1080 HD resolution camera roughly generates raw video data at 1.49 Gbit/s (30 x 1920 x 1080 x 24) for 30 FPS video. That is 178 MByte/s of data and the reason video compression is required.

Bits and Bytes

In video surveillance, bandwidth is typically measured in bits but sometimes measured in bytes, causing confusion. 8 bits equals 1 byte, so someone saying 40 Megabits per second and another person saying 5 Megabytes per second means the same thing but is easy to misunderstand or mishear.



Bits and bytes both use the same letter for shorthand reference. The only difference is that bits use a lower case ‘b’ and bytes use an upper case ‘B’. You can remember this by recalling that bytes are ‘bigger’ than bits. We see people confuse this often because at a glance they look similar. For example, 100Kb/s and 100KB/s, the latter is 8x greater than the former.

We recommend you use bits when describing video surveillance bandwidth but beware that some people, often from the server/storage side, will use bytes. Because of this, be alert and ask for confirmation if there is any unclarity (i.e., “Sorry did you say X bits or bytes”).

Kilobits, Megabits and Gigabits

It takes a lot of bits (or bytes) to send a video. In practice, you will never have a video stream of 500b/s or even 500B/s. Video generally needs at least thousands or millions of bits. Aggregated video streams often need billions of bits.

The common expression/prefixes for expressing a large amount of bandwidth are:

- Kilobits is thousands, e.g., 500Kb/s is equal to 500,000b/s. An individual video stream in the kilobits tends to be either low resolution or low frame or high compression (or all of the above).
- Megabits is millions, e.g., 5Mb/s is equal to 5,000,000b/s. An individual IP camera video stream tends to be in the single-digit megabits (e.g., 1Mb/s

or 2Mb/s or 4Mb/s are fairly common ranges). More than 10Mb/s for an individual video stream is less common, though not impossible in super-high-resolution models (4K, 20MP, 30MP, etc.). However, 100 cameras being streamed at the same time can routinely require 200Mb/s or 300Mb/s, etc.

- Gigabits is billions, e.g., 5Gb/s is equal to 5,000,000,000b/s. One rarely needs more than a gigabit of bandwidth for video surveillance unless one has a very large-scale surveillance system backhauling all video to a central site.

Bit Rates

Bandwidth is like vehicle speed. It is a rate over time. So just like you might say you were driving 60mph (or 96kph), you could say the bandwidth of a camera is 600Kb/s, i.e., that 600 kilobits were transmitted in a second.

Bit rates are always expressed as data (bits or bytes) over a second. Per-minute or hour are not applicable, primarily because networking equipment is rated as what the device can handle per second.

Video Compression and Bandwidth

Video compression is the process of encoding a video file in such a way that it consumes less space than the original file and is easier to transmit over the network/Internet. It is a type of compression technique that reduces the size of video file formats by eliminating redundant and non-functional data from the original video file.

Once a video is compressed, its original format is changed into a different format (depending on the codec used). The video player must support that video format or be integrated with the compressing codec to play the video file.

Motion JPEG

Motion JPEG (M-JPEG or MJPEG) is a video compression format in which each video frame or interlaced field of a digital video sequence is compressed separately as a JPEG image.

Originally developed for multimedia PC applications, Motion JPEG enjoys broad client support: most major web browsers and players provide native support, and plug-ins are available for the rest. Software and devices using the M-JPEG standard include web browsers, media players, game consoles, digital cameras, IP cameras, webcams, streaming servers, video cameras, and non-linear video editors

H.264

H.264, which is also called MPEG-4 AVC, is a compression standard that was introduced in 2003 and is the prevalent standard used in video surveillance cameras and many commercial media applications. In contrast to the frame-by-frame approach of MJPEG, H.264 stores the full-frame only at intervals of, for example, once a second and encodes the rest of the frames only with the differences caused by motion in the video. Full frames are called I-frame (also index frame or intra-frame) and the partial ones containing only the difference to the previous frame are called P-frame (also predicted frame or inter-frame). P-frames are smaller and more numerous than I-frames. There is also a B-frame (bidirectional frame), which refers both ways to previous and subsequent frames for changes. The recurring pattern of I-P-B frames is called a group of pictures (GOP). The time interval for I-frames varies and can range from multiple times a second to nearly a minute. The more I-frames are transmitted, the larger the video stream will be, but it makes restarting decoding of a stream easier since this can only happen at an I-frame.

H.265

High-Efficiency Video Coding (HEVC), also known as H.265 and MPEG-H Part 2, is a video compression standard designed as part of the MPEG-H project as a successor to the widely used Advanced Video Coding (AVC, H.264, or MPEG-4 Part 10). In comparison to AVC, HEVC offers from 25% to 50% better data compression at the same level of video quality or substantially improved video quality at the same bit rate. It supports resolutions up to 8192×4320, including 8K UHD, and unlike the primarily 8-bit AVC, HEVC's higher fidelity Main 10 profile has been incorporated into nearly all supporting hardware.

H.264 vs. H.265

H.265 is more advanced than H.264 because of various reasons. The biggest difference here is that H.265/HEVC allows for even lower file sizes of your live video streams. This significantly lowers the required bandwidth.

Then, another perk of H.265 is the fact that it processes data in coding tree units. Although macroblocks can go anywhere from 4×4 to 16×16 block sizes, CTUs are able to process up to 64×64 blocks. This enables H.265 to compress information more efficiently.

Additionally, H.265 also has an improved motion compensation and spatial prediction than H.264 does. That is quite helpful for your viewers in that their devices will require less bandwidth and processing power to decompress all information and watch a stream.

Constant and Variable Bit Rates (CBR and VBR)

Bitrate measures the amount of data that is transferred over a period of time. In online video streaming, video bitrate is measured in kilobits per second, or kbps. Bitrate affects the quality of a video. Streaming with higher bitrate helps you produce higher-quality streams.

Bitrate is also something that is important in the encoding or transcoding stage of the streaming process since this too deals with the transfer of data.

Constant Bitrate

When configuring a camera for CBR, the camera is set to have constant bandwidth consumption. The amount of compression applied increases as more changes are occurring. This can add compression artefacts to the image and degrade image quality. With CBR, the image quality will be sacrificed to meet the bandwidth target. If the target is reasonably set, this degradation may be hardly noticeable and it gives a stable basis for calculating storage and planning the network. For IP surveillance cameras installed in a local-area network (LAN) with low network utilization or when storage space is abundant, VBR is recommended to maintain the best image quality, whereas CBR can help control bandwidth-restricted environments.

Variable Bitrate

The strength of each compression method can be adjusted. In general, higher compression causes more artefacts, so there are different strategies to achieve the desired behaviour. When VBR compression is used, the size of the compressed stream is allowed to vary to maintain consistent image quality. Thus, VBR can be more suitable when there is motion in the scene and it tends not to be constant. The disadvantage is that the bandwidth can, to a certain extent, vary depending on the situation. So storage may be used up earlier than planned or transmission bottlenecks can appear when cameras suddenly require more bandwidth. In VBR, there is no firm cap being placed on the bitrate. The user sets a certain target bitrate or image quality level.

The VBR compression level can be set to Extra High, High, Normal, Low and Extra Low in some recorder systems.

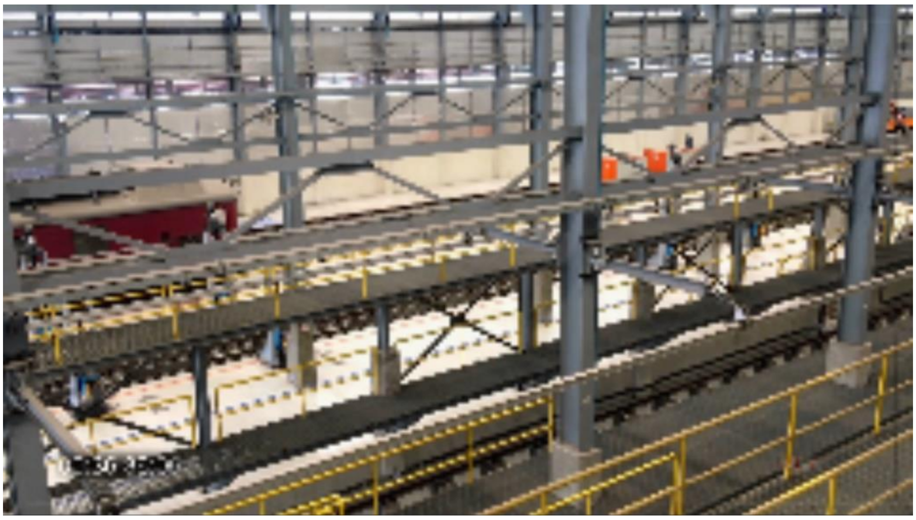


Figure 3 Video quality Extra Low, average bandwidth is 0.5 Mbit/s

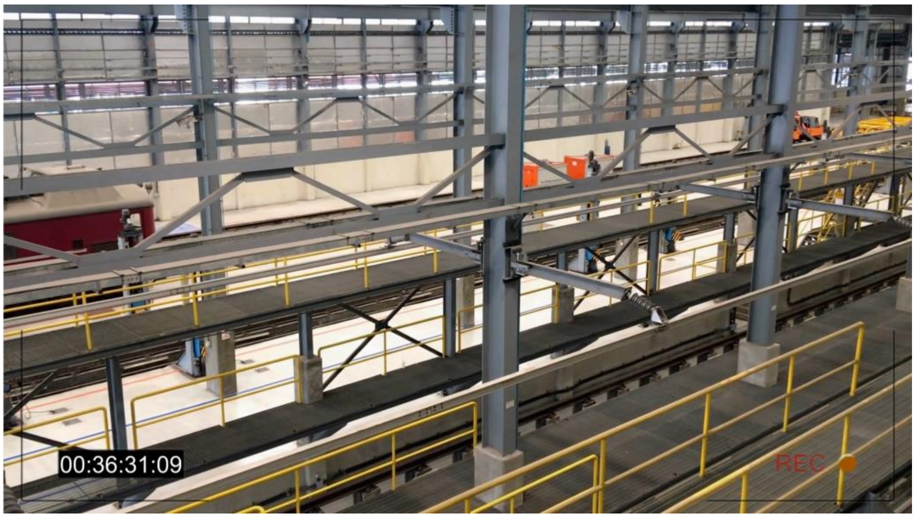


Figure 4 Video quality Extra High, average bandwidth is 1.5 Mbit/s

Camera Bandwidth Consumption

Here are a few common drivers of camera bandwidth consumption:

Resolution: The greater the resolution, the greater the bandwidth.

Frame Rate: The greater the frame rate, the greater the bandwidth

Scene Complexity: The more activity in the scene(lots of cars and people moving vs no one in the scene), the greater the bandwidth needed.

Low light: Nighttime often, but not always, requires more bandwidth due to noise from cameras.

Video Resolution

Every camera has an image sensor. The available pixels from left to right provide the horizontal resolution, while the pixels from top to bottom provide the vertical resolution. Multiply the two numbers for the total resolution of this image sensor. Assuming 24 bits for the RGB color values of a pixel:

1920(H) x 1080(V) = 2,073,600 pixels =2.0 MP x 24 bits = 48 Mbit/s

4096(H) x 2160(V) = 8,847,360 pixels =8.0 MP x 24 bits = 192 Mbit/s

Therefore, 4096 x 2160 takes more bandwidth since it contains more pixels, or simply saying, more data. But it gives clearer, sharper pictures when needed to identify a subject, a face, or a car model and its colour or license plate. Vice versa, lower resolution generates less bandwidth, but the trade-off is a less clear, blurrier image. Lower resolution usually gives surveillance operators situational awareness—seeing what is going on rather than detail.

Resolution is not the only thing that determines the clarity of an image. The optical performance of the lens, focal length (optical zoom), distance to the object, lighting conditions, dirt, and weather are also critical factors.

Frame Rate

Frame rate is measured in frames per second (FPS), which means the number of pictures being produced in a second. The higher the frame rate, the smoother the subject moves in the video. The lower the frame rate, the more jerky movements become up to the point where subjects jump from position to position with a loss of anything in between. Bandwidth increases with frame rate. Half the frame rate usually does not quite reduce bandwidth by half, though, because the encoding efficiency suffers. Modern surveillance cameras can generate up to 60 FPS. However, CPU limitations will sometimes restrict the FPS to a lower value when resolutions are set too high. Finding the optimal FPS setting for a scene is a compromise between objectives: capture all the relevant information without essential details being lost between frames versus bandwidth considerations. If a camera is monitoring a quiet overview, there is no need to go up to 30 FPS. A setting of 5 to 15 FPS is sufficient. As a rule of thumb, the more rapid change occurs or the faster subject movement is anticipated, the higher the FPS should be set. Adjust the FPS after cameras are installed and monitor whether the smoothness of the video is acceptable or not.

Scene Complexity

The complexity of a scene also affects the bandwidth a video camera generates. Generally, the more complex the scene is, the more bandwidth will be required to achieve a certain image quality. For example, scenes that have tree leaves, wire fencing, or random textures like popcorn ceilings increase the complexity of the scene. Others, like a normal, plain colour painted wall or little detail, are considered a simple scene. Similarly, motion or movement increases complexity. People walking by, cars driving across, or tree leaves in a breeze are examples.

Figure 5 Complex Scene, video bitrate is 5 Mbit/s

Number of Cameras and Clients

The number of cameras influences the bandwidth requirements for a system. If all cameras are the same, then twice the camera numbers will double the data generated. To maintain scalability of a system, it must be able to break large topologies into manageable smaller partitions. By structuring the system in a layered and distributed architecture, it is possible to maintain scalability over a large range of quantities. The key is to distribute bandwidth so bottlenecks are avoided. More will be discussed in the Bandwidth Bottlenecks section.

Number of viewing clients

The discussion above relates to camera bandwidth feeding into the recorder. This is only one side of the picture, where the other side is connecting the recorders to the clients watching live or playback video. For example, there might be a security team that constantly monitors the cameras 24 hours a day, seven days a week. This bandwidth would be equal to all the data coming from the cameras. In the case of playback, even more bandwidth is required if used in addition to live streaming. Considering there can be many clients connecting to a system at the same time, the client-side traffic can be the dominant concern.

Bandwidth and Recorder

Video surveillance consumes network bandwidth in one of the following two typical scenarios:

- Camera/encoder to recorder: Video is generally generated in different devices than recorded (e.g., a camera causes the video, a DVR / NVR / VMS server records it). In between, the video needs to be transmitted. If it goes over an IP network (e.g., IP cameras to NVR / VMS), bandwidth is required.
- Recorder to the client: Statistically, a meager percentage of video is watched by humans. Often, the person watching is on a different device on an IP network than the recorder. For example, the writer might be in a rack in an IT closet, but the viewer (i.e., client) is on a laptop, mobile phone or a monitoring station.

Because of this design, the overwhelming majority of bandwidth needed in surveillance systems is dictated by (1) camera type and (2) the relative placement of cameras and recorders.

In terms of the camera type, non IP cameras (NTSC / PAL analogue, Analog HD, HD SDI) typically do not consume network bandwidth unless the video is sent to clients. Each camera has a cable directly connected to a recorder.

For all camera types, the relative physical placement near the camera significantly impacts bandwidth needs. For example, imagine 1,000 cameras, with 100 cameras each at ten buildings on campus. If each building has a recorder, the peak bandwidth requirements will be ~90% lower than if there is only a single site for recording (i.e., each building recording its own might only need ~200Mb/s network connection compared to ~2Gb/s if they are all being sent back to one building). Each approach has pros and cons, but knowing where you will place recorders has a significant impact.

LAN and WAN

Figure 6 represents a diagram that bandwidth bottlenecks can exist in almost every connection among devices. Individual camera connections to the edge switch with a few kbps to 10 Mbit/s are unproblematic. However, the edge switch aggregates camera traffic onto its uplink port to the core switch, and a bandwidth bottleneck could occur at this point. Using the scenario shown in the illustration below, the 100 Mbit/s uplink port of the Fast Ethernet edge switch would be overloaded by the 16 x 5 Mbit/s = 80 Mbit/s traffic from the cameras. The actual throughput of Ethernet links should be limited to 50% to 70% of their nominal capacity. The solution is to upgrade the Fast Ethernet switch to a Gigabit uplink port to ease the bottleneck.

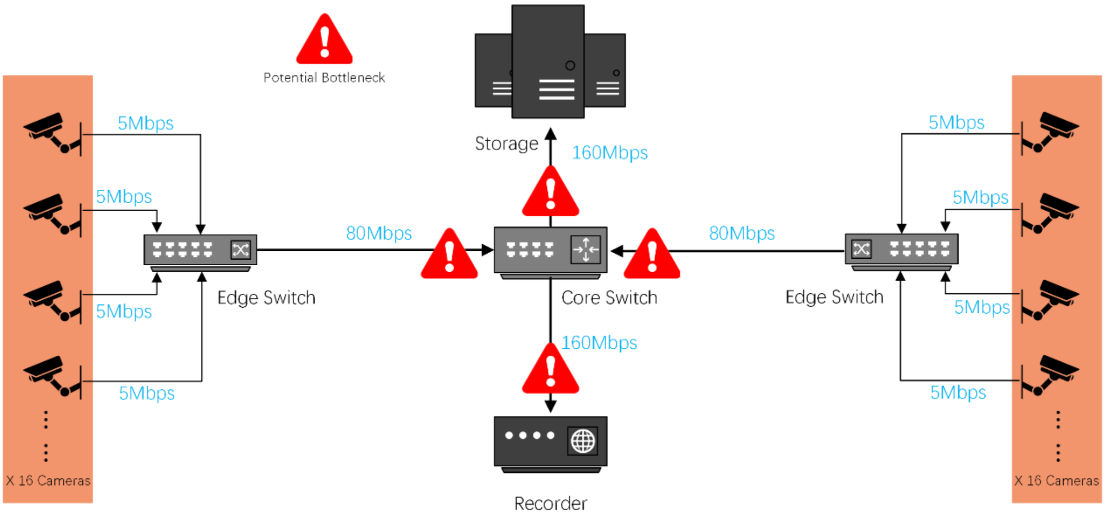


Figure 6: Bandwidth bottlenecks

WAN

WANs can deliver the same or more bandwidth as the LAN, but the costs tend to be significantly higher (in the order of 10 or 100x more expensive per bit) because these networks need to run great distances and across many obstacles. While one certainly could secure a 1 Gig WAN connection, the likelihood of doing this for surveillance is very low, given the cost, this would typically incur.

Many WAN networks/connections have asymmetric bandwidth, a problem for remote monitoring or recording of video.

- Symmetric bandwidth means the bandwidth is the same 'up' and 'down', i.e., a link can send the same amount of bandwidth as it can receive (100Mb/s up and 100Mb/s down is a classic example).
- Asymmetric bandwidth means the bandwidth up and down are not the same. Specifically, the bandwidth 'up' is frequently lower than the bandwidth 'down'. This is common in homes and small offices.

Asymmetric connections provide sufficient downstream speeds for video and general use but may only provide a fraction of the speeds needed for upload. Downstream bandwidth on common cable modem connections maybe 50 Mb/s, 100 Mb/s, 300 Mb/s or more. However, these same connections are most often rated for only 5 Mb/s or 15 Mb/s up, which may be an issue for those trying to stream video from these connections.

Asymmetric Connection Types

The most common asymmetric bandwidth connections are cable modems, by far, as DSL has fallen out of favour as cable speeds improved and residential fibre networks have increased in size. Satellite connections are typically only used in remote sites where no other options exist.

Symmetric Connection Types

The main exceptions, those that offer symmetrical bandwidth commonplace, are:

- Fiber-optic networks: In the past ten years, fibre optic internet has become common in much of the United States, offering symmetric connections (e.g., 50 Mbps Down/50 Mbps Up) at prices similar to cable modems and much lower than leased lines and commercial fibre connections. The main limitation is access to such networks. While increasing over the past decade, they tend to be limited to densely populated urban areas.

- Telecommunication/telephony networks (e.g., T1s, T3s, OC3s) but these are expensive, typically $500/month or more, and often low bit rate (e.g., respectively 1.5Mb/s and 45Mb/s for T1s/T3s).

Dedicated Switch

A dedicated switch for surveillance cameras is highly recommended. Surveillance cameras constantly generate a lot of traffic. Well-designed network infrastructure should separate surveillance traffic from existing data network traffic. This practice will make it possible to calculate the uplink bandwidth requirement and the bandwidth aggregated in your core switch.

In Figure 7, the surveillance network on the left is separated from the corporate network on the right. All models have two or more Gigabit network interfaces, except the VM version, which requires the administrator to install two Gigabit network interface cards on the server. One interface is connected to the surveillance network, and the second is connected to the corporate network.

If a dedicated switch is not practical in the existing network infrastructure, consider setting up virtual local-area networks (VLANs) in the control. A VLAN can separate the surveillance camera traffic from the data traffic. It is like splitting one physical switch into two virtual switches. From a network security standpoint, the separation of the surveillance network does also segment the area where third-party cameras are active.

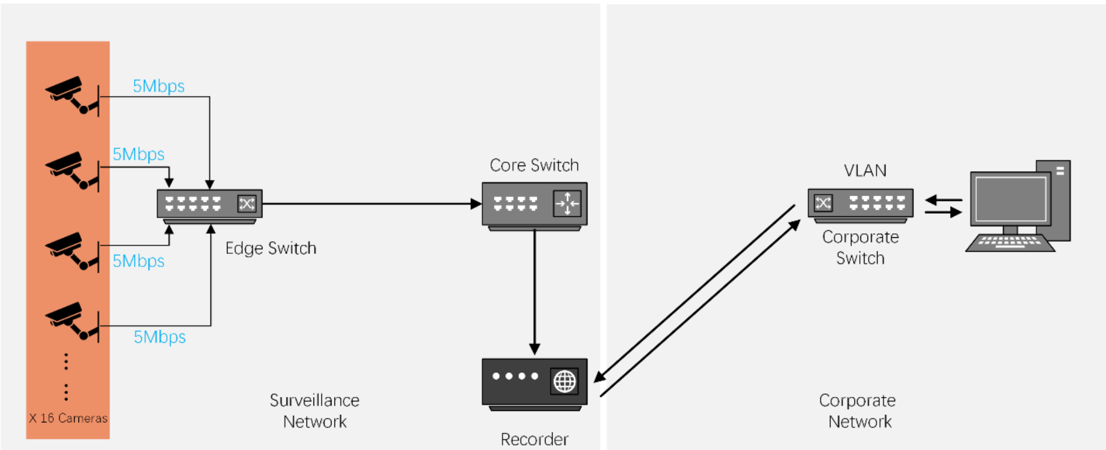


Figure 7: Dedicated Switch for Surveillance Network

Switch Recommendation

- Fast Ethernet 100 Mbit/s ports are sufficient for camera use. Their 50-70 Mbit/s effective capacity can carry the $<$ 10Mbi/s video streaming bandwidth.
- Ensure the switch fabric capacity is sufficient. Most switches today have full backplane capacity.

- The uplink port must support the sum of incoming traffic from the cameras. Gigabit Ethernet with 1,000 Mbit/s is generally sufficient for switches up to 48 ports.
- By the way, it is necessary to pay attention to the coexistence of 2K/4K video surveillance and WiFi wireless access, and it is required to upgrade to a switch with a 10G uplink port.

Sizing Networks for Video Surveillance

Putting this information together, to size a network for video surveillance, you will need to know:

- How much bandwidth each camera consumes, recognising that wide variations can exist
- How close (or far) the recorder is going to be placed to the cameras connected to it, presuming they need an IP network
- What the bandwidth of those network connections are and what pre-existing load those networks must also support.

Fiberroad Technology Co., Ltd
www.fiberroad.com
Sales Support: sales@fiberroad.com
Technical Support: tech@fiberroad.com
Service Support: service@fiberroad.com