CHAPTER 4

# Planning and Design

This chapter introduces the concepts of why an enterprise network should consider migrating from a standalone analog network to a converged IP network supporting voice, video, and data. The following topics are addressed:

- Introduction to the concepts of video resolutions, codecs, aspect ratios, frame rates and requirements for camera placement to achieve the required number of pixels per-foot to provide sufficient video quality to meet the needs of the physical security manager.
- Network deployment models and the IP transports used by IP video surveillance cameras.
- The network requirements, video traffic flows between endpoints, bandwidth requirements are discussed.
- A design checklist to help the physical security and network manager achieve the goal of integrating IP video surveillance on the IP network.
- A case study for implementing IP video surveillance on a campus deployment.

# IP Video Surveillance Fundamentals Overview

This section provides an overview of why video surveillance deployments are migrating from analog-based systems to IP-based systems. The time between 2007 and 2010 represents a market transition in the industry where sales of IP-based components began out-selling analog-based systems. While analog systems have a cost advantage in small deployments (sixteen cameras or less), when larger number of cameras are deployed, IP-based systems may be more cost-effective initially and have a lower ongoing total cost of ownership. IP-based video surveillance systems, especially the end-node (the IP camera), have several operational and technological advantages. Why implement IP video surveillance over analog-based systems? The following subsections provide the answer.

## Leveraging VoIP Adoption

Many of the advantages of implementing IP video surveillance are similar to those of VoIP adoption. The fundamental reason is the cost savings of using the IP network for both voice and data. By adding the transport of video surveillance on the existing highly-available IP network, the cost savings realized from eliminating the separate cable plant for voice extends as well to the elimination of the separate cable plant for video.

Not only the wiring for media transport can be eliminated, but also the cabling for electrical power. As is the case with VoIP in the enterprise space, where the IP phone uses PoE, so does many fixed installation IP cameras. While power to some camera deployments continue to be a requirement (Pan-Tilt-Zoom housings, wireless cameras and cameras that require fibre connectivity due to distance), PoE is a substantial cost savings.

IP video surveillance cameras, once connected to the network, may be remotely configured and managed from a central command center. The installing technician must have a laptop to focus the lens and adjust the viewpoint of the camera, but following this initial installation, the camera configuration may be completed by a technician in a central, rather than local, facility.

## Access Video Any Time, Any Place

With IP-based systems, video feeds are encoded into Motion JPEG or MPEG-4/H.264 formats and stored as a digital image on a computer disk array. This provides the ability to access the video, by way of the networked digital video recorder, through the IP network at any time, from any place. These digital images do not degrade in quality from duplication like analog recordings on magnetic tape. They can be replicated and posted on web servers, distributed to law enforcement as E-mail attachments, and sent to news outlets. When analog-based systems were the norm, loss prevention/investigations staff may have to visit the location of the incident to view the video or a tape or DVD would need to be shipped by overnight courier. These inefficiencies no longer exist with IP-based systems and WAN connectivity to the physical location.

## Intelligence at the Camera

With IP cameras, local processing of the video image may be done during capture and analysis like motion detection and tampering detection logic may raise alerts by communicating with a central server. The alert may use a variety of IP protocols, SMTP (E-mail), Syslog, File Transfer (FTP), or a TCP socket connection with a small keyword in the payload. The Cisco 4500 IP Cameras have an additional DSP capabilities specifically designed to support real-time video analytics on the camera. This option is to allow analytic vendors to develop firmware in the future to run on these resources.

## Barriers to Success

While the advantages of an IP-based system are considerable, there are some barriers to success. They mainly revolve around the human element—job responsibilities, training, and education. Typically, the physical security manager and the network manager have no overlapping job responsibilities and therefore have little need to interact with each other. The physical security manager has job responsibilities targeted at loss prevention, employee and customer/visitor safety, security and crime prevention. Because of this, the physical security manager is more confident with a dedicated, reliable, physically separate cable plant.

Many installations of physical security cameras and the accompanying components are solely or partially implemented by value added resellers (VARs) who are specialists in their field, but not yet experts in IP networking. The VAR must become more fluent in internetworking and the network manager must understand the requirements of the physical security processes and applications.

The key elements of video surveillance is the three *Rs*: resolution, retention, and reliability. For an IP video surveillance deployment to be a success on the IP network, the *reliability* element must have careful attention by the network manager for the physical security manager to be successful.

# Video Resolutions

Resolution, one of the three *Rs* of video surveillance, directly influences the amount of bandwidth consumed by the video surveillance traffic. Image quality (a function of the resolution) and frame rate are functions of the amount of bandwidth required. As image quality and frame rate increase, so does bandwidth requirements.

## Analog Video Resolutions

Video surveillance solutions use a set of standard resolutions. National Television System Committee (NTSC) and Phase Alternating Line (PAL) are the two prevalent analog video standards. PAL is used mostly in Europe, China, and Australia and specifies 625 lines per-frame with a 50-Hz refresh rate. NTSC is used mostly in the United States, Canada, and portions of South America and specifies 525 lines per-frame with a 59.94-Hz refresh rate.

These video standards are displayed in interlaced mode, which means that only half of the lines are refreshed in each cycle. Therefore, the refresh rate of PAL translates into 25 complete frames per second and NTSC translates into 30 (29.97) frames per second. Table 4-1 shows resolutions for common video formats.

*Table 4-1 Analog Video Resolutions (in pixels)*

| Format | NTSC-Based | PAL-Based |
|---|---|---|
| QCIF | 176 × 120 | 176 × 144 |
| CIF | 352 × 240 | 352 × 288 |
| 2CIF | 704 x 240 | 704 x 288 |
| 4CIF | 704 × 480 | 704 × 576 |
| D1 | 720 × 480 | 720 × 576 |

Note that the linear dimensions of 4CIF are twice as big as CIF. As a result, the screen area for 4CIF is four times that of CIF with higher bandwidth and storage requirements. The 4CIF and D1 resolutions are almost identical and sometimes the terms are used interchangeably.

**Note** IP camera vendors may use different video resolutions. The Cisco Video Surveillance Manager solution supports the format delivered by the camera.

## Digital Video Resolutions

User expectations for resolution of video surveillance feeds are increasing partially due to the introduction and adoption of high-definition television (HDTV) for broadcast television. A 4CIF resolution, which is commonly deployed in video surveillance, is a 4/10th megapixel resolution. The HDTV formats are megapixel or higher. Table 4-2 lists the typical resolutions available in the industry.

*Table 4-2 Digital Video Surveillance Resolutions (in pixels)*

| Size/ Format | Pixels |
|---|---|
| **QQVGA** | 160x120 |
| **QVGA** | 320x240 |

| Size/ Format | Pixels |
|---|---|
| **VGA** | 640x480 |
| **HDTV** | 1280x720 |
| **1M** | 1280x960 |
| **1M** | 1280x1024 |
| **2M** | 1600x1200 |
| **HDTV** | 1920x1080 |
| **3M** | 2048x1536 |

While image quality is influenced by the resolution configured on the camera, the quality of the lens, sharpness of focus, and lighting conditions also come into play. For example, harshly lighted areas may not offer a well-defined image, even if the resolution is very high. Bright areas may be washed out and shadows may offer little detail. Cameras that offer wide dynamic range processing, an algorithm that samples the image several times with differing exposure settings and provides more detail to the very bright and dark areas, can offer a more detailed image.

As a best practice, do not assume the camera resolution is everything in regards to image quality. For a camera to operate in a day-night environment, (the absence of light is zero lux), the night mode must be sensitive to the infrared spectrum. It is highly recommended to conduct tests or pilot installations before buying large quantities of any model of camera.

**Tip** Some cameras rated as megapixel cameras in Motion JPEG only offer 4CIF resolution when configured for MPEG-4.

## Video Compression CODECS

The Cisco Video Surveillance Media Server supports IP endpoints that use Motion JPEG (MJPEG) or MPEG-4 codec technology. Both types of codecs have advantages and disadvantages when implemented in a video surveillance system. A system administrator may choose to use MJPEG on certain cameras and MPEG-4 or H.264 on others, depending on system goals and requirements.

A *codec* is a device or program that performs encoding and decoding on a digital video stream. In IP networking, the term frame refers to a single unit of traffic across an Ethernet or other Layer-2 network. In this guide, *frame* primarily refers to one image within a video stream. A video frame can consist of multiple IP packets or Ethernet frames.

A video stream is fundamentally a sequence of still images. In a video stream with fewer images per second, or a lower frame rate, motion is normally perceived as choppy or broken. At higher frame rates up to 30 frames per second, the video motion appears smoother; however, 15 frames per second video may be adequate for viewing and recording purposes.

Some of the most common digital video formats include the following:

- **Motion JPEG (MJPEG)** is a format consisting of a sequence of compressed Joint Photographic Experts Group (JPEG) images. These images only benefit from spatial compression within the frame; there is no temporal compression leveraging change between frames. For this reason, the level of compression reached cannot compare to codecs that use a predictive frame approach.

- **MPEG-1 and MPEG-2** formats are Discrete Cosine Transform-based with predictive frames and scalar quantization for additional compression. They are widely implemented, and MPEG-2 is still in common use on DVD and in most digital video broadcasting systems. Both formats consume a higher level of bandwidth for a comparable quality level than MPEG-4. These formats are not typically used in IP video surveillance camera deployments.
- **MPEG-4** introduced object-based encoding, which handles motion prediction by defining objects within the field of view. MPEG-4 offers an excellent quality level relative to network bandwidth and storage requirements. MPEG-4 is commonly deployed in IP video surveillance but will be replaced by H.264 as it becomes available. MPEG-4 may continue to be used for standard definition cameras.
- **H.264** is a technically equivalent standard to MPEG-4 part 10, and is also referred to as Advanced Video Codec (AVC). This emerging new standard offers the potential for greater compression and higher quality than existing compression technologies. It is estimated that the bandwidth savings when using H.264 is at least 25 percent over the same configuration with MPEG-4. The bandwidth savings associated with H.264 is important for high definition and megapixel camera deployments.

## MJPEG

An MJPEG codec transmits video as a sequence of Joint Photographic Experts Group (JPEG) encoded images. Each image stands alone without the use of any predictive compression between frames. MJPEG is less computation-intensive than predictive codecs such as MPEG-4, so can be implemented with good performance on less expensive hardware. MJPEG can easily be recorded at a reduced frame rate by only sampling a subset of a live stream. For example, storing every third frame of a 30-frame per second video stream will result in a recorded archive at 10 frames per second.

MJPEG has a relatively high bandwidth requirement compared to MPEG-4. A 640x480 VGA resolution stream running at 30 frames per second can easily consume 5 to 10 Mbps. The bandwidth required is a function of the complexity of the image, in conjunction with tuning parameters that control the level of compression. Higher levels of compression reduce the bandwidth requirement but also reduce the quality of the decoded image. Since there is no predictive encoding between frames, the amount of motion or change in the image over time has no impact on bandwidth consumption.

## MPEG-4/H.264

An MPEG-4 codec uses prediction algorithms to achieve higher levels of compression than MJPEG while preserving image quality. Periodic video frames called I-frames are transmitted as complete, standalone JPEG images similar to an MJPEG frame and are used as a reference point for the predictive frames. The remaining video frames (P-frames) contain only information that has changed since the previous frame.

To achieve compression, MPEG-4 relies on the following types of video frames:

- **I-frames** (intraframes, independently decodable)—These frames are also referred to as *key* frames and contain all of the data that is required to display an image in a single frame.
- **P-frames** (predictive or predicted frames)—This frame type contains only image data that has changed from the previous frame.
- **B-frames** (bi-directional predictive frames)—This frame type can reference data from both preceding frames and future frames. Referencing of future frames requires frame reordering within the codec.

The use of P-frames and B-frames within a video stream can drastically reduce the consumption of bandwidth compared to sending full image information in each frame. However, the resulting variance of the video frames' size contributes to the fluctuation in the bandwidth that a given stream uses. This is the nature of most codecs because the amount of compression that can be achieved varies greatly with the nature of the video source.

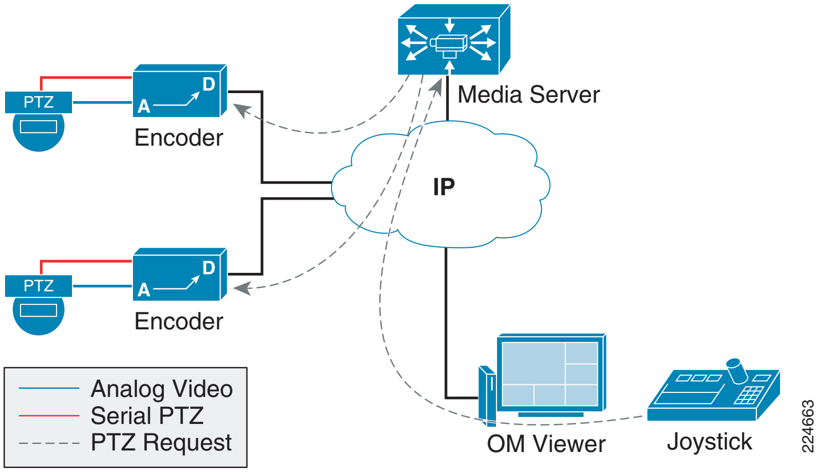
# Pan-Tilt-Zoom (PTZ)

The Cisco Video Surveillance Manager solution supports the configuration of PTZ cameras connected to encoders or as IP cameras. In order to support PTZ connectivity, the encoder should be able to connect to the camera through a serial interface. The Video Surveillance Manager solution supports the following PTZ protocols:

- Bosch
- Cohu
- J2 Vision
- Pelco D
- Pelco P

Figure 4-1 shows how an analog camera can be connected to an IP encoder to convert its video feed to an IP video format. The encoder also connects through a serial cable to the analog camera. When the OM viewer requests PTZ control through the joystick, the Media Server intercepts the request and communicates the request to the encoder. Once the request is received by the encoder, a serial communication takes place between the encoder and the analog camera.

*Figure 4-1 Pan-Tilt-Zoom Via Encoders*



# Aspect Ratio

The aspect ratio is the relationship between the number of pixels in the horizontal and vertical image dimensions. A 4:3 (1.33:1) aspect ratio is universal for standard definition cameras. For HDTV formats, 16:9 (1.78:1) is universal. In video surveillance deployments, the HDTV aspect ratio is more advantageous because the pixels at the top and bottom of the image are generally of less importance than having a wide field of view. In other words, the width of the image is more important than the height of the image. Capturing, encoding, and transporting bits that are of little value is a waste of bandwidth and disk space. In some instances, a single HDTV format video camera may be able to replace two standard definition cameras.

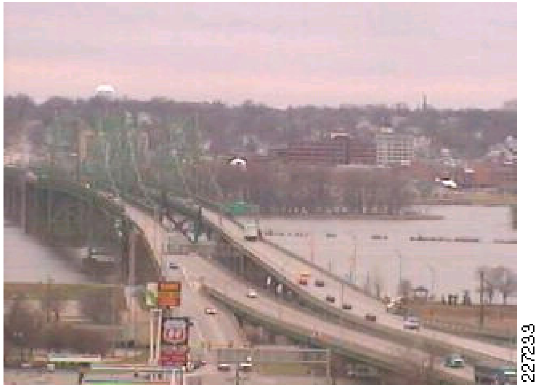
## Camera Placement

Camera placement can be characterized by either overview or detail view. The camera placement influences the resolution, frame rate and codec in use.

### Overview

A camera with an overview scene is monitoring a large area such as a parking lot or a traffic camera that is viewing vehicle congestion or the number of cars parked in the lot. Because details are not important, standard definition cameras using a wide-angle lens may be sufficient. The preferred codec may be MPEG-4 with a relatively low frame rate, 1-5 frames per second. Figure 4-2 shows an example of an overview scene.

***Figure 4-2 Overview Scene***



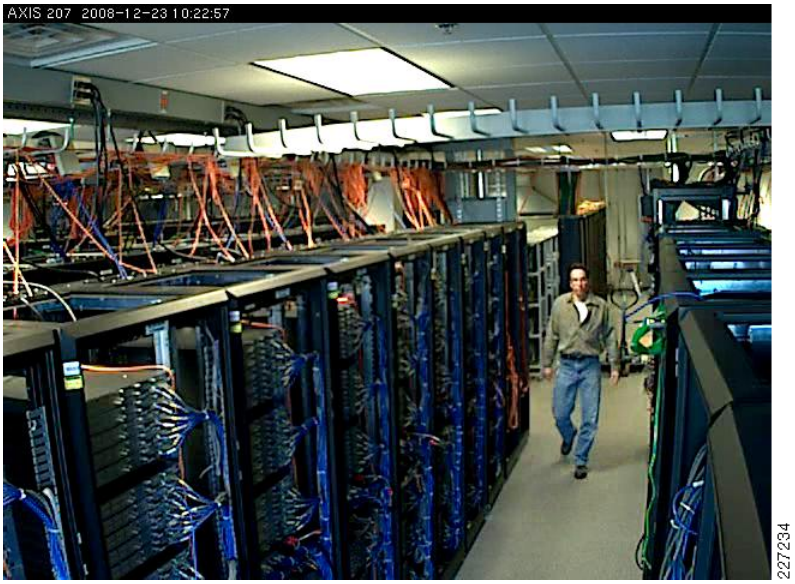
*AbbeyCam is a streaming video of the Iowa side of the I-74 bridge as seen from the Abbey Hotel in Bettendorf*

Overview cameras may be supplemented with a detail view camera focused on a key area of interest or by a PTZ camera to provide real-time analysis of areas of interest at a higher resolution.

### Detail View

The detail view placement is targeted at observing a specific area of interest at a higher resolution than the overview. Detail view is used for Point-of-sale transactions and face or license plate recognition. The detail view may have a PTZ capability, or the camera may be close to the subject area or have a long focal length lens. Megapixel or HD cameras may be deployed to provide a sufficient number of pixels per-foot to accurately represent the subject. Figure 4-3 is an example of a detail view, the camera positioned to identify a subject passing through a confined area.

*Figure 4-3 Detail View Placement*



The positioning of a camera for detail view is a function of the number of pixels per-foot required for the application.

# Detection, Recognition, Identification

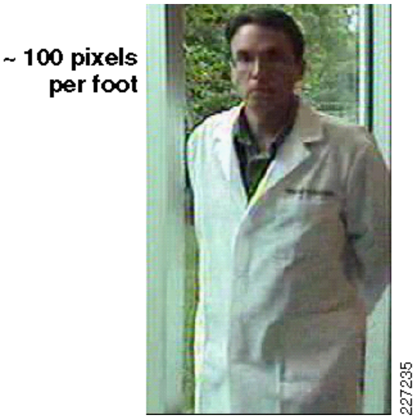
Detection, recognition, and identification are visual processes associated with the amount of detail discernable to the human eye. We detect an object when it enters the field of view. Detection means we are aware that an object (or person) now exists where previously it was not seen. Usually, this is due to movement of the object into the field of view of the surveillance camera. Detection simply means we are aware of the object, but have too little details to recognize or identify the object.

As the object moves closer, we may recognize the object from characteristics previously encountered. For example, aircraft recognition is taught to military ground troops and airmen. All aircraft have wings, engines, a fuselage, and tail assembly. They differ in size, shape, number, and position to each other. A particular model of aircraft can be recognized by recalling these characteristics from pictures, drawings or past detailed observations.

Identification is the process where sufficient details are available to uniquely discern a person or object that is previously unknown. Identification requires sufficient detail to accurately describe or recall the characteristics of the subject at a later time. For example, a mug shot (booking photograph) is taken following the arrest of a subject as a means of photographing (recording) sufficient details for later identification by a victim or witness. In video surveillance terms, sufficient detail is calibrated in pixels per foot of the area recorded by the camera.

The number of pixels per-foot to identify a subject may, at a minimum, range from 40 to over 150. If the goal, therefore, is to identify a person entering through a standard 7-foot high doorway, the camera would need to be positioned so that the pixel per-foot requirement covering the door is met. The door would then need to be covered by 1050 pixels, if the goal is to have 150 pixels per foot; *7 feet x 150* pixels per foot. Figure 4-4 provides an example of an image with approximately 100 pixels per foot for reference.

*Figure 4-4 Pixels Per Foot*



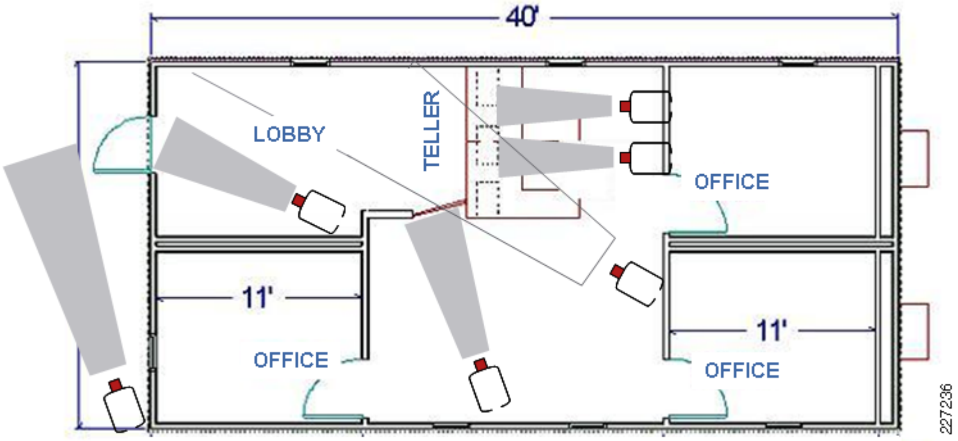
As shown in Figure 4-4, the video surveillance image is subject to uneven lighting, the subject is standing near a large window of a lab environment. There is little light from the internal space with the natural light entering from the side and rear in this scene. This image is from an analog camera that does not include a wide-dynamic range processing that would improve the image quality in this deployment. This illustrates the point that the number of pixels alone does not guarantee a high quality image.

## Number of Cameras per Location

The number of cameras at any one building or facility may vary greatly depending on the coverage requirements and the nature of the business. While there are some small office deployment scenarios where only a single IP camera is needed, in most cases even a small office will require more cameras that one might initially expect.

Using a small, two teller bank branch as an example, consider the number and placement of cameras in the example shown in Figure 4-5.

*Figure 4-5 Camera Deployment Floor Plan*



There is a camera behind each teller station, a camera on the main entrance (both inside and outside), and two cameras in the inner office area focused on the lobby and half doorway leading into the manager office areas. Additionally, the parking lot area, side, front, and rear of the branch as well as any exterior

ATM would need be covered. This small location may easily require 10 to 16 IP cameras. The Cisco Video Management and Storage System (VMSS) Network Module for the ISR router is targeted at a 16 to 32 camera deployment any may be implemented in this branch location.

Larger facilities require more cameras per location. It is not uncommon for a large retail store, home center, or warehouse retailer to need 100 to 200 IP cameras per location. Public school deployments may need 80 to 100 cameras per building.

**Tip** One advantage of deploying high definition cameras over standard definition is fewer cameras may be required to cover an area of interest with a similar number of pixels per foot.

## Frame Rates

As image quality and frame rate increase, so does bandwidth requirements. The frame rate selected must meet the business requirements, but it does not need to be higher than what is required and should be considered carefully as frame rate influences both bandwidth and storage requirements.

Motion pictures are captured at 24 frames per second (fps). The human eye/brain sees images captured at 24 fps as fluid motion. Televisions use 25 fps (PAL) or 30 fps (NTSC) as does analog video cameras. These full motion rates are not needed for all video surveillance applications and in most applications less than 12 to 15 fps is sufficient.

The following are some industry guidelines:

- Nevada Gaming Commission (NGC) standards for casinos—30 fps
- Cash register, teller stations—12 to 15 fps
- School or office hallways —5 fps
- Parking lots, traffic cameras, overview scenes —1 to 3 fps
- Sports Stadiums on non-event days, less than 1 fps

## Movement in Relation to Camera Placement

If the camera is placed where the subject moves toward the camera or vertically, the number of frames per second can be less than if the subject moves from side to side or horizontally within the field of view. The velocity of the subject is also a consideration. A cameras observing persons jogging or riding a bicycle may require higher frame rates than a person walking.

## Progressive Scanning

Analog cameras capture images using an interlaced scanning method, odd and even scan lines are done alternately. There is approximately 17 ms delay between the scanning of the odd and even lines making up the entire image. Because of this slight delay between scan passes, objects that are moving in the frame may appear blurred while stationary objects are sharp. Most IP cameras use a progressive scan that is not subject to this problem. Everything being equal, a progressive scan image has less motion blurring than an interlace scanned image.

## Wide Dynamic Range Imaging

The Cisco 2500 Series Video Surveillance IP Camera offer wide dynamic range imaging. This technology increases the image quality in harsh lighting conditions, including back lighted scenes or indoor/outdoor areas such as loading docks or stadiums.

# IP Transport

IP cameras and encoders communicate with the Media Server in different ways, depending on the manufacturer. Some edge devices may support only MJPEG over TCP, while others may also support MPEG-4 over UDP.

## TCP

MJPEG is typically transported through TCP. TCP provides guaranteed delivery of packets by requiring acknowledgement by the receiver. Packets that are not acknowledged will be retransmitted. The retransmission of TCP can be beneficial for slightly congested network or networks with some level of inherent packet loss such as a wireless transport. Live video rendering at the receiving end may appear to stall or be choppy when packets are retransmitted, but with the use of MJPEG each image stands alone so the images that are displayed are typically of good quality.

## UDP/RTP

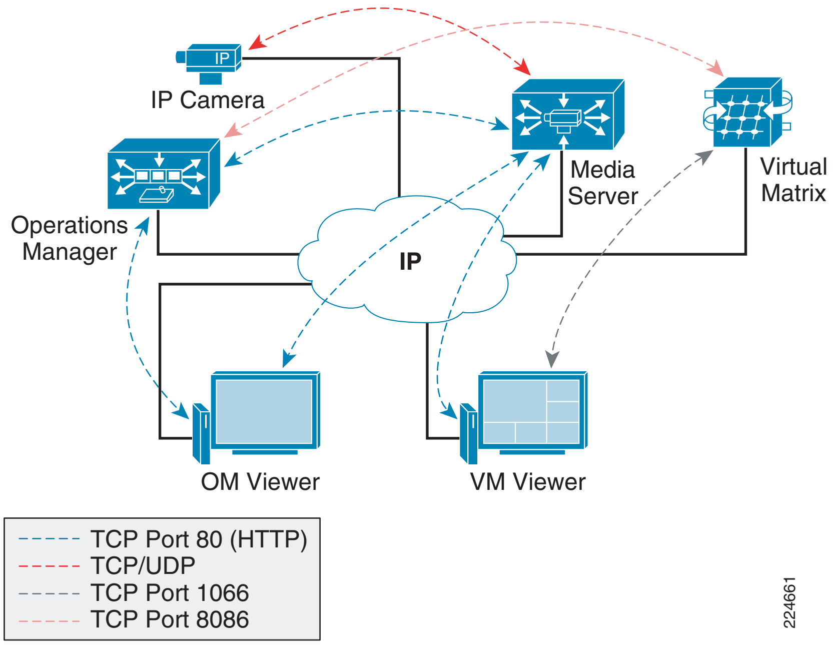
MPEG-4/H.264 video is typically transmitted over UDP or Real-time Transport Protocol (RTP). UDP does not guarantee delivery and provides no facility for retransmission of lost packets. RTP/UDP transport is most suitable for networks with very little packet loss and bandwidth that is guaranteed through QoS mechanisms. MPEG-4 over RTP/UDP is relatively intolerant to packet loss; if there is loss in the stream, there will typically be visible artifacts and degradation of quality in the decoded images. UDP transport does provide the option of IP multicast delivery, where a single stream may be received by multiple endpoints. In an IP multicast configuration, the internetworking devices handle replication of packets for multiple recipients. This reduces the processing load on the video encoder or IP camera and can also reduce bandwidth consumption on the network.

Some IP cameras and encoders also provide for TCP transport of MPEG-4. TCP encapsulation can be beneficial for networks with inherent packet loss. TCP may be useful especially for fixed cameras and streams that are only being recorded and not typically viewed live. TCP transport induces a little more latency in the transport due to the required packet acknowledgements, so may not be a desirable configuration for use with a PTZ controlled camera.

## Required TCP/UDP Ports

The example in Figure 4-6 shows that the communication between the Media Server and viewers relies on TCP port 80 (HTTP), while the communication between edge devices and the Media Server may vary. The communication between the Virtual Matrix Server and the VM monitor is typically over TCP port 1066 while the communication between the Virtual Matrix Server and the Operations Manager is typically over TCP port 8086.

*Figure 4-6 TCP/UDP Ports*

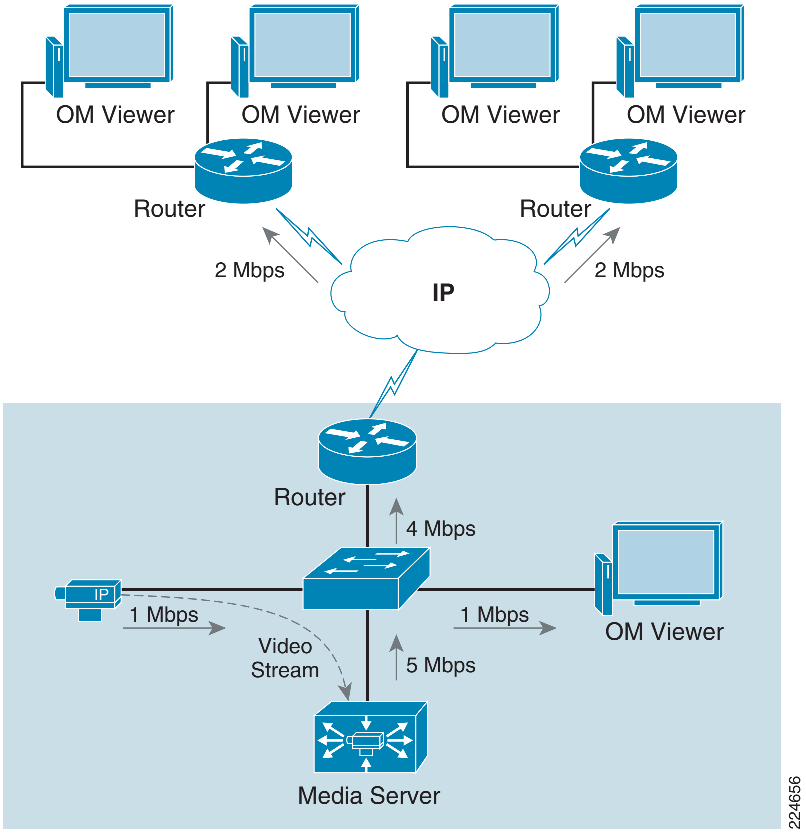


# IP Unicast

Applications that rely on unicast transmissions send a copy of each packet between one source address and one destination host address. Unicast is simple to implement but hard to scale if the number of sessions is large. Since the same information has to be carried multiple times, the impact on network bandwidth requirements may be significant.

The communication between the Media Server and the viewers is always through IP unicast, making the Media Server responsible for sending a single stream to each viewer. The example in Figure 4-7 shows five viewers requesting a single video stream from the Media Server. Assuming a single 1Mbps video feed, the bandwidth requirements are noted throughout each network link.

*Figure 4-7 IP Unicast Traffic*



**Note** The Media Server only supports IP unicast between the Media Server and the viewers.

# Network Deployment Models

This chapter provides a high-level overview of different deployment models and highlights the typical requirements of campus and wide area networks. Cisco's Enterprise Systems Engineering team offers detailed network designs that have been deployed by enterprise customers to provide enhanced availability and performance. These designs may be found at the Cisco Validated Design Program site at: http://www.cisco.com/go/cvd.

## Campus Networks

An infrastructure that supports physical security applications requires several features from a traditional campus design. A hierarchical campus design approach has been widely tested, deployed, and documented. This section provides a high-level overview and highlights some of the design requirements that may apply to a video surveillance solution. For a more detailed review of Campus designs refer to the Campus Design documents in References, page A-18.

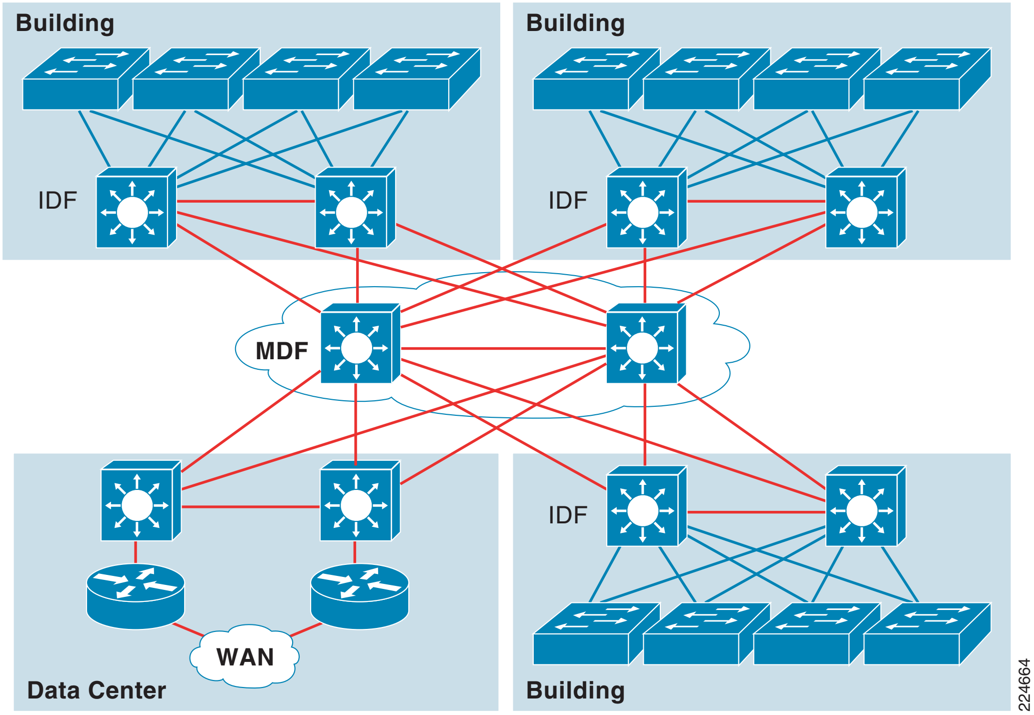
A traditional campus design should provide the following:

- **High availability**—Avoid single points of failure and provide fast and predictable convergence times.
- **Scalability**—Support the addition of new services without major infrastructure changes.
- **Simplicity**—Ease of management with predictable failover and traffic paths.

A highly available network is a network that provides connectivity at all times. As applications have become more critical, the network has become significantly more important to businesses. A network design should provide a level of redundancy where no points of failure exist in critical hardware components. This design can be achieved by deploying redundant hardware (processors, line cards, and links) and by allowing hardware to be swapped without interrupting the operation of devices.

The enterprise campus network shown in Figure 4-8 is a typical campus network. It provides connectivity to several environments such as IDFs, secondary buildings, data centers, and wide area sites. An Intermediate Distribution Frame (IDF) is the cable infrastructure used for interconnecting end user devices to the Main Distribution Frame (MDF) or other buildings and is typically located at a building wiring closet.

***Figure 4-8 Campus Network***



Quality-of-service (QoS) is critical in a converged environment where voice, video, and data traverse the same network infrastructure. Video surveillance traffic is sensitive to packet loss, delay, and delay variation (jitter) in the network. Cisco switches and routers provide the QoS features that are required to protect critical network applications from these effects.

## Hierarchical Design

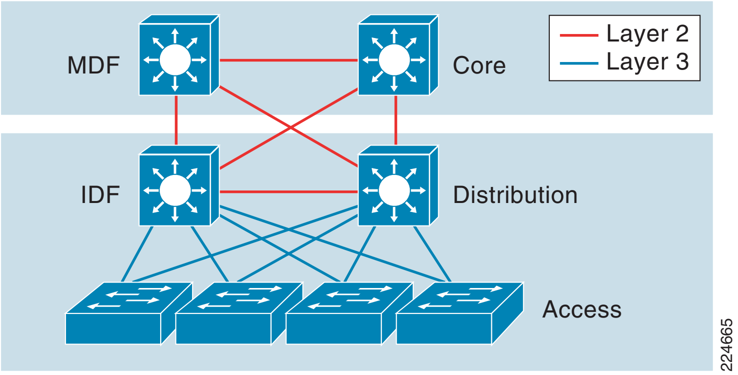
The goal of a campus design is to provide highly available and modular connectivity by separating buildings, floors, and servers into smaller groups. This multilayer approach combines Layer 2 switching (based on MAC addresses) and Layer 3 switching or routing (based on IP address) capabilities to achieve a robust, highly available campus network. This design helps reduce failure domains by providing appropriate redundancy and reducing possible loops or broadcast storms.

With its modular approach, the hierarchical design has proven to be the most effective in a campus environment. The following are the primary layers of a hierarchical campus design:

- **Core layer**—Provides high-speed transport between distribution-layer devices and core resources. The network's backbone.
- **Distribution layer**—Implements policies and provides connectivity to wiring closets. This layer provides first-hop redundancy such as Hot Standby Router Protocol (HSRP) and Gateway Load Balancing Protocol (GLBP).
- **Access layer**—User and workgroup access to the network. Security and QoS can be defined at this layer and propagated to the higher layers.

Figure 4-9 shows a typical campus design with the three main layers.

***Figure 4-9 Hierarchical Campus Design***



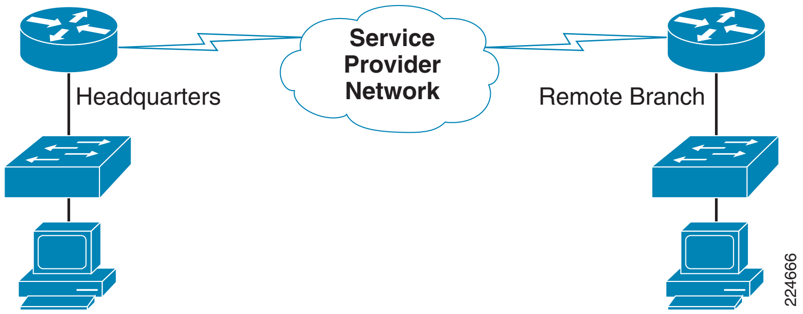
In smaller environments, it is typical to collapse the distribution and core layers into a single layer.

# Wide Area Networks

A wide-area network (WAN) is used to connect different local-area networks (LANs) and typically covers a broad geographic area. WAN services are leased from service providers who provide different speeds and connectivity options.

Figure 4-10 shows how a remote branch office relies on the connectivity provided by a WAN service provider.

*Figure 4-10 Service Provider Network*



Deploying a video surveillance solution through a WAN environment presents challenges that are not typically seen in a LAN. In a LAN environment it is common to see 1 Gbps and 10 Gbps of bandwidth, while in a WAN environment, most connections are less than 10 Mbps; many remote connections operate on a single T1 (1.544 Mbps) or less.

These inherent bandwidth constraints require careful evaluation of the placement of cameras and Media Servers and how many viewers can be supported at remote sites simultaneously. By using child proxies, bandwidth requirements can be reduced to transport video streams across WAN connections.

The placement of recording devices also becomes important. The video may be streamed to a central site using lower frame rates or resolution, but another attractive alternative is to deploy Media Servers at the remote sites and stream the traffic using the LAN connectivity within the remote site.

Table 4-3 and Table 4-4 show typical links that are offered by service providers.

*Table 4-3 Service Provider Links (1)*

| Digital Signal Level | Speed | "T" | Channels or DS0s |
|---|---|---|---|
| DS0 | 64 kbps | – | 1 |
| DS1 | 1.544 Mbps | T1 | 24 |
| DS3 | 44.736 Mbps | T3 | 672 |

*Table 4-4 Service Provider Links (2)*

| SONET Signal Level | Speed | SDH Equivalent |
|---|---|---|
| STS-OC-1 | 51.84 Mbps | STM-0 |
| STS-OC-3 | 155.52 Mbps | STM-1 |
| STS-OC-12 | 622.08Mbps | STM-4 |
| STS-OC-48 | 2488.32 Mbps | STM-16 |
| STS-OC-192 | 9.952 Gbps | |

A point-to-point or leased line is a link from a primary site to a remote site using a connection through a carrier network. The link is considered private and is used exclusively by the customer. The circuit usually is priced based on the distance and bandwidth requirements of the connected sites.

Technologies such as Multilink PPP allow several links to be bundled to appear as a single link to upper routing protocols. In this configuration, several links can aggregate their bandwidth and be managed with only one network address. Because video surveillance traffic requirements tend to be larger than other IP voice and data applications, this feature is attractive for video surveillance applications.

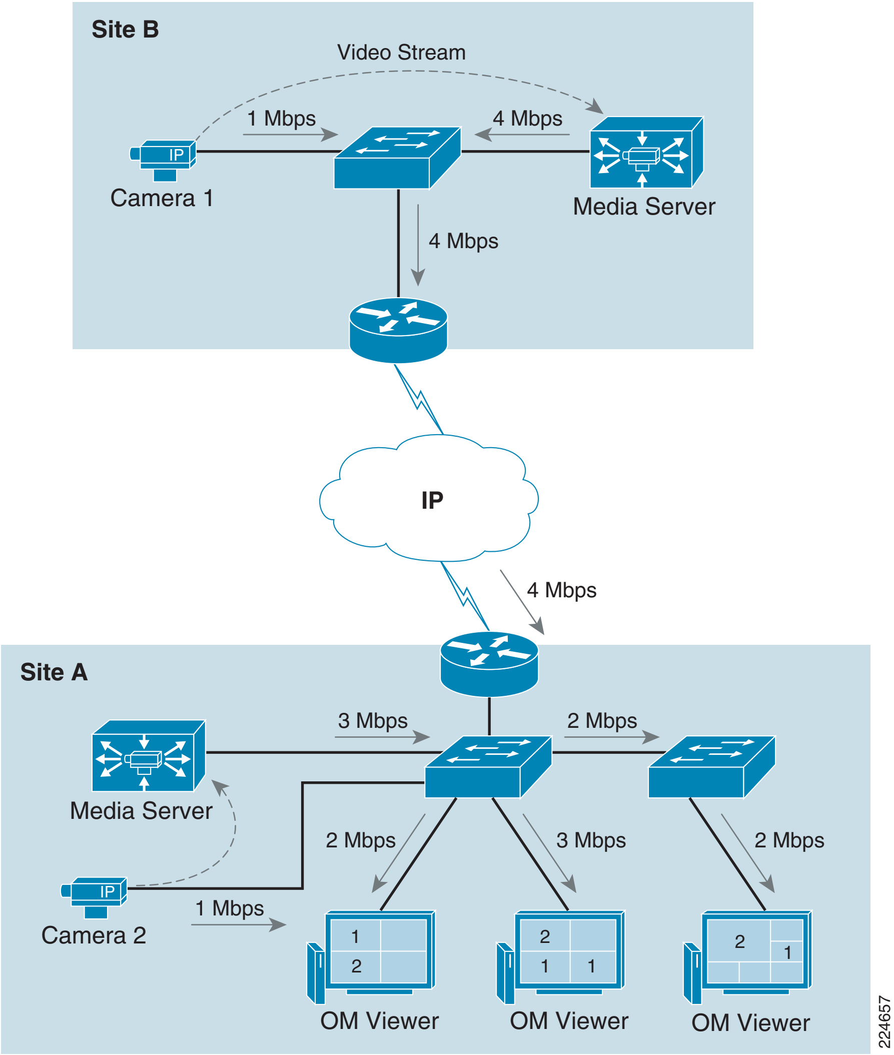
Hub-and-spoke, also known as star topology, relies on a central site router that acts as the connection for other remote sites. Frame Relay uses hub-and-spoke topology predominantly due to its cost benefits, but other technologies, such as MPLS, have mostly displaced Frame Relay.

## Example 1: Network Bandwidth Usage

Figure 4-11 shows a simple scenario with two sites. Each site has a Media Server and each Media Server is the direct proxy for an IP camera. Three OM Viewers are active in Site A and each IP cameras is generating 1Mbps of network traffic. For simplicity the Operations Manager has been removed from this graphic.

Two OM Viewers are displaying video streams from Camera 1 and Camera 2 while one OM Viewer is displaying three video streams: two streams from Camera 1 and one stream from Camera 2. The network bandwidth required to display video streams for Camera 2 in Site A are relatively small for a LAN environment, but the traffic from Camera 1 can be significant for WAN environments since four different 1Mbps streams have to traverse the WAN locations.

*Figure 4-11 Network Bandwidth Requirements*



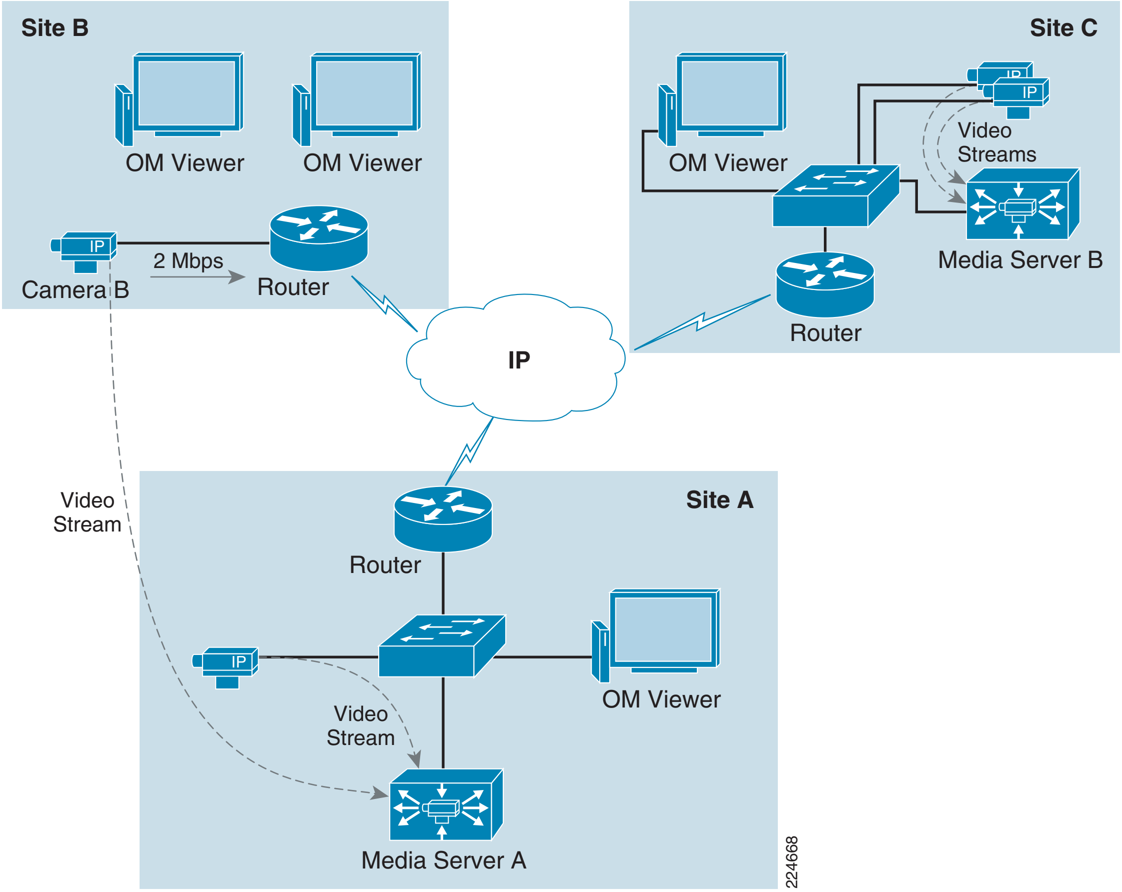
## Example 2: Sites with Remote Storage

Figure 4-12 shows how Media Servers can be deployed at different WAN locations in order to minimize the bandwidth requirements. By deploying the Media Servers close to viewers and edge devices, the network traffic remains local to each site. Archiving video streams at each location is also an attractive solution to minimize the network traffic between sites.

In this example Site A and Site C have Media Servers acting as direct proxies and archives for the IP cameras. Since both sites are archiving and distributing video to the OM Viewers locally, the network traffic remains local to each site.

Site B can function without a local Media Server, but all video streams have to traverse the WAN connections. Since Media Server A is the direct proxy for Camera B, the 1Mbps stream has to reach Media Server A before reaching any OM Viewers. A total of 3Mbps would be required in order for both OM Viewers in Site B to receive video from Camera B.

*Figure 4-12 Sites with Remote Storage*



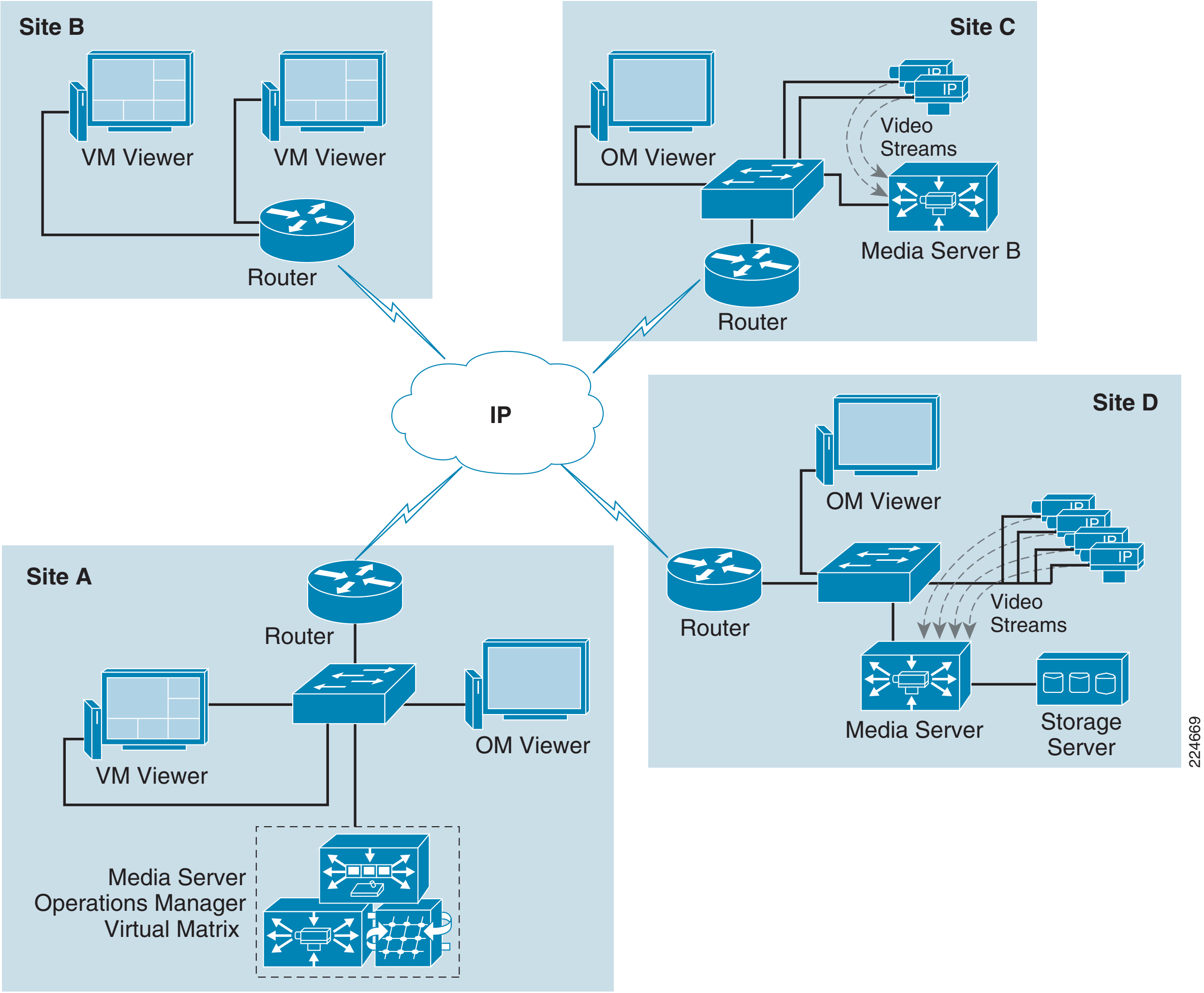
## Example 3: Virtual Matrix Scenario

Figure 4-13 shows an example that includes a Virtual Matrix Server and VM monitors located at two different sites. The Server on Site A is acting as the Media Server, Operations Manager, and Virtual Matrix for the environment. In order to reduce bandwidth traffic, Media Servers are also installed on Site C and Site D.

A single Operations Manager and a single Virtual Matrix are adequate to support this scenario. Since the cameras are located on Site C and Site D, they are able to serve the local OM Viewers at those sites.

The Media Server on Site A can also be configured with child feeds that come from the remote Media Servers and provide those feeds locally to viewers and monitors on Site A.

*Figure 4-13 Virtual Matrix Scenario*



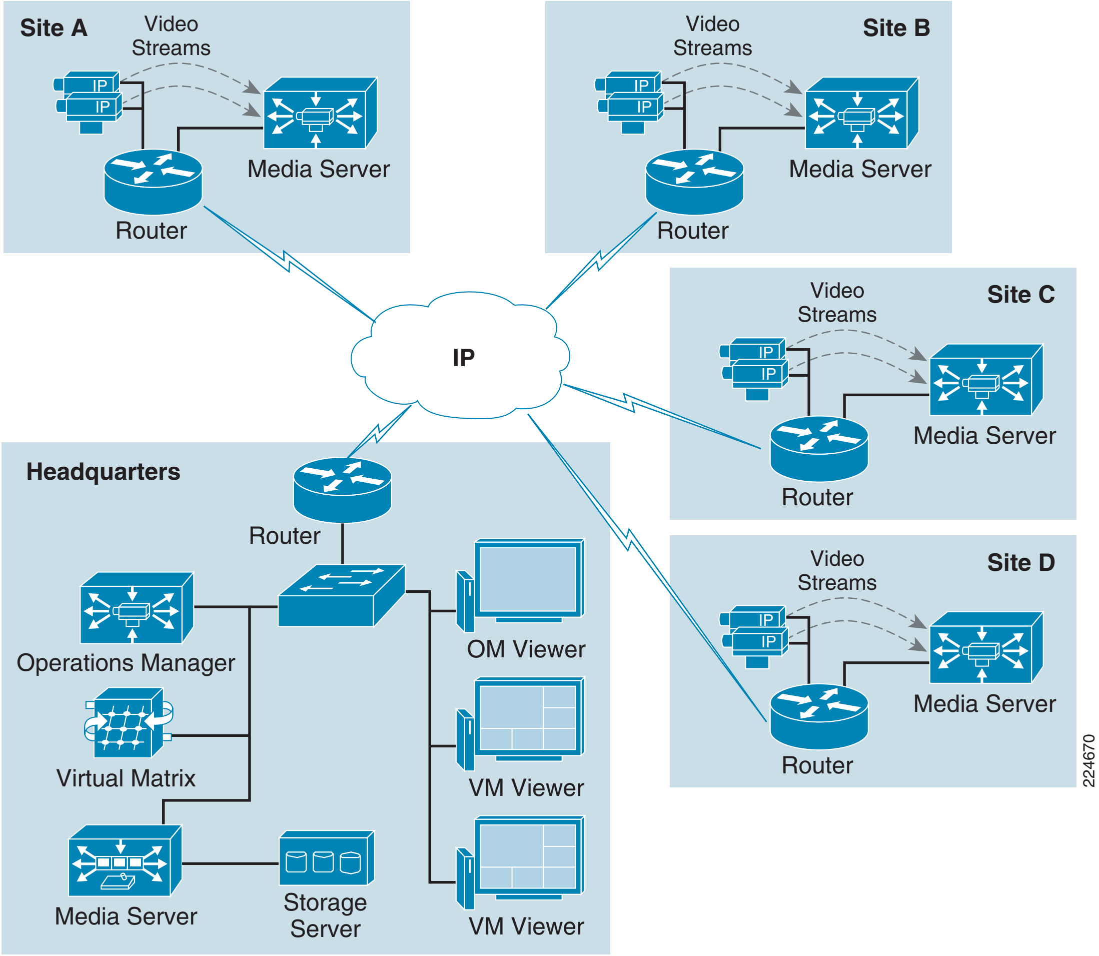
## Example 4: Distributed Media Servers

Figure 4-14 shows a deployment with several remote sites, each with a local Media Server acting as the direct proxy and archive for local IP cameras.

In this scenario, all recording occurs at the remote sites and live video streams are viewed by OM Viewers and VM monitors (video walls) at the headquarters.

The Media Server at the headquarters could also have Parent-Child proxies to each remote Media Server and request the remote streams only when required at the headquarters. This would have less bandwidth impact when the same stream is requested by more than one viewer since the traffic would be contained locally in the headquarters LAN.

*Figure 4-14 Distributed Media Servers*



# Network Requirements

This section provides an overview about the branch and campus network requirements to support IP video surveillance.

## Power-over-Ethernet (PoE)

The ability to provide electrical power to an IP camera is an important aspect of IP video surveillance that is not available in analog deployments. Analog deployments require some external power supply to meet the power requirements of the cameras. IP cameras with external PTZ housings, outdoor-rated IP cameras, wireless and IP cameras that must use fibre LAN connections due to distance limitations of copper Ethernet wiring may continue to required an external power supply. However, PoE is an important cost-savings factor for IP video surveillance.

## LAN Switches and Provisioning

In "Campus Implementation Case Study" section on page 4-31 and "Configuring Quality-of-Service (QoS) for IP Video Surveillance" section on page 6-21, LAN switching requirements are covered in the necessary detail for a successful deployment. There are several requirements for LAN switches, the primary being the ability to support the 802.1af PoE standard for those cameras that can make use of this feature. Also, aggregate backplane capacity as well as uplink capacity is important. At a minimum, switches should have 1Gigbps or 10Gigbps uplink and a 32Gbps effective backplane capacity. QoS support is also important, the ability to both trust the Layer-3 QoS markings (DSCP) and to set DSCP on ingress is critical. Most of commercially available switches support VLANs and trunking and these features are critical for segmenting IP video surveillance traffic into its own domain.

Support of features like port security, 802.1x port-based Network Access Control (NAC), Spanning Tree Protocol (STP) PortFast feature, and PortFast Bridge Protocol Data Unit (BPDU) guard are also useful. Because this design guide recommends marking video surveillance media streams as DSCP value CS5, switches that are configured by default for VoIP implementations are recommended as the media feeds will align with the default VoIP configurations.

## Storage Requirements

In general, the recommendation is to store data as close to the source as practical. This is especially true of branch location deployments. By storing video archives locally, IP video surveillance may be deployed to locations with traditional WAN links that would otherwise not have enough bandwidth to transport the archives to a central campus/data center facility. The WAN, however, may still have sufficient bandwidth to transport or view video clips to aid in investigations or other forensic analysis. By storing locally and only transporting the small amount of video surveillance data that is needed centrally, video surveillance can be network-enabled today and tied into other BMS and analytics solutions that can benefit the business.

## IP Addressing Requirements

If the network manager plans on implementing some segmentation and path isolation over the LAN/WAN by using VRF-Lite and VLANS, the IP addressing scheme may have greater flexibility than if the video surveillance networks are routed in the global routing table. However, the general recommendations for addressing IP video surveillance devices are as follows:

- Allocate an addressing scheme distinct from the end-user address space at the branch locations.
- Allocate network addressing so that IP cameras, servers, encoders, workstations, and any building management (BM) devices can be allocated from the address space at the remote locations.
- Allocate addressing to facilitate summarization of network advertisements to the core.
- Allocate addressing so that one network advertisement can encompass the entire address space for physical security (and building management) devices.
- Configure a loopback address on remote routers from the physical security address space for NTP and other management functions.

**Tip** Because the IP cameras are using static IP addresses, give careful attention to IP addressing deployed as reallocating IP addressing is more time consuming than when all end nodes use dynamically assigned IP addresses from a DHCP server.

## Requirements for Loss, Latency and Jitter Video Flows

IP video surveillance media streams based on MPEG-4/H.264 are primarily sensitive to packet loss and latency; jitter is less of an issue. One way latency in the range of 20 to 40ms, jitter 1 to 2ms with very low or no loss will provide acceptable video quality. With Motion JPEG, which is generally transported via TCP, as latency increases, the offered frame rate will need to decrease (frames must be skipped) to account for the increase in round trip time. As a best practice, loss approaching 1 percent will introduce video quality issues in MPEG-4/H.264. The Mean Opinion Score (MOS) reported by a IP SLA UDP jitter operation can be used to provide an initial assessment of the ability of the network to provide acceptable video quality. Be cautioned, however, that the IP SLA UDP jitter operation does not take into account that video requires substantially more bandwidth than VoIP.

For sample configurations, refer to the "IP SLA Probe Sample Configurations" section on page A-1.

## QoS

QoS should be implemented on both LAN and WAN and should align with the end-to-end QoS policy of the enterprise network. The "Configuring Quality-of-Service (QoS) for IP Video Surveillance" section on page 6-21 provides useful information as a baseline for enabling QoS on the network. The physical security manager and the network manger must work together to understand the bandwidth requirements of the video surveillance implementation and once sufficient LAN and WAN bandwidth is provisioned, enable QoS so both media and control plane traffic is protected during congestion.

## Performance Routing

Performance Routing (PfR) is a Cisco IOS feature that extends and enhances the capabilities of traditional routing protocols to allow for rerouting around path brownouts or temporary blackouts, packet loss, or links that are least preferred due to latency. The technology uses NetFlow and IP SLA probes to determine the characteristics of equal-cost links and select the best link, or a link that meets the stated performance characteristics. PfR can also be more effective at load sharing than IP CEF, because it takes into consideration the interface utilization, which CEF does not.

The "Performance Routing (PfR) Integration" section on page 6-50 not only uses this feature to optimize video feeds, it also provides insight on how latency, jitter, and packet loss influences video quality.

## Wide Area Application Services

The primary video encoding methods, Motion JPEG and MPEG4 / H.264, are compressed by the encoding function of the IP camera. Because they are already compressed, the compression functions of WAAS (LZ or DRE compression) does not provide as dramatic a compression-ratio as is seen with data packets. Because of this, implementing WAAS can be implemented to optimize data applications. It would benefit video by freeing up available bandwidth for transport of video surveillance. For more information, refer to "Wide Area Application Services (WAAS) Integration" section on page 6-61 and "Wide Area Application Services (WAAS) for iSCSI" section on page 6-74.

## Redundancy

Wherever possible, the path between IP cameras and the target Media Server should include as much redundancy as practical. In the "Campus Implementation Case Study" section on page 4-31, the dual uplinks from the access switches to the distribution layer switches and between distribution and core switches have at least two paths. In the "Performance Routing (PfR) Integration" section on page 6-50 and in the "Virtualization, Isolation and Encryption of IP Video Surveillance" section on page 6-87, dual LAN/WAN links are deployed. Firewalls should be deployed in an active-standby failover configuration.

Where multiple IP cameras are covering a critical viewing area, connect cameras with the overlapping vantage points to separate access-layer switches. If the access-layer switch becomes unserviceable an alternate image from the overlapping camera may be usable.

A key element of redundancy is fault management. While having an alternate path provides availability, the network management services in the enterprise must be able to detect failures and initiate remedial action in a timely manner.

## VLANs

As a best practice, where possible, segregate IP cameras, servers, and viewing stations on a VLAN separate from other devices on the network. Combined with allocating a separate IP addressing scheme for these physical security and building management endpoints, this facilitates controlling access from other users in the global routing table. Should the enterprise decide to implement end-to-end path isolation and segmentation, these VLANs can then easily be mapped to a VRF-enabled interface on the supporting router.

## Segmentation, Security, Firewalls and IPSec Encryption and Firewalls

IP video surveillance, access control systems, and related building management systems are a prime candidate for being deployed on the IP network in their own domain. The applications that enable these systems have a high likelihood for interconnecting currently on in the future. Additionally, these applications have a relatively small user group with the organization; the percentage of employees in an organization relating to facilities, loss prevention, and security is typically very low. If the systems are compromised, the end result can be a very notable event that may be highly publicized. Corporations do not want their video surveillance images published on YouTube or a someone raising or lowering the building temperature in their corporate headquarters from half a world away.

Because of these reasons, isolating and protecting these resources through segmentation techniques of VLANs and VRF-Lite should be a design consideration. Additionally, encrypting and protecting access to the address space through firewalls and access control lists can be deployed with or without the control plane segmentation. The "Virtualization, Isolation and Encryption of IP Video Surveillance" section on page 6-87 demonstrates these techniques.

# Video Traffic Flows

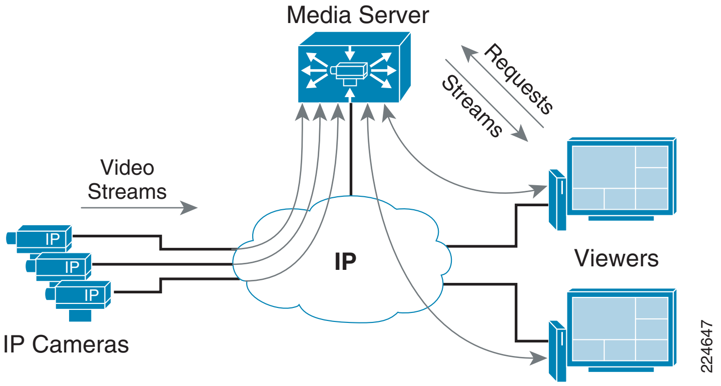
Each Video Surveillance Manager application plays a unique role in the deployment of a complete video streaming solution. When deploying and operating a Video Surveillance Manager environment, it is important to understand the video traffic flows of each application and how they interact with the system as a whole.

# Video Surveillance Media Server

The Video Surveillance Media Server is the core component of the solution, providing for the collection and routing of video from IP cameras to viewers or other media servers. The system is capable of running on a single physical server or distributed across the network, scaling to handle thousands of cameras and users.

Figure 4-15 shows how IP cameras or encoders send a single video stream to the Media Server. The Media Server is responsible for distributing live and archived video streams to the viewers simultaneously over an IP network.

***Figure 4-15 Media Server***



For archive viewing, the Media Server receives video from the IP camera or encoder continuously (as configured per the archive settings) and only sends video streams to the viewer when requested.

In environments with remote branch locations, this becomes very efficient since traffic only needs to traverse the network when requested by remote viewers. Branch office traffic remains localized and does not have to traverse wide area connections unless is requested by users other users.

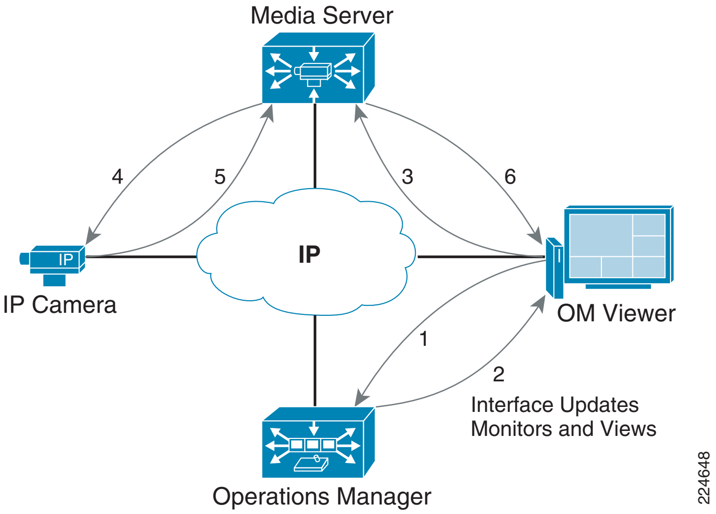
Video requests and video streams are delivered to the viewer using HTTP traffic (TCP port 80).

# Video Surveillance Operations Manager

Viewers can access the Operations Manager through their web browser. The Operations Manager is responsible for delivering a list of resource definitions, such as camera feeds, video archives, and predefined views to the viewer. Once this information is provided to the viewer, the viewer communicates with the appropriate Media Server to request and receive video streams.

Figure 4-16 shows the traffic flow of video request from the viewer to the Operations Manager. The viewer is responsible for contacting the proper Media Server to receive video streams.

*Figure 4-16 Operations Manager Traffic Flows*



When the OM Viewer requests a video stream, the following steps occur as shown in Figure 4-16:

**Step 1** The user accesses the Operations Manager screen through an ActiveX web browser. This traffic can be over TCP port 80 (HTTP) or 443 (HTTPS).

**Step 2** The OM Viewer receives a complete list of resources, such as camera feeds, views, and monitors. This information is sent each time the client starts or switches to the operator view. Since the OM Viewer has a complete list of resources, the operator may choose to view live or recorded video from any camera feed or predefined views.

**Step 3** The OM Viewers selects a video feed that is served by the Media Server and contacts the Media Server directly over TCP port 80.

**Step 4** The Media Server is the direct proxy for the IP camera and requests the video stream from the camera. This communication can be TCP, UDP, or multicast as configured by the Operations Manager.

**Step 5** The camera provides the video stream to the Media Server.

**Step 6** The Media Server replicates the requested video feed to the OM Viewer using IP unicast over TCP port 80. The connection remains active until the OM Viewer selects a different video feed.

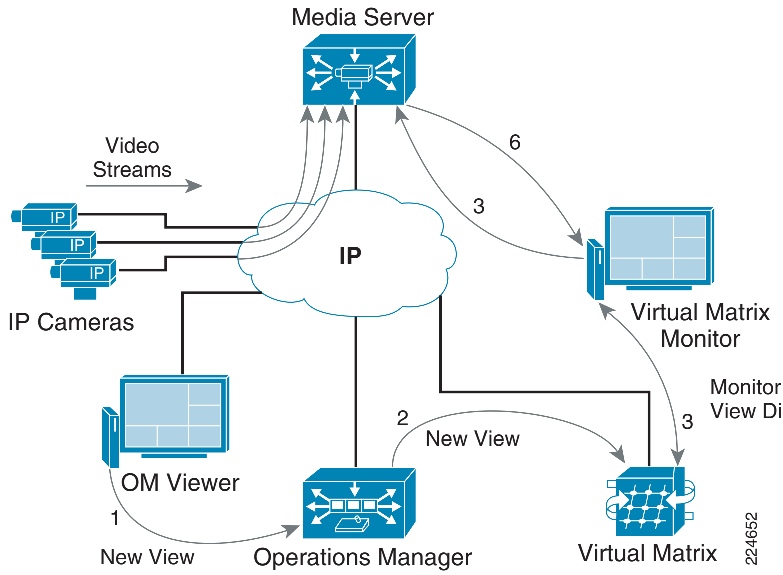
If another OM Viewer requests the video from the same IP Camera, the Media Server simply replicates the video stream as requested, and no additional requests are made to the camera.

# Video Surveillance Virtual Matrix Switch

The Virtual Matrix server is responsible for providing detailed monitor layout to the Virtual Matrix monitors and the position of each camera feed on the screen. A single Virtual Matrix server can be deployed to support a large number of Virtual Matrix monitors since the communication between the monitors and the server is required only during the initialization or when a new view is pushed to the monitor.

Once the monitor layout and views are pushed to the monitors, the monitors are responsible for contacting the appropriate Media Server(s) to request video streams.

*Figure 4-17 Virtual Matrix*



When requesting a new view for the Virtual Matrix monitor, the following steps occur as shown in Figure 4-17.

**Step 1** The OM Viewer selects a new view to be displayed by the Virtual Matrix monitor. The request is received by the Operations Manager.

**Step 2** The Operations Manager sends the layout update to the Virtual Matrix server.

**Step 3** The Virtual Matrix server pushes the new layout to the Virtual Matrix monitor.

**Step 4** Once the Virtual Matrix monitor learns the new layout and the cameras to display, it contacts the appropriate Media Servers to request video streams.

**Step 5** Video streams are sent from the Media Server directly to the Virtual Matrix monitor.

**Step 6** The Virtual Matrix server sends a keepalive message to the Virtual Matrix monitor every three minutes to confirm that the display is still active.

# Bandwidth Requirements

Compared to VoIP, video consumes considerably more network bandwidth. In the LAN environment, bandwidth is relatively inexpensive and in most cases, a LAN infrastructure that is supporting VoIP and data can also support IP video surveillance. In the "Video Traffic Flows" section on page 4-23, the sources and sinks of video data were examined. In this section, some bandwidth estimates are shown as well as tools to calculate bandwidth requirements. The two legs of interest are from the cameras to the Media Server and from the Media Server to the viewing station. The bandwidth from the control plane is trivial compared to the bandwidth consumed by the media streams. For capacity planning purposes the control plane traffic is of little significance; however, from a QoS perspective it is must be accurately marked and queued to prevent the drop of this traffic.

## Camera Feed to Media Server

The bandwidth consumption from individual IP cameras to their Media Server is going to first be determined if the camera has an active archive or operator viewing a live feed. If a camera is not being actively viewed or an archive is not running, no video output is sent from the camera.

**Tip** There is one exception. If the camera has been configured by the web interface to enable IP multicast and the configuration is complete including a multicast address, the camera will stream traffic continuously to the multicast address. The router will not forward the multicast traffic past the LAN segment unless a remote node subscribes to the group (multicast) address.

The output rate from an IP camera is dependent on the configured values for the video feed, including codec (MJPEG, MPEG-4, H.264) resolution, frame rate or bit rate, and any applicable quality factors. These configuration parameters are controlled by the physical security manager and are determined by the operational objective for implementing the camera. As resolution and frame rate increase, so does the bandwidth.

For high-level planning purposes, Table 4-5 can be used for estimating network bandwidth requirements per camera.

*Table 4-5 Per Camera Network Bandwidth Estimates*

| Camera | CODEC | Resolution | Estimated Bitrate |
|---|---|---|---|
| CIVS-IPC-2500 (Standard Definition) | MPEG-4 | D1 (720x480) | 1-2 Mbps |
| CIVS-IPC-4300 or CIVS-IPC-4500 (High Definition) | H.264 | HD (1920x1080) | 4-6 Mbps |

Both camera series can operate at higher than the estimated bitrates shown above, however these bitrates should provide acceptable video quality. Consult the appropriate camera user's guide for the maximum bitrate and other parameters. An important part of the planning process is a pilot installation of IP cameras to gauge their suitability for the intended deployment. From this pilot, more specific data points can be determined.

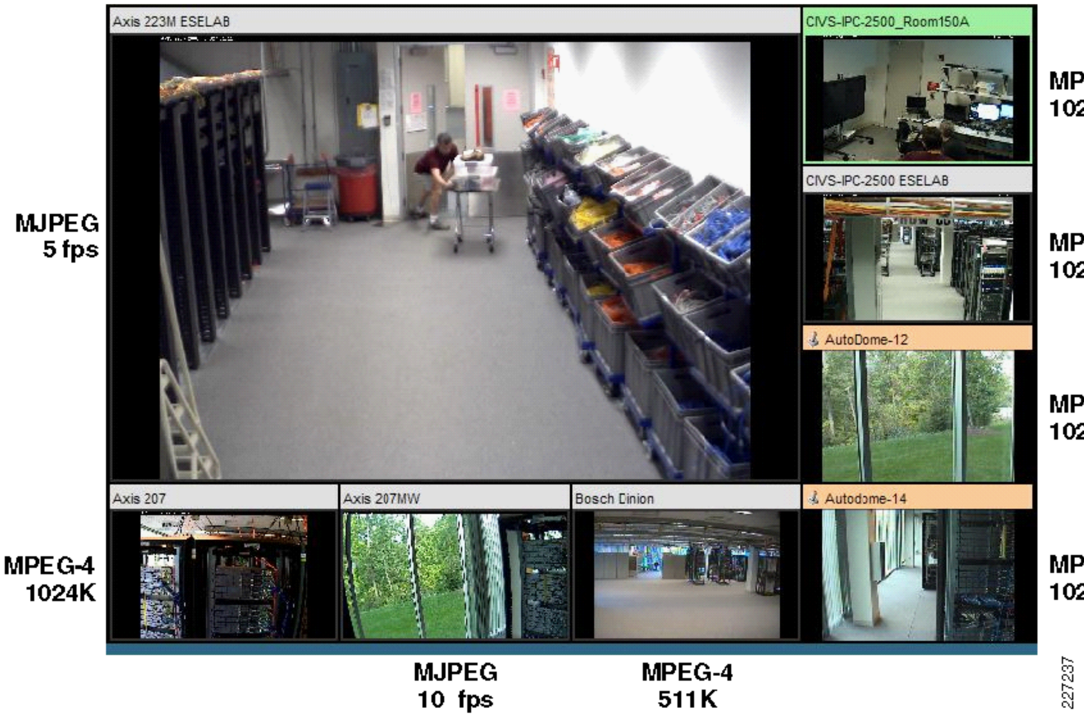
One technique for determining the amount of bandwidth required for a particular model of camera is to use a laptop and connect to the API of the camera. View the video feed directly from the camera using the CODEC, resolution and frame/bit rate of interest. In Microsoft Windows, the task manager (CTL + ALT + DEL and select Task manager from the dialog box displayed) can be used to view the image in real-time. Under the Networking tab, the network utilization and link speed of the interface can be used to estimate the bandwidth requirements of the video feed.

The "Access-layer Switch Commands" section on page A-3 demonstrates how to determine the output data rate from an IP camera when connected to a switch. It is useful for both troubleshooting and for bandwidth estimates.

## Media Server to Viewing Station

In order to view live or archived video feeds from a viewing station, the user first connects with the webserver of the Operations Manager (VSOM) while the camera feeds are sent from the appropriate Media Server (VSMS). To understand the data flow a sample operator screen is shown Figure 4-18.

*Figure 4-18 Viewing Station—Operator Panel*



The codec and bit rate/frame per second parameters are shown next to the respective video image on the operator panel in Figure 4-18. There are two cameras that have a resolution of over 1 megapixel while the remainder are standard definition cameras at D1 or 4CIF resolution.

On the operator view shown, one video image is the predominate image on the screen, covering a larger area than the seven other camera feeds. The feeds are resized on the client-viewing station, the data rate from Media Server to viewing station is not changed or adjusted by the media server based on the resolution on the screen.

Also, each video feed, regardless if the feed is Motion JPEG or MPEG-4, is transported in an individual TCP (WWW) session between viewing station and the respective Media Server. To understand the flows between the Media Server and the viewing station, a NetFlow export is captured and summarized and represented in Table 4-6.

*Table 4-6 Summarized NetFlow Export of Camera Feeds*

| Source IP Address | Src port | Destination IP Address | dst port | Packets per Second (pps) | Average Bytes per Packet | K bits per second (kbps) |
|---|---|---|---|---|---|---|
| 192.0.2.2 | WWW | 192.168.16.36 | 24638 | 97 | 1392 | 1,058 |
| 192.0.2.2 | WWW | 192.168.16.36 | 24639 | 97 | 1390 | 1,058 |
| 192.0.2.2 | WWW | 192.168.16.36 | 24649 | 364 | 1324 | 3,769 |
| 192.0.2.2 | WWW | 192.168.16.36 | 24657 | 53 | 1413 | 585 |
| 192.0.2.2 | WWW | 192.168.16.36 | 24661 | 191 | 1388 | 2,082 |
| 192.0.2.2 | WWW | 192.168.16.36 | 24665 | 49 | 1360 | 530 |
| 192.0.2.2 | WWW | 192.168.16.36 | 24668 | 55 | 1338 | 585 |

| Source IP Address | Src port | Destination IP Address | dst port | Packets per Second (pps) | Average Bytes per Packet | K bits per second (kbps) |
|---|---|---|---|---|---|---|
| 192.0.2.2 | WWW | 192.168.16.36 | 24671 | 105 | 1310 | 1,081 |
| | | | | | | 10,748 |

The source IP address of 192.0.2.2 is the address of the Media Server. The client PC IP address is 192.168.16.36. Each line in Table 4-6 represents the flow from one of the camera feeds. The flows from VSOM to the camera is not shown.

In this example, the aggregate bit rate per second is 10,748 Kbit, or over 10 Mbps. Viewing these eight feeds over a broadband or T1/E1 WAN link would not be practical. The number of concurrent video feeds would need to be limited and a reduction in the frame rate (MJPEG) or bit rate (MPEG-4) of the individual feeds to view a panel of this complexity over the WAN.

Viewing camera feeds over the WAN is not impossible, but some consideration must be given to the aggregate data rate when viewing more than one feed or individual feeds with very high resolution and bitrate or frame rate.

# Video Storage

The video surveillance storage system provides multiple options to store video and audio files. The internal storage of the Media Server may be augmented by using direct attached or SAN storage. The video surveillance storage system can store video in loops, one-time archives, or event clips triggered by alarm systems providing for redundant and remote long-term archival.

## Calculating Storage Requirements

### MJPEG

When using MJPEG streams, the frame size of each image plays a key role in estimating the storage and transmission requirements. Since each frame is unique and varies according to the image complexity, it is difficult to provide a guide that provides fixed frame sizes. An IP camera that provides images with low complexity will generate smaller frame sizes. Smaller frames will require less bandwidth and storage capacity.

The following formula is used to calculate the bandwidth requirements for MJPEG streams:

*MJPEG storage = Average Frame size x Frame rate x duration*

**Example 1**: For an 8-hour archive of a CIF video stream with 50 percent quality and 15 frames per second, the following is the calculation:

```
4 KB  x  15fps  x  3600s  =  216,000 KB/ hour
                          = 216MB /hour   x   8 hours
                          = 1.728 GB
```

**Example 2**: For a 24-hour archive of a 4CIF video stream with 100 percent quality and 5 frames per second, the following is the calculation:

```
320 KB x 5fps x 3600s  =  5,760,000 KB /hour
                       =  5,760MB /hour  =  5.76GB /hour x 24 hours
                       =  138.24 GB
```

### MPEG-4/H.264

Rather than standalone images, MPEG-4 / H.264 streams take into account video frames and the size of a given video frame varies widely between I-frames and predictive frames. Typically, H.264 is more efficient than MPEG-4. MPEG-4 is generally more efficient than Motion JPEG and requires less bandwidth and storage capacity when using higher frame rates.

The following formula is used to calculate the bandwidth requirements for MPEG-4 streams:

*MPEG4 storage = Bit rate (kbps) x duration*

The target bit rate is configured on the camera and is already expressed in bits per second.

**Example 1:** For an 8-hour video stream with target bit rate of 768kbps, the following is the calculation:

```
768kbps / 8 bits/s = 96 KB /second  x  3600 s
                   = 345,600 KB/hour  /  1000
                   = 345.6 MB/hour  x  8 hours
                   = 2.764 GB
```

## IP Camera Video Settings

When creating a new IP camera, several settings play a role in providing the appropriate image quality and frame rate. When using MJPEG video streams, the following image settings may be configured: *resolution*, *frame rate*, and *quality*. The frame rate setting determines the amount of video data generated at a given amount of time.

For MPEG-4 and H.264 videos streams, the *resolution, bit rate, and quality* may be configured. The bit rate setting specifies the amount of bandwidth required for the MPEG-4 / H.264 video stream. Higher values generate more video data every second, translating into smoother video and a more accurate representation of the field of view. A higher value also translates into larger archive file sizes.

# Design CheckList

This design checklist in Table 4-7 facilitates pre-implementation planning and the decision process.

*Table 4-7 Design Checklist*

| Design Checklist |
|---|
| Estimate the number of IP cameras required at each location. |
| Using a floor plan or exterior survey, determine cameras that can be powered by PoE and those requiring power supplies. |
| Survey existing IP or analog cameras and determine if these cameras are to be replace or migrated. |
| Estimate the CODEC, resolution, and frame rate or bit rate requirements for cameras at each location |
| Determine the retention period requirements for cameras at each location |
| Survey existing LAN switches for necessary features and available capacity |
| Based on the number of cameras per location, determine server requirements. |
| Using the *Campus Implementation Case Study* in the following section, determine what if any LAN infrastructure upgrades are required. |
| Using the estimate on the number of servers required, calculate the storage requirements for video archives based on the retention period analysis |

| Design Checklist |
|---|
| Analyze the IP addressing requirements and VLAN assignments for IP Cameras, Media Servers, routers, switches and other systems. |
| Determine if suitable Network Time (NTP) sources exist in the current network. |
| Investigate what network management servers and software are currently available for services such as Syslog and SNMP traps, TFTP/FTP for firmware download and storage. |
| Consider implementing network management servers for performance, fault and capacity planning such as *CiscoWorks Internetwork Performance Monitor (*end-to-end network performance) , *Cisco Secure Access Control Server for Windows (*authentication: TACACS+/RADIUS server*), CiscoWorks Device Fault Manager* (reporting of device faults) and *Cisco NetFlow Collector (*NetFlow analysis for capacity planning). |
| Analyze the existing QoS policies and configuration on routers and switches and incorporate the IP video surveillance requirements into the plan. |
| Determine requirements for external users to access video feeds. Analyze what level of encryption or access-control is required to meet the end-user requirements and to align with the corporate network security posture. |
| Discuss with the physical security manager and network manager the need for segmentation features such as VRF-lite, VLANS and firewalls and access-lists to limit access to end-nodes. |
| Determine the redundancy inherent in the existing network and develop a plan for meeting the physical security needs in the event of a line-card or access switch failure. |
| Consult with the physical security manager to determine what the live viewing requirements are. Determine what cameras must be viewed live and were the viewing stations are located in the network topology. |
| Determine skill set of existing staff and estimate training requirements for physical security installers, operators and managers in basic internetworking. Consider involving the network staff in day-to-day operations of the physical security operations staff. |

# Campus Implementation Case Study

This section addresses the topology considerations for deploying IP video surveillance at a campus location where a high-density of high-definition (HD) IP cameras are required. This case study is modeled after an actual customer request for design assistance in the gaming industry. Because of the high number of HD IP cameras in a dense deployment, hundreds of cameras within the maximum distance of 100 meters for Ethernet, it may be practical to implement a physically isolated LAN infrastructure for transporting, viewing, and archiving video feeds.

In other industry deployments, such as a public school system, a converged design using virtualization, path isolation, and encryption described later in this document is more applicable. In the public school system example, there may be less than 1,000 cameras deployed in 10 to 20 campus locations or 50 to 100 cameras per campus. In this case, using logical segmentation techniques, QoS on the LAN and WAN, and a converged network of voice, video, and data is the most cost-effective solution. However, the exercise of calculating the offered load from a set of cameras on a single access switch, through the distribution layer and then to the core, provides baseline information that can be leveraged elseware.

There is one key traffic engineering concept that is relevant to any IP video surveillance deployment. While there is a two-way communication between the camera and media server, the vast majority of IP traffic when using either Motion JPEG or MPEG-4/H.264 is sourced from the camera and linked to the Media Server. If large numbers of video feeds are being actively viewed, the viewing stations will likely be attached to the the interface cards in the core layer and not traversing the distribution, access, and core campus layer. Also, in many cases, the majority of video feeds are never viewed, in some cases up to 99 percent of the population of collected data.

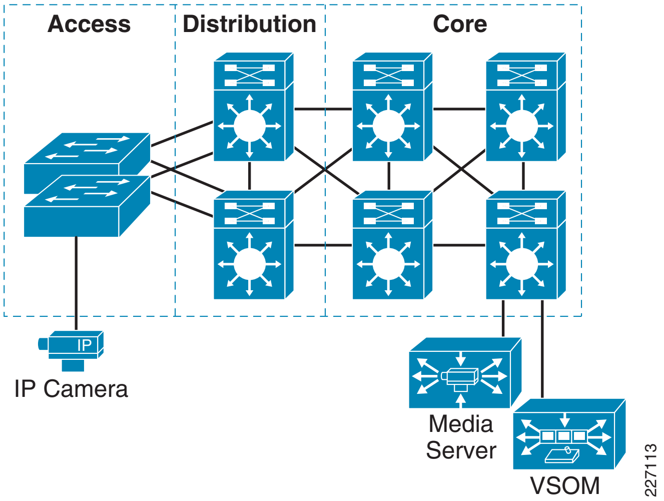
## Requirements

This case study assumes the requirement is to populate every available port on the access switch with with an HD camera. If there are PC/workstation, point-of-sale terminals, printers, or IP phones, these devices are isolated on a separate access switch. The goal is to look at the worst case deployment scenario where all cameras are generating HD video 24 hours per day.

The basic assumption is that the camera is a Cisco 4000 Series with a resolution of 1080p (1920 x 1080) using the H.264 CODEC with a configured target bit-rate of 5Mbps. This model of camera can be configured for a constant bit-rate in increments of 2, 4, 6, 8, 10, 12, and 15 Mbps. In viewing live video streams, a configured rate of 4Mbps provides generally acceptable video quality. The assumption is the cameras are configured for 4Mbps. In the calculations, 5Mbps per-camera is used to accommodate any bursts by the camera, providing for a conservative estimate of bandwidth.

This campus design is a traditional three-layer design where access switches are aggregated into distribution layer switches. The distribution layer switches are connected to the core switches. The core switches provide connectivity for the media servers. The general topology is illustrated in Figure 4-19.

*Figure 4-19 Traditional Three Layer Campus Design*



The basic assumption is that each IP camera uses Power-over-Ethernet (PoE) and attaches to the access switch at 100Mbps. The multiservice platforms (and most server platforms ) that are used for Media Servers/Video Surveillance Operations Manager (VSOM) ) have two 10/100/1000M Ethernet interfaces. Assume the servers are connected at 1000Mbps (1Gbps) for receiving live video feeds.

**Note** VSOM runs on one or more multiservice platforms, the Media Servers run on as many instances of the multiservice platforms as required to support the number of cameras and storage requirements for video archiving.

## Access Layer

One of the advantages of implementing IP video surveillance is the ability to supply electrical power to the camera using the IEEE 802.1af standard for PoE. Because of this, only PoE-capable access layer switches are considered.

Another factor is the capability of the switch to provide Layer-2 only uplinks or either Layer-2 or Layer-3 uplinks. In this design, only Layer-2/Layer-3 switches are considered because they eliminate the need to have a Layer-2 interface connecting the distribution layer switches. Additionally, a routed-access layer improves convergence times and enhances load sharing. For these and other reasons, many campus network deployments advocate using a Layer-3 connection to the access layer switch. For more information, refer to the *High Availability Campus Network Design—Routed Access Layer using EIGRP or OSPF* at the following URL: www.cisco.com/go/designzone.

For these reasons, access switches that do not support PoE and Layer-3 routing support are not considered in this campus case study.

The other produce selection considerations for access-layer switches are as follows:

- Number of ports (either 24 or 48 port models are available)
- Total switch capacity; for example, backplane capacity (32Gbps or 64 Gbps models are available)
- Uplink bandwidth available (either 1Gbps or 10Gbps is available)

Deciding which switch to deploy based on these three factors now is a matter of examining the offered load from the Cisco 4000 Series camera. The assumption is 5Mbps as a target constant bit-rate. Therefore a 24-port switch at 5Mbps per port, is 120 Mbps of offered load and a 48 port switch offers 240 Mbps. Based on this value, a switch with 48 ports, 32Gbps backplane, and 1Gbps uplinks will service these requirements. One switch that meets the requirement is the Cisco Catalyst 3560G-48PS: 48 Eth 10/100/1000 ports with PoE and 4 SFP GigE ports.

Table 4-8 represents other switches in the access switch product line. The **bolded** items meet or exceed the stated requirements.

*Table 4-8 Access Switch Features*

| Model<br>Catalyst | L2/L3 Multilayer Switch | Total Switching Capacity (Gbps) | Uplinks | Ports |
|---|---|---|---|---|
| 2975GS-48PS-L | NO | 32 | 4 SFP 1Gbps | 48 |
| 3560G-24PS<br>**3560G-48PS** | YES | 32 | 4 SFP GigE | 24<br>48 |
| 3750G-24PS<br>**3750G-48PS** | YES | 32 | 4 SFP GigE | 24<br>48 |
| 3560E-24PD<br>**3560E-48PD** | YES | 64 | 2 X2 10 GigE uplinks | 24<br>48 |
| 3750E-24PD<br>**3750E-48PD** | YES | 64 | 2 X2 10 GigE uplinks | 24<br>48 |

# Access Layer Offered Load

Given the criteria of 48 HD cameras streaming video feeds at 5mbps, the offered load from all 48 ports to the uplink ports is 48 * 5Mbps or 240 Mbps. Even if the cameras are configured at a constant bit-rate (CBF) value of 16Mbps, the total offered load to the uplink is 768 Mbps.

**Tip** A 48-port access layer switch fully populated with HD IP cameras uses 25% - 80% of a 1Gbps uplink.

A best practice for interface-level QoS, the priority queue should not exceed a third of the link bandwidth in a converged network of voice, video, and data. In this topology, using a target rate of 5Mbps per camera, less than a third of the available uplink bandwidth is a candidate for the priority queue. We can therefore, safely assume that replacing the cameras on the switch ports with IP phones and workstations will also not deviate from this guideline, as the VoIP traffic load is substantially less than the HD video stream.

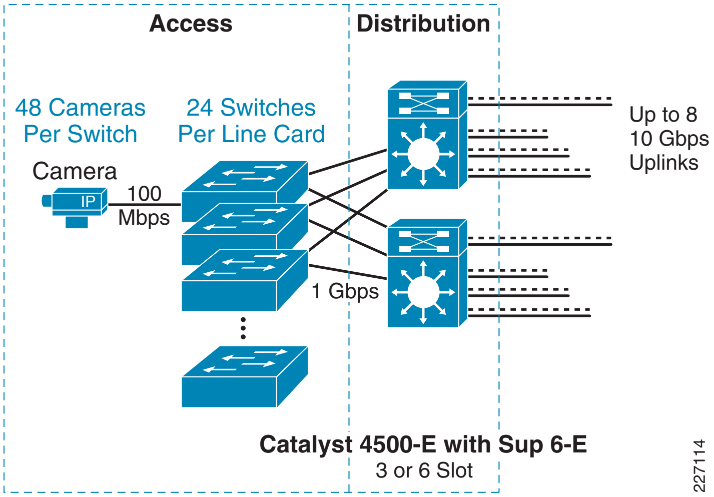
When planning for the distribution layer, assumed that each access layer switch is generating 240Mbps on one of the two uplinks, or the traffic is load sharing across the two uplinks, but does not exceed 240Mbps from a single access switch.

# Distribution Layer

The distribution layer topology design incorporates deploying two or more chassis. Each chassis is outfitted with a single supervisor module and optionally dual-power supplies. Because IP routing extends to the access layer, there is no need to implement a Layer-2 link between the two distribution layer switches. The access layer switches receive routing updates from each distribution layer switches and will select an alternate path to the core if a path fails.

One recommended solution for the distribution layer switch selection is the Cisco Catalyst 4500-E in either a 3 or 6 slot chassis. With the 3 slot chassis, the Supervisor Engine 6-E and an additional uplink link card uses two slots, while the remaining contains one 24-port line card for aggregating 24 access switches. The 6 slot chassis can houses four 24-port line cards aggregating 96 access switches. By including the uplink line card, up to eight 10Gbps uplinks are available to the core layer. The access layer and distribution layer is shown in Figure 4-20.

*Figure 4-20 Distribution Layer*



Using a Supervisor Engine 6-E and a 3 or 6 slot chassis provides for 24Gbps of per slot aggregate bandwidth. The configuration summary is as follows:

- Catalyst 4500-E with Sup 6-E (includes 2 10 Gig Uplinks)—Either 3 or 6 slot chassis

- WS-X4606-X2-E 6-Port 10 Gigabit Ethernet (X2)—Uplinks to core switches
- WS-X4624-SFP (24 Port GigE)—Downlinks to access switches

When using the three slot chassis, 24 access switches can be aggregated for a total of 1152 cameras supported on two distribution layer switches with redundancy. When using the six slot chassis, up to 96 access switches can be attached to both chassis in a redundant configuration, supporting up to 4608 cameras.

## Distribution Layer Offered Load

Assuming in a failure scenario, where one distribution switch is out of service, the offered load from the remaining switch into the core switches is in the range of either:

- 1152 cameras (5 Mbps each) approximately 6 Gbps
- 4608 cameras (5 Mbps each) approximately 24 Gbps

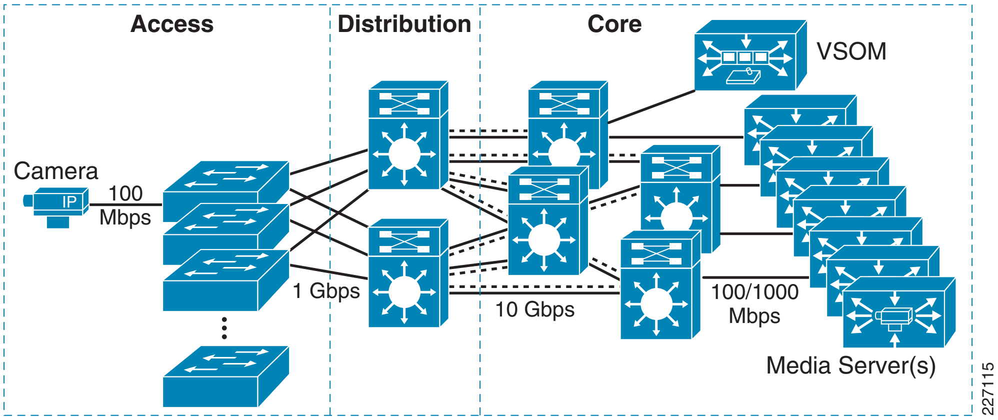
Given the above, a three slot chassis deployment would not require the installation of a WS-X4606-X2-E line card for additional uplink capacity, the two 10 Gbps uplinks on the Sup 6-E is sufficient for the offered load of 6 Gbps. The core switches, supporting the servers, are discussed next.

## Core Layer

The core layer provides LAN connectivity for the Media Servers, VSOM servers, and other network management servers. The number of Media Servers required is a function of the number of cameras which each Media Server is able to support. For planning purposes, it is assumed in this case study that each server can support 16 camera feeds. Depending on the hardware of the Media Server, the number of supported cameras may be higher or lower and the network manager must adjust the number and type of core layer switches accordingly. The general guideline for the Cisco Physical Security Multi Services platforms is no more than an aggregate of 200 Mbps of I/O per chassis.

Given a projection of 16 camera per Media Server, the low end deployment using 1152 cameras requires 72 Media Servers and the upper end deployment of 4608 cameras requires 288 Media Servers. The network topology diagram is expanded to include a representation of the core layer (see Figure 4-21).

***Figure 4-21 Core Layer***



There are several options for core layer switches. As was deployed in the distribution layer, the Cisco Catalyst 4500-E with Sup 6-E could be used. A Cisco Catalyst 6500 Series or a Cisco Nexus 7000 Series are also options. These latter two options are ideal if the access and distribution layer deployment is going to be tied into an existing enterprise network infrastructure. In that case, it is assumed that 10Gigabit line cards and line cards to support Media Servers on 10/100/1000 RJ45 ports are available or could be added to existing chassis in the core.

For the purpose of this case study, assume dedicated Cisco Catalyst 4500-E with Sup 6-E are also used for the core layer switches. The core layer switches then include the following:

- 4500-E with Sup 6-E (includes two 10Gigabit uplinks)
- WS-X4606-X2-E—6-port 10Gigabit Ethernet (X2)
- WS-X4424-GB-RJ45—24-port 10/100/1000 Module (RJ-45)

The type of chassis can be a 6, 7, or 10 slot. Because these are core layer switches, dual-power supplies are recommended as are dual supervisors. The 7 and 10 slot chassis support dual supervisor cards.

The number of core switches required depends on how many cameras, and therefore, how many Media Servers, are implemented. It is assumed the Media Servers are are equally distributed across the available ports on the core layer switches. The access layer switches are routing peers with the core layer switches and are equal cost load sharing to the core.

If the upper projection of 4,608 cameras are deployed, at 16 cameras per Media Server, then 288 media servers are required. Assuming four core switches, 72 servers on each, at a minimum each chassis must support 3, WS-X4424-GB-RJ45 24-port line cards. Either the 6 or 7 slot chassis meets this requirement.

If the lower projection of 1,152 cameras are deployed, then 72 total servers are required. Two core switches, with each switch supporting 36 servers, are necessary. Again, a 6 or 7 slot chassis meets this requirement.

The number of VSOM servers required for this installation depends if the implementation is headless (meaning little continuous viewing of live feeds) or if there are continuous viewing of many or all live feeds. It is recommended that a single VSOM manage any given Media Server. There is no absolute rule in the number of Media Servers per VSOM server. One guideline is to use 20 Media Servers per VSOM server. Allocating Media Servers to the respective VSOM server should follow some allocation scheme such as a geographical division. In other words, if there are three adjoining buildings with 100 cameras in each building, those cameras may be all controlled by 20 Media Servers and a single VSOM.

## Core Layer Offered Load

Assuming the use of a Catalyst 4500-E with Sup 6-E, each line card slot has the capacity to support up to 24Gbps of traffic. Given the Media Servers are attached on 10/100/1000 Mbps RJ45 ports on a WS-X4424-GB-RJ45 line card, there are up to 24 Media Servers per slot. At 5 Mbps from each camera, with 16 cameras streaming video to each Media Server, that is a sustained offered rate of 80 Mbps to each Media Server. Assuming the goal is to keep the aggregate below 200 Mbps of I/O per chassis, there is available capacity for both a higher bit rate from some cameras and capacity to retrieve and view live or archived feeds stored on any one Media Server. The aggregate bandwidth per slot is approximately 2Gbps, well under the 24 Gbps rate capacity.

**Tip** Verify the Media Servers are successfully auto-negotiating the 1000Mbps data rate. There is an expected sustained offered rate of 80Mbps to each Media Server.

The aggregate bandwidth from distribution layer to any one of the four core switches (given 4,608 cameras) is approximately 6Gbps, or 24Gbps in aggregate from both distribution layer switches.

## Summary

In this case study, we examined what characteristics an access-layer switch needed to support a deployment of high definition IP video surveillance cameras. The bandwidth requirement between the access and distribution layer was examined. Also, a distribution layer switch infrastructure was suggested which would aggregate a large number of access switches densely populated with cameras. Redundancy from the access layer, through the distribution layer to the core is enabled by choosing switches which support end-to-end Layer-3 routing. Finally, the core layer bandwidth and port requirements were examined.