Technical Report

Video Surveillance Solutions with NetApp E-Series Storage

Sizing Considerations

James Laing, Frank Poole, Marco Pozzoni, Ted Hayduk, NetApp

December 2017 | TR-4199

Abstract

Video surveillance solutions using E-Series storage offer the physical security integrator a highly scalable repository for video management systems supporting high camera counts, megapixel resolutions, high frame rates, and long retention periods. The architecture is designed to provide high reliability and availability to meet the demands of video surveillance deployments.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

1 Introduction

NetApp® E-Series storage arrays provide performance, efficiency, reliability, and enterprise-class support for large-scale video surveillance deployments.

All video surveillance management software shares the common feature of recording live video feeds to storage for subsequent replay to aid in forensic analysis or investigation of persons or events within the field of view of a single camera or group of cameras. These video feeds, generated by hundreds or thousands of cameras, are typically configured to record continuously, 24 hours per day, 7 days per week, with retention periods in the range of months to years.

1.1 Publication Scope

This document is intended to provide a set of guidelines for sizing a video surveillance solution based on NetApp E-Series storage arrays.

1.2 Audience

This publication is intended to provide guidance to physical security integrators, video surveillance management software engineers, network and storage system engineers, and architects responsible for integrating NetApp E-Series storage systems into existing video surveillance deployments or designing and implementing new deployments.

The content in this report is presented with the expectation that these professionals can use this information, combined with their experience and supporting documents, to build an efficient, scalable, and highly available system.

Targeted Deployments

The targeted deployments for this introduction are large (200–2,000 cameras or more) with retention periods of at least 30 days and primarily use HDTV/megapixel resolution cameras.

1.3 Training Offerings

There are a number of web-based and instructor-led training opportunities to enable successful deployment of the NetApp E-Series storage array. The classes listed in the NetApp University Customer Learning Map under Storage Systems are recommended end-user training classes.

Table 1) Training offerings.

Class Description	Duration (Hours:Minutes)	Delivery
Product Foundation		
Technical Overview of NetApp E-Series Storage Systems	1:00	Web-based
Technical Overview of NetApp EF-Series All-Flash Array	0:30	Web-based
Hands-On Skills Development		
Configuring and Monitoring E-Series and EF-Series Storage Systems	32:00	Instructor-led
Certification		
NetApp Certified Implementation Engineer: SAN Specialist, E-Series	1:30	Exam

2 Sizing Considerations

Video surveillance solutions based on NetApp E-Series provide performance, efficiency, and reliability with enterprise-class support for large-scale video surveillance deployments. The solutions can utilize the NetApp E-Series storage array in E5760, E2812, and E2860 configurations. This section addresses sizing guidance for all deployment models. A variety of video surveillance resources are available on the Field Portal.

2.1 System Requirements

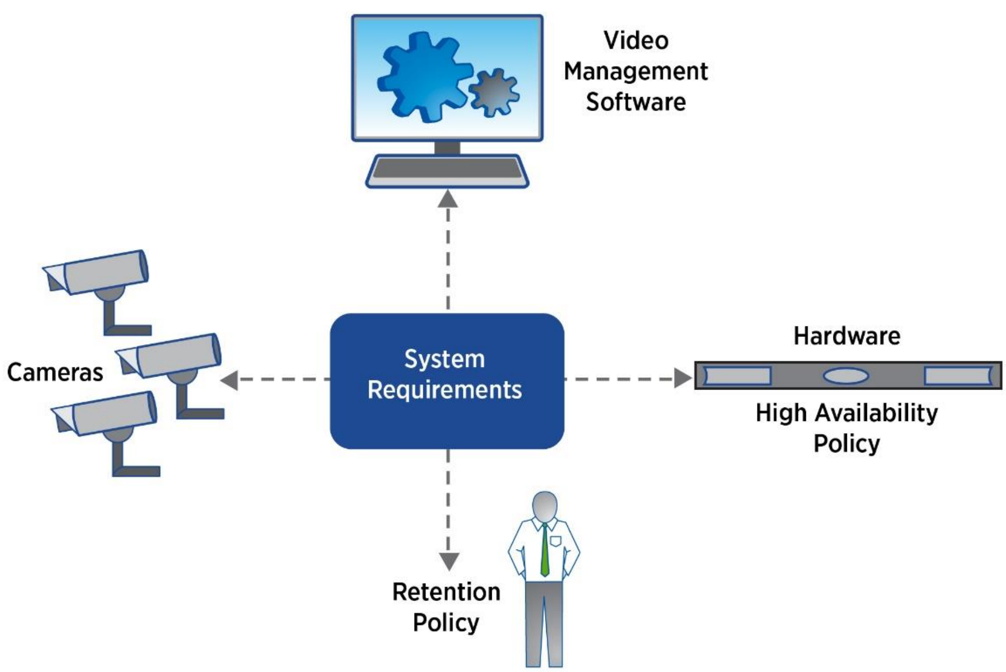
The system requirements are specified in a request for proposal/quote developed by either the end customer or a physical security consultant under contract with the end customer.

The physical security integrator must work with the physical security manager to accurately assess specific requirements, including:

- Retention period
- Number, location, and type of cameras, resolution, frame rate, and so on
- Video management software selected
- Number of cameras per recording server
- Continuous recording or record on motion
- Frequency of viewing archived video
- Failover design requirements

These requirements have dependencies that affect the total system and are illustrated in Figure 1.

Figure 1) System requirements.



The individual requirements must be discovered, analyzed, and documented to accurately size the storage array. The following sections address each of the components, provide technical background about the options, and make recommendations about industry best practices.

2.2 General Considerations

The process of sizing storage to address the video management application requirements is an exercise that balances throughput and capacity considerations. In video surveillance deployments that are characterized by minimal forensic analysis of archived video, capacity is the primary consideration. Recording servers configured to continuously record video exhibit a relatively deterministic I/O pattern. The arrival rate of video feeds from IP cameras is constant, and the video recording server in turn writes these streams to an archive at a consistent rate. Many video management systems write the video archive to a temporary or recording directory on the volume logical unit number (LUN) and subsequently read and write the temporary video archive to a permanent directory. Although this movement of video data files from one location to another adds a secondary I/O to the initial write, the workload characteristics can be quantified.

Markets that fall into this category are secondary and higher education, enterprise, and commercial or retail deployments. Video is only retrieved and analyzed if an incident requires investigation. For example, the education market might only review video from a few cameras several times a week. The I/O is mainly writes, and there is little viewing (reads) of the archived video.

The gaming market is an example where there is a high degree of forensic analysis of video archives. In these deployments, several security operators are frequently reviewing video archives to investigate suspicious activity, theft, and fraud. This activity adds a high degree of read I/O to the workload, which is exaggerated by the use of fast forward capabilities of the video management software (VMS). These I/O patterns are less deterministic than the constant influx of video feeds from camera to the server because the workload is a function of the number of operators conducting investigations, the number of camera archives being viewed, and the playback speed.

Other vendors' storage might require the use of two volumes (LUNs) per recording server. One volume is often configured for RAID 10 for the recording directory, and a second volume for storing the video archive to the target retention period is configured as RAID 5 or RAID 6. RAID 10 is used for the recording directory due to better assumed or observed write performance than RAID 5 or RAID 6 on the storage array.

While this configuration can be used with NetApp E-Series storage, the preferred configuration is to use a single Dynamic Disk Pool (DDP). DDP, optimally configured with 30 drives per pool, has been proven to satisfy high rates of both read and write activity, is easier to configure and manage, and recovery from drive failure occurs much more quickly than with traditional RAID implementations. See TR-4197: Video Surveillance Solutions with NetApp E-Series Storage: Planning and Design Considerations for more information about DDP.

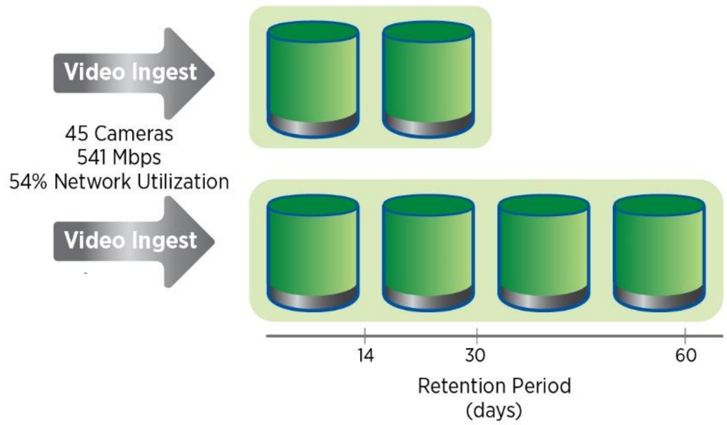
2.3 Retention Period

The retention period is usually determined by organizational or regulatory policy. For example, the Nevada Gaming Commission Regulation 5.160(2), "Surveillance Standards for Non-Restricted Licensees," specifies a minimum retention period of seven days. However, cameras deployed in nonregulated areas such as lobbies, parking lots, and other common areas might have a 30-day retention period governed by hotel policy.

State and local governments typically specify video retention periods as part of their state records retention schedules. Geographies with a history of criminal activity or terrorist threats might specify longer retention period requirements.

Increasing the retention period does not increase the arrival data rate of video feeds from IP cameras to the recording server and storage array. Longer retention policies make the sizing exercise more of a capacity calculation than a performance consideration. This is illustrated by the example shown in Figure 2.

Figure 2) Example of throughput versus retention.



There is an absolute and a practical limit to the amount of storage allocated to an individual recording server. Additionally, factors such as network interface capacity and CPU, memory, and application performance also limit the number of cameras per recording server. In most video surveillance deployments, the performance characteristics of the E-Series controllers meet or exceed the requirements of the video management application. All video management systems reach a steady state where video archives are deleted at the same rate that new video files are added to the storage array.

NetApp recommends a retention period that meets or exceeds the applicable policy.

2.4 Reserve Capacity

The amount of reserve capacity must be considered when the retention period policy is defined. As a best practice, each volume should be maintained at approximately 80% utilization. Most VMS packages implement a file deletion trigger when the configured maximum size of the archive is reached. Each volume owned by the recording server has a configurable threshold defined as the minimum free space or maximum archive size. Usually this is specified in gigabytes/terabytes in addition to the configured (days) retention period. If either of the two maximum values is reached, that is, the retention period in days or the maximum archive size, the oldest video files are removed from the system to prevent a disk full error condition.

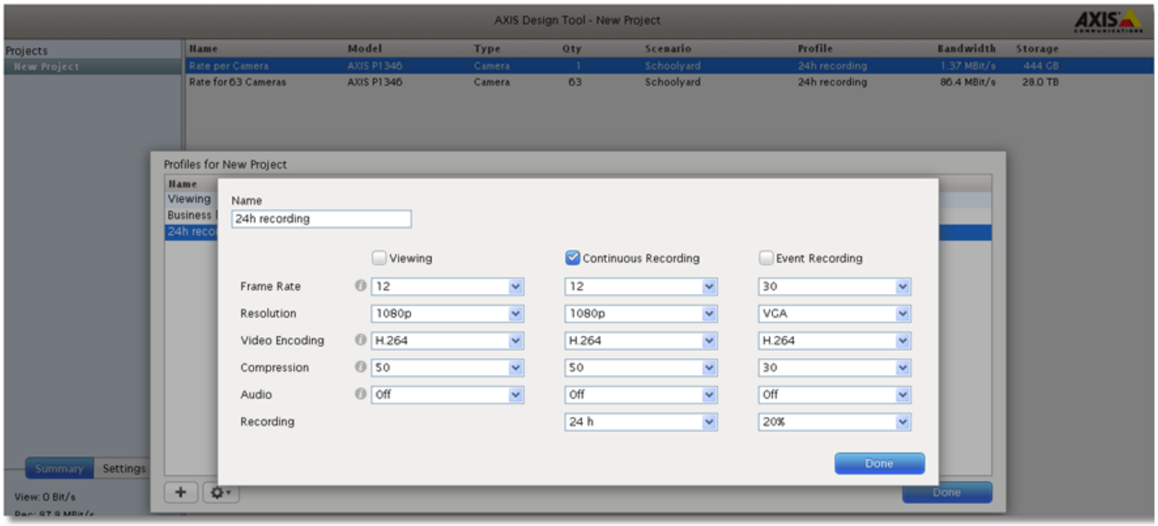
As a best practice, size the system based on a target capacity of 80% based on the stated retention period of the organization. Then, when implementing the system, configure the VMS to use all the configured capacity, maintaining only 5% unused space. This method makes sure that the video retention meets or exceeds the specified policy.

For example, if the target retention period is 30 days, during implementation specify a 38-day retention period with minimum free space at 5% of capacity. This method forces the VMS to delete files based on capacity, utilizing 95% of the volume rather than 80% of the volume.

2.5 Cameras

The number of cameras and the configured resolution, frame rate, codec type, compression factor, and image complexity must be determined to estimate the aggregate video data rate and the amount of storage required. Most camera manufacturers provide a design tool to estimate the data rate and storage requirements based on the camera model and specified values. Figure 3 illustrates an example of one such tool.

Figure 3) Axis design tool.



For more information about the Axis design tool, refer to the [Axis](#) website.

These tools provide only an estimate. The results might vary and should be verified through field trials run by the physical security integrator.

The total number of cameras at a specific location is a function of the physical security requirements of the site and how the security integrator plans to address these requirements. The total population of cameras encompasses a variety of camera models, and in some cases the cameras might be from different manufacturers. The integrator can select from several types of cameras, including indoor or outdoor cameras; fixed or pan, tilt, zoom (PTZ); and tamper-proof or vandal-proof dome cameras. From a sizing perspective, the type of camera is not important, but the resolution and number of channels (video feeds) from each camera are important.

When the total number of cameras is determined, they need to be further categorized by the resolution.

Resolution

There are a wide variety of resolutions available in the video surveillance camera industry. They are distinguished by these categories:

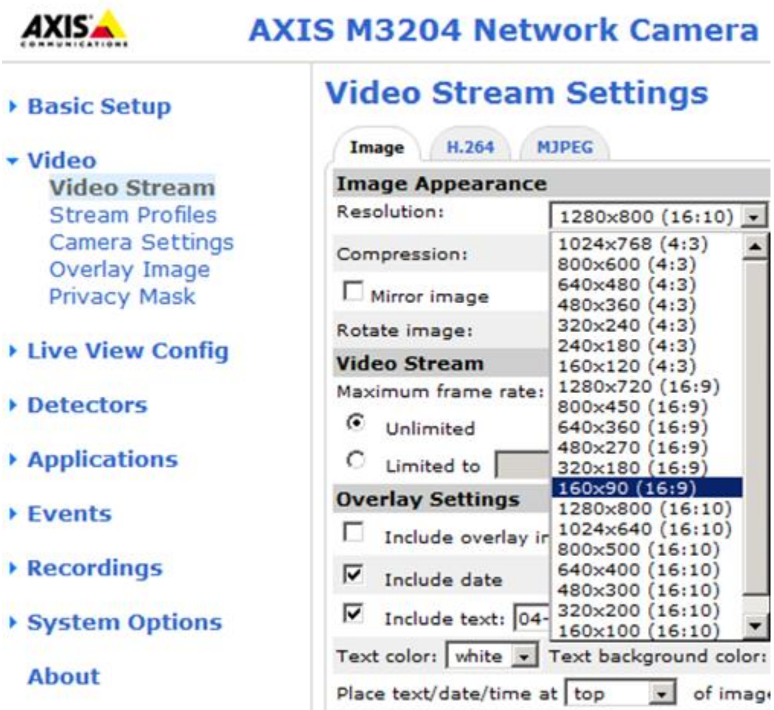
- Analog video (NTSC/PAL); 4CIF is commonly used: 704x480 pixels or 0.4 megapixels
- Video graphics array (VGA) to XVGA: 640x480 pixels to 1024x768 pixels up to 0.75 megapixels
- Megapixel SXGA to QSXGA: 1280x1024 pixels to 2560x2048 pixels or 1.3 to 5.2 megapixels
- High-definition television (HDTV): 1280x720 pixels and 1920x1080 pixels or 0.9 to 2 megapixels

E-Series storage is ideal for HDTV and megapixel resolution deployments because of the high density and performance characteristics of the E-Series. For new installations, HDTV/megapixel cameras are the preferred choice over standard definition (SD) cameras. Although the purchase price of an HDTV/megapixel camera is slightly more than that of an SD camera, the installation cost is the same. The total cost of cameras and installation might be less for HDTV/megapixel deployments over SD cameras, because fewer cameras are required to effectively cover an area.

Video surveillance camera models are selected to meet a functional requirement. These requirements are classified as detection, recognition, and identification. There are industry guidelines for the number of pixels per foot required to address each category. Detection has a lower resolution requirement than identification. One HDTV/megapixel camera might provide sufficient resolution to meet the pixel-per-foot requirements, where earlier multiple SD cameras would need to be deployed.

Network video cameras can be configured at resolutions below the specified maximum resolution. An example of the configurable resolutions for the Axis M3204 network camera is shown in Figure 4.

Figure 4) Video stream settings example.



An HDTV/megapixel camera can also be an SD camera, an HDTV format camera, or a megapixel format camera.

NetApp recommends HDTV/megapixel cameras for new deployments.

Frame Rate

The configured frame rate of the network video cameras also must be determined to accurately estimate the required storage capacity. In the United States, frame rates in the gaming industry are specified by regulations and are required to be 30 frames per second (FPS). Common frame rates for other U.S. industries are 7 to 12 FPS. Different geographies have different FPS requirements.

Cameras positioned at cash registers and teller stations usually require at least 12 to 15 FPS. In school or office hallways, 5 FPS is usually sufficient. Parking lots and other overview scenes for detecting cars, people, or objects often require only 1 to 3 FPS, while license plate capture requires higher frame rates.

Cameras positioned with horizontal movement across the field of view or with high-speed movement (highway intersections, for example) generally require higher frame rates than scenes with vertical movement or slow-moving people or objects. As a reference, motion pictures (35mm film) have traditionally used 24 FPS. Modern digital cinema cameras commonly record at 60 FPS, and modern digital cinema projection is often at 60 FPS. The human eye begins to notice choppy motion below 16 to 18 FPS.

The specified frame rate greatly influences the network bandwidth and storage requirements. In addition to categorizing the cameras by resolution, frame rate is now added to the equation.

NetApp recommends using a frame rate that meets or exceeds the regulatory or functional requirements of the camera placement.

Compression Type

The compression type (compression standard) options for video surveillance deployments are Motion JPEG, MPEG-4/H.264, H.264+, and H.265. Motion JPEG is a series of individual images that have no interdependency between frames. Motion JPEG is often required by analytic software implementations. Because there are no interframe dependencies, Motion JPEG is used in networks that exhibit rates of packet loss that would make MPEG-4 /H.264 unusable. As the frame per second rate increases, the bandwidth and storage requirements for Motion JPEG become increasingly costly compared to MPEG-4/H.264 for a given resolution.

MPEG-4 (MPEG-4 Part 2) and H.264 (MPEG-4 Part 10 AVC) are compression standards that transmit a reference frame periodically and send the changes in the scene in subsequent frames. The frequency of reference frames is configurable by the value specified for group of video/group of pictures (GOV/GOP) length. H.264 is more efficient than MPEG-4, and its use has superseded MPEG-4 for HDTV/megapixel video cameras. H.264+ and H.265 are variations on this theme that can further reduce bandwidth. For instance, H.264+ can cut the data rate in half relative to H.264.

Typically, motion JPEG uses Transmission Control Protocol (TCP) to transport video between a camera and a server. TCP is connection-oriented and provides for reliable transport. Alternately, MPEG-4 and H.264 are usually transported in Real-Time Transport Protocol (RTP)/User Datagram Protocol (UDP). The UDP transport is connectionless and does not retransmit lost packets. TCP favors reliability over timeliness. The majority of voice and video applications use RTP/UDP transport.

The network infrastructure must exhibit a very low packet loss for transport of MPEG-4/H.264 for acceptable image quality. Dropping as little as 1/20 of 1% of the IP packets for an H.264/RTP/UDP video stream is noticed as trailing artifacts in the image. The distortion of the image is corrected when the next reference frame is successfully received.

In video surveillance deployments, NetApp recommends using H.264 or later and RTP/UDP transport for network bandwidth and storage efficiency as a best practice. Video encoding technology is advancing rapidly. Newer model cameras and newer encoding are becoming more efficient, delivering better imagery using less overall bandwidth. NetApp doesn’t see this trend slowing anytime soon, so it’s important to evaluate the latest encoding technologies and weigh the trade-offs. For example, H.265 significantly reduces bandwidth requirements but increases workstation performance requirements.

Compression Ratio

The compression ratio is a configurable value on the network video camera that specifies to what factor the image is reduced before transmission to the recording server. A 10% value indicates little compression; a 90% value indicates a high degree of compression. The process of compressing the image is known as quantization. In video image processing, this is implemented as lossy compression, meaning details are lost to reduce the size of the image. Selecting a high value might result in an image with compression artifacts or pixelation. An appropriate value depends on the scene complexity, lighting, shapes of objects, and colors in the scene. Values of 30% to 50% are common.

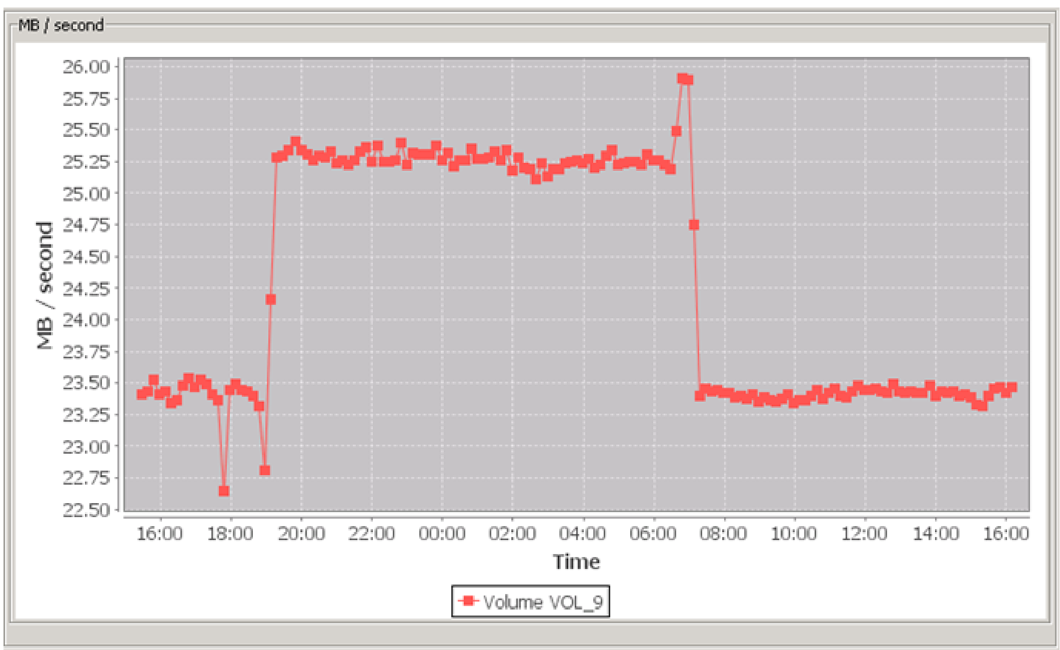
NetApp recommends leveraging the recommendations of the camera manufacturer and experience of the system integrator when selecting an appropriate compression ratio.

Variable or Constant Bit Rate

The H.264/MPEG-4 video encoder in the network video camera may be configured for either variable or constant bit rate. Constant bit rate (CBR) varies the image quality to maintain a constant output network bit rate. If there is little motion in the scene, the quality remains high. If there is a complex scene with motion (for example, trees swaying in the breeze), the image quality decreases. The decrease in image quality is recognized as noticeable pixelation of all or part of the image.

Variable bit rate (VBR) maintains the image quality but changes the output network bit rate to accommodate motion in the scene. Network video cameras deployed in areas with little or incidental motion have lower bandwidth and storage requirements with VBR. In low-light settings, the camera imager might introduce image noise, which has the same effect as motion. The sizing calculator tool of the camera manufacturer might allow the user to specify level of detail and percent motion or select an image scenario such as an intersection, stairway, and so on. A night option that returns higher bandwidth and storage requirements compared to the daylight scenario might also be available. Figure 5 illustrates the effect to ambient indoor lighting levels being adjusted from 7 a.m. through 7 p.m. with 64 HDTV cameras recording onto a single volume (LUN). In this example, there is approximately a 7% difference between the two data rates and storage required. Building lighting levels were reduced for nighttime (7 p.m.–7 a.m.), resulting in more noise in the image and thus higher data rates during the night.

Figure 5) Daylight versus nighttime data rate.



It is important to accurately estimate the selected scene complexity because it greatly influences the bandwidth and storage required. In some cases, the storage required for a scene at night might be 60% higher than during the day. Physical security integrators use infrared (IR) illuminators to provide additional light in low-light environments. The use of IR illuminators minimizes video image noise and the resulting increase in bandwidth and storage.

NetApp recommends using VBR for H.264/MPEG-4 and accurately estimating the scene complexity and percent of motion.

Audio

Most network video cameras also support an audio channel. The network bandwidth and storage required for audio are minimal compared to video. When sizing, estimate approximately 30Kbps if the audio channel is stored with the video. As with video encoding, audio encoding is becoming more efficient. In the scheme of things, the storage consumed by audio is small or negligible relative to video.

Video Management System

The VMS software is the platform for managing the network video cameras, processing of video streams from the cameras, and recording of streams to the storage array for archive and retrieval. The VMS software also integrates with access control systems, manages events, and generates alerts. Failover recording servers can be configured to continue archiving video streams in the event of a hardware failure.

Feature Set Limitations

Video management systems might be licensed by a feature set. For example, OnSSI Ocularis 5 has multiple feature sets geared to different types of installations. The Ocularis 5 feature set supports thousands of cameras at multiple locations and utilizes a central management platform for all recording servers. The recommendation for HDTV/megapixel cameras is up to 128 cameras per recording server instance with the Professional version. There is no absolute limit on the number of days of retention on the live recording volume (LUN). The distinction between multiple recording locations is further explained in the section “Tiered Storage.”

Alternately, the Ocularis 5 Enterprise and Ultimate versions have no absolute limit on the number of cameras per recording server. However, as a practical matter, most installations limit the number of cameras per server to 128 or 256 to limit the size of the failure domain. Even with failover, if a recording server fails, the cameras attached to that server are offline for a period of time.

The selection of the appropriate VMS and feature set is a collaborative effort between the video security integrator and the end user. Regardless of the VMS software package selected, any performance or hard limitations of the number of cameras per server or absolute size of volumes must be assessed to accurately size and configure the storage array.

Number of Volumes (LUNs) per Server

Most video management software can write video archives to multiple drives. For example, the recording server might have both an E:\ and F:\ drive as a target location for storing video archives. The recording server might have limitations that prevent a single camera from being written to separate drive letters. The recording server typically selects the location with the most available space when making a determination between drives.

It is more efficient to use a single volume rather than several volumes because a minimum amount of free space is required on each volume.

Continuous Recording or Record on Motion

VMS can be configured to record continuously or to record a length of time before and after a triggered event. Record on motion is the most common event to trigger recording. Because video cameras are commonly placed in areas that routinely have little motion (for example, emergency exits) or have little motion for some period of time (for example, school hallways at night), implementing record on motion can greatly reduce storage requirements.

Although configuring record on motion is efficient in the use of storage, it also has several disadvantages. Implementing server-side record on motion is CPU intensive and reduces the number of cameras that a recording server can support compared to continuous recording. Record on motion does not reduce the throughput requirements of the storage array because the recording server must temporarily store the video to disk to implement prebuffering. Prebuffering is the amount of video retained before the event. The amount of video to retain following the event is also configurable. Many video management systems store video in files that contain three to five minutes of video from each camera. It is common to implement record on motion by deleting all files except the files that contain the prebuffer and postbuffer of the event.

Record on motion also can be complicated and time-consuming to configure. Usually a detection area in the scene is selected to be the target of the record on motion algorithm, for example, a doorway. Then the degree of sensitivity is selected. A change in lighting also triggers a motion event. The algorithm detects changes in either the chroma (color) or luminance (brightness) of pixels. If the sensitivity is set too low, subtle changes do not trigger the motion event, and the video is not archived. If the sensitivity is set too high, the disk usage approaches that of continuous recording.

Other factors such as object size and percentage of object in view might also require configuration.

VMS might support camera-side motion detection with little CPU overhead on the recording server. In a camera-side implementation, the CPU on the network video camera executes the motion detection algorithm. When motion is identified, an alert is sent to the recording server to indicate that motion was detected. The server retains the video files that contain motion and deletes files that do not. Camera-side analytic algorithms might also support video analysis such as tripwire, people-counting, objects left behind, or individuals loitering events.

The NetApp recommendation is to evaluate the tradeoff between reducing storage costs against the increase in system management costs when implementing record on motion.

Video Walls

Video walls are one or more client workstations and monitors used for displaying video in a control room setting. Video can be pushed to the workstation running in a video wall mode and might also display video as the result of triggered events. From the sizing perspective, video displayed on a video wall requires the same amount of throughput as a client-viewing workstation.

Viewing Archived Video

The frequency and number of camera archives viewed concurrently incur read I/O to the storage array. Video archives may be viewed at the normal speed or at an increased playback speed. The application might support increasing the playback speed more than one thousand times the recorded speed.

Figure 6 demonstrates the performance characteristics of viewing 64 HDTV cameras at normal playback speed and transitioning to 16x playback speed.

Figure 6) 64-camera transition from 1x to 16x.

In this example, the average data rate increased from approximately 140Mbps (2.2Mbps per camera) to 400Mbps (6.2Mbps per camera) when transitioning from 1x to 16x playback. Note that the data rate increase is not a linear relationship; a 16x speed increase increased the I/O rate by a factor of 3x.

The performance characteristics associated with the forensic capabilities of the VMS-viewing client are implementation-specific and might vary between releases. Both video walls and client-viewing workstations affect the performance characteristic of the storage array but not the capacity required of the storage array. Investigative activities by the client workstations and incident reporting files might need to be considered when sizing the storage array.

2.6 Centrally Stored Video Clips

Centrally stored exported video clips (OnSSI Ocularis refers to these as bookmarks) may be stored on volumes (LUNs) on the storage array. The location is configured during the software installation.

The amount of storage required is dependent on the frequency and size of the video archives stored. These files have an infinite retention period. They are only deleted manually by the operator.

2.7 Tiered Storage

Tiered storage is the term for distinguishing between two or more types of storage that vary in price, performance, capacity, or function. Several VMS packages implement an archival function that moves video files from a temporary or recording location to an archive location. This is referred to as tiered storage because the functional aspect of the data has changed.

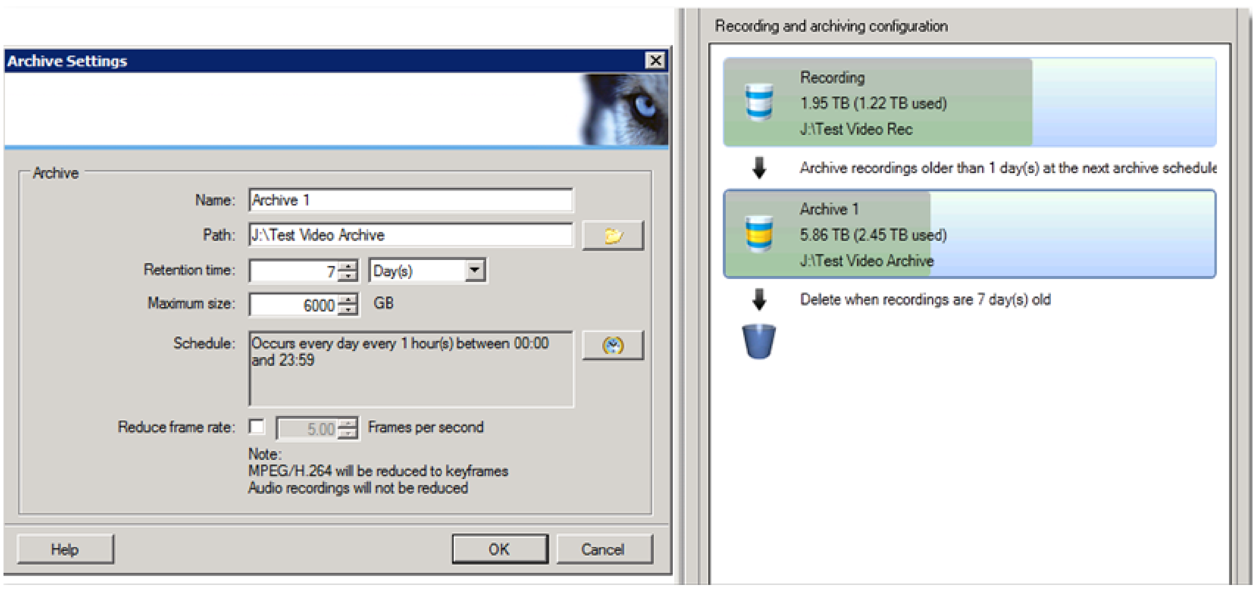
Note: Industry documentation often refers to storing video files in a database. This terminology does not imply that the video is stored in a relational database structure; rather, the individual files are stored within a directory structure of the file system.

Archiving is the term used to describe the function of moving video files from their original location on disk to an alternate location. This configuration is defined on a server-by-server basis. The configuration might have zero, one, or more archive locations. These archive locations can be defined on the same volume (LUN) as the recording location or on a separate volume (LUN) on the same or a different storage array. For example, the recording location volume (LUN) might be RAID 10, whereas the archive location could be RAID 6. RAID 10 is useful for providing protection when the recording volume is small, for instance, the size of a single high-speed HDD or a single SSD. On E-Series storage, a single DDP can be configured to support both recording volume and archive locations if desired, simplifying configuration.

When using multiple storage arrays, they need not be the same type; block-level SAN storage can be the source of the archive, and a file-level NAS device can be the destination. Files in the archive location may be backed up to tape or other media provided the archive process is not active.

During the archive process, video files may be groomed to reduce the frame rate of the original recording or may be encrypted. The archive process occurs on a user-configured schedule. It might be hourly, every four hours, or daily. This archive process affects both the performance and capacity of the storage array. Figure 7 is a sample configuration of separate directories on the same volume (LUN) for recording and archiving.

Figure 7) OnSSI Ocularis recording and archiving configuration.



In the example, both maximum size and retention time are configurable. Files in the recording directory are moved to the archive directory when the files are over 24 hours old. The archive process runs every hour. They are stored in the archive directory for seven days and then deleted.

Drive letter J: is mapped to a RAID 5 volume group with 10.914TB of capacity. The recording server does not use all of the available capacity on drive J: because of the maximum size parameters specified on the recording (1.95TB) and archive (5.86TB) configuration. This configuration can use up to 7.81TB of the 10.914TB available, or approximately 71%.

Performance of Tiered Storage

During the normal processing of video feeds from cameras and writing to the video recording storage location, the data rate is relatively constant between the server and storage array. When the archive process is initiated, the I/O characteristics change based on the source and destination of the archive. If the administrator has configured separate volumes (seen by recording servers as LUNs) for the recording location and archive location (for example, E:\RECORDING and F:\ARCHIVE), then the volume containing the recording location incurs read I/O, and the archive location incurs write I/O. The archive function in this example changes the I/O characteristics of the recording location from primarily writes to both reads and writes for the duration of the archive process.

The duration of the archive process is a function of the amount of data that must be moved. It is important to understand that the recording server writes video to storage at approximately the rate of arrival from the networked video cameras. The archive process reads and writes data as rapidly as the recording server can read, process, and write the files to the destination. In testing, the I/O rates have been observed to increase by eight times or more during the archive process.

Sizing of Tiered Storage

Every recording location and archive location must have some free space for the system to record arriving video streams. If the locations have insufficient space, the oldest video is deleted or autoarchived, if possible. If files cannot be autoarchived or deleted quickly enough to reclaim space, the recording server might not be able to write video streams to disk. It is important to allocate sufficient capacity to the respective archive locations to meet the requirements of the video retention policy. Additionally, the archive process must occur more frequently than the configured retention period for the location. For example, if the retention period is seven days, the archive function should occur at least once a day; every four hours or hourly is also commonly implemented.

2.8 High Availability

A failover server is an idle recording server that can assume the recording server role in the event a primary recording server becomes unavailable. This enhances the availability of the system; however, video is lost during the time taken to detect the server failure and initiate recording on the failure server.

The implementation deploys a keep-alive (heartbeat) TCP session between the primary and failover servers. Failover servers are defined in a group, and the primary server is configured to specify the failover group. This allows for greater flexibility and efficiency because there is no one-to-one relationship between the active server and failover server. For example, a design utilizing 10 recording servers might be configured with two or four failover servers.

Sizing of Failover Volumes

The failover servers must have sufficient disk capacity to assume the archiving function for any of the recording servers referencing the failover group. The volumes (LUNs) defined to the failover server are not used unless the failover event is triggered. When the failed server is restored and back online, the video archives on the failover server are migrated to the primary server. This recovery process adds additional I/O to the primary server and its LUNs during the migration process.

The volumes (LUNs) defined to the failover servers must be of a sufficient capacity to store video for the length of time required to detect that a primary server has failed, resolve or replace the failed component, and bring the primary server online. This length of time varies based on the geographic location of the servers, availability of system administrators at the site, and network management practices implemented.

NetApp recommends sizing the volumes (LUNs) of the failover servers to retain a minimum of three to five days of video from the recording server with the largest volume (LUN). To calculate the size of the failover volumes, divide the capacity of the primary volume by the site retention period and then multiply by the number of days to retain during failover. For example, given a 30-day retention, a volume (LUN) of 22TB (22/30*5) requires a 3.6TB volume (LUN) for the failover server.

Because of the relatively small size of the failover volume (LUN) when compared to the size of the primary recording volume (LUN), creating a volume group for each failover volume (LUN) is inefficient. Allocating one volume (LUN) per volume group eliminates disk-head contention when constantly writing to separate volumes (LUNs) in the volume group. In the example, where a 3.6TB volume (LUN) was required for the failover server, creating a separate volume group for each failover server would require the following number of drives to meet the capacity required:

- 4 drives: RAID 10 (5.5TB)
- 3 drives: RAID 5 (5.5TB)
- 5 drives: RAID 6 (8.4TB)

Failover volumes (LUNs) are idle unless a failover server is engaged. They are used only if a recording server is taken out of service for an upgrade or if a failure occurs. NetApp recommends creating a disk pool for failover volumes and allocating volumes (LUNs) to meet the requirements of the failover servers.

As an example, a disk pool with 12 6TB drives could provide for 47.227TB of capacity. From that capacity, five volumes can be created, each with a capacity of approximately 9.5TB.

```
    Volume             Capacity      DA Enabled
       POOL_2_FAILOVER_VOL2A    9,445.400 GB    No
       POOL_2_FAILOVER_VOL2B    9,445.400 GB    No
       POOL_2_FAILOVER_VOL2C    9,445.400 GB    No
       POOL_2_FAILOVER_VOL2D    9,445.400 GB    No
       POOL_2_FAILOVER_VOL2E    9,445.400 GB    No
```

Using a DDP for failover volumes provides a RAID 6 level of protection and inherent hot spare coverage with a more efficient use of disk drives than RAID 6.

Multicast and Secondary Streams

An alternate method to achieve high availability is to use multicast and secondary streams. Most IP cameras support the transport of video streams as IP multicast packets. IP multicast is a bandwidth conservation technique where the network devices replicate these packets and deliver them to multiple receivers. Some video management systems can be configured to define a single IP camera for IP multicast transport and receive the video streams on two or more recording servers. Each recording server archives the video independently. This is a means to maintain two or more copies of a video feed from a single IP camera. It is an alternative means of providing high availability without the need for primary and failover servers. Not all video management systems support this configuration.

Another means of recording the same images from an IP camera is using secondary video streams. Most IP cameras have the ability to unicast a primary video stream to one recording server and have a second recording server access a secondary stream. In some instances, the secondary stream is at a lower frame rate, resolution, or compression algorithm than the primary stream. For example, H.264 might be supported as the primary stream, but M-JPEG is only supported on the secondary stream.

When deploying cameras overlooking scenes of critical importance, using two or more cameras with overlapping field of views and defining these cameras to separate recording servers is a best practice for high availability.

The implementation of multiple unicast video streams or IP multicast transport is an alternative to providing high availability without the need for primary and failover video servers.

3 Sizing Fundamentals

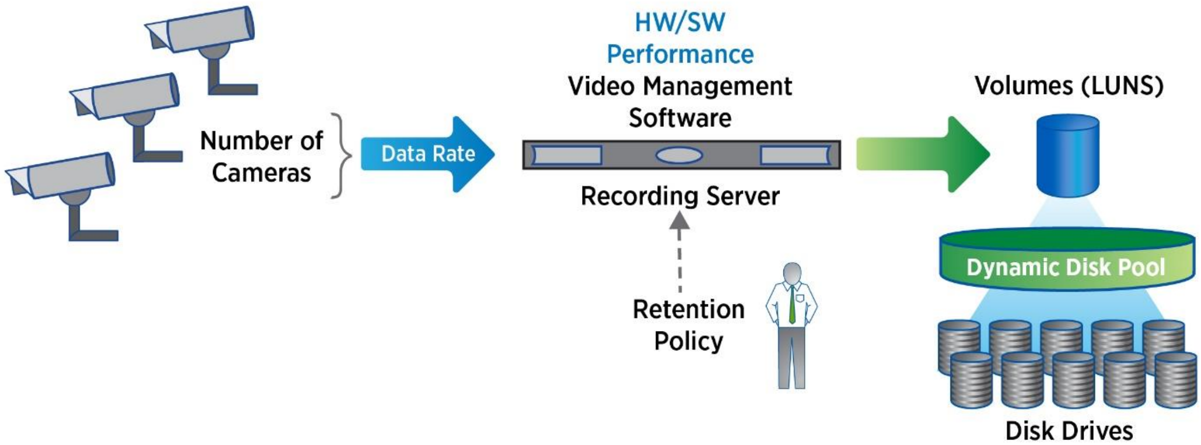
To properly size the storage array, three fundamental questions must be answered:

- What is the data rate of video received per server?
- What is the number of servers required?

- What is the retention policy?

After the data rate per server is determined, the size and number of volumes (LUNs) required by each server can be calculated based on the retention period. This process is illustrated in Figure 8.

Figure 8) Sizing fundamentals.



If all the cameras in the deployment are configured identically, the data rate is the same for all cameras, and the discussion can be simplified based on the number of cameras per server. Each server has the same number of cameras, with the resulting data rate at or below the specified maximum for the server.

	Duration (Hours:Minutes)	Delivery
Product Foundation		
Technical Overview of NetApp E-Series Storage Systems	1:00	Web-based
Technical Overview of NetApp EF-Series All-Flash Array	0:30	Web-based
Hands-On Skills Development		
Configuring and Monitoring E-Series and EF-Series Storage Systems	32:00	Instructor-led
Certification		
NetApp Certified Implementation Engineer: SAN Specialist, E-Series	1:30	Exam

5 Sizing Considerations

The VMS vendor documentation should provide a recommendation for performance and scalability. The number of cameras supported per server is a function of the version of the VMS system, the performance characteristics of the server, and the average data rate of the cameras.

Many video surveillance deployments, however, encompass a variety of camera models at different resolutions, frame rates, compression ratios, and image complexities. In such deployments, the sizing exercise becomes more complex due to the increased number of variables.

5.1 Storage, Operating System, and File System Capacity Considerations

E-Series

Video surveillance solutions based on E-Series are configurable with either an E5760, E2812, or E2860 controller configuration. Either DDP or traditional volume groups are supported. The supported drives include a variety of NL-SAS, 10K SAS, and SSD options with a range of capacity points available for each drive type. (Go to netapp.com and view the tech specs for the desired E-Series model for the latest supported drive information.)

Performance for both DDP and traditional volumes on a volume-by-volume basis is more consistent if there is one volume per volume group or disk pool. Provisioning the storage array in this manner eliminates disk drive head contention when compared to provisioning multiple volumes (LUNs) per volume group.

DDP does not require hot spare drives to be allocated. Spare capacity is reserved on all the drives in the disk pool to provide for high availability in the event that a disk drive in the pool fails. If using traditional volume groups, NetApp recommends allocating one hot spare drive for every 30 drives. A 60-drive shelf would therefore need at least two hot spare drives. Table 2 provides a reference for sizing DDP.

Table 2) DDP usable capacity (TiB).

Number of Drives	6TB	8TB	10TB	12TB	Spare Capacity*
11	38.66	50.68	63.14	75.78	2
12	42.95	56.31	70.16	84.2	2
13	47.25	61.95	77.18	92.62	2
14	51.54	67.57	84.19	101.04	2
15	51.55	67.58	84.20	101.04	3
16	55.84	73.21	91.21	109.46	3
17	60.14	78.84	98.23	117.88	3
18	64.43	84.47	105.24	126.3	3
19	68.73	90.10	112.26	134.72	3
20	73.02	95.73	119.27	143.14	3
21	77.32	101.36	126.29	151.56	3
22	81.61	106.99	133.30	159.98	3
23	85.91	112.63	140.32	168.4	3
24	90.20	118.26	147.33	176.82	3
25	94.50	123.89	154.35	185.24	3
26	98.79	129.52	161.36	193.66	3

Number of Drives	6TB	8TB	10TB	12TB	Spare Capacity*
27	103.09	135.15	168.38	202.08	3
28	107.38	140.78	175.39	210.5	3
29	111.68	146.41	182.41	218.92	3
30	115.97	152.04	189.43	227.34	3

*Spare capacity is shown in terms of drive equivalents. For example, if spare capacity is "2," that means there is 16TB of spare capacity when 8TB drives are used, 20TB of spare capacity when 10TB drives are used and 24TB of spare capacity when 12TB drives are used.

The minimum number of drives in a DDP is 11; the maximum number of disks allowed in a DDP is the total number of drives in the array. The optimal number of drives in a DDP is 30 as a best practice. The maximum number of disk drives in a volume group is 30.

After calculating the number of disks required for the deployment from Table 2, use Table 3 to determine how many controller shelves and expansion shelves are required.

Table 3) E-Series disk shelves for video surveillance deployments.

Category	E5760	E2860
Form factor	4U/60 drives	4U/60 drives
Maximum disk drives	480	180
Controller shelf	1	1
Maximum expansion shelves	7	2
Total (maximum) number of disk shelves	8	3

The following example illustrates this process. A deployment for 640 cameras, 64 cameras per server, and 10 servers is assumed. Each camera on an average generates a 1.8Mbps video feed. The physical security integrator has determined that the number of 8TB disks required to meet the recording, archive, and failover databases for 30-day retention is x. Because the E2860 can only support 180 drives, the E5760 with a DE460C controller shelf and three expansion shelves can house 240 disks.

For this deployment, the E5760 would be the recommended choice. The alternative would be to deploy two E2860 chassis, each with one expansion shelf.

New Technology File System

New Technology File System (NTFS) is the preferred file system for Microsoft Windows operating systems. The majority of VMS systems are built on Windows Server 2012 or later.

Many VMS software packages recommend formatting disks with an allocation unit size of 64KB. Given this recommendation, the maximum NTFS volume size is approximately 256TB.

Most VMS systems store video from a single camera to a file with either a maximum size or a maximum number of minutes of video. These parameters might or might not be configurable by the end user. Typically, the size of the file is three to five minutes of video from a single camera or in some cases up to 30 minutes. The documented maximum NTFS file size is approximately 16TB, which poses no practical limitation.

VMware ESXi

Implementations using VMware ESXi have an extent limit of 64TB and a file size limit of 62TB for VMFS-5 datastores.

Based on the number of cameras per server and retention period of most deployments, a recording server running in a virtual machine typically needs a volume (LUN) in the 10TB to 30TB range. VMware raw device mapping (RDM) provides access to a volume (LUN) on the storage array. RDMs require the mapped device to be a whole LUN. The documented maximum RDM volume (LUN) size is 64TB.

6 Video Surveillance Sizing Examples

This section illustrates several sizing scenarios. They increase in complexity to show various situations and how they may be sized by the video surveillance integrator.

6.1 Sizing Example 1: A Simple Deployment

The first example is for a base configuration for a video surveillance solution provisioned with two physical servers hosting four virtual machines and one E2800 storage system with 8TB drives. This sizing example assumes there are no failover servers implemented and the video management software runs under Windows 2012 R2.

The sample deployment assumes 64 Axis P1346 cameras per virtual server for a total of 256 cameras in the deployment. The Axis design tool is used to estimate the bandwidth and storage required for this model of camera. In Figure 9, a per-server estimate is shown, with line items for a single camera and then the additional 63 cameras to allow the report to calculate the totals.

Figure 9) Axis design tool bandwidth estimate.

AXIS COMMUNICATIONS

AXIS Design Tool
BANDWIDTH ESTIMATE

Project: New Project   Date: 12/1/2016   Storage: 30 days

Name	Model	Qty	View	Rec	Event	Storage
Rate per Camera	AXIS P1364	1	0 Bit/s	1.21 MBit/s	0 Bit/s	391 GB
Rate per 255 Cameras	AXIS P1364	255	0 Bit/s	308 MBit/s	0 Bit/s	99.8 TB
Project Summary		**256**	**0 Bit/s**	**309 MBit/s**	**0 Bit/s**	**100 TB**

Important notice: The AXIS Design Tool is provided for guidance only. The estimates and recommendations produced by the use of AXIS Design Tool results from a limited number of test scenarios and Axis' conclusions thereof. The user of AXIS Design Tool (the "User") acknowledges and agrees that every installation of a camera system is unique and that the actual results measured from the User's system will differ from the estimates produced by AXIS Design Tool. The estimates produced by AXIS Design Tool are thus supplied upon the condition that the User will make its own determination as to the suitability for its purposes. Axis and/or its representatives will in no event be responsible for damages of any nature whatsoever resulting from the use of or reliance upon the results produced by AXIS Design Tool. **NOTE:** The AXIS Camera Station hardware and software recommendations are only valid up to a total of one hundred (100) channels for one system, regardless of the fact that AXIS Design Tool allows more than 100 channels to be specified for a system.

This profile uses a 30-day retention, recording 24 hours per day at 10 FPS, and 1080p resolution with H.264 at 30% compression. The image complexity scenario is an intersection under low-light conditions. Note that this deployment requires approximately 0.4TB per camera per month.

Given this estimate, each virtual server requires 25TB of capacity for a total of 100TB. Referring to Table 2, configuring a 21-drive DDP provides 103.12TB of usable space. This example does not include extra capacity for failover recording servers or centrally stored video clips (bookmarks).

In this sizing example, the assumption is made that a single camera model is deployed with the same image complexity, frame rate, and compression factor. Because there is consistency in the video ingress to the server, it is very easy to determine the number of cameras per server virtual machine. There are no failover servers configured.

Many deployments use a variety of camera models at a variety of resolutions and frame rates. Failover sizing must also be included.

6.2 Sizing Example 2: Complex Deployment for a Multiuse Center

In this example, the assumption is that the video surveillance integrator is sizing the storage array to support a multiuse center containing a retail component, self-storage, and offices. The center manager provides physical security services for the complex and in conjunction with the physical security integrator has completed a site survey and determined the number and model of cameras along with the requirement and type of scene.

The physical security integrator has specified the model and manufacturer of network video cameras and their resolution, frame rate, compression standard and ratio, and image complexity. Axis Communications cameras have been selected for the project. The data rates and storage requirements have been estimated using the Axis design tool. The cameras use VBR, and the retention period required is 30 days. The values used to determine the data rates and storage required are shown in Table 4.

Table 4) Multiuse project.

Camera Model	Number	Requirement /Scene	Resolution	Frame Rate	Compression Standard/ Ratio	Image Complexity	Data Rate Mbps	Storage for 30 Days Retention
Axis Q1775	40	Identification cashier	1920x1080	30	H.264/10	Intersection	239.2	73.8TB
Axis Q1775-E	60	Overview parking lot	1920x1080	6	H.264/30	Intersection	61.2	18.8TB
Axis P3364-VE	100	Identification entrances	1280x720	12	H.264/30	Stairway	104.6	32.2TB
Axis P3214-V	200	Overview hallways	1280x720	6	H.264/50	Stairway	104.2	32.2TB
Total	400						509.6	157.4TB

Based on the recommendation of the VMS vendor and the experience of the physical security integrator, four virtual recording servers are deployed on two physical servers. Each physical server has two failover server virtual machines. In the event one physical machine fails, the failover virtual machines recover the video streams from the failed server.

In Table 5, the aggregate data rates and storage are shown on a per-camera basis. This data is used to determine the storage required per server.

Table 5) Data rate and storage per camera.

Camera Model	Data Rate per Camera	Storage per Camera for 30 Days Retention
Axis Q1775	5.9Mbps	1843GB
Axis Q1775-E	1.0Mbps	323.2GB
Axis P3364-VE	826Kbps	255.2GB

Camera Model	Data Rate per Camera	Storage per Camera for 30 Days Retention
Axis P3214-V	372Kbps	114.9GB

The cameras are distributed across the four virtual machines, as shown in Table 6. Based on the values in Table 5, the storage requirements of the number and type of camera are calculated and listed in the minimum storage required column. This column includes the assumption of 20% free space.

Table 6) Camera assignment per server.

Physical Server	Virtual Machine 1 Number/Model Camera	Minimum Storage Required (TB)	Virtual Machine 2 Number/Model Camera	Recorders	Failover 1	Failover 2
One	10 Q1775 14 Q1775-E 26 P3364-VE 50 P3214-V	43.25TB	10 Q1775 16 Q1775-E 24 P3364-VE 50 P3214-V	43.25TB	7.2TB	7.2TB
Two	10 Q1755 14 Q1755-E 26 P3364-VE 50 P3214-V	43.25TB	10 Q1775 16 Q1775-E 24 P3364-VE 50 P3214-V	43.25TB	7.2TB	7.2TB

Assuming 20% free space, 17.3TB/0.8 = 21.625TB. The failover sizing assumes five days retention based on the 43.25TB value, that is, 43.25 x (5/30), or 7.2TB.

Table 2 can be used to estimate the number of drives required to meet these storage needs. To keep things logical, the integrator decided to configure a separate DDP for each physical server. A single DDP for both would also be a workable solution. The solution using two DDPs is shown in Table 7.

Table 7) Sizing solution.

Physical Server	Minimum Required Capacity	DDP (6TB)	DDP (8TB)	DDP (10TB)
One	100.9TB	27 drives; 103.09TiB	21 drives; 103.12TiB	18 drives; 105.24TiB
Two	100.9TB	27 drives; 103.09TiB	21 drives; 103.12TiB	18 drives; 105.24TiB

This configuration uses a total of up to 54 drives, so any of the three drive capacities (6TB, 8TB, 10TB) can be accommodated in a single 60-drive shelf, leaving the system integrator free to choose the most economical configuration at the time of purchase.

7 Sizing Checklist

The following checklist describes the data that must be collected and analyzed to determine the capacity requirements of the implementation for accurate solution sizing.

Table 8) Sizing checklist.

Item	Comments
Retention period	Determines the required minimum retention period of the organization.
Reserve capacity	Estimates how much additional capacity should be allocated in excess of the minimum retention period (for example, retention period plus 20%).
DDP or RAID	Will the configuration use DDP or traditional volumes? If the latter, what RAID levels are required (for example, RAID 5 or RAID 6)?

Item	Comments
Disk characteristics	Determines the size and types of disks (for example, if 8TB NL-SAS or 1.8TB 10K RPM disks are required).
Video management software	Determines which video management system are implemented.
Volume limitations	Analyzes file system limitations of the video management system or operating system (for example, the Linux 16TB volume limitation).
Data rate per camera	Estimates the data rate expected for each type of camera deployed.
Number of cameras	Determines the total number of cameras and categorizes them by like data rates.
Continuous recording	Determines if video is recorded continuously or records on motion or on some other analytic trigger.
Frequency of viewing	Determines the frequency, number of concurrent cameras, or video streams being viewed and the number of client-viewing workstations.
Failover requirements	Determines if failover recording servers are implemented, how many servers, and the size of the volumes required per server.
Cameras per recording server	Estimates the number of servers needed for the specified cameras and the resulting data rates.
Number and size of volumes per recording server	Determines the size and number of volumes per recording server based on the VMS requirements, retention period, and reserve capacity.
Calculation of total disks required	Calculates the number of physical disks required based on the volume sizes per recording server and number of recording servers.
Additional volumes required	Determines if nonrecording volumes are required and what capacity (for example, bookmark volumes).
Hot spares	Determines if sufficient disks are available as hot spares.
Disk shelves required	Determines the number of disk shelves required based on the total number of disks required.

Summary

A video surveillance solution based on NetApp E-Series offers the physical security integrator a highly scalable repository for video management systems.

Sizing the storage array requires knowledge of policy requirements such as the retention period; it also requires knowledge of how many cameras are planned for the deployment and their configuration. Every VMS package has different performance requirements and options on how to configure storage, based on the retention and failover requirements.

Additionally, unknowns exist in each deployment that requires a sizing methodology that allows for some free space in the sizing process, in the event that the estimates for scene complexity are lower than expected. Because of the critical nature of video surveillance to loss prevention and personal safety, high availability must be a key component.

Definitions

This section contains the glossary of terms used throughout this document.

Table 9) Glossary.

Term	Definition
Controller	The controller is composed of the hardware board and firmware that manage the physical disk drives and present that capacity to a computer as logical units (LUNs).
Dynamic Disk Pool (DDP)	DDP distributes data, parity information, and spare capacity across a pool of drives. Its intelligent algorithm (seven patents pending) defines which drives are used for segment placement, making sure of full data protection. DDP dynamic rebuild technology uses every drive in the pool to rebuild a failed drive, enabling exceptional performance under failure.
H.264	A video codec—ISO/IEC Moving Picture Experts Group (Part 10)—that is more efficient than MPEG-4 and commonly used by HDTV and megapixel cameras.
HDTV	High-definition TV defines resolutions of 1920x1080 and 1280x720 pixels along with other criteria, including aspect ratio.
IP video surveillance camera	A digital video camera or network video camera is a small form factor IP networked Linux host that encodes and transports video over an IP network.
LUN	The logical unit number is an address number for how the server identifies different hard drives or, in the case of storage systems, different volumes. Most operating systems show LUNs as properties of the SCSI hard drives discovered.
Megapixel	This is any video resolution of 1 million pixels or more. However, the HDTV resolution of 1280x720 is 921,600 pixels, but is commonly referred to as a megapixel resolution.
M-JPEG	A video codec: Motion Joint Photographic Experts Group.
MPEG-4	A video codec—ISO/IEC Moving Picture Experts Group (Part 2)—and the predecessor to H.264.
NL-SAS	Near-line-SAS, or NL-SAS, drives are enterprise SATA drives with a SAS interface, head, media, and rotational speed of traditional enterprise-class SATA drives with the fully capable SAS interface typical for classic SAS drives.
RAID	RAID is an acronym for Redundant Array of Independent Disks, and it determines how data is protected from hard drive failures.
RAID 5	A striped disk with parity, RAID 5 combines three or more disks in a way that protects data against loss of any one disk. The protected storage capacity of the volume group is reduced by one disk from the raw capacity.
RAID 6	Striped disks with dual parity, RAID 6 can recover from the loss of up to two disks. The protected storage capacity of the volume group is reduced by two disks from the raw capacity.
RAID 10	RAID 10 provides high availability by combining features of RAID 0 and RAID 1. RAID 0 increases performance by striping volume data across numerous disk drives. RAID 1 provides disk mirroring, which duplicates data between two disk drives. By combining the features of RAID 0 and RAID 1, RAID 10 provides a second optimization for fault tolerance.
SAS	Serial-attached SCSI (SAS) is a computer bus used to move data to and from computer storage devices such as hard drives and tape drives. SAS depends on a point-to-point serial protocol that replaces the parallel SCSI bus technology.
Storage array	The storage array is a collection of both physical components and logical components for storing data. Physical components include drives, controllers, fans, and power supplies. Logical components include volume groups and volumes. The storage management software manages these components.
Viewing station	A high-end workstation for displaying live or archived camera feeds on a locally attached monitor.
VMS server	Video management system server is also referred to as network DVR recording server. It manages IP camera video feeds and storage media.

Term	Definition
Volume group	A volume group is a set of drives that the controller logically groups together to provide one or more volumes to an application host. All of the drives in a volume group must have the same media type and interface type.

References

The following references were used in this document:

- NetApp E-Series Storage for Video Surveillance: The advantages of simple, reliable block storage in video surveillance environments
  http://www.netapp.com/us/media/wp-7240.pdf
- TR-4196: Video Surveillance Solutions with NetApp E-Series Storage: Introduction to Video Surveillance
  http://www.netapp.com/us/media/tr-4196.pdf
- TR-4197: Video Surveillance Solutions with NetApp E-Series Storage: Planning and Design Considerations
  http://www.netapp.com/us/media/tr-4197.pdf
- TR-4198: Video Surveillance Solutions with NetApp E-Series Storage: Performance Considerations
  http://www.netapp.com/us/media/tr-4198.pdf
- Video Surveillance Storage Solution Page
  https://fieldportal.netapp.com/content/211536?assetComponentId=211635

Version History

Version	Date	Document Version History
Version 1.0	July 2013	Initial release.
Version 2.0	November 2014	Updated with E5560 and E2760 using new drive capacity.
Version 3.0	March 2017	Updated with new controller models.
Version 4.0	December 2017	Updated with new controller models and drive models

Refer to the Interoperability Matrix Tool (IMT) on the NetApp Support site to validate that the exact product and feature versions described in this document are supported for your specific environment. The NetApp IMT defines the product components and versions that can be used to construct configurations that are supported by NetApp. Specific results depend on each customer's installation in accordance with published specifications.