Specifying Storage Servers for IP security applications

The migration of security systems from analogue to digital IP based solutions has created a large demand for storage servers – high performance PCs with many hard disk drives. However, security integrators used to the older analogue technology may not have the skill sets at present to correctly specify such systems.

Many companies opt for the lowest cost PC that appears to meet the required specification without consideration of the many bottlenecks associated with high bandwidth IP video traffic. Badly specified hardware inevitably leads to poor performance of the entire system, giving the installer and the technology a bad name. The IP revolution brings new opportunities to the industry so it is incumbent upon us to make the end user experience of IP surveillance as good as it possibly can be.

From a server perspective, there are two notable pinch-points which IP video traffic must traverse. If the operating parameters of these points are exceeded by the volume of network traffic then the system will not perform as required.

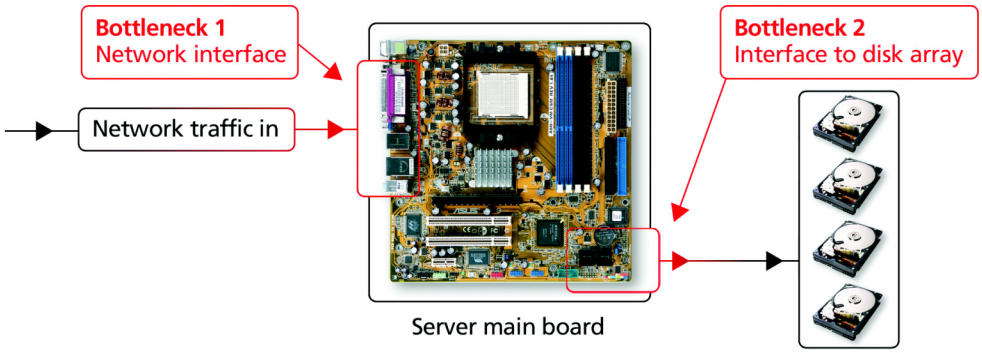


Figure 1 – Bottlenecks in an IP surveillance server

Bottleneck 1 - Network interface
Most Ethernet network interfaces are rated according to the 'wire speed' – the clock rate at which data can physically be moved along the cable. Whilst we can pump traffic into a server at 1,000,000,000 bits per second (Gigabit Ethernet), there is a limitation to how quickly any PC or server can handle the usable data further up the stack. A gigabit TCP connection could fully load a 2.4GHz Pentium 4 system in it's own right, leaving little or no resources to run the various video processing applications on the server.

With the introduction of gigabit Ethernet we have seen network speed exceed the actual data throughput capability of a typical server as the processing overhead of the network stack becomes significant. To what extent the server limits the total data throughput is a function of many parameters. Some of these attributes such as 'TCP window size' can be tweaked to

improve performance. Others, such as the quality of the NIC and associated drivers cannot readily be changed if the wrong server is specified from the outset.

TCP Offload Engines (TOE) offer a proprietary solution to TCP/IP bottlenecks by putting the entire TCP/IP stack onto the network card. This means that the NIC can handle all of the network management leaving the PC resources focussed on what they are supposed to be doing. The problem with TOE solutions is that they require proprietary hardware and do not address all of the primary causes behind data throughput bottlenecks at the TCP/IP layer.

An easier to implement and more complete solution can be achieved by using I/O Acceleration Technology (I/O AT). I/O AT is an Intel based solution to the problem of TCP/IP bottlenecks comprising several architectural improvements to server board design. Some of the complex optimisation tasks handled by this all encompassing technology are:

Enhanced DMA engine for faster data movement
Optimised TCP stack reducing PC overhead by up to 40%
Integration with existing features such as VLAN and load balancing
Platform scalability means I/O improvements scale with CPU performance

A very simple solution to data bottlenecks is to add more servers to the network and distribute the network management overhead. Most IP surveillance application software will support a maximum of 64 cameras meaning that large installations will always require a multi-server solution. As always, the cost vs. performance calculation has to be carefully considered.

Hard Drive Technology

The desired frame rate, frame size, number of cameras and length of recording will dictate the size and type of storage being used within the server. Event driven recording and increased compression (MPEG4 instead of MJPEG) will reduce the amount of storage required but IP video, by its nature, will always generate relatively large amounts of data.

Commercially available Hard Disk Drives (HDDs) are by far the most cost-effective mass storage available with a typical 750GB drive now retailing at between £100 and £150. This means that multi-Terabyte servers are the most cost-effective they have ever been and prices are still falling. As the cost per gigabyte reduces, digital CCTV becomes more affordable. It is worth mentioning that commercially available HDDs have been used in DVRs for several years and the demise of the VHS cassette recorder means that consumer HDDs have become the de facto storage medium for security systems.

The specification of appropriate hard drives is not as easy as you may think. Much of the technical specifications associated with HDDs have a fair degree of marketing spin. Some examples of this are: -

Hard drive capacity: Within the hard drive industry a kilo-byte (kB) is regarded as 1,000 bytes. However a computer works in binary, so a KiloByte is 2^{10} bytes which equates to 1024 bytes. The missing 24 bytes occurs each time you move to the next storage unit. So a typical manufacturer stated drive capacity of 500GB will be seen as 465.7GB by your Windows operating system.

Read / Write Seek Times: Often only the Sequential Rate figure is stated. However, this does not give a true indication of how the drive will perform within a security application in real world terms. The Random Seek Time will give a far more accurate indication of how quickly a hard drive will operate.

SATA, SCSI or both?
There are currently two types of HDD technology that make commercial sense for IP surveillance applications. Serial ATA (SATA) has been available for some time and now offers a theoretical, per drive, data throughput of 300 MB/s. Single SATA hard drives are now available in a 1 terabyte (TB) capacity. Storage systems utilizing these drives can achieve large amounts of cost-effective storage taking up a minimal amounts of rack space. However, many SATA drives are built on the mechanics of desktop drive technology and will not perform well within a 24/7 environment.

It is very important to choose an 'Enterprise class' drive designed for 24/7 operation as opposed to desktop grade drives that are only recommended for an 8/5 duty cycle. Whilst drives optimised for gaming and home use may offer a little more in terms of throughput, it is often achieved at the expense of reliability. In IP surveillance applications a few 10s of pounds extra spent on high quality drives are infinitely preferable to the possibility of lost video footage.

With a high camera count there is potentially 100s of Mbits each second destined to be written to the drive array. With such a high data rate, SATA drives cannot always keep up. This is when high speed drives are called for. SCSI drives have always lead the field for data read / write speed. Although high speed SATA HDDs are now available, they still fall short of SCSI performance. SCSI drives have led the field for Enterprise Class storage for many years, but the extra throughput comes at a substantially increased price.

In recent years, Serial Attached SCSI (SAS) has become available. SAS host adapters and backplanes have been developed to also accept SATA drives. This allows large storage systems to mix both technologies optimizing systems for either speed or capacity. The evolution of this 'Unified Serial' standard is captured in the diagram below.

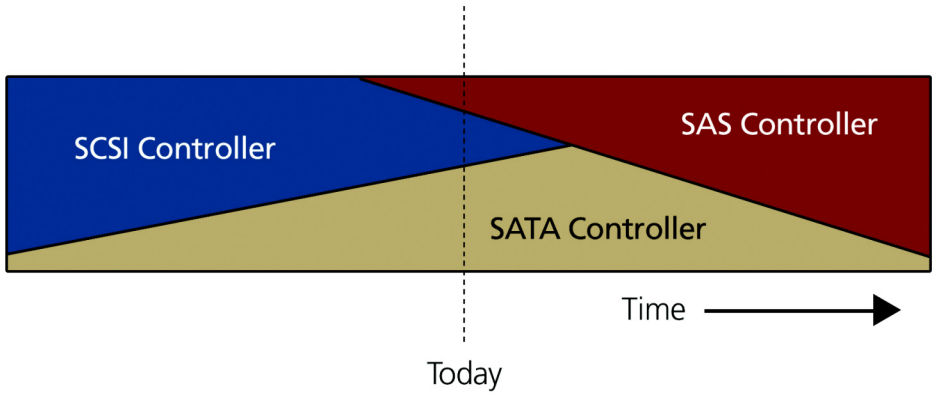


Figure 2 – Evolution of drive control technology

Bottleneck 2 – Interface to disk array

By far the greatest restriction in an IP surveillance system is the speed at which a single drive can be written to or read from (I/O). Sequential I/O is the figure that is normally quoted by drive manufacturers but this is not representative of how we use a drive for IP video. Sequential I/O is typical of large file reads and writes and typically involves operating on one data block immediately after its neighbour allowing high throughput to be achieved. In a typical IP surveillance server, Random I/O is a much more accurate description of how the drive is accessed. Many IP cameras can be writing and several viewing stations (clients) could be reading from the drive at the same time. The table below shows indicative I/O figures:

SATA disk write	55Mbyte/s	SAS disk write	70Mbyte/s
SATA disk read	55Mbyte/s	SAS disk read	70Mbyte/s
SATA disk random access	4Mbyte/s	SAS disk random access	8Mbyte/s

The figures above are not absolute and will vary from manufacturer to manufacturer and from application to application. However, they are a stark indication that the sequential I/O figures quoted by HDD manufacturers are much higher than the real world, achievable 'random' I/O in a typical storage server being used for IP video.

With the figures above, we can start to examine how the realities of drive technology could affect a real camera set-up.

Assuming a single SATA disk is used in isolation for video recording and viewing. The achievable I/O is 4MB/s = 32Mbps.
A typical MPEG4 camera can operate with a video stream of 2Mbps providing full screen, full frame rate video at reasonable quality. This equates to around 16 cameras streaming to the single disk without a degradation in performance. However, if we move to MJPEG or Megapixel cameras running H.264, each camera could use several Mbps, reducing the total possible camera count to a much lower number. Other processes that require HDD I/O will also be running in the background so the actual figure is lower than this simple, theoretical calculation.

Developments in technology to optimise random I/O capability have been taking place over a number of years. Native Command Queuing (NCQ) is designed to increase performance of SATA hard disks under certain situations by allowing the individual hard disk to internally determine the order in which received read and write commands are executed. This minimises the extent of movement of the mechanical head across the magnetic platters and provides an overall increase in data throughput. It has been suggested that drives with NCQ can improve performance by 25 to 40% but as always, the figures are very much application dependent. NCQ is a SATA based technology.

Single disk or RAID?
So far, we have only considered a system with a single hard drive to illustrate the specific limitations of HDD technology. However, storage servers rarely have a single drive and some consideration must be given to how multiple disks are required to co-operate under a typical Windows Operating System. It is practical to consider only a Windows environment as the vast majority of Video Management Software on the market is designed to run under Windows XP or Windows Server 2003.

When adding multiple drives to a system, they can be set up as autonomous 'single disks' or configured to operate as part of a 'RAID array'. Both configurations appear as a single logical volume within Windows that can be treated as an independent drive. A single disk solution is easy to implement and provides simple expandability – adding another drive provides the capability to handle another x number of cameras. Single disk bandwidth cannot be increased

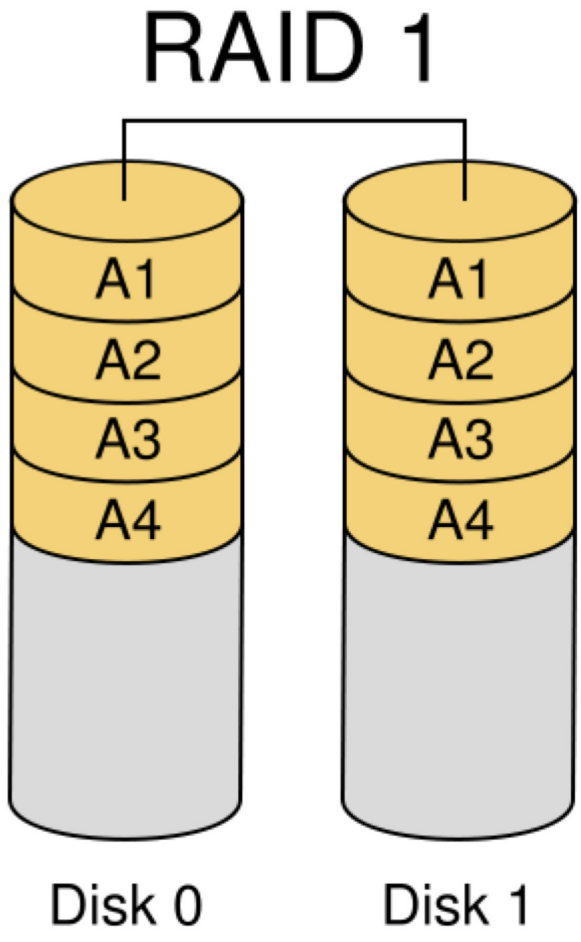
beyond the upper random I/O limit of the drive and offers no redundancy. If a drive fails, all data on that particular drive is lost. RAID offers increased reliability and the potential to increase throughput by writing to and reading from multiple drives.

RAID (Redundant Array of Inexpensive Disks) is a general term to describe data storage mechanisms that distribute data across several hard drives. Using RAID, it is possible to increase reliability and / or throughput of each logical storage unit. The usual term for a group of disks being orchestrated by a RAID controller is 'RAID array'.
A RAID array can operate at a number of different levels. These levels define the configuration of disks and the specific methods used to increase the reliability and effectiveness of the array. The core concepts that underpin RAID are:

Mirroring – writing data to two disks at once to maintain a mirror image of the primary disk that can be used in event of a disk failure.

Striping – data is written, one 'stripe' at a time, across a number of drives. This overcomes some of the throughput limitations of individual drives permitting faster total data I/O.

Parity – A calculation is performed across several drives and the resultant answer is written to an additional 'parity drive'. If any one of the drives fail, there is enough information left to recreate the lost data.



RAID level 1, mirroring - A simple example of RAID. All data is written to both drives for redundancy purposes.

What RAID level should I choose?
The question of RAID configuration is being asked ever more regularly as installers take their first steps into the world of IP security systems. Typically, only a small number of options are used, as detailed below.

One or many 'single disks'	Relatively fast I/O, no redundancy, low cost implementation. (not a true RAID level)
RAID level 0 – 'Striping'	Faster data I/O, lower reliability as any single disk failure in the array can cause a complete loss, requires a RAID controller in the system
RAID level 5 – 'Striping with parity'	Optimal trade-off of cost, bandwidth and redundancy. Operating overhead may make RAID 5 too slow for some applications
RAID level 6 – 'RAID 5 + redundancy'	Similar to RAID 5 but additional parity drive delivers increased reliability. Expensive RAID controller may be required

Unfortunately, there is no single answer to the question, "what RAID level should I choose?" It is entirely application dependent, varying with the total throughput required, budget and degree of redundancy considered 'necessary' by the client or your own design team. Basic throughput calculations and advice from knowledgeable storage server companies will help with the choice of RAID level. Expert advice should be sought from the manufacturer of the Video Management Software. Vendors of IP surveillance packages should have a detailed understanding of the hardware platform required to run their software effectively whether it is a 4 or a 64 camera system.

Other Server considerations

Redundant Power Supplies
Sourcing a high quality server with redundancy built in at the network layer as well as drive redundancy can be a pointless exercise if a cheap single input power supply is used. Power supplies represent one of the highest causes of failure of a PC / server system. Industrial grade redundant power supplies cost a little more but offer greatly improved system availability. As the total cost of a server increases, the small incremental cost of redundant power supplies makes a great deal of sense.

Self-healing RAID solutions
It is possible to configure a server's RAID array to be self-healing in the event of a drive failure. This technology can be applied to a RAID 5 array such that a 'hot spare' is automatically built into the system when another disk fails. The array is re-built over a number of hours and the system can then send an email to advise you that it has just saved the cost of an unscheduled site visit. The faulty disk can be replaced as part of routine maintenance at a more cost-effective time.

Custom configuration
Every security installation has different requirements. The final specification of server required may be a complex combination of components and technology that is not readily available as an 'off the shelf' item. Working with an experienced vendor of storage server hardware that can provide custom solutions will help you to select the right product for the job without compromise.

Software inter-operability testing

Some specialist hardware suppliers test storage servers with IP surveillance packages from key players in the IP CCTV industry to get real world figures for server performance. Relying solely on the theoretical figures can end up with disappointing results and system re-works that cost dearly in both time and money. Suppliers that understand the requirements of the software can offer a much more informed opinion than those that just provide generic server hardware.