Overcoming the IP Decoding Bottleneck
White Paper

Author: Helgi Sigurdsson (h.sigurds@matrox.com)
Date: December 19, 2008
# Table of Contents

1. **Introduction** ............................................................................................... 3

2. **The Digital Revolution** .............................................................................. 3
   2.1 Digital Camera ............................................................................................ 3
   2.2 Digital Recording ........................................................................................ 3
   2.3 Digital Monitor ............................................................................................. 4
   2.4 Digital Network ........................................................................................... 4

3. **IP Surveillance Challenges** ...................................................................... 4
   3.1 Key Advantages ......................................................................................... 4
   3.2 Well Understood Disadvantages ............................................................... 5
     3.2.1 Initial Cost .............................................................................................. 5
     3.2.2 Bandwidth .............................................................................................. 5
     3.2.3 Latency ...................................................................................................... 5
   3.3 The burden of Video Encoding/Decoding ................................................... 5

4. **The IP Surveillance Decoding Bottleneck** .............................................. 6
   4.1 Symptom .................................................................................................... 6
   4.2 Cause ......................................................................................................... 7
   4.3 Possible Solutions ...................................................................................... 7
     4.3.1 Reduce video quality ................................................................................ 7
     4.3.2 Brute force ................................................................................................ 7
     4.3.3 Multiple Computers ............................................................................... 8
     4.3.4 Hardware decoding with FPGA or ASIC .............................................. 8

5. **Matrox Video Decoding Acceleration** ..................................................... 9
   5.1 Matrox VDA-1164 ....................................................................................... 9
     5.1.1 Decoding Board ........................................................................................ 9
     5.1.2 Output Board .......................................................................................... 9
     5.1.3 Performance .......................................................................................... 10
     5.1.4 Flexibility .............................................................................................. 10
     5.1.5 Summary of Benefits ............................................................................ 10

6. **Conclusion** .............................................................................................. 11
1. Introduction

CCTV and surveillance systems are, today, going through the same transformation so many other industries did before: digitization. Initiated in the early 1990’s with the introduction of digital components in both video cameras and recorders, IP surveillance has grown steadily while becoming fully digitized. IP surveillance is now the prime option for video surveillance and monitoring.

While IP Surveillance holds several advantages over the previous generations of CCTV and surveillance systems, a few obstacles remain. Some, such as network bandwidth and latency, are well known. But other hurdles are less understood by end users and system integrators. This paper focuses on one of them: The IP Decoding Bottleneck. The symptoms of this bottleneck, including stuttering and flickering during video playback, reduced frame rate and high CPU utilization will be explained, along with the underlying cause for each. Several solutions to these limitations will be discussed and compared to a new unique solution from Matrox: the Video Decoding Accelerator (VDA).

2. The Digital Revolution

Analog Closed Circuit Television (CCTV) systems have been with us since the 1970’s, when they became popular for both security and surveillance applications. A typical system for surveillance consists of 4 components: camera, network, recorder and monitor. Over the last 20 years, these four components have been digitized; slowly at first with the integration of digital parts, but over the last 10 years, one by one, each component became fully digitized.

2.1 Digital Camera

In late 1980’s and early 1990’s, charge-coupled device (CCD) sensors were introduced into CCTV cameras replacing analog tubes. With the new CCD cameras, the input was digitized but the output remained analog, using a coaxial cable to transmit the composite video signal that was recorded on Video Cassette Recorder (VCR) tapes. The new CCD cameras and later CMOS- based cameras, offered several benefits, including a much smaller size and programmable controls for shutter speed, frame rate, resolution, etc.

2.2 Digital Recording

The next component to become digitized was the recorder. In the mid 1990’s, the Digital Video Recorder (DVR) was introduced. While it recorded all video data in a digital format, the DVR still used analog coaxial inputs to capture composite video from the cameras and featured analog outputs to connect to monitors. Despite these shortcomings, the DVR provided several benefits to the end user. With it, recording quality was consistent—no longer were numerous tapes required to store data. And searching video achieves became significantly more efficient.
2.3 Digital Monitor

DVRs have matured since they were first introduced over 12 years ago. Roughly five years after their introduction a new generation of DVRs was introduced. This allowed for monitoring of recorded video footage to take place on a standard PC. DVRs began being equipped with a network interface that enabled a PC to remotely connect to it and display recorded footage. Today, most video recorders are equipped with such a network interface for this purpose, either by default or as an option. These devices are called Network Video Recorders (NVR).

2.4 Digital Network

With three out of four CCTV components digitized, it wasn’t long until the digitization effort spread to the last remaining component of analog CCTV: the coaxial cable. Digitizing the link between the camera, recorder and monitor, the coaxial cable is replaced with standard computer network, Internet and wireless technologies. This transformation taking place today is nothing short of a revolution of the operation management of CCTV systems—offering never-before-seen functionality.

3. IP Surveillance Challenges

Now fully digitized, the 4 basic CCTV components: camera, recorder, monitor and network, are often referred to as IP Surveillance.

3.1 Key Advantages

IP Surveillance offers undeniable advantages over pure analog or hybrid digital/analog CCTV systems. The key advantages are:

- Scalability with unlimited edge devices (including cameras and other sensors) that can be added as needed
- Flexibility—with no distance or layout limitations—on where cameras are placed and from where they are viewed
- Inherent reliability and redundancy due to its distributed architecture vs. a central matrix switch
- Reuse of existing IT infrastructure vs. dedicated coaxial cables
- Easy IT integration through network standards and protocols vs. proprietary black boxes
- High-quality recordings
- Fast search and retrieval
- Easy maintenance
- Freedom from dealing with tapes
3.2 Well Understood Disadvantages

As with every new technology, IP surveillance has some disadvantages compared to analog CCTV. As this technology improves the number of disadvantages is decreasing but, as of today, the disadvantages most commented on are:

3.2.1 Initial Cost

Several analyses and studies in recent years have shown that while IP surveillance provides, in the long run, a better overall total cost of ownership, its initial costs are higher than comparable hybrid CCTV systems, unless the cable runs are very long. IP cameras are more expensive than their analog counterparts and NVRs tend to be more expensive than DVRs. Due to very high network bandwidth requirements, expensive high-performance network switches and routers need to be used. Therefore, the transition to a fully digital security surveillance solution comes with an initial high price tag.

3.2.2 Bandwidth

Network bandwidth becomes a major issue in a digital surveillance system. As the existing corporate network is the conduit used to transfer video from cameras to recording and monitoring stations, the new video traffic can overwhelm the network, slow down computers and affect the frame rate of video coming from all cameras on the network.

3.2.3 Latency

Network latency describes how much time it takes for a data packet to get from one designated point to another. On the old analog networks, latency was largely not a concern. Video would flow through dedicated coaxial cables into a large switch where it would then be re-routed to a particular monitor or recording station. An IP network is different. With IP, all devices on the network share the same medium, and network access is time-sliced between devices. As data passes over the network it goes through routers and switches, where the data packet may be buffered before it is sent out onto the next leg of the network, such as a corporate Intranet, the Internet, or wireless. Time slicing network access, and the temporary storage of data packets as the data moves across multiple network nodes, creates a visible video delay measured in milliseconds. Latency can be a significant issue, especially when someone tries to operate the camera remotely and point it at a particular object. It is like driving a car, and experiencing a delay between when you turn the steering wheel and when the car actually turns.

3.3 The Burden of Video Encoding/Decoding

The disadvantages associated with IP surveillance listed above: the initial cost and both network bandwidth and latency issues, are well known and understood by both end users and integrators. A less commonly known issue, but one that is just as important to consider is the burden of the required video encoding and decoding.
Before the camera puts the digital video data onto the network, the data is encoded. Without first being encoded, a single camera would overload the network. For example, the amount of raw data coming from a single D1 camera is around 162 Mbit/sec, which is significant considering a typical computer local area network (LAN) provides bandwidth between 100 Mbit/sec and 1Gibit. In order to send video data over a computer network, the video must first be encoded, which results in a video stream requiring as much as 100 times less bandwidth, depending on which codec is being used (i.e. MJPEG, MPEG4, H.264).

After traveling from the camera across the computer network, the video arrives at the PC monitoring or recording station. To be displayed, it is then decoded back into raw format and sent to a monitor in order to be viewed.

Significant processing power is required to both encode and decode the video signal. While encoding happens inside the camera itself, or in a dedicated encoder box that receives video from analog cameras, decoding occurs on monitoring stations and also on recording stations, if it is necessary to transcode the video into a different storage format. The large number of video streams typically processed by a monitoring or recording station can quickly surpass the capabilities of the system. While modern computers don't have a problem decoding one stream, or even a handful of streams, it is a different story when a dozen or more streams need to be decoded and monitored. Even when a computer is fully dedicated to video decoding, considerable degradation in image quality, dropped frames and a loss of data will be observed.

4. The IP Surveillance Decoding Bottleneck

So what exactly is the IP surveillance decoding bottleneck, how does it materialize, and what possible solutions are there?

4.1 Symptoms

So what are the general symptoms of the decoding bottleneck?

- Video playback of all displayed streams stutters and freezes
- Frame rate of displayed video drops
- CPU load is above 85%, and even saturated at 100%
- Opening and closing applications takes a long time
- Switching between windows and applications is sluggish
- System GUI is unresponsive
4.2 Cause

The symptoms above are caused by the single fact that video decoding is a computationally intensive task that puts a lot of load on the computer’s CPU. Video decoding is also very data intensive, putting a heavy load on the computer’s memory—placing large amounts of decoded data into memory that is then copied over the system bus to the memory of the Graphics Processing Unit (GPU).

For example, a single D1 network stream is around 2 Mbits/sec. If a monitoring station is running 10 streams, it receives 20 Mbits/sec. As every stream is decoded into raw data format and then placed into memory, each stream grows to around 20 Mbytes/sec, for a total of about 200 Mbytes/sec. This is the amount of data that needs to be processed and moved every second for the decoding task alone. Taking this into account, in addition to everything else that is running on the computer, it is clearly understood why even a high performance computer struggles to keep up.

4.3 Possible Solutions

Various solutions are available to battle the decoding bottleneck found in all IP surveillance systems today, ranging from reducing the amount of data received or decoded, to offloading the system completely from the decoding function and data transfer to and from memory all together.

4.3.1 Reduce Video Quality

The simplest solution to circumventing the IP surveillance bottleneck is to reduce the amount of data being sent and/or being decoded. For example, the workload can be cut in half by simply reducing the video resolution, or decoding only every second field in the case of interlaced video or every second frame in a progressive video format.

While simple, such a solution has significant drawbacks. Reducing video resolution or frame rate defeats the purpose of surveillance monitoring and recording—to record undisputable evidence, and to clearly see what is happening in real time.

4.3.2 Brute Force

Another solution is to throw processing power at the problem by investing in high-performance workstations that can cost in excess of US $7,000, equipped with two Quad core processors and high-end workstation graphics cards.

While such a solution could, in theory, decode and display 12 or more video streams at D1 resolution, it would come at a cost. The system would be weighed down, remaining unresponsive and sluggish to operator input, and the performance of all other applications running on the system would be impacted.
Additionally, such a system would require at least a 1,000 W power supply, and would be very loud because of all the fans that would be needed to cool the system’s components, eliminating all the hot air from the system’s chassis. All the fans required would also make the system less reliable, shortening its MTBF.

### 4.3.3 Multiple Computers

Rather than beefing up the computer’s performance, another way to overcome the IP surveillance bottleneck is to spread the load across multiple computers or PCs (a computing farm). For example, instead of using a single computer with 4 monitors for IP surveillance monitoring, four networked computers, each powering one monitor, could be used.

With such a solution, it would be possible to display 12 or more D1 resolution video streams, and more CPUs would be available to process other tasks in the background. But there are still several drawbacks with such a solution. It would require specialized software to virtualize the 4 computers as one so that the operator could still use a single keyboard and mouse to control all 4 computers simultaneously. Multiple computers also take up space and generate a lot of heat and noise. Additionally, maintenance becomes more critical and the risk of system failure increases. If a failure should occur, such as a broken hard drive, it would be much more difficult to troubleshoot and diagnose several computers as compared to only one.

### 4.3.4 Hardware Decoding with FPGA or ASIC

All possible solutions discussed so far are software based. That is, software decoding on one or more computers. But decoding can be done in hardware, as well. Field-programmable gate array (FPGA) or application-specific integrated circuit (ASIC) chips can be used to decode video streams in various formats.

Using dedicated hardware for video decoding completely offloads the system CPU. It is then free to run other tasks including those related to the network, surveillance and situation management, and access control. The drawback of an FPGA- or ASIC-based solution is flexibility. It is very hard, even at times impossible, to implement video decoding in hardware that supports IP cameras and video encoders from a variety of manufactures. While international standards exists for all main types of codecs, such as MPEG-1 (ISO 1172), MPEG-2 (ISO 13818), MPEG-4 (ISO 14496-2) and H.264 (ISO 14496-10), all camera and encoder manufactures implement codecs slightly differently. This presents a challenge that significantly reduces the likelihood of implementing a hardware decoding solution that can support IP cameras and video encoders from multiple manufacturers.
5. Matrox Video Decoding Acceleration

Building upon its long history with computer hardware—in particular display hardware, such as multi-display graphics cards and video capture cards, Matrox quickly observed the existence of the IP surveillance decoding bottleneck and the pros and cons of existing solutions.

Matrox designed a unique resolution to this limitation, taking a hybrid software/hardware approach. This technology provides the flexibility that a software solution would deliver while also alleviating the weight that video decoding places on the computer by using dedicated hardware.

5.1 Matrox VDA-1164

Designed specifically for use in workstations running Video Management System (VMS) software, the new Matrox Video Decoding Accelerator (VDA-1164) is an add-in board set, comprised of a decoding board and a quad-display output board that plugs into available PC expansion slots. It uniquely supports simultaneous decoding of multiple video streams—at multiple resolutions—from a wide variety of cameras and codecs. The VDA-1164 is ideal for use in stand-alone video surveillance systems, or in control room operator stations running video and situation management software to monitor large-scale IP security networks.

5.1.1 Decoding Board

The VDA-1164 decoding board connects to the IP camera network with a standard 1-Gig RJ-45 Ethernet jack. Incoming video streams are decoded and then sent in their raw format over the Matrox Streaming Media Bus Connector on the board to the VDA-1164 output card.

The decoding board completely offloads video decoding from the host CPU, enabling it to focus on other tasks, such as access control, system configuration, and situation management. By freeing the CPU it also maximizes system response time to alarms and operator interactions.

5.1.2 Output Board

The VDA-1164 output board receives the decoded video stream over the Matrox Streaming Media Bus Connector from the VDA-1164 decoding board and flawlessly displays multiple individual streams, in their native resolutions, at full frame rate on up to four high-definition (quad HD) displays. Streams are displayed in their captured resolution across multiple monitors—with no downscaling of imagery or dropped frames.
5.1.3 Performance

The VDA-1164 is capable of decoding and displaying up to 16 D1 or 64 CIF streams at full frame rate with superior image quality and no dropped frames. This, along with its support for IP cameras and encoders from many leading manufactures, makes the VDA-1164 a unique solution offering unparalleled performance and flexibility.

5.1.4 Flexibility

Every project is unique and each project requires different network video devices, each using a slightly different codec implementation. With this in mind, the VDA-1164 has been engineered for quick integration of new network devices to help respond to project requirements and to future-proof surveillance installations. An easy-to-use SDK is available to allow security software vendors to enable VDA-1164 support, providing even further flexibility for custom integration and the opportunity to offer additional competitive advantages within their software.

5.1.5 Summary of Benefits

<table>
<thead>
<tr>
<th>Offloads the host</th>
<th>Completely offloads the system CPU from the burden of decoding multiple IP video streams at the same time</th>
</tr>
</thead>
<tbody>
<tr>
<td>Performance</td>
<td>Simultaneous decoding and display of up to 16 D1 streams without down sampling the decoded resolution or down scaling the display resolution</td>
</tr>
<tr>
<td>Large Display surface</td>
<td>Support for 4 HD displays—enough display surface to display all incoming camera streams at their native resolutions</td>
</tr>
<tr>
<td>Great Image Quality</td>
<td>The product provides high-quality, real-time decoding and display with fast frame rate</td>
</tr>
<tr>
<td>Unparalleled Flexibility</td>
<td>The VDA provides unparalleled hardware flexibility simultaneously decoding multiple streaming formats at different resolutions from different source types</td>
</tr>
<tr>
<td>Unique Solution</td>
<td>No other comparable or competitive technology available</td>
</tr>
<tr>
<td>For demanding projects</td>
<td>Hardware-accelerated decoding offloads the effort from the workstations; providing the integrator more options to choose the right tool for the job</td>
</tr>
</tbody>
</table>

1 Performance dependent on codec and camera type
6. Conclusion

CCTV and surveillance is going through a revolution called IP surveillance. The benefits are countless, ranging from increased scalability and flexibility as compared to previous generations of analog and hybrid digital/analog CCTV systems, to overall reduced total cost of ownership. As with every new technology, there are some growing pains and challenges that still need to be addressed. One of these challenges is the decoding bottleneck found in surveillance recording and monitoring stations.

The large amounts of processing power required to decode all the incoming video streams over IP, and the amount of data that has to be moved every second, puts an overwhelming load on any IP surveillance monitoring or recording station.

The Matrox VDA-1164 is a unique solution to this bottleneck. With its hybrid software/hardware architecture, it provides numerous benefits, including the flexibility to work with any codec, IP camera or encoder and the ability to completely offload the host CPU from the daunting video decoding task.

For more info contact Matrox at: securitysolutions@matrox.com, sign up for our newsletter at securitynews@matrox.com or contact me directly at hsigurds@matrox.com.

Helgi Sigurdsson is a product manager at Matrox Graphics Inc., with a focus on IP security and surveillance solutions.