THE IBM IMAGE ADAPTER/A™

Ron J Bowater

Advanced Presentation Technology Manager
IBM United Kingdom Laboratories Ltd
Hursley Park,
Winchester,
England SO21 2JN

Introduction

The IBM Image Adapter/A™, as part of the IBM ImagePlus image processing system, is a high performance Microchannel Architecture display adapter that brings powerful image-processing capabilities to the IBM range of Personal System PS/2 computers. The adapter is a single PS/2 card that occupies the Video Feature Bus slot available on PS/2 models 50, 60, 70 and 80. The card is designed to extend the video function supplied as part of the base PS/2 machine, by providing support for large screen displays up to 1600 by 1200 pixels; image processing and display functions are performed by an on-card high performance microprocessor which has a processing throughput of 8.5 million instruction per second (mips). However, for image operations, several additional mechanisms boost the effective processor performance to several times this figure to a performance approaching 100 times that of the original IBM PC.

The card supports the attachment of an optional daughter card that allows the Image Adapter to directly drive an image scanner and printer, while requiring only a single PS/2 card slot.

The IBM Image Adapter/A can take full advantage of PS/2 Microchannel Architecture to achieve its high performance, using bus mastership for efficient access to image data stored in other parts of a PS/2 system. Bus mastership is a basic feature of the Microchannel Architecture which significantly increases adapter performance and reduces the load on the main system processor for many operations.

The IBM Hursley Laboratory

The IBM Hursley Laboratory, situated 3 miles outside Winchester, Hampshire, England, is IBM's largest development laboratory outside the United States. About 2000 people are employed in a number of development areas, producing products for manufacture and sale by IBM worldwide, for example:

- Transaction processing software
- Disk file subsystems
- OS/2 Presentation Manager
- PS/2 Display monitors
- PS/2 Display subsytems.

The IBM Image Adapter was developed as part of a family of display adapters for the PS/2 line of IBM personal computers. From initial design work in 1987 to the first product being manufactured took just over 24 months. The principal development team comprised ten engineers. The team drew on the skills and resources of many other parts of the laboratory as the product proceeded from design through development and into manufacturing test.

Image in the Office

The electronic revolution, while having a dramatic effect on the consumer industry, has been slow to affect the way in which paper is handled in the office. The majority of offices in the world still use paper as the principle means of recording and communicating information. This means the movement of very large quantities of paper, most of which is ultimately stored in mechanical filing cabinets.

The main reason for this relatively slow growth was, until recently, the lack of a number of key enabling technologies at affordable prices, for example:

- High resolution scanners - Charge Coupled Device (CCD) scanners can now provide 200 to 300 pixels per inch resolution.
- High resolution laser printers - Laser printers can provide very high quality up to 300 pixels per inch resolution.
- High resolution display screens - The use of mass- produced high-quality cathode ray tubes, together with the low cost of semiconductor storage devices required to sustain the raster scan, has made possible the display of images with quality and appearance approaching that of printing on paper.
- Optical non-volatile storage - Optical storage is the technology which really makes image processing in the office possible. A write-once-read-many (WORM) storage unit is the equivalent of an electronic filing cabinet. Many thousands of pages of scanned images can be stored on one disk for the cost of a few tens of dollars.

These technologies have now been brought together in one system that provides a complete solution for the handling of document images in the offices of large or small enterprises. It is called the IBM ImagePlus™ system.

Figure 1 shows a typical ImagePlus workstation network.

The IBM Image Adapter is a key part of ImagePlus. It provides the means for attaching:

- An image scanner
- A laser printer
- A high resolution display screen

The IBM Image Adapter also performs specialised image processing such as compression, scaling, and resolution conversion.

Requirements for the display of image

The basic video function of the PS/2 family of machines is provided by the Video Graphics Array (VGA) sub-system found on all PS/2 system boards from the model 50 upwards. This function is capable of driving a maximum of 640 by 480 pixels in the All-Points-Addressable (APA) mode, with each pixel being assigned one of 16 colours. Although this provides acceptable resolution for general purpose computing, the requirements for image processing are significantly more demanding.

The minimum size of image display is the A4 page (210 mm x 297 mm). CCD image scanners and laser printers typically operate at a resolution of 300 pixels per inch, each pixel being either on or off (black or white). The display of a full page of image at this resolution exceeds the capabilities of today's monitor technology. Fortunately, research has shown that a resolution of only 100 pixels per inch (39 per cm) can be used to display an image with a readability close to that of printing on paper, provided that each pixel is represented by one of a number of grey values. Four levels will provide good quality but 16 levels are required for excellent quality.

The display screen for a single A4 page of image should therefore support a minimum of 800 pixels wide by 1150 pixels deep. Allowing for a small area of the screen through which the user can communicate with the image system while a full A4 image is being displayed leads to the design point chosen for the IBM Image Adapter of 850 by 1200 pixels.

Although the display of a single A4 page of image is a minimum requirement, it is also useful to be able to display two A4 pages side-by-side for comparison purposes, or to enable both system alphanumeric information and the image display to be presented on the same area of glass. An image area of 1600 pixels wide by 1200 pixels deep best meets the needs of this application and would ideally be displayed on a screen of 19 inch diagonal measurement.

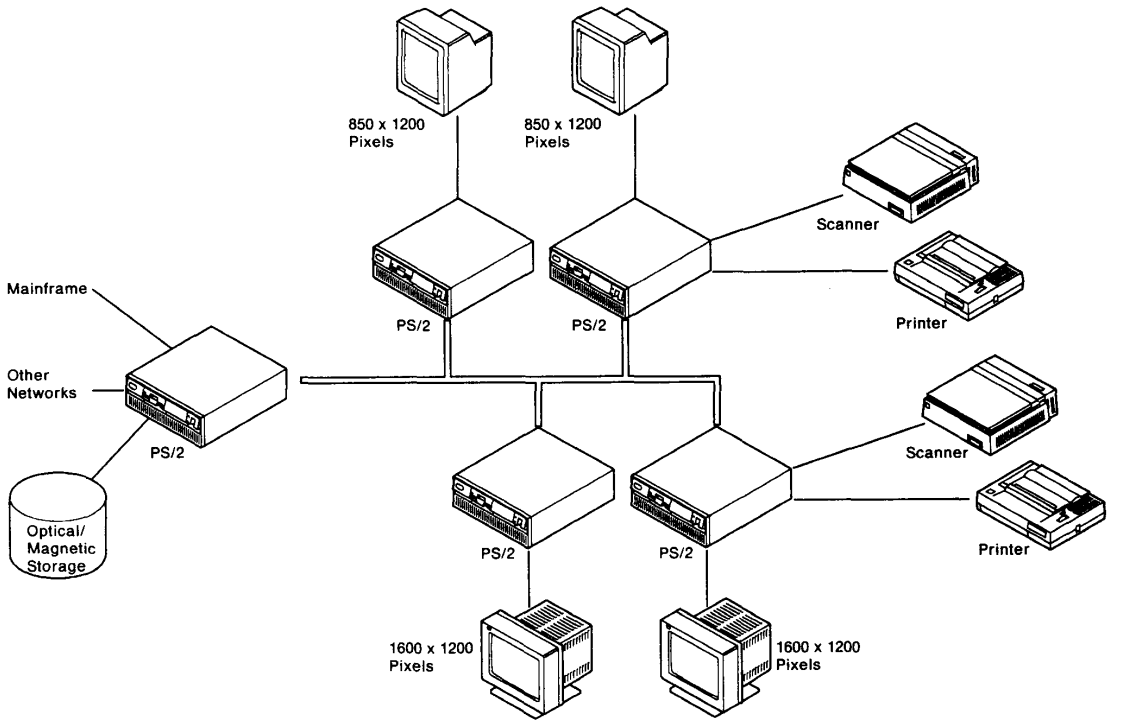

Figure 1. A typical ImagePlus workstation network

Graphics display screens have in the past operated with a black background. However, the display of image with a paper-like quality requires a black display on a white background. On a standard CRT computer display the display of a black-on-white image refreshing at 60Hz (frames per second), produces noticeable flicker. Phosphor types that elongate persistence of the image tend to suffer from reduced brightness. It is now generally accepted that a refresh rate above 70Hz is necessary to completely eliminate flicker when using standard TV prosphors. The display of image is a case where field interlace can be used to maintain the apparent refresh rate. Using this technique, even and odd lines are redrawn alternately; this reduces the pixel rate by a factor of two while preserving the refresh rate for any area of the screen.

The application requirements lead us to a calculation of the pixel rate required to drive such a display format. Assuming an interlaced refresh rate of 80Hz, the pixel rate necessary to scan 1200 lines of 1600 pixels out to a display screen is approximately 128 MHz, that is, 8 nanoseconds per pixel. This is five times faster that the standard VGA on the PS/2 system board and is extremely demanding for the display adapter. When each pixel can take on one of 16 shades of grey (4 binary bits), the adapter is required to supply data at a peak rate of over 500 million bits per second.

Data rates of this magnitude cannot be sustained by standard dynamic random access memory (RAM) modules and therefore a special device, known as a Video RAM (VRAM) must be used. The VRAM is a modification of the standard dynamic RAM (DRAM) in which a second way to access the data has been provided by means of an on-chip shift register. This shift register is typically loaded with a row of data from the RAM array during the CRT fly-back time, the data being shifted out at high speed as the electron beam flies across the CRT screen. This frees up the main parallel VRAM port for access during the active display times so that one image can be displayed while another is being created. The data rate attainable out of the shift register is much higher that that which could be achieved using the VRAM parallel port. External logic is also considerably simplified.

Requirements for the processing of image

As already mentioned, the standard resolution for high resolution scanners is 200 to 300 pixels per inch. A single A4 page at this resolution would take almost 2 megabytes of storage, making the cost of storage and transmission prohibitive. For example, at 9600 baud, the transmission of a page of A4 image would take almost 30 minutes!

Fortunately, simple techniques can be used to achieve compression ratios of up to 25:1. These use run length encoding together with an assumption of the frequency distribution of black and white runs to achieve a relatively high compression ratio.

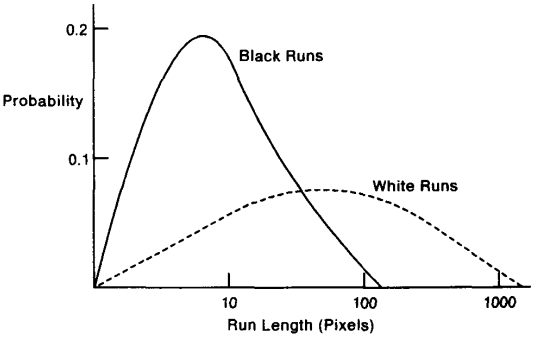

Figure 2. Probability of pixel runs in scanned image

As shown in Figure 2 typical documents have many short runs of black pixels and not many long runs. White runs tend to be much more evenly distributed. The main CCITT algorithm for Facsimile takes advantage of this characteristic in assigning smaller code points (fewer bits) to short black runs and long white runs.

A second CCITT algorithm transmits the difference between one scan line and the next. This can be very efficient, but suffers from the problem that an error in transmission can propagate down the whole screen.

Image manipulation and processing can be a complex and time-consuming process, involving one or more operations on each pixel in an image. Typical functions used to manipulate images are:

- Compression - The reduction in number of bits representing an image prior to being stored or communicated between components of a system. Compression ratios of between 5 and 25 can be achieved on typical scanned images.
- Decompression - The reverse process to compression.
- Resolution conversion - The conversion in resolution between 300 pixels per inch at black and white levels, 100 dots per inch at 4 or 16 grey levels, or any other resolution required by components of the image system.
- Scaling - The reduction or increase in the displayed size of an image.
- Changing contrast - The reduction or increase of the brightness difference between the highest and lowest intensity components of an image.

The complexity and number of different image-processing functions, together with the parametric variations possible with each function, makes a purely hardware implementation impracticable. A programmable processor with loadable microcode was considered to be the best solution to this problem.

All of the above image processing functions operate on all pixels within an image and may require between 10 and 50 processor instructions to perform the function for each pixel in the image. In order to maintain good responsiveness and interactivity, these functions should not take more than one second to complete. For an image of 800 by 1150 pixels, an algorithm with 10 instructions in the inner pixel loop requires a processor with a performance of 8.5 millions of instructions per second (mips) to completely process the image. As a reference, the 80286 processor in a PS/2 model 50 runs at approximately 1.5 mips.

Requirements for the input and output of image

For many workstation users, the display of previously stored images is adequate for their application. However, a need exists to provide the facility for reading the new image data from a CCD scanner and for printing hard copy using a laser printer. It is a sensible economy if the image-processing function for the display should be available for any scanner or printer in the system.

In addition, with current technology there are real advantages to a system design where all image data is maintained and processed within a 'sub-system' which has its own:

- Storage
- Programmable processor
- Input devices
- Output devices
- Link to the main system

Image processing is an ideal candidate for implementation within a sub-system in which the Microchannel and PS/2 main storage are not required for many operations, thus freeing up valuable system resources.

Figure 3 shows a block diagram of an image processing sub-system using the Image Adapter.

Choice of processor architecture

Alternatives

During the initial design phase of the IBM Image Adapter a number of choices were considered as the basis for its programmable display processor. The options were:

1. An existing Complementary Metal Oxide Semiconductor (CMOS) processor such as provided by the Motorola 680x0 and Intel 80x86 families
2. A bit slice processor such as the AMD 2901 or 29116
3. A new architecture specifically designed for image and graphics processing.

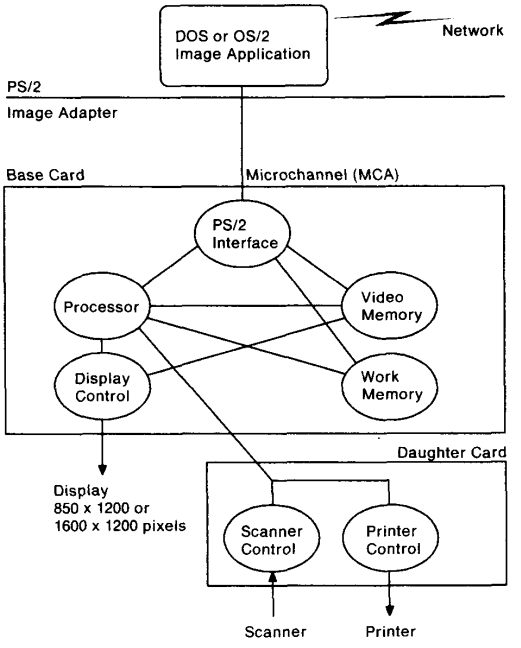

Figure 3. The Image Adapter sub-system

Requirements

The requirements of the image processor were considered to be:

1. Performance - As discussed in a previous section, a minimum processor throughput of 8.5 mips or better is required to satisfy the image processing performance design objectives.
2. Board space - The subsystem needed to fit on a single PS/2 card. After space for video memory, microchannel interface buffers, and display support logic had been allocated, the display processor had to be accommodated in a few square inches of board space. A solution had to be found which put a fully programmable 8.5 mips processing subsystem in this space.
3. Cost - The aim was to keep the total cost of the processing subsystem within a reasonable limit consistent with the target cost for the card as a whole.
4. Data width - For best performance, the processor needed a data flow width of 32 bits in order to minimise the overheads of manipulating 32-bit addresses in image- processing algorithms.
5. Power - The total power consumption of a PS/2 card should not exceed 8 watts. After memory and support circuitry had been considered, a figure of 1.5 watts remained as the processor power design objective.

CMOS standard processor:

The CMOS processor alternative was eliminated fairly quickly on account of a number of factors. The 8088 processor used in the first IBM Personal Computer had a performance of 0.27 mips and the current top-end of the PS/2 range provides around 4 mips of 32-bit processing power. Although the latter is getting close to the Image Adapter design objective, there would be no obvious performance gain for the image sub-system over the implementation of the same function on the main PS/2 processor. Furthermore, an implementation of a complete processing sub-system using the Intel 80386 (or Motorola 68020) would certainly not satisfy cost, power, or board space objectives. It was on these grounds that the standard CMOS processor alternative was eliminated.

Bit-slice processor

The processor power requirement of 8.5 mips could certainly be satisfied with a "bit-slice" implementation using a number of standard bipolar components. In fact, this is the usual industry approach for obtaining processing performance in excess of 5 mips and has been used in many products, including some made by IBM. However, this approach exceeds the available power, cost, and board space. In fact, a full 32-bit bipolar 8.5 mips bit-slice processor would probably occupy the area of three PS/2 cards, cost several hundred dollars and consume 10 watts or more of power! These considerations meant that the bit-slice approach was not appropriate.

Custom image processor

The next option considered was to design and implement a new processor architecture with special features to significantly improve the performance of image-processing algorithms.

Benchmarking of several key image-processing algorithms revealed a number of important factors that could significantly improve image performance. These are:

1. Register-based architectures (more than 8 registers) performed significantly better than stack-based architectures on account of the need, during algorithm processing, to quickly access parameters and to maintain multiple loop counters and addresses. This factor becomes significant at high processor clock rates owing to the additional delays incurred when a stack-based architecture needs to access the stack, usually held in main memory.
2. Complex memory addressing modes tended to be less efficient than simple addressing modes with overlapped processor execution.
3. A significant amount of time is spent reading and writing pixels between registers and memory. Pixels are usually represented as 1-, 2-, 4-, or 8-bit values, packed into bytes and 16- or 32-bit contiguous memory words. Pixels are then organised into bit maps which are two dimensional arrays of pixels addressed using an X,Y coordinate system. The overhead of computing the address of a pixel in memory from a base address, number of bits per pixel, X address and Y address, followed by the read and write of that pixel could take 80% or more of the instructions in an image processing algorithm.
4. There are a small number of pixel-based operations that, if implemented directly in hardware, significantly improve image-processing performance. These operations can be designed to overlap with other processor instructions so that the effective throughput can be more than doubled.

Image Adapter Description

The IBM Image Adapter is a high resolution display adapter for all IBM Personal System/2 units having the IBM Microchannel Architecture and is shown in Figure 4.

The ideal image processor

A specification gradually evolved for the ideal image processing that could be implemented using current technologies. It corresponded fairly closely with the accepted Reduced Instruction Set Computer (RISC) definition:

1. Small set of simple, fast instructions
2. Large number of registers
3. Single cycle instructions
4. Simple addressing modes with no more that one load or store per instruction.

The pixel data type

However, these processor characteristics alone did not give the performance required for the more complex image-processing operations. From the benchmark conclusions already discussed it become apparent that the addition of the pixel as a basic processor data type (in addition to bits, bytes and words), would give a significant performance improvement. Beyond the basic pixel type with 1, 2, 4 or 8 bits per pixel, an addressing structure for pixels was also introduced.

Pixel primitives

A number of generally useful operations on pixel maps can be defined and implemented directly in hardware. These include:

1. Read or write pixel
2. Read or write pixel and step X,Y to one of the eight neighbouring pixels
3. Write a sequence of pixels given its length and direction
4. Block transfer of data from one bit map to another.

Pixel 'mix' functions

A further offload of function from the programmable display processor can be achieved with the implementation of arithmetic and logic 'mix' functions in hardware.

The offloading of logic and arithmetic functions into hardware can save several processing instructions per pixel write. Hardware automatically performs a read-modify-write memory cycle on the accessed bit map, again overlapped with processor instruction execution.

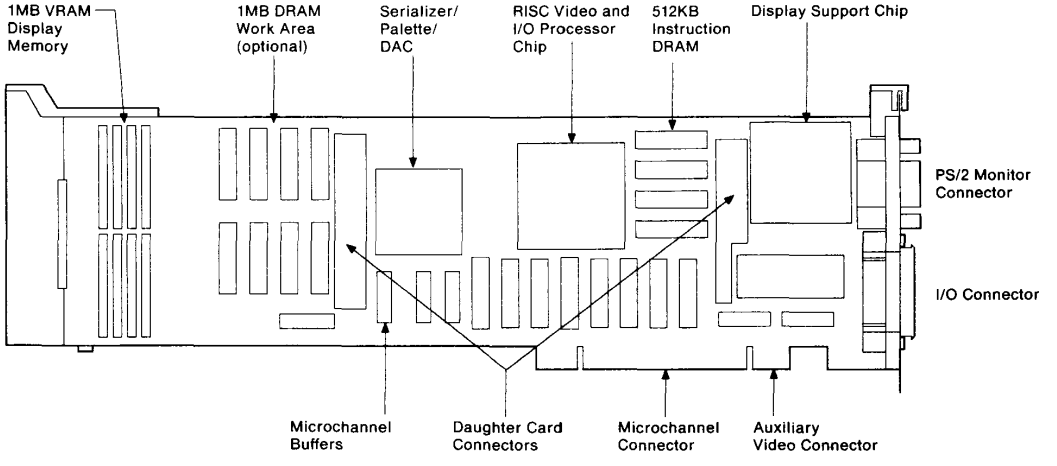

Figure 4. The IBM Image Adapter PS/2 card

The Image Adapter is capable of driving displays, up to a 128MHz pixel rate, with addressability of up to 1600 x 1200 picture elements with 16 levels of grey.

An internal block diagram of the Image Adapter is shown in Figure 5.

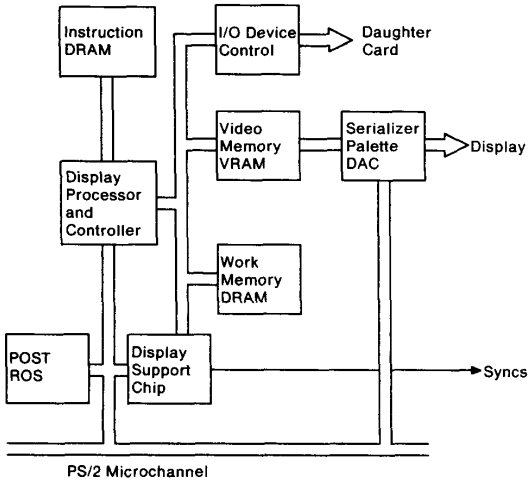

Figure 5. Image Adapter block diagram

The Image Adapter contains a 32-bit 8.5 mips display processor, a 128 MHz Serialiser, Video Palette, Digital-to-Analogue Convertors, and a CRT controller chip. It plugs into a single 16- or 32- bit Video Feature Bus system slot. Further IBM Image Adapter cards, up to a total of four, can be plugged into regular 16- bit slots of the Microchannel bus for applications that require multiple displays. Memory on the adapter comprises 1 Mbyte of Video memory (VRAM), 512 Kbytes of instruction RAM for the display processor (256K instructions), and sockets for an additional 1 Mbyte of standard dynamic memory (DRAM). The latter is available for those applications needing more data memory or improved performance.

An 'optional' daughter card plugs onto the base Image Adapter card to provide support for direct attachment of an image printer and/or scanner.

Memory access model

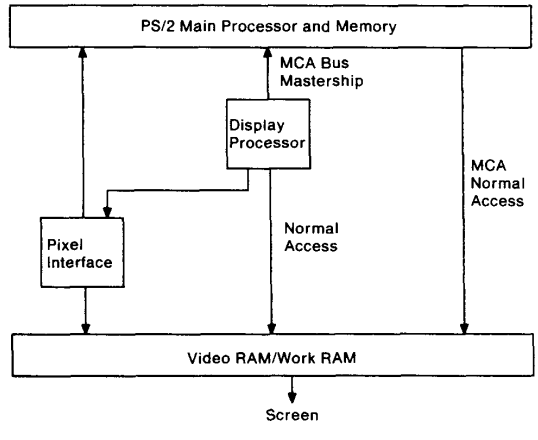

Figure 6. Image adapter memory access model

The PS/2 processor can access any area of the on-card VRAM or DRAM using normal memory access on the Microchannel, the Image Adapter's VRAM or DRAM being mapped into a window of the PS/2 engine's address space. I/O commands are used to control the display processor (start, stop, interrupt, for example). The display processor has all on-card memory in its address space and also any storage that is connected to the Microchannel, including any memory on the main PS/2 system board. This access is achieved using Microchannel bus mastership, mechanism which significantly increases performance when compared with other techniques.

Pixel bit map access

The display processor, being able to view storage as a linear array of bytes, is also able to view any area of storage as a pixel bit map (described in an earlier section). Bit maps could be either in the on-card VRAM/DRAM, or in PS/2 memory accessed using the Microchannel bus mastership mechanism.

Technology

In order to satisfy the demanding requirements discussed previously, several advanced silicon technologies have been used in the design of the IBM Image Adapter. These are:

- The Display Processor and Controller A 1 micron CMOS standard cell integrated circuit providing the equivalent of 40,000 gates on the chip using around 250,000 transistors.
- The Serialiser/palette/DAC A 1 micron CMOS mixed technology custom integrated circuit using the following circuit types integrated on a single chip:
  1. High speed static RAM (8ns access)
  2. Logic (6000 cells)
  3. Digital-to-Analogue convertors, custom designed
  4. Phase-Locked-Loop and Voltage-Controlled-Oscillators, again custom designed.
- The Display Support Chip A small gate-array VLSI chip providing synchronisation and timing pulse generation for the display and Image Adapter card.
- The Frame Buffer Video RAMs providing 1 million bits of storage per chip.

Design Methodology

A workstation-based VLSI design methodology, developed by IBM Hursley, was used to design and develop the custom silicon components employed in the Image Adapter. This design methodology offers a number of advantages when compared with conventional logic design techniques, allowing:

1. A 'right first time' VLSI chip to be produced, every time (that is, a reliable 1-pass design)
2. A rapid design and turn-around schedule
3. Maximum re-use of existing logic circuitry
4. The ability to move between types of target technologies (for example, from gate array to standard cell) with minimum logic redesign.

The VLSI design methodology used for the Image Adapter has been described by West (WEST89).

The Display Processor

The heart of the IBM Image Adapter is a single chip VLSI display controller and processor chip utilising IBM's advanced standard cell technology. This chip is 9.4 mm square and contains the standard cell equivalent of 40,000 two-input NAND gates, organised as a matrix of 27,720 wireable cells. It is packaged in a 194-pin IBM Pin Grid Array (PGA) and consumes approximately 1 watt of electrical power.

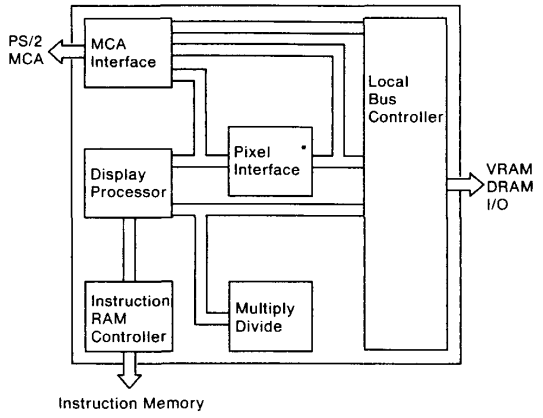

Figure 7. Display processor and controller block diagram

A block diagram of the display controller and processor chip is shown in Figure 7. It contains most of the digital logic that gives the IBM Image Adapter its very high performance and rich function.

The display processor provides a 'programmable display adapter' that can be personalised for image, text, or graphics simply by providing a new microcode load for the on-chip processor. The performance of the adapter is sufficiently high to allow algorithms such as image compression and graphics rotation to be performed at hardware rates but having the flexibility of programmable microcode.

A complete software development system is provided for the Image Adapter, to be used by systems development programmers. This toolkit will be available under license from IBM to its business partners, allowing programmers to take advantage of the on-chip processor in off-loading function from the main PS/2 processor. Advantages to be gained from such an off-load include:

- Higher performance than the PS/2 processor (2 to 5 times)
- Dedicated processing (not likely to be interrupted by other tasks)
- Reduced load on the main PS/2 processor.

The Display Processor chip has been described in more detail by Gay (GAY89).

The Serialiser/Palette/DAC chip

The IBM Image Adapter uses a fully integrated Video Serialiser, Colour Palette and Digital-to-Analogue convertor chip using 1 micron CMOS technology.

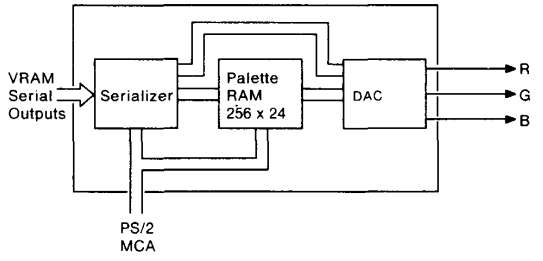

Figure 8. Serialiser/palette/DAC VLSI chip block diagram

A block diagram of the Serialiser/palette/DAC chip is shown in Figure 8. It is capable of driving high resolution colour monitors up to 1600 by 1200 pixels at 128 Mhz video frequency with 8 bits per pixel. It offers a colour palette of 256 colours selectable from a possible 16 million colours. The design represents the state of the art in mixed analogue and digital CMOS design. The VLSI design methodology described earlier in this paper was used in the development of this chip.

The Serialiser/Palette/DAC chip has been described in more detail by Chesters (CHESTERS89).

Software Support

As part of the IBM ImagePlus workstation, the IBM Image Adapter operates in the PC DOS environment, under the control of the IBM ImagePlus workstation program.

Adapter Flexibility

The IBM Image adapter has been announced specifically for the ImagePlus product but is based on a flexible design. This flexibility is due to the following:

- The Display Processor with its loadable microcode store containing a maximum of 256K instructions. Above this limit a paging scheme could be implemented which allows microcode to be moved between PC memory and the Display Processor instruction storage on a demand basis.
- Fully programmable CRT sync and VRAM address generator that allows for screen formats up to 4096 pixels per line and 4096 lines per screen, interlaced or non-interlaced. All parameters such as sync pulse widths are fully programmable.
- Programmable pixel depth - Both the Display Processor and the Serialiser/Palette/DAC chips can support bit-maps in memory of pixel depths 1, 2, 4 or 8 bits per pixel via the SPD palette and 16 bits per pixel using a palette bypass.
- Daughter card attachment - The Image Adapter base card provides for the attachment of an optional 'daughter' card which can contain additional logic to support a particular application. The daughter card connects to the main processor address and data bus and also is connected to a general purpose 36-pin connector on the rear outlet of the Image Adapter card. Possible daughter cards include:
  - Laser printer interface
  - Scanner interface
  - Video RAM extension
- PS/2 bus parameters - By means of the Microchannel Programmable Option Select (POS) mechanism, several key configuration parameters can be changed for a particular application. This include the addresses recognised by the Image Adapter in the Microchannel I/O and memory address spaces.

Summary

This paper has described the IBM Image Adapter, the company's latest and most advanced display adapter for the PS/2 family. As a part of the ImagePlus product set, it provides scanner, printer, and display attachment as well as the image processing functions that are vital to the success of an office image-system.

The IBM Image Adapter is an important part of IBM's display strategy and is likely to be used as a standard component in many new and exciting areas of workstation technology.

References

[1] WEST89 - The VLSI 'Silicon Compiler' Design Process.

[2] CHESTERS89 - A 1-Micron CMOS 128 MHZ Video Serialiser, Palette, and Digital-to-Analogue (DAC) Chip.

[3] GAY89 - A VLSI Display Controller and Processor.