

Hard disk duplications ICS – white paper

Hard drive duplication is the process of copying the contents of one hard drive to another hard drive or multiple drives. The first hard drive duplication units were introduced to the market in the late Eighties by Intelligent Computer Solutions, Inc. The technology developed for these first units was a crucial element in the process of automating the mass production of personal computers. This automation made it possible for manufacturers to offer consumers “ready to use” computers at an affordable price, and computer sales were driven up from thousands a year to 100 million a year.

Aside from the computer manufacturers’ obvious need for hard drive duplication there is an entire spectrum of non PC products that use hard drives as well. Products like: TV cable boxes, video games, medical equipment, communication equipment, etc. The manufacturers of these types of products need a way to clone the content of the internal hard drive just as the computer manufacturers do. Other applications that require an efficient way of coping data include: IT departments of organizations that maintain and update their own computers, service organizations, IT consultants, and last but not least, computer forensics investigators.

The diversity of the applications requiring hard drive duplication spawned a variety of solutions:

Hardware solutions

Ranging from low cost hardware units (including computer based devices or simple stand alone units that copy from one hard drive, sector by sector, to another similar hard drive) to more advanced units that can use intelligent copy mode (IQ copy) which only copies the occupied cluster and even scales the partitions of the target drives. Such advanced units can copy from one hard drive to multiple hard drives and from smaller to larger or from larger to smaller hard drives. Some of the most popular such devices are the ICS Image MASter line (Image MASter). The IM4008I Hard Drive Duplicator, equipped with intelligent copy (IQ Copy) or Smart Copy duplicates 8 hard drives from 1 master. The IM4004I Hard Drive Duplicator, copies to 4 hard drives and the Solo 3 (the hand held duplicator) duplicate 2 targets simultaneously.

Some of the main advantages of the hardware solutions are:

- Simple to use – usually the entire duplication process won’t require more than a couple of push buttons.
- Does not require much expertise from the operator.
- Very fast – the hard drive itself becomes the throughput bottleneck.
- No licensing fee –making it cheaper for large volume duplications.
- No network or computer to deal with.
- Very reliable and portable – suited for field operation as well.
- It is not Operation System dependent.

Some of the disadvantages of the hardware solutions are:

- It is more expensive for single or small duplication needs
- It requires direct access to the media which some times is viewed as a disadvantage

Software solutions

Similar to hardware, software solutions also range from home grown Linux or DOS applications that image a single hard drive, sector by sector, to another slave hard drive, to more advanced programs, similar to the ICS duplicators, than can clone a drive intelligently.

The latest software solutions, such as Symantec's Ghost for network, allow a simultaneous multicasting of a load or image to many computers, connected on the network,

The main advantage of the single hard drive software solution:

Its low cost, which is typically below \$50.

The disadvantages of the single drive software solution:

- User needs significant expertise.
- Requires a dedicated computer.
- One hard disk must be configured as a slave drive.
- Not suitable for forensic applications.
- For installation, the computer must be opened unless using an external USB or 1394 bridge (in which case the throughput is much slower, and there is the additional cost of the bridge).

Advantages of the **software over network** solution:

- No need to open the computer.
- **It is easy to select a different image to each computer.** This is especially attractive to large organizations that already have a network and already pay for its maintenance, and only need to update their computers once in a while.
- It is easy to select images from a data reservoir and to multi-cast over the network (ideal for build-to-order operations such as Dell computers).

The disadvantages:

- Initial cost is very high.
- Ongoing cost is high due to software licensing fees and network maintenance costs.
- Total speed is usually 10 times slower than the hardware solutions.

Selecting the appropriate duplication solution is only the first hurdle. Other issues which complicate both hardware and software duplication are the complexity of drives types, interfaces, cables, and the rates at which they are all continuously changing.

For those of us not fluent with the buzz words, here is a brief overview:

The first commercially viable interface was ST506 MFM (Modified Frequency Modulation). The hard disk had two edge connectors: 34 pins and 20 pins. Just before the “end of life” of the MFM drive, RLL (Run Length Limited) hard drives appeared. RLL was an improved coding scheme of the MFM method which resulted in a 50% capacity increase. A separate controller that resided inside the computer (or the PC later on) converted the drive analog signals to digital data. It controlled the hard drive motor and analog data and it was responsible for handling bad sector, ECC (Error Correction Code) and low level format among other tasks.

In the Seventies, another more advanced interface was proposed by disk drive manufacturer Shugart. The new interface used logical rather than physical addressing. Even though Shugart failed in its attempt to make its new SASI (Shugart Associates Systems Interface) interface an ANSI (American National Standard Interface), it paved the way for the better and more advanced SCSI (Small Computer System Interface) interface.

In the late Eighties, Western Digital and Tandon integrated the controller inside the drive and came up with the new 40 pin IDE (Integrated Drive Electronics) interface that quickly become the industry standard. To maintain software compatibility the initial IDE interface maintained the conventional commands features of the MFM/RLL drives, but with every new version of the IDE spec new features and commands appeared while others disappeared.

Some examples of famous commands that no longer exist:

- Bad sectors – In the era of MFM/RLL, the user had to manually mark the bad sectors using the provided manufacturer list. In the early IDE version they were already marked by the manufacturer, but the user could read their location and add to the list. In today’s EIDE (Enhance IDE) they are completely hidden from the user. The user can not read them or mark them. They are handled solely by the drive’s micro controller.
- Low level format – It was necessary for the user to perform this command before the first use of the drive. This process would scan the drive for bad sectors and would number the sectors and the tracks. The interleave value was entered to optimized performance. Today this operation is done during the drive’s production and is not available to the user any more.

New commands that were not available before:

- Smart Commands – Set of commands that allow the user to get statistics about drive features and condition
- HPA (Host Protected Area) commands - Set of commands that allow for a restricted area on the drive.

- DCO (Drive Configuration Overlay) command – Set of commands that allow for global setting of drive features, including a restricted area
- Password Commands – Several levels of password to protect drive access (available as global access and in HPA and DCO modes).

The following charts illustrate the progression of the main computer device's interfaces and highlight the complexity of supporting numerous standards, transfer rates and cables.

ST 506
MFM/RLL

400-600
KB/sec

20 pin data
34 pin cntl
4 pin pwr

IDE
PIO

16
MB/sec

40 pin data/cntl
4 pin pwr

EIDE
UDMA

16 – 160
MB/sec

80 pin data/cntl
4 pin pwr

EIDE
SATA

1.5
Gbit/sec

7 pin data
4 pin pwr

EIDE
SATA 2

3
Gbit/sec

7 pin data
4 pin pwr

SASI

400
KB/sec

50 pin data/cntl
4 pin pwr

SCSI

5
MB/sec

50 pin data/cntl
4 pin pwr

SCSI 2

20 – 80
MB/sec

50/68 pin data/cntl
4 pin pwr

Wide SCSI

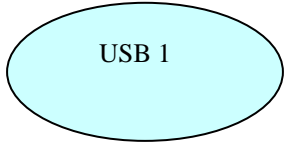
160 – 320
MB/sec

68/80 pin data/cntl
4 pin pwr

Serial SCSI -
SAS

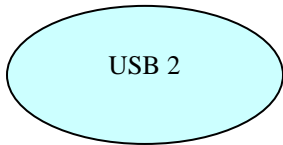
1.5 - 3
Gbit/sec

7 pin cable



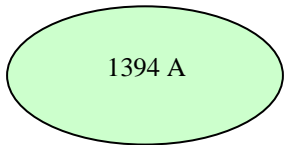
1 Mbit/sec
Full speed – 12 Mbit/sec

4 pin cable



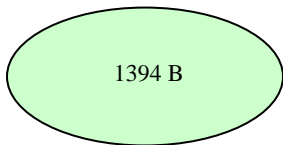
High speed – 480
Mbit/sec

4 pin cable



400
Mbit/sec

4 -6 pin cable



800
Mbit/se

6 - 8 pin cable

The charts illustrate that almost every new and improved interface created a need for new cables and connectors. The complexity of the interfaces requires that the user have more training and expertise.

The charts above do not detail the subtlety of each interface. This will be covered in a follow up whitepaper specifically aimed at that subject.