My digital data storage and backup workflow

Bob Brind-Surch
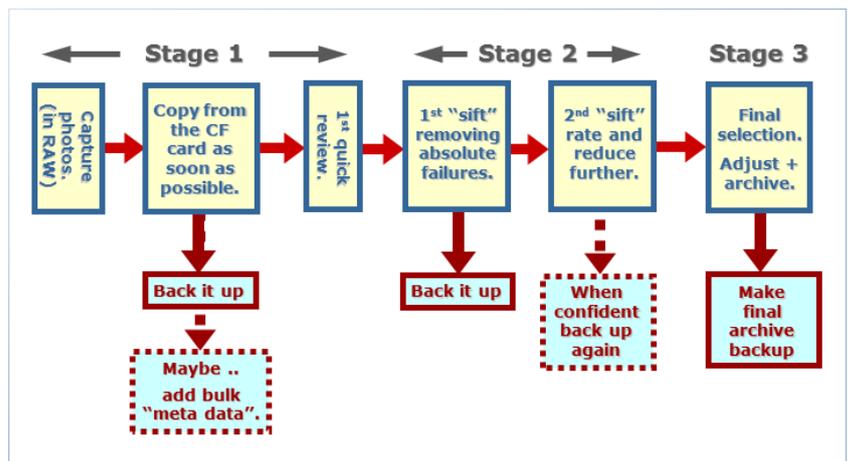**Natures Photos**

Introduction.

A great deal has been written about digital workflow and so I'm not going to attempt to cover the enormous area in detail in this article. Instead I'm going to concentrate on the areas that, from my experience working with photographers, are often omitted; namely image backup and image storage arrangements. These are not the most exciting aspects of workflow but are ones that I consider you omit 'at your peril'. I deliver a talk to camera clubs covering just this and outlining the processes I use for sorting images. As I have already said I don't intend to cover the latter here but will do so in a further article concentrating more on the tool I prefer to use namely Lightroom™.

Data backup is the most important part of your digital workflow.

Most books and articles discussing digital workflow spend a great deal of time concentrating on approaches to sorting, processing and enhancing your images. They will discuss at great length the tools you use such as Adobe Lightroom™, Photoshop™ or the myriad of the other tools currently available. They will examine their strengths, their weaknesses and the authors favourite way of improving your photos. They will talk about plug-ins, filters and all the additional tools you can use but seldom will they discuss what is, from my experience, the most important part of any digital workflow, data storage and backup. You can't enhance or post process an image which has become lost or deleted either accidentally by yourself or due to equipment failure. I once heard a definition of a hard disk as a device perpetually in a state of waiting to fail. Modern hard disks are much better than they used to be but when they fail it's catastrophic, all the data on that disc is invariably lost. If this was your holiday of a lifetime, a friend's wedding or simply just your photographs from a particular outing they are usually irreplaceable. You have probably heard horror stories of professional photographers losing their life's work or wedding photographers losing images of weddings that they have not been able to deliver to their clients yet. It happens, and it usually happens at the worst possible time too!

When it happens data loss is a very painful experience which, unfortunately, many of us will go through it at some point of our lives. In my workshops and talks I spend a great deal of time encouraging people to develop a clear workflow that incorporates good backup strategies to prevent data loss. Over time I have come across many different approaches used and promoted by photographers, some of which are quite frankly downright scary. Sadly, for many, it seems that whilst they are prepared to retell the stories of others failures is not until it happens to them that they understand the real need for sound procedures. For me it's also important that you backup at various stages in the process of sorting and managing your images. Is all very well for example to have a backup of your sorted and 'processed' images but what if you lose those which you've yet to sort? I personally keep my unsorted and sorted images separate, further details later in this article, and thus find it useful to backup each set. My digital workflow, in diagrammatic terms, follows a path similar to that opposite. I divide the process into three stages and backup as frequently as possible. There are more details of each stage and the backup process later in this article.

Stage 1 — Stage 2 — Stage 3
Capture photos. (in RAW) — Copy from the CF card as soon as possible. — 1st quick review. — 1st "sift" removing absolute failures. — 2nd "sift" rate and reduce further. — Final selection. Adjust + archive.
Back it up — Maybe .. add bulk "meta data". — Back it up — When confident back up again — Make final archive backup

When considering a backup strategy there are many excellent articles and indeed whole books discussing the many options available but these often make for tedious reading and their complexity is used by many as an excuse for not developing a comprehensive strategy of their own. What I want to do here is take you through a couple of options I would recommend for developing a good photography data storage and backup workflow. This will include a low-cost and painless process for hobbyists and a much more serious and more costly process for enthusiasts and professionals. As your data and photographic library increases in size and value you will inevitably move from one to the other and this is something you need to consider at the outset so that any change is as seamless as possible. The last thing you want to have to do in any data storage process is change the structure or approach halfway through. That would risk at the very least large amounts of work and that the most give the potential for massive data loss.

The importance of a good clear data storage and backup workflow.

Before you even start to consider your backup strategy it's important that you revisit your current data storage and backup workflow. This has served you well to a greater or lesser degree so far and thus it is important that you examine what works and what doesn't work for you. Learn from what hasn't worked and listen to advice from others but at each stage ask yourself questions such as:

- When you've finished the shoot what do you do with you images?
- How quickly do you get them from the memory card in the camera to some other location ?
- Are they all in a single place or scattered in lots of different locations ?
- Do you have a consistent, repeatable process for managing your images ? Complexity leads to confusion.
- What would happen if the place or storage medium used stored was lost or damaged ?

Every one you talk to will have our own route through this and each will tell you that theirs is the best. They might be but you need to develop your own process as you will have to live and work with it consistently – the key word here being consistently. There is however no need to completely 'reinvent the wheel'. Listen to what others find works for them and select elements that will suit your style, equipment and software. Don't oversimplify it in order to cut corners but don't overcomplicate it either. Complicated processes fail or steps get forgotten in the heat of the moment or when you are tired. Where possible try to automate the process and make it as regular as possible.

My approach and data storage workflow.

For me it's paramount that I get the files copied from the camera cards as quickly as I can after a shoot. If working in the UK and returning home every evening I make this an important part of my *'unpacking my camera equipment'* routine. If working away from home I take a lightweight laptop and a very large external USB hard drive so that I can copy them from the cards to the hard drive every evening. Further when I'm away from home I take a second hard drive and also copy them to that to provide a security backup. If you are as paranoid as I am about data security and are travelling abroad with a friend I usually ask them to take one hard disk back to the UK and I'll take the other 'just in case'.

Key for me is that all my photos are stored in one place and in a structure that is manageable. If you store photos in lots of different places and on lots of different media e.g. hard drives then you will probably end up with more than one copy of some files and eventually get confused and delete the wrong set. The primary reason I find that people use more than one hard disk for storing their data is that the first one becomes full. Some people I know solve this by having a disc per year but I prefer to have one disc for all my photos and thus buy the largest I possibly can at the outset upgrading as it becomes full. For me I also want this disc to be as fast and as reactive as possible so I elect for it to be internal to my PC. I operate on Windows and this is easy, as long as your PC will support it. Currently my

photo drive is 6 TB. If you are operating on a Mac or a laptop then this drive will probably have to be external but in order to make it as fast as possible it really ought to be on a USB3 port or if on a Mac thunderbolt. This is key as it will allow you to locate and view images faster than older USB2 or slower devices. If an internal drive then it needs to be as fast as the PC will support.

Within the hard disk I use a structure of folders, one per shoot, with names that provide some background information and include a date. I'm very well aware that some people, including me, advocate using keywords and cataloguing systems such as those available in Lightroom™ but I like to have an alternative way of finding my images especially those that I have yet to sort and catalogue. To me there is no point in storing (and backing up) images unless you are going to be able to find them again! There are many different ways of storing and organising images but you must decide on one at the outset and rigorously adhere to it. Changing the structure at a later date is massively time-consuming and risks data loss. In addition if you use a program such as Lightroom™ to catalogue your images the whole of the catalogue structure becomes meaningless. The two most frequently used folder naming conventions are chronological and thematic. I use a mixture of both – see opposite.
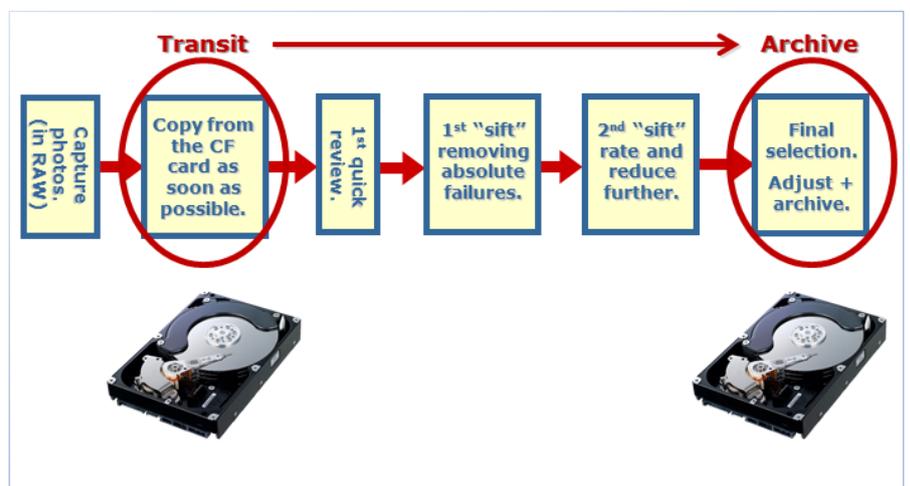

- (179) Lesser Stag beetle - 31st July 2011
- (180) Harvest Mice - Aug 2011
- (181) Wood mice
- (182) Dragonfly centre - Wicken Fen - Aug 27 2011
- (183) Cat Survival Trust - 2nd Sept 2011
- (184) G Norton Village show - 3rd Sept 2011
- (185) Snow Leopard - 11th Sept 2011
- (186) Andover falconry 18th Sept 2011
- (187) Yorkshire Wildlife Park - 3rd Nov 2011

When it comes to storing images chronologically some advocate using a "reverse date notation". This way you ensure that when the folders are sorted alphabetically they fall into date order. E.g. 5th January 2016 would be 2016-01-05 and comes before 2016-01-06 or the 6th January 2016. I personally find this somewhat confusing so number my folders with a three digit number in parentheses e.g. (001). If I subsequently find I have a shoot between for example (003) and (004) I will call it (003N). I choose 'N' rather than 'A' in case I subsequently find I have a shoot between (003) and (003N). If you're organised and copy your files from the cards you been using to shoot with as soon as possible this ought not to happen but it's always worth having an option 'just in case'.

In the field, when using temporary drives, I use a data structure that follows the trip or safari I am on and I also further separate the pictures by camera body and lens – see opposite.
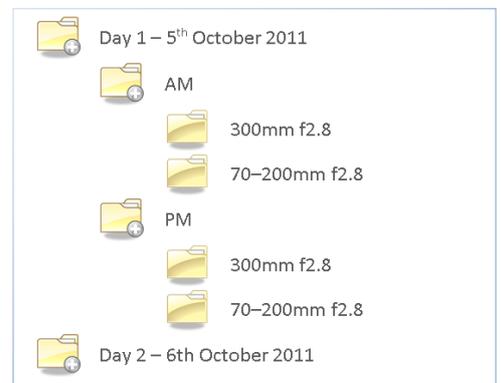
My personal approach is to separate the files I have sorted from those I have yet to work through and store each set on a separate hard drive. For ease of reference I call the first 'Transit' and the second 'Archive'. The 'Archive' drive being the one which I will keep indefinitely. This can present a couple of naming problems if using Lightroom™ catalogues but these are relatively easily resolved. To me the convenience of separating the photographs I've worked on from those I haven't yet makes the little effort required here very worthwhile. Adding this approach to my workflow diagram of earlier it thus looks something like that opposite. The files stay on the 'Transit' hard drive throughout the whole of the sorting, adjusting and cataloguing process. My backup strategy ensures that both the transit and the archive drives are backed

- Day 1 – 5th October 2011
  - AM
    - 300mm f2.8
    - 70–200mm f2.8
  - PM
    - 300mm f2.8
    - 70–200mm f2.8
- Day 2 – 6th October 2011

Transit ➝ Archive

Capture photos. (in RAW) → Copy from the CF card as soon as possible. → 1st quick review. → 1st "sift" removing absolute failures. → 2nd "sift" rate and reduce further. → Final selection. Adjust + archive.

up. If you are using an external drive plugged in perhaps to a USB three or thunderbolt port on a Mac rather than using two separate drives, one for transit and one for archive, you could use two directories on the same drive as long as it is large enough. As you complete the process of sorting, deleting and enhancing your images on the 'transit' drive you can move them to the 'archive' drive.

When choosing hard drives, either external or internal, you would be well advised to buy the largest possible. At the time of writing this article (June 2016) that's 8Tb for internal and 4TB for external if they are USB powered and much larger with separate power supply. When this drive becomes full you would be advised to exchange it for a larger one in order to stick to the principle of keeping all of your images in one place.

My approach to data backup workflow.

Whether we are a hobbyist or professional photographer we usually feel that our data backup workflow process is sound but is it ? Have you ever had the misfortune to have two to test it ?

I once knew a salesman for a large IT company who would say that the most difficult pieces of equipment to sell to customers were backup systems until, that is, they've had a failure after which they will pay almost anything. I guess the unscrupulous IT salesman would look for companies that have experienced catastrophic data failures because that's where they could make the most commission! When I used to run a large education management services business where we were reliant on data and data backups I used to insist that my IT staff rehearsed a data recovery process on a frequent basis. To me it was as important as practising the fire drill. I'm not going to advocate that we attempt to do this for our images although if you're a professional relying on them it might be worth considering. At the very least even if you are a hobbyist then it's wise to check to see that you can read and recover individual images or even sets of images from your backup on a regular basis.

The level or degree of attention you pay to your backup and the systems you use depends on how important your images are but the basic rule is to have three copies of any important file, a primary copy and two backups. Why two backups ? Well if ever you had to restore from backup and found a problem you will know the value in having a second copy. Furthermore it's very helpful to store that second copy somewhere else. Clearly having two copies and managing a process like this is time-consuming and costly so depending on the value of your photographs you may decide to adopt a single backup copy policy. I will discuss both and potential ways of storing files offsite later in this article.

A simple approach for the hobbyist photographer.

The 'bare minimum' approach to backup is relatively simple. I find many photographers don't adopt a secure backup because it's complex and costly and because a 'professional' has a bored and confused them with detail and complexity at some point in the past. What I will recommend however for the simple or the more complex process is that you store your backups on a separate medium, never ever rely on your camera cards as a backup. I've seen too many occasions where they failed or have been reused in haste resulting in irreplaceable images being lost. With respect to separate media for storing your backups there are many options - the following are just but a few:

- DVD's and rewritable disks
- Memory "sticks"
- Cloud storage
- External or internal hard drives
- RAID arrays systems (in my opinion not really necessary for the hobbyist)

Of these I would discount immediately DVD's and rewritable discs. They were a viable option some time ago when hard drives were slow and poor as they were relatively cheap and as such are described in many books on photography. They are however limited in capacity and thus you need a large number of them. Furthermore the technology used to allow you to write data to these drives is such that they can be destroyed for example by light and anyway simply deteriorate over time. I have thousands of 35mm colour slides in storage boxes which if I am honest are now of no value whatsoever because they've deteriorated so much. We were once advised to mount our transparencies between glass and I did that meticulously only to find out that advice was poor and my transparencies now have fungus growth, Newton's rings and the chemical material on the surface is sticking to the glass thus making them impossible to recover or restore even if removed from the frame. The reality with backups is that we only use them when we need to and I know many people who have returned to their DVD's only to find the surface coating flaking off and them unreadable.

Memory sticks are an option but again are limited in size and suffer in the same way as camera cards in that they of deteriorate over time potentially losing data.

Cloud storage is an option often quoted these days and is becoming relatively cheap. Many people use it and rely on it but before you do in my view you need to consider a few things:

- the process of uploading lots of images to an on-line cloud storage solution takes a very long time especially with the slow speed of broadband in the UK.
- just as it's a slow process to upload images should you need to recover or restore them it can take literally weeks to download images, that's no good if they are crucial images for a club competition or a client.
- there is a risk inherent in all such external storage systems that the provider goes out of business in which case all your valuable data is lost for ever.
- I have heard examples where cheap or free cloud-based storage companies claim in the small print that they have the right to use any of your images without consulting you first. Many of us, me included, seldom read such small print and many of us may think that our images aren't that valuable that is until you see your picture appearing on a bill holding advertising a product and you have no recourse whatsoever to the company whose used it or even sold it on!

RAID array solutions are excellent technical options but in my view costly and with a few exceptions complex to set up so really not necessary for the hobbyist hobbyist.

This leaves us with external hard drives which to me are the best and indeed excellent solution for even the most enthusiastic hobbyist. These are relatively cheap, very reliable and come in many forms. Thee most frequently used are those mounted in cases that plug into the USB port on your computer Windows or Mac. If you are using Windows it's worth buying a USB3 device and if using a Mac then Thunderbolt™, as these are much faster. You can purchase these in sizes up to 4TB (at the time of writing June 2016) if self powered and much larger if they are powered separately. As with your primary drive you would be well advised to buy the largest possible and that you can afford but when that eventually grows too small to exchange it for a single larger capacity version to ensure you have all of your backup on a single drive. If you adopt the approach I use of a 'transit' and 'archive' drive then these could simply be two separate directories on the backup hard drive even if that is not as they are stored on your PC.

External USB Powered Hard Drive

Externally USB Powered Hard Drive

The problem with external hard drive for backup is of course that if you're also using external hard drives for your primary data storage then

you need at least two USB sockets on your PC or thunderbolt sockets on a Mac. This should not be a problem unless you're working on a laptop where USB ports are limited. If so then you may have to invest in a small USB hub which when plugged into the USB port on your computer presents you with a number of additional ports.

One very cost-effective way of working with external drives is to buy what are called 'bare' 3.5 inch hard drives i.e. the sort you install inside a PC. To allow these to work and be accessible to your computer you need a docking station which is permanently plugged into a USB port. I used to use this approach and know of the many people who still do very successfully. The docking station sits on your desk permanently plugged into the computer and all you need to do is slot the relevant hard drive into the slot on the top and just as with a flash drive it is accessible to your PC as if it were internal. Furthermore if you're using a USB3 docking station the performance is very similar to that of an internal drive. The value is that 'bare' drives have considerably more capacity and are very much cheaper than enclosed USB hard drives.

The example in the upper photograph opposite was available at the time of writing this article for £20 on Amazon. If adopting this route I would however suggest spending a little bit more and going for a version which will support up to 2 'bare' hard drives. I currently still use the one in the lower photograph opposite which again at the time of writing was £36 on Amazon.

So called 'bare' drives do need more careful handling as they clearly don't have a protective case around them. In truth however so long as you are careful not to damage the side of the drive with the circuit board and store it carefully in a protective case there won't be any problem. The cheapest and easiest to use are plastic ones flip top ones which are available on the Internet for under £5.

The process of backing up.

There are many approaches to performing a backup. The simplest and cheapest is to manually copy files from your primary hard drive to the backup drive. However if you adopt this approach it is sometimes difficult to remember which files you've changed and which you haven't. A more reliable and automated approach is to use a dedicated piece of software. There are many such software applications available and most can be set to run at a pre-set time for example overnight and a copy only files that are altered. As with anything in the IT world there are those who favour one piece of software over another but the one that I've found works most reliably for me is a product called 'Syncback Pro' from a company called '2brightsparks.com'. At the time of writing this article (June 2016) this cost about £38 which to me is a small price to pay for its convenience and the knowledge that your backups will always be performed as long, that is, as your computer is turned on. An alternative approach for Mac users is to use a piece of software called 'Time Machine' which does all of this automatically for you continuously. Whilst there are products similar to Time Machine available for the PC I have personally never use them so can't comment on them.

Earlier in this article I mentioned the merit in having two separate backups and keeping one 'off-site' or away from home. This may at first seem complex but with a little thought and planning is relatively simple and well worth the effort. If you elect to backup to your second backup drive perhaps once a week and store it at a friend's house then it's simply a case of arranging for them to visit perhaps once a week or once a month for a welcome coffee or chat about photography. Whilst this second backup may at most be a week, a month or fortnight out of date it'll be really valuable if for

example your first backup becomes corrupted or the worst happens and it is stolen or otherwise lost from your house. In addition should if you delete a file on your primary computer accidentally it will of course be deleted from the backup when it runs automatically overnight night so a simple visit to your friends to recover the file from the second backup will allow you to restore it. There are more complex options which allow for what are called 'incremental backups' where files that have been deleted or changed are retained on the backup for periods of time determined at set up. In my experience this is however an unnecessary complication for most hobbyists. It should be noted however that the 'Time Machine' software approach from Apple does exactly this automatically albeit requiring considerably larger hard disks to retain all previous copies of data as well the current.

A more complex approach for the professional or enthusiastic amateur photographer.

If you are a serious amateur or a pro with a large collection of photos, creating a solid backup strategy is a key part of your professional responsibility and you should take it very seriously. If you are a professional photographer it is simply a part of your job! If you are being paid to take photographs you need to be able to produce the results and nothing will destroy your reputation faster than having to admit to a client that you've lost the photographs you've taken for their commission. Furthermore as well as handling large amounts of data you need to be a do it quickly so you need fast drives and you really do need to carefully consider having data in more than one location. Much of the process described above for data storage and management and data backup are the same it's just the media you use, its complexity and how you deploy it that needs to change. Solutions other than hard disk are really not worth considering and I personally would never look at remote cloud storage solutions because:

1. they are far too slow
2. if the supplier should cease trading you have no access to your data
3. there is always the question of data and image ownership, it would be a disaster if some of the images are client asked you to shoot was suddenly claimed by the cloud storage provider company

When considering hard disks as a solution for professional data storage and backup you need to choose your discs carefully. Not all discs are equal some are faster than others and some much more reliable. Whatever solution you choose you need to remember the old adage that every hard disk is perpetually in a state of 'preparing to fail' and one day it will and always at a critical moment when any downtime for the business is costly in financial or reputational terms.

Thankfully there are very many solutions to reducing the risk and principle amongst these are using what are called RAID (redundant arrays of independent disks). This is an arrangement where a number of separate hard drives are installed in an enclosure either internal or external to your PC and are configured in such a way that the data is written across all of them rather than to one at a time. Various different configurations can be used but in principle you gain the benefit of having more data storage and at the same time resilience such that if one or more drives fail the data is still available. RAID is sometimes described as a backup solution, it most certainly is not. What it will do however is allow you to continue working even with one or more failed hard drives but it isn't an independent backed up copy of your data. All it is doing is saving you the problems associated with loss of one or more drives. A serious professional would have a RAID array as their primary data store and another RAID array as a backup data store and maybe a third as their secondary backup. This is costly and most RAID solutions are complex to set up without some detailed IT knowledge. You do however have to consider this cost against the cost of loss of business and reputation a corrupted data problem would give you.

Fortunately there are now on the market some forms of RAID solution which require very limited IT knowledge to configure. The first was made by a company called DROBO™ but there are others now available too. The problem with a conventional RAID solution is that it needs all drives to be the same type, manufacturer and size furthermore they have to be configured in a particular

way. With DROBO™ you can use any 3.5 inch drive in each of its drive bays. They can all be different sizes, different manufacturers and even different specifications. The front of the device has a series of lights which if green means a disc is functioning correctly, amber means there is potentially a problem and red means a disk has failed. As with any RAID device one or (if configured to permit it) two hard drives can fail and your data is still accessible and usable. To replace a drive you simply take off the front cover, press a button to eject the old drive, slot a new one in - preferably of a higher capacity - and the device will sort itself out indicating the process with all the lights flashing orange. During this process your data is still available for use. The DROBO™ and indeed any RAID device simply plugs into your network e.g. your router and is visible on your computer.

DROBO™ make a number of different devices from:

- their smallest – DROBO™ mini which is almost portable and takes up to four 2.5 inch drives with a data storage capacity with redundancy of a maximum of 5.44 TB (at the time of writing – June 2016). I know of one highly successful professional photographer who uses a DROBO™ mini as his primary data drive and carries it around with him when on location thus having his data available wherever he is and on whatever computer system he has access to.
- Their largest is a massive device taking up to eight 3.5 inch drives and thus a total possible capacity of 49.93TB if configured to allow for one hard disk failure or 43.21TB if configured to allow for two hard disk failures.

This latter solution is very expensive and probably well outside requirements of a normal even professional photographer. Either solution is more expensive than some of the conventional RAID solutions but once you factored in the cost of hiring and retaining an IT professional to configure the more conventional RAID solution it can be a good deal cheaper. Further expanding the DROBO™ as your data storage requirements increase is simply a case of replacing an individual drive rather than the much more costly process of replacing all as required in a more conventional RAID solution.

If you consider a RAID device of any sort as your primary data store you need 1st to examine issues of performance. Some conventional and indeed earlier RAID and DROBO™ solutions were notoriously slow making access to data to view, sort or enhance photographs painfully slow and in business terms very costly. Modern RAID and DROBO™ solutions use an internal solid-state drive to speed the process up.

The approach I adopt.

For me, my photographs and indeed all my business data is extremely important and I cannot risk being without it for a moment. I therefore have adopted a relatively complex and costly solution which may be more than many of you would need but I will outline below as an indication.

Firstly I store all of my photographs and data on a series of internal hard drives inside my PC enclosure. These are not RAID as I didn't consider that necessary given I make frequent backups. Not using a RAID drive as my primary drive also gives me far better performance and when managing photos in Lightroom™ which is key to me. I use a powerful PC and the 'Transit' drive I use (see above) is a solid state drive or SSD drive. I also arrange that the temporary catalogue I create when initially sorting my photos is stored on this same SSD drive. Here isn't the place for me to discuss Lightroom™ or techniques to manage this - I am planning other briefings and a training workshop on this in the future.

To gain the maximum storage space I store my Archive photo library (see above) on a separate large internal hard drive (currently 6TB). My 1st Backup drive is a DROBO™ to which all my data and images are automatically backed up every 24 hours at midnight using Syncbackpro. I am thus never more than a day behind if I should have a problem. I don't use a Time Machine PC equivalent as it would take a great deal of disk space and I don't really need it. I make a secondary backup to a remote DROBO™ once a week on a particular day again automatically using Syncbackpro.