Backup and data recovery within schools is changing due to an ever-expanding amount of data. Coupled with this, schools are moving towards a model of ‘business continuity’, where networks always need to be available for the school to function and downtime is not tolerated.

This white paper proposes backup strategies that could be employed by schools; the hardware and software that can be used; and the agreements within schools around data protection.
The document is split into the following sections:

1. **Introduction**
   A summary of the problem: what is causing the change and the impact on traditional backup methods

2. **Understand your commitments**
   Highlights some of the obligations that you may have to your school

3. **Understand the nature of your data**
   Introduces the concept of separate backup strategies for different types of data

4. **Design and implement your backup solution**
   What to look out for when selecting backup software and hardware

5. **Maintain, monitor and TEST!**
   How to give yourself confidence that your strategy is working

6. **Develop a disaster recovery plan**
   How will you get your data back in the event of a disaster?

7. **Reviewing your backup strategy**
   Keeping your strategy current and up to date
Data on a school's network is increasing at a frenetic rate. With the extended use of the network throughout the school, larger numbers of people are storing more important data than ever before. The growing use of multimedia across the network and the move to community access, storage and backup requirements are set increase even faster.

This all leads to additional pressure on the backup solutions and the people responsible for running them.

The traditional strategy of a single tape drive in a single server is becoming increasingly difficult to implement, as improvement in the capacity of single tapes is rapidly being outpaced by the growth of data. This is driving the need to develop more creative solutions, that can not only back up large amounts of data, but can do so quickly.

Data held on a school's network is increasing in size at a rate of between 15 and 30% per year. * Every day at least 5% of that data is changing. In a single backup, an average secondary school is backing up a Terabyte of data.

In three years’ time, that could be 2.2 TB
In five years that, could be 3.7 TB

A true test of backup success is the restore. As the importance of data increases, the need to assure that it is safely backed up is paramount. As such, a ‘backup’ policy should be termed a ‘restore’ policy!

Before investigating the technology associated with backup, the most important thing is to understand the purpose of the backup and to whom it is important.

* From research conducted by RM in December 2009
2. Understand your commitments

In a school, the owner of the backup process is usually the Network Manager; the owners of the data are typically the senior leadership team, the staff and the learners.

It is unusual for all parties to have the same expectations for the backup and possible disaster recovery of the network. The Network Manager is inevitably familiar with the day-today problems of managing large amounts of data across a complex network, whilst senior leadership may not be familiar with the practical limitations of the technology available. Typically a SLA would define all aspects of the IT department’s interaction with the school at large, allowing for both the IT team and its clients, be it the school’s management team, the teaching staff or the learners, to understand and agree the levels of support available.

Typically a SLA would define all aspects of the IT department’s interaction with the school at large, allowing for both the IT team and its clients, be it the school’s management team, the teaching staff or the learners, to understand and agree the levels of support available.

A SLA need not be long or complicated, but it should form the basis of agreed level of service and commitment. This will allow all parties to clearly understand each other’s responsibilities and help to avoid conflict.

The generation of a complete SLA is out of the scope of this white paper. However, by way of an example, a SLA may well cover the following areas:

- Scope of agreement – this might state to whom the agreement will apply and the duration of the agreement (it might be limited to the duration of a school year to allow regular review and update of the terms of the SLA).
- Scope of support – this might define the systems and software for which the IT team are to be responsible.
- Define support events – establish an agreed definition of events, from routine maintenance events through to events that seriously disrupt system operation.
- Response times – this would establish the time taken to respond to different levels of event.
- Definition of support available – this could define the actual responsibilities of each of the parties in relation to the different elements of the IT infrastructure. For example, where a printer is located within a department, the replacement of toner cartridges and addition of paper may be designated as a customer responsibility, but the clearing of paper jams may be designated as an IT responsibility.
- Regular system maintenance windows – this would allow for the establishment of regular, mutually agreeable windows of system unavailability.
- Procurement and testing of hardware & software – this may well define the process for teachers and students to request the restoration of data
- Monitor and report on the performance against targets – a key part of having a SLA is being able to measure performance against your declared targets and where necessary, review and make changes.

The Becta developed Framework for ICT Technical Support (FITS) is a good source of information regarding best practice in school. The most up-to-date information is located at:
http://www.thefitsfoundation.org/

Your school may not have an IT Service Level Agreement (SLA). If this is the case, it is highly recommended that one is implemented.

Only 27% of primary schools and 33% of secondary schools have an agreed and documented backup strategy.*

* From research conducted by RM in December 2009
3. Understand the nature of your data

Only 12% of secondary schools rated media files as ‘critical’ or ‘moderately important’ to quickly restore in the event of a disaster.*

Rather than treat all your data as the ‘same’, it is good practice to categorise it. Good reasons for doing this include:

- as your data grows, you can ensure that the backups are completed in the most efficient way.
- when compiling a SLA, it will allow each category to have a level of service applied to it.
- when buying a backup solution, it will allow for accurate sizing and costing.
- when storing data, it will allow different levels of data security and retention to be applied.

You may want to consider limiting yourself to a small number of categories to enable easy management.

For each category that you decide upon you will need to decide:

**The Recovery Point Objective (RPO) of the data**

RPO is a term used in industry. It essentially means how recent the backup of a file or data is. It is usually measured in hours - for instance, if the data has a five hour RPO, it means that a maximum of five hours of data is lost in the event or a file being deleted or a server recovery.

As a rule of thumb, the shorter the RPO, the more costly the backup solution. Whilst a short RPO is desirable, it’s important to consider that costs can escalate dramatically as the RPO time reduces. Conversely, deciding data can have a longer RPO could help reduce the costs by reducing the amount of disk or tape storage required to store the backups.

The standard nightly backup in a school allows for an RPO of 24 hours or one day. In the majority of the cases, this is acceptable to the data owners.

**The Recovery Time Objective (RTO) of the data**

RTO is another industry term and relates to the time taken to recover any lost data from your backup solution. Again, the shorter the time, the more expensive the backup solution is likely to be. Having a different RTO for different classifications of data will allow you concentrate on returning the most important data back into use as fast as possible, with the less important data following on behind.

For the majority of schools, large media files tend to be the least important form of data, whilst schools’ management information and server operating systems tend to be the most important.

**How long you want to retain the data?**

Clearly understanding how long data is to be kept is important and will have a big impact on the sizing of your backup solution. Defining different retention periods for different types of data will allow the data owners to understand how long data is retained and allow you to optimise the use of backup resources.

Some data may need to be kept for a long period of time. Becta indicates that schools’ financial data might be stored for six years and pupil and staff data for seven. The need for this level of retention might indicate that an archiving solution would be advisable.

**How safely do you want to store the data?**

School management data containing learners’ names and addresses is understandably important. As such, this would potentially require special treatment, probably encryption and potentially the logging and tracing of media containing this data.

Removing backup media from the school could be construed as ‘external access and transfer’ and might need special treatment under the Data Protection Act.

The latest information on data protection and security can be found by searching for Data handling security guidance for schools on the Becta schools website: http://schools.becta.org.uk/

Once you have your list of categories, you may well decide, in the first instance, that the majority of them can be treated in the same manner. As time progresses you may find that your backup takes up more and more resource, you can then subsequently assign a lower level of backup service to different categories without the need to reengineer your SLA or your backup schedule.

* From research conducted by RM in December 2009
2. Understand your commitments

Category 1 Data –
Information that changes on a regular basis.
One day RPO – aim to restore data no older than 24 hours.
* Four hour RTO – aim to restore data in less than four hours.
* One month retention period.
* Data is unencrypted.

This might include the following type of data:
* Network operating system and configuration data.
* Staff home directory areas
  • Shared areas.
  • Your email server

This would give you good ‘granularity’ of restore, allowing you to restore any file or data that has changed in the last month on a day-by-day basis.

Category 2 Data –
Information that changes on a regular basis and has to be retained.
* One day RPO – aim to restore data no older than 24 hours.
* Four hour RTO – aim to restore data in less than four hours
* Three month retention period
* Data is encrypted.

This might include the following type of data:
* Your email server
* Your school management information.

This again would give you good ‘granularity’ of restore, allowing you to restore any file or data that has changed in the last month on a day-by-day basis and providing for the storage of data for up to three months. Also as learnerspecific data is being stored, the data is encrypted to prevent unauthorised use.

Category 3 Data –
Information that changes on a regular basis but is backed up twice a week
* Three day RPO – aim to restore data no older than 72 hours.
* One day RTO – aim to restore data in less than 24 hours
* Three month retention period
* Data is unencrypted.

This might include the following type of data:
* Your learner data

The bulk of data may well be learner-generated and, as such, may consume a large amount of backup resource. You may decide that a three day RPO may be acceptable for the majority of your learners.

Category 4 Data –
Large data files such as multimedia video clips.
* Five day RPO – aim to restore data no older than 120 hours.
* Two day RTO – aim to restore data in less than 48 hours
* One month retention period
* Data is unencrypted.

This might include the following type of data:
* Pupil generated video and projects.

Video and other multimedia files may well present the largest challenge to your solution in both capacity and time to backup. As such, you decide to not back up this data at all, or agree periods that do so, without impacting the backup of higher category data or impacting network performance.

* From research conducted by RM in December 2009
Once you understand the nature of your backup, you can design a solution that meets the commitments made as part of your SLA.

In any backup solution, you need to consider both the backup software and backup storage. These are often considered as one and the same; however, it is important to ensure that the features of both are considered.

**Backup software**

When selecting backup software, it is important to look for the following features:

- **Incremental & differential backup support** – the ability to support partial backups is important. A full backup of a set of data is usually followed by either an incremental or a differential backup:
  - A differential backup backs up data added or changed since the last full backup.
  - An incremental backup backs up data added or changed since the last full or incremental backup.

- **Flexible data selection** – the ability to select data based on your chosen criteria is important. This might mean the ability to select data based on:
  - Its location.
  - The type of data (video clip, document or spreadsheet).
  - The size of the data.
  - The age of the data.

- **Ability to span across media** – if you are planning to use removable media, the ability to span across media is important.

Where the amount of data exceeds the individual tape or cartridge, the backup software must have the ability to seamlessly continue the backup on another tape or cartridge either utilising a media autoloader or through manual media swap.

- **Data encryption** – if your data is likely to contain learnerspecific information, the ability to encrypt data is desirable.

- **Centralised backup and management** – when backing up more than three servers, it may be more cost-effective and allow for easier management to move to a centralised backup storage suite. This allows the backup storage to be centralised and shared amongst multiple servers.

**Backup storage**

There are a variety of different storage media for your backup data, some based on traditional tape and others based upon more modern concepts such as hard disk drive. Whatever method you choose, it is important to consider the following principles when storing your backups:

- **Always ensure that your backup data is removable.** Backup destinations located on the network, whether disk staging or some form of network storage, are always at risk of hardware failure, disaster or theft. Ensuring that you can physically remove data from the network minimises the chance that it might be irretrievably lost.

- **Always locate your backup data away from the original source.** The safe and organised storage of your backup media is critical to the success of your strategy.
  - The backup media should be stored in a separate physical location to the backup solution and servers. If your school has a separate building, then a secure room or cupboard with a purpose built media storage safe would be ideal.
  - If you have another local school that is willing to participate, you could arrange to store each other’s backup media.

- **Always store media in accordance to the manufacturer’s guidelines.** Some types of media can be susceptible to environmental factors, such as temperature and magnetic fields. Incorrect storage can lead to a total loss of data, usually discovered at the worst time.

- **Always identify and log removable media.** There is nothing worse than a stack of unlabelled tapes when trying to locate an important piece of data. Aim to label your tapes with the following information:
  - Backup set name and tape number
  - Date of first use. This allows you to make tapes obsolete when they reach the end of their working life.

Avoid using information that could identify your school to any external party that might come into possession of the tape.

Create a log of the tapes to ensure that its location is recorded in the event that it is moved around or off your site. This is particularly important if you are planning to keep a long term copy of your data.
Types of backup storage

Tape backup – The traditional method of storing backup data is using a tape unit located in a server.

Tapes come in various technologies and capacities. However, it is recommended that one of the following technologies is used:

- **LTO** – the LTO standard is widely adopted in both industry and education. It is available in five capacities:

<table>
<thead>
<tr>
<th>LTO version</th>
<th>Uncompressed/native capacity</th>
</tr>
</thead>
<tbody>
<tr>
<td>LTO-1</td>
<td>100GB</td>
</tr>
<tr>
<td>LTO-2</td>
<td>200GB</td>
</tr>
<tr>
<td>LTO-3</td>
<td>400GB</td>
</tr>
<tr>
<td>LTO-4</td>
<td>800GB</td>
</tr>
<tr>
<td>LTO-5</td>
<td>1500GB</td>
</tr>
</tbody>
</table>

LTO versions 1 and 2 are largely superseded and are not recommended.

LTO version 5 is recently released and, as such, is relatively expensive in both terms of the drive unit and the tape media.

LTO versions 3 and 4 currently offer the best price versus capacity in terms of the drive unit and tape media cost.

LTO tape drives have the capacity to compress data when writing to tape. The capacities of tapes are often listed with their compressed capacities which assume a 2:1 compression ratio. This ratio is usually based on the compression of text files which are generally much more compressible than other types of data. The data in an educational environment is not compressible to the same extent, so when sizing your backups it is advisable to use a ratio of 1.3:1.

It is important to note that there is limited compatibility between different levels. Generally a LTO tape drive can read and write a previous generation tape and can read a tape two generations down. For example, a LTO-4 drive can read and write a LTO-3 tape and read a LTO-2 tape.

A tape also has a limited number of uses and will require regular replacement to ensure that data is reliably backed up. The expected life of a LTO tape is 200 complete backups and tapes should be retired after that to minimise the chance of data loss. Provided the tape is not overused, it is expected to last 30 years.

LTO tape drives will also require cleaning with a special cleaning tape and this should be carried out at least once a month.

- **RDX** – the RDX cartridge actually contains a hard disk, but is removable in the same way as a normal tape. The cartridge is shock resistant, withstanding a metre drop onto concrete.

Unlike LTO, the cartridges come in 160, 320, 500 and 640GB sizes. The RDX drive can accommodate all sizes of cartridge, together with any future sizes without any requirement to change drive. It is expected that the cartridge capacity will increase in line with hard disk capacity.

The cartridge is designed to be used 5,000 times which typically equates to over ten years of use in an educational environment.

Tape autoloader/library solutions – Rather than situate the tape drive in a server, a tape autoloader or library solution situates the tape drive in an external enclosure. Generally, autoloaders and libraries use tapes and tape drives – the most popular and recommended format being LTO.

An autoloader is a device which houses a single tape drive, multiple cartridges and the robotics to automatically load and unload the cartridges in and out of the drive. Typically an autoloader contains up to 16 cartridges.

A library is a device which houses multiple tape drives, multiple cartridges and the robotics to automatically load and unload the cartridges in and out of the drives. Libraries typically support larger numbers of cartridges, sometimes numbering into the hundreds.

The prime purpose of an autoloader solution is to overcome the issues related to swapping cartridges. It allows the drive to carry on backing up data automatically once the first and subsequent cartridges become full. At the completion of the backup cycle, the used cartridges are then removed and replaced with cartridges ready for the next cycle. The purpose of an autoloader is not to store backup data, as the chances of data loss are greater in the event of the loss of a server room.

The purpose of an autoloader is not to store backup data, as the chances of data loss are greater in the event of the loss of a server room. Autoloaders and libraries are generally connected to a central backup server, which then backs up other servers using a remote backup process. This allows effective use of the backup capacity of the tape. Generally, this means that the only one server can back up to the device at a time, potentially extending the time taken to backup the network.

Disk staging solutions – Disk staging solutions utilise an array of hard disks as a backup destination. This increases the performance of the backup by utilising the speed of writing to disk and the ability to backup up multiple servers simultaneously. This is also known as Virtual Tape Library (VTL) technology, as the hard drives are often presented to the servers as ‘virtual’ tape drives.

In some solutions, this maybe the only backup destination and, as such, might present a single point of failure in the event of a hardware problem.

It is more usual to combine a disk staging solution with an autoloader to enable the data to be moved or copied to removable cartridge. This can be completed during the day, as it does not impact the performance of the servers and ensures that a separate copy of data can be stored elsewhere.
5. Maintain, monitor and TEST!

Selecting and implementing your backup solution is only half the battle. The solution will need to stand the test of time. To enable this, you will need to consider implementing policies, processes and logs to ensure everything runs smoothly. These will:

- make it easier to measure the success of your backup in relation to your SLA commitments.
- make it easier to train new staff.
- help with managing the time of both yourself and any staff you may have.
- demonstrate a high level of competence to your Senior Leadership Team, LA and external inspection authorities.

**A schedule for everything** – knowing what to backup when is critical to backup success. A clearly communicated schedule minimises the chances of missing a backup.

**Roles and responsibilities** – if there are multiple people responsible for the performance of your backups, you will need to clearly define each person’s responsibilities. When deciding this, it is also a good idea to split some tasks to ensure that processes are double checked. For example, the swapping of cartridges may be the responsibility of one person, whilst the reviewing of logs and media movement may be the responsibility of someone else. This could help reduce human error.

**A process for everything** – the performance of backup can be a complex series of tasks. Develop a document that details all the steps required for each of the processes that are required.

Example processes that might be considered:

- **The daily maintenance process** which could detail the steps required to:
  - swap cartridges
  - move media to storage
  - inspect backup logs and action on any failures
  - complete a system log

- **The weekly maintenance process** which could detail the steps required to:
  - swap weekly backup cartridges
  - lean the tape drive
  - remove media off site
  - inspect backup logs and attend to any failures

- **The monthly maintenance process**. This could detail the steps required to:
  - remove obsolete media
  - perform test restores of data from backup
  - perform monthly archives (if required)
  - other maintenance activities, such as defragmenting disk storage in backup staging servers.

- **A process for the recovery of deleted data**. This could cover the steps required for the data owner and the IT team to perform to recover data from backup.

- **A process to test data recovery**. This process could cover the steps required to test that the data is being kept in accordance with the SLA. For example, do you actually have copies of the data from three months ago? If you have the capability, perform a complete disaster recovery of a server.

- **A process for the adding or removing data from the backup schedule**.

**A log** - All backup activities should be logged to enable monitoring of backup success and failures. This log could be part of a wider network log or kept separate. A log might well contain:

- the process performed.
- when it was performed
- who performed it.
- if it was successful or not.
6. Develop a disaster recovery plan

In the unlikely event that you have a complete disaster, you must ensure that you are in the best possible position to recover your network.

A disaster recovery plan may not just be a sequence of technical tasks performed by IT staff to resolve a short-term problem, but it should also consider the following:

• communicating with the school a large, keeping people informed of progress and expected time for return to service. This might include the times for the different categories of data that you have to restore. For instance, this might mean the return of basic system functionality within four hours and the restoration of large multimedia files within two days.

• Long-term network unavailability. This is often termed ‘business continuity’, and concerns what happens to critical services if you cannot quickly return a system to full use.

Practical elements that should be considered as part of a disaster recovery plan are:

Hardware maintenance contacts – a full list of the various agents responsible for maintaining your equipment, together with any requirements they may have for logging a call. This might include serial numbers, for instance.

Operating system and server driver media – a full selection of the required operating installation system media and any required drivers for the server hardware should be duplicated.

Operating system installation process – the process required to install the servers operating system and backup software should be copied and stored.

Server information manifest – make a copy of the server configuration information. This should include:

• server name
• server IP information
• server disk partition sizes
• server SAN connection information (if a SAN is installed)

All of the above should be duplicated at least twice and combined together in a ‘kit’. This would minimise time required to locate the required information.
7. Reviewing your strategy

Having successfully implemented all of the above, a yearly review should be conducted. As part of the review, you may well consider:

- Testing the SLA is being met. If not, changes to the backup process or to the SLA should be agreed.
- Reviewing hardware and software maintenance required for your backup equipment.
- Reviewing that the processes are still relevant, and updating them if required.
- Formulating a purchasing plan for any new backup media requirements in the coming year.
- Reviewing any process problem or failures during the year and if any necessary action is required.

“The server platform is industry-standard hardware and provides a stable and reliable service to the clients. It’s also flexible, allowing bespoke configuration that the schools require.”
—Andrew Arnold, systems solutions team leader, Staffordshire Learning Technologies

“Everything is faster, stability is better, there’s less down-time and we are less stressed, everybody wins. We finally feel like we aren’t just coping, we are taking steps forward and progressing. There are literally no downsides to owning a SAN (Storage Area Network).”
—Alistair Hetherington, network manager, Longbenton Community College, Newcastle Upon Tyne

“We have a service level agreement with our staff that says we will retrieve lost or deleted files within an hour. Our disk-to-disk-to-tape solution means we can restore files within 5-10 minutes.”
—Darren Williams, network manager, The Abbey School, Reading

How much money could you save with better infrastructure in place? How much time could smarter systems save for all your staff and students? How long before your network collapses under the load of expansion?
8. In conclusion

The data contained on a schools network is fast becoming one of its most valuable assets. Hopefully this white paper has helped stimulate thought and debate on the disaster recovery strategies deployed in your school.

For more resources and advice about backup, storage or infrastructure visit www.rm.com/infrastructure

Discuss your backup requirements with one of RM’s experts
call 08450 700300
or email infrastructure@rm.com