

White Paper

The Online Backup Advantage

Delivery Superior Performance, Security, and Reliability

Abstract

This White Paper discusses how online backup, utilising advanced disk-to-disk technology, provides a superior alternative to tape-based solutions. Traditional tape-based backup is a cumbersome, error-prone process that drains IT resources. Online Backup delivers a better way to protect business-critical data, taking advantage of existing networks to quickly and efficiently back up all servers, desktops, and notebooks distributed throughout an organisation.

Table of Contents

IT CHALLENGES: IT'S UGLY OUT THERE	3
The problem with tape	4
NOW IS THE TIME FOR ONLINE BACKUP	6
The financial advantages of online backup	6
Factors for evaluating online backup vendors	10
GBData: CREATING VALUE THROUGH ONLINE BACKUP AND RECOVERY..	11
Advanced Compression and Transmission Technologies.....	13
Conclusion.....	15

IT CHALLENGES: IT'S UGLY OUT THERE

Businesses today face unrelenting data growth. IT staff struggle to ensure data availability under tight fiscal constraints. The unreliability and intensive management of traditional tape-based backup has forced system administrators to create manual, error-prone processes in order to protect their data. Consequently, other critical objectives, such as supporting distributed environments and complying with regulatory requirements, are not being adequately met.

External forces are adding to the backup-related challenges that organisations face today. Heightened awareness around business continuity and regulatory compliance has caused businesses to increase spending on their data protection activities. In fact, companies spent an estimated £2.9 billion to comply with Sarbanes-Oxley alone in 2004.¹ While compliance and headline-grabbing disasters are often credited for increased focus on business continuity, the truth remains that IT staff must solve difficult everyday issues if they are going to maintain optimum data availability. (See Table 1)

Table 1: Today's Backup Problems

Problem	Cause
Shrinking backup windows	Proliferation of data-intensive, high availability applications now required for around-the-clock availability.
Rapid data growth	Data growth continues to rapidly climb. It is reality, not a trend. These data-intensive applications are generally the culprits.
Lack of central control over distributed systems	IT staff managing off-site backups frequently must trust untrained co-workers to conduct backup activities such as swapping out and replacing tape media. Managing redundant hardware and software components in multiple offices also adds to the complexity.
Increasing rate of recovery failure	Business continuity is jeopardised by viruses, accidental data deletion, and data corruption. As systems grow more complex, the inability to adequately recover an organisation's business-critical data increases.
Human error	Accidentally deleting an email or crashing a server from overfilling a disk drive is among the human errors that represent some 32 percent of application downtime and data loss. ²

1 AMR research

2 Merrill Lynch/McKinsey report, 2002

The problem with tape

For more than 50 years businesses have utilised tape-based backup schemes. While proponents tout low media costs and portability, tape backup increases the amount of time and effort needed to administer backup and recovery tasks.

With a tape backup solution, how frequently are the tapes sent off-site as part of a disaster recovery strategy? Even if tapes are sent off-site on a daily basis, the tapes are usually sitting in the tape drive or a shipping container for 6-12 hours after the backup completes before they are transported to another location. This leaves an organisation exposed during that time.

Has your business ever migrated from one operating system (OS) to another or from one major release of an OS to another and lost data that was not transferred during the project? Typical tape solutions are tied to the OS and backup software version. Traditional tape solutions create an “upgrade treadmill” that forces a technology change every three to four years. The challenge that results is whether to migrate older backup data. That would be very costly. Equally costly is keeping older drives around with all the associated overhead and maintenance issues.

As tape density increases, more data is stored on each and every cartridge. While this reduces the quantity of media to manage, it increases the risk and exposure of losing much more data if there is a problem. Damaged, misplaced, lost, or stolen cartridges can present a big problem if data recovery is required.

Tape portability is always a problem. With drive and media compatibility and lead alignment to contend with, there are more potential points of failure. If your organisation has multiple offices with little or no IT staff in each remote office, then how do you ensure that backups are properly managed and guarantee consistent and reliable data protection across all offices (See Table 2)?

Table 2: Tape limitations

Problem	Cause
Slower backup and recovery speeds	Tape's linear recording format takes more time to write and restore backup data when compared to the random-access capability of disk. Tape restore times are further slowed by having to locate and mount the media to find the needed information.
Manual intervention required to get data off-site	Without manual intervention, backup tapes remain in the tape drive, leaving the data vulnerable to physical events. While disk-to-disk backup (external drives, appliances) is attempting to solve this problem, it is only adequate for short periods of time. Ultimately, data must be moved off-site.
Inability to verify backup data	Most people do not turn on the option to "verify after write" on their tape drives because this adds 30-50 percent to the time required to complete the backup.
No quick 24x7 access to data for recovery	If tapes are removed from the drives to be sent off-site, there is a significant delay in those tapes returning for recovery purposes.
Human error	Accidentally deleting an email or crashing a server from overfilling a disk drive is among the human errors that represent some 32 percent of application downtime and data loss. ²

NOW IS THE TIME FOR ONLINE BACKUP

Unless you just invested in a new tape library and media to feed it, there's no good reason to still be working with a labour intensive and error-prone method.

Why are DVD's better than VHS tapes? Why are MP3s taking over from CDs? In a sentence, there are fewer moving parts. For many of the same reasons, doing backups online is just a better all around solution for protecting your organisation's business-critical data.

First of all, a backup using an online solution is faster than a full tape backup. Secondly, it's more reliable than tape, Third, the data is off-site or centralised as soon as the backup completes, providing better DR protection. Fourth, the data will last longer with an online solution. Fifth, data can be restored much more quickly than from a tape solution. And, last but not least, your business should be able to save money if you count all the expenses associated with a tape backup solution. The more systems and sites you have, the more money you will be able to save. With three or more locations, the business case becomes a "no-brainer".

GBData's automated online backup service, utilising advanced disk-to-disk software technology, answers the challenges system administrators face by leveraging existing hardware and network infrastructures to securely and efficiently protect servers and desktops against data loss. Our online backup service is provisioned as a managed service to fit right into the existing infrastructure of any sized business. The ability to automatically and immediately move backed up copies of data securely off-site is a key differentiator. Greater security and reliability, and easier, more centralised administration, suggest that online backup is the right choice for organisations that are serious about protecting business-critical data.

The financial advantages of online backup

Online backup places significantly less data on storage arrays than the incremental methods used by tape-based software. Delta-block processing techniques within online backup are extremely fast and efficient as only altered data blocks in new and changed files are backed up. However, once an initial, full ("seed") backup has been sent to the electronic storage device, every subsequent backup appears to the administrator as a full backup. To conduct a restore, an administrator simply accesses the graphical management console and selects the file(s) and or folder(s) required. No manual retrieval and assembly of full and incremental tapes are needed.

Incremental Backups Defined

Incremental backup methodologies used by tape-based software store all files that have changed since the last full or previous incremental backup. Incremental backup tasks are a faster alternative to repeatedly conducting full backups because only the files that have changed since the most recent backup are copied. It is worth noting that entire files are backed up, not just the changed data blocks, making this technique far slower and requiring more storage than delta-block processing techniques.

Additionally, restores take longer. With incremental backups, the most recent full backup is needed as well as every incremental backup made since the last full backup. For example, a full backup on Saturday is conducted with subsequent incremental backups performed on Monday and Tuesday. If the server crashes Wednesday morning, all three backup jobs (Saturday's full backup in addition to the incremental backups on Monday and Tuesday) are needed to recover all data.

So far, this paper has discussed how online backup efficiently helps solve today's most pressing data availability and protection issues. But is there also a way to show how online backup saves money? The answer is a compelling yes.

Business Case

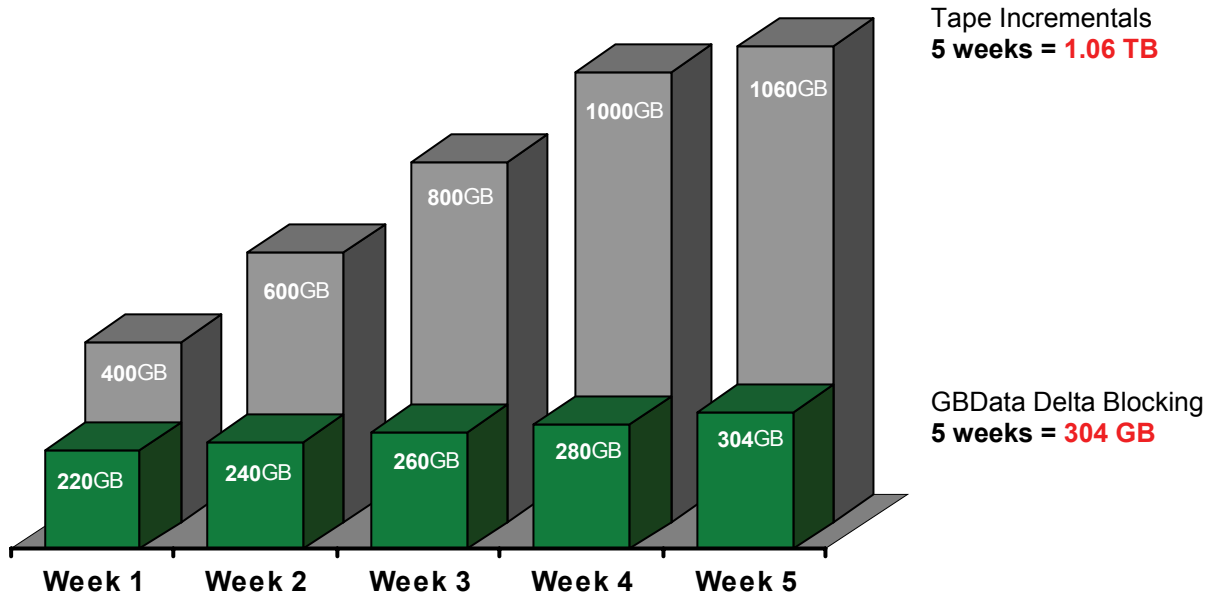
Key Assumption:

- 250GB of original data
- 20% compression + 200GB of compressed data
- Retention Policy = One month of data stored online
 - 6 daily backups online
 - 4 weekly backups online
- Daily data change rate
 - 2% average for Delta-block processing*
 - 5% average for Incremental style backup (tape or disk)

**Achievable with GBData Delta Blocking technology*

Every IT environment is different, and there are several variables in terms of hard and soft costs that make the case for online backup even more compelling. The complexity of a company's IT environment (e.g. distributed offices, heterogeneous servers, large databases) and its data retention policies also help determine the ROI from deploying online backup. This example focuses on the core difference in costs between the full/incremental approach used by traditional backup software and delta-block processing techniques used within the GBData service.

The graphic below illustrates the difference in data created. Because each week a full backup is required using incremental backups, the total after just five weeks is 1.06TB. With GBData and its Delta Blocking functionality and using the same retention policy, only 304GB has been added over the same period.



If that difference in capacity is applied to a business looking to out-task its backup and recovery processes, the cost delta from just data growth alone is quite significant. Online backup prices can vary based on the amount of data stored, length of customer contract and other factors. Average pricing for tape-based (£5/GB) and online backup (£10/5GB) was used before multiplying those prices by the amount of data stored on the respective media:

$$1.06\text{TB} \times \text{£}5/\text{GB} = \text{£}5,300$$

$$\underline{304\text{GB} \times \text{£}10/\text{GB} = \text{£}3,040}$$

Cost savings = £2,260 after just one month in data costs alone.

The cost savings from deploying online backup will only increase as data volumes grow. The £2,260 is only one month of cost savings, which can compound over the course of several months or years.

Still not convinced? Consider downtime, and the loss of money and reputation it can bring. Analyst firm International Data Corp. (IDC) estimates that companies lose an average of £42,000 for every hour of downtime. As pointed out earlier, many tapes are not readable come time to conduct a restore.

What if a single, significant restore is needed from tape and your business is down for two hours? Here are the steps and the amount of time you can expect to spend conducting tape-based versus an online backups:

Tape Recovery Steps	Online Recovery Steps
1. Declare emergency, ask tape storage provider for rush delivery of media	1. Access management console
2. Deliver media	2. Find file/directory and click “restore”
3. Mount and load media	3. Data accessed
4. Locate data on sequential-access media	
5. Data accessed	
6. Tapes rewound and unloaded	
7. Repeat process if more tapes are required for restore	
Total downtime: 2 hours	Total downtime: 5 minutes

Many businesses will need to restore data several times throughout the year. Even if your business does not experience substantial downtime, there is still a significant ROI from online backup. Automated, point-and-click restores when you need them will result in significant resource savings and keep your business running smoothly.

Factors for evaluating online backup vendors

Before choosing an online backup provider, weigh the following technology and cost considerations and ask the tough questions. If they are to be your trusted partner, learn upfront their capabilities and cost structure.

What to look for:

- Automated and unattended backups with the ability to backup open files and open databases
- Ability to centrally manage the backup and restore process from one or more locations
- Control of files and directories to be backed up, with file-filtering capabilities
- Spontaneous file restores 24x7x365 via end user or central administrator control
- No special hardware requirements or changes to your network
- Ability to restore data either over the network or via a dedicated portable storage device
- Customisable data retention schedules
- Data encryption during transmission and when stored on the storage arrays
- Automatic restart and resume capabilities for handling a variety of network conditions
- Automatic notification of exceptions and problems encountered
- Detailed usage reporting capabilities
- Secure, resilient data centre facility

Be prepared

In order to bid for your business, vendors will want to know:

- Your recovery point objectives (ie: the time between data backups) and backup window
- Detailed information on server hardware and operating systems: compile a list of servers to be backed up, including brand name, model number, nomenclature, and estimated amount of data storage
- Existing communication facilities connecting your locations: Internet, Frame Relay, ATM, etc
- Firewall management: How special ports on your firewall are opened to allow backup data to travel across your network?
- Your corporate data retention schedule
- How many people will be administering the backup/restore process and where they are located
- The location of standby systems in the event of a site outage or disaster

GBData: CREATING VALUE THROUGH ONLINE BACKUP AND RECOVERY

With more than 500 customers, including legal, financial, and health care organisations, InTechnology's GBData service is a UK leader in secure online backup and recovery offering unparalleled data protection with instant online restoration. It completely eliminates the severe limitations of traditional tape-based backup.

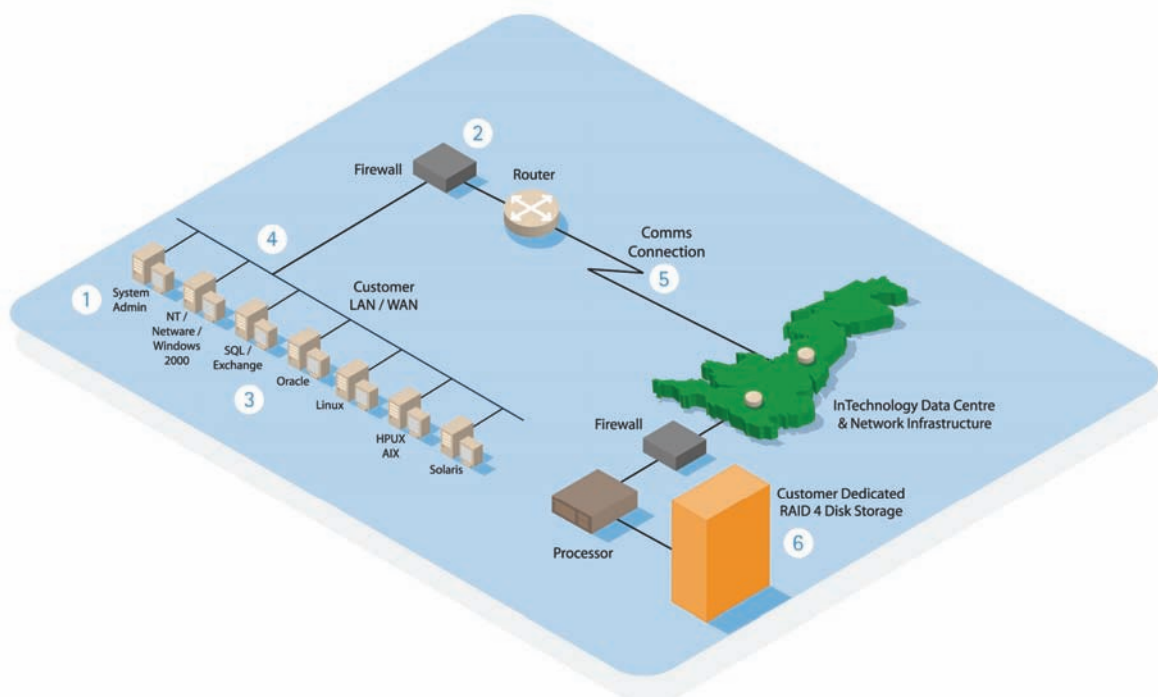
GBData is particularly effective for organisations with business-critical data, for which it has been difficult to implement consistent and coordinated backup and fast / efficient recovery procedures. Features include:

- Centralised and automated backup scheduling.
- Encrypted data sent securely and immediately off-site.
- Data held on dedicated disk for rapid restoration, with an optional second copy.
- Saves time and money and improves the backup process with a dramatic increase in data restore performance.
- Long Term Storage – available as an additional functionality to GBData – allows you to store month-end or quarter-end backup data by taking a point-in-time snapshot of GBData backup data.

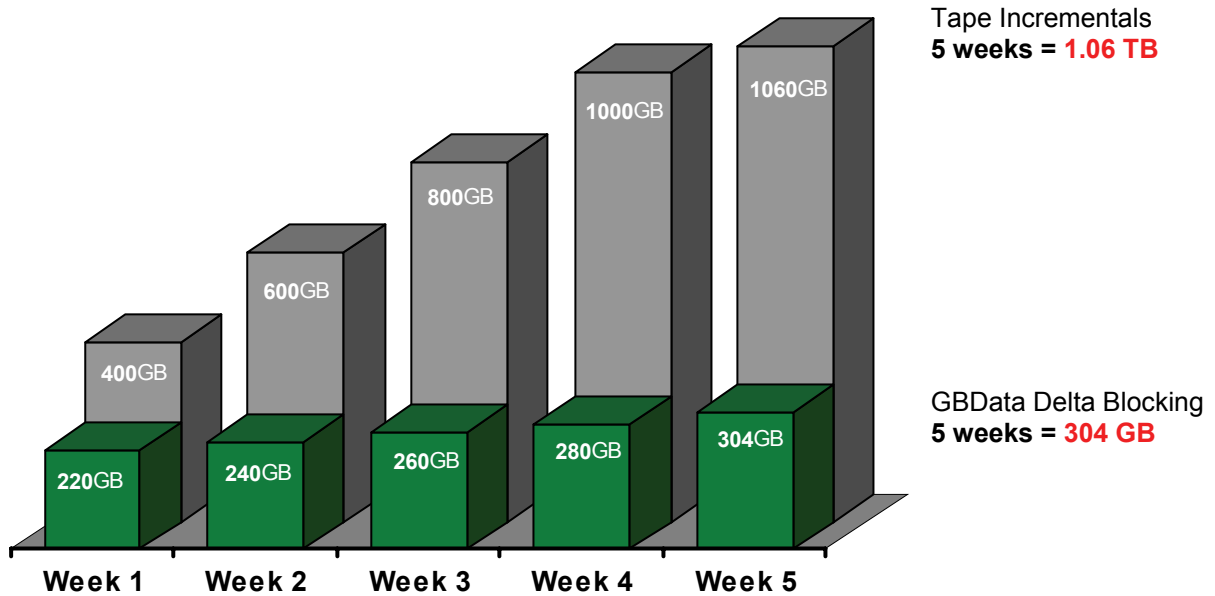
How VBAK works

Speed and security are central to the VBAK service

- 1 The VBAK user interface is installed to allow you to specify backup sets and create schedules, according to business requirements.
- 2 A Router/Switch and Firewall is installed on your network.
- 3 The backup application is installed on an InTechnology provided backup server or your target servers.
- 4 Data is encrypted and compressed prior to off-site transfer for greater security and speed of backup.
- 5 Data is transferred off-site to a resilient, remote InTechnology Data Centre via a secure private network.
- 6 Data is held on dedicated and resilient disk storage for rapid online restore when required.



The graphic below illustrates the difference in data created. Because each week a full backup is required using incremental backups, the total after just five weeks is 1.06TB. With GBData and its Delta Blocking functionality and using the same retention policy, only 304GB has been added over the same period.



If that difference in capacity is applied to a business looking to out-task its backup and recovery processes, the cost delta from just data growth alone is quite significant. Online backup prices can vary based on the amount of data stored, length of customer contract and other factors. Average pricing for tape-based (£5/GB) and online backup (£10/GB) was used before multiplying those prices by the amount of data stored on the respective media:

$$1.06\text{TB} \times \text{£}5/\text{GB} = \text{£}5,300$$

$$\underline{304\text{GB} \times \text{£}10/\text{GB} = \text{£}3,040}$$

Cost savings = £2,260 after just one month in data costs alone.

Advanced Compression and Transmission Technologies

Another technique that GBData employs is an adaptive compression algorithm. GBData will compress the data on a block-by-block basis as the data is being read and queued for transmission to the remote server. If the data cannot be compressed the GBData Agent will skip blocks, re-testing the compressibility periodically, and skipping or restarting compression depending on the results. (There are certain data types that inherently do not compress very well. These types include certain graphics files, previously compressed objects and binary executables). In this way, CPU cycles are not wasted trying to compress data that just won't cooperate!

In general, compression is a CPU intensive process completely justifiable in certain situations and can be avoided in other situations. It is justified when bandwidth is limited and it is important to shrink the data as much as possible and maximise the efficiency on the communication line. In other cases, where 100Mbit or greater networks connect the client with the server, it makes sense to push more data across the network and not spend an inordinate amount of time on compression. To put this concept into play, GBData allows for a "compression bias" setting. This setting tells the GBData Agent whether or not it should optimise compression based on network availability. By asking GBData to optimise compression you will experience a great improvement in throughput across both low- and high-speed networks.

Because only one to two percent of the data needs to be sent across the network, the elapsed time for a full backup is almost always reduced versus a traditional tape backup. And, what's the first thing you do when your full tape backups take too long? Short of buying a faster and larger capacity tape drive (and in the process rendering all of your previous backups useless), you would normally employ an incremental backup during the week and a full backup on the weekend. This reduces the elapsed time of the Monday to Friday backup, but now you have introduced significantly more complexity. If you need to restore a file from yesterday, you require the last full backup and all the intermediate incremental backup tapes. If any tape in the sequence is missing you are out of luck. And, if you run into a problem with a full backup on the weekend, then you could easily go for a full week without any ability to restore current data!

GBData online backups are more reliable because data checking is being conducted at multiple points in the process. When you write data to a tape very few people activate the "verify after write" feature because they either don't know about it, or they don't want their backups to take 30% longer to accommodate the verification step. The downside is that without verification you don't really know if any data was written properly to your tapes until you actually need the data; and then you cross your fingers! Most IT people are risk-averse and this is one risk that can easily be avoided with an online backup solution. With GBData, the data is checked during transmission for integrity, and then again when the data is stored on the remote backup server. These checks are done at the block and the file level to ensure what was backed up reflects the data that is available for restore at any given time. Further data verification can be scheduled on the back-end. This would not be practical for tape – could you imagine the effort required to read each of your backup tapes with 7 days of creation!

When you have multiple servers at a site, the traditional solution to optimise backups would be to employ a central backup server with a small to medium-sized tape library. This configuration does save some money since you don't have to purchase tape drives for each individual server on your network. However, it becomes a bottleneck and a single point of failure for all the backups on your network. Backup servers typically can handle about 40 megabytes per second (MBps) of backup traffic. Today's fast tape drives will accept 30MBps or more of data. The "gotcha" is if you can't push the rated capacity from the client servers through the backup server to the tape drives, then the drives go into a "shoe shine" mode where the tape is stopped and started continuously allowing enough data to fill the write buffer that is sent to the write head. This puts wear and tear on the drive and on the tape media. And, the throughput drops from 30MBps to maybe 15MBps or less. Some companies then go out and buy another tape drive in an attempt to solve the slow-down. This aggravates the situation even further since there is less data flow split between two drives so the shoe-shine effect is worse, and the overall elapsed time of the backup grows even longer.

With GBData, the target backup device is a high-speed disk subsystem so there is a much higher capacity to accept the data. There are also no constraints related to maintaining a constant data flow. Since the data being transmitted via an online solution is only one to two percent of the original size, the ability to handle multiple backups in parallel is greatly increased.

GBData utilises Backup Restore Transmission Protocol (BRTP) for data transmission. This is a robust protocol layered on top of TCP/IP designed to meet the following goals: complete error recovery, efficient use of bandwidth, compatibility with firewalls, and to support encryption standards to protect the contents. To achieve the error recovery goal, the client attempts to reconnect to the server at regular intervals if a network error is detected. The client and server maintain session state information to allow any interrupted transmission to continue from the point of failure. To enable flexibility with different client network configurations, the backup or restore sessions are always initiated from the client side. This simplifies firewall setup because only one port needs to be enabled for outbound connection from any client network. For security purposes, the BRTP protocol allows for negotiation of an encryption algorithm that will be used for the session. The protocol is designed to support future algorithms as they become available.

Conclusion

Business IT functions face more challenges than ever before to keep costs in check as they service employees and customers who are increasingly dependent on data access and intolerant of downtime. A convergence of regulatory and economic drivers, as well as technical innovations has led thousands of businesses to implement online backup.

The IT functions in your business are complicated enough. For small and mid-sized businesses, adding data protection expertise as a core competency simply does not make fiscal sense. For larger organisations, centralising the backup of all desktops, laptops, and remote office servers can realise enormous cost savings. GBData online backup provides organisations with the solution that best fits their needs. Partnering with the leader in online backup and recovery will save you time, money and aggravation. GBData has the right online backup solution for companies like yours that are serious about protecting their business-critical data.

For more information on GBData, visit www.gbdata.co.uk or call our expert team on **0870 479 4287**