

The Ph.D. of Data Protection

Stephen Perkins, the original guru of web backup



The Interview

Q: So NetMass is one of the original online storage providers. It shows how new this industry really is.

A: Putting it that way makes me feel old, but yes, only a handful of the original players remain. We've seen many changes. There is something about the online backup industry that makes it tumultuous. New players come and go frequently. We're still here, and are well-positioned for the future with many exciting new technologies.

Q: So you're on a mission to replace magnetic tape backup. What do you see as the best advantage online backup has over conventional tape backup?

A: I could list a variety of advantages, including continuous data protection, reliability, reduced cost, lower employee overhead, immediate availability of restores, etc. However, I believe that one of the best advantages online backup has is validation.

Here's the tape problem. Once you send a tape offsite, you lose the ability to validate the data. To "prove" that you can recover data from that tape, you would have to call the offsite facility and have them deliver the tape back to you. Only at that point, when you load it into a drive and try to recover the data can you PROVE that the data is recoverable. Nowadays, that is not acceptable.

Quality online backup providers provide the ability to "validate" a backup set. For example, in our case, an operation instructs our data center servers to open your backup set and reconstruct your original data files (this includes decrypting files, uncompressing them, and applying any delta files). Once the operations are complete and the files have been reconstructed at our end, the systems run a digital fingerprint (mathematical checksum) on the files to prove they are identical to the files originally sent by our customer.

In effect, validation allows a company to "prove" that their offsite copy is valid and can be correctly reconstructed. That process is not possible with older, conventional offsite solutions.

Executive Summary

Imagine a geek on a mission to build from the ground up the most reliable data protection infrastructure available. If the image you're getting is of a cherubic fellow with facial hair and a Harley, then maybe you know Steve Perkins, NetMass co-founder, president and CTO.

Steve has a Ph.D. in Computer Science and is an expert in electronic storage technologies. Before co-founding NetMass with CEO Mark Martin, he was a member of the technical staff in the Communications Systems Engineering branch of the Digital Signal Processing Research and Development Lab at Texas Instruments. There, he actively participated in the development of DSL proto-types, the IETF, the IEEE 802.14 working group, and the (then) Home Phone Networking Alliance (HPNA) standards group.

He is a member of the Michigan State University College of Engineering Alumni Board of Directors and a member of the Institute for Electrical and Electronics Engineers.

data protection expert



Next-generation online backup

Interview with an expert



Internet

Q: Not to mention, of course, the advantages of the internet.

A: Precisely. The internet offers immediate data restoration capabilities. There is no waiting for data to be returned to you. Our ServerSafe data protection solution offers the ability to keep a "local copy" of the backup data set onsite in addition to ALWAYS sending it offsite. If you keep a local copy, your restore will proceed at your native LAN speed (typically 100Mbps).

Q: How does NetMass handle the issue of large backups over constrained internet connections?

Many customers have large data-sets that they need to protect. However, not all of that data changes on a daily basis. Many times that data only needs to be sent across the wire infrequently or even just once.

We have two options to deal with large initial backup sets.

1. The first is to just start sending the data. While it might take days to get the initial backup set, the operation is seamless and automatic. If the customer specifies "end times" to their backup schedules, then the backups will stop at the specified time and pick up where they left off when they start back up. If a customer wants to let the backups run continuously, we also have the option to throttle the bandwidth so that the backup process does not clog their Internet pipe.
2. The second is to "pre-load" the data center with the data. Our software supports this operation and allows a customer to direct the initial backup to some alternate media (typically a portable hard drive or NAS device). They ship us that device and we can "pre-load" the data center with the data. Subsequent backups only need to send across new and changed files with respect to the initial backup.

Size Issues

Q: What about the issues with large daily backups?

A: More and more customers have increasingly large daily backup sets. These might include SQL server backups and Exchange backups. We use several approaches to cope with large backup sets.

As with any good online backup software, we apply a number of very sophisticated data optimization techniques that "reduce" the size of the data (data de-duplication, delta-block processing, and compression). These techniques are very effective in reducing the amount of data that must go across the wire.

Large Exchange implementations can benefit from our message level backup. This allows us to back-up (and restore) individual emails instead of the entire exchange store. This can significantly reduce the time it takes to back up an entire Exchange server.

Our software supports continuous data protection for Exchange and for file systems. Continuous data protection tries to immediately back up data as it changes (i.e. emails as they arrive and files as they are saved). By allowing the backup software to continually protect emails and file servers, you can benefit from a 24-hour backup window. Bandwidth throttling can be used to prevent clogging your wide area Internet connection.

Before we sign up a customer, we work closely with them to understand their backup needs. We have software that we can use to evaluate their data and give us statistics that will determine the feasibility of our services. Obviously, there are some data sets that are not appropriate for online backup.

Next-generation online backup

Interview with an expert



Security

Q: Nowadays, security is such a big issue. What are the issues with security and online backups?

Security is a multi-faceted issue, but we are never the weak link.

Physical Security. We are in Tier-1 data centers that benefit from very tight physical security. The structures are wrapped in Kevlar, are well above the 100-year flood mark, have 24x7 manned security personnel, bio-metric access, motion activated cameras, all the works. The physical access controls are typically well beyond what SME customers have at their premise. It would be easier for an intruder to gain access to the customer's data at the customer's location than at ours. Even if an intruder did gain access to our data centers, all data we have is stored encrypted so it is unreadable.

Network security. Our network infrastructure is tightly controlled and secured at many levels. We implement things like network intrusion detection, network intrusion prevention, two-factor authentication, separation of services, etc. As with item one, our infrastructure is typically much more secure than that of the average SME.

Data Security. Our software encrypts all data before it leaves the customer premise. It is encrypted "in-flight" and "at rest" using AES-256 encryption. We do not escrow a copy of the encryption key so we do not have the ability to decrypt the data on our end.

Restore security. We have systems in place that also keep data "in" the company. By only talking to authorized IP addresses and by registering the install of our software to a specific machine, we can prevent disgruntled employees from moving, transferring, restoring, and possibly stealing company data. This is true even if they have access to the encryption keys for the company's backup.

Install procedures. We have developed a set of "best-practices" for installing and operating our software. As part of our service, we help each customer set up and configure the software. By following the best-practices guide, we can assure a customer of a good install.

Awareness

Q: What are the biggest obstacles an SME faces in attempting to use online backup services in lieu of traditional tape systems?

A: Awareness. Online backup is not a new technology — NetMass has been offering online backup as a core service for over 8 years, and the software has improved significantly over time. Problems associated with older generation backup (which is still in use) including constrained bandwidth, security and reliability issues, slow restores, validation of the backup set, and infrastructure have all been solved with our next-generation technologies. However, many IT people do not realize the sophistication of the solutions now available with our ServerSafe total data protection. A five minute conversation is usually all that is necessary to convert a skeptic into a champion.

Compounding the problem is that the press tends to give a lot of attention to the very low-end providers. These providers offer very large or unlimited backups for free or almost free. However, these services are usually geared towards a home consumer. They do not meet the needs of the SME and their shortcomings increase the noise floor for those looking for a more comprehensive solution.

Q: For an SME with approximately one terabyte of data to backup monthly, what would be your company's fee

A: This list price for a terabyte of compressed storage (which typically equates to more than 2.5TB of customer data) is \$7/G for a monthly contract. Customers can earn significant discounts by agreeing to longer term contracts.

Q: How much is your average customer backing up (data size)?

I'd say 100GB compressed per physical location. With cross-site data de-duplication, optimization, and compression, this is usually greater than 250G of customer data per site.

Next-generation online backup Interview with an expert



SME Customers

Q: How do most of your SME clients connect to your service? Over public IP networks or dedicated connections? Are you seeing any trends in a specific direction?

Definitely public IP. With the availability of broadband, we rarely have any requests for private connectivity. Private connectivity requests typically come from resellers who are aggregating many backups to us. For instance, co-location providers that offer our online backup solution to all the customers in their data center.

Q: What percentage of your business is comprised of SMEs?

Over the years, we have found that one solution will rarely fit everybody's needs. One of the advantages we have over other online backup service providers, is that we have a range of online backup services. This allows us to handle the needs of a very small company that has a single desktop all the way up to franchises that have hundreds of locations and large data sets. I would say that 45% of our customers are SME (many of them with multiple physical locations). The majority of the rest would be classified as small business plus a number of SOHOs and quite a few consumer accounts to round out the count.

About NetMass

NetMass is one of the original online storage providers. As a pioneer of the industry, NetMass began leveraging the power of the internet for online backup in 1998. NetMass currently offers a full range of data backup, restoration, online storage and collaboration services. The company also has an active partner/reseller program. NetMass is located in Dallas, McKinney and Houston and has clients and partners around the world. More information is available on the web at www.netmass.com

