

Data Protection for Remote Sites Is Not Cheaper by the Dozen

By David G. Hill, Mesabi Group

Organizations are logically thought of as large integrated entities. However, practically speaking, apart from central headquarters, many organizations are typically made up of numerous geographically-dispersed remote offices and remote branches. Retail stores of all shapes and sizes, branch banks, and brokerage and insurance offices are only a few of the numerous available examples.

A significant portion of the critical data created, accessed, and used by an organization results from the work performed at remote sites. This data requires the same level of protection that is granted to data of the same caliber at a central site. Yet most discussions of data protection tend to focus on individual large locations that have the scale to house corporate data centers. A data center manages numerous backup streams for multiple applications across a large amount of data. A data center features the economies of scale as well as the professional IT staff to manage complex data protection processes.

Many remote locations do not have the scale to operate independent IT functions. While the person who is assigned the part-time role of backup administrator is often not an IT professional, the real problem lies not in his or her skills or dedication, or the quantity of the data that needs to be protected, or even in backup software and backup media drives. The real problem is inherent in the back-up process itself, which includes monitoring backups, swapping tapes, and performing restores.

Backup may only incorporate a single tape drive or optical disk drive, but adequate backup processes require that multiple copies be available over time. A copy has to be sent off-site say on a daily basis, which, whether mail or pickup, tends to be relatively expensive. Moreover, at least a second copy has to be kept onsite in case a restore is required and the last backup is on a defective piece of backup media. And, oh by the way, the latest backup is likely to be that of the previous day, which may not be good enough. All in all, most remote back-up qualifies as a Rube Goldberg manual process crying out for automation.

Straight-forward Backup Consolidation

The old dictum "think locally, act globally" applies particularly to data back-up. Backup consolidation means that backup data is sent from remote sites via a WAN to a central site (which can either be an enterprise's own data center or that of a third

party service provider). The central site provides central control and automation of the backup process. Not only is the local backup drive and media infrastructure eliminated, but so are cumbersome manual processes such as managing the offsite removable media. This all qualifies as good news for information-dependent businesses and part-time back-up administrators.

More good news is that the recovery point (i.e., the time since the last backup) is likely to be greatly improved by such processes. Typically, a backup at a remote site is done once a day. In a central site backup consolidation scheme, changes to data can be sent continuously, on an event-basis, or on a scheduled basis (depending upon the product used) to the central site. Typically, these changes can either be stored as incremental copies or reconstructed to provide a complete, up-to-date copy of a file. With those products that have this latter capability, the ability to quickly restore a file that was accidentally deleted at a remote site is possible.

Moreover, the restore time for a full system may be faster in many cases as a disk to disk transfer (taking into account the possible gating factor of the network) than with the old system. In having less data at risk and faster recovery times, the remote sites may actually have an advantage over their central office brethren in some cases! As a result, in one fell swoop remote offices move from being insecure data protection backwaters exposed to the risk of loss to the forefront of data protection technology! Not bad.

The technology for backup consolidation from hard disks at a remote site to hard disks at a central site over an IP network has actually been around for quite a while. However, two enabling technologies have made this process more cost efficient, leading to the emergence of new players in this market. The two technologies are WAN acceleration (a.k.a. WAN optimization) (see "WAN Acceleration ..." Pund-IT Review, December 21, 2005) that enables more cost effective utilization of existing bandwidth and data reduction techniques (see "Space Savers ..." Pund-IT Review, November 23, 2005) that reduce the amount of data that has to be transmitted.

Backup Consolidation as a Byproduct of Server-Storage Consolidation

This backup consolidation strategy does not touch the existing server-storage infrastructure so the remote location can continue operating as before. Depending upon an organization's business model, leaving well enough alone may be the right thing to do. Other organizations may want to consolidate its remote server-storage infrastructure into a central site to take advantage of economies of scale from both an IT infrastructure and IT organization perspective. Although backup consolidation may sweeten the pot in terms of justifying such efforts, backup consolidation is simply a natural byproduct of the overall consolidation effort.

If remote office servers and storage are consolidated into a centralized data center, eliminating enough WAN latency to enable an acceptable response time to access the data now stored at the central site is critical if the solution is to be workable. A third technology, Wide Area File Services (WAFS), replaces branch office file servers with a network appliance and adds a similar appliance at the central site. The appliance is, in effect, a surrogate file server to local applications that attaches directly to the LAN. The appliance maintains a cache copy of any file that a user opens. Meanwhile, since

the central site stores the master copy of the file standard backup practices and procedures can be implemented for the new data.

Dramatis Personae

Companies in the backup consolidation space tend to be recognizable largely by friends, relatives, and customers (Table 1). Only the trade press and industry analysts would recognize all of them. That said, please pay attention to them. In general, the technology they offer seems to be very strong and the need they fill is clearly recognizable.

As with many other technologies, large storage companies and service providers are more than willing to work with remote back-up specialists in one form or another. Asigra's Televaulting software is incorporated into HP Electronic Vaulting Services for Enterprises, a managed service that includes installation, configuration, and housing of data from remote offices in HP data centers. Asigra is also part of the EMC NAS partner program. Avamar is a part of HP's Business Continuity Services and IBM's TotalStorage Proven program. Data Domain is a technology partner with EMC Legato and IBM Tivoli. HP and IBM are LiveVault enterprise partners. Signiant counts EMC as a key partner. Tacit Networks is part of the EMC NAS partner program, the IBM TotalStorage Proven and ServerProven programs, Microsoft Storage Partner Program, and Network Storage Advantage Partner Program. Topio lists Hitachi Data Systems as one of its many partners.

The majority of companies in Table 1 focus on straight-forward backup consolidation. Tacit Networks is the only one that emphasizes backup consolidation as part of a larger server-storage consolidation strategy. Tacit Networks is not the only company in the WAFS space, but it is the leading advocate of remote office backup consolidation using a WAFS solution.

Which is the better approach? The answer is that it is not a matter of taste — vanilla vs. chocolate ice cream — but rather a matter of organizational requirements. The push-me/pull-you alternative benefits are fairly clear:

- Remotely-created data that is shared through the entire organization or needed by the central organization for something like inventory replenishment as part of supply chain provisioning favors greater central consolidation
- Remote locations which are profit centers with local responsibility for decision-making would favor straight backup consolidation as a utility service (if willing to consider anything)
- Remote locations which are cost centers theoretically should be willing to consider either alternative on a purely economic basis

The economic analysis is more likely to be a comparison between the status quo and the proposed vendor solution rather than an ROI-shootout among competing vendors. The decision has to be made by comparing the difference between the pieces removed from the remote infrastructure and the costs of moving to the new infrastructure as well as the value derived from improved manageability, accountability (especially in these days of compliance), and reduced risk.

Table 1: Remote Data Protection Product Sampler

Vendor	Product	Product Focus	Technology Foundation
Asigra	Asigra Tele-vaulting	Backup consolidation using block-level incremental backup	Agentless incremental forever backup with policy-based scheduling
Avamar	Axion E	Backup consolidation using scheduled replication	An appliance-based backup and restore solution that provides local backup services for remote offices.
Data Domain	DD400, Data Domain Repliator	Backup consolidation using replication and standard enterprise backup software	DD400 appliances mirror data over an IP WAN to a central data-center
EVault	EVault InfoStage	Backup consolidation using block-level incremental backup	Agent-based, online, automated disk-to-disk backup over a WAN
LiveVault (Iron Mountain)	LiveVault In-Control	Backup consolidation that captures changes as they occur and sends them over the network on a continuous or scheduled basis	Agent-based, online disk-to-disk backup over a WAN on either a continuous or scheduled basis
Signiant	Signiant Mobilize	Backup consolidation moving only bytes that have changed	Disk-to-disk backup of changes over an IP network on a scheduled or event-driven basis
Tacit Networks	IShared family	Backup consolidation as a byproduct of consolidation of servers and storage in a central site	Appliance using WAFS to allow users in remote offices to access and share files globally at LAN speeds over a WAN.
Topio	Topio Data Protection Suite (TDPS)	Backup consolidation using continuous replication	Asynchronous replication over an existing IP network

Source: Mesabi Group, January 2006

Mission Accomplished?

The data protection problem at remote offices and branch offices has been recognized for a long time. But while straight backup consolidation capabilities are relatively common, their evaluation and acceptance across a broad range of organizations is just starting to reach the knee of the curve. The general availability of data reduction and WAN acceleration technologies has made the backup consolidation value proposition more cost attractive as existing network infrastructures can be leveraged for the process. The overall server-storage consolidation capability that brings backup consolidation along as part of the ride has benefited from WAFS capabilities, as well.

Even though these technologies are still evolving, the increased recognition the technology is there to do the job and the question is simply one to heighten awareness of the reality of the backup consolidation space. To wit:

- Organizations, whose remote locations share much their data with other parts of the organization, probably had a server-storage infrastructure only out of previous cost/performance considerations. A new cost equation using WAFS would encourage them to consolidate everything as long as performance was satisfactory.
- Organizations whose remote locations are profit centers and depend upon fast access to local data might be willing to "outsource" the unpleasant task of backup. However, maintaining control over availability of that data (for example, they do not want to be down if the network is down) may dominant the choice in favor of straight backup consolidation unless sheer economics or dissatisfaction with the current mode of operations come into play.
- Organizations whose remote sites are cost centers would probably put great weight in a financial decision including whether or not the remote IT assets have been depreciated. Non-financial considerations, such as respect for remote site autonomy, or concerns about rocking the boat may also play a role.

Overall, the reported growth of some of the companies indicates that somebody is paying attention and a lot more is likely to happen soon. Pay attention to this space. Lights, camera, action.

David G. Hill is principal of the Mesabi Group (www.mesabigroup.com). The Mesabi Group focuses on the revolutions in Storage Networking and Storage Management, and helps clients make the best and most efficient use of information for business value.

© 2005 Pund-IT, Inc. All rights reserved.

Contact:

Web: www.pund-it.com

Phone: 510-909-0750

E-mail: charles@pund-it.com