Technology Brief

Accelerating Data Migration with WAN Optimization

Date: August, 2010 **Author:** Terri McClure, Senior Analyst

Abstract: *Sending large amounts of data over distance can be costly and time consuming. Unfortunately, it's also necessary for a myriad of IT functions including remote office backup, outsourcing, data center moves, and cloud computing models. Increasingly cost sensitive business models are forcing more and more companies to find ways to perform migrations more effectively and efficiently. Multiple options are available to get the job done and, of course, there are tradeoffs to be considered for each. One such option, WAN optimization, makes data migrations across ubiquitous IP networks viable and affordable.*

Overview

Merely uttering the phrase "data migration" can strike fear into the hearts of IT administrators. Migrations are often lengthy and painful processes; they are also a necessary evil that supports business initiatives like outsourcing, data center moves, and cloud computing.

ESG recently conducted a survey of 515 senior IT professionals concerning their organizations' data center plans and priorities for 2010 and beyond. [1] Among other findings, ESG uncovered that one in three companies are aggressively consolidating their data centers: more than one-third (35%) of respondents currently have data center reduction or consolidation activities underway. Of that group, enterprise-class organizations are more likely than midmarket firms to either be in the process of reducing or consolidating data center space. In fact, data center consolidation ranks as the third most important overall IT initiative over the next 12-18 months for enterprise-class firms.

The same survey found that 20% of respondents plan to increase use of IT outsourcing as a cost containment strategy and 17% plan to increase use of cloud computing services as an alternative to in-house applications and/or infrastructure. Whether it's data center moves, outsourcing, or subscribing to cloud services, all of these initiatives often have one thing in common: they depend on a successful migration to get new operations up and running.

A number of issues must be taken into consideration when planning a large scale data migration:

- **Time.** How long the migration takes will vary extensively based on the approach. With some methods, users also need to consider whether any additional time will be required to roll forward changes that happen between when a data copy happens and when it is up and running at the new site.
- **Data availability.** Some methods can require extensive downtime, during which data cannot be accessed.
- **The cost of bandwidth.** Depending on the amount of data and distance to the new site, implementations can become prohibitively expensive: the greater the distance or larger the data set, the more costs incurred.
- **Portability between service providers.** This concern is especially valid when leveraging cloud services, which may require programming to proprietary APIs to access data.
- **Data reduction/transport optimization.** Users need to understand the available options that reduce the overall amount of data that needs to be migrated, which can reduce both costs (media or bandwidth) and time to complete the migration.
- **Security.** Each method comes with associated risks; users must be able to ensure data is secure while in transit.

[1] Source: ESG Research Report, *Data Center Consolidation and Construction Trends*, June 2010.

Options for Large Scale Data Migrations

Options for moving a large amount of data en masse from one location to another are very limited.

Backup and Restore

Data is backed up from one system and then restored on another. The process takes applications down, is typically slow, and consumes significant server resources. With frequent delays, it is not at all unusual for a user to get only some percentage of the way through (and whether that is 1% or 99% makes no difference), exceed the operational window, and then have to start all over again. Just about the only good news is that backup operations sometimes have their own network (although that can be an expensive proposition), so network clogging can be avoided.

The biggest challenge with tape backup is that a user doesn't really know whether a job succeeded or failed until it is complete. One bad tape renders a backup useless—the process needs to be restarted. Restores from tape can also be lengthy and time consuming processes. Data reduction is available through deduplication software, which is becoming ubiquitous in the backup process, but extra time may need to be added to re-inflate the data upon restore if required. Security can be attained through encryption.

Copy to and Ship a Disk Subsystem

Users do have the option of loading the data onto a subsystem and shipping it to the new or remote site. The process is similar to regular backup in many respects, but it is also faster than backup and restore since disk-to-disk copy is faster than disk-to-tape and once the system is at the new site, it just needs to be powered up—there is no need to wait for tape restores at the remote or new site. Encryption technology is available to secure the data.

This process is also risky: disks could be damaged or lost, which could cause data loss or corruption that may not be immediately evident. Depending on distance, shipping time needs to be factored in, as well as data scrubbing and integrity checking, which could take a long time depending on the size of the data sets involved.

Network Copy

In this scenario, the server or storage array is called in to copy data across the network. Depending on the available network bandwidth, the amount of data, and the distance, this can be an expensive and lengthy process. Data is sometimes inaccessible during the copy or access is slowed significantly due to resource contention. Since migration can consume significant network resources and other applications can be affected as well, the transfer should happen during off-peak hours. The upside is that users don't have to worry about tapes or disks getting lost or damaged and risk of data loss is virtually nonexistent as writes are acknowledged at the remote site. The main downsides of network copy are time, security, and bandwidth costs, all of which can be mitigated by leveraging WAN optimization technology.

Enter WAN Optimization

WAN optimization technologies let companies do more with less, enabling them to transfer more data over a smaller network connection. Such efficiency could deliver significant improvements in both existing and new disaster recovery environments. WAN optimization leverages a host of technologies like protocol optimization, encryption, deduplication, and compression, which combine to provide consistent and secure high performance connectivity over the WAN. Additionally, policy-based quality of service (QoS) and load balancing guarantee that specific applications can be given priority for available bandwidth, minimizing the impact of the migration on other network traffic.

Leveraging WAN optimization for data migration is:

- **Faster and more efficient**. Data reduction technologies like compression and deduplication mean less data needs to be transferred, allowing the process to finish sooner. More efficient data transfer also means more data can be moved in smaller windows.
- **Safer than shipping a disk subsystem or tapes**. There is no physical piece of equipment containing data being shipped from place to place.
- **Cost effective.** Users can push more data over a smaller network connection, minimizing network costs. WAN optimization solutions can improve network traffic management and efficiency, which could reduce total network consumption needs.
- **Secure.** WAN optimization technologies provide secure communications by encrypting data in flight and authenticating users accessing applications and data. It can also help reduce the load on back-end servers by offloading processor-intensive security operations for SSL/TLS and network encryption.
- **Less disruptive than backup or network copy.** Leveraging technologies like load balancing and policy-based QoS, critical applications can receive priority bandwidth allocation to avoid disruption by the migration during the day—the migration can then take priority during slower network periods, optimizing available bandwidth to meet business needs. Additionally, protocol optimization is used to reduce the impact of chatty protocols, streamlining communication between sites.

WAN optimization technologies with multi-Gbps throughput offering deduplication, compression, and high availability configurations can create efficiencies that enable the use of standard service provider networks. This could potentially save millions in capital and operating costs that would have been spent relocating a data center due to insufficient network connectivity. At the very least, there is the potential to save tens to hundreds of thousands of dollars in reduced network costs between sites.

The Bigger Truth

Migrations are migrations. The driving force behind a migration is ultimately not important: the challenges associated with migrating data to the cloud are not all that different from the challenges many enterprises have already faced with data center moves and consolidation. There are a number of ways to tackle the challenge. Trade-offs will need to be made regarding cost, speed, and risk. Shipping equipment around can be risky from a security standpoint and restoring from tape can be a long and laborious process that does not guarantee success. Moving data across the network can take some risk out of the equation and has a much better success rate, but it can also be time consuming and expensive without technology to give it a boost. WAN optimization provides that boost. It is already a proven technology and has been used broadly for data center consolidation, remote office application acceleration, and backup/disaster recovery initiatives. WAN optimization technology can help reduce the pain of migrations so they are completed faster and cost less. Using WAN optimization, users can finish migrations faster and return to productivity sooner.