



AUGUST 21, 2008

ENTERPRISE DATA PROTECTION USING DATA DOMAIN

High-Speed Inline Deduplication Storage Systems.

OBJECTIVES OF PASADENA PILOT

- Automate backup and recovery processes
- Retain data for longer periods of time
- Accommodate restore requests faster and more reliably
- Maximize efficiency of staff resources
- Ensure business continuity by transferring data changes offsite on a nightly basis

SITUATION

One of the world's largest data storage organizations is a leading healthcare provider, encompassing over 700 physical sites. With the increasing need to meet demands for high availability, business continuity and regulatory compliance, implementing an efficient backup strategy across the organization has become very challenging in terms of both scale and complexity.

A recent pilot conducted by Trace3 for this leading healthcare provider on three production sites in the Pasadena area serves as a practical model for a better approach to enterprise data protection that is based on archiving data to disk rather than tape.

Tape backups are inconsistent, unreliable, and take substantial manual effort to maintain. Furthermore, retrieving and restoring data from tape can take up to two full days, putting business continuity at risk. Better data recovery means retaining data online locally for sufficient periods of time, which is not possible using our current model. According to Gartner, one in ten recoveries from tape fails.

Both of this leading healthcare provider's regional medical offices and national data centers need a more efficient and reliable backup solution to enable cost and labor savings, ensure data protection, and allow for improved business continuity.

This disk storage offering will be included within our data protection suite of products which encompasses backup and recovery using Symantec's Veritas Net-Backup, allowing long-term onsite retention and highly efficient network-based data replication for disaster recovery, remote office protection and tape consolidation.

BACKGROUND

Before the pilot, all the leading healthcare provider's sites in the Pasadena area were backing up individual servers to tape drives. This resulted in many servers dedicated to backups, as well as inefficiencies due to manual processes associated with tapes. Three sites—Site 1, Site 2 and Site 3—were selected for a refresh that included installing a Data Domain high-speed inline de-duplication backup appliance as network attached

SUMMARY OF PILOTS RESULTS

95% reduction in hardware footprint (3U of rack space vs. 70U)

75% less power required (648W vs. 2800W)

74% less heat to dissipate (2213 BTU/hr vs. 8568 BTU/hr)

Ability to retain data on-site for 90 days

Data restores in minutes vs. days

229% increase in ability to run concurrent backup jobs (64 vs. 28)

16 staff hours a week saved by not changing tapes

Tape costs eliminated at 2 of the 3 sites

40-to-1 bandwidth reduction for WAN transfer

Improved data security

30-to-1 data compression resulting in 48% less cost per GB of backed up data

storage at each of the three facilities for disk-to-disk backup.

Because of the data compression and de-duplication features the Data Domain devices provide, enough bandwidth is conserved so that transferring backup data over the WAN is practical. Each site is backing up 3-4 terabytes per week. The three sites replicate in a “triangle” fashion: Site 1 and Site 2 replicate to Site 3, and Site 3 replicates back out to Site 2, ensuring full redundancy and business continuity.

WHY DATA DOMAIN?

The pilot team evaluated various solutions and concluded that Data Domain’s only real competitors in the hardware deduplication market are Fiber Channel virtual tape libraries (VTLs). The leading healthcare provider’s distributed sites lack the localized fiber infrastructure to make it practical to deploy this technology.

As a hardware appliance, Data Domain offers simplicity advantages over software de-duplication solutions. It is essentially a plug-and-play unit, requiring minimal set-up and management time. What’s more, the partnership between Data Domain and Symantec has led to tighter integration with Veritas NetBackup—the leading healthcare provider’s preferred backup software solution going forward. Among these integration points is the co-developed Open Storage Standard (OST). OST enables the

NetBackup infrastructure to manage backup data replication and update the Backup Catalog to allow the seamless restoration of data at a disaster recovery (DR) location.

IMPACT ASSESSMENT OF DATA DOMAIN

Save on space, power and cooling:

To store all backup data on tape, almost two full racks would be required at each of the Pasadena sites (70U). Using the Data Domain appliance, only 3U of rack space is necessary at each site—a 95 percent reduction. This footprint reduction results in 74 percent less heat to dissipate and 75 percent less power required.

Retain data longer: The Pasadena sites are now able to retain backup data on-site for a full 90 days, so IT staff can better meet compliance requirements and accommodate restore requests over a longer period of time.

Restore data faster and more

reliably: Because administrators no longer need to drive to remote sites and locate tapes to restore data, restores can be completed in hours or minutes (1-2 gigabyte restores in under 3 minutes; 500 gigabyte restore in less than 5 hours) instead of 1-2 business days. Restores are also more reliable. In some cases, older tapes cannot be restored because sites no longer have the correct tape drive formats to load the old tapes—another drawback of a purely tape-based backup strategy.

Save administration time: Because Data Domain is a shared, centrally managed storage resource, multiple servers can write to it at any given time, vs. one server at a time per tape drive. (Estimated 16 staff hours saved per week changing backup tapes in Pasadena area)

Reduce WAN bandwidth required for replication: With an artificial throttle of 5 megabits, a terabyte of backup data can be replicated within 24 hours. This translates to roughly a 40-to-1 bandwidth reduction for the WAN transfer. This savings in bandwidth enables business continuity. Replication to the remote site begins as individual backup jobs complete and finish writing to the Data Domain.

Improve data security: With the Data Domain appliances, backup data is replicated off-site while remaining inside the network, substantially reducing the risk of data loss. Sending tapes offsite with unencrypted data is a business risk. If the tapes were to fall into the wrong hands, important, confidential data could be retrieved. Therefore, we additionally recommend adopting a tape encryption strategy for any situations where tapes are required to be sent off-site.

Reduce storage costs: With the advanced compression the Data Domain devices provide, it's actually cheaper to back up to disk than to purchase and operate multiple tape libraries. Tape costs are eliminated and hardware costs

are 48 percent lower.

DEDUPLICATION COMPRESSION STATISTICS FOR PASADENA PILOT (COMPLETED JANUARY 2008):

Site 1: 31-to-1 compression ratio (55 terabytes of backup data stored in 1.8 terabytes of RAID 6 Data Domain storage – DD565)

Site 2: 30-to-1 compression ratio (47 terabytes of backup data stored in 1.5 terabytes of RAID 6 Data Domain storage – DD565)

Site 3: 31-to-1 compression ratio (101 terabytes of backup data stored in 3.2 terabytes of RAID 6 Data Domain storage – DD580)

Note: Compression achievement based on 150 terabyte data storage requirement

CONCLUSION

Based on our success with the Data Domain deduplication appliances for disk-to-disk backup in the Pasadena pilot, we recommend that all the leading healthcare provider's distributed sites and national data centers adopt this approach. The three-tier data center model ensures business continuity and lays a solid foundation for an improved enterprise disaster recovery strategy. Due to the bandwidth savings the Data Domain appliances provide, data could be replicated to National Data Center 1, National Data Center 2 or National Data Center 3 locations to provide true DR capabilities.

**THANKS TO THE FOLLOWING PARTICIPATING CONTRIBUTORS IN THE
SUCCESS OF THE PASADENA DATA PROTECTION PRODUCTION PILOT.**

Matt Hallmark, Trace3
Tam Nguyen, Trace3
Ruben Navarro, Trace3

For more information contact your Trace3 Representative or visit us a www.trace3.com

© 2008 Trace3, Inc. All rights reserved. Specifications subject to change without notice. All other brands or products are trademarks or registered trademarks of their respective holders and should be treated as such.

