

FalconStor Addresses Top Deduplication Criteria

Date: May, 2009

Author: Lauren Whitehouse, Analyst

Abstract: There are many facets of deduplication solutions to consider prior to deployment. FalconStor VTL with deduplication delivers proof points for end-users' top purchasing criteria.

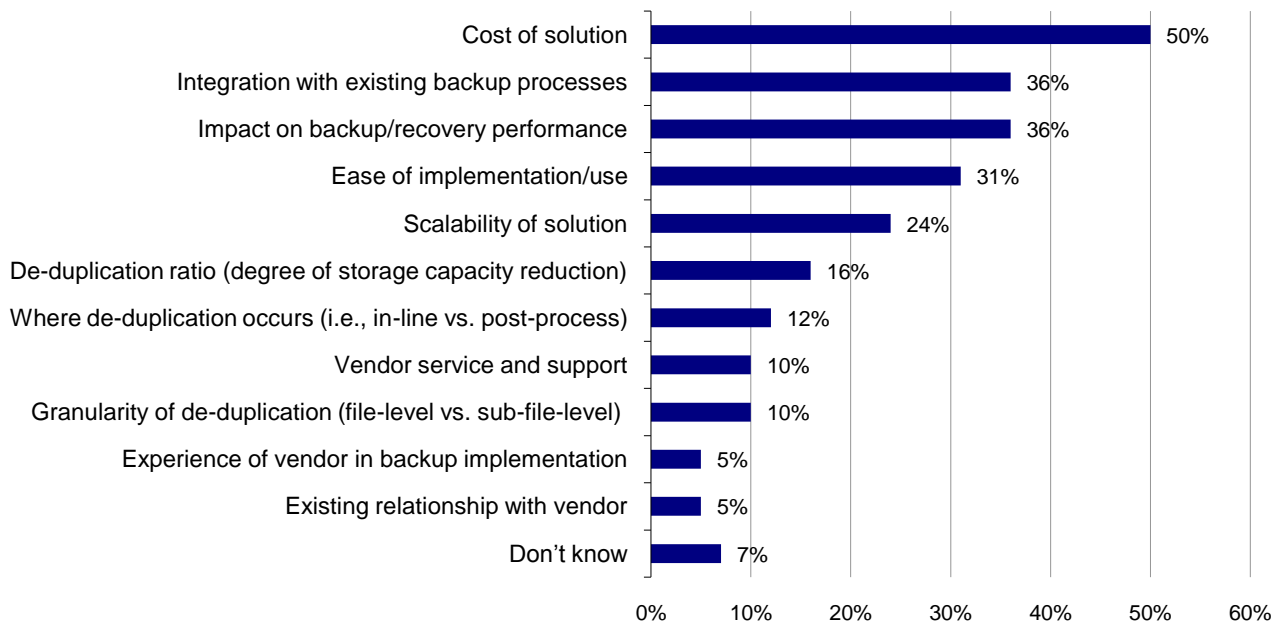
Overview

Data deduplication has quickly moved from a “nice-to-have” to a “must-have” technology for those organizations trying to keep pace with relentless data growth. Organizations surveyed by ESG are citing difficulty keeping up with storage capacity demands, while also reporting flat or decreasing IT budgets. With data deduplication, redundant data in the backup process is identified and eliminated, creating savings in valuable storage capacity and, importantly, in constrained IT budgets.

When ESG asked respondents to a recent survey to identify what they consider to be the most important criteria in the evaluation and selection of deduplication technology, several factors were revealed (see Figure 1).¹ Not surprisingly, the cost of the solution was the most frequently-cited factor (although savings garnered from capacity reduction often overcomes financial objections to deploying deduplication).

FIGURE 1. DESIRED ATTRIBUTES OF DEDUPLICATION SOLUTIONS

Which of the following considerations would you say were most important in your organization's evaluation and selection of data deduplication technology? (Percent of respondents, N = 58, multiple responses accepted)



Source: Enterprise Strategy Group, 2008

¹ Source: ESG Research Report, *Data Protection Market Trends*, January 2008.

Following cost, compatibility with backup applications, ease-of-deployment and -use, and the impact on backup/recovery performance were important considerations—more so than technical implementations such as inline or post-process approaches, level of granularity, or the deduplication ratio. Scalability of the solution also ranked high, an important factor considering the aforementioned difficulty in keeping up with storage demands.

The fact that only 5% of current data deduplication users cited “existing relationship with vendor” as a key consideration in their product evaluation and selection processes suggests that end-users would be open to considering deduplication solutions from a wide range of incumbent and emerging vendors. One vendor that has demonstrated a compelling solution is FalconStor. FalconStor Virtual Tape Library (VTL) with deduplication promises improvements in data protection effectiveness and the ability to cut costs.

Analysis

FalconStor VTL emulates physical tape via disk for faster, more reliable backup and restore. It lets users connect it to any disk array and is compatible with most major backup applications. Its integrated data deduplication reduces storage capacity requirements—saving money and/or extending the time data can be retained on disk. The solution can scale up to an eight-node cluster (with an additional eight-node deduplication cluster) and a storage pool that can accommodate more than two petabytes of actual capacity (or up to 40 petabytes of logical/deduplicated capacity), enabling considerable scalability. Its clustered architecture also ensures high availability. FalconStor VTL has the ability to send data offsite through either physical tape creation or low bandwidth data replication.

Evaluating FalconStor VTL with deduplication using the top five criteria cited by ESG research respondents as most important for selecting data deduplication technology paints a persuasive picture:

Cost. FalconStor’s solution is architected for flexibility, manageability, and high availability—delivering cost efficiency. Its multi-node, clustered appliance implementation allows for compatibility with an array of existing infrastructure components, including server platforms, network connectivity, and disk and tape storage—lowering configuration and sustainability costs. Scaling the solution is streamlined and offers greater manageability over alternative solutions that scale with additional standalone units. A single VTL/deduplication footprint is less costly to manage than several islands of deduplication. Its high availability cluster also reduces the potential costs associated with downtime.

Integration with existing backup applications. VTLs, by definition, seamlessly integrate (no changes are required to backup processes), appearing as a tape device to the backup application. However, improvements in backup processes can be made by leveraging more virtual tape devices than actual physical ones—increasing performance. Tracking duplicate media copies (created by device-to-device replication over IP or Fibre Channel, or policy-based physical media creation/export) with catalogue consistency can be accomplished via integration with NDMP4 or Symantec OpenStorage technology (OST), or by hosting backup application media server software on the FalconStor VTL.

Impact on backup and recovery performance. Disk-based backup delivers significant performance improvements over direct-to-tape-based approaches. FalconStor VTL and deduplication wrings out additional performance improvements in several ways:

- FalconStor claims performance throughput of 1.5 GB/second per node. Adding nodes to the configuration can improve aggregate performance. Restore speed of 1.2 GB/second per node from deduplicated virtual tape can similarly scale.
- FalconStor VTL enables deduplication, replication, and media creation outside the backup window. Post-process deduplication at 0.5 GB/second per node with parallel replication (over IP and/or FC) accelerates time to disaster recovery.
- The company’s Backup Accelerator option’s host-resident write-splitting technology continuously or periodically copies data to the same storage pool as the VTL, which serves as the data source for the backup application. Backup processing is offloaded from the production environment, eliminating the backup window altogether.

- FalconStor VTL's ability to host the backup application on the same system speeds backup and recovery by eliminating a "hop" in the backup data flow. By maintaining high-speed access between media server and VTL via the bus interface, backup and recovery performance is accelerated.

Ease of implementation/use. There are a few flexible deployment options for FalconStor VTL with deduplication—all with the same interface and features. A ready-to-deploy physical storage appliance delivers plug-and-play convenience. In VMware server virtualization environments, a pre-installed, pre-configured software application packaged as a virtual appliance enables quick and easy deployment, maximizes existing storage investments, and reduces infrastructure costs and complexity. A software appliance allows for the same plug-and-play convenience, but provides the added flexibility of leveraging existing hardware. After installation, global or backup set-specific policies are easily established for media duplication/export, encryption, deduplication, and more.

Scalability. Clustering enables scale—both in capacity and compute cycles—as needed. The VTL can scale up to 6 PB of data storage per node, while the deduplication component scales up to 298 TB of managed capacity per node. For throughput performance, the VTL scales up to an eight-node cluster with a single point of management, delivering up to 12 GB/second data ingest rate and managing up to 32 PB of data in a single cluster. Similarly, its deduplication scales up to an 8+1 cluster with a single point of management, delivering up to 4 GB/second of data deduplication speed and over 2 PB of managed capacity.

The Bottom Line

VTL technology is helping to modernize backup infrastructures, improving the performance and reliability of backup and recovery. Data deduplication identifies and eliminates redundant data, which dramatically improves the value proposition of disk-based data protection—a "must-have" today due to the current economic conditions forcing IT organizations to look at more efficient ways of managing data.

While the benefits of VTLs and deduplication are undeniable, IT organizations are often challenged to determine which offerings are best-suited for their environments. Technology maturity varies considerably and the vendor landscape is in flux, with new entrants and a spate of acquisitions. Vendors have focused marketing messages around technical nuances such as the reduction ratio, in-line vs. post-processing, and the granularity of deduplication; however, end-users are clearly more interested in compatibility, ease-of-implementation and use, scalability, manageability, performance, and the cost of solutions.

While the multitude of choices in FalconStor's portfolio can sometimes seem overwhelming, the company has streamlined its messaging to help end-users navigate its solution sets. In addition, the company is focused on demonstrating best-practice scenarios and validating real-world configurations. FalconStor's VTL with deduplication is a mature offering with proof points to end-users' most important evaluation criteria. By addressing the most relevant issues for end-users, FalconStor has been able to differentiate itself from the rest of the vendors in its class.