



Archive *Before* Backup — EMC *Centera*'s Prescription for the Smaller Enterprise

Analyst: Anne MacFarland

Management Summary

When faced with an onslaught of visitors, some people quickly cram the clutter in closets, and deal with reclaiming access to their wardrobe after the event has passed. Enterprises should not do this with their stored data. Yet, for the smaller enterprise, this happens more than they would care to admit. A crisis hits and - in the name of archiving, perhaps - CDs are burned - and surely there is a back-up copy somewhere. Weeks go by, and *ad hoc* becomes *de facto*.

The prime imperative for information management has always been some form of physical replication or backup, since electronically-stored information is inherently fragile. **Now that businesses of all sizes are transforming more of their processes into the more frictionless speed of electronic workflows, it is not just the application, but also the business, that is imperiled by data loss.** That fear has turned us all into packrats, saturating our email quotas and saving copies of fixed information to various directories, *just in case*. This static information, and its multiple copies, exacerbates the glut. **The answer is to archive with the same regularity, and on the same policy basis, as you back up your data. It should be routine. It should be automated.**

Use, but do not confuse, the functions of backup and archive. Backup, whether to disk or tape, and the accessory tools of snapshots and synchronous/asynchronous replication, should be for the restoration - of files, of applications, and of business units or businesses as a whole. Recovery is what these tools were developed to optimize.

Retrieval of static information should be a function of archiving. Archived data, once replicated (preferably remotely), does not have to be backed up. If you archive more production data, you backup less production data. Moreover, the more inactive data you archive, the more application performance can improve. The world of business data becomes a more rational place.

Secondary storage, including tapes, can be used as an archive, but to be useful to most enterprises, an archive must be online; must have the ability to guarantee data retention, authenticity, and integrity; must track usage; and must assist in the deletion of files as that becomes appropriate. To address these demands from their smaller customers, EMC has introduced a four-node configuration of its *Centera* Content Addressed Storage. If you wanted an archive but felt the smallest *Centera* was too big, this is good news for you. For more details, read on.

IN THIS ISSUE

➤ The Role of an Archive in the Small Enterprise	2
➤ The Centera Solution.....	2
➤ Centera Four-node as a Key Element of Solutions	3
➤ Conclusion	3

The Role of an Archive in the Smaller Enterprise

Much of business information does not change, yet it is still valuable, and there have been good reasons to keep it on primary storage. However, this is expensive and is often difficult to justify financially.

The traditional “other place” to keep information was at the far end of a backup process. “Let’s keep our backup tapes longer and use them as an archive” seems like a rational approach, but this turns out not to be the case. Retrieving information from backups is a cumbersome process, because, it is neither structured nor formatted for a simple find. Archived information, to be of the most benefit to the enterprise, must be immediately available, and of assured authenticity. Backup technology meets neither of these criteria.

Another replication approach that organizations have tried to use as a form of archiving is snapshots, but this approach also has fundamental flaws. Snapshots are incremental of point-in-time views of an information base. They are great for recalling versions of data prior to corruption, but they do nothing to ensure the continuous authenticity of the information they point to, and represent an inefficient, time-consuming way to recall static data.

Simply transferring data to secondary storage, burning CDs, or using tape requires more management and equipment to keep the data findable and available. Using a disk-based storage solution designed specifically for long-term retention and retrieval of business data is a good, and far simpler, solution.

The Centera Solution

Centera’s four-node configuration is a lower-capacity archive solution for the mid-sized enterprise. As an internally-mirrored¹ repository of fixed content, Centera does not have to be backed up, though customers may want to replicate the data to a remote site for disaster recovery. As well as being an archive

¹ If a data object becomes corrupted, there is a mirrored copy that can be accessed and then replicated to ensure data integrity. Parity protection is available on the larger 8-node Centera.

Centera 4-Node Quick Facts

- The four-node configuration is a full-function Centera, just of a lower capacity and at a lower price.
- It is targeted at enterprises that have the same needs to archive as larger enterprises, but a smaller quantity of information to be archived.
- It leverages an enhancement in the current version of the operating environment software, *storage on access nodes*, which increases the maximum quantity of information that can be stored per configuration. (Note that this is a feature available for all Centera configurations.)
- It is rack mountable. It can be placed in an EMC rack or other industry standard cabinets.
- Like all other Centera configurations, the four-node is self-configuring, healing, and managing, reducing the amount of management required and eliminating the need for emergency servicing.

solution, Centera can help optimize your production environment by removing the inactive data to reduce the bloat of full back-ups.

Centera, as content addressed storage, uses a globally unique, location-independent content address for each piece of unique content. These are, in essence, *digital fingerprints*, and the only reference needed to store and retrieve information. The use of content addresses and the storing of pertinent metadata along with the information results in an archive that:

- Eliminates management overhead
- Improves storage efficiency, because information is stored only once
- Provides absolute assurance of content authenticity, and
- Enforces any application policies intrinsic to storage.

The Centera four-node archive, like larger Centera configurations, can be used for one application, multiple applications, or as a central archive.

Centera Four-node as a Key Solution Element

Centera provides an effective archive for email archiving, enterprise content management, medical imaging, and file systems. It can be accessorized to become a solution for a variety of other archiving needs.

For email archiving, Centera creates a virtually infinite mailbox that eliminates the need for *pst* or *nsf* folders. It helps organizations ensure that their retention and disposition policies are met. It makes e-discovery easy.

In the health care industry, Centera is a popular target for both PACS (Picture Archiving Communication System) applications that manage x-rays, MRIs and other images, and for Hospital Information Systems. Some hospitals also use Centera to archive email, electronic patient records, etc. For small hospitals, Centera can be especially inviting because the skill set to support and troubleshoot tape and optical technologies is economically unsupportable.

Centera makes for an excellent long-term investment. It has almost no carrying cost to it, said Robert Terdeman, CTO and Senior Vice President, of Rogers Medical Intelligence Solutions, a large multinational based in New York City, with data hubs in New York and London. While this is a much bigger enterprise, low carrying costs are of value to an enterprise of any and every size.

Go-to-Market Plans

Resellers and systems integrators will use Centera, adding industry-specific applications, to provide resilient archive solution packages. Data service providers will be able to add Centera-based archiving to their offerings. Expect this new Centera offering to be available from a wide variety of sources.

Conclusion

An active archive for static business information does not just streamline production environments; it streamlines the organization's entire ecosystem. By offloading static information to a self-managing device like Centera, storage management administration can be focused on the

active data of an organization, where the rate of change and challenges of business require more decisions for which the policy is not as cut and dried. The Centera four-node configuration can quell many pains and let the smaller enterprise focus its resources more on non-routine demands. Think of how a Centera-based archive, with online access and assured information authenticity, could improve customer services, create new revenue streams, and make an organization work better.



About The Clipper Group, Inc.

The Clipper Group, Inc., is an independent consulting firm specializing in acquisition decisions and strategic advice regarding complex, enterprise-class information technologies. Our team of industry professionals averages more than 25 years of real-world experience. A team of staff consultants augments our capabilities, with significant experience across a broad spectrum of applications and environments.

- ***The Clipper Group can be reached at 781-235-0085 and found on the web at www.clipper.com.***

About the Author

Anne MacFarland is Director of Infrastructure Architectures and Solutions for The Clipper Group. Ms. MacFarland specializes in strategic business solutions offered by enterprise systems, software, and storage vendors, in trends in enterprise systems and networks, and in explaining these trends and the underlying technologies in simple business terms. She joined The Clipper Group after a long career in library systems, business archives, consulting, research, and freelance writing. Ms. MacFarland earned a Bachelor of Arts degree from Cornell University, where she was a College Scholar, and a Masters of Library Science from Southern Connecticut State University.

- ***Reach Anne MacFarland via e-mail at Anne.MacFarland@clipper.com or at 781-235-0085 Ext. 128. (Please dial “128” when you hear the automated attendant.)***

Regarding Trademarks and Service Marks

The Clipper Group Navigator, The Clipper Group Explorer, The Clipper Group Observer, The Clipper Group Captain's Log, The Clipper Group Voyager, and “*clipper.com*” are trademarks of The Clipper Group, Inc., and the clipper ship drawings, “*Navigating Information Technology Horizons*”, and “*teraproductivity*” are service marks of The Clipper Group, Inc. The Clipper Group, Inc., reserves all rights regarding its trademarks and service marks. All other trademarks, etc., belong to their respective owners.

Disclosure

Officers and/or employees of The Clipper Group may own as individuals, directly or indirectly, shares in one or more companies discussed in this bulletin. Company policy prohibits any officer or employee from holding more than one percent of the outstanding shares of any company covered by The Clipper Group. The Clipper Group, Inc., has no such equity holdings.

Regarding the Information in this Issue

The Clipper Group believes the information included in this report to be accurate. Data has been received from a variety of sources, which we believe to be reliable, including manufacturers, distributors, or users of the products discussed herein. The Clipper Group, Inc., cannot be held responsible for any consequential damages resulting from the application of information or opinions contained in this report.