Lab Validation Report

Catalogic ECX Copy Data Management

Designed to Intelligently Catalog, Manage, and Protect Data

By Vinny Choinski, Senior Lab Analyst and Tony Palmer, Senior Lab Analyst

October 2014

© 2014 by The Enterprise Strategy Group, Inc. All Rights Reserved.
Contents

Introduction .................................................................................................................................................. 3
Background .................................................................................................................................................. 3
Catalogic Copy Data Solutions .................................................................................................................. 4

ESG Lab Validation ..................................................................................................................................... 5
Catalogic ECX – Copy Data Management .................................................................................................. 5

ESG Lab Validation Highlights .................................................................................................................. 14
Issues to Consider ....................................................................................................................................... 14
The Bigger Truth ......................................................................................................................................... 15
Appendix .................................................................................................................................................... 16

ESG Lab Reports

The goal of ESG Lab reports is to educate IT practitioners about data center technology products for companies of all types and sizes. ESG Lab reports are not meant to replace the evaluation process that should be conducted before making purchasing decisions, but rather to provide insight into these emerging technologies. Our objective is to go over some of the more valuable feature/functions of products, show how they can be used to solve real customer problems and identify any areas needing improvement. ESG Lab's expert third-party perspective is based on our own hands-on testing as well as on interviews with customers who use these products in production environments. This ESG Lab report was sponsored by Catalogic Software.

All trademark names are property of their respective companies. Information contained in this publication has been obtained by sources The Enterprise Strategy Group (ESG) considers to be reliable but is not warranted by ESG. This publication may contain opinions of ESG, which are subject to change from time to time. This publication is copyrighted by The Enterprise Strategy Group, Inc. Any reproduction or redistribution of this publication, in whole or in part, whether in hard-copy format, electronically, or otherwise to persons not authorized to receive it, without the express consent of The Enterprise Strategy Group, Inc., is in violation of U.S. copyright law and will be subject to an action for civil damages and, if applicable, criminal prosecution. Should you have any questions, please contact ESG Client Relations at 508.482.0188.
Introduction

Catalogic Software’s goal is to improve organizations’ operational efficiency by leveraging cataloging and instant data access for restore, disaster recovery, test/dev, and big data analytics across physical, virtual, and cloud environments. This ESG Lab report focuses on Catalogic Software’s Copy Data Management platform and the technology that comprises it: copy data cataloging and search across the entire enterprise; transparent integration with NetApp and VMware snapshot technologies; detailed reporting; and protection/recovery policy workflows for data copy management.

Background

The demands on organizations to manage their data and storage requirements continue while the problem becomes increasingly complex due to data growth, virtualization, and the introduction of other emerging technologies. ESG’s 2014 IT Spending Intentions Survey revealed that, as they have been since 2010, increased use of server virtualization, improved backup and recovery, and data growth management are all in the top ten most frequently cited 2014 IT priorities. The adoption of server virtualization is nearly ubiquitous among enterprise and midmarket organizations, and increased usage of the technology was identified by 32% of respondents as one of their most important IT priorities for 2014.¹ Given this trend, it is reasonable to expect these priorities to continue to be of importance through 2015 and beyond.

Figure 1. Top-ten Most Important IT Priorities for 2014

<table>
<thead>
<tr>
<th>Top 10 most important IT priorities over the next 12 months. (Percent of respondents, N=562, ten responses accepted)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Increased use of server virtualization</td>
</tr>
<tr>
<td>Information security initiatives</td>
</tr>
<tr>
<td>Improve data backup and recovery</td>
</tr>
<tr>
<td>Manage data growth</td>
</tr>
<tr>
<td>Desktop virtualization</td>
</tr>
<tr>
<td>Use cloud infrastructure services</td>
</tr>
<tr>
<td>Regulatory compliance initiatives</td>
</tr>
<tr>
<td>Major application deployments or upgrades</td>
</tr>
<tr>
<td>Business intelligence/data analytics initiatives</td>
</tr>
<tr>
<td>Data center consolidation</td>
</tr>
</tbody>
</table>


© 2014 by The Enterprise Strategy Group, Inc. All Rights Reserved.
Catalogic Copy Data Solutions

Catalogic Software provides intelligent copy data solutions designed to deliver access to all business operational data through instantly mountable snapshots to maximize efficiency and availability. Catalogic divides this task into two essential categories, Copy Data Management and Copy Data Services, visually summarized in Figure 2.

Copy Data Management’s role is to catalog and index all of an organization’s copy data and provide deep, actionable insight. Copy Data Services leverages capabilities built into storage and virtualization platforms to do the heavy lifting, providing nearly instant backup and recovery using snapshots, vaulting, replication, and archiving. The following is a brief summary of the attributes of Catalogic’s Copy Data Management and Copy Data Services platforms:

- **Catalogic ECX**: Catalogic ECX is designed to deliver comprehensive copy data management for backup, DR, test/dev, and big data analytics. ECX provides instant visibility, operational insight, and business control of all copies of data in NetApp and VMware environments to improve business agility. By cataloging all tiers of NetApp and VMware snapshots, an organization can locate any file or object across the data center quickly, identify stale or unwanted data, validate protection compliance, and create protection/recovery policy workflows to take action on this data.

- **Catalogic DPX**: Catalogic DPX is engineered to protect all of an organization’s data quickly and reliably in virtual, physical, and cloud-based environments, irrespective of the primary storage system the data is hosted on. Catalogic DPX integrates copy data services like backup, disaster recovery, and bare metal recovery to reduce time to backup and recovery using instantly mountable snapshots to simplify data protection and reduce operational costs.
ESG Lab Validation

This report covers testing and analysis of the ECX Copy Data Management platform offering from Catalogic Software. ESG Lab performed hands-on evaluation and testing of Catalogic ECX at a NetApp facility located in Sunnyvale, California.

Catalogic ECX – Copy Data Management

The test bed was designed to simulate a corporate data center. As seen in Figure 3, multiple NetApp FAS systems are attached via iSCSI to physical and virtual servers. The virtual servers, running under a VMware ESX hypervisor on an industry-standard X64 server, hosted Microsoft Exchange, Microsoft SQL Server, and Microsoft file services as well as the Catalogic ECX virtual appliance.

Figure 3. The ESG Lab Test Bed

ESG Lab Testing

Catalogic ECX is distributed as a virtual appliance to minimize installation effort. In less than 15 minutes, ESG Lab had the appliance up and running on the network. Administrators can then immediately begin registering systems and cataloging data. ESG Lab logged into the dashboard of the Catalogic ECX appliance in our test bed.
As seen in Figure 4, the ECX dashboard provides an at-a-glance view of an organization’s data environment, including details such as the total number of objects cataloged—files, volumes, snapshots, and mirrors or clones of volumes are all cataloged as object types—as well as the status of scheduled cataloging jobs and any scheduled reports.

**Figure 4. The Catalogic ECX Dashboard**

ESG Lab clicked on Configure, which brought up a list of all the systems added to the environment. Catalogic can catalog data from multiple types of systems including NetApp FAS systems and VMware. Right-clicking on NetApp in the Resource Browser allowed ESG Lab to add a new FAS8000 Series system to be cataloged.

**Figure 5. Adding a Resource**
As Figure 6 shows, Catalogic ECX requires just an IP address to connect to the resource, and an administrator password to gain access to it. An administrator can enter a human-friendly name for each resource, which is especially helpful in larger environments.

**Figure 6. Entering Resource Details**

Once a resource is added, even before running a catalog policy, ECX instantly gives visibility into high-level configuration and statistics, seen in Figure 7.

**Figure 7. FAS8000 Volume List**

As Figure 8 shows, clicking on a volume name gives a detailed view of the volume including provisioning type, capacity utilization, file quotas, number of files, Qtrees, and snapshots.
The next step for an administrator is to create a policy for cataloging objects in the resource. A high-level policy catalogs volumes, snapshots, and clones, whereas a file-level policy will examine each file in each container. With a file-level policy, ECX catalogs every file, so if an organization keeps ten snapshots of a volume with one million files, a total of 11 million files will be cataloged. To minimize overhead, and to improve efficiency, Catalogic uses proprietary incremental methods to catalog only what’s changed between snapshots.

A high-level policy enables organizations to see all the copies of an object in a system and make intelligent decisions regarding what to catalog at the file level. Additionally, if administrators are only interested in the current data set on a volume, they can also choose to create a new snapshot and perform a file-level catalog of only that data, thereby reducing the overall number of files that need to be cataloged.
Figure 9 shows the ECX policy creation wizard. Administrators are walked through creating a policy, start to finish, in just a few simple steps. ESG Lab created a high-level policy for the new FAS8000 system in the test environment.

**Figure 9. Creating ECX Catalog Policies for a NetApp FAS System**

The catalog job completed in seconds and all objects in the system were visible to ESG Lab. At this point, ESG Lab used the ECX search engine, shown in Figure 10, to locate objects in the environment.

**Figure 10. Searching the Data Center with Catalogic ECX**
The basic search window is deliberately sparse, and allows use of familiar search operations to quickly locate objects and files anywhere in the environment. ESG Lab entered the search term “*”, which brought up results for all cataloged objects almost instantaneously. As Figure 11 shows, multiple copies of a number of files were instantly identifiable. ECX also provides advanced search capabilities to assist administrators in narrowing down results to a specific type of object or location within the enterprise, or to hide duplicate objects, in order to only see the original versions. These search functions are highly useful for customers considering hybrid clouds where data is fluid by nature.

*Figure 11. Search Results*

ESG Lab simulated typical use cases using advanced search, including searches for stale data, unused copies of data, copies spread across the entire environment, and under- and over-protected data across multiple ESX clusters and multiple file clusters. Clicking on an object’s name brings up a detailed view showing how many copies, snapshots, mirrors, and SnapVaults exist, as well as actual CIFS or NFS links to the files. These links, which provide instantaneous access to the data, have multiple use cases including: recovery, DR, test/dev analytics, and data spillage/leakage containment.
Next, ESG Lab used the Catalogic closed-loop automated workflow (or C.L.A.W.) methodology to create a VMware protection policy. As shown in Figure 12, the Lab selected the plan tab from the ECX GUI and then chose the VMware protection policy option. The Lab navigated through the four-step (e.g., Lifecycle, Notification, Schedule, and Submit) wizard to create a workflow. Workflows can be created with a VMware or storage focus. An administrator can create a policy for a folder, ESX cluster, or datastore, and then every VM that is created under that object inherits the policy and is automatically protected.

![Figure 12. Creating a VMware Protection Policy](image)

The Lab simply right-clicked on the primary storage box and selected “add a snapshot” to create a protection policy workflow. The snapshot configuration parameters were presented as check boxes and drop-down options. It should be noted that when creating snapshots, the administrator needs do nothing other than choose whether it should be simply crash consistent or application/file system consistent. Many users elect to do crash consistent snaps through the day, and an application consistent snap nightly. It should be noted that it is possible to mix and match retentions for these sub-policies. As an example, users can create multiple daily snapshots and retain them for 30 days, while retaining weekly snapshots for three months. No provisioning is required on the NetApp side at all.

The left side of Figure 12 shows the available storage lifecycle options such as snapshots, mirrors, and vaults. The right side of Figure 12 shows the VMware object that can be selected for protection because we selected a VMware-focused workflow. When a storage-focused workflow is used, a list of datastores will be presented for protection.

Adding mirrors and vaults uses the same simple right-click process the Lab used for the creation of a snapshot. Configuration parameters are again presented as check boxes and drop-down options. Only viable parameters are presented and only devices that have been cataloged are configurable. This process facilitates reliable and repeatable business operations by automating procedures and eliminating the complex and error-prone tasks of manually maintaining virtual machine to storage associations and developing custom scripts for copy data workflow management. Again, no provisioning is required on the NetApp side at all.
Next, ESG Lab initiated a VMware protection policy and, using the ECX GUI, monitored its status. The VMware protection policy workflow was designed to create a crash consistent snapshot for one group of virtual machines and an application consistent snapshot for another group of virtual machines. The monitor view available in the ECX GUI provided real-time status and detailed information about the workflow resources including VM system information, datastores, back-end storage details, and the number and frequency of snapshots. The same information was simultaneously stored in the ECX catalog for historical retrieval and analytics.

As shown in Figure 13, the Lab also initiated and monitored a VMware DR recovery policy workflow. The recovery leveraged a previous protection workflow snapshot to instantly virtualize (IV) and power on two protected VMs in an existing ESX cluster. As shown in the blue callout box on the right side of Figure 13, the recovery of the two VMs took approximately 48 seconds and, unlike most traditional protection solutions, the recovery information was also stored in the ECX catalog. It should be noted that because the IV process leverages storage snapshots, there will be very little increase in completion time as VMs are added to the workflow.

![Figure 13. Recovery Workflow Monitoring](image)

The flexibility provided to create a recovery workflow plan enabled ESG Lab to quickly and easily define key attributes. These attributes included the ability to easily select a recovery point, set up independent VLANs for different recovery scenarios, and use a drag-and-drop feature to adjust the boot order of the recovery VMs.

After the successful completion of the VMware recovery policy workflow, ESG Lab connected to the ESX cluster to confirm access to the restored VMs. At that time, the VMs were booted and running from a NetApp protection snapshot, and could be used for purposes like granular restore, data protection validation, test and development, or analytics. At this point, and as shown in the blue callout box in the middle of Figure 13, the administrator has two options available: End IV or Rapid Return to Production (RRP). If the recovery was used for temporary purposes like test and development or analytics, the End IV process can be selected to automatically stop the recovery and clean up all storage and VMware dependencies. If the IV recovery was a validation step before full restore, the RRP process can be selected to automatically and seamlessly move the VMs back to production on the desired infrastructure.

© 2014 by The Enterprise Strategy Group, Inc. All Rights Reserved.
Lastly, ESG Lab explored the analytics and reporting capabilities of ECX. Figure 14 shows a key chart from the VMware and datastore compliance report. The report shows that in this environment, two of our seven VMs and one of two datastore workflows were out of compliance with regard to protection. In addition, the Lab audited the Protection Usage report, which shows the cost of data protection for the environment by reporting on the capacity utilization of snaps, mirrors, and vaults associated with it.

**Figure 14. Catalogic ECX VMware and Datastore Compliance Report**

![Figure 14](image)

### Why This Matters

According to ESG research, two of the top IT priorities for 2014 are increasing the use of server virtualization and managing data growth. Traditional enterprise data management solutions are complex systems that require high levels of expertise to integrate and use. Organizations need simpler solutions to catalog, manage, and automate business operations around the growing volumes of data in their environments.

As virtualization grows, so does the number of orphaned snapshots and VMs. These are objects that are no longer being used but are taking up space and often degrading overall system performance. Also, without proper visibility, it becomes almost impossible to determine if an environment is over or under protected. Administrators need to be able to optimize copies of data retained for business purposes while locating and removing stale, unused, and unnecessary copies as well as be able to leverage existing copies for multiple use cases.

ESG Lab has confirmed that Catalogic ECX was easy to set up and use for copy data management, with the ability to quickly add and index storage systems with a few clicks, locate all copies of files and volumes across the enterprise, automate typical business operational workflows, and leverage that copy data to efficiently meet multiple business needs.

---

ESG Lab Validation Highlights

- Catalogic ECX was easy to set up and use, with the ability to quickly add and catalog storage systems and virtual environments. No agents were required to deploy ECX and installation was completed in minutes with just a few clicks of the mouse.
- ESG Lab was able to locate all copies of specific files, volumes, and objects across multiple systems and platforms rapidly using a simple, intuitive search interface.
- Catalogic ECX provided information on all data as well as protection/recovery policy workflows in the test environment via detailed reports. The reports were easy to run and interpret and helped make auditing the environment quick and easy.
- The Lab leveraged snap data in protection and recovery policy workflows in the test environment the way that an administrator would in the real world for purposes such as granular recovery, D/R, analytics, and test and development.
- We found the workflows intuitive to work with. Configuration was wizard driven with options presented in simple drop-down, click-to-select, and drag-and-drop functionality.
- Scheduling was also robust and easy to use and can eliminate the need to write and maintain custom scripts for copy data management.

Issues to Consider

- ESG Lab observed very rapid cataloging by Catalogic ECX in a controlled test environment. It is important to note that file-level cataloging will provide extremely rich detail for every object in a system, but will require more time to complete than high-level scans. ESG Lab spoke with a customer with a large deployment who cataloged one FAS system using both Catalogic and a competing product in a proof of concept. Both solutions accurately cataloged all objects in the system, but the competing product took more than a week to catalog all objects whereas Catalogic completed the catalog in two days. The target was a FAS6000 series HA pair with more than 90 TB of data and hundreds of millions of files.
The Bigger Truth

According to a recent ESG survey of IT spending intentions for 2014, increasing the use of server virtualization, managing data growth, and improving backup and recovery remain in the top five most-cited IT priorities for the year. Requirements for software that can provide effective tools for cataloging, managing, and protecting all the disparate data in large, complex physical and virtual environments will be a high priority.

These challenges are motivating IT to find solutions that can provide copy data management—the ability to catalog all tiers and types of data, quickly locate any file or object across the enterprise, identify stale data, validate protection compliance, define automated workflows, and leverage existing copy data for multiple business purposes by leveraging advanced storage capabilities native to solutions like NetApp FAS and VMware. In fact, ESG research indicates that 55% of data protection capabilities are due to the copies that people keep on disk/tape and their movement.

Catalogic ECX, combined with NetApp FAS storage systems, delivers highly scalable copy data management solutions that are easy to deploy, scale, and manage. Integration with NetApp snapshot, clone, and SnapVault technology enables block-level snapshots to be leveraged as base data set images, enabling near-instant point-in-time data access to critical applications and files without transferring data.

ESG Lab installed and configured Catalogic ECX through an easy-to-understand web interface in minutes, then added and indexed storage systems with just a few mouse clicks. Using ECX search, all copies of files and volumes across multiple systems and platforms were located in a fraction of a second, and protection/recovery policy workflows allowed existing copy data to be quickly and easily leveraged for multiple business purposes like granular restores, DR, analytics, and test and development.

Catalogic’s simple-to-use management interfaces and integration with both NetApp and VMware provides a seamless solution for managing all copies of an organization’s data in highly virtualized environments. ESG Lab was quite impressed with Catalogic’s ability to leverage a tight integration with NetApp and VMware for cataloging, search, detailed reporting, and copy data workflow management.

It is clear that Catalogic is continuing to execute on its vision to provide users with a comprehensive copy data solutions infrastructure. In our opinion, Catalogic ECX enables enterprises to reduce costs, simplify infrastructures, and do more with the copy data they already have.

---

# Appendix

**Table 1. Test Environment Details**

<table>
<thead>
<tr>
<th>Catalogic</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Catalogic ECX Virtual Appliance</td>
<td>Version 2.0</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Servers, Virtualization, Operating System, Applications</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>IBM Blade Server H</td>
<td>Windows 2008 R2</td>
</tr>
<tr>
<td>VMware ESX, Version 5.1</td>
<td>SQL Server 2008</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Storage</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>NetApp FAS3210</td>
<td>Data ONTAP 8.1</td>
</tr>
<tr>
<td>NetApp FAS8020</td>
<td>Data ONTAP 8.1</td>
</tr>
<tr>
<td>NetApp Virtual FAS</td>
<td>Data ONTAP 8.1</td>
</tr>
</tbody>
</table>