Redbooks

ibm.com/redbooks

IT Modernization using Catalogic ECX Copy Data Management and IBM Spectrum Storage

Christian Burns

Peter Eicher

Kamlesh Lad

Prashant Jagannathan

Jon Tate

Storage

IBM

Redbooks

International Technical Support Organization

IT Modernization using Catalogic ECX Copy Data Management and IBM Spectrum Storage

April 2016

SG24-8341-00

Note: Before using this information and the product it supports, read the information in “Notices” on page vii.

First Edition (April 2016)

This edition was produced using Version 2.3 of Catalogic ECX, Spectrum Virtualize 7.5.0.3, Spectrum Protect Snapshot for Virtual Environments 4.1.3.0.

Contents

Notices

This information was developed for products and services offered in the US. This material might be available from IBM in other languages. However, you may be required to own a copy of the product or product version in that language in order to access it.

IBM may not offer the products, services, or features discussed in this document in other countries. Consult your local IBM representative for information on the products and services currently available in your area. Any reference to an IBM product, program, or service is not intended to state or imply that only that IBM product, program, or service may be used. Any functionally equivalent product, program, or service that does not infringe any IBM intellectual property right may be used instead. However, it is the user's responsibility to evaluate and verify the operation of any non-IBM product, program, or service.

IBM may have patents or pending patent applications covering subject matter described in this document. The furnishing of this document does not grant you any license to these patents. You can send license inquiries, in writing, to:
IBM Director of Licensing, IBM Corporation, North Castle Drive, MD-NC119, Armonk, NY 10504-1785, US

INTERNATIONAL BUSINESS MACHINES CORPORATION PROVIDES THIS PUBLICATION “AS IS” WITHOUT WARRANTY OF ANY KIND, EITHER EXPRESS OR IMPLIED, INCLUDING, BUT NOT LIMITED TO, THE IMPLIED WARRANTIES OF NON-INFRINGEMENT, MERCHANTABILITY OR FITNESS FOR A PARTICULAR PURPOSE. Some jurisdictions do not allow disclaimer of express or implied warranties in certain transactions, therefore, this statement may not apply to you.

This information could include technical inaccuracies or typographical errors. Changes are periodically made to the information herein; these changes will be incorporated in new editions of the publication. IBM may make improvements and/or changes in the product(s) and/or the program(s) described in this publication at any time without notice.

Any references in this information to non-IBM websites are provided for convenience only and do not in any manner serve as an endorsement of those websites. The materials at those websites are not part of the materials for this IBM product and use of those websites is at your own risk.

IBM may use or distribute any of the information you provide in any way it believes appropriate without incurring any obligation to you.

The performance data and client examples cited are presented for illustrative purposes only. Actual performance results may vary depending on specific configurations and operating conditions.

Information concerning non-IBM products was obtained from the suppliers of those products, their published announcements or other publicly available sources. IBM has not tested those products and cannot confirm the accuracy of performance, compatibility or any other claims related to non-IBM products. Questions on the capabilities of non-IBM products should be addressed to the suppliers of those products.

Statements regarding IBM's future direction or intent are subject to change or withdrawal without notice, and represent goals and objectives only.

This information contains examples of data and reports used in daily business operations. To illustrate them as completely as possible, the examples include the names of individuals, companies, brands, and products. All of these names are fictitious and any similarity to actual people or business enterprises is entirely coincidental.

COPYRIGHT LICENSE:

This information contains sample application programs in source language, which illustrate programming techniques on various operating platforms. You may copy, modify, and distribute these sample programs in any form without payment to IBM, for the purposes of developing, using, marketing or distributing application programs conforming to the application programming interface for the operating platform for which the sample programs are written. These examples have not been thoroughly tested under all conditions. IBM, therefore, cannot guarantee or imply reliability, serviceability, or function of these programs. The sample programs are provided “AS IS”, without warranty of any kind. IBM shall not be liable for any damages arising out of your use of the sample programs.

Trademarks

IBM, the IBM logo, and ibm.com are trademarks or registered trademarks of International Business Machines Corporation, registered in many jurisdictions worldwide. Other product and service names might be trademarks of IBM or other companies. A current list of IBM trademarks is available on the web at “Copyright and trademark information” at http://www.ibm.com/legal/copytrade.shtml

The following terms are trademarks or registered trademarks of International Business Machines Corporation, and might also be trademarks or registered trademarks in other countries.

AIX®
DB2®
FlashCopy®
FlashSystem™
IBM FlashSystem®
IBM Spectrum™
IBM Spectrum Protect™
IBM Spectrum Virtualize™
IBM®
Redbooks®
Redbooks (logo)  ®
Storwize®
Tivoli®

The following terms are trademarks of other companies:

Adobe, the Adobe logo, and the PostScript logo are either registered trademarks or trademarks of Adobe Systems Incorporated in the United States, and/or other countries.

SoftLayer, and SoftLayer device are trademarks or registered trademarks of SoftLayer, Inc., an IBM Company.

Linux is a trademark of Linus Torvalds in the United States, other countries, or both.

Microsoft, Windows, and the Windows logo are trademarks of Microsoft Corporation in the United States, other countries, or both.

Other company, product, or service names may be trademarks or service marks of others.

Find and read thousands of IBM Redbooks publications

- Search, bookmark, save and organize favorites
- Get personalized notifications of new content
- Link to the latest Redbooks blogs and videos

Get the latest version of the Redbooks Mobile App

iOS **Download Now** Android

Promote your business in an IBM Redbooks publication

Place a Sponsorship Promotion in an IBM® Redbooks® publication, featuring your business or solution with a link to your web site.

Qualified IBM Business Partners may place a full page promotion in the most popular Redbooks publications. Imagine the power of being seen by users who download millions of Redbooks publications each year!

ibm.com/Redbooks
About Redbooks → Business Partner Programs

THIS PAGE INTENTIONALLY LEFT BLANK

Preface

Data is the currency of the new economy, and organizations are increasingly tasked with finding better ways to protect, recover, access, share, and use data. Traditional storage technologies are being stretched to the breaking point. This challenge is not because of storage hardware performance, but because management tools and techniques have not kept pace with new requirements. Primary data growth rates of 35% to 50% annually only amplify the problem.

Organizations of all sizes find themselves needing to modernize their IT processes to enable critical new use cases such as storage self-service, Development and Operations (DevOps), and integration of data centers with the cloud. They are equally challenged with improving management efficiencies for long established IT processes such as data protection, disaster recovery, reporting, and business analytics.

Access to copies of data is the one common feature of all these use cases. However, the slow, manual processes common to IT organizations, including a heavy reliance on labor-intensive scripting and disparate tool sets, are no longer able to deliver the speed and agility required in today's fast-paced world.

Copy Data Management (CDM) is an IT modernization technology that focuses on using existing data in a manner that is efficient, automated, scalable, and easy to use, delivering the data access that is urgently needed to meet the new use cases. Catalogic ECX, with IBM® storage, provides in-place copy data management that modernizes IT processes, enables key use cases, and does it all within existing infrastructure.

This IBM Redbooks® publication shows how Catalogic Software and IBM have partnered together to create an integrated solution that addresses today's IT environment.

Authors

This book was produced by the following team of specialists.

Christian Burns is a Senior Product Manager with Catalogic Software, with primary responsibility for ECX, Catalogic's software-defined copy data management platform. Before joining Catalogic, Christian was a Storage Solution Architect with IBM, where he worked with clients, IBM Business Partners, and IBM employees worldwide, designing and implementing storage solutions that included various IBM products and technologies. He has over 15 years of industry experience in the areas of sales engineering, solution design, and software development and product management. Christian holds a BA degree in Physics and Computer Science from Rutgers College.

Kamlesh Lad is a Technical Architect with the Research and Development group at Catalogic Software in the San Francisco Bay Area. Before joining Catalogic Software, he worked at CommVault Systems as a Software Engineering Manager. Kamlesh has over 20 years of experience with designing and developing storage management software.

Prashant Jagannathan is a Technical Manager at Catalogic with 10 years of experience in the IT industry. He is part of the Business Development team at Catalogic and is responsible for working with the technical teams of strategic business partners and identifying joint solutions and potential integration in products. Before Catalogic, Prashant was an APAC Team lead in Syncsort Inc, where he managed the technical sales operations in the Asia-Pacific region.

Peter Eicher is Senior Product Marketing Manager at Catalogic Software. He has worked in multiple facets of enterprise software for over 20 years, including data protection, security, and network management. He is a frequent contributor to industry discussions through blogging and trade journal publications.

Jon Tate was the Project Manager for this book. He is a Project Manager for IBM Storage at the International Technical Support Organization (ITSO), San Jose Center. Before joining the ITSO in 1999, he worked in the IBM Technical Support Center, providing Level 2/3 support for IBM mainframe storage products. Jon has 30 years of experience in storage software and management, services, and support, and he is an IBM Certified IT Specialist, an IBM SAN Certified Specialist, and Project Management Professional (PMP) certified. He is also the UK Chairman of the Storage Networking Industry Association.

Thanks to the following people for their contributions to this project:

Paul Bell
Raghuram Devarakonda
Jay Hamilton
Mitch Light
Ron So
Sathya Sankaran
Virendra Singh

Wonyoung Kim
Siddharth Bhatt
Catalogic Software

Bert Dufrasne
IBM Redbooks

Now you can become a published author, too!

Here's an opportunity to spotlight your skills, grow your career, and become a published author—all at the same time! Join an ITSO residency project and help write a book in your area of expertise, while honing your experience using leading-edge technologies. Your efforts will help to increase product acceptance and customer satisfaction, as you expand your network of technical contacts and relationships. Residencies run from two to six weeks in length, and you can participate either in person or as a remote resident working from your home base.

Find out more about the residency program, browse the residency index, and apply online at:

ibm.com/redbooks/residencies.html

Comments welcome

Your comments are important to us!

We want our books to be as helpful as possible. Send us your comments about this book or other IBM Redbooks publications in one of the following ways:

- Use the online **Contact us** review Redbooks form found at:

  ibm.com/redbooks
- Send your comments in an email to:

  redbooks@us.ibm.com
- Mail your comments to:

  IBM Corporation, International Technical Support Organization
  Dept. HYTD Mail Station P099
  2455 South Road
  Poughkeepsie, NY 12601-5400

Stay connected to IBM Redbooks

- Find us on Facebook:

  http://www.facebook.com/IBMRedbooks
- Follow us on Twitter:

  http://twitter.com/ibmredbooks
- Look for us on LinkedIn:

  http://www.linkedin.com/groups?home=&gid=2130806

- Explore new Redbooks publications, residencies, and workshops with the IBM Redbooks weekly newsletter:

  https://www.redbooks.ibm.com/Redbooks.nsf/subscribe?OpenForm

- Stay current on recent Redbooks publications with RSS Feeds:

  http://www.redbooks.ibm.com/rss.html

1

IT modernization, Copy Data Management, and Catalogic ECX

This chapter provides an overview of the IT modernization challenges facing many organizations today and how Copy Data Management can address those issues. It examines the merits of the different approaches to Copy Data Management available in the industry and provides an overview of Catalogic Software's ECX Copy Data Management solution.

This chapter includes the following sections:

- Introduction
- The shift to bimodal IT
- The copy data solution
- Catalogic ECX overview

1.1 Introduction

Data is the currency of the new economy, and organizations are increasingly tasked with finding better ways to protect, recover, access, share, and use it. Traditional storage technologies are being stretched to the breaking point, not because of storage hardware performance, but because management tools and techniques have not kept pace with new requirements. Primary data growth rates of 35% to 50% annually only amplify the problem.

Organizations of all sizes find themselves needing to modernize their IT processes to enable critical new use cases such as storage self-service, DevOps, and integration of data centers with the cloud. They are equally challenged with improving management efficiencies for long established IT processes such as data protection, disaster recovery, reporting, and business analytics.

Access to copies of data is the one common feature of all these use cases. However, the slow, manual processes common to IT organizations, including a heavy reliance on labor-intensive scripting and disparate tool sets, are no longer able to deliver the speed and agility required in today's fast-paced world.

Copy Data Management (CDM) is an IT modernization technology that focuses on using existing data in a manner that is efficient, automated, scalable, and easy to use, delivering the data access that is urgently needed to meet the new use cases. Catalogic ECX, with IBM storage, provides in-place copy data management that modernizes IT processes, enables key use cases, and does it all with existing infrastructure.

1.2 The shift to bimodal IT

The need to modernize the IT environment is cogently summed up in the concept of *bimodal IT*. Traditional IT, referred to as *Mode 1*, is focused around existing enterprise and mission critical systems where operational reliability is paramount. In these environments, change is complex, slow, and often cost prohibitive. The ultimate charter is to preserve and protect critical business assets, notably organizational data. In today's business world, this is a cost center under constant and intense scrutiny.

Newer, agile IT, referred to as *Mode 2*, is characterized by a high degree of innovation and dynamic environments that allow organizations to rapidly deliver new and modern services to a range of consumers, third parties, and business units. Typically, agile IT helps top line revenue, so it often has a large budget.

As an example of the shift from Mode 1 to Mode 2, consider the DevOps method of software development that requires rapid access to data by using programmable interfaces. Traditional models that depend on a "submit a request for data and wait days or weeks" process cannot support a DevOps methodology. As application developers become primary consumers of copies of production data sets, organizations must move away from traditional "slow" IT models and towards "fast" IT models that enable greater agility and responsiveness.

To remain competitive, organizations must find ways to modernize their traditional IT processes to enable newer, more agile use cases. All too often the Mode 2 groups in an organization drive their own IT infrastructure outside of data center control, which is sometimes referred to as *Shadow IT*. This structure is both inefficient and can create risk because these shadow IT resources might not have the same resiliency, recoverability, and security commonly provided for data center assets. But organizations struggle to bridge this

gap in operational modes because existing enterprise systems do not speak “the right language”.

The missing piece in this puzzle is an abstraction layer that enables the automated, self-service oriented provisioning of data sets for Mode 2 operations while allowing traditional IT teams to maintain governance and service level agreement (SLA) accountability. In effect, organizations require an “infrastructure as code” operating model, where they can spin up entire copies of production systems on demand without the bottlenecks typical of storage provisioning processes in traditional IT environments. Ideally, this system would make use of data center IT resources.

1.2.1 ECX bridges the bimodal IT gap

ECX bridges the bimodal IT gap through a RESTful API that provides an abstraction layer between the Mode 1 data center resources and the immediate access requirements of Mode 2. ECX frees users to consume traditional IT through an API and in an on-demand fashion that requires no intervention from the data center staff, but also allows them to maintain control. With this model, data users such as application developers or data analytics teams no longer must wait weeks for new environments to be made available. Copies of entire systems envelopes (the virtual machines and related storage that represent a specific application set) can now be available in a fully isolated environment, on demand, by using API calls.

This capability provides the following benefits to traditional IT:

- Infrastructure modernization without the need for “rip and replace”
- The ability to allow users to self-service without sacrificing resource control
- A data management interface into existing infrastructure that simplifies complexity through abstraction, automation, governance, and templates
- Broad-based automation that simultaneously drives down OpEx and increases agility
- Role-based access control (RBAC) that allows granting of access to a limited set of ECX capabilities across a limited set of IT resources

Application teams (such as DevOps) and business units (such as analytics) get “infrastructure as code” that provides these benefits:

- Timely access to application consistent infrastructure that can be called “on demand” by using whatever tools they see fit
- Self-service provisioning to access data whenever needed without the need to submit requests
- Hybrid cloud automation to automate moving data from the enterprise storage environment into the cloud, making it available to cloud compute resources
- DevOps automation through the ECX REST API

1.3 The copy data solution

As stated in the Introduction, the common factor across multiple IT modernization use cases is gaining access to copies of data. How one gains access can be considered the “copy data management challenge.” Typically, access to data (such as on an enterprise storage array) or systems that connect to data (such as a virtual machine) requires a lengthy request and

provisioning process passed through gatekeepers in the IT organization. Data consumers can wait days or even weeks to get the data they request, which can limit their effectiveness.

CDM technology addresses this challenge by bringing modern IT approaches including self-service, automation, and APIs to existing systems where the vast bulk of organizational data still resides. At the same time, CDM simplifies data protection and recovery, and provides the orchestration needed for reliable disaster recovery.

However, not every CDM approach is the same.

1.3.1 Approaches to CDM

History has shown that IT challenges can be addressed with varied approaches. Commonly, as new challenges are identified, vendors develop products to address those challenges that differ greatly in both their complexity and value. When considering the CDM solutions from a holistic perspective, several factors must be considered:

- Ease of implementation
- Up front and ongoing costs (CapEx and OpEx) to maintain the solution
- Ability to use existing technology investments
- Quantity and location of existing data copies, and the ability to use those copies

Off-host CDM

In CDM, a common approach has been to address the challenge by creating redundant systems (both hardware and software) to manage copies. Although on the surface this approach appears to address the challenge of managing the creation and reuse of new data copies, it neglects to address the problem of pre-existing data copies present in primary storage systems. Further, it fails to use the investments organizations have already made in their existing infrastructure and the copy services capabilities they already include. Typically, a full replica of production data, or “gold copy”, must be moved to some new system. Lastly, such solutions are typically complex and expensive to deploy, and require organizations to learn and maintain entirely new (and redundant) systems. Such approaches are commonly referred to as Copy Data Management 1.0.

In-place CDM

The ideal approach to addressing the CDM challenge is a solution that can deploy quickly and easily (think “software only”), integrate with current infrastructure, maximize the return on existing storage investments, and allow existing data copies to be used, all without the need to implement costly new hardware or migrate data. As such, this solution must be able to automate and orchestrate the copy services of the existing storage and hypervisor assets in an environment. This “in-place” approach is what is now known as CDM 2.0.

1.4 Catalogic ECX overview

An effective copy data solution begins with a comprehensive view of data across physical, virtual, and cloud-based environments. ECX is a storage management application that delivers automation, orchestration, self-service, and analytical tools for the data in your environment through a comprehensive catalog.

ECX catalogs all of the primary storage copy data in your supported IBM environments in one centralized management console. This process allows you to orchestrate, analyze, search, and report on all your data, allowing you to take full advantage of your data assets.

By cataloging all tiers of primary, snapshot, replica and VMware objects, search tasks are completed in seconds, enabling you to identify stale or unwanted data, ensure protection compliance, and much more.

Additionally, with an intuitive point-and-click interface, organizations can automate and orchestrate Copy Data and Use Data workflows. ECX's automated workflows allow organizations to streamline data copy operations for a number of business operations such as enhanced recovery, DR, forensics, test/dev, DevOps, and analytics.

Furthermore, ECX's storage workflow templates allow organizations to remove the provisioning bottlenecks that are typical of most IT environments. Application owners or VM administrators no longer must wait weeks for storage teams to provision new environments. With ECX, provisioning can be a self-service operation, performed on demand and as frequently as the business demands.

1.4.1 ECX solution architecture

ECX is built on a scalable and extensible platform and includes a full set of robust RESTful APIs for enterprise or managed service provider (MSP) integration. ECX is built based on a provider model. Each storage array or hypervisor that is supported is built as a provider and snaps into the architecture. This approach ensures that the user can drive all of the core ECX policies and services consistently across each supported platform.

In-place copy data management with ECX

Using a provider-based model makes ECX unique among the other CDM solutions available in the market today that require either a "rip and replace" approach or a redundant hardware infrastructure dedicated solely to managing data copies. ECX allows organizations to use their existing infrastructure investments while putting their data copies to use. With ECX, there is no need to purchase a redundant set of "copy services" functions (think snapshotting and remote copy). The primary storage systems already contain these capabilities. ECX just brings the necessary automation, orchestration, and self-service to those services to allow CDM in-place.

ECX allows the IT team to manage copy data within the existing production storage environment. They can use the APIs of both VMware and the storage systems to manage the creation and orchestrate the use of data copies (snapshots, replicas, and so on). This approach requires no new hardware, allowing for simplified deployments and superior return on investment (ROI) and return on assets (ROA).

Off-host copy data management with ECX

Although the fundamental architecture of ECX is based on an in-place model, with CDM performed directly on the existing storage system, customers with virtual environments on heterogeneous storage can still take advantage of ECX. ECX offers Copy Data Services for VMware, delivering the full value of ECX for all VMware environments. With this approach, ECX uses VMware VADP to capture images of the virtual machines, copying them from any production storage environment to an ECX-supported storage system, creating an off-host CDM environment. ECX uses VMware Changed Block Tracking (CBT) technology to efficiently keep the copies up to date. The full power of ECX allows the IT team to manage, orchestrate, and analyze the VM copies in support of the IT and business functions that rely on them.

1.4.2 Integration with IBM products

As mentioned in 1.4.1, “ECX solution architecture” on page 5, ECX uses a unique, in-place approach to Copy Data Management that uses a provider model for supporting specific storage and hypervisor platforms. ECX 2.2 introduced two new storage providers that expanded the supported system to include a significant portion of the IBM storage portfolio. Specifically, providers were added to support the following IBM products:

- Storage systems built with IBM Spectrum™ Virtualize software
- Spectrum Protect Snapshot (formerly IBM Tivoli® Storage FlashCopy® Manager)

Note: For more information about the versions of IBM products that are supported by ECX, see the *Catalogic ECX 2.3 User’s Guide.*

Storage systems built with Spectrum Virtualize software

ECX illustrates the advantage of a software defined architecture like the ones provided by IBM. Through a single provider, ECX is able to bring CDM capabilities to a large cross-section of the IBM storage portfolio. The following is a comprehensive list of IBM storage systems that are supported by Catalogic ECX:

- IBM SAN Volume Controller
- IBM Storwize® V7000, V7000F, and V7000 Unified* (*support for block volumes only)
- IBM Storwize V5000 and V5000F
- IBM Storwize V3500 and V3700
- IBM FlashSystem™ V9000
- IBM VersaStack

IBM SAN Volume Controller, the Storwize Family of Products, FlashSystem V9000, and IBM VersaStack are all powered by the same Spectrum Virtualize software, and as such, all of these products benefit from the same world class copy services capabilities of FlashCopy and Global Mirror. ECX is able to support in-place CDM across all of these products by delivering orchestration, automation, and self-service on top of these copy services through this single provider.

By combining the ECX software with existing infrastructure investments (the IBM storage systems and VMware vCenter servers/ESXi hosts already in place), organizations immediately gain the benefit of four new and powerful use cases:

- Self Service for Application Teams and VM Owners

  ECX enables storage administrators to create templates for the provisioning of new environments from data copies. In turn, application teams can use these templates to self-provision new environments and have faster access to the data they need to do their job. Robust RBAC ensure that application teams get access to the appropriate data sets. This access enables the application teams to gain speed while the IT team maintains control and accountability of the SLAs.

- Simplified Storage Management

  ECX provides template-based provisioning that makes heterogeneous storage management easy, repeatable, and scalable. There is no longer any need for complex scripting with high maintenance costs.

- Copy Data Orchestration and Automation

  ECX allows for the automation and orchestration of the storage and application consistent snapshots at the hypervisor layer. Most organizations already have application consistent copies of production environments in a separate location just waiting to be used. ECX enables those copies to be used for these use cases. among others:

  - Automated daily disaster recovery testing
  - Enhanced test/dev
  - Real-time data access for these capabilities, among others:
    - Operational recovery
    - Business analytics
    - Reporting

- Next Generation Protection and Recovery

  Traditional host-based backup models are no longer viable in many environments due to data growth and ever more stringent recovery needs. As a result, more organizations are turning to disk array protection technologies, using snapshot-and-replication methods to protect and recovery data. ECX provides complete management and cataloging for these environments, while also delivering instant mount and recovery.

- Hybrid Cloud

  Organizations continue to struggle to use enterprise data that uses limitless cloud compute power. ECX provides automation and orchestration for making array-based data available to cloud compute engines. This system allows for easy spin-up and spin-down of resources on demand.

- DevOps or “Infrastructure as Code”

  The DevOps model of application development promises greater speed and agility, but enterprises are often challenged with connecting DevOps methods to existing systems that contain valuable data. ECX provides a bridge between DevOps and data center IT resources though the use of cloud-ready template, roles-based access, and a complete REST API that allows all ECX functions to be easily programmable.

IBM Spectrum Protect Snapshot

IBM customers who run virtual environments on IBM storage commonly use IBM Spectrum Protect™ Snapshot to protect those virtual machines. Its rich support for application consistency and integration with hardware snapshot capabilities make it a valuable tool for backing up these virtual environments. For environments where IBM Spectrum Protect Snapshot is used to create application consistent FlashCopies of virtual machines running on any of the supported storage systems mentioned above, ECX can provide further added value. With ECX, these copies created with IBM Spectrum Protect Snapshot can now be used not just for backup, but for all the same use cases mentioned above. ECX makes this flexibility possible by cataloging the copies created by IBM Spectrum Protect Snapshot and making them available for use in Copy Data Use workflows.

2

Catalogic ECX deployment and usage

This chapter provides an overview of the deployment options for Catalogic ECX, highlights the advantages of those options, and describes the steps that are required to complete some of the more common configuration tasks. For comprehensive instructions on deploying, configuring, and using ECX, see the *Catalogic ECX 2.3 Quick Start Guide* and the *Catalogic ECX 2.3 User's Guide*, available on the Catalogic Software "MySupport" site at:

`https://mysupport.catalogicsoftware.com`

Note: Access to the Catalogic Software MySupport website requires a MySupport account. If you do not have a MySupport account, you can request one here:

`https://mysupport.catalogicsoftware.com/request.php`

This chapter includes the following sections:

- ECX installation
- ECX configuration and management by using the management portal

2.1 ECX installation

ECX offers flexible deployment options. The typical deployment model is to install ECX as a VMware virtual appliance. ECX can also be deployed in a clustered mode that provides enhanced application availability. Regardless of which deployment model is chosen, the full feature set of the product is available.

2.1.1 ECX deployment using VMware

Deployment of Catalogic ECX as a VMware virtual appliance is a quick and easy process. ECX is delivered as a single “OVA” file, and typical deployment time is under 15 minutes. After it is deployed, ECX immediately begins creating an actionable catalog of all of the IT environment’s copy data using the public APIs of the underlying infrastructure. IBM storage and VMware hypervisor assets are easily cataloged without the need to deploy any agents.

Virtual machine installation

Before deploying ECX to the VMware host, ensure that you have the following prerequisites:

- The correct OVF template, which has an OVA extension, and is approximately 1.2 GB
- vSphere 5.1, 5.5, or 6.0
- Network information and VMware host information
- Either an available static IP address to use or access to DHCP

For initial deployment, configure your virtual appliance to meet the following minimum requirements:

- 64-bit dual core machine
- 32 GB memory

The appliance has three virtual disks that total 380 GB storage:

- 30 GB for operating system and application, which includes 16 GB for the swap partition, 256 MB for the boot partition, and the remainder for the root partition
- 100 GB for configuration data related to policies, jobs, events, and logs
- 250 GB for catalog data

After the preceding requirements have been validated, ECX can be deployed.

Important: It is assumed that the reader has familiarity with deploying VMware virtual appliance templates. For detailed instructions on how to deploy the ECX OVA template, see the *Catalogic ECX 2.3 User’s Guide*.

Browser support

Access the ECX management portal from a computer that has access to the installed virtual appliance. The following web browsers are supported:

- Internet Explorer 11
- Microsoft Edge 20 or later
- Firefox 10 or later
- Chrome 19 or later

If your resolution is below 1024 x 768, some items might not fit in the browser window.

Note: Pop-up windows must be enabled in your browser to access the Help system and some ECX operations.

Virtual appliance configuration by using the VAMI

After the initial deployment of the ECX virtual appliance, the following one-time configuration tasks must be performed by using the VMware virtual appliance management interface (VAMI):

- Install a valid ECX license key
- Configure the ECX appliance time zone setting

Note: If you do not have a valid ECX license file, contact Catalogic support to obtain one.

Use the steps in the following sections to perform these tasks:

Installing an ECX license key

To install an ECX license key, complete these steps:

1. From a supported browser, enter the following URL to access the management (VAMI) console of the ECX virtual machine:

   https://<HOSTNAME>:5480/

   where <HOSTNAME> is the IP address of the virtual machine where the application is deployed.

2. In the login window, enter a VAMI user name and password and click **Login** to access the management console of the virtual machine, as shown in Figure 2-1. The default VAMI user name is ***administrator*** and the default password is ***ecxadLG235***.

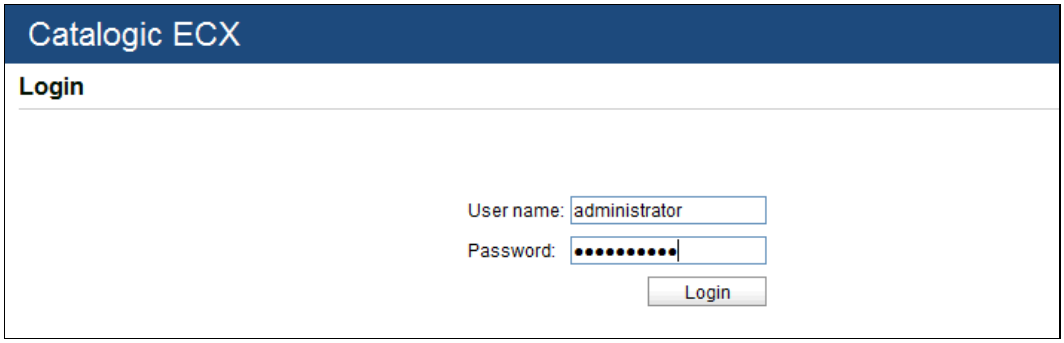


Figure 2-1 The ECX VAMI login window

3. Click the License tab. In the **License** field, browse for your ECX license file on your computer, then click **Upload**, as shown in Figure 2-2.

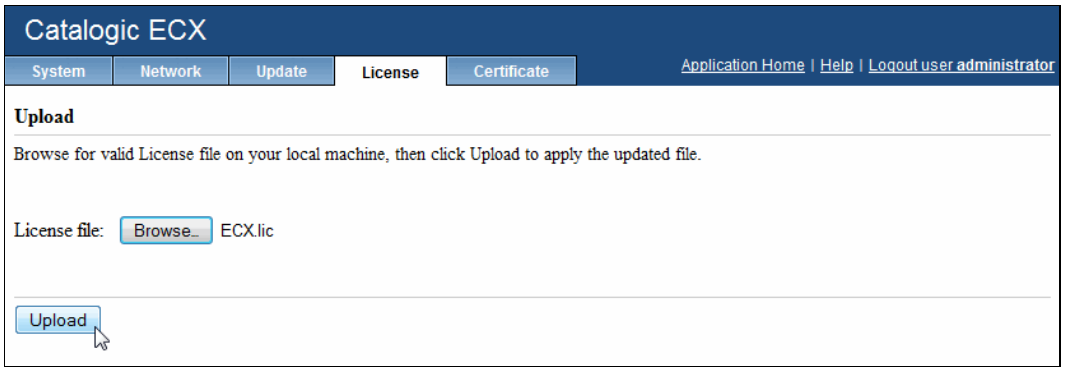


Figure 2-2 Uploading a license file in the ECX VAMI console

Note: After you click **Upload** in step 3, you might be prompted to enter a user name and password again. If so, use the same credentials from step 2.

If a valid license file is uploaded, a successful upload message appears, as shown in Figure 2-3.

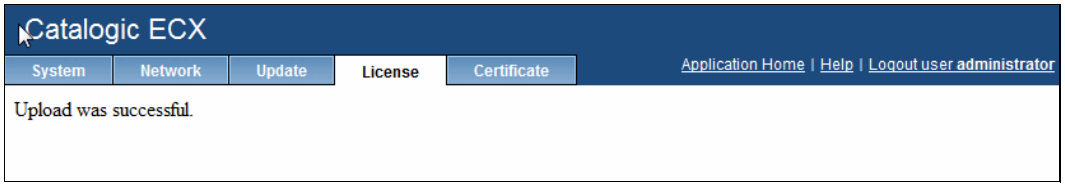


Figure 2-3 Successful license file upload in the ECX VAMI console

Configuring the appliance time zone

To configure the appliance time zone, complete these steps:

1. In the VAMI console, click the System tab at the top of the window.
2. Click **Time Zone** and use the drop-down menu to select the appropriate time zone for your ECX appliance, as shown in Figure 2-4.

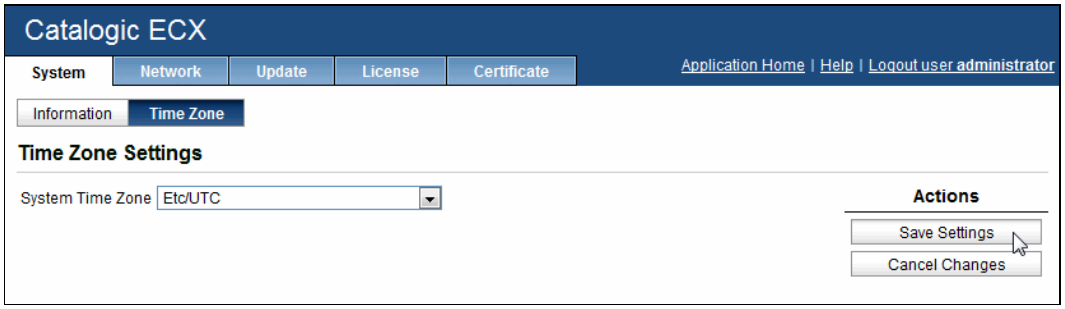


Figure 2-4 Configuring the time zone in the ECX VAMI console

3. Click **Save Settings** to apply the time zone selection.

2.1.2 ECX Cluster Option (ECO) deployment

ECX can also be deployed in a clustered mode that provides enhanced application availability. This enhanced availability is realized by running ECX as a set of Docker containers under the control of the Apache Mesos cluster management software.

The ECX Cluster Option can be deployed over 1, 4 or 6 VMware VMs, and includes these components:

- Zookeeper: Enables fault tolerance by coordinating the election of a leader Mesos node.
- Mesos: The “master” node that controls the cluster. It receives resource offers in the form of CPU and memory from the several “worker” nodes, and offers these resources to the Marathon framework so that it can run containerized tasks using these resources.
- Marathon: Starts and monitors containerized tasks across the cluster, ensuring that the tasks are always running.
- Docker: Automates the deployment of containerized applications and manages their environment and resource usage.
- Weave: Creates a virtual overlay network across which containers, regardless of their location, can communicate by using standard network protocols.

Further description of ECO is beyond the scope of this book. For more information about the ECO deployment option, contact Catalogic Support.

2.2 ECX configuration and management by using the management portal

After the ECX appliance has been deployed and VAMI configuration tasks have been completed, the ECX appliance web-based management portal (GUI) is the primary interface for using the product. Using the portal, you can perform the following common configuration tasks, among others:

- Configure sites
- Configure storage and hypervisor providers
- Define or import user accounts
- Configure role-based access controls

The ECX management portal can also be used to perform common day-to-day tasks, including:

- Create ECX Copy Data and Use Data policies
- Create ECX Report policies
- Monitor ECX Jobs

2.2.1 Accessing the ECX management portal

To access the ECX web management portal, complete these steps:

1. From a supported browser, enter the following URL:

   `https://<HOSTNAME>:8443/portal/`

   where <HOSTNAME> is the IP address of the virtual machine where the application is deployed.

2. In the login window, enter a user name and password and click **Sign In**, as shown in Figure 2-5. The default user name is ***admin*** and the default password is ***password***.

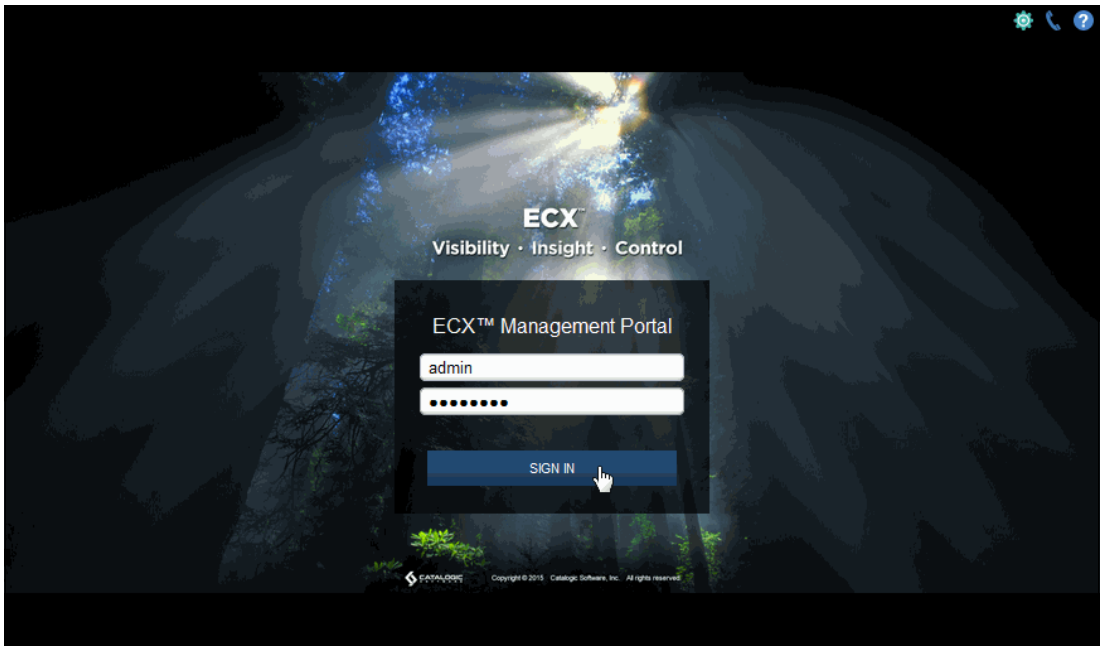


Figure 2-5 Logging in to the ECX management portal

Note: You will be prompted to change the password for the *admin* user account upon initial login.

3. If prompted, confirm the current user password and enter a new password, as shown in Figure 2-6.

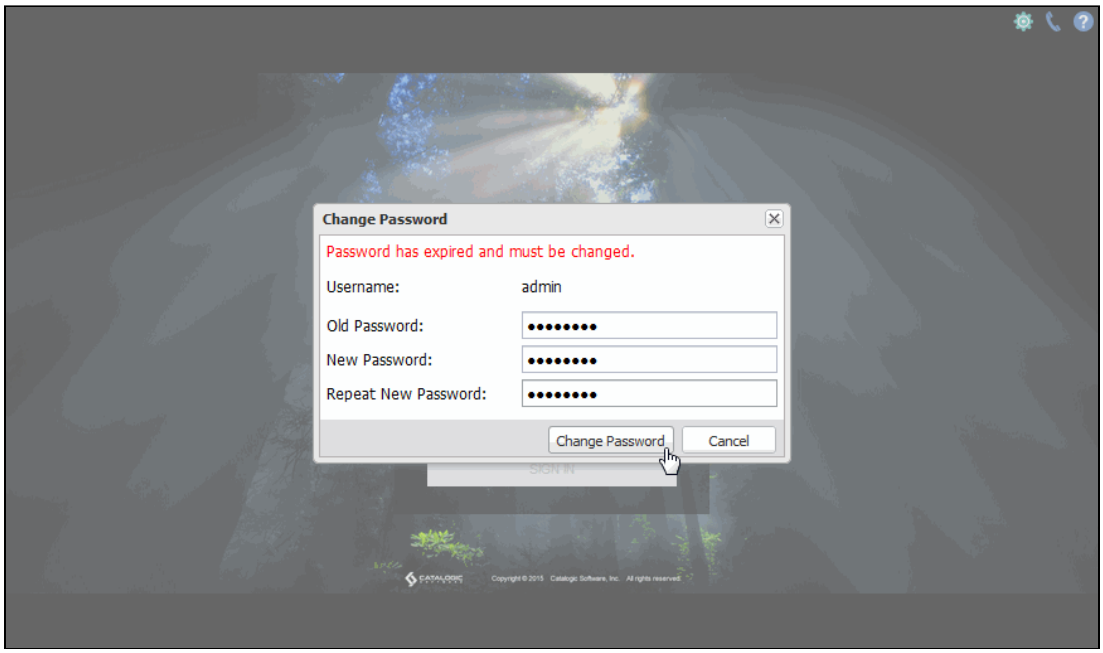


Figure 2-6 Changing the user password in the ECX management portal

Important: Be sure to record the new password and store it securely for future use.

4. Upon successful login, the ECX management portal Home tab loads, as shown in Figure 2-7.

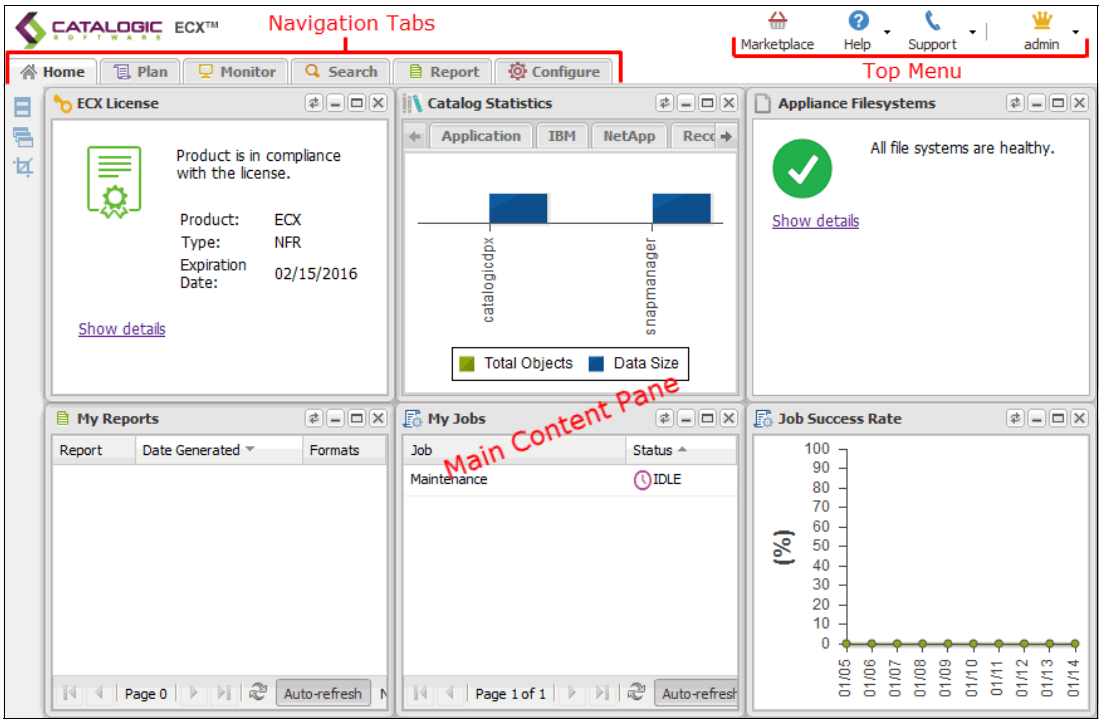


Figure 2-7 The ECX management portal Home tab

2.2.2 The ECX management portal layout

The management portal layout consists of a set of navigation tabs at the upper left, a top menu at the upper right, and the main content pane below these two. All three of these portal elements are identified in red in Figure 2-7.

Top menu

The top menu in the ECX management portal consist of these navigation items:

- Marketplace
- Help
- Support
- User Account

Marketplace

The **Marketplace** button provides a link to the Catalogic Software Marketplace, an online portal where you can download and install additional product features and tools, on-demand. Clicking **Marketplace** opens a new browser window and loads the Marketplace portal, as shown in Figure 2-8.

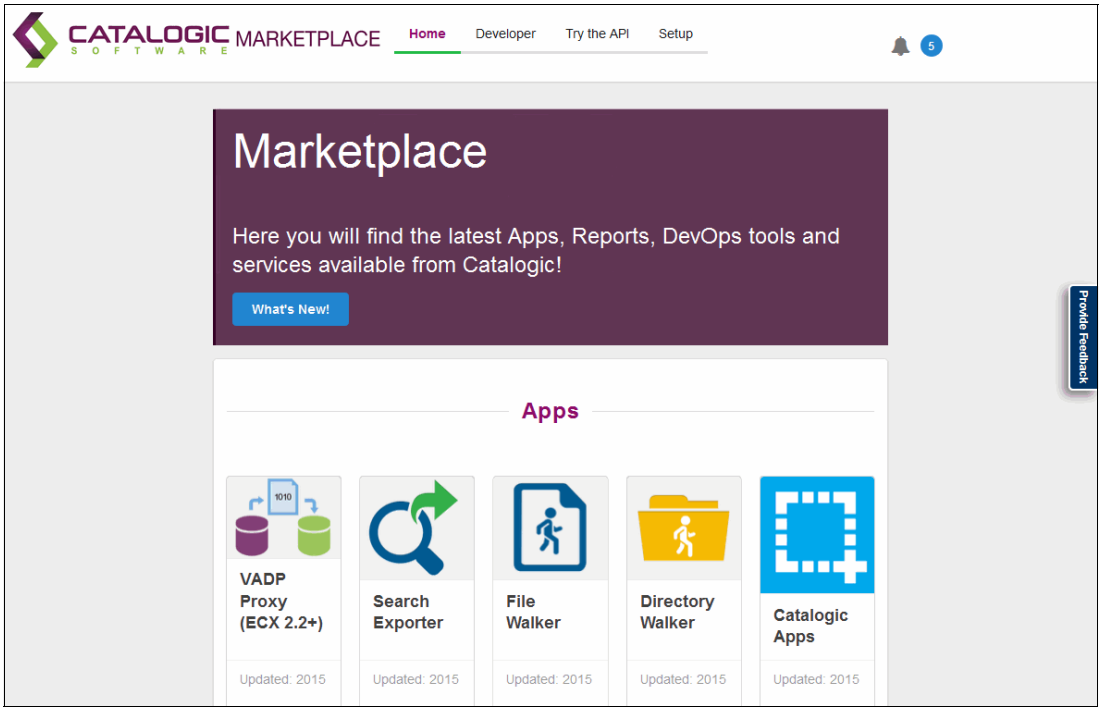


Figure 2-8 The ECX Marketplace portal

Catalogic follows an agile software development methodology, with three product software releases per year. The Marketplace offers you the option to get the latest product enhancements without waiting for a scheduled product software update.

Help

ECX provides an online help system to ensure users have access to the latest product documentation. Click **Help** to open a new browser window and load the ECX product documentation site, as shown in Figure 2-9.

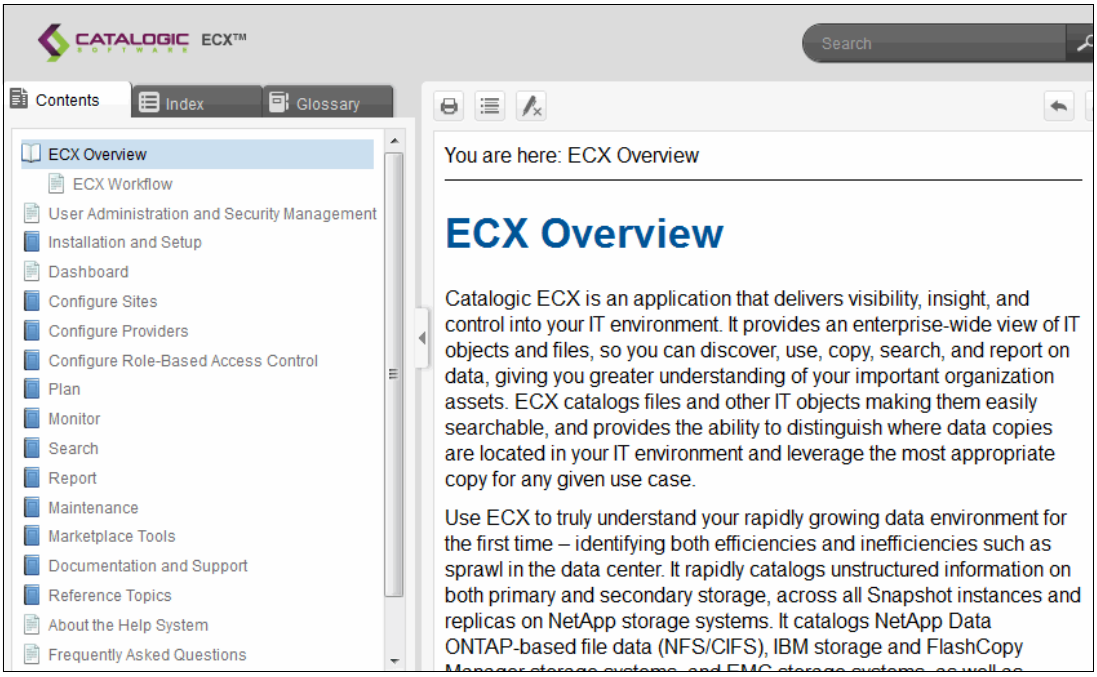


Figure 2-9 The ECX product documentation

Tip: The ECX **Help** button is contextual. When you click the button, it opens the online product documentation page that corresponds to your current location within ECX.

Support

Clicking **Support** opens a new browser window and loads the Catalogic Software MySupport site, as shown in Figure 2-10.

Figure 2-10 The Catalogic Software MySupport site

User Account

The **User Account** button displays the user name of the currently logged-in user, and allows you to log out of ECX, as shown in Figure 2-11.

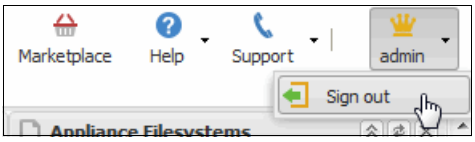


Figure 2-11 Logging out of ECX by using the Top menu

Navigation Tabs

The navigation tabs are used to navigate between the functional areas of ECX. The complete set of tabs consists of these items:

- Home
- Plan
- Monitor
- Search
- Report
- Configure

Home

The **Home** tab is loaded automatically upon login to the ECX management portal. As shown in Figure 2-7 on page 15, it consists of several widgets that display information about the ECX appliance:

- License information
- Catalog statistics
- Appliance file system usage details
- Links to ECX reports
- Summarized ECX job status
- ECX job success metrics

Plan

The **Plan** tab is used to create, modify, and delete ECX policies. By default, it loads the All Policies pane, which displays all of the policies that the logged in user has permissions to see, as shown in Figure 2-12 on page 19. You can control which policies a user has access to with role-based access controls (RBACs). For more information about RBAC, see 5.2, “Role-based access control (RBAC)” on page 76.

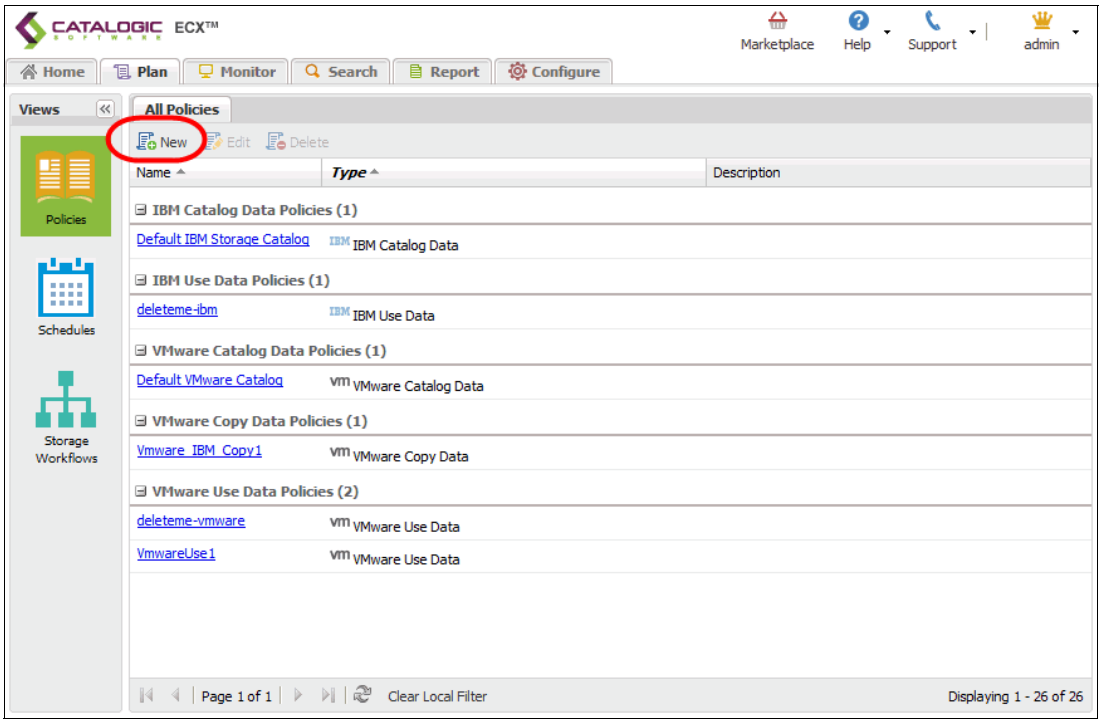


Figure 2-12 The Plan Tab, All Policies view in the ECX management portal

To create a new policy, click **New**, and click the **Plan** tab to load the **New** policy pane, as shown in Figure 2-13.

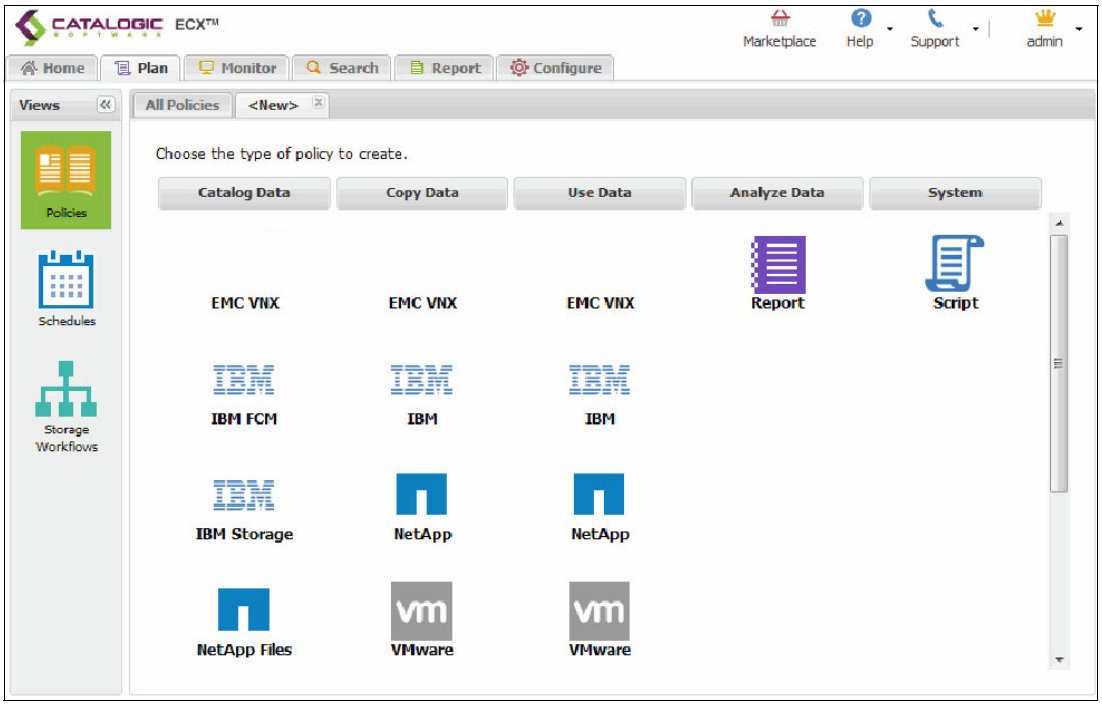


Figure 2-13 The Plan tab New policy view in the ECX management portal

For more information in creating ECX policies, see Chapter 3, “Copy data lifecycle management with Catalogic ECX” on page 37, Chapter 4, “Copy data orchestration and

automation with Catalogic ECX” on page 55, and Chapter 6, “Copy data catalog and analytics with Catalogic ECX” on page 107.

Monitor

The Monitor tab is used to monitor the jobs corresponding to the policies that have been configured in ECX. By default, the Monitor tab displays the All Jobs pane, as shown in Figure 2-14. Similar to the Plan tab, the specific jobs visible in the All Jobs pane can be controlled with RBAC. For more information about RBAC, see 5.2, “Role-based access control (RBAC)” on page 76.

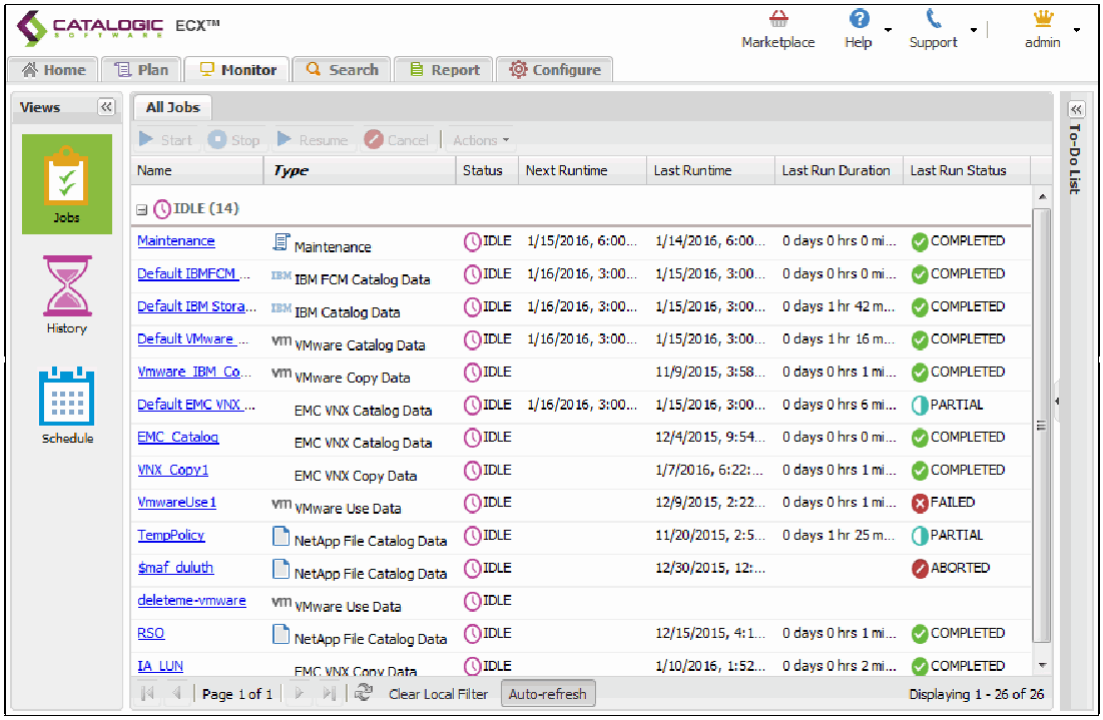


Figure 2-14 The Monitor tab All Jobs view in the ECX management portal

Search

The Search tab provides the ability to perform simple and advanced searches across the contents of the ECX catalog, as shown in Figure 2-15 on page 21. For more information about the search functions in ECX, see 6.3, “Search” on page 112.

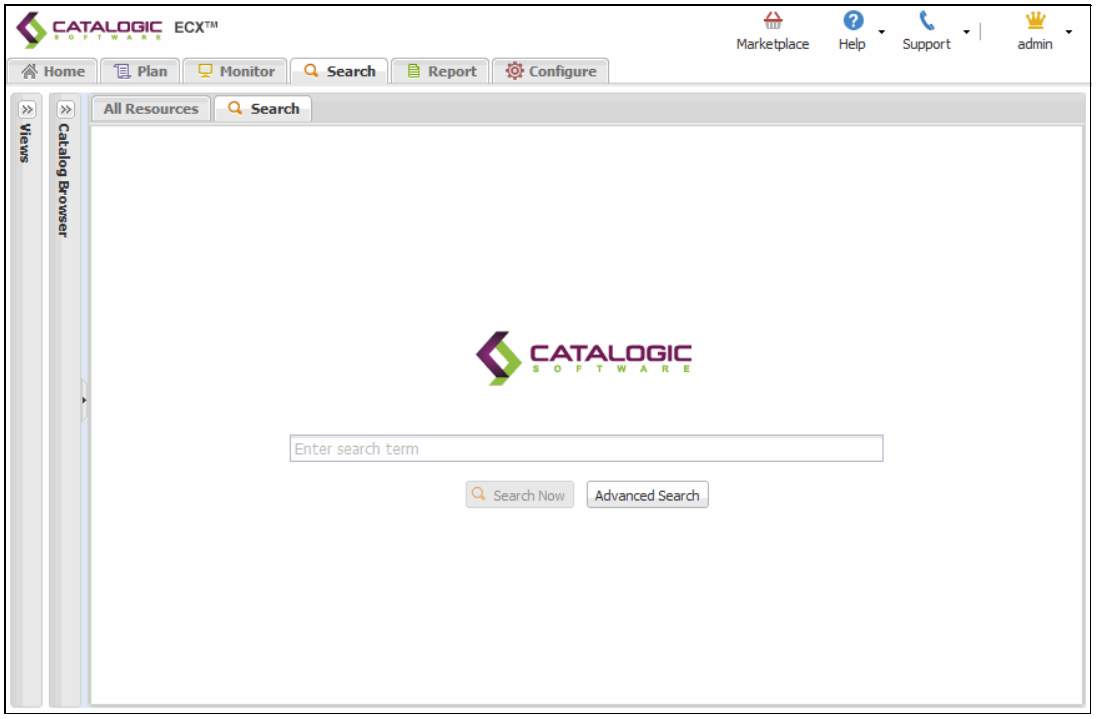


Figure 2-15 The Search tab in the ECX management portal

Report

The Report tab provides access to the rich reporting and data analytics capabilities of ECX, as shown in Figure 2-16. For more information, see 6.4, “Data analytics” on page 115.

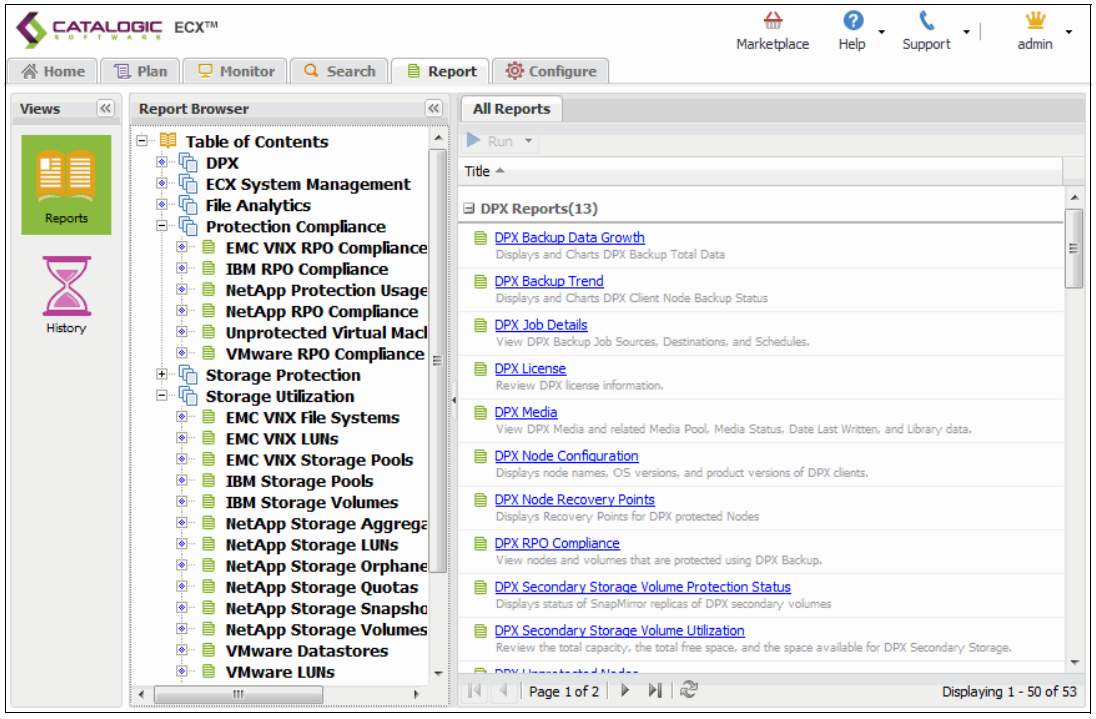


Figure 2-16 The Report tab in the ECX management portal

Configure

The Configure tab, shown in Figure 2-17, is used to configure sites and providers, allowing you to integrate ECX with your storage and hypervisor resources. Optionally, you can use the Configure tab to integrate ECX with your LDAP and SMTP servers, for directory authentication and email notification. The Configure tab also provides access to the configuration settings necessary to configure RBAC. For more information about configuring RBAC, see 5.2, “Role-based access control (RBAC)” on page 76.

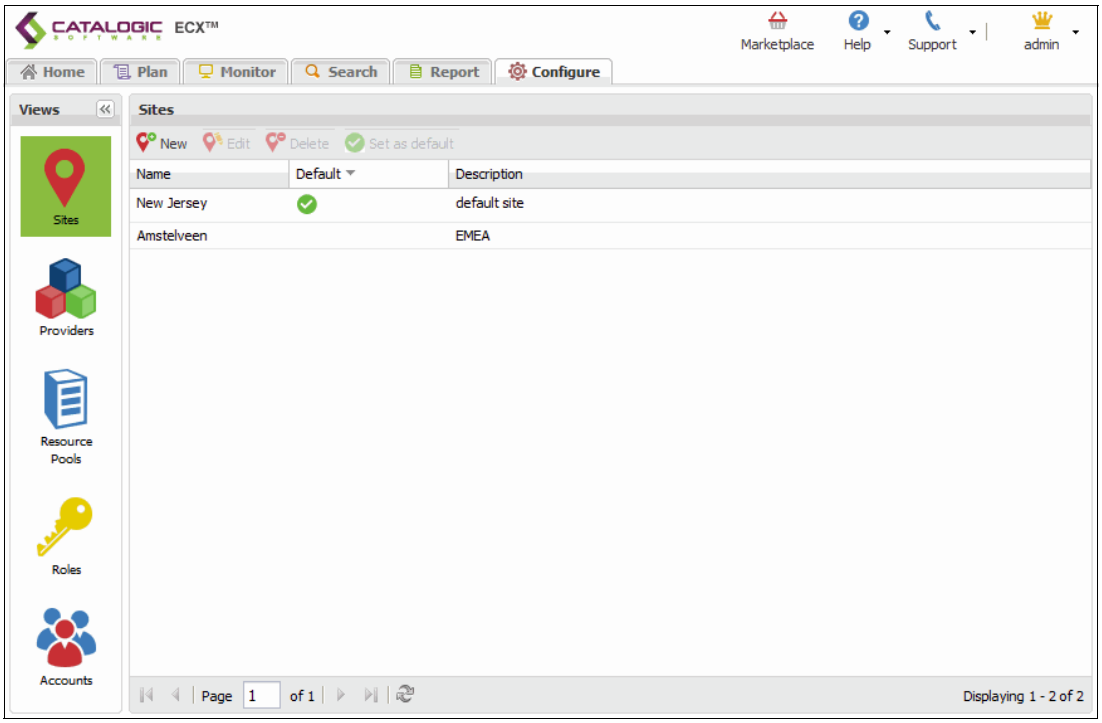


Figure 2-17 The Configure tab in the ECX management portal

2.2.3 Configuring sites and providers

This topic explains sites and providers.

Sites

A site is a user-defined grouping of providers that is generally based on location to help quickly identify and interact with data created through Copy Data Management policies.

Sites are assigned when registering providers. When creating Copy Data and Use Data policies, sites clearly identify where your data is replicated by location.

Upon initial installation, ECX is configured with a single site called **Default** that is flagged as the *default* site. This default designation means that new storage and hypervisor resources are associated with this site automatically when they are created, unless a different site is selected.

ECX allows you to create sites with logical, meaningful names. To create a new site, complete these steps:

1. Click **New** as shown in Figure 2-18.

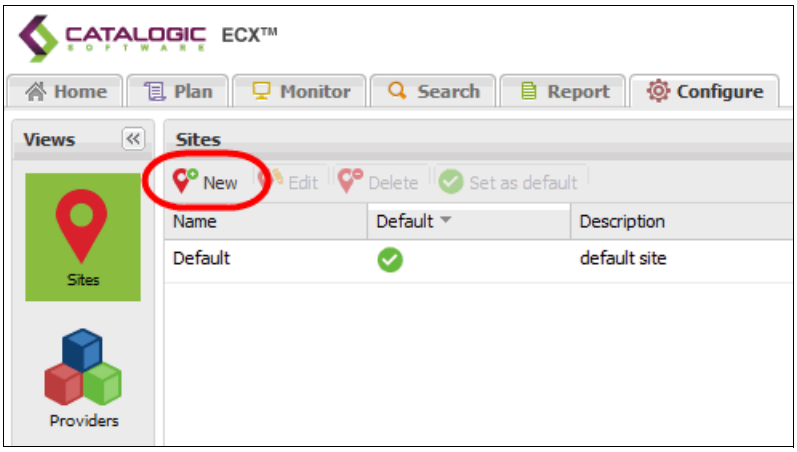


Figure 2-18 Creating a site in the ECX management portal

2. In the dialog window, enter a name and description for the new site, and optionally, check the box to make this new site the *default* site.
3. Click **OK** to save the new site as shown in Figure 2-19.

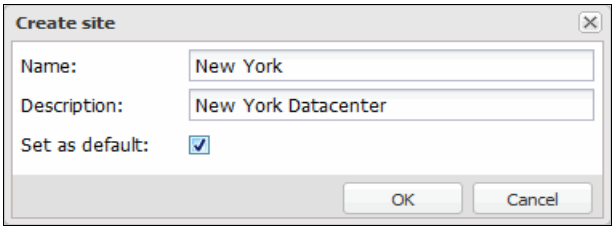


Figure 2-19 Providing the details for a new site in the ECX management portal

Providers

Providers are software abstraction layers that allow ECX to communicate with other platforms. The following are examples of provider types:

- Storage providers
- Hypervisor providers
- Directory service providers
- Mail service providers

Storage and hypervisor providers allow ECX to present a common set of copy data management, automation, and self-service functions across a heterogeneous infrastructure by interfacing directly with the APIs of the specific storage and hypervisor platforms. ECX storage and hypervisor providers obviate the need for manual scripting, and provide a mechanism to catalog the data copies in the configured systems, and automate and orchestrate the copy services of those systems.

As mentioned in 1.4.2, “Integration with IBM products” on page 6, ECX 2.2 introduced two new storage providers that enable support for these items:

- Storage systems built with IBM Spectrum Virtualize™ software
- IBM Spectrum Protect Snapshot (formerly IBM Tivoli Storage FlashCopy Manager)

This book will demonstrate the registration of an IBM Storwize V7000 storage system and a FlashCopy Manager server instance in ECX using these new providers, as well as a VMware vCenter server. It will also demonstrate the registration of an LDAP server and an SMTP server in ECX.

Configuring an IBM Storwize V7000 in ECX

To configure an IBM Storwize V7000 in ECX, gather the following information about that system:

Site	A user-defined provider location, created in the Sites view on the Configure tab.
Name	A user-defined name for the IBM provider. This name can be the same as the host name or it can be a meaningful name that is used within your organization to refer to the provider. Provider names must be unique.
IP address	A resolvable IP address or a resolvable path and machine name.
Username	The name that is used to access the provider.
Password	The password associated with the user name.
Comment	Optional provider description.

To configure the system in ECX, complete these steps:

1. Navigate to the Configure tab and click the **Providers** icon in the left side navigation, as shown in Figure 2-20.

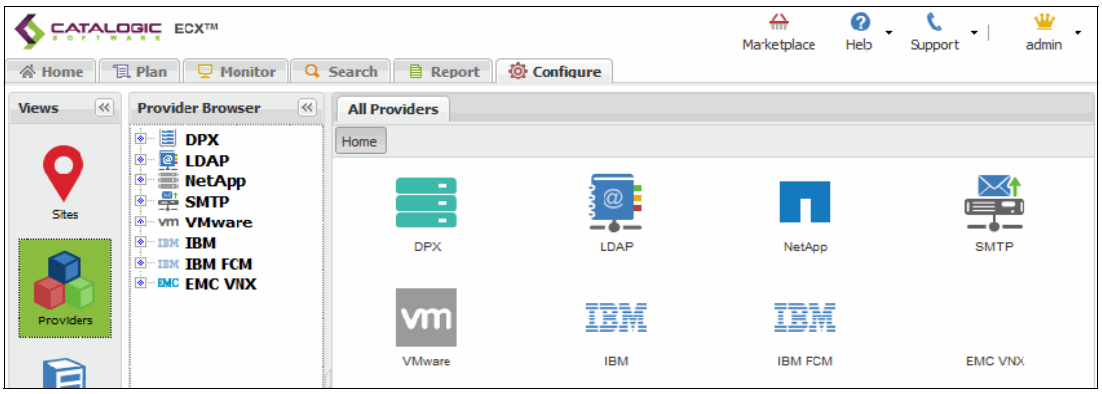


Figure 2-20 The Providers window in the ECX management portal

2. Right-click the **IBM** item in the Provider Browser pane, as shown in Figure 2-21, and click **Register**.

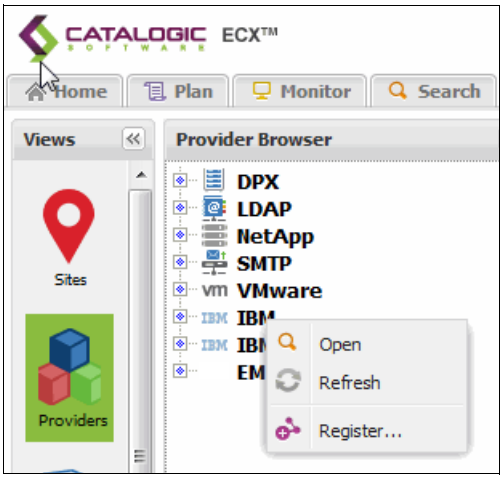


Figure 2-21 Registering a new IBM system in the ECX management portal

3. In the dialog window, enter the details about the IBM Storwize V7000 system, as shown in Figure 2-22, and click **OK** to register the system.

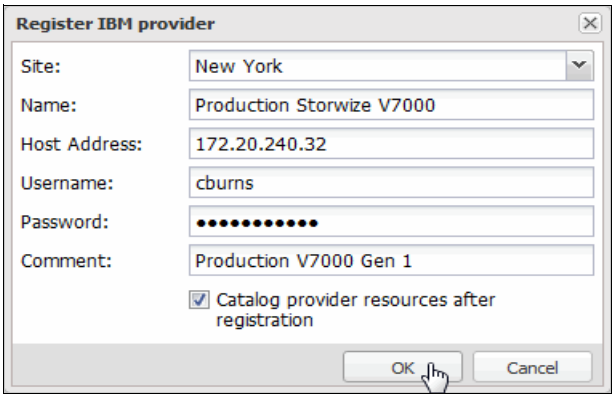


Figure 2-22 IBM provider registration details in the ECX management portal

Tip: Leave the **Catalog provider resources after registration** check box selected. Doing so ensures that the system you are registering is added to the appropriate ECX default storage catalog policy.and. ECX will then automatically catalog the resources on this system.

The IBM Storwize V7000 system is now registered in ECX and available for use in ECX Copy Data and Use Data policies.

Configuring a FlashCopy Manager server instance ECX

To configure an IBM FlashCopy Manager or IBM Spectrum Protect Snapshot server instance in ECX, gather the following information about that system:

Site	A user-defined provider location, created in the Sites view on the Configure tab.
Name	A user-defined name for the IBM provider. This name can be the same as the host name or it can be a meaningful name that is used within

	your organization to refer to the provider. Provider names must be unique.
IP address	A resolvable IP address, or a resolvable path and machine name.
Username	The name that is used to access the provider.
Password	The password associated with the user name.
Comment	Optional provider description.

To configure the system in ECX, complete these steps:

1. Click the Configure tab.
2. Click the **Providers** icon in the left side navigation, as shown in Figure 2-20 on page 24.
3. Right-click the **IBM FCM** item in the Provider Browser pane, as shown in Figure 2-23, and select **Register**.

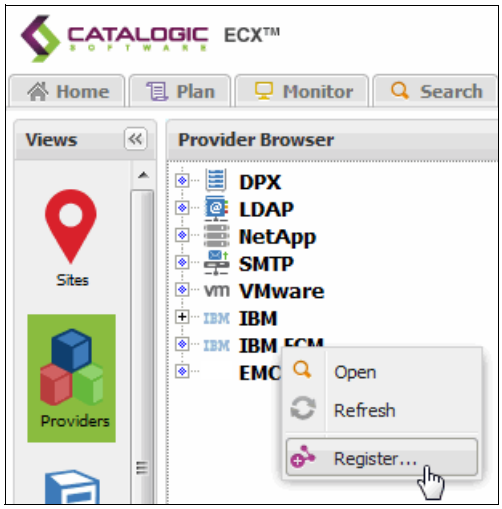


Figure 2-23 Registering a new IBM FlashCopy Manager system in the ECX management portal

4. In the dialog window, enter the details about the IBM FlashCopy Manager or IBM Spectrum Protect Snapshot server, as shown in Figure 2-24, and click **OK** to register the system.

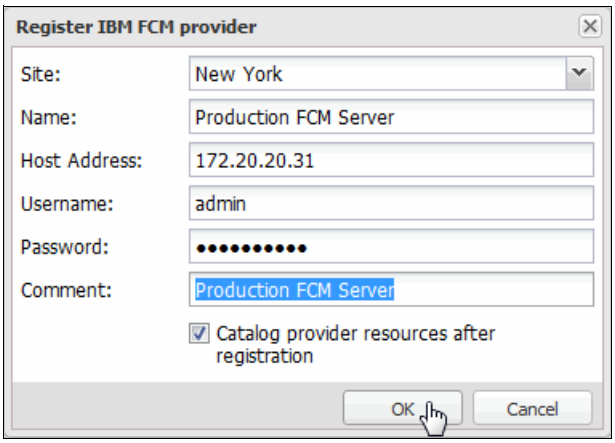


Figure 2-24 IBM FCM provider registration details in the ECX management portal

Tip: Leave the **Catalog provider resources after registration** check box selected. Doing so ensures that the FCM server you are registering is added to the appropriate ECX default FCM catalog policy. ECX will automatically catalog the resources on this server.

The IBM FlashCopy Manager server is now registered in ECX and available for use in ECX Copy Data and Use Data policies.

Configuring a VMware vCenter server instance ECX

To configure a VMware vCenter server instance in ECX, gather the following information about that system:

Site	A user-defined provider location, created in the Sites view on the Configure tab.
Name	A user-defined name for the VMware server. This name can be the same as the host name, or it can be a meaningful name that is used within your organization to refer to the provider. Provider names must be unique.
Host address	A resolvable IP address, or a resolvable path and machine name.
Port	The communications port of the provider you are adding. The typical default port is 80 for non-SSL connections or 443 for SSL connections.
Use SSL	Enable an encrypted Secure Sockets Layer connection.
Username	The name that is used to access the provider.
Password	The password associated with the user name.
Comment	Optional provider description.

To configure the system in ECX, complete these steps:

1. Navigate to the Configure tab and click the **Providers** icon in the left side navigation, as shown in Figure 2-20 on page 24.
2. Right-click the **VMware** item in the Provider Browser pane, as shown in Figure 2-25, and select **Register**.

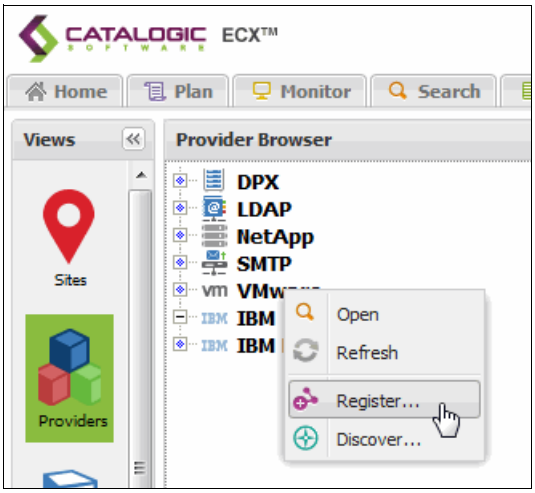


Figure 2-25 Registering a new VMware vCenter server in the ECX management portal

3. In the dialog window, enter the details about the VMware vCenter server, as shown in Figure 2-26, and click **OK** to register the system.

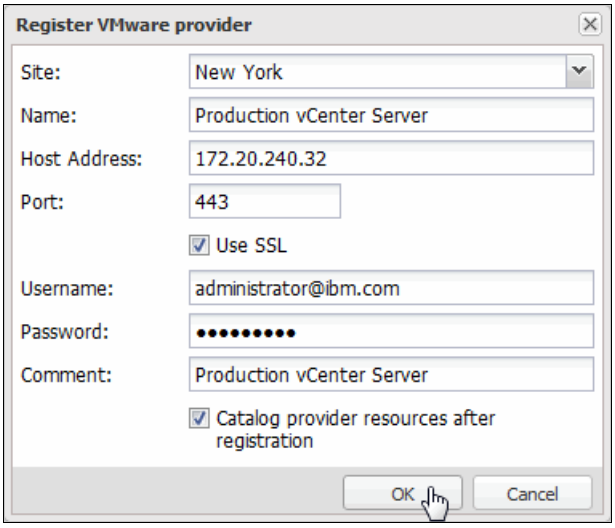


Figure 2-26 VMware provider registration details in the ECX management portal

Tip: Leave the **Catalog provider resources after registration** check box selected. doing so ensures that the VMware vCenter server you are registering is added to the appropriate ECX default VMware catalog policy. ECX will automatically catalog the resources on this server.

The VMware vCenter server is now registered in ECX and available for use in ECX Copy Data and Use Data policies.

Configuring an LDAP server in ECX

To configure an LDAP server in ECX, gather the following information about that server:

Name	A user-defined name for the LDAP Server. Provider names must be unique.
IP address	The IP address or resolvable logical node name of the LDAP server.
Port	The port on which the LDAP server is listening. The typical default port is 389 for non-SSL connections or 636 for SSL connections
Use SSL	Enable in order to establish a secure connection to the LDAP server.
Bind DN	The name used for authenticating the connection to the LDAP server. ECX supports simple bind.
Password	The password associated with the Bind Distinguished Name.
Base DN	The location where users and groups can be found.
User Filter	A filter to select only those users under the Base DN that match certain criteria. An example of a valid default user filter is: cn={0} This entry controls the type of user name that appears in ECX display of users.
User RDN	The relative distinguished path for the user. Specify the path where user records can be found. An example of a valid default RDN is: cn=Users

Group RDN	The relative distinguished path for the group. Specify the path where group records can be found if the group is at a different level than the user path.
Comment	Optional description.

To configure the system in ECX, complete these steps:

1. Navigate to the Configure tab and click the **Providers** icon in the left side navigation, as shown in Figure 2-20 on page 24.
2. Right-click the **LDAP** item in the Provider Browser pane, as shown in Figure 2-27, and select **Register**.

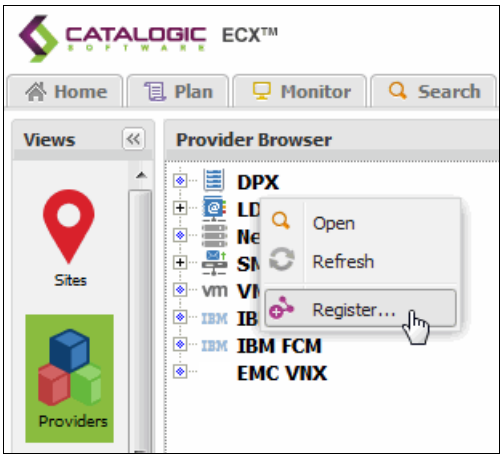


Figure 2-27 Registering a new LDAP server in the ECX management portal

3. In the dialog window, enter the details about the LDAP server, as shown in Figure 2-28, and click **OK** to register the server.

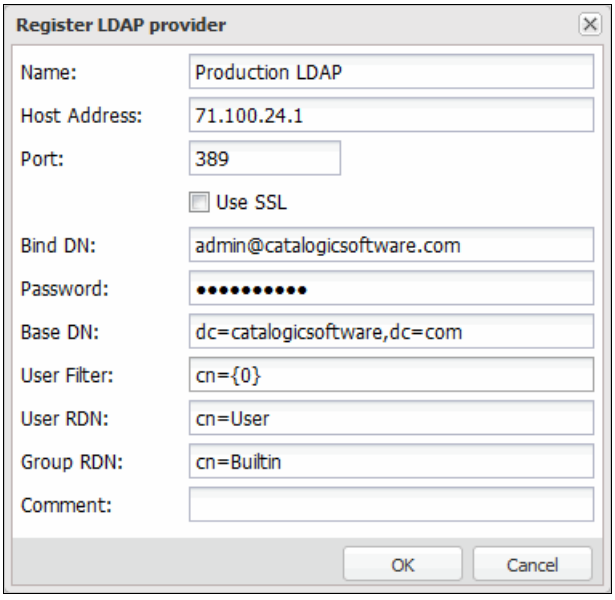


Figure 2-28 LDAP provider registration details in the ECX management portal

The LDAP server is now registered in ECX and available to use for importing LDAP user groups.

Configuring an SMTP server in ECX

To configure an SMTP server in ECX, gather the following information about that server:

Name	A user-defined name for the SMTP server. This name can be the same as the host name or it can be a meaningful name that is used within your organization to refer to the provider. Provider names must be unique.
IP address	A resolvable IP address, or a resolvable path and machine name.
Port	The communications port of the provider that you are adding. Select the **Use SSL** check box to enable an encrypted Secure Sockets Layer connection. The typical default port is 25 for non-SSL connections or 443 for SSL connections.
Username	The name that is used to access the provider.
Password	The password associated with the user name.
Comment	Optional provider description.
From Address	Optional address to be associated with email communications from ECX.
Subject Prefix	An optional prefix to add to the email subject lines sent from ECX.
Timeout (msec)	The email timeout value in milliseconds.

To configure the system in ECX, complete these steps:

1. Navigate to the Configure tab.
2. Click the **Providers** icon in the left side navigation, as shown in Figure 2-20 on page 24.
3. Right-click the **SMTP** item in the Provider Browser pane, as shown in Figure 2-29, and select **Register**.

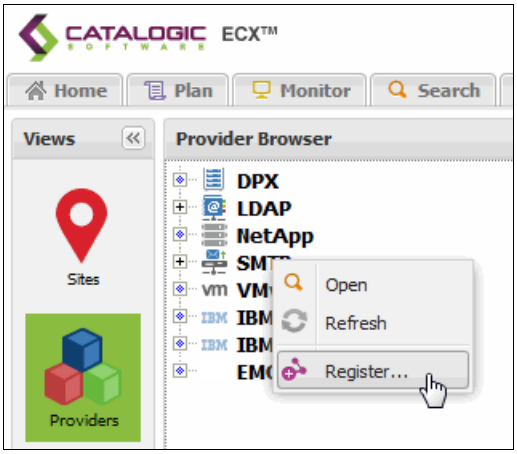


Figure 2-29 Registering a new SMTP server in the ECX management portal

4. In the dialog window, enter the details about the SMTP server, as shown in Figure 2-30, and click **OK** to register the server.

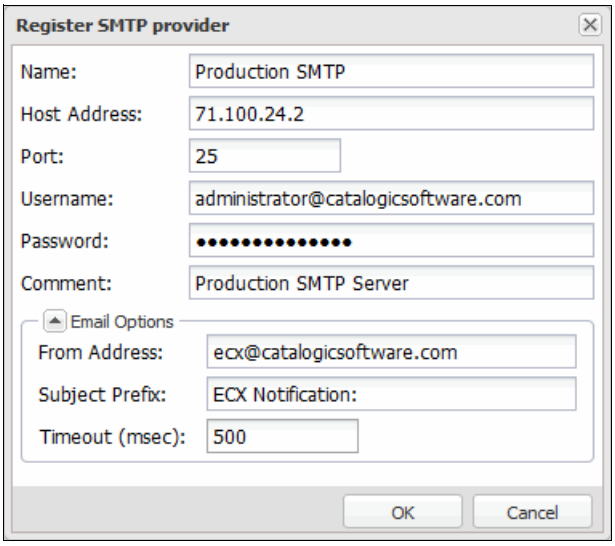


Figure 2-30 SMTP provider registration details in the ECX management portal

The SMTP server is now registered in ECX and available to use for policy notifications.

2.2.4 Monitoring jobs in the ECX management portal

The Monitor tab, described in “Monitor” on page 20, can be used to monitor the status of various ECX jobs. Many of the Copy Data policies and Use Data policies that are created in ECX run jobs automatically according to the schedules defined when the policies are created. From the Monitor tab, you can gain immediate insight into the status of the jobs, and drill down into specific jobs to gain more detail.

The All Job pane

As described in “Monitor” on page 20, the **Monitor** tab displays the **All Jobs** pane by default. The **All Jobs** pane displays the following job information, as shown in Figure 2-31 on page 32:

- Name
- Type
- Status

The following list describes the various possible job statuses in ECX:

Completed	Indicates the job session completed successfully. All tasks associated with the job session were completed.
Partial	Indicates the job session completed, but one or more tasks failed or were skipped.
Failed	Indicates that the job session did not successfully complete due to mixed task statuses.
Aborted	Indicates that the job session did not successfully complete due to a reset, restart, or shutdown of the virtual appliance server.
Idle	Indicates that the job session is idle.

Skipped	Indicates that a volume was not cataloged. See the Task tab for more information about skipped jobs.
Stopped	Indicates that the job was stopped by using the **Stop** button.

- Next runtime
- Last runtime
- Last run duration
- Last run status

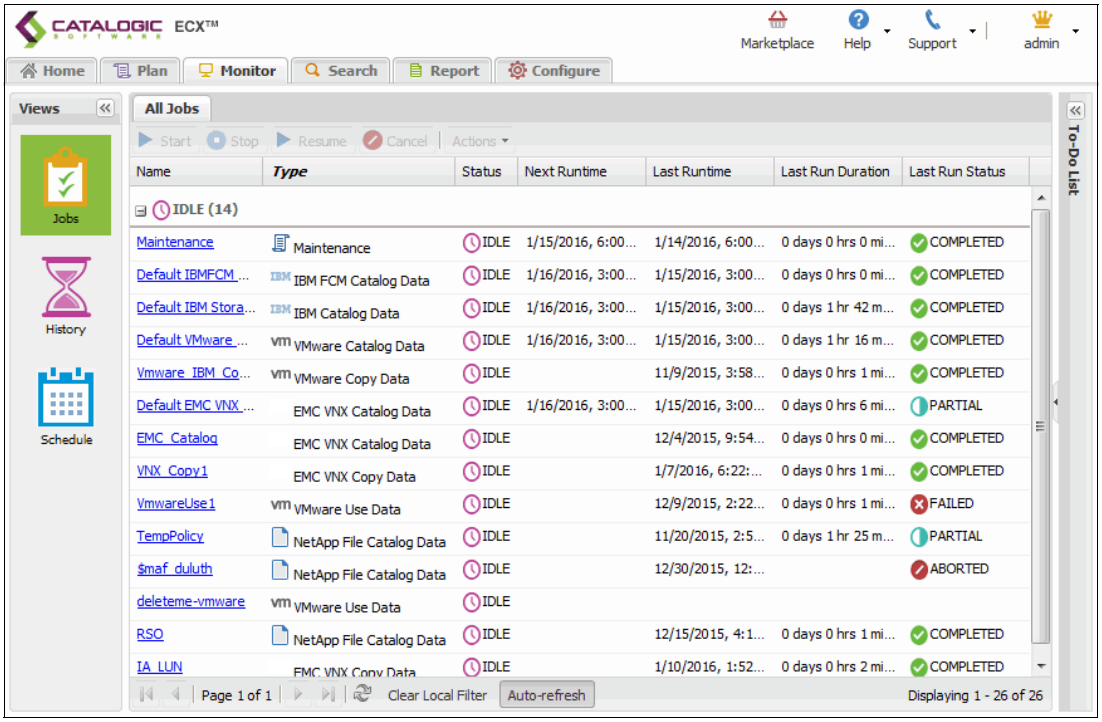


Figure 2-31 The All Job pane in the Monitor tab in the ECX management portal

You can customize the specific columns that are displayed by hovering the cursor over any column heading and clicking the triangle icon as shown in Figure 2-32.

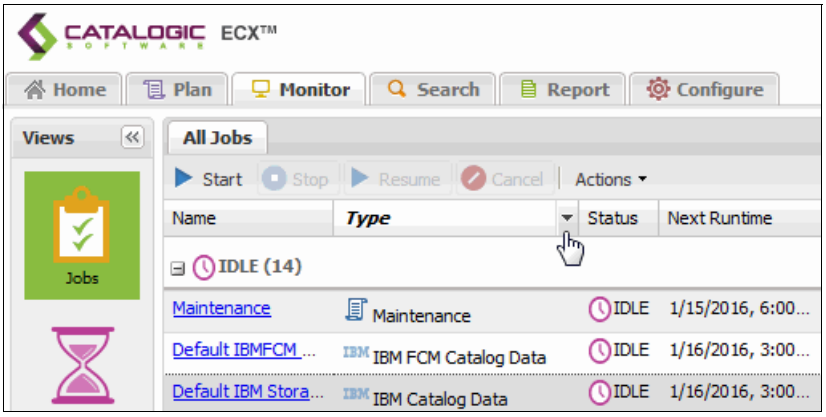


Figure 2-32 Accessing the heading menu in the ECX management portal

Select or clear column headings as wanted as shown in Figure 2-33.

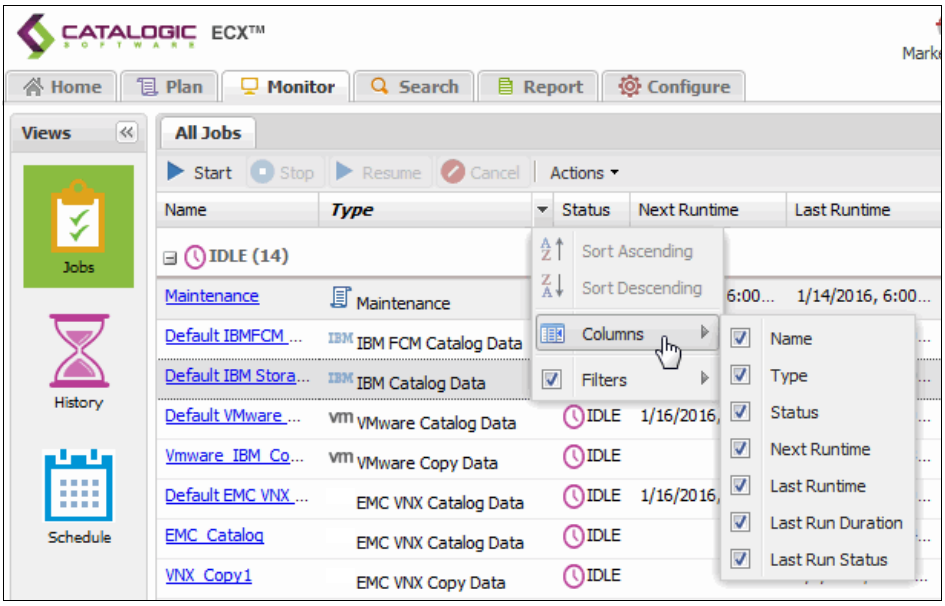


Figure 2-33 Customizing the column headings in the ECX management portal

In addition, you can filter the jobs displayed on the All Jobs pane by job type. To do so, hover the cursor over the **Type** column heading and click the triangle icon as shown in Figure 2-32 on page 32, then select or clear the wanted job types, as shown in Figure 2-34.

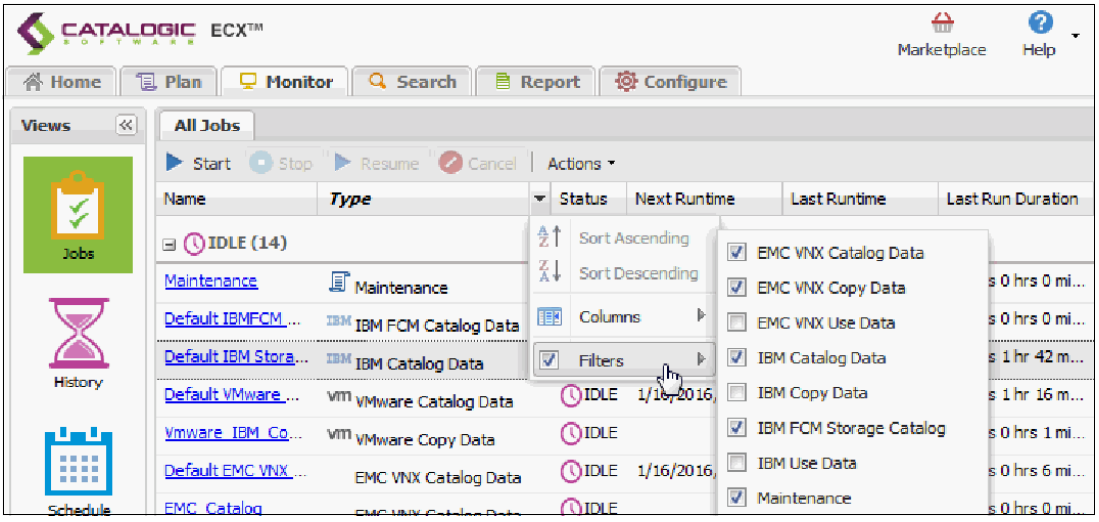


Figure 2-34 Filtering the All Jobs pane by job type in the ECX management portal

Tip: The view customizations that are shown in Figure 2-33 and Figure 2-34 are applicable in many screens throughout the ECX management portal.

Monitoring a job session

You can view the details of a job session that is running or one that has finished.

To monitor a running job session, from the All Jobs pane in the Monitor tab, click the name of the corresponding job, as shown in Figure 2-35.

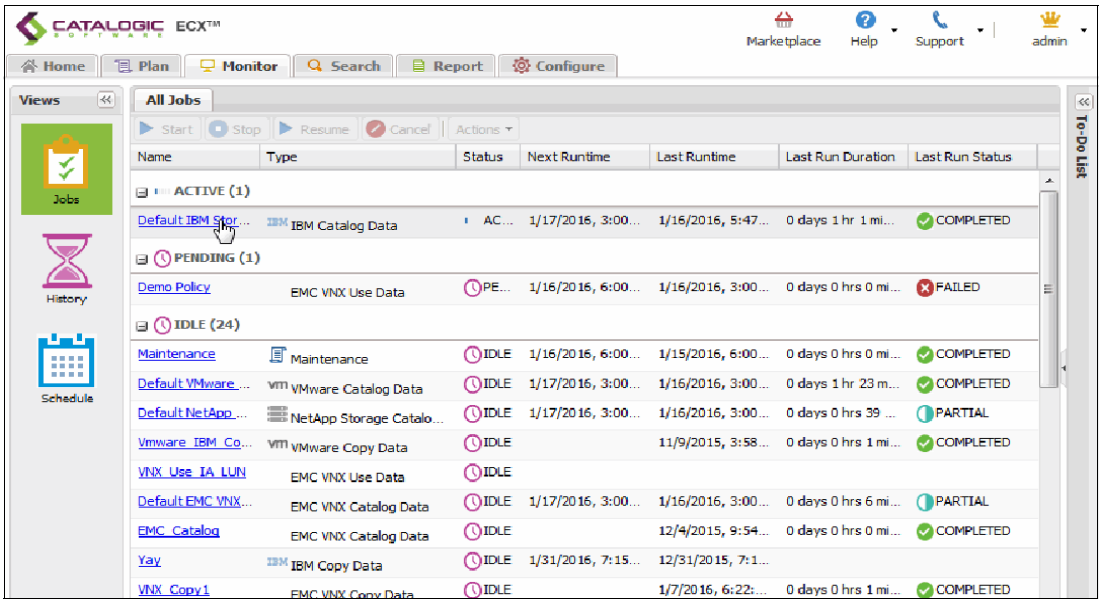


Figure 2-35 Accessing the job details of a running job in the ECX management portal

To monitor a job that has finished, from the Monitor tab, click the **History** icon in the Views pane, then click the name of the corresponding job, as shown in Figure 2-36.

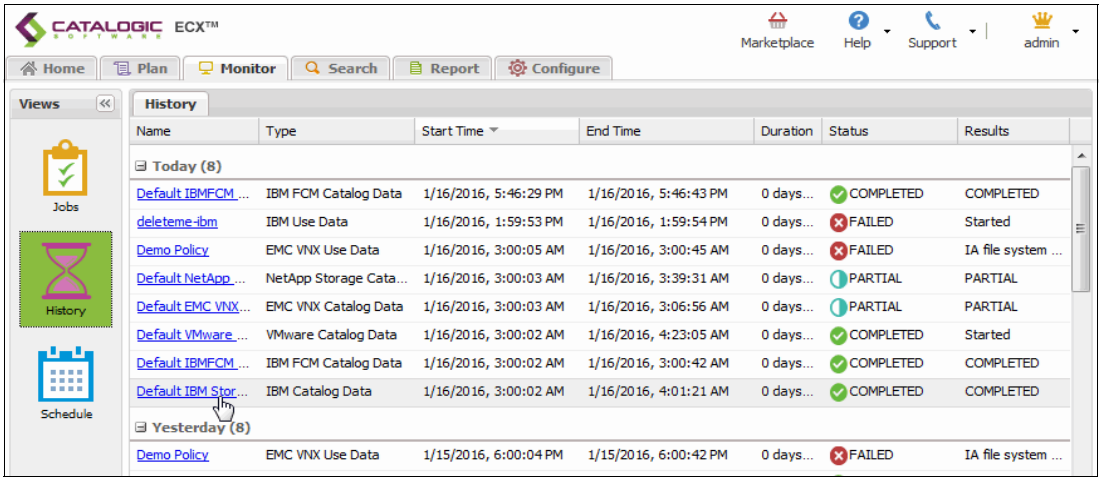


Figure 2-36 Accessing the job details of a job that has finished in the ECX management portal

Either method loads the Job Details pane for the selected job as shown in Figure 2-37.

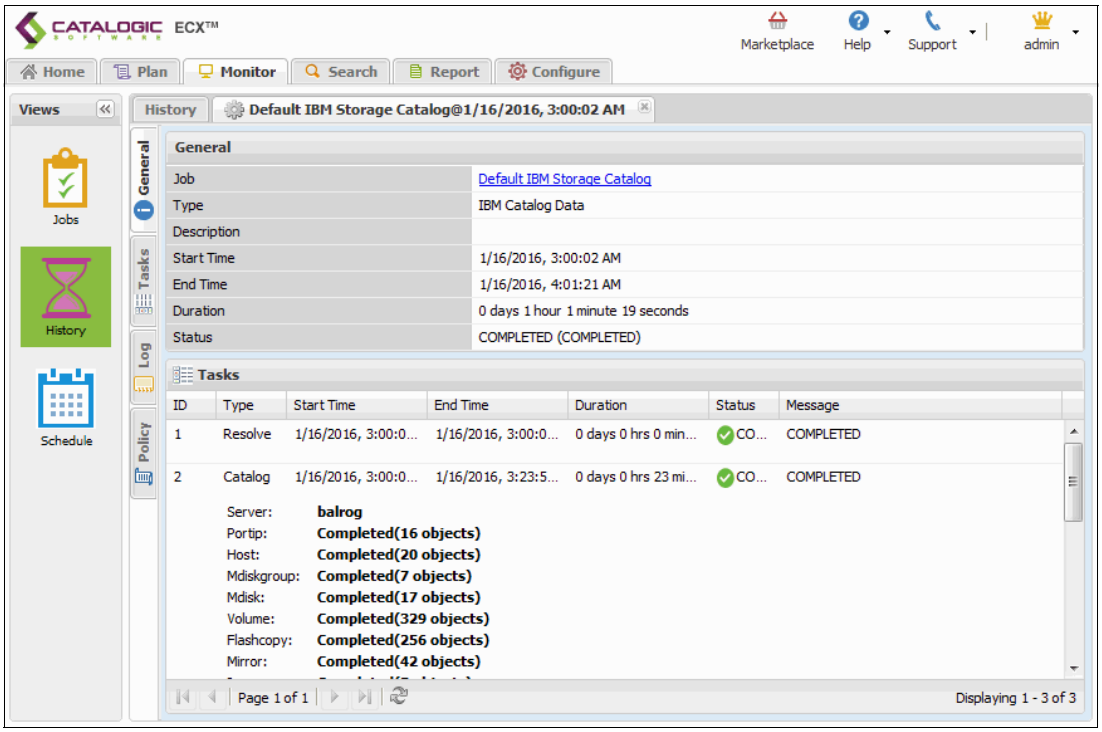


Figure 2-37 The Job Details pane in the ECX management portal

The Job Details pane provides detailed information about the job session. This information is broken down between four tabs. Click each tab to view its details.

General	Displays an overview of the job session, including start and end times, duration, and status. It also displays details of the underlying tasks that take place during the job session, including the task's type, duration, and status.
Tasks	Displays a task-by-task view of the job session, including start and end times, duration, and status.
Log	Displays the job log, which can be used for troubleshooting purposes. Job logs are automatically removed from the ECX Catalog six months after they are generated.
Policy	Displays an overview of the policy definition, including the policy name, sources, options, and notification settings.

3

Copy data lifecycle management with Catalogic ECX

This chapter describes Catalogic's ECX Copy Data Lifecycle Management with IBM Spectrum Virtualized storage.

This chapter includes the following sections:

- ECX copy data lifecycle management
- ECX Copy Workflow technical overview
- ECX Copy workflow examples
- Application consistency and custom scripting

3.1 ECX copy data lifecycle management

ECX manages the full copy data lifecycle that includes Planning, Creating, Managing, and Retention. ECX incorporates the many aspects of the copy data lifecycle by using ECX Copy Data workflows. ECX enables users to quickly and easily establish workflows that create snapshots and replica copies for storage volumes and VMware virtual machines.

With ECX, snapshots and replicas are deployed across the environment and are aligned with the service level agreements (SLAs) established by the lines of business and the IT team. Workflows can range from a simple one that creates a single local snapshot once per day, to more complex workflows that involve multiple local and remote data copies of every volume or virtual machine multiple times per day. Creating these workflows through ECX saves administrators time, drives repeatability of best practices, and reduces copy data sprawl of data copies that never get used, which drives out complexity and significantly reduces CapEx and OpEx. ECX eliminates the dependency on manual procedures or scripts that are hard to maintain, often fail, and do not provide IT with any insight into their service levels. ECX provides service-level and exception-based reporting, and detailed logging, allowing IT to manage and diagnose failures faster and help ensure SLAs. This chapter shows how to perform copy data lifecycle management by using ECX and IBM storage systems built with Spectrum Virtualize software.

3.2 ECX Copy Workflow technical overview

ECX Copy Workflows allow users to create point-in-time application consistent snapshots and replicas of VMware applications or stand-alone storage volumes. ECX uses the IBM FlashCopy and Global Mirror Change Volume (GMCV) technologies to create snapshots and replicas. For snapshots, the ECX workflow automatically identifies source volumes, quiesces applications, and provisions FlashCopy target volumes. For replicas, ECX provisions Global Mirror target volumes, creates the necessary GMCV relationships, pauses and syncs the relationships, and creates a FlashCopy of the mirror target volume on the destination system. The ECX Copy workflow logs all aspects of each operation performed and the resulting status codes. It provides real-time status and history in the ECX Job Monitor. See 2.2.4, “Monitoring jobs in the ECX management portal” on page 31 for more details. ECX also provides compliance reports for workflows by using the ECX Reporting functions. For more information about reporting in ECX, see 6.4, “Data analytics” on page 115.

ECX workflows are represented in ECX as Copy Data policies. There are two different types of ECX Copy Data policies: VMware Copy Data policy and IBM Copy Data policy. The VMware Copy Data policy is used to capture VMware vSphere applications, whereas the IBM Copy Data policy is used to capture applications that are directly using IBM storage volumes.

Note: In this chapter, snapshots are defined as a point-in-time hardware snapshot of a data storage volume, specifically a FlashCopy of a volume on primary storage. A replica is defined as a combination of Global Mirror of the source volume and FlashCopy on the mirror target volume.

3.2.1 ECX VMware Copy Data Policy

The ECX VMware Copy Data Policy manages snapshots and replication for applications that run on VMware vSphere hypervisor. The VMware virtual machines can be captured in both an application consistent or crash-consistent manner. VM snapshot and application-specific

operations such as quiesce and log truncation are also performed as part of the copy data policy. This system ensures that both the virtual machine and applications are consistent before a hardware snapshot FlashCopy is taken. The policy can capture VMware virtual machines on both IBM storage and non-IBM storage.

For the VMware Copy Data Policy, two major parameters should be specified:

- The source of copy data
- The ECX Storage Workflow

There are also options related to application consistency, pre/post scripts, notifications, and activation time. The source of copy data in the VMware copy data policy is specified as VMware vSphere objects such as vCenter, Datacenter, VM Folders, vApps, Datastores, and VMs.

ECX Storage Workflows are templates that specify how the source virtual machines are captured and replicated. It can be selected from a list of Storage Workflows that have already been created by the storage administrator. Users should choose the Storage Workflow that best meets the SLA requirements for those particular virtual machines. The Storage Workflow defines how virtual machines are captured and the target location for replication. Retention parameters are also specified in the Storage Workflow. There are two different methods of capturing virtual machines: FlashCopy and “VM Copy”.

If FlashCopy is selected as the capture method, then the copy data policy automatically determines on which IBM storage volumes the virtual machines are placed. Then it dynamically creates FlashCopy target volumes for each source volume and creates a FlashCopy of the source volumes. If the option for VM snapshot is specified in the copy data policy options, a VMware vSphere VM snapshot is performed before the FlashCopy and then deleted after the FlashCopy is complete.

If “VM Copy” is selected as the capture method, ECX uses VMware’s vSphere Storage APIs - Data Protection (VADP) protocol to copy the virtual machines to an IBM target storage volume. The “VM Copy” method is useful when the VM primary storage is not IBM storage. VMware VADP is a software framework that enables block level incremental backup of VMware virtual machines. After the VMs are copied to an IBM target storage volume, a FlashCopy of the target IBM storage volume is created to preserve that point in time. The “VM Copy” is stored on the IBM target storage volume in native VMware VMDK format, and not wrapped in any backup or archive file formats. For subsequent captures, incremental Change Block Tracking (CBT) is used to only copy changed blocks since the last ECX policy run. After each incremental capture, a FlashCopy is performed on the target storage volume to capture the incremental point in time. This particular method of using FlashCopy to capture the incremental point in time allows you to perform incremental capture without needing to merge incrementals as FlashCopies are aged.

Note: In a single storage workflow, both FlashCopy and VM Copy can be specified as a capture method. If both capture methods are specified, a FlashCopy is performed for any VMs on IBM storage volumes and a VM Copy is performed for any VMs not on IBM storage volumes.

The Storage Workflow also specifies whether a replica of the virtual machines is created on an auxiliary storage node. ECX achieves replicas of virtual machines by orchestrating Global Mirror with change volume and FlashCopy functions. See “Storage workflows” on page 96 for more details. At the end of the copy data policy execution, all information about VM configuration, VM Networks, FlashCopy targets, and Global Mirror destination volumes are cataloged as recovery points within the ECX catalog. The recovery points in the ECX catalog are used when application copies are used with ECX Use Data policies.

3.2.2 ECX IBM Copy Data Policy

The IBM Copy Data Policy captures and replicates storage volumes. The storage volumes can be captured in both an application-consistent and crash-consistent manner. Primarily, the IBM Copy Data Policy is meant to be used for storage volumes where applications are directly using the storage volumes. An example might include a Linux database running on a physical or virtual server where the database data disks are directly on IBM storage volumes. The IBM storage volumes would be mounted on the physical or virtual client by using a storage protocol such as iSCSI or Fibre Channel. If application consistency is required, ECX provides pre/post snapshot script options. If no pre/post scripts are specified in policy options, a crash-consistent capture is performed by the IBM Copy Data Policy.

Two major parameters should be specified for the policy:

- The source of the copy data
- The storage workflow

Options are available related to pre/post scripting for application consistency, notifications, and activation time. The source of copy data can be the entire storage system or individual storage volumes. The storage workflow is selected from a list of storage workflows that have already been created by the storage administrator. The storage workflow can also specify whether a replica of the IBM storage volumes is created on an auxiliary storage node. ECX achieves replicas of storage volumes by orchestrating Global Mirror with Change Volume and FlashCopy functions. See “Copy policies” on page 101 for more details.

3.3 ECX Copy workflow examples

In this section, a detailed walkthrough is provided for both VMware Copy Data Policy and IBM Copy Data Policy.

3.3.1 Creating a VMware Copy Data Policy example

This example assumes that VMware hypervisor and IBM storage nodes have been configured in ECX, as described in “Configuring an IBM Storwize V7000 in ECX” on page 24 and “Configuring a VMware vCenter server instance ECX” on page 27. Storage workflows should also be configured before creating a VMware Copy Data Policy. Usually these workflows are created by your storage administrator. See “Storage workflows” on page 96 for more details.

1. Log in to your ECX web GUI. Open a compatible web browser and go to the following URL:

   `http://<ECX hostname>:8443/portal`

2. Log in with your ECX user ID and password. The default administrator user name is *admin* and password is *password*.
3. Click **Plan** → **Policies** → **New**.

4. In the **Copy Data** column, click **VMware**, as shown in Figure 3-1.

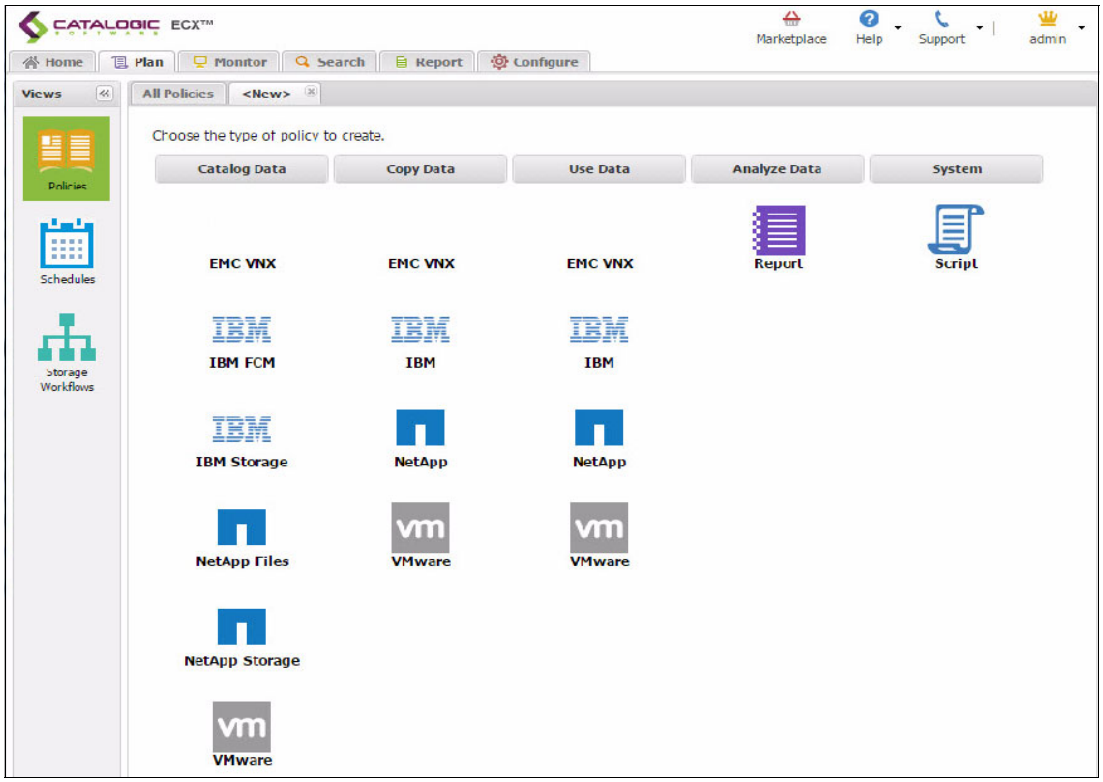


Figure 3-1 The New Policy window in the ECX management portal

5. The **New VMware Copy Data Policy** wizard opens as shown in Figure 3-2.

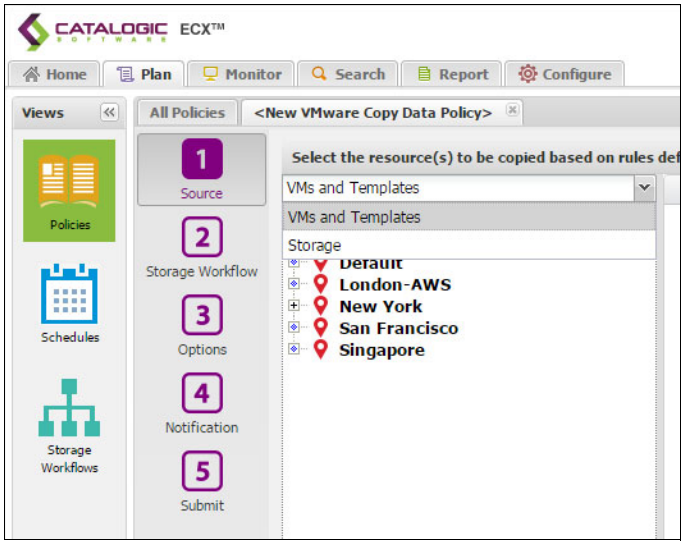


Figure 3-2 VMware Copy Data Policy source selection in the EVC management portal

6. The first step, Source, requires you to select the resources of the policy, VMware objects. For resource selection, users can use the **VMs and Templates** view or **Storage** view.
7. The **VMs and Templates** view allows users to select vCenters, DataCenters, Virtual Machine Folders, vApps, and Virtual Machines.

8. The **Storage** view allows selection of Datastores. This example uses the **VMs and Templates** view. In both views, the objects are grouped by ECX **Sites**. For more information about sites, see “Sites” on page 22.
9. As shown in Figure 3-3, the site **New York** is expanded to reveal the VMware vCenters at that location. Datacenters **PSvCenter1** and **PS_SE-Datacenter** are expanded. Within the datacenter, this environment has several applications grouped in VM Folders. The VM Folder **2_Marketing** is selected. By selecting the VM Folder, all virtual machines inside that folder are selected for the policy. In the future, if any virtual machines are created within this particular VM folder, ECX will automatically, at policy run time, include these virtual machines into the policy.

Note: ECX does not require users to specify where the virtual machines are on the IBM storage volume. Virtual machines can also move to different VMware datastores and thus move to different IBM storage volumes. ECX VMware Copy Data Policy, at run time, automatically determines the location of virtual machine storage volumes and perform the appropriate FlashCopy or Global Mirror operations. ECX also captures the configuration of the VM at a specific point in time and stores this data in the ECX catalog. This information is used by the ECX Use Data Policy.

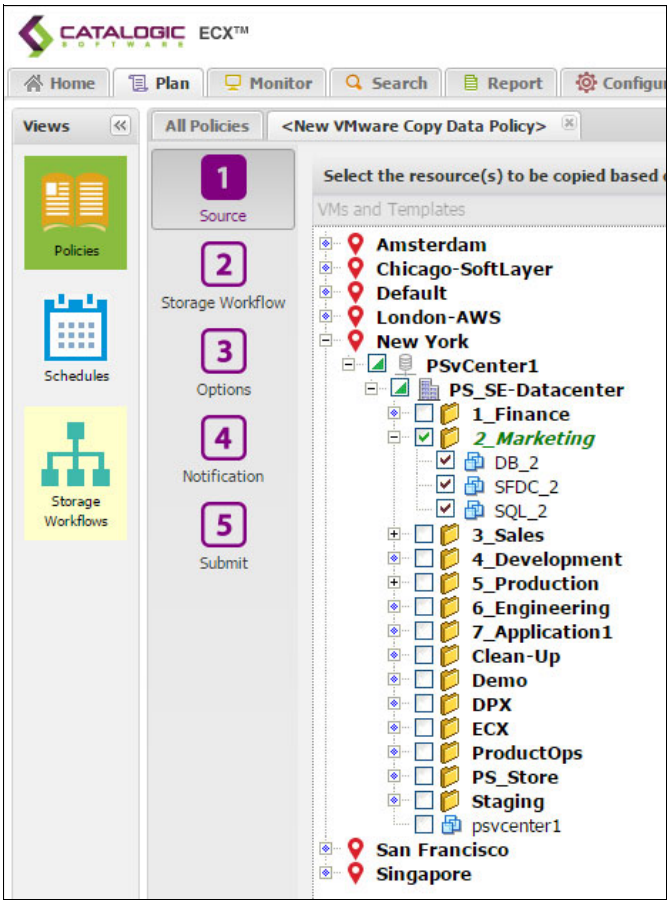


Figure 3-3 VMware Copy Data Policy VM and Templates Tree source selection

10. Figure 3-4 shows the Storage Workflow step in the wizard, where a storage workflow must be selected. Select the storage workflows that are most appropriate for your requirements. The storage workflows determine how the virtual machines in the policy are captured, the type of replication if any, the frequency of FlashCopies, and the retention of the FlashCopies. See “Storage workflows” on page 96 for more details. At minimum, one storage workflow must be added to the policy. Multiple storage workflows can be added. The **Specify activation time** option allows you to select the time of day when the policy will run. The actual frequency of policy execution is specified in the storage workflow itself. Select a storage workflow and click **Add** to add the storage workflow to the policy, then click **Next** to proceed to step 3 in the wizard, **Options**. In this example, the **IBM Gold Daily** storage workflow is selected.

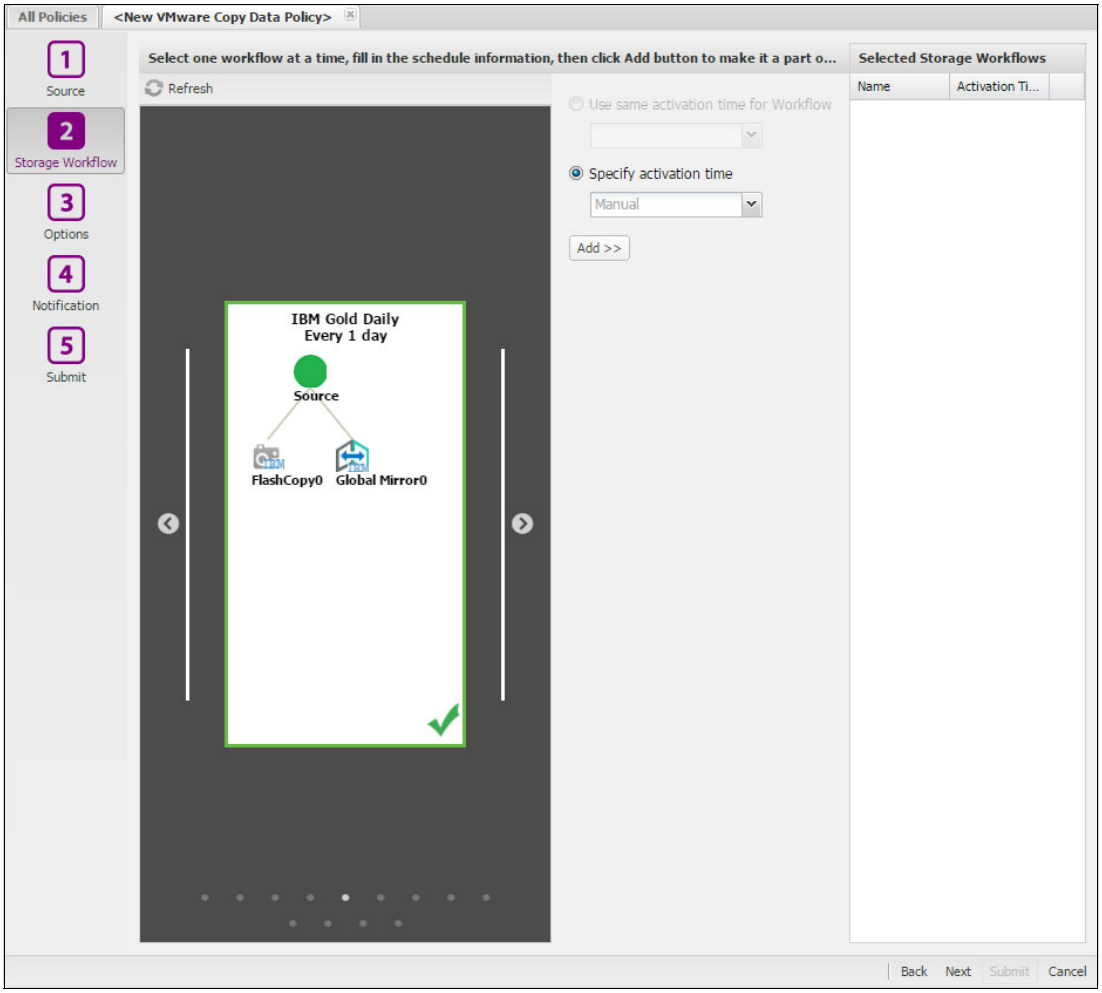


Figure 3-4 VMware Copy Data Policy Storage Workflow selection

11. The **Options** step, which is shown in Figure 3-5 on page 45, presents several policy options:

    a. The **Maximum concurrent tasks** option limits the number of parallel operations that can be executed during the policy run.

    b. The **Create VM snapshots for all VMs** option creates VM Snapshots before a FlashCopy is performed on the storage volumes. This process results in a VM consistent snapshot. If not selected, only the FlashCopy is executed at policy run time, resulting in a crash consistent copy.

Note: The VM snapshot is transient, and will be deleted after the FlashCopy is complete. The FlashCopy performed by ECX is a “NOCOPY” FlashCopy.

c. The **Maximum concurrent snapshots on ESX** option specifies the maximum number of VMware snapshots that occur in parallel. Tune this parameter based on the load and configuration of your VMware ESX server. Increasing this number might decrease the run time of the copy data policy. However, it might also put more load on the VMware ESX server.

d. The **Make all VMs application/file system consistent** option uses Microsoft Volume Shadow Copy Service (VSS) and creates an application consistent capture as part of the VM snapshots. Any applications that support the Microsoft VSS framework will be made consistent. In addition, the file system itself will be made consistent.

e. The **Make these VMs application/file system consistent** option allows you to select individual virtual machines within the copy data policy for application consistency.

f. The **Skip readonly datastores** option allows you to skip datastores that are mounted as read-only in VMware VCenter.

g. The options **Pre-Script** and **Post-Script** allow you to specify scripts that run during the VMware Copy Data Policy job. Both options can take Linux shell commands or the absolute path to a Linux shell script on the ECX server. The task specified as **Pre-Script** runs as the first task in the policy job, whereas the **Post-Script** is the last task to run during the policy job. This option is useful for quiescing and unquiescing applications. This option is also useful for integrating with third-party tools. See 3.4, “Application consistency and custom scripting” on page 51 for more detail.

h. The **Continue operation if Pre script fails** option allows you to specify whether the policy job should fail if the command or script that is specified in the **Pre-Script** option returns a nonzero exit status code. The exit status of the Pre and Post scripts is also logged in the ECX job history.

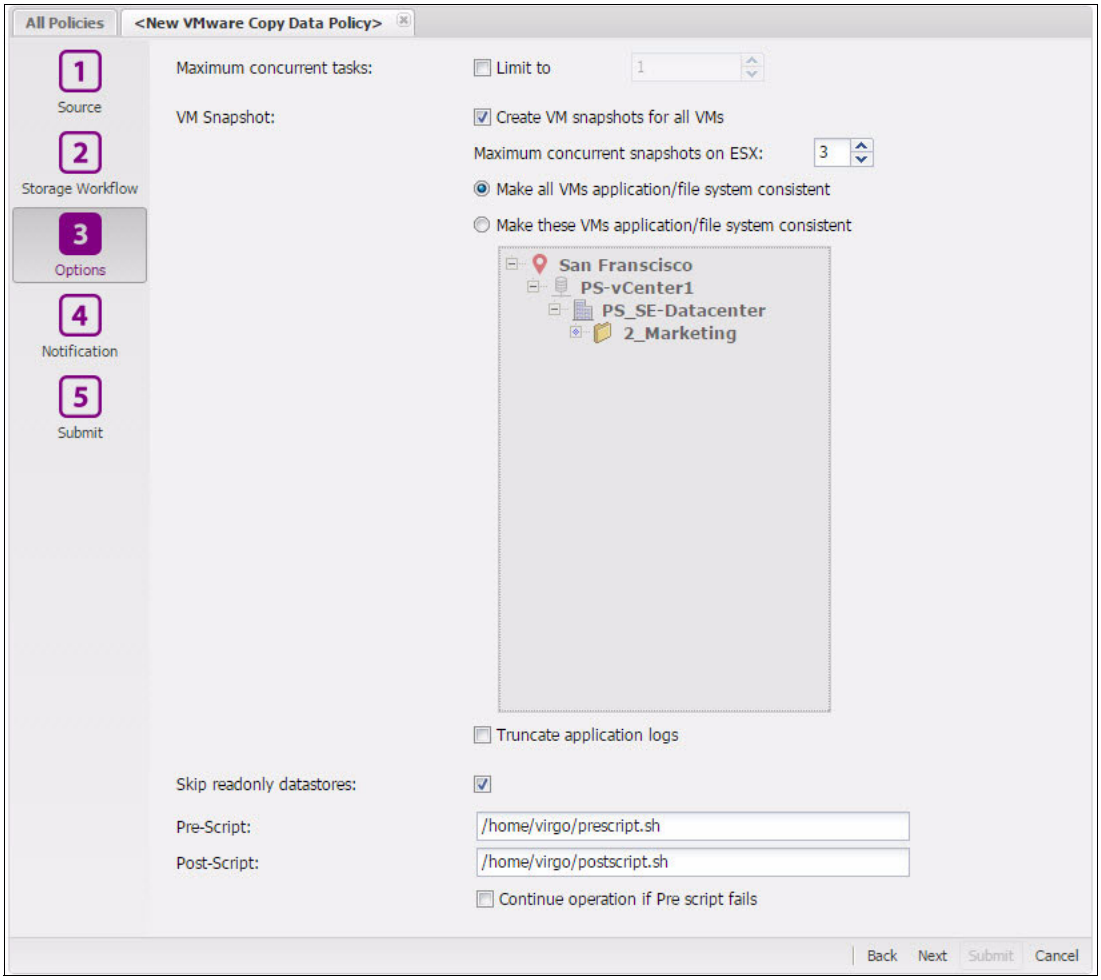


Figure 3-5 VMware Copy Data Policy Options

This example has VM snapshots with application/file system consistency for all VMs, and skips readonly datastores. It specifies to run the pre script at `/home/virgo/prescript.sh`, and the post script at `/home/virgo/postscript.sh`. Select the appropriate options and click **Next** to proceed to step 4, Notification.

12. On the **Notification** window of the wizard (Figure 3-6), you can specify an SMTP server and email recipients who receive notifications about job status for this policy. STMP server should be configured before this policy is created. See “Configuring an SMTP server in ECX” on page 30 for details about configuring an SMTP server in ECX. The example does not use email notifications, so these options are left blank. Click **Next** to proceed to step 5 in the wizard.

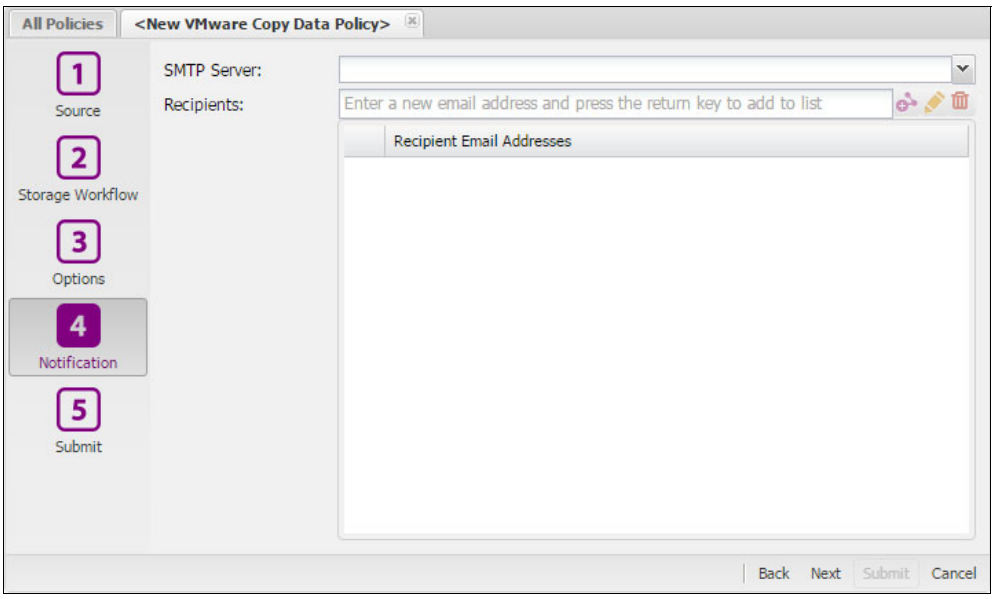


Figure 3-6 VMware Copy Data Policy Notification settings

13. Step 5, **Submit**, the final step in the wizard, is shown in Figure 3-7. Here, enter a name for the policy. An optional description can be specified. This window provides a summary of the policy, including the source, storage workflow, and options selected. Click **Submit** to create the policy.

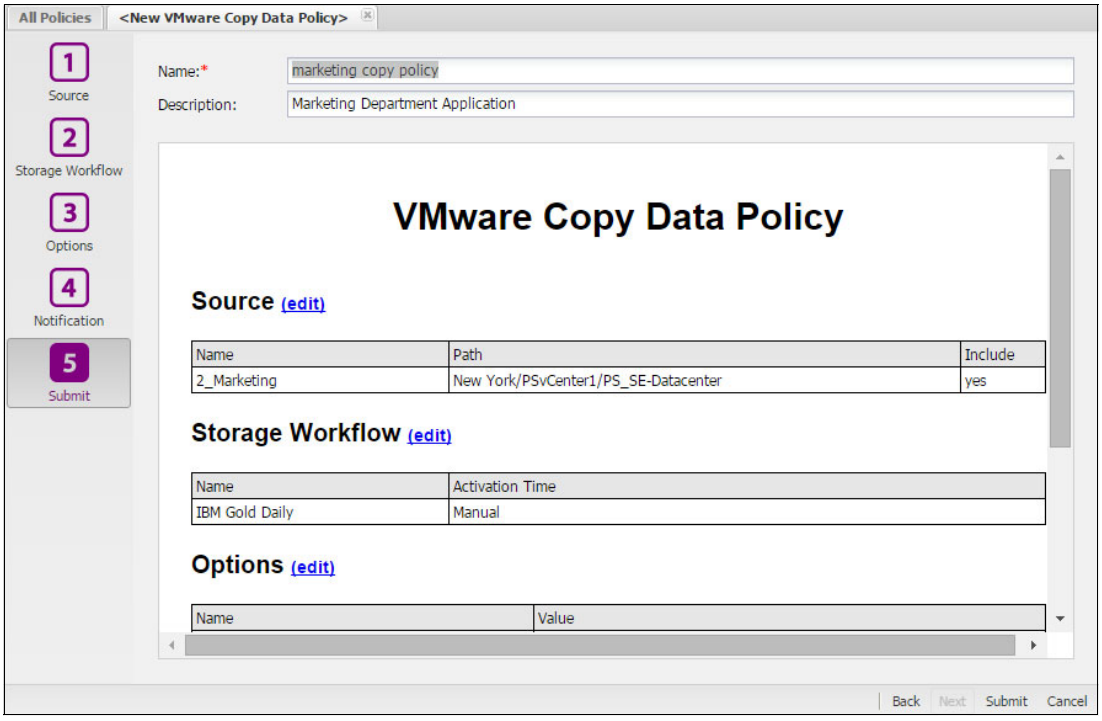


Figure 3-7 VMware Copy Data Policy Summary and Submit Screen

14. The VMware Copy Data Policy is now complete. The policy executes at the activation time that is specified, and the frequency of execution is determined by the storage workflows selected.

3.3.2 Creating an IBM Copy Data Policy example

This example assumes that IBM storage nodes have been configured in ECX, as described in “Configuring an IBM Storwize V7000 in ECX” on page 24. Storage workflows should also be configured before creating an IBM Copy Data Policy. Usually these workflows are created by your storage administrator. See “Storage workflows” on page 96 for more details.

1. Log in to ECX web GUI. Open a compatible web browser and go to the following URL:

   `http://<ECX hostname>:8443/portal`

2. Log in with your ECX user ID and password. The default administrator user name is *admin* and password is *password*.

3. Click **Plan** → **Policies** → **New**.

4. In the **Copy Data** column, click **IBM**, as shown in Figure 3-8.

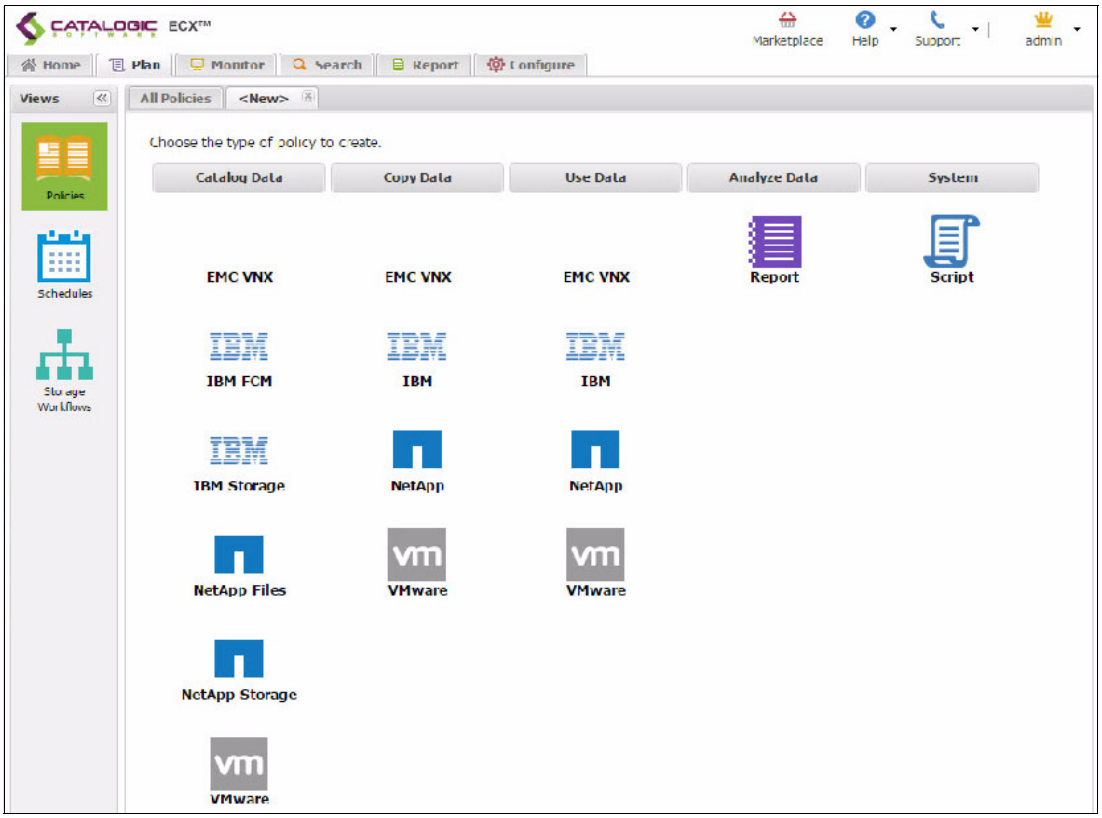


Figure 3-8 ECX New Plan Screen

5. The New IBM Copy Data Policy wizard, shown in Figure 3-9, opens. The first step, **Source**, requires you to select the resources of the policy, IBM storage objects. For resource selection, you can select entire storage nodes or individual storage volumes. The IBM storage nodes are grouped by ECX **Sites**. For more information about sites, see “Sites” on page 22. By default, FlashCopy target volumes are automatically filtered from the resource view. You can toggle this filter off by clearing **Exclude FlashCopies**.

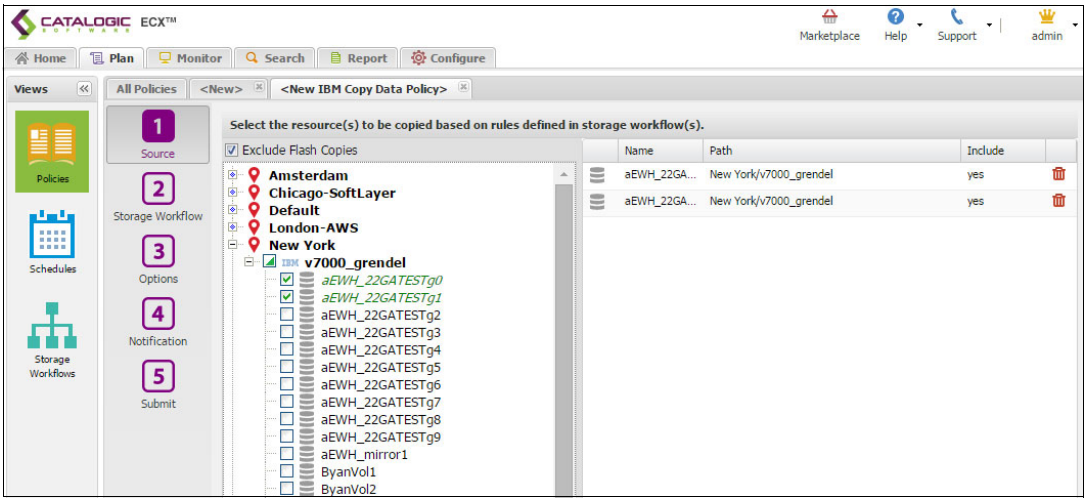


Figure 3-9 IBM Copy Data Policy Source Selection

6. Expand the resource tree, select the wanted IBM volumes, and click **Next** to proceed to step 2 in the wizard. In this example, the **New York** site is expanded, as shown in Figure 3-9. Within this site, select two volumes under a Storwize V7000.
7. In step 2 in the wizard, **Storage Workflow**, as shown in Figure 3-10, a storage workflow must be selected. Here select the storage workflows that are most appropriate for your requirements. The storage workflows determine how the virtual machines in the policy are captured, the type of replication if any, the frequency of FlashCopies, and the retention of the FlashCopies. See “Storage workflows” on page 96 for more details. At minimum, one storage workflow must be added to the policy. Multiple storage workflows can be added. The **Specify activation time** option allows users to select the time of day when the policy will run. The actual frequency of policy execution is specified in the storage workflow itself. Use the carousel storage workflow selector to select a storage workflow and click **Add** to add the storage workflow to the policy, then click **Next** to proceed to step 3 in the wizard, **Options**. In this example, the **IBM Gold Daily** storage workflow is selected.

Important: Any pre/post scripts that are needed for application consistency should be specified in the storage workflows.

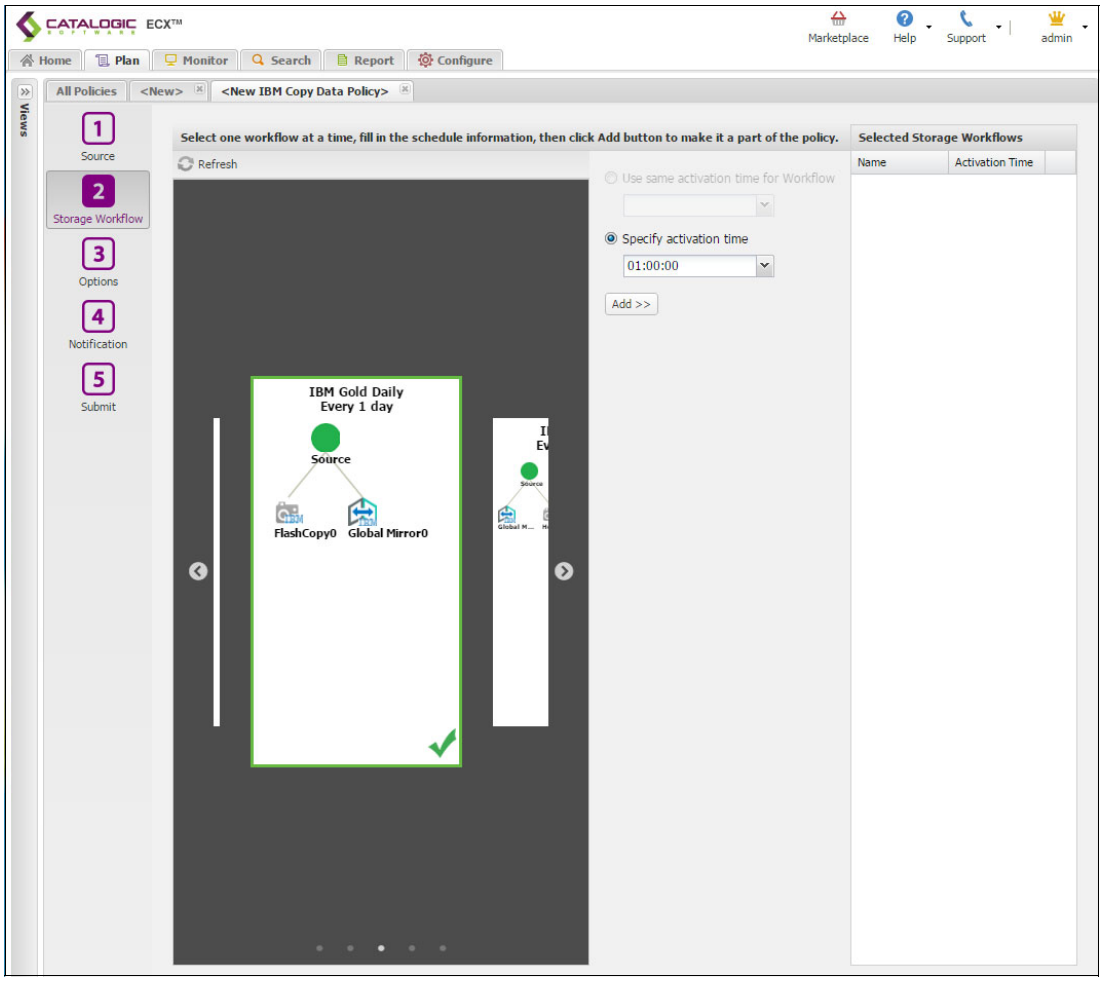


Figure 3-10 IBM Copy Data Policy Storage Workflow selection

8. The **Options** step, which is shown in Figure 3-11, presents several policy options:
   a. The **Skip the FlashCopy Target Volumes** option is selected by default. This option causes the policy to skip any volumes that are FlashCopy target volumes.
   b. The **Maximum concurrent tasks** option limits the number of parallel operations done during the policy run. By default, this option is not selected.
   c. The options **Pre-Script** and **Post-Script** allow you to specify scripts that run during the policy job. Both options can take Linux shell commands or the absolute path to a Linux shell script on the ECX server. The task sepcified in **Pre-Script** is run as the first task in the policy job, and the one specified in **Post-Script** is the last task to run during the policy job. See 3.4, “Application consistency and custom scripting” on page 51 for more detail.
   d. The **Continue operation if Pre-Script fails** option allows you specify whether the policy job should fail if the command or script that is specified in the **Pre-Script** option returns a nonzero exit status code. The exit status of the Pre and Post scripts will also be logged in the ECX job history.

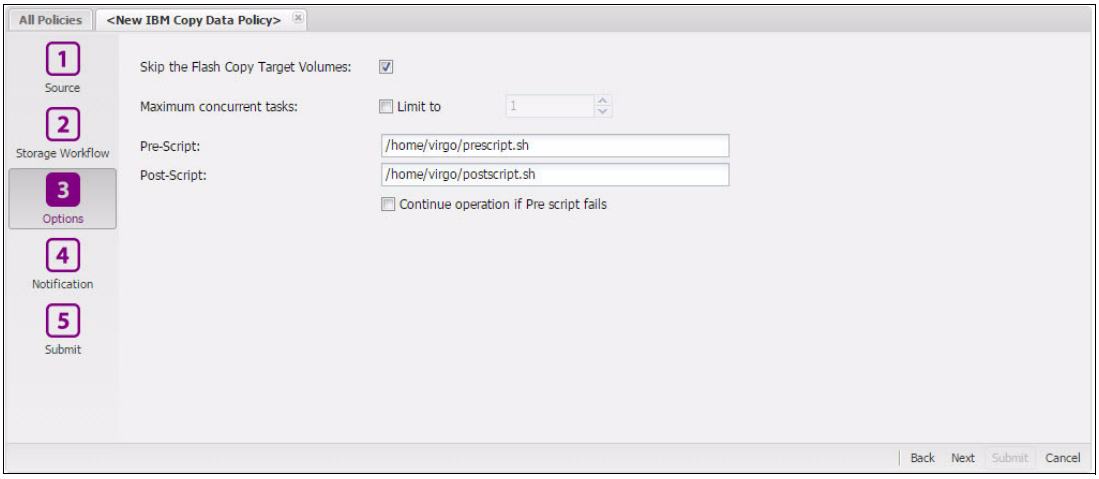


Figure 3-11 IBM Copy Data Policy Options window

9. On the **Notification** window of the wizard (Figure 3-12), you can optionally specify an SMTP server and email recipients who will receive notifications about job status for this policy. The STMP server should be configured before this policy is created. Refer to “Configuring an SMTP server in ECX” on page 30 for details about configuring an SMTP server in ECX. This example skips email notifications, and so both these options are blank. Click **Next** to proceed to step 5 in the wizard.

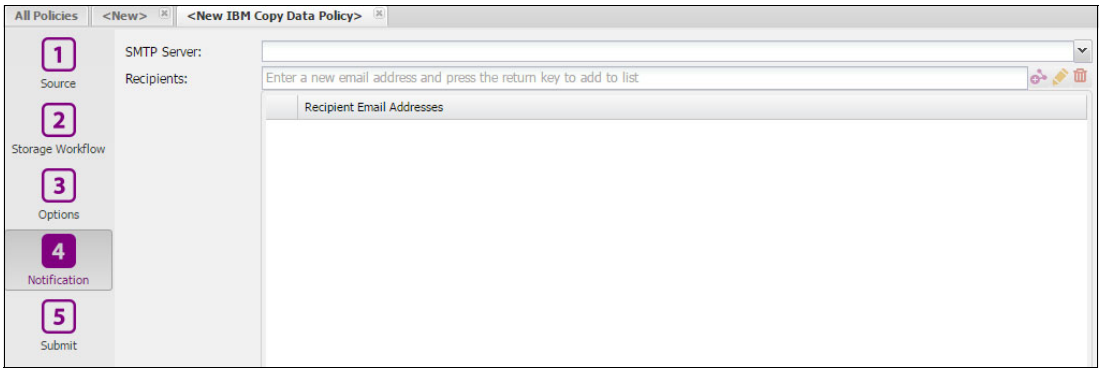


Figure 3-12 IBM Copy Data Policy Notification window

10. Step 5, **Submit**, the final step in the wizard, is shown in Figure 3-13. Enter a name for the policy. An optional description can be specified. This window also provides a summary of the policy, including the source, storage workflow, and options selected. Click **Submit** to create the policy.

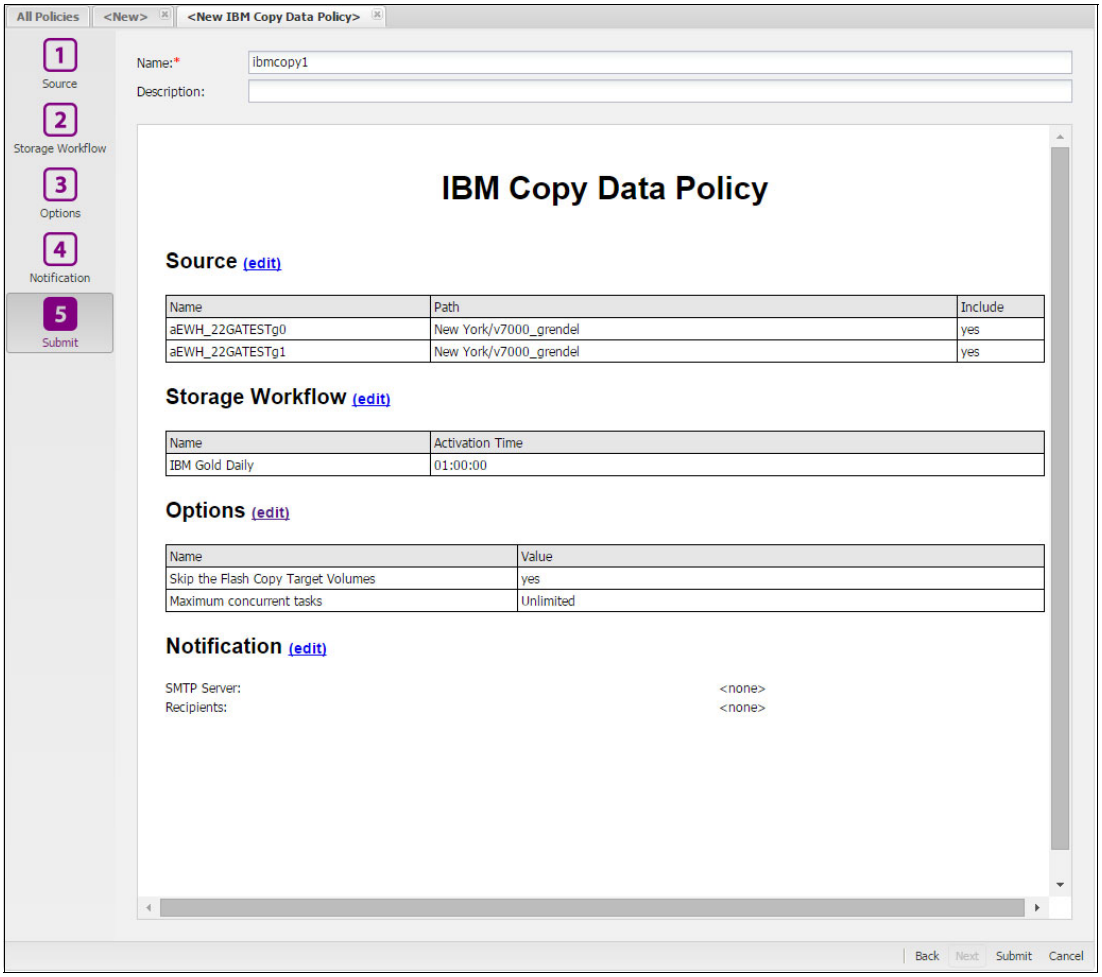


Figure 3-13 IBM Copy Data Policy Summary and Submit window

11. The IBM Copy Data Policy is now complete. The policy executes at the activation time that is specified, and frequency of execution is determined by the storage workflows selected.

3.4 Application consistency and custom scripting

This section provides an overview of how ECX achieves application consistency, both natively and through integration with third-party tools. Some examples of applications include Microsoft SQL Server (MSSQL), Microsoft Exchange, Microsoft SharePoint, Oracle, and IBM DB2®. To achieve application consistency, ECX must ensure that when a Copy Data Policy performs a capture of applications that are online and performing I/O on storage volumes, that the capture represents a logically consistent point in time. If application consistency is not enabled, a crash-consistent capture is performed by ECX Copy Data policies. ECX can support application consistency on both virtual and physical environments. In addition to application consistency, ECX handles transaction log management and custom scripting. This section provides detail on the different methods used for application consistency, transaction log management, and custom scripting.

3.4.1 VMware consistency

For VMware vSphere environments, ECX uses vSphere VM Snapshot functions to provide VM and application consistency. During an ECX VMware Copy Data Policy, ECX takes a VM Snapshot before a FlashCopy or VM Copy is performed. On virtual machines running Windows operating systems, a VM snapshot that uses VMware Tools enables Microsoft VSS and engages the necessary VSS writers for any application that provides VSS writers. This process ensures that the VM snapshot contains a logically consistent point-in-time copy of application data. Some popular applications that provide VSS writers include MSSQL, Exchange, SharePoint, Active Directory, and Oracle. On virtual machines that run Linux, a VM Snapshot, using VMware Tools, runs the freeze and thaw scripts during the VM snapshot process.

Microsoft SQL transaction log management

For MSSQL, an ECX VMware Copy Data Policy performs log truncation. ECX does not require the installation of agents inside the VM running MSSQL. Instead, ECX will, at run time, inject into the virtual machine the relevant code to perform log truncation. This process removes the administrative processor usage of installing, patching, and maintaining an agent inside each VM. You must provide credentials for the virtual machine in ECX so that code injection and execution can occur.

3.4.2 ECX Policy script integration

ECX provides the ability to specify pre and post scripts to run during the execution of the ECX Copy and Use Data policies. These scripts offer an alternative mechanism for achieving application consistency. See Figure 3-14 on page 53, where pre and post scripts are used to quiesce Oracle in an IBM Copy Policy.

Tip: ECX Pre and Post script integration can be used in both VMware Copy policies and IBM Copy policies, providing benefits for both physical and virtual servers.

The policy options **Pre-Script** and **Post-Script** are fields that can take Linux shell commands or the absolute path to a Linux shell script on the ECX server. The scripts run as the user “virgo” on the ECX server. The task marked as **Pre-Script** is run as the first task in the policy job, and the one marked **Post-Script** is the last task to run during the policy job. During Copy Data policies, the pre and post scripts are useful for quiescing and unquiescing applications before and after ECX performs VMware snapshots and IBM FlashCopies. In the case of Use Data policies, they can be used to perform custom recovery procedures or to integrate with DevOps or data masking tools.

The pre and post scripts are integrated with the ECX Policy execution and job status. ECX records the script output and exit status code of both the pre script and post script in the ECX job monitor and history. If the ECX Policy option **Continue operation if Pre script fails** is selected, the ECX Policy continues execution and sets failure status on the policy job if a nonzero exit status code is return by the pre script. The post script will not affect the execution of the policy job.

Note: For IBM Copy policies, applications that run on physical servers can specifically benefit from ECX pre and post script integration. The following section, “Oracle Application Consistency” on page 53, demonstrates how to use this feature to achieve application consistency for Oracle running on Linux. However, similar benefits can be achieved across many combinations of operating systems and applications. Examples include, but are not limited to, Oracle or DB2 running on Linux, IBM AIX®, Solaris, and HP-UX.

Oracle Application Consistency

ECX provides application integration with Oracle Real applic ation Clusters (RAC) and Automatic Storage Management (ASM) configurations on Linux and AIX. An example is an Oracle RAC configuration that consists of several compute nodes and shared storage on Spectrum Virtualized storage. During the ECX Copy Policy job, the Oracle database is quiesced by putting the database into backup mode. While the database is in backup mode, it remains online and operational. During the backup mode, extra information is written to Oracle redo logs so that the database can be recovered to a consistent point in time from the FlashCopy that is created by the ECX Copy Policy Job. After the FlashCopy is completed, the database is taken out of backup mode.

Note that a NOCOPY FlashCopy is performed, which is near instantaneous, allowing the Oracle database to remain in backup mode for a short time.

In the ECX Copy Policy, you can specify the IP address of the Oracle server node, instance parameters, and credentials by using the ECX Oracle Pre/Post scripts, as shown in Figure 3-14. Additionally, you can customize the Oracle quiesce and unquiesce scripts. Oracle pre and post scripts are available through the ECX Marketplace or directly from Catalogic Software Support. For more information about the Marketplace, see “Marketplace” on page 16.

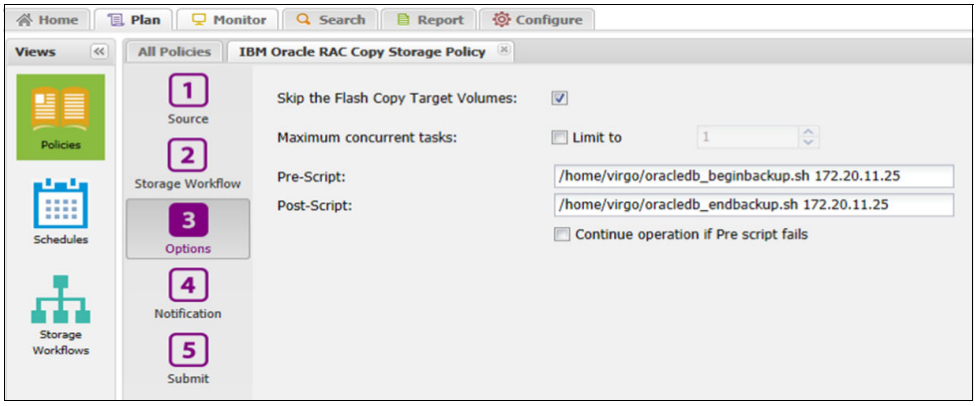


Figure 3-14 Oracle Pre-Script and Post-Script options

Note: The ECX Copy Policy job log and verifies status about the success or failure of Oracle quiesce/unquiesce pre and post scripts. The ECX Job fails and shows the appropriate status in job monitor and job history if the **Pre-Script** fails to put the Oracle database in backup mode. This integration ensures that ECX Copy Policy will always create a logically consistent point in time FlashCopy of the Oracle database.

Recovery of the Oracle database can be done using the ECX Use Data Policy by employing the **Instant Access** option. ECX Instant Access automatically locates the correct FlashCopy, depending on the particular point in time that is selected, and mount the Oracle database LUNs to the Oracle database servers. The advantage of ECX using IBM FlashCopy technology is that the recovery can be instantaneous. Even if you have a large database, it can be recovered from a local FlashCopy in minutes versus hours. After the LUNs are recovered, a database recovery procedure can be applied to the database to bring it online.

3.4.3 VMware Script Integration

For the VMware Copy Data policy, in addition to the Pre/Post script options described in “ECX Policy script integration” on page 52, VMware provides granular Pre/Post scripts at the VM level. This granular VMware scripting is activated if the option **Make all VMs application/file system consistent** or **Make thes VMs application/file system consistent** are selected in the VMware Copy Data Policy, as shown in Figure 3-5 on page 45. The scripts are specified by using the VMware Tools pre-freeze-script and post-thaw-script.

For more information, see:

http://kb.vmware.com/kb/1006671

4

Copy data orchestration and automation with Catalogic ECX

This chapter describes Catalogic's ECX Copy Data Orchestration and Automation with IBM Spectrum Virtualize storage.

This chapter includes the following sections:

- Orchestration and automation
- ECX Use Data Workflow technical overview
- Use cases
- Automated disaster recovery

4.1 Orchestration and automation

The ability to use data copies for multiple use cases is one of the key value drivers of copy data management. ECX automates and orchestrates otherwise complex data reuse scenarios for multiple business solutions such as DevTest /DevOps, Automated DR, and Data Analytics quickly and easily. ECX abstracts underlying hypervisor and heterogeneous storage arrays, and provides a unified web GUI and REST API to use the copy data in the organization. For GUI users, a web interface provides a single plane of glass of all hypervisor and storage infrastructure. The ECX REST API provides infrastructure as code to agile development teams. In both API and GUI interfaces, ECX role-based access control (RBAC) provides security at a granular level.

ECX allows IT to deliver services that are often otherwise failing or have been deemed next to impossible to implement. ECX drastically improves DevTest/DevOps workflows by giving development teams access to the hypervisor and storage by using a REST API. Business operations such as disaster recovery can be automated so that they can be tested and validated every day. The ECX platform enables the resources that are used for these business operations to be brought up in a fenced environment, promoted to production quickly, and cleaned up after a test. Orchestration provides repeatability and auditability, allowing you to use a single copy of your data for multiple purposes to reduce data sprawl and lower costs, reducing both CapEx and OpEx.

4.1.1 Using copy data

Every line of business requires a “copy” of production data in order to function. How ECX Copy Data workflows are used to create and replicate copies of your application was described in 3.2, “ECX Copy Workflow technical overview” on page 38. This chapter focuses on several prominent use cases for copy data to exemplify the impact that ECX can have on an IT environment: DevTest/DevOps, automated DR, and business analytics. ECX enables these use cases on IBM Spectrum Virtualize storage. Currently, IT supports each of these use cases by using data copies with similar time-consuming and archaic workflows that end up creating data proliferation. This chapter demonstrates how Catalogic ECX, in combination with IBM Spectrum Virtualize storage, can provide automation and orchestration to these workflows. This system allows IT to take advantage of a single copy of data for multiple use cases.

Copy data management is as much about controlling the creation of the data copies for the proper business solution as it is about controlling the use of those copies to gain efficiency and agility. The primary foundation for all copy data use cases is “Copy Data Leverage.” This is the ability to automate and orchestrate data using a common set of storage services to create the right copy of the data, and then use that single copy of data for multiple business solutions. The quickest, most effective way to provide instant data access is using storage technologies such as IBM FlashCopy, which allow instant access to a point in time of a storage volume without any data movement. Catalogic ECX can orchestrate VMware hypervisor and IBM Spectrum Virtualize Storage to allow applications to be available, instantly, for any number of business solutions.

4.2 ECX Use Data Workflow technical overview

In ECX, the method of using copy data in your environment is with ECX Use Data policies. Use Data policies automate the process of instantly standing up applications by using FlashCopies on primary or auxiliary storage. No movement of data is required. After the

application copy is used and no longer needed, ECX can clean up all resources on a scheduled or ad hoc basis, or move the application to production. As described in 3.2, “ECX Copy Workflow technical overview” on page 38, ECX catalogs all necessary information about the application during the ECX Copy Data Policy execution. Even if the primary servers and IBM storage nodes no longer exist, ECX can recover applications during a disaster from the secondary IBM storage node. ECX Use Data policies can be scheduled to start applications and clean up the resources on an hourly or nightly basis. Two different types of ECX Use Data policies are available: VMware Use Data policy and IBM Use Data policy. VMware Use Data policy uses copies of VMware vSphere applications on IBM storage, whereas IBM Use Data policy uses copies of applications that directly use IBM storage volumes.

The VMware Use Data policy provides two different workflow templates that it can run:

- Instant Access
- Instant Virtualization

In both cases, ECX orchestrates the hypervisor and storage operations that are needed to access FlashCopies on IBM storage systems. *Instant Access* allows users to mount an individual virtual disk (vmdk) onto an existing virtual machine. Here, the policy job finds the correct FlashCopy by searching the ECX catalog based on the point in time that is required. ECX then creates a FlashCopy of a FlashCopy to make sure that the original point in time snapshot is not disturbed. The VMware datastore on the copy of the FlashCopy is signed again and mounted onto the appropriate VMware ESX server. ECX then makes VMware vSphere API calls to attach the vmdk to the selected virtual machine. In the case of *Instant Virtualization*, ECX creates a new virtual machine with the same settings as the source virtual machines, and then mounts and signs the data store from the FlashCopy of the FlashCopy again. Before powering the virtual machines in the order specified in the policy, ECX also remaps the network settings at both the vSphere and operating system level.

The IBM Use Data policy provides two different workflow templates that it can run:

- Instant Access
- Restore Volume

With Instant Access, ECX locates the correct FlashCopy at the required site and creates a FlashCopy of a FlashCopy that can be mounted on a host through iSCSI or Fibre Channel. With Restore Volume, ECX can restore the contents of the FlashCopy onto the original source volume or a new volume.

Note: Because ECX Use Data Policy uses a FlashCopy of a FlashCopy, minimal space is used for each instance of an application because only new data writes use additional storage space. This system enables users to provide copies of applications to multiple users quickly and efficiently. ECX catalog knows about replicas that are created by Global Mirror, so the ECX Use Data policy uses the correct FlashCopies based on the location of your VMware ESX server. This process is accomplished by using ECX’s site concept where resources such as VMware vCenter and IBM Storage nodes are categorized by sites. For more information about sites, see “Sites” on page 22.

4.3 Use cases

This section provides detailed examples of how ECX, in combination with IBM Spectrum Virtualize storage, can be used for DevTest/DevOp and Automated Disaster Recovery. Use cases such as business analytics and providing infrastructure as code are also described.

4.3.1 DevTest/DevOps

DevOps, or a test-driven development (TDD) cycle, is a software development process that relies on the repetition of a short development cycle:

1. Write an (initially failing) automated test case that defines a wanted improvement or new function.
2. Produce the minimum amount of code to pass that test.
3. Promotes the code to production.
4. Refactor the new code to acceptable standards.

DevTest is a key component of the DevOps process. It involves developers performing defect prevention and detection during the development process. This process ensures that the code contains fewer defects before it enters QA and production.

As part of DevTest, it is important that developers are able to develop and test against an environment that is as close to production as possible. Traditionally, developers test applications in environments that might have different versions of operating systems and updates, and no production data, or old version of the production data. To improve on this system, developers can ask their virtual and storage admins to provision a virtual environment with a copy of the most recent production database. However, this procedure is not only tedious and lengthy for the system administrators, but also uses valuable resources.

If you have many developers who need their own copy of an application environment, the time, compute, and storage resources needed are multiplied. Due to the cost and complexity of setting up these environments, developers might not have the latest version of the application and data to test with. If developers can be provided with the latest version of their applications on-demand or on an hourly or daily basis, the code quality can be increased significantly. With ECX Use Data policies and IBM Spectrum Virtualize storage, this process can be accomplished quickly, automatically, and with resource efficiency. This section describes a walk-through on how to use ECX to make DevTest more efficient.

ECX Use Data Policy example for DevTest

In this example, developers need access to the latest copy of a marketing application daily. It consists of three virtual machines. A VMware Use Data policy is created that produces a copy of the application daily in an isolated test environment. On subsequent executions of the policy, all resources such as virtual machines and FlashCopies are cleaned up before the next daily copy of the virtual machines is created.

To create the example Use Data Policy, complete these steps:

1. Log in to the ECX web GUI. Open a compatible web browser and go to the following URL:

   `http://<ECX hostname>:8443/portal`

2. Log in with your ECX user ID and password. The default administrator user name is *admin* and password is *password*.
3. Click **Plan** → **Policies** → **New**.

4. In the Use Data column, select **VMware**, as shown in Figure 4-1.

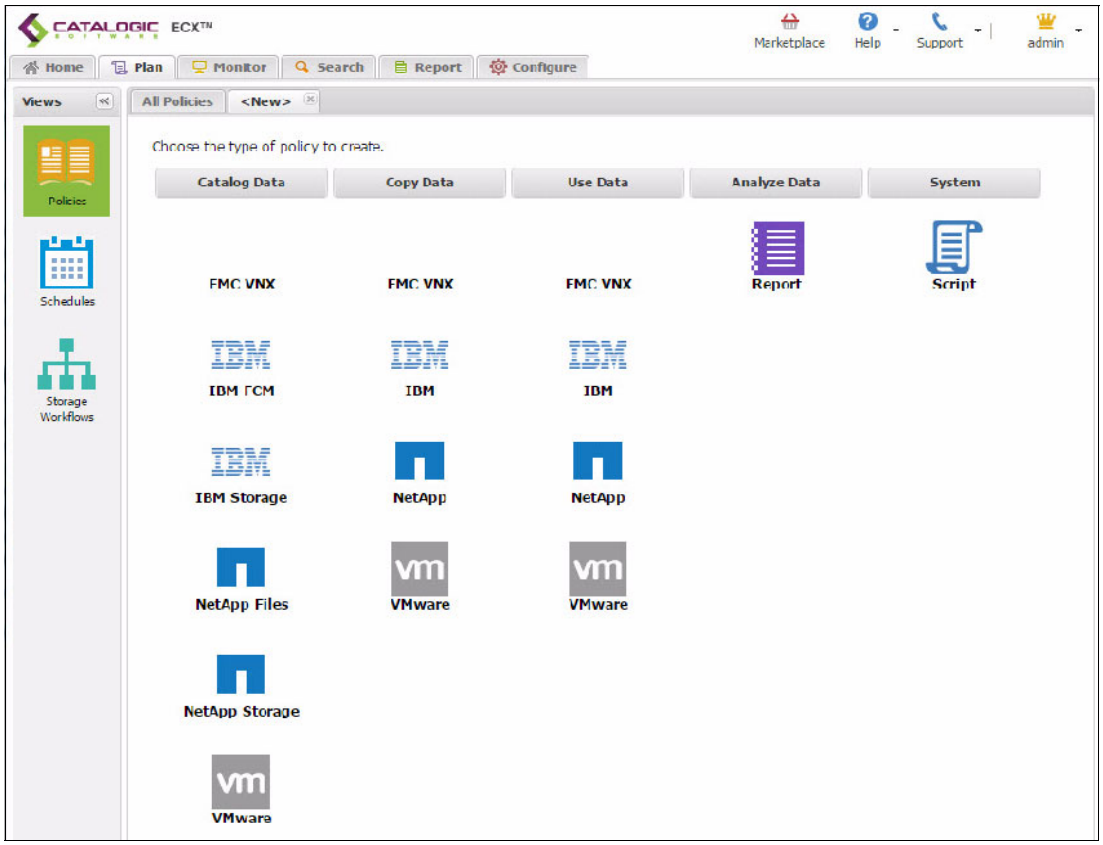


Figure 4-1 ECX New Plan Tab

The **New VMware Use Data Policy** wizard loads as shown in Figure 4-2.

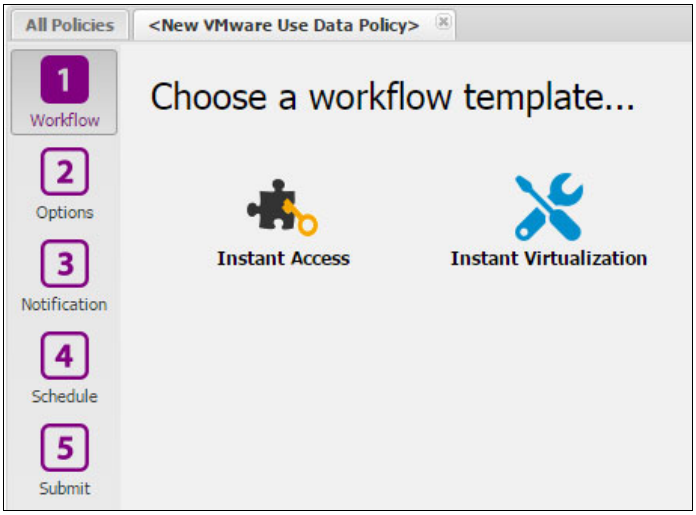


Figure 4-2 VMware Use Data Policy workflow template

5. The first step, **Workflow**, requires you to select a workflow template, as described in 4.2, “ECX Use Data Workflow technical overview” on page 56. **Instant Access** instantly mounts a vmdk from a FlashCopy to an existing data store. **Instant Virtualization**, in test mode, creates a virtual machine with the same VM properties as the source VM. However,

it creates the VM under a different name and either places it in the same vCenter and ESX host, or places it onto a different vCenter and ESX host, depending on your selection in step 9 on page 61. For this example, **Instant Virtualization** is chosen.

6. After the workflow template is chosen, the **New VMware Use Data Policy** wizard allows you to select VMware VMs that will be part of the Instant Virtualization, as shown in Figure 4-3. Here the wizard shows a **VMs and Templates** view. A tree browser shows only the VMs that have been captured for ECX VMware Copy Data policies. The VMware objects are categorized by the ECX site at the top level. In the tree selection, VM folders or individual VMs can be selected. For this example, the **New York** site is expanded, along with the vCenter **PSvCenter1**, data center **PS_SE-Datacenter**, and the **2_Marketing** folder. Within the **2_Marketing** folder, three VMs are selected.

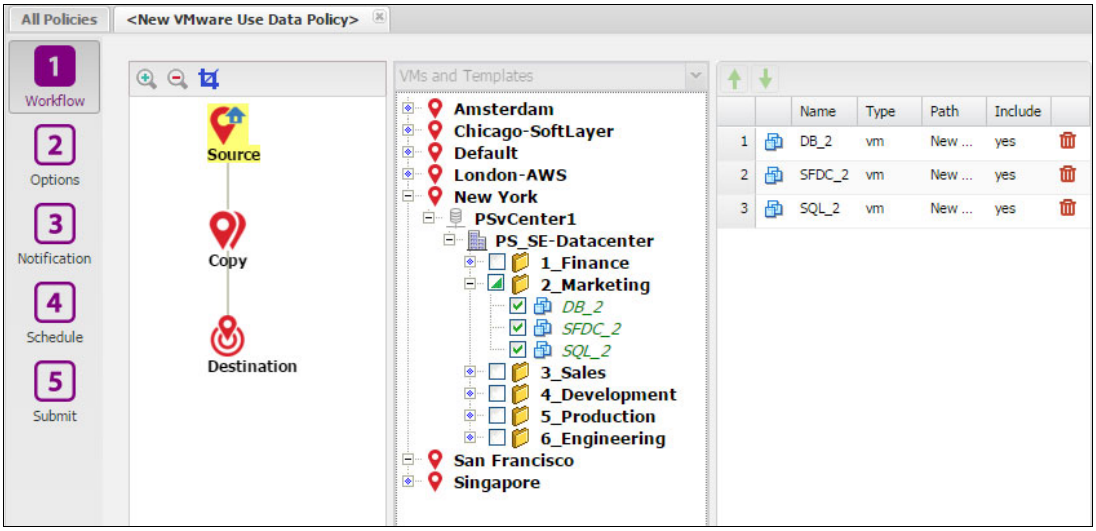


Figure 4-3 VMware Use Data Policy source selection

7. After the source selection is complete, the selected virtual machines that will be part of the Instant Virtualization are shown in the right pane, as shown in Figure 4-3. The order of the list determines the order in which the virtual machines are started. To change the order, highlighting the VM in the list and click the **Up** or **Down** arrows.

8. Click the **Copy** icon in the left pane to show the sites where FlashCopies of the selected virtual machines are available, as shown in Figure 4-4 on page 61. Select a site for ECX to use when this policy runs. In this example, **New York** and **Amsterdam** are available with FlashCopies of the virtual machines selected in step 6 on page 60. In this example, the **New York** copies are selected. By default, the FlashCopy with the latest version of each virtual machine is used. Optionally, a previous point in time copy can be selected by using the **Select Version** button.

> **Note:** Any FlashCopies that were created by using FlashCopy Manager for VMware are also contained in the ECX catalog if an IBM Tivoli Storage FlashCopy Manager Catalog Policy is configured in ECX. Thus, Tivoli Storage FlashCopy Manager FlashCopies can be used in the VMware Use Data Policy. See “IBM Spectrum Protect Snapshot” on page 7 for more information.

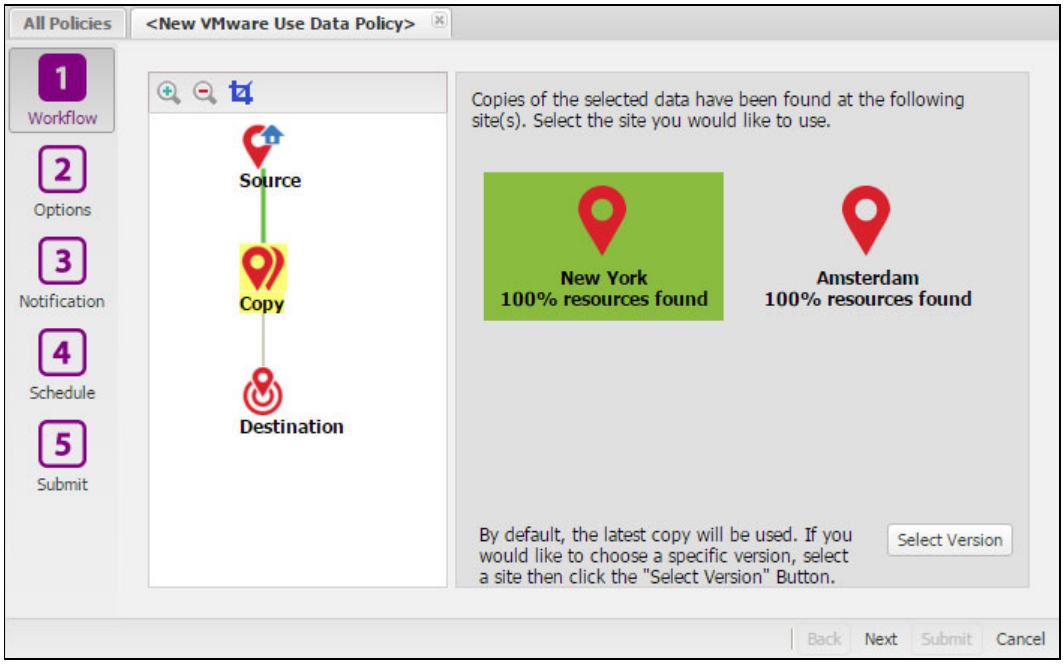


Figure 4-4 VMware Use Data Policy Copy Selection

9. Click the **Destination** icon in the left pane, as shown in Figure 4-5, to select the VMware ESX host or cluster that the test virtual machines will be created on. By default, the **Use original host or cluster** check box is selected, which creates virtual machines with different names from the source on the original ESX host or cluster. When this check box is cleared, you can not only select a different VMware ESX host or cluster, but also perform IP address and VM network remapping and set an alternate data store destination for clone operation. In this example, the **Use original host or cluster** check box is cleared so you can select an alternative ESX host.

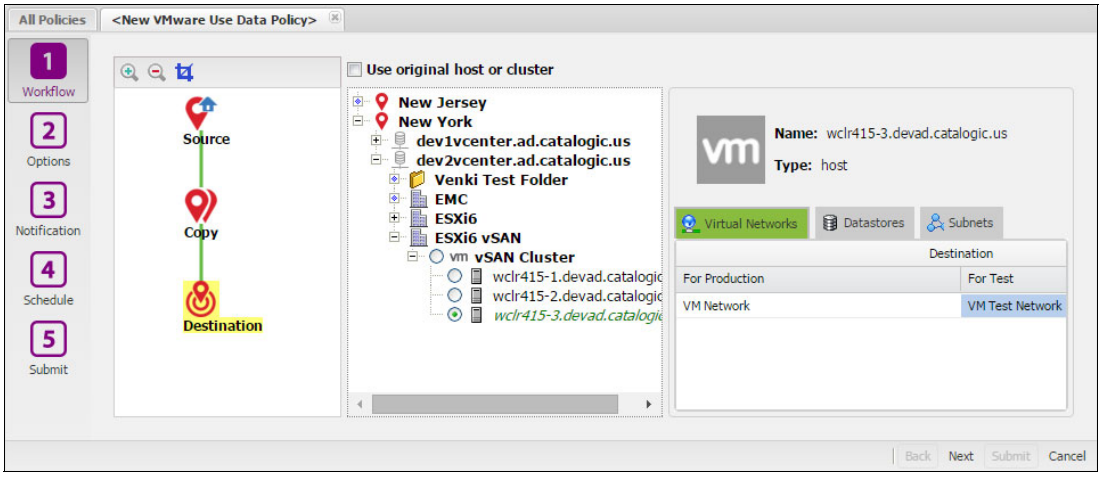


Figure 4-5 VMware Use Data Policy Destination Selection

10. When the **Use original host or cluster** check box is cleared, three tabs are visible in the right pane:
    - Virtual Networks
    - Datastores
    - Subnets

The Virtual Networks tab, as shown in Figure 4-5 on page 61, enables you to select the virtual networks for both Production and Test modes. Even though Instant Virtualization will start in a particular mode, after it is running, it can be changed from Test mode to Production or Clone mode. To do so, you must select the VM production and test networks to use in those modes. In this example, the **VM Network** is selected for production and **VM Test Network** is selected for test. The **Network Selection** box displays the currently available networks on the destination VMware ESX host or cluster.

11. The Datastores tab (Figure 4-6) enables you to select destination data stores for production and clone modes. When the Instant Virtualization job starts (or transition to) production mode or clone mode, the virtual machines run from a FlashCopy of a FlashCopy. While the VM is running, the Instant Virtualization job uses vMotion to migrate the virtual machine storage from the FlashCopy to this selected data store. This process allows the FlashCopy to be aged, and the production or clone of virtual machines to be on permanent storage. While the vMotion process is active, the virtual machines can remain powered on and functioning. In test mode, the data store destination will not be used.

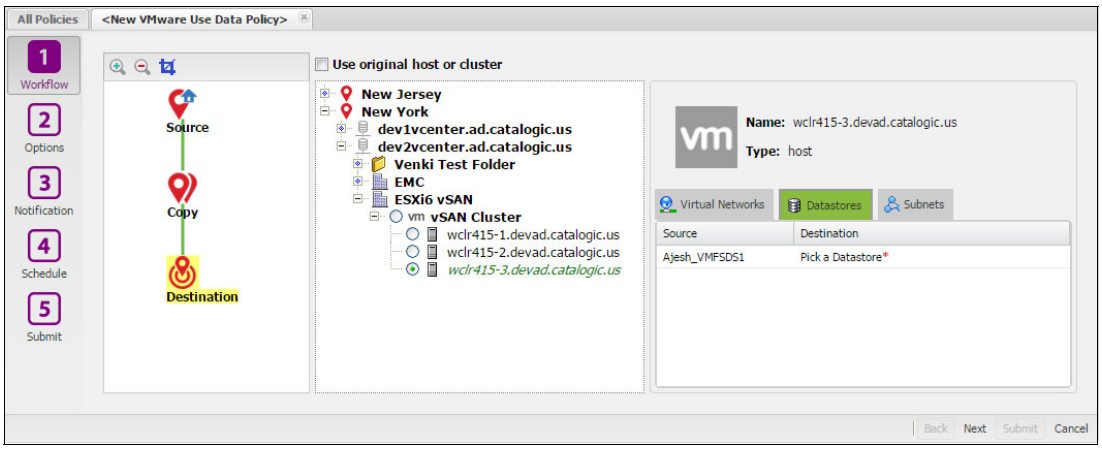


Figure 4-6 VMware Use Data Policy Destination Datastore selection

12. The Subnets tab (Figure 4-7) allows you to remap IP addresses and operating system network settings within the VM. During the Instant Virtualization job, ECX agentlessly accesses the network properties of the operating system within each virtual machine when performing Instant Virtualization to sites that can have different network topology. In this example, virtual machines in subnet **192.154.0.0** are remapped to the **192.11.0.0** subnet. Users can also specify individual IP addresses instead of subnets. Remapping by subnet is less tedious if you have many virtual machines in a VMware Use Data Policy. In addition to IP addresses, the DHCP option, Gateway, and DNS entries can be remapped.

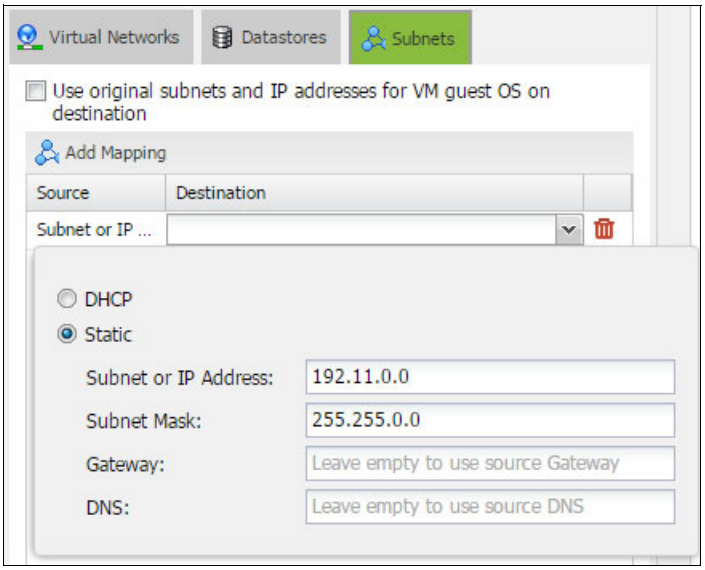


Figure 4-7 VMware Use Data Policy Subnet details

After you complete the destination selections, click **Next** to proceed to step 2 in the wizard. Step 2, Options, loads as shown in Figure 4-8.

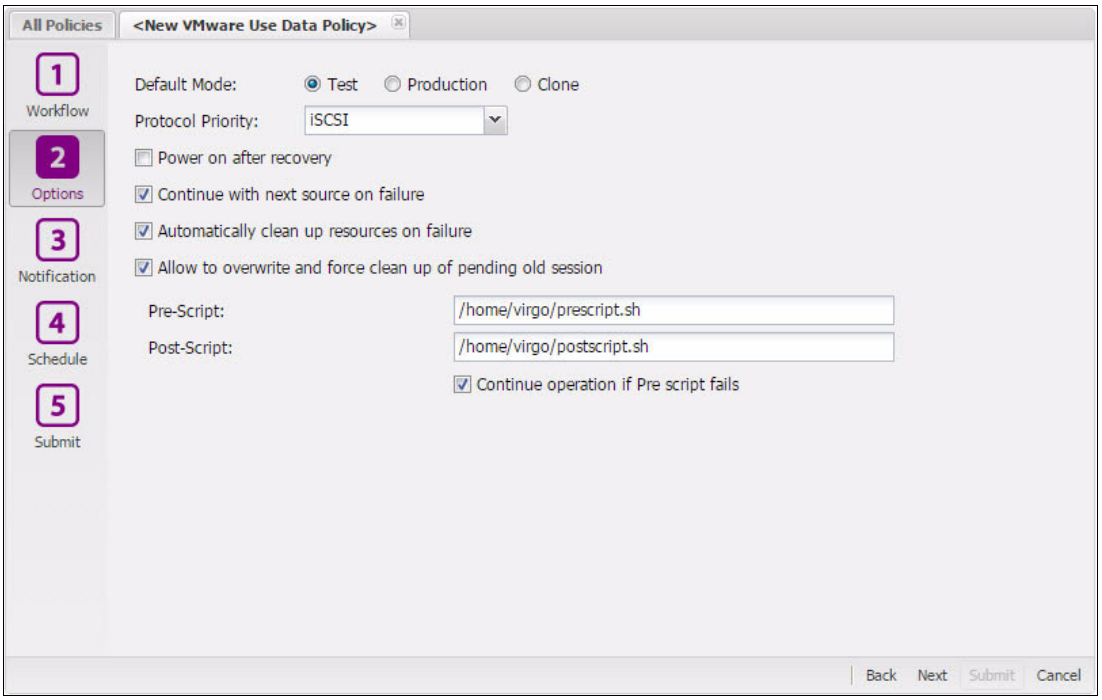


Figure 4-8 VMware Use Data Policy Options

13. The Options window provides the following options:

    a. The first option, **Default Mode**, specifies the default mode of the policy when it runs. The modes are **Test**, **Production**, and **Clone**. Test mode creates virtual machines with the original name of the source virtual machine appended with a "_test" string. Your production source virtual machines will not be affected. In test mode, virtual machines are attached to the test network. In Production mode, the Instant Virtualization job first deletes the original source VM, and a new VM with the same name and configuration is created. The virtual machine storage is on the FlashCopy, and a vMotion is started to migrate the virtual machine hard disks to a production data store. In Clone mode, a new virtual machine is created with the name of source virtual machine appended with "_clone". Agai,n the virtual machine storage is on the FlashCopy and a vMotion is run to migrate the virtual machine to permanent storage.

    b. The **Protocol Priority** option specifies the protocol to take priority in the policy. Available protocols include iSCSI and Fibre Channel.

    c. The **Power on after recovery** option specifies whether the virtual machines are powered on during the Instant Virtualization job. If not selected, the virtual machines can be manually powered on by using the VMware vSphere interface.

    d. The **Automatically clean up resources on failure** option enables the automatic cleanup of allocated resources as part of a restore if the virtual machine recovery fails. This option is selected by default.

    e. The **Allow to overwrite and force clean up of pending old sessions** option allows the Instant Virtualization job to clean up resources held by currently pending sessions. This option is selected by default.

    f. The **Pre-Script** and **Post-Script** options allow you to specify scripts that will run during the policy job. Both the **Pre-Script** and **Post-Script** options can take Linux shell commands or the absolute path to a Linux shell script on the ECX server. The **Pre-Script** task is run as the first task in the policy job. The **Post-Script** task is the last task to run during the policy job. For more information, see 3.4, "Application consistency and custom scripting" on page 51.

    g. The **Continue operation if Pre-Script fails** option specifies whether the policy job should continue to run if the command or script that is specified in **Pre-Script** option returns a nonzero exit status code. The exit status of the Pre and Post scripts are also logged in the ECX job history.

    After selecting the wanted options, click **Next** to proceed to step 3 in the wizard.

14. In Step 3, Notification, you can optionally specify an SMTP server and email recipients, who will receive notifications about job status for this policy. The STMP server should be configured before this policy is created. See "Configuring an SMTP server in ECX" on page 30 for details about configuring an SMTP server in ECX. This example does not use email notifications, so these options are left blank. Click **Next** to proceed to step 4 in the wizard.

15. In Step 4, Schedule, you can select the schedule settings for the policy. Schedules can be created by pressing the **Schedules** icon. While you are in the policy wizard, you can change to the Schedules window, create a schedule, and come back to the Policies window. Press the **Refresh** icon in the Available Schedules pane to refresh the list. In this example, select the existing schedule named **Daily_9pm** and click **Next** to proceed to the final step in the wizard.

Figure 4-9 shows the Schedule window.

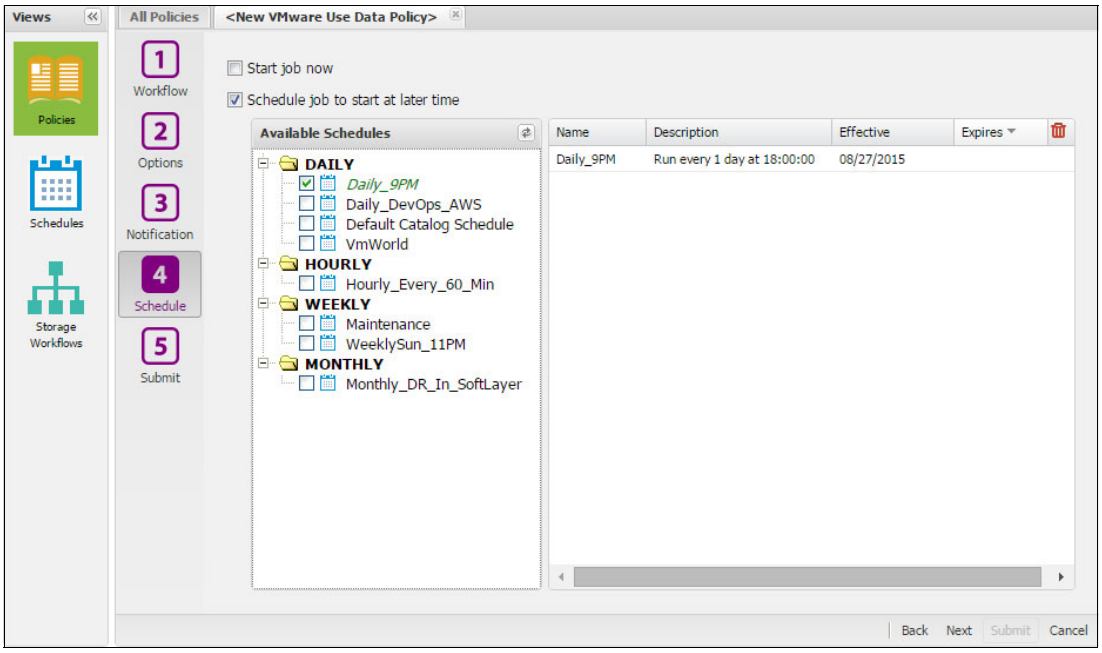


Figure 4-9 VMware Use Data Policy Schedule

16. On the **Submit** window, specify a name for the policy and an optional description. This window also provides a summary of the policy options. Click **Submit** to complete policy creation, after which the policy will run on the specified schedule.

4.4 Automated disaster recovery

One part of the business no one likes to think about is “What happens when a disaster strikes?” and “How do I ensure uptime or access to my data?” Many IT shops try to enable some level of efficiency in their disaster recovery (DR) process, but more typically, DR environments have manual processes. Clients use a number of scripts (that all too often fail) and run books that reference these brittle scripts. Catalogic’s ECX solution allows the IT administrator to create *Use Data* workflows that can start an application copy in a remote location for the sole purpose of automated DR. The concept of location or destination is key to the orchestration of the Use Data workflows to ensure that data can be used off-site during a disaster. Catalogic ECX can orchestrate data copies to any device that has been registered to the ECX software, regardless of location.

IT selects the application or combination of applications that need to be recovered during a disaster, and can start these applications individually or in a group. Commonly, dependencies between applications necessitate that they be recovered in a particular order. As an example, when recovering an email server, the directory server might need to be recovered first. With ECX, applications can be orchestrated to start in any order, allowing the recovery process to automatically accommodate these dependencies. The advantages that ECX offers through the automation and orchestration of these workflows are even greater when considering DR testing. Typically, IT runs DR tests over a weekend. All the appropriate personnel are assembled and must be onsite and available during the tests.

The test typically includes these tasks:

- Find where all the data is
- Ensure all the data is in the correct location
- Ensure all the scripts are prepared to run
- Fence off the ‘test’ environment
- Run the scripts in the proper order
- Test the applications to see that they were recovered properly
- When and if the test is successful, tear the environment down

This process can take all weekend, and in many instances these tests fail and more time is spent troubleshooting scripts that fail than recovering data. With Catalogic ECX and IBM Spectrum Virtualize storage, all of these tasks are automated. No scripts are required. The Copy Data workflow puts the data at the destination where it would be needed during a disaster. Additionally, the Use Data workflow can fetch the data with three different methods: Test, Clone, and Production. These methods bring data and applications up in a protected environment so it cannot overwrite other production data during testing. After testing is complete, the user can have ECX tear down the environment, avoiding unnecessary data sprawl, or can promote the environment to production, as IT would do during a real disaster. The workflow to test the DR environment can be scheduled to happen every evening so that the IT team can come in the morning to find that the DR environment is successfully running. If something failed, Catalogic ECX lets administrators know what failed, and why, through exception-based reporting.

4.4.1 Automated DR example

This example has two data centers: A primary data in New York, and a secondary data center in San Francisco. Assume that you have set up VMware Copy Data policies to capture your applications daily at the New York site and replicate to the San Francisco data center. A disaster has damaged the New York data center and all applications are down. This example demonstrates how to create an ECX VMware Data Policy that can be run during a disaster. This policy can also be used in Test mode and scheduled for daily automated DR testing.

To create an automated DR policy, complete these steps:

1. Log in to ECX web GUI. Open a compatible web browser and go to the following URL:

   `http://<ECX hostname>:8443/portal`

2. Log in with your ECX user ID and password. The default administrator user name is *admin* and password is *password*.
3. Click **Plan** → **Policies** → **New**.

4. In the Use Data column, select **VMware** as shown in Figure 4-10.

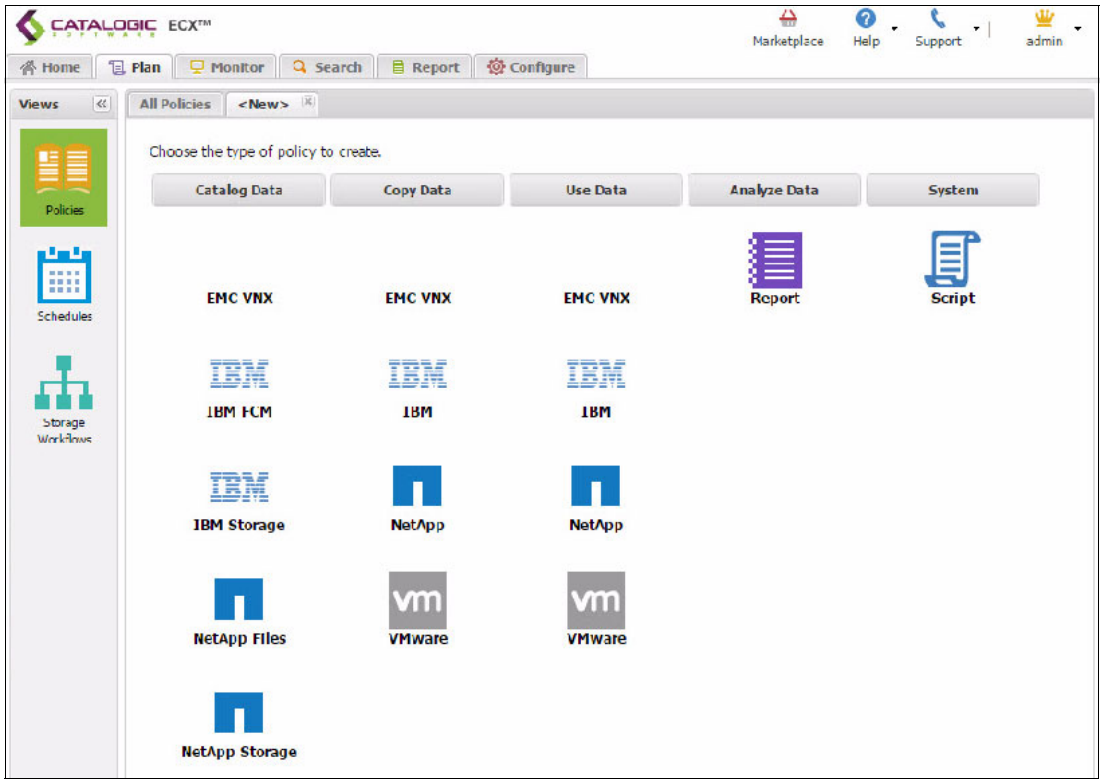


Figure 4-10 ECX Plan Tab

5. The **New VMware Use Data Policy** wizard loads as shown in Figure 4-11. The first step, **Workflow**, requires you to select a workflow template, as described in 4.2, “ECX Use Data Workflow technical overview” on page 56. **Instant Access** instantly mounts a vmdk from a FlashCopy to an existing data store. **Instant Virtualization**, in test mode, creates a virtual machine with the same VM properties as the source VM. However, it creates the VM under a different name and either places it in the same vCenter and ESX host, or places it onto a different vCenter and ESX host, depending on your selection in step 9 on page 70. For this example, **Instant Virtualization** is chosen.

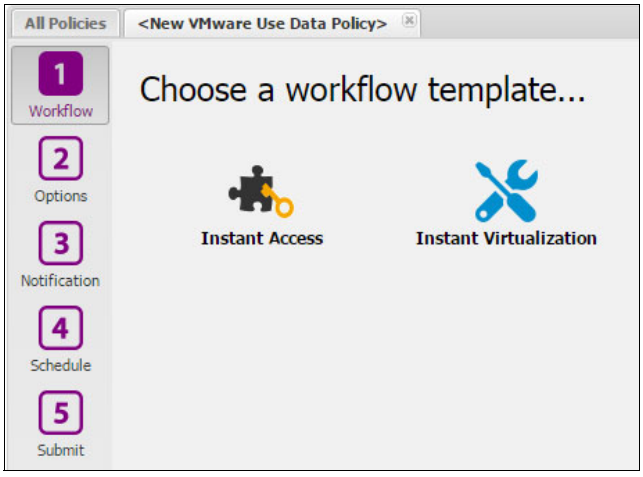


Figure 4-11 VMware Use Data Policy workflow template

6. After the workflow template is chosen, the New VMware Use Data Policy wizard allows selection of VMware VMs that will be part of the Instant Virtualization, as shown in Figure 4-12. Here the wizard shows a VMs and Templates view. A tree browser shows only the VMs that have been captured for ECX VMware Copy Data policies. FlashCopy Manager for VMware FlashCopies is also available if the ECX FCM Catalog policy is configured. The VMware objects are categorized by the ECX site at the top level. In the tree selection, VM folders or individual VMs can be selected. For this example, the site **New York** is expanded, along with the vCenter **PSvCenter1**, and data center **PS_SE-Datacenter**. All applications at the New York Center are selected.

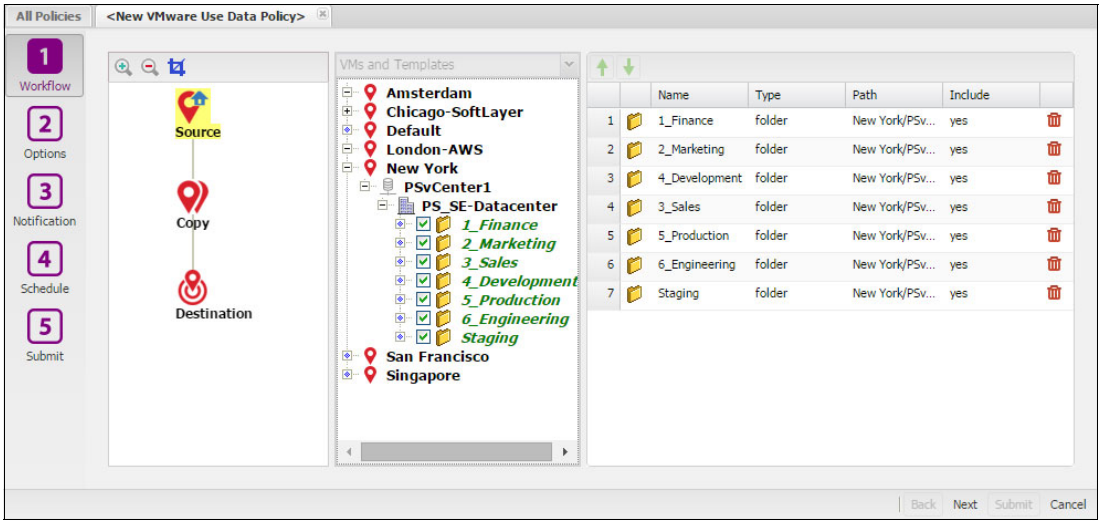


Figure 4-12 VMware Use Data Policy Source selection

7. Click the **Copy** icon in the left pane to display the sites where FlashCopies of the selected virtual machines are available, as shown in Figure 4-13. You must select a site for ECX to use when this policy runs. This example has the sites **New York** and **San Francisco** available with FlashCopies of the virtual machines selected in step 6 on page 68. In this example, because the applications are to be recovered to the secondary data center, the **San Francisco** copy is selected. By default, the FlashCopy with the latest version of each virtual machine is used. Optionally, a previous point in time copy can be selected by using the **Select Version** button.

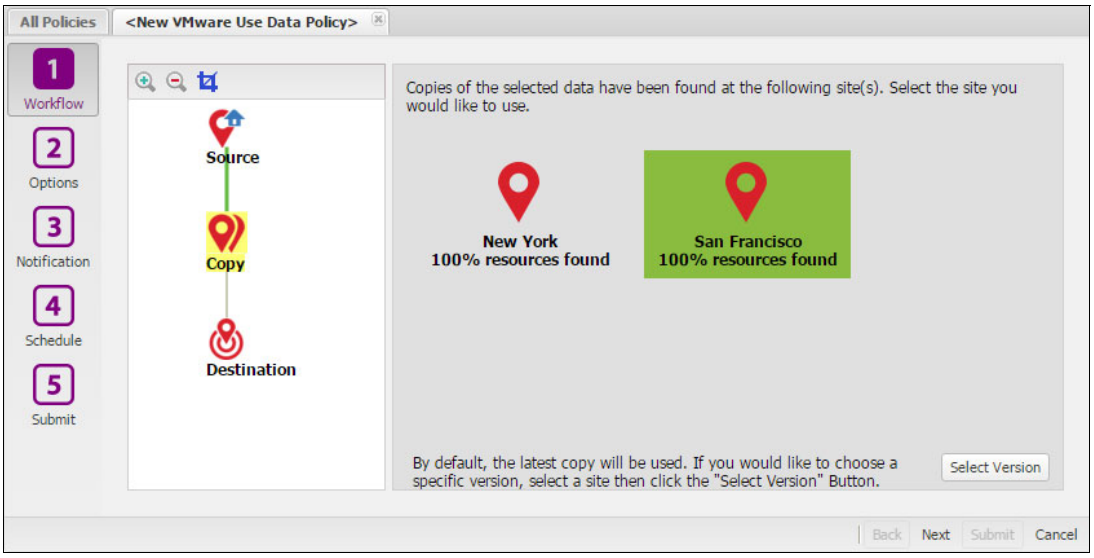


Figure 4-13 VMware Use Data Policy Copy Selection

8. Click the **Destination** icon in the left pane, as shown in Figure 4-14, to select the VMware ESX host or cluster that the test virtual machines will be created on. For the example, clear the **Use original host or cluster** check box to select an ESX host at the secondary data center in San Francisco.

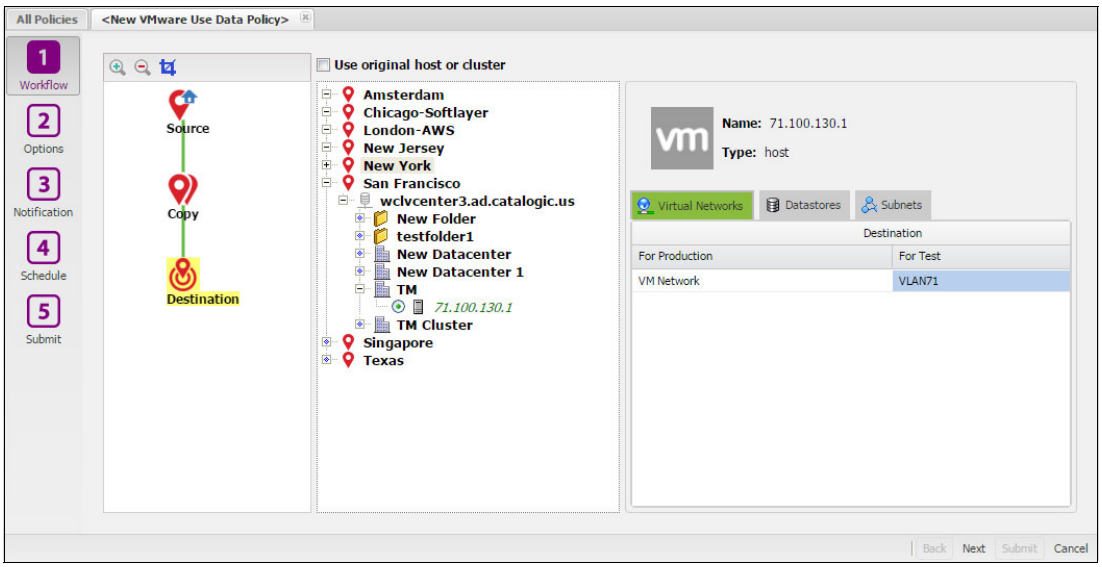


Figure 4-14 VMware Use Data Policy Destination Selection

9. When the **Use original host or cluster** check box is cleared, three tabs are visible in the right pane:
   - Virtual Networks
   - Datastores
   - Subnets

   The **Virtual Networks**, as shown in Figure 4-14 on page 69, enables users to select the virtual networks for both **Production** and **Test** modes. In this example, the production network is used because **Production** mode will be selected later for the policy.

   The Datastores tab, as shown in Figure 4-15, enables you to select destination data stores for production and clone modes. When the Instant Virtualization job starts in (or is changed to) production mode or clone mode, the virtual machines run from a FlashCopy of a FlashCopy. While the VM is running, the Instant Virtualization job uses vMotion to migrate the virtual machine storage from the FlashCopy to this selected data store. This process allows the virtual machines to start immediately without any data movement. While the vMotion is active, the virtual machines can remain powered on and functioning.

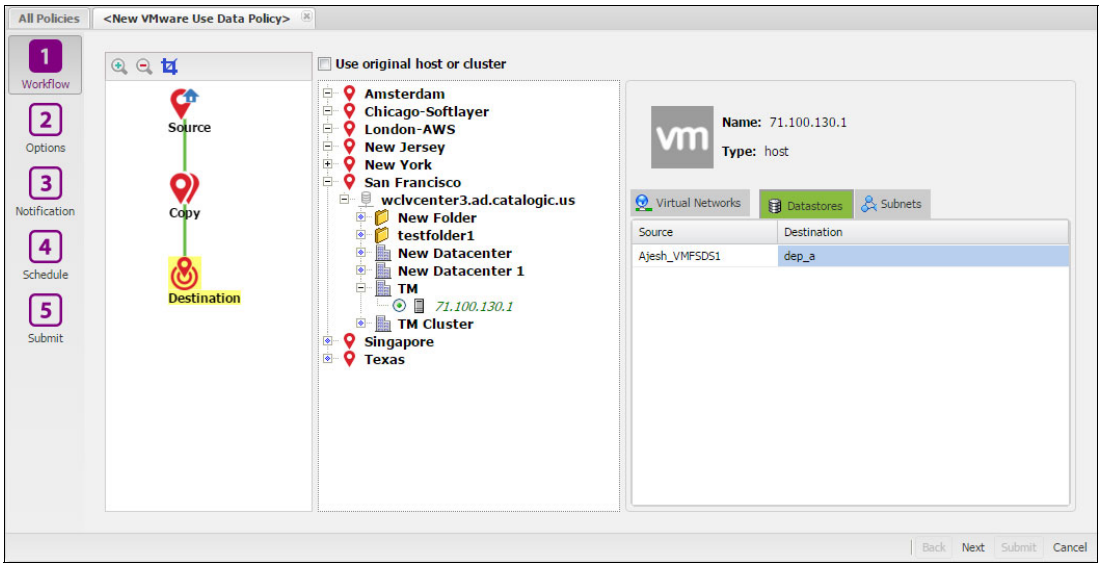


Figure 4-15 VMware Use Data Policy Destination Datastore selection

The Subnets tab (Figure 4-16) allows you to remap IP addresses and operating system network settings within the VM. During the Instant Virtualization job, ECX agentlessly accesses the network properties of the operating system within each virtual machine. This process can be used when performing Instant Virtualization to different sites that can have different network topology. In this example, any virtual machines in the primary data center subnet **192.154.0.0** are remapped to the secondary data center **192.55.0.0** subnet. Remapping by subnet is less tedious if many virtual machines are part of a particular VMware Use Data Policy. In addition to the IP addresses, the Gateway and DNS server are remapped.

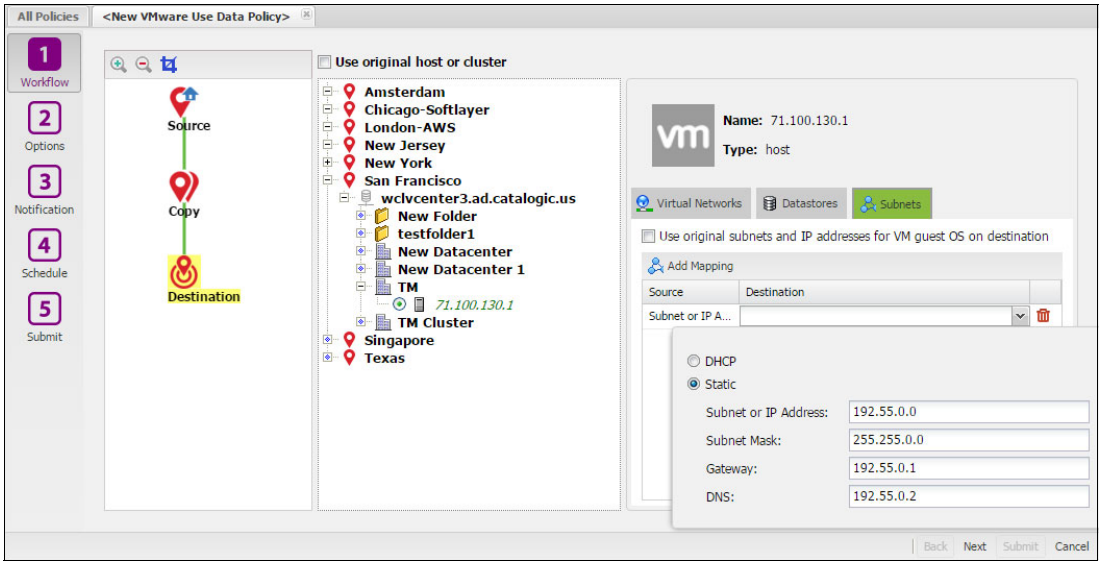


Figure 4-16 VMware Use Data Policy Subnet remapping

After you complete the selections, click **Next** to go to Step 2, **Options** (Figure 4-17).

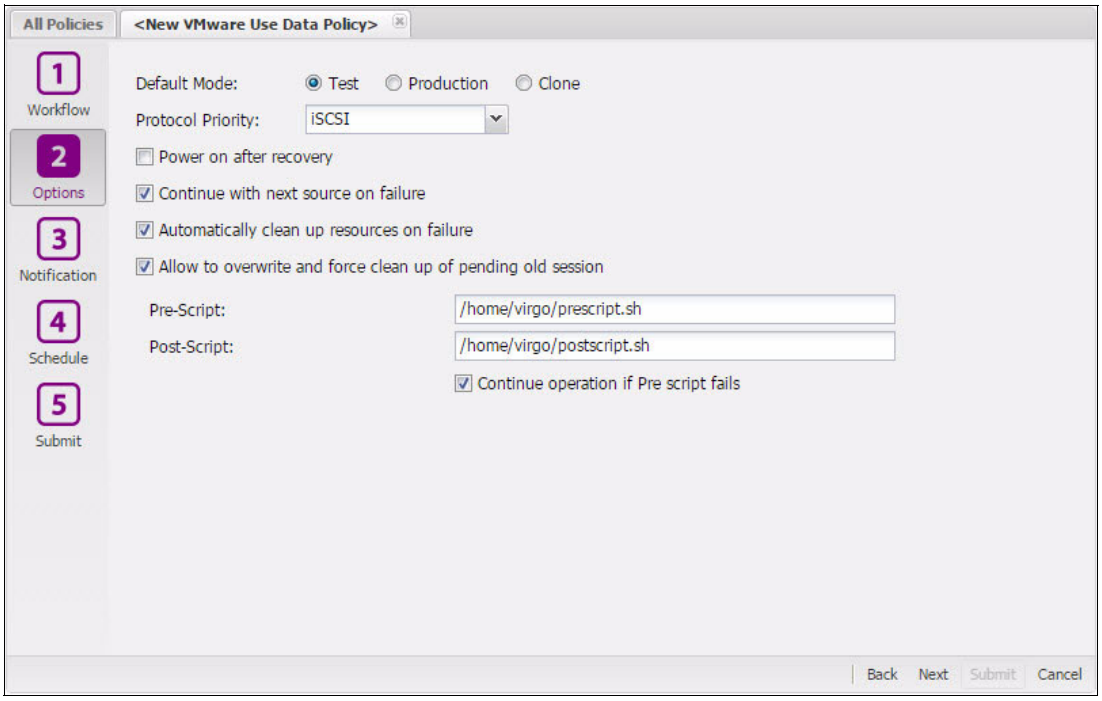


Figure 4-17 VMware Use Data Policy options

10. Select among these options for your Use policy:

    a. The first option, **Default Mode**, specifies the default mode of the policy when it runs. There are three modes: **Test**, **Production**, and **Clone**. For this example, **Production** mode is selected. In this mode, the Instant virtualization job first deletes the original source VM and a new VM with the same name and configuration is created. The virtual machine storage is on the FlashCopy, and vMotion is started to migrate the virtual machine hard disks to a production data store.

    b. The **Protocol Priority** option specifies the protocol to take priority in the policy. Available protocols include iSCSI and Fibre Channel.

    c. The **Power on after recovery** option specifies whether the virtual machines are started during the Instant Virtualization job. If not selected, the virtual machines can be manually powered on by using the VMware vSphere interface.

    d. The **Automatically clean up resources on failure** option enables the automatic clean up of allocated resources as part of a restore if the virtual machine recovery fails. This option is selected by default.

    e. The **Allow to overwrite and force clean up of pending old sessions** option allows the Instant Virtualization job to clean up resources held by currently pending sessions. This option is selected by default.

    f. The **Pre-Script** and **Post-Script** options allow you to specify scripts that will run during the policy job. Both the **Pre-Script** and **Post-Script** option can take Linux shell commands or the absolute path to a Linux shell script on the ECX server. The **Pre-Script** task is run as the first task in the policy job. The **Post-Script** task is the last task to run during the policy job. See 3.4, “Application consistency and custom scripting” on page 51 for more details.

    g. The **Continue operation if Pre-Script fails** option specifies whether the policy job should continue to run if the command or script that is specified in **Pre-Script** option returns a nonzero exit status code. The exit status of the Pre and Post scripts is also logged in the ECX job history.

    After selecting the options that you want, click **Next** to proceed to step 3 in the wizard.

11. Step 3, **Notification**, loads. Here you can optionally specify an SMTP server and email recipients who will receive notifications about job status for this policy. This example skips email notifications. Click **Next** to proceed to step 4 in the wizard.

12. Step 4, **Schedule,** loads, as shown in Figure 4-9 on page 65. Here you can select the schedule settings for the policy. In this example, both the **Start job now** and **Schedule job to start at later time** check boxes should be cleared, as shown in Figure 4-18. When a disaster occurs, this policy will be ready to be run from the Monitor tab.

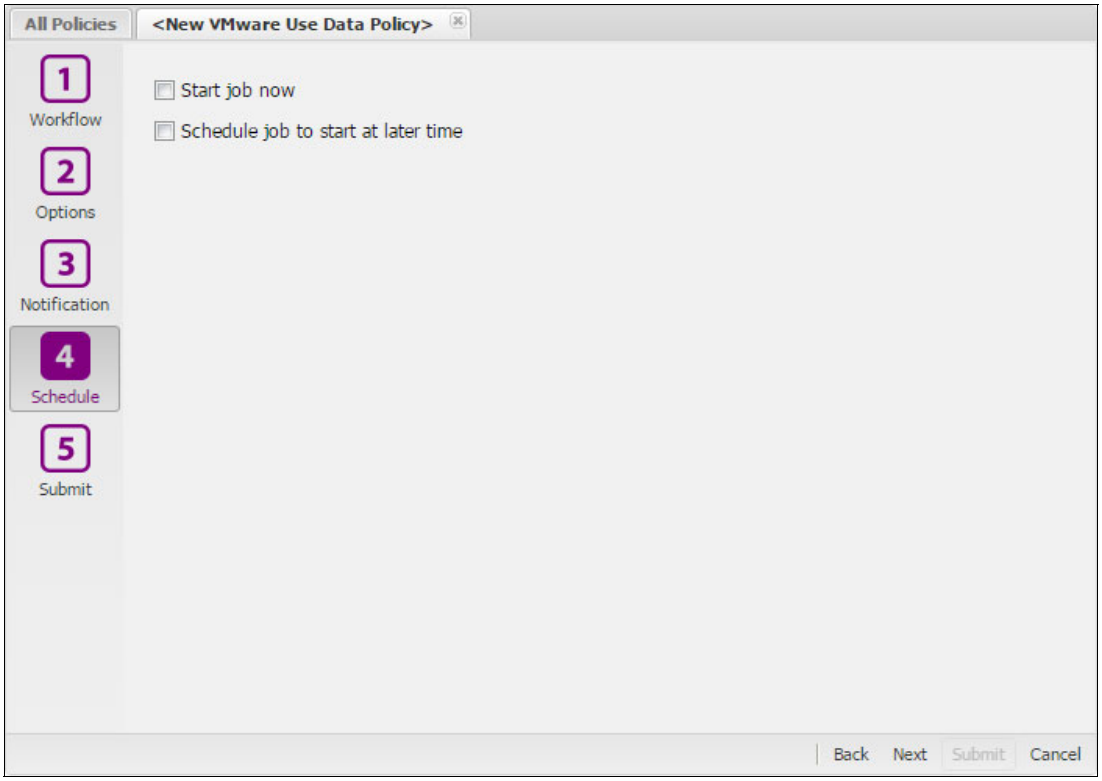


Figure 4-18 VMware Use Data Policy Schedule

13. On the **Submit** window, specify a name for the policy. Optionally, a description can be specified. This window also provides a summary of the policy options. After **Submit** is clicked, the policy creation is complete.

During a disaster, a VMware Use Data Policy is now ready to run to automate the disaster recovery of your application to the secondary data center. To run the policy, go to the Monitor tab, and right-click on the policy and select **Start.** A prompt is displayed for mode. Specify **Test** for a DR test, or **Production** to perform automated disaster recovery.

4.4.2 Business analytics

The latest trend in driving competitive advantage is the use of analytics within the business. Gaining insight into the business through its data can help businesses make better decisions on where to focus, how to save money, and how to develop a competitive business strategy. Traditional analytic operations use an extract, transform, and load (ETL) process. These three database functions are combined into one tool to pull data out of one database and place it into another database.

The purpose behind this process is to put data in a common location (database), allowing analysis software to run against it. Because this process can be time-consuming, it only grabs the parts of each data set that the analytics team believes that they need. Currently, this data is typically loaded manually or by using scripting, which is time consuming, resource intensive, and error prone.

After the data is brought up for the analytics team, they run their analytic applications against the data to gain the valuable insights they need to drive the business. Because the process of getting the data set ready for the analytics team is difficult, this data cannot be refreshed every day. As a result, the accuracy of the data that is extracted from the analysis can become stale and less relevant. Not analyzing the latest data means that the business might not be making decisions as fast as they could to stay competitive.

By taking advantage of Catalogic's ECX and IBM Spectrum Virtualize Storage, IT can mount any data sets quickly for analysis. This process builds in a great deal of flexibility for the analytics team. Instantly mountable data access provides these benefits:

- No more scripts that fail, nor the timely process of troubleshooting these failed scripts
- Automated process that ensures data is where it needs to be, when it needs to be there
- Ability to fetch entire data sets so data scientists can make decisions dynamically about new analyses
- Daily updates to the latest data sets to analyze, which yields higher analytic accuracy
- Higher accuracy drives a more competitive business
- Clean up process in the ECX policy to ensure that analytics do not generate data sprawl
- ECX replicates data copies to multiple locations, enabling analysis in different geographic locations

4.4.3 Infrastructure as code

ECX abstracts hypervisor and storage infrastructure and allows IT to provide Infrastructure as a Service or Infrastructure as Code in the realm of copy data management to your developers.

Traditionally, developers do not have access to hypervisor and storage, and must manually request the IT team to perform operations such as provisioning new virtual machines or mounting FlashCopies. The use cases that are covered in this chapter demonstrated using the ECX web GUI. In today's DevOps world, automating by using code is key. ECX provides a rich REST API to bring IT infrastructure into the DevOps workflow. All functions that are available in the GUI are also available by using the ECX REST API. The GUI actually uses the ECX REST API to perform all operations, ensuring that the REST API covers all the functions of ECX.

5

Self-Service for application teams with Catalogic ECX

This chapter describes Self-Service capabilities in Catalogic ECX. It provides an overview of what self-service means for an application owner and describes the different features that enable these capabilities. It also provides a step-by-step implementation of providing role-based access control (RBAC) for administrators and users. In addition, this chapter illustrates the roles of an application owner and a storage administrator in Copy and Use workflows.

This chapter includes the following sections:

- Overview of self-service capabilities in ECX
- Role-based access control (RBAC)
- Configuring custom roles
- Self-service in ECX

5.1 Overview of self-service capabilities in ECX

An IT organization is made up different teams, each with its own roles and responsibilities. A common division is that *application teams* create and manage applications and an *infrastructure team* is responsible for supporting the application team by providing them the required hardware and software resources. In ECX, you can create distinct granular roles for application owners and the infrastructure owners. The infrastructure owners can create templates that enforce appropriate service levels, and the application owners can apply these templates to the resources that they manage. In this way, both groups achieve their organizational goals.

5.2 Role-based access control (RBAC)

RBAC in ECX allows you to set the resources and permissions available to ECX accounts. Using RBAC, you can tailor ECX for individual users, giving them access to the features and providers (or resources) that they need. After providers are associated with a site, they can be added to a resource pool along with high-level ECX features such as policies, reports, and windows. Roles are then configured to define the actions that can be performed by the user of the account associated with the resource pool. These parameters are then associated with one or more user accounts, which can be native to ECX or imported as part of an LDAP group.

Note: Users that register providers, such as storage devices, or add resources to ECX, such as policies or customized reports, will have full access to interact with those providers or resources regardless of RBAC restrictions. For example, if users' permissions allow them to register IBM providers, they will also be able to view, edit, and unregister the IBM providers that they registered, even if the necessary permissions are not assigned to them through RBAC.

5.2.1 Resource pools

A resource pool defines the resources that are made available to an account. Every provider added to ECX, such as storage devices and LDAP servers, can be included in a resource pool, along with individual ECX functions and screens. This configuration allows you to fine-tune the experience of a user. For example, a resource pool might include only storage devices that are associated with a single vendor, with access to only the ECX search and reporting functions. When the resource pool is associated with a role and an account, the account user only sees the windows that are associated with search and reporting, and only has access to the storage devices defined in the resource pool.

5.2.2 Roles

Roles define the actions that can be performed on the resources defined in a resource pool. A resource pool defines the providers that are made available to an account, such as storage devices, and resources, such as ECX functions and screens. A role sets the permissions to interact with the resources defined in the resource pool. For example, if a resource pool is created that includes ECX Copy Data and Use Data policies, the role determines how a user can interact with the policies.

Permissions can be set to allow a user to create, view, and run the Copy Data and Use Data policies that are defined in a resource pool, but not delete them. Similarly, permissions can be set to create administrator accounts, allowing a user to create and edit other accounts, set up sites and resources, and interact with all of the available ECX features.

5.2.3 Accounts

An account associates a resource pool with a role. To enable a user to log on to ECX and use its functions, add the user to ECX as a native user or as part of an imported group of LDAP users. Then assign resource pools and roles to the user account. The account then has access to the resources and features defined in the resource pool as well as the permissions to interact with the resources and features defined in the role.

5.3 Configuring custom roles

This section covers how to use the RBAC described in 5.2, “Role-based access control (RBAC)” on page 76 to create a Storage administrator role and an Application administrator role.

5.3.1 Storage administrator

A storage administrator should have access to all the IBM storage resources, and be able to create and modify storage workflows and IBM Copy and Use policies.

Note: The following example applies to an IBM storage administrator only.

Resource pool

To configure a resource pool for a storage administrator, complete these steps:

1. In the Configure tab in the ECX GUI, click **Resource Pools** in the **Views** pane. The All Resource Pools pane opens, as shown in Figure 5-1.

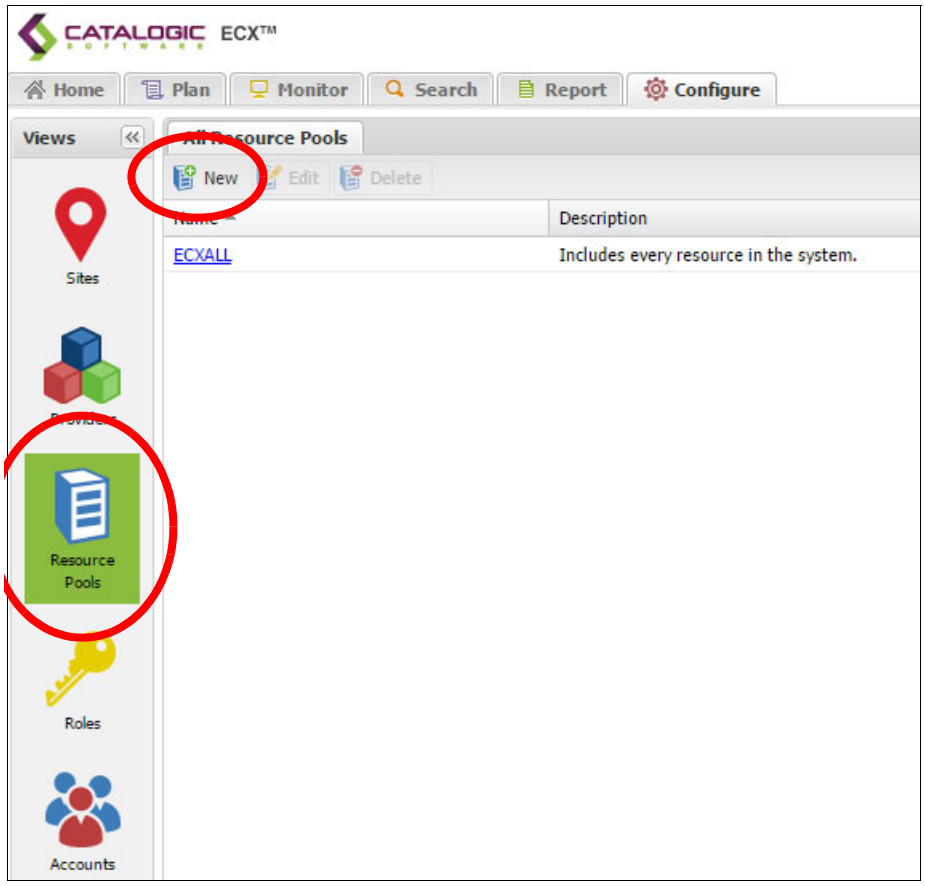


Figure 5-1 Configuring Resource Pools in the ECX management portal

2. In the All Resource Pools pane, click **New**.

3. The **New Resource Pool** wizard opens, as shown in Figure 5-2. In step 1 of the wizard, Providers**,** select one or more providers to add to the resource pool from the list of available sites and providers. Click the **+** icon next to a site to display all the resources and select the IBM resources. Repeat this process for all the sites.

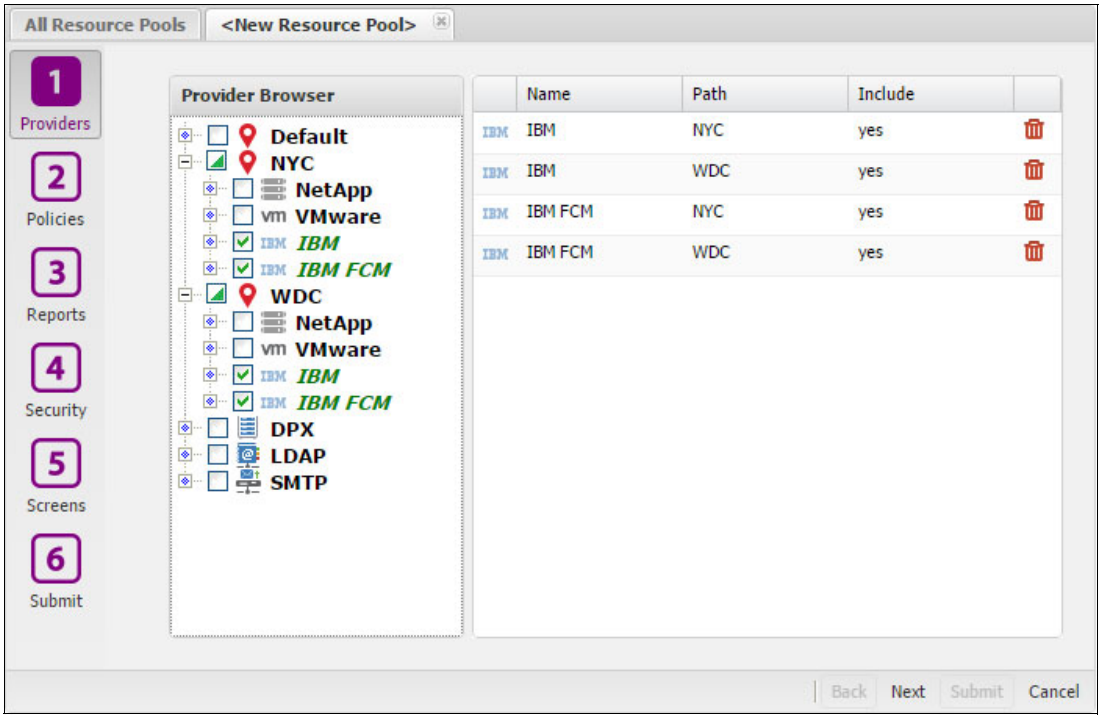


Figure 5-2 IBM Providers selection in the ECX management portal

Note: You can select individual storage controllers within a site, as well.

4. Click **Next** to proceed to step 2 in the wizard, Policies.

5. Expand **All Policies** and select the following policies, as shown in Figure 5-3:
   - IBM Catalog Data
   - IBM Copy Data
   - IBM FCM Storage Catalog
   - IBM Use Data
   - Report

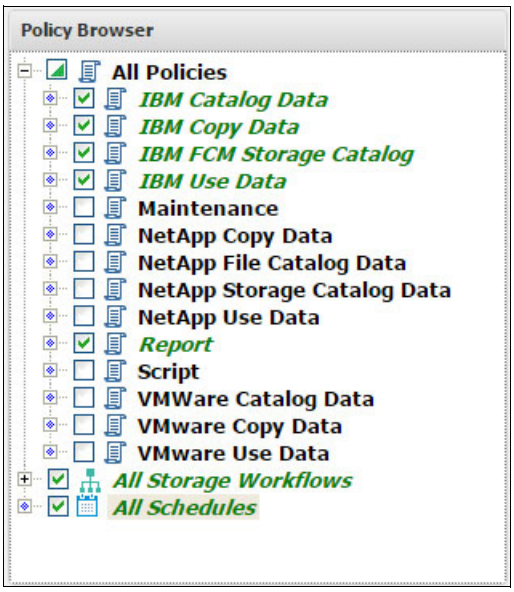


Figure 5-3 Policy Selection in the ECX management portal

6. Select **All Storage Workflows** and **All Schedules** to include these items in the resource pool .
7. Click **Next** to proceed to step 3 in the wizard, Reports.

8. Expand the **Protection Compliance** and **Storage Utilization** items and select the following reports, as shown in Figure 5-4:
   - IBM RPO Compliance
   - IBM Storage Pools
   - IBM Storage volumes

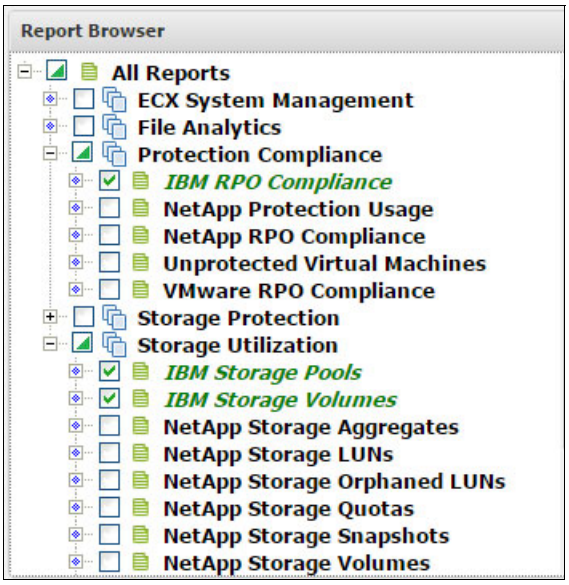


Figure 5-4 Report Selection in the ECX management portal

9. Click **Next** to proceed to step 4 in the wizard, Security.
10. Select a security option that is configurable by accounts associated with this resource pool. The following options are available as shown in Figure 5-5:
    - All Roles
    - All Accounts
    - All Resource Pools

    For this example, no selection is made on this tab.

Figure 5-5 Security Browser selection in the ECX management portal

Note: If **All Resource Pools** is selected in this step, users who are associated with this resource pool can create, view, edit, and delete Resource Pools, if paired with the necessary **resourcepool** permission, set on the Roles window.

11. Click **Next** to proceed to step 5 in the wizard, Screens.

12. Select the following ECX screens to include in the resource pool, as shown in Figure 5-6:
    - Home
    - Plan
    - Monitor
    - Report
    - Configure
    - Logs

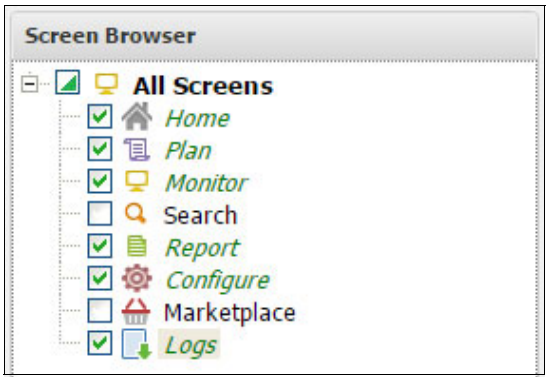


Figure 5-6 Screen selection in the ECX management portal

13. Click **Next** to proceed to step 6 in the wizard, Submit.
14. Enter a name for your resource pool and a meaningful description, as shown in Figure 5-7, and click **Submit** to save the resource pool. The resource pool creation is complete, and it can be applied to new and existing ECX user accounts.

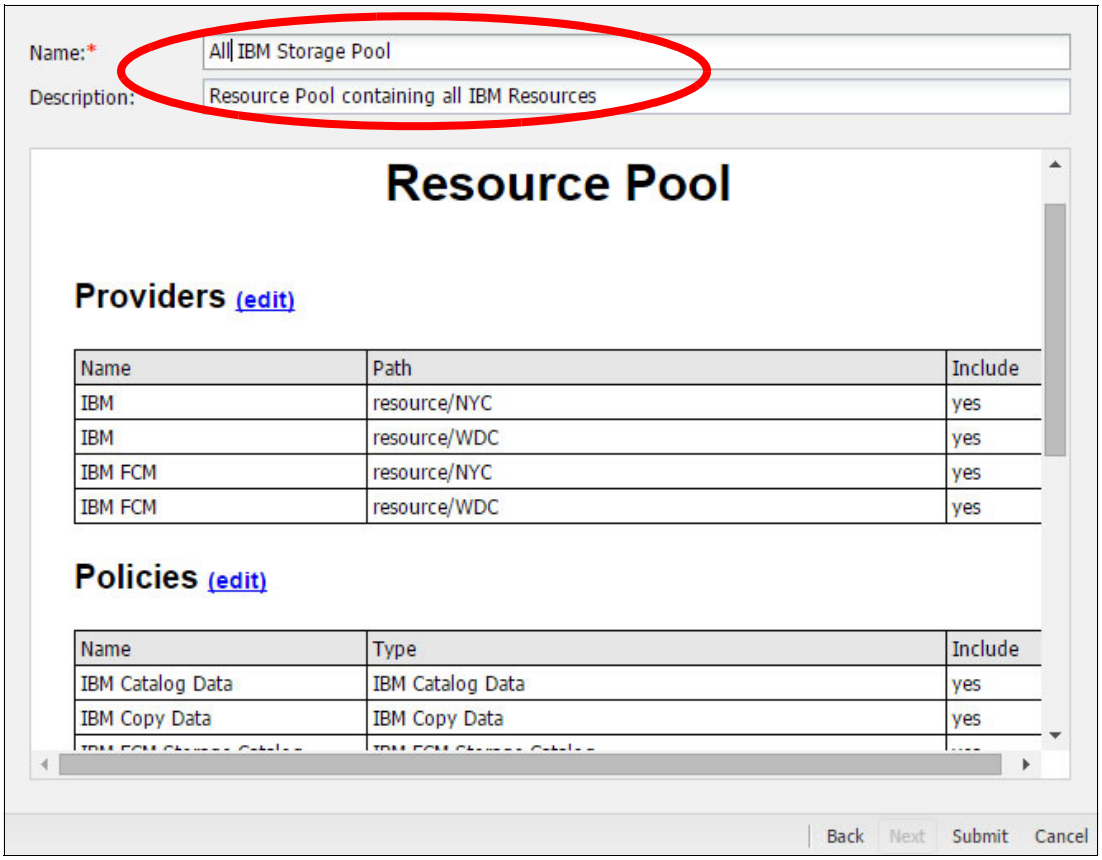


Figure 5-7 Final Review page to enter name and description

Role

Now that the appropriate resource pool has been created and configured, complete these steps to create the corresponding role for the storage administrator:

1. Click the **Configure** tab.
2. On the Views pane, select **Roles**. The All Roles pane opens as shown in Figure 5-8.

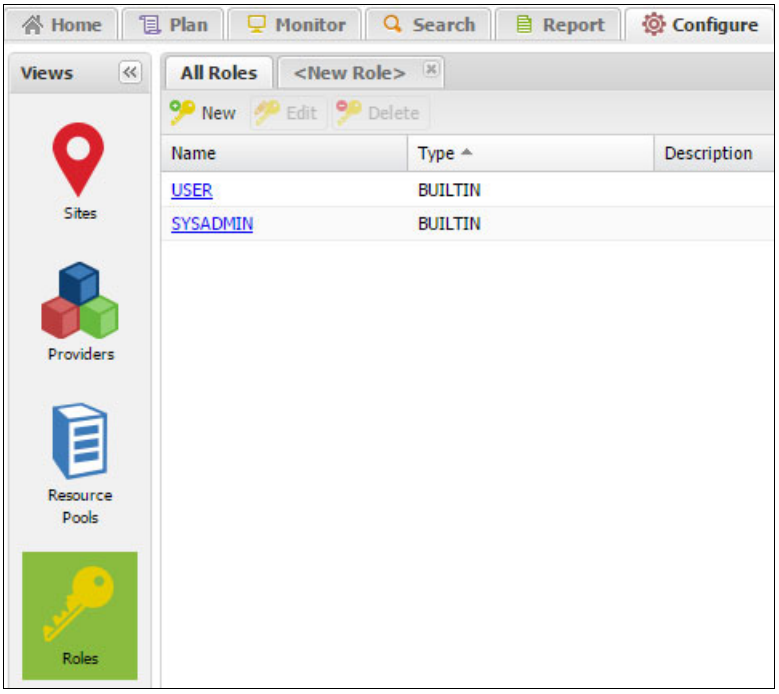


Figure 5-8 Roles Pane showing default User and Sysadmin role

3. In the All Roles pane, click **New**. The New Role window opens, as shown in Figure 5-9. Enter a name and a description for the role, and select the following ECX features to add to the role:
   - ibmfcm
   - ibmsvc
   - report
   - policy
   - storageworkflow
   - screen

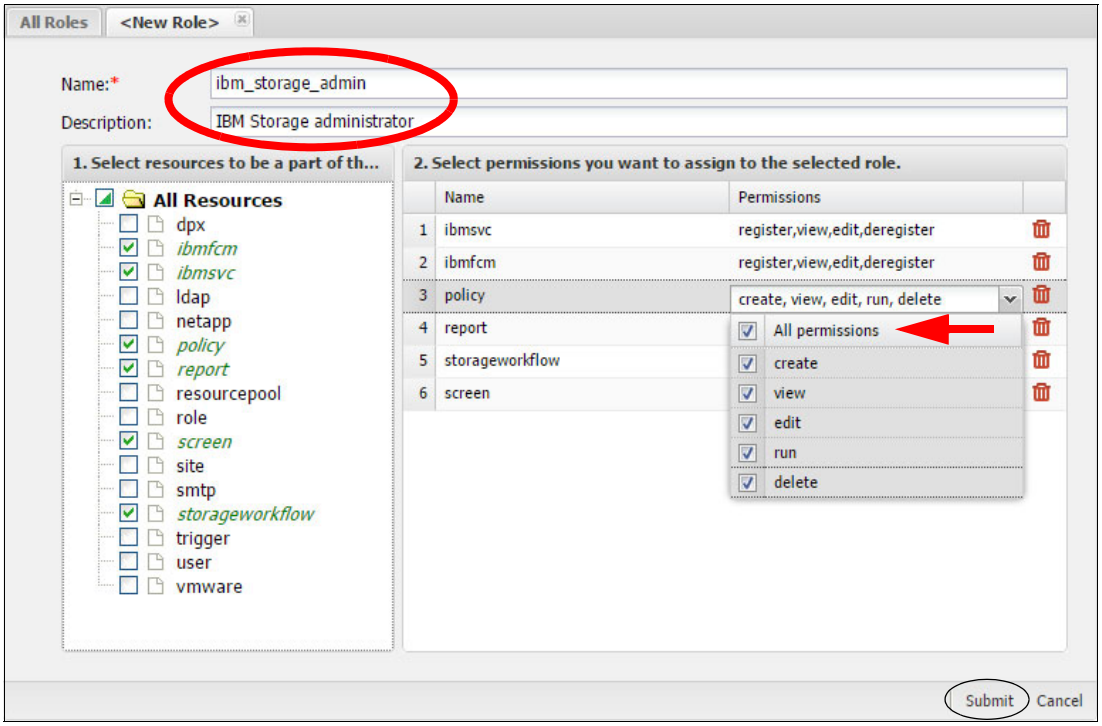


Figure 5-9 Setting up permissions for each role

When each feature is added to the role, it is displayed in the Permissions pane on the right of the console.

4. Click the **Permissions** field next to each of the features selected in step 3, and for each one, select the **All Permissions** check box as shown in Figure 5-9.
5. Click **Submit** to save the new role.

Account

You can add an ECX user account by creating a native user or importing users by using an LDAP user group. This section provides an example of each method.

Note: Importing users using an LDAP user group requires that an LDAP provider has been registered. For more information about registering and LDAP server in ECX, see “Configuring an LDAP server in ECX” on page 28.

Creating a native ECX user

Complete these steps to add a native user:

1. Click the Configure tab.
2. In the Views pane, select **Accounts**. The Accounts pane opens, as shown in Figure 5-10.

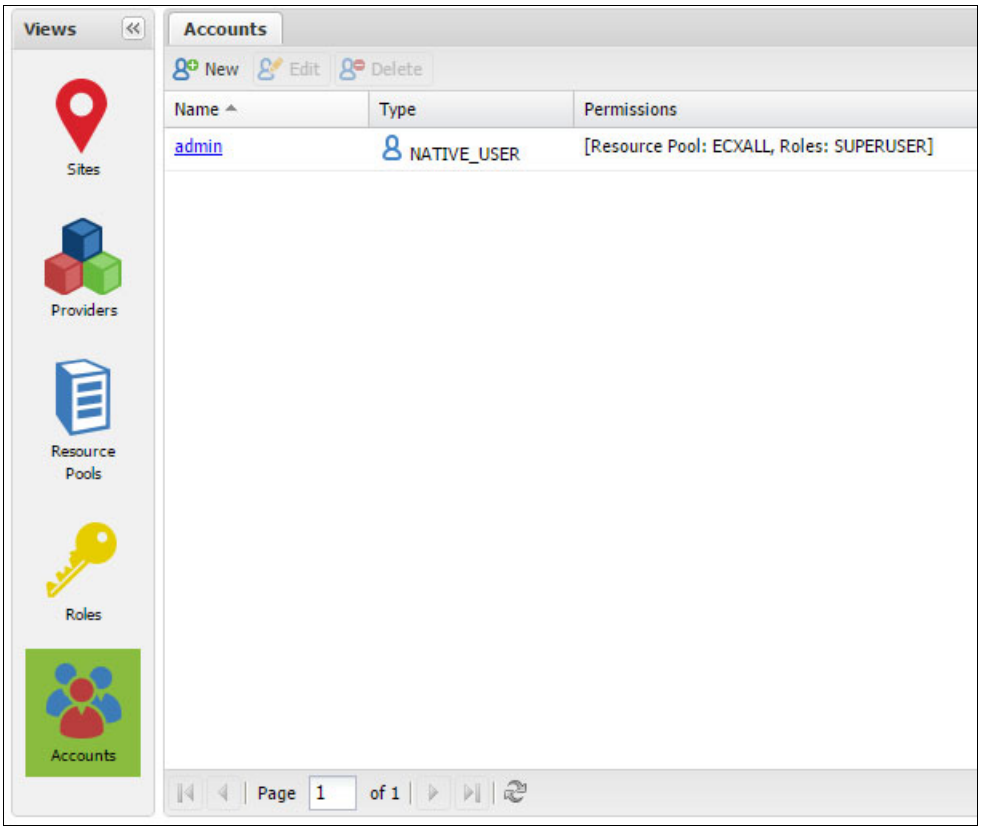


Figure 5-10 Configuring New Accounts in the ECX management portal

3. Click **New**.
4. In the New Account pane, click **Create Native User** as shown in Figure 5-11**.** The New Account window opens.

Figure 5-11 Create Native User

5. Enter a user name and a password for the account. This example uses the user name *ibm_admin*.

6. Select the resource pool that you created in “Resource pool” on page 78 to add to the account as shown in Figure 5-12.

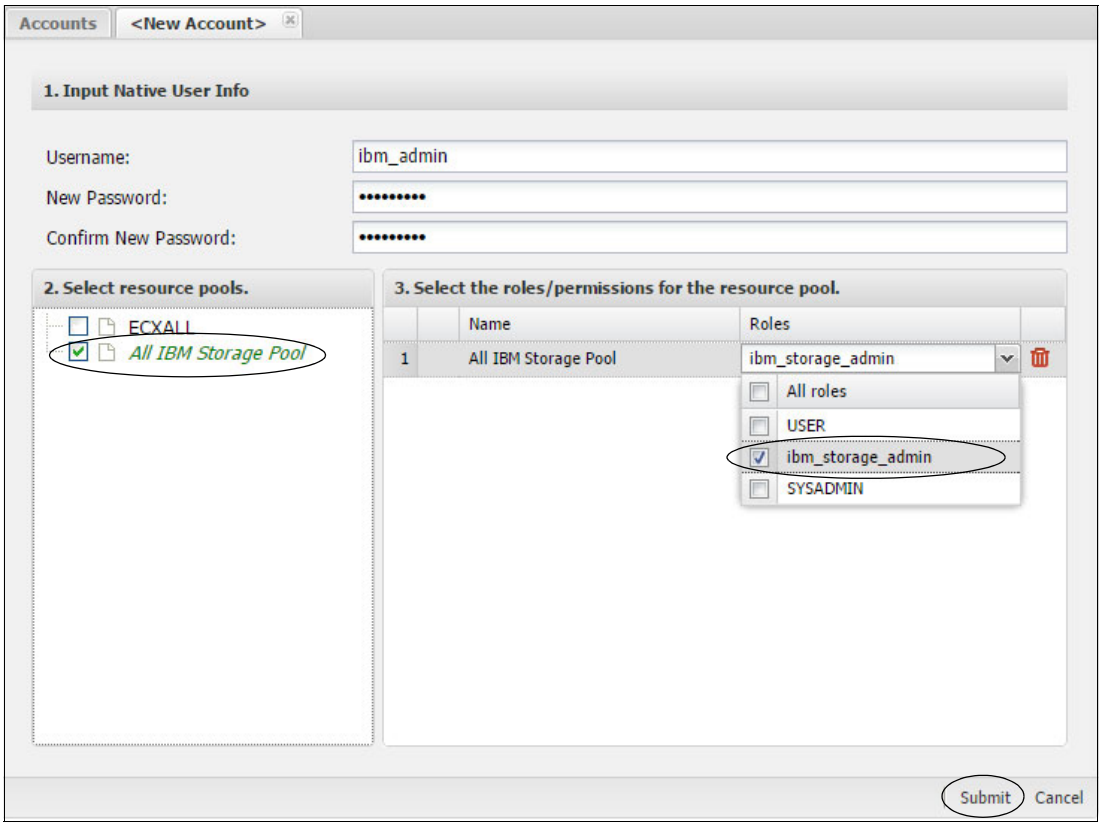


Figure 5-12 New account with Resource pool and Role assigned

7. Select the role that you created in “Role” on page 83 to associate with the resource pool as shown in Figure 5-12.
8. Click **Submit** to save the new user account. The account then appears on the Accounts pane.

Importing user accounts by using an LDAP group

Complete these steps to import an LDAP user:

1. Click the Configure tab.
2. On the Views pane, select **Accounts.**
3. Click **New**.
4. In the New Account pane, as shown in Figure 5-11 on page 85, click **Import LDAP Group**.
5. The **New Account** dialog opens and a list of available LDAP groups displays, as shown in Figure 5-13 on page 87. Select one or more LDAP groups to assign to the selected account.

> **Note:** One more LDAP providers must be registered in **Configure** → **Provider** → **LDAP** before importing a user.

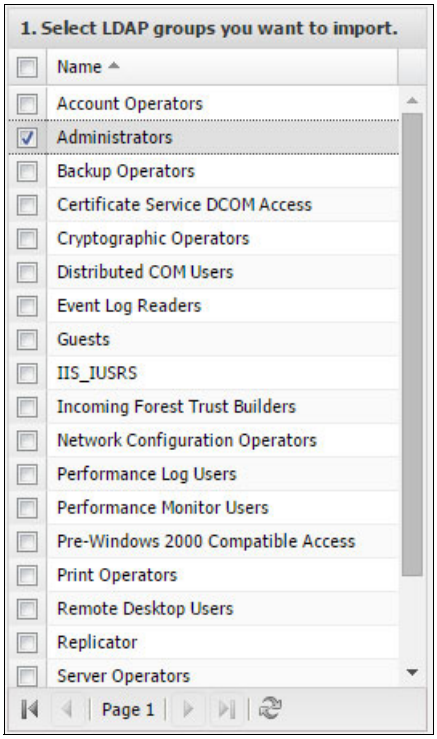


Figure 5-13 LDAP Groups to Import

6. Select the resource pool that you created in “Resource pool” on page 78 to add to the account.
7. Select the role that you created in “Role” on page 83 to associate with the resource pool.
8. Click **Submit**. The account appears on the Accounts pane.

Storage administrator login

After the account is created with appropriate resources and roles, the storage administrator can log in to the ECX management portal with the appropriate credentials, as shown in Figure 5-14. The administrator will be able to access only the screens and resources that were granted permissions for.

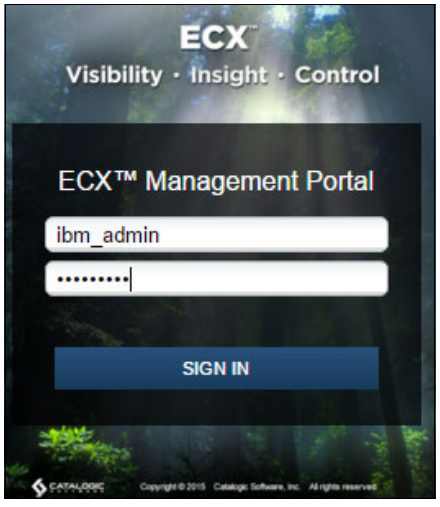


Figure 5-14 Storage admin login to the portal

Figure 5-15 shows that only the screens that were assigned in Figure 5-6 on page 82 are visible and only the policies assigned in the resource pool created in Figure 5-3 on page 80 are accessible.

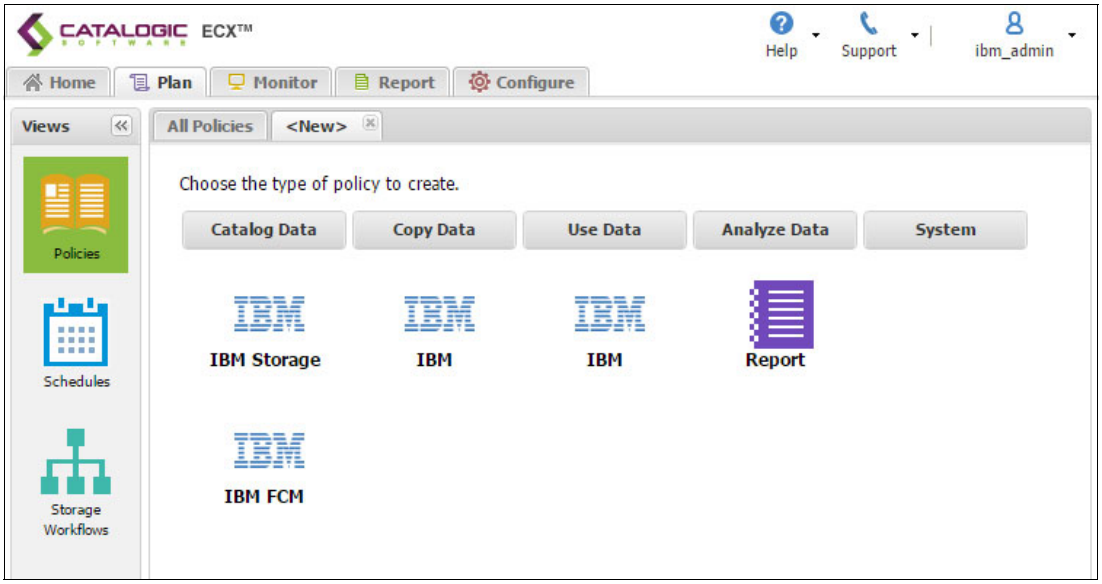


Figure 5-15 ECX Plan window when logged in as IBM storage administrator

5.3.2 Application administrator

Application administrator should have access to all the virtual machine resources that the administrator maintains, and can create and modify VMware Copy and Use policies.

Resource pools

To create resource pools, complete these steps:

1. Click the Configure tab.
2. On the Views pane, select **Resource Pools**. The **All Resource Pools** pane opens as shown in Figure 5-16.

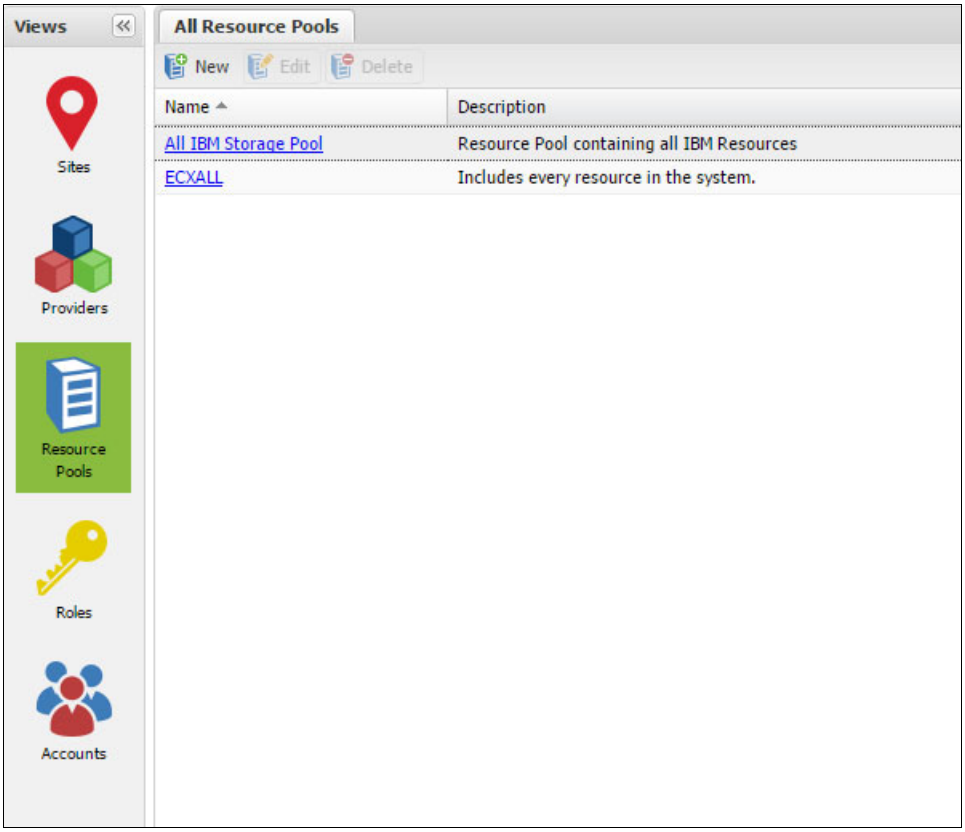


Figure 5-16 Resource Pools in the ECX management portal

3. In the All Resource Pools pane, click **New**. The New Resource Pool wizard opens as shown in Figure 5-17.

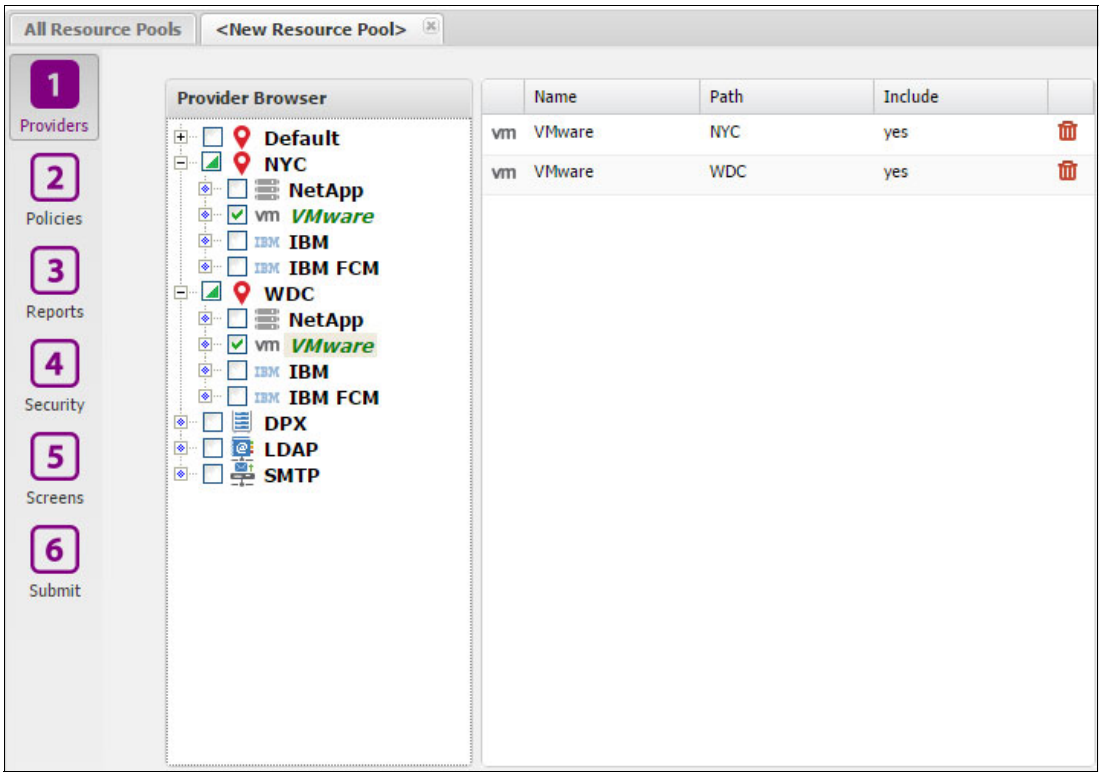


Figure 5-17 VMware Provider selection in the ECX management portal

4. In step 1, Providers**,** select one or more providers from the list of available sites and providers to add to the resource pool. Click the **+** icon next to a site to display all the resources and select the **VMware** resource. Repeat this procedure for all the sites as shown in Figure 5-17, then click **Next** to proceed to the next step in the wizard.

> **Tip:** You can select individual virtual machines within a vCenter instance.

5. In step 2, Policies, expand **All Policies** and select the following policies:
   - VMware Catalog Data
   - VMware Copy Data
   - VMware Use Data
   - Report

6. Select **All Storage Workflows** and **All Schedules** to include in the resource pool, as shown in Figure 5-18, then click **Next** to proceed to the next step in the wizard.

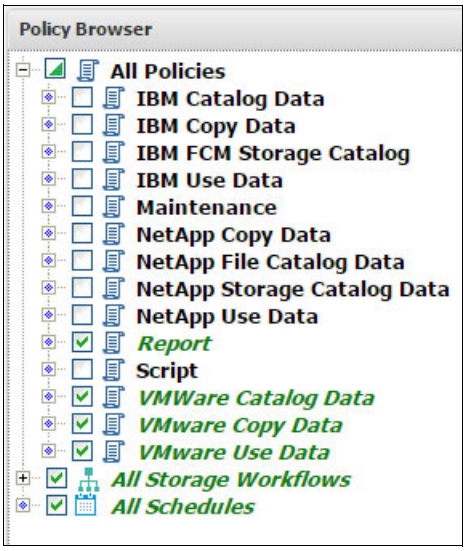


Figure 5-18 Policy and Workflow selection

Tip: Select individual storage workflows for users who are not administrators.

7. In step 3, Reports, select all the following IBM based reports as shown in Figure 5-19 on page 92:
   - In Protection Compliance, select the following reports:
     - VMware RPO Compliance
     - Unprotected Virtual Machines
   - In Storage Utilization, select the following reports
     - VMware Datastores
     - VMware LUNs
     - VMware Orphaned LUNs
     - VMware VM Snapshot Sprawl
     - VMware VM Sprawl

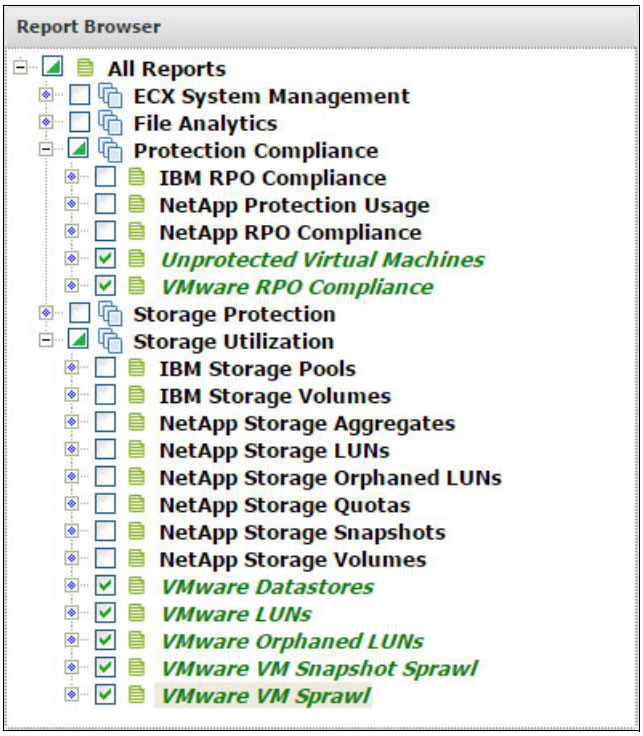


Figure 5-19 Report Browser in the ECX management portal

8. Click **Next** to proceed to the next step in the wizard.
9. In step 4, Security, click **Next**. No selections need to be made on this tab.
10. In step 5, Screens, select the following ECX screens to include in the resource pool, as shown in Figure 5-20, then click **Next** to proceed to the final step in the wizard:
    - Home
    - Plan
    - Monitor
    - Report
    - Configure
    - Marketplace
    - Logs

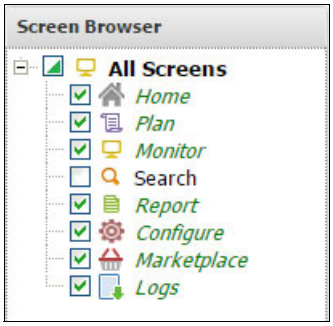


Figure 5-20 Screen Browser in the ECX management portal

11. In the final step of the wizard, Submit, enter a name for your resource pool and a meaningful description and click **Submit** to save the new resource pool. The resource pool appears on the All Resources pane and can be applied to new and existing accounts.

Roles

To create roles, complete these steps:

1. Click the Configure tab.
2. On the Views pane, select **Roles**.
3. In the All Roles pane, click **New**. The New Role window opens, as shown in Figure 5-21.

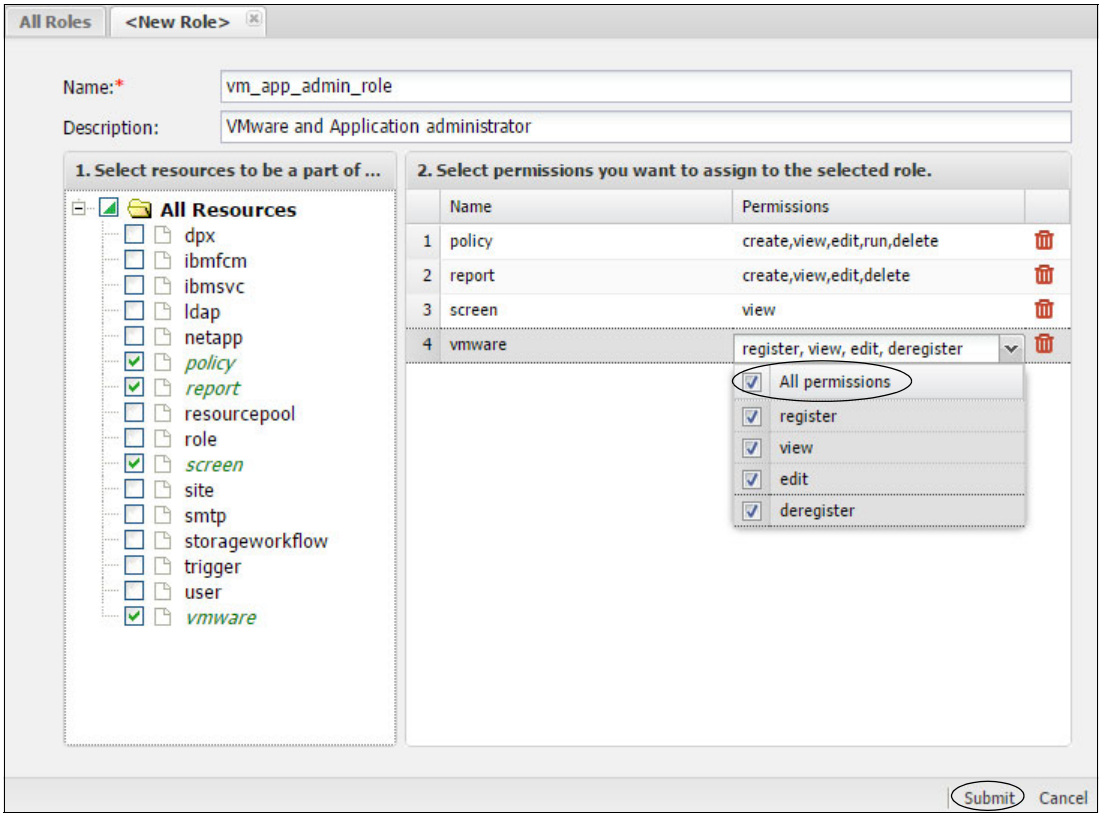


Figure 5-21 App admin role with features and permissions

4. Enter a name and description for the role. This example uses the name **vm_app_admin_role** and the description **VMware and Application admininistrator**.
5. Select the following ECX features to add to the role:
   - Policy
   - Report
   - Screen
   - VMware

   When a role is selected, it displays in the Permissions pane on the right of the console.
6. Click the permissions cell next to each of the roles selected in step 5 and select the **All Permissions** check box as shown in Figure 5-21.
7. Click **Submit** to save the new role.

Account

There are two ways to add an EC user account, as described in “Account” on page 84. This example shows how to create a native user account.

Complete these steps to add a native user:

1. Click the Configure tab.
2. In the Views pane, select **Accounts**. The Accounts pane opens, as shown in Figure 5-10 on page 85.
3. Click **New**.
4. In the New Account pane, click **Create Native User** as shown in Figure 5-11 on page 85**.** The **New Account** dialog opens.
5. Enter a user name and a password for the account. The example uses the user name **vm_app_admin**.
6. Select resource pool that you created in “Resource pools” on page 89 to add to the account. as shown in Figure 5-22.

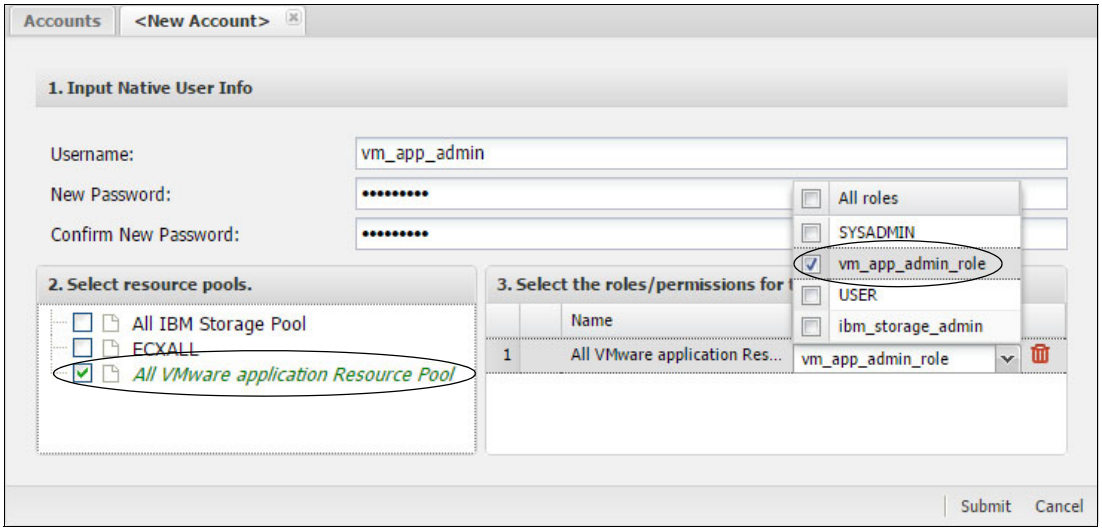


Figure 5-22 New account creation with resource pool and role assigned

7. Select the role that you created in “Roles” on page 93 to associate with the resource pool as shown in Figure 5-22.
8. Click **Submit** to save the new user account. The account then appear on the Accounts pane.

Application administrator login

After the account is created with appropriate resources and roles, the application administrator can log in to the ECX management portal with the appropriate credentials as shown in Figure 5-23. The administrator will be able to access only the screens and resources that were granted permission for.

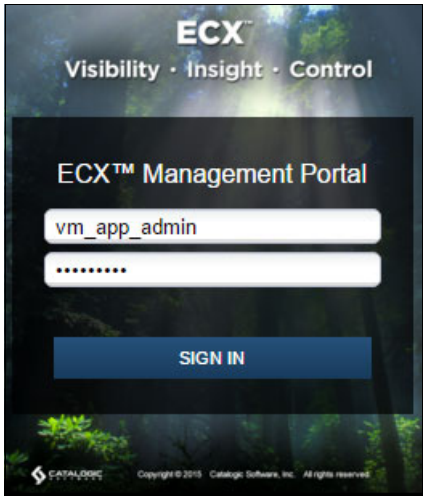


Figure 5-23 Application admin login to the portal

Figure 5-24 shows that only the screens that were assigned in Figure 5-20 on page 92 are visible and only the policies that were assigned in the resource pool created in Figure 5-18 on page 91 are accessible.

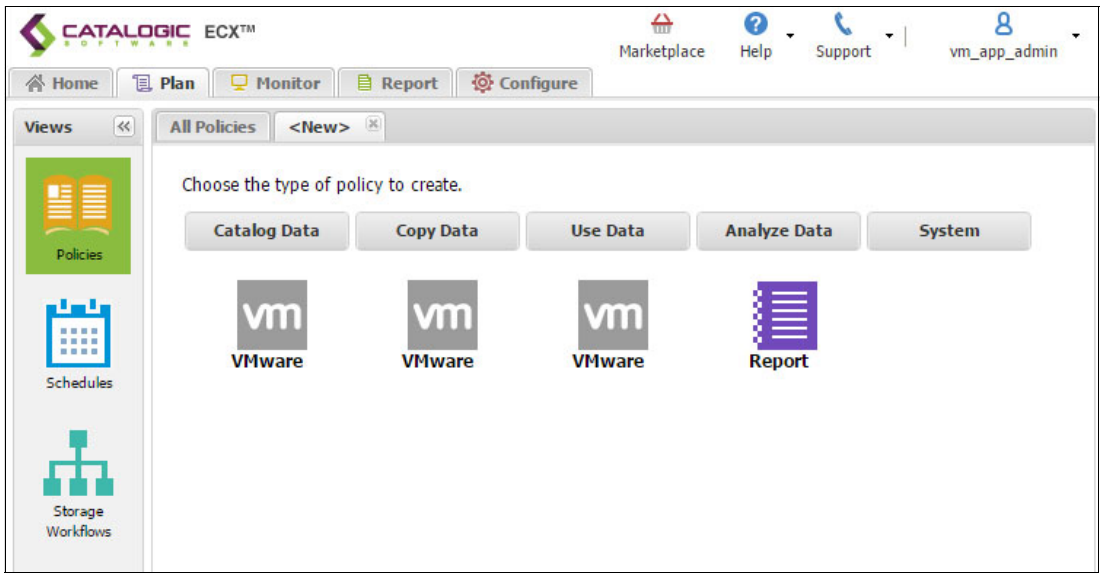


Figure 5-24 ECX Plan window while logged in as application administrator

5.4 Self-service in ECX

This section covers how the roles of the storage administrator and the application administrator enable self-service in the ECX console.

5.4.1 Role of storage administrator

In ECX, the role of a storage administrator is simplified by ability to create storage workflows. This capability eliminates manual storage provisioning and monitoring of storage resources.

Storage workflows

As described in 3.2, “ECX Copy Workflow technical overview” on page 38, storage workflows allow the storage administrators to create multiple templates, as shown in Figure 5-25 that define the sequence of operations. The administrator can then assign the service levels to each of the templates.

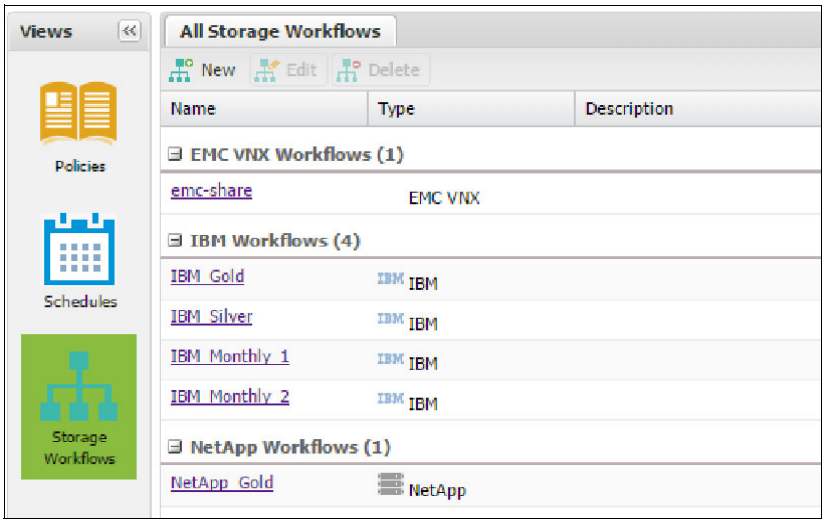


Figure 5-25 Storage Workflows in the ECX management portal

Complete these steps to create a storage workflow:

1. Click the Plan tab.
2. In the Views pane, select **Storage Workflows**.
3. In the All Storage Workflows pane, click **New.** The New Storage Workflow pane opens, as shown in Figure 5-26.

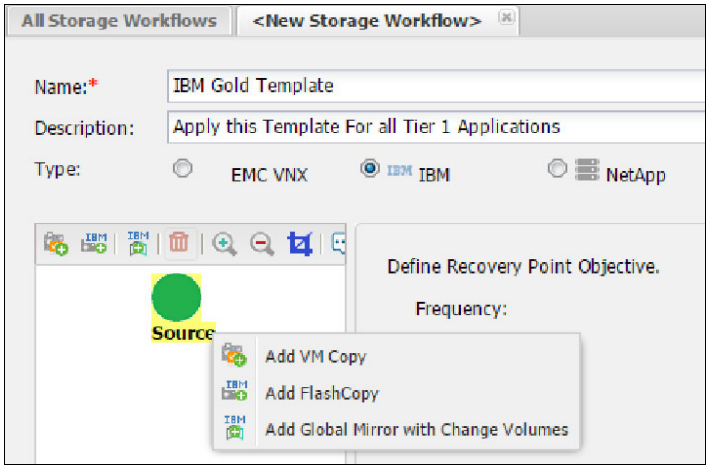


Figure 5-26 New Storage Workflow and subpolicies

4. Enter a name and a meaningful description for the Storage Workflow. The example uses the name **IBM Gold Template** and the description **Apply this Template For all Tier 1 Applications**.
5. In the **Type** field, select a type of workflow to create based on your storage provider. Select **IBM** to create an IBM storage workflow that allows you to define FlashCopies, Global Mirror with Change Volumes relationships, and VM Copies, as shown in Figure 5-26 on page 96.
6. The **Define Recovery Point Objective** options allow you to dictate the minimum frequency and interval at which copy operations must be made for the source resources that this storage workflow will be applied to in Copy Data policies. In the **Frequency** field, select **Hourly**, **Daily**, **Weekly**, or **Monthly**, and then set the **Interval** value.

> **Note:** The **Define Recovery Point Objective** options that were selected in step 6 on page 97 are applied to all copy subpolicies included in the storage workflow.
>
> For an ECX Copy Data policy that uses a GMCV relationship, ECX creates FlashCopies of the destination volume and then catalogs these copies for subsequent use in ECX Use Data policies. The **Define Recovery Point Objective** options of the related storage workflow determine the frequency and retention for these remote FlashCopies.

7. To add a **FlashCopy** subpolicy to an IBM Storage Workflow, complete these steps:
   a. Right-click the green source icon and select **Add FlashCopy**.
   b. In the Options pane, set these FlashCopy subpolicy options as shown in Figure 5-27:
      i. **Keep Snapshots:** After a certain number of snapshot instances are created for a resource, older instances are purged from the storage controller. Enter the age of the snapshot instances to purge in the **Days** field, or the number of instances to keep in the **Snapshots** field.
      ii. **Name**: Enter an optional name to replace the default FlashCopy subpolicy name displayed in ECX. The default initial name is **FlashCopy0**.
      iii. **FlashCopy Volume Prefix**: Enter an optional label to identify the FlashCopy. This label is added as a prefix to the FlashCopy name created by the policy.

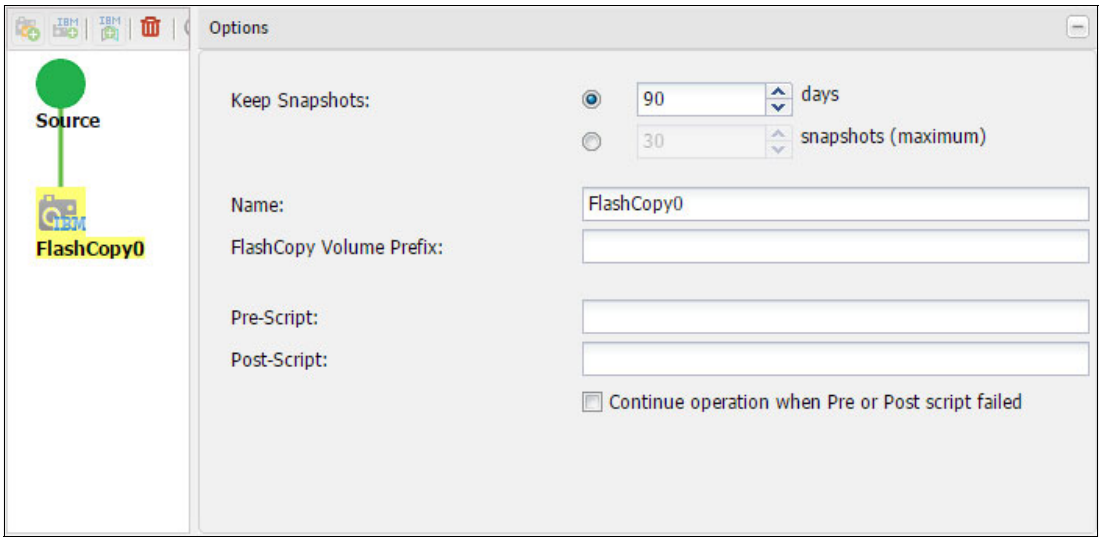


Figure 5-27 FlashCopy subpolicy and options

Note: FlashCopy labels must contain only alphanumeric characters and underscores.

8. To add a **Global Mirror** subpolicy to an IBM Storage Workflow, complete these steps:
   a. Right-click the green icon and select **Add Global Mirror with Change Volumes**.
   b. In the Global Mirror with Change Volumes Destination pane, select an IBM host destination from the list of available resources as the Global Mirror destination, along with an associated storage pool. If no storage pool is selected, the storage pool with the largest amount of space available is chosen by default.
   c. In the Options pane, set the following Global Mirror with Change Volumes subpolicy options as shown in Figure 5-28 on page 99:
      i. **Keep Snapshots:** After a certain number of snapshot instances are created for a resource, older instances are purged from the storage controller. Enter the age of the snapshot instances to purge in the **Days** field, or the number of instances to keep in the **Snapshots** field.
      ii. **Name**: Enter an optional name to replace the default Global Mirror subpolicy name displayed in ECX. The default initial name is Global Mirror0.
      iii. **Keep Source Volume name for target volume:** Enable this option to retain the source volume name for copy data generated by ECX.
      iv. **Volume Prefix Label:** Enter an optional label to identify the volume. This label is added as a prefix to the volume name created by the policy and cannot be edited after the policy is submitted.

      Note: Volume prefix labels must contain only alphanumeric characters and underscores. Labels cannot begin with numeric characters

      v. **Cycle Period** (seconds): Specify the time in which the change volumes are refreshed with a consistent copy of the data. If a copy does not complete in the cycle period, the next cycle period will not start until the copy is complete. The range of possible values is 60 - 86400. The default is 300.
      vi. **Global Mirror Volume Prefix**” Enter an optional label to identify the Global Mirror. This label is added as a prefix to the Global Mirror name created by the policy.

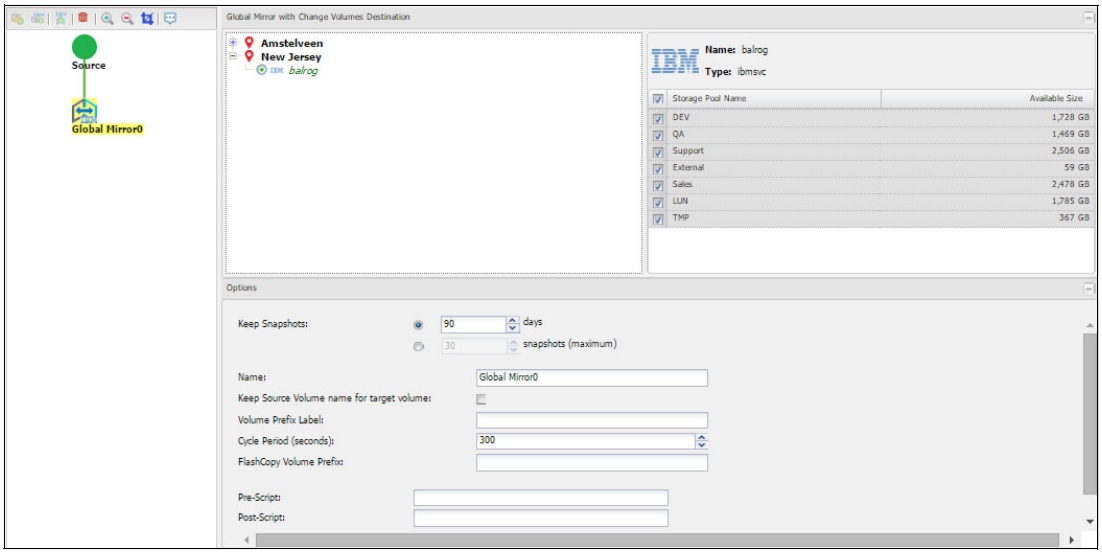


Figure 5-28 Global Mirror subpolicy, destination, and options

9. To add a **VM Copy** subpolicy to an IBM Storage Workflow, complete these steps:
   a. Right-click the green icon and select **Add VM Copy**.
   b. In the VM Copy Destination pane, select an IBM host destination from the list of available resources as the VM Copy destination, along with an associated storage pool. If no storage pool is selected, the storage pool with the largest amount of space available is chosen by default.
   c. In the Options pane, set the following VM Copy subpolicy options as shown in Figure 5-29 on page 100:
      i. **Keep Snapshots:** After a certain number of snapshot instances are created for a resource, older instances are purged from the storage controller. Enter the age of the snapshot instances to purge in the **Days** field, or the number of instances to keep in the **Snapshots** field.
      ii. **Snapshot Prefix Label:** Enter an optional label to identify the snapshot. This label is added as a prefix to the snapshot name created by the policy.

      > **Note:** Snapshot labels must contain only alphanumeric characters and underscores.

      iii. **Name:** Enter an optional label to replace the default snapshot subpolicy label displayed in ECX. The default initial label is **VM Copy0**.
      iv. **Protocol:** Select the transport protocol to use when copying data to IBM storage. Available protocols include **iSCSI** and **Fibre Channel**.

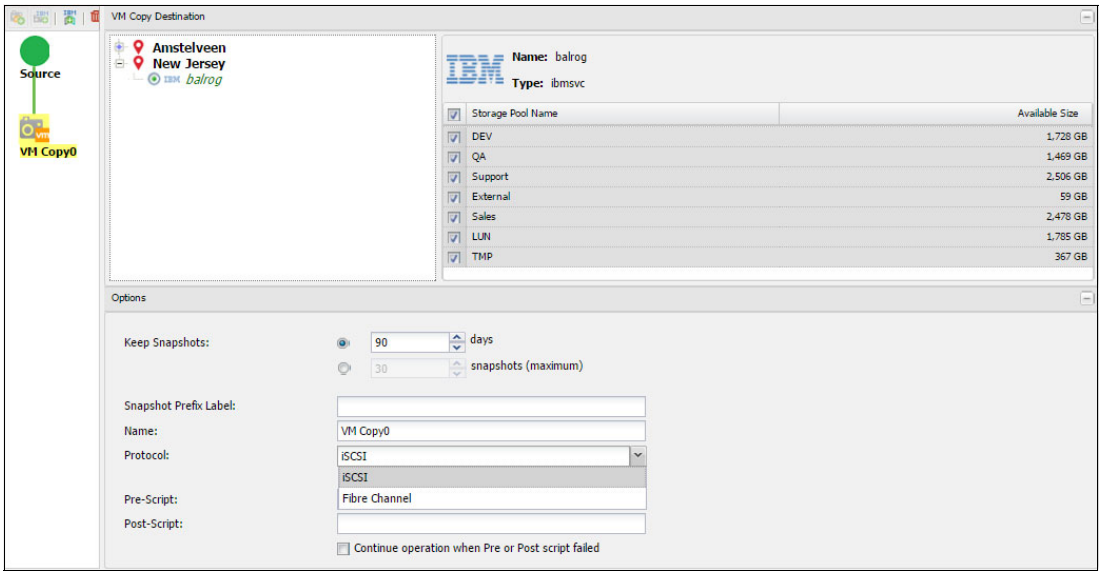


Figure 5-29 VM Copy subpolicy, destination, and options

10. When you are satisfied that the Storage Workflow-specific information is correct, click **Submit**. The Storage Workflow appears on the All Storage Workflows pane and can be applied to new and existing Copy Data policies.

Storage Workflows can be created with just one subpolicy or a combination of subpolicies. Multiple such Storage workflows can be created to suit the service level requirements in each enterprise. Figure 5-30 shows a typical “Gold” IBM Storage Workflow.

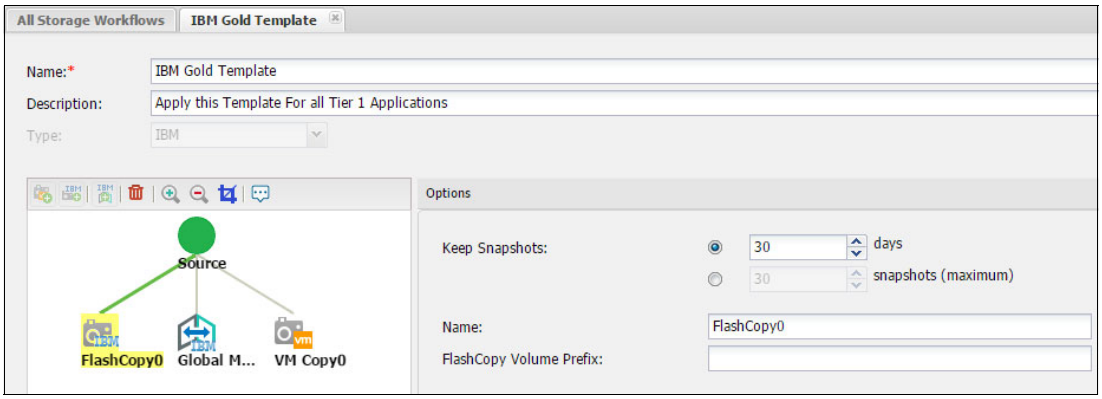


Figure 5-30 IBM Gold Storage workflows template

5.4.2 Role of application owner

With ECX, application administrators and VMware administrators can quickly create Copy and Use policies without waiting for storage to be provisioned. This section shows how the application administrator can take advantage of self-service capabilities in copy and use policies. For more information about ECX Copy Policies and Use Policies, see 3.2, “ECX Copy Workflow technical overview” on page 38 and 4.2, “ECX Use Data Workflow technical overview” on page 56.

Copy policies

In a VMware Copy policy, as an application owner, select the resources belonging to the application group, as shown in Figure 5-31.

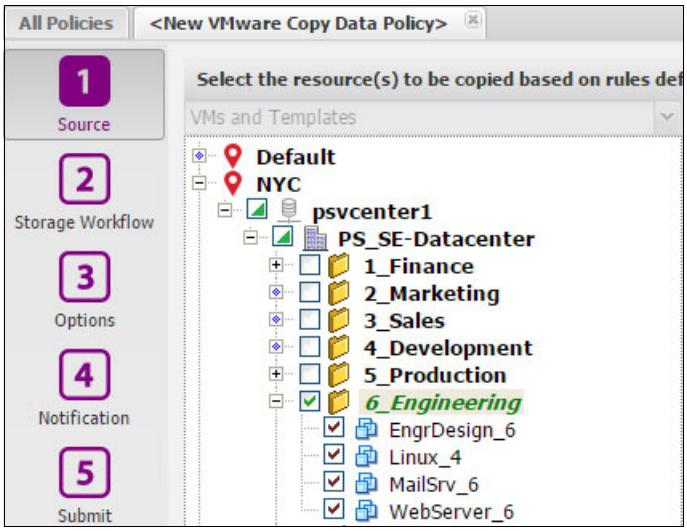


Figure 5-31 Select the application resources

Apply the wanted storage workflow created by the storage administrator, as shown in Figure 5-32.

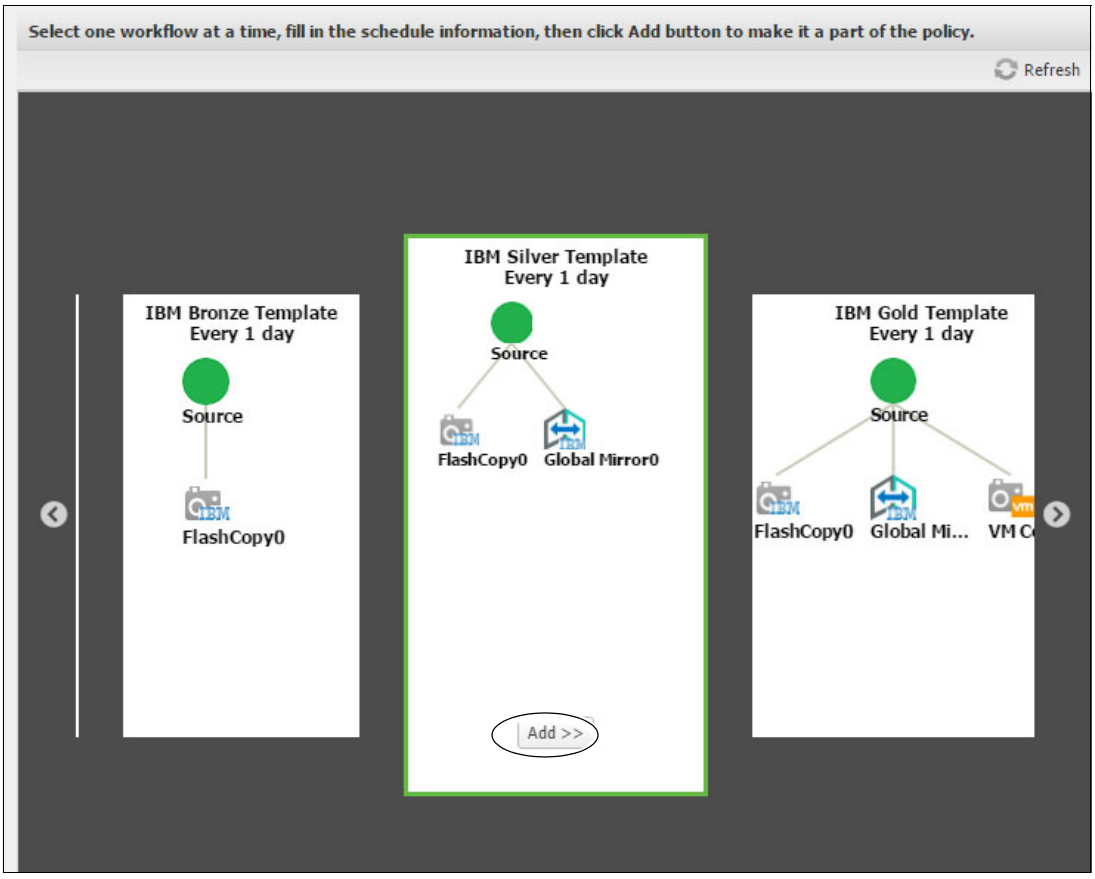


Figure 5-32 Template selection in Copy policy

Note: Application owners only see the Storage workflows assigned to them through ECX role-based access controls.

Use policies

Use Policies start an entire VMware or Application infrastructure in a few minutes, either in a production or a fenced off environment. Application owners can quickly use any of the predefined policies. Each policy contains the VMs and the sequence in which they are started, as seen in Figure 5-33, in a repeatable fashion by scheduling Use policies to start the application group.

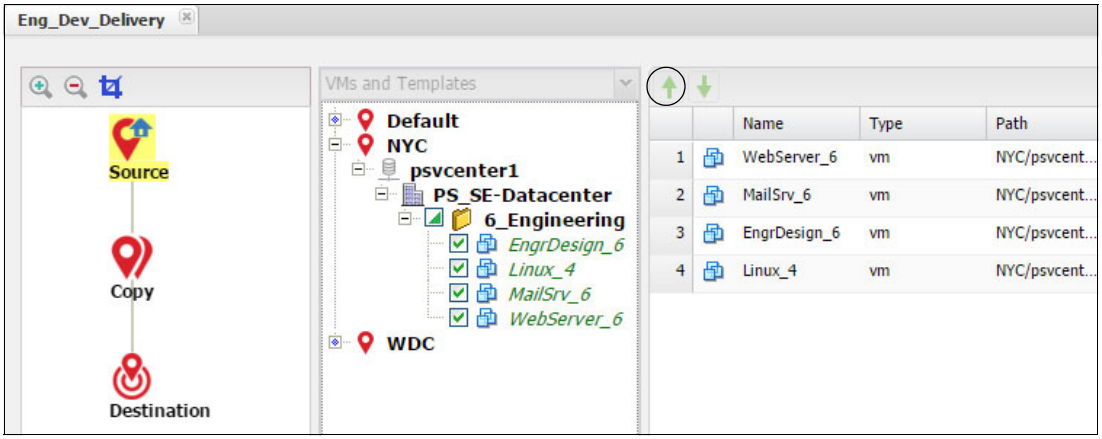


Figure 5-33 Use Policy with VMs in the application group and their start sequence

The application resources can also be spun up in an ad hoc fashion in test, production, or clone mode whenever the application owner desires. The following example demonstrates how to spin up and clean up environments with ECX Use policies.

Start a job session to spin up Test/Dev environments

To start a job session to start Test/Dev environments, complete these steps:

1. Click the Monitor tab in the Views pane and select **Jobs**. You will see all the policies that the Application owner has access to as shown in Figure 5-34.

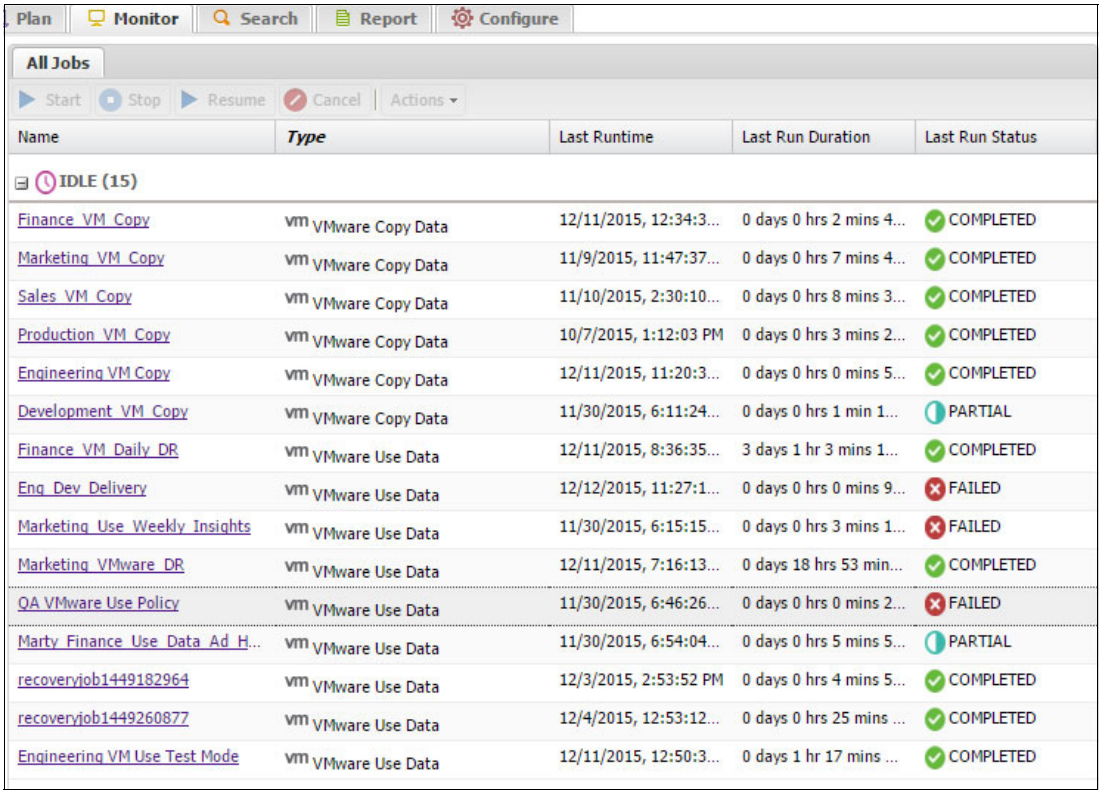


Figure 5-34 All Copy and Use Workflows accessible to application owner

2. The jobs automatically run based on the schedules applied to them. You can bypass a schedule that is associated with a policy and run a job session on demand from the **Jobs** pane on the Monitor tab. To start a job, click the row containing the job name and click **Start**.

3. A confirmation dialog box opens, and you are prompted to select a job mode, as shown in Figure 5-35.

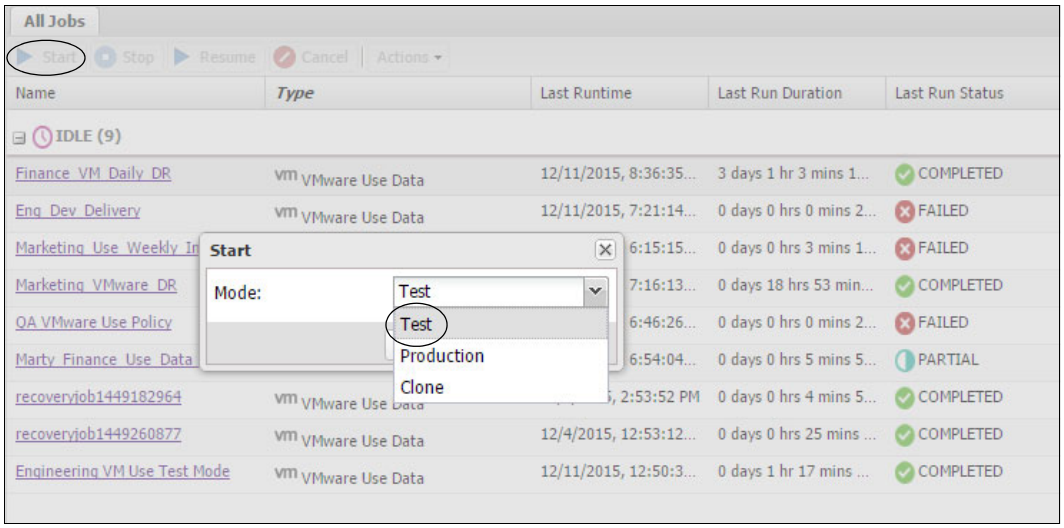


Figure 5-35 Start a job in test mode

4. Select **Test** and click **OK** to run the job session.
5. Click the job name to view the job session details, including the job session's start date and time, duration, description, status through a progress bar, and associated messages, as shown in Figure 5-36.

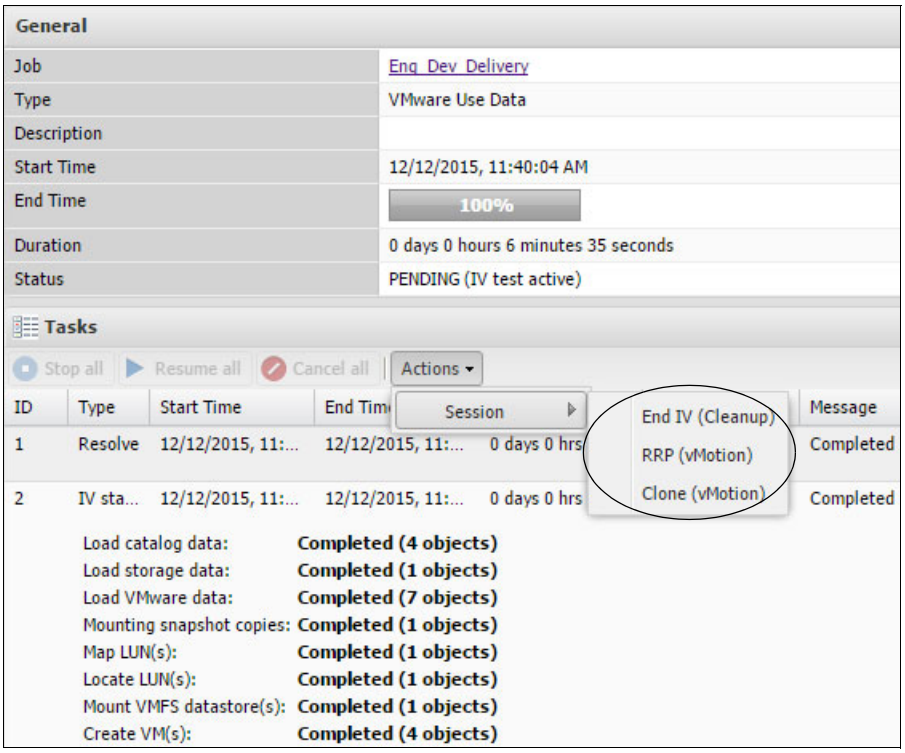


Figure 5-36 Active jobs session actions

Active job actions

After the policy completes successfully, select one of the following options shown in Figure 5-36 on page 104 from the **Actions** menu on the General tab of the job session on the Monitor tab:

- **End IV (Cleanup)** : This option removes all the test Virtual machine created by “Start a job session to spin up Test/Dev environments” on page 103 and cleans up the mount points to the storage
- **RRP (vMotion):** This option refers to Rapid Return to Production, and creates copies of virtual machines for use cases that require permanent or long-running copies back to the original Production environment
- **Clone (vMotion):** Creates copies of virtual machines for use cases that require permanent or long-running copies for data mining or duplication of a test environment on a fenced network.

6

Copy data catalog and analytics with Catalogic ECX

This chapter describes the powerful metadata catalog in ECX, and the search and analytics capabilities that can be driven from the cataloging operation.

This chapter includes the following sections:

- Introduction
- Copy data catalog
- Search
- Data analytics

6.1 Introduction

IT departments need tools that can help them gain visibility into storage and virtual infrastructure objects, increase efficiency of utilization, cope with the increased scale of data, and keep costs low. Catalogic ECX software delivers in all four areas. ECX software provides an enterprise-wide view of data objects and assets, allowing you to discover, record, search, and report on data, giving you greater understanding of your most important organizational assets.

6.2 Copy data catalog

Catalogic ECX Enterprise Catalog gathers, centralizes, and catalogs information about your IBM and VMware environments, allowing you to search, report, analyze, and take action on that information. By cataloging all tiers of your infrastructure, search tasks can be completed in seconds, protection compliance validated, stale or unwanted data identified, and much more. By having a central catalog of information, users gain unprecedented IT insight that can lead to new levels of efficiency and accelerate the use of snapshot technology for efficient data protection.

6.2.1 IBM Storage Catalog Data policy

An IBM Storage Catalog Data policy provides the framework to catalog and collect information about high-level IBM objects on your IBM storage systems. You can select one or more IBM providers in a single policy for cataloging.

Run an IBM Storage Catalog Data policy to analyze your IBM environment in real time and navigate to and correlate the objects from across the Enterprise into a single view. Additionally, you can discover sprawl, storage overutilization, and other storage inefficiencies.

To create an IBM Catalog Data policy, complete these steps:

1. Click the **Plan** tab.
2. On the Views pane, select **Policies**.
3. Click the **All Policies** tab.
4. Click **New**, then select **IBM Storage** in the **Catalog Data** column. The IBM Storage Catalog Data Policy wizard opens.

5. In step 1 of the wizard, Source, select one or more providers to catalog, as shown in Figure 6-1, then click **Next**.

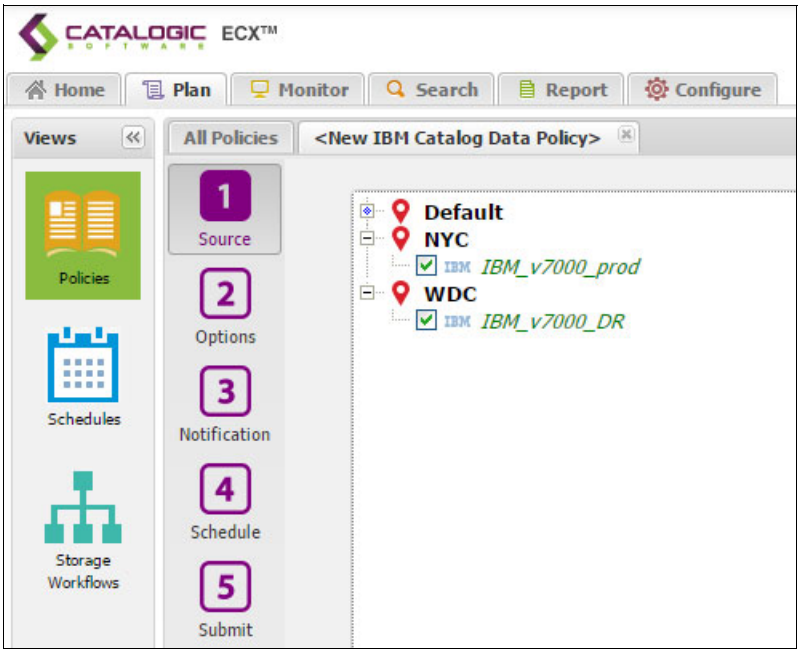


Figure 6-1 IBM Storage selection in Catalog Policy

6. In step 2, Options, select the options for your policy, as shown in Figure 6-2, and click **Next**. The **Number of catalog instances to keep** option sets the number of job runs for a policy after which older IBM objects for that job are purged from the Catalog.

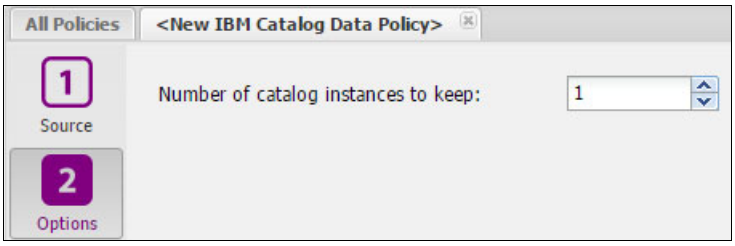


Figure 6-2 Number of Catalog Instances to retain

7. In step 3, Notification, you can select the notification options for your policy:
   - **SMTP Server:** From the list of available SMTP resources, select the SMTP Server to use for job status email notifications. If an SMTP server is not selected, no email will be sent.
   - **E-mail Address**: Enter the email addresses of the status email notifications recipients. Click **Add** to add it to the list.

   Click **Next** to proceed to the next step in the wizard.
8. In step 4, Schedule, select **Start job now** to create a policy that starts the job immediately. Select **Schedule job to start at later time** to view the list of available schedules. Optionally, select one or more schedules for the job. As each schedule is selected, the schedule's name and description displays. A policy that is paired with a schedule is a job. After your schedule selections are made, click **Next** to proceed to the next step in the wizard.

Note: To create and select a new schedule, click **Views**, then select **Schedules**. Create a schedule, then return to the policy editor, refresh the Available Schedules pane, and select the new schedule.

9. In step 5, Submit, enter a name for your policy and a meaningful description, as shown in Figure 6-3, and click **Submit** to save the policy.

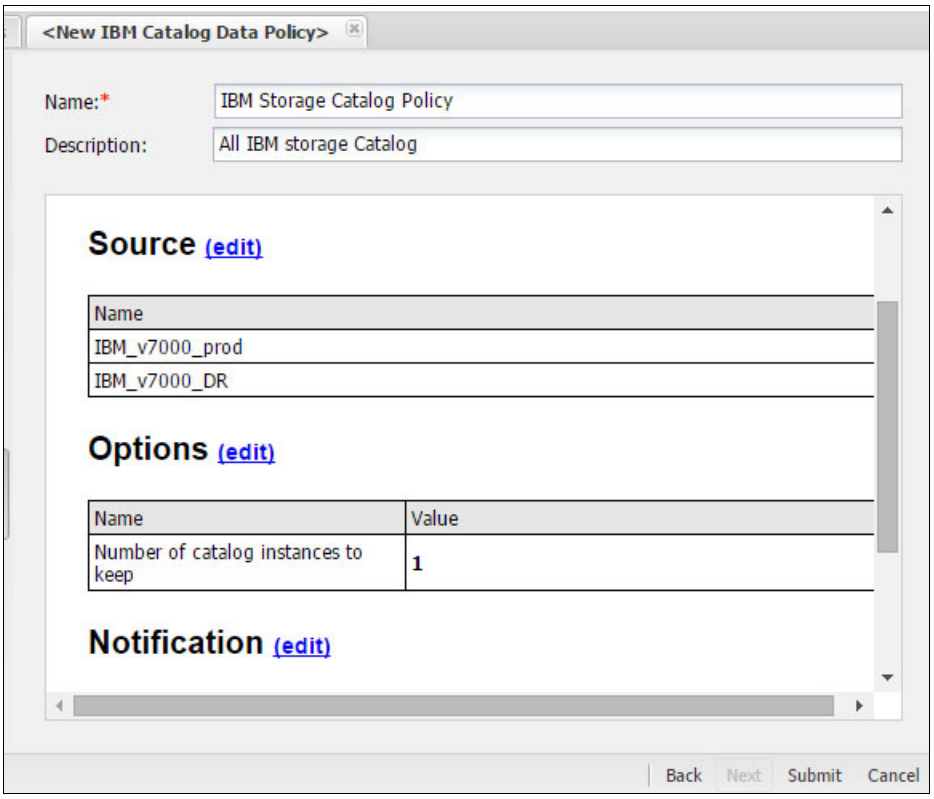


Figure 6-3 Name and Description of Policy

10. Click the All Policies tab. Your new policy appears in the policy list.

6.2.2 VMware Catalog Data policy

A VMware Catalog Data policy provides the framework to catalog and collect information about VMware objects. You can select one or more VMware providers in a single policy for cataloging.

Run a VMware Catalog Data policy to analyze your VMware environment in real time, and navigate to and correlate the objects from across the Enterprise in a single view. Additionally, you can detect VM sprawl, storage over utilization, and other storage inefficiencies.

To create a VMware Catalog Data policy, complete these steps:

1. Click the **Plan** tab.
2. On the Views pane, select **Policies.**
3. Click the **All Policies** tab.

4. Click **New**, then select **VMware** in the **Catalog Data** column. The **VMware Catalog Data Policy** wizard opens, as shown in Figure 6-4.

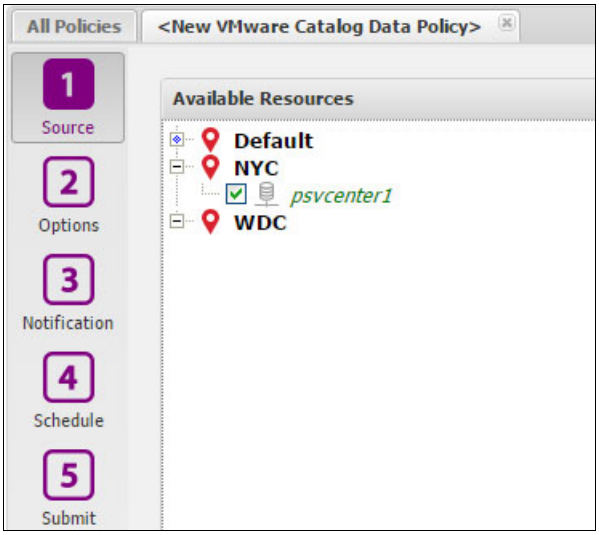


Figure 6-4 VMware resource selection

5. In step 1 of the wizard, Source, select one or more providers to catalog from the list of available providers, as shown in Figure 6-4, then click **Next** to proceed to the next step in the wizard.
6. In step 2, Options, select the following options for your policy as shown in Figure 6-5:
   - **Connection time-out (secs)**: To run a catalog job, the application needs to connect with the resource. If there is no response within a certain time limit, it times out and the job session fails. Enter the number of seconds to wait before timing out.
   - **Number of catalog instances to keep**: After a certain number of job runs for a policy, older VMware objects for that job are purged from the Catalog. Enter the number of job runs for which VMware objects are to be retained.

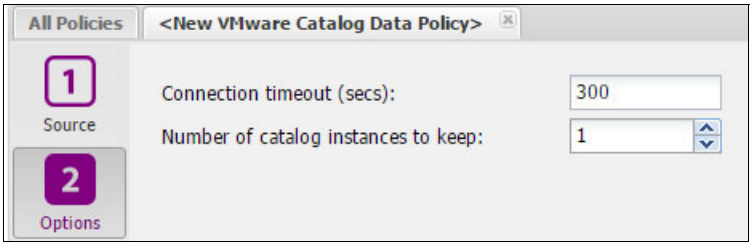


Figure 6-5 Policy options

   Click **Next** to proceed to the next step in the wizard.
7. In step 3, Notification, you can select the notification options for your policy:
   - **SMTP Server:** From the list of available SMTP resources, select the SMTP Server to use for job status email notifications. If an SMTP server is not selected, no email is sent.
   - **E-mail Address:** Enter the email addresses of the status email notifications recipients. Click **Add** to add it to the list.

   Click **Next** to proceed to the next step in the wizard.

8. In step 4, Schedule, select **Start job now** to create a policy that starts the job immediately. Select **Schedule job to start at later time** to view the list of available schedules. Optionally, select one or more schedules for the job. As each schedule is selected, the schedule's name and description is displayed. A policy paired with a schedule is a job. After your schedule selections are made, click **Next**.

   Note: To create and select a new schedule, click **Views**, then select **Schedules**. Create a schedule, then return to the policy editor, refresh the Available Schedules pane, and select the new schedule.

9. In step 5, Submit, enter a name for your policy and a meaningful description, as shown in Figure 6-6, and click **Submit** to save the policy.

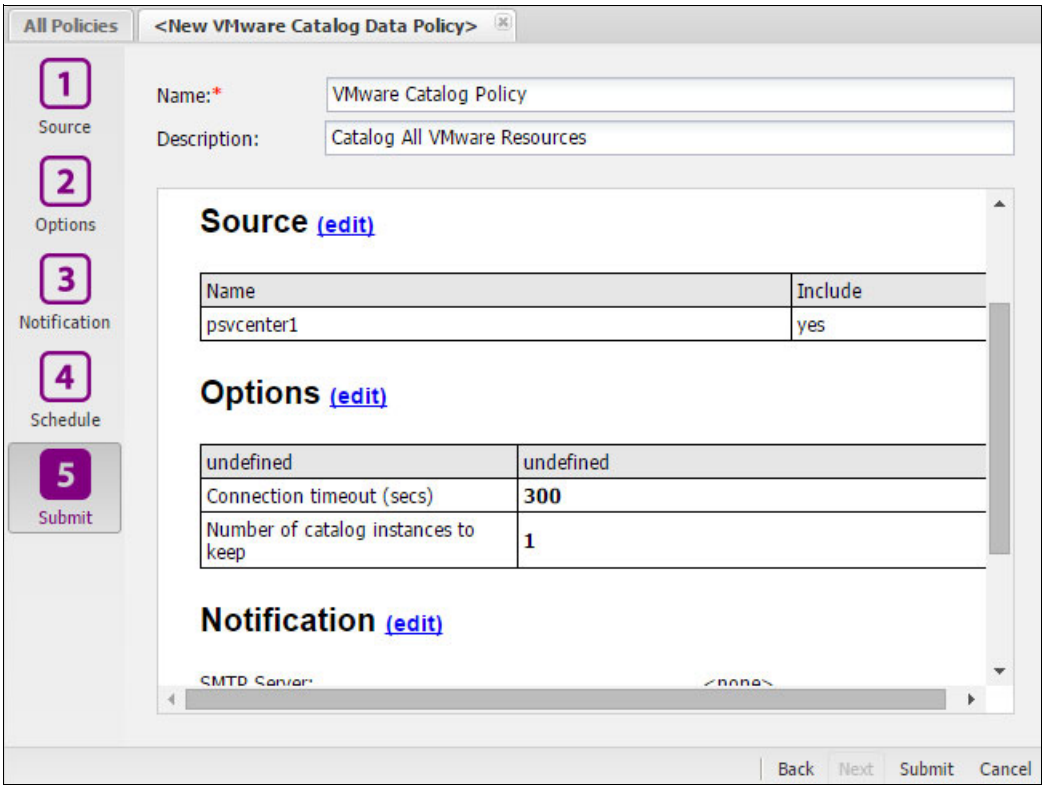


Figure 6-6 Name and description of policy

10. Click the All Policies tab. Your new policy appears in the policy list.

6.3 Search

ECX includes a powerful search capability that returns rich information in seconds. With the Search feature, you can easily search for and rapidly find all objects that match certain criteria. Searches can be filtered by date ranges, host names, and so on, allowing you to refine searches with precision. Search tasks that used to consume hours of valuable IT time are now completed in seconds.

The IBM search Catalog includes these categories:

- FlashCopies
- Hosts
- IOGroups
- MDisks
- Mirrors
- Node Canisters
- PortIPs
- Volumes

The VMware Catalog can include these categories:

- Datastores
- ESX hosts
- LUNs
- Virtual disks
- Virtual machines
- VMware hosts
- Virtual snapshots

The Recovery Catalog can contain these categories:

- Datacenters
- Datastores
- ESX hosts
- LUNs
- Folders
- Recovery points
- vApps
- vDisks
- vSnapshots
- vSpheres

6.3.1 Searching for objects

Use the search feature to find objects on providers that are cataloged in ECX. Examples of objects are volumes, files, snapshots, and virtual machines. You can tailor your search by applying filters.

There are two types of search: Basic and advanced.

Basic search examine all text fields. Enter a character pattern, which can including wildcards and inline search strings for more advanced searches. ECX searches the entire Catalog and returns all objects with names that match or contain the search entry.

Advanced search is similar to basic search, with an additional function. You can search and filter by object name, category, object type, and location. When searching for NetApp files, you can also filter by the last modified time, creation time, last accessed time, and file size.

To perform a basic search for objects, complete these steps:

1. Click the **Search** tab.
2. From the Search pane, as shown in Figure 6-7, you can perform a basic search or an advanced search.

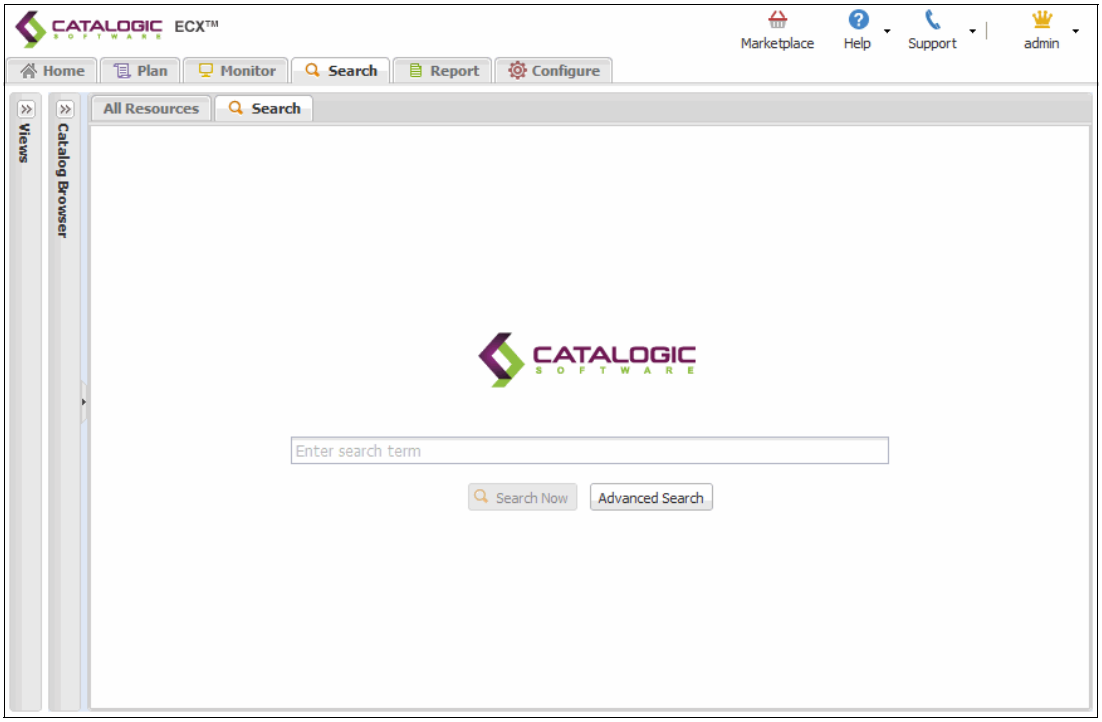


Figure 6-7 The ECX Search pane

3. In the **Enter search term** field, enter the character pattern to search on.

Tips: When searching, follow these guidelines:

- Enter a character string to find objects with a name that matches or contains the character string.
- Partial character strings are permitted.
- Character strings are case insensitive.
- An asterisk can be used as a wildcard.

4. Click **Search Now** to perform a search. The list of objects that meet all the criteria is displayed, as shown in Figure 6-8.

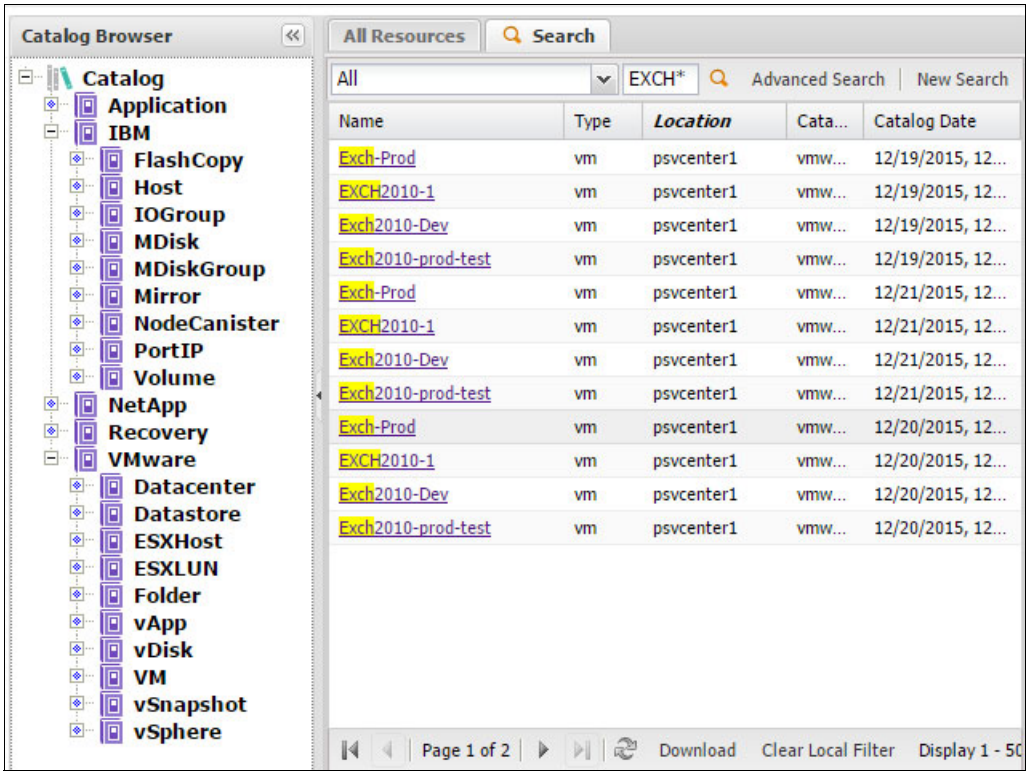


Figure 6-8 Search results and object browser

5. Click an object name to display the properties of the object in a new tab. The specific properties vary by type of object.

6.4 Data analytics

Analytics provide insight into storage consumption across all tiers, and data protection reporting quickly identifies unprotected objects to help you ensure compliance.

ECX provides a number of predefined reports that you can tailor to meet your specific reporting requirements. Reports are based on the data that is collected by the most recently run catalog policy. You can generate reports after all cataloging jobs and subsequent database condense jobs complete.

Click the Reports tab to display the Report Browser, as shown in Figure 6-9.

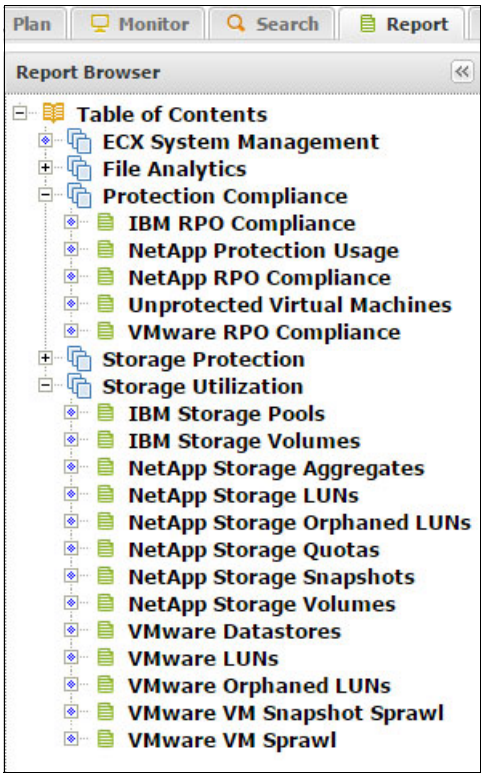


Figure 6-9 Report Browser

You can run reports with predefined default parameters, or run and save customized reports driven by custom parameters as shown in Figure 6-10.

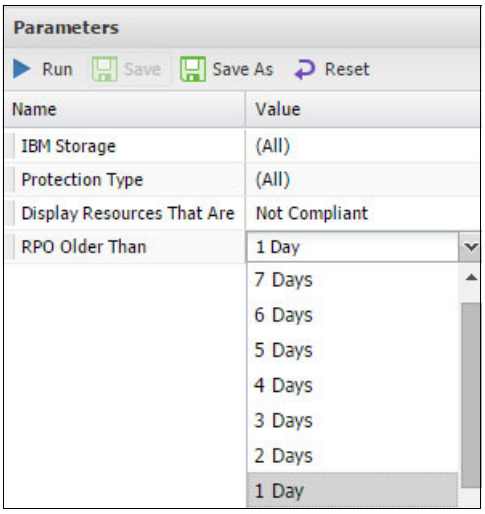


Figure 6-10 Report Parameters

The information in these reports is presented in a chart-based Quick View section, or tabular Summary View and Detail View sections.

You can add a Report policy to summarize information about cataloged providers and the data and other resources that are on them. You can also set the Report policy to run as defined by the parameters of the schedule.

To further analyze the data or print a hardcopy, use the export function to save the data from the generated report to an Adobe PDF, Microsoft Word file, Microsoft Excel file, or HTML file.

ECX reports are divided into different categories. This section describes some of them and offers a few examples.

6.4.1 Protection compliance

Protection Compliance reports help ensure that your data is protected through user-defined recovery point objective parameters.

VMware RPO Compliance report

This report displays the primary and replicated snapshot protection for virtual machines and data stores. Figure 6-11 shows a bar chart of VMs and data stores that are protected by primary snapshot, and ones that have a mirrored copy. All the VMs and data stores that are not compliant to primary protection are listed below the chart.

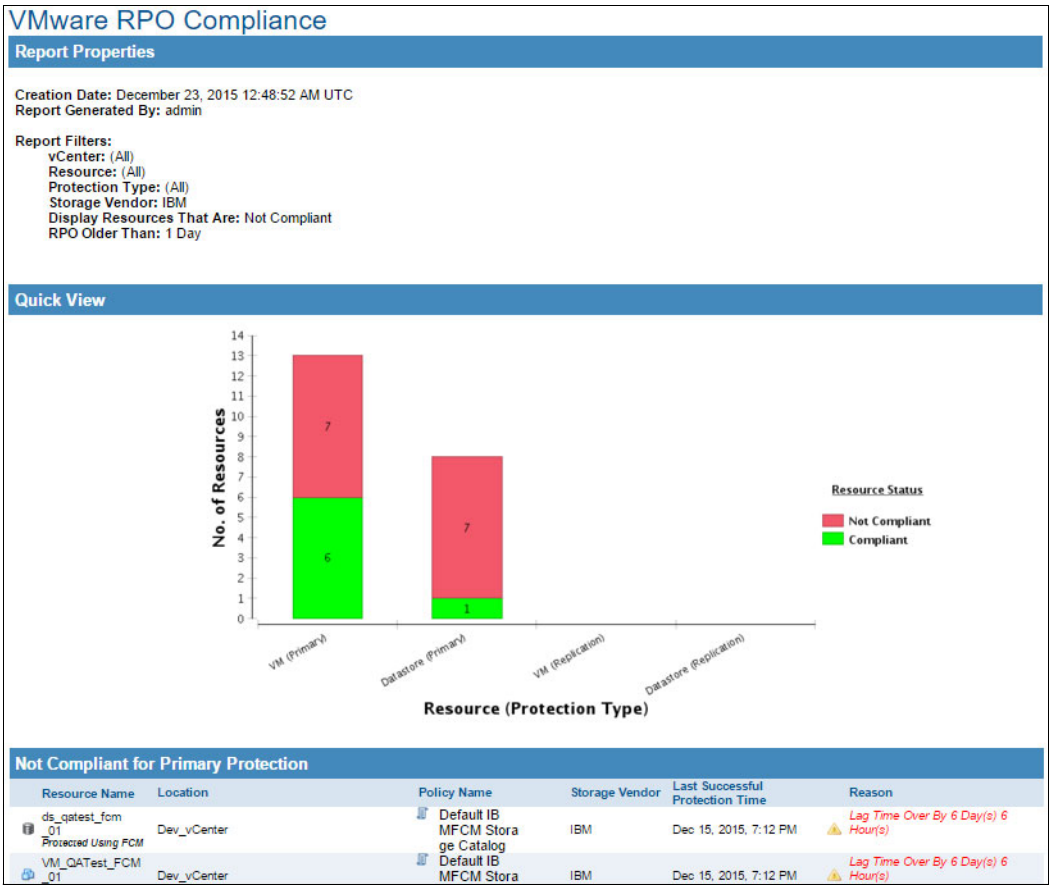


Figure 6-11 VMware RPO compliance report

Note: Compliance is defined in the report parameters before running a report.

IBM RPO Compliance report

This report displays IBM storage systems in relation to your recovery point objective parameters. Figure 6-12 shows a bar chart of all the IBM storage volumes that have a primary snapshot protection, and volumes that are replicated by using Global Mirror Changed Volume protection. The volumes that are not compliant are listed below the bar chart.

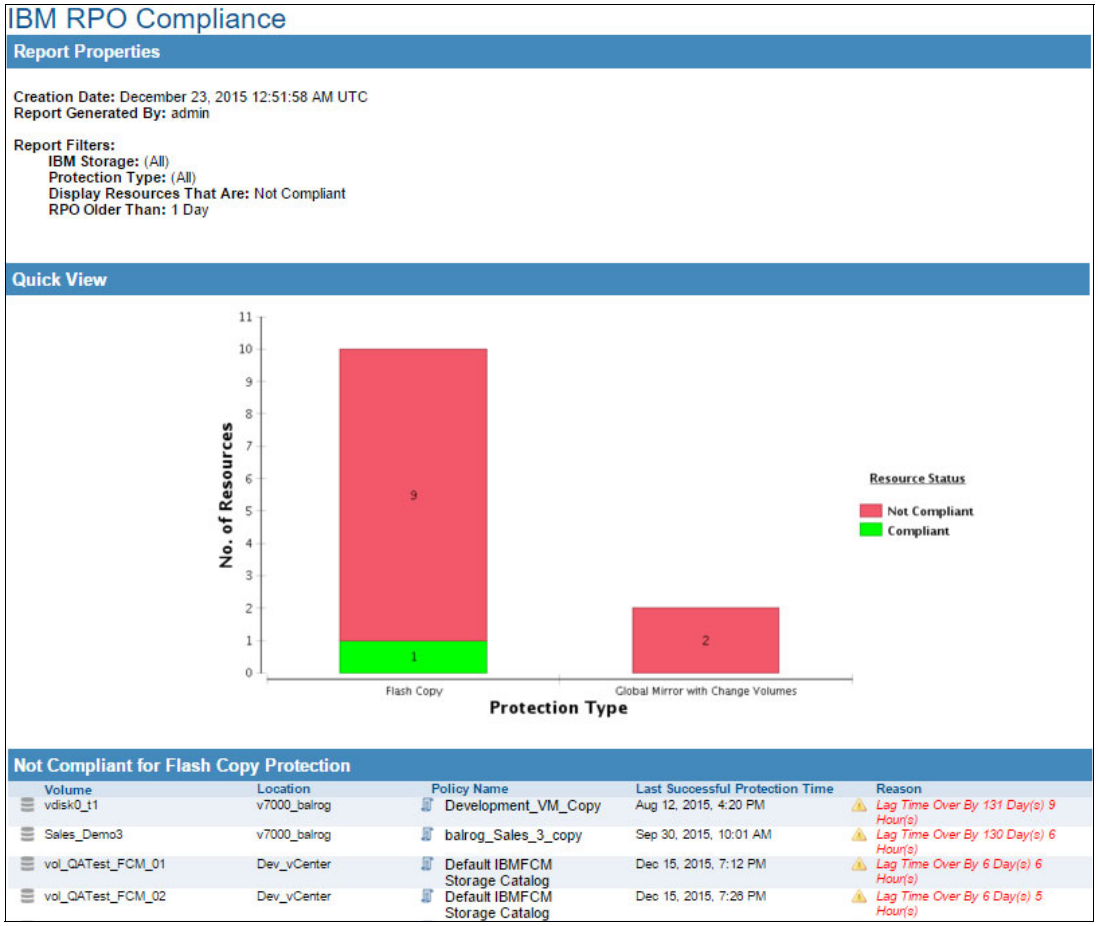


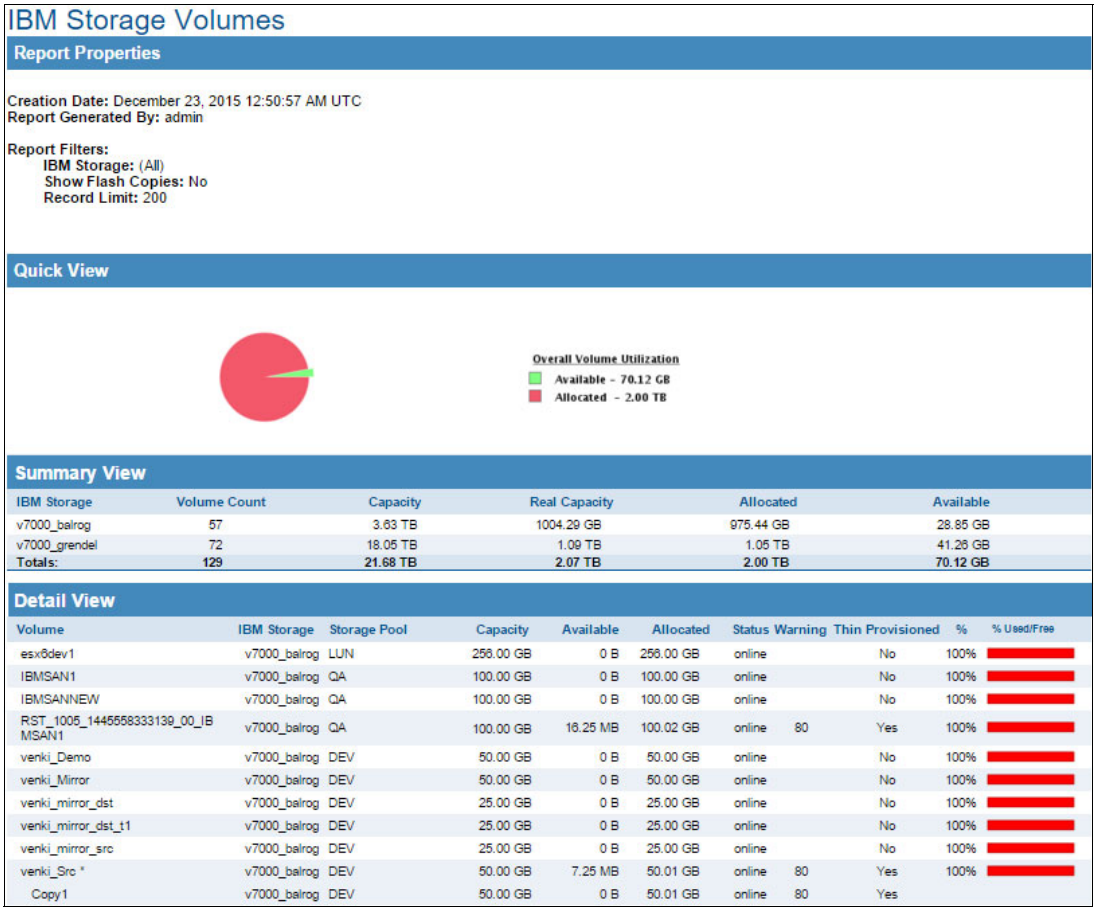
Figure 6-12 IBM Storage RPO Compliance report

6.4.2 Storage utilization

Storage Utilization reports are designed to help improve efficiency. These reports provide information about IBM and VMware storage utilization, including threshold warning reports, reports on potentially wasted resources (for example, VMware sprawl, orphaned LUNs), and informational reports. The Storage Utilization Reports help you review your storage needs and examine your storage capacity, and view the total and free space available as well as the total capacity of your volumes and storage pools.

IBM Storage Volumes report

This report helps you review the storage utilization of your IBM volumes. Figure 6-13 displays the total space consumed and the free space available on your IBM volumes.



IBM Storage Volumes

Report Properties

Creation Date: December 23, 2015 12:50:57 AM UTC
Report Generated By: admin

Report Filters:
IBM Storage: (All)
Show Flash Copies: No
Record Limit: 200

Quick View

Overall Volume Utilization
Available - 70.12 GB
Allocated - 2.00 TB

Summary View

IBM Storage	Volume Count	Capacity	Real Capacity	Allocated	Available
v7000_balrog	57	3.63 TB	1004.29 GB	975.44 GB	28.85 GB
v7000_grendel	72	18.05 TB	1.09 TB	1.05 TB	41.26 GB
Totals:	**129**	**21.68 TB**	**2.07 TB**	**2.00 TB**	**70.12 GB**

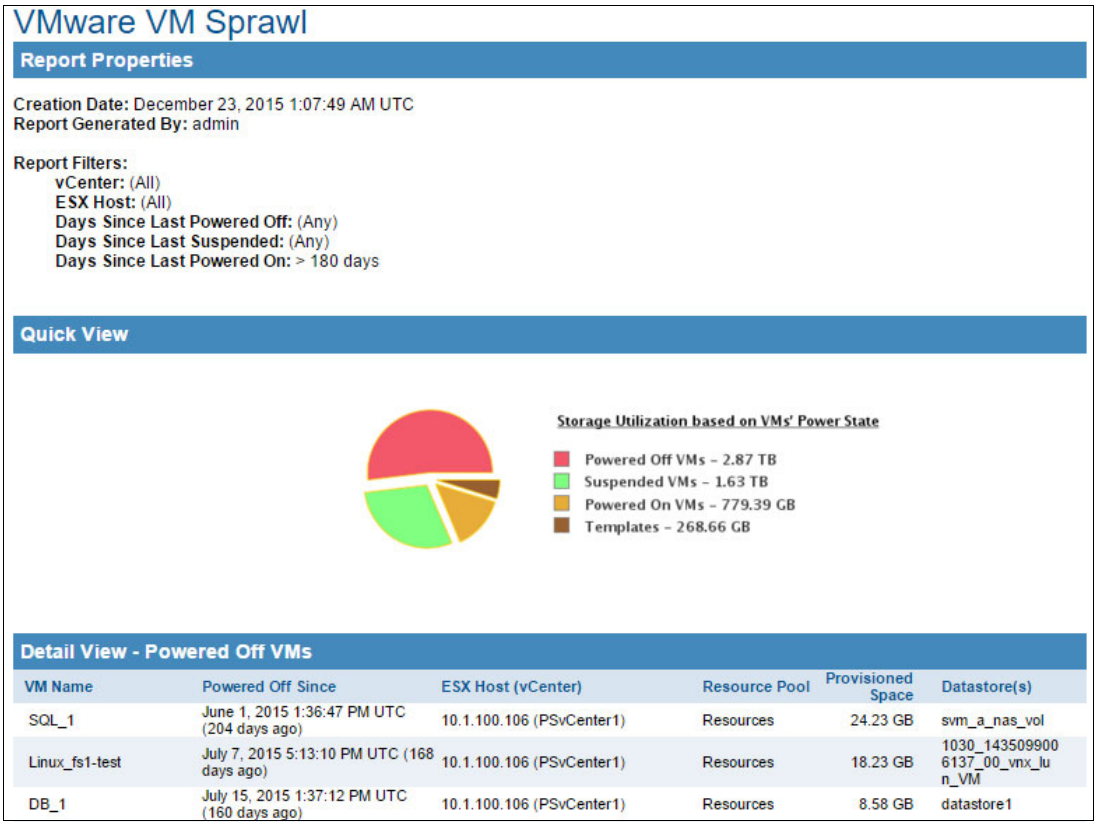
Detail View

Volume	IBM Storage	Storage Pool	Capacity	Available	Allocated	Status	Warning	Thin Provisioned	%	% Used/Free
esx6dev1	v7000_balrog	LUN	256.00 GB	0 B	256.00 GB	online		No	100%	
IBMSAN1	v7000_balrog	QA	100.00 GB	0 B	100.00 GB	online		No	100%	
IBMSANNEW	v7000_balrog	QA	100.00 GB	0 B	100.00 GB	online		No	100%	
RST_1005_1445558333139_00_IBMSAN1	v7000_balrog	QA	100.00 GB	16.25 MB	100.02 GB	online	80	Yes	100%	
venki_Demo	v7000_balrog	DEV	50.00 GB	0 B	50.00 GB	online		No	100%	
venki_Mirror	v7000_balrog	DEV	50.00 GB	0 B	50.00 GB	online		No	100%	
venki_mirror_dst	v7000_balrog	DEV	25.00 GB	0 B	25.00 GB	online		No	100%	
venki_mirror_dst_t1	v7000_balrog	DEV	25.00 GB	0 B	25.00 GB	online		No	100%	
venki_mirror_src	v7000_balrog	DEV	25.00 GB	0 B	25.00 GB	online		No	100%	
venki_Src *	v7000_balrog	DEV	50.00 GB	7.25 MB	50.01 GB	online	80	Yes	100%	
Copy1	v7000_balrog	DEV	50.00 GB	0 B	50.01 GB	online	80	Yes		

Figure 6-13 IBM Storage Volume report

VMware VM Sprawl report

This report displays storage utilization across virtual machines based on their power state and storage utilization across virtual machine templates, as shown in Figure 6-14.



VMware VM Sprawl

Report Properties

Creation Date: December 23, 2015 1:07:49 AM UTC
Report Generated By: admin

Report Filters:
vCenter: (All)
ESX Host: (All)
Days Since Last Powered Off: (Any)
Days Since Last Suspended: (Any)
Days Since Last Powered On: > 180 days

Quick View

Detail View - Powered Off VMs

VM Name	Powered Off Since	ESX Host (vCenter)	Resource Pool	Provisioned Space	Datastore(s)
SQL_1	June 1, 2015 1:36:47 PM UTC (204 days ago)	10.1.100.106 (PSvCenter1)	Resources	24.23 GB	svm_a_nas_vol
Linux_fs1-test	July 7, 2015 5:13:10 PM UTC (168 days ago)	10.1.100.106 (PSvCenter1)	Resources	18.23 GB	1030_143509900 6137_00_vnx_lu n_VM
DB_1	July 15, 2015 1:37:12 PM UTC (160 days ago)	10.1.100.106 (PSvCenter1)	Resources	8.58 GB	datastore1

Figure 6-14 VMware VM Sprawl report

7

Automation and DevOps with Catalogic ECX

This chapter describes the DevOps methodology in IT, and how ECX is used for the automation of complex tasks and enhanced DevOps.

This chapter includes the following sections:

- DevOps introduction
- Automation with ECX
- REST API and Scripting in ECX
- DevOps Automation with Puppet

7.1 DevOps introduction

DevOps, or a test-driven development (TDD) cycle, is a software development process that relies on the repetition of a short development cycle. First, the developer writes an (initially failing) automated test case that defines a wanted improvement or new function, produces the minimum amount of code to pass that test, promotes it to production, and finally refactors the new code to acceptable standards.

DevOps workflows can be improved by integrating and testing code against application environments that are as close to the production state as possible. To set up a test environment with production environment and data sets, developers would have to submit requests to their virtual and storage administrators to provision compute, network, and storage resources and perform copy operations. This process could take days or weeks, and is difficult to repeat reliably. Storage arrays with snapshot and cloning technology, such as IBM Spectrum Virtualize, can alleviate the need to copy data sets to the test environment. However, the process of finding the correct FlashCopy, creating a FlashCopy of the FlashCopy, and mounting the FlashCopy to the proper server is still a manual process and must be handled by the system administrators. Then there is also the work that is needed to provision new storage, networks, and virtual machines for the test environment.

Figure 7-1 shows the traditional DevOps practices.

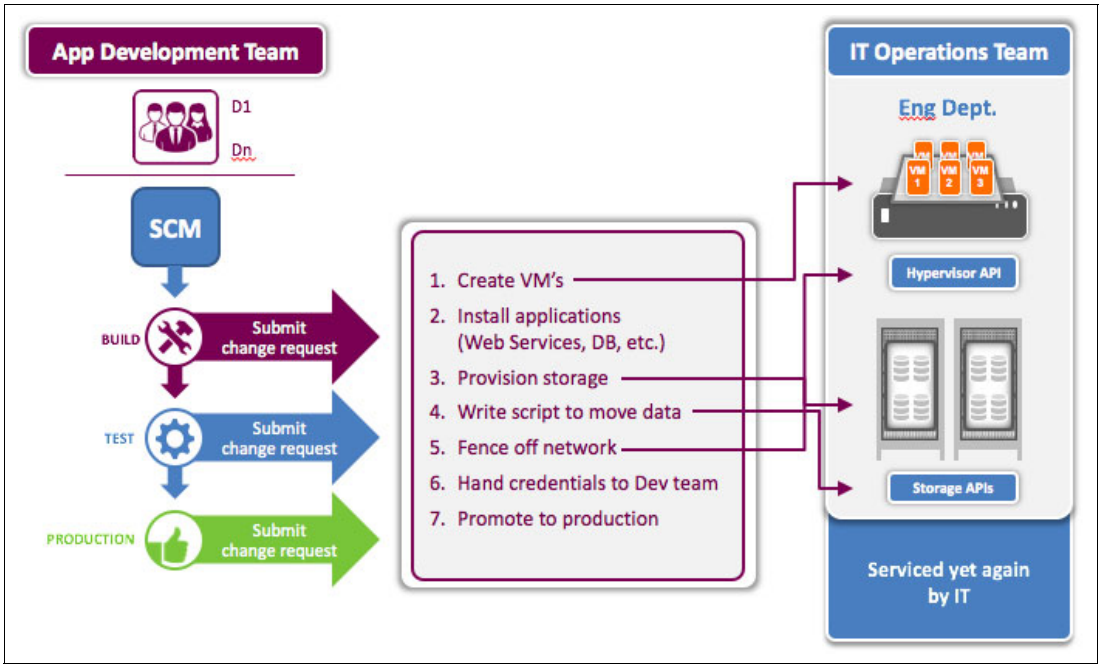


Figure 7-1 Traditional DevOps practices

Because this process is so time-consuming and error prone, providing the development organization with consistent access to current data sets and environments becomes unachievable. Instead of having the most relevant data every morning, for example, the process might only happen once a month. This stale environment drives down the quality of the new code being developed. In addition, in many global companies, testing of the new code might be performed by overseas personnel whose workday begins after the developers' has ended. In these scenarios, at the end of the development workday, the development team needs to "set up" the code they have worked on, such that it can be easily tested by these overseas resources. Inevitably, the test team has difficulty locating and setting up the proper

data at test time. This challenge wastes time and money, and usually places some burden back on IT again, after hours, to help solve the data access problems. The next section describes how ECX along with IBM Virtualized Storage can significantly improve your DevOps workflow.

7.2 Automation with ECX

Catalogic ECX and IBM Storage Systems built with IBM Spectrum Virtualize help users drastically improve the DevOps process by instantly providing copies of application environments and data sets in a virtualized test environment. With ECX automation, users can automate and orchestrate VMware, FlashCopies, and replicas of the latest production environment and data, to be instantly stood up in multiple locations. This process facilitates true DevOps efficiency, in a consistent and timely manner. The process provides these benefits:

- The ECX REST API provides a unified abstraction of the compute and storage infrastructure, also known as Infrastructure as Code
- No more scripts that fail, nor the need to troubleshoot failed scripts
- Automated workflows for accessing environments and data globally
- Moves organizations closer to a DevOps model
- Continuous development enables higher-quality code, faster
- Ability to implement new technology quicker, which provides a competitive edge

Figure 7-2 shows automation with ECX.

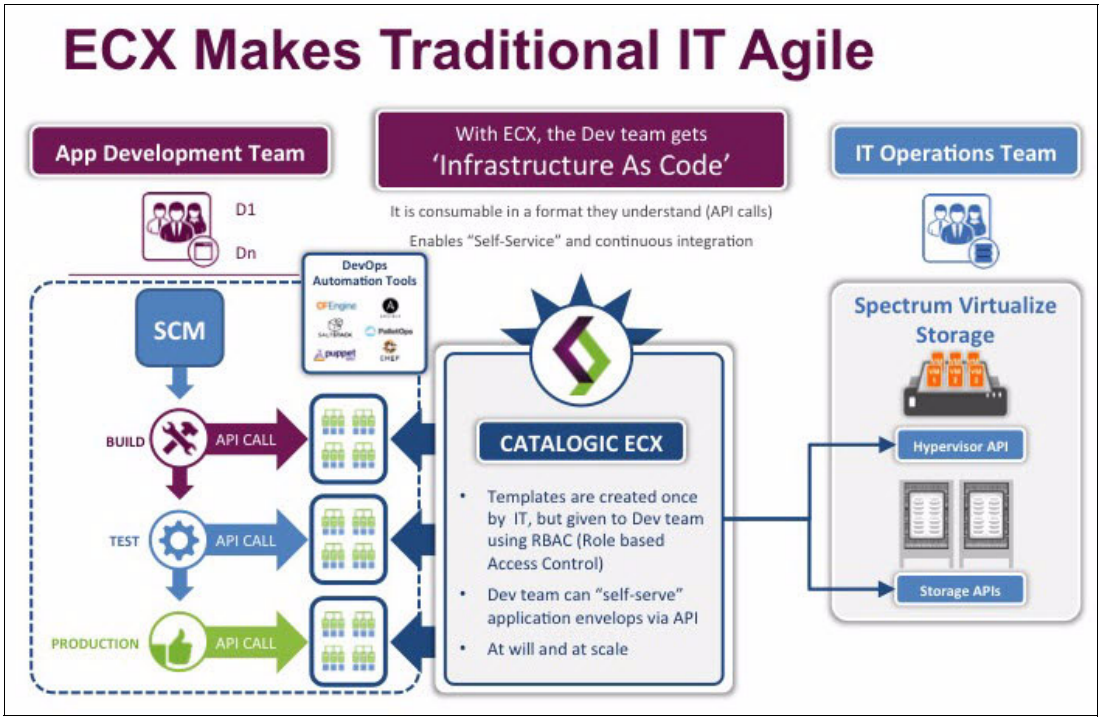


Figure 7-2 ECX Automation

The ECX workflow now can enable a true DevOps approach to Test/Dev with the ability to bring data up, in a protected environment, repeatedly with ease. DevOps is the next

generation of Test/ Dev and does require a close working relationship between the IT and the development teams. The ability to automate the process of making data available for Test/Dev provides for smooth interactions with development lines of business. ECX automates and abstracts this process of standing up copies of entire applications by using its Use Data Workflows. These workflows allow developers, by using the ECX REST API, to instantly stand up copies of production databases and applications. For more information about Use Data Workflows, see 4.2, “ECX Use Data Workflow technical overview” on page 56.

7.3 REST API and Scripting in ECX

DevOps toolchains, third-party tools, custom applications, and scripts can integrate with ECX by using the ECX REST API and Pre/Post Scripting integration.

7.3.1 ECX REST API

The ECX REST API covers all functions of ECX. The API requests and responses are in JSON format. The ECX REST APIs are organized by functionality:

- Authentication
- Account Management
- Resource Management
- Job Management
- Event/Log Management
- Alert Management
- Catalog Management
- Schedule Management
- Policy Management
- Search
- Report Management
- License Management

The full specification and examples of the ECX REST APIs are available from the ECX Marketplace that is accessible from within the ECX GUI. The **ECX Marketplace** button is in the upper right in the GUI next to the **Help** and **Support** buttons.

7.3.2 Pre/Post scripting

ECX Copy and Use policies can run Pre and Post scripts. This feature allows users to automate processes that they need to perform as part of the ECX policy jobs. The scripts can be specified in the **ECX Copy** and **Use workflow** options. The scripts run on the ECX virtual appliance, and can consist of one or many commands or the full path to a shell script. Command entries can include any entry that is valid for a Linux CentOS command shell. The ECX policy job checks the exit status from the scripts. If the scripts return a nonzero exit status value, the policy job fails with error. You can set an option to ignore the exit status of the scripts.

7.4 DevOps Automation with Puppet

ECX is built on top of a RESTful framework through which all actions and queries can be started by using REST commands. This configuration means that developers can continue to use the tools they prefer, such as Chef, Puppet, and Ansible. These automation tools can call on ECX templates as defined by the storage administrator. When that API is called, ECX

orchestrates all of the work necessary to create application consistent copies in a protected environment. This process allows the application developer to get work done faster and more effectively by using near real-time data. When used with a copy data management platform like ECX, developers can define workflows that provision complete environments through tools like Puppet and Chef, and use copies of those environments through ECX.

7.4.1 Example of automating Test/Dev environment by using Puppet and ECX

This example shows how to use Puppet with ECX workflows to configure a test virtual machine or multiple virtual machines. This example does not go through the installation of the Puppet master or agent.

Operations in ECX can be performed by making REST API calls. The example involves a Python script called **recovervm.py.** This script automatically creates a Use workflow within ECX by taking the name of the VM to be restored along with the designated vCenter server the VM belongs to and recovers the VM from the latest snapshot copy cataloged by ECX. Additional parameters such as the ECX appliance IP address, user name, and password are also specified at the command line.

Puppet uses the following terms:

- Master /Server: Central configuration server.
- Agent: The client. In this example, the Puppet agent is deployed in the ECX appliance.
- Modules: Bundles that contain code and data files.
- Manifests: Contains the code in a module.

To automate a Test/Dev environment using Puppet and ECX, complete these steps:

1. Create a module to keep the ECX manifests by using the Puppet module generate**.** Example 7-1 shows the creation of the ECX module. A series of questions are displayed when you run the command. Leave them blank and press Enter.

   Example 7-1 Create ECX Puppet module

   ```
   puppet module generate ecx-module
   ```

2. Set the node in the Puppet master where the commands are to be run and include the specific module in the `site.pp` file. The content of the file is shown in Example 7-2.

   Example 7-2 Contents of Puppet manifest site.pp

   ```
   #etc/puppet/manifests/site.pp
   node ecx231 {
   include ecx-recovervm
   }
   ```

3. Call the **recovervm** script in the ECX module in the `init.pp` manifest, as shown in Example 7-3.

   Example 7-3 Contents of init.pp file to invoke recovervm script

   ```
   #/etc/puppet/modules/ecx-recovervm/manifests/init.pp
   class ecx-recovervm {

   exec { 'run_my_script':
   command => '/usr/bin/python /opt/ECX/tools/scripts/recovervm.py 172.19.15.85
   admin catalOgic psvcenter1 WebServer_6',
      path  => '/usr/bin:/bin:/usr/sbin:/sbin:/opt/ECX/tools/scripts',
   ```

```
   logoutput => true,
   returns => ['0','1','2'],
        }
}
```

The syntax of the script is shown in Example 7-4.

Example 7-4 Script syntax

```
python recovervm.py <ecx ipaddress> <user> <password> <vcenter> <vmname>.
```

Note: Variables can be passed in a separate parameters file. For sake of simplicity, in Example 7-3 the variables are included in the command itself.

4. In the agent, run the Python agent command as shown in Example 7-5 to start the recovery operation.

Example 7-5 Running the python agent command

```
puppet agent --test
```

5. You should see the script perform its operation as shown in Example 7-6.

Example 7-6 Manifest initiation from Puppet agent

```
[root@ecx231 ~]# puppet agent --test
/usr/lib/ruby/site_ruby/1.8/puppet/defaults.rb:214: warning: Insecure world
writable dir /opt/ECX/tools/scripts in PATH, mode 040777
Info: Retrieving pluginfacts
Info: Retrieving plugin
Info: Loading facts
Info: Caching catalog for ecx231
Info: Applying configuration version '1452415000'
Notice: /Stage[main]/Ecx-recovervm/Exec[run_my_script]/returns: getsessionid:
Notice: /Stage[main]/Ecx-recovervm/Exec[run_my_script]/returns: Searching for
VM Recovery Points
Notice: /Stage[main]/Ecx-recovervm/Exec[run_my_script]/returns: creating policy
Notice: /Stage[main]/Ecx-recovervm/Exec[run_my_script]/returns: Calling POST
url = https://172.19.15.85:8443/api/endeavour/policy
Notice: /Stage[main]/Ecx-recovervm/Exec[run_my_script]/returns: Created Use
Data Policy with Id=1021
Notice: /Stage[main]/Ecx-recovervm/Exec[run_my_script]/returns: Calling POST
url = https://172.19.15.85:8443/api/endeavour/job
Notice: /Stage[main]/Ecx-recovervm/Exec[run_my_script]/returns: Created Job
with Id=1021
Notice: /Stage[main]/Ecx-recovervm/Exec[run_my_script]/returns: Calling POST
url = https://172.19.15.85:8443/api/endeavour/job/1021?action=start
Notice: /Stage[main]/Ecx-recovervm/Exec[run_my_script]/returns: Started Job
with Id=1021
Notice: /Stage[main]/Ecx-recovervm/Exec[run_my_script]/returns: executed
successfully
Notice: Finished catalog run in 5.77 seconds
```

6. A VMware Use Policy job is created in ECX, as shown in Figure 7-3, which will start a VM specified in the command from a snapshot copy. This process remains pending for development and test activities until the Instant Virtualization is ended through the ECX console or started through another REST API command.

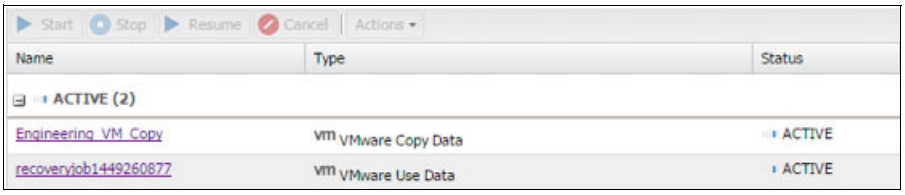


Figure 7-3 Recovery job created in ECX

8

Copy Data Management for a hybrid cloud with Catalogic ECX

This chapter describes how an organization can use Catalogic ECX software with IBM Storage systems built with Spectrum Virtualize software and the SoftLayer® cloud to achieve hybrid cloud use cases. It describes a hybrid cloud automated DR example, compares that example to a traditional DR scenario, and describes the benefits that it can offer.

This chapter includes the following sections:

- Introduction
- Hybrid cloud use case: Automated DR testing

8.1 Introduction

As discussed in 1.1, “Introduction” on page 2, organizations are increasingly exploring new ways to modernize their IT operations. Speeding time to market, increasing reliability, and simplifying operations are long standing IT goals, and the cloud promises to deliver on all of them. However, for many the challenge has been how to automate and orchestrate existing business processes to take advantage of nearly limitless cloud-scale compute power without overloading staff or busting budgets.

Organizations now have a clear path to these goals by deploying Catalogic ECX software with IBM Storage systems built with Spectrum Virtualize software and the SoftLayer cloud. Such a deployment can offer significant hard and soft dollar savings, and result in these advantages:

- Easier data access
- Improved data availability
- Increased competitive advantage
- Lower capital and operating expenses

Industry trends show that a vast data center transformation towards a hybrid cloud model is inevitable due to the overwhelming promise of reduced costs and increased business agility. Although IT organizations with traditional data center operations will be reluctant to move mission-critical production environments to the cloud too rapidly, business operations that rely on copies of production data are ideal for a hybrid cloud model.

The acute pain of today’s copy data management challenges and the promise of the hybrid cloud help justify modifying data center operations to incorporate hybrid cloud environments to drive a more agile organization.

8.2 Hybrid cloud use case: Automated DR testing

Section 4.4, “Automated disaster recovery” on page 65 describes how ECX can be used, along with IBM storage, to achieve automated DR testing. By including SoftLayer into the configuration, it is not only possible to do automated DR testing in the cloud. It is amazingly simple. By combining SoftLayer with Catalogic ECX and IBM Storage, it becomes easy to remove the operational barriers that face IT organizations so that they can easily get the benefits of a hybrid cloud infrastructure.

8.2.1 Using Copy Data Management and snapshots for the hybrid cloud

Using application consistent array snapshots and mirrors in the hybrid cloud provides a superior disaster recovery solution. Combining IBM FlashCopy snapshots with Global Mirror replication capabilities, the Storwize platform lets you quickly and easily move data in and out of different locations, which is essential to get greater value from a hybrid cloud strategy. By automating and orchestrating FlashCopy and Global Mirror functions, ECX replaces the need for scripting and the mix of various tools that are currently required to take advantage of the hybrid cloud. With a few clicks, ECX enables policy-driven workflows to be built on top of the IBM and VMware APIs that use the hybrid cloud.

ECX workflows define these snapshot settings:

- Which snap copies are used
- Where they are located
- How often they are refreshed

- Who has access to them
- How long they are left in place (retention)

Snapshots can be mounted to servers, promoted to production, and properly torn down to minimize any orphaned or forgotten resources that were being used temporarily. These features help lower operational costs and prevent snapshot duplication, which in turn reduces complexity, management effort, and costs in a hybrid cloud environment.

8.2.2 Example configuration

The example production environment that was configured for testing included online production applications and data sets that were created by using live application data. By simulating an actual user environment, the results can be extrapolated to demonstrate the value of the solution in a range of real IT environments and use cases.

The simulated production environment was housed in a test lab facility in New Jersey, where a number of virtual machines were connected to a Storwize V7000 system. The complete environment consisted of over 100 virtual machines connected to 20 TB of production storage. Testing ran over the course of 60 days to make sure that the results were consistent over that time. The test could have been completed in a day.

Copy and Use Data policies for the VMs were created within Catalogic ECX. Daily replication of the VM data was performed by using IBM Global Mirror between the New Jersey based Storwize V7000 and a target array in the San Jose data center of Equinix, a collocation facility that allows customers to deploy their own infrastructure within its data center. For more information about configuring Copy and Use Data policies in ECX, see 3.3, “ECX Copy workflow examples” on page 40 and 4.4.1, “Automated DR example” on page 66.

The neighboring SoftLayer data center, physically next to the Equinix data center, served as the location for the VMware ECX hybrid cloud environment that was used to instantiate the replicated VMs to spin up and test the DR environment.

A diagram of the implementation is shown in Figure 8-1.

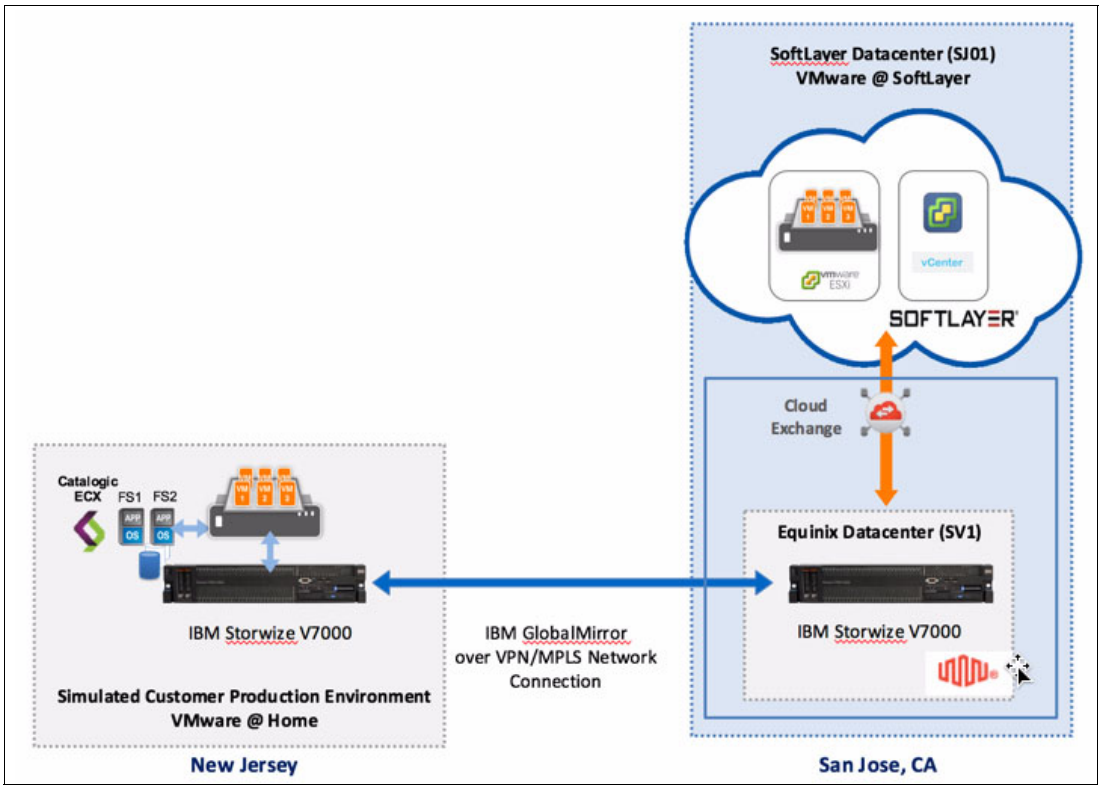


Figure 8-1 Example hybrid cloud DR environment using SoftLayer and Equinix

The detailed setup had these characteristics:

- Simulated customer production environment (New Jersey):
  - VMware ESX 5.5 with 4 cores and 64 GB of RAM
  - IBM Storwize V7000
  - Catalogic ECX, deployed as a virtual appliance
  - Application VMs hosted on Storwize V7000:
    - Microsoft Exchange 2010
    - Oracle database server
    - Microsoft SQL database server
    - Microsoft Active Directory server
- Collocation site (Equinix, San Jose, CA):
  - IBM Storwize V7000
- SoftLayer (San Jose, CA):
  - VMware ESX 5.5 with 4 cores and 32 GB of RAM

In addition to establishing policies for automated replication on a per-VM basis, the testing also included the automated use of the replicated VM data in a simulated DR test scenario. Testing included full validation of the ability to instantiate the production applications within the SoftLayer/Equinix environment. For more information about such automated ECX Use policies, see 4.4.1, “Automated DR example” on page 66.

8.2.3 Traditional DR testing versus automated DR testing in a hybrid cloud

This section compares traditional and automated DR testing.

Traditional DR testing

To put a typical disaster recovery workflow that doesn't use ECX into more perspective, you must complete the following steps:

1. Pre-synchronize the storage resources between the production and recovery sites.
2. Shut down the virtual machines at the production site and prepare them for migration to the recovery site.
3. Synchronize storage resources between production and recovery sites.
4. Suspend non-critical virtual machines at the recovery site to make room for migrated virtual machines from the production site.
5. Change storage access at the recovery site to "writeable."
6. Power on the virtual machines at the recovery site (these virtual machines can be powered on in a pre-determined priority order if necessary)

For a customer of similar size to the test environment, DR testing is a process that typically takes a full weekend and is administered using many scripts and runbook procedures, both of which must be continually updated. Recovery tests often fail, and significant IT work is needed to diagnose and correct the problem.

IT teams know that testing more frequently leads to greater success, but because DR testing requires such a monumental effort, identifying the best way to tackle "how" to make that happen is a challenge. Moreover, the overall DR operation can be one of the most costly responsibilities for IT and typically requires a second data center. With IT budgets constantly under pressure, painful trade-offs are often made that negatively influence the overall effectiveness of the DR strategy.

Automated DR testing in a hybrid cloud

Having the ability to quickly iterate through challenges and fix them as they come up is key to ensuring a successful disaster recovery environment. By testing in SoftLayer and using IBM storage services, IT can test, fix, and test again, iterating through the process quickly as they pay only for the time that the services are running. Clients can streamline the DR testing process and perform it daily, something that would be unheard of in the traditional style of DR testing.

Modern DR testing can also use the different geographic locations of the SoftLayer infrastructure. By using different locations, an organization can avoid being affected by large, regional events such as hurricanes. This configuration in turn improves the effectiveness of the company's disaster recovery strategy.

8.2.4 Results summary

The testing and the results validate the ability to simply and predictably deliver an automated DR process in the SoftLayer IaaS cloud infrastructure. Using the results from this testing and contrasting them to a traditional DR approach, such as one that would test off-site DR readiness twice per year, this use case can save as many as 200 staff hours.

The test results for the example environment show how Catalogic ECX was able to automate and orchestrate an instantiation of application consistent Mirrors in SoftLayer, taking full

advantage of IBM for the automated DR use case. DR testing used a subset of the simulated client environment. This testing focused on seven VMs across a range of applications:

- Five VMs with their data stored on the IBM V7000 system:
  - Three application servers: Exchange, Oracle, and SQL
  - Two IT admin servers: Active Directory and a web server
- Two VMs were IT administrator servers that stored their data on non-IBM storage. For these servers, ECX orchestrated the creation of VM image copies and replicated them to the IBM platform through the vCenter API as described 8.2.2, “Example configuration” on page 131.

With the simulated production environment up and running, establishing DR connectivity to the Equinix and SoftLayer environments was quick and easy. The first VM replication jobs began running within two hours of the beginning of our effort to establish the hybrid cloud environment. The more significant time savings came from ECX’s automated process of instantiating VMs at the DR site and validating their readiness. When compared to the traditional “brute force” method of testing DR, the operational savings were significant. For a customer similar in size to the simulated environment, the traditional process would typically require upwards of 200 staff hours per year. This estimate assumes a DR test effort twice per year, with multiple people spending the weekend planning and running each DR test.

In contrast, the Catalogic ECX-powered hybrid cloud solution took one person less than an hour to run. The tests therefore can validated DR readiness daily.

8.2.5 Financial impact

In addition to the operational expense savings from the significant reduction in labor to plan and manage the DR testing, a hybrid cloud automated DR testing deployment can also deliver significant financial savings. For a detailed financial analysis that compares the CapEx savings of a traditional DR approach to one that is powered by the joint SoftLayer-Catalogic-IBM solution, including a detailed total cost of ownership (TCO) financial analysis, see *Copy Data Management for a Hybrid Cloud* at:

http://data.catalogicsoftware.com/hubfs/Automated_Disaster_Recovery_to_Softlayer.pdf

8.2.6 Alternative hybrid cloud use cases

The example use case shows the significant benefits that can be realized for automated disaster recovery when using Catalogic ECX along with IBM Storage systems built with Spectrum Virtualize software and SoftLayer in a hybrid cloud deployment. It is important to note that the advantages of such a deployment model are not limited to just this automated DR scenario. Similar benefits can be achieved with such a model for many other use cases. For example, in the DevOps scenario described in 7.4.1, “Example of automating Test/Dev environment by using Puppet and ECX” on page 125, test and development resources could be spun up with API calls in SoftLayer as easily as they were in a private data center. ECX offers an enhanced level of flexibility to almost any use case. It allowing organizations to spin up near real-time copies of production data in any environment, regardless of whether it is “on premises” or in the cloud.

8.2.7 Conclusion

By using a combination of SoftLayer, IBM and Catalogic, IT has a new, more powerful way of harnessing the value of the hybrid cloud, without adding complexity or requiring additional

expertise. In fact, this solution dramatically simplifies operations and reduces the required labor to manage the flow of data to and from the cloud. The ECX copy data management software platform, with IBM SVC/Storwize and SoftLayer, can automate and orchestrate all of the IT organization's key operations that rely on copy data. These operations include automated DR, test/dev, and development and business analytics. The automation of a DR solution provides clients with these advantages:

- The ability to take advantage of flexible hybrid cloud resources
- An automated process for getting the correct data to the cloud, saving hundreds of IT hours per year in DR validation testing
- The ability to move data into the cloud quickly, overcoming one of the primary obstacles of cloud adoption today
- The ability to orchestrate data copies in the cloud for multiple use cases beyond automated DR testing
- Visibility into what data is in the cloud
- The operational control of copy data across the organization, including the cloud, thus ensuring that the IT team is only using (and paying for) the resources that are needed

The impact of Copy Data Management in the hybrid cloud is transformational, and as demonstrated, implementation is quite simple. Having a Copy Data Management solution such as ECX is the best way to understand where snapped and mirrored data lives, allowing IT to better use those copies for business operations like automated DR. IBM SVC/Storwize allows IT to use the same processes that they use today to create a seamless, secure process for migrating data into the cloud. SoftLayer provides on-demand ECX compute resources, and eliminates the need for the client to own a physical DR environment.

As IT looks for solutions to reduce operational and capital costs, while improving its ability to deliver mission critical services to the business, use of the hybrid cloud holds tremendous potential. This chapte has shown that using IBM SVC/Storwize and Catalogic ECX with SoftLayer overcomes many of the barriers to cloud adoption today, allowing IT to provide simplified access to copy data for multiple business operations. Doing so saves significant spending while delivering superior value to business.

Related publications

The publications listed in this section are considered particularly suitable for a more detailed discussion of the topics covered in this book.

IBM Redbooks

The following IBM Redbooks publications provide additional information about the topic in this document. Note that some publications referenced in this list might be available in softcopy only.

- *IBM FlashSystem V9000 Product Guide*, REDP-5317
- *IBM FlashSystem V9000 in a VersaStack Environment*, REDP-5264
- *IBM FlashSystem V9000 and VMware Best Practices Guide*, REDP-5247
- *Implementing the IBM Storwize V5000*, SG24-8162
- *Implementing the IBM Storwize V7000 V7.4*, SG24-7938
- *Implementing the IBM Storwize V7000 Unified Disk System*, SG24-8010
- *Implementing the IBM System Storage SAN Volume Controller V7.4*, SG24-7933
- *Introducing and Implementing IBM FlashSystem V9000*, SG24-8273
- *VersaStack Solution by Cisco and IBM with IBM DB2, IBM Spectrum Control, and IBM Spectrum Protect*, SG24-8302
- *VersaStack Solution by Cisco and IBM with SQL, Spectrum Control, and Spectrum Protect*, SG24-8301

You can search for, view, download or order these documents and other Redbooks, Redpapers, Web Docs, draft and additional materials, at the following website:

ibm.com/redbooks

Other publications

These publications are also relevant as further information sources:

- Catalogic ECX 2.3 Quick Start Guide (requires registration and login)
- Catalogic ECX 2.3 User's Guide (requires registration and login)

Online resources

These websites are also relevant as further information sources:

- Catalogic Software site

  https://www.catalogicsoftware.com
- Catalogic Software MySupport site

  https://mysupport.catalogicsoftware.com/

- Catalogic Software Knowledgebase

  http://doc.catalogicsoftware.com/kb/index.htm
- VMware Knowledge Base article “Running custom quiescing scripts inside Windows virtual machine with ESX 3.5 Update 2 and later (1006671)”

  http://kb.vmware.com/kb/1006671
- Gartner IT Glossary

  http://www.gartner.com/it-glossary
- IBM Storage home page

  http://www.storage.ibm.com
- SAN Volume Controller supported platform

  http://ibm.co/1FNjddm
- SAN Volume Controller IBM Knowledge Center

  http://www.ibm.com/support/knowledgecenter/STPVGU/welcome

Help from IBM

IBM Support and downloads

ibm.com/support

IBM Global Services

ibm.com/services

Redbooks

IT Modernization using Catalogic ECX Copy Data Management and IBM Spectrum Storage

(0.2”spine)
0.17”<->0.473”
90<->249 pages

SG24-8341-00

ISBN 0738441589

Printed in U.S.A.