IBM
Spectrum
Discover

IBM Spectrum Discover
Metadata Discovery, Insights and Management

Stephen Edel
edel@us.ibm.com

Spectrum Discover Screenshot Overview

IBM Spectrum Scale

TOPICS

Spectrum Discover Overview and Capabilities

How to connect Data sources

Datasource Overview

Viewing different Capacity categories

Metadata Tagging

Creating Policies and Action Agents

Metadata Searches and Results

Reporting

Licensing

PoC/Trial VM and Hardware/OS/Networking Requirements

Spectrum Discover - Metadata Management Software

IBM Spectrum Scale

IBM
Spectrum
Discover

Provides unified metadata management and insights for heterogeneous file and object storage, on-premises and in the cloud.

Discover	Classify	Label	Find
Automatically ingest & index system metadata from multiple file & object storage systems on-prem & in the cloud	Automatically identify and classify data, including sensitive and personally identifiable information	Enrich data with system & custom metadata tags that increase the value of that data	Find data quickly and easily by searching catalogs of system & custom metadata

Spectrum Discover can Increase business value

Analytics

Uncover hidden data value

- Accelerate data identification for large-scale analytics
- Efficiently curate large-scale unstructured data and create custom datasets for AI / ML / Analytics workflows

Governance

Help mitigate risk / improve quality

- Automatically identify certain kinds of PII & sensitive data, and map this data to the right storage location
- Help reduce risk buried in unstructured data stores
- Tag / index data for eDiscovery & legal hold, helping speed up investigations

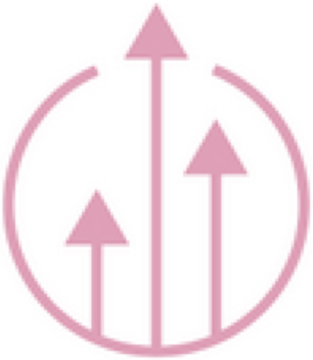
Optimization

Improve storage utilization

- Decrease storage CAPEX by facilitating data movement to colder, cheaper storage
- Increase storage efficiency by eliminating ROT data
- Reduce storage OPEX by improving storage administrator productivity

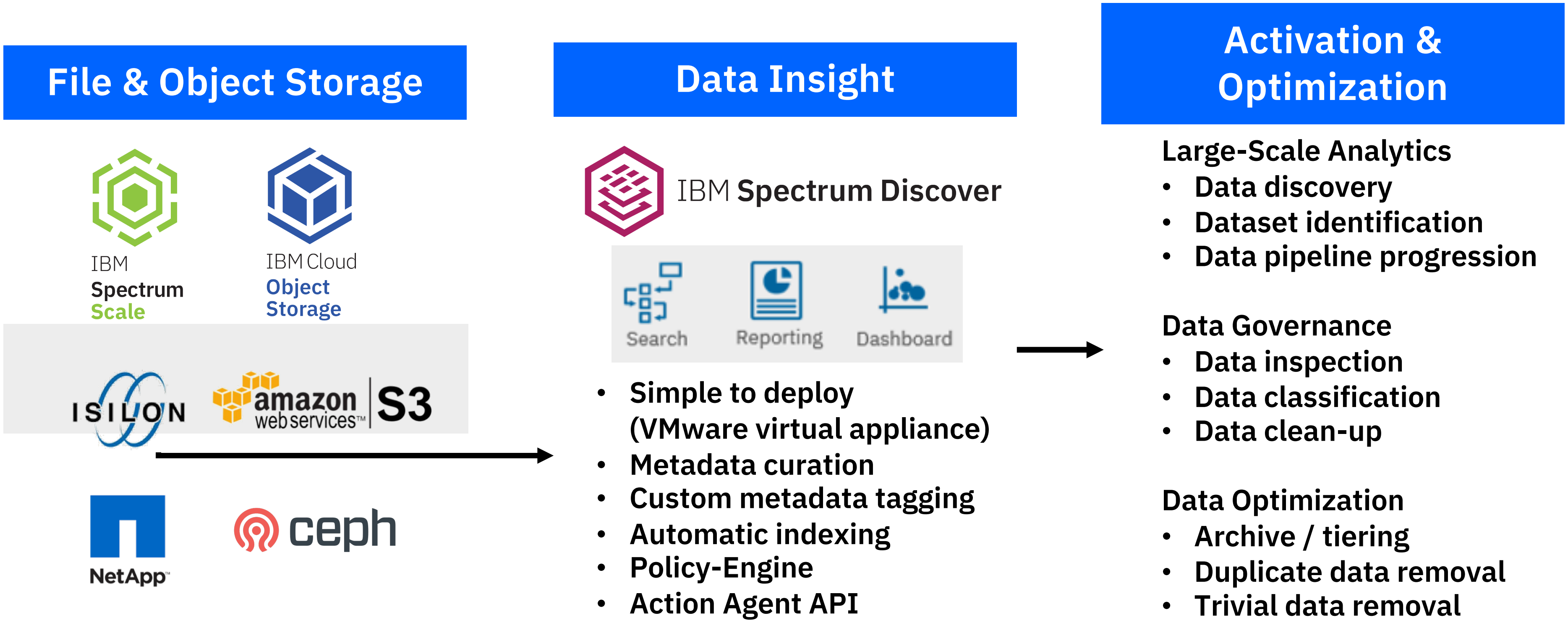
IBM Spectrum Discover Highlights

IBM Spectrum Scale



- ✓ Unified Metadata Solution for IBM COS, Spectrum Scale, and 3rd-party NFS storage:
  - Ingest metadata from Spectrum Scale filesystem(s), Isilon NFSv3 exports, NetApp NFSv3 exports
  - Ingest metadata from Amazon S3 buckets, Ceph S3 buckets
  - Scanning & live event notifications for COS, Spectrum Scale and NFS
  - Flexibility -Custom tagging for data classification on a wide range of parameters in system metadata or with custom metadata
- ✓ Scalable to index & tag exabyte-scale repositories
- ✓ Advanced, intuitive GUI dashboard for critical insights, analytics, & reporting
- ✓ No plug-ins or server agents required for 3rd party data sources

IBM Spectrum Discover Environment



File & Object Storage

Data Insight

- Simple to deploy (VMware virtual appliance)
- Metadata curation
- Custom metadata tagging
- Automatic indexing
- Policy-Engine
- Action Agent API

Activation & Optimization

Large-Scale Analytics

- Data discovery
- Dataset identification
- Data pipeline progression

Data Governance

- Data inspection
- Data classification
- Data clean-up

Data Optimization

- Archive / tiering
- Duplicate data removal
- Trivial data removal

Spectrum Discover is deployed as a virtual appliance in Vmware ESXi 6.0 or later environments.

https://www.ibm.com/support/knowledgecenter/en/SSY8AC_2.0.0/com.ibm.spectrum.discover.v2r00.doc/pln_planning_landing.html

Metadata-Fueled Data Analysis

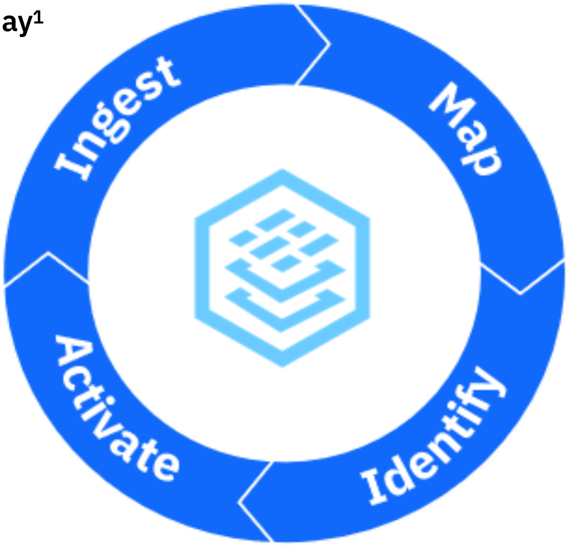
IBM Spectrum Scale

Large Scale Data Ingest

- Scan billions of records per day[1]
- Live event notifications
- Capture system-level tags
- Automatic indexing

Data Visualization

- Query billions of records in seconds
- Multi-faceted search
- Drilldown dashboard
- Customizable reports



Business-Oriented Data Mapping

- Custom data tagging
- Content-inspection via Action Agent API
- Policy-driven workflows

Data Activation

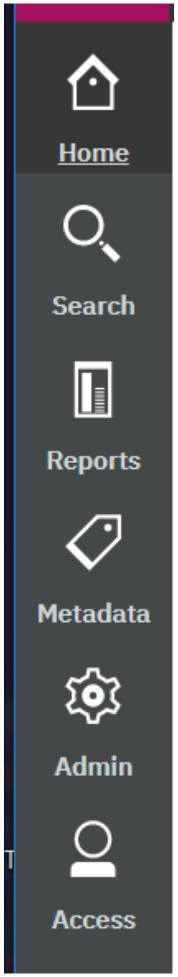
- Data movement via Action Agent API
- Extensible architecture
- Solution blueprints

Types of metadata

- System: information about file and object types, their sizes, when they were last modified, etc.
- Custom: user/organization-defined based on unique taxonomy
- Derived: derived from analytics and applied to your data enriching the metadata model with additional meaning

[1] Up to 30K/sec; ~2.5 billion rec/day w/ IBM Spectrum Scale; ~432 million/day w/ IBM Cloud Object Storage

Exploring the Spectrum Discover GUI



- Overview of Data Sources, monitor storage utilization and data recommendations
- Preview capacity use by data facet. Identify duplicate files

- Search by Datasource, Platform, Temperature, File Size range, Project, Owner, Time since last access

- Generate Reports based on Search Results

- Automatic cataloguing of system-level metadata, create custom metadata field names and tags
- Specify value of your choice, enforce only defined values to be used, or create characteristic tags
- Policies to automate the classification of large-scale data based on both system and customized metadata
- Either AUTOTAG (add custom metadata values to all or a subset of the records based on filter criteria, or DEEPINSPECT policy, allows you to enrich metadata through content inspection of source data

- Capacity used by Project, Datasource, Owner, Filesize, Last Access
- Adding and Viewing Data Source Connections

- Create Users, Groups, Collections

IBM Spectrum Discover

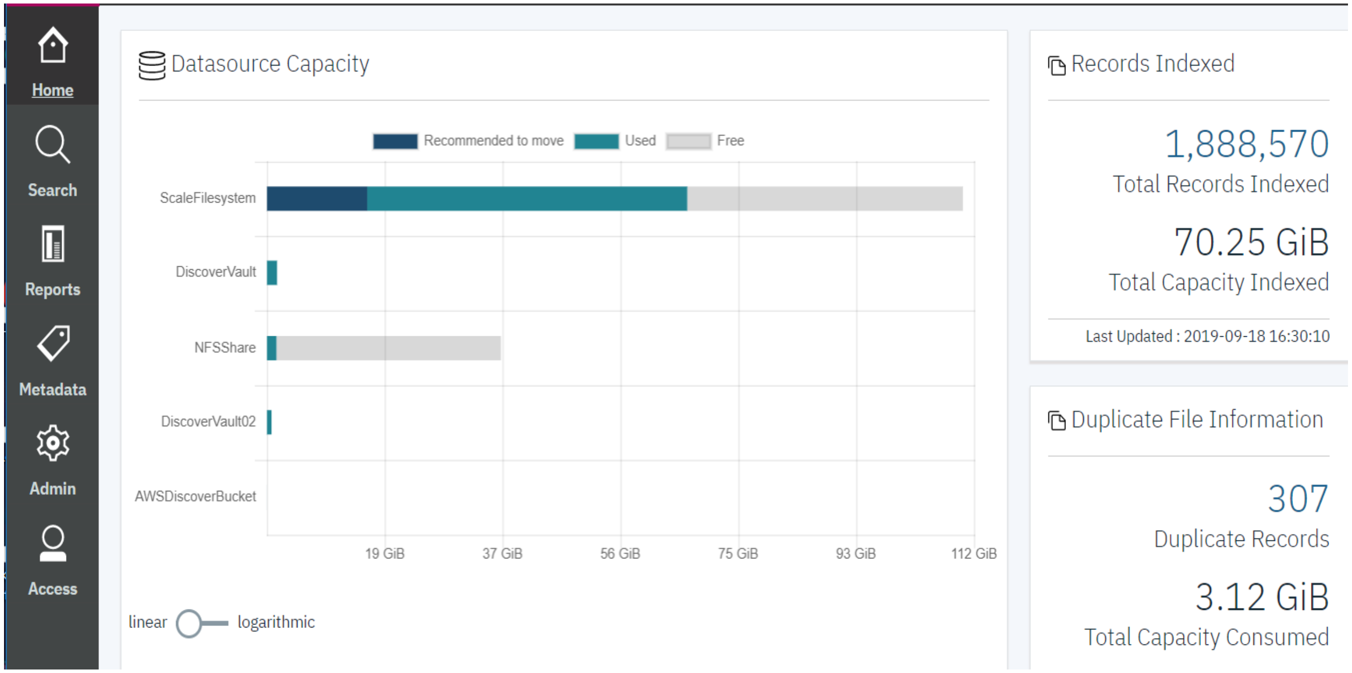
Spectrum Discover Dashboard

IBM Spectrum Scale

Monitor storage utilization and data recommendations (Move/Archive)

Preview capacity use by data facet
- Classification
- Owner
- File Type
- Etc.

Mouse over elements of the Data Source Capacity to view capacity used by each data source



1,888,570
Total Records Indexed

70.25 GiB
Total Capacity Indexed

Last Updated : 2019-09-18 16:30:10

307
Duplicate Records

3.12 GiB
Total Capacity Consumed

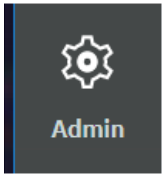
Total indexed data and capacity

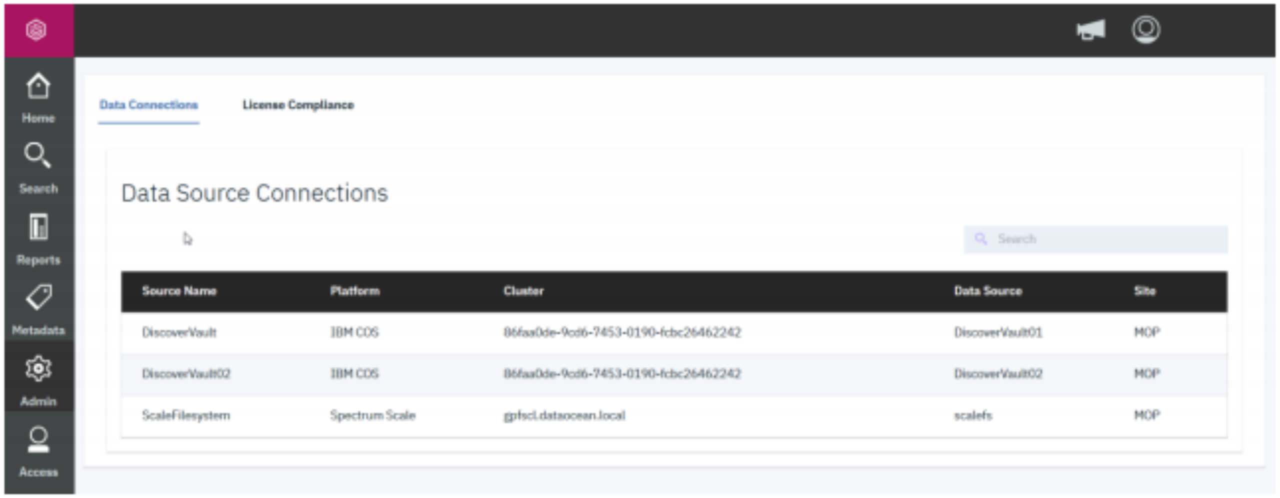
Duplicate file or object candidates
- Number
- Capacity used

Data capacity by group/collection
- Customer defined
- Lab/Project/etc.

Viewing & Adding Data Source Connections

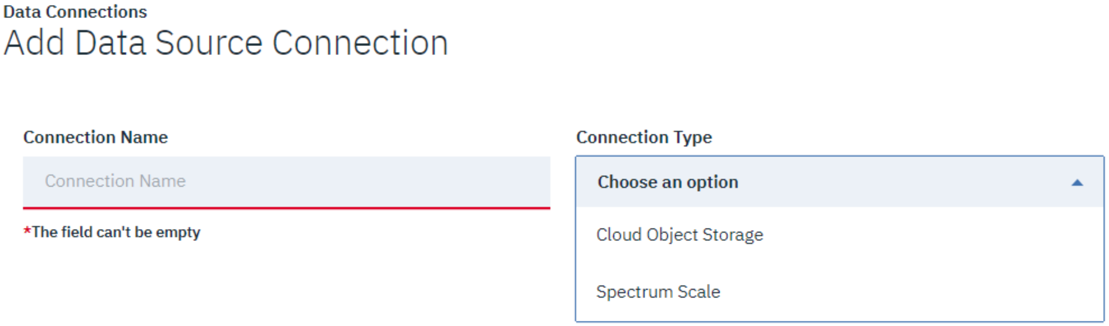
IBM Spectrum Scale





Data Source Connections

Source Name	Platform	Cluster	Data Source	Site
DiscoverVault	IBM COS	86faa0de-9cd6-7453-0190-fcbc26462242	DiscoverVault01	MOP
DiscoverVault02	IBM COS	86faa0de-9cd6-7453-0190-fcbc26462242	DiscoverVault02	MOP
ScaleFilesystem	Spectrum Scale	gpfscl.dataocean.local	scalefs	MOP



Data Connections

Add Data Source Connection

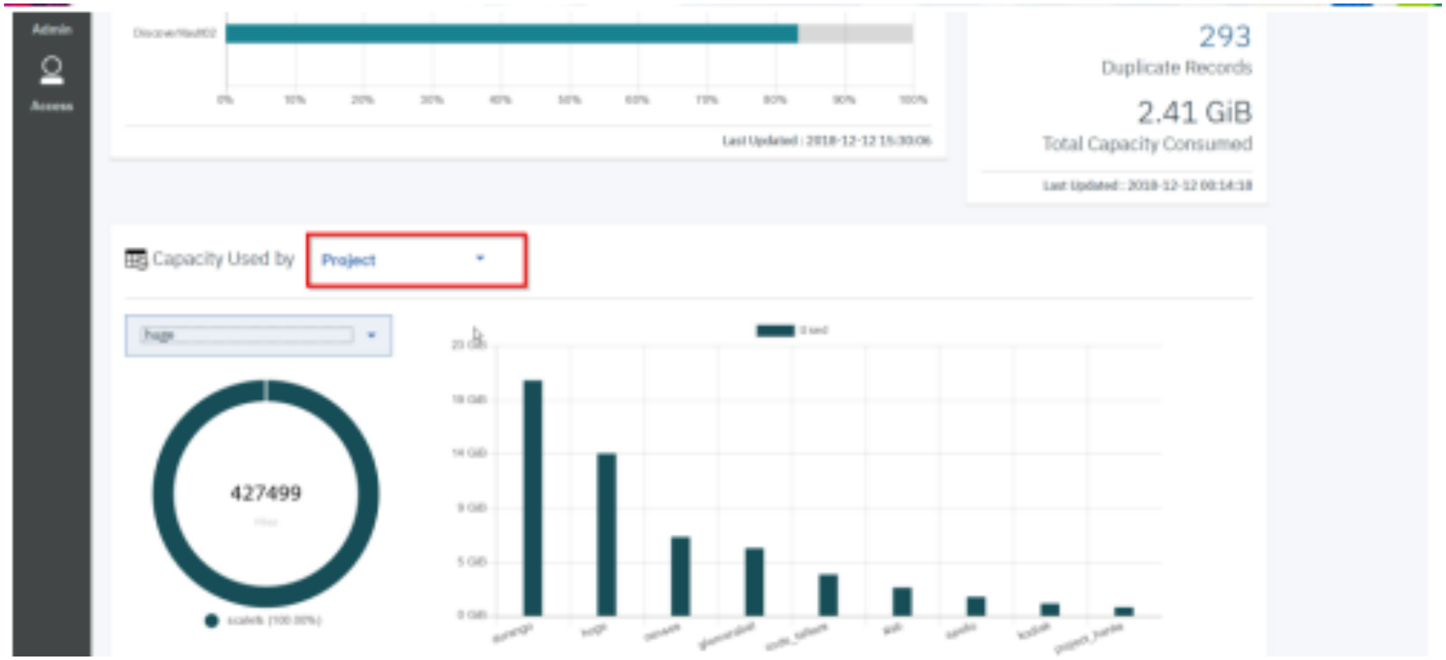
Connection Name

*The field can't be empty

Connection Type

Capacity Used by Project

- For example, users can view utilization according to a project custom metadata tag across the data sources that was set leveraging the Spectrum Discover policy engine.



293 Duplicate Records

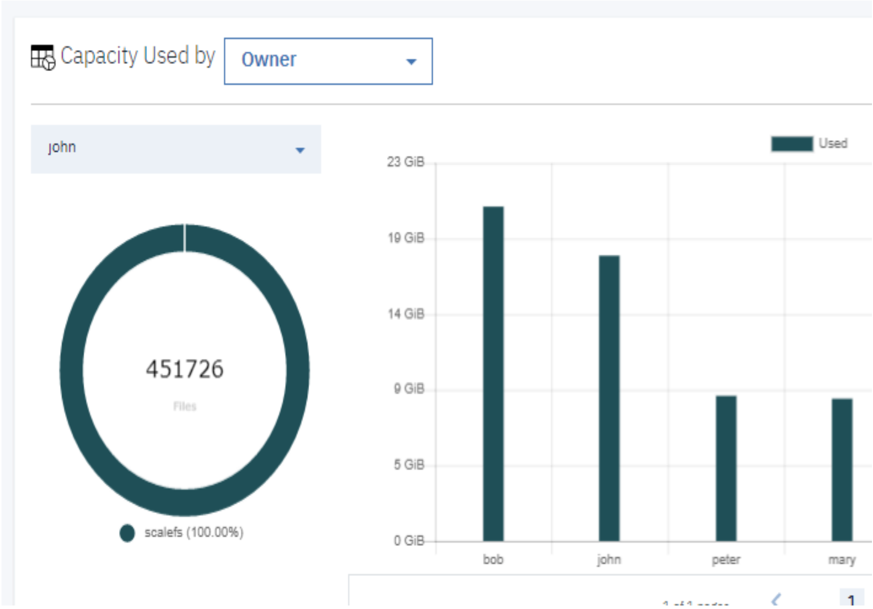
2.41 GiB Total Capacity Consumed

Last Updated : 2018-12-12 00:14:18

Capacity Used by: Project

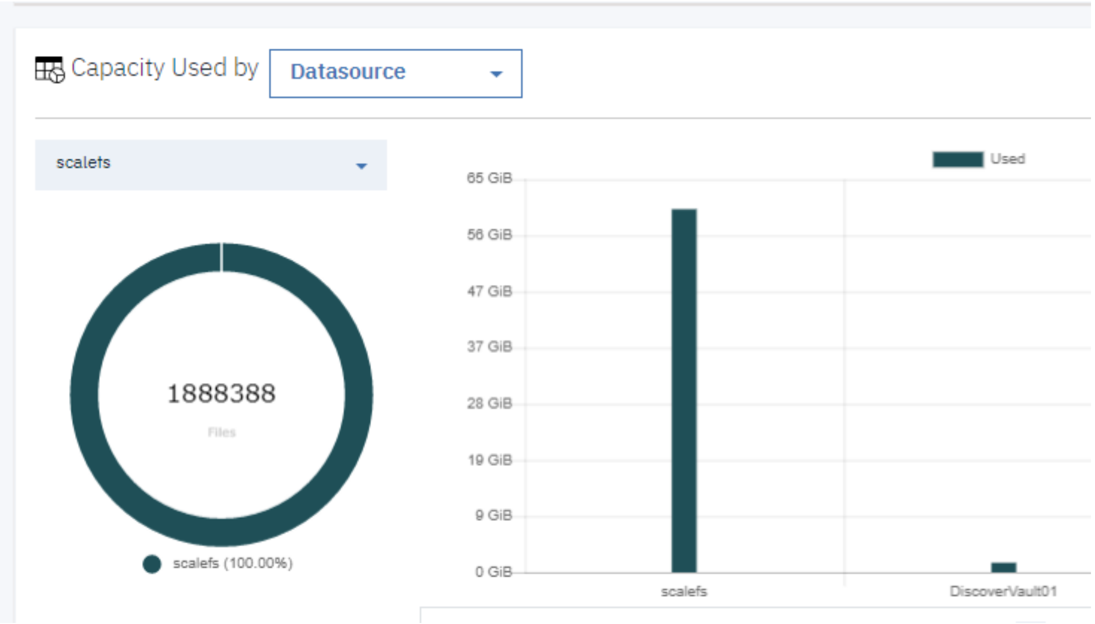
Various Projects

Displaying Capacity by Owner & Data Source



Capacity Used by Owner

john

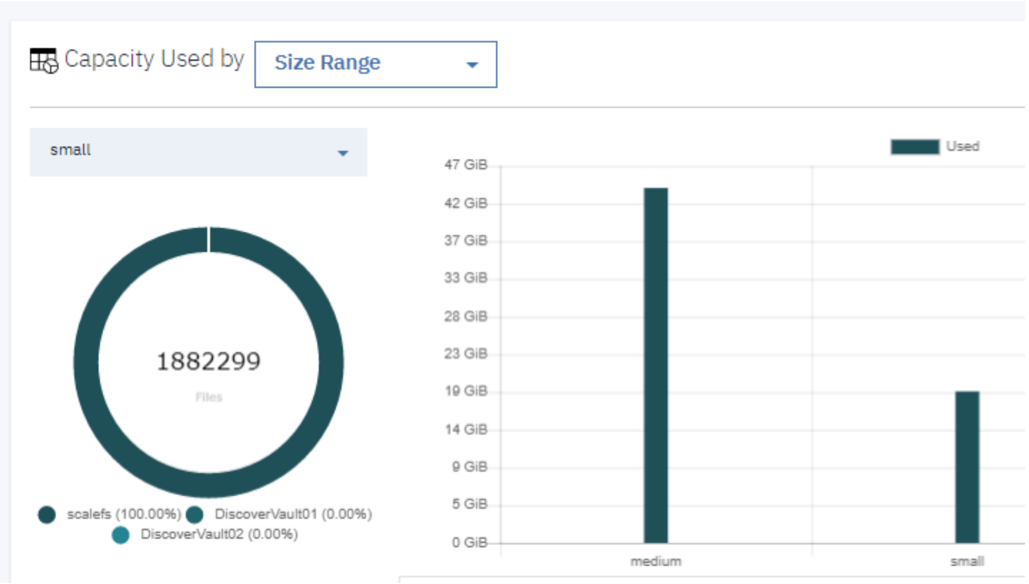


Capacity Used by Datasource

scalefs

Displaying Capacity by File Size Range and Temperature

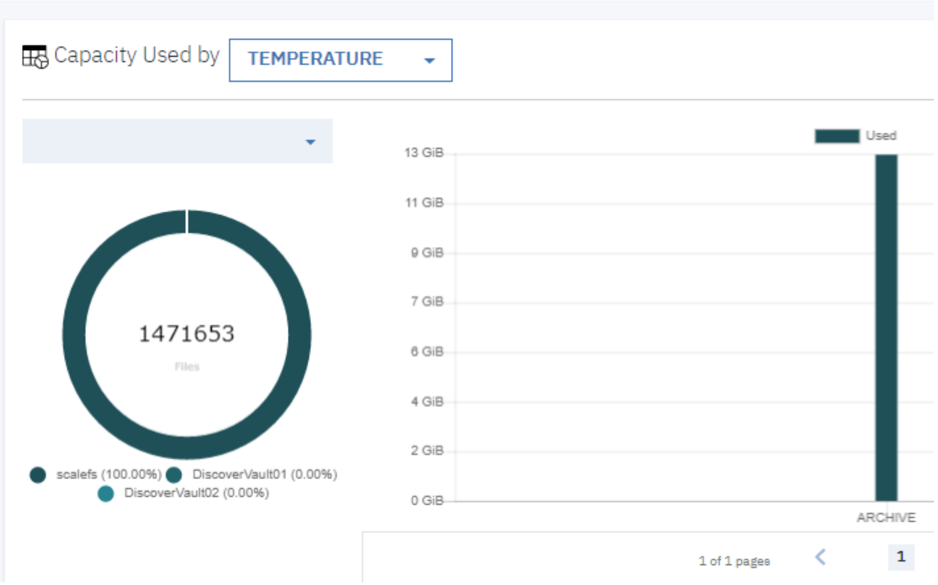
IBM Spectrum Scale



Capacity Used by: Size Range

small

1882299 Files

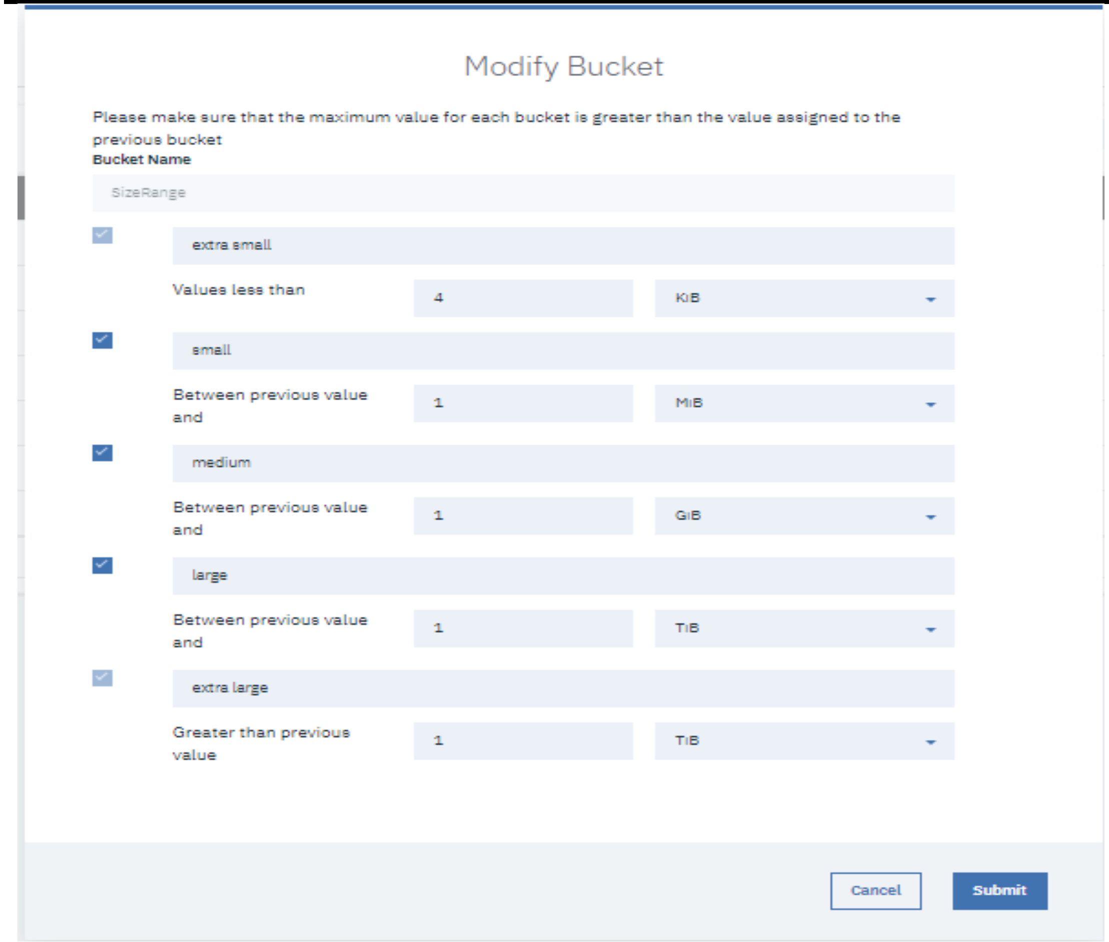


Capacity Used by: TEMPERATURE

1471653 Files

Setting File Size Range

IBM Spectrum Scale



Modify Bucket

Please make sure that the maximum value for each bucket is greater than the value assigned to the previous bucket

Bucket Name

SizeRange

☑ extra small

Values less than 4 KiB

☑ small

Between previous value and 1 MiB

☑ medium

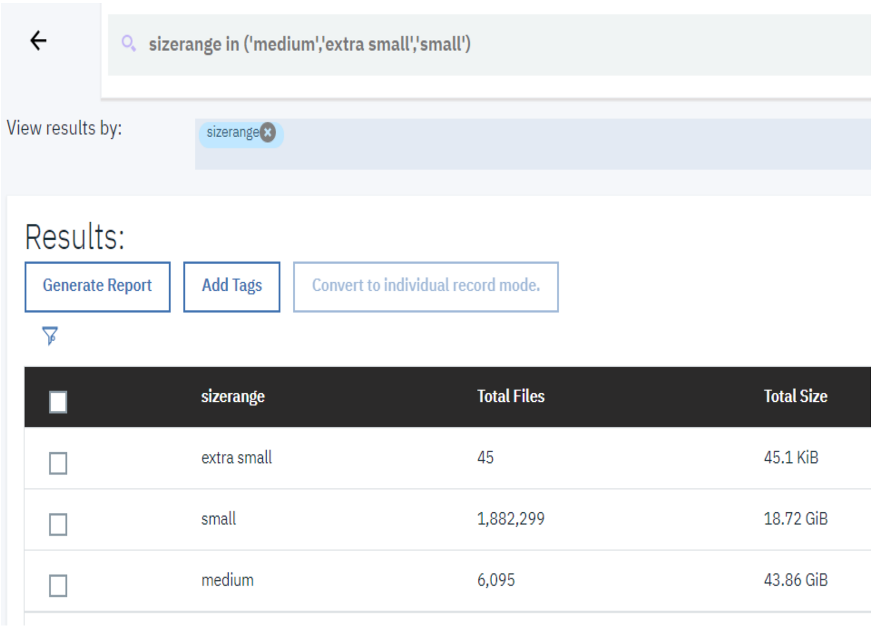
Between previous value and 1 GiB

☑ large

Between previous value and 1 TiB

☑ extra large

Greater than previous value 1 TiB



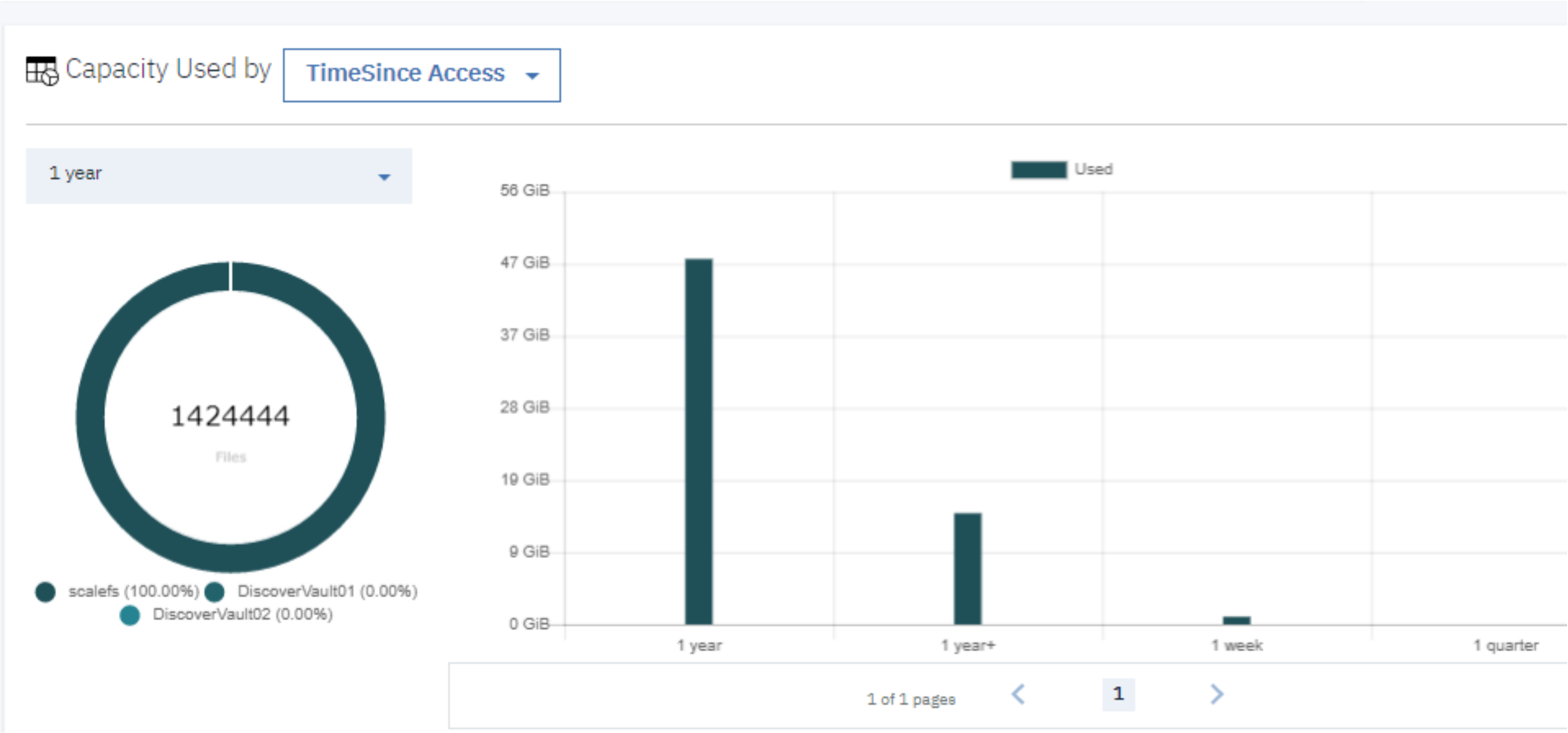
sizerange in ('medium','extra small','small')

View results by: sizerange

Results:

□	sizerange	Total Files	Total Size
□	extra small	45	45.1 KiB
□	small	1,882,299	18.72 GiB
□	medium	6,095	43.86 GiB

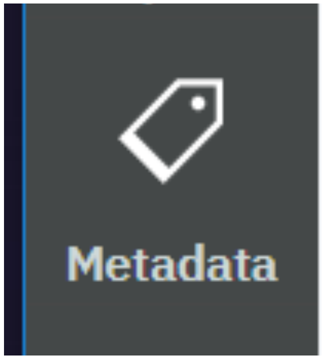
Capacity Ranked by Last Access Time



Capacity Used by TimeSince Access

1 year

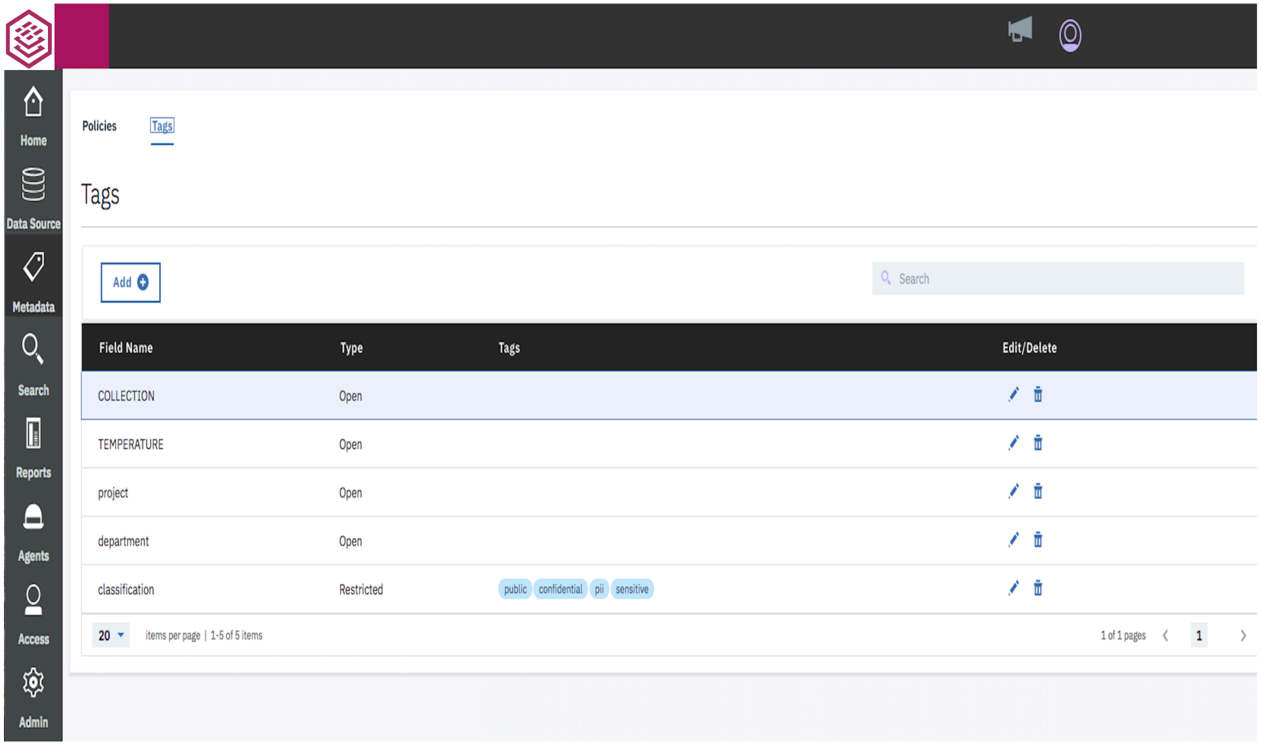
Metadata Tagging



Customize Metadata – Defined through Tags

IBM Spectrum Scale

- Metadata > Tags page:
- Create custom metadata field names and tags
  - Unique to organizational schema/ taxonomy
  - Manual and/or via API for auto insertion
  - Enables organizations to describe data with more meaningful tags
- Metadata tags can be Open, Restricted, or Characteristic in nature
- Open tags allow user to specify value of their choice
- Restricted tags enforce only defined values to be used
- Characteristic tags, which may be used to extract meaning from the data itself.
  - (For example, a characteristic tag that extracts the estimated value or other information from the documents themselves.)



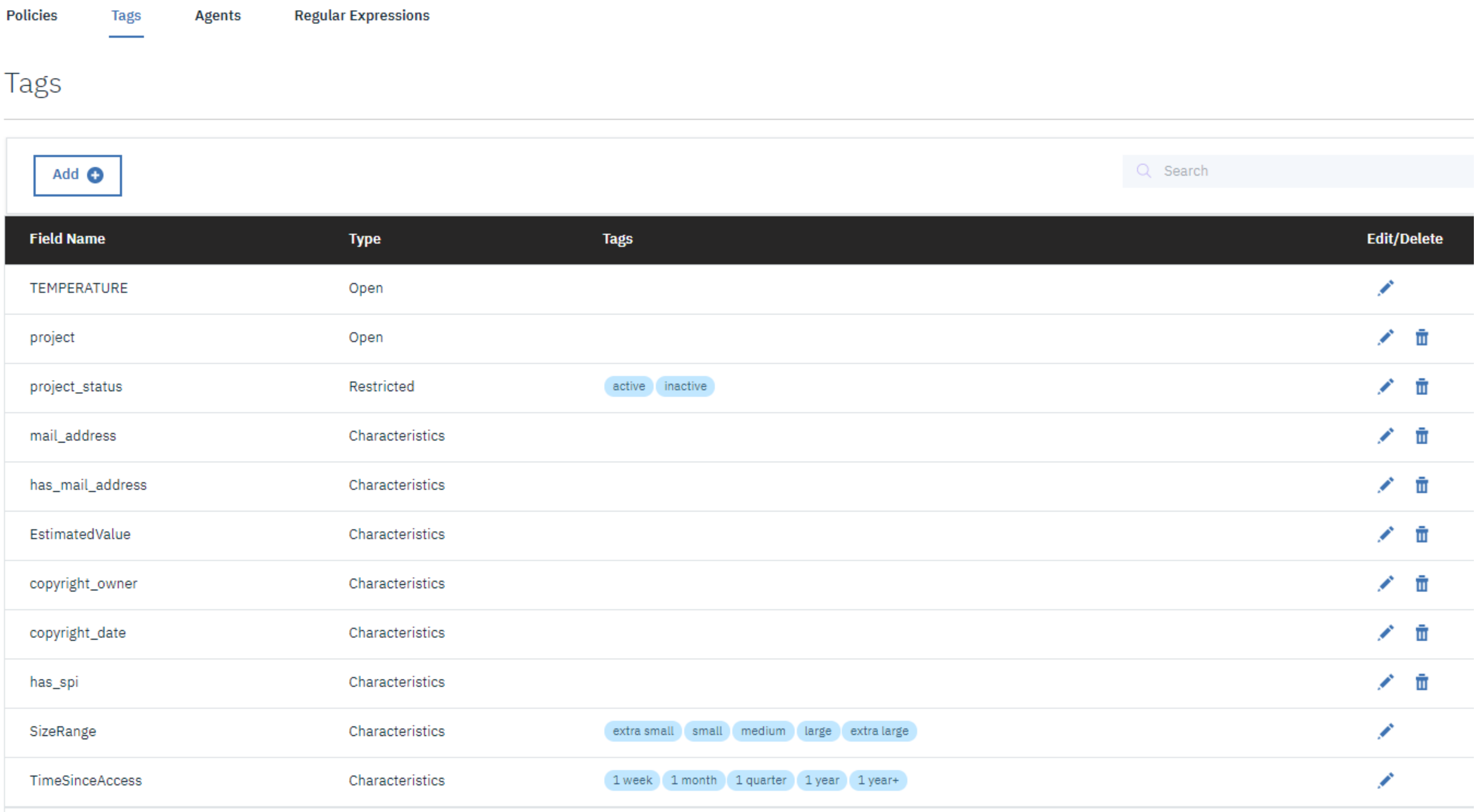
Tags

Field Name	Type	Tags	Edit/Delete
COLLECTION	Open		
TEMPERATURE	Open		
project	Open		
department	Open		
classification	Restricted	public confidential pii sensitive	

20 items per page | 1-5 of 5 items 1 of 1 pages 1

Tag Type Examples

IBM Spectrum Scale



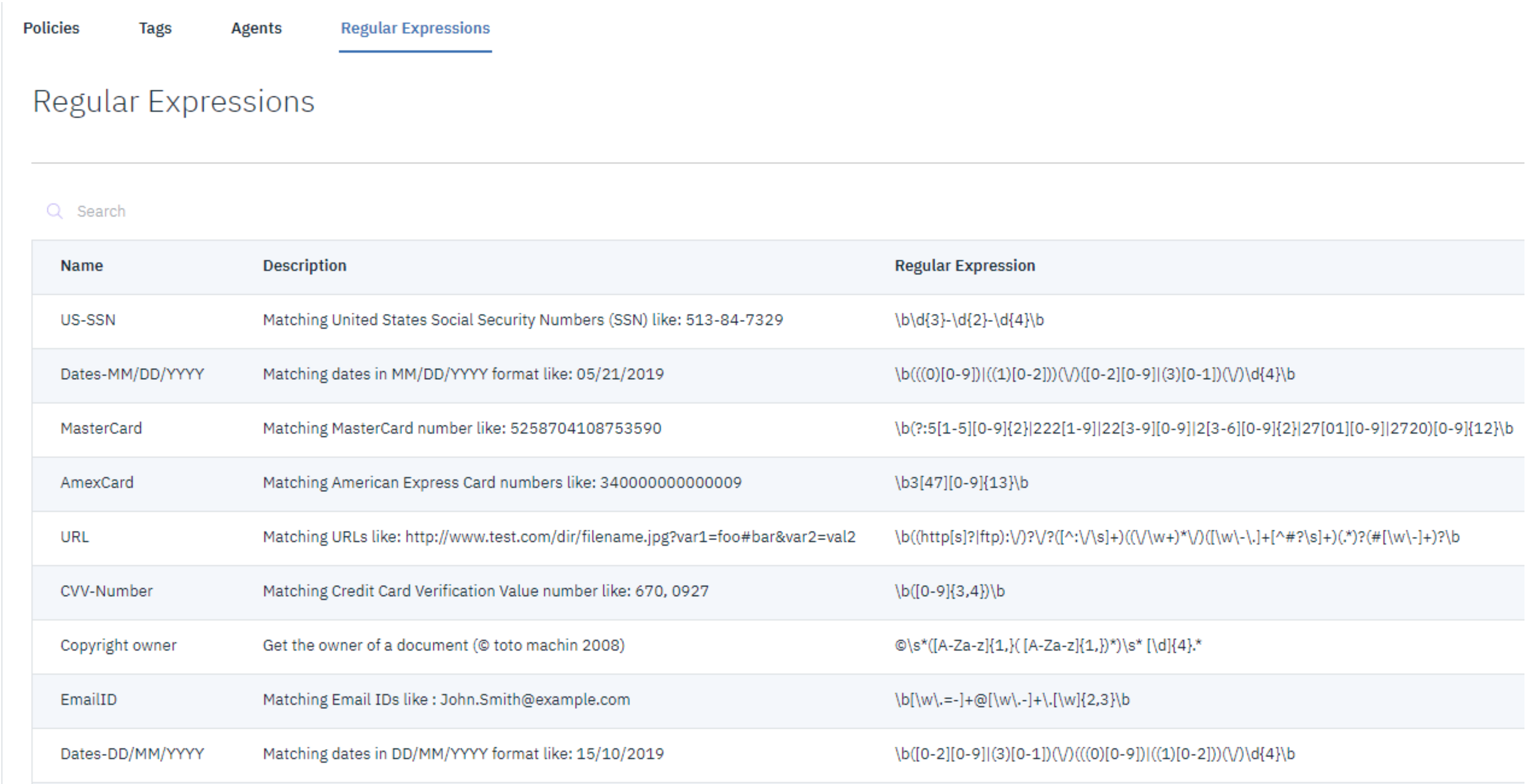
Field Name	Type	Tags	Edit/Delete
TEMPERATURE	Open		✎
project	Open		✎ 🗑
project_status	Restricted	active inactive	✎ 🗑
mail_address	Characteristics		✎ 🗑
has_mail_address	Characteristics		✎ 🗑
EstimatedValue	Characteristics		✎ 🗑
copyright_owner	Characteristics		✎ 🗑
copyright_date	Characteristics		✎ 🗑
has_spi	Characteristics		✎ 🗑
SizeRange	Characteristics	extra small small medium large extra large	✎
TimeSinceAccess	Characteristics	1 week 1 month 1 quarter 1 year 1 year+	✎

Note flexibility to
Tag by Last Access
and File Size

->

Expressions

IBM Spectrum Scale



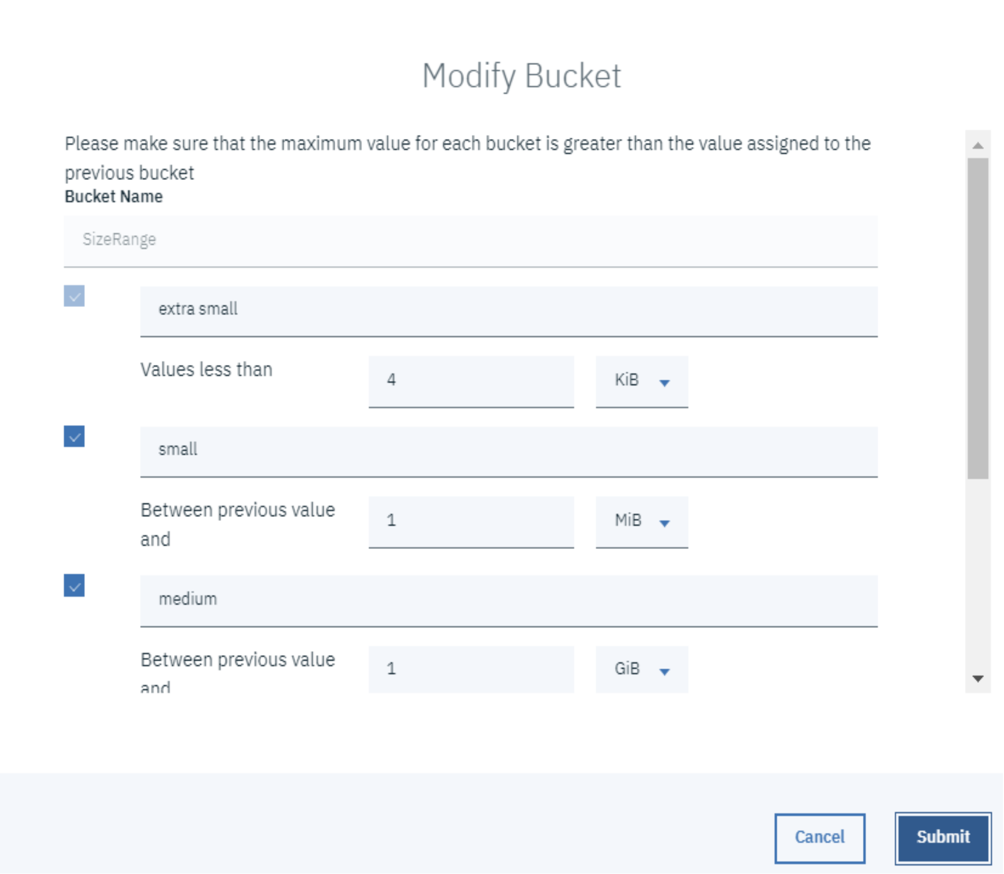
Regular Expressions

Name	Description	Regular Expression
US-SSN	Matching United States Social Security Numbers (SSN) like: 513-84-7329	\b\d{3}-\d{2}-\d{4}\b
Dates-MM/DD/YYYY	Matching dates in MM/DD/YYYY format like: 05/21/2019	\b(((0)[0-9])|((1)[0-2]))(\/)([0-2][0-9]|(3)[0-1])(\/)\d{4}\b
MasterCard	Matching MasterCard number like: 5258704108753590	\b(?:5[1-5][0-9]{2}|222[1-9]|22[3-9][0-9]|2[3-6][0-9]{2}|27[01][0-9]|2720)[0-9]{12}\b
AmexCard	Matching American Express Card numbers like: 340000000000009	\b3[47][0-9]{13}\b
URL	Matching URLs like: http://www.test.com/dir/filename.jpg?var1=foo#bar&var2=val2	\b((http[s]?|ftp):\/)?\/?([^:\/\s]+)((\/\w+)*\/)([\w\-\.]+[^#?\s]+)(.*)?(#[\w\-]+)?\b
CVV-Number	Matching Credit Card Verification Value number like: 670, 0927	\b([0-9]{3,4})\b
Copyright owner	Get the owner of a document (© toto machin 2008)	©\s*([A-Za-z]{1,}([A-Za-z]{1,})*)\s* [\d]{4}.*
EmailID	Matching Email IDs like : John.Smith@example.com	\b[\w\.=-]+@[\w\.-]+\.[\w]{2,3}\b
Dates-DD/MM/YYYY	Matching dates in DD/MM/YYYY format like: 15/10/2019	\b([0-2][0-9]|(3)[0-1])(\/)(((0)[0-9])|((1)[0-2]))(\/)\d{4}\b

IBM Spectrum Discover

Modifying Tags

IBM Spectrum Scale



Modify Bucket

Please make sure that the maximum value for each bucket is greater than the value assigned to the previous bucket

Bucket Name

SizeRange

☑ extra small

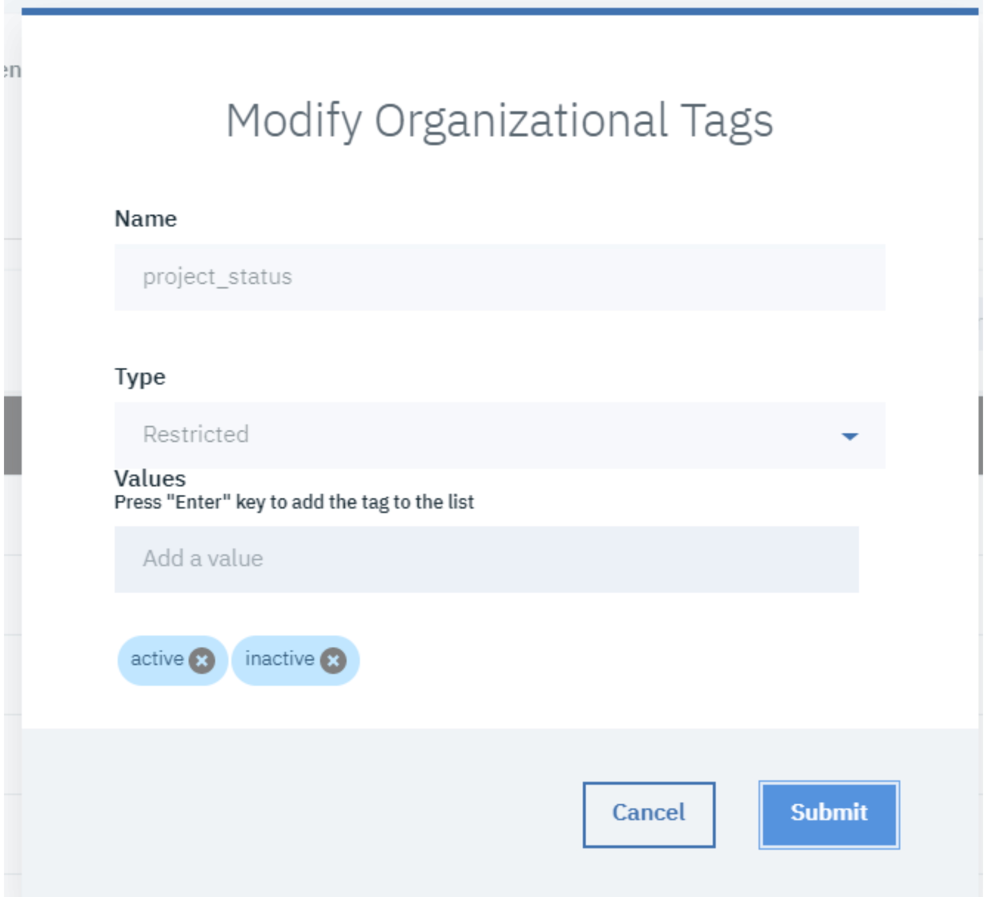
Values less than 4 KiB

☑ small

Between previous value and 1 MiB

☑ medium

Between previous value and 1 GiB



Modify Organizational Tags

Name

project_status

Type

Restricted

Values

Press "Enter" key to add the tag to the list

Add a value

active ⊗ inactive ⊗

Policies

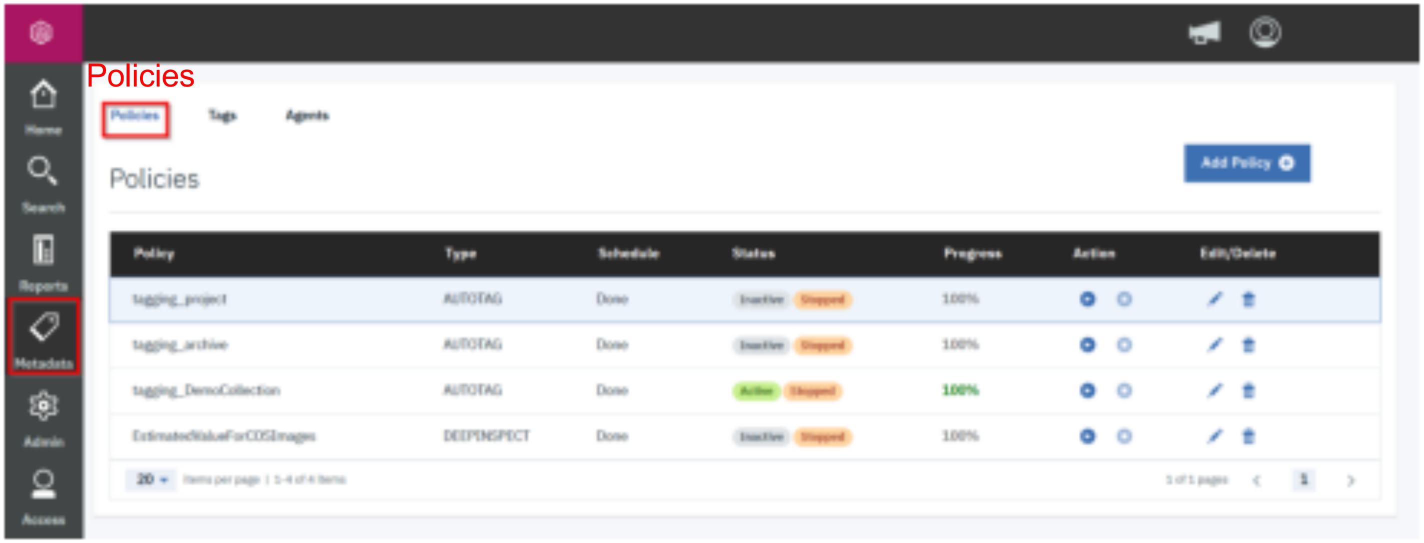
IBM Spectrum Scale

Policy Tagging

IBM Spectrum Scale

Leveraging the Spectrum Discover policy engine to apply the tags to the appropriate metadata records.
Go to Metadata > Policies

Metadata ->



Policies

Policy	Type	Schedule	Status	Progress	Action	Edit/Delete
tagging_project	AUTOTAG	Done	Inactive Stopped	100%		
tagging_archive	AUTOTAG	Done	Inactive Stopped	100%		
tagging_DemoCollection	AUTOTAG	Done	Active Stopped	100%		
EstimatedValueForCOSImages	DEEPINSPECT	Done	Inactive Stopped	100%		

20 Items per page | 1-4 of 4 items 1 of 1 pages 1

There are 2 types of policy tagging:

- AUTOTAG policy, which will add custom metadata values to all or a subset of the records based on filter criteria, and if necessary, the value can be extracted from the record path.
- DEEPINSPECT policy, allows you to enrich metadata through content inspection of source data. using an external agent, according to a filter.
- Filters are similar to the "where" clause in an SQL query. The filter is constructed using standard SQL syntax.
- Policies can be scheduled daily, weekly, monthly (or run on demand) and will be applied only if the policy is set active.

Deep Inspect Policies and Action Agents

IBM Spectrum Scale

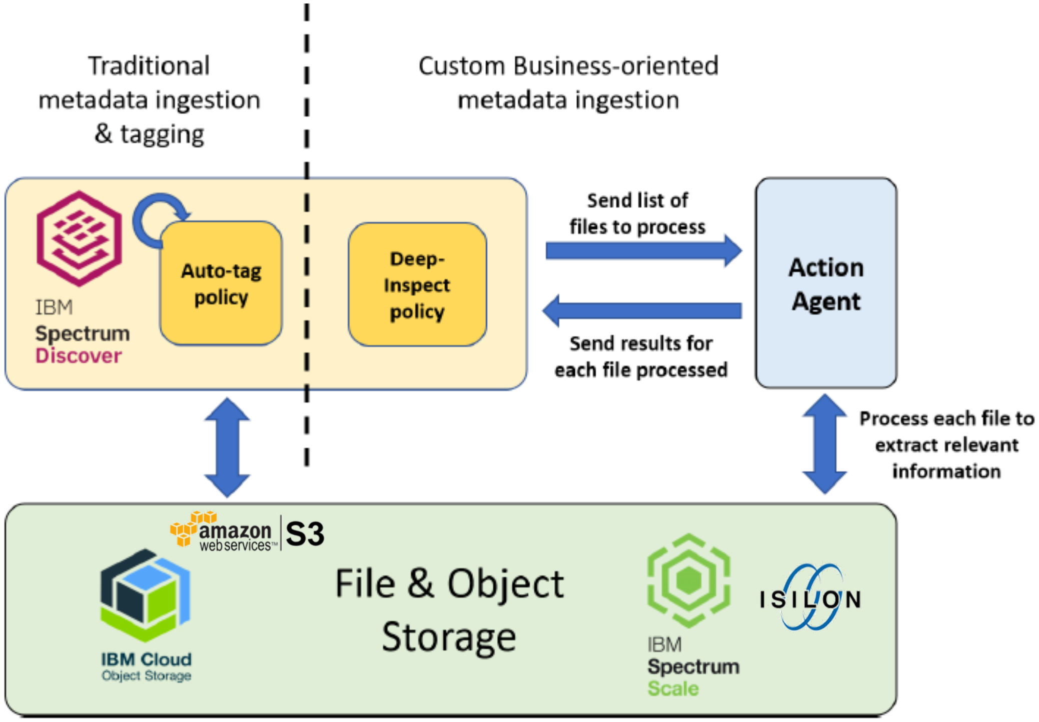
- Action agent - program interfacing with IBM Spectrum™ Discover and has **access to the source** storage.

- Use cases for action agents, including data content inspection for enriching metadata, data movement/migration, data scrubbing/sanitation

- Data is identified by IBM Spectrum Discover **by policy filter** and passed to the action agent as pointers through a messaging queue.

- Action agent performs whatever work is appropriate on the source data and returns a completion status back to IBM Spectrum Discover.

- If it does include enriched metadata, IBM Spectrum Discover **catalogs that metadata and makes it immediately searchable.**


Traditional metadata ingestion & tagging
Custom Business-oriented metadata ingestion
IBM Spectrum Discover
Auto-tag policy
Deep-Inspect policy
Send list of files to process
Send results for each file processed
Action Agent
Process each file to extract relevant information
amazon web services S3
IBM Cloud Object Storage
File & Object Storage
IBM Spectrum Scale
ISILON

Content-based Keyword Search & Tagging

IBM Spectrum Scale

FEATURE

The objective of the content-based keyword search use case is to:

- Provide **o**ut-of-the-box support for content search

- Enable end users to easily set up policies to automatically identify, classify and categorize data, which could be leveraged for specific business needs.

BENEFITS

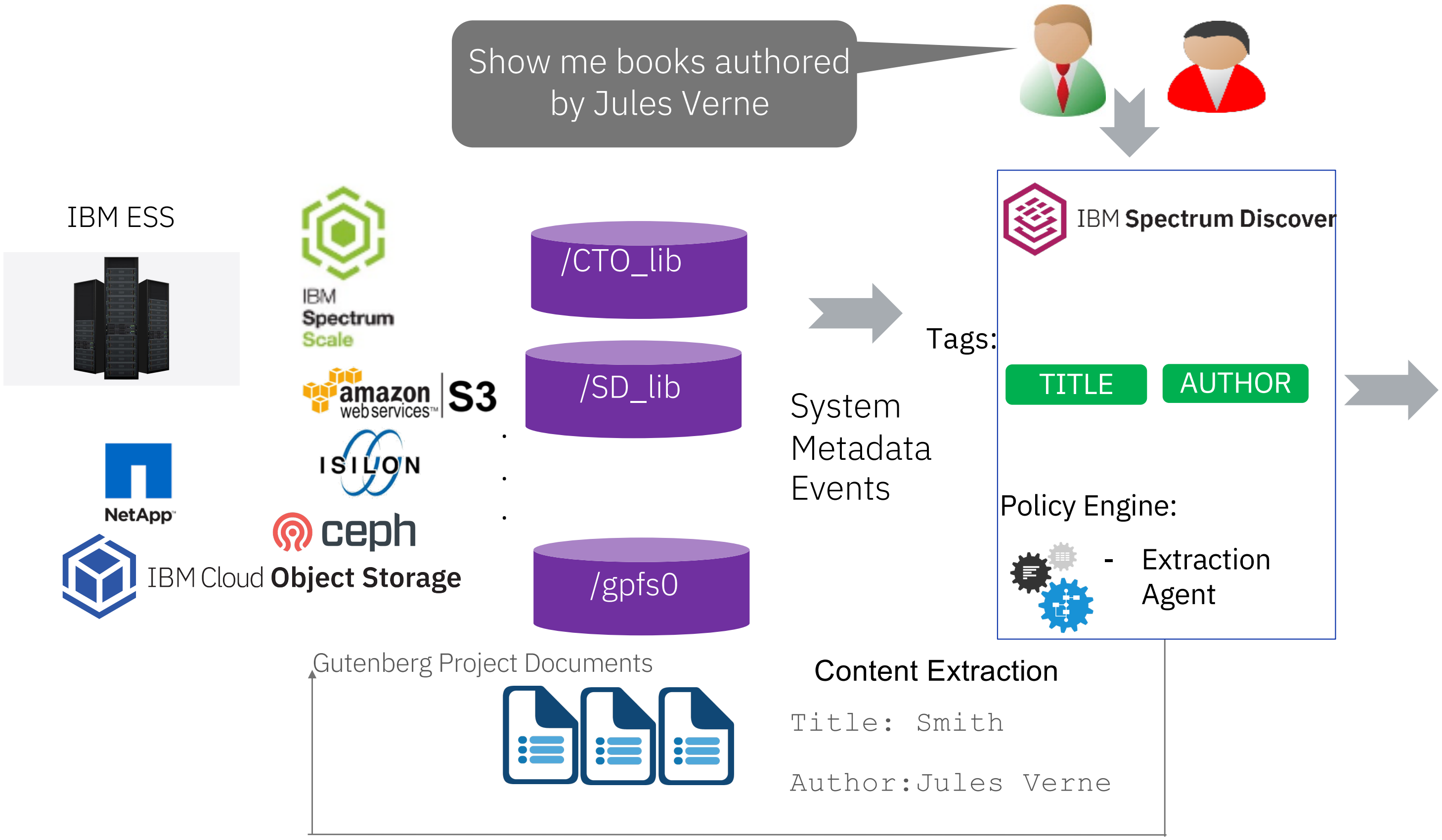
- For the Data Scientist, CIO and the Data Analyst, the ability to curate, extract and gather data containing specific keywords is critical in large scale analytics involving vast amounts of unstructured data.

- For the Data Steward and the CIO the ability to find and organize documents based on content greatly helps with their data administration efforts – for example, identifying data that may be subject to specific governance policies and/or compliance regulations.

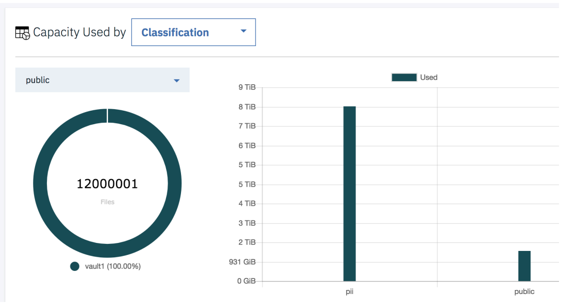
- For the Administrator, the ability to create and manage collections (logical groups of metadata) that share a common member access list along with leveraging the role based access controls (RBAC) feature

Use Case: Identifying relevant data based on content enrichments

IBM Spectrum Scale



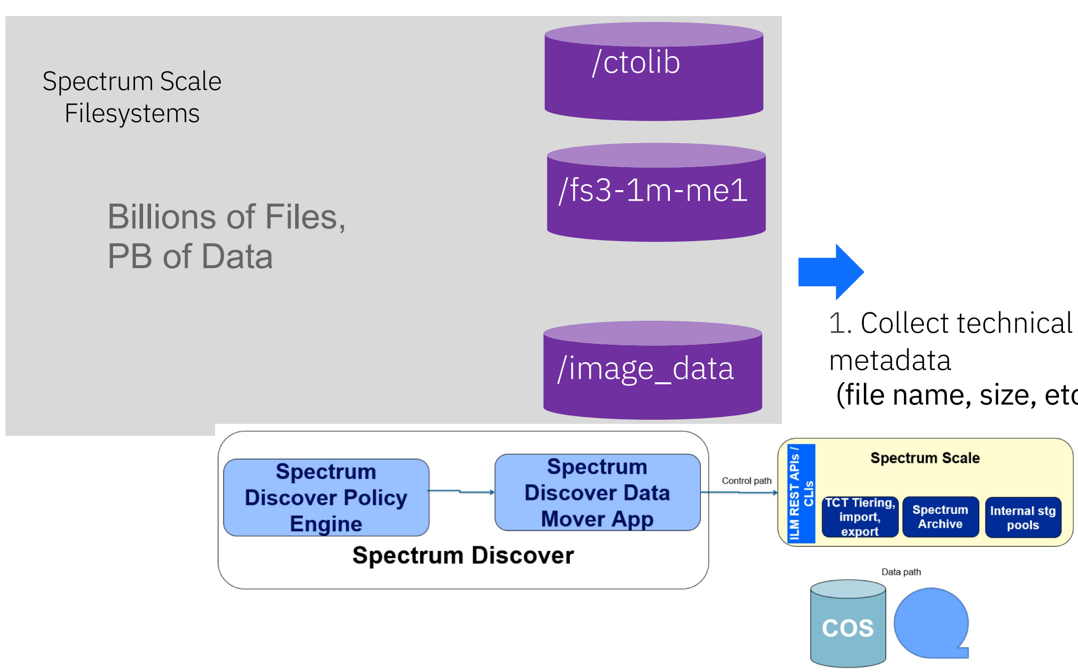
```
Title: Smith

Author:Jules Verne
```



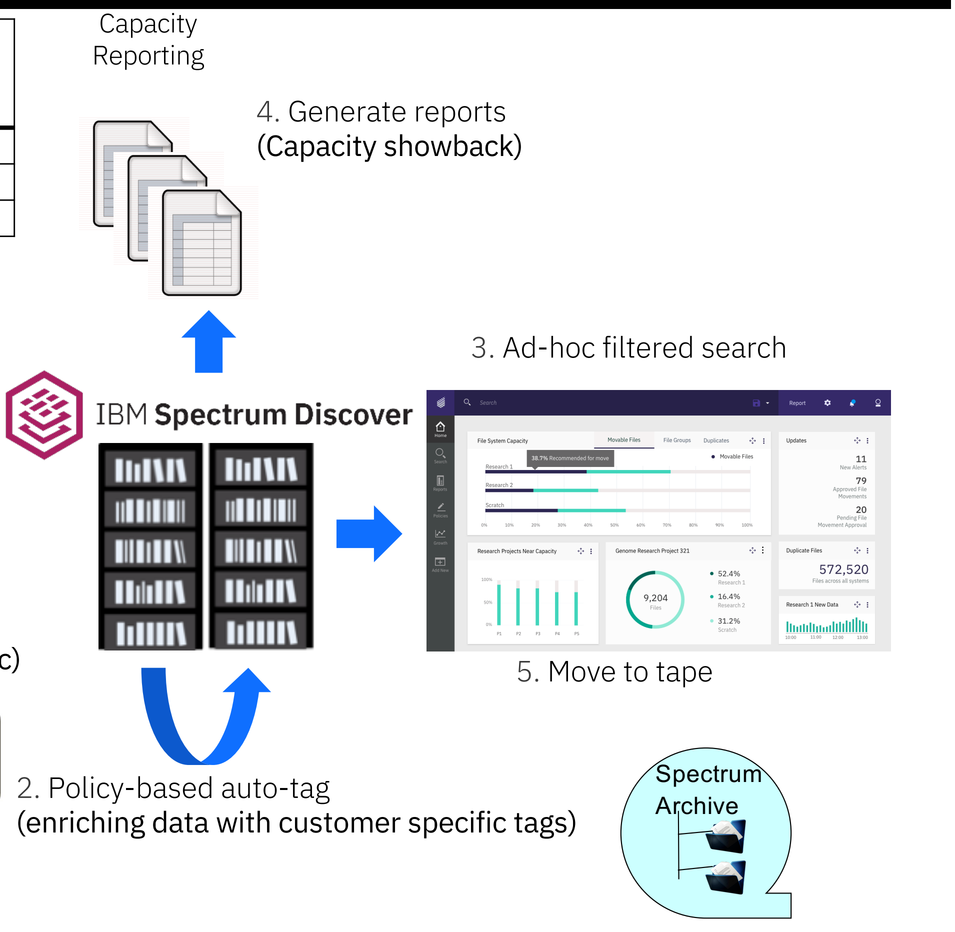
Use Case: Curating the Research Data for Placement Optimization

IBM Spectrum Scale

User	Department	Project	Project State	Spectrum Scale Fileset / Base Directory
ibmuser1	staff	phase1	active	/whole_cell
ibmuser2	postdoctoral	phase2	inactive	/nucleus
ibmuser3		phase3	active	/polysomes


Spectrum Scale Filesystems
Billions of Files, PB of Data
/ctolib
/fs3-1m-me1
/image_data
Spectrum Discover Policy Engine
Spectrum Discover Data Mover App
Spectrum Discover
Control path
ILM REST APIs / CLIs
Spectrum Scale
TCT Tiering, import, export
Spectrum Archive
Internal stg pools
Data path
COS

1. Collect technical metadata (file name, size, etc)



4. Generate reports (Capacity showback)

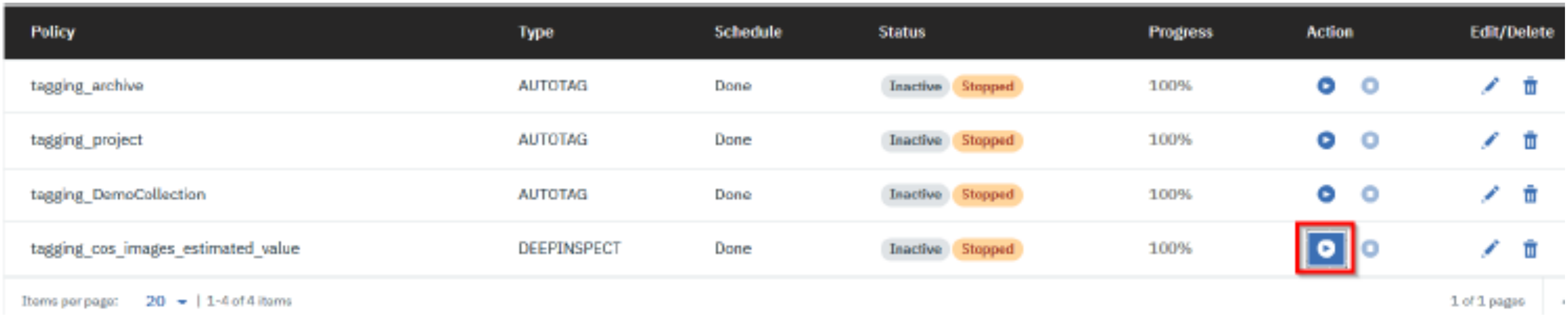
3. Ad-hoc filtered search

5. Move to tape

2. Policy-based auto-tag (enriching data with customer specific tags)

Example of the Results of Deep Inspect Policy

In this case, the agent is retrieving the headers of each IBM COS object and retrieve the requested S3 custom metadata.



Policy	Type	Schedule	Status	Progress	Action	Edit/Delete
tagging_archive	AUTOTAG	Done	Inactive Stopped	100%	▶ ■	✎ 🗑
tagging_project	AUTOTAG	Done	Inactive Stopped	100%	▶ ■	✎ 🗑
tagging_DemoCollection	AUTOTAG	Done	Inactive Stopped	100%	▶ ■	✎ 🗑
tagging_cos_images_estimated_value	DEEPINSPECT	Done	Inactive Stopped	100%	▶ ■	✎ 🗑

Items per page: 20 ▾ | 1-4 of 4 items

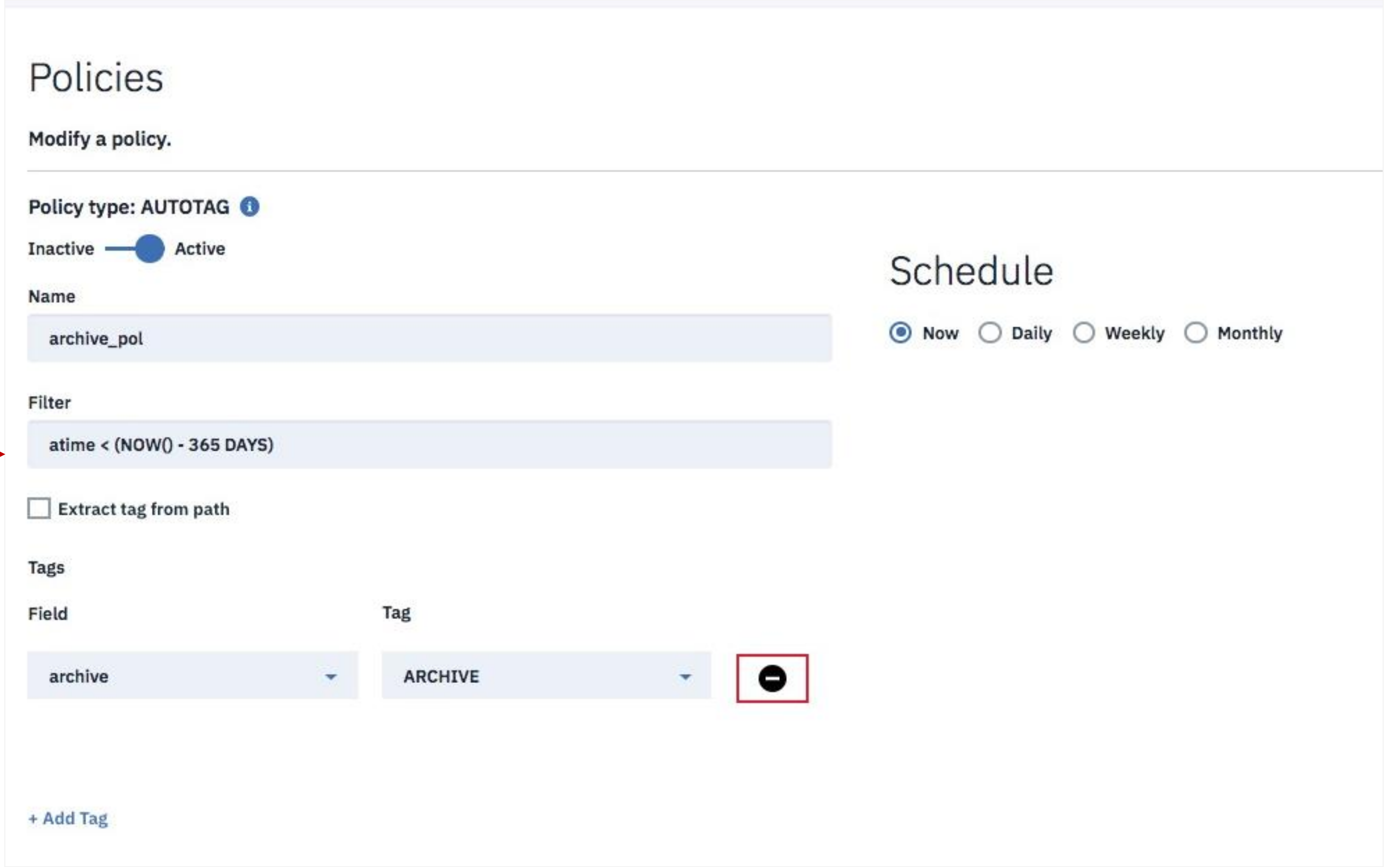
1 of 1 pages

From the Metadata > Policies page, clicking on the start button on the “tagging_cos_images_estimated_value” line , the status of the policy will change from **Stopped** to **Running.**

Example of a Temperature Policy

IBM Spectrum Scale

- This policy will select all the files that have not been accessed for 1 year and set the TEMPERATURE tag to ARCHIVE.
- This tagging is reflected in the Home Dashboard where you can see that some files are recommended to move.



Policies

Modify a policy.

Policy type: AUTOTAG

Inactive ● Active

Name: archive_pol

Filter: atime < (NOW() - 365 DAYS)

□ Extract tag from path

Tags

Field	Tag
archive	ARCHIVE

+ Add Tag

Schedule

● Now ○ Daily ○ Weekly ○ Monthly

IBM Spectrum Discover

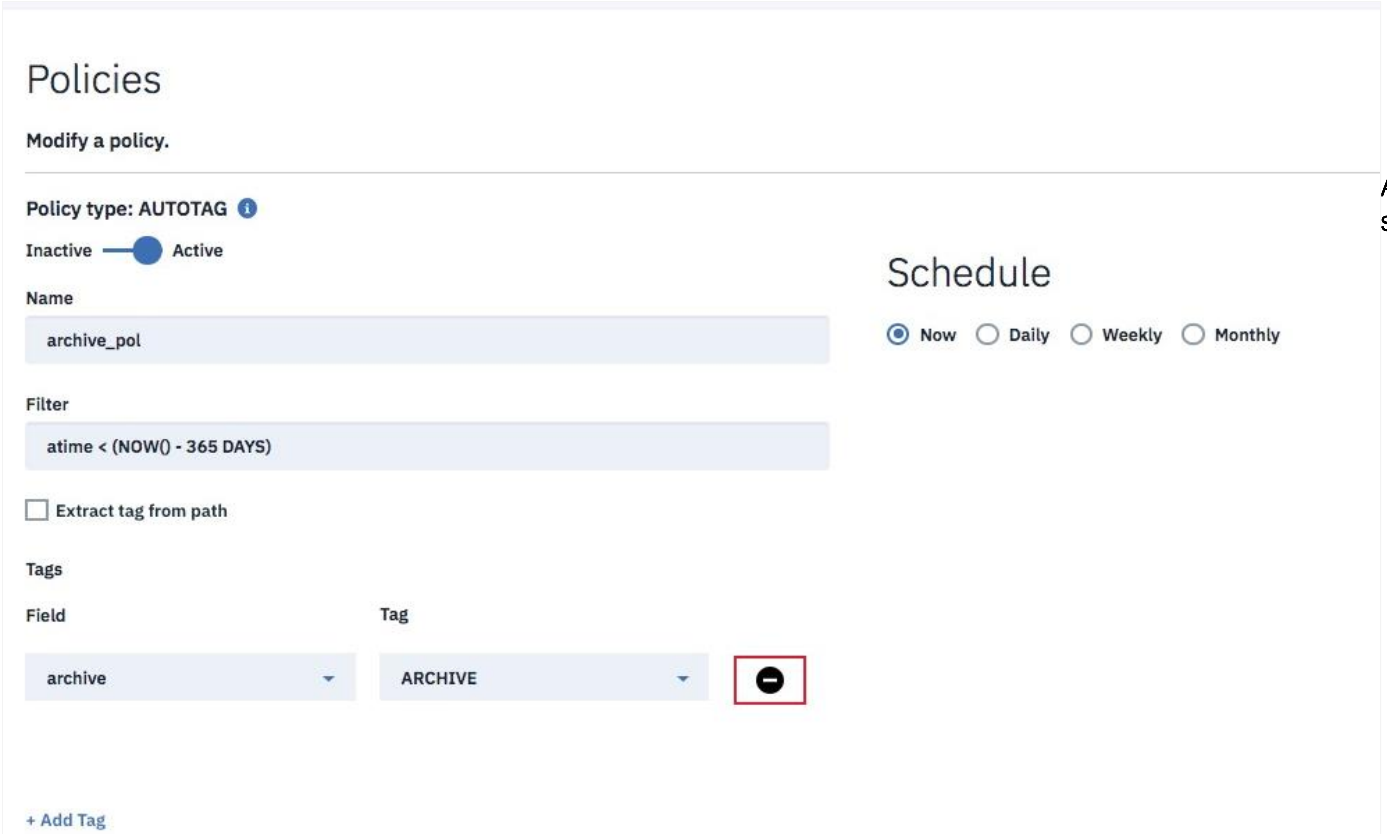
Expanding on Temperature Policy to Execute an Archive

IBM Spectrum Scale

- Policy schedules can be immediate execution, daily, weekly, monthly.

Define/Modify policies

Identify tag applied by policy



Policies

Modify a policy.

Policy type: AUTOTAG

Inactive ● Active

Name: archive_pol

Filter: atime < (NOW() - 365 DAYS)

□ Extract tag from path

Tags

Field	Tag
archive	ARCHIVE

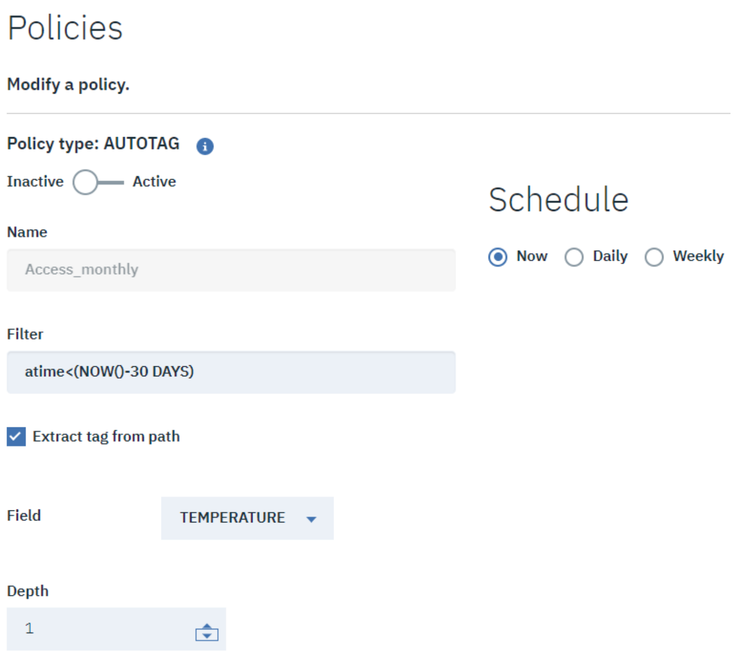
+ Add Tag

Schedule

● Now ○ Daily ○ Weekly ○ Monthly

Automate execution by setting schedule

Creating a Monthly Policy & an Estimated Value Policy



Policies

Modify a policy.

Policy type: AUTOTAG

Inactive ○— Active

Name: Access_monthly

Filter: atime<(NOW()-30 DAYS)

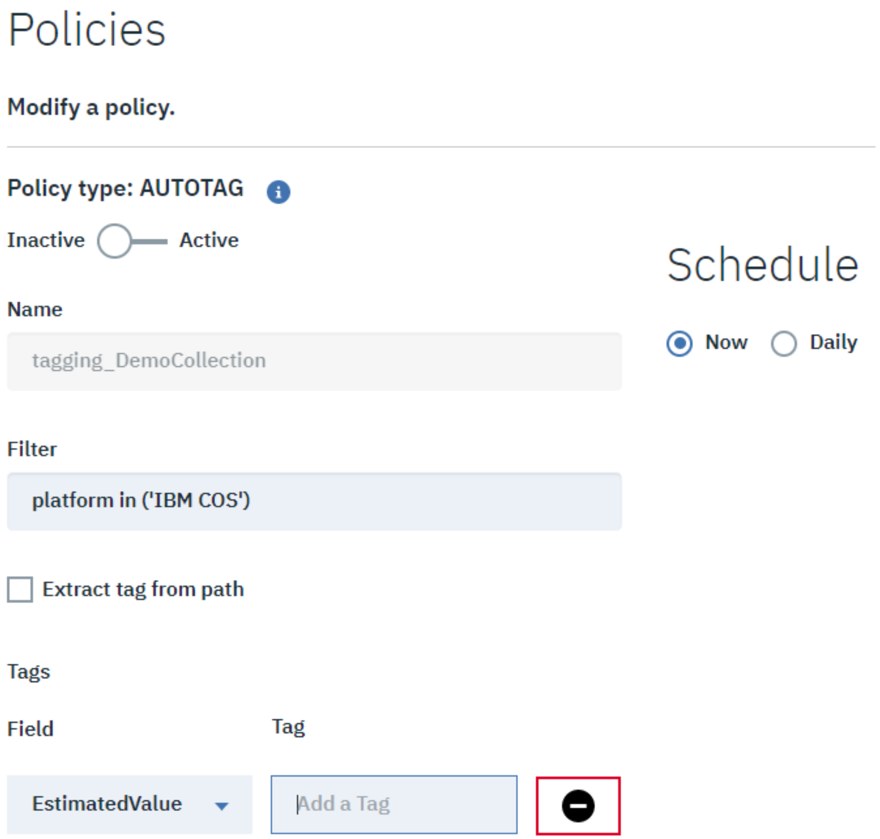
☑ Extract tag from path

Field: TEMPERATURE

Depth: 1

Schedule

● Now ○ Daily ○ Weekly



Policies

Modify a policy.

Policy type: AUTOTAG

Inactive ○— Active

Name: tagging_DemoCollection

Filter: platform in ('IBM COS')

□ Extract tag from path

Tags

Field	Tag
EstimatedValue	Add a Tag

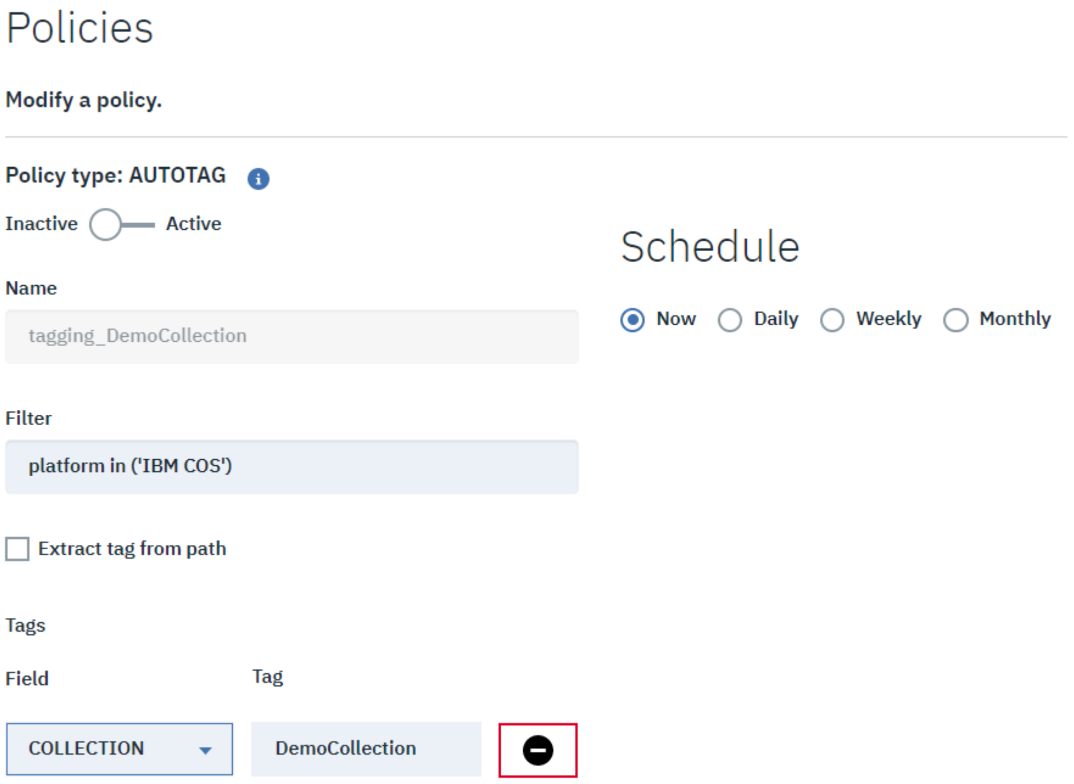
Schedule

● Now ○ Daily

Data Collection Policy

IBM Spectrum Scale

Collections may be used to restrict access to files based on a tag, and values assigned to that tag



Policies

Modify a policy.

Policy type: AUTOTAG

Inactive ○— Active

Name

tagging_DemoCollection

Filter

platform in ('IBM COS')

□ Extract tag from path

Tags

Field	Tag
COLLECTION	DemoCollection

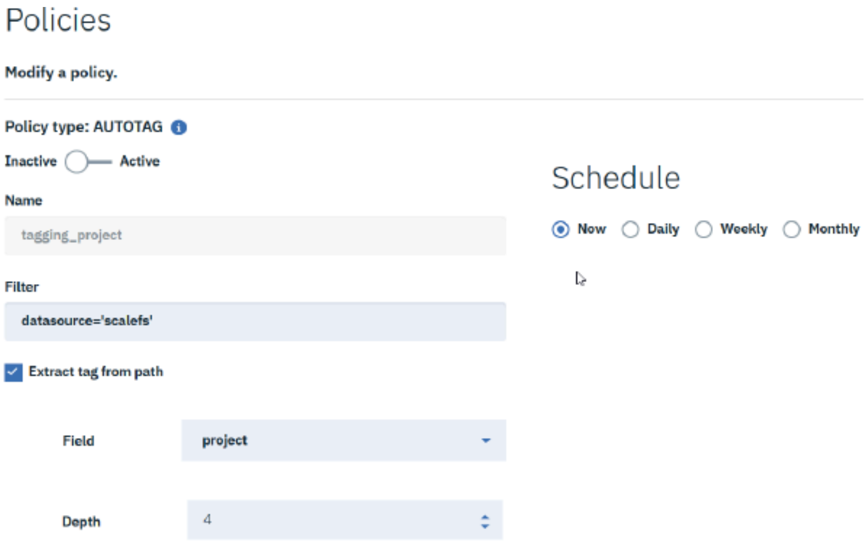
Schedule

● Now ○ Daily ○ Weekly ○ Monthly

Example of a Sub-directory Policy

IBM Spectrum Scale

Select all the metadata belonging to the datasource named **scalefs** and will extract from the path of each record the **project** value, which will be the 4th field of the path.



Policies

Modify a policy.

Policy type: AUTOTAG

Inactive ○— Active

Name: tagging_project

Datasource -> Filter: datasource='scalefs'

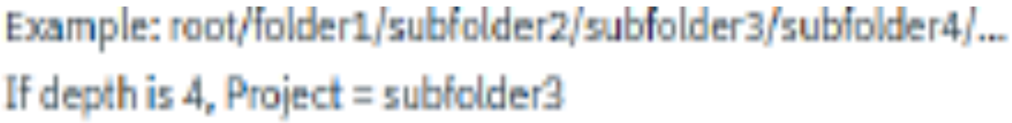
☑ Extract tag from path

Field: project

Depth of directory-> Depth: 4

Schedule

● Now ○ Daily ○ Weekly ○ Monthly

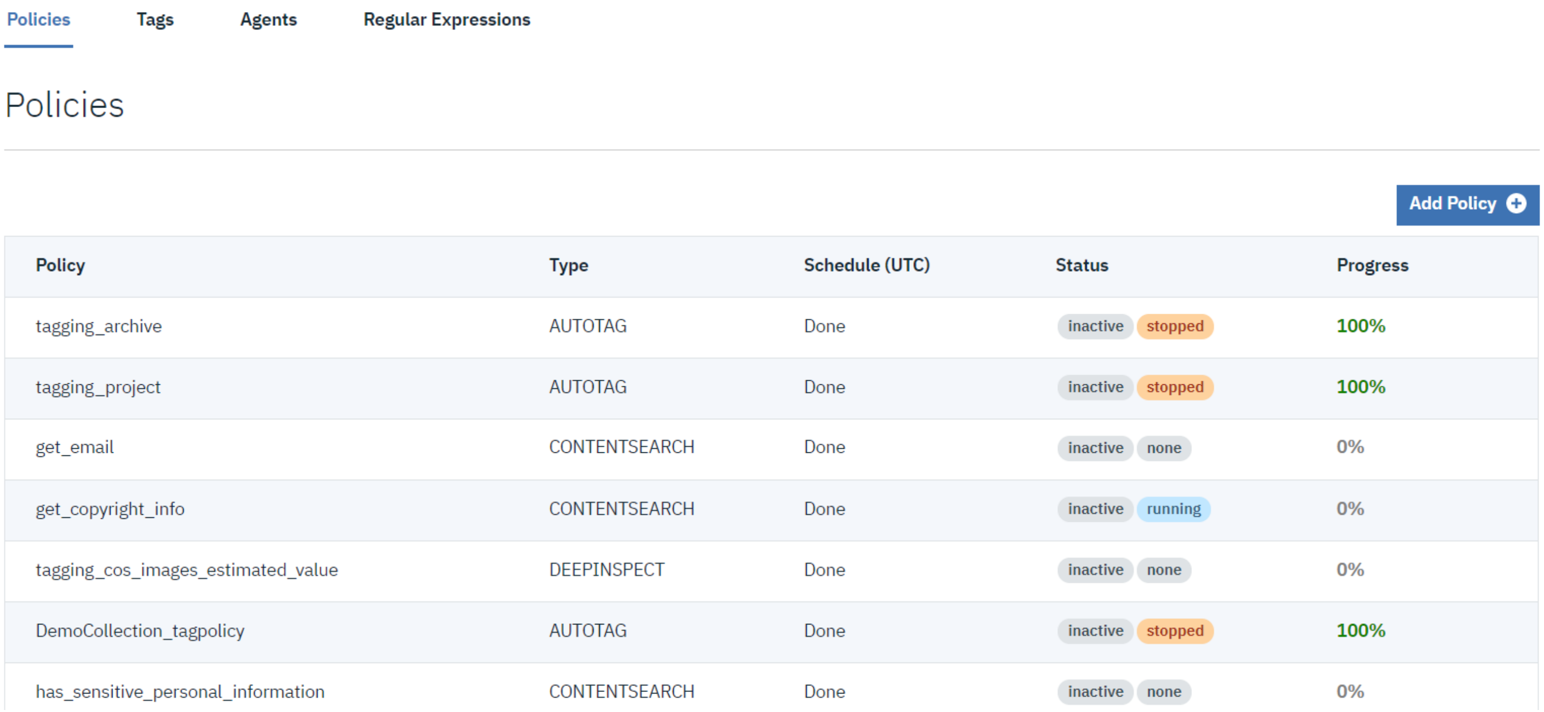


Example: root/folder1/subfolder2/subfolder3/subfolder4/...

If depth is 4, Project = subfolder3

Creating & Executing Multiple Policies

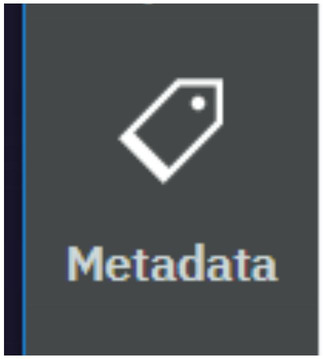
Monitor policies, change, edit or delete



Policies

Policy	Type	Schedule (UTC)	Status	Progress
tagging_archive	AUTOTAG	Done	inactive stopped	100%
tagging_project	AUTOTAG	Done	inactive stopped	100%
get_email	CONTENTSEARCH	Done	inactive none	0%
get_copyright_info	CONTENTSEARCH	Done	inactive running	0%
tagging_cos_images_estimated_value	DEEPINSPECT	Done	inactive none	0%
DemoCollection_tagpolicy	AUTOTAG	Done	inactive stopped	100%
has_sensitive_personal_information	CONTENTSEARCH	Done	inactive none	0%

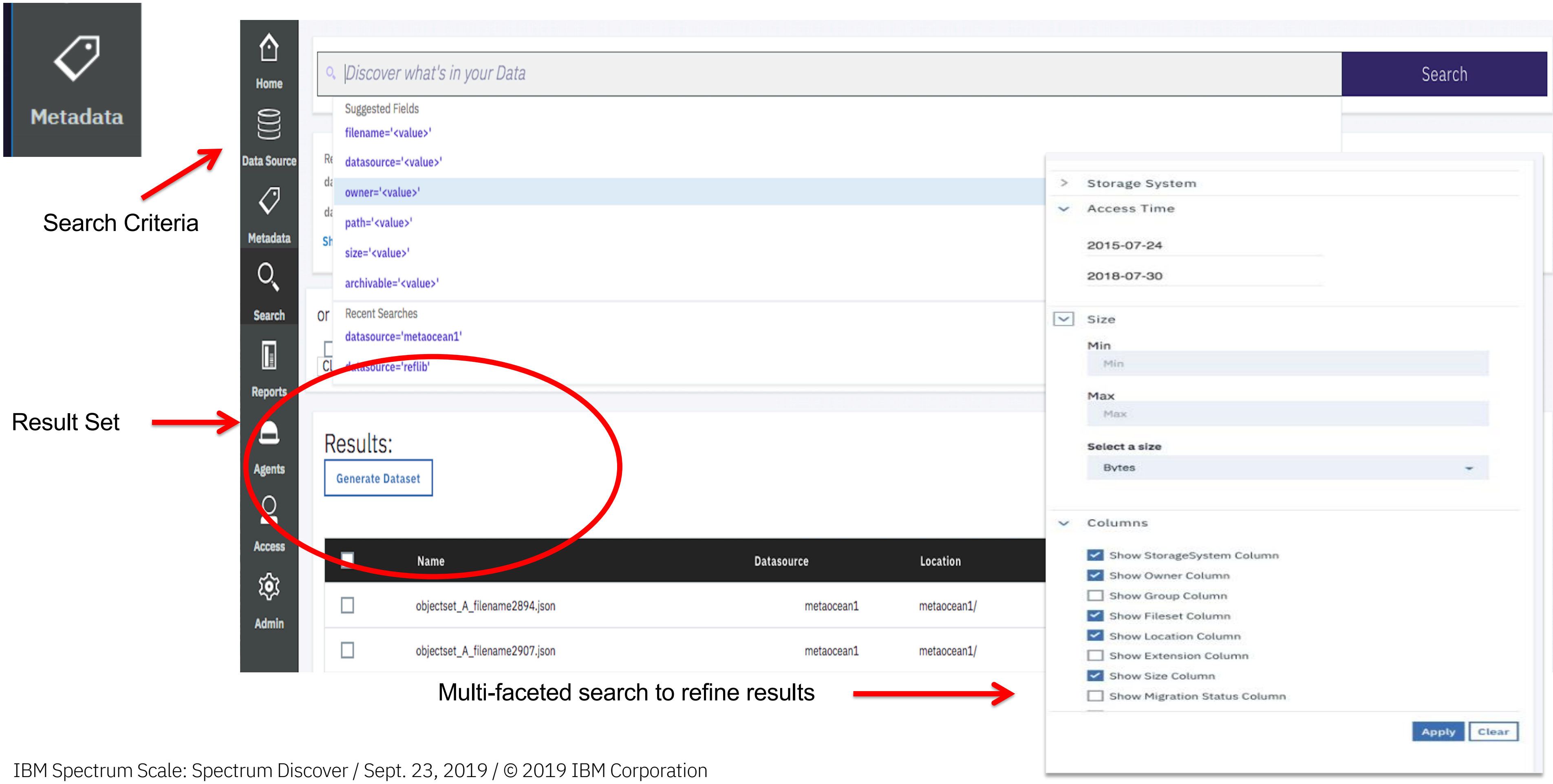
Metadata Search



IBM
Spectrum
Discover

Metadata Search

IBM Spectrum Scale



Search Criteria

Result Set

Multi-faceted search to refine results

Suggested Fields:

- filename='<value>'
- datasource='<value>'
- owner='<value>'
- path='<value>'
- size='<value>'
- archivable='<value>'

Recent Searches:

- datasource='metaocean1'
- datasource='reflib'

□	Name	Datasource	Location
□	objectset_A_filename2894.json	metaocean1	metaocean1/
□	objectset_A_filename2907.json	metaocean1	metaocean1/

Access Time: 2015-07-24 to 2018-07-30

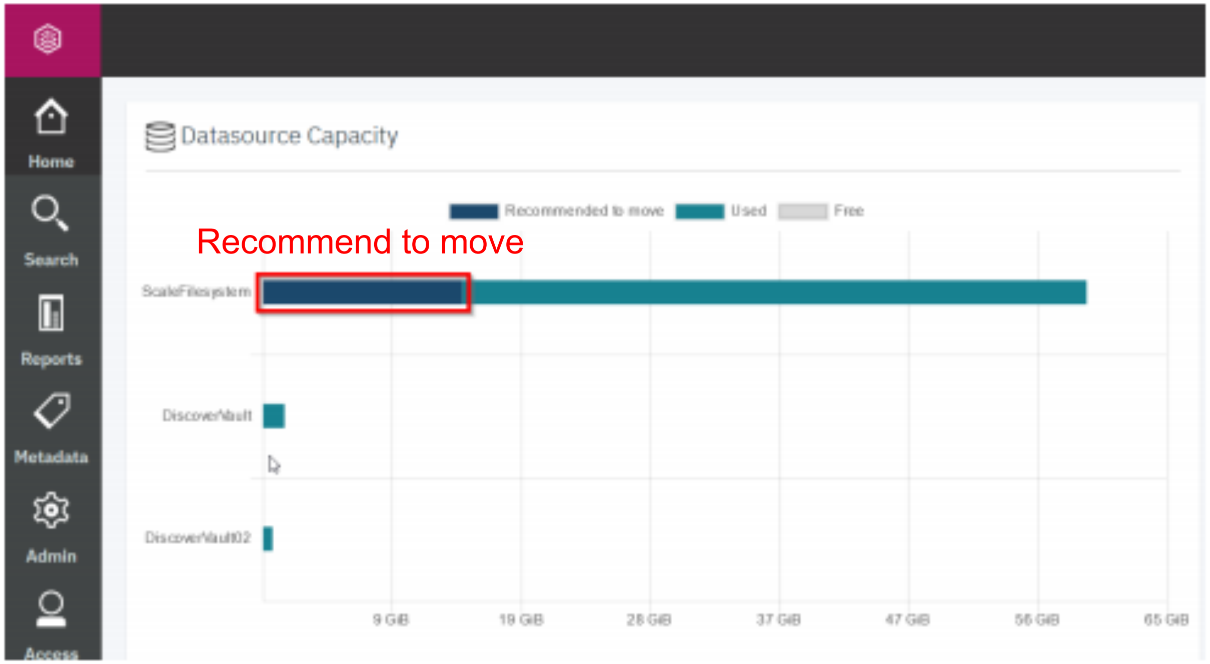
Size: Min, Max; Select a size: Bytes

Columns:

- ☑ Show StorageSystem Column
- ☑ Show Owner Column
- □ Show Group Column
- ☑ Show Fileset Column
- ☑ Show Location Column
- □ Show Extension Column
- ☑ Show Size Column
- □ Show Migration Status Column

Take action on specific metadata/policies

- Leverage the drilldown faceted search capability of Spectrum Discover to perform visual analytics on the metadata and take action manually or by policy, based on the insights specifically.

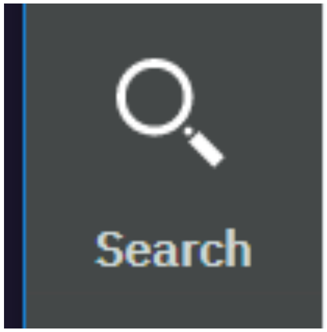


Recommend to move

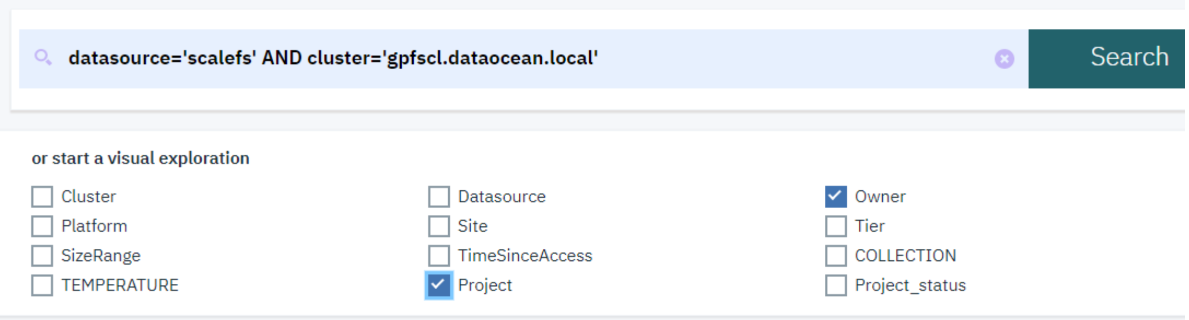
IBM Spectrum Discover

Search Results

IBM Spectrum Scale



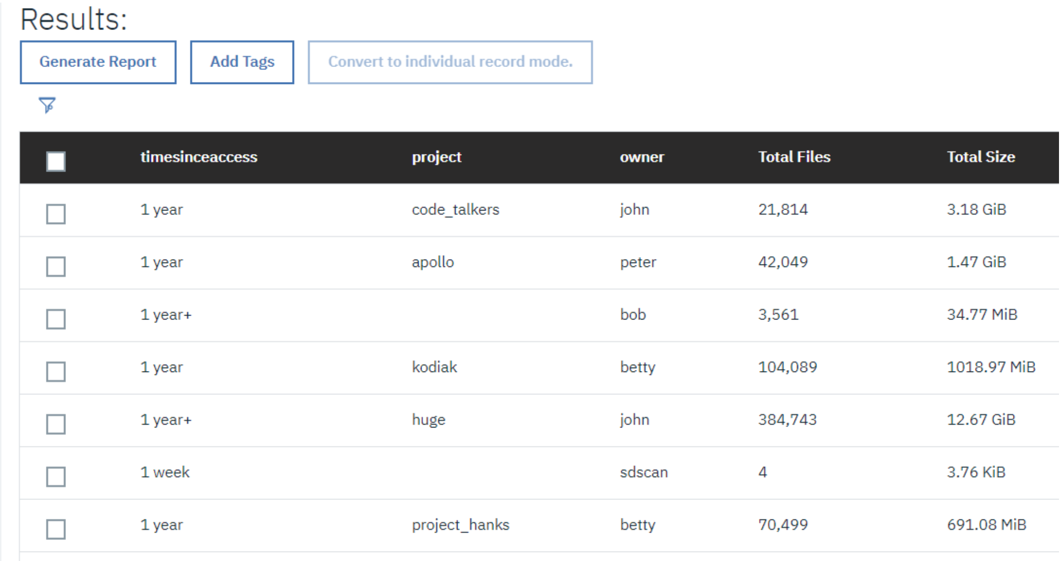
Searching Results by Owner and Project



datasource='scalefs' AND cluster='gpfscl.dataocean.local'

or start a visual exploration

□ Cluster	□ Datasource	☑ Owner
□ Platform	□ Site	□ Tier
□ SizeRange	□ TimeSinceAccess	□ COLLECTION
□ TEMPERATURE	☑ Project	□ Project_status



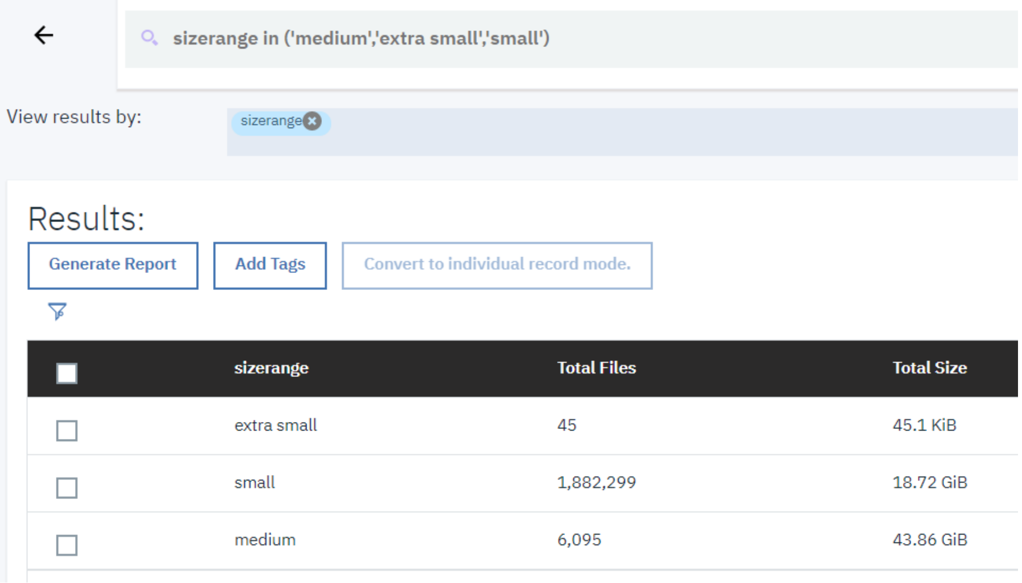
Results:

□	timesinceaccess	project	owner	Total Files	Total Size
□	1 year	code_talkers	john	21,814	3.18 GiB
□	1 year	apollo	peter	42,049	1.47 GiB
□	1 year+		bob	3,561	34.77 MiB
□	1 year	kodiak	betty	104,089	1018.97 MiB
□	1 year+	huge	john	384,743	12.67 GiB
□	1 week		sdscan	4	3.76 KiB
□	1 year	project_hanks	betty	70,499	691.08 MiB

IBM Spectrum Discover

Example: Search Results by Range of File Sizes and Time last Accessed

IBM Spectrum Scale

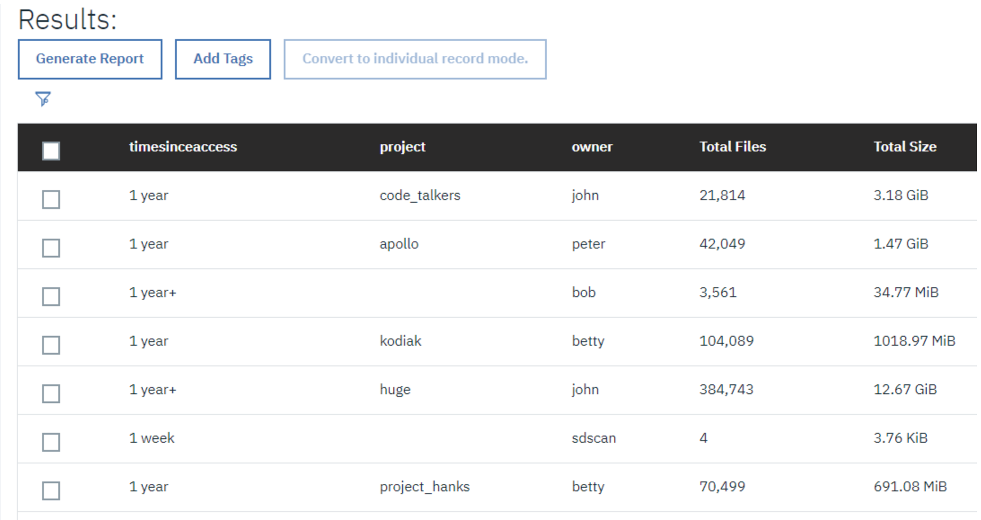


sizerange in ('medium','extra small','small')

View results by: sizerange

Results:

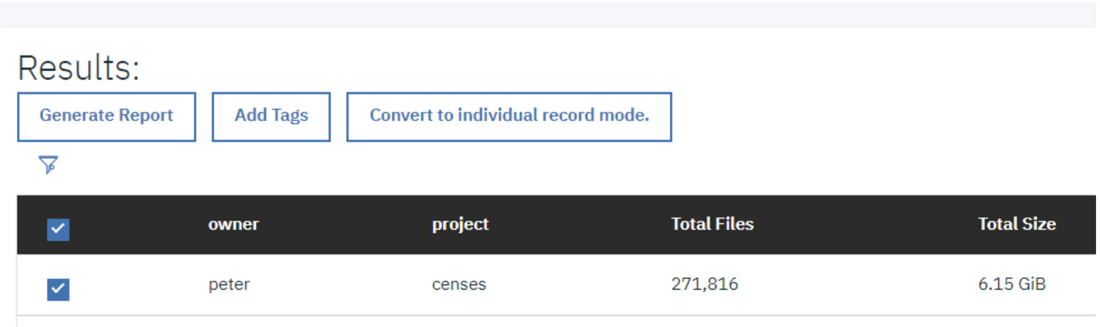
□	sizerange	Total Files	Total Size
□	extra small	45	45.1 KiB
□	small	1,882,299	18.72 GiB
□	medium	6,095	43.86 GiB



Results:

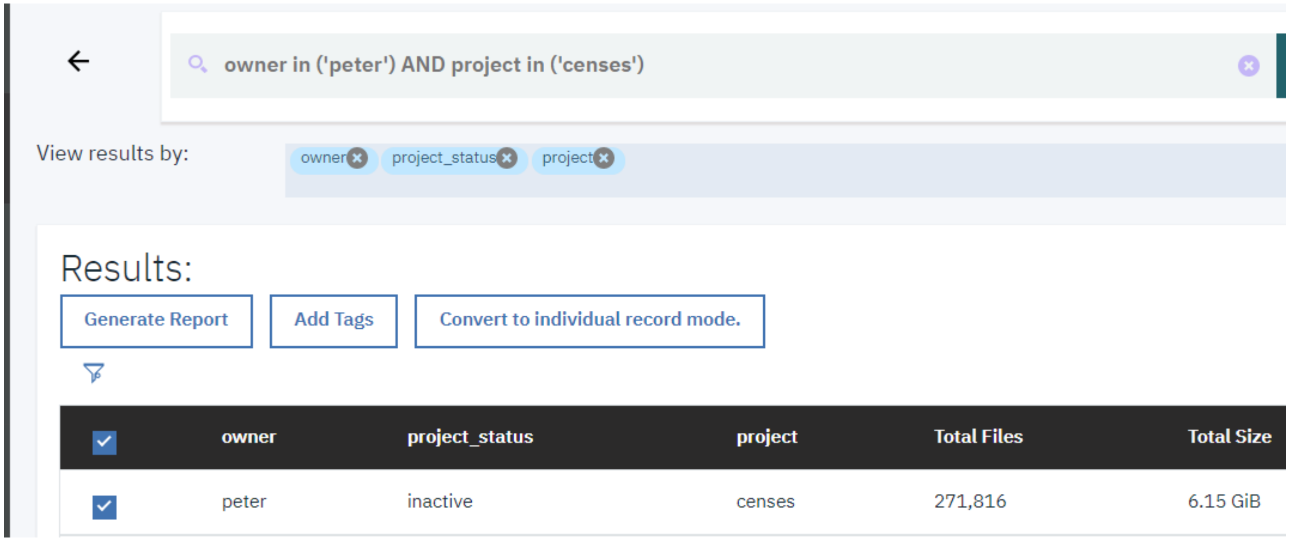
□	timesinceaccess	project	owner	Total Files	Total Size
□	1 year	code_talkers	john	21,814	3.18 GiB
□	1 year	apollo	peter	42,049	1.47 GiB
□	1 year+		bob	3,561	34.77 MiB
□	1 year	kodiak	betty	104,089	1018.97 MiB
□	1 year+	huge	john	384,743	12.67 GiB
□	1 week		sdscan	4	3.76 KiB
□	1 year	project_hanks	betty	70,499	691.08 MiB

Using Results to Select for Archiving



Results:

☑	owner	project	Total Files	Total Size
☑	peter	censes	271,816	6.15 GiB



owner in ('peter') AND project in ('censes')

View results by: owner ⊗ project_status ⊗ project ⊗

Results:

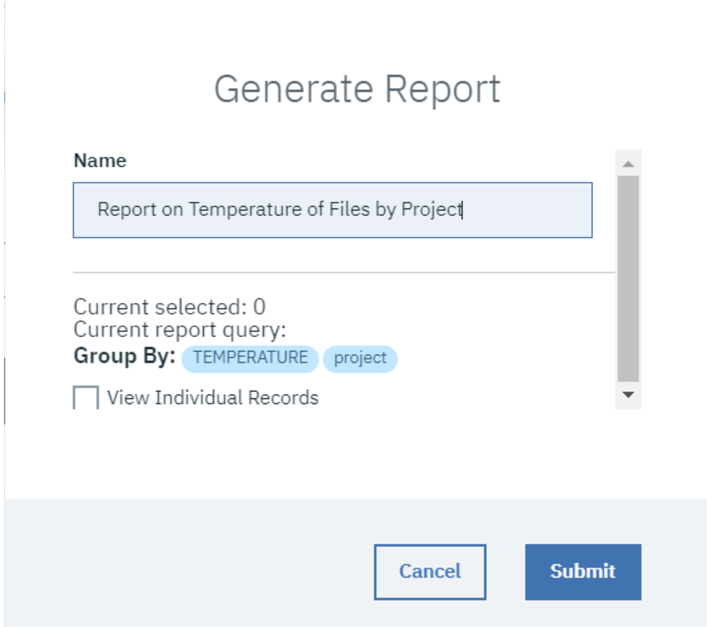
☑	owner	project_status	project	Total Files	Total Size
☑	peter	inactive	censes	271,816	6.15 GiB

Showing Owner and Project now Inactive

IBM Spectrum Discover

Example: Results by Project and Temperature

IBM Spectrum Scale



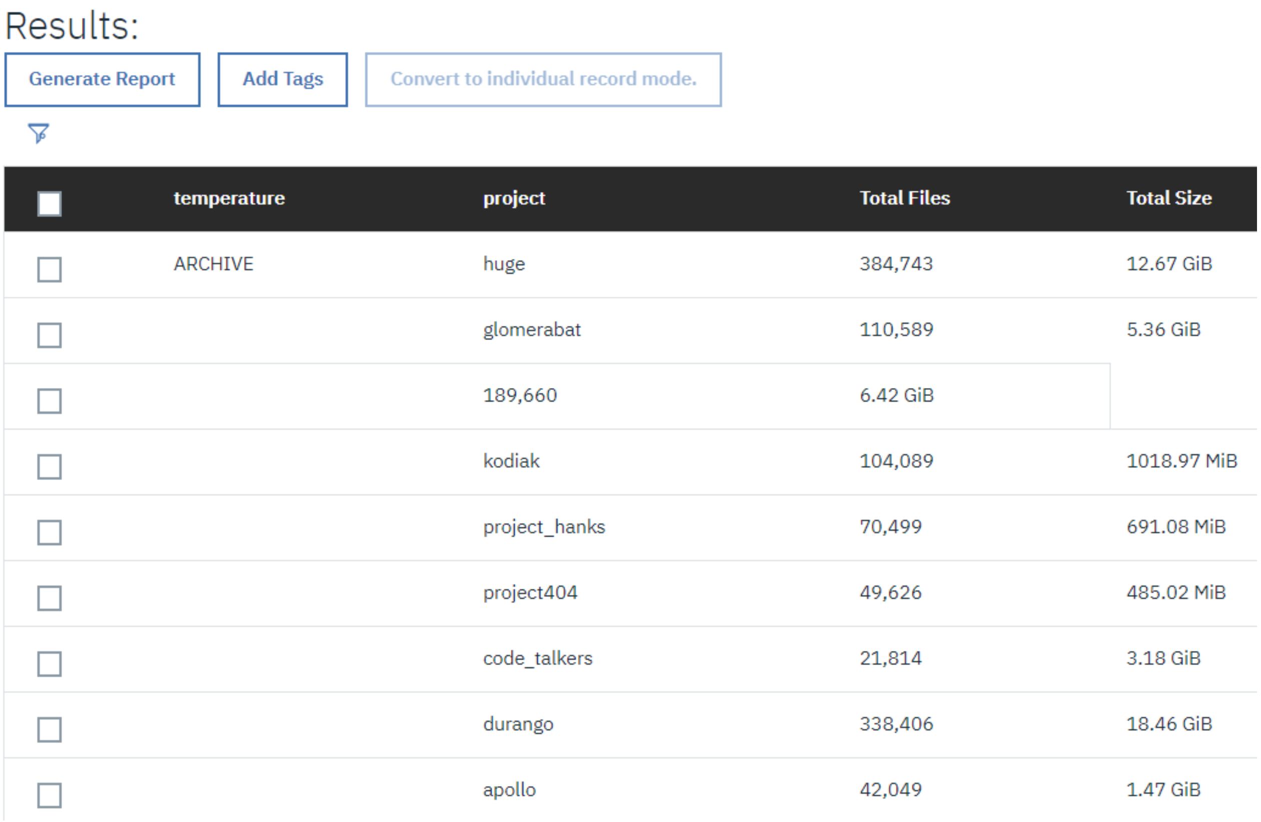
Generate Report

Name

Report on Temperature of Files by Project

Current selected: 0
Current report query:
Group By: TEMPERATURE project

□ View Individual Records



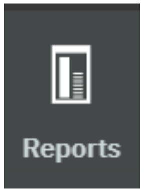
Results:

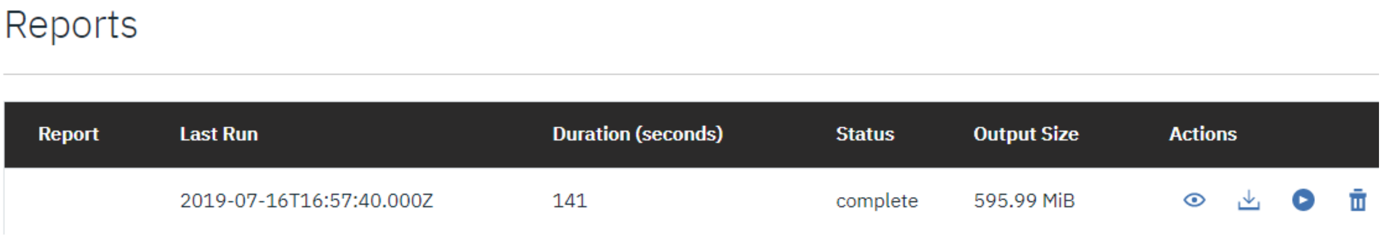
□	temperature	project	Total Files	Total Size
□	ARCHIVE	huge	384,743	12.67 GiB
□		glomerabat	110,589	5.36 GiB
□		189,660	6.42 GiB	
□		kodiak	104,089	1018.97 MiB
□		project_hanks	70,499	691.08 MiB
□		project404	49,626	485.02 MiB
□		code_talkers	21,814	3.18 GiB
□		durango	338,406	18.46 GiB
□		apollo	42,049	1.47 GiB

IBM
Spectrum
Discover

Generating Reports: Selecting Owner and Project

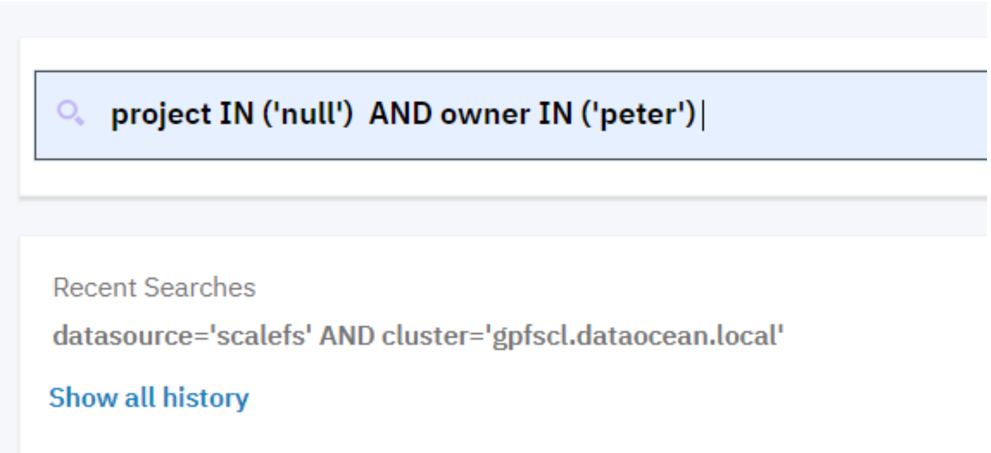
IBM Spectrum Scale





Reports

Report	Last Run	Duration (seconds)	Status	Output Size	Actions
	2019-07-16T16:57:40.000Z	141	complete	595.99 MiB	👁 ⤓ ▶ 🗑

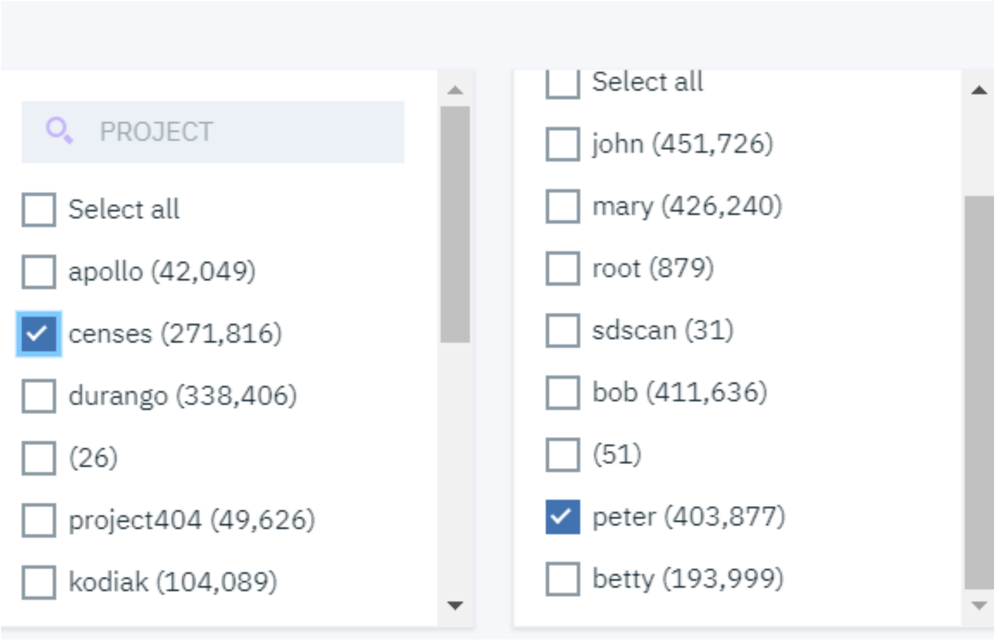


project IN ('null') AND owner IN ('peter')|

Recent Searches

datasource='scalefs' AND cluster='gpfscl.dataocean.local'

Show all history



PROJECT

- □ Select all
- □ apollo (42,049)
- ☑ censes (271,816)
- □ durango (338,406)
- □ (26)
- □ project404 (49,626)
- □ kodiak (104,089)

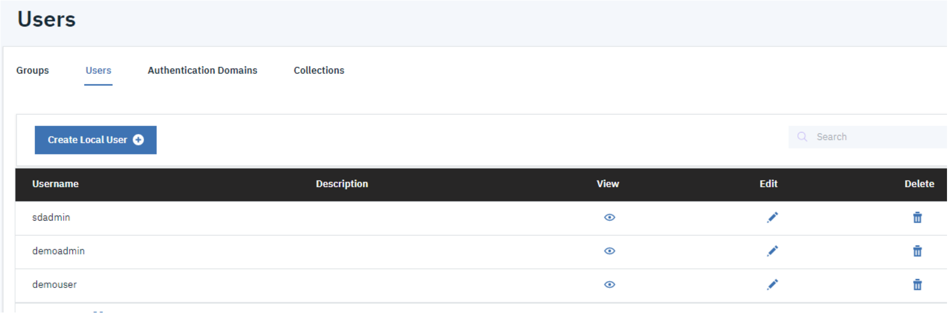
- □ Select all
- □ john (451,726)
- □ mary (426,240)
- □ root (879)
- □ sdscan (31)
- □ bob (411,636)
- □ (51)
- ☑ peter (403,877)
- □ betty (193,999)

Access Options

IBM Spectrum Scale

Access

Create users, groups, collections



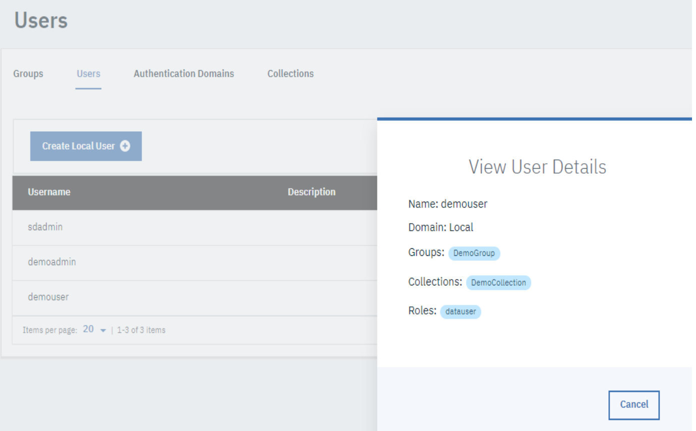
Users

Groups | Users | Authentication Domains | Collections

Create Local User ⊕

Search

Username	Description	View	Edit	Delete
sdadmin		👁	✎	🗑
demoadmin		👁	✎	🗑
demouser		👁	✎	🗑



Users

Groups | Users | Authentication Domains | Collections

Create Local User ⊕

Username	Description
sdadmin	
demoadmin	
demouser	

Items per page: 20 ▾ | 1-3 of 3 items

View User Details

Name: demouser

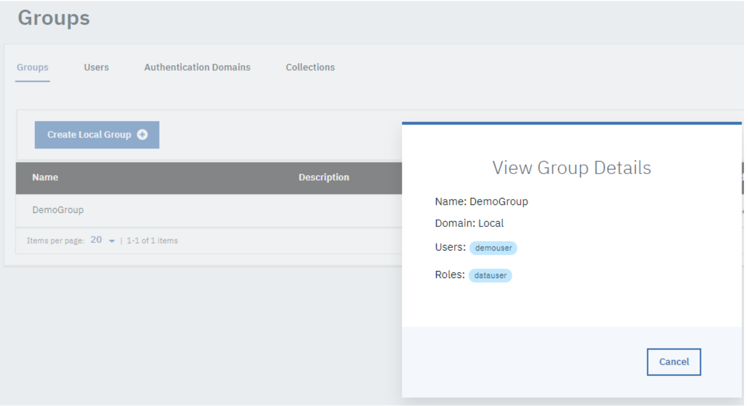
Domain: Local

Groups: DemoGroup

Collections: DemoCollection

Roles: datauser

Cancel



Groups

Groups | Users | Authentication Domains | Collections

Create Local Group ⊕

Name	Description
DemoGroup	

Items per page: 20 ▾ | 1-1 of 1 items

View Group Details

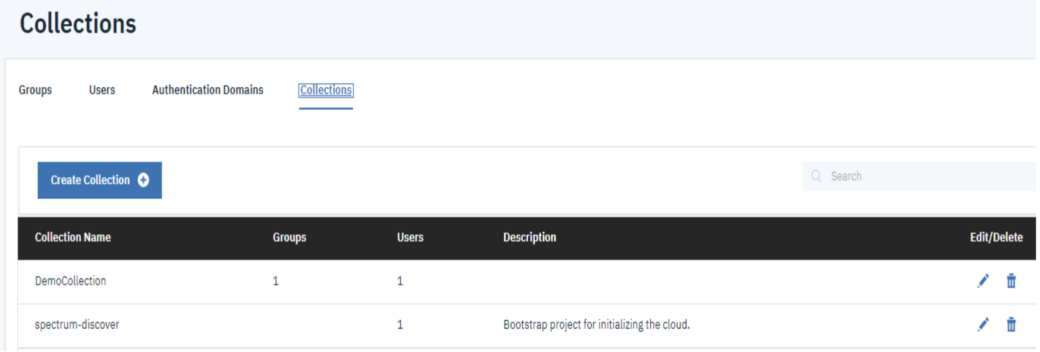
Name: DemoGroup

Domain: Local

Users: demouser

Roles: datauser

Cancel



Collections

Groups | Users | Authentication Domains | Collections

Create Collection ⊕

Search

Collection Name	Groups	Users	Description	Edit/Delete
DemoCollection	1	1		✎ 🗑
spectrum-discover		1	Bootstrap project for initializing the cloud.	✎ 🗑

Licensing

IBM Spectrum Scale

Pricing Summary

- Licensed based on data managed by the Program (L-GMVS-BU26FM)
- Aggregate size of all the files that Spectrum Discover indexes and/or scans
- Ability to report on that size (whatever is indexed, logical size)
- For Spectrum Scale: can be configured to manage data on a specific file-system(s) or fileset(s) on that file-system.
- For Cloud Object Storage: can be configured to manage data on a specific vault(s).
- 90-day FREE trial

Licensing Summary

- Licensed on a managed terabyte basis; flat pricing, no tiering; customers can manage as little or as much as they want.
- Orderable through either PPA or AAS; also available via eConfig for IBM Cloud Object Storage

Spectrum Discover POC process

IBM Spectrum Scale

Spectrum Discover POC process

IBM Spectrum Scale

1. Identify use cases to test during POC process
   a. Test cases and success criteria are provided for the following use cases:
      i. Installation and configuration
      ii. Metadata harvesting
      iii. Storage optimization
      iv. Content inspection and data classification
      v. Content based keyword search
      vi. Role Based Access Control
      vii. Custom Tagging
      viii. Security Analytics
      ix. Researcher / data scientist portal
   b. Custom use cases based on unique customer requirements

2. Complete Spectrum Discover POC pre-planning worksheet
   a. This worksheet specifies the pre-requisites for the POC
3. Review Spectrum Discover pre-planning worksheet with Spectrum Discover technical team (either via webex or email)
4. Install and configure Spectrum Discover in your environment
   a. Conduct remote support webex with Spectrum Discover technical team during installation and configuration
5. Harvest metadata from data sources
   a. Conduct remote support webex with Spectrum Discover Development to verify scans successfully started
6. Spectrum Discover workshop (on site meeting or webex)
   a. Transfer of knowledge / lab training for how to use the Spectrum Discover GUI
   b. Discussion on custom tagging policy based on organizational constructs
7. POC Test plan execution
8. Evaluate results and summarize customer value

Deploy and Configure a Single Node Spectrum Discover Instance

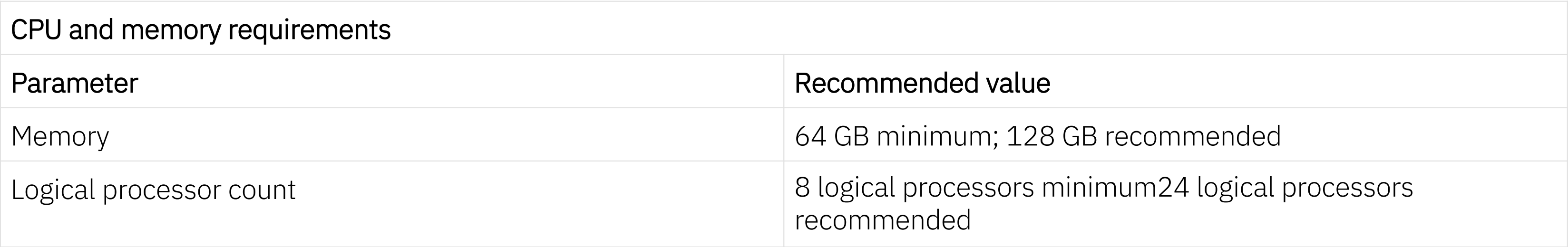
IBM Spectrum Scale

- The IBM Spectrum Discover SW is an OVA (open virtualization appliance) file that is deployed on a VMware ESXi 6.0 or later server by using the VMware vSphere Client.
- The trial and production virtual appliance node requires two additional VMDK storage devices.
- Configure storage for a single node trial IBM Spectrum Discover virtual appliance
- Adding Virtual disk for IBM Spectrum Discover persistent message queues
- Adding a virtual disk for the IBM Spectrum Discover database
- Configuring CPU and memory allocation for the IBM Spectrum Discover virtual appliance
- Configuring networking and performing provisioning of a single node trial IBM Spectrum Discover virtual appliance.
- After the installation completes, log into the Spectrum Discover GUI by entering the host name of the Spectrum Discover appliance in a web browser and provide the following default credentials:
- Username: sdamin
- Password: Passw0rd (with a zero)

https://www.ibm.com/support/knowledgecenter/en/SSY8AC_2.0.0/com.ibm.spectrum.discover.v2r00.doc/ins_deploying_configuring_single_node_trial_single_node_production_ibm_spectrum_discover_virtual_appliance.html

CPU and Memory Requirements - PoC

- A single node **production** IBM Spectrum Discover virtual appliance requires 128GB RAM and 24 logical processors.
- 128GB RAM and 24 logical processors is also recommended for the single node trial IBM Spectrum Discover virtual appliance, but A PoC can be executed with less memory/processors.
- If using 64GB of RAM, no more than 25 million files may be indexed into IBM Spectrum Discover.
- IBM recommends reserving all memory assigned to the IBM Spectrum Discover virtual appliance to avoid running out of physical memory and swapping, which severely impacts database performance and stability.



CPU and memory requirements

Parameter	Recommended value
Memory	64 GB minimum; 128 GB recommended
Logical processor count	8 logical processors minimum24 logical processors recommended

https://www.ibm.com/support/knowledgecenter/en/SSY8AC_2.0.0/com.ibm.spectrum.discover.v2r00.doc/pln_nwrequirements.html

Storage Requirements

Parameter	Recommended value
Base OS SW VMDK	500 GB thick provision, lazy zero SSD / flash
Persistent message queue VMDK	Persistent message queue minimum (without action agent): 50 GB minimum + 1 GB per 2 million indexed files, thick provision, lazy zero HDD or SSD / flash
	Persistent message queue recommended (without action agent): 3.2 TB, thick provision, lazy zero SSD / flash
	Persistent message queue minimum (with action agent): 50 GB minimum + 2 GB per 2 million indexed files, thick provision, lazy zero HDD or SSD / flash
	Persistent message queue recommended (with action agent): 3.2 TB +1 TB per action agent thick provision, lazy zero SSD / flash
Database VMDK	Database minimum (does not include capacity for database backup) 100 GB minimum, 1 GB per 2 million indexed files, thick provision, lazy zero SSD/flash VMDK
	Database minimum (includes capacity for database backup) 100 GB minimum, 2 GB per 2 million indexed files, thick provision, lazy zero SSD/flash VMDK

Networking Requirements

IBM Spectrum Scale

- **The minimum BW - 1 GbE. For source data inspection BW= 10GbE.**
- **Domain Name (FQDN) that is registered in a customer supplied DNS.**
- **Network Time Protocol (NTP) server IP or host name**

Parameter	Value format	Recommended value	Example
<hostname>	host.domain.com	Fully qualified domain name of the node	node.example.com
<interface>	ensXXX	The Ethernet interface to use for the virtual appliance networking	ens192
<ip>	xxx.xxx.xxx.xxx	The IP address of the node	10.10.200.10
<netmask>	xxx.xxx.xxx.xxx	Network mask for the IP range of the node	255.255.254.0
<gateway>	xxx.xxx.xxx.xxx	IP address of the network gateway	10.10.200.1
<dns>	xxx.xxx.xxx.xxx	The IP address of a single DNS server	10.10.200.35
<ntp>	xxx.xxx.xxx.xxxorhost.domain.com	Fully Qualified Domain Name or IP address of NTP server.	Pool1.ntp.org

https://www.ibm.com/support/knowledgecenter/en/SSY8AC_2.0.0/com.ibm.spectrum.discover.v2r00.doc/pln_nwrequirements.html

Storage Optimization

Test cases associated with performing analytics associated with the storage optimization use case

- Leverage the TimeSinceAccess bucketing feature to identify cold data and understand data aging in your environment
- Leverage file and object size bucketing feature to perform analytics on the size distribution of the data
- Leverage File and object data consumed by owner analytics to understand the distribution of data by data owner across heterogeneous storage environments
- Combine file and object grouping criteria for advanced analytics
- Map file and object data to business constructs
- Generate file type distribution capacity showback reports
- Generate default data curation / capacity showback reports
- Generate custom data curation reports from the Spectrum Discover GUI
- Identify potentially duplicate data

It is recommended to schedule a webex session with IBM Spectrum Discover technical support for a transfer of knowledge session when executing the storage optimization POC tasks.

Adding a Connection to Spectrum Scale and Initiating Scans

IBM Spectrum Scale

- The Spectrum Scale scanner uses the policy engine via the mmapplypolicy command to harvest metadata.
- A sudo user with access to this command must be created on the Spectrum Scale source system.
- For optimal performance, Kafka client and python dependencies should be installed on the Spectrum Scale node.
  - It is possible to still scan a Spectrum Scale filesystem without these dependencies, but these dependencies optimize the metadata scan and ingest into Spectrum Discover performance.
- Log in to the IBM Spectrum Discover web interface with a user id that has the data admin role associated with it., Select admin, then select "Add Connection:, then define a Connection Name.
- Click the down arrow for Connection Type to display a drop-down menu and select the connection type IBM Spectrum Scale, then click "Submit Connection"
- Follow the procedures documented in the best practices for scanning Spectrum Scale file systems from the Knowledge Center links below

https://www.ibm.com/support/knowledgecenter/en/SSY8AC_2.0.0/com.ibm.spectrum.discover.v2r00.doc/ins_scale_creating_data_source_connection.html

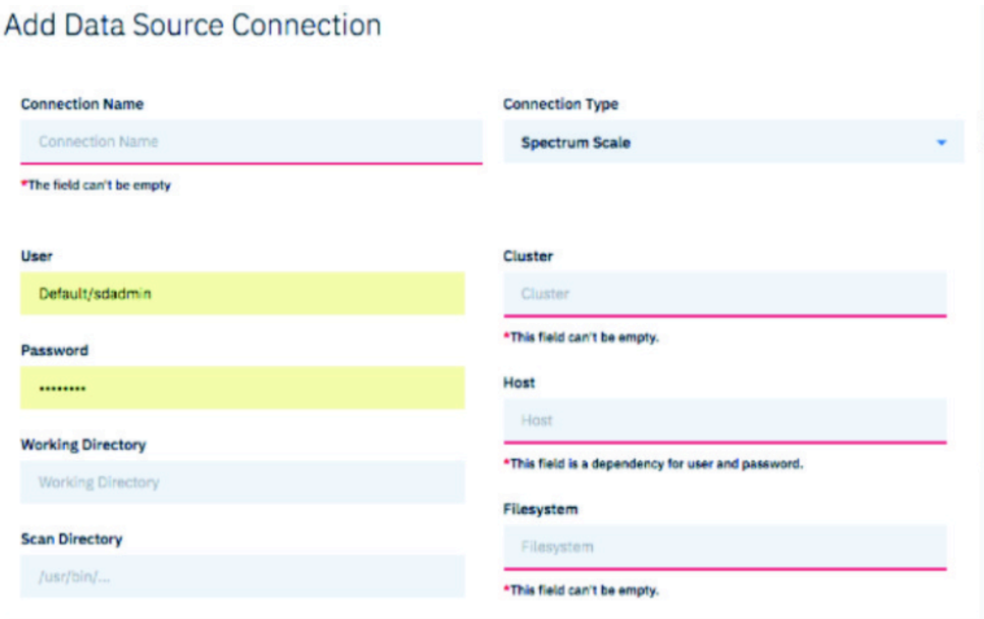
https://www.ibm.com/support/knowledgecenter/en/SSY8AC_2.0.0/com.ibm.spectrum.discover.v2r00.doc/ins_scan_guidelines.html

Adding a Connection to Spectrum Scale and Initiating Scans

IBM Spectrum Scale

- Password - Enter the password for the user id specified in user.
- Working Directory - A scratch directory on the source data system where Discover can put its temporary files.
- Scan Directory - The root directory of the scan. All files and directories under this one will be scanned. Typically, this is the base directory of the filesystem, for example /gpfs/fs1.

- Connection Type - The type of source storage system this connection represents.
- Site - An optional physical location tag that an admin can provide if they want to see the physical distribution of their data.
- Cluster - The Scale/GPFS cluster name, found at: /usr/lpp/mmfs/bin/mmlscluster.
- Host - The hostname or IP address of an IBM Spectrum Scale node from which a scan can be initiated, for example a quorum-manager node.
- Filesystem - The short name (omit /dev/) of the filesystem to be scanned, for example fs1
- Node list -The list of nodes or node classes participating in the scan of theSpectrum Scale file system.



Add Data Source Connection

Connection Name: Connection Name

*The field can't be empty

Connection Type: Spectrum Scale

User: Default/sdadmin

Password: ********

Working Directory: Working Directory

Scan Directory: /usr/bin/...

Cluster: Cluster

*This field can't be empty.

Host: Host

*This field is a dependency for user and password.

Filesystem: Filesystem

*This field can't be empty.

Deploy and configure a single node trial or single node production IBM Spectrum Discover virtual appliance

IBM Spectrum Scale

The IBM knowledge center provides worksheets and step-by-step instructions to deploy IBM Spectrum Discover

- **Deploying a single node trial or single node production IBM Spectrum Discover virtual appliance**
  The IBM Spectrum Discover software is available as an OVA (open virtualization appliance) file. You can deploy it on your VMware ESXi server by using the VMware vSphere Client.
- **Configuring storage for a single node trial or single node production of IBM Spectrum Discover virtual appliance**
  The IBM Spectrum Discover trial and production virtual appliance node requires two additional VMDK storage devices.
- **Configuring CPU and memory allocation for the IBM Spectrum Discover virtual appliance**
  It is required to increase the default allocations of CPU and memory for each IBM Spectrum Discover virtual appliance.
- **Configuring networking and performing provisioning of a single node trial or single node production IBM Spectrum Discover virtual appliance**
  After virtual appliance in the IBM Spectrum Discover is deployed, and storage, CPU, and memory are configured, you need to configure networking and then provision the virtual appliances by using a provisioning tool.

More detailed information can be found at this link:

https://www.ibm.com/support/knowledgecenter/en/SSY8AC_2.0.0/com.ibm.spectrum.discover.v2r00.doc/ins_deploying_configuring_single_node_trial_single_node_production_ibm_spectrum_discover_virtual_appliance.html

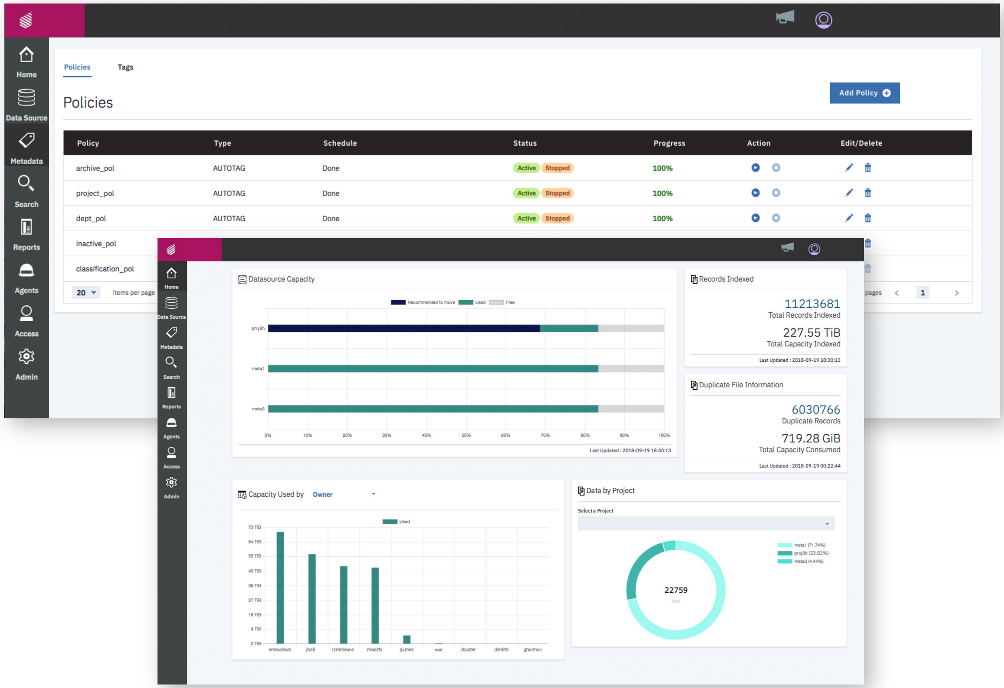
Free Trial Software Download

- 90 Day Free Trial
  - ➢ At end of 90 days, code accessible by client w/approved extension or purchase of a full license
- Full function version of code
  - ➢ Not limited in scale or function set
  - ➢ At termination of trial, access terminates
- Restriction(s)
  - ➢ Cannot upgrade from single node trial to multi-node production
- Support for trial: spdiscov@us.ibm.com

IBM’s metadata management solution is the answer

IBM Spectrum Scale

IBM Spectrum Discover



Policies Tags

Policies

Add Policy

Policy	Type	Schedule	Status	Progress	Action	Edit/Delete
archive_pol	AUTOTAG	Done	Active Stopped	100%		
project_pol	AUTOTAG	Done	Active Stopped	100%		
dept_pol	AUTOTAG	Done	Active Stopped	100%		
inactive_pol						
classification_pol						

20 items per page … pages < 1 >

Datasource Capacity

Last Updated : 2018-09-19 18:30:13

Records Indexed

11213681 Total Records Indexed

227.55 TiB Total Capacity Indexed

Last Updated : 2018-09-19 18:30:13

Duplicate File Information

6030766 Duplicate Records

719.28 GiB Total Capacity Consumed

Last Updated : 2018-09-19 00:22:44

Capacity Used by Owner

Data by Project

Select a Project

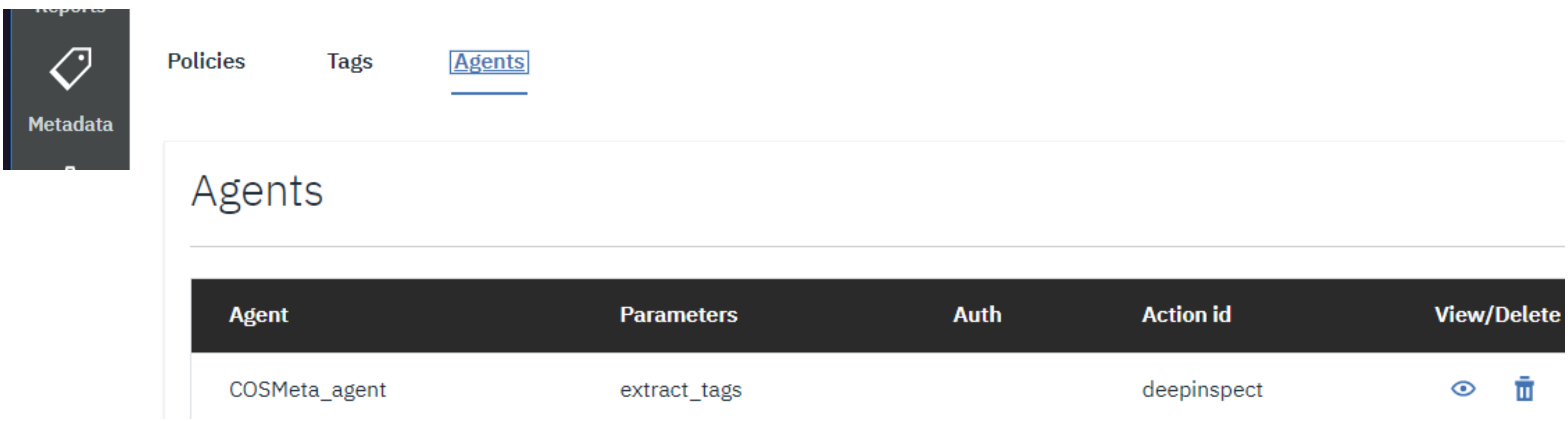
Data Insight for Analytics, Governance, & Optimization

- **Automate cataloging** of unstructured data by capturing metadata as it is created
- **Enable comprehensive insight** by combining system metadata with custom tags to increase storage admin & data consumer productivity
- **Leverage extensibility** using the API, custom tags, and policy-based workflows to orchestrate content inspection & activate data in AI, ML, & analytics workflows

BACKUP

IBM Spectrum Scale

Action Agents



Agents

Agent	Parameters	Auth	Action id	View/Delete
COSMeta_agent	extract_tags		deepinspect	👁 🗑

After filtering the metadata. it will select all the images from any IBM COS datasource and send the files list to the **COSMeta agent** to retrieve the EstimatedValue metadata for each record.

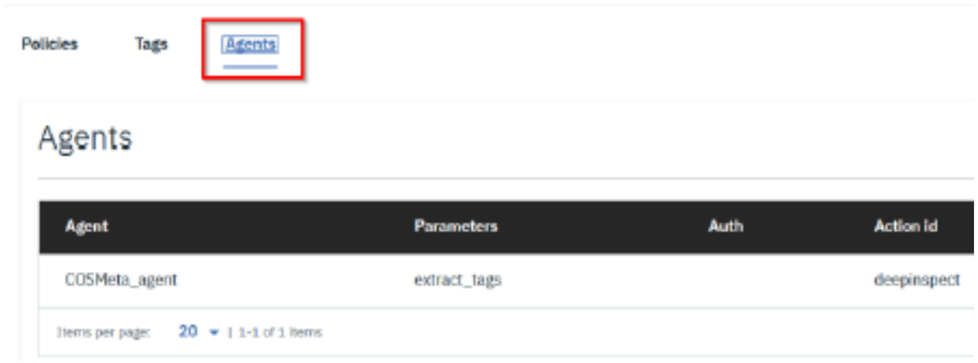
The agent will then extract the requested metadata and send the result back to Spectrum Discover which will apply the value to the record’s tag.

More on Action Agents

This agent is running, (preferably as a daemon/service), on a server which can access the files that will need to be inspected. It is waiting to receive messages from Spectrum Discover with work to do.

- In this case, when asked to work by Spectrum Discover, this action agent will get COS file headers and retrieve the requested S3 metadata. DEEPINSPECT policy allows the external use of an action agent.
- From the Metadata > Policies page, check the **tagging_cos_images_estimated_value**. .

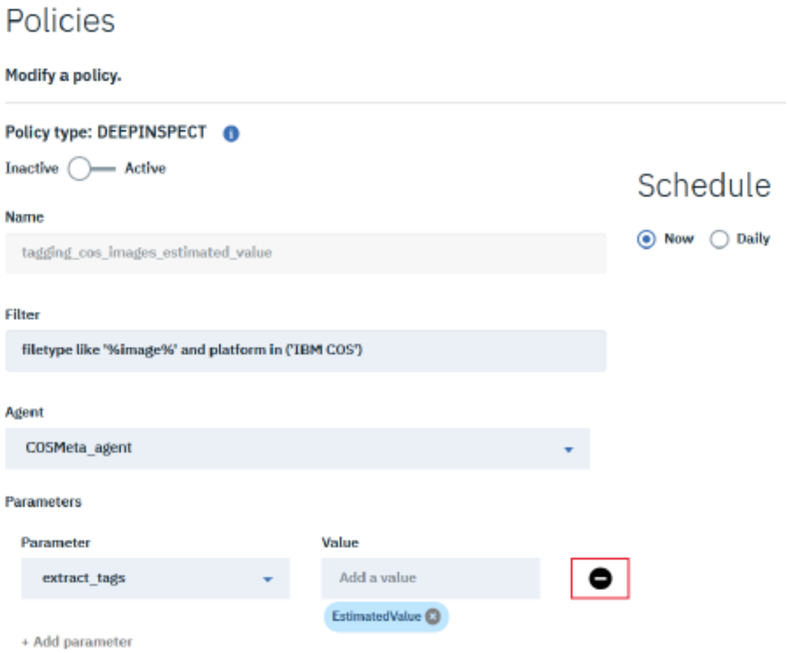
Select Metadata
Then Agents



Agents

Agent	Parameters	Auth	Action id
COSMeta_agent	extract_tags		deepinspect

Items per page: 20 ▾ | 1-1 of 1 items



Policies

Modify a policy.

Policy type: DEEPINSPECT

Inactive ○— Active

Name: tagging_cos_images_estimated_value

Filter: filetype like '%image%' and platform in ('IBM COS')

Agent: COSMeta_agent

Parameters

Parameter	Value
extract_tags	Add a value EstimatedValue

+ Add parameter

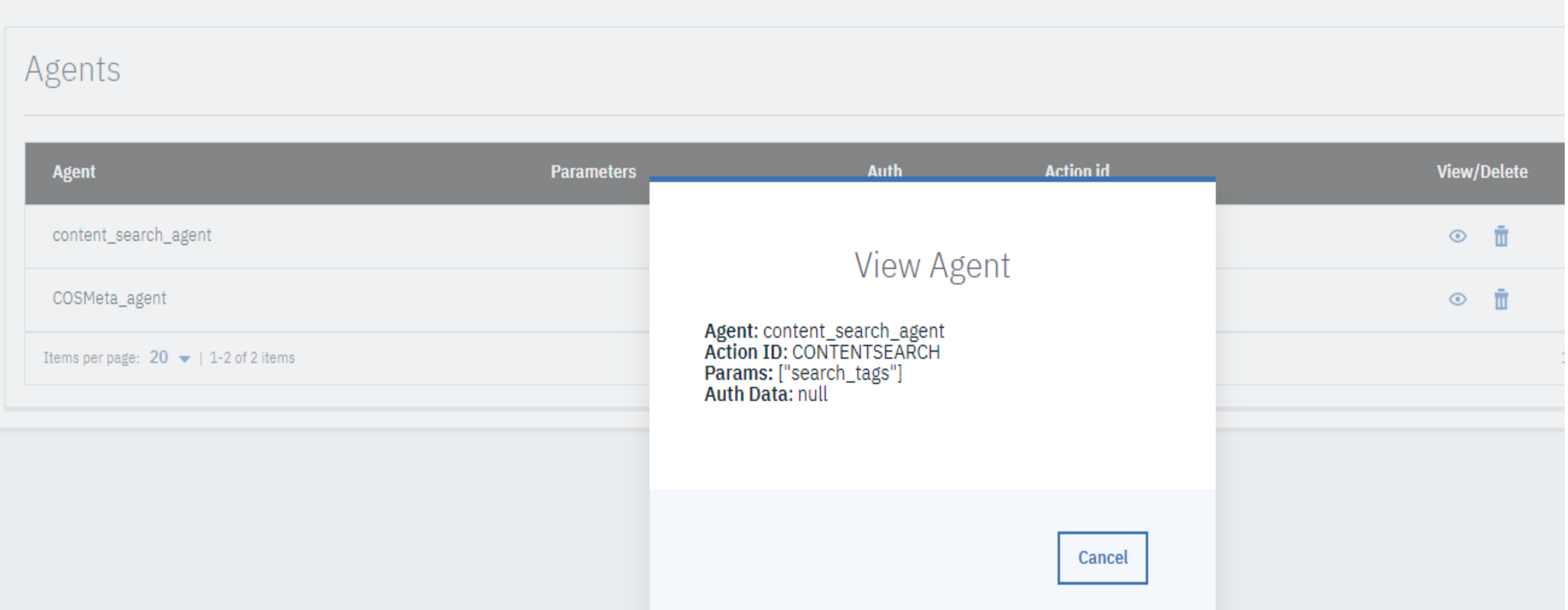
Schedule

◉ Now ○ Daily

Deep Inspect Policies and Action Agents

IBM Spectrum Scale

You can also view the action agent details



Agents

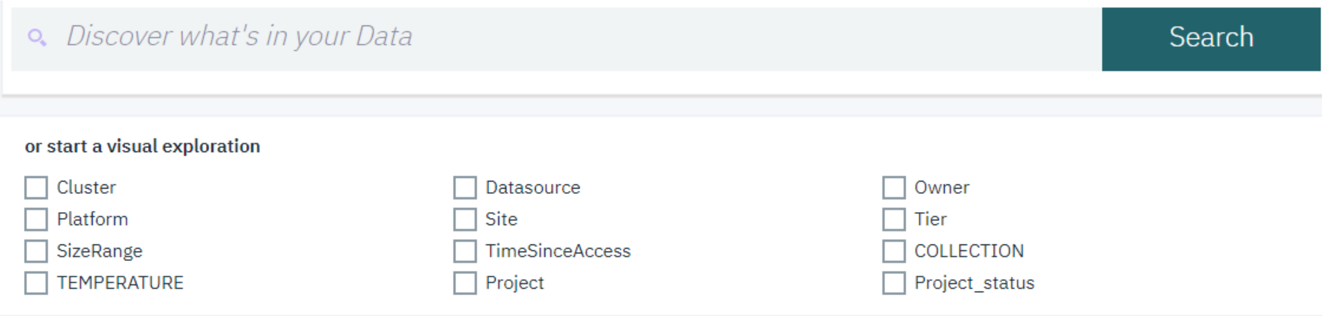
Agent	Parameters	Auth	Action id	View/Delete
content_search_agent				
COSMeta_agent				

View Agent

Agent: content_search_agent
Action ID: CONTENTSEARCH
Params: ["search_tags"]
Auth Data: null

Results – Mine your Data

Search



Discover what's in your Data

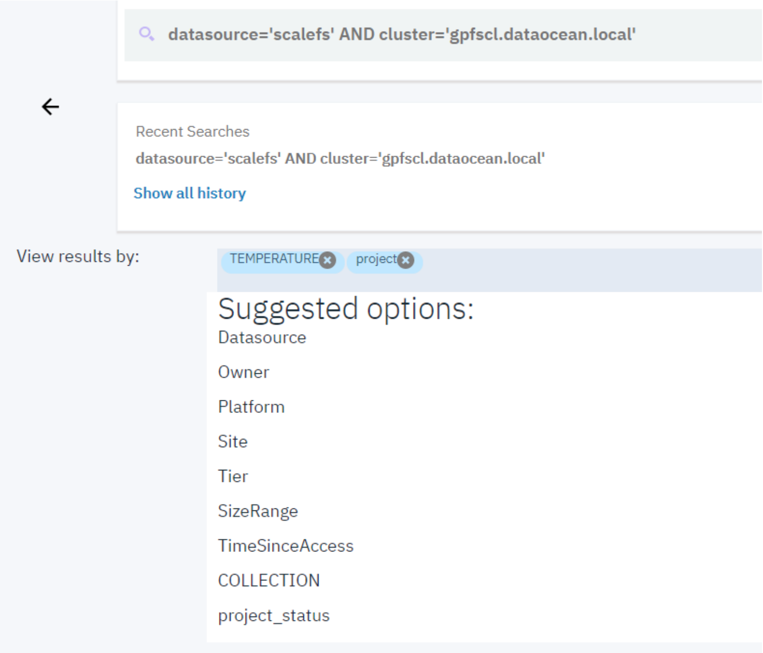
Search

or start a visual exploration

□ Cluster
□ Platform
□ SizeRange
□ TEMPERATURE

□ Datasource
□ Site
□ TimeSinceAccess
□ Project

□ Owner
□ Tier
□ COLLECTION
□ Project_status



datasource='scalefs' AND cluster='gpfscl.dataocean.local'

Recent Searches

datasource='scalefs' AND cluster='gpfscl.dataocean.local'

Show all history

View results by: TEMPERATURE ⊗ project ⊗

Suggested options:

Datasource

Owner

Platform

Site

Tier

SizeRange

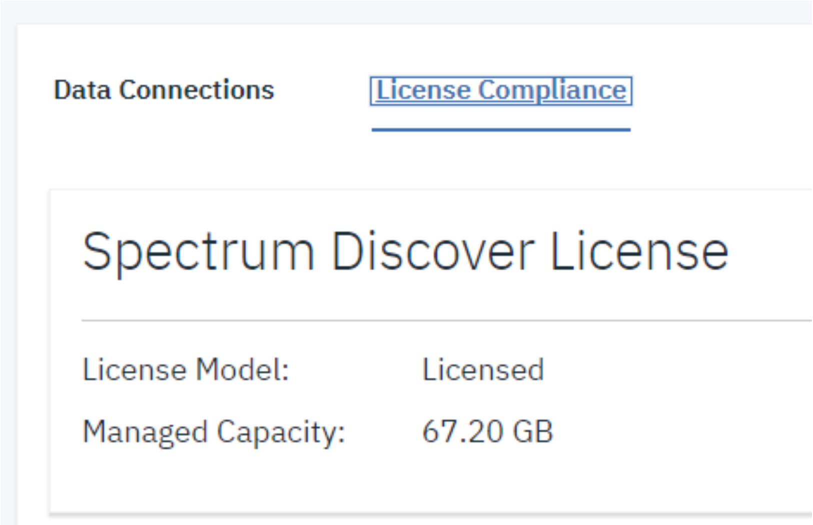
TimeSinceAccess

COLLECTION

project_status

Data Source License

IBM Spectrum Scale



Spectrum Discover License

License Model: Licensed

Managed Capacity: 67.20 GB

Legal notices

Copyright © 2019 by International Business Machines Corporation. All rights reserved.

No part of this document may be reproduced or transmitted in any form without written permission from IBM Corporation.

Product data has been reviewed for accuracy as of the date of initial publication. Product data is subject to change without notice. This document could include technical inaccuracies or typographical errors. IBM may make improvements and/or changes in the product(s) and/or program(s) described herein at any time without notice. Any statements regarding IBM's future direction and intent are subject to change or withdrawal without notice, and represent goals and objectives only. References in this document to IBM products, programs, or services does not imply that IBM intends to make such products, programs or services available in all countries in which IBM operates or does business. Any reference to an IBM Program Product in this document is not intended to state or imply that only that program product may be used. Any functionally equivalent program, that does not infringe IBM's intellectually property rights, may be used instead.

THE INFORMATION PROVIDED IN THIS DOCUMENT IS DISTRIBUTED "AS IS" WITHOUT ANY WARRANTY, EITHER OR IMPLIED. IBM LY DISCLAIMS ANY WARRANTIES OF MERCHANTABILITY, FITNESS FOR A PARTICULAR PURPOSE OR NONINFRINGEMENT. IBM shall have no responsibility to update this information. IBM products are warranted, if at all, according to the terms and conditions of the agreements (e.g., IBM Customer Agreement, Statement of Limited Warranty, International Program License Agreement, etc.) under which they are provided. 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 in connection with this publication and cannot confirm the accuracy of performance, compatibility or any other claims related to non-IBM products. IBM makes no representations or warranties, ed or implied, regarding non-IBM products and services.

The provision of the information contained herein is not intended to, and does not, grant any right or license under any IBM patents or copyrights. Inquiries regarding patent or copyright licenses should be made, in writing, to:

IBM Director of Licensing
IBM Corporation
North Castle Drive
Armonk, NY 1 0504- 785
U.S.A.

Information and trademarks

IBM, the IBM logo, ibm.com, IBM System Storage, IBM Spectrum Storage, IBM Spectrum Control, IBM Spectrum Protect, IBM Spectrum Archive, IBM Spectrum Virtualize, IBM Spectrum Scale, IBM Spectrum Accelerate, Softlayer, and XIV are trademarks of International Business Machines Corp., registered in many jurisdictions worldwide. 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 are trademarks or registered trademarks of other companies.
Adobe, the Adobe logo, PostScript, and the PostScript logo are either registered trademarks or trademarks of Adobe Systems Incorporated in the United States, and/or other countries.
IT Infrastructure Library is a Registered Trade Mark of AXELOS Limited.
Linear Tape-Open, LTO, the LTO Logo, Ultrium, and the Ultrium logo are trademarks of HP, IBM Corp. and Quantum in the U.S. and other countries.
Intel, Intel logo, Intel Inside, Intel Inside logo, Intel Centrino, Intel Centrino logo, Celeron, Intel Xeon, Intel SpeedStep, Itanium, and Pentium are trademarks or registered trademarks of Intel Corporation or its subsidiaries in the United States and other countries.
Linux is a registered trademark of Linus Torvalds in the United States, other countries, or both.
Microsoft, Windows, Windows NT, and the Windows logo are trademarks of Microsoft Corporation in the United States, other countries, or both.
Java and all Java-based trademarks and logos are trademarks or registered trademarks of Oracle and/or its affiliates.
Cell Broadband Engine is a trademark of Sony Computer Entertainment, Inc. in the United States, other countries, or both and is used under license therefrom.
ITIL is a Registered Trade Mark of AXELOS Limited.
UNIX is a registered trademark of The Open Group in the United States and other countries.
* All other products may be trademarks or registered trademarks of their respective companies.

Notes:
Performance is in Internal Throughput Rate (ITR) ratio based on measurements and projections using standard IBM benchmarks in a controlled environment. The actual throughput that any user will experience will vary depending upon considerations such as the amount of multiprogramming in the user's job stream, the I/O configuration, the storage configuration, and the workload processed. Therefore, no assurance can be given that an individual user will achieve throughput improvements equivalent to the performance ratios stated here.

All customer examples cited or described in this presentation are presented as illustrations of the manner in which some customers have used IBM products and the results they may have achieved. Actual environmental costs and performance characteristics will vary depending on individual customer configurations and conditions.
This publication was produced in the United States. IBM may not offer the products, services or features discussed in this document in other countries, and the information may be subject to change without notice. Consult your local IBM business contact for information on the product or services available in your area.
All statements regarding IBM's future direction and intent are subject to change or withdrawal without notice, and represent goals and objectives only.
Information about non-IBM products is obtained from the manufacturers of those products or their published announcements. IBM has not tested those products and cannot confirm the 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.

Prices subject to change without notice. Contact your IBM representative or Business Partner for the most current pricing in your geography.

This presentation and the claims outlined in it were reviewed for compliance with US law. Adaptations of these claims for use in other geographies must be reviewed by the local country counsel for compliance with local laws.

Special notices

This document was developed for IBM offerings in the United States as of the date of publication. IBM may not make these offerings available in other countries, and the information is subject to change without notice. Consult your local IBM business contact for information on the IBM offerings available in your area.

Information in this document concerning non-IBM products was obtained from the suppliers of these products or other public sources. Questions on the capabilities of non-IBM products should be addressed to the suppliers of those products.

IBM may have patents or pending patent applications covering subject matter in this document. The furnishing of this document does not give you any license to these patents. Send license inquires, in writing, to IBM Director of Licensing, IBM Corporation, New Castle Drive, Armonk, NY 10504-1785 USA.

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

The information contained in this document has not been submitted to any formal IBM test and is provided "AS IS" with no warranties or guarantees either expressed or implied.

All examples cited or described in this document are presented as illustrations of the manner in which some IBM products can be used and the results that may be achieved. Actual environmental costs and performance characteristics will vary depending on individual client configurations and conditions.

IBM Global Financing offerings are provided through IBM Credit Corporation in the United States and other IBM subsidiaries and divisions worldwide to qualified commercial and government clients. Rates are based on a client's credit rating, financing terms, offering type, equipment type and options, and may vary by country. Other restrictions may apply. Rates and offerings are subject to change, extension or withdrawal without notice.

IBM is not responsible for printing errors in this document that result in pricing or information inaccuracies.

All prices shown are IBM's United States suggested list prices and are subject to change without notice; reseller prices may vary.

IBM hardware products are manufactured from new parts, or new and serviceable used parts. Regardless, our warranty terms apply.

Any performance data contained in this document was determined in a controlled environment. Actual results may vary significantly and are dependent on many factors including system hardware configuration and software design and configuration. Some measurements quoted in this document may have been made on development-level systems. There is no guarantee these measurements will be the same on generally-available systems. Some measurements quoted in this document may have been estimated through extrapolation. Users of this document should verify the applicable data for their specific environment.

© IBM Corporation