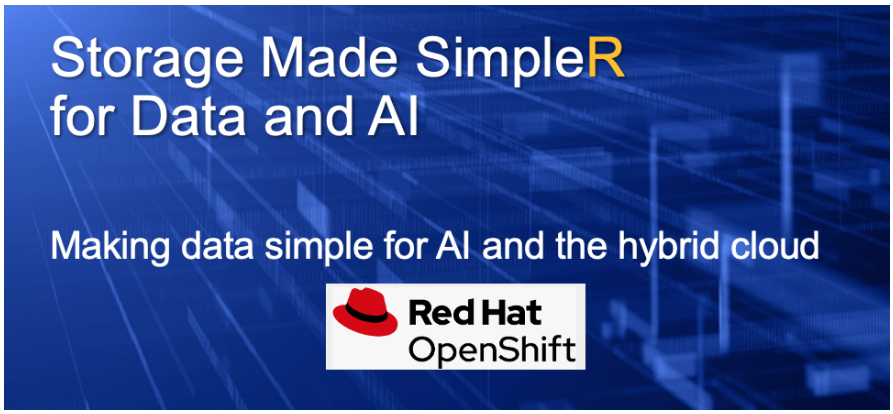
IBM Spectrum Discover with Red Hat OpenShift Container Storage (OCS)

Expand data governance, AI workflows, data organization and data analysis to container native storage

Highlights

- Expand use cases for container native storage
- Continuous data updates to data catalog for real-time analysis
- Leverage policy engine for additional data services
- Auto tags and custom tags make AI data set selection easy
- View details of container data to improving data management
- Search for content patterns to improve data governance
- OCS data is linked to IBM Watson data services with Spectrum Discover



Storage made simpleR

Modern data architectures are substantially more complex than past generations of data management. The complexities begin with the data itself. Data is not all structured and easily organized with relational concepts. Much of the data is not SQL-based and sometimes not SQL compatible. It is not all slow-moving data where batch processing is sufficient. It is not entirely internal data that is known and within the span of the enterprise's control. It is not able to meet the needs of all use cases with a one-size-fits-all schema such as a data warehouse. External data and big data sources can be filled with governance or security issues. Some of these issue occur when the data contains unanticipated Personally

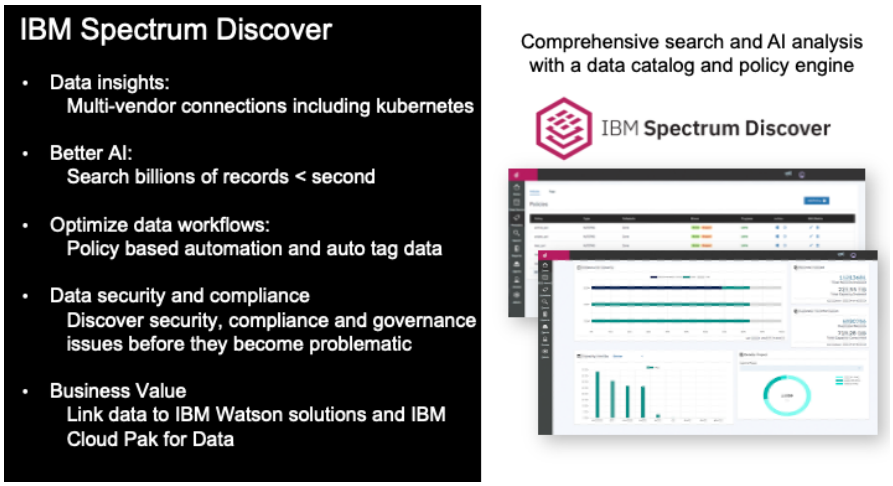
Identifying Information (PII), Payment Card Information (PCI), Protected Health Information (PHI), or other security and privacy sensitive data.

With multiple types of users and use cases for today's data resources—data scientists, data analysts, and business users and analysts many times need a self-service data friendly environment. They need to find data for analysis, understand the data, evaluate its suitability to their purpose, and access the data that they need—all without compromising the privacy of protected data.

Now comes Spectrum Discover

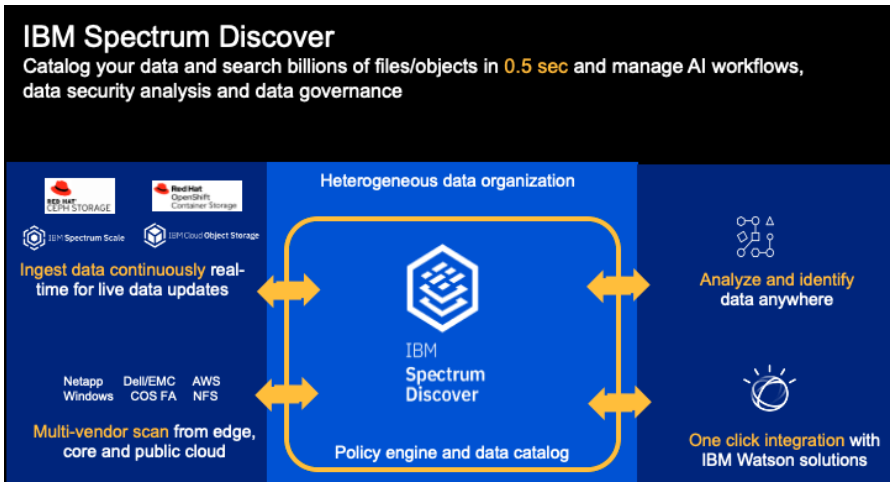
IBM Spectrum Discover

Spectrum Discover can be summarized as a data catalog and policy engine to organize the AI infrastructure and help solve the data and AI puzzle faster. Spectrum Discover is the best solution when customers require a solution to optimize the AI workflows or manage data governance or provide insight and data analysis for IT or data users.



IBM Spectrum Discover

IBM Spectrum Discover is a multi-source data catalog that automatically and continuously indexes objects and files whenever changes are made using the metadata in real-time. The result is a powerful and customizable database with a user-friendly interface that allows users to locate and identify the most relevant data regardless of its type or location. Using either a simple SQL query command or actionable API scripts or commands, users are empowered with comprehensive insight into the data in a fast and efficient manner. Spectrum Discover can also be used to create custom tags and policy-based workflows to orchestrate content inspection and activate data in artificial intelligence (AI), machine learning (ML), and analytics workflows. Spectrum Discover can be used for faster AI analysis, compliance classification, image and video indexing, identifying personal data, AI data pipeline integration, real-time data discovery, and providing new insights to optimize data and find bad or duplicate data. Data sources include IBM Spectrum Scale, IBM COS, AWS S3, NFS or SMB data sources including Netapp and Isilon and Red Hat (Ceph and OCS).



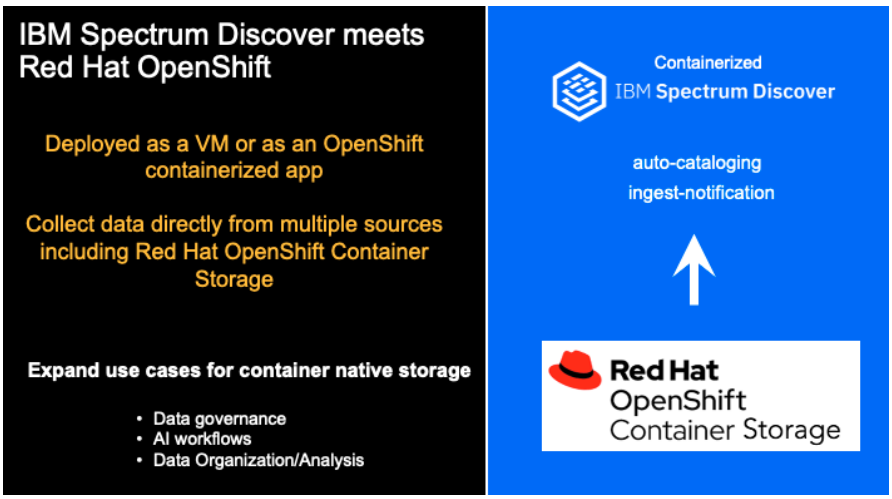
A comprehensive storage data catalog

Spectrum Discover can ingest from multiple types of unstructured data sources. Spectrum Scale and ESS, IBM Cloud Object Storage (IBM COS), and Red Hat Ceph external and Red Hat OpenShift Container Storage (OCS) can all ingest data **in real-time**, which means that a data scan only needs to occur initially and all subsequent changes (writes, updates or delete) to the data are automatically updated in Spectrum Discover. This creates a very powerful tool for analyzing data in real-time.

Spectrum Protect and Spectrum Archive are two IBM backup and archive applications that have been recently updated to enable integration with Spectrum Discover.

Finally, Netapp, Dell EMC Isilon (PowerScale), and Amazon Simple Storage Service (S3) along with any SMB storage can all be scanned by Spectrum Discover to provide a multi-cloud multi-storage metadata repository.

IBM Spectrum Discover can also be used to automatically export data to IBM Watson solutions to leverage all the unstructured data for new insights using IBM CLoud Pak for Data and IBM AI tools and analysis.



Containerized IBM Spectrum Discover

IBM Spectrum Discover can be deployed as a VM or as a containerized application. Deploying in a container also brings Red Hat OpenShift OCS storage into the data catalog. This containerized solution helps bridging the gap with enterprise storage and Red Hat OpenShift and allows IBM Spectrum Discover to expand use cases for container native storage and the expanding growth of

containerized applications. Now Red Hat OCS can leverage the power of IBM Spectrum Discover for data governance, AI workflows and data organization and analysis.

Why IBM?

IBM Storage for Data and AI is more than storage products or even storage solutions. It consists of a storage strategy to help customers on their journey to AI and the hybrid cloud data center. IBM continues to drive leadership for scalable high-performance workloads as well as efficient, secure, scalable, capacity storage for file and object-based solutions. Our products provide an enhanced strategy for AI and the hybrid cloud. We provide a foundation for the future for the edge, the core data center and the public cloud including kubernetes containers and the Red Hat OpenShift platform. IBM Storage for Data and AI lowers complexity and cost with increase integration to an AI information architecture for the hybrid cloud that can be infused to the entire organization. Our message is easy “Storage Made SimpleR for Data and AI”

For more information

For more information please visit: https://www.ibm.com/products/spectrum-discover