Fragmentation Results in Data Sprawl and Lack of Visibility into Enterprise Data

Today's fragmented approach to data storage has made it difficult to establish a complete picture of all the information stored across different data centers. With distinct silos of data for different operational workflows, including primary or tier one applications, data protection, and test and development, organizations often end up with multiple copies of the same information in different locations and formats. Vendors benefit from the vast sprawl of data created by numerous workloads running across different environments, but these silos create enormous challenges for companies that want to run meaningful analytics across the organization.

Separate silos create data sprawl, hindering the ability to perform comprehensive analyses across these isolated infrastructure stacks. In order to extract meaningful information from their data, organizations are forced to migrate data from their current infrastructure stack to an isolated analytics cluster, resulting in another infrastructure stack that must be managed and supported. Due to the increased cost, complexity, and overhead of multiple isolated infrastructure stacks, enterprises are forced to seek out alternative solutions to consolidate and increase their visibility into data.

Turn Data into Information with Cohesity Analytics

Secondary storage, which consists of use-cases like data protection and file services, already have mounds of raw data that can be mined for useful information. However, in most cases, it is sitting on passive infrastructure, largely unstructured, and waiting idle until something needs be restored. Rather than maintaining expensive secondary storage solutions to be used as an insurance policy, Cohesity Analytics can unlock the vast potential of this data and provide meaningful information back to the organization. Cohesity Analytics is powered by Cohesity's Indexing Engine, which runs against all data stored on the platform. This index can immediately provide some very interesting high-level storage information:

Storage Metrics: Detailed storage-level information, including storage utilization, available capacity, and data growth trend analysis.

File-Level Metrics: Comprehensive breakdown of file-level information, such as file-type and user access history to understand how secondary storage is being used in a particular environment.

VM-Level Metrics: Dashboards that show storage consumption by application and data change rates to better anticipate future storage needs.

The Power of Built-In MapReduce

MapReduce is a powerful way to efficiently process large sets of data on a cluster. Cohesity, the industry's first web-scale, intelligent secondary storage company, has natively integrated MapReduce into its OASIS (Open Architecture for Scalable, Intelligent Storage) operating environment to enable the efficient and quick processing of large amount of data. Cohesity Clusters make it possible for even the most complex MapReduce computations to run natively within the Cohesity Data Platform, never requiring external compute or data migrations that are typical with Hadoop-like analytics implementations.

Leveraging the power of Cohesity's Indexing Engine, finding the needle-in-the-haystack of data becomes very simple. Cohesity is able to generate detailed storage, file-level, and VM-level reports on the fly to remove unnecessary complexity of having to leverage an external business intelligence (BI) tool to extract information from data. To provide immediate value, queries, which are used to generate these comprehensive reports, run in the background as data is streamed onto the Cluster, delivering real-time insights into the data that's being collected.

Key Benefits

- In-place analytics running directly on the Cohesity Data Platform
- Native storage, file, and VM-level metrics and trend analyses for capacity planning
- Create vertical specific apps with Cohesity Analytics Workbench
- Quickly derive intelligence from raw data with just a few clicks

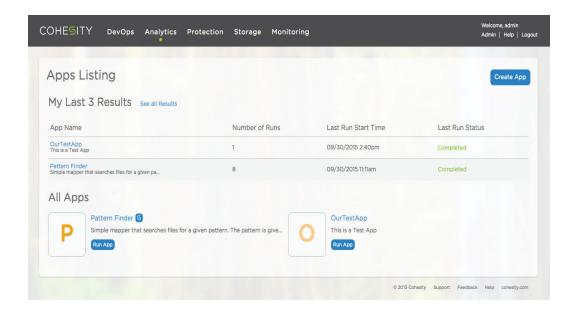## One Platform. Infinite Possibilities.

COHESITY

COHESITY

Supercharge and Customize Analytics with Cohesity Analytics WorkBench

Further extending the native analytics capaibilities of the Cohesity Data Platform, Cohesity Analytics Workbench (AWB) opens up the platform for customized analytics that can be tailored to fit unique business needs. AWB allows users to upload and execute custom code to efficiently and quickly process large datasets across a Cohesity Cluster. AWB jobs run in the background and can be prioritized in conjunction with other concurrent workloads using Cohesity's built-in Quality of Service (QOS) capabilities. AWB jobs can be scheduled to run on pre-defined intervals or on a one-time only basis.

With Cohesity Analytics Workbench, organizations have the ability to leverage their secondary storage solutions to become information hubs. AWB enables businesses across different industries to build tailored Apps that provide the reporting and analysis that fits their individual needs. Examples of how AWB is being used across different verticals include:

eDiscovery: Rapid content analysis to find relevant case information for legal requests or holds.
Compliance: Ensure compliance with Personally Identifiable Information (PII) requirements with Cluster-wide content scans for names, phone numbers, and credit card information that may have been stored in clear text.
Threat Analysis: Log correlation across disparate packet capture solutions to identify potential threats or origin of a security breach.

Leveraging the Past to Change the Future

Cohesity Analytics is a simple and effective way for businesses to transform their idle secondary storage into a powerful tool to derive insight from data. The combination of Cohesity Analytics and Analytics Workbench gives organizations the flexibility to extract detailed information via built-in or custom Apps that are tailored to their individual needs. With Cohesity, businesses can eliminate data sprawl, widen the visibility into their data, and get more from their secondary storage.

@cohesity