Accelerating Data Growth Necessitates Better Data Management

Silverton Consulting, Inc. StorInt™ Briefing

Silverton Consulting
Strategy, Storage & Systems

Introduction

Today, governments around the world are scrutinizing organizational data practices like never before. Ransomware is also running rampant throughout the world, hitting local governments, utilities, universities and financial services firms throughout 2020. Ransomware targets data because it represents the keys to the IT kingdom. Without data, IT may as well close shop. As a result, many organizations have no choice but to pay for access (or recover from backups). And as we encounter the first pandemic in a century, we are finding that data about the infection, transmission, and progression are essential to help reduce the toll of this and future pandemics.

Industry analysts estimate 40 ZB (10**21 bytes of) data has been created thus far in 2020[1].The ways in which IT manages, controls, and secures data has become more critical than ever before to organizations around the globe. Commvault® has business solutions that IT can use to increase storage efficiency, enable faster and more streamlined responses to compliance requests and reduce data risks with analytics, reporting, and search capabilities.

The state of data today

New applications generate and consume more data than ever before. Data growth really began with big data analytics used to gain insight from customer experiences and corporate practices. Then when mobile took off, data generation grew exponentially. In addition, AI, machine learning (ML) and deep learning (DL) devour considerable amounts of data to create accurate classification and recommendation engines. The combination of data analytics, AI-ML-DL and mobile has given rise to massive data growth. Furthermore, exponential data growth far and beyond what we've seen to date will likely occur when IoT begins widespread deployment.

Cyber security remains a serious concern. Anti-virus software offered a respite from an onslaught of desktop attacks, but then threats just migrated to other targets in IT. Keeping systems and applications updated helped, but phishing attacks continue to target humans. Zero-day exploits will also expose vulnerabilities in software or hardware, requiring constant vigilance and ongoing analysis of data and systems.

Then the pandemic hit, and the world's workforce transformed. According to an October Gallup poll, 33% of U.S. workers are now always working remotely and another 25% are sometimes working remotely.[2] As a result, supporting millions of new offsite workers with corporate data and applications has become a critical new IT requirement.

Another impact of COVID-19 was the destabilization of existing business models. People stuck at home shopped online more often. While this shift has been great for scalable online businesses, businesses reliant on person-to-person, brick-and-mortar sales saw revenues plummet. The pandemic's business model disruption ultimately led to budget freezes and cuts and corporate downsizing. At times like these, effective data management can be vital to an organization's present and future success. When IT knowledge and experience walk out the door, ensuring unencumbered access to corporate data and applications is essential for corporate survival.

---

[1] https://www.linux-magazine.com/Online/News/Digital-Universe-Projected-to-Reach-40ZB-by-2020
[2] https://news.gallup.com/poll/321800/covid-remote-work-update.aspx

Silverton Consulting
Strategy, Storage & Systems

Against this backdrop, organizations have faced unrelenting pressure to deploy applications and services on the cloud. While the cloud supports ready scalability and on-demand services, IT and other groups within a business must proceed with care in the use of cloud offerings in order to retain corporate control over cloud data and applications. Running applications on the cloud doesn't absolve IT from managing and controlling data that resides there.

Finally, European, Asian, and other state regulations regarding how data is used, protected and transferred came about due to long-standing and substantially differing views on data privacy around the world. Dealing with separate regulatory environments regarding data privacy has become and will remain a price to be paid for doing business across regions.

Data vectors

Four vectors are driving IT data use today:[3]

1. **Data volume**: The exponential growth of organizational data makes the management of that data extremely important.
2. **Data performance**: Data access and performance often define an offering's customer experience as well as help determine the ultimate success or failure of new applications or services.
3. **Data availability**: Continued data availability becomes yet another crucial factor impacting corporate revenues. When systems and applications go down, customers go elsewhere.
4. **Data risk**: An organization's misuse of data can result in fines, prosecution or, at a minimum, a tarnished public image. Around the world, litigation is increasingly commonplace, and data and services can undergo intense scrutiny by almost any legal entity.

Data Intelligence with Commvault

Commvault offers a number of solutions ranging from Commvault Complete™ Data Protection, a backup, recovery and disaster recovery solution, Commvault HyperScale™ X, a scale-out backup and recovery appliance, Metallic™, a SaaS backup and recovery service, Hedvig®, a software defined storage system, Commvault eDiscovery & Compliance, a compliance solution, Commvault Data Governance, a sensitive data management solution, and Commvault File Storage Optimization, a storage efficiency solution.

Commvault eDiscovery & Compliance, Commvault Data Governance and Commvault File Storage Optimization solutions can help bring order to the chaos of an organization's data or information environment. These three different solutions can operate independently or can be used together to make data readily visible, more manageable and less risky wherever it may be stored.

---

[3] A vector indicates an amount or magnitude, as well a direction of a force.

Silverton Consulting
Strategy, Storage & Systems
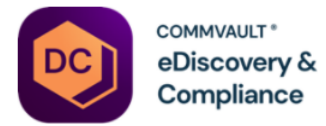
Commvault eDiscovery and Compliance (DC) is for organizations facing periodic litigation or working in environments tightly controlled by government regulations. Litigation has become a cost of doing business worldwide.

DC can help organizations manage litigation better by providing legal self-service portals to generate custodial data sets that retain and preserve relevant information. DC uses advanced content indexing and search capabilities to discover data of interest. Further, DC can help organizations manage and implement government-defined data subject rights and obligations over file data by accessing online file data and data residing in an organization's Commvault backup sets.


Commvault Data Governance (DG) is for organizations that have valuable and sensitive data. An organization's valuable data needs to be protected, secured and made continuously available. Sensitive data goes one step further, requiring restricted use of and access to that data over its lifetime. Controlled data falls under regulatory or legal compliance regimens and must be carefully used.

DG can help better manage different classes of data by indexing the contents of all file data. With an index in place, DG can help admins assess the value, sensitivity and controllable characteristics of data flowing through and stored by IT. DG can also help organizations control data by creating workflows that can be used to identify and classify incoming data to determine its handling requirements. DG works by accessing online file systems and optionally using Commvault backup sets.

Commvault File Storage Optimization (FSO) is for organizations that are struggling with data growth and overwhelmed by storage costs. With FSO, admins can easily see which data is consuming which storage and whether or not that data has recently been accessed. By using FSO, organizations can better control and manage their data in order to minimize storage consumption, protection overhead and administrative effort. FSO depends on metadata provided by all online file data. Where optionally available, it can also scan Commvault backup sets.

## Integrated solutions that work well together
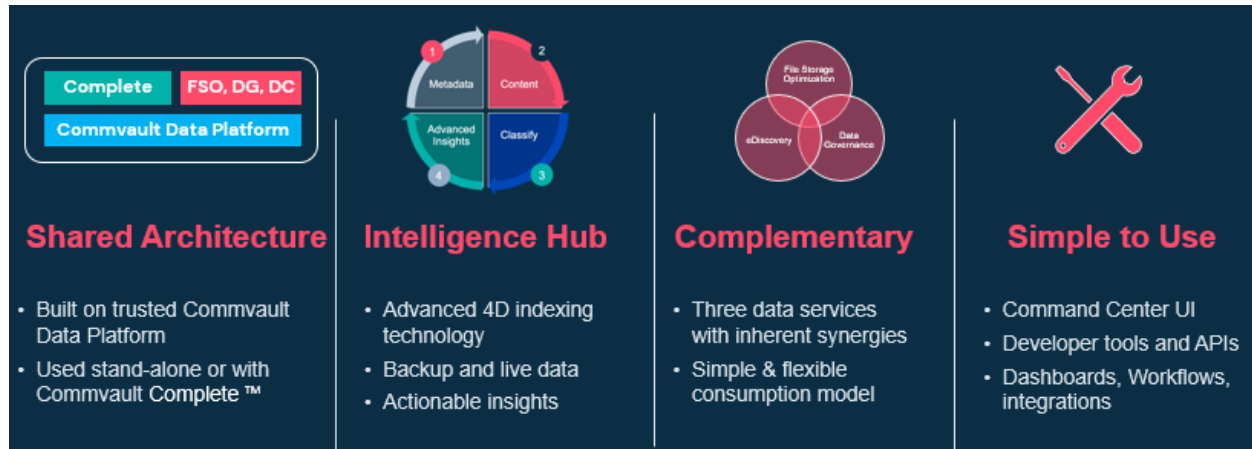
Commvault DC, DG and FSO are all built on top of the **Commvault Data Platform**, which has formed the foundation of the Commvault Complete Data Protection suite for more than two decades. These solutions perform **4D indexing of data**, including metadata, content index, data classification, AI-driven insights and other data platform services. To index data, these solutions draw upon Commvault technologies that have been developed and vetted over thousands of customer deployments. Organizations that deploy all three solutions can effectively optimize and continuously control their data environment.

Furthermore, all solutions, including Commvault Complete, operate under Commvault's **Command Center** GUI, which enables admins to protect, recover and manage their organization's data from one pane of glass. Moreover, DC, DG, FSO and Complete all offer APIs and software development tool kits for organizations that want to automate control over these services. These solutions can also be easily integrated into industry-standard operations automation systems to provide automated workflows to manage data.

Examples of Commvault Data Intelligence in operation

Commvault can help organizations manage data, as shown in the three examples below.

File Storage Optimization
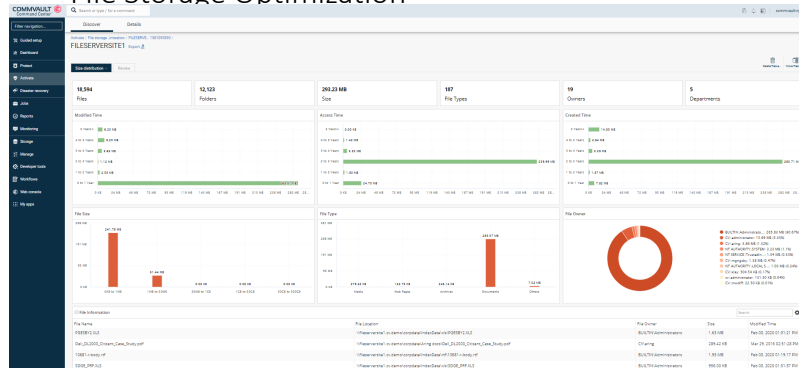
An **oil and gas company** had 10 PB of live data, hindering IT's ability to reduce costs and create new services. The company's initial goal was to better manage 1B files and reduce the $12M/year spend on data storage. With Commvault Activate FSO, they were able to archive 30% of their primary data to secondary storage, reducing storage costs, protection expenses

and data management time and saving an estimated $2.7M.

A **financial services company** managed 25 PB of market data used to drive financial models and algorithms. By using DG, they were able to create a new workflow where data stewards could determine whether massive amounts of incoming data needed to be retained, archived or jettisoned. The company has estimated that they will reduce their online storage from 25PB of active data down to 10PB by year's end.

Data Governance Dashboard: Risk, Sensitive data
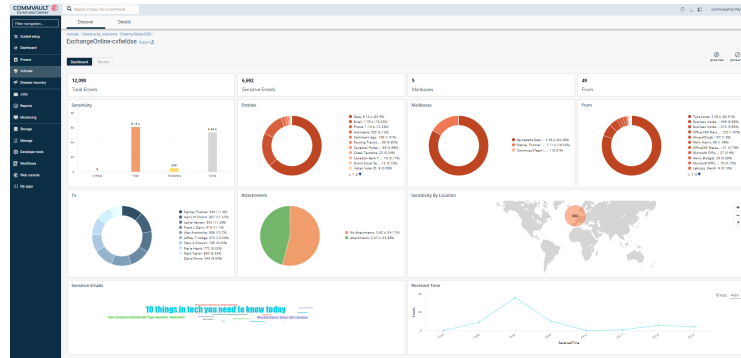
EDiscovery
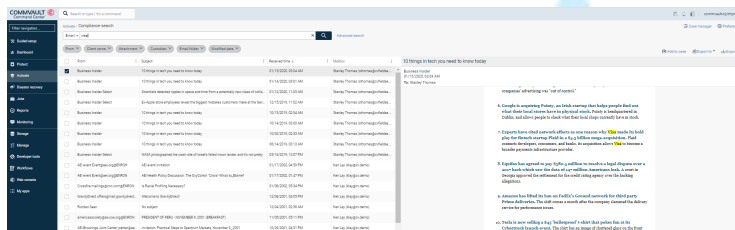
A **global consulting firm** was reaching end of lease on thousands of laptops. They operate worldwide, but the end of lease was occurring mostly under the jurisdiction of the EU's General Data Protection Regulation (GDPR). The firm wanted to retain relevant laptop information by migrating it to the cloud. In addition to the end-of-lease action, the firm was also under eDiscovery and GDPR data subject requests. By using DC, they were able to handle the immediate end-of-lease challenge by migrating only important data to the cloud. Moreover, with all this data indexed, cataloged and searchable, they were able to speed up their response to GDPR data subject and eDiscovery requests.

Summary

Today, the vast majority of ZB of data resides in data centers, and the amount of data generated, along with associated risks, will undoubtedly increase as analytics, AI-ML-DL, mobile applications and IoT continue to take off in coming years. Commvault Data Intelligence solutions are meeting a critical need by helping organizations better manage their data.

The choice is clear: IT can either throw more storage, admins and money at the growing data problem or bring in tools like Commvault eDiscovery & Compliance, Commvault Data Governance and Commvault File Storage Optimization to provide more proactive and effective organizational data management.

Silverton Consulting, Inc., is a U.S.-based Storage, Strategy & Systems consulting firm offering products and services to the data storage community.