

## **IBM Storage: Scaling Up and Scaling Out for AI Analytics**

**By Jean S. Bozman**

IBM's Storage group is focusing on the fast-growing AI and data analytics market to continue its revenue growth for flexible configurations of storage systems – and IBM Spectrum storage-management software.

These technical bets appear to be paying off. The IBM Storage group, along with IBM's zSystems group, showed revenue growth in 2Q20, the most recent financial quarter reported by IBM. Storage revenue grew more than 2 percent year-over-year. This is the third consecutive quarter that IBM Storage has grown year-over-year.

### **IBM Storage and Software for AI and Data Analytics**

IBM intends to build on that storage growth, linking it with the rapid rise in customer adoption of AI, Big Data, and automation software. In a series of [July 2020 announcements](#), IBM Storage focused on AI and data analytics workloads.

The reason is clear: AI is a fast-growing market segment, and customers are turning to data analytics to mine large data repositories to find – and use – actionable data. Customers in enterprise data centers and cloud customers (CSPs, MSPs, co-los and and hosters) are all potential buyers for the IBM Storage system and software solutions.

IBM placed its new products in the context of a progression toward AI that leads to actionable business outcomes. The steps are labeled as: Collect, Organize, Analyze and Infuse, something that IBM calls the AI Ladder.

The IBM ESS 5000 and Spectrum Scale address the data collection-process; the IBM ESS 3000 – announced in October 2019 - addresses the analytics process. The Infuse step refers to business customers increasingly leveraging data analytics to find patterns in the data, leading to insights and better business outcomes. For AI purposes, the IBM ESS 5000 can be co-installed with the IBM Elastic Storage System 3000 (IBM ESS 3000), as they both run IBM Spectrum Scale storage software.

As the last step in the AI ladder of related business processes, data scientists work with the metadata to perform analytics with the relevant datasets. They can use a variety of software tools to do that analysis, often choosing to visualize the data to speed pattern recognition.

### **A Closer Look at the Announcement**

IBM's scale-up and scale-out storage portfolio represents a business strategy that supports a wide range of use-cases, providing a broader base for future growth. IBM makes and sells the IBM ESS series of systems solutions – and has extended IBM Spectrum storage-management software to support the IBM products – and more than 550 storage devices and arrays worldwide.

For systems, IBM Storage is fielding HDD-based storage arrays that collect data for rapidly growing data-stores (IBM ESS 5000) – and pairing them with IBM ESS 3000 arrays that organize data for data-analytics. It enhanced IBM Spectrum storage-management software for ease of use, and support for mixed-vendor shops.

The IBM Elastic Storage System (IBM ESS 5000) is a high-capacity storage array built with hard-disk-drive (HDD) devices, supporting more than 450 drives in a single system. The result is a storage system that can support PB to EB of data – all within a single physical space – and a single, global name-space for searching.

In recent years, high-capacity hard-disk-drives (HDDs) have gotten less notice than all-flash arrays with flash SSDs and high-speed NVMe-based interconnects. But HDDs – known for their high-capacity, high-density drives, plug into very high-capacity, rackable storage devices, making it possible to pack Exabyte-levels of data into highly dense rack-based spaces.

That tells a story that is attractive to customers who are working to reduce their operational expenses (Opex) by becoming more efficient. The high-capacity storage, housed within a unit that takes up less floor space in a data center, will interest customers looking to build up large data lakes, inside a corporate data center.

## **The Software for Data Management**

### ***IBM Spectrum Scale***

IBM Spectrum Scale (formerly called GPFS for its parallel file system) is used to collect large datasets as customers build large data lakes. The IBM ESS 5000 ships with a new version of IBM Spectrum Scale software that has an enhanced accelerator for faster data movement. This allows object storage to be moved and handled by a single global namespace for simplified navigation.

Additionally, answering customer requests for more ease-of-use features, IBM Spectrum Scale (formerly called GPFS for its parallel file system) has new GUIs and wizards – for storage admins who are configuring the software.

### ***IBM Spectrum Discover***

Data scientists work with the metadata to perform analytics with the relevant datasets. They can use a variety of software tools to do that analysis, often choosing to visualize the data to speed pattern recognition.

IBM Spectrum Discover is used to locate, identify, and prepare data for data-analytics processing. With these announcements, Spectrum Discover gained broader support for a wide variety of storage devices, including those made by storage competitors Dell/EMC and NetApp.

IBM Spectrum Discover now supports IBM OpenShift containers (formerly called Red Hat OpenShift). This aligns with customer demands to combine file-based data with cloud-native object-based data. IBM Storage customers can use either VMs or containers – or both.

### ***IBM Cloud Object Storage (IBM COS 3.1)***

IBM Cloud Object Storage (COS) gets a new release, 3.1. This is a software-only update that supports exabyte scale. It is available as a free upgrade to current IBM COS customers, provided they have paid maintenance.

## **Key Takeaways**

IBM Storage has seen its revenue grow, year over year, for three quarters in a row. Now, as a contributor to IBM's overall revenue growth, IBM Storage plans to optimize its product portfolio for use in fast-growing AI deployments. This growth is linked to enterprise data centers – and to cloud providers, including CSPs, MSPs, colos and hosters.

Looking forward, IBM realizes that AI, data analytics and ability to automate the metadata sorting of large data repositories will become an almost universal need for large companies that want to create a unified view of their actionable stored data. Combining its storage portfolio with support for Red Hat OpenShift containers and multiple non-IBM storage products will widen IBM's appeal as a storage provider.