**In the pursuit of digital transformation, organizations focus on the creation, retention, and analysis of data. With much of the data unstructured, including video and video-like images, organizations must deploy flexible and cost-effective storage infrastructure environments capable of keeping up with growing data sets.**

# Managing Unstructured Data Growth Requires a Fresh Approach

*January 2020*

**Written by:** Amita Potnis, Research Director, Infrastructure Systems, Platforms, and Technologies Group

The digitization of business and life means that more data is injected into our day-to-day business workflows and personal lives. As organizations pursue digital transformation, they are reshaping how business gets done — from how they develop and market products to how they interact and retain customers. The key to this digital transformation is data creation, retention, and analysis to drive insights — and much of this data is unstructured.

Sensors, Internet of Things (IoT) devices, and real-time analytics increase the continuous capture, analysis, and delivery of data for the purpose of driving personalization, efficiency, and 24 x 7 operations. While data is exceedingly valuable, its growth is a serious concern for IT organizations charged with storing it. IDC's research highlights just how much data growth is taking place. IDC's *Data Age 2025* study forecast that 80% of data will be video or video-like by 2025. IDC's *Worldwide Global DataSphere IoT Device and Data Forecast, 2019–2023* projects that there will be 79.4ZB of data created by connected IoT devices by 2025, growing from 13.6ZB at the start of the forecast period. Almost 90% of the data generated by IoT devices in 2018 was generated by video surveillance applications through a growing number of Smart Cities projects. However, other categories such as industrial and medical applications are expected to generate more data over time.

IDC's *Worldwide Global DataSphere Forecast, 2019–2023* indicates that content related to entertainment (creation, production, distribution, and consumer consumption) continues to be the largest category of data within the Data GlobalSphere; from 2018 to 2023, entertainment-related content will have a CAGR of 21% and will represent 35% of the overall 102ZB of data. While the installed base of storage capacity is growing at different rates based on system application or device, the video surveillance storage installed base is growing the fastest at a CAGR of 31% over the same period.

As video and video-like data sets continue to grow along with the requirements of storing those data sets for extended periods of time due to regulatory requirements, it is safe to assume that a significant amount of data will be retained on cost-effective cold storage offerings, such as tape or object storage. The need for mining valuable insights from large data

## AT A GLANCE

### KEY STATS

IDC's research highlights just how much data growth is taking place. By 2025:

» 80% of data will be video or video-like.

» 79.4ZB of data will be created by almost 42 billion connected IoT devices.

### WHAT'S IMPORTANT

While data is exceedingly valuable, its growth is a serious concern for IT organizations charged with storing it. As the variety of workloads consuming the data and the endpoints where data is generated continues to increase, there is a need for end-to-end integrated solutions.

sets to improve business processes, products, or services will also grow, and therefore, these data sets should be stored on an active archive platform, such as object-based storage. Consequently, the infrastructure deployed in support of such data sets needs the flexibility of scaling performance and capacity. Thus, infrastructure must be both cost efficient and deliver on insights via analytics and artificial intelligence. As the variety of workloads consuming the data and the endpoints where data is generated continues to increase, there is a need for end-to-end integrated solutions.

## Data Growth Market Trends

Unstructured data, such as video and video-like data, is created in many ways; once retained, this data undergoes a typical life cycle. Examples of where and how video and video-like data is generated include media and entertainment (M&E)–related content for television and devices; corporate training and informational material; digitized educational materials across schools and universities; healthcare and life sciences–related high-resolution images; seismic and geospatial imagery for the oil and gas industry; video surveillance at disparate locations across many verticals; and sensor data from edge devices to predict traffic patterns or in support of autonomous cars.

The following data points from IDC research indicate specific market trends that are driving unstructured data growth across market segments:

» **Overall unstructured data growth.** IDC's *Worldwide File- and Object-Based Storage Forecast, 2018–2022* indicates that capacity shipped in support of file and object storage will reach 881EB by 2022, growing at a 40.3% CAGR from 2017 to 2022.

» **Capacity for media workloads.** IDC's *Worldwide Composite Media Workloads (Compute and Storage) Infrastructure Forecast, 2018–2022* predicts that the infrastructure spend for compute and storage will reach $24.3 billion by 2022, growing at a 14.6% CAGR from 2017 to 2022. *Note: Composite workloads in this case refers to a mechanism to size complex workloads that consist of several subtypes. For example, M&E has many subtypes (and infrastructure needs) for the various stages of content acquisition, production, and delivery.*

» **IoT devices.** IDC estimates that the overall number of "connected" IoT devices worldwide will increase to over 35.2 billion by 2023 and almost 42 billion by 2025. While the amount of data produced by any individual "thing" at this point in time will most often be measured in bytes, cumulatively and over time, connected things will produce vast quantities of data that will need to be collected (via a network) and analyzed (on some type of compute). Some of this data can be transient in nature, but a significant amount of data will still need to be collected, stored, managed, and analyzed.

» **Edge and core.** According to IDC's *Worldwide Composite Media Workloads (Compute and Storage) Infrastructure Forecast, 2018-2022,* 133.5EB of capacity will be shipped in support of edge IT environments. The same forecast also states that core infrastructure spend was expected to decline by 6.3% in 2019, and continue to decline 0.6 in 2020, but edge infrastructure spend was forecast to grow by 5.9% in 2019 and by 14.5% in 2020. To better serve these customers, suppliers should take a holistic solutions approach rather than try to present standalone piece parts for edge deployments.

## Key Infrastructure Considerations to Optimize Data Storage

Most video or video-like data, regardless of where it is generated, eventually undergoes ingestion, processing, finishing, distribution and, finally, archiving. The workflow of any video or video-like content is as follows and bears specific infrastructure requirements:

» **Content creation and ingestion.** Video and video-like content is typically created at the edge. For example, raw footage for films or sitcoms is generated at the site of the shoot, and similarly, video surveillance is generated at disparate locations. These raw large files need to be ingested at high speeds into the core IT infrastructure for further processes within the workflow such as cataloguing, editing, transcoding, or analyzing.

» **Content cataloguing.** Large video or video-like files generated at the edge have limited system-related metadata such as time and date of creation and size of file. More information regarding the content of the file needs to be added to increase the efficiency of finding and retrieving relevant large files based on specific search filters in an efficient manner. Much like video content, metadata should also be treated as critical information — created when content is generated/ingested and backed up appropriately. Typically, the storage infrastructure houses the metadata database in a dedicated server to improve performance.

» **Content finishing and distribution.** After content is ingested and catalogued, the post-production (editing or rendering in M&E) or analysis (e.g., as with healthcare or video surveillance) process begins with raw footage. In M&E, this process ends with a finished product ready for distribution. Every step of the process in M&E — sharing media within and outside the post-production house, ensuring security, maintaining performance as well as delivering ease of management — has specific requirements from the underlying storage infrastructure. To accelerate workflows within this category, organizations need to carefully plan and deploy the file-based storage infrastructure (across on-premises, private, or public cloud), keeping in mind performance and capacity requirements. Additionally, partnerships with industry-specific independent software vendors (ISVs) are important to consider.

» **Content preservation and protection.** Regulatory and business requirements mean that content will be retained for extended periods of time. Tape- or object-based storage provides a cost-effective storage tier for long-term preservation. When cold data is stored on cheap and dense tiers, it is equally important that the storage infrastructure supporting it have capabilities to protect against data corruption with regular checks and self-healing mechanisms.

Organizations have specific requirements of their infrastructure supporting high-resolution video-related workloads. IDC believes that the key storage infrastructure requirements needed to handle these workloads are as follows:

» End-to-end solutions that span edge to core to cloud are few in the marketplace and are actively considered for several reasons, such as ease of use, cost of deployment, and better integration.

» Ease of use is one of the top 3 requirements of storage infrastructure across organizations of any size, according to IDC research. Reducing management overhead and thus redirecting existing IT personnel to higher-value tasks is an important objective for any CIO. A solution that spans all deployment locations (edge, core, and cloud) with simplicity at scale is an easy choice for end users.

» Partnerships with strategic media workflow providers, digital asset and content management solution providers, and data storage providers are important to ensure a seamless experience for organizations dealing with production and finishing of video and video-like content.

» High-performance infrastructure is key for editing, rendering, and transcoding media workflows. Infrastructure supporting NVMe or all-flash arrays is a must-have for modern M&E and other editorial workflow environments.

» Cost effectiveness can be achieved with intelligent data management capabilities. For example, solutions that offer automated data tiering to low-cost object storage free up space on expensive high-performance primary storage tiers, resulting in resource optimization and cost management.

» Content management across heterogeneous systems enabling visibility and control is a very important piece of managing unstructured data at scale for ease of retrieving data anytime from anywhere. Solutions offering such capability that is integrated with various storage solutions and services are a key requirement in the modern datacenter.

## *Considering Quantum*

Founded in 1980, Quantum Corporation has established a strong  presence in the M&E market with its StorNext file system. The company also offers tape solutions to a number of verticals. Over the years, Quantum has built a portfolio of offerings that span file system (StorNext), tape storage, and OEM-ed disk-based backup and archive storage platforms.

Recently, Quantum has focused on organic innovation and redefining a unique road map for itself that will leverage the company's 20+ years of expertise working in the M&E (especially Hollywood) industry and major government agencies. With a brand-new leadership team, Quantum has primarily focused its strategy on building products using its own intellectual property, which has reduced operational costs by $70 million. In 2020, Quantum announced plans to acquire the ActiveScale object storage business from Western Digital, adding object storage software and erasure coding technology to the company's portfolio to help speed their expansion into other vertical markets.  Today, the company aims to extend its offerings to many verticals in addition to M&E. According to Quantum, the company is poised for growth and boasts a customer base of more than 30,000 active customers.

### *Product Portfolio: Quantum's Innovation and Heritage*

Quantum's current portfolio includes several products announced over the past 18 months. The company envisions a product portfolio that allows its customers to deploy the fastest image processing solution as well as the cheapest cold storage offering. Quantum has been cranking on the innovation engine; its portfolio includes the following products that target image processing for various use cases and deployment locations:

» **StorNext Scale-Out File Storage** is a scalable, high-performance networked storage offering that is specifically designed for large unstructured data sets. At the core of the StorNext product line is the StorNext file system and data management software. Customers build a StorNext SAN, or cluster, based on a combination of Quantum appliances, including F-Series NVMe storage servers and QXS hybrid storage arrays. The StorNext policy engine enables users to move data between pools of storage, including file, object, tape, and cloud.

» **VS-Series** is an appliance series purpose built for video recording and management. The VS-Series is value-engineered to capture feeds from hundreds of cameras in a very efficient server design and is able to host VMS applications on the same server hardware used for recording servers.

» **Tape storage** is available for massive-scale archives, on-premises and in the cloud, with Quantum's Scalar tape offering. The company believes that given the rate of data growth and need for cost efficiency, its tape offering can provide a cold storage tier in a hyperscale public cloud infrastructure environment.

» **R-Series** is designed to capture, upload, manage, and store data generated in rugged environments at the edge across various use cases, including in military vehicles and autonomous research vehicles.

» **Backup appliances** include Quantum's DXi Series for scalable, high-performance disk-based backup requirements. The DXi Series addresses the needs of small to very large environments and integrates into the widest range of backup software in the industry.

» **Distributed cloud services** are available through Quantum's pay-per-use storage services with the control of on-premises storage for the company's tape, StorNext, and DXi products as well as custom offerings. With this service, Quantum installs and integrates its products in its customers' datacenters and supports capacity monitoring, metering, and upgrades as well as technology refreshes and data migration. Quantum also leverages cloud-based analytics (CBA) to collect telemetry data and logs from the as-a-service deployment to Quantum Cloud for monitoring and analysis.

Quantum's new and existing offerings combine to make a portfolio that caters to the modern, demanding infrastructure requirements of organizations undergoing digital transformation initiatives. The company's current solutions target verticals such as M&E, healthcare, and local and federal government as well as the hyperscale public cloud market.

Quantum is predominantly focused on making infrastructure deployment and management a seamless and user-friendly experience. Today, Quantum has established elements within its storage portfolio to support a software-defined storage version of the company's StorNext file system and is poised to potentially bring a hyperconverged operating mode to market.

> Quantum's portfolio of offerings caters to the modern, demanding infrastructure requirements of organizations undergoing digital transformation initiatives.

### Challenges

One of the immediate challenges for Quantum is to leverage its success in media and entertainment in order to grow in adjacent markets. The capabilities of StorNext within the M&E space can be replicated to many adjacent markets across verticals such as healthcare and finance and use cases such as video surveillance. The company's outbound marketing activities will help in implementing this shift. Despite the announcements of several new products that target data collection and processing at the edge and also video surveillance, the biggest task at hand for the company is to shed its old cloak of predominantly being a tape storage vendor and emerge as a new, foundationally solid company that serves many verticals with its end-to-end solutions. This change will come about through constant customer education, outbound marketing activities, increased testimonials of success stories across other verticals, and

continued organic development. Quantum should continue to stress to potential customers both its innovation and its ability to address the storage needs of many verticals beyond the M&E industry.

In 2017, the SEC launched an investigation into the company's historical accounting practices and internal controls under a former management team. These events triggered an internal investigation, a restructuring, and a delisting from the NYSE. Quantum rectified its financial missteps, implemented new business priorities, standards, and governance practices focused on innovation and profitable growth, and is now listed on the NASDAQ exchange as QMCO.

## Conclusion

Quantum has an established presence in the marketplace and in recent months has made advances in product innovation, resulting in an expanded portfolio. Quantum's clear vision and road map are the basis for the company's accelerated innovation across the edge, core, and cloud.

IDC believes that any organization, regardless of industry (healthcare, life sciences, autonomous vehicles, government, M&E), looking for storage solutions that are geared toward video or video-like data should evaluate Quantum's offerings.

# About the Analyst

*Amita Potnis,* *Research Director, Infrastructure Systems, Platforms, and Technologies Group*

Amita Potnis is a Research Director within IDC's Enterprise Infrastructure Practice, covering research on infrastructure for content and cloud-native apps, artificial intelligence (AI), and security. Ms. Potnis specifically focuses on early-stage and emerging vendors providing advisory services in areas of product management, marketing, and market positioning. She takes a keen interest in assessing the impact of new technologies and products on established markets. Ms. Potnis also engages with clients on various custom research and consulting services.

## MESSAGE FROM THE SPONSOR

**About Quantum**

Quantum technology and services help customers capture, create and share digital content — and preserve and protect it for decades at the lowest cost. Quantum's platforms provide the fastest performance for high-resolution video, images, and industrial IoT, with solutions built for every stage of the data lifecycle, from high-performance ingest to real-time collaboration and analysis and low-cost archiving. Every day the world's leading entertainment companies, sports franchises, research scientists, government agencies, enterprises, and cloud providers are making the world happier, safer, and smarter on Quantum. See how at www.quantum.com.

**IDC** Custom Solutions

**The content in this paper was adapted from existing IDC research published on** www.idc.com**.**

**IDC** | ANALYZE THE FUTURE

www.idc.com