Western Digital®

Qumulo and IHME Deploy Ultrastar® NVMe™ SSDs and Hard Drives to Combat Covid-19

Institute for Health Metrics and Evaluation

Qumulo

Challenge

- COVID-19 pandemic resulted in an immediate demand for large-scale data modeling and visualization
- Dramatic increase in data requires a scalable file storage system
- Need to store large data sets, while simultaneously requiring performance for analysis and visualization

Solution

- A scalable solution for healthcare forecasts, models and visualization
- Western Digital Ultrastar DC HC550 18TB[1] HDDs with HelioSeal®
- Western Digital Ultrastar DC SN640 3.2TB, 2 DW/D NVMe SSDs
- Qumulo C-432T Cached Performance platform and Qumulo file data platform
- Microsoft Azure for publishing results as data visualizations to the public

Key Results

- IHME gained the ability to rapidly store and analyze large databases from multiple customers
- IHME increased the amount of terabytes read, analyzed and visualized for agencies and organizations to create data-driven, science-based, actionable plans to combat COVID-19
- Qumulo provided a scalable, on-prem file storage system to enable analysis of massive data sets to understand trends that impact public health.

Improving Healthcare with Data

The Institute for Health Metrics and Evaluation (IHME) is an independent global health research organization based at the University of Washington (UW) School of Medicine. IHME's mission is to improve the health of the world's populations by providing the best public health information available. IHME makes their research freely available so that policymakers and healthcare institutions have the evidence they need to make informed decisions on allocating healthcare resources—critical decisions that save lives. IHME initially deployed Qumulo's file data platform in 2017, enabling a dynamically scalable architecture and real-time visibility of data that IHME needed to manage a wide variety of incoming data from hospitals, universities and governments around the world.

Ramped Data Analysis to Combat a Pandemic in Record Time

In February 2020, a month before the World Health Organization (WHO) announced the COVID-19 pandemic, the UW School of Medicine requested IHME's immediate help with pandemic modeling. Other hospital systems and multiple state governments quickly followed with requests to IHME for population models for their communities. They needed to know when COVID-19 would overwhelm their ability to care for patients and what they could do to avoid that. Within days, similar requests poured in from Puerto Rico, Canada and European Economic Area (EEA) countries.

To respond to these urgent requests, IHME needed to produce large-scale data modeling for forecasts and include daily and cumulative COVID-19 death reports, infection and testing numbers, and social distancing information.

Almost overnight, IHME needed massive additional data resources. IHME turned to Qumulo to help them respond to the new influx of data and create a rapid release cadence of new visualizations without putting existing projects on hold. Qumulo provided the Cached Performance platform powered by the Qumulo file data platform for on-prem data processing with massive storage capacities enabled by Western Digital Ultrastar NVMe SSDs and HDDs.

"Visualizations are core to IHME communications with policymakers for the scientific papers that are rigorously peer-reviewed by journals. Qumulo is critical to enabling us to distill hundreds of millions of data points into a single visualization, which allows policymakers to easily view the results and communicate them to their teams."

Serkan Yalcin
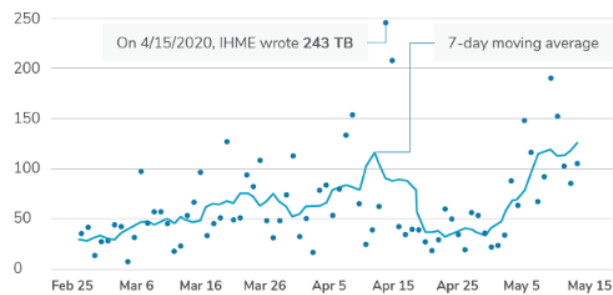Director of IT Infrastructure, IHME

Along with Qumulo and Western Digital, IHME used Microsoft Azure for hosting data visualizations. These platforms enabled IHME to deliver COVID-19 visualizations within days of the initial requests and rapidly scale to meet needs from governments and healthcare officials tasked with keeping communities safe around the globe. Observations and projections included data based on the number of deaths, infections and testing, mask use, social distancing, and hospital resource use including ICU beds and ventilators.

Prior to COVID-19, the Qumulo system was already capturing and processing large amounts of data for existing research deliverables. COVID-19 research dramatically increased the amount of data analyzed, reading up to 2PB/day. The agility of Qumulo's architecture enabled IHME to accommodate the sudden growth in both data ingest and processing demands without the need to re-architect. Extremely large amounts of unstructured data were stored on Qumulo's file storage system using Ultrastar NVMe SSD flash for high throughput and low latency.

IHME: Terabytes read per day from Qumulo

On 4/15/2020, IHME read **2,073 TB**

7-day moving average
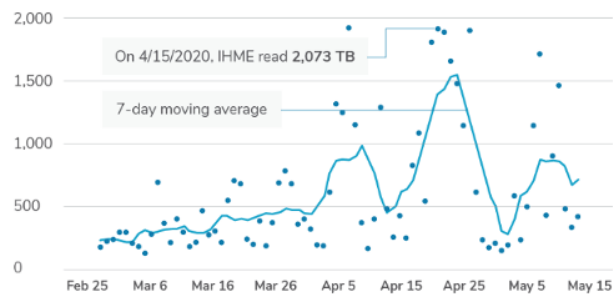
IHME: Terabytes written per day to Qumulo

On 4/15/2020, IHME wrote **243 TB**

7-day moving average

Going forward, IHME has announced a multi-stage hybrid modeling approach that includes estimating COVID-19 deaths, infections and viral transmission in multiple stages. A new component will capture not only fatality rates but the rates at which individuals move from being susceptible to exposed to infected to recovered. This information will drive quantifiable evidence for social distancing policies by location, a major advance in protecting the population against COVID-19 while jumpstarting local and regional economies.

Ultrastar NVMe SSDs and HDDs for High-Performance and Large Capacity Storage

With a broad portfolio of data center products, Western Digital Ultrastar NVMe SSDs and HDDs were the ideal choice for Qumulo's Cached Performance file storage system.

NVMe SSDs are optimal for the high-performance data analytics that the Qumulo C-432T file storage system is designed for. Ultrastar DC SN640 NVMe SSDs deliver 6X improvement in sequential read performance compared to SATA SSDs. This translates into faster data analysis on large data sets. With endurance of 2 DW/D, a 6×3.2TB SSD configuration can write 70PB of data over the 5-year warranty of the drives.

Large-scale analytics depend on massive data sets, which benefit from HDDs as primary storage. The Ultrastar DC HC550 18TB HDD is the ideal solution to store the maximum amount of data within the 2U form factor of the Qumulo C-432T Cached Performance platform. Ultrastar HelioSeal-based HDDs offer large capacities, low power and exceptional reliability.

"Radical simplicity is critical for customers to be successful with unstructured data. Our software-driven file data platform offers enterprise-level capabilities that radically simplify analytics of massive data sets," said Ben Gitenstein, VP of Product, Qumulo. "With fast reliable storage from Western Digital supporting the file data platform, our customers can simplify the complexity of their infrastructure, accelerate innovation and unleash the power of their data, wherever it resides."

---

[1] One gigabyte (GB) is equal to 1,000MB (one billion bytes) and on terabyte is equal to one trillion bytes. Actual user capacity may be less due to operating environment.

Western Digital.