FOUR STORAGE CONSIDERATIONS FOR SCIENTIFIC RESEARCH

Digital data is critical to scientific research today more than ever. New technologies—such as next-generation genomic sequencers, unmanned vehicles and spacecraft, and advanced sensors and image capture devices—are enabling tremendous breakthroughs in the understanding of our world, our universe, and our bodies. These new technologies also create tremendous amounts of data. To maximize the value of this data, and of the associated research teams, research institutions must have a properly designed storage infrastructure and data management plan. The following four considerations will help institutions evaluate the storage requirements for supporting these new research workflows and enable more efficient processes—ultimately, accelerating time to discovery while better managing storage and data management costs.

1. ACCESS

Obviously researchers need access to their data. And research today typically involves collaboration within a team. Therefore, it's necessary that all team members have shared access to all their data, all the time—regardless of IT details like operating system (Windows, Mac, Linux, UNIX), protocol (LAN, SAN, IP), or storage location (primary, archive, cloud). And not just shared access, but simultaneous access to the same file to enable the efficiencies of a truly collaborative, parallel workflow. Another access consideration depends on the location of the research team. Is the whole team working in the same building, on the same campus, in the same country? The storage infrastructure must accommodate the entire team—whether local or distributed—and perhaps provide flexibility as the environment may change over time. To summarize, scientists just want their data available whenever the analysis needs it—no searching, no IT tickets, no waiting.

2. PERFORMANCE

Performance needs vary depending on where the data is in the workflow. For example, as data is created or captured it must be quickly stored or ingested. The opportunity to recreate the data may be impractical or perhaps even impossible. And running analytics on the data requires streaming storage performance to support the computational requirements. Optimizing performance for a particular analysis workload depends on both the data characteristics and how the files are accessed. For example, is there a lot of random I/O with small files or sequential streaming of large files? What about read vs. write? It's important to understand these parameters, and deploy a solution that delivers the best storage performance for the workload. If the parameters vary across workloads, then the storage solution must provide flexibility to deliver optimum performance accordingly. Another consideration impacting performance is data protection. As research data sets become large, the feasibility—due to both cost and time required—breaks down for traditional IT data protection processes. What's needed is an automated, non-disruptive data protection process to ensure the valuable research content is secure and protected without interfering with the research process.

3. COST

Let's face it, there's competition for funds within the research budget. Money saved on storage means more money available for fun things—like faster compute, or genomic sequencers, HD cameras, or sensors—which then drives the need for more storage. Or maybe trimming the storage budget means securing a new research grant. At any rate, prudent use of storage resources without penalizing the research workflow just makes sense. A multi-tiered data management solution is the best way to manage costs because it allocates the right storage performance to the right process in the workflow—high-performance storage for active data and low-cost, high-capacity storage for inactive data—while maintaining accessibility to the data from the same access point or directory location, regardless of which storage tier it resides on. Solutions that can combine different types of storage—such as flash, disk, object, tape, and cloud—enable institutions to balance their needs for performance, scale, access, and budget. The data management solution should also include data protection policies to automate the copying of data to appropriate storage tiers for security and protection of the data, including disaster recovery (DR)—without impacting the research workflow.

4. FUTURE

Finally, research institutions must invest in storage solutions that are flexible enough for the needs of an uncertain future. One thing we know the future will bring is more data, and as the volume of data is getting so large, the traditional "forklift" upgrade of yesterday isn't feasible anymore; it's important to scale low-cost capacity without disruption. Additionally, the ability to scale performance independent of capacity helps as new algorithms or new high-performance computing (HPC) systems are brought on line. Unanticipated changes could drive the need for deploying cloud storage, or changing a cloud deployment from hybrid to private due to cost, security, or some other concern. The bottom line is, the needs of the future aren't etched in stone, and today's data management solutions must be flexible enough to support unanticipated future needs, including perhaps even new storage technologies.

READY TO GET STARTED?

Scientists today need intelligent data management throughout the entire research workflow, from ingest to HPC processing to archive. With multi-tier storage from Quantum, teams can better harness their data and transform the world.

To learn more, visit www.quantum.com/scientific-research.

ABOUT QUANTUM

Quantum is a leading expert in scale-out tiered storage, archive, and data protection, providing solutions for capturing, sharing, and preserving digital assets over the entire data lifecycle. From small businesses to major enterprises, more than 100,000 customers have trusted Quantum to address their most demanding data workflow challenges. Quantum's end-to-end, tiered storage foundation enables customers to maximize the value of their data by making it accessible whenever and wherever needed, retaining it indefinitely and reducing total cost and complexity. See how at **www.quantum.com/customerstories**.

Quantum®

www.quantum.com
800-677-6268

ST02047A-v01 Apr 2017