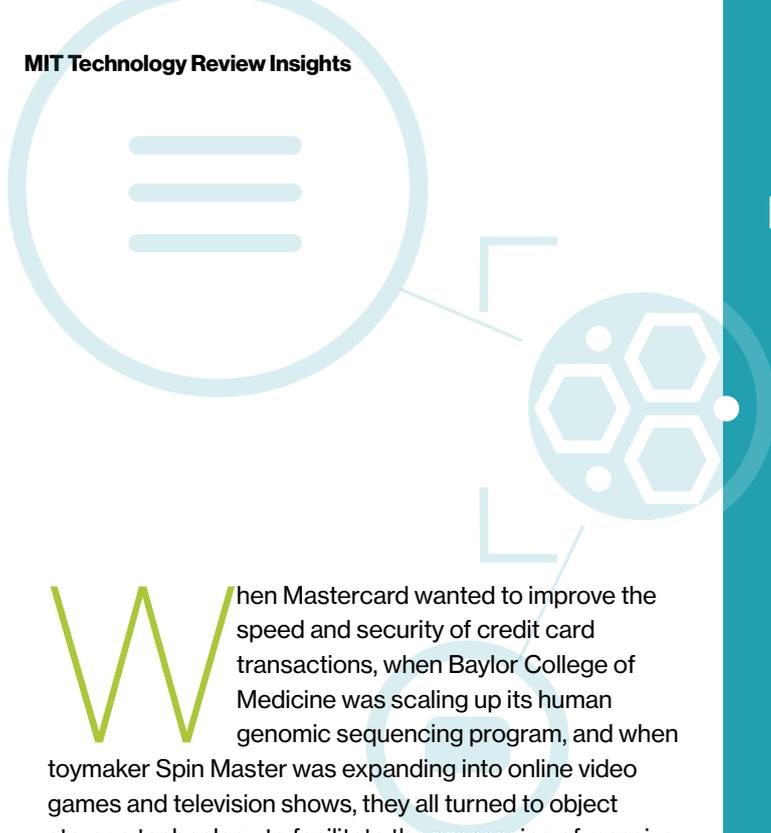


Once used primarily for archiving and backup, the technology is gaining traction for its mammoth capacity, processing power, and cybersecurity capabilities.

Object storage for digital-age challenges





When Mastercard wanted to improve the speed and security of credit card transactions, when Baylor College of Medicine was scaling up its human genomic sequencing program, and when toymaker Spin Master was expanding into online video games and television shows, they all turned to object storage technology to facilitate the processing of massive amounts of data.

Object storage, with its virtually infinite capacity and low cost, has a long history of being deployed for backup, archiving, disaster recovery, and regulatory compliance. But the demands of today's data-centric organizations have brought the technology from the dusty storage closet to the center stage of digital transformation.

For any tech decision-maker thinking about an overall data strategy, having a large central repository, also known as a data lake, is the preferred approach – it helps break down silos and aggregate data from multiple sources for the type of data analysis that delivers value to the business. Object storage is the most effective underlying technology for applying data analytics, machine learning, and artificial intelligence (AI) to those vast data stores, says Scott Sinclair, storage analyst at market researcher Enterprise Strategy Group.

“The biggest advantage of object storage is to add more value to primary data. It doesn't just store files; it adds context,” says Paul Schindeler, a former IDC analyst and currently CEO of the Dutch consultancy Data Matters. An object store includes metadata, or labels, which enables companies to easily search vast volumes of data, determine the origin of the data, whether it has been altered and, more important, to set policies and keep auditable records on who can see the file, who can open it, and who can download data.

Key takeaways

- 1 Organizations have long used object storage for backup, archiving, and regulatory compliance. Now, in the digital age, tech leaders are looking to the technology to undergird data-centric applications such as analytics, machine learning, and artificial intelligence.
- 2 Object storage enables forward-thinking organizations to build critical applications on vast stores of aggregated data, analyze that data to fuel innovation, and protect the organization from ever-worsening threats of ransomware.
- 3 The technology won't deliver on its promises without careful thought and commitment. Organizations need to plan to store data, add metadata to it – to figure out what data they have and how valuable it is – and spark innovation. Finally, they need the right skills and technology to do the job.

The demands of today's data-centric organizations have brought object storage from the dusty storage closet to the center stage of digital transformation.

Most organizations today use a mix of storage types: file storage, block storage, and object storage (see sidebar, “Three types of data storage”). But the use of object storage is surging for a number of reasons: speed, scalability, searchability, security, data integrity, reliability, and protection against ransomware. And it's the wave of the future when it comes to big data analytics.

Object storage, then and now

Object storage was developed in the 1990s to handle data stores that were simply too large to be backed up with file and block storage, says Sinclair. When introduced, the almost infinite scalability, low cost, and immutability of object storage made it ideal for backup and recovery and long-term archiving and compliance with regulations such as the Health Insurance Portability and Accountability Act, in health care, and Sarbanes-Oxley, in banking.

The next watershed event in the evolution of object storage was the ascendance of cloud storage. Cloud services vendor Amazon Web Services chose object storage architecture as the foundation for its popular

Simple Storage Service (S3), and object storage has become the standard platform for all cloud storage, whether from Google, Microsoft, or others. In addition, S3 protocols have become the industry standard for modern data-centric applications, whether they run in the cloud or in a corporate data center.

More recently, organizations have come to the realization that they need to do more than just park and protect their data; they need to extract value from vast troves of historical data, as well as from new data sources and data types, such as internet-of-things sensor data, video, and images. That's where object storage really shines. It has become the platform organizations are building their data analytics capabilities on to modernize their computing environments, create innovation, and drive digital transformation.

An object storage play drives business

Spin Master is a Canadian company that started out selling children's toys – everything from old-school Etch A Sketch and Erector Sets to the Air Hogs line of remote-controlled toy airplanes and Hatchimals robotics toys. Through both innovation and acquisition, Spin Master has transformed into a diversified creator of digital applications, video games, movies, and television shows. It owns the Angry Birds franchise and produces a number of animated television shows.

Pravine Balkaran, vice president for global IT, information, cybersecurity, and AI at Spin Master, says his pre-existing, on-premises block storage system could not keep pace with the company's rapidly expanding data storage

“Since moving over to the object storage platform, we worry less about file corruption, and we worry more about keeping up with the business.”

Pravine Balkaran, Vice President for Global IT, Information, Cybersecurity, and AI, Spin Master

Three types of data storage

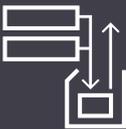
Here are the similarities and differences of file storage, block storage, and object storage.



File storage is the simplest form of data storage. Think of file storage as a digital equivalent of a filing cabinet. Files are organized into directories and sub-directories. The problem with file storage is that it doesn't scale.



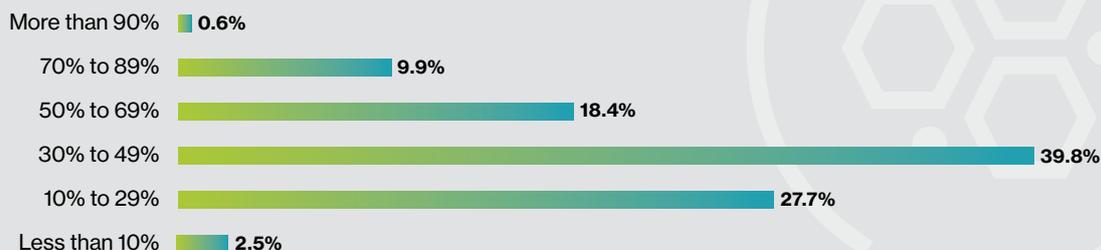
Block storage technology splits up large volumes of data into blocks of equal size. Block storage is used for a wide range of database applications, but, as with file storage, assuring data integrity becomes a challenge when volumes grow.



Object storage manages data not as files or blocks of files, but as objects. Objects can include all types of structured and unstructured data. Each object contains the data, metadata (contextual information about the data) and a unique identifier, which enables distributed data to be easily retrieved. The rich metadata provides valuable information about the use and function of the data. Object storage has infinite scalability, quick retrieval, and is the technology of choice for data analytics.

A data tsunami

Most organizations will see their data growth swell in the next two years.



Source: IDC's "Unstructured Data Storage Study 2019" of 515 IT leaders, June 2019

requirements. "We needed data integrity to be spot on, configuration to be spot on. We had to do all of this monitoring for file corruption," Balkaran says. "Since moving over to the object storage platform, we worry less about file corruption, and we worry more about keeping up with the business."

He points out that Spin Master generates massive amounts of rich data. "Object storage gave us a bottomless pit where we just keep dumping data, whether we need it now or whether it needs to be retrieved in six months." Given the size of these data sets, Spin Master decided that it would be easier and faster to keep high-value data that needs to be accessed quickly on-premises rather than in the cloud.

"How quickly can I access a 5-gigabyte file of video that I did a commercial for, that I need to chop up and put into megabytes on social media – and I need to get it done in the next two hours. If we had to pull from the cloud, it could take more than two hours to do that," Balkaran explains.

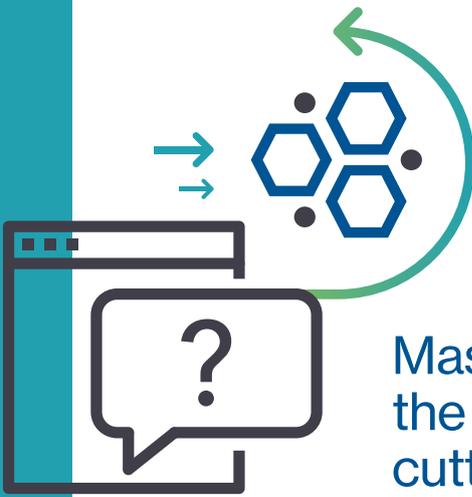
Spin Master also uses cloud storage for traditional object storage jobs, backups, and archiving, and those backups provide protection against ransomware, an increasing concern for companies and government agencies. One in four incidents remediated by the IBM Security X-Force Incident Response team is a [ransomware attack](#), and attacks are increasing in number and in the amount of money perpetrators are seeking. Not only that – attackers are shifting their focus to manufacturers, which can least afford any downtime.

Balkaran says object storage allows Spin Master to verify the integrity of its data, add policies on confidentiality and availability, and produce an audit trail that shows who has accessed what. The object store is backed up, so if the primary data store is hit with a ransomware attack, the backup is still available. But the most significant benefit of object storage is it enables Spin Master to move at the speed of business.

Mastercard's massive processing needs

For 50 years, Mastercard has been driving the digital economy by building a global payments system that is safe, secure, and fast. As the volumes of transactions go through the roof and customers expect lightning-fast digital experiences, contactless transactions increase during the 2020 coronavirus pandemic, and attackers intensify their efforts to disrupt the smooth functioning of digital commerce, Mastercard constantly updates and upgrades its architecture. The company is also committed to delivering innovative, data intelligence-based services on top of its core transactions-processing function.

Mastercard chief technology officer Steve Bagby explains that customers paying with a credit card at a bricks-and-mortar store aren't that concerned about waiting an extra second or two for the transaction to complete, but when you're driving your car through an automated tolling gate or interacting inside a video or virtual-reality game, latency becomes critical. The solution is to do processing closer to the customer, which entails a seismic shift from large, centralized data centers to a fully distributed system that includes cloud storage.



Mastercard is using object storage as the key enabling technology behind cutting-edge information services.

“And that’s where different storage architectures come in play,” says Bagby. “We use a lot of containerized workloads that basically are around for maybe a few seconds or a few hours at a time, as opposed to the old school where you had an application that would sit there and run for a time. So for us, to the existing suite of files and blocks, we’ve added object storage to make those ephemeral applications work a lot better.”

With object storage, Bagby adds, “there’s no cruising around trying to find either a shared block storage or shared file system. With the object store you can essentially pull a copy of the data through an HTTP request, and you’re off to the races.”

Mastercard has deployed object storage to make sure that “card not present” e-commerce transactions are conducted securely and efficiently. In this scenario, a customer attempts to buy something online, which means the card isn’t handed over to a sales associate or cashier, as would occur in a face-to-face transaction.

On the back end, a number of complex, interrelated actions occur, all enabled by an object store. For example, all of the card information that the customer enters needs to be verified, and sophisticated fraud detection mechanisms kick in to make sure the card isn’t being used in a suspicious manner.

Assuring the security of credit card transactions is a cornerstone of Mastercard’s business. Bagby says, “We are investing pretty heavily on the security front to help combat fraud. And not just the fraudulent use of credit cards, but things like money laundering and other emerging issues that are part of electronic payments.”

Mastercard is also using object storage as the key enabling technology behind cutting-edge information services the company is providing to retailers on top of the transaction stream. These live data services are predominantly either direct-stream processing or working on data that is being pulled out of an object storage device. “Data insight has been particularly important in our

Multiple uses for object storage

Some ultramodern uses include supporting the development of cloud apps and metadata enrichment – improving metadata quality so content is easily found online.



current situation because consumer behavior has shifted so tremendously” in light of the ongoing coronavirus pandemic, says Bagby.

Gene sequencing and city management

Human genome sequencing is one of the more exciting research activities taking place today – and one of the most data-intensive. The Human Genome Sequencing Center at Baylor College of Medicine has a massive 5-petabyte object store that supports a variety of research projects, including “All of Us,” a federally sponsored research project aimed at creating a genome-based health database of more than 1 million volunteers across the country.

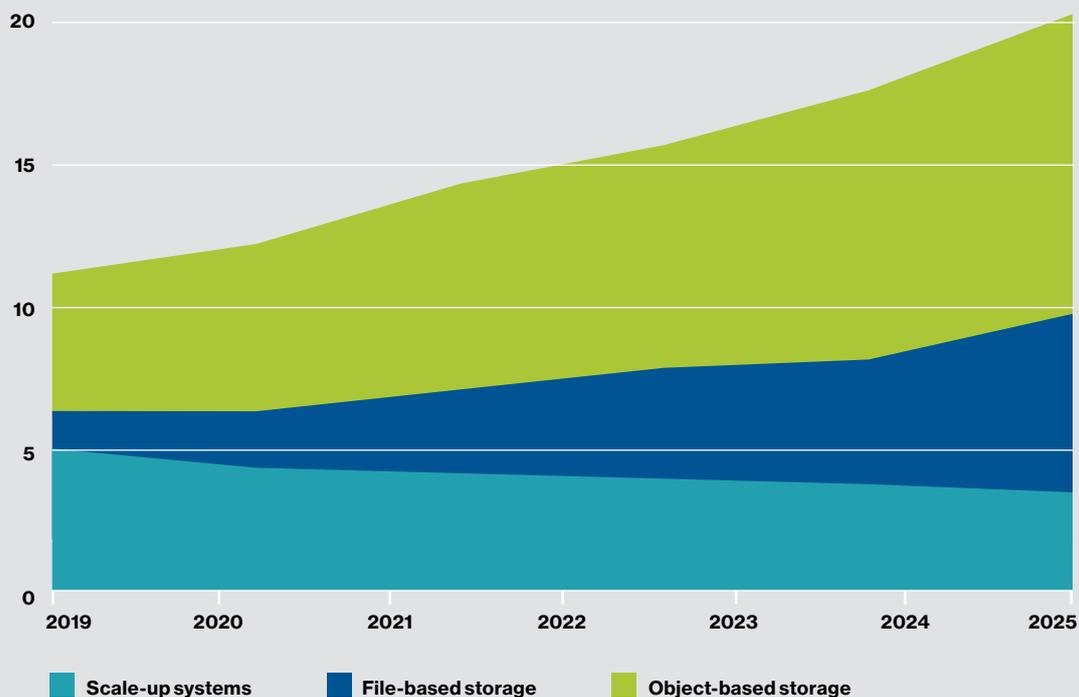
IT director Kim Andrews is taking full advantage of object storage technology to conduct genomic sequencing, a complex series of processes that requires intensive data

analytics. For example, each DNA sample is broken up into 12 parts, which are all processed independently, then re-assembled. “So it presents a unique storage challenge for both the cloud and locally, because you can’t control when you kick off these computes,” says Andrews. Since DNA samples are essentially images, it would be virtually impossible to use file or block storage, which don’t easily handle “unstructured” image data.

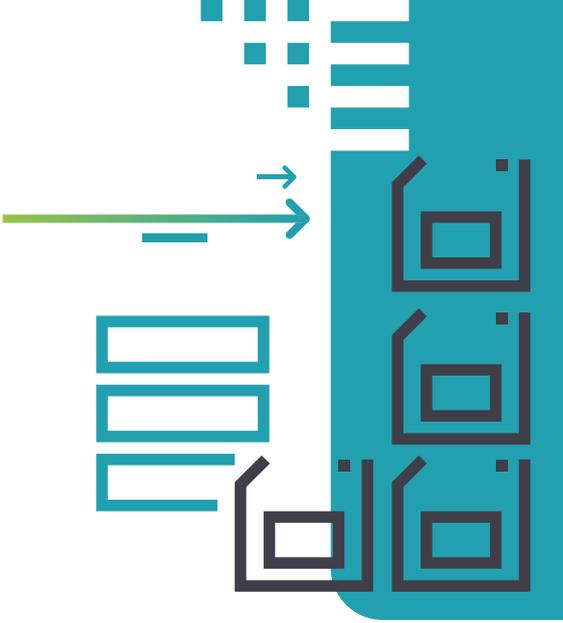
Object storage not only provides the virtually limitless volume and ability to process the data that human genomic sequencing requires – it helps Andrews secure sensitive documents. “We have a security system that we maintain evidence of compliance in, and that evidence needs to be protected and encrypted but easily accessible,” Andrews says. “So I can send a URL to our security auditors and they can open that secure URL and use that as evidence.”

Storage market forecast 2019

Object storage revenue is expected to grow over the next four years – an acceleration of cloud storage adoption in the wake of the coronavirus pandemic – while the costly “scale-up” file servers and appliances market is projected to fall at a compound annual growth rate of 7%.



Source: IDC's "Worldwide File- and Object-Based StorageForecast, 2020-2024," June 2020



Object storage “allows you to keep multiple copies automatically. You don’t have to manage backups. That makes it resilient for ransomware and data loss.”

Jason Carolan, Chief Innovation Officer, Flexential

And object storage is enabling other types of innovative projects that couldn’t have been accomplished with file or block storage. For example, colocation provider Flexential, which offers object storage-as-a-service, is using the technology to process video files for the City of Denver in a project aimed at real-time traffic management.

Jason Carolan, chief innovation officer at Flexential, explains that the city government has three separate systems from three separate vendors that run different parts of the traffic management system. Using object storage as the foundation, the project involves putting together a video canopy to triangulate what’s happening upstream and downstream from traffic lights, and then building an integration platform to better control lights, smooth out traffic congestion, and make roads safer for emergency vehicles.

‘It keeps sucking in the data’

Object storage provides a range of benefits for organizations that want to aggregate vast amounts of data, layer essential applications on top of that data store, conduct advanced data analytics that drives innovation, and protect against ransomware attacks.

“The nice thing with object storage is that it keeps sucking in the data and just gets out of the way. It allows you to keep multiple copies automatically,” Carolan says. “You don’t have to manage backups. That makes it resilient for ransomware and data loss.”

Analyst Sinclair adds that the long-term, strategic goal for organizations should be to eliminate data siloes and deploy a consistent architecture across on-premises and cloud environments – because data will inevitably be moving back and forth. Object storage enables companies to create a single view of its data so that it can apply AI and machine learning for business-critical data analytics.

But pursuing a data strategy that incorporates object storage requires careful planning and deep commitment. Such a strategy should include the capabilities to store, enrich, and monetize data, says Jason Hardy, global chief technology officer for Hitachi Vantara’s content portfolio. Storing means also managing data wherever it resides. Transactional data must be protected, kept private and secure, and governed. And to understand the value of data, organizations need to enrich it, assigning metadata that describes the context under which it was created. Only then can data can be cataloged. Object storage has such capabilities, Hardy says, but to deliver on them, organizations need to shape “ecosystems of talent, technology, and information,” consisting of engineers, consultants, and trusted technology partners.

“Delivering on an innovative data strategy requires a diverse set of skills to support planning, execution, integration, governance,” Hardy says. But skills alone aren’t enough. “Monetizing on the value of data is a momentous undertaking that requires commitment – and a new mindset – at all levels of the organization.”

“Object storage for digital-age challenges” is an executive briefing paper by MIT Technology Review Insights. It is based on research and interviews conducted in July and August 2020. We would like to thank all participants as well as the sponsor, Hitachi Vantara. MIT Technology Review Insights has collected and reported on all findings contained in this paper independently, regardless of participation or sponsorship. Jason Sparapani and Laurel Ruma were the editors of this report, and Nicola Crepaldi was the publisher.

About MIT Technology Review Insights

MIT Technology Review Insights is the custom publishing division of MIT Technology Review, the world’s longest-running technology magazine, backed by the world’s foremost technology institution – producing live events and research on the leading technology and business challenges of the day. Insights conducts qualitative and quantitative research and analysis in the US and abroad and publishes a wide variety of content, including articles, reports, infographics, videos, and podcasts. And through its growing MIT Technology Review Global Panel, Insights has unparalleled access to senior-level executives, innovators, and thought leaders worldwide for surveys and in-depth interviews..

From the sponsor

Hitachi Vantara, a wholly-owned subsidiary of Hitachi, Ltd., guides our customers from what’s now to what’s next by solving their digital challenges. Working alongside each customer, we apply our unmatched industrial and digital capabilities to their data and applications to benefit both business and society. More than 80% of the Fortune 100 trust Hitachi Vantara to help them develop new revenue streams, unlock competitive advantages, lower costs, enhance customer experiences, and deliver social and environmental value. Visit us at hitachivantara.com.

HITACHI
Inspire the Next

Illustrations

Cover illustration by BloomIcon. Page 3 icons: CreativeStall; Page 5 icons: BloomIcon and 32 Pixels; Page 7 Icons. 32 Pixels. All icons from Shutterstock, assembled by Scott Shultz Design.

While every effort has been taken to verify the accuracy of this information, MIT Technology Review Insights cannot accept any responsibility or liability for reliance on any person in this report or any of the information, opinions, or conclusions set out in this report.

© Copyright MIT Technology Review Insights, 2020. All rights reserved.



MIT Technology Review Insights

 www.technologyreview.com

 @techreview @mit_insights

 insights@technologyreview.com