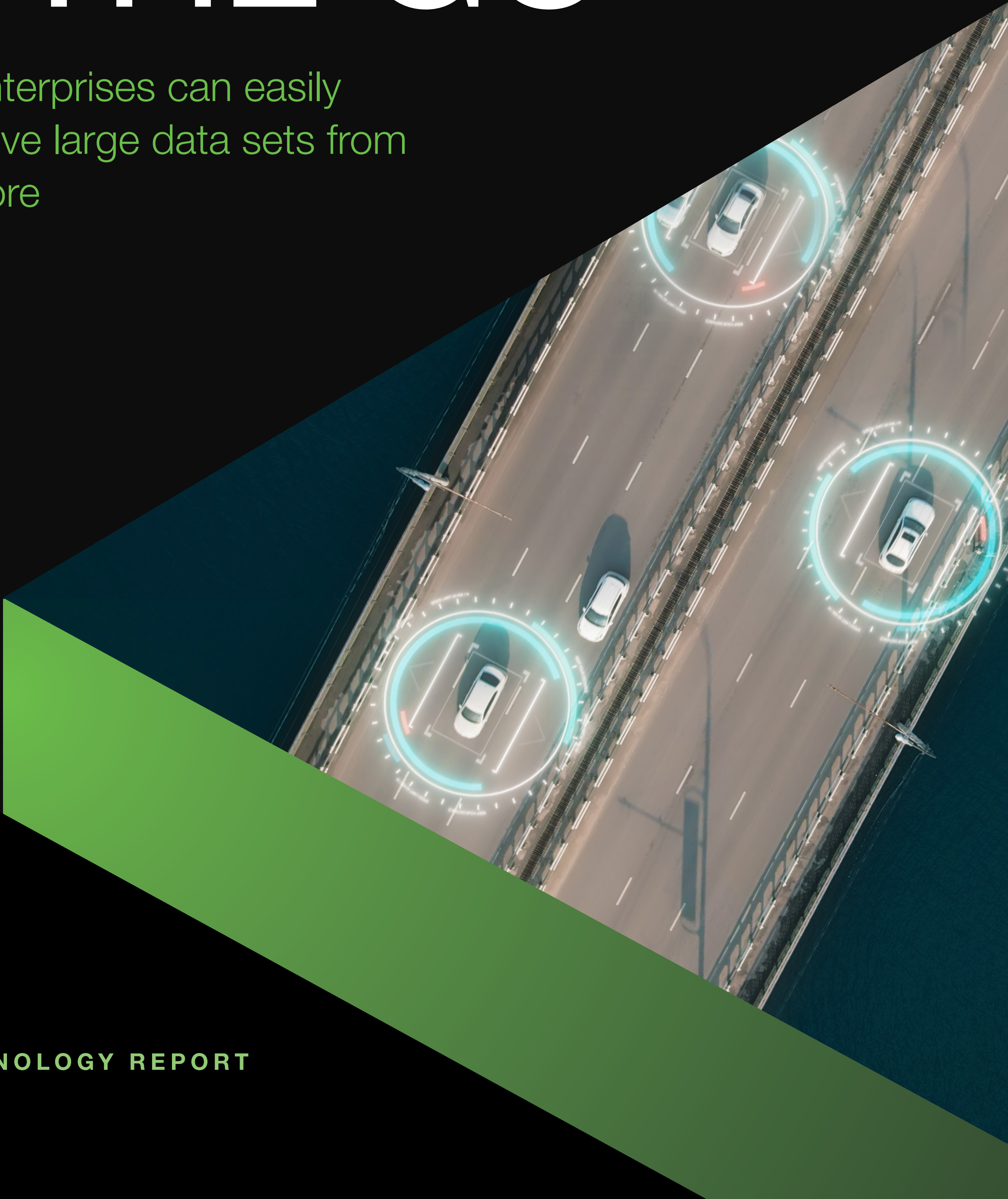
SEAGATE

MASS DATA ON THE GO

How today's enterprises can easily access and move large data sets from endpoints to core



A SEAGATE TECHNOLOGY REPORT

CONTENTS

INTRODUCTION

Mass data is going places, now more than ever.

The current shifts in data gravity usher extraordinary amounts of data in two directions—toward the edge and toward the multicloud core. This means not only that data resides at these two locations, but that the movement of data streams between them needs support.

Whether supporting data at rest or in flight, storage architectures need to be data-centric—prioritizing what's best for a particular data set and its goals. Without an end-to-end data operations strategy that boosts the value of data between endpoints and core, businesses cannot scale.

Companies that can harness the power of their data will be the leading companies of tomorrow. To do that, they need to overcome the cost and complexity challenges that come with data sprawl.

What new opportunities for getting more value from data arise in the distributed world? What kinds of storage strategy can reduce impediments to the movement of large data sets? These are the questions that the winning organizations are asking—and the questions that this report answers.

Jeff Fochtman
Senior Vice President of Business and Marketing
Seagate Technology

HIGHLIGHTS

- The more data can be employed in multiple environments, the more dynamic and fluid it can be—and the higher its business value.
- Mass data is sprawling, and shifts in data gravity are tugging at it. Increasing waves of data are moving toward the multicloud core and the edge.
- In order to be useful, large data sets must be easy to access and move from endpoints, through three types of edge, to the multicloud core.
- Every industry handling massive data sets (100TB or more) faces transport challenges.
- The impediments to mass data's full value include: network and capacity constraints; slow speeds; limited access to fiber-optic networking; cost; security and compliance concerns; and limited storage capacities.
- But these same factors are often also reasons why mass data is propelled into motion—and why enterprises increasingly choose physical data transport. Physical data shuttles can distribute data to the right locations much faster than network uploads.

Section One

FLOWING FROM THE EDGE

Water is vital. With nothing to drink, the body fails. When water sits still, it stagnates. Data is like fresh water: essential. It is nearly infinite and can be captured indefinitely, but it must flow.

Business success demands a constant stream of new information. As with stagnant water, excessive data at rest, unless properly contained, can turn into a resource sinkhole. That's when data silos form and innovation stifles.

Like nature's water cycle, data needs to flow from its local points of creation out to edge and on-premises data centers, the cloud(s), and sometimes back again. When this happens, data turns from something complex and costly into an asset whose value grows.

The philosopher Heraclitus observed that you cannot step in the same river twice. This is especially true when it comes to enterprise data. The streams that flow from factory-floor sensors are both acted on immediately (making real-time adjustments to preempt quality deterioration) and fed into longer-term analysis with artificial intelligence (AI) and machine learning (ML) tools, which results in improved processes. Those processes then, containing new insights, affect how and what data is gathered on factory floors.

Data from local points of generation flows into edge aquifers. From there, it might return to its start or travel to the outer edge data centers or the multicloud core, where it can be easily shared and activated around the world. The more data can be employed in multiple environments, the more dynamic and fluid it can be—and the higher its business value.

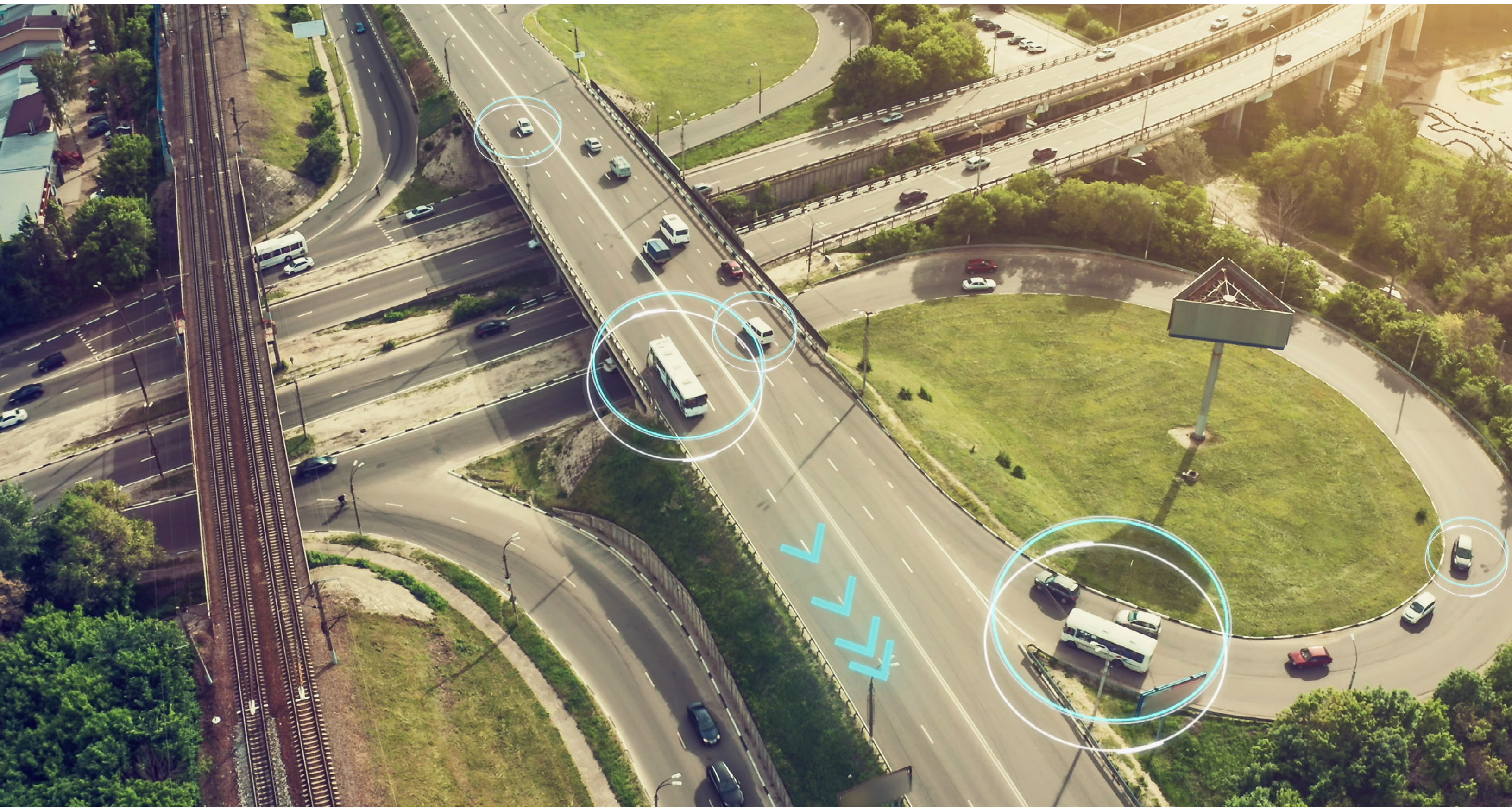
WHY MASS DATA FLOW MATTERS: A USE CASE ON WHEELS

Let's picture the journey of an enormous enterprise data set through the lens of assisted driving innovation—and the masses of data that make it happen.

In the automotive world, advanced driver-assistance systems (ADAS) began decades ago with features such as anti-lock brakes and traction control.

Today, cars with the most evolved ADAS systems can park themselves and use radar to help avoid collisions. This is only one milepost on a long road that leads to fully autonomous highway driving. Most current ADAS technologies are only one step (SAE Level 1) beyond fully manual driving. Still, IDC expects SAE Level 1 vehicle shipments to experience an 10.1% CAGR through 2024. In 2024, 1 in 5 cars will in developed regions will offer SAE Level 2, with partial driving automation features. Honda recently revealed a SAE Level 3 vehicle.

Getting from SAE Level 1 to driverless, fully autonomous vehicles (AVs) will take many years and incremental advances of individual features. Those advances depend on software. What's the difference between a safe, reliable autonomous vehicle and one that acts as if a



tunnel painted on a brick wall were a thoroughfare? Software—and specifically the AI algorithms that govern the functionality of ADAS and automated driving.

But what does this software rely on? Dependable AI algorithms require countless terabytes (TB), and even petabytes (PB), of data derived from untold hours of real-world driving. (One petabyte is equal to 1000 terabytes).

Giant carmakers, academic labs, and eager startups are all making this future a reality. NVIDIA, for instance, has its DRIVE Hyperion platform for collecting data from an array of cameras, radars, and other inputs. The more autonomy researchers aim for, the more data they need. Aiming for SAE Level 5 (complete autonomy) could require up to 20TB per hour per vehicle used for AI data recording, with data from all training vehicles pooling into a training data set of at least 20PB, according to *Automotive World*. When gathering enough data to train for levels 2 through 4 of autonomy, an ADAS research vehicle can record up to 150TB per day.

This industrial imperative—make AVs a reality ASAP—creates a data flow bottleneck. AI challenges of this magnitude rely on processing in hyperscale cloud data centers. The petabytes of data generated by vehicle fleets must reach those repositories. But how? Sending just 1.5PB—equivalent to data collected by 10 to 20 research vehicles—over an “enterprise-class” gigabit (~1000 Mbps) connection would require over 150 days. It is possible for data pipelines to become so backlogged that the outflow can never catch up, and most of that effort and expense is for nothing.

Imagine trying to get a complete education when missing over half of the classes. Essentially, this is what happens when most AI training data, such as that collected in our ADAS/automated driving example, goes unused.

Every industry handling massive data sets (often defined as 100TB or more) faces transport challenges.

Having more data is like having more water. It serves little purpose (and can even become dangerous) unless it’s put to good use. Successful enterprises realize that if their mass data sets can’t move in an agile, cost-effective manner and if the data cannot be easily accessed, the business value suffers.

THE GREAT FLOOD: 40 DAYS AND 40 PETABYTES (AT LEAST)

ADAS research is part of an ocean of next-generation applications that require data flows exceeding typical local-to-cloud conduit capacities. Other common use cases include media and entertainment, public safety video imaging, critical healthcare data transport, smart manufacturing, etc.

The MIT-published *Production Engineering Solutions* notes that "the average smart factory produces 5PB (that's 5 million GB) of data every week." Cisco confirms machine-to-machine (M2M) systems as the main driver of current data growth, growing to 50% of all connected devices by 2023. By then, there will be 3.6 devices for every person in the world, all collectively pouring out hundreds of petabytes per day.

The production of a television series generates roughly 14TB of digitally recorded content every week, ranging in resolution from 2K to 5K. That's roughly 2.2PB per season. With 8K on the horizon, a single series could easily generate 100TB per week or more. Productions are expensive. In order to produce the show and ready it for viewers, these masses of data cannot afford to wait days or weeks for bandwidth transfers.

Is all this data being used, though? Or is much of it either stagnating or evaporating into nothing?

Up to present, the answer seems to be mostly the latter. The 2020 Seagate report *Rethink Data* revealed that enterprise data is growing, in 2021 alone, at the average annual rate of 42%. IDC predicts an astonishing increase in total data generated from 64ZB in 2020 to 180ZB by 2025. Internet of Things (IoT) devices, especially cameras, and automated M2M interactions, which span from utility smart meters to healthcare device management systems, play key roles in this growth. Consumer data that starts small often aggregates into massive enterprise-level waves. Trying to manage this data can feel like filling shot glasses with a fire hose.

The *Rethink Data* report found that only 32% of that enterprise data gets activated, or put to use. This is in part because capturing, storing, and managing that deluge can be tricky.

And much of this difficulty has to do with the access to and the transport of mass data.

WHERE MASS DATA GOES

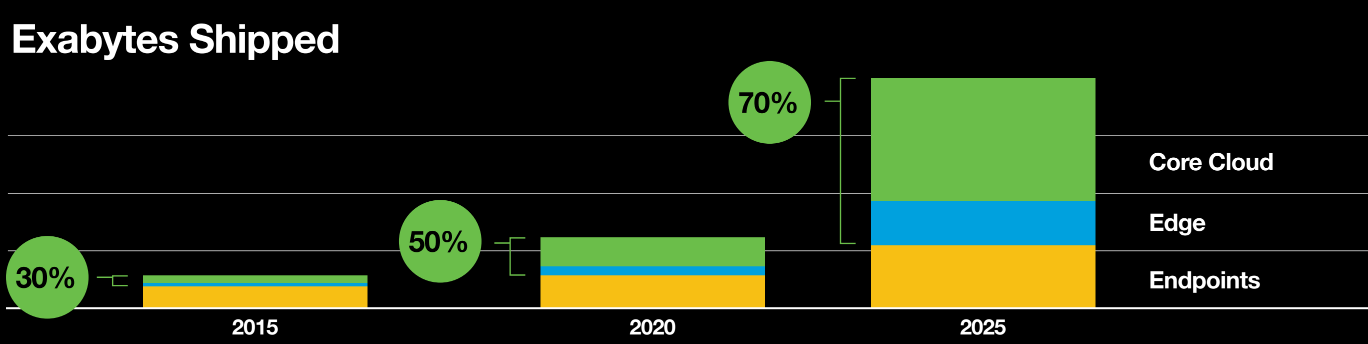
Ten years ago, enterprises debated between storing data in the public or private cloud. Now, the situation is far more nuanced. Data flows through local, edge, and cloud systems, but the pace and rapidly rising volumes of that movement must be accounted for. Enterprises now employ multicloud and hybrid cloud models, which can help optimize where data gets stored and how to best distribute, access, and use it.

In the modern data economy, more than ever, mass data sprawls. According to the 2021 *IDC Storage Systems & Infrastructure Trends Survey* and the Future-Proofing Storage report that this survey informed, 47% of enterprises use a centralized cloud storage architecture. In two years, that number will fall to 22%. Conversely, 25% of respondents currently have a hybrid storage architecture (a combination of both centralized and edge locations); that number will rise to 47% in two years. The shifts are provoking more mass movement of data. Modern enterprises centralize two-thirds of their data into data centers and public cloud sites, according to IDC's survey.

The IDC Cloud Data Storage & Infrastructure Trends Survey revealed that 40% of respondents already store 10PB to 49PB in the public cloud. Another 12% fall in the 50PB to 99PB range. Enterprise data grows fast (42% annual growth rate projected this year alone, according to *Rethink Data*). It is often the backbone of enterprise and consumer data-driven services. According to IDC's Global DataSphere 2021, the worldwide digital universe will grow from 64ZB in 2020 to 180ZB in 2025—that's an annual growth rate of 23% over that span.

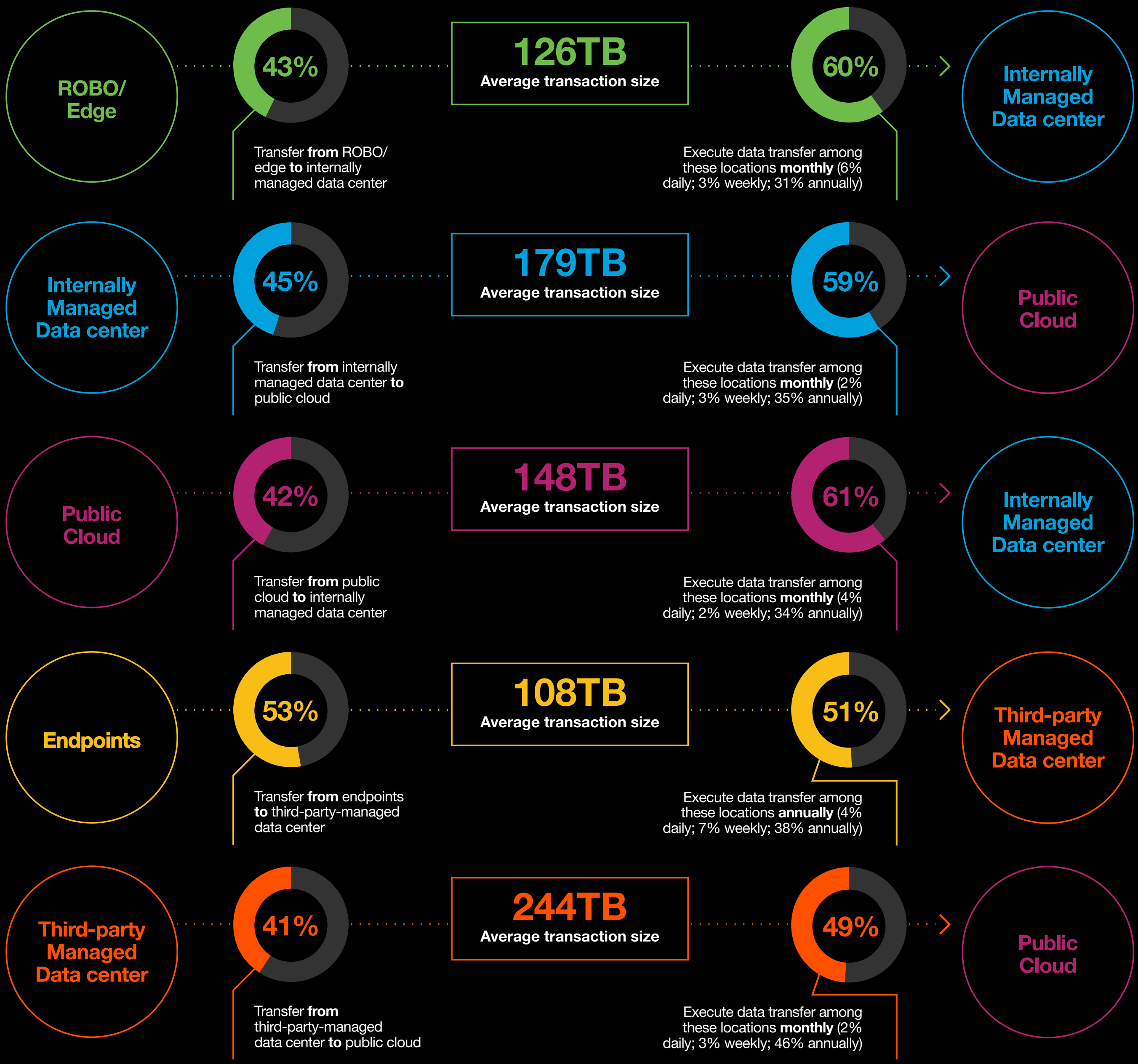
Take a look at how this trend is reflected by the ratio of storage exabytes needed by enterprises and shipped by Seagate Technology in 2020 versus in 2015, and the forecasting for mass-capacity demand for 2025. The following graph drives home two points: The amount of data is rising exponentially, and that data is headed for the local, edge, and core domains, with the last two growing the fastest.

An increasing proportion of data is shifting to the core cloud and the edge. In 2015, that amount was 30%. In 2020, 50%. In 2025, the projected percentage is 70%.



Source: Seagate Technology

Enterprises are regularly transferring large sets of stored data among storage locations. Take a look at average data transaction sizes among the various parts of the data infrastructure, as reported by IDC in *Future-Proofing Storage*.



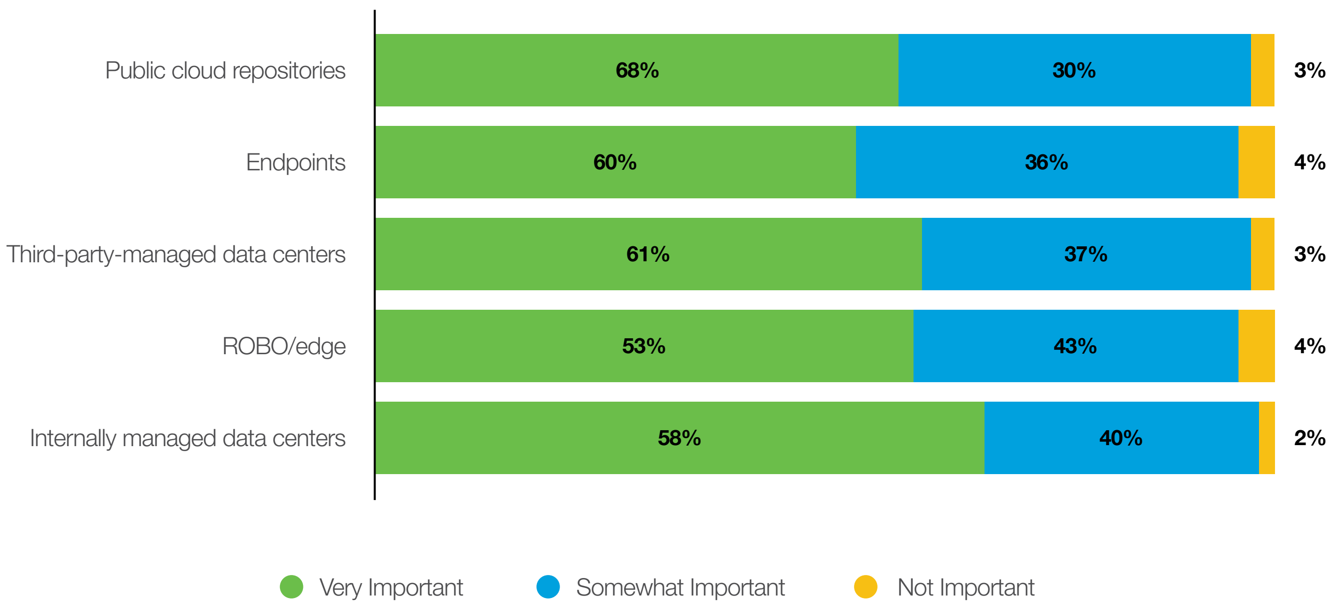
Source: IDC Cloud Data Storage & Infrastructure Trends Survey, sponsored by Seagate Technology, January 2021

The total capacity available for an average organization's physical data transfer is 473TB. The transferred data sets regularly exceed the amount at which massive data sets start: 100TB. The average size of a data set physically sent from ROBO/edge locations to an internally managed data center is 126TB. The average data set shipped from endpoints to a third-party-managed data center is 108TB. (Remarkably, 4% of companies send this transfer *daily*.)

All these vibrant data flows among different storage locations happen partly in order to put data to use, which means security proximity to applications.

Data Collection Must Migrate Quickly to Data Applications for Maximum Utility

Q. For enterprise data stored in the following locations, how important is it that this data is collected adjacent to applications?



Source: IDC Cloud Data Storage & Infrastructure Trends Survey, sponsored by Seagate Technology, January 2021

THE FASTEST MOVER

Because enterprise data sets keep increasing while **network capacities** do not keep up, the conventional method of transport—over the network—may no longer suffice.

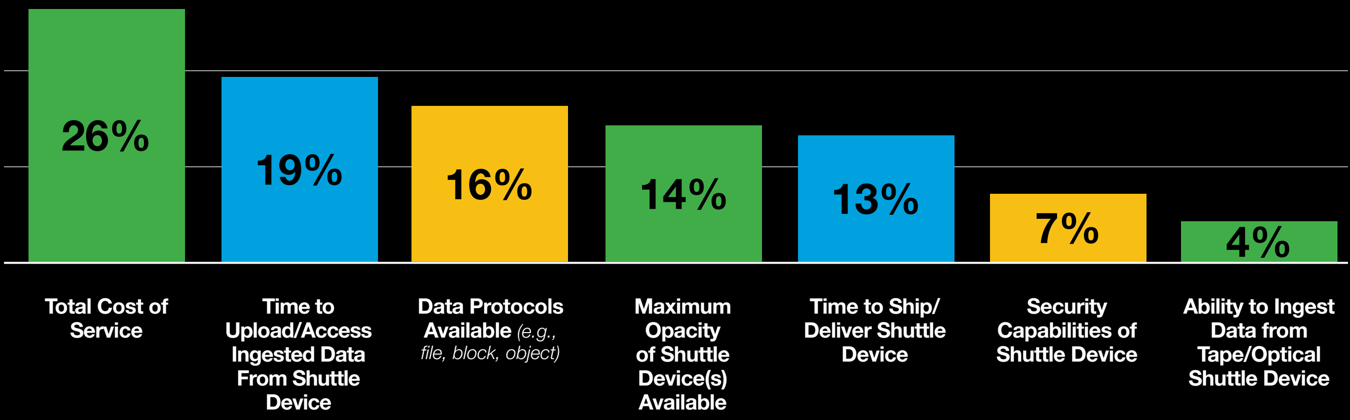
In the world where streaming content for work and relaxation is taken for granted, it's easy to forget how long enterprise-scale data transfers take. **Latency** impediments are common. Consider the binary alignment/map (BAM) files, which contain the sequences, base qualities, and reference alignments used in genomic sequencing. While BAM sizes can vary considerably, 100GB per file is not uncommon, and a modest sample collection for analysis might easily require 80TB. Moving such databases requires hours if not days. With top-grade infrastructure, an enterprise might be lucky to move 1PB (1000TB) across the U.S. in one full day. Now recall that many modern organizations are managing dozens of petabytes.

In addition to network and latency constraints, another barrier to data access and movement is the insufficient amount of **fiber**. Access to direct fiber-optic networking can vary widely by region, country, and local area. For example, Staten Island, New York has a lot of fiber availability; Main and Puerto Rico have almost no fiber access. Lack of access to top-speed connectivity will obviously impair a company's ability to move rising amounts of data in a timely manner.

In what is a mystery to no business leader, **cost** is another big factor that constrains the access to and the movement of enormous data sets.

Many storage providers offer a services menu loaded with costly complexities and caveats. According to the surveys that informed the *Future-Proofing Storage* report, in 2020 enterprises spent an average of $460,000 annually on storage solutions and services. Nearly all (99%) incurred egress fees as they moved their data from storage providers to other resources—a necessary action if data is to offer business value. When asked, "What was the top factor influencing your organization's choice of data transport/migration solution?," the greatest percentage of respondents, 26%, pointed to total cost of service.

Q. What was the top factor influencing your organization's choice of data transport/migration solution?



Source: IDC Cloud Data Storage & Infrastructure Trends Survey, sponsored by Seagate Technology, January 2021
N=683; those that use physical data transfer solutions

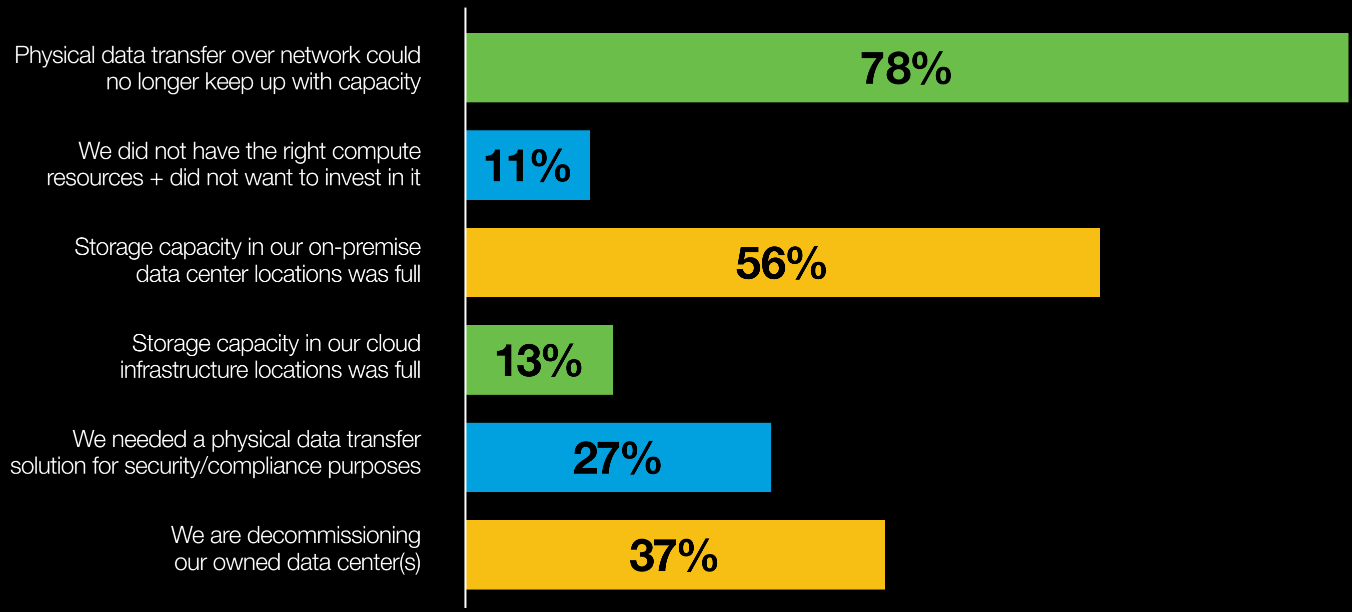
Data security and compliance also plays a key role in organizations' decisions regarding data access and movement. Security ranked as the top response for public cloud challenges and physical data transportation concerns. Four of every five respondents felt concerned about their ability to comply with privacy laws. As many as 27% of respondents in the *Cloud Data Storage & Infrastructure Trends Survey* noted that they opted for physical data transport "for security/compliance purposes." Data ownership and compliance policies often restrict where data can be sent and processed. Country-specific laws often constrain how data may move across borders, and some cities mandate that all municipal data remain in their municipality.

In addition to network and capacity constraints, slow speeds, limited access to fiber-optic networking, cost, security and compliance concerns, **storage capacity** limitations also propel the movement of data.

In the same survey, 78% respondents indicated that data transfer over their networks could no longer keep up with capacity.

Over half—56%—of survey takers reported that "storage capacity in our on-premises data center locations was full." Data volumes are rising so fast that local storage tanks are literally overflowing, compounding issues presented by network constraints and inefficiencies. As a result, both access to and the movement of mass data are impeded.

Q. Which factor(s) drove your organization to use physical data transport/ migration solutions?



Source: IDC Cloud Data Storage & Infrastructure Trends Survey, sponsored by Seagate Technology, January 2021.; N=683; those that use physical data transfer solutions

The IDC survey also revealed what types of massive workloads enterprises manage, with three dominating the list:

- Data warehouse: 70%
- Multimedia content: 66%
- Transactional data: 64%

Note that these workloads, or parts of them, may need high accessibility and responsiveness in the short term, but be more appropriate for cost-effective backup or archival storage over the long term. For example, an organization may need very fast access to new data within the first days of its creation but then may not touch the data again for months. This may mean that some data is sent to a storage-as-a-service cloud for backup. A more granular storage strategy can cost-optimize for these changing priorities.

To recap, network and capacity constraints, slow speeds, limited access to fiber-optic networking, cost, security and compliance concerns, and limited storage capacity are reasons why mass data is on the move—and why enterprises increasingly choose physical data transport.

MORE DATA, MORE EDGES

In the 21st century, data creation is particularly vibrant at the edge, making it, on the one hand, a key catalyst of data flow. On the other hand, the edge also complicates this flow.

The decade-old model of choosing whether to keep data local or in the cloud has proven too simplistic. Today, enterprises have a broader diversity of data types, exploding data volumes, more clearly staged data workflows, and varying performance demands at each stage. Increasingly, edge systems are a vital part of data flow and storage strategy.

The edge can be found at the periphery of any network, from highway intersections to manufacturing floors. It is where data gets made and data-driven decisions take place. The edge has become more granular as specific application needs have developed. We can now view the edge as three concentric rings surrounding the network core, each with its own attributes and benefits.

This report adapts its edge definitions from a joint white paper by NVIDIA and Equinix, "Artificial Intelligence: From the Public Cloud to the Device Edge." The paper defines the edge as comprising enterprise-hardened systems and devices that aggregate, distribute, and process data from sensors and devices. And it divides it into three sub-edges.

Mega Trends Driving Need for Mass Capacity Storage at the Edge

Source: Seagate Technology

Today's edge can be seen as three differentiated concentric rings surrounding the network core, each with its own attributes and benefits.

MICRO EDGE

This is where most node data collection happens. Micro edge—situated closest to the outer boundary of the network and the endpoints—typically offers sub-5 ms latency, which is needed for true real-time performance. The need for such immediacy is clear in fields like machinery failure monitoring and traffic control. Micro edge systems can reside anywhere from cellular tower bases to office closets to military field stations, or even the back of a Humvee. Corporate possibilities include enterprise getaways, branch offices, and remote operations. Micro edge data collection devices are often external storage drives connected via edge servers to endpoint nodes either wirelessly (private 5G, Bluetooth, Wi-Fi, etc.) or by wireline (especially USB or Ethernet). Throughput will be as high as device interfaces allow.

METRO EDGE

As its name suggests, the metro edge exists at the major city-level scale, not necessarily at the city's physical boundary. A metro edge system will provide 5 to 10 ms responsiveness, which is still sufficient for real-time performance in fields including cloud gaming and telemedicine. (Nobody wants remote surgery with network lag.) A metro edge facility might take the form of a small data center at a headquarters office building or a limited number of resources at a private colocation (colo) facility. At the opposite end of the metro scale, large colos, interconnected colos, multitenant data centers (MTDCs), regional data centers, content delivery networks (CDNs), and telco data centers are all possibilities. Metro edge sites will have much larger storage capacities than micro edge locations. As shown on the infographic toward the end of this report, the metro edge data centers are ideal for long-term backup needs. Their combination of proximity and capacity makes them strong choices for applications including transactional databases and streaming media. For businesses pursuing a tiered storage strategy, the metro edge offers a balance between local storage and nearline or archival storage.

MACRO EDGE

If the network core is a nationally or globally distributed data center service with its effectively infinite storage scaling, then the macro edge is one step removed from that. Macro edge sites generally provide 10 to 20 ms responsiveness and are large facilities, only hosting 10 or fewer tenants with larger deployments—roughly between 5 and 100 miles from endpoints. Depending on the distance and network conditions between points, access times from local nodes to macro edge sites may run higher. These sites are more likely to be colocation facilities or full data centers with redundant backbone lines rather than a metro edge system, which would prioritize connections to other metro sites. Macro edge sites may be large-scale private cloud facilities with application performance and storage space close to that of major core clouds, which can be essential when handling workloads bound by compliance restrictions. Such workloads could include AI training, online transaction processing (OLTP) databases, large-scale ecommerce hosting, and scale-up mass storage.

Given the push to enable organizations and applications with real-time decision capabilities, sending data back to the core for processing is often not feasible. It's easy to understand why in cases like anomaly detection in smart manufacturing and automated driving research, which call for split-second decisions. That work must be done at the edge, with the specific edge resources chosen to reflect factors from storage capacity to budget and infrastructure bandwidth. Later, of course, as in the case of ADAS research, some of the data already processed at the edge is often transported to the cloud for ML/AI analysis and learning.

Hyperscale data centers at the macro edge and core continue to excel for centralized applications, including large-scale archiving, massive content distribution, scale-out application storage, and big data analytics. However, as more data is collected, processed, and applied outside the traditional data center, a complementary model makes sense: cloud with edge. This can take the form of compact metro edge data centers assisting with some of the data load from larger sites, especially when lower latency to endnodes is advantageous.

DRIVERS OF THE EDGE

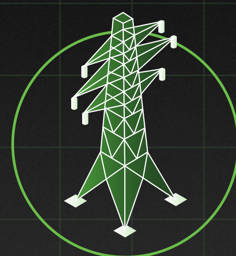
We noted at the outset of the previous section that the edge is a catalyst of data's growth. That growth—of the edge and of the data at the edge—owes a lot to five tech drivers. Together, they have ushered masses of data to the edge, and enabled the movement of data between the edge and the multicloud core.

Artificial intelligence (AI)

AI automates and accelerates cognitive tasks normally handled by humans, often with higher accuracy. The degree of AI accuracy achieved improves with the training data set's size. This training will benefit from running on the larger, aggregated resources of a data center. However, obtaining those results in real-time may require the trained AI to execute on-device at the edge. If a tree suddenly falls in the road, approaching AVs don't have time to send data to the cloud and back. Split-second, AI-processed results need to happen under the hood or at the micro edge. Of course, automotive is only one of countless segments AI benefits. PwC expects AI to add up to 26% to local economies by 2030, boosting the global economy by $15.7 trillion.

The Internet of Things (IoT)

Many of the previously-mentioned machine-to-machine (M2M) systems involve IoT devices, which can be anything from air pressure sensors to security cameras and smart refrigerators. The more sensors a system uses, the more data it collects—and the better the predictive model built that can result. IoT data sets may become quite large, though. Statista figures show total IoT data volumes growing by nearly 500% from 2019 to 2025. Thus, having performant compute and storage resources at the edge will be needed to keep IoT-driven applications running at peak levels and to allow IoT-sensor and log-file data to constantly stream into the cloud for AI/ML and DevOps processing.



5G

4G LTE mobile networks were good in their day, but they were not made for a world filled with billions of IoT devices. Beyond offering improvements in latency and throughput, 5G networks provide connectivity for up to 1 million devices per square kilometer—an order of magnitude denser than 4G LTE. (6G is expected to achieve a 10X density increase over 5G.) This 5G leap will be essential for environments packed with high-bandwidth IoT sensors and systems capturing real-time data. IDC expects private LTE/5G network deployments to exhibit a 43.4% CAGR through 2024, and all those networks will depend on edge servers. 5G doesn't solve mass data needing transport (it is often insufficient for this need). But 5G has contributed to the rise of masses of data created and consumed at the edge.

IT/OT Convergence

Traditional manufacturing has relied on operational technologies (OT) often decades old for good reasons: it works, it's secure, and it's installed. But only recently have OT and IT systems begun to mingle, largely thanks to IoT advances. By blending these two worlds, manufacturing floors can begin to harness the analytics that IT has enjoyed without sacrificing security. Bringing IT/OT convergence to the manufacturing floor requires micro and often metro edge investment across many industries. This convergence leads to higher production efficiency, improved maintenance and failure prevention, and greater profitability. One Cisco report points to 49% lower defect rates, 48% lower unplanned downtime, 17.5% lower energy costs, and 23% improvement in new product introduction times.

Edge Data Centers

To augment the cloud at the edge, a new class of data center has emerged. The edge data centers exist at the micro edge, the metro edge, and the macro edge. The data centers at the edge are both a result of data creation at the endpoints—and the multiplier of that data's growth. A private cloud data center can be found at the macro edge; a medium-sized storage-as-a-service data center at the metro edge; and a small, micro-regional edge data center in locations such as the base of cell towers.

The unprecedented convergence of AI, IoT, 5G, IT/OT, and edge data centers creates a unique mix of technology and economics, making it practical to assemble, store, and process vast amounts of data at the edge. As we'll see in the following section, these factors conspire to shift the ecosystem's data gravity—providing one more good reason to keep mass data agile.

Section Two

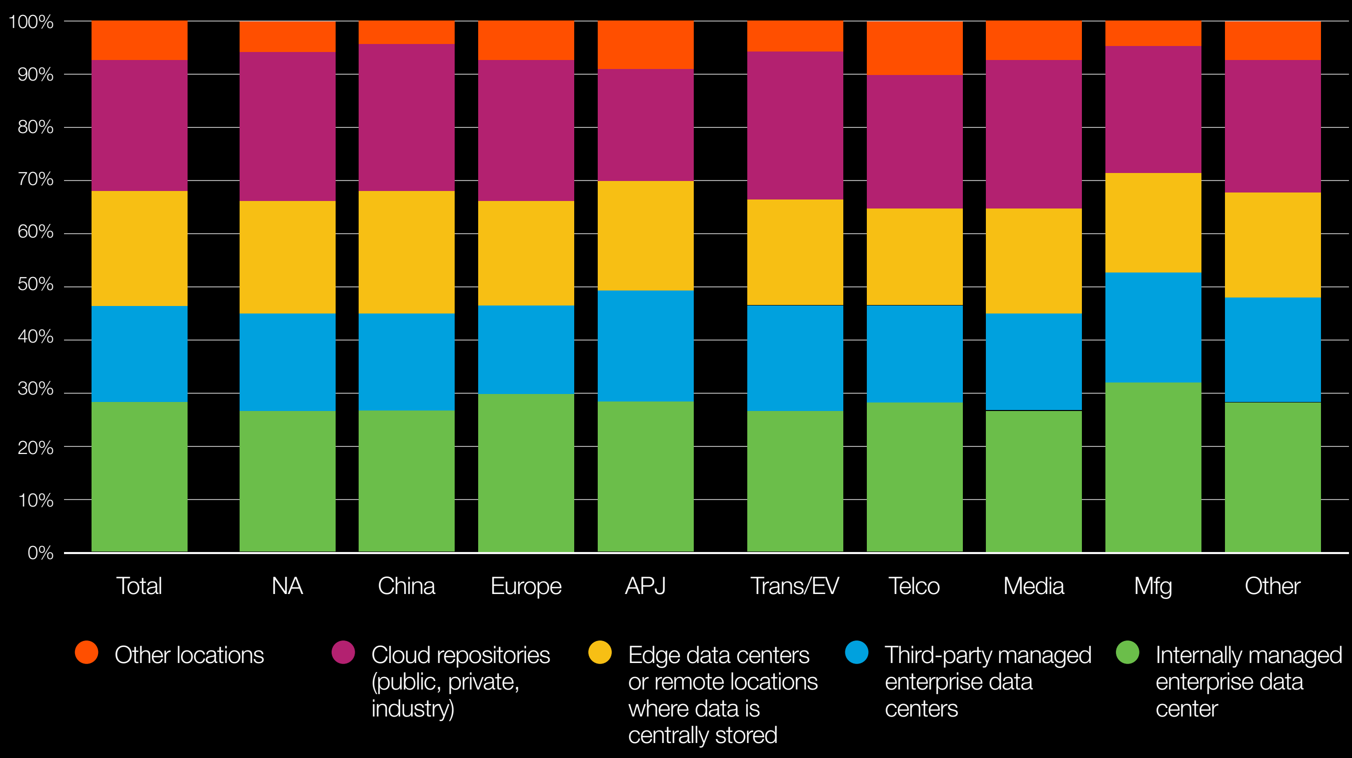
THE PULL OF DATA GRAVITY

Clearly, edge networks can help address enterprise needs, and trends show edge adoption is booming. Fueled in part by multi-access edge computing in the telecom space for 5G enablement and widespread growth in AI applications (such as ADAS and automated driving research), the global edge computing market is ballooning. Grand View Research estimates its CAGR at 37.4% through 2027.

This shift impacts storage, as well. As noted in the “Where Mass Data Goes” section, while 47% of enterprises today use a centralized cloud storage architecture, in two years that number will fall to 22%. Conversely, 25% of respondents currently have a hybrid storage architecture spanning centralized and edge locations; that number jumps to 47% in two years.

The *Rethink Data* report arrives at a similar conclusion. While variances exist across geographies and data types, large data sets will be spread broadly across different cloud and edge resources by 2022.

Q. Where Data Will be Stored in 2 Years (on Average)



Source: IDC, The Seagate Rethink Data Survey, 2020

The above graph also speaks to local storage. As we can see, "other locations," which include node devices and client systems, occupy at most 10% of total storage. Edge occupies nearly 20%, but let's not forget that edge—per Equinix and NVIDIA's report cited above—also spills over into metro edge's colocation data centers and the macro edge's data centers as well.

Specifically identifying what facility falls into which category is less important than the overall point: Data is flowing across the endpoint-to-core ecosystem in more ways than ever before, which helps to put that data next to applications and keep those applications running at peak performance. Let's turn now to one other force that threatens this flow.

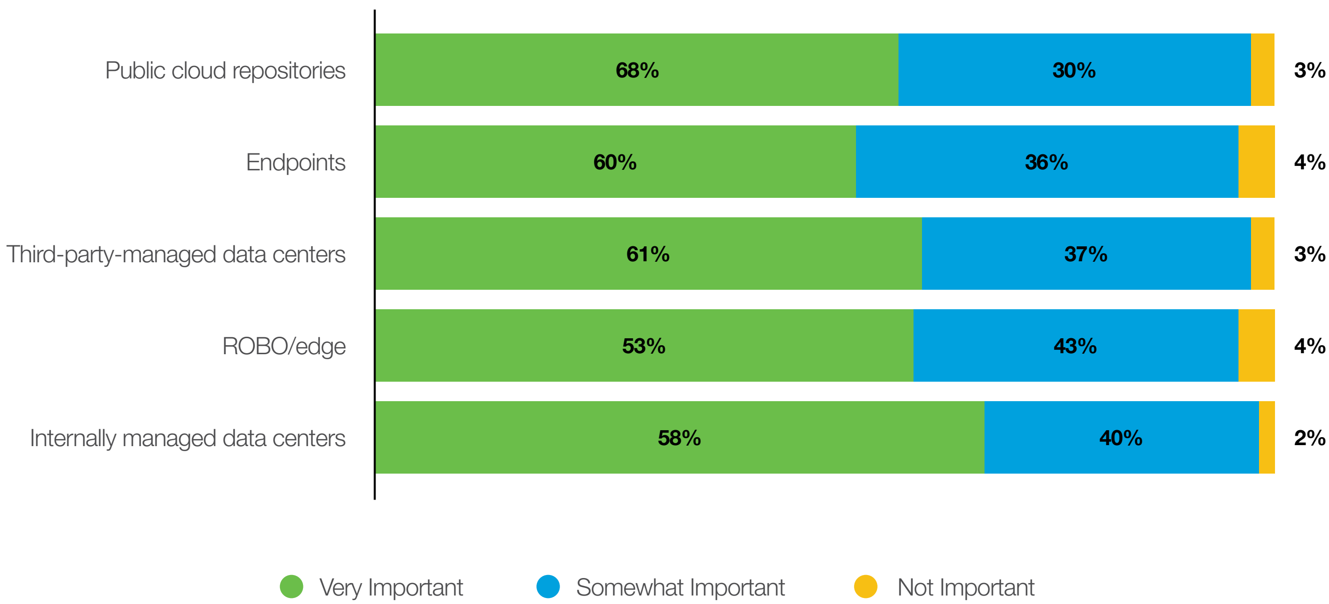
DEFYING GRAVITY

Just as stars form from scattered nebula dust particles that accrete over time, data exhibits its own gravity. The larger the data mass, the greater its gravitational force in attracting applications, services, and ever more data. This can be useful.

But sometimes masses can grow too large, inadvertently clogging up data highways with even the most beneficial of applications and services. As IDC notes in the *Future-Proofing Storage* report, “massive data sets risk becoming ‘black holes,’ trapping stored data, applications, and services in a single location.”

To prevent data black holes, IDC advises companies to collocate data with its associated applications, no matter where those applications reside—rather than drawing all those applications to the edge in all cases. This can happen, for example, at the metro edge in storage-as-a-service (STaaS) data centers, which are located closer to where data is created, ensuring ease of access and favorable latency. IDC’s surveying reveals that over half of respondents already pursue this strategy.

Q. For enterprise data stored in the following locations, how important is it that this data is collected adjacent to applications?



Source: IDC Cloud Data Storage & Infrastructure Trends Survey, sponsored by Seagate Technology, January 2021

To offset the excessive pull of data gravity, IDC advises enterprises to keep that gravity's presence in mind and "ensure that no single data set exerts uncontrollable force on the rest of the IT and application ecosystem."

One doesn't have to be a rocket scientist to know that escaping a gravity well requires a suitable spaceship. In this case, that ship may take the shape of a company vehicle or security van carrying a payload of petabytes. As noted earlier, connection bandwidth can be a serious bottleneck in data flow between storage sites and tiers. This will likely make data volumes continue to expand.

This, again, brings us to fault-tolerant, rugged, vendor-agnostic portable storage shuttles.

Portable data storage shuttles are increasingly favored as the best means of moving mass data. Fast and convenient, they migrate large data much more quickly than a wide-area network.

SAFETY, PRIVACY, AND TRUST

A key caveat remains: While physical data transport plays a key role in thwarting excessive data gravity, that transportation must be handled with all the safety precautions as data managed inside the firewall. In fact, managing security and throughout transportation can be more convoluted than managing data at rest. Data must remain encrypted throughout the transportation sequence, from export onto portable media through ingestion and management at the receiving site. More broadly, this is also true as data moves from nodes to edge to core.

Some compliance and data sovereignty requirements may have specific restrictions on data movement, though. These will likely be driven by protection of privacy and proprietary information. Organizations may have strict limitations on how or even if they can shuttle data sets off-premises or across borders. According to the IDC survey, 80% of enterprises are concerned with their organization's ability to comply with existing privacy laws.

Understandably, businesses' concern when evaluating storage solutions and services is whether those resources have sufficient security and compliance capabilities. In particular, any good solution will need to address three security aspects based on the idea of computational trust:

- As devices enter and exit the network, they must be trusted to participate in the data domain.
- The data itself must be trusted.
- Data in motion is more vulnerable than data at rest, so nodes must communicate securely.

Establishing a trusted network of devices and data from local capture to the edge and on to the core will ensure that edge applications can continue to develop and provide greater organizational value by keeping the data flowing.

Section Three

LYVE RESHAPES THE EDGE

Against this backdrop of rapid ecosystem change and mounting storage challenges that set mass data on its journeys, Seagate engineers saw unmet needs. To meet those needs, the company created the Lyve portfolio, a range of edge-to-cloud mass storage products and services that address the challenges associated with mass data sets. The challenge is to get enormous quantities of data from collection, through ingest, and distribution across the range of edge layers, often all the way to the core. Along the way, that data must meet a spectrum of requirements concerning security, speed, compliance, orchestration, and cost-effectiveness. Lyve was designed from the ground up to succeed on all fronts—enabling enterprises to frictionlessly store, move, access, and activate great volumes of data.

MASS DATA ON THE GO

A Data Set on Wheels

What are the best ways to transfer massive volumes of data? That's the 100TB question. In order to get the most value out of business data, enterprises need more efficient ways to enable its flow. Every industry handling massive data sets—from 100TB to multiple petabytes—faces transport challenges.

Below we show the data journey in real-life research required to achieve levels 2 through 4 of automated driving as well as in driver assistance systems (ADAS) use cases. The transport and infrastructure solutions apply to other data-rich workflows.

Endpoints

Sensors on research and development vehicles capture loads of data. That data can't stay captured on a Lyve Mobile Array drive in the trunk of the car: In order to be converted into valuable insights, it needs to move.

20 & MORE sensors in each research vehicle

Each car can record up to **150 TB** per day

>150 days That's how long it would take to send 1.5PB from 10-20 cars via an enterprise-class connection.

Micro Edge

The more autonomy researchers aim for, the more data they need. The garages and data centers that house autonomous and advanced driving research fleets need to facilitate at least 5PB of data each. The more relevant data sets are sorted before heading by cargo trucks and freight airplanes to 3 locations for storage and high-performance processing.

The vibration-resistant, rugged Lyve Mobile solution connects to the data logger in the car's trunk.

It provides easy, fast ingestion and extraction of **TBs** of data.

Lyve Mobile enables quick detachment of the data storage shuttle.

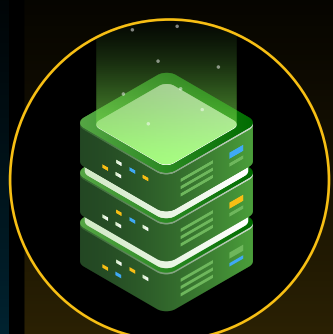
From the garage, data can go to any of these 3 locations:

Metro Edge

Some data can be stored in a STaaS repository like Lyve Cloud for long-term backup.

Data that is not immediately valuable is often useful later. In the Lyve Cloud, it's stored cost-effectively and frictionlessly, without vendor lock-in.

Macro Edge



In on-prem data centers, data relevant to the business and needing extra security can be kept on TCO-optimizing storage systems powered by mass-capacity Exos and Nytro drives.

The Multicloud Core

After getting to the multicloud by freight or by air via Lyve Mobile, the unstructured data captured by the vehcile sensors is ready to deliver insights. It is processed in the high-performance compute layer by ML/AI algorithms—and converted to lessons that are critical to the development of safer and more efficient autonomous vehicles.

About 30% of the data from the research vehicles ends up in the multicloud for ML/AI learning.

The ML tools process road-testing information to identify critical solutions.

The insights are then fed back into the fleet, improving AI & the safety of future AVs on the roads.

LYVE MOBILE

Physical data shuttles like the ones from the Seagate® Lyve™ Mobile portfolio can provide a more cost-effective and faster solution for large data ecosystems.

The data journey begins with the countless clients and devices generating often-unstructured data—picture a public-safety camera, a warehouse desk, or a battery-powered Arctic field station. In order to deliver value, the terabytes accumulated at the edge on media cards and client drives often need to be aggregated into a highly rugged collection container. Moving the data efficiently and turning that data into insights can provide a distinct competitive advantage. Physical data transfer is often the most efficient way to do this. Seagate developed unique enterprise-grade data mobility solutions that are delivered as a service. Designed to be modular with monthly and annual service plans available, the Lyve Mobile solution enables enterprises to scale up or down.

For data sets that are smaller in size and require NAS connectivity, the Lyve Mobile Shuttle provides the right combination of performance and capacity without the need of a host PC; it's available in capacities at least up to 16TB (as of mid 2021). Equipped with a quad-core CPU, AES-based Seagate Secure Encryption, and a low-power, E Ink touchscreen, this is more stand-alone system than ordinary portable drive. The drive works well with other external storage without the need for a host PC. Lyve Mobile Shuttles connect via USB 3.1 Gen 2 or 10 GbE, allowing I/O to flow without connection bottlenecks.

For larger data sets at the outer edge, the Lyve Mobile Array offers a quick bridge between the field and rackmount storage. The easy-to-transport enclosures contain six enterprise-grade Seagate SAS hard drives or SSDs, with RAID protection, capacities at least up to 96TB (as of mid 2021), Seagate Secure AES encryption, and connectivity via Thunderbolt 3 (40Gb/s), USB 3.2 (10Gb/s), as well as PCIe Gen3. Lyve Mobile Arrays do require a host for operation, but they're designed to plug and play into a Lyve Mobile Rackmount Receiver, a 19" rackmount frame that accepts two Lyve Mobile Array units. The Lyve Rackmount Receiver features redundant power as well as SAS, iSCSI, and Fibre Channel connectivity for extremely fast ingest of these large data sets. Seagate also offers the Lyve Mobile Array Shipper—a wheeled, rugged shipping case designed for extreme conditions.

In the ADAS and automated driving research use cases (pictured in the infographic that puts it all together), data from onboard sensors is recorded directly onto the shuttle. All an operator needs to do at the end of the recording day is to detach Lyve Mobile from the trunk, and ship it to wherever it needs to go—for example, to an on-prem data center or to the multicloud core for AI/ML analysis. The information is protected, encrypted, and ready for fast, efficient offloading. Lyve Mobile Shuttle and Array enclosures are purpose-built for physical transportation to ingest facilities.

MASS-CAPACITY DRIVES AND SYSTEMS

Data's future business value is not always fully known today. For this and other reasons, it's risky to discard potentially valuable data. In the AI world, all other things being equal, the companies with the largest training data sets tend to emerge with superior solutions. In addition to potential future research use, it is savvy for businesses to save data on premises for reasons related to archiving, compliance, intellectual property protection, business value, and other purposes.

For enterprises that want to scale their data centers with industry-leading density, Seagate recommends enterprise-class hard drives (HDDs) and solid state drives (SSDs), as well as the storage systems powered by these devices. The self-healing, high-density CORVAULT and other data storage systems are built on trust, affordability, and ease of use.

Mass-capacity Exos® HDDs and high-performance Nytro® SSDs are engineered for modularity, capacity, and performance. They enable private data centers to integrate powerful, scalable storage within traditional environments or build new ecosystems from the ground up in a secure, cost-effective manner.

LYVE CLOUD

Seagate created Lyve™ Cloud as a storage-as-a-service (STaaS) complement to enterprises' other storage options. Some data center needs are universal: multiple levels of security, compute performance options for S3 workloads, unlimited scale-out capability, and exceptional availability. Thanks to Seagate's collaboration with Equinix, Lyve Cloud is located closer to where data is created—at the metro edge. At its simplest, the Lyve Cloud is a world-class object storage service. This is where data can find permanent, cost-effective residence, be activated for a host of applications, and be instantly available for flowing to edge locations via high-speed backbone links.

Cost structure simplicity and predictability may be Lyve Cloud's most compelling value proposition. Unpredictability of billing and avoidance of unforeseen fees are among the key drives motivating companies to move out of the public cloud. With Lyve Cloud, there is only one cost: the amount of storage used. All other charges—retrieval, ingestion, I/O, export, egress, and so on—are included. Lyve Cloud does not seek to replace existing storage services. Rather, it provides a more cost-friendly object storage service that can easily interact with other clouds and data centers, using its own strengths to complement.

Lyve Cloud will be conducive to always-on, longer-term storage priorities while also satisfying advantageous TCO needs, availability, privacy, and data security needs. (In the infographic on page 26, you'll find Lyve Cloud providing backup to research data in a colo center at the metro edge). What Lyve Cloud can offer is desired by enterprises complementing their data lakes while transforming their IT and building an active archive for data that is accessed by different or specialized applications. Concurrently, those raw data sets can flow to the macro edge and core services for compute-intensive operations, such as automatic point of interest labeling and post-processing.

Together, the Lyve portfolio of solutions work to remove impediments that stand in the way of enterprises putting their data to use. They exist to seamlessly support data activation, access to data, its frictionless transport, and increased security.

From node collection of terabytes through gathering petabytes at the micro edge to moving exabytes into metro, macro, and core data centers over time, Lyve covers every essential step. Data can flow freely to where applications need it, and where it can operate with minimal latency and resistance. Because of Lyve's easy scalability, volume bottlenecks cease to be an issue. All costs are known and predictable. And Lyve's alacrity to both public and private cloud deployment make data sovereignty and compliance simple and efficient to accommodate.

All this because mass data has places to go.

SEAGATE

seagate.com