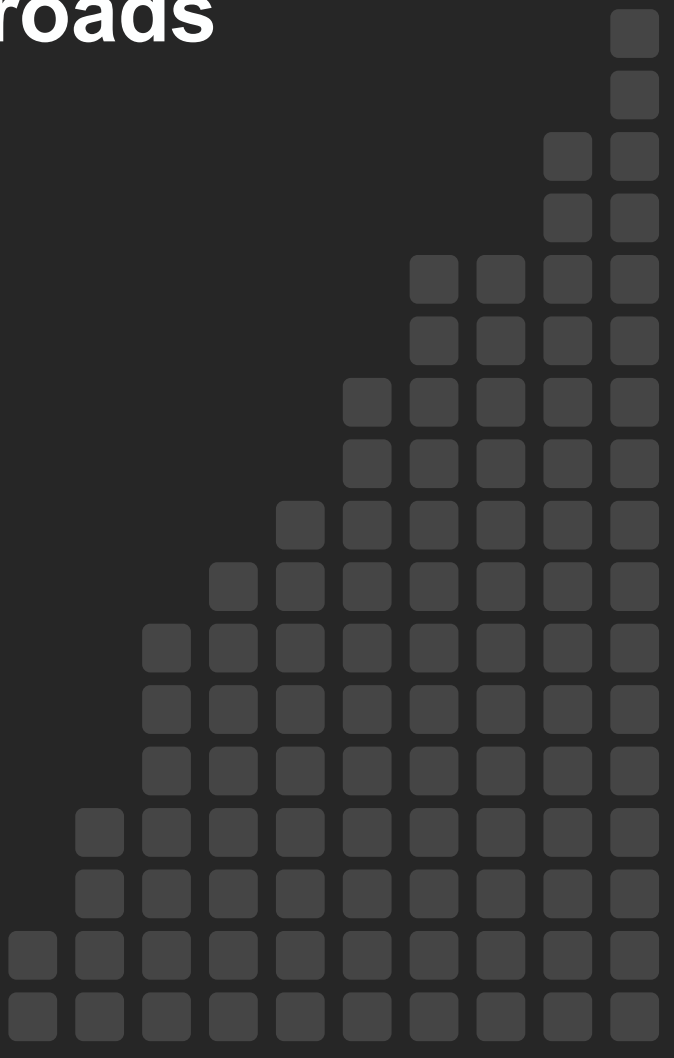
Enterprise Strategy Group™
by TechTarget

RESEARCH REPORT

Cloud Data Protection Strategies at a Crossroads

By Christophe Bertrand, Practice Director and Christian Perry,
Director of Syndicated Research
Enterprise Strategy Group

August 2023



Contents

Executive Summary

Report Conclusions

TechTarget's Enterprise Strategy Group conducted an in-depth survey of 397 IT professionals in North America (US and Canada) familiar with and/or responsible for data protection technology decisions for their organization, specifically around data protection and production technologies that may leverage cloud services as part of the solution. Based on the data collected as part of this project, the report illustrates:

- **Against a backdrop of ballooning data volumes, organizations continue to struggle with data protection.** Data volumes continue expanding across both on-premises and public cloud environments, putting data protection strategies to a demanding test. In fact, much of the data in these volumes is secondary, which can include backup and archive data, as well as other collected data that organizations use (or plan on using) to derive insights as their analytics capabilities evolve. Despite this data deluge, 53% of organizations only reassess these strategies every two to five years, potentially leaving themselves exposed to data loss risks. This behavior is somewhat different when focusing exclusively on cloud data protection strategies, which see 16% of organizations using an ad hoc approach based on new infrastructure or platform deployments. This will likely press organizations to reassess or rearchitect their data protection strategies more often moving forward, as ad hoc methods can easily introduce complexity, if not chaos spawned by cybersecurity events.
- **Data stored on public cloud services is not safer than on-premises data.** A longstanding argument around the safest deployment location for data—on premises or public cloud—continues with no clear winner in sight. In fact, organizations need to recover cloud-based data more often, as 44% of organizations said they had to recover data four or more times in public cloud environments in the last 12 months, compared with just 21% of organizations that recovered data that often on premises. The cloud picture does not improve when examining the percentage of successfully recovered data, which shows more organizations successfully recovering on-premises data than organizations recovering cloud data. Cloud-based data protection, backup, and disaster recovery services can be more limited than on-premises, as organizations may be forced to use the cloud service provider (CSP)'s tools. Other differences, including distribution of data across cloud servers or storage devices, also challenge cloud data recovery efforts.
- **Cloud service adoption is strong despite skepticism.** Most organizations either use or plan to use backup-as-a-service (BaaS) and disaster recovery as-a-service (DRaaS) to protect on-premises *and* public cloud-based applications, workloads, and data. Although adoption and plans are strong for both types of services, they are somewhat stronger for BaaS, which can be attributed partly to ongoing confusion around the services' capabilities. More than two-thirds of organizations feel BaaS and DRaaS are roughly the same, differentiated only by unique attributes delivered by vendors, and most feel a better term for the services would be data protection-as-a-service. Further, over 40% of organizations feel that public cloud data protection just serves as another potential target for ransomware, suggesting these organizations are using these services regardless due to the specific requirements of their environments (e.g., data sovereignty).
- **Multi-cloud is "the Wild West" of data protection.** As organizations increasingly deploy applications, workloads, and data on multiple cloud services, it is clear they struggle to apply holistic data protection to the highly diverse multi-cloud ecosystem. Most organizations use a best-of-breed, stovepipe approach for multi-cloud backup and data protection whereby specialized tools are used for each of their public cloud environments rather than a single, unified platform that protects all environments holistically and consistently. Agility, performance, and reliability are critical for modern, cloud-native environments, but this common siloed approach does not fit that mold, likely due to multi-cloud limitations of available platforms. Further, this approach leads to a larger overall footprint, which increases the potential for damaging cybersecurity events. Organizations are likely to push vendors for increased support of all their multi-cloud environments to avoid the resource, management, and complexity challenges inherent with patchwork data protection.

Introduction

Research Objectives

The broad adoption of public cloud services and containers as sources and repositories of business-critical data puts the onus on data owners to deliver on data protection SLAs for cloud-resident and container-based applications and data. Users are confused about the data protection levels that public cloud and Kubernetes environments deliver and about the changing protection options (DIY in the cloud, cloud-native third-party solutions, hyperscalers' built-in features, as-a-service, etc.). As vendors and the cloud ecosystem evolve and add as-a-service consumption options, end-users are making incorrect comparisons and assumptions as well as failing to select the key data protection capabilities they need to maximize their cloud technology investments. This confusion leads to lasting challenges, and the market is now at a crossroads.

To assess the state of cloud-based data protection and the as-a-service market (e.g., in cloud/to the cloud, BaaS, and DRaaS), TechTarget's Enterprise Strategy Group (ESG) surveyed 397 IT professionals in North America (US and Canada) familiar with and/or responsible for data protection technology decisions for their organization, specifically around data protection and production technologies that may leverage cloud services as part of the solution. This study sought to answer the following questions:

- How do organizations define backup-as-a-service (BaaS) and disaster recovery-as-a-service (DRaaS)?
- What is the adoption status of BaaS, DRaaS, and cloud backup/disaster recovery targets?
- What groups/roles within organizations are involved with the evaluation of and influence the purchase of public cloud-based data protection solutions? Which group/role typically makes the final purchase decision?
- How many times in the last 12 months have organizations had to recover data from on-premises and/or public cloud environments? What percentage was recovered on average in those cases?
- What were the reasons for data recovery efforts in the last 12 months?
- Would organizations consider a public cloud-based data protection solution that includes an on-premises cache or storage for local recovery to improve data recovery SLAs (e.g., RPO)?
- What approaches currently protect applications/workloads/data in public cloud infrastructure services?
- What types of data protection technologies are used in these approaches, and which assets are protected?
- How is critical public cloud-based unstructured data protected, and what are acceptable recovery times?
- What is the impact on teams of the daily management and maintenance of public cloud data?
- How many full-time staff are allotted for data protection objectives associated with cloud data?
- What methods do organizations use to protect data within virtual machines on public cloud infrastructure?
- What are organizations' preferred approach to protecting multiple unique public CSP environments?
- How do organizations estimate the costs of their cloud backups and recoveries for hyperscalers?
- What approaches do organizations take to ensure cost-efficient data tiering for the data protection storage supporting their public cloud infrastructure-resident applications?
- Does organizations' backup software handle the appropriate tiering of data written to object storage?
- How important is it to have a container backup and recovery management solution that works across multiple disparate public cloud infrastructure services going forward?
- Do organizations' container backup schemas integrate with their current data protection environment?

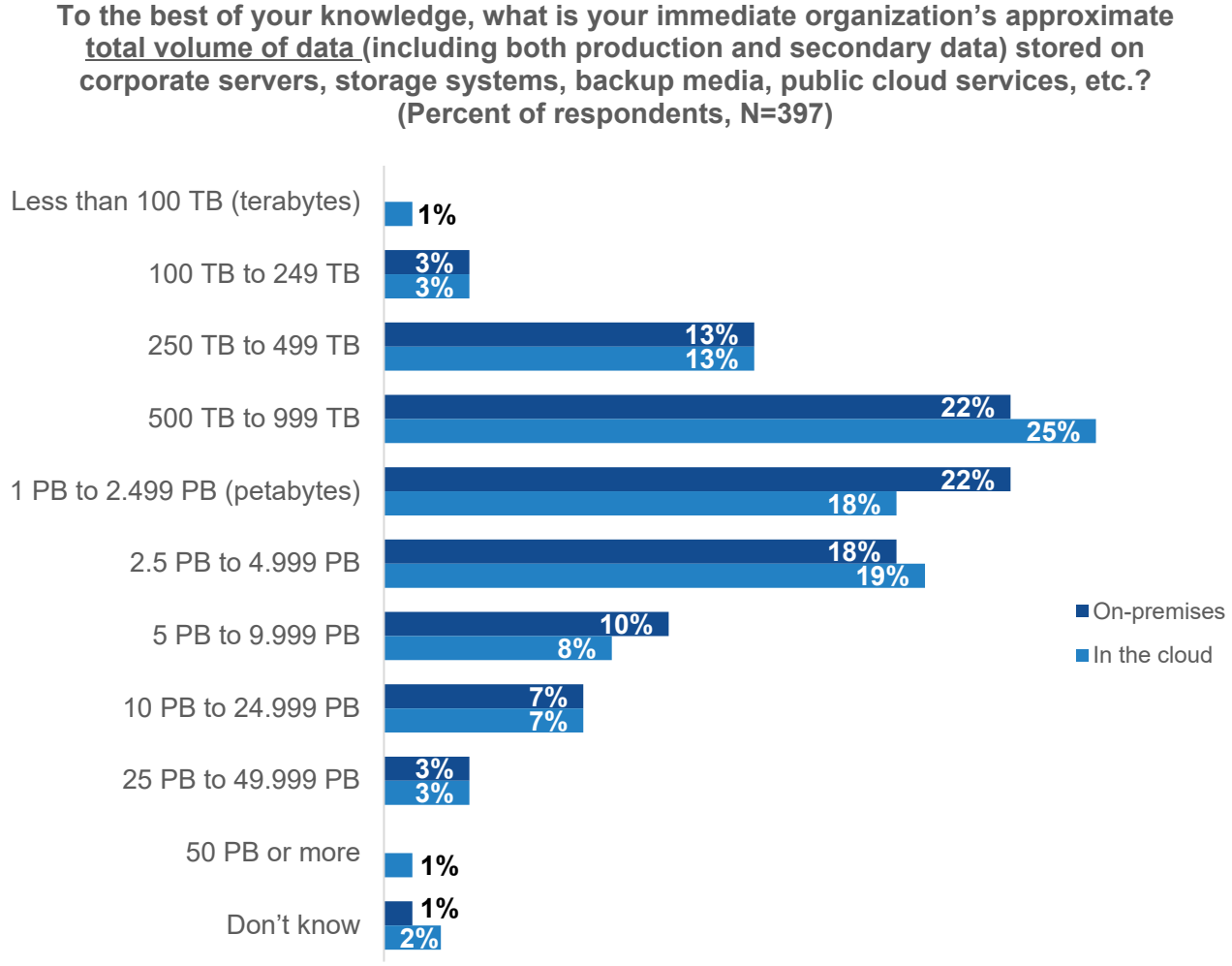
Survey participants represented a wide range of industries, including financial, manufacturing, retail/wholesale, and healthcare, among others. For more details, please see the *Research Methodology* and *Respondent Demographics* sections of this report.

Research Findings

Organizations Struggle With Data Protection Strategies as an Ongoing Data Deluge Continues

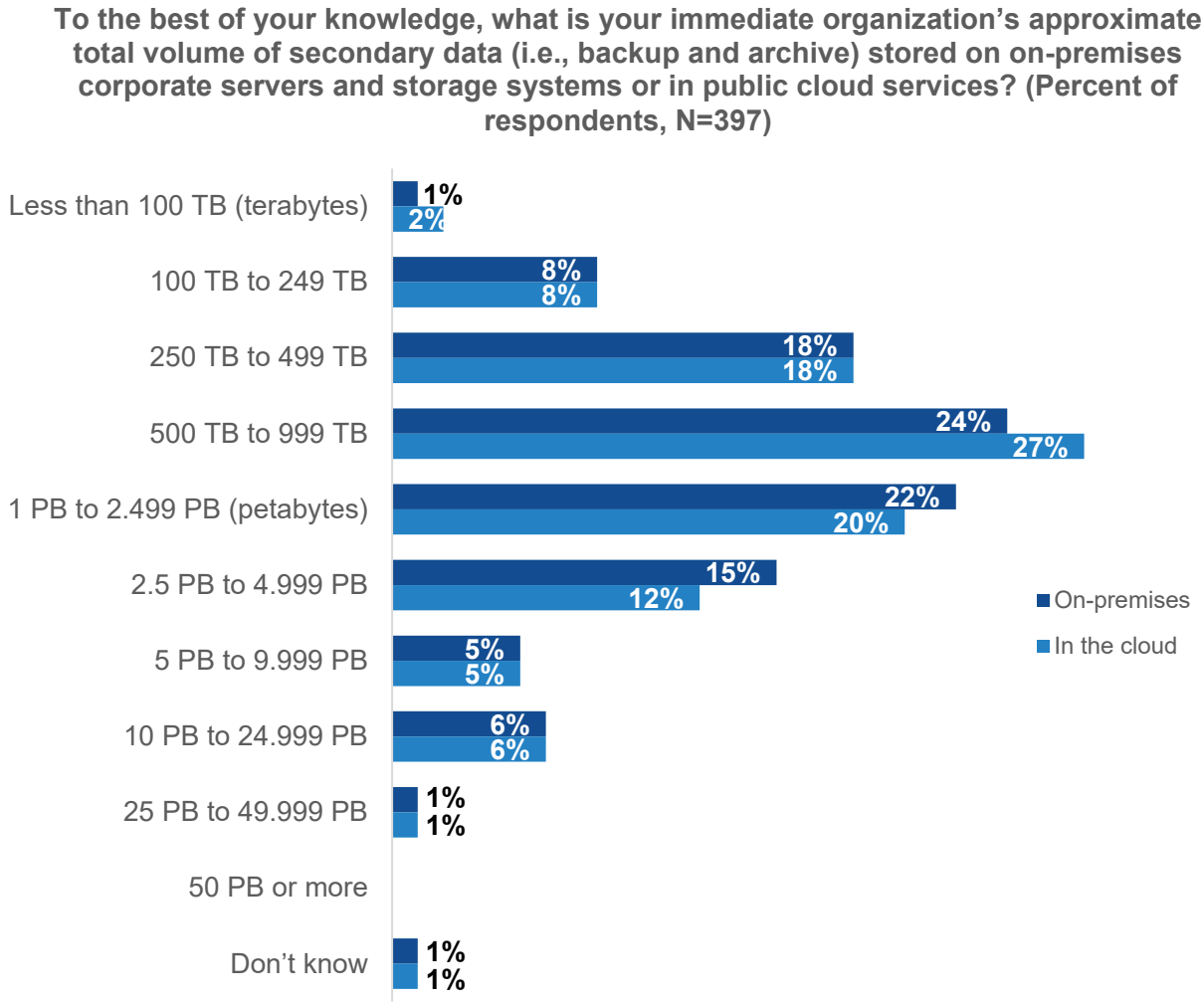
Data saturation continues in earnest across the entire ecosystem of organizations, with relatively even distribution between on-premises and public cloud environments (see Figure 1). As seen in Figure 2, most of this data is composed of secondary data, which can include backup and archive data, or data collected by organizations for use with analytics platforms. Traditionally, rising data primarily led to the increase of storage infrastructure in organization-owned data centers, where more storage capacity, servers, databases, and associated platforms would support more challenging availability and reliability requirements. Modern businesses, on the other hand, typically distribute data both on- and off-premises, in turn ramping up the potential for data breaches, privacy and compliance violations, and the likelihood of data loss in the event of system or service failures. In turn, secondary data now accounts for the bulk of stored data, as organizations recognize the value of collected data for future business insights and must also now protect throngs of data in both their own data centers and in the cloud.

Figure 1. Data Volumes Consistent Across On-premises and Cloud Environments



Source: Enterprise Strategy Group, a division of TechTarget, Inc.

Figure 2. Secondary Data Accounts for a Massive Share of Overall Data Volumes



Source: Enterprise Strategy Group, a division of TechTarget, Inc.

Despite the prevalence of data recovery efforts both on- and off-premises, a surprising percentage of organizations appear to underestimate the importance of reassessing data protection strategies annually or more often to minimize risks of data loss. For example, 53% of organizations reassess or rearchitect their general data protection strategies every two to five years (see Figure 3). However, this drops to 43% when specifically reassessing cloud data protection strategies. Further, while the percentages of teams that reassess general and cloud strategies annually is nearly the same, far more organizations (16%) reassess cloud strategies on an ad hoc basis compared with those using an ad hoc approach with their broader data protection strategy.

This difference is likely driven by the same factors that challenge organizations attempting to recover data from cloud-based data protection deployments. As previously stated, organizations typically have less control over cloud data protection management than with on-premises data protection (to be fair, this also is a key driver behind cloud-based deployments, as it frees organizations from infrastructure maintenance, software deployments, and other requirements that can heavily consume resources). Further, whereas organizations generally have a strong understanding of their own IT environments, cloud-based data protection deployments depend on cloud service providers' stated shared responsibility model, security, and reliability claims. Compliance can also be a larger

concern in the cloud, particularly as organizations expand their international presence and are more challenged by data residency and related requirements. This creates a highly dynamic environment that requires more regular strategic attention for some organizations.

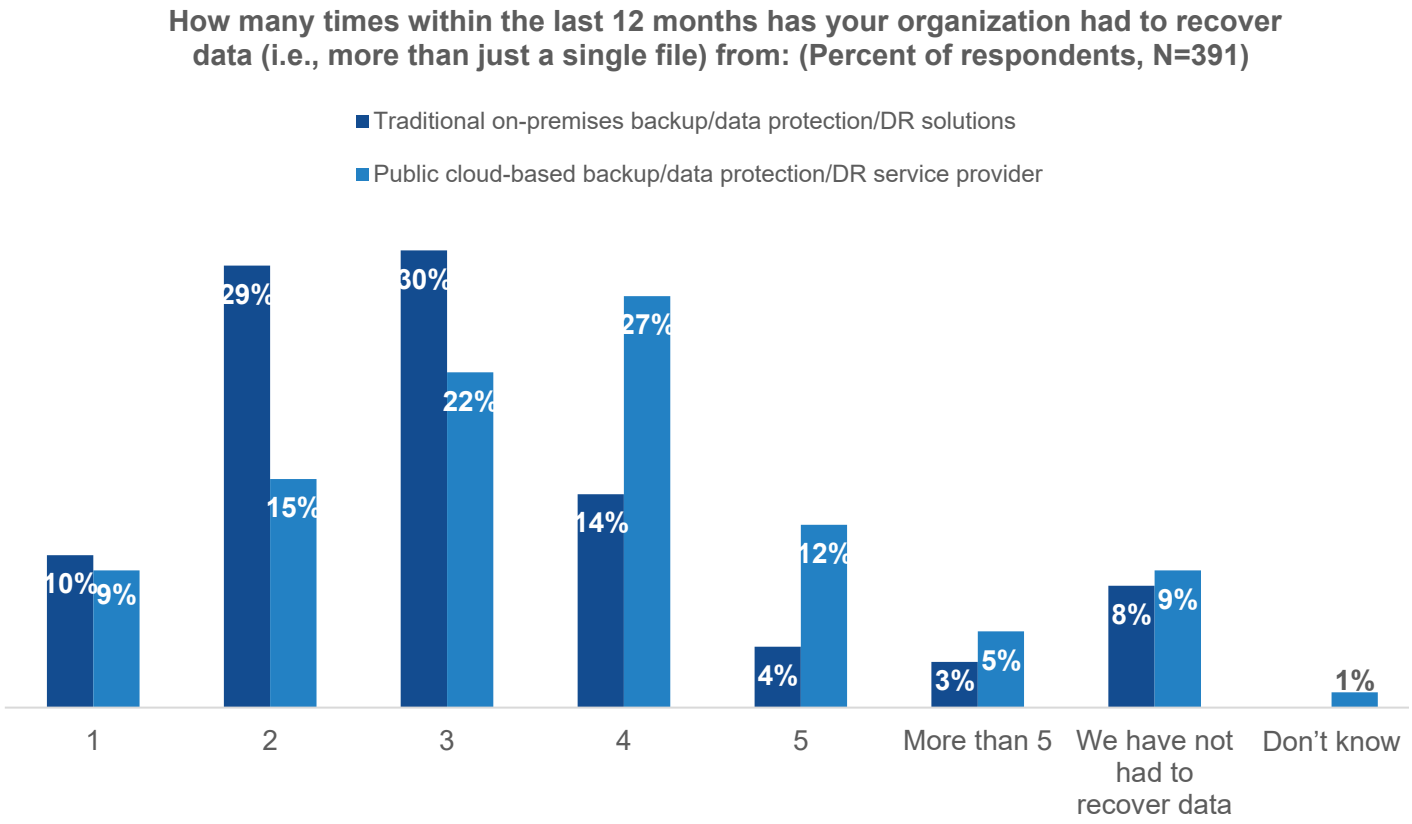
Figure 3. Many Organizations Remain Overly Lax on Data Protection Assessment

Source: Enterprise Strategy Group, a division of TechTarget, Inc.

Public Cloud Data Recovery Is Common But Leaves Much To Be Desired

Is data safer when stored in the cloud? The answer is a resolute "no." Although the need to recover data from traditional on-premises backup, data protection, and disaster recovery solutions remains common, organizations recover public cloud-based data more often than they recover on-premises data. For example, 44% of organizations said they had to recover data four or more times in public cloud environments in the last 12 months, compared with just 21% of organizations that recovered data that often on premises (see Figure 4). Considering the similarity of data volumes stored on premises and in public cloud environments, this recovery frequency indicates that public cloud environments may be less stable than on-premises data centers. However, this might also be a sign that data recovery from these environments is easier. In the latter case, organizations may be recovering this data more often because the process is less painful than what they are typically accustomed to.

Figure 4. Organizations Recovering Data Regularly Across All Environments

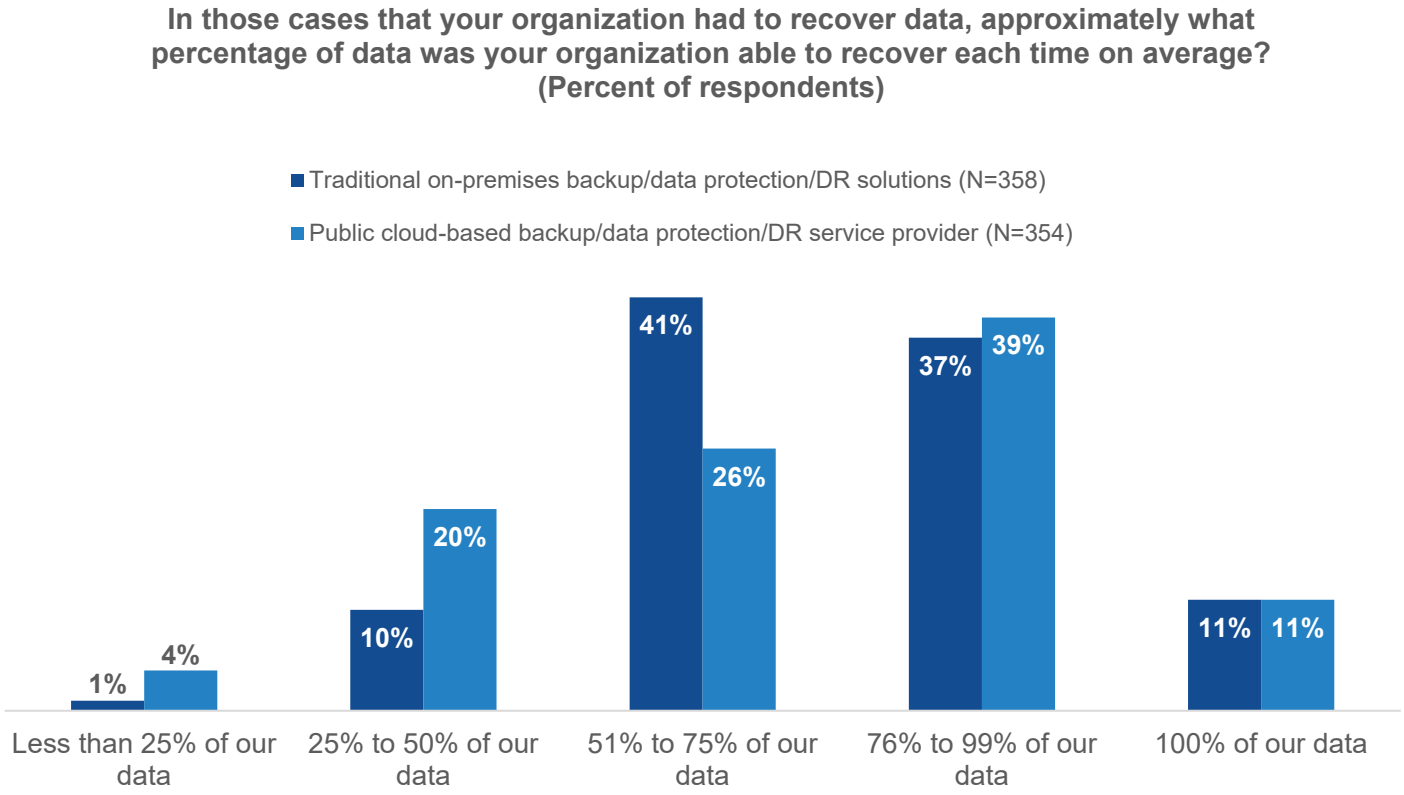


Source: Enterprise Strategy Group, a division of TechTarget, Inc.

Public cloud data protection metrics do not improve when looking closer at these recoveries. For example, 89% of organizations said they recovered more than half of their on-premises data during the last 12 months, compared with just 76% of organizations recovering data from public cloud (see Figure 5). So not only are organizations needing to recover cloud data on a more frequent basis than on-premises data, but they are less likely to recover more than half of their data.

Organizations generally have less control over their public cloud infrastructure when compared to on-premises infrastructure, including the freedom to use their preferred backup and recovery tools, because cloud service providers can limit these options to their own technologies. Further, depending on the cloud service in use, data might be distributed across several servers or storage devices, or even across different storage types (e.g., object, file, or block). These and other differences with cloud data storage can not only delay recovery efforts but also lower the feasibility of recovering all or most data.

Figure 5. Data Recovery Efforts in the Cloud Lag On-premises Effectiveness



Source: Enterprise Strategy Group, a division of TechTarget, Inc.

Cybersecurity events represent the most common driver of organizations' data recovery efforts in the last 12 months, at 54% of respondents, followed by system failures (44%), malicious deletion (37%), power outages (36%), and accidental deletions (35%) (see Figure 6). When cybersecurity events occur, all existing processes, best practices, and typical recovery efforts can require significant workflow adjustments, depending on the severity of the event. Unlike system failures or power outages, cybersecurity events are highly unpredictable (in form, function, and timing) and carry the potential for massive disruption in IT environments.

In turn, ransomware and cyber-attacks are most likely to be the biggest data protection concern for organizations, with 29% identifying those events as their top concern, compared with 18% each of compliance, accidental overwrite/deletion, and meeting employee/customer SLAs (see Figure 7). Organizations will increasingly push data protection vendors to build or expand solutions with integrated ransomware preparedness and recovery capabilities.

Figure 6. Cybersecurity Leads Reasons for Recent Data Recovery Efforts

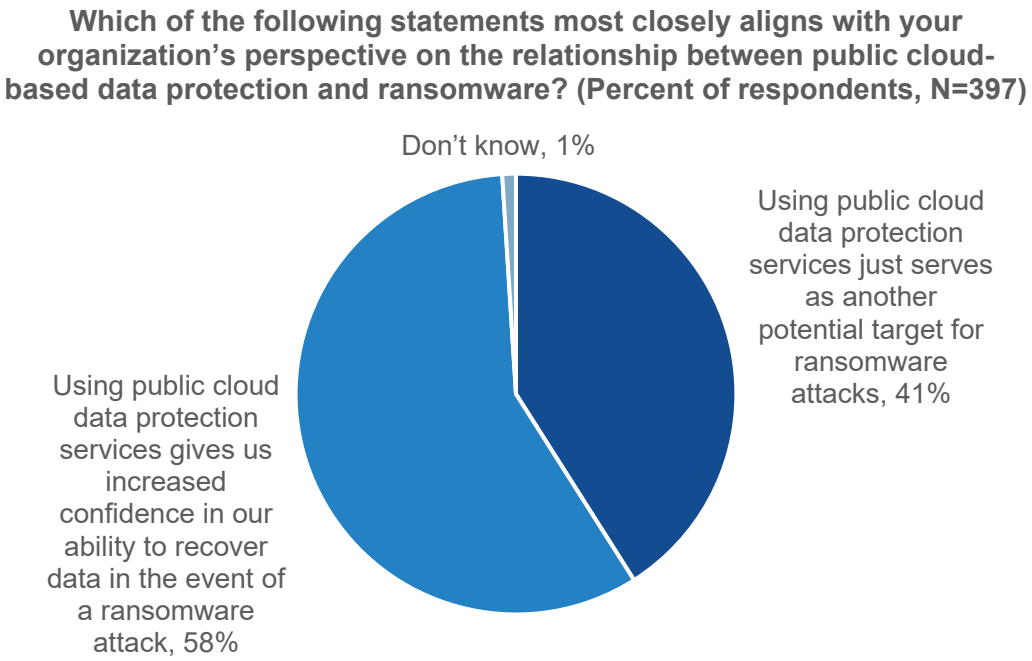
Source: Enterprise Strategy Group, a division of TechTarget, Inc.

Figure 7. Cybersecurity Events Also Represent the Most Common Top Concern

Source: Enterprise Strategy Group, a division of TechTarget, Inc.

In terms of ransomware, public cloud-based data protection remains a work in progress. While 58% of organizations said this data protection increases their confidence in recovering data in the event of a ransomware attack, a rather alarming 41% said it just serves as another potential ransomware target (see Figure 8). However, this is not necessarily a knock on cloud-based data protection because in theory, any application or service deployment—regardless of its location—serves as a target for ransomware. The larger the application and service footprint, the larger the risk.

Figure 8. Public Cloud Data Protection Services Not a 'Slam Dunk' for All



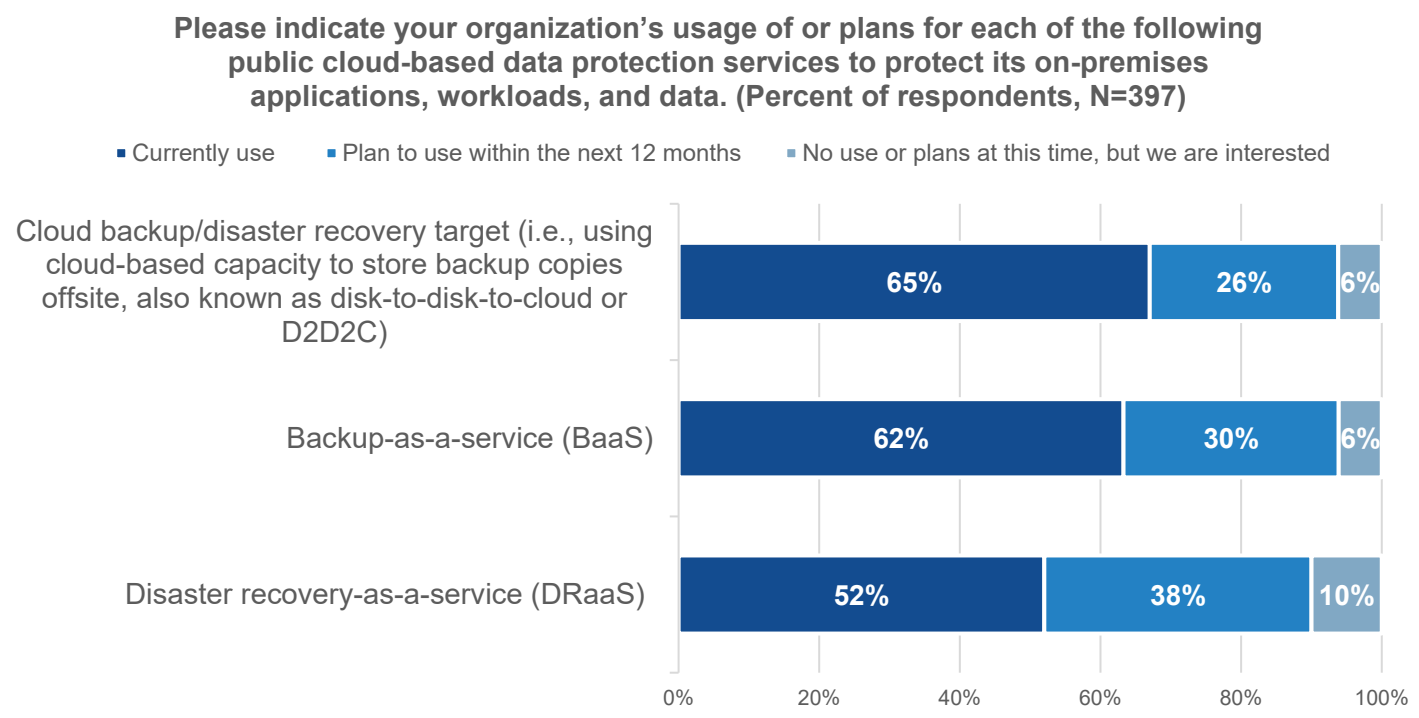
Source: Enterprise Strategy Group, a division of TechTarget, Inc.

Despite Skepticism, Cloud-based Data Protection Adoption Continues in Earnest

Despite the lack of trust evident in two of every five organizations in terms of the relationship between cloud-based data protection and ransomware, current adoption and plans for cloud-based data protection services are solid. To protect on-premises applications, workloads, and data, nearly two-thirds (65%) of organizations currently use a public cloud service as a backup and/or disaster recovery target, with another 26% planning to use one in the next 12 months (see Figure 9). Similar percentages were seen with BaaS. DRaaS has a lower percentage of current adopters (52%) but a larger percentage of planners (38%), suggesting that DRaaS will reach the adoption levels of the other two services for on-premises purposes in the near future.

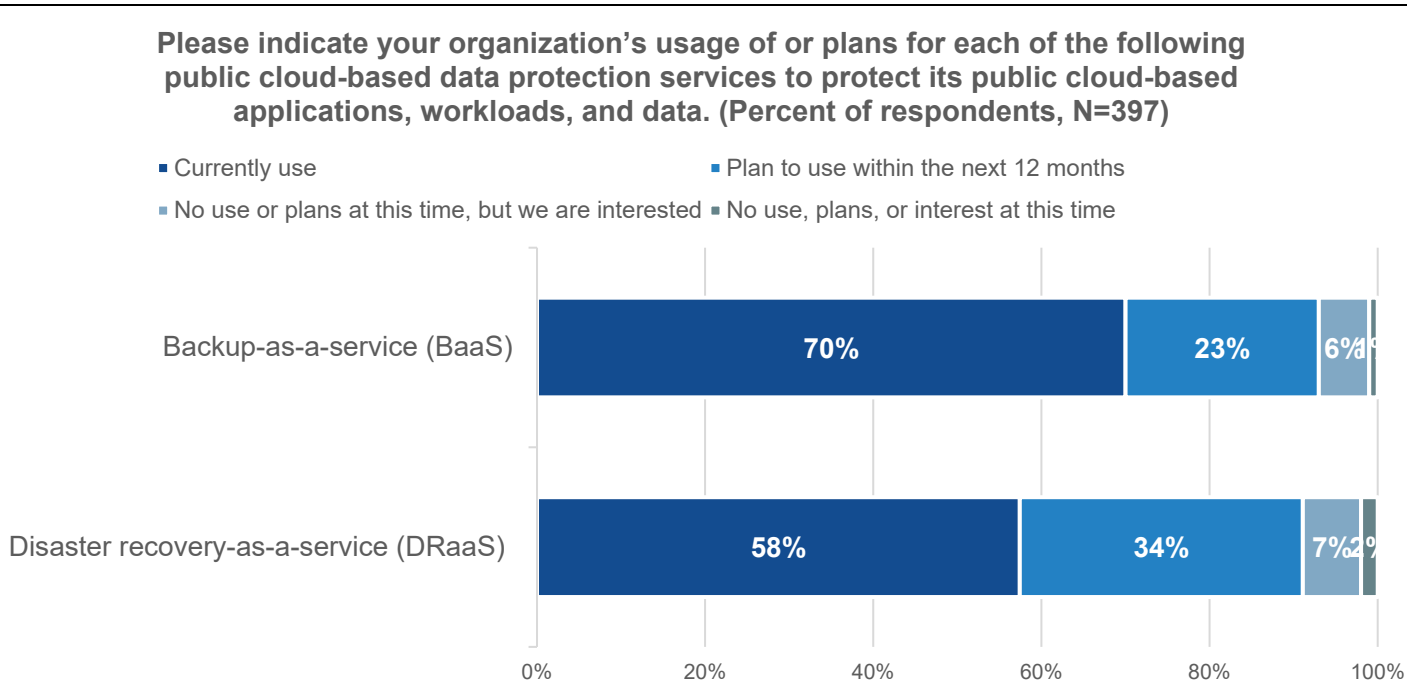
BaaS and DRaaS are also widely used to protect public cloud-based applications, workloads, and data, with an even larger delta between the two compared with the use of the services to protect on-premises data. To protect cloud-based data, 70% of organizations are currently using BaaS (with 23% planning to use it within 12 months), compared with 58% currently using DRaaS (with 34% planning to use it within 12 months) (see Figure 10).

Figure 9. Public Cloud-based Data Protection Services Enjoy Wide Adoption



Source: Enterprise Strategy Group, a division of TechTarget, Inc.

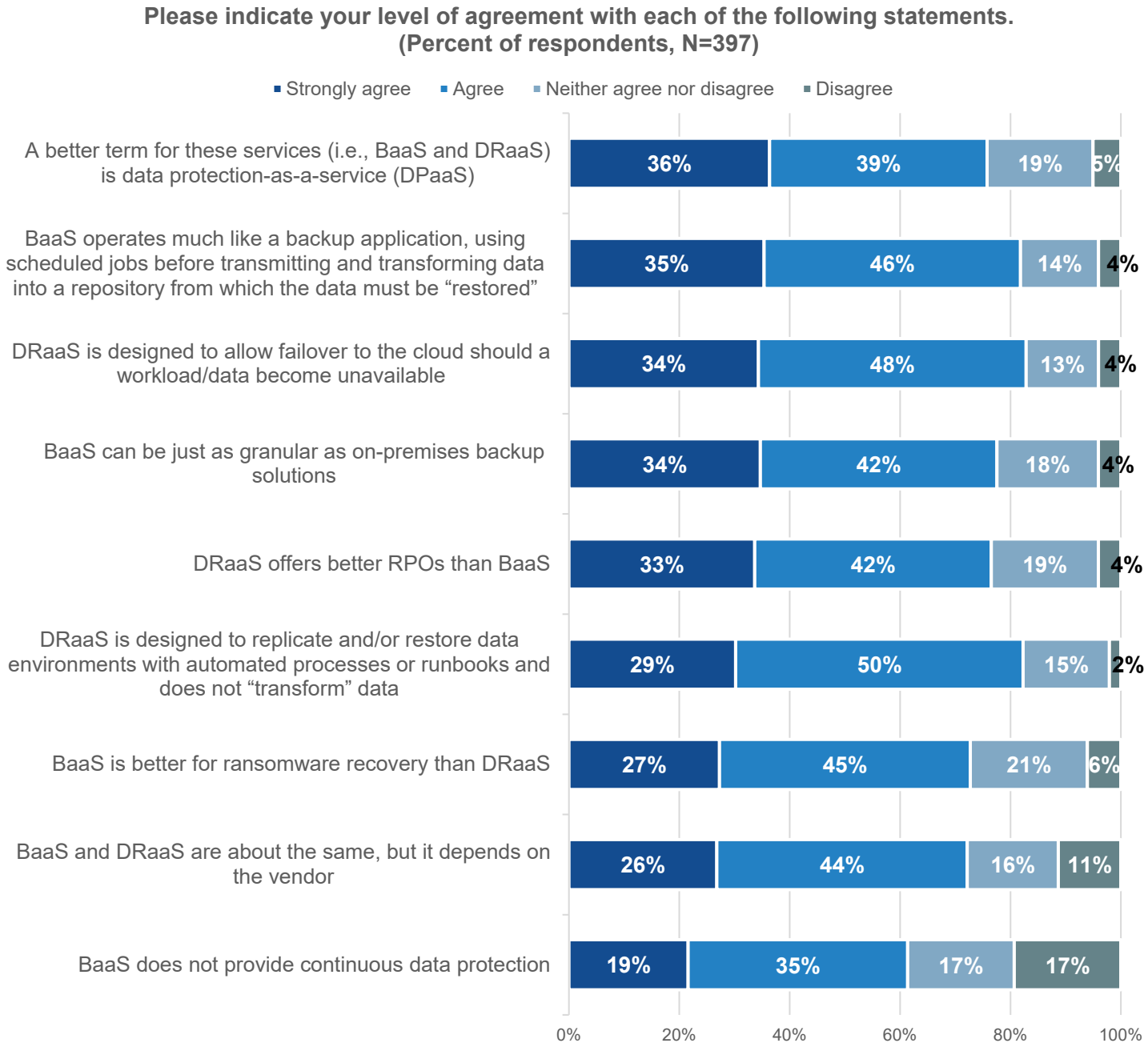
Figure 10. DRaaS Adoption Lags BaaS But Future Deployments Are Coming



Source: Enterprise Strategy Group, a division of TechTarget, Inc.

A potential explanation for the adoption disparity between BaaS and DRaaS is that there is no clear-cut definition for these services. For example, 70% of organizations said BaaS and DRaaS are about the same, depending on the vendor delivering the service (see Figure 11). However, 75% said DRaaS offers better RPOs than BaaS, while 72% said that BaaS is better for ransomware recovery than DRaaS. Meanwhile, 75% agree that a better term for these services is data protection-as-a-service (DPaaS). Considering more than two-thirds of organizations feel these services are roughly the same depending on the vendor, DPaaS could serve as a far less confusing term, as services can be evaluated solely on the capabilities they deliver without being thrown in a perceived limited bucket such as BaaS or DRaaS.

Figure 11. Organizations Unclear on Purposes of DRaaS and BaaS

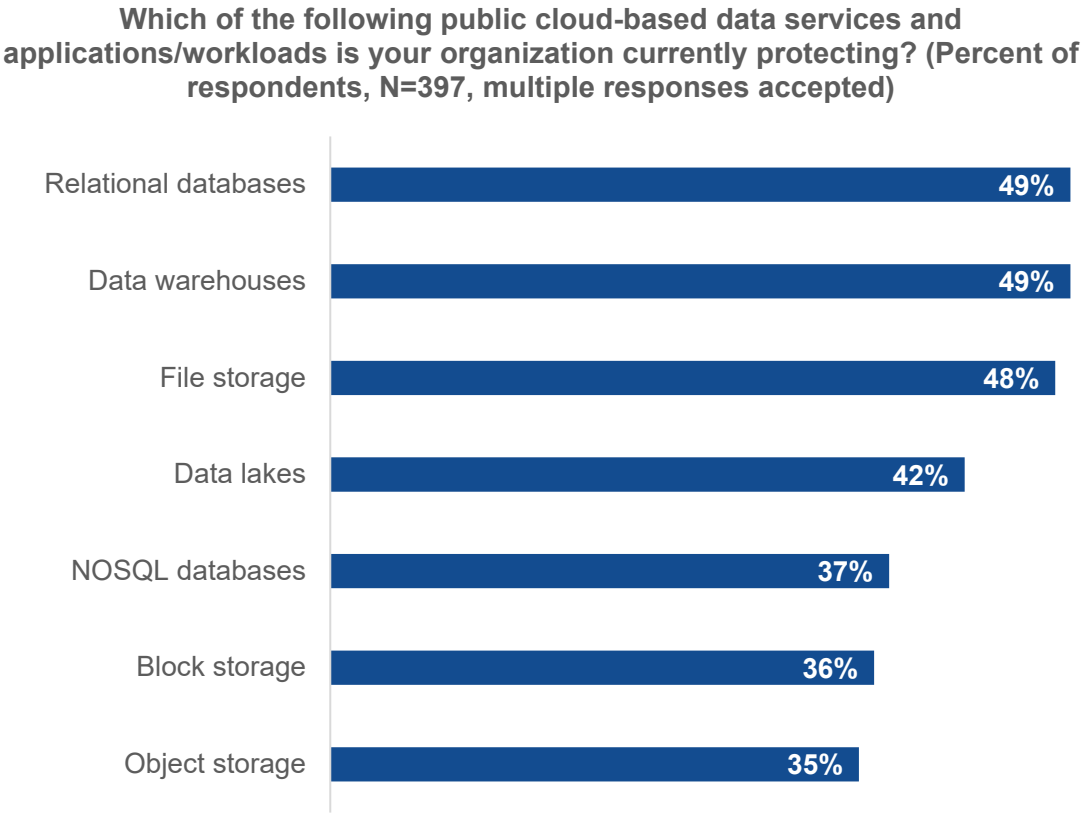


Source: Enterprise Strategy Group, a division of TechTarget, Inc.

Organizations Protect a Broad Range of Cloud-based Data Services and Applications

In terms of cloud-based data services and applications currently being protected, relational databases (49% of respondents), data warehouses (49%), file storage (48%), and data lakes (42%) are the most common (see Figure 12). The top selections here are typically mission-critical and data-intensive, reinforcing the need for highly scalable data protection infrastructure. Organizations can help accelerate initiatives designed to ensure optimal scalability, including regular assessment of their data protection infrastructure to ensure it meets the needs of their current and future data protection strategy. Another proven method for increasing scalability is data classification to ensure the right data is protected and in compliance.

Figure 12. Data and Services Being Protected Reinforce Need for Scalable Infrastructure

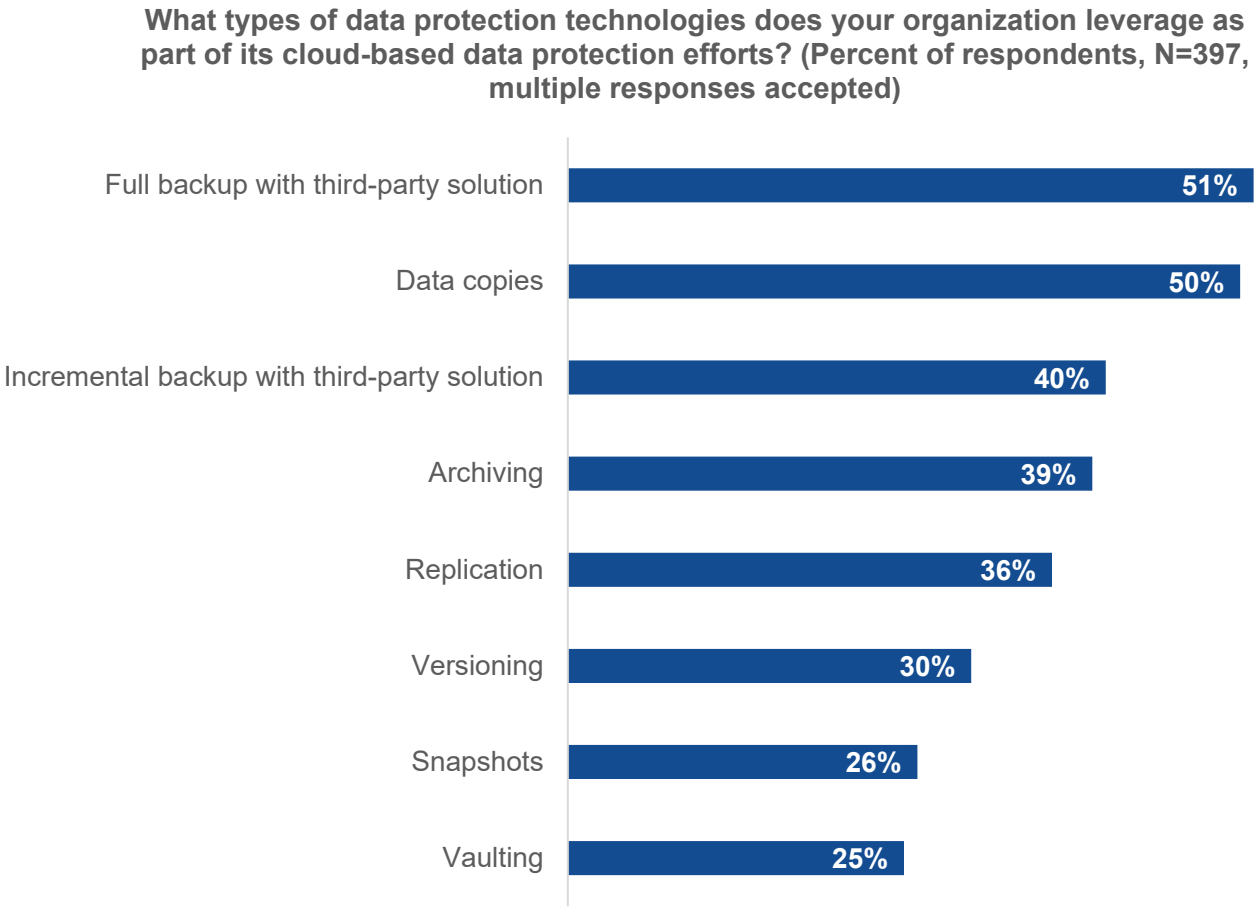


Source: Enterprise Strategy Group, a division of TechTarget, Inc.

Core Data Protection Principles and Expectations Are Carried to Cloud

Organizations are using a broad assortment of technologies in their cloud-based data protection efforts, including full backup with a third-party solution (51% of respondents), data copies (50%), incremental backup with a third-party solution (40%), and archiving (39%) (see Figure 13). These and other identified technologies are all present in various traditional data protection methodologies, providing ample evidence that organizations expect the same capabilities with cloud-based data protection services. Just as many organizations do not have sufficient resources to refactor legacy applications for cloud services, many also are hesitant to adopt cloud versions of on-premises technologies with steep learning curves.

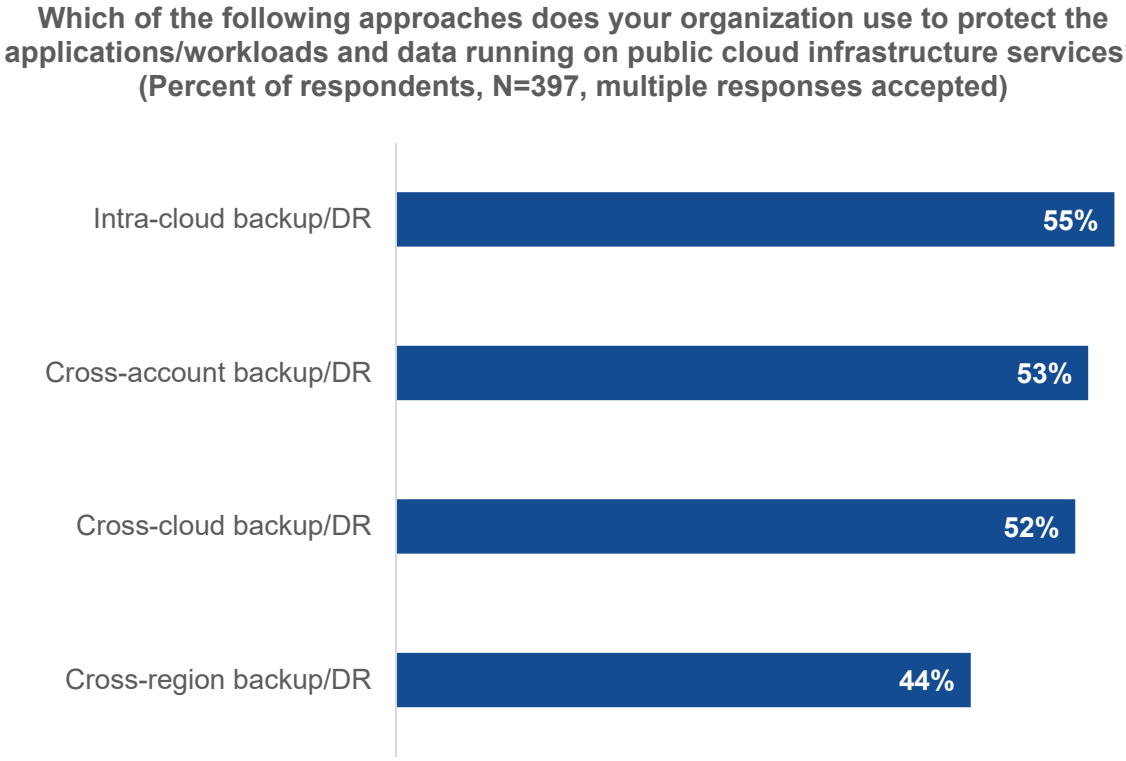
Figure 13. Core Data Protection Capabilities Heavily in Play on Cloud Services



Source: Enterprise Strategy Group, a division of TechTarget, Inc.

Largely due to the extensive diversity of applications, workloads, and data in the cloud, there is no dominant approach to data protection on public cloud infrastructure services. More than half (55%) of organizations use intra-cloud backup/DR, which leverages APIs, third-party solutions, or in-cloud snapshots on the cloud service where the production applications or data reside (see Figure 14). More than half (53%) also use cross-account backup/DR, which leverages one or more accounts within a single hyperscaler to securely copy backups. With this approach, if the original backup is inadvertently deleted, the backup can be copied from its destination account to its source. Again, 52% of organizations use cross-cloud backup/DR, which leverages a cloud-based data protection service to protect applications and data residing on one vendor's cloud service by using a second vendor's cloud service. Finally, a slightly smaller contingent of organizations (44%) reported using cross-region backup/DR, which leverages a cloud-based data protection service within one hyperscaler that supports failing over or moving data to more than one region.

Figure 14. Extensive Cloud Diversity Drives Multiple Approaches to Data Protection

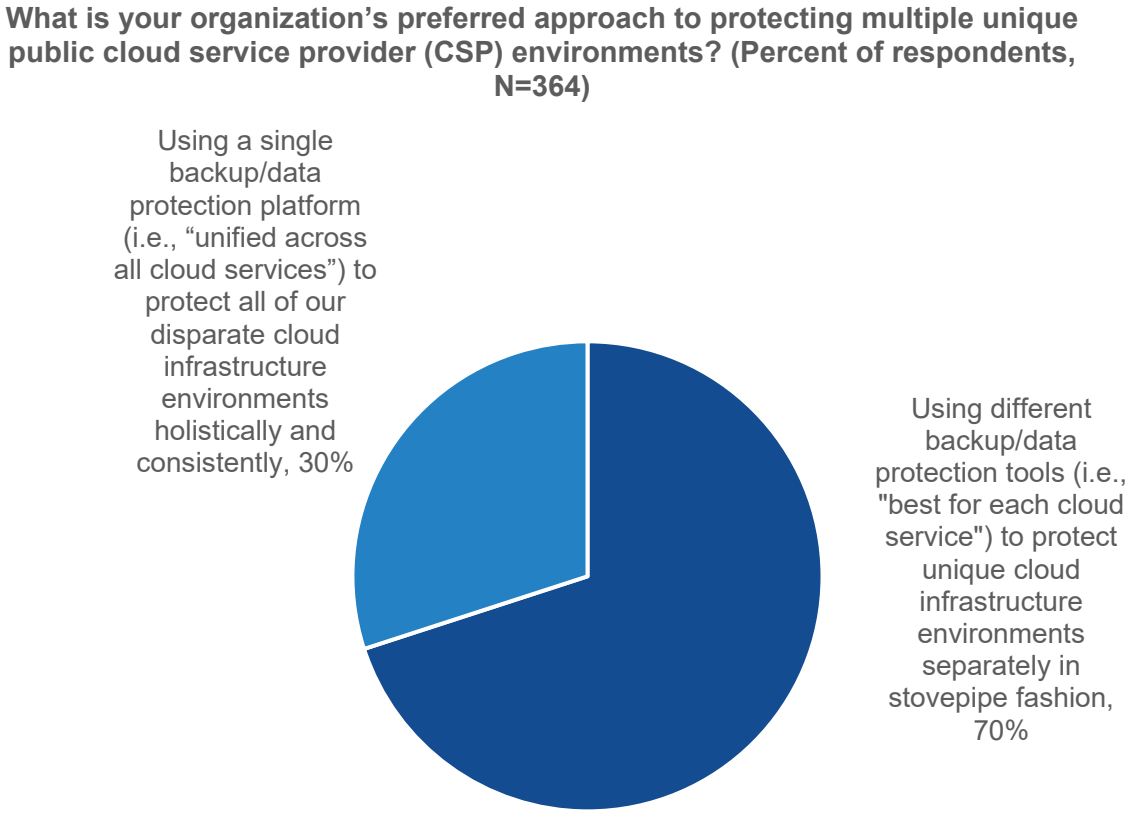


Source: Enterprise Strategy Group, a division of TechTarget, Inc.

As multi-cloud strategies become standard across the market, organizations increasingly face the challenge of protecting multiple unique public cloud service provider environments. The approach most often used (by 70% of organizations) is utilizing specialized backup and data protection tools to protect unique cloud infrastructure environments in stovepipe fashion, while 30% use a single, unified platform to protect all environments holistically and consistently (see Figure 15). While stovepipe methods allow organizations to protect their environments with a best-of-breed approach that theoretically delivers the best performance and reliability for each of their cloud infrastructure deployments, they also can lead to tool isolation that increases complexity and resource requirements. Each tool must be managed independently, which weakens an overall data protection strategy, as it increases the potential for a tool not being updated or patched.

Holistic, single-platform approaches can eliminate that isolation and provide a single source of truth for data protection processes, but availability of effective holistic platforms is somewhat limited, forcing many organizations to continue with the stovepipe approach. These firms should continue to push vendors to broaden capabilities in existing platforms to better support multi-cloud deployments. Even organizations that may never find a perfect single platform for their unique data protection requirements will still seek opportunities to consolidate at least some of those requirements within a single platform while using specialized tools for the rest.

Figure 15. Stovepipe Method Most Common for Protecting Multi-cloud Environments

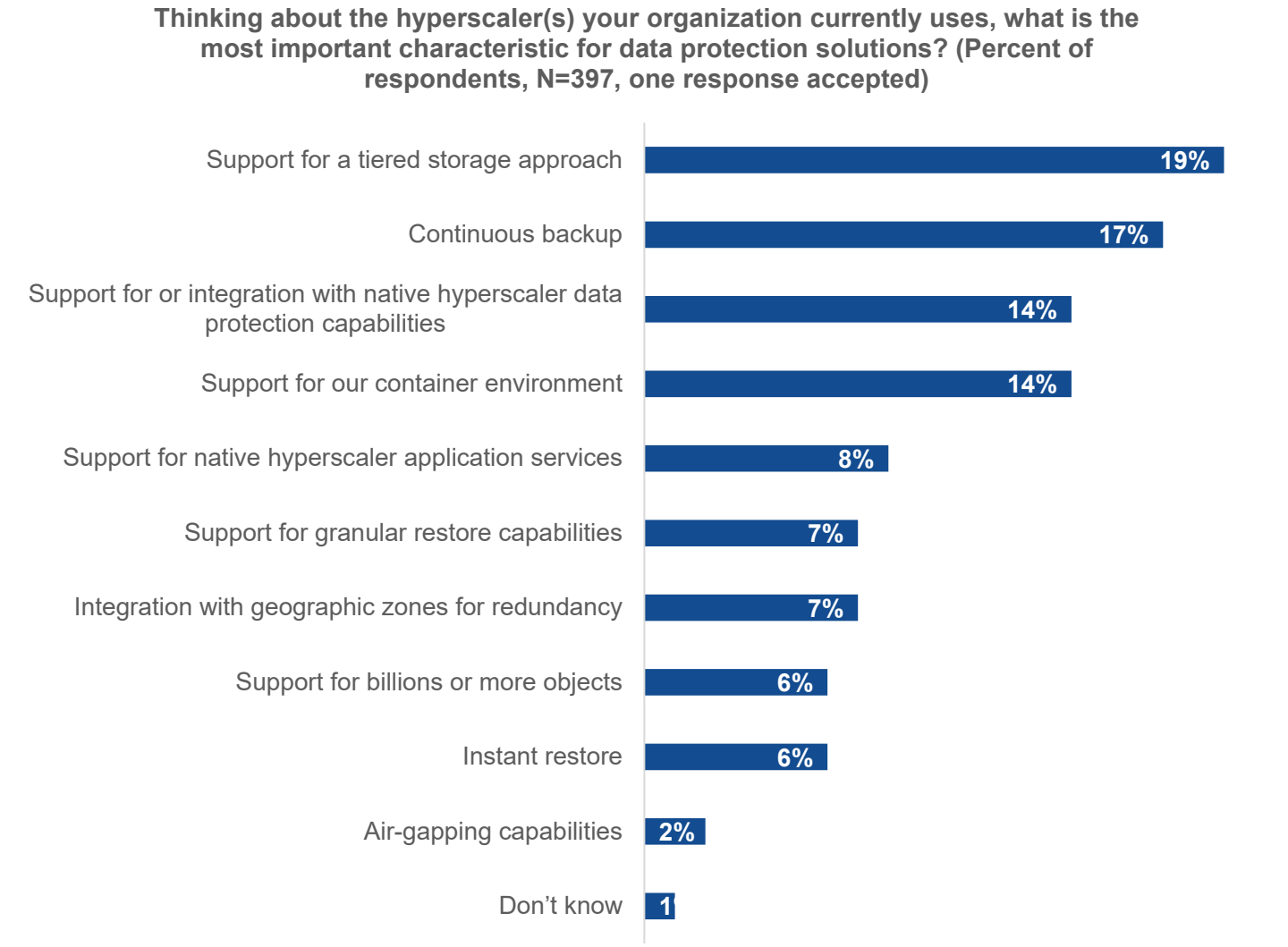


Source: Enterprise Strategy Group, a division of TechTarget, Inc.

Organizations Lean on Cloud Data Protection to Support Stringent SLAs

Organizations are becoming more adept at managing their overall data protection environment, which includes data and applications on public cloud services. For example, when asked to identify the most important characteristic of data protection solutions used to protect their hyperscaler environments, 19% of organizations said support for a tiered storage approach, an approach already well established in on-premises environments (see Figure 16). They also are likely to value stringent SLAs with continuous backup (17%) and/or support for (or integration with) native hyperscaler data protection capabilities (14%). While tiered storage is crucial for optimizing data protection environments, it is evident that many organizations place a strong focus on capabilities that support stringent data protection SLAs.

Figure 16. Organizations Prioritize Data Protection Capabilities That Ensure SLA Compliance



Source: Enterprise Strategy Group, a division of TechTarget, Inc.

Conclusion

As cloud data protection services evolve and provide improved support for data, applications, and workloads deployed on public cloud services, organizations face the considerable challenge of integrating these technologies and processes into their existing data protection strategies. If general data strategies were universally effective and trusted, this process might roll seamlessly into existing architectures. But those strategies are too often executed and then left relatively untouched for years, leading to the rise of ad hoc adjustments that spawn complexity and security risks. Nonetheless, organizations are moving forward with cloud data protection services as they continue to deploy data-heavy, critical applications on public cloud infrastructure.

To ensure their cloud data protection strategies continue to improve and evolve to support stringent SLA requirements, organizations should:

- **Increase the frequency of assessments.** Too many organizations wait years before reassessing or rearchitecting their data protection strategies, which is far too long to support the extraordinarily dynamic nature of public cloud. Annual—or more frequent—assessments of general and cloud-specific strategies are imperative to avoid falling into the resource-consuming trap of changing data protection processes and policies on an ad hoc basis, particularly in the new norm of multi-cloud deployments. While more frequent assessments can be perceived as an unnecessary burden, they are a far better alternative to the unpredictable, unsafe practice of making changes on the fly.
- **Lean on best practices developed in on-premises deployments.** Because most or all of the services and applications protected in cloud environments are mission-critical and data-intensive, it is crucial not to overestimate any security benefits of cloud services, nor underestimate the potential for data loss or cybersecurity events. Organizations are constantly undergoing data recovery operations in the public cloud, and those efforts can be less successful than data recovery on premises. In turn, all cloud data protection strategies—large or small—should stem from established best practices from on-premises data protection to ensure that all required capabilities are covered across backup, archiving, replication, versioning, and others.
- **Seek platforms that enable consolidation of data protection processes.** The reality of modern, intricate, and invariably unique IT ecosystems nearly eliminates the likelihood of finding a data protection platform that meets every need in any given environment. With that unfortunate truth in hand, organizations are advised to avoid using time and resources seeking all-inclusive platforms and instead investigate platforms that enable consolidation of at least some cloud data protection requirements. Vendors in this market recognize the massive opportunity tied to holistic cloud data protection, so while a platform might not deliver all capabilities now, the vendor likely will integrate more functions moving forward. While this approach can initially disrupt teams accustomed to using a best-of-breed approach, the improvements to consistency, expended resources, and security will be worth the effort.
- **Keep all employees in the loop.** Data protection is a far larger, more menacing beast when cloud enters the picture. Choosing cloud data protection services that meet requirements and support holistic management is only part of the equation, as even the most well-regarded services will be completely ineffective in the absence of employee training and best practices. All employees should be regularly educated on the unique dangers of deploying and using data and applications on public cloud services. Organizations should apply a regular, annual (at a minimum) cadence to employee training that parallels their systematic data protection strategy assessments.

Research Methodology

To gather data for this report, TechTarget's Enterprise Strategy Group conducted a comprehensive online survey of IT professionals from North America's (United States and Canada) private- and public-sector organizations between March 9, 2023, and March 15, 2023. To qualify for this survey, respondents were required to be familiar with and/or responsible for data protection technology decisions for their organization, specifically around those data protection and production technologies that may leverage cloud services as part of the solution. All respondents were incentivized to complete the survey through cash awards and/or cash equivalents.

After filtering out unqualified respondents, removing duplicate responses, and screening the remaining completed responses (on several criteria) for data integrity, we were left with a final sample of 397 professionals.

Please see the Respondent Demographics section of this report for more information on these respondents.

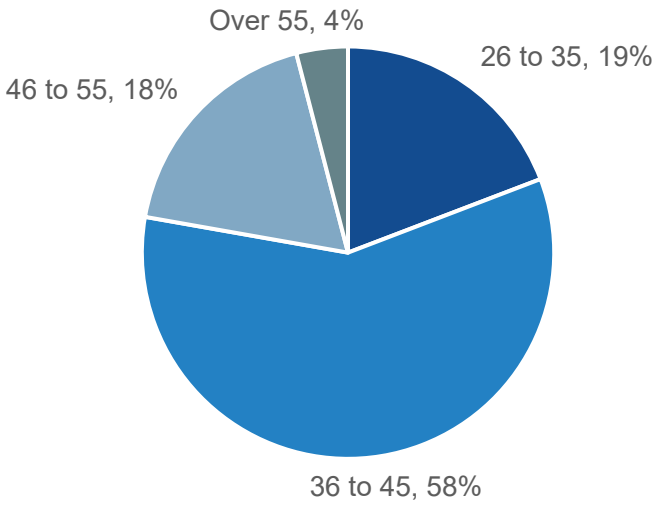
Note: Totals in figures and tables throughout this report may not add up to 100% due to rounding.

Respondent Demographics

The data presented in this report is based on a survey of 397 qualified respondents. Figure 17 through Figure 21 detail the demographics of the respondent base at an individual and organizational level.

Figure 17. Respondents by Age Group

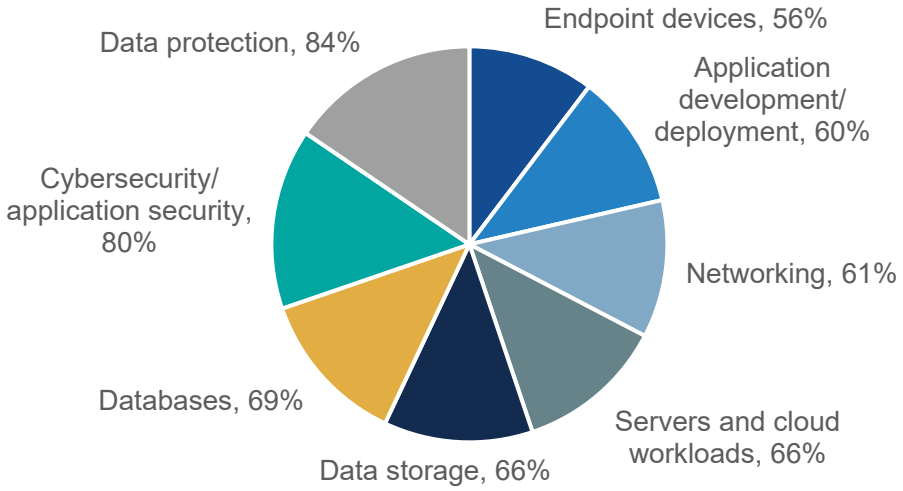
Please select your age group. (Percent of respondents, N=397)



Source: Enterprise Strategy Group, a division of TechTarget, Inc.

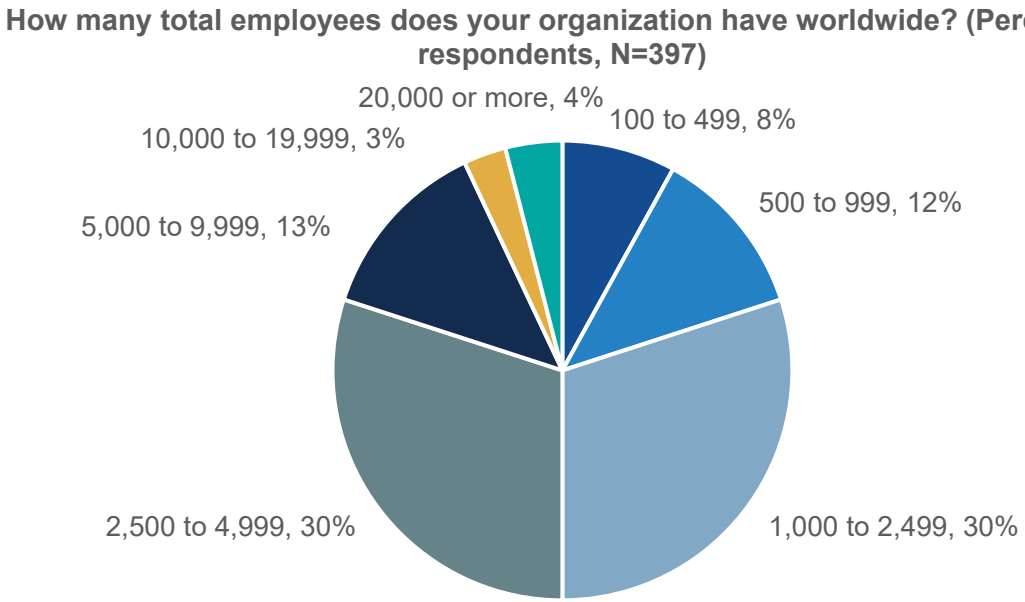
Figure 18. Respondents by Day-to-day Responsibility

In which of the following areas of IT do you have significant day-to-day involvement and/or responsibility? (Percent of respondents, N=397, multiple responses accepted)



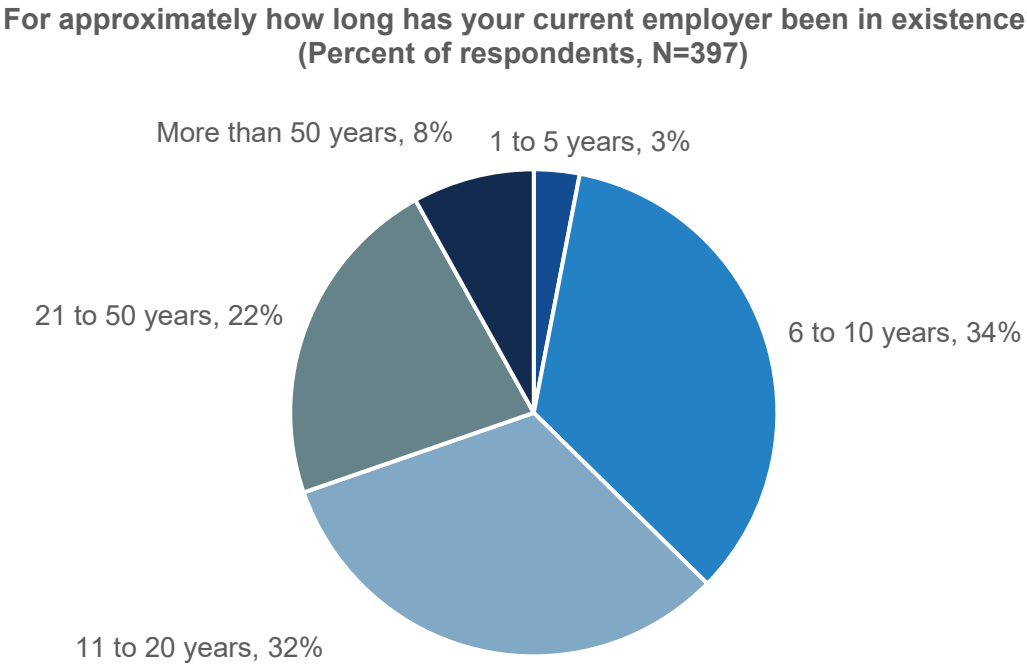
Source: Enterprise Strategy Group, a division of TechTarget, Inc.

Figure 19. Respondents by Number of Employees



Source: Enterprise Strategy Group, a division of TechTarget, Inc.

Figure 20. Respondents by Age of Organization

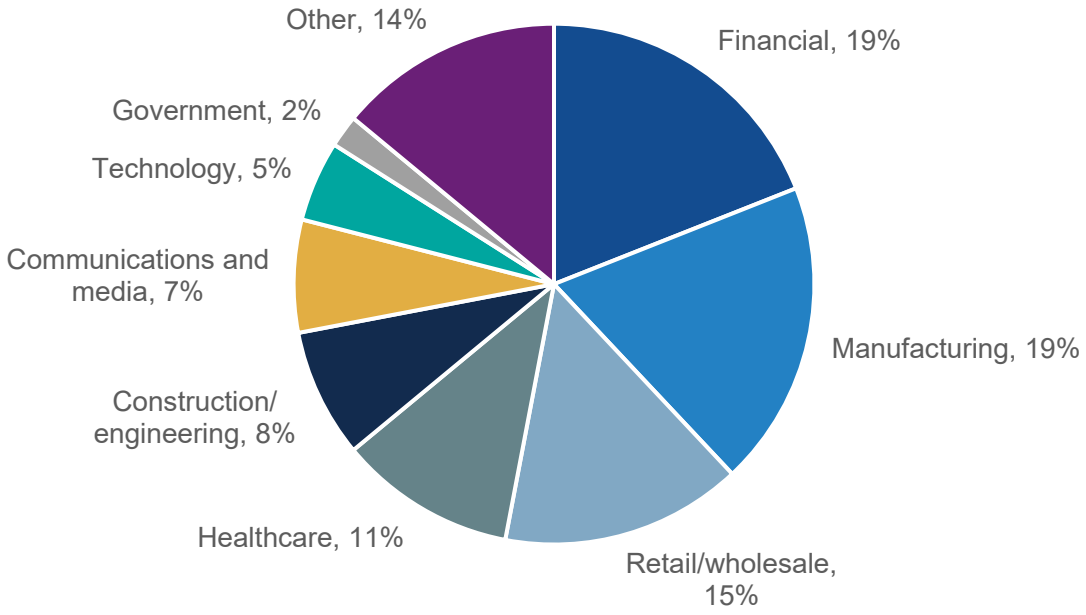


Source: Enterprise Strategy Group, a division of TechTarget, Inc.

Respondents were asked to identify their organization's primary industry. ESG received completed, qualified responses from individuals in 22 distinct vertical industries, plus an "Other" category. Respondents were then grouped into the broader categories shown in Figure 21.

Figure 21. Respondents by Industry

What is your organization's primary industry? (Percent of respondents, N=397)



Source: Enterprise Strategy Group, a division of TechTarget, Inc.