IDC TECHNOLOGY SPOTLIGHT

Sponsored by: Dell Technologies, Intel, and VMware

IT organizations must support workloads in the private cloud, public cloud, and hybrid cloud. Far from being static, these workloads are highly dynamic as they evolve to meet business objectives. Data protection must be equally dynamic to ensure that data is continuously available to key decision makers in data-driven organizations.

Deploying Flexible Data Protection to Support Cloud Workload Placement

July 2020

Written by: Eric Burgener, Research Vice President, and Phil Goodwin, Research Director

Introduction

Digital transformation (DX) projects are designed to help organizations better utilize data for competitive advantage. This advantage can be realized in two forms. First, it is axiomatic to say that organizations with greater data availability will have a relative advantage over organizations that suffer data outages. Companies that cannot process customer transactions may permanently lose both the revenue and the customer. Second, organizations that can leverage data analytics into better customer insights, identify cost savings, or discover superior merchandizing will find market opportunities where others do not.

IDC research over recent years shows that 60% of organizations either have embarked on or have completed a DX project. Based on research conducted by IDC for this paper, sponsored by Dell Technologies, Intel, and VMware, 91% of respondents surveyed consider infrastructure modernization either "very" or "extremely" important to DX success. Furthermore, 70.1% of respondents plan to conduct and deploy a data protection refresh as part of their DX initiative, a rate slightly higher than server refresh (67%) and storage refresh (68.2%), although these similar rates indicate an "all of the above" need for most organizations. The goal of data protection refreshes is to protect the business' ability to operate and, when needed, to recover data faster and more completely — specifically, to achieve service-level agreements (SLAs), and improve metrics such as recovery point objective (RPO) and recovery time objective (RTO), lowering them to as near to zero as possible. In plain language, a zero RPO and a zero RTO would deliver zero downtime with zero data loss. While this may not be an entirely realistic goal today, technology is continually getting closer and closer to reaching it.

The current era of cloud computing offers both opportunities and challenges for data protection transformation. Our research shows that 95.7% of organizations have a mandate to migrate workloads to cloud-based environments.

AT A GLANCE

KEY STATS

According to IDC's 2020 *Modernized Infrastructure Survey*:

- » 70.1% of respondents plan a data protection refresh as part of their digital transformation initiative.
- » 76.4% of respondents cited security (and therefore data security) as their top driver for workload placement.

WHAT'S IMPORTANT

Organizations that have completed a data protection refresh showed substantial improvement in almost every key SLA.

KEY TAKEAWAYS

Data availability and cyber-recovery, both key parts of data protection refresh efforts, are foundational to successful digital transformation.

Cloud deployments offer the agility of on-demand resources and reduced manual labor in the datacenter but can also carry data management and protection challenges in cloud environments where IT has less direct control. Other IDC research, separate from this particular effort, shows that data is almost evenly divided between on-premises and off-premises repositories, with the trend toward off-premises such as cloud. We believe that multicloud management is inevitable for nearly every organization as applications are deployed on different clouds and consumed by IT organizations wherever they are hosted. This scattering of applications can lead to gaps in data management and protection. For example, IT managers may erroneously assume that data in the cloud is protected. This is not always the case, and even when data protection is provided in the cloud, it is usually very basic (i.e., 24-hour backup [24-hour RPO] with 30-day retention) and thus rarely in conformance with corporate retention and governance standards, including recovery. As a best practice, IDC recommends that organizations review all cloud deployments, whether their own or SaaS applications, to ensure conformance with governance and protection standards. In cases where cloud policies are deficient, organizations should deploy their own data protection and management resources, either on premises or in the cloud, to capture and govern data appropriately.

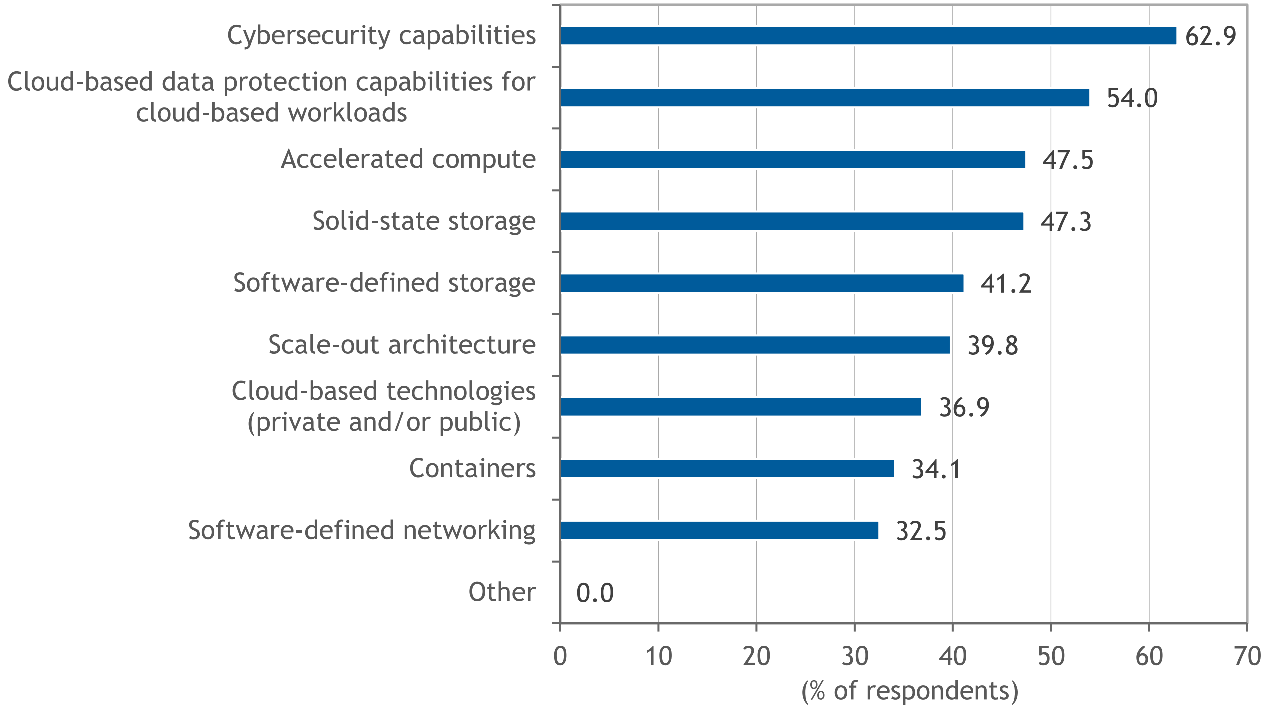
Risks and Benefits

Although cloud computing is the current focus of organizations, large-scale enterprises (and even many smaller enterprises) must deal with a necessary mix of workload deployments, including on premises, cloud native, and SaaS in the cloud, plus edge workloads such as remote office/branch office and Internet of Things (IoT) and endpoint devices. In some cases, data that is derived at the edge through analytics must be replicated (i.e., from edge cloud to public cloud or private cloud), while in other cases, data is stored where it is (at the edge) or moved to an archive. These diverse workload deployment models result in data being scattered geographically across repositories and data types. The possible permutations of data movement and protection are numerous, and such fragmentation can result in siloed and isolated data. These silos may cause organizations to deploy data management and protection point products that are unique to a specific silo. The result is duplicate products, redundant training, and inconsistent data policies across the organization. Silos can also lead to significant security and governance risks. To negate these risks, organizations need an enterprisewide perspective on data management policies and data protection solutions that deliver policy consistency and have the agility to meet widely different scenarios and evolving requirements.

Our research bears out this concern among IT organizations (see Figure 1). When we asked respondents to tell us which technologies were most desirable for modernizing infrastructure, the top 2 responses were cybersecurity capabilities (62.9%) and cloud-based data protection capabilities for cloud-based workloads (54%).

FIGURE 1: ***Most Desired Technologies for Infrastructure Modernization***

Q What key technologies do you want to leverage for your infrastructure modernization initiative?



n = 900

Source: IDC's Modernized Infrastructure Survey, 2020

Having workload agility benefits the business greatly but requires organizations to determine the optimum placement for a given workload. From the survey, we learned 76.4% of respondents believe that security is the top priority for determining workload placement, followed by 70% citing performance and 62.4% selecting ease of management. Applications with higher security and governance requirements are more likely to be hosted on premises where IT organizations have more direct control of data protection and management policies. In fact, security was again the top driver for 60% of respondents when repatriating applications from the public cloud to on premises.

According to our survey, the need for data protection refreshes extends beyond workload placement and agility. We learned that 56.2% of respondents see data growth — which has led to unacceptable backup times and missed backup/recovery windows — as the biggest driver for data protection refresh efforts. The second driver of data protection refresh was the need to improve service levels to meet shorter RPOs and RTOs, with 54.4% of respondents reporting it as a key technology. As a further indicator for the need for agility, the ability to support new workload deployments — specifically containers — was the driver of data protection refresh for 50.7% of respondents.

From organizations that had completed data protection refresh projects, we learned that the benefits were significant. Highlights from these findings include:

- 80.5% of respondents reduced unplanned downtime by 25% or more.
  - 11.7% of respondents reported a reduction of more than 75%.
- 84.6% of respondents of respondents reported an RPO improvement of at least 25%.
  - 14.7% of respondents reported an RPO improvement of more than 75%.
- 84.8% of respondents reported an RTO improvement of at least 25%.
  - 15.2% of respondents reported an RTO improvement of more than 75%.
- 82.9% of respondents reported backup and recovery job failures were reduced by 25% or more.
  - 18.3% of respondents reported a reduction of more than 75%.
- 84.5% of respondents reported backup personnel efficiency improvements of at least 25%.
  - 13.3% of respondents reported an efficiency improvement of 75% or more.

Trends

IDC forecasts that as many applications will be deployed in the next five years as have been deployed in the previous 40 years. The majority will be cloud-native or SaaS applications but will also include IoT applications at the edge. We believe that to properly protect, manage, and govern the data from so many independent applications in the IT portfolio, organizations will implement data management platforms. These platforms not only will enable data protection (i.e., backup/recovery, snapshots, replication, and cloud tiering) but also will incorporate cyber-recovery and governance functionality. They will provide visibility into private, public, and edge clouds across all repositories, instances, and applications so that protection, retention, and governance policies can be broadly and consistently applied and thus break down the silo barriers.

In many cases, these new applications will also be container based. Given that Kubernetes has become the de facto container orchestration platform for many organizations, any data management platforms must support Kubernetes workloads. This requirement is clear from the results of our survey, in which 86.1% of respondents indicated that data protection infrastructure for containers is either "very" or "extremely" important to them. Container support goes beyond data protection because containers themselves are often deployed in a persistent manner, especially for enterprise workloads. Thus container support must include recovery of whole systems where the data volumes are persistent. Because container-based workloads often require data protection to be built into the application, DevOps teams need to be included in data protection strategy development where they have not been previously involved. Moreover, containers require event-based mechanisms (i.e., triggers) rather than the time-based methods of traditional backup systems to take the appropriate protective action.

To successfully address cyberthreats, organizations need to adopt a "cyber-recovery first" mentality and approach to architecting their data protection environment. We believe that cyberattacks are the leading risk to data loss and data breaches, more than infrastructure failure, natural disasters, or even internal threats. To effectively counter these threats, organizations should first build in an "air gap" between the production data and the backup data. This air gap breaks any physical connection between the backup system and the production system so that bad actors cannot corrupt both. IDC's best practice also prescribes that the data and control planes be separate to make it more difficult or impossible for both to be compromised at the same time. Second, organizations should keep immutable copies of backup data to provide a fail-safe recovery option if other defenses are defeated. This immutability goes beyond retention locks but dictates the need for an immutable data vault that is a last known safe copy, free of any malware. To ensure this is the case, organizations must use machine learning/artificial intelligence (ML/AI) methods to scan not only the file metadata but also the document metadata and content.

Many organizations are turning to backup data sets as a source to enhance digital information, which in turn can be used for other secondary purposes. No longer are backup data sets being stored idly in case something goes wrong. In fact, these data sets are rich in information and enablers of DevOps. Key secondary use cases include data analytics, test/dev, and staging. These activities significantly contribute to the value that organizations derive from their data.

Considering Dell Technologies

The Dell EMC data protection portfolio includes data protection and backup appliances (referred to as purpose-built backup appliances [PBBAs]) and data protection and backup software that allows organizations to protect and recover their data across private, public, and edge cloud-based infrastructure as well as address the needs of container backup. Backup appliance offerings include the PowerProtect DD Series Appliances, Integrated Data Protection Appliances (IDPAs), and PowerProtect DD Virtual Edition (DDVE) for cloud-based virtual appliances. Software solutions include Data Protection Suite (NetWorker for backup/recovery, Avamar for edge-to-core or cloud data copy) and PowerProtect Data Manager. Dell EMC PowerProtect Cyber Recovery delivers an air-gapped solution that complements the base immutability offering through retention locks. In addition, running analytics on the data in the vault is an important component to enable a rapid recovery after an attack. Analytics help determine whether a data set is valid and usable for recovery or has been improperly altered or corrupted, making it "suspicious" and potentially unusable. CyberSense analytics assist with assured data recovery because CyberSense reads through the backup set so there is no need to restore data just to determine if it is clean, thereby avoiding the risk of opening harmful data in the vault. In addition, CyberSense can evaluate the full contents of the file, not just its metadata, to deliver advanced analytics.

These solutions, available as integrated appliances, software only, hardware only, or comprehensive bundles, have the integrated features and technologies organizations need to satisfy the top drivers of data protection refresh uncovered in IDC's recent primary research in this area. Easy scalability into the petabyte range, assisted by inline compression and deduplication, cloud tiering capabilities, integrated solid state storage options that speed both backup ingest and data recovery, and the flexibility to accommodate various deployment models (bare metal, virtual machines, containers), enables these systems to deliver on the promise of modernized infrastructure to meet evolving DX requirements.

Dell EMC data protection solutions are fully integrated with vSphere to enable VM administrators to manage data protection directly from the native vSphere UI. With Dell EMC advanced VMware integration, VMware administrators are empowered to more efficiently control their own data protection, resulting in faster backups and restores for virtualized mission-critical applications. Additionally, PowerProtect Data Manager delivers support for vSphere 7 and Tanzu, paramount for protecting container-based applications and Kubernetes workloads.

Challenges

Dell Technologies arguably has one of the broadest and most agile data protection portfolios in the industry. According to IDC data, the company was the market share leader (by revenue) for data replication and protection software and purpose-built backup appliances in 2019. Dell's data protection platform can address the majority of enterprise data protection and cyber-recovery requirements, but such a broad range of products has its inherent challenges. First, some of Dell's products were introduced over a decade ago, well before the cloud became ubiquitous and containers were a serious application development choice. Thus Dell must continually update not just product features but also core architectural designs to meet evolving requirements never contemplated by the original designers. Similarly, Dell must constantly counter new entrants to the market that have the advantages of a single-purpose product focus without concern for prior architectures. Dell must deliver continued innovations in the area of Kubernetes, cyber-recovery, and significant VMware integration. Thus the key challenge for Dell is to remain competitive and innovative across numerous products and against dozens of competitors.

Conclusion

Digital transformation is seen as an imperative for many organizations to gain competitive advantage in the marketplace. These data-driven organizations must have the highest levels of data availability with rapid and certain cyber-recovery via systems that deliver comprehensive data management policies across private, public, and edge clouds. Data protection platforms must have the built-in flexibility to support various workload deployment models and to support applications as they evolve through various implementations. This includes air gaps and immutable data copies to ensure data survival in the face of cyberattacks. To meet these requirements, most organizations will need to make data protection transformation foundational to their other efforts. The results can be significant: better service-level agreements, more data availability, and increased personnel efficiency to gain time for other projects.

Data-driven organizations must have the highest levels of data availability with rapid and certain cyber-recovery via systems that deliver comprehensive data management policies across private, public, and edge clouds.

About the Analysts

	Eric Burgener, *Research Vice President, Infrastructure Systems, Platforms and Technologies* Eric Burgener is Research Vice President within IDC's Enterprise Infrastructure practice. Mr. Burgener's core research coverage includes storage systems, software and solutions, quarterly trackers, and end-user research as well as advisory services and consulting programs. Based on his background covering enterprise storage, Mr. Burgener's research includes a particular emphasis on solid state technologies in enterprise storage systems as well as software-defined infrastructure.
	Phil Goodwin, *Research Director, Infrastructure Systems, Platforms, and Technologies* Phil Goodwin is a Research Director within IDC's Enterprise Infrastructure practice, covering research on data management. Mr. Goodwin provides detailed insight and analysis on evolving industry trends, vendor performance, and the impact of new technology adoption.

IDC Custom Solutions

The content in this paper was adapted from existing IDC research published on www.idc.com.

IDC Research, Inc.
5 Speen Street
Framingham, MA 01701, USA
T 508.872.8200
F 508.935.4015
Twitter @IDC
idc-insights-community.com
www.idc.com

This publication was produced by IDC Custom Solutions. The opinion, analysis, and research results presented herein are drawn from more detailed research and analysis independently conducted and published by IDC, unless specific vendor sponsorship is noted. IDC Custom Solutions makes IDC content available in a wide range of formats for distribution by various companies. A license to distribute IDC content does not imply endorsement of or opinion about the licensee.

External Publication of IDC Information and Data — Any IDC information that is to be used in advertising, press releases, or promotional materials requires prior written approval from the appropriate IDC Vice President or Country Manager. A draft of the proposed document should accompany any such request. IDC reserves the right to deny approval of external usage for any reason.