IDC ANALYZE THE FUTURE

IDC TECHNOLOGY SPOTLIGHT

Sponsored by: Zerto

Data availability and application availability can no longer be considered separately. Modern workload migration technology allows applications to move seamlessly from on premises to multicloud environments. Data must move in tandem to ensure the greatest agility and meet availability service-level agreements (SLAs).

Continuous Data Protection: A New Era of Backup and Recovery

October 2020

Written by: Phil Goodwin, Research Director, Infrastructure Systems, Platforms, and Technologies Group

Introduction

IT focus is no longer on backup/recovery and disaster recovery (DR) as separate tasks but on the availability of data and application. Data availability and application availability must be considered together to deliver the business services that both businesses and their customers expect. Companies cannot risk downtime or data loss that could impact customer relations and loyalty.

IDC research shows that 90% of organizations utilize the cloud for data protection and that 70% of CIOs have a "cloud first" strategy for application deployment. For most applications, this invokes a hybrid cloud strategy, which keeps at least one backup copy on-premises with a second copy replicated, or tiered, to the cloud. The secondary copy is stored for data survival — similar to vaulting tapes — or other value-added purposes such as DR staging, test/dev, and analytics.

AT A GLANCE

KEY STATS

- Sub-minute RPO is quickly becoming a best practice.
- Best practice RTO is 15 minutes.

WHAT'S IMPORTANT

Organizations must apply data recovery and workload migration simultaneously to drive the highest levels of business availability. Backup/recovery, disaster recovery, and workload migration cannot be separate tasks if this level of business availability is to be achieved.

IDC forecasts that as many applications will be deployed in the next five years as have been in the previous 40 years. We expect that more than 80% of these applications will be cloud based at the edge, either SaaS or cloud native. These applications will be hosted on more than one public cloud, making multicloud data management and protection a requirement for almost every organization. Organizations are deploying new applications in the cloud to achieve greater agility, enabled by on-demand resources and cloud scaling. In protecting those applications with backup and DR, cloud is similarly leveraged to lower the cost associated with these tasks. Cloud is an ideal platform for rapid recovery of both applications and data due to on-demand infrastructure.

In response to the need for ever greater application availability with less data loss, a new generation of continuous data protection (CDP) technology is emerging to significantly reduce recovery point objectives (RPOs). Prior generations of CDP generated too much processing and storage overhead to gain mainstream use. Other data protection techniques aimed at minimizing RPOs — primarily snapshots — may be used to deliver a 15-minute RPO but can generate significant storage overhead (i.e., 2% per snapshot) and induce a noticeable delay in system response during each snapshot. Although snapshots can theoretically be taken every five minutes, few organizations do so because of the overhead; typically, snapshots are taken only every 4 hours, 8 hours, or nightly. Modern CDP solutions, in contrast, can offer sub-minute RPOs with very little overhead or impact to production systems.

Ransomware and malware have emerged as the primary threats to data and application availability. In response, ransomware recovery has become a primary use case for CDP. Using CDP, data can be recovered to the point in time just prior to the attack to hold data loss to the absolute minimum for application consistency. In most cases, CDP makes sub-minute RPO practical.

CDP is equally applicable to on-premises and cloud recoveries. It converges data protection and disaster recovery into a single, seamless management framework that automates recovery regardless of cause, source, or target. Other use cases are more commonplace data loss occurrences, such as database corruption and accidental file deletions. Being able to recover a database at a very granular level and a recent point of consistency speeds database recovery and application restart, while accidentally deleted files can be recovered with minimal data loss. CDP can contribute to better business results, such as improved customer service, better employee productivity, and reduced data protection costs.

Benefits

CDP can change the way IT organizations protect data because it facilitates the lowest possible RPO without going to expensive synchronous solutions (i.e., highly available systems). Whereas most recovery solutions offer RPOs measured in minutes at best — and more likely hours — CDP solutions can offer sub-minute or near-zero RPOs because the systems utilize a journaling feature that captures each write to disk. Journaling ensures that every data change is captured and recoverable down to the individual write. Such granularity allows data to be restored to a point just prior to a disruption, which is especially useful for recovering from malware and ransomware.

CDP solutions also offer the potential for consolidation of backup/recovery and DR solutions to reduce redundant tools and the need for separate backup and DR infrastructure, not to mention that both tasks can be accomplished by one team rather than two teams. By not having to implement separate backup and DR infrastructure, organizations can more easily leverage this secondary infrastructure for the previously mentioned value-added functions. Moreover, data and applications can be restored directly into production, thereby also reducing recovery time objective (RTOs). Faster recovery time is obviously important from a general business perspective, but it is financially important as well. IDC research shows that the average downtime cost is $10,000 per hour per application workload. Thus, any reduction in downtime provides direct return on investment for any CDP solution.

When combined with cloud tiering, CDP allows data to be moved to the cloud automatically and by policy for any use case. Data can be moved geographically, either between different datacenters or from one cloud zone to a secondary cloud zone. Based on these policies, data movement can be configured to comply with data sovereignty requirements. Moreover, data can be stored for long-term retention directly in cloud storage locations such as AWS S3 buckets or Azure blobs to optimize cost.

Organizations now commonly deal with multicloud environments. Because of the way applications are deployed, most organizations will find themselves with a combination of the following: applications on-premises and SaaS applications on many of the major hyperscalers (e.g., Amazon, Azure, GCP) as well as industry-specific clouds (e.g., healthcare, financial services). The permutations of possible recoveries can be complex, with omni-directional recovery possibilities to, from, and between any of these clouds. Organizations needed the greatest agility to ensure recovery and attainment of service-level agreements (SLAs). DR teams should not architect solutions for just one kind of recovery (e.g., on premises to cloud); rather, they need solutions that have the flexibility to respond to unforeseen situations that could involve cloud to cloud or even cloud to on premises; applications and data are spread across core, cloud, and edge repositories and must be able to be recovered in any other location.

Trends

Among the biggest current trends is the accelerating deployment of containers, which goes hand in hand with the deployment of cloud-native applications. Container deployments are now moving from the development environment into production, where Kubernetes has become the dominant container orchestration tool. For the purposes of data protection, containers have some important differences from traditional application deployments. First, containers and their data can be either transient or persistent; backup applications must be closely integrated with Kubernetes and the application to know the difference and protect both the container and the data as needed. Second, whereas traditional data protection is time based, containers are event based. Data protection products must be capable of handling event-based triggers. Third, Kubernetes is a highly dynamic environment under almost constant evolution. The Kubernetes version of today will not be the same as even the recent past, meaning that entire systems — Kubernetes, containers, and data — must be recovered to the event needed, with time being a possible event.

IDC research also shows that 60% of organizations are striving to be "data driven." Being data driven has the ultimate goal of creating competitive advantages in the market. Data that is timely and accurate enables better, more informed decisions by business leaders. Being data driven also means using data to discover usable information such as market trends, customer needs, and other insights before the competition has time to react. Clearly, 60% of organizations cannot succeed equally; organizations that do a better job of using data are more likely to become market leaders. Data availability, completeness, and accuracy are foundational to being data driven. Organizations that can reduce downtime and data loss using tools such as CDP will be better able to set that foundation of competitive advantage.

Considering Zerto

The Zerto Platform is designed to deliver converged data protection and disaster recovery, enabled by continuous data protection, workload mobility, recovery orchestration, and analytics, in a single experience. It integrates all aspects of data protection and disaster recovery so that organizations can consolidate operations, reduce the number of products to manage, and simplify end-to-end data protection and recovery. The software-only platform is designed to work in on-premises, hybrid cloud, or multicloud environments.

The platform's continuous data protection and journal-based recovery allow very granular recoveries to provide near-zero RPOs. Combined with near-synchronous replication that sits at the hypervisor level, the platform can also deliver very rapid recovery times for both traditional and cloud-native environments. Its journal-based recovery also allows recovery from cyberattacks, such as ransomware, to the point in time just seconds prior to an attack.

Built-in real-time and historical analytics allow organizations to track and monitor application and data protection status and health. These analytics include resource planning and forecasting to predict the storage and infrastructure requirements when moving applications to the cloud or protecting applications with Zerto.

The flexibility of the Zerto licensing model enables customers to utilize Zerto for all applications, both mission critical and lower tier, in a single scalable platform. Options include:

- **Zerto Data Protection (ZDP) license:** Local continuous backup and recovery, alongside long-term retention, whether on premises and/or to the public cloud, with analytics and easy restore workflows.
- **Zerto Enterprise Cloud Edition (ECE) license:** All the functionality of the Zerto Data Protection license, plus disaster recovery for on-premises and public cloud, with full orchestration, automation, and one-to-many support for hybrid and multicloud environments

Zerto claims that its licensing and packaging can deliver a 50% reduction in total cost of ownership (TCO), largely by allowing the consolidation of tools and functionality, compared with older, noncontinuous architectures with multiple point solutions. Other factors contributing to this savings are reduced time spent on managing separate processes and reduced downtime costs as a result of shorter recovery points and faster recovery times.

Challenges

IT organizations are accustomed to and familiar with the traditional separation of backup/recovery and DR infrastructure. Zerto's architecture of combining functions into a seamless cloud data management and protection platform is different from the usual paradigm and requires IT managers to think differently about how they approach application availability. Consequently, Zerto must educate IT practitioners on the benefits of its approach.

To achieve the full TCO benefits of Zerto, organizations must be willing to replace their existing backup software infrastructure with Zerto. However, making such a change will not be an easy undertaking for all organizations. Data already in long-term retention will require an organization to retain its legacy environment until all of the related data sets expire before it can be fully transitioned to one platform.

Conclusion

The need for the highest levels of application availability will become only more urgent. Near-zero RPOs and RTOs are achievable using modern continuous data protection technology. Because applications can be migrated and recovered across physical repositories, organizations will no longer need to maintain separate backup and DR infrastructure. This is what we mean by "application availability" — rapid recovery from any cause to any location where there is no longer any distinction between data recovery and disaster recovery. To the extent that Zerto can address the challenges described in this paper, the company is well positioned to be a key supplier as this trend gains momentum in the marketplace.

The need for the highest levels of application availability will become only more urgent. Near-zero RPOs and RTOs are achievable using modern continuous data protection technology.

About the Analyst

Phil Goodwin, *Research Director, Infrastructure Systems, Platforms, and Technologies Group*

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. He is responsible for producing and delivering timely, in-depth market research with a specific focus on cloud-based and on-premises data protection, business continuity and disaster recovery, and data availability. Mr. Goodwin takes a holistic view of these markets and covers risk analysis, service-level requirements, and cost/benefit calculations in his research.

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