IDC
Analyze the Future

I D C T E C H N O L O G Y S P O T L I G H T

Taking Data Protection to the Next Level with a Unified Data Storage, Management, and Protection Platform

June 2017

Adapted from *Market Analysis Perspective: Worldwide Storage Systems, 2016*, by Phil Goodwin, et al., IDC # US40357815

Sponsored by Reduxio

Businesses increasingly demand 100% uptime with no data loss, or what IDC refers to as the "new race to zero" – 0 RPO and 0 RTO (recovery point objective and recovery time objective). While such a scenario is still unrealistic for all but the most expensive and demanding applications that can afford fully redundant systems and failover mechanisms, IT organizations are gradually driving down their RPO and RTO. Unfortunately, in most cases this requires a complex strategy of interleaved products including snapshots, mirrors (clones), backup software, and asynchronous and synchronous replication, as well as stand-alone cloud based solutions for data protection and disaster recovery. Moreover, these tools are often vendor and platform-specific, meaning that tools are both manually interleaved and redundantly overlaid. To make matters worse, and despite the deployment of all this technology, IDC research shows that nearly two-thirds of organizations are not fully confident that they could recover their data in the event of a significant disaster. This Technology Spotlight examines the evolution of data protection in a virtualized and cloud world and how Reduxio's unified data storage, management and protection platform is designed to deliver near-zero RPO and RTO as a feature of its primary storage system while significantly simplifying the data protection and disaster recovery.

Introduction: The Evolution of Data Protection

Since the beginning of the client-server era nearly 30 years ago, the emphasis for data protection has been on the re-creation of physical infrastructure, such as data blocks, files, LUNs, volumes, and the like. Regardless of methodology, a copy of a physical element is made at a specific point in time so that it can be restored to its original condition if the production version is lost or damaged. Originally, this was limited to backup/recovery (B/R) software that could make a copy of the complete system but had no better than a 24-hour RPO/RTO.

IDC research has found that the most common target for RPO is one hour, and the most common target for RTO is four hours. Clearly, B/R software alone cannot meet this service level requirement. To meet the business-driven service level requirements, IT organizations have implemented multiple products, all of which have attributes that overcome the weaknesses of related products as illustrated in Table 1. This is not intended as an exhaustive list of pros/cons, but rather keys that impact service levels.

Table 1

Common Current Data Protection Schemes with Pros & Cons

Technology	Pros	Cons
Backup/recovery software	● Protects an entire environment ● System agnostic ● Simple long-term retention	● 24 hour RPO/RTO typical ● Often only 85% effective
Mirrors/clones	● Protect entire systems ● Low RTO	● RPO usually no better than 12 hours ● 100% overhead ● System specific
Synchronous replication	● Low RPO ● Can be low RTO	● Expensive ● High overhead ● Distance-limited ● May be system specific
Asynchronous replication	● Lower RPO ● Lower overhead ● Unlimited distance	● Not as good as synchronous ● May be system specific
Snapshots	● Lower RPO ● Lower RTO	● Often system-specific ● Can be a complicated recovery ● Short-term data retention ● Potentially high cumulative overhead
Cloud replication	● Offsite and can facilitate DR ● Low storage cost	● Bandwidth may significantly impact RTO ● Sometimes high data egress costs

Source: IDC, 2017

From the table above, it's clear that no one product or technology can deliver data protection to meet the data recovery and availability needs of the business—each technology compensates for the weakness of another. IT professionals also well know that each of these products has its own user interface, reporting mechanism, and learning curve. Completing a data recovery may require significant human ingenuity to get the right data from the right point in time based on the appropriate combination of tools.

Data Protection as a Continuum

Older, traditional data protection schemes approach high availability (HA), backup/recovery, and disaster recovery as distinctly different endeavors. This is one of the reasons that disaster recovery (DR) remains underserved in many organizations. Too often, DR infrastructure is separate and distinct from the primary infrastructure, leading to costly duplication plus complex planning and testing to move multiple workloads from one set of infrastructure to another. Our research indicates that DR failovers are successful on the first try less than 25% of the time. In most cases, multiple tries are needed to get all components of the infrastructure coordinated so that the failover is successful. This

real time, trial-and-error method to recovery can increase the recovery time 3X-5X (or more), usually well beyond the stated service level requirement.

When one thinks of the implications of 0 RTO and 0 RTO, it is the very definition of high availability. Achieving this level of service previously required duplicate infrastructure, complex failover mechanisms, and applications written to take advantage of HA. However, the current trend toward self-healing devices means that near-zero RPO/RTO can be accomplished without these complex, expensive mechanisms and be available to all applications using those systems. Self-healing does not necessarily imply a single device; it may use multiple interrelated devices, yet in such a way that is transparent to the operator and all applications utilizing it. Self-healing technology may be both proprietary (unique to the device) or based on open standards such as some object storage methodologies. Neither is inherently better, but buyers should be cognizant of the difference and weigh each for their own requirements.

Storage Inefficiency

In addition to the complexity of managing and maintaining the various data protection products described above, these products have an additional detrimental impact: storage inefficiency. Having overlapping data replication products means that the same data is copied numerous times by different tools. IDC calls these cumulative copies "copy data" and estimates that IT organizations will spend $50 billion per year into the foreseeable future just for the infrastructure needed to store these copies. In most cases, this data simply lays idle, waiting to be used in the event something goes wrong with the primary copy.

In response to this nearly ubiquitous problem, IT organizations are implementing copy data management solutions. It is important to note that copy data management does not refer to the management of copy data, which would simply be orchestrating the use of those copies. Instead, copy data management (CDM) is the virtualization of data, so that a single copy may be used for multiple purposes simultaneously. Thus, CDM tools reduce the total amount of storage needed, while making copy data usable for such things as dev/test, analytics, and more.

CDM products were initially implemented as third-party utilities. More recently, storage vendors have begun adding CDM capabilities to their systems, transparently to the user. Thus, data is automatically virtualized and virtual copies (views) may be available on-demand for DBAs, developers, and analysts. Because data may now be available to a wide variety of users, security is critical. Sensitive data such as personally identifiable information (PII) and health-related data (such as HIPAA) may be included in these data sets and must not be disclosed to unauthorized users. Thus, CDM tools should be equipped with role-based masking and/or data scrubbing to assure appropriate data governance.

Considering Reduxio

Reduxio is a U.S. company with headquarters in South San Francisco. It also has other US offices plus offices in Europe and Israel. It's unified data storage, management and protection platform, designed to unify the management of primary and secondary data is centered around its HX-series of Flash arrays. In addition, it is provides seamless integration with public cloud storage for both immediate data availability as well as disaster recovery.

The HX-series arrays are controlled by the company's TimeOS storage operating system. This system has several unique features and other capabilities built in, as described below:

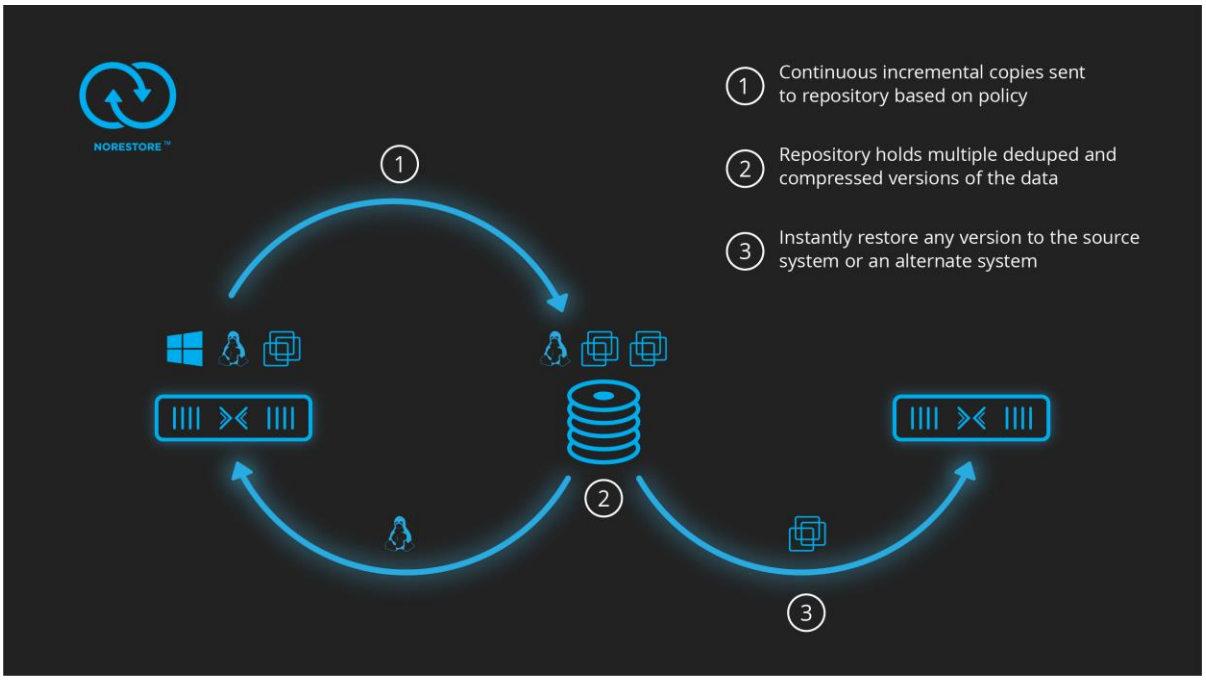
- Metadata managed separately from the data. This separation allows data and volumes to be moved independently, making data movement simpler and enhancing recoverability.

- Time-stamped data writes. Facilitates RPO as low as one second, making near-zero data loss possible using BackDating while eliminating the need for snapshots.
- NoRestore recovery. Designed to eliminate the need for backup/recovery software, NoRestore supports storing continuous increment point-in-time copies to an iSCSI or object storage device with "near 0" RTO restore. The company claims that even 100TB of data can be restored in a matter of minutes.
- Cloud as DR target. Continuous incremental replication to an S3-compatible cloud target moves data offsite for very low RPO in the event of disaster.
- NoMigrate. NoMigrate is the ability to instantly migrate volumes from third-party storage. The volumes are presented instantly on Reduxio, while data is moved in the background.
- NoDup. NoDup is the company's implementation of deduplication and compression, implemented globally across all stored data.

The Reduxio user interface uses HTML-5 that included a customizable dashboard with drag-and drop widgets. The innovative user interface allows users to easily provision and manage the storage, perform point-in-time restore, and set policies for data protection. The interface, using a vCenter plug-in, allows data store-level restores, plus a dashboard for reporting purposes. Figure 1 captures key elements of Reduxio's NoRestore architecture.

Figure 1

Reduxio NoRestore Architecture



Source: Reduxio, 2017

Referring to Table 1 and comparing it to the capabilities of the Reduxio system above, it's evident that the Reduxio system has many of the capabilities built in to construct a data protection infrastructure. It should not be overlooked that the HX-series systems are designed from ground-up to support the most demanding application workloads; they are not data protection devices; they just have data protection built in. BackDating allows recovery of data at one second granularity instantaneously

without the complexity of managing snapshots and snapshot schedules. The NoRestore functionality allows customers to set up a DR environment with very low RPO and near 0 RTO without the need for any additional software solutions. It is also worth noting that the Reduxio system would generally eliminate the need for a separate purpose-built backup appliance (PBBA), further reducing cost and complexity for many organizations.

Challenges

As an emerging vendor, Reduxio can be expected to face all the usual challenges. Moreover, it would not be reasonable to expect them to have a fully-developed system as vendors who have been in the market for 20 years. However, Reduxio also has the advantage of starting with a clean slate into which they can design state-of-the-art capabilities without legacy concerns.

As Reduxio continues to evolve and mature its product, we would like to see two elements added. The first is an "air gap" to its data replication capabilities. While the company advertises that B/R software can be eliminated - it is technically true - some users will still want to retain a backup process to insert an air gap into the backup copies. This air gap can help to assure that malware does not have the opportunity to infect the entire set of data copies and that at least most of the data could be recovered, even if the RPO is not optimal. Second, we would like to see copy data management data masking or scrubbing. Reduxio has indicated that this capability is on the roadmap.

Conclusion

IDC believes that the evolution of data protection is moving away from discrete functional endeavors and becoming integrated into the primary storage stack. In doing so, service levels, especially RPO and RTO, can be driven lower and lower. It is our forecast that backup/recovery will be eliminated as an IT task by 2025 and that seamless, built-in data protection will replace it. As organizations drive to 0 RPO/ 0 RTO, applications will have the benefit of inherent high availability and significantly reduced risk of data loss. In addition, IT staff will be freed up to perform higher-value tasks.

The Reduxio system represents an interesting move toward these highly available, self-healing and self-recovering systems. Certainly, there is more than one technological way to accomplish the desired goals and we expect more vendors to move in that direction. However, Reduxio's unique approach to the problem gives IT organizations another capability and product to consider when architecting their storage environment for maximum effectiveness.

Global Headquarters: 5 Speen Street Framingham, MA 01701 USA P.508.872.8200 F.508.935.4015 www.idc.com