WRITING PAPER

Taking Data Protection to the Next Level

Sponsored by: EMC

Ashish Nadkarni
January 2013

IDC OPINION

Traditional approaches to data protection are no longer sufficient given our economy’s emerging dependence on data-driven strategies. Businesses are relying on data more and more to improve their value proposition and widen their competitive edge. Suppliers are learning from businesses that their data must always be available. Companies are investing in tools and technologies that provide faster access to data, and they are also investing in products that protect their data and reduce the risk of data loss. For over 25 years, businesses everywhere have taken comfort in knowing that their data and applications are protected in once-a-day backup routines using a traditional tape-based or disk-to-tape-based data protection solution. But today, businesses are realizing they need more than a daily full backup. This is a major pain point that has existed for a long time, and the changing nature of infrastructure only makes it worse:

- **Virtualization is ubiquitous.** IT infrastructure is becoming densely virtualized. What was once reserved for tier 2 and 3 environments is now the new norm for mission-critical environments. Businesses are virtualizing even the most critical environments, and these environments need to be protected beyond the standard daily backup.

- **The new economy is driven by data.** IT infrastructure is becoming more information centric. Businesses are relying more and more on data as their primary asset, and data volumes are growing rapidly. This places a higher level of demand on the infrastructure to support near-instant recovery time objectives (RTOs) and tiered recovery point objectives (RPOs). Traditional data protection environments with single-threaded recovery mechanisms fall short in their ability to support this new paradigm of data-centric computing.

- **IT demands self-service.** Hypervisor and application administrators are now performing tasks that were once reserved for backup and storage administrators. They are demanding access to the storage and data protection infrastructure, which is forcing storage and data protection teams to rethink how they can delegate access while retaining ownership and accountability.

- **Users require granular recovery.** Users, for their part, are demanding granular recovery capabilities wherein they can quickly restore a single mailbox or a single item in the mailbox. For a mobile workforce that demands information on the go, lack of email access due to an incredibly long data recovery process means loss of productivity — and this translates to loss of revenue for the business.
Keeping such demands in mind, companies such as EMC — with vast experience in the storage and data protection market — have focused on optimizing their software management products to address these specific requirements. With their virtualization experience, they are building software products that offer advanced data protection capabilities in a simplified user interface. Vendors such as EMC understand that VMware administrators want to protect and restore their own data, and in an IT-as-a-service deployment model, service-level agreements (SLAs) will be unique across applications and tenants. The result is that products such as EMC AppSync provide businesses with the ability to:

- Easily create a robust data protection infrastructure where application data is continuously protected using storage-based point-in-time snapshots and replication solutions.
- Deploy a tiered RPO model fitting into an IT-as-a-service deployment wherein applications are protected based on policies and service plans and not forced into a one-size-fits-all approach.
- Offer predictable and reliable operational and disaster RTOs wherein backup or storage administrators do not become bottlenecks for restores. The operational burden for performing restores now shifts to application and hypervisor administrators. Leverage advanced protection functionality while remaining very simple to use and deploy by masking some of the complexity of the underlying replication technologies and providing an intuitive easy-to-use interface.

Products such as AppSync are forcing businesses to rethink operational and disaster recovery processes. They are creating a delegated data protection model that will surely gain more traction in the market as businesses continue their data-centric and densely virtualized transformation.

IN THIS WHITE PAPER

This white paper provides an overview of the data protection and recovery management solutions market. It explains how traditional recovery management solutions are slowly being replaced by application-centric solutions in the wake of dense server virtualization. As businesses transition their mission-critical applications to virtual environments, they are increasingly looking to applications such as EMC AppSync that offer a simple, self-service, and SLA-driven approach for protecting virtualized applications.

SITUATION OVERVIEW

In traditional IT datacenters, data protection and recovery environments have always been managed as silos — separate from the storage and compute infrastructure. In many cases, such environments offer very little flexibility on how data is backed up and recovered. Naturally, this rigidity has lent itself to IT organizations. Backup and recovery procedures are often rigid and require manual intervention, integration with applications is tedious and requires extensive testing, and more important, users or even systems administrators have very little control over the actual backup and recovery function. For the longest time, backups have been considered a once-a-day ritual that is "mysteriously" triggered at some point at night and is expected to be over
before the business day starts the following morning. During the day, the backup systems often are either down for maintenance or undergo self-maintenance activities — and therefore are unavailable for any user-initiated restores or backups that occur more than a day. For the most part, IT organizations have learned to live with this situation and “adjusted” their business processes to deal with this rigidity. However, the status quo is not sustainable.

This lack of sustainability is largely driven by the transformation of datacenter infrastructure to support private cloud environments — one that features service-oriented, data-driven, densely virtualized applications and has built-in process-based automation. In this new model, traditional tape-based data protection solutions with fixed backup windows and limited recovery options can no longer provide the level of granularity and control needed to protect the data and the applications it houses:

- Backups that occur once a day can provide rollback or recovery capabilities until the last known good backup, which could be on tape; in the worst-case scenario, this backed-up data could be a day old. In a virtualized environment, datasets residing in applications running on virtual machines have their own different change rates. Therefore, a wholesale restore to a point in time is not a feasible idea.

- In a densely virtualized environment, hypervisor administrators, administrators of individual guest operating systems (OSs), and administrators of applications running on these guest OSs need granular control over when and how frequently data is backed up. Furthermore, they need on-demand restore capabilities that allow them to quickly roll back the application or the guest OS in question. In other words, they need a backup/recovery solution that offers tiered recovery point objective/recovery time objective capabilities for the entire infrastructure and not just some components.

- In a traditional IT environment, application integration with storage data management capabilities such as point-in-time copies and replication is handled externally via scripts, and by and large, such integration is either independent of or runs parallel to the data protection and recovery solution. In most cases, this integration also assumes some kind of a privileged or super user role — with an assumed handshake agreement between administrators. With a data-centric approach in a densely virtualized environment, this handshake agreement is no longer viable or secure.

- Some environments still are not virtualized. For example, many businesses have not yet virtualized their mission-critical database server environments. Given the requirements around availability of such environments, an abstraction layer that automates backup and recovery interaction with the storage layer is essential.

In this "new world," data is one of the most important assets a business owns. For the new breed of businesses that harness value from Big Data and those that base their business on information research, application downtime is simply not an option. Therefore, such businesses are complementing their daily backup strategies with application-aware snapshots, clones, and continuous data protection (CDP) and/or replication solutions. This kind of approach provides several tiers of operational and
disaster recovery RPOs and, more important, quick restore or rollback capabilities that support reduced RTOs:

- Continuous data protection solutions often come in the form of an appliance and can be used to replicate data locally or remotely in order to reduce recovery points. They are excellent for always-on protection of mission-critical environments.

- Array-based snapshots that are taken several times a day allow rollback to the last known good state of the environment at the local site. They can be used to quickly recover databases, logs, and single mailbox items.

- Daily backups are still an important insurance policy should a restore be needed at a third location or on a separate infrastructure and can also double as a mechanism for long-term archiving and retention.

What binds these solutions together is management software that works with the underlying replication technologies and integrates with the applications. It ensures that all of these array- and appliance-based solutions work in perfect harmony with each other — and in an application-aware manner. This orchestration layer also allows multiple applications residing on the same infrastructure to interface with storage components independent of each other, allowing them to be governed by different but reliable and predictable recovery objectives. Furthermore, it provides transparent control over the storage layer — essentially making the recovery process the same whether the application is recovered from a local snapshot or from a remote replica.

This lack of reliability and predictability is one of the biggest shortcomings of traditional data protection solutions. It essentially forces all applications to adhere to a single recovery point objective. With resources such as tape drives that are single threaded, the infrastructure would force applications into separate recovery time objectives. For example, when restore jobs for one application are running, restore jobs for another application have to wait in queue, especially if the data they need to restore resides on the same tape or if all of the tape drives are used up in the restore. Thus, even with common recovery point objectives, applications can have widely varying recovery time objectives — in both operational and disaster recovery scenarios. Therefore, administrators of applications such as Exchange and databases perform disk-based backups in addition to tape backups. Depending on the resources used to restore data (i.e., from tape or from a local database backup or a combination of the two), the process for restoring for applications such as Exchange and SQL Server can become very tedious, error prone, and resource intensive.

One of the chief attributes of such orchestration applications is the ability to empower application owners and users to perform their own restores. These solutions need role-based access control in addition to deep application integration, regardless of whether they are virtualized or not. User-privileged functions need to be made available via a native but intuitive user interface (UI) so that specific restores such as Exchange mailbox recoveries can be self-serviced. Such capabilities not only offer unprecedented levels of granularity and control over recovery point and time objectives but also relieve the operational burden placed on backup administrators in servicing multi-application restores.
THE EMC APPSYNC SOLUTION

EMC is no newcomer to the challenges faced by businesses with traditional data protection solutions. These challenges have led EMC to develop a vast storage and data protection portfolio — built on array- and appliance-based, point-in-time cloning and replication capabilities that complement a full set of data protection software packages.

With AppSync, EMC is seeking to transform the approach taken to protect applications that reside in virtual environments. EMC AppSync provides a simple, self-service-based and SLA-driven approach for virtualized applications deployed on its VNX platform along with RecoverPoint. Applications can be protected in an application-consistent manner (i.e., the snapshot or replica copy is created to prevent any transaction losses) or a crash-consistent manner (i.e., all data is captured at the same time but may have lost in-flight transactions). Critical applications such as Microsoft Exchange or SQL Server can be protected in an application-consistent manner, which means that you "quiesce" the application and create an image that allows you to quickly restart the application once you have recovered the data. Other applications that reside on VMware data stores, for example, can be protected in a crash-consistent manner, which means you are making a copy without any host shutdown or quiesce process.

In developing AppSync, EMC has acknowledged that virtual environments are dynamic by design; therefore, a rigid host or hypervisor-centric backup solution does not fit the bill. Instead, AppSync follows the application that resides inside the guest OS, keeping track of its location as it moves throughout the infrastructure to ensure that the application SLAs are met no matter where the application resides. This "Recovery Assurance" functionality can be tailored to fit the SLA for each application and forms a fundamental building block for modern data protection infrastructure. This is in stark contrast to the one-size-fits-all approach in traditional data protection environments.

Because AppSync is SLA driven, its functionality is essentially policy based. Furthermore, such policies are tied to the capabilities of the storage system deployed in the infrastructure. For example, a Platinum policy could translate to the need to leverage synchronous replication. Once these policies are customized to fit the needs of the application and the capabilities of the storage infrastructure, they are made available as items in a "shopping cart." Application owners select the service plan that meets the recovery objectives of their applications. With one or two clicks, they are well on their way to protecting data. Furthermore, they can make space-efficient copies of production data and, if necessary, recover data quickly from the last restore point — AppSync takes care of the necessary processes behind the scenes.

AppSync makes protecting data easy by providing a customizable service catalog with prebuilt service plans and a workflow-like editing tool to call out actions that should be taken before and after the replication. Today, AppSync ships with four user-customizable service plans: Platinum, Gold, Silver, and Bronze. Each plan is tied to differing RPO capabilities offered by the supported replication capabilities of RecoverPoint (CDP) and to VNX snapshots.

AppSync has an extensible, feature-rich and yet easy-to-use management interface. EMC heavily leveraged the "Empower users but retain centralized control" philosophy in designing the AppSync UI. AppSync provides storage and application
administrators with four levels of administrative access. Storage administrators can launch AppSync directly from Unisphere. VMware administrators can launch AppSync from vCenter using the add-on to the VMware Virtual Storage Integrator (VSI). In addition, service providers can integrate AppSync into their management frameworks or create their own interface using its REST API.

**Benefits of Using Solutions Such as AppSync**

Solutions such as AppSync define a new paradigm for applications residing in virtual environments. For example, AppSync is an essential tool for anyone who has deployed or is considering deploying Microsoft applications in a virtualization infrastructure on a VNX platform. By deploying tools such as AppSync in their environments, businesses can:

- **Eliminate backup windows.** Solutions such as AppSync integrate with Microsoft application protection frameworks such as Microsoft Volume Shadow Copy Service (VSS). Microsoft VSS is a set of interfaces that implement a framework to allow volume backups to be performed while applications on a system continue to write to the volumes. Thus, production environments can be backed up without impacting performance at any time during the day, even during peak production hours. This eliminates the concept of a "backup window," especially in 24 x 7 environments where there are no "downtimes" or "off-peak hours." Such capabilities also allow the production volumes to be "snapped" multiple times a day and then backed up as secondary or tertiary copies. AppSync also supports integration with EMC and third-party tape/disk backup environments.

- **Take advantage of instant restore capabilities.** Businesses no longer have to struggle with restore lag times because of system availability. For example, AppSync's transparent integration with the storage systems means that businesses can restore data quickly and on demand. Data is restored in minutes using the SAN and not over the LAN. There is no longer a need to copy huge files over a shared IP network. Furthermore, businesses can choose to restore this data to an alternate location for test, development, or reporting purposes.

- **Deploy tiered application RTOs and RPOs.** Solutions such as AppSync are much more than snapshot-based data protection solutions. For example, because of its tight integration with RecoverPoint and VNX, AppSync offers multiple SLAs from continuous data protection solutions, clones, and snapshots. AppSync can be configured to manage zero data loss synchronous replication and asynchronous replication capabilities of the RecoverPoint platform as well as snapshots and clone capabilities of the VNX platform.

- **Implement automated RPO-based protection.** Such tools allow application owners to easily assign a service plan to an application. Service plans are based on a defined schedule that is tied to the recovery point objective of the application (such as Microsoft Exchange). The tool takes care of automating the execution of the service plan so that copies of application data are created on a prescheduled basis.
Gain confidence in application-based recovery. Such tools allow application owners to verify data before it is offloaded to disk or tape. They can also set alerts that inform them when the RPO goals for the application are exceeded. This provides a level of assurance in the ability of the infrastructure to meet or exceed stated RPOs for all applications — a task that was nearly impossible with traditional backup solutions.

AppsSync Deployment Scenarios

EMC is positioning AppSync as a comprehensive recovery assurance solution for Microsoft applications running in virtualized environments deployed on a VNX platform. Deploying AppSync in a data protection infrastructure means that the following features are added:

- Application-aware monitoring, reporting, and alerting. AppSync monitors the data protection state for VMware vSphere environments and applications such as Microsoft Exchange and SQL Server residing in it or on standalone servers.

- Recoverability of application copies. AppSync can provide a quick update on the "health" of application copies. This allows application owners and storage administrators to proactively identify and resolve issues that can impact a potential restore.

- Self-service application recovery status. AppSync provides an update on which application is being restored and which resources are being consumed for such restores.

- Application protection reports. The AppSync UI has built-in features to report on the protection status of every application under its control. Such features are essential in operational and disaster recovery process auditing.

Exchange Mailbox Restore as a Service

Many times, Exchange administrators need the ability to restore a single mailbox rather than the entire Exchange data store. Granular recovery of single items provides very effective mechanisms for rapid recovery of critical data for Exchange. EMC's ItemPoint software can be used in conjunction with AppSync to provide a fully integrated single mailbox and mail item recovery for Exchange 2010. EMC ItemPoint can perform search and restores using the Exchange replicas created by AppSync. The key to EMC ItemPoint is its ability to directly read EDB files located on the Exchange replicas.

SQL Server Protection with a Single Click

EMC AppSync is the ideal solution for database administrators to protect their SQL Server instances on their own. Storage administrators can preconfigure the AppSync service plan catalog for database protection and allow database administrators to select the appropriate plan depending on the database instance. Database administrators no longer have to worry about creating scripts for disk-to-disk backups or storage snapshots and clones.
CHALLENGES/OPPORTUNITIES

Applications such as AppSync are designed to help IT organizations ease into the “Empower users but retain centralized control” model. Their success largely depends on whether they live up to the promise of providing a self-service-based, SLA-driven model to application owners and administrators. EMC, with its reputation of standing by its products, will need to ensure that any inadequacies are promptly addressed. Because EMC will be selling AppSync largely to existing VNX customers, it has an opportunity to improve their customer experience and potentially expand its reach within these organizations.

While AppSync is a great solution for VNX environments, it currently provides application-consistent protection only for Exchange and SQL Server environments. Businesses looking to expand the distributed data protection management model to other applications may have to make do with a crash-consistent level of support for VMware vSphere environments. Similarly, lack of support for other storage platforms, including EMC's own offerings, may force businesses to look elsewhere.

Third-party solution providers are trying to make a comeback in this market with innovative solutions designed for Microsoft applications. Data protection vendors that compete with EMC in the data protection software market are also trying to make a play to unseat AppSync as the solution of choice for VNX- and RecoverPoint-based environments, using hooks built into their own data protection software.

EMC should seize this opportunity to quickly expand the reach of AppSync beyond VNX-, RecoverPoint-, and Microsoft-only environments. It should examine commonly deployed applications such as other database solutions from Oracle and SAP, alternative hypervisor platforms, and productivity application suites. It should also seriously consider tying AppSync to other commonly deployed portfolio solutions such as Avamar, Data Domain, and NetWorker.

CONCLUSION

Products such as AppSync are here to stay because they offer businesses the ability to offload the bulk of their operational burdens from their storage and data protection mechanisms. Businesses thus can empower their application owners to replicate and restore their own data sets without changing the balance of power, which frees up storage and data protection teams to focus on strategic initiatives. Businesses should seriously consider making AppSync a part of their arsenal if they happen to have VNX arrays in their infrastructure or are considering deploying them. Similarly, businesses should consider investing in application and hypervisor administrators to ensure the skills are in place to handle these additional tasks. They should also foster collaboration between the various “backup and recovery” groups in order to close any gaps that may be exposed as a result of process change. Products such as AppSync also fit well in IT-as-a-service and cloud deployments because they provide the tenants of the service the ability to perform self-service on backups/restores of their application environments. By making policies available to such tenants, service providers can essentially allow their tenants to configure and deploy applications quickly.
EMC — and vendors like EMC — should continue to monitor this progression in environments in which AppSync or similar products are deployed. They should learn from the experiences of their customers to improve the user experience and add more functionality — potentially extending it to other EMC (and non-EMC) platforms.

Copyright Notice

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.

Copyright 2013 IDC. Reproduction without written permission is completely forbidden.