

SANscreen® Replication Assurance
Technical Overview
March 2006

Executive Summary

Data replication is vital to the success of any storage strategy. Widely used throughout the data center, replicated data enables backup, testing, upgrades, data mining, disaster recovery, and business continuity. Replication, however, brings its own set of challenges. While application owners request specific levels of replication service, storage managers must translate these service-centric requests into device and script configurations. A gap exists between the replication services requested by application owners and the device-centric nature of storage resource management (SRM) applications used to implement, monitor and manage these services. To fill this gap, storage teams typically rely on tedious spreadsheet reporting, error-prone employee memory or hand-written status tables – all of which falls short of the job. Onaro's SANscreen® Replication Assurance bridges the analysis gap by providing a service view of replication that results in improved storage quality, return on storage and compliance.

Challenges of Replication Implementation

In this paper, replication refers to the copying of data from one volume to another to support data warehousing, test and development environments, business continuity or disaster recovery. Organizations replicate data through various means including snapshot-based replication as well as synchronous and asynchronous replication. All of these processes involve a degree of complexity that grows as environments change and expand. The following scenarios illustrate the challenges faced during today's typical replication implementations.

Maintaining the Required Recovery Point Objective (RPO)

To maintain the required RPO for an application, a storage team must ensure that:

- All volumes for an application are properly replicated
- The replication occurs in the correct timeframe
- The RPO meets the required standards

The primary contributors to RPO violations are inconsistencies between changes to the production storage environment and the required replication configuration. Other triggers include erroneous scripts that fail to create copies on schedule.

Monitoring RPO is an extremely challenging and manually intensive task. First, storage teams must verify that all volumes for an application are replicated. Then these teams must check to ensure that the scripts which manage the replication have executed correctly. Finally, the time of last replication must be noted to calculate the RPO for the application. Due to the manual nature of the job, it is nearly impossible to determine continuous RPO compliance. Yet this is precisely what is needed in order to recover from data corruption or other incidents in a timely manner.

High-growth applications that are constantly demanding additional storage capacity make RPO calculations even more difficult. To correctly implement replicated capacity, an organization must bring additional storage online in accordance with the existing replication strategy and RPO policy. This requires manually reviewing and verifying the device configurations after each storage addition. With some organizations executing 50 or more changes per week, this process becomes extremely tedious and error prone given the manual nature of the job.

Generating Chargeback Reports for Replicated Storage

In many organizations, storage managers must issue detailed financial reports to every application manager. These reports must document the total cost of the storage service including production storage, replicated storage and tier. In many cases, the bulk of the cost resides in replicated storage.

The challenge is to generate an up-to-date report that accurately reflects the full cost of replicated storage in a dynamically changing environment. Device-level reports show which volumes are replicated. But tracking this information back to the application is a tedious and manual process. As a result, most storage teams are unable to easily cross-charge for replicated storage. This lack of chargeback capability is costly given that replicated data can represent up to three times the storage requirements of production data.

Migrating Replicated Storage Arrays

Storage arrays and switch devices are typically replaced every two to three years. Each such change opens the door to improvements and efficiencies but creates huge implementation challenges. These challenges multiply when the array is replicated.

In a typical array migration, a storage team must map all the host services enabled by the production array, all replication policies, and the array's potential role as a target for remote replications. Once this discovery is complete, the team must then eliminate any existing problems to the current array services prior to the migration. The challenge lies in determining the status of all array-provided services and their compliance with the requested service levels. Once again, Excel spreadsheets, human memory and white boards serve as the knowledge repository for this critical information. The process is manually intensive, time consuming and prone to error.

Assessing the Impact of Changes to Inter-array Links

Links between arrays can fail or change for a variety of reasons. Without a functioning inter-array link connection, the RPO starts to increase dramatically. However, tying the change in link status back to an application is an extremely challenging task. Once again, storage teams depend upon human memory, spreadsheets and tedious white-board analysis to fill the gap. For real-time impact analysis, this solution is less than optimal at best.

The Analysis Gap

Solving these replication challenges demands strict operating procedures, massive amounts of self-auditing and rigorous record keeping. Unfortunately, most storage teams are lean organizations that simply cannot afford to dedicate the time, cost and resources needed to meet these demands. As a result, most organizations rely on error-prone manual processes that ultimately trigger sporadic incidents and create a firefighting environment.

But why is this still the case given today's expensive SRM applications? While SRM solutions provide significant amounts of information about device events, they say nothing about how a particular device-related incident affects the overall service delivered to an application. There is a gap between the device-centric view presented by SRM applications and the service view required to meet the needs of business application owners. IT storage teams are forced to plug this gap with spreadsheets, human memory, white boards and lots of manual labor. What's needed is a solution that can correctly monitor replication from a service-level standpoint, rather than from a device-centric viewpoint.

SANscreen Replication Assurance Fills the Gap

SANscreen Replication Assurance bridges the analysis gap between device management and service management. Device managers do an adequate job of monitoring device status and managing networked storage elements. But the output of their data is not expressed as services. In contrast, SANscreen Replication Assurance provides storage teams with real-time visibility into their storage services.

SANscreen Replication Assurance starts by unobtrusively collecting element information throughout the storage area network (SAN). Agentless in design, SANscreen Replication Assurance quickly discovers the entire SAN and all arrays. As the element information is assembled, SANscreen's service intelligence engine pieces together the true SAN topology – determining which sources of information are “trusted” and, even more importantly, evaluating the interdependencies between the elements to create a service-level view of the environment.

The service intelligence engine operates not only during initial installation, but on a continuous basis. So as new elements are added to the system or as existing elements give off alarms, SNMP traps and other alerts, the service intelligence engine processes this information and reports how the incident affects service delivery.

In addition, the service intelligence engine has a simulation component that enables users to model future changes to the environment and predict their impact on service quality before they actually occur.

Impact on the Replication Environment

SANscreen Replication Assurance delivers a service view of replication by enabling storage teams to visualize the entire replication environment:

- From an application to its source data
- From the source data to the replicated data
- From replicated data back to the disaster recovery, business continuity or data warehouse servers, if installed

Once the replication environment is completely visible, SANscreen Replication Assurance enables verification that the environment is complying with established replication policies for:

- The application's RPO
- Local and remote copy count for the application
- Asynchronous, synchronous or point-in-time copies
- Redundancy for the remote hosts
- Access paths for the remote hosts

Any deviation from these policies results in a violation.

Beyond steady-state operations, SANscreen Replication Assurance also helps plan for the future. With SANscreen, storage teams can simulate changes to the replication environment and identify their impact on replication service levels at each step of the change process. In this manner, storage managers can proactively anticipate and avoid any reduction in replication services due to a planned change – and decide the best course of action on how to proceed.

Throughout this entire process, SANscreen Replication Assurance also provides robust auditing and service analytics reporting.

Impact on Replication Challenges

Specifically, how does SANscreen Replication Assurance address the replication challenges previously defined? Let's review each in detail.

Maintaining the Required Recovery Point Objective (RPO)

Maintaining the RPO for an application requires that all volumes are replicated properly, the replication occurs in the correct time frame, and the RPO meets the required policies. SANscreen Replication Assurance understands the RPO policies for each application and monitors the environment to identify any RPO violations. In addition, SANscreen Replication Assurance

knows the most current RPO point for rapid recovery after data corruption or disaster recovery events.

For newly added volumes, SANscreen Replication Assurance recognizes their presence and verifies that their replication policies are in accordance with their application. Any new volumes which do not meet the established replication policies will trigger a violation for the application.

Generating Chargeback Reports for Replicated Storage

Since SANscreen Replication Assurance provides full visibility into an application's original data, replicated data and entire replicated environment, it is simple to identify and charge back any replicated storage resources used by a particular application.

Migrating Storage Arrays Using Replications

SANscreen Replication Assurance jump starts the migration process by automatically discovering and mapping all services enabled by the array for production data, replications and the array's role as a source or target for remote replications. Any problems with the current services are highlighted. In addition, SANscreen Replication Assurance's planning function optimizes the migration and tracks each step of the change process to ensure its compliance with the approved change plan.

Summary

Implementing a replication strategy adds complexity to the storage environment. Storage teams must manually translate the device-level information from SRM applications to determine if the service-level targets for replicated environments are being met. SANscreen Replication Assurance bridges the gap between device-centric SRM applications and the need to manage replication as a service. The technical benefits of SANscreen Replication Assurance directly translate into reduced operating labor costs, reduced capital costs and improved service quality for applications.