Technical White Paper

Dell EMC PowerProtect Data Manager Software: Managed File Replication(MFR)

Abstract

This white paper explains Power Protect Data Manager Replication technology. It discusses Architecture, workflow, topologies and configuration of PowerProtect Data Manager MFR Replication.

July 2019

Revisions

Date	Description
July 2019	Initial release

Acknowledgements

This paper was produced by the following:

Author: Sonali Dwivedi

Table of Contents

Executive summary

Data protection has become an integral and essential part of any successful business. The need to provide a powerful, scalable and yet simple disaster and operational recovery solution is at all-time high.

IT teams are also looking for a solution that is scalable, easy and efficient to implement and handle the workload of their small and mid-size environment. To meet the Mid-Market industry demands Dell EMC has come up with a new offering called PowerProtect Software. It is Software-Defined with built-in deduplication, Cloud Optimized, Self Service, Saas based monitoring capability and has Modern Service based architecture.

PowerProtect Data Manager also offers Managed File Replication (MFR) when it is integrated with Data Domain using ddboost based replication technology. With Replication enabled it offers complete Disaster Recovery solution meeting customer's satisfaction and SLA.

Audience

This white paper is intended for customers, partners, and employees who want to better understand, evaluate, and explore offering with PowerProtect Data Manager Replication. Familiarity with PowerProtect Software and Data Domain is required.

1 Component Overview

1.1 Power Protect Software

Dell EMC PowerProtect Software is next generation Data Management Software Platform. It is Software-Defined, Cloud Optimized, Self Service, Saas based monitoring capability and has Modern Service based architecture. It helps in minimizing data protection complexity with a simplified user interface, workflow, centralization and protection life cycle. It also provides complete automation for Storage configuration and management, application agents configuration and compliance verification.

PowerProtect Software is integrated with one or more Data Domain for data storage making it a complete data protection solution.

1.2 Data Domain

Data Domain systems are disk-based inline deduplication appliances that provide data protection and disaster recovery (DR) in the enterprise environment. Data Domain system features ensure data integrity, reliable restoration, efficient resource usage, and ease of management. Licensed features allow you to scale the system feature set to match your needs and budget. DD OS data deduplication identifies redundant data during each backup and stores unique data just once. The storage of unique data is invisible to backup software and independent of data format. Data can be structured, such as databases, or unstructured, such as text files. Typical deduplication ratios are 20-to-1, on average, over many weeks.

Data Domain Boost (DD Boost) provides advanced integration with backup and enterprise applications for increased performance and ease of use. DD Boost distributes parts of the deduplication process to the backup server or application clients, enabling client-side deduplication for faster, more efficient backup and recovery.DD Boost is an optional product that requires a separate license to operate on the

Data Domain system. You can purchase a DD Boost software license key for a Data Domain system directly from Data Domain.

1.3 Managed File Replication

Managed file replication, which is used by DD Boost, is a type of replication that is managed and controlled by backup software. PowerProtect is using Managed File Replication to replicate data between Data Domain.

With managed file replication, backup images are directly transferred from one DD system to another, one at a time, at the request of the backup software. The backup software keeps track of all copies, allowing easy monitoring of replication status and recovery from multiple copies. Managed file replication offers flexible replication topologies including full system mirroring, bi-directional, many-to-one and one-to-many enabling efficient cross-site deduplication. Here are some additional points to consider about managed file replication:

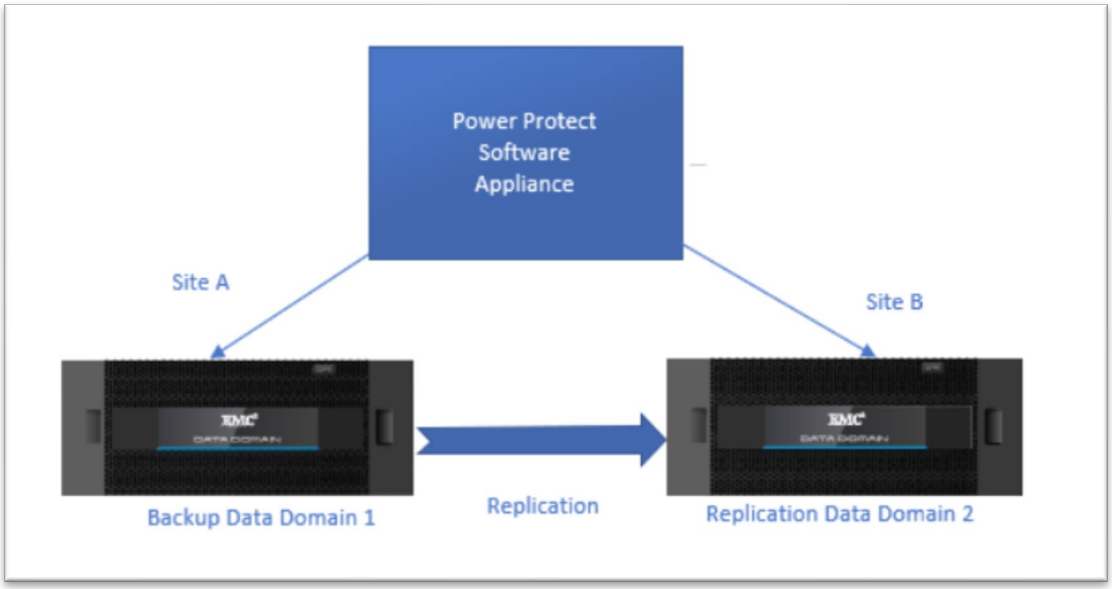
- Replication contexts do not need to be configured.
- Lifecycle polices control replication of information with no intervention from the user.
- DD Boost will build and tear down contexts as needed on the fly.

2 Overview of Replication

The PowerProtect replication feature allows transfer of data from source Data Domain where backup is taken to a Destination Data Domain where second copy called replica is stored for Disaster Recovery purpose.

PowerProtect Software stores the metadata of backups in Elastic Search Node which is not visible in the PowerProtect Data Manager UI and data is stored directly on Data Domain. When the replication protection lifecycle schedule runs for each client, data moves from source to destination Data Domain.

This operation is only possible when you use minimum of two Data Domain system with PowerProtect Software, then the replication process transfers backup data from the source Data Domain system to a destination Data Domain system.



2.1 Replication Architecture Workflow

In this use case we have minimum architecture of two Data Domains registered with same PowerProtect Data Manager Appliance. Data Domain in Site A stores the backup data and Data Domain in Site B is configured to store the second copy of the backup called replica copy. Following steps explain the replication workflow:

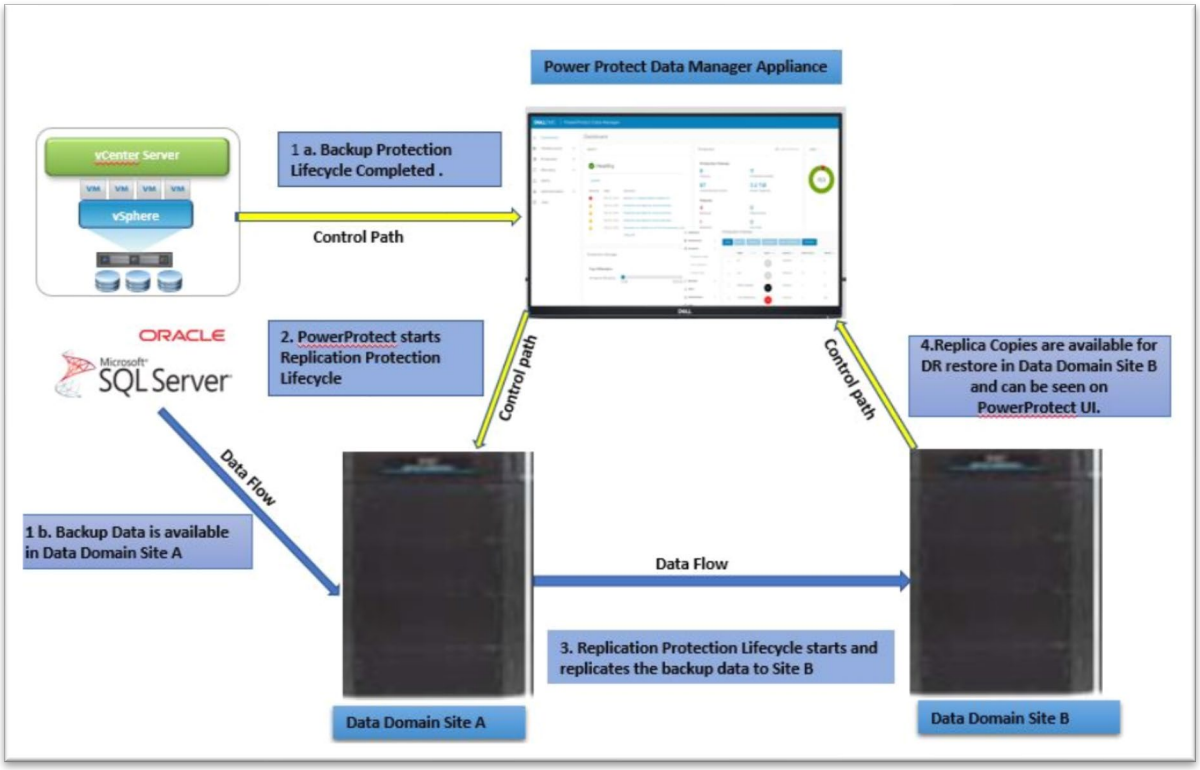
1. As per backup schedule, first Backup protection lifecycle (PLC) stage completes.

Note: You should not schedule replication window at the same time as backup window or overlapping with Backup window as it can cause replication issues.

2. Now Backup copy is available in Data Domain in Site A to replicate.
3. MFR Replication schedule starts from Data Domain in Site A to Site B and runs the Replication protection lifecycle (PLC) using ddboost replication and replicates the backup.

Once the replication is successfully completed, the replica copy is available for restore in Data Domain in Site B.

The replica copies can now be seen in Destination Data Domain in PowerProtect Data Manager Infrastructure and Recovery Assets Tab.

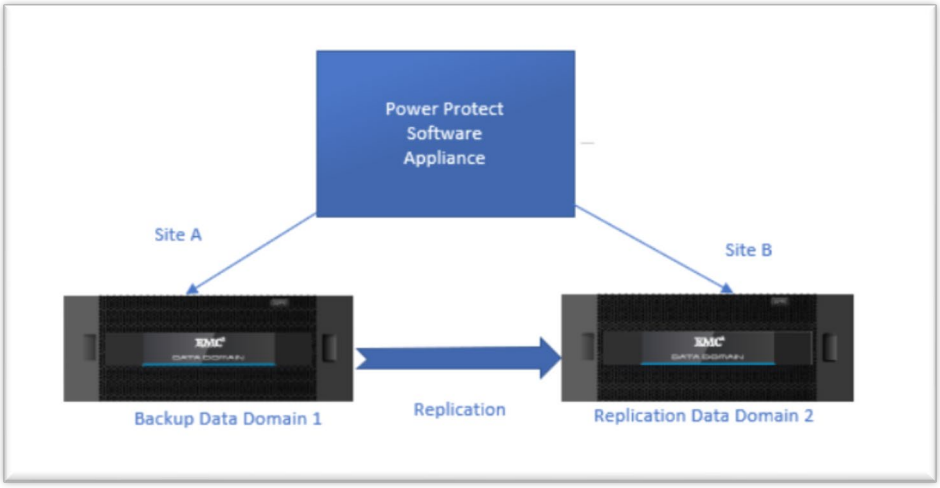


2.2 Replication Topologies

- If the PowerProtect Data Manager uses more than two Data Domain system, then you can use either a single destination Data Domain system or multiple destination systems for replica copies.
- We can also have more than two Data Domain added to same PowerProtect Data Manager appliance and we can backup or replicate from one Data Domain to multiple Data Domain and vice versa using same PowerProtect Data Manager appliance i.e cross replication is allowed as well.
- All of the data is replicated using DD Boost and deduplication is taken care by Data Domain.
- Following are the topologies that are currently possible with PowerProtect Data Manager. These replication topologies can be achieved by defining various policies or schedules and targets in Replication Protection Lifecycle(PLC):

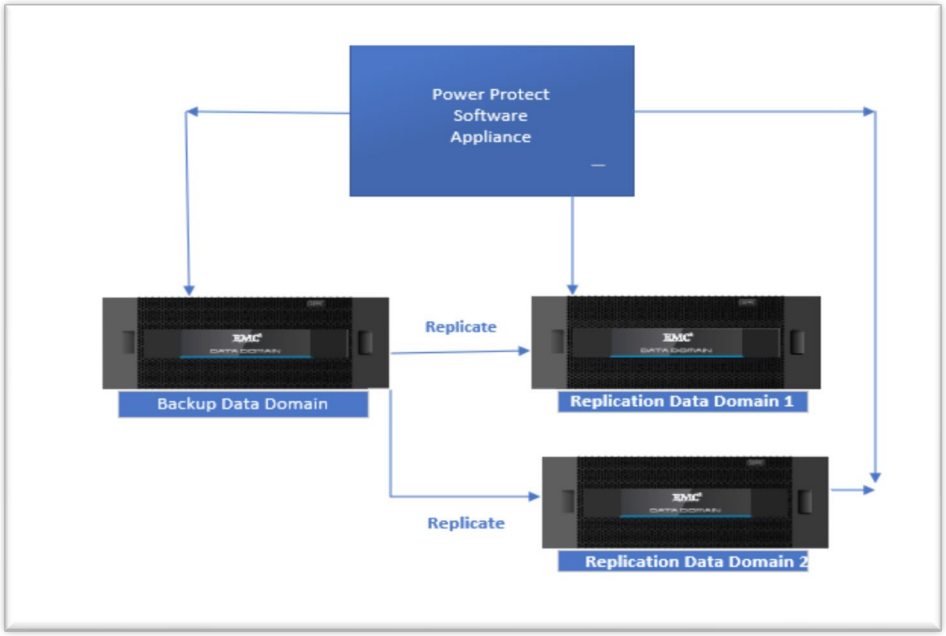
2.2.1 One to one replication

The following figure illustrates an environment where backup data replicates from a single PowerProtect appliance having two Data Domain system. Replication runs backup data from Data Domain 1 to Data Domain 2. Data Domain 1 hold the backup copy and Data Domain 2 holds the replica copy.



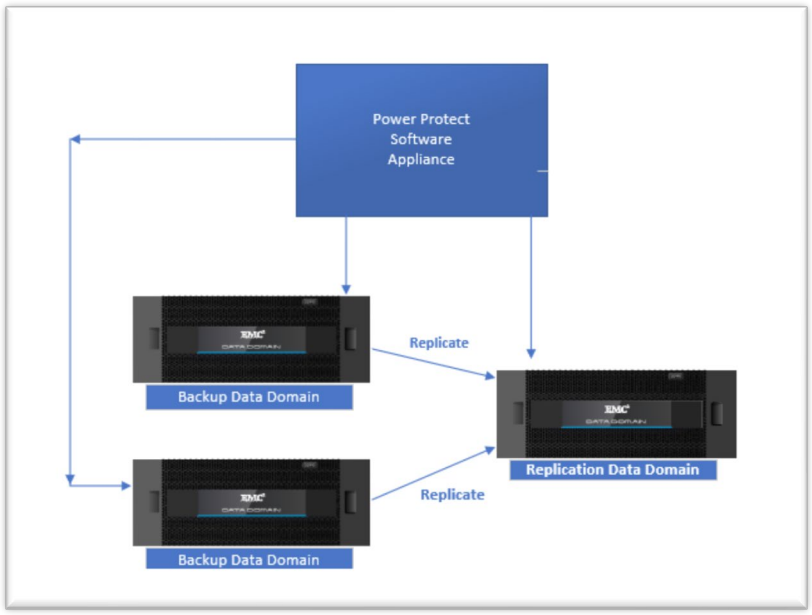
2.2.2 One to many replication

The following figure illustrates an environment where backup data replicates from a single source Data Domain system to multiple destination Data Domain systems.



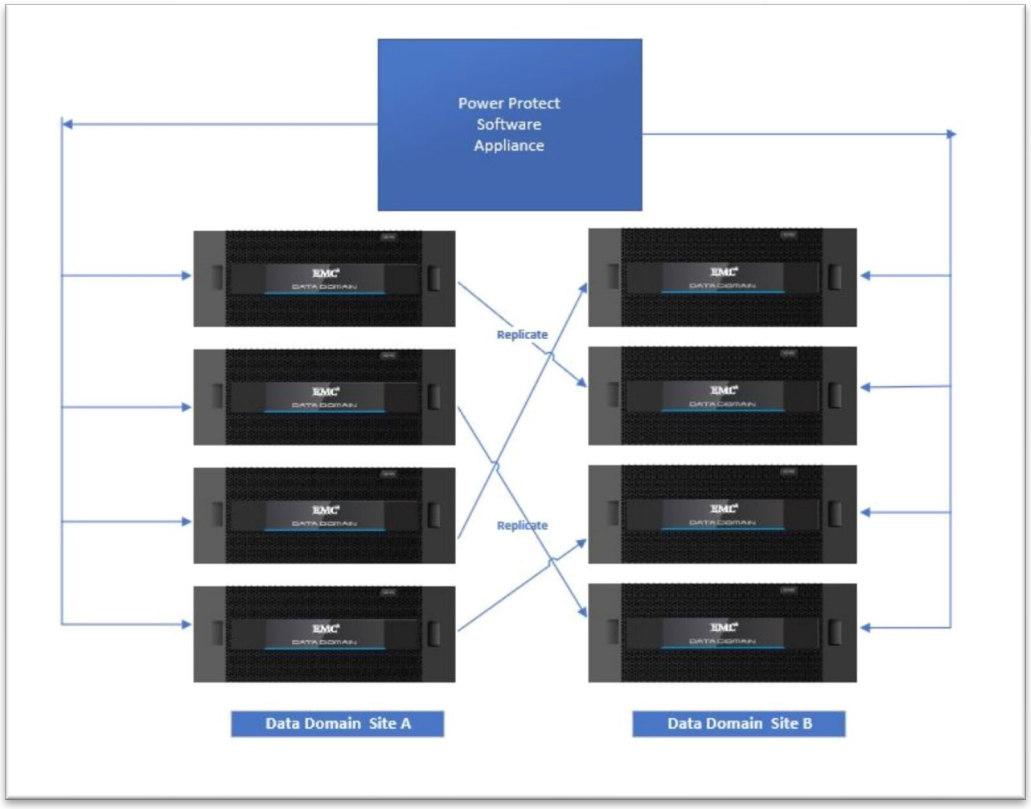
2.2.3 Many to one replication

The following figure illustrates a PowerProtect Software appliance with two backup Data Domain systems. It replicates the backup data on the two source Data Domain systems to a single destination Data Domain system.



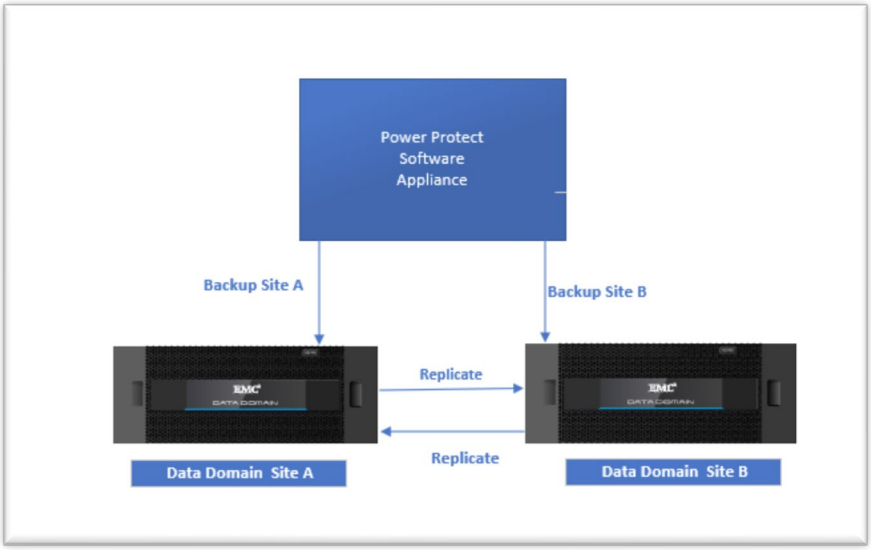
2.2.4 Many to Many replication

The following figure illustrates an environment with multiple destination Data Domain systems replicating to multiple destination Data Domain systems.



2.2.5 Cross Replication

The following figure illustrates an environment with Data Domain on each site that backups the data and replicate to each other so both Data Domain have backup and replication data of cross sites.

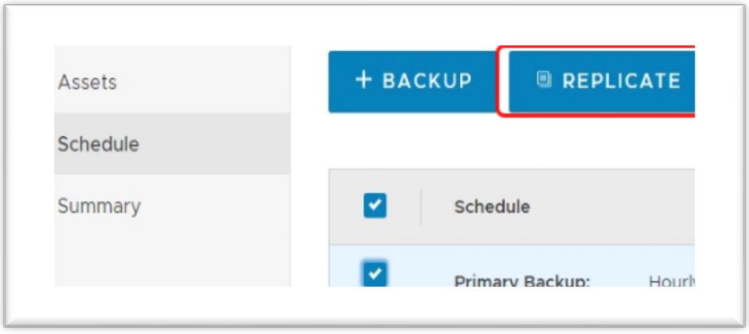


3 Configuring Replication

- To configure replication login to PowerProtect Data Manager Interface as user admin or root.
- Add one or more Data Domain where you want the replica copies.
- You need to ensure that backup protection lifecycle policy is created for the client asset and one or more backups are available.
- Encryption for Replication needs to be enabled from PowerProtect Data Manager UI and it needs to set for source as well as target Data Domain otherwise replication will fail.
- Number of replication stream is set at Data Domain level.
- You can have same or different network for replication, but it also needs to be set in Data Domain. Please refer to guide Data Domain Operating System 6.2 Administration Guide for more details.

  Note until the Backup Protection Lifecycle stage(PLC) schedule is created you will not be able to select the replicate tab as shown in figure:

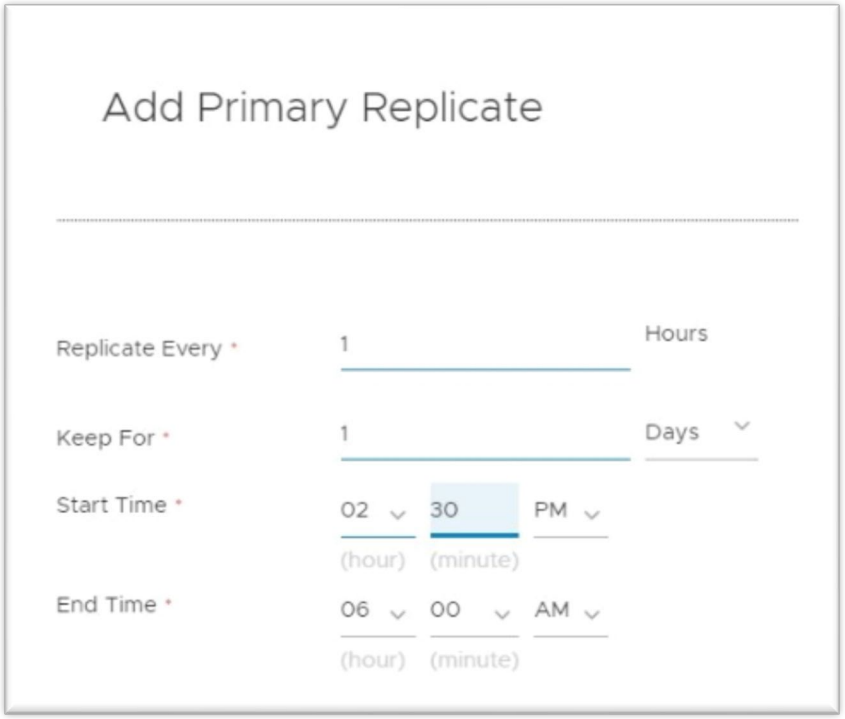
Backup Protection Lifecycle stage is not completed



Backup Protection Lifecycle stage is completed and now replicate option is available

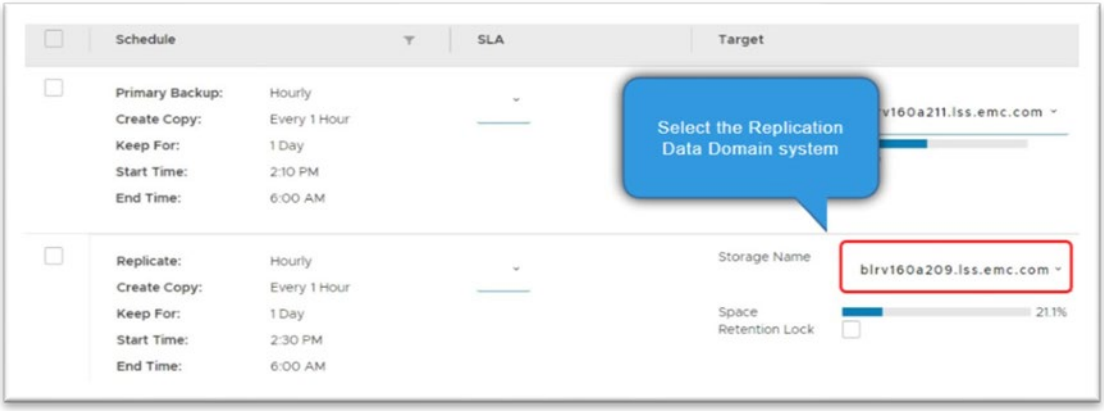
3.1 Replication Schedule and Retention

- When you select the replicate window following settings are available. Please add the start time of replication schedule after the backup window gets over otherwise replication will fail.
- In case there are many backup copies available to replicate and replication windows get over after replicating few copies then replication will not start from scratch in next schedule, but it will replicate from where it left off in last schedule.
- Replication retention needs to be carefully set for all backups as per the SLO and SLA of customers.



3.2 Destination Data Domain Selection

After this is done we can change the destination data domain for the replication using the drop-down window from replication protection lifecycle policy schedule. If you have multiple data domain please be careful to select the correct destination data domain.



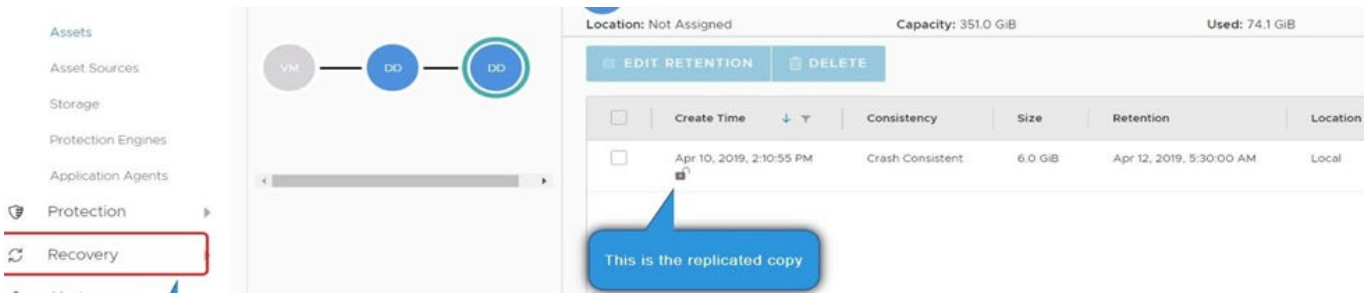
3.3 Replication Monitoring

To check if the scheduled replication completed successfully you can check Jobs tab and it will show if replication was successful, running, failed, queued, partial success or cancelled.

You will not see any errors or warnings if only few replication copies were able to replicate in the scheduled replication window. If this happens you can either increase the replication schedule window or wait for the next replication schedule to complete the data replication.

3.4 Verify Replicated Data

After replication shows successful or partial success, the replica copy will show under the replication Data Domain in Recovery→Asset→DD window during restore as well as under Infrastructure→Assets→DD.



Interoperability

PowerProtect Software Appliance v 19.1 supports Data Domain OS version 6.1.2 and higher. For more information on version compatibility please refer to link <compatibility guide link>.

One PowerProtect Software Appliance requires minimum two Data Domain and maximum 20 Data Domain for replication.

Conclusion

This paper provided information on PowerProtect Software enabled MFR replication leveraging Data Domain technology.

This paper thoroughly discussed the concepts of managed file replication with Data Domain integration workflow, configuration as well as the rich use-cases and topologies this solution enables.

References

The following documents were used in writing this whitepaper. All documents are available on https://support.emc.com.

Data Domain Operating System 6.2 Administration Guide

Data Domain Operating System 6.1 Administration Guide

PowerProtect_Data_Manager_Administration_and_User_Guide 19.1