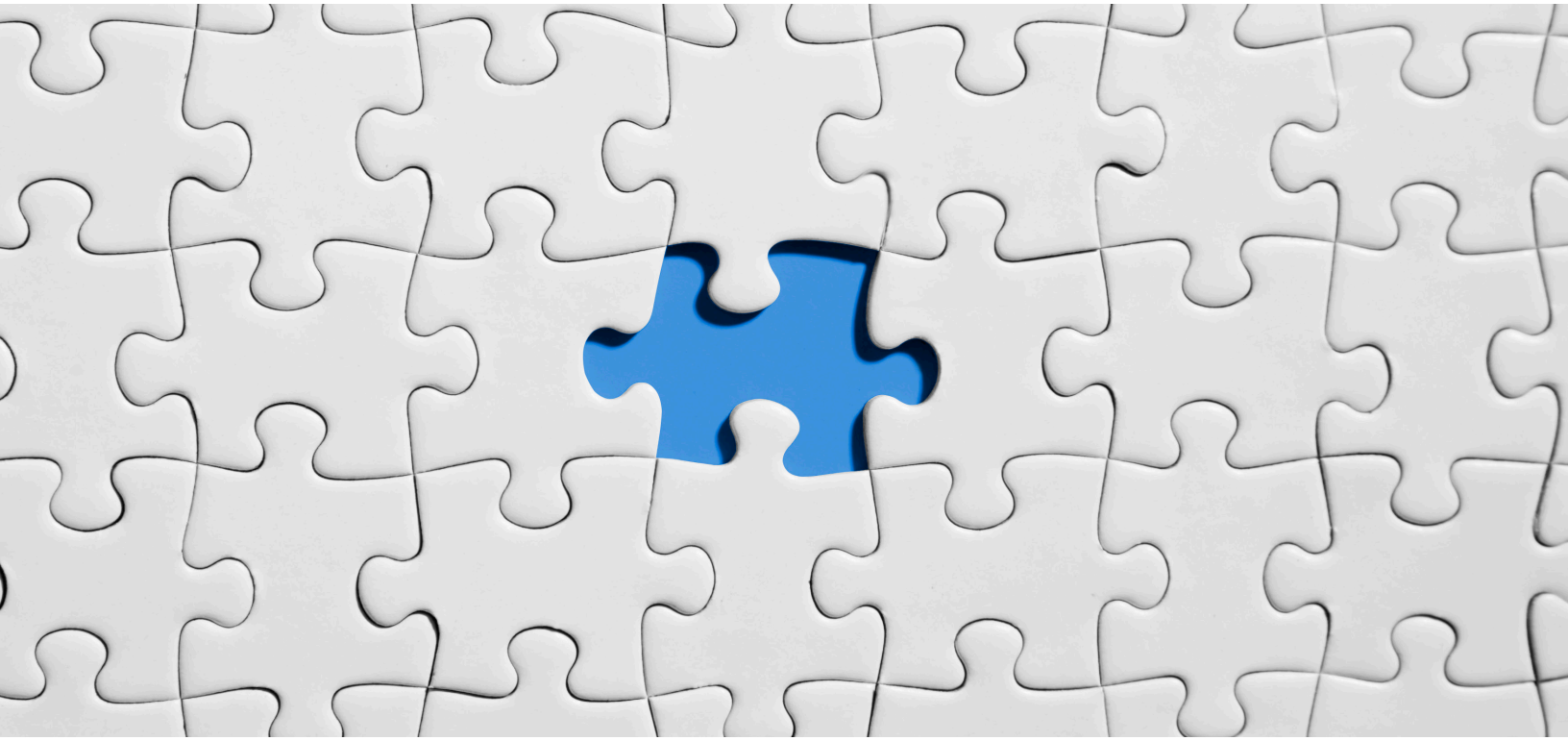
DELL Technologies

GUIDELINES TO SIZE DD IN A CYBER RECOVERY VAULT



Yashaswini N

Manager1, Inside Sales Management
Dell Technologies
Yashaswini.n@dell.com

Shrishti Shetty

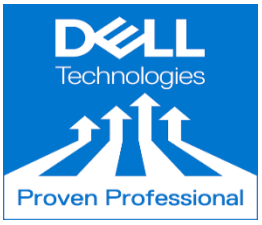
Specialist 2, Inside Product
Dell Technologies
Shrishti.shetty@dell.com

Abhijith M

Specialist 2, Inside Product
Dell Technologies
Abhijith.m@dell.com

DELL Technologies Proven Professional

Knowledge Sharing Article



The Dell Technologies Proven Professional Certification program validates a wide range of skills and competencies across multiple technologies and products.

From Associate, entry-level courses to Expert-level, experience-based exams, all professionals in or looking to begin a career in IT benefit from industry-leading training and certification paths from one of the world's most trusted technology partners.

Proven Professional certifications include:

- Cloud
- Converged/Hyperconverged Infrastructure
- Data Protection
- Data Science
- Networking
- Security
- Servers
- Storage
- Enterprise Architect

Courses are offered to meet different learning styles and schedules, including self-paced On Demand, remote-based Virtual Instructor-Led and in-person Classrooms.

Whether you are an experienced IT professional or just getting started, Dell Technologies Proven Professional certifications are designed to clearly signal proficiency to colleagues and employers.

Learn more at www.dell.com/certification

Table of Contents

Disclaimer: The views, processes or methodologies published in this article are those of the authors. They do not necessarily reflect Dell Technologies' views, processes or methodologies.

Executive Summary

In response to heightened Cybersecurity risks, several years ago leaders in the U.S. financial industry created sheltered harbor to protect customer confidence in the financial system. Sheltered harbor is not a vendor, product or service. It is a non-profit industry-led initiative to enhance financial sector stability in the U.S.

How do they do that? They have come up with a complex and important specification to implement by their participants to increase Cyber Resilience. Dell Technologies is the first compliant participant in the Solution Technology Provider Program and the Sheltered Harbor Alliance Partner Program helping customers achieve this level of resilience.

For over 20 years Dell Technologies has helped customers recover data when accidentally deleted or compromised through some form of database corruption. Today we can use that same technology to protect our customers against new attack vectors where files are deliberately corrupted or deleted.

Dell Technologies achieves this is by deploying Data Domain in a physical, isolated, air gapped vault. That vault holds a base level of compute, but it is a very simple subset. We work with the clients to identify the critical subset of data and take a daily copy of that data and push it to the vault. We only connect the vault to the production network for a minimum amount of time; when the data is transferred and, once the transfer is complete, terminating the connection.

Once the data is secured in the vault, we take a snapshot of it and have the option of leveraging retention lock, a feature that can make the data immutable so that in the event that bad actors get access to the vault, that data is not compromised by them. Additionally, we can run analytics everyday looking for indicators of compromise in the backup image. If malware has made its way through the vault; we can flag it up to the security operations center and take actions. Hence, when we are recovering the data, we know that it is valid.

Dell Technologies Cyber Recovery Overview

Today, a customers' business is data-driven and they need to be ready for anything that happens to their data. Destructive attacks such as ransomware is rising. It is more important than ever for customers to protect their mission critical information. Due to the new types of attacks, an enticing target for an attacker – after a production environment – is their backup and disaster recovery data. Clearly, businesses cannot continually run backups. So how can precious data be protected?

Cyber Resilience: the ability to withstand and recover efficiently from an attack. Though organizations may have implemented disaster recovery and backup in case of physical disaster, this does not help in the case of Cyber-attacks.

One of the best ways to do this is in an air-gapped cyber recovery vault. As illustrated in Figure 1, the vault is connected to the network only when transferring data. It is then disconnected making it virtually invisible to attackers.

The shift to remote and virtual operations by many organizations has led to exponential growth in Cybercrime: in fact, the FBI reports cybercrime has quadrupled since the beginning of COVID-19. The impact can be crushing, both to individuals and to companies.

Figure 1

To combat this, Cyber Recovery has become even more important than we can imagine. Keep in mind, Cyber Recovery is not the same as Disaster Recovery. which is all about physical location and the events that can happen there. For example, any natural disaster or accidents where a secondary copy to recover from would be needed.

A Cyber event on the other hand can span the entire environment. It can even affect Disaster Recovery capability. It is important to position Cyber Recovery because it is targeted at a much different recovery point objective (RPO) and recovery time objective (RTO). We need to help customer understand that these are two distinct things that they will require in their line of business.

Figure 2 highlights the importance of each while still underlining the difference in scenarios that they can be used in.

Category	Disaster Recovery	Cyber Resilience
Recovery Time	Close to instant	Reliable & fast
Recovery Point	Ideally continuous	1 day average
Nature of Disaster	Flood, power outage, weather	Cyber attack, targeted
Impact of Disaster	Regional; typically contained	Global; spreads quickly
Topology	Connected, multiple targets	Isolated, in addition to DR
Data Volume	Comprehensive, all data	Selective, includes foundational services
Recovery	Standard DR (e.g., failback)	Iterative, selective recovery; part of CR

Figure 2

General Guidelines to Size the Vault

Growth and change are the primary deciding factors for sizing a PowerProtect Data Domain in the Vault. Along with data retention, these variables can really dictate what is going to happen over a span of time.

Below are the factors that determines the size of Data Domain in the Cyber Recovery Vault

- **Growth** (Estimated backup capacity increase annually)
- **Change** (The rate of change on daily backup data)
- **Retention** (The period or duration for which the backup is retained)
- **Workloads** (The type of data or application that is going to be backed up)
- **RPO** (how frequently data should be sent to the vault to ensure a specific recovery point can be achieved)
- **RTO** (how quickly the data must be recovered from the vault to achieve business requirements around downtime)

Collectively, these variables when it comes to Cyber Recovery, can have a dramatic impact.

Now let's discuss effective retention with the help of a use case.

Consider a retention of 105-days in the Active Tier and 14-day vault retention applied to the MTree with daily syncs (replications) into the vault. Let’s examine why the vault ends up being or must be bigger in size than the production. We will frame out a Data Domain system based on the days and not on the model and capacity.

Figure 3 shows the 14-day mark for the vault retention. That is the data we want to protect and care about. We are going to bring in the workload and see what it looks like in a steady state.

So, in the steady state using the below example we just described the effective retention would be 105 x 14 which leads to 1,470 copies that would be retained in the vault. But there is only 14-days’ worth of data that we wish to protect. Let us now look at it with a sizing example.

When it comes to Cyber Recovery, the only kind of replication that is supported is MTree replication. So, each MTree which is being retained for 105 days in production will be replicated to the vault where this 105-day retained backup copy would again be retained for 14 days in the vault. As a result of this, data will be retained for durations more than what is required, hence resulting in a larger capacity requirement.

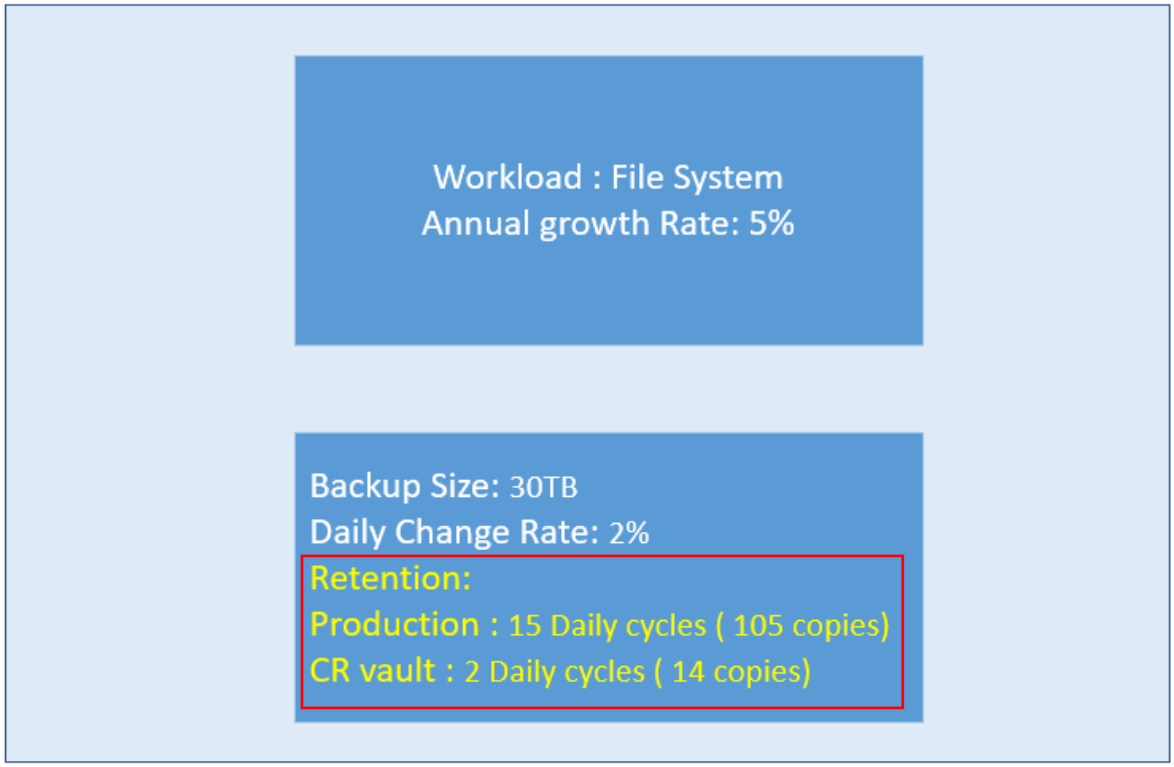


Figure 3 (Solution builder inputs)

Since we are dealing with MTree, it cannot be deleted. Thus, the KPI that we care most about is that effective retention. There are ways to control retention within the primary backup solution's design. We can use a host that can direct a clone operation or we can set up a replication job to a single node Avamar server (which will be the initiator). An initiator is simply a replication target. So the Avamar server (initiator) integrated with the Data Domain in the production can control which data goes into the vault. The MTree associated with the initiator can be smaller than the corresponding data stored on the production Data Domain because many backup utilities allow primary retention and specific workloads to be modified and filtered out during the incoming replication process.

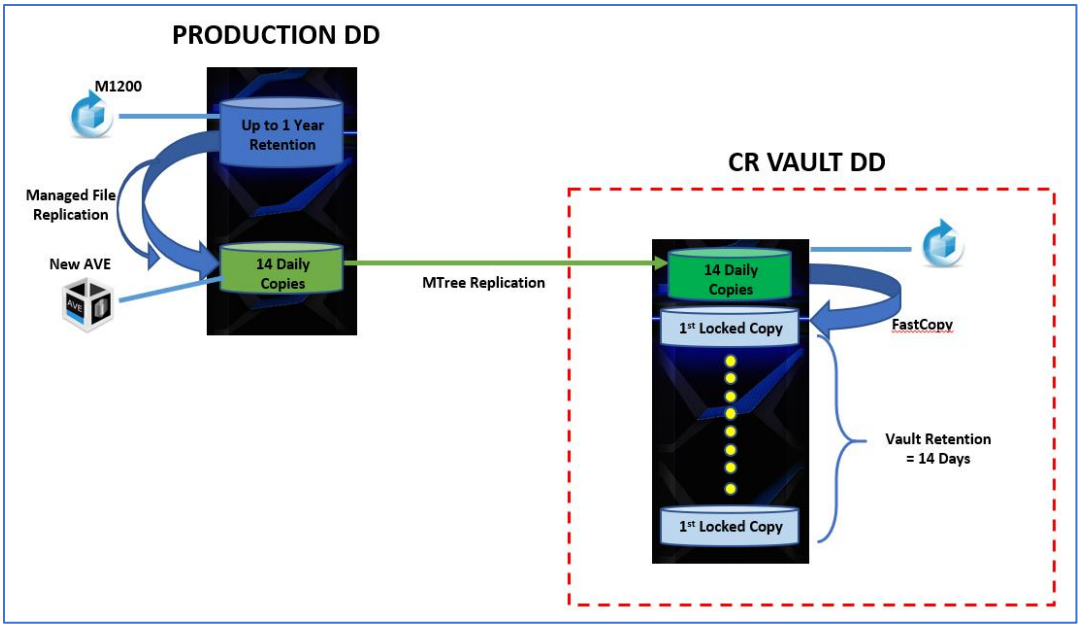


Figure 4

When dealing with larger retention in a Avamar / Data Domain environment as shown in Figure 4, we have a Data Domain and an M1200 that is attached to it in production. It represents the first MTree with all the active tier retention up to 105 days. It is very important to put what is critical in the vault so that we can make efficient use of this highly protected storage area. The workaround is the initiator that was mentioned earlier. In this case it is the new AVE in the production. The initiator can create a copy of the data with a much smaller retention than that of the 105-day production MTree, more reflective of what our customer wants which is just 14 days in the vault. So, we use Managed File Replication on the same production Data Domain and create a second MTree and associate it with the initiator. While this new copy coexists on the production Data Domain, it consumes minimal extra space because of the global deduplication happening on the Data Domain. We replicate data into the initiator and the new MTree would be a subset of the production and we change the active tier retention from 105 days to 14 days to match the vault. This way the whole system would be the right size. We can use the AVE node as the replication target to shrink the production MTree by controlling the number of hosts and effective retention inside each MTree. Controlling these two factors can dramatically lower capacity requirements.

What starts happening now is that the MTree replication is kicked off by the Cyber Recovery software and the MTree is copied into the vault on a set schedule where ports are activated and de-activated and retention lock applied and orchestrated by the Cyber Recovery software.

One of the ways to achieve this i.e., retain just the 14 days' worth of data and size the right Data Domain for this data in the Solution Builder would be to use values as shown in Figure 5.

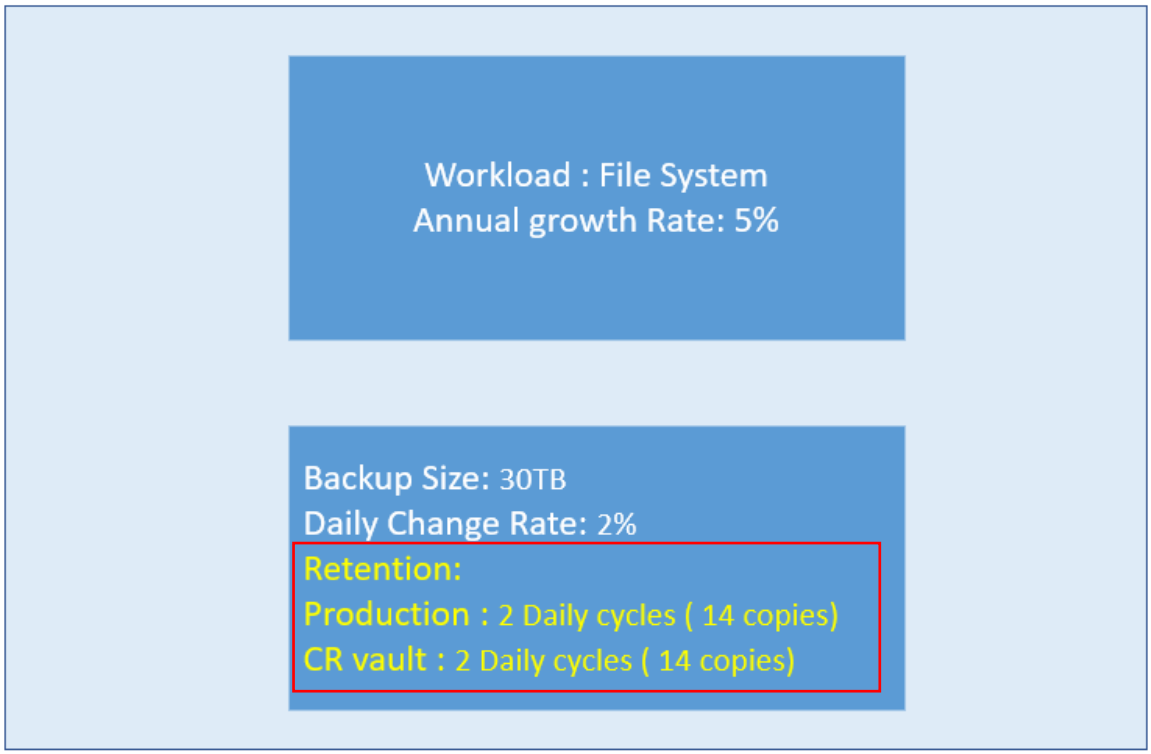


Figure 5 (Solution Builder inputs)

What we are trying to do here is that we clone the Production Site to create a Cyber Recovery Vault and edit the retention in the new site as per our requirements as shown above.

It is important to identify what data is critical for our customers and build Cyber Recovery Vault only for that. If we were to send the entire data to the vault without considering the replication context that is cloning and creating another site for only mission critical data with required retention, the vault Data Domain capacity would be somewhere between 50-60TB for the above example.

However, since we will create a separate MTree to reflect the required retention, the resulting Data Domain capacity will be 20.94TB

Figure 6

This example we've considered has a small amount of data and an entry level Data Domain. However, for environments with 100s of terabytes of data, the storage space savings would be tremendous.

Other Important Considerations

- The Data Domain systems must be running DD OS 6.0.2.20 and later
- The Cyber Recovery software does not support Cloud Disaster Recovery, Cloud Tier (inside the Vault) or Data Domain High Availability
- Neither the DP5300 nor the DP5800 are supported as targets in the vault although they are supported in production as source
- At the time of this writing, Avamar Multi-Node Grids, or RAIN configurations, unlike Avamar Single Nodes, do not have the ability to write a usable checkpoint into the associated Data Domain's MTree and therefore single node initiators will be required for these systems in production
- DP4400 is supported in both the production and vault side although some features of the DP4400 in the target would have to be disabled for e.g., Retention Lock Compliance.
- Avamar cascaded replication is not supported for data intended to be stored in the Cyber Recovery Vault
- An Air Gap on the data path between the production Data Domain and the vault Data Domain is mandatory for the solution to be deemed as a Cyber Recovery vault
- Cyber Recovery and Disaster Recovery cannot be shared on the same Data Domain unit.
- It is important to request that a VMware assessment be run on the customer environment so that we can identify and eliminate the white spaces, consequently sizing the right Data Domain.
- Ensure that every Cyber Recovery solution is validated by the Backup and Recovery Design Center (BRDC). BRDC can also assist in designing solutions. (Refer to this link for the same: https://inside.dell.com/docs/DOC-48562)

Conclusion

The number of bad actors attempting and succeeding in compromising business networks is increasing every year. This trend is not expected to subside. The only way to truly protect data is to disconnect it from the network. The good news is that the Dell Technologies Cyber Recovery Solution does just that and protects our customers from this reality by creating and protecting an offline, encrypted backup of data and orchestrating network availability. Only Dell Technologies provide a true air-gapped backup copy. The solution can be paired with early intrusion detection and scanning capabilities provided by Cyber Sense scanning inside the vault. The intention of this article is to size a near-perfect Cyber Recovery Vault Data Domain in the Avamar/Data Domain environment.

References

Sizing examples discussed in this article and some of the important considerations mentioned above have been reviewed by the BRDC team.

https://www.delltechnologies.com/en-us/video-collateral/demos/microsites/mediaplayer-video/2018/isolated-recovery-services-video.htm

https://inside.dell.com/docs/DOC-464147

https://inside.dell.com/docs/DOC-473010

https://h-isac.org/joint-cybersecurity-advisory/

https://www.delltechnologies.com/asset/en-us/products/data-protection/briefs-summaries/isolated-recovery-solution-overview.pdf

https://dl.dell.com/content/docu103173

https://www.delltechnologies.com/asset/en-us/products/data-protection/technical-support/h17670-cyber-recovery-sg.pdf