Avamar 7.4 Technical Differences: Data Domain Integration Enhancements

Welcome to Avamar 7.4 Technical Differences: Data Domain Integration Enhancements.

Copyright ©2016 EMC Corporation. All Rights Reserved. Published in the USA. EMC believes the information in this publication is accurate as of its publication date. The information is subject to change without notice.

THE INFORMATION IN THIS PUBLICATION IS PROVIDED “AS IS.” EMC CORPORATION MAKES NO REPRESENTATIONS OR WARRANTIES OF ANY KIND WITH RESPECT TO THE INFORMATION IN THIS PUBLICATION, AND SPECIFICALLY DISCLAIMS IMPLIED WARRANTIES OF MERCHANTABILITY OR FITNESS FOR A PARTICULAR PURPOSE.

Use, copying, and distribution of any EMC software described in this publication requires an applicable software license. The trademarks, logos, and service marks (collectively "Trademarks") appearing in this publication are the property of EMC Corporation and other parties. Nothing contained in this publication should be construed as granting any license or right to use any Trademark without the prior written permission of the party that owns the Trademark.

Revision Date: December 2016

Course Number: MR-7WN-AV74TDWW
This course covers Avamar 7.4 enhancements when integrated with Data Domain. This includes replication enhancements and using Avamar with Data Domain Cloud Tier.
Module: Enhancements with Data Domain

Upon completion of this module, you should be able to:

- Describe enhancements to Instant Access
- Identify range replication and its benefits
- Recognize recipe replication and its benefits

This module focuses on various enhancements with Data Domain and how Avamar can take advantage of them.
Lesson: Compatibility and Increased Instant Access

This lesson covers the following topics:

• Data Domain compatibility with Avamar 7.4
• Increased instant access sessions

This lesson covers Data Domain compatibility and the increased number of Instant Access sessions.
As before, Avamar can use Data Domain for back-end storage. For Avamar 7.4, this requires that the Data Domain be at version 5.7, or 6.0. Data Domain Virtual Edition 3.0 is also supported, since it runs 6.0 code. As usual, DD Boost must be licensed on these servers.
One improvement with Avamar 7.4 is an increased number of Instant Access instances.

The Instant Access feature (IA) in pre-7.4 releases supported only one IA instance at a time.

In the Avamar 7.4 release, up to 32 IA instances are supported at a time. This is achieved by utilizing the FSS feature of Avamar as well as some random I/O performance improvements in both Avamar and the Data Domain system.
The prerequisites for these enhancements include the following:

- Avamar server 7.4
- Data Domain system 6.0 (SSD is required to get better performance)
- A compatible vCenter with Avamar proxy deployed
Currently, the Instant Access feature is available as an option in the Avamar Administrator’s Backup and Restore menu. This IA option can be seen if a VM backup has been stored on a Data Domain system. In pre-7.4 releases, only 1 instance at a time is supported. Now, up to 32 instances are supported.
With the Instant Access feature, users can quickly access and restore virtual machines within a few minutes. This can be achieved by the use of an NFS Share. On the Data Domain system, the CUR folder is where all the backups are stored. In addition, there is a new folder called VMDS. This folder is created to store the images of the virtual machines during Instant Access. When IA triggers, a new folder named Avamar-MOD followed by an ID, is created, which is used by NFS.

On the vCenter, the NFS Share is mounted and the virtual machine’s data files can be accessed.
As best practices for this feature, it is recommended to change the NFS.MaxVolumes values on the ESXi server to a value greater than 32. This is because each IA creates an NFS volume on the server. Within vCenter, the setting is under ‘Configuration’->‘Advanced Settings’ of an ESXi server. By default, the value is set to 8.

Another recommendation is to perform a post-restore migration and clean-up after a system restore. This is because Avamar resources need to be released for later IA operations. This also protects the virtual machines’ data. vMotion can be used to migrate a virtual machine from one datastore to another.
If there is a NFS or IA failure, administrators can check the logs for details. Administrators can also check for failures in the activity log.

According to the IA workflow, it is recommended to check the NFS share data first, followed by the DD System, and then finally vCenter.
Please refer to the Avamar 7.4 for VMware User Guide for more details.
Lesson: Replication Enhancements

This lesson covers the following topics:

- Data Domain Range and Recipe Replication

This lesson covers replication performance enhancements when using Avamar with Data Domain.
To describe the Avamar 7.4 Range Replication feature, you'll need to be able to explain some of the features of existing Avamar replication technologies; Virtual Synthetic Replication (VSR) and Ninja Chop Replication (NCR).

Virtual Synthetic Replication (VSR) is an existing replication technology that uses base-file relationships to ensure only new or modified data is copied to the destination.

VSR uses a base-file relationship to compare the copy on the destination to the copy on the source. This ensures only new data is replicated. This, of course, improves replication performance for VS workloads.

Ninja Chop replication (NCR) is another existing replication technology that chops (divides) files into smaller file segments. One of the advantages of this replication technology is that it uses multiple streams to transfer data from the source Data Domain system to the destination. Using multiple streams reduces the replication time for large files.

Unfortunately, chopping files removes the base-file relationship, so it cannot be used with VSR. This means that VSR and NCR are mutually exclusive.
Range replication is a new technology introduced in Avamar Software version 7.4 to replace NCR.

Like NCR, range replication uses multiple streams to transfer data to the destination Data Domain system. Of course, using multiple streams shortens the replication time. It also has the ability to maintain the base-file relationship, so it can be used in conjunction with VSR. The use of VSR reduces the amount of data sent over the network.

After range replication transfers the data and the systems establish a base-file relationship, VSR is used for subsequent data transfers.
Let’s take a look at the existing Ninja Chop Replication. Its architecture is setup to follow this process.

First, the source Data Domain system identifies the file to be replicated.

Next, the source system chops the file into segments. The system uses memory and CPU resources to create these segments. Chopping the file prohibits the use of Virtual Synthetic replication because base-file relationships between the source and destination are lost.

Then, the source system transfers the segments using Managed File Replication (MFR) across multiple streams. When compared to using a single stream, multiple streams increase the efficiency of the transfer and reduces the duration of the overall replication.

On the target device, the destination Data Domain system receives the file segments.

Finally, the destination system reassembles the segments to recreate the file.
The new Range Replication architecture follows this process to replicate files.

First, the source Data Domain system identifies the file to be replicated.

Next, the source system divides the file into regions. It does not chop the file or copy file segments so it is much faster than NCR.

Then, the source system transfers the segments using Virtual Synthetic Replication (VSR) across multiple streams. Since VSR is used, the base file relationship is maintained. Only the new data needs to be transferred.

On the target device, the destination Data Domain system receives the incoming data.

Finally, the destination system reassembles the segments to recreate the file.
There are a couple of prerequisites for using the range replication feature.  
First, the Avamar Server must be running software version 7.4 or above.  
Second, the source and target Data Domain systems must be running DDOS 6.0 or above.
The system determines which replication method to use by examining two criteria; the availability of range replication and associated base files on all systems.

If all base files are available, then the system uses Virtual Synthetic Replication (VSR) regardless of the software version.

If all base files are not available, the system must run either NCR or range replication.

If range replication is available, the system will use it. This would typically be on an Avamar system running software version 7.4 or above.

If range replication is not available, the system uses NCR. This would typically be the case for systems running Avamar software version 7.3 or below.
You can configure the Avamar system to use either range replication or VSR. To do so, you must manually configure the ddr-repl-method-control setting through the Avamar Administrator interface.

First, navigate to the replication section of Avamar Administrator user interface and select the replication group you wish to edit. Next select the edit icon from the top menu.

When the Edit Replication Group dialog box appears, select Overview from the list of steps. Then, select the More Options button.

When the More Options dialog box appears click the More button.

Enter ddr-repl-method-control in the Enter Attribute field.

In the Enter Attribute Value field, enter zero (0) or ten (10). A value of 0 forces the system to use Range Replication. A value of 10 forces the system to use VSR.

Select the plus (+) symbol to add the attribute.

Select the OK button.

When the Edit Replication Group dialog box appears, select the Finish button.
You can check the replication log file to determine if the configuration is forcing the replication method to be range replication or VSR.

The replication log files are stored in the /usr/local/avamar/var/client directory on the Avamar utility or single node. The replication log file name ends with the string Replicate-avtar.log. Example log entries are shown on the screen.

Search the log file for the string ddr-repl-method-control. The presence of this string indicates the system's replication selection logic is being circumvented and either range replication or VSR is being forced. If the value is zero (0), the system is configured to force range replication. If the value is ten (10) the system is configured to force VSR.
Check Forced Settings in Log File (2 of 2)

- **Range Replication**
  2016-06-27 00:24:38 avtar Info <10609>: Data Domain Destination login to griffin-dd10.asl.lab.emc.com complete
  2016-06-27 00:24:39 avtar Info <0000>: id:1 Range-Replication selected as per user specified control method bits 0 (container.1.cdsf)
  2016-06-27 00:24:39 avtar Info <40160>: - Establishing a connection to the Data Domain system with basic authentication (Connection mode: A:0 E:0).

- **Virtual Synthetic Replication (VSR)**
  2016-06-27 00:55:19 avtar Info <10609>: Data Domain Destination login to griffin-dd10.asl.lab.emc.com complete
  2016-06-27 00:55:20 avtar Info <40193>: id:1 VSR selected as per user specified control method bits 10 (container.1.cdsf)
  2016-06-27 00:55:20 avtar Info <40160>: - Establishing a connection to the Data Domain system with basic authentication (Connection mode: A:0 E:0).

```
/usr/local/avamar/var/client/MOD-XXXX#X-XXXX-Replicate-avtar.log
```

When the replication process starts, an entry is written to the replication log if range replication or VSR is being forced. Example log entries are shown on the screen.
You can disable the use of range replication on the Avamar system. To do so, you must manually configure the private, support-only x29 setting through the Avamar Administrator interface.

First, navigate to the replication section of Avamar Administrator user interface and select the replication group you wish to edit. Next select the edit icon from the top menu.

When the Edit Replication Group dialog box appears, select Overview from the list of steps. Then, select the More Options button.

When the More Options dialog box appears click the More button.

In this next dialog box, enter x29 in the Enter Attribute field.

In the Enter Attribute Value field, enter 0x20000000 - that's 0x2 followed by seven zeros.

Select the plus (+) symbol to add the attribute.

Select the OK button.

When the Edit Replication Group dialog box appears, select the Finish button.
Performance tests were run with the parameters shown on the screen. As you can see, Range replication produced reductions in the number of bytes transmitted and the amount of time elapsed.
Known Limitations

- Range replication not supported:
  - With Data Domain Extension Retention
  - Source file is Fixed Sized Segment (FSS) backup
- The system automatically uses MFR/VSR

Range replication is not supported if the Data Domain system is an Extension Retention unit or if the source file is a Fixed Sized Segment (FSS) backup. In this case, the system automatically uses MFR/VSR. When using Data Domain Extended Retention, the user should modify the `ddr-repl-method-control flag` to a value of 10 in order to ensure that the MFR/VSR method is used.
Data Domain 6.0 has also improved replication performance with its new recipe replication method. Recipe replication will be used for any Avamar data that is stored on a Data Domain system, so the benefits of this feature get passed on to Avamar servers that use replication.

Recipe replication improves the efficiency of replication by doing a better job at identifying data already at the target Data Domain system. This leads to less data being transferred over the network which leads to faster replication. Recipe replication accomplishes this by adjusting two factors: the size of region that is tracked and the number of base files that are tracked.
Previously, Data Domain would track changes in a file in regions that were multiples of 4 MB in size. This means that when a second backup with only a small change occurs, the entire 4 MB region with the change would be replicated. That region would then be merged with previous unmodified data to form the new backup.

With recipe replication, the minimum size of the region is now 16 KB. This allows for a greater granularity in detecting small changes between one backup and another. If only a small amount of data has changed between one Avamar backup and another, then at a minimum, only 16 KB need to be transferred.
Recipe Replication also improves efficiency by increasing the number of base files that can be tracked.

With previous versions, Data Domain would track the changes of up to 8 files. Consider a SQL client with 9 databases. Only 8 of them would be able to be tracked as base files when replicated to the target Data Domain system while the ninth would not. When the Data Domain system replicates a second backup with a few changes, only the first 8 databases would be able to use the existing data at the target as a base and only their changes would be replicated. But the 9th database would have to be replicated every time, since only 8 base files could be tracked.

With recipe replication, the number of base files tracked has increased to 256. Since more base file are tracked, less data needs to be transferred. This greatly improves the amount of data transferred.
Since recipe replication is a DD OS 6.0 feature, it required to have Data Domain OS version 6.0 at both the source and target sites. The feature is enabled by default. Also, the Avamar servers at both sites must be version 7.2 or above, since that is the minimum Avamar version for DDOS 6.0.
Module Summary

Key points covered in this module:

- Enhancements to Instant Access
- Benefits of range replication and recipe replication

This module covered various enhancements with Data Domain and how Avamar can take advantage of them.
Module: Avamar with Data Domain Cloud Tier

Upon completion of this module, you should be able to:

- Describe Data Domain Cloud Tier
- Explain how Avamar uses DD Cloud Tier
- Describe how to configure and use Avamar with DD Cloud Tier

This module focuses on Avamar’s integration with Data Domain Cloud Tier.
Lesson: Data Domain Cloud Tier Overview

This lesson covers the following topics:

- DD Cloud Tier architecture and functions

This lesson covers an overview of Data Domain Cloud Tier.
Avamar 7.4 introduced support for Data Domain Cloud Tier. Before describing how Avamar uses DD Cloud Tier, let’s first examine how DD Cloud Tier works on its own.

Data Domain 6.0 introduced DD Cloud Tier to allow data to be automatically moved from the Data Domain itself onto lower cost cloud storage for long term retention. This creates a two tiered system with the Data Domain system as the Active Tier, and cloud based storage as the Cloud Tier. The Active Tier is local and is used for immediate access, while the Cloud Tier is used for long term retention. Data is first stored on the local Data Domain for quick access. Once the data has passed the configured age limit, the Data Domain system moves it onto the cloud tier. This frees up storage space on the primary Data Domain storage. The archived data can still be accessed by first bringing it back to the active tier. Once on the active tier, the client or application can read the data again.
Before using DD Cloud Tier, you must create at least one Cloud Unit on the Data Domain. A Cloud Unit represents storage on a cloud provider. Cloud units can be configured for EMC Elastic Cloud Storage, Virtustream, or Amazon Web Services. When a Cloud Unit is created on a Data Domain system, it not only stores access credentials to the cloud provider, but it also reserves space on the local Data Domain system to store metadata for any files stored on the cloud. This metadata allows the files to be quickly browsed without having to perform slower access calls to the offsite cloud storage.
Within Data Domain, cloud tiering policies define what data gets moved to the cloud and to what Cloud Unit. Policies are defined per MTree. The user selects an age that a file must be. Any file within the configured age range becomes eligible to be moved to the configured Cloud Unit.

Data can be moved manually, or, more often, it can be configured to be moved automatically on a scheduled basis.
Recalling Data from the Cloud Tier

It is important to note that data that has been moved to the cloud tier cannot be accessed directly. Either a Data Domain user, or an application using the Data Domain, must recall the data back to the Active Tier. Only then can the data be read again by a user or application.
This lesson covers an overview of how Avamar interacts with DD Cloud Tier.
If an Avamar 7.4 server is configured to use Data Domain for backup storage, it can take advantage of DD Cloud Tier to move backup data to the cloud. This allows for longer retention periods by moving older backups that are less likely to be accessed to cheaper cloud storage. The tiering process is configured to automatically select backups and move them to the cloud tier on a scheduled basis. Also, restoring data from a backup on the cloud tier is seamless to the user. There are no changes to the restore procedure.
DD Cloud Tiering with Avamar is driven by user configured Tier Groups. A Tier Group is a new item configured within Avamar. It defines which backups will be eligible to be moved to the cloud, where they will be moved to, and how often to move data to the tier.

Backups are selected based on the client that they originated from. The backups can be further filtered based on their retention tags and the age of the backup. You may also limit the number of backups moved per client. For example, you may choose to send no more than 5 backups to the cloud tier.

The tier group specifies which Data Domain cloud unit will be used as the cloud tier. Any backups selected by the tier groups filter will be sent to the configured cloud unit.

Tier Groups will run according to an Avamar Schedule.
Let’s take a look at how Avamar works with Data Domain to send data to the Cloud Tier.

At the beginning of the process, Avamar is storing backups on the Data Domain system, or active tier. The Avamar metadata for these backups tracks their tier status as “Active” meaning that the backup data is stored on the active tier.

Once its scheduled time arrives, the Tier Group on the Avamar server will determine which backups are eligible to be moved according to its settings. The Avamar will mark the backup data on the Data Domain system as eligible to be moved. During the mark process, the backup will show a tier status of “Indeterminate”, but once marked, the status will change to “Marked.”

Avamar does not initiate the data movement to the cloud. Instead, it waits until the time scheduled by the Data Domain system’s movement policy. Once this scheduled time arrives, the Data Domain system moves the data to the configured cloud unit.

Notice that the Avamar still shows the backup’s tier status as “Marked.” The Data Domain system does not notify the Avamar that the data has been moved. However, during the Avamar server’s next garbage collection process, the Avamar metadata is synchronized with the data moved on the cloud, and the backup’s tier status is changed to “Cloud.” Attempting to access the backup before garbage collection would also cause this synchronization to happen.

At this point, the backup data has been moved to the cloud tier, and the Avamar server is aware of its location on the cloud.
It is important to realize that there are two schedules involved in the process. The schedule on the Avamar Tier Group determines how often backups are selected and marked. The Data Domain system’s data movement schedule determines when those marked backups are actually moved to the cloud.

These two schedules should be set to run at similar intervals so that marked backups do not wait too long before being moved. If the Data Domain system’s schedule does not run often enough, the Avamar Tier Group will be marking many backups without them actually being moved right away. When the Data Domain system’s schedule finally does run, there will be many backups marked for movement. The Data movement will be very large and slow.
Restoring data from the cloud tier is a two step process. First the data is recalled to the active tier, then it is restored back to the client. These two steps happen automatically when the user performs a restore, so the user experience is unchanged.

When a user performs a restore, the Avamar server checks to see what the tier status of the backup is. If it is located on the cloud tier, then Avamar instructs the Data Domain system to recall the backup data from the cloud back to the active tier. Note that the entire backup will be recalled, even if only one file from the backup was selected for restore.

While the data is being transferred back, the backup’s tier status is “Indeterminate.” Once the transfer to the active tier completes, the backup’s tier status is “Active.” Since the backup data is now on the Data Domain system, Avamar performs the restore in the same manner as a regular restore.
To use Avamar with DD Cloud Tiering, you must have Avamar version 7.4. Data Domain OS version 6.0 or higher is needed, since that is the version that introduced cloud tiering. Both physical or virtual Avamar server can be used, and both physical and virtual editions of Data Domain can be used.
This lesson covers the process of configuring and using Avamar with DD Cloud Tier.
Prior to configuring Cloud Tier for Avamar, the Data Domain system needs to be properly configured for the cloud. First, enable the DD Cloud Tier feature. Also, create a cloud unit on the Data Domain system in order to provide connection information to the cloud provider. Be sure that the Data Domain system is configured to automatically move data to the cloud unit on a scheduled basis. More information on configuring DD Cloud Tiering on a Data Domain system can be found in the *Data Domain Operating System Administration Guide*. 
On the Avamar server, edit the Data Domain system from the Server window in Avamar Administrator. Under the new **Tiering** tab, enable Cloud Tier, and select the desired cloud unit that will be used as the cloud tier. The list of cloud units is automatically populated from the Data Domain settings.
Next, create a tier group. Navigate to the Replication/Tiering window and select Actions> New Group> Tier. This will open the New Tier Group wizard. Provide a name for the tier group. In the source tab, select which clients will have their backups moved to cloud storage. You may also add filters to only send certain backups to the cloud. Here you can filter based on the backup tag. In most cases, you will want to use a date restriction to move only older backups to the cloud. Data Domain requires that any data be stored on its active tier for a minimum of 14 days before being moved to the cloud tier. Therefore, this setting must be set to at least 14 days. You may also limit the number of backups per client that are moved.

Next, select the destination Data Domain system. When this is selected, the cloud unit associated with that Data Domain will be selected for use.

The expiration of the backups cannot be modified. Next, select or create a schedule for the tier group. By default, the tier group runs once a day at midnight.

You are given an opportunity to confirm all the choices and create the tier group.
Once the tier group has been created, it will run automatically according to its schedule.

Remember when a tier group runs, it only marks backups as eligible for movement. The actual movement to the cloud will not occur until the time scheduled on the Data Domain system.
Performing a restore from a backup that has been moved to the cloud tier is no different than performing a regular restore. Select the backup from the restore interface. You may notice its tier status in the GUI. A tier status of “Active” means that it is on the local Data Domain system. A status of “Cloud” means it has been moved to the cloud. A status of “Marked” means that it has been marked as eligible for movement and may or may not have already been moved to the cloud.

When performing the restore, you may be reminded that restoring from the cloud tier will cause the entire backup to be recalled back to the active tier, even if you are only restoring a single file. Since some cloud providers may charge a transfer cost or impose limits to the amount of data transferred out, it is important to consider the full size of the backup.

The backup data is only stored on the active tier temporarily. Once the tier group runs again, the data will be eligible to be moved back to the cloud tier.
Tiering operations will appear in the Activity Monitor. Whenever a tier group runs, it will show in the activity monitor as a “Tier” operation. When performing a restore of data from data on the cloud tier, the activity monitor will show both steps of the process. A “Recall” operation is listed when the data is recalled to the active tier, and a “Restore” operation shows the restoration of data from the Data Domain system to the client.
Troubleshooting and Logs

MCS Log:
/usr/local/avamar/var/mc/server_log/mcserver.log.X

AvTier Logs:
/usr/local/avamar/var/client/*Tiering.log

ddrmaint Logs:
/usr/local/avamar/var/ddrmaintlogs/ddrmaint.log

As usual, you may double click the entries in the Activity monitor to open the relevant logs. You may also find information on tiering in various logs on the Avamar server: the mcserver log, the tiering.log files, and the ddrmaint log files, as shown.
When using the Cloud Tier feature, be sure that data is replicated before it is moved to the cloud tier. Replication of data must be performed from the active tier. If a backup is on the cloud tier when Avamar attempts to replicate it, the backup will first be recalled back to the active tier. This uses unnecessary resources. To ensure that backups are replicated before being moved to the cloud, be sure that replication runs every day and that tier groups are configured to allow backups to remain on the active tier for a few days before being moved to the cloud tier.

Also, when creating tier groups, do not use very infrequent schedules. If the tier group does not run on a regular basis, then the amount of backups that will need to be moved to the tier becomes very large. This results in large amounts of data being moved to the cloud at once.
Currently, there are a few limitations with this feature. Unlike backups and restores, you cannot cancel the tier and recall operations from the Activity monitor.

Also, file level restore and instant access for VMware virtual machine backups, and granular level recovery for various databases and applications, are not supported when the backup is on the cloud tier. The backup must already be located on the active tier in order for these activities to take place since they require the backup data to be accessed using protocols that cloud storage does not support. Therefore, it is important to keep in mind that these features will not be available for older backups that have been sent to the cloud until that backup has been recalled to the active tier.
Web Object
Address:
https://edutube.emc.com/Player.aspx?
vno=ulbt+nbx7muiX7gAwRSLGw==&autoplay=true

Load time may vary depending upon bandwidth available.
This video will be available in eduTube at product GA.
This module covered Avamar’s integration with Data Domain Cloud Tier.
Course Summary

Key points covered in this course:

- Improvements with Avamar integration with Data Domain
- Replication enhancements
- Avamar integration with Data Domain Cloud Tier

This course covered Avamar 7.4 enhancements when integrated with Data Domain.

This concludes the training.