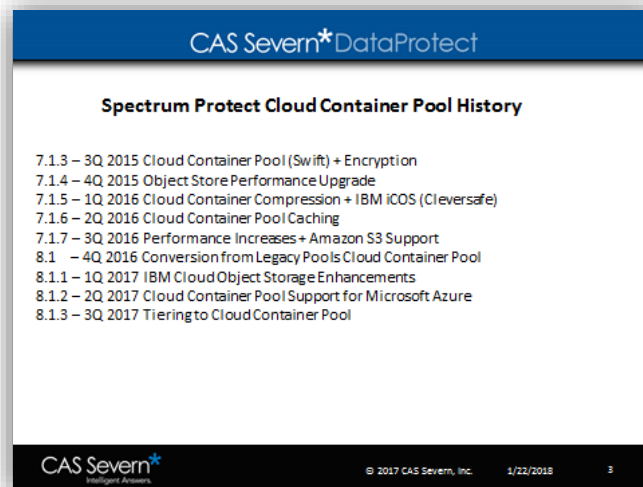




Click here to see the video: <http://bit.ly/2E2MN6h>

Introduction: Welcome. My name is Joseph King. I'm the Chief Technology Officer at CAS Severn, and today's topic is Spectrum Protect cloud tiering to Amazon S3. We're going to walk you through how to leverage this new capability in Spectrum Protect.

00:17 This capability came out in the third quarter of 2017 in the 8.1.3 release and the specific name of the capability is “tiering to a cloud container pool.”



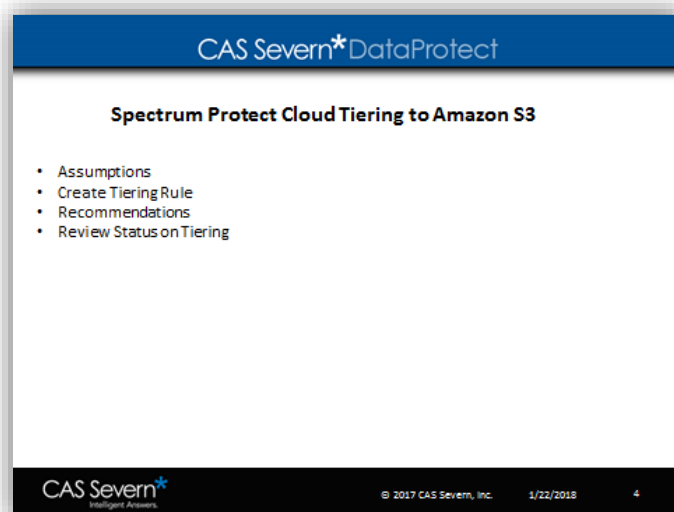
00:30 We'll walk through some assumptions, we'll create the cloud tiering role which is straightforward. We'll spend most of the time talking about recommendations and talk a little bit about how to review your status on the tiering operations. Your assumptions are you have to be at 8.1.3 or above, you have to have a functional directory container pool, you have to have a functional cloud container pool.

00:54 We'll just briefly walk through the steps here, if you want to more details, see our previous video at this link:

Creating a Spectrum Protect Cloud Container Pool on Amazon S3

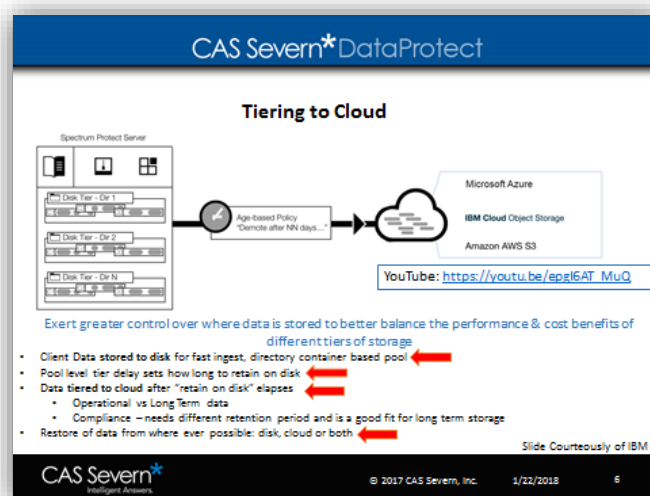
https://www.youtube.com/watch?v=tJLtasjIG_8&t=2s

For tiering to the cloud, this is specifically tiering a local directory container pool to a cloud container pool.



01:10 So this is the ability to take data that you want to be able to access quickly, and have it locally. We'll have any data that is less likely to be accessed, and have that stored in a cloud without you as an administrator needing to worry about where it is. Or, your users doing a recovery having to worry about where the data is. The data is tiered to the cloud after a retention period is hit.

01:37 This allows you to split your operational versus long-term data. We'll go into some specifics on the type of data that is appropriate moving forward, but really this should be used for compliance data. It's the data that needs a different retention period and it's a good fit for long-term storage. In the end when you're doing a recovery, it'll do the recovery from wherever it's possible either from the local directory container pool, the cloud, or both.

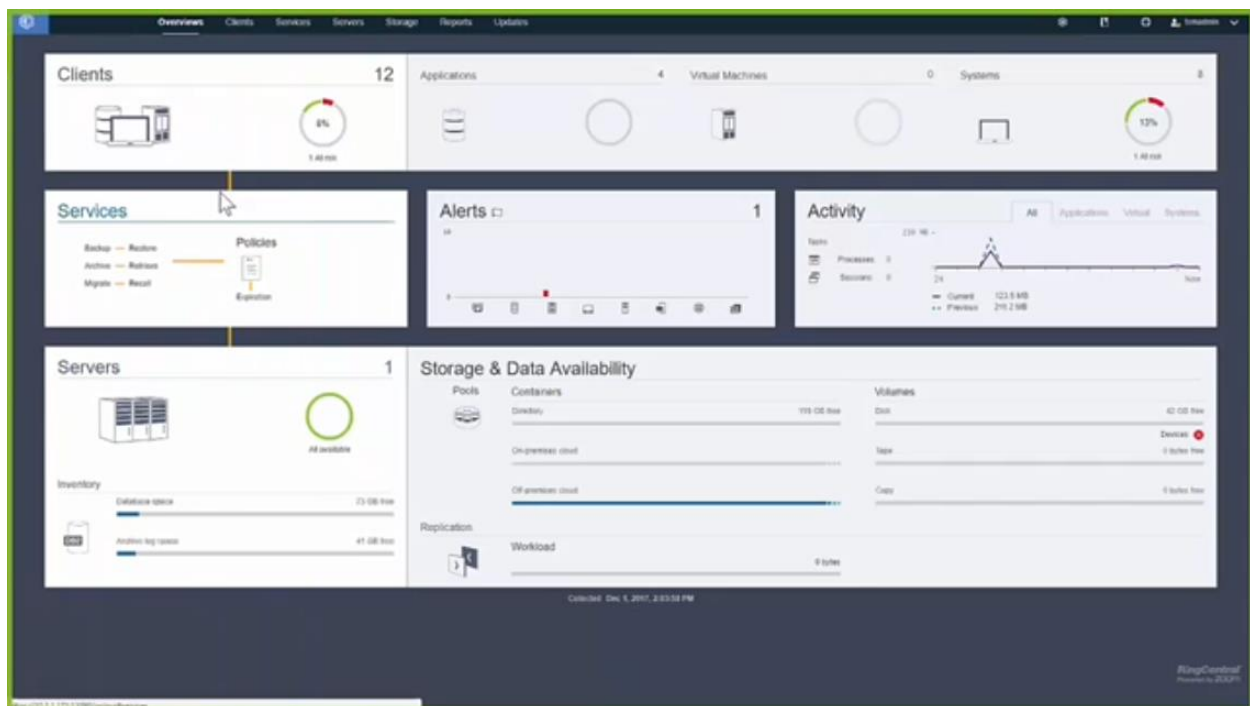


02:09 So briefly, you need to create a cloud tiering rule and this is used doing the STG rule, rule name in the command line. Or, we're going to show you how to do it in the Operations Center and there's some things that you need to keep in mind.

02:29 The first thing you're going to have to do is create a cloud tiering rule. This is going to be done either in a command line or in the Operations Center, and you need to set a few variables. Mainly how many days the object will stay in the source directory container pool before being tiered, how many processes you want to run simultaneously for each source storage pool, how long do you want that process to run and when do you want that process to start.

03:04 Once you do that, then when the start time is hit every day it will run the job for the duration that you set for, up to the maximum number of processes that you set. It's important to know that the cloud tiering action is between a local directory container pool and a cloud container pool. Individual nodes will participate if their data is being stored to a particular directory container pull that's in scope.

03:35 So let's switch over to the Operations Center and look at creating a cloud container pool.



So here you'll see the main screen of the Operations Center, and we're going to look at the storage pools that we have defined.

Type	Name	Server	Status	Capacity Used	Device Class	Container Type	% Savings
Primary	ARCHEVPOOL	RHEL7SP	Normal	No capacity	DISK	---	---
Primary	BACKUPPOOL	RHEL7SP	Normal	No capacity	DISK	---	---
Container	DEDUPPOOL	RHEL7SP	Normal	1.0B / 100.0 GB	---	Directory	55
Primary	SPACESPOOL	RHEL7SP	Normal	No capacity	DISK	---	---
Container	AMAZONSPPOOL	RHEL7SP	Normal	53.0 GB	---	OP-presses cloud	48
Primary	RM_DEPLOY_CLIENT	RHEL7SP	Normal	1.0B / 40.0 GB	RM_DEPLOY_CLIENT	---	---
Container	TIERPOOL	RHEL7SP	Normal	0.0B / 47.0 GB	---	Directory	9

03:49 You can see that we already have a cloud container pool created. This one is called Amazon S3 pool. And, we have a local directory pool called tiered pool that we're going to use as our local directory pool that we are going to tier to the Amazon S3 pool. If we go and take a look here, you go up to the storage tab down to tiering rules, you can see we already have an existing rule in place. Before we take a look at that existing rule, let's create a new rule.

04:25 Click on create rule, you'll see that the target pool is going to be a cloud container pool, in our case Amazon S3 pool, and then the source pools. You can see that the D pool here is greyed out, that pool is already defined in another roll. That's the one that we'll take a look at next.

04:47 So here we're going to check the tiered pool we're going to give it a rule name, and we'll call this test tier 2s3. We're going to move our data to our target pool. We'll say 15 days we're going to start it at 8 p.m., and we're not going to give it a limit on its run time. But if you needed to limit it in order to keep within your schedules, you could set that, and once it hits the maximum run time, that job will be turned off and we'll start again tiering the next day.

05:24 We're going to hit cancel here and we're going to go into the existing rule d-dupe to S3 rule, and take a look at the details. This d-dupe to S3 rule is going to move data to the target after 5 days, it's going to have a daily start time here of 10:36 a.m., and it's going to run until complete. Once you take a look at this after it has some data running to it after a few days, this window on the right is going to populate, and you're going to see whether that job was successful with a green checkmark or filled with a red X mark.

06:09 This is our test server so we're not expecting this to work every day. If you look at the lower left side you're can hover over and take a look at the transfer speed that you're getting over the last two weeks.

06:29 Let's look at some recommendations that you need to follow when you're using cloud tiering. The most important aspect when you're looking at cloud tiering is that objects are grouped together, and this is very important.

06:44 Let's walk through some examples. If you look at jobs that run at weekly full and daily incrementals, and this could be Spectrum Protect for databases.

07:01 An example of what a tiered delay for 4 days looks like.

The slide is titled "CAS Severn*DataProtect" and "Grouped Objects Together". It contains the following text:

- Cloud tiering treats groups as one object, for clients & data types that group objects together
 - Such as: Windows System State, data protection modules, etc
- Tier to cloud decision, is dependent on the eligibility of the oldest object in the group
 - Oldest object is typically the FULL backup
 - WATCH OUT : Spectrum Protect Virtual Environments / Incremental Forever (Single Full)

For Instance:

- Oracle backup using Spectrum Protect for Databases
- Doing weekly fulls + daily differentials
- TIERDELAY=4 days

Day 1 - Full backup	No tier action
Day 2 - Differential	No tier action
Day 3 - Differential	No tier action
Day 4 - Differential	Full + 3 Diffs (day 2, day 3, day 4) tiered to the cloud
Day 5 - Differential	Diff (day 5) tiered to the cloud
Day 6 - Differential	Diff (day 6) tiered to the cloud
Day 7 - Full backup	No tier action

Slide Courtesy of IBM

CAS Severn* Intelligent Answers. © 2017 CAS Severn, Inc. 12/1/2017 8

At the bottom of the slide you'll see a seven-day cycle where day 1 and day 7 have full backups, day two four days, six have differential backups, day 1, day 2 and day 3 have no tiering action, day 4 will tier not only the full, but the first three differentials. This is important because when you're doing your decision making on tiering the cloud, it is dependent on the eligibility of the oldest object in the group. So day 5 and day 6 will also tier to the cloud, again, because the oldest object in the group, which in this case would be the full backup of that database, is eligible.

07:51 So those differentials will still tier to cloud, and this is very important to understand for jobs that are set up using incremental forever. This could be jobs that include Spectrum Protect for virtual environments, or just a regular incremental forever set up with a backup archive client. So take note that cloud tiering will tier objects as soon as that first full backup is eligible.

08:25 For items that might never have a full back up, or might only have a full back up every month or two, this might not be the best type of data to use in cloud tiering. Our recommendations are to keep your data local to meet your likely recovery time objective.

08:47 For instance, if you're most likely going to recover your databases within the first two weeks of them being backed up, then you keep at least 15 days of that data local. That way, in our example here, that would be two weekly data sets are kept local and then anything from there would go to the cloud tier.

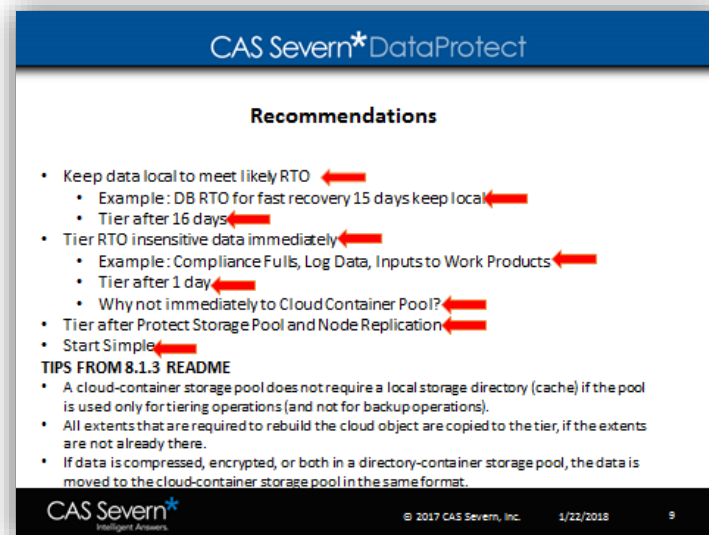
09:16 So in this example, we would tier after 16 days, and then tier recovery time objective in sensitive data to the cloud immediately. So these could be full backups that you're taking for compliance reasons, log data that has a high likelihood of never needing to be recovered, or say inputs to work products where it's the work product that you need to keep local.

09:41 But the inputs used to make that work product are no longer needed to have a short recovery time objective; in that case setting you're tiering to be 1 day would be appropriate. But if you take a look it, might just make more sense just to send that data directly to a cloud container pool using the local cache directory we talked about in our video:

Creating a Spectrum Protect Cloud Container Pool on Amazon S3

https://www.youtube.com/watch?v=tJLtasjIG_8&t=2s

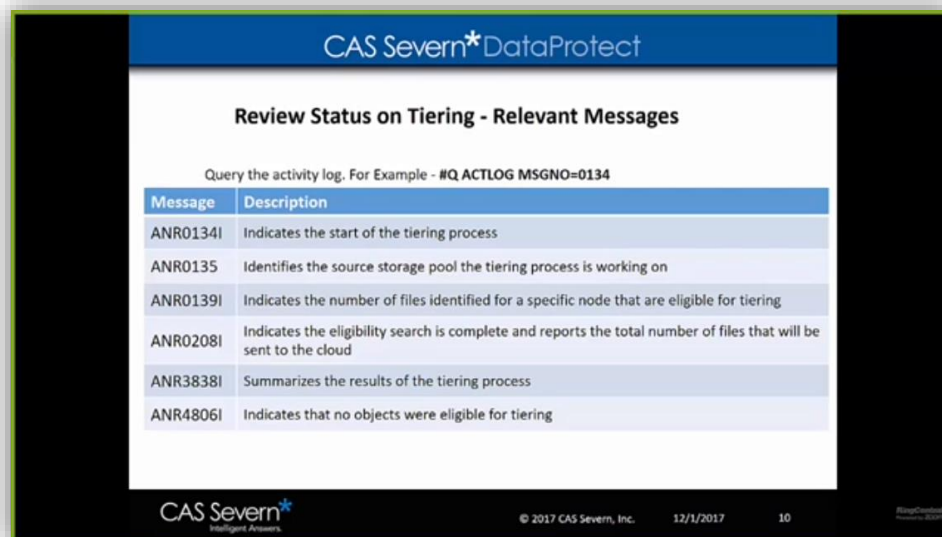
10:08 You want to time your tiering to occur after your protect storage pool and node replication jobs are done. And really, as we always recommend, start simple, understand what data will be sent to the cloud container pool in a tiering relationship, and what data won't.



10:29 There's three really good tips from the 8.1.3 readme that you need to be aware of. The one that I think is very important is if the data is compressed encrypted in the directory container pool, it will maintain that compression encryption in the cloud container pool. All extents that are required to rebuild the cloud object are copied to the tier, if the extents were not already there. So this is the idea we're again going back to our previous database example, why those day 5 and day 6 differentials were sent to the cloud tier.

11:10 All extents that are required to rebuild the cloud object are copied to that tier. There's some relevant messages they need to keep in mind when taking a look at the operation of a cloud container

pool and these individual messages will walk you through the different processes that are done behind the scenes in order to do the cloud tiering.



The screenshot shows the CAS Severn*DataProtect interface. At the top, it says 'CAS Severn*DataProtect'. Below that, the title is 'Review Status on Tiering - Relevant Messages'. Under the title, it says 'Query the activity log. For Example - #Q ACTLOG MSGNO=0134'. There is a table with two columns: 'Message' and 'Description'. The table contains the following rows:

Message	Description
ANR0134I	Indicates the start of the tiering process
ANR0135	Identifies the source storage pool the tiering process is working on
ANR0139I	Indicates the number of files identified for a specific node that are eligible for tiering
ANR0208I	Indicates the eligibility search is complete and reports the total number of files that will be sent to the cloud
ANR3838I	Summarizes the results of the tiering process
ANR4806I	Indicates that no objects were eligible for tiering

At the bottom of the interface, there is a footer with the CAS Severn* logo, the text '© 2017 CAS Severn, Inc.', the date '12/1/2017', and the page number '10'.

11:38 Look at message 134, it indicates the start of the securing process, 135 identifies the source 3rd storage pool, and the tiering process that it's working on. It's 139 and on work it's interesting. 139 is the object identification for a specific node. 208 just shows that the eligibility search is complete and it reports the total number of files that will be sent to the cloud. 3838 summarizes the results of the process.

12:09 Then if you see a message 4806, that indicates that no objects are eligible for tiering and that's really just an informational message. So let's take a look what that messages will look like when you query them. So let's go back to operations center and go to command. Take a look at that and you could see here that we're querying the accounting log beginning with today and looking back 10 days. We're looking at message 139 for a specific file space on a specific node. In this case, the node my surface and the C drive on that particular system, and we can see that here on 11/22 that 503 files were identified and then later on at 11/25, 528 files were identified.

13:13 If you go back and if you remember you saw that those errors were in place there and you can see that the number of files aren't incrementing, this is just showing that those files weren't necessarily sent to that cloud container pool.

13:29 If you take a look at your occupancy table, you can see the amount of data that is in each pool for this individual node. Take a look at my surface and we can see that that C Drive is split across two pools where the majority of the data has been tiered out to that Amazon S3 pool, and a little bit of that data, 121 files, is in the local directory container pool called DD pool. This is a way when you're taking a look at what the impact is to cloud tiering on an individual node, you can take a look at what data and how much is at each tier and this is why we recommend to start simple. To take a look at what is the impact of cloud tiering on your particular workload.

14:24 So to recap, when using cloud tiering the setup is very easy. It's run using administrative process that's typically going to run nightly. You set the time in the duration, it's going to send data from a directory container pool to a cloud container pool using delay that you set in the rule.

14:44 Then you're going to use these relevant Status Messages to take a look at its operation. We think that cloud tiering is a great way of shifting data that has a low likelihood of ever needing to be recovered, and shifting it to the cloud containers without having to do any special setup when it comes to identifying what data makes sense to send.

15:10 Just let the RTO be set appropriately on your nodes so the right data is sent to the right tier at the right time, taking into account how items are grouped together.

15:29 We thank you, if you have any questions feel free to email us at the email addresses included. We look forward to seeing you in our next video, take care.

To schedule a 1:1 with Joe and his technical team, contact us at:
sales@cassevern.com
800-252-4715