

The Restore Point Simulator

The hidden manual

Restore Point Simulator - The hidden manual

In this document, we will describe how you can use the restore point simulator and how it maps to Backup & Replication.

Please send all remarks you might have to @tdewin

Current version: Draft 0.0.1

Contents

Understanding the input parameters	4
Quick Presents.....	4
Style Parameter	4
Used Size GB.....	5
Retention Points.....	6
Change Rate	7
Data Reduction.....	8
Interval	8
Increment Specific.....	10
Reverse Specific.....	12
Backup Copy Job Specific	12
Understanding the result	13
Manual Run	15
URL Parameters.....	16

Understanding the input parameters

Quick Presents

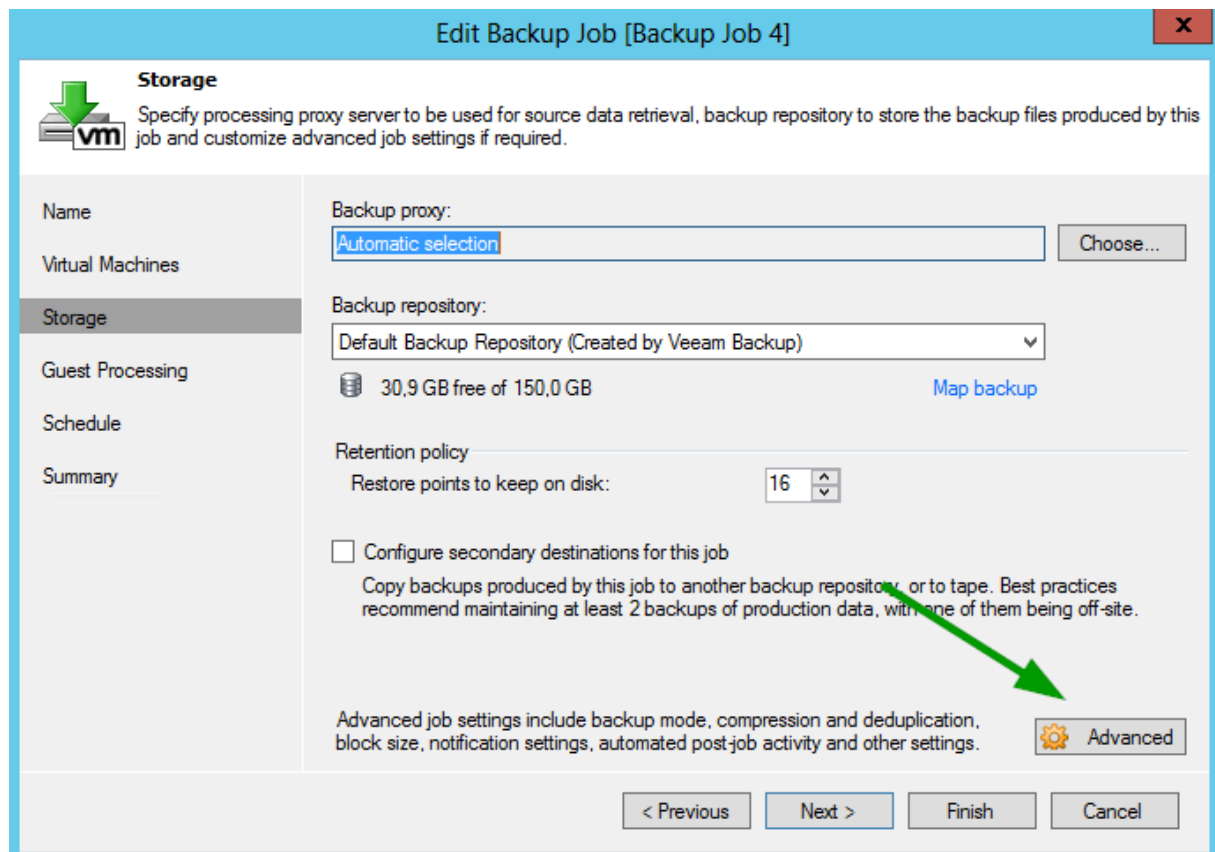
Quick Present allow you to quickly select predefined scenario's. This could be handy if you have an idea what you want to configure and quickly want to preconfigure the settings. Notice that you can also take a scenario close to what you like, and then adapt it to your specific needs.

For backup jobs, the default scenario in v8, would be "Forever Incremental" except for jobs defined before upgrading to v8. In v7, for such jobs, the default would be "Incremental weekly synthetic".

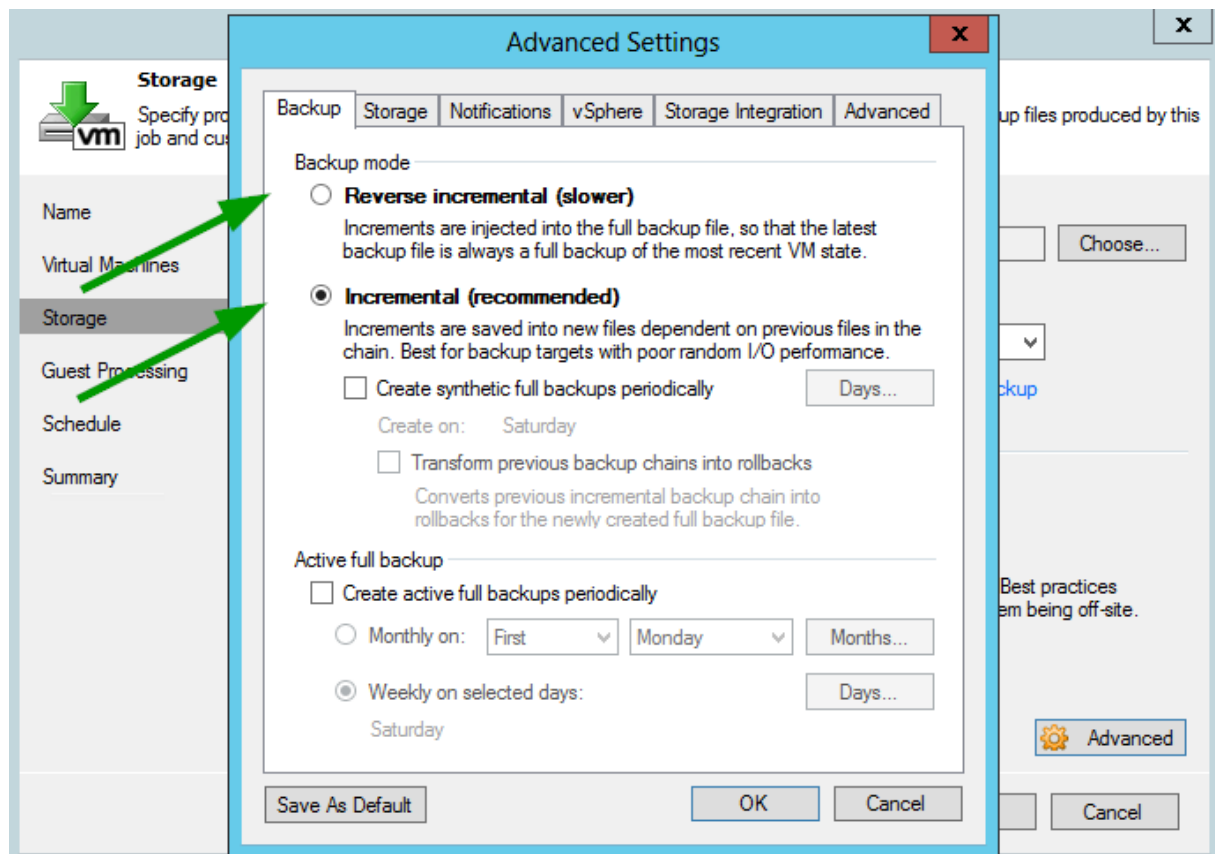
For backup copy jobs, the default scenario would be "Backup Copy Job"

Style Parameter

The style parameter corresponds to the job type you choose. "Backup Copy Job" of course matches the backup copy job. For regular jobs, you have 2 choose: Incremental or Reverse. These match directly with the GUI of backup & Replication. In the storage section you can select the "Advanced Button".



Here you can select the corresponding style



What is important, the new Forever Incremental style is not a separate option. Instead, use Incremental, uncheck “Create synthetic full backups periodically” and uncheck “Create active full backups periodically”. This configuration is not possible in v7.

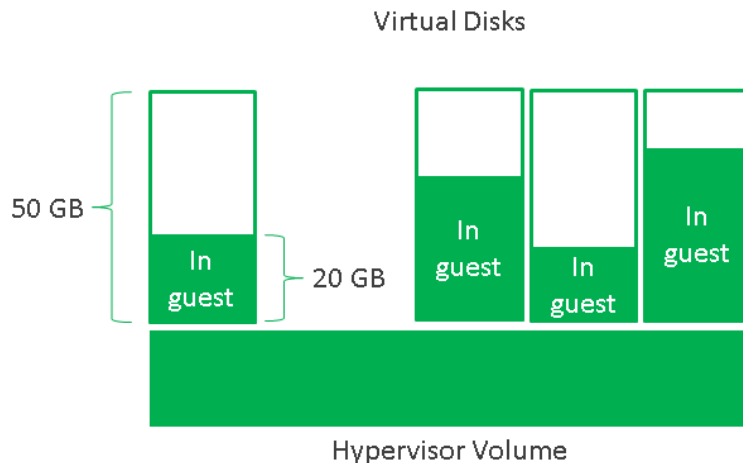
Generally, this forever incremental style is a balance between requiring a lot of I/O performance and storage space. However if you are backing up to a slow NAS device, it might be better to run an “Incremental” job with a weekly active full as this will create a very sequential I/O stream. The downside is that this type of configuration will consume a lot more space.

To read more about different style, please consult the manual:

http://helpcenter.veeam.com/backup/80/vsphere/backup_methods.html

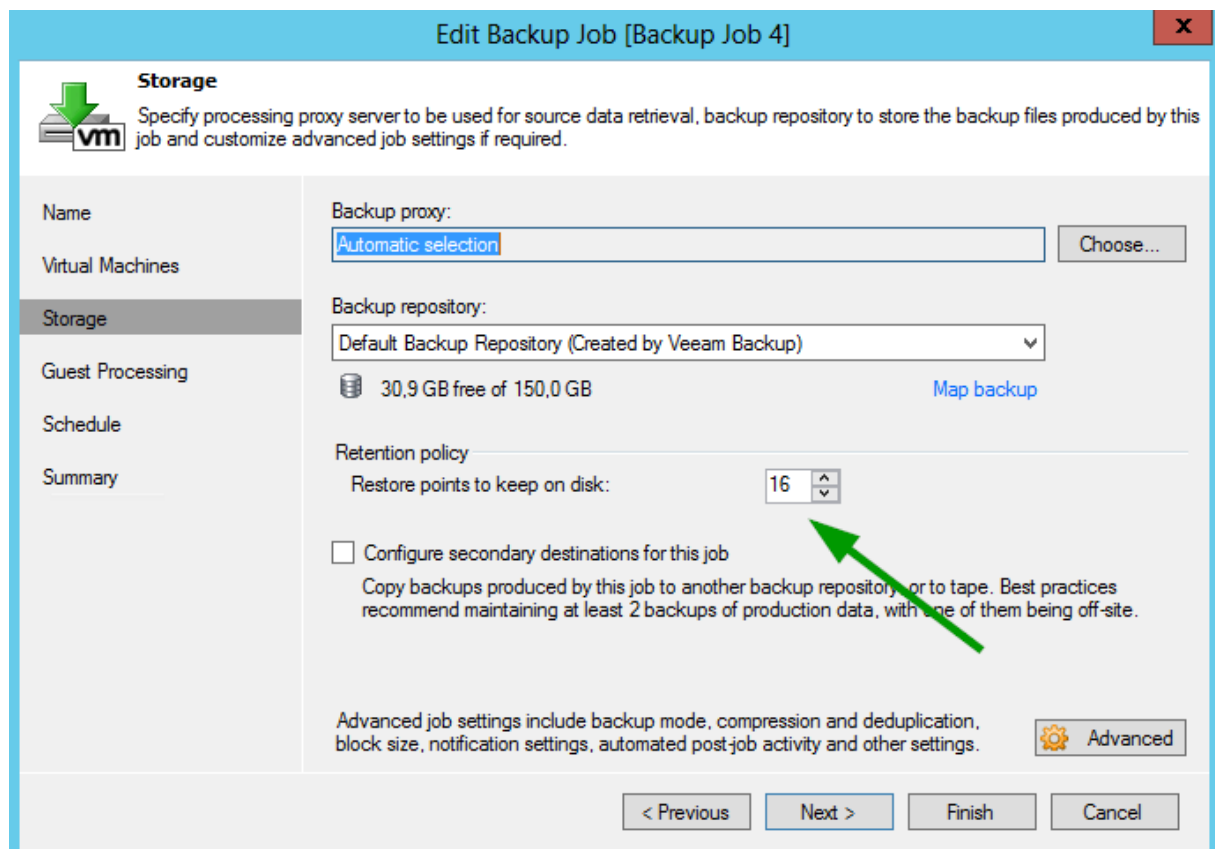
Used Size GB

Used size is the data that is being consumed. Imagine you have a 20 VM's and each VM has a virtual disk size of 50 GB. However inside the guest, only 20 GB is being used by the OS. In this case, the used size is not 50GB x 20, but 20GB x 20 = 400GB. When you have thin provisioned virtual disks, this would match how big the actual virtual disks are since they only consume space when a guest writes to a certain array of the disk.



Retention Points

“Retention points” matches exactly the Backup & Replication GUI. It is the amount of points you want to keep on disk. Veeam might keep more points because of dependency. This behavior can be simulated by choosing “Incremental Weekly Synthetic” scenario. If you choose 14 points, you will end up with 20 points on disk.

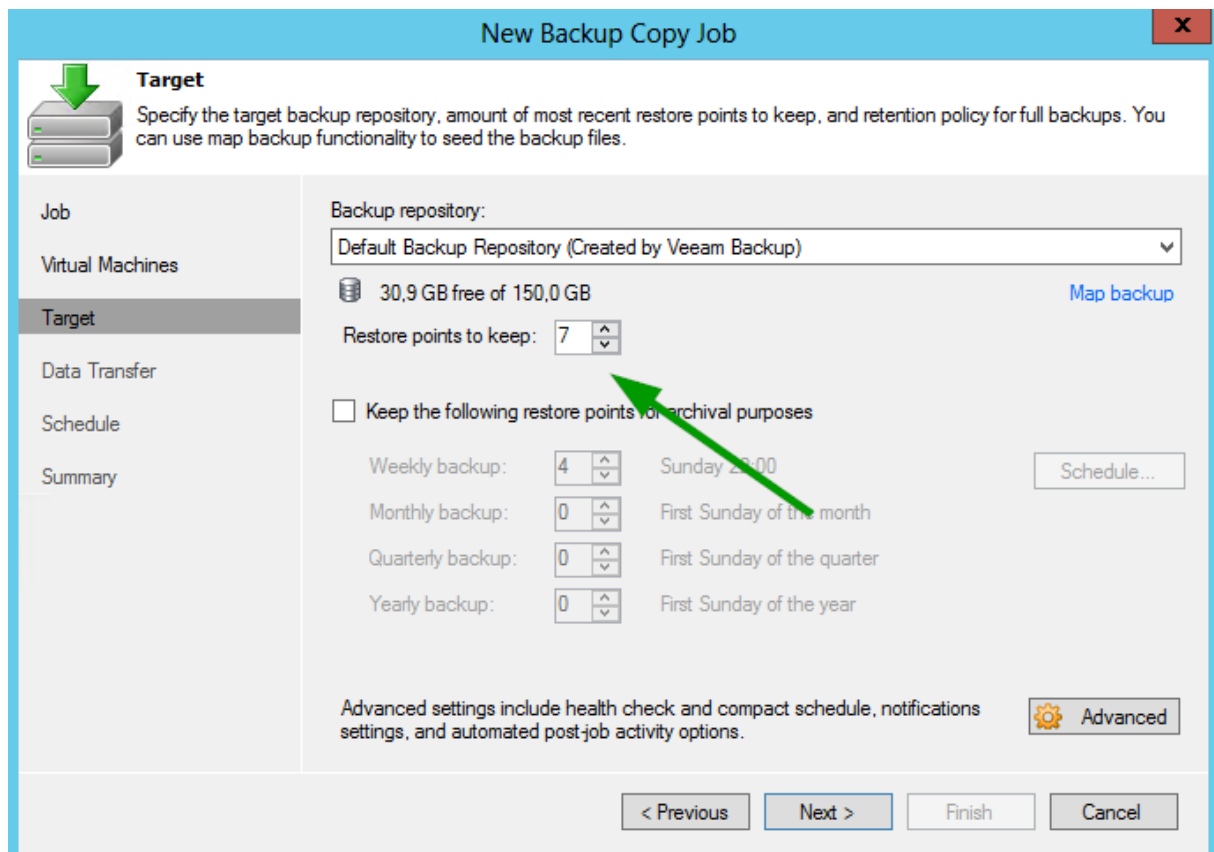


Important is also that “Retention points” does not match with a specific time frame. If you configure 24 points, and run the backup job every hour, you will have around 1 day worth of backups. Sometimes people expect that if they configure 14 retention points, that by default this equals to 14 days of backup. This would match only when you run daily backup.

Finally, if you run a manual backup, this will also count as a retention point. Imagine if you run 1 backup during the night, but somewhere during the day, you decided to run a manual backup. If you configured 20 backups, you will end up with 19 days' worth of backups since you have 1 day that is linked to 2 restore points. Thus, it might be useful, to configure a couple of extra restore points. If, a backup job fails, please consider using the retry option, which does not impact retention policy. If you want to back up a single VM with impacting retention policy, consider using the new quick backup option. This could be useful if you want to run a backup before patching or upgrading a VM.

Quick Backup : http://helpcenter.veeam.com/backup/80/vsphere/quick_backup.html

For backup copy jobs, this matches the simple retention policy



New Backup Copy Job

Target
Specify the target backup repository, amount of most recent restore points to keep, and retention policy for full backups. You can use map backup functionality to seed the backup files.

Job
Virtual Machines
Target
Data Transfer
Schedule
Summary

Backup repository:
Default Backup Repository (Created by Veeam Backup) Map backup

30,9 GB free of 150,0 GB

Restore points to keep: 7

☐ Keep the following restore points for archival purposes

Weekly backup: 4 Sunday 23:00 Schedule...

Monthly backup: 0 First Sunday of the month

Quarterly backup: 0 First Sunday of the quarter

Yearly backup: 0 First Sunday of the year

Advanced settings include health check and compact schedule, notifications settings, and automated post-job activity options. Advanced

< Previous Next > Finish Cancel

Change Rate

Change rate is a parameter that defines how much data will be changed between backups. The first day, all blocks will be read, but the second days, Veeam will identify which blocks have been updated via Changed Block Tracking (CBT). This way, only the changed data needs to be backed up and stored. This impacts the incremental size significantly.

For active directory servers or web servers, the change rate might be as low as 3% on a daily basis. In case of database, due to reorganizations of the data structure the number might go up to 30%. The general consensus is 10%. You can set a more aggressive 5%, but remember that this really depends on the data you are backing up

What is also important is that Veeam use CBT from outside the guest. For example, for every update, the VMware CBT drive will flag at least a 64kb block as changed. So even if the guest only changes 8kb, a lot more data will be flagged as changed in the hypervisor.

Predicting the change rate might be difficult. This is why we encourage the customer to run Veeam One, which has a report that can help you predict the change rate. You can download a free trial from the web site and run it for 30 days.

Free trial : <http://www.veeam.com/virtualization-management-one-solution.html>

The change rate report :

http://helpcenter.veeam.com/one/80/reports/vm_change_rate_estimation.html

Data Reduction

This is the amount of data reduction due to data deduplication and compression. The number is expressed as a percentage of the original data. The default value is 50%. For example if you would enter 20 GB of used data, the full backup will be calculated by taking $20\text{GB} \times 50\% = 20\text{GB} \times (50/100) = 10\text{GB}$. An increment would incorporate the change rate factor. So if you would have 10% change rate, an increment would be around $20\text{GB} \times 50\% \times 10\% = 20\text{GB} \times (50/100) \times (10/100) = 1\text{GB}$.

50% compression might be conservative but know that you should also account for data growth. 40% might be acceptable. With more aggressive values you might end up with a scaling that is only valid for the next 6 months.

Interval

The interval is how often the backup job should be run. This corresponds with the scheduler. For the backup copy job, there is straightforward match

New Backup Copy Job

Job
Backup copy job efficiently creates local and remote copies of your backups, making it easy to maintain multiple copies of your data. Type in a name and description for the job, and specify backup copy interval.

Job

Virtual Machines

Target

Data Transfer

Schedule

Summary

Name: Backup Copy Job 3

Description: Created by VLAB\timo at 6-3-2015 11:19.

Copy every: 1 Day starting at 00:00

Controls how often backup copies are created. Backup Copy job creates a new backup file for each copy interval and starts copying the most recent restore point of each processed VM into this backup file immediately, or as soon as the new restore point appears in the source backup repository.

< Previous Next > Finish Cancel

For a regular backup job, the restore point simulator only allows a subset of the actual possible configurations in Veeam Backup & Replication, mainly daily backup + hourly backup.

Edit Backup Job [Backup Job 4]

Schedule
Specify the job scheduling options. If you do not set the schedule, the job will need to be controlled manually.

Name

Virtual Machines

Storage

Guest Processing

Schedule

Summary

☒ Run the job automatically

☒ Daily at this time: 22:00 Everyday Days...

☐ Monthly at this time: 22:00 Fourth Saturday Months...

☐ Periodically every: 1 Hours Schedule...

☐ After this job: Backup Job Linux (Created by VLAB\timo at 10-11-2014 19:16.)

Automatic retry

☒ Retry failed VMs processing: 3 times

Wait before each retry attempt for: 10 minutes

Backup window

☐ Terminate job if it exceeds allowed backup window Window...

If the job does not complete within allocated backup window, it will be terminated to prevent snapshot commit during production hours.

< Previous Next > Finish Cancel

Edit Backup Job [Backup Job 4]

Schedule
Specify the job scheduling options. If you do not set the schedule, the job will need to be controlled manually.

Name

Virtual Machines

Storage

Guest Processing

Schedule

Summary

☒ Run the job automatically

☐ Daily at this time: 22:00 Everyday Days...

☐ Monthly at this time: 22:00 Fourth Saturday Months...

☒ Periodically every: 1 Hours Schedule...

☐ After this job: Backup Job Linux (Created by VLAB\timo at 10-11-2014 19:16.)

Automatic retry

☒ Retry failed VMs processing: 3 times

Wait before each retry attempt for: 10 minutes

Backup window

☐ Terminate job if it exceeds allowed backup window Window...

If the job does not complete within allocated backup window, it will be terminated to prevent snapshot commit during production hours.

< Previous Next > Finish Cancel

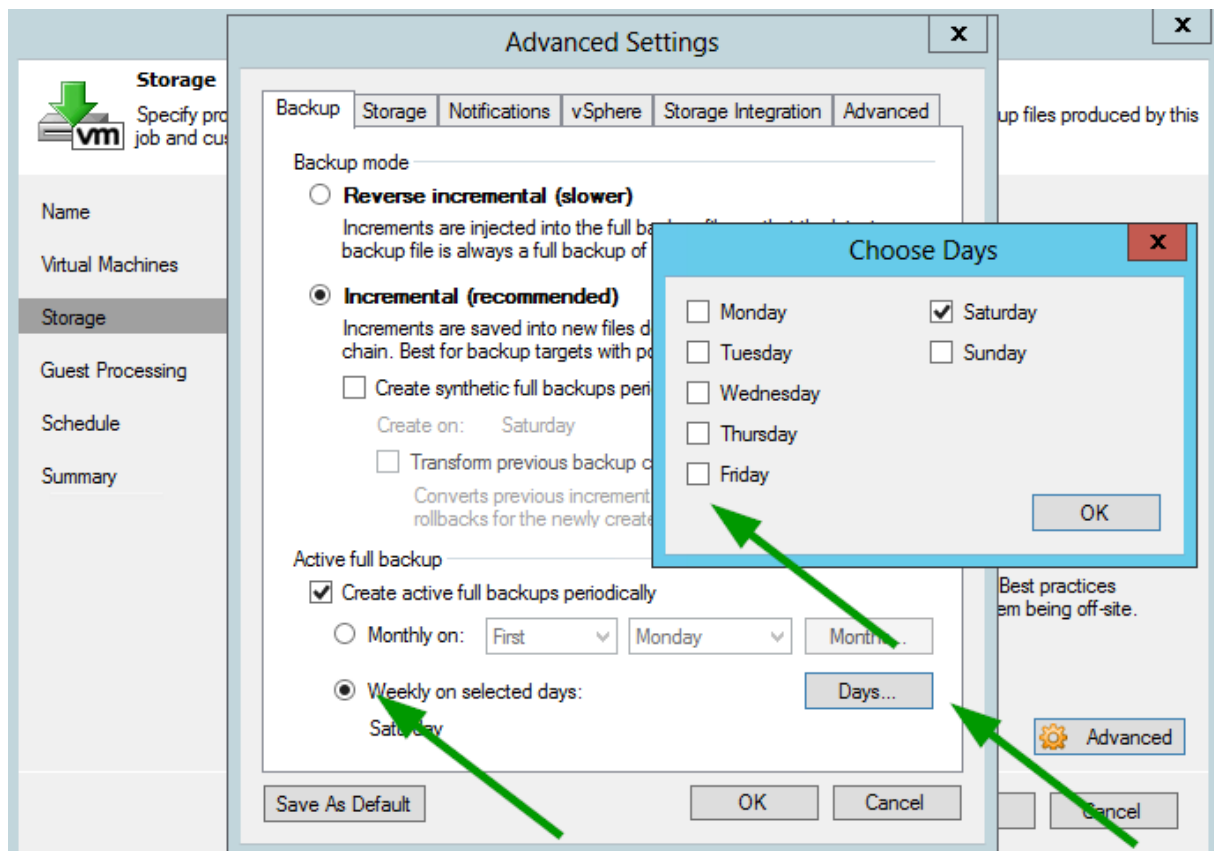
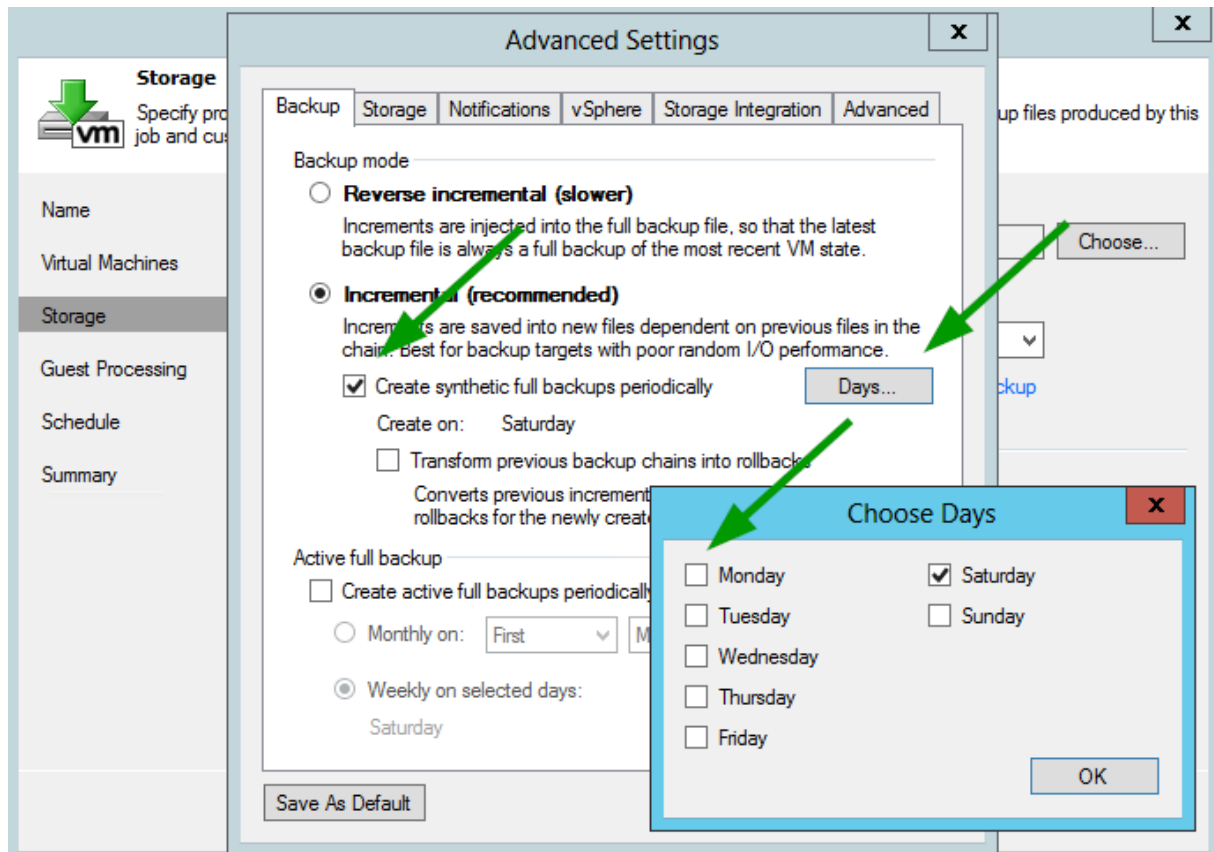
Increment Specific

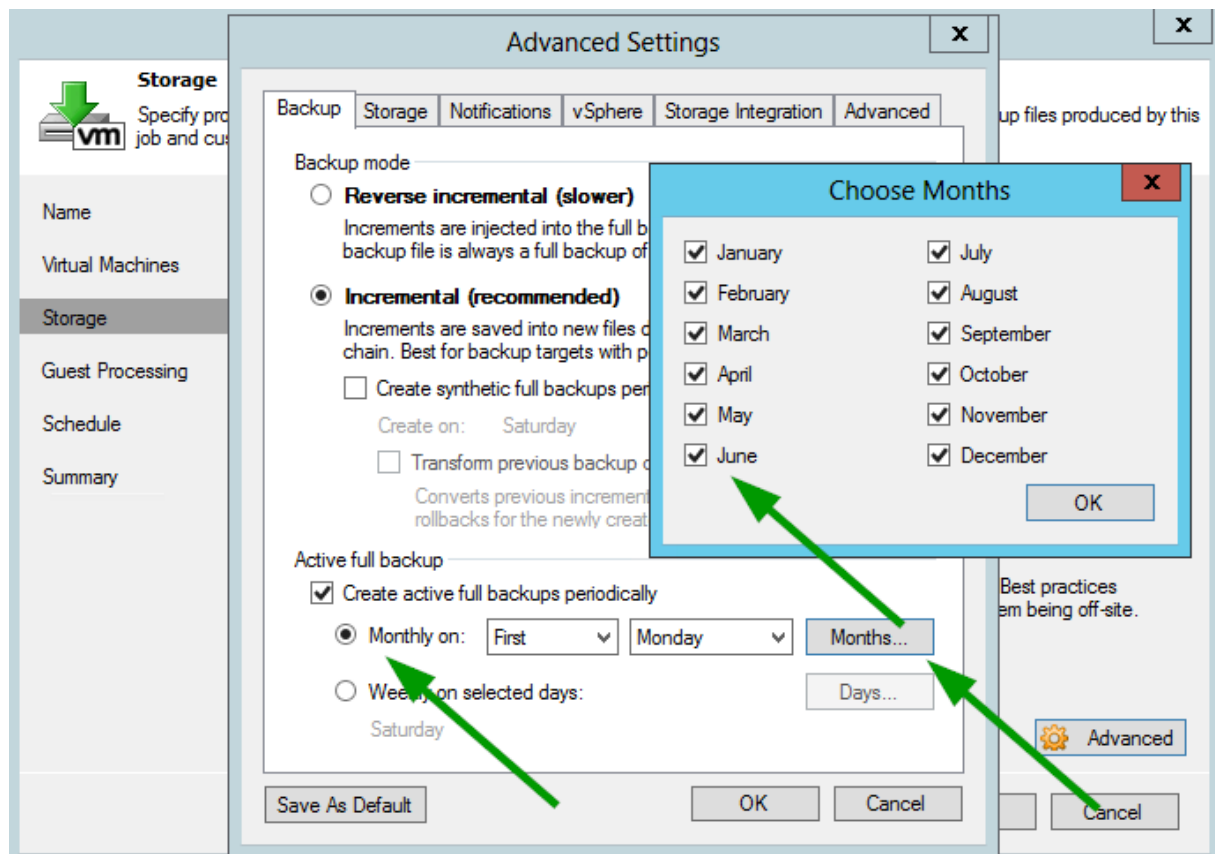
Synthetic backups allow Veeam to create full backup files without requesting all the data from production. Instead it runs an incremental backup first and then in a post process, Veeam will synthetically create a full backup based on the data that is already stored in backup repository. Synthetic full puts fewer loads on production but does require more I/O throughput then Active Full. However different jobs can create synthetic backups on different days, even during production days because it does not create load on production.

Active full reads all data from the production environments similar to the first backup ever ran. However, the data is more sequentially streamed to disk.

These settings allow you to specify when an active full or a synthetic full needs to run. This is a one to one mapping with the backup job which you can find in the advanced windows of the backup job "storage" section. To enable synthetic backup, select one of the days. Similar, for weekly backups select one of the days and for monthly, check one of the months. If you select both active weekly full days and active monthly full months, the active monthly full configuration will be used.

The "Transform previous backup's chains into rollback" option is not available in the simulator.





If you want to use monthly active full, please select the corresponding scenario in the preset so you don't have to check 12 individual checkboxes

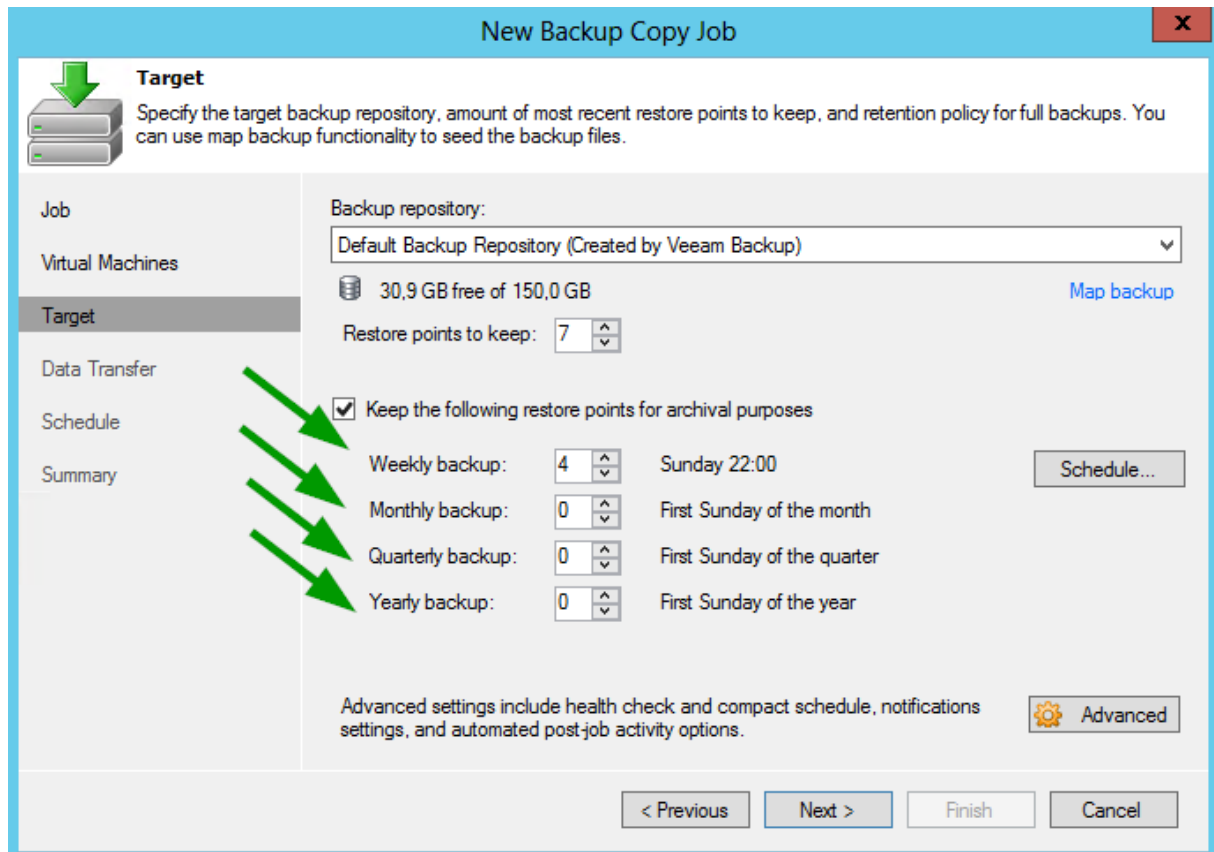
Reverse Specific

Reverse specific is similar in incremental specific. However it does not allow you to create a synthetic full. This is because a Reverse Incremental backup always updates the VBK (Copy on Write, 3 Random I/Os), thus the latest backup is always a full.

You can however configure an active full every 2 months to clean up the VBK file. However over time, it is expected that forever incremental backup will supersede the reverse incremental backup

Backup Copy Job Specific

This specific setting allows you to define the GFS settings. Notice that if you configure 4 weeklies backups, you will end up with 4 weeklies + 1 Working full VBK (5x VBKs). However, monthly backups and weekly backups can be shared. For example, if you configure, 6 weeklies and 1 monthly, you might up with only 6 weeklies + 1 working VBKs. This is because the first Saturday of the month copy might be counted as weekly and a monthly backup. This is also the case for quarterly backups and yearly backups.



New Backup Copy Job

Target
Specify the target backup repository, amount of most recent restore points to keep, and retention policy for full backups. You can use map backup functionality to seed the backup files.

Job
Virtual Machines
Target
Data Transfer
Schedule
Summary

Backup repository: Default Backup Repository (Created by Veeam Backup) [Map backup](#)
30,9 GB free of 150,0 GB

Restore points to keep: 7

☒ Keep the following restore points for archival purposes

Weekly backup: 4 Sunday 22:00 [Schedule...](#)
Monthly backup: 0 First Sunday of the month
Quarterly backup: 0 First Sunday of the quarter
Yearly backup: 0 First Sunday of the year

Advanced settings include health check and compact schedule, notifications settings, and automated post-job activity options. [Advanced](#)

< Previous Next > Finish Cancel

Understanding the result

The result shows the output chain of your job configuration. VBK files are full backup files. VIB/VRB files are (reverse) incremental files.

The first column shows a number that represents the restore point. However in some case you might up with 2 numbers. If you take for example 18 (14), you will see that although we configured 14 restore points, more points are kept. This 18th point is kept, because the 14th point is dependent on a chain and this 18th point is or the parent (VBK) or a sibling dependent (VIB) on the same parent. Because backups are incremental and not differential, all the siblings need to be kept. You will see this happen with incremental style but not with reverse incremental, forever incremental or backup copy style configurations.

With Backup Copy Style you might see “-1 -1W -1M 2Q 1Y” this kind of output if you configure GFS. the -1W specifies that this point was once marked as a weekly backup but is no longer considered as a weekly restore point. 2Q means that it is the second point that is considered as a quarterly back up. In this case you can see that the yearly backup and the quarterly backup are shared in one full backup

The second column shows a visual representation of the file. Orange full backup files are active backups. Green full backups files are synthetic backups or active fulls modified by reverse incremental.

The third column specifies the file extension you might see.

The forth column specifies the file size as calculated by source size and compression factor. In case of incremental files, the change rate factor is also considered

Modify date is the last date the file was changed but does not represent the restore file it is storing. The restore point is shown in the last column.

Finally, the total amount is added and the working space is added. Quite often people have asked what this working space represents. Well the job output shows the result after the job has run. However during the backup, the job might temporarily need extra space because the retention policy is only executed after backup. Also, in some case, it might be required to run an active full manually, which will take extra space. Thus the working space is in correlation with this temporary space.

ONE THING THE SIMULATOR DOES NOT CONSIDER IS THE AMOUNT OF JOBS. IT ONLY REPRESENTS ONE GLOBAL JOB. IN GENERAL IT IS RECOMMENDED TO CONFIGURE 5TB PER SOURCE DATA.

Although possible in the simulator, in reality 100TB VBK files might not be possible. Also it is quite reasonable to say that you will not run an active full on every job at the same time or all jobs are not running at the exact time. This is why an easing function is built in starting from 0.2.8. So working space will not grow linear with the source configuration but will grow slower if the source size becomes bigger.

The idea is built with a bucket system. Once the bucket overflows, the next bucket is used.

0-10 TB = source data will be compressed and then multiplied with a factor of 1.05
10-20 TB = source data will be compressed and then multiplied with a factor of 0.66
20 - 100 TB = source data will be compressed and then multiplied with a factor of 0.4
100 - 500 TB = source data will be compressed and then multiplied with a factor of 0.25
500 TB+ = source data will be compressed and then multiplied with a factor of 0.10

If you have for example 5 TB of source data and 50% compression configured, this would result in:
 $5\text{TB} \times 50/100 \times 1.05 \approx 2.6 \text{ TB Workspace}$

If you have for example 50 TB of source data and 50% compression configured, this would result in:
 $10 \text{ TB} \times 50/100 \times 1.05 + 10 \text{ TB} \times 50/100 \times 0.66 + 30 \text{ TB} \times 50/100 \times 0.4 \approx 5 + 3 + 6 = 14\text{TB Workspace}$

If you have for example 500 TB of source data and 50% compression configured, this would result in:
 $10 \text{ TB} \times 50/100 \times 1.05 + 10 \text{ TB} \times 50/100 \times 0.66 + 80 \text{ TB} \times 50/100 \times 0.4 + 400 \text{ TB} \times 50/100 \times 0.25 \approx 5 + 3 + 16 + 50 = 74\text{TB Workspace}$

You can see that even though 500TB is 100x 5TB, the work space does not grow linearly ($74\text{TB} \approx 28 \times 2.6$)

Manual Run

When you enable this checkbox and you click Simulate, you will be able to run the backup day by day or hour by hour and evaluate the result. By default, the RPS tries to guess what will be the worst case scenario. To do this, the RPS will first predict an end date. The worst case should occur in between the start date and this predicted end date. Once the end date is defined, the engine will start running as per configured interval. Every time, it will evaluate the backup file set as if a successful backup occurred. After a successful “backup” and all the processes like synthetic full creation, retention policy, etc. have been executed, the engine will evaluate if this specific case is the worst case scenario (if it uses more space than previous runs)

The manual mode is very interesting if you want to understand how the retention policy works. Also it will give you a window with the executed actions. In the example below, you will see that because retention point 14 is a full backup and is no longer dependent on the previous chain, the previous chain is being deleted.

Click the 1x to run one interval. The 7x interval might be useful if you have configured a daily backup and you want to skip one week ahead.

Finally, you can also drag the “Manual Run” window by dragging the green bar to anywhere you want. The position will be reset every time you click simulate

Last Steps Fr 2015-03-27 22 h

Created VIB / SEQ 1x I/O Write / 1x 50 GB

Deleting point F >> Sa 15-03-07 22 h full.vbk << 500 GB 21 (14) MOD Sa 15-03-07 22 PARENT <null>

Deleting point I >> Su 15-03-08 22 h incremental.vib << 50 GB 20 (14) MOD Su 15-03-08 22 PARENT Sa 15-03-07 22 h full.vbk

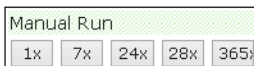
Deleting point I >> Mo 15-03-09 22 h incremental.vib << 50 GB 19 (14) MOD Mo 15-03-09 22 PARENT Sa 15-03-07 22 h full.vbk

Deleting point I >> Tu 15-03-10 22 h incremental.vib << 50 GB 18 (14) MOD Tu 15-03-10 22 PARENT Sa 15-03-07 22 h full.vbk

Deleting point I >> We 15-03-11 22 h incremental.vib << 50 GB 17 (14) MOD We 15-03-11 22 PARENT Sa 15-03-07 22 h full.vbk

Deleting point I >> Th 15-03-12 22 h incremental.vib << 50 GB 16 (14) MOD Th 15-03-12 22 PARENT Sa 15-03-07 22 h full.vbk

Deleting point I >> Fr 15-03-13 22 h incremental.vib << 50 GB 15 MOD Fr 15-03-13 22 PARENT Sa 15-03-07 22 h full.vbk



Result

Retention		File	Size	Modify Date	Point Date
14		full.vbk	500 GB	2015-03-14 Sa 22	2015-03-14 Sa 22
13		incremental.vib	50 GB	2015-03-15 Su 22	2015-03-15 Su 22
12		incremental.vib	50 GB	2015-03-16 Mo 22	2015-03-16 Mo 22
11		incremental.vib	50 GB	2015-03-17 Tu 22	2015-03-17 Tu 22
10		incremental.vib	50 GB	2015-03-18 We 22	2015-03-18 We 22
9		incremental.vib	50 GB	2015-03-19 Th 22	2015-03-19 Th 22
8		incremental.vib	50 GB	2015-03-20 Fr 22	2015-03-20 Fr 22
7		full.vbk	500 GB	2015-03-21 Sa 22	2015-03-21 Sa 22
6		incremental.vib	50 GB	2015-03-22 Su 22	2015-03-22 Su 22
5		incremental.vib	50 GB	2015-03-23 Mo 22	2015-03-23 Mo 22
4		incremental.vib	50 GB	2015-03-24 Tu 22	2015-03-24 Tu 22
3		incremental.vib	50 GB	2015-03-25 We 22	2015-03-25 We 22
2		incremental.vib	50 GB	2015-03-26 Th 22	2015-03-26 Th 22
1		incremental.vib	50 GB	2015-03-27 Fr 22	2015-03-27 Fr 22
			1600 GB		
			Work Space		
			+550 GB		
			2150 GB		

URL Parameters

One of the possibilities with RPS is that you can pass the result via get parameters, for example if you want to share the result with a colleague.

For example:

<http://rps.dewin.me/?m=1&s=500000&r=14&c=50&d=10&i=D&e>

When you click the run, a new scenario will be created called "URL parameters". Here is a subset of possible parameters:

Parameter	Activates
m	Style parameter. Incremental = 1, Reverse Incremental = 2 & Backup Copy Job = 3
s	Source data
r	Retention points
c	Compression
d	Delta or change rate
i	Interval: D for daily or number of hours e.g. i=4 will set the interval to every 4 hours
e	Execute, instantly calculates, essentially clicking the Simulate button
x	Manual execute mode
sy	Select synthetic days as an array e.g. sy=0,0,0,0,0,1,0 will select saturday
iw,im	Similar as sy but for incremental active weekly or monthly full e.g. im=1,0,1,0,1,0,1,0,1,0,1,0 to do a bi-monthly backup
rw,rm	Similar as iw,im but for reverse incremental
g	GFS retention. For example g=4,3,2,1 will configure 4 Weekly, 3 Monthly, 2 Quarterly and 1 Yearly backup

There is a hidden dev mode where you can preconfigure everything and then click the URL>> parameter to generate this for you. Use the dev parameter to enable this. You can use no value or random value for the dev parameter.