

Frequency Based Backup – It's not a bad thing

You know, I still get a lot of people asking me to re-publish this article that I wrote for public consumption a several years ago. How to use frequency based scheduling with NetBackup...so I thought I'd post it here in my blog.

Frequency Based Backup - Does it deserve a bad "rep"?

Do you remember that kid on your block that all the parents loathed? It seemed like everyone who hooked up with him got into trouble. Mom and Dad just finally had to say "no" when you asked to play with him.

In most cases these assumptions and positions are made on very little if any first hand knowledge or worse pure heresy. I think Frequency Based Backup is a lot like that kid on the block with the bad "rep", the one everyone "thought" they knew. In reality the kid ended up being a great asset, trustworthy, prompt and a good friend – just a bit misunderstood; much like Frequency Based Backups.

VERITAS NetBackup has been using frequency based backups for as long as I can remember and that goes back to version 1.6. Frequency based backup is very powerful if you understand how to unleash it. I'll explain the details later in this document, but right now let's understand the basics.

The Backup Policy acts as the control panel for any backup job that runs under NetBackup. It controls the "who, what, where, when and how", or clients, file list, storage units, backup window and backup type. The Backup Window and the Backup Type are components of the Schedule. The Backup Window defines the time to open and close the window and the Backup Type defines whether it will be a full, incremental, differential or user directed backup.

Another part of the schedule is the frequency of this backup. In most customer environments I have visited, I see incremental schedules with windows opened up for 5 or 6 of the 7 days in the week and the frequency set to daily. And full backup schedules on the "open" days. So it looks something like this:

Incremental Schedule – Frequency 1 day

Sunday	Monday	Tuesday	Wednesday	Thursday	Friday	Saturday
	Inc	Inc	Inc	Inc	Inc	
	7:00pm	7:00pm	7:00pm	7:00pm	7:00pm	
	6:00am	6:00am	6:00am	6:00am	6:00am	

Full Schedule – Frequency 1 week

Sunday	Monday	Tuesday	Wednesday	Thursday	Friday	Saturday
Full						Full
3:00pm						3:00pm
4:00am						6:00am

This works well, but it doesn't unleash the power of the Frequency Based Backup schedule. As you can see, in this example the Admin starts the incremental window at 7:00pm and closes it at 6:00am, opens the Full on Saturday and Sunday at 3:00pm and closes it at 6:00am Sunday morning and 4:00am Monday morning. The problem with this is if full backups happen to fail on Saturday, it will try again on Sunday and if successful will re-run 1 week later, which is on a Sunday. You run the risk of having the entire backup set choke up on one particular day especially if this is how ALL of your policies are configured. If you cannot complete the backups on Sunday then the next FULL backup will not start until Saturday and you are in the same predicament as previously stated.

When you understand frequency based backups it will provide you so much more leverage in how you architect your backup strategies.

With NetBackup the schedules are prioritized within a Backup Policy based on a hierarchy of frequency not backup type as some have wrongfully assumed. For example an incremental backup schedule with a frequency of 30 days will have a higher priority than a full backup schedule with a frequency of 7 days. Here's an application of frequency based backups that hopefully will help convey the message.

A customer needs to backup their environment and they want to ensure that their full backups are staggered throughout the week for each policy all totaled about 24 backup policies need to be balanced. After a little homework reviewing the backup policies it is determined that four policies each day will run their full backup schedules. It is possible to have 12 of the 24 policies run and successfully backup their schedules, but it is very

tight and sometimes runs into the production hours so the plan is to avoid this at all costs. To simplify this illustration we will only use one Backup Policy and their schedules to show how Frequency Based Backup scheduling works.

Backup Policy “DB-Servers” needs to run its full backups on Monday and if by chance the backups fail, we will allow a “rain-date” to be played on Tuesday and Wednesday. Since the incremental backup jobs need to run daily that schedule will look like this:

Incremental Schedule – Frequency 1 day

Sunday	Monday	Tuesday	Wednesday	Thursday	Friday	Saturday
Inc	Inc	Inc	Inc	Inc	Inc	Inc
7:00pm	7:00pm	7:00pm	7:00pm	7:00pm	7:00pm	7:00pm
6:00am	6:00am	6:00am	6:00am	6:00am	6:00am	6:00am

The full backup schedule is the one that needs to be “tweaked”. For this to take place the full backup schedule will need to look like the following:

Full Schedule – Frequency 5 days

Sunday	Monday	Tuesday	Wednesday	Thursday	Friday	Saturday
	Full	Full	Full			
	7:00pm	7:00pm	7:00pm			
	4:00am	6:00am	6:00am			

This ensures that the full backup will ALWAYS run on Monday. For example if the backup runs on Monday but fails, it will be attempted on Tuesday if that is successful the schedule waits five days before it will restart. Five days from Tuesday is Sunday which doesn’t have an open window so the very next window of opportunity that opens for this schedule will allow it to run which is Monday. Same goes for a failed backup on Tuesday that runs on Wednesday, five days from that is Monday.

Now the next question that people typically ask is okay, well if I have an Incremental schedule with an open window on the same day as the Full Backup window, how does NetBackup know which one to run? Very good question and I hope I can illustrate it well enough for you to understand.

NetBackup will make its selection based on the hierarchy of the frequency. For example, you have two very good friends one lives next door to you and the other 1000

miles away who you rarely see. When the friend 1000 miles away and the friend next door come to visit, which one do you talk to first? Hopefully you said the one 1000 miles away. This is exactly how NetBackup works, the schedule that has been seen the least is the schedule that is selected. Perhaps a picture will help solidify this concept.

Sunday	Monday	Tuesday	Wednesday	Thursday	Friday	Saturday
I-1	I-1	I-1	I-1	I-1	I-1	I-1
	F-5	F-5	F-5			
	F-26	F-26	F-26			

If the backup schedule starts on Sunday, naturally the incremental (1-day) will run, but on Monday there are three schedules, NetBackup looks at the available schedule options and deducts based on “last seen” which one it will select. Based on all three it “basically” says, I’ve seen I-1 yesterday, but haven’t seen F-5 ever, but then there is this F-26 which I also haven’t seen so I will run this because it has the greater frequency. Next week it will go through the same process, but this time it will look at I-1 and say “saw it yesterday, F-5 haven’t seen this for 7 days and it is a 5 day frequency, F-26 saw this 7 days ago and is a 26 day frequency, so I will select F-5.”

So taking a look at what schedules would look like that are spread across the entire week, as this customer wants to do would look something like this (for the sake of space and time we have only elected to show 6 policies out of the 24 outlined in our example):

Sunday	Monday	Tuesday	Wednesday	Thursday	Friday	Saturday
All:I-1	All:I-1	All:I-1	All:I-1	All:I-1	All:I-1	All:I-1
	DB:F-5	DB:F-5	DB:F-5	APP:F-5	APP:F-5	APP:F-5
	DB:F-26	DB:F-26	DB:F-26	APP:F-26	APP:F-26	APP:F-26
CL1:F-5	CL1:F-5	CL1:F-5	CL2:F-5	CL2:F-5	CL2:F-5	NFS:F-5
CL1:F-26	CL1:F-26	CL1:F-26	CL2:F-26	CL2:F-26	CL2:F-26	NFS:F-26
NFS:F-5	NFS:F-5	HOM:F-5	HOM:F-5	HOM:F-5		
NFS:F-26	NFS:F-26	HOM:F-26	HOM:F-26	HOM:F-26		

So you can see how this will eventually end up, very nicely spread out with “rain-dates” for all of the full backup schedules. Incidentally, the 26 day full backup schedule is a monthly, some companies say it has to run on a particular DATE and if that is the case,

then the calendar scheduling found in NBU version 4.5 and later is appropriate. For 95% of your backup scheduling needs, I think frequency based backup is powerful enough to meet your expectations.

My advice, get to know the kid before you pass judgment, I think you will be pleasantly surprised.

So that was the original article that I wrote. I still think Frequency based backup is pretty powerful, if you know how to expose its potential.

-Chapa signing off

David A. Chapa

CTE, Quantum Corporation

January, 2012