

The Data Protection Framework

symform

AUTHOR:

Praerit Garg

6/24/2010

ABOUT THE AUTHOR

Praerit Garg, President and Co-founder

Praerit Garg is the President and Co-founder of Symform. Prior to Symform, Praerit was a Senior Director in Microsoft's Server and Tools division where he built and managed the Dynamic Systems Platform & Tools team. Under his leadership, the team grew from 0 to over 70 people and drove a broad industry initiative to expand W3C's XML standards to incorporate modeling enhancements called Service Modeling Language (SML). The team also delivered technology and solutions across several Microsoft products, including the distributed modeling system in Visual Studio 2007, Group Policy tools in Windows Vista/Server 2008, Server Manager in Windows Server 2008, and Desired Configuration Management in System Center Configuration Manager 2007. Praerit also led an acquisition and integration of a software company developing Group Policy tools.

Earlier in his 12+ year tenure at Microsoft he held various engineering leadership positions in Windows, ranging from a software developer driving US and European security certifications of Windows NT 3.51/4.0 products, to Group Program Manager for Windows Security delivering security infrastructure, such as Active Directory, Kerberos, PKI, EFS and Group Policy in Windows 2000, XP and 2003 releases.

Praerit is a co-inventor on 14 U.S. and international patents and a co-author of IETF RFC 3645, Web Services WS-Trust and WS-Secure Conversation specifications. He holds an M.S. in Computer Science from Purdue University and a B.E. (with Honors) from Birla Institute of Technology & Science, Pilani, India.

The Data Protection Formula

Achieving best possible coverage, speed and efficiency

In an earlier whitepaper, we discussed the “3 C’s” that any data protection solution must meet to be effective, particularly in the SMB market. The 3 C’s are: **comprehensive**, **convenient** and **cost-effective**. Combining a good automated local backup process with the Symform Cooperative Storage Cloud results in a *set-n-forget* solution. The creation of onsite backups + offsite mirroring of the backups is the only way to achieve all 3 C’s in the market today. In this paper, we will discuss the specifics of configuring such a system in order to achieve the 3 C’s of this framework and then apply the framework to specific use cases.

Convenience is all about the solution being set-n-forget – something the customer or their IT service provider does not need to constantly babysit and will send alerts when things go wrong. Any serviceable local backup software application allows you to setup a backup schedule to ensure backups are happening regularly, as well as send alerts to inform you if anything goes wrong. Installing the Symform Node software on one of the local computers so that it can participate in the Symform Cooperative Storage Cloud ensures that those backup files are quickly and automatically mirrored into the storage cloud. This provides disaster protection. The Symform Cooperative Storage Cloud keeps the IT service provider informed via periodic reports as well as via alerts when something goes wrong.

Cost-effectiveness is about the solution fitting a customer’s budget. For local backup, customers have many choices that fit their budget and needs. These range from using the built-in backup software in Windows that is already paid for as part of the Windows license to more sophisticated and full featured software applications like StorageCraft’s ShadowProtect™ or Acronis’ TrueImage™. The Symform Cooperative Storage Cloud is offered at a flat monthly fee per server or per PC for *unlimited* online storage through an IT service provider (a reseller). There are no per GB fees. Depending on the customer’s IT service contract, the service provider may bundle the monthly fee in their contract or bill separately for it. This disruptive pricing model enables customers to get disaster protection for all their business critical data at fraction of the cost relative to other options in the market that charge on a per GB basis. Furthermore, customers and IT service providers do not have to worry about costs increasing as data grows over time. Equally important is the fact that the solution is installed and supported through a trusted, local service provider – it is not just an online website + 1-800 number. This means that in the event of a disaster, the customer is serviced by someone they already know and trust -- someone who can physically be there to help. This is a level of commitment that not many online services can provide.

Unlike convenience and cost-effectiveness, achieving **comprehensive** data protection coverage involves a bit more thinking and planning. It is best addressed by a trained IT service provider. Full coverage requires the understanding of many variables and constraints around storage, bandwidth, and the preferred backup software. Only then can an IT service provider create a plan and an automated process that provides the necessary coverage against any kind of data loss ranging from “oops -- I

deleted a file” to a power surge frying all the equipment including any local backup storage. Most importantly, it must also balance the critical need for a speedy restore as well as the efficient use of storage and bandwidth. We express this as a function:

F: [backup size, upload bandwidth, backup software] → Data Protection [quality & cost of coverage]

We call this **the Data Protection Framework™**. It enables us to take the input constraints around the required backup capacity, available upload bandwidth, and backup software options to achieve the optimal data protection coverage for a particular customer with a given budget.

Storage constraints involve the minimum amount of data that needs to be backed up as well as the rate of change of this data. Bandwidth constraints are created around the maximum Internet bandwidth that the business can afford and the Internet services available in its locale. Upload bandwidth constraints impact the speed at which the data can be transferred into the cloud.

Together, these issues will shape the best backup strategy to employ in order to ensure there is at least one fairly recent *offsite* backup set for disaster recovery needs. It is worth noting that when using the Symform Cooperative Storage Cloud, the only relevant bandwidth constraint is the customer’s upload bandwidth. Any alternative online storage solution would also require factoring in the maximum per-connection bandwidth available on the data center side. This is due to the fact that data center based services need to manage their bandwidth costs and capacity. The Symform system does not suffer from this limitation – the aggregate storage and bandwidth naturally grows as more nodes join the network. This powerful property enables Symform nodes to mirror as fast as they can, simplifying the data protection framework. Symform is typically 10X faster than traditional online storage services.

The Data Protection Framework

Let us start by understanding the core components of the data protection framework and then apply them to three customer scenarios.

The data protection framework consists of following 3 basic steps:

1. **Identify an appropriate backup schedule** -- the backup schedule is based upon how much data can be successfully uploaded given the customer’s data set size, rate of change, available bandwidth, and backup software options.
2. **Configure local backups** -- implement an optimal backup schedule for local backups.
3. **Configure offsite mirroring** of the local backup sets by configuring the Symform node to get disaster protection.

The following tables provide some rules of thumb for determining:

- An estimate of daily and weekly data transfer rates given particular upload bandwidth constraints.

- A conservative backup schedule based on the size of backup data set, rate of change of data, and available bandwidth.

These two tables are used as a starting place when determining the best data protection strategy for the customer. Subsequent examples will demonstrate how to use them.

Data Transfer Based on Bandwidth Availability

BW during business hours (10 hours)	BW during non business hours (14 hours)	Data uploaded during 1 day (max.)	Data uploaded during 1 week (max.)
384Kbps	768Kbps	4GB	32GB
1.5Mbps	2Mbps	12GB	90GB
4Mbps	5Mbps	32GB	230GB
9Mbps	10Mbps	50GB	350GB

Backup Schedule Matrix (when using built-in free, simple backup software)

Data Upload bandwidth →	≤1Mbps	<10Mbps	≥10Mbps
≤50GB total, <50MB/day	Monthly	Weekly	Weekly
≤500GB total, <1GB/day	Quarterly	Monthly	Weekly
500GB or more, <5GB/day	3 C's not possible	Quarterly	Monthly/Weekly*

*It will depend on the customer situation. Scenario 3 discussed below provides some insights.

A Backup Schedule

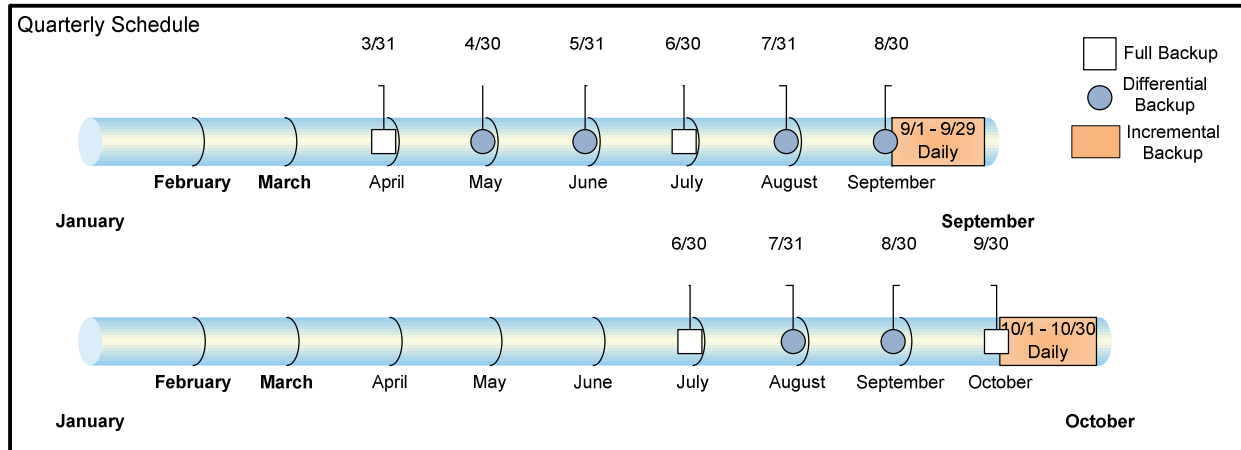
A backup schedule consists of three backup sets that are generated based on a specified periodicity such as weekly, monthly, or quarterly. One backup set is generated for each period.

This configuration is often referred to as the “Grandfather” – “Father” – “Son” (G-F-S) scheme. The backup set currently in progress is referred to as the “Son”. Last period’s set is referred to as the “Father” and a set for the period before that is the “Grandfather.” Together, these three sets provide the desired level of data protection:

- A Grandfather set contains the end-of-period full backup only. It represents the oldest point in time that a restore could be completed.
- A Father set typically contains one or more differentials relative to the Grandfather and an end-of-period full backup.
- A Son set contains one or more differential backups relative to the Father full backup and one or more incremental backups relative to the last differential backup to minimize the data loss time window.

Let us look at three common backup schedules to help understand the G-F-S scheme.

Quarterly Backup Schedule



For example, on Sept 11, a quarterly G-F-S scheme would look as follows:

- Grandfather: Full Backup on 3/31.
- Father: Differential Backups on 4/30, 5/31 and a Full Backup on 6/30.
- Son: Differential backups on 7/31 and 8/31 and daily incremental backups for Sept 1-10th.

A disaster on Sept 11 could be compensated by a restore to state on Sept 10th using the full backup on 6/30 and combining the differential from 8/31 and incrementals from Sept 1-10th. If for any reason the full backup on 6/30 was corrupted, the customer could still be restored to the state on 5/31 using the full backup on 3/31 and the differential from 5/31. For a quarterly schedule, the best case data loss window is 1 day and worst case could be as much as 6 months if by sheer bad luck, disaster occurred on 9/30 and the full backup on 7/31 as well as differentials on 6/30 and 5/31 were all corrupted.

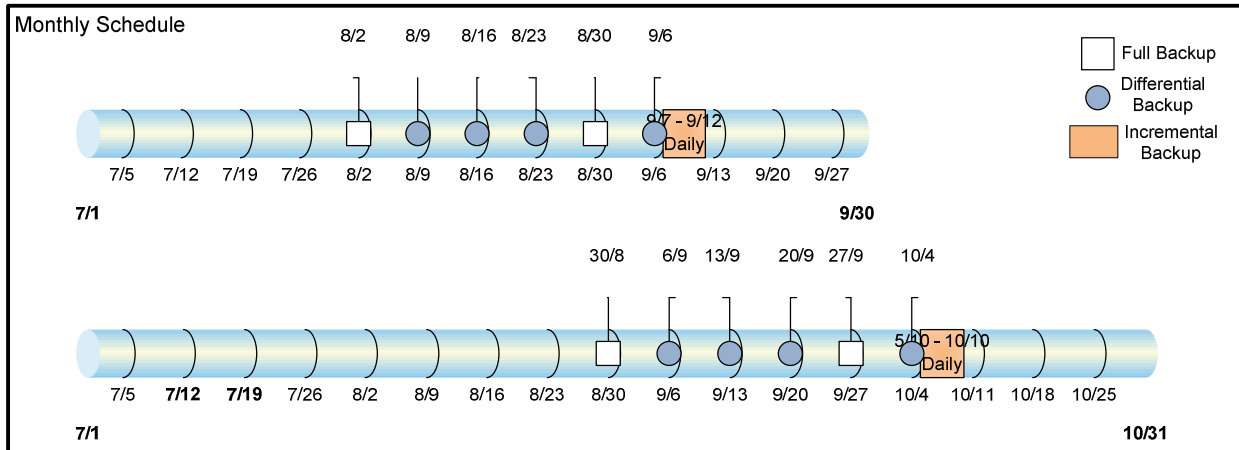
On October 11, a quarterly G-F-S scheme would look as follows:

- Grandfather: full backup on June 30.
- Father: differential backups on July 31st, August 31st, and a full backup on Sept 30th.
- Son: daily incremental for Oct 1-10th.

Once the full backup on 9/30 is complete, the backups from 3/31, 4/30 and 5/31 are dropped; making the 6/30 backup set the grandfather and the September set now the father.

NOTE: incremental backups may be reduced to a weekly schedule if the data rate needs to be managed. It also means that in the event of a restore, the best case would be to the last Sunday.

Monthly Backup Schedule



NOTE: because months are not uniform, it is usually better to configure the monthly schedule as a 4-week schedule as shown here rather than sticking to months.

On Sept 11, a monthly G-F-S scheme would look as follows:

- Grandfather: full backup on Sunday, Aug 2nd.
- Father: differential backups on each Sunday in August and a full backup on Sunday, Aug 30th.
- Son: differential for Sunday Sept 6th and daily incremental backups for Sept 7th-10th.

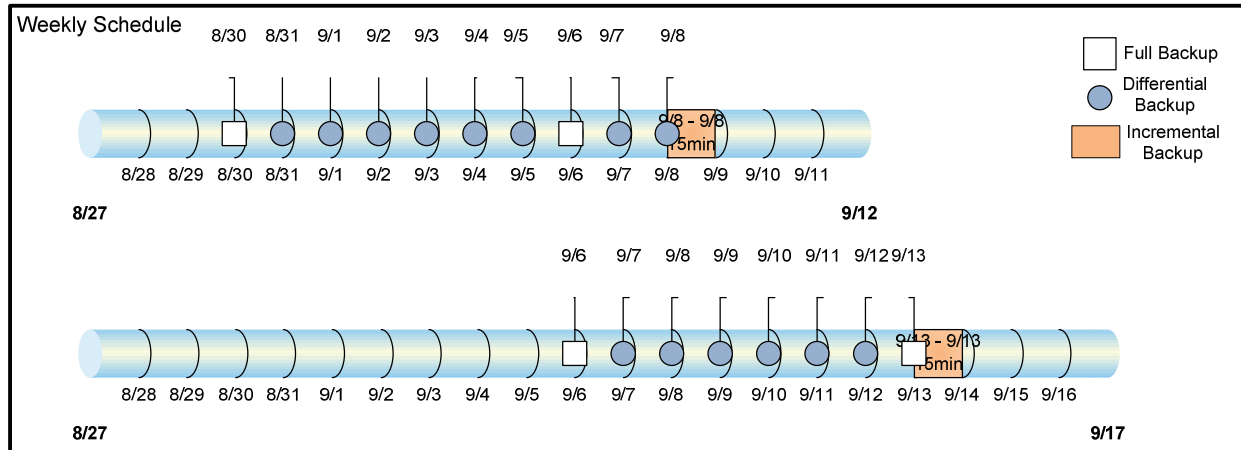
A disaster on 9/11 could be compensated by a restore to the state on 9/10 using the full backup on 8/30, the differential on 9/6, and the dailies from 9/7-10. If, for any reason, the full backup on 8/30 is bad, the customer still has the ability to restore to the state on 8/23 using the full backup on 8/2 and differential on 8/23. In the best case, monthly schedule is as good as the quarterly schedule but in the worst case, it is much better than quarterly schedule where the customer might have to suffer data loss of at most 8 weeks, not 6 months.

On Oct 11, a monthly G-F-S scheme would look as follows:

- Grandfather: full backup on Sunday, August 30st.
- Father: differential backups on Sept 6, 13 and 20th and full backup on Sept 27th.
- Son: differential for Oct 4 and daily incrementals for Oct 5-10.

When the full backup for 9/27 is complete, the full backup on 8/2 and the differentials on 8/9, 16, and 23 are dropped. The full backup on 8/30 takes on the grandfather status and backups in September make up the father set.

Weekly Backup Schedule



On Tues Sept 8th, a weekly G-F-S scheme would look as follows:

- Grandfather: full backup at 12am on Sun Aug 30th.
- Father: daily differentials taken at 12am Mon Aug 31st – Sat Sept 5 and Full Backup at 12am on Sun Sept 6th.
- Son: daily differentials at 12am on Mon Sept 7th and Tue Sept 8th and 60 min. (or smaller) incrementals during Tue Sept 8th starting at 1am

On Sun Sept 13th, a weekly G-F-S scheme would look as follows:

- Grandfather set: full backup at 12am on Sun Sept 6th.
- Father: daily differentials taken at 12am Mon Sept 7th – Sat Sept 12th and full backup at 12am on Sept 13th.
- Son: Hourly (or smaller) incremental during Sun Sept 13th.

As can be seen, the weekly backup schedule is much more granular compared to the quarterly or the monthly schedule. In the best case, it enables restore to the state of the system as recent as an hour (or less) back and in the worst case two weeks old.

Rationale for a G-F-S scheme

One obvious question that arises when looking at various G-F-S schemes for a backup schedule is – why are so many backups needed and how to decide which schedule makes sense for a customer?

For starters, it is worth observing that the amount of data stored is in the same order of magnitude irrespective of the schedule used – two full backups, a few differentials relative to each full backup and a few incremental backups relative to the last differential. They are primarily different in time granularity -- the quarterly schedule is less granular than the monthly which is less granular than the weekly.

Why so many backups?

The G-F-S scheme optimizes for the following key data protection objectives:

1. Ability to restore from the most recent version. Most often this is what happens.
2. Ability to restore from a few intervals back. In some cases where the data loss or corruption is not discovered immediately, it might be necessary to restore from a few versions back. Again the objective is to get to the most recent good state.
3. Protect against backups themselves going bad. Just like live data, backups can be bad or get corrupted as well, therefore protecting against that is important as well.

The most recent full backup (Father set) and differentials and incremental since then (Son set) provide coverage for first two objectives. Previous full backup (Grandfather set) and differentials to it in the Father set provide coverage for the third objective.

Which backup schedule?

As noted, the key difference between backup schedules is the granularity of backup. While in theory, the quarterly schedule would let a customer go back many months and the weekly would only enable a customer to go back three weeks. In practice, however, one is unlikely to want to go back too far. The Grandfather full backup is mostly an insurance plan against the more recent backups getting corrupted. It is highly recommended that a test restore be performed using the Grandfather backup periodically. Ideally, the period should align with the backup schedule.

In a quarterly schedule, in the best case, you could restore to the last day's state and in the worst case 6 months old. On the other hand, a weekly schedule enables you to restore to as recent as an hour ago in the best case and in the worst case just 2 weeks back.

In general, the more granular the schedule the better it is. The constraint that may cause you to pick a less granular schedule, such as monthly or even quarterly, is primarily the amount of time it takes to get any given backup automatically mirrored offsite for disaster protection (or the hassle involved in manually taking the backup tapes offsite). As noted earlier, unlike alternative online services, when using the Symform Cooperative Storage Cloud, this is gated only by the upload bandwidth available at the customer site. The Backup Schedule Matrix provides the recommended schedule to use based upon the initial backup set that needs to be mirrored and the upload bandwidth available. Note that the Backup Schedule Matrix assumes a correlation between the initial backup set that needs to be mirrored and the rate of daily change. In general, that will hold true because higher rate of change typically implies a larger data set. There might be exceptions such as

- A large initial backup set but with a low rate of change such as a long running business. In this case, the customer could be more aggressive and use a more granular schedule and use an alternative high bandwidth mechanism to get the initial backup set into the storage cloud. An alternative high bandwidth mechanism may be their IT Service provider's premises or collocation facility.
- A small initial backup set but with a high rate of change such as a newly created graphic design firm. In this case, we recommend being conservative and going with a schedule that can sustain the rate of change. Otherwise, the mirror to the storage cloud may not be able to keep up with the local changes over time.

Configuring a local backup schedule

Let us take two different examples to explain the process of setting up a backup schedule. We will use Microsoft's NT Backup and StorageCraft's ShadowProtect. We will also discuss Windows 2008 Backup because of its significant departure from the standard backup policies supported by most backup software.

- Weekly GFS scheme using NT Backup utility
 - Using the backup scheduling capability in NT Backup, configure following:
 - Two full backups happening on alternate Sundays at midnight.
 - Daily differentials for 6 remaining days of each week relative to the full backups generated at midnight. Each of these is set up as two schedules. The first day differential overwrote the previous file and the subsequent 5 daily differentials appended to it.
 - Hourly incrementals capturing changes during the day. This is set up as 2 schedules. The first schedule creates the first 1 hr incremental by overwriting the last day's incremental. The second schedule appends the incremental every hour to it for the next 22 hrs. The differential backup would kick in at the midnight to capture the entire day.
- Quarterly GFS scheme using StorageCraft's ShadowProtect
 - Create two backup sets, each spanning three months.
 - Each set starts with a full backup on the first day of the quarter.
 - The daily incrementals are taken on all remaining days of the month.
 - On the first of month of the next two months, a differential backup is taken and replaces all the daily incrementals of the previous month.

Additional technical details and guidance can be found at: <https://control.symform.com/forums/>

Special Considerations when using Windows 2008 Built-in Backup

The above backup schedule recommendation is based on a fairly standard practices in the industry around full, differential, and incremental backup support in most backup software.

In Windows 2008 Server, Microsoft replaced the NT Backup utility with a new backup utility which is significantly simpler to use for the basic local backup needs. However, it does not have the necessary sophistication to manage the backup data set for efficient and automatic mirroring to an online store service using limited bandwidth broadband connections. It creates one large VHD file containing all backup versions of the live system. This single file changes significantly on every backup so mirroring it to the online store cannot be accomplished quickly. We recommend following adjustments to the backup schedules when using Windows 2008 Server's built-in backup software:

- Separate the local backups from the online backups and schedule them independently.

- Use a more aggressive local backup schedule such as twice a day. This backup is not impacted by storage or bandwidth constraints and will be used in the common restore scenarios when it is not a real disaster.
- Use the command line tool “wbaadmin.exe” and task scheduler to implement an independent backup schedule for offsite disaster protection, as follows:
 - Pick one of the schedules based on the size of backup and available bandwidth:
 - Quarterly – the backup task runs once at the end of each quarter. This is appropriate when the data set is large (250GB+) and the bandwidth is low (768Kbps up). In the case of a disaster, the customer would be able to go back to the last quarter in the best case and one before the last in the worst case (if other one gets corrupted).
 - Monthly – the backup task runs once at the end of each month. This is appropriate when the data set is modest (100GB or less) and/or the bandwidth is modest (2Mbps up). In the case of a disaster, the customer would be able to go back to end of last month in best case or to the end of month prior to that in the best case.
 - Weekly – the backup task runs once at the end of each week. This is appropriate when the data set is low/modest (50GB) and/or the bandwidth is good (5Mbps up). In the case of a disaster, the customer would be able to go back to end of last week in the best case or to the week prior in the best case.
 - Daily – the backup task runs once a day. This is appropriate when the data set is small (50GB or less) and the bandwidth is good (10Mbps or more). In this case, the customer would be able to go back to yesterday’s state in the best case or to the day before that in the worst case.
 - Create a batch file with following command in it:

```
rd /s \\server\share\grandfather
ren \\server\share\WindowsImageBackup \\server\share\grandfather
wbaadmin start backup -backupTarget:\\server\share -include:C,D,... -noVerify -vssFull
```

NOTE: the remote server path as well as the volumes to include must be specified based on the system being backed up.

- Using Task Scheduler (Start->Programs->Accessories->System Tools->Task Scheduler) schedule the command to run at appropriate schedule (daily, weekly, monthly or quarterly).

Configuring Offsite Mirroring using the Symform Cooperative Storage Cloud

The last step in the data protection framework is to install and configure a Symform Node at the customer site to automatically mirror the local backup sets generated into the storage cloud. This involves few simple steps:

- Create a node in the Symform Cooperative Storage Cloud’s Partner Dashboard by logging on to <https://control.symform.com>
- Download the Symform Node software from the dashboard and install it on the designated Windows computer at the customer site.
- Configure the node as you desire (sync only, sync-contribute, contribute only).
- Add the folder paths to various backup sets under Synchronization Folders in the configuration tool. If these are remote, ensure that these folders and the shares exposing them allow access to the machine account where the Symform Node is running.
- Provide the folder path to a contribution folder, if appropriate. We recommend that this be on a drive (internal or external) with sufficient free capacity to compensate for 2X the size of the backup sets. If these are remote, ensure that this folder and the share exposing it allow access to the machine account where the Symform Node is running.
- Pick and configure a random port and the firewall to enable inbound access on that port, if appropriate (for contribution).
- Throttle and schedule the bandwidth appropriately to ensure that it does not interfere with other Internet access during business and off hours.
- Press Apply and ensure that configuration is successful. Contact support@symform.com if anything fails.

Applying the Data Protection Framework

We will use three different examples to demonstrate how to optimally apply the Data Protection Framework in each case using various tools at our disposal. The three examples we’ll use are as follows:

Customer	Initial Backup Set	Daily rate of change	Local Backup Software Options	Upload Bandwidth Options
Risky Business	10-20GB	10-50MBs	NT Backup or Windows 2008 Backup	768Kbps DSL or up to 2Mbps Cable
Gummy Teeth	100-500GB	50-200MBs	NT Backup or Windows 2008 Backup	768Kbps DSL, up to 10Mbps Cable
Rough Scribbles	500GB+	~1GB or more	NT Backup, Windows 2008 Backup or StorageCraft ShadowProtect	Up to 10Mbps Cable, 20-30Mbps Fiber

Risky Business Insurance Agency

It is primarily a documents based business. Their backup needs are fairly modest with the initial backup around 10-20GB and the rate of change ranging from 10-50MBs per day. The agency is located in rural Midwest, focused on providing insurance services to local farming community. The only Internet

connectivity options are DSL or basic Cable. Being twister country, backup and disaster protection is deeply understood and appreciated.

Gummy Teeth Dentist

This is a dental clinic which is fairly modern and generates a fair amount of data due to extensive use of imaging. Beyond that, it is a fairly standard document and transaction processing business. Due to the images, the initial backup set is large and the daily rate of change is also higher. This is a neighborhood clinic in a suburb of Dallas where the Internet connectivity options include DSL or Cable. Cable offers higher upload capacity but also costs more.

Rough Scribbles Design firm

This design firm is a heavy data generator. Their initial data set is over 500GB and they are typically generating or modifying several 100MBs every day for their client projects. The design firm is centrally located in a San Francisco which gives them a variety of Internet access options including higher tier of Cable as well as fiber connectivity.

Let us now take the framework and apply it to the three examples.

Risky Business

Scenarios	Initial Backup Set	Daily rate of change	Local Backup Software Options	Upload Bandwidth Options
Risky Business	10-20GB	10-50MBs	NT Backup or Windows 2008 Backup	768Kbps DSL or up to 2Mbps Cable

The Backup Schedule Matrix suggests using a monthly G-F-S schedule for Risky Business based on its data and bandwidth profile.

A monthly schedule will require the following backups to get mirrored into the cloud:

- A full backup created at the end of the previous month (Father set). Ideally, this must mirror into the cloud within the first week. The time it takes to mirror the full backup into the cloud represents a vulnerability window where a disaster during that week would mean starting with a month old Grandfather full backup.
- A weekly differential created at the end of each week should make it into the cloud within the first day of the week.
- A daily incremental created at the end of each day should make it into the cloud within a few hours of the day.

Per the upload rate table, @4GB available per day,

- 20GB of full backup would take 5 days if all the data blocks being uploaded are new. However, the full backup in the father set is likely to share commonality with the grandfather full backup.

Assuming 50% blocks are unchanged, week to week, the subsequent full backups should get mirrored within a couple of days.

- Weekly differentials of about 200-300MB would make it within few hours.
- Daily incrementals of about 50MB should make it in within the hour.

Based on this, a monthly GFS scheme is a reasonable conservative backup schedule for Risky Business if they only have 768Kbps DSL.

If on the other hand, Risky Business can change their subscription and get 2Mbps upload bandwidth through local cable operator, they could move to a more aggressive Weekly Backup Schedule.

A weekly schedule will require following backups to make it into the cloud:

- A full backup created at the end of previous week (Father set). Ideally, this must mirror into the cloud within the day. The time it takes to mirror the full backup into the cloud represents a vulnerability window where a disaster during that day would mean starting with a week old Grandfather full backup.
- A daily differential created at the end of each day should make it into the cloud within the first hour of the day.
- An hourly incremental backup should make it in within a few minutes.

Per the upload rate table, @12GB available per day,

- 20GB of full backup would take 2 days in the initial week and get mirrored within the day subsequently assuming at least 50% blocks are unchanged week to week.
- 50MB of daily differentials will get mirrored within the hour.
- <1MB of hourly incremental will get mirrored in about 5 minutes.

Based on this analysis, getting the 2Mbps cable bandwidth and weekly backup schedule is the recommended option. It will cost slightly more than the 768 Kbps DSL but ROI may be justified both by better restorability as well as other productivity gains with increased bandwidth helping the overall Internet experience.

Gummy Teeth Dental Clinic

Customer	Initial Backup Set	Daily rate of change	Local Backup Software Options	Upload Bandwidth Options
Gummy Teeth	100-500GB	50-200MBs	NT Backup or Windows 2008 Backup	768Kbps DSL, up to 10Mbps Cable

Gummy Teeth has an order of magnitude more data to be backed up than Risky Business. It is also generating more data everyday due to things like x-ray images, etc. However, it is a small clinic and wants to manage its cost as much as possible.

There are two levers that can be pulled to get the optimal results – Internet bandwidth and local backup software.

Let’s look at the bandwidth first. 786Kbps DSL is not really an option for getting the initial full backup into the cloud. However, that may be solved by the service provider seeding the data into the cloud from another location (with better bandwidth). At that point, the customer’s site will need to be able to keep up. With 4GB/day capacity, hourly incremental are likely to be too aggressive. Daily incremental backups seem easy to absorb. In the same vein, weekly differentials should work as these differentials should be in the range of 5-10GB max. Subsequent monthly full backups could also be absorbed if the number of blocks changed from older full backups is within 50GB or so. This is where more sophisticated backup software that efficiently handles creating full backups by appending or applying changes to an existing full backup file is useful. Sophisticated backup software such as StorageCraft’s ShadowProtect would however cost few hundred dollars of initial investment and a little more to get upgrades and maintenance.

Next, the best bandwidth option would be 2Mbps Cable. Again, this is not particularly great for getting the initial full backup into the cloud, but good enough for daily and even hourly incrementals. At 12GB/day and 90GB/week capacity, it should be able to handle the daily differentials and subsequent weekly full backups provided full backups are done efficiently. Monthly full backups should be doable using the free backup software such as NT Backup or even Windows 2008 Backup.

Finally, the best bandwidth option would be to go for 10Mbps Cable. It would be more than sufficient bandwidth to handle the weekly backup schedule once the initial full backup is uploaded. The initial full backup may be seeded from the customer site. It should take about 2-10 days depending on the data set. It may be done from another facility with more bandwidth to upload it quicker. A weekly backup schedule would be no problem with 10Mbps bandwidth even with less sophisticated built-in backup software such as NT Backup or even Windows 2008 backup.

The above analysis results in several choices for Gummy Teeth Clinic. Each has a different cost structure with different data protection results:

Bandwidth	Backup Software	Schedule	Comments
DSL 768Kbps	Built-in free versions	Quarterly	Most economical traded with potential loss of up to 6 months of data in a disaster.
DSL 768Kbps	Smarter ISV versions	Monthly	Initial investment, economical on on-going basis. Potential loss of up to 8 weeks of data in a disaster
Cable 2Mbps	Built-in free versions	Monthly	No initial investment, slightly more expensive ongoing basis. Potentially additional productivity benefits of good bandwidth to justify the ROI.
Cable 2Mbps	Smarter ISV versions	Weekly	Initial investment + slightly higher on-going costs provide the most effective coverage.
10Mbps	Built-in free version	Weekly	No initial investment, significantly higher on-going costs for the most effective coverage.

The above analysis suggests that the greatest ROI could be achieved with the option #4 – using 2Mbps Cable and a more sophisticated version of backup software. While this would require some initial investment, it would enable the most aggressive weekly backup schedule and would provide Gummy Teeth the ability to restore to a backup as recent as one hour ago in the best case or two weeks ago in the worst case.

Rough Scribbles

Customer	Initial Backup Set	Daily rate of change	Local Backup Software Options	Upload Bandwidth Options
Rough Scribbles	500GB+	~1GB or more	NT Backup, Windows 2008 Backup or StorageCraft ShadowProtect	Up to 10Mbps Cable, 20-30Mbps Fiber

As a successful design shop, Rough Scribbles has a lot of data that needs protection and is generating much more data on a daily basis. Our experience with the previous two examples already gives us the ability to eliminate a few options for Rough Scribbles. Built-in backup software is not really an option because of inefficiency of resulting backup sets which grows the amount of data that needs to be backed up. Given the larger data set, the impact on Rough Scribble is that much more. Similarly, lower bandwidth options such as DSL or 2Mbps Cable aren't really an option either considering the throughput needed here. This narrows the options to only a couple:

Bandwidth	Backup Software	Schedule	Comments
10Mbps Cable	Smarter ISV versions	Weekly	Initial backup will need to be seeded from another higher bandwidth location.
20-30Mbps Fiber	Smarter ISV versions	Weekly	Initial backup may be done from the customer site itself.

With 10Mbps Cable connectivity and more sophisticated backup software, there is enough bandwidth to absorb the daily rate of change and the weekly changes to full backups. 20-30Mbps Fiber only makes things better. Based on location availability, the pricing of these two options are similar based on price comparisons on services like Comcast Business Class and Verizon FiOS. The key difference is that with 20-30Mbps fiber, there is enough capacity for the initial backup to be uploaded from the customer site – an alternative facility wouldn't yield any better results.

Therefore, either option would work for Rough Scribbles. However, given that they cost about the same, it is always recommended to go with more bandwidth because it creates more headroom to handle the growth and also creates more buffer to handle downtimes.

Summary

If you are an IT Service provider, we invite you to register as a Symform Partner and learn about the benefits of providing these solutions to your customers. If you are an SMB, we encourage you to ask

your IT Service Provider to create a comprehensive, convenient, and cost-effective data-protection solution using Symform's Cooperative Storage Cloud. You can stop worrying about data loss forever.



999 N. Northlake Way, Suite 211

Seattle, WA 98103

(206) 906 9212

www.symform.com

email: info@symform.com