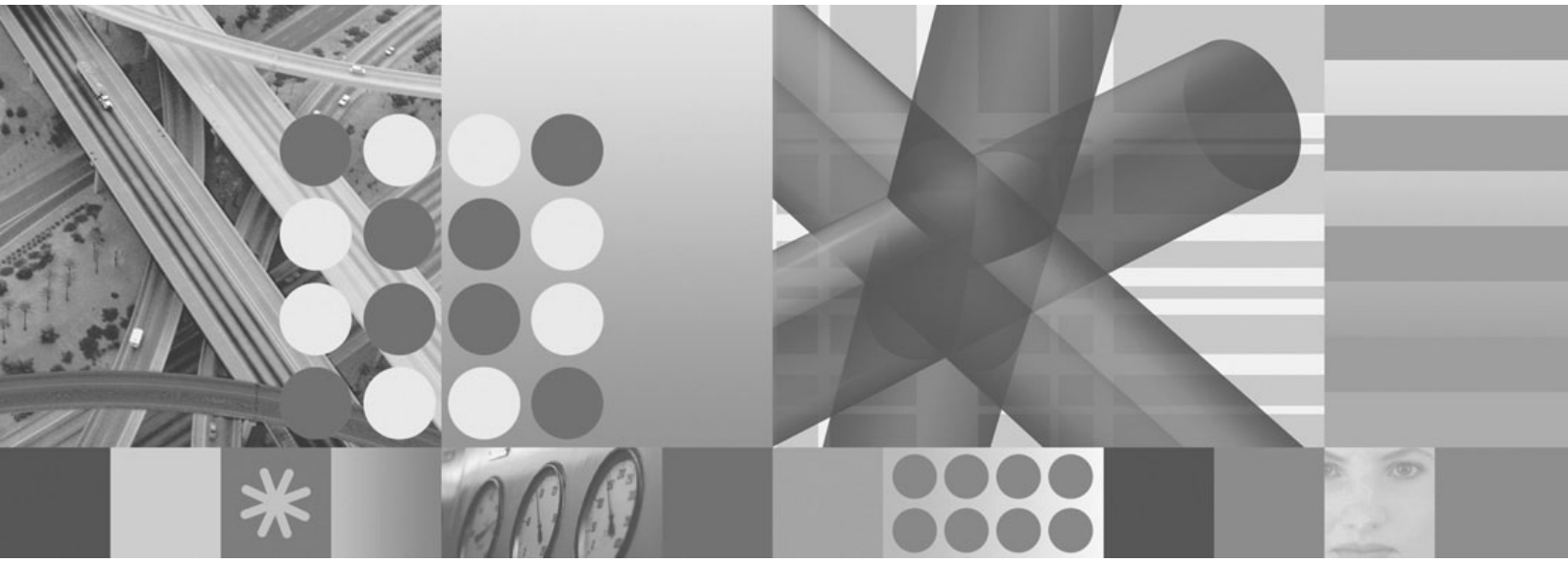
Tivoli®

IBM



IBM Data Deduplication Strategy and Operations

Contents

Executive Overview

With rapid growth in data creation and increasing retention requirements, today's businesses need to control and minimize the amount of data that is created and stored across the Information Infrastructure.

Data reduction is needed to minimize the total amount of storage and network bandwidth required, to improve availability, and to lower Total Cost of Ownership (TCO) in hardware, administration, and environmental costs. Deduplication and other forms of data reduction (compression, Single Instance Store, and so on) are features, not products, and thus can exist in many places in the Information Infrastructure stack. IBM® offers a comprehensive set of data reduction and deduplication solutions for the entire Information Infrastructure.

IBM has been the industry leader in data reduction techniques for decades. IBM invented Hierarchical Storage Management (HSM) and the progressive incremental backup model, greatly reducing the primary and backup storage needs of its customers. Today, IBM continues to provide its customers the most efficient data management and data protection solutions available. ProtecTIER®, the Storage Industry's fastest and highest scaling deduplication solution, and Tivoli® Storage Manager, with its HSM, tape, and progressive incremental efficiencies combined with built-in deduplication, are excellent examples of IBM's continued leadership.

IBM data reduction and deduplication strategy

IBM's Information Infrastructure strategy delivers efficient storage and data management solutions. These efficiencies necessarily include employing a number of data reduction techniques in different parts of the Information Infrastructure to lower TCO. Data deduplication is one of the newer techniques for achieving data reduction. As with any technology, there are benefits and costs associated with different deduplication deployment options. IBM offers coordinated data deduplication capabilities in multiple parts of its storage hardware and software portfolio to enable customer choice through more flexible deployment options:

As a Virtual Tape Library

IBM ProtecTIER's unique, patented deduplication technology is unmatched in the industry in terms of its scalability, performance and data integrity characteristics. ProtecTIER is offered as a gateway or disk-based appliance. It is accessed today as a Virtual Tape Library (VTL). ProtecTIER offers global deduplication across a wide domain of IBM and non-IBM backup servers, applications, and disk. Tivoli Storage Manager works very effectively with ProtecTIER and can exploit ProtecTIER's efficient network replication capability available in ProtecTIER version 2.3. Tivoli Storage Manager operations with ProtecTIER are the subject of the latter part of this document.

In the data protection application

Another option for server side deduplication is Tivoli Storage Manager Version 6 native storage pool deduplication which offers reduction of backup and archive data. Native deduplication helps customers store more backup data on the same disk capacity, thereby enabling additional recovery points without incurring additional hardware costs. Tivoli Storage Manager deduplication is especially applicable in smaller environments where ultimate scalability is not required or where an additional deduplication appliance is not economically feasible. Tivoli Storage Manager deduplication can be used in larger environments if appropriate CPU, memory, and I/O resources are available on the server.

In the collaboration or content application

Lotus® Domino® and IBM Content Manager both deliver application based data deduplication solutions as part of their core features. This helps reduce the amount of primary data stored and created by these applications which in turn reduces the amount of backup data needed to protect them.

As a NAS appliance

IBM n-series appliances offer Single Instance Store (SIS) and fixed block deduplication.

In the network

IBM also partners (for example, with Juniper and Riverbed) to deliver deduplication in WAN optimization appliances which minimize network traffic by deduplicating data before transferring across the network.

As can be seen by these offerings, IBM has a strong suite of data reduction and deduplication solutions available today. IBM is enhancing its data reduction leadership with delivery of a variety of additional deduplication options for reduction of both primary as well as backup data.

IBM's strategy focuses on improving overall storage efficiencies through a combination of data reduction techniques. While data deduplication is a valuable data reduction technique and a key part of IBM's data management solutions, it is just one means of data reduction. For example, minimizing data creation or deleting unwanted data are perhaps the most effective data reduction techniques partly due to the cascading reductions in backup and archive copies. Other data reduction methods used by Tivoli Storage Manager and Tivoli Storage Manager FastBack, such as file and block level progressive incremental backup paradigms, help minimize the operational impacts of backup as compared with other backup paradigms that require frequent full backups.

IBM offers a combination of coordinated data reduction techniques to maximize benefits to the customer. IBM will continue to deliver data reduction solutions, including deduplication, wherever they bring customer value.

How to choose between native Tivoli Storage Manager and ProtecTier deduplication

ProtecTier and Tivoli Storage Manager native deduplication provide two options for server side deduplication of data. Many users ask when they should use one versus the other.

There is no single answer to this question since it depends on the customer's environment and objectives. However, there are a number of decision points and best practices that provide guidance towards the best solution for your circumstances. Here is a summary decision table followed by more detailed discussions of the decision points.

When to use ProtecTIER deduplication	**When to Use Tivoli Storage Manager Native Deduplication**
For medium to large, enterprise environments requiring highest deduplication performance (over 1 GB/sec) and scaling (up to 1 PB storage representing 20+PB of data prior to deduplication)	For large or small environments requiring deduplication to be completely incorporated within Tivoli Storage Manager without separate hardware or software. Sufficient server resources must be available for required deduplication and reclamation processing
For global deduplication across multiple backup servers (Tivoli Storage Manager and others)	For environments where deduplication across a single Tivoli Storage Manager server is sufficient (for example, small or single server environments)
When a VTL appliance model is desired	When customer does not wish to pay any additional licensing costs to enjoy the benefits of deduplication
For backup environments with mostly small files being backed up	For backup environments with mostly large files being backed up

When to choose ProtecTier deduplication

Most large, enterprise environments demanding high deduplication performance and scaling should choose ProtecTIER's industry leading capabilities. ProtecTIER is the best choice in the market today for data deduplication in very large environments. In addition to its unmatched speed and scalability, ProtecTIER's unique technology is extremely efficient in its utilization of memory and I/O, allowing it to sustain performance as the data store grows. This is a major area of concern with most other deduplication technologies.

So, what is a large environment? This also is subjective, but here are some guidelines to consider. If you have 10 TB of changed data to backup per day, then you would need to deduplicate at 115 MB/sec over the full 24 hours to deduplicate all that data. 115 MB/sec is moving towards the practical upper limits of throughput of most deduplication technologies in the industry with the exception of ProtecTIER. Some vendors claim higher than 115 MB/sec with various configurations, but the actual, experienced deduplication rates are typically much lower than claims, especially when measured over time as the data store grows. As a result, most other vendors avoid focusing on direct deduplication performance rates (even with their deduplication calculators) and instead make capability claims based on broad assumptions about the amount of duplicate data (deduplication ratio) processed each day. ProtecTIER, on the other hand, can deduplicate data at 500 MB/sec (1 GB/sec with a 2 node cluster), and can sustain those types of rates over time.

Another consideration is that it is very unlikely that any environment will have a full 24 hour window every day to perform deduplication (either inline during backup data ingest, or post processed after data has been stored). Typical deduplication windows will be more like 8 hours or less per day (for example, during, or immediately after, daily backup processing). For a specific environment, you can calculate the required

deduplication performance rate by dividing the total average daily amount of changed data to be backed up and deduplicated, by the number of seconds in your available daily deduplication window:
`Deduplication_Rate = Amount_of_MBs_Daily_Backup_Data / Number_Seconds_in_Deduplication_Window`

For example, we saw 10 TB of data in a full 24 hour window requires a deduplication rate of 115 MB/sec. Another example is 5 TB of daily data in a 8 hour deduplication window would require a 173 MB/sec deduplication rate (5,000,000 MB / 28,800 sec = 173 MB/sec). Although deduplication performance rates can vary widely based on configuration and data, 100 to 200 MB/sec seems to be about the maximum for most deduplication solutions except ProtecTIER. The scenario of 5 TB in a 8 hour window would lead immediately to a decision of ProtecTIER.

Another view of this is that ProtecTIER can deduplicate 10 TB of data in 2.8 hours. Under very ideal conditions, it would take the fastest of other deduplication solutions 13.9 hours to deduplicate 10 TB of data. Rather than utilize ProtecTIER, you could deploy multiple, distinct deduplication engines to handle a larger daily load like this, but that would restrict the domain across which you deduplicate and minimize your deduplication ratios.

The equation above can assist in determining if you need to go with ProtecTIER for highest performance. The above discussion only considers deduplication processing; note that you also need to consider that Tivoli Storage Manager native deduplication will introduce additional impact to reclamation processing. Also, remember to plan for growth in your average daily backup amount.

- ProtecTIER supports data stores up to 1 PB, representing potentially up to 25 PB of primary data. ProtecTIER deduplication ratios for Tivoli Storage Manager data is lower due to Tivoli Storage Manager data reduction efficiencies, but some ProtecTIER customers have seen up to 10-12:1 deduplication on Tivoli Storage Manager data. Environments needing to deduplicate PBs of represented data should likely choose ProtecTIER.
- Environments that require global deduplication across the widest domain of data possible, should also use ProtecTIER. ProtecTIER deduplicates data across many Tivoli Storage Manager (or other) backup servers and any other tape applications. Tivoli Storage Manager's native deduplication operates only over a single server storage pool. So, if you desire deduplication across a domain of multiple Tivoli Storage Manager (or other backup product) servers, then you should employ ProtecTIER.
- ProtecTIER is the right choice also if a VTL appliance model is desired.
- You should consider using ProtecTIER for environments where mostly very small files (10 KB or less) are backed up (see discussion below).
- Different deduplication technologies use different approaches to guarantee data integrity. However, all technologies have matured their availability characteristics to the point that availability is no longer a salient decision criteria for choosing deduplication solutions.

When to choose native Tivoli Storage Manager deduplication

- Another way of deciding between ProtecTIER and Tivoli Storage Manager deduplication is to evaluate Tivoli Storage Manager server system resources (CPU, memory, I/O bandwidth, database backup size, and potential available window for deduplication). If sufficient server resources can be made available for daily deduplication and additional reclamation processing, Tivoli Storage Manager deduplication is a great option.
- Tivoli Storage Manager deduplication is ideal for smaller environments and for customers who don't want to invest in a separate deduplication appliance. Tivoli Storage Manager can also be used in larger environments if appropriate CPU, memory, and I/O resources are available on the server.

  Like most deduplication technologies, Tivoli Storage Manager deduplication performance rates vary greatly based on data, system resources applied, and other factors. In our labs, we have measured server deduplication rates of 300 to 400 MB/sec with large files (greater than 1 MB) on 8 processor AIX® systems with 32 GB of memory. On 4 processor systems we've seen rates around 150 to 200 MB/sec on large files. Rates with mostly small files or running on single processor systems were much lower. If your environment has mostly large files, our benchmarking would suggest that 100 to 200 MB/sec deduplication rates are possible with Tivoli Storage Manager using 4 or 8 processors.

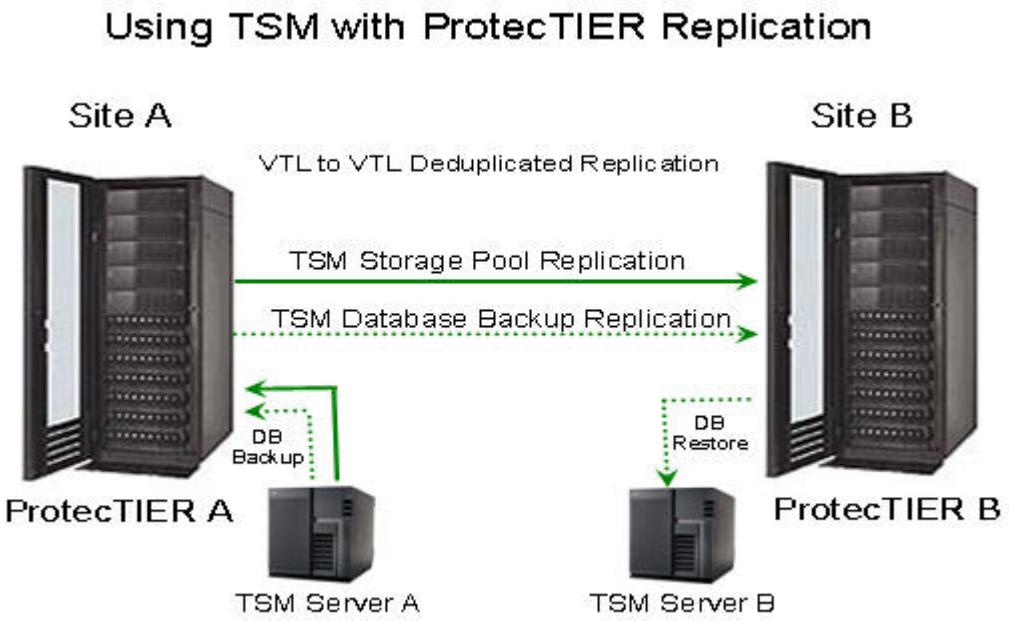
Note: Tivoli Storage Manager native deduplication increases the size of the database, limiting the number of objects that can be stored in the server.

- Native Tivoli Storage Manager is also the right choice if you desire the benefits of deduplication completely integrated within Tivoli Storage Manager, without separate hardware or software dependencies or licenses. Native deduplication provides minimized data storage completely incorporated into Tivoli Storage Manager's end to end data lifecycle management.
- Another important reason for choosing native deduplication is that it comes as part of Tivoli Storage Manager Extended Edition, at no additional cost.

Using Tivoli Storage Manager with ProtecTier replication

Tivoli Storage Manager exploits the new efficient network replication features of ProtecTIER version 2.3. This section outlines the Tivoli Storage Manager and ProtecTIER operations, configurations, and best practices needed to establish electronic Disaster Recovery (DR) solutions using the new ProtecTIER replication capabilities. A general operational flow is given here. Please see the ProtecTIER User's Guide and the Tivoli Storage Manager Information Center for more details on specific commands and configurations.

Many businesses require efficient vaulting and recovery of their data to and from an offsite Disaster Recovery site. This can be achieved electronically through Tivoli Storage Manager use of ProtecTIER replication. As shown in the diagram below, a primary Tivoli Storage Manager server A uses ProtecTIER A as its primary (deduplicated, VTL) storage pool. Regular Tivoli Storage Manager database backups are also performed by Tivoli Storage Manager server A to volumes in ProtecTIER A. ProtecTIER A is configured as the source repository for replication, and ProtecTIER B as the destination repository. Replication is done continuously on Tivoli Storage Manager server A storage pool volumes and on Tivoli Storage Manager database backup volumes. Only unique deduplicated data is replicated between ProtecTIER A and B. Network efficient, automated electronic vaulting of Tivoli Storage Manager server A data is thus achieved across sites. With proper synchronization provided by the operational steps described below, Tivoli Storage Manager server B can then serve as a warm or cold standby server for Tivoli Storage Manager operations in the case of loss of Tivoli Storage Manager server A and/or loss of ProtecTIER A.



Tip: Have Tivoli Storage Manager server A and recovery server B at the same release levels.

Required configuration

1. Using the Grids Management View of ProtecTIER Manager, establish a cross site ProtecTIER replication grid between the primary (A) and recovery (B) ProtecTIER systems.
2. Next, you need to define a ProtecTIER Replication Policy. A ProtecTIER Replication Policy is the only means by which to transfer deduplicated data from a source ProtecTIER repository to a destination ProtecTIER repository. Using the Systems Management view of the ProtecTIER Manager, define a ProtecTIER Replication Policy between the source ProtecTIER system (A) and a destination ProtecTIER system (B). Include in this policy the barcode ranges for all tape cartridges that will be used as Tivoli Storage Manager primary storage pool and database backup volumes. We recommend setting the replication priority to at least Normal.

   Replicated volumes will be introduced at the destination ProtecTIER system in the ″Shelf″ category. Do not worry about the Library or Visibility Switching features for our scenario which has a second Tivoli Storage Manager server B at the disaster recovery site. Visibility Switching is applicable for a scenario where a single Tivoli Storage Manager server is using multiple ProtecTIER systems (for example, perhaps one local and one at a remote site).

   You can monitor ProtecTIER replication policies and activities through the Systems Management view of the ProtecTIER Manager.
3. Do not set the Replication Timeframe unless you need to strictly control when ProtecTIER replication operations occur. Not setting the Replication Timeframe will enable replication to be run automatically whenever a volume changes. Changes result in there being updates on that ProtecTIER A volume which are not yet replicated to its corresponding ProtecTIER B volume (for example, if a primary storage pool volume gets updated during a nightly backup). When a volume changes, ProtecTIER recognizes that and searches for a Replication Policy. If a Replication Policy is found for that volume, it causes a replication of that cartridge to occur automatically to the DR site.
4. Configure sufficient import/export slots on the ProtecTIER systems to handle all storage pool volumes and database backup volumes. The source and target ProtecTIER systems should be configured the same in terms of libraries, drives, import/export slots, and so on.
5. Configure Tivoli Storage Manager server A to have client backups and database backups go to ProtecTIER A storage pool.
6. Tivoli Storage Manager provides a parameter for delaying the reuse of volumes within sequential access storage pools (such as ProtecTIER VTL storage pools). When files are expired, deleted, or moved from a volume, they are not actually erased from the volumes until the specified days have passed. Delaying reuse of volumes can be very helpful for recovery scenarios like the one we are discussing. It ensures no data is overwritten on the storage pool tape volumes and the database backup tape volumes for a minimal number of days. If a database is restored during a recovery or disaster scenario, this **REUSEDELAY** parameter guarantees the integrity of data references. This process will ensure that each storage pool volume's contents are valid as it relates to that database backup.

Tip: Using the DEFINE STGPOOL or UPDATE STGPOOL commands, set the **REUSEDELAY** parameter for the storage pools which use ProtecTIER (A and B). We recommend the **REUSEDELAY** value be set to at least 7 days, but minimally it should be set to 3 Recovery Point Objective (RPO) cycles. For example, if you run daily database backups, your RPO is 1 day and, in that case, the **REUSEDELAY** parameter should be set to 3 (days) at the very minimum. Be careful setting the **REUSEDELAY** parameter too high, as this prevents the server from releasing free space on volumes and a larger **REUSEDELAY** value will increase the size of your ProtecTIER systems. When performing daily database backups, a **REUSEDELAY** value of 7 days is reasonable.

Backup operations

1. Perform regular daily Tivoli Storage Manager backups to the ProtecTIER A system. This is recommended to be done on a daily schedule, but can also be done manually. Whichever volumes

Tivoli Storage Manager updates for backup will be flagged as changed by ProtecTIER. ProtecTIER will then generate replication events for each of the updated volumes.

2. After the daily backups are complete, perform a BACKUP DB command on Tivoli Storage Manager server A to volumes in ProtecTIER A. This can be done automatically through a schedule, or manually, but needs to be done after the daily backups have completed. The database backup time is the key synchronization time to be aware of. We will call this database backup time, T1. Volumes used by the Tivoli Storage Manager database backup will be flagged as changed by ProtecTIER. This will cause replication events to be initiated for each of the updated database backup volumes.

   Tip: Perform full database backups. With the Tivoli Storage Manager 6 database, full database backups are almost just as fast as incremental backups; plus full database backups will facilitate recovery operations. Also, your Recovery Point Objectives (RPOs) will be determined by how often you perform database backups. Once a day is normal and recommended. However, to achieve more granular RPOs, you will need to run database backups more often.
3. After the database backup is complete, save the volume history and device configuration information to files. This is done using the BACKUP VOLHISTORY and BACKUP DEVCONFIG commands. The volume history, device configuration, and server options files should be sent to the Tivoli Storage Manager server B system at the remote site or to some other remote system that will be accessible from remote site B. The database backup, the volume history file, the DEVCONFIG file, and the server options file should all be considered part of a consistency group needed to restore Tivoli Storage Manager appropriately.
4. As each updated storage pool or database backup volume finishes its replication to ProtecTIER B, its last synchronization point times are updated, and a replicated copy of the volume is introduced in the Shelf category of the remote ProtecTIER B. When all storage pool volumes and database backup volumes are replicated, they are then available for recovery as needed at the remote site.

Cold recovery operations

Recovering a Tivoli Storage Manager server in the event of a disaster at the primary location when using ProtecTIER Native Replication for primary storage pool volumes is a matter of entering ″DR Mode″ on the secondary ProtecTIER, recovering the Tivoli Storage Manager server itself using the latest database backup, synchronizing both the defined devices and the storage pool volumes, and, optionally, adding scratch volumes to the inventory if backup services need to be provided at the recovery site.

The steps detailed here cover all of the tasks required to complete this process from scratch (that is, a cold recovery). Many of these steps can be performed in advance resulting in a much smoother overall process. These steps can also be modified to provide a ″warm site″ Tivoli Storage Manager server running and ready to take over the workload. Variations needed for Warm Recovery procedures are described in the next section.

The following information is needed to begin the recovery process:

- IP addresses and system names of the recovery Tivoli Storage Manager and ProtecTIER systems
- Administrator IDs and passwords for the recovery Tivoli Storage Manager and ProtecTIER systems
- Library name on destination ProtecTIER to be used for recovery

Use the following table to document information gathered and required during recovery operations.

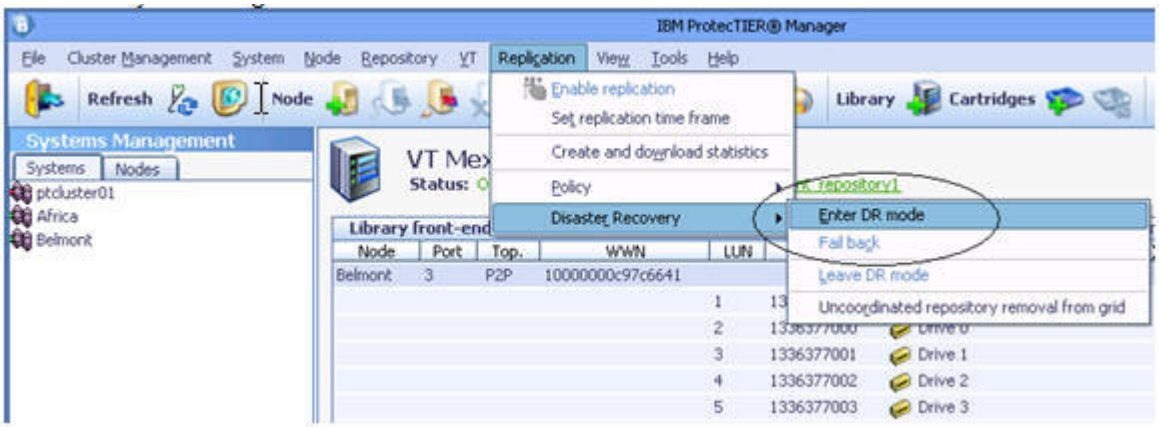
Type of information	Step number	Example	Your values
CSV filename	Step 5c on page 12	Belmont_Aug_29.csv	
Library serial number	Step 6c on page 13	0013363779990402	
Library device address	Step 7c on page 14	dev/smc1	
Date/time of last valid database backup	Step 9b on page 15	2009/08/26	

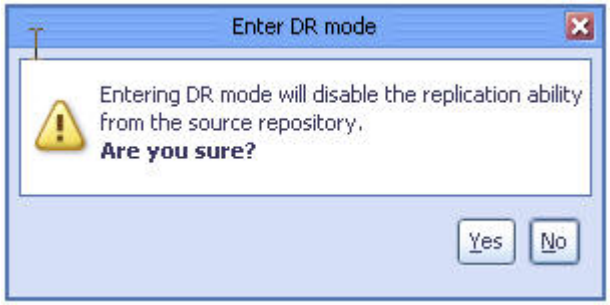
Type of information	Step number	Example	Your values
database backup volume	Step 9b on page 15	AHB116L3	
database backup volume element location and ProtecTIER slot number	Steps 10f on page 17 and 10i on page 17	1142	
ProtecTIER Drive Device Addresses	Step 15b on page 19	/dev/rmt0, /dev/rmt1, and so on.	

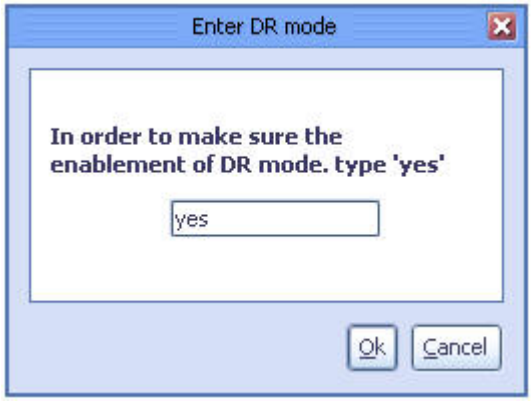
If you encounter a situation where you have lost access to the ProtecTIER A system, you can continue operations from the DR site (B) until the primary site (A) becomes available again. To start working with the remote site as the primary site, you need to do the following.

1. **Log in to the recovery ProtecTIER Manager** Start the IBM ProtecTier Manager at the remote recovery site by logging in as ptadmin with the appropriate password.

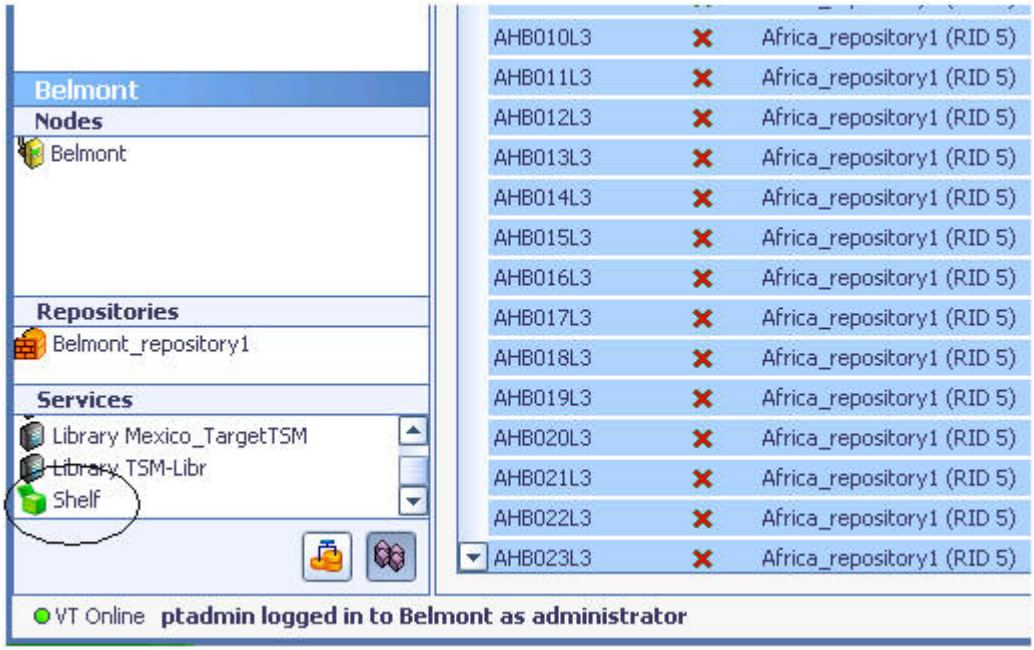
2. **Enter Disaster Recovery mode** Enter Disaster Recovery mode on the ProtecTIER B system at the DR site (B). This is done in order to enable you to failback cartridges to a replacement of the original primary site when it's restored. This is done through the Systems Management view by selecting **Replication** from the menu bar, then **Disaster Recovery**, then **Enter DR mode**. Entering DR mode blocks all incoming replication and visibility switching activities.



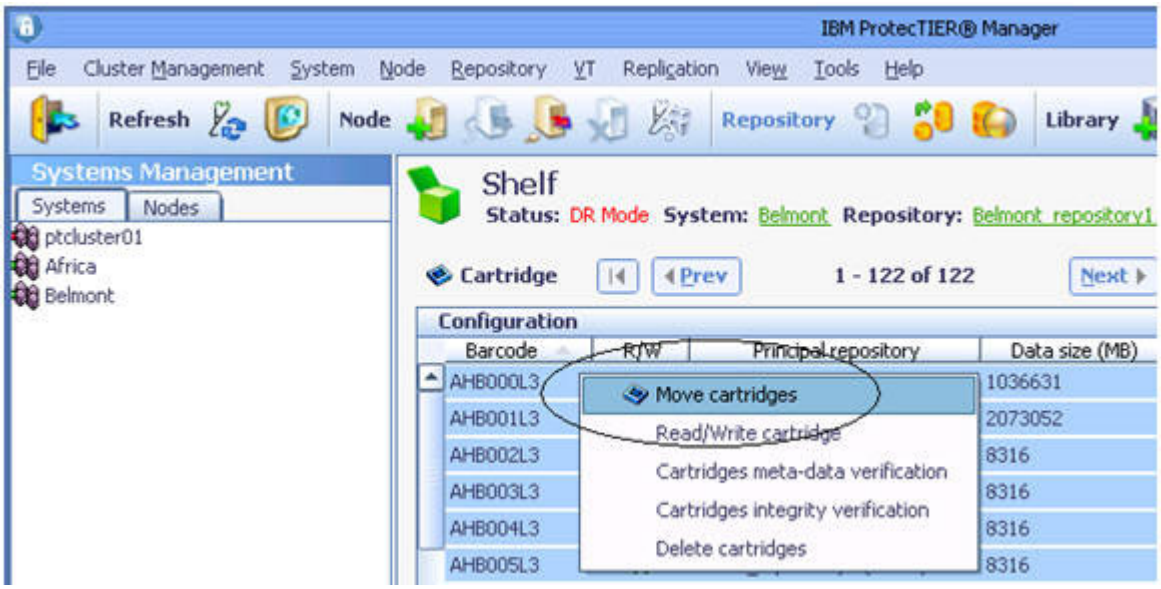




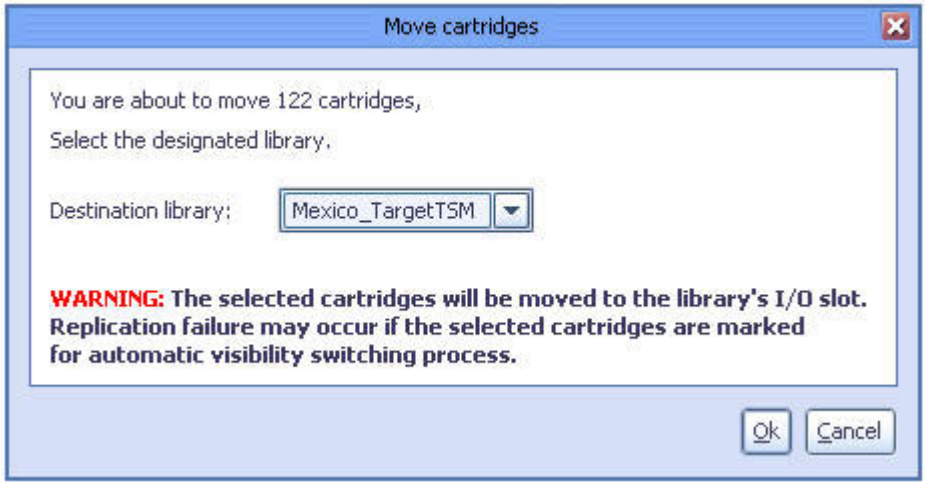
3. **Restore the latest Tivoli Storage Manager control files**
   a. Obtain the latest volume history, device configuration, and server options files from after the last database backup. If you are using Disaster Recovery Manager (DRM), you can locate the newest plan file and explode it.
   b. Place the server options file in the server instance directory.
   c. Place these latest copies of the volume history and device configuration files into the respective locations specified in the server options file.
   d. Overwrite previous versions of these files. You can first save copies of the older versions if you wish.
4. **Move all cartridges from the ProtecTIER shelf into the library import/export station for use by Tivoli Storage Manager**
   a. Select **Shelf** from **Services** on lower left of the ProtecTIER manager.



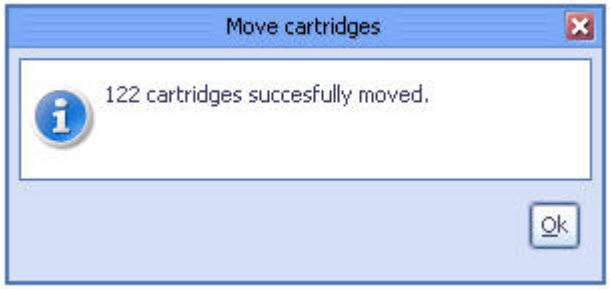
b. Select all cartridges.
c. Right click one of the selected cartridges.
d. Select **Move cartridges**.



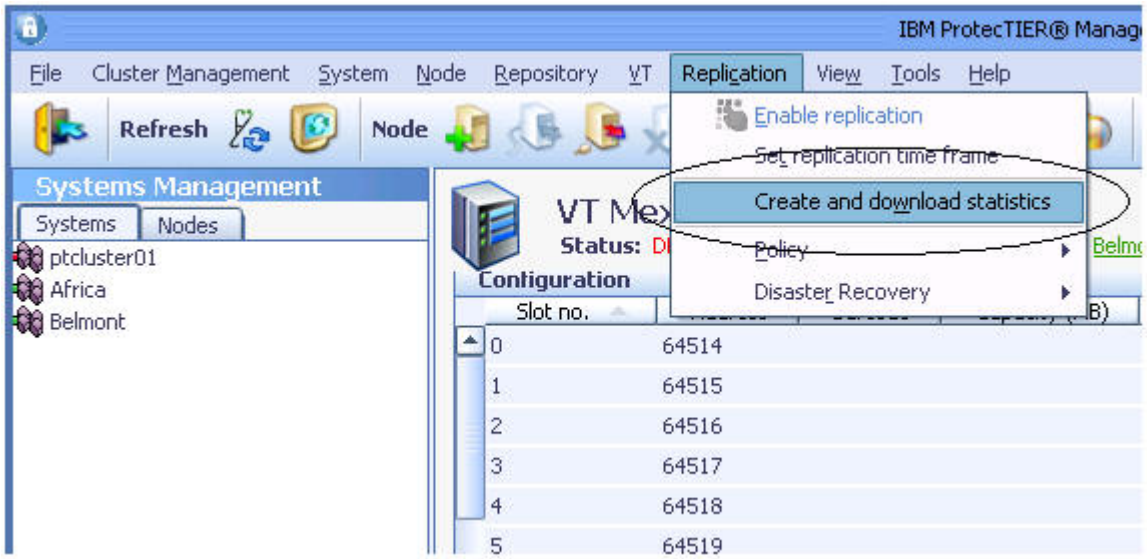
e. In the pop-up window, select the target library to be used, and click **Ok**.

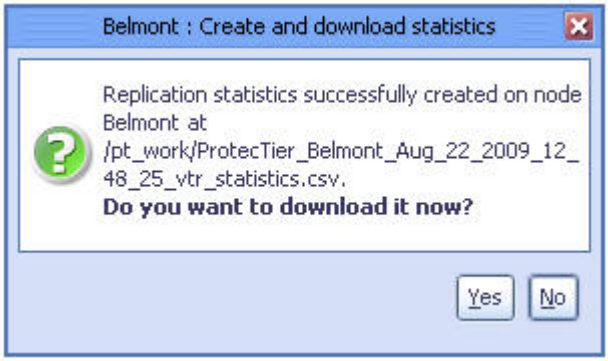


f. Wait for the pop-up window showing all cartridges have been moved.

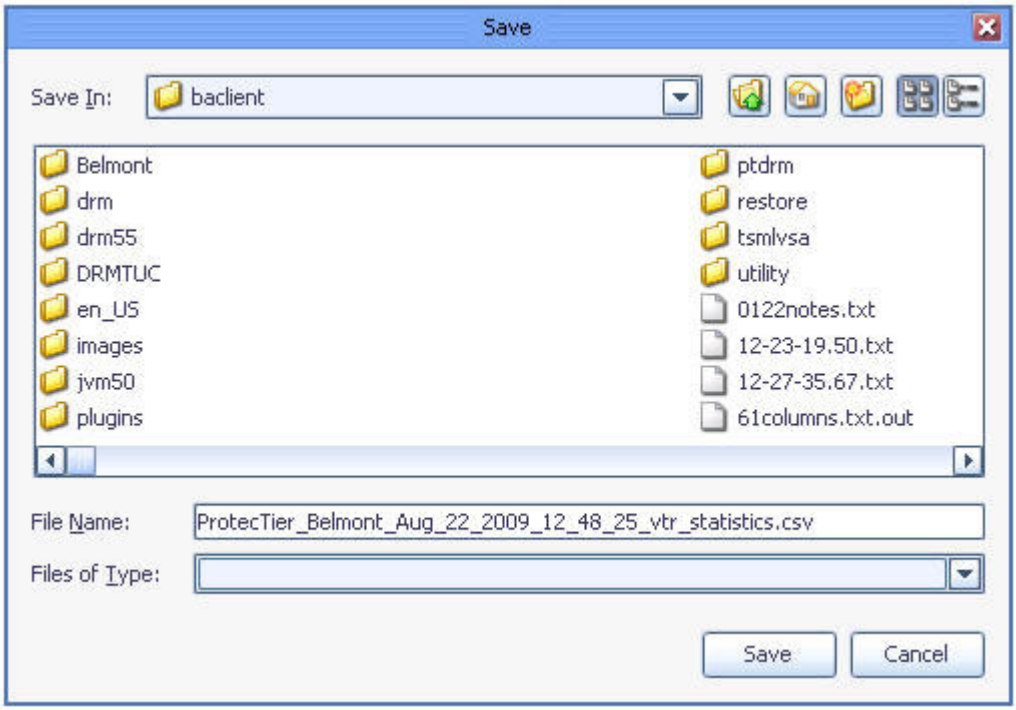


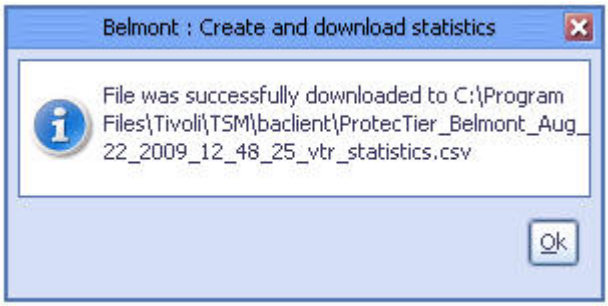
5. **Create a .csv file with replication statistics.**
   a. Obtain cartridge information from ProtecTIER B by doing the following. Use the ProtecTIER Manager on the destination ProtecTIER system, B.
   b. From the menu bar, click **Replication** and **Create and download statistics**.





c. Enter a file name for the information to be saved into, and click **Save**. This saves the cartridge information to a .csv file. Note the name of the file and the destination where you saved it.

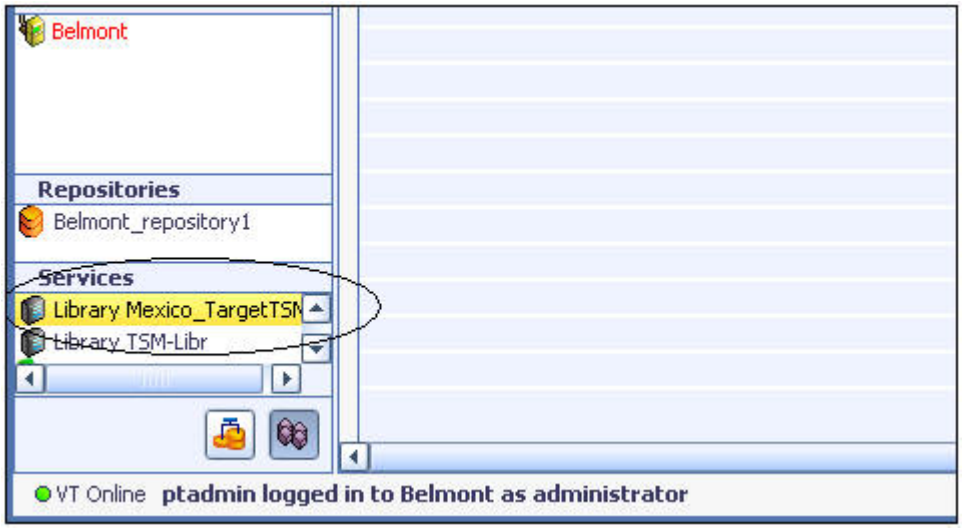




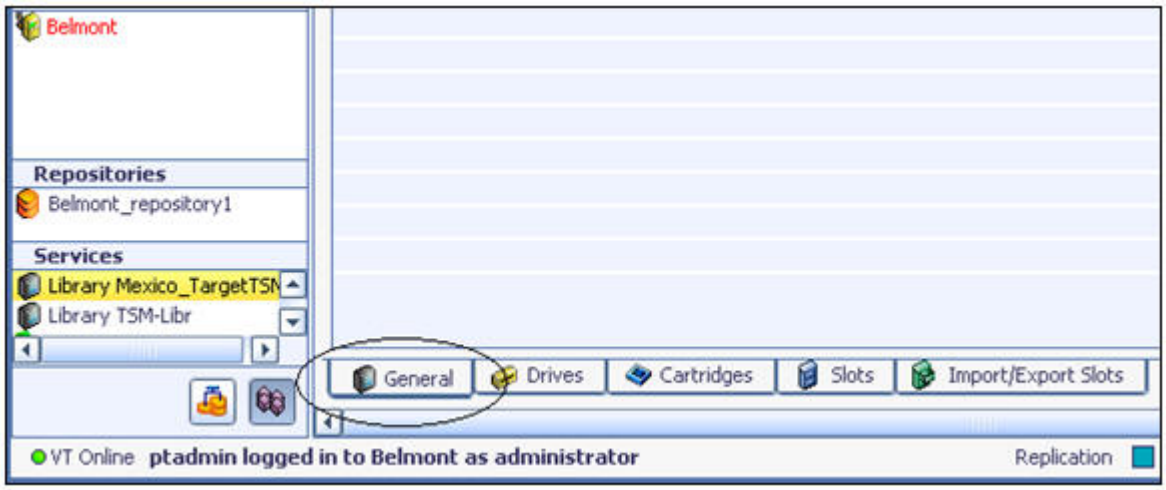
d. Open this newly created .csv file with MS Excel.

6. **Find the serial number for ProtecTIER robotics library**

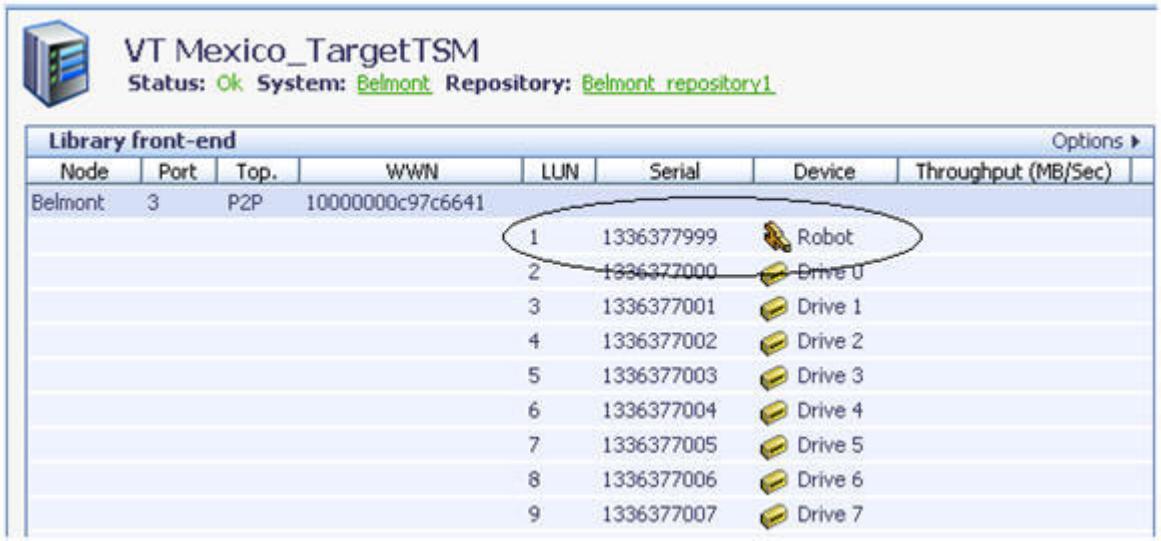
a. Select the recovery library.



b. Open the ProtecTIER **General** tab of the recovery library.



c. Find the value in the **Serial** column for the **Robot** device and note the serial number of the robot.



7. **Find the device address of the Robotics**
   a. This example shows how to find the device address of the robotics on an AIX system. From an AIX command prompt enter:

      ```
      lsdev -Cc tape|grep smc
      ```

   b. For each smc device listed enter:

      ```
      tapeutil -f /dev/smcx inquiry 83
      ```

      where *x* is suffix for each device. The last 16 bytes of the inquiry string contain the device serial number. For example:

      ```
      $ tapeutil -f /dev/smc1 inquiry 83
      Issuing inquiry for page 0x83...

      Inquiry Page 0x83,  Length 48

                0 1  2 3  4 5  6 7  8 9  A B  C D  E F   0123456789ABCDEF
        0000 - 0883 002C 0201 0028 4942 4D20 2020 2020  [.â.,...(IBM     ]
        0010 - 3033 3538 344C 3332 2020 2020 2020 2020  [03584L32        ]
        0020 - 3030 3133 3336 3337 3739 3939 3034 3032  [0013363779990402]
      Serial number ==>                                  -- serial no ---
      ```

   c. Find the serial number that matches the serial number of the robot found in a previous step 9 above. Make note of the device name and serial number. In this example:
      The serial number is **0013363779990402** and the
      The device address is **/dev/smc1**

      .

8. **Update the device configuration file with the Device Address and Serial Number of Recovery Library**
   a. Change to home directory For example

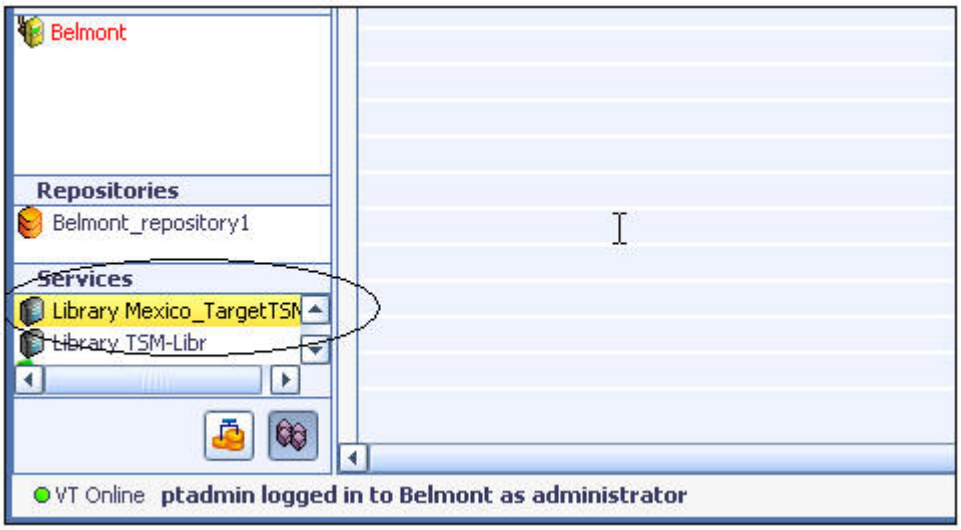
      ```
      *$ cd /home/tsminst1
      ```

   b. Edit the device configuration file using the robotics serial number and device address obtained in the previous two steps. For example:
      1) vi devconf.dat
      2) Find ″DEFINE LIBRARY *lib_name_b* where *lib_name_b* is the library name for the ProtecTIER robot.
      3) Update the ″SERIAL=″ parameter with the serial number found above. For example:
         `SERIAL=0013363779990402`

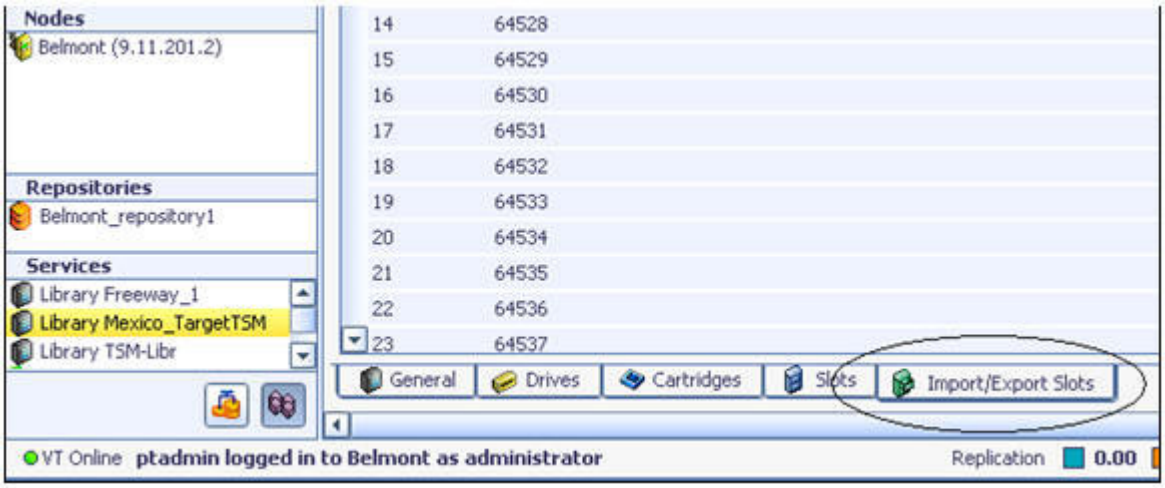
4) Find ″DEFINE PATH *tsm_server_b lib_name_b* where *tsm_server_b* is the name of the server and *lib_name_b* is the library name for the ProtecTIER robot.
5) Update the ″DEVICE=″ parameter with the device address found above For example:
```
DEVICE=/dev/smc1
```
6) Save device configuration file (for now).

9. **Determine the latest valid database backup volumes in the volume history file**
   a. View the latest volume history file and compare that with the volumes that are in the remote ProtecTIER B system (as can be viewed in the newly created .csv file). In the volume history file, identify the volumes involved with the latest database backup. This is done by working backwards through the volume history file searching for the last stanza with a Volume Type of BACKUPFULL. Note the ″Operation Date/Time:″ and the ″Volume Name″.
   b. Compare the ″Operation Date/Time:″ for the volume to the ″source time for last sync point″ field in the .csv file for this volume. If ″source time for last sync point″ for the volume is more recent than the ″Operation Date/Time:″ value (for the database backup time), then this volume is fully synchronized (that is, it has all the updates from the database backup at time T1) and can be used. Note this latest database backup volume names and its ″Operation Date/Time:″
   c. If all of the latest database backup volumes for this database backup are in ProtecTIER B, and all of their ″source time for last sync point″ fields are more recent than the time of the last database backup, use this database backup (in this case, at time T1) for recovering Tivoli Storage Manager server B.
   d. If one or more of the Tivoli Storage Manager database backup volumes is not in ProtecTIER B, or the ″source time for last sync point″ is earlier than the time of the Tivoli Storage Manager database backup (that is, replication of those volumes occurred before this database backup), then use the volume history file to go back to the previous database backup cycle and repeat the process above to determine if that database backup (for example, database backup at time T0) is usable for recovery.
   e. After determining the most recently completely replicated database backup, use that database backup to restore the Tivoli Storage Manager database at site B. From the ProtecTIER Manager interface at site B, select the Tivoli Storage Manager database backup volumes that will be used for recovery.

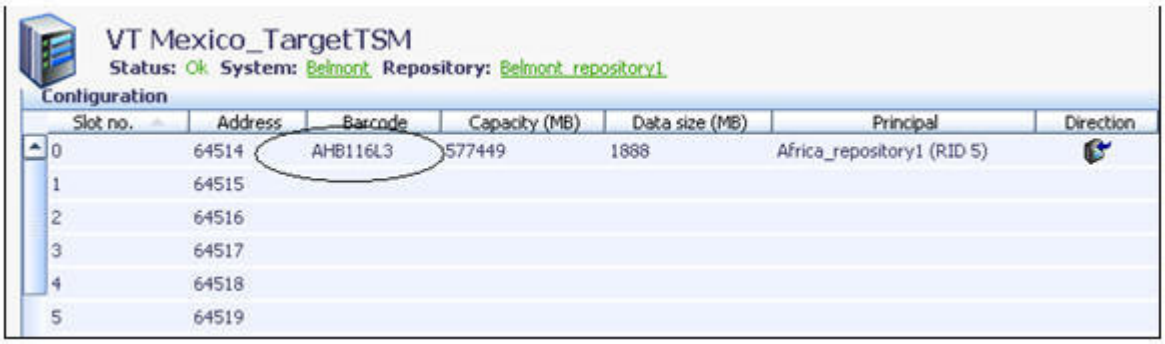
10. **Move database backup volumes from the import/export station to a library slot**
    a. This database backup volumes then needs to be moved from the import/export station to a slot in the recovery library. Note the address of the slot to which this cartridge is moved. This can be done by using the ProtecTIER Manager.
    b. Select the recovery library from **Services** on lower left of ProtecTIER manager.



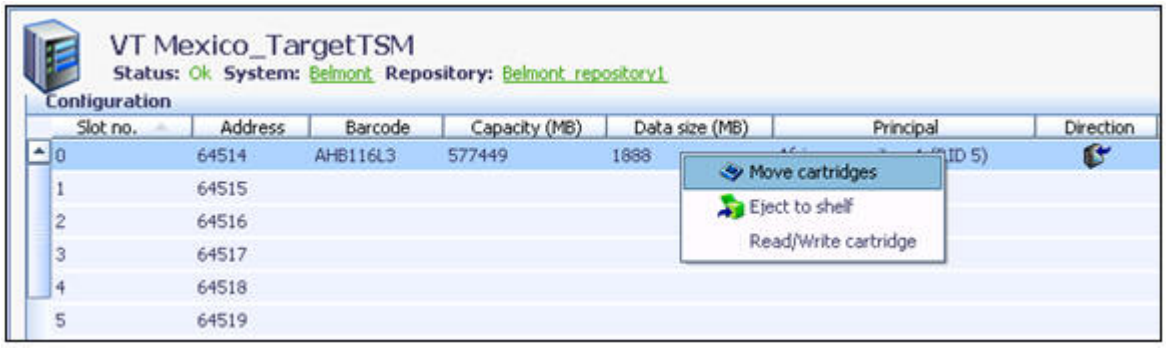
c. Select **Import/Export Slots** from the tab at the bottom of the main pane.



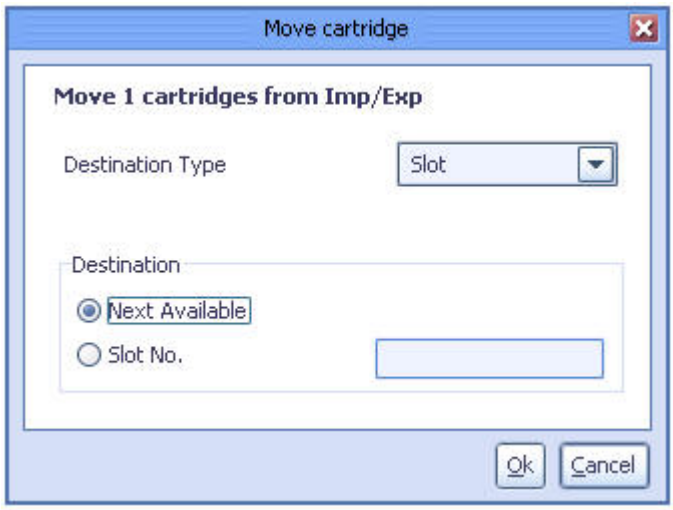
d. Find the database backup volumes selected in the previous step.



e. Right click this cartridge and select **Move cartridges**.



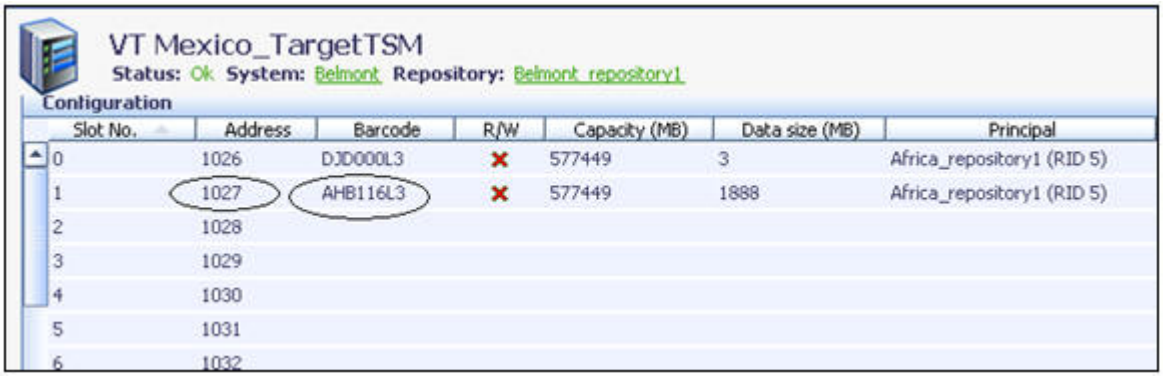
f. In the pop-up window, select a **Destination Type** of **Slot** and **Destination** of **Next Available**.



g. Click **Ok** and wait for pop-up window showing all cartridges have been moved.

h. Select the **Slots** tab from ProtecTIER Manager.

i. Find the database backup volume that was just moved, and note the **Address** column for this volume. This is the element number within the library that the volume was moved to.



11. **Update the element number in the device configuration file** Edit the device configuration file to update the element number for the database backup cartridge to the address value determined in the previous step. This element number is located in the LIBRARYINVENTORY line for this cartridge and is the next to last value in that line.
    a. Edit the device configuration file

       ```
       *$ vi devconf.dat
       ```

    b. Find ″/* LIBRARYINVENTORY SCSI LTO_7 xxxxxxxx nnnn 101*/″ (Where *xxxxxxxx* is the volume selected and *nnnn* is the old element number)
    c. Change *nnnn* to the element number obtained in the previous step.
    d. Save the device configuration file.
12. **Restore the Tivoli Storage Manager database**
    a. At this point, you are ready to recover the database using the database backup from time T1. Ensure you are in the home directory for the server instance. Use the server utility command, DSMSERV RESTORE DB, specifying the date and time from the ″Operation Date/Time:″ found in the volume history file for this volume.

       ```
       dsmserv restore db todate=mm/dd/yyyy totime=hh:mm:ss source=dbb
       ```

13. **Start Tivoli Storage Manager** After the Restore DB completes, start the server using the DSMSERV utility command.

    ```
    dsmserv
    ```

14. **Update the library definitions for the recovery library**
    a. One of the last things that needs to be done is to update the ProtecTIER B library and drive information to Tivoli Storage Manager. This is needed because the newly restored database has different library and drive inventory information (for example, serial numbers) from the ProtecTIER A system at the primary site. This is the cause of the error message when starting the server:

       ```
       ANR8441E Initialization failed for SCSI library xxxxx
       ```

    b. The DSMSERV RESTORE DB utility also changed the server name to the primary site name. To avoid confusing the source and target servers, change the server name back to the recovery site server name and update the library information with the following commands. Substitute your server name, library name, library serial number, and device names as appropriate (for values shown in italics below).

       ```
       set servername tsm_server_b
       ```

    c. Take the library path offline and then update the library name:

       ```
       update path tsm_server_b lib_name_b srct=server destt=libr online=no
       ```

    d. Update the library to avoid relabelling a cartridge if it goes to scratch, and specify the new library serial number.

       ```
       update library lib_name_b relabelscratch=no
       update library lib_name_b serial=0013363779990402
       ```

e. Update the library path with the device name obtained in a previous step and bring the library path back online:

```
upd path tsm_server_b lib_name_b srct=server destt=libr device=/dev/smc1 autod=yes online=yes
```

15. **Delete and define drives and paths in the recovery library** All drives in the recovery library should be deleted and redefined as the serial numbers have changed and possibly the device addresses as well.

a. For each drive in the library, delete the path and drive using these commands:

```
del path tsm_server_b drivex srct=server destt=drive libr=libname
del dr lib_name_b drivex
```

Where *tsm_server_b* is your server name, *lib_name_b* is your library name, and *drivex* is the name of a drive.

b. You can list all available tape drives using operating system commands such as the example shown here for AIX:

```
$ lsdev -Cc tape|grep rmt
rmt0 Available 0A-08-02 IBM 3580 Ultrium Tape Drive (FCP)
rmt1 Available 0A-08-02 IBM 3580 Ultrium Tape Drive (FCP)
rmt2 Available 0A-08-02 IBM 3580 Ultrium Tape Drive (FCP)
rmt3 Available 0A-08-02 IBM 3580 Ultrium Tape Drive (FCP)
rmt4 Available 0A-08-02 IBM 3580 Ultrium Tape Drive (FCP)
rmt5 Available 0A-08-02 IBM 3580 Ultrium Tape Drive (FCP)
rmt6 Available 0A-08-02 IBM 3580 Ultrium Tape Drive (FCP)
rmt7 Available 0A-08-02 IBM 3580 Ultrium Tape Drive (FCP)
```

c. For each *rmt* device do the following:

1) Use the tapeutil command to find the serial number of that drive:

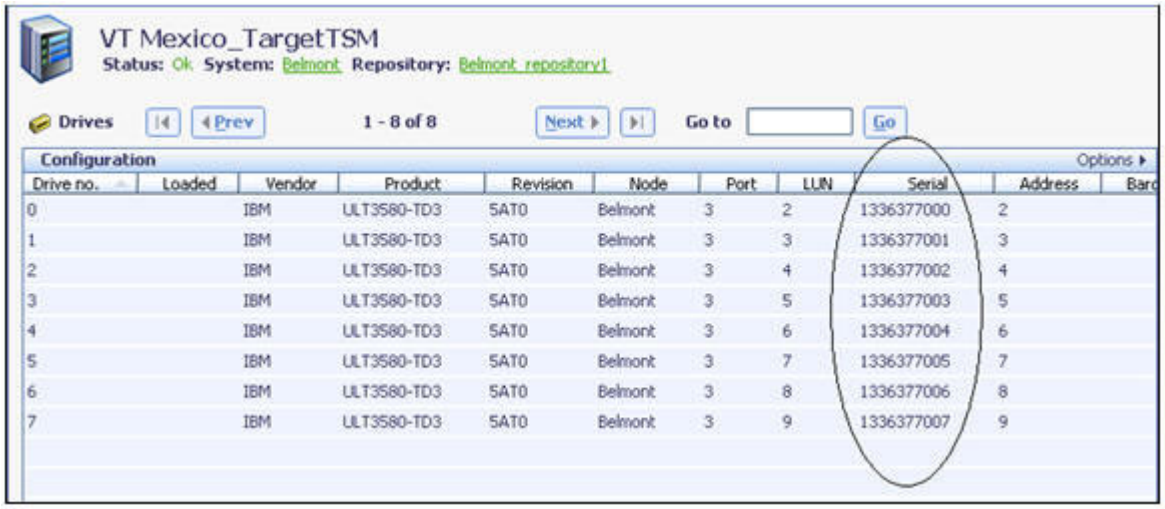
```
$ tapeutil -f /dev/rmt0 inquiry 83
Issuing inquiry for page 0x83...

Inquiry Page 0x83,  Length 74

          0 1  2 3  4 5  6 7  8 9  A B  C D  E F   0123456789ABCDEF
   0000 - 0183 0046 0201 0022 4942 4D20 2020 2020  [.â.F..."IBM     ]
   0010 - 554C 5433 3538 302D 5444 3320 2020 2020  [ULT3580-TD3     ]
   0020 - 3133 3336 3337 3730 3030 0183 0008 2000  [1336377000.â.. .]
   SERIAL NUMBER >                                 1336377000
```

2) Note the serial number returned in bytes x20 through x29 as highlighted above.

3) If this serial number matches a drive serial number listed in ″Drives″ tab for the recovery library using the ProtecTIER Manager.



d. Add this drive (drivex and device /dev/rmty) to Tivoli Storage Manager:

```
def dr lib_name_b drivey
def path tsm_server_b drivey srct=server destt=drive libr= lib_name_b devi=/dev/rmty autod=yes
```

Where *tsm_server_b* is your server name, *lib_name_b* is your library name, and *drivey* is the name of a drive

16. **Restart the Tivoli Storage Manager server and synchronize library volume contents** After the library and all drives have been updated, restart the server. At this point, the existing library inventory reflects the contents of the primary ProtecTIER library, not the recovery library.
    a. Perform an AUDIT LIBRARY command to synchronize the library contents with the database.

    ```
    audit library lib_name_b checkl=b
    ```

    b. Next, issue the CHECKIN LIBVOLUME command to check in all the remaining storage pool volumes from the import/export station as private. This will avoid their being written on:

    ```
    checkin libvol lib_name_b search=bulk checkl=b status=private
    ```

    Alternatively, you can use CHECKLABEL=YES parameter to have the server read the label written on each cartridge, not just the barcode. This will verify that the drives have been properly defined in the previous step.
17. **Mark all existing volumes as read-only** All existing volumes need to be marked as read-only. The cartridges in the recovery library are write protected by ProtecTIER at this time, but may be needed for client restores. Marking all the volumes as read-only allows them to be read by the server when a client restore needs files stored on them, but prevents them from being written on for any reason.

    ```
    upd vol * acc=readonly
    ```

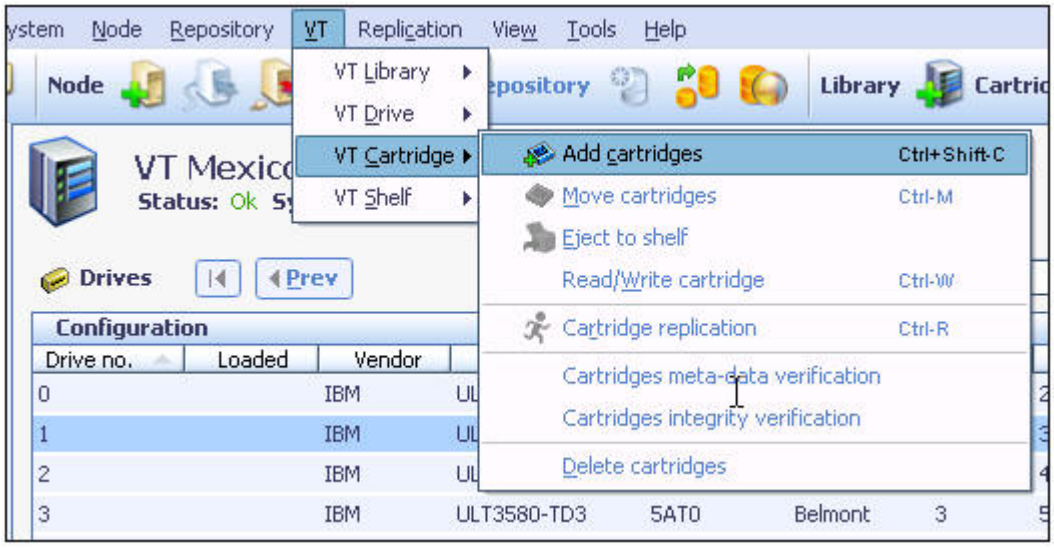
18. **Mark all missing volumes as destroyed** Find all volumes that were not replicated to the recovery server and mark them as destroyed. This will include all disk volumes. For each missing volume, use the UPDATE VOLUME command:

    ```
    Upd vol xxxxx acc=destroyed
    ```

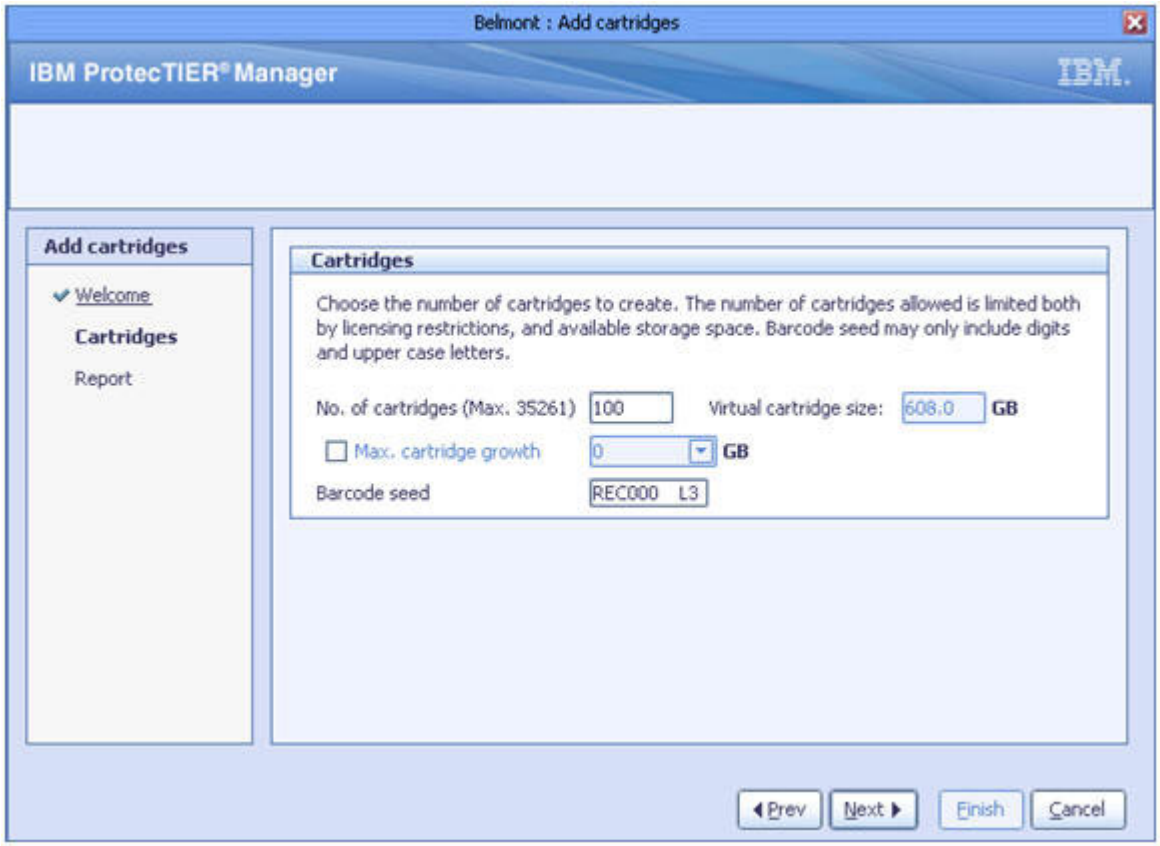
    Where *xxxxx* is a missing volume name.
19. **Find all ″at-risk″ volumes and audit them**
    a. Compare the ″last_write_date″ in the ″volumes″ table to ″source time for last sync point″ in the .csv file.
    b. If a volume was written to by Tivoli Storage Manager at a time later than it was replicated, this volume will need to be audited using the following server command:

    ```
    audit volume xxxxxxxx fix=yes
    ```

    Where xxxxxxxx is the volume name
    c. The audit volume process will mount the cartridge in a drive and read its contents. You should start fewer audits than you have drives available in the recovery library so that drives will still be available for client restore activities.
20. **Prepare new scratch volumes for backup activity on the recovery Tivoli Storage Manager server**
    a. If you intend on providing backup services, not simply recovery services, on this system you will need to add additional scratch volumes for that activity.
        1) From the menu bar on the ProtecTIER Manager, select **VT**.
        2) Select **VTCartridge** .
        3) Select **Add Cartridges**.



b. Select the number of cartridges to create, and specify the barcode seed. The barcode seed should start a range of cartridges that does not conflict with any existing cartridges. It will also be used in the failback policy when returning to the production site. You should also not create more cartridges than you have available slots in the recovery library to hold them.



c. Complete the wizard, the new cartridges appear in the slots in the library.

d. Label these volumes for scratch using the LABEL LIBVOLUME command:

```
label libvol lib_name_b search=yes labels=b checkin=scratch
```

Where lib_name_b is your library name

At this point, your Tivoli Storage Manager server B is ready to be used for client recoveries and new backup operations.

Considerations for Warm Recovery Operations

The process to build a Tivoli Storage Manager server in preparation for a disaster at the primary location when using ProtecTIER Native Replication for primary storage pool volumes is a subset of the tasks shown in the preceding section on Cold Recovery. There are three basic parts to the warm recovery process: 1) define the static components of the recovery configuration, 2) perform daily rebuild of the Tivoli Storage Manager server from the latest database backup, and 3) take over the workload when a disaster occurs. The following steps can be used to accomplish each of the stages of a warm recovery setup. Most of these steps are described in more detail in the previous section on Cold Recovery.

Define the static components

In this stage, define the unique characteristics of the recovery configuration such as device addresses and serial numbers, build a minimal device configuration file to be used for the database restore task, and create a script of administrator commands to adjust the server to match this recovery configuration.

1. Find serial number for the ProtecTIER robotics.
2. Find device address of the Robotics.
3. Find the drives in the recovery library.
4. Determine library slot address for database backup tape.
5. Create a minimal device configuration file with minimal contents of a device class, a library and path, a drive and path, and the location and serial number of the latest database backup volume.
6. Create a script to delete and define drive and path definitions in the recovery library.

Perform daily rebuild of the Tivoli Storage Manager server

Using the output from the steps above, restore the latest database backup, start the restore server, and perform necessary updates to the server configuration to reflect the recovery environment.

1. Connect to the recovery system.
2. Find and explode the newest disaster recovery manager plan file.
3. Restore the latest server control files.
4. Determine the latest valid database backup volumes in the volume history file, and ensure it is fully replicated.
5. Update the minimal device configuration file.
6. Move the database backup volumes from Import/Export station to a slot in the library.
7. Restore the database.
8. Start the server.
9. Execute prepared script to update the library and drive definitions.

Take over the workload when a disaster occurs

Using the output from the steps above, restore the latest database backup, start the restore server, and perform necessary updates to the configuration to reflect the recovery environment.

1. Connect to the recovery system.
2. Find and explode the newest disaster recovery manager plan file.
3. Restore the latest server control files.
4. Determine the latest valid database backup volumes in the volume history file and ensure it is fully replicated.
5. Update the minimal device configuration file.
6. Move database backup volumes from Import/Export station to a slot in the library.

7. Restore the database.
8. Start the server.
9. Execute prepared script to update the library and drive definitions.

Summary

Tivoli Storage Manager can be used effectively with ProtecTIER replication to employ efficient electronic vaulting and disaster recovery operations. Tivoli Storage Manager and ProtecTIER will continue to advance their replication solutions to enable additional efficiencies in customer environments.

With ever increasing amounts of data to be managed and protected, data reduction (including data deduplication) is a critical feature – one that IBM has already deployed in many places in the Information Infrastructure. IBM will continue to deliver additional data reduction capabilities that can be used independently, or in conjunction with each other, for maximum efficiencies and customer value.

Trademarks

IBM, the IBM logo, and ibm.com® are trademarks or registered trademarks of International Business Machines Corporation in the United States, other countries, or both. If these and other IBM trademarked terms are marked on their first occurrence in this information with a trademark symbol (® or ™), these symbols indicate U.S. registered or common law trademarks owned by IBM at the time this information was published. Such trademarks may also be registered or common law trademarks in other countries. A current list of IBM trademarks is available on the Web at "Copyright and trademark information" at http://www.ibm.com/legal/copytrade.shtml.