Josh Krischer & Associates GmbH
Enterprise Servers, Storage and Business Continuity

Native Replication for IBM System Storage™ TS7650 ProtecTIER Gateway & Appliance

Lab Validation Report

Josh Krischer is an expert IT advisor with 40 years of experience in high-end computing, storage, disaster recovery, and data center consolidation. Currently working as an independent analyst at Josh Krischer & Associates GmbH, he was formerly a Research Vice President at Gartner, covering mainframes, enterprise servers and storage from 1998 until 2007. During his career at Gartner he was responsible for high-end storage-subsystems and disaster recovery techniques. He spoke on these topics and others at a multitude of worldwide IT events, including Gartner conferences and symposia, industry and educational conferences, as well as major vendor events.

Josh Krischer

January 2010

Native Replication for IBM System Storage™ TS7650 ProtecTIER Gateway & Appliance

Table of Contents

Native Replication for IBM System Storage™ TS7650 ProtecTIER Gateway & Appliance

Introduction

During the third week of September, Josh Krischer & Associates together with the IBM development lab in Tel-Aviv performed a test of their ProtecTIER Virtual Tape Library solution and the new Native Replication technology. This document outlines the tests performed as well as the specifics of ProtecTIER itself.

On July, 14th, 2009 IBM announced a native IP-based replication capability for IBM System Storage™ TS7650 ProtecTIER Gateways and Appliances. This new enhancement functionality replicates virtual tape cartridges created at a primary site to a remote location. It offers the flexibility to replicate or move all or some virtual tape cartridges at a preset time and at the required priority. ProtecTIER is a leading enterprise-suitable deduplication technology which IBM required from Diligent in 2008 and continues to develop and enhance at a remarkable pace. Unlike many instances where a company's acquisition was followed by an exodus of people, all Diligent staff were integrated into IBM, which explains the smooth and even accelerated road map execution.
This announcement combines one of the currently "hottest" storage technologies with an increased focus on disaster recovery protection and business availability.

Disaster Recovery

The Scandinavian countries were the first to deploy disaster recovery infrastructures, followed by the Benelux countries and central Europe. Because Europe is less vulnerable to natural disasters, synchronous remote copy techniques were selected initially; even today, this is probably still the most commonly used scheme. The US initially lagged Europe in the number of disaster recovery deployments. For many American companies, awareness of the need for business continuity/disaster recovery plans were partially triggered by the tragic events of September 11th, Hurricane Katrina, and the Northeast Blackout of August 2003 (a massive power outage that occurred throughout parts of the Northeastern and Midwestern US, and Ontario, Canada). These examples of extreme regional disasters (sometimes called "global disasters") are relatively rare in comparison to "local disasters", such as power grid failures, fires, building water damage, flooding, and extreme weather conditions. Another trigger was the introduction of various compliance regulations, such as Sarbanes-Oxley or SEC 17a(4).

The 2004 Hurricane season in Florida presented a number of challenges that had not been experienced before anywhere in North America. Given the number and force of the storms, government agencies and businesses were in a constant state of emergency. Recovering from the storms required resources from across the United States.

The era of storage replication in dispersed locations ("remote copy") started in 1994. It was IBM who first introduced remote copy technology on March 1st with the announcement of enhancements to the IBM Storage Control Unit 3990-6. Among the features introduced were IBM's Extended Remote Copy (XRC) and Peer-to-Peer Remote Copy (PPRC), both for the MVS operating system.

IBM's announcement raised interest in disaster recovery planning, in particular among financial institutes; however, IBM was late to deliver the products on time, postponing the general availability by almost a year. Hitachi, which followed IBM compatibility closely with its own versions of comparable features, delayed its delivery as well, and hence, it was EMC that reaped the greatest befits from IBM's marketing efforts. EMC already pioneered its own version of Symmetrix Remote Data Facility (SRDF), which was not compatible to IBM, in1994.

While remote copy techniques debuted on high-end enterprise storage systems in the 1990's, the current decade has produces similar techniques for mid-range storage arrays and virtual tape subsystems.

Recently, a three-site solution has slowly been gaining popularity, particularly within US financial institutes, wherein the third site is located hundreds or even thousands of miles from the main production site. Still today, a common practice of many organizations with limited budgets is to transfer backup tapes to offline vaults or to use third-party disaster recovery service providers such as IBM, Iron Mountain, or SunGard. Manual transportation of cartridges usually causes administration overhead, possible media damage due to shocks or temperature fluctuation, and potential security exposures.[1]

[1] In February of 2005, Bank of America disclosed that it had lost computer tapes containing account information on 1.2 million federal employee credit cards, among them those of U.S. senators, potentially exposing them to theft or hacking.
In June of 2005, Citigroup disclosed that personal information on 3.9 million consumer lending customers of its CitiFinancial subsidiary was lost by UPS while in transit to a credit bureau.

Virtual Tapes and Data Deduplication

IBM's ProtecTIER is a proven enterprise-class virtual tape library with in-line embedded data deduplication. For many years tape was the only media used to store backups. However, in the last decade, disk technologies have been gaining wider acceptance with many organizations. While tapes force to access their data sequentially, disk drives allow direct random access, enabling multiple concurrent backup sessions and better performance. The resulting productivity gains and lowered operating expense can lead to significant time and cost savings.

Virtual Tapes Libraries

The first virtual tape was EMC's CopyCross for MVS (currently called Mainframe VTF), which debuted in 2000 and was developed by EMC's Israeli research lab. In 2002 this lab was spun-off to become Diligent (Diligent's deduplication project started with a new team after the spin-off).
A Virtual Tape Library (VTL) is a virtualization feature that gives disk storage the look and feel of a tape library. Standard backup applications cannot distinguish between a VTL and a physical tape library, which simplifies operation and makes the backup process fully transparent. There are two types of VTLs: stand-alone, which includes only the emulation on disk subsystems, and integrated VTLs containing physical tapes library connected to the VTL.
Regular daily backups can be directed to the VTL to exploit disk technology advantages, but when the data on a virtual tape needs to be archived, the backup application or the VTL control program copies the data from the virtual tape to a physical tape (cloning operation). For restore operations, the major backup applications are able to access backups on the virtual tape or a physical tape mounted in the library. Backup applications track tapes by barcode labels and therefore are confused if two "tapes" have the same barcode label. Hence, the physical tape usually carries a different barcode than the virtual one despite that both contain the same data. If the original virtual tape is deleted, the physical tape can be imported and the backup application's catalog should be updated.

Two major disadvantages of VTL technology are the relative high cost of disks in comparison to tapes and its limited scalability. Two technologies address these problems: high-capacity, low-price SATA type disks and data deduplication.

There is no doubt that disk subsystems are faster and inherently more reliable than tape solutions; therefore, using virtual tapes for backup reduces failure rates, allows for smaller backup windows, and significantly speeds up recovery. Still, tape retains a vital role in backup and recovery operations, especially when it comes to portability and long term data retention requirements.

Data Deduplication

Deduplication aims to eliminate redundant data. Currently it is mainly used in backups and archiving. While it can be employed for any data, this is not recommended due to performance considerations. The deduplication process divides the data into blocks (or chunks) and builds a catalog based on the contents of these blocks, which is stored as metadata. This metadata indexes individual blocks of unique information. Upon new writes (subsequent backup, for example), the deduplication mechanism identifies which data elements are unique, stores them on disk, and updates the metadata. For the non-unique data elements, only references are created without actually storing the data.

There are two basic implementations: inline and post processing. With the inline technique, as data is received by the target device, it is deduplicated in real-time. Post processing, on the other hand, first allows the entire data stream to be stored on disk temporarily, which is then read back off-line and then processed by a deduplication engine.

Data deduplication vendors deploy different methods to detect unique information. Most vendors create a data-dependent fingerprint by applying a hashing algorithm on data blocks and comparing the result with previously calculated hashes. The hash results are usually stored on a disk as well. ProtecTIER uses a unique pattern matching and differencing algorithm (HyperFactor®) that identifies duplicate data. HyperFactor, IBM's patented deduplication technology, first identifies "similar data" using a small key that fits into a deduplication appliance server's memory-resident index and is not stored on disk. If an element looks similar, HyperFactor then performs a bit-level comparison between the new data and the similar data, storing only the bit-level differences. This unique method is more efficient because it dramatically reduces disk accesses for indexing, thus maintaining consistently high performance. In addition, ProtecTIER was designed to deliver 100% data integrity by avoiding the risks associated with hash collisions.

ProtecTIER uses a unique pattern matching and differencing algorithm (HyperFactor®) that identifies duplicate data. HyperFactor, IBM's patented deduplication technology, first identifies "similar data" using a small key that fits into a deduplication appliance server's memory-resident index and is not stored on disk.

Users' Requirements for Deduplication

Users require several important properties from a deduplication subsystem: data integrity, performance (sustained throughput in Mbyte/s per node, number of nodes), scalability, application transparency, product maturity (customer references), price, and ease-of-use.

As mentioned above, most deduplication subsystems use hashing algorithms to develop the unique chunk signature. The risk of *hash collision* (the same signature for two different data chunks) is very low statistically, but such risk may be realized in very large environments; when it does, there is no way of knowing about it until the data needs to be retrieved/restored. For example, in August 2005 a test of SHA-1 (a commonly used 160 bit hash algorithm) detected collisions in 2^{63} operations. Bit comparison techniques such as the one used by ProtecTIER ensure no collisions and were engineered to deliver 100% data integrity.

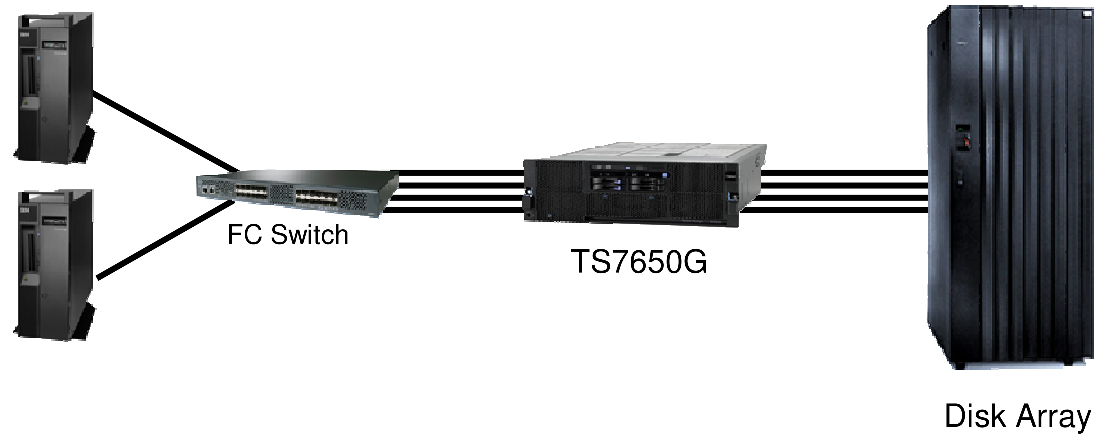
Deduplication performance is the result of two processes: identification of duplicate data, which requires a database or index look-up, and a store operation in the repository.

Scalability depends mainly on the deduplication technique. For example, indexing using SHA-1 hashing requires 20 bytes for the hash signature of each 8 Kbyte chunk; this means that a 100 TByte repository requires ca. 306 Gbytes for its signatures, which is too large to keep in the deduplication appliance's memory and must therefore be stored on the disk subsystem. ProtecTIER's HyperFactor maps 1 PByte of data onto a 4 GByte index, which allows the deduplication appliance to keep the index in its memory, thus increasing performance and scalability as well as enabling overall system throughput to run at speeds greater than 1000 MByte/s.

ProtecTIER's replication works transparently with any of the major backup applications and is flexible as well as easy to use.

IBM System Storage™ TS7650 ProtecTIER Gateway & Appliance

IBM ProtecTIER is available either as a gateway (IBM System Storage™ TS7650G ProtecTIER Deduplication Gateway) that connects to an external (also third party) disk array, or as a fully-bundled appliance (IBM System Storage TS7650 ProtecTIER Deduplication Appliance) with internal storage (see **Error! Reference source not found.**). It supports full, incremental, and differential backups while integrating seamlessly with real physical tape library resource management. When it is a target of the backup application, a ProtecTIER subsystem presents itself as a single tape library (or several libraries) to the system. The backup application manages the "cartridges" within a ProtecTIER system as if they were real cartridges. It can perform read and write operations as with physical tape. In addition, it can import and export cartridges, track the cartridges with barcodes, and perform many other tape library operations.



- **High performance (>1000MB/s sustained)**
- **True active-active Cluster available**
- **Highly available**
- **High capacity**
- **Highly scalable**
- **Easily manageable**

Figure 1: IBM System Storage TS7650G ProtecTIER Deduplication Gateway Solution

The combination of deduplication, which dramatically saves disk space and leads to significantly lower transmission bandwidth requirements, with Native Replication, presents a powerful disaster protection solution at reduced costs compared to regular disk-based replication.

Native Replication for IBM ProtecTIER Deduplication Solutions Overview

In September 2009 IBM launched IP-based asynchronous replication to allow ProtecTIER users to deploy disaster recovery schemes at lower costs. As mentioned above, disaster recovery solutions based on physical tapes with human intervention provide relatively low reliability and introduce security risks. Performing disaster recovery testing from tapes is complicated and slow. In many cases, the costs of storage and particularly data transfer, prevents users from deploying disk-based replication. The combination of deduplication, which dramatically saves disk space and leads to significantly lower transmission bandwidth requirements, with *Native Replication,* presents a powerful disaster protection solution at reduced costs compared to regular disk-based replication. See Figure 2 for a sample ProtecTIER replication solution. The replication ability is a licensed feature included in the latest version (2.3) of the software. Currently installed ProtecTIER Gateway and Appliance systems only require a software upgrade; older systems may need more NICs for the replication connection (for example, 6 Ethernet ports are required for a clustered configuration). The licensing is based on a tiered pricing structure. The replication process is managed from a user-friendly GUI with the ability to create policies, select individual, or ranges of virtual cartridges and to set priorities. The virtual tapes can be cloned to physical tapes at the remote site for long term vaulting.

ProtecTIER's Native Replication has many advantages compared to disk array hardware or host-based disk replications. ProtecTIER's Native Replication understands and manages virtual tapes, whereas disk replications replicate LUNs. A virtual tape can be spread over several LUNs. Not every asynchronous replication technique is able to re-synchronize and to continue operation after a network outage; in many disk subsystems, the replication process has to be re-started in such a case. ProtecTIER's Native Replication supports internal and external heterogeneous storage arrays with a single management interface.

The tiered licensing scheme is based on deduplicated data capacity, which in most cases is significantly lower than replication licenses based on native storage capacity.

Replication management allows a high degree of automation by employing user-defined policies. The fine granularity allows users to choose the time for replication and to set policies for an individual cartridge, a pool of cartridges, or even an entire virtual tape library. Replication is performed asynchronously, at the logical cartridge level, and replication progress is tracked and reported at the cartridge level through the ProtecTIER management interface. To ensure data integrity, data validation of the transferred data is performed at the second site as part of the replication operation and prior to making the virtual cartridge available.

Because replication at the ProtecTIER level is transparent to backup applications, ProtecTIER's replication function allows synchronization with the backup application similar to normal tape management methodologies. It is important that a particular cartridge at any given time is only visible to the backup application in one location regardless of how many replicas exist. To ensure that, ProtecTIER uses a "visibility control" function that allows the user to determine in which location the cartridge should be mounted and visible to the backup application. This is achieved by utilizing the import/export slots of the virtual libraries and exactly mimics the operation of physical tape management.

Like many major replication techniques, ProtecTIER‘s Native Replication supports fail-over and fail-back operations. In the case of fail-over, the ProtecTIER subsystem allows fast restoration of the production data at the secondary site. Once the operational data has been restored, the ProtecTIER continues its usual tasks such as backup or archiving at the secondary site. When the primary site becomes operational again, the users may replicate any new data created at the secondary site back to the primary and return the primary to its original status.

It is quite a common practice to place backup or archiving tapes in a vault at a remote location for longer periods of months or even years. ProtecTIER users can replicate from the primary site to the secondary site, and then move/copy the data from the disk-based repository onto physical tape cartridges in a process called *cloning*. The *visibility control function* of ProtecTIER simplifies this operation. Once the cartridges complete replication to the second site, users may clone these cartridges to real physical tape using the native backup application tape copy or vaulting functions. This allows the backup application to remain in control of the end-to-end process, and maintain its catalog of all cartridges, the associated data, and the storage location.

A unique capability of ProtecTIER's replication is priority setting and policy management. A user can create a cartridge-level replication policy. The priority mechanism lets the user determine the order in which policies are carried out

Unique Features

A unique capability of ProtecTIER's replication is priority setting and policy management. A user can create a cartridge-level replication policy. The priority mechanism lets the user determine the order in which policies are carried out. Finally, a user can set a time-window for when the replication is to take place (for all policies). The policies can be set for different virtual cartridges or ranges of barcodes, allowing setting of different priority levels for the replication.

Users can monitor the network status, the replication relations between the sites, and the data throughput rate of replication by using the ProtecTIER Manager GUI. In addition to management functions, the GUI also provides statistical information and a graphical representation of the cartridges participating in the replication process, replication progress, etc.

Future developments

The IBM development lab in Tel-Aviv continues to develop additional options for ProtecTIER Native Replication. The next enhancement scheduled for early 2010 should include "many-to-one" replication. This option is aimed at organizations with remote offices or branches. It will allow replicating the backups of their remote sites to their central/DR location. Following developments will deliver a "many-to-many" mesh replication configuration and bi-directional replication, which will allow multiple sites or data-centers to protect each other.

In parallel to these replication developments, IBM labs continue to improve the integration of ProtecTIER with major backup applications and with other IBM subsystems. The unique HyperFactor deduplication engine software used in IBM's TS7650 ProtecTIER family is one of the important building blocks in IBM's comprehensive storage strategy.

Economical considerations

By using the native replication and deduplication capabilities of ProtecTIER, users can lower their CapEx and OpEx investments significantly. The major savings come from smaller disk space and lower transmission bandwidth requirements. Smaller disk capacity means lower investment, service, energy, and storage management costs. In addition to financial savings, carbon-dioxide emissions are also lowered, yielding a better environmental footprint. ProtecTIER replicates just the deduplicated data, which can be up to 25 times smaller than the raw data, meaning that the required transmission bandwidth can be significantly lower and cheaper. The replicated data is compressed, which again reduces data traffic. The replication feature license scheme is based on tiered deduplicated capacity, which in most cases is much lower than a disk replication license based on raw capacity. The transmission costs in data replication are sometimes the biggest part of disaster recovery OpEx. Table 1 compares the transmission costs between a traditional VTL and a ProtecTIER deduplication solution with Native Replication (a deduplication factor of 10 is assumed, which in my experience is a typical value).

	Traditional VTL	**ProtecTIER**
Daily backup (Tbytes)	4	4
Backup window (hrs)	16	16
Data transferred (Tbytes,daily)	4	0.4
Required Bandwidth (Mbps)	582.4	58.2
Required connection	OC12	OC3
Monthly costs (3 years contract) VERIO-NTT)	$27,999	$7905

36 months	$1,007,964	$284,580
Delta communication costs, 36 month period	**+$ 723,384**	

Table 1: Bandwidth costs comparison

Test Configuration

The test environment was comprised of 3 servers, one with external IBM DS8000 disks and two with internal disks, running Red Hat Linux and Symantec NetBackup. Two IBM TS7650 ProtecTIER appliance clusters were connected by four 1Gbps links. ProtecTIER replication can run between any combination of single or clustered gateways in the primary and the secondary, however, the test was performed on a cluster configuration since that is the most complicated scheme. Several tests evaluating the full functionality of native replication were conducted.

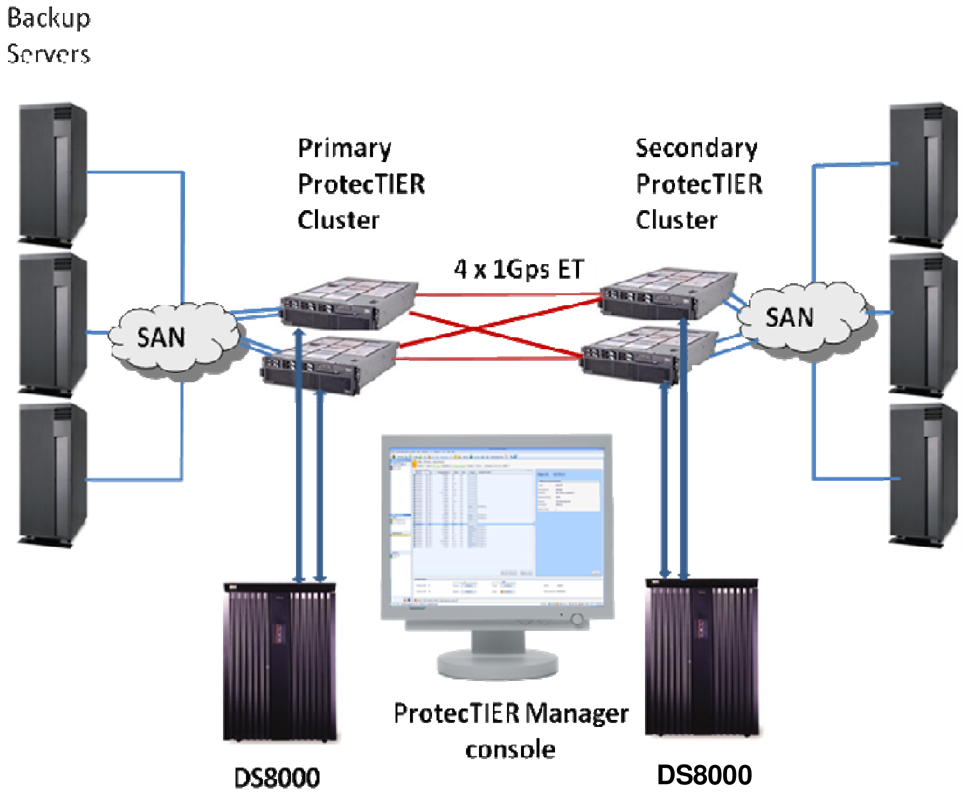
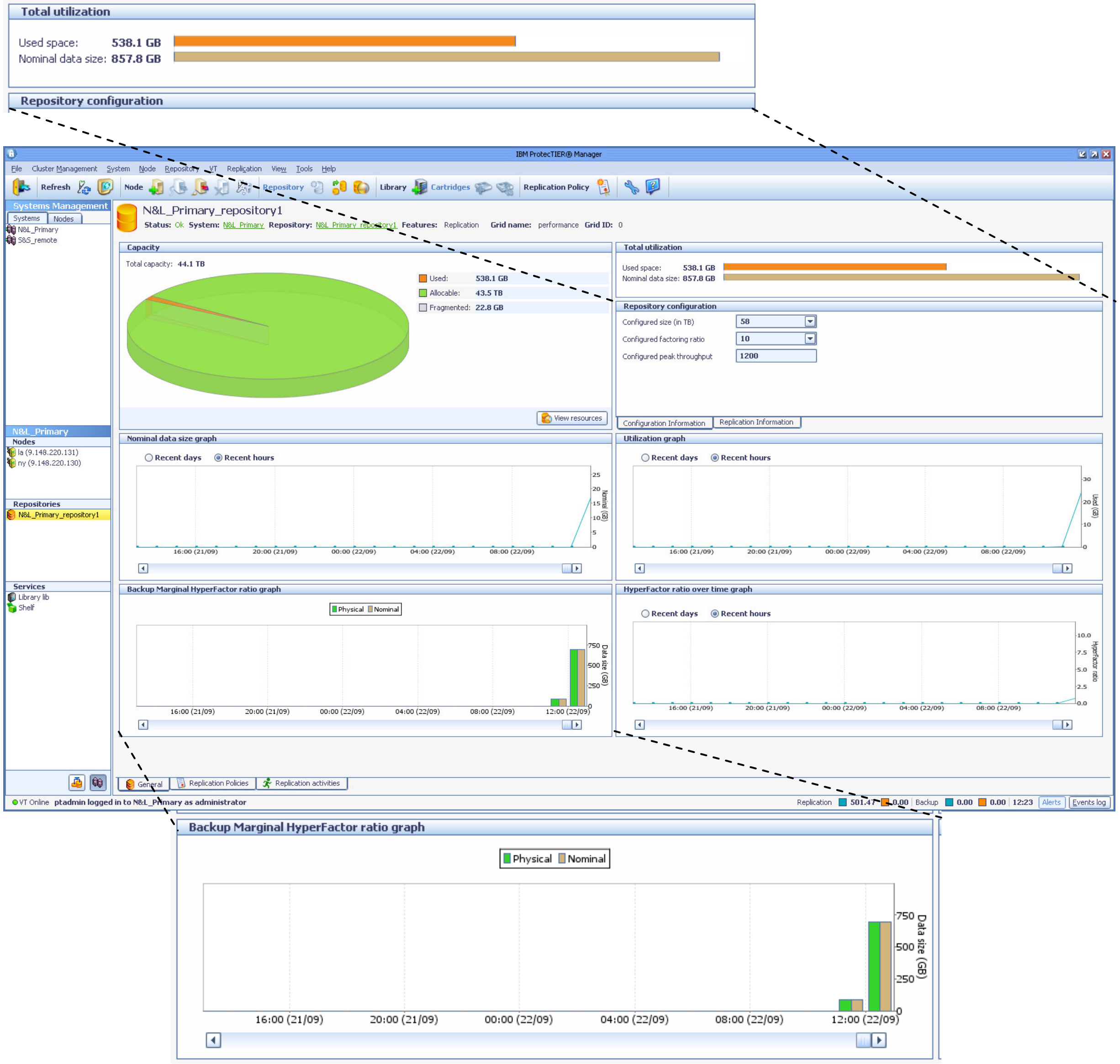

Backup Servers
Primary ProtecTIER Cluster
Secondary ProtecTIER Cluster
4 x 1Gps ET
SAN
SAN
ProtecTIER Manager console
DS8000
DS8000

Figure 2: ProtecTIER replication (test configuration)

Test1 - Creating a First-Generation Backup at the Primary Site.

In this test the primary was backed up for the first time. The screen shot in **Error! Reference source not found.** represents the ProtecTIER manager screen. It shows that the nominal data was 857 GByte, but due to compression, only 538.1 GByte of disk capacity was used. Because this was the initial backup, no deduplication took place, which can be seen in the bottom part of the figure–the HyperFactor ratio graph shows that the physically stored data equals the nominal data.



Total utilization

Used space: 538.1 GB
Nominal data size: 857.8 GB

Repository configuration

N&L_Primary_repository1

Status: Ok System: N&L_Primary Repository: N&L_Primary_repository1 Features: Replication Grid name: performance Grid ID: 0

Capacity

Total capacity: 44.1 TB

Used:	538.1 GB
Allocable:	43.5 TB
Fragmented:	22.8 GB

Repository configuration

Configured size (in TB)	58
Configured factoring ratio	10
Configured peak throughput	1200

N&L_Primary
Nodes
la (9.148.220.131)
ny (9.148.220.130)

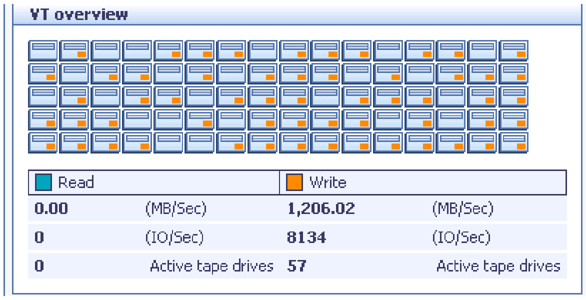
Repositories
N&L_Primary_repository1

Services
Library lib
Shelf

VT Online ptadmin logged in to N&L_Primary as administrator

Replication 501.47 0.00 Backup 0.00 0.00 12:23

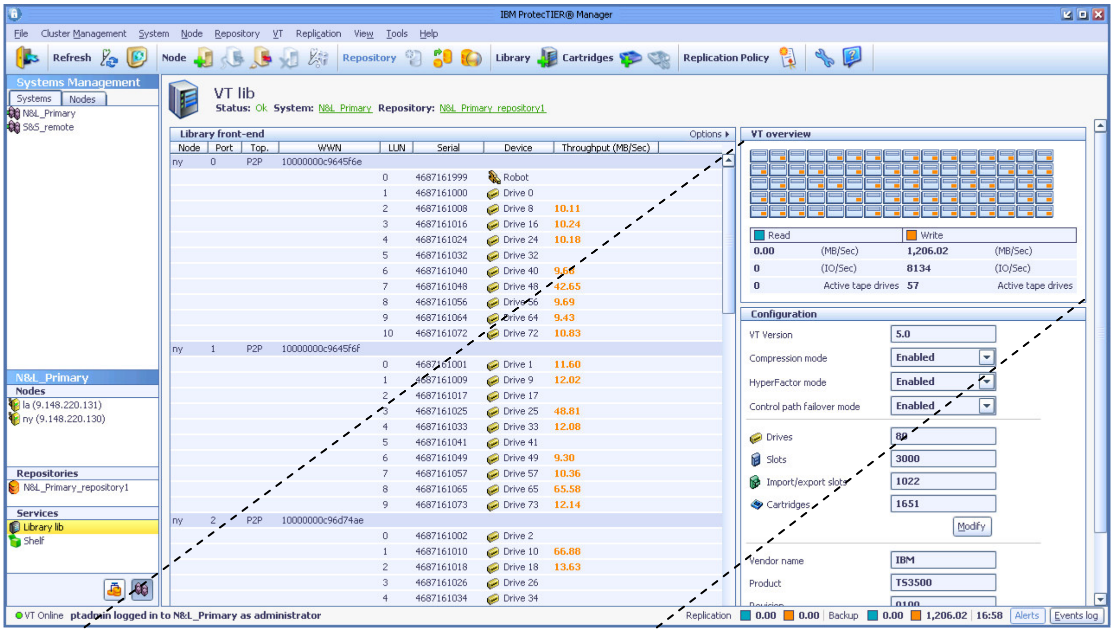
Figure 3: ProtecTIER management screen after first backup.



VT overview

Read		Write	
0.00	(MB/Sec)	1,206.02	(MB/Sec)
0	(IO/Sec)	8134	(IO/Sec)
0	Active tape drives	57	Active tape drives

Figure **5** shows the emulated tape library tape drives and virtual tapes. Note the actual data inline backup/deduplication throughput of 1206MB/s shown in the blown-up box and in the bottom right of the screen.



VT lib
Status: Ok **System:** N&L_Primary **Repository:** N&L_Primary_repository1

Library front-end

Node	Port	Top.	WWN	LUN	Serial	Device	Throughput (MB/Sec)
ny	0	P2P	10000000c9645f6e				
				0	4687161999	Robot	
				1	4687161000	Drive 0	
				2	4687161008	Drive 8	10.11
				3	4687161016	Drive 16	10.24
				4	4687161024	Drive 24	10.18
				5	4687161032	Drive 32	
				6	4687161040	Drive 40	9.66
				7	4687161048	Drive 48	42.65
				8	4687161056	Drive 56	9.69
				9	4687161064	Drive 64	9.43
				10	4687161072	Drive 72	10.83
ny	1	P2P	10000000c9645f6f				
				0	4687161001	Drive 1	11.60
				1	4687161009	Drive 9	12.02
				2	4687161017	Drive 17	
				3	4687161025	Drive 25	48.81
				4	4687161033	Drive 33	12.08
				5	4687161041	Drive 41	
				6	4687161049	Drive 49	9.30
				7	4687161057	Drive 57	10.36
				8	4687161065	Drive 65	65.58
				9	4687161073	Drive 73	12.14
ny	2	P2P	10000000c96d74ae				
				0	4687161002	Drive 2	
				1	4687161010	Drive 10	66.88
				2	4687161018	Drive 18	13.63
				3	4687161026	Drive 26	
				4	4687161034	Drive 34	

VT overview

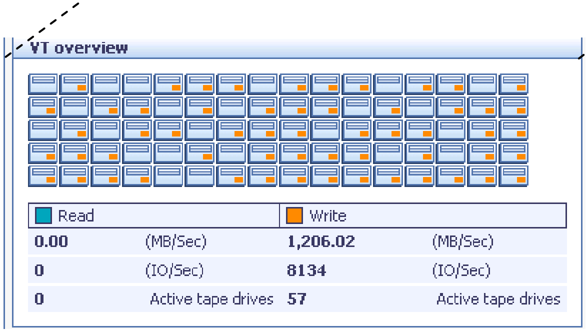
Read		Write	
0.00	(MB/Sec)	1,206.02	(MB/Sec)
0	(IO/Sec)	8134	(IO/Sec)
0	Active tape drives	57	Active tape drives

Configuration

VT Version: 5.0
Compression mode: Enabled
HyperFactor mode: Enabled
Control path failover mode: Enabled
Drives: 80
Slots: 3000
Import/export slots: 1022
Cartridges: 1651
Modify
Vendor name: IBM
Product: TS3500
Revision: 0100

VT Online ptadmin logged in to N&L_Primary as administrator

Replication 🟦 0.00 🟧 0.00 | Backup 🟦 0.00 🟧 1,206.02 | 16:58 | Alerts | Events log



VT overview

Read		Write	
0.00	(MB/Sec)	1,206.02	(MB/Sec)
0	(IO/Sec)	8134	(IO/Sec)
0	Active tape drives	57	Active tape drives

Figure 5: Display of an emulated IBM 3500 tape library

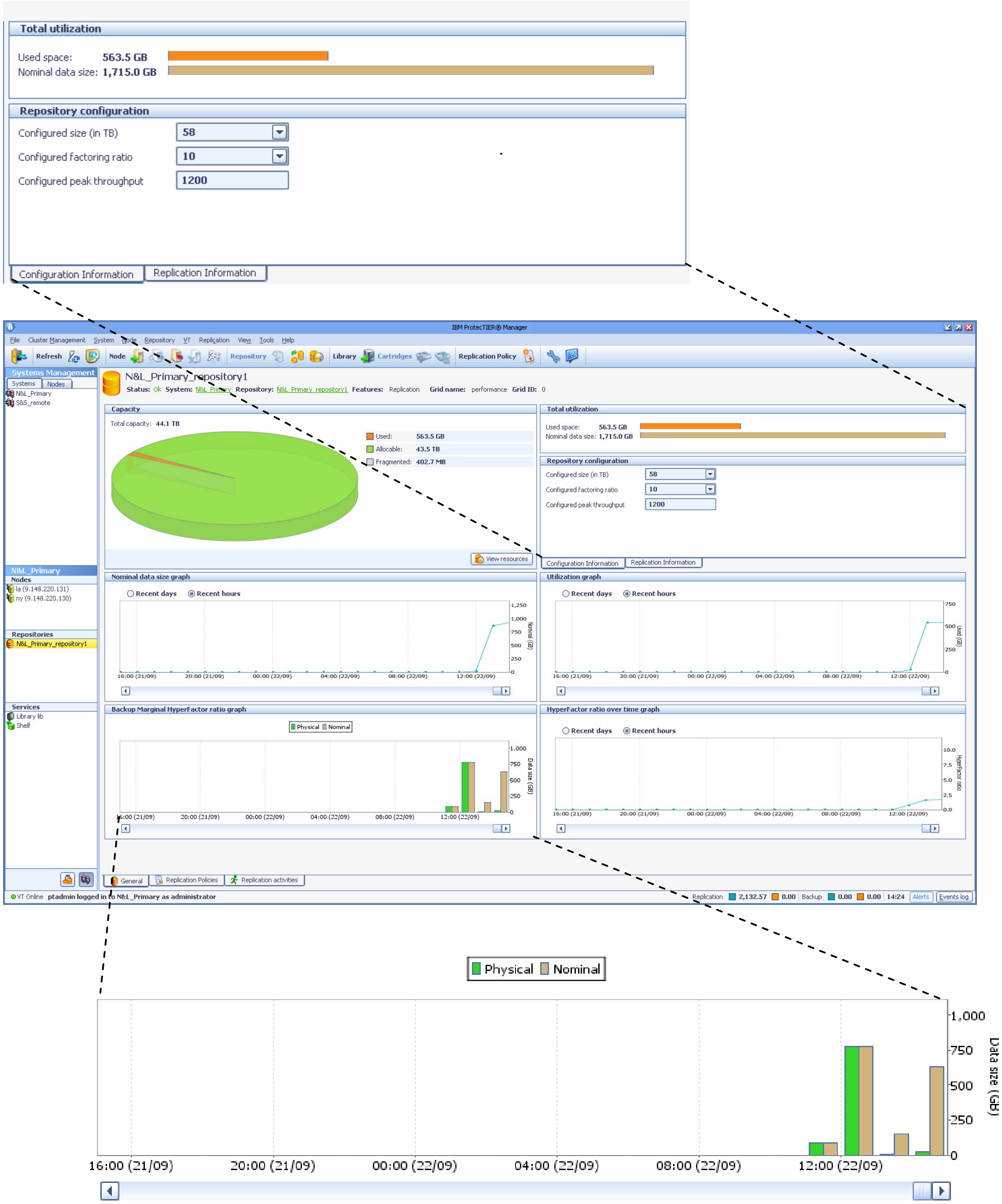
Test 2 - Replication of the Primary Repository Virtual Tapes to the Secondary

The replication was performed after backup has completed. Replication can run in parallel to a backup; however, in such cases, it competes with the backup application for system resources. Therefore, if possible, it is recommended to run it after backup completion which adheres with most users' physical tape practices, to create reliable SLAs and predictable RPOs. Another option is to cap or throttle the performance of the replication. See also test 5.

Test 3 - Simulation of a 3% Change in Primary Data and Performing a Full Backup.

Performing another backup to a new set of virtual tapes created backup generation 2. As can be seen in

Figure **5**; the nominal data doubled but the used space grew by less than 5%.



Total utilization

Used space: 563.5 GB
Nominal data size: 1,715.0 GB

Repository configuration

Configured size (in TB): 58
Configured factoring ratio: 10
Configured peak throughput: 1200

IBM ProtecTIER® Manager

N&L_Primary_repository1
Status: Ok System: N&L_Primary Repository: N&L_Primary_repository1 Features: Replication Grid name: performance Grid ID: 0

Capacity
Total capacity: 44.1 TB
Used: 563.5 GB
Allocatable: 43.5 TB
Fragmented: 402.7 MB

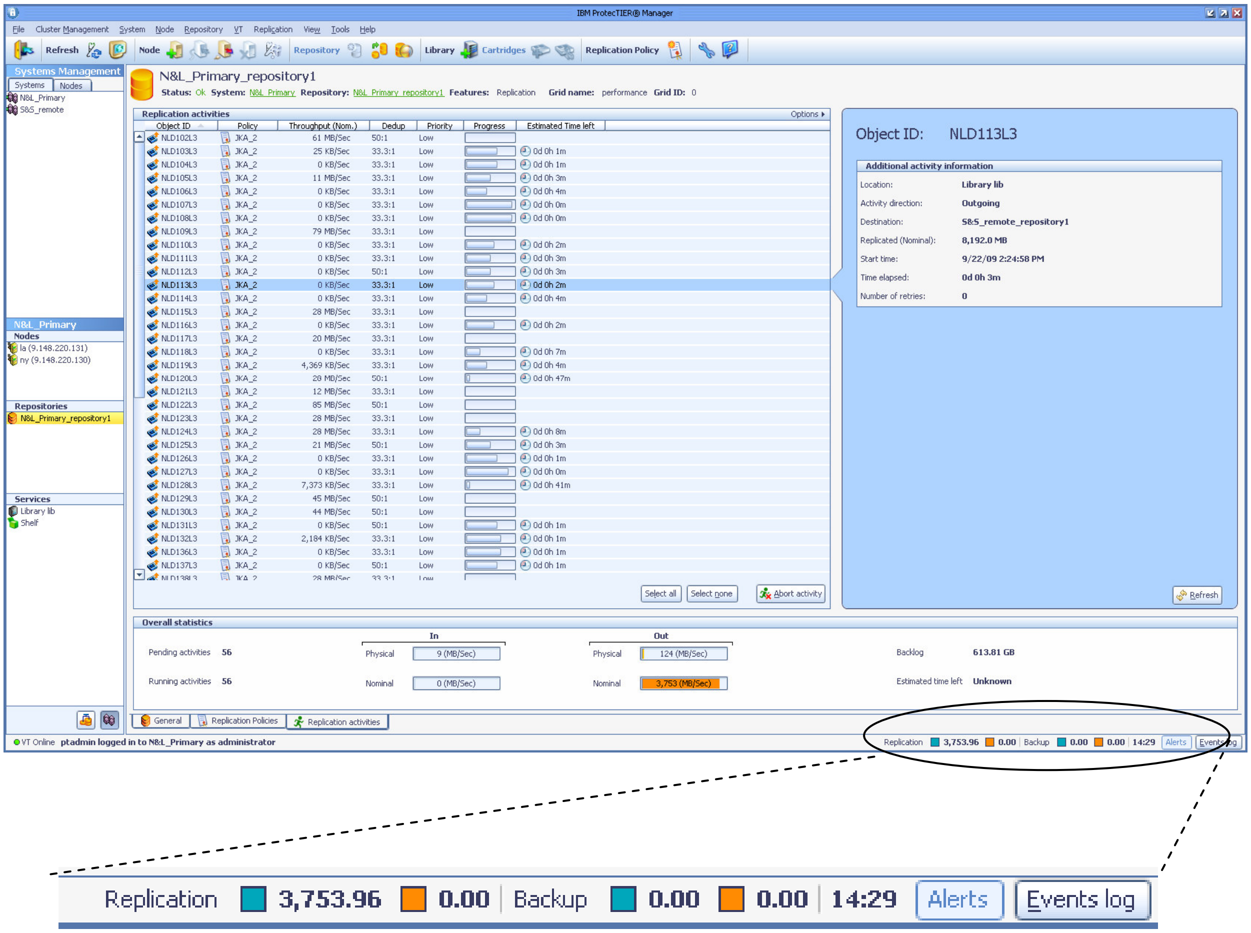
VT Online ptadmin logged in to N&L_Primary as administrator

Replication 2,132.57 0.00 Backup 0.00 0.00 14:24

Figure 5: Deduplication ratio and required capacity after second backup

Test4 - Replication of 2nd Generation Virtual Tapes to the Remote Site.

Figure 6 shows the progress of replicating the 2nd generation virtual tapes to the remote site. The first column represents the barcode of the virtual tape, the second the policy chosen for this action, the rest of the columns show the performance, progress, and estimated time to completion, respectively. Note the replication nominal throughput of 3753MB/s shown at the right bottom of the screen.



N&L_Primary_repository1

Status: Ok **System:** N&L_Primary **Repository:** N&L_Primary_repository1 **Features:** Replication **Grid name:** performance **Grid ID:** 0

Replication activities

Object ID	Policy	Throughput (Nom.)	Dedup	Priority	Progress	Estimated Time left
NLD102L3	JKA_2	61 MB/Sec	50:1	Low		
NLD103L3	JKA_2	25 KB/Sec	33.3:1	Low		0d 0h 1m
NLD104L3	JKA_2	0 KB/Sec	33.3:1	Low		0d 0h 1m
NLD105L3	JKA_2	11 MB/Sec	33.3:1	Low		0d 0h 3m
NLD106L3	JKA_2	0 KB/Sec	33.3:1	Low		0d 0h 4m
NLD107L3	JKA_2	0 KB/Sec	33.3:1	Low		0d 0h 0m
NLD108L3	JKA_2	0 KB/Sec	33.3:1	Low		0d 0h 0m
NLD109L3	JKA_2	79 MB/Sec	33.3:1	Low		
NLD110L3	JKA_2	0 KB/Sec	33.3:1	Low		0d 0h 2m
NLD111L3	JKA_2	0 KB/Sec	33.3:1	Low		0d 0h 3m
NLD112L3	JKA_2	0 KB/Sec	50:1	Low		0d 0h 3m
NLD113L3	JKA_2	0 KB/Sec	33.3:1	Low		0d 0h 2m
NLD114L3	JKA_2	0 KB/Sec	33.3:1	Low		0d 0h 4m
NLD115L3	JKA_2	28 MB/Sec	33.3:1	Low		
NLD116L3	JKA_2	0 KB/Sec	33.3:1	Low		0d 0h 2m
NLD117L3	JKA_2	20 MB/Sec	33.3:1	Low		
NLD118L3	JKA_2	0 KB/Sec	33.3:1	Low		0d 0h 7m
NLD119L3	JKA_2	4,369 KB/Sec	33.3:1	Low		0d 0h 4m
NLD120L3	JKA_2	28 MB/Sec	50:1	Low		0d 0h 47m
NLD121L3	JKA_2	12 MB/Sec	33.3:1	Low		
NLD122L3	JKA_2	85 MB/Sec	50:1	Low		
NLD123L3	JKA_2	28 MB/Sec	33.3:1	Low		
NLD124L3	JKA_2	28 MB/Sec	33.3:1	Low		0d 0h 8m
NLD125L3	JKA_2	21 MB/Sec	50:1	Low		0d 0h 3m
NLD126L3	JKA_2	0 KB/Sec	33.3:1	Low		0d 0h 1m
NLD127L3	JKA_2	0 KB/Sec	33.3:1	Low		0d 0h 0m
NLD128L3	JKA_2	7,373 KB/Sec	33.3:1	Low		0d 0h 41m
NLD129L3	JKA_2	45 MB/Sec	50:1	Low		
NLD130L3	JKA_2	44 MB/Sec	50:1	Low		
NLD131L3	JKA_2	0 KB/Sec	50:1	Low		0d 0h 1m
NLD132L3	JKA_2	2,184 KB/Sec	33.3:1	Low		0d 0h 1m
NLD136L3	JKA_2	0 KB/Sec	33.3:1	Low		0d 0h 1m
NLD137L3	JKA_2	0 KB/Sec	50:1	Low		0d 0h 1m
NLD138L3	JKA_2	28 MB/Sec	33.3:1	Low		

Object ID: NLD113L3

Additional activity information

Location:	Library lib
Activity direction:	Outgoing
Destination:	S&S_remote_repository1
Replicated (Nominal):	8,192.0 MB
Start time:	9/22/09 2:24:58 PM
Time elapsed:	0d 0h 3m
Number of retries:	0

Overall statistics

			In		Out		
Pending activities	56	Physical	9 (MB/Sec)	Physical	124 (MB/Sec)	Backlog	613.81 GB
Running activities	56	Nominal	0 (MB/Sec)	Nominal	3,753 (MB/Sec)	Estimated time left	Unknown

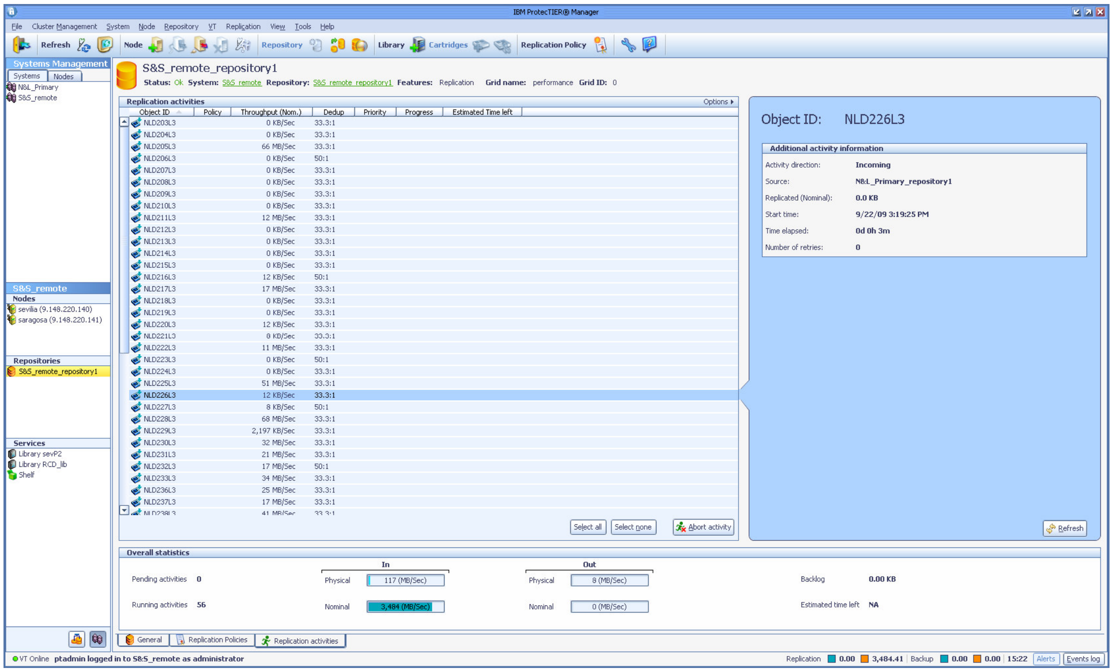
ptadmin logged in to N&L_Primary as administrator

Replication 🟦 3,753.96 🟧 0.00 | Backup 🟦 0.00 🟧 0.00 | 14:29

Replication 🟦 3,753.96 🟧 0.00 | Backup 🟦 0.00 🟧 0.00 | 14:29

Figure 6: Replication progress as seen from the primary

Figure 7 shows the virtual tapes at the remote site.



S&S_remote_repository1

Status: Ok **System:** S&S_remote **Repository:** S&S_remote_repository1 **Features:** Replication **Grid name:** performance **Grid ID:** 0

Replication activities

Object ID	Policy	Throughput (Nom.)	Dedup	Priority	Progress	Estimated Time left
NLD203L3		0 KB/Sec	33.3:1			
NLD204L3		0 KB/Sec	33.3:1			
NLD205L3		66 MB/Sec	33.3:1			
NLD206L3		0 KB/Sec	50:1			
NLD207L3		0 KB/Sec	33.3:1			
NLD208L3		0 KB/Sec	33.3:1			
NLD209L3		0 KB/Sec	33.3:1			
NLD210L3		0 KB/Sec	33.3:1			
NLD211L3		12 MB/Sec	33.3:1			
NLD212L3		0 KB/Sec	33.3:1			
NLD213L3		0 KB/Sec	33.3:1			
NLD214L3		0 KB/Sec	33.3:1			
NLD215L3		0 KB/Sec	33.3:1			
NLD216L3		12 KB/Sec	50:1			
NLD217L3		17 MB/Sec	33.3:1			
NLD218L3		0 KB/Sec	33.3:1			
NLD219L3		0 KB/Sec	33.3:1			
NLD220L3		12 KB/Sec	33.3:1			
NLD221L3		8 KB/Sec	33.3:1			
NLD222L3		11 MB/Sec	33.3:1			
NLD223L3		0 KB/Sec	50:1			
NLD224L3		0 KB/Sec	33.3:1			
NLD225L3		51 MB/Sec	33.3:1			
NLD226L3		12 KB/Sec	33.3:1			
NLD227L3		8 KB/Sec	50:1			
NLD228L3		68 MB/Sec	33.3:1			
NLD229L3		2,197 KB/Sec	33.3:1			
NLD230L3		32 MB/Sec	33.3:1			
NLD231L3		21 MB/Sec	33.3:1			
NLD232L3		17 MB/Sec	50:1			
NLD233L3		34 MB/Sec	33.3:1			
NLD236L3		25 MB/Sec	33.3:1			
NLD237L3		17 MB/Sec	33.3:1			
NLD238L3		41 MB/Sec	33.3:1			

Object ID: NLD226L3

Additional activity information

Activity direction:	Incoming
Source:	N&L_Primary_repository1
Replicated (Nominal):	0.0 KB
Start time:	9/22/09 3:19:25 PM
Time elapsed:	0d 0h 3m
Number of retries:	0

Overall statistics

			In		Out		
Pending activities	0	Physical	117 (MB/Sec)	Physical	8 (MB/Sec)	Backlog	0.00 KB
Running activities	56	Nominal	3,484 (MB/Sec)	Nominal	0 (MB/Sec)	Estimated time left	NA

VT Online ptadmin logged in to S&S_remote as administrator

Replication 🟦 0.00 🟧 3,484.41 | Backup 🟦 0.00 🟧 0.00 | 15:22

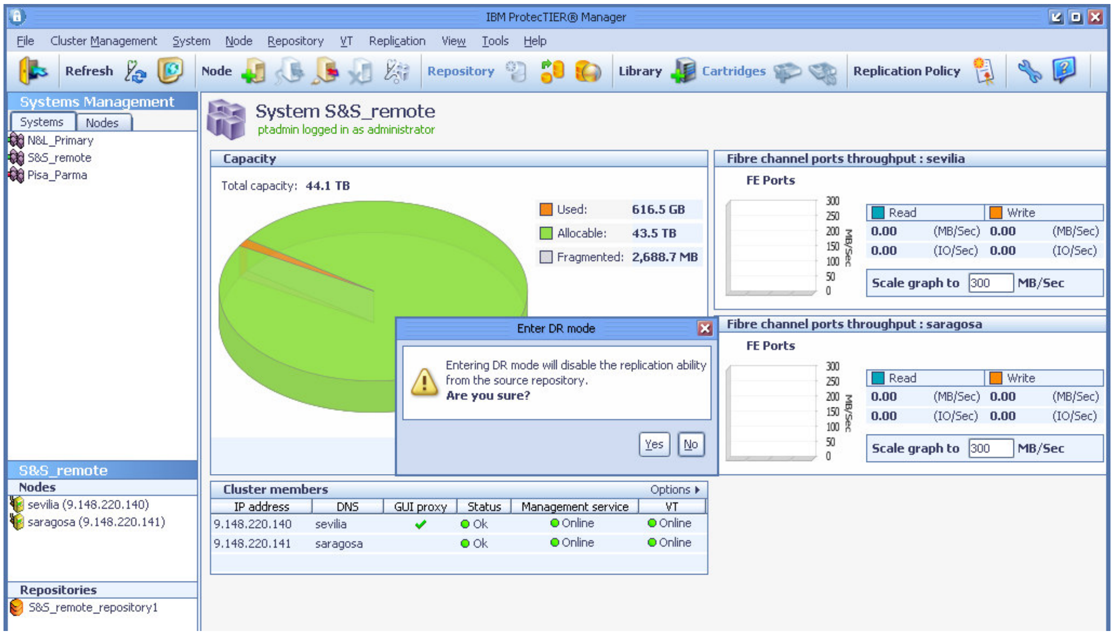
Figure 7: Virtual tapes in the secondary repository

Test 5 - Simulation of Additional 3% Change and Replicating in Parallel

Again 3% of the data was changed and the result replicated, this time in parallel to the backup. Replication commenced as soon as data was written to the primary.

Test 6 - ProtecTIER Native Replication Fail-over and Fail-back

This test was the most important part of the lab evaluations. As expected, declaring a disaster correctly froze the replication process, which is shown in the screen shot in Figure 8. Then, a file was restored from the previous generation of the backup tape at the secondary site using the backup application. See the file selection screen of NetBackup in Figure 9. The operation then moved to the secondary site, so new tapes were created on the secondary only. Then, fail-back to the primary was initiated (Figure 10). Finally, the new tapes created on the secondary were replicated to the primary. As can be seen from the screen shots, all the operations described above worked smoothly.



System S&S_remote
ptadmin logged in as administrator

Capacity

Total capacity: 44.1 TB

🟧 Used:	616.5 GB
🟩 Allocable:	43.5 TB
□ Fragmented:	2,688.7 MB

Fibre channel ports throughput : sevilia

FE Ports

🟦 Read		🟧 Write	
0.00	(MB/Sec)	0.00	(MB/Sec)
0.00	(IO/Sec)	0.00	(IO/Sec)

Scale graph to 300 MB/Sec

Fibre channel ports throughput : saragosa

FE Ports

🟦 Read		🟧 Write	
0.00	(MB/Sec)	0.00	(MB/Sec)
0.00	(IO/Sec)	0.00	(IO/Sec)

Scale graph to 300 MB/Sec

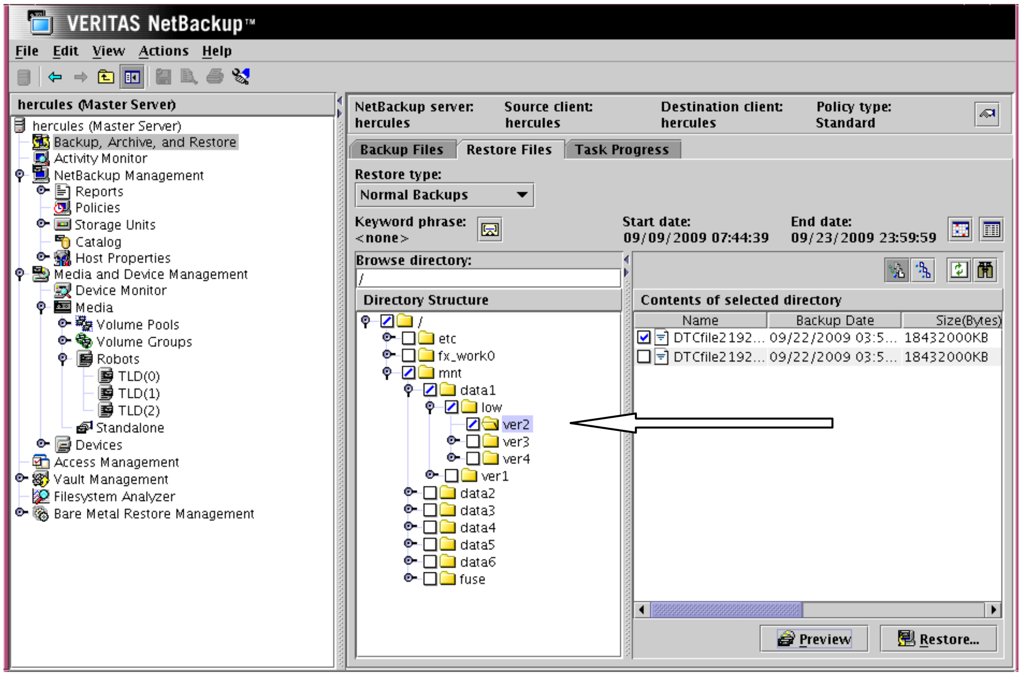
Enter DR mode

Entering DR mode will disable the replication ability from the source repository.
Are you sure?

Cluster members

IP address	DNS	GUI proxy	Status	Management service	VT
9.148.220.140	sevilia	✓	● Ok	● Online	● Online
9.148.220.141	saragosa		● Ok	● Online	● Online

Figure 8: Declaration of disaster



NetBackup server: hercules
Source client: hercules
Destination client: hercules
Policy type: Standard

Restore type: Normal Backups

Keyword phrase: <none>

Start date: 09/09/2009 07:44:39

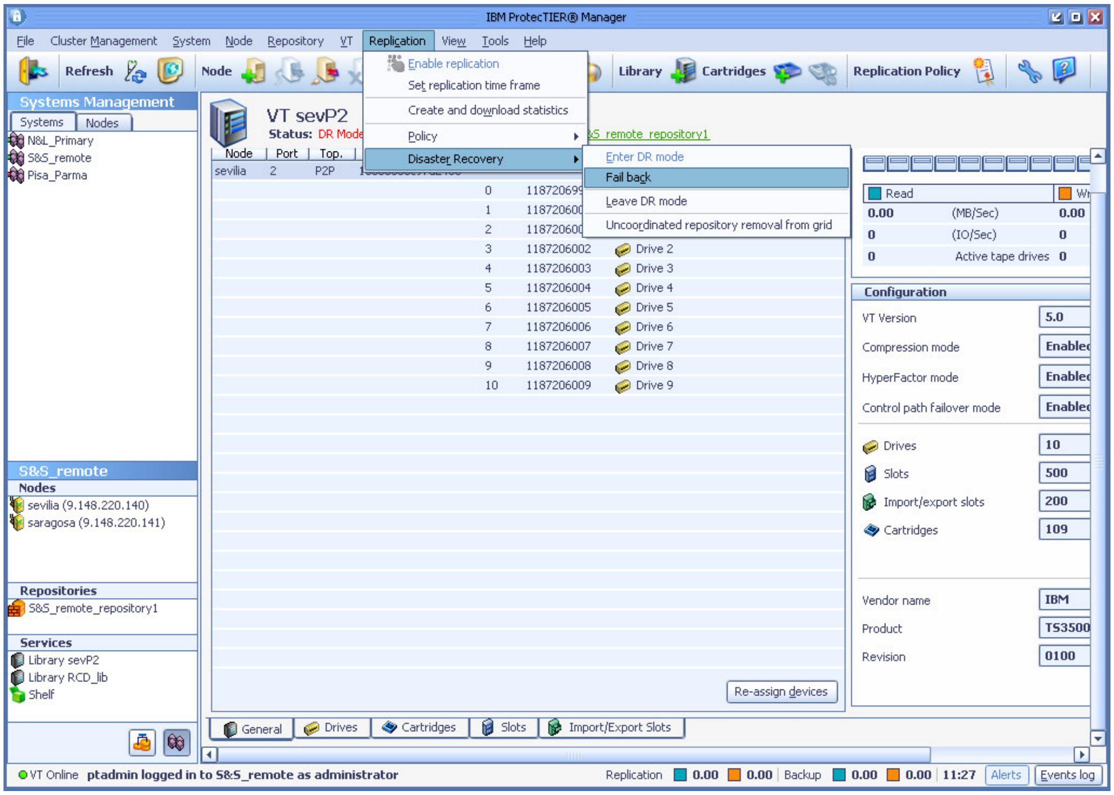
End date: 09/23/2009 23:59:59

Browse directory: /

Contents of selected directory

	Name	Backup Date	Size(Bytes)
☑	DTCfile2192...	09/22/2009 03:5...	18432000KB
□	DTCfile2192...	09/22/2009 03:5...	18432000KB

Figure 9: Selection of file to restore



VT sevP2
Status: DR Mode
S&S_remote_repository1

Node	Port	Top.
sevilia	2	P2P

0	118720699	
1	118720600	
2	118720600	
3	1187206002	Drive 2
4	1187206003	Drive 3
5	1187206004	Drive 4
6	1187206005	Drive 5
7	1187206006	Drive 6
8	1187206007	Drive 7
9	1187206008	Drive 8
10	1187206009	Drive 9

Read		Wr
0.00	(MB/Sec)	0.00
0	(IO/Sec)	0
0	Active tape drives	0

Configuration

VT Version	5.0
Compression mode	Enablec
HyperFactor mode	Enablec
Control path failover mode	Enablec
Drives	10
Slots	500
Import/export slots	200
Cartridges	109
Vendor name	IBM
Product	TS3500
Revision	0100

VT Online ptadmin logged in to S&S_remote as administrator

Replication 0.00 0.00 Backup 0.00 0.00 11:27

Figure 10: Fail-back activation

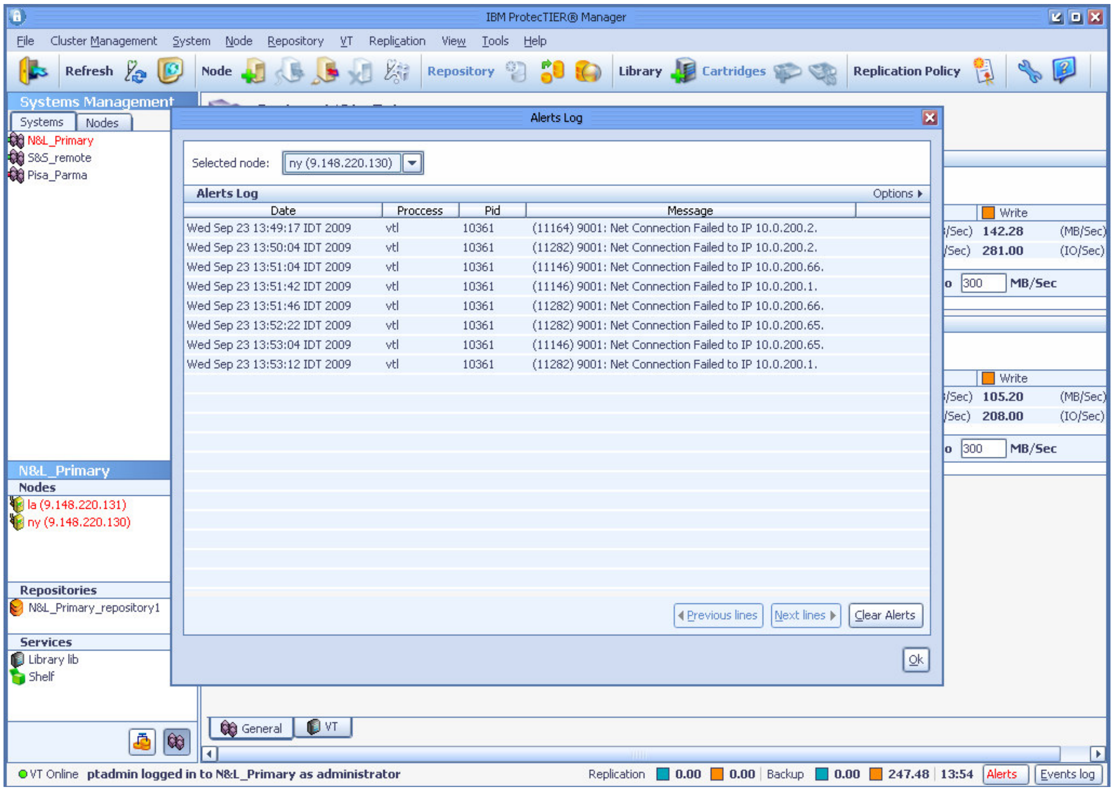
Test 7 - Testing ProtecTIER Replication's Reaction to Network Outages

A backup with parallel replication was performed. The 4 links were then physically disconnected. The replication stopped presenting an alert (the alerts can be seen in Figure 11, with the nodes related to the alert in red), but the backup continued as seen in Figure 12. The link was then restored 5 minutes later. The replication resumed from the point it had been interrupted. A comparison of the replicated tape with the original is performed in order to ensure data integrity. ProtecTIER replication has a default setting of 7 days for retries in case of network outages; after this threshold is reached, the job is aborted.

Test 8 - Virtual Tape Cloning at Remote Site

The two sites were defined as a single domain, which means that both shared the same NetBackup catalog.[2]. A backup with parallel replication was performed at a combined nominal throughput of over 1600MB/s, as shown in Figure 13 below. The tape was then placed in the "vault" using NetBackup's *eject function.*

A copy from virtual tape to physical tape was performed, and a different barcode was assigned to the physical tape to avoid a conflict of two "tapes" with identical barcodes. The physical tape was ejected and moved to the export window.



Alerts Log

Selected node: ny (9.148.220.130)

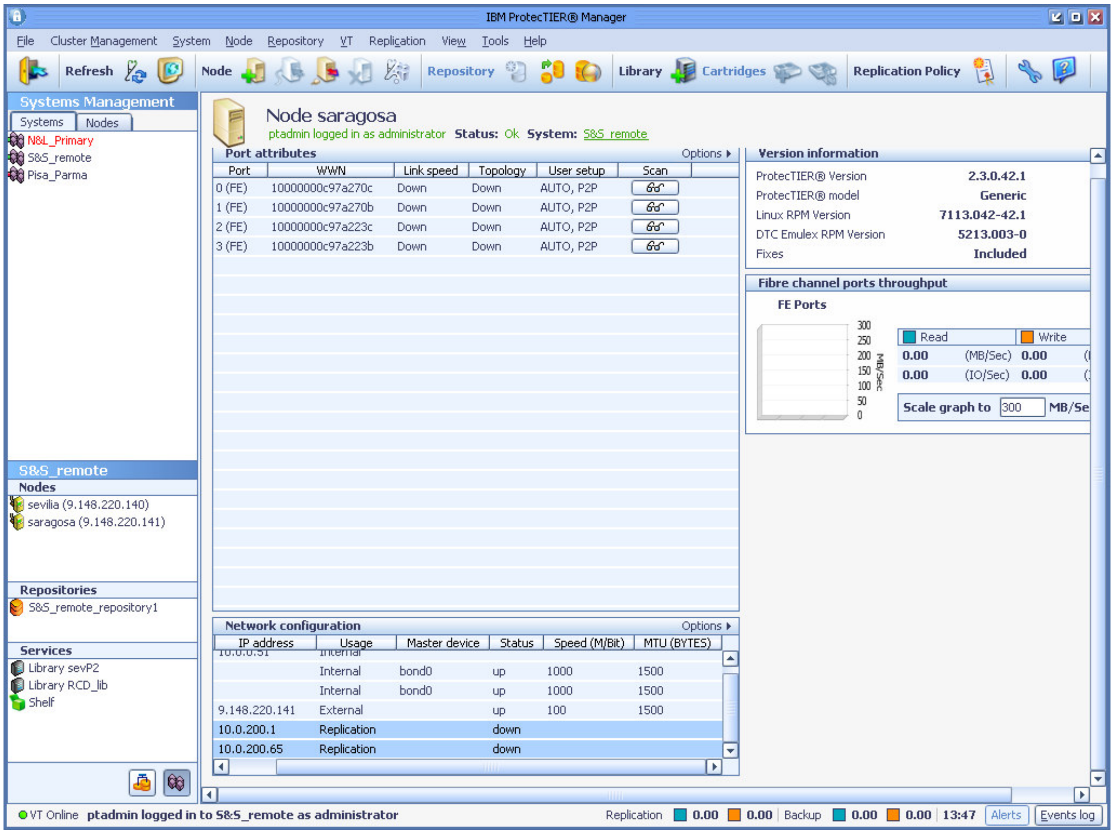
Alerts Log — Options

Date	Proccess	Pid	Message
Wed Sep 23 13:49:17 IDT 2009	vtl	10361	(11164) 9001: Net Connection Failed to IP 10.0.200.2.
Wed Sep 23 13:50:04 IDT 2009	vtl	10361	(11282) 9001: Net Connection Failed to IP 10.0.200.2.
Wed Sep 23 13:51:04 IDT 2009	vtl	10361	(11146) 9001: Net Connection Failed to IP 10.0.200.66.
Wed Sep 23 13:51:42 IDT 2009	vtl	10361	(11146) 9001: Net Connection Failed to IP 10.0.200.1.
Wed Sep 23 13:51:46 IDT 2009	vtl	10361	(11282) 9001: Net Connection Failed to IP 10.0.200.66.
Wed Sep 23 13:52:22 IDT 2009	vtl	10361	(11282) 9001: Net Connection Failed to IP 10.0.200.65.
Wed Sep 23 13:53:04 IDT 2009	vtl	10361	(11146) 9001: Net Connection Failed to IP 10.0.200.65.
Wed Sep 23 13:53:12 IDT 2009	vtl	10361	(11282) 9001: Net Connection Failed to IP 10.0.200.1.

ptadmin logged in to N&L_Primary as administrator

Figure 11: Alerts as result of the network outage

[2] Other backup applications use the term database in their nomenclature. It is a file containing data such as when the backup was taken, which files were backed up on each tape, the barcode of the tape, the retention date, client name, etc.



IBM ProtecTIER® Manager
File Cluster Management System Node Repository VT Replication View Tools Help
Refresh Node Repository Library Cartridges Replication Policy
Systems Management
Systems Nodes

N&L_Primary
S&S_remote
Pisa_Parma

S&S_remote

Nodes

sevilia (9.148.220.140)
saragosa (9.148.220.141)

Repositories

S&S_remote_repository1

Services

Library sevP2
Library RCD_lib
Shelf

Node saragosa
ptadmin logged in as administrator **Status:** Ok **System:** S&S_remote

Port attributes Options ▸

Port	WWN	Link speed	Topology	User setup	Scan
0 (FE)	10000000c97a270c	Down	Down	AUTO, P2P	
1 (FE)	10000000c97a270b	Down	Down	AUTO, P2P	
2 (FE)	10000000c97a223c	Down	Down	AUTO, P2P	
3 (FE)	10000000c97a223b	Down	Down	AUTO, P2P	

Version information

ProtecTIER® Version	2.3.0.42.1
ProtecTIER® model	Generic
Linux RPM Version	7113.042-42.1
DTC Emulex RPM Version	5213.003-0
Fixes	Included

Fibre channel ports throughput

🟦 Read		🟧 Write
0.00	(MB/Sec)	0.00
0.00	(IO/Sec)	0.00

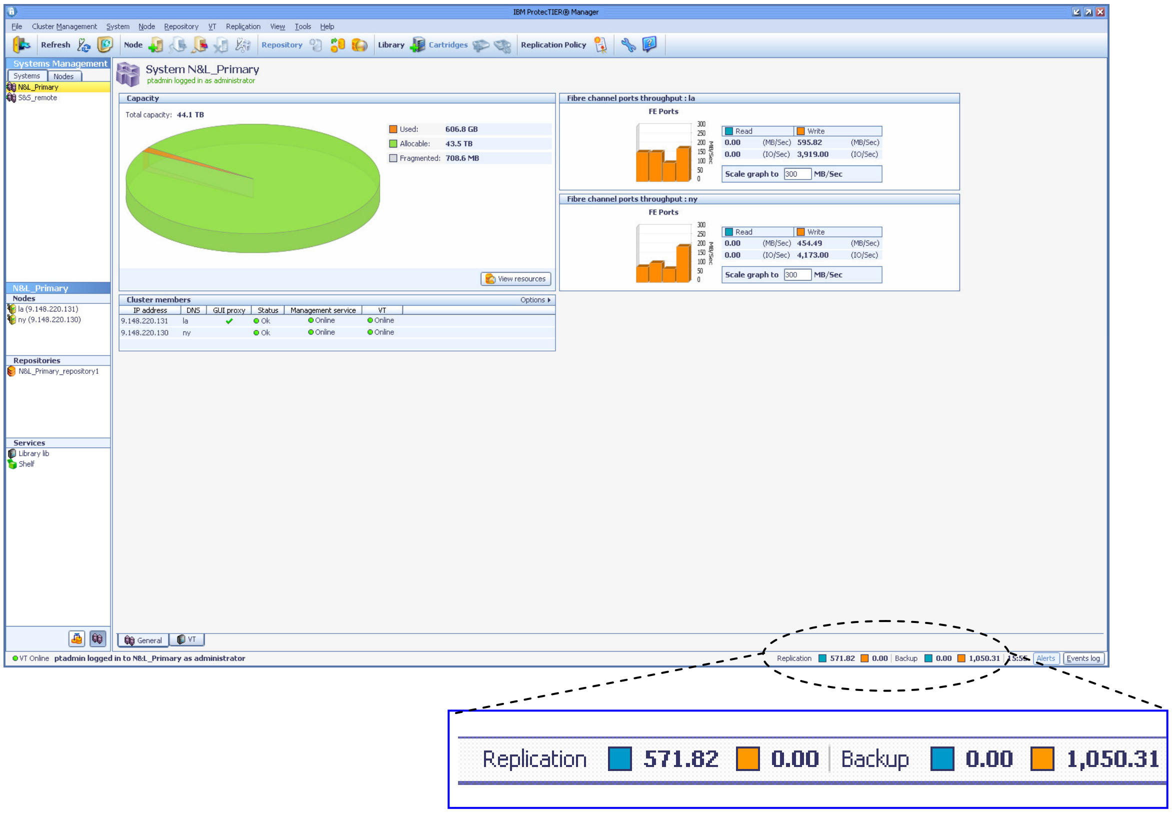
Scale graph to 300 MB/Se

Network configuration Options ▸

IP address	Usage	Master device	Status	Speed (M/Bit)	MTU (BYTES)
10.0.0.51	Internal				
	Internal	bond0	up	1000	1500
	Internal	bond0	up	1000	1500
9.148.220.141	External		up	100	1500
10.0.200.1	Replication		down		
10.0.200.65	Replication		down		

● VT Online **ptadmin logged in to S&S_remote as administrator** Replication 🟦 0.00 🟧 0.00 Backup 🟦 0.00 🟧 0.00 13:47 Alerts Events log

Figure 12: The replication stopped but the backup on the primary continues



System N&L_Primary
ptadmin logged in as administrator

Capacity

Total capacity: 44.1 TB

🟧 Used:	606.8 GB
🟩 Allocable:	43.5 TB
□ Fragmented:	708.6 MB

Fibre channel ports throughput : la

🟦 Read		🟧 Write	
0.00	(MB/Sec)	595.82	(MB/Sec)
0.00	(IO/Sec)	3,919.00	(IO/Sec)

Scale graph to 300 MB/Sec

Fibre channel ports throughput : ny

🟦 Read		🟧 Write	
0.00	(MB/Sec)	454.49	(MB/Sec)
0.00	(IO/Sec)	4,173.00	(IO/Sec)

Scale graph to 300 MB/Sec

Nodes: la (9.148.220.131), ny (9.148.220.130)

Repositories: N&L_Primary_repository1

Services: Library lib, Shelf

Cluster members

IP address	DNS	GUI proxy	Status	Management service	VT
9.148.220.131	la	✓	● Ok	● Online	● Online
9.148.220.130	ny		● Ok	● Online	● Online

● VT Online ptadmin logged in to N&L_Primary as administrator

Replication 🟦 571.82 🟧 0.00 | Backup 🟦 0.00 🟧 1,050.31 | 15:55

Figure 13: Replication is running concurrently with backup activity, the bottom right corner of the screen shows the in-progress read/write throughput (in MB/s) per activity at the primary site ProtecTIER server

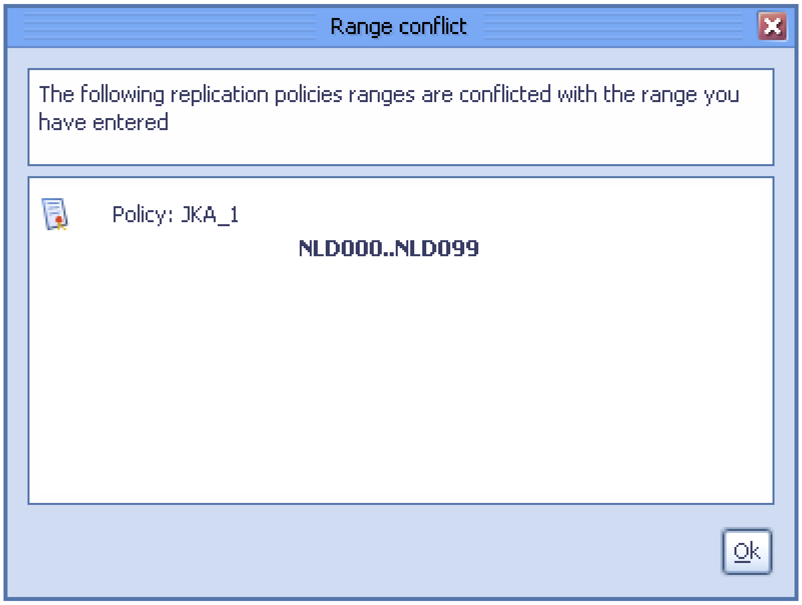
IBM System Storage™ TS7650 ProtecTIER Replication Management

The initial setting of the configurations, connections, translation between repository IDs and IP addresses, *etc.,* is done by the *Replication Grid Manager* software installed on one of the ProtecTIER nodes. This software does not participate in the replication itself but needs to be available for the fail-back operation.

Replication management is performed from the IBM ProtecTIER management screen, which is a user-friendly, easy to understand GUI as can be seen from the screen shots presented so far. In general, the ProtecTIER manager screen provides detailed information such as used and nominal capacity, performance figures, HyperFactor (deduplication) ratios, statistical information, alerts, *etc.*

There are built-in “fool proof” validation utilities such as avoiding problems due to a potential setting of conflicting policies. Figure 13 shows a message that is generated when a user mistakenly tries to set a new policy for a virtual tape that already belongs to another policy.

ProtecTIER’ replication policies allow for easy replication control and setting of priorities, and enable a high degree of automation. The native replication feature is transparent to the backup application, behaves just like a physical tape library, and is storage-subsystem agnostic.


Range conflict

The following replication policies ranges are conflicted with the range you have entered

Policy: JKA_1
NLD000..NLD099

Figure 13: Example of conflict detection

Bottom Line

All the tests were conducted with the most complicated configuration: two ProtecTIER clusters, backup servers of different performance, and a combination of external and internal storage. The application servers ran many backup streams in parallel, while at times, performing replication activity concurrently, at sustainable throughputs that in most cases were consistently and considerably higher than what is claimed by IBM. All possible scenarios such as replication, recovery on the secondary, running on the secondary, fail-back, moving new tapes created on the secondary to the primary, reaction to network outages, and creating (cloning) physical tape from virtual tape to store away in an external vault were successfully performed and tested. All these tests ran successfully and the IBM System Storage™ TS7650 ProtecTIER Native Replication delivered on its promises.

ProtecTIER’ replication policies allow for easy replication control and setting of priorities, and enable a high degree of automation. The native replication feature is transparent to the backup application, behaves just like a physical tape library, and is storage-subsystem agnostic. It saves communication costs and replication license costs on top of the savings from much smaller disk capacity requirements. The combination of high-speed inline deduplication, replication to remote sites, and external off-site vaulting of data provides tremendous value and answers all users' requirements at acceptable costs. In addition to lower data transfer costs it may save vaulting costs and make tape operations more secure, reliable and efficient,

On top of its comprehensive feature set, IBM System Storage™ TS7650 ProtecTIER fully is the only deduplication technique with a proven usage record by large organizations. By mid 2009, IBM had hundreds of ProtecTIER customers spanning Global 500 and mid-sized businesses across all industries. Tier-1 ProtecTIER customers include large corporations such as Verizon and Panasonic.