IBM's *TS7650 ProtecTIER®*: Value Proposition and TCO

Enterprise-class data protection systems with data deduplication such as IBM's TS7650 ProtecTIER®, have a tremendous value proposition. This paper thoroughly details why you need to incorporate data deduplication into your backup and disaster recovery practices and the benefits you can receive, including reducing the cost of backup operations on the order of 40%.

Unless yours is one of the organizations that have already implement data deduplication[1], backup is still considered to be the most costly storage related IT practice in the datacenter[2]. It is hard to imagine that IT still fights with backup after all these years, but to illustrate, look at the storage inefficiencies incumbent in the practice. Just six years ago, we thought that sticking a disk array into the backup data stream as a target was a wonderful and revolutionary improvement that would change the face of backup. It did and it didn't. Soon, the race was on all across the industry to move to disk-based backup. Current research estimates that about two-thirds of all datacenters have now obtained the benefits of disk as the backup target. Paradoxically though, backup-to-disk has many hidden costs incompatible with the shift to energy and cost reduction as top business priorities. Here is the dilemma. Just inserting a disk array or virtual tape

Hidden Costs of Backup

- Cost of backup is 30%-50% of IT storage budget
- 90%+ of your backup pool is made up of duplicate data
- 20%-25% of your data is expired
- Only 20-30% of your data is active

[1] Data Deduplication: is defined as the replacement of multiple copies of data with references to a shared copy in order to save storage space and/or bandwidth. The granularity of data deduplication varies based on the specifics of the technologies and processes used. Sub-file data deduplication is a form of data deduplication that operates at finer granularity than an entire file. Examples of subfile objects include data sets, data objects, or even the I/O stream. Source: SNIA "*Building a Terminology Bridge: Guidelines for Retention and Preservation Practices*", Sept. 2009

[2] Source: Gartner, 2009: "Backup practices cost 30% to 50% of the IT storage budget"

library into the backup process adds hidden expenses for reasons like the following:

- 90%+ of the backup stream is duplicate or redundant data requiring large amounts of expensive disk capacity to accommodate the redundancy
- A disk array typically runs well below 70% utilization; backup arrays are just as bad or worse as over-allocation is common to accommodate rapid capacity growth.
- And, typical RAID redundancy (RAID 5 for example) requires at least an additional 20% overhead

These three numbers mean that the effective "data efficiency" in the backup-to-disk storage pool is approximately 6%-8%. Said the other way around, with 90% redundancy in the backup pool you have to purchase and install ~15-20 times more disk (and tape) to store your backups than your primary storage you are trying to protect. Think about it. If it takes $20,000 to $50,000 of disk array cost to protect 1 TB of primary storage due to these inefficiencies, something is wrong.

Poor Data Efficiency with Backup-to-Disk

The data efficiency of Backup to Disk is less than 10%

Cost Savings Rule the Day

The point of this exercise is to make it painfully obvious that eliminating redundancy in the backup pool is one of the easiest, most transparent, and most effective cost reduction approaches available. Improved 'data efficiency' isn't the only reason to consider adding data deduplication or the only opportunity to improve your storage architecture and reduce operating costs and expense. Look at some of the other dimensions of backup. For example, a "backup everything" mentality also has huge inefficiencies. Consider the fact that only 20%-30% of the primary storage pool is in an 'active state'[3]. Continuing to backup content that is static, not changing and has already been backed up is a waste of effort, time, and resources. Even worse, continuing to backup the 25% of information and data in the typical datacenter that is no longer needed and has 'expired' is a waste and actually risky from a risk management perspective. The solution is to take a holistic business approach to data protection and look at both the opportunity and applicability of cost reduction technologies.

The sidebar on the next page presents what Strategic Research calls the "Hierarchy of Available IT Practice Cost Savings" ordered relative to the capital expense required to gain the benefits.

[3] For an explanation of information or data state and how to use it to reduce operating costs, refer to the SNIA report: "*Building a Terminology Bridge: Guidelines for Retention and Preservation Practices*", Sept. 2009

Value Proposition

Deletion of expired data is the first on the list because of risk. Deletion has huge benefits from removing the ~25% expired data and thereby immediately reducing the resource load with a low implementation cost. Deletion has the added benefit of reducing legal, security, and compliance risk; costs that may easily outweigh all others. Second on the list, and easier to achieve, is data deduplication within the backup and disaster recovery process. Most backup streams have over 90% redundancy. Reducing that load saves storage cost significantly and improves recovery responsiveness because more data can cost-effectively be saved for longer on disk where it is accessible for recovery and discovery purposes. Deduping content before sending it offsite for disaster recovery is an equally important opportunity to reduce overhead and improve recovery capabilities. Dedupe is so efficient at reducing the data load and the required bandwidth that remote vaulting becomes less expensive and more practical than shipping tapes to a disaster recovery (DR) site[4]. Dedupe also applies to reducing potential redundancy in datacenter archives or preservation stores for all the same reasons. Tiering and the automation of tiering through virtualization are next on the list, but be aware that the capital expense for these practices goes up significantly even though the ROI may be very good. Tiering has the added benefit of allowing more intelligence to be applied to backup schemas, potentially reducing the backup and DR load and cost even further.

Hierarchy of Available IT Cost Savings

1. **Deletion**
   - Reduce storage and backup cost
     - 20%-25% expired information and data
     - Free up capacity, reduce resource utilization…
2. **Data Deduplication**
   - Reduce Backup cost:
     - 90%+ redundancy
     - Improve recovery
   - Dedupe Test/Dev Clones
   - Reduce DR cost
   - Dedupe archive-preservation repositories
3. **Tiering**
   - Reduce storage cost
     - 75% of stored data is inactive, reference, or expired
     - Dedupe secondary tiers
   - Reduce backup cost
     - Align data protection strategy to requirements by tier
4. **Storage Virtualization**
   - Reduce storage cost
     - Thin provisioning
     - Automate tiering

Data Deduplication Value Proposition

Data deduplication is a consistent thread through all these 'cost-reduction' practices because of its strong value proposition. The benefit algorithm is straightforward. The higher the level of redundancy and the greater the cost or overhead of maintaining, transferring, or protecting redundant data, the greater the benefits of data deduplication.

The important elements of the business value propositions for data deduplication in backup and disaster recovery include its impact on the actual processes as well as cost reduction, data restoration, and operational recovery improvements. These are strong and compelling business and operations benefits. Here is a comprehensive list for comparison.

[4] See the TCO analysis in this report for clarification and evidence of this point

Data Deduplication Value Propositions

Transparency and Normalcy:

- Drops into the backup process transparently and non-disruptively
- Enables the benefits of backup to disk (performance, reduced backup time, reduced backup errors, rapid recovery capabilities, etc.) along with improved storage efficiency
- Tape may still be utilized as a second tier in the backup repository, though far less media and drives will be needed than before

Improved IT Operations:

- Improved manageability of the backup pool (it is physically smaller and still managed normally through the backup management console)
- 10x or better improvement in recovery and restoration times compared to tape-base backup (better able to meet Recovery Time Objectives, RTO)
- Higher reliability and data integrity in backup and recovery processes
- Enables improved DR practices through remote replication, resulting in improved operational recovery and business continuity protection

Cost Reduced Operations, Improved Agility:

- Whether backing up to a disk array or a virtual tape library, a deduplicated disk pool is ~10-25x smaller per TB protected and this reduced size contains more history, providing the ability to retain backup data online for longer periods for a given storage capacity, aiding in recovery and data restoration
- Improved disk and tape utilization efficiencies
- 50%-80% storage cost savings compared to disk-based backup without deduplication
- Improved ability to delete expired data with better disposition and litigation hold controls
- Reduced operator and administrative costs to manage backup and restoration processes at both the primary backup site(s) and the DR site(s).
- Technology and media migration pain will be greatly reduced as the load and reliance on tape media is reduced.
- Reduced storage demand, slowing hardware growth requirements, and resulting savings, including reduced storage management costs, bandwidth, and resource loads
- Better ability to respond to backup requirements from capacity growth driven by new projects

Deduplication Value Proposition

Transparency and Normalcy

Improved IT Operations

Cost Reduced Operations

Improved Agility

Lower TCO

Rapid Payback

TCO Analysis

- Improved capabilities and reduced costs for centralizing remote site(s) backup by utilizing global deduplication practices
- Improved, simplified, and cost effective disaster recovery processes by deduplicating before replicating data to or from remote DR sites. (Dramatically reduces WAN bandwidth requirements and cost.)
- Reduced power consumption and fewer storage and media resources provides 'green' benefits

Deduplication Value Proposition

The value proposition is so strong that failing to add data deduplication to a disk or tape-based backup or disaster recovery strategy can no longer be justified

Hopefully, the ripple effect of data deduplication is visible. It impacts just about every process it touches, reducing cost and improving agility. Bottom line, the business and operations benefits outweigh the direct cost savings. The value proposition is so strong that failing to add data deduplication to a disk or tape-based backup or disaster recovery strategy can no longer be justified. But, just in case you are not convinced, let's now look specifically at the cost benefits of data deduplication.

TCO Analysis Methods for Data Deduplication

A TCO analysis approach is useful in that it quantitatively assesses and validates the claims of the value proposition. The analysis method is to compute the gross operating costs of several different approaches to providing data protection services and compare them quantitatively over a five-year time period[5]. While no standard methodology for Total Cost of Ownership, TCO, analysis for data deduplication of the backup and disaster recovery process exists, the approach taken in this one attempts to be as neutral and as inclusive of good practices as possible. This TCO analysis tool was developed collaboratively by IBM and Strategic Research, pooling our collective experience focused on making the model realistic and comprehensive. The model's key differences to other published TCO analysis of deduplication in backup are summarized by the following points and are further detailed in the Appendix:

TCO Model Characteristics

- It is a consistent, uniform model, from which normalized data can be derived. The key variable is capacity under management and derivatives from that variable. It is not based on varying customer specific implementations or varying implementation parameters that confound the data. Rather, it is based on normalized, complete, "best"

[5] TCO measurement over a 5-year period is long enough to take into account technology retirements and migration expense, unlike a 3-year TCO.

practices consistently applied across all methods being compared for each scenario.

- These TCOs are computed on 'complete' processes, not selective subsets, including factors categorized as hardware, software, maintenance fees, expansion licenses, bandwidth costs, administrative and operator costs, media handling and transportation, media purchases and replacements, equipment retirements, etc.
- The model does not include soft-dollar expenses such as downtime or administrative or user productivity losses nor does it include the costs of deferred expenses. Deferred capital expenses are often used in TCOs to offset or justify TCOs. That thinking is misleading and effectively double counting. Each storage configuration needs to be evaluated independently on its own merits as a new acquisition on top of and incremental to existing practices.
- Resource loads and equipment requirements are determined by performance and scale factors such as available backup or transfer windows, retention periods, equipment performance capabilities, retirements, backup and DR schedules, compression and deduplication ratios, etc. and costs are based on publicly available market pricing. The model used scales based on these types of deterministic factors to accommodate different installation sizes. Loads, equipment, labor, and supply requirements are computed using industry norms, not guessed at.
- The analysis includes backup and disaster recovery (DR) processes, but not archive[6]. In each scenario, it is assumed that the organization already has an existing tape-library-based backup and DR practice in place. The DR analysis compares the required expansion of the existing tape-based DR operation to new implementations of VTLs with and without deduplication-based DR practices that utilize remote replication.

Use Cases Evaluated

In this 5-Year TCO analysis, the efficiency and cost of three backup targets are compared across three different sized use-cases. The three targets are tape libraries, VTLs without deduplication, and IBM's *TS7650 ProtecTIER* data protection platform with data

Analysis Methods

TCO Model Characteristics

Consistent application of 'best practices'

End-to-end perspective with full and equivalent equipment loading

No soft dollars

No deferred capital expenses

Dedupe HW & SW costs added to existing platforms

Market pricing

Backup and Disaster Recovery

Does not include Archive

[6] From a "best practices" and a risk management perspective, archive should not be part of the backup or DR domain. It should be a separate, independent process. Consequently it was left out of this analysis. Accordingly, backup and DR retention periods were defined as 90 days. For reference, an archive is a specialized preservation repository for the retention and preservation of digital information and data. Source: SNIA's "*Building a Terminology Bridge: Guidelines to Digital Information Retention and Preservation Practices in the Datacenter*", September 2009

Use-Cases Evaluated

deduplication. The three use-case configurations analyzed vary in size consistent with many real-world situations as illustrated in the following chart.

Use Cases Evaluated

	Starting Primary Capacity	Ending Primary Capacity	Backup Retention Period	DR Retention Period	Storage Growth Rate
Case 1	25 TB	134 TB	90 day	90 day	40%
Case 2	100 TB	538 TB	90 day	90 day	40%
Case 3	250 TB	1345 TB	90 day	90 day	40%

Backup & DR Solutions Analyzed

- **Tape library-based backup** including DR to a remote site achieved by transporting tapes where 3 months of data is held online in a duplicate system.
- **VTL based backup** utilizing compression. DR is accomplished via remote replication of compressed data to a duplicate VTL in the remote site.
- **Backup to IBM's *TS7650 ProtecTIER*** data protection platform with data deduplication. DR is achieved by remote replication to a duplicate *ProtecTIER* system at the DR site.

5 Yr. TCO Comparison

Dedupe is very cost-effective compared to Tape-based Backup

Backup: 45%-50% lower

DR: 15%-20% lower

Key TCO Observations

The data derived from the analysis is full of interesting and important points and conclusions. Begin by looking at some of the key metrics. For example, by normalizing the scenarios and including all CapEx and OpEx costs, the relative annual cost of backup and DR can now be gauged. We can clearly say that in the general use-case, backup to a *ProtecTIER*-class deduplication system has a 45%-50% lower 5-year TCO than backup to a tape-only system. Similarly, a remote replication process for disaster recovery using the deduplicated backup pool has a 15%-20% lower TCO than transporting tapes to some remote facility for disaster recovery.

To fully see and understand the entire results, study the following summary charts and review the detailed data in the Appendix.

Comparing Backup TCO: The 5-year TCO comparison metrics are presented in Figure 1. Both *ProtecTIER*-to-tape and VTL-to-tape are summarized in this figure. The details of the analytical model are presented in the appendix, but a quick summary of some key assumptions will provide a useful perspective before looking at results:

- **Analytical Model Key Assumptions that Reduce Bias**
  - To reduce potential bias in the model against tape-based practices three important assumptions were intentionally made:
    - **Capital Expenses**: Tape Library capital costs were minimized by assuming that each use-case started with an exisiting and adequately sized tape library based solution for "Year 1" needs. In comparison, both VTL and ProtecTIER based solutions were completely new buys – and both incurred substantial upfront capital expense which draws out the payback period. (Remember as well, that no offsetting expenses were used to justify TCO.)
    - **Bandwidth Costs**: Tape-based DR practices use relatively inexpensive physical pickup and delivery mechanisms; no capital expense, but a fair amount of labor. In comparison, both VTL and ProtecTIER DR practices are based on remote replication to duplicate systems and bandwidth costs that grow in the large capacity installations to be huge.
    - **Labor**: The normal, industry experience and ratio of management cost of tape-based backup systems compared to the management cost of disk arrays is a ratio that says tape is 6-10 times more costly to manage than disk. The actual metrics are[7]
      - **Tape**: 10-15 TB/FTE
      - **Disk**: 60-100 TB/FTE

      This analysis could have used the 6-10 times ratio to compute the labor load for each scenario. But, it did not. Instead, it used a factor of 3x to force erring on the low side, if at all. This model also compensated the ratios used based on the total capacity under management. For example, in the small site, use-case 1, it used these ratios: tape 20TB/FTE and disk 60TB/FTE. In the largest site, it used tape 40TB/FTE and disk 110TB/FTE.

TCO Observations

Labor Cost Ratios

IT management efficiency metrics

Tape: 10-15 TB/FTE

Disk: 60-100TB/FTE

[7] Sources: numerous industry studies published over the last several years

TCO Comparisons

Figure 1 compares *ProtecTIER*-to-tape and VTL-to-tape across the three use cases and then provides a summary view of the payback period and the total cost reduction at net present value, NPV.

Figure 1

5-Year TCO SAVINGS COMPARISON

	Case 1 70TB	Case 2 275TB	Case 3 700TB	Weighted Average
PROTECTIER TO TAPE-BASED BACKUP				
Savings as Percent of Annual Cost				
Backup	53%	54%	50%	52%
DR	31%	25%	12%	17%
Labor	69%	66%	62%	63%
Total	45%	45%	38%	41%
VTL TO TAPE-BASED BACKUP				
Savings as Percent of Annual Cost				
Backup	22%	14%	-3%	3%
DR	-44%	-104%	-138%	-123%
Total	-3%	-24%	-45%	-37%
ProtecTIER compared to Tape Libraries				
Payback Period (mo)	21	6	18	
Total Cost Reduction (@NPV)	48%	49%	43%	45%

- **ProtecTIER-to-Tape Libraries**:
  - 50% average lower TCO compared to traditional tape-based backup when comparing annualized costs across all three use cases
  - A weighted average 17% reduction in DR TCO
  - 60% reduction in labor required for Backup and DR administration and operations[8].
  - The payback period varies widely, from 6 to 21 months, driven by the amount of upfront new capital equipment required for both backup and DR since one of the key assumptions is to place identical systems onsite and offsite.

- **VTL-to-Tape Libraries**:
  - VTLs for backup have other benefits than just TCO otherwise as this TCO study reveals cost is not a positive driver. As noted in Figure 1, the cost savings is

[8] For more on why this is the case, see the discussion in the Appendix regarding the ratio of tape labor to ProtecTIER labor.

only 3% over 5 years – effectively no different. The cause in this analysis centers on the large capital expense and relatively low cost savings.

- VTLs for DR also do not compare favorably to tape-based DR operations that physically transport media due to the capital expenses and significant bandwidth expenses of remote replication.

Lessons Learned

Many great lessons reside in the model and its analysis. These range from better understanding the cost impact of making process decisions like how much data to keep online in the backup pool for recoveries to understanding cost tradeoffs of using disk and tape in concert. The following lessons discuss how various backup or DR practices impact TCO.

- These **increase TCO**: longer online retention periods, inefficient backup practices[9] with retention periods that retain more data than needed, more data held online, transmitting more data offsite (increased bandwidth requirement), increased capacity growth rate, poor equipment choices that increases the time to process and move data or require more equipment to be purchased to handle peak loads
- These **reduce TCO**: (In addition to the inverse of the 'increase TCO' list consider these observations from the model.) performance is key (while more expensive up front, it will save dollars in the long run), data deduplication is essential both for backup and DR, become more efficient in limiting the amount of data retained in the backup and DR pool and the amount being retained online, and also carefully analyze the types and amount of data being backed up to begin with.
- The total cost of ownership of a deduplicated disk-based backup target compared to a non-deduplication-based VTL is 50% lower for backup and 60% lower for DR. The reasons deduplication comes out as less expensive center on proportional reductions in hardware, software, bandwidth costs, and labor costs due to the space reduction advantages of dedupe. When including DR practices via remote replication, then deduplication-based practices have an additional large advantage due to the lower annual WAN bandwidth cost.

Lessons Learned

To Reduce TCO

Reduce backup retention periods

Use high performance systems

Use deduplication

Be smart about what gets backed up

[9] For example, excess redundancy caused by backing up previously backed-up data or by backing up inactive or reference data and information causes inefficiencies

Lessons Learned

- The model conservatively used a 95% deduplication plus compression ratio (20:1 dedupe ratio including compression) to determine what effect a relatively low ratio would have on TCO. The results were quite positive. Obviously, if your organization has a higher dedupe ratio, your savings will only be greater.
- To add reality in comparing DR practices, it was assumed that each site already had a tape library-based DR process in place. Consequently, the disk-based systems required entirely new equipment purchases whereas the tape-based use case only had to expand as it grew. Even with this added overhead, remote replication and the ensuing efficiencies of disk-based processes with deduplication resulted in great savings. Alternatively, the VTL-only solution went the other way and ended up costing an average of 120% more.
- This model does not include archival storage. It does not assume that the IT practice of retaining old backups (usually held on tape) and calling them 'archives' is used. Nor does it include any tiering practices called "archiving." Rather, the model uses a short internal retention period (90 days) for both backup and DR and nothing more. It assumes long-term retention of business important or business critical information or data classes are held in a proper preservation-class archive not in a backup repository[10]. Nor does it track any data protection methods for the archive repository in this model. This decision to separate archive away from backup and DR allows the cost model to more accurately portray and compare active operating costs. This is useful considering that practices vary tremendously when it comes to dealing with archival repositories and protection of archives.

How to Look at Data Deduplication with *ProtecTIER*

Huge Value Proposition for Backup and DR

50% lower TCO for Backup than Tape

40% lower TCO for DR than Tape

Conclusions and Recommendations

It is important to point out again that by using a comprehensive, normalized model the benefits and value proposition of deduplication for backup and disaster recovery have been validated. Even though this data is a specific comparison of IBM's *TS7650 ProtecTIER*, we can finally say, "Data duplication's value proposition is now quantified and understandable. Data dedupe is 40% to 50% less expensive to own and operate for backup and DR as compared to tape-based practices."

[10] For more on the discussion of what a proper archive is and what it is not, see the SNIA publication "*Building a Terminology Bridge: Guidelines to Digital Information Retention and Preservation Practices in the Datacenter*" 2009

Recommendations

The *TS7650 ProtecTIER* data protection platform is on one hand a deduplication appliance like many others with similar benefits. On the other, it has unique enterprise-class attributes that are very important such as cost-effective scalable performance as verified by this model. Use Cases #1 to #3 spanned a huge range of size. Use Case #3 is multi-Petabyte in size[11] and the $11 million in savings really stands out as an opportunity for critical cost control in a rapidly growing large environment.

Most published TCO analyses conclude that labor-savings alone overwhelm all other costs and justify movement to data deduplication. Sure, the labor savings benefit is very large, but there is so much more to the equation and so many places in this complex process to reduce costs and improve efficiencies. Consequently, we recommend that you do not look at data deduplication as a point solution. Rather, take a balanced, holistic approach. Look at the big picture as there are many places to save money and improve efficiency in data protection and disaster recovery practices. Hopefully, you will find many of the observations and recommendations made in this analysis pertinent to your organization's situation and of help in your planning.

--- end ---

[11] Use case 3 has 6 PB online in the backup pool at the end of year 5.

Appendix

TCO Model description:

The TCO model compares three different sized use-cases with three backup and DR methods: tape automation-only, VTL-assisted by tape, and an IBM *TS7650 ProtecTIER* data protection platform with data deduplication. The performance characteristics of IBM's *TS7650 ProtecTIER* platform were used to compute its loading and resulting system cost in the analysis. The backup solution is a traditional tape-based solution using LTO4 technology drives and media with a grandfather-father-son rotation schema.

Scenarios Evaluated

	Starting Primary Capacity	Ending Primary Capacity	Ending Backup Pool Capacity	Backup/DR Retention Period	Storage Growth Rate
Case 1	25 TB	134 TB	620TB	90 day	40%
Case 2	100 TB	538 TB	2,440TB	90 day	40%
Case 3	250 TB	1345 TB	6,025TB	90 day	40%

Backup & DR Solutions Analyzed

- **Tape library-based backup** with physical media transported to a remote site for DR. 13 week retention with a grandfather-father-son rotation schema.
- **VTL based backup** utilizing compression. DR is accomplished via remote replication of compressed data to another VTL in the remote site.
- Backup to an **IBM *TS7650 ProtecTIER* appliance** with data deduplication and remote replication to an identical *TS7650 ProtecTIER* system at the remote site.

Key Assumptions for all use cases

Key Assumptions

- **Labor**: As discussed in the report on page 8, all published studies of tape management labor and disk administrative labor identify a factor of 6-10 between the two. These numbers are typically reported as:
  - **Tape**: 10-15TB/FTE
  - **Disk**: 60-100TB/FTE

  This analysis chose to use a 3x ratio for the difference between disk and tape and a higher starting point for tape in an effort to reduce possible bias. Here is the table of administrative labor used for each backup method based on site capacity.

Backup: Primary TBs managed per FTE

	Small Sites	Medium Sites	Large Sites	Xtra Lg
Tape Library	20	25	40	60
VTL	30	40	55	75
ProtecTIER	60	80	110	150

- **Tape-based backup & DR platform**:
  - Configured with LTO-4 technology and performance.
  - Scalable libraries with extensible frames allowing thousands of slots and hundreds of drives.
  - Compression is in use.
  - 13 week backup pool (with 3 monthly and 1 week of data) and a duplicate configuration at the DR site.
  - No archival activity based on backup. The role of backup is recovery and data availability. Long-term retention is conducted in a proper preservation store outside the domain of this study.

- **VTL backup & DR platform**:
  - The disk space requirements were adjusted annually to support the required backup loads .
  - Remote replication for DR is to an identical system that is kept asynchronously coordinated over the WAN.

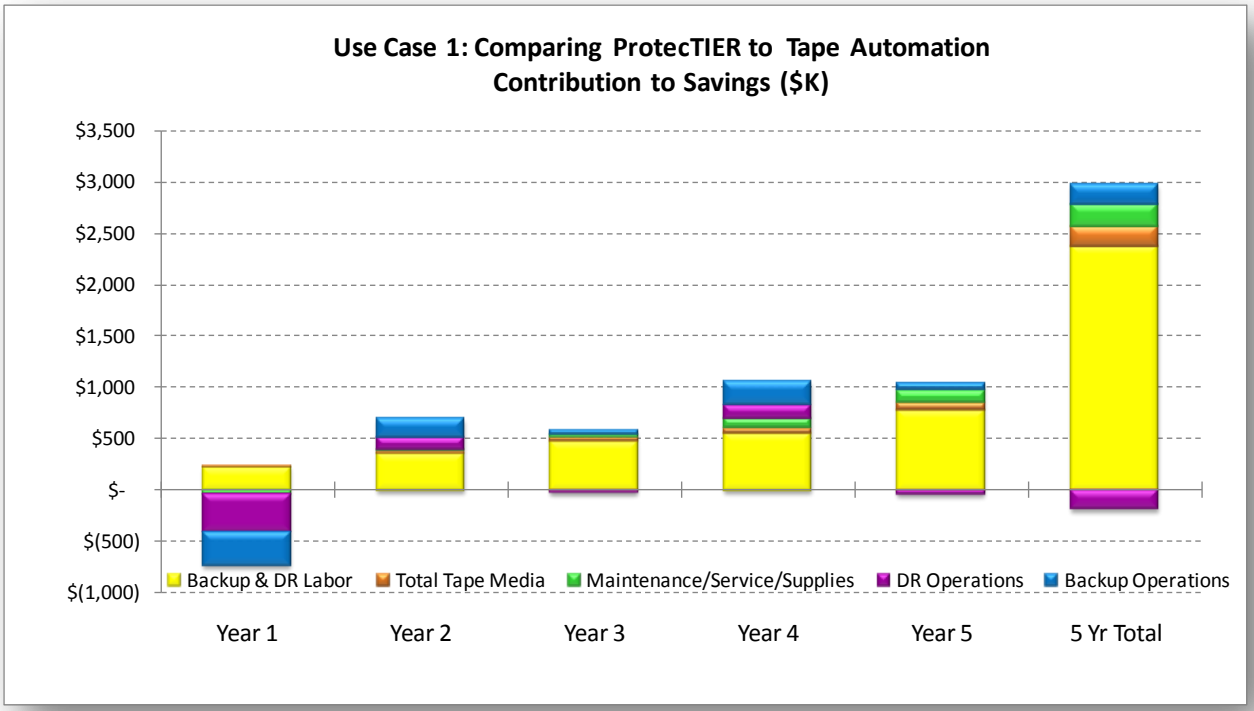
- ***TS7650 ProtecTIER* deduplication data protection platform**:
  - A 20:1 or 95% deduplication plus compression ratio is assumed for the model.
  - Enterprise-class disk arrays were used – specifically, IBM's XIV product line. Market pricing was used throughout the study.

Analysis

Use-Case 1 Analysis:

This is a medium sized site with 25TB capacity under management at the start of the 5 year period. By the end of 5 years the backup repository holds 620TB and it costs $1.8Million in year 5 to operate and manage the backup and DR practices with tape.

- The *TS7650 ProtecTIER*'s payback period is 21 months. This is caused by the relatively substantial capital expense required to purchase *ProtecTIER* solutions for backup and DR in the first year. In comparison, the tape operations have minimal capital expense in year 1 since they start with pre-existing equipment. You'll note in the 'Savings' comparison chart below that operating costs offset the up-front capital in year 2.
- Total 5 year *TS7650 ProtecTIER* backup and DR cost savings over tape at net present value, NPV, are $2.7 million representing a 48% reduction in cost.
- *TS7650 ProtecTIER* DR cost savings over tape are $0.7 million and backup savings are $2.1 million for a total savings of $2.8 million over the 5 years. This is equivalent to a 50% cost savings for backup and a 31% cost savings for DR over tape.
- *TS7650 ProtecTIER* saves 19 man years of labor compared to a tape-only solution over the 5 year period.
- Labor is the top savings category. Only DR operations costs (not including DR labor) is negative compared to tape-based backup and DR due to the capital costs.

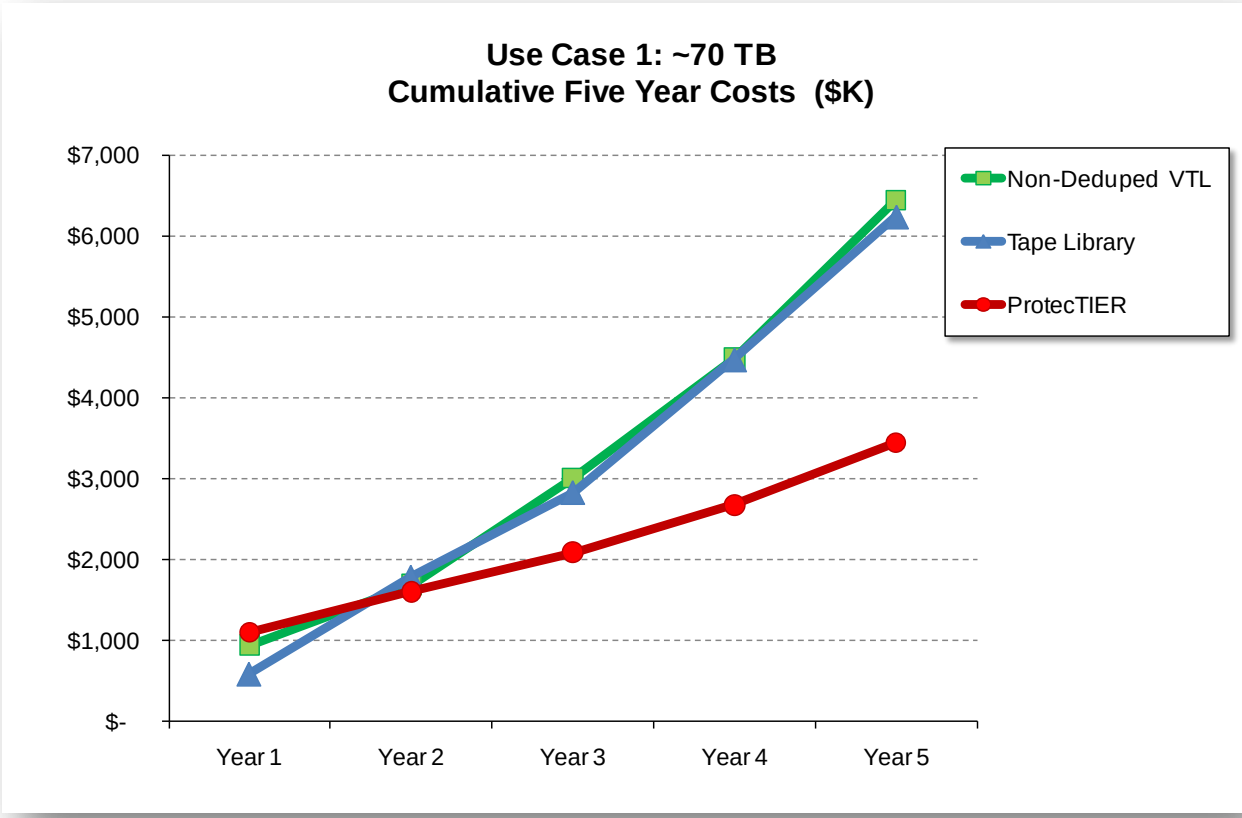


Use-Case 1 Charts:

5 Year TCO COMPARISON

Use Case 1 -- 70 TB (avg.)

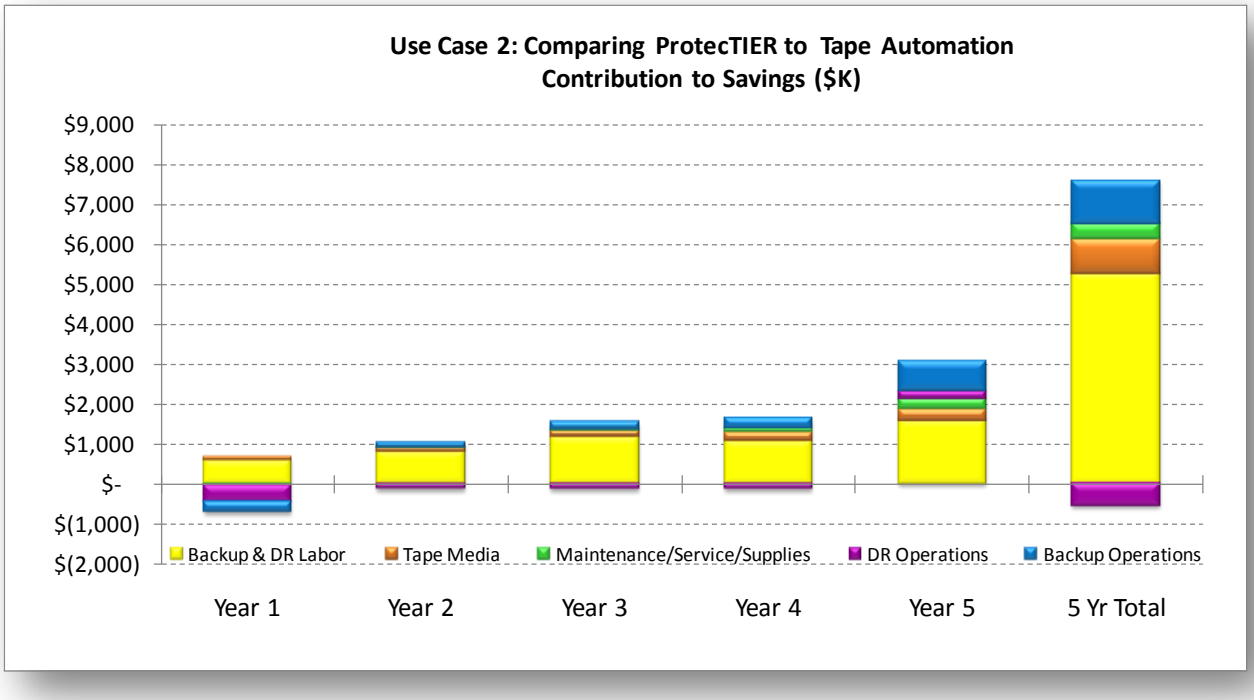
		Tape	VTL	ProtecTIER	Savings
METRICS					
	Backup Cost $/TB/Yr. Managed (Avg.)	$ 10,418	$ 8,492	$ 6,172	$ 4,245
	DR Cost $/TB/Yr. Managed (Avg.)	$ 6,939	$ 9,787	$ 5,948	$ 991
SAVINGS					
	Cumulative 5 Year Costs (TCO @NPV)	$ 5,639,113	$ 5,295,905	$ 2,933,928	$ 2,705,186
	Total Labor (5 man-years)	27.5	16.3	8.5	19.0
	Deferred CapEX				$ 1,636,500
	Deferred Supplies/Maintenance				$ 944,646
	ProtecTIER Payback Period (mo.)	21 months			



Use-Case 2 Analysis:

This is a large sized site with 100TB capacity under management at the start of the 5 year period. By the end of 5 years the backup repository holds 2.4PB and it costs $5.2Million in year 5 to operate and manage the backup and DR practices with tape.

- *TS7650 ProtecTIER* payback period is 6 months. In this use-case the labor savings rapidly offset the initial capital expense.
- The total 5 year *ProtecTIER* backup and DR cost savings over tape at NPV is $6.9Million representing a 49% reduction in cost.
- The *TS7650 ProtecTIER* DR cost savings over tape are $1.3Million and backup savings are $5.8Million for a total savings of $7Million annualized. DR savings are 25% of tape costs and backup savings are 54% of tape cost.
- The *TS7650 ProtecTIER* saves 21.5 man years of labor compared to a tape-only solution over the 5 year period.
- Labor and the cost of backup operations continue as the top two savings categories with $842,000 in tape media offset as the next greatest savings category.

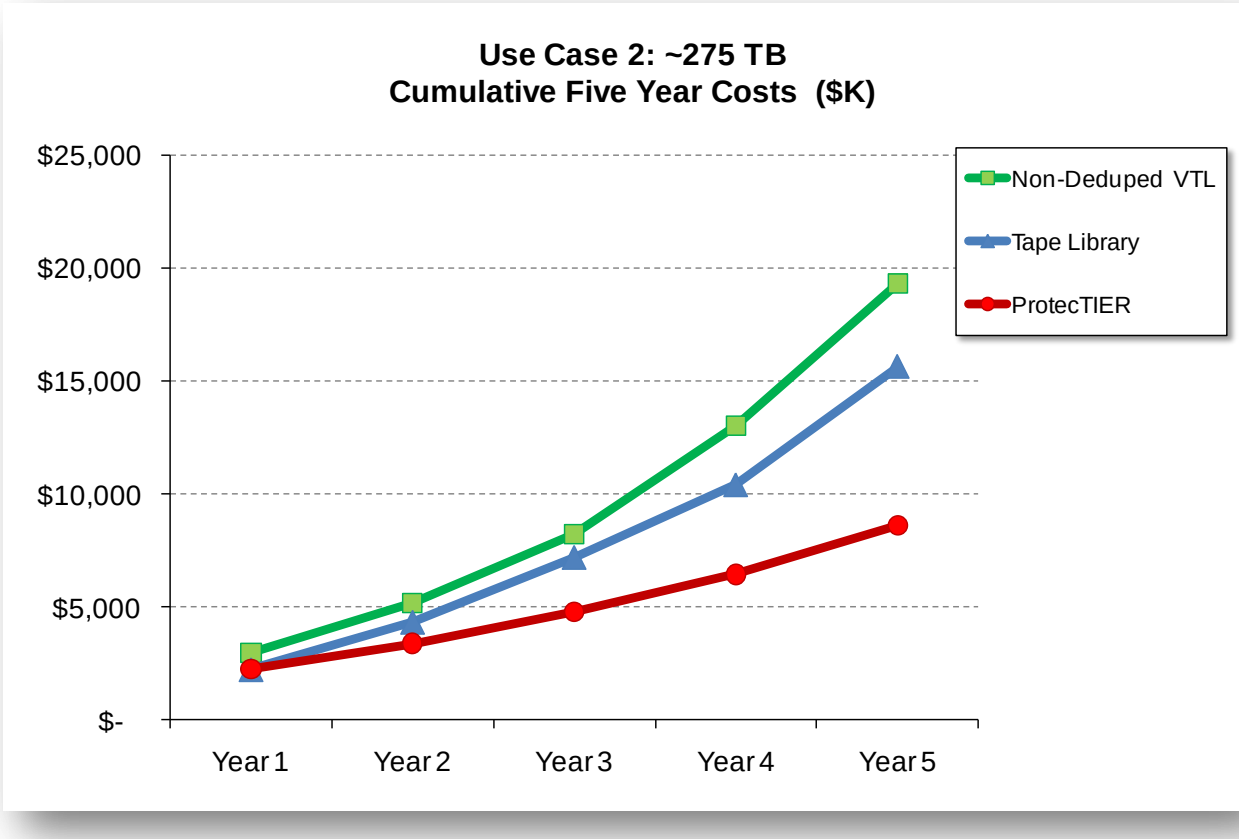


Use-Case 2 Charts:

5 Year TCO COMPARISON

Use Case 2 -- 275 TB (avg.)

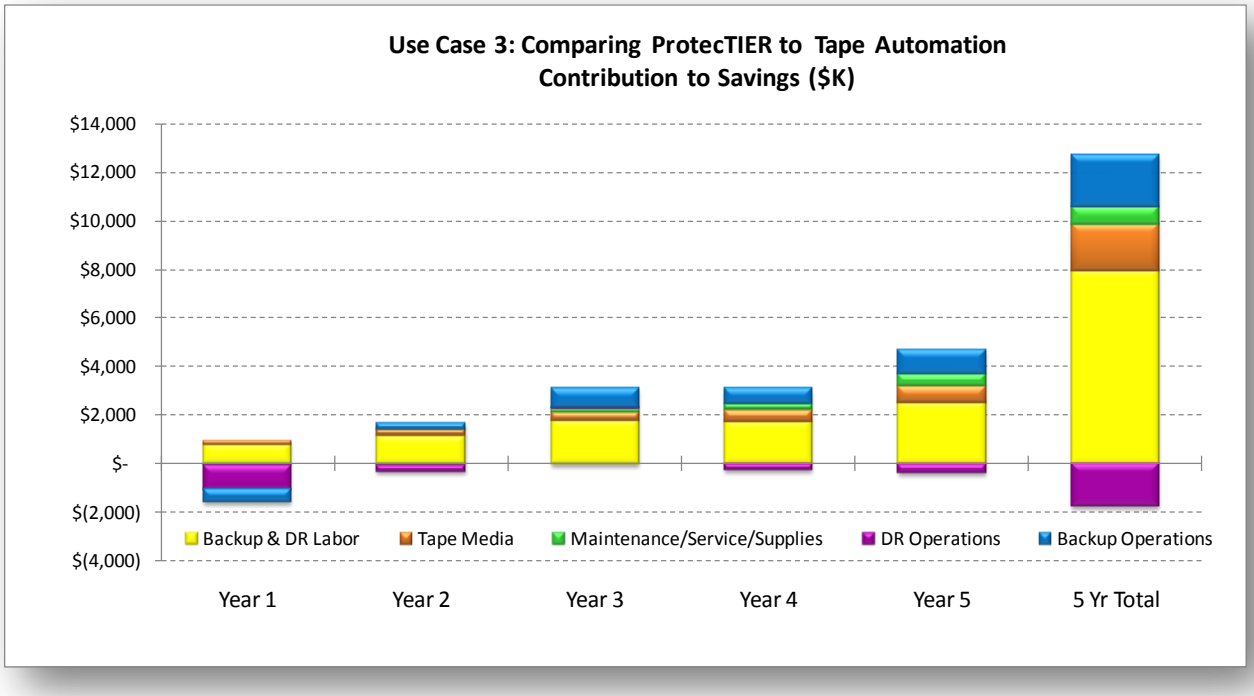
	Tape	VTL	ProtecTIER	Savings
METRICS				
Backup Cost $/TB/Yr Managed (Avg.)	$ 7,453	$ 6,528	$ 3,917	$ 3,536
DR Cost $/TB/Yr Managed (Avg.)	$ 3,582	$ 7,012	$ 3,125	$ 457
SAVINGS				
Cumulative 5 Year Costs (TCO @NPV)	$ 14,125,756	$ 15,820,579	$ 7,219,623	$ 6,906,134
Total Labor (5 man-years)	64.2	45.2	21.5	42.7
Deferred CapEX				$ 4,560,779
Deferred Supplies/Maintenance				$ 2,870,431
ProtecTIER Payback Period (mo.)	6			



Use Case 3 Analysis:

This is a large sized site with 250TB capacity under management at the start of the 5 year period and 1.3PB at the end. By the end of 5 years the backup repository holds 6PB and it costs $9.5Million in year 5 to operate and manage the backup and DR practices with tape.

- This is a Petabyte size backup repository from year 1, growing at 40% per year.
- *TS7650 ProtecTIER* payback period is 18 months because of the large upfront capital expense. Once again, labor is the dominant cost savings.
- Total *TS7650 ProtecTIER* backup and DR cost savings over tape at NPV are $11.1Million representing a 43% reduction in cost.
- *TS7650 ProtecTIER* DR cost savings over tape are $1.1Million and backup savings are $9.9 million for a total savings of $11 million annualized. The backup savings are a 50% cost savings over tape whereas the DR savings are only 12% due principally to the large capital expense with so much capacity online.
- *TS7650 ProtecTIER* saves 40 man years of labor compared to a tape-only solution over the 5 year period.
- Labor and the cost of backup operations continue as the top two savings categories.

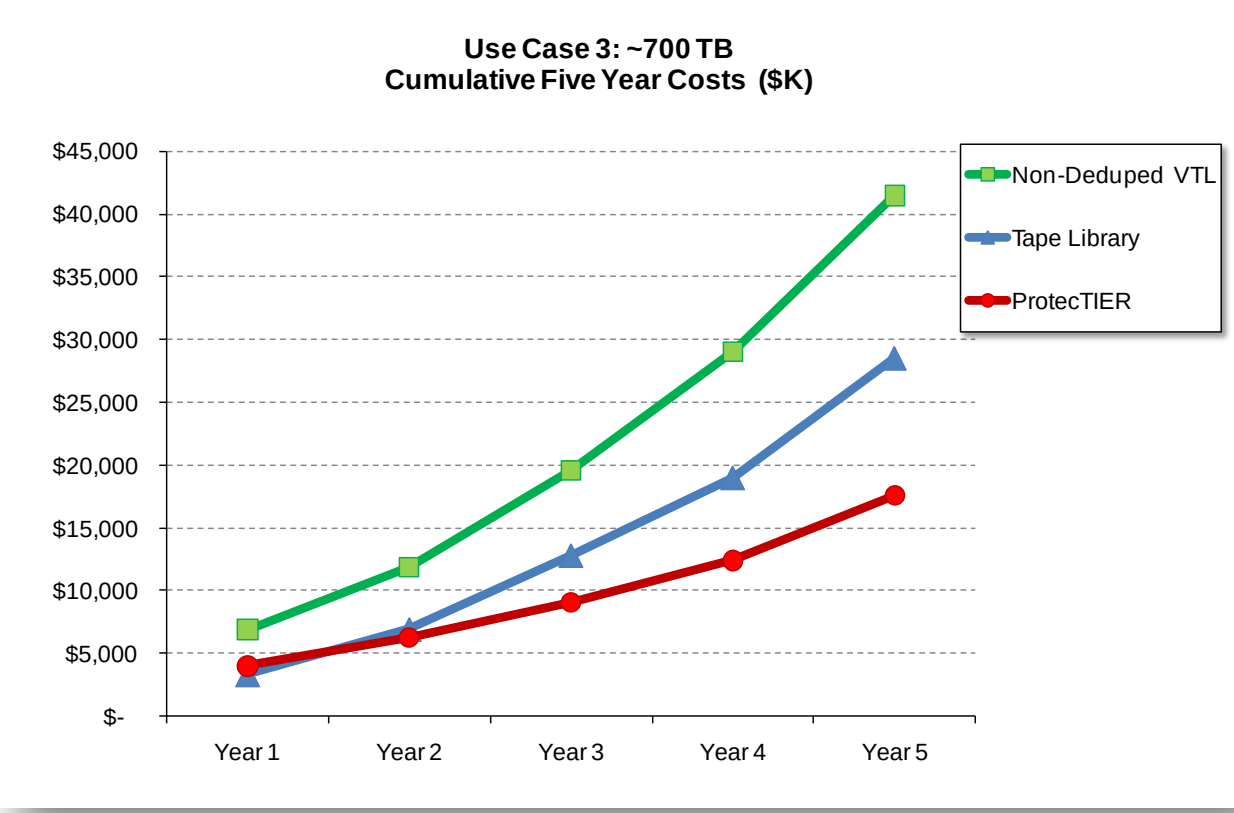


Use-Case 3 Charts:

5 Year TCO COMPARISON

Use Case 3 -- 700 TB (avg.)

	Tape	VTL	ProtecTIER	Savings
METRICS				
Backup Cost $/TB/Yr Managed (Avg.)	$ 5,322	$ 5,727	$ 2,990	$ 2,333
DR Cost $/TB/Yr Managed (Avg.)	$ 2,470	$ 6,309	$ 2,495	$ (25)
SAVINGS				
Cumulative 5 Year Costs (TCO @NPV)	$ 25,728,608	$ 34,194,123	$ 14,651,092	$ 11,077,515
Total Labor (5 man-years)	103.8	87.7	39.8	64.0
Deferred CapEX				$ 9,368,668
Deferred Supplies/Maintenance				$ 6,025,158
ProtecTIER Payback Period (mo.)	18			



IBM System Storage *TS7650 ProtecTIER®* Deduplication Data Protection Platform

IBM Storage and Deduplication

IBM System Storage *TS7650G ProtecTIER* Deduplication Gateway

IBM System Storage *TS7650 ProtecTIER* Deduplication Appliance

The IBM System Storage™ *TS7650G ProtecTIER®* Deduplication Gateway is designed to meet the disk-based data protection needs of the enterprise data center while enabling significant infrastructure cost reductions. The solution offers industry leading inline deduplication performance and scalability up to 1 Petabyte (PB) of raw physical storage capacity per system. Combined with IBM storage, the *ProtecTIER®* Gateway solution provides a powerful disk-based repository to improve the retention and availability of backup and recovery data. *ProtecTIER®* is also available as an appliance equipped with the following features:

- Pre-configured for rapid deployment into existing backup environments
- IBM *TS7650 ProtecTIER®* software with patented *HyperFactor*™ deduplication technology
- IBM System x Server with 18 2.66GHz processor cores, 24 GB RAM & 4GB HBAs
- IBM Storage Controller with 15k 450GB Fibre Channel drives
- IBM TS3000 System Console for call home and remote support
- Complete solution that includes rack, cables, switches, and everything that is needed

ProtecTIER®'s native replication technology enables virtual tape cartridges to be replicated to a remote location for enhanced disaster recovery and business continuity. By eliminating the need to transport physical tape cartridges, data can be recovered faster and more reliably enabling systems to get back online quicker in the event of a disaster or major system outage.

Strategic Research Corporation

PO Box 5365
Santa Barbara, CA. 93150
(805)201-3178
www.sresearch.com

Strategic Research Corporation

Michael Peterson is President of Strategic Research Corporation based in Santa Barbara California. For the past 22 years he has been an energetic leader and catalyst for the storage industry, publishing insightful books and industry reports, consulting with the entire industry in business and market development and IT practices, pioneering IT research on storage and management practices, creating innovative conferences, speaking internationally as an industry visionary, forming industry trade groups, and even developing new solutions and companies. Michael is the founder of the SNIA and was the past president from 1998 to 1999. He continues to work with SNIA and is currently Chief Strategy Advocate guiding market education and development programs. He also consults on IT, long-term retention, preservation, and risk management practices through TechNexxus, a risk management consultancy.