Implementing ProtecTIER 3.3 and IBM FlashSystem
IBM Redbooks Solution Guide

This IBM® Redbooks® Solution Guide describes the benefits that are gained by integrating IBM FlashSystem™ storage with the IBM System Storage® TS7650G ProtecTIER® Deduplication Gateway (TS7650G Gateway), and its core value, benefits, and technological advantages.

FlashSystem products are designed to deliver the lowest latency and the highest IOPS in the market today at an economical price. FlashSystem products deliver fast response time to accelerate critical applications by using IBM MicroLatency™ (that is, roughly 100-microsecond access time), which enables faster decision making by facilitating an extreme-performance data path to accelerate critical applications. FlashSystem technology helps you achieve a true market-based competitive advantage to uncover business insights using traditional data analytics and new, big data technologies.

The TS7650G Gateway is designed to meet the disk-based data protection needs of the enterprise data center while enabling significant infrastructure cost reductions. ProtecTIER is the first virtual tape product to contain the data-factoring technology IBM calls IBM HyperFactor® for data deduplication. The solution offers industry-leading in-line deduplication performance and scalability up to 1 PB of physical storage capacity per system, and can provide up to 25 PB or more backup storage capacity.

IBM FlashSystem is used to accelerate ProtecTIER deployments. Figure 1 shows the key features of both products.
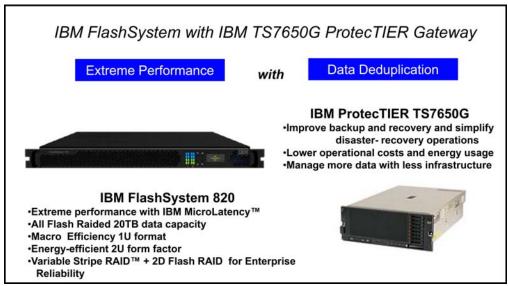

Figure 1. Key features of IBM FlashSystem with TS7650G Gateway

Did you know?

Flash technology enhances the performance, efficiency, reliability, and design of essential enterprise applications and solutions by addressing the bottleneck in the IT process (data storage), enabling a truly optimized information infrastructure.

The IBM HyperFactor deduplication technology that is used by ProtecTIER offers a patented deduplication algorithm (that is not hash-based) that is designed to reduce storage needs by up to 25 times or more, radically reducing costs while maintaining enterprise-class data integrity.

Integrating the IBM FlashSystem with IBM ProtecTIER optimizes storage utilization by using ProtecTIER Deduplication technology, and implementing robust disaster recovery environments and data replication by using the ProtecTIER capabilities.

Integrating IBM FlashSystem with IBM ProtecTIER provides additional capability to work with different applications in your environment, such as Virtual Tape Libraries (VTL), OpenStorage (OST), or File System Interface (FSI).

Business value

IBM FlashSystem shared flash storage systems offer affordable, high-density, ultra low-latency, high-reliability, and scalable performance in a storage device that is both space and power-efficient. IBM FlashSystem products, which can either augment or replace traditional hard disk drive storage systems in enterprise environments, empower applications to work faster and scale further.

In addition to optimizing performance, the IBM FlashSystem family brings enterprise reliability and macro efficiency to the most demanding data centers, allowing businesses to accomplish the following goals:

- Reduce customer complaints by improving application response time.
- Service more users with less hardware.
- Reduce I/O wait and response times of critical applications.
- Simplify solutions.
- Reduce power and floor space requirements.
- Speed up applications and enhancing the pace of business.
- Improve utilization of existing infrastructure.
- Compliment the existing infrastructure.
- Mitigate risk.

The TS7650G Gateway is designed to meet the disk-based data protection needs of the enterprise data center while enabling significant infrastructure cost reductions and offers the following competitive advantages:

- HyperFactor uses a non-hash based approach that performs bit for bit comparison of data before deduplication for the highest data integrity.

- In-line deduplication eliminates redundant data in real time for optimal performance,

- Two-node clustering for high availability.

- Up to 1 PB physical capacity per node.

- Daily operations in-line deduplication eliminates the need for significant secondary processing.

- ProtecTIER technology deduplicates all data (global deduplication).

IBM FlashSystem storage that is integrated with the TS7650G Gateway can provide additional improvements and benefits:

- With the addition of data deduplication, the amount of storage (in this case, FlashSystem) that is required to store data and keep it online can be reduced.

- Performing restores from FlashSystem is much faster than restoring from disk, or from tape for VTL clients.

- Adding ProtecTIER to a running environment with FlashSystem adds the capability to implement strongest disaster recovery scenarios by using the ProtecTIER Replication and Disaster Recover functions.

- Running an environment with ProtecTIER and FlashSystem as back-end storage adds the ability to use applications, such as Virtual Tape Library (VTL), OpenStorage technology (OST), or the File System Interface (FSI), which supports Common Internet File System (CIFS) and Network File System (NFS) as backup targets.

From the client business perspective, an IBM FlashSystem provides focus benefits and value in four essential areas:

- **Extreme performance**: Enable business to unleash the power of performance, scale, and insight to drive services and products to market faster.

- **MicroLatency**: Achieve competitive advantage through applications that enable faster decision making because of microsecond response times.

- **Macro Efficiency**: Decrease costs by getting more from efficient usage of IT staff, IT applications, and IT equipment because of the efficiencies flash brings to the data center.

- **Enterprise reliability**: Durable and reliable designs that use enterprise class flash and patented data protection technology.

Solution overview

IBM FlashSystem can be integrated with the TS7650G Gateway to make backup environments faster and more efficient. In this solution, the IBM FlashSystem is attached as back-end storage for the TS7650G Gateway, adding its high performance, efficiency, and reliability for shared enterprise storage environments to the backup environment that ProtecTIER provides.

The TS7650G Gateway also provides disaster recovery capabilities that can be used by FlashSystem clients by replicating backup data to a different location. The ProtecTIER IP replication function provides a powerful tool that can be used to design robust disaster recovery architectures. Clients electronically place backup data into vaults with much less network bandwidth, thus changing the paradigm of how data is taken off-site for safe keeping.

For VTL clients, the ProtecTIER IP replication feature can eliminate some of the expensive and labor-intensive handling, transport, and securing of the real tapes for disaster recovery purposes. For OST and FSI clients, they can implement enhanced functionality that is possible only with FlashSystem products and avoid the limitations of tape emulation.

The TS7650G Gateway back-end storage array is configured with two types of LUNs: metadata and user data. Metadata LUNs are used to record where data is kept and user data LUNs are used to store the actual data. Metadata LUN performance is critical and the high performance that is provided by the IBM FlashSystem in this solution is the key element to achieve a higher performance.

Solution architecture

In this solution, the IBM FlashSystem works as the back-end storage for the TS7650G Gateway. The clients can use any backup application that is supported by ProtecTIER. These clients send the data through a Fibre Channel (FC) SAN for VTL environments, or a 1 Gbps / 10 Gbps Ethernet network for OST and FSI environments. This data is received by the TS7650G Gateway, deduplicated and compressed by using the ProtecTIER capabilities, and finally sent to the IBM Flash System for Storage. Figure 2 shows an overview of this integration.
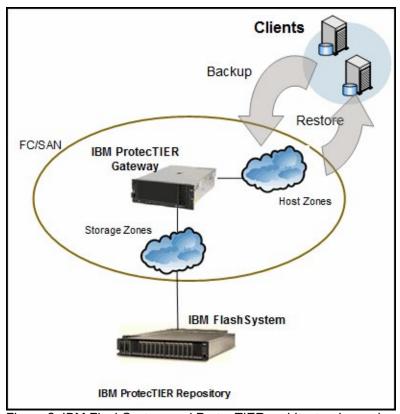

Figure 2. IBM FlashSystem and ProtecTIER architectural overview

The process to set up the environment to integrate the IBM ProtecTIER system and the IBM FlashSystem is not different from other back-end storage systems; the general process consists of the following steps, which are also shown in Figure 3.
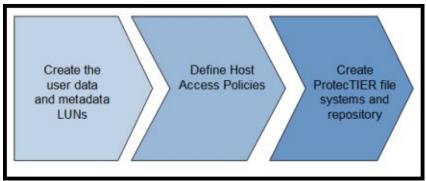

Figure 3. IBM FlashSystem and ProtecTIER configuration steps

1. Create the user data and metadata logical units in the IBM FlashSystem.
2. Define Host Access Policies for each one of the logical units that are created.
3. Create the file systems and build the ProtecTIER repository

Usage scenarios

IBM FlashSystem can be used by the TS7650G Gateway as the back-end storage device for metadata and user data LUNs for optimal performance.

A common use case for IBM FlashSystem is for the TS7650G Gateway metadata LUNs. Compared to the cost of dozens of 15 K or 10 K HDD spindles, it can be more cost-effective to use IBM FlashSystem products for the TS7650G Gateway metadata LUNs. It might also be more cost-effective to use IBM FlashSystem for the entire TS7650G Gateway repository if high performance and small capacity are needed. The FlashSystem can be connected directly to the TS7650G Gateway and be used as the back-end storage. It can also be integrated with the IBM SAN Volume Controller and configured for the following scenarios.

Scenario 1: IBM SAN Volume Controller and IBM FlashSystem as the back -end storage for the TS7650G Gateway for all FlashSystem storage systems

An all FlashSystem SAN Volume Controller solution, in which the FlashSystem storage system is the only storage device being virtualized by SAN Volume Controller, and IBM Easy Tier®, is not enabled.

The main consideration if you use FlashSystem storage systems in an environment where Easy Tier is not enabled is to treat the storage system as any other virtualized storage device, and place it in its own storage pool with a dedicated SAN Volume Controller I/O group.

Scenario 2: IBM FlashSystem with SAN Volume Controller with Easy Tier as back -end storage for the TS7650G Gateway

FlashSystem storage is used with Easy Tier to improve performance across all storage pools within an existing environment. In this scenario, FlashSystem storage can be used as tier 0 to store the most frequently used (hottest) data. Here are some general recommendations for this scenario:

● Use FlashSystem storage when your capacity requirements for Easy Tier exceed five SSDs. At this point, FlashSystem storage systems should be considered for cost efficiency and performance impact reasons.

● The recommended maximum MDisk allocation for a single FlashSystem is 16 MDisks. Use your Easy Tier capacity requirements, along with this recommendation, to determine the number of FlashSystem storage systems that are required.

For more preferred practices about these scenarios, see *IBM SAN Volume Controller and FlashSystem 820*, SG24-8172.

Integration

Integrating IBM FlashSystem and TS7650G Gateway in your backup environment allows you to use different backup applications to manage your backups and restore needs. Here are some of the supported backup applications:

● IBM Tivoli® Storage Manager

   IBM Tivoli Storage Manager is a data protection platform that gives enterprises a single point of control and administration for storage management needs. It is the flagship product in the IBM Tivoli Storage Manager family. This advanced, highly scalable product helps increase the efficiency of your IT operations and helps cut costs that are related to storage management by providing a wide range of data protection, recovery management, movement, retention, reporting, and monitoring capabilities using policy-based automation.

   IBM Tivoli Storage Manager enables outstanding efficiency, simplicity, and scalability for virtual, physical, and cloud backup environments of all sizes.

● Symantec NetBackup

   Symantec NetBackup (NBU) is an Open Systems Enterprise backup software solution. Its architecture includes the following main building blocks:

   o   Clients: The machines with the data that requires backing up
   o   Media servers: The machines that are connected to the backup devices
   o   Master server: The machine that is controlling the backups

   Collectively, master, media, and clients are known as an NBU Domain.

● Backup, Recovery, and Media Services for IBM i

   Backup, Recovery, and Media Services (BRMS) for IBM i is the IBM strategic solution for planning and managing the backup of your IBM i server.

BRMS provides the IBM i server with support for policy-oriented setup and execution of backup, recovery, archive, and other removable-media-related operations. BRMS uses a consistent set of intuitive concepts and operations that can be used to develop and implement a backup strategy that is tailored to your business requirements. The user interface is menu-driven, with a significant number of functions that are enabled through the optional IBM Systems Director Navigator for i.

● EMC NetWorker

EMC NetWorker (NetWorker), formerly Legato NetWorker, is a centralized, automated backup and recovery product for heterogeneous enterprise data. The NetWorker Server runs on all major operating systems, such as IBM AIX®, Linux, Windows, SUN Solaris, and HP-UX.

For more information about the specific version that is supported, you can review the TS7650G Gateway Interoperability Matrix found at the following website:
http://public.dhe.ibm.com/common/ssi/ecm/en/ivl12348usen/IVL12348USEN.PDF

Supported platforms

For the latest information about the supported systems and platforms, see the IBM System Storage Interoperation Center (SSIC) at the following website:
http://www.ibm.com/systems/support/storage/ssic/interoperability.wss

Ordering information

Ordering information, including program numbers for the IBM products that are used in this solution, is shown in Table 1.

Table 1. Ordering part numbers and feature codes

Program name	Machine-type/ model or PID	Description
ProtecTIER Enterprise Edition V3.3 Base Software	5639-PTA	5639-PTA ProtecTIER Enterprise Edition V3.3 Base Software.
TS7650G Server IBM System Storage TS7650G Gateway	3958 DD5	Gateway server - x3850X5 (MT 7143) - Seven 5.0 Gb PCIe I/O (one x16, five x8, one x4) sockets. Four 10-core Intel Xeon E7-4870 processors, 2.26 GHz, 64 GB of memory.
IBM FlashSystem 820	9381-AE2	The IBM 9831 FlashSystem 820 Model AE2 is a storage system that occupies 1U of standard 19-inch rack space and provides 10 or 20 TB of usable RAID 5 protected capacity (12 or 24 TB with RAID 0), based on eMLC Flash technology.
SAN Volume Controller	2145-CG8	IBM 2145 System Storage SAN Volume Controller Storage Engine, Quantity 2

Related information

For more information, see the following documents:

- *IBM ProtecTIER Implementation and Best Practices Guide*, SG24-8025
  http://www.redbooks.ibm.com/redpieces/pdfs/sg248025.pdf

- IBM System Storage TS7650G ProtecTIER Deduplication Gateway
  http://www-03.ibm.com/systems/storage/tape/ts7650g/index.html

- IBM Flash Storage and Solutions
  http://www-03.ibm.com/systems/storage/flash/?lnk=mprST-flsy-usen

- *IBM FlashSystem 720 and IBM FlashSystem 820 Product Guide*
  http://www.redbooks.ibm.com/technotes/tips1003.pdf

- *Implementing the IBM SAN Volume Controller and FlashSystem 820*, SG24-8172
  http://www.redbooks.ibm.com/redbooks/pdfs/sg248172.pdf

- IBM Offering Information page (announcement letters and sales manuals)
  http://www.ibm.com/common/ssi/index.wss?request_locale=en

  On this page, enter TS7650G for ProtecTIER, or FlashSystem 820 for the FlashSystem product, select the information type, and then click **Search**. On the next page, narrow your search results by geography and language.

Notices

This information was developed for products and services offered in the U.S.A.

IBM may not offer the products, services, or features discussed in this document in other countries. Consult your local IBM representative for information on the products and services currently available in your area. Any reference to an IBM product, program, or service is not intended to state or imply that only that IBM product, program, or service may be used. Any functionally equivalent product, program, or service that does not infringe any IBM intellectual property right may be used instead. However, it is the user's responsibility to evaluate and verify the operation of any non-IBM product, program, or service. IBM may have patents or pending patent applications covering subject matter described in this document. The furnishing of this document does not give you any license to these patents. You can send license inquiries, in writing, to:

IBM Director of Licensing, IBM Corporation, North Castle Drive, Armonk, NY 10504-1785 U.S.A.

The following paragraph does not apply to the United Kingdom or any other country where such provisions are inconsistent with local law: INTERNATIONAL BUSINESS MACHINES CORPORATION PROVIDES THIS PUBLICATION "AS IS" WITHOUT WARRANTY OF ANY KIND, EITHER EXPRESS OR IMPLIED, INCLUDING, BUT NOT LIMITED TO, THE IMPLIED WARRANTIES OF NON-INFRINGEMENT, MERCHANTABILITY OR FITNESS FOR A PARTICULAR PURPOSE. Some states do not allow disclaimer of express or implied warranties in certain transactions, therefore, this statement may not apply to you. This information could include technical inaccuracies or typographical errors. Changes are periodically made to the information herein; these changes will be incorporated in new editions of the publication. IBM may make improvements and/or changes in the product(s) and/or the program(s) described in this publication at any time without notice.

Any references in this information to non-IBM Web sites are provided for convenience only and do not in any manner serve as an endorsement of those Web sites. The materials at those Web sites are not part of the materials for this IBM product and use of those Web sites is at your own risk.IBM may use or distribute any of the information you supply in any way it believes appropriate without incurring any obligation to you. Information concerning non-IBM products was obtained from the suppliers of those products, their published announcements or other publicly available sources. IBM has not tested those products and cannot confirm the accuracy of performance, compatibility or any other claims related to non-IBM products. Questions on the capabilities of non-IBM products should be addressed to the suppliers of those products. This information contains examples of data and reports used in daily business operations. To illustrate them as completely as possible, the examples include the names of individuals, companies, brands, and products. All of these names are fictitious and any similarity to the names and addresses used by an actual business enterprise is entirely coincidental.

Any performance data contained herein was determined in a controlled environment. Therefore, the results obtained in other operating environments may vary significantly. Some measurements may have been made on development-level systems and there is no guarantee that these measurements will be the same on generally available systems. Furthermore, some measurement may have been estimated through extrapolation. Actual results may vary. Users of this document should verify the applicable data for their specific environment.

COPYRIGHT LICENSE:

This information contains sample application programs in source language, which illustrate programming techniques on various operating platforms. You may copy, modify, and distribute these sample programs in any form without payment to IBM, for the purposes of developing, using, marketing or distributing application programs conforming to the application programming interface for the operating platform for which the sample programs are written. These examples have not been thoroughly tested under all conditions. IBM, therefore, cannot guarantee or imply reliability, serviceability, or function of these programs.

This document was created or updated on March 12, 2014.

Send us your comments in one of the following ways:
- Use the online **Contact us** review form found at:
  ibm.com/redbooks
- Send your comments in an e-mail to:
  redbook@us.ibm.com
- Mail your comments to:
  IBM Corporation, International Technical Support Organization
  Dept. HYTD Mail Station P099
  2455 South Road
  Poughkeepsie, NY 12601-5400 U.S.A.

This document is available online at http://www.ibm.com/redbooks/abstracts/tips1140.html .

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. These and other IBM trademarked terms are marked on their first occurrence in this information with the appropriate symbol (® or ™), indicating US 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 http://www.ibm.com/legal/copytrade.shtml.

The following terms are trademarks of the International Business Machines Corporation in the United States, other countries, or both:

AIX®
Easy Tier®
FlashSystem™
HyperFactor®
IBM FlashSystem™
IBM®
MicroLatency™
ProtecTIER®
Redbooks®
Redbooks (logo)®
System Storage®
Tivoli®

The following terms are trademarks of other companies:

Intel Xeon, Intel, Intel logo, Intel Inside logo, and Intel Centrino logo are trademarks or registered trademarks of Intel Corporation or its subsidiaries in the United States and other countries.

Linux is a trademark of Linus Torvalds in the United States, other countries, or both.

Windows, and the Windows logo are trademarks of Microsoft Corporation in the United States, other countries, or both.

Other company, product, or service names may be trademarks or service marks of others.