

WHITE PAPER

Protecting SQL Server with FlashBlade

A framework for achieving rapid backup and recovery with native SQL Server tooling.

Contents

Introduction	3
Pure Storage FlashBlade	3
Solution Architecture	5
Background	5
Backup and Recovery with FlashBlade and Native SQL Server T-SQL Scripts	6
T-SQL Backup and Recovery with FlashBlade Validation	9
Specifications	9
Validation & Analysis	11
Limitations	11
Single Server Validation	12
Conclusion	16
Additional Resources	16



Introduction

In today's digital era, data protection is no longer just meeting legal compliance, it's essential for success. Strong data protection keeps your information safe from unauthorized access and empowers business continuity by mitigating data loss. Prioritizing data security fosters trust, innovation and resilience in the digital age.

Microsoft SQL Server is the cornerstone of data management for countless organizations across the globe. Its powerful features and scalability make it ideal for storing and managing critical business data, supporting mission-critical applications and driving valuable insights. However, in today's dynamic and data-driven world, ensuring the speedy recovery of vital information in the event of data loss is paramount.

Pure Storage® FlashBlade® integrates with SQL Server's native backup capabilities, offering comprehensive data protection. It empowers faster data recovery, enhances resilience through replication, and facilitates long-term data retention. FlashBlade's flexible access, and advanced compression technology ensure continuous availability, resilience, and efficient management for SQL Server environments.

FlashBlade provides unmatched flexibility for SQL Server's native backup by offering options to use SMB file shares or S3 object storage as backup targets. FlashBlade with on-board hardware compression maximizes your storage efficiency. But the real magic lies in Rapid Restore, delivering data recovery at very high speeds.

This technical white paper explores the capabilities of Pure Storage FlashBlade in solving large scale SQL Server database data protection challenges. This document serves as a roadmap for safeguarding Microsoft SQL Server environments with concise overview of baseline performance and provides insights into integrating Pure Storage FlashBlade with Microsoft SQL Server for optimal database backup and recovery performance using SQL Server native backup.

Pure Storage FlashBlade

The Pure Storage FlashBlade is an advanced storage solution designed for modern data protection needs. With its unique blade architecture prioritizing speed and efficiency. FlashBlade offers exceptional performance and scalability. Its simplicity and adaptability simplifies and enables organizations to scale storage infrastructure to satisfy needs of data-intensive workloads like analytics, AI and machine learning.

Featuring an all-flash design, FlashBlade consistently delivers high performance, supporting various workloads from backup to analytics and AI. The system's hyper-scalable, all-flash performance is facilitated by a purpose-built distributed file system, supporting diverse data types and simplifying storage management.

FlashBlade seamlessly integrates into different environments, providing robust security measures for sensitive information. Environment-friendly models cater to specific business requirements, optimizing downtime, maximizing data value, and ensuring efficient data backup and recovery for seamless business continuity.



This unified file and object powerhouse lets you craft backup sets for diverse needs, like SafeMode™ snapshots, Multi-protocol support and greater resiliency to replicate your backups to another FlashBlade or even AWS S3, securing your data across three locations when disaster strikes.

FlashBlade//S™ represents the evolution in enterprise scale-out storage, offering a blend of high density, capacity, performance, and scalability to cater to the requirements of contemporary applications. With a distributed metadata architecture, FlashBlade//S provides multidimensional performance on a unified platform, supporting access through NFS, SMB, and S3 protocols.



FlashBlade//E™ emerges as a revolutionary force in enterprise storage, aiming to make cost-efficient performance, uncompromising reliability, and all-flash capabilities accessible to a broader audience, particularly those with budget considerations. It is a cost-effective storage solution within the FlashBlade product family, bringing unparalleled value to data protection and unstructured data scenarios.



Solution Architecture

This solution architecture provides performance capabilities of SQL Server Native backup with FlashBlade using SMB and S3 as backup targets.

Using Server Message Block (SMB) as a backup target provides a versatile and widely supported protocol, enabling seamless integration with various operating systems. It allows for straightforward configuration, making it user-friendly for administrators. SMB's user-level authentication enhances security, ensuring that backup access is controlled and authenticated.

Using S3-compatible object storage as a backup target provides a scalable and versatile solution. It allows seamless integration with systems supporting the S3 protocol, offering greater scalability and flexibility. The compatibility ensures a straightforward transition for users familiar with S3, enabling efficient backup and recovery processes. Additionally, S3-compatible storage inherits the durability and reliability associated with the S3 standard, making it a reliable choice for secure and scalable backup storage.

Introducing the collaborative power of SQL Server and Pure Storage FlashBlade, an advanced storage solution that takes data protection to unprecedented levels, providing straightforward and efficient data services. FlashBlade's high-performance design seamlessly integrates with and improves SQL Server's capabilities, establishing a robust synergy to fortify against accidental data loss and unexpected system issues. Additionally, its fast recovery technology adds extra value to the solution.

Background

Traditional data protection methodologies employed for Microsoft SQL Server are ill equipped to cope with rapidly growing capacity and performance needs. A close look at these situations reveals limitations that act as strong barriers to success.

Embarking on an exploration of these constraints, it is clear that these older methods have difficulty adapting to and handling the complex issues brought by today's dynamic business environments. The ensuing analysis sheds light on these critical limitations, providing a basis for understanding the obstacles that affect the efficiency of traditional data protection methods for SQL Server.

Extended recovery times:

Conventional backup systems focus on rapid data protection, often overlooking recovery. With the acceleration of digital transformation, organizations now concentrate intensely on minimizing downtime for business-critical applications to maximize the value derived from their data. Slow recovery adversely affects business operations, introducing a bottleneck that hampers efficiency and responsiveness.

Complicated and Inefficient storage:

The absence of data reduction in traditional methods creates storage management issues, resulting in inefficient use of storage space and other related problems. This not only wastes storage resources but also increases operational costs, posing significant challenges in terms of scalability and resource optimization.



Limited data services:

Traditional storage solutions fall short in providing performance storage solutions with comprehensive data services such as instant recovery, multi-site replication, cloud integration, compression, and granular application recovery exposes data to heightened vulnerability in the face of unforeseen events.

These identified challenges underscore a compelling need for a sophisticated data protection storage solution specifically tailored for SQL Server's evolving requirements.

Backup and Recovery with FlashBlade and Native SQL Server T-SQL Scripts

FlashBlade accelerates Microsoft SQL Server backups with its parallel architecture and flash technology, ensuring rapid data transfer and reduced backup times. The combination of native SQL Server backups with FlashBlade provides the following benefits for SQL Server environments. The following outcomes can be realized when combining SQL Server T-SQL backup and recovery with FlashBlade file and object storage:

- FlashBlade Single Chassis provides rapid restore for SQL Server workloads, achieving a data recovery performance of 60 TB/hr for scalable, parallel operations
- Multi-protocol support for file and object allowing for flexible data protection and disaster recovery strategies.
- Allows greater resiliency for file shares and objects to replicate to another FlashBlade or AWS S3 buckets that offers disaster recovery capabilities.

Figure 1 shows high-level information of SQL Server setup and FlashBlade configured to store database backups on SMB and S3.

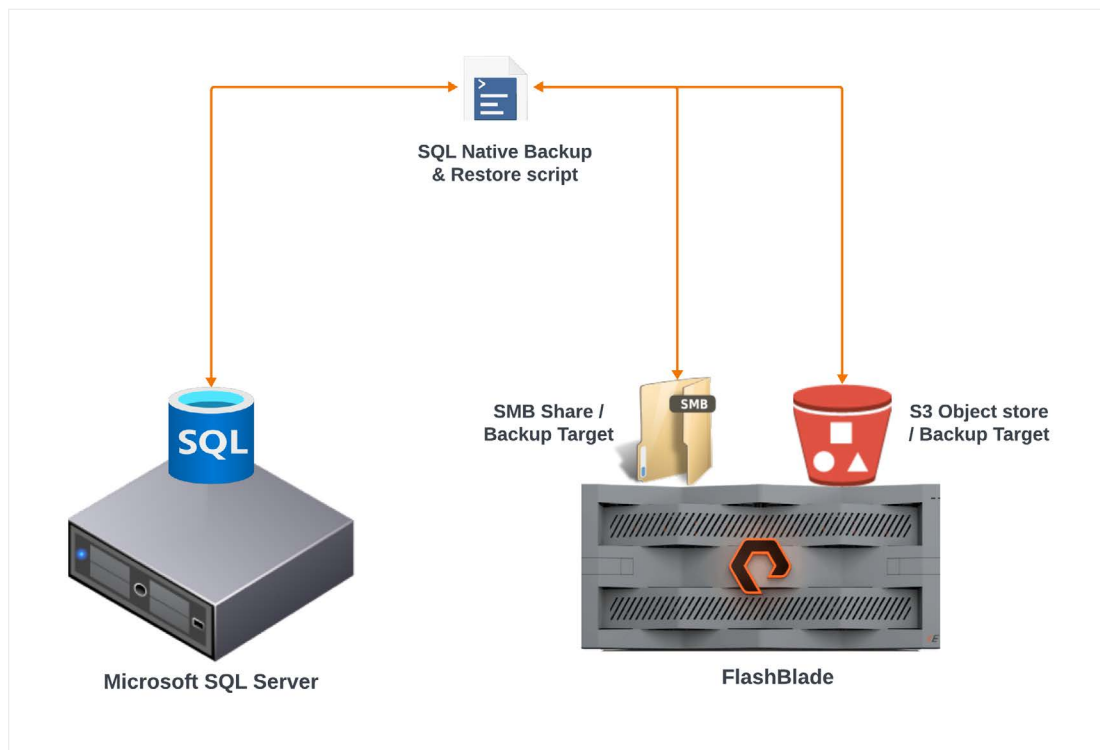


FIGURE 1 High-level view of Microsoft SQL Server and FlashBlade setup



Quick Start : Using Windows File Shares (SMB):

The combination of SQL Server Native Backup capabilities and FlashBlade file storage (SMB) represents a significant advancement in Windows storage efficiency and speed.

For additional information on creating SMB share on FlashBlade, see the knowledge article [How to create SMB on FlashBlade](#)

The choice between a single data VIF and multiple data VIFs depends on factors such as simplicity, throughput requirements, load balancing needs, concurrency goals, and the desire for redundancy and scalability.

Multiple data VIFs are often favored when seeking to optimize performance, efficiency, and resilience in backup and restore operations on a FlashBlade SMB Share.

The following T-SQL command can be used to create a full database backup to an SMB file share using multiple data VIFs.

```
BACKUP DATABASE <database_name> TO
    DISK = '\\data-vif1\<fileshare_name>\db_01.Bak',
    DISK = '\\data-vif2\<fileshare_name>\db_02.Bak',
    DISK = '\\data-vif3\<fileshare_name>\db_03.Bak',
    DISK = '\\data-vif4\<fileshare_name>\db_04.Bak',
    DISK = '\\data-vif5\<fileshare_name>\db_05.Bak',
    DISK = '\\data-vif6\<fileshare_name>\db_06.Bak',
    DISK = '\\data-vif7\<fileshare_name>\db_07.Bak',
    DISK = '\\data-vif8\<fileshare_name>\db_08.Bak'
WITH INIT, [NO_COMPRESSION | COMPRESSION, BLOCKSIZE=<block_size>, BUFFERCOUNT=<buffer_count>]
```

The following T-SQL command can be used To restore/replace a database with full backup from SMB file share using multiple data VIFs.

```
RESTORE DATABASE <database_name> FROM
    DISK = '\\data-vif1\<fileshare_name>\db_01.Bak',
    DISK = '\\data-vif2\<fileshare_name>\db_02.Bak',
    DISK = '\\data-vif3\<fileshare_name>\db_03.Bak',
    DISK = '\\data-vif4\<fileshare_name>\db_04.Bak',
    DISK = '\\data-vif5\<fileshare_name>\db_05.Bak',
    DISK = '\\data-vif6\<fileshare_name>\db_06.Bak',
    DISK = '\\data-vif7\<fileshare_name>\db_07.Bak',
    DISK = '\\data-vif8\<fileshare_name>\db_08.Bak'
WITH REPLACE
```

Usage of SQL Server backup parameters MAXTRANSFERSIZE, BLOCKSIZE, and BUFFERCOUNT will be infrastructure specific and appropriate configurations for customer environment will be determined through testing.

[The Microsoft SQL Server documentation](#) provides additional details on backup parameters to further tune backup and restore operations.



Quick Start : Using Object Storage (S3):

SQL Server 2022 seamlessly integrates with FlashBlade object storage, ensuring excellent scalability for growing data center needs. The introduction of S3-compatible Object Storage for SQL Server 2022 facilitates backup and recovery workflows through a URL syntax, leveraging the S3 REST API for integration with any provider of S3-compatible object storage. Unlike traditional methods relying on CPU or faster disks, the S3 API scale-out backups enhance performance by adding compute nodes as the system scales. With rapid recovery speeds, off-site resiliency options, and tamper-proofing features such as Object lock, SQL Server 2022 on FlashBlade offers a comprehensive solution for efficient and secure data management.

For additional information on creating S3 Buckets on FlashBlade, refer to the knowledge article [How to create S3 bucket in FlashBlade](#)

For large databases, using FlashBlade object store as backup target for SQL Server is simple to configure and enables seamless communication using HTTPS protocols to transfer data securely. It improves performance by dividing the backup into multiple files, allowing concurrent operations. This capability is linked to endpoints (URLs) associated with the object store.

The following T-SQL command can be used to create a full database backup in S3 bucket

```
BACKUP DATABASE <database_name> TO
    URL= 's3://data-vif/<bucket_name>/<folder_name>/db_01.Bak',
    URL= 's3://data-vif/<bucket_name>/<folder_name>/db_02.Bak',
    URL= 's3://data-vif/<bucket_name>/<folder_name>/db_03.Bak',
    URL= 's3://data-vif/<bucket_name>/<folder_name>/db_04.Bak',
    URL= 's3://data-vif/<bucket_name>/<folder_name>/db_05.Bak',
    URL= 's3://data-vif/<bucket_name>/<folder_name>/db_06.Bak',
    URL= 's3://data-vif/<bucket_name>/<folder_name>/db_07.Bak',
    URL= 's3://data-vif/<bucket_name>/<folder_name>/db_08.Bak'
WITH INIT, [NO_COMPRESSION | COMPRESSION, MAXTRANSFERSIZE=<maxtransfersize>]
```

The following T-SQL command can be used to restore/replace a database with full backup from S3 bucket.

```
RESTORE DATABASE <database_name> FROM
    URL= 's3://data-vif/<bucket_name>/<folder_name>/db_01.Bak',
    URL= 's3://data-vif/<bucket_name>/<folder_name>/db_02.Bak',
    URL= 's3://data-vif/<bucket_name>/<folder_name>/db_03.Bak',
    URL= 's3://data-vif/<bucket_name>/<folder_name>/db_04.Bak',
    URL= 's3://data-vif/<bucket_name>/<folder_name>/db_05.Bak',
    URL= 's3://data-vif/<bucket_name>/<folder_name>/db_06.Bak',
    URL= 's3://data-vif/<bucket_name>/<folder_name>/db_07.Bak',
    URL= 's3://data-vif/<bucket_name>/<folder_name>/db_08.Bak'
WITH REPLACE
```

Usage of SQL Server backup parameters MAXTRANSFERSIZE, BLOCKSIZE, and BUFFERCOUNT will be infrastructure specific and appropriate configurations for customer environment will be determined through testing.

[The Microsoft SQL Server documentation](#) provides additional details on backup parameters to further tune backup and restore operations.



T-SQL Backup and Recovery with FlashBlade Validation

In this solution validation, the performance capabilities of FlashBlade//S200, //S500 and //E are assessed for potential outcomes. The primary database(s) are hosted on high performing all-flash storage with 2 x Pure Storage FlashArray™. The connection between FlashArray and SQL Server hosts use Fibre Channel connectivity, while data backup and restore operations on FlashBlade occur over a separate ethernet network. This configuration ensures dedicated resource paths for both database storage and backup storage, ultimately optimizing performance. In this validation, each FlashBlade is configured to store SQL Server backups in both SMB share and S3 object store (bucket).

Figure 2 illustrates the high-level architecture of the environment used for validation.

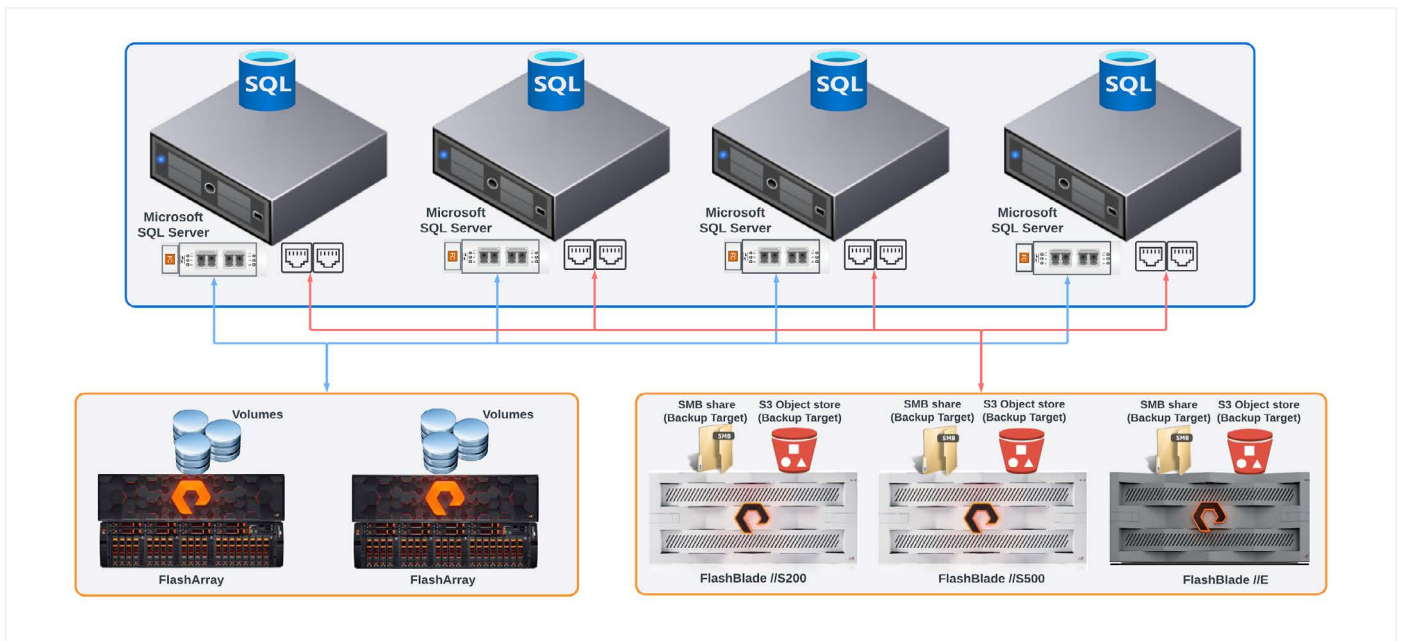


FIGURE 2 Overview of the validation environment architecture

Specifications

The potential performance outcome of database backup and restore operations using SQL Server native backup relies on a variety of factors. Below are the specifications for this environment:

Primary Storage: FlashArray is connected to Windows-SQL Server(s) via Fibre Channel to allocate primary storage.

Backup target(s): A single chassis, fully configured (10×4) FlashBlades (//S500, //S200) are used to create SMB share & S3 object as backup targets with 8 FlashBlade data VIF's (to get network concurrency while using SMB as backup target).



Server configuration: Each database host has 32 CPUs and 128 GB of memory, running SQL Server 2022 Enterprise Edition on Windows Server 2022. For an optimal CPU configuration on the SQL Server host, it is advised to have a number of cores greater than the count of both data files and backup files.

Database configuration: Configuration of 1 data/log file per volume/disk enables SQL Server to spawn a reader thread per volume/disk while conducting backup or restore operations, and leads to enhanced performance.

Considering SQL Server parallelism, 9 volumes are created each of 2 TB of size with 64KB NTFS partitions to place 8 data and 1 log files ensuring 1 file per volume. The exact number of volumes and data files can be customized based on the customer's environment.

In implementing this solution, the configuration entails setting up 9 volumes, 8 datafiles, 1 log file and 8 backup files. Single database with 8 primary data files, 1 log file

- Database memory setting to 100 GB
- Database size of 1 TB
- Optimized backup parameters:

The constraints associated with SQL Server backups to object storage, such as the maximum file size and the maximum number of objects in a bucket, are linked to the MAXTRANSFERSIZE parameter. This can be customized based on the customer 's environment.

- MAXTRANSFERSIZE = 20M (For S3) & 4M (For SMB)
- BUFFERCOUNT = 512

Backup configuration: Backing up to a single file means using only one writer thread used to write to the backup file, which can act as a bottleneck and impede backup performance. To optimize backup speed, split the backup to multiple files on target, enabling SQL Server to allocate multiple writer threads to run in parallel. This approach enhances backup performance significantly.

The backup and restore operations were carried out using a Transact-SQL (T-SQL) script, directing the backups to SMB share through eight network paths to achieve network concurrency.

SQL Server customers often manage numerous databases, and in the case of enterprise customers, simultaneous backups and restores of multiple databases are a common requirement. The purpose of validations with single and multi-servers is to showcase FlashBlade's ability to handle multiple streams concurrently.



Validation & Analysis

The performance capabilities of FlashBlade were thoroughly assessed revealing FlashBlade capacity to meet the demands of enterprise SQL Server deployments.

Below validations are performed to assess the speed of Microsoft SQL Server backup (with and without compression) and restore process when using FlashBlade as backup target (SMB & S3).

Single Server: Single server with 32 cores hosted with 1 database in full recovery model with 8 data files and 1 log file distributed across 9 disks to ensure ample concurrency during database backup and restore operations. The SQL Server launches one backup reader thread for each logical volume hosting data files for the source database.

Multi-Servers: To Exhibit FlashBlade's concurrency, efficiency and performance, A total of 8 Servers were deployed in this solution each with identical configuration that's used in single server validations.

The number and configuration of servers may differ depending on the customer's environment, yet FlashBlade consistently demonstrates superior performance.

The charts below provide a backup and restore throughput comparison obtained from validations performed on database(s) using SMB and S3 as backup targets on FlashBlades //E, //S200, //S500.

Limitations

The performance of SQL Server backup and restore operations can vary depending on below factors in the environment.

- Host specific bottlenecks such as CPU and memory overhead.
- Storage bottlenecks such as array configuration, resource limits can degrade backup and restore performance.
- Network bottlenecks such as number of hops, bandwidth can affect overall performance.
- SQL Server features such as compression, encryption require balanced compute settings.



Single Server Validation

Single database validations demonstrate that the FlashBlade has performance capabilities to meet the demands of enterprise SQL Server database deployments. The most optimal performance can be easily obtained with minimal tuning for SQL Server databases.

Figure 3 & 4 illustrates the single database validation results across 3 different FlashBlade models using SMB as Backup target.

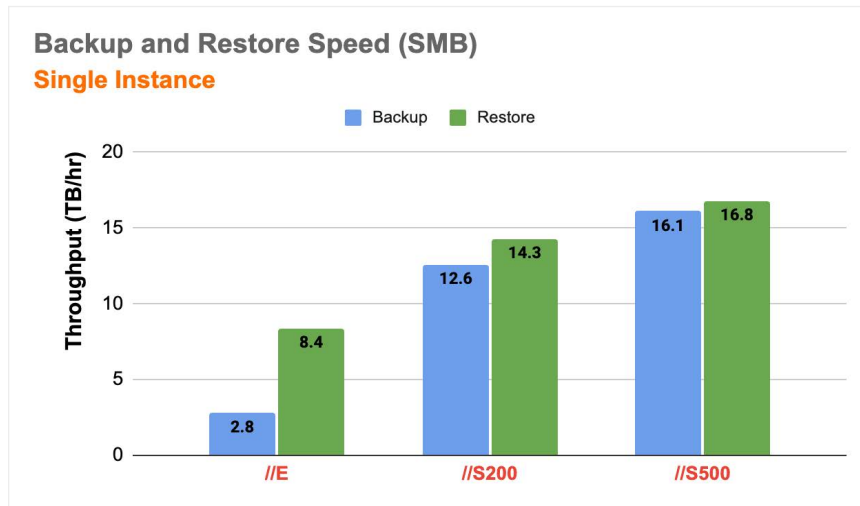


FIGURE 3 Summary of single database evaluation using SMB

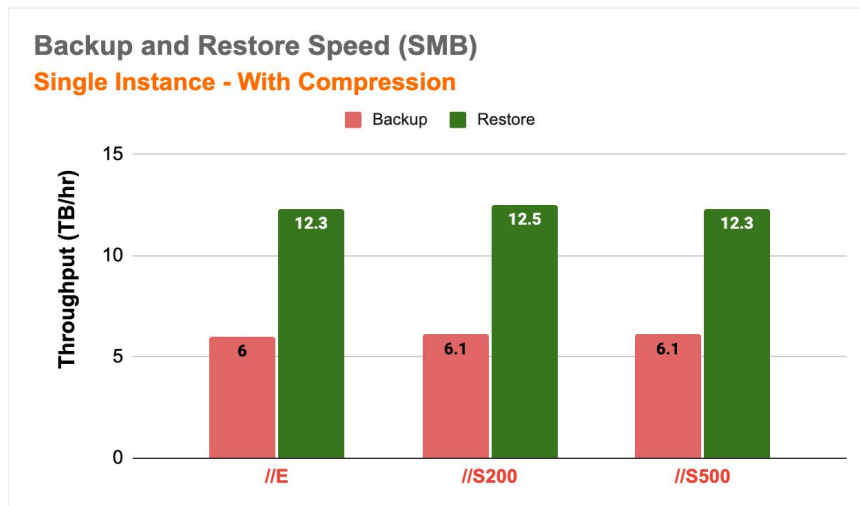


FIGURE 4 Summary of single database evaluation using SMB (With Compression)

SQL Server backup compression is a widely adopted practice to minimize storage needs or decrease network load, or both. Although backup compression helps save storage space, it does incur host CPU cycle costs. Thus, finding the right balance is a trade-off for customers, and there isn't a one-size-fits-all recommendation.

FlashBlade addresses this challenge by featuring built-in native compression, ensuring data is consistently stored in an optimized compressed format to enhance Flash Storage efficiency.

This validation helps to understand how FlashBlade can easily handle backup and restore operations for large database instances ensuring consistent performance and relieving CPU workload on the host.



In figure 4, Backup and restore throughput rates, especially with SQL Server compression (host level) exhibit consistent performance across various FlashBlade models when using SMB as the backup target, in contrast to conventional storage solutions.

The speed of SQL Server backup and restore operations differs between SMB and S3 backup targets. S3, being a higher-level protocol compared to SMB, demands more effort for data transfer between the source and target. FlashBlade excels in high-performance scenarios, especially when handling concurrency with multiple instances.

Figure 5 & 6 illustrates the single database validations result across 3 different FlashBlade models using S3 as Backup target.

Backup throughput is varying from 5.7 TB/hr to 10 TB/hr based on the FlashBlade model for single database validations. While restore throughput is consistent across FlashBlade models.



FIGURE 5 Summary of single database evaluation using S3



FIGURE 6 Summary of single database evaluation using S3 (With Compression)

Single server validation helps to understand baseline performance of SQL server backup and restore on FlashBlade SMB and S3 contrasting it with conventional storage solutions.

Multi-Server Validation

During the multi-server validation phase, simultaneous backup and restore operations were conducted across 8 servers, each with a database size of 1TB. All the tuning techniques acquired from the single database validation were applied during this process.

The multi-server validations clearly showcases FlashBlade's ability to provide high performance, scalability, and handle significantly higher workloads simultaneously with optimized backup and restore times.

The charts below provide a comparison between the results obtained from validations performed on multi-servers using SMB and S3 as backup targets.

Figure 7 & 8 illustrates the multi server validations result across 3 different FlashBlade models using SMB as Backup target.

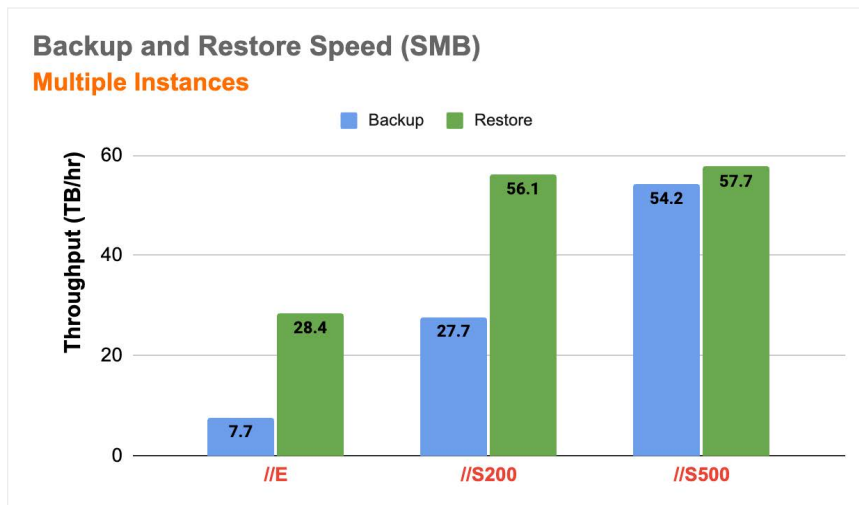


FIGURE 7 Summary of Multiple server (8) evaluation using SMB

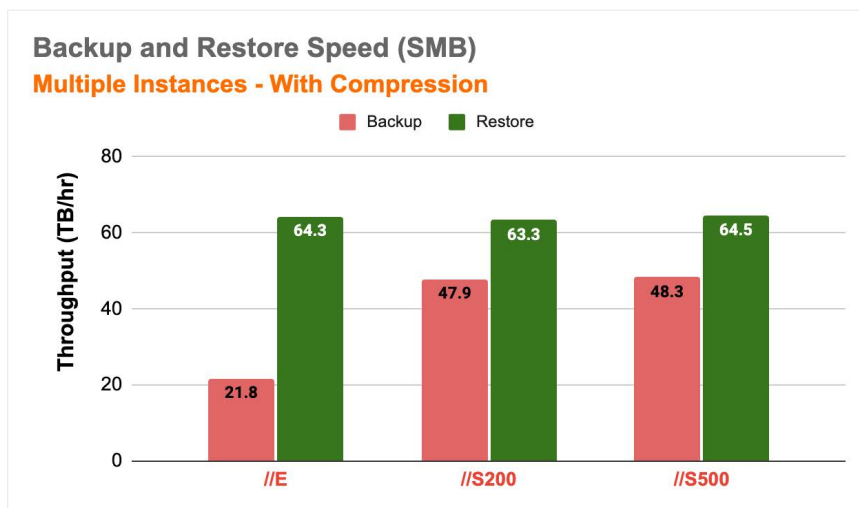


FIGURE 8 Summary of multiple server (8) evaluation using SMB (With Compression)



Using SQL Server compression, Restore throughputs are high when host-level compression is enabled and consistent across all FlashBlade models.

Figure 9 & 10 illustrates the multi server validations result across 3 different FlashBlade models using S3 as Backup target.

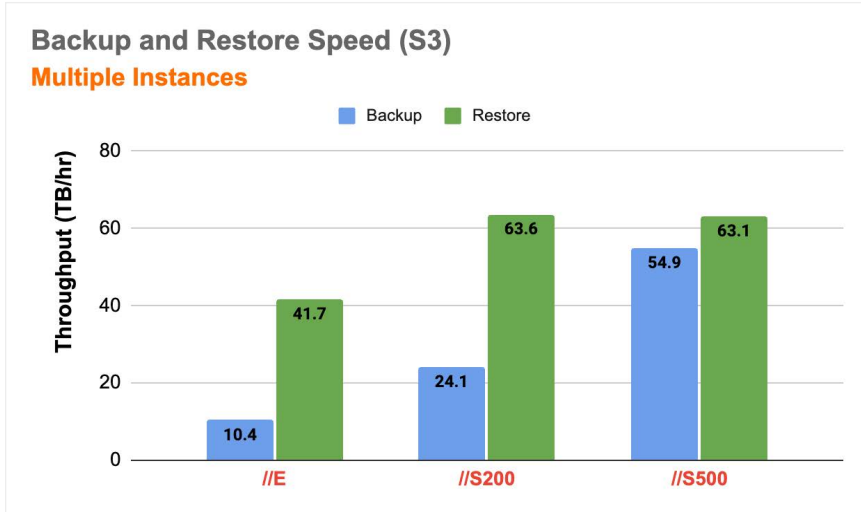


FIGURE 9 Summary of multiple server evaluation using S3

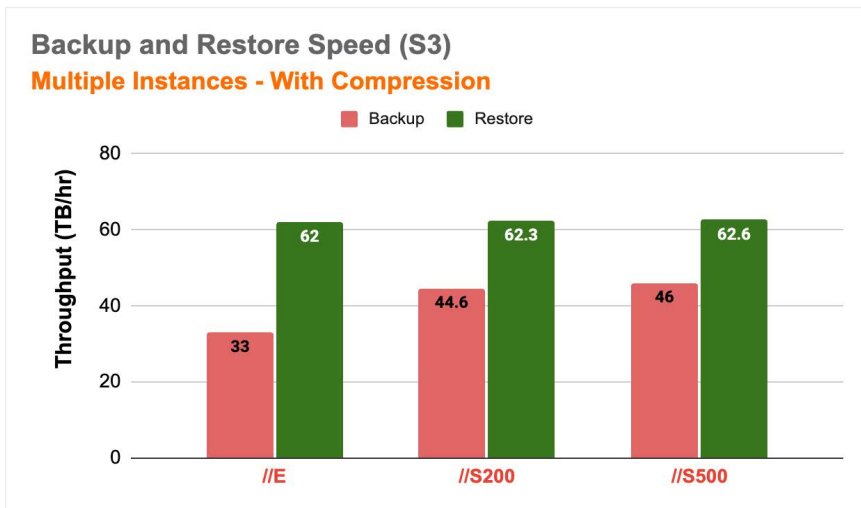


FIGURE 10 Summary of multiple server evaluation using S3 (With Compression)



Conclusion

Different FlashBlade models offer various benefits to enterprises that use SQL Server Databases for Business/Mission critical applications to protect data that ensures business continuity to meet SLA's and RPO's.

FlashBlade//S500 is a high-performance all-flash storage system designed for demanding enterprise workloads. It offers exceptional performance, scalability, reduced storage costs and agility, making it an ideal and optimal choice for organizations that need to store and access large amounts of data quickly and efficiently, especially those with critical applications necessitating high availability.

FlashBlade//S200 proves that powerful storage doesn't have to come in a large package. Its compact design, impressive performance, and flexible scalability make it a strategic choice for organizations seeking to optimize their data infrastructure for mid-range workloads. It is best fit for organizations seeking scalable, efficient backup, quicker recovery performance, and optimized storage space with an advanced compression mechanism.

FlashBlade//E breaks down the barriers to high-performance storage, making it an ideal entry point for organizations facing increasing demands for unstructured data within constrained budgets, making it an environmentally friendly option with cost-efficient performance.

Based on SQL Server Backup and restore validations using FlashBlade as backup target offers multiple benefits.

- Enhanced backup and recovery performance across different FlashBlade models and maintains consistent results even with increasing the number of instances.
- Rapid Restore on FlashBlade delivers modern data protection that's fast, reliable, and effortless.
- Organizations can utilize storage compression to efficiently manage space on backup targets. This approach also reduces CPU and memory load on SQL Server Instance.
- Super-fast native SMB and SQL Server's native tools let enterprises consolidate not just backups, but file services, analytics, and AI, all on a single, powerful platform.
- Upgrade your backup game with FlashBlade object store (S3). It's the high-performance, resilient, and cost-effective way to keep your SQL Server data safe and accessible.

Additional Resources

- Learn about [Best Practices for Microsoft SQL Server on FlashArray](#).
- Read more about SQL Server backup to URL for [S3-compatible object storage](#).

purestorage.com

800.379.PURE

