PURESTORAGE®

TECHNICAL WHITE PAPER

Microsoft SQL Server Backup and Restore with FlashArray//C

Optimize capacity and performance storage for SQL Server backups.

Contents

Introduction

In this document, we discuss the test results that demonstrate Pure Storage® FlashArray™ is an excellent SQL Server backup target for file and block backup, and is a reliable part of a robust backup strategy. As well, it highlights tuning parameters that can be used to optimize backup performance of a Pure FlashArray//C storage system when used as a direct network attached storage (NAS) target for SQL Server backups.

FlashArray//C is the first, all-quad-level cell (QLC) flash array with capacity-optimized NAND flash memory technology that delivers a per-terabyte cost that matches or beats hard-disk drives (HDDs). FlashArray//C leverages QLC for optimal NVMe performance, hyper-consolidation, and simplified management of your data for capacity-oriented workloads.

Solution Overview

Organizations that use SQL Server databases for their business-critical applications must be able to protect and recover their data or run the risk of business interruption. SQL Server comes with a built-in Transact-SQL Backup command tool, as well as an agent for scheduling and automating backups. Both allow database administrators (DBAs) to easily back up and restore data, ensuring data integrity and retaining sufficient data to restore the database to any supported point in time. A target storage device is used to backup and recover files. We tested three SQL Server instances for backup and restore using block and file services on Pure Storage FlashArray//C hardware and Purity data storage software as the target.

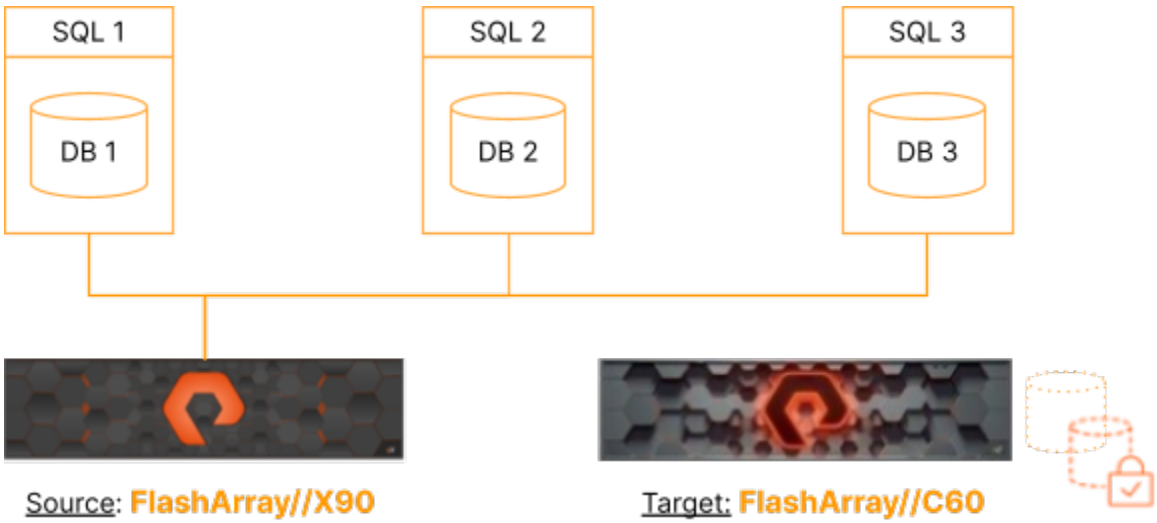

Figure 1: Solution overview

Pure Storage FlashArray//C

Pure Storage FlashArray systems are all solid-state systems, providing DirectFlash®, NVMe-oF, FC, and iSCSI access, as well as file access with SMB and NFS. The FlashArray family has three members: FlashArray//X for performance-oriented workloads, FlashArray//XL for large applications on fewer arrays, and FlashArray//C for capacity-oriented workloads. All data is both deduplicated and compressed by FlashArray prior to being written. There is no performance penalty for deduplicating and compressing data in-line, which enables an effective capacity of 3.3PB for //X, 5.5PB for //XL, and up to 7.3PB for //C.

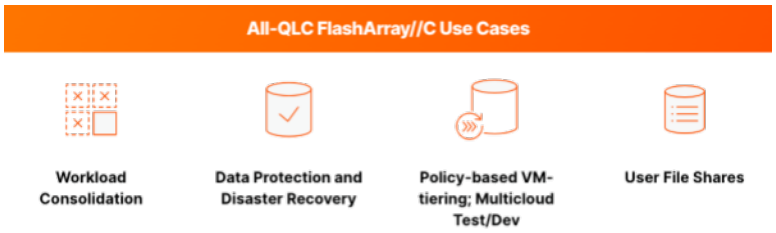

Figure 2: FlashArray//C Highlights

FlashArray//C was introduced to the Pure FlashArray family to provide a capacity-optimized platform based on QLC while retaining all of Purity's ease-of-use, robust data management, and data protection features.

FlashArray//C takes advantage of the success of FlashArray//X by providing customers with high-capacity arrays, consistent performance and lower latency when compared to traditional disk or even hybrid storage. FlashArray//C has a higher latency profile because Purity updates took advantage of native QLC NAND while maintaining high reliability. Generally, QLC NAND requires more care when writing or deleting data because of its complex design, low cell durability, and increased energy requirements. As a result, FlashArray//C's performance can be measured by latency as low as 1 millisecond (Purity 6.1.9 and higher).

The read latency on FlashArray//C gradually increases from 1 to 3.5 milliseconds as workloads increase from low thread counts and low throughput to high thread counts and high throughput, while the write latency remains consistent. Because of the increase in read latency, it is crucial to have multiple parallel read I/O stream workload applications—such as backups—to ensure maximum performance.

The improvements in FlashArray//C's crawling speed and efficiency have lowered the gap between FlashArray//X and FlashArray//C in terms of overall data reduction rates. With this deduplication improvement, FlashArray//C has data reduction rates that are about 80–85% of FlashArray//X.

FlashArray//C is currently the only Pure Storage FlashArray to offer file services. FlashArray file services provide native, multi-protocol file storage and access to FlashArray//C for a unified block and file platform. You now have the option of selecting file or block as a backup destination for SQL backups, depending on your infrastructure capability. FlashArray//C file services provide an option to create a central NAS repository that can serve as a backup destination for business critical databases, providing easy access, and simplicity without sacrificing features, or performance. Your file-based data and block-based data will benefit from one unified data layer, a common pool of capacity, and global data reduction. NAS file storage services using FlashArray//C provide customers with the same reliability, data reduction, and simplicity as FlashArray//X block storage.

Realizing consistent all flash performance at 99.9999% availability with non-disruptive upgrades for the most demanding 24x7 operations and lower costs than hybrid storage, FlashArray//C enables IT organizations to simplify data protection operations by eliminating complex siloed approaches to backups across multiple disparate hybrid disk-based solutions.

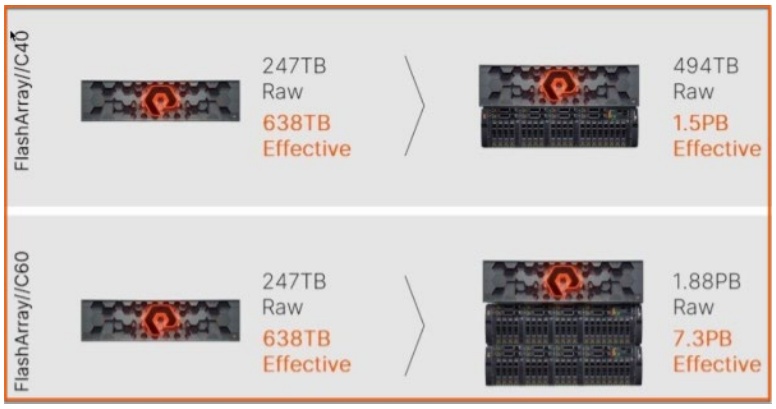

Figure 3: FlashArray//C Scalability

Scalability, another aspect of FlashArray//C (shown above in Figure 3), delivers consistent performance via its end-to-end all NVMe architecture and density configurations that span 1.5PB in a 3U to more than 7.3PB effective capacity in 9U. Simple hardware, software, and cloud management tools are designed to make everything work together seamlessly. This inherent simplicity is characterized by:

- One box with only six cables and a 30-minute installation with no manual (with available Pure Professional Services or partner installation)
- Inline data-reduction and end-to-end encryption
- Seamless expansion with a flexible buying program (Evergreen Storage™ subscription) for non-disruptive rapid upgrades and expansion
- No performance tuning required
- APIs for automation
- AI-driven cloud management
- All array software included

Purity

The Purity operating environment powers FlashArray//C to deliver comprehensive data services. Purity data storage software provides the confidence of knowing that your data is stored in a way that ensures data efficiency, security, and availability due to its key components:

- **Data reduction**: Purity's granular and complete data reduction provides unmatched storage efficiency. Considering granular and complete data reduction, Purity's storage efficiency is unparalleled. Repetitive binary patterns are removed before data is deduplicated and compressed, streamlining the data reduction process.
- **Always-on encryption**: Purity's "encrypt everything" approach provides built-in enterprise-grade data security without user intervention or key management. Pure Storage EncryptReduce™ encrypts data in-flight without affecting data compression and deduplication. It also provides physical security through Key Management Interoperability Protocol (KMIP).
- **High availability**: Purity ensures business continuity by reducing risk of downtime while keeping mission-critical applications and data online and accessible. Designed from the ground up for flash, Purity RAID-HA protects against

concurrent dual-drive failures, initiates rebuilds automatically within minutes, and detects and heals bit-errors. Purity also treats performance variability as a failure and uses parity to work around bottlenecks to deliver consistent latency.

- **Non-disruptive everything**: When hundreds of applications are hosted on the storage array, downtime is not an option. Purity allows you to expand flash capacity, upgrade controllers, and replace failed components—all without taking your storage offline or negatively impacting application performance. That makes Purity truly nondisruptive.
- **Intelligent quality of service (QoS)**: Purity continuously optimizes infrastructure using always-on QoS to prevent workload disruptions and resource hogging. Array utilization is maximized without artificial limits on workloads, and full performance is delivered for all workloads.
- **Snapshots**: Purity snapshots are simple, space efficient, with no performance impact. The snapshot technology encapsulates metadata with data, making the snapshot portable. This means you can offload it from a Pure FlashArray to the cloud in a format that is recoverable by any FlashArray. Purity also allows free movement copies of the snapshots between FlashArray, to FlashBlade®, third party NFS servers. Data portability is central to a unified data management strategy. Snapshots offer infinitely configurable policies and a near-infinite means of offloading snapshots for replication and mobility.
- **Ransomware protection**: SafeMode™ Snapshots can be enabled to provide ransomware protection. Pure snapshots have always been immutable, high-performant, and highly storage-efficient. SafeMode adds additional protection to snapshots to prevent hackers from compromising admin credentials to a storage array to irrevocably delete snapshots. Snapshots are key to a data protection strategy with the ability to recover from ransomware attacks.

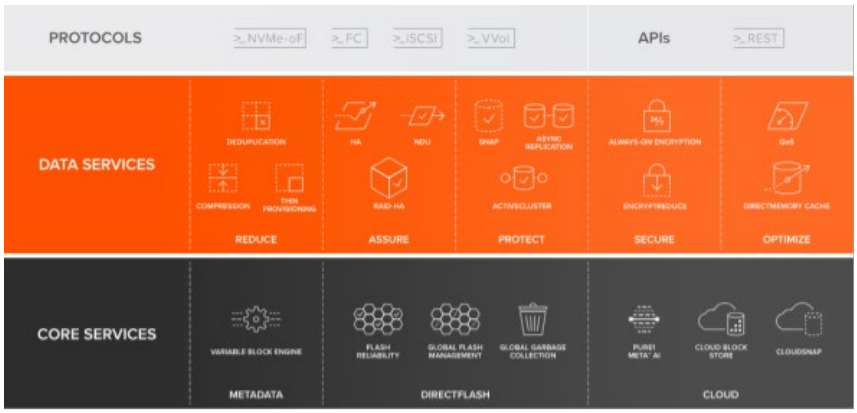

Figure 4: Purity//FlashArray Features

A FlashArray//C powered by Purity storage software is enterprise-ready and provides up to 99.9999% availability, ransomware protections, non-disruptive upgrades, and consistent performance for the most demanding environments.

Solution Benefits

Microsoft SQL Server backup and restore on FlashArray//C solution has been verified by Pure to operate efficiently, and effectively, the benefits include:

- Consistently high performance backup and restore. Purity always-on QoS continuously tunes FlashArray//C for optimum performance.
- Cost optimization, and scalability with all-flash performance at the economics of hybrid storage.

- Deployment flexibility and simplicity with a single solution, incorporating block and file.
- High availability with protection against logical corruption, and ransomware. Purity RAID-HD protects against concurrent dual-drive failures, detects, and heals bit-errors. SafeMode snapshots offerprotection against ransomware with quick, and fast recovery on block & file.

Test Environment

Server Details

All servers had identical hardware configurations, as shown in Table 1, including Microsoft Windows Server 2022, Build 20348.169.210806, as the operating system.

Component	Description
Processor	2 x Intel Xeon E5-2697 v2 @ 2.70 GHz (24 cores total) HyperThreading enabled
Memory	512 GB RAM
Ethernet (iSCSI)	2 x Mellanox MT27500 family network adapter @ 40 Gbps, in LACP team
Host Bus Adaptor (HBA)	Qlogic 2692 Dual Port 16Gb Fibre Channel HBA PCIe

Table 1 : Server Details

SQL Server Configuration

All servers had identical hardware configurations, as shown in Table (2).

Component	Description
SQL Server Version	15.0.4188.29(CU15)
SSMS	18.10 (Build: 15.0.18390.0)
SQL Server Memory	Restricted : Min: 0, Max: 16384MB (16GB)
Max Degree of Parallelism	MAXDOP = 0
Max Transfer Size	MAXTRANSFERSIZE: Default*
Number of Data Disks	8 Disks (NTFS, Block Size:64K)
Number of Data Files	8 Data Files

Number of Log Disks	1 Disk (NTFS, Block Size 64K)
Number of Log File	1 File
Number of Backup Disks	8 Backup Disks

*Automatically set by SQL Server

Table 2: SQL Server configuration

FlashArray//C Configuration

Table 3 below illustrates the details of the FlashArray//C used in our testing.

Component	Description
Purity Version	6.1.11
Capacity	320.09 TB (351.94 TB usable)
Controllers	2 Controllers (Active/Passive)
Ethernet (iSCSI, File)	8 (25 Gb) Active Ports
Number of Backup Disks	24 Volumes
Number of SMB Shares	1 SMB Share
Number of Virtual Interfaces	4 Virtual interfaces

Table 3: FlashArray//C configuration

Connectivity

As with any high-performance infrastructure, multiple network connections should be configured between the SQL server and the FlashArrays to maximize throughput and provide redundancy. The same recommendation applies to the backup workload profile, it is also recommended that these network interfaces be dedicated, in the case of iSCSI, a dedicated VLAN can provide backup traffic isolation. Figure 5 provides details on how our test environment was connected. The frontend SQL Server data was connected to the fibre channel on FlashArray//X90. For a backup workload we opted to use dedicated vLAN for iSCSI, and files services.

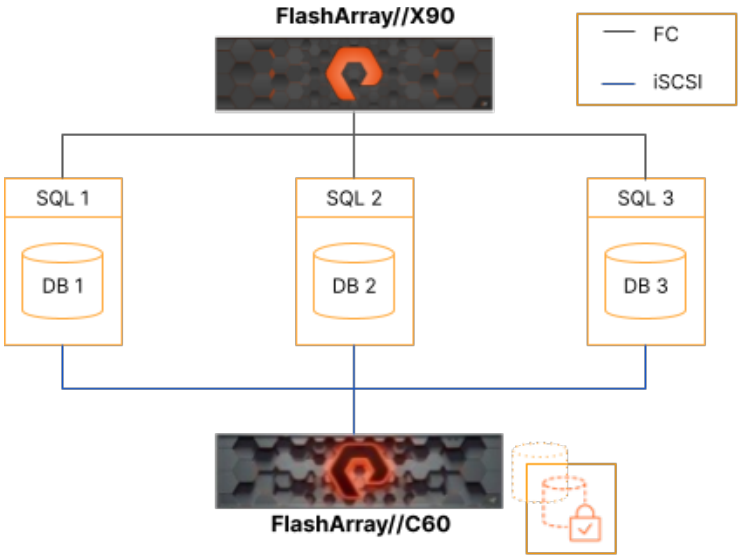

Figure 5: Test Environment connectivity

Test Data

An internal tool was used to generate 1TB test data for each SQL Server in the test environment. Table 4 shows the details.

Component	Description
Test Data Source	Wikipedia
Randomness	60% random data
Size	1TB on each SQL Server Host

Table 4: Test data

Backup/Restore to FlashArray//C Using Block Storage

Tests were performed to stress the FlashArray//C for best performance possible, within the context of the available hardware.

Testing was conducted using only full database backups. Incremental backups were not examined. During our analysis, we examined parallelism impact backup to a single versus multiple backup file, as well as single and multiple SAN disks to identify how FlashArray//C handled the increased backup workload. Additionally, we evaluated SQL server compression settings and how they impacted the overall solution performance. We used duration, IOPS, latency, and throughput to measure the FlashArray//C performance.

During backup, and restore, SQL Server defaulted to using 256KB block size.

Parallelism

As recommended by Microsoft, we know parallelism drives better performance. However, the purpose of this test was to identify whether having multiple disks from the FlashArray//C for SQL server backups/restores would significantly improve the performance. We found on the FlashArray//C that backups to multiple disks did not significantly improve backup duration, compared to backups to multiple backup files on a single disk. Figure 6 and Table 5 show slightly better IOPS in favor of using multiple backup files on just a single disk. In any case, the difference is negligible.

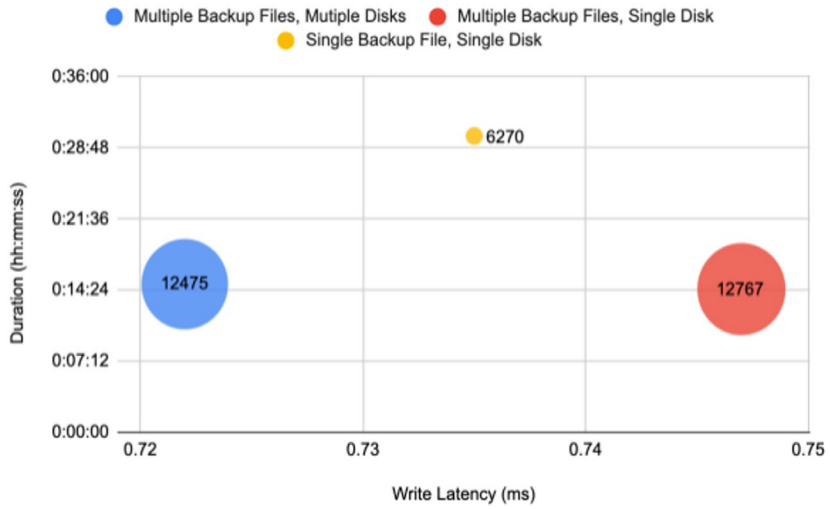

Figure 6: Backup Parallelism (Block)

Backup to	Duration (hh:mm:ss)	Write Latency (ms)	IOPS
Multiple Disks	0:15:00	0.72	12,475
Single Disk	0:30:00	0.74	6,270
Single Disk, Multiple Backup Files	0:14:30	0.75	12,767

Table5 : Backup Parallelism (Block)

Restore performance for the same configuration showed again a negligible improvement in restore duration in favor of using multiple disks. Figure 7 and Table 6 illustrate our findings.

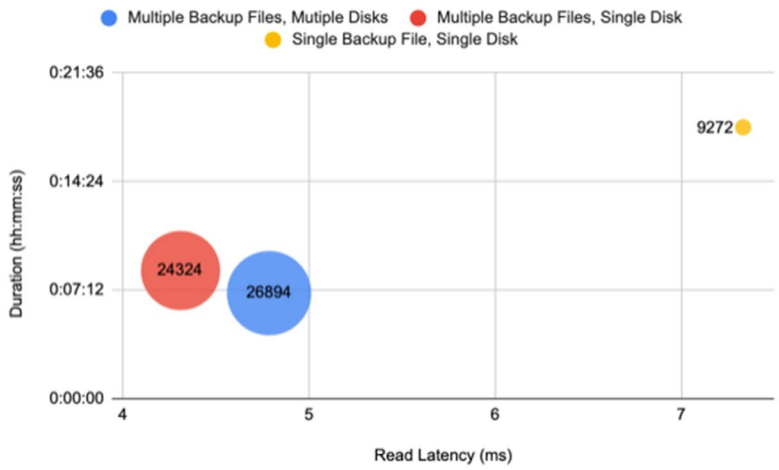

Figure 7: Restore parallelism (Block)

Restore From	Duration (hh:mm:ss)	Read Latency (ms)	IOPS
Multiple Disks	0:07:00	4.8	26,894
Single Disk	0:18:00	7.3	9,272
Single Disk, Multiple Backup Files	0:08:30	4.3	24,324

Table 6: Restore Parallelism (Block)

It is the reader's choice whether to use a single disk for SQL server backups, or several disks. Our recommendation is to use multiple backup files. Testing showed creating 1:1 data files, to backup files, will yield the shortest backup duration.

Compression

The Pure FlashArray//C storage system comes with thin provisioning and always-on data reduction by default; turning off SQL server backup compression can lead to considerable CPU savings on a SQL Server host. Nevertheless, there are environments where it may not be possible to dedicate Network infrastructure for backups, have large databases, and tight backup SLA, SQL server backup compression may be the only option.

FlashArray//C inline, and post processing data reduction can still be beneficial while compression is enabled for SQL server backups, your mileage may vary, depending on the variety of databases being backed up. If you have the same production database being used for test and development, and still need to back them up, you're probably going to see better data reduction. However, if you have a variety of very different SQL Server databases being backed up, data reduction will not be as good. Using our test data, where we are using three identical, one-terabytes databases, FlashArray//C was still able to provide a 1.2:1 data reduction, compared to 3:1 data reduction when SQL server compression is turned off. Most importantly, backup duration was reduced by nearly 40%, with very low latency. Figure 8 and Table 7 provides more details on the findings.

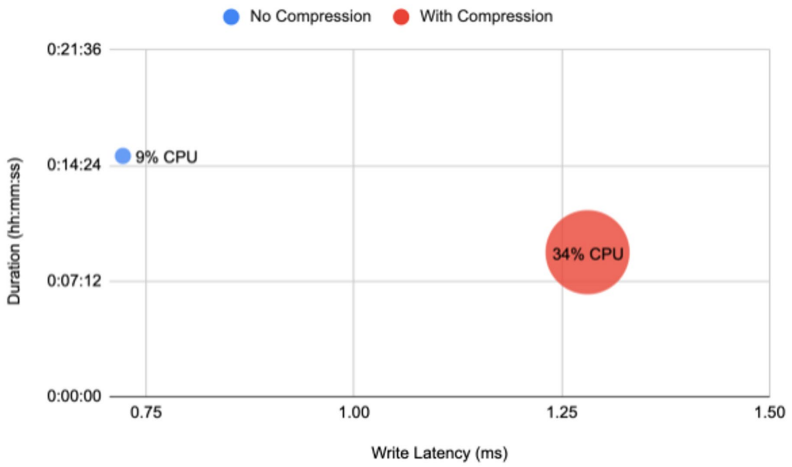

Figure 8: Compression Performance (Block)

Backup	Duration (hh:mm:ss)	Write Latency (ms)	SQL Server CPU	FlashArray//C DRR
No Compression	0:15:00	0.72	9%	3:11*
With Compression	0:09:00	1.3	34%	1.2:1*

* Not Guaranteed Data reduction, based on the dataset used in the test.

Tale 7: Compression performance (Block)

Backup/Restore Throughput

Table 8 breaks down the FlashArray//C impressive processing throughput performance on both backup and restores of our three terabytes database.

Operation	Duration (hh:mm:ss)	Throughput (TB/h)	IOPS
Backup (Compression: OFF)	0:15:00	11	12,767
Restore (Compression: OFF)	0:07:00	23	26,894
Backup (Compression: ON)	0:09:30	6.3	7,402
Restore (Compression: ON)	0:06:30	9	10,481

Table 8: Overall Throughput Performance (Block)

Backup/Restore to FlashArray//C using SMB Shares

FlashArray//C File Services use virtual interfaces (VIFs) for client connectivity. Clients connect using the VIF DNS name, or IP address, which internally to FlashArray//C maps to an SMB shared directory. VIFs should be set up with two physical interfaces

from each controller for redundancy. A minimum of one VIF is required; however, FlashArray//C can have more than one VIF, improving performance, and reducing network hops. As with testing on the block storage side of the FlashArray//C, testing was conducted using only full database backups. During our analysis, we examined parallelism impact backup to a single versus multiple backup file, as well as single and multiple SMB shares to identify how FlashArray//C handled the increased backup workload.

Additionally, we evaluated SQL server compression settings. We have used duration, IOPS, latency, throughput, and number of SMB connections to measure the overall solution performance. During backup, and restore, SQL Server defaulted to using 1024KB block size.

Parallelism

We started out testing a single backup file, on a single SMB share, then branched out to multiple files on a single SMB share, and finally to multiple backup files on multiple SMB shares. Compared to block storage, having multiple backup files on the same share did not have a significant impact. What really impacted performance was using multiple backup files on multiple FlashArray//C VIFs, which proved to be comparable to using multiple backup files on multiple disks on block storage.

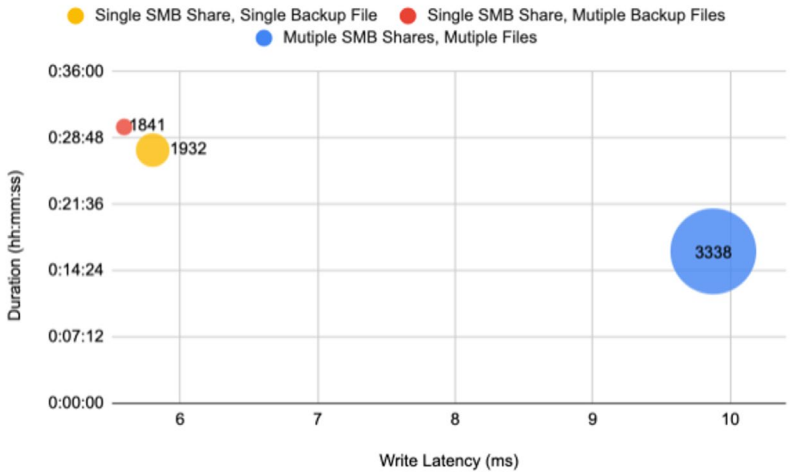

Figure 9: Parallelism Performance (File)

Backup to	Duration (hh:mm:ss)	Write Latency (ms)	IOPS	SMB connections
Single VIF, Single Backup Files	0:27:30	5.8	1,932	8
Single VIF, Multiple Backup Files	0:30:00	5.6	1,841	24
Multiple VIFs, Multiple Backup Files	0:16:30	9.8	3,338	24
Single VIF, Single Backup Files	0:27:30	5.8	1,932	8

Table 9: Parallelism performance (file)

Compression

As we stated earlier both file, and block services on the FlashArray//C benefit from the underlying data reduction architecture and logic.

During testing with multiple FlashArray//C SMB shares as a target for SQL server backup files, we've experienced consistent performance characteristics compared to similar configuration on FlashArray//C block services. From a latency standpoint, FlashArray//C-File performed slightly better than block, though both finished at the same time. Table(10) details our findings.

Backup with Compression: ON	Duration (hh:mm:ss)	Write Latency (ms)	IOPS	Block Size (KB)
Multiple Backup Files, Multiple Block Disks	0:09:00	6.3	7,402	256
Multiple Backup Files, Multiple VIFs	0:09:00	4.0	1932	1024

Table 10: Compression Performance (File)

Backup/Restore Throughput

Table 11 is a breakdown of FlashArray//C-File processing throughput performance of our 3TB test databases.

Operation	Duration (hh:mm:ss)	Throughput (TB/h)	IOPS
SMB Backup (Compression: OFF)	0:16:30	11.4	3,338
Restore (Compression: OFF)	0:13:30	11.8	3,432
Backup (Compression: ON)	0:09:30	6.6	1932
Restore (Compression: ON)	0:06:30	6.9	2026

Table 11: Overall throughput performance (file)

Recovery with FlashArray//C SafeMode Snapshot

We've performed hardware-based snapshots for both FlashArray//C-File, and FlashArray//C-Block backup files volumes, and the SMB share, respectively. We followed Pure storage best online documentation and did not experience any issues, allowing us to recover our database.

Additional Resources

- Discover Pure Storage Data Protection Solutions.
- Learn about FlashArray//C.
- Lean about Pure Storage solutions for Microsoft.

Pure Storage, Inc.
650 Castro Street, #400
Mountain View, CA 94041