

WHITEPAPER

BACKUP OPTIMIZATION



Business Challenge

Backing up data is a process that no one is aware of until they need a restore. Hopefully, you rarely do. The reality of the backup environment is that it costs your organization money. The more data you have, the more you have to backup and the costs go up accordingly.

Costs are accrued in several places.

- Licensed Backup software
- Man-hours spent to manage the backup
- Media costs
- Media storage costs

There are also hidden costs in your infrastructure and it's management that go unaccounted for many times.

- Disk space and customer's time to get through their data
 - Ex: I have 500,000 files to go through and the shared drive takes too long to find my data
- Risk exposure due to redundant or outdated files with sensitive information.
 - EX: Files from old employees.
 - If sensitive data is exposed, there is hefty cost in terms of a lawsuit and regulatory violation.

This whitepaper takes you through storage optimization and backup efficiency with StorageX.

Deep Dive on the Problem

Let me start off by saying that ALL of us

1. Create a file
2. Save it to our home drive
3. Email it
4. Lastly, forget about it.

Right off the bat, it seems harmless, and it's not a big deal. However, the cost can get very large and very quickly. What if I told you after 5 years, you could have over 20,000 copies of that file stored in your backup infrastructure? Many files stay on their file system for 10 years or longer. At 10 years, there could be 40,000+ copies. It's a staggering number. How does this happen?

Having been an Systems Admin and Architect, I've designed disaster recovery, backup, and archive systems. Here is the scenario I used to figure out how many copies could be out there.

- There is dedupe but for the purposes of this scenario, each file stored is unique on the file system.
- Your file systems are backed up in full once a month.
- There is a D/R site that is also backed up.
- No file purges are done.
- Backups stored for 6 years & 10 months (total of 82 months).
- Email is managed to eliminate old data automatically.

Jan 1, 2000

At 8am the CEO issues a Welcome to the 2000's memo via email. It contains 12 people's names, 4 departments, and pictures of 6 events with different teams.

At 9:15 am

12 people whose names are mentioned, 23 people from 4 departments, and 35 people in the pictures save the file to their home drives. The marketing department also makes a copy. All copies are moved to the D/R site. That's 144 copies to start.

Jan 2, 2000

Backups pick up the new files on primary and d/r.

Feb 1, 2000

Monthly backups make another copy on primary and d/r. We're up to 720 copies now.

Jan 1, 2001

We're now at ~ 3600 copies.

Jan 1, 2005

Presently, storing 17,000+ copies.

Jan 1, 2006

The CEO resends the original file to show 'how far they've come'. This time 75 people save the file. Now we're at 21,000+ copies

Oct 1, 2006

Backup Purge begins. Copies now at 26,000+

Jan 1, 2009

Ten year mark. Copies now at 34,000+.

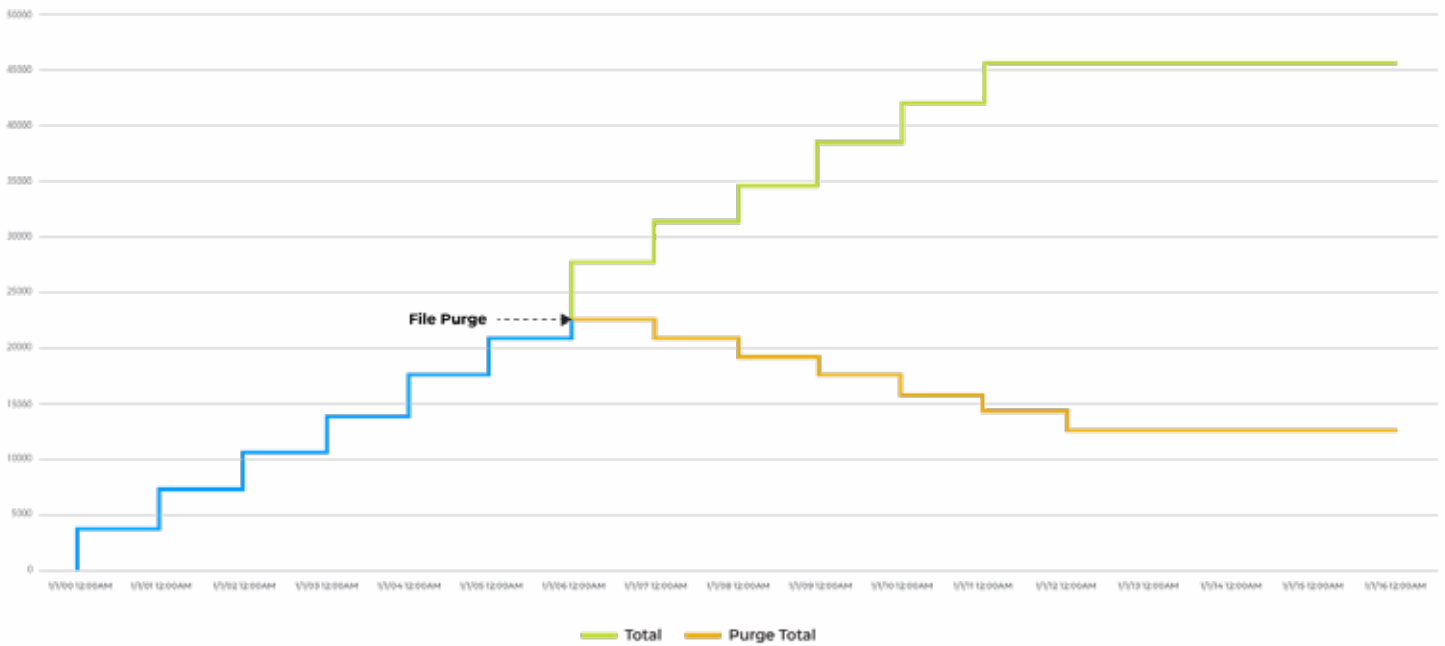
Jan 1, 2012

Because of the purge, the copies now level off at 45,660.

These numbers are staggering, and they represent only one of many backup possibilities. We need to think about is the cost of backup infrastructure and the cost of operations. Remember, this was one file, created by the CEO. The people who saved it? Are they still there? Is there a policy to remove them after they leave? I ask this because the average person now stays with a company for 3 years. What happens when you have billions of files?

How do you manage this to be a more realistic number? Simply put, archive the data to object storage and delete it. With the same infrastructure and a purge on Jan-1-06, the number of copies peaks to 12,444. This is a greatly reduced percentage of number of files.

10 Year File Count



Data Dynamics' StorageX is designed to tackle this situation. More and more companies are realizing that they're wasting space and time. Many companies are investing in object storage that is typically, 10% the cost of standard file servers. Data can be replicated and policies are setup to purge the data totally at the configured time and date.

The problem is how and what should be moved onto object storage. StorageX gives you the visibility into your file storage systems and files that can highlight the datasets which should be moved to the right tier of storage. The key is being able to use multiple points of data. Examples include:

- Identify the file's modified time and locate files that were owned by people who have left the company
 - StorageX can analyze your file systems, find the data, then allow you to use that analysis to archive the data. The ROI on this kind of activity is typically measured in months from the storage terms.
- Identify files with a certain extension that was created more than 5 years ago. These could be archived to a bucket that automatically deletes the data after 1 year.
 - Many people never .log files is an example of this use case.
 - Database dumps, that admins perform 'out of process'.
- Files that were created more than 7 years ago, not modified for 5, and not accessed for 3 years.
 - This would cover my example above.

Benefits StorageX for your Backup Environment

Benefit	Description
Analytical Process	Analytics takes the guesswork out and utilizes proven, efficient methods to identify the right data and move it to the right location within your environment - either local or cloud storage.
Scalability	A software solution that utilizes a scale out architecture that allows you to spin up or down depending on the current migration requirements of the factory. Our UDE design allows the factory to easily move data movement and analytics services to where they are needed, when they are needed.
Performance	The solution can scale up based on the performance required to meet SLAs to support an active hybrid cloud environment.
Use Case Coverage	Broad use case coverage makes StorageX an integral part of your overall data management and budgetary goals
Integration	StorageX provides a robust set of APIs to drive integration with key orchestration and reporting tools that are a must for enterprise deployments. These integrations deliver a true factory functionality by automating many of the tasks involved to drive efficiency and reduce human errors.
Data Mobility	Moving your data that's been backed up to a place that is no longer used is key to saving space, time, and budget across the board.
Licensing	StorageX subscription based capacity license is designed to support your environment, providing a true cost basis that does not limit your ability to scale up or down

Summary

The benefits of moving data that's already backed up to an object storage are from budget constraints and efficiency, to workload optimization. Your backup environment is a target component of the overall IT strategy and the corporate data governance goals. Through automation, scheduling, and usage of 100's of StorageX APIs, you can implement a strategy built specifically for your requirements.

Using StorageX, you can implement a large scale, coherent methodology to continually evaluate, optimize, and reorganize your unstructured data. While this benefits your backup infrastructure, it helps your entire company strategy to reduce expenditure, mitigate risk, and improve data governance.

Data Dynamics, a global leader in enterprise data management, stands at the forefront of the industry-wide shift towards Digital Trust & Data Democracy. Trusted by 300+ organizations, including 25% of the Fortune 20, the company is recognized for its commitment to creating a transparent, unified, and empowered data ecosystem. Whether addressing data risk, privacy, sovereignty, optimization, sustainability, or facilitating seamless, policy-driven data migration across hybrid and multi-cloud environments, the company is ushering in a new era where data ownership, control, & actionability reside with the data owners.



Contact Sales

Book a Demo