5 Ways to Stay Cyber Resilient Amidst the Unstructured Data Explosion

Why you can't get by with legacy solutions and what must-have capabilities you need to prioritize now

rubrik

TABLE OF CONTENTS

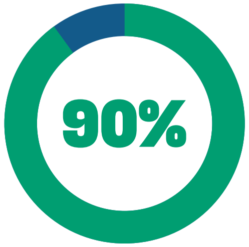
THE UNSTRUCTURED DATA EXPLOSION IS HERE

The volume of unstructured data—or data that exists outside of conventional data tools, such as relational databases—is exploding. According to IDC, unstructured data accounts for 90% of all data. But here's the kicker: A quarter of organizations say that data is growing faster than they can keep up with.[1] That's worrisome because unstructured data includes tons of sensitive information like intellectual property (IP), personal data, research findings, and more.

So, if your most valuable intellectual property—including research data, product designs, source code, engineering models, and scientific discoveries (all of which are unstructured data)—gets attacked, you could lose years of work, your competitive advantage, and irreplaceable business knowledge.

You need a way to protect all that unstructured data, but that's a huge challenge. Not only is the volume of unstructured data huge, but it's also extremely fragmented and scattered across different storage systems and locations (on-premises network attached storage, object stores, etc.), which makes manually keeping track of it difficult.

Traditionally, organizations have used legacy backup approaches to try to protect their unstructured data. But those methods just don't give you the speed, visibility, or control you need to do the job.



According to IDC, unstructured data accounts for 90% of all data.

[1] IDC: The Untapped Value of Unstructured Data, https://www.box.com/resources/unstructured-data-paper

For instance:

Legacy backup solutions weren't designed to back up heavy datasets. Backing up petabyte-scale datasets with legacy backup solutions can take weeks. But even if the data is backed up relatively quickly, the process can create lag and performance issues in the production environment. Both of these issues can lead to teams leaving their data unprotected.

Traditional backup methods also lack cross-platform visibility, which means you don't have a unified view across different storage types. This fragmented approach makes it difficult to manage protection across your vast landscape of unstructured data.

They don't give you insights into where your sensitive data lives and who has access to it, which makes it difficult to know where your risk lies and can prevent you from meeting regulatory compliance obligations.

They lack adequate security controls to defend against rising threats like ransomware.

And legacy backup methods make recovering your files and folders a slow, manual process that can keep you from meeting your SLAs.

The writing is on the wall: As unstructured data proliferates, sticking with legacy backup solutions is a recipe for disaster. Let's examine five ways a modern unstructured data protection solution can reduce your risk, make your data resilient, and keep your business running in the face of an attack.

FIVE MUST-HAVES

IN AN UNSTRUCTURED DATA PROTECTION SOLUTION

1 Performance at Petabyte Scale for Unstructured Data Protection

Legacy backup systems start to falter when unstructured data reaches the petabyte scale. Here's why: These outdated tools simply weren't built to handle the massive volume and throughput requirements needed to protect billions of files.

Faced with this amount of data, legacy backup systems start to buckle. Performance starts to lag—both within the backup system itself and the systems it's backing up. Your recovery point objectives (RPOs) will begin to slip along with your backup window, possibly leaving critical information at risk. But throwing more hardware at the problem is a costly exercise in futility. Network data management protocol (NDMP) and other legacy approaches lack the application awareness and scalability needed to manage such vast data volumes efficiently. So, even with additional hardware, you'll still end up dealing with the same issues.

So, what's the solution? You need a modern approach that can efficiently protect billions of files with highly parallelized scanning, indexing, and data movement. By maximizing throughput and dynamically throttling for optimal network utilization, you can shrink backup windows and comprehensively protect your data without breaking the bank.

PRO TIP

Look for unstructured data protection solutions that provide parallel streaming and dynamic quality of service (QoS) for high-performance backup. These tools can efficiently protect petabytes of data while eliminating the impact on production workloads.

2 Holistic Visibility Across All Your Unstructured Data

We talked about how unstructured data is scattered across multiple NAS platforms and locations, making it difficult to know exactly where everything is, who has access to it, and if you're adequately protecting that data.

Legacy backup solutions don't make seeing and protecting all that data any easier. These solutions often have siloed architectures that require their own policies, schedules, and oversight, making it nearly impossible to maintain a consistent, SLA-driven data protection strategy. This situation also makes it extremely difficult to get a holistic view of your global unstructured data footprint. Without centralized insights into data growth, usage patterns, and trends, you can't effectively manage your data lifecycle or make informed decisions about what to keep and what to archive. And if you're backing up data to the cloud, keeping large amounts of data in the wrong tier could have a significant impact on your cloud bill.

It's time to tame your unstructured data with a unified approach that provides global oversight through a single pane of glass. Centralized dashboards should deliver actionable insights for proactive data management while enabling consistent policy enforcement.

PRO TIP

Insist on solutions that provide a unified control plane for managing unstructured data protection across all NAS systems and locations. Centralized, holistic visibility is key to developing an optimized and compliant enterprise data lifecycle.

3 Enterprise-Grade Data Security with Anomaly Detection at Scale

Cybercriminals are upping their game, and your data is in their crosshairs. 93% of organizations reported malicious actors attempting to impact data backups during a cyberattack, and 73% of those organizations said the attempts were at least partially successful.[2]

At least part of the problem is legacy solutions often leave your backup data online and accessible and fail to provide robust logical air gaps. **Air gaps** isolate and hide your backups, making them invisible to threats. Without air gaps, your backups are left vulnerable, compromising your ability to recover if those backups are attacked.

That's also why **immutable backups**—i.e., backups that can't be modified, deleted, or changed—are so important. But legacy tools often lack this feature, too. If your backups aren't immutable, an attack could overwrite your backups, leaving no clean recovery point in sight.

To make matters worse legacy tools often offer limited **role-based access controls** and overly broad permissions. This security gap leaves the door open for malicious insiders and credential-stealing hackers. Without granular, role-based access controls enforcing least privilege access, users (or threat actors using stolen credentials) can access more data than they need, increasing the risk of something happening to your data.

You need to modernize your unstructured data protection with solutions that have integrated security controls. Your checklist should include: air gaps, backup immutability, role-based access controls, and finally **anomaly detection**.

Anomaly detection helps you figure out the scope of a cyberattack by detecting deletions, modifications, and encryptions for optimal ransomware investigation. With that information, you can go in and surgically recover exactly the data you need rather than initiating a wholescale (and lengthy) recovery process— meaning you can get back on your feet quicker with your data intact.

PRO TIP

Prioritize unstructured data protection solutions that have robust logical air gaps, immutable backups, granular role-based access controls, and anomaly detection. These layered defenses work together to shield your backup data from unauthorized access and alteration and can help you get back up and running if an attack does happen.

[2] The State of Data Security: The Hard Truths, https://www.rubrik.com/zero-labs/2023-spring

4 Sensitive Data Visibility and Classification

Sensitive data refers to any information that, if compromised, could harm individuals or organizations. This can include personally identifiable information (PII), financial data, intellectual property, health records, or any other confidential information that requires protection to maintain privacy, security, and compliance.

Governing and securing sensitive data is a top priority for maintaining regulatory compliance and customer trust. However, as unstructured data grows into petabytes, locating and identifying sensitive or regulated data becomes a monumental challenge.

Legacy backup tools provide no native capabilities for data discovery and classification. With no insights into what types of sensitive data exist or where they reside, you're left with risky blind spots in your compliance and security postures.

Not being able to identify sensitive data, like PII, PHI, and PCI, lurking in your unstructured datasets leaves you vulnerable to inadvertent exposure and non-compliance with privacy mandates, like GDPR, HIPAA, and CCPA. You can't apply appropriate protections when you don't know where your critical data lives.

Lack of data awareness also leads to inefficient use of backup resources and higher costs. Without insight into data value and sensitivity, you're forced to treat all data the same. Backing up and replicating non-critical information unnecessarily will ultimately lead to increased storage costs.

To reduce sensitive data exposure and streamline compliance, you need modern solutions that combine data protection with deep content intelligence. By automatically discovering, classifying, and reporting on sensitive data, you can prioritize protecting your most critical data first, tiering that data based on how at-risk it is, and better controlling costs.

PRO TIP

Demand data protection solutions that have built-in sensitive data discovery and classification. By identifying regulated data and assessing privacy and security risks, these insights help you strengthen data governance and prove compliance. You'll also have a better understanding of what you have, so you can better protect your most critical information and back up just what you need.

5 Granular and Efficient Data Recovery

The average cost of a single hour of downtime now exceeds $300,000 for over 90% of mid-size and large enterprises, according to Information Technology Intelligence Consulting's *2024 Hourly Cost of Downtime Survey*.[3] Every second counts during a cyber incident, and you need to be able to recover quickly and precisely to minimize your losses. But if you're still relying on legacy backup solutions, it may be difficult to make that happen.

Using legacy backup tools to recover your data is like searching for a needle in a haystack. Without granular search capabilities, IT admins are forced to wade through tons of data scattered all across your environment to try to find what was affected during an attack or a disaster.

If you can't pin down the affected data, you have no other choice but to recover entire systems. This option clogs your networks, overwhelms your storage devices, and ultimately slows down your recovery. But can you even be sure the data you're backing up is actually clean?

Here's another area where legacy solutions often fall short: They don't let you verify backup integrity, which means you could be unknowingly restoring corrupted or infected data. Suddenly, you're back at square one, with the bonus of having wasted precious time and resources. On top of that, legacy solutions also lack automated recovery processes, which only slows you down further.

To reduce data loss and recovery times, you need modern tools that enable rapid, targeted recovery at scale. Deep search and granular restore options ensure you can quickly get back the data you need. Backup verification and automated recovery workflows accelerate the entire process.

PRO TIP

Ensure you adopt tools that enable rapid, targeted recovery at scale, allowing you to quickly locate and retrieve specific data without having to perform full system restores. In addition, look for backup verification and automation features that streamline the recovery process, reduce human error, and improve recovery efficiency.

[3] ITIC 2024 Hourly Cost of Downtime Report Part 1, https://itic-corp.com/itic-2024-hourly-cost-of-downtime-report/

MODERNIZE YOUR UNSTRUCTURED DATA PROTECTION

Legacy backup solutions have served their purpose well, but the landscape of data and threats has evolved. Today, unstructured data demands the same rigorous protection and management as all other critical data.

To meet these challenges, it's essential to look for a solution designed to handle petabyte-scale unstructured data. Rubrik NAS Cloud Direct delivers key capabilities to help you stay ahead in safeguarding your most valuable assets.

Here's how Rubrik NAS Cloud Direct addresses the gaps in unstructured data protection:

Rapid Backup and Performance at Scale

Rubrik NAS Cloud Direct delivers backup and recovery of petabytes of data and billions of files faster and more efficiently than NDMP, significantly outpacing legacy backup approaches. Parallel scanning, indexing, and data movement maximize throughput, while intelligent QoS eliminates impact on production workloads.

> Easy backup configuration, really fast backup, and super easy restore procedures.
>
> **Sr. IT Infra Admin**
> **Amusement and Recreation Services Company**

Unified Management and Visibility

NAS Cloud Direct provides a centralized control plane for unified management of your unstructured data. Global dashboards give you actionable insights and reports for optimizing protection, archiving, and compliance. Simplify operations with consistent policies across your entire footprint.

> Adopting Rubrik NAS Cloud Direct has improved overall efficiency in terms of speed and delivery.
>
> **Lead Product Designer**
> **Financial Company**

Advanced Data Security with Anomaly Detection

Rubrik NAS Cloud Direct hardens data security with logical air gaps, immutable backups, and role-based access controls. Plus, with Rubrik Anomaly Detection, you'll get notified when something goes awry and be able to assess the potential impact of an attack. With Rubrik, you can see how the attack happened, identify malicious activities, and quickly respond and recover.

> We can now monitor and analyze our NAS backups and get alerts anytime something looks out of the ordinary so we can investigate right away.
>
> **Kevin Mortimer**
> **Head of Operations, University of Reading**

Sensitive Data Monitoring and Compliance

NAS Cloud Direct combines data protection with content intelligence to discover, classify, and report on sensitive data. Automate discovery of PII, PHI, PCI, and other regulated data to assess compliance risks and easily find and safeguard your most critical data assets.

> NAS Cloud Direct monitors the system and lets me know if anything is happening.
> It helps me sleep at night.
>
> **Travis Spurley**
> **Sr. Systems Engineer, Quantum Spatial**

Rapid, Granular Recovery at Scale

Rubrik NAS Cloud Direct enables instant search and granular recovery to minimize data loss and downtime. During an attack, NAS Cloud Direct searches across billions of files rapidly and orchestrates recovery to help you get back to business sooner.

> The ability to search across millions, if not billions of files is huge. The Rubrik solution definitely improved our ability to do our job.
>
> **Carl Lucas**
> **VP of Information Technology, Quantum Spatial**

Don’t let your legacy backup vendor jeopardize your unstructured data. Rubrik equips you to efficiently protect and manage massive unstructured datasets while ensuring resilience against ever-evolving threats.

The time to modernize is now.
Learn more about Rubrik and NAS Cloud Direct today.