

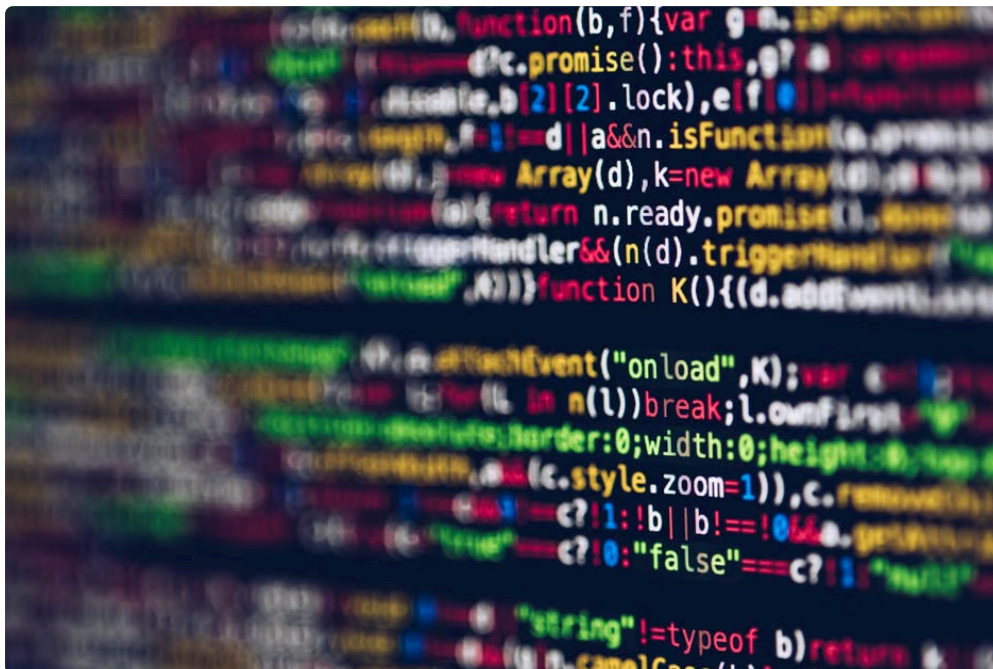
For IT teams wrestling with sprawling file servers and ballooning storage costs, understanding and managing **inactive file data** is mission-critical. Inactive files—sometimes dubbed *dark data*—can constitute anywhere from **60%-80%** of file server capacity, silently driving up expenses and risk without delivering value.

Let's break down how to measure the inactive file data percentage using *file server analytics* and *last accessed reports*. We'll explore definitions, why dark data accumulates, visible tools and techniques, and the costly consequences of ignoring this challenge—including compliance and security pitfalls. Along the way, we'll naturally introduce industry-leading resources like **Komprise**, **Gartner**, and **DataStackHub**, and tools such as **Komprise Analysis** and **Komprise Intelligent Data Management**.

What Is Dark Data and Why Does It Accumulate on File Servers?

Dark data is unstructured data that organizations store but do not actively use. It lives in file shares, NAS devices, and cloud buckets—essentially out of sight and mind, yet consuming resources. Common forms include old project files, legacy backups, duplicates, and orphaned data left after employee departures or reorganizations.

Why does dark data accumulate?



- **Lack of ownership:** Files left behind when employees move on, managers change, or projects end.
- **No visibility:** IT doesn't know what's there or how critical it is.
- **Lengthy retention policies:** Mandates to keep data years beyond active use "just in case."
- **Backup and replication blind spots:** Copying and storing inactive data multiples times.
- **Fear of deleting "important" data:** Avoiding cleanup due to uncertain impact.

The result? Inefficient storage, longer backups, more expensive replication, and higher risk exposure due to forgotten sensitive files.

Why Measuring Inactive File Data Percentage Matters

Multiple Gartner reports reinforce that up to **80%** of stored file data is rarely or never accessed after initial creation. This puts a crucial focus on measuring inactivity to better control costs and risks associated with

unstructured data.

- **Cost drivers:** Storage, backup, replication, disaster recovery (DR), and e-discovery expenses swell when inactive files occupy valuable resources.
- **Compliance risks:** Inactive files may contain personally identifiable information (PII) or protected health information (PHI), exposing organizations to breach fines.
- **Security exposure:** Unknown stale files are prime candidates for unauthorized access or ransomware encryption.

How to Gain Visibility into Inactive Data: File Server Analytics and Discovery Tools

Getting a handle on dark data starts with visibility and understanding—what do we have, who owns it, and how old is it?

Traditional methods like manual folder audits or scripts are error-prone and unsustainable at scale. Instead, modern tools leverage metadata analytics to provide real-time insights on file activity patterns.



Komprise Analysis and Intelligent Data Management

Komprise offers a robust solution designed specifically for measuring and managing inactive file data:

- **Komprise Analysis:** This module scans file servers across NAS and object storage, generating last accessed reports that identify files untouched for months or years. It calculates the percentage of inactive data by user-defined thresholds (e.g., files not accessed in 6 months).
- **Komprise Intelligent Data Management:** Once identified, Komprise automates optimization workflows—archiving inactive data transparently to cheaper storage tiers or cloud, without disrupting user access.

The combination gives IT clear, data-driven insight and action plans to slash dark data footprint effectively.

Other Tools and Platforms

Besides Komprise, platforms like **DataStackHub** provide centralized metadata aggregation and discovery that complements file server analytics. Gartner advises integrating such tools into broader data governance and risk

management strategies. Key capabilities to look for include:

1. Automated metadata collection and classification
2. Configurable inactivity criteria based on last accessed timestamps
3. Role-based access to reports to identify true data owners—one of my cardinal principles is asking “**who owns the data?**” before deciding what to do with it.
4. Dashboards showing cost impact of dark data (storage, backup, DR)

Step-by-Step: Measuring Inactive File Data Percentage on a File Server

Step Description Tools/Actions 1. Identify File Servers and Storage Inventory all file shares, NAS devices, and cloud buckets where unstructured data resides. Use asset management, storage monitoring tools 2. Gather File Metadata Collect key metadata, especially *last accessed* timestamps, file size, and ownership. Use Komprise Analysis or scripts that query file systems 3. Define Inactivity Threshold Agree on what “inactive” means for your org (e.g., files not accessed in past 6 or 12 months). Policy discussions with security, legal, and business units 4. Generate Last Accessed Reports Run analytics jobs to list inactive files, total size, and share by department or owner. Komprise last accessed reports, DataStackHub dashboards 5. Calculate Inactive Data Percentage Divide inactive file data size by total file data size for each share/server to get percentage. Automated report outputs or custom calculations 6. Validate and Engage Data Owners Confirm inactivity with file owners and validate retention requirements. Collaboration tools, emails, or governance platforms 7. Plan Data Management Actions Decide which inactive data to archive, migrate, or delete to reduce storage and risk. Use Komprise Intelligent Data Management workflows

The Cost of Ignoring Inactive Data: Beyond Storage Expenses

Many organizations run blind when it comes to inactive data. This oversight can manifest in [dark data](#) various hidden costs:

- **Storage costs:** Money wasted on expensive primary storage housing datasets that users no longer need actively.
- **Backup & replication impact:** More inactive data means bigger backups, longer windows, bigger disaster recovery footprints.
- **Security risks:** Stale files often include PII and PHI, forgotten but fully exposed—raising breach costs and regulatory fines significantly.
- **Compliance headaches:** Without clear lifecycle tracking, data retention policies can be violated, inviting audits and penalties.

Gartner emphasizes that leveraging intelligent analytics like Komprise’s can move organizations from reactive firefighting to proactive data governance.

Final Thoughts: Make Data Visibility Your First Step

Before you jump into cloud migration or new technologies, get clear visibility into your existing file server environment. Measuring inactive file data percentage with real analytics tools—not guessing—is foundational. You can then confidently engage data owners, implement compliant data lifecycle governance, and use intelligent solutions like **Komprise Analysis** and **Komprise Intelligent Data Management** to automate data movement and reduce costs.

Remember, a file server isn't just a "data dump" or "cloud candidate" —it's a living, breathing attic filled with boxes (files) you may no longer need to carry around. As always, start by asking "*who owns this box?*" before deciding what to do with it.

Further Resources

- [Komprise Official Website](#)
- [Gartner IT Research](#)
- [DataStackHub Metadata Discovery](#)

Want to know something interesting? taking control of inactive file data unlocks significant value, cost savings, and risk reduction for your enterprise.