Technical White Paper

Miria for Migration
Storage Migration
Process Overview

Table of contents

Executive Summary

The 'Miria for Migration' solution by Atempo enables businesses to securely and reliably migrate any existing file storage system to any file storage solution.

Many companies have experience using generic free tools ; however, they are not scalable, require numerous time-consuming checks and can lead to critical data loss. The impact on production storage is often not controllable.

Far from being just a tool to simply switch users to a NAS or a file storage environment, data migration requires several stages of work including a preparatory phase that is crucial to a successful migration. Just as crucial is the migration tool's ability to guarantee data integrity as well as reduce down time thereby minimizing the impact on production workflows.

This paper describes not only how the Miria software solution is efficient but the key benefits, best practices and the solution-specific configuration steps to deploy Miria for Migration.

1. Introduction

1.1. About Atempo

Atempo is a leading independent European-based software vendor with an established global presence providing solutions to protect, store, move and recover all mission-critical data sets for thousands of companies worldwide. With over 25 years of experience in data protection, Atempo offers a complete range of proven solutions for physical and virtual servers, enabling backup, archive and data migration between different storage volumes and scale. Atempo's three flagship solutions, Lina, Miria and Tina are labeled "As used by French Armed Forces" and "France Cybersecurity" and are considered the de facto standard.

Selected to join the initial selection of the French Tech 120, a government program designed to nurture 25 and accelerate growth of the top 25 companies by 2025, Atempo is headquartered in Paris with a Global footprint in Europe, the US and Asia with an expanding network in excess of 100 partners, integrators and managed service providers.

1.2. Storage Migration Challenges

Migrating unstructured file data between storages is complex.

Many companies have deployed very large file-based storage systems that will eventually require replacement, and their data migrated. There was a time where simple off the shelf tools could handle the task of migrating a 100 TB or less, but today's appliances are hundreds of TB's or many PB's requiring high performance throughput and ease of use to complete the migrations within acceptable time frames. All while not impacting the user's experience and workflows.

Free Copy or Sync tools may work for modest data sets when volumes do not exceed tens of TB and the data itself is not considered critical for business continuity.

Customers across the globe have reported some of the same challenges and or constraints which many other solutions cannot address:

- Transfer large volumes of data as fast as possible
- Copy data when the source storage is still in production
- Minimize the impact on the production during the synchronization
- Minimize the duration of the production downtime for the cutover (last step of the migration process)
- Retain file permissions and extended attributes
- Supervising data transfers

Given the explosion in the growth of data volumes and the number of files to handle, organizations require a professional, enterprise level solution guaranteeing the integrity of data but also ensuring production workflows remain unencumbered during all facets of data migration.

Atempo's experience in managing large volumes of data, led to the development of the Miria solution for Migration which enables the synchronization of data between NAS appliances, Cloud and or parallel file systems.

2. Miria Software Presentation

Full Control of your Unstructured Data
Miria engine which delivers performance and manages petabytes of data and billions of small files on site and in hybrid environments.

Miria for Migration
Migrate very large data volumes and billion files efficiently between heterogeneous storages and file systems.

Miria for Archiving
Free up storage space on high performing primary storages and manage storage growth requirements.

Miria for Backup
Rapidly back up data from damage and loss and ensure lasting protection all from a single centralized platform.

Miria is an open solution platform which encompasses Archiving, Backup and Migration of unstructured data for heterogeneous storages.

Built around Atempo's Data Mover agents, Miria is designed for high performance data protection and combines a flexible integration to business workflows with the most efficient data movement technologies across all our market verticals.

2.1. Miria for Migration Solution Overview

Miria for Migration is a solution that empowers organizations to migrate large data volumes and high numbers of files while preserving their associated ACLs between heterogeneous storages (NAS, Cloud, distributed or parallel file servers).

Powerful and reliable migration tools to keep older legacy storage in production until the new appliance is ready:

During the migration of a production file system from one storage to another, Miria detects any modified or deleted objects and manages the entire incremental synchronization process. Storage migration suddenly just became a lot easier.

Highly paralleled processes for optimal performance:

Miria implements a technology which separates the scan of objects from the actual data movement. Unlike many other solutions, Miria starts moving data very early on and uses a powerful parallel processing mechanism and multi-tasking environment adapted to all types of file sizes and workflows.

**Vendor-agnostic solution enables migration between different types of storages:**

Using standard SMB and NFS protocols, Miria for Migration collects files with their access rights and ensures full compatibility with the target storage appliance.

2.2. Architecture Overview

Miria's infrastructure is composed of the following software components:

- Miria Server
- Miria Data Movers
- Miria Web Interface
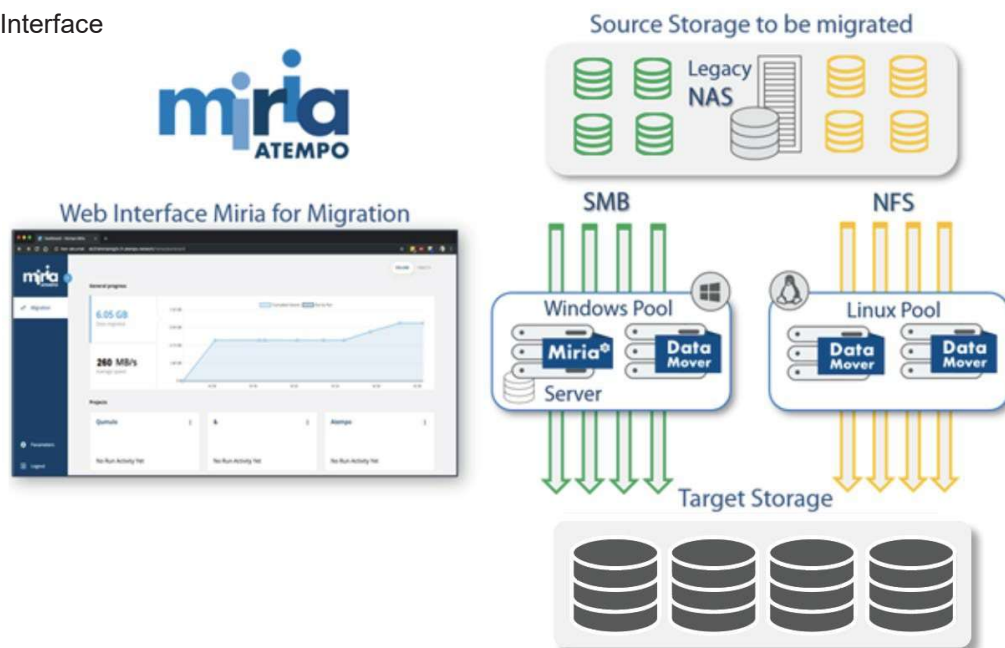
Miria Server

The Miria Server manages all the Data Movers and can be installed on a dedicated physical server or on a virtual machine.

The server hosts Miria's database which contains:

- Configuration settings
- Necessary information to perform the various steps of the migration project
- Report information

Miria Data Movers

Miria Data Movers facilitate or manage the end-to-end data movement. They are the key components with respect to global performance of the migration process, as such they should be properly sized for network, CPU, and memory configurations.

Miria Web interface

The interface of Miria for Migration is designed to facilitate the migration project(s) and allows for:

- Configuration and launching the synchronization tasks
- Supervising operations
- Viewing Statistics reports
- Viewing and sending Activity reports by email

2.3. Miria for Migration Key Features

Miria uses a multi-task environment to separate the detection of objects from the data movement process while enabling parallel processing across all Data Movers. This technology ensures data movement as soon as Miria has scanned a defined volume. Instead of waiting for a complete scan of the entire file system.

Optimized Datapath:

Data Movers communicate directly with one other, if required, and use the shortest communication path available from one storage appliance to another. Additionally, Data Movers use a multi-threading process that is robust and insensitive to network latency issues to achieve the highest throughput even on slow links (such as WAN).

Load Balancing:

Miria automatically balances the sub-tasks across pooled Data Movers in order to achieve the highest throughput possible to multiple destination storage platforms.

Seamless Failover:

If for any reason a sub-task fails on a particular Data Mover, it is automatically restarted on another Data Mover. This is particularly important when moving several hundred Terabytes which could potentially take days.

Data integrity check:

Miria guarantees that data and metadata written on the target storage appliance(s) match exactly those read from the source storage using a signature on each file called a digest utilizing xxHash64 algorithm.

Advanced storage integration

Miria leverages the advanced capabilities offered by each supported storage (through proprietary APIs) such as those currently employed for Dell-EMC² Isilon, Qumulo, and the shared file system IBM Spectrum Scale (GPFS).

Snapshot:

Snapshot technology is a marker or image of a file system at a particular point in time (e.g.: VSS for Windows). The Snapshot is employed at every step of the migration project when available. It freezes a consistent image of the data at that point in time for usage by FastScan.

FastScan:

FastScan optimizes the detection of modifications made on the file system between two snapshots. This technical feature of each supported NAS avoids the need for full scan of the entire file system on incremental runs.

The retrieved list of modifications contains the following information:

- Modified, Deleted, Created objects (files, directories, and links)
- Objects whose type has changed (e.g. a directory has become a file or vice-versa, a directory has become a link or vice-versa, etc.)

Automatic Synchronization modes

Miria offers 4 different modes of synchronization.

- **Echo:** unidirectional, replicates all changes including deletions (by default recommended for most migrations)
- **Contribute**: same as Echo but keeps deleted files on the target
- **Subscribe:** target to source replication on updated files only
- **Combine:** bi-directional on added and updated files only

2.4 Miria for Migration Key Benefits

Using processes that can be run in parallel and that are multi-threaded for optimum performance, Miria adapts to all file sizes (> 1KB up to several TB) and maintains ACL integrity, thereby providing:

- Incremental and automatic file migration between heterogeneous architectures
- Minimum impact on production and a fast cutover to the new storage
- Automatic integrity checks on all migrated files
- Scalability and Performance adjustments by simply adding or removing a Data Mover.
- Simple and user-friendly web interface to supervise the migration process
- New storage can be backed up using the same solution if necessary

The solution is accompanied by a full range of Atempo Professional Services (or by a Certified Partner) from the initial audit of the technical requirements to the final storage cutover:

3. Migration Project: Requirements and Best Practices

In today's 24x7 Global environments IT projects are seeing more and more restrictions being placed on workflows from: user demand, ever more complex environments, minimal downtime and other time constraints, mitigation of risks to data and rising budgets… Storage migration projects multiply these challenges. As data migration is a time limited project, a successful migration strategy requires a well-prepared and executed plan.

PLAN
- Audit your data sets: asset scope
- Construct a file migration plan
- Provision servers & resources
- Build the infrastructure: Server and Data Movers
- Configure migration jobs
- Test & calibrate

MIGRATE
- Migration cycles: Run & Verify
- Daily/weekly progress status
- Migration cutover
- Project validation
- Uninstall the infrastructure or leverage for backup
- Celebrate!

3.1. Migration Preparation

Collecting all relevant information to estimate the duration of the migration project which depends on:

- The volume of data and the number of files to migrate
- The modification rate on the source storage
- The appropriate duration of the production interruption on the source storage for cutover
- The global performance of the current infrastructure (performance of source and target storage, network bandwidth, etc.)

Audit the data sets

- Identify volumes / number of files / file-size distribution
- List shares, departments owning files and associated network protocols
- Ensure having all required admin / root logins of machines and the shares access rights

Construct the migration plan

- Divide data sets into several projects (similar files, by department, etc.)
- Define validation criteria for each project
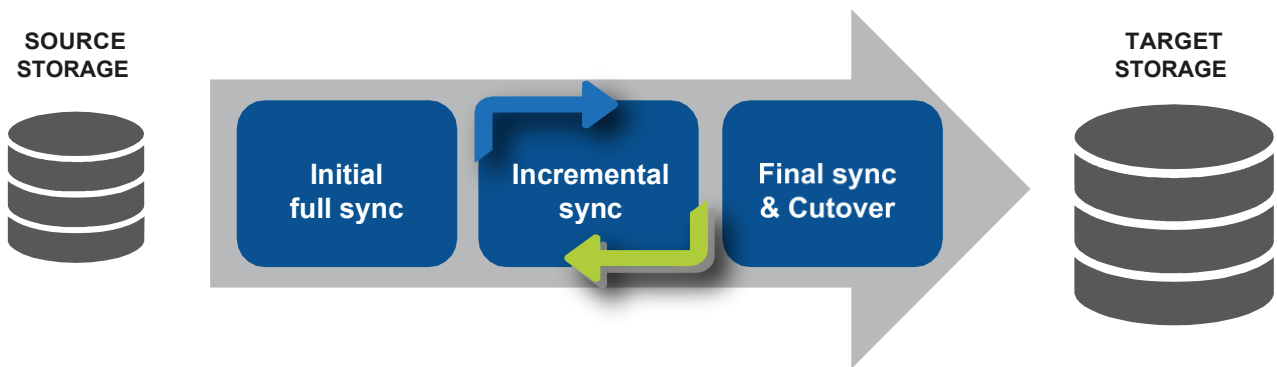- Test migration performance on each project

Tune the migration jobs

- Adjust infrastructure according to the test results
- Configure Data Movers including job parallelization, multi-threading, network, FastScan etc.
- Prepare jobs schedule
- Estimate cutover date or dates

The migration cycles

The synchronization consists of 3 cycles:

1. **First Synchronization:** an initial full transfer of data between the Source and the Target is used as a reference.
2. **Incremental Synchronization:** automatic incremental synchronization enables Miria to propagate any changes that have been made on the source storage since the previous synchronization to the target storage.
3. **Final Synchronization** (Cutover): a final incremental synchronization is performed to complete the migration to the target storage while the production is momentarily suspended on the source storage.

These steps may happen fairly quickly or take days or even weeks depending on the amount of data, customer workflow, testing and timeline for cutover.

3.2. File size and performance consideration

Using Miria software, transfer rates from the source storage to the target storage may be close to the speed of the hardware network equipment. However, performance throughput is commonly limited by filesystem read or write operations when dealing with large numbers of small files or directories.

- **For large files (>10 MB)** data movement, one Miria Data Mover can achieve a sustained performance throughput of 3GB/s or greater.
- **For very small files (<10 KB)** data movement, one Miria Data Mover can achieve a sustained performance throughput of 1-3 GB/s.

3.3. Deploying Miria for Migration - Environment

Below is an example of a typical Miria deployment and configuration migrating both SMB and NFS exports.
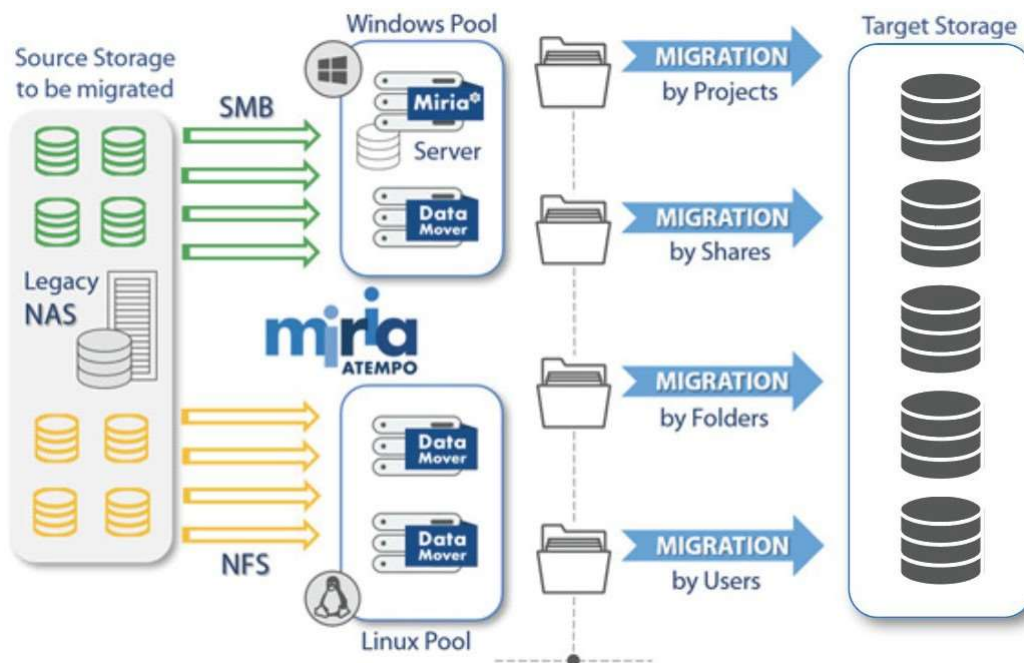
WINDOWS Pool	Miria Server 3.10	DataMover
Hostname	miriawinsrv10	miriawindm20
IP address	10.16.34.95	10.16.34.93
Operating System	Windows 2016 standard	
System	Dual-CPU 8 cores ; RAM 32 GB	
Disk	300 GB for O.S. and Miria Database	

LINUX Pool	DataMover	DataMover
Hostname	mirialindm10	mirialindm20
IP address	10.16.34.94	10.16.34.91
Operating System	CentOS 7.x	
System	Dual-CPU 8 cores ; RAM 32 GB	
Disk	100 GB for O.S.	

NAS File Server	Source: ISILON 8.2	Target: NAS, GPFS, etc.
Hostname	Atempo-isi	Atempo-target
IP address	10.16.34.89	10.16.34.50
SMB share/export	/ifs/Atempo-isi-smb_Source	/Atempo_SMB
NFS share/export	/ifs/Atempo-isi-nfs_Source	/Atempo_NFS

3.4. Miria Installation Prerequisites

Before installing Miria Software, take into account the prerequisites below.

Operating System

Miria Data Movers must run on:

- Linux (for example Red Hat, CentOS etc.) when the source storage exports shares through an NFS protocol.
- Windows (for example 2016, 2019, etc.) when the source storage exports shares through an SMB protocol.

The complete list of supported operating systems and storages is detailed in the Miria Compatibility Guide which can be downloaded from: http://www.atempo.com/resources

Linux System

Miria installer needs to be run as root with a graphical environment installed:

- Configure X Windows on the local host or set the DISPLAY variable to use the graphical environment from a remote host.
- X11 forwarding is mandatory. Some Linux distributions may also require installing Xauth for the display to be exported:
- Install XTERM package and dependencies: xorg-x11-server-Xorg xorg-x11-xauth xorg-x11-apps.

It is also mandatory to install these packages: libXext, libXtst, bc, psmisc.

Windows System

Miria must be installed as administrator.

The Windows Update process may perform automatic reboots when a scheduled task is due or while jobs are in progress. Ensure that the Windows Update process does not automatically restart the platforms hosting Miria components.

Date and Time

The Miria scheduling system requires that each platform used by Miria is on the same time and configured according to its local time zone.

Product Download

Ensure that you download the latest GA release from the Atempo Client Portal: https://support.atempo.com

License Request

Gather host MAC Address to generate the license and send the request to: lks@atempo.com

3.5. Miria Installer overview

The same Miria version must be installed on all Miria platforms.

Miria for Migration installer utilizes a single binary labeled **installMiria** (.exe on Windows and .bin on Linux).

The same binary is used for the Server and or data Mover components:

- Miria Server, including the Web interface, Apache Web Server, MaxDB database.
- Data Mover, including the agent for each platform that manages the source and target storage.

The installation process is also the same for all platform types and Operating System (Server or Data Mover and Windows or Linux). During the Miria Setup the following are the specific recommendations:

- Chose the type of installation Server or Agent mode.
- Select **Typical** option (recommended for most users).
- In the Access Server window, check secured HTTPS Port and type a number (443 is recommended).
- In the Firewall Configuration window, select **Yes** to configure Firewall automatically

3.6. Firewall Configuration

The typical installation automatically sets the default port numbers that have to be opened on the Firewall Rule. General Miria port numbers are:

- TCP/IP port: 2524
- HTTP Apache port: 80 and/or, HTTPS Apache port: 443

The TCP/IP port number can be changed by selecting **Custom** option during the Miria installation.

For more details and other ports, see the **Port Number Matrix** in the Appendix section.

3.7. Preparing source and target storages

Below is a checklist of the main steps to be configured on the target environment:

An Admin user dedicated to the migration

- If using Windows based Miria Server and Data Movers, they need to be in the same domains as your source and target storage.
- The migration needs read only admin privileges on the source and full read/write on the target to move/create all objects (files and directories) and their associate ACLs.
- Create and verify a dedicated user(s) in your Active Directory. This user is then used by Miria Data Movers when migrating data.
- Declare the associated user / role into your source and target storage platform. It is common to utilize a **Backup Admin** for this role.

Source storage:

- **NFS Share/Export -** Mount your NFS exports on each Miria Data Mover individually. The mount recommended command is:

      # mount -t nfs -o nfsvers=3,tcp,rw,hard,intr,timeo=600,retrans=2,bg nfsserver:/export /mnt

- **SMB/CIFS Share/Export -** Should be browsable from each of your Windows Data Mover (s). No need to mount the share when accessing through SMB/CIFS.
- **Multiprotocol Share/Export -** If your storage offers support for multiprotocol from the same share, that is the capability to enable simultaneous SMB and NFS clients read-and-write access to the same share and data, you will need to make a choice regarding the network protocol to be used for migration of those share(s).

Target storage:

- Create an NFS export and ensure the admin user created earlier has the required access.
- Create an SMB share and ensure the admin user created earlier has the required access.
- If your storage offers support for multiprotocol from the same share, you can of course create these types of shares they are fully supported by Miria. Just keep in mind that individual files are migrated and written to the share using either SMB or NFS.
- Make sure SMB/CIFS Share/Export are fully accessible in read/write from each Windows Data Mover.

# 4. Miria Configuration and Monitoring Overview

The Miria Web interface is used to configure Miria infrastructure and migration projects. The following steps are used:

The same Web interface is also used for monitoring each migration project task(s), as well as the overall migration project.
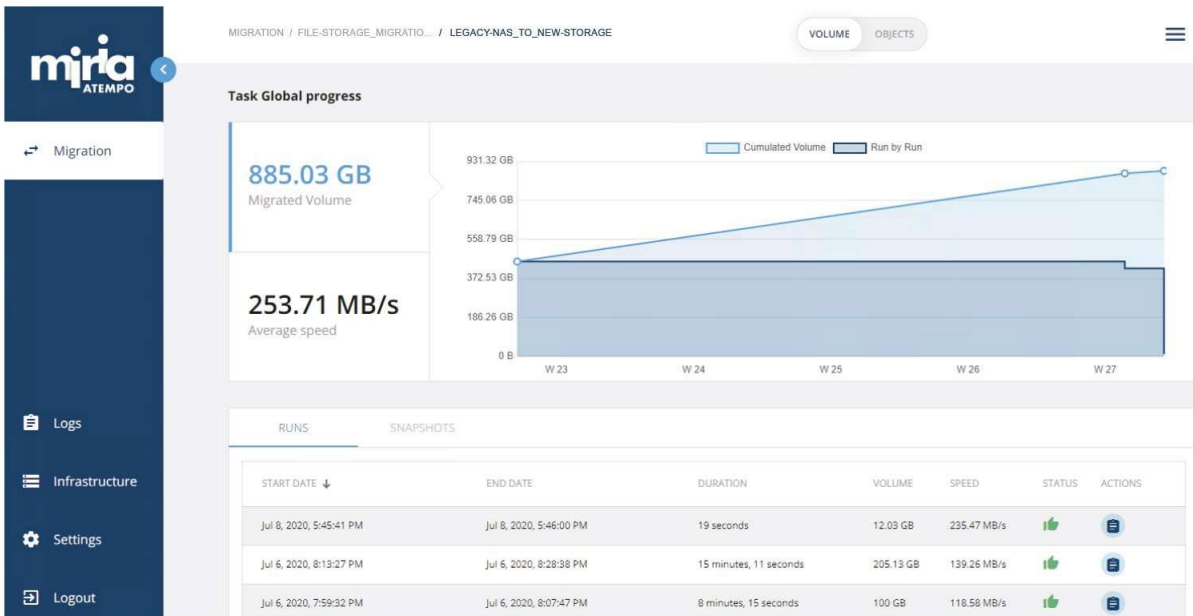
The above screen capture is a global view of a migration project showing the status of each incremental run performed.
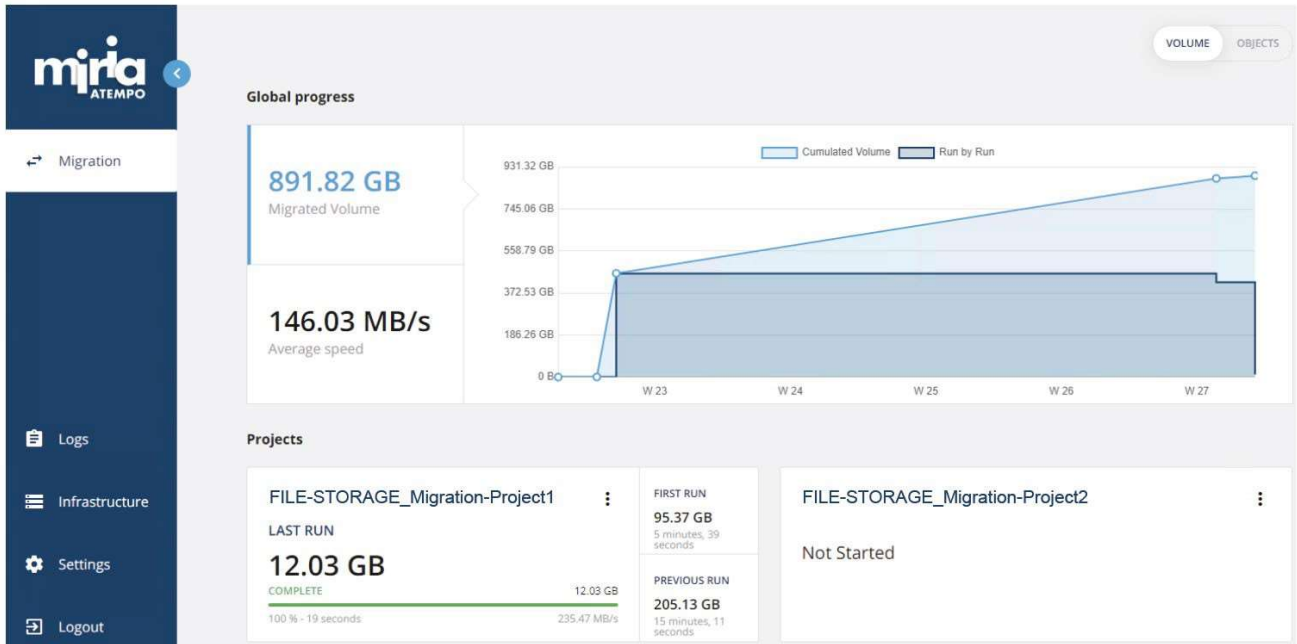
The above screen capture gives a global view of the progress of all the migration projects in progress.

Conclusion

Migrating large volumes of unstructured data (hundreds of terabytes or even petabytes) from a large scale-out NAS to a more capacitive File Storage Platform demands fast and reliable data transfers to the higher capacity, more flexible storage, while successfully managing growing data volumes. The right solution must attain the seamless migration of billions of files, allow for ongoing automatic backups, improve security, and ultimately lower costs.

This is where Atempo's Miria for Migration can help.

Atempo's Miria Data Management software enables customers and partners to synchronize, move, backup and archive unstructured data from different types of source storages (NAS, any NFS/SMB/S3, and PFS) to a wide range of storage destinations (disk, object, optical disk, tape, cloud).

Miria's FastScan avoids time consuming file tree walks during filesystem scanning which means petabyte data movement is easy to manage without impacting users with continues filesystem tree walks. In addition, the Miria Data Movers enable parallel data streams and multithreading to all data volumes removing workflow constraints.

Atempo's Professional Service Experts work with businesses to securely and reliably migrate their data from legacy storage systems to a newer storage platforms. Providing support and industry vetted knowledge for a successful migration project(s), by:

- Identifying business requirement
- Mapping the migration project
- Preparing the data for migration
- Deployment of Miria as the best tool to meet a complex migration project requirement
- Running a data migration pilot test
- Implementing the file migration project
- Validating the successful data migration