



CloudSoda
Cloud Data Management

WHITE PAPER

Media Workflow Data Orchestration with CloudSoda

Data Management Leveraging
Dell PowerScale and ECS



In the dynamic ecosystem of media workflows, companies, including industry leaders like Dell, face the daunting challenge of managing a diverse array of storage environments. From the high-speed demands of on-premises PowerScale file storage to the scalable realms of private Elastic Cloud Storage (ECS) and public cloud options, and even the nuances of direct attached workstation storage, the complexity of orchestrating data across these varied platforms is a significant industry pain point.

Addressing these challenges, CloudSoda's Global Data Management Platform stands out as a key innovator—a modern, cloud-native solution that's acutely aware of on-premises requirements. This platform empowers media companies to orchestrate sophisticated data movements with ease, offering a unified, intuitive interface designed to streamline and simplify the complexities of diverse storage management.



Media Workflow Data Orchestration with CloudSoda's Cloud Data Management Platform

Media production has evolved into a global enterprise, with productions dispersed across multiple facilities in different regions. This international spread often leads to a patchwork of storage systems, such as high-performance PowerScale file storage at each location. At the heart of this network lies a central object store, which could be a private ECS or a public cloud, serving as the organizational core for storage. Alongside these shared-storage solutions, there is also data located in remote areas on internal direct-attached storage or small NAS based devices.. These diverse storage elements constitute a complex data ecosystem.

The challenge of orchestrating data movement within this ecosystem is monumental. Ensuring that the correct files are available at the necessary locations at the right times is crucial. Effective management and transfer of production assets are not just technical tasks; they bear significant consequences on production timelines and budgets. The ability to navigate this intricate web of storage and transfer requirements is vital for the smooth operation and financial health of media production endeavors.

CloudSoda's innovative platform is expertly crafted to address the complexities of diverse storage environments. It features a central console that offers comprehensive visibility into all storage types, including object, file, and local. This console facilitates the configuration of data transfers, employing a policy-based engine that handles copies, moves, and syncs across different data formats. Additionally, it seamlessly communicates with local agents at each site, enabling WAN-optimized data movement directly among all locations.

Deployed as a fully managed solution, the CloudSoda controller simplifies user interaction. Clients can easily log into their dedicated instance and download agents compatible with Windows, Linux, and macOS for deployment adjacent to storage resources. These local agents have the capability to communicate with various storage targets simultaneously. The versatility of CloudSoda allows for multiple agents to connect to the same storage target in distinct ways – for example, a Windows agent might use SMB, while a Linux agent opts for NFS. This intelligent platform distinguishes between these different scenarios, providing tailored agent-based access to shared storage and ensuring efficient and effective data management.



Managing a Mix of Storage Types: Concepts and Architecture

CloudSoda's platform, tailored for media workflows, offers an intuitive orchestration of data movement across PowerScale and ECS storage systems. This platform comprises two key components: a cloud management console and local agents.

Cloud Management Console: Also known as the CloudSoda Controller, this interface is a fully managed service provided by CloudSoda. It allows administrators to configure storage targets, local agents, and accessors (the methods through which agents connect to storage). This console is central for setting up and monitoring data movement policies and managing one-time transfers.

Local Agents: These are installed on machines with access to various storage targets, which could include direct-attached blocks, shared NAS, or object storage. Operating as background processes on local machines, these agents execute the data movement tasks, with all management handled through the cloud console.

CloudSoda's architecture is modular, providing flexibility for a range of scenarios, particularly when accessing the same storage through different protocols. It supports both file storage (shared or local direct-attached drives) and object storage (private like Dell ECS or public cloud).

Agents form a crucial layer in this architecture. They are essentially computers with access to storage targets responsible for data movement. CloudSoda supports agents for Windows, Linux, and macOS. Once an agent is configured in the console, a configuration file is generated for download.

The platform also introduces the concept of storage accessors. These define how agents connect to storage targets. For instance, a Linux client may connect to an NFS share, while a Windows client uses SMB for the same storage. This approach allows multiple agents to access the same storage device, simplifying connectivity and eliminating the need for complex path mapping.



CloudSoda Data Movement and Policies

How CloudSoda Moves Data

CloudSoda's data management system features advanced agents that facilitate efficient data transfer. These agents employ encrypted, WAN-optimized UDP transfers, ensuring data integrity with checksum verification at both ends. They are adept at interacting with object storage using the S3 API and can handle multipart uploads for large files. The transfer speeds are influenced by the storage performance, the compute resources available on the agent host, and the network link speed between agents. This system dynamically adjusts the number of threads based on file size, optimizing performance across WAN or LAN links, regardless of whether the data is file or object-based.

The accessibility of data is a key advantage of CloudSoda's platform. File data transferred to an object store remains natively accessible via S3 APIs, and objects moved back to file storage are equally accessible in their native format. This platform does not rely on any proprietary gateways for its operation.

Policies and Quick Transfers

Within the CloudSoda console, a policy-based engine controls data movement. Users can configure policies to run on a schedule, create drop folders, or initiate one-time transfers. This engine also includes features for error and conflict resolution. A unique aspect is the ability to perform a dry run of policies before actual implementation on production datasets, an integral part of CloudSoda's patented technology.

CloudSoda's policies are designed for precision, moving only necessary data. They offer multiple filtering options, including the differentiation of hot and cold data or the identification of specific file types and names, ensuring that data management is both efficient and tailored to specific needs.



API and Webhook Support

CloudSoda's platform enhances data orchestration with robust API and webhook support. It features a REST API that allows external systems, like Media Asset Management platforms, to monitor and manage large-scale data movements. While it doesn't replace the management console, the API leverages CloudSoda's capabilities into broader media workflows.

The API enables the programmatic creation of transfer jobs, complete with all necessary parameters. Users can define the nature of the job (move, sync, or copy) and set filters based on file type or name. Additionally, the API offers real-time monitoring capabilities for both new and existing transfer jobs.

Complementing the API, CloudSoda's platform also includes webhook support. This feature allows for proactive notifications to be sent to systems upon job completion, potentially triggering further data processing activities. Configurable within the console, webhooks can be set for various job conditions, offering flexibility and responsiveness to dynamic workflow requirements.

In media workflows, especially remote productions, the platform proves invaluable. It facilitates the rapid transfer of camera data to post-production facilities, overcoming traditional logistical hurdles. With CloudSoda agents deployed on data transfer machines at remote locations, data can be efficiently moved from local storage (like RAID or hard drives) to object storage or PowerScale NAS back at the post-production site, streamlining the entire production process.

Media Workflow Use Cases

Remote Operations

In the world of remote production, content creation often spans multiple facilities, each requiring access to essential media assets. CloudSoda's platform efficiently addresses this need with its policy-based engine. This engine is capable of synchronizing files to an object repository while also facilitating direct data transfers between file storage at various remote locations. This dual approach ensures that stable media assets are securely stored in object storage while more dynamic, project-specific files are swiftly copied between sites.



Remote Users and Main Production

Additionally, the platform addresses scenarios where remote users, such as those working from home, require access to production data. It also provides an automated solution for transferring newly created media back to the main facility. CloudSoda's platform ensures secure, automated data movements to remote users, with complementary policies available to move data back to either an object core or high-performance file storage at the main studio.

Remote Users and Main Production Facility

For multi-site productions that involve both remote users and a main production facility, CloudSoda's platform is particularly beneficial. It supports an object-core approach, facilitating seamless collaboration and data access regardless of location.

Media Asset Management Data Mover

The platform's API enables external applications to initiate data movements or archives programmatically. A media asset management system, for instance, can use this API to instruct CloudSoda's platform to execute large-scale data transfers. The platform not only carries out these file or object movements but also communicates with the asset management system via webhook upon job completion, ensuring a smooth and coordinated workflow.



Get Your **Free 14 Day Trial Today!**

Seeing is believing, let us give you a detailed tour of CloudSoda and answer any questions you have about the powerful tools its technology has to offer you and your organization.

VISIT [CLOUDSODA.IO](https://cloudsoda.io) FOR DETAILS