ObjectMatrix

ProcessinPlace

Problems Processing Petabytes of Data

As more data is generated than ever before the ability to process it in an efficient and intelligent manner becomes more complex. How to process petabytes of high definition CCTV footage to spot the moment an event occurs? How to automatically classify and log content with little or no human interaction? How to reformat or transcode petabytes of video data into different formats without moving that content between the storage and the servers needed to performs the processing?

Ingest > Move Data >

Processing Servers < Move Data <

Big Pool of Dumb Storage

Some Scalable NAS

Processing Data Where it Lives

MatrixStore from Object Matrix has been designed to process data, in place, where it lives. Object Matrix MatrixStore cluster combines multiple nodes, consisting of CPU and storage, providing a self-managing digital preservation platform. The added advantage of using intelligent nodes is that the CPU can be used for other tasks when the core digital preservation work has been completed. Those tasks can range from automated metadata extraction to detailed data analysis. MatrixStore PiP (Process in Place) is the framework provided to perform tasks on data without moving the data around. PiP utilises the power of the MatrixStore cluster to perform the processing where the data lives.

Ingest > Move Data >

Processing Results <

Transcode Vault | MetaData Extraction Vault

PiP PiP PiP PiP PiP

MatrixStore Nearline Storage

Broadcast Archiving Example
MatrixStore PiP can automatically extract metadata from AS11 video files. Once extracted users can search on that metadata using applications integrated with the MatrixStore API.

The first implementation of the PiP framework is based on the extraction of metadata from standard video formats. The first of which is the AS11-DPP format.
*   Organisations simply store AS11 files into the AS11 MatrixStore vault
*   MatrixStore PiP extracts the metadata automatically when the cluster is not performing digital preservation tasks
*   The metadata is stored as part of the object within MatrixStore and made available for search
*   The data is not delayed for pre-processing nor moved for post-processing. It is processed in place, where it lives

MatrixStore PiP Benefits

*   Extracts metadata from ingested assets and makes it available for search via a fast and powerful distributed database
*   Flexible, Scalable and Extensible framework that allows advanced data analytics and advanced post processing algorithms to be carried out
*   Makes the metadata available via an API and shares metadata with technology partner's applications*
*   MatrixStore PiP technology uses the power of the cluster's CPU farm to quickly analyse assets without putting an extra load on the client CPUs
*   Because the archive's content is processed by local CPUs within direct attached servers there is no need to move petabytes of content in and out of the archive to perform post-processing
*   AS11-DPP and MXF metadata extraction algorithms included in the core package

* Where Integrated