

# Sync&share solution for massive multimedia data

Maciej Brzeźniak, Krzysztof Wadówka, Paweł Woszek, Filip Blicharczyk, Mikołaj Płaczek,  
HPC Department

Maciej Głowiak, Maciej Stróżyk, New Media Department  
Poznań Supercomputing and Networking Center,

*Keywords: cloud storage, sync&share, high resolution video, multi-media, collaboration*

Handling 100s TBs of data containing uncompressed video files locally is common need in the 8K media recording, processing and delivery domain. In the same time large, virtually unlimited storage capacity, high performance and ease of use that includes convenient data sharing over the long-distance networks rarely go together in today's storage systems.

In the presentation we will demonstrate how a scalable sync & share storage system based on robust synchronisation mechanisms and equipped with functional and well-optimised user tools and interfaces can help in harnessing such an highly demanding workload.

The discussed solution is based on Seafile, a scalable and reliable sync&share software deployed in PSNC's HPC department servers and storage infrastructure. These components comprise box@psnc service, provided internally, and box@pionier<sup>1</sup> service that PSNC offers to academic community in Poland since 2015 through the PIONIER network.

The service initially targeted mostly regular users who store their documents, graphics and other multimedia as well as researchers who handle computations input and output files. However, over time, the system proved to be able to serve much more demanding use-cases including addressing needs of power users who use sync & share as their interface to HPC cluster storage as well as PR department that treats sync & share service as the back-end for the digital content serving platform for PSNC web portals.

However the real challenges were yet to come. In autumn 2017, Immersify EU project<sup>2</sup> started to use Seafile as the basic platform for storing TBs of 8K video files and sharing them among project partners located across Europe - in France, Germany, Austria, Poland and Sweden. Within this project raw versions of the data generated using PSNC's equipment are to be used for developing, optimising and testing video encoding solutions at partner's sites.

This use case added brought new aspects to the already interesting mixture of requirements against our service. This includes the need to handle very large data volumes so that they appear as stored locally at sites, which enables immediate access to arbitrary fragment of the dataset for ad-hoc video preview or playback without the need of moving the data around. In the same time ease of use and convenience of access methods is expected.

---

<sup>1</sup> box.pionier.net.pl

<sup>2</sup> www.immersify.eu

In our presentation we will analyse the specific requirements of 4K and 8K video data serving, exchange, distribution and sharing as well as demonstrate how we exploited features of Seafiler in order to address them. We will explain what aspects of synchronisation engine and algorithms helped to achieve required performance and responsiveness of the solution. We will also show that clients such as SeaDrive, an equivalent of Dropbox's Project Infinite, enable transparently accessing large files despite the fact that they are massive in volume and are handled over long-distance network. We will also analyse the impact of the configuration and state of the network at the service back-end on the overall service performance as well as the impact of the long-distance network latency on the ad-hoc data access efficiency. We will also show how Seafiler's ability to cache data locally can help in overcoming the limits of the physical infrastructure.

*Status of the work:* At the current stage we finished a conceptual design of the solution architecture and started scalability, performance and reliability analysis based on limited-scale laboratory tests conducted using representative samples of the video content. In 1Q2018 we will gather measures observe system behaviour and user experience in a truly distributed setup.

## **Vitae:**

**Maciej Brzeźniak** leads the Data and Storage Management Group in HPC Department of PSNC. He is also a senior administrator of the PSNC storage systems. His interests include scalable data management and storage technologies, storage systems performance and data security. He co-lead development of the National Data Storage system well as its deployment as the Popular Archiving Service for academic community in Poland. He co-chairs the SIG-CISS group in Geant and participates in several initiatives and projects related to data management and cloud services including EUDAT/2020, Indigo-DataCloud, EOSC-hub and Geant4.

**Krzysztof Wadówka** is a storage and data management systems specialist at HPC Department of PSNC. His professional interests include high-availability, performance analysis and optimisation and scalability in cloud and HPC storage systems. He plans, develops and maintains the storage services offered internally at PSNC and within the country-wide and international collaborations. Since 2 years he manages our sync & share service.

**Paweł Wozuk** is a storage systems administrator in HPC Department of PSNC. His interest include reliability and performance in cloud storage systems. Currently he manages Ceph clusters as well as cluster filesystems, that provide back-end for PSNC's OpenStack and other cloud computing and storage services and HPC systems.

**Filip Blicharczyk** is Linux and Open Source systems enthusiast. He works as cloud and storage sysadmin in HPC Department of PSNC with main focus on OpenStack and Ceph deployments and management. He is involved in activities related to providing, designing, operating and monitoring cloud storage back-ends for PSNC's data management services.

**Mikołaj Płaczek** is an experience network and IT services administrator. He recently joined HPC Department at PSNC. His professional interests include scalable, high-performance networking. He developed and maintained several services used in mass scale at one of the largest universities in Poland including deployment and integration of Office 365, as well as development and maintenance of e-learning and content serving platforms.

**Maciej Głowiak** is Head of New Media Department at PSNC's. His technical background include network software engineering and parallel computing. His research interests include new multimedia technology (e.g. stereoscopy), Ultra High Definition (4K, 8K) video hardware and software development, uncompressed and compressed video streaming, HD and 3D videoconferencing, high-capacity optical networks, network monitoring and new protocols (e.g. IPv6, experimental networks). Recently he has been working on designing and building 4K and 8K node in Poland and leading an Ultra High Definition & Network work package in the European VISIONAIR project. He is a coordinator of the Immersify project and he is involved into ImmersiaTV (H2020) project.

**Maciej Stróżyk** is new media technologies specialist in New Media Department at PSNC. His interests focus on the ultra high resolution (4K, 8K) multimedia streaming technologies and real time communications systems. Maciej has been participating in many research projects in collaboration with European and national partners: 6NET, PHOSPORUS, Clusterix, Platon, Future Internet Engineering, GEANTN2, GEANT3. He was involved in designing and building 4K and 8K visualisation laboratory in PSNC and currently he works in Immersify, ImmersiaTV and GEANT4 projects.