



SAGE for European EsDs (Extreme Scale Demonstrators)

Highlights of your project?

SAGE is addressing the problems of storage and I/O for the extreme scale HPC community, which is one of the primary roadblocks in achieving Exascale. Storage and I/O has traditionally lagged behind innovations in compute, and this problem will be exacerbated as HPC systems scale into the Exascale Regime. It is critical for the storage community to co-design storage and the I/O stack with the user community. Such a strong co-design is being pursued within the SAGE project. We have a hardware prototype of the storage system now available for continued evaluation within the use cases within the evaluation environment (Juelich Supercomputing Center). We have also completed the architecture and design of the SAGE software stack (Object storage software, data management tools, programming models, data analytics solution and runtimes).

The anticipated technologies suggested for inclusion in EsDs

- (a) Exascale Object Storage Software (Mero) ; and its ecosystem below
 - Data Management Tools (HSM, Data Integrity Checking, Performance Analysis)
 - Programming models, runtimes and visualization services
 - Apache Flink suitable for Exascale HPC,and,
- (b) Storage servers with in-built compute capability

The SAGE system and the software innovations will be at **TRL5** (Validated in relevant environment) towards the end of the project:

https://ec.europa.eu/research/participants/data/ref/h2020/wp/2014_2015/annexes/h2020-wp1415-annex-g-trl_en.pdf

How should this technology be used / integrated (I/F, APIs)?

The Exascale Object storage will be used through its API Clovis. The ecosystem innovations can be separately exploited as they are built on top of Clovis.

Are there any pre/co-requisite items?

Availability of NVRAM technologies will be critical for the exploitation of the SAGE software components

Any extra work/interaction (on top of current project roadmap) needed to make them ready?

Getting more user communities outside of the SAGE project aware of the capabilities of Mero, Clovis and its ecosystem. SAGE also needs to understand new HPC use cases and workloads (Eg: AI/Deep learning).

What information/actions are needed to best prepare for EsD projects?

We need to know what are the plans for incorporating existing FETHPC innovations into EsDs without losing critical innovations.