Project[1] Number: 671500

Project Acronym: SAGE

Project title: Percipient StorAGe for Exascale Data Centric Computing

Periodic Technical Report

Part B

Period covered by the report: from 01/09/2015 to 28/02/2017

Periodic report: *[*1st*]*

[1] The term 'project' used in this template equates to an 'action' in certain other Horizon 2020 documentation

1. Explanation of the work carried out by the beneficiaries and Overview of the progress

The work carried out during the first 18 months of the SAGE project defined the overall architecture of the SAGE system including the architecture and designs of the individual software components of the SAGE system (including the Mero Object store and its API, the ecosystem tools, use case access, programming models and runtimes). We also completed the formal co-design activity with all the use cases. The SAGE prototype hardware is implemented and is now functional.

The SAGE project dissemination targets included HPC communities, system administrators, storage and I/O communities, large research institutions, application communities, big data community and the general public via events, publications, website updates including the SAGE video, social media, SAGE white paper and press coverage. Collaboration has been sought both within and outside the H2020 programme. Market exploitation of the SAGE components has already begun through on going discussion and early evaluations with large potential customers.

Summary of deliverables and Milestones

Project results have been delivered in all the work areas as defined and planned in the Description of Action (DoA).

The project has defined and delivered the requirements for the SAGE architecture based on co-design inputs from the use cases, including a detailed study of the use cases in lieu of the SAGE architectural assumptions. The project has also provided and delivered an exploration of the use of performance analysis tools to characterize the application requirements especially in the areas of storage and I/O.

The project has defined and delivered the architecture and design of the key concepts of the percipient storage software components based on Mero and its API Clovis. An analysis of potential Non-Volatile RAM (NVRAM) technologies has been done including its emulation possibilities in the absence of these NVRAM devices within the timeframe of the project.

The project has delivered the detailed designs and architecture of a tool (Hierarchical Storage Management, or HSM) to manage the hierarchical movement and locality of data in the SAGE system. Legacy data access methods have also been defined, architected and delivered. Integrity checking mechanisms for data have also been defined and delivered. Detailed designs and architecture of the performance profiling and debugging tools for the SAGE platform have been delivered.

Architecture and designs for programming models exploiting the SAGE architecture (MPI/IO), the SAGE runtime system exploiting the SAGE hierarchy, the SAGE visualization system working with the multiple tiers, and the Apache Flink based data analytics system have been delivered.

The SAGE hardware has been co-designed, defined, architected implemented and delivered as a prototype - being ready for shipping into Juelich Supercomputing Center for validations with the SAGE use cases. Plans for integration of the SAGE prototype into Juelich have also been defined and delivered.

The project website (incl. social media presence) has been delivered very early on in the project. The detailed strategy for dissemination, exploitation, collaboration and standardization has been delivered. Two snapshots of the activities in these areas have been delivered during Month 9 and Month 18 – in alignment with that strategy

The project has completed, all on time, 9 of the 18 milestones, all indicated in Part A of the periodic report.

Summary of exploitable results and how it will be exploited

Partner	Exploitable Result	Plan for exploitation
Seagate	**Mero**	(In Progress) Seagate is in discussion with various organisations to evaluate early products and feature sets based on Mero. We refer to D6.5 (Section 4.3) on the update on commercial exploitation activity of Mero by Seagate. Discussions on Mero with National labs in the US, large supercomputing centres, start-ups, weather forecast and weather services organisations have started A Mero based beta product installation is being commercially evaluated by a large research organisation, which is also a partner in SAGE. IP generation activity is also in progress (function shipping).
Seagate	**Clovis**	(Planned) Plans will be made in the next 18 months to move Clovis to become the de-facto Object Storage API for Exascale.
Seagate	**OEM Storage Enclosures (Tier 2,3,4 of SAGE)**	(Planned) Seagate OEM Storage enclosures in Sage is expected to provide a pathway for Seagate OEM storage solutions in extreme computing.
Seagate	**Storage Technologies (NVRAM)**	(Planned) Storage technologies used as part of the Sage tiers is expected to provide a pathway for seagate device technologies, namely hard disks and the Flash solution, in Extreme Scale HPC market place.

KTH	**MPI I/O Objects, MPI windows for I/O**	(Planned) KTH is expected to actively participate in the MPI Forum (the MPI standardization committee) and discuss new concepts in MPI I/O, such as storage objects, and extension of existing concepts, such as MPI windows, based on work done in SAGE
	MPI Storage Windows in MPICH	MPICH is one of the most used MPI implementations. During the first part of the SAGE period, KTH implemented MPI storage windows in MPICH. It is expected that SAGE will contribute to MPICH implementation.
	Parallel I/O Teaching/ Training Material	KTH is using the results of SAGE project as material for the KTH courses "Methods in High Performance Computing" and "Introduction to High-Performance Computing". In particular, the material from "state of the art" deliverables, such as D4.1, is used to prepare lectures on parallel I/O.
ATOS	**NVMe standard implementation**	(Planned) Technical preview mode NVMe Standard implementation to allow early usage exploration
DFKI	**Data Analytics in SAGE**	(Planned) Efficient support of iterative data flows and efficient recovery
Juelich	**New Memory and Storage**	(Planned) Exploitation of new storage and memory technologies through suitable software layers
Juelich	**Visualization**	(Planned) Link large capacity storage systems and high-performance visualisation nodes
CEA	**pNFS**	(Planned) Support of object storage backend for use in HPC
CEA	**Information Lifecycle Management (ILM)**	(Planned) Implement ILM concepts on storage objects for use in HPC with HSM in SAGE
CCFE	**Scientific Productivity**	(Planned) Deconstruct large monolithic data structures into simpler data objects with finer granularity and higher performance and to reengineer the IMAS API interfaces to modelling and analysis codes - for use in ITER
STFC	**Usage of new memory and storage technologies**	(Planned) Exploitation of new storage and memory technologies for enhancing performance in their service to provide supercomputing for data-intensive

		applications.

1.1 Objectives

The Primary Objectives defined in section 1,.1 of the DoA are.

- *Provide a next-generation multi-tiered object-based data storage system (hardware and enabling software) supporting current and future-generation persistent storage media, (solid-state and disc) within an I/O hierarchy. We term this "Percipient Storage".*
- *The project;*
    - *Redefines the storage subsystem as an integral part of the computational infrastructure.*
    - *Provides integrated computational capability anywhere in the storage system.*

The project has defined the architecture and designs of the function shipping & compute –in-storage capabilities within the Mero Object store as part of work in WP2 (defined later below). The project has provided the hardware architecture that is suitable to handle computations in WP5 (defined later below). The project has defined the designs and architecture of the runtime system that will do pre & post processing of specific tasks within the storage system as part of work in WP5 (defined later below).

At this stage the results have been in the area of architecture & design definitions that are in-line with the objectives as can be seen in the deliverables in WP2 and WP5.

- *Significantly improves the overall scientific output through advancements in systemic I/O performance and latency, and drastically reduces data movements hence improving energy efficiency by:*
    - *providing the ability to flexibly move appropriate computational workloads to where the data resides*
    - *providing a storage architecture built from the ground up to handle Exascale I/O*
    - *providing a potential to use resources in the computational cluster as part of the storage system*

In addition to the project outcomes defined above, the project has implemented a hardware architecture from the ground up (WP5) with multiple tiers of storage that can gracefully handle I/O requests either for performance (exploiting higher tiers for higher bandwidth and lower latency) or capacity (exploiting the lower tiers). Tools such as HSM (WP3), and, Runtimes and visualization (WP4) have been defined that intelligently move data between the tiers for enhanced I/O performance. By defining function shipping (WP2) and runtime mechanisms (WP4), the project is on the path towards drastically reducing data movements between the compute and storage. Further detailed performance analysis and debugging tools (WP3) have been defined to understand the behaviour of such a complex system.

Further, the project has provided the design of Mero Object storage features such as Layouts that can provide visibility to data in NVRAM pools through Object storage software. These pools could reside in the compute nodes in the future.

At this stage the results have been in the area of architecture & design definitions that are in-line with the objectives as can be seen in the deliverables in WP, WP3 and WP4.

- *Provides a roadmap of technologies supporting data access for both Exascale/Exabyte and High Performance Data Analytics (HPDA) requirements:*
    - *Targeting scalability to 500-1000PBytes, with bandwidth in the order of 60TB/sec with a storage system energy footprint of less than approximately 5KW/petabyte;*
    - *With flexible and efficient usage of HPDA application environments regardless of the compute node's architecture and implementation.*

The project has defined and provided architectures for the technologies namely:

- Mero Object Store and its exascale components (WP2)
    - Layouts
    - Function Shipping
    - Containers
    - Data format interfacing
    - High Availability at Scale
    - Distributed Transaction Management
- The Clovis API (WP2)
- HSM tool(WP3)
- Performance analysis and debugging at scale (WP3)
- Object storage integrity checking at extreme scale (WP3)
- Apache Flink data analytics for Mero (WP4)
- MPI/IO for Objects (WP4)
- Runtimes & Visualization mechanisms for multiple tiers with in-storage compute (WP4)

Each of these software components are uniquely exploitable to be part of Extreme scale/HPDA roadmap as we head into Exascale. Also all the software components are agnostic to compute node architecture.

At this stage the results have been in the area of architecture & design definitions that are in-line with the objectives as can be seen in the deliverables in WP2, WP3 and WP4.

- *Investigates and documents the requirements of relevant HPC applications and their storage use cases as part of a co-design approach.*

The project has provided a detailed investigation of all the use cases, namely from CCFE (3 use cases), Juelich (2 use cases), KTH (2 use cases), Diamond (1 use case) and DFKI (1 use case) and provided co-design inputs into the SAGE platform (WP1).

At this stage the results have been in the area of co-design inputs and detailed application documentations that are in-line with the objectives as can be seen in the deliverables in WP1.

At this stage the results have been in the area of architecture & design definitions that are in-line with the objectives as can be seen in the deliverables in WP2, WP3 and WP4.

- *Provides programming models and access methods for the SAGE architecture and validates their usability, including (but not limited to) legacy applications and 'Big-Data' data access and analysis methods.*

The project has provided the design and architecture for MPI/IO for Objects (WP4), data format access for SAGE from ubiquitous data formats such as HDF5 (WP2), parallel file system access for Objects (WP3) from legacy applications and Big Data access hooks from Apache Flink (WP4).

At this stage the results have been in the area of architecture & design definitions that are in-line with the objectives as can be seen in the deliverables in WP2, WP3 and WP4.

- *Validates the the full system in a relevant environment, for a relevant set of applications and benchmarks on a SAGE prototype integrated into an HPC data centre, validating performance, scalability, energy efficiency and the reduction in data transport requirements. The validation will enable the projection of an order of magnitude improved overall system efficiency in the Exascale regime compared to existing methods, showing increased efficiency with scale.*

At this stage the results have been in the area of architecture & design definitions of hardware, prototype availability and plans for integration into Juelich, that are in-line with the objectives as can be seen in the deliverables in WP5.

1.2 Explanation of the work carried per WP

1.2.1 Work Package 1

The initial objective of work package WP1 has been to investigate and characterise the performance features of the applications within the SAGE Application portfolio (described in D1.1) and to derive requirements for the SAGE architecture. Within this work package these applications are now being ported to the SAGE prototype in order to validate usability and efficiency of this architecture. The work package is organised in 4 tasks concerned with application characterisation (T1.1), requirements specification (T1.2), application porting (T1.3), and tools-based application characterisation (T1.4).

The status of specific tasks, deliverables and contributions are as shown below:

Tasks as per DoA (Status)	Deliverables completed as per DoA, associated with the task	Contributions
T1.1 (Completed)	**D1.1**	For the application characterisation within task T1.1 methodologies had been developed to create, in a systematic way, application-derived knowledge that can be used as input for the co-design process. Based on a questionnaire the same type of information has been

		collected for all applications within the SAGE application portfolio. Additional methods included a scaling analysis based on a performance modelling approach and a data retention time analysis. Both methods have been applied to a selected set of applications. Juelich lead the task with contributions from Diamond, KTH, DFKI and CCFE.
T1.2 (Completed)	**D1.2**	This task of provided a blue print of the SAGE infrastructure based on co-design inputs. The blue print provided a framework for the overall architecture of the SAGE platform feeding into the other technical Work Packages mainly on the software side. Seagate lead this task with contributions from CEA, Juelich and CCFE. D1.1 and D1.2 contributed to timely achievement of Milestone MS4 ("Co-Design activities complete").
T1.4 (Completed)	**D1.3**	The more "manual" application characterisation performed in T1.1 has been complemented by an automatised application characterisation using tools in task T1.4. The tools used were Allinea MAP, Allinea Performance Reports and Darshan. While the application characterisation as performed within T1.1 requires a significant level of expertise as well as efforts, performance analysis tools relieve the application developer from such burden. All tools have been used to explore a selected set of applications and results were documented in deliverable D1.3 ("Tools Based Application Characterisation"). While the tools cannot provide all information collected through the methods used in T1.1, it could be shown that they can provide very valuable information, which in parts was complementary to what has been obtained within task T1.1 and in other parts served as a cross-check of the results obtained in T1.1. Furthermore, the work of task T1.4 helped to improve the Allinea performance tools and its use on different platforms. Allinea lead this task with contributions from Juelich,KTH,CCFE and Diamond.
T1.3 (Ongoing)	-	For all applications a strategy for porting them to the SAGE prototype have been performed and implementation work was started. As the SAGE prototype had not been deployed at the time of writing this report and some of the required technologies developed within work packages WP2, WP3 and WP4 are still work-in-progress, only initial porting efforts could be performed, yet.

1.2.2 Work package 2

WP2 focussed on the design and architecture development of the software components related to Mero as required by the SAGE prototype, during the first 18 months of the project.

WP2 also worked on analysing potential emerging NVRAM technologies and their hardware aspects and provided methods for software emulation of these devices during this reporting period. The status of specific tasks, deliverables and contributions are as shown below:

Tasks as per DoA (Status)	**Deliverables completed as per DoA, associated with the task**	**Contributions**
T2.2 (a), (b), (c), (d), (e) ,(f) and (g) (Completed)	**D2.1**	Seagate contributed to the architecture and development of the following components described in D2.1, namely, HDF5 interfacing with Mero, pNFS and HSM Interfacing, Object I/O Scalability, Application Services for Data Centric Compute, enabling efficient multi-tiering, extreme resiliency for applications and system availability at scale. CEA contributed to the WP3 interfacing work in T2.2 (b) – HSM and pNFS interfacing D2.1 contributed to the timely achievement of Milestone MS5 ("Percipient Storage Concept and Supporting Architectures, Initial version defined").
T2.3 (a), (b), (c), (d), (e), (f) and (g) (Ongoing)	-	Seagate is involved in the implementation of all the components architected in T2.2.
T2.1.1 (Ongoing)	**D2.2**	ATOS contributed to the analysis of various potential NVRAM technologies available at this time as well as its hardware aspects (including the possibilities of emulation) D2.2 contributed to the timely achievement of Milestone MS09 ("NVRAM Devices, Techniques and Emulation defined").

1.2.3 Work package 3

WP3 covers four domains for the SAGE project:

1. Hierarchical Storage Management between the SAGE tiers: Extreme HSM Tools, named HSM
2. Performant legacy access to object store: pNFS services on object storage, named pNFS
3. Integrity service: Object Storage Integrity Services, named OSIC
4. Performance Profiling and debugging Tools for SAGE: Systemware and Software Tools, named SST

The status of specific tasks, deliverables and contributions are as shown below:

Tasks as per DoA (Status)	Deliverables completed as per DoA, associated with the task	Contributions
T 3.1	**D3.1, D3.5**	CEA and Seagate completed a design of HSM and how it will use certain features of Mero such as "FDMI" and "Pool Selection". CEA and Seagate also co-designed the "composite layout" API needed by HSM. This is described in D3.1. A HSM Proof Of Concept has been realized by CEA and is described in D3.5.
T3.2	**D3.3**	CEA has done the design of pNFS and has implemented a prototype of nfs-Ganesha over Mero object store. CEA created a new open source project, named libkvns, which implements a Posix name space over a Key/Value Store. Libkvns is used in nfs-ganesha to interface over Mero KVS. pNFS design is described in D3.3. D3.1 and D3.3 contributed to the timely achievement of the milestone MS06 ("Architecture for Storage Services and System Software Defined").
T3.3	**D3.4**	Seagate has done the OSIC architecture design and how to implement it over Mero. It has performed studies to evaluate design choices. Seagate has started a Proof Of Concept of OSIC. This work is described in D3.4
T3.4 (a)(b)	**D3.2, D3.6**	Allinea made a high level design for implementing SAGE related functions and features for debugging and profiling tools. This is described in D3.2. Later they were focused on low-level design. The development is described in D3.6. KTH used the Allinea deliverables against iPIC3D to see how they can implement the application over Mero.

1.2.4 Work package 4

WP4 worked on the preparation of a programming model, a runtime and visualization system, and a data analytics framework for the SAGE prototype.

Tasks as per DoA (Status)	**Deliverables completed as per DoA, associated with the task**	**Contributions**
T4.1 (a),(b)	**D4.1**	KTH led the work on parallel I/O programming models. This work resulted in the preparation of D4.1 on the state of the art and gap analysis in MPI I/O. A first implementation of PGAS I/O has been completed as a library atop MPI and in MPICH. The first performance results have been collected and presented in two conference papers (one accepted at CCGrid and one under review at Euro-Par).
T4.2	**D4.3**	Juelich led the work on the development of a runtime system. This work resulted in the preparation of D4.3 presenting the run-time architecture. The implementation of the runtime system together with the first tests is an on-going effort.
T4.3	**D4.5**	Juelich led the work on the development of visualization support at runtime system. This work resulted in the preparation of D4.5 presenting the design specification of the virtual memory hierarchy for visualization.
T4.4	**D4.2, D4.4**	DFKI led the work on the development of the Flink data analytics framework to run on the SAGE prototype using Mero. The Flink parts that have to be modified to exploit the SAGE prototype have been analysed and presented in D4.2. D4.4 extended D4.2 focusing on the design of Flink sub-components to use SAGE for data access and output. This task is now focusing on porting Flink to use the Mero object storage via a Java interface of the Clovis API.

1.2.5 Work package 5

WP5 was primarily focussed on the development of the SAGE prototype hardware during the first 18 months of the project. The hardware was designed working closely with the use cases in WP1 as part of the co-design exercise. The list of materials for the prototype was then prepared and the actual hardware was assembled in Seagate premises. The prototype hardware was brought to a working status with the necessary system software installation. The prototype is shipped and reassembled and connected to the Juelich computing cluster during the time of this periodic reporting window.

WP5 has also been working on the performance evaluation strategy based on feedback from the commission in the last review.

Tasks as per DoA (Status)	**Deliverables completed as per DoA, associated with the task**	**Contributions**
T5.1 (a) **(Completed)**	**D5.1 and D5.2**	Seagate lead the hardware design of the prototype, including contribution to the hardware components as described in the list of Materials as described in D5.2. ATOS contributed to the hardware design of the prototype including contribution to the hardware components as described in the list of materials in D5.2
T5.1 (b) (Completed)	**D5.3 and D5.4**	Seagate lead the prototype build and bring-up in Seagate premises following hardware assembly as described in D5.3. Seagate also contributed to the hardware integration plan in Juelich as described in D5.4 ATOS contributed to the Build and bring-up of SAGE hardware prototype in Seagate premises. Jueilich contributed to the Hardware integration plan as described in D5.4 D5.1, D5.2 and D5.3 contributed to the timely achievement of the milestone MS10 (“System functional”).
T5.4 (On-going)	**None**	STFC is working on performance evaluation strategy for the SAGE prototype and is leading the performance evaluation working group following recommendations from the commissions on having ambitious performance targets as an internal goal. More details are described in section 4.

1.2.6 Work package 6

WP6 focussed on dissemination, innovation and exploitation, and collaboration activities (which includes standardization). The following are the status of the various tasks and associated deliverables.

Tasks as per DoA [Status]	Deliverables completed as per DoA, associated with the task	Contributions
T6.1	**D6.1, D6.2, D6.3, D6.4, D6.5**	All partners contributed to the dissemination task (including developing the strategy). The details of dissemination activities and the overall strategy is available in D6.2, D6.4 and D6.5. D6.1 covers the website and D6.3 covers the initial dissemination materials generated.
T6.2	**D6.2, D6.4, D6.5**	All partners contributed to developing the initial strategy for collaboration. Seagate contributed to the external collaboration activities. All partners contributed to the internal collaboration task. The latest details on collaboration can be obtained from D6.5.
T6.3	**D6.2, D6.4, D6.5**	All partners contributed to developing the initial strategy for exploitation. Seagate contributed to exploitation activities. D6.1, D6.2 and D6.3 contributed to the timely achievement of the milestone MS01 ("Initial Dissemination Activities Complete"). D6.4 and D6.5 contributed to the timely achievement of the milestone MS06 ("First Review report and plans complete Innovation, Dissemination and Exploitation").

1.2.7 Work package 7

WP7 focussed on overall project management. The following are the status of the various tasks.

Tasks as per DoA (Status)	Deliverables completed as per DoA, associated with the task	Contributions
T7.1	**D7.1, D7.2**	All partners contributed to this task with Seagate leading

(Ongoing)	**(Referencing the periodic reports)**	the task. D7.1 indicates the management processes followed. Please note that D7.2 references the periodic reports to prevent duplication of information across multiple reports (after our consultation with the Project Officer).
T7.2 (Ongoing)	**D7.1, D7.2 (Referencing the periodic reports)**	All partners contributed to this task with Seagate leading the task. D7.1 indicates the Quality Control process for deliverables followed.
T7.3 (Ongoing)	**D7.1, D7.2 (Referencing the periodic reports)**	All partners contributed to this task with Seagate leading the task. D7.1 describes the risk management process and the risk update for Month 9. The periodic reporting activity provides updates on the latest risks as of Month 18.
D7.4 (Ongoing)	**D7.1, D7.2 (Referencing the periodic reports)**	All partners contributed to this task with Seagate leading the task. Financial management procedures and the status of finances for Month 9 was described in D7.1. The periodic reporting process provides updates on the state of finances as of Month 18.
D7.5 (Ongoing)	**D7.1, D7.2 (Referencing the periodic reports)**	All partners contributed to this task with Seagate leading the task. Communication procedures and the status for Month 9 was described in D7.1. We have continued to follow the internal and external communication procedures effectively without any issues.

1.3 Impact

All the information described in section 2.1 of the DoA continue to the highly relevant. Our deliverables D6.4 and D6.5 describe how we are tracking against it.

In D6.4 (section 2.2) we introduced further evolving European objectives with regards to the European Cloud Initiative. More details can be obtained in that deliverable.

At the time of writing this report there have some more activity in the European HPC ecosystem, namely the setting up of EuroHPC[2] for Exascale/Pre-Exascale procurements and further information on the development of the European Low Power Processor technology for Exascale, on which one of the Exascale systems will be based on. We are closely following these developments to identify the relevant target audience for the outcomes of SAGE in this regard. We are very keen for SAGE technology outcomes to make it to the EsDs in the shorter term and pre-exascale/Exascale systems in the longer term.

2. Update of the plan for exploitation and dissemination of result (if applicable)

We presented the latest dissemination and exploitation strategy and results in D6.5. There are no further updates beyond this.

3. Update of the data management plan (if applicable)

Not Applicable

4. Follow-up of recommendations and comments from previous review(s) (if applicable)

[Recommendation 1] – As part of the validation process, we recommend the consortium to include criteria to evaluate the performance of the proposed storage system. These criteria will serve to monitor internally the progress of the project and will not entail additional obligations towards the European Commission.

A Performance Evaluation Working Group has been set up with representation from a subset of the consortium of partners. A review of traditional benchmarking tools highlights a lack of readily available tools for benchmarking SAGE considering its unique characteristics namely, a new extended object based storage interface (Clovis) in-built compute and a deep I/O hierarchy with many device types. SAGE at this stage is thus a new architecture, which has been implemented as a prototype with a size that is small compared to typical high-end storage systems. Furthermore, performance tuning of such storage systems requires significant efforts. Considering these issues, the group believes that running benchmarks tuned for current HPC architectures on the SAGE prototype would not provide a fair comparison in the performance of the systems. However, the group has identified methods for useful tracking of the performance of the prototype and its development through the course of the project. Methods for tracking the performance which have been identified are:

1. Adapting Allinea tools for analysing application performance in SAGE; this is ongoing work being carried out by Allinea in partnership with the consortium members and the output from this work will be used to track performance using SAGE for a set of the project's use cases as the project progresses.

[2] https://ec.europa.eu/digital-single-market/en/news/eu-ministers-commit-digitising-europe-high-performance-computing-power

2. CEA is porting IOR, which is a standard I/O benchmark, for Clovis, which will enable performance evaluation of Clovis/Mero over the course of the project, setting a baseline and helping to track performance, including demonstrating performance with SAGE storage tiering enabled/disabled to demonstrate the benefits of tiering.

[Recommendation 2] – As part of the dissemination activities concerning future work, we recommend the consortium to better advertise the SAGE project to the scientific community, for instance through the publication of scientific papers. In order to facilitate user engagement, we also suggest exploring the possibility to give access to the prototype to users outside the consortium.

We are actively working on increasing the number of scientific publications. Between the time of the Month 9 (M9) review and Month 18, we have 6 accepted publications. We did not have anything in M9.

Further, taking the recommendation of the commission, we decided to have an overarching technical paper that succinctly covers all the key technical concepts of SAGE in one place. The paper was rigorously worked on by the consortium, including a very thorough Quality Review process from Seagate. The white paper is now openly available from the Sage website and has been very useful in discussions with the scientific community as well as potential Exploitation targets/Customers of SAGE components such as Mero.

We are discussing the usage of the SAGE prototype outside of the consortium starting with the weather and climate community (through EsiWACE CoE) and have had discussions on this with the EsiWACE partners. More details, as well as more potential users will be identified and worked out during the course of the project.

[Recommendation 3] - A better connection of this project with other HPC projects is expected.

We had very little collaboration activity at the time of the M9 review. We have initiated and are working on identifying opportunities for collaboration with a few other FP7, FETHPC and CoE projects.

We are exploring the use of the Clovis API with the NextGenIO project and understanding the open source software tools developed NextGenIO. We are keen to understand more about the 3DxPoint technology from NextGenIO. However this has been affected due to the very strong NDA EPCC has with Intel, which prevents EPCC from deeper information sharing in this area. We are nevertheless continuing to discuss. We have had 1 joint workshop at SC'16 and are organising a joint workshop again at the European HPC Summit in May 2017.

Since we don't have many weather & climate use cases in SAGE, we have had many discussions with members of the EsiWACE CoE including DKRZ, Met Office and ECMWF. We have presented SAGE to them separately and are in the process of exploring the possibility of the weather and climate use cases providing co-design inputs into Mero which is the core technology in SAGE. We are also developing a common I/O middleware (termed "Common Metadata Format" or CMF) that can be utilized by the EsiWACE project as well. As mentioned earlier we are also exploring if the weather and climate community can use the SAGE prototype when it is operational in Juelich.

We have also provided inputs to the DEEP-ER FP7 project on the opportunities for I/O and storage beyond DEEP-ER with technologies coming out of SAGE. We have had a joint BoF with them in ISC'16.

We have also planned many upcoming collaboration activities with HPC initiatives as described in Section 5.2 of D6.5.

[Recommendation 4] - It is a minor recommendation but we suggest the consortium to consolidate all the quarterly management reports to produce only a single one for the entire reporting period.

We do not have quarterly management reports and further, we are also in the process of eliminating redundant information, and are streamlining our communication with the commission and reviewers. Our D7.2 Management Report, for example, now entirely points to the periodic reports – after consultation with the project officer.

5. Deviations from Annex 1 and Annex 2 (if applicable)

5.1 Tasks

There have not been any deviation in objectives for any of the Work Package tasks.

5.2 Use of resources

The table below shows the usage of resources for partners until now (Actuals) and the budget person months over the whole project.

The first 18 months focussed on co-design, and concept and architecture of the SAGE components. The next 18 months will require significant additional effort considering that we are now in the implementation and testing phase. The overall efforts are hence back end loaded towards the latter half of the project.

Actual Person months	SEA	JUELICH	KTH	CEA	Allinea	Bull	UK Atomic Ener	DFKI	Diamond	S&T Facilities	Total
Work Package											
1	4.64	22.06	1.5	1.32	2.7	2	8.2	2	3.72	2.34	50.48
2	34.37	0	0	1.85	0	17.22	0				53.44
3	5.59	0	0	40.84	9.04	0	0	1			56.47
4	0.34	12.51	23.3	0	0	0	0	22.9			59.05
5	3.78	0		0.8	0.09	20.26	1		0.11	3.53	29.57
6	4.87	0		0.6	0.27	0.8	1.5	1.5	0.01		9.55
7	14.5	0.56		0.8	0.26	1	1	2	0.23	0.06	20.41
Total	68.09	35.13	24.8	46.21	12.36	41.28	11.7	29.4	4.07	5.93	278.974366

Budget Person months	SEA	JUELICH	KTH	CEA	Allinea	Bull	UK Atomic Ener	DFKI	Diamond	S&T Facilities	Total
Work Package											
1	7	23	4	2	4	2	10	2	9	2	65
2	180		7			47					234
3	20		3	70	57			3			153
4	10	29	36					51			126
5	33	15	9	6	6	18	9	11	10	17	134
6	12	3	3	3	3	3	3	3	3	3	39
7	36	2	2	2	2	2	2	2	2	2	54
Total	298	72	64	83	72	72	24	72	24	24	805

As can be seen, significant deviations exist for Diamond, Allinea and STFC. The following are the reasons:

Diamond's efforts were mainly in application characterization (WP1) until this stage, which did not require significant effort from them. Diamond has now started developing their use case (SAVU) to work with Mero/Object stores (as part of ongoing Task 1.3) and they will be doing benchmarking around this in WP5. This will take significantly more effort going forward.

Most of Allinea's/ARM's efforts are in WP3, which was waiting on SAGE platform work (mainly Mero and the associated interface readiness to access telemetry records) to reach sufficient maturity for them to continue. WP3 work from Allinea/ARM will take up significantly more effort in the second period.

Most of STFC's efforts are in "Task 5.4: Percipient Storage in SAGE: System strengths, weaknesses and comparisons with existing technologies". This work relies heavily on Mero and the SAGE platform to reach sufficient level of maturity to do the comparison studies. This work will significantly increase in the second half of the project.

5.2.1 Unforeseen subcontracting (if applicable)

Not Applicable. The subcontracting work has been planned from the start of the project and we are executing according to that plan.

For the sake of completeness, we include the activities performed by **planned** subcontracting below, initiated by Seagate through subcontractor EURL Tweag:

Tasks (as per DoA) performed by Subcontract	**Deliverables (as per DoA) involved in by Subcontract**	**Contributions**
T2.2(c) Research and development Object I/O Scalability concepts for Percipient Storage (completed)	**D2.1 Percipent storage: First version of the concept and architecture report**	Defined the architecture for the "Containers" component as needed by D2.1
T2.2 (d) Architectural concepts for enabling	**D2.1 Percipient storage: First version**	Defined the architecture for the "function shipping/in-compute-storage" component as needed by D2.1

application services for data centric compute (completed)	**of the concept and architecture report**	
T2.2(g) Architectural concepts for enabling system availability at scale (completed)	**D2.1 Percipient storage: First version of the concept and architecture report**	Defined the architecture for the "High Availability" component as needed by D2.1
T2.3(c) Validation of object I/O scalability in Percipient Storage (in progress)	**None**	Currently working on the validation/implementation of Containers for percipient storage.
T2.3 (d) Validation of enabling data centric computing for applications (in progress)	**None**	Currently working on the validation/implementation of data centric computing working closely with the use cases
T2.3 (f) Validation of extreme resiliency for applications and data sets	**None**	Currently working on models for the feature as part of validation/implementation

5.2.2 Unforeseen use of in kind contribution from third party against payment or free of charges (if applicable)

Not Applicable

HISTORY OF CHANGES		
VERSION	PUBLICATION DATE	CHANGE
0.1	01/03/2017	First draft inputs
0.2	10/03/2017	Work Package Descriptions
0.3	31/03/2017	Impacts
1.0	18/04/2017	Incorporation of partner inputs