EuroHPC-01-2019

IO-SEA

IO – Software for Exascale Architectures

Grant Agreement Number: 955811

D7.1

Periodic Report Year 1

Final

Version:	1.0
Author(s):	M. Gilliot (CEA)
Contributor(s):	P. Deniel (CEA)
Date:	31.03.2022

Project and Deliverable Information Sheet

IO-SEA Project	**Project ref. No.:**	955811
	Project Title:	IO – Software for Exascale Architectures
	Project Web Site:	https://www.iosea-project.eu/
	Deliverable ID:	D7.1
	Deliverable Nature:	Report
	Deliverable Level: PU *	**Contractual Date of Delivery:** 31 / March / 2022
		Actual Date of Delivery: 31 / March / 2022
	EC Project Officer:	Daniel Opalka

*– The dissemination levels are indicated as follows: **PU** - Public, **PP** - Restricted to other participants (including the Commissions Services), **RE** - Restricted to a group specified by the consortium (including the Commission Services), **CO** - Confidential, only for members of the consortium (including the Commission Services).

Document Control Sheet

Document	**Title:** Periodic Report Year 1	
	ID: D7.1	
	Version: 1.0	**Status:** Final
	Available at: https://www.iosea-project.eu/	
	Software Tool: LaTeX	
	File(s): IO_SEA_D7_1.pdf	
Authorship	**Written by:**	M. Gilliot (CEA)
	Contributors:	P. Deniel (CEA)
	Reviewed by:	E. Gregory (FZ Jülich) J. Hawkes (ECMWF)
	Approved by:	Exec Board/WP7 Core Group

Document Status Sheet

Version	Date	Status	Comments
0.1	01.03.2022	outline approved	completed
0.2	04.03.2022	draft with diffrent contributions	completed
0.3	11.03.2022	draft for internal review	completed
0.4	14.03.2022	internal review comments integrated	completed
0.5	21.03.2022	post-review edits complete	completed
1.0	31.03.2022	final draft ready for EU submission	completed

Document Keywords

Keywords:	IO-SEA, HPC, Exascale, Software

Contents

List of Figures

List of Tables

Executive Summary

This deliverable reports on the activities performed as part of WP7 related to the "Management and the Scientific Coordination" of the IO-SEA project. This report covers the first 12 months of the project, that is, the period from April 1, 2021 until March 31, 2022.

All internal management bodies have been set up, such as the Project Management Office (also referred to as the WP7 Core Group), the Executive Board and the Project Board. These instances are operational and are cornerstones of our internal communication and running the project on a daily basis. In addition, tools for facilitate the collaboration have been set up (mailing lists, shared code repositories, shared workspaces).

The Executive Board, which brings together all WP leaders, the Scientific Coordinator and the Project Management Office is a central element: it supervises the review process for all deliverables, regularly reviews the risks that could hinder the smooth execution of the project, and tracks the use of resources to avoid financial imbalances.

Jointly with the Scientific Coordinator, the Project Management Office also serves as a liaison element towards the other R&I projects funded under the same EuroHPC call[1]. It acts as contact point for all matters related to the Supplementary Grant Agreement and all other joint activities. It also takes care the financial management, the risk management, the quality assurance procedures, and the data management.

[1] EuroHPC-01-2019

1 Introduction

This deliverable reports on the activities performed as part of WP7 related to the "Management and the Scientific Coordination" of the IO-SEA project. This report covers the first 12 months of the project, that is, the period from April 1, 2021 until March 31, 2022. The activities of WP7 are divided into six tasks, for each of which we will present the work performed.

- The management and communication activities are covered by the Task 7.1 and 7.5, respectively. These activities (as presented in Chapter 2) cover all aspects for running the project on a day-to-day basis and ensure the progress of the work within the project with respect to DoA. These tasks also cover the project-internal communication (tools and methods), as well as the communication with partners in the European HPC ecosystem, including for example other R&I projects funded under the same EuroHPC call. And, in particular, this includes our collaboration with the DEEP-SEA and RED-SEA projects, the communication with the EuroHPC Project Officer, and with organisations such as ETP4HPC and PRACE ([1], [2], [3].
- Task 7.2 focuses on controlling the quality of the project's output - in particular its deliverables. In Chapter 3 we will present details of IO-SEA's quality control process and report on the status of the deliverables.
- The third task in WP7 is about risks to the project, their probabilities and possible mitigation actions. The process and the details for updating and reviewing this list of risks regularly is described in Chapter 4.
- In Chapter 5 we present details on the financial status of the project. It compares the actual versus planned effort, as well as the actual versus planned expenditures.
- The work related to the common tasks for the complementary grants (Task 7.6) is presented in Chapter 6.

This deliverable closes with a short summary and describes the next steps related to the management and the coordination of the IO-SEA project (Chapter 7).

2 Management and Communication

The management of IO-SEA is based on the “management by exception” approach with delegated responsibility. In the context of IO-SEA, this means that the decision-taking body — the Project Board — is only solicited when the project deviates significantly from the planned activities and tasks, encounters difficulties in fulfilling the project’s objectives, or if significant deviation with respect to effort, cost or budget occur. On a daily basis, the management of the project is taken care of by the Executive Board and the Project Management Office (also referred to as WP7 Core Group). The management bodies (overview given in Figure 1[1]) and their activities in the first 12 months of the project are detailed below.

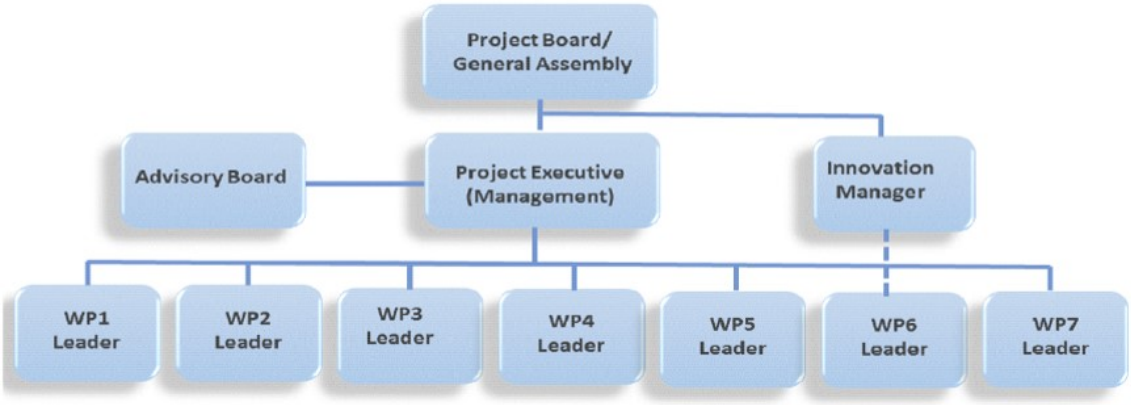

Figure 1: IO-SEA Management Structure

2.1 Project Board

The Project Board is the ultimate authority to run the project. The Board directs the project’s strategic orientation and safeguards the interests of all of its participants. The Board assesses the project’s performance and takes appropriate measures if the performance is found to be lacking. It comprises one representative from each partner and meets every six months[2] (cf. Annex 1). The first meetings were held on April, 17, 2021 and on October 6, 2021. As outlined in the DoA, at least one face-to-face meeting is to be held per year. Given the current pandemic however, both meetings took place via video-conference. During these meetings, the members of the Project Board are updated on the technical achievements within the project and on organisational and administrative aspects of the project and approved the composition of the Executive Board.

[1] The role of the Innovation Manager will not be detailed in this deliverable. Details on innovation management are given in deliverable D6.1.

[2] Technically speaking, ParTec is a Linked Third Party to FZJ. ParTec being our only Third Party, the project parnters decided to consider ParTec from an organisational point of view as a partner of the project.

Partner	Project Board Representative	Project Board Proxy
CEA	Jacques-Charles Lafoucrière	Jean-Philippe Nominé
Atos-Bull	Cornel Crisan	Jean-Robert Bacou
FZJ	Wolfgang Frings	Dominik Gottwald
ECMWF	James Hawkes	Tiago Quinto
Seagate	Sai Narasimhamurthy	Ganesan Umanesanjan
ICHEC	Buket Benek Gursoy	Venkatesh Kannan
IT4I	Jan Martinovic	Martin Golasowski
KTH	Artur Podobas	Stefano Markidis
CEITEC	Jirka Novacec	
JGU	André Brinkmann	
ParTec	Hugo Falter	Ina Schmitz

Table 1: Members of the Project Board (as of March 2022)

The next meeting of the Project Board is scheduled for April, 5, 2022 (from 14:00 - 16:00) and will also be held remotely. The main agenda items for this meeting are (1) update on the project (technical achievements, collaborations, outreach, training, ...), (2) feedback from the informal review, and (3) a pending decision regarding the changes on the Executive Board.

2.2 Project Executive (or "Executive Board")

The Project Executive comprises the Scientific Coordinator, the Project Management Office and the work package leaders. The list of work package leaders appointed by the Project Board is given in Figure 2[3]. The Executive Board meets currently once a month online. Face-to-face meetings can be planned in the future, for example jointly with an All-hands-meeting. The main tasks of the Executive Board are to oversee all quality-assurance procedures and guidelines, to ensure the progress towards the technical objectives of the project, and an effective cooperation amongst the project partners, with the EU, our partner projects and the wider HPC ecosystem.

2.3 Project Management Office

The Project Management Office assures the day-to-day activities for running this project, under the control of the Executive Board. The WP7 Core Group consists of

- Philippe Deniel (CEA, Scientific Coordinator)
- Maike Gilliot (CEA, WP7 Leader)

[3] Due to changes on Atos side, it is planned that Philippe Couvée takes over the leadership of WP2, whereas Céline Lemarinier will lead the WP3. This will be submitted for approval by the Project Board on April 5, 2022.

Role	Main	Proxy
Scientific Project Coordinator	Philippe Deniel (CEA)	
Project Management Team	Maike Gilliot (CEA)	
	Cornel Crisan (Atos)	
	Jean-Robert Bacou (Atos)	
Lead WP1	Eric Gregory (FZJ)	
Lead WP2	Alexandre Lopez (Atos)	Philippe Couvée (Atos)
Lead WP3	Philippe Couvée (Atos)	Alexandre Lopez (Atos)
Lead WP4	Thomas Leibovici (CEA)	Philippe Deniel (CEA)
Lead WP5	James Hawkes (ECMWF)	Tiago Quinto (ECMWF)
Lead WP6	Mark Wiggins (Seagate)	Sai Narasimhamurthy (Seagate)
Lead WP7	Maike Gilliot (CEA)	Cornel Crisan (Atos)

Figure 2: Members of the Executive Board

- Jean-Robert Bacou (Atos)
- Cornel Crisan (Atos)

The WP7 Core Groups meets every 3-4 weeks and is involved in all tasks of WP7. Its main activities are related to the work of the project management bodies. To this end, it prepares the meetings and the minutes of the different bodies (Project Board, Executive Board, Advisory Board), including our All-hands-meetings and the regular Cross-WP-sessions, which focus on the topics of common interest between the different work packages.

It is also in charge of the quality assurance procedures, controls that the internal deadlines of the review process are met, and manages the appointment of the reviewers. As described in Chapter 5, the WP7 Core Group also collects the financial statements of the partners and prepares the analysis for the Executive Board. It also provides logistic support for all IO-SEA related mailing lists and shared working spaces.

Moreover, the WP7 Core Group is also the liaising element towards the Project Officer and to our partner projects. In particular, the members of the WP7 Core Group take part in all joint All-SEA-coordinator calls[4].

2.4 Advisory Board

IO-SEA's Advisory Board is composed of people from other EuroHPC projects and of qualified person in the HPC data management domain. It gives advice on the orientations and the implementation of the IO-SEA project, requirements, solutions and outside perspectives throughout the project. This input is considered carefully, aiming at maximising the effectiveness and the impact of the IO-SEA project. The IO-SEA Advisory Board is composed by:

[4]This regular call brings together the coordinators from the DEEP-SEA, the RED-SEA and the IO-SEA project in order to define joint activities.

- **Estela Suarez**, a senior scientist and leader of the DEEP series of EU-funded projects she has driven the development of the Cluster-Booster and the Modular Supercomputing Architectures.
- **Johann Lombardi**, Senior Principal Engineer in the Extreme Scale Architecture & Development Division at Intel, and leading the development of the Distributed Asynchronous Object Store (DAOS).
- **Jalil Boukhobza**, Professor at the ENSTA-Bretagne, with focus on storage system design, performance evaluation and energy optimisation.
- **Robert Ross**, a Senior Computer Scientist at Argonne National Laboratory and the Director of the DOE SciDAC RAPIDS Institute for Computer Science, Data, and Artificial Intelligence.
- **Soraya Zertal** is Associate Professor in Computing at University of Paris Saclay-UVSQ and member of Li-PaRAD lab, with research interests in data storage systems and performance evaluation.

A first meeting of the Advisory Board took place on-line on September 25, 2021. According to the Advisory Board, the paradigm of Ephemeral Services seems a promising path to tackle exascale challenges. They may act as kind of proxy. This intermediate level between client and storage server will bring flexibility for managing storage workload, which will tend to become more and more diversified. It also discussed the role of AI. The board members encouraged us to look into the work done to understand IO behaviors (for example based on Darshan or Lustre log files). AI is not only of interest for the data placement policy, but will also allow to optimise sizing of the ephemeral services. Regarding outreach and dissemination, the members of the board encouraged the IO-SEA consortium to have a well-identified and common location for our publications and the software developed in IO-SEA. They also underlined the importance of openness and transparency regarding the IP strategy to facilitate the uptake of open-source solutions.

The next meeting with the IO-SEA Advisory Board is planned for May 2022. By then, all technical work packages will have closed the first design phase. At this point, the members of the Advisory Board can challenge the chosen approach and the suggested solutions. This will provide valuable input before starting the implementation phase of the IO-SEA software solution.

2.5 Cross-WP sessions

These regular cross work package sessions were not planned for in the proposal, but were set up in a ad-hoc fashion as the need emerged. Probably amplified by the lack of face-to-face meetings, during this architecture-definition phase the need for regular exchanges between the work packages on topics of joint interest became clear.

This effort was kicked off by a full day workshop and has now become a regular meeting (every other Wednesday from 16:00 - 17:00). Depending on the need from the project, the Executive Board decides on the topic for each session in advance and appoints for each session a leader (in most cases one of the WP leader or the Scientific Coordinator takes this role). The full list of sessions and their topic is given in Table 2. Today, these cross-WP sessions fulfil two roles. First, they allow discussion and exchange on topics of common interest between the WPs. For example, concepts

such as ”namespace” or ”datasets” are key to the project, and the cross-WP sessions have proven a good format for the discussions on these topics.

Date	Title	Leading WPs
21/06/2021	Full-day workshop for kicking off the cross-wp-sessions	All WPs
07/07/2021	Concept of "namespace" for IO-SEA	Scientific Coordinator
21/07/2021	Concept of "dataset" for IO-SEA	Scientific Coordinator
15/09/2021	IO-SEA’ VM	Scientific Coordinator
29/09/2021	HSM API	WP4
13/10/2021	Application Instrumentation	WP3
27/10/2021	Implications of namespace and dataset definition for WP4	WP4
10/11/2021	Datasets and namespaces: common definition	Scientific Coordinator
24/11/2021	IO-SEA’s Void prototype	Scientific Coordinator
08/12/2021	Data sets and namespaces: the LQCD use case as example	WP1, WP2, WP3
05/01/2022	Synchronisation on upcoming deliverables	Scientific Coordinator
02/02/2022	DASI presentation	WP5
16/02/2022	Void prototype	Scientific Coordinator
02/03/2022	Recommendation Systems	WP3, WP4
16/03/2022	API for datasets	WP3, WP4, WP5
30/03/2022	Ganesha Request Handler	WP4

Table 2: Overview of cross-WP sessions.

Second, over time these meeting have also established themselves as "mini-all-hands-meetings". They serve as short, but regular synchronisation points for the Project Management Team with all the project participants. This simplifies communication and has proven very useful for example in preparing the joint SEA checkpoint with EuroHPC (February, 9, 2022). Thus, these cross-WP sessions also serve as a complement to the All-hands meeting that take place regulary every 6 months.

3 Quality Control

Within IO-SEA, the quality management focuses on the quality assessment of the work and the deliverables produced within the project. The process is overseen by the WP7 Core Group, but relies on all partners, all work package leaders, and all project participants. To this end, the WP7 Core Group:

- Established a review process for the deliverables and the milestones
- Coordinates the interaction with the project-internal reviewers
- Ensures that the process is respected by all project participants
- Controls the final deliverables before submission to EuroHPC

The WP7 Core Group is also in charge of overseeing the **appointment of reviewers**. Beginning of each year (M01, M12 and M24), the WP7 Core Group identifies potential reviewers for each deliverable in the 12 months to come. This suggestion takes into account possible conflicts of interests and also the technical background of the potential reviewer. Moreover, the WP7 Core Group seeks to share the work load equally amongst the partners. These suggestions are checked by the Executive Board and amended upon its recommendations. Once the list is approved, the potential reviewers are contacted in order to get their approval for performing the review.

Approximately six weeks before the due date for the deliverable, the WP7 Core Group contacts the reviewers again to make sure that they will still be available for the review. At this point in time, changes can occur, and other reviewers might have to be identified. This process avoids — as much as possible — last minute hassles in identifying reviewers.

The list of possible reviewers for the deliverables due after M12 will be presented to the Executive Board in April 2022.

The **review process for deliverables** foresees a timeline with intermediate steps in order to assure that the final deliverables are submitted on time. Its main steps are:

1. The outline (table of content) of the document is to be provided 2 months before the due date for submission.
2. 15 working days before the due date: a first complete version has to be submitted into IO-SEA's shared storage space and made available to the project-internal reviewers.
3. 10 working days before the due date: the project-internal reviewers have completed their reviews and submitted their comments to the author, who prepares a new version of the document, taking into account the reviewer's comments and suggestions. The document annotated by the reviewers has been uploaded to the shared storage space.
4. Five working days before due date: the Scientific Coordinator and the WP7 Core Group review the document. The main author takes the comments and suggestions into account and prepares the final version of the document.
5. One to two working days before the due date: the WP7 Core Group submits the document to EuroHPC via the EC portal.

Review process for milestones: For the IO-SEA project, the choice was made to reduce the number of milestones since most of them occur at the same point in time. Moreover, most of the software components developed in the technical work packages have strong dependencies. Therefore, we have opted for common milestones between the technical work packages WP1-WP5. In the same way, WP6 (on dissemination and outreach)and WP7 (on management) share three common yearly milestones.

IO-SEA Milestone Assessment[1]

Milestone Name and Milestone Number	
Description from Project Proposal	To do
Expected date:	DD/MM/YYYY
Delivery date	DD/MM/YYYY
Assessed during the IO-SEA Executive Board Meeting on	DD/MM/YYYY

Assessment

Criteria	Milestone Assessment	
	Assessment[2]	Comments
Have the MILESTONE objectives been achieved?		
Have the documents and deliverables related to the MILESTONE been reviewed according the the IO-SEA internal rules and submitted to the EC in time?		
Have the milestone results been presented to the IO-SEA consortium?		

Figure 3: Template for assessing the project milestones

Regarding the review of the milestones, we have opted for a light-weight process, presented in the milestone assessment document (cf. also Figure 3), focusing on three main elements for assessing whether a milestone is achieved:

- Have the milestone's objectives been achieved?
- Have the documents and deliverables related to the milestone been reviewed according the the IO-SEA internal rules and submitted to the EC in time?
- Have the milestone results been presented to the IO-SEA consortium?

The possible assessments are "YES", "NO" or "PARTIALLY". The second column allows the Executive Board to provide some more details or comments on its assessment. All milestone assessments will be annexed to the interim report (M18) and the final report (M36).

The table in Figure 4 gives an overview of the milestones and deliverables submitted or currently under preparation.

#	Deliverable Title	WP	Leader	Type	Due Date
D1.1	Application co-design input	WP1	FZJ	Report	4
D7.4	Collaboration plan with definition of common objectives and activities including milestones	WP7	CEA	Report	6
D7.5	Data Management Plan	WP7	CEA	Report	6
MS1	Intial Co-Desgin input collected	WP1	FZJ	Milestone	6
D1.2	Application use cases and traces	WP1	FZJ	Report	9
MS2	IO-SEA Benchmark suites defined	WP1	FZJ	Milestone	9
D2.1	Ephemeral Data Access Environment: Concept and architecture	WP2	Bull	Report	10
D1.3	Application I/O strategy	WP1	FZJ	Report	12 → 15
D3.1	Instrumentation and Monitoring: Concept and architecture	WP3	Bull	Report	12
D4.1	Hierarchical Storage Management Feature: Concept and architecture	WP4	CEA	Report	12
D5.1	First version of the Data Access and Storage Interface	WP5	ECMWF	Report	12
D6.1	Dissemination, Exploitation and Training Report and Future Plans Year 1	WP6	Seagate	Report	12
D7.1	Periodic report Year 1	WP7	CEA	ORDP	12
MS6	Description of the concepts	All	CEA	Milestone	12
MS9	Year 1 activities and reporting completed	All	CEA	Milestone	12

Figure 4: Overview of Deliverables and Milestones (M1-M12)

As of now, all deliverables and milestones have been submitted on time. The deliverable D1.3 has been postponed from M12 to M15 in order to account for the dependency of D1.3 on the deliverables D3.1, D4.1 and D5.1. These deliverables are due in M12 and close (jointly with D2.1, submitted in M10) the architecture design phase. The technical work packages describe the architecture of their software bricks and the interaction between the different bits and pieces. This input is needed by WP1 for describing how the use cases will interact with the IO-SEA storage solution.

4 Risk Management

Within the IO-SEA project, the Executive Board oversees risk management. The starting point of the risk assessment is the list of risks with their associated risk-mitigation measures as presented in the DoA. Three times a year, the Executive Board re-assesses the risks identified so far and adds newly-identified risks to the list. The full list and the assessment of the Executive Board in M4, M8 and M12 is a living document [4].

Currently, four risks are tracked more closely:

Under-spending: The project partners submit regular updates on the used resources: the number of PMs for each Work Package, as well as the cost associated to this effort. Within IO-SEA, it was planned to collect data twice a year. The M6 report however showed strong discrepancies between the actual and the expected use of resources. As a mitigation measure, an additional reporting on the used resources in M9 has been requested from the partners as to trace in more detail the evolution of the use of resources. Details of the use of resources are given in Chapter 5.

Overall, there is some severe under-spending in some organisations, but this seems due to organisation-internal issues (such as hiring problems) and — at this point in time — do not affect the overall progress of the project. The work package leaders did not identify any kind of missing engagement from any of the partners and report on excellent collaboration with all partners in all work packages. The WP7 Core Group will continue to monitor this issue closely.

Ongoing COVID pandemic: With respect to project management and to dissemination, the past two years taught us a lot on tools and formats to better engage with each other online. And it also showed the limits of on-line events compared to face-to-face events. From a project management point of view, the first year is crucial for building some team spirit, for allowing the partners to connect and to get jointly engaged into the project. Joint face-to-face meetings are excellent means to foster collaboration. Due to the effort of the work package leaders the IO-SEA team managed to engage all partners. In particular, cross work package interaction was strongly supported by the work package leaders with additional meetings (either on a regular basis or on-demand).

A second topic strongly impacted by the restrictions are the dissemination and outreach activities. We have learned that online booths, workshops and Birds-of-a-feather sessions at conferences gather overall less people and less attention than physical events. Also in training activities it is more difficult to build interactive sessions and to foster exchange between the participants. For the first year of the IO-SEA project, we mainly focused on on-line events (cf. Deliverable D6.1). For 2022, more physical dissemination and training activities are scheduled, starting with ISC22 [5] end of May 2022 and Forum Teratec [6] in June 2022.

Other activities however, as for example the one-day All-SEA workshop (with plenary parts and with topic specific break-out sessions) did not take place and have been postponed into the second year of the project.

Complementary Grant Agreement: The initial deadline for having the Complementary Grant Agreement signed (Milestone MS12) was in M6. However, very soon after some first discussions amongst the project coordinators and with our Project Officer it became clear that such an agreement would take more time (cf. Chapter 6). The progress on this agreement is tracked closely and is

planned to be completed in time by the end of March 2022. The scientific coordinator, along with the WP7 Core Group and CEA's legal department act as main contact points for all comments or questions from the IO-SEA partners and assure that all input from the IO-SEA side is provided in time to the legal department of FZJ, who is coordinating this effort on behalf of all ten projects funded under this call EuroHPC-01-2019.

Lack of a common IO-SEA system: The DoA did not foresee any physical HPC infrastructure for integrating and testing the IO-SEA software stack, and to deploy it in order to benchmark the performance of our use cases. This was identified during the first risk review performed by the Executive Board in M4 as possible risk: It could hinder seamless integration into a solid and coherent IO-SEA storage solution and also be a hurdle for assessing the performance of the IO-SEA solution. Over the past months, the IO-SEA partners have jointly developed solutions.

In a first step, an image for deployment in a Virtual Machine (VM) has been configured, so that all project participants can execute their developments in a shared and common environment. Based on this common VM, the work packages are currently developing a "void prototype". This void prototype does not provide any functionality or any services, but it makes sure that the interfaces between the work packages are well defined and compatible. In particular, the void prototype helps in identifying potential conflicts in the software dependencies required by different components from different Work Packages. Moreover, some IO-SEA partners are able to provide access to HPC systems via in-kind contributions.

- An OpenStack cluster at IT4I, with dedicated data nodes where the image of the IO-SEA-VM can be launched, containing the basic services from IO-SEA stack MOTR, Phobos, and nfs-ganesha). This system will mainly be used by the IO-SEA partners for testing and integrating the IO-SEA software ([7], [8]).The previously described virtual machine will here be used as a "seed" for building an IO-SEA test and development system on the OpenStack cluster.
- Access to the DEEP-SEA system in Jülich, mainly for benchmarking the IO-SEA solution and for integration purposes with the developments of the DEEP-SEA project. To this end, the hardware of the SAGE2 prototype would be re-purposed as IO-SEA data nodes on the DEEP system [9].

Whereas the OpenStack cluster at IT4I is operational, the details for accessing the DEEP system still need some last validation steps before being fully set up.

5 Financial Management

The WP7 Core Group also traces the actual effort declared on the project, thus identifying deviations at an early stage in order to alert and to suggest — if appropriate — counter-measures. To do so, the actual declared efforts are compared to an expected consumption of PMs. For the “expected consumption”, a linear distribution of the effort over time is assumed. For example, after M9 (25% of the project duration elapsed), 25% of the resources are expected to be consumed). This very basic assumption gives some indication on the effort put into the project.

5.1 Efforts per work package as of M9

The project partners submit regular updates on the used resources: the number of PMs for each work package, as well as the cost associated to this effort. Within IO-SEA, it was planned to collect this data twice a year. The M6 report however showed strong discrepancies between the actual and the expected use of resources, thus leading to some additional reporting in M9 as to trace in more detail the evolution of the use of resources.

WP	Expected consumption (PM)	Consumed (PM)	Deviation (%)
WP1	41,25	26,04	63,13
WP2	34,75	25,51	73,40
WP3	34,00	25,64	75,40
WP4	39,25	24,94	63,54
WP5	39,50	13,31	33,69
WP6	8,75	4,71	53,79
WP7	9,00	12,55	139,43
	206,50	**132,69**	**64,26**

Figure 5: Effort - PM per WP (as of M9).

In Figure 6 third column shows actual “consumed” effort (a total of 132,69 PM) declared by the partners, where the second column (“expected consumption”) indicates the number of PMs which would be due in M9 if the resources would be used in a linear way throughout the project life time. The column “deviation” expresses the discrepancy. More precisely, it indicates the % of the expected resources that were actually consumed. Overall, as of M9, under-spending is observed. In total, 64% of the effort expected (compared to a linear distribution) has been used. For WP7, which relates to project management, some over-spending can be observed. This can be linked to the additional effort in setting up all bodies and the infrastructure in the first months of the project. In WP 6, which is about communication and outreach, we observe the strong under-spending. This can be explained by the ongoing Covid pandemic with less on-site events and less exploitation-related activities in these

first months of the project. The development of the DASI, which is at the heart of WP5, is expected to ramp up quickly now with the implementation phase starting now in M13. Figure 6 gives a graphical overview of the data presented in Figure 6

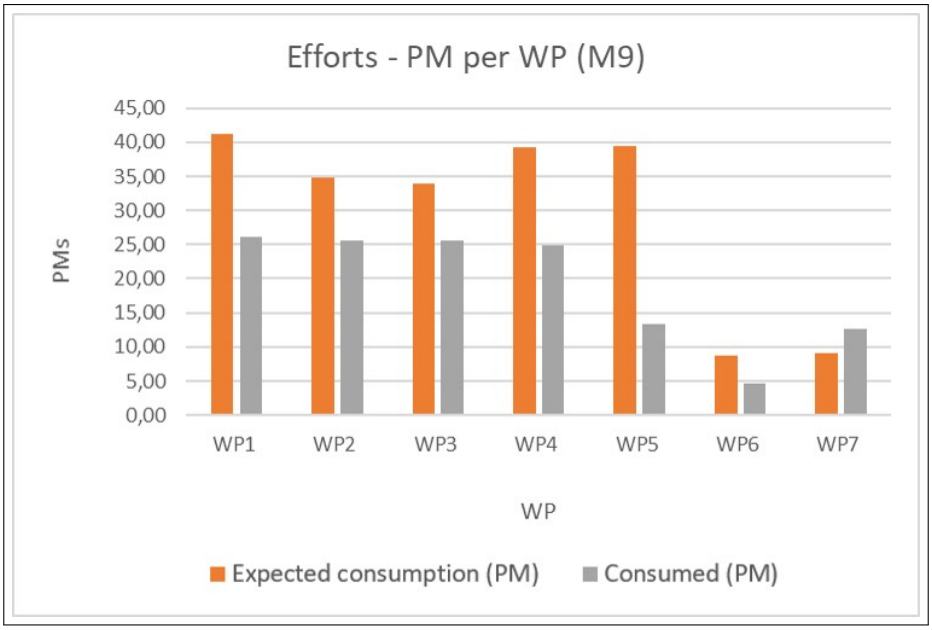

Figure 6: Effort - PM per WP (as of M9)

5.2 Efforts declared per partner as of M9

When looking at the efforts declared by the partners as indicated in Figure 7, we observe strong discrepancies: Overall, all partners are currently below the expected consumption. The strongest deviation can be observed for ECMWF, ParTec, and KTH. ECMWF and KTH report on hiring issues, in particular the problem of finding appropriate candidates.

These figures have also been discussed at the Executive Board with the work package leaders. As of now, they did not identify any kind of missing engagement from any of the partners and report on excellent collaboration with all partners in all work packages. This may indicate that the observed discrepancies are due to some organisation-internal issues (such as hiring). It also shows that the assumption of a linear distribution over time might not be fully adequate.

Partner	Expected consumption (PM)	Consumed (PM)	Deviation (%)
CEA	32,50	28,30	87,08
Atos-Bull	41,25	28,80	69,82
FZJ	28,00	22,20	79,29
ECMWF	16,25	4,91	30,22
Seagate	24,00	13,54	56,42
ICHEC	9,00	6,14	68,22
IT4I	20,50	13,70	66,83
KTH	9,00	1,10	12,22
CEITEC	5,00	4,35	87,00
JGU	12,00	6,25	52,08
ParTec	9,00	3,40	37,78
SUM	**206,50**	**132,69**	**64,26**

Figure 7: Effort - PM per partner (as of M9)

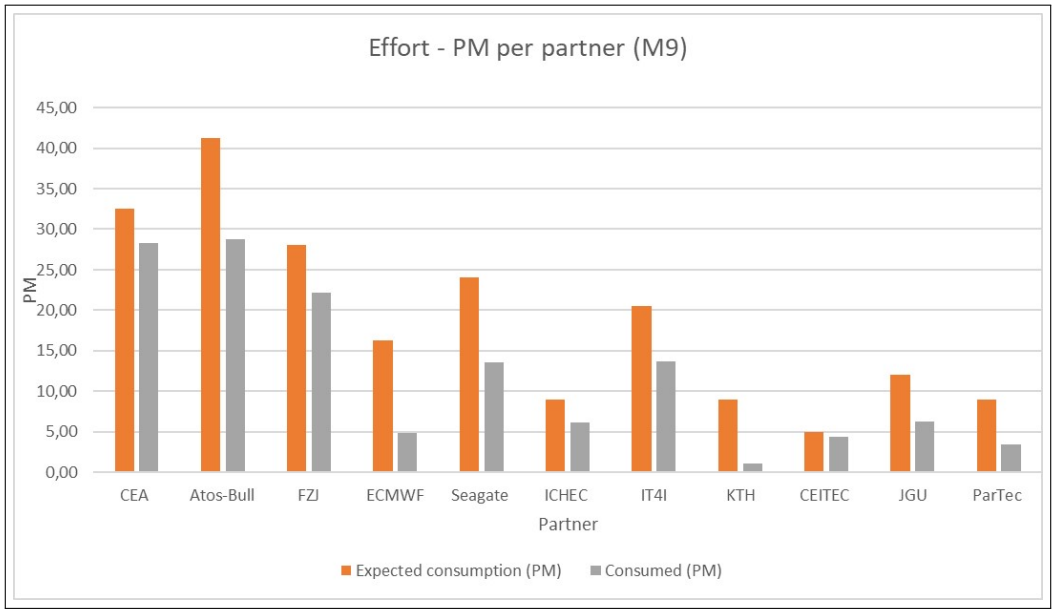

Figure 8: Effort - PM per partner (as of M9)

6 Common task for complementary grants

The initial deadline for having the **Complementary Grant Agreement** signed (Milestone MS12) was in M6. However, very soon after some first discussions amongst the project coordinators and with our Project Officer it became clear that such an agreement would take more time. FZJ (coordinating the DEEP project) kindly accepted to take the lead in this matter:

- In September 2021, a first draft (prepared by FZJ's legal department) has been shared among all project coordinators and their legal departments.
- This resulted in December 2021 with a version approved by all the coordinating parties.
- All project coordinators shared this consolidated version with their project partners and clarified open questions from their consortium members.
- Currently, a second iteration including all organisations of all concerned projects has been closed on March 8, 2022.
- The signature phase has started and is expected to be finalised by March 31, 2022.

This progress on this Complementary Grant Agreement is closely tracked. The scientific coordinator, along with the WP7 Core Group and the legal department, act as main contact point for all comments or questions from the IO-SEA partners.

As depicted in D7.4, IO-SEA some points of interest for **collaboration and joint activities with DEEP-SEA and IO-SEA** have been identified, and has led to an active collaboration between the three SEA projects (DEEP-SEA, IO-SEA and RED-SEA). This collaboration has different parts. First, there are natural synergies and natural cross-project relations, which happen on an operational level on a day-to-day basis. In addition, some concerted actions take place, driven by the Coordinators and the Project Management Teams from all three projects. The common system allowing full integration is a third part of this collaboration.

- **Natural synergies between the projects on a day-to-day basis:** Topics such as benchmarking, continuous integration, or monitoring and relevant to all three projects. The different teams working on these topics collaborate today closely, discuss common tools and exchange on best practices, leading also to topic-specific calls between participants of the different projects. Information on technical workshops are shared by default with the partners of the other SEA-projects, so that a natural emulsion takes place in our daily activities.
- **Concerted actions:** Examples of concerted activities are the joint communication and outreach activities, including the joint Birds-of-a-feather session at the ISC22 conference in Hamburg (in June 2022), the joint booth at ISC, joint effort on social media, and the joint web page. Also part of these activities is the All-SEA-workshop, which took place on December 3, 2021, focusing on the following two topics:
  - The **exploitation of traces** for analysing the system behaviour plays an important role in all three projects. It showed that a multitude of tools are used within the three projects, providing different kind of traces to different kind of "users". In some cases, the “users” of the trace analysis may be system administrators (eager to monitor/improve the overall system), or application developers (in order to identify bottlenecks in their applications),

or end users (wishing to understand the status of their computing jobs). This discussion will be continued in order to understand which traces generated by one of the SEA components can be used by one of the other SEA components for its own needs. For example, some information on congestion at network level may allow for a “smarter” allocation of data nodes.

  - A second mechanism to share information between the storage, the compute and the network parts of a system could be based on the **exploitation of hints**. In this context, hints are to be understood as all kind of metadata, tags or annotations to objects. Whereas the three projects agree that hints can be an excellent vehicle to convey information, the definition of the “objects” they relate to are not fully clear yet. This raised the more general question of how the hardware-centric outcome of RED-SEA interfaces with the software layers provided by DEEP-SEA and IO-SEA project. This will be the topic of a workshop planned for May 2022.

- **Common prototype:** All three SEA projects have their “in-house system” for testing and integrating their developments and their benchmarking activities (the Dibona system for the RED-SEA team and the nodes IT4I cluster for the IO-SEA project). The partners of the three consortia are currently investigation the option to use the DEEP-system for integrating the developments of all three projects, thus leading to a fully integrated prototype system, which would allow to demonstrate the compatibility of the developments in the different projects. The organisational and the technical conditions for such an integrated system are currently under investigation.

Moreover, IO-SEA is also engaging in collaboration with the **ADMIRE** project. Items for collaboration were identified with the ADMIRE project as part of D7.4. The idea is to extend this collaboration and to include the **MAELSTROM project**. Those collaborations have lead and will - in an even more intense way - lead to remote one-day workshops where experts from different projects share ideas, concepts on items identified as common points of interests. Whereas during the first year, some effort was necessary to get the IO-SEA project and the All-SEA collaboration up and running, the second year will pay more attention to collaboration in a wider sens.

7 Summary

This deliverable reports on the activities performed as part of WP7 related to the “Management and the Scientific Coordination” of the IO-SEA project. This report covers the first 12 months of the project from April 1, 2021 until March 31, 2022.

All internal management bodies, such as the Project Management Office, the Executive Board and the Project Board, have been set up. These instances are operational for running the project on a daily basis. In addition, tools to facilitate the collaboration have been set up (mailing lists, shared code repositories, shared work spaces). The Executive Board, which brings together all WP leaders, the Scientific Coordinator and the Project Management Office is a central element: it supervises the review process for all deliverables, regularly reviews the risks that could hinder the smooth execution of the project, and tracks the use of resources to avoid financial imbalances.

Jointly with the Scientific Coordinator, the Project Management Office serves also as liaison element towards the other R&I projects funded under the same EuroHPC call. It acts as contact point for all matters related to the Supplementary Grant Agreement and all other joint activities. It also takes care of the Data Management, in particular it oversees the release of project related data to the wider community.

The focus of the past months has been to put all bodies on track and to set up the right tools and mechanisms to (1) steer the project in an efficient and effective way, (2) support collaboration between the work packages, and (3) to establish a sound structure for the collaboration with RED-SEA and DEEP-SEA.

From April 2022 on, the project will enter its second phase with the implementation of the defined architecture. In this second phase, first tangible results are expected soon. For the WP7, the focus might shift to supporting dissemination, outreach, and training. Another focus will be on collaboration. With the (hopefully now) better travel conditions, joint activities such as an All-SEA workshop can now be planned for. Moreover, collaboration with a wider circle than “just” our sister projects will be strengthened. Another important element of the second year for the WP7 Core Group is related to data management and to support the technical WPs in their efforts for releasing data openly.

List of Acronyms and Abbreviations

A

AI	Artificial Intelligence.

D

DAOS	Distributed Asynchronous Object Store, developed by Intel.
Darshan	Characterization tool provides snapshot of application I/O behavior, developed at the Argonne National Lab.
DASI	Data Access and Storage Interface developed in Work Package 5.
DoA	Description of Action: legally binding description of the work performed in the IO-SEA project.
DOE	US Departement of Energy.

E

ECMWF	European Centre for Medium-Range Weather Forecasts.
ETP4HPC	The European Technology Platform for High Performance Computing is an industry-led think-tank promoting European HPC research and innovation [3].
EU	European Union.
EuroHPC	The European High Performance Computing Joint Undertaking (EuroHPC JU) is a legal and funding entity, allowing the European Union and the EuroHPC JU participating countries to coordinate their efforts and pool their resources for HPC related R-and-I and investments [1].

F

FZJ	Forschungszentrum Jülich, in Jülich, Germany, is one of the largest research centres in Europe and a member of the Helmholtz Association.

H

HPC	High-Performance Computing.

I

IT4I	IT4Innovations National Supercomputing Centre at VSB Technical University of Ostrava, Czech Republic.

K

KTH KTH Royal Institute of Technology, Stockholm, Sweden.

L

Li-PaRAD lab Laboratoire d'informatique parallélisme réseaux algorithme distribués.

LQCD Lattice quantum-chromodynamics is a numerical framework for calculating physical properties of hadrons, composite particles composed of quarks.

Lustre Lustre is an open-source parallel file system.

M

MOTR Motr is a distributed object storage system developed by Seagate, open source.

N

nfs-ganesha A NFS server with HPC features, developed by CEA, open source.

O

OpenStack OpenStack is a cloud operating system that controls large pools of compute, storage, and networking resources throughout a datacenter [8].

P

ParTec ParTec is one of the leading SMEs in the HPC domain in Europe.

Phobos Phobos stands for "Parallel Heterogeneous Object Store", developed by CEA, open source [7].

PRACE The mission of PRACE (Partnership for Advanced Computing in Europe) is to enable high-impact scientific discovery and engineering research and development across all disciplines to enhance European competitiveness for the benefit of society [2].

R

R&I Research and Innovation.

RAPIDS A SciDAC Institute for Resource and Application Productivity through computation, Information, and Data Science.

S

SAGE2 R-and-I project funded under H2020, whose outcome is used and developed further in IO-SEA.

SciDAC The US Department of Energy's (DOE's) Scientific Discovery Through Advanced Computing (SciDAC) program.

U

UVSQ Université de Versailles-Saint-Quentin-en-Yvelines.

V

VM Virtual Machine.

W

WP Work Package.

Bibliography

[1] EuroHPC JU web site. https://eurohpc-ju.europa.eu/.

[2] PRACE web site. https://prace-ri.eu/.

[3] ETP4HPC web site. https://www.etp4hpc.eu/.

[4] IO-SEA's risk table living document. https://tinyurl.com/3u6dynxc.

[5] ISC web site. https://www.isc-hpc.com/.

[6] Teratec web site. http://www.teratec.eu/gb/forum/index.html.

[7] Phobos code repository. https://github.com/cea-hpc/phobos.

[8] OpenStack web site. https://www.openstack.org/software/.

[9] SAGE2. https://cordis.europa.eu/project/id/800999/fr.