Figure 1: Integration of PCP pilots.

The PCP pilot systems JULIA (Cray) and JURON (IBM and NVIDIA) have been integrated into the High Performance Analytics & Computing Platform (see T7.6.3, section 7.7.3.2). They are already used by HBP scientists for their research projects.

Project Number:	720270	Project Title:	Human Brain Project SGA1

Document Title:	SP7 High-Performance Analytics and Computing Platform - Results for SGA1 Period 1
Document Filename:	D7.6.2 (D42.2 D24 SGA1 M12) ACCEPTED 180709.docx
Deliverable Number:	SGA1 D7.6.2
Deliverable Type:	Report
Work Package(s):	WPs 7.1, 7.2, 7.3, 7.4, 7.5, 7.6
Dissemination Level:	PU
Planned Delivery Date:	SGA1 M12 / 31 Mar 2017
Actual Delivery Date:	SGA1 M14 / 23 May 2017 RESUBMISSION 28 Mar 2018; ACCEPTED 09 Jul 2018

Authors:	Thomas LIPPERT, JUELICH (P20), SP Leader Thomas SCHULTHESS, ETHZ (P18), SP Leader Anna LÜHRS, JUELICH (P20), SP Manager
Compiling Editors:	Anna LÜHRS, JUELICH (P20), SP Manager
Contributors:	Julita CORBALÁN, BSC (P5) Benjamin CUMMING, ETHZ (P18) Timo DICKSCHEID, JUELICH (P20) Markus DIESMANN, JUELICH (P20) Stefan EILEMANN, EPFL (P1) Giovanni ERBACCI, CINECA (P13) Andreas FROMMER, BUW (P7) Colin MCMURTRIE, ETHZ (P18) Cristian MEZZANOTTE, ETHZ (P18) Ralph NIEDERBERGER, JUELICH (P20) Boris ORTH, JUELICH (P20) Dirk PLEITER, JUELICH (P20) Hans Ekkehard PLESSER, NMBU (P44) Raül SIRVENT, BSC (P5) Bernd SCHULLER, JUELICH (P20) Dennis TERHORST, JUELICH (P20) David VICENTE, BSC (P5) Benjamin WEYERS, RWTH (P46) Gabriel WITTUM, UFRA (P35) Felix WOLF, TUDA (P85)

SciTechCoord Review:	EPFL (P1): Jeff MULLER, Martin TELEFONT, Marie-Elisabeth COLIN UHEI (P47): Martina SCHMALHOLZ, Sabine SCHNEIDER
Editorial Review:	EPFL (P1): Guy WILLIS, Colin McKINNON, Annemieke MICHELS

Abstract:	This Deliverable summarises the progress made and results achieved by the High Performance Analytics & Computing Platform Subproject (SP7) in the first twelve months of SGA1.
Keywords:	High Performance Analytics and Computing Platform, High Performance Computing (HPC), Platform, research infrastructure, simulation technology, data-intensive supercomputing, interactive visualisation, dynamic resource management, progress update

Table of Contents

List of Figures

List of Tables

1. SP Leader's Overview

1.1 Key Personnel

Subproject Leader: Thomas LIPPERT (JUELICH)

Subproject Deputy Leader: Thomas SCHULTHESS (ETHZ)

Subproject Managers: Anna LÜHRS (JUELICH)

Boris ORTH (JUELICH)

Meredith PEYSER (JUELICH)

1.2 Progress

Work in SP7 in the first year of SGA1 went well and according to plan. SP7 actively contributed to the Data Planning and Implementation Team (DPIT; led by Thomas Lippert), that developed a new work plan for the Neuroinformatics Platform. Since then, SP5 and SP7 have been working very closely together in order to develop and enhance the basis of the research infrastructure for the Human Brain Project. The base infrastructure provided by SP7 and the SP5 services running on top are now well aligned and designed to work together seamlessly. This is a key improvement compared to the Ramp-Up Phase.

The Federated Data Pilot Project (FeDaPP) was started by ETHZ-CSCS, JUELICH-JSC and CINECA in the Ramp-Up Phase. In autumn 2016, this project delivered a working Proof of Concept of a federated data infrastructure, which is based on neuroscience use cases that were collected and analysed by T7.6.4. This will be the foundation for Fenix, a federated infrastructure for data and scalable compute services planned to be built by a consortium of five major European HPC and data centres. The infrastructure services supported by Fenix are foreseen to accommodate the inherent diversity of data sources and relevant processing workflows. This project is currently in the design phase; BSC and recently CEA joined the efforts as well so that now all of the five PRACE hosting members are involved in Fenix.

All SP7 Work Packages contributed to the overall goal of SP7 for SGA1, which is the design and operation of the High Performance Analytics and Computing (HPAC) Platform as well as the design, implementation and deployment of novel software capabilities and algorithms for brain simulation, dynamic resource management, data-intensive supercomputing and interactive visualisation. WP7.5 has extended the HPAC Platform with virtualisation capabilities based on Docker. This is an important feature required by other HBP Platforms such as SP10, that need to deploy their software services on scalable compute resources in order to be able to provide a software infrastructure that is scalable with respect to the number of concurrent users. WP7.5 has also started migrating the HBP Collaboratory services to ETHZ-CSCS and JUELICH-JSC infrastructure. This will allow a tighter integration of the Collaboratory with the HBP base infrastructure.

The four research and development Work Packages in SP7, WP7.1-WP7.4, made significant progress as well. Details are described in the respective chapters below.

WP7.1 advanced the simulation technology, in some cases even faster than expected. T7.1.1 provided a prototype simulation engine with rate-based model neurons to collaborators in SP2 and CDP4, who successfully tested these developments. T7.1.2 worked on the interfacing of different simulators using ug4 and MUSIC. T7.1.3 continued the development of a code generation mechanism from modelling languages, which disentangles the description of a neuroscientific model from the simulator code and programming language used. T7.1.4 worked on the compact specification of large-scale point neuron networks and the efficient generation of these networks on HPC architectures, which is important for efficiently simulating such large-scale networks on supercomputers. T7.5.5 successfully linked the development of the NEST simulator and related activities (WP7.1) with the community, e.g. by organising user and community workshops.

WP7.2 worked on different aspects in the context of data-intensive supercomputing, in particular the exploitation of hierarchical storage architectures (T7.2.1), the coupling of data analytics and visualisation to simulation (T7.2.2), as well as the co-allocation of compute and storage resources (T7.2.3). Additionally, this WP comprises a Task (T7.2.4) working on end-to-end implementations of data analytics workflows for human brain atlases, i.e. it provides a focal use case for this Work Package.

WP7.3 continued the development of visualisation components that was started in the Ramp-Up Phase, strongly driven by neuroscience use cases. Hence, a new focus on in situ visualisation has been established (T7.3.1). Progress was also made in the development of general visualisation methods for large neural networks (T7.3.2) and the multi-view paradigm (T7.3.3). WP7.3 is in close contact with the (neuro)scientists in the HBP who benefit from the technology developed in this WP.

WP7.4 achieved first important results related to dynamic resource management. T7.4.1 delivered an initial set of dynamic resource management capabilities and integrated malleability in Neuron, so that the resources allocated to this application can now be dynamically changed at runtime. First steps in the same direction for NEST have also been made. T7.4.2 started the integration of dynamic resource management into the scheduler SLURM. T7.4.3 developed first policies for the different levels of a holistic dynamic resource management system. T7.4.4 worked on a concept for the transfer of knowledge gained related to dynamic resource management to vendors of HPC technology for the time after the Pre-Commercial Procurement.

Last but not least, WP7.6 took care of the coordination and management of SP7 (T7.6.1 and T7.6.2), established and maintained contacts with HPAC Platform users, which is an important prerequisite for co-design and the validation of technology (T7.6.4), and developed a concept for the co-design knowledge management with relevant external stakeholders (T7.6.3). WP7.6 is the entity in SP7 that has the overview of all developments inside SP7. It makes sure that all these developments are well aligned, that all relevant information is properly and timely communicated within SP7, between SP7 and the HBP governance bodies, and between SP7 and the other SPs. This WP was also responsible for Deliverable D7.6.1, which was submitted in M6 as scheduled.

1.3 Deviations

Overall, work in SP7 is on track. Some partners had difficulties in the first months due to the delayed signature of SGA1. They were all able to keep the staff already hired in the Ramp-Up Phase, but the hiring of new personnel in some cases had to be delayed until after the signature. However, all delays caused by this will have been compensated by the end of SGA1, e.g. by shifting work between team members.

Most of the Milestones scheduled for the first year of SGA1 have been accomplished on time or even earlier than planned. A few Milestones have been or will be achieved with some delay, but in all cases this does not have a severe impact on other parts of SP7 or the HBP as whole. Detailed reports can be found in the respective sections below.

So far, we do not see the need to change any elements of the SP7 work plan as laid out in the Description of Action.

1.4 Impact of work done to date

The impact of the work done in SP7 on the HBP is high, as the HPAC Platform provides the base infrastructure for the HBP. Therefore, it is very important to let all developments in SP7 be driven by important, representative use cases of the HBP and the neuroscience community in general, and to adapt the plans where necessary. A good example of such a change is the migration of the Collaboratory services to ETHZ-CSCS and JUELICH-JSC infrastructure, which was not planned in the beginning. Also, the work done in the R&D Work

Packages has an impact on other parts of the HBP, as they deal with central topics like simulator technology and visualisation.

The infrastructure building of the HBP is a unique effort and will thus likely have a large impact on the neuroscience communities, since it seems that no other larger brain initiative or project aims at building a research infrastructure at such scale for the entire community. These pioneering activities are challenging, but also promise to be beneficial for the world-wide neuroscience research.

1.5 Priorities for the remainder of the SGA1 phase

The SP7 Work Packages will continue the activities started in the first year of SGA1. It will be a challenge to finalise the already started planning of SGA2, while at the same time working on the tasks planned for SGA1. The SP7 management will try to keep as much planning overhead away from the other SP7 members as possible, so that they can concentrate on SGA1 activities.

The planning of Fenix, including the alignment with the SGA1 work plan and the integration into the SGA2 work plan, will continue, so that the implementation of key services at production quality can start as soon as possible. Since BSC and CEA have joined the Fenix planning, it has recently been agreed that CEA will also join the HPAC Platform in SGA2. The integration of their systems and services into the HPAC Platform will already start in the next months, so that, ideally, they will be fully integrated by the start of SGA2.

SP7 would like to intensify the collaboration with the other Platforms in order to let the HBP Platforms grow together into a uniform research infrastructure that transparently offers its services jointly to the users. These activities will become even stronger under SGA2. Also, SP7 will work even more closely with the users to support them in using the SP7 infrastructure and services on the one hand, and to validate the HPAC Platform against their requirements on the other hand. The SP5 and SP7 managers are in close contact to support the assessment of the users' requirements with respect to storage and storage-related services, since these activities are highly relevant for both SPs and thus synergies should be exploited.

We are confident that SP7 will have achieved all its goals for this project phase by the end of SGA1.

2. WP7.1 Simulation Technology

2.1.1 Key Personnel

Work Package Leader: Markus DIESMANN (JUELICH)

Deputy Work Package Leader: Hans Ekkehard PLESSER (NMBU)

2.2 WP Leader's Overview

The work on simulation technology for rate-based model neurons (T7.1.1) advanced faster than expected such that a prototype simulation engine could be delivered to SP2 partners in CDP4 already after eight months. The successful testing of the new framework in turn led to improvements that will enter the first release and generalisations that are planned to be advanced in SGA2. Thus, a full iteration cycle between technology development and neuroscience will be completed within a single SGA period. Another result of the interaction is the plan of a joint publication of SP2 and SP7 members on the reproduction of a historically important finding with modern tools, to practically enable neuroscientists to build upon this work. The multi-simulator integration (T7.1.2) bears many interesting aspects on technical level and the multi-scale integration on a neuroscience level. The code generation from modelling languages (T7.1.3) will on the one hand disentangle the neuroscientific models from the actual simulator code and improve the maintainability of the code. On the other hand, it provides an interface for modellers that is focused on the theoretical model content, rather than its expression in a programming language. Expressing the network structure and efficiently loading it on HPC architectures is handled by T7.1.4, which is important for large scale simulations of anatomically realistic networks.

An even more important positive aspect is that all WP7.1 Tasks received overwhelmingly positive feedback on presentations of their work in several workshops and tutorials (for example, the NEST user workshop in Karlsruhe and the NESTML community workshop in Jülich). This feedback and the number of people interested were highly motivating especially for the young researchers and this helped to form the WP7.1 participants into a team.

Simulator technology development in WP7.1 is tightly integrated with high-level user support provided by T7.5.5 "Simulator NEST as a Service" that serves as a central point for directing requests to the corresponding Tasks and thus acts as the external interface of WP7.1. This close collaboration ensures that technology development is driven by user requirements in SP2, SP3, SP4 and SP10 and that users can exploit new technologies as part of their neuroscientific research as soon as they become available. Integration is ensured through bi-weekly developer video conferences, online discussions and review on GitHub and overlap in personnel.

The main challenge during the first year of SGA1 was the reporting and communication required inside SP7 and between SPs for planning SGA2, while SGA1 is only half way in its two-year period, which was even more demanding than expected.

2.3 Priorities for the remainder of the phase

The priorities for the remainder of the phase are to publish the two already planned peer-reviewed papers, one on simulation technology for rate-based model neurons (T7.1.1) and the other one on technology for network creation with massively parallel threading (T7.1.4).

The results of the NESTML Workshop will be summarized in the report on neuronal modelling languages by T7.1.3 and work will mainly be focused on the abstracted model representations in the NEST code.

Task 7.1.2 will focus on the implementation of simulator-simulator interfaces and will work on the report on multi-scale challenges.

2.4 Milestones

Table 1: Milestones for WP7.1 Simulation Technology

MS No.	Milestone Name	Lead Partner	Task(s) involved	Expected Month	Achieved Month	Comments
MS7.1.1	Community contacts established	P20 JUELICH	All in WP7.1	M06	M06	This Milestone has been achieved as scheduled. All WP7.1 Tasks contributed to this Milestone: T7.1.1: Very fruitful contacts with SP2 partners on the use of T7.1.1 technology on reproduction of historically important findings with modern tools. T7.1.2: As projected, we have established contacts with project partners, especially T7.1.1, T7.1.4 and T6.4.3, either through site visits or telecommunication. Interaction will be deepened during the Simulation in Technology Workshop 2016, the HBP Summit 2016 and the NEST workshop 2016. T7.1.3: We established the first contact with NESTML, which is our developing modelling language for neuron and synapses, users in T7.1.3 during the Community Workshop (see below) in Aachen. We got comprehensive scientific feedback on the NESTML approach by publishing a paper (see below) and holding a presentation in a computer science conference in March 2016. The NESTML community works jointly on NESTML in an open-source repository on GitHub (https://github.com/nest/nestml). Interaction will be deepened during the "Code Generation from Model Description Languages II" Workshop in December 2016. T7.1.4: We have established contacts with users in SP3, especially T3.2.5, T3.3.3, T3.4.2, T4.2.1 and T4.5.1, either through site visits or videoconferences. Interaction will be deepened during the "Network Simulations with NEST" parallel session at the HBP Summit 2016. NESTML paper: https://emdesk.humanbrainproject.eu/shared/59b10b8c10bd9-ada84a6cb7dbcf2c98edd4458af0225d

						NESTML workshop: Website: http://www.fz-juelich.de/ias/jsc/EN/Expertise/SimLab/slns/news_events/2015/nestml-ws/_node.html Tutorial: https://emdesk.humanbrainproject.eu/shared/59b10b9d085a4-dd5f67cc19a17872f1d9443b7c01e4a2 Modelling: https://emdesk.humanbrainproject.eu/shared/59b10b957b9a2-732de3732b97a8a469b351cc51cc758e Workshop: https://emdesk.humanbrainproject.eu/shared/59b10b99330d3-7a06a68a890525dd4230e0c5daa1a9fc
MS7.1.2	Reports on new technology written	P20 JUELICH	All in WP7.1	M18		
MS7.1.3	Reference implementation in NEST version released	P20 JUELICH	All in WP7.1	M24		
MS7.1.4	Simulator-simulator interfaces available	P20 JUELICH	T7.1.2	M24		

2.5 T7.1.1 Integration methods for continuous-time population models

2.5.1 Key Personnel

Task Leader: Andreas FROMMER (BUW)

Other Researcher: Markus DIESMANN (JUELICH)

2.5.2 SGA1 DoA Goals

The goal of T7.1.1 is to research appropriate numerical integration techniques for continuous-time population methods and to implement highly parallelizable algorithms in the framework of spiking neuronal network simulation engines. The Task also makes sure that sufficient overlap exists with the simulation engine dedicated to the population level for cross-validation and to facilitate multi-scale simulations.

2.5.3 Progress summary

During the first year of SGA1, Task 7.1.1 created a prototype implementation of the previously evaluated and chosen numerical scheme to integrate stochastic differential equations, which describe the continuous-time population models. This implementation was tested for various neuroscientifically relevant network models.

Furthermore, Task 7.1.1 supported the partners involved in the Co-Design Project "Visuo-Motor Integration" (CDP4) in implementing the population model used in CDP4 (Grossberg, 1973, Stud. Appl. Math. 52 (3), 213).

2.5.4 Component Progress

2.5.4.1 Report on numerical techniques for stochastic equations of population dynamics

Description of Component

Report on numerical techniques to integrate stochastic equations of population dynamics in spiking network simulator.

CDPs to which the Component contributes: None

Progress on Component:

BUW and JUELICH worked together on an initial draft.

2.5.4.2 Continuous dynamics code in NEST

Description of Component

High parallel algorithm integrated in NEST that enables the interaction of population dynamics with spiking neurons in the same simulation engine

CDPs to which the Component contributes: CDP4

Progress on Component

During the first year of SGA1, T7.1.1 created a prototype implementation of the previously evaluated and chosen numerical scheme to integrate stochastic differential equations, which describe the continuous-time population models. This implementation was tested for various neuroscientifically relevant network models.

2.5.4.3 NEST – The Neural Simulation Tool

Description of Component

NEST - The Neural Simulation Tool is a highly scalable simulator for networks of point or few-compartment spiking neuron models. It includes multiple synaptic plasticity models, gap junctions, and the capacity to define complex network structures.

CDPs to which the Component contributes: CDP1-P3, CDP1-P4, CDP4

Use cases:

- CDP1-P3: A virtual imaging lab app
- CDP1-P4: A virtual behaviour lab app
- CDP2-UC-002 - Multi-scale validation
- comparative analysis of experimental and simulated data
- generic models and algorithms
- Generic Neuronal Network Simulation
- Manipulation experiments with humanoid robots and human avatars
- Mouse rehabilitation experiment in the Neurorobotics Platform
- SP6 - In silico microcircuit experimentation
- SP9-UC-006 Validation of neuromorphic results with respect to software simulations

Progress on Component

New features and fixes of bugs found by the community are continuously being included into the software. Coordination between the developers is ensured by a fortnightly open developer videoconference. The latest available release is NEST 2.12.0.

Links

- Information about the open NEST developer videoconferences: https://github.com/nest/nest-simulator/wiki/Open-NEST-Developer-Video-Conference
- NEST in the HPAC Platform Guidebook: https://hbp-hpc-platform.fz-juelich.de/?hbp_software=nest-the-neural-simulation-tool

2.5.4.4 Community contacts established related to NEST

For a comprehensive description of this Component, please see T7.5.5, section 6.9.3.1.

Very fruitful contacts with SP2 partners on the use of T7.1.1 technology for the reproduction of historically important findings with modern tools have been established.

2.6 T7.1.2 Integration methods for multi-simulator multi-scale simulations

2.6.1 Key Personnel

Task Leader: Gabriel WITTUM (UFRA)

Other Researcher: Jan HAMAEKERS (FG)

2.6.2 SGA1 DoA Goals

To accommodate the multi-scale structure of the brain, multi-scale approaches are necessary for modelling and computation. In order to bring these approaches efficiently to large-scale computing, they need to be integrated into a numerical multi-scale strategy. This allows for a full multi-scale computation where all relevant scales are computed in a coupled way. To set up this numerical multi-scale strategy, a careful multi-scale analysis has to be carried out. T7.1.2 explores how the HBP simulation engines MIND, NEST, NEURON, STEPS and potentially also engines at intermediate levels can be engaged in multi-scale simulation. The suitability of MUSIC to mediate the communication is investigated and appropriate interfaces are discussed with the engine developers.

2.6.3 Component Progress

2.6.3.1 Simulator-simulator interfaces

Description of Component

Interfaces should be created in order to couple different simulators, which use their own modelling paradigms. Point neuron simulators, especially NEST, should be coupled with cable equation simulators, like NEURON or NestMC, and these in turn with a fully resolved three-dimensional model.

CDPs to which the Component contributes: None

Progress on Component

In the multi-scale multi-simulator approach of Task T7.1.2, the highly resolved level of signal processing in three dimensions and the connection with spatial components like the extracellular potential will be accomplished with the HBP simulation engine UG4. NEST and NEURON are simulation engines on the level of point and cable neurons, respectively. We use MUSIC to bridge these different simulation engines to an integrated multi-simulator platform which takes into account all relevant scales.

The Task-internal Milestone (M12) for this period "the development of interface concepts between the simulators NEST, NEURON and UG4" has been accomplished. Therein UG4 describes the simulation platform, in which the three-dimensional model for the membrane potential propagation is realised. Since MUSIC is a communication platform between simulation engines and because connections between MUSIC and NEURON as well as NEST are already available, we concentrate on making it usable and on establishing a coupling of MUSIC with UG4. The mathematical analysis, which leads to a validation of the standard cable equation on the basis of the three-dimensional model, is one part of this component. A manuscript about this topic is already quite advanced and we will be able to submit it soon. Work in this component is mainly done by UFRA.

Links

- ug4 in the HPAC Platform Guidebook:

  https://hbp-hpc-platform.fz-juelich.de/?hbp_software=ug4

2.6.3.2 Community contacts established related to NEST

For a comprehensive description of this Component, please see T7.5.5, section 6.9.3.1.

T7.1.2 members visited the NEST user workshop in November 2016 in Karlsruhe and established first contacts with NEST, NestMC and MUSIC developers. This interaction is very fruitful. NEST is an essential part in the multi-simulator framework and thus this interaction is very important. Out of these active conversations, the Task will be able to determine the cornerstones for the development of a multi-simulator multi-scale simulator, whose development we aim to work on during SGA2. On the basis of these contacts, a new idea has emerged concerning a new modelling approach to bring the simulation of neuronal networks on a new level. In this respect, UFRA worked with FG on creating a model using machine learning algorithms in order to predict basic parameters (i.e. mean firing frequency) of an active neuronal network. This latter investigation has just started and uses NEST in order to create the training sets. T7.1.2 is keen to see first results, which will be reported on in greater detail in the next report.

2.6.3.3 Report on multi-scale challenges in the HBP simulator hierarchy

Description of Component

This Component examines the main HBP simulation engines and creates a plan for a framework to investigate the suitability of each simulation engine to represent the respective scales. Moreover, it attempts to develop a model on the scale of neuronal networks which describes the signal processing.

CDPs to which the Component contributes: None

Progress on Component

UFRA works on developing a tool for interactively creating networks of neurons in hoc-file formats. This tool will facilitate to put up benchmarks for the evaluation of the integration methods and the examination of each simulator's limits.

An additional element of the multi-scale approach is a model which describes network activity as a function of its basic parameters. In collaboration with the FG, this component tries to develop such a model using algorithms of Machine Learning and Deep Learning.

2.7 T7.1.3 Code generation for neuron and synapse models for software and neuromorphic backends

2.7.1 Key Personnel

Task Leader: Markus DIESMANN (JUELICH)

2.7.2 SGA1 DoA Goals

This Task works under the assumption that domain specific languages like NeuroML and PyNN for the description of complex neuronal networks are developed outside of this Task. T7.1.3 concentrates on a lower layer, where neuron and synapses models are formally described together with hints on appropriate solvers. The primary aims are to research concepts for the reduction of the code base of neuronal network simulators, by an automatic generation of C++ code of concrete neuron and synapse classes, and to create a framework where different mechanisms such as synapse types, plasticity, and adaptation can easily and reliably be assembled to create compatible neuron and synapse implementations for different simulators. This Task will review existing approaches described in literature and continuously interact with the community. The code generation will also be optimised for target architectures such as neuromorphic systems (SP9).

2.7.3 Component Progress

2.7.3.1 Community contacts established related to NEST

For a comprehensive description of this component, please see T7.5.5, section 6.9.3.1.

For the better dissemination of the NESTML approach, T7.1.3 presented its first results during poster sessions of the NEST User Workshop in Karlsruhe and the IAS Symposium in Jülich. During the NEST User Workshop, a NESTML tutorial talk with a practical exercise was also given.

2.7.3.2 Report on neuronal modelling languages and corresponding code generators

Description of Component

Report on neuronal modelling languages and corresponding code generators.

CDPs to which the Component contributes: None

Progress on Component

T7.1.3 organised and realised the "Code Generation from Model Description Languages II" workshop that took place on December 7-9, 2016 in Jülich. More than 30 participants from different institutions and countries attended. It is planned to publish a scientific contribution based on the experiences gained during the workshop in the next months.

Links

- Website of "Code Generation from Model Description Languages II" workshop: https://hbp-hpc-platform.fz-juelich.de/?tribe_events=code-generation-from-model-description-languages-ii

2.7.3.3 NEST code with abstracted neuron model representations

Description of Component

Automatic code generation integrated into NEST with concrete neuron and synapse classes and to create a framework where different mechanisms like synapse types, plasticity, and adaptation can easily and reliably be assembled to compatible neuron and synapse implementations. The code generation should also be optimisable for target architectures like neuromorphic hardware.

CDPs to which the Component contributes:

- CDP1-P3: A virtual imaging lab app
- CDP5

Progress on Component

JUELICH improved and stabilised the developed code generator for the NEST Modelling Language (NESTML). The validation of the generated code from neuron models has started, which are defined as NESTML artifacts. During the first year of SGA1, T7.1.3 made several releases of the NESTML tools which are publicly available on Github.

It is planned to complete the validation of the generated code in the next months. T7.1.3 will start integrating the NESTML tools with the NEST simulator environment.

Links

- NESTML repository on Github:

  https://github.com/nest/nestml
- LEMS Showcase:

  https://github.com/OpenSourceBrain/NESTShowcase

2.7.3.4 NEST the neural simulation tool

See section 2.5.4.3, related to T7.1.1.

2.8 T7.1.4 Massively parallel methods for network construction from rules and data

2.8.1 Key Personnel

Task Leader: Hans Ekkehard PLESSER (NMBU)

2.8.2 SGA1 DoA Goals

The goals of T7.1.4 are:

- Research concepts required to compactly specify structured and spatially extended large-scale networks of point neurons so that they are easy for the neuroscientist to review, and at the same time can efficiently be instantiated in a massively parallel fashion.
- Strategies, algorithms and data representations to minimise data exchange among parallel processes in the network construction phase.
- Review of existing approaches in the literature and the continuous interaction with the community are essential components of the Task.

2.8.3 Component Progress

The delayed signature of SGA1 led to delayed hiring of staff for this Task (start at the end of September instead of April), which led to delays in the progress of T7.1.4.

2.8.3.1 Rule- and data-based connectivity generation in NEST

Description of Component

Brain-scale network models exhibit spatial organisation on multiple levels with individual neurons potentially projecting to anywhere in the brain. These models will be built from descriptions combining connectome data with rules governing connectivity details depending on characteristics like cell type, the placement in layers and substructures, and neighborhood relations. The systematised connectivity generation code is integrated into NEST.

CDPs to which the Component contributes: CDP1-P4, CDP5

Progress on Component

T7.1.4 has drafted approaches for flexible connectivity specification and had a first round of discussions with the NEST user community at the NEST User Workshop 2016. The Task is also collaborating closely with T6.3.5 on improved massively parallel connectivity generation; this work has already resulted in a publication (Ippen et al, 2017). The systematic performance measurements done in that collaboration will inform design choices for further work in T7.1.4. Work is executed by NMBU in close collaboration with JUELICH.

2.8.3.2 Community contacts established related to NEST

For a comprehensive description of this component, please see T7.5.5, section 6.9.3.1.

Connectivity concepts were presented to and discussed with the user community inside and beyond the HBP during NEST Developer Video Conferences and at the NEST User Workshop 2016 by partner NMBU.

2.8.3.3 Report on neuronal modelling languages and corresponding code generators

See section 2.7.3.2, related to T7.3.1.

2.8.3.4 NEST the neural simulation tool

See section 2.5.4.3, related to T7.1.1.

3. WP7.2 Data-Intensive Supercomputing

3.1 Key Personnel

Work Package Leader: Dirk PLEITER (JUELICH)

3.2 WP Leader's Overview

A broad range of different aspects that emerge in the context of data-intensive supercomputing could be covered in this Work Package and progress has been made in each of the Tasks. The aspects addressed by this Work Package are the exploitation of hierarchical storage architectures (T7.2.1), the coupling of data analytics and visualisation to simulation (T7.2.2) as well as the co-allocation of compute and storage resources (T7.2.3). Additionally, this Work Package comprises a Task (T7.2.4) working on end-to-end implementations of data analytics workflows for human brain atlases, i.e. it provides a focal use case for this work package. As requirements from neuroscience applications in this context are still evolving, their detailed definition continues to be a major challenge. More efforts thus were needed to establish a common understanding by exploring use cases (in particular, the use case of Task T7.2.4) and features of emerging architectures, which are, e.g., realised in the HBP PCP pilot systems (RUP-WP7.1). The work focuses on three use cases: two are related to processing of high-resolution brain image data and one is related to scalable brain simulations.

3.3 Priorities for the remainder of the phase

Within the Task on hierarchical storage architectures (T7.2.1) the next step will be a comparative evaluation of different software stacks that allow exploitation of hierarchical storage architectures and in particular the use of high-performance storage devices based on new dense memory technologies. By exploiting the capabilities of the HBP PCP pilot systems, different approaches and technologies will be evaluated. Based on the results, suitable abstraction layers will be designed and implemented using the applications from T7.2.4 as main use case.

During this phase the Task on coupling of data analytics and visualisation to simulation (T7.2.2) focused on identifying the types of data objects that need to be exchanged between simulators and visualisation pipelines. The focus for the remaining part of SGA1 is on the expansion of this work to support a broader scope of simulators and to provide implementations that can leverage multiple interfaces to hierarchical storage architectures in collaboration with T7.2.1. Finally, software components needed for in situ data analytics and data extraction will be developed.

Based on a formal problem definition of the problem of co-allocation of compute and storage resources, the focus of T7.2.3 was on setting-up a simulator environment that allows for exploring different solution strategies. The input profiles should be as realistic as possible, although the architectures, for which the scheduler is designed, are not yet in production. The workloads will thus be modelled based on profiles collected on available HPAC systems. Based on this work, different scheduling algorithms will be implemented and evaluated.

The work on data analytics workflows for the human brain atlas (T7.2.4) so far addressed only a part of the larger workflow focusing on cell segmentation. In the remainder of this phase this part of the workflow will be integrated with the data store interfaces provided by T7.2.1 and furthermore expanded to the full workflow. The focus will be on preparing for significantly higher acquisition rates of approximately 1 TB/day.

The Work Package as a whole will work towards an assessment of performance and usability of the proposed solutions for real-life applications. Furthermore, requirements and challenges for integrating the solutions into the HPAC Platform as well as the ICEI infrastructure will be analysed and documented in collaboration with WP7.5 and the ICEI project.

3.4 Milestones

Table 2: Milestones for WP7.2 Data-intensive supercomputing

MS No.	Milestone Name	Lead Partner	Task(s) involved	Expected Month	Achieved Month	Comments
MS7.2.1	Identification of basic data types for coupling data analytics and visualisation to simulation	P20 JUELICH	T7.2.2	M06	M09	A delay in the definition of use cases had an impact on the identification of relevant data types. Therefore, the Milestone has been achieved with delay. The basic data types that have been identified for coupling data analytics and visualisation to simulation are described in the following I/O library documentation http://bluebrain.github.io/Brion-1.10/data.html , which is part of the following open source repository https://github.com/BlueBrain/Brion/tree/master/doc/datamodel.
MS7.2.2	User requirements analysis for hierarchical data store software	P20 JUELICH	T7.2.1	M10	M16	Analysis of use cases and the resulting requirements at lower architectural levels is taking longer than expected due to need of extrapolation of current user requirements. Milestone MS7.2.2 "User requirements analysis for hierarchical data store software" has been achieved in July 2017 (M16). The supporting documentation is available as an internal report: https://emdesk.humanbrainproject.eu/shared/5a83f3dc331eb-a2a91d037d5852bec3bec4cba39b69bc
MS7.2.3	Resource co-allocation strategies defined	P20 JUELICH	T7.2.3	M12	M14	This Milestone was postponed as evaluation of different simulators was prioritised. The milestone report is available through the following publicly accessible link: https://emdesk.humanbrainproject.eu/shared/59b10b906f7f7-2ecbbfa21b220f33b442154c4cc9054c
MS7.2.4	Independent prototype of data store sources for	P20 JUELICH	T7.2.1	M14		This milestone was delayed due to additional efforts needed for evaluation of the software stacks.

	interactive visualisation for selected data types					
MS7.2.5	Prototypical implementation of registration pipeline running on HPAC platform	P20 JUELICH	T7.2.4	M16	M16	We prototypically deployed and evaluated a complex image registration algorithm on the JURON pilot system, which is used for the 3D reconstruction of ultra-high resolution fibre architecture models of rodent and human brains based on 3D Polarized Light Imaging. The alignment uses free form deformations implemented with a cubic b-spline model, and considers an extremely large number (~1e10) of possible shifts of b-spline grid control points. The best deformation is found by computing the Mutual Information (MI) metric for each possible shift, efficiently exploiting JURON's modern P100 GPUs. The performance of the code is significantly better than on JURECA nodes, which offer two K80 GPUs each. We observed an overall speedup of 2.4 of the workflow compared to JURECA. The image registration workflow reads image data that are stored on JUELICH's GPFS system JUST, which will become part of the Fenix infrastructure. To this end, we have investigated to what point neuroscientists from an HBP-associated lab are able to upload and provide their data as input for such workflows, using only their central HBP accounts. We report the current status of such upload in an extended document (https://emdesk.humanbrainproject.eu/shared/59b10b80c7c40-c7665ca941034d040a8b8e98060dc13b).

3.5 T7.2.1 Hierarchical data stores for data analytics frameworks

3.5.1 Key Personnel

Task Leader: Dirk PLEITER (JUELICH)

Other Researcher: Yolanda BECERRA (BSC)

3.5.2 SGA1 DoA Goals

The goal of this Task is to enable the use of hierarchical storage infrastructures for asynchronous data analysis on online and offline extreme scale data volumes. Existing software components as well as newly developed software components will be used with focus on the use of high-performance storage-class storage devices, which are integrated into HPC architectures and allow for near-storage data analytics. BSC and JUELICH work together on this Task, see below for details.

3.5.3 Component Progress

Traditional storage technologies continue to feature impressive growth in terms of capacity. Compared to the increase of compute performance, speed of these storage devices is improving at a much slower pace. This problem can be mitigated by integrating new types of storage devices based on non-volatile memory. Due to the high costs in terms of a price over capacity metric, one is, however, forced to compromise on capacity. To meet both performance as well as capacity requirements, it is thus unavoidable to move towards hierarchical storage architectures. A wide range of solutions with different architectures are expected to emerge over the next years, ranging from already announced products like DDN's Infinite Memory Engine to prototypes realised within current H2020 FETHPC projects like SAGE. In addition to these solutions, which provide a file system interface to high-performance data stores, this Task also enables the use of modern database technologies running on top of these high-performance data stores. Using databases allows to significantly speed-up cell segmentation for high-resolution brain images and related data analysis (e.g. determination of cell densities). The strategy is to start with an analysis of the application requirements as well as the technology roadmaps. The solutions must contribute to enabling portability of the applications within the HPAC Platform, which comprises a federation of different systems. A final step within T7.2.1 consists of an evaluation of the implementation of the use cases on relevant architectures.

BSC and JUELICH both contributed to all components listed below.

3.5.3.1 Evaluation of the hierarchical data store prototypes

Description of Component

This Component summarises a first evaluation of the hierarchical data store prototypes.

CDPs to which the Component contributes: None

Progress on Component

First evaluation steps using synthetic test cases have been performed using the HBP PCP pilot systems. A more realistic evaluation based on brain image data is in preparation.

While BSC worked on Apache Cassandra using the Hecuba abstraction layer, JUELICH-JSC focused on technologies created or introduced in the context of HBP PCP, namely DSA/DSS and Ceph.

3.5.3.2 Hierarchical data store software components

Description of Component

This Component comprises the software technology required for making hierarchical data stores usable.

CDPs to which the Component contributes: None

Progress on Component

Possible realisations of the components have been identified and deployed on the HBP PCP pilot systems.

3.5.3.3 Report on hierarchical data store software components design specification

Description of Component

This report lays out the design specification of the hierarchical data store software components.

CDPs to which the Component contributes: None

Progress on Component

Requirements started to be collected and architectural options using different kinds of technologies have been explored as pre-requisite for a design specification.

3.6 T7.2.2 Coupling data analytics and visualisation to simulation

3.6.1 Key Personnel

Task Leader: Stefan EILEMANN (EPFL)

3.6.2 SGA1 DoA Goals

The goal of this Task is to enable coupling simulators on the one hand to data analytics applications and visualisation pipelines on the other hand. To enable interactive steering of future brain simulators, tight coupling must be possible to present results to users within a short time to allow for interactive feedback loops. Furthermore, support for asynchronous data processing and transport of the remaining data is required.

3.6.3 Component Progress

3.6.3.1 Data store sources for interactive visualisation for selected data types

Description of Component

To enable interactive steering of future brain simulators, it is crucial to couple these to data analytics applications and visualisation pipelines such that results can be presented to the user within a sufficiently short amount of time to allow for interactive feedback loops. In a first step, basic data types for coupling data analytics and visualisation to simulation are identified. In a next step the software components needed for in situ data analytics and data extraction will be developed. A first prototypical implementation on relevant architectures, e.g. the HBP PCP pilot systems, will enable a first evaluation of the approach.

CDPs to which the Component contributes: None

Progress on Component

Available input from the upstream components was used to implement the current version of the BRION I/O library.

Links

- BRION repository:

  http://bluebrain.github.io/Brion-1.10/

3.6.3.2 Identification of basic data types for coupling data analytics and visualisation to simulation

Description of Component

To enable interactive steering of future brain simulators, it is crucial to couple these to data analytics applications and visualisation pipelines such that results can be presented to the user within a sufficiently short amount of time to allow for interactive feedback loops. In a first step, basic data types for coupling data analytics and visualisation to simulation are identified.

CDPs to which the Component contributes: None

Progress on Component

Data models and types have been identified and documented for the Neuron simulator.

3.7 T7.2.3 Compute and data resource co-allocation

3.7.1 Key Personnel

Task Leader: Felix WOLF (TUDA)

3.7.2 SGA1 DoA Goals

The goal of this Task is to enable concurrent scheduling of compute and storage resources. Storage based on storage-class memory will be a precious resource and thus needs to be taken into account for resource scheduling. This Task will focus on laying the foundations of new schedulers by defining suitable data resource allocation schemes and identifying resource co-allocation strategies. For suitable simulators, these strategies will be implemented in order to facilitate their evaluation. The final goal is to prepare the integration of these new scheduling algorithms into relevant resource managers like SLURM. During the time between first and second submission of this deliverable, a simulation setup was implemented, which is being used for the evaluation of different architectures, scheduling strategies and workloads.

3.7.3 Component Progress

3.7.3.1 Compute and data resource co-allocation

Description of Component

Based on the results of the report on data resource allocation schemes, the strategies will be implemented in simulators to facilitate simulations on the proposed co-allocation strategies. The foundations will be laid to adapt the simulator also to real resource managers like SLURM.

CDPs to which the Component contributes: Indirect contribution to CDP3 via T7.2.4

Progress on Component

Different simulators have been evaluated and a test environment based on a selected candidate is being prepared. A simulation framework has been implemented based on the Batsim [1] simulator. The report of the now achieved milestone MS7.2.3 (https://emdesk.humanbrainproject.eu/shared/59b10b906f7f7-2ecbbfa21b220f33b442154c4cc9054c) discusses the algorithms.

3.7.3.2 Evaluation of co-allocation prototype

Description of Component

The prototypes that facilitate the co-allocation strategies started to be evaluated and compared in simulations, which will be used to model the real conditions of HPC systems and tested under large workloads inside the simulations. Workloads have been modelled

[1] Pierre-François Dutot et al. “Batsim: a Realistic Language-Independent Resources and Jobs Management Systems Simulator”. In: 20th Workshop on Job Scheduling Strategies for Parallel Processing. Chicago, United States, May 2016.

following the workload modelling methodology of Feitelson *et al*[2]. Simulations are based on an architectural model, which is inspired by the architecture of the HBP PCP Pilot System complex, which integrates different storage tiers. For the evaluation of the co-allocation prototype two storage tiers will be assumed: A Large Capacity Storage Tier (LCST), which is optimised for storage capacity, and a High Performance Storage Tier (HPST), which is optimised for performance, based on dense memory technologies and more tightly integrated with the compute system. Depending on the outcome of this evaluation, next steps towards integration into relevant resource managers like SLURM will be analysed.

CDPs to which the Component contributes: Indirect contribution to CDP3 via T7.2.4

Progress on Component

Preparatory work to enable the evaluation based on simulators and job histories from HPAC systems was completed.

3.7.3.3 Report on data resource allocation schemes and resource co-allocation strategies (report)

Description of Component

The HPAC Platform needs to support both scalable computing as well as processing extreme volumes of data. This will require the integration of precious compute and storage resources and the ability to schedule both types of resources such that efficient utilisation is maximised. This requires a significant change to the current paradigm of scheduling resources in supercomputers, which almost exclusively focuses on compute resources.

As only limited work has been performed in the past to address this challenge, this Task will initially focus on laying the foundations by defining suitable data resource allocation schemes and identify resource co-allocation strategies.

CDPs to which the Component contributes: Indirect contribution to CDP3 via T7.2.4

Progress on Component

The definition of co-allocation problem has been completed and different strategies have been explored.

3.8 T7.2.4 Data analytics workflows for the human brain atlas

3.8.1 Key Personnel

Task Leader: Timo DICKSCHEID (JUELICH)

3.8.2 SGA1 DoA Goals

The goal of this Task is to realise full data analytics workflows in the context of the creation of brain atlases exploiting, in particular, components developed within WP7.2. During the first part of this phase the Task will work on workflows for well-established processes for image registration, segmentation, and low-level analysis in the context of the Big Brain and high-resolution PLI data. The goal for the remaining parts is to expand this to workflows for extracting quantitative data, e.g. cell counts.

3.8.3 Component Progress

3.8.3.1 Data analytics workflows for the human brain atlas

Description of Component

[2] Dror G Feitelson. “Packing schemes for gang scheduling”. In: Workshop on Job Scheduling Strategies for Parallel Processing. Springer. 1996, pp. 89–110.

To establish an early end-to-end solution using technologies developed within this Work Package, a full analysis workflow will be implemented that will exploit both, the HPAC Platform HPC resources and its federated data stores. This effort will also push cross-Platform efforts as it requires interaction with metadata stored in the Neuroinformatics Platform and contributes to the Co-Design Project “Multi-level Human Brain Atlas” (CDP3).

T7.2.4 will initially focus on workflows for well-established processes for image registration, segmentation, and low-level analysis in the context of the Big Brain and high-resolution PLI data. As part of this effort specific requirements to the HPAC Platform will be formulated, provided solutions tested and gaps identified.

More details are provided in the supporting documentation for milestone MS7.2.5 (https://emdesk.humanbrainproject.eu/shared/59b10b80c7c40-c7665ca941034d040a8b8e98060dc13b).

CDPs to which the Component contributes: CDP3

Progress on Component

We prototypically deployed and evaluated three different algorithms on the HPC systems: 1) a cell segmentation algorithm for histological sections using UNICORE on JURECA, 2) a complex GPU-based image registration pipeline on the JURON pilot system, and 3) a Deep Learning approach for brain mapping using a high-capacity convolutional neural network architecture on the JULIA pilot system. The results of these experiments have been documented in a written report, submitted with the achievement of MS7.2.5 (available here: https://emdesk.humanbrainproject.eu/shared/59b10b80c7c40-c7665ca941034d040a8b8e98060dc13b).

Furthermore, we worked out a data management plan and workflow setups for data management with high throughput microscopy, which currently circulates between JUELICH-INM1 and JUELICH-JSC and is used for future storage hardware planning. A preliminary version of this document has been shared with T7.2.3 (S. Lackner) for planning purposes.

We also carried out a practical evaluation of federated storage access patterns for human atlas data providers from SP2 to JUELICH’s GPFS systems. We formalised this access pattern and documented it in the MS7.2.5 report as well. With the recent supply of object storage for SP5 based on SWIFT by ETHZ-CSCS, we are now working to transfer this workflow to the object stores for use with strategic human atlas data.

4. WP7.3 Interactive Visualisation

4.1 Key Personnel

Work Package Leader: Torsten KUHLEN (RWTH)

Work Package Co-Leader: Benjamin WEYERS (RWTH)

4.2 WP Leader's Overview

Despite the postponed signature of SGA1, the work in this Work Package is on track. Developments of components that were started in the RUP have been continued and consolidated (T7.3.1-ETHZ/EPFL, T7.3.2-UPM/URJC). A new focus on in situ visualisation has been established in a series of workshops and working meetings (T7.3.1-RWTH).

In situ visualisation in context of this Work Package addresses on the one hand the visualisation of simulation results from, e.g. a NEST simulation during runtime to overcome the problem of not being able to store the whole generate data but to gain insight into the simulated neural network. When the SGA1 proposal was written (prior to December 2015) ADIOS, HDF5, MPI and ZeroEQ were established frameworks for *in situ* visualisation and communication. In the meantime, Conduit (announced in February 2015, release 0.1.0 in March 2016) and SENSEI (release 1.0.0 in February 2017) have evolved as major tools in the *in situ* visualisation community. RWTH is currently evaluating both, yielding a tendency towards Conduit as basis for the *in situ* pipeline to be developed by RWTH in T7.3.1 (see 4.5.3.7 and 4.5.3.8). RWTH continues the collaboration with JUELICH and Cray to additionally co-design the *in situ* framework that will run on the PCP pilot system JULIA and to address NEST as first simulator. To stick to the project goals, we had regular videoconferences with these partners, which are continued on a regular basis.

As outlined in the DoA, we use the term *in situ* also for the visualisation of large data sets, which do not fit in main memory and thus need dynamic handling of data access, which is similar to the *in situ* visualisation of simulation data and strongly relates to the needed collaboration between WP7.3 and Fenix. Beside the organisation of a workshop on image services (see below), we started working on PLI-Vis (see 4.5.3.9, addressing a further neuroscience use-case, see below) addressing the "large data" aspect of the *in situ* pipeline, while supporting co-design with SP5 on data access and image services.

Furthermore, the goal of the current work has been shifted from the initial idea of establishing interactive supercomputing (main topic of this Work Package in the RUP) as a scientific endeavour. The new focus is to develop and to support visualisation tools and services that are strongly related to well-selected neuroscientific use cases, as for the PLI-Vis component (use case from SP2, see above), and to address the overall goal of the HBP to implement a research infrastructure. RWTH, UPM, URJC and JUELICH started to implement a use case on modelling and creation of neural networks (for running in NEST) by means of developing and deploying visualisation and interactive analysis tools based on previous developments done in the RUP, such as NeuroScheme (see 4.6.3.4) and NeuroLots (see 4.6.3.3). This use case will further drive the co-design needed for in situ visualisation, general visualisation methods for large neural networks, and the developed multi-view paradigm (T7.3.1-RWTH, T7.3.2-UPM/URJC, T7.3.3-RWTH). As part of the visualisation-centric programming framework (SGA1-T7.3.2), and of work on accessing distributed data resources for visualisation that started during the benchmark development in the RUP (necessary for the access of large multi-scale data (SGA1-T7.3.2)), UPM started the prototypical development of a collaborative visualisation middleware (ColViz, see 4.6.3.2) during the first months of SGA1. Due to the request for the development of an interactive tool for the creation of neural network models (see above), UPM decided to put further development of ColViz on hold and moved its resources to the implementation of this use case. The developments on the ZeroEQ library (see 4.5.3.2) started in the RUP and were continued in SGA1 regarding an improvement of its maturity as well as additional functionality (T7.3.1-EPFL). For instance, it has been extended with automated proxy object generation, which enables ZeroEQ to be used for remote steering of ZeroEQ-enabled applications through a Jupyter notebook as it is integrated into the Collaboratory. To provide this functionality, ZeroBuf (see 4.5.3.3) has been extended with an optional http server that provides a web service API in C++ applications. Furthermore, the HPX backend to VTK-m (see 4.5.3.4 and 4.5.3.5) has been improved (T7.3.1-ETHZ) to enable it to be used in production-like environments, rather than purely for testing. This includes the distributed network handling of HPX to make it usable in larger simulations and visualisations in context of the HBP and on the PCP machines. The visualisation software developed in WP7.3 is also facilitated for the HPAC Platform validation (T7.6.4) by UPM, since use cases based on visualisation play a central role to validate neuroscientist workflows in the HBP infrastructure. Finally, components developed in the RUP and not further extended in SGA1 are still maintained, such as Monsteer (EPFL, see 4.5.3.1) and MSPViz (UPM, see 4.6.3.1). All developments are an integral part of the HPAC Platform either as software components (visualisation application and tools), as frameworks support for the integration of visualisation tools (ZeroEQ, multi-view framework) or for *in situ* visualisation (pipeline, HPX, large data visualisation).

The work in WP7.3 was successfully presented in form of scientific publications, as well as talks and poster-based presentations at research conferences and during the HBP Summit 2016. WP7.3 presented an interactive and immersive visualisation of the JuBrain Atlas (T7.3.3-RWTH) at the STOA exhibition at the European Parliament in Brussels. Furthermore, members of the Work Package (T7.3.3-RWTH) are involved in activities of CDP3 in an advisory role, specifically focusing on atlas viewer developments as well as the development of image services as a software entity. The image service is planned to connect viewers as well as analysis software to large data sets accessible through the Fenix infrastructure. This co-design activity defines a bridge between SP5 and SP7. The image service was a major subject of a workshop organized by WP7.3 (Benjamin WEYERS, T7.3.3-RWTH) together with members of SP5 (Jeff MULLER, Timo DICKSCHEID). The major outcome of the workshop is a list of action items outlining a co-design process between users and developers of the image service. The image service will have high relevance for the development of in situ capabilities (T7.3.1-RWTH) as well as for the developed visualisation tools in WP7.3 (T7.3.2-UPM/URJC). Finally, WP7.3 members experimentally installed visualisation software on the PCP machines and wrote a short review on the experiences made as additional feedback for the PCP wrap-up workshops that took place in March 2017 (T7.3.1-RWTH, T7.3.1-EPFL, and T7.3.2-UPM/URJC). The Work Package organised and held its regular face-to-face meeting in Aachen as a two-day workshop in January 2017.

As mentioned in the beginning, the initial six-months delay caused by the postponed signature of SGA1 hindered hiring scientific staff. Nevertheless, due to a huge effort invested in order to bridge financing, it was possible to keep scientists from the RUP in the project. Only this enabled the Work Package to do basic work for the planned components and investigations as defined in the SGA1 work plan. By M12, all partners were able to complete the hiring procedures, and the Work Package is now fully functional. Due to this development, we currently do not foresee any delays in achieving the Milestones scheduled for M18 and M24.

4.3 Priorities for the remainder of the phase

As previously mentioned, a use case has been identified for which a work plan has been extracted for the second half of SGA1, which is in line with the overall SGA1 work plan of WP7.3. The use case specifies the need for interactive tools supporting the creation of large neural networks with a focus on the interactive definition of the connectivity in such networks. The main motivation for this work is the variety of neural simulators and with this the large variety of existing modelling formats. It is hypothesised that neuroscientists will accept an easy-to-use visual description language as a general modelling tool. From this description, the individual input formats can then be extracted algorithmically. This work will be mainly conducted by UPM and URJC (in context of T7.3.2, also involving T7.3.3-RWTH for the creation of multiple-view analysis of the created networks) and in close collaboration with JUELICH.

The created models need to be validated, which can be done by simulating the created model and subsequent analysis of the simulation results. Therefore, a second and closely related use case will be implemented on the basis of the developed/deployed in situ framework. The framework will be used to analyse the model at runtime, e.g. by means of an interactive visualisation of spikes or membrane potentials. In the long run, this work shall include steering capabilities to change the connectivity at runtime. A prototype for this capability has been developed (T7.3.1-RWTH) and a publication is currently under review. Ultimately, this will enable fast turn-around times for the modelling process and help computational neuroscientists to reach their goals more efficiently. This work will be mainly conducted by RWTH (in context of T7.3.1) in close collaboration with UIO and JUELICH.

Both, the creation and validation will first be targeting NEST as a simulation backend. In general, the implementation of the above use cases comprises work from all Tasks in the Work Package, and they are planned to be continued in SGA2 broadening the developed software stack towards NestMC and according to the SP7 visualisation use cases. This work directly contributed to the infrastructure development.

Furthermore, work in T7.3.1 on the flexible application coupling with ZeroEQ including the Jupyter notebook integration as well as the task-parallel visualisation framework developments based on HPX and including VTK-m will be continued. The maintenance and support of the various visualisation facilities offered by the Work Package will be continued and extended by integration of RWTH’s tiled display cluster with seven GPUs into the UNICORE infrastructure of the HPAC Platform heading towards a close integration with Fenix and the NIP, which directly contributes to the infrastructure development.

Finally, WP7.3 was asked in the last review to work on a solution towards the harmonisation of visualisation software development in the HBP. As SP7 representative in the HBP Software Development Committee, Benjamin WEYERS (RWTH) is driving this activity in close collaboration with the SP7 coordination and Anna LÜHRS (JUELICH-JSC) in particular.

4.4 Milestones

Table 3: Milestones for WP7.3 Interactive visualisation

MS No.	Milestone Name	Lead Partner	Task(s) involved	Expected Month	Achieved Month	Comments
MS7.3.1	Event driven streaming integrated in visualisation tools	P46 RWTH	T7.3.1	M18		
MS7.3.2	Improved neuroscience-specific visualisations available	P46 RWTH	T7.3.2	M24		
MS7.3.3	Library for synchronous, multiple view visualisation available	P46 RWTH	T7.3.3	M24		

4.5 T7.3.1 In situ visualisation for neuronal network simulations

4.5.1 Key Personnel

Task Leader: Benjamin WEYERS (RWTH)

Other Researcher: Stefan EILEMANN (EPFL)

Other Researcher: John BIDDISCOMBE (ETHZ-CSCS)

Other Researcher: Tom VIERJAHN (RWTH)

4.5.2 SGA1 DoA Goals

Due to storage and I/O bandwidth constraints that are especially critical for large scale neural network simulations, the major goal of T7.3.1 is to initiate the developments of in-situ visualisation capabilities and components as part of the HPAC Platform. As this is a long-term, multi-phase activity, work in SGA1 focuses on extracting data from live runs of the two major simulation codes in HBP, i.e. Neuron and NEST, and stage this data to a dedicated visualisation/analysis machine.

4.5.3 Component Progress

4.5.3.1 Monsteer (EPFL)

Description of Component

Monsteer is a library for Interactive Supercomputing in the neuroscience domain. Monsteer facilitates the coupling of running simulations (currently NEST) with interactive visualisation and analysis applications. Monsteer supports streaming of simulation data to clients (currently only spikes) as well as the control of the simulator from the clients (also known as computational steering). Monsteer's main components are a C++ library, a MUSIC-based application and Python helpers.

CDPs to which the Component contributes: None

Progress on Component

In the reporting period EPFL did not do any significant changes to this component. T7.3.1 focused on the API design and refactoring of the existing spike and compartment report classes in the underlying Brion IO library, which will allow to implement a simpler streaming plugin, as well as providing a simpler, unified API to files and streams for developers.

Links

- Monsteer in the HPAC Platform Guidebook:

  https://hbp-hpc-platform.fz-juelich.de/?hbp_software=monsteer

4.5.3.2 ZeroEQ (EPFL)

Description of Component

A cross-platform C++ library to publish and subscribe for events. Applications communicate using ZeroMQ, discover each other automatically through the integrated ZeroConf protocol or through explicit addressing using hostname and port. A defined vocabulary defines semantics for the published events, provided by ZeroBuf or using a simple Serializable interface. An optional http server provides a web service API in C++ applications.

CDPs to which the Component contributes: None

Progress on Component

In the reporting period EPFL focused on maturing the ZeroEQ-based ecosystem by (re-) introducing event-based communication in ZeroEQ, moving the vocabulary to the separate Lexis repository, and adding introspection of the exposed data to the ZeroEQ http server.

Furthermore, functionality for steering remote applications from Jupyter notebooks in the Collaboratory based on automatic proxy object generation for any ZeroEQ-enabled application has been added to the ZeroEQ-based ecosystem.

Links

- ZeroEQ in the HPAC Platform Guidebook:

  https://hbp-hpc-platform.fz-juelich.de/?hbp_software=zeroeq

4.5.3.3 ZeroBuf (EPFL)

Description of Component

ZeroBuf implements zero-copy, zero-serialise, zero-hassle protocol buffers. It is a replacement for FlatBuffers, resolving the following shortcomings: Direct get and set functionality on the defined data members. A single memory buffer storing all data members, which is directly serialisable. Usable, random read and write access to the data members. Zero copy of the data used by the (C++) implementation from and to the network.

CDPs to which the Component contributes: None

Progress on Component

For event-based communication as developed with ZeroEQ, we extended the ZeroBuf implementation to allow zero-copy instantiation of immutable objects from within the event handler. This introspection provides a vocabulary endpoint, implemented by the ZeroEQ http server, as well as a JSON schema for each exposed object. We extended ZeroBuf to generate the JSON schema from the input data schema as well as the C++ API to access the schema. Furthermore, functionality for steering remote applications from Jupyter notebooks in the Collaboratory based on automatic proxy object generation for any ZeroEQ-enabled application has been added to the ZeroEQ-based ecosystem. For this, we extended the existing JSON support in ZeroBuf to support partial updates on objects, which is used by the Python API to perform a http PUT with partial data.

Links

- ZeroBuf in the HPAC Platform Guidebook:

  https://hbp-hpc-platform.fz-juelich.de/?hbp_software=zerobuf

4.5.3.4 VTK-m (ETHZ)

Description of Component

VTK-m is a toolkit of scientific visualisation algorithms for emerging processor architectures. VTK-m supports the fine-grained concurrency for data analysis and visualisation algorithms required to drive extreme scale computing by providing abstract models for data and execution that can be applied to a variety of algorithms across many different processor architectures.

CDPs to which the Component contributes: None

Progress on Component

Progress on the task-based libraries in visualisation has been made on two fronts. The VTK-m library has been extended with an HPX backend that allows multithreading using a platform and compiler independent library complying to upcoming C++ standards for parallelism and concurrency. The implementation achieves similar performance to the OpenMP and TBB implementations, but opens up the possibility of creating asynchronous distributed visualisation and analysis pipelines without using additional libraries such as MPI.

4.5.3.5 Task based analysis based on HPX, VTK and VTK-m (ETHZ)

Description of Component

Task based analysis based on HPX, VTK and VTK-m.

CDPs to which the Component contributes: None

Progress on Component

The HPX backend to VTK-m has been improved by ETHZ to enable it to be used in more production-like environments, rather than purely for testing, including the distributed network handling of HPX to make it feasible to use it in larger simulations and visualisations.

The network communication layer in HPX has been extended to support direct RDMA features such as those found in Infiniband verbs and is now capable of higher performance with lower latency. The Infiniband network layer demonstrates that higher performance is possible than with the previous MPI based implementation and work on a new libfabric port of the layer has started which will allow scalable communications on nearly all modern HPC machines, including those provided by Cray, IBM (BlueGene) and Infiniband based vendors.

4.5.3.6 NEST in situ framework (RWTH)

Description of Component

The NEST in situ framework developed by RWTH implements in situ capabilities for the NEST simulator. Main functionality comprises the streaming of spiking and membrane potential data out of a running NEST simulation. Furthermore, it offers functionality to steer a running NEST simulation. The framework is built on the basis of well-established in situ libraries for compatibility and extensibility reasons.

CDPs to which the Component contributes: None

Progress on Component

With regard to the identification of use cases for interactive streaming and steering for the NEST simulator, a prototypic implementation of an interactive parameter analysis for NEST simulations has been conducted by RWTH. This work is ongoing with respect to an extension of the Use Case and a publication is in preparation. It has been presented in a poster during the HBP Summit 2016. Furthermore, current work comprises research on existing in situ infrastructures. A set of frameworks has been assessed in terms of their applicability, requirements and data handling mechanisms. This assessment yielded a tendency to narrow the focus on two platforms to be closely examined. This examination is ongoing in order to decide on a framework serving as the basis for the later implementation as part of the HPAC Platform.

4.5.3.7 NEST-simulated spatial-point-neuron data visualisation (RWTH)

Description of Component

Complementary to other viewer and visualisation implementations for NEST simulations, this component developed by RWTH offers a ray-casting-based rendering of activity and membrane potentials in a neural network simulated with NEST. The renderer is based on high fidelity ray-casting libraries such as OSPRay, which can be used on HPC resources such as the PCP machines. It is meant to be equipped with remote rendering capabilities for an integration into web-based resources. As such, this solution allows for on-node (strict), staged (in transit), and hybrid *in situ* scenarios, as specified in the DoA. Beside the driving use case on the visualisation of membrane potentials emerging from a running NEST simulation, this component represents the initial application connected to the in-situ framework developed in Task SGA1-T7.3.1 (see section 4.5.3.6).

CDPs to which the Component contributes: None

Progress on Component

Fundamental work on high-quality and interactive server-side in situ rendering has been conducted by RWTH. On the one hand, this aims at visualising neuronal data, e.g., membrane potentials, emerging from a NEST simulation. On the other hand, this aims at visualising geometric structures. For both applications, prototypical software has been developed, based on a modern CPU ray tracing framework. This work is closely related to the implementation of an in situ framework (see section 4.5.3.6). It is driven by a use case on the interactive visualisation of membrane potentials emerging from a NEST simulation.

4.5.3.8 Large-scale PLI data explorer (RWTH)

Description of Component

A tool for visualising fibre datasets created with 3D-Polarized Light Imaging (PLI) as well as diffusion Magnetic Resonance Imaging (dMRI). It supports viewing of scalar and vectorial PLI data such as transmittance, retardation and fibre orientation maps. Aside from inspecting the original data, it provides methods for calculating and visualising orientation distribution functions from the fibre orientation maps, through projecting blocks of vectors into the spherical harmonics basis. This provides an effective way to downscale the data contained in the microscopic to macroscopic resolutions targeted with dMRI. Furthermore, the tool supports the visualisation of analyses resulting from spherical harmonics based dMRI methods such as Constrained Spherical Deconvolution and Diffusion Orientation Transform. This enables multimodal comparison of microscopic PLI vector data with macroscopic dMRI tensor data. The Orientation Distribution Functions (ODFs) calculated from microscopic resolution PLI datasets can then be utilised for validation against established dMRI methods.

This tool also offers several new visualisation techniques. The first method is based on calculating a hierarchical level-of-detail structure where the orientation distribution functions are recursively summed to represent larger and larger parts of the volume. The structure serves as a multiresolution representation of the input dataset, where the different resolutions can be browsed individually. It will further be utilised to provide programmatic smooth transition across different scales of data. Several clustering methods are also considered to run on this structure, with the intent of grouping similar orientation distributions together in order to reduce visual clutter when viewing large volumes of interest. Finally, the tool will be extended to support local and global fibre tractography in order to visualise the structural connectivity in a continuous manner, which is not yet available for PLI datasets.

CDPs to which the Component contributes: None

Progress on Component

RWTH conducted initial research on and development of techniques to visualise fibre data-estimates from large-scale PLI data. Two software libraries are under development: The first is for high performance spherical harmonics calculation/sampling, which targets the mathematical requirements of this type of representation and a second for reading from and writing to 3D-PLI datasets. Using these libraries, a tool for calculating orientation distribution functions from 3D-PLI fibre orientation maps is under development. Furthermore, the development of a GUI application for visualizing both the original and processed datasets has started. The latter is planned to make the functionality easily accessible and usable as a standalone application.

4.5.3.9 Streaming & steering between NEST and Livre (EPFL)

Description of Component

This component assembles a set of applications and libraries to enable a workflow to stream data from the NEST simulator to the Livre volume rendering application for observation, and a user interface to steer the simulation based on the observed behaviour. At the high level, the simulator and visualisation engines are linked using Monsteer and MUSIC. On one side, Monsteer provides a MUSIC proxy application hooking into the simulator. This proxy speaks the ZeroEQ protocol to the outside, for which Monsteer delivers a Qt GUI component for steering the simulator, as well as a spike report plugin for Brion to read the data. This setup allows applications based on Brion to read the report data as if it comes from a file, and to integrate the steering UI into a larger application, which also manages the rendering engine.

CDPs to which the Component contributes:

Progress on Component

Based on work in the Ramp-Up Phase, we concentrated in the first year of SGA1 on making the interactive supercomputing components accessible to a wider user base. Most notably, all our applications can now be controlled from Jupyter notebooks in the Collaboratory. Regular quarterly releases of all software components are available.

4.5.3.10 Streaming & steering between NEST and RTNeuron (EPFL)

Description of Component

This component is similar to the previous one, but using a different rendering engine to visualise different aspects of the simulation. RTNeuron is a mesh-based rendering application with parallel rendering support for larger circuits. The work on this component consists of linking RTNeuron to a NEST simulation as in the previous component, testing and optimising the setup.

CDPs to which the Component contributes: None

Progress on Component

Based on work in the Ramp-Up Phase, we concentrated in the reporting period on making the interactive supercomputing components accessible to a wider user base. Most notably, all our applications can now be controlled from python Jupyter notebooks in the Collaboratory. Regular quarterly releases of all software components are available.

Links

- RTNeuron in the HPAC Platform Guidebook:

  https://hbp-hpc-platform.fz-juelich.de/?hbp_software=rtneuron

4.6 T7.3.2 Neuroscience-specific visualisation

4.6.1 Key Personnel

Task Leader: Luis PASTOR (URJC)

Other Researcher: Vicente MARTÍN (UPM)

4.6.2 SGA1 DoA Goals

T7.3.2 will cover the design of visual data representations and the implementation of interactive visualisation methods. Work on data representations will focus on schematic and realistic visual representations of structural and functional data of neural circuits at different levels of abstraction.

4.6.3 Component Progress

4.6.3.1 MSPViz (UPM)

Description of Component

MSPViz is a visualisation tool for the Model of Structural Plasticity visualisation developed together with UPM (main contributor), RWTH, and JUELICH. It uses a novel visualisation technique based on the representation of the neuronal information through the use of abstract levels and a set of representations for each level. These hierarchical representations let the user interact and change the respective visual representation of the data by modifying the degree of detail of the information to be analysed in a simple and intuitive way. This includes functionality for the navigation of different views at different levels of abstraction. The designed representations in each view only contain the necessary variables to achieve the desired tasks, thus avoiding providing the user with an overwhelming amount of information. The multilevel structure and the design of the representations constitute an approach that provides organised views facilitating visual analysis tasks.

CDPs to which the Component contributes: None

Progress on Component

- Dynamic loading of data at visualisation time, from a dCache instance as well from an internal server.
- Upgrade hardware and software of the dCache instance for distributed visualisation.
- Starting the integration of a visualisation node with 3D remote hardware acceleration using different visualisation tools of WP7.3 to facilitate visual analysis of massive distributed data.

Links

- MSPViz in the HPAC Platform Guidebook:

  https://hbp-hpc-platform.fz-juelich.de/?hbp_software=mspviz

4.6.3.2 ColViz Middleware (UPM)

Description of Component

ColViz (Collaborative Visualisation) is a middleware developed by UPM oriented on the application of collaborative visualisation techniques inside the HBP. This middleware focuses on the analysis of huge neuroscientific data sets. Since the size of datasets constantly grows, this also increases the difficulty of knowledge extraction. Visualisation techniques could help in the analysis by taking advantage of the human visual system. A visual analysis using collaborative visualisation techniques is another advantage, as it allows for groups of neuroscientists from different labs to work collaboratively. ColViz enables the neuroscientists to apply collaborative visualisation techniques to ease the analysis of neuroscientific data, using different collaborative visualisation paradigms, both in space and time. This gives the neuroscientist the possibility to analyse the data using co-located and distributed visualisation approaches as well as synchronous or asynchronous mechanisms for knowledge extraction from large and complex data sets.

CDPs to which the Component contributes: None

Progress on Component

Design and first implementation of the middleware ColViz to provide collaborative visualisation techniques inside the visualisation pipeline of WP7.3**.** For testing purposes, a first prototype of a plugin-based visualisation framework (UniVi) has been integrated with ColViz. Multi-view capabilities have been tested using both OpenGL-based and web-based (WebGL and SVG) rendering visualisation plugins.

4.6.3.3 NeuroLOTs (URJC)

Description of Component

NeuroLOTs is a set of tools and libraries that allow creating neuronal meshes from a minimal skeletal description. It generates soma meshes using FEM deformation and allows interactive adaptation of the tessellation level using different criteria (user-defined, camera distance, etc.).

CDPs to which the Component contributes: None

Progress on Component

- Improvements of NeuroLOTs for better soma generation.
- New functionality that allows the user to get a neuronal mesh with the chosen level of detail.

- Improvements in the neurite meshes generation.
- Added the possibility of on-demand loading of morphologies to cope with larger datasets.
- First prototype for the generation of spines at different levels of detail.

Links

- NeuroLOTs in the HPAC Platform Guidebook:

  https://hbp-hpc-platform.fz-juelich.de/?hbp_software=neurolots

4.6.3.4 NeuroScheme (URJC)

Description of Component

NeuroScheme is a tool that allows users to navigate through circuit data at different levels of abstraction using schematic representations for a fast and precise interpretation of data. It also allows filtering, sorting and selections at the different levels of abstraction. Finally, it can be coupled with realistic visualisation or other applications using the ZeroEQ event library developed in T7.3.1.

CDPs to which the Component contributes: None

Progress on Component

- Refactored NeuroScheme in order to be able to create schematic representations from an ontology-based input data.
- Added multi-panel views in NeuroScheme to allow performing side-by-side comparison tasks and provide focus-and-context.
- Added different layouts: grid, 3D camera-based, scatterplot-based, circular.
- Developed ViSimpl, a first prototype for the abstract and schematic visualisation of simulation data, as part of the NeuroScheme framework, enabling a joint visual analysis from temporal, spatial and structural perspectives.
- Started adapting ViSimpl for the visual analysis of correlations between input patterns and produced activity.

Links

- NeuroScheme in the HPAC Platform Guidebook:

  https://hbp-hpc-platform.fz-juelich.de/?hbp_software=neuroscheme

4.7 T7.3.3 Software infrastructure for multi-view analysis of neuroscientific data

4.7.1 Key Personnel

Task Leader: Benjamin WEYERS (RWTH)

4.7.2 SGA1 DoA Goals

Task 7.3.3 will develop a software library for synchronisation of multiple views (CMV) to analyse data from *in vivo*, *in vitro* and *in silico* experiments in a semantic fashion. The overall goal is to facilitate the integrated analysis of various heterogeneous types of data using data-specific, interactive, visual representations.

4.7.3 Component Progress

4.7.3.1 Software library for multi-view analysis of neuroscientific data (RWTH)

Description of Component

This software library will support the synchronisation of multiple views to analyse data from *in vivo*, *in vitro* and *in silico* experiments in a semantic fashion. The overall goal is to facilitate the integrated analysis of various heterogeneous types of data using data-specific, interactive, visual representations. To this end, the target library will aim at a flexible selection and linking of views. This library will use a semantic data description to enable users to interactively create and couple multiple views in very flexible and transparent way. It will be based on an asynchronous communication API for the integration of distributed software components, not necessarily restricted to visualisation applications. The resulting networks of Components could include applications for data analysis. The used approach will follow the event-driven communication scheme which is based on the publish-subscribe software pattern.

CDPs to which the Component contributes: None

Progress on Component

In the reporting period, RWTH focused on the development of a lightweight messaging library which enables flexible implementation of the intended coordinated multiple view library (CMV) including an ontology-based reasoning for gathering networks of various applications and views. Based on this initial development, existing monolithic visualisation components (VisNEST) were split up into single views. Furthermore, the conceptualisation of the semantic support has been finalised for a first basic semantic representation of central components, such as data sources and interaction concepts. The initial implementation of the CMV library has been used in Task 7.3.1 for the implementation of the parameter analysis use case including streaming and steering capabilities for NEST (see section 4.5.3.6). Work on the identification of use cases for the CMV library is ongoing, such as in the context of the flexible instantiation of analysis tools for complex data analysis frameworks, e.g., including the Elephant toolkit.

Links

- VisNEST in the HPAC Platform Guidebook:

  https://hbp-hpc-platform.fz-juelich.de/?hbp_software=visnest

4.7.3.2 Components for visualisation and rendering infrastructure support and service (RWTH/EPFL)

Description of Components

The following components are summarised under this headline:

- CAVE (visualisation system) (RWTH): RWTH provides a fully immersive visualisation environment in the form of a CAVE installation. The aixCAVE is a five-sided virtual reality environment.
- Immersive analysis session (RWTH): An immersive analysis session in a virtual environment, which can be booked through a booking system offered by the HPAC Platform.
- Visualisation systems: booking service (RWTH): This service allows to send a booking request for the visualisation systems at RWTH and EPFL.
- Tiled display wall at RWTH Aachen (RWTH): Tiled display at RWTH Aachen University. The installation is equipped with a seven-node rendering cluster, which is planned to be integrated into the HPAC Platform and the UNICORE infrastructure. This enables the installation to be used through the HPAC Platform as remote rendering site and interactive analysis sessions on site.
- Tiled display wall at EPFL (EPFL): Tiled display wall at EPFL.

All components offer visualisation infrastructure and services to the HBP community.

CDPs to which the Component contributes: None

Progress on Component

All sub-components are continuously maintained by RWTH and EPFL, and are provided to the HBP.

Links

- CAVE, RWTH and tiled display wall, EPFL in the HPAC Platform Guidebook:
  https://hbp-hpc-platform.fz-juelich.de/?page_id=53
- Visualisation booking service
  in the HPAC Platform Guidebook: https://hbp-hpc-platform.fz-juelich.de/?page_id=442
  in the Collaboratory: https://collab.humanbrainproject.eu/#/collab/264/nav/4604

5. WP7.4 Dynamic Resource Management

5.1 Key Personnel

Work Package Leader: Raül SIRVENT (BSC)

Work Package Co-Leader: Julita CORBALÁN (BSC)

5.2 WP Leader's Overview

The main objective of WP7.4 is to enable malleability in both Neuron and NEST (i.e. modify the amount of computing resources they use during run time) as representative applications of the neuroscience field, and advance in the dynamic resource management research topic by enhancing the system capabilities to offer functionalities related to that topic and to deal with them, since systems today do not offer this dynamicity. The dynamic management of resources can be useful to correct non-optimal situations while using computing resources, i.e. improve load balance, increase system utilisation, decrease wait time of applications, etc. As a WP focused on research, WP7.4 has comparably uncertain impact and results as compared to other WPs in SP7. This means that it is difficult at M12 to measure the impact of the work done so far, also because this WP only planned to contribute indirectly to different CDPs, by creating secondary branches of Neuron and NEST with enabled malleability. Anyway, the work plan includes the benchmarking of the solution presented at the end of the SGA1, thus in the end of the phase we will be in the position to be more specific about the impact and benefits of WP7.4 contributions.

The requirements capturing from both NEST and Neuron applications have been proven to be harder than expected, as well as the specification of use cases requiring dynamic resource management capacities. Here, we find a sort of the chicken or the egg problem, since it is very unlikely that users are going to request a capability that does yet not exist. This has made it even more difficult to convince users to adopt this technology, because we had to show the potential benefits one could get, before starting to adopt dynamic resource management. We are happy to say that in both cases the answer has been positive and both NEST and Neuron are willing to include implementations with dynamic resource management.

Another difficulty was when we tried to identify further use cases that could benefit from dynamic resource management. Our initial idea was to interview NEST and Neuron developers to obtain such use cases from them. Anyway, while some ideas appeared, the interviews showed that they were probably not the correct actors to be asked, since their focus is primarily on their software development and not so much in what others do with it. Therefore, we had to completely change the strategy and look for real use cases, opening the scope not only to the rest of SP7, but to the whole HBP project. In particular, a clear use case that could benefit from dynamic resource management is the *in-situ* visualisation use case (with both Neuron and NEST), planned in WP7.3, and also used in WP7.2 as a reference, since the resources for running the visualisation task need to be obtained somehow.

Besides that, CDP2 "Mouse-Based Cellular Cortical and Subcortical Microcircuit Models" (Egidio D'Angelo/Michele Migliore) has shown interest in this dynamic resource management capacity provided in WP7.4. Their plan is to provide Neuron and NEST simulators as a service to their users, and when they are available, groups of different users are going to submit many different simulations of different sizes that will compete for resources, causing a lot of contention to the system. Dynamic resource management will become very useful in this scenario, increasing the utilisation of the system and decreasing the response time for user jobs, making their experience much more satisfactory.

During this first year, and despite some initial delays, we have been able to accomplish almost all the objectives initially fixed for M12, except one (to have NEST with enabled malleability). We have been able to deliver an initial set of dynamic resource management capabilities in the holistic system designed, and to enable Neuron to be malleable to exploit these capacities. The initial delays mentioned have been mainly on the one hand to get answers from the questionnaires from the use cases, and on the other hand due the implementation of dynamic resource management techniques in the SLURM plug-in system, which turned out to be more complex than initially estimated. Nevertheless, for Neuron we have been able to catch up with the delays and to accomplish the M12 deadlines. In the NEST case, a need for deeper investigation meant that we did not achieve to have a first version of NEST with enabled malleability in M12. We have set a new work plan between JUELICH and BSC to speed up the achievement of this goal. In any case, the delay in implementing malleability in NEST does not have an impact on any other WP or Task for the rest of the project. The mitigation strategy has been to increase communication between JUELICH and BSC, both by e-mail and with more videoconferences between the partners. For the SLURM issues also mentioned, BSC assumes this extra effort in SLURM developments as an in-kind contribution, so without any impact on the resources used and planned for the WP.

It is also worth mentioning that we have been in contact with T7.2.3, dedicated to compute and data resource co-allocation, regarding targeting also dynamicity in the resource allocations for the jobs in the system, but while WP7.4 focuses on co-allocating computing resources, T7.2.3 focuses on co-allocating data and compute resources, so no overlap is seen between the work currently planned for these two WPs. Our discussions show that both efforts share many commonalities strategically speaking, and that synergies could be exploited. In particular, the work done with job scheduling simulators can be used in WP7.4 to simulate the implementations done and explore different scenarios, while the expertise gained in WP7.4, especially in working with SLURM plug-ins, can be beneficial for T7.2.3 when moving from a simulated to an actual system. We will continue exploiting this collaboration at least for the rest of the SGA1.

Finally, we would like to remark that the dynamic resource management capacity is an innovative and experimental feature that we plan to intensively test during SGA1, and as such, the TRL is not high enough to consider its integration in a production system such as the HPAC Platform. This is the main reason why there is no plan of integrating WP7.4 results in the HPAC Platform during SGA1, therefore the integration is to be considered for SGA2, or any other future phases of the project.

5.3 Priorities for the remainder of the phase

The first clear priority is to catch up with the delay on implementing malleability in NEST. We plan to have sorted out the situation in three months from now (M15), highlighting again that this delay has no impact on any other Tasks or WPs in HBP.

For the rest of SGA1, we have one more Milestone to accomplish, but we internally defined some others to continue the progress with intermediate objectives. They are:

- MS7.4.3: Scheduling policies (generic and for NEURON) implemented (M22) (official Milestone).
- MS7.4.4-i: Generic policies defined and implemented (M18) (internal Milestone).
- MS7.4.5-i: Implementation and testing of malleability for NEST and NEURON (M24) (internal Milestone).
- MS7.4.6-i: Dynamic Resource Management capabilities as a Service (M24) (internal Milestone).

Therefore, the next Milestone we need to accomplish is the definition and implementation of generic scheduling policies for dynamic resource management in M18. This work is already progressing well, so we do not foresee any delays or problems delivering this, and it will feed the official Milestone MS7.4.3 (the generic part). Thus, from M19 to M22 we will focus on the definition of specific policies for NEURON.

We need to continue our development efforts steadily, trying to advance possible delays to avoid any deviations from the work plan established as it happened with the work related to NEST.

5.4 Milestones

Table 4: Milestones for WP7.4 Dynamic resource management

MS No.	Milestone Name	Lead Partner	Task(s) involved	Expected Month	Achieved Month	Comments
MS7.4.1	Malleability APIs defined	P5 BSC	T7.4.1, T7.4.2	M12	M12	• Job Level API (for job schedulers, e.g. SLURM): Based on the questionnaires to the users and the previous experiments, no new functions have been required in the API. However, considering the scheduling policies that will be introduced, some job hints to the scheduler regarding performance penalty, malleability hints, etc. are needed. This advanced API will be targeted to improve the execution of the application when executed in the context of a workload with dynamic scheduling policies. • Node Level API (used by a node scheduler, e.g. DLB): The API has been refactored to make it extensible to include new and more complex metrics, because when executing jobs in much more complex environments such as shared nodes, the computation of these metrics is more complex. BSC is currently working on refining DLB metrics already specified, and in re-designing the DLB API for monitoring to make it extensible in order to support requirements from T7.4.3. • Application Level API (divided in User and Runtime level): a first version was produced, and a refinement with the questionnaires input has also taken place (T7.2.1). At this level a new function is included which must be used by OpenMP applications, where the concept of malleability is not clearly integrated in the API (there is no function available to inform the application that resources have changed). – User Level API (a user indicates in the code which resources should be used): targets simplicity, so a very small set of calls is selected.

						– Runtime Level API (to be used by programming model runtime, e.g. OpenMP, OmpSs…). At this level, a few set of DLB functions are used to coordinate different processes from the same or different applications inside nodes. BSC is designing and implementing the same basic functionality but asynchronously to support applications not fully parallelized or malleable only at certain specific moments, such as NEURON. • Kernel Level API (at the scheduling level of the Operating System kernel): BSC has re-assessed the needs of this kernel level API with the inputs from the use cases, particularly because the efforts for modifying the Unix/Linux kernel can be high. Since T7.4.2 does not see any special need in the use cases of WP7.4, it has been discarded to tackle this level by now. An internal report describing the details of all APIs will be made available here: https://emdesk.humanbrainproject.eu/shared/59b10b86b2e6c-d329ab08281a67adadb4a165057c7bb9
MS7.4.2	Strategy for work with relevant vendors defined	P5 BSC	T7.4.4	M12	M12	The Pre-Commercial Procurement started in RUP-WP7.1 has produced two pilot systems that serve as a basis of the work to be done in T7.4.4. The future plan for T7.4.4 includes to continue using the pilots, since they will still be available in Jülich. Therefore, further tests (both functionality testing and benchmarking) and developments will be done on these systems. In addition, the end of the PCP RUP-WP means that no more monitoring visits will be arranged, which in turn implies that future communications between WP7.4 and the vendors will have to be peer-to-peer, and of particular importance the communication with the vendor that results selected in the procurement. We will not only interact with PCP vendors, but also with other companies active in the topic, such as Bull and Adaptive Computing. The plans for future interactions are outlined in the report of T7.4.4 in this Deliverable, see section 5.8.

MS7.4.3	Scheduling policies (generic and for NEURON) implemented	P5 BSC	T7.4.3	M22		

5.5 T7.4.1 Co-design of applications to enable dynamic resource management

5.5.1 Key Personnel

Task Leader: Raül SIRVENT (BSC)

Other Researcher: Fabien DELALONDRE (EPFL)

Other Researcher: Markus DIESMANN (JUELICH)

5.5.2 SGA1 DoA Goals

To implement Dynamic Resource Management, applications that are run on a system need to be malleable, i.e. the number of parallel threads used are suitable to be changed at runtime. In this Task, malleability for the NEST and Neuron applications are explored and implemented. Besides that, other representative applications of the HBP will be considered as well. The simulation code experts work together with the dynamic resource management experts to instruct the latter on the organisation of the simulation code and their requirements on the changes.

One of the APIs defined in Task 7.4.2 (the application-level API) is used to indicate whether an application is entering or leaving a parallel computational region, a sequential part or a communication region. In all cases changes in the applications are simple and precise, and targeting the indication of the application structure in terms of parallelism.

5.5.3 Component Progress

5.5.3.1 Study of malleability options in NEST and Neuron

Description of Component

The main objective of WP7.4 is to achieve a novel approach for dynamic resource management in HPC facilities, by combining different tools and techniques to reach a next step in the dynamicity implemented in modern supercomputers, therefore having a direct impact on how neuroscience applications are executed. In order to achieve this objective, a first requirement is that applications that are run on a system need to be malleable, i.e. the number of parallel threads used is suitable to be changed at runtime. It will be explored how malleability for the NEST and NEURON can be realised.

CDPs to which the Component contributes: None (indirectly by contributing to NEST and NEURON)

Progress on Component

The initial work planned for this Component was twofold: on one hand to introduce the dynamic resource management concepts to the use cases participating in T7.4.1 and on the other hand to create a questionnaire that both use cases owners could fill in to identify requirements for dynamic resource management from the applications involved in the use cases. BSC prepared the questionnaire and circulated it to the other T7.4.1 partners, JUELICH and EPFL, in M4 and answers were received in months M7 and M8, respectively. The questions were grouped in different sections depending on the topic asked: describe the experiments in more detail (i.e. what applications may need to run together on the resources), details on the applications involved, the metrics important/needed for these cases, and more general questions. In both cases the most obvious requirement for dynamic resources was when running more than one instance of NEST or Neuron, respectively, also due to the lack of maturity of the analytics codes, that had been selected as the first scenario to be considered to apply dynamic resource management. The main objective in this scenario will be to avoid long waiting times of small simulations when long simulations are running in the system, and to increase the utilisation of the system by exploiting more resources at the same time.

5.5.3.2 NEST with enabled malleability

Description of Component

This NEST branch enables malleability, i.e. to change the number of parallel threads used at runtime.

CDPs to which the Component contributes: None (indirectly by contributing to NEST)

Progress on Component

The key way of enabling malleability in NEST was first to start with a performance analysis together with a code analysis to identify the parts where parallelism is exploited at OpenMP level. Both BSC and JUELICH have been working together in these activities, and one of the first observations is that the NEST application isolates the operations of threads as far as possible to increase performance. This is reflected in a large *parallel* region in the code and that limits the possibility of dynamically changing the resources dedicated to the calculation. Therefore, the first modification was to break down the code in smaller parallel constructs to enable the change of the threads in different parts of the application. Anyway, we have not been able to deliver a first version with enabled malleability for M12. We have established a new deadline for M15.

An initial version of a malleable version of NEST was implemented by M15 as the newly planned deadline. Malleability tests showing how two instances of NEST can share resources have been performed (M18) and are pending to be shown to NEST developers. On the other hand, we have identified a memory utilisation increase in the NEST malleable version intrinsic to the code design (i.e. the usage of the VP, virtual processor structure). A memory utilisation study is ongoing, and a discussion with the developers to find alternative possible solutions is pending.

5.5.3.3 Neuron with enabled malleability

Description of Component

This Neuron branch enables malleability, i.e. to change the number of parallel threads used at runtime.

CDPs to which the Component contributes: None (indirectly by contributing to Neuron)

Progress on Component

When enabling malleability for Neuron, the possibilities were clearer since the beginning, in part because of previous work done on the performance analysis for Neuron from BSC during the Ramp-Up Phase. In this case the main problem was that the workload for each MPI rank and each thread was perfectly balanced before running the application, so a change of resources would cause a load imbalance between threads. After several discussions, both partners agreed that the best approach was to provide a simple API call for the application to notify the system where is the best point in the application to accept a change of resources, and another API from the node scheduler to the application to notify how many resources have been granted to the application, so it can adapt the workload when needed. NEURON has been modified to include this new functionality and it is now ready to be included in a workload to test the dynamic scheduling policies that are defined in T7.4.3.

5.6 T7.4.2 Holistic dynamic resource management

5.6.1 Key Personnel

Task Leader: Julita CORBALÁN (BSC)

5.6.2 SGA1 DoA Goals

Resource management for modern supercomputers must be able to deal with resource sharing not only inside a single application, but also between different applications to support the concept of user/project sessions. To achieve this, the different schedulers inside the system will have to work cooperatively towards the same objective, e.g. to increase the performance of the application, or node utilisation.

This Task defines and implements different levels of malleability APIs in the schedulers used in the supercomputer, so they can exchange information towards obtaining a holistic resource management. The schedulers considered are the kernel scheduler, an application scheduler (i.e. a runtime system in a parallel application), a node scheduler (scheduling resources between different applications inside a computing node) and a job scheduler (in charge of assigning nodes for the execution of applications).

The main difference of the dynamic resource management technique developed in this Task with any other work done in the scientific workflow community is that here we include the job manager as the main actor for implementing the scheduling decisions, while in the literature about workflow scheduling, resources are assumed to be available and job scheduling metrics are never taken into account (e.g. wait time, response time …).

5.6.3 Component Progress

5.6.3.1 Job scheduler with dynamic resource management capabilities

Description of Component

Job scheduler with dynamic resource management capabilities implemented. It is based on high-level malleability APIs for Job Schedulers and Application Schedulers and low-level malleability APIs for Node Schedulers and Kernel Schedulers. This piece of work refers to enabling dynamicity in the Job scheduler, not about its scheduling policies (work done in T7.4.3).

CDPs to which the Component contributes: None

Progress on Component

As described in the SGA1-DoA for T7.4.2, BSC is working on the integration of the whole design in the SLURM job scheduling and queuing system. SLURM provides support for evolving jobs, but this functionality was based on the re-utilisation of an already existing functionality, not specifically designed for that objective, therefore malleability was not supported. The plan is to support malleability by integrating it using the SLURM philosophy of plug-ins and with minimum modifications in data structures. First, a new command has been defined that allows creating a reservation and submitting the job with a single command (sbatch + scontrol). This command is based on the SLURM API and it was a first approach to force resource sharing, making possible to execute a first simple set of evaluation tests for T7.4.3 (scheduling policies). However, using SLURM as the platform for fast prototyping turned out to be more challenging than expected, so the implementations require more efforts than initially planned.

In addition, the SLURM configuration file has been extended with a new threshold, the *SharingFactor*. This argument will be used by all the malleability policies we define to limit how many CPUs per node are granted to be shared with new jobs from already running jobs. These arguments, together with the new command, are going to be used for performance evaluation.

Currently, more parameters are being defined that will allow to define new scheduling policies which will be more aggressive compared to the more conservative approach followed until now. These new parameters will specify, for instance, the maximum penalty allowed in terms of performance introduced to jobs. T7.4.2 is in the process of including these new arguments and updating some core functions used by SLURM resource selection plug-ins to take them into consideration. All work in SLURM is based on SLURM plug-ins, so it is clearly encapsulated, which facilitates the potential adaptation to another queuing system.

5.6.3.2 Definition of APIs for Job schedulers

Description of Component

This report defines high-level malleability APIs for job schedulers and application schedulers, and low-level malleability APIs for node schedulers and kernel schedulers.

CDPs to which the Component contributes: None

Progress on Component

Four different levels of malleability APIs have been defined:

- Job Level API: BSC has tested this API with a number of synthetic benchmarks (PILS, BT, STREAM, …). Based on the questionnaires and the previous experiments, no new functions have been required in the API. However, considering the scheduling policies that will be introduced, some job hints to the scheduler regarding performance penalty, malleability hints, etc. are needed. This advanced API will be targeted to improve the execution of the application when executed in the context of a workload with dynamic scheduling policies.
- Node Level API: The initial tests with SLURM and DLB-DROM API (DLB: Dynamic Load Balancing library, see https://hbp-hpc-platform.fz-juelich.de/?hbp_software=dynamic-load-balancing; DROM stands for Dynamic Resource Ownership Management, where the system requests the application to change its resources) revealed a misbehaviour when two or more jobs entered the system at the same time, therefore a new call DLB_pre_register was defined to avoid deadlocks in the registration process. The API has been refactored to make it extensible to include new and more complex metrics. Previous versions of DLB offer very interesting metrics, such as *node load* to be usable by scheduling policies. However, this metric was calculated when a single job was executing in a node and with "simple" runtime configurations. When executing jobs in much more complex environments such as shared nodes, the computation of these metrics is more complex. BSC is currently working on refining the DLB metrics already specified, and in re-designing the DLB API for monitoring to make it extensible in order to support requirements from T7.4.3.
- Application Level API: a first version was produced, and a refinement based on the questionnaire results has also taken place (T7.2.1). At this level a new function is included which must be used by OpenMP applications, where the concept of malleability is not clearly integrated in the API (there is no function to inform the application that resources have changed).
  - User Level API: targets simplicity, so a very small set of calls is selected.
  - Runtime Level API: to be used by runtimes (OpenMP, OmpSs, …). At this level, a few set of DLB functions are used to coordinate different processes from the same or different applications inside nodes. BSC is designing and implementing the same basic functionality but asynchronously to support applications not fully parallelized or malleable only at certain specific moments, such as NEURON.
- Kernel Level API: BSC has re-assessed the needs of this kernel level API with the inputs from the use cases, particularly because the efforts for modifying the Unix/Linux kernel can be high. Since T7.4.2 does not see any special need in the use cases of WP7.4, it has been discarded to tackle this level by now.

5.7 T7.4.3 Scheduling policies for dynamic resource management

5.7.1 Key Personnel

Task Leader: Julita CORBALÁN (BSC)

Other Researcher: Fabien DELALONDRE (EPFL)

5.7.2 SGA1 DoA Goals

In this Task, different scheduling policies are defined for the different levels of the holistic dynamic resource management system that are defined in T7.4.2. In order to chase a general optimisation objective (e.g. minimise the application execution time), the definition and interaction of the scheduling policies defined at the four different scheduling levels foreseen is studied. Different specific policies need to be defined at the different levels, but their combination must be dealt in a holistic and coordinated manner to ensure the four levels work towards the same objective.

T7.4.3 considers mainly but not exclusively the following generic policies: optimisation of the execution time for a single application or for a group of applications, minimisation of data movements when executing an application to exploit data locality, maximisation of node utilisation or maximisation of the whole system's throughput. Policies defined from the specific use cases of neuroscience applications are also defined.

5.7.3 Component Progress

5.7.3.1 Job scheduler with new scheduling policies

Description of Component

The component represents a finished implementation of the new scheduling policies capabilities in a job scheduler, starting from the generic policies and finishing with the specific policies for the applications provided by the use cases. Note that this piece of work specifically refers to the work done implementing new scheduling policies in the job scheduler, and not about making the job scheduler able to dynamically change the resources for a job (which is handled by T7.4.2).

CDPs to which the Component contributes: None

Progress on Component

During the first six months of SGA1, T7.4.3 implemented a simple generic scheduling policy: when a new job is submitted to a reservation, it gets all the resources it needs, but taking them equally from all existing (running) jobs. This means that the new job has a higher priority than the running ones. For this implementation, it was needed to define a new parameter *priority of jobs* in a reservation, in order to correctly deal with new jobs against jobs already running in the reservation.

Therefore, that priority identified before evolved to what is called the *sharing factor*. A new basic policy has been defined and implemented in T7.4.3 that enables to control the *sharing factor* of resources when multiple jobs are executed in the same node. This argument allows the scheduler to prioritise already running jobs or new jobs. Several synthetic workloads have been executed to evaluate the performance penalty/speedup of applications when varying the sharing factor and considering the application characteristics. This evaluation is the starting point for new scheduling policy proposals. The *sharing factor* has been implemented as a SLURM extension in T7.4.2.

What remains to be done in the next period is to study specific scheduling policies for the use cases while the malleability adaptation of the applications takes place (as defined in T7.4.1). So, by now the generic policies have been defined and implemented by BSC, and next steps will involve EPFL to propose scheduling policies useful for Neuron.

5.8 T7.4.4 Evaluation of dynamic resource management and technology transfer

5.8.1 Key Personnel

Task Leader: Raül SIRVENT (BSC)

5.8.2 SGA1 DoA Goals

Once adaptation of the applications to the new malleability concept has been completed, a systematic evaluation will be performed to establish the benefits of this approach. This Task is furthermore responsible for working with at least one relevant vendor on how to transfer these technologies into its product stacks. When selecting a vendor, T7.4.4 will consider the vendors that participated in the Pre-Commercial Procurement (PCP) of the Ramp-Up Phase of the HBP, but not only them. This effort will allow to make the R&D efforts within this Work Package sustainable and to establish a path towards the developed technologies becoming integral part of future supercomputing solutions.

5.8.3 Component Progress

5.8.3.1 Strategy for technology transfer related to dynamic resource management

Description of Component

This report describes how the knowledge transfer with HPC vendors is established that is required to transfer the new technologies into their product stacks.

CDPs to which the Component contributes: None

Progress on Component

This Task feeds from the Pre-Commercial Procurement (PCP) Tasks that started in the RUP and that still continued their activities during the first months of SGA1 (RUP-WP7.1), where the Dynamic Resource Management topic has been investigated by the participating vendors. Communication and alignment with this PCP Tasks of the RUP has continued. In particular, the PCP pilots have been built and made available by Cray and IBM-NVIDIA (JULIA and JURON, respectively) and integrated by RUP-WP7.1 and RUP-WP7.5. BSC has requested accounts on these systems and started running tests to check for incompatibilities and fixes needed to work with these new architectures. More specifically, it is required to deploy SLURM as a queuing system on the Cray pilot, so that BSC is able to test the modifications made to SLURM also on this machine. Besides, BSC has worked closely with IBM to run the DJSB benchmark on JURON in order to include its results in the PCP Phase III final reports. The results of the benchmark provided by both vendors have been analysed and commented on in the PCP wrap-up workshop that took place in Jülich on the 15th and 16th of March 2017, together with possible future collaborations on the topic.

Since the PCP is now finished, the future plan for T7.4.4 needs to be defined. The pilots will still be available in Jülich, so further tests (both functionality testing and benchmarking) and developments will be done on these systems. In addition, the end of the PCP RUP-WP means that no more monitoring visits will be arranged, therefore future communications will have to be peer-to-peer between BSC and the vendors, especially with the vendor that results selected in the PCP.

During the PCP wrap-up workshops, T7.4.4 has taken the opportunity to explore possible future plans with both IBM-NVIDIA and Cray. In the first case, with IBM-NVIDIA, two scenarios were tested: oversubscription and checkpoint-restart. Now that an environment is ready for testing, new scenarios will be explored that can be tested on JURON (e.g. introducing dynamicity by modifying the number of OpenMP threads) and new versions of the benchmark will be tested with the objective of publishing them as research results.

Cray reports that they would like to have a real Use Case that motivates the dynamic resource management work, rather than the DJSB benchmark, since they do not have other clients or projects requesting these functionalities, as it happens in the case of IBM-NVIDIA). T7.4.4 has agreed to share the progress of WP7.4 (in particular T7.4.1) to show Cray a real use case with Neuron and NEST, and from there analyse the delta in terms of implementation effort to achieve dynamicity with Cray's proposed solution implemented in JULIA.

Besides the PCP vendors, our plans include to contact other vendors or companies active in the topic, such as Bull and Adaptive Computing. Also, an alignment with T7.6.3 is needed. The plan in T7.6.3 is to gather general information and a set of benchmarks publicly accessible to facilitate the contact with commercial operators, as some sort of interface to them, avoiding duplicities.

6. WP7.5 High Performance Analytics & Computing Platform

6.1 Key Personnel

Work Package Leader: Thomas SCHULTHESS (ETHZ)

Work Package Co-Leader: Colin MCMURTRIE (ETHZ)

6.2 WP Leader's Overview

The M1-M12 period (April 2016 to March 2017) was a fruitful one but was not without its challenges. There were significant challenges at the start of the period, due mainly to the delays in the signing of the SGA1. Some partnering institutions risked the loss of key staff but were thankfully able to absorb the costs associated with the delay and press on with the necessary work. As a result, the Federated Data Pilot Project (FeDaPP) was able to deliver a working Proof of Concept (PoC) and a foundation architecture for the Fenix follow-on federated data infrastructure. This work directly supported the requirements of the HBP use cases for data management and allowed the achievement of key milestones, as reported below. WP7.5 also contributed to the crucial outcomes of the DPIT working group which helped pave the way for a reinvigorated and revitalised SP5. In recent months, there has been very positive interactions relating to the identification of use cases and a path forward for the building of an effective infrastructure for neuroscience research. Current efforts focus upon the migration of Collaboratory Web Services to Infrastructure as a Service (IaaS) resources at ETHZ-CSCS and JUELICH-JSC. Good progress has also been made in enabling light-weight virtualisation through the use of Docker containers for software environments coming from other SPs within the Project.

6.3 Priorities for the remainder of the phase

Work in the next period will focus on consolidation of all these endeavours in order to bring an operational infrastructure into the SGA2 timeframe. The PoC for the data management infrastructure, created as part of FeDaPP, will be hardened and brought into a fully operational state. To this end, work on the Fenix architecture will be continued and plans will be enacted to augment the infrastructure with hardware and software capabilities that support interactive supercomputing for neuroscience. The OpenStack IaaS will similarly be refined and augmented in order to meet the needs of the other SPs and CDPs; support for workflows based on Jupyter notebooks will be a key focus as well as provisioning of VMs for needed services. On the HPC systems the support of containerised workloads will be brought into production status at ETHZ-CSCS and effort will be put on expanding the service to other sites, in support of the needs coming from the other SPs and CDPs. Migration of all sites to the new PRACE network configuration will be completed and reporting, monitoring and accounting processes will be hardened. Support for Pay-As-You-Go services will also be deployed at some sites in support of a more flexible method for delivery of compute and data analytics services.

6.4 Milestones

Table 5: Milestones for WP7.5 High Performance Analytics & Computing Platform

MS No.	Milestone Name	Lead Partner	Task(s) involved	Expected Month	Achieved Month	Comments
MS7.5.1	Project Implementation Plan	P18 ETHZ	T7.5.1, T7.5.2, T7.5.3, T7.5.6	M03	M20	As part of the work to define WP7.5 components for the PLA, a Service Breakdown for WP7.5 was created which included a comprehensive breakdown of service components and linkages. This work is summarised in section 2.8 "HPAC Platform: Architecture Overview" of D7.6.1. In addition, a migration plan for moving the back-end services (DevOps and IaaS) of the Collaboratory to one of the host sites of the HPAC Platform (i.e. ETHZ-CSCS) has been developed and agreed. Work is now underway to migrate the Collaboratory Web Services to OpenStack IaaS at ETHZ-CSCS, with a timeline for completion in late M12 or M13. The associated planning documents along with the FeDaPP T6.1 "Architecture Overview" report (https://emdesk.humanbrainproject.eu/shared/5a4f337f7606f-705547810b210f4704711b2d4546be77) and the Fenix Implementation Plan (which will build upon the Services Portfolio Management Proposal for Fenix, the follow-on project from FeDaPP) will form the basis for a comprehensive Project Implementation Plan (PIP) for the HPAC Platform. More time is needed to formulate the PIP for the HPAC Platform which must also take into consideration the outcomes of the DPIT working group, which presented its findings to the EC reviewers on 24th of October 2016. Discussions are underway with the HBP Technical Coordinator and SP5 Leader in order to fully align the PIP with the requirements coming from SP5 and the Collaboratory. The Milestone is therefore further delayed to M13 with the intention to add the PIP as an Annex to D7.6.1. A first version of the Project Implementation Plan is available. It will also be attached as an annex (see Annex A: Project Implementation Plan and the path to Use Case driven Co-Design) to Deliverable D7.6.2, when this is going to be resubmitted. The PIP is also available as a stand-alone document here: https://emdesk.humanbrainproject.eu/shared/5a604e5cf037b-57ef4b5ec370baf5684523bdedbf2a0f

MS7.5.2	Identify concrete use cases for data management	P18 ETHZ	T7.5.1, T7.5.2	M04	M12	The needs of a subset of HBP users have been extensively assessed within FeDaPP T6.2 with major contributions from the partners within the HBP, where CDP1, CDP3, CDP5 and the NEST-Elephant-SpiNNaker use case were examined. Several meetings were held with representatives of these use cases and use case description documents were written that accentuate the needs of these groups for data management. The activities in this area are closely tied to MS7.6.2 and the DPIT working group. Discussions are now underway with SP5 in order to identify all potential data management use cases. However, whilst it is likely that additional use cases will be found, we consider the Milestone as having been fulfilled because, contrary to the period of the RUP, concrete use cases for data management have been identified and described. The FeDaPP use case descriptions are available here: CDP1: https://emdesk.humanbrainproject.eu/shared/59b77f3654d33-a23b47918ed56964d1a937e8cfc788ad CDP3: https://emdesk.humanbrainproject.eu/shared/59b77f3c30e62-7f6fd62c8af32c3f6f150740e68ae0c1 CDP5: https://emdesk.humanbrainproject.eu/shared/59b77f4115d49-ce4141dfe60b3fe22194be68530e8cea (NM UMAN) https://emdesk.humanbrainproject.eu/shared/59b77f2bc29fc-0b6c5523fa534096b6500e7a96fcae4b (NM UHEI) https://emdesk.humanbrainproject.eu/shared/59b77f265d155-5ae7a24476bd6754a3aba6855d2cf9e4 (NM software) https://emdesk.humanbrainproject.eu/shared/59b77f30c6a15-ce3982026f8d0dd564b309ea2be926af (NEST)
MS7.5.3	Storage quota manager and browser	P18 ETHZ	T7.5.2, T7.5.6	M06	M11	This Milestone has a strong overlap with FeDaPP Task 1.1 (Resource Usage Reporting and Accounting Tools), which produced a deliverable outlining the options for resource usage and accounting tools, including a Storage Quota Manager and Browser. As an extension of the work done in FeDaPP, a storage quota manager has been developed providing a set of tools able to store information about the HBP projects storage quota available in the HPAC sites. All reporting data from the four sites (BSC, CINECA, ETHZ-CSCS and JUELICH-JSC) is collected

						and stored in an ElasticSearch DB protected with a combination of username/password per site. The storage quota information can be fetched by any HBP application using the ElasticSearch's native API, as well as using the Grafana interface to pull pre-populated informational charts. Some work was additionally done by the Collaboratory team (HBP-PCO) to allow rudimentary file-system browsing within the Collaboratory. This work will need augmentation based on the use case descriptions from FeDaPP and DPIT, concerning interaction with file catalogues, different storage repositories and features such as initiating and monitoring file transfers. Work is ongoing and will be dealt with in the remainder of SGA1. However, we see this work as now decoupled from the requirements of a Storage Quota Manager and therefore we regard the Milestone as completed. The related documentation can be found here: https://hbp-hpc-platform.fz-juelich.de/?page_id=1448 Deviation from the initial deadline was due to the dependencies on the FeDaPP activities and the development of the Storage Browser, that happened in parallel. The coordination of all partners, implementation of scripts, assessment and enforcing of the authentication process, required more time than initially estimated.
MS7.5.4	Full deployment of user support infrastructure	P18 ETHZ	T7.5.7	M08	M06	T7.5.7 is responsible for the development, deployment and maintenance of a user support infrastructure for the HPAC Platform, which includes a central trouble ticket system for the HPAC Platform, the establishment of a user support team, documentation and an appropriate training programme. A central trouble ticket system (component “HPAC Ticket System”) for the HPAC Platform has been released by BSC in M3 and configured and connected to the ticket systems at BSC, CINECA, ETHZ-CSCS and JUELICH-JSC in the following months. The central ticket system has been connected with the central HBP OIDC system, so that all HBP users can use their HBP accounts for the ticket system as well. A user guide for the ticket system has been written and made available to the HBP (https://emdesk.humanbrainproject.eu/shared/5a4f367b1e037-0aec2f335247fdca30137e83cdcd1e34). The central ticket system is fully operational since M6. The HPAC Platform user support team (component “HPAC User Support Services”) has members at all four HPC sites, and it can be reached through the new email address hpac-

						support@humanbrainproject.eu. Emails sent to this email address automatically open a ticket in the central ticket system, that is either answered by the support team or routed to one of the 2nd level ticket systems. All online contact forms of the HPAC Platform are connected to this new email address since M6. The documentation of the HPAC Platform is available in the "HPAC Platform Guidebook" (component "HPAC Platform Guidebook", available from the HPAC Platform Collab: https://collab.humanbrainproject.eu/#/collab/264/nav/2378), which was deployed already in the Ramp-Up Phase and is continuously updated. This website also contains the HPAC Platform training programme. With the central ticket system, the user support, the new central email address and Guidebook, the user support infrastructure for the HPAC Platform is fully deployed and operational since M6, which is two months earlier than originally scheduled.
MS7.5.5	Key functionalities of UNICORE Portal integrated into the Collaboratory UI	P18 ETHZ	T7.5.6	M09	M09	This milestone includes the following key functionalities: 1) Access to HPC storage 2) HPC job submission and management features 3) Information about a user's compute allocations, storage quotas and availability of the various sites in the infrastructure The three functionalities are available in a first version and thus this milestone is achieved. Work on enhancing the functionality continues: 1) Access to HPC storage The Collaboratory app "HPC Storage Browser" provides basic access to the HPC storage. Basic features such as file browsing, upload, download, rename etc. are supported. The application internally uses the UNICORE REST APIs. 2) HPC job submission and management Jupyter notebooks have emerged as the preferred way for users to realise their custom computing applications in the Collaboratory. Thus, HPC job submission and management was made available to users via Python bindings for the UNICORE REST APIs that can be loaded and used from Jupyter notebooks. The Python bindings are available as a module "pyunicore" that is being developed on GitHub:

						https://github.com/HumanBrainProject/pyunicore An example Jupyter notebook is under development and will be released by the end of January. This can serve as a convenient starting point for users and will include code examples and links to documentation. 3) Information about HPC allocations Functions for getting basic information about a user's available compute allocations and storage space at the HPC sites are part of the UNICORE Python client bindings described above. Specifically, users can check at which sites they do have compute allocations, but not (yet) how much compute budget is available to them. Additionally, for quickly checking general site availability and service status, a monitoring app was developed, which is part of the HPAC Collab.
MS7.5.6	Proof of concept of container-based SW packaging and deployment for HBP	P18 ETHZ	T7.5.1	M11	M11	A Docker-based deployment of the simulation package for the Neurorobotics Platform (SP10) has been setup on a cluster at CSCS. Specifically, the work involved moving the code that was natively compiled for the BBP's *vizcluster* to be executed directly from a container runtime. The deployment includes the installation of all the software services required to support the software stack of the Neurorobotics Platform, e.g., user accounts, network configuration, Docker runtime, etc. The software stack contains two main components, namely the *Backend* and the *Frontend*. Their deployments were tested and documented in order to replicate them from a freshly-installed CentOS 7.10 distribution on the above-mentioned cluster. Moreover, an encrypted SSH tunnel was created in order to enable the Neurorobotics developers external access to the infrastructure. Its installation and setup was also documented. The installation of the Neurorobotics services is documented here (snapshot of continuously updated document, status 15 May 2017): https://emdesk.humanbrainproject.eu/shared/59b10b59af6a3-95d6200b3987b0460401849c9397bf09
MS7.5.7	Deployment of prototype data management infrastructure that meets the use-case requirements	P18 ETHZ	T7.5.2, T7.5.3, T7.5.6	M12	M7	A PoC data management infrastructure was successfully demonstrated as part of FeDaPP. This PoC was based on the needs of the Polarized Light Imaging use case identified as part of FeDaPP T6.2. The infrastructure created in FeDaPP is still in place and will form the core of the Fenix infrastructure. However, it needs to be hardened in the remainder of SGA1 in order to also fulfil MS7.5.8.

						A presentation of the FeDaPP PoC is available here: https://emdesk.humanbrainproject.eu/shared/59b10b469dc7d-dfef4f77580d41f3fe5df93bfe0fd1b4
MS7.5.8	Federated data management infrastructure operational	P18 ETHZ	T7.5.2, T7.5.3, T7.5.6	M24		
MS7.5.9	Production-level SW packaging and deployment on production HPC systems as well as related services	P18 ETHZ	T7.5.1	M24		Work is ongoing and linked to the work done on a PoC of such deployments in MS7.5.6. In order to feature full production deployments of the PoC workflow at scale, the following milestones will need to be accomplished in the M12-M24 period: 1) Containerise a headless *gazebo* server, finding a way to run it using GPU acceleration through the *nvidia-docker* facility; 2) Deploy the containers from Item 1) using Shifter on CSCS' production Test and Development System (TDS); 3) Prototype the setup using the R&D system to run both *Frontend* and *Backend*, initially launching the compute-intensive jobs remotely on the TDS using Slurm; 4) Update the remote job submission to the TDS to use UNICORE; 5) Move from the R&D system to HPAC-hosted OpenStack VMs for deploying both *Frontend* and *Backend* services, keeping the compute-intensive jobs as an independent component.

6.5 T7.5.1 HPC and Cloud Services

6.5.1 Key Personnel

Task Leader: Colin MCMURTRIE (ETHZ)

Other Researchers: Cristian MEZZANOTTE, Lucas BENEDICIC (ETHZ)

Other Researchers: Jutta DOCTER, Dorian KRAUSE, Ralph NIEDERBERGER, Bernd SCHULLER, Björn HAGEMEIER (JUELICH)

Other Researcher: Giovanni ERBACCI (CINECA)

Other Researcher: Javier BARTOLOMÉ (BSC)

Other Researcher: Marcus HARDT (KIT)

Other Researcher: Jeff MULLER (EPFL-PCO)

6.5.2 SGA1 DoA Goals

The SGA1 goals of T7.5.1 are to build on and leverage existing capabilities of the sites that host the HPC systems in order to better support the computational science workflows of the HBP. Key areas of focus are:

1) HBP-specific developer and user environments will be integrated with data centre services;
2) Leveraging of existing development activities at the sites and their ongoing collaborations with vendors, work will be done in providing support for container-based software packaging and deployment technologies (e.g. Docker images) on production HPC systems;
3) Support will be provided for common developer services such as bug tracking and source control, as well as continuous integration and development services via tools such as Jenkins for test-driven development;
4) Support for the augmentation of cluster-based Cloud infrastructure at participating data centres and providing monitoring services and VM configuration systems.

6.5.3 Component Progress

6.5.3.1 Components and their Descriptions

- A: HPC systems at CSCS (DoA goal 1): Piz Daint
- B: HPC systems at Cineca (DoA goal 1): Marconi and Pico
- C: HPC systems at BSC (DoA goal 1): MareNostrum
- D: HPC systems at JSC (DoA goal 1): JUQUEEN and JURECA; PCP Pilots: JULIA and JURON
- E: KIT S3 Storage (DoA goals 1, 4): Cloud storage with S3 interface
- F: Container-based software packaging and deployment technologies (DoA goal 2): Providing support for container-based software packaging and deployment technologies (e.g. Docker, which can be considered as light-weight VMs) on production HPC systems
- G: Common developer services (DoA goal 3): Support for common developer services such as bug tracking and source control, as well as continuous integration and development services via tools such as Jenkins for test-driven development
- H: Lightweight VM Service (DoA goal 2): Provision of virtual machines for the users to setup their own services

- I: VM Services (DoA goal 4): Providing virtual machine resources and container support at KIT.
- J: CSCS Computing Service (DoA goal 1): The computing services at the Swiss National Supercomputing Centre
- K: JSC Computing Service (DoA goal 1): Computing services offered by Jülich Supercomputing Centre
- L: CINECA Compute Services (DoA goal 1): Compute services provided by CINECA
- M: BSC Computing Service (DoA goal 1): Computing services provided by Barcelona Supercomputing Centre
- N: CSCS Workload Manager (DoA goal 1): CSCS Workload Manager
- O: JSC Workload Manager (DoA goal 1): JSC Workload Manager
- P: CINECA Workload Manager (DoA goal 1): CINECA Workload Manager
- Q: BSC Workload Manager (DoA goal 1): BSC Workload Manager
- R: KIT VM System Service (DoA goal 4): KIT VM System Service
- S: KIT OpenStack Deployment (DoA goal 4): KIT OpenStack Deployment
- T: CSCS Docker Service (DoA goal 2): CSCS Docker Service
- U: Shifter (DoA goal 2): Shifter
- V: SP7 Federated HPAC Computing Services (DoA goals 1-4): SP7 Federated HPAC Computing Services
- W: HPAC Computing Services (DoA goals 1, 2, 4): HPAC Computing Services
- X: HPAC Monitoring Service (DoA goal 4): Monitoring the status of the HPAC Platform infrastructure

CDPs to which the Components contribute: The computing infrastructure is used by several CDPs directly or indirectly (by using data and results produced using the computing infrastructure).

Progress on Components:

Components A, B, C, D, relating to the operations of the HPC systems, in conjunction with Components N, O, P, Q relating to the site-specific workload managers, enable and support Components J, K, L, M which comprise Component W and thereby provide the service offered by Component V, namely the SP7 Federated HPAC Computing Services. During the first year of SGA1 these services have been sustained by the sites in normal operations. In addition, as reported in T7.5.3 below, work has been underway to connect all Components A, B, C, D to the new PRACE network infrastructure. The latter work is on-going with a target completion before the end of SGA1.

Work on Component F (Container-based packaging and deployment technologies), in conjunction with Component T (CSCS Docker Service) and Component U (Shifter) has been ongoing in the period with good progress being made in understanding a complex workflow coming from SP10 (Neurorobotics Platform). A working PoC using Docker containers running on an R&D system has been created at ETHZ-CSCS, thereby fulfilling MS7.5.6. Work is now underway to move the workflow to a production Test and Development System (TDS) where Shifter has been deployed to run the Docker containers on a production HPC architecture, as part of the work to fulfil MS7.5.9. It is important to point out that the deployment of

Docker containers is not performed directly but rather using Shifter due to the scalability and security issues of standard Docker images[3].

Component I (VM Services) has been augmented with the deployment of two OpenStack IaaS environments at ETHZ-CSCS. One is an R&D environment that has been created to help understand the complexities of OpenStack IaaS and to furthermore help specify a new production OpenStack environment. The other is a production-ready OpenStack IaaS that has been created to be the initial host for the Collaboratory Web Services. The target is to have started the migration of production Collaboratory Web Services to this environment in the late M12 or M13 timeframe. Once this is done, work will start on federating the OpenStack IaaS in order to host the Collaboratory Web Services at, at least, two sites (i.e. ETHZ-CSCS and JUELICH-JSC).

Component X (HPAC Monitoring Service) has been created as part of T7.5.3, see below, using the *Icinga* monitoring server and a simple viewer has been created in the Collaboratory.

Component S (KIT OpenStack Deployment) is up and running and is available at https://oscloud-1.scc.kit.edu/. The service provides OpenStack-based IaaS, which enables users to configure and launch virtual machines and Docker containers. The KIT cloud was integrated with the HBP AAI, allowing users to easily connect using their HBP accounts. The procedure for requesting access to this service is documented at https://collab.humanbrainproject.eu/#/collab/264/nav/16576.

Component E (KIT S3 Storage Service) remains up and running. It is available via the S3 Protocol at https://s3.data.kit.edu. Usage by HBP continues to be on a low level (1.55TB of 500TB used). The reasons for the low usage needs to be investigated, especially since it is well integrated with the HBP AAI and the various Platforms via UNICORE. Furthermore, Component E was integrated with Component S to the extent that VM images may now be stored in the Storage Service, thereby supporting the delivery of Component R (KIT VM System Service).

Links

- Infrastructure section in the HPAC Platform Guidebook:

  https://hbp-hpc-platform.fz-juelich.de/?page_id=32

6.6 T7.5.2 Data Services

6.6.1 Key Personnel

Task Leader: Giovanni ERBACCI (CINECA)

Other Researcher: MUCCI, Roberto (CINECA)

Other Researcher: MEZZANOTTE, Cristian (ETHZ)

Other Researcher: BARTOLOME, Javier (BSC)

Other Researcher: SCHULLER, Bernd (JUELICH)

Other Researcher: GUDU, Diana (KIT)

6.6.2 SGA1 DoA Goals

This Task is focused on the testing and deployment of tools and technologies for data management (e.g. AFM, dCache, iRODS, OpenStack, WOS, S3, CDMI, OpenNebula, etc.). Furthermore, a data management infrastructure will be created that is aligned with the workflows and use cases of the HBP. In particular, storage and quota management will be

[3] Benedicic, Lucas, Felipe A. Cruz, and Thomas C. Schulthess. "Shifter: Fast and consistent HPC workflows using containers.", Cray Users Group Conference (CUG'17), 2017.

supported. Interoperability with external Clouds (e.g. AWS, federated Clouds like EGI) will be investigated and enabled, as applicable.

6.6.3 Component Progress

6.6.3.1 Data, network and AAI services

Components summarised in this section

- SP7 Federated HPAC Data and Computing Services: Federated data and compute infrastructure built, provided and maintained by WP7.5, including: supercomputers, HPC clusters, storage at HPC centres, cloud storage, high-speed network between the SP7 HPC sites, account creation and management. There are, by design, strong links to the FeDaPP components.
- SP7 Federated HPAC Data Service: The data part of the SP7 infrastructure (closely linked to FeDaPP/FENIX).
- Data federation and management protocols: Data federation and management protocols of the federated data infrastructure
- CSCS Data Service: CSCS Data Services
- JSC Data Service: JSC Data Services
- BSC Data Service: BSC Data Services
- CINECA Data Service: CINECA Data Services

CDPs to which the Components contribute:

- CDP2-UC-001 - single cell modelling
- CDP2-UC-002 - Multi-scale validation
- CDP3-P2 3D interactive big data viewer
- CDP3-P3 Provision and maintenance of template datasets with labelled parcellations in the Human Brain Atlas
- CDP3-P5 Interactive spatial alignment tools for human brain data

Progress on Components

Work driven by Federated Data Pilot Project (FeDaPP) has aimed at building an efficient federated data infrastructure based on current infrastructure services of the centres involved in the HPAC Platform, starting with CINECA, ETHZ-CSCS and JUELICH-JSC.

A proof of concept (PoC) has been defined to show that the most important functionalities produced by FeDaPP are in place and work as expected. The PoC also demonstrates the capabilities of the data services infrastructure.

The services investigated in the context of FeDaPP are currently extended to BSC and, later on to KIT for the Cloud services.

A description of the different activities is reported in the following paragraphs.

The consolidation of the data infrastructure and the assessment of new technologies will be the objective of the FENIX project, currently under design.

6.6.3.2 Active data repositories

Components summarised in this section

- Active data repositories
- CSCS Active Repository Service
- JSC Active Repository Service

- CINECA Active Repository Service

Description of Components

The following functionalities distinguish an Active Data Repository from the other data store types (in SP7):

1) Data store localised close to computational or visualisation resources such that high performance access to data is enabled (in terms of high bandwidth and/or high IOPS rates);
2) Used for storing temporary slave replicas of large data objects, with master copy of data kept in archival data repository, for improving access performance;
3) Availability of resources similar to attached computational or visualisation resources. If data gets lost then it needs to be replicated again (no data loss, but performance impact).

CDPs to which the Components contribute: CDP1-P5: A data explorer and importer app

Progress on Component

In order to set up a fast and reliable cross-site data transfer service the GPFS Active File Management (GPFS AFM) technology has been investigated and assessed.

Through GPFS AFM a part of the GPFS can be shared among sites allowing data replication and improving data availability: this is particularly useful when data has to be moved efficiently between computing facilities and does not have to be shared with external actors.

Each HPAC site (ETHZ-CSCS, CINECA, JUELICH-JSC, BSC) set up a GPFS AFM instance and performed data replication tests: local directories have been exported between partners in a read-only mode. During the tests, data produced in one site was transparently replicated in a second one, making the same data available across the sites.

6.6.3.3 Archive data repositories

Components summarised in this section

- Archive data repositories
- CSCS Archive Repository Service
- JSC Archive Repository Service
- CINECA Archive Repository Service
- BSC Storage

Description of Components

The following functionalities distinguish an Archival Data Repository from the other data store types (in SP7):

1) Data store optimised for capacity, reliability and availability (tier-1 level availability);
2) Used for storing large data objects permanently;
3) Data is not replicated and thus will be unavailable when the repository used for storing data is not available.

CDPs to which the Component contributes:

- CDP1-P5: A data explorer and importer app
- SP9-UC-007b Long-term storage of a simulation result e.g. for publication in a paper

Progress on Component

Archival data refers to datasets that require reliable long-term accessibility and cannot be easily generated (e.g. resulting from computations).

The identified technology for the archive data repositories is IBM TSM (backup/archive) which is already deployed in production environments at all HPAC sites (ETHZ-CSCS, CINECA, JUELICH-JSC, BSC).

Alternative solutions, such as Object Storage services, are under investigation among the involved partners and will be further assessed in order to avoid a vendor lock-in and to overcome POSIX file systems limitations in terms of data access policies.

6.6.3.4 Upload, download and transfer services

Components summarised in this section

- Upload caches & services
- CSCS Upload/Download Service
- JSC Upload/Download Service
- CINECA Upload/Download Service
- CSCS Data Transfer Service
- JSC Data Transfer Service
- CINECA Data Transfer Service
- Download services

Description of Components

The following functionalities distinguish the Upload Cache (also known as Data Staging Repository) from the other data store types in SP7:

1) Data store located next to a data source outside of the HPC data centre infrastructure with limited capacity, optimised for reliability and performance;
2) Used for keeping temporary copies of large data objects, before these are moved to an archival data repository;
3) Availability must be as good as required by the attached data source;
4) Note that the classification is largely functional, i.e. the performance requirements are to be defined in a post-classification step.

Download services enable users to download datasets stored in the federated data infrastructure given that they have the necessary access permissions.

CDPs to which the Components contribute:

Progress on Components

The technology investigated and identified to provide data transfer services (upload and download) is UNICORE file transfer protocol (UFTP). UFTP provides high-performance data staging and data movement functionalities, taking advantage of the UNICORE authentication mechanisms.

In the first year of SGA1, UFTP has been installed and configured at all HPAC sites (ETHZ-CSCS, CINECA, JUELICH-JSC, BSC) and the service has been demonstrated to work as part of the FeDaPP Proof of Concept (PoC). In the PoC demonstration, data was efficiently moved across the HPAC sites using the UFTP command-line client. UTFP is also available as transfer protocol in the Data Manager of the UNICORE Portal, allowing data movement directly from the HBP Collaboratory.

Besides UFTP, GridFTP has been evaluated due to its adoption as data transfer service at different HPC sites. At the moment, GridFTP is not supported as a data transfer service as it needs integration to the UNICORE environment.

6.6.3.5 Storage quota manager and browser

Description of Component

A tool showing the user's storage quota at the HPAC sites.

CDPs to which the Component contributes: None

Progress on Component

This Milestone has a strong overlap with FeDaPP Task 1.1 (Resource Usage Reporting and Accounting Tools), which produced a deliverable outlining the options for resource usage and accounting tools, including a Storage Quota Manager and Browser.

A storage quota manager has been developed providing a set of tools to store information about the HBP projects storage quota available at the HPAC sites. All reporting data from the four sites (BSC, CINECA, ETHZ-CSCS and JUELICH-JSC) is collected and stored in an ElasticSearch DB protected with a combination of username/password per site.

The storage quota information can be fetched by any HBP application using the ElasticSearch's native API, as well as using the Grafana interface to pull pre-populated informational charts.

Some work was additionally done by the Collaboratory team (EPFL-PCO) to allow rudimentary file-system browsing within the Collaboratory. This work will need augmentation based on the use case descriptions from FeDaPP and DPIT, concerning interaction with file catalogues, different storage repositories and features such as initiating and monitoring file transfers. Work is ongoing and will be dealt with in the remainder of SGA1. However, this work is now considered as decoupled from the requirements of a storage quota manager.

Links:

- More information about the storage quota manager and browser:

  https://hbp-hpc-platform.fz-juelich.de/?page_id=1448

6.6.3.6 Monitoring Services

Components summarised in this section

- HPAC Monitoring Service: Monitoring the status of the HPAC Platform infrastructure.
- BSC Monitoring Service
- CINECA Monitoring Service
- JSC Monitoring Service
- CSCS Monitoring Service
- KIT Monitoring Service
- ICINGA: Icinga is a resilient, open source monitoring and metric solution based on a new object-based, rule-driven configuration

CDPs to which the Components contribute: None

Progress on Components

T7.5.2 developed an Incinga-based monitoring system of the UNICORE services, which is in place in each HPAC site. A page showing the HPAC Platform status is available from within the Collaboratory.

Links

- Monitoring page in the HPAC Platform Collab:
  https://collab.humanbrainproject.eu/#/collab/264/nav/4307

6.6.3.7 KIT Cloud Storage Services

Description of Component

The Cloud storage services provided by KIT.

CDPs to which the Component contributes: None

Progress on Component

The KIT Cloud Storage Service put in place in the Ramp-Up Phase continues to be up and running and is maintained by KIT. It is available via the S3 Protocol at https://s3.data.kit.edu.

The integration with the HBP AAI and UNICORE was completed. The KIT Cloud Storage Service was integrated with the KIT Cloud Computing Service to the enable storing VM images in the Cloud Service.

Links

- KIT Cloud Storage documentation:
  https://hbp-hpc-platform.fz-juelich.de/?page_id=1326

6.7 T7.5.3 Low-level Infrastructure Services

6.7.1 Key Personnel

Task Leader: Ralph NIEDERBERGER (JUELICH)

Other Researchers: Stephan GRAF, Carsten KARBACH, Bastian TWEDDELL (JUELICH)

Other Researchers: Javier BARTOLOMÉ, Sergi MORÉ (BSC)

Other Researchers: Marco ALBERONI, Stefano Claudio GORINI, Roberto MUCCI (CINECA)

Other Researchers: Massimo BENINI, Nicola BIANCHI, Chris GAMBONI, Miguel GILA, Giuseppe LO RE, Christian MEZZANOTTE, Marco PASSERINI (ETHZ)

Other Researcher: Stefano ZANINETTA (EPFL)

6.7.2 SGA1 DoA Goals

This Task deals with the operation and maintenance of the low-level federated infrastructure, including the network, Authentication and Authorisation Infrastructure (AAI), accounting, monitoring, and middleware. Security aspects relating to the infrastructure also form an important part of the work. Assistance and recommendations for the integration of new computing systems into the infrastructure are also undertaken in this Task.

6.7.3 Component Progress

6.7.3.1 Components and their Descriptions

Site-specific implementations of many of the components below exist as well.

- A: HPAC Network Service: The secure high-speed network connects the HPC & data centres of SP7: BSC, CINECA, CSCS, JUELICH and KIT.
- B: SP7 Federated HPAC Data and Computing Services: Federated data and compute infrastructure built, provided and maintained by SP7: supercomputers, HPC clusters,

storage at HPC centres, cloud storage, high-speed network between the SP7 HPC sites, account creation and management. It does also link to the FeDaPP components.

- C: HPAC Reporting and Accounting Service: Mechanisms for reporting and accounting of the resource usage
- D: Upload caches & services: see section 6.6.3.4
- E: Download services: see section 6.6.3.4
- F: Transfer services: see section 6.6.3.4
- G: Federated user database: A federated user database for the federated infrastructure
- H: HPAC Authentication & Authorisation Infrastructure Services: HPAC Platform Authentication & Authorisation Infrastructure (AAI) provides single sign-on mechanism. The underlying LDAP accounts are linked to the user's central HBP account.
- I: HPAC Monitoring Service: Monitoring the status of the HPAC Platform infrastructure
- J: Middleware: Maintenance and operation of the HPAC Platform middleware
- K: HPAC Security Services: The low-level infrastructure of the HPAC Platform
- L: Low level infrastructure: The low-level infrastructure of the HPAC Platform
- M: Accounting service: Accounting of the HPAC Platform access
- N: HBP LDAP: The LDAPs set up at the HPC & data centres of SP7 are used for the federated user management.
- O: ICINGA: see section 6.6.3.6
- P: Grafana with ElasticSearch: Grafana is the leading open source project for visualizing metrics.
- Q: PRACE Network / Internet Connection: The dedicated network operated by PRACE / Connectivity to the internet
- R: Unity: http://unity-idm.eu/

CDPs to which the Components contribute: None

Progress on Components

Components A, B, K and Q: Within the first year of SGA1 plans for a migration of the old dedicated PRACE network infrastructure to a more flexible virtual infrastructure have been made. Technical requirements for a transition and IT security implications have been analysed and harmonised, so that HBP IT security policies can be guaranteed. Together with the PRACE project a decision has been made to start the migration process in October 2016. So, parallel to the dedicated PRACE network infrastructure, a virtual private network (VPN) solution has been installed. The HPAC Platform partners BSC, CINECA, ETHZ-CSCS, and JUELICH-JSC have been connected to the new virtual infrastructure already. A link between the old dedicated and new virtual infrastructure has been set up, so that a smooth migration of the remaining partners can be fulfilled. Furthermore, an IT security mailing list of security personal of each partner organisation has been created, where together with the PRACE project partners, security questions, policies etc. can be discussed and incidents solved. Also, networking activities in the area of the IT security have been taken forward by cooperating with other e-infrastructures in the WISE community. The areas addressed here are “strengthen of trust” between collaborating partners and security in big and open data.

Components B, I and O: Basic monitoring of the HPAC Platform has been set up using the Icinga monitoring server including its web frontend. Plugins for checking the status of the UNICORE services using the UNICORE REST API have been implemented and deployed. A simple viewer application in the Collaboratory developed by JUELICH provides an overview of the HPAC services status. Each HPAC site (ETHZ, CINECA, JUELICH, BSC, KIT) has configured its UNICORE installation for access by the monitoring server (see also section 6.6.3.6).

Components C, G, H, I, Q and R: This Task has a strong overlap with activities in FeDaPP related to monitoring and Authentication and Authorisation. The HBP AAI has been demonstrated to work as part of the FeDaPP Proof of Concept (see T7.5.2, section 6.6).

Components N: The central LDAP server (master server) for HBP HPC projects and accounts was migrated from EPFL to a new virtual machine at JUELICH-JSC. A REST API was developed and installed on that server to allow central management of HPC accounts for the entire HPAC Platform. This setup implements identical project groups and accounts on all member site HPC systems. Access to that API layer was granted to system administrators at all sites to enable automated project and account creation and propagation throughout the infrastructure.

Components C and P: An architecture design and implementation plan was developed for a storage quota manager and browser. This tool collects quota information for all HPC users of the HBP in a central database based on ElasticSearch, which is hosted by CINECA. In regular intervals quota information is pushed by all HPC and storage sites to the central database. The collected data can be visualized by Grafana, which eventually is to be integrated into the Collaboratory for HBP users. An adapter software was developed at JUELICH-JSC for converting available quota information into the required data format for ElasticSearch. The current quota information is parsed and pushed into the database, so that on-line plots of the usage of storage resources of the distributed HPC infrastructure can be generated.

Components B, D, E, and F: A testbed for a GPFS-AFM (Active File Management) based data infrastructure has been setup. Here, local directories have been exported between partners in a read-only mode. A workflow has been created which demonstrated the usability of the tested environment.

6.8 T7.5.4 Application Software Services

6.8.1 Key Personnel

Task Leader: Benjamin CUMMING (ETHZ)

Other Researcher: Alexander PEYSER (JUELICH)

Other Researcher: Raül SIRVENT (BSC)

6.8.2 SGA1 DoA Goals

This Task focuses on providing support for the migration of simulation codes to hybrid and/or accelerator-enabled architectures.

6.8.3 Component Progress

6.8.3.1 Application Software Services

Description of Component

The work on this Component is described in the "NestMC" and "Code Migration Support" components, which are described below in 6.8.3.2 and 6.8.3.3 respectively.

6.8.3.2 NestMC

Description of Component

Simulator designed from the ground up for large networks of multi-compartment neurons on hybrid/accelerated/many-core computer architectures.

CDPs to which the Component contributes: None

Progress on Component

This Component is responsible for the development of the NestMC simulator software. This is progressing well, with our prototype of the software already available in an open repository online: https://github.com/eth-cscs/nestmc-proto

The prototype software will be used as the basis for the first, full relase of NestMC in M24.

Most of the development has been performed by the partners at ETHZ-CSCS and JUELICH-JSC, with a small contribution from BSC. Contributions to NestMC as of M12 can be summarised as follows:

- Prototype library for multi-compartment simulation (CSCS)
- Task-based multithreading support with TBB and HPX (CSCS), pthread pools (JSC) and OpenMP (BSC)
- Concise recipes for describing models (CSCS)
- Implemented back ends for the PCP systems
- The GPU back end has been validated on JURON (not yet optimised) (CSCS)
- The KNL back end has been validated on JULIA (partially optimised) (CSCS & JSC)
- An efficient communication framework for large networks has been implemented (CSCS & JSC)
- Currently being tested and optimised for very large networks on future exascale systems (JSC)
- Developed a "miniapp" for testing and benchmarking the library (CSCS)
- Developed a suite of validation tests to test against known analytic solutions and other simulators (CSCS)
- Exploration of OmpSs support (BSC)

The work has also been disseminated to the broader community by presentations and posters:

- Bernstein Conference 2016, Berlin: workshop presentation and poster.
- HBP Summit 2016, Firenze: poster presentation.
- GPU Hackathon 2017, Jülich: NestMC developers participated as one of the teams

Links

- Poster presented at the Bernstein Conference 2016:

  https://eth-cscs.github.io/nestmc/
- Open repository for NestMC:

  https://github.com/eth-cscs/nestmc-proto

6.8.3.3 Code Migration Support

Description of Component

Support for the migration of simulation codes to hybrid and/or accelerator-enabled architectures.

CDPs to which the Component contributes: None

Progress on Component

This work is being performed by JUELICH. Preliminary benchmarking and application tuning has been completed for Elephant: http://neuralensemble.org/elephant/

Elephant is a Python application that is single threaded, and the developers approached T7.5.4 to help with parallelisation on HPC nodes. A preliminary investigation into the performance and algorithms of the Python implementation was performed. From this it was decided that there was scope for two levels of optimisation:

- Optimization of the Python code using Cython, which gained a 60% speedup.
- Parallelisation using MPI, which has been prototyped.

So far testing has been on a multi-core cluster, with early work on the Intel KNL based PCP system JULIA. In the second half of SGA1 work will be focused on improving the MPI and multithreading support, and optimising for KNL.

6.9 T7.5.5 Simulator NEST as a Service

6.9.1 Key Personnel

Task Leader: Hans Ekkehard PLESSER (JUELICH)

Other Researcher: Dennis TERHORST (JUELICH)

6.9.2 SGA1 DoA Goals

Maintenance of the NEST simulation engine and user support with focus on user level documentation, an increase of the test coverage, and other measures to achieve and maintain the required software quality. Support mechanisms are periodic reviews of the user mailing list, management of the issue tracking software, and the coordination of problem solving with the community. The Task is also concerned with active outreach to promote the use of the simulation code as a service with the aim to increase the effectiveness of research.

6.9.3 Component Progress

6.9.3.1 Components and their Descriptions

Since the Components of this Task are tightly coupled, we report on them together.

A. Community Contacts: Community contacts have been established that are related to the NEST.

B. NEST Support for Modellers: Support for Modellers provides support to groups using NEST as simulation tool, including advice on porting models to and implementing models in NEST.

C. NEST Requirements Management: Requirements management provides a central contact point for scientists requiring extensions and adaptations of the NEST simulator for their work. Requirements are systematized and forwarded to the NEST developers.

D. NEST Support for Providers: Support for Providers provides support to computing centres and other service providers in issues related to installing and providing NEST as a simulation tool.

Effort in these components was mainly by JUELICH with in-kind effort by NMBU and the NEST developer community. Efforts related to Component A contributed significantly to the achievement of MS7.1.1.

CDPs to which the Components contribute:

- CDP1 - Development of Whole Mouse Brain Model and Mouse Brain Atlas (Components A-C)

- CDP4 – Visuo-motor integration (Components A-C)
- CDP5 - Plasticity, Learning and Development: Modelling the Dynamic Brain (Components A-C)

Progress on Components

- Established community contacts in HBP
  - SP2 (A-C), SP3 (A-C), SP4 (A-C), SP6 (A, C, D), SP9 (A, B), SP10 (A-C)
  - CDP1, CDP4, CDP5
  - NEST User Workshop, 3-4 November 2016, FZI Karlsruhe, over 60 participants (A, B)
- Support for model porting (Component B), e.g.,
  - Hill-Tononi network model (SP3: Storm, Massimini)
  - WaveScalES (SP3: Paolucci)
  - Cerebellum models (SP6: d'Angelo)
- NEST User Support via mailing list (all Components):
  - 50% of inquiries answered in less than five hours
  - 75% answered in one day
  - 90% answered within one week
- NEST Community-Based Development (Components A, C, D):
  - 17 Open NEST Developer Video Conferences (new in SGA1, first conference took place on 13 June 2016)
  - Over 80 issues closed, mean 48 days to solution
  - 139 pull requests reviewed and quality controlled, mean 15 days to merge
  - Over 1000 commits to NEST Github repository
  - Over 40 community developers inside and outside the HBP
  - See also https://github.com/nest/nest-simulator

Links

- NEST website:

  http://www.nest-simulator.org/
- NEST repository:

  https://github.com/nest/nest-simulator
- NEST in the HPAC Platform Guidebook:

  https://hbp-hpc-platform.fz-juelich.de/?hbp_software=nest-the-neural-simulation-tool

6.10 T7.5.6 Platform Integration Services

6.10.1 Key Personnel

Task Leader until M09: Daniel MALLMANN (JUELICH)

Task Leader from M10: Bernd SCHULLER (JUELICH)

6.10.2 SGA1 DoA Goals

This Task deals with various aspects of the HPAC Platform integration including the integration of new services with the UNICORE middleware, integration of services with the existing single sign-on (SSO) facilities (currently based on OpenID Connect) and the development of web-based user interfaces and integration with the Collaboratory, including a storage quota manager and storage browser app. This Task also covers the ongoing development of the underlying UNICORE middleware itself, dealing with new requirements by users and other tasks within the project. This will include support for automatic provenance tracking for UNICORE jobs and workflows.

6.10.3 Component Progress

6.10.3.1 UNICORE components

- UNICORE/X
- UNICORE Workflow Framework
- UNICORE Command-line Client
- Transfer service

Description of Components

UNICORE is a general-purpose federation software suite providing seamless access to compute as well as data resources. Within HBP, UNICORE is used as integration layer and interface technology for accessing HPC machines, cloud storage and other resources. UNICORE provides services for job and workflow execution, data access, data upload/download and data transfer.

UNICORE contains several sub-components, which are derived from a common code base:

- UNICORE/X is a server component providing basic services such as HPC job submission;
- The UNICORE Workflow System is a set of services providing workflow enactment and resource brokering;
- Clients such as the UNICORE Command-line Client provide tools for user access to the services independent of or in addition to the Collaboratory.

The UNICORE services provide RESTful APIs, which are the preferred means of interfacing the Collaboratory user interface to the resources in the HPAC Platform.

CDPs to which the Components contribute:

- CDP1 - Development of Whole Mouse Brain Model and Mouse Brain Atlas
- CDP5 - Plasticity, Learning and Development: Modelling the Dynamic Brain.

Progress on Components

The work was focused on the UNICORE RESTful APIs. In the core server (UNICORE/X) Component, the implementation of the features related to data and metadata management was finished. In the Workflow Framework, a RESTful API for submitting and managing workflows was implemented. These features were released in UNICORE version 7.7.0.

To ease the use of the RESTful APIs and thus simplify and improve the integration of the HPAC Platform with the Collaboratory, development of a Python client library has started on GitHub, and a first version is already available. This library includes helper functions for job submission and management, as well as functions for file upload and download.

The development work is done by JUELICH, while important user feedback, requests for enhancement and testing was done by ETHZ.

Links

- UNICORE 7.7.0 core servers release:
  https://sourceforge.net/projects/unicore/files/Servers/Core/7.7.0/
- UNICORE 7.7.0 workflow framework release:
  https://sourceforge.net/projects/unicore/files/Servers/Workflow/7.7.0/
- UNICORE 7.7.0 command-line client release:
  https://sourceforge.net/projects/unicore/files/Clients/Commandline%20Client/7.7.0/
- REST API documentation:
  https://hbp-hpc-platform.fz-juelich.de/?page_id=1153
- UNICORE Python client library:
  https://sourceforge.net/p/unicore/wiki/REST_API/
- Repository of the Python client library:
  https://github.com/HumanBrainProject/pyunicore

6.10.3.2 Upload/Download Services

See also section 6.6.3.4

Description of Components

These services enable users to upload/download datasets stored in the federated data infrastructure given that they have the necessary access permissions.

Upload and download services are currently realised using both the built-in UNICORE file transfer tools (based on the default HTTPS transport that UNICORE uses) and the UNICORE File Transfer Protocol (UFTP) which is an add-on that provides high-performance data transfer. This section describes the work done in UFTP.

CDPs to which Components contributes: None

Progress on Components

UFTP consists of a server component (uftpd) which provides access to a file system, and a client library that is used by UNICORE tools as well as a standalone client application.

The uftpd server was released in version 2.4.0 with a number of changes and enhancements. The ftp-like protocol for controlling the data transfer has been made more compliant to the usual FTP protocol. This allows using standard Unix tools (e.g. curl or the ftp program) for a data transfer.

The development work is done by JUELICH, while user feedback, requests for enhancement and testing was provided by other sites in the HPAC Platform, notably ETHZ-CSCS and CINECA.

Links

- UFTPD 2.4.1 release:
  https://sourceforge.net/projects/unicore/files/Servers/UFTPD/2.4.1/

6.11 T7.5.7 User Support Services

6.11.1 Key Personnel

Task Leader: David VICENTE (BSC)

Other Researcher: Cristian MEZZANOTTE (ETHZ-CSCS)

Other Researcher: Cristiano PADRIN (CINECA)

Other Researcher: Rajalekshmi DEEPU, Anna LÜHRS (JUELICH-JSC)

6.11.2 SGA1 DoA Goals

This Task is responsible for the coordination and implementation of an effective ticketing system, as well as user support for users of the HPAC Platform, federated across the participating data centres (BSC, CINECA, ETHZ-CSCS, JUELICH-JSC). Another important activity is to provide support to projects that wish to apply for resources allocations within the HPAC Platform. Coordination of the documentation and user engagement related to the HPAC Platform services are undertaken. To this end, the Task is responsible for maintaining a repository of links to applicable documentation, including underlying APIs and adopted standards.

6.11.3 Component Progress

6.11.3.1 HPAC Platform Guidebook

Description of Component

General information, technical and user documentation, contact forms, etc. of the HPAC Platform

CDPs to which the Component contributes: None

Progress on Component

The Component is managed by JUELICH. It provides general information, documentation and the training programme of the HPAC Platform in the HPAC Platform Guidebook, which is continuously updated. Relevant news and information are also tweeted through the HPAC Platform Twitter channel @HBPHighPerfComp, i.e. announcement of meetings, major software releases etc.

Links

- HPAC Platform Guidebook:

  https://collab.humanbrainproject.eu/#/collab/264/nav/2378

  https://hbp-hpc-platform.fz-juelich.de/

6.11.3.2 User Support

Description of Components

- HPAC User Support Services: Infrastructure to support users of the HPAC Platform and SP7 as a whole. This in particular also includes supporting users in the process of applying for computing time allocations.
- CSCS User Support Service: CSCS User Support Service
- JSC User Support Service: JSC User Support Service
- CINECA User Support Service: CINECA User Support Service
- BSC User Support Service: BSC User Support Service

CDPs to which the Components contribute: None

Progress on Components

The Component HPAC User Support Service is managed by BSC, and the other Components are managed by the respective sites (CINECA, ETHZ, JUELICH and BSC).

The current HPAC Platform user support infrastructure is full deployed and in production. Also, all other contact forms on the HBP websites for submitting queries to the HPAC Platform support have been connected with our ticketing system (see below) to receive all the questions of the HPAC Platform users or future users. With that, Milestone MS7.5.4 (Full deployment of user support infrastructure) was reached earlier than scheduled. The current system is already receiving tickets from users. In a first instance the tickets are managed by the first level support (provided by BSC) and when needed, the tickets are moved to the second level support provided by each of the sites involved in the Task.

Links

- HPC access and support request form:

  https://collab.humanbrainproject.eu/#/collab/264/nav/6676

- HPAC Platform contact form:

  https://collab.humanbrainproject.eu/#/collab/264/nav/6677

6.11.3.3 HPAC Ticket System

Components described in this section:

- G: HPAC Ticket System: HPAC Ticket System
- H: CSCS Ticket system: CSCS Ticket system
- I: JSC Ticket System: JSC Ticket system
- J: CINECA Ticket System: CINECA Ticket system
- K: BSC Ticket System: BSC Ticket system

CDPs to which the Components contribute: None

Progress on Components

The Component HPAC Ticket System is managed by BSC, and the other Components are managed by the respective sites (CINECA, ETHZ, JUELICH and BSC).

During this first twelve months of SGA1 the main goal was to define and deploy the support structure for all the HPC sites. The current Component released by BSC has been the HPAC Ticketing System for the first level support of HPAC Platform and the connection with all the second level support sites providing compute resources in the HPAC Platform. The HPAC ticketing system is connected with the central HBP OIDC system used for the Collaboratory, so all the HBP users can use the same account and password for the ticketing system as for the other HBP services. BSC has written a user guide for all users of the ticketing system in order to use it in a proper and efficient way and also to inform the first and second level support about the guidelines for the user support service.

Links

- HPAC ticketing system manual and guidelines for first and second level support for HBP HPAC Platform:
  https://emdesk.humanbrainproject.eu/shared/5a4f367b1e037-0aec2f335247fdca30137e83cdcd1e34

7. WP7.6 Management and Coordination

7.1 Key Personnel

Work Package Leader: Thomas LIPPERT (JUELICH)

Work Package Co-Leader: Boris ORTH (JUELICH)

7.2 WP Leader's Overview

Work in WP7.6 went very well in the first year of SGA1. WP7.6 had no large deviations and did not face any significant difficulties in the first year of SGA1. The four Tasks closely collaborated to manage and coordinate SP7, to establish and maintain contacts and coordinate the collaboration with the other HBP Subprojects. Members of WP7.6 participated in all coordination committees of the HBP to actively shape the HBP and to contribute to moving it forward. The SP7 Managers were in regular contact with the other SP Managers and Coordinators and closely collaborated with them at working level in order to support the HBP scientists in reaching important milestones of the HBP. The SP7 Managers and Leaders are the linking element between SP7 members, the HBP governance bodies and the other Subprojects. The contacts of the SP7 management and leadership with the SP7 partners already established in the Ramp-Up Phase are very close. The SP7 members approach their SP management in case of questions, to identify relevant contacts in other SPs or if they need other kinds of support. This very good relationship between all SP7 partners is the basis of a collaborative and productive working atmosphere and key to jointly achieving the common goals.

7.3 Priorities for the remainder of the phase

Apart from the regular activities to manage and coordinate SP7, T7.6.1 and T7.6.2 will focus in the next months on the planning for SGA2, both SP7-internally and in close collaboration with the other Subprojects, in particular with the Neuroinformatics Platform. An important task in this context is the lead and coordination of a working group that aims at defining a High-Level Support Team for the Human Brain Project infrastructure, that should start its activities with the beginning of SGA2. T7.6.3 will implement the co-design knowledge management process defined in the first year of SGA1. T7.6.4 will continue to systematically collect use cases from HBP users and to validate the HPAC Platform against them.

7.4 Milestones

Table 6: Milestones for WP7.6 Management and coordination

MS No.	Milestone Name	Lead Partner	Task(s) involved	Expected Month	Achieved Month	Comments
MS7.6.1	Co-design knowledge transfer process established	P20 JUELICH	T7.6.3	M06	M08	The co-design knowledge management process has been defined and is described in brief in an internal report which is available in the HPAC Platform Collab: https://emdesk.humanbrainproject.eu/shared/59b10b547b54e-5ed9944cffd9549e0556bc242d39f9c5 In the context of defining this process we propose a slight change of the subtask organisation without changing the content of this task: A. Augment use case descriptions for co-design B. Select, implement and deploy mini-applications or application benchmarks C. Architectures and technologies roadmap assessment D. Setup and maintain technology test-beds
MS7.6.2	HPAC Platform validation use cases defined	P20 JUELICH	T7.6.4	M06	M06	T7.6.4 defined initial use cases for validating the HPAC Platform with users from CDP1, CDP3, CDP5 and the NEST-Elephant-SpiNNaker workflow. The set of use cases is likely to be complemented with additional use cases in the future as other parts of the HBP evolve. The use cases are described in internal reports and also in excerpts in D7.6.1. Use cases related to SP5 and visualisation will be collected and analysed in the next weeks, since we had to wait for the DPIT process to be finished. The internal reports are available here: CDP1: https://emdesk.humanbrainproject.eu/shared/59b77f3654d33-a23b47918ed56964d1a937e8cfc788ad

						CDP3: https://emdesk.humanbrainproject.eu/shared/59b77f3c30e62-7f6fd62c8af32c3f6f150740e68ae0c1 CDP5: https://emdesk.humanbrainproject.eu/shared/59b77f4115d49-ce4141dfe60b3fe22194be68530e8cea (NM UMAN) https://emdesk.humanbrainproject.eu/shared/59b77f2bc29fc-0b6c5523fa534096b6500e7a96fcae4b (NM UHEI) https://emdesk.humanbrainproject.eu/shared/59b77f265d155-5ae7a24476bd6754a3aba6855d2cf9e4 (NM software) https://emdesk.humanbrainproject.eu/shared/59b77f30c6a15-ce3982026f8d0dd564b309ea2be926af (NEST) See also Annex B: Use cases and quantitative requirements analysis for more details.
MS7.6.3	SP7 Roadmap for SGA2	P20 JUELICH	T7.6.1, T7.6.2	M13	M13	The SP7 Roadmap for SGA2 is dominated by the planning of the infrastructure that will be created as part of the ICEI proposal, assuming the funding is approved. ICEI will largely take-over the responsibility for putting infrastructure services in place, while SP7 will focus on further development, deployment, operation and enhancement of platform services as well as the exascale enablement of brain simulations. The SP7 work plan for SGA2 has been developed and provided for the SGA2 proposal.

7.5 T7.6.1 Subproject management

7.5.1 Key Personnel

Task Leader: Thomas LIPPERT (JUELICH)

Deputy Task Leader: Boris ORTH (JUELICH; SP7 Manager)

Other Researcher: Anna LÜHRS (JUELICH; SP7 Manager)

Other Researcher: Meredith PEYSER (JUELICH; SP7 Manager)

WP Leader:

Other Researcher: Dirk PLEITER (JUELICH)

Other Researcher: Markus DIESMANN (JUELICH)

Other Researcher: Torsten KUHLEN (RWTH)

Other Researcher: Thomas SCHULTHESS (ETHZ)

Other Researcher: Raül SIRVENT (BSC)

7.5.2 SGA1 DoA Goals

T7.6.1 is responsible for the management and coordination of SP7, in particular for

- Resource allocation and use,
- Performance and risk management,
- Internal review and quality control,
- Reporting and Deliverables,
- Internal communication within SP7, between SP7 and the central project coordination offices and other relevant HBP bodies,
- Outreach and dissemination,
- Planning of the next project phase, and
- Participation in the Ethics Rapporteur Programme.

7.5.3 Component Progress

7.5.3.1 SP7 management and coordination

Description of Component

Management and coordination of SP7, the High Performance Analytics & Computing Platform

CDPs to which the Component contributes: None

Progress on Component

This Component summarises all activities related to the management and coordination of the High Performance Analytics & Computing Platform Subproject (SP7). The work is mainly performed by the SP7 Managers (JUELICH); strategic decisions are made by the SP Leaders (JUELICH, ETHZ). The WP leaders (BSC, ETHZ, JUELICH, RWTH) are involved in the reporting and they contribute to Deliverables, dissemination, outreach and to the planning of the next project phase. SP7 Managers and WP leaders are in close contact with each other to ensure an efficient information exchange.

The most important activities in the first twelve months of SGA1 were:

- Collecting information from SP7 for and writing of Deliverable D7.6.1 (submitted in M6) in close collaboration with T7.6.2

- Supporting SP7 members in the definition of their components for the Project Lifecycle App, and integrating them into this system
- Collecting information from SP7 for and writing of the SP7 section of the Semester Report covering M1-M6 of SGA1
- Organising quarterly SP7 videoconferences and of the SP7 kick-off meeting for SGA1, that took place 19 May 2016 in Frankfurt, Germany
- Progress monitoring through regular contact with the WP leaders
- Dissemination and outreach: Preparation of material for and presenting SP7 at
  - HBP Summit and Open Day 2016, 12-15 October 2016, Florence, Italy
  - Society for Neuroscience Annual Meeting, 12-16 November 2016, San Diego, California, USA
  - STOA Exhibition at the European Parliament, 29-20 November 2016, Brussels, Belgium
  - Annual Meeting of the US Brain Initiative, 12-14 December 2016, Bethesda, Maryland, USA
- Maintaining and updating the SP7-related Collabs and the general parts of the HPAC Platform Guidebook
- Disseminating news, upcoming events and other relevant information through the HPAC Platform Twitter channel
- Collecting and writing contributions of SP7 for the HBP Newsletter
- The planning for the next project phase (SGA2) started in December 2016 and is coordinated by the SP7 Managers
- Collecting of internal reports and documentation, and making these documents available in the HPAC Platform Collab

The SP7 Managers are in regular contact with the central project coordination offices and other relevant bodies. They also act as Ethics Rapporteurs for SP7 and participate in the different HBP-wide coordination committees.

Links

- HPAC Platform Twitter Channel: @HBPHighPerfComp
- HPAC Platform Guidebook: https://collab.humanbrainproject.eu/#/collab/264/nav/2378

  or

  https://hbp-hpc-platform.fz-juelich.de/
- HPAC Platform Collab: https://collab.humanbrainproject.eu/#/collab/264/nav/1973

7.6 T7.6.2 Technology and infrastructure coordination

7.6.1 Key Personnel

Task Leader: Colin MCMURTRIE (ETHZ)

Other Researcher: Cristian MEZZANOTTE (ETHZ)

Other Researcher: Anna LÜHRS (JUELICH)

Other Researcher: Boris ORTH (JUELICH)

Other Researcher: Dirk PLEITER (JUELICH)

Other Researcher: Björn HAGEMEIER (JUELICH)

Close contact with:

Jeff MULLER (EPFL-PCO)

Benjamin WEYERS (RWTH)

Bernd STAHL (DMU)

Jan BJAALIE (UIO)

7.6.2 SGA1 DoA Goals

T7.6.2 coordinates the technology and infrastructure development in SP7. The responsibilities of the SP7 Technology and Infrastructure Coordinator include the:

- Assessment of TRLs of software and services;
- Coordination of infrastructure planning, including sizing estimates and justifications;
- Monitoring and reporting of technological and infrastructure progress to the HBP Science and Technology Coordination;
- Documentation and dissemination of the HPAC Platform's technical and operational standards;
- Testing and quality control of technologies and infrastructure developed in SP7;
- Implementation of infrastructure components according to the project lifecycle as described in the FPA;
- Representation of the SP on the Technical Coordinators Committee and the Infrastructure Coordinators Committee;
- Interaction with European infrastructure providers if necessary, coordinating this work with the HBP Technical Coordinator and the SP Software Coordinator;
- Aligning of SPs technical activities on-going technical developments, HBP external emerging trends and the HBP Co-Design Projects;
- Attendance of conferences/events on behalf of the HBP, participation in HBP booths.

7.6.3 Component Progress

7.6.3.1 Components and their Description

- A: SP7 technology & infrastructure coordination: Coordination of the technology and infrastructure development in SP7, see the SGA1 goals above for more details.
- B: HBP Software Engineering and Quality Assurance Approach: WPs involved: WP11.2, WP11.3, WP5.6, WP6.5, WP7.5, WP8.6, WP9.5 & WP10.7. This will describe the overall software engineering and quality assurance approach (covering both agile and co-design processes) This will update related content in System Engineering Package sent by HBP Project Coordination Office (PCO; Jeff Muller) to the EC in March 2016 and address feedback points received from the EC.
- C: HBP IT Architecture: Overall HBP IT architecture focusing on base infrastructure provided for data repositories with a discussion of how this infrastructure will support novel neuroscience workflows.

CDPs to which the Components contribute: None

Progress on Components

- Participation in the Platform Coordination meetings led by the HBP PCO;
- Helped set up the Software Development Committee (SDC) including helping write the remit approved by Software Development Director and the HBP SIB. Participation in the meetings thereof;
- Participation in the Data Governance WG (DGWG) lead by Bernd Stahl, the HBP Ethics Director, including contributions to the Policy Manual;
- Advocated for and helped set up the Infrastructure Development Committee (IDC) including helping write the related remit. Participation in the meetings thereof;
- Co-managed FeDaPP and lead of WP6 within FeDaPP (pilot project for Fenix);
- Involved in the planning of Fenix (which forms an integral part of Component C, above);
- Leadership role in the planning and management of the migration of the Collaboratory DevOps and production infrastructure to CSCS OpenStack IaaS;
- Met with EGI to discuss a potential interaction between SP7 and FeDaPP (held at JUELICH-JSC);
- Attended HBP Open Day at the SP7 booth;
- Participation in DPIT, WG, PCP & EC meetings;
- Contributed to Deliverable D7.6.1, in particular related to the HPAC Platform architecture description;
- Participation and contributions to the HBP SGA ICEI proposal to the EC;
- Participation and contributions to the SGA2 planning, including the Science Planning meeting in Malaga (19-22 March 2017);
- Participation in the Software Planning meeting in Düsseldorf (15 March 2017).

7.7 T7.6.3 Co-design knowledge management

7.7.1 Key Personnel

Task Leader: Dirk PLEITER (JUELICH)

Other Researcher: Colin MCMURTRIE (CSCS)

Other Researcher: Fabien DELALONDRE (EPFL)

7.7.2 SGA1 DoA Goals

The main goal of this Task is to facilitate the sharing of knowledge with external stakeholders. In this context, relevant external stakeholders are developers of relevant HPC technologies and architectures. Until M10 the parallel running HBP PCP was used for this purpose. After completion of the PCP this task will take-over this role.

7.7.3 Component Progress

7.7.3.1 SP7 co-design knowledge management

Description of component

Management of knowledge created within the Subproject that is relevant for co-design, in order to facilitate the sharing of this knowledge with external stakeholders.

CDPs to which the Component contributes: None

Progress on Component

The process for the knowledge management has been defined.

Links

Description of Co-Design Knowledge Management Process:

https://emdesk.humanbrainproject.eu/shared/59b10b547b54e-5ed9944cffd9549e0556bc242d39f9c5

7.7.3.2 PCP Pilot systems

Description of component

Pilot systems delivered by Cray and IBM-NVIDIA in the third phase of the Pre-Commercial Procurement (WP7.1 in the RUP). The systems are hosted at JSC.

CDPs to which the Component contributes: None

Progress on Component

The systems have been deployed in M6 and M7. The contacts with the vendors will be maintained by this Task after the end of the Pre-Commercial Procurement; the integration of the systems into the HPAC Platform was done by SGA1-WP7.5. Maintenance contracts with both Cray and IBM for about two years after the end of the PCP are being prepared by T7.6.3.

Links

- HPAC Platform Guidebook page about JULIA (Cray pilot):

  https://hbp-hpc-platform.fz-juelich.de/?page_id=1063
- HPAC Platform Guidebook page about JURON (IBM-NVIDIA pilot):

  https://hbp-hpc-platform.fz-juelich.de/?page_id=1073
- General information, wiki and information on getting access:

  https://trac.version.fz-juelich.de/hbp-pcp/wiki

7.8 T7.6.4 Platform validation

7.8.1 Key Personnel

Task Leader: Alexander PEYSER (JUELICH)

Other Researcher: Vicente MARTÍN (UPM)

Other Researcher: Jean-Denis COURCOL (EPFL)

7.8.2 SGA1 DoA Goals

This Task takes care of

- Ensuring the usability of the HPAC Platform for the concerned research area;
- Establishing a validation process for a selected set of relevant use cases; and
- Validation by implementing end-to-end solutions for these use cases.

The work is organised in the following sub-tasks:

1) Selection of initial use cases and detailed definition of requirements
2) Implementation of use cases on the HPAC Platform
3) Evaluation of the results

7.8.3 Component Progress

7.8.3.1 HPAC Platform Validation

Description of Component

SP7 will develop and implement a variety of ICT technologies with the goal of enabling neuroscience research by integrating these technologies into an ICT platform, which also comprises architectures and components from other ICT solution providers, such as providers of supercomputing solutions. T7.6.4 has the goal of ensuring the usability of the HPAC Platform for the concerned research area. The strategy for achieving this goal is to establish a validation process for a selected set of relevant use cases.

Validation is achieved by implementing end-to-end solutions for these Use Cases. This approach will not only stimulate the use of the Platforms by the neuroscience Subprojects, but also provide important guidance for the R&D activities within SP7 as well as the ongoing build-up and operation of the HPAC Platform.

The work will be organised in the following sub-tasks:

1) Selection of initial use cases and detailed definition of requirements;
2) Implementation of use cases on the HPAC Platform;
3) Evaluation of the results.

CDPs to which the Component contributes:

- CDP1: Development of Whole Mouse Brain Model and related Mouse Brain Atlas
- CDP3: Multi-Level Human Brain Atlas
- CDP5: Functional Plasticity for Learning in large-scale Systems
- NEST-Elephant-SpiNNaker Workflow

Progress on Component

JUELICH collected and analysed use cases emerging from CDPs 1, 3 and 5 and the NEST-Elephant-SpiNNaker workflow.

- Reports on use cases written for CDPs 1, 3, 5 and for NEST-Elephant-SpiNNaker Workflow:
  - CDP1: https://emdesk.humanbrainproject.eu/shared/59b77f3654d33-a23b47918ed56964d1a937e8cfc788ad
  - CDP3: https://emdesk.humanbrainproject.eu/shared/59b77f3c30e62-7f6fd62c8af32c3f6f150740e68ae0c1
  - CDP5: https://emdesk.humanbrainproject.eu/shared/59b77f4115d49-ce4141dfe60b3fe22194be68530e8cea (NM UMAN)
  - https://emdesk.humanbrainproject.eu/shared/59b77f2bc29fc-0b6c5523fa534096b6500e7a96fcae4b (NM UHEI)
  - https://emdesk.humanbrainproject.eu/shared/59b77f265d155-5ae7a24476bd6754a3aba6855d2cf9e4 (NM software)
  - https://emdesk.humanbrainproject.eu/shared/59b77f30c6a15-ce3982026f8d0dd564b309ea2be926af (NEST)
- Scaling report written for NestMC: https://emdesk.humanbrainproject.eu/shared/5a4f3e7b87ace-80229786a8331be0234c50cbc1d971b9

UPM is concerned with HPAC Platform validation related to visualisation. UPM focused on remote visualisation and the design of metrics for visual applications developed in the HBP. The validation process will be centred on real Use Cases. UPM has successfully ported the visualisation tools over to both PCP pilot systems (JURON and JULIA), and VirtualGL as well as TurboVNC are used for remote visualisation. Due to the new architectures of the machines, UPM will have to invest some effort in tuning the applications for these platforms.

T7.5.4 delivered code for the validation against the PCP pilots and related scaling tests. This activity is still on-going. A report on the scaling behaviour for up to 28,000 nodes and suggestions for technical improvements has been delivered.

T7.6.2 delivered technical requirements for the definition of use cases. As the Use Cases are added and the FeDaPP/Fenix infrastructure is defined, T7.6.4 will continue to extend and modify the use case templates for validation.

T7.6.1 gave incentives to T7.6.4 regarding additional topics and groups that potentially need to be additionally targeted by use cases. Both Tasks are in regular contact with each other.

EPFL has not contributed to this Task in the first year of SGA1.

8. Publications

The contributing HBP-SP7 members are highlighted using **bold** font.

WP7.1

- **Hahne, J.**, Helias, M.,Kunkel, S., Igarashi, J., Kitayama, I., Wylie, B., **Bolten, M.**, **Frommer, A.**, **Diesmann, M.** Including Gap Junctions into Distributed Neuronal Network Simulations. Brain-Inspired Computing / Amunts, Katrin (Editor) ; Cham : Springer International Publishing, 2016, Chapter 4 ; ISSN: 0302-9743=1611-3349 ; ISBN: 978-3-319-50861-0=978-3-319-50862-7 ; doi:10.1007/978-3-319-50862-7. International Workshop on Brain-Inspired Computing, BrainComp 2015, Cetraro, Italy, 6 Jul 2015 - 10 Jul 2015. Lecture Notes in Computer Science 10087, 43 - 57 (2016) [10.1007/978-3-319-50862-7_4].
- **Schuecker, J.**, van Albada, S.J., **Diesmann, M.**, Helias, M., Schmidt, M. Fundamental Activity Constraints Lead to Specific Interpretations of the Connectome. PLoS Computational Biology 13(2), e1005179 - (2017) [10.1371/journal.pcbi.1005179].
- Martínez-Cañada, P., Morillas, C., **Plesser**, **H.E.,** Romero, S., Pelayo, F. Genetic algorithm for optimization of models of the early stages in the visual system. Neurocomputing (2017), online first. http://dx.doi.org/10.1016/j.neucom.2016.08.120.
  - *This publication covers results of the Ramp-Up Phase.*
- **Dahmen, D**., **Bos, H**., Helias, M. Correlated Fluctuations in Strongly Coupled Binary Networks Beyond Equilibrium. Physical review / X 6(3), 031024 (2016) [10.1103/PhysRevX.6.031024].
- Torre, E., **Quaglio, P.**, Denker, M., Brochier, T., Riehle, A., Grun, S. Synchronous Spike Patterns in Macaque Motor Cortex during an Instructed-Delay Reach-to-Grasp Task. The Journal of Neuroscience / Rapid communications 36(32), 8329 - 8340 (2016) [10.1523/JNEUROSCI.4375-15.2016].
- Maksimov, A., van Albada, S., **Diesmann, M.** [Re] Cellular And Network Mechanisms Of Slow Oscillatory Activity (<1 Hz) And Wave Propagations In A Cortical Network Model. Zenodo 2(1), 1-18 (2016) [10.5281/zenodo.161526].
- **Bos, H., Diesmann, M**., Helias, M. Identifying Anatomical Origins of Coexisting Oscillations in the Cortical Microcircuit. PLoS Computational Biology 12(10), e1005132 - (2016) [10.1371/journal.pcbi.1005132].
- Hagen, E., **Dahmen, D.**, Stavrinou, M.L., Lindén, H., Tetzlaff, T., van Albada, S., Grün, S., **Diesmann, M**., Einevoll, G.T. Hybrid Scheme for Modeling Local Field Potentials from Point-Neuron Networks. Cerebral cortex 26(12), 4461-4496 (2016) [10.1093/cercor/bhw237].
- Ippen, T., Eppler, J.M., **Plesser, H.E., Diesmann, M.** Constructing neuronal network models in massively parallel environments. Frontiers in Neuroinformatics 11:30 (2017) [doi: 10.3389/fninf.2017.00030].
  - Related to T6.3.5 with contributions of T7.1.4
- Posters
  - **Blundell, I.** , Plotnikov, D. , Eppler, J.M., Rumpe, B., Morrison, A. NESTML- A modeling language for NEST. IAS-Symposium, Juelich, Germany, 5 Dec 2016 - 6 Dec 2016.

- **Yegenoglu, A., Senk, J**., Amblet, O., Brukau, Y., Davison, A., Lester, D., **Lührs, A.**, Quaglio, P., Rostami, V., Rowley, A., **Schuller, B.**, Stokes, A., van Albada, S., Zielasko, D., **Diesmann, M.**, **Weyers, B**., Denker, M., Grün, S. Embedding Elephant in a Simulation-Validation Workflow within the HBP Collaboratory. Human Brain Project Summit 2016, Florence, Italy, 12 Oct 2016 - 15 Oct 2016.

  WP7.3, WP7.5 and WP7.6 were also involved in this poster.

- **Quaglio, P.**, Torre, E., Denker, M., Brochier, T., Riehle, A., Grün, S. DETECTION OF SPIKE-SYNCHRONY PATTERNS IN MASSIVELY PARALLEL SPIKE TRAINS. 2016 HBP Summit, Florence, Italy, 11 Oct 2016 - 15 Oct 2016.

- **Quaglio, P.**, Torre, E., Denker, M., Brochier, T., Riehle, A., Grün, S. Behavior specific spike patterns in macaque motor cortex during an instructed-delay reach-to-grasp task. IAS symposium, Juelich, Germany, 5 Dec 2016 - 6 Dec 2016.

WP7.2

- **Rinke, S**., Naveau, M., **Wolf, F**., Butz-Ostendorf, M. (June 2017). Critical periods emerge from homeostatic structural plasticity in a full-scale model of the developing cortical column. In A. van Ooyen & M. Butz-Ostendorf (Eds.), *The Rewiring Brain - A Computational Approach to Structural Plasticity in the Adult Brain*. Academic Press. ISBN: 9780128037843.
  - Related Task: T7.2.3
- Spitzer, H., Amunts, K., Harmeling, S., **Dickscheid, T**. (2017) Parcellation of visual cortex on high-resolution histological brain sections using convolutional neural networks. *IEEE International Symposium on Biomedical Imaging, Melbourne.* In press.
- Bludau, S., **Dickscheid, T.**, Iannilli, F., Amunts, K. (2016) 3D-reconstruction of cell distributions in the human subthalamic nucleus at 1 micron resolution. Annual Meeting of the Organization of Human Brain Mapping (OHBM).
- Spitzer, H., Stibane, D., Caspers, S., Zilles, K., Amunts, K., **Dickscheid, T.** (2016) Feasibility of deep learning for automatic parcellation of cortical regions in histological sections. Annual Meeting of the Organization of Human Brain Mapping (OHBM).

Related publications (without HBP acknowledgement)

- **Rinke**, R., Butz-Ostendorf, M., Hermanns, M.-A., Naveau, M., **Wolf, F.** A Scalable Algorithm for Simulating the Structural Plasticity of the Brain. In *Proc. of the 28th International Symposium on Computer Architecture and High Performance Computing (SBAC-PAD)*, Los Angeles, CA, USA, pages 1-8, October 2016. Doi: https://doi.org/10.1109/SBAC-PAD.2016.9.
  - Related Task: T7.2.3

WP7.3

- Hänel, C., **Weyers**, B., Hentschel, B., **Kuhlen**, T.W. "Visual Quality Adjustment for Volume Rendering in a Head-Tracked Virtual Environment," in IEEE Transactions on Visualization and Computer Graphics, vol. 22, no. 4, pp. 1472-1481, April 21 2016. doi: 10.1109/TVCG.2016.2518338.
- Zielasko, D., Horn, S., Freitag, S., **Weyers, B.**, **Kuhlen, T.W.** (2016, March). Evaluation of hands-free HMD-based navigation techniques for immersive data analysis. In 3D User Interfaces (3DUI), 2016 IEEE Symposium on (pp. 113-119). IEEE.
- Hänel, C., Khatami, M., **Kuhlen, T.W.**, **Weyers**, B. "Towards Multi-user Provenance Tracking of Visual Analysis Workflows over Multiple Applications", in Proceedings of

EuroVis Workshop on Reproducibility, Verification, and Validation in Visualization (EuroRV3), 2016.

- Zielasko, D., **Weyers, B**., Hentschel, B., **Kuhlen, T.W.** (2016, June). Interactive 3D Force-Directed Edge Bundling. In Computer Graphics Forum (Vol. 35, No. 3, pp. 51-60).
- **Vierjahn**, T., Zielasko, D., van Kooten, K., Messmer, P., Hentschel, B., **Kuhlen**, T.W., **Weyers, B.** Towards a Design Space Characterizing Workflows That Take Advantage of Immersive Visualization In: IEEE Conf. Virtual Reality 2017 — Posters. Accepted for publication.
- Galindo, S.E., **Toharia, P.**, **Robles, O.D.**, **Pastor, L**. ViSimpl: Multi-View Visual Analysis of Brain Simulation Data, Frontiers in Neuroinformatics Date: 07.10.2016, Article 44.
- **Garcia-Cantero**, **J.J., Brito, J.P., Mata, S., Bayona, S., Pastor, L.** Neuron Mesh Generation and Adaptive On The Fly Refinement. Frontiers in Neuroinformatics. *In Press.*

WP7.4

- Lopez, V., Jokanovic, A., D'Amico, M., Garcia, M., Sirvent, R., Corbalan, J. "DJSB: Dynamic Job Scheduling Benchmark", 21st Workshop on Job Scheduling Strategies for Parallel Processing (JSSPP17) (in conjunction with IPDPS 2017).
- Artigues, A., **Cugnasco, C., Becerra, Y.,** Cucchietti, F., Houzeaux, G., Vazquez, M., Torres, J., Ayguadé, E., Labarta, J. "ParaView + Alya + D8tree: Integrating High Performance Computing and High Performance Data Analytics". Submitted to the International Conference on Computational Science 2017 (ICCS '17).
- Submitted:
  - **Valero-Lara, P.,** Martínez-Pérez, I., **Sirvent, R.,** Martorell, X., Peña, A.J."NVIDIA GPUs Scalability to Solve Multiple (Batch) Tridiagonal Systems. Implementation of cuThomasBatch". International European Conference on Parallel and Distributed Computing (Euro-Par 2017). Submitted.
  - **Cugnasco, C., Becerra, Y.,** Torres, J. Ayguadé, E. "Exploiting key-value data stores scalability for HPC". Submitted to the 46th International Conference on Parallel Processing (ICPP '17).

WP7.5

- Benedyczak, K., **Schuller, B**., Petrova-El Sayed, M., Rybicki, J., Grunzke, R. UNICORE 7 — Middleware services for distributed and federated computing. In IEEE International Conference on High Performance Computing & Simulation (HPCS) 2016. 18-22 July 2016. DOI: 10.1109/HPCSim.2016.7568392.
- *Accepted poster*:
  Valero-Lara, P., Martínez-Pérez, I., Peña, A.J., Martorell, X., **Sirvent, R.,** Labarta, J. "Simulating the Behavior of the Human Brain on NVIDIA GPUs (Human Brain Project). cuHinesBatch Implementation", GPU Technology Conference 2017, San Jose (USA).
- **Valero-Lara**, **P.,** Martínez-Pérez, I., Peña, A.J., Martorell, X., **Sirvent, R.,** Labarta, J. "cuHines: Solving Multiple (Batched) Hines systems on NVIDIA GPUs", ICCS 2017. International Conference on Computational Science (ICCS 2017).

WP7.6

- **Lippert, T., Lührs, A., McMurtrie, C., Orth, B., Pleiter, D., Schulthess, T**. Human Brain Project: Towards a European infrastructure for brain research. Innovatives Supercomputing in Deutschland 14(2), 88-92 (2016)
- Amunts, K., Ebell, C., Muller, J., Telefont, M., Knoll, A., **Lippert, L.** The Human Brain Project: Creating a European Research Infrastructure to Decode the Human Brain. Neuron 92(3), 574 - 581 (2016) [10.1016/j.neuron.2016.10.046].

9. Dissemination

The contributing HBP-SP7 members are highlighted using **bold** font.

WP7.1

- **Quaglio, P**. Detection of spatio-temporal spike patterns in massively parallel electrophysiological recordings. Seminar, Italy, 21-22 December 2016.
- **Blundell, I.**, Eppler, J.M., Plotnikov, D., Morrison, A. Code Generation from model description languages II. Seminar, Germany.
- **Senk, J., Yegenoglu, A.**, Amblet, O., Brukau, Y., Davison, A., Lester, D., **Lührs, A**., **Quaglio, P.**, Rostami, V., Rowley, A., **Schuller, B.**, Stokes, A., van Albada, S., Zielasko, D., **Diesmann, M**., **Weyers, B**., Denker, M., Grün, S. Integrating HPC into a Collaborative Simulation-Analysis Workflow for Computational Neuroscience. JARA-HPC Symposium, Aachen, Germany, 4-5 October 2016.
  - WP7.3, WP7.5 and WP7.6 also contributed to this activity.
- **Jordan, J.**, Kunkel, S., Ippen, T., Helias, M., **Diesmann, M.**, Morrison, A. NEST 5g: a simulation kernel for the exascale. Seminar, Germany.
- **Plesser, H.E**. The Human Brain Project Research Infrastructure. Technologies for Digital Life Workshop, Centre for Digial Life Norway, Bergen, Norway, 21 October 2016.
- **Plesser, H.E**. NEST, HBP, & Large-Scale Simulation. Seminar lecture, Center for Integrative Neuroplasticity, University of Oslo, Norway, 31 August 2016.
- **Quaglio, P.** Detection of Spatio-Temporal patterns in massively spiketrains. Seminar, France.

WP7.2

- **Pleiter, D.** "Human Brain Project Unifying our understanding of the human brain", keynote at EaPEC2016 conference, Tiblisi (Georgia), October 2016.
- "Human Brain Project: Pilot Systems for Interactive Supercomputer Launched"/ "Human Brain Project: Pilotsysteme für interaktiven Superrechner gestartet". Press release by JUELICH on 28 September 2016. http://www.fz-juelich.de/SharedDocs/Pressemitteilungen/UK/EN/2016/16-09-27hbp_pilotsysteme.html.

WP7.3

- Research stage of **Óscar David Robles** at Universidad Militar Nueva Granada (Colombia), from November 28th to December 2nd, within their visiting professors program. Amongst other activities, he gave a conference on "Experience on neuroscientific data visualization – "Cajal Blue Brain" and "Human Brain Project"**.**
- Interview with **Luis Pastor**, broadcasted on TVE - Programme "La aventura del Saber" on 20 December 2016.

- InCytBrain Demo (**Benjamin Weyers**) at the STOA Event, 29-20 November 2016, European Parliament, Brussels, Belgium.
- Presentation of the HBP on the RWTH transparent (university-internal).

WP7.4

- Supercomputers: a key for understanding the human brain. Press release published on 9 February 2017 by BSC. https://www.bsc.es/news/bsc-news/supercomputers-key-tool-understanding-the-human-brain.
- SP7 Documentary. Movie about the contribution of BSC to the HBP/SP7. Published on YoutTube on 6 February 2017. https://www.youtube.com/watch?v=2Yk-DrsfnBU.
- Implementing Multi-Scale Simulation Workflows with COMPSs. Conference presentation by **Raül Sirvent** on 27 February 2017.
- Facebook posts and Tweets by BSC related to HBP. https://www.facebook.com/BSCCNS and https://twitter.com/BSC_CNS/
- Supercomputers: a key tool for understanding the human brain. In: BSC Newsletter March 2017. https://www.bsc.es/news/newsletter/bsc-cns-newsletter-march-2017
- BSC: Supercomputación para el Desarrollo de una Nación (Mateo Valero). Invited talk on 3 October 2016, Puebla (Mexico).
- PRACE PATC courses:
  - PATC: Programming Distributed Computing Platforms with COMPSs, 2 February 2017, Barcelona Supercomputing Centre (BSC), Barcelona, Spain.
  - Heterogeneous Programming on GPUs with MPI + OmpSs, 11-12 May 2016, Barcelona Supercomputing Centre (BSC), Barcelona, Spain.
  - Performance Analysis and Tools, 9-10 May 2016, Barcelona Supercomputing Centre (BSC), Barcelona, Spain.

WP7.5

- WISE - Going global on Security - An Overview on WISE - Security in Big and Open Data – WG, **R. Niederberger**, presentation at TNC16, Prague, Czech Republic, 14 June 2016, https://tnc16.geant.org/core/event/21.
- WISE Community - Updates from Working groups - Security in Big and Open Data – WG, **R. Niederberger**, Remote presentation at WISE Workshop at XSEDE Conference 2016, 18 July 2016, Prague, Miami, USA, https://wiki.geant.org/display/WISE/WISE+@XSEDE.
- Security in PRACE & HBP, **R. Niederberger**, presentations at 3rd WISE Workshop, 27 September 2016, Krakow, Poland, https://wiki.geant.org/display/WISE/WISE+@DI4R.
- An Overview on WISE - SBOD, **R. Niederberger**, presentations at 3rd WISE Workshop, 27 September 2016, Krakow, Poland, https://wiki.geant.org/display/WISE/WISE+@DI4R.
- Security and the Human Brain Project, **R. Niederberger**, presentation at the DI4R conference, 29 September 2016, Krakow, Poland, https://www.digitalinfrastructures.eu/sites/default/files/DI4R-SecurityAndTheHumanBrainProject%20_1.pdf.
- WISE and the WISE-WG on SBOD, **R. Niederberger**, presentation at the DI4R conference, 29 September 2016, Krakow, Poland, https://www.digitalinfrastructures.eu/sites/default/files/DI4R-WISE-and-The-WISE-WG-on-SBOD_0.pdf.

- The PRACE and HBP VPN – The migration phase, **R. Niederberger**, presentation at PRACE 4IP AHM and and 5IP-Kickoff Meeting, 31 January 2017, Athens, Greece.
- Bernstein Conference 2016, Berlin: workshop presentation and poster presented by Wouter Klijn.
- Booths of the HPC sites at the ISC High Performance 2016 (19-23 June 2016, Frankfurt am Main, Germany) and SC16 (13-18 November 2016, Salt Lake City, Utah, USA).
- CINECA: a Pillar for the HPAC Platform in HBP, presentation during HBP Summit 2016 Open Day by **G. Erbacci**, Florence, 12 October 2016.
- PRACE PATC course: Systems Workshop: Programming MareNostrum III, 11-12 April 2016, Barcelona Supercomputing Centre (BSC), Barcelona, Spain.

WP7.6

- Presentation of the HPAC Platform (**A. Lührs**) at the US Brain Initiative Meeting, 11-15 December 2016, Washington D.C, USA.
- “Human Brain Project specifies its Research Goals”/ “Human Brain Project konkretisiert seine Forschungsziele”. Press release by JUELICH on 11 November 2016. http://www.fz-juelich.de/SharedDocs/Pressemitteilungen/UK/EN/2016/2016-11-07-hbp-paper-neuron.html.
- Presentation of the HPAC Platform (**A. Lührs, B. Orth**) at the STOA Event, 29-30 November 2016, European Parliament, Brussels, Belgium.
- Presentation of the HPAC Platform at the HBP Open Day 2016, 12 October 2016, Florence, Italy.

10. Education

10.1 Participation in HBP Education Programme activities as lecturer

- HBP Young Researchers Event, 12 April 2016, Budapest, Hungary
  - Benjamin WEYERS (RWTH) as a member of the organising committee and speaker
- FENS Satellite Event, 1 July 2016, Copenhagen, Denmark
  - Hans Ekkehard PLESSER (NMBU)
  - Luis PASTOR (URJC)

10.2 Participation in HBP Education Programme activities as students

- HBP Student Conference, 8-10 February 2017, Austria
  - Sandra DIAZ-PIER (JUELICH)
  - Wouter KLIJN (JUELICH)
  - Alper YEGENOGLU (JUELICH)
- HBP Young Researchers Event, 12 April 2016, Budapest, Hungary
  - Philipp WEIDEL (JUELICH)
  - Victor LOPEZ (BSC)

11. Ethics

The SP7 Managers also act as Ethics Rapporteurs. In this role, they were in regular contact with the HBP Ethics Management and participated in related videoconferences, meetings and workshops. Together with the two members of the Ethics Advisory Board assigned to SP7, they collected, analysed and regularly reviewed potential ethical and social issues related to the research, development of and infrastructure provided by SP7, both as an individual Subproject as well as in the interplay with the other parts of HBP. Topics discussed are in particular the access to the SP7 infrastructure, data protection, security in an HPC environment and dual use or misuse of technology.

12. Innovation

SP7 has established close contacts with industry in the context of the Pre-Commercial Procurement (PCP), which started in the Ramp-Up Phase and overlapped with the first year of SGA1. Task 7.6.3 *Co-design knowledge management* will sustain the industry contacts that were established in the context of the PCP, but also with other relevant external stakeholders. Therefore, T7.6.3 has defined a co-design knowledge management process (see also report on MS7.6.1, section 7.4) to facilitate the sharing of knowledge relevant for co-design with external stakeholders, i.e. developers of relevant HPC technologies and architectures. In order to support the project-wide coordination of innovation-related activities, the leader of T7.6.3 regularly attends the meetings of the HBP's Innovation and Tech Transfer Committee (ITTC).

Annex A: Project Implementation Plan and the path to Use Case driven Co-Design

Preamble

The original intent, at the time the SGA1 DoA was created, was that WP7.5 would create a waterfall-style Project Implementation Plan (PIP) following a classical project management methodology. Hence, as part of the work to define WP7.5 components for the Project Lifecycle Application (PLA), a Service Breakdown for WP7.5 was created which included a comprehensive breakdown of service components and linkages. This work is summarised in Section 2.8 “HPAC Platform: Architecture Overview” of D7.6.1.

There was also the intention to build upon the FeDaPP T6.1 “Architecture Overview” report (https://emdesk.humanbrainproject.eu/shared/5a4f337f7606f-705547810b210f4704711b2d4546be77) and to utilise the Fenix Implementation Plan (which was intended to build upon the Services Portfolio Management Proposal for Fenix, the follow-on project from FeDaPP). The need to take into consideration the outcomes of the DPIT working group, which presented its findings to the EC reviewers on 24th of October 2016, was also made a priority. In the last quarter of 2016, as a first step in this direction, a project plan was made for the migration of the *Collaboratory* Web Services to the Infrastructure-as-a-Service (IaaS) resources at ETHZ-CSCS. In parallel, discussions were underway with the HBP Technical Coordinator and SP5 Leader in order to fully align the PIP with the requirements coming from SP5 and the *Collaboratory* development team. However, as the work progressed it soon became apparent that a different approach was needed.

Adoption of the Agile Methodology

At around the time of the above activities, work ramped up on the *HBP Software Engineering and Quality Assurance Approach* document (D11.3.3). Members of SP7, and in particular WP7.5, contributed to this document because they had training and experience with the Agile methodology for software development. It became clear that the methodology can also be applied to small system installation and configuration activities and hence the bring-up of IaaS resources at ETHZ-CSCS began to be managed in this way, in particular using *Scrum* procedures (see D11.3.3, pg.6ff). These activities proved to be very successful and to the direct benefit of the HBP, especially the *Collaboratory* Web Services hosting.

By May 2017, members of SP6 were starting to make extensive use of the *Collaboratory’s* Jupyter Notebooks services hosted at CSCS. As they exercised the service and introduced more complexity to their workflows they began to experience problems that needed help from the *Collaboratory* development team and the WP7.5 infrastructure team to resolve; regular (but somewhat *ad hoc*) meetings between SP6, SP7 (WP7.5) and the *Collaboratory* Development team were established.

By June 2017 it became apparent that these meetings were closely aligned with the activities of the HBP Software and Infrastructure Development Committee and hence, following a request from the SP7 Technical Coordinator and in agreement with the HBP Technical Coordinator and the Software and Infrastructure Development Directors, these *ad hoc* SP6-SP7 meetings became a regular weekly meeting which then grew to include SP5 and SP10, the latter having been added because the development efforts of WP7.5 and the *Collaboratory* had grown to also include use cases from these SPs.

For SP5 the work had grown to include the provisioning of archival-class Object Storage services using Swift within the OpenStack IaaS provided by ETHZ-CSCS. For SP10 the work had focused on the containerisation of the software environment of the Neurorobotics Platform (NRP) using Docker containers with the ultimate goal of deployment on the HPC resources at ETHZ-CSCS via Shifter.

Use Case Driven Co-design

The semi-formal minutes of these regular SP5-SP6-SP7-SP10-Collaboratory coordination meetings can be found within the Software and Infrastructure Development Committee *Collab* (see https://collab.humanbrainproject.eu/#/collab/325; this is a "private" Collab to which only *Collab* members have access). As can be seen, the meetings are somewhat freeform and some more formalism is needed but they follow a regular format where the development efforts are tracked and aligned with the requirements coming from the use cases of the users; SP5, SP6 and SP10 are regularly represented by a designated spokesperson of the senior users of these SPs.

As a consequence, the development efforts have become aligned with the direct needs of the users and can therefore be regarded as Use Case Driven Co-design.

Tracking Progress with Kanban Metrics

The Agile Scrum methodology was applied to the IaaS work at ETHZ-CSCS as well as the development of *Shifter* and the containerisation work done to migrate the NRP to HPC environments. In each case a product owner, stakeholders and a scrum master were defined. Small teams of up to eight technical personnel were created to work on the backlog of tasks with a prioritisation set in sprint planning meetings, sprints being mostly two weeks in duration. In order to track progress a Kanboard was created for each product.

The following figures show the Cumulative Flow diagrams taken from the KanBoards for the respective activities. These KanBoards are hosted at ETHZ-CSCS and are private but the plots show the level of activity in the backlog over time and give an indication of the amount of work done and the timeframe in which the work took place.

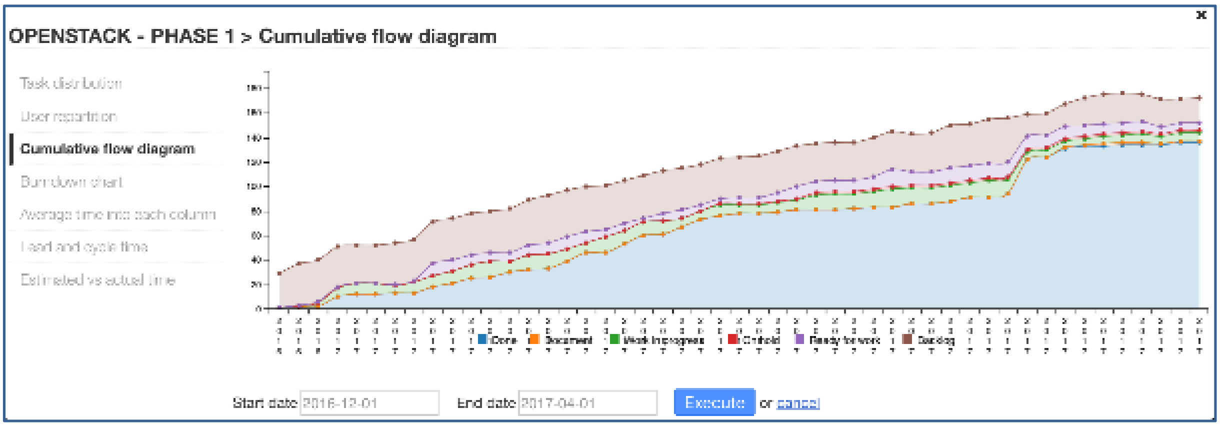

Figure 2: Kanban cumulative flow diagram: first deployment phase

Kanban cumulative flow diagram for the first phase of OpenStack IaaS deployment at ETHZ-CSCS during the period December 2016 through March 2017. This work relates to an early system (aka *Tambo*) setup at ETHZ-CSCS as a Proof of Concept on recycled hardware.

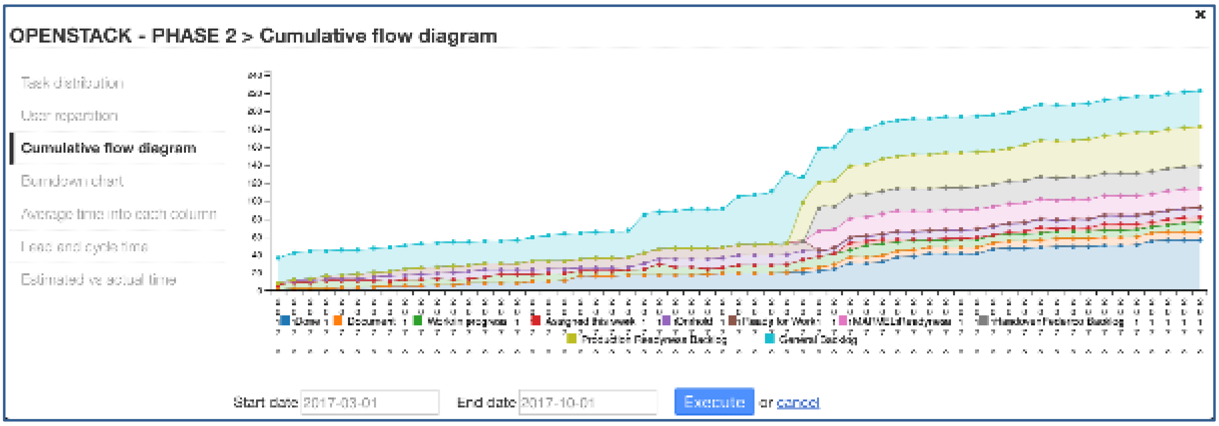

Figure 3: Kanban cumulative flow diagram: second deployment phase

Kanban cumulative flow diagram for the second phase of OpenStack IaaS deployment at ETHZ-CSCS during the period March to October 2017. This work relates to a new production system (aka *Pollux*) at ETHZ-CSCS which is used to host the *Collaboratory* web services as well as other non-HBP customer platforms.

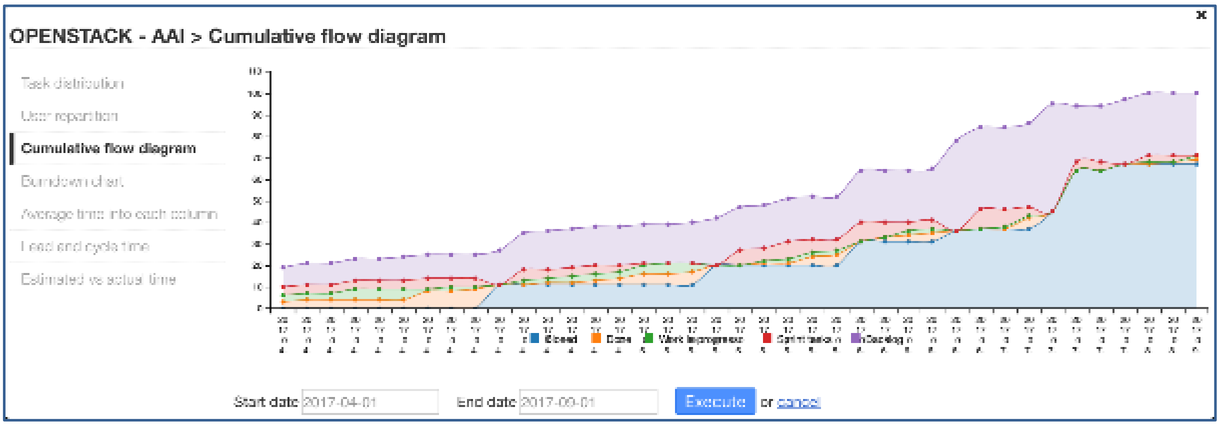

Figure 4: Kanban cumulative flow diagram: OpenStack AAI tasks

Kanban cumulative flow diagram for the OpenStack AAI tasks at ETHZ-CSCS during the period April to August 2017. These tasks were important for the full integration of the *Pollux* OpenStack IaaS system into the AAI of the Centre and the integration of the KeyStone components of the system with those of the external GPFS-hosted Swift.

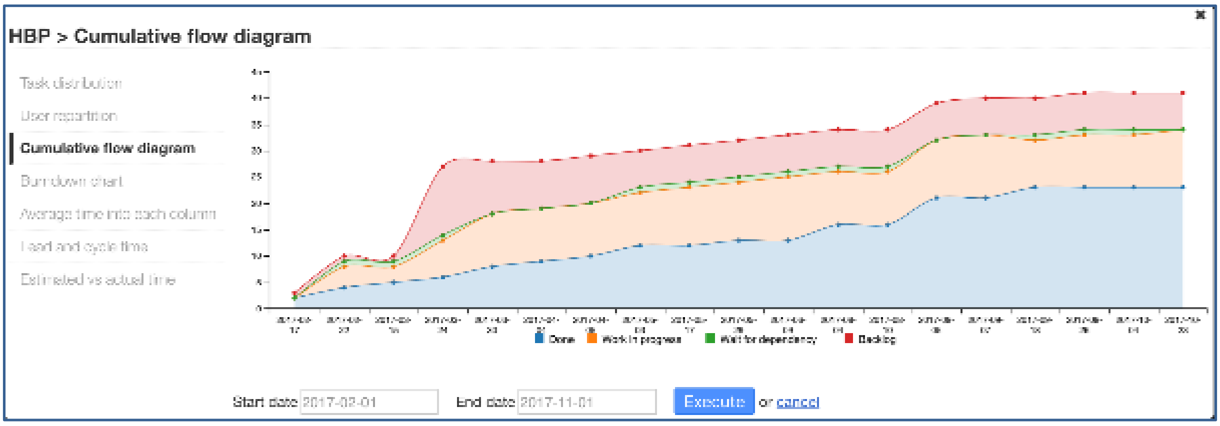

Figure 5: Kanban cumulative flow diagram: HBP NRP-related tasks

Kanban cumulative flow diagram for the HBP NRP-related tasks. Work involves the containerisation of the *Gazebo* virtual robotics environment and migration of the platform to ETHZ-CSCS HPC ecosystem (including OpenStack IaaS). Note that there is a low level of granularity in the tasks and they are more *meta-tasks* comprised of many smaller subtasks.

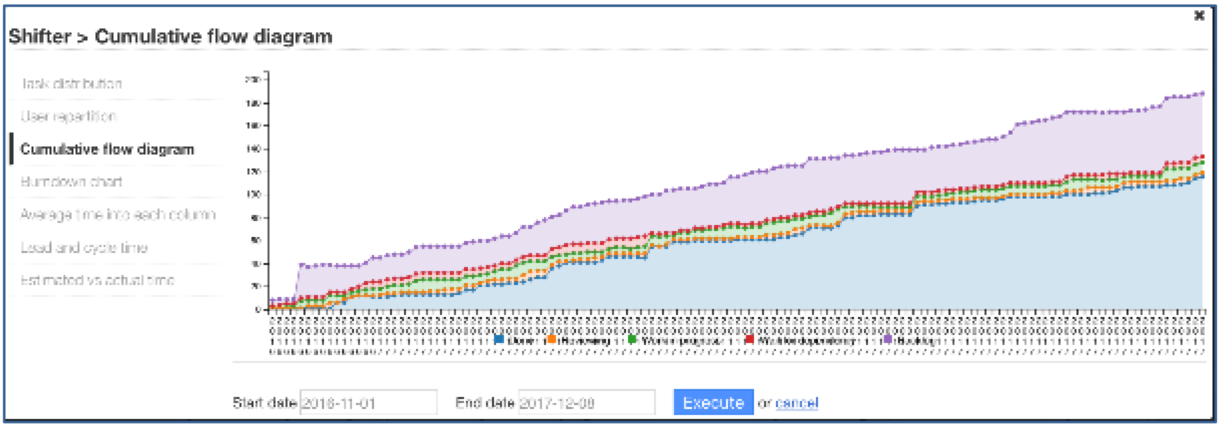

Figure 6: Kanban cumulative flow diagram: *Shifter* software development

Kanban cumulative flow diagram for the *Shifter* software development at ETHZ-CSCS during the period November 2016 to November 2017. Note that not all tasks were directly applicable to the HBP. Nonetheless the *Shifter* work is crucial to the work being done on the use of Docker containers in the NRP and the subsequent hosting of the platform in the HPC ecosystem of ETHZ-CSCS.

Another development effort undertaken within WP7.5 (T7.5.4) is that of *Arbor* (https://github.com/eth-cscs/arbor), a multi-compartment neural network simulation library. The development team has two full time developers in the same office, and part time contributions from partners in other locations (JUELICH-JSC and BSC). Standard agile processes have to be adapted to work with a small team that is not collocated. For example, the clearly-defined roles like product owner and scrum master predicated by scrum cannot be applied by the book to such a small team. The development team is experimenting with agile methods and online tools, including:

- Maintaining a Kanban board on GitHub https://github.com/eth-cscs/arbor/projects/13.
- Posting daily stand up status on a slack channel, so that distributed team members can participate with minimal time overheads.
- Developing KPIs that relate to the completion of tasks on the Kanban board, to measure progress.

On-going Work and Next Steps

The above use case driven co-design activities have proven to be highly productive over the past twelve months and as a consequence work will continue in this vein. There is still room for improvement and to this end, in addition to the on-going development activities, the following items will be actioned up to the end of SGA1:

1) Continue to improve the format of the HBP Software and Infrastructure Development Committee meetings. It would be beneficial to have more structure in the way the prioritised tasks are tracked and reported in the meetings;
2) Improve the quality of the reporting tools. So far free tools have been used but ETHZ-CSCS are working on migrating their projects to a newly setup Jira environment (https://www.atlassian.com/software/jira).
3) Investigate what tools can be provided by the *Collaboratory* for tracking such activities.

Integration of results from other SP7 Work Packages

WP7.5 operates and supports the High Performance Analytics and Computing (HPAC) Platform. In the formal sense of a Service Oriented Architecture (SOA), the simulation technology developed in WP7.1 and the interactive visualisation methods and tools of WP7.3 belong to the Software as a Service (SaaS) layer and as such are not formally part of the

HPAC Platform, which operates at the Infrastructure as a Service (IaaS) layer and provides Infrastructure Services to the other platforms and the Collaboratory. In both cases, however, a tight collaboration at the technical level ensures that the simulations and visualisation methods and tools run efficiently on the HPAC infrastructure. WP7.2 and WP7.4 are concerned with R&D on data-intensive supercomputing and dynamic resource management, respectively. The results of these WPs, where applicable, will be integrated and made available as part of the HPAC Platform within the SGA2 timeframe.

Integration with Pre-existing Structures

The approach of enabling use case driven co-design using agile methods has developed in parallel to the on-going operational aspects of SP7 (especially WP7.5) and the DoW as defined for SGA1. Coordination of these activities has been overseen by the SP7 Technical Coordinator in conjunction with the SP7 management and the WP7.5 leadership; in particular the SP7 Technical Coordinator and the co-leader of WP7.5 are the same person. Regular interactions between the SP7 Technical Coordinator and the HBP Technical Coordinator have been facilitated primarily via the HBP Software and Infrastructure Development Committee meetings. The SP7 Technical Coordinator has also represented SP7 on the Data Governance Working Group and has helped to specify the Data Management Plan and the Data Policy Manual for the HBP. Furthermore, key technical and management staff from SP7 were instrumental in the specification and subsequent approval of the ICEI project, which will enable the first instantiation of the federated Fenix infrastructure.

Agile methods allow the use case driven co-design to proceed quickly and having key SP7 and WP7.5 staff involved means that the transition to operations can also happen as quickly and seamlessly as possible. At the same time refinement of the operational aspects of the infrastructure can happen concurrently. The primary vehicle for WP7.5 coordination has been the monthly WP meetings (for an overview see the WP7.5 *Collab* at the following (private; i.e. only Collab members have access) link: https://collab.humanbrainproject.eu/#/collab/3561). Activities have focused on hardening the infrastructure for operations (especially monitoring and reporting) and migrating new environments and tools into the operational infrastructure. Cross pollination between the use case driven co-design and the operational teams has highlighted potential and real shortcomings in the way in which certain aspects of the operational service catalogue were configured (e.g. UFTP) and work has proceeded to rectify these shortcomings. This then has also ensured that the operational environment is aligned with user needs.

The SP7 activities are summarised in the (public; i.e. everyone with a Collaboratory account has read permissions) SP7 *Collab* (see https://collab.humanbrainproject.eu/#/collab/1798).

Annex B: Use cases and quantitative requirements analysis

Status of SP7 use cases

The use cases below are listed by title, owners and document id, followed by a short summary either directly from the document description or generated for this document. Documents are available here:

- SGA1 folder (tar.gz): https://emdesk.humanbrainproject.eu/shared/5a536cef0f608-f9693f5ac4f159b47b3b4511a4481a8e
- Fenix folder (tar.gz): https://emdesk.humanbrainproject.eu/shared/5a536d07b8b97-fe5aec20ddb941c283d042af128ef220

Six use cases were developed early in SGA1 as part of the FeDaPP/Fenix prototypes (Fenix folder), and five use cases have been developed more recently (SGA1 folder). Several of these documents belong to clusters of "use cases", looking at different aspects and subgroups of the same CDP or different elements of common workflows.

New SGA1 use cases

Jupyter notebooks for single cell model optimization and building

Involved SP(s): SP6

Related document: *SGA1/SP6-lupascu.docx*

Jupyter notebooks allowing a user-friendly environment for interactive generation and analysis of model neurons, are used to create a single cell model optimized to reproduce a specific set of electrophysiological features. The model can then be used as a stand-alone template, to run single cell simulations, or be included into the more complex workflow for circuit building. The particular use case illustrated below "Rebuild an existing single hippocampal cell model" allows a user to configure BluePyOpt to run an optimization, using HBP or own files for morphology, channel kinetics, features, and parameters. The optimizations are carried out through the Neuroscience Gateway or on one of the HPC centres supporting Brain Simulation Platform activities (currently CINECA and JUELICH-JSC). So far 45 optimizations have been generated, with a total size of 1.5 GB. A HBP-internal user can access these data through the Jupyter notebooks kernels provided by the Collaboratory to generate new optimizations. HBP-external users can run simulations only through the Neuroscience Gateway.

Jupyter notebooks with access to circuit and simulation files generated on HPC file system

Involved SP(s): SP6

Related document: *SGA1/SP6-Jupyter-notebooks.docx*

Jupyter notebooks are user-friendly environments for interactive exploration of data. Scientists can get insights with a non-linear, interactive, trial and error style of exploration of their data, and then share the analysis code and the generated figures. The analysis software the scientists are using needs to access circuit and/or simulation data. These data are generated on one of the HPC centre's systems (currently JUQUEEN). For a circuit, there are about 1000s of files with a total size of more than 200 GB. For a simulation, there are 10s of files with a total size of about 10 GB (some can go to TB range, but this is out of scope for this use case). Unfortunately, as of today, the end user cannot access to these data through the Jupyter notebooks kernels provided by Collaboratory. The files are too big to be downloadable, too.

GUI based experiment workflow control of automated parameter exploration in neural network model with connections structured in a 2d topology

Involved SP(s): SP7-internal use case (Simulation Laboratory Neuroscience at JUELICH-JSC, Visualisation group at RWTH, UPM and URJC)

Related document: *SGA1/Use Case Description and Specification 2d dynamics.docx*

Infrastructure for neuronal simulation, connectivity generation, live visualisation and job submission on HPC resources. Currently primarily a template/experiment for further development of use case documentation templates.

Image acquisition and 3D cellular analysis for at 1-micron resolution

Involved SP(s): SP2, SP5

Related document: *SGA1/Storage_UseCase_3D_Density_Analysis.docx*

Coupling laboratory scopes and temporary storage with supercomputing storage, down sampling, texture analysis, visualisation, scan packaging and region of interest analysis and packaging; large scale permanent, archival storage at HPC sites is also an element of this use case, which is a subset of the use case CDP3 Multi-Level Human Brain Atlas.

Ilastik through HBP infrastructure

Involved SP(s): SP5

Related doc.: *SGA1/UseCaseDescription_and_Specification_Ilastik_UseCase1_v2_0.docx*

The user is going to upload high resolution microscopic imaging data and has the objective to classify/segment objects in the microscopic image. In this use case, the user will train the classifier locally on its machine. The optimized classifier will be transferred to where the images are located. The images of the dataset will be processed in a batch manner on the server. The output of the batch processing, i.e. segmented images, will be stored together with the high resolution microscopic images and will be used in another use case. This is also related to the use case CDP3 Multi-Level Human Brain Atlas.

Imaging Analytics

Involved SP(s): SP2

Related document: *SGA1/UseCaseDescription_Imaging_Analytics_JURECA.docx*

MRI analysis toolkits are used for the analysis of imaging data. The data is currently distributed between image acquisition labs, stored on institute infrastructure and also on storage at an HPC centre (JUELICH-JSC) for final analysis and archival; depending on the analysis stage. The main goal is to optimise the analytics workflow.

FeDaPP/Fenix use cases

CDP5: Plasticity, Learning and Development: SpiNNaker module

Involved SP(s): SP5, SP6, SP9

Related document: *Fenix/CDP-5-AR-FeDaPP_UseCaseDescription_and_Specification.pdf*

This project involves the co-simulation of neural networks involving plasticity on supercomputers using NEST and on neuromorphic hardware at least partially outside of the FeDaPP environment. SpiNNaker is an ARM-based neuromorphic project designed to simulate very large neural networks at high speed, using a routing-based protocol to map connectivity between simulated neurons. The neural models and networks to be simulated on different platforms will be stored in git repositories, either public or private. The models will be in a NEST-based domain-specific language (DSL) such as NESTML or a NEURON-based modelling language, while networks may be PyNN scripts. These network-model combinations will be pulled to Manchester using a Collaboratory portal and presented to the Manchester system for batch processing that will be authorized by a server in Andrew Davison's group (AD server) using the OIDC infrastructure and the RESTful API designed for the Collaboratory. A similar process may occur with the Heidelberg neuromorphic hardware designed by Karlheinz Meier's group. The simulation will stage the results of time-varying spiking and weight changes locally in Manchester. Using the same RESTful API, the location at a supercomputing site (JUELICH-JSC) for uploading will be exchanged with the AD server, and direct upload will follow. All data will then be pushed to be used in an Elephant analysis pipeline, with the same model then being run at the supercomputing site with NEST and fed as well into this pipeline, for comparison, where the results will be shared among a Collab group and later disseminated by that group. This process and tools will be run from Jupyter notebook on the Collab server and will be available to other HBP consortium Collab work groups.

CDP5 Plasticity, Learning and Development: Neuromorphic server

Involved SP(s): SP5, SP6, SP9

Related doc.: *Fenix/CDP5-AD-FeDaPP_UseCaseDescription_and_Specification_v3.4.pdf*

The AD server links the SpiNNaker, Heidelberg and INM-6 workflows together. Using the Unity/Collaboratory authentication system, a Collab Jupyter Notebook will use a REST API to submit a job reservation to the AD server. AD will then queue a batch job to Heidelberg or Manchester, which will at termination update the AD server and request information needed to upload to JUELICH-JSC storage. For Manchester, this is through the Collab REST API requiring transport through the AD server currently, which then uses this transport to JSC. For Heidelberg, the authorization keys for a UNICORE workflow are downloaded, and Heidelberg directly uses the UNICORE tools to directly upload results to JSC.

CDP3 Multi-Level Human Brain Atlas

Involved SP(s): SP2, SP5

Related document: *Fenix/CDP3-FeDaPP_UseCaseDescription_and_Specification.pdf*

Brain atlas data in the form of [tif, metadata] is produced at an experimental lab, transmitted to an HPC centre for image optimization, combination and tagging to produce a three-dimensional atlas. Approved subsets will then be made available for global visualisation. Superset of Image acquisition and 3D cellular analysis for at 1-micron resolution.

CDP5 Plasticity, Learning and Development: Heidelberg Neuromorphic Hardware Workflow

Involved SP(s): SP5, SP6, SP9

Related document: *Fenix/CDP5-Heidelberg-FeDaPP_UseCaseDescription_and_Specification_v3.4.pdf*

As in the SpiNNaker workflow (see above), Jupyter workbooks will be used to control the workflow, and time allocation management and project tracking will go through Andrew Davison's servers (AD) and the Neuromorphic Computing Collab (NCC). Jobs will be directed to the Heidelberg equipment (as well as to SpiNNaker, NEST running at supercomputer at JUELICH-JSC, …) for processing within their SLURM batch system. Input data will be PyNN scripts to build and run simulations; spike trains and voltage traces will be uploaded to the HPC system in a location determined by AD in order to integrate with other inputs into INM6's Elephant analysis pipeline which will feed back to the Jupyter notebook.

CDP1: Development of Whole Mouse Brain Model and related Mouse Brain Atlas

Involved SP(s): SP1, SP5, SP6

Related document: *Fenix/CDP1-FeDaPP_UseCaseDescription_and_Specification.pdf*

Empirical data for mice will be collected from several sources: high resolution data from microscopy, MRI and such; the Allen Institute database; and functional mouse data including live calcium imaging and mouse limb positions and responses. The high-resolution data and Allen Institute data will be used to create a Mouse Brain Atlas which will then be used to develop a point neuron model of the mouse brain for NEST which will be simulated in conjunction with a robotic simulator, as well as a neural mass model simulation. The same experiments will be repeated both in the lab with live mice to produce the functional mouse data, and as simulations. The uploaded results will then be compared to validate the simulations and that data will be archived as well. The atlas will be available within the consortium and after publication to the public. NEST/VirtualBrain simulations using the Atlas data will be available through the Collaboratory to the community.

CDP5 Plasticity, Learning and Development: Elephant/NEST/SpiNNaker workflow

Involved SP(s): SP5, SP6, SP9

Related doc.: *Fenix/CDP5-INM6-FeDaPP_UseCaseDescription_and_Specification_v3.4.pdf*

This subsection of CDP5 focuses on the NEST data production, and the Elephant analysis. A Collab Jupyter script will send both a request to the AD server as well as to an HPC system, using a task id produced by the Collaboratory to track different parts of the experiment. On the HPC system, the needed directories will be created, a PyNN script will be downloaded and a batch job will be executed to run NEST through that PyNN script. The NEST simulation produces spike trains in an internal format (gdf, sionlib) that will then be converted to a standard HDF5 container (NEO or NIX, for example). These results will then be relayed to the Jupyter notebook, where the neuroscientist will be able to continue a UNICORE workflow using the Elephant analysis toolkit, such as a comparison with SpiNNaker results. A similar workflow will be constructed to compare experimental results with NEST results, with the addition that the NEST simulations would then be iterative with the analyst rerunning the simulation with different parameter sets after the Elephant analysis comparing the internal simulation results with the external results.

Quantitative analysis of required resources

As part of Fenix and ICEI projects, a preliminary quantitative resource analysis has been performed. A summary of those findings follow; a fuller version is part of the ICEI grant proposal and agreement:

GUI-based interaction with extreme scale network models: The Brain Simulation Platform (BSP) is an internet-accessible collaborative Platform, which comprises a suite of software tools and workflows for collaborative brain research to allow researchers to reconstruct and simulate detailed multi-level models of the brain, displaying emergent structures and behaviours. BSP users can define and start simulations from Jupyter notebooks inside the Collaboratory. Most of the BSP use cases, including the simplest CDP2 product (the Hodgkin-Huxley Neuron Builder, planned in SGA2-T6.1.3), require HPC resources. Hence, the BSP has the need to run non-trivial HPC workloads from within the interactive Jupyter notebooks.

Extreme scale network models: The bottleneck of computational demands for this class of use cases is the network simulation, using tools like NEST, Neuron, Arbor or The Virtual Brain (TVB). To estimate the scale of the problem: the "record" NEST simulation on the K supercomputer in 2013 used approximately 1.1 PB of memory while taking up most of a 10 PFlop/s supercomputer. To move from 10^9 neurons to 10^{10} neurons using the current class of software architectures would take on the order of 10 PB and 100 PFlop/s.

For morphologically detailed simulations, initial estimates for "Arbor" indicate that peak performance for a GPU-based architecture occurs with 10k cells/GPU. Since 10k P100 GPUs is at the order of 50 PFlop/s, to sustain such activity we would need

10k sockets * 10k cells/socket * (1~1000) k bytes/cells = 0.1~100 TB for 100 * 10^6 cells, which is on the order of the size of the human hippocampus. Other simulators may require more memory. In practice, such simulations will use a variety of kinds of hardware resources (CPUs, GPUs, FPGAs …) across a variety of platforms (within ICEI, national resources, PRACE resources). The previous analysis is intended to indicate, to an order of magnitude, the relative balance of various resources for simulations of networks of morphologically detailed neurons.

Visualisation: Demands of visualisation itself (such as visualizing 1000 morphologically detailed neurons) can be handled with desktop systems with moderate amount of memory. However, the post-processing leading to such visualisation is highly dependent on the details of the post-processing, but is usually flexible to trade-offs between different levels of the memory hierarchy.

This use case is relevant for reaching the SP6 SGA2 objective *6d) Package and operate the capabilities developed for 6c as Software and Platform as a Service (SaaS/PaaS) and make the Brain Simulation Platform accessible to the community*.

Enrichment of the human brain atlas with qualitative and quantitative datasets: A range of highly diverse qualitative and quantitative datasets need to be spatially and semantically registered to the human brain atlas by the Neuroinformatics Platform (SP5), a direct contribution to the HBP Flagship Objective *FO2 Gather, organise and disseminate data describing the brain and its diseases*. These datasets have different file formats, come from different partners, and need different kinds of user interface functionality. Data is collected within various HBP Subprojects (mainly SP2 for the Human Brain Atlas), but also come from external repositories like the Human Connectome Project (HCP).

Current pipelines are implemented on JURECA at JUELICH-JSC, a commodity-type HPC cluster with 1872 nodes producing about 2 PFlop/s and using about 0.2 PB. For microscopic image analysis steps, typical usage is from ⅓ to less than 100% of total hyper-threads available, so the ratio of locally available memory to Flops is less than three-fold the current ratio available on JURECA. I/O performance is the current bottleneck for whole brain histological sections and deep learning analysis for automatic image classification. More than overall memory demand being an issue, since much of the analysis is only loosely coupled above the level of a single node, "fat" nodes with 1 TB of local memory would be helpful. The total memory usage is limited by the number of physical scanners (8 currently) handling one section per day, where computationally each is an independent job using 1-10 nodes.

Validation of neuromorphic results: Analysis of the similarities and differences of results obtained through software simulation on supercomputers and from neuromorphic systems, e.g. using Elephant developed in SGA1/SGA2-T7.5.1, require significant resources behind a Jupyter notebook interface. First activities started in SGA1 (e.g. related to milestone MS9.1.2 *Neuromorphic-HPC comparison workflow*); work on a tighter integration of the Neuromorphic Computing Platform (SP9) and SP7 is planned to continue in SGA2. This is an important topic for HBP, as also pointed out by the reviewers of the mid-term SGA1 review.

We do not have sufficient information for a careful assessment of the memory footprints for Elephant analysis, which is the key bottleneck for cross-validation between standard simulators and neuromorphic results. This would also require a definition of a validation benchmark suite. However, given the limited extent of neuromorphic resources available, demands of traditional simulators are likely to far outstrip computational uses needed for this use case. Python analysis pipelines of features such as spike series are also likely to be dominated by random and sequential access times to permanent I/O resources.

Neurorobotics: the HBP Neurorobotics Platform (NRP; SP9) is an internet-accessible simulation system that allows the simulation of robots controlled by spiking neural networks. Prospective users include, but are not limited to, neuroscientists wanting to validate brain models in the context of closed action-perception loops as well as robotics researchers wanting to develop new neuro-inspired controllers. The NRP will need to spawn an interactive session which, in turn, spawns the virtual robotics environment and a NEST simulation, the latter being connected to the virtual robot that is embedded within the virtual robotics environment. This activity is an important contribution to the HBP Flagship Objective *FO5 Develop brain-inspired computing, data analytics and robotics*.

For this use case a 10 GB of memory per Gazebo container is required plus the memory requirements of the NEST simulation. The latter cannot yet be estimated as the problem sizes have not yet sufficiently well defined. It is, however, expected that currently available systems with O(100) TB of main memory provide sufficient memory capacity.