Human Brain Project

Co-funded by the European Commission

Project Number:	284941	Project Title:	Human Brain Project

Document Title:	High Performance Computing Platform v1 - specification document
Document Filename:	HBP_SP7_EPFL_140502_D7.7.2_Final.docx
Deliverable Number:	D 7.7.2
Deliverable Type:	Platform Specification
Work Package(s):	WP 7.1, WP 7.2, WP 7.3, WP 7.4, WP 7.5, WP 7.6, WP 7.7
Planned Delivery Date:	M 6 / 31 March 2014
Actual Delivery Date:	M 8 / 02 May 2014

Author:	Luc CORBEIL, ETHZ, P15	Use Cases, Requirements, Integration & Editing
Author:	Boris ORTH, JUELICH, P17	Goals, Parallel Activities, Integration & Editing
Contributor:	Rosa M BADIA, BSC, P4	Use Cases, Software, Functions
Contributor:	Javier BARTOLOME, BSC, P4	Physical Architecture, Software
Contributor:	Matthias BOLTEN, BUW, P5	Software, Functions
Contributor:	Giovanni ERBACCI, CINECA, P10	Physical Architecture, Software, Use Cases
Contributor:	Giuseppe FIAMENI, CINECA, P10	Software
Contributor:	Diana GUDU, KIT, P30	Software
Contributor:	Marcus HARDT, KIT, P30	Software
Contributor:	Thomas HEINIS, EPFL, P1	Use Cases, Requirements, Software, Functions
Contributor:	Jesus LABARTA, BSC, P4	Use Cases, Requirements, Software, Functions
Contributor:	Thomas LIPPERT, JUELICH, P17	Goals, Prerequisites, Parallel Activities
Contributor:	Daniel MALLMANN, JUELICH, P17	Parallel Activities
Contributor:	Cristian MEZZANOTTE, ETHZ, P15	Use Cases, Software
Contributor:	Bernd MOHR, JUELICH, P17	Use Cases, Software
Contributor:	Roberto MUCCI, CINECA, P10	Software
Contributor:	Ralph NIEDERBERGER, JUELICH, P17	Network and Security, Functions
Contributor:	Dirk PLEITER, JUELICH, P17	Requirements, Software
Contributor:	Bernd SCHULLER, JUELICH, P17	Relations to other Platforms, UP Integration

Contributor:	Anna WESTHOFF, JUELICH, P17	Functions, Key Performance Indicators
Contributor:	Benjamin WEYERS, RWTH, P42	Use Cases, Software
Editor:	Celia Luterbacher, EPFL, P1	*Copy editing*
	Richard Walker, EPFL, P1	*Final review and copy edit*
	Guy Willis, EPFL, P1	*Copy editing*

Abstract:	This report provides a detailed specification of the High Performance Computing (HPC) Platform Version 1.0, which the HPC Platform Subproject (SP7) of the Human Brain Project (HBP) will set up and make available during the HBP's ramp-up phase. A key feature of the HPC Platform will be its integration with the HBP's Unified Portal (HBP-UP), a web portal where scientists from around the world can work collaboratively to improve, validate and use neuroscientific models at all levels of detail. After an overview of the goals and layout of the HPC Platform, the report describes typical examples of how the Platform will be used. The report describes a set of functional and non-functional requirements derived from these "use cases," and then describes the software and hardware components that will be used to meet these requirements. The report also describes necessary parallel activities, in particular, the preparation of the next version of the Platform. It concludes with the definition of Key Performance Indicators (KPIs) that can be used to track Platform development and adoption by the target community, as well as target values for the indicators.

Keywords:	High Performance Computing (HPC), Platform, Unified Portal (UP), use cases, functional requirements, non-functional requirements, software, physical architecture, component functions, Key Performance Indicators (KPI)

Document Status

Version	Date	Status	Comments
0.1	12 Feb 2014	Draft	Initial draft, based on detailed outline
0.2	13 Feb 2014	Draft	Inputs integrated
0.3	14 Feb 2014	Draft	Additional input
0.4	18 Feb 2014	Draft	Mainly corrections, responsibilities added
0.5	20 Feb 2014	Draft	Additional input, corrections
0.6	24 Feb 2014	Draft	Additional input, corrections
0.7	25 Feb 2014	Draft	Additional use cases, corrections
0.8	26 Feb 2014	Draft	More use cases and corrections
0.9	28 Feb 2014	Draft	For internal review
1.0	31 Mar 2014	Draft	Second version for internal review
1.1	14 Apr 2014	Draft	Pre-final version, for internal review
1.2	15 Apr 2014	Draft	Copyedit, CL
1.3	22 Apr 2014	Draft	Final review and edit, RW
1.4	24 April 2014	Draft	Author review & approval
1.5	28 Apr 2014	Final	Final Copyedit, CL

Table of Contents

List of Figures and Tables

1. Executive Summary

The Human Brain Project (HBP) is a large-scale, European-led, international scientific undertaking that aims to integrate all existing knowledge about the human brain in unifying models, and supercomputer-based simulations of these models. The resulting "virtual brain" offers the prospect of a fundamentally new understanding of the human brain and its diseases, as well as novel, brain-like computing technologies.

The High Performance Computing (HPC) Platform Subproject (SP7) is one of the HBP's twelve scientific subprojects. Over the 10-year duration of the HBP, SP7 will build, integrate and operate the HBP's HPC Platform. The Platform comprises all hardware and software components of the supercomputing, data and visualisation infrastructure required to:

- Run large-scale, data intensive, interactive brain simulations up to the size of a full human brain,
- Manage the large amounts of data used and produced by the simulations, and
- Concurrently manage workloads and workflows, data processing and visualisation.

SP7 will make version 1.0 of the HPC Platform available to the HBP Consortium in project month 18 and to the wider scientific community in month 30. In this version, the HPC Platform will comprise existing supercomputing, data storage and visualisation capabilities at Forschungszentrum Jülich (Germany), the Swiss National Supercomputing Centre in Lugano (CSCS, Switzerland), Barcelona Supercomputing Centre (BSC, Spain), Cineca in Bologna (Italy), Karlsruhe Institute of Technology (KIT, Germany) and RWTH Aachen University (Germany). As part of the integrated HPC Platform, each of the sites will play a specific role.

A key feature of the HPC Platform will be its integration with the HBP Unified Portal (HBP-UP), developed by the HBP's Brain Simulation Platform Subproject (SP6). The HBP-UP will be a web portal integrating all web-enabled components of the different HBP Platforms, where scientists from around the world will work collaboratively to improve, validate and use neuroscientific models at all levels of detail. The HPC Platform will contribute to the UP services such as authentication and authorisation for single sign-on to all HPC Platform resources, job submission and monitoring, file/data transfer and storage of large data sets, resource quota accounting and management, and hardware and software monitoring.

This report specifies in detail the High Performance Computing Platform Version 1.0. It is intended for a technical and scientific audience. After an overview of the goals and layout of the Platform, it describes typical examples of how the HPC Platform will be used. From these use cases, a set of functional and non-functional requirements is derived. We then describe the software and hardware components that will be used to meet the requirements. The report also describes necessary parallel activities, in particular, the preparation for the next version of the Platform. The report concludes with the definition of Key Performance Indicators (KPIs) that will be used to track Platform development and adoption by the target community, as well as target values for the indicators.

2. The High Performance Computing (HPC) Platform

2.1 HPC Platform: Overall Goals

2.1.1 The Human Brain Project

The goal of the Human Brain Project (HBP)[1] is to gather all existing knowledge about the human brain and to reconstruct the brain, piece by piece, in multi-scale models and supercomputer-based simulations of these models. The resulting "virtual brain" offers the prospect of a fundamentally new understanding of the human brain and its diseases, as well as novel, brain-like computing technologies.

The HBP is developing six Information & Communication Technology (ICT) Platforms dedicated to Neuroinformatics, Brain Simulation, High-Performance Computing (HPC), Medical Informatics, Neuromorphic Computing and Neurorobotics. This research infrastructure will make it possible to federate neuroscience data from all over the world, to integrate the data in unifying models and simulations of the brain, to validate the results against empirical data from biology and medicine, and to make the data available to the world scientific community.

The resulting knowledge on the structure and connectivity of the brain will open up new perspectives for the development of "brain-like" computing systems incorporating unique characteristics of the brain such as energy-efficiency, fault-tolerance and the ability to learn. The HBP's models and simulations will also enable researchers to carry out *in silico* experiments on the human brain that cannot be done *in vivo* for practical or ethical reasons.

In October 2013, the European Commission (EC) started supporting this vision through its Future & Emerging Technologies (FET) Flagship Initiative. The HBP's ramp-up phase, which will last until 2016, is funded by the EU's 7th Framework Programme (FP7). This phase should be followed by a partially overlapping operational phase, which will be supported under the next Framework Programme, Horizon 2020.

The HBP as a whole will be implemented in three phases, spread over ten years, with an estimated total budget of more than EUR 1 billion. The project, which is coordinated by the École Polytechnique Fédérale de Lausanne (EPFL) in Switzerland, already brings together 80 European and international research institutions. In April 2014, more partners joined the Consortium via the HBP's Competitive Call Programme.

2.1.2 The HPC Platform

The HPC Platform Subproject (SP7) is one of the HBP's twelve operational subprojects. Coordinated by Forschungszentrum Jülich, it currently involves 15 partner institutions, including four leading European supercomputing centres and several universities. Its 10-year mission is to build, integrate and operate the hardware and software components of the supercomputing and data infrastructure required to:

- Run large-scale, data intensive, interactive brain simulations up to the size of a full human brain,
- Manage the large amounts of data used and produced by the simulations, and
- Concurrently manage workloads and workflows, data processing and visualisation.

This infrastructure as a whole will form the HPC Platform. SP7 will make version 1.0 of the HPC Platform available to the HBP Consortium from project month 18 and to the wider scientific community from month 30. The HPC Platform will be accessible in a seamless and intuitive manner through the HBP's Unified Portal (UP)[2].

In the HBP's ramp-up phase, the HPC Platform will comprise existing supercomputing, data management and visualisation capabilities at Forschungszentrum Jülich (Germany), the Swiss National Supercomputing Centre in Lugano (CSCS, Switzerland), Barcelona Supercomputing Centre (BSC, Spain), Cineca in Bologna (Italy), Karlsruhe Institute of Technology (KIT, Germany) and RWTH Aachen University (Germany). As part of the HPC Platform, each of the sites will play a specific role.

2.1.2.1 HBP Supercomputer

The HBP Supercomputer located at Jülich Supercomputing Centre (JSC) will be the main production system for the HBP. It will continuously integrate hard- and software technologies developed in SP7 and related subprojects once they have reached a sufficient level of maturity, and offer them as production-quality services to users from the HBP and beyond.

In the ramp-up phase, the existing 5.9 PFlops IBM Blue Gene/Q system JUQUEEN at JSC is available for HBP production runs. The HBP Supercomputer for the following phases will be built in stages. A "pre-exascale" system of the order of 50 PFlops is planned for the 2016-18 timeframe. Multi-scale brain simulations up to the size of a full human brain as required by the HBP are expected to require exascale capabilities. According to most suppliers' roadmaps, these capabilities are likely to be available in approximately 2021-22. In addition to providing sufficient computing performance, the HBP Supercomputer will need to support the large memory footprints of brain simulations as well as interactive supercomputing. A pre-commercial procurement (PCP), which SP7 is conducting during the ramp-up phase in parallel with building the Platform, will drive the research and development of suitable HPC technology on the supply side, so that it will be commercially available at the time of the planned procurement of the pre-exascale version of the HBP Supercomputer.

2.1.2.2 HBP Development System

The HBP Development System located at CSCS will provide a platform for the development and testing of brain models and their simulation.

In the ramp-up phase, the existing 840 TFlops IBM Blue Gene/Q system Blue Brain 4, hosted by CSCS for the EPFL's Blue Brain Project (BBP), is available for this purpose. This should be followed by a system with 5-10 PFlops and 2-5 PB of memory in the 2017-20 time frame. In 2021-24, the 5-10 PFlop/2-5 PB memory system should be followed by a system with 50-100 PFlops and 10-20 PB of memory. Storage clusters and interactive visualisation equipment will complement all versions of the HBP Development System.

2.1.2.3 HBP Molecular Dynamics Supercomputer

The HBP Molecular Dynamics Supercomputer at BSC will support HBP simulations in the subcellular domain.

In the ramp-up phase, the existing 1.1 PFlops IBM iDataPlex system MareNostrum and the 200 TFlops GPU cluster Minotauro are available for HBP molecular dynamics simulations. An upgrade to a machine with a performance of about 3 PFlops is planned for the 2016-18 timeframe, an upgrade to about 50 PFlops in the 2018-21 timeframe.

2.1.2.4 HBP Massive Data Analytics Supercomputer

The HBP Massive Data Analytics Supercomputer located at Cineca will provide the data-centric HPC resource for managing, processing, analysing and storing the large volumes of data, which will be generated by the HBP.

In the ramp-up phase, the 2.1 PFlops IBM Blue Gene/Q system FERMI (200 TB main memory), the 300 TFlops GPU cluster PLX, and another data-intensive computing system (currently being procured), are available at Cineca for massive data analytics and post-processing. The three systems are integrated with a Data Facility of 20 TB fast storage SSD, around 5 PB on-line disk storage repository and more than 10 PB for long-term data preservation, providing efficient data life-cycle management of structured and un-structured data generated by the HBP.

In the following phases, the HBP Massive Data Analytics Supercomputer will evolve in three main steps: In 2015-16 the HPC system will evolve into the range of 10 PFlops, integrated with a data facility of the order of 100 TB fast storage SSD.

In 2017-18, the HPC system will evolve further into the range of 50 PFlops, integrated with an enhanced fast storage data facility suitable to manage the Big Data analytics and visualisation activity related to the data produced by the HPB for the primate-scale brain and possibly larger brain samples, up to the human brain.

In the 2020-23 time frame, it is foreseen that the HPC system will reach a performance of 500 PFlops and that the Data Facility will increase its capacity and performance accordingly.

2.1.2.5 HBP Cloud Service

The HBP Cloud Service based at KIT will provide users of the HBP Platforms with easy, secure and transparent access to Cloud-based storage and computing resources, complementing the storage provided locally by the HPC centres.

In the ramp-up phase, KIT provides Cloud-based storage with a capacity of 3 PBytes, accessible through an Amazon S3-compatible interface. The capacity may be increased in the following phases, depending on resource usage. In general, the HBP will directly benefit from the on going storage and archive federation activities at KIT. The capabilities of the HBP storage system at KIT will be enhanced as the underlying open source system develops.

2.1.2.6 HBP High Fidelity Visualisation Systems

The HBP High Fidelity Visualisation Systems located at RWTH Aachen University will provide users of the HBP Platforms with high resolution and fully immersive visualisation resources.

In the ramp-up phase, RWTH provides a fully immersive visualisation environment in the form of a CAVE installation. The aixCAVE at RWTH is a five-sided virtual reality environment with a floor size of 5.25x5.25 m² and a height of 3.3 m. All walls are equipped (rear projection) with 4 projectors each; the floor is equipped with 8 projectors. The 24 HD-DLP projectors support active stereo (NVIDIA 3DVision pro shuttering at 120 Hz) and are driven by a 24-node visualisation cluster. Each node contains 2 Intel Xeon Westmere CPUs (6 Cores at 2.7 GHz), 24 GB RAM, and runs 2 NVIDIA Quadro 6000 GPUs each. The CAVE volume is fully tracked by an ART tracking system involving 8 infrared cameras.

An additional high-resolution visualisation system—a tiled display wall—will be implemented at RWTH by project month 6 as a reference system of an existing system at EPFL. The display wall at RWTH will consist of 6 full HD screens of 46'' size. It is driven by a 7-node cluster (1 master and 6 slaves) each equipped with an Intel Xeon E5 CPU (4 Cores at 3.7 GHz), 16 GB RAM, and 1 NVIDIA GTX 780Ti GPU each.

The four distributed HPC systems at Jülich, CSCS, BSC and Cineca will be connected via a high-speed communication network, allowing the partners to share massive data sets ranging from petabytes to exabytes of data. The HBP Cloud storage at KIT and the visualisation system at RWTH will be connected to the HPC systems through their regular Internet links.

Advanced software for interactive visualisation, resource and data management will be developed within the subproject and integrated into the Platform as it becomes available.

The different hardware and software components will be integrated with regard to job submission and data management, user authentication, authorisation, accounting, and resource monitoring with the explicit goal of providing unifying and easy-to-use interfaces to the diverse and heterogeneous compute and data resources that constitute the HPC Platform.

2.2 HPC Platform: Use Cases

In this section we give some typical examples ("Use Cases") of how we expect the HPC Platform to be used during the HBP ramp-up phase (up to project month 30).

As is the case for all HBP Platforms, the default method to interact with the HPC Platform will be the Unified Portal (UP), which will be developed by the Brain Simulation Platform Subproject (SP6). It is described in detail in the Brain Simulation Platform Specification[2].

Roles

Computational Scientific User (CSU) – A user with scientific development skills and comfort in launching command line HPC jobs.

Biological Scientific User (BSU) – A user with scientific expertise, but limited technical computing skills.

Scientific User (SU) – A scientific user, either a CSU or a BSU.

Scientific Developer (SCIDEV) – A user who is developing software to directly realise the scientific objectives. This user is usually working in close collaboration with scientists, both CSUs and BSUs.

Developer (DEV) – A user who is developing software to realise engineering, operational and/or scientific objectives.

Portal User (PU) –A user who accesses Platform functions through the Web GUI.

Service User (ServU) – A user who accesses Platform functions through a programmatic Service Client API.

Infrastructure Personnel (INFRA) - An infrastructure system administrator or developer, typically responsible for deploying and monitoring Platform services that are offered directly to customers.

2.2.1 Simulation

Use Case ID: SP7-UC-001

Primary Actors: Two Scientific Users, Abigail and Bill

Success Scenario:

1) Abigail finds a released brain model in the UP using the integrated search functionality.
2) Abigail adds the brain model to a new project, Project1.
3) Abigail configures a simulation task using the UP simulation configuration system.
4) Abigail then launches the simulation task as a job to run on an HPC Platform compute resource. During the launch process she may be asked (depending on the task definition) to make a decision about which computing resource to run on and where to store final output data. The HPC Platform could provide recommendations based on data locality, data transfer volume and compute resource availability. However, regardless of Abigail's decision, the simulation results will be equivalent.
5) The job registers some or all of its output (as determined by the task definition) in Project1 in the UP once it is complete. The output may be marshalled to its final storage location by the HPC Platform data transfer service.
6) Abigail adds Bill to the project with write permissions to help her analyse the simulation output data.
7) Bill can run analysis on the simulation output and share the analysis results with Abigail.

2.2.2 Interactive Simulation with Dynamic Resource Management

Use Case ID: SP7-UC-002

Primary Actors: Three Scientific Users, Abigail, Bill and Chris

Success Scenario:

1) Abigail finds a released brain model in the UP using the integrated search functionality.
2) Abigail adds the brain model to a new project, Project2.
3) Abigail configures a simulation task using the UP simulation configuration system.
4) Abigail then launches a long lasting simulation task as a job to run on an HPC Platform compute resource.
5) Bill launches an analytics computational workflow and/or a visualisation pipeline during the long simulation run and for a duration that may be controlled by him (or by Abigail as well). This analytics job can be launched a few hours after the simulation was launched to analyse what has been simulated so far and how the simulation evolves. Each of these previous jobs will consist potentially of multiple MPI processes, each of them potentially multithreaded (OpenMP, OmpSs, pthreads). The balance of computational demand between the different jobs (components of the multi-scale simulation, analytics and visualisation) may change along time with fine granularity (on the scale of minutes). It will thus be necessary to support job requisites. Each of them will be a different application run within the context of the session**.** Simulations and analytic workflows can be launched by the same or by different users.
6) Chris configures a molecular dynamic simulation that computes new values to update some of the basic properties of the neurons used in Abigail's model. Alternatively, Chris can also use the output of Abigail's simulations to feed another simulator.
7) Chris launches another job in Abigail's session, since she granted him the permission to do this. One or more sessions may be running simultaneously on the system. The instantaneous partitioning of resources among sessions may be dynamic in order to optimise the utilisation of the system, minimise response times and ensure that each

user gets a fair share of the resources on a medium-term basis (hour) in proportion to her/his total allocation by an Access Committee.

8) Abigail will be called the main user of the session and she will be the only one allowed to request extension or reduction of the total amount of resources allocated to the session.

9) Bill or Chris can connect to the on-going session of Abigail, if they have granted rights. They can launch jobs in the same session. All resource usage will be charged to Abigail.

10) Since the balance of computational demand between the different jobs may change along time, within a session the dynamically redistribution of resources will be enabled.

11) Abigail can change the total allocation of her session, assigning new resources or releasing them based on her requests. This will be possible either through command line utilities or from within a running job by means of an API.

12) Even if all resources are already allocated to running sessions, urgent sessions may be started in specific cases by a properly authorised system manager who can decide on the fraction of resources withdrawn from already running sessions and allocated to the urgent session.

2.2.3 Workflow Creation with PyCOMPSs

Use Case ID: SP7-UC-003

Primary Actor: Scientific Developer, Daniel

Success Scenario:

1) Scientific Developer Daniel would like to implement a scientific application as a PyCOMPSs workflow.

2) Daniel programs the code in Python and adds the corresponding PyCOMPSs task definitions to implement the PyCOMPSs workflow.

3) Daniel tests the new workflow locally with a small data set on his own machine.

4) Daniel creates an HBP Task description that contains the PyCOMPSs library as a dependency.

5) Daniel will use the Task registration tool to publish his PyCOMPSs workflow as an HBP Task in the UP Task Repository for everyone to use.

2.2.4 Deployment of New Application Code

Use Case ID: SP7-UC-004

Primary Actor: Scientific Developer, Harry

Success Scenario:

1) A new release version of an application (e.g. simulation) code is ready to be deployed for production on the HPC Platform.

2) Harry transfers the application source code to the system.

3) Harry configures, builds and installs the software according to the software installation guidelines.

4) Harry creates user documentation and help pages.

2.2.5 Correction of Faulty Application Code

Use Case ID: SP7-UC-005

Primary Actors: Scientific Developer Harry and Scientific User Anna

Success Scenario:

1) Anna reports a failure or wrong behaviour of an application code via the central UP ticket system.
2) The ticket is assigned to developer Harry, who is responsible for this application.
3) Harry logs in to the system and uses a debugging tool to locate the error and fix it.
4) Harry re-installs the corrected application and closes the ticket.

2.2.6 Optimisation of Inefficient Application Code

Use Case ID: SP7-UC-006

Primary Actors: Scientific Developer Harry and Scientific User Anna

Success Scenario:

1) Anna detects and reports inefficient execution of an application code via the central UP ticket system.
2) The ticket is assigned to developer Harry, who is responsible for this application.
3) Harry logs in to the system and uses performance analysis tools to locate the performance bottlenecks and fix them.
4) Harry re-installs the corrected application and closes the ticket.

2.2.7 HBP Cockpit Request

Use Case ID: SP7-UC-007

Primary Actor: Scientific User, Maria

Success Scenario:

1) Maria enters the URL of the UP.
2) Maria selects the HPC Web GUI in the UP.
3) Maria selects the HBP Cockpit Hardware (CAVE/Tiled Display Wall) section as part of the HPC Web GUI.
4) Maria selects the project in which the data are defined for visualisation or creates a new one.
5) Maria defines the tools for visualising the selected data and the cockpit implementation:
   - CAVE
   - Tiled Display Wall
6) Maria visits the booked site and works with the visualisation in immersive 3D or on the high-resolution display wall.

2.2.8 Provenance Tracking in Interactive Visualisation

Use Case ID: SP7-UC-008

Primary Actor: Scientific User, Maria

Success Scenario:

1) Maria starts the UP on the tiled display wall and selects a certain project with data for the visualisation she wants to start.
2) Maria starts VisTrails on the tiled display wall.
3) Maria visualises the data using Vtk, embedded into VisTrails, on the tiled display wall.
4) After finishing the work with VisTrails, Maria stores the trails she created.
5) Maria goes back to the UP and connects the generated trails to the project in the UP.

2.2.9 Navigation through Multi-scale Representations

Use Case ID: SP7-UC-009

Primary Actor: Scientific User, Maria

Success Scenario:

1) Maria wants to visualise a neuroscientific data set consisting of a large set of morphologies (either independent or arranged in a circuit).
2) Maria chooses the data set she is going to visualise.
3) Maria selects and launches the rendering service to visualise the data using the UP.
4) The rendering service runs inside the UP or as a cockpit component in a resource booked by a HBP Cockpit request (2.2.7).
5) Using the interaction devices, Maria can navigate through realistic representations using traditional 3D camera-based techniques and select different levels of detail.
6) Maria can also visualise complex data using expressive representations at different levels of abstraction.
7) Maria can link both the realistic and the abstract representations.
8) When the views are linked, semantic navigation and zooming in one view updates the rendering on the other view.

2.2.10 Search and Filter

Use Case ID: SP7-UC-010

Primary Actor: Scientific User, Ana

Success Scenario:

1) Ana would like to explore a large and complex set of neuroscientific data, using search and filter techniques complementary to visualisation and interaction.
2) Ana loads the data set.
3) Ana selects and launches the rendering service to visualise the data using the UP.
4) The rendering service runs inside the UP or as a cockpit component in a resource booked by a HBP Cockpit request (2.2.7).
5) Ana can visualise the input data and she can interactively filter out data using both automatically extracted metadata or/and manually annotated metadata.
6) Ana can select one or more items within the filtered input data.

7) Ana can use the selected items to start a query in order to find similarities in the data set.
8) Ana can visualise the results selecting different representations to help her extracting knowledge from the data set.
9) Ana can interact with the results in different ways. For example, she could use data filtering to ease the interpretation of the results, or she could select or discard query items to add relevance feedback and guide the search process by obtaining new results. She could also annotate the results by adding user-defined metadata.
10) When Ana is happy with the results she can save the session and/or the analysis results.

2.2.11 High-Volume Experimental Data Management, Processing and Visualisation

Use Case ID: SP7-UC-011

Primary Actor: Scientific User, Abigail

Success Scenario:

1) Abigail produces high-resolution experimental brain tomography data. The data set is composed of thousands of small files with a total size of the order of 5 TB. The data are stored on her local system connected to the experimental tomography acquisition tool, outside of the HPC Platform.
2) Abigail transfers the data to the HPC Platform to:
    - Preserve the data
    - Share the data with other users
    - Perform pre-processing activities
    - Perform data analytic activities
    - Perform remote visualisation of the data stored and new data produced by the processing activity.
3) Abigail or her colleagues can exploit a dedicated visualisation facility equipped with specialised tools and instruments (i.e. GPUs) that are not always available at the local facilities.
4) Abigail's data are stored on the HPC Platform together with other data produced by the analytics post-processing. The data can then be analysed and visualised directly on the HPC Platform without the need to transfer back the data to Abigail's own site.

2.2.12 Spatial Join of Two Dense Neuron Data sets

Use Case ID: SP7-UC-012

Primary Actor: Scientific User, Abigail

Success Scenario:

1) Abigail would like to perform a brain model-building task by joining two neuron data sets on a Cloud storage resource. Both of the data sets are very dense.
2) The spatial-join task is computationally expensive with existing index methods and algorithms.
3) The HPC Platform loads both data sets into memory, builds in-memory indices on one or both data sets and applies a scalable spatial-join algorithm to execute the task.

4) The HPC Platform returns the result to Abigail within a reasonable amount of time.

2.2.13 Spatial Join of Neuron Data sets with Different Densities

Use Case ID: SP7-UC-013

Primary Actors: Scientific User, Abigail

Success Scenario:

1) Abigail would like to perform a brain model-building task by joining several sparse neuron data sets with a dense neuron data set.
2) The spatial-join task on two data sets with contrast density is computationally expensive with existing index methods and algorithms.
3) The HPC platform indexes the dense data set, traverses on the sparse data sets and applies a scalable algorithm to determine the elements to be joined.
4) The HPC Platform returns the results to Abigail within a reasonable amount of time.

2.2.14 Query of the Structural Connectivity of a Dense Data set

Use Case ID: SP7-UC-014

Primary Actors: Scientific User, Abigail

Success Scenario:

1) Abigail would like to perform a brain simulation task by querying a data set of spatial brain models in an external data centre.
2) Abigail queries the structural connectivity, i.e., the places where two neuron branches touch, of the data set.
3) The entire data set is extremely dense so that readily available index methods do not scale due to the page-reading overhead.
4) The HPC Platform builds a scalable index on the data set to decouple the page-reading overhead from the density of the data set.
5) The HPC Platform executes the simulation task and returns the results to Abigail within a reasonable amount of time.

2.2.15 Interactive, Exploratory Querying

Use Case ID: SP7-UC-015

Primary Actor: Scientific User, Abigail

1) Abigail creates a database project with the objective of building a repository of simulation data, clinical data, and background.
2) Abigail provides an initial (integrated) scheme with strict user privileges and roles.
3) Abigail identifies the location of several file repositories for integration into the system.
4) The meta-data are extracted from the file repository, or the CSV-data sources are injected straight away.
5) Abigail deploys any of the SQL-based graphical user interfaces to explore both the consolidated information and the data transparently accessible from the file store, e.g. using in-situ processing. The SciQL variant of SQL is used to manipulate the mostly geometric data directly and declaratively.

6) Having had a glimpse of a scientific discovery, she defines an in-lined R table producing an SQL function to process the data selected and aggregated with SQL predicates.
7) Upon inspection of the R-visualisation she notices a need to convert the data.
8) The underlying database optimiser exploits the data-parallel semantics of SQL to execute the code on the Platform in the most efficient way.
9) Coming back to the main workflow, the results produced are stored in the database to leverage the work for others.

2.2.16 Exascale Data Analytics

Use Case ID: SP7-UC-016

Primary Actor: Scientific User, Abigail

Success scenario:

1) Abigail performs a large-scale brain simulation.
2) The time-series data resulting from the simulation is massive, both in terms of the velocity of individual time-series readings and the number of time series generated, making it impossible to handle/analyse the data using conventional data management technology.
3) The HPC Platform analyses the resulting massive streams of simulation data in real time, building concise synopses and approximate indexing structures over the streaming time-series data.
4) The HPC Platform employs the synopses and indexing structures to allow Abigail to accurately monitor (in real time):
   a) Strong correlation and anti-correlation patterns in the streaming time series
   b) The validity of the simulation readings against externally provided models and/or data from clinical studies of the human brain.

2.3 HPC Platform: Functional Requirements

Here we list the functional requirements to the HPC Platform that follow from the use cases descibed in the previous section.

2.3.1 HPC, Data Management and Visualisation Capabilities

Requirement ID: SP7-FR-001

The HPC Platform should provide:

- Production-quality high performance computing capabilities for brain simulations, molecular dynamics (subcellular) simulations, and massive data analytics.
- A high performance computing system for code development, testing and optimisation.
- Capabilities for data storage and fast Platform-internal data transfer.
- High-resolution, 3D and fully immersive visualisation capabilities.

2.3.2 Single Sign-on to All HPC Platform Resources

Requirement ID: SP7-FR-002

- Users must be uniquely identifiable and have explicit access permissions for all services of the Platform.
- User management and access to individual Platform resources should be governed by a common, HBP-wide access policy.
- The diverse tasks executed by the end users will need to interact with authenticated UP services at job completion. This requirement will be contingent upon possible authenticated session timeouts and the potential unavailability of authentication tokens in various tiers of the HPC Platform job service.
- A Computer Emergency Response Team (CERT) (see Section 2.4.6) should be in place to revoke a user's authentication token in case it was compromised.

2.3.3 HPC Job Management

Requirement ID: SP7-FR-003

2.3.3.1 Submission of HPC Jobs from a Single, Common Entry Point

- The Platform should allow the submission of HPC jobs from a single, common entry point accessible through the UP.
- The Platform should allow for interactive, command-line based access for Developers.

2.3.3.2 Dispatch of HPC Jobs Automatically to the Best-Suited HPC Resource

- The UP should list all available HPC resources and services and their technical characteristics.
- The UP should make use of execution estimation services in the HPC Platform.
- Tasks published in the UP should contain information on suitable HPC resources.
- The UP should provide meta-scheduling capabilities such as cross-site synchronised scheduling and shortest time-to-execution scheduling.
- The Platform should provide a data staging mechanism to move data, if necessary, to the location of the best-suited HPC resource before job execution.

2.3.3.3 Monitoring of HPC Resources Availability

- Each site should monitor the usage of compute, network and storage resources and provide this information in a commonly agreed-upon format.
- The Platform should provide a central registry of HPC Platform compute resources that can be queried by the UP.
- Resources available to the HPC Platform should be registered in the central registry.
- The Platform should allow the UP to report on health checks performed on various HPC resources. Health checks should report at least the following:
  - Resource availability
  - Resource degradation/failure
  - Upcoming scheduled downtime.

2.3.3.4 Accounting of Used HPC Resources

- The HPC Platform should support resource management for executed tasks according to project-specific quotas and data permissions. Quotas will be per group and/or per user and will target the following resources:

– Computing hours and memory usage
– Data storage
– Network transfers.

2.3.4 Software Management

Requirement ID: SP7-FR-004

- The HPC Platform should provide a software distribution mechanism across the HPC sites for Task Framework components.
- The distribution should be done securely, consistently and in a repeatable manner.
- Runs of the same task at different sites of the HPC Platform should always deliver the same result.
- Software distribution mechanisms should exist for distributing the UP Task Framework and Analysis Toolkit to Developers. License restrictions should be considered.
- Users should be able to create new versions of existing software and register their new software in the UP.

2.3.5 Development Platform/Environment for Developers

Requirement ID: SP7-FR-005

- The Platform should allow (command-line based) interactive access for selected HBP developers.
- The Platform should provide a full software development environment, including but not limited to:
  – Parallel programming environment
  – Compilers
  – Debuggers
  – Performance analysis tools.

2.3.6 Data Storage, Retrieval and Processing via Traditional Repositories or Cloud-Based Storage

Requirement ID: SP7-FR-006

- The HPC Platform will store, retrieve, process and manage HBP data in traditional file system repositories or Cloud-based storage.
- The storage for various data files should be accessible through the UP.
- The storage should be suited to accommodate:
  – Low-density, i.e., large numbers of small files
  – High-density, i.e., smaller numbers of large files.
- The HPC Platform should provide a metadata service. Functionality for marking data as hidden to other users is required.
- The HPC Platform should provide a mechanism for the UP to establish links to data.
- The HPC Platform should allow for seamless duplication of linked data (link “upgrade”) to various HPC resources when required.

- The HPC Platform should allow for staging of linked data, i.e., the automatic migration of data to the HPC resource where it will be needed.
- Granular permission granting scenarios must be supported, and can be managed by the UP.
- The HPC Platform should allow data files to be explicitly shown to or hidden from other users/groups.
- The HPC Platform should allow read/write/delete permissions to be explicitly granted to users/groups for individual files and/or folders.
- Data administrators should be able to prevent certain data from being modified/deleted.
- The HPC Platform should provide a regularly updated list of potential destinations for output data.
- The HPC Platform should provide a service that facilitates the reliable transfer of data between sites using high performance bulk transfer protocols. The UP will be able to interact with this service.
- Authentication and authorisation for access to data are required. This should be integrated with the authentication to the UP and should support appropriate delegation schemes.

2.3.7 Data Management, Querying and Analysis

Requirement ID: SP7-FR-007

- The HPC Platform should enable users to query indexed model data based on spatial predicates and to perform spatial operations (e.g., spatial join).
- Spatial queries should allow scientific users to query different types of spatial objects used in the models (e.g., neurons, synapses, etc.).
- The HPC Platform should enable users to query time series data resulting from a simulation using simple predicates (based on time but also on other attributes including spatial attributes).
- The HPC Platform should allow for the complex analysis of a simulation result, i.e., it should enable scientific users to mine the simulation result for interesting time series patterns.
- The HPC Platform should support basic database services that provide a relational interface to users.
- The database part of the HPC Platform should feature *in-situ* processing and analysis of data so that data do not need to be imported into the database first.
- The database should also provide the user with the functionality to define R code for data analysis in-line.

Table 1 below shows the relationship between functional requirements and use cases. The grey fields indicate for each functional requirement the use case(s) it is intended to address.

	SP7-FR-001	SP7-FR-002	SP7-FR-003	SP7-FR-004	SP7-FR-005	SP7-FR-006	SP7-FR-007
SP7-UC-001	■	■	■			■	
SP7-UC-002	■	■	■			■	
SP7-UC-003		■	■	■	■		
SP7-UC-004	■	■		■	■		
SP7-UC-005		■		■	■		
SP7-UC-006		■		■	■		
SP7-UC-007	■	■				■	
SP7-UC-008	■	■				■	
SP7-UC-009	■	■				■	
SP7-UC-010	■	■				■	
SP7-UC-011	■	■	■			■	
SP7-UC-012	■	■	■			■	■
SP7-UC-013	■	■	■			■	■
SP7-UC-014	■	■	■			■	■
SP7-UC-015	■	■	■			■	■
SP7-UC-016	■	■	■			■	■

Table 1: Relationship between Functional Requirements and Use Cases

2.4 HPC Platform: Non-Functional Requirements

In this section we list non-functional requirements for the HPC Platform as a whole.

2.4.1 Scalability and Flexibility

Requirement ID: SP7-NFR-001

Although the current use cases described in Section 2.2 are helpful in describing what the Platform will do, it is currently difficult to derive quantifiable requirements such as:

- Expected number of concurrent users for each use case
- Expected resource consumption per user for each use case
- Data volumes and transfer speeds (although orders of magnitude are available for some use cases).

Consequently, the HPC Platform should be constructed such that it has, in principle, the flexibility to grow and accommodate current and future quantitative requirements. Any upgrade of the Platform will of course be constrained by technological feasibility and available budgets.

2.4.2 Interoperability Between Sites

Requirement ID: SP7-NFR-002

The HPC Platform should provide the necessary tools and infrastructure to allow the participating sites to cooperate efficiently and to increase scientists' productivity. In particular:

- Common user environments should be deployed on the various HPC sites.
- Software should be deployed in a consistent manner on the various HPC sites.
- Resolutions to problems identified at one site should be communicated appropriately to the other HPC sites for consistent system improvement.

2.4.3 Network Topologies, Entry Points to the HPC Platform

Requirement ID: SP7-NFR-003

- The main sites hosting HPC resources for the HPC Platform should be connected via a high throughput network.
- Users will access the HPC Platform through their regular Internet links, via the UP.

2.4.4 Performance

Requirement ID: SP7-NFR-004

- It should be stressed that optimising HPC Platform performance is not simply about maximising compute performance in terms of Flops; it is also about improving the overall time-to-solution, and taking into account all hardware and software components.
- All the hardware and software components of the HPC Platform should be balanced out to offer fair resource allocation to the various workflows and to avoid bottlenecks.
- Services should be allowed to run with the required Quality of Service to ensure successful completion.

2.4.5 Data Protection (Backup/Retention Policies, Disaster/Recovery)

Requirement ID: SP7-NFR-005

- The HPC Platform should offer various data protection levels (backups, RAID, data retention policies).
- Users should be able to classify their data with regard to the required protection level (e.g., short term storage vs. long term archiving).
- A disaster recovery strategy should be provided.

2.4.6 IT Security (CERT/CSIRT Team, Security Forum, Security Policies)

Requirement ID: SP7-NFR-006

The HPC Platform should have an Information Technology Security Policy, comprising of the following elements:

- Each multi-domain infrastructure lacks information about the security level that the other partners have implemented. Mutually exchanging security polices or at least providing information on security policies implemented will help to generate a web of trust between the participating sites. Only with this "Web of Trust" close collaboration between these sites can be initiated. The HPC Platform should establish such trust between sites by exchanging security policies and working together in setting up and handling the infrastructure in a secure manner.
- An HBP-AUP/ToU (Acceptable Use Policy/Terms of Use) should be elaborated, so that the users of the HPC Platform are aware under which conditions they are allowed to use the infrastructure provided. Users not adhering to these regulations may be sanctioned by revoking their user account and/or access permissions.
- HPC Platform sites should implement an operational IT Security Team (CERT/CSIRT) to handle any security incidents coming up during operation of the Platform. The team should consist of at least one staff member of every site. Optimally, the CERTs of all sites work closely together and exchange any security relevant information.
- Any service provided by the UP should be evaluated for security risks by a Security Forum composed of IT Security experts of the partner sites. Services should only be made operational, if the "Security Forum" has concordantly agreed on its production readiness.

2.4.7 Documentation and User Training

Requirement ID: SP7-NFR-007

Documentation and training should help users with different backgrounds and experience to make effective use of the HPC Platform.

- The HPC Platform should be documented in such a way that new users are able to execute simple workflows in less than two hours.
- Advanced or more specific HPC-related topics, e.g. for Scientific Developers, should be the subject of suitable training courses.

2.5 HPC Platform: Software

In this section we describe the software components of the HPC Platform. The focus here is on software requirements for which one or more solution(s) already exist. For each functional requirement of Section 2.3, a summary table lists a description of the required software functions and proposes one or more existing software packages. The table also indicates at which sites of the HPC Platform the software is or should be installed. Then, a more detailed explanation follows on how the software should be deployed in order to meet the various requirements.

2.5.1 HPC, Data Management and Visualisation Capabilities

Related functional requirement: SP7-FR-001

All of the software categories described in the rest of this section will be necessary to meet this requirement.

2.5.2 *Single Sign-on to all HPC Platform Resources*

Related functional requirement: SP7-FR-002

Function	Description	Proposed Software	Site
Single sign-on	Global User Administration System (UAS) on top of existing local UASs *First prototype by M10*	Lightweight Directory Access Protocol (LDAP)	All sites

Table 2: Software for Single Sign-on to all HPC Platform Resources

Traditionally, any HPC system has its own user database and uses its own authentication and authorisation mechanisms. Implementing single sign-on would require that every user could be mapped by a general authentication system to the special user ID required for the specific HPC resource. Since the HBP HPC resources are part of the HPC systems installed at the different HBP HPC sites, it cannot be anticipated that the login mechanisms will be changed to fit the HBP single sign-on prerequisites. Therefore, an alternative approach has been chosen.

A User Administration System (UAS), which does not make any assumptions about these systems, will be built on top of existing local user administration systems. This will make the UAS as non-invasive as possible with a minimum of local adaptations required. Every site can perform administrative tasks for its own users. Policies define which tasks can be performed.

All HBP sites have read access to the UAS of other sites, so that if users are added or attributes are changed, other sites are able to verify this and make corresponding changes. The UAS will be implemented in a secure way, so that unauthorised persons cannot tamper with the information available. It will use a uniform namespace, with a strictly defined set of allowed attributes and values. This way, all information in the UAS is stored in exactly the same way at each site.

To provide this kind of UAS, the Lightweight Directory Access Protocol (LDAP) has been selected for the implementation. LDAP is widely used and accepted as a user administration system and directory service in general. LDAP services can also be adapted quite easily to interface with different information systems (back-end databases). The LDAP protocol is used for communication, but the repository itself can be almost any system of choice.

The information can be hosted by a set of distributed servers, each serving a partition, which is a complete subtree of the authoritative domain for which a site is responsible. These servers can access each other and update information on a periodic time schedule. It is also possible to use only one server. Sites/partners will only have write access to their own site-specific subtrees. Everything below such a site domain suffix will be managed by that site. Other sites/partners will have read access to the subtrees if the information is needed for the local administration of users or internal services.

The UAS will not be queried directly for authentication or authorisation purposes. The update process of the local site administration of users operates in two steps. In the first step, information is extracted from all the LDAP servers. In the second step, this information is used to create or modify local user accounts and to update local authorisation information. This second step will differ between sites and must be implemented by each site separately.

Within the first step, the client queries the main LDAP server. The main LDAP server sends back referrals of all LDAP servers in use. The client queries all LDAP servers, including the main server. All LDAP servers send back the requested data.

For the second step, updating the local user administration, sites will have to create tools and procedures to incorporate newly created and modified information of the LDAP servers into their local user administration and authorisation facilities. Examples of how to implement this are provided by the PRACE project, which has chosen a similar approach.

2.5.3 HPC Job Management

Related functional requirement: SP7-FR-003

Function	Description	Proposed Software	Site
HPC workload management	Uniform Interface to Computing Resources providing a seamless interface for preparing and submitting jobs to a variety of computing resources. First prototype by M10	UNICORE	All sites
	Registry server providing information about available HPC services First prototype by M10	UNICORE Registry	At least two sites for high-availability configuration
	Monitoring tool to detect infrastructure problems by executing periodic, automated, user-level testing of software and services First prototype by M10	Nagios/INCA	All sites
	Graphical monitoring of batch system controlled clusters	LLview	All sites
Job submission	Job scheduling and management (site-specific)	LoadLeveler - IBM parallel job scheduling system	JSC, Cineca
		PBS - Portable Batch System: job scheduling computer software	Cineca

Function	Description	Proposed Software	Site
Job submission cont'd		Slurm - Open-source resource manager	CSCS, BSC (MinoTauro)
		LSF - Platform Load Sharing Facility: commercial job scheduler	BSC (MareNostrum)

Table 3: Software for HPC Job Management

The management of the computational resources provided by the HBP HPC machines will be handled by job scheduler software to maximise the overall productivity. The scheduler is responsible for managing jobs (program execution) on the machines by allocating partitions for the user on the compute nodes, returning job output and error files to the users. Because of site-specific local technical or contractual requirements, each site will continue to use its own job scheduling system. However, through the adoption of a common workload manager such as UNICORE, users will have a single entry point for the submission of computational jobs and the use of the HPC machines without entering the complexity underneath. Complementing job submission, the common workload manager (UNICORE) will provide data staging functionality, both for data residing on the Platform (i.e. other HPC sites or the Cloud storage), as well as storage that is external to the HPC Platform. The centralised monitoring system will monitor the workload of each individual machine as well as the status of its service.

2.5.4 Software Management

Related functional requirement: SP7-FR-004

Function	Description	Proposed Software	Site
Software management	Dynamic modification of a user's environment	Environment Module packages	All sites
	Software repository	Git, apt, yum	All sites

Table 4: Software for Software Management

The availability of scientific software for the users will be managed by the Environment Modules commands. On a complex HPC system, where it is necessary to make available a wide choice of software packages in multiple versions, it can be difficult to set up the user environment to allow users to always find the required executables and libraries. The Modules package provides for the dynamic modification of a user's environment via *modulefiles*: each *modulefile* contains the information needed to configure the shell for an application.

To guarantee the constant update of the scientific software on the various machines, tools for Continuous Integration will be investigated in collaboration with end-users. Examples of software distribution mechanisms might be a git, apt or yum repository.

See also Section 2.5.8 on application software.

2.5.5 Development Platform/Environment for Developers

Related functional requirement: SP7-FR-005

Function	Description	Proposed Software	Site
Programming	General purpose, high-level programming language	Python	All sites
	Python modules	NumPy, SciPy	All sites
	Fortran77 and Fortran90 compilers	ifort, gfortran	All sites
	C and C++ compilers	Icc, icpc, gcc, g++	All sites
	Blue Gene/Q Cross-compiler	bgq-xl	Cineca, JSC
Parallel programming	Standardised and portable message-passing system	Message-Passing Interface (MPI) library	All sites
	API supporting multi-platform shared-memory parallel programming	OpenMP	All sites
	Extension of OpenMP with new directives to support asynchronous parallelism and heterogeneity	OmpSs	All sites
	Python bindings for OmpSs First implementation M6 Integration with HBP Platform M12 Advanced features to support multi-scale simulations M18	PyCOMPs	All sites
	Dynamic Load Balancing (DLB) and resource management	DLB and resource management libraries	All sites
Debugging	GUI-based source code defect analysis tool to control over processes and thread execution and visibility into program state and variables	TotalView	All sites
	Profiling tool	gprof	All sites
	Tool to detect memory management and threading bugs, code profiler	Valgrind	All sites

Function	Description	Proposed Software	Site
Profiling and performance analysis	Highly-Scalable call-path profiling and trace analysis	Score-P, Scalasca	All sites
	Sophisticated trace measurement and visualisation	Extrae, Paraver	All sites
Collaborative code development	Free and open source distributed version control system	Git	All sites
	Open-source, centralised version control system	Subversion	All sites

Table 5: Software for Developers

In order to provide Developers with a complete environment for the development of their software, a set of tools and libraries will be made available on the machines:

- Compilers for the most used programming languages
- Debuggers, profilers and performance analysis tools for testing, correction and improvement of code
- Tools for collaborative code development.

2.5.6 Data Storage, Retrieval and Processing via Traditional Repositories or Cloud-Based Storage

Related functional requirement: SP7-FR-006

Function	Description	Proposed Software	Site
Data storage, retrieval and processing	Data streaming library and file transfer tool First prototype by M10	uFTP	All sites
	High-speed, reliable, and secure data transfer protocol First prototype by M10	GridFTP	All sites
	Distributed parallel file system: provides online storage management, scalable access, and integrated information lifecycle management tools capable of managing petabytes of data and billions of files	GPFS	All sites

Function	Description	Proposed Software	Site
Data storage, retrieval and processing (cont'd)	Remote file caching technology that ensures seamless and high performance access to files no matter geographical location First prototype by M12	GPFS/AFM	The adoption of the AFM technology will be evaluated during the course of the project.
	Integrated Rule-Oriented Data-management System, a community-driven, open source, data grid software solution	iRODS	The adoption of the iRODS technology will be evaluated during the course of the project.
	System for storing and retrieving huge amounts of data, distributed among a large number of heterogeneous server nodes, under a single virtual file system tree with a variety of standard access methods	dCache	The adoption of the dCache technology will be evaluated during the course of the project.
	Scalable I/O library for parallel access to task-local files	SIONlib	All sites
Cloud Storage	Amazon-S3 compatible interface	CEPH	KIT
Data Analytics	Software programming language and software environment for statistical computing and graphics	R	All sites
	An interface to MPI (Message-Passing Interface) for statistical computing	RMPI	All sites
	A framework allowing for the distributed processing of large data sets across clusters of computers using simple programming models First prototype by M12	Hadoop	Cineca

Table 6: Software for Data Storage, Retrieval and Processing via Traditional Repositories or Cloud-Based Storage

Starting with (independent) GPFS file systems at the different HPC systems, software components will be installed that allow to access data independently of the location (i.e. HPC storage system). Starting with file transfer tools like GridFTP and Unicore uFTP, in a second phase, distributed file system technologies will be used.

AFM is a powerful remote file caching technology that ensures high performance access to file-based information no matter where it might be geographically located. This feature is useful for the HBP since large volumes of file-based data are shared and require high levels of information availability no matter where it resides. A new version of AFM will be available in summer this year. This new version will be tested as access method to HBP data distributed over the different HPC file systems.

The Cloud storage located at KIT will be accessible through an Amazon S3-compatible interface. The Cloud storage will be used as repository for secondary data: the scientist user will be able to upload/download data to/from the cloud storage through a standard Internet connection.

Scientific user(s) will be able to preserve and share data in the Cloud storage using a persistent URL. This feature is essential for use cases where persistence of data localisation and content is required.

The Cloud storage will be based on Ceph, a highly scalable and performing distributed storage system. Ceph provides a RESTful interface, allowing a range of basic operations such as uploading, downloading and deleting objects, as well as sharing data with HTTP URLs.

The authentication and authorisation will be integrated with the Platform's infrastructure in order to have a single sign-on common to all HBP resources. There are many technical approaches to reach this goal, and a thorough evaluation of them will dictate which one will be chosen.

Ceph uses replication to ensure data reliability and recovery. Furthermore, Ceph offers a block device interface, which can be used as back-end for a Cloud computing system like OpenStack. The users can have access to on-demand computing resources by launching virtual machines at the KIT Cloud when needed.

2.5.7 Data Management, Querying and Analysis

Related functional requirement: SP7-FR-007

Function	Description	Proposed Software	Site
Data management, querying and analysis	Column-store database management system First prototype by M12	MonetDB	tbd
	Relational database First prototype by M12	PostgreSQL	tbd
	Libraries for the analysis of massive spatial model data	FLAT	tbd
	Tools for the mining and analysis of simulation time-series on the exascale		tbd

Table 7: Software for Data Management, Querying and Analysis

To provide users with the ability to analyse data beyond data on the file level, a complete suite for storing, indexing and querying model and simulation data will be made available:

- Database System: A database system for storing, indexing and querying general scientific data.
- Libraries for the analysis of spatial and spatio-temporal data: To analyse the brain models used as input and generated as output from the simulations based on spatial and spatio-temporal predicates.

- Time series mining tools: Efficient tools for finding complex patterns in massive time series resulting from simulations. The tools will run distributed to analyse the temporal data from simulations.

2.5.8 Application Software and Libraries

Related functional requirement: SP7-FR-004

The following list of software is not comprehensive, but it should demonstrate the wide range of codes that will be required to run on the HPC Platform for simulation, visualisation, image processing and other applications. We expect that the user community will request a lot of additional software. As already indicated in Section 2.5.4, proper distribution mechanisms, common installation methods and adequate documentation are required so users can find and reliably use the right tool.

Function	Description	Proposed Software	Site
Simulation	Spiking neural network simulator	NEST	All
	Empirically-based simulator of neurons and networks of neurons	NEURON	All sites
	Exact stochastic simulator of reaction-diffusion systems in arbitrarily complex 3D geometries	STEPS	All sites
	Molecular dynamics simulator for simulations of proteins, lipids and nucleic acids	GROMACS	All sites
	Molecular dynamics simulator	Tremolo-X	All sites
	Flexible numerical software environment for solving partial differential equations on huge clusters	UG 4	All sites
Libraries	Basic numerical libraries	BLAS, LAPACK, FFT, GSL, GMP, SPRNG	All sites

Function	Description	Proposed Software	Site
Libraries cont'd	Parallel numerical libraries	ScaLAPACK, ARPACK, PARPACK, PETSc	All sites
	Library for Accelerated Math Applications, framework for building efficient, extensible and flexible solvers for sparse linear systems and in application domains that involve - sparse and dense - numerical linear algebra. Supports shared and distributed memory compute architectures, including accelerators, through an easy to use interface.	LAMA	All sites
	Multi-GPU, Distributed Memory 3D Fast Fourier Transform	FFT	All sites
	Hyperbolic Cross Fast Fourier Transform, a high dimensional pseudospectral methods sparse grid based library (serial version)	HCFFT	All sites
Visualisation	Handy, fast, and versatile 3D/4D/5D Image Visualisation & Analysis System for Bioimages & Surface Objects	Vaa3D	Cineca
	Virtual Reality Toolkit for the development of virtual reality applications with a focus on immersive, scientific visualisation.	ViSTA	RWTH Aachen University
	An open-source software system for 3D computer graphics, image processing and visualisation.	Vtk	All sites
	An open-source scientific workflow and provenance tracking system that supports data exploration and visualisation.	VisTrails	RWTH Aachen University
Image Processing	A tool for fast automatic 3D-stitching of teravoxel-sized microscopy images	Terastitcher	Cineca

Table 8: Application Software and Libraries

2.5.9 Interoperability among Sites

Related non-functional requirement: SP7-NFR-002

The HPC sites will provide a common, interoperable environment for end-user jobs, for example through the provision of a common HBP module environment. Although no specific software will achieve interoperability, close collaboration among the various sites and rigorous deployment methods will be required.

2.5.10 Documentation

Related non-functional requirement: SP7-FR-007

The documentation will be created using a Wiki or any other tool that will permit the consultation of the materials either online or via PDF documents.

2.6 HPC Platform: Physical Architecture

In this section we describe the hardware components of the HPC Platform.

2.6.1 HPC, Data Management and Visualisation Capabilities

Related functional requirement: SP7-FR-001

During the HBP's ramp-up phase, high performance computing capabilities will be provided by the systems at four different locations:

2.6.1.1 HBP Supercomputer

JUQUEEN, located at the Jülich Supercomputing Centre (JSC) of Forschungszentrum Jülich, comprises 28 IBM BlueGene/Q racks (28,672 nodes - 458,752 IBM PowerPC® A2, 1.6 GHz cores) with an overall peak performance of 5.9 PFlops, 448 TB of main memory and an 8 PB file system with an aggregated performance of about 160 GB/sec. Nodes are interconnected using a 5D Torus network allowing 40 Gb/s transfers with a 2.5 µsec latency (worst case).

2.6.1.2 HBP Development System

Blue Brain 4, located at the Swiss National Supercomputing Centre (CSCS) and composed of four IBM BlueGene/Q racks (4,096 nodes - 65,536 IBM PowerPC® A2, 1.6 GHz cores), is able to provide 836 TFlops of peak computing performance, 65 TB of main memory and 4 PB of storage. The system also includes an iDataPlex 40-node x86 visualisation cluster, connected via a FDR Infiniband network providing 40 Gb/s node-to-node.

2.6.1.3 HBP Molecular Dynamics Supercomputer

Two supercomputers located at Barcelona Supercomputing Centre (BSC) will be used:

- MareNostrum, a 1.1 PFlops machine of 36 IBM iDataPlex racks (3,056 nodes – 48,896 Intel SandyBridge-EP 2.6 GHz cores). It provides 100 TB of RAM, 1.9 PB of short-term storage with an aggregated performance of 144 Gb/s, and 4 PB of long-term storage.
- Minotauro, a GPU cluster providing 200 TFlops and 3 TB of main memory composed of 4-rack Bull B505 (128 nodes - 768 Intel E5649 2.53 GHz cores - 256 M2090 NVIDIA GPU Cards). The nodes are interconnected through a two-link QDR Infiniband network, capable of 64 Gbp/s per node.

2.6.1.4 HBP Massive Data Analytics Supercomputer

Three integrated supercomputer systems located at Cineca will be used:

- FERMI, a 10-rack IBM BlueGene/Q system (10,240 nodes – 163,840 IBM PowerPC® A2, 1.6 GHz cores) with a peak performance of 2.1 PFlops and 165 TB of main memory. Internally connected with a 5D torus network and integrated with a 5 PB storage system suitable for data analytics.
- PLX, a GPU cluster composed of 274 IBM iDataPlex M3 nodes (2 six-core Intel Westmere 2.40 GHz per node, 48 GB/node RAM and 2 NVIDIA GPU Tesla M2070 per node) with an overall peak performance of 300 TFlops. The nodes are interconnected through Infiniband with 4x QDR switches.
- A data-intensive system for massive data analytics and post-processing. This machine, currently under procurement, will support various data access patterns either for structured or unstructured data providing high performance to I/O bound applications.

In addition to the storage available locally at the HPC centres, the HPC Platform will provide Cloud-based storage:

2.6.1.5 HBP Cloud Service

See 2.6.4 below.

A high throughput network for fast data exchange between the sites will connect the HPC systems of the Platform to each other:

2.6.1.6 High-throughput network

The high throughput data network will connect the HPC systems at Jülich, BSC, Cineca and CSCS as well as the data storage system located at EPFL in Lausanne. Jülich, BSC and Cineca are already connected via a dedicated 10 Gb/s network connection implemented and owned by the EU project PRACE.

Discussions have been started with PRACE to use the PRACE network for HBP data traffic between the HPC systems of the Platform. A Memorandum of Understanding (MoU) is in preparation and usage of the network should start soon. Connecting the HPC system at CSCS and the high-volume data storage located at EPFL will require additional efforts. CSCS and EPFL are already connected to each other via a 10 Gb/s link used for the Blue Brain project. This link can also be used for the HBP. Thus, only a dedicated connection from EPFL to Geneva, the GÉANT Point of Presence (PoP), and from there via GÉANT to the central network router of PRACE in Frankfurt is still needed. Discussions with Switch, the Swiss NREN, and DANTE, the European Research network provider, will start soon.

This setup will allow cost-effective testing of the usability of such a network as long as the PRACE network is available. In parallel, alternative network designs can be evaluated and prepared for future changes in user requirements.

Access to the Cloud storage at KIT will be made through regular Internet links, as no requirement for a high speed network is expected.

The HPC Platform will also include high-end visualisation environments provided by partners:

2.6.1.7 CAVE / Tiled Display Wall

Two installations at RWTH Aachen University will be used:

- AixCAVE, a 5.25 x 5.25 x 3.3 m^3 fully immersive virtual reality environment using 24 HD-DLP projectors with a resolution of 1920x1200 pixels for rear projection (4 for each wall, 8 for the floor plane). The projectors are driven by a 24-node visualisation cluster, each containing 2 Intel Xeon Westmere CPUs (6 Cores at 2.7 GHz), 24 GB RAM, 2 x NVIDIA Quadro 6000 GPUs. The complete volume of the AixCAVE is tracked by an infra-red-based tracking system with 8 ARTTRACK1 and 2 tracking cameras.

- Tiled Display Wall, a display composed of 6 tiles (Full HD resolution and 46'' size each, total resolution of 5760 x 2160 pixels) driven by a 7-node cluster equipped with one Intel XEON E5 (4 Cores at 3.7 GHz), 16 GB RAM, and one NIVIDIA GTX 780Ti GPU each. The Display Wall is further equipped with 2 ARTTRACK2 tracking cameras for interaction.

One installation at EPFL will be used:

- Tiled Display Wall, a display composed of 12 tiles (Full HD resolution for each tile, total resolution of 7680 x 3240 pixels), driven by a 2-node cluster equipped with 3 GPUs each. The Display Wall is further equipped with a PQLab multi-touch interface supporting up to 32 touch points.

Figure 1 shows a schematic overview of the physical HPC Platform architecture.

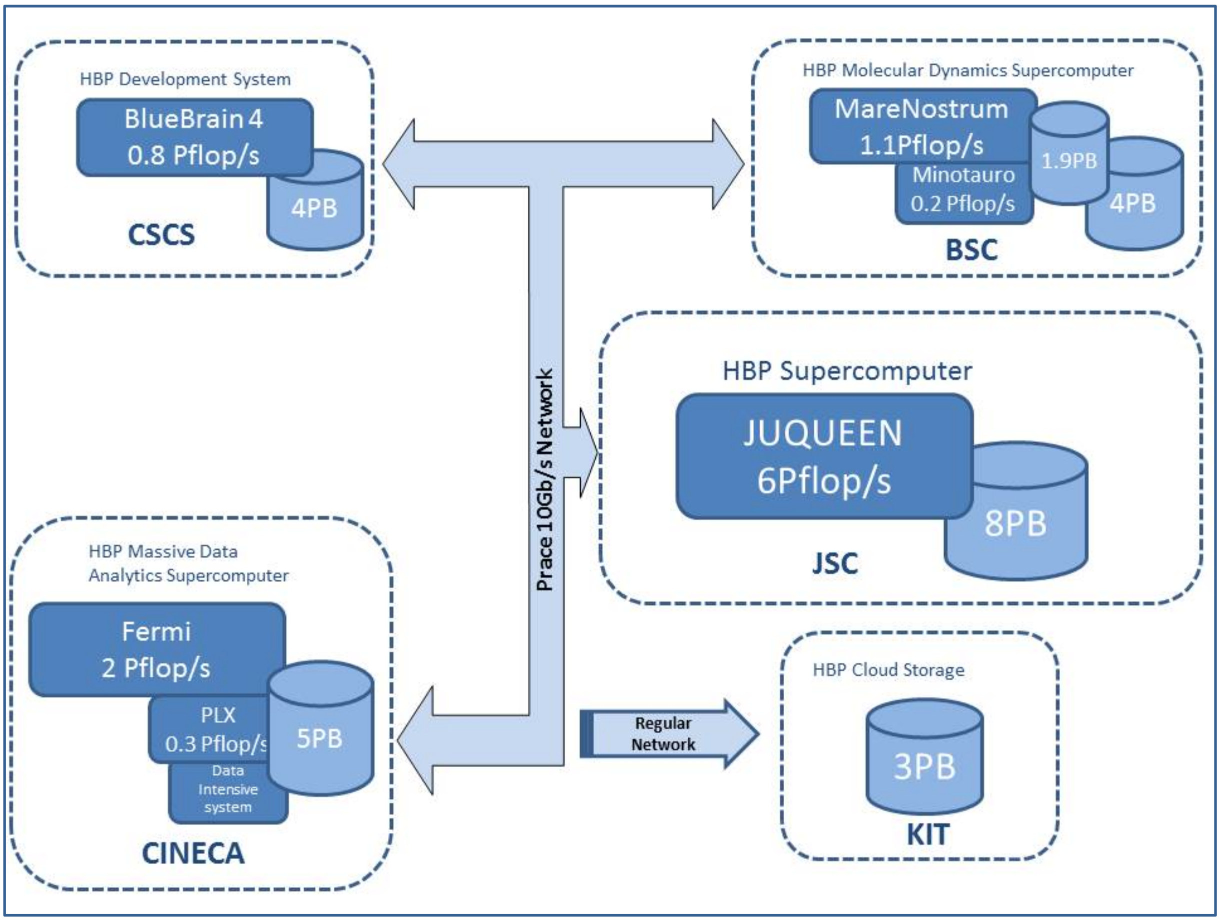

Figure 1: Physical Architecture of the High Performance Computing Platform

2.6.2 *Single Sign-on to all HPC Platform Resources*

Related functional requirement: SP7-FR-002

Each site of the HPC Platform should provide a (virtual) machine to host the distributed LDAP service.

2.6.3 HPC Job Management

Related functional requirement: SP7-FR-003

The following hardware will be required to fulfil the requirement of a HPC job management system for the HBP with capabilities of submission, dispatch, monitoring and accounting:

- A central UNICORE registry server, hosted by one partner, used as a common entry point
- A central Unity server (UNICORE authentication service)
- Each site will need to host one or several (virtual) machines to deploy an instance of the UNICORE gateway, UNICORE/X, XUUDB, UNICORE TSI and UFTPD daemons.

2.6.4 Data Storage, Retrieval and Processing via Traditional Repositories or Cloud-Based Storage

Related functional requirement: SP7-FR-006

KIT will provide Cloud storage with a capacity of 3 PB, which will be accessible through an Amazon S3-compatible interface.

2.7 HPC Platform: Relations to other Platforms

The HPC Platform is expected to serve all other HBP Subprojects/Platforms, in particular the

- Neuroinformatics Platform (SP5)
- Brain Simulation Platform (SP6)
- Neuromorphic Computing Platform (SP9)
- Neurorobotics Platform (SP10)

The HPC Platform provides a number of HPC-related services to these platforms:

- HPC user management (user authentication, authorisation, group membership)
- HPC job management (submission and management of batch jobs)
- HPC data management (management of high volumes of experimental and simulation data, and making the data available for subsequent computational tasks)
- HPC tools (visualisation, programming tools, performance analysis tools)
- HPC systems monitoring
- HPC resource usage monitoring and accounting (including quotas for compute time and storage space).

All these services must be made available for simple consumption, for example through the provision of easy-to-use programming APIs and libraries (for example, RESTful APIs). Especially important is the integration and alignment with the HBP UP in the areas of HPC user management, authentication, single sign on and delegation of rights.

2.8 HPC Platform: Prerequisites

A basic prerequisite for any significant usage of HPC Platform resources in the ramp-up phase is that these are made available to the HBP by the hosting sites. Envisaged scenarios are:

- A member site of the HPC Platform directly commits available capacity to the HBP, subject to local access regulations.
- Programme/Community Access is granted to the HBP via the peer-review instruments of the European Partnership for Advanced Computing (PRACE).

The second scenario requires negotiations with PRACE, which are taken care of by Task 7.6.2. All members of the HPC Platform are Tier-0 centres of PRACE.

In the next phase of the HBP, the availability of the ("pre-exascale") supercomputers themselves will largely depend on external funding from national funding sources. This national funding must be complemented by EU funding through e.g. a PPI as discussed in the EC's HPC strategy.

2.9 HPC Platform: Necessary Parallel Activities

2.9.1 Access Committee

As a general rule, any substantial resource allocation on the HPC Platform must always be based on a peer-reviewed project proposal. This requires the implementation of a suitable application, reviewing, ranking and allocation procedure. Under the precondition that the HBP will have certain amounts of resources at the hosting sites at its disposal (see Section 2.8), the project-internal distribution of these resources will be organised as follows:

Requests to use the Platform are selected via a competitive peer review process, supervised by a special board—the Access Committee—with eleven members: a chair (the leader of the HPC Platform), three members appointed by the HBP Board of Directors, four appointed by the hosting partners of the HPC Platform, and three independent scientists appointed by the HBP External Advisory Board. The membership of the Access Committee, which will be subject to approval by the HBP Board, will be defined to ensure an appropriate balance between the countries and disciplines present in the HBP and the broader user community. Members of the board will remain in office for five years.

2.9.2 Continuous Discussions with HBP Communities to Capture Changing HPC Requirements

The requirements of the HBP with regard to HPC resources are likely to evolve as the project progresses. New requirements may arise with advances, e.g. in brain simulation technology. Capturing and adapting to such changing HPC requirements will require continuous interaction and exchange between the HPC Platform and its consumers.

2.9.3 Coordination with other European Research Initiatives and Infrastructures

Three of the four supercomputing centres of the HPC Platform, BSC, Cineca and Jülich, are hosting members of the European Partnership for Advanced Computing (PRACE); CSCS is a so-called non-hosting member of PRACE. Therefore, to be in alignment with the EC's strategy for HPC, the provision of computing resources by these centres and access to these resources should be coordinated with PRACE. Existing and proven infrastructures and procedures should be used and leveraged wherever possible. Discussions have been started with the PRACE Council to trigger programmatic access in PRACE, and to determine how the PRACE Access Committee (proposed by the PRACE Scientific Steering Committee and appointed by the Council) can be involved in the allocation process (see Section 2.9.1).

HPC Platform members BSC, Cineca and Jülich are also members of EUDAT, the European Data Infrastructure. Through these partners, the HBP is in close contact with EUDAT to work out which EUDAT services might be useful for the HBP, and how they could be exploited. In a first meeting of HBP and EUDAT representatives in January 2014, the EUDAT services B2SHARE for data sharing and B2FIND for metadata harvesting were identified as first candidates for further evaluation.

2.9.4 Technology Evaluation: Preparing the Next Version of the Platform

In this section, we describe what the HPC Platform will look like at the end of the HBP's ramp-up phase (M30). Large-scale brain simulations running on the future pre-exascale and exascale versions of the HBP Supercomputer will need to be interactively visualised and controlled by experimenters, locally and from remote locations. "Interactive Supercomputing" capabilities will allow the HBP Supercomputer to be used like a scientific instrument, enabling *in silico* experiments on virtual human brains. These requirements will affect the whole system design, including the hardware architecture, run-time system, mode of operation, and resource management.

The HBP pre-commercial procurement (PCP), conducted by SP7 in parallel with the setup of the initial HPC Platform, aims to obtain architectural solutions that will permit such an interactive use of large-scale supercomputers. More specifically, it will procure R&D of HPC system components that allow interactive visualisation and steering of large-scale, data-intensive brain simulations on HPC architecture capable of providing a floating-point peak performance up to 50 PFLOPs. Suppliers will be required to deliver pilot systems, demonstrating the readiness of the developed technologies and their integration into a scalable HPC architecture for a representative set of HBP use cases. The pilots should be deployed and operated as "pre-production" test systems at Jülich Supercomputing Centre. The results of the HBP PCP will form the technological basis for the planned procurement of a pre-exascale production supercomputer at Jülich in the 2016-17 time frame, i.e., in the next phase of the HBP. This will be a commercial procurement, e.g. a procurement of innovative solutions (PPI), which should be supported by EC funding.

3. Software/Hardware Functions (Components of Tasks)

In this section we specify Key Performance Indicators (KPIs) for each Task in SP7, which can be used to track the progress of the HPC Platform development towards the goals defined in the Description of Work[3]. KPIs will be reported regularly to the HBP Board of Directors.

To define the KPIs, for most Tasks in SP7 we break down the Task into a distinct set of component "Functions". Implementing each Function will typically take a few months and be verifiable via one or two simple test cases. When the test cases are satisfied, the development of that Function is completed. Table 9 shows which functional and non-functional requirements are addressed by each Function. The Functions for each Task are planned so that their expected completion dates are distributed over the 30-month duration of the ramp-up phase. The number of completed Functions then represents the KPI for that Task. The KPIs defined in this way are summarised in Table 10 at the end of this section, together with their target values.

In a few exceptional cases, where Tasks do not easily lend themselves to being split into Functions, we have chosen alternative methods to define KPIs for these Tasks. These cases are also listed in Table 10.

WP7.1: Technology Evaluation (PCP)

See Deliverable D 7.7.1 (PCP Planning Document) for a detailed description of the steps involved in the PCP process and schedule.[4]

WP7.2: Mathematical Methods, Programming Models and Tools

T7.2.1: Parallel Programming Models for Interactive Brain Modelling and Brain Simulation

T7.2.1 will develop general purpose, fine-grained programming models and runtime interfaces for interactive, massively parallel, highly asynchronous applications. These models and interfaces, which will subsequently be incorporated in the High Performance Computing Platform, are essential for brain modelling and simulation. They will, however, be completely independent of any specific model or simulation and thus be applicable in a broad range of other domains. T7.2.1 will derive its development strategy from recent task-based proposals based on directionality annotations and from PGAS (Partitioned Global Address Space)-based runtime models already under development by the partners. The resulting programming environment will support high programmer productivity, as well as automatic performance optimisation and mapping to available resources in a dynamic multi-scale, multi-application framework.

Task No:	7.2.1	Partner:	BSC (P4)
Function No:	7.2.1.1	Leader:	Jesus LABARTA
Function Name:	OmpSs @ Linux based clusters		
Test Case A:	Task based OmpSs programming model can be used by application developers on Linux based clusters to increase programmability and performance of brain modelling and simulation apps		
Planned Start Date:	October 2013	Planned Completion Date:	February 2014
Requires Functions:			

Task No:	7.2.1	Partner:	BSC (P4)
Function No:	7.2.1.2	Leader:	Jesus LABARTA
Function Name:	OmpSs @ BG/Q systems		
Test Case A:	Task based OmpSs programming model can be used by application developers on the compute node kernel based BG/Q system		
Planned Start Date:	March 2014	Planned Completion Date:	May 2014
Requires Functions:			

Task No:	7.2.1	Partner:	BSC (P4)
Function No:	7.2.1.3	Leader:	Jesus LABARTA
Function Name:	NEURON and NEST @ OmpSs		
Test Case A:	First version of NEURON and NEST programmed in OmpSs can be used on Linux and BG/Q systems		
Planned Start Date:	March 2014	Planned Completion Date:	May 2014
Requires Functions:			

Task No:	7.2.1	Partner:	BSC (P4)
Function No:	7.2.1.4	Leader:	Jesus LABARTA
Function Name:	Optimised NEURON and NEST @ OmpSs		
Test Case A:	Second version of NEURON and NEST programmed in OmpSs using nesting and interfacing with the Dynamic Load Balancing Library (DLB) can be used on Linux systems		
Planned Start Date:	June 2014	Planned Completion Date:	August 2014
Requires Functions:			

Task No:	7.2.1	Partner:	BSC (P4)
Function No:	7.2.1.5	Leader:	Jesus LABARTA
Function Name:	Two other applications @ OmpSs		
Test Case A:	Two other applications from the HBP outside of the Brain Simulation subproject analysed and running in the OmpSs environment		
Planned Start Date:	September 2014	Planned Completion Date:	January 2015
Requires Functions:			

Task No:	7.2.1	Partner:	BSC (P4)
Function No:	7.2.1.6	Leader:	Jesus LABARTA
Function Name:	OmpSs runtime implementation optimised for the needs of NEURON and NEST		
Test Case A:	Optimised version of the NANOS runtime (and Mercurium compiler if needed) that improves the performance of the OmpSs versions of NEURON and NEST		
Planned Start Date:	September 2014	Planned Completion Date:	March 2015
Requires Functions:			

T7.2.2: Workflow and Distributed Programming Models for Brain Modelling

T7.2.2 will define and implement a general-purpose, medium grained workflow model that will later be incorporated in the High Performance Computing Platform, where it will enable brain modelling and simulation. The model will make it possible to specify workflows in ways that support efficient data exchange between workflow components. The implementation will take the form of extensions to existing scripting languages, and will include functionality that allows programmers to annotate the directionality of data access. This will allow the runtime system to manage locality and optimise computation and data transfer schedules. Workflows will be specified, not as graphs, as in existing systems, but as programs. This will allow the system to provide more sophisticated intelligence than current workflow management systems.

Task No:	7.2.2	Partner:	BSC (P4)
Function No:	7.2.2.1	Leader:	Rosa M BADIA
Function Name:	PyCOMPSs first version		
Test Case A:	A developer can implement PyCOMPSs apps in Python that are parallelised by the COMPSs runtime at task level in a cluster		
Planned Start Date:	October 2013	Planned Completion Date:	March 2014
Requires Functions:			

Task No:	7.2.2	Partner:	BSC (P4)
Function No:	7.2.2.2	Leader:	Rosa M BADIA
Function Name:	PyCOMPSs integrated with the HPC Platform		
Test Case A:	PyCOMPSs apps can be launched from the HPC Platform		
Planned Start Date:	April 2014	Planned Completion Date:	September 2014
Requires Functions:	7.2.2.1		

Task No:	7.2.2	Partner:	BSC (P4)
Function No:	7.2.2.3	Leader:	Rosa M BADIA
Function Name:	PyCOMPSs integrated with the HPC Platform and with advanced features to support multi-scale simulations		
Test Case A:	PyCOMPSs apps that involve multi-scale simulations can be launched from the HPC Platform.		
Planned Start Date:	October 2014	Planned Completion Date:	March 2015
Requires Functions:	7.2.2.2		

T7.2.3: Middleware for Resource and I/O Management

T7.2.3 will develop the highly dynamic resource management capabilities required for interactive supercomputing. The resources managed in this way will include processors, memory, interconnects and I/O. Specific specialised components will make it possible to capture the characteristics and demands of specific applications, to determine appropriate resource allocations and mappings, and to enforce these allocations. The resulting system will provide global optimisation for multi-scale simulation in a multi-application environment.

Task No:	7.2.3	Partner:	BSC (P4)
Function No:	7.2.3.1	Leader:	Jesus LABARTA
Function Name:	DLB support on Linux systems		
Test Case A:	An MPI + OmpSs application can be linked with the dynamic runtime library such that automatic Dynamic Load balancing is achieved on Linux based clusters		
Planned Start Date:	October 2013	Planned Completion Date:	March 2014
Requires Functions:			

Task No:	7.2.3	Partner:	BSC (P4)
Function No:	7.2.3.2	Leader:	Rosa M BADIA
Function Name:	DLB support on BG/Q systems		
Test Case A:	MPI + OmpSs applications can be linked with the dynamic runtime library such that automatic Dynamic Load balancing is achieved on BG/Q systems		
Planned Start Date:	March 2014	Planned Completion Date:	June 2014
Requires Functions:			

Task No:	7.2.3	Partner:	BSC (P4)
Function No:	7.2.3.3	Leader:	Jesus LABARTA
Function Name:	Interactive supercomputing microbenchmark to test dynamicity in resource management policies		
Test Case A:	A microbenchmark can be used to measure the utilisation and response time of the envisaged usage pattern of interactive supercomputing. In this vision, a user can launch a simulation task and, after a while, an analytic workflow can be interactively launched to analyse the results available and to monitor the progress of the simulation.		
Planned Start Date:	December 2013	Planned Completion Date:	June 2014
Requires Functions:			

Task No:	7.2.3	Partner:	BSC (P4)
Function No:	7.2.3.4	Leader:	Jesus LABARTA
Function Name:	Extend basic DLB mechanism to multiple applications		
Test Case A:	The mechanism to shift cores between processes can be used between processes belonging to different MPI applications sharing nodes.		
Planned Start Date:	June 2014	Planned Completion Date:	October 2014
Requires Functions:	7.2.3.1 and 7.2.3.2		

Task No:	7.2.3	Partner:	BSC (P4)
Function No:	7.2.3.5	Leader:	Jesus LABARTA
Function Name:	Basic mechanism to enable interactive simulation		
Test Case A:	A first integrated environment of job submission and fine grained resource node level management (DLB). The user will be able to launch a simulation task and, after a while, an analytic workflow can be interactively launched to analyse the results available and to monitor the progress of the simulation. Both applications will share common nodes. Cores will be shifted among applications as needed to optimise the usage of the resources.		
Planned Start Date:	September 2014	Planned Completion Date:	December 2015
Requires Functions:			

Task No:	7.2.3	Partner:	BSC (P4)
Function No:	7.2.3.6	Leader:	Jesus LABARTA
Function Name:	Interactive simulation demonstrator and evaluation		
Test Case A:	Demonstration and evaluation of the interactive session submission and resource management environment integrated in the UP and thoroughly evaluated with the microbenchmark developed in 7.2.3.4.		
Planned Start Date:	January 2014	Planned Completion Date:	March 2015
Requires Functions:	7.2.3.4		

T7.2.4: Tools for Performance Analysis and Prediction

T7.2.4 will develop application- and system-level tools to model application performance, interactions among applications, and system performance under resource constraints. The tools will extend existing performance analysis environments and address critical scalability issues. The application-level tools will consider both the sequential and the parallel levels. System performance models will focus on the interaction between workflow components, system resources and the persistent storage subsystem.

Task No:	7.2.4	Partners:	JUELICH (P17), BSC (P4)
Function No:	7.2.4.1	Leader:	Bernd MOHR
Function Name:	Installation of performance analysis tools		
Test Case A:	Performance analysis teams at JUELICH and BSC install their tools (Extrae/Paraver, Score-P, Scalasca, Cube) on the HPC Platform and test basic functionality		
Planned Start Date:	October 2013	Planned Completion Date:	February 2014
Requires Functions:			

Task No:	7.2.4	Partners:	JUELICH (P17), BSC (P4)
Function No:	7.2.4.2	Leader:	Bernd MOHR
Function Name:	Performance analysis of simulation codes		
Test Case A:	Code developers use installed performance tools to measure, analyse and optimise their simulation code		
Test Case B:	Code developers team up with performance analysis teams at JUELICH and BSC to extensively measure, analyse and optimise their simulation code		
Planned Start Date:	March 2014	Planned Completion Date:	March 2016
Requires Functions:	7.2.4.1		

Task No:	7.2.4	Partners:	JUELICH (P17), BSC (P4)
Function No:	7.2.4.3	Leader:	Bernd MOHR
Function Name:	Basic functional test of performance analysis tools installation		
Test Case A:	Code developers of NEST and NEURON are able to use installed performance tools to measure and analyse their simulation code on a restricted test case (minimal test input, restricted test execution on one rack of BG/Q) in cooperation with performance analysis teams at JUELICH and BSC.		
Planned Start Date:	March 2014	Planned Completion Date:	June 2014
Requires Functions:	7.2.4.1		

Task No:	7.2.4	Partners:	JUELICH (P17), BSC (P4)
Function No:	7.2.4.4	Leader:	Bernd MOHR
Function Name:	Gap analysis of performance analysis tools		
Test Case A:	Code developers (of the NEST and NEURON teams, but not limited to these) determine and document missing, non-effective, and broken functionality of performance tools, which keeps code developers from successfully analysing the performance of production-runs of their simulation codes on the HPC Platform. Performance analysis teams at JUELICH and BSC implement missing features, fix broken components, and optimise ineffective tool functions.		
Planned Start Date:	July 2014	Planned Completion Date:	June 2015
Requires Functions:	7.2.4.3		

Task No:	7.2.4	Partners:	JUELICH (P17), BSC (P4)
Function No:	7.2.4.5	Leader:	Bernd MOHR
Function Name:	Production functional test of performance analysis tools installation		
Test Case A:	Code developers of NEST and NEURON are able to use improved and enhanced performance tools to measure and analyse their simulation codes on production test case in cooperation with performance analysis teams at JUELICH and BSC.		
Planned Start Date:	July 2015	Planned Completion Date:	December 2015
Requires Functions:	7.2.4.4		

T7.2.5: Numerical Methods for Neuroscientific High Performance Computing

Multi-scale brain simulation requires numerous numerical algorithms; for instance, methods for solving systems of differential equations, for molecular dynamics simulations and for data analysis.

T7.2.5 will therefore develop scalable numerical methods, which application developers can use as building blocks for domain specific tools (e.g. tools for neuroinformatics, brain simulation tools and tools for the design of neuromorphic hardware). Typical examples include solvers for linear systems and multi-scale methods that make it possible to represent multiple temporal and spatial scales in a single simulation. The algorithms developed in this task will be designed for use in the interactive computing environments designed in WP7.3

Task No:	7.2.5	Partner:	BUW (P5)
Function No:	7.2.5.1	Leader:	Andreas FROMMER
Function Name:	Integration of coupled systems of ODEs in NEST		
Test Case A:	Within NEST, coupled systems of ODEs can be integrated on parallel computers where the different equations reside on different nodes		
Planned Start Date:	October 2013	Planned Completion Date:	July 2014
Requires Functions:			

Task No:	7.2.5	Partner:	BUW (P5)
Function No:	7.2.5.2	Leader:	Andreas FROMMER
Function Name:	Integration of coupled systems of ODEs on parallel computers		
Test Case A:	A user can integrate coupled systems of ODEs on parallel computers where the different equations reside on different nodes		
Planned Start Date:	August 2014	Planned Completion Date:	March 2016
Requires Functions:			

Task No:	7.2.5	Partner:	FG (P18)
Function No:	7.2.5.3	Leader:	Michael GRIEBEL
Function Name:	High dimensional approximation tools		
Test Case A:	A developer of high dimensional data analysis tools for neuroscience can use core modules e.g. for uncertainty quantification, data mining, Big Data analysis and parameter fitting of reduced order/coarse grained models.		
Planned Start Date:	October 2013	Planned Completion Date:	September 2014
Requires Functions:			

Task No:	7.2.5	Partner:	UFRA (P29)
Function No:	7.2.5.4	Leader:	Gabriel WITTUM
Function Name:	Simulation tool based on the three-dimensional cable equation: Discretisation		
Test Case A:	Deriving and implementing a finite volume/finite element discretisation for the three-dimensional cable equation		
Planned Start Date:	October 2013	Planned Completion Date:	September 2014
Requires Functions:			

Task No:	7.2.5	Partner:	UFRA (P29)	
Function No:	7.2.5.5	Leader:	Gabriel WITTUM	
Function Name:	Simulation tool based on the three-dimensional cable equation: Solver			
Test Case A:	Developing and implementing a fast multi-grid based solver for the discrete three-dimensional cable equation			
Planned Start Date:	October 2014	Planned Completion Date:		September 2015
Requires Functions:	7.2.5.4			

Task No:	7.2.5	Partner:	UFRA (P29)	
Function No:	7.2.5.6	Leader:	Gabriel WITTUM	
Function Name:	Simulation tool based on the three-dimensional cable equation: Parallelisation and interface			
Test Case A:	Signal transduction in neurons and small networks of neurons can be simulated using the three-dimensional cable equation			
Planned Start Date:	October 2015	Planned Completion Date:		March 2016
Requires Functions:	7.2.5.4 and 7.2.5.5			

Task No:	7.2.5	Partner:	FG (P18)	
Function No:	7.2.5.7	Leader:	Michael GRIEBEL	
Function Name:	Optimisation of simulation codes on the molecular scale			
Test Case A:	A user (e.g. from SP6) of simulation tools on the molecular scale is able to treat systems larger than the typical initial size at the beginning.			
Planned Start Date:	March 2015	Planned Completion Date:		March 2016
Requires Functions:				

Task No:	7.2.5	Partner:	UFRA (P29)	
Function No:	7.2.5.8	Leader:	Gabriel WITTUM	
Function Name:	Simulation tool based on the three-dimensional cable equation			
Test Case A:	Signal transduction in neurons and small networks of neurons can be simulated using the three-dimensional cable equation			
Planned Start Date:	October 2013	Planned Completion Date:		March 2016
Requires Functions:				

WP7.3: Interactive Visualisation, Analysis and Control

T7.3.1: Visualisation and Analysis Component Execution Framework

T7.3.1 will build on the experience of the Blue Brain Project to develop generic frameworks for resource discovery, allocation and scheduling; dynamic processing pipelines; and data staging. This work will prepare the way for production-quality software (to be released in the second phase of the project). The new frameworks will apply modern software engineering principles, supporting task-parallel and data-parallel execution as well as multi-threading with memory sharing. In this way, the frameworks will offer improved flexibility, ease of use and robustness, and a smaller memory footprint. The processing pipelines they provide will link HPC resources to workstations and cockpits, allowing researchers to interact with simulations in real time.

Task No:	7.3.1	Partner:	EPFL (P1)
Function No:	7.3.1.1	Leader:	Felix SCHÜRMANN
Function Name:	Simulation data streaming to visualisation		
Test Case A:	A user can launch a visualisation application, connect it to a running simulation, and visualise live, streamed simulation data.		
Planned Start Date:	October 2013	Planned Completion Date:	Mar 2014
Requires Functions:			

Task No:	7.3.1	Partner:	EPFL (P1)
Function No:	7.3.1.2	Leader:	Felix SCHÜRMANN
Function Name:	Simulation data steering		
Test Case A:	A user can influence simulation data parameters while streaming simulation data		
Planned Start Date:	Mar 2014	Planned Completion Date:	Sep 2014
Requires Functions:	7.3.1.1		

Task No:	7.3.1	Partner:	EPFL (P1)
Function No:	7.3.1.3	Leader:	Felix SCHÜRMANN
Function Name:	Process simulation data during streaming		
Test Case A:	A user can launch a dataflow-processing pipeline that connects a running simulation to a filter transforming the data for a visualisation application.		
Planned Start Date:	Sep 2014	Planned Completion Date:	Sep 2015
Requires Functions:	7.3.1.2		

T7.3.2: Neuroscience-Specific Visualisation and Interfaces

T7.3.2 will build on advanced software from the Spanish Cajal Blue Brain Project to address visualisation issues specific to neuroscience, in particular the need to accommodate a huge range of spatial and temporal scales, mixed data representations, and a high level of geometrical complexity. The work will cover data representations, visualisation algorithms and human-computer interfaces. Work on data representations will focus on symbolic and realistic visual representations of structural and functional data and on visualisation-specific data structures. An important challenge will be to mix different representations and scales in coherent visualisations. Humans rely heavily on visual and auditory feedback to rapidly make sense of complex environments, and the project aims to provide this feedback to guide in-depth analysis. T7.3.2 will develop initial building blocks for domain-specific visualisation techniques adapted to HBP requirements, and efficient methods to render results from very large multi-scale simulations accessible to the observer. Finally, T7.3.2 will work on human-computer interaction, developing user interfaces, interaction metaphors, and virtual laboratory instruments specially designed for use by neuroscientists. Development work will follow well-proven principles of user-centred design.

Task No:	7.3.2	Partner:	URJC (P60)
Function No:	7.3.2.1	Leader:	Luis PASTOR
Function Name:	Multi-scale representation		
Test Case A:	A user is able to select a relevant set of ontologies for interpreting data at different levels of abstraction.		
Test Case B:	A user is able to access different levels of abstraction for each selected ontology.		
Planned Start Date:	June 2014	Planned Completion Date:	July 2014
Requires Functions:			

Task No:	7.3.2	Partner:	URJC (P60)
Function No:	7.3.2.2	Leader:	Luis PASTOR
Function Name:	Navigation		
Test Case A:	A user can control the camera within a network of neurons, including camera position and orientation, and geometrical zoom.		
Test Case B:	A user can navigate using semantic zoom interacting with both realistic and abstract representations.		
Planned Start Date:	August 2014	Planned Completion Date:	October 2014
Requires Functions:	7.3.2.1		

Task No:	7.3.2	Partner:	URJC (P60)
Function No:	7.3.2.3	Leader:	Luis PASTOR
Function Name:	Search and Filtering		
Test Case A:	A user can perform efficient and interactive filtering of input data.		
Test Case B:	A user can start search tasks using data contents as well as metadata.		
Test Case C:	A user can visualise and interact with the obtained results in a user-guided iterative process on top of the search and filtering stages.		
Planned Start Date:	January 2015	Planned Completion Date:	March 2015
Requires Functions:			

T7.3.3: Hardware Technology, Benchmarking and Optimisation for Visualisation and Rendering towards the Exascale

T7.3.3 will provide the initial design for all visualisation-related components (software, graphics hardware, interaction devices, displays) that interact with HBP data, models and simulations. The task will adapt current ultra-scale visualisation paradigms to meet specific requirements for local and remote visualisation. The task will pay particular attention to bandwidth, required computational capabilities, input devices and output devices. The end result will be a unified hardware design ready for integration into the HBP Simulation Platform (see T6.5.1).

Task No:	7.3.3	Partner:	UPM (P59)
Function No:	7.3.3.1	Leader:	Vicente MARTIN
Function Name:	PCP Visualisation subsystem benchmarks		
Test Case A:	A bidder in the Pre-Commercial Procurement runs the benchmarks in their system to asses its performance compared to a base machine.		
Test Case B:	A Developer of the HBP HPC platform runs the benchmarks in a reference machine to set the base performance for the PCP.		
Planned Start Date:	October 2013	Planned Completion Date:	June 2014
Requires Functions:			

Task No:	7.3.3	Partner:	UPM (P59)
Function No:	7.3.3.2	Leader:	Vicente MARTIN
Function Name:	Optimisation of the visualisation tools (1)		
Test Case A:	A user of the visualisation tool is able to explore data sets with a size significantly larger than the initial size typical at the beginning. Target visualisation: Detailed neuron representations. Target tool: RTNeuron. Target performance: 20K neurons, as compared to the 10K typical of today.		
Planned Start Date:	June 2014	Planned Completion Date:	December 2014
Requires Functions:			

T7.3.4: Integrative Visualisation and Analysis Tools for the HBP Cockpits

T7.3.4 will bring together the methods and interfaces developed in T7.3.2 in early software prototypes ready for integration in the HBP Simulation Platform (see T6.5.1). The concepts will include a broad range of advanced functionality, including synchronised multi-view and session management techniques and techniques to combine geometrical and symbolic information.

Task No:	7.3.4	Partners:	RWTH (P42), EPFL (P1)
Function No:	7.3.4.1	Leader:	Torsten KUHLEN, Felix SCHÜRMANN
Function Name:	High resolution visualisation		
Test Case A:	A user is able to initialise a tiled display wall using the software package DisplayCluster.		
Test Case B:	A user is able to run a simple visualisation application on a high-resolution tiled display wall.		
Planned Start Date:	January 2014	Planned Completion Date:	March 2014
Requires Functions:			

Task No:	7.3.4	Partner:	RWTH (P42)
Function No:	7.3.4.2	Leader:	Torsten KUHLEN
Function Name:	Tracked interaction with high resolution displays		
Test Case A:	A user is able to interact with a visualisation application using infrared tracking.		
Test Case B:	A user is able to execute simple selecting operations on the high-resolution visualisation.		
Planned Start Date:	June 2014	Planned Completion Date:	July 2014
Requires Functions:	7.3.4.1		

Task No:	7.3.4	Partner:	RWTH (P42)
Function No:	7.3.4.3	Leader:	Torsten KUHLEN
Function Name:	Session Management		
Test Case A:	A user is able to track a visualisation process using Vtk and VisTrails.		
Test Case B:	A user is able to load a VisTrail and reproduce a previous applied visualisation process.		
Planned Start Date:	November 2014	Planned Completion Date:	December 2014
Requires Functions:			

Task No:	7.3.4	Partner:	RWTH (P42)		
Function No:	7.3.4.4	Leader:	Torsten KUHLEN		
Function Name:	First multi-view architecture				
Test Case A:	A user is able to execute a multi-view visualisation using data backend organising the view synchronisation.				
Test Case B:	The architecture is prototypical implemented using an event-driven architecture.				
Planned Start Date:	January 2015	Planned Completion Date:	February 2015		
Requires Functions:					

WP 7.4: Exascale Data Management

T7.4.1: Scalable Querying of Peta to Exascale Data Sets

T7.4.1 will develop novel spatial indexes, allowing scalable querying of massive neuroscience data sets located in external data centres, cloud systems and supercomputer storage. Indexes will be designed to support queries that occur frequently in neuroscience research, decoupling query execution time from the size and density of the data sets. Critically, indexes will be designed for use in memory systems, significantly speeding up access to data.

Task No:	7.4.1	Partner:	EPFL (P1)
Function No:	7.4.1.1	Leader:	Anastasia AILAMAKI
Function Name:	Scalable execution of range queries on spatial models		
Test Case A:	A user can efficiently identify objects (neurons, somas etc.) within a given range of the model. The query is efficiently executed using an index decoupling query execution from data set size.		
Planned Start Date:	October 2013	Planned Completion Date:	March 2014
Requires Functions:			

Task No:	7.4.1	Partner:	EPFL (P1)
Function No:	7.4.1.2	Leader:	Anastasia AILAMAKI
Function Name:	Structure-aware methods for executing spatial query sequences		
Test Case A:	To analyse models, a user can accelerate an interactive spatial query sequence using a structure-aware method for prefetching the spatial data required by the sequential queries.		
Planned Start Date:	April 2014	Planned Completion Date:	September 2014
Requires Functions:	7.4.1.1		

Task No:	7.4.1	Partner:	EPFL (P1)
Function No:	7.4.1.3	Leader:	Anastasia AILAMAKI
Function Name:	Scalable in-memory spatial join		
Test Case A:	To build models users can detect intersections/place synapses efficiently using an efficient spatial join method. The method performs the join scalable and efficiently in memory.		
Planned Start Date:	October 2014	Planned Completion Date:	March 2015
Requires Functions:			

Task No:	7.4.1	Partner:	EPFL (P1)
Function No:	7.4.1.4	Leader:	Anastasia AILAMAKI
Function Name:	Scalable execution of disk-based, large-scale spatial joins		
Test Case A:	Users can efficiently perform a large-scale spatial join of several data sets to find intersections during model building.		
Planned Start Date:	April 2015	Planned Completion Date:	March 2016
Requires Functions:	7.4.1.3		

Task No:	7.4.1	Partner:	EPFL (P1)
Function No:	7.4.1.5	Leader:	Anastasia AILAMAKI
Function Name:	Querying time series with spatio-temporal predicates		
Test Case A:	A user can efficiently query massive time series resulting from simulations for simple patterns. The queries are based on spatio-temporal predicates.		
Planned Start Date:	April 2015	Planned Completion Date:	March 2016
Requires Functions:	7.4.1.2		

T7.4.2: Exascale Data Analytics

T7.4.2 will specify requirements and develop a prototype for a system that can perform deep, real-time analytics on data from experiments, brain atlases, models and simulations, while conserving the data in its original, natural format. The full version of the system will be developed in the second phase of the project.

Task No:	7.4.2	Partner:	TUC (P51)	
Function No:	7.4.2.1	Leader:	Minos GAROFALAKIS	
Function Name:	Survey State-of-the-Art			
Test Case A:	Study existing, state-of-the-art time-series data mining algorithms and architectures.			
Planned Start Date:	January 2014	Planned Completion Date:		March 2014
Requires Functions:				

Task No:	7.4.2	Partner:	TUC (P51)	
Function No:	7.4.2.2	Leader:	Minos GAROFALAKIS	
Function Name:	Time-series data mining in STORM			
Test Case A:	Implementation of basic tools for massive streaming time-series data mining in the STORM parallel stream-processing architecture.			
Planned Start Date:	April 2014	Planned Completion Date:		June 2014
Requires Functions:	7.4.2.1			

Task No:	7.4.2	Partner:	TUC (P51)	
Function No:	7.4.2.3	Leader:	Minos GAROFALAKIS	
Function Name:	Horizontal and vertical scaling			
Test Case A:	Extending the algorithms and implementation with tools (e.g., synopses, indexing) that allow them to scale both in the volume/velocity of the data and the number of observed time-series.			
Planned Start Date:	July 2014	Planned Completion Date:		September 2014
Requires Functions:	7.4.2.2			

Task No:	7.4.2	Partner:	TUC (P51)	
Function No:	7.4.2.4	Leader:	Minos GAROFALAKIS	
Function Name:	Initial release of T-STORM			
Test Case A:	Initial release of the STORM-based data-mining engine for massive streaming time-series data.			
Planned Start Date:	October 2014	Planned Completion Date:		December 2014
Requires Functions:	7.4.2.3			

Task No:	7.4.2	Partner:	TUC (P51)
Function No:	7.4.2.5	Leader:	Minos GAROFALAKIS
Function Name:	Enhancing T-STORM functionality		
Test Case A:	Extend STORM-TS core data mining functionalities based on user and application requirements for HBP massive brain simulations.		
Planned Start Date:	January 2015	Planned Completion Date:	March 2015
Requires Functions:	7.4.2.4		

Task No:	7.4.2	Partner:	TUC (P51)
Function No:	7.4.2.6	Leader:	Minos GAROFALAKIS
Function Name:	External models/data sources in T-STORM		
Test Case A:	Integration of external data source (e.g., clinical studies, brain atlases) and models in the T-STORM architecture to allow real-time correlations with streaming brain-simulation time series.		
Planned Start Date:	April 2015	Planned Completion Date:	July 2015
Requires Functions:	7.4.2.5		

Task No:	7.4.2	Partner:	TUC (P51)
Function No:	7.4.2.7	Leader:	Minos GAROFALAKIS
Function Name:	Extending T-STORM functionality and HPC platform integration		
Test Case A:	Enhance system functionality and scaling based on HBP user and application requirements, and integrate within the HPC platform.		
Planned Start Date:	August 2015	Planned Completion Date:	December 2015
Requires Functions:	7.4.2.6		

T7.4.3: Data Provenance and Preservation

This Task starts in October 2014 (M13). Functions will be defined before the start of the Task.

T7.4.4: Array-Based Data Processing Models

T7.4.4 will extend current database technology with arrays, making it possible to perform statistical analyses inside the DBMS core and avoiding the need to export massive amounts of data to external statistical packages. The results will consist of new data management algorithms, supporting the HBP data analysis process.

Task No:	7.4.4	Partner:	CWI (P48)	
Function No:	7.4.4.1	Leader:	Martin KERSTEN	
Function Name:	Release version 0.9 of MonetDB/R			
Test Case A:	A user can exploit the functionality using a separate package at CRAN. More control is provided to interact with multiple database servers through R.			
Planned Start Date:	January 2014	Planned Completion Date:		March 2014
Requires Functions:				

Task No:	7.4.4	Partner:	CWI (P48)	
Function No:	7.4.4.2	Leader:	Martin KERSTEN	
Function Name:	BAMloader			
Test Case A:	A user can integrate scientific file formats in the database using a modern database loader, illustrated through provisioning of loading sizable gnome files.			
Planned Start Date:	April 2014	Planned Completion Date:		June 2014
Requires Functions:	7.4.4.1			

Task No:	7.4.4	Partner:	CWI (P48)	
Function No:	7.4.4.3	Leader:	Martin KERSTEN	
Function Name:	Release version 1.0 of MonetDB/R			
Test Case A:	The official release of MonetDB/R with callback functionality to provide a symbiotic execution of SQL queries in an R environment.			
Planned Start Date:	July 2014	Planned Completion Date:		September 2014
Requires Functions:	7.4.4.2			

Task No:	7.4.4	Partner:	CWI (P48)	
Function No:	7.4.4.4	Leader:	Martin KERSTEN	
Function Name:	Alpha 0.2 release of MonetDB/Python			
Test Case A:	A user can define Python code fragments as in-lined UDFs, which provides direct access to the variety of simulation packages written in Python and a fast data-exchange interconnect between DBMS kernel and Python interpreter.			
Planned Start Date:	October 2014	Planned Completion Date:		March 2015
Requires Functions:	7.4.4.3			

Task No:	7.4.4	Partner:	CWI (P48)
Function No:	7.4.4.5	Leader:	Martin KERSTEN
Function Name:	Beta release of MonetDB/SciQL		
Test Case A:	A user can experiment with declarative specification of computational models and rely on the system to optimise and translate it into executional code using the various language bindings.		
Planned Start Date:	April 2015	Planned Completion Date:	June 2015
Requires Functions:	7.4.4.4		

Task No:	7.4.4	Partner:	CWI (P48)
Function No:	7.4.4.6	Leader:	Martin KERSTEN
Function Name:	Alpha 0.2 release of MonetDB/Javascript		
Test Case A:	The user can embed his NoSQL and web-based applications with a fast track path into the database system. This provides a stepping-stone for fusion of structured data and non-structured data.		
Planned Start Date:	July 2015	Planned Completion Date:	September 2015
Requires Functions:	7.4.4.5		

Task No:	7.4.4	Partner:	CWI (P48)
Function No:	7.4.4.7	Leader:	Martin KERSTEN
Function Name:	Packaging of in-lined language bindings		
Test Case A:	Users can freely use snippets written in their favourite programming language (Python, R, JavaScript) to enhance data processing within the kernel of the DBMS, which can take care of data-parallelism decisions.		
Planned Start Date:	October 2015	Planned Completion Date:	December 2015
Requires Functions:	7.4.4.6		

T7.4.5: Data Platform Dissemination and Integration

T7.4.5 is responsible for deploying the algorithms and methods developed in T7.4.1-T7.4.4 across the HBP. The task team will bring the research prototypes produced in WP7.4 to product-grade quality, and will include personnel to resolve bugs, manage installations and provide technical support. Given that algorithms and methods have to be developed before they can be deployed, this task will start slowly, ramping up as new results become available.

Task No:	7.4.5	Partner:	EPFL	
Function No:	7.4.5.1	Leader:	Anastasia AILAMAKI	
Function Name:	Deployment and release of software for range queries on spatial models			
Test Case A:	The index decoupling data model size and query execution time is deployed and released, making the functionality available to users.			
Planned Start Date:	April 2014	Planned Completion Date:		September 2014
Requires Functions:				

Task No:	7.4.5	Partner:	EPFL	
Function No:	7.4.5.2	Leader:	Anastasia AILAMAKI	
Function Name:	Deployment and release of structure-aware methods for the efficient execution of spatial query sequences			
Test Case A:	The data structures and algorithms needed for the efficient prefetching of spatial data are deployed and released to users.			
Planned Start Date:	October 2014	Planned Completion Date:		March 2015
Requires Functions:	7.4.5.1			

Task No:	7.4.5	Partner:	EPFL	
Function No:	7.4.5.3	Leader:	Anastasia AILAMAKI	
Function Name:	Deployment and release of scalable in-memory spatial join			
Test Case A:	The algorithm for the scalable in-memory spatial join are re-implemented and released to the users of the platform.			
Planned Start Date:	April 2015	Planned Completion Date:		March 2016
Requires Functions:				

Task No:	7.4.5	Partner:	EPFL	
Function No:	7.4.5.4	Leader:	Anastasia AILAMAKI	
Function Name:	Deployment and release of algorithm for disk-based, large-scale spatial joins			
Test Case A:	Algorithms for the efficient large-scale spatial join of several data sets are deployed and released to the users of the platform.			
Planned Start Date:	April 2015	Planned Completion Date:		March 2016
Requires Functions:	7.4.5.3			

Task No:	7.4.5	Partner:	EPFL
Function No:	7.4.5.5	Leader:	Anastasia AILAMAKI
Function Name:	Deployment of algorithms for querying time series with spatio-temporal predicates		
Test Case A:	Re-implementation and deployment of methods for the efficient querying of time-series resulting from simulations.		
Planned Start Date:	April 2015	Planned Completion Date:	March 2016
Requires Functions:	7.4.5.1		

WP 7.5: HPC Platform: Integration and Operations

T7.5.1: The HBP Supercomputer for Brain Modelling and Simulation

T7.5.1 will operate the supercomputer systems used for "production runs" of HBP simulations. During the ramp-up phase, these will consist of existing Jülich supercomputers. The first step in the work will be to analyse resource and software requirements. T7.5.1 will then implement and configure the necessary software and resource management tools, establishing a valid use model. After all technical requirements are implemented and the Access Committee (see Section 2.9.1) has assigned projects, users will be able to run HBP simulation jobs on the Jülich supercomputers. In parallel with this work, members of T 7.5.1 will contribute to discussions concerning the specification of the hardware and architectures for later stages of the project.

Task No:	7.5.1	Partner:	JUELICH (P17)
Function No:	7.5.1.1	Leader:	Klaus WOLKERSDORFER
Function Name:	Integration of the HBP Supercomputer into the Single Sign-On (SP7-FR-002)		
Test Case A:	Any HBP user authorised to access the HBP Supercomputer will be able to use their unique HBP credentials to log into the system.		
Planned Start Date:	June 2014	Planned Completion Date:	October 2014
Requires Functions:	7.5.6.5		

Task No:	7.5.1	Partner:	JUELICH (P17)
Function No:	7.5.1.2	Leader:	Klaus WOLKERSDORFER
Function Name:	Integration of the HBP Supercomputer within the PRACE network (see Section 2.6.1.6)		
Test Case A:	A physical network interface allowing the HBP Supercomputer to connect to the PRACE network is active		
Planned Start Date:	April 2014	Planned Completion Date:	July 2014
Requires Functions:	7.5.6.1		

Task No:	7.5.1	Partner:	JUELICH (P17)	
Function No:	7.5.1.3	Leader:	Klaus WOLKERSDORFER	
Function Name:	Usage of the HPC Platform to submit workloads to the HBP supercomputer (SP7-FR-003)			
Test Case A:	HPC jobs can be submitted via a HBP specific UNICORE configuration and executed on the HBP Supercomputer. Data produced can be stored locally.			
Planned Start Date:	July 2014	Planned Completion Date:		December 2014
Requires Functions:				

T7.5.2: The HBP Development System

T7.5.2 will make available existing supercomputing resources dedicated to brain simulation and hosted at CSCS. These resources, based on current commercial technology, will be dedicated to the testing and development of brain simulation codes performed by other HBP subprojects. Findings will contribute to the specification of the HBP Supercomputer (see T7.5.1) and future versions of the HBP Development System at CSCS and elsewhere.

Task No:	7.5.2	Partner:	ETHZ/CSCS (P15)	
Function No:	7.5.2.1	Leader:	Colin McMURTRIE	
Function Name:	Integration of the HBP Development System into the Single Sign-On (SP7-FR-002)			
Test Case A:	Any HBP user authorised to access the HBP Development System will be able to use their unique HBP credentials to log into the system.			
Planned Start Date:	June 2014	Planned Completion Date:		October 2014
Requires Functions:	7.5.6.5			

Task No:	7.5.2	Partner:	ETZH/CSCS (P15)	
Function No:	7.5.2.2	Leader:	Colin McMURTRIE	
Function Name:	Integration of the HBP Development System within the PRACE network (see Section 2.6.1.6)			
Test Case A:	A physical network interface allowing the HBP Development System to connect to the PRACE network is active			
Planned Start Date:	April 2014	Planned Completion Date:		October 2014
Requires Functions:	7.5.6.1			

Task No:	7.5.2	Partner:	CSCS	
Function No:	7.5.2.3	Leader:	Colin McMURTRIE	
Function Name:	Usage of the HPC Platform to submit workloads to the HBP Development System (SP7-FR-003)			
Test Case A:	HPC jobs can be submitted via a HBP specific UNICORE configuration and executed on the HBP Development System. Data produced can be stored locally.			
Planned Start Date:	July 2014	Planned Completion Date:		December 2014
Requires Functions:				

T7.5.3: The HBP Supercomputer for Molecular Dynamics

T7.5.3 will begin the design of a supercomputing facility for molecular dynamics simulations. This work, which will be completed in the second phase of the project, will involve the identification of user requirements, the specification of performance benchmarks, and the procurement, installation, validation and operation of the system.

Task No:	7.5.3	Partner:	BSC (P4)	
Function No:	7.5.3.1	Leader:	Javier BARTOLOME	
Function Name:	Integration of the HBP Supercomputer for Molecular Dynamics into the Single Sign-On (SP7-FR-002)			
Test Case A:	Any HBP user authorised to access the HBP Supercomputer for Molecular Dynamics will be able to use their unique HBP credentials to log into the system.			
Planned Start Date:	June 2014	Planned Completion Date:		October 2014
Requires Functions:	7.5.6.5			

Task No:	7.5.3	Partner:	BSC (P4)	
Function No:	7.5.3.2	Leader:	Javier BARTOLOME	
Function Name:	Integration of the HBP Supercomputer for Molecular Dynamics into HBP within the PRACE network (see Section 2.6.1.6)			
Test Case A:	A physical network interface allowing the HBP Supercomputer for Molecular Dynamics to connect to the PRACE network is active			
Planned Start Date:	April 2014	Planned Completion Date:		July 2014
Requires Functions:	7.5.6.1			

Task No:	7.5.3	Partner:	BSC (P4)	
Function No:	7.5.3.3	Leader:	Javier BARTOLOME	
Function Name:	Usage of the HPC Platform to submit workloads to the HBP Supercomputer for Molecular Dynamics (SP7-FR-003)			
Test Case A:	HPC jobs can be submitted via a HBP specific Unicore configuration and executed on the HBP Supercomputer for Molecular Dynamics. Data produced can be stored locally.			
Planned Start Date:	July 2014	Planned Completion Date:		December 2014
Requires Functions:				

T7.5.4: The HBP Supercomputer for Massive Data Analytics

T7.5.4 will implement and operate a data-centric HPC facility providing efficient storage, processing and management of large volumes of data generated by the HBP. During the ramp-up phase of the project, T7.5.4 will use existing CINECA HPC and data storage resources. The task will begin by collecting and analysing HBP requirements, installing software tools and providing services based on *ad hoc* service policies. It will then go on to finalise the specification for the data analytics platform to be developed in the second phase of the project.

Task No:	7.5.4	Partner:	CINECA (P10)	
Function No:	7.5.4.1	Leader:	Roberto MUCCI	
Function Name:	Integration of the HBP Supercomputer for Massive Data Analytics into the Single Sign-On (SP7-FR-002)			
Test Case A:	Any HBP user authorised to access the HBP Supercomputer for Massive Data Analytics will be able to use their unique HBP credentials to log into the system.			
Planned Start Date:	June 2014	Planned Completion Date:		October 2014
Requires Functions:	7.5.6.5			

Task No:	7.5.4	Partner:	CINECA (P10)	
Function No:	7.5.4.2	Leader:	Roberto MUCCI	
Function Name:	Integration of the HBP Supercomputer for Massive Data Analytics into HBP within the PRACE network (see Section 2.6.1.6)			
Test Case A:	A physical network interface allowing the HBP Supercomputer for Massive Data Analytics to connect to the PRACE network is active			
Planned Start Date:	June 2014	Planned Completion Date:		October 2014
Requires Functions:	7.5.6.1			

Task No:	7.5.4	Partner:	CINECA (P10)	
Function No:	7.5.4.3	Leader:	Roberto MUCCI	
Function Name:	Usage of the HPC Platform to submit workloads to the HBP Supercomputer for Massive Data Analytics (SP7-FR-003)			
Test Case A:	HPC jobs can be submitted via a HBP specific UNICORE configuration and executed on the HBP Supercomputer for Massive Data Analytics. Data produced can be stored locally.			
Planned Start Date:	June 2014	Planned Completion Date:		October 2014
Requires Functions:				

Task No:	7.5.4	Partner:	CINECA (P10)	
Function No:	7.5.4.4	Leader:	Roberto MUCCI	
Function Name:	Visualisation and post-processing			
Test Case A:	A user can locally and remotely perform visualisation and post-processing activities on the data stored on the Cineca data repository.			
Planned Start Date:	October 2014	Planned Completion Date:		March 2015
Requires Functions:	7.5.4.3			

T7.5.5: HBP Cloud Services

T7.5.5 will design technological solutions that will provide HBP platform users with easy, secure and transparent access to cloud-based storage and computing resources. The task will set up interface abstractions making it possible to mix and match resources from academic and commercial providers. This task will be carried out in close collaboration with other tasks intending to use high-level cloud services (T7.4.1, T7.4.2, T8.1.1).

Task No:	7.5.5	Partner:	KIT (P30)	
Function No:	7.5.5.1	Leaders:	Diana GUDU, Marcus HARDT	
Function Name:	Basic S3 Storage service			
Test Case A:	A user can create buckets and store and retrieve objects using the Cloud storage resources at KIT in an Amazon compatible way. Client software can either be external or may be included in future UNICORE releases.			
Comment	Since this is a test deployment, limited storage space will be available.			
Planned Start Date:	December 2013	Planned Completion Date:		May 2014
Requires Functions:				

Task No:	7.5.5	Partner:	KIT (P30)
Function No:	7.5.5.2	Leaders:	Diana GUDU, Marcus HARDT
Function Name:	S3 Storage service		
Test Case A:	A user can create buckets and store and retrieve objects using the Cloud storage resources at KIT in an Amazon compatible way. Client software can either be external or may be included in future UNICORE releases.		
Comment	Same use-case, but larger amount of storage as well as regular backups of stored data will be performed.		
Planned Start Date:	June 2014	Planned Completion Date:	September 2014
Requires Functions:	7.5.5.1		

Task No:	7.5.5	Partner:	KIT (P30)
Function No:	7.5.5.3	Leader:	Diana GUDU, Marcus HARDT
Function Name:	KIT Cloud Storage user authentication		
Test Case A:	Any HBP user authorised to access KIT Cloud Storage will be able to access their storage using the common HBP credentials or an alternative method.		
Planned Start Date:	June 2014	Planned Completion Date:	October 2014
Requires Functions:			

T7.5.6: Supporting Infrastructure - Networking, Storage and Monitoring

T7.5.6 will develop and operate a dedicated HPC network infrastructure and a parallel HBP file system that will provide archiving facilities, monitoring, and secure access to HPC resources. Work in T7.5.6 will include the development of a framework for managing cross-site operational IT security, an Acceptable Use Policy (AUP), and information exchange and trust building among partners. It will also include the development of an interface providing engineers and users with information on the current status and compute performance, storage and network resources.

Task No:	7.5.6	Partner:	JUELICH (P17)
Function No:	7.5.6.1	Leader:	Ralph NIEDERBERGER
Function Name:	MoU with PRACE for network usage		
Test Case A:	A Memorandum of Understanding (MoU) between the HBP and PRACE allows HBP users to use the dedicated PRACE high-speed network to transfer large amounts of data between the different HBP HPC systems and data storage sites.		
Planned Start Date:	March 2014	Planned Completion Date:	April 2014
Requires Functions:			

Task No:	7.5.6	Partners:	JUELICH (P17), ETHZ/CSCS (P15), EPFL (P1)	
Function No:	7.5.6.2	Leader:	Ralph NIEDERBERGER	
Function Name:	Integration of CSCS and EPFL into the PRACE dedicated network			
Test Case A:	After negotiations with GÉANT and SWITCH, the Swiss NREN, and after the technical realisation of dedicated links, CSCS and EPFL can use the PRACE high-speed network in the same way as their HPC Platform partners BSC, CINECA and JUELICH.			
Planned Start Date:	March 2014	Planned Completion Date:		July 2014
Requires Functions:	7.5.6.1, 7.5.2.2			

Task No:	7.5.6	Partners:	JUELICH (P17), KIT (P30)	
Function No:	7.5.6.3	Leader:	Ralph NIEDERBERGER	
Function Name:	Integration of Cloud storage system at KIT into the network infrastructure			
Test Case A:	The Cloud storage facility provided by KIT is integrated into the network infrastructure.			
Planned Start Date:	May 2014	Planned Completion Date:		July 2014
Requires Functions:				

Task No:	7.5.6	Partners:	BSC (P4), CINECA (P10), ETHZ/CSCS (P15), EPFL (P1), JUELICH (P17)	
Function No:	7.5.6.4	Leader:	Ralph NIEDERBERGER	
Function Name:	HBP storage infrastructure			
Test Case A:	A distributed or shared storage infrastructure for the HBP is designed and implemented. This follows a first step, in which explicit data transfers via UNICORE/GridFTP have been used between the existing local installations at all sites. In a second step, alternative technologies such as AFM/GPFS enabling seamless data exchange between the various HPC sites are evaluated and considered for implementation.			
Planned Start Date:	July 2014	Planned Completion Date:		December 2014
Requires Functions:	7.5.6.2			

Task No:	7.5.6	Partners:	BSC (P4), CINECA (P10), ETHZ/CSCS (P15), EPFL (P1), JUELICH (P17)
Function No:	7.5.6.5	Leader:	Ralph NIEDERBERGER
Function Name:	HBP AAI infrastructure		
Test Case A:	HBP users are able to access their allocated HBP resources in a simple, fast and secure manner. A User Administration System (UAS) built on top of existing local user administration systems at the HBP sites provides the necessary basis for an LDAP based exchange of AAI user information. Users registered at one HBP site are made known to the other partners via this UAS system.		
Planned Start Date:	April 2014	Planned Completion Date:	July 2014
Requires Functions:			

Task No:	7.5.6	Partners:	BSC (P4), CINECA (P10), ETHZ/CSCS (P15), EPFL (P1), JUELICH (P17)
Function No:	7.5.6.6	Leader:	Ralph NIEDERBERGER
Function Name:	HBP CERT/CSIRT		
Test Case A:	An HBP CERT/CSIRT group composed of staff of the local HBP sites CERT members is established. Security incident information is shared among this group on a need to know basis and action will be taken together if needed.		
Planned Start Date:	May 2014	Planned Completion Date:	June 2014
Requires Functions:	7.5.6.2		

Task No:	7.5.6	Partners:	BSC (P4), CINECA (P10), ETHZ/CSCS (P15), EPFL (P1), JUELICH (P17)
Function No:	7.5.6.7	Leader:	Ralph NIEDERBERGER
Function Name:	HBP Security Policy		
Test Case A:	A common HBP IT security policy is designed to create a net of trust, so that each site is aware of the policies of the other sites and can handle/provide services on the basis of these policies.		
Planned Start Date:	May 2014	Planned Completion Date:	August 2014
Requires Functions:			

Task No:	7.5.6	Partners:	BSC (P4), CINECA (P10), ETHZ/CSCS (P15), EPFL (P1), JUELICH (P17)
Function No:	7.5.6.8	Leader:	Ralph NIEDERBERGER
Function Name:	HBP AUP / ToU		
Test Case A:	An Acceptable Use Policy (AUP) and Terms of Use (ToU) are defined, so that each HBP user knows the rules under which HBP resources are provided.		
Planned Start Date:	May 2014	Planned Completion Date:	August 2014
Requires Functions:			

Task No:	7.5.6	Partners:	BSC (P4), CINECA (P10), ETHZ/CSCS (P15), EPFL (P1), JUELICH (P17)
Function No:	7.5.6.9	Leader:	Ralph NIEDERBERGER
Function Name:	HBP Security Forum		
Test Case A:	Installing services within the HBP infrastructure may violate the HBP IT security policy. Each time a new service is planned to be installed, it has to be checked if it is compliant with this policy. A team of security experts is created which will be responsible for evaluating those services.		
Planned Start Date:	March 2014	Planned Completion Date:	September 2014
Requires Functions:			

T7.5.7: HPC Platform Website Construction and Maintenance

T7.5.7 will design, build and operate an integrated internet-accessible portal as a single point of access to the distributed supercomputer, storage and network services of the High Performance Computing Platform, operated by T7.5.1 – T7.5.6. T7.5.7 will use the monitoring interface developed in T7.5.6 to continuously inform Platform users of the status, usage and performance of different Platform components.

The first internal release of the portal will take place in month 18, when the HPC platform is fully opened to the Consortium. The first full version, open for use by external researchers, will become available in month 30. From this time onwards (in phase 2 of the project), T7.5.7 will operate and maintain the Platform for the benefit of the community, ensuring that it is accessible and continuously upgrading its capabilities.

Task No:	7.5.7	Partner:	JUELICH (P17)
Function No:	7.5.7.1	Leader:	Daniel MALLMANN
Function Name:	First version of the HBP HPC Platform Website		
Test Case A:	A Portal User accesses HPC resources through the UP or the HPC Platform Website in order to submit a job, monitor the job status (queued, running, failed, successfully finished, etc.), access results of her/his previously submitted jobs, access and manage data stored on a resource of the HPC Platform.		
Planned Start Date:	April 2014	Planned Completion Date:	September 2014
Requires Functions:			

Task No:	7.5.7	Partner:	JUELICH (P17)
Function No:	7.5.7.2	Leader:	Daniel MALLMANN
Function Name:	Launch of certified and registered applications on different HPC resources		
Test Case A:	A Portal User can select applications that are registered in the UNICORE IDB and start them on the resources where they are installed through the UP or the HPC Platform Website. In absence of certified HBP applications, a demo test case will be used to demonstrate the functionality.		
Planned Start Date:	September 2014	Planned Completion Date:	October 2014
Requires Functions:	7.5.7.1		

Task No:	7.5.7	Partner:	JUELICH (P17)
Function No:	7.5.7.3	Leader:	Daniel MALLMANN
Function Name:	Upload and download to and from HPC resources and HBP data repositories		
Test Case A:	A user can upload and download data to and from the HPC Platform through the UP		
Planned Start Date:	October 2014	Planned Completion Date:	November 2014
Requires Functions:	7.5.7.1, 7.5.6.4		

Task No:	7.5.7	Partner:	JUELICH (P17)
Function No:	7.5.7.4	Leader:	Bernd SCHULLER
Function Name:	Integration of all HPC systems of the participating sites (JUELICH, CSCS, BSC and CINECA) with the HPC Platform		
Test Case A:	The HPC systems can be accessed and used by a HBP user through the HPC Platform		
Planned Start Date:	March 2014	Planned Completion Date:	January 2015
Requires Functions:	7.5.6.2, 7.5.6.4, 7.5.6.5		

Task No:	7.5.7	Partner:	JUELICH (P17)
Function No:	7.5.7.5	Leader:	Bernd SCHULLER
Function Name:	RESTful API to HPC Platform		
Test Case A:	The user can access the fundamental HPC Platform functions (e.g., job submission and management, data access and movement) from a non-interactive JSON REST client authenticated using OpenID-Connect		
Planned Start Date:	May 2014	Planned Completion Date:	March 2015
Requires Functions:	7.5.7.1		

WP7.6: HPC Platform: User Support and Community Building

T7.6.1: HPC Platform Documentation and User Training

The goal of T7.6.1 is to assist and train users with different backgrounds and experience to make effective use of the High Performance Computing Platform. T7.6.1 will implement and manage technical user documentation for the Platform and devise and implement a training programme. To this end, it will establish cross-site, cross-service documentation standards, implement a common document management system, produce training materials, and organise training events for different user groups.

Task No:	7.6.1	Partner:	JUELICH (P17)
Function No:	7.6.1.1	Leader:	Boris ORTH
Function Name:	Technical framework and guidelines for documentation		
Test Case A:	A technical framework and general guidelines for documentation are defined in agreement with HBP-wide standards. Prospective HPC Platform user groups are classified according to their background and experience.		
Planned Start Date:	October 2013	Planned Completion Date:	March 2014
Requires Functions:			

Task No:	7.6.1	Partners:	JUELICH (P17), SP7
Function No:	7.6.1.2	Leader:	Boris ORTH
Function Name:	HPC Platform Specification		
Test Case A:	This document provides a detailed specification of the HPC Platform, including overall goals, use cases, functional and non-functional requirements, software, physical architecture, prerequisites and parallel activities.		
Planned Start Date:	January 2014	Planned Completion Date:	March 2014
Requires Functions:	7.6.1.1		

Task No:	7.6.1	Partners:	JUELICH (P17), SP7
Function No:	7.6.1.3	Leader:	Boris ORTH
Function Name:	Guidebook v0.5		
Test Case A:	This version of the HPC Platform Guidebook comes with the HBP-internal release of the HPC Platform. It contains technical user documentation provided by the different tasks and groups of SP7, based on the documentation guidelines.		
Planned Start Date:	April 2014	Planned Completion Date:	March 2015
Requires Functions:	7.6.1.1, 7.6.1.2		

Task No:	7.6.1	Partners:	JUELICH (P17), SP7
Function No:	7.6.1.4	Leader:	Boris ORTH
Function Name:	Guidebook v1.0		
Test Case A:	This version of the HPC Platform Guidebook comes with the official release of the HPC Platform. It contains updated and improved technical user documentation provided by the different tasks and groups of SP7, based on the documentation guidelines and taking into account the different user groups identified in 7.6.1.1.		
Planned Start Date:	April 2015	Planned Completion Date:	March 2016
Requires Functions:	7.6.1.1, 7.6.1.2, 7.6.1.4		

Task No:	7.6.1	Partner:	JUELICH (P17)
Function No:	7.6.1.5	Leader:	Boris ORTH
Function Name:	Overview of existing training courses		
Test Case A:	All relevant trainings already offered by the HPC Platform members are collected and published in a single place. Links to available training material are provided.		
Planned Start Date:	April 2014	Planned Completion Date:	June 2014
Requires Functions:			

Task No:	7.6.1	Partner:	JUELICH (P17)
Function No:	7.6.1.6	Leader:	Boris ORTH
Function Name:	Evaluation of training requirements		
Test Case A:	Services and components of the HPC Platform are identified for which appropriate training courses and/or documents are missing.		
Planned Start Date:	July 2014	Planned Completion Date:	September 2014
Requires Functions:	7.6.1.2		

Task No:	7.6.1	Partners:	JUELICH (P17), SP7, other platforms
Function No:	7.6.1.7	Leader:	Boris ORTH
Function Name:	Design of training material and events		
Test Case A:	Based on the requirements, concepts for training material and events are designed in collaboration with the members of the HPC Platform and its users.		
Planned Start Date:	October 2014	Planned Completion Date:	December 2014
Requires Functions:	7.6.1.6, 7.6.1.2		

Task No:	7.6.1	Partner:	JUELICH (P17), SP7, other platforms
Function No:	7.6.1.8	Leader:	Boris ORTH
Function Name:	Production and publication of training material		
Test Case A:	Training material is produced and kept up to date in collaboration with the members of the HPC Platform and its users. The material or links to the material are made available via a single access point.		
Planned Start Date:	January 2015	Planned Completion Date:	March 2016
Requires Functions:	7.6.1.7		

Task No:	7.6.1	Partner:	JUELICH (P17), SP7
Function No:	7.6.1.9	Leader:	Boris ORTH
Function Name:	Organisation of a training event		
Test Case A:	A training event for a specific user group is organised.		
Planned Start Date:	January 2015	Planned Completion Date:	March 2016
Requires Functions:	7.6.1.7, 7.6.1.8		

T7.6.2: Coordination with PRACE and other RIs

See **Error! Reference source not found.** for the KPI for this Coordination Task.

WP7.7: HPC Platform: Scientific Coordination

T7.7.1: Scientific coordination and support

See **Error! Reference source not found.** for the KPI for this Coordination Task.

Table 9**Error! Reference source not found.** below lists for each function the requirements it will address.

	SP7-FR-001	SP7-FR-002	SP7-FR-003	SP7-FR-004	SP7-FR-005	SP7-FR-006	SP7-FR-007	SP7-NFR-001	SP7-NFR-002	SP7-NFR-003	SP7-NFR-004	SP7-NFR-005	SP7-NFR-006	SP7-NFR-007
7.2.1.1	■				■									■
7.2.1.2	■				■									■
7.2.1.3	■				■									■
7.2.1.4	■		■		■									■
7.2.1.5	■				■									■
7.2.1.6	■				■									■
7.2.2.1	■				■									■
7.2.2.2	■		■		■				■					■
7.2.2.3	■		■		■				■					■
7.2.3.1	■				■									■
7.2.3.2	■				■									■
7.2.3.3	■		■											■
7.2.3.4	■		■											■
7.2.3.5	■													■
7.2.3.6	■		■											■
7.2.4.1					■									■
7.2.4.2					■									■
7.2.4.3					■									■
7.2.4.4					■									■
7.2.4.5					■									■
7.2.5.1	■				■									■
7.2.5.2	■				■									■
7.2.5.3	■				■	■	■							■
7.2.5.4	■													■
7.2.5.5	■													■

	SP7-FR-001	SP7-FR-002	SP7-FR-003	SP7-FR-004	SP7-FR-005	SP7-FR-006	SP7-FR-007	SP7-NFR-001	SP7-NFR-002	SP7-NFR-003	SP7-NFR-004	SP7-NFR-005	SP7-NFR-006	SP7-NFR-007
7.2.5.6	■													■
7.2.5.7	■													■
7.2.5.8	■													■
7.3.1.1	■													■
7.3.1.2	■													■
7.3.1.3	■						■							■
7.3.2.1	■						■							■
7.3.2.2	■						■							■
7.3.2.3	■						■							■
7.3.3.1	■				■									■
7.3.3.2	■													■
7.3.4.1	■													■
7.3.4.2	■													■
7.3.4.3	■			■										■
7.3.4.4	■													■
7.4.1.1							■							■
7.4.1.2							■							■
7.4.1.3							■							■
7.4.1.4							■							■
7.4.1.5						■	■							■
7.4.2.1	■						■							■
7.4.2.2	■						■							■
7.4.2.3	■						■							■
7.4.2.4	■						■							■
7.4.2.5	■						■							■

	SP7-FR-001	SP7-FR-002	SP7-FR-003	SP7-FR-004	SP7-FR-005	SP7-FR-006	SP7-FR-007	SP7-NFR-001	SP7-NFR-002	SP7-NFR-003	SP7-NFR-004	SP7-NFR-005	SP7-NFR-006	SP7-NFR-007
7.4.2.6	■						■							■
7.4.2.7	■			■			■							■
7.4.4.1							■							■
7.4.4.2							■							■
7.4.4.3							■							■
7.4.4.4							■							■
7.4.4.5							■							■
7.4.4.6							■							■
7.4.4.7							■							■
7.4.5.1							■							■
7.4.5.2							■							■
7.4.5.3							■							■
7.4.5.4							■							■
7.4.5.5							■							■
7.5.1.1		■												■
7.5.1.2										■	■			■
7.5.1.3	■		■		■						■	■	■	■
7.5.2.1		■												■
7.5.2.2										■	■			■
7.5.2.3	■		■		■						■	■	■	■
7.5.3.1		■												■
7.5.3.2										■	■			■
7.5.3.3	■		■		■						■	■	■	■
7.5.4.1		■												■
7.5.4.2										■	■			■

	SP7-FR-001	SP7-FR-002	SP7-FR-003	SP7-FR-004	SP7-FR-005	SP7-FR-006	SP7-FR-007	SP7-NFR-001	SP7-NFR-002	SP7-NFR-003	SP7-NFR-004	SP7-NFR-005	SP7-NFR-006	SP7-NFR-007
7.5.4.3	■		■		■						■	■	■	■
7.5.4.4	■													■
7.5.5.1						■								■
7.5.5.2						■								■
7.5.5.3		■									■			■
7.5.6.1										■				■
7.5.6.2										■				■
7.5.6.3										■				■
7.5.6.4						■		■	■					■
7.5.6.5		■												■
7.5.6.6													■	■
7.5.6.7													■	■
7.5.6.8													■	■
7.5.6.9													■	■
7.5.7.1														■
7.5.7.2				■					■					■
7.5.7.3						■			■	■				■
7.5.7.4	■	■	■	■		■			■	■				■
7.5.7.5	■	■	■	■		■								■
7.6.1.1														■
7.6.1.2														■
7.6.1.3														■
7.6.1.4														■
7.6.1.5														■
7.6.1.6														■

	SP7-FR-001	SP7-FR-002	SP7-FR-003	SP7-FR-004	SP7-FR-005	SP7-FR-006	SP7-FR-007	SP7-NFR-001	SP7-NFR-002	SP7-NFR-003	SP7-NFR-004	SP7-NFR-005	SP7-NFR-006	SP7-NFR-007
7.6.1.7														
7.6.1.8														
7.6.1.9														

Table 9: Relationship between Functions and Requirements

Task	KPI	ID	Target values
T7.1.1-3	Percentage of PCP budget spent	SP7-SKPI-001	M15: 8%; M21: 23%; M36: 100%
T7.2.1	Number of completed Functions	SP7-SKPI-002	M5: 1; M8: 3; M11: 4; M16: 5; M18: 6
T7.2.2	Number of completed Functions	SP7-SKPI-003	M6: 1; M12: 2; M18: 3
T7.2.3	Number of completed Functions	SP7-SKPI-004	M6: 1; M9: 3; M13: 4; M18: 5; M27: 6
T7.2.4	Number of completed Functions	SP7-SKPI-005	M5: 1; M9: 2; M21: 3; M27: 4; M30: 5
T7.2.5	Number of completed Functions	SP7-SKPI-006	M10: 1; M12: 3; M24: 4; M30: 8
T7.3.1	Number of completed Functions	SP7-SKPI-007	M6: 1; M12: 2; M24: 3
T7.3.2	Number of completed Functions	SP7-SKPI-008	M10: 1; M13: 2; M18: 3
T7.3.3	Number of completed Functions	SP7-SKPI-009	M9: 1; M15: 2
T7.3.4	Number of completed Functions	SP7-SKPI-010	M6: 1; M10: 2; M15: 3; M17: 4
T7.4.1	Number of completed Functions	SP7-SKPI-011	M6: 1; M12: 2; M18: 3; M30: 5
T7.4.2	Number of completed Functions	SP7-SKPI-012	M6: 1; M9: 2; M12: 3; M15: 4; M18: 5; M27: 7
T7.4.3	Number of completed Functions	SP7-SKPI-013	*tbd*
T7.4.4	Number of completed Functions	SP7-SKPI-014	M6: 1; M9: 2; M12: 3; M18: 4; M21: 5; M24: 6; M27: 7
T7.4.5	Number of completed Functions	SP7-SKPI-015	M12: 1; M18: 2; M30: 5
T7.5.1	Number of completed Functions	SP7-SKPI-016	M10: 1; M13: 2; M15: 3
T7.5.2	Number of completed Functions	SP7-SKPI-017	M13: 2; M15: 3

Task	KPI	ID	Target values
T7.5.3	Number of completed Functions	SP7-SKPI-018	M10: 1; M13: 2; M15: 3
T7.5.4	Number of completed Functions	SP7-SKPI-019	M13: 3; M18: 4
T7.5.5	Number of completed Functions	SP7-SKPI-020	M8: 1; M12: 2; M13: 3
T7.5.6	Number of completed Functions	SP7-SKPI-021	M7: 1; M9: 2; M10: 5; M11: 7: M12: 8; M15: 9
T7.5.7	Number of completed Functions	SP7-SKPI-022	M12: 1; M13: 2; M16: 3; M18: 4
T7.6.1	Number of completed Functions	SP7-SKPI-023	M6: 2; M9: 3; M12: 4; M15: 5; M18: 6; M30: 9
T7.6.2	Number of fields for collaboration identified	SP7-SKPI-024	More is better
T7.7.1	Number of SP7 Subproject Committee Meetings held	SP7-SKPI-025	More is better

Table 10: Key Performance Indicators (KPIs) for SP7 Tasks

4. Glossary

Artifact - a high data-density discrete data element, primarily meant to denote a file larger than ~10kB, which is not human readable or editable.

AUP – Acceptable Use Policy

CAVE – a high resolution, 3D fully immersive virtual reality environment where projectors are directed to three, four, five or six of the walls of a room-sized cube.

CERT – Community Emergency Response Team

Cockpit - desktop, display wall or CAVE visualisation resource with a mechanism for good data locality.

Compute resource - a computer or collection of computers where a Job can be executed.

CSIRT - Computer Security Incident Response Team

DANTE - plans, procures, builds and operates large-scale, advanced high-speed networks for global research and education, including the 500 Gbps pan-European GÉANT network.

Datatype – A datatype is a semantically enriched mimetype. For example, the mimetype of a particular data file might be XML, but the datatype would be CircuitML, implying that the data file can be interpreted in a richer way. This allows the selection of editing interfaces and input data much more user friendly in the Unified Portal.

GÉANT - the pan-European research and education network that interconnects Europe's NRENs.

GUI – Graphical User Interface

HBP Unified Portal (HBP-UP) – the unifying web interface through which the web accessible components of the 6 HBP Platforms and all other HBP activity are delivered.

Host - A single operating system instance, running on virtualised or real hardware.

Job - an instance of task execution on a compute resource. For some tasks, the user in the UP on job launch will select the compute resource. For other Tasks, the task will decide the execution.

NREN – National Research and Education Network

Parameter - a low data-density discrete data element, primarily meant to denote a value that one might enter into a single form element. It might also be used to refer to a richer configuration document containing a group of settings.

Platform - software components: libraries, services, APIs and their documentation that are to be used to build portals or cockpits

RAID - a data storage technology that combines multiple disk drive components into a logical unit for the purposes of data redundancy and performance improvement.

Resources - Parameters, Artifacts, services, or compute capacity

REST - acronym for REpresentational State Transfer, for a definition see http://en.wikipedia.org/wiki/Representational_state_transfer

SAN - acronym for Storage Area Network, http://en.wikipedia.org/wiki/Storage_area_network

Service - a software function performed by a third party for a User or other Service. Services consume Parameters, Artifacts and compute capacity. Services produce Artifacts and parameters.

Site - A collection of hosts collected together in a single location. The grouping is potentially arbitrary.

Task - A logical software unit. A Task takes Artifacts and Parameters as input. A Task produces Artifacts and Parameters as outputs. It may or may not be visible as a Service. A Task identifies its dependencies and its default parameters. Specifically, it is a software component that combines:

- A Python-based task entry point
- A Git repository or Python package index URL for the task
- A repository revision or package content specified by sha1
- A requirements file specifying all required dependencies. Tasks can have dependencies in non-Python languages, but these dependencies must be packaged for reproducible deployment.

Task definition - the collection of data that defines an individual Task.

Task repository - a database of Task definitions.

ToU – Terms of Use

Workflow - A tree of decision structures and Tasks. A Workflow takes Artifacts and Parameters as input. A Workflow produces Artifacts and Parameters as output. It may or may not be visible as a Service.

5. References

[1] https://www.humanbrainproject.eu

[2] Brain Simulation Platform Specification (Deliverable D 6.7.1)

[3] Annex I to the HBP Grant Agreement – “Description of Work”

[4] PCP Planning Document (Deliverable D 7.7.1)