Grant agreement no. 675451

CompBioMed

Research and Innovation Action
H2020-EINFRA-2015-1
Topic: Centres of Excellence for Computing Applications

D5.3 Report on access mechanisms to HPC systems

Work Package: 5

Due date of deliverable: Month 12

Actual submission date: 29 / September / 2017

Start date of project: October, 01 2016 Duration: 36 months

Lead beneficiary for this deliverable: *UVA*
Contributors: SARA, *UEDIN, BSC*

Project co-funded by the European Commission within the H2020 Programme (2014-2020)		
Dissemination Level		
PU	Public	YES
CO	Confidential, only for members of the consortium (including the Commission Services)	NO
CI	Classified, as referred to in Commission Decision 2001/844/EC	NO

*"This project has received funding from the European Union's Horizon 2020 research and innovation programme under the Grant Agreement No **675451**"*

Disclaimer

The content of this deliverable does not reflect the official opinion of the European Union. Responsibility for the information and views expressed herein lies entirely with the author(s).

Table of Contents

List of Tables and Figures

*"This project has received funding from the European Union's Horizon 2020 research and innovation programme under the Grant Agreement No **675451**"*

1 Version Log

Version	Date	Released by	Nature of Change
V1.0	10/08/2017	Marco Verdicchio (SURFsara)	First Draft
V1.1	6/9/2017	Marco Verdicchio (SURFsara) Victor Azizi (UvA) Gabor Zavodszky (UvA) Nuno Ferreira (SURFsara)	Added sections on HPC resources, Cloud computing and Computational biomedicine access.
V1.2	7/9/2017	Toni Collis (UEDIN) Jazmine Aguado Serra (BSC)	Added details on HPC centres and review.
V1.3	8/9/2017	Marco Verdicchio (SURFsara) Victor Azizi (UVA) Nuno Ferreira (SURFsara)	Additions and review
V1.4	20/9/2017	Jonas Latt (UNIGE) Andrew Narracott (USFD)	Review
V1.5	27/9/2017	Emily Lumley (UCL) Peter Coveney (UCL)	Final review and Release

*"This project has received funding from the European Union's Horizon 2020 research and innovation programme under the Grant Agreement No **675451**"*

2 Contributors

Name	Institution	Role
Marco Verdicchio	SURFsara	Author
Nuno Ferreira	SURFsara	Co-Author
Victor Azizi	UvA	Co-Author
Gabor Zavodszky	UvA	Co-Author
Toni Collis	UEDIN	Contributor
Jazmin Aguado Serra	BSC	Contributor
Jonas Latt	UNIGE	Reviewer
Andrew Narracott	USFD	Reviewer
Emily Lumley	UCL	Reviewer
Peter Coveney	UCL	Reviewer

“This project has received funding from the European Union’s Horizon 2020 research and innovation programme under the Grant Agreement No ***675451****“*

3 Definition and Acronyms

Acronyms	Definitions
CoE	Centre of Excellence
HPC	High Performance Computing
SME	Small and Medium Enterprise
WP	Work Package
CPU	Central Processing Unit
GPU	Graphics Processing Unit

*“This project has received funding from the European Union’s Horizon 2020 research and innovation programme under the Grant Agreement No **675451**“*

4 Executive summary

The adoption of HPC technologies to perform a wide range of biomedical numerical simulation activities, which investigate complex phenomena and study new prototypes, is crucial to help CompBioMed users in their research and to help clinical and industrial partners to innovate products, processes and services and thus to be more competitive. To this end, in this report we have characterized the various compute patterns of the applications used within CompBioMed and identified typical access models that allow users to run their codebases efficiently.

We have also analysed the main access mechanisms used by Cloud providers and HPC centres to access their resources. With these findings, we set up a survey to investigate possibilities for emerging compute patterns that are currently not (widely) available, namely distributed computing and urgent computing.

This report is based on the results of the work conducted during months M7-M12 for the task T5.2:

Task 5.2: Access Mechanisms for Computing and Data Resources (M7-M12)
Leader: UvA (4 PM);
Partners: SARA (4), UEDIN (2 PM), BSC (2), UCL (2). Here we will conduct the technical investigation of access models currently in use at large HPC centres, as well as those commonly adopted by the biomedical computing but not (yet) widely integrated with HPC. This will cover the usual batch systems, as well as methods to provide on- demand and urgent access to users and applications, and methods to provide automated access for large workflows. Then, relying on the results from task 5.1 we will propose access models to the shared computing and data e-infrastructure. We will conduct feasibility studies on the extension of and addition to the current access mechanisms and select requirements of high priority, based on underlying needs of the community and the upfront possibilities at the partners' facilities. We will then test these new access and execution methods through pilot projects on the partners' infrastructure. We anticipate that the results of these pilot projects will lead to firm and mature new access models, suited for the biomedical community at large, and will push for widespread adoption through the activities in task 5.4.

The outcome of this report will be used as input for other tasks in WP5 (Task 5.4 "Increasing Uptake of Existing e-infrastructures and Alignment with International Projects" and Task 5.5 "Integrating Compute and Data Infrastructure") as well as for the work of WP2 and WP6 on the deployment and development of CompBioMed applications and workflows.

*"This project has received funding from the European Union's Horizon 2020 research and innovation programme under the Grant Agreement No **675451**"*

5 Introduction

Today's generic HPC environments, characterized by highly parallel and memory efficient applications that run on many cores, occupy only the lowest layers of the vertically integrated application stacks that the CompBioMed Centre of Excellence (CoE) will develop. For example, these applications may be characterized by multiple applications linked together and complex multi-physics/multiscale workflows.

Many of the applications used in computational biomedicine require access to high-end computational resources and data storage services, as well as to specialized hardware and dedicated computing services (for data analytics, visualization, etc.). In the deliverable D5.2 "Report on computing and data needs of the biomedical community" we presented an analysis of the main computational strategies adopted within the CompBioMed project, showing that the ability to access both supercomputers and cloud computing infrastructure is essential for the users within this community.

CompBioMed users require not only high computing power, but also flexible and secure access to high performance resources including advance reservation and on-demand access to these resources. Such forms of access are crucial for clinical applications (e.g. to support clinical decision-making) as well as for industrial partners whose requirements are also stricter in terms of access security and quality of service. CompBioMed has the technical tools that allow us to provide such services, and the combined use of cloud computing and HPC infrastructures through the use of specialized services for the management and reservation of dedicated computing resources will be of great benefit for the CompBioMed user community. This will be essential to build next generation biomedical applications that can run on emerging exascale architectures.

In this report, we provide a technical analysis of the main mechanisms adopted to access the services offered by HPC centres and cloud providers. We also analyse the requirements of CompBioMed users in terms of access topologies and provide a basis for a more detailed study to provide insight into the needs for non-conventional access models (such as advance reservations, on-demand access) and into the feasibility of implementing such mechanisms on the side of the HPC providers.

*"This project has received funding from the European Union's Horizon 2020 research and innovation programme under the Grant Agreement No **675451**"*

6 Access mechanisms to HPC resources

In this section, we present an analysis of the access models currently used at HPC centres, as well as those adopted by cloud computing providers. We describe the service components which characterize the different resources (supercomputers and cloud computing) and the access models identified for the use of the specific service.

6.1 Supercomputers

The term ‘supercomputer’ is often applied to a broad range of architectures from the most powerful computers in the world (e.g. the pan-European HPC supercomputers - Tier 0, National HPC supercomputers - Tier 1) to computer clusters (Regional HPC facilities - Tier 2, Local infrastructures - Tier 3). Supercomputers are usually massively parallel with distributed resources, powerful processor units and large memory, and they depend on specialist programming methods to make effective use of their high-end components.

6.1.1 Service Models

Supercomputers are dedicated resources with restricted user permissions, to ensure optimal use of the computational resource for software execution. Supercomputers provide dedicated CPU time on systems consisting of a collection of specialized computing nodes. These nodes are usually equipped with multiple processors (CPUs and/or GPUs) and fast memory access, interconnected by ultrafast and reliable network systems which allow efficient exploitation of the high level of parallelism by the modern software. The multiple nodes of a supercomputer are generally constructed with similar components, facilitating tuning and optimisation of codes across the entire machine for the architecture provided. Specialized file systems (e.g. LUSTRE, GPFS) are usually coupled with the compute elements to allow fast access to the temporary files created and required during the simulation.

In many cases, longer-term raw or processed data needs to be archived in a safe and secure environment. Data storage services are also provided by many HPC centres. These services are designed for long-term storage of large quantities of files, and are commonly interfaced with tools for the backup and preservation of archived data.

Supercomputing centers support a diverse range of applications and simulation software. The full software distribution is centrally managed and provided by specialists in order to run optimally on the target hardware. In many cases, HPC centres also provide access to specialized tools for remote data visualization and high-performance data analysis which may not be available at the user’s site.

The expertise of HPC centre staff also plays a key role in allowing researchers to exploit the computing power available to drive their investigations forward. Users can request general support as well as consultancy for their computational problems: from support for improving software performance or implementation of new algorithms to improved hardware usage and runtime tuning.

6.1.2 Access to HPC resources

In a supercomputer, users do not interact directly with the computing resources, but they interface with a batch system (e.g. SLURM [1], PBS [2], LoadLeveler [3], etc.) which manages the scheduling of jobs, the reservation of resources and monitor the job execution/error

*“This project has received funding from the European Union’s Horizon 2020 research and innovation programme under the Grant Agreement No **675451**“*

handling. A schematic representation of this type of access mechanism is depicted in Figure 1. Batch systems are extensively adopted because of their ability to track resource usage for user accounts and budgeting control, mediate the access to specific resources (specialized computing nodes, software licenses, etc.) and ensure that all users get a fair share of the compute resources.

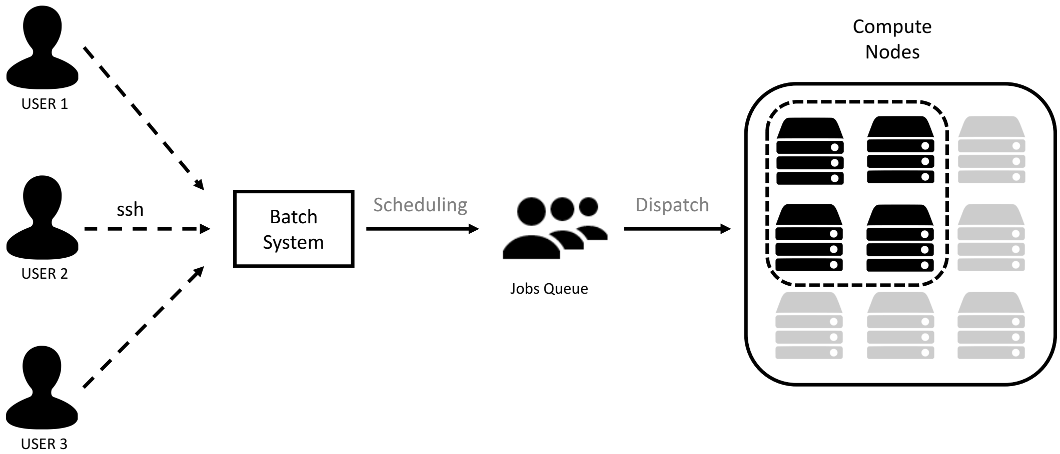

Figure 1. Schematic representation of HPC resources access mechanism.

Moreover, because the scheduler manages the queuing, execution and input/output stream of a user's job, by using batch systems it is possible to submit large, long calculations or many simultaneous jobs, leaving the system to schedule their execution when the requested resources become available. The users connect to these systems mainly through SSH (Secure Shell Access) which relies on the traditional username and password or public key encryption authentication mechanism. Many computing sites also provide dedicated middleware (e.g. QCG [4], Unicore [5], etc.), often available through a graphical interface, to enhance and simplify the interaction with the resources. These services may require the use of digital certificates (i.e. X.509 certificate) for the authentication and authorization of their users, with the additional difficulties of obtaining and correctly manage such a type authentication mechanism.

HPC centres offer, sometimes, the possibility to access computing nodes dedicated to "on-demand" interactive jobs. Machines of this type are usually dedicated to a specific service such as visualization, post-processing, data storage and analysis etc. On these nodes, users can directly interact with the computing hardware, but are usually limited in terms performance, available software, libraries and level of parallelism.

HPC centres provide access to their resources through different allocation mechanisms. Research groups and institutes can apply for compute time and data storage using national funding schemes (e.g. NWO [6], EPSRC [7], NERC [8], RES [9]). Compute resources are granted through competitive calls where the user submits a proposal with a description of the resource requested and the research to be performed. These proposals are evaluated by technical and scientific committees which decide whether or not to grant the allocation on the requested system for a specific timeframe. In addition to national schemes, users can submit proposals

*"This project has received funding from the European Union's Horizon 2020 research and innovation programme under the Grant Agreement No **675451**"*

for R&D research purposes, to one of the many calls from European projects dedicated to the HPC community, such as PRACE [10], DECI [11], EGI [12], etc., which allow access to world class computing including associated data management resources and services.

Ad hoc service agreements can also be established between the resource provider and specific scientific communities (e.g. Life Science GRID [13], HECBioSim [14], etc.) to allow their users preferential access routes to the HPC resources. The CompBioMed framework, for example, is designed to support biomedical users in their access to large-scale computational facilities, providing dedicated computing time on partner infrastructures (Tier 0 and Tier 1 systems) as well as easy access to data storage and cloud computing services.

Many computing centres also offer the possibility to purchase compute time, as well as specialized support, following a pay-per-use business model which is especially convenient for industry and SMEs who cannot afford to buy and operate their own in-house HPC capability.

6.2 Cloud computing

The National Institute of Standards and Technology (NIST) [15] defines Cloud computing as a model for enabling ubiquitous, convenient, on-demand network access to a shared pool of configurable computing resources that can be rapidly provisioned and released with minimal management effort or service provider interaction.

6.2.1 Service Models

The traditional breakdown of the service models offered by Cloud service providers is depicted in Figure 2 [16].

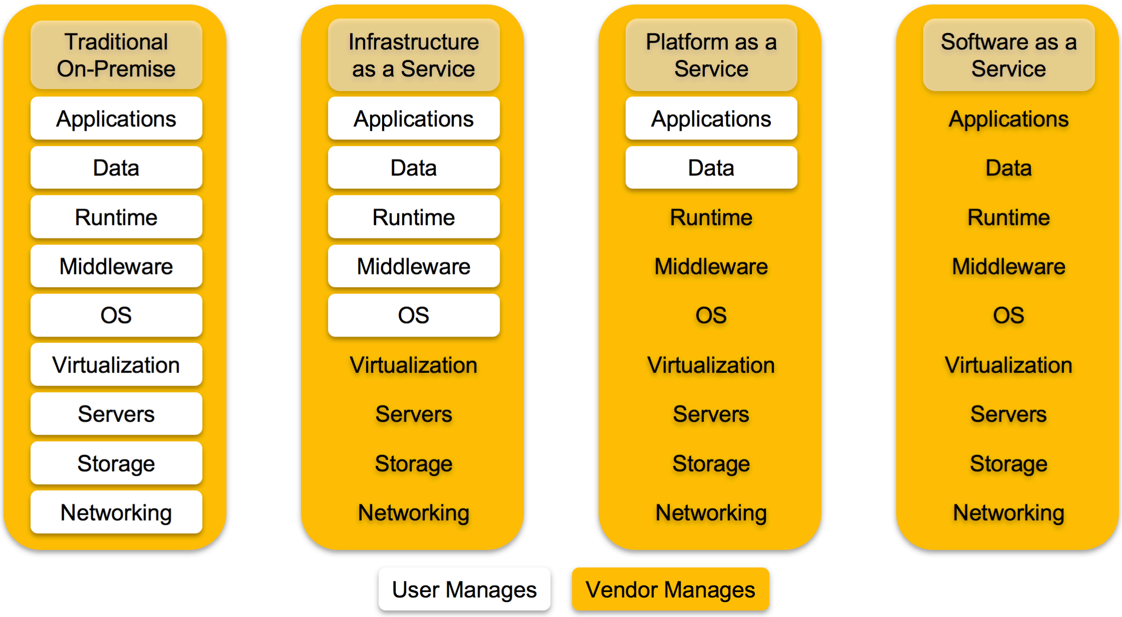

Figure 2. Traditional Cloud computing service models.

In comparison with traditional on premise resources (e.g. personal computer, private cluster), where the customer manages the full stack, the different cloud models offer different degrees of control. As an example, the Infrastructure as a Service (IaaS) model offers the customer the capability to provision processing, storage, networks, and other fundamental computing resources where the customer is able to deploy and run arbitrary software (from the

*"This project has received funding from the European Union's Horizon 2020 research and innovation programme under the Grant Agreement No **675451**"*

Operating System to the application level). Nevertheless, the customer does not manage the underlying cloud infrastructure (computing and storage nodes, network and virtualization technology) as is the case at most supercomputing centres. IaaS is the most flexible cloud model and is the core business of most of the big players acting as Cloud resource providers (Amazon's AWS, Microsoft's Azure, Google Cloud Platform, etc.) which own the resources provided to the final user.

Iaas cloud computing enables on-demand access to resources, without the overhead of their ownership, by offering the required services through the use of one or more virtual machines. This can have a negative impact on performance, depending on the particular problem under consideration. If an application has tightly coupled processes, requiring significant communication, cloud computing is likely to perform less well than dedicated supercomputing facilities. Although virtualisation is now a minimal overhead, due to the use of multiple virtual machines and the lack of fast interconnections, communication latency in cloud computing is higher than in a dedicated HPC resource, resulting in reduced performance. Despite the performance gap is closing, as cloud service providers adopt hardware capable of the performance of HPC (i.e. Azure now uses Infiniband, a networking standard commonly used in HPC system), cloud computing remains extremely expensive for compute intense applications.

At the other end of the service model spectrum, Software as a Service (SaaS) offers the customer the capability to use applications pre-deployed by the provider running on cloud infrastructure (e.g. apps on a smartphone). The customer does not have control of any components of the stack, with the possible exception of limited user specific application configuration settings.

6.2.2 Access to cloud services

The access mechanism to Cloud resources is dependent on the service model offered by the Cloud provider and/or being explored by the customer (see Figure 3). The more flexible the cloud computing model is in terms of customer control over the provisioned resources, the more access mechanisms are allowed. For example, in a Software as a Service (SaaS) model, the applications are accessible from various client devices through either a thin client interface (e.g. web browser), or an Application Programming Interface (API). In an Infrastructure as a Service (IaaS) model, the customer can also have access to the resources, encapsulated inside the so-called Virtual machines, via Secure Shell Access and/or remote desktop software.

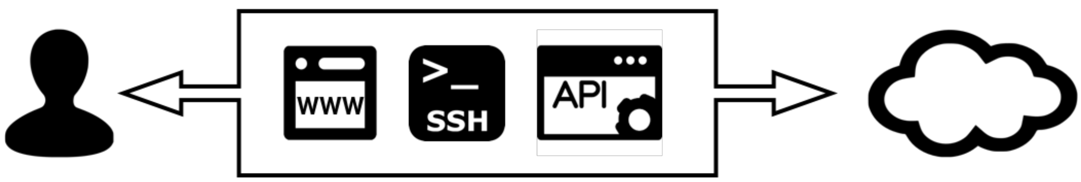

Figure 3. Generic access mechanisms to Cloud services.

Commercial cloud providers (e.g. Amazon Web Services [17], Microsoft Azure [18]) provide a pay-per-use access to their resources and/or services. Within the IaaS model of cloud

*"This project has received funding from the European Union's Horizon 2020 research and innovation programme under the Grant Agreement No **675451**"*

computing, resources are owned and maintained by the cloud provider which offer access to compute resources, storage, network capabilities, etc. over the Internet. In this way, the user has access to low cost infrastructures (because usage from a high number of customers is aggregated in the cloud, providers can offer cheaper services to their users) which can scale up and down as required.

Commercial cloud providers, can also provide a Platform as a Service (PaaS) which rely on the infrastructures made available by others. Examples of these are cloud providers such as Alces Flight [19] and DNANexus [20] both of whom are Associate Partners within CompBioMed. PaaS providers abstract most of the work of dealing with the underlying resources and give users a specialized environment (including for example additional security layers or efficient resource management tools). These type of cloud providers can either include in their billing models the price of the resources used together with the delivered services (i.e. DNAnexus access model) or leave the users the responsibility of obtaining an account with the IaaS provider and set up their own virtual environments (i.e. Alces Flight access model) charging the customers only for the delivered services and/or support.

Users from academy and research institutions can benefit from dedicated access channel to specific cloud environments developed to support scientists, multinational projects and research infrastructures. Examples of this type of cloud providers are the EGI Federated Cloud [21] and the SURFsara HPC Cloud [22], which provide an IaaS-type Cloud made of academic and/or private clouds. Their development is driven by the requirements of the scientific community and users from universities and/or research institutes. These academic clouds cannot compete with commercial providers in terms of resource availability and quality of the services (security, incidents management, etc.), but they provide highly specialized support, usually with competence in the specific scientific domain, and ease of access for researcher and academics. Users can obtain allocations to these cloud infrastructures by submitting an application through the dedicated channels. Applications are reviewed on their scientific merit and technical requirements match towards the infrastructure being requested. Once approved, a work environment is setup by the Cloud provider with agreed quota resources. Initial consultancy support is also often provided to the customer.

7 Access models for computational biomedicine

In deliverable 5.2 of the CompBioMed project we conducted a survey which collected the computing needs of the users in the project. With these results, we want to identify the existing needs of this community. Furthermore, we are interested in what partners with available computing power (compute centres and commercial cloud computing companies) can offer to these users right now, and what they might offer to meet unfulfilled needs of the community.

From the survey conducted for deliverable D5.2, we identified the target systems and the execution patterns for typical applications that are being used in the CompBioMed project. Table 1 summarizes the typical modes of operation and target systems for the codes used in the three main research exemplars present in CompBioMed.

Table 1. Summary of the main compute patterns adopted in CompBioMed.

CompBioMed research exemplar	Typical mode of operation in CompBioMed	Simulation codes used	Target System

"This project has received funding from the European Union's Horizon 2020 research and innovation programme under the Grant Agreement No **675451**"

Cardiovascular/ Respiratory	**Highly parallel single run**	Alya CHASTE FLIBRA HemeLB Hemocell Palabos Ansys	- Tier 0
	Intermediate parallel concurrent runs	CHASTE Palabos Ansys	- Tier 1 - Tier 2 - Cloud
Molecular-based medicine	**Highly parallel single run**	GROMACS NAMD AMBER Desmond	- Tier 0 - Tier 1 - GPUs
	Intermediate parallel concurrent runs	GROMACS NAMD AmberTools ACEMD GAMESS-US Gaussian	- Tier 1 - Tier 2 - GPUs - Cloud
Neuro-musculoskeletal medicine	**Serial concurrent runs**	MATLAB Ansys	- Tier 2 - Cloud
	Weakly parallel concurrent runs	MATLAB Ansys	- Tier 1 - Tier 2 - Cloud

From these findings, we propose four different types of access for an application; cloud computing, batch computing, distributed computing and urgent computing:

1. Cloud computing: Rent the hardware, run your own software on it. Instant access to machines most of the time at a higher cost. All software maintenance is handled by the user if making use of a Cloud IaaS implementation. No software maintenance required on a SaaS Cloud model.

2. Traditional Compute Centre Access: Resource access is based on queues handled by a batch scheduler (e.g. SLURM [1]), access time can fluctuate based on the load of the HPC system and it can vary from few seconds, when the requested resources are available at the submission time, to several hours on heavy loaded machines. This type of access mechanism includes also the possibility to reserve in advance the resources for the job being executed by the user.

3. Distributed Computing: Higher level abstraction of HPC centres, where the actual HPC machine is not selected by the end user. Rather, the submitted task is dynamically allocated where resource is available and the quality of service is guaranteed by the middleware (e.g. QCG [4], Mapper [23]). The use of a middleware allows also users to request for advanced reservation of resources and co-reservation of multiple infrastructures. Because of the availability of multiple HPC systems, access time for distributed computing can be drastically reduced compared to traditional compute access (2) targeting the centre with less workload at the moment of the submission. However, if a job needs to have simultaneous access to multiple HPC centres, the access time can increase drastically.

*"This project has received funding from the European Union's Horizon 2020 research and innovation programme under the Grant Agreement No **675451**"*

4. Urgent Computing: There is a hard deadline for the start of the execution of a job. This can be implemented in two ways, either other jobs are stopped in favour of this one (pre-emptive) or a job gets the next available slot, with a maximum on the waiting time (next-to-run). This does not include advance reservation as possible within pattern 3. It is not known in advance when it might be necessary to run an urgent computing task.

The first two options already exist whereas the last two are emerging compute patterns that are not yet (widely) available.

Cloud computing and traditional compute centre access are the current access mechanisms with which it is possible to gain access to computing resources. However Distributed computing is only just now being developed (within the ComPat [24] project for example).

The last compute type needed, Urgent Computing, is a new requirement where time to completion plays an important role in the applicability of the results. One such example is patient-specific medical simulations that primarily target clinical application as part of a treatment protocol. An application example could be a vascular flow simulation after a medical device implantation on a patient-specific vessel geometry to aid the decision-making process. In this case the medical practitioner needs the simulation results within a clinically relevant timeframe, which imposes a hard limit on the latest possible start of execution and total run-time for the computation.

Urgent computing can currently be done in several other ways, with the objective that the resources for a simulation should be available within a specified time frame. It is possible to reserve the resources for the entire duration within which an urgent computing task might occur, however this is generally too costly. Another way would be to use a commercial cloud computing provider, within these it is possible to request resources whenever needed, however a hard guarantee for availability is never given. For that reason, we want to explore the possibilities within traditional compute centres to implement urgent computing without the need of pre-reserving and with the insurances that the job will start within a specified timeframe.

The current stage of the CompBioMed project represents an appropriate point to conduct such a survey as most partners have now set up their codebases and are running them with the desired resource providers within the project's fast track. Thus, investigating these "additional" computing patterns (urgent computing and distributed computing) is a logical next step. Furthermore, the exploration of these issues in the context of the codebases used within the CompBioMed will provide a good understanding of the issues relevant for the computational biomedicine community in general.

8 Conclusions

In order to run computationally intensive calculations CompBioMed users need flexible and easy access to high performance computing infrastructure as well as to data storage facilities capable of correctly managing the high quantities of data produced (or needed as input) by biomedical simulations. For this reason, classic access models adopted by HPC centres are not always suitable for CompBioMed users, in particular for industrial and medical partners who require tools that provide on-demand access and flexible mechanisms. Therefore, the use of cloud computing infrastructure offers an alternative to traditional HPC access which may be attractive for this user community, keeping in mind, although, that the cost of this type of

*"This project has received funding from the European Union's Horizon 2020 research and innovation programme under the Grant Agreement No **675451**"*

services can be very expensive for the compute intense applications frequently used in this project.

In this report, we conducted a technical investigation of the main access models adopted by large HPC centres and cloud providers to access their resources. This analysis shows that, whereas HPC centres provide a high-level service for users to run on their system (e.g. optimized software, support, dedicated hardware, etc.), the access mechanisms to these resources can be a disadvantage for applications which require immediate availability of resources and the possibility to interactively run on the target system (i.e. on-demand and advance reservation access mechanisms). Cloud computing, on the other hand, enables on-demand access to resources, that can be accessed over the network through standard mechanisms, without the overhead of their ownership. In this way, users can request compute capabilities as they need and can acquire and release capabilities to scale rapidly with their demands in terms of compute power and desired time. However, as shown in Section 6, cloud computing resources are provided to the customer through virtualized computing environments (i.e. Virtual Machines). Despite the fact that the usage of VM ensure a high reproducibility of the conducted computations, the software does not run natively on the hardware, resulting in reduced performance, depending on the particular problem under consideration.

Based on the outcomes of D5.2, we also analysed the main access models required by CompBioMed users to run their simulations. We identified four main access models: Cloud computing, Traditional Compute Centre Access, Distributed Computing and Urgent Computing. While the first two access typologies are commonly used in current HPC and Cloud environments, the latter two are not at all widely adopted and have only been used in special cases, one-off studies. For this reason, we will complete a more extensive study of the need for distributed and urgent access mechanisms by surveying the CompBioMed user community (both users and HPC resource providers) to establish improved understanding of the technical requirements and possible limitations of new access mechanisms. The results of the survey will be used to identify the types of application that would benefit from more sophisticated access mechanisms. A suitable use case within CompBioMed will be the subject of a pilot project to exploit and demonstrate the advantage of new tools to facilitate this more sophisticated HPC access.

The outcomes of the survey and the results of the pilot project will be used as input to establish access mechanisms and application modes that fulfil the requirements of the CompBioMed community and to assess and improve usage of the e-infrastructures that will be targeted within the deliverable D5.4 “Report on best practices for e-infrastructure application usage”.

*“This project has received funding from the European Union’s Horizon 2020 research and innovation programme under the Grant Agreement No **675451**“*

9 Annex

Urgent Computing Survey

User Questionnaire

Name	Code name and version number if applicable
Use Case	Why is urgent computing required for this task?
Minimal job resource requirements	What are the minimal constraints to get an acceptable answer?
Improvement of result with scaling	Does the result improve if larger amount of resources is used for the task? Explain scalability?
Maximum waiting time	What is the maximum amount of time that a job can wait on resources before the job becomes unnecessary.

Resource provider Questionnaire

System name	Name of the computing resource and subsystem when applicable
Resource type	The kind of resource (computing, data, visualization)
Resource size	The minimum guaranteed size of the resources
Resource duration	The maximum time the resource could be reserved by an urgent computing job
Maximum waiting time	The maximum waiting time for a job for the resources to become available
Work needed to implement	The hypothetical work involved in adding urgent computing to these resources
Extra cost for urgent computing tasks	If an urgent computing task delays or resets other jobs, these costs should be paid by the urgent computing task owner. What would the resource owner hypothetically bill a typical urgent computing task on top of the normal billing?

*"This project has received funding from the European Union's Horizon 2020 research and innovation programme under the Grant Agreement No **675451**"*

Distributed Computing Survey

User Questionnaire

Name	Code name and version number if applicable
Use Case	Why is distributed computing required for this task?
Minimal job resource requirements	What are the minimal constraints to get an acceptable answer?
Co-Allocation requirements	Does the job require at any point the co-allocation of two different resources?
Other requirements	Is the job special in any other way?

Resource provider

System name	Name of the computing resource and subsystem when applicable
Resource type	The kind of resource (computing, data, visualization)
Resource size	The minimum guaranteed size of the resources
Co-allocation possibility	Is it possible to reserve the resource ahead of time?
Automation possibility	Is there a possibility to reserve through an automated system?
Links with other centres	Is there a system in place to do distributed computing over multiple providers?
Work needed to implement	The hypothetical work involved in adding distributed computing to these resources.
Extra cost for distributed computing tasks	What would the resource owner hypothetically bill a typical distributed computing task with co-allocation on top of the normal billing?

*"This project has received funding from the European Union's Horizon 2020 research and innovation programme under the Grant Agreement No **675451**"*

10 Bibliography

[1] https://slurm.schedmd.com/.

[2] http://www.pbspro.org/.

[3] https://www-03.ibm.com/systems/power/software/loadleveler/.

[4] http://www.qoscosgrid.org/trac/qcg.

[5] https://www.unicore.eu/.

[6] https://www.nwo.nl/en.

[7] https://www.epsrc.ac.uk/.

[8] http://www.nerc.ac.uk/.

[9] https://www.res.es.

[10] http://www.prace-ri.eu/.

[11] http://www.prace-ri.eu/deci-projects/.

[12] https://www.egi.eu/.

[13] http://lsgc.org/.

[14] http://www.hecbiosim.ac.uk/.

[15] http://nvlpubs.nist.gov/nistpubs/Legacy/SP/nistspecialpublication800-145.pdf.

[16] https://www.linkedin.com/pulse/20140730172610-9679881-pizza-as-a-service.

[17] https://aws.amazon.com/.

[18] https://azure.microsoft.com/.

[19] http://alces-flight.com/.

[20] https://www.dnanexus.com/.

[21] https://www.egi.eu/federation/egi-federated-cloud/ .

[22] https://userinfo.surfsara.nl/systems/hpc-cloud .

[23] http://www.mapper-project.eu/.

[24] http://www.compat-project.eu/.

*"This project has received funding from the European Union's Horizon 2020 research and innovation programme under the Grant Agreement No **675451**"*