BalticLSC Environment Vision

Results of the end-user requirements elicitation workshops

Version 1.0

Priority 1: Innovation

Warsaw University of Technology, Poland
RISE Research Institutes of Sweden AB, Sweden
Institute of Mathematics and Computer Science, University of Latvia, Latvia
EurA AG, Germany
Municipality of Vejle, Denmark
Lithuanian Innovation Center, Lithuania
Machine Technology Center Turku Ltd., Finland
Tartu Science Park Foundation, Estonia

BalticLSC Environment Vision

Results of the end-user requirements elicitation workshops

Work package	WP3
Task id	A3.1
Document number	O3.1
Document type	Vision Document
Title	BalticLSC Environment Vision
Subtitle	Results of the end-user requirements elicitation workshops
Author(s)	Henri Hanson and Sven Lilla (TSP) from input provided by Baltic LSC partners
Reviewer(s)	Maj Held (Vejle)
Accepting	Michał Śmiałek (WUT)
Version	1.0
Status	**Final version**

Interreg Baltic Sea Region
EUROPEAN UNION
EUROPEAN REGIONAL DEVELOPMENT FUND

History of changes

Date	Ver.	Author(s)	Change description
20.03.2019	0.01	Krzysztof Marek (WUT)	Document creation and initial contents from Poland
14.06.2019	0.02	Agris Šostaks (IMCS)	Content from Latvia
14.06.2019	0.03	Birute Leipute and Vilma Vilutyte (LIC)	Content from Lithuania
14.06.2019	0.04	Ann-Christin Eriksson (RISE)	Content from Sweden
18.06.2019		Henri Hanson (TSP)	Notes from workshop in Germany
20.06.2019	0.8	Maj Held (Vejle)	Content from Denmark
20.06.2019	0.9	Henri Hanson and Sven Lilla (TSP)	Content from Estonia and compilation of content from partner reports
25.06.2019	0.9	Maj Held (Vejle)	Review of version 0.9
03.07.2019	1.0	Henri Hanson (TSP)	Improved as per review
04.07.2019	1.0	Maj Held (Vejle)	Final Review and confirmation

Executive summary

All scientific disciplines are becoming "computational" today. Modern scientific discovery requires very high computing power and capability to deal with huge volumes of data. At the same time, industry and SMEs are increasingly relying on the power of supercomputers to invent innovative solutions, reduce cost and decrease time to market for products and services.

HPC (High-Performance Computing) is a strategic resource for Europe's future as it allows researchers to study and understand complex phenomena while allowing policy makers to make better decisions and enabling industry to innovate in products and services. Societal, scientific and economic needs are the drivers for the next generation of HPC - computing with exascale performance (computers capable of performing 10 to the power of 18 floating point operations per second).

HPC is part of a global race. Many countries (USA, Japan, Russia, China, Brazil, India) have announced ambitious plans for building the next generation of HPC with exascale performance and deploying state-of-the-art supercomputers. [1]

According to Partnership for Advanced Computing (PRACE) Report, EU loses 2-3% of GDP per year due to being behind in the area of large-scale computing, while the situation in the Baltic Sea region is even worse.[2] With the support of EU Horizon2020 and other support programs, universities and other research organizations are investing heavily into High Performance Computing (HPC) resources for scientific calculations. For SMEs, who could benefit from HPC services, approaching HPC centers at universities is usually not the perfect option because of the high expenses involved as well as their computational needs being smaller than is optimal for HPC level.

From private sector, the computing market today is dominated by heavyweight players, e.g. Amazon, Google, Microsoft, etc. who leverage their market power and assets to ensure hefty margins, resulting also in inefficiently priced computing services.

Baltic Large-Scale Computing (BalticLSC, from 01.01.19 – 30.06.21) is an Interreg BSR project dedicated to solve the issue of smaller companies and institutions suffering from the lack of proper access to super computing resources and necessary expertise to use them in the Baltic Sea region. In the project framework, 8 partners from countries around the Baltic Sea are working on developing and providing a platform for truly affordable and easy to access LSC Environment for end-users to significantly increase capacities to create new innovative data-intense and computation-intense products and services by a vast array of smaller actors in the region.

Two main target groups for Baltic LSC:

- **LSC service providers:** incl. SMEs, research HPC and LSC centers, science parks, etc.
- **LSC service users:** innovators (start-ups, etc.), small/medium R&D centers in various domains (product engineering, weather forecasting, biotechnology, etc.) and companies in the tourism sector

This current Environment Vision Document is an important output of the Baltic LSC project, based on end-user requirements obtained during Baltic LSC project's international and local workshops as well as individual meetings with potential cooperation partners and end-users from March – June 2019. It contains a list of features that reflects the needs of the BalticLSC Environment end-users.

The Vision Document is complemented by two additional documents describing in detail hardware and software visions for the platform set-up:

- Baltic LSC Platform Architectural Vision (Baltic LSC Output 4.1)
- Baltic LSC Software Architectural Vision (Baltic LSC Output 5.1)

[1] Source: https://ec.europa.eu/programmes/horizon2020/en/h2020-section/high-performance-computing-hpc, (18.06.19)

[2] Source: http://www.prace-ri.eu/IMG/pdf/prace-annual-report-2018_LOW-RES.pdf (19.06.19)

Table of contents

1. Introduction

1.1 Objectives and scope

The current Environment Vision document contains a list of features that reflects the needs of the BalticLSC Environment end-users. It summarises all the ideas formulated by participants of various workshops and individual meetings with experts organised by BalticLSC partners during Q1-Q2 2019. Special care was taken to identify features at trans-national level. This included the features that allow for building various instances of the BalticLSC Environment across the Baltic region, forming a distributed network of LSC nodes. Also, it focuses on the features that will allow for networking through sharing and reuse of computing power, distance management of the environment, and trans-national collaboration on computation applications within the Baltic region.

Based on this list, it presents functional features of the system, including functionalities of the BalticLSC Platform and BalticLSC Software. These features describe functionalities from the point of view of typical users of the environment. Furthermore, the vision addresses all the non-functional issues that include performance, security, safety, reliability and other.

Therefore, this vision document seeks to summarize end-user requirements for the BalticLSC system. Based on results of the end-user requirements elicitation workshops, it delineates the BalticLSC Environment Vision.

1.2 Relations to Other Documents

The current document is a compilation of presentations, minutes and other input from different workshops and individual meetings held by BalticLSC partners during Q1-Q2 2019. Individual partner overviews are available at Baltic LSC document Dropbox depository: Dropbox\BalticLSC\WP3_User_requirements\A3.1 End-user requirements elicitation workshops

The Vision Document is complemented by two additional documents describing in detail hardware and software visions for the platform set-up:

- Baltic LSC Platform Architectural Vision (Baltic LSC Output 4.1)
- Baltic LSC Software Architectural Vision (Baltic LSC Output 5.1)

1.3 Intended Audience and Usage Guidelines

This Vision Document is the Output 3.1 as described in the Baltic LSC Application Form and intended for internal use within BalticLSC consortium as basis for future software development activities as well as for reporting purposes for local First Level Control and Baltic Sea Region Program.

2. Positioning

2.1 Business opportunity

Baltic LSC partnership consists of technology creators, business development organizations and science parks from 8 European countries. The partners are working on developing a large-scale computing (LSC) environment that will consist of computing platform and computing software accessible online by professionals and amateurs to create and use complex data processing templates for the benefit of smaller companies and other institutions. As a result, computations would become more efficient, less time-consuming and would save costs for the SMEs in the region.

The BalticLSC Environment will impact the Baltic Sea Region (and beyond) by significantly supporting two major groups of actors:

- First, it will significantly facilitate emergence of new LSC service providers. Even small start-ups will be able to take up the BalticLSC specifications and create their own local LSC centers. Moreover, several such local centers will be able to network and combine their computing capacities, distributed trans-nationally.
- Second, even very small innovation and design centers will be able to develop and test easily their own computing applications within the emerging BalticLSC environment. This way, they will significantly reduce time-to-market of innovative engineering products through easy access to LSC resources that were previously not accessible and too difficult to use.

During the Baltic LSC project an initial network of BalticLSC centers will be implemented that will facilitate building an active community of BalticLSC providers and users. Based on interactions and feedback from innovators, engineers and researchers from across the BSR, the project will work out a solution that is self-sustainable, replicable, transferable and easy to disseminate, assuring scalability, affordability, ease-of-use, and efficient distribution across the BSR. In the future, expanding the network between universities and linking to a broader range of European companies is foreseen.

2.2 Problem statement

SMEs, who could benefit from LSC services, approaching HPC centres at universities is usually not the perfect option because of the high expenses involved as well as their computational needs being smaller than HPCs can provide. HPCs at universities usually develop models that often just end up in their own environments and are only being used by Academia. However, these are usually open source and could be useful for private companies to make use of. One of Baltic LSC tasks would be mapping the possibilities of providing access to these models/applications and offering feedback to their developers of how the models/applications work in "real life" (the usage and performance).

From private sector, the computing market today is dominated by heavyweight players, e.g. Amazon, Google, Microsoft, etc. who leverage their market power and assets to ensure hefty margins, resulting also in inefficiently priced computing services as well as specific expertise in building custom software for SME-specific tasks.

The problem of available HPC resources affects both public institutions and private companies that deal with large amounts of data. Inefficient computations cost these entities extra time and expenses as well as significantly limit opportunities for their growth. A successful solution would establish a working service model and would satisfy companies' needs in specific areas and for easier access, offer predetermined computation models with assisted customization.

The solution should be accessible online, offer sufficient bandwidth for large amounts of data to be transferred fast, be highly reliable and secure.

2.3 Product position statement

For SMEs that require complex computations, BalticLSC is an online platform that provides easy access computation services in specific areas. Unlike private HPC clusters, BalticLSC unites computation resources in one network and can provide the best performance using several preselected models. Unlike competitors, who accept all resource providers, the network would include only pre-selected partners to minimize security risks.

3. Stakeholder and user descriptions

3.1 User summary

Overall, the BalticLSC Environment users can be divided into four categories: end-user, resource provider, developer and administrator. Below detailed description of different users.

Name	**End-user**
Description	End user is using BalticLSC to perform large scale computing with provided ready-to-use applications they can find in the BalticLSC Marketplace. They are paying for used computing resources and applications with credits purchased at BalticLSC.
Name	**Resource provider**
Description	Resource provider provides computing resources to the BalticLSC Network. In exchange they receive credits from end user for computations performed on provided resources.
Name	**Developer**
Description	Developer provides applications and modules to the BalticLSC Marketplace using BalticLSC Dockerfiles. In exchange for use of their applications and modules they receive credits from the end users.
Name	**Administrator**
Description	Administrator manages computing resources within the BalticLSC Network. Administrator is also responsible for approving new resource providers, new modules and applications for the BalticLSC Marketplace. The Administration of the Baltic LSC network after the end of the project needs to be discussed during set-up and initial operation of the Network.

3.2 Stakeholder summary

A wide list of different stakeholders will be involved to promote, utilise and further develop the Baltic LSC Environment in the future. Sustainability of the service is of key importance as is partnership with stakeholders. Below an overview of different stakeholders and their roles during preparation and after launch of Baltic LSC Environment.

Name (Type)	**Description**	**Members**	**Role in the development effort (responsibilities; interests)**
Governmental institutions	Governmental offices, e.g. Departments of Statistics, Business Registries, National Land Service, etc. as well as ministries and/or agencies that are responsible for regional development, innovation support, etc.	Ministries; Agencies for Science, Innovation and Technology, Departments of Statistics; State Enterprise Registries, State Tax and Audit offices, National Banks, etc.	Enablers as well as end-users. Describe use cases; provide input for end-user requirements, funding and other support (e.g. change or enact laws concerning various aspects relevant for supercomputing or HPC (e.g. GDPR).

Type	**Description**	***Members***	***Role in the development effort (responsibilities; interests)***
State enterprises (incl. Research Institutions)	Other large public institutions	Hospitals/Clinics; Research Institutes; National Radios and Televisions; etc.	Resource providers and/or end-users. Might be interested in small scale testing before approaching HPC centers.
Public Business support organisations	Organizations that have the capacity to provide shared R&D infrastructure for companies and/or startups. That integrate services and connect research institutes, higher education institutions, business enterprises and other organizations.	R&D centers, Clusters/Open Access Centers/Science and Technology Parks/Integrated Science, Studies and Business Centers ('Valleys' in Lithuania)/ Business Incubators and Accelerators	They are generally interested in the topic, can provide insights in the development of sustainable business models. Also play a substantial role in networking and promotion of the service. Can host servers or computers. May act as resource providers (e.g. infrastructure hosts), brokers or as business development supporters.
Non-governmental organisations	Organizations that have a mission to promote certain goals. In this case – education in the field of HPC.	e.g. Turing Society (NGO), governmental initiatives "Kurk Lietuvai" ("Create for Lithuania"), , student communities and NGOs etc.	Are interested to use the system for experimentation and education purposes. Potential supporters, promoters testers and/or end-users of the system that will be developed. Can offer multiple perspectives from different points of view (regarding hardware, software, usability, etc.)
Associations	Organisations uniting ICT professionals and companies, and/or companies in knowledge and data-intensive sectors, in which supercomputing or HPC have clear benefits and would be extremely useful.	e.g. Vilnius Industry and Business Association (associated partner from Lithuania); Lithuanian ICT association "INFOBALT"; associations of identified sectors (listed under the section of user profiles)	Can consult on technical aspects, provide use cases and formulate end-user requirements. Also help to identify institutions and companies that specialize in supercomputing or HPC, can recommend relevant people for cooperation.
Type	**Description**	***Members***	***Role in the development effort (responsibilities; interests)***

Public HPC service providers (resource providers; developers; administrators)	Usually a supercomputer administrator at university or Open Access Center	Universities and R&D centers specialised in HPC and similar services	Compatibility of end-user requirements with computing capabilities and owned resources. It is probable that service providers will be involved in the technological development of the system and may use it as one of the tools among others in the service package.
Small and medium-sized enterprises (SMEs)		e.g. any SME approaching BalticLSC partners either through direct contacts, workshops held or other channels	Describe use cases; provide input for end-user requirements. Main target group and future users. Their input is most important because the future success of the system depends on how well it will be able to retain these stakeholders.
Private HPC cluster owners	Usually a company or an organization that has its own computation resources.	e.g. Geomatrix, Ltd.; Turing Society (NGO) in Lithuania, the coming AI House in Vejle, Denmark	Since these entities can act both as service providers and end-users, their input is especially valuable. They also have a more defined interest to establish certain service quality.
Large Enterprises		e.g. Banks, telecommunication groups, Energy companies, etc.	End-users; private HPC clusters owners; potential sponsors?

3.3 Key stakeholder problems

As with all emerging fields in business there is still confusion within the stakeholders on what is already possible in the field of LSC and how to fully utilise the rising opportunities for improving their daily business needs. Besides own computers, supercomputers, custom architecture HPC clusters are often used by companies, as well as different cloud services (Amazon, Google, Microsoft, etc.).

Below a list of key stakeholder problems discussed during the Baltic LSC workshops. The different aspects described cover a wide range from security issues (data access), availability and price of LSC resources, access to competencies, algorithms and data etc.

Short name	**Lack of computational resources**		
Id	IMCS_PR1	Priority	6
Description	Computations require more resources than the stakeholder owns. As a consequence, the problem solving takes too long or can't be done at the required scale.		
Related Stakeholders	End-users as problem owners, resource providers as problem solvers		
Current solution	Renting or purchasing appropriate resources.		
Competition	Using cloud resources, Purchasing a new resource		

Short name	**Lack of appropriate algorithms (software)**		
Id	IMCS_PR2	Priority	4
Description	Computations require specific algorithms (software) which are hard to build (lack of knowledge, skills). As a consequence, the solution to the computation problem is expensive to build. A possible solution - standardized solutions or modifiable templates.		
Related Stakeholders	End-users as problem owners, developers as problem solvers		
Current solution	Renting or purchasing appropriate software. Building own software.		
Competition	Using cloud services, Building own software		

Short name	**Lack of appropriate data**		
Id	IMCS_PR3	Priority	3
Description	Computations require specific data which is hard to collect and maintain properly. As a result, the solutions to the computation problem is not accurate and not competitive.		
Related Stakeholders	End-users, developers as problem owners, data providers as problem solvers		
Current solution	Renting or purchasing appropriate data or services that use such data.		
Competition	Using cloud services		

Short name	**Lack of appropriate knowledge**		
Id	IMCS_PR4	Priority	6
Description	Computations require specific solutions and a stakeholder is not aware of how to solve it. As a result, the computation problem can't be solved. A possible solution - standardized solutions or modifiable templates.		
Related Stakeholders	End-users, developers as problem owners, outsourced specialists as problem solvers		

Current solution	Consultancy in HPC centers of Universities for HPC problems. Consultancy enterprises, on-line courses and materials for smaller problems.
Competition	Consultancy services

Short name	**Data uploading and processing**		
Id	LIC_PR1	Priority	4
Description	Uploading the data in real-time increases drastically the load on the network and can be considered a suspicious activity by network administrators		
Related Stakeholders	IT companies, Manufacturers		
Current solution	In-house LSC data processing or manual transfer of data from end-user to HPC service provider		
Competition	Using EuroHPC Cluster, local HPC service providers		

Short name	**Not enough in-house resources**		
Id	LIC_PR2	Priority	5
Description	If it is a larger company or a company that specializes in IT related fields, then there are more people involved and project teams are formed. Task cycles are not discussed (whether time cycles have shortened or lengthened). However, time-consuming activities, such as data preparation, model development, its optimization should be taken into account. Baltic LSC could provide consultancy on optimal allocation of resources, testing, debugging, model validation, data interpretation, etc.		
Related Stakeholders	IT companies, Manufacturers		
Current solution	Outsourcing complete service		
Competition	Using EuroHPC Cluster, local HPC service providers		

Short name	**Simulation Type Problems**		
Id	WUT_PR1	Priority	5
Description	Many computing problems with commercial potential have been enumerated. Then they were divided into two groups: Simulation and Data Processing. This division significantly differentiates tasks based on the amount of data that should be transmitted into computation nodes. Simulation: • Engineering optimization - Hull optimization, CFD (air flow, aerodynamics, turbines, metal casting); • Simulation of natural processes (weather forecast, watershed simulation, traffic/crowd simulation, customer behavior, customer path in the shop); • Road planning (logistics-optimization, transport optimization/scheduling, cosmic logistic) • Molecular modeling (chemistry, biology, material science …); • 3D Rendering.		
Related Stakeholders	Technical University, Engineering Company, Chemical Company, etc.		
Current solution	In-house computations for smaller companies, using large HPC centres for big companies		

Competition	Using AWS (requires programming knowledge), Building an own cluster (big investments, resources not always needed)

Short name	**Data Processing Type Problems**		
Id	WUT_PR2	Priority	5
Description	Many computing problems with commercial potential have been enumerated. Then they were divided into two groups: Simulation and Data Processing. This division significantly differentiates tasks based on the amount of data that should be transmitted into computation nodes. Data Processing: • Image processing - medical image segmentation (microscope, MRI, CT, 2D and 3D); • Text processing (NLP); • Video analysis (trans-coding, theft detection, customer classification); • Machine Learning - various applications.		
Related Stakeholders	Emerging Technology Company, Technical/Medical University, General Company, …		
Current solution	In-house computations for smaller companies, using large HPC centres for big companies		
Competition	Using AWS (requires programming knowledge) Building an own cluster (big investments, resources not always needed)		

Short name	**Security for sensitive data**		
Id		Priority	6
Description	Some end users operate with sensitive data (IP, personal data, etc.). The problem for using LSC service is how to secure the data is remaining safe and secure. Utilise clear security policy and high security standards.		
Related Stakeholders	End-users		
Current solution	Own servers closed to internet Manual transfer of data to HPC service providers		
Competition	Local HPC service providers		

Short name	**Forecasting calculation time/cost**		
Id		Priority	3
Description	End-users are sensitive to costs and as LSC service resource utilization is hard to forecast, ensuring approximations and providing opportunity for testing with smaller data set is of key importance. Potentially this could be overcome with individual consultancy and customization, but this would significantly add to the cost of the service.		
Related Stakeholders	End-users		
Current solution	Consultancy in HPC centers of Universities for HPC problems		
Competition	Local HPC service providers		

Short name	**The HPC and storage capacity available is more or less just access and might be a bit too advanced for SMEs without technical expertise**		
Id	Vejle_PR1	Priority	4
Description	There is a need for guidance on to how to use the acquired time/access to LSC the most. Maybe a supervisor or the like that guide you through the process – and assist you on knowing what is doable? Cases that can serve as inspiration – how do you compare task and computing power…		
Related Stakeholders	SMEs – e.g. Givskud Zoo, start-ups		
Current solution	In-house capacity – not LSC		
Competition	Amazon AWS (but requires programming knowledge)		

Short name	**Students and professors developing modules/applications that could benefit companies in the "real world" but is stored "away" in universities.**		
Id	Vejle_PR2	Priority	4
Description	Could be more universities than Aarhus University, Department of BD and Technology that could provide free of charge modules/applications and perhaps even services in return for feedback on performance and usage and citation.		
Related Stakeholders	Universities, end-users		
Current solution	They keep the modules/applications stored locally in the university		
Competition	Other service providers (however these will often charge for theirs)		

4. Product features

The workshops held at BalticLSC partners' premises with local stakeholders provided a lot of input regarding needed features for the BalticLSC Platform.

Firstly, the highest priority should be given to the functionality of the system – what kind of problems it can solve. Below are a set of solutions developed with predetermined models that represent reproducible use cases. Identifying and defining these cases, translating them to be accessed and modified virtually is the key issue during the initial phase of setting up the BalticLSC environment.

Another related issue refers to selection of replicable use cases that have the greatest commercial potential. Which means that the market should be analysed and cases that are currently utilised should be the starting point.

Finally, one of the most frequently mentioned issue is ensuring security and reliability of the service, so checking and applying the appropriate standards is one of the main challenges for the development. The problem is, it was not disclosed during the workshops, what kind of standards that entails, as there was an opinion expressed that such standards are not widely established in the field of supercomputing/ HPC. On the other hand, these standards may apply more for the application and network parts rather than for other operational layers, therefore common standards are relevant and must be implemented.

One key regulation to be followed on all levels is the EU General Data Protection Regulation (GDPR). Relevant measures should be taken to protect the personal data on all Baltic LSC user levels as well as protecting the end-user data during different LSC operations.

In terms of reliability, it is important to ensure uninterrupted operations day and night. There must be external staff approachable to solve issues, such as debugging. Resources available must be made transparent and their allocation should be possible across different service providers.

Regarding quality, it is important to ensure a robust performance that can be sustained 24/7.

The product features have been divided into functional and non-functional.

4.1 Functional Product Features of the Baltic LSC Platform

Short name	**Computation Application Management**		
Id	WUT_F1	Type	Functional
Description	System should allow for management of computation applications, including support for the application store, creating computation tasks from applications, setting schedule for task execution, monitoring tasks progress, interacting with tasks and aborting tasks.		
Required resources	BalticLSC Network composed of nodes of various size		

Short name	**Computation Application Development**		
Id	WUT_F2	Type	Functional
Description	System should support functions associated with computation applications development including creating and editing computation applications and computational modules, testing them and approving before placing them in the store.		
Required resources	BalticLSC Network composed of nodes of various size		

Short name	**Computation Resource Management**		
Id	WUT_F3	Type	Functional
Description	System should provide computation resource management for the resource owners, including creating them, submitting them to be used by other users, checking their status and parameters and safely deactivating them.		
Required resources	BalticLSC Network composed of nodes of various size		

Short name	**Computation Resource Supervision**		
Id	WUT_F4	Type	Functional
Description	System should provide functions for supervising computational resources like checking their statuses and parameters, approving them to use for all users, conducting benchmarks or deactivating them.		
Required resources	BalticLSC Network composed of nodes of various size		

Short name	**Computation Application Requests**		
Id	WUT_F5	Type	Functional
Description	System should allow for submitting requests for new applications and cosigning them by other users.		
Required resources	BalticLSC Network composed of nodes of various size		

Short name	**Computation execution**		
Id	IMCS_F1	Type	Functional
Description	System should allow execution of various computations (in a broad sense). Computation consumes input data and parameters and provides output data. System should provide the specific computation resources in appropriate amount.		
Required resources	BalticLSC Network composed of nodes of various size		

Short name	**Computation definition**		
Id	IMCS_F2	Type	Functional
Description	System should allow the definition of computation. There should be an easy way to define a computation which is based on the existing software.		
Required resources	Easily understandable UI, with integrated pre-determined computational platforms		

Short name	**Computation integration**		
Id	IMCS_F3	Type	Functional
Description	System should allow the execution of computations from the external systems – non-human users.		
Required resources	Possibility of integrating AI based computations and a possibility to use VPN like systems		

Short name	**Computation transparency**		
Id	IMCS_F4	Type	Functional
Description	System should provide precise details on what resources and how long have been used.		
Required resources	Status counter that integrates all needed information (CPU / GPU usages, core hours, memory usage, storage)		

Short name	**Computation monitoring**		
Id	IMCS_F7	Type	Functional
Description	System should provide real-time data on the state of computation execution. This may include computation-specific data.		
Required resources	Computation owner's UI shows progress and others necessary parameters		

Short name	**Computation billing**		
Id	IMCS_F8	Type	Functional
Description	System should provide billing of resources used.		
Required resources	UI has integrated billing system, which shows user's paid balance compared to usage		

Short name	**Predefined models**		
Id	LIC_F1	Type	Functional
Description	The system should offer a number of predetermined templates or cases that are easily modified. Their descriptions must be well-written.		
Required resources	Use case descriptions converted to sets of algorithms		

Short name	**Expert support form**		
Id	LIC_F2	Type	Functional
Description	Users should be able to describe tasks and get in touch with system administrators (or consultants) that would help to reformulate and adapt tasks for the system, if needed		
Required resources	Online form for contacting administrator		

Short name	**Transparency: Real-time monitoring**		
Id	LIC_F3	Type	Functional
Description	A user should know where and what calculations will take place (at which points in time).		
Required resources	Resource allocation prediction and/or recommendation		

Short name	**Enabled reservation**		
Id	LIC_F4	Type	Functional
Description	The resources could be reserved through the system for a time window in the future (even if infrastructure remains idle)		
Required resources	Information regarding resource availability; universal scheduling system		

Short name	**Automatic calculation/estimation of costs**		
Id	LIC_F5	Type	Functional
Description	In best case scenario, calculation costs are known in advance or at least there is an indication of expected costs		
Required resources	Costs calculation formula (during the meeting in Hamburg a few different billing models were proposed; a hybrid model can also be considered);		

Short name	**Task execution control**		
Id	LIC_F6	Type	Functional
Description	Being able to stop calculations midway or perform them in stages (an option)		
Required resources	Administrator rights		

Short name	**Automated debugging**		
Id	LIC_F7	Type	Functional
Description	The system is able to find and fix common mistakes		
Required resources	Testing of algorithms beforehand; code fixing in real-time		

Short name	**Informational resources**		
Id	LIC_F8	Type	Functional
Description	The system should also contain informational resources and relevant documentation as well as FAQ		
Required resources	All relevant documentation; FAQ; online form for contacting administrator; training material with examples, step-by-step guides, etc.		

Short name	**Batch processes – ability to run both parallel and series**		
Id	Vejle_F1	Type	Functional
Description	The ability to run both series and parallel batch processes.		
Required resources	Hardware/software (e.g. apps on BalticLSC Platform) available supporting running of series and parallel batch processes.		

Short name	**Coding language – multi input possibility**		
Id	Vejle_F2	Type	Functional
Description	Start-ups developing modules need the possibility to use open source, non-expensive coding language like Python as this is accessible for them – thus, the solution should allow multi-input for coding language.		
Required resources	Hardware/software (e.g. apps on BalticLSC Platform) available supporting multi-input for coding language.		

Short name	**Computation integration**		
Id	Vejle_F3	Type	Functional
Description	Being able to integrate with data accessible on different platforms like Copernicus (satellite data etc.)		
Required resources	Hardware/software (e.g. apps on BalticLSC Platform) available supporting data from different platforms.		

4.2 Non- Functional Product Features of the Baltic LSC Platform

Short name	**Integration with External Systems**		
Id	WUT_NF1	Type	Non-functional
Description	It is important that the user can use the environment (s)he already knows. Therefore, the computing system should be able to accept orders from other applications via plug-ins. Integration with external systems through a BalticLSC API. Some calculations can be performed locally, and others commissioned to the BalticLSC system. For example, for visualization or data confidentiality reasons. Cooperation between external (micro-)services could be possible.		
Required resources	External systems provided by their supervisors		

Short name	**"Fire and Forget"**		
Id	WUT_NF2	Type	Non-functional
Description	"Fire and forget" – the end-users prefer to start an app and forget about the rest until the computation results are available.		
Required resources	BalticLSC Network composed of nodes of various size		

Short name	**Computation Packages**		
Id	WUT_NF3	Type	Non-functional
Description	Grouping of Computation Modules into "Computation Packages" – such packages must be executed on single Computation Resources.		
Required resources	BalticLSC Network composed of computation nodes with higher internal throughput		

Short name	**Distance from Data**		
Id	WUT_NF4	Type	Non-functional
Description	"Strength" of data connections between Computation Modules and Data Sources. "Strong" connection means that the connected modules should be "close" when computation is made (e.g. on a single Computation Resource). Parallelization dependent on physical distribution of computations (optimization of overhead for data transmission).		
Required resources	BalticLSC Network composed of nodes of various size; access to external data stores of various size and throughput		

Short name	**Crash Handling**		
Id	WUT_NF5	Type	Non-functional
Description	The system has to handle Computational Module crashes/freezing. Errors should be expected situations.		
Required resources	BalticLSC Network composed of nodes of various size		

Short name	**Action Logging**		
Id	WUT_NF6	Type	Non-functional
Description	Don't forget about action logging (for debugging, billing, code profiling, fraud detection).		
Required resources	BalticLSC Network composed of nodes of various size		

Short name	**Application Flow**		
Id	WUT_NF7	Type	Non-functional
Description	Computational application is a combination of Data Flow and Control Flow. Thus, the language used to create applications has to define Data Sources and Actions + Flow Blocks. - A Data Source contains: reference to the data location, data format, other metadata; - An Action converts input into output - Pause is a special 'empty' action; - Conversion Actions are to transform data between different formats. - Number and type of CMs executed in a given Stage depends on previous computations. Advanced conditional processing (conditional modules, conditional expressions?).		
Required resources	BalticLSC Network composed of nodes of various size		

Short name	**Computational Gauging**		
Id	WUT_NF7	Type	Non-functional
Description	Evaluation of computation resources (CPU, GPU, RAM, etc.) – translate into a common unit. Standard/custom "exchange rate" for various resources.		
Required resources	BalticLSC Network composed of nodes of various size		

Short name	**Computational Market Exchange**		
Id	WUT_NF8	Type	Non-functional
Description	Pekao SA expertise could be used to solve billing or market trading problems. Advertisement exchange market mechanisms could be used for calculating optimal pricing for computations (computations are performed on nodes that offer best value for money).		
Required resources	BalticLSC Network composed of nodes of various size		

Short name	**Computation security**		
Id	IMCS_F5	Type	Non-Functional
Description	System should assure the privacy of the input and output data and of the computation itself. The access to the data is granted to the computation executor, but access to the computation execution metadata is granted also to the resource owner in the context of the resource usage.		
Required resources	Encrypted communication between computation node and external user		

Short name	**Computation reliability**		
Id	IMCS_F6	Type	Non-Functional
Description	System should assure the availability of the provided resources while computation executes. System should assure that the provided resources perform accordingly to their specification.		
Required resources	Possibility to enter computation need and match it with available resources		

Short name	**Easy access**		
Id	LIC_NF1	Type	Non-Functional
Description	Making sure that users can access the system online and that models have their own adapted environments (like popular modelling suites)		

Required resources	Single point of contact; entry through a website; a convenient user interface (a need for GUI is arguable, although adopting visual programming would partially solve that)

Additionally, initial ideas for User Interface have been provided by WUT, including computation cockpit and app store:

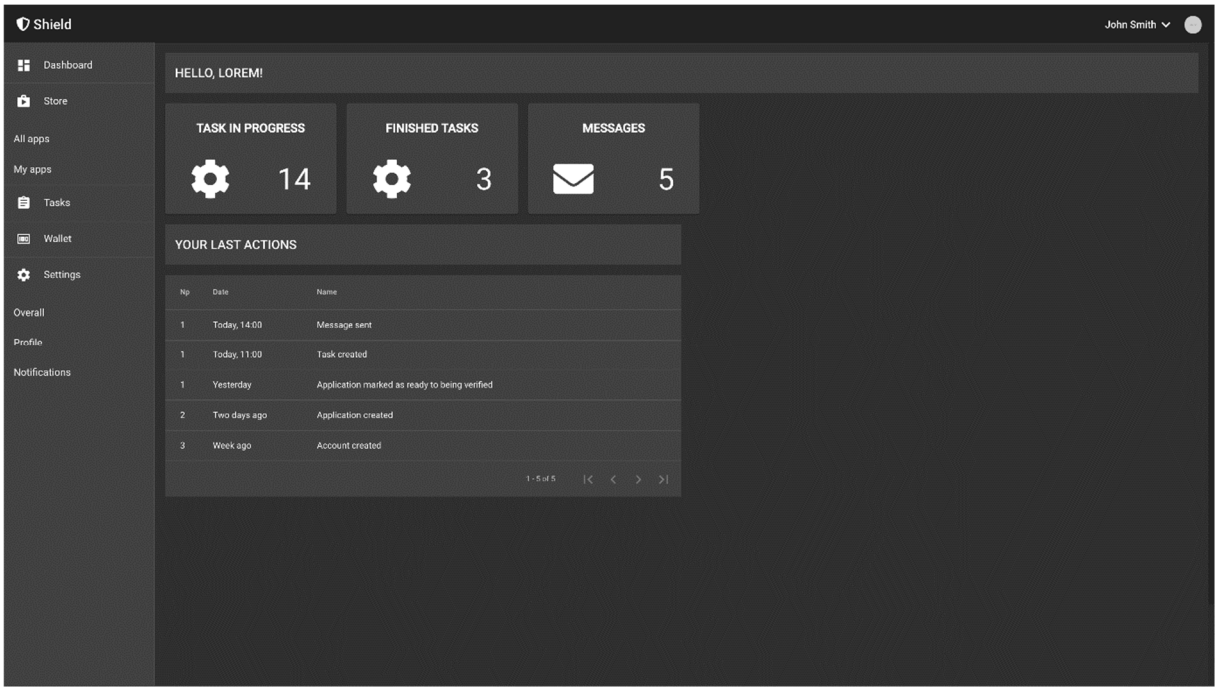

Figure 1 Computation cockpit of BalticLSC Environment

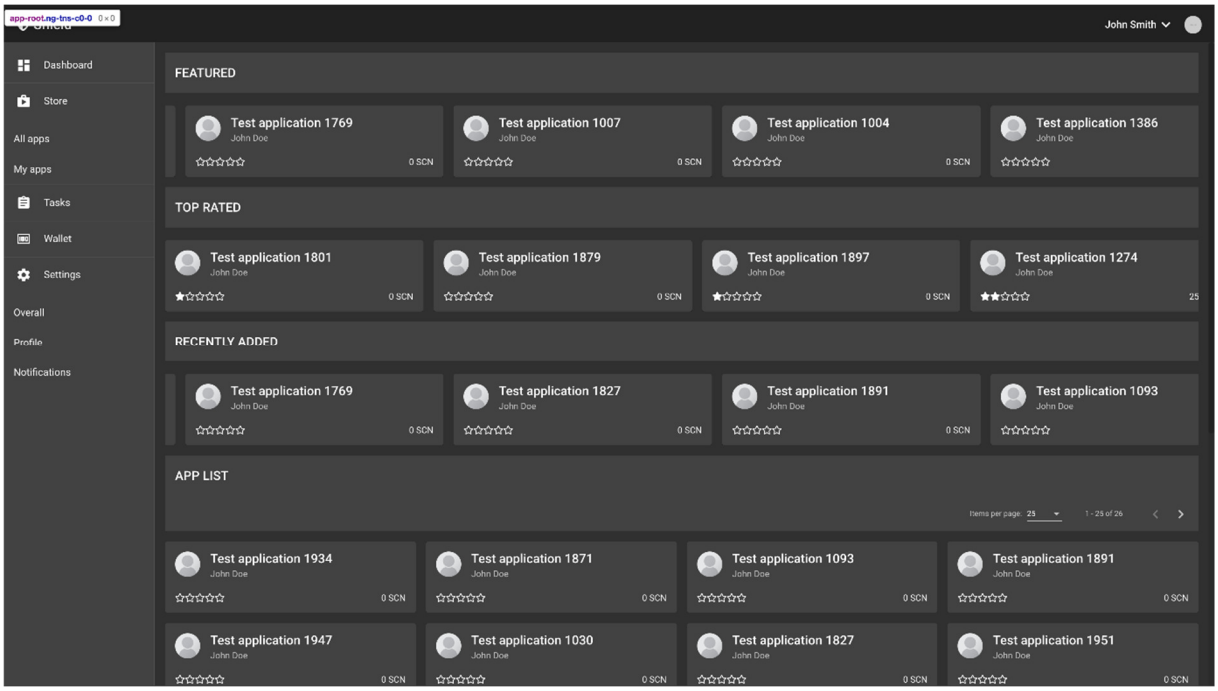

Figure 2 App Store UI of Baltic LSC Environment

5. Case study proposals

During desk research activities, qualitative interviews, and Baltic LSC workshops a number of case studies have been described. Additionally, several SMEs and experts have provided insights into currently ongoing activities in their respective fields, including several prototypes as potential case studies for the Baltic LSC Environment. These would be available for testing as soon as Baltic LSC service will be operational. The case studies are proposed from a variety of sectors, so significant effort is needed to test, analyse and choose the ones with highest potential in the future.

Name	**Video processing framework**
Author	Agris Šostaks (IMCS, Latvia)
Description	System to process video, as a sequence of images. Segmentation, recognition, filtering.
Contact	Ask Baltic LSC partner representative for direct contacts

Name	**Digital Weather Modeling**
Author	Lithuanian Hydrometeorological Service
Description	Possible cooperation on testing weather modeling. They have access to digital weather modeling system-model HARMONIE (installed on a supercomputer SGI ICE X).
Contact	Ask Baltic LSC partner representative for direct contacts

Name	**Molecular modeling**
Authors	Golem (Poland), Vilnius University Faculty of Physics (VU FF); "HPC Saulėtekis" OAC (Lithuania)
Description	Molecular modeling is a computer technique to mimic structure and interactions between 3D molecules. The tool has wide applications from computational chemistry, drug design, computational biology and materials science. The complexity relates to high number of cases that are typically analyzed. Potential services: modelling and simulation of complex molecules, molecular mechanics/dynamics, quantum chemistry; modelling of organic, metal-organic and inorganic compounds and analysis of their properties with quantum chemistry methods.
Contacts	Ask Baltic LSC partner representative for direct contacts

Name	**Fraud detection and taxation in banking**
Author	(Pekao SA, Poland), Monese (Estonia)
Description	Artificial Intelligence algorithms could be used to analyze bank data (transfers, deposits, withdrawals) to identify and highlight possible frauds. Potential approach: test different modelling strategies, giving value to data and comparison of models – static learning (GLM, XGBoost, Multilayer perceptron (MLP)) vs learning of sequences, RF vs LSTM. Graph learning (GraphSAGE). Combine models.
Contact	Ask Baltic LSC partner representative for direct contacts

Name	**Theory of algorithmics**
Author	Krzysztof Hryniów (WUT)
Description	Theory of optimal control problems solved by the parallel computational environment
Contact	Ask Baltic LSC partner representative for direct contacts

Name	Optimal dose for radiation
Author	Jacek Starzyński (WUT)
Description	The problem is with medical origins. Task is to choose optimal pattern of X-ray radiation for cancer treatment.
Contact	Ask Baltic LSC partner representative for direct contacts

Name	Analysis of satellite data with artificial intelligence and deep learning algorithms
Author	Geomatix (Lithuania), Kappazeta OÜ (Estonia), DataDuctus (Sweden)
Description	Earth Observation satellite data processing (data available for free by European Space Agency). Analysis of Earth observation images for measurements of grasslands cutting in time, agricultural gains with AI and DL. As well as for the assessment of various air, water and soil properties, environment monitoring, forest and environment changes, to decrease for example risk for bark beetle diseases, etc.
Contact	Ask Baltic LSC partner representative for direct contacts

Name	GPS Location calculations
Author	Precision Navigation Systems OÜ (Estonia)
Description	Task is to run software algorithm that takes any positioning signal and combines it with ground correction stations and company's own reference stations
Contact	Ask Baltic LSC partner representative for direct contacts

Name	Media multi-language processing: video transcript, translation, and voiceover.
Author	Agris Šostaks, IMCS and SIA Mindflux (Latvia)
Description	Media multi-language processing: video transcript, translation, and the voiceover is a task when a video is automatically transcribed (audio to text), translated (text to text) and voiceover (text to audio). SIA Mindflux has its own software frontend, but it uses three external services (from different providers) to perform the mentioned tasks. Their own software is used to orchestrate the computations. Thus, Mindflux buys both resources and algorithms as services to perform their tasks.
Contact	Ask Baltic LSC partner representative for direct contacts

Name	Video processing: person's appearance detection.
Author	Agris Šostaks, IMCS from SIA *Noscos (Latvia)*
Description	Video Processing: Person's Appearance is a task when one should recognize a known person(s) in a video and find time-frames when the person(s) has appeared. SIA Noscos possesses its own software which is used to store, edit (cut, merge, enhance, crop, etc…), process, post-process and publish videos. The resource-consuming step is the video processing when a video is processed frame-by-frame and the detection of a known person is done. This and only this step can be delegated to the BalticLSC Network. The reason to delegate is getting access to the appropriate amount of computing resources.
Contact	Ask Baltic LSC partner representative for direct contacts

Name	**Logistics: route planning and resource optimisation.**
Author	Agris Šostaks, IMCS from SIA Smart Routing Services (Latvia); VU MIF and an undisclosed company from Lithuania
Description	Logistics: route planning is a task when one should determine a sequence in which a fleet of vehicles should visit the clients in order to satisfy their demand and do it in an efficient way. The task involves different set-ups and various constraints (CVRP, VRPTW, etc. …). Another case in logistics involves tracking movements of rental cars. A growing number of start-ups are exploring tracking applications in transport. SIA Smart Routing Services (SRS) plans to provide API for solving various planning tasks. The first use-case is route-planning for garbage trucks for small and medium waste collection enterprises. SRS uses an external service to determine point-to-point routes (because SRS has no detailed road topology data) and rents resources for other tasks (because SRS has no owned computation resources). SRS has its own software to compute and visualize routes. The most computation-intensive task is route planning.
Contact	Ask Baltic LSC partner representative for direct contacts

Name	**Increasing efficiency for engineering/production sector**
Author	Protolab OÜ (Estonia)
Description	During the preparation of CAM programs for the CNC machines, the calculation and optimisation of toolpaths for details with higher complexity might take several hours and even days. Increasing this calculation would siginificantly decrease engineering and production times for engineering industry.
Contact	Ask Baltic LSC partner representative for direct contacts

Name	**Design optimization**
Author	LIC (Lithuania)
Description	Calculations for optimizations in construction and/or engineering, e.g. optimize material use for the manufacturing of different products of, e.g. plastic bottles.
Contact	Ask Baltic LSC partner representative for direct contacts

Name	**Image analysis in Medicine**
Author	LIC (Lithuania)
Description	Analysis of digital images. Analytical solutions for the identification of cancerous lesions in digital photos. Analysis of X-ray images, e.g. determination of bone age from X-ray images of children limbs. A start-up Oxipit also works in this area and performs automated processing of medical X-ray images with AI (has developed a tool for computerized lung scanning).
Contact	Ask Baltic LSC partner representative for direct contacts

Name	**Minimizing downtime in process industry**
Author	SSAB, Ferruform from Sweden
Description	In process industry downtime scenarios of machines can be analysed.
Contact	Ask Baltic LSC Swedish partner representative for direct contacts

Name	**Risk reduction in work environment**
Author	Samoosa from Sweden
Description	Modelling of Work Environment data to improve the work environment and reduce risks.
Contact	Ask Baltic LSC Swedish partner representative for direct contacts

Name	**Risk reduction in mining operations**
Author	Mobilaris from Sweden
Description	Modelling of peoples locations in case of fire, to rescue lives. Modelling of intelligent decision support for increased safety and productivity in underground mines and tunnel constructions.
Contact	Ask Baltic LSC Swedish partner representative for direct contacts

Name	**Data analysis**
Author	TokenMill (Lithuania)
Description	Data analysis employing various algorithms (incl. artificial intelligence, machine learning, deep learning, etc.; Internet of Things (IoT)). A start-up TokenMill performs textual data analysis (natural language processing and generation) and perhaps could be considered for the formulation of a case.
Contact	Ask Baltic LSC Lithuanian partner representative for direct contacts

Name	**3D rendering and virtual reality**
Author	Gluk Media, Ltd. (Lithuania)
Description	A 3D rendering and/or virtual reality case in which videos acquired with multiple camer-as must be combined in one and such a rendering takes a month to perform.
Contact	Ask Baltic LSC Lithuanian partner representative for direct contacts

Name	**Transcoding of the internet video archive**
Author	VU FF; VU FF in collaboration with National Radio and Television (LRT) (Lithuania)
Description	Transcoding of the internet video archive
Contact	Ask Baltic LSC Lithuanian partner representative for direct contacts

Name	**Weather forecasting and other readymade modules/applications supplied by external**
Author	Aarhus University, Department of Business Dev. and Technology & Municipality of Vejle
Description	The University of Aarhus has itself supercomputing facilities; however, if a unit like the Department of Business Development and Technology (DBVT) can access LSC cheaper or faster elsewhere, it is entitled to do so. The Department of Business Development and Technology has strong competences within wind forecasting, storage etc. and is developing models that often just ends up in their own database and are only being used by academia. These are based on open source and could be useful for private companies to make

	use of. The Department of Business Development and Technology would like to contribute to the BalticLSC project with the possibility of using these models/applications without any fee but requesting feedback though of how the models/applications work in "real life" (the usage and performance). Furthermore, citation is needed. DBVT is producing a lot of start-ups as the students have the right combination of technology and business development knowledge to be suitable for starting innovative companies. Also, students can prolong their education with one year and start up own company in this period e.g. with focus on machine learning. However, access to LSC is a barrier with regards to the expenditures and an introduction price for students or newly registered start-ups could be an attractive business opportunity in the BalticLSC project. Aarhus University, the Department of Business Development and Technology has expressed interest in participating as pilot/use case in the BalticLSC project – and probably both as a service provider and end-user.
Contact	Ask Baltic LSC partner representative for direct contacts

Name	**Vibrations data processing using AI, requiring LSC**
Author	**Vertikal AI** & Municipality of Vejle
Description	Vertikal AI is a company that uses vibrations data from generator, gearboxes etc. on wind turbines and analyzing them with AI which enables them to predict damages 3 years before they occur.
Contact	Ask Baltic LSC partner representative for direct contacts

Name	**People tracking and analyzing (sensor & picture) data to enhance customer experience and optimal use of resources in animal park/tourism attraction**
Author	**Givskud Zoo &** Municipality of Vejle
Description	The tourism industry handles a lot of data and could enhance their businesses with the use of AI and LSC/HPC. The Danish animal park, Givskud, has the ambition to enhance the user experience and use resources best possible by using their data smarter. They got 35.000 season passes and all the day to day visitors and this requires some knowledge and access to computing power to use the acquired data. Furthermore, the tracking of people' and animals' movements in the park could provide valuable information as well. However, Givskud has not the in-house computing power nor the expertise to make properly use of their data and has expressed interest in becoming a pilot case/use case in the BalticLSC project to develop the skills to make use of their data and become a super-user of BalticLSC for future use when the project expires.
Contact	Ask Baltic LSC partner representative for direct contacts

6. Conclusion

The current BalticLSC Environment Vision Report is a compilation of input from 8 Baltic LSC project partners during the first 6 months of the project, starting from January 1st 2019. Following the initial contacts with local experts in LSC/HPC (incl. innovative SMEs, R&D centres, et al.) and desk research of current status at participating countries, a series of workshops were held to gather input from a wide range of stakeholders involved in setting up the Baltic LSC Environment.

The activities were focused on collecting information from the target groups regarding Visions for:

- BalticLSC Environment, including functional and non-functional features of the BalticLSC Environment as well as potential case-studies and initial applications
- BalticLSC Platform (hardware and operating systems), including a wide range of technical possibilities to launch the network with flexibility to include additional service providers in the future
- BalticLSC Software (administrative and computation tools). This is an important basis for the software development activities during the most part of the project

The main objective of this document is to provide an overview of business and technology needs of the BalticLSC Environment's potential end-users and to ensure that these needs are successfully met. It also includes a general vision for the functional and non-functional features of the BalticLSC Environment.

During the following months, the project team will turn this general vision into detailed user requirements specifications and input for programmers. This will be done separately for the Baltic LSC Platform (hardware and operating systems) and BalticLSC Software (administrative and computation tools).

Additionally, training programs and promotional events will be held after launching the BalticLSC Environment, to test different technological aspects, receive feedback from stakeholders and fine-tune the service. As a result, new supercomputing solutions, developed for SMEs will be piloted, jointly tested, transferred or adapted into day-by-day business.

Finally, different business models will be tested for the platform to offer a self-sustainable service to target groups in the future, fostering innovation in the area of Large Scale Computing leading into the strong community of BalticLSC developers and engineers resulting in investments of companies in new products and services in the field.