

# ETP4HPC – BDVA cPPP's positioning

## 1 The context: development of the data economy

It is widely accepted that we are at the beginning of a new industrial revolution. This new revolution has many names: some call it the digitisation of industry, others the 'Internet of Everything', or 'the Sensor Economy'. All these refer to the use of sensors to collect data, and the use of computing power to analyze the data and make useful judgments from it.

The data economy, drawing on data provided by the Internet of Things, and other sources, will transform many aspects of our lives and businesses. Gartner forecasts that by 2020, 25 billion "things" will be connected, with each producing data for analysis in various ways. This will include data from infrastructures, data from cities, health data, logistics data, performance data, and agricultural data and so on. This new economy will have the potential to improve the way we use resources, and deliver services. Consumers, citizens, businesses and governments will all benefit.

One of Europe's main challenges in the coming years will be to seize all the opportunities of the digital economy. This will be essential for the competitiveness of its economy, for maintaining its position in the world and for the well-being of its citizens. One dimension of this challenge is a significant increase of the IT deployment by the European industry. Without a strong action towards the digitalisation of the industry, existing (and future!) industries could become less competitiveness and be leapfrogged by new industrial regions, even within the domestic European markets. The second dimension is related to the development of the new data economy. Europe must be at the forefront of the creation of the data economy that will generate new areas of activities and will also support the digitalised industry.

In order to meet both dimensions of the above challenge, Europe must identify the strategic assets to develop and the levels of the value chain that need to be controlled in order for the continent to be in the driving seat. Looking at these priorities, it appears that having cutting edge IT infrastructures in Europe and the ability to deal with big data are of the utmost importance. Without excellent computing infrastructures located in Europe and the right skills to extract value from data, we will be unable to successfully develop the new digitalised industry and the data economy that go with it. The data will flow to and be stored where the best computing infrastructures and skills are located. So without excellent, secure infrastructures and the ability to extract value from the data, European data will be exploited elsewhere and will thus generate value outside Europe. The successful development of this new industry also requires the use of simulation and the processing of large amounts of data. The existence of high-performance computing (HPC) infrastructures and expertise in big data are of essential importance to the modernisation and the continued competitiveness of industry.

In this context, the BDVA (Big Data Value Association) and the ETP4HPC (European Technology Platform for High-Performance Computing) will work together in order to strengthen Europe's position in these two domains and will interact to exploit all the synergies between them. In line with the EC priority of a connected Digital Single Market, they will develop Europe's capacity to seize the opportunities offered by the data economy. This paper briefly presents the way the two organisations plan to cooperate.

# ETP4HPC – BDVA cPPPs positioning

## 2 Position of the associations

### 2.1 BDVA main goal

The Big Data Value Association (BDVA) is an industry-led contractual counterpart to the European Commission for the implementation of the Big Data Value PPP cPPP. As of Dec 2015,, the BDVA has over 120 members including large and SME industry together with research institutions and academia. The Big Data Value PPP is a partnership between the European Commission and the BDVA which aims to strengthen the data value chain, cooperate in data research and innovation, enhance community building around data and set the grounds for a thriving data-driven economy in Europe. The BDV cPPP is driven by the conviction that research and innovation, focusing on a combination of business and usage needs, is the best long-term strategy to deliver value from Big Data and create jobs and prosperity.

The BDV PPP activities will address technology and applications development, business model discovery, ecosystem validation, skills profiling, regulatory and IPR environment and social aspects. The EU has earmarked over €500 million of investment over 5 years (2016-2020) from Horizon 2020. The BDV PPP will lead to a comprehensive innovation ecosystem for achieving and sustaining European leadership on Big Data, and for delivering maximum economic and societal benefit to Europe – its business and its citizens.

### 2.2 ETP4HPC main goal

ETP4HPC is an industry-led think tank and advisory group made up of companies and research centres involved in High-Performance Computing (i.e. Supercomputing) technology research in Europe. It was formed in 2011 with the aim to build a world-class HPC Technology Supply Chain in Europe, increase the global share of European HPC and HPC technology vendors as well as maximising the benefits that HPC technology brings to the European HPC user community. ETP4HPC is also the EC's partner in the HPC contractual Public-Private Partnership (cPPP<sup>1</sup>) which helps to define, monitor and manage the European HPC research investment programme supported by a €700M investment by the EC within the Horizon 2020 programme. As of end of 2015 ETP4HPC has 70 members from industry and research; 35% of the total number of members are SMEs.

ETP4HPC believes that a competitive European HPC technology value chain will help Europe achieve economic leadership through the use of HPC for product creation and service delivery, for the development and supply of world-class technologies and sub-systems, and for the supply, and administration of HPC systems themselves.

The strategy adopted by the European Commission in 2012<sup>2</sup> highlights the need for intensified HPC technology provision in Europe and calls for an integrated and concerted effort by the entire European

---

<sup>1</sup> <https://ec.europa.eu/digital-agenda/en/high-performance-computing-contractual-public-private-partnership-hpc-cppp>

<sup>2</sup> European Commission, 'High-Performance Computing: Europe's place in a Global Race', 15 Feb 2012

# ETP4HPC – BDVA cPPPs positioning

HPC industry. HPC is also one of the pillars of the Digital Single Market (DSM)<sup>3</sup> strategy adopted by the EC in 2015, which includes Big Data and Cloud and aims to build a basis for the long-term growth of the European ‘digital’ economy.

## 3 Cooperation axis

The cooperation between ETP4HPC and BDVA has two dimensions: the basis for any kind of positioning of their relative focus areas is an in depth understanding of the two technical roadmaps with their overlapping sections, interdependencies and complementary aspects. Then, on top of this base understanding, a set of interactions, joint events, supportive actions, workshops and conferences can be defined. Cooperation between the two organisations is based on a mutual understanding of the corresponding technical roadmaps especially interdependencies, complementary aspects and possible areas of overlap. From this understanding joint actions can be defined, including joint: events, workshops, training and conferences.

### 3.1 Technical topics

#### 3.1.1 Technology roadmap interaction

BDVA and ETP4HPC have already held two working sessions with technical representatives of both parties, focusing on defining common ground for their mutual technical roadmap. The base understanding is that HPC technology is the underpinning compute infrastructure for High-Performance Data Analytics.

BDVA has a multitude of non-technical challenges (e.g. addressing legal and social issues, building a business model). Development of the compute infrastructure is a long-term element and HPDA is now using the currently available infrastructure, and the process of optimisation for high demand use cases is only starting now.

The increasing use of HPDA will impose additional, in some cases much stricter, requirements on the design of future generations of systems. Although there are no immediate bottlenecks with current HPC design points identified, today, the Big Data community is using whatever IT infrastructure is available to them. HPC technology providers, ISVs and application software vendors need to prepare for the handling of massive amounts of extremely diverse types of data coming from a multitude of sources in the next 5-10 years. The HPC ecosystem needs to be ready for this change driven by Big Data.

BDVA and ETP4HPC have agreed to synchronise BDVA’s Strategic Research and Innovation Agenda (SRIA) and ETP4HPC’s SRA using a top-down approach by analysing highly demanding Extreme Data use cases and extracting HPC research priorities.

---

<sup>3</sup> European Commission, ‘A Digital Single Market Strategy for Europe’, 6 May 2015

# ETP4HPC – BDVA cPPPs positioning

## 3.1.2 Development of HPC usage for extreme data exploitation

HPC systems are now dealing with extreme data situations with increasing frequency. This is due to two main reasons:

Traditional HPC simulations (e.g. fluid dynamics, computational chemistry and physics, cosmology among others) have been benefitting from the tremendous increases in computational capacity of HPC systems and are now using models of unprecedented realism and accuracy. These models represent the real world using trillions of discrete degrees of freedom, which require huge memory and scale out systems. On the other hand, researchers need to apply advanced and highly complex analytics and processing (including visualisation) on this data, which simply means that off-loading this data to remote platforms is simply not an option. Thus, data analytics needs to take place in-situ, and perhaps in synchrony with tightly coupled synergistic computing platforms (e.g. visualisation engines).

New applications are becoming potential HPC clients. Big Data applications, in which data is not generated from a model but rather is collected, accumulated or even streamed, and has a complexity that means its computational needs are at the petascale or even the exascale, even after local pre-processing. Big Data systems, that have been primarily developed for scale out to distributed, non-reliable resources, are simply too coarse in efficiency and cost effective to cope with such computational complexity. Thus, HPC solutions, with lots of memory and very fast networks are starting to be very appealing and are the next natural step for Big Data users. Indeed, Big Data systems have already started to be influenced by HPC architectures and practices (e.g. multi-threading, parallelism, memory modelling etc.), and it is also clear that data will often be highly distributed as it originates from distributed sources. Thus Big Data systems and tightly coupled HPC systems may form larger, hybrid, computing resources that combine local and global processing to serve the needs of new Big Data applications.

## 3.1.3 Democratisation of HPC for data intensive applications

Progress in HPC technologies is also creating new opportunities for data intensive applications. Efforts to develop exascale computing systems will produce top-end supercomputers and will also drive the cost of mid-size systems down thus opening up the use of these systems to a broader range of applications. Once we have 'Exascale systems consuming 20 MW' we will also have 'Petaflops systems consuming 20KW', costing only a few hundred thousand Euros. This level of performance/Watt could be very interesting for some big data applications and the potential Return on Investment could justify dedicating certain systems to only one application.

So we foresee the development of the usage of mid-size HPC systems dedicated to big data applications. The democratisation of HPC systems will be a driving force for the development of big data in industry, the city, medicine and agriculture. BDVA and ETP4HPC will work together to accelerate this trend and to exploit this opportunity.

# ETP4HPC – BDVA cPPPs positioning

## 3.2 Interaction

### 3.2.1 Exchange between the two associations

The BDVA and ETP4HPC have agreed to collaborate on a number of strategic issues:

- Align their SRIAs to highlight the complementary nature of their technical roadmaps
- Provide an update to the relevant governance body of each association at least once a year
- Actively participate in each other's events (e.g. BDVA Summit, HPC Summit) including workshops to focus on stakeholder interactions
- Appoint liaisons to support interaction and to increase cooperation
- Annual Technical interlock workshop between both associations

### 3.2.2 Common discussion with EC

Both associations are willing to support an interlock with the Commission, if requested.

### 3.2.3 Joint communication on the importance of data and computing

BDVA and HPC will issue a joint communication on the potential benefits of data and computing for Europe. The communication will target decision makers and policy makers and will set out a vision for the role of data and computing in the context of the Digital Single Market and relevant Societal Challenges. It will provide a clear understanding of how both PPPs are driving change and it will emphasise the impacts of both PPPs needed for Europe to establish and maintain a leadership position in both of these areas. The communication will detail the key synergies between the technical and non-technical agendas of the two PPPs and how they both contribute to realise the vision. The alignment of the strategic agendas will be detailed in the communication together with clear actions for common activities between the associations and their stakeholder communities, possible collaboration opportunities between funded projects of both PPP, and identification of possible joint calls in future Work Programmes.

## 4 Conclusion

Big Data and High Performance Computing are both pivotal means and technologies to boost the prosperity of societies in Europe and worldwide. By establishing a tight cooperation and coordination the two associations BDVA and ETP4HPC are committed to accelerating the growth of Europe's digital society with a global vision of data and computing.