1 Publishable Summary

The objective of the RETHINK big Project is to bring together the key European hardware, networking, and system architects with the key producers and consumers of Big Data to identify the industry coordination points that will maximize European competitiveness in the processing and analysis of Big Data over the next 10 years. Specifically, RETHINK big will deliver a strategic roadmap for how technology advancements in hardware and networking can be exploited for the purpose of data analytics while also taking into consideration advancements in applications, algorithms and systems.

In this project, we will not carry out actual research on Hardware optimizations for Big Data, but rather survey the landscape of opportunities and propose a strategic roadmap from that landscape. The outcome of the project will be a series of detailed (mentioning specific technologies), realistic (considering our 10-year timeline), verifiable (including target metrics) and coordinated technology development recommendations that would be in the best interest of European Big Data companies to undertake in concert as a matter of competitive advantage.

Practically speaking, the roadmap will be produced as a result of area specific and cross-functional working groups meetings and congresses. We will initially identify and evaluate the existing competencies across European Big Data application domains and technology providers in Europe and then identify the key European stakeholders, or the established and up-and-coming institutions that possess or are developing the technologies, processes or services that map to these competencies. From these stakeholder institutions, we will select technology and business experts that will chart the technological advancements, their respective challenges and the potential business opportunities that they present. These experts will not only share an interest in defining a credible roadmap, but also hold the decision-making power within their respective institutions (and collectively) to implement that roadmap.

At the highest level, this project will identify and evaluate the existing competencies across European Big Data Hardware and Networking technology sectors and application domains and prioritize the complementary interests and the shared opportunities that allow all key industrial stakeholder companies to unlock the highest return on their respective investments; it will result in a roadmap that would be irrational not to follow.

1.1 Motivation and goals

In future decades, the ability to process and analyze Big Data will have a greater and greater impact on the European Union's productivity and competitiveness. Big Data describes a massive volume of both structured and unstructured data that is too large to be processed using traditional database and software techniques. It is often seen as the **"oil of the 21th century"**; it is estimated that by the year 2020 there will be 35 zettabytes of data. As Big Data analytics is an important part of the modern ICT society, many European companies rely on it to describe consumer behavior by real-time figures, and to use that information as a guideline for business decisions, effectively reducing risk and increasing profit. Businesses and public institutions are currently being challenged to unlock the potential for the most added-value by making use of Big Data. Offering the right information extracted from Big Data, to the right people, is fast become the key to competitive advantage.

- The objective of the RETHINK big Project is to bring together the key European hardware, networking, and system architects with the key producers and consumers of Big Data to identify the industry coordination points that will maximize European competitiveness in the processing and analysis of Big Data over the next 10 years. Specifically, RETHINK big will deliver a strategic roadmap for how technology advancements in hardware and networking can be exploited for the purpose of data analytics while also taking into consideration advancements in applications, algorithms and systems. The resulting report will be used as a direct reference for technology innovation in Europe.
- In a multi-phase approach, the project will identify critical stakeholders in Europe in the area of hardware and networking support for Big Data, engage top stakeholder technical experts with access to decision makers, organize these experts in working groups in applications, hardware, and networking, consider opportunities for cross-functional co-design across working groups and synthesize the working group findings into a strategic roadmap defined as a series of detailed, realistic, verifiable and coordinated technology development undertakings that would be in the best interest of EU Big Data companies to undertake in concert as a matter of competitive advantage.

As an indirect outcome, it is hoped that the roadmap will facilitate the coordination between the Big Data ecosystem and public authorities (EU and the Member States) responsible for Big Data research and dissemination programs. It will foster joint initiatives among the main stakeholders in the area of research and innovation programs and RETHINK big members.

The majority of the investigation and analysis will be performed in well-focused, well-prepared surveys and meetings in which the Working Groups play the central role. The Working Groups are composed of 2 tiers of participation. Tier 1 comprises the Project Partners which includes some of the major players in hardware, networking and software support in addition to European Big Data practitioners. The Partners will be directly responsible for the acquiring, discussing and synthesizing input from key industrial and academic players or stakeholders from Tier 2 into the roadmap. Tier 2 is comprised of the representatives of the stakeholders in the outcome of the project; these are Big Data practitioners and technology providers who will directly benefit from a clear roadmap for advancements in technology support for Big Data in Europe.

Technology roadmapping is a critical activity for a global economy when a near-future technology investment decision is not entirely straightforward, meaning that it is not clear which alternative to pursue and how quickly the technology is needed / will be available or when there is a need to coordinate the development of multiple related technologies. Each of these ideas is relevant in the case of Big Data. The RETHINK big Roadmap must define *which* technical capabilities the European ICT industry needs to develop in order to stay on top of Moore's Law and other trends, as well as defining *when* each of these capacities will be needed.

For this reason, the RETHINK big Project will only be deemed successful having achieved the following outcomes:
- A roadmap that is detailed (mentioning specific technologies), realistic (considering our 10-year timeline), verifiable (including target metrics);

- A roadmap that includes coordinated technology development recommendations that would be in the best interest of European Big Data companies to undertake in concert as a matter of competitive advantage;
- A roadmap that takes into consideration the known Hardware and Networking challenges of performance and scalability, bandwidth and latency limitations, energy efficiency, reliability, and security;
- A roadmap that takes into consideration the known Big Data-related challenges of variety, velocity, volume and veracity;
- A roadmap that accounts for other key enabling technologies from non-Big Data industry and research that might bring disruptive solutions into Big Data;
- A roadmap that facilitates the creation of new international standards, as required;
- A Big Data Technology Ecosystem that includes the key European hardware, networking, and system architects, the key European producers and consumers of Big Data that cooperates with key European existing initiatives in the area of Big Data, HPC, and ICT in Europe, such as the NESSI, the European Data Forum, The European Technology Platform for High Performance Computing and the like…

It is only through this successful roadmap that we can achieve a world-class, globally competitive, European Big Data Technology value chain that provides the key industry decision makers with a mutually beneficial strategy for the long term development for Big Data in Europe.

1.2 Technical work performed and main results

In the first year of the RETHINK big Project, the RETHINK big Project Team focused on setting up the project internally, identifying the Key Stakeholder Companies based on European Competencies and bringing together the representatives of these companies for a series of workshops and meetings.

From a Management (WP1) perspective, the RETHINK big Coordinator (BSC) Management Team spent the first few months of P1 putting the tools and procedures required for strong communication within the Project Team into place. This work included releasing the Project Portal for internal document exchange and project-related information and presentations in addition to creating distribution lists, document templates and a Quality Plan. Additionally, the Management Team set up monthly General Assembly teleconferences (with Minutes and assigned actions) as well as monthly and later nearly weekly Work Package Leader and Working Group Leader teleconferences. At the same time, the Management Team organized and facilitated several face-to-face meetings including the Kick-off meeting held on 19 March 2014 in conjunction with the EDF in Athens, Greece (24 Attendees), the Dissemination Meeting held on 25 April in Barcelona Spain (10 Attendees), the first Working Group Meeting held in Madrid on 18 and 19 September in Madrid, Spain (70 Attendees), the first Synthesis Meeting held on 9 and 10 December in Barcelona (30 Attendees), Spain and finally a SWOT Meeting held on 18 March in London, U.K (10 Attendees). The output from each of these meetings has been analyzed within the context of the technical Work Packages (WP3 Applications Challenges, WP4 Key Enabling Technologies and WP5 Strategy and Roadmapping); it has also served as the baseline content for all of the technical deliverables delivered in P1.

In P1, the Dissemination Team (WP2) led by UPM focused largely on establishing initial online presence. In the initial months of the project, the Dissemination Team worked to launch a preliminary version of the www.rethinkbig-project.eu, LinkedIn Group and Twitter feed. The team then followed up with additional releases of more detailed content on the website in the months following the Kick-off Meeting. Moreover, the team released both a Press Release coinciding with the Kick-off and subsequently a bi-yearly (twice yearly) newsletter highlighting the project motivation, goals and Project Kick-off meeting. At the end of the first year, the team was working toward the release of the second newsletter in addition to making the initial preparations for the final project Congress to take place at the end of P2.

The work of WP3 Applications Challenges, WP4 Key Enabling Technologies and WP5 Strategy and Roadmapping focused primarily on gathering and processing input from industry and academia in a series of meetings, surveys and interviews. They began by identifying the key European Competency Areas in Big Data and used these areas to define a list of target key stakeholders (in the form of experts) from both public and private sectors but with a push to engage industry as much as possible. This culminated in an initial working group meeting at the Universidad Politécnica de Madrid (UPM) on 18 and 19 September 2014. The objectives were to identify challenges across European Big Data-related application sectors and to develop a shared language to discuss these and to identify key strategists within each organization. Nearly seventy experts participated in a structured brainstorming session to facilitate communication across sector, disciplinary and cultural boundaries. Participants were asked to provide concrete, practical and measurable examples which were real rather than hypothetical, and to communicate them in objective, rather than subjective, language. The results from the lively, engaged debates were analyzed with the results of a pre-meeting survey in order to provide a high level assessment of the most common perceived Big Data problems, particularly in industry in the D3.1 Application Challenge and D4.1 Enabling Technologies (both) Target Metrics and Success Criteria documents. As a follow on activity to this initial work, Work Packages 3, 4 and 5 then brought key internal team contributors together with some of the experts from the Working Group Meeting to the Initial Synthesis in which they performed an initial assessment of Strengths, Weaknesses, Opportunities and Threats (SWOT). These results were later fed into a more detailed analysis at the SWOT Meeting in London; this combination of meetings led to the division of labor for writing the first version of the Roadmap recommendations.

Finally, the RETHINK big Project engaged with the most important projects and organization in the current Big Data Ecosystem. RETHINK big submitted a position paper in response to the public consultation for the SRIA sponsored by the EC and Big Data Value and were subsequently invited to the NESSI Summit. We collaborated with associated Big Data projects BIG and BYTE to consider an initial Key Stakeholder Platform Strategy and continue to work together and engage with the Big Data cPPP.