

## 1. Contributor names and short CVs

**Adrian Cristal is the coordinator of the Rethink Big project.** He co-leads the BSC (Barcelona Super Computing) Computer Architecture for Parallel Paradigms research group. His expertise is in hardware and computer architecture. His team is active in the BSC Microsoft Research Center and Intel-BSC Exascale Lab. He has worked on the coordination of the VELOX FP7 project on Transactional Memory, and are partners in the AXLE FP7 project on Advanced Analytics for Extremely Large European Databases and they coordinate the ParaDIME FP7 project on Parallel Distributed Infrastructure for Minimization of Energy. (Dedication = 30% / for each)

**Ernestina Menasalvas.** Responsible of the dissemination of the project. Professor at Universidad Politecnica de Madrid (UPM). Her subject area is Data Mining. Coordinator of the MIDAS “Data Mining and data simulation group” at the Center of Biotechnology in UPM. Her research activities are on various aspects of data mining project development and in the last years her research is focused on data mining on the medical field.

## 2. Type of the presentation proposed

Impact contribution

## 3. Title of the presentation

RETHINK BIG PROJECT: strategic roadmap for hardware, software and networking integration for data analytics.

## 4. Summary of the presentation

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. This presentation will present the advances so far of the roadmap.

## 5. Extended abstract of the presentation

In future decades, the ability to process and analyze Big Data will have a greater impact on the European Union's productivity and competitiveness. The technology progress has focused mainly on developing new paradigms for scalable data storage and processing on commodity hardware and lead to a whole new range of technology which is available in large portions in open-source.

Europe has a very strong industry base of SMEs as well as large Enterprises which are global leaders in their area of expertise /vertical. Global leaders are present in various verticals: i) Automotive: VW, Fiat, Peugeot, BMW ...ii) Telecommunication: Vodafone, Telefonica, Deutsche Telekom, Orange ...iii) Networking / Telecommunication Equipment: Ericsson, Nokia Networks; iv) Pharmacy: Sanofi, Bayer, Roche ...v) Chemical: BASF, Bayer, Total ...vi) Mobile Gaming: King.com, Rovio, SuperCell ... Besides that another strong base of European industry are SMEs and start-ups, which are leaders in their specific area of expertise. This ranges from mechanical engineering, security sensors, glass-manufacturing, etc. Europe also has a strong base in fundamental sciences. While one might argue why those are relevant w.r.t to the roadmap, we only need to look into the past to find good examples of technology derived from fundamental research. There are huge potentials in transforming existing industrial strengths, domain expertise and research strength of EU companies to enable future growth of existing business models, but even more so enabling new business models in the future.

In this context, RETHINK big Project aims 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 is working to 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.

Consequently, along the two years, this project has identified and evaluated 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. Altogether will be the input for the roadmap that will be presented in the proposed presentation.

In fact by the celebration of EDF 2015 the project is in the last semester of the RETHINK BIG project and it is a perfect moment to discuss the results of the roadmap so far with the participants to this event.

The objective of the presentation is to present results so far of the roadmap bringing together researchers and key industrial players.

The presentation will bring two perspectives:

1. Application domains: The perspective from different domains will be presented to do it possible Rethink Big has been investigating future Big Data application requirements from a wide spectrum of application domains that make use of Big Data. Even that the stress has been put on four initial domains which include classical science applications such as Physics, Chemistry, and Engineering in addition to cutting-edge Life Sciences applications such as -omics (genomics, proteomics and proteomics, etc.) that are leading the march toward more personalized medicine during the session we will also consider the demands of data produced by the ubiquitous internet and internet of things and “always on” social media as well as the critical role data now plays in business and finance applications.
2. Enabling technologies: we aim here to cover the challenges for technology with special emphasis on Hw, Sw and network integration that all the application poses. In fact we will present future hardware directions and how they may be applied to meet the specific needs of Big Data applications. Once again we will put special emphasis on trends in conventional and unconventional processor technologies, memory and storage architectures, distributed systems, and

networking but we will also cover the software aspects will cover developments in software frameworks, programming models, data structures and algorithms

The session that will be focused on the 10 key Roadmap points would start with an introduction covering the methodology followed in the project by the coordinator, this will be followed by one speaker that will present the two perspectives covered by the project: applications and enabling technologies. This would then be followed by the coordinator presenting the roadmap and the main conclusions so far. To end with we propose an open discussion with the audience.

The main outcomes of the presentation will be:

- Diffusion of the roadmap
- Detailed comments on the roadmap as presented to the audience.
- A shared understanding on the challenges and opportunities for the Big Data companies in Europe.

Consequently the targeted audience will be:

- Big Data technologists from research, academia and industry
- Industry leaders
- Newly created companies (i.e.: Start-Ups, Incubators, ...)

We will in fact be inviting all the partners of the project together with all those (more than 100) that have contributed in the process of roadmap definition.