



# EUBrasilCloudFORUM

Fostering an International dialogue between Europe & Brazil

## EUBrasilCloudFORUM concertation meeting

18 March, CBPF, Rio de Janeiro, Brazil

# Session 1: EU-Brazil ICT collaboration today

*Wagner Meira Jr., Universidade Federal de Minas Gerais*

*Project acronym*

*Europe – Brazil Collaboration of BIG Data Scientific Research  
through Cloud-Centric Applications*

# EUBra-BIGSEA objective & focus

- ▶ The development of **advanced QoS cloud** to support **Big Data**
- ▶ The development of innovative **Big Data services** for capturing, federating and annotating large volumes of data
- ▶ The use of **innovative** and **efficient** technologies for guaranteeing the fulfilment of the **security** and **privacy** policies
- ▶ The **transfer** of this technology to a **real user scenario** with high social and business impact, and of high interest for both EU and BR
- ▶ The **dissemination** and subsequent **uptake** of the results and identification of business opportunities

# EUBra-BIGSEA expected results

- ▶ A BigData Analytics platform
  - ▶ A QoS layer for IaaS, agnostic to the cloud provider (including containers), with automatic horizontal and vertical elasticity
  - ▶ Customized for BigData applications
  - ▶ A privacy and security framework suitable for performing efficiently on Big Data applications
  - ▶ A programming model bridging to data analytics and QoS functionality, independent of the IaaS provider
- ▶ A real use case implemented on top of the platform for the analysis of data from massively connected societies
  - ▶ Transfer of the technology to external projects, institutions or entities
  - ▶ The technology will be released under Open-Source licenses
- ▶ An international research community on Cloud and Big Data

# EUBra-BiGSEA Exploitable background

| Name of the assets                   | Short Description  | Exploitation plans  | Source                                   |
|--------------------------------------|--|---|--|
| Multi-level Dynamic Elastic Cloud    | A container & VM-based IaaS with Horizontal and Vertical elasticity            | Include along on-premise Cloud Management Platform ecosystems | EU & national projects (INDIGO, CLUVIEM) |
| Big Data Application Models          | Models of data quality, descriptive and predictive Data Analysis applications. | Not yet identified  | New                                      |
| Big Data Analytics Framework         | Multi-platform data analytics service and API                                  | Not yet identified  | Dice                                     |
| Programming Model                    | Programming model integrating general purpose languages with Big Data services | Not yet identified  | EUBrazilCC                               |
| Big data security framework          | Security and privacy-preserving framework tailored for big data applications   | Not yet defined   | New                                      |
| Traffic Data Recommendation Use Case | A system to predict best routes based on environmental and social conditions.  | Services for the municipality                                 | New                                      |

# EUBra-BIGSEA & Cloud computing challenges

- *Personal data protection & privacy*
  - *Privacy annotation of data and processing*
  - *Efficient privacy protection for BigData*
- *Flows of data & portability*
  - *Compromise between on-the fly access and replication*
  - *Support of a variety of standard formats*
  - *Data Quality as a Service*
- *Vendor lock-in*
  - *Infrastructure-agnostic solutions*
- *API-centric project for Data Scientists and Application developers*
- *Integration of multiple programming models*



# Hurdles & recommendations

- We anticipate a high complexity on the modelling of applications to adjust resources on deployment
  - Dynamic readjustments will correct divergences to achieve the QoS
  - Learning process
- Scattered legal framework in Europe and differences with Brazil
- Priorities between EU and Brazil are not always aligned in terms of security / privacy / infrastructure needs
- No standardised way to represent the data – this might limit the applicability of the solutions
  - In the use case working with data from Curitiba municipality, but thinking wider
- Create a self-sustainable community and impact on standards