



Plataformas Big Data Eficientes y Escalables para Machine Learning

Guillermo López Taboada

Guillermo.Lopez@Torusware.com

Taboada@udc.es

Machine Learning Workshop Galicia 2016

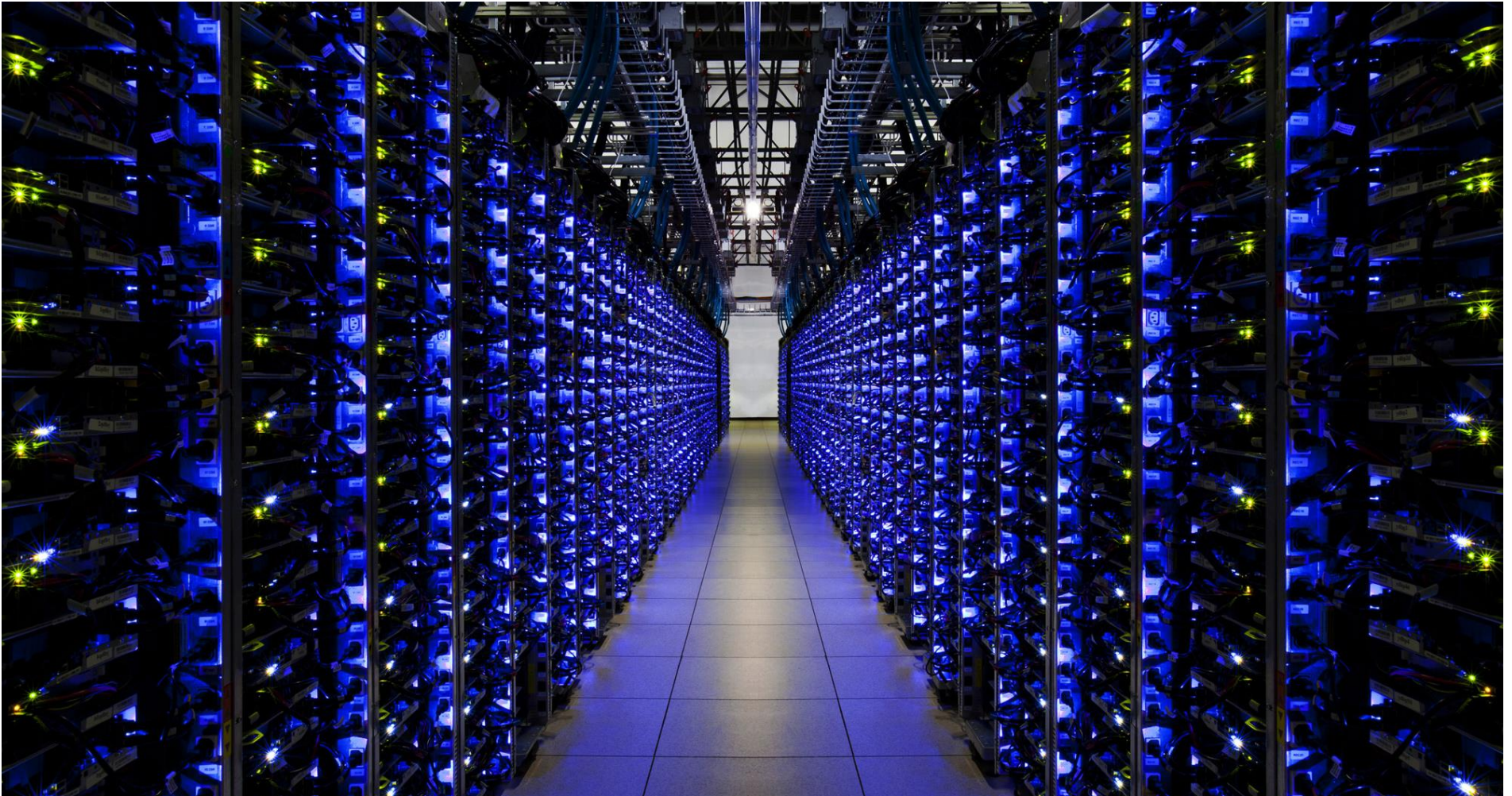


Rede Galega de Tecnoloxías

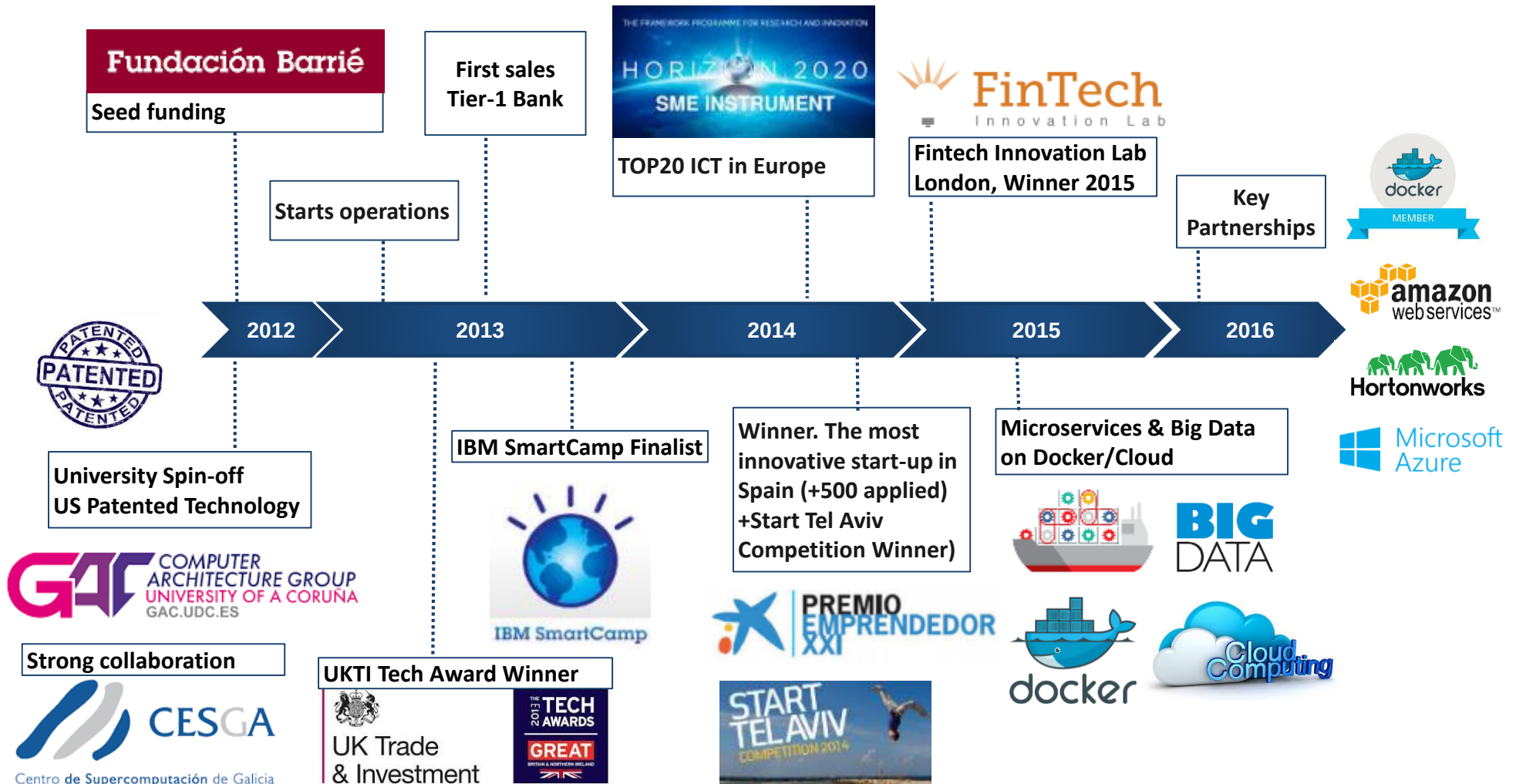
Cloud e Big Data para



Torusware: *Enabling Applications for the Data Centric World*



Torusware assists organizations adopting microservices and Big Data to unlock their transformational power through agile and scalable software



Product Company



Available in <http://speedus.torusware.com>

AWS and Azure Market Places & Docker Registry



Professional Services Business Area



Microservices/Cloud/Docker and Performance Optimization



Big Data Architecture

BIG
DATA



Torusware addresses the need for fast and scalable Big Data through:



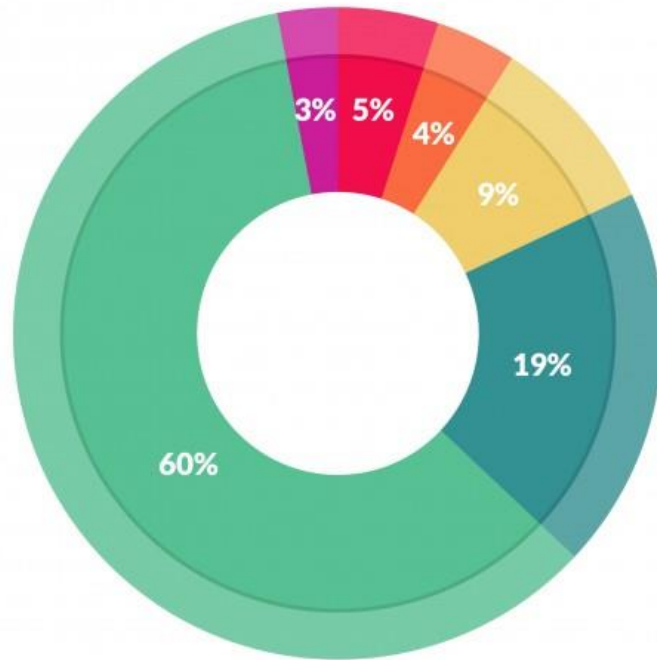
Professional Services in Big Data:

- Analyze, Design, Build, Deploy, Run & Optimize Big Data Architectures
- Discovering the “Business Case” in Big Data
- Assessments and Proof-of-Concept projects
- Training on Big Data



- ## Professional Services in Big Data:
- Analyze, Design, Build, Deploy, Run & Optimize Big Data Architectures
 - Discovering the “Business Case” in Big Data
 - Assessments and Proof-of-Concept projects
 - Training on Big Data
- 

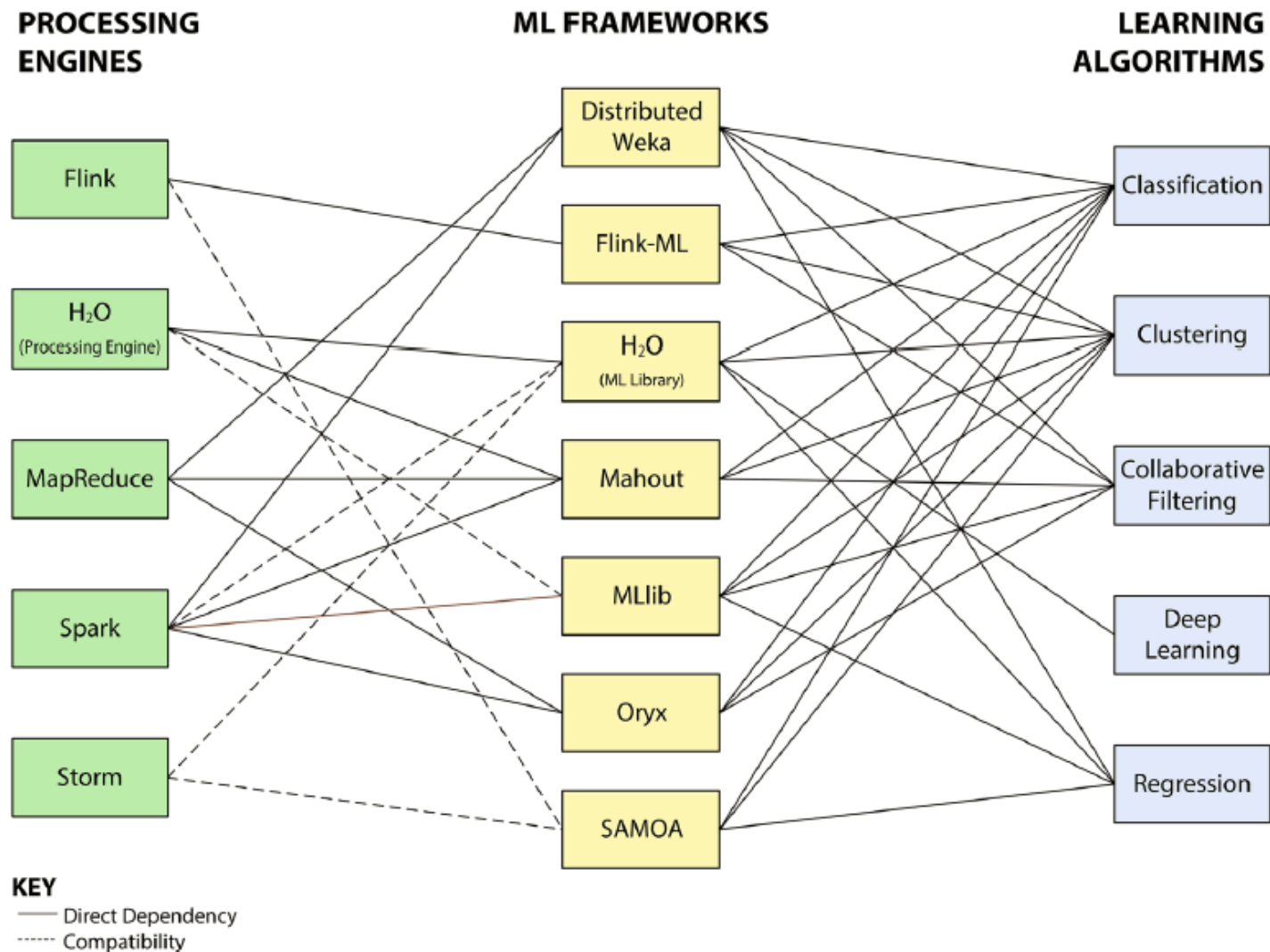




What data scientists spend the most time doing

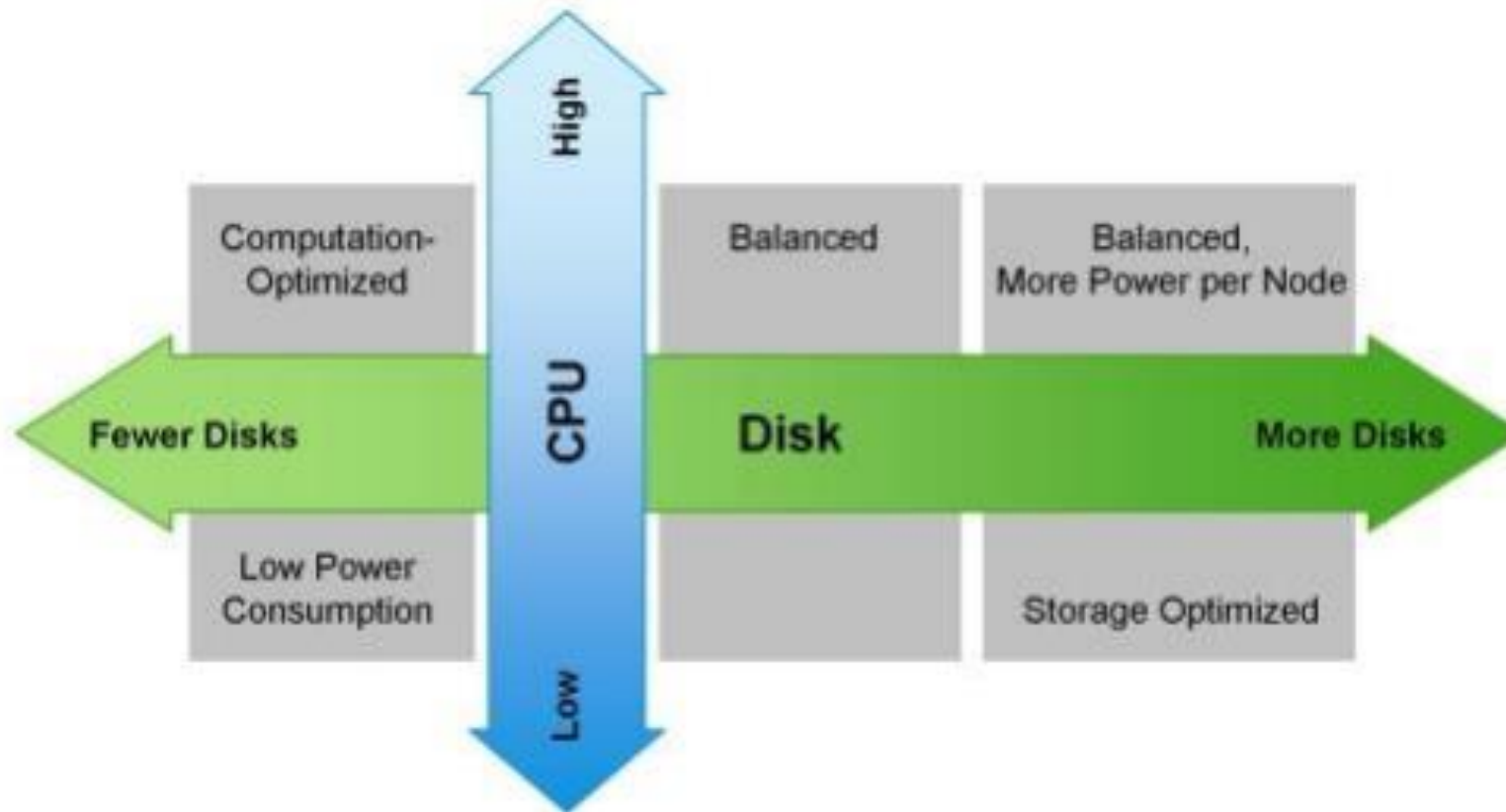
- Building training sets: 3%
- Cleaning and organizing data: 60%
- Collecting data sets; 19%
- Mining data for patterns: 9%
- Refining algorithms: 4%
- Other: 5%

Los Data Scientist dedica menos de 4 horas semanales a Machine Learning



Que configuración escoger: *CPU vs Disk*

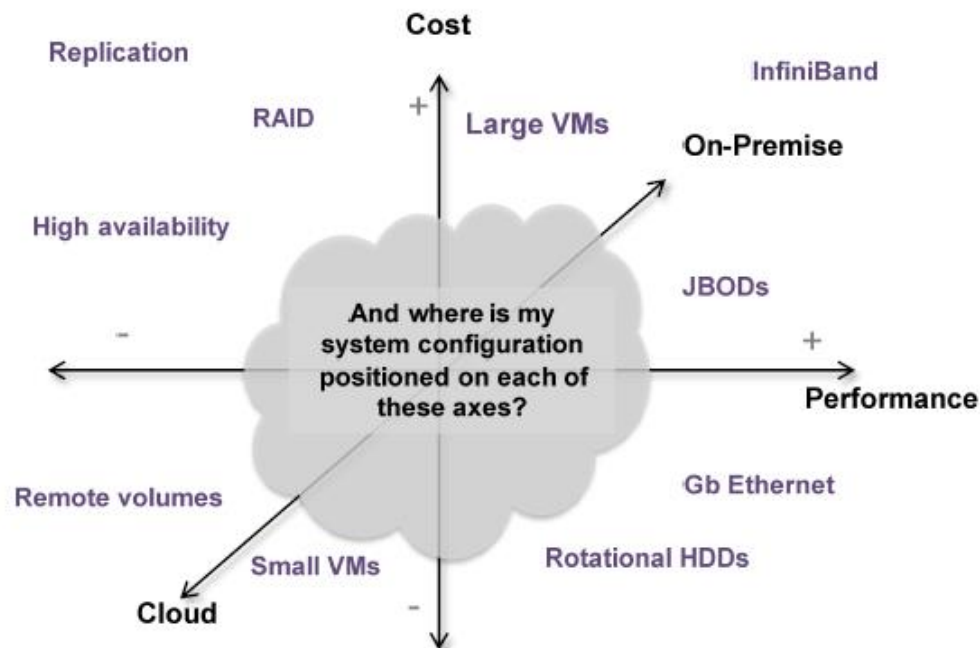
Figure 9. Workload based configuration chart



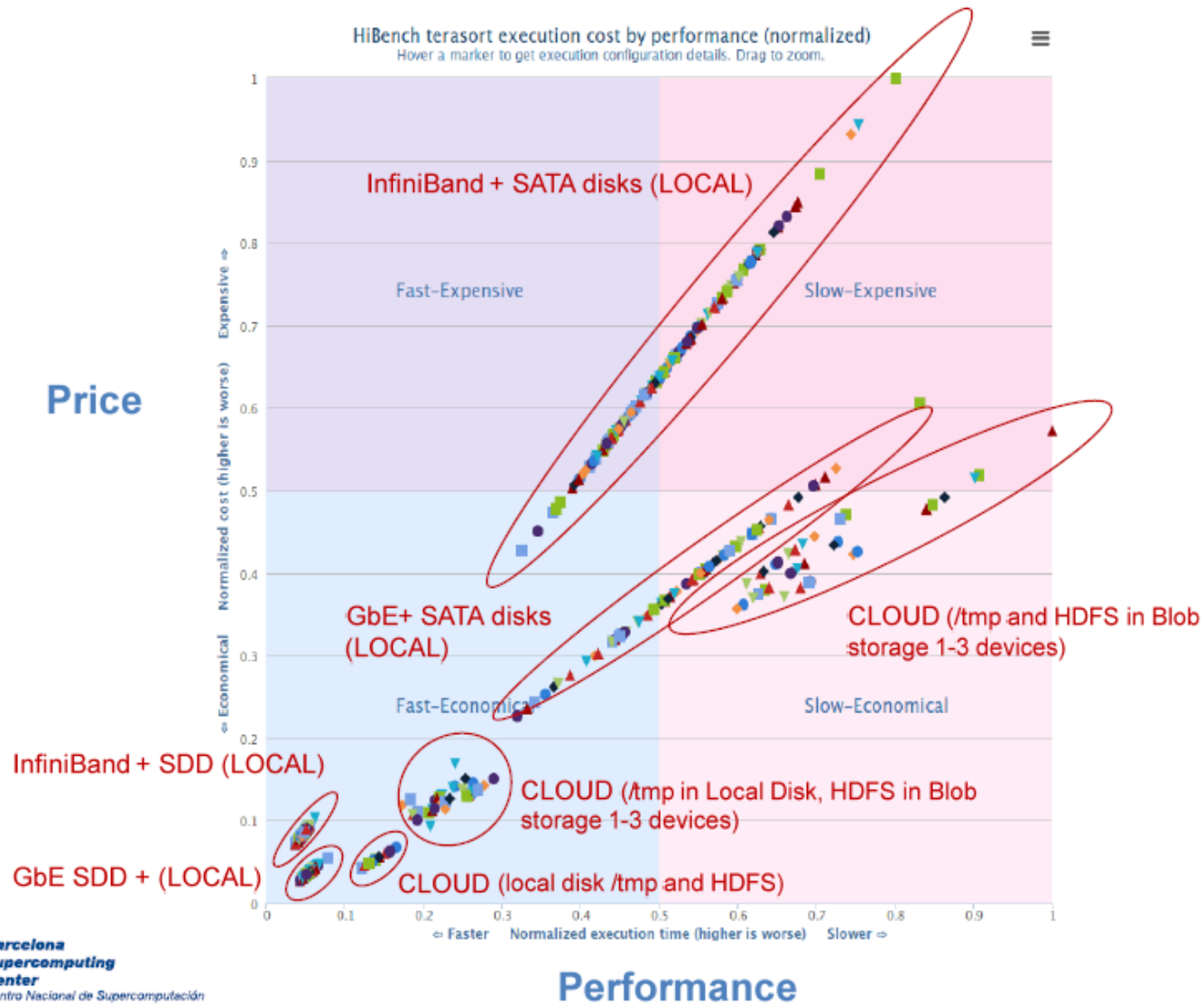
Source: Cloudera blogs: <http://blog.cloudera.com/blog/2013/08/how-to-select-the-right-hardware-for-your-new-hadoop-cluster/>

« What is the most cost-effective configuration for my needs?

- Multidimensional problem

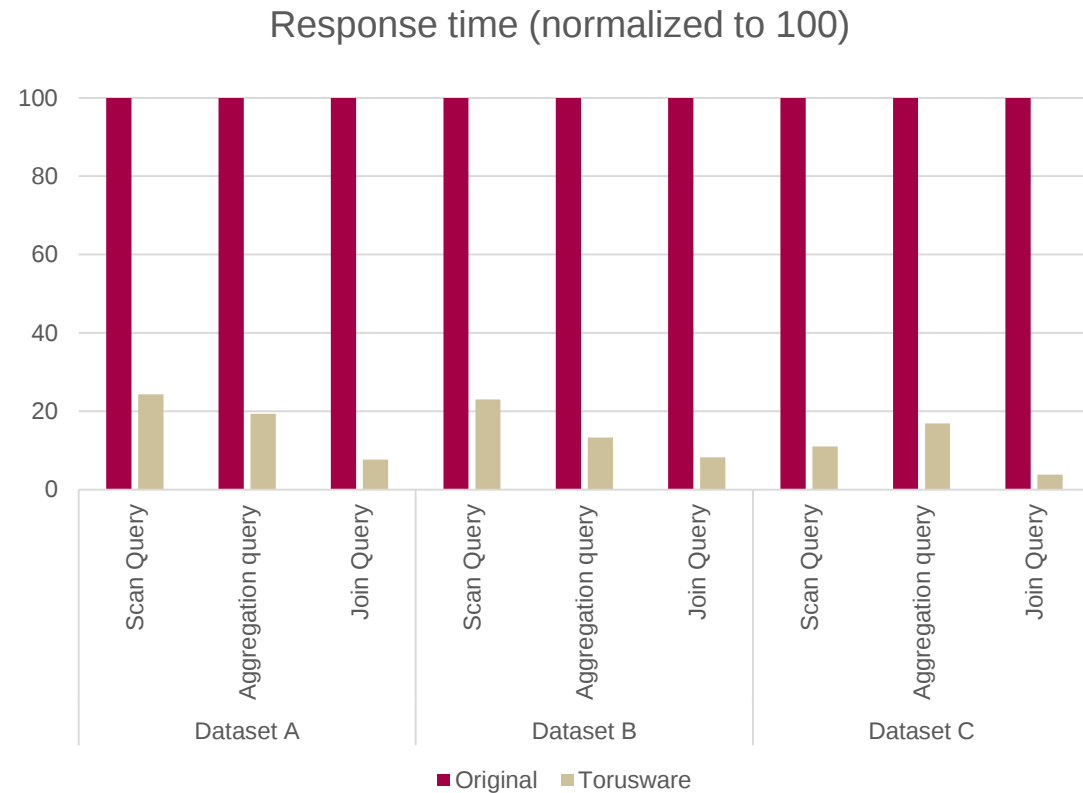


Que configuración escoger: *Rendimiento y coste: almacenamiento + red*





Caso de éxito: reducción del tiempo de respuesta un 75%-90% con un 50% menos de infraestructura



Contribución de Torusware

Cliente



Diseño del “Data Lake” (el corazón) de la Plataforma Big Data

Tier-1 Bank



Diseño y mantenimiento de plataforma Big Data

Transport leader



Diseño de la nueva plataforma Big Data más rápida y económica

Retail leader

Marketing

Diseño y despliegue de plataforma analítica que permitió duplicar la tasa de éxito de las campañas

Media leader



Diseño de una arquitectura Big Data 10 veces más rápida

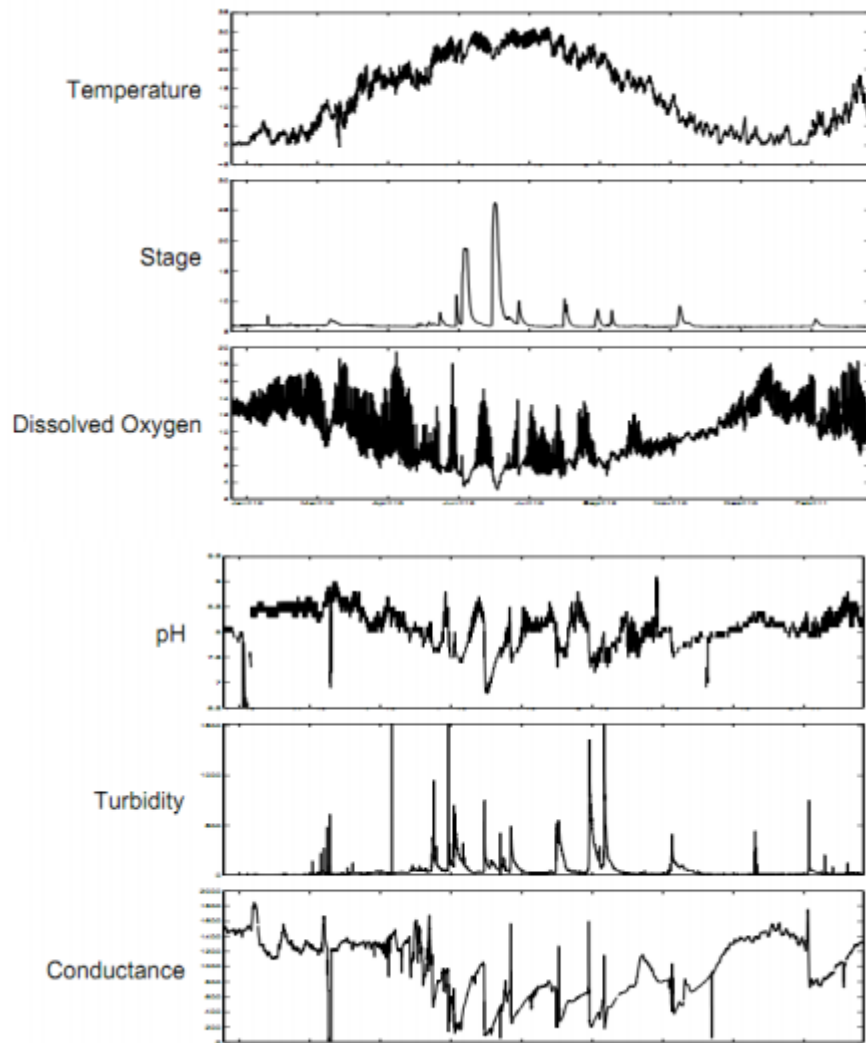


WATERSIG

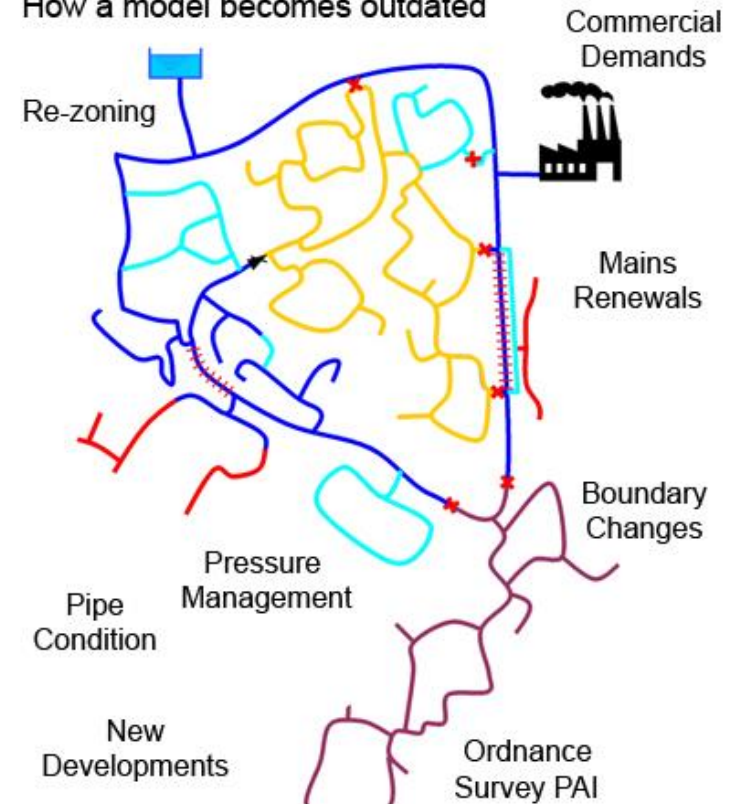
Aplicación de las **Tecnologías de Información** a la Gestión Integrada de servicios de **Agua** Potable y Saneamiento bajo paradigma **Big Data** y empleo de Estrategias de **Paralelización**

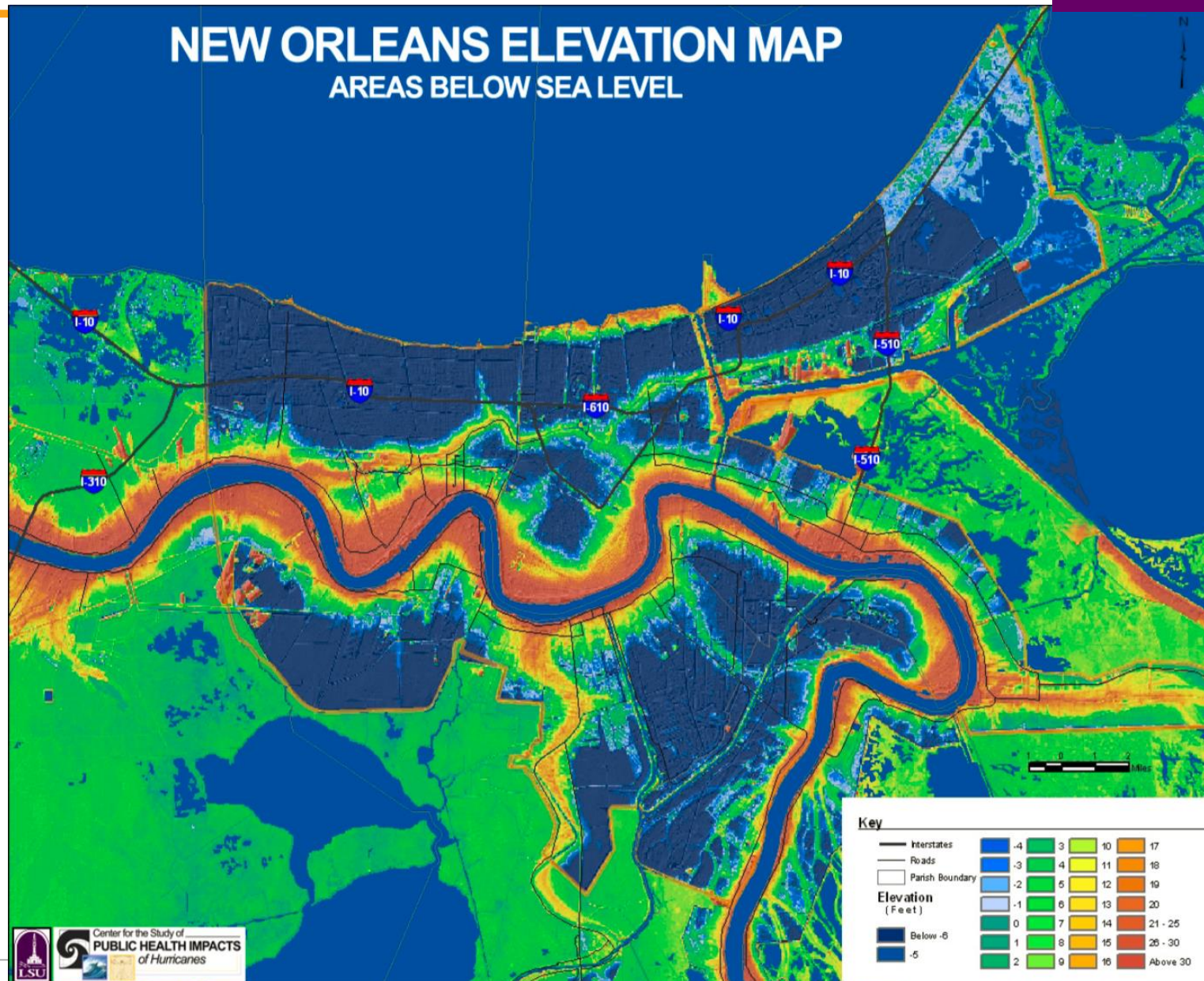


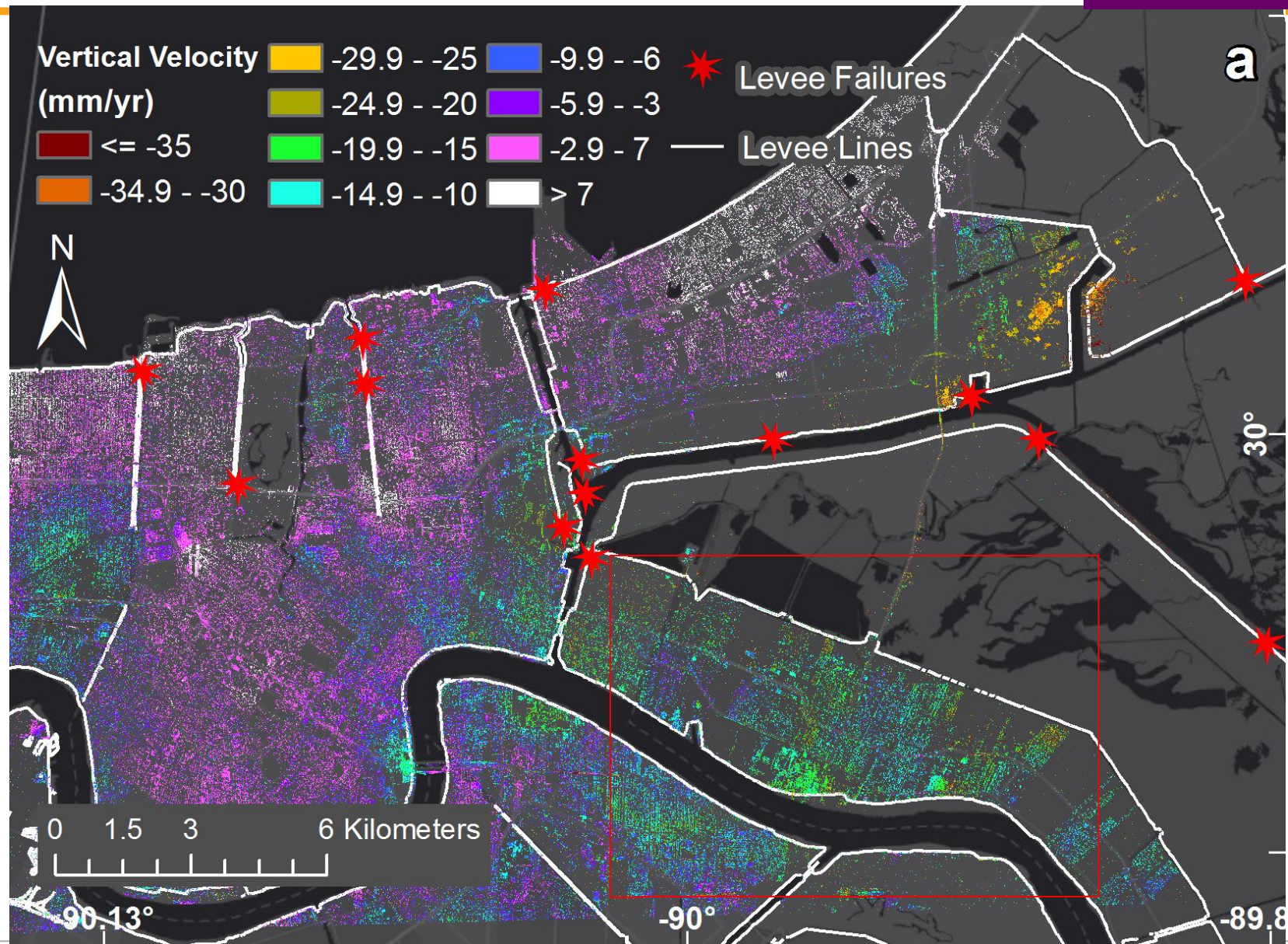
Programa CONECTA PYME – Convocatoria 2015



Model Maintenance How a model becomes outdated









Contact:

Email: Info@Torusware.com

[Torusware HQ]
68 1ºB, Ronda Outeiro
15009 A Coruña, SPAIN
Phone: +34 981 97 82 86

[Torusware@UK]
L39, One Canada Square, Canary Wharf
E14 5AB London
Phone: +44 7756 788791

