

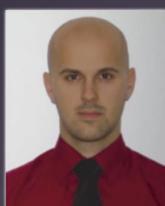
BIG DATA

WORKSHOP: BIG DATA PROCESSING WITH APACHE SPARK

CIC - IPN; Evento organizado por el laboratorio de Ciencia de los Datos y Tecnología de Software

Objective:

The interest in Apache Spark surpassed Apache Hadoop, as it provides a more convenient and faster way to process large data sets and currently used at a wide range of organizations to build big data applications. The goal of the course is to provide an introduction to distributed computing as well as an in-depth practical knowledge to the most powerful, general analytical engine, Spark. Students will be able to write applications quickly in Java, Scala and Python by using Spark's sophisticated libraries to process unstructured, structured, graph data or data streams.



Lecturers: **Zoltán Zvara**, **Mátyás Manninger** and **Frederick Ayala**

Date: May 23 – Jun 03; 9:00-12:00

Place: Sala de usos múltiples CIC IPN

Requisites: Background in networks, Linux administration and Java; and Laptop (preferable core i5 or AMD A8 or higher).

LIST OF TOPICS:

1. Introduction to a powerful language, Scala (*theoretical and practical*)
2. Scala classes, objects (*theoretical and practical*)
3. Advanced Scala Spark (*theoretical and practical*)
4. Distributed challenges, concepts, tools (*practical*)
5. Introduction to Apache Spark (*practical*)
6. Data exploration with Apache Spark 1 (*practical*)
7. Data exploration with Apache Spark 2 (*practical*)
8. Data exploration with Apache Spark 3 (*practical*)
9. Big Data processing with Apache Spark 1 (*practical*)
10. Big Data processing with Apache Spark 2 (*practical*)

More information and details about the workshop: 5729-6000

ext. 56600

Dr. Martínez Luna

llunameztli@gmail.com

ext. 56615

Dr. Olivares Ceja

jesusoc@hotmail.com

Inscriptions: <http://www.capacitacion.cic.ipn.mx>

5729600

ext. 56605

Lic. Alfonso García González

About Instructors:

<http://www.sztaki.hu/en>

<http://www.sztaki.hu/department/INFOLAB/>

<http://www.sztaki.hu/department/WebSearch/>

<http://www.sztaki.hu/department/DWBIG/>



Av. Juan de Dios Bátiz, Esq. Miguel Othón de Mendizábal Col. Nueva Industrial Vallejo Delegación Gustavo A. Madero C.P 07738. México, D.F. www.cic.ipn.mx



Centro de Investigación
en Computación
Instituto Politécnico Nacional