



Scuola Superiore
Sant'Anna

Course on “Big-Data and Analytics”

Prof. Tommaso Cucinotta

Goal

This course provides an overview of the challenges to face, and the technical solutions to embrace, when building large-scale distributed systems that need to process huge amounts of data, often in presence of industrial-grade reliability, availability and real-time processing constraints. In order to cope with the huge amounts of data to process, as well as the volume of the submitted workload, these systems make use of heavily distributed algorithms spanning across several machines, therefore scalability of architectures and protocols becomes a primary concern.

Program at a glance

- Big Data and Analytics
 - Basic concepts
 - Real-time data streaming and analytics
 - Storage services for big-data
 - Distributed messaging/queueing systems
 - Map Reduce
 - Big-Data and the Internet of Things
- Platforms
 - Apache Hadoop, Storm, Spark
 - Apache ActiveMQ and alternatives
 - AWS Lambda and Google Cloud Functions

Requirements

Students need a basic understanding of software, computer architectures, distributed systems and communication protocols.

Why to attend

Students will acquire a unique insight into the world of big-data related technologies, and will be able to

master key concepts behind them. This is a fundamental brick in the background of a software engineer / computer scientist who will deal with modern distributed software systems in industry or academia, spanning across high-performance, cloud and even (increasingly connected) embedded systems.

About the course

Duration and format: 10 hours of front lessons

Credits: 1 CFU

Schedule: flexible, to be agreed with students

Exam: oral + project (optional)

About the lecturer

Prof. Tommaso Cucinotta has a MSc in Computer Engineering from University of Pisa and a PhD from Scuola Superiore Sant'Anna. He spent more than 10 years at the Real-Time Systems Laboratory (ReTiS) of Scuola Superiore Sant'Anna carrying out research in security and smart-card based authentication, adaptive deadline-based scheduling in the Linux kernel for embedded, soft real-time and multimedia applications, temporal isolation in virtualized cloud services and novel OS designs for massively parallel and distributed systems. He has been MTS at Bell Labs in Dublin, carrying out industrial research on security and confidentiality, and real-time performance of cloud systems, with a focus on Telco applications. He has also been a Software Development Engineer in AWS DataBase Services in Dublin, Ireland, working on scalability and performance enhancements to the AWS DynamoDB NoSQL real-time data-base. Since 2016, he is back at the ReTiS of Scuola Superiore Sant'Anna as associate professor.