

Emerging tools and techniques for massive data analysis

December 15--16, 2014 -- Bologna

Final agenda

	Monday, December 15	
09.00 -- 09.20	Registration	
09.20 -- 09.30	Introduction/logistics/agenda	Giuseppe Fiameni
09.30 -- 11.00	Big Data Challenges	Prof. Sonia Bergamaschi Dip. of Engineering "Enzo Ferrari"
11.00 -- 11.20	Coffé break	
11.20 -- 12.30	Introduction to Map--Reduce and Hadoop (Hadoop, HDFS, Yarn)	Giovanni Simonini Dip. of Engineering "Enzo Ferrari"
12.30 -- 14.00	Lunch break	
14.00 -- 14.40	Data analytics -- other approaches (Pig, Hive, Pregel, etc.)	Giovanni Simonini Dip. of Engineering "Enzo Ferrari"
14.40 -- 15.40	Set--up the working environment [1]	Paolo D'Onorio De Meo
15.40 -- 16.00	Coffé Break	
16.00 -- 17.30	Exercices: First Map--Reduce algorithms (NGS) [2]	Paolo D'Onorio De Meo, Giuseppe Fiameni, Giovanni Simonini Dip. of Engineering "Enzo Ferrari"
17.30	End of the first day	

	Tuesday, December 16	
09.00 -- 9.30	Using Hadoop on Cineca resources (PICO)	Giuseppe Fiameni
09.30 -- 11.00	In memory Map--Reduce: the case of Apache Spark	Giovanni Simonini Dip. of Engineering "Enzo Ferrari"
11.00 -- 11.20	Coffé Break	
11.20 -- 12.30	Exercices: <ul style="list-style-type: none"> • SPARK algorithms [3] 	Paolo D'Onorio De Meo, Giuseppe Fiameni, Giovanni Simonini Dip. of Engineering "Enzo Ferrari"
12.30 -- 14.00	Lunch break	
14.00 -- 17.30	Exercices: <ul style="list-style-type: none"> • Matrix--matrix product • Submitting a Job on PICO [4] 	Paolo D'Onorio De Meo, Giuseppe Fiameni