Distributed Enabling Platforms (PAD)

• Teacher(s) name: Nicola Tonellotto
  email: nicola.tonellotto@isti.cnr.it
  tel: 050 315 2967
  web: http://hpc.isti.cnr.it/~khast/

• Semester: 1st
• Exam mode: project + oral examination
Syllabus

• Design issues and solutions in very-large-scale distributed systems

• The objectives of this course are:
  • to develop an understanding of the typical issues of very large scale distributed systems;
  • to equip students with tools, best practices and common procedures to design, implement and program such systems, through understanding of algorithms and suitable theoretical models.

• List of topics:
  1. Introduction to large scale distributed systems.
  2. Cloud computing: introduction, service and deployment models, solutions.
  3. Infrastructure: virtualization, coordination, scalability, availability
  4. Programming: mapreduce model, APIs, patterns
  5. Data: data management, consistency, replication, fault tolerance.
Expected Outcomes

• Distributed Computer System Engineer
  • Analyze requirements
  • Understand design choices
  • Propose and implement solutions

• Large-scale Data Manager
  • Data warehousing
  • Analysis of solutions

• Big Data Analyst
  • Program applications to crunch terabytes of data
  • From numbers to texts to structured data
Thesis available

- Web search algorithms for efficient processing
- Distributed and replicated architectures for Web search
- Energy-efficient Web Search
- Green Information Retrieval
- Dynamic modeling of distributed systems
- Large-scale algorithms for data processing
- Efficient large-scale machine learning algorithms