Overview: This course is an introduction to architectures to enable scalable computing for senior under-graduates. The course will cover essential concepts (including mapreduce and in-memory processing) while presenting a series of examples, such as redis/memcached, hadoop, graphlab and spark.

Pre-requisites: EE 422C or equivalent.

Course website: The Canvas site. This website will have homeworks, solutions for you to download.

Course Policy: Attendance is expected. You are responsible for material covered in the reading assignments (even if not covered in class) as well as material covered in class that is not in the book. Homework will be assigned roughly every week and will be due in class at the end of the week. **Late homeworks will not be accepted.**

Grading:

 (i) Class Participation: 15%

 (ii) Assignments: 35%

(iii) Final Project: 50%

<div align="center">

Syllabus

</div>

1. **Introduction:** Introduction to architectural concepts in scalable computing. Mapreduce.

2. **Distributed Storage:** Distributed disk-storage based architectures - motivation and examples.

3. **Learning:** Basics of Machine Learning. Unsupervised and supervised learning, examples.

4. **To distribute or not to distribute:** Algorithmic considerations in distributed architectures, architectural considerations for distributed computing.

5. **Distributed in-memory Computation:** Examples where distributed in-memory architectures are desired and/or even essential for scalable computing. Spark as example.