Seminar: Topics in Cloud-Based Architectures (1851450)
October 2022 (סימסטר א תשפ"ג)

OVERVIEW

The past two decades have experienced a significant shift in the way large-scale systems are designed, due to cloud computing and big data processing. These systems have been built by what we call today "hyper-scale companies", or "Cloud companies". Large companies such as Google, Microsoft, Amazon, Facebook, LinkedIn, Netflix, Twitter and others had to create highly scalable and highly available systems to support hundreds of millions of users simultaneously. They had to efficiently and reliably store and process huge amounts of new type of data (photos and movies, indexes, social graphs). This led to the creation of new technologies and development paradigms.

In this seminar students will read, analyze and present to their peers key papers in this area that were published in major scientific Systems conferences. The selected papers cover a broad range of topics in cloud-based architectures, such as
- New architectures for storage & data (object stores, GFS/HDFS, Data Reduction, BigTable)
- Distributed computing (Dynamo, Zookeeper, MapReduce, Spark, Mesos)
- Analytics support for BigData (Graph processing, Data formats such as Parquet)
- Reliability (Replication systems and codes)

Time/Place: Tuesdays, 17:15-19:45 , Weston 031

Instructor:
Dalit Naor
office hours: By appointment
email: dalit.naor@mta.ac.il

GOALS

The objective of the seminar is twofold:

- Expose students to the architectural shift of cloud-based architectures; understand the motivation for building these new highly scalable and distributed data-intensive systems, and how they are structured
- To practice the art of reading a research Systems paper, comprehending it and presenting it to others

GRADING

- Critical reading of an academic Systems paper (from a list of proposed papers) and a presentation of the paper in class (60%)
- Participation and contribution to the discussions (10%)
- Exercises (30%)

Note: this class requires physical attendance. This is important for the presenter and for follow-up discussions in class. Attendance is required in 10 (out of the 13) meetings.

LEARNING OUTCOMES

- Understand how to properly read a computer science Systems paper and how to present it.
- Understand in depth some of the key concepts that prevailed in Cloud Computing such as consistency, replication, NoSQL, MapReduce.