

שם הקורס: סמינר במערכות אחסון מידע
Seminar on Storage Systems: Deduplication
מספר הקורס 236834

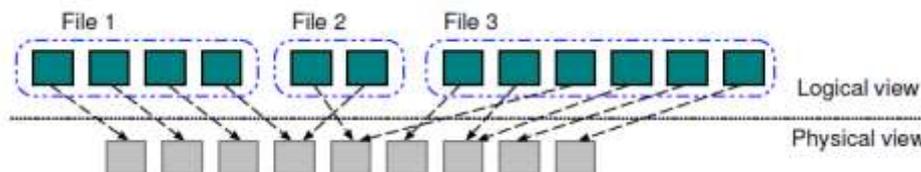
מרצים:	פרופ' ישראל צידון ודר' גלה ידגר
שעות הרצאה:	רביעי 9:30-11:30
דרישות קדם:	מערכות הפעלה (234123)

תאור הקורס

“From 2005 to 2020, the digital universe will grow by a factor of 300, from 130 exabytes to 40,000 exabytes, or 40 trillion gigabytes (more than 5,200 gigabytes for every man, woman, and child in 2020). From now until 2020, the digital universe will about double every two years.”

[IDC study for EMC, 2012](#)

Data deduplication is one for the most effective ways to reduce the size of data stored in large scale systems, and is widely used to date. The process of deduplication consists of identifying duplicate chunks of data in different files (including backup versions, virtual machine images, etc.), storing a single copy of each unique chunk, and replacing the duplicate chunks with pointers to this copy.



The goal of this seminar is to study the basic concepts of deduplication, the challenges and tradeoffs it introduces in the design of large scale storage systems, and the most recent advances in addressing these challenges. We will cover the various aspects of system design, including chunking, duplicate identification, metadata and reference management, inline vs. background deduplication, centralized vs. distributed deduplication, hard disk vs. SSD storage, and more.

דרישות הקורס

The course will combine lectures by the instructors with independent reading in a seminar format. The students will read important papers in the field, will reason about the results in a critical way, and will present them in class along with their own ideas for extending the results.

רישום

This is a joint seminar of the Computer Science and Electrical Engineering departments. CS students who wish to register: please email your name and ID to Gala Yadgar. Please state whether you are a graduate or undergraduate student (which semester), and other information you think is relevant, such as related courses you took.