שם הקורס: סמינר על קידודים למחיקות במערכות אחסון מבוזר
Seminar on Erasure Coding in Distributed Storage Systems
מספר הקורס 236804

מרצה:	פרופ' איתן יעקובי
שעות הרצאה:	ראשון 12:30-14:30
דרישות קדם:	ראה תאור הקורס

תאור הקורס

"Our warehouse stores upwards of 300 PB of Hive data, with an incoming daily rate of about 600 TB. In the last year, the warehouse has seen a 3x growth in the amount of data stored. Given this growth trajectory, storage efficiency is and will continue to be a focus for our warehouse infrastructure."

(Facebook, April 2014)

The traditional approach to maintain data integrity in large scale storage systems is to replicate each data item several times, so that if an item is corrupted or lost, its content can be retrieved by accessing one of its replicas. However, as storage volume in such systems continues to grow in meteoric speed, replication becomes too costly and inefficient. **Erasure codes** enable reconstruction of corrupted or unavailable data by algebraic manipulations on the data items that remain in the system. Their storage overhead is an order of magnitude smaller than that of replication, which makes them ideal candidates for protecting data at large scale.

Erasure coding is an important research area in information theory. Current objectives in constructing new codes and techniques are to minimize their storage overhead and reconstruction time without interfering with the system's normal operation. The goal of this seminar is to understand the benefits, limitations, and tradeoffs of erasure codes in the context of real world, large scale, distributed storage systems. We will survey promising new coding techniques, as well as experience in employing them in large scale data centers and distributed storage systems.

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.

דרישות קדם

The course combines techniques from the fields of coding theory, linear and modern algebra, and operating and storage systems. There are no formal pre-requisites, but the following courses provide good background:

236309 Introduction to Coding Theory

234322 Information Storage Systems

To register, email your name and ID to the instructor. Please state whether you are a graduate or undergraduate student, and other information you think is relevant, such as related courses you took.

ספרות

The course will survey recent literature with the state-of-the-art works on erasure coding, for example:

- **Erasure coding theory**

1. I. Tamo, Z. Wang, and J. Bruck, *Zigzag Codes: MDS Array Codes with Optimal Rebuilding*, *IEEE Transactions on Information Theory*, vol. 59, no. 3, pp. 1597-5313, March 2013.
2. M. Blaum, J. Brady, J. Bruck, and J. Menon, *EVENODD: An Efficient Scheme for Tolerating Double Disk Failures in RAID Architectures*, *IEEE Transactions on Computers*, vol. 44, no. 2, pp. 192-202, February 1995.
3. A.G. Dimakis, P.B. Godfrey, Y. Wu, M.J. Wainwright, and K. Ramchandran, *Network Coding for Distributed Storage Systems*, *IEEE Transactions on Information Theory*, vol. 56, no. 9, pp. 4539-4551, September 2010.
4. I. Tamo and A. Barg, *A Family of Optimal Locally Recoverable Codes*, *IEEE Transactions on Information Theory* vol. 60, no. 8, pp. 4661-4676, August 2014.
5. P. Gopalan, C. Huang, B. Jenkins, and S. Yekhanin, *Explicit Maximally Recoverable Codes With Locality*, *IEEE Transactions on Information Theory*, vol. 60, no. 9, pp. 5245-5256, *September* 2014.
6. M. Blaum, J. L. Hafner, and S. Hetzler, *Partial-MDS Codes and Their Application to RAID Type of Architectures*, *IEEE Transactions on Information Theory*, vol. 59, no. 7, pp. 4510-4509, *July* 2013.
7. M. Blaum, J. Plank, M. Schwartz, and E. Yaakobi, *Construction of Partial MDS (PMDS) and Sector-Disk (SD) Codes with Two Global Parity Symbols*, *IEEE Transactions on Information Theory*, vol. 60, no. 5, pp. 2673-2681, May 2016.
8. W. Huang and J. Bruck, *Secure RAID Schemes for Distributed Storage*, *Proc. IEEE Int'l Symp. on Information Theory*, Barcelona, Spain, July 2016.

- **Storage systems research**

9. O. Khan, R.l Burns, J. Plank, W. Pierce, C. Huang, *Rethinking Erasure Codes for Cloud File Systems: Minimizing I/O for Recovery and Degraded Reads. FAST 2012.*
10. J. Plank, M. Blaum, and J.L. Hafner, *SD Codes: Erasure Codes Designed for How Storage Systems Really Fail*, *FAST* 2013.
11. M. Li and P. Lee, *STAIR Codes: A General Family of Erasure Codes for Tolerating Device and Sector Failures in Practical Storage Systems*, *FAST* 2014.
12. M. Xia, M. Saxena, M. Blaum, and D.A. Pease, *A Tale of Two Erasure Codes in HDFS*, *FAST* 2015.
13. J. Kim, E. Lee, J. Choi, D. Lee, and S.H. Noh, *Chip-Level RAID with Flexible Stripe Size and Parity Placement for Enhanced SSD Reliability*, to appear *IEEE Transactions on Computers*, 2015.
14. L. Pamies-Juarez, F. Blagojevic, R. Mateescu, C. Gyuot, E. En Gad, Z. Bandic, *Opening the Chrysalis: On the Real Repair Performance of MSR Codes*, *FAST* 2016.

- **Real world storage system deployment**

15. J. Colgrove, J.D. Davis, J. Hayes, E.L. Miller, C. Sandvig, R. Sears, A. Tamches, N. Vachharajani, and F. Wang, *Purity: Building Fast, Highly-Available Enterprise Flash Storage from Commodity Components*, *SIGMOD* 2015.
16. M. Sathiamoorthy, M. Asteris, D. Papailiopoulos, A.G. Dimakis, R. Vadali, and S. Chen, *XORing Elephants: Novel Erasure Codes for Big Data*, *VLDB* 2013.
17. C. Huang, H. Simitci, Y. Xu, A. Ogus, B. Calder, P. Gopalan, J. Li, and S. Yekhanin, *Erasure Coding in Windows Azure Storage*, *ATC* 2012.