



The Repair Problem in Distributed Storage Systems

Rhodes Hall 310: April 10, 2013 @ 12:00PM



ISN Seminar Speaker:

Itzhak (Zachi) Tamo

Institute for Systems Research - University of Maryland

◇

Abstract

One of the characteristics of the Information era is the enormous amounts of data being generated and stored each day. Large data centers and distributed storage systems (DSS) are becoming more widespread, and they will have an increasing role in our everyday computational tasks.

Storage efficiency is of high importance in data centers, hence large scale DSS are currently transitioning to erasure codes. One of the prevalent problem that DSS are facing is the repair problem. Namely, repairing the lost data of a single failed disk. Classical codes like Reed-Solomon are suboptimal for distributed environments, since they provide poor performance during the repair process. Thus, new codes that better address this problem need to be constructed.

In the first part of the talk I will introduce formally the repair problem and an optimal repair code construction. In the second part I will address the limitations of data protection in such codes. Namely, for a given amount of redundancy, what is the maximum number of disks that can be protected in an optimal repair code. I will present an upper bound for the general case, and two tight bounds in the special cases of two important families of codes.

◇

Biography

Itzhak Tamo received his B.A. and B.Sc. degrees in 2008 from the Mathematics Department and the E.E. Department, respectively, Ben-Gurion University, Israel. Recently he submitted his Ph.D. thesis also to Ben-Gurion University. He is now a post-doc with the Institute for Systems Research at the University of Maryland. His research interests include algebraic coding with applications to storage systems, combinatorial structures, and nite group theory.

This is a joint work with Zhiying Wang and Prof. Jehoshua Bruck, both from Caltech.