Instructor: Alex Dimakis (www–rcf.usc.edu/~dimakis)
Lectures: Two 1hr and 20min lectures weekly

Course Description

Information theory has been phenomenally successful in finding the ultimate performance limits for the fundamental point-to-point communication problem formulated by Claude Shannon. The developed coding techniques have had a tremendous impact on centralized storage systems ranging from coding for magnetic storage to codes for RAID systems. This course is aimed at exposing the recent information theoretic methods developed for distributed storage problems, moving beyond point–to–point coding methods. This research has been fueled by technological innovations enabling massive data centers and distributed cloud storage.

The lectures will build up knowledge on advanced topics in coding theory developed for constrained systems and distributed RAID architectures including array codes, fountain codes and other sparse graph codes. Subsequently, graph theory and network coding will be introduced and the recent results on distributed storage coding will be developed. Theory for security and privacy in distributed storage will be developed, using both cryptographic and information theoretic methods.

Finally, the recent breakthrough technique of interference alignment will be introduced for both wireless and wired problems and the surprising applicability for distributed storage problems will be explained. Many open problems in this emerging area will be formulated and discussed.
Prerequisites
Information Theory (EE 565a) or instructor approval. The course will focus on proving techniques and analytical tools for theoretical research.

Grading
Homework: 40% Research Project: 60%
Students will be expected to perform a research project based on recent paper preprints on theoretic methods for distributed storage. The project will require a report and a short presentation. The choice of papers shall be agreed upon together with the instructor.

Textbooks
The course will be based in parts on recent research papers, and lecture notes. General overview papers:

Z. Zhang, A. Deshpande, X. Ma, E. Thereska, D. Narayanan
Does erasure coding have a role to play in my data center?

A.G. Dimakis, K. Ramchandran, Y. Wu, C. Suh,
"A Survey on Network Codes for Distributed Storage"

A.G. Dimakis, P.B. Godfrey, Y. Wu, M. Wainwright and K. Ramchandran
Network Coding for Distributed Storage Systems

On Secure Distributed Data Storage Under Repair Dynamics
S. Pawar, S. E. Rouayheb, and K. Ramchandran

Full online bibliography: http://tinyurl.com/storagecoding.