Read this paper:


You can focus on sections 1-5, but read section 6 if you are interested in experimental results, and section 7 if you are interested in related work.

Answer these questions:

1. Consider the “two-tiered bin packing” algorithm described in the paper, and the “greedy extended” algorithm described in the paper. What is one advantage of the two-tiered bin packing algorithm over the greedy extended algorithm? What is one disadvantage?

2. Consider a blogging application. Each user can create one or more blogs, and add postings to a blog over time. Some blogs turn out to be significantly more popular than others, but all blogs have some readers. When a new posting is added to a blog, it may be read multiple times, but once it becomes older, it will almost never be read. Imagine we are using E-store with the “first fit” algorithm, and we are trying to decide between two partitioning schemes. In option 1, we partition the data by blog, so that all the postings of a given blog are stored in the same partition. In option 2, we partition the data by posting, so that the postings of a given blog may be stored in separate tuples. Which option is likely to be better? Why?

3. Consider a social networking application. This application has two tables: Users, which lists individual users and is keyed by userid, and Friends, which lists pairs of users that are friends with each other. Thus, Users might contain a tuple with key “Mary” and another tuple with Key “Bob”, and if they are friends, the Friends table contains a tuple with key (“Mary”, “Bob”). Would E-store be a useful database system for this application? Why or why not?