Read the following three papers (on the course web page): Mutlu and Moscibroda, Ahn et al. and Isen and John.

(1) [40] Briefly summarize the papers first. When summarizing a paper, be specific about the key points made by the authors. For example, rather than saying “I find the paper interesting and it has lots of information”, say “The authors compare two branch prediction schemes: single-bit and two-bit. The single-bit scheme ... The two-bit scheme differs from the single-bit scheme in that ...”.

(2) [20] Mutlu and Moscibroda.
   (a) What does “parallelism-awareness” mean in the paper? What about “quality of service”?  
   (b) How are the concepts of memory level parallelism and multicore architecture related with the number of DRAM banks?  
   (c) What is the technical problem the paper is trying to solve? What are relevant metrics to evaluate solutions to the problem? Can you think of an alternative technique that directly or indirectly addresses the same problem? How does it compare with the proposed solution?  
   (d) Discuss strengths and weaknesses of the work.

(3) [20] Ahn et al.
   (a) What is “rank subsetting”?  
   (b) Without hardware support (i.e., in current commodity systems), would smart OS-level memory allocation help address the problem that this paper is concerned about or not? Justify your answer.  
   (c) Discuss strengths and weaknesses of the work.

(4) [20] Isen and John.
   (a) The authors propose to “explicitly” track memory allocation and de-allocation events. Compare in detail this explicit tracking and an “implicit” strategy to extract the same information from the OS (memory manager).  
   (b) Discuss how the system reliability could be compromised with the proposed ISA changes.  
   (c) Discuss strengths and weaknesses of the work.

Submit your work at the class or directly to the mailbox of the instructor (box #276), located in the mail room on 5th floor, SENSQ.