

In the paper “Memory Scaling: a Systems Architecture Perspective” Mutlu describes that the memory system in a computer has been a major bottleneck for performance. They also accompany this with having an explanation of the problems that DRAM technology faces and some potential solutions for better memory systems. Ultimately the authors describe that the large issue at hand with DRAM comes with scaling them into smaller technology standards, the second issue being that emerging technologies to replace DRAM also come with various shortcomings.

They continue with solutions for current DRAM technologies. One of their ideas being that since an issue of scaling is that a lot more energy is required simply refreshing memory which require higher refresh rates than other portions of the DRAM. However, this is easier said than done as doing so would require information based on Data Pattern Dependence and Variable Retention time. A few other ideas they propose are: exploiting the inherent structure that is found within DRAM as multiple access can be pipelined from the same block but different words with the inclusion of some Logic and a relatively small area overhead, creating different subsections within DRAM which run at different speeds and consume different amounts of area like a Cache hierarchy.

The second solutions that they see being implemented in order to overcome the Memory bottleneck revolving around emerging technologies. They go on to further explain and describe the benefits of some of the emerging technologies as well as the issues that each one has. One solution of course is to create a Hybrid system which uses older DRAM technology along with new technologies. In this same front they explore the ideas of new possibilities with using new technologies but explain that they can come with their own shortcomings such as security issues.

One thing which I think is important to note as a reader is that in their section regarding Predictable performance they mention concepts similar to those from the approximate computing paper we read a while ago. That being, the system must not be perfect but rather we need to optimize the system to have a good enough performance in which users wouldn't be concerned with the final outcome.

Overall, while I think that this paper is good and has some interesting points regarding the bottleneck which is the memory system in a computer, I wish that there were more information about how these concepts may even be conducted. A lot of the paper presents different ideas which maybe used to but lacks in figures and hard data which is understandable

given that this paper explores ideas more than understanding one idea very well. Thus i think I would not change anything about this paper.