Abstract

Name of the student: **Subisha V** Roll No: **17210107**

Degree for which submitted: **Master of Technology**

Department: **Computer Science and Engineering**

Thesis title: **Prefetching in Hybrid Main Memory Systems**

Thesis supervisor: **Manu Awasthi**

Month and year of thesis submission: **June 28, 2019**

Hybrid memory systems, the integration of DRAM and non volatile memory (NVM) technologies is a promising replacement to the current DRAM system. Data placement in such integrated memory systems is really important, as every miss in DRAM might be serviced by NVM at a latency much higher than DRAM. A cache based hybrid main memory system can perform even worse than a DRAM and a Non Volatile Memory system in the case of poor data placement. Alloy cache is a cache based hybrid main memory system, where the DRAM acts as a direct mapped cache to the larger NVM. Previous proposals like Alloy cache, which uses DRAM as a cache for NVM can suffer from higher latencies due to tag serialization and have high space overheads.

We propose a DRAM cache based hybrid main memory design that prefetches relevant data at both page and cacheline granularity from PCM to DRAM. This design improves the utilization of DRAM, by placing prefetched pages to locations that are otherwise not used due to its direct mapped structure. Additionally every 4KB sized prefetched pages in DRAM eliminate the loss of 8 cachelines per page on tag

storage. Our evaluation on PARSEC benchmark suite provides on an average 68% improvement for single threaded applications and 22% for multithreaded applications. The architecture is able to reduce the average memory latency by 30% for single threaded and 5% for multithreaded applications as compared to the state of the art Alloy cache model.