**DINH DUY KHA**
**2019712308**

## Traffic Management: A Holistic Approach to Memory Placement on NUMA Systems

This paper discovered that congestion on memory controllers and interconnects is the bigger bottleneck in NUMA systems compared to remote access. The authors also propose *Carrefour*, a memory placement algorithm that is optimized for traffic congestion. The algorithm consists of four components: page co-location, page interleaving, page replication and thread clustering. Evaluation results show that the algorithm boosts performance up to 3.6 times and never hurts the performance by more than 4%.

The cost of remote wireless delay which the main focus in earlier NUMA-aware systems, is no longer a concern in modern systems. Instead, the authors of this paper focus on optimizing memory for traffic congestion which yield much better performance compared to previous approaches. The authors also back up this claim with sufficient evaluations and experiments.

Even if remote access is not the main performance bottleneck, researches had shown that it is the biggest bottleneck in synchronization scalability. Placing memory only based on avoiding congestion could hurt performance on synchronization intensive workloads.