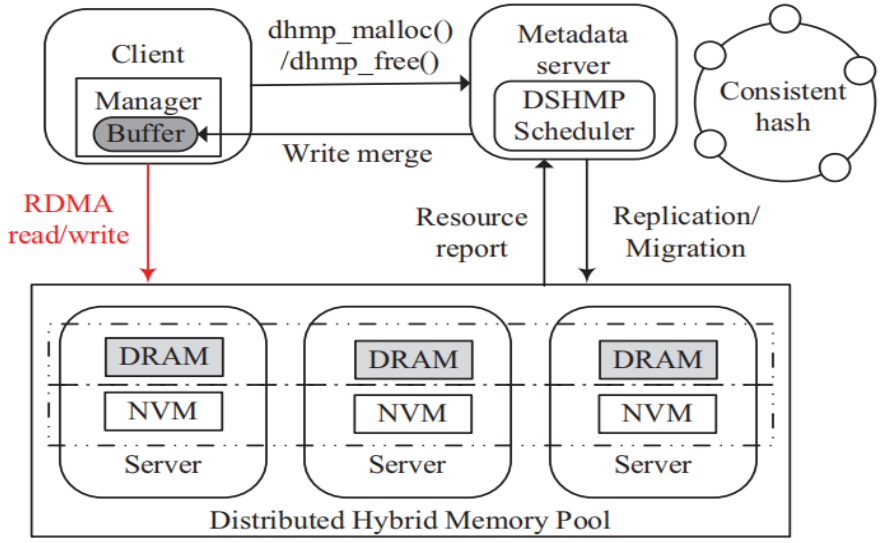
Resource abstraction and data placement for distributed hybrid memory pool

Tingting CHEN, Haikun LIU , Xiaofei LIAO, Hai JIN

Frontiers of Computer Science, DOI: 10.1007/s11704-020-9448-7

Problems & Ideas

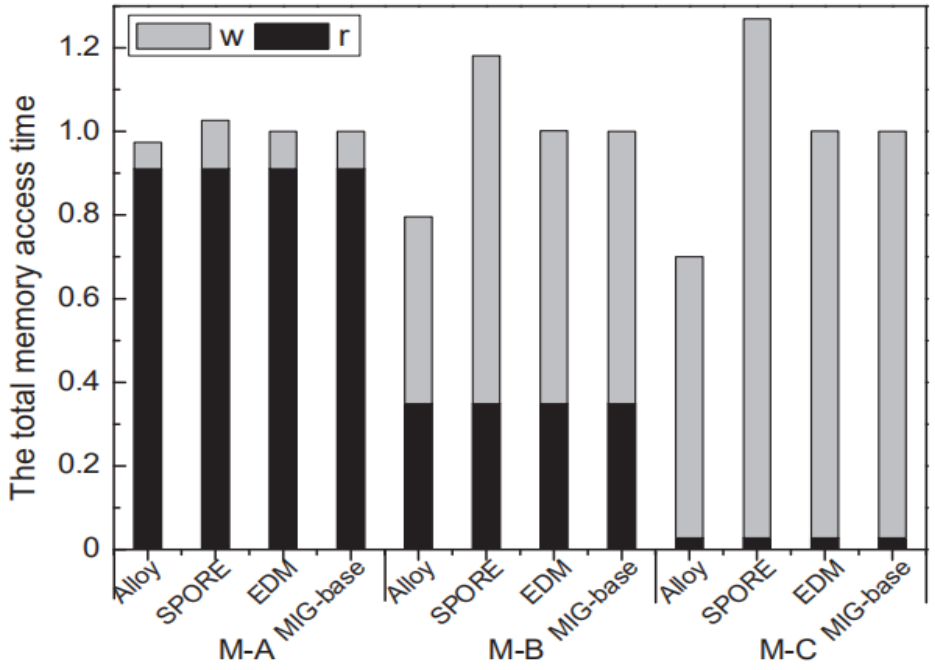
- Hybrid NVM/DRAM management poses several challenges in a distributed environment.
  - Since the access latency and bandwidth are diverse for different memories in a Distributed Hybrid Memory Pool (DHMP), applications have many choices to place data in a DHMP.
  - Previous hybrid memory management strategies on a single server do not consider data placement across servers and load balancing, and are not applicable to a DHMP.
- Ideas: A memory resource abstraction and data placement strategy for a DHMP
  - Alloy manages disturbed memory resources in the global memory space through a high-level abstraction in the metadata server.
  - Objects are classified into fine-grained categories according to memory access behaviors, and placed in the DHMP according to a hotness-aware data placement scheme which combines hot data replication, write merging and global greedy based data migration strategies together.



Architecture of Alloy

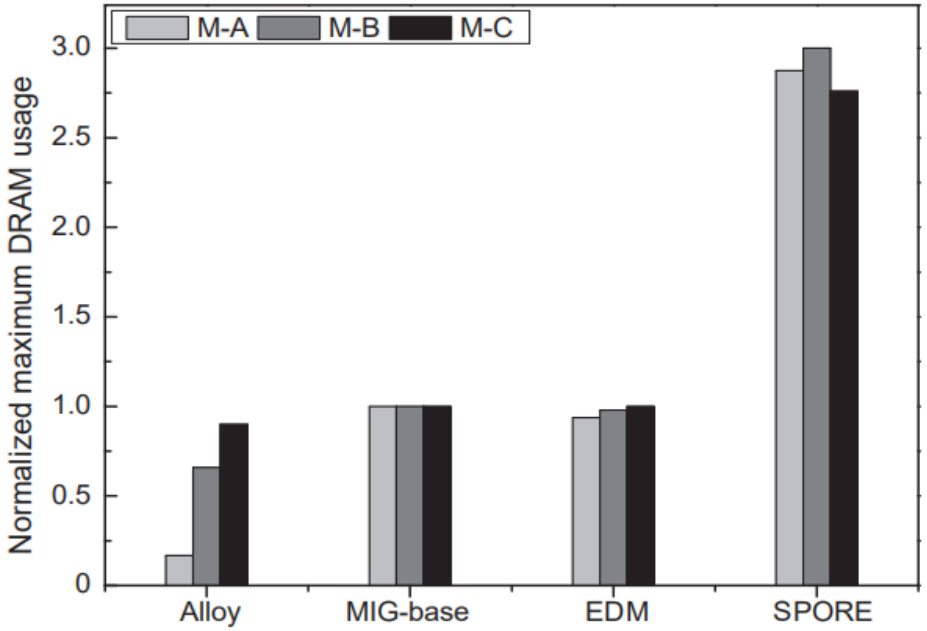
Main Contributions

- **Performance improvement**



When the proportion of NVM writes increase, Alloy introduces more performance improvement. Alloy can reduce the total memory access time by up to 57%.

- **DRAM usage reduction**



Alloy reduces the DRAM usage by up to 84% compared to MIG-base and EDM, and by 69%-95% compared to SPORE.