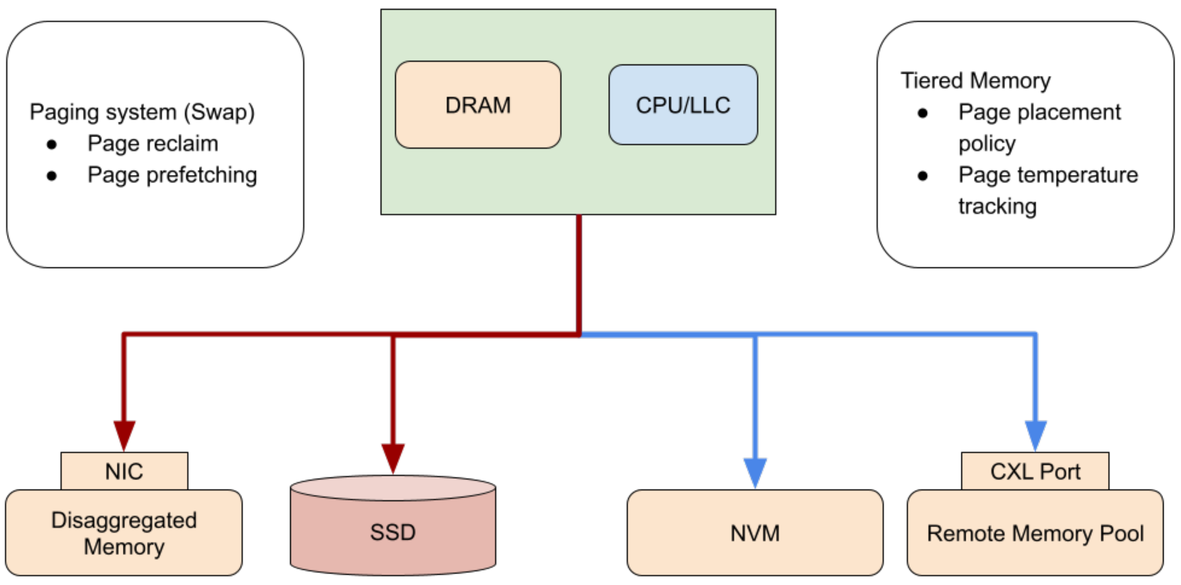
µswap: A Semi-microkernel for Memory Disaggregation

Sepehr Jalalian, Margo Seltzer, Alexandra Fedorova
University of British Columbia

Problem

- DRAM accounts for more than 30% of datacenter server cost [1]
- Disaggregated memory improves resource utilization in datacenters
- Leveraging disaggregated memory imposes many challenges on the system software [2]



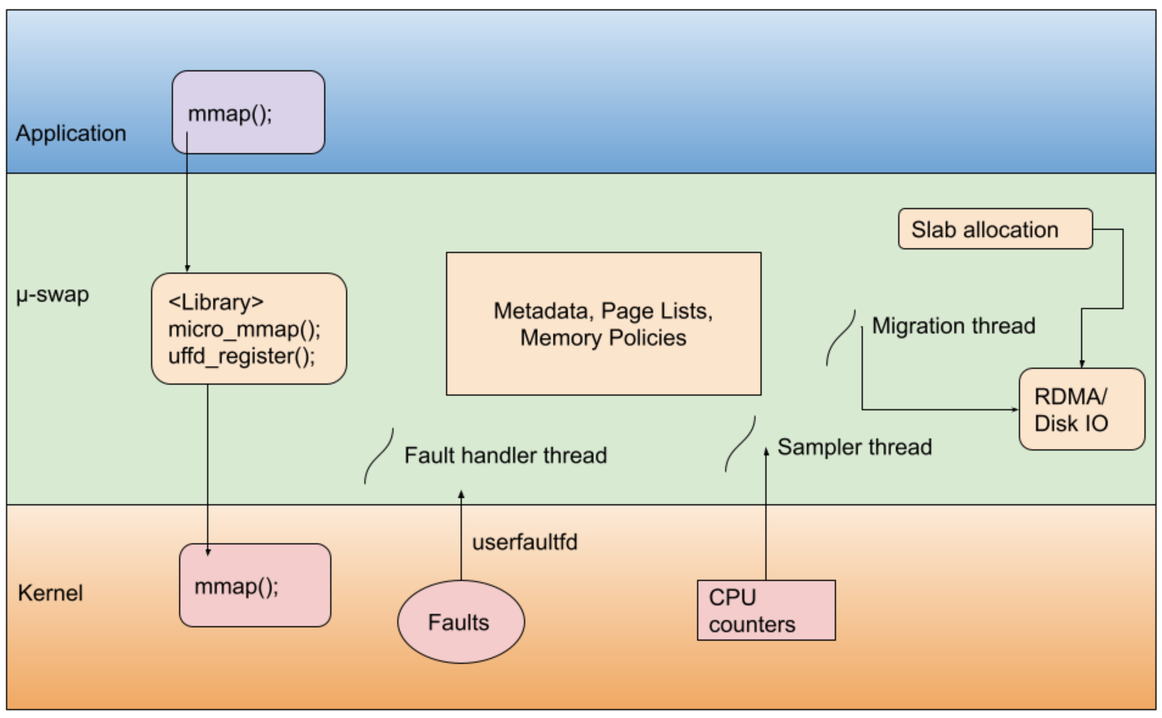
- Two main types of systems developed for memory disaggregation have architectural differences:
  - **Tiered memory systems** (NVM, CXL pools) [3]
  - paging systems based on **swap** (network, disk) [1]

Research Questions

- Do tiered memory systems and paging systems require two separate kernel subsystems?
  - Both do page migration
  - Both need page placement policies [4]
- **How should system software support disaggregated memory?**

Our Approach

- **µswap**: a semi-microkernel prototype for disaggregated memory management.
  - Enables user-level memory management in cooperation with a monolithic kernel.
  - Enables fast development and testing of memory management strategies.
  - A unifying framework for page migration that supports different types of second-tier and remote memory.



➢ Memory semi-microkernel: Inspired by the microkernel architecture, we offload memory management to a user-space server. Semi-microkernel (1) intercepts application memory requests, (2) keeps page information and (3) applies memory policies.

[1] Weiner et al, Tmo: transparent memory offloading in datacenters. ASPLOS'22
[2] Yizhou Shan et al. LegoOS: A disseminated, distributed OS for hardware resource disaggregation. OSDI'18
[3] Amanda Raybuck et al. Hemem: Scalable tiered memory management for big data applications and real NVM. SOSP'21
[4] Zi Yan et al. Nimble page management for tiered memory systems. ASPLOS'19