Intelligent and Cost-Efficient Data Management for Hybrid Memory Systems

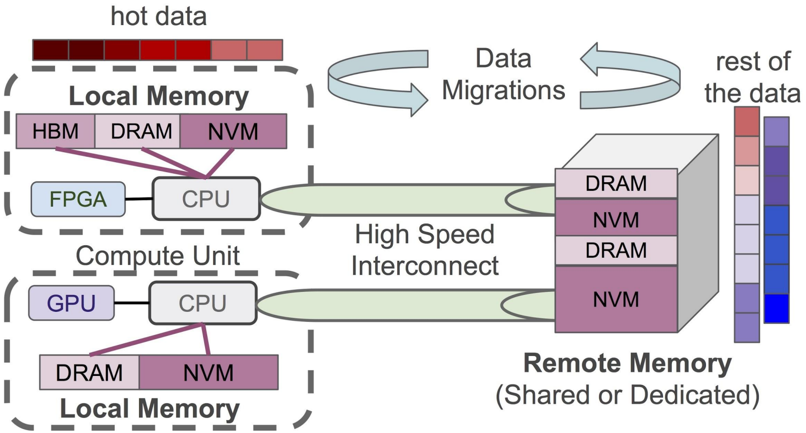
Thaleia Dimitra Doudali

@ EuroDW 2020

Problem Space

1. Hybrid Memory Systems:

Provide Big Data analytics with high memory capacity at a reasonable cost using emerging memory hardware or by disaggregating the available resources.



Example of hybrid (disaggregated) memory system.

2. Challenge:

The high variability in the access speeds of the heterogeneous memory units, leads to significant application performance loss.

3. Data Management Approach:

Dynamically move frequently accessed data in the fastest memory component.

4. Problem:

- Which data to move?
- When and how much data to move?

Georgia Tech
CREATING THE NEXT

Thesis Contributions

Problem 1:
Which data to move?

Solution 1:
Kleio: *a Hybrid Memory Page Scheduler with Machine Intelligence*.
[Best Paper Award Finalist at HPDC ’19]

Problem 2:
When and how much data to move?

Solution 2:
Terpsichore: *Cost-Efficient Data Movements for Hybrid Memory Systems*.
[Ongoing Work]

Kleio [Completed Work]

Problem : Which data to move?

Existing Solutions:
Use history-based information to predict which pages are frequently accessed and move them to a fast memory. That can lead up to 50% performance loss!

Approach:

- Find a Machine Learning (ML) method that can better predict which pages are frequently accessed.
- Practically integrate ML into a system-level data management solution.

Contributions:

- Kleio deploys *Recurrent Neural Networks* to learn memory access patterns.
- Kleio uses a hybrid of ML and historical predictions to identify frequently accessed pages.
- Kleio reduces by 80% the existing performance gap between historical and oracular solutions.
- Kleio practically integrates ML by cleverly selecting a small page subset whose ML management significantly boosts application performance.

Terpsichore [Ongoing Work]

Problem : When and how much data to move?

Existing Solutions:

Selection of data movements:

- Which: frequently accessed pages.
- When: periodically.
- How much: as many as the available bandwidth allows during that period.

Significant performance loss due to possible stalls by migration ordering and shared bandwidth use.

Approach

- Cost-benefit analysis of data movements.
- What system-level information is needed to allow for cost-efficient migrations?
  - Capture that into a utility metric used upon deciding whether to perform a data movement.
- What system-level support is currently missing?

Thesis Publications

CoMerge: *Toward Efficient Data Placement in Shared Heterogeneous Memory Systems.*
[MemSys ‘17]

Mnemo: *Boosting Memory Cost Efficiency in Hybrid Memory Systems.*
[HPBDC @ IPDPS ‘19]

Kleio: *a Hybrid Memory Page Scheduler with Machine Intelligence*.
[HPDC ’19]

Terpsichore: *Cost-Efficient Data Movements for Hybrid Memory Systems*.
[Ongoing Work]