Lecture 3 of the
MLArchSys Seminar

Instructor: Thaleia Dimitra Doudali
Assistant Professor at IMDEA Software Institute
Universidad Politécnica de Madrid (UPM)
March 2023
Outline of Today’s Lecture

Today’s Paper:

**Kleio: A Hybrid Memory Page Scheduler with Machine Intelligence**

Thaleia Dimitra Doudali
Georgia Institute of Technology
thdoudali@gatech.edu

Sergey Blagodurov
Advanced Micro Devices, Inc.
Sergey.Blagodurov@amd.com

Abhinav Vishnu
Advanced Micro Devices, Inc.
Abhinav.Vishnu@amd.com

Sudhanva Gurumurthi
Advanced Micro Devices, Inc.
Sudhanva.Gurumurthi@amd.com

Ada Gavrilovska
Georgia Institute of Technology
ada@cc.gatech.edu

Lecture Outline:

- Hybrid Memories
- Hybrid Memory Management
- How to Integrate ML?
- System Design
- Evaluation
- Lessons Learned
Outline of Today’s Lecture

Today’s Paper:

**Kleio: A Hybrid Memory Page Scheduler with Machine Intelligence**

<table>
<thead>
<tr>
<th>Authors</th>
<th>Affiliations</th>
</tr>
</thead>
<tbody>
<tr>
<td>Thaleia Dimitra Doudali</td>
<td>Georgia Institute of Technology</td>
</tr>
<tr>
<td><a href="mailto:thdoudali@gatech.edu">thdoudali@gatech.edu</a></td>
<td></td>
</tr>
<tr>
<td>Sergey Blagodurov</td>
<td>Advanced Micro Devices, Inc.</td>
</tr>
<tr>
<td><a href="mailto:Sergey.Blagodurov@amd.com">Sergey.Blagodurov@amd.com</a></td>
<td></td>
</tr>
<tr>
<td>Abhinav Vishnu</td>
<td>Advanced Micro Devices, Inc.</td>
</tr>
<tr>
<td><a href="mailto:Abhinav.Vishnu@amd.com">Abhinav.Vishnu@amd.com</a></td>
<td></td>
</tr>
<tr>
<td>Sudhanva Gurumurthi</td>
<td>Advanced Micro Devices, Inc.</td>
</tr>
<tr>
<td><a href="mailto:Sudhanva.Gurumurthi@amd.com">Sudhanva.Gurumurthi@amd.com</a></td>
<td></td>
</tr>
<tr>
<td>Ada Gavrilovska</td>
<td>Georgia Institute of Technology</td>
</tr>
<tr>
<td><a href="mailto:ada@cc.gatech.edu">ada@cc.gatech.edu</a></td>
<td></td>
</tr>
</tbody>
</table>

Lecture Outline:

- Hybrid Memories
- Hybrid Memory Management
- How to Integrate ML?
- System Design
- Evaluation
- Lessons Learned
Big Data hits the Memory Wall

Applications generate huge amounts of data.

Processing speed grows faster than memory and data transfer speed.

Need for larger and faster memory configurations.
New Memory Technologies

We are in the era of Hybrid Memory Systems. A mix of different technologies at different speeds / capacities / costs.
In today’s paper we assume a hybrid memory system with **DRAM** and **NVM** (Non Volatile Memory).

The NVM actual product was released by Intel in 2019, after the paper was published. So, the paper had to assume various possible configurations, e.g., capacity ratios.

Intel Optane is packaged together with DRAM. Guess what? The product got discontinued in 2022..

**Intel kills off Optane Memory, writes off $559 million inventory**

In a terrible quarter for the chip giant

July 29, 2022  By: Sebastian Moss  1 Comment

But this is not the end for hybrid memory configurations. It is a mix of any number of fast vs. slow, small vs. big memories.
Outline of Today’s Lecture

Today’s Paper:

**Kleio: A Hybrid Memory Page Scheduler with Machine Intelligence**

Thaleia Dimitra Doudali  
Georgia Institute of Technology  
thdoudali@gatech.edu

Sergey Blagodurov  
Advanced Micro Devices, Inc.  
Sergey.Blagodurov@amd.com

Abhinav Vishnu  
Advanced Micro Devices, Inc.  
Abhinav.Vishnu@amd.com

Sudhanva Gurumurthi  
Advanced Micro Devices, Inc.  
Sudhaaev.Gurumurthi@amd.com

Ada Gavrilovska  
Georgia Institute of Technology  
ada@cc.gatech.edu

Lecture Outline:

- Hybrid Memories
- Hybrid Memory Management
- How to Integrate ML?
- System Design
- Evaluation
- Lessons Learned

ML for Systems

Machine Learning

for Hybrid Memory Management (HMem Management)
OS-level Memory Management

Application

Data = 1 GB

Operating System (OS)

Memory Hardware (HW)

Virtual Address Space

Physical Address Space

Page Table

Page = 1
Page = 2
Page = N

Page = 4 KB

1 GB

Virtual Address Space

Physical Address Space

Ongoing
Hybrid Memory Page Allocation

Example: Let’s assume each application uses 3 GB. And hybrid memory consists of:

The OS allocates the pages in memory using the “first touch” policy.

If the memory allocation doesn’t change throughout time, then we get no use out of the fast memory.
Hybrid Memory Page Scheduling

The OS should move pages dynamically across hybrid memory to maximize the efficiency.

The page *hotness* changes through time. Keep the hot pages in fast memory through time.

The page *hotness* changes through time.
Page Scheduling as a Prediction Problem

1. Page Access Monitor
   Keep track of page hotness.
   In every time period

2. Page Hotness Predictor
   Predict future page hotness, based on past access history.
   Past Page Hotness

3. Page Movement Selector
   Choose which pages to move across hybrid memory.
   Future Page Hotness

Applications

OS-level Page Scheduler

Hardware

Hybrid Memory

Page Movements
OS uses a very simple **history**-based predictor, to minimize operational overheads.

Predicts for all pages that page hotness at period $p_N = \text{page hotness at period } p_{N+1}$

Need something more clever, to close this big gap. How effective would Machine Learning be?
Outline of Today’s Lecture

Today’s Paper:

**Kleio: A Hybrid Memory Page Scheduler with Machine Intelligence**
Thaleia Dimitra Doudali
Georgia Institute of Technology
thdoudali@gatech.edu

Sergey Blagodurov
Advanced Micro Devices, Inc.
Sergey.Blagodurov@amd.com

Abhinav Vishnu
Advanced Micro Devices, Inc.
Abhinav.Vishnu@amd.com

Sudhanya Gurumurthi
Advanced Micro Devices, Inc.
Sudhanya.Gurumurthi@amd.com

Ada Gavrilovska
Georgia Institute of Technology
ada@cc.gatech.edu

**LSTM**

Lecture Outline:

- Hybrid Memories
- Hybrid Memory Management
- How to Integrate ML?
- System Design
- Evaluation
- Lessons Learned

ML for Systems

Machine Learning

for Hybrid Memory Management (HMem Management)
The Vision: Integrate ML in Page Scheduling

Applications

OS-level Page Scheduler

Hardware

Data Access Monitoring

Re-train?

Memory Management

ML training

prediction

Trained Models

Hybrid Memory

Memory Management

practical

no

yes

Re-train?

Applications

OS-level Page Scheduler

Hardware

Data Access Monitoring

Re-train?

Memory Management

ML training

prediction

Trained Models

Hybrid Memory

practical

no

yes

Re-train?
Toward Realizing the Vision: Questions to Ask

Applications

Exploded Data Sizes

Which data needs ML?

Where to use ML?

Which ML method?

What to Predict?

ML training

Trained Models

Yes

No

Hybrid Memory

OS-level Page Scheduler

Memory Management

How to practically use ML?
Where to Use ML? (1)

Replace the Page Scheduler with a Reinforcement Learning (RL) agent.

Learn the Action: Learn from moving pages across hybrid memory.
Learn from mistakes (e.g., cold pages in DRAM).

Why it is not a good fit:
• Exponential Action Space = $2^N$, when moving N pages across 2 memories.
• Need to re-train if configuration of hybrid memory changes.
  • Number of memory units.
  • Difference in access speeds / capacities.

Not practical / scalable.
Don’t learn the action!
Learning Memory Access Patterns

Don’t learn the action.. Learn the memory access pattern?

LSTMs (Long Short Term Memory) Networks are used in timeseries forecasting.

Can we do something similar?
Applications

1. Page Access Monitor
   Keep track of page hotness.

2. Page Hotness Predictor
   Use LSTM models to predict which pages will be accessed in the next period. Then, calculate page hotness.

3. Page Movement Selector
   Choose which pages to move across hybrid memory.

OS-level Page Scheduler

Past Page Hotness

Future Page Hotness

Hardware

Page Movements

Hybrid Memory
Learning the Page Access Sequence

Learn **which** pages will be accessed in the next period of time, given a window of history.

**input**

Page ID $t_{N-1}$, Page ID $t_{N-2}$, ...

Page ID $t_{N-h}$

**LSTM-based Deep Neural Network**

**output**

Page ID $t_N$

<table>
<thead>
<tr>
<th>No. models</th>
<th>Overheads</th>
<th>Accuracy</th>
</tr>
</thead>
<tbody>
<tr>
<td>1 per app</td>
<td>Huge Memory Trace Files. Days to train. Months to fine-tune.</td>
<td>Low, because millions of pages. Top-k predictions not useful.</td>
</tr>
</tbody>
</table>
Learning the Page Hotness

Learn **how hot** a page will be in the next period of time, given a window of history.

**input**

- Page Hotness $t_{N-1}$
- Page Hotness $t_{N-2}$
- ...  
- Page Hotness $t_{N-h}$

**LSTM-based Deep Neural Network**

**output**

Page Hotness $t_N$

<table>
<thead>
<tr>
<th>No. models</th>
<th>Overheads</th>
<th>Accuracy</th>
</tr>
</thead>
<tbody>
<tr>
<td>Many per app. 1 per page.</td>
<td>Smaller models. Parallel Training. Faster to train and fine-tune.</td>
<td>High, because hotness value is smaller, depending on the period length.</td>
</tr>
</tbody>
</table>

**Challenge:** 1 model per page, means millions of models.

Flip the way you look at a problem!
Do all Pages Need ML?

Probably the pages that are part of a pattern, need ML-based management. For the rest, the simple history page scheduler works well (page hotness at period $p_N = \text{page hotness at period } p_{N+1}$).

Which pages really need ML?

Page suddenly becomes very hot. History scheduler will misplace it (not in DRAM when hot).

Use ML for subset of pages, and the existing history scheduler for the rest.
Proposed ML Integration

- **Page Selector**
  - Priority
  - Page hotness predictions
  - Pool of Pages
  - History

- **ML Integration**
  - Apply ML on a small page subset.
  - Foundations for practical use of ML.
  - Carefully select pages for ML.
  - Application performance boost.
Outline of Today’s Lecture

Today’s Paper:

**Kleio: A Hybrid Memory Page Scheduler with Machine Intelligence**

Thaleia Dimitra Doudalidou
Georgia Institute of Technology
thdoudalidou@gatech.edu

Sergey Blagodurov
Advanced Micro Devices, Inc.
Sergey.Blagodurov@aamd.com

Abhinav Vishnu
Advanced Micro Devices, Inc.
Abhinav.Vishnu@aamd.com

Sudhanva Gurumurthi
Advanced Micro Devices, Inc.
Sudhanva.Gurumurthi@aamd.com

Ada Gavrilovska
Georgia Institute of Technology
ada@cc.gatech.edu

**Lecture Outline:**

- Hybrid Memories
- Hybrid Memory Management
- How to Integrate ML?
- System Design
- Evaluation
- Lessons Learned

**Today's Paper: Kleio: A Hybrid Memory Page Scheduler with Machine Intelligence**

- **ML for Systems**
- **Machine Learning**

- **LSTMs**

- **Hybrid Memory Management (HMem Management)**
System Design of Kleio

### Offline Phase
1. Run Page Selector.
2. Train LSTMs.

### Online Phase
1. **Page Access Monitor**
   - Keep track of page hotness.
2. **Page Hotness Predictor**
   - LSTM inference
   - History predictions
   - In every time period
3. **Page Movement Selector**
   - Choose which pages to move across hybrid memory.

### Hardware
- Page Movements
- Hybrid Memory

---

**Greek Trivia:** According to the ancient Greek mythology, Kleio was the muse of history, daughter of Mnemosyne, goddess of memory.
Which Pages to Select?

Need to create a metric to select pages.

Per Page Benefit Factor

\[ \text{benefit} = \# \text{ accesses} \times \# \text{ misplacements} \]

100% Accurate ML

History

Page Hotness Prediction

Select a **small subset** of pages in the **order** that brings the desired **performance** level.
Page Selector Design

Page Selector calculates internally the performance curve, using a performance estimate analytical model.

It is not a lightweight process, but necessary to deliver the desired application performance levels.
Outline of Today’s Lecture

Today’s Paper:

**Kleio: A Hybrid Memory Page Scheduler with Machine Intelligence**

Thaleia Dimitra Doudali  
Georgia Institute of Technology  
thedoudali@gatech.edu

Sergey Blagodurov  
Advanced Micro Devices, Inc.  
Sergey.Blagodurov@amd.com

Abhinav Vishnu  
Advanced Micro Devices, Inc.  
Abhinav.Vishnu@amd.com

Sudhanva Gurumurthi  
Advanced Micro Devices, Inc.  
Sudhanva.Gurumurthi@amd.com

Ada Gavrilovska  
Georgia Institute of Technology  
ada@cc.gatech.edu

---

Lecture Outline:

- Hybrid Memories
- Hybrid Memory Management
- How to Integrate ML?
- System Design
- Evaluation
- Lessons Learned
Effect on Application Performance

- **Best** Solution
  - 100% Accurate ML for the Selected Pages
  - LSTMs for the Selected Pages
  - Baseline Solution
  - History-based for all Pages

- **Kleio**
  - History-based for all Pages
  - Rest of Pages

For half of the applications, Kleio reaches **95%** the possible performance levels!

*For 100 selected pages.
Effect on Quality of Page Placement

Kleio reduces more than 80% of the page misplacements, due to the improved page movement decisions, via the more accurate page hotness predictions.
LSTM Prediction Accuracy

e.g. MAE = 50 means that the RNN predicted on average 50 more accesses per scheduling epoch per page.

High prediction error does not impact application performance, when it does not affect the quality of page movement decision.
Outline of Today’s Lecture

Today’s Paper:

**Kleio: A Hybrid Memory Page Scheduler with Machine Intelligence**

Thaleia Dimitra Doudali  
Georgia Institute of Technology  
thedoudali@gatech.edu

Sergey Blagodurov  
Advanced Micro Devices, Inc.  
Sergey.Blagodurov@amd.com

Akhinav Vishnu  
Advanced Micro Devices, Inc.  
Akhinav.Vishnu@amd.com

Sudhanya Gurumurthi  
Advanced Micro Devices, Inc.  
Sudhanya.Gurumurthi@amd.com

Ada Gavrilovska  
Georgia Institute of Technology  
ada@cc.gatech.edu

LSTMs

**for** Hybrid Memory Management (HMem Management)

**Lecture Outline:**

- Hybrid Memories
- Hybrid Memory Management
- How to Integrate ML?
- System Design
- Evaluation
- Lessons Learned
Lessons Learned

1. Understand what would be the benefit from using ML.

“History” = x%

Performance Gap

“Oracle” = 0%
Perfect Predictions

2. Learn the Behavior, not the Action, for the use case of hybrid memory management.

Replacing the Page Scheduler with an RL agent would not be practical, nor scalable.
Lessons Learned

3. Look at the same problem from a different angle.

Learn **which** pages will be accessed next.

4. High prediction error of page hotness **does not impact** application performance, when it does not affect the quality of page movement decision.

Learn **how hot** a page will be next.
Lessons Learned

5. Use existing solutions to the best of their ability, and deploy ML where necessary.

Apply ML on a small page subset.
- Foundations for practical use of ML.

Carefully select pages for ML.
- Application performance boost.
Improvements on Kleio

Can we reduce the number of LSTMs via clustering?

**Coeus**: Clustering (A)like Patterns for Practical Machine Intelligent Hybrid Memory Management. [CCGrid 2022]

Can we accelerate the page selection process?

**Cronus**: Computer Vision-based Machine Intelligent Hybrid Memory Management. [MEMSYS 2022]
Report Due April 4 at 18.00

Answer / expand upon these 4 questions:

1. What problem is the paper addressing and why is it important?
2. How do they approach to solve the problem?
3. What are the main evaluation results?
4. What are 2 things you will remember from this paper?

Contact
• Via email: thaleia.doudali@imdea.org

Website
https://thaleia-dimitradoudali.github.io/

Teaching

Spring 2023
MLArchSys Seminar Series.
At the MUSS and EMSE Master Programs of the School of Computer Science at Universidad Politécnica de Madrid. MUSS Link EMSE Link
Seminar 2: Machine Learning for Cache Prefetching. Slides Paper
Seminar 3: Machine Learning for Hybrid Memory Management. Slides Paper