CHEOPS Workshop at EuroSys 2025

Workshop on Challenges and Opportunities of Efficient and Performant Storage Systems

CHEOPS '25

Monday, March 31ˢᵗ, 2025

Rotterdam, Netherlands

< EURO / SYS'25 >

CHEOPS'25 Website

General Chair

Suren Byna, *The Ohio State University, USA*

Program Chairs:

Amelie Chi Zhou, *Hong Kong Baptist University, Hong Kong*

Kira Duwe, *École Polytechnique Fédérale de Lausanne (EPFL), Switzerland*

Welcome!

The main objective of CHEOPS to provide a forum for researchers, developers of scientific applications, and engineers to discuss state-of-the-art R&D, innovative ideas, and experiences that focus on the design and implementation of storage systems.

CHEOPS'25 Website

< EURO / SYS'25 >

CHEOPS'25:
The fifth Workshop on Challenges and Opportunities of Efficient and Performant Storage Systems
https://cheops-workshop.github.io

Program highlights!

▶ **Keynote speech**
  ▶ *Gustavo Alonso (ETH Zurich, Switzerland)*

▶ **4 full paper presentations and 2 Work-in-Progress papers**

▶ **3 Invited Speakers**
  ▶ *Jean-Thomas Acquaviva (DDN Storage)*
  ▶ *Yang Zheng (Huawei Technologies)*
  ▶ *Shadi Ibrahim (Inria, Rennes)*

▶ ***ACM proceedings***
  ▶ *https://go.osu.edu/cheops25*

CHEOPS'25 presentations

< EURO / SYS'25 >

CHEOPS'25:
The fifth Workshop on Challenges and Opportunities of Efficient and Performant Storage Systems
https://cheops-workshop.github.io

Time	Speaker / Authors (Affiliation)	Content
8:30-9:00		**Coffee and Registration**
9:00-9:10	Suren Byna	CHEOPS Welcome
9:10-10:00	**Keynote - <u>Gustavo Alonso</u> (ETH Zurich, Switzerland)**	**Vertically integrated storage systems**
10:00-10:30	*Louis-Marie NICOLAS (Lab-STICC, CNRS UMR 6285, ENSTA, Institut Polytechnique de Paris)*, Salim MIMOUNI (Atos BDS R&D Data Management), Philippe COUVEE (Atos BDS R&D Data Management), Jalil BOUKHOBZA (Lab-STICC, CNRS UMR 6285, ENSTA, Institut Polytechnique de Paris)	Characterizing the Use of DVFS for HPC I/O Optimization: A Microbenchmarking Approach
10:30-11:00		**Coffee break**
11:00-11:30	*Robin Vonk (Delft University of Technology)*, Joost Hoozemans (Voltron Data), Zaid Al-Ars (Delft University of Technology)	GSST: Parallel string decompression at 191 GB/s on GPU
11:30-12:00	*Shadi Ibrahim (Inria, Rennes)*, Jad Darrous (Inria, Rennes)	Erasure Coding Aware Block Placement for Data-Intensive Applications
12:00-12:30	**Invited talk - <u>Jean-Thomas Acquaviva</u> (DDN Storage)**	From HPC to AI: A Data Journey
12:30-14:00		**Lunch break**
14:00-14:30	Zebin Ren (Vrije Universiteit Amsterdam), Krijn Doekemeijer (Vrije Universiteit Amsterdam), Tiziano De Matteis (Vrije Universiteit Amsterdam), Christian Pinto (IBM Research Europe), Radu Stoica (IBM Research Europe), Animesh Trivedi (IBM Research Europe)	An I/O Characterizing Study of Offloading LLM Models and KV Caches to NVMe SSD
14:30-15:00	**Invited talk - <u>Yang Zheng</u> (Huawei Technologies)**	**Reliability challenges and opportunities for AI infra: from industry perspective**
15:00-15:30	**Invited talk - <u>Shadi Ibrahim</u> (Inria, Rennes)**	**Scalable and Efficient Big Data Processing in Clouds: Addressing Performance Variability**
15:30-16:00		**Coffee break**
16:00-16:15	Joost Hoozemans (Voltron Data, Delft University of Technology), Robin Vonk (Delft University of Technology), Johan Peltenburg (Voltron Data), Felipe Aramburu (Voltron Data), Zaid Al-Ars (Delft University of Technology)	Using GPU Direct Storage with High-Performance Distributed Filesystems
16:15-16:30	*Pınar Tözün (IT University of Copenhagen)*, Karl B. Torp (Samsung Denmark Research Center), Simon A. F. Lund (Samsung Denmark Research Center)	A Quest to Reduce Dependency on CPUs in Deep Learning Data Pipelines
16:30-16:50	All participants	Discussion
16:50-17:00	CHEOPS Organizers	Closing remarks
18:00-19:30		**Welcome Reception**

Thanks to the combined efforts of this year's workshop team.

- **Jean-Thomas Acquaviva**, *DDN Storage, France*

- **Jalil Boukhobza**, *ENSTA Bretagne, Lab-STICC CNRS UMR 6285, France*

- **Konstantinos Chasapis**, *DataDirect Networks (DDN) Storage, France*

- **Kira Duwe**, *EPFL, Switzerland*

- **Shadi Ibrahim**, *Inria, Rennes*

- **Michael Kuhn**, *Otto von Guericke University Magdeburg (OVGU), Germany*

Big thanks to the program committee.

- Anastasios Papagiannis, Isovalent at Cisco

- Animesh Trivedi, IBM Research Europe, Zurich

- Christos Kozanitis, FORTH-ICS

- Diana Moise, Hewlett Packard Enterprise (HPE)

- Jalil Boukhobza, ENSTA Bretagne

- Jay Lofstead, Sandia National Laboratories

- Jean Luca Bez, Lawrence Berkeley National Laboratory

- Jerry Chou, National Tsing Hua University

- Marc-André Vef, DDN

- Marcus Paradies, LMU Munich

- Matthieu Dorier, Argonne National Laboratory

- Thomas Lambert, Université de Lorraine

- Yusen Li, Nankai University

Thanks to everyone who contributed research papers for sharing their work with the community!

Community discussion - Use cases

- AI applications – training and inference
  - benchmarks
  - Memory management
- Simulations (bursty I/O), experiments and observations (streaming)
- Real-time analytics
- Reliability, fault tolerance,
- Integrated data systems – vertical, horizontal, distributed
- Cloud storage systems
- Object-focused data management
- Edge computing

Community discussion - Challenges

- Data management and storage system challenges
    - Management of memory is an open question
    - Unclear whether GPU-centric I/O is beneficial?
    - reproducibility of experiments, results, performance,
    - Sustainability, energy efficiency,
    - Managing heterogeneity,
    - Managing / utilizing computational storage
    - Interfaces for efficient / transparent usage of storage devices, systems
    - Optimization of data movement
    - Trust in executing computations and storing data
    - Storage optimization targeting virtualization
    - Data management software supply chain management - for security
    - Data sharing – between tasks, jobs, workflows, – performance
    - Application-specific file / data management systems for large-scale, AI
    - Interfaces for heterogeneous storage hardware and software systems
    - Metadata, provenance management - interfaces for storing, querying, and accessing
    - Multi-tenancy in using storage systems efficiently
    - performance and power monitoring / tracing, characterizing, and fixing issues

Community discussion - Potential solutions

- Interfaces and data management systems that unify heterogeneous storage systems and hide the complexity of hardware and software
- Quantum storage
- Using computational storage for performing various tasks (compression, encryption, etc.)
- DNA storage
- Finding the right place for persistent storage - next-gen Optane
- Storage in extreme edge environments - satellites, underwater, etc.
- Giving storage system an ability to forget, remove, hide, - autonomy for identifying only useful information and forget unnecessary data
  - Autonomous garbage collection in storage
  - Smart, situational-aware data
- Anomaly detection in data and in storage system usage (probably to improve performance)

Closing notes

CHEOPS'26 @ EuroSys 2026

Edinburgh, Scotland, UK
April 13th—16th, 2026

https://2026.eurosys.org/index.html