Reduction of Data Sets Through High-Level I/O Interfaces

Yevhen Alforov, Michael Kuhn, Thomas Ludwig

Deutsches Klimarechenzentrum GmbH, Hamburg, Germany
Universität Hamburg, Hamburg, Germany

September 26, 2017

Big Storage project

- BigStorage is a European Training Network (ETN)
  - Main goal is to train future data scientists
- Focus on performance and energy consumption
- Consortium with ten partners
  - Seven research centers/universities
  - Three large companies
- Three associated partners from industry
- 15 Early Stage Researchers (ESRs)

Big Storage project

SKA
SQUARE KILOMETRE ARRAY

Use cases

Data Science
Big Data, Statistics, Machine Learning, Visualization, Data Bases, HPC

HPC and Cloud Convergence
I/O middleware, code-data co-location, elasticity, relaxed semantics, guided I/O

Storage Devices
NVRAM, High capacity Flash, Large Disk, Integrated Photonics

ENERGY

Introduction

- **Motivation and problem statement**:
  - Enormous volumes of data sets are produced
  - Additional disks for more storage space are needed
  - Hardware has to be procured and operated – additional costs
  - Energy consumption increase and CO_2 emission
- **Straight-forward solution**:
  - Data reduction techniques leveraging:
    - Data compression and deduplication
    - Discrete cosine transform and Fourier transform
  - Benefits:
    - Storage capacity optimization
    - Network bandwidth reduction
    - Operation costs minimization and energy saving
    - Saving the environment

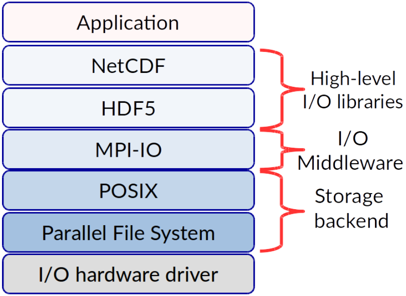
LESS DATA ⇒ LESS STORAGE HARDWARE ⇒ LESS ENERGY COSTS

German Climate Computing Center (DKRZ) example

- Investment costs for 1PB of storage ≈100 000 €
- 1 PB of storage needs 3 kW of power
- 1 kWh of energy costs ≈0.14 €
- Annual electricity costs ≈ 3 680 €
- almost 200 000 € per year for electricity alone (for 54 PiB)
- Not included costs for:
  - Maintenance (approximately 15 % of investment costs)
  - Tapes for long term archives

Data reduction deployment in common HPC I/O stack

You can deploy data reduction techniques on **lower (system)** or **higher (application)** level of HPC I/O stack:



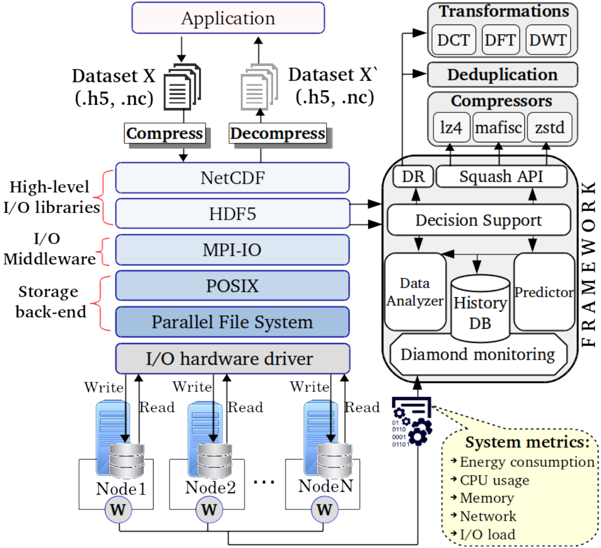
- APPLICATION LEVEL
  - Benefit: **Flexibility** (access to meta information)
  - Drawback: **Clarity** (insight into code is needed)
- SYSTEM LEVEL
  - Benefit: **Transparency** (no modifications of app.)
  - Drawback: **Uncertainty** (only lossless reduction)

Data Reduction usage on **higher levels** of HPC I/O stack is **advantageous**!

Drawbacks and benefits of DR deployment in I/O stack

	SYSTEM LEVEL	APPLICATION LEVEL
DRAWBACK	**Uncertainty** due to the lack of access to application-specific semantical information (e.g., data structures, important variables, etc.) only lossless reduction can be considered	**Clarity** insight into the code and requirements of applications is needed for tuning the performance of data reduction technique
BENEFIT	**Transparency** no needs to modify applications, even if they are very diverse or don't use a common I/O software stack	**Flexibility** semantical information is easily accessible, hence more reduction techniques can be leveraged (even for specific portions of data)

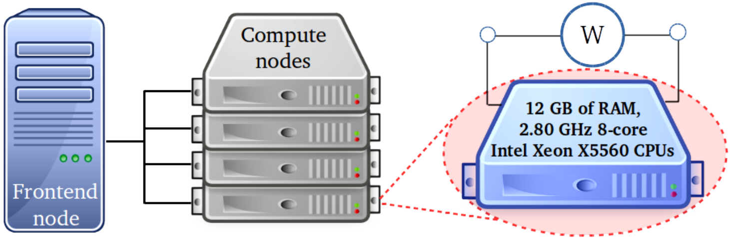
Design of EEDR framework (under development)



Evaluation setup

Cluster operated by FS **Lustre**

ArduPower wattmeter

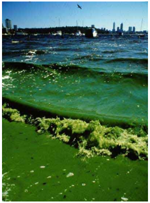


Maximum **throughput** - 110 MB/s

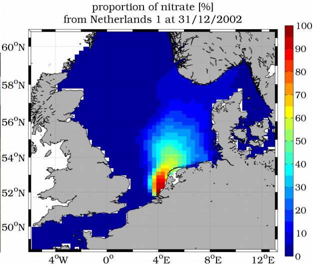
1 node usage

Dataset and workload

- **17 GB** data set of 3D ecosystem model for the North Sea **ECOHAM** (from Climate Science)

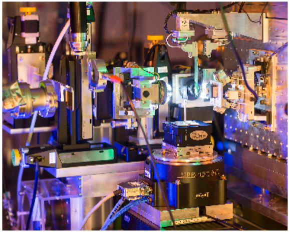


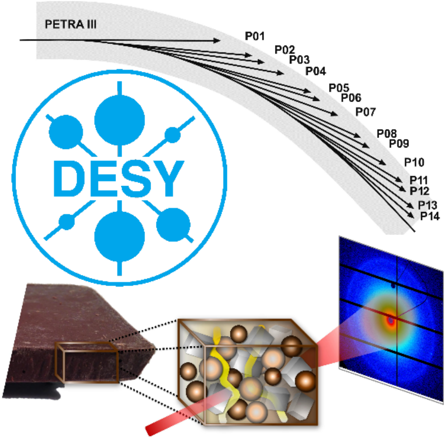
ECOHAM
(ECOsystem Model Hamburg)
Version 5



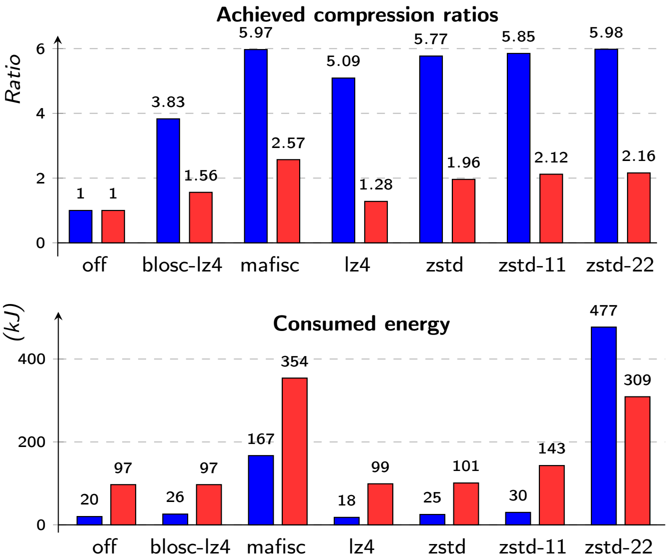
Dataset and workload

- **14 GB** data set of tomography experiments from **PETRA III**
  PCO 4000 detector (from High Energy Physics)



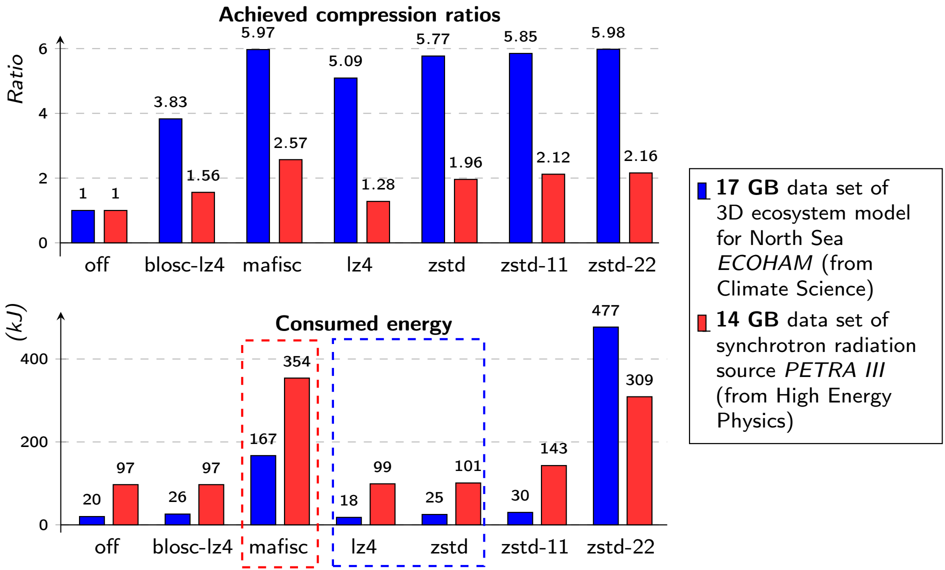


Preliminary evaluation and results for HDF5 filters

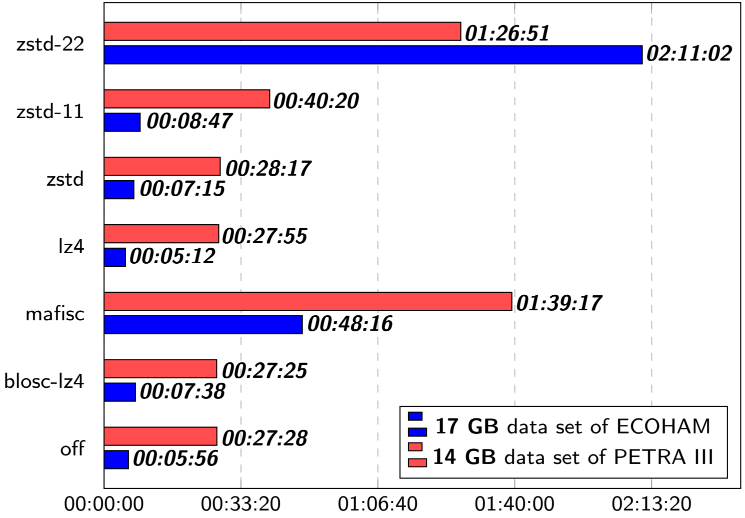


- **17 GB** data set of 3D ecosystem model for North Sea *ECOHAM* (from Climate Science)
- **14 GB** data set of synchrotron radiation source *PETRA III* (from High Energy Physics)

Preliminary evaluation and results for HDF5 filters



Runtime of evaluated HDF5 filters



Preliminary evaluation and results for HDF5 filters

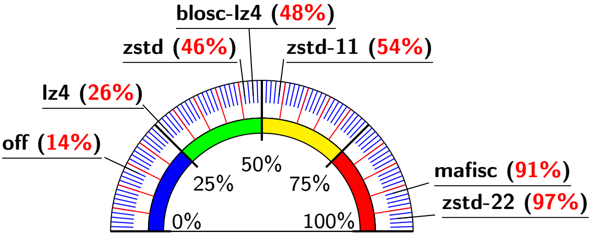


Figure: Average CPU utilization with **ECOHAM** data set

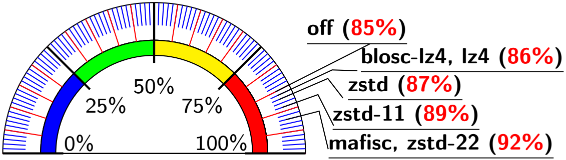


Figure: Average CPU utilization with **PETRA III** data set

Outcome

Significant algorithms for the framework are:

- MAFISC (when only compression ration matters)
  - 0.013 € - consumed energy costs for ECOHAM5
  - 0.0065 € - consumed energy costs for PETRA III
- LZ4, ZSTD (when runtime, energy or CPU are also important)
  - 0.0039 € - consumed energy costs for ECOHAM5
  - 0.0007/0.00098 € - consumed energy costs for PETRA III

Summary and future research work

- **Summary**:
  - **Amount** of reduced data **depends** on the **structure of data**
  - **Trade-off** between compression ratio and energy consumption
  - Different approaches are appropriate depending on the use case
    - **Archival** with **slower algorithms**
    - **Parallel I/O** should be handled as **fast** as possible
- **Future plans**:
  - Experimenting with **application-specific** techniques
    - Using **semantic information** available at the application level
    - Leveraging techniques that are **complex and/or expensive to deploy** at the system level such as deduplication
  - Provide high-level I/O extensions to reduce energy consumption
  - Taking into account **Big Data** applications