Compression for Extreme-Scale Computing: A Case Study of SPH-EXA's Data Analysis Architecture

Yiqing Zhu, Osman Seckin Simsek, Rubén Cabezón, Florina Ciorba
University of Basel

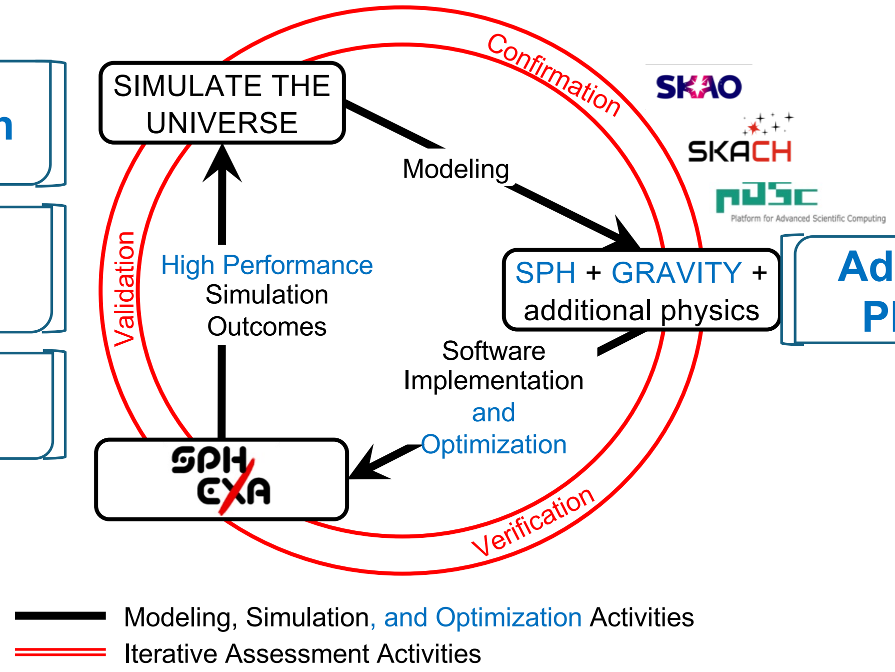
SKACH Winter Meeting, January 22, 2024

Modeling, Simulation and Optimization through Interdisciplinary Co-Design

Parallelization

Portability + Scalability

Analysis + Visualization



Additional Physics

Parallelization
Domain decomposition
Multiprocess
Multithreaded
Acceleration
Vectorization
Overlapping comm. & comp.
Scheduling and load balancing
Parallel I/O
…

Portability + Scalability
Mixed precision
Multiple accelerator models
Minimal SW dependencies
Cross-platform compilation, testing, benchmarking
Fault tolerance
Reproducibility
…

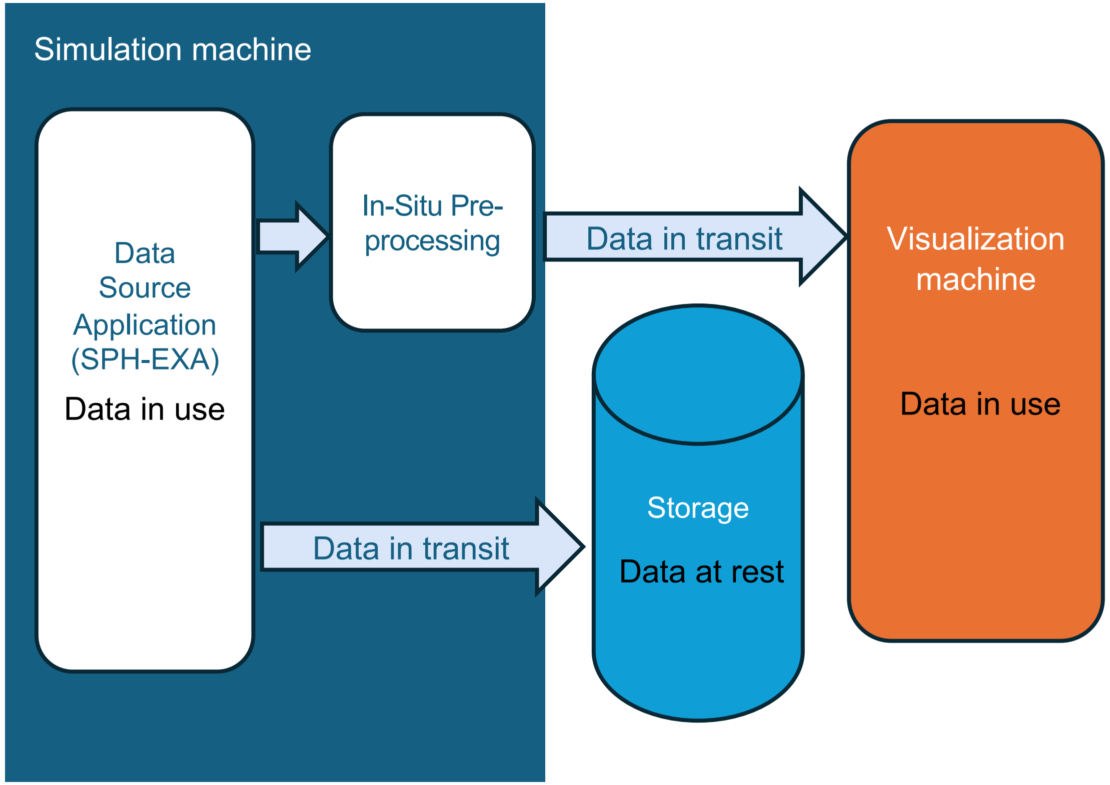
Additional Physics
Nuclear reactions
Radiative cooling
Radiation pressure
Magnetism
Energy diffusion
Neutrino transport
General relativity
…

Analysis + Visualization
Performance profiling, tracing, visualization
Scientific visualization via snapshots and in-situ
…

What is the role of I/O in supporting data analysis at extreme scale?

[Adapted from: Schlesinger, S., “Terminology for Model Credibility,” Simulation, Vol. 32, No. 3, 1979.]

Extreme-Scale Data Analysis: Post-Hoc and In-Situ



SPH-EXA data analysis architecture: Post-hoc and in-situ

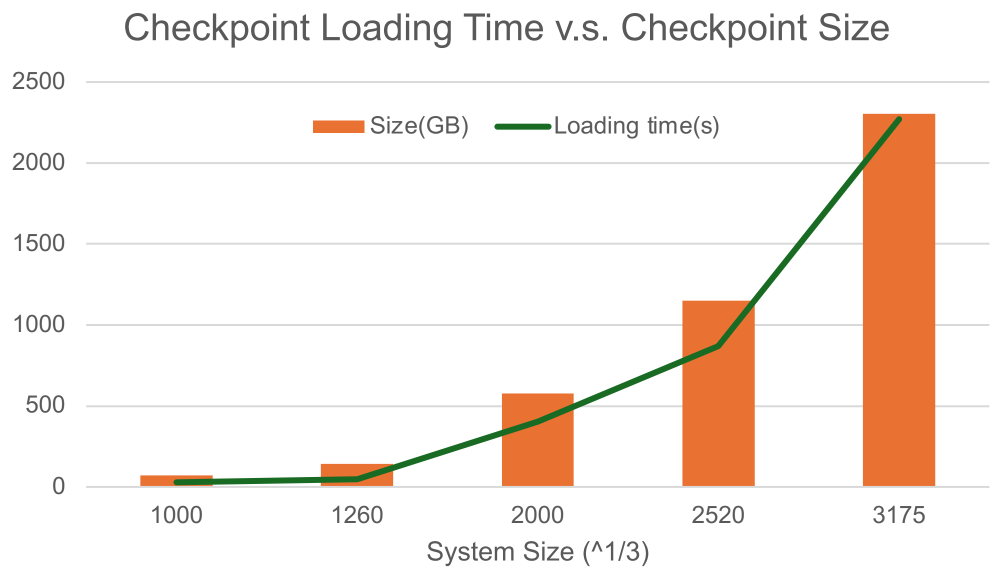
Post-hoc Analysis

- From simulation machine directly output to storage or visualization pipeline
  - High I/O
  - Needs a powerful visualization machine

In-situ Analysis

- Runtime co-process the simulation data, then outputs to visualization pipeline
  - Can extract features in runtime
  - Low I/O
  - Low HW requirement (visualization can even be in-place)

Extreme-Scale Data Analysis: Post-Hoc and In-Situ



SPH-EXA Turbulence, 128 MPI ranks, 128 nodes on Piz Daint XC50

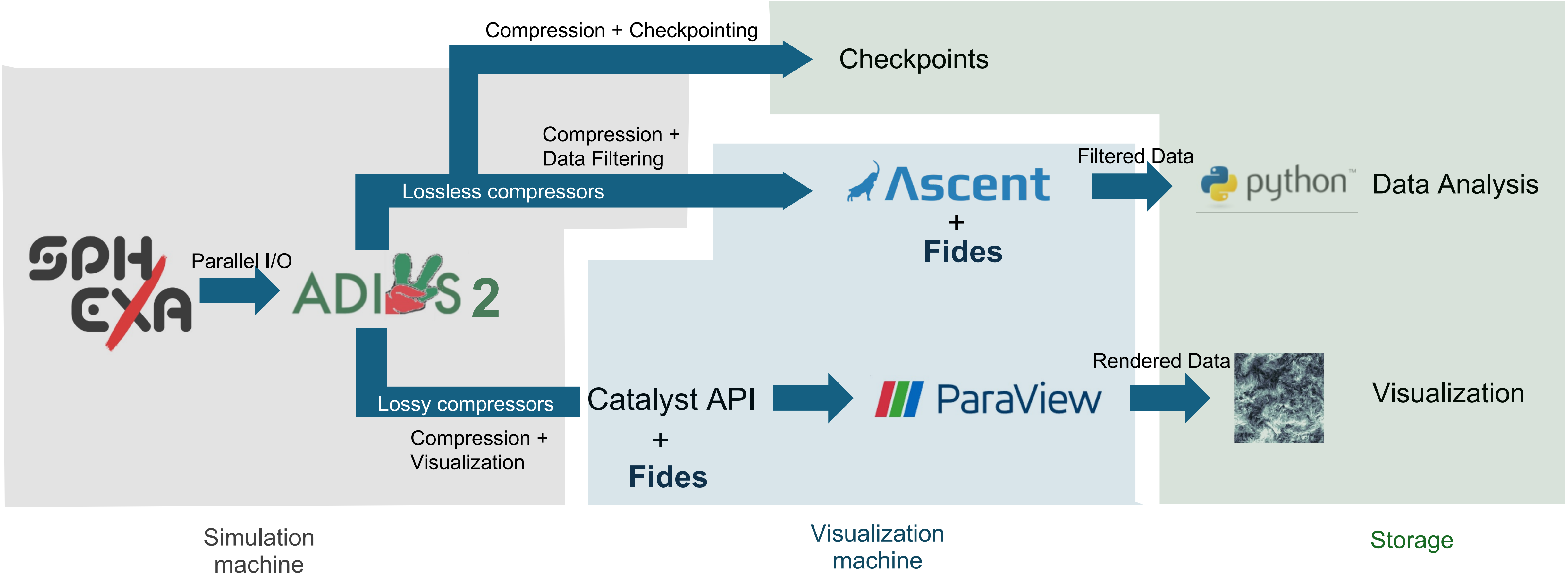
Challenges

- Reduce checkpointing time during simulation
- In-situ process of large datasets
- Parallel non-blocking I/O

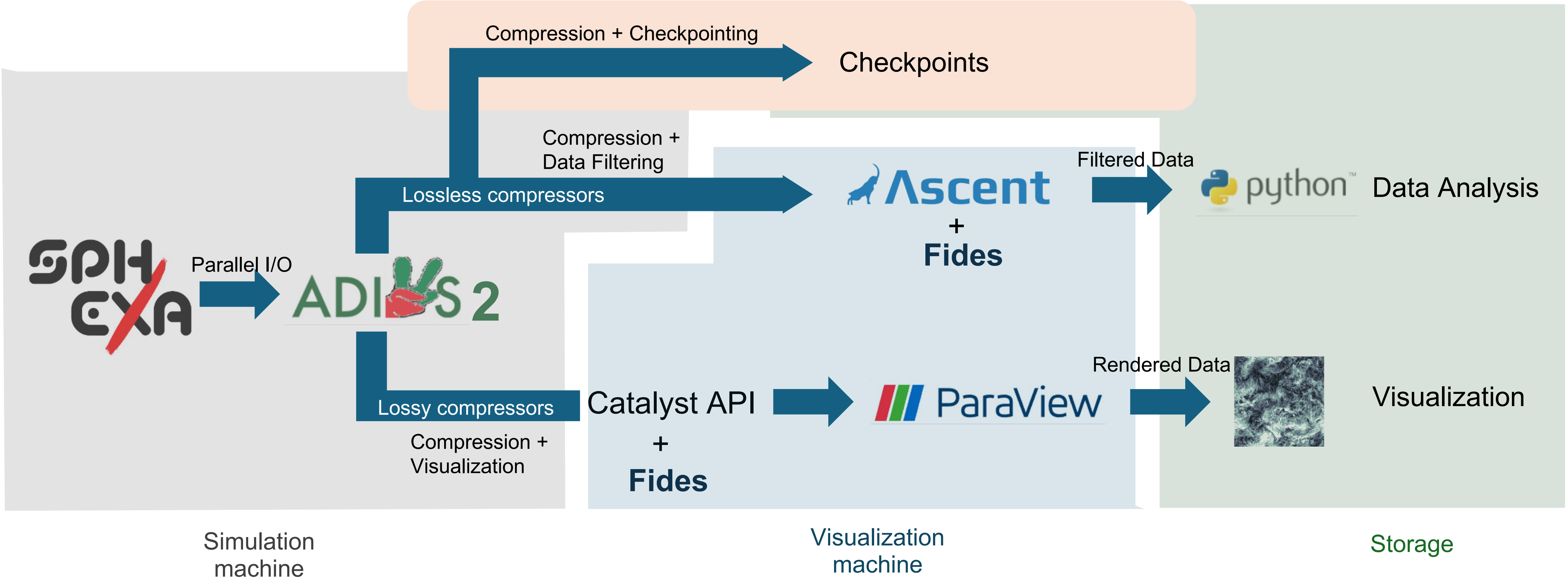
Possible Solution: Compression

Data storage, loading, transfer…

SPH-EXA Checkpointing and Data Analysis

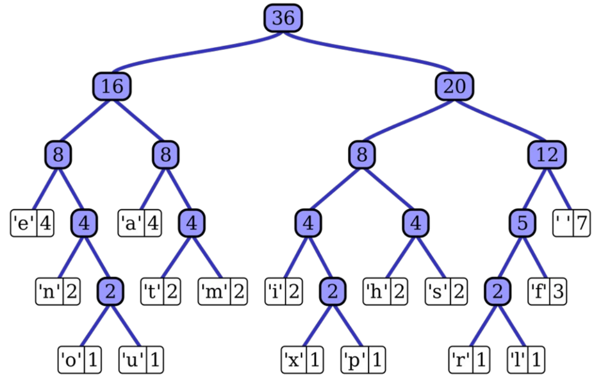


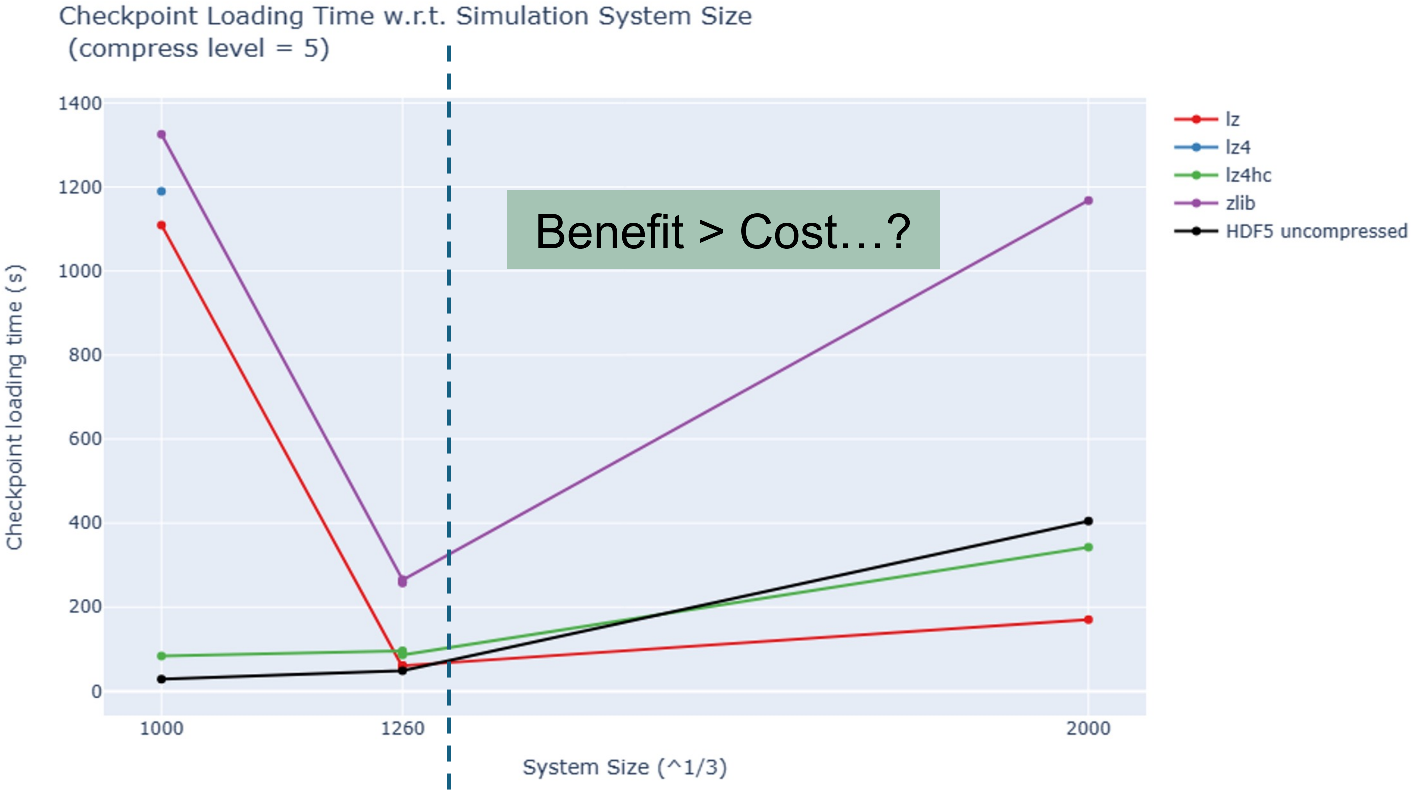
Checkpointing: Lossless Compression



Checkpointing: Lossless Compression

- Lossless compressors
  - Zlib
  - LZ
  - LZ4
  - LZ4HC

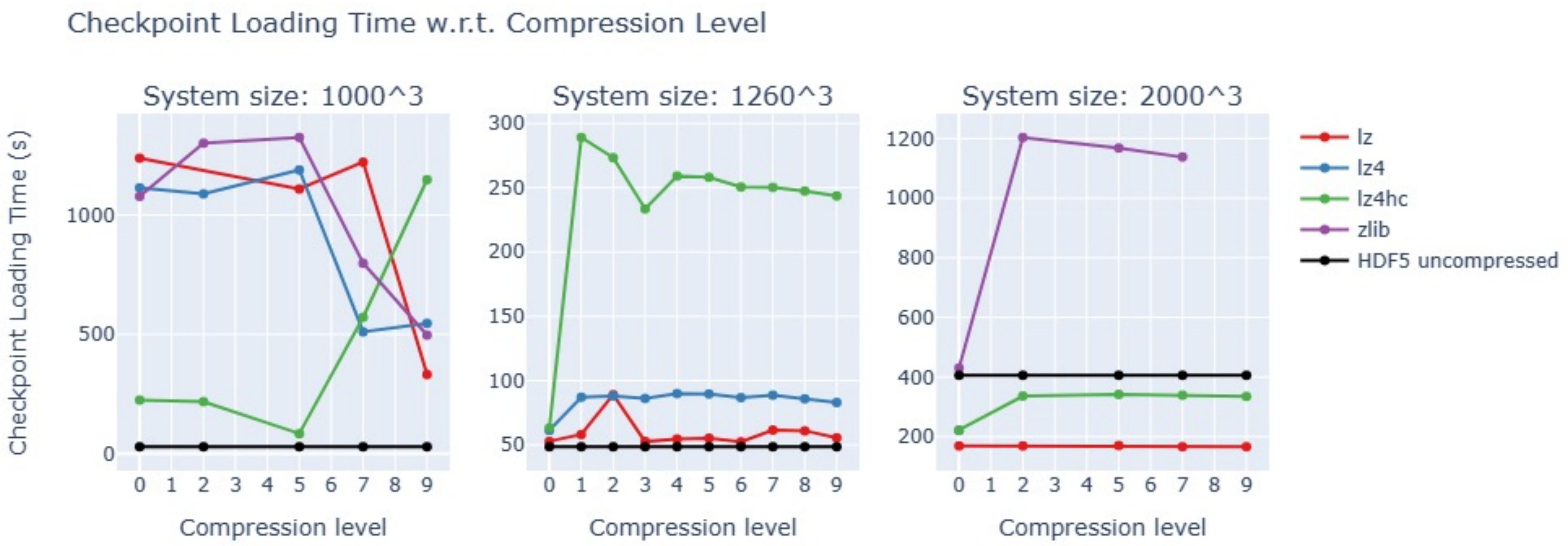




SPH-EXA Turbulence, 128 MPI ranks, 128 nodes on Piz Daint XC50

[Meteficha. (2023, December 2). *Huffman coding*. Wikipedia. https://en.wikipedia.org/wiki/Huffman_coding.]

Checkpointing: Lossless Compression

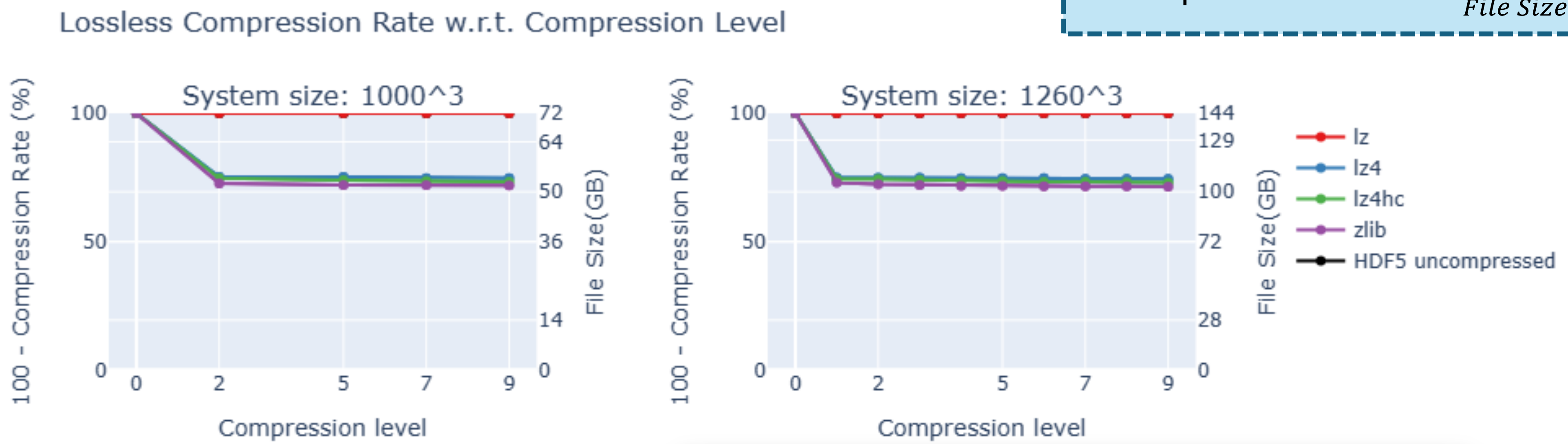


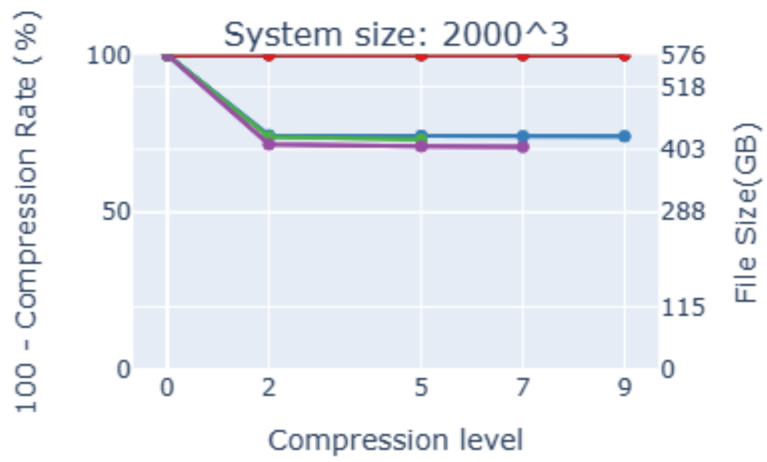
SPH-EXA Turbulence, 128 MPI ranks, 128 nodes on Piz Daint XC50

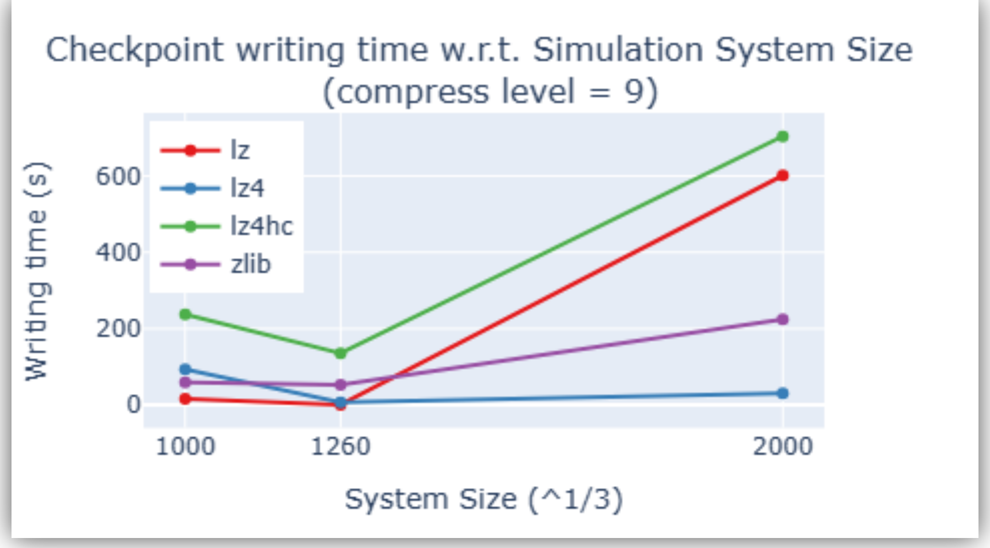
For most lossless compressors, decompression cost decreases with compression level

Checkpointing: Lossless Compression

Compression rate: $1 - \frac{File\ Size_{compressed}}{File\ Size_{uncompressed}}$

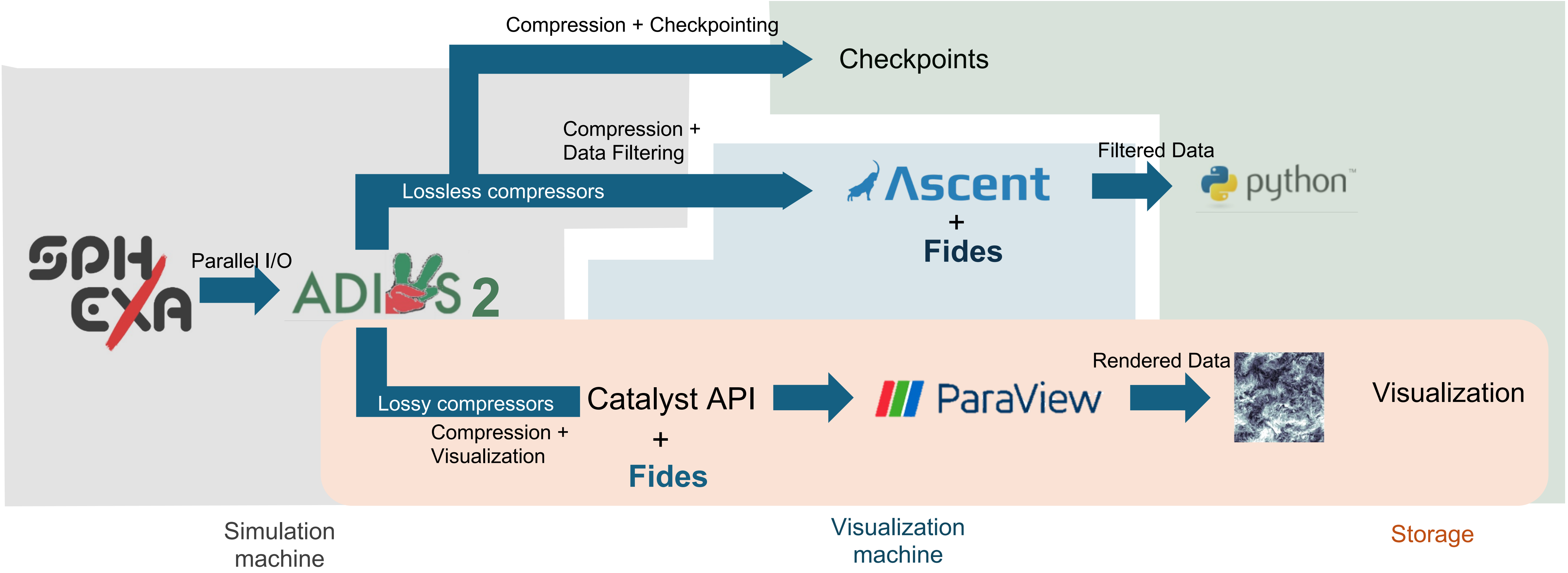






SPH-EXA Turbulence, 128 MPI ranks, 128 nodes on Piz Daint XC50

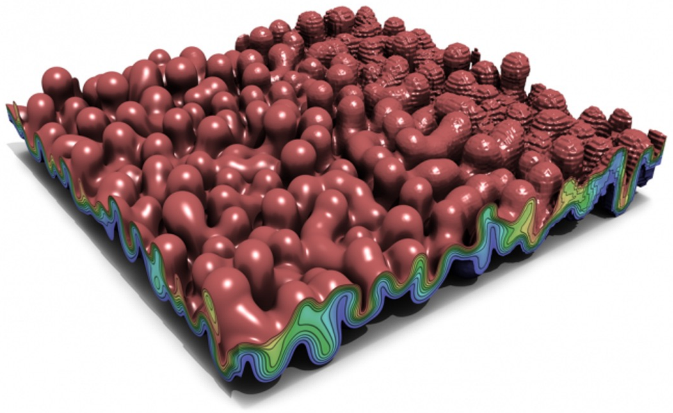
Visualization: Lossy Compression



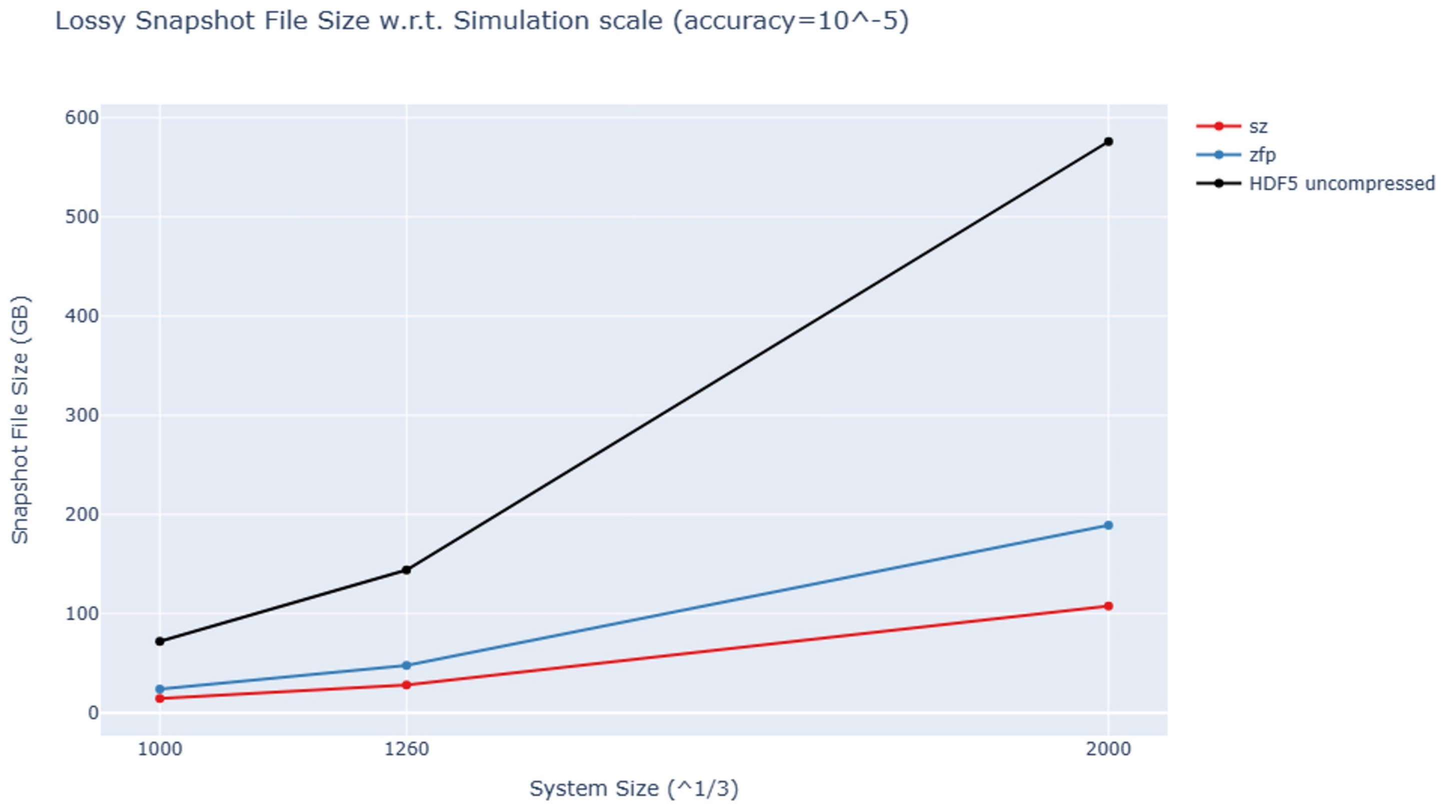
Visualization: Lossy Compression

- Lossy Compressors
  - SZ
  - ZFP

Given uncompressed value f and reconstructed value g, it is guaranteed that
$|f - g| \leq accuracy$



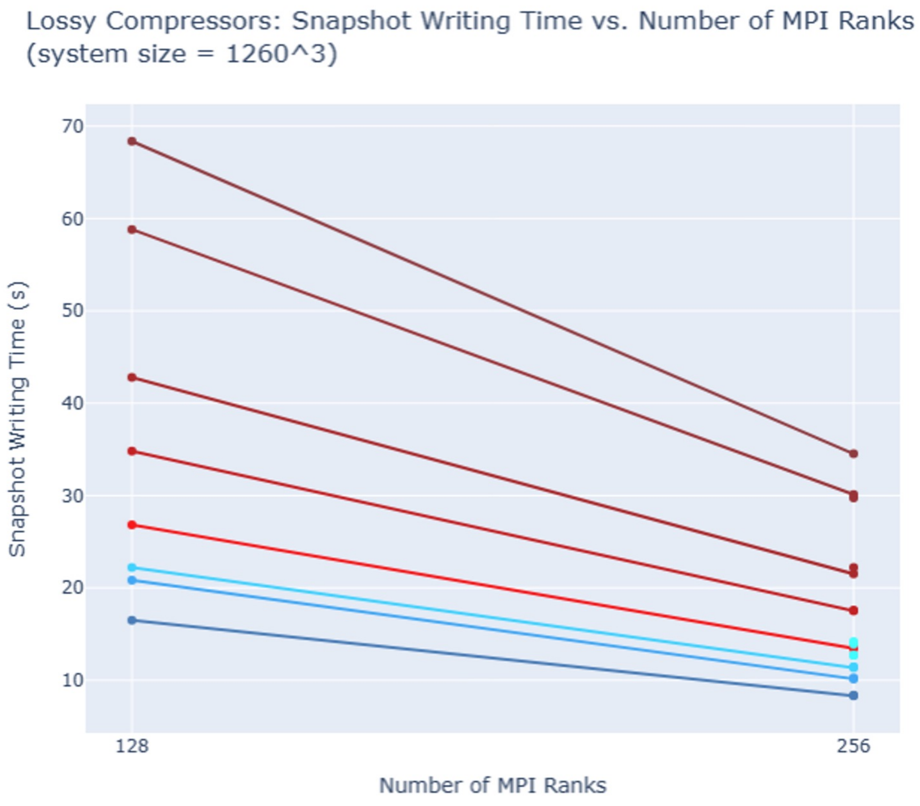
[Lindstrom, P. (2014). Fixed-rate compressed floating-point arrays. *IEEE transactions on visualization and computer graphics*, *20*(12), 2674-2683.]

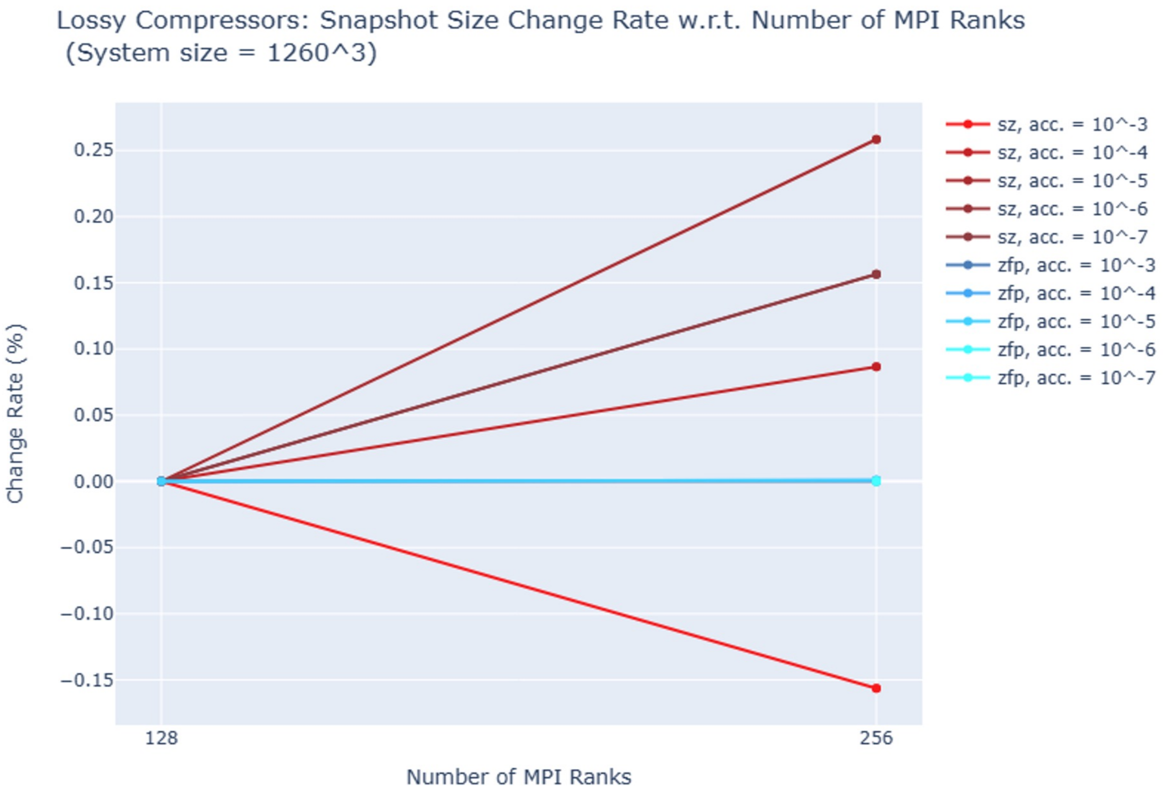


SPH-EXA Turbulence, 128 MPI ranks, 128 nodes on Piz Daint XC50

Visualization: Lossy Compression

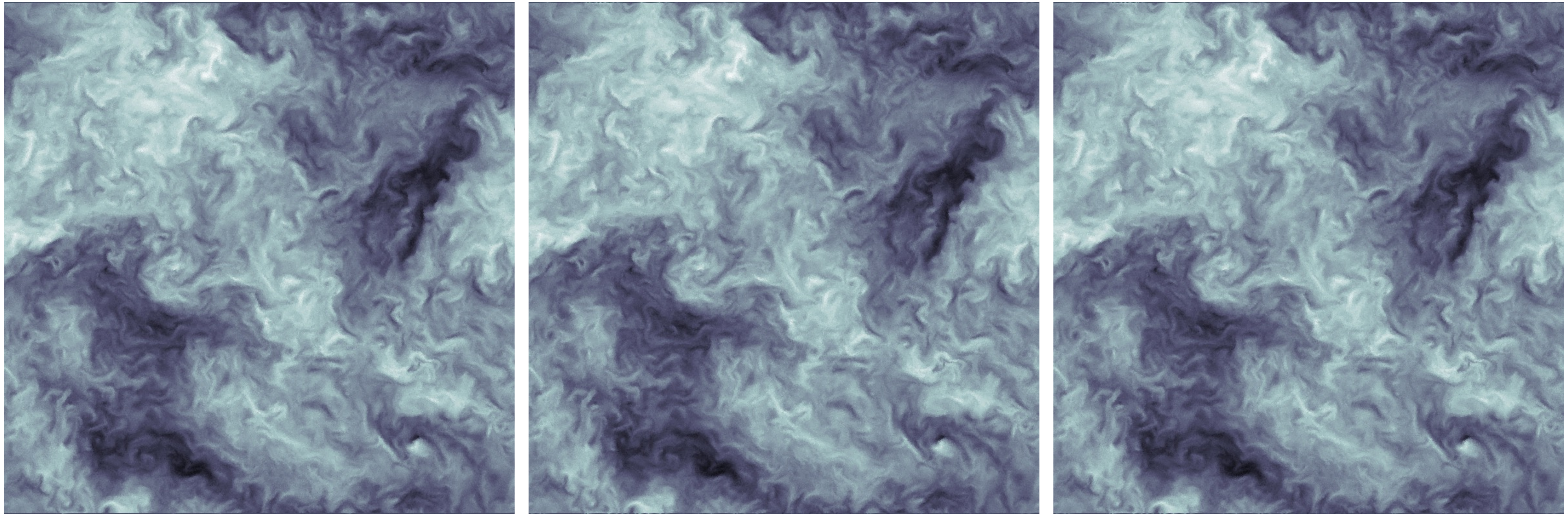
Snapshot Size Change Rate: $\frac{File\ Size_{256\ Ranks}}{File\ Size_{128\ Ranks}} - 1$





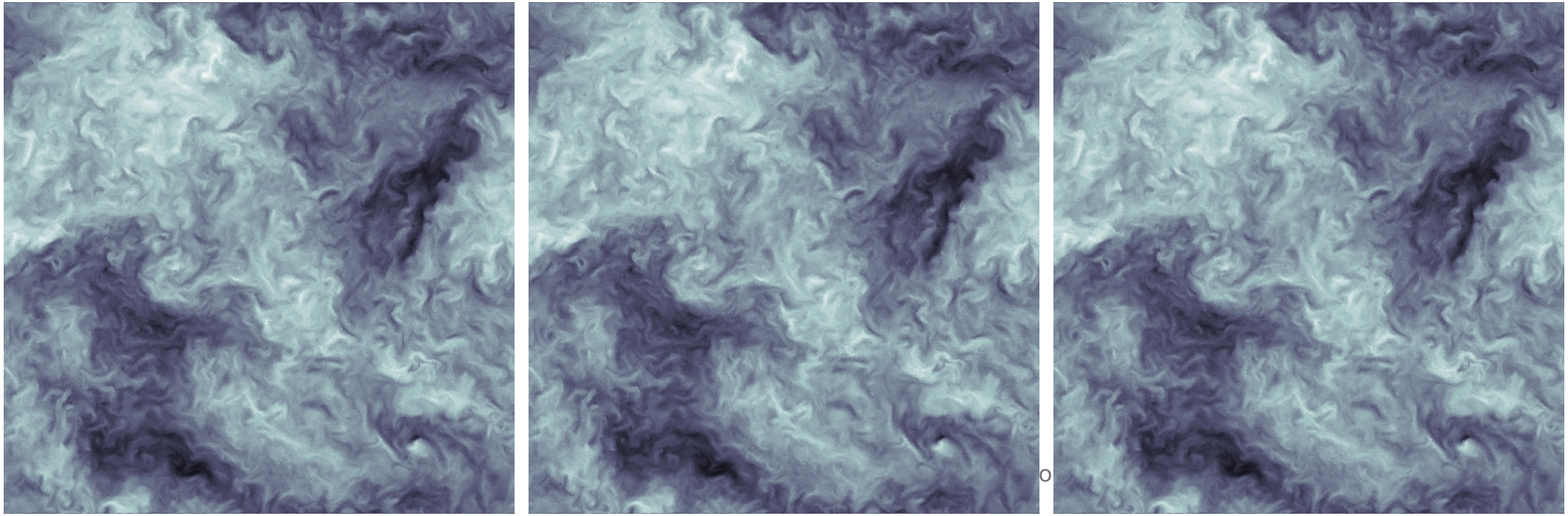
SPH-EXA Turbulence, 128 MPI ranks, 128 nodes on Piz Daint XC50

Visualization: Lossy Compression



Compressor: SZ

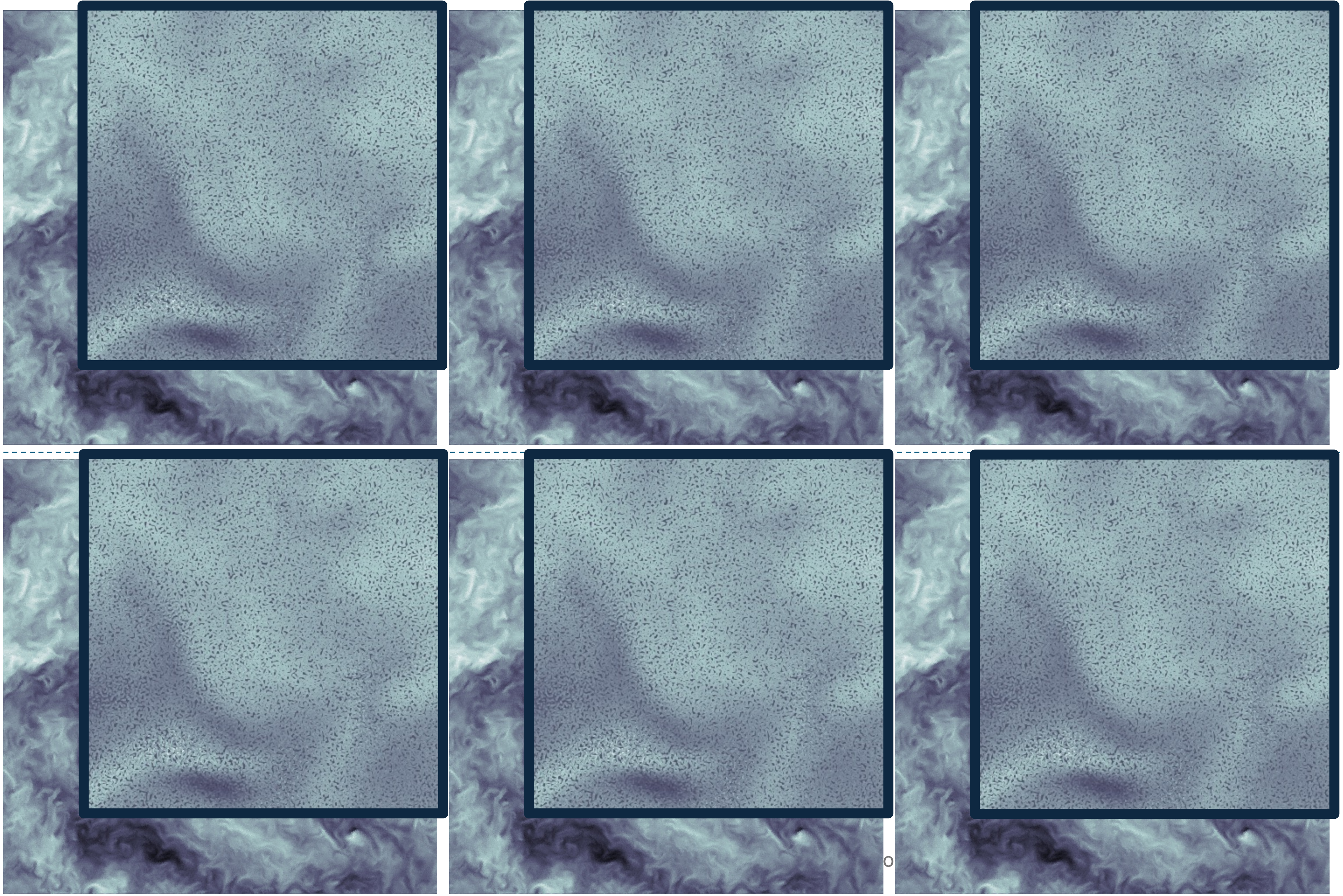
Compressor: ZFP



SPH-EXA Turbulence, system size 1000^3 at time t=18.0s. Colored by velocity, with different lossy compression accuracy

Zhu, Simsek, Cabezon, Ciorba

Visualization: Lossy Compression



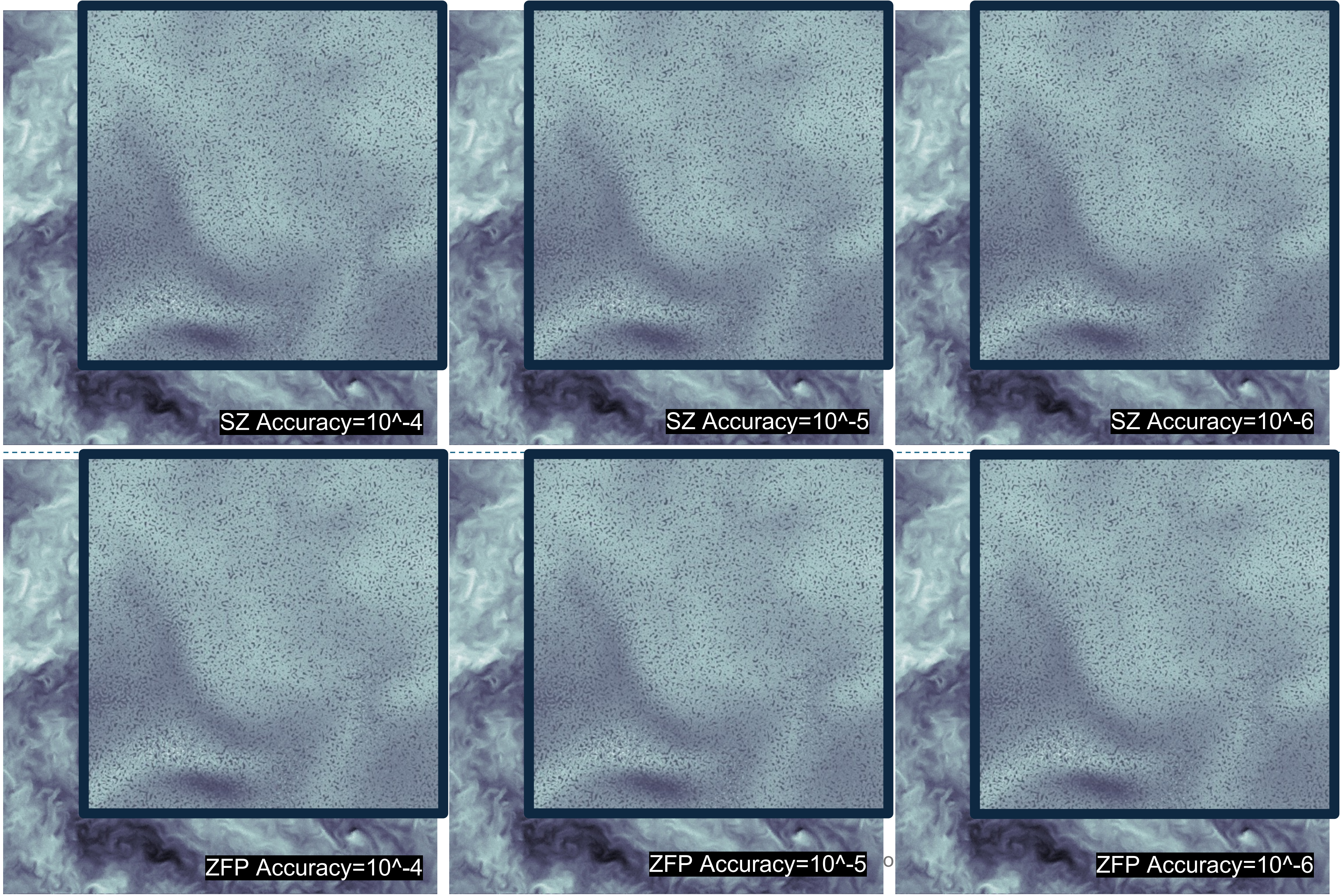
Compressor: SZ

Compressor: ZFP

SPH-EXA Turbulence, system size 1000^3 at time t=18.0s. Colored by velocity, with different lossy compression accuracy

Zhu, Simsek, Cabezon, Ciorba

Visualization: Lossy Compression



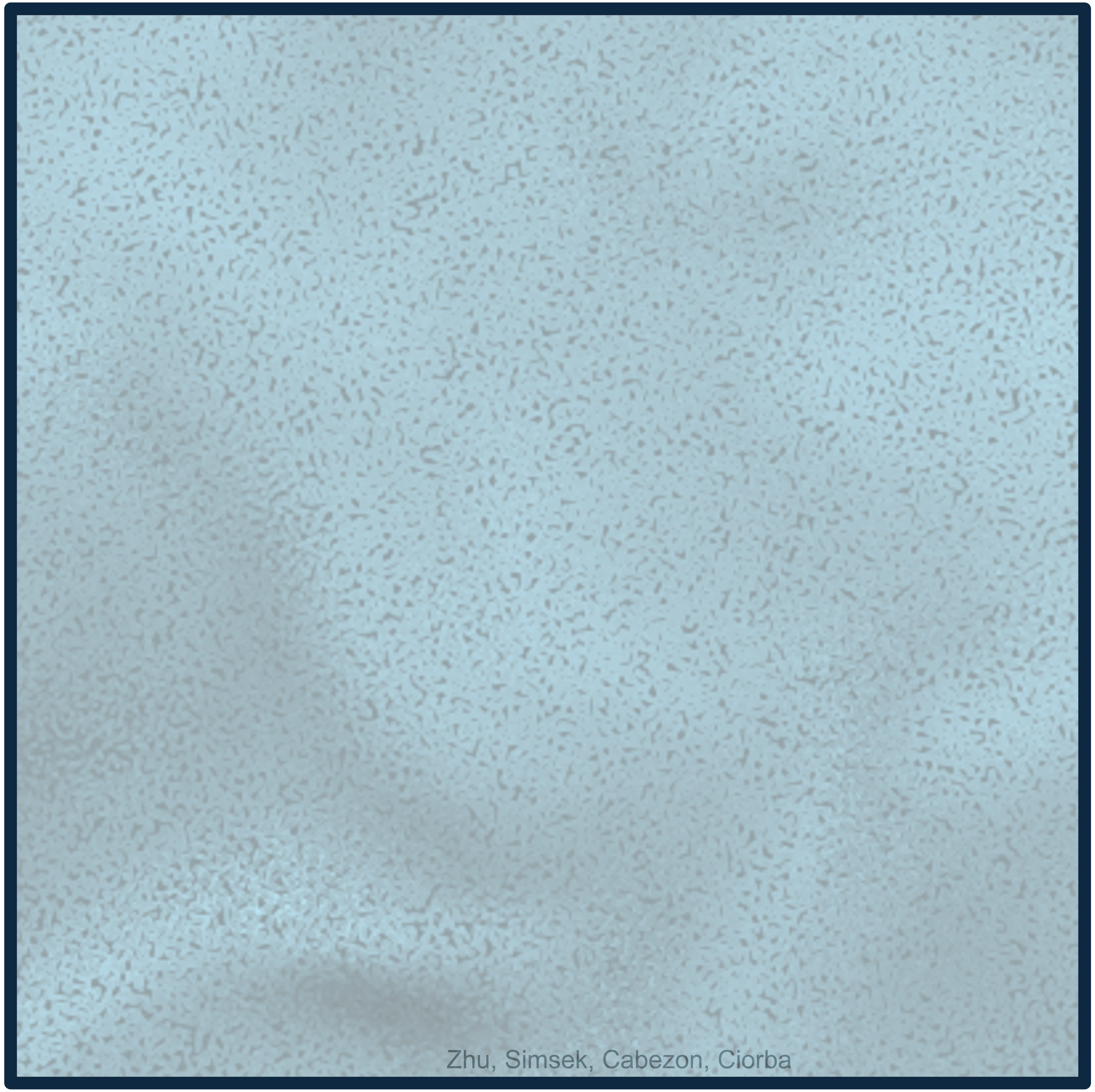
Compressor: SZ

Compressor: ZFP

SPH-EXA Turbulence, system size 1000^3 at time t=18.0s. Colored by velocity, with different lossy compression accuracy

Zhu, Simsek, Cabezon, Ciorba

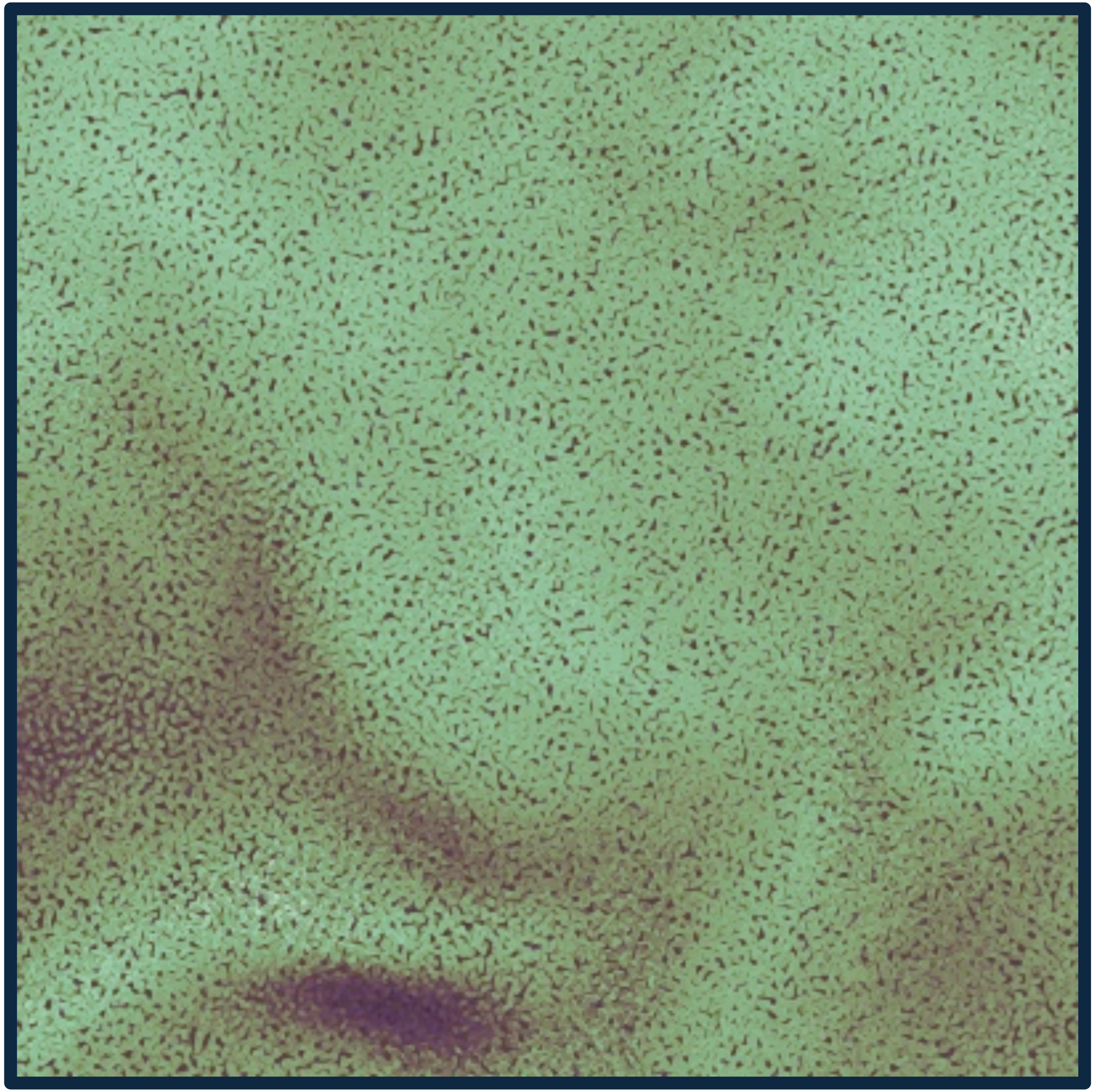
Visualization: Lossy Compression



SZ accu. = 10^-4 (yellow)
SZ accu. = 10^-6 (blue)

SPH-EXA Turbulence, system size 1000^3 at time t=18.0s. Colored by velocity, with different lossy compression accuracy

Zhu, Simsek, Cabezon, Ciorba



Visualization: Lossy Compression

SZ accu. = 10^-4 (yellow)
SZ accu. = 10^-6 (blue)

SPH-EXA Turbulence, system size 1000^3 at time t=18.0s. Colored by velocity, with different lossy compression accuracy

Summary

- Exploration of compression for checkpointing
- Lossy compression for in-situ visualization
- Provide reference parameters for tuning compressors for extreme-scale simulations

What if we have higher I/O bandwidth…?

In Progress

- Mixed-precision
- Larger scale visualization
- Verification of compression
- …

Summary

- Exploration of compression for checkpointing
- Lossy compression for in-situ visualization
- Provide reference parameters for tuning compressors for extreme-scale simulations

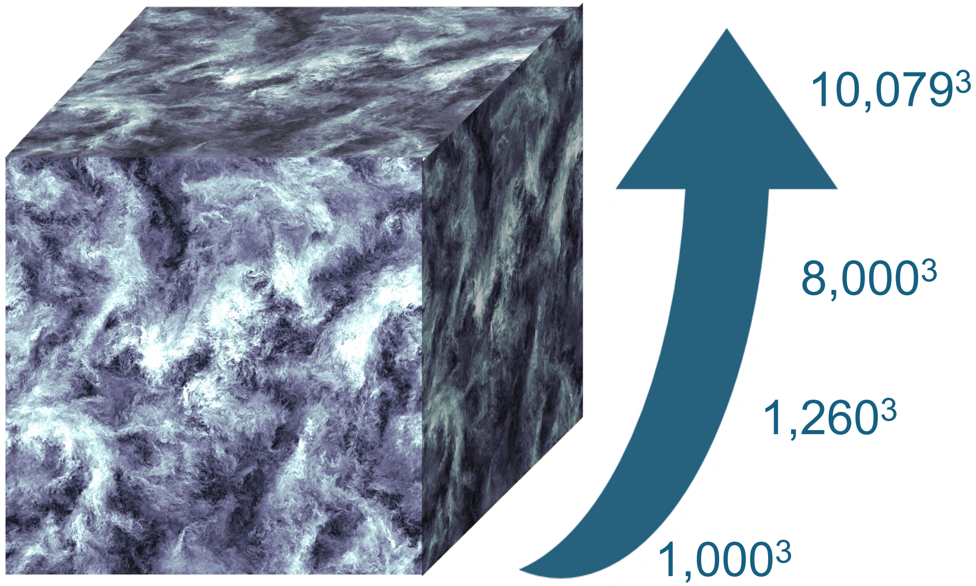
What if we have higher I/O bandwidth…?

In Progress

- Mixed-precision
- Larger scale visualization
- Verification of compression
- …

Projects We Support

TGSF: The Role of Turbulence and Gravity in Star Formation, Unveiling the Sonic Scale with Smoothed Particle Hydrodynamics



For $10,079^3$ particles, relaxing the ICs already consumes 500’000 node hours on LUMI-G

Summary

- Exploration of compression for checkpointing
- Lossy compression for in-situ visualization
- Provide reference parameters for tuning compressors for extreme-scale simulations

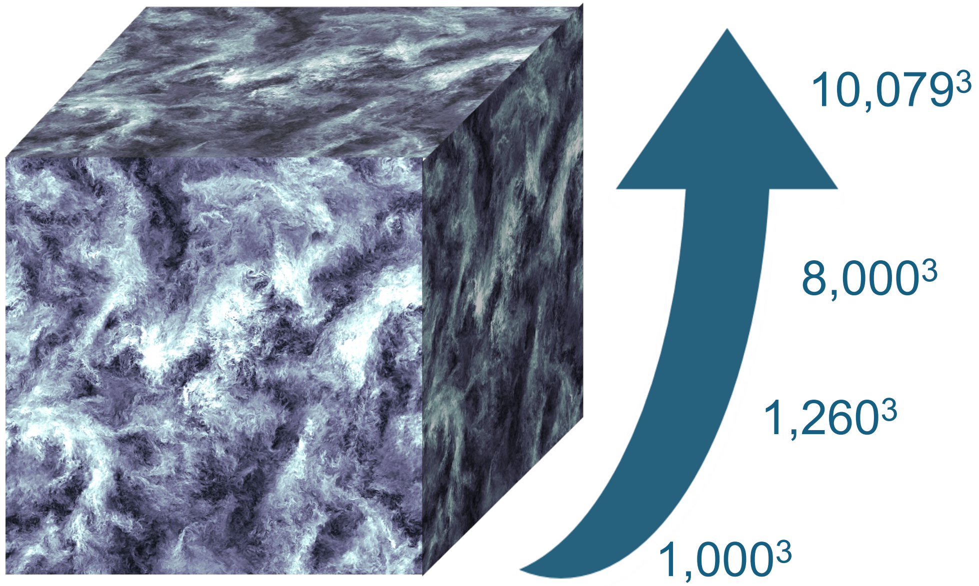
What if we have higher I/O bandwidth…?

In Progress

- Mixed-precision
- Larger scale visualization
- Verification of compression
- …

Projects We Support

TGSF: The Role of Turbulence and Gravity in Star Formation, Unveiling the Sonic Scale with Smoothed Particle Hydrodynamics



For $10{,}079^3$ particles, relaxing the ICs already consumes 500’000 node hours on LUMI-G

Take away
We explore the potential of an efficient data analysis architecture to support extreme-scale simulations, such that storage, data loading and data transfer are not a bottleneck.

Compression for Extreme-Scale Computing: A Case Study of SPH-EXA's Data Analysis Architecture

Yiqing Zhu, Osman Seckin Simsek, Rubén Cabezón, Florina Ciorba
University of Basel

SKACH Winter Meeting, January 22, 2024