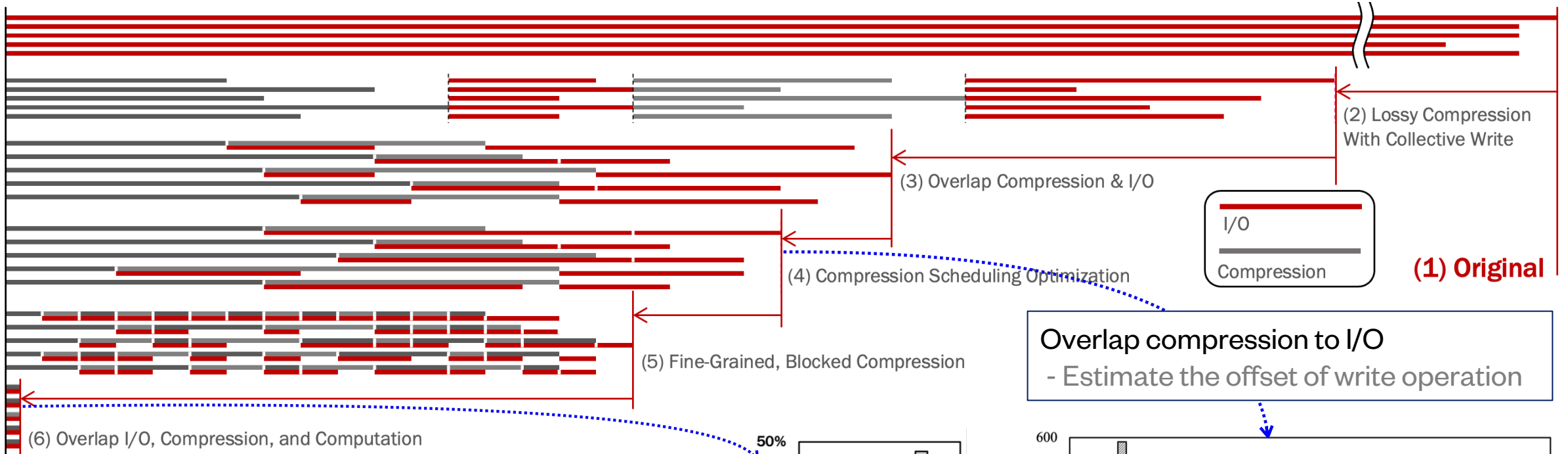




Accelerating Parallel Write via Deeply Integrating Predictive Lossy Compression with HDF5

Sian Jin^{*}, Dingwen Tao^{*}, Houjun Tang [‡], Sheng Di[†], Suren Byna[‡], Zarija Lukic[‡], Franck Cappello[†]
^{*}Luddy School of Informatics, Computing, and Engineering, Indiana University, Bloomington, IN, USA
[†]Mathematics and Computer Science Division, Argonne National Laboratory, Lemont, IL, USA
[‡]Computational Research Division, Lawrence Berkeley National Laboratory, Berkeley, CA, USA

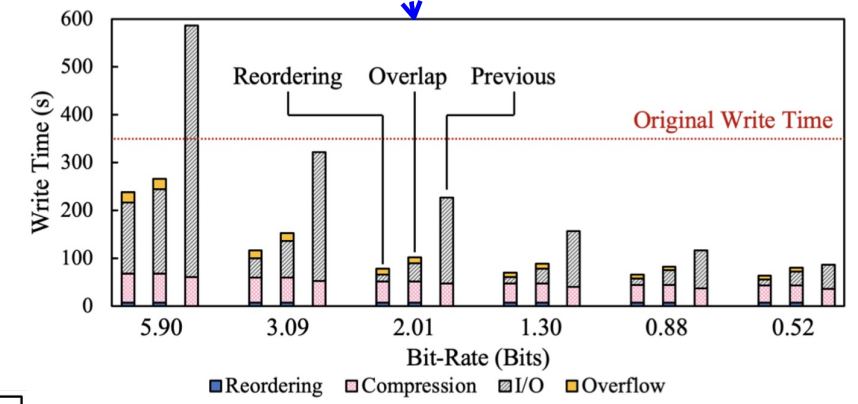
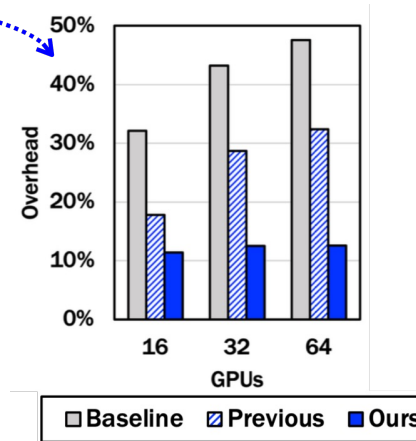
Accelerating Parallel Write via Deeply Integrating Predictive Lossy Compression with HDF5



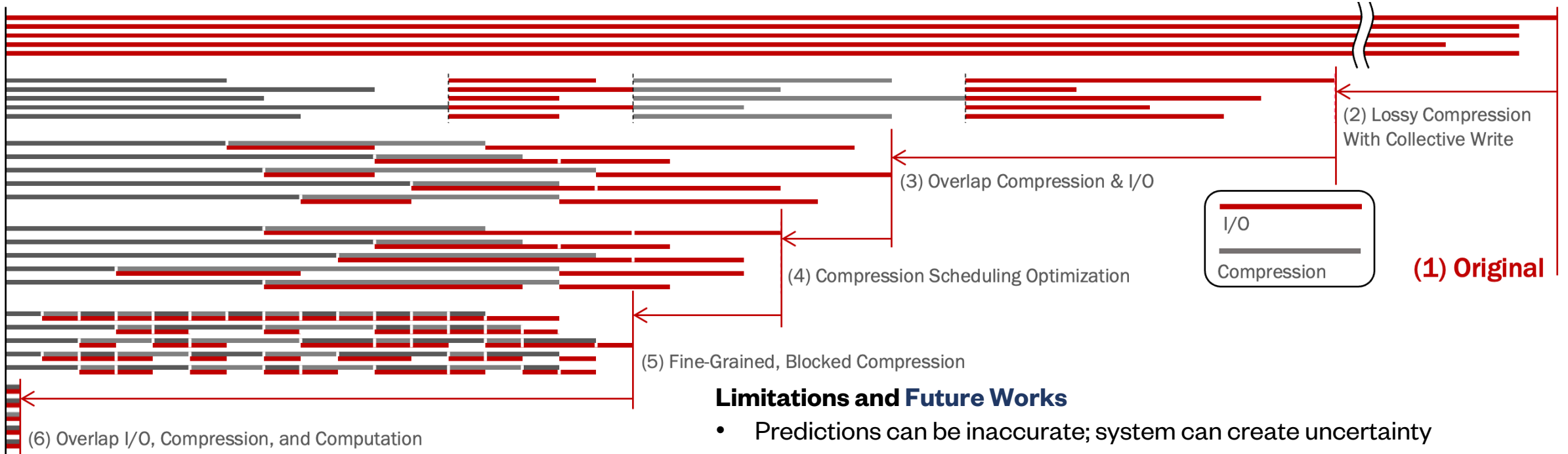
Overlap compression to I/O
- Estimate the offset of write operation

Key Components

- HDF5 VOL Connector
- Prediction model for lossy compressors
- Scheduling algorithms
- Other specific designs: overflow control, fine-grained compression, buffering, shared Huffman tree



Accelerating Parallel Write via Deeply Integrating Predictive Lossy Compression with HDF5



Key Components

- HDF5 VOL Connector
- Prediction model for lossy compressors
- Scheduling algorithms
- Other specific designs: overflow control, fine-grained compression, buffering, shared Huffman tree

Limitations and Future Works

- Predictions can be inaccurate; system can create uncertainty
 - Accurate prediction or more tolerant system design
 - Uncertainty analysis
- Not user friendly

