HDF5 VOL Connector to Apache Arrow

Jie Ye
Illinois Institute of Technology

Suren Byna
Lawrence Berkeley National Laboratory

Quincey Koziol
Lawrence Berkeley National Laboratory

Anthony Kougkas
Illinois Institute of Technology

October 12th, 2021

Outline

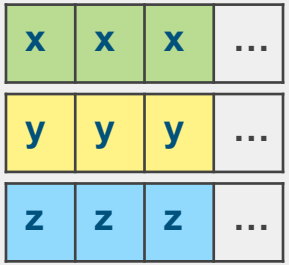
- Problem Statement
- Background
- Project Goals
- Design
- Initial Performance Results
- Conclusion

Problem Statement

- Scientific data is often organized as array of structures (AOS) or structure of arrays (SOA)

Structure of Arrays (SOA)

```
struct point3D {
    float x[3];
    float y[3];
    float z[3];
};
struct point3D points;
```



Array of Structures (AOS)

```
struct point3D {
    float x;
    float y;
    float z;
};
struct point3D points[3];
```

Problem Statement (cont.)

- HDF5 works well with structure of arrays (SOA)
- However, it performs suboptimal when dealing with array of structures (AOS) and table-like data structures
  - Data accesses are column-oriented
  - They would cause transformations when columnar data is required

(AOS)

- Recently, Apache Foundation released an in-memory column store called Apache Arrow
  - It is considered as an efficient in-memory, column store data management in cloud environment

Background

- Apache Arrow
  - A popular platform for columnar in-memory data representation offering efficient data processing and transfer
  - Specifies a standard columnar in-memory format for representing structured and table-like datasets
  - Libraires are available for many languages, like C, C++, Go, Java, Python and others.
  - Official website - https://arrow.apache.org/

Background (cont.)

- Advantages of Apache Arrow
  - Column is Fast
  - Reduce the overhead of copy and convert when moving data from one system to another

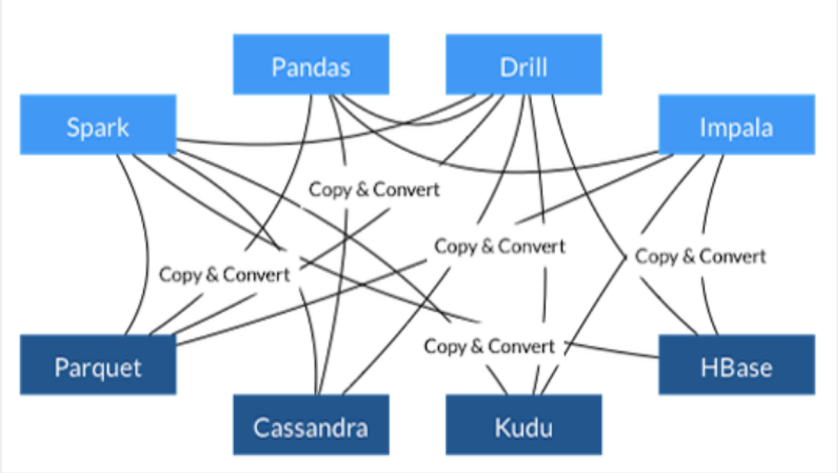


Fig1. a) without Apache Arrow

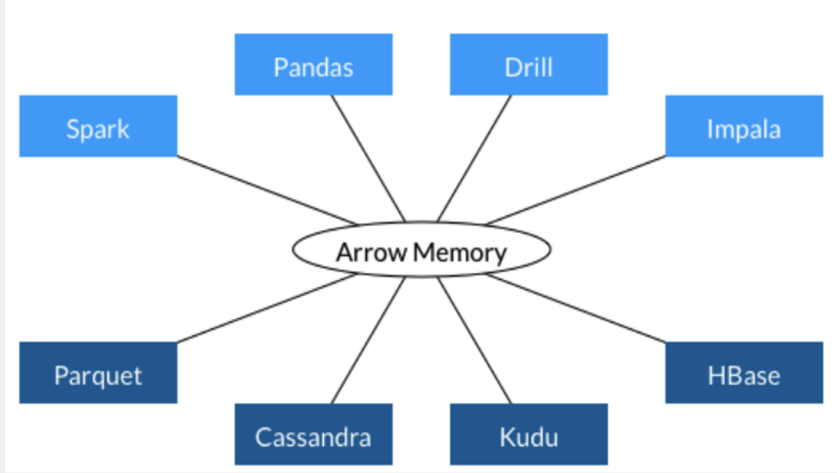


Fig1. b) with Apache Arrow

Goals

- Developed a terminal HDF5 VOL connector to Apache Arrow
- Evaluate how big data technologies that offer new capabilities work in HPC systems
- Bridge the gap between science applications and analytics tools that use HDF5 and Apache Arrow data

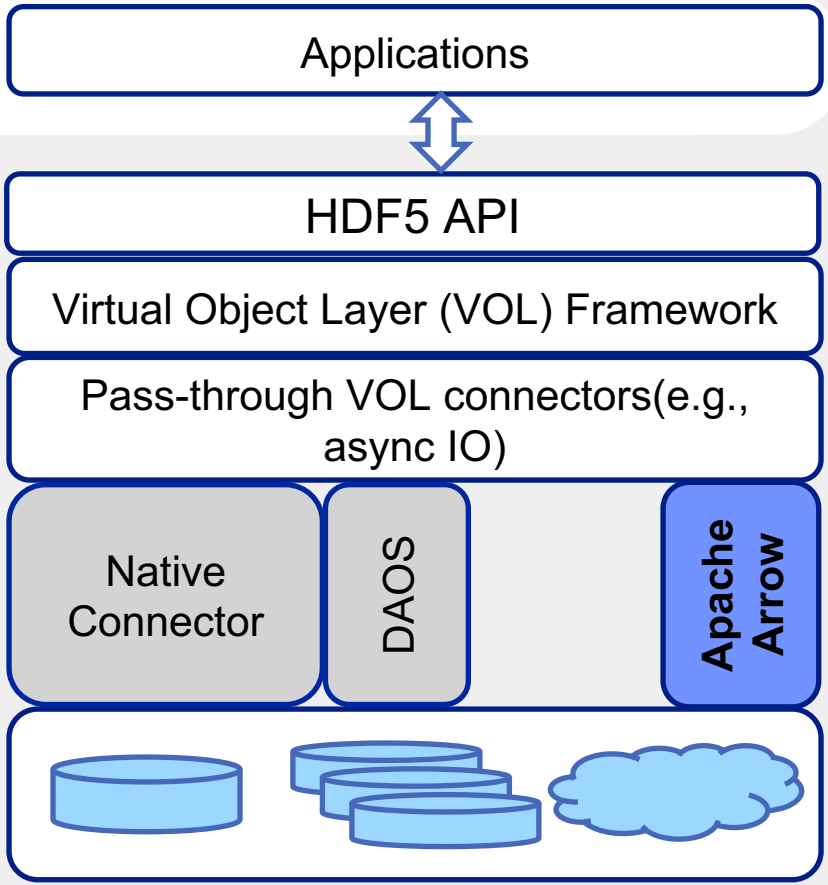


Fig2. The location of Apache Arrow within VOL

Design

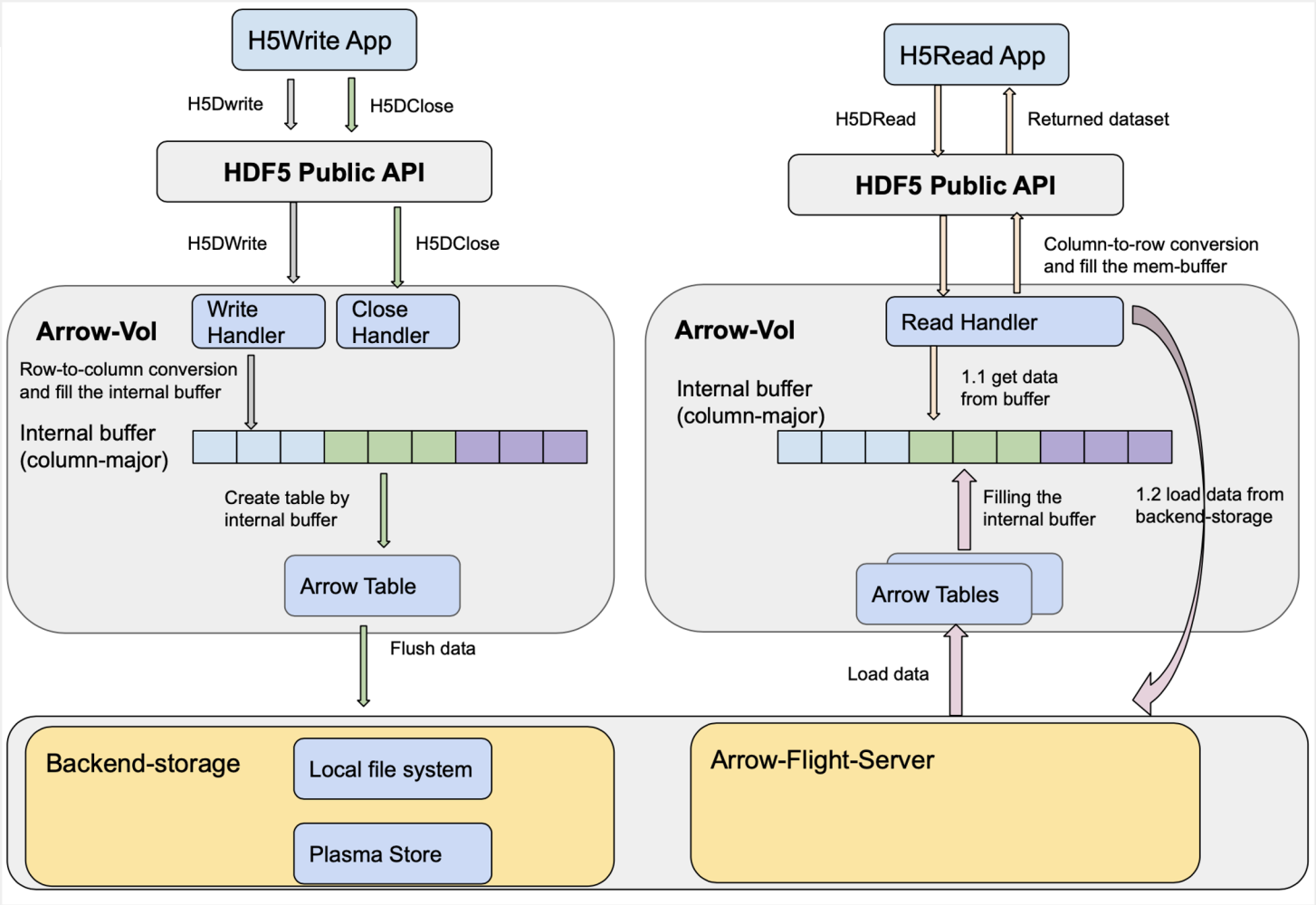


Fig3. The internal workflow in Arrow-VOL connector

Testbed

- **Cori Supercomputer**
  - 8 nodes and 128 processes
- **Lustre file system configuration for native hdf5 and arrow-vol**
  - Stripe-size: 16M
  - Stripe-count: 1 for arrow-vol; 8 for native hdf5
- **Dataset**
  - Test three different 2D-Array compound-type physical particles dataset
    - 4M: dataspace (2048, 2048)
    - 8M: dataspace (4096, 2048)
    - 16M: dataspace (4096, 4096)
- **Benchmark Tool**
  - H5bench_write & H5bench_read

I/O patterns tested

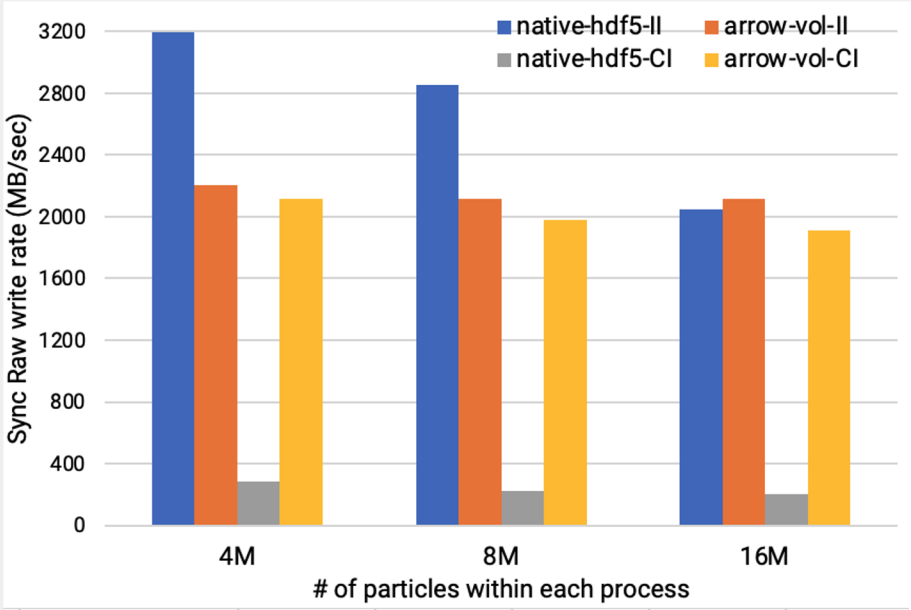
- I/O Patterns
  - H5bench_write – CI, II
  - H5bench_read
    - Read the entire data – CI, II
    - Read a subset of data – CI, II

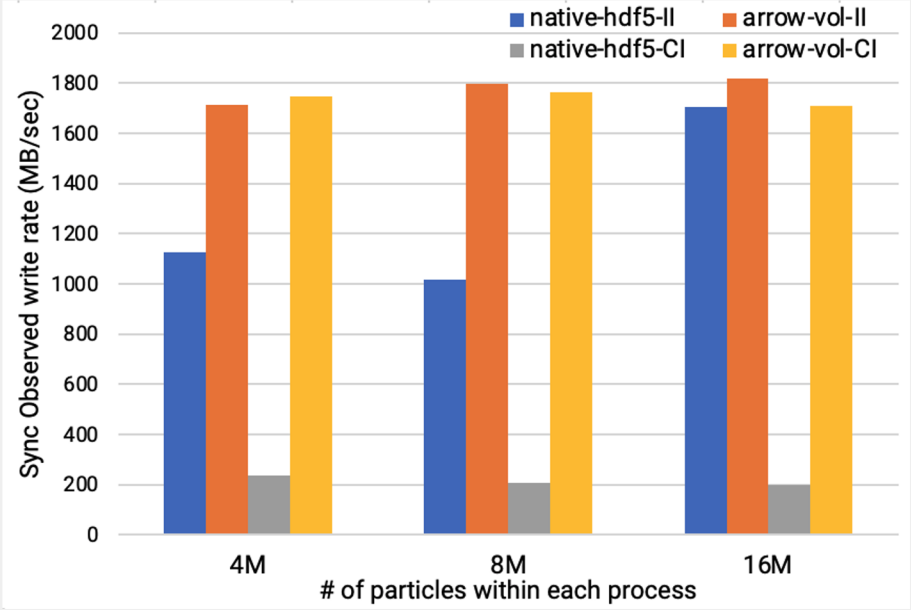
I/O patterns	memory representation	file representation
contiguous in memory and compound in file (CI)		
compound in memory and compound in file (II)		

Write Performance

- In left figure, arrow-vol-CI is always better than native-hdf5-CI
- In right figure, arrow-vol is always better than native-hdf5

CI: Contiguous in memory and Compound in file
II: Compound in memory and Compound in file

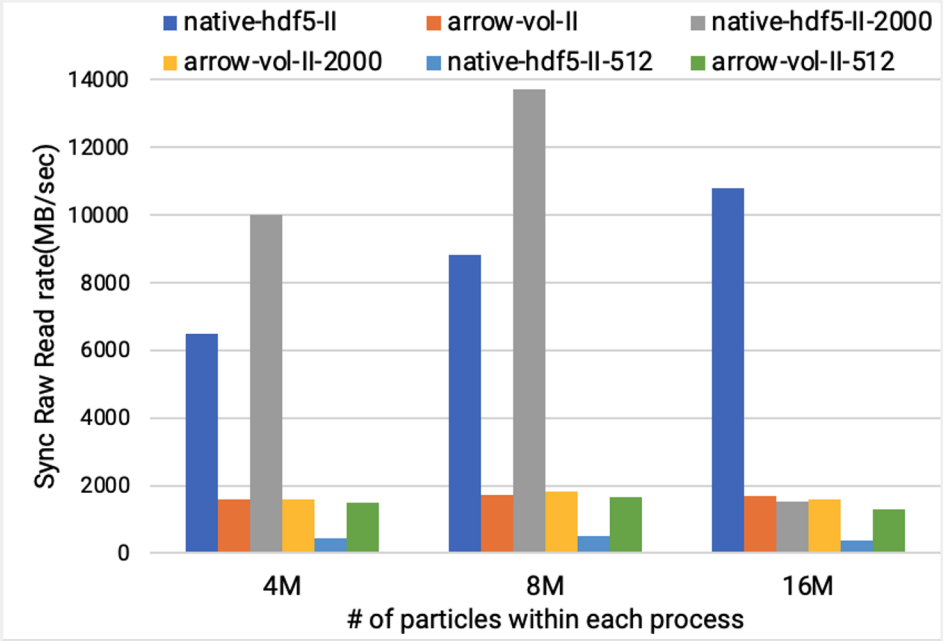


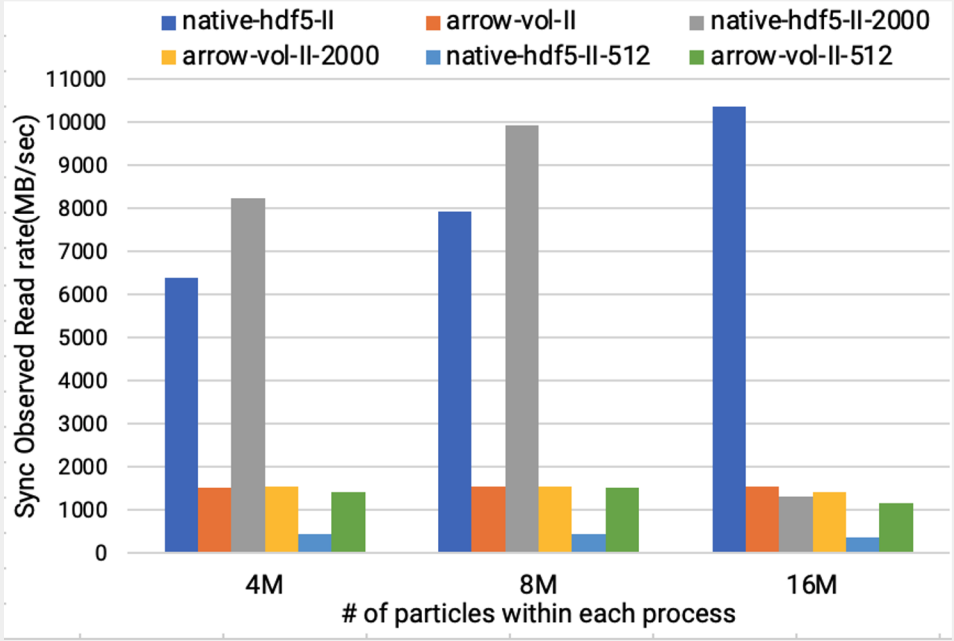


Read Performance

- The performance of arrow-vol is almost same for each dataset
- Only arrow-vol-II-512 is better than native-hdf5

II: Compound in memory and Compound in file

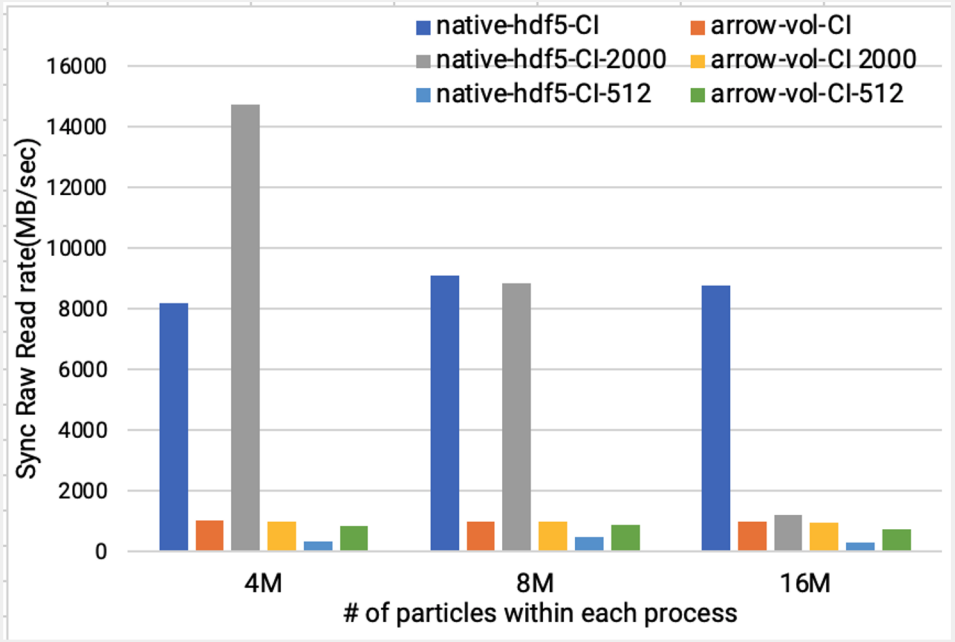


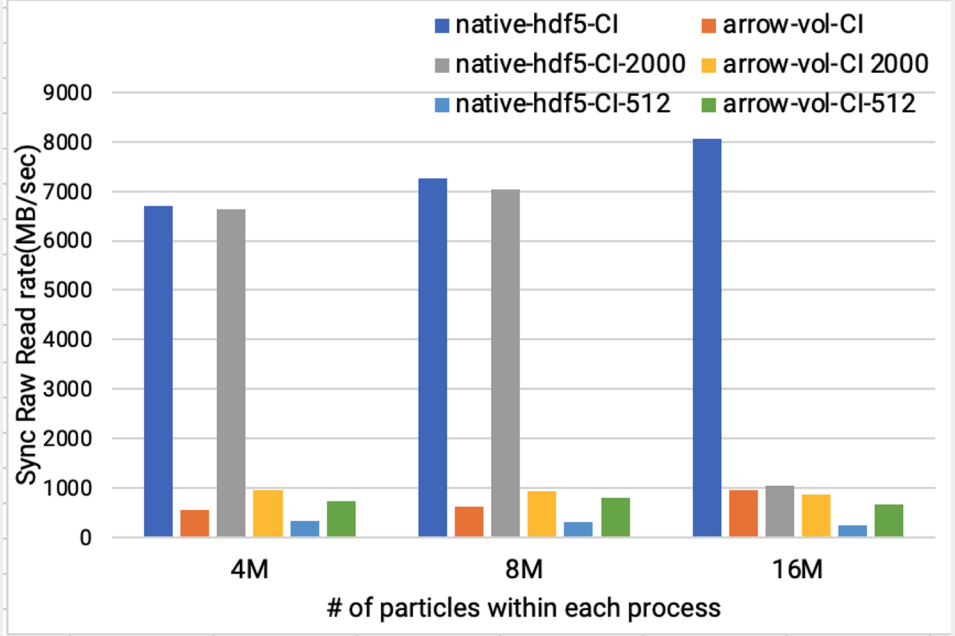


Read Performance (Con.)

- The performance of arrow-vol is almost same for each dataset
- Only arrow-vol-CI-512 is better than native-hdf5

CI: Contiguous in memory and Compound in file





Conclusion

- Developed a terminal HDF5 VOL connector to Apache Arrow
  - Enable science applications to access Apache Arrow data through native HDF5 calls
- Verified how big data technologies that offer new capabilities work in HPC systems
- Show some initial performance result when using Arrow-VOL connector
  - There is still a lot of room for optimization
  - Laid the foundation for our future work

Thank You
Q & A