



## FasTensor: Pain-free HDF5 data analysis at large scale

Bin Dong, PhD, Research Scientist, Lawrence Berkeley National Lab

Other Contributors: Kesheng Wu, Suren Byna, Houjun Tang, Weijie Zhao, Florin Rusu

HDF5 USER GROUP MEETING 2020 (HUG 2020)



## Mountains of Scientific Data Wait for Analysis



#### Light Source 180 PB/year (ALS-U at Berkeley Lab)



#### Genomics 10 PB/year



#### High Energy Physics 200 PB/year



#### Climate 100 EB/year

Library usage on Cori and Edison in 2017

Libraies

Most are multidimensional arrays, stored in file formats like HDF5, PNetCDF, ADIOS, etc

Sources: L. Nowell, D., Ushizima, S. Byna, JGI and ALS at LBNL, etc.



## Parallel Data Programming Model is a Lifesaver

Data programming model: an programming abstraction, hiding complexities of hardware/software and being generic to many advanced analysis tasks





#### MapReduce's Limitations in Tensor Data Analysis



# FasTensor (previously ArrayUDF): A new data programing model on array



Inspiring by: a **Tensor** can be defined as a multidimensional **array** and proper **transformation** rules

- 1, An Array-native Data Programming Model namely SLOPE
- 2, A Stencil-based Abstract Data Type
- 3, A single Transform operation
- 4, HPC friendly
  - MPI based Single Program Multiple Data (SPMD) Pattern
  - Directly on scientific data formats, e.g., HDF5, ADIOS, PNetCDF
  - Manual/auto-chunking & ghost zone management
  - Distributed array cache
  - Async ghost zone exchange
  - Check-point
  - In-place modification semantic
  - Multi-arrays supports

Bin Dong, Kesheng Wu, Surendra Byna, Jialin Liu, Weijie Zhao, Florin Rusu, **"ArrayUDF:** User-Defined Scientific Data Analysis on Arrays", HPDC 2017





#### Stencil: Abstract Data Type

#### Stencil

- An abstract data structure to represent a neighborhood of an Array
- Definition: **S**(Base Cell, Neighbor Cells -- relative offsets)





Flexible geometric shapes/size to break an array into small units for atomic, out-of-core, or parallel processing



## SLOPE Programming Model of FasTensor

<u>'Structural Lo</u>cality'-aware <u>Programming model</u>



## An Example of 3-point Moving Average





#### FasTensor in Earth Science (DAS)

Octorial Room

lating attend

-----



Gr

**RICE** 

BERKELEY

Berkeley m4.4 01/04/2018

## FasTensor in Plasma physics (VPIC)





Bin Dong, Patrick Frank Heiner Kilian, Xiaocan Li, Fan Guo, Suren Byna and Kesheng Wu, "Terabyte-scale Particle Data Analysis: An ArrayUDF Case Study", SSDBM 2019, July 23, 2019,

$$\begin{split} E_{px} =& (1-dY)(1-dZ)E_x(i\times d_x,(j-0.5)\times d_y,(k-0.5)\times d_z)/4+\\ & (1+dY)(1-dZ)E_x(i\times d_x,(j+0.5)\times d_y,(k-0.5)\times d_z)/4+\\ & (1-dY)(1+dZ)E_x(i\times d_x,(j-0.5)\times d_y,(k+0.5)\times d_z)/4+\\ & (1+dY)(1+dZ)E_x(i\times d_x,(j+0.5)\times d_y,(k+0.5)\times d_z)/4 \end{split}$$

- Field Data Analysis
- Particle Data Analysis
- Fusing Particle Data and Field Data

FastTensor can express all these diverse analysis



Without Considering Particle Locality



## Comparing FasTensor, TensorFlow, Spark

2 layers CNN with forward steps on a 64K by 64K 2D array





## FasTensor Scales Perfectly on CNN Computation





#### Summary

- A formal release of FasTensor is at https://bitbucket.org/berkeleylab/fastensor/
- We are looking for more use cases to provide dedicated support and make FasTensor better
- HDF5 feature request
  - Parallel virtual dataset

#### Acknowledgement

- U.S. Department of Energy (DOE), Office of Science, Office of Advanced Scientific Computing Research under contract number DE-AC02-05CH11231
- National Energy Research Scientific Computing Center (NERSC)

## **Thank You & Questions**

#### **FasTensor: Pain-free HDF5 data analysis at large scale**

https://bitbucket.org/berkeleylab/fastensor/

Bin Dong dbin@lbl.gov https://crd.lbl.gov/dong