Benchmarking Using AirSpeed Velocity

Parth Tripathi Mentors: Josh Moore and Sanket Verma

Introduction

Zarr is a NumPy-dependent format for chunked, compressed, N-dimensional array storage that takes design cues from HDF5, h5py, and bcolz.It also deals with matrices and vectors

ASV is a tool for benchmarking Python packages over their lifetime. Runtime, memory consumption and even custom-computed values may be tracked. The results are represented in the form of web portal.

How does ASV work with Zarr?

- This project works on the implementation of benchmarking in Zarr using Airspeed Velocity(ASV) in order to catch any changes that introduce a performance regression.
- The project began from the ground up, but with time, benchmarks were updated to show the trend of various metrics over time and define a number of simple metrics for reading and writing Zarr files.

Outcomes

- 1. Zarr now has a benchmark folder which is a valuable addition to the Zarr organization.
- 2. An asv.conf.json file which is very important to keep the benchmarks updated.
- 3. Benchmarks for the creation module with all it's function.
- 4. We are able to locate regressions.

Results









Google Summer of Code

References:

- 1. Project Details
- 2. GSoC Journey Blogs
- 3. Benchmark Zarr Implementations project updates
- 4. A complete guide to work with Zarr in ASV
- 5. GitHub
- 6. ASV Documentation
- 7. Zarr Documentation