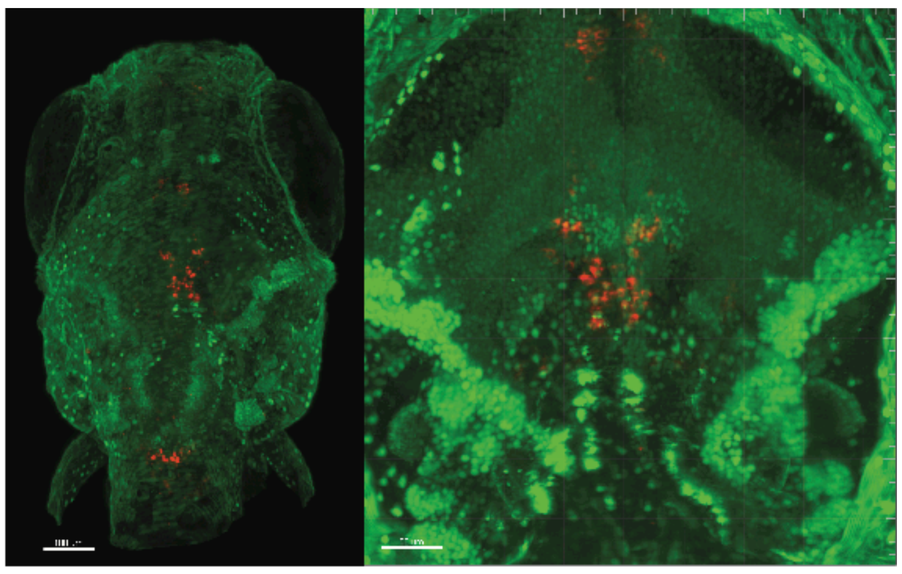
Calculation of Complete Zebrafish Brain Neural Activities on ABCI

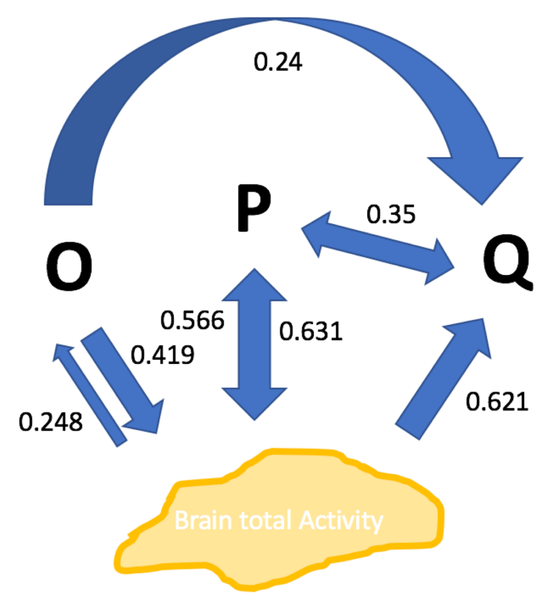
Wassapon Watanakeesuntorn, Kohei Ichikawa, Keichi Takahashi, Jason Haga, Gerald Pao
Email: wassapon.watanakeesuntorn.wq0@is.naist.jp

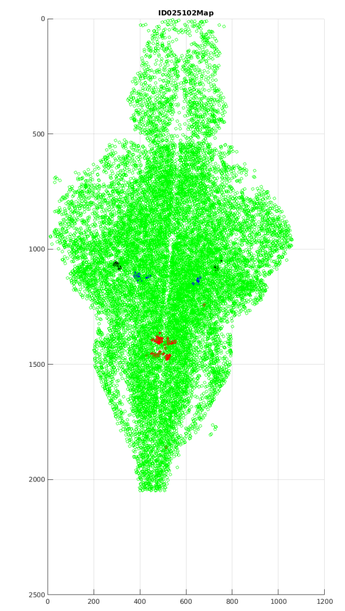
Introduction

Zebrafish Neural Activity Maps for Novel Neuromorphic Deep Learning Architectures is a project from Scripps Institution of Oceanography, UC San Diego, and Salk Institute collects and analyzes Zebrafish neural activity for using with CCM and other tools from EDM to find the relationships and generate the neural activity network of the transparent larval fish brain. EDM (Empirical Dynamic Modeling) is a suite of mathematical tools developed by Sugihara lab at UCSD based on the generalized Takens embedding theorem for the analysis of nonlinear time series. Among these, CCM (Convergent Cross Mapping) allows the inference of causation from nonlinear time series even with substantial noise and complete absence of correlation. Multiple datasets of neural activity were collected. Each dataset is around 1,600 timesteps in length sampled at 2 hertz containing 70,000-80,000 active neurons in each case which makes this becomes a big data problem for the EDM framework.

We aim to accelerate the EDM framework by using CUDA architecture for vector and matrix calculation on EDM and CCM. The goal is to adapt these calculations with the AI Bridging Cloud Infrastructure (ABCI) architecture which is the most powerful supercomputer in Japan. ABCI is operated by National Institute of Advanced Industrial Science and Technology (AIST).





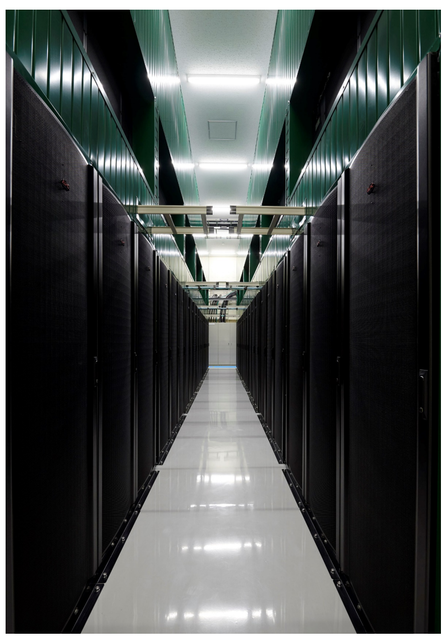


Approaches

Our goal is to run the CCM calculations of the complete zebrafish brain datasets using ABCI supercomputer. Our proposed development milestones are:

1. Profiling and CUDA enabling the EDM calculation code
2. Singularity containerizing and testing the code on ABCI
3. Evaluate the performance and optimize the code
4. Execute CCM calculation with zebrafish brain dataset

ABCI
AI Bridging Cloud Infrastructure



Development and Progress

Sugihara Lab develops 3 EDM library includes with rEDM, pyEDM, and cppEDM. We aim to use one of these libraries calculating the CCM on ABCI. We have to profile the code for finding hotspots to optimize and accelerate the code with CUDA. The optimized code is in development and currently testing on the local machine only.

- rEDM: EDM library for R
  - Using R code as an interface of input-output and C++ code as the calculate code that causes an overhead between languages
  - Some part of the code may not compatible with CUDA behavior
- pyEDM: EDM library for Python
  - Python has a slow execution compared to other languages
- cppEDM: EDM library for C++
  - Profiler shows a distance calculation takes the most CPU time
  - `valarray` datatype is cause an allocate memory overhead
  - Tried to CUDA enabling on find nearest neighbor function