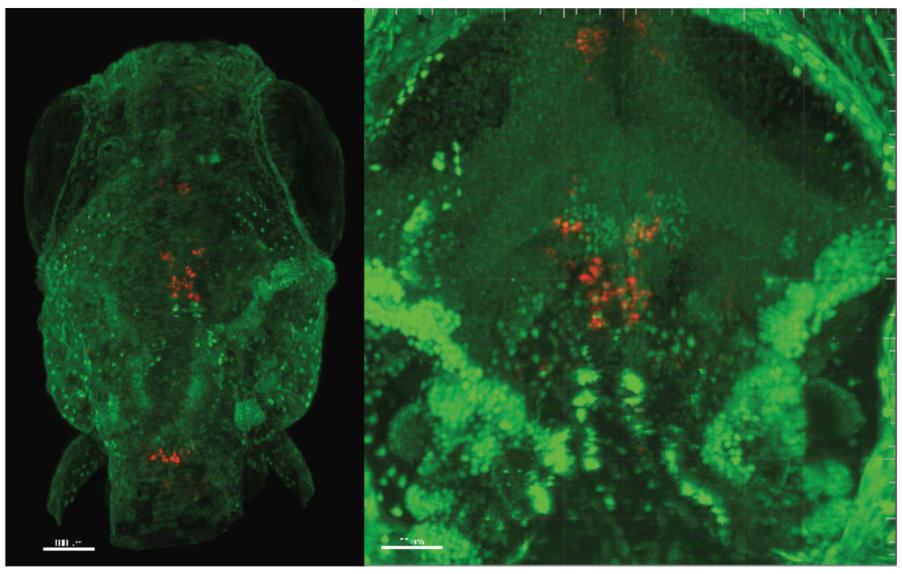
Modeling of Complete Zebrafish Brain Neural Activities on ABCI

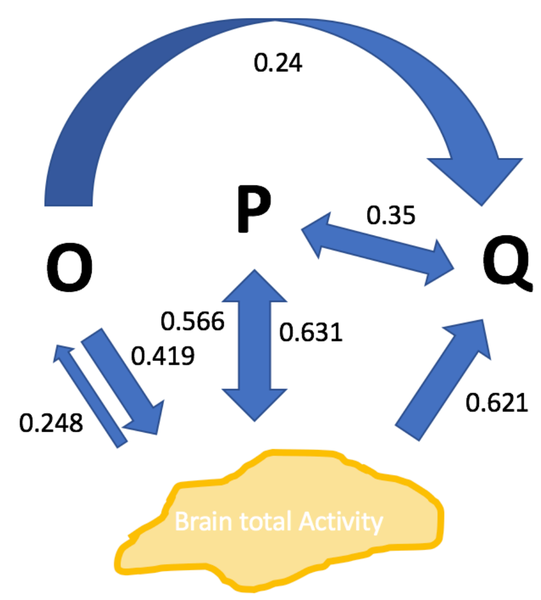
Wassapon Watanakeesuntorn, Kohei Ichikawa, Keichi Takahashi, Jason Haga, Joseph Park, Gerald Pao
Email: wassapon.watanakeesuntorn.wq0@is.naist.jp

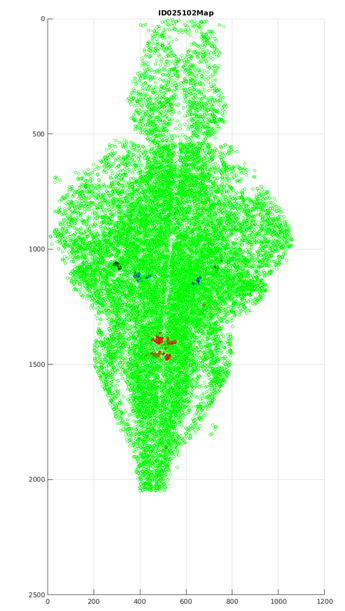
Introduction

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).





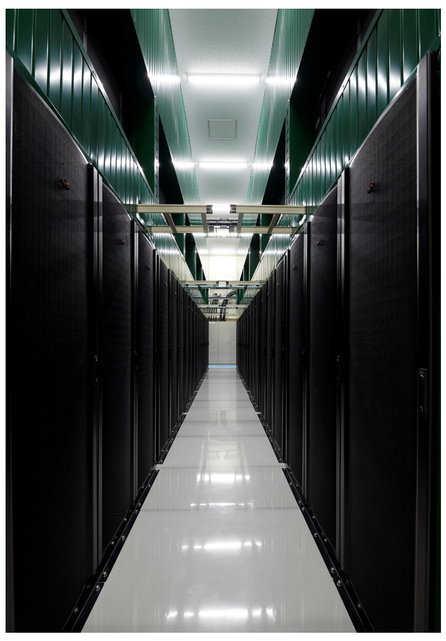


Approaches

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 three EDM libraries: rEDM, pyEDM, and cppEDM. We aim to use cppEDM for calculating the CCM on ABCI. We have profiled the code to find hostspots and accelerate the code using CUDA. The optimized code is under development and currently being testing on a local workstation.

- cppEDM: EDM library for C++
  - Developed an application for running the CCM calculation with a real dataset
  - Profiling with Intel VTune Amplifier revealed that:
    - Nearest neighbor calculation takes most of the CPU time
    - Memory allocation functions use high CPU time (`__GI_, operator new`)
  - Optimized the code from profiling result and achieved a 2.8E9 cross map run completion in 8.5 hours, instead of 85 hours, with minimal memory consumption on ABCI
  - Trying to apply CUDA for accelerating the nearest neighbor calculation using GPU
  - Trying to apply MPI for distributing the calculation across multiple nodes