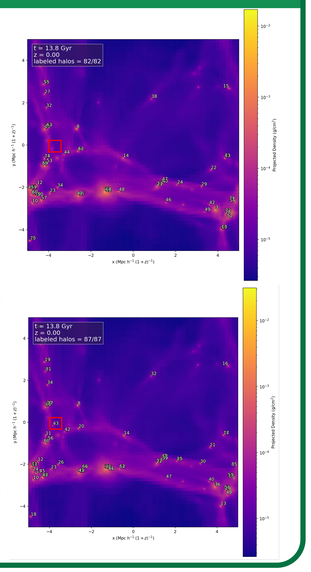
ND-Locate: Function-Level Attribution of Non-Determinism in MPI-Based Scientific Simulations

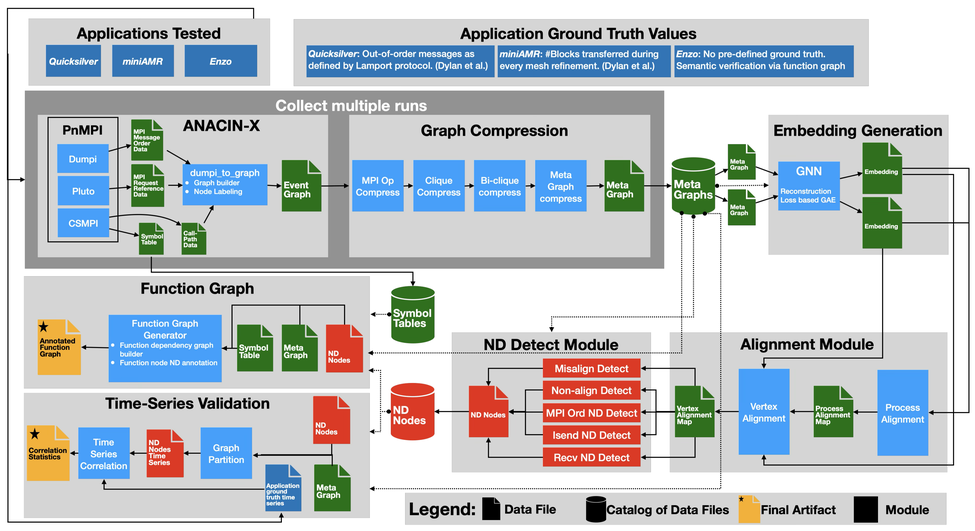
Dhroov Pandey (Student), Jack Marquez, Sanjukta Bhowmick (Advisor), Michela Taufer (Advisor)

1. Overview

Large scientific simulations leveraging MPI suffer from non-determinism (ND) induced by asynchronous communication which can have uncertain downstream effects such as out-of-order floating point calculations (unlike integers, floating point arithmetic is not commutative), which can compound and result in varying outputs for equally seeded runs. To study and quantify the regions of ND, these executions can be traced and modeled as graphs known as event graphs. We propose a framework to analyze large event graphs traced and generated by ANACIN-X via (i) graph compression preserving underlying semantics, (ii) defining 5 distinct ND node categories, (iii) formulating a workflow to compare event graphs pairwise and detect ND nodes, and (iv) generating an annotated function graph characterizing the ND within the scope of each function in a program, hence, enabling us to isolate high ND regions/functions.



2. Our Methodology



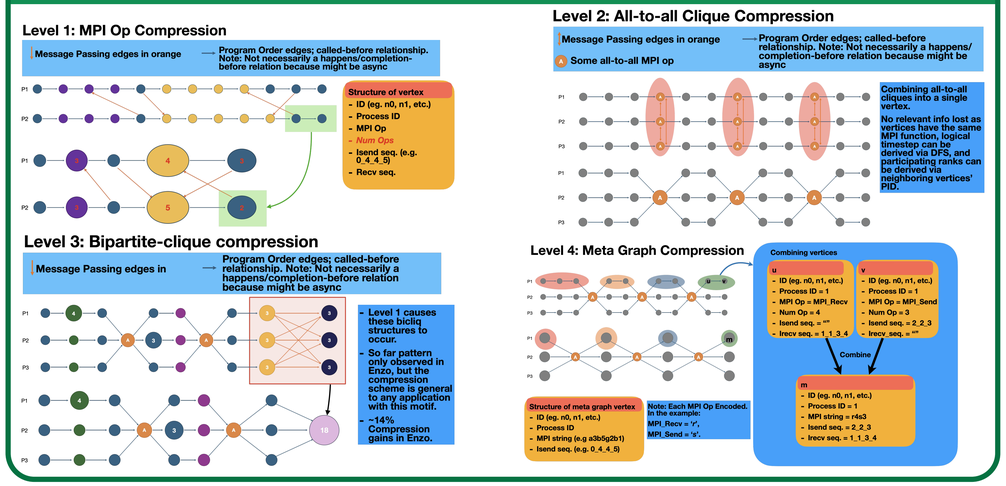
3. Application Descriptions

Enzo: An adaptive mesh refinement based astrophysics application for simulating cosmological structures.

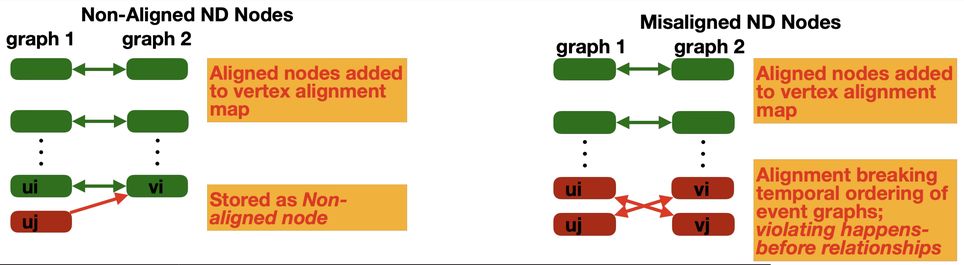
Quicksilver: A proxy application that solves a simplified dynamic Monte Carlo particle transport problem.

MiniAMR: A benchmarking application for adaptive mesh refinement that models 3D stencil calculations.

4. Graph Compression



5. ND Patterns



MPI Ord ND	Mismatch in *mpi string* of aligned meta graph nodes.
Isend ND	Mismatch in *isend sequence* of aligned meta graph nodes.
Recv ND	Mismatch in *recv sequence* of aligned meta graph nodes.

Conclusion

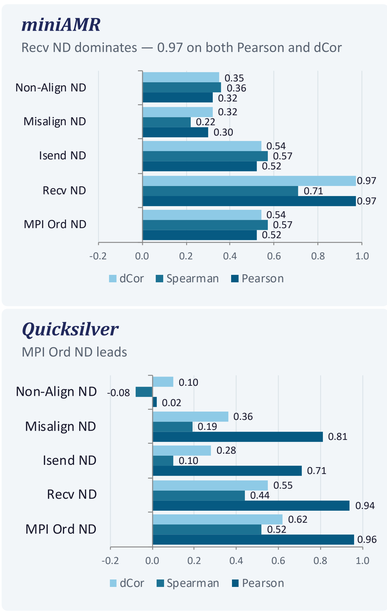
We have developed a methodology that can analyze large event graphs via semantically-smart graph compression and a multi-class ND categorization to produce an annotated function graph that is able to capture and highlight high ND functions in an HPC program. We validate our ND detection module via time-series correlation on miniAMR and Quicksilver. We also demonstrate the effectiveness of our annotated function graph generator by highlighting the high ND functions captured in miniAMR, Quicksilver, and Enzo.

Acknowledgement

The work is partially funded by NSF grants #1900765 and #1900888.

References

6. Time-Series Validation



7. Function Graph Results

