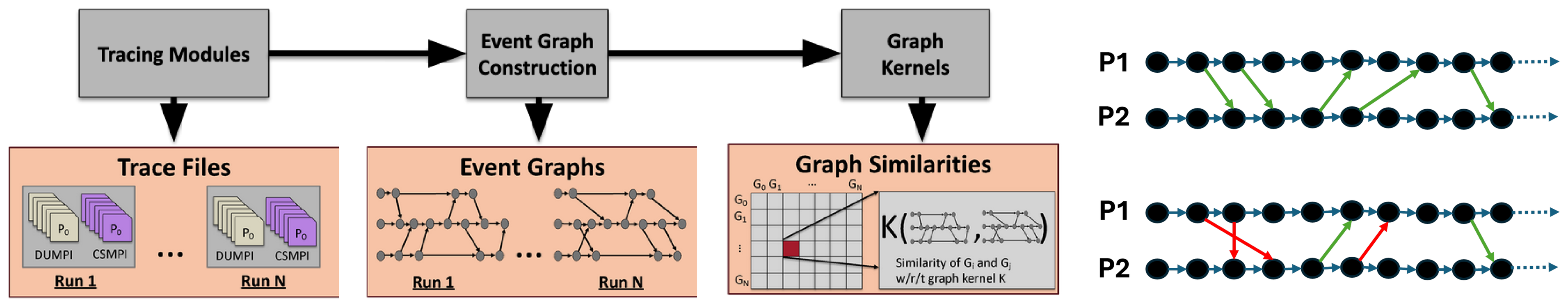
Novel Graph Alignment Algorithms for Identifying Non-Determinism in Large-Scale Simulations

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

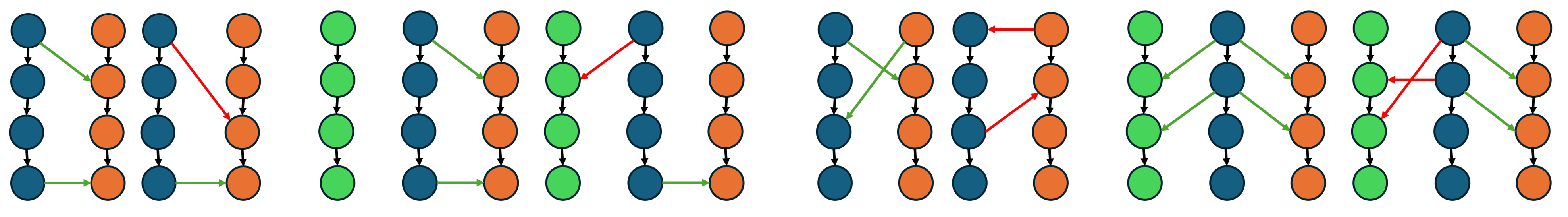
University of North Texas ; University of Tennessee at Knoxville

1. Overview

Increasing complexity of HPC simulations pose several challenges to their reproducibility and reliability. One critical issue is the non-determinism (ND) induced by asynchronous MPI communication. Locating the sources of ND in large codes is difficult. This problem can be addressed by comparing event graphs (graphs mapping MPI communication) across multiple runs of the application by using tools like ANACIN-X[2] (to trace the event graphs) and network alignment (to locate areas of ND). We expand ANACIN-X's point-to-point tracing capabilities by adding collective communication tracing, and propose a novel network alignment algorithm to effectively compare event graphs.



4. Message ND Patterns



(i) Point to point ND (ii) Varying neighbours ND (iii) MPI order ND (iv) Collective ND

5. Experimental Results

Meta-Process Graph Heuristic

MPI Application	#Procs	#Nodes	NetAlign	GAE	Meta Graph Heuristic
AMG 2013	16	1.6k	0.10	0.59	1
AMG 2013	32	2.2k	0.04	0.52	1
AMG 2013	64	330k	0.02	0.28	1
Message Race	16	660	0.14	0.87	1
Message Race	32	1.1k	0.11	0.72	1
Message Race	64	2.8k	0.10	0.70	1
MCB Grid	16	320k	0.005	0.21	1
MCB Grid	32	770k	0.001	0.17	0.875
MCB Grid	64	154k	0.02	0.19	0.938
c-amg2013	16	2.2k	0.08	0.46	1
c-mcb grid	16	270k	0.02	0.15	0.94

Experiment Settings

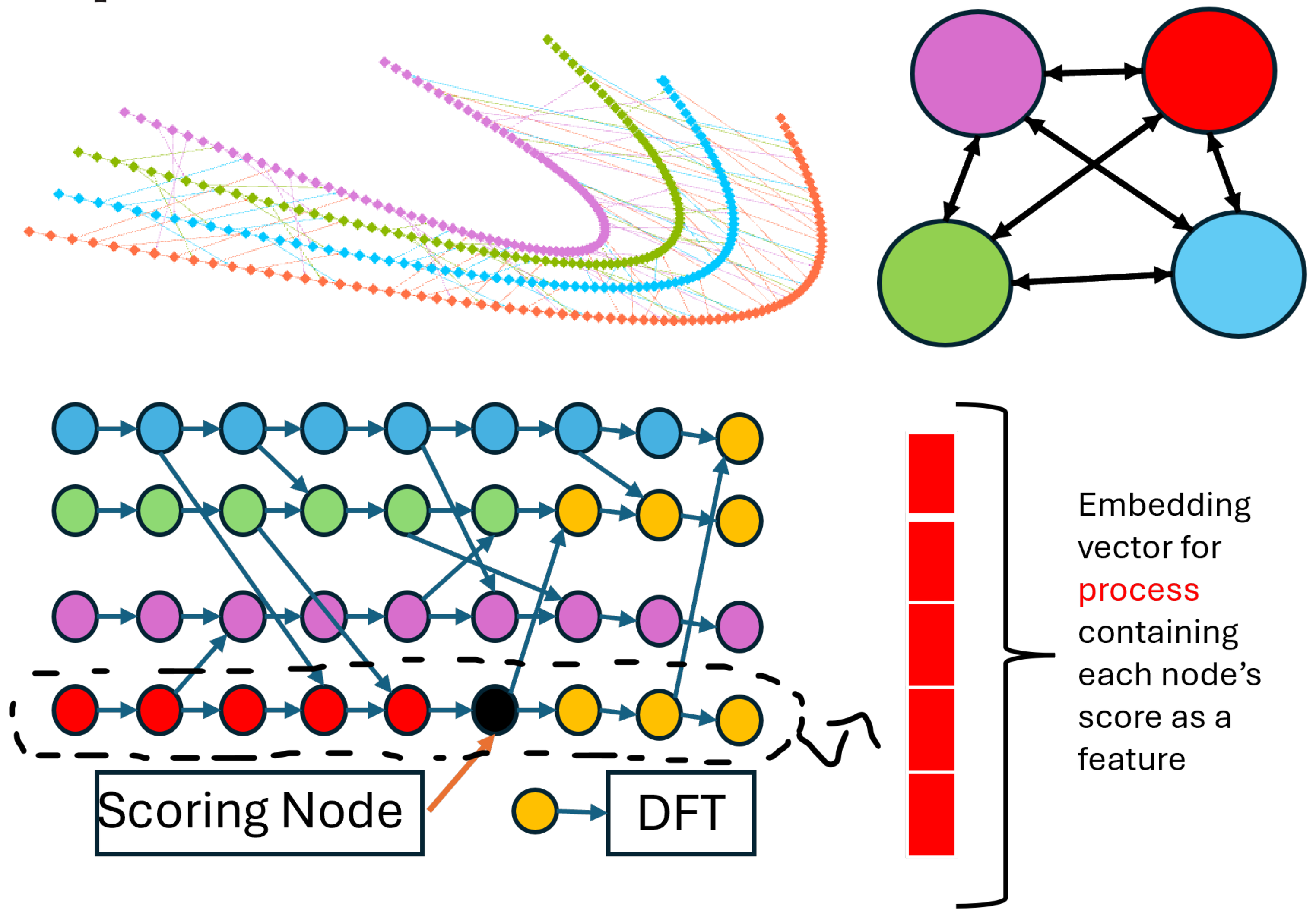
1. If the Meta-Graph alignment matched more than one node, we counted it as an incorrect matching.
2. For NetAlign, we shuffled the node labels before measuring the accuracy.

6. Conclusion

We incorporated collective communication tracing into ANACIN-X's event graph generation module. We developed an algorithm that can compare event graphs with message ND and it outperforms existing state-of-the-art. The identified ND nodes can be traced back to their function call using ANACIN-X's CSMPI module.

3. Our Contribution: Meta Process-Graph Heuristic

AMG2013 event graph with 4 processes Corresponding Meta-Process Graph



Meta-Process Graphs across event graphs are *isomorphic*. We need to find the isomorphism mapping for alignment.

Methodology

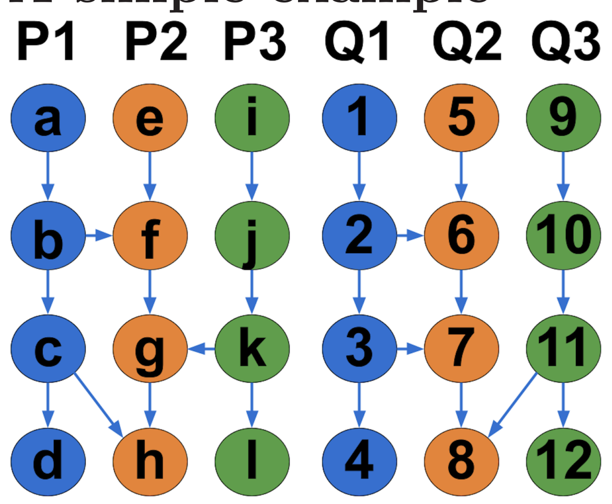
1. Score each node by encapsulating the DFS Tree numerically.
2. Use a threshold function to prune the alignment space of the meta graph.
3. Use node scores as features for embedding vectors of meta graph nodes.
4. Compare the meta graph nodes across diff. graphs which pass the threshold similarity using vector similarities. E.g. Euc Norm, Cosine, etc.

Mathematical Scoring Function

Let $S: V(G) \to \mathbb{R}$ *denote the tree scoring function defined as*

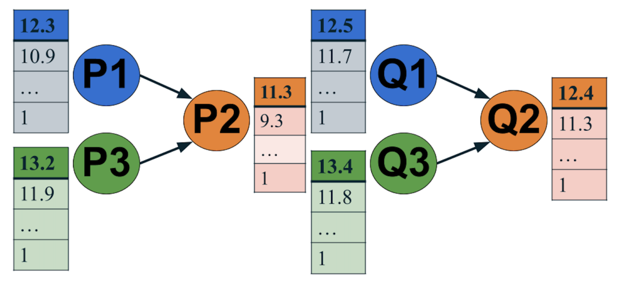
$$S(v) = \ln[Poly(lvl(v)) + \sum_{u \in Children(v)} S(u)], \text{where } lvl(v) \text{ is the reverse topological level of } v$$

A simple example

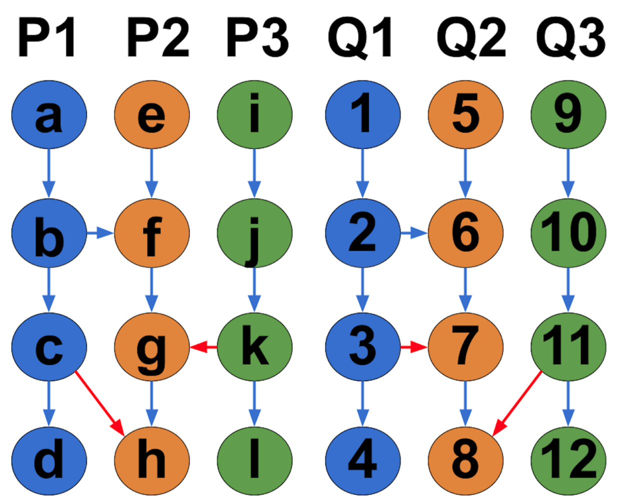


Node	Score
a	12.37
b	10.91
…	…
l	1

Node	Score
1	12.41
2	11.34
…	…
12	1



	P1	P2	P3
Q1	0.95	0.64	0.31
Q2	0.45	0.99	0.73
Q3	0.17	0.55	0.94



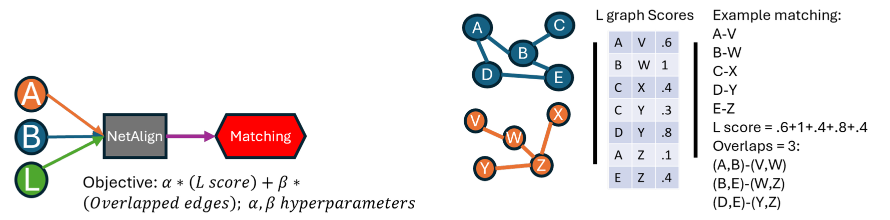
Acknowledgement

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

References

2. State-of-the-art Baselines

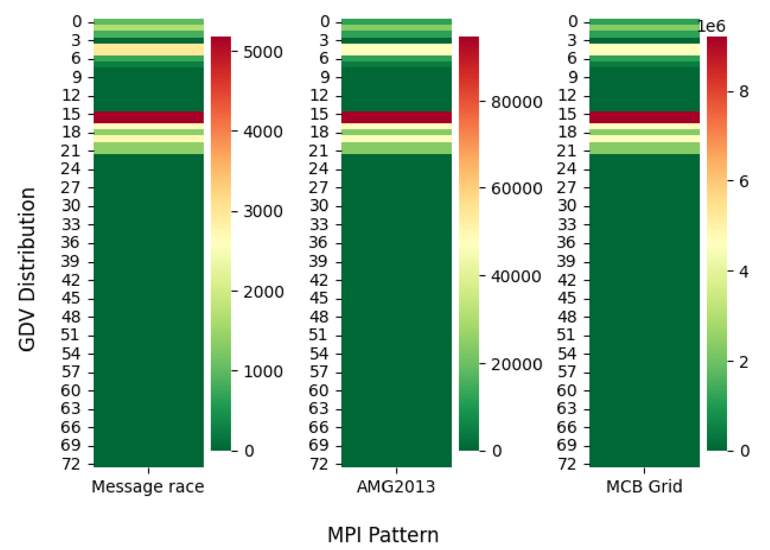
NetAlign



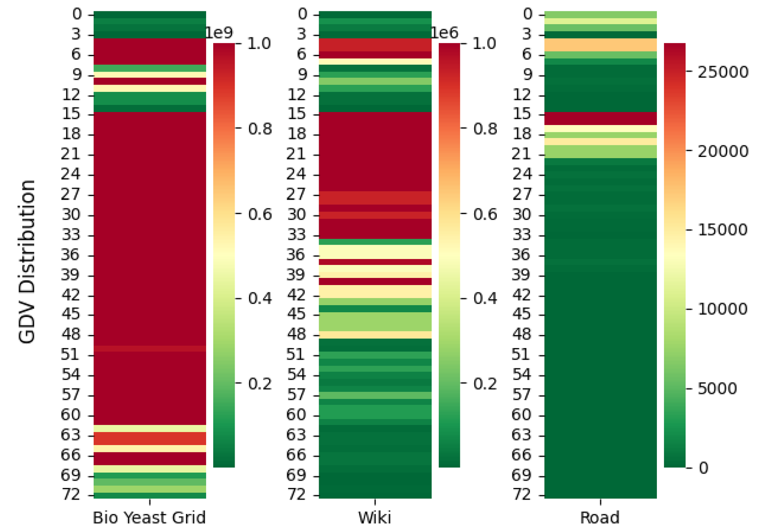
Jaccard Index Scores for Label Sensitivity

Network	#Nodes	#Edges	$\frac{\lvert E \rvert}{\lvert V \rvert^2}$	NetAlignMC	FINAL
Message Race	1K	1.2K	1.2e-3	0.1 ± 0.04	0.015 ± 0.01
AMG2013	3K	4.4K	4.9e-4	0.05 ± 0.04	0.04 ± 0.02
MCB	38K	57K	3.9e-5	0.01	0.01 ± 0.01
Bio-Yeast-Protein	2K	4K	1e-3	0.52 ± 0.02	0.54 ± 0.01
Road	1174	1417	1e-3	0.59 ± 0.04	0.63 ± 0.02
Wiki	889	2914	3.7e-3	0.93 ± 0.02	0.90 ± 0.01
Bio-Grid-Yeast	6K	314K	8.7e-3	0.95 ± 0.02	0.93 ± 0.03
Karate	34	78	0.067	1.0	1.0
Dolphin	62	159	0.041	0.99 ± 0.01	1.0

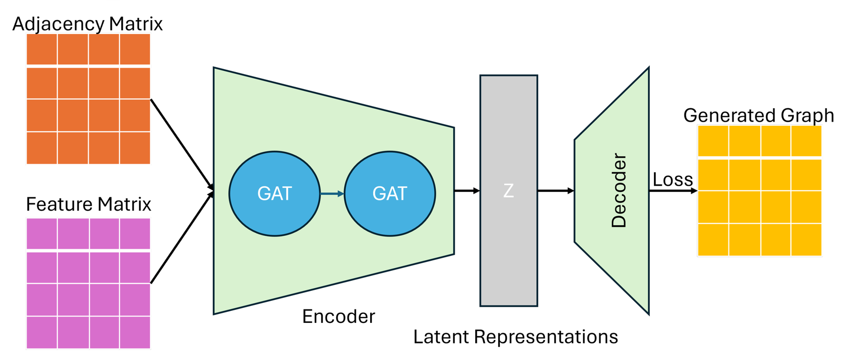
Event Graph GDV Distribution



Dense network GDV Distribution



Graph Auto Encoders



Methodology: Use latent representations to measure node similarity

Hyperparameters:
Learning Rate: 1e-4, Hidden Dim: 512, Output Dim: 16, Training Epochs: 1e3
Input Vector: [topological level] concat [One Hot MPI Function]

Limitation: Can only aggregate information from k-hop neighborhood, $k << diameter$.