A Research-Based Course Module to Study Non-determinism in High Performance Applications

Patrick Bell, Kae Suarez, Barbara Fossum, Dylan Chapp, Sanjukta Bhowmick, Michela Taufer
Communication Non-Determinism

Example of a simple MPI parallel application and its communication pattern

MPI process

Logical order of communication events
The MPI applications can run on the same computer and yet have different executions (i.e., order in which the messages are exchanged).
Communication Non-Determinism

Example of a simple MPI parallel application and its communication pattern

The MPI applications can run on the same computer and yet have different executions (i.e., order in which the messages are exchanged)

This is communication non-determinism.
Impacts of Non-Determinism

HYPRE solves linear systems of equations.

A non-deterministic catastrophic bug occurred in only 2% of executions.

Impact:
- 18 months of scientists’ time
- 9560 node-hours

Expensive Bug: HYPRE [1]

Impacts of Non-Determinism

Expensive Bug: HYPRE [1]

- HYPRE solves linear systems of equations.
- A non-deterministic catastrophic bug occurred in only 2% of executions.
- Impact:
  - 18 months of scientists’ time
  - 9560 node-hours

Variable Results: Enzo [2]

- Enzo detects galactic halos.
- Detection of galactic halo 49 is nondeterministic, harming trustworthiness of the result.

Motivation

We present the key components (software and platform; training and tests; and reasoning) for a course module to introduce computer science students to the concept in non-determinism.
Motivation

We present the key components (software and platform; training and tests; and reasoning) for a course module to introduce computer science students to the concept in non-determinism.

The course module is currently in development. We seek feedback on ways to use the material in class.
Software and Platforms

MPI Mini-Applications

ANACIN-X: analysis of non-deterministic executions*

Software and Platforms

**ANACIN-X**: A Software Framework for Studying Non-determinism in MPI Applications.

Software and Platforms

ANACIN-X: analysis of non-deterministic executions*

- Visualization of communication patterns
- Measurement of non-determinism

MPI Mini-Applications

Software and Platforms

MPI Mini-Applications

ANACIN-X: analysis of non-deterministic executions*

- Visualization of communication patterns
- Measurement of non-determinism
- Identification of non-determinism sources

Software and Platforms

- GitHub\(^1\): Open source access to code and visualization tools.
- Code Ocean\(^2\): Reproducible tests in a cloud environment when a student can’t use other HPC platforms.
- XSEDE Jetstream cloud\(^3\): Dependencies all installed into an image


\(^1\) ANACIN-X GitHub: [https://github.com/TauferLab/ANACIN-X](https://github.com/TauferLab/ANACIN-X)


\(^3\) Jetstream image: [https://use.jetstream-cloud.org/application/images/1056](https://use.jetstream-cloud.org/application/images/1056)
MPI Mini-Applications

Message race

Simple receiver-side non-determinism
MPI Mini-Applications

Message race

AMG 2013[1]

Simple receiver-side non-determinism

Non-determinism propagates to sender-side
MPI Mini-Applications

Message race

AMG 2013[1]

Unstructured Mesh[2]

Simple receiver-side non-determinism

Non-determinism propagates to sender-side

Non-determinism extends beyond permuted message order

Non-deterministic process topology
Overview of Course Modules

MPI application components:
Overview of Course Modules

MPI application components:

Course module components:

<table>
<thead>
<tr>
<th>Level</th>
<th>Task</th>
</tr>
</thead>
<tbody>
<tr>
<td>Beginner Level:</td>
<td>Visualize non-determinism in different communication patterns</td>
</tr>
<tr>
<td>Intermediate Level:</td>
<td>Measure non-determinism associated to different input parameters</td>
</tr>
<tr>
<td>Advanced Level:</td>
<td>Identify the root sources of non-determinism in different runs</td>
</tr>
</tbody>
</table>
Beginner Level - Visualize Non-determinism

MPI
Mini-Applications

ANACIN-X: analysis
of non-deterministic
executions

- Visualization of communication patterns
- Measurement of non-determinism
- Identification of non-determinism sources

* Figures generated by the ANACIN-X software. Test settings listed in paper.
Visualize Non-determinism - Event Graphs

\[ P_0 \]

\[ P_1 \]

\[ P_2 \]
Visualize Non-determinism - Event Graphs

Vertices:
- Blue circles represent MPI Send
- Red circles represent MPI Receive

P₀

P₁

P₂

Vertices: represent MPI communication calls
Visualize Non-determinism - Event Graphs

Vertices:
- represent MPI communication calls

Within-Process Edges:
- represent logical precedence

- $P_0$
- $P_1$
- $P_2$

- $\text{MPI}_\text{Send}$
- $\text{MPI}_\text{Receive}$
Visualize Non-determinism - Event Graphs

Vertices: represent MPI communication calls
Within-Process Edges: represent logical precedence
Between-Process Edges: represent MPI point-to-point messages

= MPI_Send
= MPI_Receive
An event graph models an MPI communication pattern in a graph format.
Beginner Level - Visualize Non-determinism

**ANACIN-X: analysis of non-deterministic executions**

- MPI Mini-Applications
- Visualization of communication patterns
- Measurement of non-determinism
- Identification of non-determinism sources

* Figures generated by the ANACIN-X software. Test settings listed in paper.
Beginner Level - Visualize Non-determinism

**ANACIN-X: analysis of non-deterministic executions**

- Visualization of communication patterns
- Measurement of non-determinism
- Identification of non-determinism sources

Students visualize communication patterns over multiple runs of the same code on the same platform, identifying the order in which messages are exchanged.

* Figures generated by the ANACIN-X software. Test settings listed in paper.
Intermediate Level - Measure Non-determinism

MPI
Mini-Applications

ANACIN-X: analysis of non-deterministic executions

Visualization of communication patterns
Measurement of non-determinism
Identification of non-determinism sources

* Figures generated by the ANACIN-X software. Test settings listed in paper.
Quantify Non-determinism - Graphs Kernels

\[ K(G, G') \rightarrow \text{Similarity Score} \]

**Intuition:** \( K \) counts matching substructures

- A **match** increases the score
- A **difference** does not increase the score
Quantify Non-determinism - Graphs Kernels

\[ K(G, G') \rightarrow \text{Similarity Score} \]

\( K(G, G' \cup G') = \sqrt{(K(G, G) + K(G', G') - 2K(G, G'))} \)

**Intuition:** \( K \) counts matching substructures
- A **match** increases the score
- A **difference** does not increase the score

Graph kernel \( K \) **induces a metric of non-determinism**

Intermediate Level - Measure Non-determinism

- **MPI Mini-Applications**
- **ANACIN-X: analysis of non-deterministic executions**
  - Visualization of communication patterns
  - Measurement of non-determinism
  - Identification of non-determinism sources

Violin plots generated over 20 runs

* Figures generated by the ANACIN-X software. Full test settings listed in paper.
Intermediate Level - Measure Non-determinism

MPI Mini-Applications

ANACIN-X: analysis of non-deterministic executions

- Measurement of non-determinism
- Identification of non-determinism sources
- Visualization of communication patterns

Students experience first hand the impact of parameters, such as number of processes, on the amount of non-determinism across multiple executions.

* Figures generated by the ANACIN-X software. Full test settings listed in paper.
Advanced Level - Identify Root Sources

* Figures generated by the ANACIN-X software. Test settings listed in paper.
Multiple runs of a Mini-AMR application

Code slices divide code into sections containing callstacks for which non-determinism is quantified and compared.
Advanced Level - Identify Root Sources

MPI Mini-Applications

ANACIN-X: analysis of non-deterministic executions

- Visualization of communication patterns
- Measurement of non-determinism
- Identification of non-determinism sources

* Figures generated by the ANACIN-X software.
Advanced Level - Identify Root Sources

MPI
Mini-Applications

ANACIN-X: analysis of non-deterministic executions

Visualization of communication patterns
Measurement of non-determinism
Identification of non-determinism sources

Students identify possible root sources by counting functions with high amount of non-determinism in code slices.

* Figures generated by the ANACIN-X software.
Opportunities for Collaborations

The course module presented introduces students to three levels of understanding in non-determinism:

• visualize non-determinism
• measure impact of parameters on application non-determinism
• identify root causes of non-determinism
Opportunities for Collaborations

The course module presented introduces students to three levels of understanding in non-determinism:
• visualize non-determinism
• measure impact of parameters on application non-determinism
• identify root causes of non-determinism

We are looking for collaborations
• Help and advice in developing this course module
• Additional mini-applications for students to study in this course module
Additional questions?
Patrick Bell:  pbell12@utk.edu
Michela Taufer: taufer@gmail.com

Find out more at:
https://globalcomputing.group