

# Interactive Visual Analysis Tool for Anomaly Provenance Data

Alicia Guite, Tanzima Islam, Christopher Kelly, Wei Xu

Texas State University, Brookhaven National Labs

Chimbuko is a framework for detecting real-time performance anomalies for large-scale applications. Understanding the source of anomalous behaviors is difficult due to the high volume of information stored by the Chimbuko framework in a provenance database. In this project, we design and develop an interactive analysis and visualization tool to intuitively display this high volume of information. This project aims to facilitate user investigation of the provenance data generated by Chimbuko.

## Goal

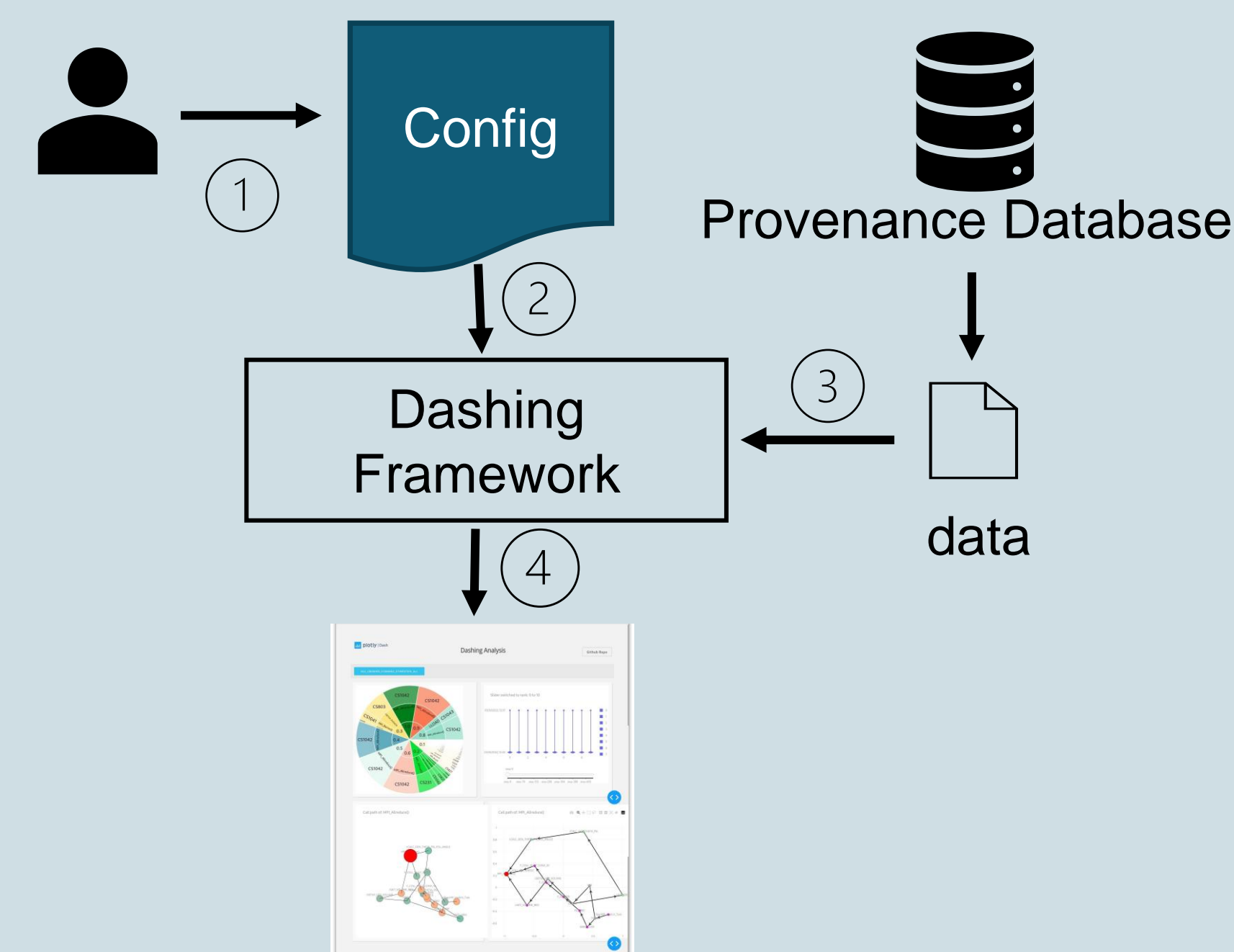
Analyze and visualize anomaly provenance data to understand root cause of performance anomalies.

## Background

- Chimbuko flags anomalies using function execution time through HBOSS algorithm then stores that data in provenance database
- Workflow applications run with many processes called "ranks"
- Chimbuko's current visualizations do not show the potential sources.
- Although function execution time is just one of metric, as Chimbuko expands their metrics, our visual analysis methods can still apply.

## Approach

- Challenge: Displaying large amount of data in an intuitive manner.
- Export provenance database into a JSON file
- Import the json file into the Dashing framework, which provides a visually intuitive interactive dashboard.



- Example config file:

```
NWChem:
data: 'data.json'
reader: util.pipeline.DataLoader.read_chimbuko_json
tasks:
- viz.chimbuko_sunburst.chimbuko_sunburst
- viz.chimbuko_violin.chimbuko_violin_symmetric
- viz.chimbuko_callgraph.chimbuko_callgraph3D
name: 'NWChem'
target: runtime_exclusive
callgraph_root: 'MPI_Barrier()'
compute_target: modules.compute_target.compute_normalized_target

main:
tasks:
- NWChem
```

## Top-Down View of Anomalies

