Granula: Toward Fine-grained Performance Analysis of Large-scale Graph Processing Platforms

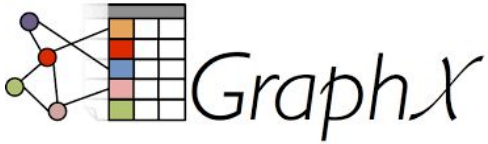
Wing Lung Ngai, Tim Hegeman, Stijn Heldens, and Alexandru Iosup

@Large Research
Massivizing Computer Systems

VU UNIVERSITY AMSTERDAM

Large-scale Graph Processing

Towards trillion vertices and edges



OpenG

PGX

GraphMat

Powergraph

Graph processing platforms

Performance Evaluation

Performance evaluation is the process of deepening the understanding of the performance by

- quantifying performance,
- explaining differences,
- identifying overheads/bottlenecks
- recommending improvements.

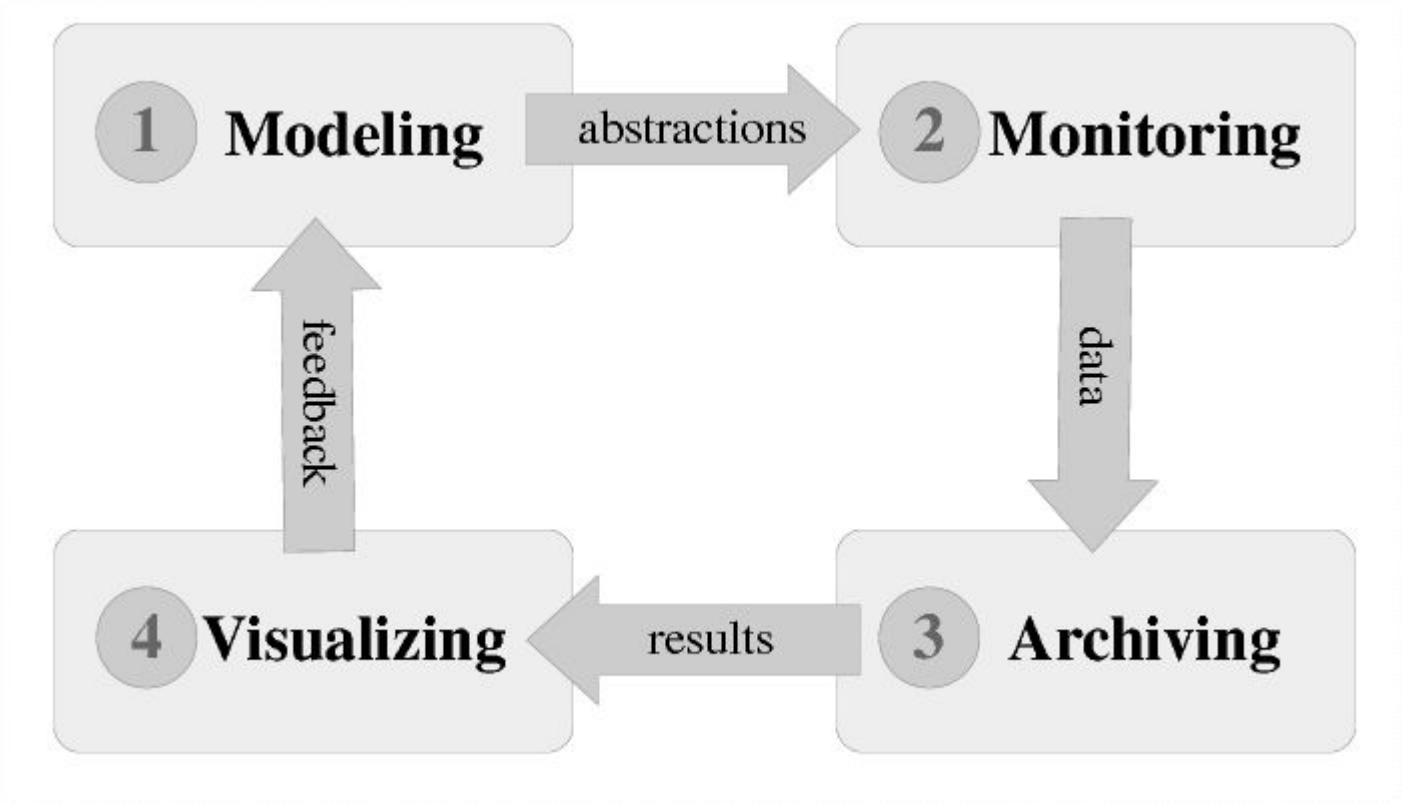
Identified Issues

- Lack of an end-to-end process
- Limited reusability of studies
- Shortcomings in coarse-grained evaluation
- Inefficiency of fine-grained evaluation

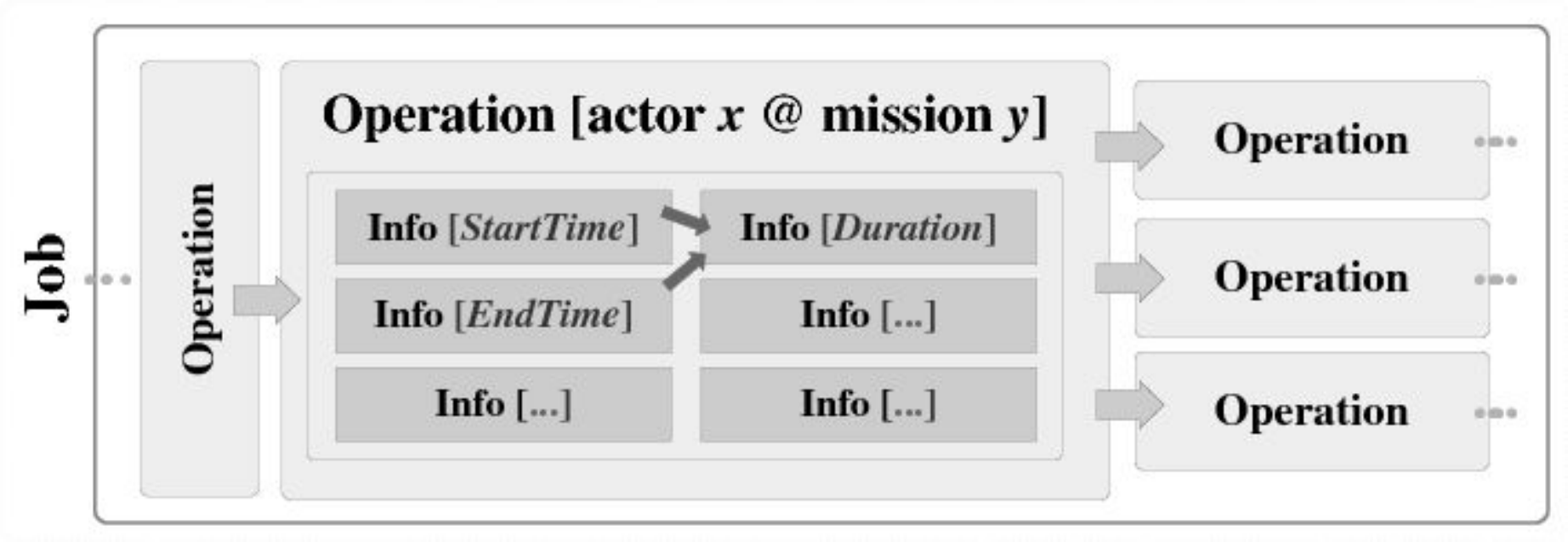
Fine-grained Performance Analysis

- Comprehensive: end-to-end evaluation process.
- Standardized: reusable across platforms.
- Incremental: iterative performance modeling.
- Automatable: automated repetitive procedures.

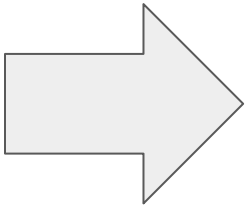
End-to-end evaluation process

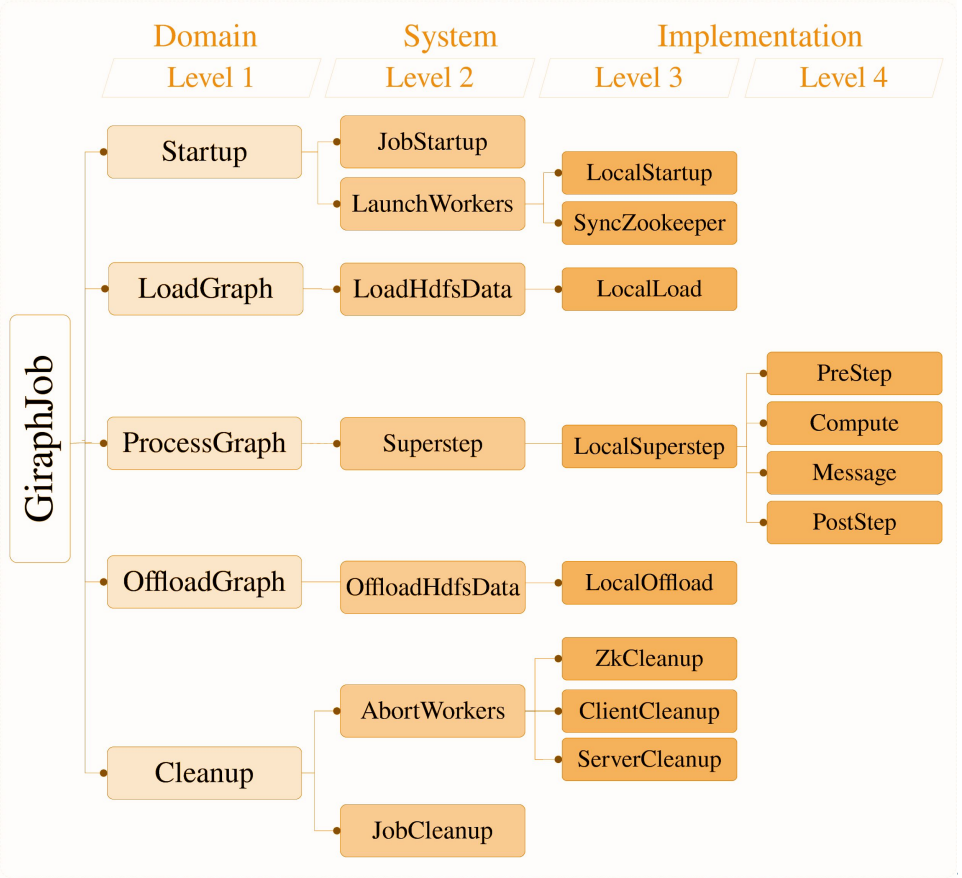


Performance Modeling Language



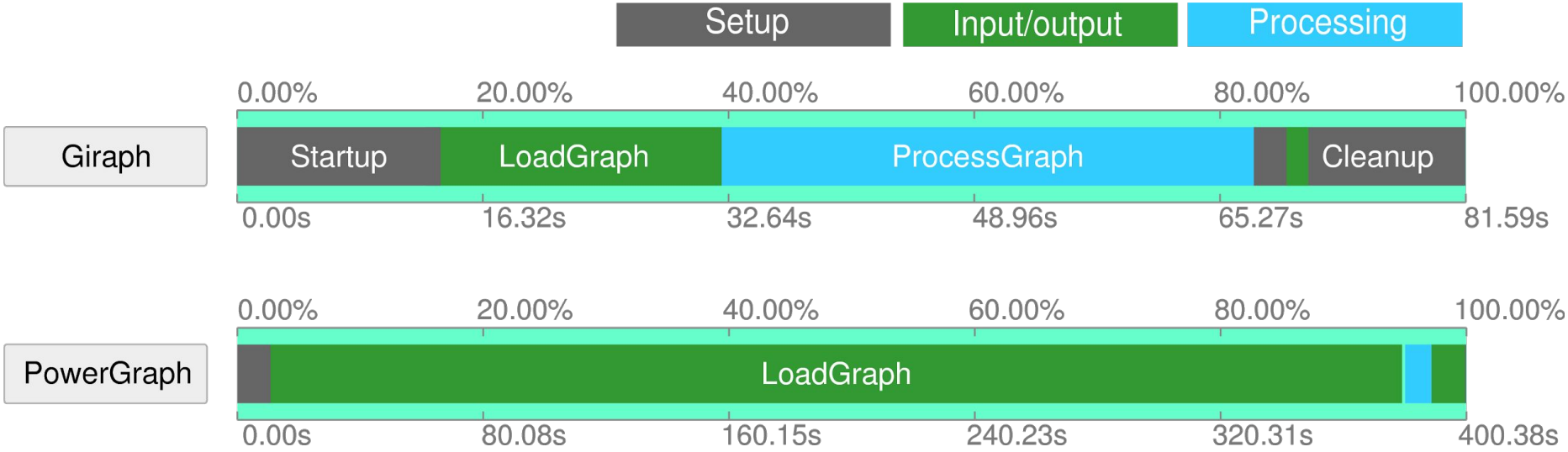
Building Performance Model





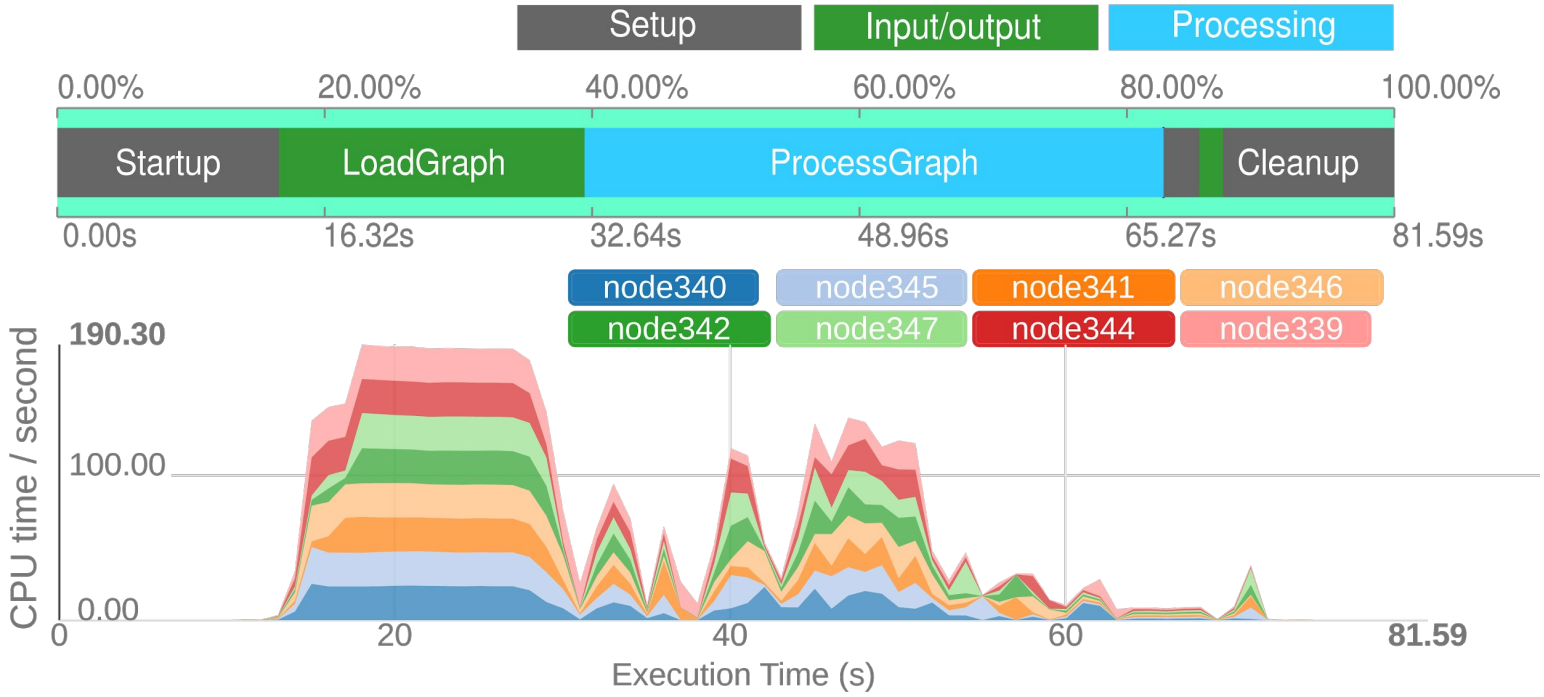
Quantifying System Performance

Running BFS on dg1000 (giraph and powergraph)



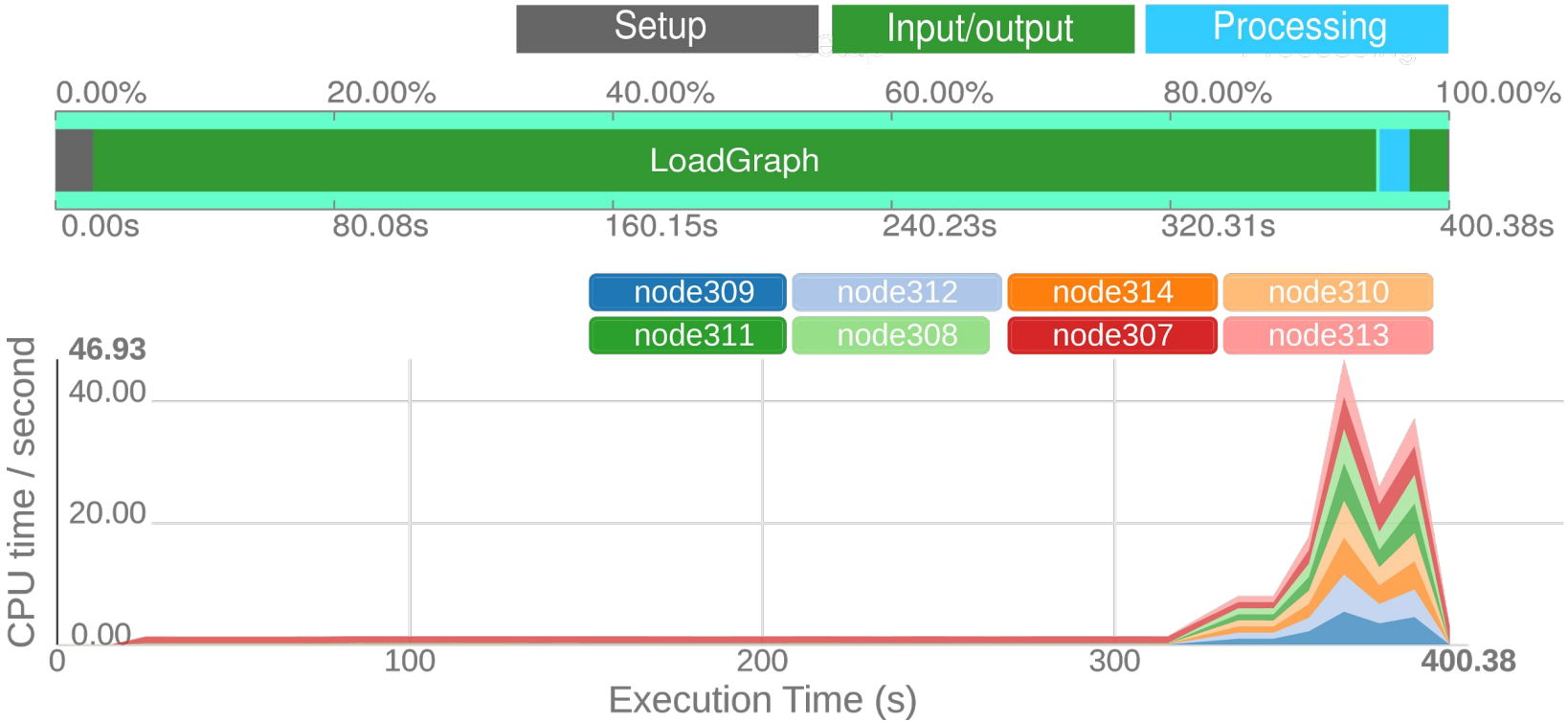
Monitoring Resource Usage

Running BFS on dg1000 (giraph)



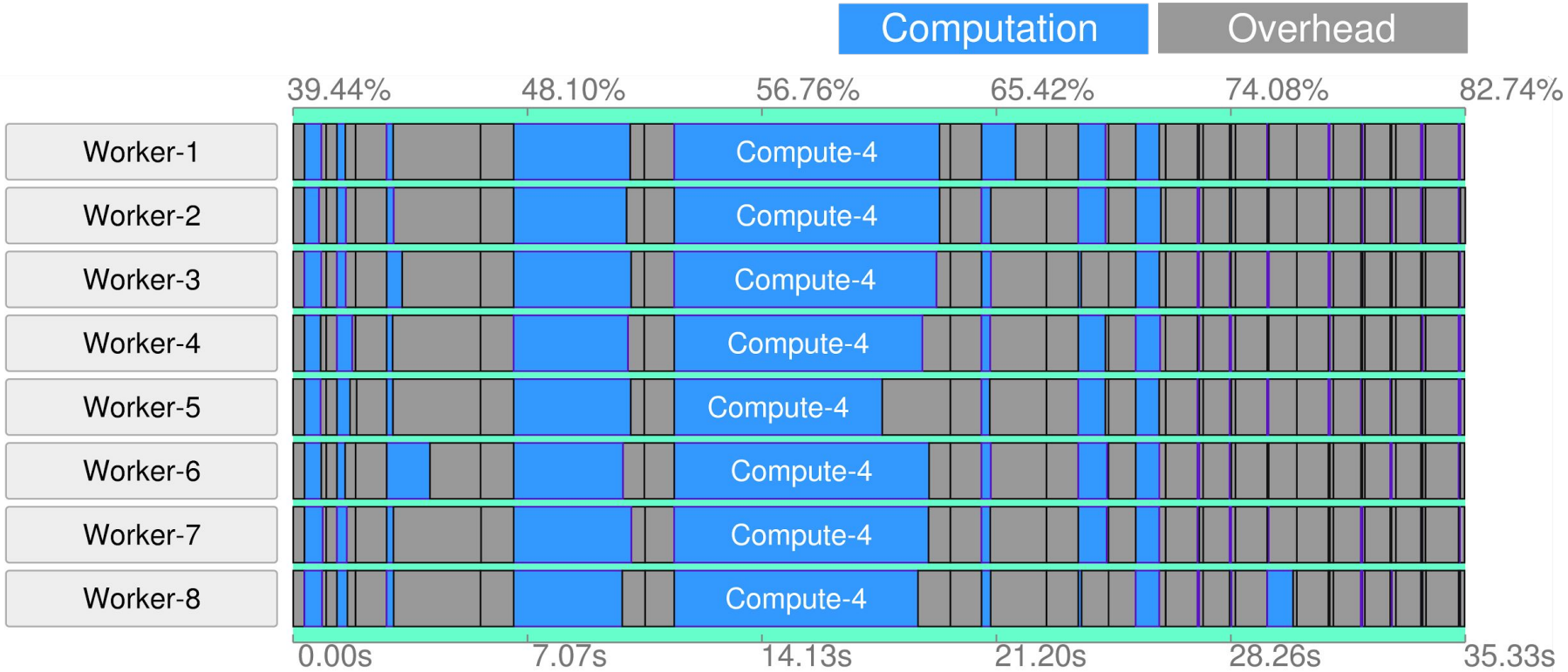
Monitoring Resource Usage

Running BFS on dg1000 (powergraph)



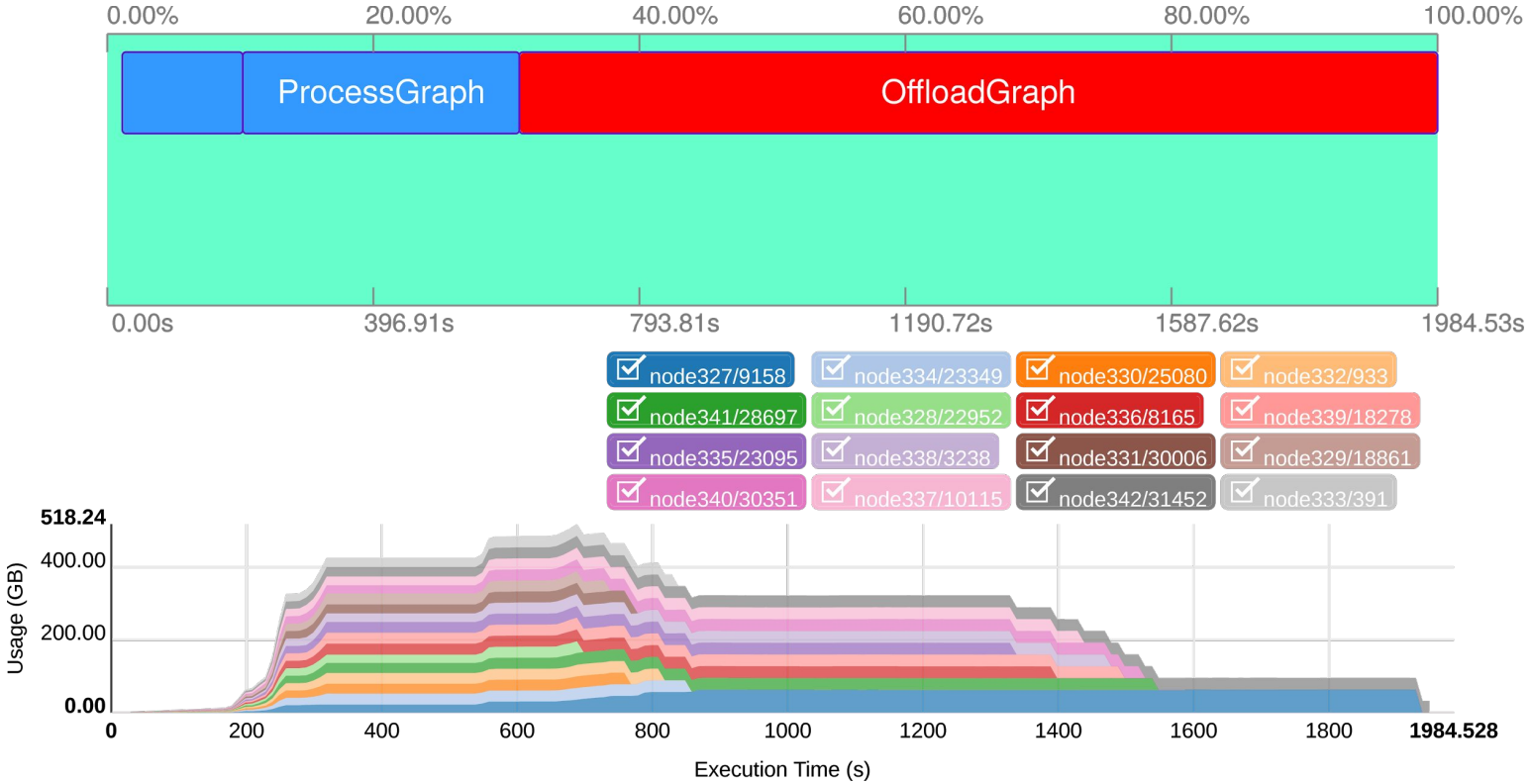
Visualizing System Behavior

Running BFS on dg1000 (giraph)



Diagnosing Failure

Running LCC on dg1000 (powergraph)



Conclusion

We propose **Granula**:

- a fine-grained performance analysis system for Big Data platforms
- that facilitates modeling, monitoring, archiving, and visualization

Future Work:

- continue the development of our research prototype,
- apply our techniques on other types of Big Data platforms,
- better support for analysts on, e.g, failure diagnosis, regression tests,
- integrate performance analysis into standard software engineering practices