Multi-Phase Task-Based HPC Applications: Quickly Learning how to Run Fast

3rd Workshop of the LIG SRCPR Axis (2022)

Lucas Leandro Nesi, Lucas Mello Schnorr, Arnaud Legrand

May 12, 2022

POLARIS - Performance evaluation and Optimization of LARge Infrastructures and Systems
Institute of Informatics, Federal University of Rio Grande do Sul (UFRGS), Brazil
University Grenoble Alpes, CNRS, Inria, Grenoble INP, LIG, France

Heterogeneity in HPC – At System-Level

System-Level Heterogeneity: Hybrid Nodes with Different Computational Power

Santos Dumont

Five Partitions/Systems:

- Base CPU: 504 nodes
- Base Hybrid: 54 nodes
- Base GPU: 198 nodes
- BS CPU: 282 nodes
- BS GPU: 94 nodes

Jean Zay

© Photothèque CNRS/Cyril Frésillon

Four Partitions/Systems:

- CPU Only: 1508 nodes
- CPU + GPU 1: 612 nodes
- CPU + GPU 2: 31 nodes
- CPU + GPU 3: 3 nodes

Heterogeneity in HPC – At System-Level

System-Level Heterogeneity: Hybrid Nodes with Different Computational Power

Santos Dumont

Five Partitions/Systems:

- Base CPU: 504 nodes
- Base Hybrid: 54 nodes
- Base GPU: 198 nodes
- BS CPU: 282 nodes
- BS GPU: 94 nodes

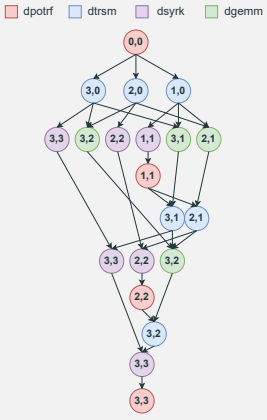
Jean Zay

© Photothèque CNRS/Cyril Frésillon

Four Partitions/Systems:

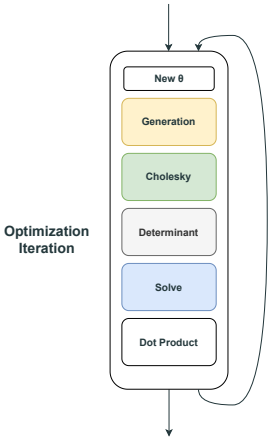
- CPU Only: 1508 nodes
- CPU + GPU 1: 612 nodes
- CPU + GPU 2: 31 nodes
- CPU + GPU 3: 3 nodes

Modern HPC Applications



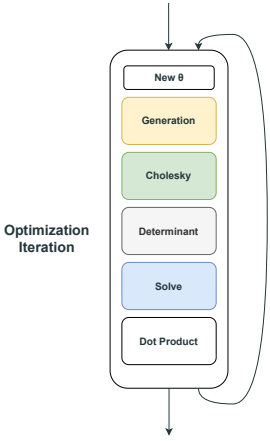
ExaGeoStat - Application Structure

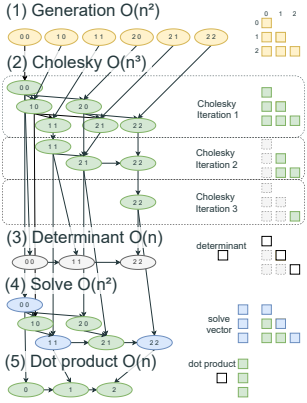
Asynchronous Phases

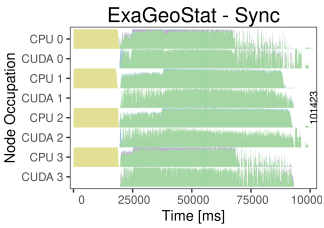


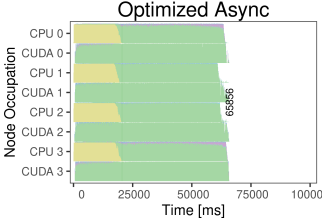
ExaGeoStat - Application Structure

Asynchronous Phases



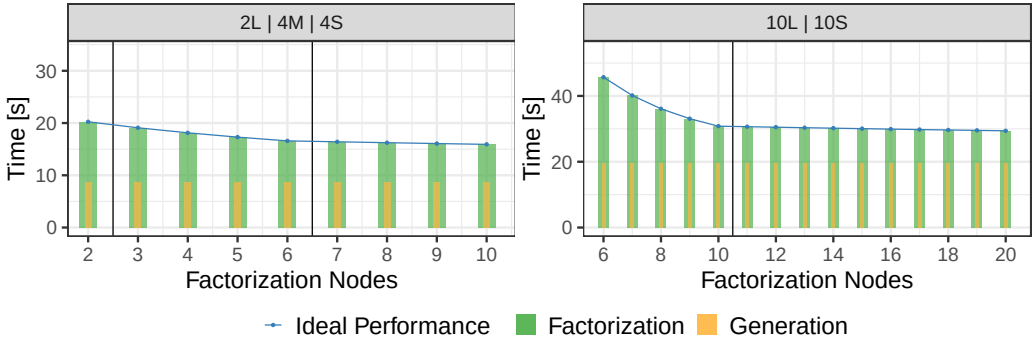






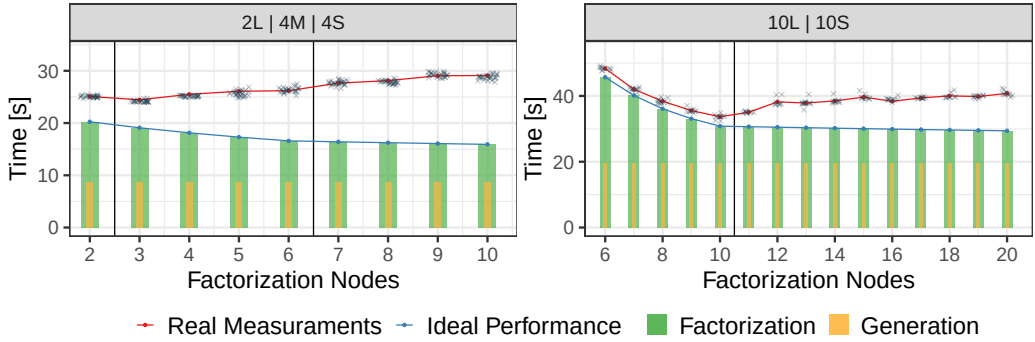
Behavior varying the number of nodes

On a different number of machines:



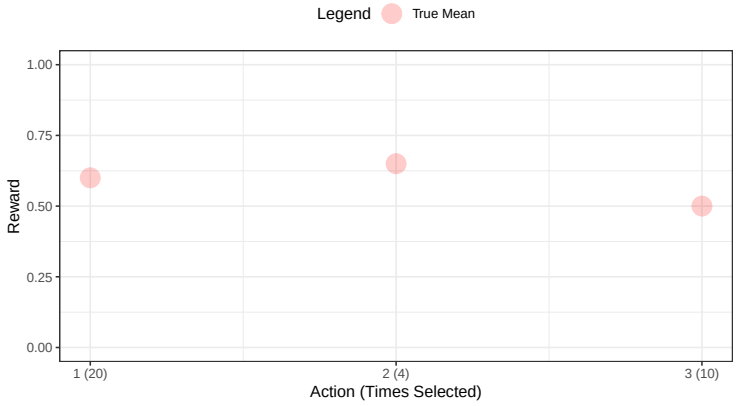
Behavior varying the number of nodes

On a different number of machines:



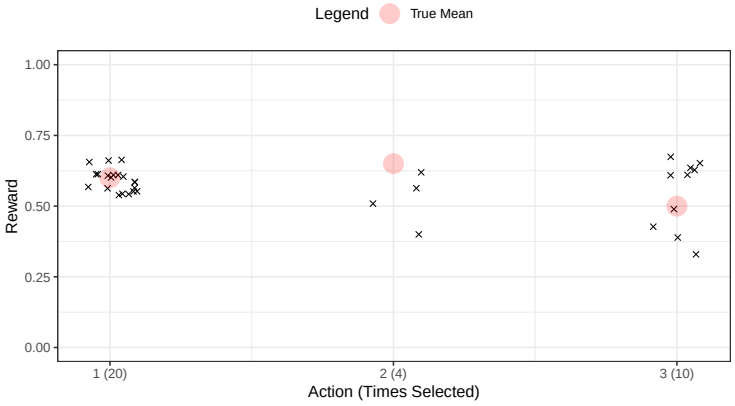
Traditional Multi-armed bandit overview

- Independent actions (arms)



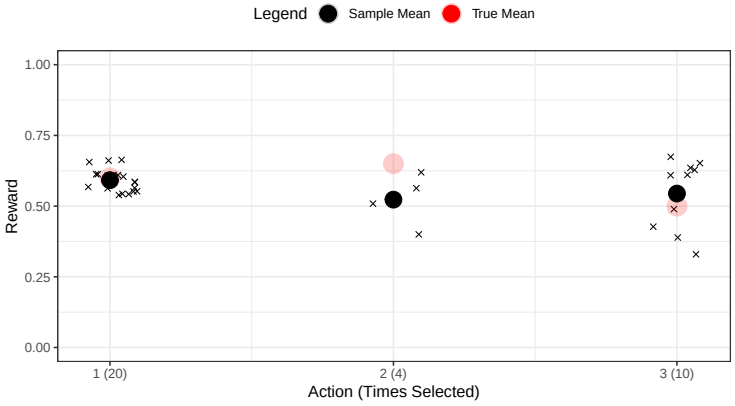
Traditional Multi-armed bandit overview

- Independent actions (arms)



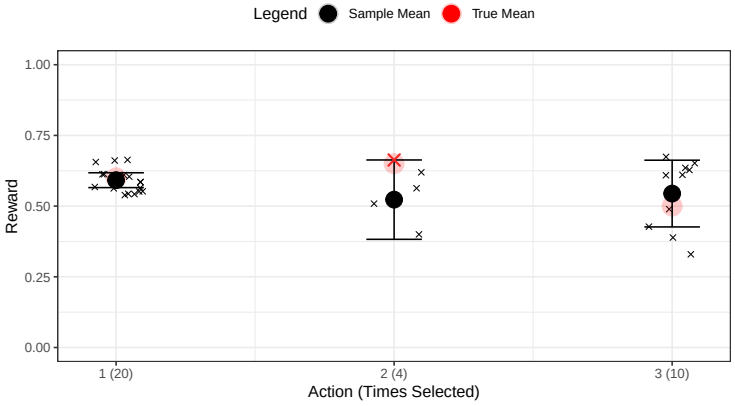
Traditional Multi-armed bandit overview

- Independent actions (arms)



Traditional Multi-armed bandit overview

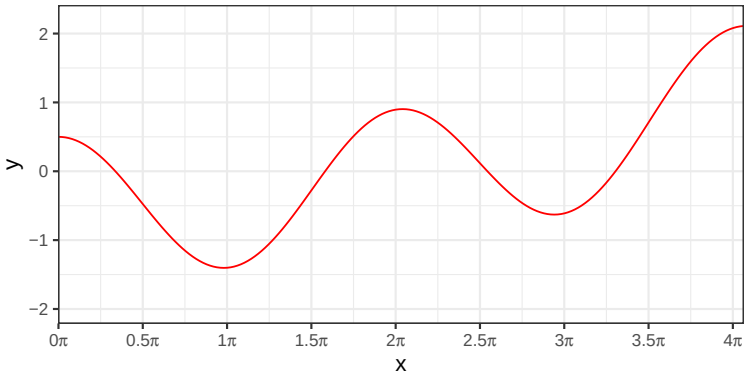
- Independent actions (arms)



- UCB
  - Select the action with the higher reward plus confidence
  - Good regret

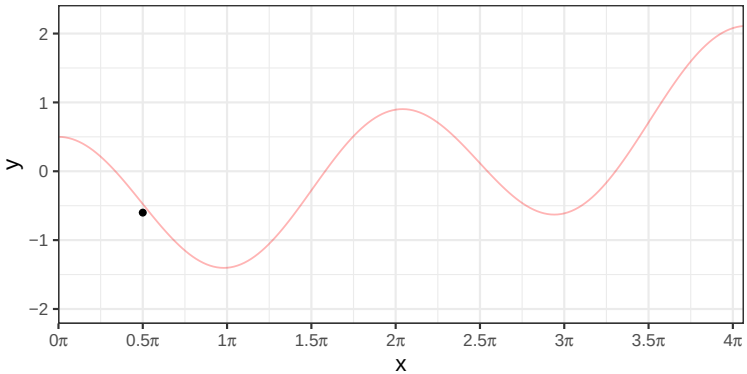
Gaussian Process overview

- Assumes a form of smoothness over the data



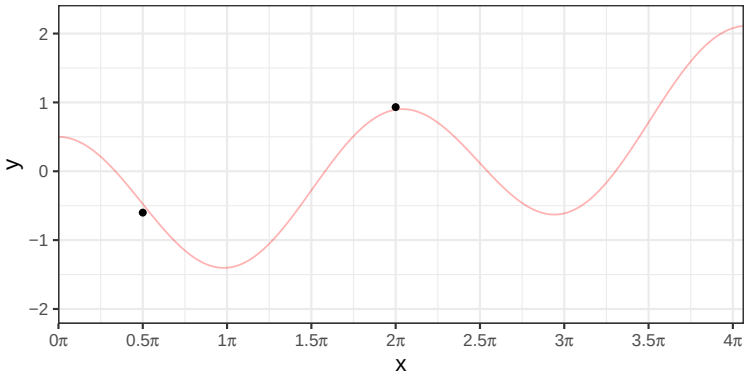
Gaussian Process overview

- Assumes a form of smoothness over the data



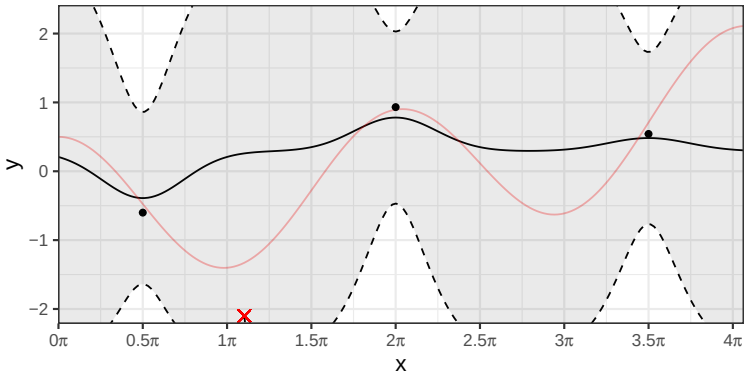
Gaussian Process overview

- Assumes a form of smoothness over the data



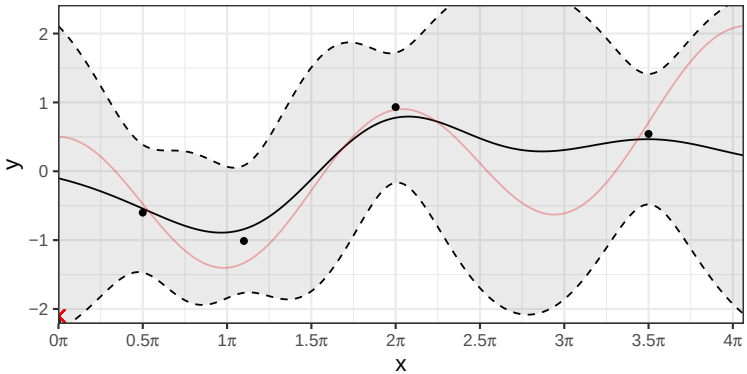
Gaussian Process overview

- Assumes a form of smoothness over the data



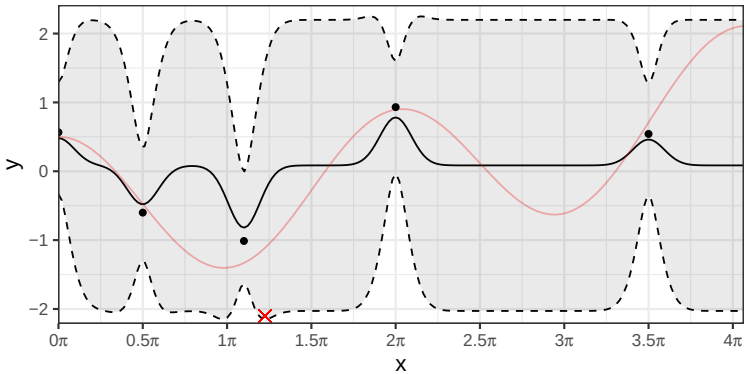
Gaussian Process overview

- Assumes a form of smoothness over the data



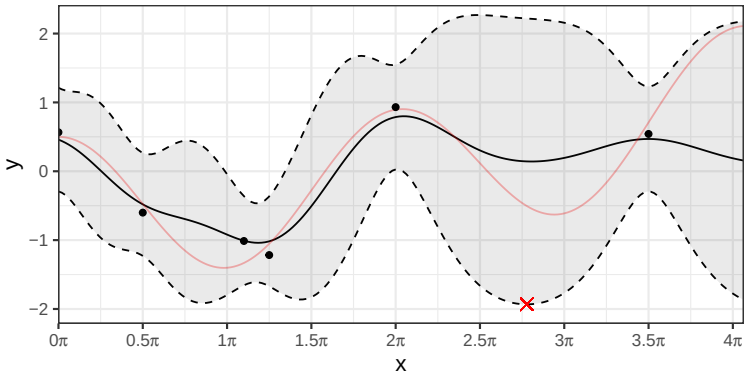
Gaussian Process overview

- Assumes a form of smoothness over the data



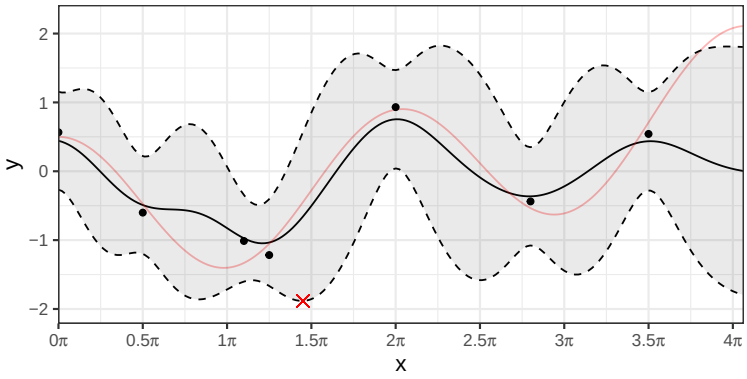
Gaussian Process overview

- Assumes a form of smoothness over the data



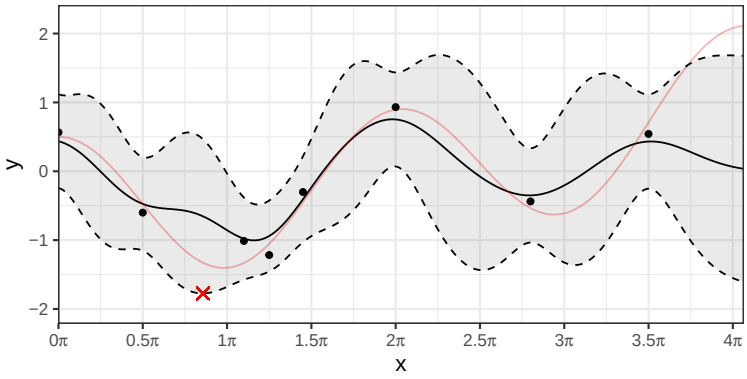
Gaussian Process overview

- Assumes a form of smoothness over the data



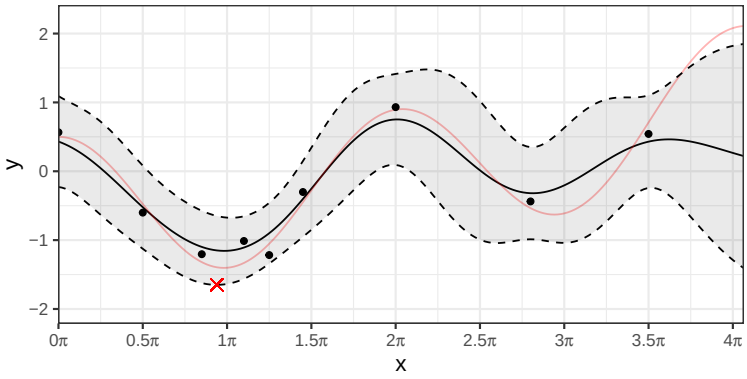
Gaussian Process overview

- Assumes a form of smoothness over the data



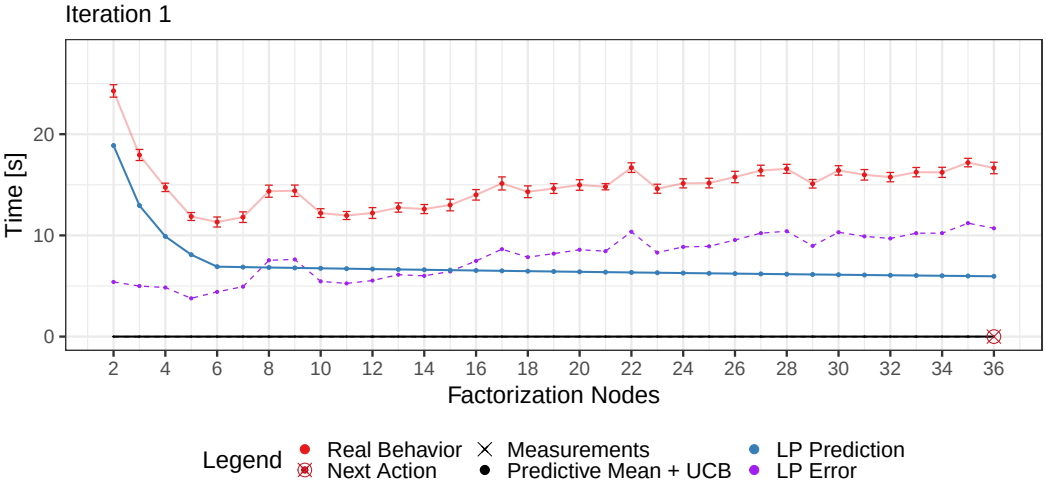
Gaussian Process overview

- Assumes a form of smoothness over the data



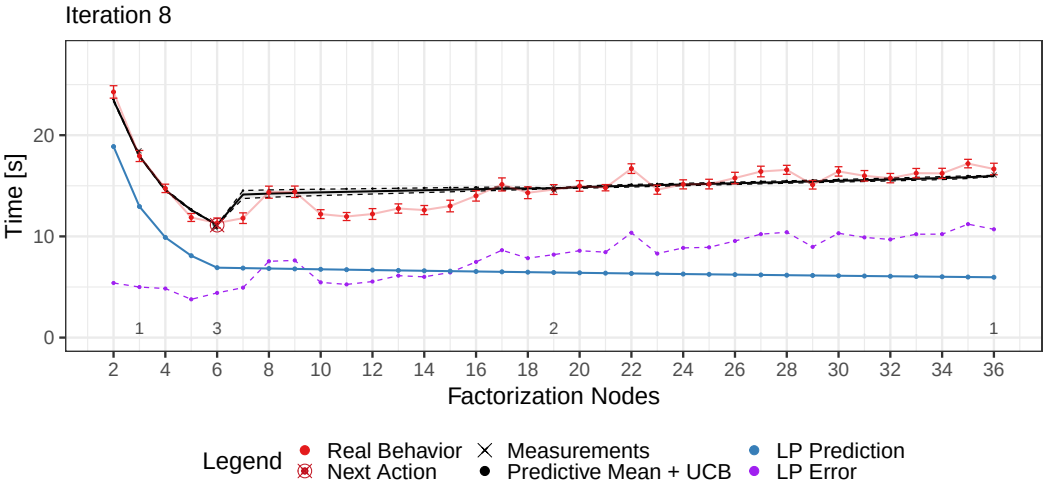
Experiments - GP-Discontinuous Step by Step (6L-30S)

GP-Discontinuous (Model overhead, Search-space limit, Discontinuities) evolution:



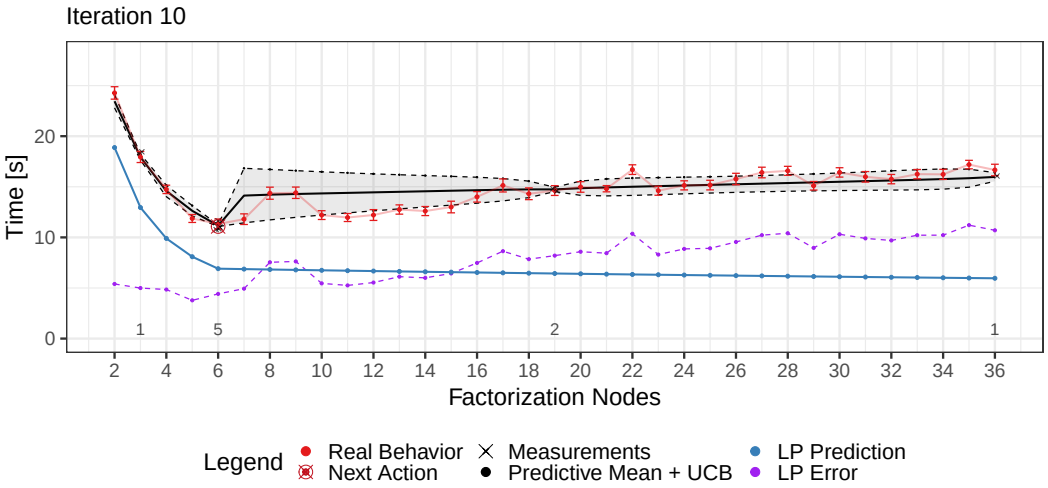
Experiments - GP-Discontinuous Step by Step (6L-30S)

GP-Discontinuous (Model overhead, Search-space limit, Discontinuities) evolution:



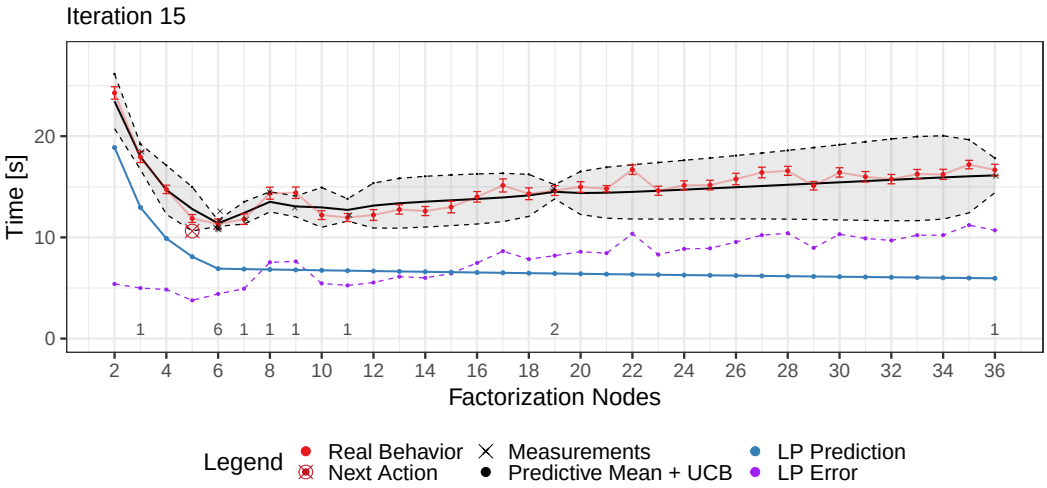
Experiments - GP-Discontinuous Step by Step (6L-30S)

GP-Discontinuous (Model overhead, Search-space limit, Discontinuities) evolution:



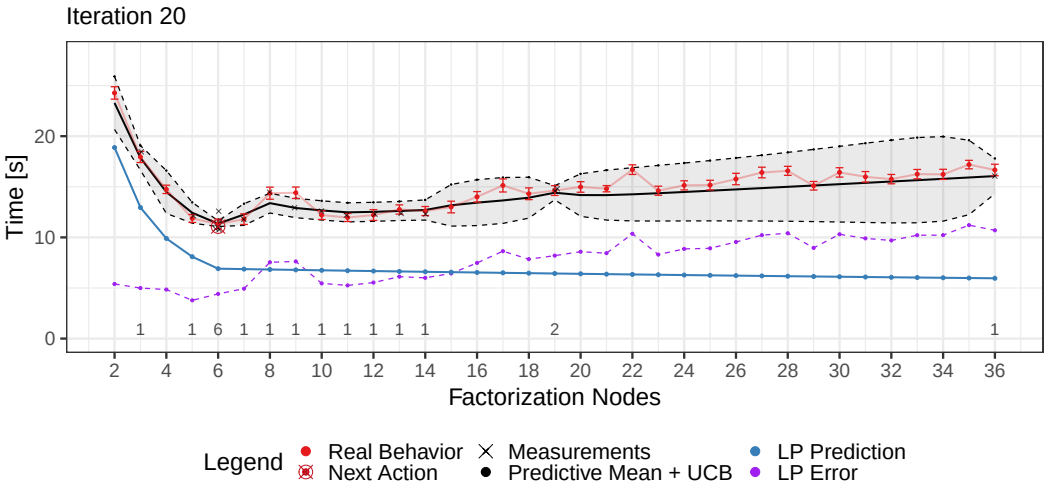
Experiments - GP-Discontinuous Step by Step (6L-30S)

GP-Discontinuous (Model overhead, Search-space limit, Discontinuities) evolution:



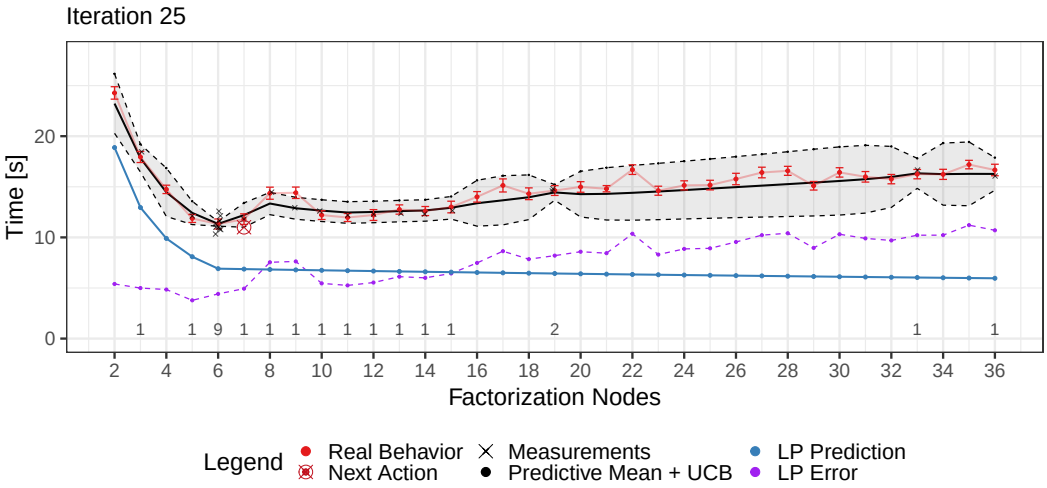
Experiments - GP-Discontinuous Step by Step (6L-30S)

GP-Discontinuous (Model overhead, Search-space limit, Discontinuities) evolution:



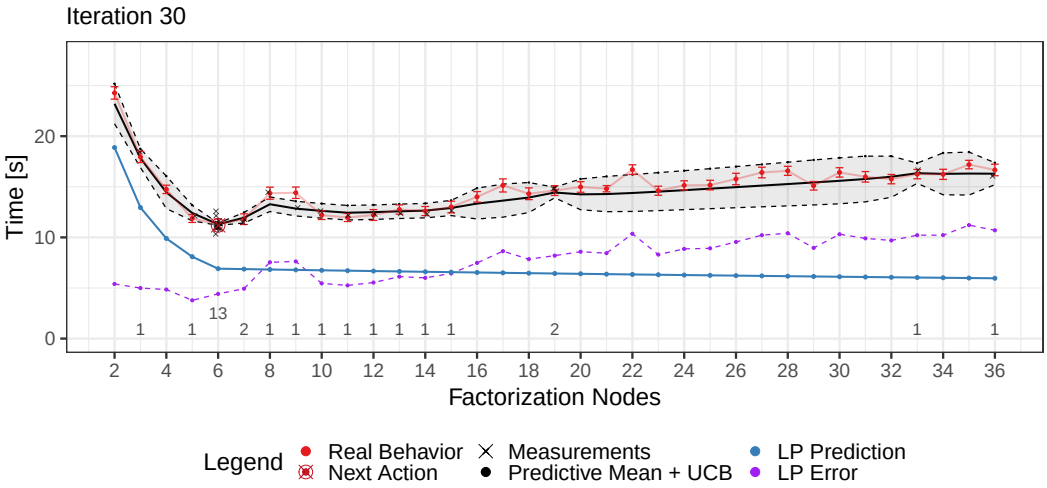
Experiments - GP-Discontinuous Step by Step (6L-30S)

GP-Discontinuous (Model overhead, Search-space limit, Discontinuities) evolution:



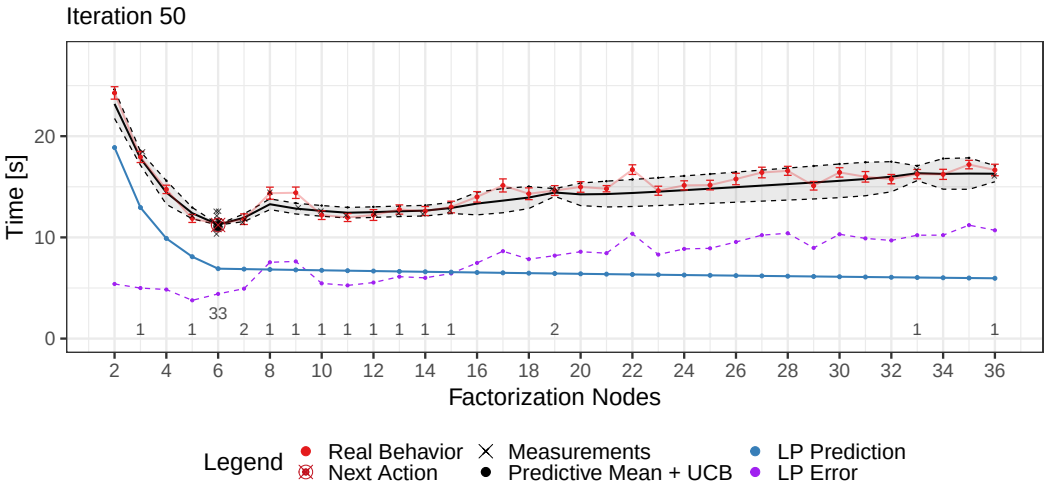
Experiments - GP-Discontinuous Step by Step (6L-30S)

GP-Discontinuous (Model overhead, Search-space limit, Discontinuities) evolution:



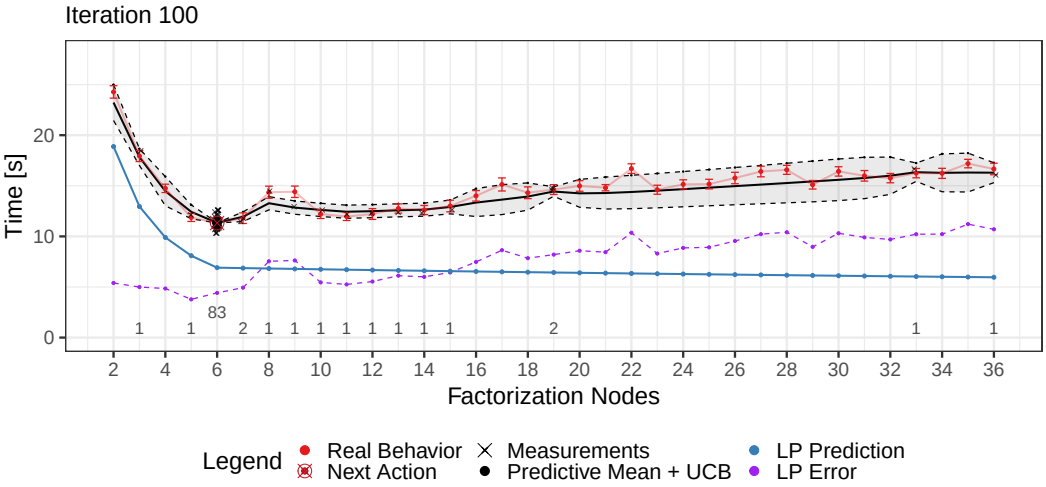
Experiments - GP-Discontinuous Step by Step (6L-30S)

GP-Discontinuous (Model overhead, Search-space limit, Discontinuities) evolution:



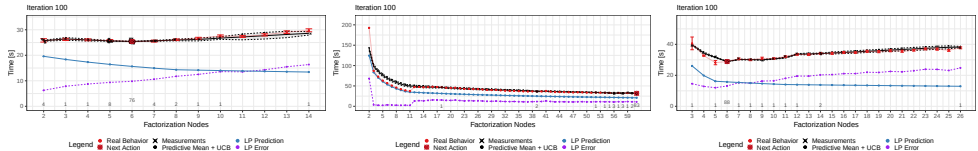
Experiments - GP-Discontinuous Step by Step (6L-30S)

GP-Discontinuous (Model overhead, Search-space limit, Discontinuities) evolution:



Conclusion

- Predicting the ideal number of resources to use is complex
- The application can learn it online and adapt
- We have a collection of setups and comparisons with other algorithms
  - Available online: `https://adaphetnodes.shinyapps.io/shiny/`
- We hope this strategy can be helpful for your problems as well



Acknowledgments

This study was financed in part by the Coordenação de Aperfeiçoamento de Pessoal de Nível Superior - Brasil (CAPES) - Finance Code 001, the National Council for Scientific and Technological Development (CNPq), grant no 141971/2020-7 to the first author, the projects: FAPERGS (Data Science – 19/711-6), CAPES/Cofecub (04/2017) and the INRIA Associated Team ReDaS.

Some experiments were carried out using Grid'5000, supported by a scientific interest group hosted by Inria and including CNRS, RENATER and several Universities as well as other organizations (`https://www.grid5000.fr`). The authors acknowledge the National Laboratory for Scientific Computing (LNCC/MCTI, Brazil) for providing HPC resources of the SDumont supercomputer, which have contributed to the research results reported in this paper (`http://sdumont.lncc.br`).

Thank you!

Thanks for your attention!

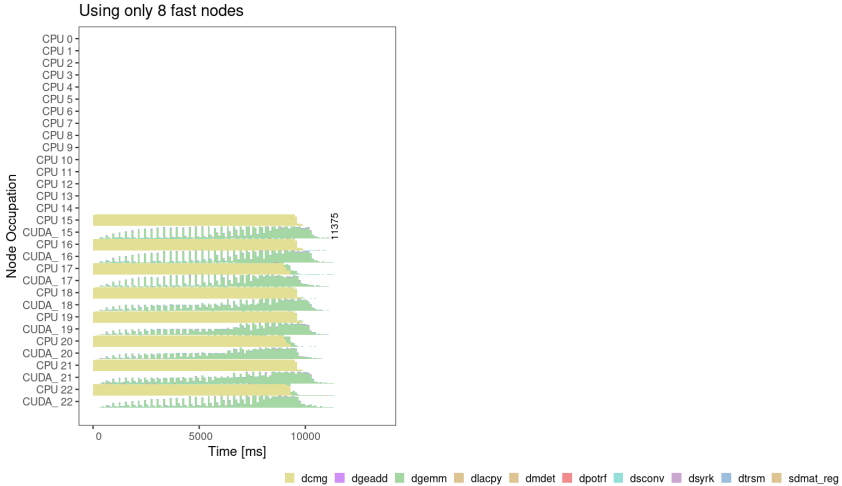
Contact: lucas.nesi@inf.ufrgs.br

Lucas Leandro Nesi, Lucas Mello Schnorr, Arnaud Legrand

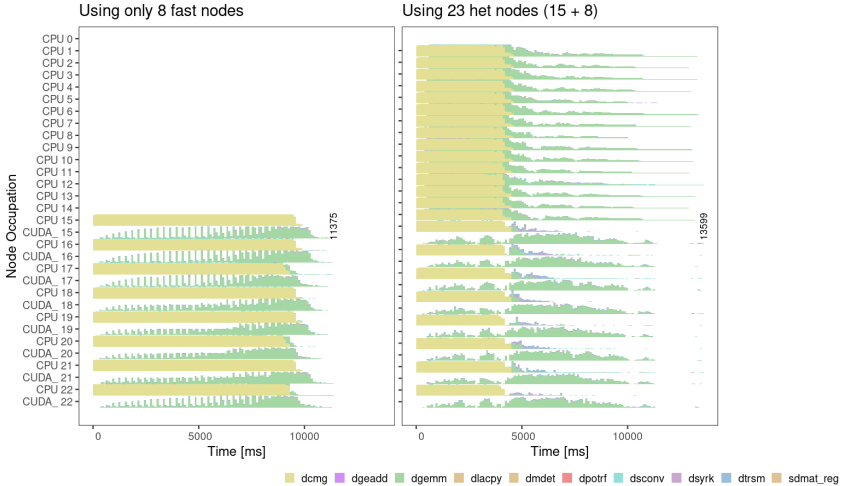
May 12, 2022

POLARIS - Performance evaluation and Optimization of LARge Infrastructures and Systems
Institute of Informatics, Federal University of Rio Grande do Sul (UFRGS), Brazil
University Grenoble Alpes, CNRS, Inria, Grenoble INP, LIG, France

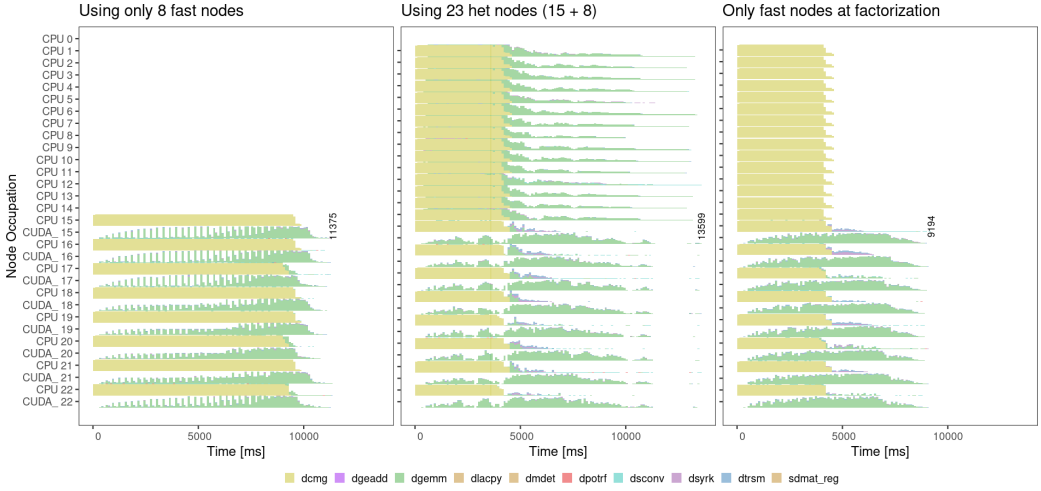
ExaGeoStat - Different number of machines



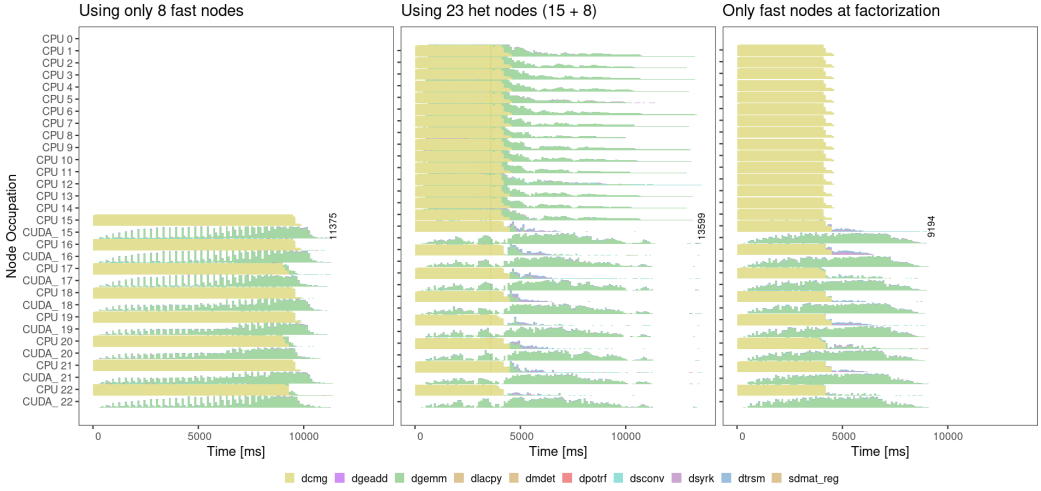
ExaGeoStat - Different number of machines



ExaGeoStat - Different number of machines



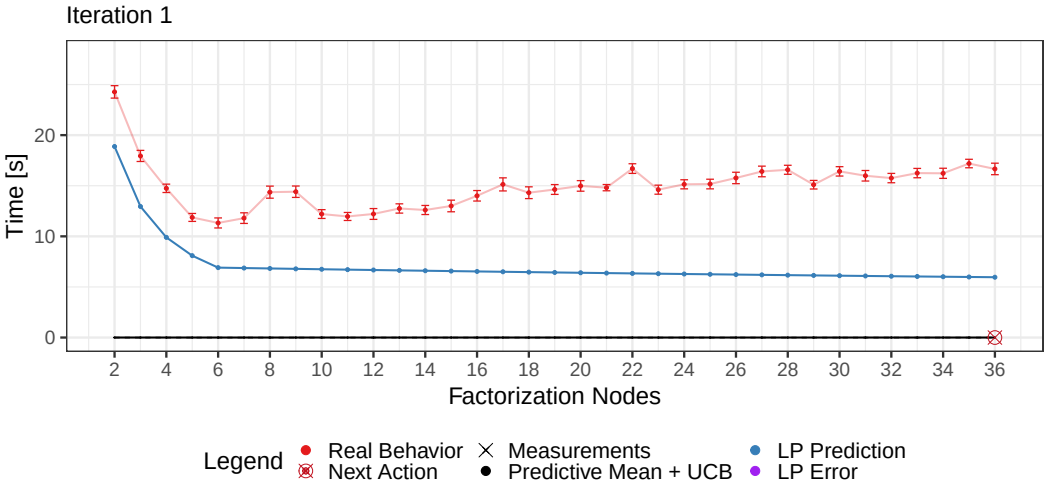
ExaGeoStat - Different number of machines



- How many nodes a application should use?

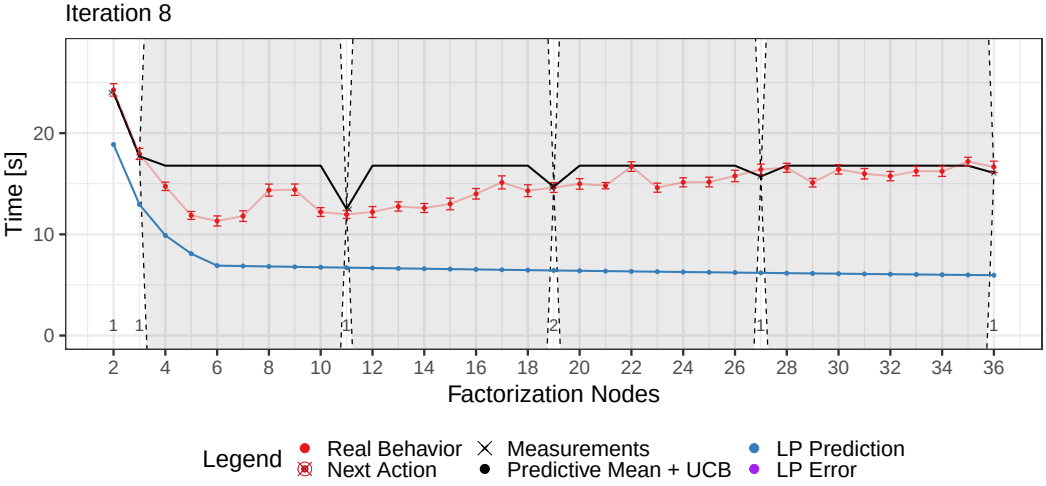
Experiments - GP-UCB Step by Step (6L-30S)

GP-UCB evolution:



Experiments - GP-UCB Step by Step (6L-30S)

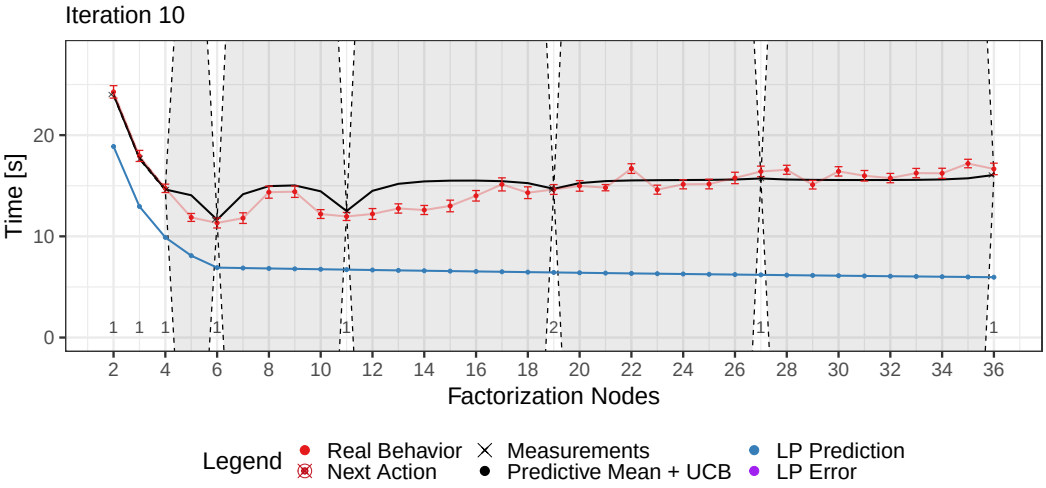
GP-UCB evolution:



- Initial measurements

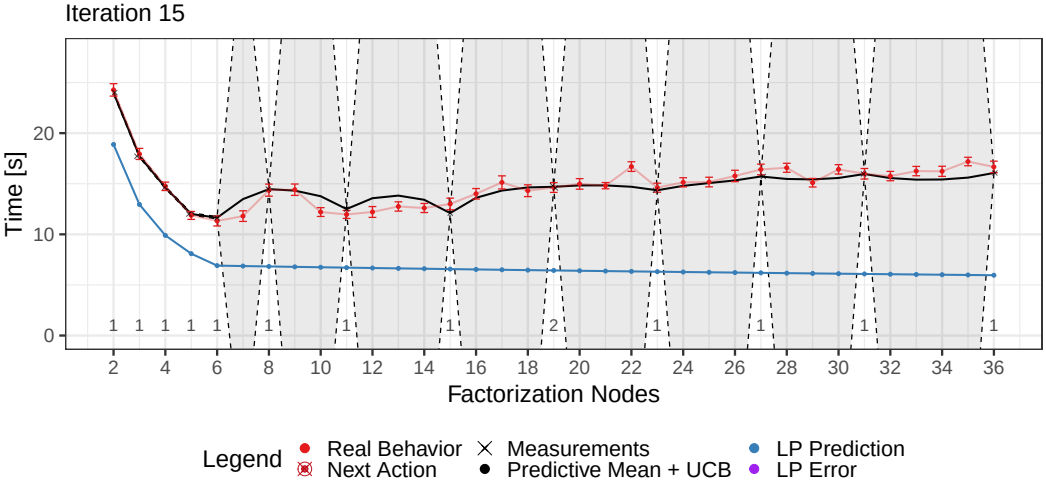
Experiments - GP-UCB Step by Step (6L-30S)

GP-UCB evolution:



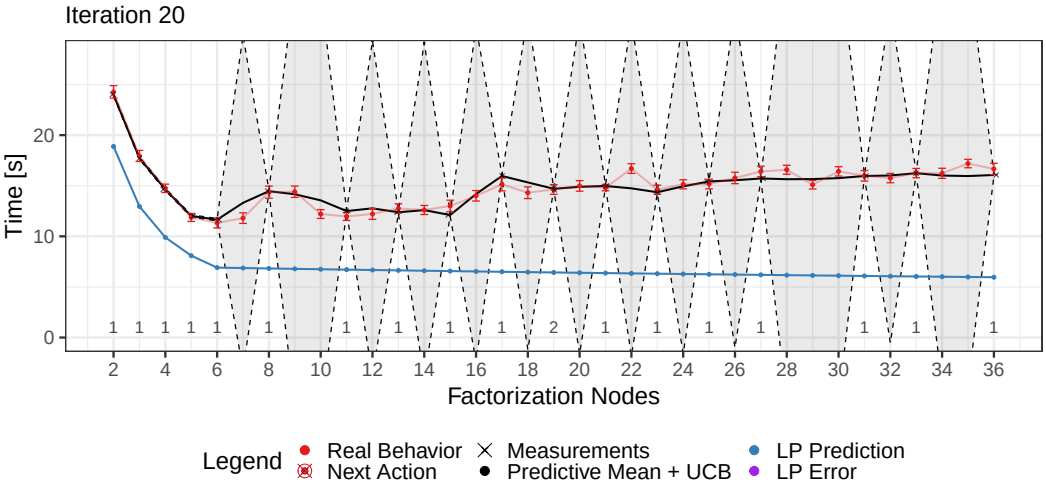
Experiments - GP-UCB Step by Step (6L-30S)

GP-UCB evolution:



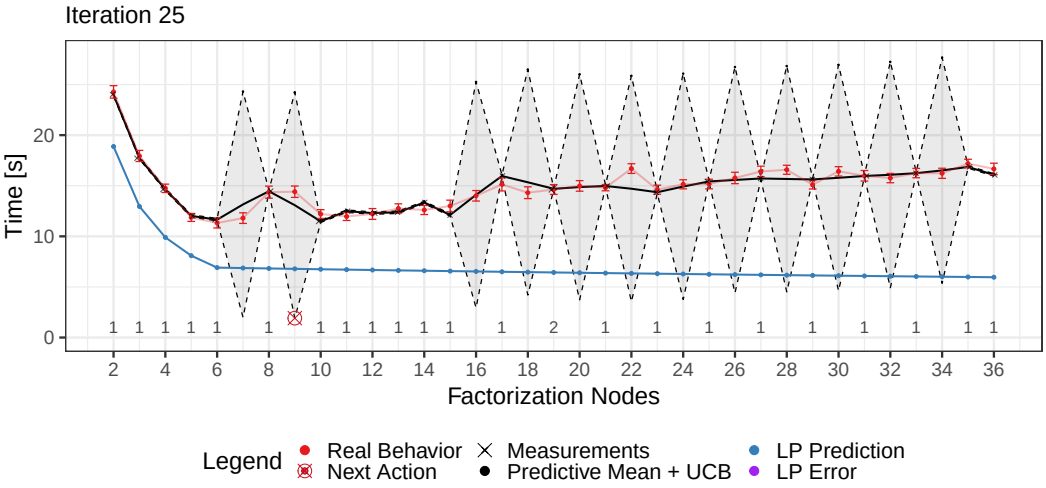
Experiments - GP-UCB Step by Step (6L-30S)

GP-UCB evolution:



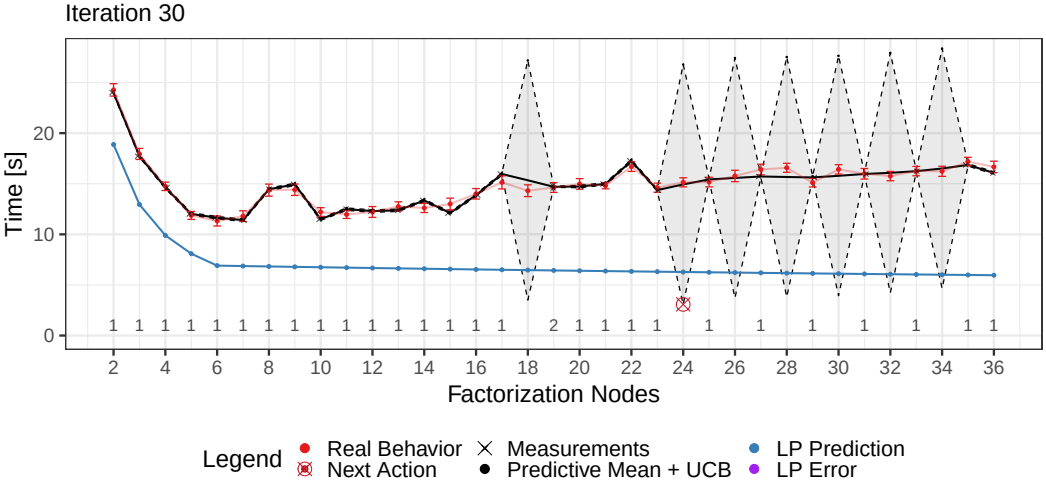
Experiments - GP-UCB Step by Step (6L-30S)

GP-UCB evolution:



Experiments - GP-UCB Step by Step (6L-30S)

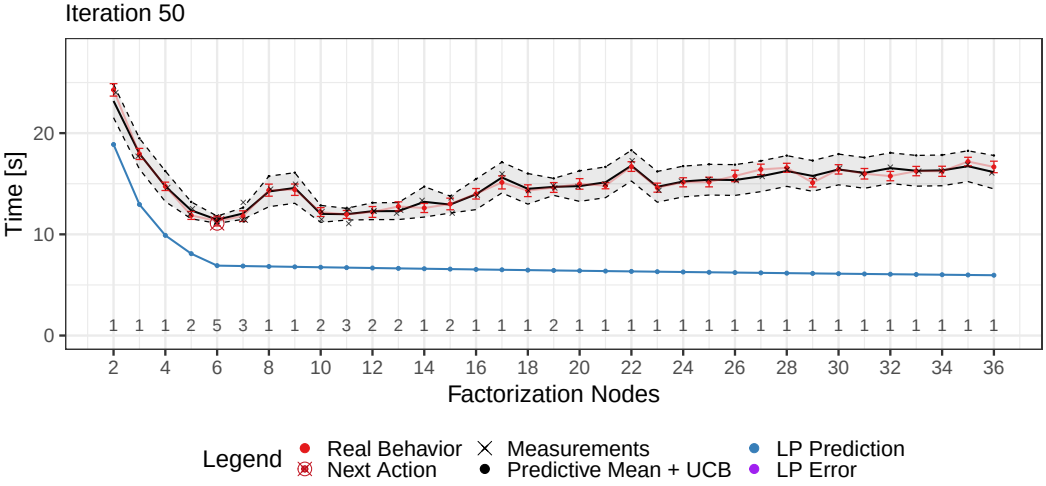
GP-UCB evolution:



- Few repetitions $\Rightarrow$ Poor confidence estimation

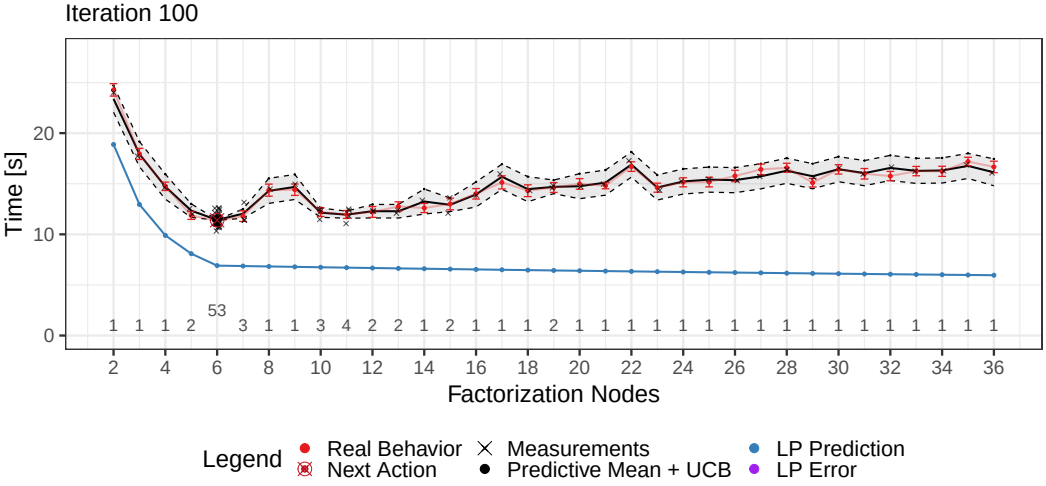
Experiments - GP-UCB Step by Step (6L-30S)

GP-UCB evolution:



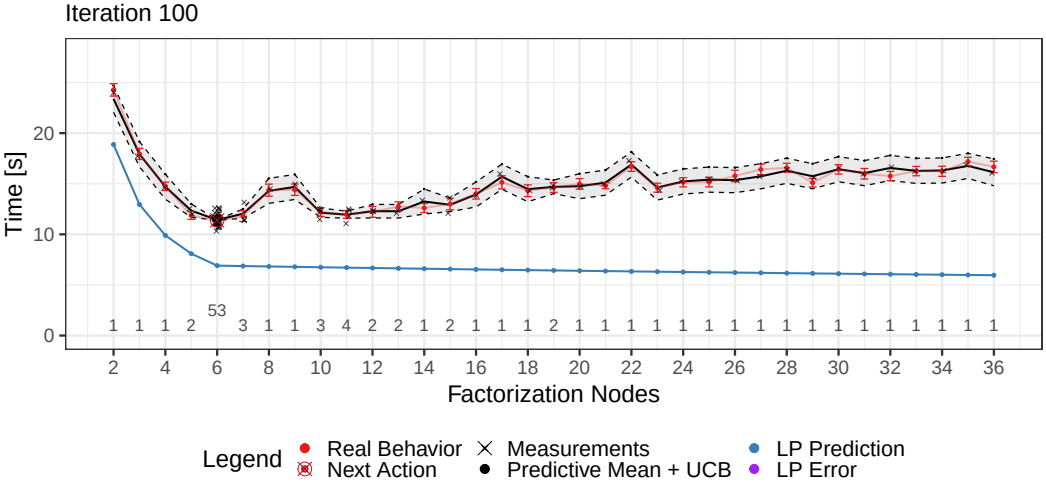
Experiments - GP-UCB Step by Step (6L-30S)

GP-UCB evolution:



Experiments - GP-UCB Step by Step (6L-30S)

GP-UCB evolution:



- Measuring every possibility is bad!