Benchmarking Industrial Computing Environments - Pharmacometrics Study (BICEPS)

Authors: Francois P Combes[1], Matt Fidler[2], David Ochsenbein[3], William Illis[1], Keerty Subrahmanyam[4], Yu-Yun Ho[1]

[1]Novartis Pharmaceuticals Corporation, East Hanover, New Jersey , USA; [2]Novartis Pharmaceuticals Corporation, Fort Worth, Texas, USA; [3]Novartis Pharmaceuticals Corporation, Basel, Switzerland; [4] Novartis Services Inc, East Hanover, New Jersey, USA.

Introduction

The pharmacometrics (PMX) community uses High Performance Computing (HPC) platforms for analyzing clinical data and performing larger-scale computations. An HPC environment allows users to submit jobs to the cluster, enabling parallelization ("splitting" the work across several CPU). PMX-specific software benefits greatly from this parallelization, allowing to significantly reduce computation time, though consuming more immediate resources.

Evaluating and improving the cluster efficiency is crucial to allow a rapid execution of analyses but it is rather difficult to characterize how long it takes for a model or analysis to run. Factors, such as dataset size and model complexity may explain a long run time, regardless of the resources requested

To our knowledge, there is no benchmark set of PMX test cases that quantifies expectations about runtime which could be used to monitor performance over time. We would like to call for volunteers in the BICEPS project to establish an anonymous map of HPC efficiency among industry, institutional and consulting partners.

Objectives

To motivate participation in the BICEPS project, we present here HPC performance results of a proof-of-concept study performed on Novartis HPC.

Methods

- A simultaneous PKPD model of ribociclib PK with absolute neutrophils count (PK-ANC) model with Bayesian estimation (Lu *et al*, 2021) and a PD model of tumor growth inhibition (TGI) (Ribba *et al*. 2012) were used as case example. Further demo examples and results are available in the github platform.
- Two sets of algorithm options were used: default or expert (with increased number of iterations for all estimation steps and without auto-stop criteria), with default seed value
- Each combination of algorithm options, software version and CPUR was run in triplicate, for a total of 504 runs submitted simultenaously
- Several CPU Resources (CPUR, defined by the nodes and threads per nodes requested) were considered. Memory requested was fixed at 500 MB per CPU (1000 MB if only 1 CPU was used), with a maximal run time of 7 days.

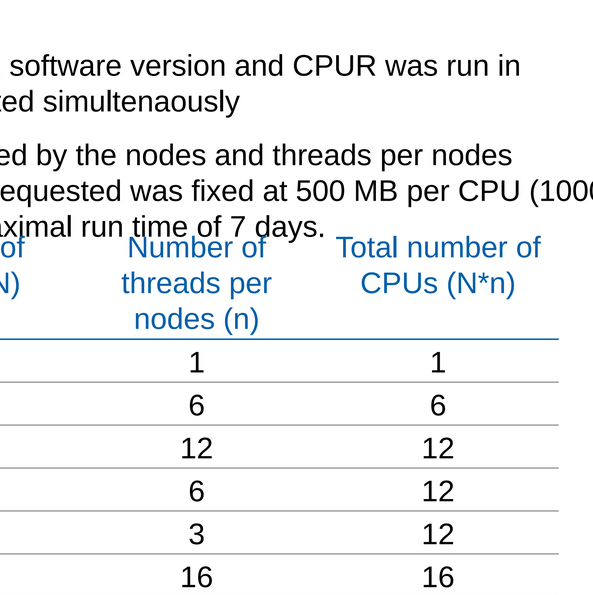
CPU resources (CPUR)	Number of nodes (N)	Number of threads per nodes (n)	Total number of CPUs (N*n)
Minimal resources	1	1	1
Default resources	1	6	6
High resource (MPI 1)	1	12	12
High resources (MPI 2)	2	6	12
High resources (MPI 4)	4	3	12
Max resources (MPI 1)	1	16	16
Max resource (MPI 4)	4	12	48

Results

Parameter estimation

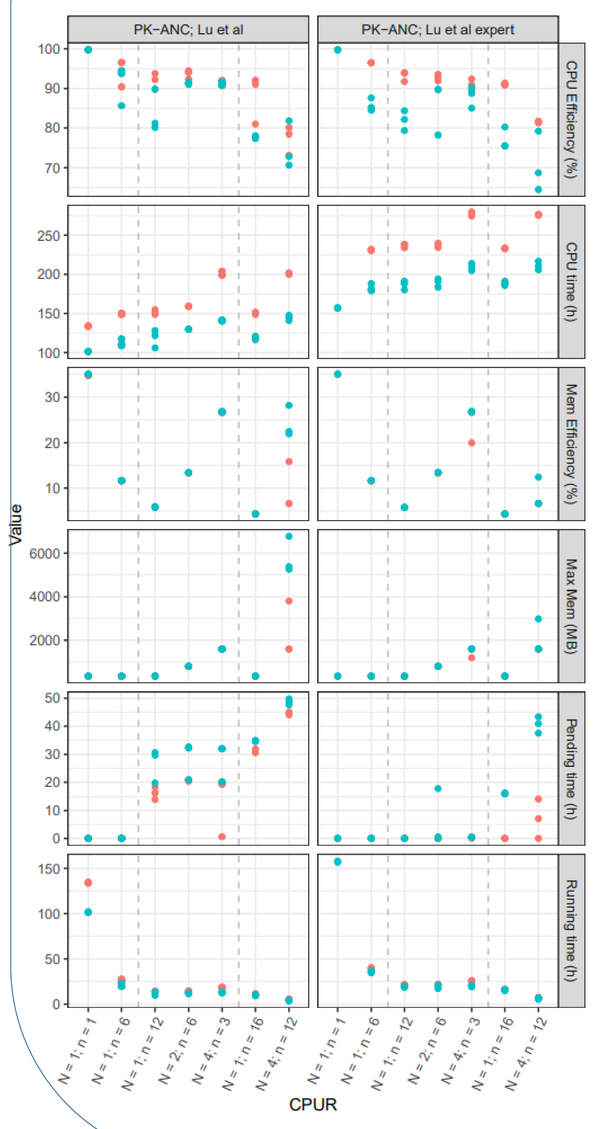
- Successful estimation except for Lu *et al***.** model in Monolix 2021 with minimal resources (reached time limits)
- Overall, similar estimates and RSE within the same version for TGI models
- Mean difference in estimates/initial values and RSE were consistent across models and versions

Figure 1. Comparison of parameter estimates and RSEs



Cluster performance (PK-ANC model)

Figure 2. BICEPS Benchmarking results for PK-ANC model



- Decreased CPU efficiency from 100% (N1-n1) to 70%(N4-n12), with increased CPUR
- Slight variability in CPU efficiency between replicates
- CPU time slightly increased with CPUR
- Memory request was overestimated: memory efficiency always below 40%
- Increase in maximal memory used with increase in CPUR
- A setting of N4-n12 led to an overall shorter runtime, though pending time was higher
- Increase in CPUR lead to decrease in runtime from about 100h to 10h
- For scenarios using 12 threads (within vertical dashed lines), parallelization led to slightly higher runtimes and lower CPU efficiency

Conclusions

- Benchmark performed on Novartis HPC helps users to tailor job requirements to model complexity and dataset size, allowing mitigation of the total shared resources vs the importance and urgency of activities
- This project will benefit from inclusion of results from other partners to create the benchmarking distribution which will help the organizations to further maintain, improve or replace HPC assets
- Further developments are needed to improve this analysis:
  - Inclusion nonmem and nlmixr2,
  - Adjustment of the programs to partner organizations both for run automatization and results collection from log files
  - Evaluation of the interaction between type of structural model, dataset size and MPI performances
- We invite the audience to register for the BICEPS study and participate in the code development

References: Lu Y, *et al*. J Clin Pharmacol. 2021 Aug;61(8):1054-1068; Ribba B *et al*. Clinical Cancer Research, 18(18), 5071-5080, 2012; Monolix 2021R2, Lixoft SAS, a Simulations Plus company; R Core Team (2021). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. URL https://www.R-project.org/.

Acknowledgments: The authors would like to thanks Pr S Schmidt and Dr S Kim from the Center of Pharmacometrics and Systems Pharmacology, University of Florida for their involvement in the project, along with Dr Jonathan Chauvin and the Simulation Plus support team for their help and suggestions

This study was sponsored by Novartis Pharmaceutical corporation

Scan this QR code to access the github repository

Materials (R codes and monolix models necessary to reproduce this analysis are available publicly on github
https://github.com/combefr1/BICEPS

Scan this QR code to email the authors

francois_pierre.combes@novartis.com