Productivity and Software Development Effort Estimation in HPC

Sandra Wienke

IT Center & Chair for High Performance Computing, RWTH Aachen University, Germany

Introduction

Serving the ever increasing demands for computational power, expenses of HPC centers increase in terms of acquisition, energy, and programming. Thus, a quantifiable productivity metric is needed in HPC procurements to make an informed decision on how to invest available budgets.

$$\text{productivity} = \frac{\text{value}}{\text{cost}\,[\$]}$$

$\$$ = HW + energy + development cost + …

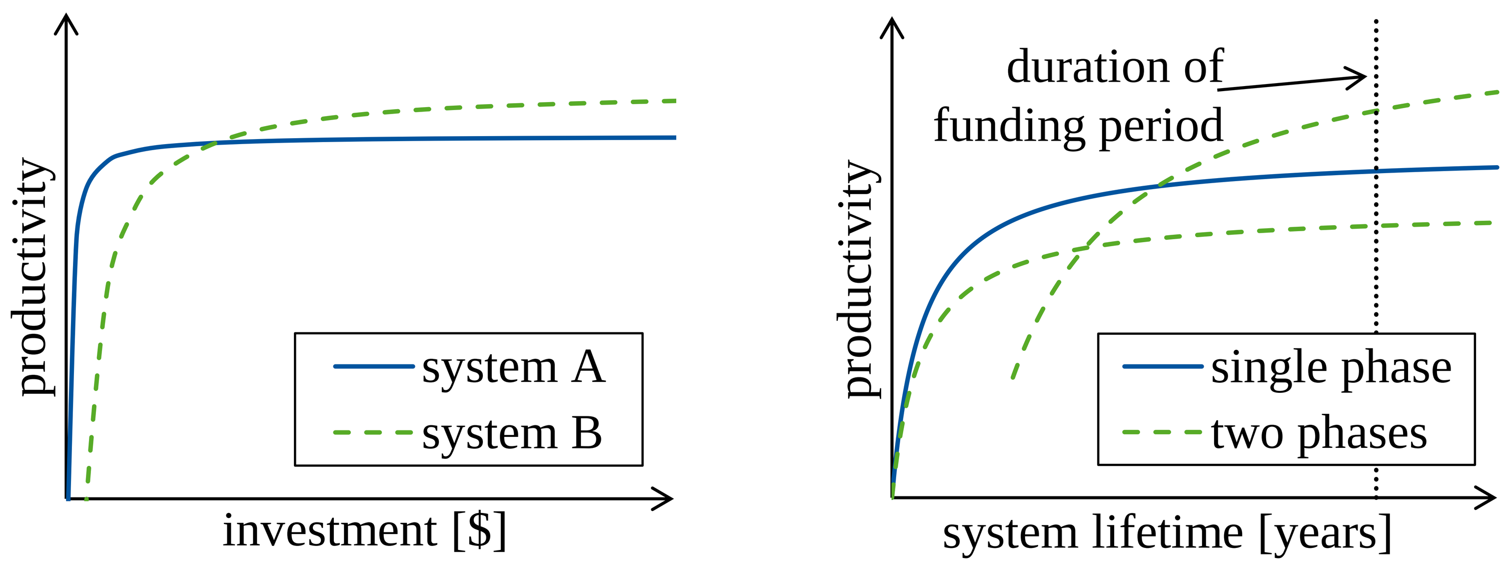
dev. effort [day] * salary $\left[\frac{\$}{\text{day}}\right]$

Challenges

- Real-world applicability w/ multi-job setups
- Quantification of value
- Prediction of all productivity parameters
- Especially, estimation of development effort
- Intangible character of various impact factors
- Data collection through human-subject research

Productivity

I introduce a productivity model with predictive power to foster informed decision making in HPC procurements.



For instance, it can be used to compare various systems, or to argue for single- or two-phase procurements.

$$\text{productivity} \quad \psi(n,\tau) = \frac{\sum r_{app,i}(n,\tau)}{\text{TCO}(n,\tau)}$$

n: system size [#nodes]
τ: system lifetime [years]

Value: Number of Application Runs

- Number of runs of (relevant) simulation applications
- Focus on application runtime $t_{app,i}$ for application i running on n_i compute nodes
- Overarching metric for real-world job mix setups

$$r_{app,i}(n,\tau) = p_i \cdot \frac{q_{app,i}(n_i)\cdot\alpha\cdot n\cdot\tau}{n_i\cdot t_{app,i}(n_i)}$$

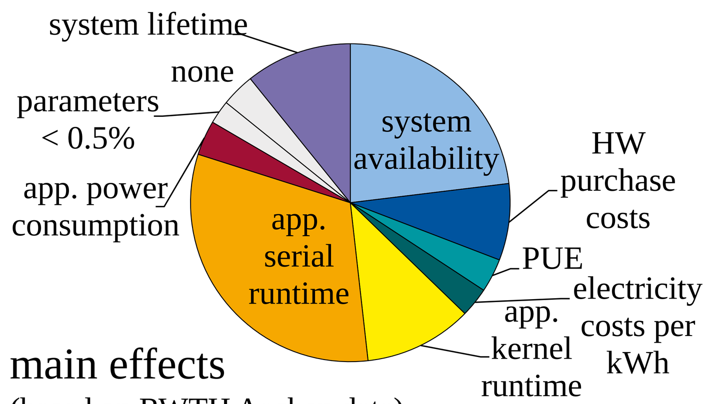
α: system availability
$q_{app,i}$: quality weighting factor
p_i: capacity weighting factor

Cost: Total Cost of Ownership (TCO)

- One-time costs C^{ot} and annual costs C^{pa}
- Categorization: costs per node and per node type
- Inclusion of costs for, e.g., programming, energy, hardware

$$\text{TCO}(n,\tau) = C^{ot}(n) + C^{pa}(n)\cdot\tau$$

Sensitivity Analysis



main effects
(based on RWTH Aachen data)

- Variances in productivity due to errors in assumptions
- Only few (well-understood) parameters must be accurately predicted → robust

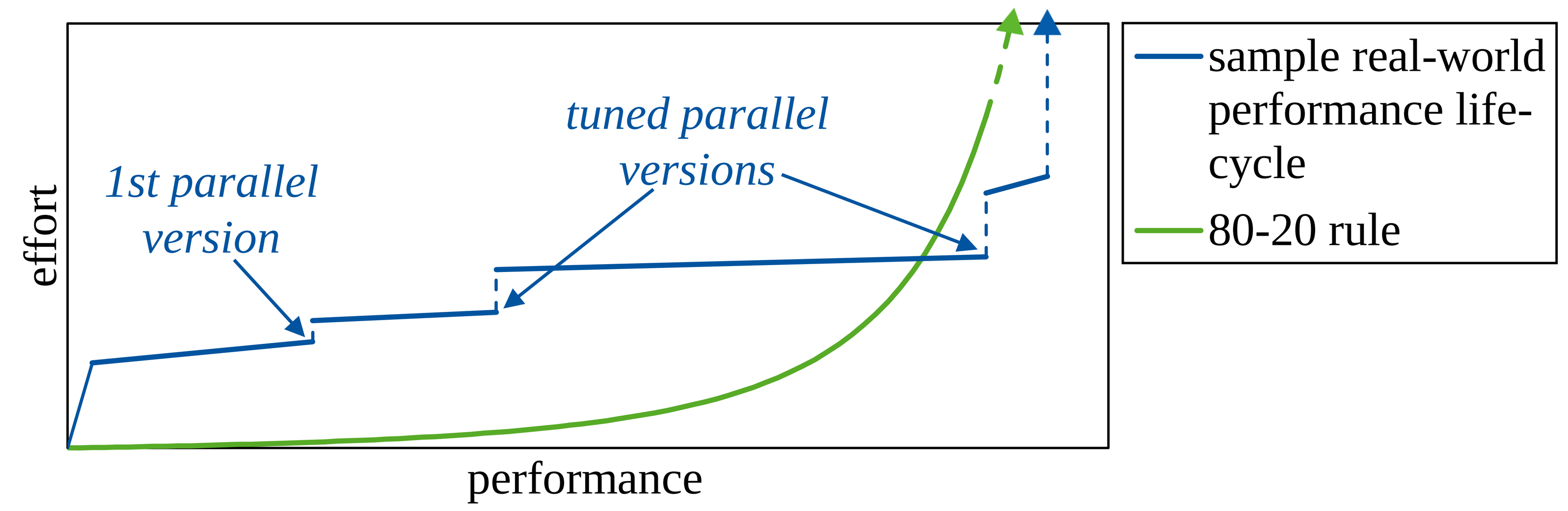
Development Effort

As part of a sound TCO and productivity model, I introduce a methodology to estimate development efforts needed to parallelize, port and tune simulation codes to efficiently exploit (novel) HPC hardware.

Performance Life-Cycle

- Model of relationship of performance and effort needed to achieve this performance

$$\text{effort} = S\cdot f(\text{performance})^R + Q$$



- Numerous factors impact effort (captured in S, R, Q)
- Regression analysis from collected data

Impact Factors on Effort

more impact ↑ … ↓ less impact

- Pre-knowledge on HW & parallel prog. model
- Code work
- Pre-knowledge on numerical algorithm used
- Parallel prog. model & compiler/ runtime system
- Performance
- Architecture/ hardware
- Tools
- Kind of algorithm
- Code size
- Portability & maintainability over code's lifetime
- Energy efficiency

Identification of key drivers

- Focus on most influencing factors
- Factor ranking based on surveys

Quantification of "pre-knowledge"

- Knowledge surveys (confidence rating)

Quantification of "parallel programming model & numerical algorithm"

- Pattern-based suitability mapping

Data Collection

- Basis of statistically reliable results are sufficient data sets
- Targeting at community effort: tools and material online, e.g., effort-performance pairs by our tool *EffortLog*

Conclusion

My research covers a productivity figure of merit for HPC procurements that is applicable to real-world multi-job setups. To embrace an estimation of software development effort into the productivity model, I provide a methodology that is based on performance life-cycles and statistical analysis of data collected in human-subject research.

In future, I will continue to collect corresponding data sets and investigate conditional refinements of the productivity model.

References

S. Wienke, D. an Mey, and M. S. Müller, "Accelerators for Technical Computing: Is It Worth the Pain? A TCO Perspective," in Supercomputing, ser. Lecture Notes in Computer Science, J. M. Kunkel, T. Ludwig, and H. W. Meuer, Eds. Springer Berlin Heidelberg, 2013, vol. 7905, pp. 330–342.

S. Wienke, H. Iliev, D. an Mey, and M. S. Müller, "Modeling the Productivity of HPC Systems on a Computing Center Scale," in High Performance Computing, ser. Lecture Notes in Computer Science, J. M. Kunkel and T. Ludwig, Eds. Springer International Publishing, 2015, vol. 9137, pp. 358–375.

S. Wienke, J. Miller, M. Schulz, and M. S. Müller, "Development Effort Estimation in HPC," in SC16: International Conference for High Performance Computing, Networking, Storage and Analysis, 2016, pp. 107–118.

S. Wienke, T. Cramer, M. S. Müller, and M. Schulz, "Quantifying Productivity-Towards Development Effort Estimation in HPC," Poster at the International Conference for High Performance Computing, Networking, Storage and Analysis (SC15), 2015.

S. Wienke, "Development Effort Methodologies," Chair for High Performance Computing, RWTH Aachen University, http://www.hpc.rwth-aachen.de/research/tco, 2016.