

## COLLECTIVE KNOWLEDGE: A FRAMEWORK FOR SYSTEMATIC PERFORMANCE ANALYSIS AND OPTIMIZATION

We present the outcome of a technology transfer project between the non-profit cTuning Foundation, France (Grigori Fursin) and ARM, UK (Anton Lokhmotov, Ed Plowman). The six-month project supported by the TETRACOM Coordination Action has resulted in developing from scratch the Collective Knowledge framework and validating it on realistic use cases, as well as forming a startup called dividiti.



Designing, modeling and benchmarking of computer systems in terms of performance, power consumption, size, reliability and other characteristics is becoming extraordinarily complex and costly. This is due to a large and continuously growing number of available design and optimization choices, a lack of common performance analysis and optimization methodologies, and a lack of common ways to create, preserve and reuse vast design and optimization knowledge. As a result, optimal characteristics are achieved only for a few ad-hoc benchmarks, while often leaving real-world applications underperforming. Eventually, these problems lead to a dramatic increase in the development, optimization and maintenance costs, increasing time to market for new products, eroding return on investment (ROI), and slowing down innovation in computer engineering.

Since 2012 the non-profit cTuning Foundation and ARM have been engaging in discussions on systematic performance analysis and optimization using statistical analysis, machine learning and crowd-tuning techniques. In November 2014, we started a six-month technology transfer project supported by the FP7 TETRACOM Coordination Action. The cTuning technology comprises a framework and repositories for collaborative and reproducible experimentation combined with predictive

analytics. This technology has been successfully used in several EU projects including the FP6 MILEPOST project.

The TETRACOM grant has allowed us to completely design and develop from scratch the fourth version of cTuning technology which we called Collective Knowledge (or CK for short). CK is a Python-based framework, repository and web service, supporting JSON interfaces and standard Git services such as GitHub and Bitbucket. CK allows engineers and researchers to organize, describe, cross-reference and share their code, data, experimental setups and meta information as unified and reusable components. CK users can assemble from components various experimental workflows, quickly prototype ideas, crowdsource experiments using spare computer resources such as mobile phones, and more. Importantly, CK allows experimental results to be exposed to powerful predictive analytics packages such as scikit-learn and R in order to speed up decision making via statistical analysis, data mining and machine learning.

During the project, we have successfully applied the Collective Knowledge framework to perform systematic analysis, data mining and online/offline learning on vast amounts of benchmarking data available at ARM. Our technology showed good potential to automatically find various important correlations between numerous in-house benchmarks, data sets, hardware, performance, energy and run-time state. Such correlations can, in turn, help derive

representative benchmarks and data sets, quickly detect unexpected behavior, suggest how to improve architectures and compilers, and speed up machine-learning based multi-objective autotuning. Furthermore, our technology has also showed potential to enable collaborative research and development within and across groups. Therefore, we have released the Collective Knowledge framework under a permissive BSD license, and expect to grow the user community.

Finally, our positive results have motivated us to establish a UK-based startup called dividiti to accelerate computer engineering and research by further developing our technology and applying it to real-world problems.

**Acknowledgments:** We would like to thank the TETRACOM Steering Committee for accepting the project proposal and the TETRACOM manager Eva Haas for simplifying the paperwork. We also would like to thank our ARM colleagues Marco Cornero, Alexis Mather and Jem Davies for encouraging and supporting the project.

Further resources:

- <http://tetracom.eu>
- <http://ctuning.org>
- <http://github.com/ctuning/ck>
- <http://hal.inria.fr/hal-01054763>
- <http://www.dividiti.com>

*Grigori Fursin, cTuning Foundation*