CerCo

INFORMATION AND COMMUNICATION
TECHNOLOGIES
(ICT)
PROGRAMME

Project FP7-ICT-2009-C-243881 CerCo

Deliverable 6.4
Organization of an Event Targeted to Potential Industrial Stakeholders

Deliverable 6.5
Organization of an Event Targeted to the Scientific Community

Version 1.0

Main Authors:
Brian Campbell, Ilias Garnier, James McKinna, Ian Stark

Project Acronym: CerCo
Project full title: Certified Complexity
Proposal/Contract no.: FP7-ICT-2009-C-243881 CerCo

Executive Summary

CerCo Work Package 6 on *Dissemination and Exploitation* includes two deliverables that take the form of event organization.

D6.4: Organization of an Event Targeted to Potential Industrial Stakeholders

Realised as:

CerCo: Certifying Costs in a Certified Compiler
Workshop at HiPEAC 2013: 8th International Conference on High-Performance and Embedded Architectures and Compilers
Wednesday 23 January 2013, Berlin.

D6.5: Organization of an Event Targeted to the Scientific Community

Realised as:

CerCo/PROARTIS Technical Day on Innovative Techniques on Timing Analysis
Workshop at ETAPS 2013: European Joint Conferences on Theory and Practice of Software
Sunday 23 March 2013, Rome, Italy

This report describes the completion of these deliverables, their outcome and impact on future dissemination. We note in particular the following.

- Interaction with other projects and research groups. Both events had invited talks from notable researchers, and the workshop targeted to the scientific community at ETAPS was presented jointly with the PROARTIS project.
- Good fit with the other activities at both the industrially-targeted and scientific events: the CerCo event at HiPEAC was part of a workshop track on compilers which ran throughout the conference; and the CerCo day at ETAPS collaborated with the parallel workshop QAPL (Quantitative Aspects of Programming Languages on Systems) through three joint talks.
- Additional scientific impact and further directions, including:
  - Potential of CerCo technology to carry out early-stage timing analysis not reachable with current WCET object-code tools (identified by invited speaker Björn Lisper).
  - Parameterisation of CerCo analyses with respect to different cost algebras (identified in interaction with QAPL speakers).
  - Application of probability distributions over costs to tame cache unpredictability (arose from presentation of Vardanega from PROARTIS).

  All of these are discussed in detail in the report.
- New links to industrial researchers and other European projects: PROARTIS, COST action TACLe, parMERASA, and T-CREST.

Contents

1 Task

CerCo Work Package 6 specifies the following two deliverables under *Dissemination and Exploitation*.

> **D6.4) Organization of an Event Targeted to Potential Industrial Stakeholders**: We will organize a public event opened to industries and other potential stakeholders and we will invite a few potentially interested industries to be identified in D6.2 and during the project development. The event could be affiliated to an international conference relevant to the project and could involve a tutorial on the use of the software developed in CerCo. The deliverable date is only indicative, since we need to identify a suitable conference for affiliation. The event could be co-located and partially overlap with D6.5. [month 34]

> **D6.5) Organization of an Event Targeted to the Scientific Community**: We will organize a public event aimed at presenting the CerCo compiler to the scientific community. The event could be affiliated to an international conference relevant to the project and it could involve a tutorial on the use of the software developed in CerCo. Alternatively, it could consist in a course give in an international summer school on the use and implementation of the CerCo compiler. The deliverable date is only indicative, since we need to identify a suitable conference or summer school for affiliation. The event could be co-located and partially overlap with D6.4. [month 34]

2 Report

The Consortium identified two potentially fruitful destinations at which to hold such events, taking into account suitable candidates, and the opportunity to hold workshops during, or shortly after the end of the Project lifetime, given our requested extension to end month 39, Mar. 2013.

Beyond each Conference call for participation, invitations to the two meetings were sent out to the following researchers and research groups, all world leaders in cost/timing analysis for programming languages, compilers and embedded systems: prof. Kevin Hammond (St Andrew's University, UK), dr. Björn Franke (Edinburgh University, UK), prof. Peter Sewell (Cambridge University, UK), prof. dr. Heiko Falk (Ulm University, DE), prof. dr. Martin Hofmann, dr. Steffen Jost (LMU Munich, DE), dr. Christian Ferdinand (CEO, AbsInt GmbH, DE), prof. Germán Puebla, COSTA team (TU Madrid, E), prof. Björn Lisper (Mälardalen University, SE).

D6.4 Before receiving the Project extension, we had identified for month 37 (Jan. 2013) the workshop track of the HiPEAC Conference series on High Performance and Embedded Architectures and Compilers as a suitable destination, comprising as it does a wide range of sessions bringing together industrial and academic research.

HiPEAC 2013 was held in Berlin, with the CerCo event a half-day workshop (see Table 1): `http://www.hipeac.net/conference/berlin/workshop/CerCo`

Prof. Kevin Hammond (St. Andrews University, UK) gave an invited talk on his work on methods for WCET analysis developed in the EmBounded project (EU FP6 IST-510255),

in particular for Hume programs at the source-level, and their application in the autonomous vehicle guidance and aerospace domains.

Time	Speaker	Title
14:00-14:30	Claudio Sacerdoti Coen	Certified complexity: a general introduction
14:30-15:15	Yann Régis-Gianas	Certifying and reasoning on cost annotations of C programs
15:15-16:00	Paolo Tranquilli	Dealing with loop optimizations and pipelines
16:00-16:30		break
16:30-17:00	Kevin Hammond (Invited talk)	Hume: a Functionally-based Domain Specific Language Targeting Real-time Embedded Systems
17:00-17:30	Roberto Amadio	Certifying and reasoning on cost annotations of functional programs
17:30-18:00	Brian Campbell	A certified proof based on structured traces

Table 1: HiPEAC workshop on Certified Costs: programme 2013-01-23

D6.5 We identified for month 39 (Mar. 2013) a workshop at ETAPS, the pre-eminent European federated conference on programming languages, systems and tools.

ETAPS 2013 was held in Rome, with the full-day CerCo event a "Technical Day on Innovative Techniques on Timing Analysis" (see Table 2):
`http://cerco.cs.unibo.it/innovative_techniques_on_timing_analysis_technical_day`

It ran in parallel with a two-day workshop on Quantitative Aspects of Programming Languages (QAPL'11), sharing three sessions with that meeting. The CerCo workshop also included presentations from Tullio Vardanega, representing the PROARTIS Consortium (FP7-ICT-2009.3.4), and an invited talk from Prof. Björn Lisper (Mälardalen University, SE) on Parametric WCET analysis.

Organization details and attendance The ETAPS event was one of 20 workshops organised over the 4 days either side of the main conference. The event also had three sessions shared with the QAPL workshop (Quantitative Aspects of Programming Languages and Systems), and was the better attended, and scientifically more successful, meeting. The main introduction to the project had an audience of around 35 people, and some QAPL participants have expressed interest in future collaborations.

The HiPEAC workshop was one of 24 such meetings in a programme organised in parallel with the 3 days of the main conference. Attendance was limited (at most 10), which can partially be explained by the workshop being held in parallel with the main conference. There have been practically no industrial attendance to any of the workshops talks. Nevertheless, the main conference also hosted an Industry Session and an Industrial Showcase and the days that preceded our workshop posed several good occasions to get in touch with representatives of the high performance hardware industry and of European projects involved in advanced real time architectures (projects parMERASA, T-CREST and PROARTIS).

Scientific Outcomes Several fruitful discussions and suggestions emerged at both meetings. We try here a brief overview of the most interesting ones.

Time	Speaker	Title
09:00-09:30	Claudio Sacerdoti Coen	Certified Complexity (CerCo)
09:30-10:30	Alessandra di Pierro (Joint talk with QAPL)	Probabilistic static analysis and security trade-offs
10:30-11:00	Yann Régis-Gianas	Certifying and reasoning on cost annotations of imperative programs
11:00-11:30		break
11:30-12:00	Paolo Tranquilli (Joint talk with QAPL)	Cost prediction for loop optimizations
12:00-12:30	Björn Lisper (Invited talk)	Parametric WCET Analysis
12:30-14:00		lunch
14:00-14:30	François Bobot	Frama-C: an Extensible Software Analysis Framework Example of an Automatic Cost Annotation Plugin
14:30-14:45	Björn Lisper	The COST Action TACLe (Timing Analysis on Code-Level)
14:45-15:30	Tullio Vardanega	Probabilistic Time Analysis (PROARTIS)
15:30-16:00		break
16:00-16:15	Gabriele Pulcini	Is source level cost prediction for pipelines and caches possible?
16:15-17:00	Round table	Should certified complexity meet probabilistic time analysis?
17:00-17:30	Claudio Sacerdoti Coen Tullio Vardanega (Joint talk with QAPL)	Innovative techniques in timing analysis: cost prediction on high level languages and probabilistic time analysis
17:30-18:00	Brian Campbell Paolo Tranquilli	A certified proof based on structured traces

Table 2: ETAPS Technical Day on Timing Analysis: programme 2013-03-23

The existence of two different approaches to source-level cost reasoning emerged at the HiPEAC event. The first one, embraced by the EmBounded project, does not try to maximize performance of the code, but is interested only in full predictability and simplicity of the analysis. The second approach, embraced by CerCo, tries to avoid any performance reduction, at the price of complicating the analysis. It is therefore closer to traditional WCET. Technology transfer between the two approaches seem possible. Kevin Hammond's group use amortized analysis techniques to connect local costs about embedded programs in the Hume language to global costs, technology which it may be possible to transfer to the CerCo setting. A key difference in our approaches is that their Hume implementation uses the high predictability of their virtual machine implementation to obtain local cost information, whereas CerCo produces such information for a complex native-code compiler. Replacing their virtual machine with out compiler seems also possible.

At the ETAPS workshop Björn Lisper drew attention to the many points of common interest and related techniques between the work on CerCo and his own on Parametric WCET analysis. The most interesting difference between what we do and what is done in the WCET community is that we use (automated) theorem proving to deal with the control-flow (i.e. to put an upper bound to the executions). The standard technique in parametric WCET

consists in using polyhedral analysis to bound the number of loop iterations. That analysis produces constraints which are solved with the aid of off-the-shelf linear programming tools. Comparing the effectiveness of theorem proving with the effectiveness of polyhedral analysis in computing precise costs interested him.

In addition to his own technical talk, he took the opportunity to advertise, and solicit interest in, the recently formed COST Action IC1202 Timing Analysis and Cost-Level Estimation (TACLe), of which he is Chair. Members of CerCo are going to join the COST Action. This offers very promising potential for future collaborations and the wider communication of results from CerCo.

In particular, during the round-table there clearly emerged an immediate and significant application of the CerCo technology that we missed during the project. WCET analysis has traditionally been used in the verification phase of a system, after all components have been built. Indeed, state-of-the-art WCET techniques all work on the *object* code, which is available only after compilation and linking. Since redesigning a software system is very costly, designers usually choose to over-specify the hardware initially and then just verify that it is indeed sufficiently powerful. However, as systems' complexity rises, these initial safety margins can prove to be very expensive. Undertaking lightweight (but less precise) analysis in the early stages of the design process has the potential to drastically reduce total hardware costs. To perform this analysis, a new generation of early-stage timing analysis tools that do not need the object code are required. The CerCo Trusted and Untrusted Prototypes already fill this niche by working on the *source* code and giving the user the possibility to axiomatize the cost of external calls or that of computing a WCET that is parametric in the cost of unimplemented modules. Before industrial exploitation becomes possible, however, we would need to establish greater levels of predictability, robustness and automation of the analysis.

A common theme emerged from the shared sessions with QAPL, and in particular the invited talk there from prof. Alessandra di Pierro on *probabilistic* timing analysis: the parameterisation of a given timing analysis with respect to different cost *algebras*. In the case of probabilistic analyses, costs are taken with respect to given probability distributions, with *expected* costs being computed. A quick analysis of the labelling approach of CerCo reveals that the method assumes very weak conditions on the cost model that it is able to transfer from the object code to the source code. In particular, probabilistic cost models like the one used by di Pierro satisfy the invariants. Prof. Vardanega's talk emphasised a radical approach to probabilistic analyses, by turning the processor/cache architecture into a probabilistic one to yield an essentially predictable analysis. To exploit the CerCo methodology in case of caches, embracing probabilistic analysis may be the key ingredient. This idea already emerged in the discussions at ETAPS and was investigated during the last period of CerCo.

In the deterministic case studied in CerCo, we have taken a given, fixed, cost algebra of natural numbers (obtained from Siemens data-sheet clock timings) under addition, but already Tranquili's work on *dependent labelling* suggests a move to computing costs in algebras of *functions* (in the case of his analysis of loop unrolling, of cost expressions parameterised with respect to valuations of the loop index variables). The wider implications of such a move are yet to be explored, but probabilistic analysis fits in the model, as well as the computations of costs that are parametric on the hardware state. At both events we presented preliminary results on the time analysis of systems with pipelines obtained by exposing the hardware state in the source code. Some members of the audience were skeptical because of fear of exposing a level of complexity difficult to tame. However, before we get a working implementation and we test the behaviour of invariant generators on the obtained source code, we honestly believe

that it is difficult to come to conclusions.

The feedback obtained from discussions with industrial representatives and with representatives from the parMERASA and T-CREST projects was less significant and describes a bleak future for static time analysis. Microprocessor and embedded systems developers are in a race to provide the largest amount of computing power on a single chip, with systems-on-chip at the centre of the scene during the industrial showcase. The major issue in the design of safety-critical and non-safety-critical systems is now on how to exploit this additional power to optimize the average case or, simply, because the additional power is present anyway and it is a pity to waste it. The timing behaviour of programs running on a computing unit of such multi-core processors or systems-on-chip is potentially greatly affected by the other units. Buses and caches are also shared, often in non uniform ways, and different computations also interfere through the states of these shared components.

Statically analysing a program for WCET in isolation yields totally useless bounds, because ignorance about the behaviour of the other computing nodes forces to always assume the worst possible behaviour, hence hopelessly imprecise bounds. The mentioned EU projects, among others, and a large part of the scientific community is working on the design of alternative hardware that could make the worst case statically predictable again, but at the moment attempts to influence microprocessor manufacturers have been totally unsuccessful. The CerCo technology, which implements a form of static analysis, suffers from the same such problems and does not contribute to its solution. On the other hand, it is likely that, if a solution to the problem emerges, it could be exploited in CerCo too.