



**School of Electronics, Electrical  
Engineering and Computer Science**

Computer Science Building  
14-18 Malone Road  
Belfast BT9 5BN  
Tel +4428 9097 4620

[www.qub.ac.uk/schools/eeecs](http://www.qub.ac.uk/schools/eeecs)

# **INFORME EXPERT EXTERN**

*INFORME EXPERTO EXTERNO*

*EXTERNAL REFEREE REPORT*

Nom de l'expert/ Nombre del experto/ Name of the referee: Prof Dimitrios S Nikolopoulos

Categoria/Categoría/Category Professor and Head of School

Departament, Universitat a què pertany/Departamento, Universidad a la que pertenece/ Department, University to which s/he belongs Electronics, Electrical Engineering and Computer Science, Queen's University Belfast

Títol de la tesi/ Título de la tesis/ Title of the thesis Energy Optimising Methodologies on Heterogeneous Data Centres

Nom del doctorand que presenta la tesi/ Nombre del doctorando que presenta la tesis/ Name of the student presenting the thesis: Rajiv Nishtala

**Especifiqueu les raons que avalen la qualitat de la tesi per a la seva defensa pública :**

**Especificar los motivos que avalan la calidad de la tesis mencionada para su defensa pública:**

**Specify reasons endorsing the quality of the above-mentioned thesis for its public reading:**

Quins objectius s'han assolit amb la tesi ?

¿Qué objetivos se han logrado con la tesis presentada?

What objectives have been achieved with the thesis?

The thesis has demonstrated methods to predict the power consumption and performance of individual processor cores and homogeneous multi-core processors in response to changes in voltage/frequency of each core, contention between cores for shared resources, and core consolidation interventions. The thesis has also demonstrated schedulers for heterogeneous multi-core processors that use voltage/frequency scaling and task-to-core assignment methods to preserve quality of services for latency critical workloads while improving overall system utilization by consolidating latency critical and batch workloads.

Originalitat del treball :

Originalidad del trabajo:

Originality of the work:

The work is of appropriate originality for the award of the PhD. The work uses known statistical and machine learning methods, but the novelty lies in that it formulates new prediction and optimization problems for power and performance scaling. The novelty of the scheduling work presented in the third part of the thesis is particularly high and the overall work in this part is excellent. Its acceptance in a premier computer architecture conference is testament to its quality.



Metodologia emprada / hipòtesis contrastades :  
Metodología usada / hipótesis contrastadas:  
Methodology used / hypotheses tested:

The methodology and hypotheses used in the thesis are broadly appropriate. I have had a number of concerns about the methodology which I would like to explore during examination of the candidate. Specifically:

- Whilst the work targets datacenters the implementation and evaluation in their totality are conducted in standalone servers. The thesis notably misses the evaluation of workloads that are distributed between servers and where servers need to communicate to exchange data. Comments on how the proposed methods would apply or would be extended to address distributed workloads would be necessary
- Because of the above, it is unclear how the proposed methods would work at scale. If the thesis does not explore large-scale datacenter workloads this should be clarified upfront.
- The term 'scalable' is used repeatedly for the prediction methods proposed in the thesis but it is never the case that scalability is defined or evaluated for the methods. The fact that the methods have low time overhead does not mean that they are also scalable. Scalability needs to be defined against a parameter (e.g. number of processors, utilization, etc.) before being measured and claimed as a contribution.
- The term dynamic datacenter is used extensively in the thesis without being defined and without clarity in what it might mean. The term gives the impression that the hardware of the datacenter is dynamic but this is not the case. On the other hand if the term refers to the workload then all datacenter workloads are dynamic, so the term might not be of good use.
- The of PMCs for power modelling is not motivated sufficiently in the thesis. Power prediction in response to DVFS, core consolidation or other interventions can be trained and conducted also with direct power measurements instruments, such as RAPL or voltage sensors. The thesis should clarify why modeling of power using PMCs is necessary in the context of the work.
- The thesis occasionally has claims of optimality in configuration selection, scheduling, etc. I would caution against using the term as optimality is not proven and may even be impossible to prove for the problems explored in the thesis.
- The thesis uses extensively multi-linear regression techniques to predict parameters in response to readings from PMCs. This is a well established method but not the only one. There are several alternatives, including classification methods (SVMs), deep learning methods using neural networks, probabilistic methods, etc. Polynomial methods may also be used and indeed have proven to improve accuracy for such models. I would expect the thesis to provide stronger motivation for the use of multiple linear regression methods and whether it is sufficient for the purposes of the work.
- Throughout the thesis an issue that is not discussed adequately is if the magnitude of the error in predictions forces mispredictions of the optimal state that the cores should be in, in order to achieve the performance or power objective. Prediction inaccuracy varies from very low to very high in the presented experiments but it is hard to know whether the low/high prediction accuracy matters without knowing the impact that (mis)prediction has on optimizing performance and power during execution (not training) of the workloads. Chapter 4 indicates that the predictors are effective in scheduling but it gives no indication on how far from optimal performance and energy the predicted states come.

Valoració absoluta i/o ponderada de la tesi en relació amb altres treballs d'investigació:

Valoración absoluta y/o ponderada de la tesis presentada en comparación con otros trabajos de investigación:

Absolute and/or relative assessment of the thesis in comparison with other works of research:

The thesis advanced the state of the art in scheduling mixed workloads on heterogeneous processors and provides a range of experimental tools and methods to develop better schedulers for such systems, based on prediction. The prediction methods and approaches could have been better motivated and the implications of their accuracy could have been explored further experimentally or analytically. Nevertheless, the use of the proposed framework in scheduling of mixed workloads for heterogeneous systems has been effective. The thesis meets the quality criteria of a PhD dissertation.

Considera que la tesi esmentada és apta per al tràmit de lectura i la defensa pública?

Considera que la tesis anteriormente mencionada es apta para su lectura y defensa pública?

In consideration of all the above, is the thesis judged to be suitable for public reading?

☒ Si / Yes ☐ No

Observacions:

Observaciones:

Observations:

None.

Signatura i data

Firma y fecha

Signature and date



Professor Dimitrios S. Nikolopoulos

May 12<sup>th</sup>, 2017

(\*) Si és necessari adjunteu les respostes en els fulls annexos/ Si es necesario adjunten las respuestas en las hojas anexas/ If necessary, enclose the answers in the annexed sheets