Accepted Manuscript

Exploring the relationship between architectures and management policies in the design of NUCA-based chip multicore systems

Sandro Bartolini, Pierfrancesco Foglia, Cosimo Antonio Prete

PII: S0167-739X(17)31179-2
DOI: http://dx.doi.org/10.1016/j.future.2017.06.001
Reference: FUTURE 3492

To appear in: *Future Generation Computer Systems*

Received date : 23 May 2016
Revised date : 5 June 2017
Accepted date : 5 June 2017

Please cite this article as: S. Bartolini, P. Foglia, C.A. Prete, Exploring the relationship between architectures and management policies in the design of NUCA-based chip multicore systems, *Future Generation Computer Systems* (2017), http://dx.doi.org/10.1016/j.future.2017.06.001

This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

ACCEPTED MANUSCRIPT

Exploring the Relationship between Architectures and Management Policies in the design of NUCA-based Chip Multicore Systems

Sandro Bartolini□, Pierfrancesco Foglia[ψ1], Cosimo Antonio Prete[ψ]
[ψ]*Dipartimento di Ingegneria dell'Informazione, Università di Pisa*
Via Diotisalvi, 2 – 56100 Pisa (PI) Italy
{foglia, prete}@iet.unipi.it
□*Dipartimento di Ingegneria dell'Informazione e Scienze Matematiche*
Università degli Studi di Siena, via Roma 56 – 53100 Siena (SI), Italy
bartolini@dii.unisi.it

Abstract

The last level on-chip cache (LLC) is becoming bigger and more complex to effectively support the various localities coming from multiple cores and threads running concurrently in modern processors. Furthermore, LLC design can be constrained by various restrictions that limit the freedom in their organization, for example in the relative positioning and clustering of processing cores and cache banks. Non Uniform Cache Architectures (NUCAs) offer a hierarchy of access times, which can be usefully exploited by the NUCA management policies (i.e. the ways in which data are either mapped to cache banks and/or moved among them upon access) to achieve high performance and low power consumption. The objective of the work is to single out the optimal combination of data management policies and cache-core layouts and to highlight which is the most performing one. With this aim, we compare two basic layouts for NUCA based systems, the first with cores connected to only one side of the shared NUCA cache (one-side), the second with half of the cores on one side and the others at the opposite side of the NUCA (two-sides). For all the configurations, we evaluate the effectiveness of both static and dynamic NUCAs and, where applicable, we consider also optimizations based on profile-guided bank remapping and replication of shared copies. As overall design guidelines, our results show that the one-side layout achieves the best performance and the lowest power consumption with the considered hw-sw optimizations. Then, similar results can be achieved in the two-sides layout only by introducing more sophisticated copy replications. Lastly, software based profile driven optimization allows the system to achieve the lowest usage of network resources.

Keyword: multicore architecture, sub-banked last level cache structure, layout, design space exploration, data replication, profile-based memory layout optimizations.

[1]Corresponding author

Download English Version:

https://daneshyari.com/en/article/4950187

Download Persian Version:

https://daneshyari.com/article/4950187

Daneshyari.com