Analysis of Hybrid Cache Coherence Schemes in Large Multicore Systems

Alan Mujumdar and Simon Moore

UNIVERSITY OF CAMBRIDGE
Computer Laboratory

Motivation

Memory management is critical in any computer architecture design. Numerous schemes have been proposed in the past, but few were tested in hardware.

Mechanisms such as snoopy coherence are popular in most modern processor designs. They provide strong coherency and are easy to implement. It is generally considered that these schemes do not scale well beyond a few tens of processors.

Directory based schemes have been demonstrated in architectures such as the Stanford DASH processor. Tiled chip multiprocessors use networks on chip and directory based schemes for inter-core communication. Both of these methods have been known to scale above several thousands of cores, however, very few examples have been shown in hardware. Most designs based on directory schemes use a relatively small number of cores and scalability statistics are mostly derived from software simulations.

Hybrid, snoopy-directory coherence schemes have also been proposed but tests have been limited to simulations.

Research Aim

Designing a hybrid coherence system and implementing it in hardware. This hybrid system will blend a local snoopy mechanism with a global directory scheme. We aim to test the scalability and performance of such a system and compare it with other large-scale systems currently in operation. Research has shown that most of the inter-core communication within a chip multi processor occurs between neighbouring cores, hence, a snoopy scheme is very effective. On the other hand, transferring large amounts of data between distant cores in cases such as task migration, a directory scheme is more favourable.

Results

We have succeeded in creating a multi-core system based on the BERI MIPS processor. The BERI processor has been written in Bluespec System Verilog. We have been able to fit a quad-core BERI processor on a single Stratix-IV FPGA chip (Figures 1 & 2).

The cores communicate through a shared L2 cache. A variation of the MSI coherence protocol is currently implemented. Simple parallel applications can be run using this setup.

Free BSD is compatible with this architecture, however, only one of the cores is used by the OS in its present state. We are currently modifying the design to allow the OS to utilize all cores.

In the near future we will be linking multiple FPGA boards using high-speed links. Communication between boards and processors will be channelled over these links and a directory based scheme will be implemented for board to board inter-core communication.

As a proof of concept a parallel bubble sort test was implemented on a dual-core BERI processor in hardware. These results were compared to a single core BERI running the same test. A comparison between the two designs has been shown in terms of speed-up in Figure 3. These results demonstrate that our system is comparable in operation to commercial multi-core processors.

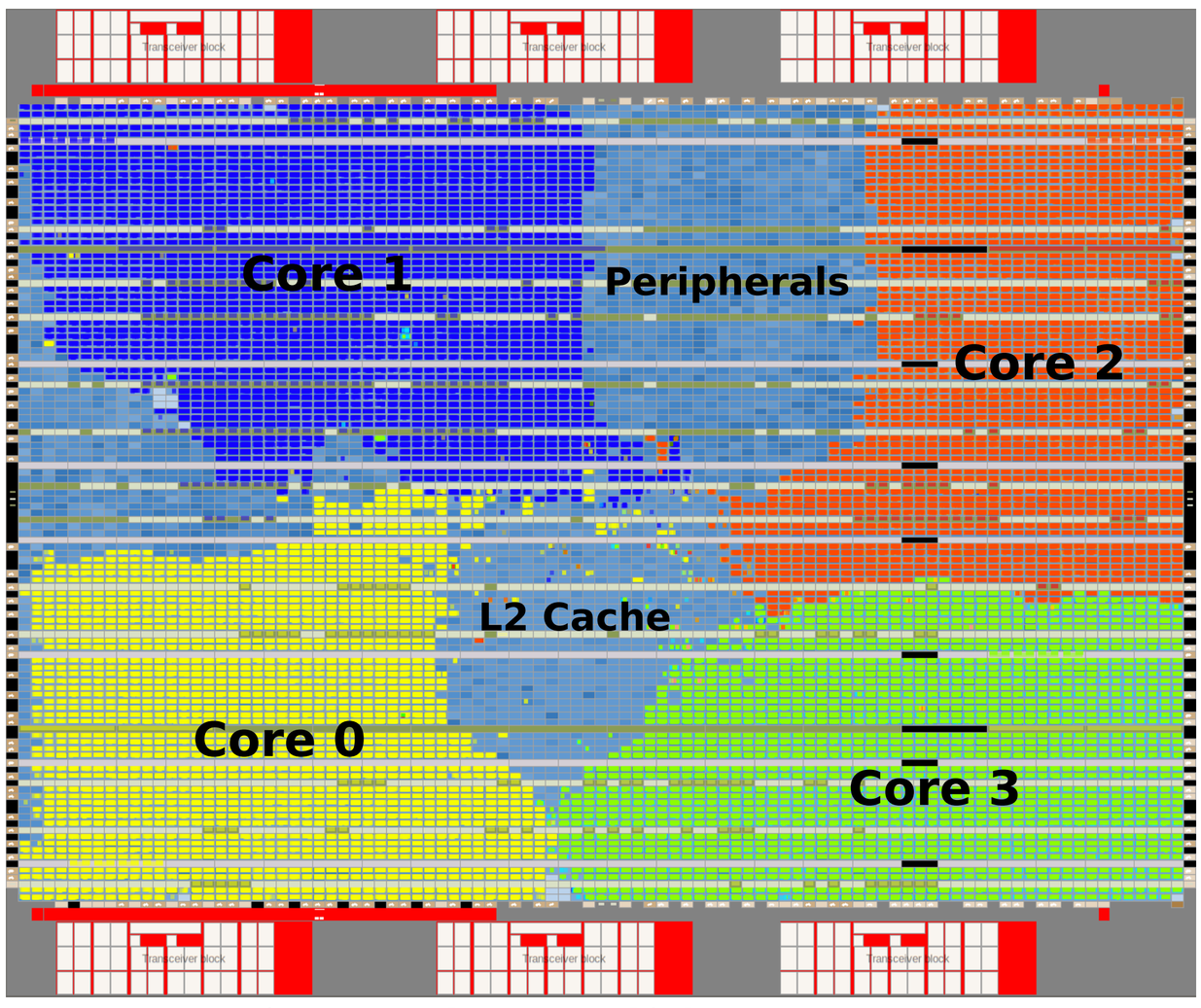

Figure 1: Quad-core BERI chip layout on an FPGA

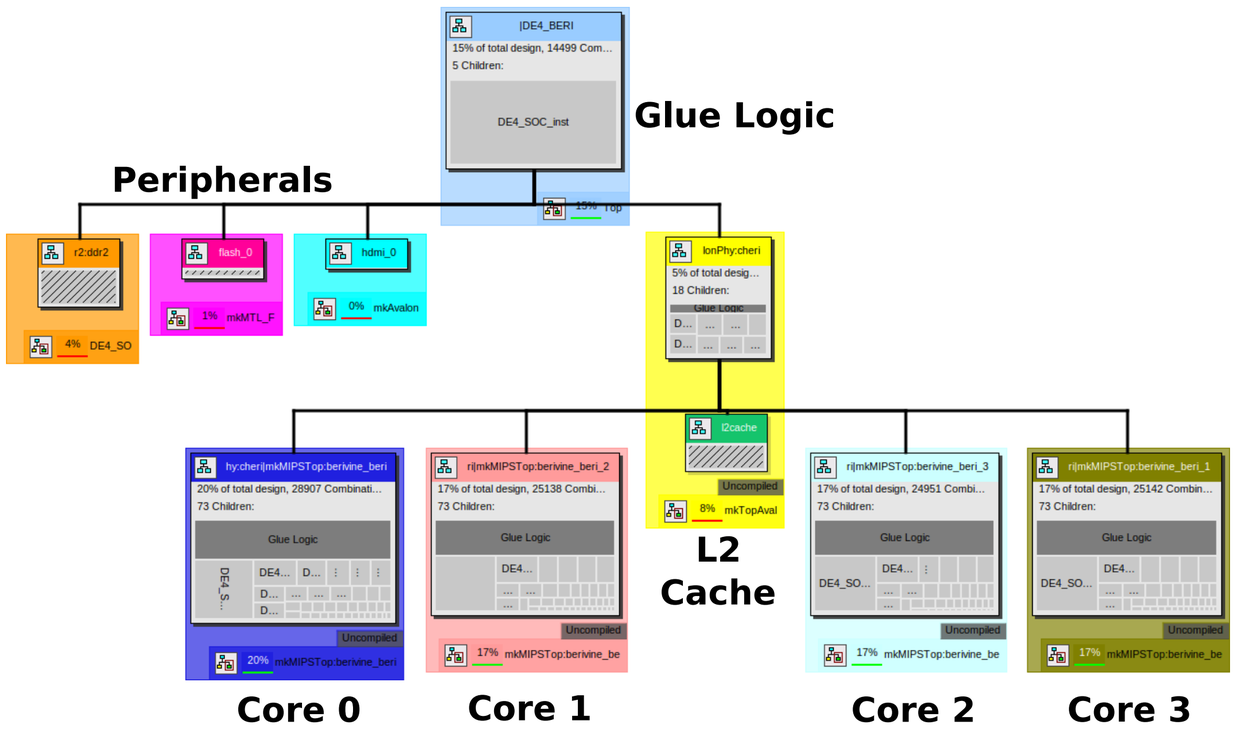

Figure 2: Quad-core BERI interconnection diagram

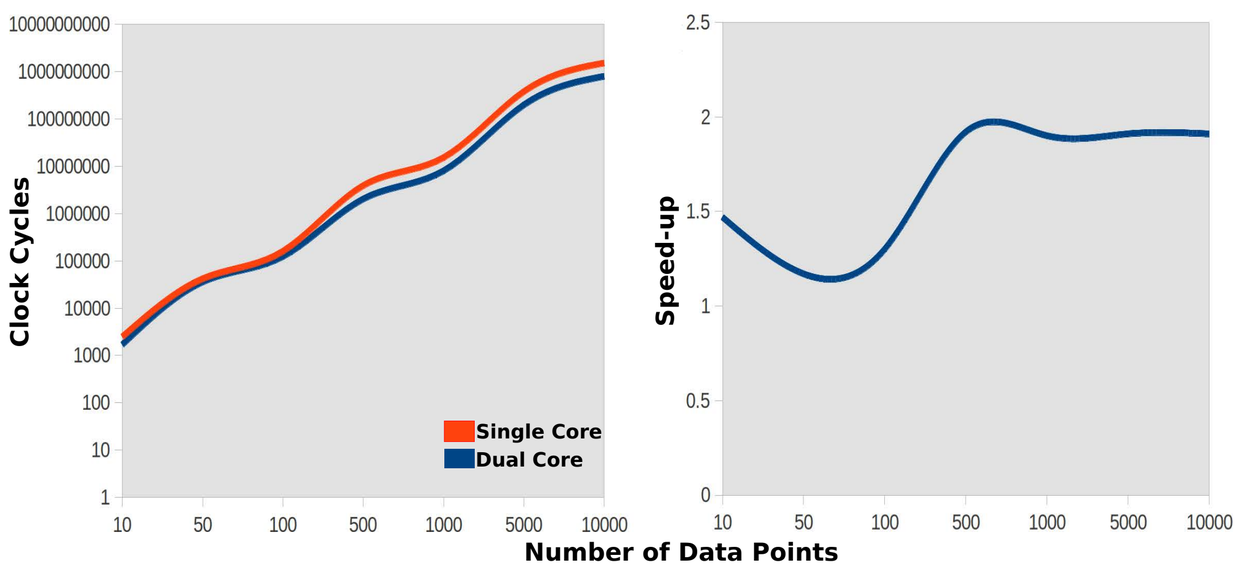

Figure 3: Performance of Dual vs Single core BERI processor

BERI platform funded by DARPA
Approved for public release. This research is sponsored by the Defense Advanced Research Projects Agency (DARPA) and the Air Force Research Laboratory (AFRL), under contract FA8750-10-C-0237. The views, opinions, and/or findings contained in this article/presentation are those of the author/presenter and should not be interpreted as representing the official views or policies, either expressed or implied, of the Defense Advanced Research Projects Agency or the Department of Defense.

alan.mujumdar@cl.cam.ac.uk simon.moore@cl.cam.ac.uk

Computer Architecture Group
http://www.cl.cam.ac.uk/research/comparch/