CEA LETI - DÉPARTEMENT DACLE

TSARLET Programming Interface

Version 0.1

November, 2017

César Fuguet Tortolero

Contents

1 Introduction

The TSARLET Integrated Circuit implements the 16-cores TSAR (Tera-Scale Architecture). These cores are physically and logically organized in four different clusters, interconnected by a Network-on-Chip (NoC).

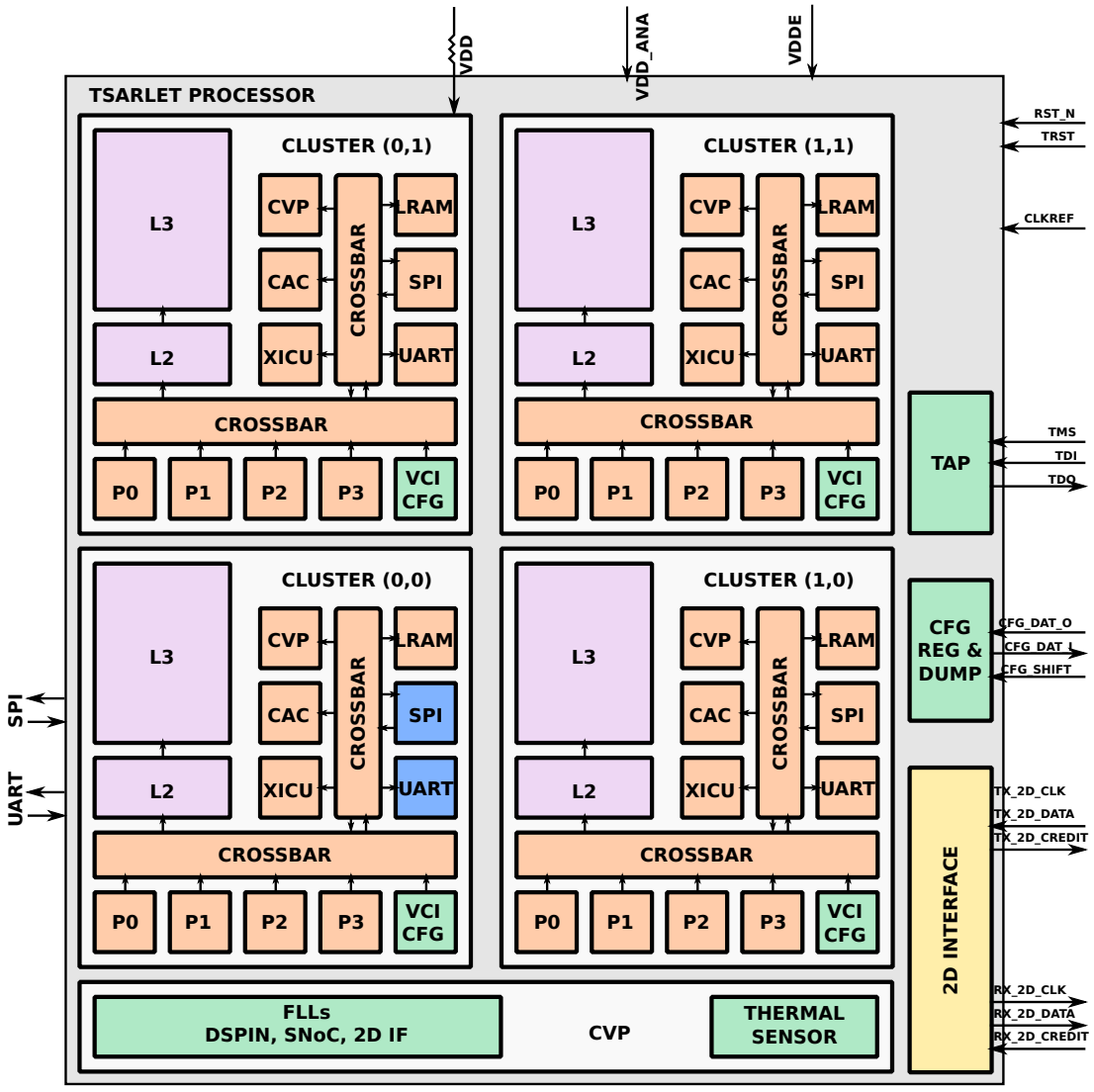

Figure 1: TSARLET Architecture

2 TSARLET General Description

There are three level of caches at TSARLET: two private first level caches (L1 data and L1 instructions), and a shared second and third levels of caches (L2 and L3, respectively). The cache coherency between L1 and L2 caches is implemented by the hardware. Each cluster has: four cores each with its own L1 cache (data and instruction), a L2 cache and a L3 cache.

Parameters of caches are the following:

L1 Data Cache	
Number of sets	64
Number of ways	4
Number of 32-bits words per set	16
Total capacity	16 KBytes
L1 Instruction Cache	
Number of sets	64
Number of ways	4
Number of 32-bits words per set	16
Total capacity	16 KBytes
L2 Cache Controller	
Number of sets	256
Number of ways	16
Number of 32-bits words per set	16
Total capacity	256 KBytes
L3 Cache Controller	
Number of sets	512
Number of ways	16
Number of 32-bits words per set	32
Total capacity	1 MByte

3 TSARLET Cores

TSARLET's cores are simple, single-issue, RISC processors implementing the MIPS32 architecture.

All cores boot at address @0x00000000. As the Memory Management Unit (MMU) is disabled at reset, this corresponds to the @0x0000000000 physical address, which is mapped on the external memory. This means, that a software procedure, shall be written at this address by another entity (external to TSARLET) before the TSARLET reset.

As an example, in the case of the STELLAR board, the boot procedure at address @0x0000000000 is loaded by the Power, Reset, Clock Management Microcontroller Unit (PRCMCU) at the FPGA before resetting TSARLET.

4 TSARLET Memory Management Unit

The MMU (Memory Management Unit) is a hardware component implemented into every L1 cache controller of TSARLET.

The MMU implements two separate TLBs (Translation Look-aside Buffers) which are in charge of performing the virtual to physical address translation, and access rights verification. One TLB is used for data accesses, and the other for instruction accesses.

The parameters of both ITLBs and DTLBs are as follows:

MMU TLBs	
Number of sets	16
Number of ways	4

The TLB miss events are handled by an hardwired "Table Walk" state machine.

4.1 Virtual Memory

The TSARLET's virtual address space size is 4 GB, but the physical address space size is 1 TB. Two different page sizes are implemented: 4 KBytes, and 2 MBytes.

The running operating system is responsible for building the page tables for each process. These page tables are built as shown on Figure 2.

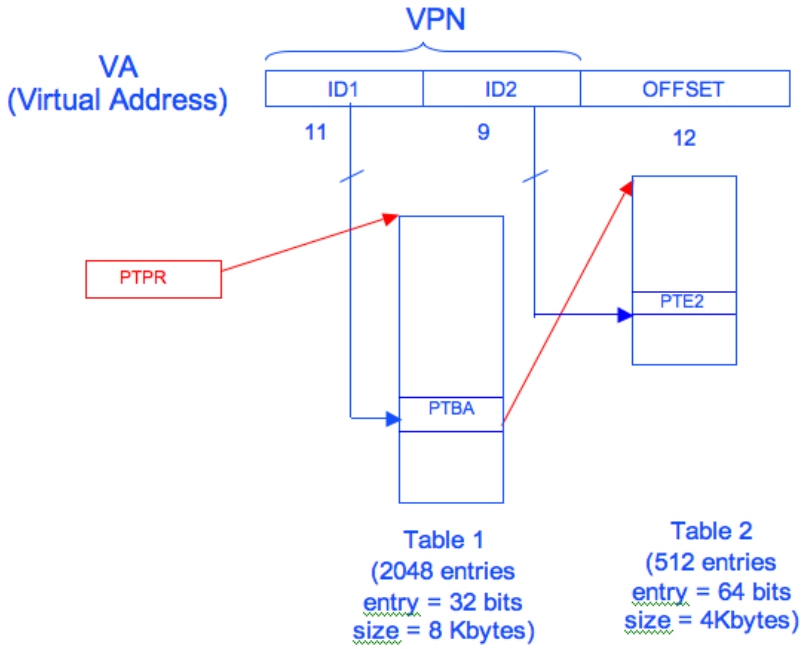

Figure 2: TSARLET's Page Tables (image from the LIP6 TSAR website)

4.1.1 First Level Page Table

An entry of a first level page table is a 32-bits word that contains either a 2MBytes page descriptor (called PTE1), or a second level page table descriptor (called PTD1).

In order to point to the corresponding page tables of a process, the MMU implements a PTPR (Page Table Pointer Register). This register shall be modified by the operating system at each process switch and it should point to the physical base address of a first level page table. The PTPR contains only the 27 most signifiant bits of the physical address, thus a first level page table shall be aligned to 8 KBytes.

The content of a first level page table entry is as follows:

- In case of PTE1:

PTE1 Bit Mapping		
Bit	Name	Description
31	V / Valid	When set, the entry is valid (set by the OS)
30	T / Type	When set, the entry is a PTD1 (set by the OS)
29	L / Local	Set by the hardware when a local core access the page
28	R / Remote	Set by the hardware when a remote core access the page
27	C / Cacheable	When set, the page is cacheable in the L1 cache (set by the OS)
26	W / Writable	When set, writes are allowed on the page (set by the OS)
25	X / eXecutable	When set, the page can contain instructions (set by the OS)
24	U / User	When set, the page is accessible in user mode (set by the OS)
23	G / Global[1]	When set, the page is kept in the TLB during process switch
22	D / Dirty	When set, the page has been modified
21 – 19	Reserved	
18 – 0	PPN1 / Physical Page Number	Concatenated to the 21 LSb of the virtual address

- In case of PTD1:

PTD1 Bit Mapping		
Bit	Name	Description
31	V / Valid	When set, the entry is valid (set by the OS)
30	T / Type	When unset, the entry is a PTD1 (set by the OS)
29 – 28	Reserved	
27 – 0	PTBA / Page Table Base Address	28 MSb of a second level page table base address

4.1.2 Second Level Page Table

An entry of a second level page table is a 64-bits word. An entry is as follows:

[1] Global bit shall not be set as there is a known bug

PTE2 (First Word) Bit Mapping		
Bit	Name	Description
31	V / Valid	When set, the entry is valid (set by the OS)
30	T / Type	Shall be 0 (set by the OS)
29	L / Local	Set by the hardware when a local core access the page
28	R / Remote	Set by the hardware when a remote core access the page
27	C / Cacheable	When set, the page is cacheable in the L1 cache (set by the OS)
26	W / Writable	When set, writes are allowed on the page (set by the OS)
25	X / eXecutable	When set, the page can contain instructions (set by the OS)
24	U / User	When set, the page is accessible in user mode (set by the OS)
23	G / Global[2]	When set, the page is kept in the TLB during process switch
22	D / Dirty	When set, the page has been modified
21 – 8	Reserved	
7 – 0	Can be used freely by the OS	

PTE2 (Second Word) Bit Mapping		
Bit	Name	Description
31 – 28	Reserved	
27 – 0	PPN2 / Physical Page Number	Concatenated to the 12 LSb of the virtual address

4.1.3 MMU Registers

TODO: define MMU registers

5 TSARLET Internal Peripherals

TSARLET Internal Peripherals	
Base Address	Peripheral
0xXYE0000000	L2 Configuration Interface (L2CFG)
0xXYF0000000	eXtended Interrupt Controller Unit (XICU)
0xXYF2000000	Serial Peripheral Interface (SPI)
0xXYF4000000	Teletypewriter Controller (TTY)
0xXYF5000000	CVP
0xXYF6000000	L3 Cache Access Controller (CAC)
0xXYFF000000	Scratchpad Memory (MEM)

The most signifiant bits of the physical address designates the X and Y coordinates of the cluster. All these peripherals are replicated in all clusters. However, only the TTY and SPI controllers of cluster 0 are usable as only those outputs are connected to the IO ring of the circuit.

[2]Global bit shall not be set as there is a known bug

5.1 Interrupts Routing

This subsection describes the routing of all internal distributed peripherals to the distributed interrupt controllers. The following table resumes the routing of hardware interrupts to the XICU component:

TSARLET XICU Hardware Interrupts Routing	
Hardware Interrupt Index	Peripheral
$0-7$	Reserved
8	Memory Cache Configuration (Segmentation Fault Signalling)
9	SPI
10	TTY/UART
11	Cache Access Controller (CAC)
12	CVP

Regarding the routing of interrupts between the XICU and the cores, it is the following:

TSARLET XICU Output Interrupts		
XICU Interrupt Output Index	Cores	Cores Inputs
$3-0$	Core 0	$3-0$
$7-4$	Core 1	$3-0$
$11-8$	Core 2	$3-0$
$15-12$	Core 3	$3-0$

Finally, **regarding the number of XICU mailboxes (WTI) and programmable timers (PTI), both are sixteen (16)**.

As a reminder, this interrupt routing is the same in all 4 clusters of TSARLET, as all internal peripherals are replicated. However, regarding SPI and TTY controllers, only those of cluster 0 are actually usable interrupts.

6 TSARLET Known Bugs

Global Bit in the Page Tables Entries

The global bit in page tables shall not be set by the operating system, as it can cause a complete stall of the system.

A particular scenario makes that an entry with the global bit, can be replicated in 2 or more slots of the TLB, which reduces the capacity of the TLB. In the worst case, the entry is replicated in all slots, and this causes a live-lock, as this affects, at the end, the victim selection function of the TLB.

Unhandled Race Condition on L3 External Accesses

An unexpected behavior can arise on L3 caches when intense traffic is received, and the L3 to XRAM communication channel bandwidth is completely used. In those cases, a race condition can arrive where during a L3 cache writeback, an obsolete version of the data is written on external memory.

Through the reconfiguration mechanism of the L3 caches, this can be avoided with a negative impact on L3 cache capacity and bandwidth. The capacity is actually divided by 2, so the per-L3 cache capacity is 512 KByte instead of 1 MByte.

With the workaround, the L3 cache external (XRAM) access bandwidth is, in the worst case, also divided by 2.