15.3 A 477mW NoC-Based Digital Baseband for MIMO 4G SDR

Fabien Clermidy[1], Christian Bernard[1], Romain Lemaire[1], Jerome Martin[1], Ivan Miro-Panades[1], Yvain Thonnart[1], Pascal Vivet[1], Norbert Wehn[2]

[1]CEA-LETI-Minatec, Grenoble, France
[2]Technical University of Kaiserslautern, Kaiserslautern, Germany

Baseband processing for advanced Telecom applications have to face two contradictory issues [1]. The first one is the flexibility required, with the exploding number of modes for a single protocol (e.g. 63 for 3GPP-LTE), the number of protocols to be supported by a single chip (>10 in 2010) and new applications requiring a handover between protocols. The second concern is related to performance and power consumption: performance demands are exploding (up to 100GOPS are now required) with decreasing power consumption constraints (roughly 500mW).

To tackle these issues, previous solutions have used heterogeneous NoC-based platforms [2-4]. This work uses a digital baseband chip built on a semi-homogeneous NoC-based architecture. The main innovations are fast run-time reconfiguration (<10µs for configuring a new telecom physical layer) and distributed power management for fine adaptation to performance requirements (leading to 477mW power consumption). We use 5 programmable VLIW cores, where each core is capable of 3.2GOPS performance using 50mW; we also use a 17Gb/s throughput-per-link communication framework.

Our chip is organized around a 15-router asynchronous NoC that connects 22 processing units. The chip structure and topology are shown in Fig. 15.3.1. Flexibility comes from programmable units replicated as hard-macro clones: the OFDM (instantiated 4 times), specialized complex-number computing units (MEPHISTO) (5 times), and Data and Configuration Memory (DCM) (4 times). The OFDM core is an IFFT/FFT engine associated with a framing/deframing phase for pilots/data association/split. It can compute from a 64 up to a 2K-point FFT while loading the following data in the good shape. The MEPHISTO core is a VLIW structure organized around a MAC dedicated to complex numbers signal processing such as channel estimation or MIMO decoding. MEPHISTO reduces the control power consumption compared to a straightforward implementation from 70mW down to 7mW thanks to reconfigurable profiles and cache instructions strategy. The DCM core supports complex intra-chip communications. It can manage up to 4 parallel data flows with fully programmable data re-ordering and merging/splitting of flows. It is also used as configuration server for the computing cores. While the standalone DCM contains 1Mb embedded SRAM, a special version can offer up to 128Mb of external memory. The general manager of the chip is an ARM1176 core with a floating-point unit and a 2Mb SRAM shared with a DCM. It can handle interruptions from the NoC and directly send configurations to the cores. It also supports the debug for the whole chip, thanks to specialized NoC instructions, and 2 direct wires to each unit: a soft reset and a stepping control signal. Finally, 2 specialized cores handle bit-level computing such as mapping and interleaving, and 3 efficient channel-decoding cores perform variable types of LDPC and turbo-decoding functions.

Each processing unit is associated to a Communication and Configuration Controller (CCC) that supports the NoC protocol and offers a unified distributed programming model. The chip supports full dynamic reconfigurability between telecom protocols, thanks to 5 DCM servers and the possibility to manage distributed configuration flows. Each CCC runs a micro-program sequencing the required configurations of the core and the data-flows. After receiving a start instruction, it points on the first configuration identifier of the micro-program. Then, it checks in all the configuration slots of the processing unit if the corresponding id is already loaded. If the configuration is missing, it generates a request to its associated DCM (association is programmable), which returns the corresponding configuration. An end-of-configuration signal from the unit side is used to allow a new configuration to be loaded, and finally the CCC can point to the next configuration. This scheme assures as-soon-as-possible pre-loading policy and is compatible with double-buffering used to hide latency. Taking into account the configuration time required for each processing unit (Fig. 15.3.2), the whole reconfiguration cannot take more than 10µs when distributed over 4 DCMs.

The chip power consumption is controlled at different levels (Fig. 15.3.3). The asynchronous routers are event-driven, which naturally limits dynamic consumption to a minimum (<1mW per router on a typical application). A comparison with an equivalent synchronous router (same performance) shows a consumption gain of more than 7× due to event-based feature. Inside each unit, a programmable local clock, implemented as a looped delay line packed in a hard macro, generates frequencies between 320 and 790MHz distributed over 16 operating points (124ps steps). Slower frequencies are reachable through a clock divider (up to 16×). They are tuned to limit the power consumption for underused units. Additionally, classical gated-clock techniques are used during the back-end process. In our chip this solution is efficient to cut between 70 and 85% of the dynamic consumption, depending on the processing units. To further reduce the power, a high-level clock-gating mechanism is added into the CCC to totally cut the dynamic consumption of the core in idle phases. This general clock gating strategy is triggered on-line by the CCC when inactivity phases are detected: as the CCC has a full view of the communications of the units with the outside, it can cut the core clock when no data are forecasted to be sent or received by the core. For better efficiency, an optional gate-enable signal from the core allows cutting the clock as soon as data processing is achieved.

The chip layout, implemented in a 65nm CMOS technology is presented in Fig. 15.3.7. The semi-homogeneous structure is connected through the asynchronous NoC routers. The beneficial properties of the Quasi-Delay Insensitive (QDI) asynchronous technique used for the NoC routers and links allow a natural pipelining of long wires. As a consequence, 17Gb/s throughput is achieved in each NoC link/router in conjunction with low latencies: 1.82ns per router, 1.98ns including the links (with 2 QDI half buffers). The NoC routers (78Kgates/router) and the GALS interfaces (6.5Kgates) do not exceed 10.2% of the whole core area (27mm²). The chip complexity represents 96M transistors including 8.7Mb of embedded SRAM including 6Mb of shared memory in DCM.

The circuit has been programmed to process the digital baseband of a MIMO 4×2 3GPP-LTE receiver (Fig. 15.3.4). It achieves the hard real-time constraint of 1ms for a frame decoding and a 10.8Mb/s throughput (receiver using 10 Resource Blocks) and a total power consumption of 219mW for the physical layer, under 1.2V. For the full load of 54Mb/s (receiver using 50 RBs), the power is estimated at 668mW.

In Fig. 15.3.5, our chip is compared with state-of-the-art digital baseband circuits. Results show its good features, namely an efficient NoC that gives twice the throughput of the most efficient previously reported solution and low power consumption due to its special management technique. Finally, this chip is the only one that uses a fast configuration scheme between physical layers, allowing its use for new applications such as cognitive radio.

References:

[1] C.H. Van Berkel, "Multi-Core for Mobile Phones", invited paper, *Design, Automation and Test in Europe DATE'09*, March 2009.

[2] D. Lattard, E. Beigne, F. Clermidy, Y. Durand, R. Lemaire, P. Vivet, F. Berens, "A Reconfigurable Baseband Platform Based on an Asynchronous Network-on-Chip", *IEEE Journal of Solid State Circuits*, Vol. 43, Issue 1, pp. 223-235, Jan. 2008.

[3] D. Arditti I. Jeffrey D. Hoffman, A. Chun, B. Perez Esparza, "Architecture of the Scalable Communications Core's Network On Chip", *IEEE Micro*, pp. 62-74, Sept.-Oct. 2007.

[4] T. Limberg, M. Winter, M. Bimberg, R. Klemm, E. Matus, M. Tavares, G. Fettweis, H. Ahlendorf, P. Robelly, "A Fully Programmable 40 GOPS SDR Single Chip Baseband for LTE/WiMAX Terminals", *European Solid-State Circuits Conference ESSCIRC'08*, Sept. 2008.

[5] A. Nilsson, E. Tell, D. Liu, "An 11mm² 70mW Fully-Programmable Baseband Processor for Mobile WiMAX and DVB-T/H in 0.12µm CMOS", *2008 IEEE International Solid-State Circuits Conference,* Feb. 2008.

 978-1-4244-6034-2/10/$26.00

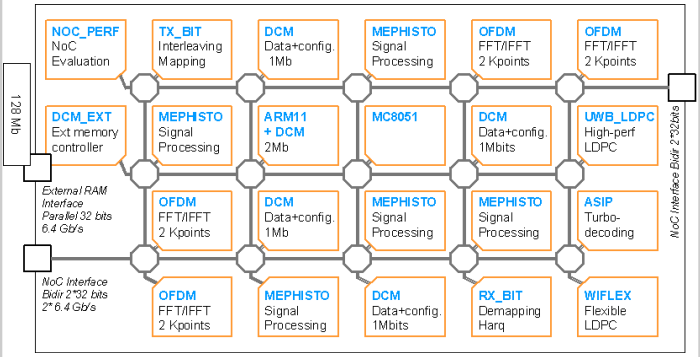

Figure 15.3.1: Structure and processing cores.

		Area (mm2)	Operating Frequency (MHz)	Performance (GOPS)	Configuration Time (ns)	Total power LTE Rx (mW)	Instances Total	Instances Used in LTE Rx
Processing cores	MEPHISTO	0.63	400	3.2	2580	14.3	5	4
	OFDM	0.75	400	3.2	330	12.7	4	2
	RX_BIT	0.72	333	1.3	1560	10.6	1	1
	TX_BIT	0.26	400	1.2	213	not used	1	0
	UWB_LDPC	0.58	400	1.2	8	not used	1	0
	WIFLEX	1.17	400	2.4	3010	not used	1	0
	ASIP	0.7	400	2.4	16675	66.0	1	1
Other units	ARM11	4.19	313	MAC support	—	150.0	1	0
	DCM	1.64	350	—	—	7.2	4	2
	DCM_EXT	0.54	350	—	—	3.7	1	1
	NODE	0.16	530	—	—	0.7	15	10
	NOCIF	0.013	200	—	—	0.5	2	1
	I/O	2.39	200	—	—	34.6	1	1
TOTAL (with all instances)				37.3 Gops	36 µs			
TOTAL (instances LTE RX)						219 mW		

Figure 15.3.2: Main features of each unit.

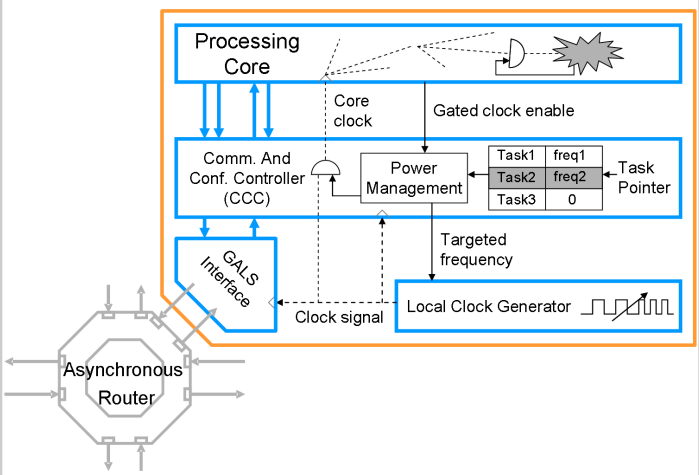

Figure 15.3.3: Distributed power management scheme.

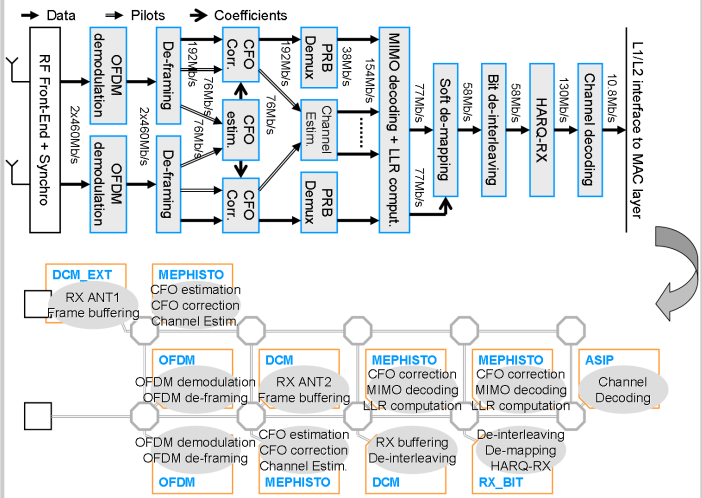

Figure 15.3.4: 3GPP-LTE 4×2 MIMO scheme on the receiver side.

	MAGALI	FAUST [2]	SCC [3]	TOMAHAWK [4]	LINKOPING [5]
Platform scope	MIMO 3GPP-LTE, WiMAX, 802.11n, Cognitive Radio	SISO 3GPP-LTE, WiMAX	SISO WiMAX, 802.11-04	MIMO 3GPP-LTE, WiMAX, H264	SISO WiMAX, DVB-T/H
Configuration scheme	Fully dynamic	Static	Multiple static	Static	Static
Power scheme	GALS, Local DFS, application-driven gated clock	GALS, Centralized DFS	Global frequency	Global frequency	Global frequency
Memory organization	Fully distributed	3 centralized units	Local	Centralized and Local	Fully Distributed
Peak performance	37.3GOPS	6GOPS	Partial data	40GOPS	9.6GOPS
Application for power measurements	4x2 MIMO 3GPP-LTE Rx, 2x2 MIMO Tx, MAC, 10.8Mb/s	SISO 3GPP-LTE Tx+Rx, MAC	No data	Estimation for MIMO 3GPP-LTE Tx+Rx	SISO DVB-T-H, 31.67Mb/s
Power consumption	477mW@1.2V	1.2W@1.2V	No data	1.2W* @1.3V	70 mW@1.2V
NoC throughput (per link)	17Gb/s	5.12Gb/s	8Gb/s	5.47Gb/s	No data
Die size	29.6 mm²	80mm²	25mm² (estimated)	100mm²	11mm²
Technology	CMOS65 nm	CMOS130 nm	CMOS 65nm	CMOS 130nm	CMOS 120nm

*without units specialized for H264

Figure 15.3.5: Comparison to previous baseband chips.

NoC Architecture	Topology	2-D regular mesh
	Number of routers / units	15 routers / 22 units
	Network protocol	Packet-switching, source routing
	Flit size	34 bits (32 payload + control)
	Reconfiguration	Control distributed, up to 5 parallel configuration flows
	Flowcontrol	Distributed in each unit, credit-based synchronization
	Links with external world	2 direct bidirectional NoC access (4*36 bits)
	NoC implementation	Asynchronous Quasi-Delay-Insensitive
	Clock domains / Freq.	22 programmable local clocks / 320-790 Mhz, 16 points
Technology	Process	65 nm STMicroelectronics, LowPower, 3 voltage thresholds (Low, Standard and High)
	Logic gates / Memory	11 MGates / 8.7 Mb SRAM
	MOS count	96 MTransistors
	Core size	5.21 mm x 5.17 mm = 27 mm2
	I/O count	363 (255 logical)
	Supply voltage	1.2 V Core, 2.5 V I/O
Power consumption	Rx MIMO 4×2	219 mW
	Tx MIMO 2×2	60 mW (estimated)
	MAC layer	150 mW
	I/O	48 mW
	Total	477 mW

Figure 15.3.6: Main features and performance metrics.

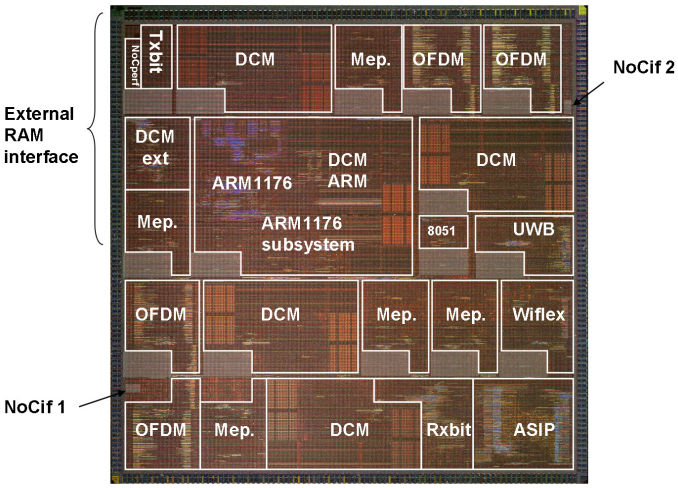

Figure 15.3.7: MAGALI micrograph.

978-1-4244-6034-2/10/$26.00