

Towards a New Baseband Processing Architecture for Next Generation Software-Defined Radio

Nelson Silva, Arnaldo S. R. Oliveira, Nuno Borges de Carvalho

Abstract – The necessity for better radios with increased flexibility, easier design and verification led a paradigm shift in favour of the Software-Defined Radio (SDR). On the other hand, the SDR implementation of Next Generation Wireless Networks (NGWN) will require significantly higher power efficiency than current processors can provide.

In this paper, we present a baseband processing architecture designed to shorten the gap between the achievable and the NGWN processing requirements. In this sense, the challenge is to develop a computational architecture with inherent high performance while maintaining low power consumption.

By matching the processing architecture with the physical layer of NGWN, it is possible to achieve higher power efficiency. Thereby, each processing unit of the proposed architecture is optimized for specific operations, as also the interconnection between processing units is designed to match to the NGWN processing chain.

Resumo – A necessidade de melhores rádios, dotados de maior flexibilidade e que facilitem o projecto e verificação levaram a uma mudança de paradigma a favor do Rádio Configurável por Software (SDR). Por outro lado, a implementação de redes sem fio de próxima geração (NGWN) através de SDR irá requerer unidades de processamento com uma eficiência energética muito superior à actualmente permitida.

Neste artigo apresentamos uma nova arquitectura de processamento de banda base, criada para encurtar a distância entre a actual capacidade de computação e os requisitos de processamento das redes de próxima geração. Neste sentido, o desafio consiste em desenvolver uma arquitectura computacional com inerente alto desempenho, sendo o consumo de energia reduzido ao mínimo. Ao corresponder a arquitectura de processamento com a camada física das NGWN é expectável que se obtenha uma maior eficiência energética. Assim, cada unidade de processamento da arquitectura proposta é optimizada para operações específicas, bem como a interligação entre as várias unidades de processamento é projectada para corresponder com a cadeia de processamento das NGWN.

Keywords – Baseband Processing, DSP, NGWN, SDR

Palavras chave – Processamento de Banda Base, DSP, Redes Sem Fio de Próxima Geração, SDR

I. INTRODUCTION

The growing necessity of higher flexibility and reduced Time-to-Market (TTM) make traditional wireless radios deprecated and fosters new approaches such as the Software-Defined Radio (SDR) [1]. In an SDR, the major baseband processing operations (e.g. filtering, modulation, error correction) are carried out by software instruc-

tions running over a Digital Signal Processor (DSP).

Compared with traditional radios where its functioning is based on Application Specific Integrated Circuits (ASICs), SDRs can provide several important benefits. In fact, due to the high flexibility that is allowed by such radio, a software update can be enough to support new standards or to improve existing features. Moreover, since the same hardware can perform communication over multiple wireless standards (e.g. GSM, Wi-Fi, LTE), it fosters interoperability with other radios as also mass IC manufacturing, which may allow cheaper devices with reduced size and weight.

Due to its advantages, SDRs are expected to be among the key techniques to serve the future needs of the wireless communications market. Future radios would be able to observe the environment and automatically select the adequate bands, standards or applications in order to meet the desired needs. For that, it is fundamental to have a highly flexible radio, which is very hard (if not impossible) to achieve with traditional hardware-based architecture radios.

On the other hand, such full-featured radio also has a considerable price to pay. In fact, a SDR that has the flexibility to steer to any band, to tune to one or more channels of any bandwidth and to receive any modulation [2] will certainly require powerful hardware, such as processors with a huge Digital Signal Processing (DSP) capacity.

A radio with such characteristics may still be far away from being feasible. However, even in a near future it is expected that Next Generation Wireless Networks (NGWN) will require much more processing capacity when compared with previous standards. In fact, Fourth Generation (4G) wireless networks will require about one to three orders of magnitude more computational capacity when compared with Third Generation (3G) wireless networks, while maintaining a reduced power consumption [3]. Such performance gap must be reduced. In this sense, innovative processing architectures with inherent high computational capacity must be explored [4].

The remainder of this paper is organized as follows. Section II presents the physical layer basics of a possible 4G wireless system. Section III summaries several high performance DSP architectures. Section IV introduces the proposed wireless baseband processing architecture as an innovative computing model for reducing the gap between the processing requirements of NGWN and the achievable processing capacity. At last, Section V presents the main conclusions and future work.

II. 4G WIRELESS PHYSICAL LAYER BASICS

4G networks gained importance due to the increasing demand for wireless systems with improved mobility and data rate. The expected throughput of 100Mbps up to 1Gbps, for high and low mobility situations, respectively, requires new approaches for implementing the 4G physical layer [4].

By using transceiver arrays (see Fig. 1), it is possible to increase data rate and signal robustness, which seems to be a possible approach for implementing 4G systems.

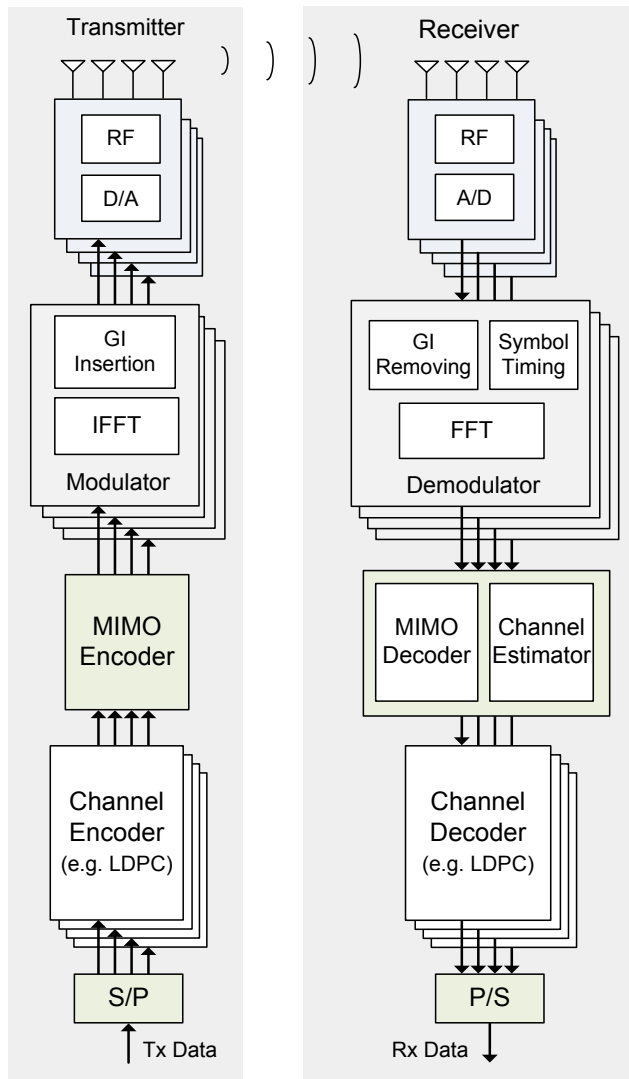


Fig. 1 - Physical layer block diagram of a possible 4G wireless system.

Fig. 1 depicts the physical layer block diagram of a possible 4G wireless system [4], already used in the field for evaluation purposes [5]. The major DSP-intensive blocks of the transceiver chain are the Orthogonal Frequency Division Multiplexing (OFDM) modulator/demodulator (modem), the Multiple Input Multiple Output (MIMO) encoder/decoder and the channel encoder/decoder. The demodulator converts the incoming amplitude and phase time domain signals to data in the frequency domain. Due to its efficient computation, the Fast Fourier Transform (FFT) algorithm is usually used to perform the discrete time-to-frequency conversion. The modulator transmits amplitude

and phase time domain signals by performing operations similar to the demodulator but in reverse order.

The MIMO decoder is typically used for two different purposes: i) combine the received signals from the multiple antennas to generate a signal with higher robustness, ii) multiple incoming signals are used to increase the data rate. The MIMO encoder performs the reverse operation by multiplexing data signals over multiple antennas.

At last, the Forward Error Correction (FEC) is implemented by the channel encoder/decoder pair. Currently, high performance FEC algorithms with closer Shannon capacity are the Low Density Parity Check (LDPC) and the Turbo Code. LDPC has higher performance however, Turbo Code requires less computational capacity. Due to its superior power efficiency, LDPC and Turbo Code are expected to be among the key FEC algorithms for use in NGWN [6].

III. HIGH PERFORMANCE DSP ARCHITECTURES SURVEY

The growing demand for improved DSP led researchers to develop new architectures, capable of delivering high performance on specific applications. The presented DSP solutions can be categorized in the following application domains: i) baseband processing solutions, where processors are optimized for supporting baseband processing of current generation wireless networks; ii) multimedia processing solutions, where processors are optimized for other DSP-intensive applications, such as graphics rendering.

A. Wireless Baseband Processing Solutions

The Montium tile processor [7] is an example of an energy-efficient, coarse-grained reconfigurable architecture, suitable for wireless baseband processing. The Montium processor is comprised by five identical Arithmetic Logic Units (ALUs) and ten local memories, all interconnected by ten configurable global buses, similar to a crossbar switch. Montium has a regular architecture, which makes easier to increase the processing capacity. However, due to the higher number of ALUs, the efficiency of such architecture is very dependent on the intelligence of the compiler to provide high code optimization. In addition, mapping wireless baseband algorithms on such architecture may also be a challenging task.

The Embedded Vector Processor (EVP) [8] is specialized in supporting 3G standards. EVP is comprised by a general purpose processor, a programmable vector processor, a configurable filter processor, a conventional DSP and a configurable channel decoder, all interlinked by a shared bus.

Vector processors have superior performance for streaming applications. However, traditional vector processors have performance limitations due to the complexity and size of its centralized vector register [9].

Sandbridge Sandblaster [10] is a commercial processor developed for baseband processing of current wireless networks protocols. It comprises four architecturally identical DSP units, each one providing scalar and vector processing, all interconnected through a shared bus. Sandblaster also supports multithreading which allows to improve perfor-

	Imagine	IBM Cell	NVIDIA G70	Philips EVP	Sandbridge Sandblaster	SODA	ARM Ardbeg
Architecture	Stream	SIMD+Scalar	Graphics Pipeline	Vector	Multithread+SIMD	SIMD+Scalar	SIMD+Scalar
Manufacturing Tech. (nm)	150	90	110	90	90	180	90
Clock Speed (MHz)	400	4000	430	300	600	400	350
# DSP Cores	1	8	-	1	4	4	2
Core Type	VLIW+SIMD	SIMD	SIMD	VLIW+SIMD	SIMD	VLIW+SIMD	LIW+SIMD
Power Consumption (mW)	3000	50000	81000	300	70	4000	210
Peak Performance*	32	520	176	6	3	50	9
Power Efficiency**	10.6	10.4	2.2	20.0	42.9	12.5	42.9
Hardware Coprocessor	No	No	No	Yes	No	No	Yes
Scratchpad Memory	Yes	Yes	Yes	Yes	Yes	Yes	Yes

TABLE I

ARCHITECTURAL COMPARISON BETWEEN SEVERAL HIGH PERFORMANCE DSP ARCHITECTURES. *PERFORMANCE METRICS FOR IMAGINE, IBM CELL, SODA AND ARDBEG ARE IN GOPS AND FOR NVIDIA G70, PHILIPS EVP AND SANDBRIDGE SANDBLASTER ARE IN GFLOPS, GIPS AND GMACS, RESPECTIVELY. **POWER EFFICIENCY FOR IMAGINE, IBM CELL, SODA AND ARDBEG ARE IN MOPS/MW AND FOR NVIDIA G70, PHILIPS EVP AND SANDBRIDGE SANDBLASTER ARE IN MFLOPS/MW, MIPS/MW AND MMACS/MW, RESPECTIVELY.

mance by exploiting ILP. However, such technique requires additional hardware overhead, such as cache coherency, besides not being fully exploited due to the reduced ILP of wireless baseband algorithms [11].

SODA [12] is a high performance and low power processor. It has one General Purpose Processor (GPP) and four identical Processing Elements (PEs), all interconnected through a shared bus. Each PE has dedicated memory and specialized hardware for improving the performance of common wireless baseband operations. Wide SIMD architectures are used for exploring the high DLP of typical wireless baseband algorithms.

SODA has a processing architecture specialized for baseband processing of current generation wireless networks. However, the global shared bus may impose scalability restrictions for achieving the NGWN requirements.

Ardbeg [11] appears as an evolution of the SODA processor. It has one control processor, two PEs, a high performance interconnect bus and a coprocessor for turbo code acceleration. Similarly with SODA, each execution unit has a 512-bit wide SIMD for exploiting the high DLP of baseband processing operations. The coprocessor addition together with other architectural modifications, such as the Long Instruction Word (LIW) support and the implementation technology change of 180nm in SODA to 90nm in Ardbeg, allowed Ardbeg to achieve a 3.4x average speedup over SODA and about 7x lower power.

Ardbeg is a high performance processor, well suited for wireless baseband processing of current generation networks. However, both Ardbeg and SODA PEs only have one memory port which makes serialized memory accesses a performance bottleneck for certain algorithms, such as the Turbo Code and the serial architecture LDPC.

B. Multimedia Processing Solutions

Imagine [13] is a stream processor tailored for media processing applications. Imagine has a 128-Kbyte Stream Register File (SRF) for data temporary storage, eight arithmetic clusters controlled by a microcontroller, a streaming memory system and a stream controller. Since multimedia applications usually have high Instruction-Level Parallelism (ILP) and Data-Level Parallelism (DLP), Imagine provides Very Long Instruction Word (VLIW) and Single Instruction

Multiple Data (SIMD) support for performance improvement, respectively, for exploring both types of parallelism.

Cell [14] is a high performance processor, optimized for multimedia and vector processing applications. It combines eight Synergistic Processor Units (SPUs), used for data processing, with a general-purpose IBM Power architecture, used for control tasks, all interconnected by a coherent bus. Each SPU combines scalar and SIMD processing.

Graphics Processing Units (GPUs) are another type of high performance DSP architectures, specially developed for speeding up graphics processing. Unlike conventional processors that use the classical von Neumann architecture, GPUs employ a different computational model that is more adequate to the graphics processing pattern. Nowadays, GPUs are capable of very high throughput by exploiting massive parallel processing over a programmable graphics hardware pipeline [15]. Moreover, by creating a processing architecture that is closer with the typical operations and requirements of graphics processing, it is possible to improve performance as also power efficiency.

The presented architectures achieve superior performance by exploring different strategies of parallel execution (e.g. multi-core, VLIW, SIMD). However, they are not optimized for wireless baseband processing, which naturally leads to computational and power inefficiencies [12].

At last, an extended architectural comparison between several high performance DSP architectures is presented on Table I. These architectures use extensive parallelism to achieve the required performance. However, NGWN have power efficiency requirements of about 500 to 25000 MOPS/mW, which far exceed the capabilities of current processors.

IV. PROPOSED ARCHITECTURE

NGWN will require an estimated processing capacity increase up to three orders of magnitude over the existing wireless networks while keeping a low power consumption [4]. As discussed above, current processors are far from scaling with the next generation wireless requirements, which leads an opening for exploring innovating approaches with inherent higher power efficiency.

It is well known that clock frequency is reaching a boundary and computational performance no longer scales with

the clock frequency as well the power consumption no longer scales with the lithography. Other approaches such as the Multi-Processor System-on-Chip (MPSoC) are also used for obtaining superior performance. However, the addition of a high number of processors significantly increases the complexity of the hardware, compiler, application mapping and power consumption, which is not compatible with the NGWN requirements.

On the other hand, by matching the processing architecture with the desired application, it is possible to achieve higher power efficiency, which seems to be a feasible solution for NGWN. In fact, current major wireless baseband processors implement MPSoC with hardware support for wide SIMD operations, which allow to significantly improve performance by exploiting the high DLP of common wireless baseband operations.

The proposed architecture goes one step forward by optimizing each processing unit to specific DSP kernels and by matching the interconnection between processing units with the next generation wireless baseband processing chain, Fig. 2.

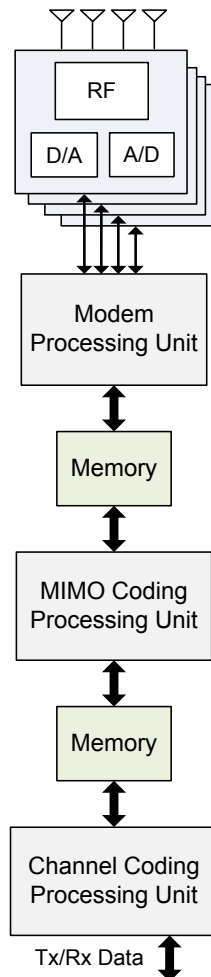


Fig. 2 - High level block diagram of the proposed 4G wireless physical layer processing architecture.

A similar approach was already implemented in traditional GPUs, where the high achievable throughput results from the extensive DLP exploitation and by adapting the processing

chain to match to the graphics pipeline. As in graphics, streaming computation is also very well suited for wireless baseband processing [12]. In that sense, vector and stream processing architectures seem to be an interesting approach for next generation wireless baseband processing. However, using vector and stream processing approaches such as the EVP and the Imagine stream architecture may not be desirable, since the centralized, large and complex register file may impose performance limitations [9].

The proposed architecture also breaks from traditional MPSoC designs for SDR baseband processing by avoiding a global bus, shared by all processing units. Instead, it was adopted a Point-to-Point (P2P) topology where each processor is connected only with its neighbours through a scratchpad memory. Due to the local data spatiality, reduced complexity and power consumption when compared with cache, scratchpad memory is very well suited for SDR baseband processing.

We believe that the P2P interconnect topology is a key element in redefining the concept of high performance wireless baseband processing architectures. In fact, such interconnect scheme does not suffer from scaling limitations inherent to the use of a shared bus as also it allows to reduce the interconnection complexity among processing units, which in turn leads to area and power consumption savings. In addition, P2P permits communication parallelism as also allows to reduce unnecessary data movement by reducing the interconnection path among consecutive processing units. By travelling shorter paths, the inherent reduced latency fosters higher throughput as also the shorter data movement will require less energy.

In order to achieve higher power efficiency, all processing units of the proposed architecture must be optimized for effective execution of the next generation wireless DSP kernels. However, dissimilar DSP kernels require different hardware solutions. For instance, a Finite Impulse Response (FIR) filter is well handled by SIMD processing architectures, while Turbo Code is better handled by application specific hardware, eventually offloaded on a coprocessor [8], [11]. Thus, achieving higher power efficiency requires specific algorithm optimization on each processing unit of the hierarchical chain.

On the other hand, due to the architecture high specialization, the support for concurrent execution of dissimilar wireless protocols may lead to computational and power inefficiencies. However, this may be considered a small price to pay for enabling NGWN at a near future. Moreover, there are several strategies that allow to overcome these issues. In fact, the implementation of an Advanced Power Management (APM) system in each processing unit may reduce power inefficiencies by adjusting clock frequencies and by disabling unused hardware. Furthermore, implementing the processing architecture on reconfigurable hardware, such as Field Programmable Gate Arrays (FPGAs), may also provide additional flexibility by adjusting the hardware to the requirements of the executing protocols.

On the software side, the application mapping becomes easier since each processing unit is optimized for specific DSP kernels. Contrarily to many of the previously dis-

cussed MPSoC that require specific programming models (such as Synchronous Data Flow - SDF) for enabling efficient computation, the proposed architecture requires one program for each processor, which can be efficiently done by using conventional and widely known programming languages, such as C or C++. Since software is a key element in this type of radios, simplifying its development is crucial for enabling a broader SDR adoption.

V. CONCLUSION AND FUTURE WORK

The above discussion summarized several DSP architectures, capable of delivering high performance by exploiting parallel execution. However, NGWN have performance and power consumption requirements that are far from being allowed by current processing architectures.

Due to its inherent high performance, we believe that the proposed baseband processing architecture goes one step forward, allowing to shorten the gap between the achievable and the NGWN requirements. In addition, by matching the processing architecture with the physical layer chain of the NGWN, it is possible to achieve higher power efficiency. Thereby, the proposed architecture has each processing unit optimized for specific DSP kernels, as also the hierarchical interconnection between processing units was designed to match to the NGWN processing chain. Moreover, by avoiding a global bus shared by all processing units, it is possible to reduce hardware interconnection complexity, silicon area and power consumption as well to improve data throughput.

Future work involve the development of a prototype based on the proposed architecture, followed by an extensive evaluation which will allow to quantify the achievements made on performance and power consumption.

REFERENCES

- [1] J. Mitola, "The software radio architecture", *IEEE Communications Magazine*, vol. 33, no. 5, pp. 26–38, May 1995.
- [2] A. A. Abidi, "The path to the software-defined radio receiver", *IEEE Journal of Solid-State Electronics*, vol. 42, no. 5, pp. 954–966, May 2007.
- [3] Mark Woh, Sangwon Seo, Hyunseok Lee, Yuan Lin, Scott Mahlke, Chaitali Chakrabarti, and Krisztian Flautner, "The Next Generation Challenge for Software Defined Radio", *Lecture Notes in Computer Science*, vol. 4599, pp. 343–354, 2007.
- [4] M. Woh, Yuan Lin, Sangwon Seo, T. Mudge, and S. Mahlke, "Analyzing the scalability of SIMD for the next generation software defined radio", in *Proc. IEEE International Conference on Acoustics, Speech and Signal Processing ICASSP 2008*, 2008, pp. 5388–5391.
- [5] H. Taoka, K. Higuchi, and M. Sawahashi, "Field Experiments on Real-Time 1-Gbps High-Speed Packet Transmission in MIMO-OFDM Broadband Packet Radio Access", in *Proc. VTC 2006-Spring Vehicular Technology Conference IEEE 63rd*, 2006, vol. 4, pp. 1812–1816.
- [6] T. Lestable, et al., "Block-LDPC Codes vs Duo-Binary Turbo-Codes for European Next Generation Wireless Systems", in *Proc. VTC-2006 Fall Vehicular Technology Conference 2006 IEEE 64th*, 2006, pp. 1–5.
- [7] Paul Heysters, Gerard Smit, and Egbert Molenkamp, "A Flexible and Energy-Efficient Coarse-Grained Reconfigurable Architecture for Mobile Systems", *The Journal of Supercomputing*, vol. 26, pp. 283–308, 2003.
- [8] Kees van Berkel, Frank Heinle, Patrick P. E. Meuwissen, Kees Moerman, and Matthias Weiss, "Vector Processing as an Enabler for Software-Defined Radio in Handheld Devices", *EURASIP J. Appl. Signal Process.*, vol. 2005, pp. 2613–2625, 2005.
- [9] Christos Kozyrakis and David Patterson, "Overcoming the Limitations of Conventional Vector Processors", in *Proceedings of the 30th Annual International Symposium on Computer Architecture (ISCA)*, 2003, pp. 399–409.
- [10] J. Glossner, K. Chirca, M. Schulte, Haoran Wang, N. Nasimzada, D. Har, Shenghong Wang, Jr. Hoane, A. J., G. Nacer, M. Moudgill, and S. Vassiliadis, "Sandblaster low power DSP [parallel DSP arithmetic microarchitecture]", in *Proc. Custom Integrated Circuits Conference the IEEE 2004*, 2004, pp. 575–581.
- [11] Mark Woh, Yuan Lin, Sangwon Seo, Scott Mahlke, Trevor Mudge, Chaitali Chakrabarti, Richard Bruce, Danny Kershaw, Alastair Reid, Mladen Wilder, and Krisztian Flautner, "From SODA to scotch: The evolution of a wireless baseband processor", in *MICRO '08: Proceedings of the 2008 41st IEEE/ACM International Symposium on Microarchitecture*, Washington, DC, USA, 2008, pp. 152–163, IEEE Computer Society.
- [12] Yuan Lin, Hyunseok Lee, M. Who, Y. Harel, S. Mahlke, T. Mudge, C. Chakrabarti, and K. Flautner, "SODA: A Low-power Architecture For Software Radio", in *Proc. 33rd International Symposium on Computer Architecture ISCA '06*, 2006, pp. 89–101.
- [13] B. Khailany, W. J. Dally, U. J. Kapasi, P. Mattson, J. Namkoong, J. D. Owens, B. Towles, A. Chang, and S. Rixner, "Imagine: media processing with streams", vol. 21, no. 2, pp. 35–46, 2001.
- [14] H. Peter Hofstee, "All About the Cell Processor", in *IEEE Symposium on Low-Power and High-Speed Chips(COOL Chips VIII)*, April 2005.
- [15] John D. Owens, David Luebke, Naga Govindaraju, Mark Harris, Jens Krüger, Aaron E. Lefohn, and Timothy J. Purcell, "A Survey of General-Purpose Computation on Graphics Hardware", *Computer Graphics Forum*, vol. 26, no. 1, pp. 80–113, 2007.

URL: <http://www.blackwell-synergy.com/doi/pdf/10.1111/j.1467-8659.2007.01012.x>