



Deliverable – D17

Feasibility study for using GALS NoC in the GALS system implementation

| | |
|---|--|
| Grant Agreement No: | 214364 |
| Project acronym: | GALAXY |
| Project title: | GALS InterfAce for CompleX Digital System Integration |
| Funding Scheme: | STREP |
| Date of latest version of Annex I against which the assessment will be made: | 28.10.2008. |
| Contractual Date of Delivery to the EC: | 31. Oct. 09 |
| Actual Date of Delivery to the EC: | 31. Oct. 09 |
| Author(s): | Milos Stanisavljevic (EPFL) |
| Participant(s): | EPFL, UNIBO |
| Work Package: | WP8 |
| Security: | Public |
| Nature: | Report |
| Version: | 1 |
| Total number of pages: | 16 |

Abstract:

This report investigates the effectiveness of use of NoCs in the target hardware accelerator design and assesses the applicability of the NoC solution. The system planned to be GALSified is investigated in the scope of trade off between NoC implementation difficulties and benefits. An alternative hardware accelerator design that could benefit more from NoC implementation is discussed. Finally, the optimal GALS interface architecture for targeted application is analyzed.

- **Keyword list: point-to-point topology, multi-point topology, OFDM baseband processor, deinterleaver, Viterbi decoder, LDPC decoder, Reed-Solomon decoder**



GALAXY

GALS InterfAce for CompleX Digital
SYstem Integration

Confid. Level: Public
Date : 02/11/2009
Issue: 1

| Function | Responsibility | Date | Signature |
|---------------------|-----------------------|-------------|------------------|
| Written by: | Milos Stanisavljevic | 28.10.2009 | |
| Checked by: | | | |
| Approved by: | | | |

Reserved to EC

| | | | |
|---------------------|--|--|--|
| Approved by: | | | |
|---------------------|--|--|--|





GALAXY

GALS InterfAce for CompleX Digital
SYstem Integration

Confid. Level: Public
Date : 02/11/2009
Issue: 1

CHANGE RECORDS

| <i>ISSUE</i> | <i>DATE</i> | <i>§ : CHANGE RECORD</i> | <i>AUTHOR</i> |
|--------------|-------------|--------------------------|----------------------|
| 1 | 28-Oct-09 | Main evaluation | Milos Stanisavljevic |
| | | | |



BIBLIOGRAPHIC RECORD

| | |
|---|--|
| Project Number: | 214364 |
| Project Title: | GALAXY |
| Deliverable Type: | Report |
| Deliverable Number: | D17 |
| Contractual Date of Delivery: | 31. Oct. 2009 |
| Actual Date of Delivery: | 31. Oct. 2009 |
| Title of Deliverable: | Feasibility study for using GALS NoC in the GALS system implementation |
| Work package contributing to the Deliverable: | WP8 |
| Authors: | Milos Stanisavljevic |
| Abstract | This report investigates the effectiveness of use of NoCs in the target hardware accelerator design and assesses the applicability of the NoC solution. The system planned to be GALSified is investigated in the scope of trade off between NoC implementation difficulties and benefits. An alternative hardware accelerator design that could benefit more from NoC implementation is discussed. Finally, the optimal GALS interface architecture for targeted application is analyzed. |
| Keywords | point-to-point topology, multi-point topology, OFDM baseband processor, deinterleaver, Viterbi decoder, LDPC decoder, Reed-Solomon decoder |
| Confidentiality Level | Public |
| Name of Client: | EC |
| Distribution List: | GALAXY, EC, internet |
| Authorised by: | |
| Issue: | 1 |
| Document ID: | D17 |
| Total Number of Pages: | 16 |
| Contact Details: | krstic@ihp-microelectronics.com |



TABLE OF CONTENTS

| | | |
|----------|---|-----------|
| 1 | INTRODUCTION | 6 |
| 2 | REFERENCES | 8 |
| 2.1 | ACRONYMS | 8 |
| 2.2 | REFERENCE DOCUMENTS..... | 8 |
| 3 | GALS HARDWARE ACCELERATOR | 10 |
| 3.1 | THE CHOICE OF DESIGN | 10 |
| 3.2 | BENEFITS AND DIFFICULTIES IN IMPLEMENTING NOC IN DESIGN..... | 12 |
| 3.3 | AN ALTERNATIVE TARGET DESIGN | 13 |
| 4 | FINAL DISCUSSION AND CONCLUSION | 16 |

LIST OF FIGURES

| | | |
|------------------|---|-----------|
| Figure 1: | Typical NoC architecture | 7 |
| Figure 2: | Top level block diagram of OFDM processor | 11 |
| Figure 3: | NoC optimal deinterliver/decoder part of ODFM receiver | 14 |

LIST OF TABLES

| | | |
|-----------------|---|-----------|
| Table 1: | Complexity of FPGA realization of OFDM processor | 11 |
|-----------------|---|-----------|



1 INTRODUCTION

Very Large Scale Integrated (VLSI) circuits designed using modern Computer Aided Design (CAD) tools are becoming faster and larger, incorporating millions of smaller transistors on a chip. VLSI designs can be divided into two major classes: Synchronous and Asynchronous circuits. Synchronous circuits use global clock signals that are distributed throughout their sub-circuits to ensure correct timing and to synchronize their data processing mechanisms [SU02, SA04]. Asynchronous circuits contain no global clocks. Their operation is controlled by locally generated signals [OH02]. Asynchronous circuits [EBE04] have many potential advantages over their synchronous equivalents including lower latency, low power consumption, and lower electromagnetic interference [EBE04][SAK07].

Synchronous design methodologies have been plagued by the necessity to distribute the centralized clock all throughout the chip with an acceptably low skew. This task has become even more difficult with the ever decreasing feature size of modern technologies. The recent trend to integrate more and more functions on a single chip, called System-on-Chip (SoC), provides additional challenges. SoCs require a great deal of modularity, where previously designed hardware modules can easily be embedded in a larger design. This typically results in a system that consists of several hardware blocks that have been designed to operate with different clock frequencies. To derive synchronous local clocks from a centralized clock in such a system is an involved task without a trivial solution [GUR02].

Asynchronous design techniques have always attracted attention as an alternative, especially for SoC integration, as they do not rely on a centralized clock. This alleviates problems related to clock distribution and enables the integration of hardware blocks with different clock domains. However, asynchronous design techniques also have a number of well known shortcomings, most notably complicated design methodologies and lack of reliable tools and test methodologies [GUR02].

The Globally-Asynchronous Locally-Synchronous (GALS) is a relatively new VLSI system design methodology, that promises to combine the advantages of both synchronous and asynchronous operations [CHA84]. In GALS, rather coarse grained synchronous functional blocks are surrounded by self-timed wrappers that include a local clock generator and asynchronous communication ports. The functional blocks operate synchronous to the local clock but communicate asynchronously with similar blocks [GUR02].

Globally-asynchronous locally-synchronous (GALS) systems may become a solution for nowadays challenges in the field of VLSI design. The ITRS road-map [SIA06] predicts that, as a solution to the clock distribution problem, GALS will become mainstream design technique in the near future. In a GALS system, a number of synchronous islands of logic communicate asynchronously using a suitable interconnect. Fully synchronous chips are becoming not feasible anymore due to clock distribution and power consumption problems. The value of GALS lies in combination of well-known synchronous design methods and relative simple asynchronous communication channels. The key components are the communication control ports around the synchronous modules and the stretchable clock also called a wrapper. This clock has an unbound delay and is controlled by events the asynchronous channel [BL02]. GALS is a solution to combine the advantage of asynchronous and synchronous circuits design.



Recently, Networks on Chip (NoC) [DEM06] have emerged as one of the important techniques to develop complex System on Chip designs. The NoC paradigm seems to be very attractive solution for the future chip interconnects since NoC aims to provide scalable communication bandwidth through a modular interconnect design. The typical NoC architecture is illustrated in Figure 1.

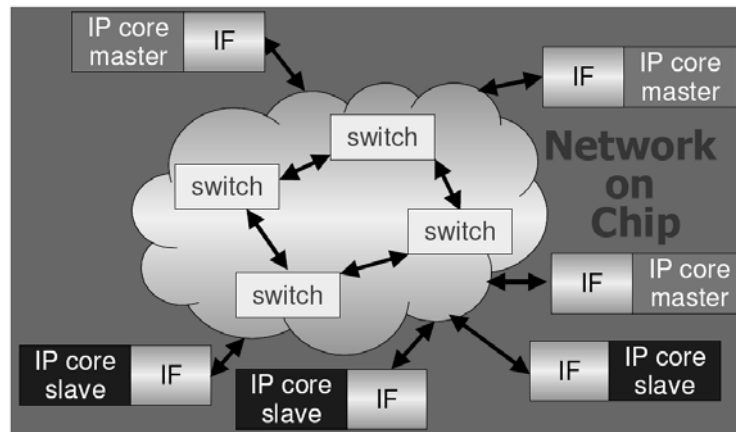


Figure 1: Typical NoC architecture

The lesson learned from state-of-the-art NoC prototypes is that almost 50% of total NoC power is drained by the clock tree (clock distribution and flip-flops). Moreover, the problem of distributing the global clock in a chip with minimal clock skew is getting difficult to solve due to increases in clock frequencies, smaller feature sizes and growing design complexities, which indicates that a global timing notion will not be feasible for highly integrated nanoscale designs.

GALS technology is generally viewed to be a breakthrough technology for future system interconnect designs. There are many parallels between the design approach of GALS systems and NoCs are generally accepted as a great application for GALS systems. GALS-based NoCs would allow the synchronous design of network nodes at their optimum clock frequency, while facilitating asynchronous communication between modules. This would reduce the timing convergence constraints during back-end physical design steps, remove the power-hungry global clock tree and pave the way for new variation-tolerant on-chip interconnection schemes. Unfortunately, although these considerations would appear to be common sense, they have not been enough to remove the barriers to the adoption of GALS interconnects due to the lack of convincing analysis and exploration frameworks, crossbenchmarking with synchronous solutions, proven robustness against nanoscale physics effects and tool support.

Even though there majority of GALS implementations are point-to-point GALS chips, there are already a few NoC platforms based on the GALS architecture [DOB05, BEI05, BJE05].

In the following text we will investigate the effectiveness of use of NoCs in the target hardware accelerator design and assesses the applicability of the NoC GALSification. We will also discuss an alternative hardware accelerator design that supports NoC implementation and make the final discussion about an optimal design for the project.



2 REFERENCES

2.1 ACRONYMS

| | |
|-------------|---|
| CAD | Computer Aided Design |
| CP | Content Protection |
| FEC | Forward Error Correction |
| FIFO | First-in First-out |
| GALS | Globally Asynchronous Locally Synchronous |
| IP | Intellectual Property |
| LDPC | Low-density Parity-check Code |
| MIMO | Multiple-Input and Multiple-Output |
| NI | Network Interface |
| NoC | Network on Chip |
| OFDM | Orthogonal Frequency Division Multiplex |
| QoS | Quality of Service |
| RS | Reed-Solomon |
| SoC | System on Chip |

2.2 REFERENCE DOCUMENTS

* Where we put the reference outside of the sentence at the end of a paragraph that means that the whole paragraph is cited from the referenced document

| Ref. | Document Title |
|-------------|--|
| [BEI05] | E. Beigne, et al. An Asynchronous NOC Architecture Providing Low Latency Service and its Multi-Level Design Framework. In Proc. of the 11th IEEE Int. Symp. on Asynchronous Circuits and Systems, 2005. |
| [BEI06] | E. Beigne, P. Vivet, Design of On-chip and Off-chip Interfaces for a GALS NoC Architecture, Proceedings of 12th IEEE International Symposium on Asynchronous Circuits and Systems (ASYNC'06), Grenoble, France, pp. 172-181, March 2006. |
| [BEI08] | Edith Beigné, Fabien Clermidy, Sylvain Miermont, Pascal Vivet: Dynamic Voltage and Frequency Scaling Architecture for Units Integration within a GALS NoC. NOCS 2008: 129-138 |



| | |
|---------|---|
| [BJE05] | T. Bjerregaard, et al, An OCP Compliant Network Adapter for GALS-based SoC Design Using the MANGO Network-on-Chip, Proc. of the International Symposium on System-on-Chip (SoC'05), pp. 171-174, 2005. |
| [BL02] | Blaauwendraad, B., Student thesis, Title: TIR, design and testing of a Simple GALS, Linköping University, Department of Electrical Engineering, 2002 |
| [DOB05] | R. Dobkin, et al, An Asynchronous Router for Multiple Service Levels Networks on Chip, Proc. of 11th IEEE Int. Symp. on Asynchronous Circuits and Systems (ASYNC 2005), New York, Mar 2005. |
| [EBE04] | Aristides Efthymiou, John Bainbridge, Douglas A. Edwards, Adding Testability to an Asynchronous Interconnect for GALS SoC, Asian Test Symposium 2004: 20-23 |
| [CHA84] | D. M. Chapiro, Globally Asynchronous Locally Synchronous. Systems. PhD thesis, Stanford University, 1984. |
| [DEM06] | G. DeMicheli, L. Benini, Networks on Chips: Technology and Tools (Systems on Silicon), Morgan Kaufmann, ISBN 978-0123705211, 2006 |
| [GRA07] | E. Grass, et al, 60 GHz SiGe-BiCMOS Radio for OFDM Transmission, The IEEE International Symposium on Circuits and Systems (ISCAS), New Orleans, USA, pp. 1979-1982, 27 - 30 May 2007. |
| [GUR02] | F. Gürkaynak et al., A Functional Test Methodology for Globally-Asynchronous Locally-Synchronous Systems, <i>Proc. of ASYNC</i> , pp. 181-189, 2002. |
| [GUR06] | F. Gürkaynak, GALS System Design: Side Channel Attack Secure Cryptographic Accelerators, PhD thesis, Hartung-Gorre Verlag, 2006 |
| [KRS06] | M. Krstic, E. Grass, C. Stahl, M. Piz, System Integration by Request-driven GALS Design, IEEE Proc. Computers & Digital Techniques, Vol. 153, Issue 5, September 2006, pp 362-372. |
| [LAT07] | D. Lattard, et al, "A Telecom Baseband Circuit-Based on an Asynchronous Network-on-Chip", Proc. of the International Solid State Circuits Conference, ISSCC'2007, San Francisco, USA, Feb 2007. |
| [NEX04] | Lines, A., Asynchronous interconnect for synchronous SoC design, IEEE Micro, 2004, 24, (1), pp. 32-41 |
| [OH02] | Eunjung Oh, Soo-Hyun Kim, Dong-Ik Lee, Ho-Yong Choi, High-Level Test Generation for Asynchronous Circuits from Signal Transition Graph, IEICE Trans on Fundamentals of Electronics, Communications and Computer Sciences Vol.E85-A No.12, 2002, pp. 2674-2683 |
| [SA04] | Mathew A. Sacker, Andrew D. Brown, Andrew J. Rushton, Peter R. Wilson, A behavioral synthesis system for asynchronous circuits, IEEE Trans. VLSI Syst. 12(9): 978-994 (2004) |
| [SIA06] | SIA, The International Technology Roadmap for Semiconductors: 2006 Update, Semiconductor Industry Association, San Jose, CA (http://public.itrs.net), 2006. |
| [SU02] | Yin-He Su, Ching-Hwa Cheng, Shih-Chieh Chang, Novel Techniques for Improving Testability Analysis, IEICE Trans on Fundamentals of Electronics, Communications and Computer Sciences Vol.E85-A No.12, 2002, pp. 2901-2912 |



3 GALS HARDWARE ACCELERATOR

3.1 THE CHOICE OF DESIGN

The main ASIC implementation with GALAXY project is a complex target system using both a standard synchronous CMOS design flow and the GALS design flow developed in this project (WP3, WP4, WP5) in a cutting edge 40 nm CMOS process. Selecting a suitable target platform is important. Our main goal is to design a complex GALS system for the high-end communication application with supporting datarates up to 1 Gbps [GRA07]. The purpose of this system implementation is evaluation of the proposed novel GALS design flow, exploration of the GALS interfaces, and power and EMI reduction possibilities. To properly address this, in parallel, the purely synchronous version of the same system will be implemented.

Until now there have been only relatively few complex GALS demonstrators developed, such as the WLAN baseband processor from IHP (0.25um CMOS) [KRS06], the FAUST processor developed at LETI (0.13um CMOS) [LAT07] and circuits fabricated in GALS group at ETHZ (0.25um CMOS) [GUR06]. Even though this GALS system implementation is not the ultimate goal of the GALAXY project, we regard the GALS system implementation in this project a necessary vehicle to evaluate achieved improvements in the GALS design flow and architecture. However, the planned GALAXY GALS chip implementation will definitely have many competitive features in comparison with the previous approaches.

Since wireless communication is an application domain which is gaining more and more popularity and which poses significant technical challenges to system designers (performance- and power-wise), and since several group members have prior experience in this field, we consider target platforms from this field. Our main candidate, on which this research is based, is an accelerator for an OFDM baseband processor with data rates up to 1 Gbps for communication systems in 60 GHz range which is currently under development in IHP [GRA07b]. Currently, the full processor integrates several Viterbi decoders, 256-FFT and IFFT, many complex interleavers/deinterleavers. The top level block diagram is depicted in Figure 1. The processor consists of transmitter and receiver, each of them having up to 12 parallel streams and multiple complex blocks.

Previously this processor was implemented in FPGA, achieving 1 Gbps throughput, with 100MHz operating frequency and 10 streams. The complexity of FPGA realization, and some specific blocks is given in Table 1.

The goal is to increase the throughput with ASIC solution, reduce the power, and area. New standards (802.15.4c) target throughputs of 5 Gbps (up to 10 Gbps). Standards have different modes with different decoding techniques, for example, using deinterleaving together with Viterbi decoding, using LDPC (and/or Reed-Solomon) decoding, combining deinterleaving together with LDPC (and/or RS) decoding. In order to increase throughput we could add more streams (instances) and we could use the higher operating frequency.

The system should support different data-rates, frame-sizes, streaming, MIMO, etc. It is possible to focus on particular decoding technique, and limited set of the datarates and optimize the system for such implementation. In this case point-to-point GALS would be probably the best solution.

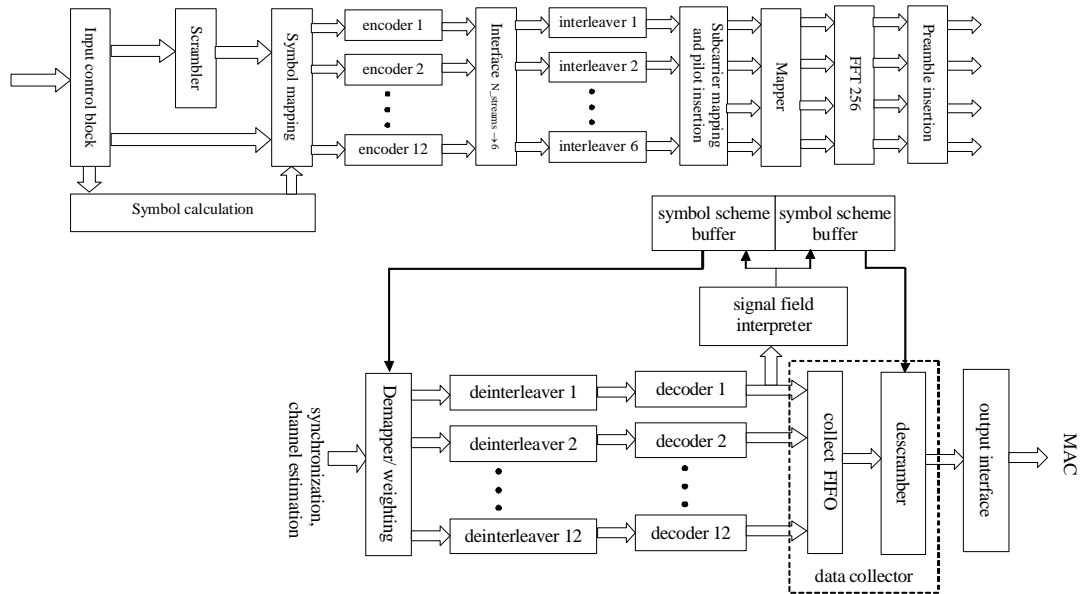


Figure 2: Top level block diagram of OFDM processor

Table 1: Complexity of FPGA realization of OFDM processor

| block | gates (Mil.) | slices | flip-flops | 4-input LUT | BRAM | MULT |
|-------|--------------|--------|------------|-------------|------|------|
| Tx | 7.8 | 12430 | 18069 | 16375 | 111 | 41 |
| Rx | 15.5 | 29142 | 27450 | 46632 | 225 | 56 |

The planned throughput is 3 Gbps with 12 streams and 250 MHz operating frequency (which is limited only by I/O pad frequency). The preliminary results (core logic area) for 40nm Infinition library:

- Transmitter 1803013 μm^2
- Rx_frontend 3209173 μm^2
- Rx_datapath 5254238 μm^2

This is in total 10.3 mm^2 and the area is dominated by RAM structures (70%). These are reported results for the synchronous version of the design. The amount of area is quite large and already over the specification dictated by technology fabrication. The design is currently being GALSified. GALSification will bring additional increase in area.



3.2 BENEFITS AND DIFFICULTIES IN IMPLEMENTING NoC IN DESIGN

We are exploring the effectiveness of GALS approaches to deep sub-micron processes, and therefore, our approach will provide completely different additional merits to the GALS subject. The planned application in our case seems to be very challenging because of its ad-hoc architecture with very demanding processing requirements. There are several difficulties in designing such a large and complex design as proposed OFDM baseband processor:

- The desired throughput is high and difficult to achieve.
- The occupied area is very large even in such an advanced technology as 40nm.
- Performing GALS partitioning of the complex synchronous design is block and function dependent.
- Despite maturing of the GALS design flow on course of this project, this is the largest and most complex GALS design reported up to date and is very hard to predict all the possible difficulties in the backend and simulation phase of the designing process.

Majority of GALS system implemented so far ('99-'06) are using point-to-point communication. Classical GALS implementation in GALAXY project is challenging It will have significant merits, due to the interesting application field.

However, there are also at least three GALS NoC implementation:

- NEXUS chip [NEX04], and two LETI implementation,
- FAUST chip [BEI06], and
- ALPIN chip [BEI08].

In the last few years the community is mainly focused on GALS NoC system architectures and implementations. Main "competitor", LETI, is focused on GALS NoC chip for wireless communications. The real crossbenchmarking between the GALS NoC and sync NoC solution has never been made. LETI chips, although having the telecom functionality, were never competitive with the state-of-the-art telecom solutions. With the proper application and implementation, the impact of GALS NoC demonstrator can be extremely high.

In the design like baseband processor the main challenge is to generate general purpose decoding accelerator that covers different techniques/modes/datarates in the system. This architecture is difficult to implement using classical topologies. The point-to-point topology is hard to implement due to the reconfigurable system structure. The bus topology is not useful due to the extreme throughput and wiring. Promising option for reconfigurable/programmable decoding hardware accelerator can be NoC interconnect. With NoC, we can achieve full reconfigurability, and high data throughput. The main parameter is decoding throughput of such system. Achieving decoding throughput of 10 Gbps for several configurations and modes would have significant impact. For this NoC resources with high throughput and QoS are necessary.

All the potential benefits of NoC implementation should be considered in with respect to GALS implementation. For example GALSification of NoC recourses also significantly reduces their power consumption, which is one of the main drawbacks of conventional NoC design.



Since it is very hard to provide large and complex GALS design that is in the same time fully reconfigurable the focus has to be on particular decoding technique, and limited set of the datarates and system should be optimized accordingly. In this case point-to-point GALS would be probably the best solution since NoC implementation has disadvantages, namely:

- Point-to-point processor architecture is optimal, and NoC brings no benefit.
- Achieving planned throughput implies large NoC structures that should provide sufficient QoS.
- Additional overhead introduced by NoC architecture is rather large and implies the final chip area much larger then planned, increasing the necessary cost planned for chip fabrication.
- Additional power consumption due to NoC is not negligible even though it is minimized with GALSification.
- In the process of GALSification, optimal partition size for GALS modules is dictated by the functionality of the processor design and is suboptimal for typical blocks used in NoC, i.e. optimal GALS modules are smaller then typical cores used in NoC architectures.
- Even though GALSification of all NoC building blocks is done according to project agreement, integration of NoC on the large and complex system such as OFDM processor brings a lot of additional hazards. Moreover, optimal GALSification of switches is still an open problem.

Due all these drawbacks, and despite the benefits our suggestion is that in case of large and complex design such as OFDM processor as target hardware accelerator, implementation of NoC is *not* beneficial.

Despite this conclusion, we will briefly investigate an alternative target design in the following section and provide the definitive choice in the final discussions section.

3.3 AN ALTERNATIVE TARGET DESIGN

As discussed in the previous section, a large and complex design as the baseband processor is not an optimal design for application of NoC architecture. However, a smaller part of that design, namely deinterleaver/decoder part of the receiver offers a lot of potential reconfigurability is therefore suitable for NoC application. The basic structure is shown in Figure 3, and consists of deinterleaver, Viterbi, LDPC and Reed-Solomon decoders. As there are several potential modes of operation, a multi-point topology is an optimal one, and in the same time highly suitable for NoC application.

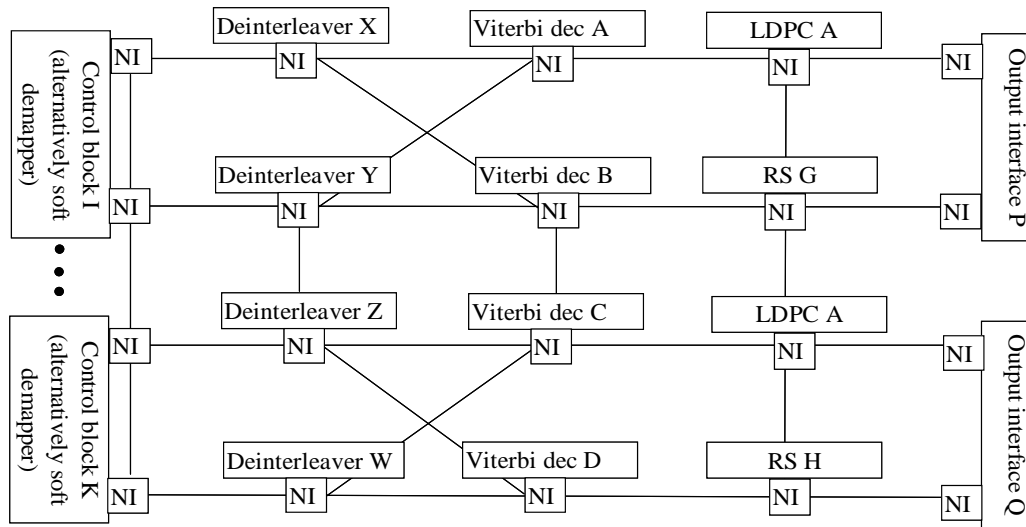


Figure 3: NoC optimal deinterleaver/decoder part of OFDM receiver

This system should be packet based. The size of data packets is variable and even the same block can in different modes generate different packet sizes. In our 1 Gbps demonstrator, data packets between demapper and deinterleaver, and deinterleaver and Viterbi decoder, are 60-540 Bytes. All packets will have destination address. Throughput of NoC node is defined by the data rate of the soft demapper. Usual number of soft bits is 5, and FEC redundancy is 2. With 2 coders in series complexity increase by 2. If we multiply this with the output data throughput (5-10Gbps), the desired throughput of the NoC node should be around 100-200 Gbps. This is a very large value difficult to achieve in today's state-of-the-art NoC architectures.

With more than one LDPC/RS decoder, the NoC node throughput may be reduced, maximally by the factor of the parallelization. However, parallelization introduces significant overhead not only in processing elements, but also in I/O interfaces and I/O pads, which are one of the limiting factors in the final chip design.

Regarding QoS, it would be good to guarantee that some average NoC data throughput can be reached. The transient delays can be tolerated.

IP cores are parametrical/reconfigurable VHDL cores. There are five different reconfigurable modes of operation:

- Convolution decoding involving deinterleaver and Viterbi decoders;
- LDPC decodings;
- Combined LDPC and Reed Solomon decoding;
- Deinterleaving with LDPC decoding; and
- Deinterleaving with LDPC and RS decoding.



GALAXY

GALS InterfAce for CompleX Digital
SYstem Integration

Confid. Level: Public
Date : 02/11/2009
Issue: 1

Currently the designs supports only simple token-flow interface (data valid in - data valid out). CP interface extension for all cores is needed. The IP cores will be equipped with input/output buffers to accommodate for delays in data transfer.

Regarding implementation of this design as the final target design, there are few implementation drawbacks, such as missing IP elements that need to be designed and NoC integration GALSification problems that need to be solved. However, the more serious drawback lays in the potential hazards during whole system integration and in potential impact that the smaller target design such as this one can have.



4 FINAL DISCUSION AND CONCLUSION

As discussed in the previous sections, two possible designs can be implemented as the target hardware accelerator, however only a smaller design (part of the receiver) is suitable for NoC integration. There are two significant advantages of larger design (OFDM processor) without NoC, despite the benefits of GALS NoC implementation and potential impact that this kind of demonstration can have:

- Since there is only one demonstrator target design in the project, the safer approach is to use structures that are already tested and proven.
- Since the main goal of the whole project is to demonstrate potential benefits of GALS designs, implementation of such a large GALS design that has never before been demonstrated will have the larger potential impact and significance in the research community (and industry) than smaller GALS NoC implementation that has already demonstrated implementations.