Chapter 17
Silicon Validation of GALS Methods and Architectures in a State-of-the-Art CMOS Process

Milos Krstic
IHP, Germany

Xin Fan
IHP, Germany

Eckhard Grass
IHP, Germany

Luca Benini
University of Bologna, Italy

M. R. Kakoei
University of Bologna, Italy

Davide Bertozzi
University of Ferrara, Italy

Christopher Heer
Intel Mobile Communications, Germany

Birgit Sanders
Intel Mobile Communications, Germany

Alessandro Strano
Intel Mobile Communications, Germany

Gabriele Miorandi
University of Ferrara, Italy

Alberto Ghiribaldi
University of Ferrara, Italy

ABSTRACT

The GALS methodology has been discussed for many years, but only a few relevant implementations in silicon have been done. This chapter describes the implementation and test of the Moonrake Chip – complex GALS demonstrator implemented in 40 nm CMOS technology. Two novel types of GALS interface circuits are validated: point-to-point pausable clocking GALS interfaces and GALS NoC interconnects. Point-to-point GALS interfaces are integrated within a complex OFDM baseband transmitter block, and for NoC switches special test structures are defined. This chapter discloses the full structure of the respective interfaces, the complete GALS system, as well as the design flow utilized to implement them on the chip. Moreover, the full set of measurement results are presented, including area, power, and EMI results. Significant benefits and robustness of our applied GALS methodology are shown. Finally, some outlook and vision of the future role of GALS are outlined.

DOI: 10.4018/978-1-4666-6034-2.ch017
INTRODUCTION

Globally Asynchronous Locally Synchronous (GALS) technology has been proposed many years ago as an alternative to the traditional synchronous paradigm for chip synchronization (Krstic, 2006). Although significant potential was reported by the academia, the GALS methodology has never taken off in the industry. However, the growing challenges, imposed by the unrelenting pace of technology scaling to the nanoscale regime, urge for an efficient and safe system-level integration methodology. Consequently, this section provides the overview of the recent implementation of a chip, named Moonrake, in advanced 40 nm CMOS process, aiming at the assessment of GALS technology for nanoscale designs.

The intention of this implementation was to evaluate GALS vs. standard synchronous technology on the same die, by fabricating synchronous and GALS counterparts of the same baseline designs, both in the point-to-point as well as in the network-on-chip (NoC) scenarios for on-chip communication. The two scenarios are very different, hence motivating the different choice of baseline designs for their analysis. In point-to-point communication, once an optimized GALS interface is selected, the focus is on the implications of redesigning an entire system around these links. With this respect, a state-of-the-art multi-million gate synchronous system, an OFDM baseband transmitter developed for a 60 GHz transceiver with a gigabit throughput as presented by Krstic in 2008 has been taken, and re-implemented it with GALS methodology, using the optimized interfaces for pausible (stoppable) clocking as defined by Fan in 2009. One major goal was to explore Electromagnetic Interference (EMI) and switching noise properties of GALS designs and special algorithms and circuits for noise reduction based on the GALS methodology, initially analyzed by Fan in 2010. Within the chip, the switching noise (and correspondingly EMI) is caused by simultaneous switching activity of the digital circuits and it can lead to various problems including ground bounce, power integrity, IR drop, substrate noise etc.

For on-chip networking applications, the communication landscape is more heterogeneous since it results from the interconnection of domains with different synchronization assumptions. Therefore, the focus was on the provision of flexible and cost-effective interfaces for arbitrary composability. In this direction, the novel synchronization interfaces presented by Strano (2010) and Ludovici (2010), aiming at low-area/power/latency overhead while preserving timing robustness, were integrated into NoC test structures exposing (and comparing) a range of flexible GALS solutions.

The main intentions of this chapter are as follows:

- The GALS partitioning criteria for a state-of-the-art OFDM transmitter is presented, highlighting the optimized asynchronous link crossing scheme and the partitioning granularity and strategy at the system level.
- The design flow followed for different GALS systems is illustrated: from pausible clocking to mesochronous synchronization to mixed-timing systems. Compatibility with mainstream standard cell libraries and design toolflows is discussed.
- The feasibility of GALS NoCs linking sub-systems with heterogeneous timing assumptions by means of area/power/latency optimized interfaces while preserving timing margins has been demonstrated.
- Synchronous and GALS counterparts of the same baseline designs (the OFDM transmitter and a NoC sub-set), implemented in the same demonstrator chip, have been compared in terms of area, pointing out counterintuitive benefits of the GALS design style.
- Finally, the test and measurement results of Moonrake chip are presented and analyzed, with the focus on EMI and power measure-