

Toward an Integrated Design Methodology for Fault-Tolerant, Multiple Clock/Voltage Integrated Systems*

Radu Marculescu, Diana Marculescu, Larry Pileggi

Department of Electrical and Computer Engineering

Center for Silicon System Implementation

Carnegie Mellon University

Pittsburgh, PA 15213

Email: {radum,dianam,pileggi}@ece.cmu.edu

Abstract - *This paper describes a communication-centric design methodology that addresses the fundamental challenges induced by the emergence of truly heterogeneous Systems-on-Chip (SoCs). For such systems, the globally asynchronous design paradigm seems to be the most promising (if not the only) solution for providing an underlying substrate for cost-effective and power efficient on-chip communication among diverse, mixed technology IPs. Additional challenges are related to reliability and error resilience of on-chip communication architectures. The proposed on-chip communication methodology targets all levels of abstraction, from circuit, to microarchitecture and system-level by seamlessly integrating solutions for robust and efficient globally asynchronous communication among diverse IPs.*

1. Introduction

As next generation electronic systems scale to higher levels of integration, we expect them to look very different from today's integrated circuits and systems in terms of both *complexity* and *heterogeneity*. As complexity increases, there is a corresponding increase in heterogeneity due to the expanded range of applications and the higher levels of design abstraction. Component heterogeneity relates to the very different functionality, performance and reliability constraints that are likely to be enforced for system components as the system size increases. Providing reliable communication among such blocks, therefore, is a potential showstopper for reaching higher scales of integration. Moreover, since power density and thermal considerations are already limiting what we can affordably integrate on a single die or within a single package, the communication problem is further complicated by the dynamic power control techniques that are required for any complex ICs.

This paper introduces the concept of *fault-tolerant communication* for *Multiple Clock/Voltage* (MCV) integrated systems. As CMOS technology continues toward the nanoscale domain, the circuit behavior becomes more susceptible to process parameter variation and noise disturbances, thereby necessitating some degree of fault-tolerant communication. In addition, due to power considerations, we cannot expect that all components will be able to run at full power concurrently.

Instead, further increases in scale of integration will be realized by voltage/frequency island-based design, whereby different blocks and regions operate at different supply voltages and frequencies, as controlled either dynamically or statically by the system-level constraints.

From a theoretical perspective, the optimal communication strategy for such ICs may be one based on Globally Asynchronous Locally Synchronous (GALS) communication. Our preliminary circuit-level work has shown some theoretical advantages to asynchronous global communication schemes for accommodating process variation [8]. In addition, our system-level analysis indicates that fault tolerance is most readily incorporated via asynchronous protocols. While these results demonstrate great promise, only a complete problem construction and analysis from circuit- to system-level can truly assess such a design methodology.

To this end, this paper introduces a design methodology that targets all levels of design abstraction, from system, to microarchitecture and circuit-level by seamlessly integrating solutions for achieving robust, error resilient and efficient globally asynchronous communication among diverse IPs. It is further expected that multiple clock and/or voltage islands will be considered to provide fine-grain dynamic power management. Such a design and modeling methodology addresses not only some of the short-term challenges related to IP-reuse and computation-communication separation of concerns, but also provides a robust solution for the long-term challenges imposed by the on-chip integration of mixed-technologies and mixed-design styles in next generation systems. To support the proposed methodology, two major issues are of interest:

- Fault-tolerant communication schemes and protocols suitable for on-chip implementation and tools for their performance evaluation.
- Design methodologies for asynchronous communication schemes that are applicable to multiple voltage/frequency islands, as well as their cycle-accurate modeling for GALS architectures.

We believe that having a unified design methodology for fault-tolerant communication of MCV integrated systems brings a significant long-term contribution by providing a natural solution to the issue of integrating, in a seamless manner, very diverse design styles and cutting-edge technologies.

* This research has been supported in part by Semiconductor Research Corporation under Contract No. 2004-HJ-1189.

2. Existing Communication Schemes for Integrated Systems

To increase the productivity of system design, semiconductor companies usually offer libraries of pre-designed and reusable blocks or “cores” and an architecture for easily interconnecting them via a set of buses with different performance characteristics. In the beginning, a system may include one processor core with a high-performance bus to access memory and a slower speed bus to communicate with peripherals. Experienced designers have little trouble selecting a configuration to meet their needs. Today and in the future, much larger integrated systems are possible, not only because multiple processor designs are becoming more common, but due to the growing use of IP core libraries in complex SoCs and high-performance processors. In both cases, cores with different capabilities, performance and power dissipation are integrated together on the same die. System designers are confronted with a much more difficult task of selecting the best cores and interconnect strategy to meet their overall performance, power and cost objectives.

The use of shared-buses simplifies the task of interconnecting cores and allows them to be reused widely. However, as integrated systems become larger and increasingly complex, it is clear that what is being sacrificed in power dissipation and performance, might no longer be acceptable. Large busses present a significant load on drive circuits, leading to significant delays and power dissipation. In addition, wide busses can cause major congestion problems during physical implementation. Other approaches to interconnection may offer significant advantages, such as point-to-point links which enable increased concurrency. Such links can be customized to meet specific performance needs with reduced power and chip size, but they would require changes in the IP cores.

To hide the implementation details and make the communication possible among heterogeneous modules that reside on the same chip, a scalable solution consists of a grid-like architecture, whereby various IP modules are placed on a grid of point-to-point communicating tiles. Such a regular structure is very attractive because it can offer well-controlled electrical parameters, which further enable high performance circuits by reducing latency and increasing the available bandwidth [3].

Defining communication protocols for such regular structures is not an easy matter, as the resources used in traditional data networks are not available on-chip. In addition, coping with various types of failures (induced either by manufacturing-driven defects, process and system parameter variation, soft errors or synchronization errors) necessitates a paradigm shift in the communication design methodology. Thus, a communication scheme for fault-tolerant, MCV systems must address the following issues:

- *On-chip diversity*, in order to allow for heterogeneity that combines existing technologies and design methodologies with the strength of newly emerging technologies.

- *Scalability*, in order to cope with complexity of future integrated systems.
- *Asynchronous communication* among multiple voltage/frequency islands residing on the same chip.
- *Resilience to complex failure mechanisms* for reducing the costs of design and verification and enabling the design of complex systems in the next technologies.

The trend of integrating multiple design styles or technologies will increase in importance for future generations of integrated systems. Despite the technical difficulties inherent to such an endeavor, this trend will ultimately lead to a completely new dimension in the design of tomorrow's electronic systems. Indeed, combining different architectures, design styles and underlying technologies is perhaps the most profound change in design practice to achieve the highest levels of flexibility and performance. We call this feature of next generation MCV systems *on-chip diversity* [5].

3. Emerging Solutions for MCV Systems

While the design of MCV systems allowing for on-chip diversity can be partially based on existing CAD tools and design practice for the IPs themselves, the same cannot be said about the on-chip communication mechanism. We believe that on-chip diversity is possible only by defining a completely new paradigm for on-chip communication. The foundation this new communication scheme relies on the Network-on-Chip (NoC) communication approach: routing packets (instead of wires) among heterogeneous modules or IPs that implement the desired functionality.

Since reuse of IP (digital or analog) is imperative for the future levels of integration we envision, it follows that the IP blocks must be conceived as being surrounded by a *communication wrapper* which abstracts, for the other IPs and the underlying communication architecture, the actual details of module's operation and provides a unified interface for on-chip communication. Therefore, the communication among wrappers would happen via packets, not wires. We consider next some design issues related to the NoC-based communication infrastructure.

3.1 Wrapper Design - Interfacing Bus-based IPs and On-chip Routers

An important problem involved in the utilization of the NoC-based communication is the reuse and migration of the IPs designed with traditional bus-based interfaces to this new architecture. More precisely, because most IPs are developed by major semiconductor companies with certain bus-interface capabilities (e.g., CoreConnect, AMBA, etc.), they cannot be directly reused within a NoC architecture because of the incompatibility between legacy bus interfaces and the on-chip router protocols. Re-implementation and customization of all these IPs may prove to be too expensive, if not infeasible and, more importantly, not appealing to core designers. Therefore, the bus-based legacy is a serious issue to consider with respect to the NoC approach.

We believe that a new design methodology for NoC wrappers which can efficiently interface the existing bus-based IPs and the NoC communication infrastructure is required. As shown in Figure 1, by creating an adaptation

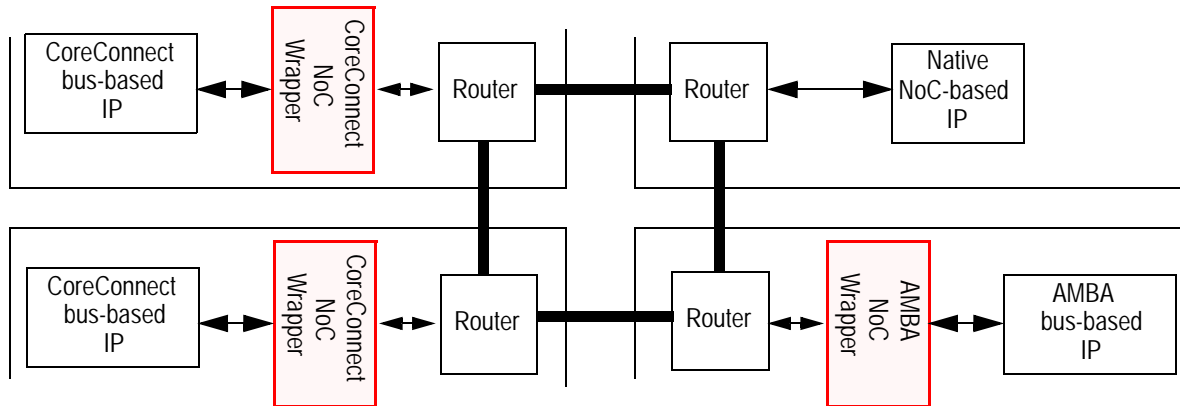


Figure 1. Application of NoC wrapper to reusing legacy bus-based IPs. The main idea is to use this wrapper instead of redesigning the cores from an already existing library.

layer between the interface of the bus-based IPs and the network-interface of the NoC, the legacy bus-based IPs can be reused in NoC designs with minimal or no modifications. The advantage of the NoC wrapper approach in reusing existing bus-based IPs is obvious. Instead of spending manpower to customize and modify legacy IPs to make them usable in NoCs, the NoC wrapper design appears to be a one-time investment. Specifically, for each standard bus interface, only one wrapper needs to be designed in order to enable the reuse of all the IPs designed to communicate using that bus interface. On the other hand, since only one wrapper is needed for each standard bus interface, it is possible to highly optimize and fine tune the wrapper design such that the performance degradation and area overhead incurred by the conversion logic of the wrapper can be kept at minimum. Thus, the main issues that become essential for the wrapper design problem are:

- *Specification of the NoC services/protocols* exported to the wrapper. For each bus interface under consideration, one needs to decide a subset of functions that need to be supported by NoC and derive efficient ways (in terms of area/performance/power) to support these functions.
- *Quality of Service (QoS)-based strategy* for on-chip routing which exploits the routing paths that are less affected by congestion or run-time errors [6]. Such a routing strategy can be motivated by differences generated by parameter variation during manufacturing, elevated levels of soft error rates during runtime, etc.

3.2 On-Chip Stochastic Communication

In terms of a supporting communication paradigm, we believe that *stochastic communication* must lie at the heart of the new design methodology for fault-tolerant MCVs. This would consist of using a simple *probabilistic broadcast* scheme for inter-node communication, similar to the randomized gossip protocols used in distributed databases. The motivation behind using a probabilistic rather than a deterministic approach is to provide built-in system-level fault tolerance that can better adapt to the context of heterogeneous systems. Indeed, for the MCVs we envision, malfunctions can only be characterized by stochastic models, as they are either non-deterministic in

nature or too complex to be described by simple failure models [7][2].

Stochastic communication belongs to the class of randomized broadcast primitives called *gossip algorithms*. The behavior of such a communication scheme is similar to the spreading of an epidemics across a large population (i.e., exponentially fast). Simply speaking, for a regular grid-based architecture where each IP on a tile represents a node in the network (Figure 2), if a node has a packet to send, it will forward the packet to a randomly chosen subset of the nodes in its neighborhood. This way, the packets are *diffused* from tile to tile. Every IP then selects from the set of received messages only those that have their own ID as destination [4].

A typical tile which implements a node in such a communication scheme is shown in Figure 2(a). Each IP core is wrapped in a unified communication interface that enables compatibility across different technologies or design styles. On the four edges of the tile there are buffers that hold messages sent and received by the IP. A Cyclic Redundancy Check (CRC) decoding circuit checks all the received messages and when an error is discovered, the message is discarded before being fed into the IP. The tile keeps a list of messages that have to be sent to the output buffer².

Using stochastic communication in a grid-like architecture represents, however, a huge departure from the classical deterministic, bus-based communication. As we know, bus-based communication is very efficient when only a few communicating IPs are connected or when the application requires a significant number of message broadcasts. On the other hand, stochastically communicating NoCs are very scalable, can include a large number of IPs and their performance does not degrade significantly under the influence of on-chip failures. However, most SoC applications do not have a uniform structure; thus, in order to provide a solution for on-chip diversity, a combination of these structures would be more appropriate (see Figure 2). Such a structure supports heterogeneity through the use of communicating

²The functionality of the random number generator (RND) module is related to the on-chip stochastic communication mechanism.

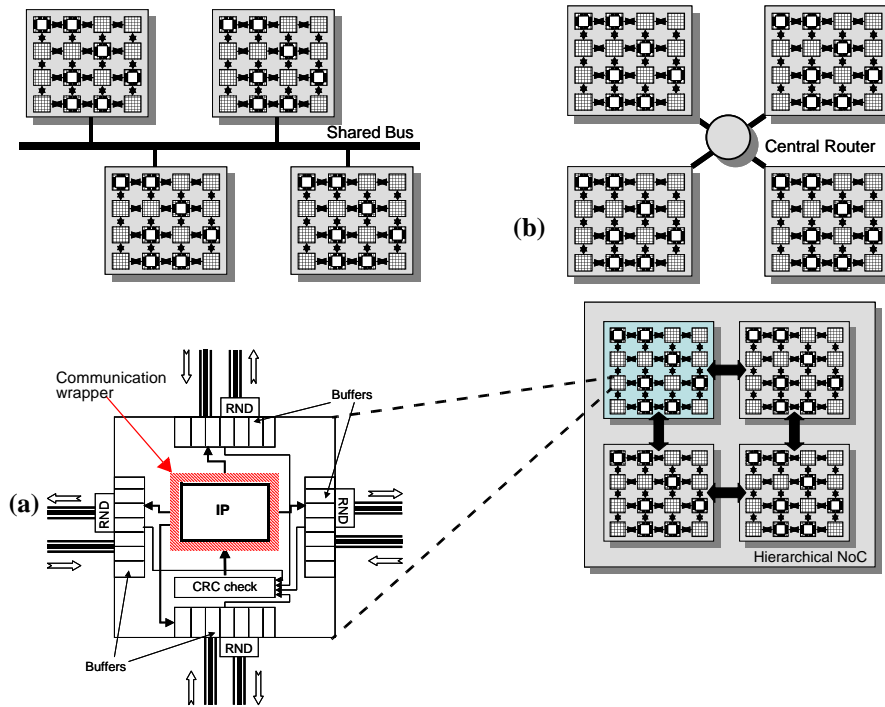


Figure 2. On-chip diversity: (a) Tile structure; (b) Possible heterogeneous communication structures.

islands that can be connected with a traditional bus or can be assembled in a hierarchy, according to the requirements of the application (see Figure 2(b)). This enables the design of truly heterogeneous systems where properties of the different communication architectures can be combined hierarchically to obtain the best of all possible worlds. Assuming that such a tile-based design methodology is used for MCVs, the overall design methodology becomes possible only by being supported by several key mechanisms, as shown in Figure 3.

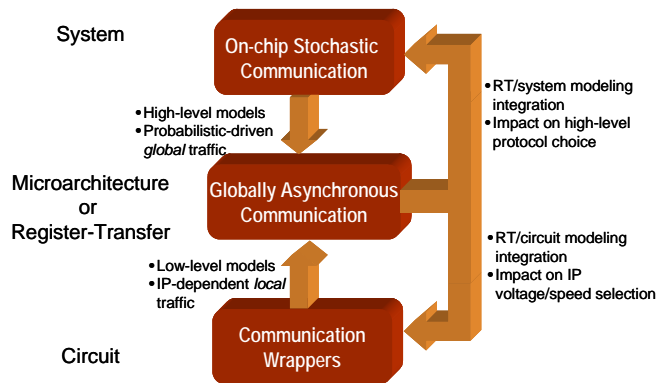


Figure 3. The proposed design methodology for on-chip communication of multiple clock/voltage integrated systems.

As shown in the figure, the proposed communication methodology relies on tight integration between circuit-, microarchitecture- and system-level models for taking the right decisions for system-level protocol selection, or circuit-level enablers for globally asynchronous

communication or fine-grain power management. To make such a communication-centric design methodology a reality, several things have to happen: First, at circuit-level, we need a completely new communication infrastructure that can accommodate probabilistic behavior besides just the deterministic one. This can be based on new design techniques for communication wrappers and routers that enable a wide spectrum of adaptability to various application requirements, ranging from performance-oriented routing, to QoS-based and stochastic routing for applications where fault-tolerance is the metric of interest.

Moving up to the highest level of abstraction, a change in the communication paradigm can facilitate an affordable, scalable and fault-tolerant communication scheme. Indeed, stochastic communication has not only very good latency, but also excellent fault-tolerance [4]. As messages are transmitted multiple times in the network, this redundancy can be exploited to protect communication against failures. This is very useful in a heterogeneous environment when complex failure mechanisms can hinder the propagation of correct messages between different communication nodes [2].

At an intermediate level, integrating different technologies and design styles is best accomplished via a globally asynchronous communication mechanism. Our analyses have considered communication mechanisms ranging from low latency asynchronous FIFOs, or asynchronous wrappers around asynchronous blocks, to pausable clocks mechanisms that pause the passive phase of the producer/consumer clock to avoid timing violations. We next consider some of these design issues.

3.3 Globally Asynchronous On-Chip Communication

Since communication seemingly becomes the major source of power consumption, the application must be partitioned such that it can perform most of the operations locally, without the need to communicate across the entire system. In addition, IP cores should be placed such that the communication cost is minimized. Preliminary system-level analysis has shown that properly dividing the network into several separate communicating islands can potentially lead to a drastic reduction in the volume of communication and therefore, to significant energy savings. However, some of our circuit-level analyses indicate that the theoretical benefits of asynchronous busses can be deceiving [8].

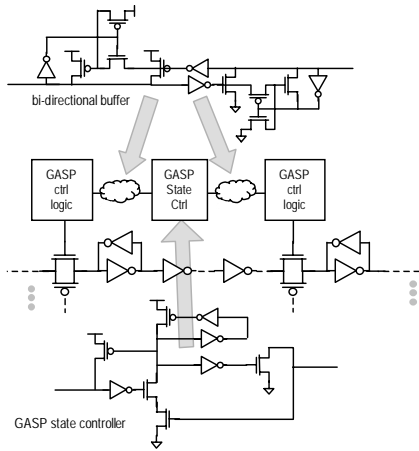


Figure 4. GasP-based asynchronous bus design.

For example, our circuit level design and experiments using a GasP-based [10] asynchronous bus (Figure 4) demonstrate that the asynchronous control power is larger than the local synchronous clock power by an amount that exceeds the savings in latch power along the data bits of the bus for small bus sizes. Only as we increase the number of bits and share the asynchronous control line does the total asynchronous power become marginally better. Based on these preliminary results, asynchronous communication for *global* interconnect signals seems to provide little advantage for signal transmission. Thus, in the context of globally asynchronous communication, a

tile-based grid architecture, relying heavily on *local* communication, seems to be the architecture of choice from the perspective of communication cost in power.

When we consider mixed-clock domains, or frequency islands, however, our same circuit-level analyses suggest that an asynchronous bus scheme between islands can provide several advantages for local communication in terms of both performance and power. A potentially viable communication scheme would be one based on mixed-clock/asynchronous buffers. For such systems, in the context of using multiple voltage islands, the total power consumed by the system can be managed on a fine-grain basis by monitoring incoming and outgoing traffic to the buffer, as we describe in the next section.

It is worth noting that none of the existing mixed-clock communication paradigms consider the case of producer and consumer data that may not only be part of different clock domains, but also of different voltage domains. Mixed-clock/mixed-voltage interface design becomes thus an important ingredient for MCV systems that rely on various voltage/frequency islands independently optimized for speed and energy efficiency. Such mixed-clock/mixed voltage interfaces become intrinsic components for IP-router or inter-router communication, as in the case of voltage island-based design not only IPs can operate at different speeds or voltages, but the same can be assumed for the routers.

Since communication schemes based on mixed-clock/asynchronous buffers can not only do the job of synchronizing data transfers, but also replace repeaters that are used for dealing with increased wire delays [1], they are also good candidates for providing mixed-clock/mixed-voltage interfaces between IPs and routers or between different routers in case of multi-hop routes. As shown in Figure 5, such mixed-clock/mixed-voltage buffers or FIFOs are envisioned to facilitate communication among IPs and routers even if they are operating at different speeds or voltages.

One of the natural applications of this design paradigm is the design of *mixed-clock issue queues* [9] for high-end processors which have the potential of hiding the additional latency associated to using synchronizers for dealing with timing violations. Early results for such a design [11] suggest that performance may not need to be sacrificed when a globally asynchronous design style is used.

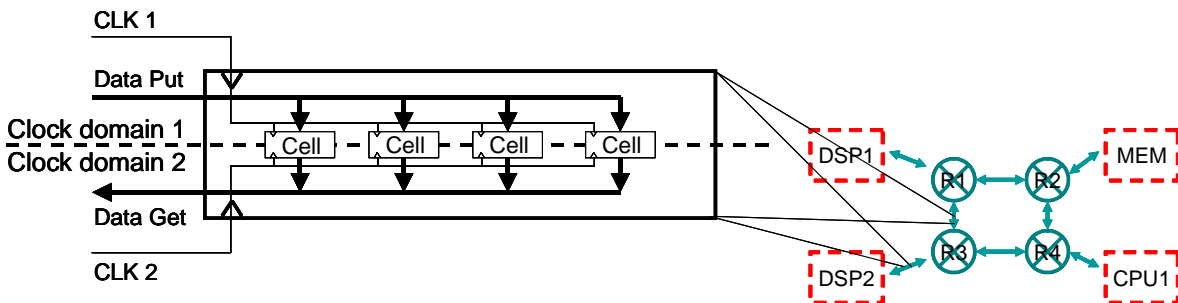


Figure 5. A mixed-clock(/mixed-voltage) buffer used for communication between IPs and adjacent routers, or between routers in a tile-based architecture. To support bidirectional channels, two buffers would be needed for each such channel.

4. Energy Efficiency for MCV Systems

One of the main motivations for the use of a MCV design style is the ability to better control power consumption via fine-grain, local power management. For instance, when using stochastic communication, application components that do not require a high level of performance can be mapped onto a stochastically communicating islands running at a lower voltage and possibly at a lower frequency. Alternatively, for such systems, communication mechanisms for supporting workload-driven adaptation for optimal energy efficiency under performance constraints must also be considered. However, schemes that use the traffic information in the communication buffers to adapt the supply voltage of various components are not scalable and do not account for complex communication patterns that may occur in MCV integrated systems. For example, in a hierarchical organization such as the one presented in Figure 2, the decision of using a dynamically adaptable supply voltage for one of the IPs cannot be taken in isolation as it may affect and be affected by the overall traffic and communication patterns. In addition, some IPs may tolerate changes in the communication traffic, while others may not. Therefore, in the context of using multiple voltage islands, the total power consumed by the system can be managed on a fine-grain basis, but care must be taken as to what are the best local decisions that may lead to overall energy efficient operating points.

Typical integrated systems are designed to operate at a maximum throughput rate and thus, a decrease in the input sampling rate would lead to a non-zero available slack. An algorithm which optimally distributes this available slack between the IPs and routers would lead to overall energy savings maintaining the throughput as governed by the input sampling rate. It is our belief that the best way to address such challenges is to enable accurate modeling of the globally asynchronous communication mechanisms based on lower levels (such as communication wrappers) that offer information about *locally generated traffic* patterns, as well as higher levels of abstraction that provide information about *globally generated traffic* patterns due to specific routing decisions. To this end, *cycle-accurate modeling* at the microarchitectural level for mixed-clock/mixed-voltage communication schemes seems to be an essential ingredient for the proposed design methodology. While efforts exist that deal with microarchitectural modeling of multiple clock or GALS designs, none of them includes accurate accounting for the extra circuitry needed for adaptive voltage and speed scaling.

In addition, models that accurately include the effects of clocking a smaller area on the overall system performance have not been considered so far. It is our belief that a GALS design style for MCV systems lends itself to lower process and system parameter variation, and thus lower clock skew and higher clock speed. However, a detailed accounting of such effects at microarchitectural levels or above has not been addressed so far and would represent a major achievement. Furthermore, since asynchronous/mixed-clocking buffers

with IP and router voltage scaling capabilities have a behavior that is strongly dependent on the incoming/outgoing traffic, the information about locally or globally generated traffic should be used for defining the proper power management mechanisms for multiple voltage/multiple frequency island designs.

5. Conclusion

This paper has presented an integrated design methodology for MCV systems as a possible solution for upcoming challenges imposed by increased integration in conjunction with the use of multiple design styles or technologies. Essential ingredients for this methodology include fault-tolerant communication and reliance on a globally asynchronous, locally synchronous design style. In addition to increased error resilience and scalability, such a system organization seems to also offer the advantage of better power efficiency through the use of fine grain, local voltage or speed scaling in a static or dynamic manner.

Acknowledgements: The authors would like to acknowledge their graduate students contributions, some of which have been included in this paper.

References

- [1] T. Chelcea and S. Nowick, "Robust Interfaces for Mixed-Timing Systems with Application to Latency-Insensitive Protocols," in Proc. ACM/IEEE Design Automation Conference, June 2001.
- [2] C. Constantinescu, "Impact of Deep Submicron Technology on Dependability of VLSI Circuits," in Proc. IEEE Intl. Conference on Dependable Systems and Networks, June 2002.
- [3] W. Dally and B. Towles, "Route packets, not wires: On-chip interconnection networks," in Proc. ACM/IEEE Design Automation Conference, June 2001.
- [4] T. Dumitras and R. Marculescu, "On-Chip Stochastic Communication," in Proc. IEEE/ACM Design, Automation and Test Conference in Europe, Munich, March 2003.
- [5] T. Dumitras, S. Kerner and R. Marculescu, "Enabling On-Chip Diversity Through Architectural Communication Design," in Proc. ACM/IEEE Asia-Pacific Design Automation Conference, Tokyo, Japan, Jan. 2004.
- [6] J. Hu and R. Marculescu, "DyAD: Smart Routing for Networks-on-Chip," in Proc. ACM/IEEE Design Automation Conference, June 2004.
- [7] T. Karnik, S. Borkar and V. De, "Sub-90nm Technologies-- Challenges and Opportunities for CAD," in ACM/IEEE Proc. Intl. Conf. on Computer-Aided Design, Nov. 2002.
- [8] E. Malley, A. Salinas, K. Ismail and L. Pileggi, "Power Comparison of Throughput Optimized IC Busses," in Proc. IEEE Symposium on VLSI, Feb., 2003.
- [9] V.S.P. Rapaka and D. Marculescu, "A Mixed-Clock Issue Queue Design for Globally Asynchronous, Locally Synchronous Processor Cores," in Proc. ACM Intl. Symposium on Low Power Electronics and Design, Aug. 2003.
- [10] I. Sutherland and S. Fairbanks, "GasP: A Minimal FIFO Control," in Proc. Symposium on Asynchronous Circuits and Systems, 2001.
- [11] E. Talpes, V.S.P. Rapaka and D. Marculescu, "Mixed-Clock Issue Queue Design for Energy Aware, High-Performance Cores," in Proc. ACM/IEEE Asian-South Pacific Design Automation Conference, Jan.2004.