

Technology-Driven Limits on DVFS Controllability of Multiple Voltage-Frequency Island Designs: A System-Level Perspective*

Siddharth Garg Diana Marculescu Radu Marculescu
Dept. of ECE
Carnegie-Mellon University
{sgarg1,dianam,radum}@ece.cmu.edu

Umit Ogras
Strategic CAD Labs
Intel Corp.
umit.y.ogras@intel.com

ABSTRACT

In this paper, we consider the case of network-on-chip (NoC) based multiple-processor systems-on-chip (MPSoCs) implemented using multiple voltage and frequency islands (VFIs) that rely on fine-grained dynamic voltage and frequency scaling (DVFS) for run-time control of the system power dissipation. Specifically, we present a framework to compute theoretical bounds on the performance of DVFS controllers for such systems under the impact of three important technology driven constraints: (i) reliability and temperature driven upper limits on the maximum supply voltage; (ii) inductive noise driven constraints on the maximum rate of change of voltage/frequency; and (iii) increasing manufacturing process variations. Our experimental results show that, for the benchmarks considered, any DVFS control algorithm will lose up to 87% performance, measured in terms of the number of steps required to reach a reference steady state, in the presence of maximum frequency and maximum frequency increment constraints. In addition, increasing process variations can lead to up to 60% of fabricated chips being unable to meet the specified DVFS control specifications, irrespective of the DVFS algorithm used.

Categories and Subject Descriptors

B.7.2 [Hardware]: Integrated Circuits—*Design Aids*

General Terms

Performance

Keywords

Networks-on-chip, power management, performance bounds

1. INTRODUCTION

With increased levels of integration in scaled technologies, novel on-chip communication architectures that use a Network-on-Chip approach have emerged as a scalable alternative to traditional bus-based or point-to-point communication solutions. Furthermore, due to increased power density and energy consumption, NoCs implemented using a multiple Voltage Frequency Island design style have become an attractive alternative to single-clock, single-voltage designs [2, 7]. Each island in a VFI system is locally clocked and has an independent voltage supply, while inter-island communication is orchestrated via mixed-clock, mixed-voltage FI-FOs. The power savings result from the fact that the voltage of each island can be independently tuned to minimize the system power dissipation under performance constraints.

To cope with *run-time* variations in the workload or power characteristics of VFI systems, the voltage and frequency of each island can be dynamically scaled to exploit the slack in the application

to reduce the power dissipation under a performance budget [6]. Not surprisingly, designing appropriate dynamic voltage and frequency scaling (DVFS) control algorithms for run-time control of VFI systems is a matter of great importance. While this problem has been addressed before by a number of authors [7, 9, 6], no attention has been given to analyzing the *fundamental limits* of the capabilities of DVFS controllers for multiple VFI systems. Starting from these overarching ideas, in this paper, we specifically focus on three technology driven constraints that we believe have the most impact on DVFS controller characteristics: (1) reliability and temperature constrained upper-limits on the maximum voltage and frequency at which any VFI can operate; (2) inductive noise driven limits on the maximum rate of change of voltage and frequency; and (3) the impact of manufacturing process variations. We note that each of these factors is becoming increasingly important with technology scaling.

Given the broad range of proposed DVFS control algorithms proposed in literature, we believe that it is insufficient to merely analyze the performance limits of a specific control strategy. The only assumption we make, which is common to many of the DVFS controllers proposed in literature, is that the goal of the control algorithm is to ensure that the occupancies of a pre-defined set of mixed-clock, mixed-voltage queues in the NoC are controlled to remain at pre-specified reference values¹. We then define the performance of a controller to be its ability to bring the queues, starting from an arbitrary initial state, back to their reference utilizations in a *desired but fixed number of control intervals*. Given the technology constraints, our framework is then able to provide a *theoretical guarantee* on the existence of a controller that can meet this specification.

2. RELATED WORK AND PAPER CONTRIBUTIONS

Power management of multiple VFI-based MPSoCs has been the subject of extensive prior research. [6] presents a Lagrange optimization based approach to perform DVFS in multiple VFI systems, while in [9], the authors propose a PID DVFS controller to set the occupancies of the queues in a multiple clock-domain processor to reference values. Finally, [7] presents a state-space model of the queue occupancies in an NoC with multiple VFIs and proposes a formal feedback control algorithm to control the queues based on the state-space model. We note that, compared to prior work, we focus on the fundamental limits of controllability of DVFS enabled multiple VFI systems and not on a specific control algorithm. Our results are, therefore, equally applicable to any of the control techniques proposed before.

As compared to previous work we make the following novel contributions: (i) we propose a computationally efficient framework to analyze the impact of three major technology-driven constraints on the performance of DVFS control algorithms for multiple VFI networks-on-chip; and (ii) starting from a formal state-space representation of the queues in an NoC, we provide theoretical bounds on the performance capabilities of *any* DVFS control algorithm (i.e., the proposed bounds are *DVFS control algorithm agnostic*).

¹While we explicitly control the NoC queues in this work, it can easily be applied to other systems with queue-based communication, such as point-to-point queues or even software queues.

*This research was funded in part by SRC Grant 2008-HJ-1800

3. PRELIMINARIES AND ASSUMPTIONS

We begin by briefly reviewing the state-space modeled developed in [7] to model the controlled queues in a multiple VFI system. We start with a design with N interface queues and M VFIs. An example of such a system is shown in Figure 1, where $M = 3$ and $N = 2$. Furthermore, without any loss of generality, we assume that the system is controlled at discrete intervals of time, i.e., the k^{th} control interval is the time period $[kT, (k+1)T]$, where T is the length of a control interval.

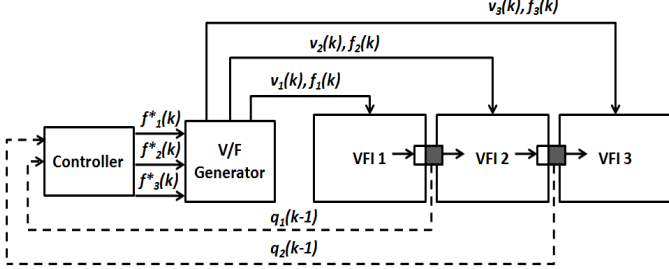


Figure 1: VFI system with three islands and two queues.

The following notation can now be defined:

- The vector $Q(k) \in \mathbb{R}^N = [q_1(k), q_2(k), \dots, q_N(k)]$ represents the queue occupancies in the k^{th} control interval.
- The vector $F(k) \in \mathbb{R}^M = [f_1(k), f_2(k), \dots, f_M(k)]$ represents the frequency of each VFI in the k^{th} control interval.
- λ_i and μ_i ($i \in [1, N]$) represent the average arrival and service rate of queue i , respectively. These parameters are workload dependent and can be obtained by probabilistic workload characterization.
- The system matrix $B \in \mathbb{R}^{M \times N}$ is defined such that the $(i, j)^{th}$ entry of B is the rate of write (read) operations at the input (output) of the i^{th} queue due to the activity in the j^{th} VFI. We refer the reader to [7] for a detailed example on how to construct the system matrix.

The state-space equation that represents the queue dynamics can now simply be written as [7]:

$$Q(k+1) = Q(k) + TBF(k) \quad (1)$$

The key observation is that given the applied frequency vector $F(k)$ as a function of time, this equation describes completely the evolution of queue occupancies in the system. In this equation, we have implicitly assumed that the occupancy feedback from the queues is available to the centralized controller within a single control interval - this is a reasonable assumption, since, in practice, a control interval is of the order of micro- to milli-seconds [7].

As shown in Figure 1, we also introduce an additional vector $F^*(k) = [f_1^*(k), f_2^*(k), \dots, f_M^*(k)]$, which represents the *desired* control frequency values at control interval k . For a perfect system, $F^*(k) = F(k)$, i.e., the desired and applied control frequencies are the same. However, due to the technology driven constraints, the applied frequencies may deviate from the frequencies desired by the control, for example if there is a limit on the maximum frequency at which a VFI can be operated. The technology driven deviations between the desired and actual frequency will be explained in greater detail in the next section.

4. LIMITS ON DVFS CONTROL

We now present the proposed framework to analyze the limits of performance of DVFS control strategies in the presence of technology driven constraints. To describe more specifically what we mean by performance, we define $Q_{ref} \in \mathbb{R}^N$ to be the desired *reference queue occupancies* that have been set by the designer. Furthermore, we assume that as a performance specification, the designer also sets a *limit*, J , that specifies the *maximum* number of control intervals that the control algorithm should take to bring

the queues back from an arbitrary starting vector of queue occupancies, $Q(0)$, back to their reference occupancy values. Note that the time index 0 here refers to a time instant at which the queue occupancies deviate from their reference values due to workload variations, and not to the start-up state of the system.

Given this terminology, using Equation 1, we can write the queue occupancies at the J^{th} control interval as [4]:

$$Q(J) = Q(0) + TB \sum_{k=0}^{J-1} F(k) \quad (2)$$

Since we want $Q(J) = Q_{ref}$, we can write:

$$(TB) \sum_{k=0}^{J-1} F(k) = (Q_{ref} - Q(0)) \quad (3)$$

4.1 Limits on Maximum Frequency

In a practical scenario, reliability concerns and peak thermal constraints impose an **upper limit** on the frequencies at which the VFIs can be clocked. For now, let us assume that each VFI in the system has a maximum frequency constraint f_{MAX}^i ($i \in [1, M]$). Therefore, we can write:

$$f_i(k) = \min(f_{MAX}^i, f_i^*(k)) \quad \forall i \in [1, M] \quad (4)$$

Consequently, the system can be returned to its required state Q_{ref} in at most J steps *if and only if* the following system of linear equations has a feasible solution:

$$(TB) \sum_{k=0}^{J-1} F(k) = (Q_{ref} - Q(0)) \quad (5)$$

$$0 \leq f_i(k) \leq f_{MAX}^i \quad \forall k \in [0, J-1], \forall i \in [1, M] \quad (6)$$

Note that this technique only works for a *specific* initial vector of queue occupancies $Q(0)$; for example, $Q(0)$ may represent an initial condition in which all the queues in the system are full. However, we would like the system to be controllable in J time steps for a *set* of initial conditions, denoted by R_Q .

Let us assume that the set of initial conditions for which we want to ensure controllability is described as follows: $R_Q = \{Q(0) : A_Q Q(0) \leq B_Q\}$, where $A_Q \in \mathbb{R}^{P \times N}$ and $B_Q \in \mathbb{R}^P$. Clearly, the set R_Q represents a bounded *closed convex polyhedron* in $\mathbb{R}^N$. We will now show that to ensure controllability for all points in R_Q , it is sufficient to show controllability for each vertex of R_Q . In particular, without any loss of generality, we assume that R_Q has V vertices given by $\{Q^1(0), Q^2(0), \dots, Q^V(0)\}$. First, due to the Krein-Milman theorem [8], we get the following result ²:

LEMMA 1. *Any $Q(0) \in R_Q$ can be written as a convex combination of the vertices of R_Q , i.e., $\exists \{\alpha_1, \alpha_2, \dots, \alpha_V\} \in \mathbb{R}^N$ s.t. $\sum_{i=1}^V \alpha_i = 1$ and $Q(0) = \sum_{i=1}^V \alpha_i Q^i(0)$.*

LEMMA 2. *The set of all $Q(0)$ for which Equation 5 and Equation 6 admit a feasible solution is convex.*

Finally, based on Lemma 1 and Lemma 2, we can show that:

THEOREM 1. *Equation 5 and Equation 6 have feasible solutions $\forall Q(0) \in R_Q$ if and only if they have feasible solutions $\forall Q(0) \in \{Q^1(0), Q^2(0), \dots, Q^V(0)\}$.*

Significance Theorem 1 establishes necessary and sufficient conditions to *efficiently* verify the ability of a DVFS controller to bring the system back to its reference state, Q_{ref} , in J control intervals starting from a large set of initial states, R_Q , without having to independently verify that *each* initial state in R_Q can be brought back to the reference state; instead, it is sufficient to verify the controllability for only the set of initial states that form the vertices of R_Q . This significantly reduces the computational cost of the proposed framework.

²For brevity, all results in the paper are stated without proof. Detailed proofs can be found in [1].

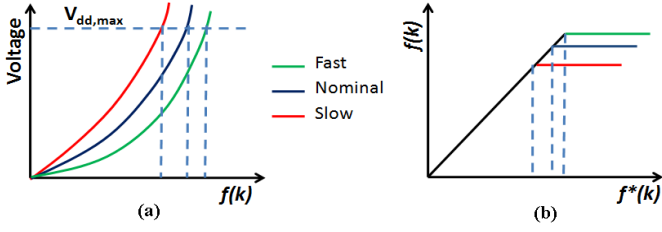


Figure 2: Process variation impact on DVFS controller.

In practice, the region of initial states R_Q will depend on the behavior of the workload, for example, a bursty read or a bursty write. While it is possible to obtain R_Q from extensive simulations of real workloads, R_Q can be defined conservatively as follows: $R_Q = \{Q(0) : 0 \leq q_i(0) \leq q_{MAX}^i, \forall i \in [1, N]\}$, where q_{MAX}^i is the physical queue length of the i^{th} queue in the system.

Finally, we note that minimum frequency constraints are not considered in this work because we assume that the frequency of each VFI can safely be reduced to zero, for example, using clock gating. However, minimum frequency constraints can easily be included in the framework if required.

4.2 Frequency Increment Constraints

A major consideration for the design of systems that support dynamic voltage and frequency scaling is the resulting inductive noise in the power delivery network due to sudden changes in the power dissipation and current requirement of the system. While there exist various circuit-level solutions to the inductive noise problem, it may be necessary to additionally constrain the maximum frequency increment within a control interval to minimize inductive noise. Furthermore, the limited slew rate of off-chip voltage controllers also constrains the maximum frequency change possible in a control interval [3]. We note that while there exist techniques to scale frequency on the order of a few clock cycles, both voltage and frequency scaling are required to obtain maximum energy efficiency from the system.

Frequency increment constraints can be modeled in the proposed framework as follows:

$$|f_i(k+1) - f_i(k)| \leq f_{step}^i \quad \forall i \in [1, M], \forall k \in [0, J-1] \quad (7)$$

where f_{step}^i is the maximum frequency increment allowed in the frequency of VFI i . Equation 7 can further be expanded as linear constraints as follows:

$$f_i(k+1) - f_i(k) \leq f_{step}^i \quad \forall i \in [1, M], \forall k \in [0, J-1] \quad (8)$$

$$-f_i(k+1) + f_i(k) \leq f_{step}^i \quad \forall i \in [1, M], \forall k \in [0, J-1] \quad (9)$$

Together with Equations 5 and 6, Equations 8 and 9 define a linear program that can be used to determine the existence of a time-optimal control strategy.

Finally, we note that for Theorem 1 to hold, we need to ensure that Lemma 2 is valid with the additional constraints introduced by Equation 7. We show that this is indeed the case.

LEMMA 3. *The set of all $Q(0)$ for which Equation 5, Equation 6 and Equation 7 admit a feasible solution is convex.*

Significance Lemma 3 ensures that Theorem 1 still remains valid after the inductive noise constraints given by Equation 7 are added to the original set of linear constraints.

4.3 Process Variation Impact

To demonstrate the effect of process variations on DVFS control, we plot the voltage-frequency curves for the *slow*, *nominal* and *fast* process corners of a VFI in Figure 2(a). As mentioned before, we assume that due to reliability concerns, the process allows a maximum voltage of $V_{dd,max}$. As it can be seen, for a given $V_{dd,max}$, we obtain three different values of f_{MAX} for the *slow*, *nominal* and *fast* VFIs. Figure 2(b) shows the relationship between desired and applied frequency values in the presence of process variations - from the figure it is clear that under the impact

of process variations, we must think of f_{MAX}^i as *random variables*, not fixed upper limits on the operating frequency of each VFI.

As a result, the linear programming framework described in the previous sections will now have a certain probability of being feasible, i.e., there might exist values of f_{MAX}^i for which it is not possible to bring the system back to steady state within J control intervals. We will henceforth refer to the probability that a given instance of a multiple VFI system can be brought back to the reference queue occupancies in J time steps as the *probability of controllability* (PoC).

In this paper, we use Monte Carlo simulations to estimate the PoC, i.e., in each Monte Carlo run, we obtain a sample of the maximum frequency for each VFI, f_{MAX}^i , and check for the feasibility of the linear program defined by Equations 5, 6, 8 and 9. Furthermore, we are able to exploit the specific structure of our problem to *speed up the Monte Carlo simulations*. In particular, we note that if a given vector of upper bounds, $f_{MAX}^{i,1}$ ($i \in [1, M]$), has a feasible solution, then another vector, $f_{MAX}^{i,2}$ ($i \in [1, M]$), where $f_{MAX}^{i,2} \geq f_{MAX}^{i,1} \forall i \in [1, M]$ must also have a feasible solution. Therefore, we do not need to explicitly check for the feasibility of the upper bound $f_{MAX}^{i,2}$, thereby saving significant computational effort. This provides significant speed-up over a naive Monte Carlo implementation.

5. EXPERIMENTAL RESULTS

To validate the theory presented in this paper, we experiment on two benchmarks: (1) *MPEG*, is a distributed implementation of an MPEG-2 encoder with six ARM7-TDMI processors that are partitioned to form a 3 VFI system, as shown in Figure 3(a); and (2) *Star*, a 5 VFI system organized in a star topology as shown in Figure 3(b). The MPEG encoder benchmark was profiled on a cycle-accurate MPSoC simulator to obtain the average rates at which the VFIs read and write from the queues, as tabulated in Figure 3(a). The arrival and service rates of the *Star* benchmark are randomly generated.

To begin, we first compute the nominal frequency values f_{NOM}^i of each VFI in the system to ensure stable queues for the nominal workload values. The maximum frequency constraint, f_{MAX}^i is then set using a parameter $\gamma = \frac{f_{MAX}^i}{f_{NOM}^i}$ - in our experiments we use three values of $\gamma = \{1.1, 1.25, 1.5\}$. Finally, we allow the maximum frequency increment per control interval to vary from 5% to 20% of the nominal frequency.

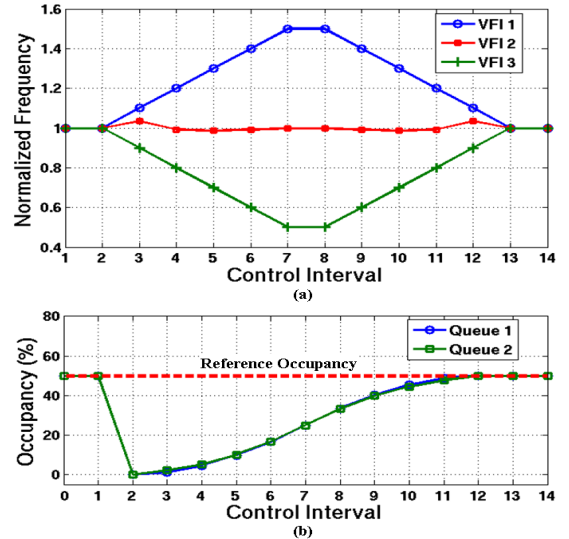


Figure 4: (a) Response of proposed controller to a deviation from the reference queue occupancies. (b) Evolution of queue occupancies in the system.

Figure 3(c) shows the obtained results as γ and the maximum frequency step are varied for the *MPEG* benchmark. The results for *Star* benchmark are quantitatively similar, so we only show the graph for *MPEG* benchmark in Figure 3(c). As it can be seen, the

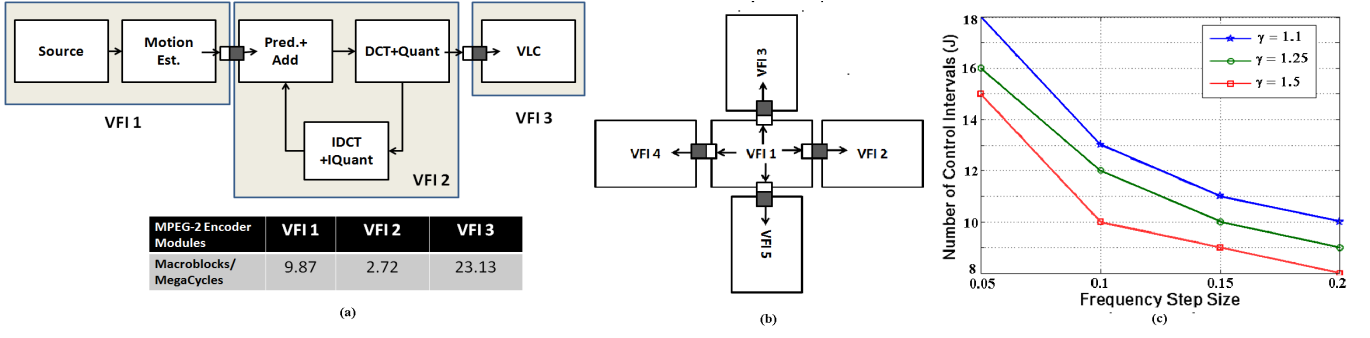


Figure 3: (a) Topology and workload characteristics of the *MPEG* benchmark. (b) Topology of the *Star* benchmark. (c) Impact of γ and maximum frequency increment on the minimum number of control intervals, J .

frequency step size has a significant impact on the controllability of the system, in particular, for $\gamma = 1.5$ we see an 87% increase in the number of control intervals required to bring the system back to reference queue occupancies, J , while for $\gamma = 1.1$, J increases by up to 80%. The impact of γ itself is slightly more modest - we see a 20% to 25% increase in J as γ increases from 1.1 to 1.5.

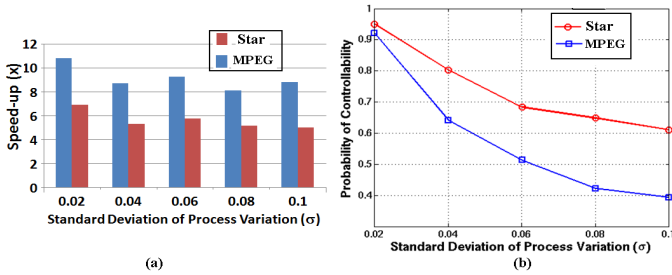


Figure 5: (a) Speed-up ($\times$) of the efficient Monte Carlo technique compared to naive Monte Carlo. (b) PoC as a function of increasing process parameter variations.

In Figure 4, we plot the response of the time-optimal control strategy for the *MPEG* benchmark when the queue occupancies of the two queues in the system drop to zero at the control interval 2. As a result, the applied frequency values are modulated to bring the queues back to their reference occupancies within $J = 10$ control intervals. From Figure 4(a), we can clearly observe the impact of both the limit on the maximum frequency, and the limit on the maximum frequency increment, on the time-optimal control response. Figure 4(b) shows how the queue occupancies change in response to the applied control frequencies, starting from 0% occupancy until they reach their reference occupancies.

Next, we investigate the impact of process variations on the PoC of DVFS enabled multiple VFI systems. For this experiment, we model the maximum frequency of each VFI as an independent normal distribution, and increase the standard deviation (σ) of the distribution from 2% to 10% of the maximum frequency. Finally, we use 5000 runs of both naive Monte Carlo simulations and the proposed efficient Monte Carlo simulations (see Section 4.3) to obtain the PoC for various values of σ and for both benchmarks. From Figure 5(a), we can see that the proposed efficient version of Monte Carlo provides **significant speed-up** over the naive Monte Carlo implementation - on average, a $9\times$ speed-up for the *MPEG* benchmark and a $5.6\times$ speed-up for the *Star* benchmark.

From the estimated PoC values in Figure 5(b), we can see that the PoC of both *MPEG* and *Star* benchmarks are significantly impacted by process variations, though *MPEG* sees a greater degradation in the PoC, decreasing from 92% for $\sigma = 2\%$ to only 40% for $\sigma = 10\%$. On the other hand, the PoC of *Star* drops from 95% to 62% for the same values of σ . To explain the significance of these results, we point out that a PoC of 40% implies that, on average, 60% of the fabricated circuits will *not* be able to meet the DVFS control performance specification, irrespective of the control algorithm that is used. Of note, while the specific parameters used in the Monte Carlo simulations (for example, the value of γ at various technology nodes) are implementation dependent and

may cause small changes in the PoC estimates in Figure 5, the fundamental predictive nature of this plot will remain the same.

We now briefly discuss the integration of the proposed framework in a power-aware NoC design methodology. First, given a specific implementation of a DVFS controller, for example, the ones proposed in [7] or [9], the framework can be used to determine how far away a particular implementation of the controller in Figure 1 is from an optimal solution. Second, based on the application, designers may be willing to trade-off reliability, peak temperature or maximum inductive noise for a better DVFS controller. In this scenario, the proposed approach can be used to efficiently explore the trade-offs between these technology driven factors and the upper bound on performance of DVFS control, similar to the analysis presented in Figure 3(c).

6. CONCLUSION

In this paper, we presented a theoretical framework to efficiently obtain the limits on the controllability and performance of DVFS controllers for multiple VFI based networks-on-chip. Using a computationally efficient implementation of the framework, we present results, using both real and synthetic benchmarks, that explore the impact of three major technology driven factors - temperature and reliability constraints, maximum inductive noise constraints and process variations - on the performance bounds of DVFS control strategies. Our experiments demonstrate the importance of considering the impact of these three factors on DVFS controller performance, particularly since all three factors are becoming increasingly important with technology scaling. As future work, we plan to consider explicit control of power consumption instead of purely time optimal control.

7. REFERENCES

- [1] S. Garg et al. Technology-Driven Limits on DVFS Controllability of Multiple Voltage-Frequency Island Designs. Technical report, CSSI, CMU, 2009.
- [2] W. Jang, D. Ding, and D.Z. Pan. A Voltage-Frequency Island aware energy optimization framework for networks-on-chip. In *Proceedings of ICCAD*, 2008.
- [3] W. Kim et al. System Level Analysis of Fast, Per-Core DVFS using On-Chip Switching Regulators. In *Proceedings of HPCA*, 2008.
- [4] B.C. Kuo. *Digital Control Systems*. Oxford University Press, Inc. New York, NY, USA, 1992.
- [5] D. Marculescu and S. Garg. System-level process-driven variability analysis for single and multiple voltage-frequency island systems. 2006.
- [6] K. Niyogi and D. Marculescu. Speed and voltage selection for GALS systems based on voltage/frequency islands. In *Proceedings of ASP-DAC*, 2005.
- [7] U.Y. Ogras et al. Variation-Adaptive Feedback Control for Networks-on-Chip with Multiple Clock Domains. In *Proceedings of DAC*, 2008.
- [8] H.L. Royden. *Real analysis*. Macmillan New York, 1968.
- [9] Q. Wu et al. Formal online methods for voltage/frequency control in multiple clock domain microprocessors. *Proceedings of ASPLOS*, 2004.